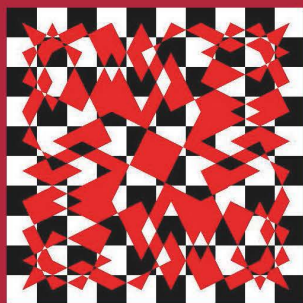


Grzegorz
Andrzejczak



KONSTRUKCJE II

Wybrane konstrukcje matematyki
teoretycznej Teoria całki i probabilistyka

Politechnika Łódzka
Wydział Fizyki Technicznej, Informatyki i Matematyki
Stosowanej
Matematyka Stosowana

Konstrukcje II

**Wybrane konstrukcje matematyki
teoretycznej
Teoria całki i probabilistyka**

Grzegorz Andrzejczak

Łódź sierpień 2018 – czerwiec 2021

Recenzenci:
prof. dr hab. Aleksander Błaszczuk
prof. dr hab. Andrzej Łuczak

Redaktor Naukowy Wydziału Fizyki Technicznej,
Informatyki i Matematyki Stosowanej:
dr hab. inż. Aneta Poniszewska-Marańda

©Copyright by Politechnika Łódzka 2021

WYDAWNICTWO POLITECHNIKI ŁÓDZKIEJ

90-924 Łódź, ul. Wólczańska 223

tel. 42 631-29-52, 42 631-20-87

fax 42 631-25-38

e-mail: zamowienia@info.p.lodz.pl

www.wydawnictwo.p.lodz.pl

ISBN 978-83-66741-22-5

DOI: 10.34658/9788366741225

Monografie Politechniki Łódzkiej, Nr 2370
Wydanie pierwsze

Spis treści

Wstęp	1
1 Miary i całka Lebesgue’a	5
1.1 Elementy teorii miary	5
1.1.1 Przestrzenie mierzalne	5
1.1.2 Czym mierzyć zbiory?	11
1.1.3 Co zapożyczamy od Carathéodory’ego?	14
1.1.4 Lokalnie skończone miary borelowskie	16
1.1.5 Słaba zbieżność w przestrzeni miar	19
1.2 Całka Lebesgue’a	23
1.2.1 Podejście aksjomatyczne	23
1.2.2 Wnioski, konsekwencje...	25
1.2.3 Zamiana zmiennych w całce Lebesgue’a	27
1.2.4 ...i rozszerzenia	28
1.2.5 Całka względem produktu miar	31
1.3 Twierdzenie Lebesgue’a–Radona–Nikodyma	34
1.3.1 Absolutna – relatywna ciągłość miary	34
1.3.2 Twierdzenia o rozkładzie	36
1.3.3 ...i dowody	39
1.3.4 Całka względem σ –ciała	44
1.3.5 Warunkowe miary wg. Billingsleya	49
2 To nieprawdopodobne	56
2.1 Probabilistyczny punkt widzenia	56
2.1.1 Odświeżanie tradycji	56
2.1.2 Zmienne losowe i ich rozkłady	62
2.1.3 Najważniejsza jest niezależność	66

2.1.4	Warunkowe rozkłady i wartość oczekiwana	74
2.1.5	Charakterystyki wektorów losowych	78
2.2	Prawo wielkich liczb	81
2.2.1	Dokład prowadzi błądzenie losowe	81
2.2.2	Mocne prawo wielkich liczb	89
2.2.3	Podejście empiryczne	96
2.2.4	Twierdzenie Gliwienki–Cantelliego	100
2.2.5	Nieskończone próby losowe	104
2.3	Modelowanie rzeczywistości czyli nieco statystyki	107
2.3.1	Jednostajna losowość	107
2.3.2	Skończone ciągi pseudolosowe	109
2.3.3	Nieparametryczne testy losowości	113
2.3.4	Jaką informację wnosi skończony zbiór obserwacji?	119
2.3.5	Statystyki dostateczne czyli dobre	122
2.4	Rozkład normalny	128
2.4.1	Skąd TO się wzięło?	128
2.4.2	...i od strony rozkładów ciągłych	134
2.4.3	Rozkład wielowymiarowy	136
2.4.4	Jak „wyprostować” rozkład normalny?	139
2.5	Przestrzenie rozkładów prawdopodobieństwa	141
2.5.1	Transformata Fouriera	141
2.5.2	Ciągłość	144
2.5.3	Czas na Centralne twierdzenie graniczne	148
2.5.4	...i wariant wektorowy	151
2.5.5	Metryzowalność	154
2.6	Krok w stronę procesów stochastycznych	157
2.6.1	Definicja – czyli twierdzenie Kołmogorowa	157
2.6.2	Trajektorie i ośrodkowość	160
2.6.3	Ciągłość trajektorii jako problem	164
2.6.4	Trajektorie jednostronnie ciągłe	168
2.7	Ruchy Browna	170
2.7.1	...czyli miara Wienera	170
2.7.2	Konstrukcja procesu Wienera wg. Ciesielskiego	173
2.7.3	Twierdzenie graniczne Donskera	177
3	Całki jako miary	182
3.1	Abstrakcyjna konstrukcja całki Daniella–Stone’a	182
3.1.1	Od funkcjonału do miary zewnętrznej	182
3.1.2	Co można pominąć?	187

3.1.3	Funkcje całkowalne wg Daniella i Stone'a	189
3.1.4	Klasyczne twierdzenia graniczne dla całek	192
3.1.5	Powrót do całki Lebesgue'a	193
3.2	Miary Radona	196
3.2.1	Topologie zbieżności (niemal) jednostajnej	197
3.2.2	Mierzalność oparta na ciągłości	201
3.2.3	Struktura rzeczywistych i zespolonych całek Radona	203
3.3	Regularne miary i całki	206
3.3.1	Miary wyznaczone poprzez zbiory zwarte	206
3.3.2	Charakteryzacje miar regularnych	208
3.3.3	Nośnik miary regularnej	212
3.3.4	Całka względem miary regularnej	214
Spis symboli i skrótów		219
Skorowidz		224
Bibliografia		229

Wstęp

Ze wstępu do Konstrukcji I [2]:

Pomysł i szczegółowa koncepcja prezentowanego tekstu „Wybrane konstrukcje matematyki teoretycznej...” są wynikiem obserwacji stanu wiedzy niezbyt licznej grupy studentów matematyki studiów doktoranckich prowadzonych od 2012 r. na Wydziale FTIMS Politechniki Łódzkiej. Absolwenci matematyki, po studiach z różnych uczelni, prezentują na ogół dość zaawansowaną wiedzę dotyczącą z zasady wąskich dyscyplin matematyki, pojmowanych jako odrębne i właściwie niezależne – niepowiązane w istotny sposób ze sobą. Zasadniczym celem tekstu jest zatem pokazanie czytelnikowi wzajemnego przenikania wybranych, wskazanych w tytule działów matematyki lokujących się w pobliżu szeroko pojętej analizy matematycznej. Topologia, teoria miary i całki oraz rachunek prawdopodobieństwa na średnio zaawansowanym poziomie stanowią w ramach matematyki na studiach politechnicznych dość kłopotliwy materiał wykładowy: dwie pierwsze dyscypliny – ze względu na wymagany właściwy poziom abstrakcji, natomiast probabilistyka – z uwagi na wyraźną i pilną potrzebę zastosowań, co skutkuje istotnymi uproszczeniami w zakresie teorii. (...) Przygotowywane równolegle opracowanie „Konstrukcje II...” zaczyna się od twierdzenia Radona–Nikodyma, a kończy na ruchach Browna oglądanych także z perspektywy miary Wienera w przestrzeni funkcji ciągłych $\mathcal{C}(\mathbb{R}_+)$.

Układ prezentowanej właśnie czytelnikowi drugiej części wielotematycznej monografii poświęconej *Konstrukcjom w Matematyce* mieści się w naszkicowanych ramach. Spis treści podpowiada jednak, że w rozdziale 1 dwa pierwsze podrozdziały stanowią dodatkowe wprowadzenie – w zakresie pojęć (i oznaczeń) do elementarnej teorii miary i podstaw całki Lebesgue’a. Terminologia i treść twierdzeń odpowiadają w zasadzie wybranym podrozdziałom Konstrukcji I, ale nie są tylko zapożyczeniem, co czytelnik znający

pierwszą z monografii z pewnością zauważy.

Podrozdział 1.3 obejmuje twierdzenia wiązane z nazwiskami Lebesgue’a oraz Radona i Nikodyma, dotyczące wzajemnych relacji pomiędzy miarami w tej samej przestrzeni mierzalnej. Umieszczenie całości rozważań w omawianym rozdziale uzasadnia fakt, iż uzyskane zależności przybierają szczególnie czytelną postać dla całek. Wybrany sposób prezentacji (i dowodu) nie wymaga odwoływania się do pojęć bardziej ogólnych niż klasyczne pojęcie miary. Jako szczególny, ważny w probabilistyce wniosek wyprowadzamy stąd twierdzenia o istnieniu warunkowej wartości oczekiwanej, scharakteryzowanej jako *relatywna całka* – względem ustalonego σ -ciała złożonego ze zbiorów mierzalnych. Wprawdzie naturalny zestaw założeń przyjęty w podrozdziałach 1.3.4–1.3.5 nie odwołuje się do pojęć probabilistycznych, uzyskane pod koniec rozdziału *warunkowe miary*, czyli regularne rozkłady warunkowe stały się dla autora monografii jednoznaczną motywacją, by zmienić planowaną wcześniej kolejność rozdziałów.

Rozdział 2 kieruje teraz uwagę czytelnika na wybrane konstrukcje specyficzne dla *Probabilistyki*. Podrozdział 2.1 przypomina podstawy (i nie tylko) intensywnie wykorzystując dostępne pojęcia z teorii miary i całki. Czytelnik znajdzie tu zatem charakteryzacje zbieżności wg. rozkładów, omówienie pojęcia niezależności σ -ciał, a także szczegółowe, korzystające z poprzednich podrozdziałów definicje warunkowych rozkładów zmiennych losowych – w ujęciu wielowymiarowym.

Głównym założeniem przyjętym w podrozdziale 2.2 jest niezależność zmiennych losowych o tym samym rozkładzie. Tytułowe *prawo wielkich liczb* jest dowodzone w wariancie silnym, a poprzedza je dowód klasycznej alternatywy dotyczącej rozbieżności prawie na pewno (zbieżności do $\pm\infty$) ciągu sum opisujących błędzenie losowe – bez założeń dotyczących istnienia wartości oczekiwanej zmiennych. Po sprecyzowaniu pojęcia próby losowej dowolnej zmiennej losowej prezentujemy podstawowe zastosowanie prawa wielkich liczb dla estymacji – wartości oczekiwanej, wariancji i kowariancji rozważanych zmiennych. Klasyczne twierdzenie Gliwienki–Cantellego przyjmuje w tym kontekście rozszerzoną postać obejmującą także szczególną zbieżność ciągu empirycznych funkcji kwantylowych. Podrozdział kończy twierdzenie *O regularnej różnorodności prób losowych* stanowiące wprowadzenie do intuicyjnej koncepcji *losowości ciągu danych*.

Bardziej systematyczny opis konstrukcji związanych z losowością danych zawiera podrozdział 2.3. Omawiamy tu w szczególności klasyczną rodzinę kongruencyjnych algorytmów generujących ciągi liczb pseudolosowych. Z kolei istotną dla wiarygodności ciągów danych empirycznych *losowość*

poddaje próbie test **runs** Walda–Wolfowitza, którego uniwersalny charakter podkreśla twierdzenie 2.3.13. Dla skończonych danych empirycznych konstruujemy przykładowe drogi estymacji bayesowskiej wybranych rozkładów parametrycznych. Podrozdział kończy układ twierdzeń (i przykładów) charakteryzujących statystyki dostateczne.

Podrozdział 2.4 zawiera szczegóły dwu konstrukcji prowadzących w granicy do klasycznego rozkładu Gaussa – także w wariancie wielowymiarowym, w oparciu o pojęcie wektora losowego typu Gaussa. Czytelnik znajdzie tu również stosowne twierdzenie graniczne dla rozkładu Walda–Wolfowitza, a także algorytmy Boxa–Mullera i Marsaglii symulujące rozkład normalny na podstawie ciągu danych o rozkładzie jednostajnym.

Teraz, skoro mamy już rozkład normalny, w podrozdziale 2.5 formułujemy i dowodzimy *centralne twierdzenie graniczne* – w oparciu o własności transformaty znanej jako funkcja charakterystyczna rozkładu. Zbieżność rozważanych wcześniej ciągów zmiennych losowych i dystrybuant formalizowana jest następnie jako zbieżność w metryzowalnej przestrzeni rozkładów.

Wprowadzenie do procesów stochastycznych omawia podrozdział 2.6. Podajemy tu pochodzącą od Dooba konstrukcję redukcji przestrzeni funkcji rzeczywistych $\mathbb{R}^{\mathbb{R}^+}$ pozwalającą na wyróżnienie jako zbioru mierzalnego podprzestrzeni funkcji ciągłych, co ściśle wiąże się z problemem istnienia ciągłych trajektorii procesu. Osłabienie założeń prowadzi do konstrukcji przestrzeni $\mathcal{D}$ procesów jednostronnie ciągłych.

Rozdział poświęcony probabilistyce kończy podrozdział 2.7, w którym znajdziemy jednoznaczność i konstrukcję *ruchów Browna* i miary Wienera – wyróżniającej najbardziej prawdopodobne, *typowe* własności funkcji ciągłych na półprostej $\mathbb{R}_+$ – w tym ich nieróżniczkowalność. Pochodzący od Kołmogorowa bardziej abstrakcyjny dowód istnienia procesu Wienera uzupełnia (niemal) jawna konstrukcja ruchu Browna zaproponowana przez Ciesielskiego. Rozdział kończy tw. graniczne Donskera będące uniwersalnym, ciągłym odpowiednikiem centralnego twierdzenia granicznego.

W rozdziale 3, ostatnim rozdziale prezentowanej monografii, wracamy do całek – nie jest to jednak „ogólna teoria całki”. Rozdział zawiera dwie klasyczne (i czasem zapominane) koncepcje, w których całka czyli miara funkcji poprzedza miarę zbioru. Konstrukcja całki pochodząca od Daniella i rozwinięta przez Stone’a stanowi treść podrozdziału 3.1. Punktem wyjścia jest tu nieujemny, monotonicznie ciągły funkcjonał liniowy na wskazanej kracie skończonych funkcji rzeczywistych. Zarówno całka jak i mierzalność funkcji wraz z miarą zbiorów powstają w oparciu o jednoznaczne rozszerzenie wyjściowego funkcjonału. Omawiany podrozdział zawiera także cha-

rakteryzację funkcji całkowalnych w sensie Daniella i zbiorów *pomijalnych*, czyli miary 0. W przypadku kraty spełniającej dodatkowy warunek Stone'a otrzymana w ten sposób całka Daniella–Stone'a jest pojęciem równoważnym z całką Lebesgue'a – względem miary Daniella–Stone'a

Koncepcja miary i całki Radona, rozwijana w podrozdziale 3.2, dotyczy przestrzeni lokalnie zwartych i ma ścisły związek z twierdzeniem Radona o dualności, charakteryzującym funkcjonały liniowe i ciągłe na przestrzeni funkcji (zespolonych) o zwartych nośnikach. Ponieważ nieujemne funkcjonały są tożsame z całką Lebesgue'a względem miary regularnej na σ –pierścieniu zbiorów Baire'a, ostatni podrozdział przedstawia konstrukcję wiążącą takie właśnie miary z regularnymi miarami borelowskimi w przestrzeniach lokalnie zwartych.

Życzę uważnej lektury – Grzegorz Andrzejczak.

Rozdział 1

Miary i całka Lebesgue'a

1.1 Elementy teorii miary

1.1.1 Przestrzenie mierzalne

Definicja 1.1.1. Niepustą rodzinę $\mathfrak{T}$ podzbiorów zbioru T nazywamy ciałem (lub algebrą zbiorów), jeśli jest zamknięta ze względu na operacje dopełnienia oraz sumy zbiorów. Ciałło $\mathfrak{T}$ jest σ -ciałem, jeśli jest zamknięte ze względu na przeliczalne sumy zbiorów. Podzbiory T należące do ustalonego σ -ciała są zbiorami mierzalnymi.

Rodzinę wszystkich podzbiorów zbioru T będziemy oznaczać $\mathfrak{P}(T)$.

Definicja 1.1.2. Dla dowolnego σ -ciała $\mathfrak{T}$ w zbiorze T parę $(T, \mathfrak{T})$ nazywamy przestrzenią mierzalną. Odwzorowaniem $f: (S, \mathfrak{S}) \rightarrow (T, \mathfrak{T})$ przestrzeni mierzalnych jest każda funkcja $f: S \rightarrow T$ $\mathfrak{S}/\mathfrak{T}$ -mierzalna, tzn. spełniająca warunek

$$f^{-1}B \in \mathfrak{S}, \quad \text{dla } B \in \mathfrak{T}.$$

Zbiór pusty i – w tej konwencji – cała przestrzeń są mierzalne podobnie jak przeliczalne przekroje zbiorów mierzalnych.

Stwierdzenie 1.1.3. Złożenie $f \circ g$ dowolnych odwzorowań mierzalnych $f: (S, \mathfrak{S}) \rightarrow (T, \mathfrak{T})$ i $g: (R, \mathfrak{R}) \rightarrow (S, \mathfrak{S})$ jest $\mathfrak{R}/\mathfrak{T}$ -mierzalne.

Przestrzenie i odwzorowania mierzalne stanowią kategorię. $\square$

Definicja 1.1.4. Dla dowolnej rodziny podzbiorów $\mathfrak{C} \subset \mathfrak{P}(T)$ zbioru T , najmniejsze σ -ciało w T zawierające $\mathfrak{C}$ nazywamy σ -ciałem generowanym przez $\mathfrak{C}$ i oznaczamy przez $\sigma_T(\mathfrak{C})$.

Poprawność definicji wynika z faktu, że przekrój dowolnej rodziny σ –ciał w T jest również σ –ciałem.

Pojęcie σ –ciała jest *relatywne* – zależy od otaczającej przestrzeni. Wskazuje na to również użyty w oznaczeniu $\sigma_T(\mathfrak{C})$ symbol przestrzeni T , w której rodzina $\mathfrak{C}$ generuje σ –ciało.

Definicja 1.1.5. σ –ciałem zbiorów borelowskich w przestrzeni topologicznej (T, τ) nazywamy rodzinę $\mathfrak{B}_\tau = \sigma_T(\tau)$ generowaną przez topologię. Parę $(T, \mathfrak{B}_\tau)$ będziemy nazywać borelowską przestrzenią mierzalną przestrzeni T . W sytuacji, gdy topologia przestrzeni T wynika z kontekstu, korzystamy z uproszczonego oznaczenia $\mathfrak{B}_T \equiv \mathfrak{B}_\tau$.

Podane wyżej definicje są wzajemnie zgodne w odniesieniu do podprzestrzeni, jeśli przyjmujemy oznaczenie

$$\mathfrak{C}|A = \{C \cap A; C \in \mathfrak{C}\} \quad \text{dla dowolnej rodziny zbiorów } \mathfrak{C} \quad (1.1)$$

i przestrzeń mierzalną $(A, \mathfrak{T}|A)$ będziemy uważać za *podprzestrzeń* przestrzeni $(T, \mathfrak{T})$, dla $A \subset T$.

Stwierdzenie 1.1.6. Dla dowolnej rodziny podzbiorów $\mathfrak{C}$ w T i zbioru $A \subset T$ zachodzi równość

$$\sigma_A(\mathfrak{C}|A) = \sigma_T(\mathfrak{C})|A. \quad (1.2)$$

W szczególności, $\mathfrak{B}_A = \mathfrak{B}_T|A$, jeśli T jest przestrzenią topologiczną, a $A \subset T$ podprzestrzenią.

Dowód. Proste, klasyczne ćwiczenie. □

Zakładając, że rozszerzona prosta rzeczywista $\overline{\mathbb{R}} = \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$ ma naturalną topologię wyznaczoną przez porządek, przyjmujemy

Definicja 1.1.7. Odwzorowania $\mathfrak{T}/\mathfrak{B}_{\overline{\mathbb{R}}}$ –mieralne z dowolnej przestrzeni mierzalnej $(T, \mathfrak{T})$ w prostą $\overline{\mathbb{R}}$ nazywamy skończonymi funkcjami mierzalnymi na przestrzeni $(T, \mathfrak{T})$ (lub $\mathfrak{T}$ –mierzalnymi). Za funkcje $\mathfrak{T}$ –mieralne będziemy uważać ogólniejsze odwzorowania $\mathfrak{T}/\mathfrak{B}_{\overline{\mathbb{R}}}$ –mieralne $T \rightarrow \overline{\mathbb{R}}$.

Stąd – dwa analogiczne oznaczenia zbiorów: $\mathcal{M}(\mathfrak{T})$ – dla przestrzeni liniowej skończonych funkcji mierzalnych i $\overline{\mathcal{M}}(\mathfrak{T})$ – dla zbioru wszystkich mierzalnych funkcji na $(T, \mathfrak{T})$.

Ćwiczenie 1.1.1. Dla dowolnej przestrzeni mierzalnej $(S, \mathfrak{S})$ i dowolnej rodziny zbiorów $\mathfrak{C} \subset \mathfrak{P}(T)$ odwzorowanie $f: S \rightarrow T$ jest $\mathfrak{S}/\sigma_T(\mathfrak{C})$ –mieralne wtedy i tylko wtedy, gdy $f^{-1}A \in \mathfrak{S}$ dla $A \in \mathfrak{C}$.

Każde odwzorowanie ciągle przestrzeni topologicznych $f: S \rightarrow T$ jest zatem $\mathfrak{B}_S/\mathfrak{B}_T$ -mierzalne, a mierzalność funkcji rzeczywistej $f: T \rightarrow \overline{\mathbb{R}}$ oznacza mierzalność przeciwobrazów $f^{-1}(-\infty, a)$ dla $a \in \mathbb{R}$.

Definicja 1.1.8. Produktem dowolnej rodziny $\{(S_a, \mathfrak{S}_a)\}_{a \in \mathcal{A}}$ przestrzeni mierzalnych nazywamy przestrzeń

$$\prod_{a \in \mathcal{A}} (S_a, \mathfrak{S}_a) = \left(\prod_{a \in \mathcal{A}} S_a, \bigotimes_{a \in \mathcal{A}} \mathfrak{S}_a \right), \quad (1.3)$$

w której produkt kartezjański $\prod_{a \in \mathcal{A}} S_a$ rozważany jest wraz σ -ciałem $\bigotimes_a \mathfrak{S}_a$ generowanym przez przeciwobrazy $\pi_\alpha^{-1} \mathfrak{S}_\alpha, \alpha \in \mathcal{A}$.

Zgodnie z [2, stw. 2.2.2] podana w definicji konstrukcja wyróżnia produkt w kategorii przestrzeni mierzalnych (i odwzorowań mierzalnych) i zapewnia mierzalność wszystkich rzutowań $\pi_\alpha: \prod_{a \in \mathcal{A}} S_a \rightarrow S_\alpha$.

Przykład 1.1.2. Zbiór $\mathbb{R}^T$ funkcji rzeczywistych określonych na (dowolnym) zbiorze $T \neq \emptyset$ posiada naturalną topologię produktową, w której wszystkie projekcje $\pi_t: \mathbb{R}^T \ni f \mapsto f(t) \in \mathbb{R}, t \in T$, są ciągłe. Jednakże, jeśli zbiór T jest nieprzeliczalny, każda z funkcji $f \in \mathbb{R}^T$ wyznacza jednopunktowy zbiór domknięty $\{f\}$, który nie jest zbiorem mierzalnym w produkcie $(\mathbb{R}, \mathfrak{B}_\mathbb{R})^T = (\mathbb{R}^T, \mathfrak{B}_\mathbb{R}^T)$ zdefiniowanym przez słabszy warunek mierzalności projekcji $\pi_t, t \in T$ (por. [2, podrozdz. 2.2.1]). Preferowane, produktowe σ -ciało

$$\mathfrak{B}_\mathbb{R}^T = \sigma\{\pi_t^{-1}A; A \in \mathfrak{B}_\mathbb{R}, t \in T\} \quad (1.4)$$

zapewnia naturalną charakteryzację mierzalności po współrzędnych dowolnego odwzorowania $X = (X_t)_{t \in T}: S \rightarrow \mathbb{R}^T$ utworzonego przez rodzinę odwzorowań mierzalnych $X_t: (S, \mathfrak{S}) \rightarrow (\mathbb{R}, \mathfrak{B}_\mathbb{R}), t \in T$, i jest wykorzystane np. w obrębie probabilistyki do opisu procesów stochastycznych.

Jeśli T jest zbiorem skończonym lub przeliczalnym, przestrzeń topologiczna $\mathbb{R}^T$ jest ośrodkową przestrzenią metryczną (metryzowalną), a produktowe σ -ciało $\mathfrak{B}_\mathbb{R}^T$ jest tożsame z σ -ciałem zbiorów borelowskich $\mathfrak{B}_{\mathbb{R}^T}$ [2, Wn. 2.2.8].

Dla opisu sytuacji, w których cała przestrzeń jest „zbyt duża”, by można ją było realnie mierzyć, przydatne okazują się struktury spełniające słabsze lub alternatywne układy warunków.

Definicja 1.1.9. Klasa (rodzina) zbiorów $\mathfrak{C} \subset \mathfrak{P}(S)$ jest π -systemem, jeśli jest zamknięta ze względu na skończone przekroje zbiorów, tzn.

$$A, B \in \mathfrak{C} \Rightarrow A \cap B \in \mathfrak{C}.$$

- Klasa zbiorów $\mathfrak{D} \subset \mathfrak{P}(S)$ jest λ -systemem w S , jeśli zawiera S , a wraz z każdym zbiorem $A \in \mathfrak{D}$ zawiera dopełnienie $S \setminus A$, oraz jeśli suma każdego wstępującego ciągu klasy $\mathfrak{D}$ również należy do $\mathfrak{D}$.

Rodzina zbiorów $\mathfrak{D} \subset \mathfrak{P}(S)$ o własności

$$A \cap B = \emptyset \implies A \cup B \in \mathfrak{D}, \quad A \subset B \implies B \setminus A \in \mathfrak{D} \quad (1.5)$$

dla dowolnych $A, B \in \mathfrak{D}$, jest odpowiednio

- σ -systemem, jeśli każda suma wstępującego ciągu klasy $\mathfrak{D}$ należy do $\mathfrak{D}$,
- δ -systemem, gdy przekrój każdego zstępującego ciągu klasy $\mathfrak{D}$ należy do $\mathfrak{D}$.

Uwaga 1.1.3. W literaturze przedmiotu można znaleźć nierównoważne definicje λ -systemu jako rodziny zbiorów $\mathfrak{D} \subset \mathfrak{P}(S)$ zawierającej S i spełniającej alternatywnie w stosunku do podanych wyżej warunki

- (i) $A \setminus B \in \mathfrak{D}$ dla $A, B \in \mathfrak{D}$ takich, że $B \subset A$, oraz (lub)
- (ii) suma dowolnej przeliczalnej rodziny rozłącznych zbiorów $A_n \in \mathfrak{D}$, $n \in \mathbb{N}$, także należy do $\mathfrak{D}$.

Ćwiczenie 1.1.4. Sprawdzić, że rodzina zbiorów $\mathfrak{D} \subset \mathfrak{P}(S)$ jest σ -ciałem wtedy i tylko wtedy, gdy jest jednocześnie π - i λ -systemem.

Definicja 1.1.10. Niepustą rodzinę zbiorów $\mathfrak{D}$ zamkniętą względem operacji sumy i różnicy zbiorów nazywamy pierścieniem. Pierścień $\mathfrak{D}$ jest σ -pierścieniem, jeśli jest zamknięty ze względu na przeliczalne sumy zbiorów, a δ -pierścieniem, jeśli jest zamknięty ze względu na przeliczalne przekroje.

Stwierdzenie 1.1.11. (i) Rodzina zbiorów $\mathfrak{D} \subset \mathfrak{P}(S)$ jest σ -pierścieniem wtedy i tylko wtedy, gdy jest jednocześnie π - i σ -systemem, a δ -pierścieniem – gdy jest jednocześnie π - i δ -systemem.

(ii) Każdy σ -pierścień jest także δ -pierścieniem. Zarówno σ -pierścień jak i δ -pierścień jest σ -ciałem wtedy i tylko wtedy, gdy zawiera całą przestrzeń. $\square$

Definicja 1.1.12. Dla dowolnej niepustej rodziny zbiorów $\mathfrak{C}$, najmniej σ -pierścień zawierający $\mathfrak{C}$ nazywamy σ -pierścieniem generowanym przez $\mathfrak{C}$ i oznaczamy $\sigma(\mathfrak{C})$. Najmniejszy δ -pierścień zawierający $\mathfrak{C}$ nazywamy δ -pierścieniem generowanym przez $\mathfrak{C}$ i oznaczamy $\delta(\mathfrak{C})$.

Stosowana tu konwencja dotycząca oznaczeń nie jest w pełni zgodna z powszechnie przyjętą stosowaną właściwie tylko dla σ -ciał. Równość

$$\sigma_S(\mathfrak{C}) = \sigma(\mathfrak{C} \cup \{S\}), \quad \text{jeśli } \mathfrak{C} \subset \mathfrak{P}(S),$$

odpowiada intuicji, by dla prostszego i bardziej ogólnego pojęcia σ –pierścienia korzystać z prostszego oznaczenia. Ani σ –pierścień $\sigma(\mathfrak{C})$ ani δ –pierścień $\delta(\mathfrak{C})$ nie zależą od otaczającej przestrzeni. Przypominamy za [2, tw. 1.2.13] proste twierdzenie strukturalne

Twierdzenie 1.1.13. (i) *Dla dowolnego σ –pierścienia $\mathfrak{F}$ w S takiego, że $S \notin \mathfrak{F}$, zachodzi równość*

$$\sigma_S(\mathfrak{F}) = \mathfrak{F} \cup \mathfrak{F}^{\complement}, \quad \text{gdzie } \mathfrak{F}^{\complement} = \{S \setminus C; C \in \mathfrak{F}\},$$

przy czym rodziny $\mathfrak{F}$ i $\mathfrak{F}^{\complement}$ są rozłączne. Co więcej, nie istnieje para zbiorów $A \in \mathfrak{F}$ i $B \in \mathfrak{F}^{\complement}$ taka, że $B \subset A$.

(ii) *Dla dowolnej (niepustej) rodziny $\mathfrak{C} \subset \mathfrak{P}(S)$ zbior*

$$\mathfrak{F} = \{A \in \sigma_S(\mathfrak{C}); \exists \{C_n\}_{n \in \mathbb{N}} A \subset \bigcup_{n \in \mathbb{N}} C_n\}$$

jest σ –pierścieniem $\sigma(\mathfrak{C})$ generowanym przez $\mathfrak{C}$. W szczególności σ –pierścień $\sigma(\mathfrak{C})$ jest σ –ciałem wtedy i tylko wtedy, gdy zbiór S jest sumą ciągu zbiorów $(C_n)_{n \in \mathbb{N}}$ w $\mathfrak{C}$.

(iii) *Jeśli $\mathfrak{C}$ jest δ –pierścieniem, to rodzina $\mathfrak{F}$ utworzona z sum ciągów wstępujących*

$$\mathfrak{F} = \{\bigcup_n C_n; \forall_{n \in \mathbb{N}} C_n \in \mathfrak{C} \text{ i } C_n \subset C_{n+1}\}$$

jest σ –pierścieniem $\sigma(\mathfrak{C})$ generowanym przez $\mathfrak{C}$. □

Podane wyżej pojęcia łączy także twierdzenie Sierpińskiego [2, tw. 1.2.12]:

Twierdzenie 1.1.14 (O monotonicznych klasach zbiorów). *Niech $\mathfrak{C}$ będzie dowolnym π –systemem w przestrzeni (zbiorze) S .*

- (i) *najmniejszy λ –system zawierający $\mathfrak{C}$ jest σ –ciałem $\sigma_S(\mathfrak{C})$;*
- (ii) *najmniejszy σ –system zawierający $\mathfrak{C}$ jest σ –pierścieniem $\sigma(\mathfrak{C})$;*
- (iii) *najmniejszy δ –system zawierający $\mathfrak{C}$ jest δ –pierścieniem $\delta(\mathfrak{C})$.* □

Pojęcie δ –pierścienia ma raczej znaczenie techniczne, natomiast σ –pierścień prowadzi do naturalnych uogólnień. Rozszerzając definicję 1.1.2, przyjmujemy zatem

Definicja 1.1.15. *Dowolną parę $(T, \mathfrak{F})$ złożoną ze zbioru $T \neq \emptyset$ i σ –pierścienia $\mathfrak{F} \subset \mathfrak{P}(T)$ będziemy nazywać przestrzenią półmierzalną.*

Za odwzorowanie dwu przestrzeni półmierzalnych $f: (S, \mathfrak{G}) \rightarrow (T, \mathfrak{F})$ będziemy uważać każdą funkcję $\mathfrak{G}/\mathfrak{F}$ –mierzalną $f: S \rightarrow T$, czyli taką, że $f^{-1}\mathfrak{F} \subset \mathfrak{G}$.

Przeciwbraz $f^{-1}\mathfrak{F} = \{f^{-1}C; C \in \mathfrak{F}\}$ dowolnego σ -pierścienia jest zawsze σ -pierścieniem. W szczególności, każdy niepusty podzbiór $A \subset T$ wyznacza *podprzestrzeń*.

Definicja 1.1.16. Podprzestrzenią przestrzeni półmierzalnej $(T, \mathfrak{F})$ jest dowolny podzbiór $A \subset T$ wraz z σ -pierścieniem $\mathfrak{F}|A := \{A \cap C; C \in \mathfrak{F}\}$.

Definicja podprzestrzeni zapewnia mierzalność włożenia $\iota_A: A \hookrightarrow T$ niezależnie od tego, czy podzbiór $A \subset T$ jest mierzalny (tzn. $A \in \mathfrak{F}$), czy nie. Korzystamy tu z oznaczenia „ $\hookrightarrow$ ” i nazwy *włożenie* przyjętych w nieco szerszym kontekście, m.in. dla przestrzeni topologicznych.

Produkt przestrzeni półmierzalnych istnieje dla dowolnych przeliczalnych rodzin [2, par. 2.2.4].

Podążając za Halmosem [9], przedstawiamy nieco ogólniejszą koncepcję mierzalności funkcji rzeczywistych.

Definicja 1.1.17. Dla dowolnego σ -pierścienia $\mathfrak{F}$, rozszerzoną funkcję rzeczywistą f określoną na zbiorze $S \supset \bigcup \mathfrak{F}$ będziemy nazywać $\mathfrak{F}$ -mierzalną, jeśli spełnia warunek

$$\{x; f(x) \geq c\} \in \mathfrak{F} \quad \text{i} \quad \{x; f(x) \leq -c\} \in \mathfrak{F} \quad \text{dla } c > 0. \quad (1.6)$$

Zbiór funkcji $\mathfrak{F}$ -mierzalnych będziemy oznaczać $\overline{\mathcal{M}}(\mathfrak{F})$.

Z warunku (1.6) wynikają analogiczne, równoważne własności

$$\{x; f(x) > c\} \in \mathfrak{F} \quad \text{i} \quad \{x; f(x) < -c\} \in \mathfrak{F} \quad \text{dla } c \geq 0.$$

Ćwiczenie 1.1.5. Sprawdzić, że funkcja $f: S \rightarrow \mathbb{R}$ jest $\mathfrak{F}$ -mierzalna wtedy i tylko wtedy, gdy jest $\mathfrak{F}/\mathfrak{B}_{\mathbb{R} \setminus \{0\}}$ -mierzalna.

Stwierdzenie 1.1.18. Zbiór $\mathcal{M}(\mathfrak{F})$ skończonych funkcji $\mathfrak{F}$ -mierzalnych jest przestrzenią liniową. Z dokładnością do naturalnego izomorfizmu, przestrzeń $\mathcal{M}(\mathfrak{F})$ nie zależy od wspólnej dziedziny funkcji $\mathfrak{F}$ -mierzalnych. $\square$

Wspomniany tu naturalny izomorfizm wiąże się z faktem, iż poza zbiorem $S_0 := \bigcup \mathfrak{F}$ każda funkcja $\mathfrak{F}$ -mierzalna f jest równa 0 i w tym sensie stanowi jednoznaczne rozszerzenie ograniczenia $f|_{S_0}$. Jeśli $\mathfrak{F}$ nie jest σ -ciałem, przeciwbraz $f^{-1}(0) \notin \mathfrak{F}$ nie jest zbiorem mierzalnym.

Ćwiczenie 1.1.6. Funkcja $f: S \rightarrow \overline{\mathbb{R}}$ jest $\mathfrak{F}$ -mierzalna wtedy i tylko wtedy, gdy jest mierzalna względem σ -ciała $\sigma_S(\mathfrak{F})$ generowanego przez $\mathfrak{F}$, a zbiór $\{x; f(x) \neq 0\}$ należy do $\mathfrak{F}$.

Przykład 1.1.7. Funkcji rzeczywiste na (dowolnym) zbiorze $T \neq \emptyset$ mogą być uważane za mieralne – o ile wspólną dziedzinę traktować jako przestrzeń mieralną i rozważać wraz z odpowiednim σ –ciałem lub σ –pierścieniem. Dla ustalonej przestrzeni liniowej $\mathcal{A} \subset \mathbb{R}^T$ funkcji rzeczywistych na T największym takim σ –ciałem jest oczywiście rodzina $\mathfrak{P}(T)$ wszystkich podzbiorów T , a najmniejszym σ –pierścieniem – rodzina

$$\mathcal{A}^\sigma = \sigma(\mathcal{A}^f), \quad \text{gdzie } \mathcal{A}^f = \{f^{-1}[1, \infty); f \in \mathcal{A}\}; \quad (1.7)$$

w szczególności zachodzi inkluzja $\mathcal{A} \subset \mathcal{M}(\mathcal{A}^\sigma)$.

Definicja 1.1.19. Rodziny zbiorów $\mathcal{A}^\sigma$ oraz $\sigma\mathcal{A} = \sigma_T(\mathcal{A}^f)$ nazywać będziemy odpowiednio σ –pierścieniem i σ –ciałem generowanym przez przestrzeń $\mathcal{A}$.

1.1.2 Czym mierzyć zbiory?

Definicja 1.1.20. Miarą w przestrzeni (pół-)mieralnej $(T, \mathfrak{F})$ nazywamy każdą przeliczalnie addytywną funkcję $\mu: \mathfrak{F} \rightarrow [0, \infty]$ równą 0 na zbiorze pustym $\emptyset$. Trójkę $(T, \mathfrak{F}, \mu)$ nazywamy w tym przypadku przestrzenią z miarą.

Definicja 1.1.21. Zbiór mieralny A nazywamy σ –skończonym (względem miary μ), jeśli jest sumą ciągu zbiorów o mierze skończonej. Miarę μ uznajemy za σ –skończoną, gdy istnieje zbiór σ –skończony $S \in \mathfrak{F}$, taki że dla każdego zbioru $A \in \mathfrak{F}$ zachodzi równość $\mu(A \setminus S) = 0$.

W analogiczny sposób definiuje się miary skończone – równe 0 poza zbiorem o skończonej mierze. Oczywiście, miara na σ –ciele w T jest skończona (σ –skończona) wtedy i tylko wtedy, gdy przestrzeń T ma miarę skończoną (odp. σ –skończoną).

Przykład 1.1.8. Podstawową i raczej banalną miarą skończoną jest *miara Diraca* δ_x o wartościach 1 lub 0 zależnie od tego, czy mierzony zbiór zawiera ustalony punkt x . Nieco ogólniejsza i niekoniecznie skończona jest *miara zliczająca* ρ_S (względem zbioru S), dla której

$$\rho_S(A) \equiv \#(A \cap S),$$

a S jest ustalonym zbiorem mieralnym. Indeks S równy całej przestrzeni pomijamy.

Uznając za zdegenerowane zbiory mieralne typu $0/\infty$, czyli zbiory o mierze nieskończonej niezawierające podzbiorów o skończonej mierze niezerowej, przyjmujemy także

Definicja 1.1.22. Zbiór $A \in \mathfrak{T}$ niezawierający żadnego zbioru typu $0/\infty$ będziemy uważać za pólskończony względem miary μ .

Przestrzeń z miarą $(T, \mathfrak{T}, \mu)$, a także samą miarę μ określoną na σ -pierścieniu $\mathfrak{T}$ nazywamy pólskończoną, jeśli dziedzina $\mathfrak{T}$ miary składa się wyłącznie ze zbiorów pólskończonych.

Dla dowolnej przestrzeni z miarą $(T, \mathfrak{T}, \mu)$ przyjmujemy oznaczenia

$$\mathfrak{T}_\mu^f := \{A \in \mathfrak{T}; \mu(A) < \infty\}, \quad \text{oraz} \quad \mathfrak{T}_\mu^\sigma := \sigma(\mathfrak{T}_\mu^f) \quad (1.8)$$

$$\mathfrak{T}_\mu^{loc} := \{A \in \mathfrak{T}; \forall C \in \mathfrak{T}|_A (\mu(C) = \infty \implies \exists B \in \mathfrak{T}|_C 0 < \mu(B) < \infty)\} \quad (1.9)$$

dla δ -pierścienia zbiorów o mierze skończonej $\mathfrak{T}_\mu^f$, σ -pierścienia zbiorów o mierze σ -skończonej (σ -skończonych) $\mathfrak{T}_\mu^\sigma$ i, odpowiednio, σ -pierścienia zbiorów o mierze pólskończonej (*pólskończonych*) $\mathfrak{T}_\mu^{loc} \supset \mathfrak{T}_\mu^\sigma$.

Definicja 1.1.23. Dla dowolnej miary μ w przestrzeni $(T, \mathfrak{T})$, składową σ -skończoną będziemy nazywać miarę $\mu^\sigma: \mathfrak{T}_\mu^\sigma \rightarrow [0, \infty]$ na σ -pierścieniu zbiorów σ -skończonych, będącą ograniczeniem miary μ do $\mathfrak{T}_\mu^\sigma \subset \mathfrak{T}$.

Ograniczenie $\mu^{loc}: \mathfrak{T}_\mu^{loc} \rightarrow [0, \infty]$ miary μ do σ -pierścienia zbiorów pólskończonych jest pólskończoną składową miarą.

Naturalna własność lokalizacji dla równości prawie wszędzie

$$f \stackrel{\text{pw}}{=} g \iff \forall K \in \mathfrak{T}_\mu^f \quad f|_K \stackrel{\text{pw}}{=} g|_K, \quad (1.10)$$

prawdziwa dla funkcji $f, g \in \mathcal{M}(\mathfrak{T}_\mu^{loc})$, uzasadnia przyjęte niżej wyróżnienie.

Definicja 1.1.24. Dla dowolnej przestrzeni z miarą $(T, \mathfrak{T}, \mu)$ podzbiory

$$\begin{aligned} \overline{\mathcal{M}}(T, \mu) &= \{f \in \overline{\mathcal{M}}(\mathfrak{T}); \{x \in T; f(x) \neq 0\} \in \mathfrak{T}_\mu^{loc}\} \\ \mathcal{M}(T, \mu) &= \mathcal{M}(\mathfrak{T}_\mu^{loc}) \end{aligned}$$

nazywamy odpowiednio zbiorem funkcji μ -mierzalnych oraz przestrzenią liniową skończonych funkcji μ -mierzalnych.

Stwierdzenie 1.1.25. Jeśli $g: (S, \mathfrak{S}) \rightarrow (T, \mathfrak{T})$ jest odwzorowaniem mierzalnym przestrzeni półmierzalnych, to złożenie $g^*: \overline{\mathcal{M}}(\mathfrak{T}) \ni f \mapsto f \circ g \in \overline{\mathcal{M}}(\mathfrak{S})$ zachowuje funkcje skończone, a dla dowolnej miary μ w $(S, \mathfrak{S})$ wzór

$$g_*\mu(A) := \mu(g^{-1}A), \quad \text{dla } A \in \mathfrak{T} \quad (1.11)$$

określa miarę $g_*\mu$ w przestrzeni $(T, \mathfrak{T})$.

Dla miary pólskończonej zachodzi inkluzja $g^*\overline{\mathcal{M}}(T, g_*\mu) \subset \overline{\mathcal{M}}(S, \mu)$, przy czym

$$g^*: \mathcal{M}(T, g_*\mu) \rightarrow \mathcal{M}(S, \mu) \quad (1.12)$$

jest homomorfizmem przestrzeni liniowych. $\square$

Definicja 1.1.26. *Miarę $g_*\mu$ w przestrzeni $(T, \mathfrak{T})$ spełniającą warunek (1.11) nazywamy miarą indukowaną (z μ) przez odwzorowanie $\mathfrak{S}/\mathfrak{T}$ –mieralne $g: S \rightarrow T$. Homomorfizm g^* nazywamy homomorfizmem indukowanym.*

Definicja 1.1.27. *Dla dowolnej miary μ w (pół-)przestrzeni $(T, \mathfrak{T})$ i podprzestrzeni $S \subset T$ takiej, że $\mathfrak{T}|S \subset \mathfrak{T}$, za ograniczenie miary do podprzestrzeni S przyjmujemy obcięcie $\mu|S \equiv \mu|_{\mathfrak{T}|S}$.*

Uwaga 1.1.9. Obliczenia związane z miarą korzystają w istocie z jej części skończonej, określonej na δ –pierścieniu $\mathfrak{T}_\mu^f$. Rola symbolu ∞ przyjęta dla oznaczenia granicy ciągu liczbowego lub sumy szeregu – wprawdzie rozbieżnego, ale w „kontrolowany” sposób – wynika z naturalnej monotoniczności rozważanej operacji i oznacza wartość *większą* od każdej liczby rzeczywistej. Dla miary σ –skończonej możemy zatem mówić o *efektywnym* pomiarze całości badanego zbioru – włączając w proces pomiaru przeliczalne ciągi *skończonych* operacji. Wyróżniając σ –pierścień $\mathfrak{T}_\mu^\sigma$ zbiorów o mierze σ –skończonej, traktujemy zbiory należące do dopełnienia $\mathfrak{T} \setminus \mathfrak{T}_\mu^\sigma$ jako *nieosiągalne*, których mierzalność jest częściowo deklaratywna, a *formalna* miara jest z konieczności nieskończona. Warunek ten dopuszcza do rozważań zbiory półskończone (1.9) i związane z nimi prawie wszędzie skończone funkcje mieralne o własności lokalizacji (1.10), natomiast σ –skończona składowa miary definiuje całkę Lebesgue’a i rozstrzyga o całkowalności funkcji mieralnych.

Stwierdzenie 1.1.28. *Dowolna miara jest funkcją addytywną i σ –subaddytywną. Każda miara μ jest monotonicznie ciągła (dokł. półciągła) z dołu – w sensie tożsamości*

$$\mu(\bigcup_{k \in \mathbb{N}} A_k) = \lim_{k \rightarrow \infty} \mu(A_k) = \sup_k \mu(A_k), \quad (1.13)$$

dla dowolnego wstępującego ciągu zbiorów mierzalnych $(A_k)_{k \in \mathbb{N}}$. Miara skończona jest także (pół-)ciągła z góry, tzn. analogiczna własność

$$\mu(\bigcap_{k \in \mathbb{N}} A_k) = \lim_{k \rightarrow \infty} \mu(A_k) = \inf_k \mu(A_k), \quad (1.14)$$

zachodzi dla ciągów zstępujących – jeśli $\mu(A_k) < \infty$ dla pewnego $k \in \mathbb{N}$. $\square$

Uwaga 1.1.10. Zgodnie z [2, lem. 1.3.12], zarówno ciągłość z dołu jak i ciągłość z góry (dla przekroju $\bigcap A_k = \emptyset$) jest także warunkiem dostatecznym σ –skończoności miary addytywnej.

Podstawową, klasyczną własność dotyczącą zgodności miar o wspólnej dziedzinie zawiera [2, tw. 1.3.14]:

Twierdzenie 1.1.29 (O jednoznaczności miar). *Niech μ będzie miarą w przestrzeni mierzalnej $(T, \mathfrak{T})$, a $\mathfrak{C} \subset \mathfrak{T}$ – dowolnym π –systemem złożonym ze zbiorów o skończonej mierze. Dla każdej miary ν w $(T, \mathfrak{T})$ przyjmującej na zbiorach rodziny $\mathfrak{C}$ te same wartości co μ zachodzi równość*

$$\mu(A) = \nu(A) \quad \text{dla } A \in \sigma(\mathfrak{C}). \quad \square$$

W przypadku, gdy obie miary są zupełne, równość rozszerza się także na pojmowane w opisany niżej sposób uzupełnienie σ –pierścienia $\sigma(\mathfrak{C})$.

Definicja 1.1.30. *Przestrzeń z miarą $(T, \mathfrak{T}, \mu)$ i samą miarą μ nazywamy zupełną, jeśli podzbiory dowolnego zbioru $S \in \mathfrak{T}$ o mierze $\mu(S) = 0$ są wszystkie zbiorami mierzalnymi (również o mierze 0).*

Dla dowolnej miary $\mu: \mathfrak{T} \rightarrow [0, \infty]$ najmniejsze rozszerzenie do miary zupełnej $\mu^c: \mathfrak{T}^c \rightarrow [0, \infty]$ nazywamy jej uzupełnieniem. Przestrzeń z miarą zupełną $(T, \mathfrak{T}^c, \mu^c)$ jest wówczas uzupełnieniem przestrzeni $(T, \mathfrak{T}, \mu)$.

Istnienie i jednoznaczność uzupełnienia miary precyzuje elementarne

Stwierdzenie 1.1.31. *Dla dowolnej przestrzeni mierzalnej $(T, \mathfrak{T}, \mu)$ istnieje dokładnie jedna przestrzeń z miarą zupełną $(T, \mathfrak{T}^c, \mu^c)$ określoną na σ –pierścieniu zawierającym $\mathfrak{T}$ i taką, że każda miara zupełna będąca rozszerzeniem miary μ jest także rozszerzeniem miary μ^c .*

Uzupełnienie $\mathfrak{T}^c$ składa się ze zbiorów $S \subset T$, dla których istnieją mierzalne $A, B \in \mathfrak{T}$ takie, że $A \subset S \subset B$ i $\mu(B \setminus A) = 0$. $\square$

Przykład 1.1.11. Na prostej $\mathbb{R}$ istnieje dokładnie jedna miara ℓ na σ –ciele zbiorów borelowskich $\mathfrak{B}_{\mathbb{R}}$ spełniająca warunek

$$\ell[a, b] = b - a \quad \text{dla } a, b \in \mathbb{R}, \ a \leq b,$$

czyli określona jako długość dla skończonych przedziałów domkniętych. Ten sam symbol rezerwujemy także dla miary Lebesgue'a będącej jednoznacznym uzupełnieniem miary $\ell: \mathfrak{B}_{\mathbb{R}} \rightarrow [0, \infty]$.

Klasyczna konstrukcja miary Lebesgue'a [2, par. 1.3.2] wykorzystuje analogiczny wzór dla miary przedziałów postaci $[a, b)$, $a < b$, rozszerzony następnie do miary na pierścieniu złożonym ze skończonych sum przedziałów domknięto–otwartych.

1.1.3 Co zapożyczamy od Carathéodory'ego?

W opracowanej przez Carathéodory'ego konstrukcji za punkt wyjścia przyjmujemy dowolną rodzinę $\mathfrak{A} \ni \emptyset$ zbiorów zawartych w ustalonej przestrzeni (zbiorze) T oraz funkcję $\nu: \mathfrak{A} \rightarrow [0, \infty]$, będącą wstępnym przybliżeniem miary wybranych zbiorów.

Definicja 1.1.32. Dla dowolnej funkcji $\nu: \mathfrak{A} \rightarrow [0, \infty]$ takiej, że $\nu(\emptyset) = 0$, funkcję $\nu^*: \mathfrak{P}(T) \rightarrow [0, \infty]$ określoną poprzez formułę

$$\nu^*(B) = \inf \left\{ \sum_{n \in \mathbb{N}} \nu(A_n); \forall n \in \mathbb{N} A_n \in \mathfrak{A} \text{ i } B \subset \bigcup_n A_n \right\} \quad \text{dla } B \subset T, \quad (1.15)$$

nazywać będziemy miarą zewnętrzną w T wyznaczoną przez ν . Zakładamy przy tym, że kresem górnym zbioru pustego (w przypadku braku stosownego pokrycia zbioru B) jest nieskończoność ∞ .

Definicja 1.1.33. Każdą funkcję σ -subaddytywną $\mu: \mathfrak{P}(T) \rightarrow [0, \infty]$ spełniającą warunek $\mu(\emptyset) = 0$ będziemy nazywać miarą zewnętrzną Carathéodory'ego lub – po prostu – miarą zewnętrzną.

Uwaga 1.1.12. Równość $\nu^*(A) = \nu(A)$ dla wszystkich zbiorów $A \in \mathfrak{A}$ zachodzi wtedy i tylko wtedy, gdy funkcja ν jest σ -subaddytywna. Miara zewnętrzna ν^* wyznaczona przez funkcję ν jest zawsze σ -subaddytywna. Każda miara zewnętrzna jest zatem postaci (1.15), aczkolwiek sposób otrzymania konkretnej miary może być odmienny.

Definicja 1.1.34. Dla dowolnej miary zewnętrznej μ w zbiorze T podzbiór $A \subset T$ jest μ -mierzalny (w sensie Carathéodory'ego), jeśli spełnia warunek

$$\mu(Z \cap A) + \mu(Z \setminus A) = \mu(Z) \quad \text{dla każdego podzbioru } Z \subset T. \quad (1.16)$$

W warunku Carathéodory'ego (1.16) można się ograniczyć do zbiorów Z o skończonej mierze zewnętrznej – w przeciwnym wypadku odpowiednia równość ($\infty = \infty$) wynika z subaddytywności μ . Co więcej, prawdziwe jest

Stwierdzenie 1.1.35. (i) Dowolny zbiór $A \subset T$ jest mierzalny względem miary zewnętrznej $\nu^*: \mathfrak{P}(T) \rightarrow [0, \infty]$ wyznaczonej przez σ -subaddytywną funkcję $\nu: \mathfrak{A} \rightarrow [0, \infty]$ wtedy i tylko wtedy, gdy spełnia warunek

$$\nu^*(B \cap A) + \nu^*(B \setminus A) \leq \nu(B) \quad \text{dla } B \in \mathfrak{A}.$$

(ii) Miara zewnętrzna ν^* wyznaczona przez dowolną miarę $\nu: \mathfrak{F} \rightarrow [0, \infty]$ na pierścieniu $\mathfrak{F} \subset \mathfrak{P}(T)$ jest jej rozszerzeniem, a wszystkie zbiory należące do $\mathfrak{F}$ są ν^* -mierzalne. $\square$

Pora przywołać także – bez dowodu (por. [9], [18]) – klasyczne

Twierdzenie 1.1.36 (Carathéodory). Rodzina $\mathfrak{E}_\mu$ wszystkich zbiorów mierzalnych względem dowolnej miary zewnętrznej $\mu: \mathfrak{P}(T) \rightarrow [0, \infty]$ jest σ -ciałem, przy czym ograniczenie miary zewnętrznej

$$\mu_{\mathfrak{E}} = \mu|_{\mathfrak{E}_\mu}: \mathfrak{E}_\mu \rightarrow [0, \infty] \quad (1.17)$$

jest miarą zupełną.

Definicja 1.1.37. Dla dowolnej nieujemnej funkcji ν określonej na niepustej rodzinie podzbiorów zbioru T i równej 0 na zbiorze $\emptyset$, ograniczenie $\nu_{\mathfrak{C}}$ miary zewnętrznej ν^* do σ -ciała $\mathfrak{C}_{\nu} := \mathfrak{C}_{\nu^*}$ będziemy nazywać miarą Carathéodory'ego wyznaczoną przez ν .

Uwaga 1.1.13. π -system $\mathfrak{I}$, dla którego każda z różnic $A \setminus B$, $A, B \in \mathfrak{I}$, jest skończoną sumą rozłącznych zbiorów z $\mathfrak{I}$, nazywamy *półpierścieniem*. W takiej sytuacji rodzina $\mathfrak{I}$ utworzona ze skończonych sum zbiorów z $\mathfrak{I}$ jest pierścieniem, a każda addytywna funkcja $\mu: \mathfrak{I} \rightarrow [0, \infty)$ rozszerza się do skończenia addytywnej miary na $\mathfrak{I}$. Zgodnie z [2, tw. 1.5.12–1.5.13] zachodzi

Twierdzenie 1.1.38 (Hahn–Kołmogorow). Każda σ -addytywna miara μ na półpierścieniu $\mathfrak{I}$ w przestrzeni T rozszerza się do miary $\tilde{\mu}$ na σ -ciele $\sigma_T(\mathfrak{I}) \supset \mathfrak{I}$. Miara $\tilde{\mu}$ jest wyznaczona jednoznacznie, jeśli T jest sumą przeliczalnej rodziny zbiorów z $\mathfrak{I}$. $\square$

Półpierścień zbiorów postaci $A \times B$, gdzie $A \in \mathfrak{S}$, $B \in \mathfrak{T}$, jest także techniczną podstawą dla definicji i konstrukcji produktu miar w przestrzeni mierzalnej $(S, \mathfrak{S}) \times (T, \mathfrak{T})$. Ogólnie, dla dowolnych przestrzeni z miarą $(S_i, \mathfrak{S}_i, \mu_i)$, $i \leq n$, przyjmujemy

Definicja 1.1.39. Miarą produktową lub produktem miar σ -skończonych μ_i , $i \leq n$, nazywamy jedyną miarę μ w przestrzeni $\prod_{i \leq n} (S_i, \mathfrak{S}_i)$ spełniającą warunek

$$\mu(A_1 \times \cdots \times A_n) = \mu_1(A_1) \cdot \cdots \cdot \mu_n(A_n), \quad (1.18)$$

dla dowolnych zbiorów o skończonej mierze $A_i \in \mathfrak{S}_i$, $i = 1, \dots, n$. Produkt miar będziemy oznaczać $\mu_1 \otimes \cdots \otimes \mu_n$.

Dodatkowy warunek podany w [2, tw. 2.3.11] i postulujący, by rodziną zbiorów σ -skończonych dla μ był σ -pierścień $\bigotimes_{i \leq n} \mathfrak{S}_i^{\sigma}$, pozwala zrezygnować z założenia σ -skończoności miar.

Przyjmujemy oznaczenie $\ell^n = \bigotimes_{i \leq n} \ell$ dla miary Lebesgue'a w $\mathbb{R}^n$.

1.1.4 Lokalnie skończone miary borelowskie

Definicja 1.1.40. Miarą borelowską w przestrzeni topologicznej T nazywamy dowolną miarę określoną na σ -ciele zbiorów borelowskich $\mathfrak{B}_T$.

Miarę borelowską uważa się za lokalnie skończoną, jeśli każdy punkt przestrzeni ma otoczenie o skończonej mierze.

W przypadku przestrzeni euklidesowej możemy się spodziewać bardziej jawnej charakterystyki miar.

Definicja 1.1.41. Dla dowolnej lokalnie skończonej miary borelowskiej μ na prostej $\mathbb{R}$ funkcję niemalejącą $F_\mu: \mathbb{R} \rightarrow \mathbb{R}$ taką, że $\mu(a, b] = F_\mu(b) - F_\mu(a)$, dla dowolnych liczb $a < b$, nazywamy dystrybuantą (ang. cumulative distribution function) miary μ .

Dystrybuanty miar na prostej charakteryzuje prawostronna ciągłość: $F(x+) := \inf \{F(t); t > x\} = F(x)$ dla $x \in \mathbb{R}$. Rozszerzając [2, Tw. 1.3.20], proponujemy

Stwierdzenie 1.1.42. (i) Związek lokalnie skończonej miary μ z jej dystrybuantą spełniającą dodatkowy warunek $F_\mu(0) = 0$ jest wzajemnie jednoznaczny:

$$F_\mu(x) = \begin{cases} \mu((0, x]), & \text{dla } x \geq 0, \\ 0, & \text{dla } x = 0, \\ -\mu((x, 0]), & \text{dla } x < 0. \end{cases} \quad (1.19)$$

(ii) Każda prawostronnie ciągle niemalejąca funkcja $F: \mathbb{R} \rightarrow \mathbb{R}$ jest dystrybuantą miary $\mu = Q_*^F \ell$ indukowanej z miary Lebesgue'a ℓ przez funkcję kwantylową określoną wzorem

$$Q^F(u) = \inf \{x \in \mathbb{R}; u \leq F(x)\} \quad \text{dla } u \in (\inf F, \sup F). \quad (1.20)$$

Dowód. Z monotoniczności F wynika, że dla każdego $u \in (\inf F, \sup F)$ zbiór $\{x; u \leq F(x)\}$ jest przedziałem nieskończonym o początku równym $Q^F(u)$. Ponieważ prawostronna ciągłość zapewnia, że jest to przedział domknięty, stwierdzamy równoważność

$$Q^F(u) \leq x \iff u \leq F(x) \quad \text{dla } u \in (\inf F, \sup F), x \in \mathbb{R}. \quad (1.21)$$

Stąd $a < Q^F(u) \leq b \iff F(a) < u \leq F(b)$, a więc

$$\ell\{u; a < Q^F(u) \leq b\} = \ell((F(a), F(b)]) = F(b) - F(a)$$

dla $a < b$, $a, b \in \mathbb{R}$. □

Dla miar skończonych za odpowiedni jednoznaczny „punkt zaczepienia” przyjmuje się najczęściej własność $\lim_{x \rightarrow -\infty} F(x) = 0$ (zamiast $F(0) = 0$). W ogólniejszej sytuacji przyjmujemy zatem

Definicja 1.1.43. Dla $p \in \mathbb{N}$ dystrybuantą skończonej miary borelowskiej μ w $\mathbb{R}^p$ nazywamy funkcję

$$F_\mu(x) = \mu((-\infty, x_1] \times \dots \times (-\infty, x_p]), \quad \text{dla } x = (x_1, \dots, x_p) \in \mathbb{R}^p. \quad (1.22)$$

Podstawową własność dystrybuanty podaje

Lemat 1.1.44. *Dla dowolnego niepustego p -wymiarowego przedziału postaci $A = (a_1, b_1] \times \dots \times (a_p, b_p]$ zachodzi równość*

$$\mu(A) = \sum_{c \in \{a_1, b_1\} \times \dots \times \{a_p, b_p\}} \text{sign}(c) F_\mu(c), \quad (1.23)$$

gdzie znak $\text{sign}(c)$ odpowiada parzystości liczby indeksów $i \leq p$, dla których $c_i = a_i$.

Dla danej funkcji $F: \mathbb{R}^p \rightarrow [0, 1]$ i dowolnego otwarto-domkniętego przedziału $A \subset \mathbb{R}^p$ naprzemienną sumę stojącą po prawej stronie równości (1.23) oznacza się symbolem $\Delta_A F$.

Twierdzenie 1.1.45. *Funkcja $F: \mathbb{R}^p \rightarrow [0, 1]$ jest dystrybuantą skończonej miary borelowskiej w $\mathbb{R}^p$ wtedy i tylko wtedy, gdy spełnia następujące warunki:*

- (i) *F jest niemalejąca ze względu na każdą ze zmiennych.*
- (ii) *$\lim_{x_i \rightarrow -\infty} F(x) = 0$ dla każdego $i \leq p$, $\lim_{x \rightarrow \infty} F(x) < \infty$.*
- (iii) *$F(x) = \lim_{y \searrow x} F(y)$ dla każdego $x \in \mathbb{R}^p$, gdzie zbieżność ogranicza warunek $y_i \geq x_i$ dla każdej współrzędnej $i \leq p$.*
- (iv) *$\Delta_A F \geq 0$ dla każdego otwarto-domkniętego przedziału $A \subset \mathbb{R}^p$.*

Dowód. Ważne przeglądowe ćwiczenie z teorii miary. Monotoniczność (i) jest prostą konsekwencją nierówności (iv) i własności (ii). $\square$

Struktura zbioru punktów ciągłości p -wymiarowej dystrybuanty F_μ jest bardziej skomplikowana niż w przypadku 1-wymiarowym. Niewątpliwie, dla każdej współrzędnej $i \leq p$ wśród hiperpłaszczyzn

$$H_c^i = \{x \in \mathbb{R}^p; x_i = c\}, \quad c \in \mathbb{R}, \quad (1.24)$$

co najwyżej przeliczalna podrodzina ma miarę $\mu(H_c^i) > 0$. Własność ta jest wykorzystywana w kolejnym podrozdziale – w kontekście słabej zbieżności.

Stwierdzenie 1.1.46. *Dystrybuanta F_μ dowolnej skończonej miary borelowskiej w $\mathbb{R}^p$ jest ciągła poza sumą przeliczalnej rodziny hiperpłaszczyzn (1.24), dla których $\mu(H_c^i) > 0$.*

Dowód. Załóżmy, że punkt $x = (x_1, \dots, x_p)$ nie należy do żadnej z wyróżnionych hiperpłaszczyzn, a mimo to funkcja F_μ jest nieciągła w x . Oznacza to istnienie ciągu $(y_n)_{n \in \mathbb{N}}$ zbieżnego do x i takiego, że $F_\mu(x) - F_\mu(y_n) \not\rightarrow 0$.

Przechodząc ewentualnie do podciągów możemy założyć, że każdy z ciągów utworzonych przez kolejne współrzędne $y_{ni} \rightarrow x_i$ mieści się w jednym przedziałach $(-\infty, x_i)$ lub $[x_i, \infty)$, dla $i \leq p$. Zmieniając odpowiednio numerację zmiennych zakładamy zatem, że dla pewnego $k \leq p$ zachodzą nierówności

$$y_{n1} < x_1, \dots, y_{nk} < x_k, y_{n,k+1} \geq x_{k+1}, \dots, y_{np} \geq x_p$$

dla wszystkich $n \in \mathbb{N}$, przy czym oczywiście $F_\mu(x) - F_\mu(y_n) \rightarrow 0$. Przyjęte oznaczenia i nierówności pozwalają na oszacowanie

$$\begin{aligned} |F_\mu(x) - F_\mu(y_n)| &\leq |F_\mu(x) - F_\mu(x_1 \dots x_k, y_{n,k+1} \dots y_{n,p})| \\ &\quad + |F_\mu(x_1 \dots x_{k-1}, x_k, y_{n,k+1} \dots y_{n,p}) - F_\mu(x_1 \dots x_{k-1}, y_{nk} \dots y_{n,p})| \\ &\quad + \dots + |F_\mu(x_1, y_{n2} \dots y_{n,p}) - F_\mu(y_{n1}, y_{n2} \dots y_{n,p})| \\ &\leq \mu\{\xi; x_i < \xi_i \leq y_{ni} \text{ dla } i > k\} + \mu\{\xi; y_{nk} < \xi_k \leq x_k\} \\ &\quad + \dots + \mu\{\xi; y_{n1} < \xi_1 \leq x_1\} \rightarrow 0 + \mu(H_{x_k}^k) + \dots + \mu(H_{x_1}^1) = 0. \end{aligned}$$

Otrzymana sprzeczność wyklucza nieciągłość funkcji F_μ w punkcie x . $\square$

W punktach nieciągłości (i – oczywiście – w punktach ciągłości) dystrybuanta F miary μ ma pełen zestaw *granic lewostronnych* zdefiniowanych jako stosowny kres górny:

$$F(\dots, t_{i_1}^-, \dots, t_{i_k}^-, \dots) = \sup_{t'_{i_1} < t_{i_1}, \dots, t'_{i_k} < t_{i_k}} F(\dots, t'_{i_1}, \dots, t'_{i_k}, \dots) \quad (1.25)$$

dla dowolnych podzbiorów indeksów $1 \leq i_1 < \dots < i_k \leq p$ i dowolnych $t = (t_i)_{i \leq p} \in \mathbb{R}^p$.

Ćwiczenie 1.1.14. Sformułować (i wykazać) analogiczny do (1.23) wzór na miarę $\mu(A)$ dowolnego p -wymiarowego przedziału $A = J_1 \times \dots \times J_p$. Wzór powinien wykorzystywać lewostronne granice dystrybuanty F w tych przypadkach, gdy multiindeks $c = (c_1, \dots, c_p)$ korzysta z prawego końca $c_i = b_i$ przedziału prawostronnie otwartego lub – z lewego końca $c_i = a_i$ przedziału lewostronnie domkniętego.

1.1.5 Słaba zbieżność w przestrzeni miar

W zbiorze wszystkich miar skończonych określonych w ustalonej przestrzeni mierzalnej $(M, \mathfrak{M})$ dowolna liniowa kombinacja miar o nieujemnych współczynnikach jest także miarą. Przestrzenie rozszerzające tę strukturę i obejmujące uogólnienia miar badane są – w różnych aspektach – w dalszych

rozdziałach kursu, w szczególności w ramach teorii Lebesgue’a–Radona–Nikodyma. Przypadek miar borelowskich, gdy M jest przestrzenią topologiczną i $\mathfrak{M} = \mathfrak{B}_M$, rozwijany jest dla miar Radona w rozdziale 3.2, a w obrębie probabilistyki – dla rozkładów zmiennych losowych. Stosowna topologia w przestrzeniach skończonych, znormalizowanych miar borelowskich opiera się na tzw. *słabej* zbieżności ciągu miar, co w klasycznych kursach oznacza zbieżność *ciągu całek Lebesgue’a*. Proponujemy bardziej elementarne podejście – bez odwoływania się do niedostępnego w tym miejscu (jeszcze) pojęcia całki, z wykorzystaniem regularności (1.26). Przytaczamy za [2, par. 1.4.3] specyficzny układ własności, jakie rozważana rodzina miar ma w przestrzeniach metrycznych.

Twierdzenie 1.1.47 (Regularność miar borelowskich). *Niech μ będzie dowolną lokalnie skończoną miarą borelowską w przestrzeni metrycznej M .*

(i) *Jeśli M jest sumą ciągu zbiorów otwartych o skończonej mierze, to*

$$\mu(A) = \inf_{A \subset U} \mu(U) = \sup_{F \subset A} \mu(F) \quad \text{dla } A \in \mathfrak{B}_M, \quad (1.26)$$

gdzie U przebiega rodzinę wszystkich nadzbiorów otwartych A , a F – rodzinę podzbiorów domkniętych.

(ii) *Jeśli M jest ponadto ośrodkową przestrzenią zupełną, to także*

$$\mu(A) = \sup_{K \subset A} \mu(K) \quad \text{dla } A \in \mathfrak{B}_M, \quad (1.27)$$

gdzie K przebiega rodzinę podzbiorów zwartych. □

Założenie podane w (i) spełnione są, w szczególności, gdy miara μ jest skończona lub gdy przestrzeń M jest ośrodkowa. Temat regularności miar borelowskich wraca w końcowej części niniejszej monografii w podrozdziale 3.3.

Definicja 1.1.48. *Niech (M, d) będzie dowolną przestrzenią metryczną. Ciąg skończonych miar borelowskich $(\mu_n)_{n \in \mathbb{N}}$ nazywamy słabo zbieżnym do miary borelowskiej μ , co zapisujemy jako $\mu_n \xrightarrow{w} \mu$, jeśli spełnia warunki*

$$\left. \begin{aligned} &\bullet \liminf_n \mu_n(U) \geq \mu(U), \quad \text{dla podzbiorów otwartych } U \subset M \\ &\bullet \limsup_n \mu_n(F) \leq \mu(F), \quad \text{dla podzbiorów domkniętych } F \subset M. \end{aligned} \right\} \quad (1.28)$$

Z warunków (1.28) wynika zbieżność $\mu_n(M) \rightarrow \mu(M)$, a dla miar o stałej wartości $\mu(M) = \mu_n(M) < \infty$ podane warunki są równoważne. Brak równości w (1.28) oznacza, że sprawdzenia wymaga pożądana jednoznaczność miary granicznej.

Definicja 1.1.49. Zbiór borelowski $C \subset M$ nazywamy zbiorem ciągłości miary $\mu: \mathfrak{B}_M \rightarrow [0, \infty]$, jeśli $\mu(\partial C) = 0$.

Przypominamy, że $\partial C = \overline{C} \setminus \text{int}(C)$ oznacza brzeg dowolnego zbioru $C \in \mathfrak{B}_M$. Z własności operatora ∂ wynika, że zbiory ciągłości miary tworzą ciało – na tyle duże, że z porównania $\mu(\text{int } C)$ i $\mu(\overline{C})$ wynika

Stwierdzenie 1.1.50. Zbieżność $\mu_n \xrightarrow{w} \mu$ zachodzi wtedy i tylko wtedy, gdy

$$\lim_{n \rightarrow \infty} \mu_n(C) = \mu(C), \quad (1.29)$$

dla każdego zbioru ciągłości C miary μ . $\square$

Twierdzenie 1.1.51. Niech (M, d) będzie ośrodkową przestrzenią metryczną, a μ dowolną skończoną miarą borelowską w M .

(i) W przestrzeni M istnieje przeliczalna baza utworzona z kul otwartych będących zbiorami ciągłości miary μ .

(ii) Dla dowolnej przeliczalnej bazy $\mathfrak{V}$ topologii w M złożonej ze zbiorów ciągłości miary μ i dowolnego ciągu skończonych miar borelowskich $(\mu_n)_{n \in \mathbb{N}}$ w M

$$\mu_n \xrightarrow{w} \mu \iff \mu_n(U) \rightarrow \mu(U) \text{ dla } U \in \mathfrak{U} \cup \{M\}, \quad (1.30)$$

gdzie $\mathfrak{U} = \{V_1 \cap \dots \cap V_n; V_1, \dots, V_n \in \mathfrak{V}, n \in \mathbb{N}\}$ jest π -systemem generowanym przez bazę $\mathfrak{V}$. $\square$

Dołączona do $\mathfrak{U}$ przestrzeń M może być traktowana jako przekrój skończonej pustej rodziny zbiorów z bazy.

Dowód. Ćwiczenie – por. [2, tw. 1.4.30]. $\square$

Ponieważ dla dowolnej funkcji ciągłej $h: M \rightarrow N$ przeciwobrazy zbiorów otwartych są otwarte, a domkniętych – domknięte, zachodzi

Stwierdzenie 1.1.52 (Twierdzenie o odwzorowaniu dla miar). Ciągłe odwzorowanie h przestrzeni metrycznej (M, d) w przestrzeń (N, ρ) zachowuje słabą zbieżność ciągu miar, tzn.

$$\mu_n \xrightarrow{w} \mu \implies h_* \mu_n \xrightarrow{w} h_* \mu$$

dla dowolnych skończonych miar borelowskich $\mu, \mu_n, n \in \mathbb{N}$, w M . $\square$

W przypadku skończonych miar borelowskich w $\mathbb{R}^p$ przyjmujemy

Definicja 1.1.53. Dla dowolnego $p \in \mathbb{N}$ ciąg p -wymiarowych dystrybuant $(F_n)_{n \in \mathbb{N}}$ nazywamy słabo zbieżnym do dystrybuanty F , co zapisujemy

$$F_n \xrightarrow{w} F,$$

jeśli jest zbieżny w każdym punkcie ciągłości funkcji F .

Ograniczenie zbieżności do punktów ciągłości wyjaśnia np. fakt, iż dla każdego ciągu $(c_n)_{n \in \mathbb{N}}$ zbieżnego w $\mathbb{R}^p$ do 0 i dowolnej dystrybuanty F słaba zbieżność $F(\cdot + c_n) \xrightarrow{w} F$ pozwala omijać punkty, w których nieciągłość implikuje brak zbieżności punktowej. Przypominamy, że zbiór punktów ciągłości dystrybuanty jest gęsty (stw. 1.1.46), a jego dopełnienie zawiera się w sumie przeliczalnego zbioru hiperpłaszczyzn.

Twierdzenie 1.1.54 (O równoważnej zbieżności). *Dowolny ciąg rozkładów prawdopodobieństwa $(\mu_n)_{n \in \mathbb{N}}$ w $\mathbb{R}^p$ jest słabo zbieżny do rozkładu μ wtedy i tylko wtedy, gdy ciąg dystrybuant $F_n := F_{\mu_n}$ jest słabo zbieżny do F_μ , tzn.*

$$\mu_n \xrightarrow{w} \mu \iff F_n \xrightarrow{w} F_\mu.$$

Dowód. „ $\Rightarrow$ ” Jeśli funkcja F_μ jest ciągła w $x = (x_1, \dots, x_p)$, to dla zbioru domkniętego $A = (-\infty, x_1] \times \dots \times (-\infty, x_p]$ zachodzi równość

$$\mu(A) = \mu((-\infty, x_1) \times \dots \times (-\infty, x_p)) = \mu(\text{int } A),$$

a więc $\mu(\partial A) = 0$ i $F_n(x) = \mu_n(A) \longrightarrow \mu(A) = F_\mu(x)$.

„ $\Leftarrow$ ” Przyjmując zgodne z (1.24) oznaczenie $D_i = \{c; \mu(H_c^i) > 0\}$, rozważmy dowolny zbiór postaci $A = \prod_{i \leq p} (a_i, b_i]$, gdzie $a_i, b_i \notin D_i$ dla wszystkich $i \leq p$. Z założenia wynika, że poszczególne składniki sumy $\Delta_A F_n$ dążą do $\Delta_A F_\mu$, a zatem $\mu_n(A) \rightarrow \mu(A)$, zgodnie z (1.23). Ponieważ dopełnienia $\mathbb{R} \setminus D_i$ są zbiorami gęstymi, każdy zbiór otwarty $U \subset \mathbb{R}^p$ jest sumą $U = \bigcup_{k \in \mathbb{N}} A_k$ co najwyżej przeliczalnej rodziny rozłącznych przedziałów rozważanej postaci. Nierówności

$$\mu_n(U) \geq \sum_{k \leq m} \mu_n(A_k)$$

prowadzą w granicy do oszacowania

$$\liminf_{n \rightarrow \infty} \mu_n(U) \geq \sum_{k \leq m} \mu(A_k) \quad \text{dla } m \in \mathbb{N},$$

a to implikuje słabą zbieżność ciągu miar. □

1.2 Całka Lebesgue'a

1.2.1 Podejście aksjomatyczne

Pojęcie całki jest ściśle związane z definicją dziedziny – przestrzeni liniowej skończonych funkcji całkowalnych. Jako wprowadzenie przypominamy zatem

Definicja 1.2.1. Funkcjonał liniowy $\chi: \mathcal{A} \rightarrow \mathbb{R}$ określony na przestrzeni liniowej $\mathcal{A} \subset \mathbb{R}^T$ nazywamy nieujemnym, jeśli $\chi(f) \geq 0$ dla wszystkich $f \geq 0$.

Funkcjonał nieujemny jest także monotoniczny, tzn. $f \leq g \Rightarrow \chi(f) \leq \chi(g)$.

Definicja 1.2.2. Przestrzeń liniowa $\mathcal{A}$ złożona z funkcji rzeczywistych określonych na ustalonym zbiorze T jest kratą, jeśli spełnia (równoważne) warunki $\max(f, g) \in \mathcal{A}$ i $\min(f, g) \in \mathcal{A}$ dla $f, g \in \mathcal{A}$.

Własność kraty jest równoważna implikacji

$$f \in \mathcal{A} \implies f^+, f^- \in \mathcal{A} \quad (1.31)$$

dla części dodatniej $f^+ = \max(f, 0)$ i części ujemnej $f^- = \max(-f, 0)$ dowolnej funkcji mierzalnej, przy czym prawdziwe są zależności

$$f = f^+ - f^- \quad \text{ i } \quad |f| = f^+ + f^- \quad \text{ dla } f \in \mathcal{A}. \quad (1.32)$$

Twierdzenie 1.2.3 (O całce Lebesgue'a [2, tw. 3.1.4]). Niech $(T, \mathfrak{T}, \mu)$ będzie dowolną przestrzenią z miarą. Istnieje dokładnie jeden nieujemny funkcjonal liniowy $\int d\mu: \mathcal{I}(T, \mu) \rightarrow \mathbb{R}$, $f \mapsto \int f d\mu$, o dziedzinie $\mathcal{I}(T, \mu) \subset \mathcal{M}(\mathfrak{T})$ stanowiącej kratę w przestrzeni skończonych funkcji mierzalnych, spełniający warunki:

(i) Przestrzeń $\mathcal{I}(T, \mu)$ zawiera wszystkie funkcje charakterystyczne zbiorów o skończonej mierze – i tylko te, tzn.

$$1_A \in \mathcal{I}(T, \mu) \iff \mu(A) < \infty \quad \text{ dla } A \in \mathfrak{T}. \quad (1.33)$$

Jeśli miara zbioru $A \in \mathfrak{T}$ jest skończona, to $\int 1_A d\mu = \mu(A)$.

(ii) [Monotoniczna ciągłość] Dla każdego monotonicznego ciągu $(f_n)_{n \in \mathbb{N}}$ w $\mathcal{I}(T, \mu)$, takiego że ciąg $(\int f_n d\mu)_{n \in \mathbb{N}}$ jest ograniczony, jeżeli granica $f = \lim_{n \rightarrow \infty} f_n$ jest funkcją skończoną, to $f \in \mathcal{I}(T, \mu)$ i zachodzi równość

$$\int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu. \quad (1.34)$$

(iii) [Warunek Stone'a] Jeśli $f \in \mathcal{I}(T, \mu)$, to również $\min(f, 1) \in \mathcal{I}(T, \mu)$.

Jeśli miara μ jest σ -skończona, jednoznaczność funkcjonału $\int d\mu$ zapewniają warunki (i)–(ii). $\square$

Definicja 1.2.4. W przestrzeni z miarą $(T, \mathfrak{T}, \mu)$, całką Lebesgue’a względem miary μ nazywamy funkcjonal liniowy $\int d\mu: \mathcal{I}(T, \mu) \rightarrow \mathbb{R}$ scharakteryzowany jednoznacznie w twierdzeniu 1.2.3. Skończone funkcje mierzalne tworzące przestrzeń $\mathcal{I}(T, \mu)$ nazywamy całkowalnymi lub, dokładniej, μ -całkowalnymi.

Znaczenie warunku Stone’a (iii) przybliża

Lemat 1.2.5. (i) Dla dowolnej nieujemnej funkcji mierzalnej $f \in \mathcal{M}(\mathfrak{T})$ i liczby $c > 0$, ciąg funkcji mierzalnych

$$f_n = \min(n(f - c \min(c^{-1}f, 1)), 1), \quad n \in \mathbb{N},$$

jest niemalejący, ograniczony z góry przez $c^{-1}f$ i dąży do funkcji charakterystycznej zbioru $\{x; f(x) > c\}$.

(ii) Jeśli funkcja f jest μ -całkowalna, to dla każdej liczby $c > 0$ zachodzi nierówność

$$\mu\{x; |f(x)| > c\} \leq c^{-1} \int |f| d\mu.$$

Dowód. Ćwiczenie. Sprawdzić, że

$$f_n(x) = \begin{cases} 0 & \text{gdy } 0 \leq f(x) \leq c \\ n(f(x) - c) & \text{gdy } c \leq f(x) \leq c + \frac{1}{n} \\ 1 & \text{gdy } c + \frac{1}{n} \leq f(x), \end{cases}$$

dla $x \in T$, $n \in \mathbb{N}$. □

Wniosek 1.2.6. Zachodzi inkluzja $\mathcal{I}(T, \mu) \subset \mathcal{M}(\mathfrak{T}_\mu^\sigma)$, tzn. dla każdej funkcji μ -całkowalnej f zbiór $\{f \neq 0\}$ ma miarę σ -skończoną. □

Zgodnie z [2, Stw. 3.1.9] dodatkowe, słabsze założenie μ -mierzalności funkcji μ -całkowalnych, w postaci

$$\mathcal{I}(T, \mu) \subset \mathcal{M}(T, \mu) = \mathcal{M}(\mathfrak{T}_\mu^{loc}),$$

pozwała ograniczyć warunki charakteryzujące całkę Lebesgue’a do (i)–(ii).

Pominięte tu szczegóły dowodu (i klasyczna konstrukcja całki) kierują naszą uwagę na problemy związane z nieskończoną miarą zbiorów lub nieograniczonymi wartościami całkowanej funkcji. Jako ćwiczenie proponujemy zatem wykazać

Stwierdzenie 1.2.7. Dla dowolnej miary μ przestrzeń funkcji μ -całkowalnych zawiera wszystkie ograniczone funkcje mierzalne, równe 0 poza zbiorem o mierze skończonej. □

Uprzedzając stosowne konstrukcje, *całką Lebesgue’a względem miary μ* nazywamy także opisane w twierdzeniu 1.2.17 jednoznaczne rozszerzenie funkcjonału $\int d\mu$ na nieujemne funkcje mierzalne zawarte w $\overline{\mathcal{M}}(\mathfrak{T})$, oraz funkcjonał (1.47) określony dla klas równoważnych funkcji całkowalnych.

Uwaga 1.2.1. Z twierdzenia 1.2.3 wynika, że przestrzeń funkcji μ –całkowalnych zawiera *przestrzeń funkcji prostych* $\mathcal{P}(T, \mu)$, czyli podprzestrzeń liniową generowaną przez funkcje charakterystyczne zbiorów mierzalnych o mierze skończonej. Monotoniczna ciągłość

$$\mathcal{P}(T, \mu) \ni f_n \nearrow f \in \mathcal{P}(T, \mu) \implies \tilde{\mu}(f_n) \rightarrow \tilde{\mu}(f). \quad (1.35)$$

zgodnego z (i) rozszerzenia miary do funkcjonału $\tilde{\mu}: \mathcal{P}(T, \mu) \rightarrow \mathbb{R}$, prowadzi do kolejnego rozszerzenia – na zbiór $\overline{\mathcal{M}}_+^\sigma \equiv \overline{\mathcal{M}}_+(\mathfrak{T}_\mu^\sigma)$ nieujemnych funkcji $\mathfrak{T}$ –mierzalnych *osiągalnych* z $\mathcal{P}(T, \mu)$. Przestrzeń funkcji μ –całkowalnych powstaje zatem jako specyficzne *domknięcie* przestrzeni funkcji prostych – w przestrzeni liniowej $\mathcal{M}(\mathfrak{T}_\mu^\sigma)$.

W odniesieniu do konkretnej funkcji $f \geq 0$ osiągalność oznacza σ –skończoność zbioru $\{f > 0\}$. Istotnie, jeśli dla nieujemnej funkcji mierzalnej $f: T \rightarrow [0, \infty]$ istnieje wstępujący ciąg zbiorów o skończonej mierze $T_n \in \mathfrak{T}$, $n \in \mathbb{N}$, sumujący się do zbioru $\{f > 0\}$, to f jest granicą niemalejącego ciągu funkcji prostych, np.

$$f_n = \sum_{m=0}^{n \cdot 2^n - 1} \frac{m}{2^n} \cdot 1_{\{x \in T_n; \frac{m}{2^n} < f(x) \leq \frac{m+1}{2^n}\}} + n \cdot 1_{\{x \in T_n; n < f(x)\}} \nearrow f, \quad (1.36)$$

dla $n \rightarrow \infty$. Odwrotna implikacja jest oczywista i wynika z tożsamości $\{f > 0\} = \bigcup_n \{f_n > 0\}$, jeśli tylko $f_n \nearrow f$.

1.2.2 Wnioski, konsekwencje...

Bezpośrednio z twierdzenia 1.2.3 wynika

Wniosek 1.2.8. *Jeśli miara ν w przestrzeni $(T, \mathfrak{H})$ jest rozszerzeniem miary μ określonej w $(T, \mathfrak{T})$, tzn. $\mathfrak{T} \subset \mathfrak{H}$ i zachodzi równość $\mu = \nu|_{\mathfrak{T}}$, to przestrzeń $\mathcal{I}(T, \mu)$ jest podprzestrzenią w $\mathcal{I}(T, \nu)$, a całka $\int d\nu$ jest rozszerzeniem funkcjonału $\int d\mu$. $\square$*

Przestrzeń funkcji całkowalnych jest najmniejszą kratą w $\mathcal{M}(\mathfrak{T})$ zawierającą podprzestrzeń $\mathcal{P}(T, \mu)$ oraz wszystkie skończone granice niemalejących ciągów funkcji mierzalnych $(f_n)_{n \in \mathbb{N}}$, dla których ciąg całek $\int f_n d\mu$, $n \in \mathbb{N}$, jest ograniczony. Wynika stąd, że dla każdego zbioru mierzalnego $S \in \mathfrak{T}$

i dowolnej funkcji całkowalnej $f \in \mathcal{I}(T, \mu)$ funkcja $f \cdot 1_S$ równa 0 poza S jest także μ -całkowalna. Możliwość ograniczenia całki Lebesgue’a do podprzestrzeni precyzuje

Wniosek 1.2.9. *Dla dowolnego zbioru mierzalnego S w przestrzeni z miarą $(T, \mathfrak{T}, \mu)$ zachodzi równość całek*

$$\int f|_S d(\mu|_S) = \int f \cdot 1_S d\mu \quad \text{dla } f \in \mathcal{I}(T, \mu). \quad (1.37)$$

przy czym ograniczenie $f \mapsto f|_S$ zachowuje całkowalność funkcji, podobnie jak trywialne rozszerzenie $\mathcal{I}(S, \mu|_S) \ni f \mapsto \tilde{f} \in \mathcal{I}(T, \mu)$ do funkcji znikających poza S . $\square$

Oznaczając przez $f \cdot 1_S$ funkcję równą $f(x)$ dla $x \in S$ i równą 0 dla $x \notin S$, przyjmujemy

Definicja 1.2.10. *Dla dowolnej przestrzeni z miarą $(T, \mathfrak{T}, \mu)$ i zbioru mierzalnego $S \in \mathfrak{T}$, całką Lebesgue’a względem miary μ po zbiorze S nazywamy odwzorowanie*

$$\int_S d\mu: f \mapsto \int f \cdot 1_S d\mu$$

określone dla funkcji mierzalnych f takich, że $f|_S \in \mathcal{I}(S, \mu|_S)$, czyli całkowalnych na S ,

Wiemy już, że przypadku skończonych miar borelowskich wszystkie ograniczone funkcje ciągłe są całkowalne (por. stwierdzenie 1.2.7). Niech $\mathcal{C}_b(M)$ oznacza przestrzeń metryczną ograniczonych funkcji ciągłych na przestrzeni topologicznej M . Ze względu na ścisły związek całki z pojęciem słabej zbieżności miar przywołujemy za [2, tw. 3.1.16]

Twierdzenie 1.2.11. *Niech (M, d) będzie dowolną przestrzenią metryczną, a $\mu: \mathfrak{B}_M \rightarrow \mathbb{R}_+$ skończoną miarą borelowską w M .*

(i) *Miarę μ jednoznacznie charakteryzuje funkcjonal $\int_\mu: \mathcal{C}_b(M) \rightarrow \mathbb{R}$ będący obcięciem $\int d\mu|_{\mathcal{C}_b(M)}$ całki Lebesgue’a.*

(ii) *Dla dowolnego ciągu $(\mu_n)_{n \in \mathbb{N}}$ skończonych miar borelowskich w M słaba zbieżność $\mu_n \xrightarrow{w} \mu$ jest równoważna własności*

$$\lim_{n \rightarrow \infty} \int f d\mu_n = \int f d\mu, \quad \text{dla } f \in \mathcal{C}_b(M).$$

$\square$

1.2.3 Zamiana zmiennych w całce Lebesgue'a

Z zawartej w twierdzeniu 1.2.3 jednoznaczności całki Lebesgue'a wynika charakteryzacja całki względem miary indukowanej:

Stwierdzenie 1.2.12 (Twierdzenie o zamianie zmiennych). *Dla dowolnej miary μ w przestrzeni $(S, \mathfrak{S})$ i dowolnego odwzorowania mierzalnego $g: (S, \mathfrak{S}) \rightarrow (T, \mathfrak{T})$ zachodzi równość całek*

$$\int f d(g_*\mu) = \int (f \circ g) d\mu \quad \text{dla } f \in \mathcal{I}(T, g_*\mu), \quad (1.38)$$

oraz równoważność ich dziedzin, tzn.

$$\mathcal{I}(T, g_*\mu) = \{f \in \mathcal{M}(T, g_*\mu); f \circ g \in \mathcal{I}(S, \mu)\},$$

co oznacza, że funkcja f jest $g_*\mu$ -całkowalna wtedy i tylko wtedy, gdy jest $g_*\mu$ -mierzalna i funkcja $f \circ g$ jest μ -całkowalna. $\square$

W przypadku całki względem miary Lebesgue'a ℓ^n korzysta się z oznaczenia $\int f d\ell^n = \int_{\mathbb{R}^n} f(x) dx$ lub, dla całki po odcinku, $\int_a^b f(x) dx$ – podając końce przedziału jako granice całkowania i zacierając różnicę pomiędzy całką Lebesgue'a a całką Riemanna (przykład 3.1.1).

Ogólniej, gdy potrzebny jest wzór opisujący wartości całkowanej funkcji powszechnie używana jest także alternatywna notacja

$$\int f d\mu = \begin{cases} \int f(x) \mu(dx) & \text{lub} \\ \int f(x) d\mu(x). \end{cases} \quad (1.39)$$

Mamy zatem

Wniosek 1.2.13. *Dla dowolnej miary μ i funkcji μ -całkowalnej g zachodzi równość*

$$\int g d\mu = \int x d\mu_g(x),$$

gdzie $\mu_g = g_*\mu$ jest miarą indukowaną. $\square$

Borelowska miara μ_g jest lokalnie skończona, jeśli $\mu\{x < g \leq y\} < \infty$, dla $x, y \in \mathbb{R}$. Sformułowana w stwierdzeniu 1.1.42 charakteryzacja lokalnie skończonych miar borelowskich na prostej $\mathbb{R}$ – poprzez dystrybuantę F i pseudoodwrotną funkcję kwantylową Q^F – pozwala wyprowadzić ogólne

Stwierdzenie 1.2.14. (i) *Całka Lebesgue'a względem lokalnie skończonej miary borelowskiej μ w przestrzeni $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ wyraża się wzorem*

$$\int f d\mu = \int_{\inf F}^{\sup F} (f \circ Q^F)(u) du,$$

przy czym $f \in \mathcal{I}(\mathbb{R}, \mu) \iff$ funkcja $f \circ Q^F$ jest całkowalna względem miary Lebesgue'a ℓ .

(ii) Całka Lebesgue'a względem miary μ o dystrybuancie F jest jedynym funkcjonałem liniowym $\int dF: \mathcal{I} \rightarrow \mathbb{R}$, nieujemnym i określonym na kracie $\mathcal{I} \subset \mathcal{M}(\mathfrak{B}_{\mathbb{R}})$, spełniającym warunek monotonicznej ciągłości sformułowany w twierdzeniu 1.2.3(ii), oraz warunek:

(i') Przestrzeń liniowa $\mathcal{I}$ zawiera wszystkie funkcje charakterystyczne przedziałów skończonych $(a, b]$, przy czym $\int 1_{(a,b]} dF = F(b) - F(a)$ dla $a < b$, $a, b \in \mathbb{R}$. $\square$

Prosty dowód (por. [2, stw. 3.1.18]) korzysta z twierdzenia Sierpińskiego o monotonicznych klasach zbiorów. Występujący w (ii) funkcjonał $\int dF$ oznacza ogólnie *całkę Lebesgue'a–Stieltjesa* i jest uogólnieniem całki Lebesgue'a – przy słabszych założeniach dotyczących funkcji F .

Przypominamy także (por. [2, tw. 3.1.32]) jawną postać wzoru (1.38) – dla dyfeomorfizmów.

Twierdzenie 1.2.15 (O zamianie zmiennych w $\mathbb{R}^n$). *Dla dowolnego zbioru otwartego $U \subset \mathbb{R}^n$ i C^1 –odwzorowania $F: U \rightarrow \mathbb{R}^n$ założymy, że F jest różnowartościowe poza zbiorem $U_o = \{x; J(x) = 0\}$, gdzie $J = J_F$ jest jacobianem. Wówczas obraz $W = F(U)$ jest zbiorem mierzalnym względem miary Lebesgue'a i zachodzi równość całek*

$$\int_W f(y) dy = \int_U f(F(x)) |J_F(x)| dx \quad \text{dla } f \in \mathcal{I}(W, \ell^n|_W). \quad (1.40)$$

Ponadto, funkcje $f: W \rightarrow \mathbb{R}$ są całkowalne wtedy i tylko wtedy, gdy iloczyn $(f \circ F)|J_F|$ jest całkowalny względem miary Lebesgue'a na zbiorze U . $\square$

Wniosek 1.2.16 (Liniowa zamiana zmiennych w $\mathbb{R}^n$). *Jeśli $A \in \mathbb{R}^{n \times n}$ jest dowolną macierzą odwracalną i $a \in \mathbb{R}^n$, to*

$$\int_{\mathbb{R}^n} f(Ax + a) dx = |\det(A)|^{-1} \int_{\mathbb{R}^n} f(x) dx \quad (1.41)$$

dla dowolnej ℓ^n –całkowalnej funkcji f . $\square$

1.2.4 ...i rozszerzenia

Podstawą rozszerzonej definicji jest

Twierdzenie 1.2.17. *Dla każdej przestrzeni z miarą $(T, \mathfrak{T}, \mu)$ istnieje dokładnie jedna funkcja $\int d\mu: \overline{\mathcal{M}}_+(\mathfrak{T}) \rightarrow \overline{\mathbb{R}}_+$ równa całce $\int d\mu: \mathcal{I}(T, \mu) \rightarrow \mathbb{R}$ na wspólnej części dziedzin, niemalejąca i spełniająca warunki:*

(i) $\int 1_A d\mu = \mu(A)$ dla dowolnego $A \in \mathfrak{T}$.

(ii) Dla każdego niemalejącego ciągu $(f_n)_{n \in \mathbb{N}}$ w $\overline{\mathcal{M}}_+(\mathfrak{T})$ zachodzi równość

$$\int \lim_{n \rightarrow \infty} f_n d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu. \quad (1.42)$$

(iii) Funkcja $\int d\mu$ jest dodatnio jednorodna i addytywna.

(iv) Dla dowolnych $f, g \in \overline{\mathcal{M}}_+(\mathfrak{T})$, jeśli $f \stackrel{\text{pw}}{=} g$, to $\int f d\mu = \int g d\mu$; dla dowolnej funkcji mierzalnej f prawdziwe są równoważności

$$\int |f| d\mu = 0 \iff f \stackrel{\text{pw}}{=} 0 \quad (1.43)$$

oraz

$$\int |f| d\mu < \infty \iff \exists_{g \in \mathcal{I}(T, \mu)} f \stackrel{\text{pw}}{=} g. \quad (1.44)$$

Dowód. Ćwiczenie – por. [2, tw. 3.1.23]. $\square$

Definicja 1.2.18. Funkcję mierzalną $f \in \overline{\mathcal{M}}(\mathfrak{T})$ nazywamy μ -całkowalną, jeśli jest równa prawie wszędzie skończonej funkcji μ -całkowalnej; podzbiór

$$\overline{\mathcal{I}}(T, \mu) = \{f \in \overline{\mathcal{M}}(\mathfrak{T}); \int |f| d\mu < \infty\}$$

będziemy nazywać przestrzenią funkcji μ -całkowalnych. Wzór

$$\int f d\mu := \int f^+ d\mu - \int f^- d\mu \quad (1.45)$$

określa na dziedzinie

$$\mathcal{D}(T, \mu) = \{f \in \overline{\mathcal{M}}(\mathfrak{T}); f^+ \in \overline{\mathcal{I}}(T, \mu) \text{ lub } f^- \in \overline{\mathcal{I}}(T, \mu)\} \quad (1.46)$$

całkę Lebesgue'a względem miary μ w przestrzeni mierzalnej $(T, \mathfrak{T})$.

Wniosek 1.2.19. Całka Lebesgue'a (1.45) istnieje wtedy i tylko wtedy, gdy funkcja f jest ograniczona przez funkcję całkowalną – z dołu lub z góry.

Funkcja mierzalna f spełniająca nierówność $h \leq f \leq g$ (pw), dla pewnych całkowalnych funkcji $h \leq g$, jest również całkowalna. $\square$

Wniosek 1.2.20. (i) Funkcja mierzalna f jest μ -całkowalna wtedy i tylko wtedy, gdy obie całki $\int f^+ d\mu$ i $\int f^- d\mu$ są skończone lub, równoważnie, gdy funkcja $|f|$ jest całkowalna.

(ii) Dla dowolnych funkcji całkowalnych $f, g \in \overline{\mathcal{M}}(\mathfrak{T})$, funkcje $\max(f, g)$ i $\min(f, g)$ oraz określone prawie wszędzie $f + g$ i $f - g$ są całkowalne.

(iii) Na zbiorze funkcji mierzalnych ograniczonych z dołu przez funkcję całkowalną całka Lebesgue'a spełnia warunek

$$\int (af + bg) d\mu = a \int f d\mu + b \int g d\mu \quad \text{dla } a, b \in \mathbb{R}_+,$$

oraz ma własność (1.42). $\square$

Zgodnie z twierdzeniem 1.2.17(iv), konstrukcja i definicja całki Lebesgue'a przenoszą się na klasy funkcji całkowalnych określonych prawie wszędzie, czyli klasy abstrakcji względem relacji $\stackrel{\text{pw}}{=}$.

Wniosek 1.2.21. *Dla dowolnej przestrzeni z miarą $(T, \mathfrak{T}, \mu)$, przestrzeń liniowa ilorazowa $L^1(T, \mu) = \mathcal{I}(T, \mu)/[0]$ jest izomorficzna z przestrzenią klas abstrakcji $\overline{\mathcal{I}}(T, \mu)/\stackrel{\text{pw}}{=}$ funkcji całkowalnych względem relacji równości prawie wszędzie. Całka Lebesgue'a $\int d\mu$ wyznacza funkcjonal liniowy*

$$\int d\mu: L^1(T, \mu) \rightarrow \mathbb{R}, \quad (1.47)$$

ograniczony względem normy $\|f\|_1 = \int |f| d\mu$. $\square$

Dla każdego niemalejącego ciągu $(f_n)_{n \in \mathbb{N}}$ w $L^1(T, \mu)$ zbieżnego do funkcji całkowalnej $f = \sup_n f_n$ własność monotonicznej ciągłości oznacza zbieżność według normy

$$\|f - f_n\|_1 = \int f d\mu - \int f_n d\mu \longrightarrow 0.$$

Ćwiczenie 1.2.2. Korzystając z twierdzenia 1.2.17 sprawdzić, że dla dowolnej nieujemnej funkcji $\mathfrak{T}$ -mierzalnej f , odwzorowanie $\mathfrak{T} \ni A \mapsto \int_A f d\mu \in \overline{\mathbb{R}}_+$ jest miarą – skończoną wtedy i tylko wtedy, gdy f jest funkcją całkowalną.

Proponujemy w tym miejscu uzupełnienie stwierdzenia 1.2.7, które można uznać za specyficzną *przeliczalną całkowalność* każdej(!) skończonej funkcji nieujemnej.

Twierdzenie 1.2.22. *Dla dowolnej przestrzeni $(T, \mathfrak{T}, \mu)$ z miarą σ -skończoną i dowolnej skończonej funkcji mierzalnej $f \geq 0$ istnieje ciąg zbiorów o skończonej mierze $(S_n)_{n \in \mathbb{N}}$ takich, że*

$$T = \bigcup_{n \in \mathbb{N}} S_n \quad \text{ i } \quad \int_{S_n} f d\mu < \infty, \quad \text{ dla } n \in \mathbb{N}.$$

Definicja 1.2.23. *W przestrzeni z miarą $(T, \mathfrak{T}, \mu)$ funkcję $f \in \overline{\mathcal{M}}(\mu)$ nazywać będziemy przeliczalnie μ -całkowalną, jeśli istnieje ciąg zbiorów $S_n \in \mathfrak{T}$ takich, że $\int_{S_n} |f| d\mu < \infty$ dla $n \in \mathbb{N}$, przy czym $f \stackrel{\text{pw}}{=} 1_S f$ dla $S = \bigcup_n S_n$.*

Dowód. Bez istotnego zmniejszenia ogólności możemy ograniczyć dowód do przypadku miary skończonej $\mu(T) < \infty$; ustalmy $f \in \mathcal{M}_+(\mathfrak{T})$. Zgodnie z twierdzeniem 1.2.17 całka $\nu(A) = \int_A f d\mu$ traktowana jako funkcja zbioru $A \in \mathfrak{T}$ jest miarą, a teza twierdzenia oznacza σ -skończoność ν . Wykażemy najpierw półskończoność miary ν . W tym celu w σ -pierścieniu

$$\mathfrak{F} = \{C \in \mathfrak{T}; \forall B \in \mathfrak{T}|_C \nu(B) \in \{0, \infty\}\}$$

zbiorów typu $0/\infty$ lub miary zero względem ν rozważamy skończony kres górny $\alpha = \sup(\mu|_{\mathfrak{F}})$. Niech $F = \bigcup_n F_n \in \mathfrak{F}$ będzie zbiorem, na którym miara μ osiąga maksimum $\mu(F) = \alpha = \lim_n \mu(F_n)$. W analogiczny sposób w σ -pierścieniu

$$\mathfrak{G} = \{C \in \mathfrak{T}|F; \nu(C) = 0\} \subset \mathfrak{F}$$

znajdujemy zbiór G , dla którego $\mu(G) = \gamma = \sup(\mu|_{\mathfrak{G}})$. Z maksymalności miary $\mu(G)$ wynika, że dla każdego zbioru mierzalnego $B \subset F \setminus G$ zachodzą implikacje

$$\nu(B) = 0 \implies \mu(B) = 0, \quad \text{a zatem także} \quad \mu(B) > 0 \implies \nu(B) = \infty.$$

Jeśli zatem $\mu(F \setminus G) > 0$, to podzbiór $W_M = \{x \in F \setminus G; f(x) \leq M\}$ miałby dla pewnego $M > 0$ dodatnią miarę $\mu(W_M)$, a wynikająca stąd równość $\nu(W_M) = \infty$ jest sprzeczna z nierównością

$$\nu(W_M) = \int_{W_M} f d\mu \leq M\mu(W_M).$$

Z równości $\mu(F \setminus G) = 0$ wynika $\nu(F \setminus G) = 0$, a więc $\nu(F) = \int_F f d\mu = 0$ i funkcja $f|_F$ jest prawie wszędzie równa 0.

Dla dowodu σ -addytywności miary ν (na $F^c = T \setminus F$) możemy teraz rozważyć rodzinę

$$\mathfrak{S} = \{C \in \mathfrak{T}|F^c; \text{miara } \nu|_C \text{ jest } \sigma\text{-skończona}\},$$

w ramach której kres górny $\beta = \sup(\mu|_{\mathfrak{S}})$ jest osiągany dla pewnego zbioru $S \in \mathfrak{S}$. Zauważmy, że dla każdego podzbioru $B \subset F^c \setminus S$ o skończonej mierze $\nu(B)$ z ograniczenia $\beta \geq \mu(S \cup B) = \beta + \mu(B)$ wynika równość $\mu(B) = 0$, a więc także $\nu(B) = 0$. Skoro zatem zbiór $F^c \setminus S$ rozłączny z F należy do rodziny $\mathfrak{F}$, z oszacowania

$$\alpha \geq \mu(F \cup (F^c \setminus S)) = \alpha + \mu(F^c \setminus S)$$

wynika równość $\mu(F^c \setminus S) = 0$, a więc także $\nu(F^c \setminus S) = 0$, co kończy dowód σ -skończoności miary ν : $A \mapsto \int_A f d\mu$ na $T = F \cup S \cup (F^c \setminus S)$. $\square$

Analogiczna do przedstawionej w dowodzie konstrukcja stanowi istotny, końcowy fragment dowodu twierdzenia Radona–Nikodyma 1.3.5.

1.2.5 Całka względem produktu miar

Przypominamy podstawowy zespół ogólnych wyników w postaci sformułowanej i udowodnionej np. w [2, par. 3.1.5].

Twierdzenie 1.2.24. Niech $(S, \mathfrak{S}, \mu_S)$ i $(T, \mathfrak{T}, \mu_T)$ będą dowolnymi przestrzeniami z miarą.

(i) Fubini: Dla dowolnej funkcji całkownej $f \in \overline{\mathcal{L}}(S \times T, \mu_S \otimes \mu_T)$ zachodzi równość

$$\int f d(\mu_S \otimes \mu_T) = \begin{cases} \int (s \mapsto \int f(s, \cdot) d\mu_T) d\mu_S \\ \int (t \mapsto \int f(\cdot, t) d\mu_S) d\mu_T \end{cases} \quad (1.48)$$

czyli, z wykorzystaniem notacji (1.39),

$$\int f(s, t) \mu_S(ds) \mu_T(dt) = \begin{cases} \int (\int f(s, t) \mu_T(dt)) \mu_S(ds) \\ \int (\int f(s, t) \mu_S(ds)) \mu_T(dt). \end{cases} \quad (1.49)$$

Funkcje podcałkowe występujące po prawej stronie sformułowanych tu równości są określone i całkowne dla prawie wszystkich wartości parametrów.

(ii) Tonelli: Jeśli f jest funkcją $(\mathfrak{S} \otimes \mathfrak{T})$ -mierzalną, a zbiór $\{f \neq 0\}$ jest σ -skończony i istnieje przynajmniej jedna z całek iterowanych

$$\int \int |f(s, t)| \mu_T(dt) \mu_S(ds) < \infty \quad \text{lub} \quad \int \int |f(s, t)| \mu_S(ds) \mu_T(dt) < \infty,$$

wówczas funkcja f jest całkowna i zachodzą równości (1.48)–(1.49). $\square$

Wniosek 1.2.25. Dla dowolnych przestrzeni z miarą $(S, \mathfrak{S}, \mu_S)$ i $(T, \mathfrak{T}, \mu_T)$ oraz dowolnego zbioru σ -skończonego $C \in (\mathfrak{S} \otimes \mathfrak{T})^\sigma = \mathfrak{S}^\sigma \otimes \mathfrak{T}^\sigma$ zachodzi równość

$$(\mu_S \otimes \mu_T)(C) = \begin{cases} \int \mu_T(C_s) \mu_S(ds) \\ \int \mu_S(C^t) \mu_T(dt), \end{cases} \quad (1.50)$$

gdzie $C_s = \{t \in T; (s, t) \in C\}$ dla $s \in S$ i $C^t = \{s \in S; (s, t) \in C\}$ dla $t \in T$. $\square$

Z twierdzenie Fubini–Tonellego otrzymujemy klasyczną geometryczną interpretację całki – bez założenia σ -skończoności miary.

Stwierdzenie 1.2.26. Dla dowolnej przestrzeni z miarą $(S, \mathfrak{S}, \mu)$ i nieujemnej funkcji $f: S \rightarrow [0, \infty]$ niech $W_f = \{(s, t) \in S \times \mathbb{R}; 0 < t < f(s)\}$ oznacza zbiór pod wykresem.

(i) Prawdziwa jest równoważność $f \in \overline{\mathcal{M}}(\mathfrak{S}) \iff W_f \in \mathfrak{S} \otimes \mathfrak{B}_{\mathbb{R}}$.

(ii) Jeśli funkcja $f \geq 0$ jest mierzalna, zachodzi równość

$$\int f d\mu = (\mu \otimes \ell)(W_f), \quad (1.51)$$

gdzie ℓ jest miarą Lebesgue'a w $\mathbb{R}$.

Funkcja f jest μ -mierzalna wtedy i tylko wtedy, gdy $W_f \in \mathfrak{S}_\mu^\sigma \otimes \mathfrak{B}_{\mathbb{R}}$, a μ -całkowna wtedy i tylko wtedy, gdy miara zbioru W_f jest skończona.

Dowód. (i) Jeśli zbiór pod wykresem W_f jest $\mathfrak{S} \otimes \mathfrak{B}_{\mathbb{R}}$ -mierzalny, to równości

$$W_f \cap (S \times \{c\}) = \{f > c\} \times \{c\} \quad \text{dla } c > 0 \quad (1.52)$$

pociągają za sobą mierzalność zbiorów $\{f > c\}$, a więc także i funkcji f . Dla dowodu implikacji przeciwnej wystarczy zauważyć rozkład

$$W_f = \bigcup_{c>0} \{x; f(x) > c\} \times (0, c), \quad (1.53)$$

gdzie sumę można ograniczyć do indeksów c będących liczbami wymiernymi.

(ii) Z twierdzenia Tonellego wynika równość $\int 1_{W_f} d(\mu \otimes \ell) = \int f d\mu$ – prawdziwa, gdy $W_f \in (\mathfrak{S} \otimes \mathfrak{B}_{\mathbb{R}})_{\mu \otimes \ell}^{\sigma} = \mathfrak{S}_{\mu}^{\sigma} \otimes \mathfrak{B}_{\mathbb{R}}$. Wskazany rozkład zbioru W_f sprawia, że równoważnym warunkiem jest $\mathfrak{S}_{\mu}^{\sigma}$ -mierzalność funkcji f . Skończona wartość obu całek oznacza wówczas całkowalność funkcji f .

Jeśli dla funkcji $\mathfrak{S}$ -mierzalnej $f \geq 0$ zbiór W_f nie jest σ -skończony, również zbiór $\{f > 0\}$ nie jest σ -skończony – i sformułowana w tezie równość sprowadza się do $\infty = \infty$. $\square$

Przykład 1.2.3. Rozważmy dowolne dwie lokalnie skończone miary borelowskie μ, ν na prostej $\mathbb{R}$, o dystrybuantach, odpowiednio, F i G – zdefiniowanych zgodnie z (1.19). Jeśli przynajmniej jedna z dystrybuant (np. G) jest funkcją ciągłą, to dla dowolnych liczb $a < b$ twierdzenie Fubiniego implikuje równości

$$\begin{aligned} (\mu \otimes \nu)\{a < x \leq y \leq b\} &= \int_{(a,b]} \left(\int_{(a,y]} d\mu(x) \right) d\nu(y) = \int_{(a,b]} (F(y) - F(a)) d\nu(y) \\ &= \int_{(a,b]} F d\nu - F(a)(G(b) - G(a)) \end{aligned}$$

oraz

$$\begin{aligned} (\mu \otimes \nu)\{a < x \leq y \leq b\} &= \int_{(a,b]} \left(\int_{[x,b]} d\nu(y) \right) d\mu(x) = \int_{(a,b]} (G(b) - G(x-)) d\mu(x) \\ &= (F(b) - F(a))G(b) - \int_{(a,b]} G d\mu. \end{aligned}$$

Porównując skończone wartości miary produktowej otrzymujemy

$$\int_{(a,b]} F d\nu + \int_{(a,b]} G d\mu = FG|_a^b = F(b)G(b) - F(a)G(a), \quad (1.54)$$

czyli wzór na całkowanie przez części.

1.3 Twierdzenie Lebesgue'a–Radona–Nikodyma

1.3.1 Absolutna – relatywna ciągłość miary

Definicja 1.3.1. W przestrzeni mierzalnej $(T, \mathfrak{T})$, iloczynem dowolnej nieujemnej funkcji mierzalnej $f \in \overline{\mathcal{M}}_+(\mathfrak{T})$ i miary $\mu: \mathfrak{T} \rightarrow [0, \infty]$ nazywamy miarę $f\mu$ taką, że

$$(f\mu)(A) = \int_A f d\mu \quad \text{dla } A \in \mathfrak{T}.$$

Po sprawdzeniu, że istotnie mamy do czynienia z miarą, z twierdzenia 1.2.3 otrzymujemy

Wniosek 1.3.2. Całki Lebesgue'a względem miar μ i $f\mu$ związane są zależnościami

$$\int g d(f\mu) = \int gf d\mu \quad \text{dla } g \in \mathcal{I}(T, f\mu), \quad (1.55)$$

przy czym $\mathcal{I}(T, f\mu) = \{g \in \mathcal{M}(T, f\mu); gf \in \mathcal{I}(T, \mu)\}$.

Dowód. Ćwiczenie. W przypadku, gdy f przybiera wartości nieskończone, przyjmujemy $0 \cdot \infty = 0$. $\square$

Stwierdzenie 1.3.3 (Zależność całki od dziedziny). Dla dowolnej przestrzeni z miarą $(T, \mathfrak{T}, \mu)$ i funkcji całkowalnej $f \in \mathcal{I}(T, \mu)$

- (i) jeśli $\int_A f d\mu \geq 0$ dla wszystkich $A \in \mathfrak{T}$, to $\mu\{f < 0\} = 0$;
- (ii) całka $\int_A f d\mu$ jest jednostajnie ciągłą funkcją dziedziny $A \in \mathfrak{T}$, tzn.

$$\forall \varepsilon > 0 \exists \delta > 0 \mu(A) < \delta \implies \left| \int_A f d\mu \right| < \varepsilon. \quad (1.56)$$

Dowód. (i) Gdyby $\mu\{f < 0\} > 0$, to dla pewnych $n \in \mathbb{N}$ i $\varepsilon > 0$ zachodziłaby nierówność $\mu\{f \leq -\frac{1}{n}\} \geq \varepsilon > 0$. Wówczas

$$\int_{\{f \leq -1/n\}} f d\mu \leq -\frac{\varepsilon}{n} < 0.$$

(ii) Korzystając z rozkładu $f = f^+ - f^-$, możemy założyć, że funkcja f jest nieujemna. Jeśli implikacja (1.56) nie jest prawdziwa, to istnieje liczba $\varepsilon > 0$ i ciąg zbiorów $A_n \in \mathfrak{T}$ o mierze $\mu(A_n) < 2^{-n}$ takich, że $\int_{A_n} f d\mu \geq \varepsilon$, dla $n \in \mathbb{N}$. Wówczas zbiory $B_n = \bigcup_{k \geq n} A_k$ tworzą ciąg zstępujący, przy czym $\mu(B_n) < 2^{1-n}$, dla $n \in \mathbb{N}$, a zatem przekrój $B = \bigcap_n B_n$ ma miarę 0. Z drugiej strony, założona własność zbiorów A_n implikuje nierówności $\int_{B_n} f d\mu \geq \varepsilon, n \in \mathbb{N}$. Przechodząc do granicy (por. ćwiczenie 1.2.2) otrzymujemy sprzeczność $0 = \int_B f d\mu \geq \varepsilon$, co kończy dowód. $\square$

Podstawowym celem niniejszego podrozdziału jest blok twierdzeń dotyczący porównywania dwu miar na tej samej przestrzeni mierzalnej $(T, \mathfrak{T})$, prowadzący m.in. do charakteryzacji miar postaci $\nu = f\mu$. W dowodzie stwierdzenia 1.3.3, gdy ν jest miarą skończoną, implikację (1.56) sprowadziliśmy do szczególnego przypadku postaci $\mu(A) = 0 \implies \nu(A) = 0$, dla $A \in \mathfrak{T}$. Sformułowana niżej klasyczna definicja nie jest w pełni równoważna rozważanej tu zależności miar, jednak o jej powszechnej akceptacji zadecydowała niewątpliwie prostota sformułowania.

Definicja 1.3.4. *Miarę ν w przestrzeni $(T, \mathfrak{T})$ nazywamy absolutnie ciągłą względem miary μ określonej na tej samej przestrzeni, co zapisuje się symbolicznie $\nu \ll \mu$, jeśli zachodzi implikacja*

$$\forall A \in \mathfrak{T} \quad \mu(A) = 0 \implies \nu(A) = 0. \quad (1.57)$$

Funkcję mierzalną $f: T \rightarrow [0, \infty]$ nazywamy gęstością miary ν absolutnie ciągłej (względem μ), jeśli $\nu = f\mu$, czyli

$$\nu(A) = \int_A f \, d\mu \quad \text{dla } A \in \mathfrak{T}. \quad (1.58)$$

Ćwiczenie 1.3.1. Dla miary $\nu = f\mu$ zbiór $Z = \{f > 0\}$ sprowadza implikację (1.57) do równoważności

$$\forall A \in \mathfrak{T} \quad \nu(A) = 0 \iff \mu(A \cap Z) = 0.$$

Uwaga 1.3.2. Gęstość miary ν względem μ bywa oznaczana $\frac{d\nu}{d\mu}$ i nazywana również *pochodną Radona–Nikodyma*. Uzasadnia to m.in. niezbyt formalna równość $d\nu = f \, d\mu$ zawarta w tożsamości (1.55) stanowiącej treść wniosku 1.3.2. Pozostawiamy czytelnikowi wybór interpretacji, czy przytoczone tu stwierdzenie oznacza *łączność mnożenia* $g(f\mu) = (gf)\mu$, czy też jest to *reguła łańcuchowa* $\frac{d\lambda}{d\nu} \frac{d\nu}{d\mu} = \frac{d\lambda}{d\mu}$ dla pochodnych miary $\lambda = (gf)\mu$ względem $\nu = f\mu$ i μ .

Przykład 1.3.3. Absolutna ciągłość nie oznacza istnienia gęstości. W szczególności, każda miara jest absolutnie ciągła względem miary zliczającej ρ . Miary postaci $\mu = f\rho$ na przestrzeni mierzalnej $(T, \mathfrak{T})$ wyróżnia np. własność

$$\mu(A) = \sum_{x \in A} f(x) = \sum_{x \in A} \mu(\{x\}) \quad \text{dla } A \in \mathfrak{T},$$

o ile tylko zbiory jednopunktowe są mierzalne. Miarę μ nazywamy *dyskretną* jeśli jest absolutnie ciągła względem miary zliczającej postaci ρ_S , gdzie S jest przeliczalnym zbiorem mierzalnym.

Przykład 1.3.4. Miara μ przyjmująca skończone wartości $\mu(S_n) > 0, n \in \mathbb{N}$, na zbiorach wypełniających przestrzeń jest absolutnie ciągła względem miary skończonej $\tilde{\mu}(A) = \sum_n \frac{\mu(A \cap S_n)}{2^n \mu(S_n)}$.

Warunkiem wystarczającym dla istnienia gęstości jest naturalne założenie σ –skończoności miary *dominującej*.

Twierdzenie 1.3.5 (Radon–Nikodym). (i) *Dla miary σ –skończonej μ i dowolnej miary ν w przestrzeni mierzalnej $(T, \mathfrak{T})$ następujące dwa warunki są równoważne:*

- *Miara ν jest absolutnie ciągła względem μ , tzn. $\nu \ll \mu$.*
- *Istnieje nieujemna funkcja mierzalna $f \in \overline{\mathcal{M}}_+(\mathfrak{T})$, taka że $\nu = f\mu$.*

(ii) *Miara $\nu \ll \mu$ jest σ –skończona wtedy i tylko wtedy, gdy jest postaci $\nu = f\mu$, gdzie $f \geq 0$ jest dowolną skończoną funkcją mierzalną.*

Wyróżniona w (i)–(ii) funkcja gęstości f wyznaczona jest jednoznacznie z dokładnością do relacji $\stackrel{\text{pw}}{=}$ względem μ .

Dowód twierdzenia Radona–Nikodyma zawiera podrozdział 1.3.3.

Uwaga 1.3.5. Niech μ będzie dowolną miarą w przestrzeni produktowej $(S \times T, \mathfrak{S} \times \mathfrak{T})$. Jeśli indukowana przez rzutowanie miara brzegowa $\mu_T: B \mapsto \mu(S \times B)$, dla $B \in \mathfrak{T}$, jest σ –skończona, to dla każdego zbioru $A \in \mathfrak{S}$ miara $\mu^A: B \mapsto \mu(A \times B)$ jest absolutnie ciągła względem μ_T , a więc istnieje funkcja $f_A \in \mathfrak{M}_+(\mathfrak{T})$ taka, że $\mu^A = f_A \cdot \mu_T$,

$$\mu(A \times B) = \int_B f_A d\mu_T \quad \text{dla } B \in \mathfrak{T}.$$

Lewa strona równości jest – przy ustalonym $B \in \mathfrak{T}$ – miarą na $(S, \mathfrak{S})$. W ogólnym przypadku zadanie, by ustalić *zgodną* rodzinę funkcji $f_A, A \in \mathfrak{S}$ w taki sposób, aby tworzyły rodzinę miar $A \mapsto f_A(t)$, dla $t \in T$, nie ma pozytywnej odpowiedzi. Pozytywny wyjątek stanowią np. miary borelowskie w $\mathbb{R}^{n+m}$. Do tematu wracamy w podrozdziale 1.3.5.

1.3.2 Twierdzenia o rozkładzie

Podana wyżej charakteryzacja miar absolutnie ciągłych pojawiła się w pracy Johana Radona w 1913 r. dla miar borelowskich na $\mathbb{R}^n$, po czym została uogólniona przez Ottona Nikodyma (1930 r.) na dowolne przestrzenie mierzalne. Całość może być uważana za istotne rozszerzenie rozkładu uzyskanego przez Lebesgue’a w 1910 r., wykorzystującego symetryczną relację singularności (wzajemnej osobliwości) miar.

Definicja 1.3.6. *Miary λ i μ w przestrzeni $(T, \mathfrak{T})$ są wzajemnie singularne, jeśli istnieje rozkład przestrzeni $T = L \cup M$ na rozłączne zbiory mierzalne takie, że*

$$\lambda(A \cap M) = \mu(A \cap L) = 0 \quad \text{dla } A \in \mathfrak{T},$$

co zapisuje się jako $\lambda \perp \mu$. Pomimo symetrii relacji używa się także równoważnych określeń: jedna (każda) z miar jest osobliwa względem drugiej.

Twierdzenie 1.3.7 (Lebesgue'a o rozkładzie). *Dla dowolnej miary μ w przestrzeni mierzalnej $(T, \mathfrak{T})$, każda miara σ –skończona λ w $(T, \mathfrak{T})$ posiada jednoznaczny rozkład na sumę miar λ_o, λ_{ac} w $(T, \mathfrak{T})$ takich, że*

- (i) $\lambda_o \perp \mu$ – miara λ_o jest osobliwa względem μ i
- (ii) $\lambda_{ac} \ll \mu$ – miara λ_{ac} jest absolutnie ciągła względem μ .

Dowód. Dla miary skończonej λ przyjmijmy $c = \sup \{\lambda(A); \mu(A) = 0\}$. Ponieważ równość $c = 0$ oznacza $\lambda \ll \mu$, załóżmy, że c jest skończoną liczbą dodatnią i ustalmy ciąg zbiorów $A_n \in \mathfrak{T}$ takich, że $\mu(A_n) = 0$, dla $n \in \mathbb{N}$, oraz $c = \lim_n \lambda(A_n)$. Łatwo sprawdzić, że ciąg wstępujący zbiorów $B_n = A_1 \cup \dots \cup A_n, n \in \mathbb{N}$, ma tę samą własność, a suma $L = \bigcup_n B_n$ ma miary $\mu(L) = 0, \lambda(L) = c$. Ponieważ poza L nie ma zbiorów $A \in \mathfrak{T}$, dla których $\mu(A) = 0$ i $\lambda(A) > 0$, miary

$$\lambda_{ac}(A) := \lambda(A \setminus L), \quad \lambda_s(A) := \lambda(A \cap L), \quad \text{dla } A \in \mathfrak{T}, \quad (1.59)$$

tworzą szukany rozkład.

W przypadku ogólnym, gdy $T = \bigcup_{n \in \mathbb{N}} T_n$ jest sumą wstępującego ciągu zbiorów o mierze skończonej $\lambda(T_n), n \in \mathbb{N}$, można w rozkładzie (1.59) miar $\lambda|_{T_n}$ na każdym ze zbiorów zadbać o to, by stosowne zbiory $L_n \subset T_n$ również tworzyły ciąg wstępujący. Wówczas suma $L = \bigcup_n L_n$ wyznacza stosowny rozkład Lebesgue'a miary λ . Jednoznaczność rozkładu wynika z oczywistego faktu, iż jedyną miarą jednocześnie absolutnie ciągłą i osobliwą jest miara zerowa, $\square$

Wniosek 1.3.8 (Lebesgue–Radon–Nikodym). *Jeśli miary μ i λ w przestrzeni $(T, \mathfrak{T})$ są σ –skończone, to istnieje funkcja $f \in \mathcal{M}_+(\mathfrak{T})$ i zbiór $L \in \mathfrak{T}$ taki, że $\mu(L) = 0$ oraz*

$$\lambda(A) = \lambda(A \cap L) + \int_A f d\mu \quad \text{dla } A \in \mathfrak{T}.$$

Opisany w ten sposób rozkład miary λ jest jednoznaczny. $\square$

Przygotowując się do dowodu twierdzenia Radona–Nikodyma i konstrukcji funkcji f spełniającej (1.58), zauważmy, że każda funkcja prosta postaci $g = c \cdot 1_B \leq f$, dla $c > 0$ i $B \in \mathfrak{T}$, powinna spełniać warunek

$$\nu(A) \geq \int_A c \cdot 1_B d\mu = c \cdot \mu(A) \quad \text{dla } A \in \mathfrak{T} \text{ zawartych w } B. \quad (1.60)$$

Charakteryzację zbiorów, na których pomiędzy rozważanymi miarami zachodzi określona nierówność podaje

Twierdzenie 1.3.9 (Rozkład Hahna). *Dla dowolnych miar μ i ν w przestrzeni mierzalnej $(T, \mathfrak{T})$, z których przynajmniej jedna jest skończona, istnieje rozkład $S = P \cup N$, $P, N \in \mathfrak{T}$, taki, że $P \cap N = \emptyset$ oraz*

$$\begin{aligned} \mu(A) &\geq \nu(A) & \text{gdy } A \subset P \\ \mu(A) &\leq \nu(A) & \text{gdy } A \subset N, \end{aligned} \quad (1.61)$$

dla dowolnego zbioru $A \in \mathfrak{T}$.

Dowód. Niech $\nu(T) < \infty$. Szukając zbioru N , w którym funkcja $\varphi := \nu - \mu$ osiąga skończoną wartość maksymalną $\omega := \sup\{\nu(A) - \mu(A); A \in \mathfrak{T}\} \geq 0$, korzystamy z jawnej konstrukcji podanej w [3] i rozważamy ciąg zbiorów $\{A_m\}_{m \in \mathbb{N}}$ w $\mathfrak{T}$ taki, że $0 \leq \varphi(A_m) \rightarrow \omega$. Dla każdego $m \in \mathbb{N}$, niech $\mathfrak{B}_m$ oznacza rodzinę

$$\mathfrak{B}_m = \left\{ \bigcap_{k \leq m} A'_k; A'_k = A_k \text{ lub } A'_k = T \setminus A_k \text{ dla } k \leq m \right\}$$

zawierającą co najwyżej 2^m zbiorów rozłącznych, przy czym A_m jest sumą podrodziny $\{A \in \mathfrak{B}_m; A \subset A_m\}$. Odrzucając składniki, na których funkcja φ jest ujemna, otrzymujemy nierówności $\varphi(A_m) \leq \varphi(B_m)$, dla

$$B_m = \bigcup \{A \in \mathfrak{B}_m; \nu(A) \geq \mu(A)\}, \quad m \in \mathbb{N}.$$

Skoro dla $n > m$ każdy ze zbiorów rodziny $\mathfrak{B}_n$ zawiera się w jednym ze zbiorów rodziny $\mathfrak{B}_m$, zachodzą nierówności

$$\varphi(B_m) \leq \varphi(B_m \cup B_{m+1}) \leq \dots \leq \varphi(B_m \cup \dots \cup B_n).$$

Z ciągłości obu miar wynikają zatem nierówności $\varphi(B_m) \leq \varphi(C_m) \leq \omega$ dla $C_m = \bigcup_{n \geq m} B_n$, $m \in \mathbb{N}$; stąd równość

$$\varphi\left(\bigcap_m C_m\right) = \lim_{m \rightarrow \infty} \nu(C_m) - \lim_{m \rightarrow \infty} \mu(C_m) = \omega.$$

Zbiór $N = \bigcap_m C_m$, na którym funkcja $\nu - \mu$ osiąga maksimum, ma szukaną własność: nie zawiera żadnego podzbioru A na którym $\mu(A) > \nu(A)$, a dopełnienie $P := T \setminus N$ nie zawiera podzbiorów B na których $\mu(B) < \nu(B)$. $\square$

Ćwiczenie 1.3.6. Sprawdzić, że ciąg $(\varphi(C_m))_{m \in \mathbb{N}}$ jest niemalejący. Wykazać, że dla każdego $n > m$ różnicę

$$\bigcup_{k=m}^n B_k \setminus \bigcup_{k>m}^n B_k$$

można przedstawić jako sumę zbiorów $A \in \mathfrak{B}_n$, dla których $\varphi(A) < 0$.

1.3.3 ...i dowody

Podstawą zastosowania rozkładu Hahna do dowodu twierdzenia Radona–Nikodyma jest

Lemat 1.3.10. *Jeśli ν oraz μ są skończonymi miarami w przestrzeni $(T, \mathfrak{T})$ przy czym $\nu \ll \mu$ i $\nu(T) > 0$, to istnieje liczba $c > 0$ i zbiór $B \in \mathfrak{T}$ o mierze $\mu(B) > 0$, dla których zachodzi nierówność (1.60).*

Dowód. Dla dowolnego $k \in \mathbb{N}$ rozważmy rozkład Hahna $T = P_k \cup N_k$, taki że

$$\frac{\mu(A)}{k} \geq \nu(A) \text{ dla } A \subset P_k \quad \text{ i } \quad \frac{\mu(A)}{k} \leq \nu(A) \text{ dla } A \subset N_k.$$

Zbiór $A = \bigcap_{k \in \mathbb{N}} P_k$ spełnia każdą z nierówności $\nu(A) \leq \mu(A)/k$, $k > 0$, a zatem ma miarę $\nu(A) = 0$. Dla dopełnienia $A^c = T \setminus A = \bigcup_{k \in \mathbb{N}} N_k$ wynika stąd nierówność $\nu(A^c) > 0$, a z absolutnej ciągłości – także $\mu(A^c) > 0$. Dla zakończenia dowodu pozostaje przyjąć $c = \frac{1}{k}$ i $B = N_k$, wybierając z sumy A^c dowolny składnik o dodatniej mierze. $\square$

Dowód twierdzenia Radona–Nikodyma 1.3.5 – cz. I. Dla miar skończonych ν i μ spełniających warunek absolutnej ciągłości $\nu \ll \mu$ rozważmy zbiór funkcji

$$\mathcal{F} = \{g \in \mathcal{M}_+(\mathfrak{T}); \int_A g d\mu \leq \nu(A) \text{ dla } A \in \mathfrak{T}\},$$

zawierający (ewentualnie) szukaną gęstość ν względem μ . Zauważmy, że rodzina $\mathcal{F}$ jest zamknięta ze względu na kres górny skończonych i przeliczalnych ciągów w $\mathcal{F}$. Istotnie, jeśli $g, h \in \mathcal{F}$, to dla dowolnego zbioru mierzalnego $A \in \mathfrak{T}$ mamy

$$\begin{aligned} \int_A \max(g, h) d\mu &= \int_{A \cap \{g \geq h\}} g d\mu + \int_{A \cap \{g < h\}} h d\mu \\ &\leq \nu(A \cap \{g \geq h\}) + \nu(A \cap \{g < h\}) = \nu(A), \end{aligned}$$

a więc $\max(g, h) \in \mathcal{F}$. Analogiczna własność dla ciągów nieskończonych wynika z indukcji i monotonicznej ciągłości całki i miary. Jeśli zatem wybrać ciąg funkcji $f_n \in \mathcal{F}$, $n \in \mathbb{N}$, dla którego całki $\int f_n d\mu$ dążą do wartości

$$\omega := \sup\{\int g d\mu; g \in \mathcal{F}\} \leq \nu(T),$$

to funkcja $f = \sup\{f_n; n \in \mathbb{N}\}$ należy do rozważanej rodziny, przy czym $\int f d\mu = \omega$.

Różnica $\tilde{\nu} = \nu - f\mu$ jest miarą absolutnie ciągłą względem μ . Wykażemy, że jest to miara zerowa, a więc funkcja f jest szukaną gęstością. Istotnie, jeśli $\tilde{\nu}(T) > 0$, to z lematu 1.3.10 wynika istnienie zbioru $B \in \mathfrak{T}$ i liczby $c > 0$ takich, że $\tilde{\nu}(B) > 0$ i $\tilde{\nu}(A) \geq c\mu(A)$ dla $A \subset B$. Stąd

$$\nu(A) - \int_A f d\mu \geq c \int_A 1_B d\mu \quad \text{dla } A \in \mathfrak{T},$$

a więc rodzina $\mathfrak{F}$ zawiera funkcję $\tilde{f} = f + c1_B$ o całce $\int \tilde{f} d\mu > \omega$. Otrzymana sprzeczność potwierdza równość $\nu = f\mu$.

Jeśli funkcja $g \in \mathcal{F}$ także jest gęstością, to, zgodnie z wykazaną wyżej własnością rodziny $\mathcal{F}$, można ograniczyć się do przypadku $f \leq g$. Na mocy (1.43), z równości $\int (g - f)d\mu = 0$ wynika wówczas $f \stackrel{\text{pw}}{=} g$.

W sytuacji, gdy obie miary są σ -skończone, istnieje wstępujący ciąg zbiorów mierzalnych $T_n \nearrow T$, na których miary μ i ν są skończone. Ponieważ $\nu|_{T_n} \ll \mu|_{T_n}$, dla każdego $n \in \mathbb{N}$ istnieje gęstość $f_n \in \mathcal{M}_+(\mathfrak{T}|_{T_n})$ taka, że $\nu|_{T_n} = f_n \cdot \mu|_{T_n}$. Funkcja

$$f(x) = \begin{cases} f_1(x) & \text{dla } x \in T_1 \\ f_n(x) & \text{dla } x \in T_n \setminus T_{n-1}, n > 1, \end{cases}$$

jest szukaną gęstością miary ν względem μ . Teza twierdzenia została tym samym udowodniona dla miar σ -skończonych.

Wykazana wyżej część twierdzenia wystarcza dla wyprowadzenia twierdzenia Lebesgue'a–Radona–Nikodyma – wniosku 1.3.8.

Klasycznym zastosowaniem twierdzenia Radona–Nikodyma (dla miar σ -skończonych) jest kończący [2] dowód twierdzenia Riesza o dualności w przestrzeniach L^p .

Dowód twierdzenia Radona–Nikodyma 1.3.5 – cz. II. W końcowej części dowodu odrzucamy założenie σ -skończoności miary ν , pozostając przy relacji absolutnej ciągłości $\nu \ll \mu$. Zgodnie z założeniem dla miary μ przestrzeń T jest sumą wstępującego ciągu zbiorów $(T_m)_{m \in \mathbb{N}}$ o skończonej mierze $\mu(T_m)$ dla $m \in \mathbb{N}$. Ta część dowodu rozwija koncepcje i konstrukcje przedstawione wcześniej w dowodzie twierdzenia 1.2.22.

Brak dodatkowych założeń dla ν oznacza, że przestrzeń mierzalna $(T, \mathfrak{T})$ może zawierać zbiory typu $0/\infty$ względem tej miary. Dla każdego $m \in \mathbb{N}$ niech $\mathfrak{F}_m \subset \mathfrak{T}$ oznacza rodzinę

$$\mathfrak{F}_m = \{A \in \mathfrak{T}|_{T_m}; \forall B \in \mathfrak{T}|_A \nu(B) \in \{0, \infty\}\}$$

zbiorów mierzalnych $A \subset T_m$, dla których miara $\nu|_A$ przybiera wyłącznie wartości 0 lub ∞ . Skonstruujemy wstępujący ciąg zbiorów $F_m \in \mathfrak{F}_m, m \in \mathbb{N}$, taki że

$$\mu(F_m) = \sup \{ \mu(A); A \in \mathfrak{F}_m \} \quad \text{dla } m \in \mathbb{N},$$

przy czym dopełnienia $F_m^c = T_m \setminus F_m, m \in \mathbb{N}$, tworzą ciąg wstępujący, co jest równoważne (sprawdzić!) własności $F_{m+1} \cap T_m = F_m$ i implikuje tożsamość

$$T \setminus \bigcup_{m \in \mathbb{N}} F_m = \bigcup_{m \in \mathbb{N}} (T_m \setminus F_m). \quad (1.62)$$

Zauważmy, że dla dowolnego $m \in \mathbb{N}$, liczba

$$\alpha_m := \sup \{ \mu(A); A \in \mathfrak{F}_m \} \leq \mu(T_m)$$

pozwala kontrolować rozmiar „trywialnej” części miary ν . Istotnie, wychodząc od dowolnego ciągu zbiorów $(A'_n)_{n \in \mathbb{N}}$ w T_m , dla których $\mu(A'_n) \rightarrow \alpha_m$, stwierdzamy równość $\mu(F'_m) = \alpha_m$ dla sumy $F'_m := \bigcup_{n \in \mathbb{N}} A'_n \in \mathfrak{F}_m$. Aby użyć rozkład (1.62), możemy – dla $m > 1$ – kolejne zbiory $A'_n \in \mathfrak{F}_m$ zastąpić przez $A_n = (A'_n \setminus T_{m-1}) \cup F_{m-1} \in \mathfrak{F}_m$, gdzie

$$\begin{aligned} \mu(A_n) &= \mu(A'_n \setminus T_{m-1}) + \alpha_{m-1} \\ &\geq \mu(A'_n \setminus T_{m-1}) + \mu(A'_n \cap T_{m-1}) = \mu(A'_n), \end{aligned}$$

co zapewnia równość $\mu(F_m) = \alpha_m$ dla $F_m := \bigcup_{n \in \mathbb{N}} A_n$.

Wykażemy, że na każdym z dopełnień $F_m^c = T_m \setminus F_m, m \in \mathbb{N}$, miara ν jest σ –skończona. W tym celu rozważamy σ –pierścień $\mathfrak{S}_m \subset \mathfrak{T}|_{F_m^c}$ utworzony z podzbiorów σ –skończonych względem ν i w standardowy sposób (jako sumę ciągu) znajdujemy zbiór $S_m \in \mathfrak{S}_m$, taki że

$$\mu(S_m) = \sup \{ \mu(A); A \in \mathfrak{S}_m \} \leq \mu(T_m).$$

Zauważmy, że zbiór $F_m^c \setminus S_m$ należy do rodziny $\mathfrak{F}_m$. Istotnie, jeśli $A \subset F_m^c \setminus S_m$ jest dowolnym podzbiorem o mierze skończonej $\nu(A) < \infty$, to dla zbioru $A \cup S_m \in \mathfrak{S}_m$ zachodzi równość $\mu(A \cup S_m) = \mu(S_m)$, co oznacza $\mu(A) = 0$, a więc także $\nu(A) = 0$ – z założenia absolutnej ciągłości miar. Skoro tak, to suma zbiorów rozłącznych $F_m \cup (F_m^c \setminus S_m)$ również należy do $\mathfrak{F}_m$, przy czym z maksymalności liczby $\mu(F_m)$ otrzymujemy $\mu(F_m^c \setminus S_m) = 0$ i $\nu(F_m^c \setminus S_m) = 0$. Wykazaliśmy tym samym, że miara ν na zbiorze $F_m^c = S_m \cup (F_m^c \setminus S_m)$ jest σ –skończona.

Istnieje zatem zbiór $F = \bigcup_{m \in \mathbb{N}} F_m \in \mathfrak{T}$ taki, że miara $\nu|_F$ przybiera wyłącznie wartości 0 lub ∞ , natomiast dopełnienie $F^c = T \setminus F = \bigcup_{m \in \mathbb{N}} F_m^c$ jest σ –skończone. Z udowodnionej wcześniej I części twierdzenia wynika

istnienie skończonej funkcji mierzalnej $f: F^c \rightarrow \mathbb{R}_+$ takiej, że $\nu(A) = \int_A f d\mu$ dla $A \subset F^c$. Dla prawidłowego rozszerzenia funkcji f na T pozostaje jeszcze uzgodnić zachowanie obu miar na zbiorze F . Czytelnik z łatwością sprawdzi, że dla każdego $m \in \mathbb{N}$ kres górny

$$\gamma_m := \sup \{ \mu(A); A \in \mathfrak{T}|F_m \text{ i } \nu(A) = 0 \}$$

jest osiągany dla pewnego zbioru $G_m \subset F_m$ takiego, że $\nu(G_m) = 0$, przy czym $G_m \cap F_{m-1} = G_{m-1}$, jeśli $m > 1$. Z maksymalności miary $\mu(G_m) = \gamma_m$ wynika, że dla każdego zbioru mierzalnego $A \subset F_m \setminus G_m$ zachodzi implikacja $\nu(A) = 0 \Rightarrow \mu(A) = 0$, a więc także $\mu(A) > 0 \Rightarrow \nu(A) = \infty$. Przyjmujemy

$$\tilde{f}(x) = \begin{cases} f(x) & \text{dla } x \in F^c \\ \infty & \text{dla } x \in F \setminus G \\ 0 & \text{dla } x \in G, \end{cases}$$

gdzie korzystamy z własności

$$F \setminus G = \bigcup_{m \in \mathbb{N}} (F_m \setminus G_m) \quad \text{dla } G := \bigcup_{m \in \mathbb{N}} G_m$$

i rozszerzamy f do funkcji $\tilde{f} \in \overline{\mathcal{M}}_+(\mathfrak{T})$, takiej że

$$\int_A \tilde{f} d\mu = \begin{cases} \int_A f d\mu & \text{dla } A \subset F^c \\ \infty & \text{jeśli } A \subset F \setminus G \text{ i } \mu(A) > 0 \\ 0 & \text{jeśli } A \subset G \text{ lub } \mu(A) = 0. \end{cases}$$

Zatem $\int_A \tilde{f} d\mu = \nu(A)$ dla $A \in \mathfrak{T}$.

Jeśli $g: T \rightarrow [0, \infty]$ jest dowolną (inną) funkcją gęstości miary ν względem μ , to na zbiorze $T \setminus (F \setminus G) = F^c \cup G$ obie miary są σ -skończone, a funkcje $\tilde{f}$ i g równe pw. Aby rozszerzyć tę własność na T , rozważmy zbiory $A_n = \{x \in F \setminus G; g(x) \leq n\}$, $n \in \mathbb{N}$. Dla każdego n , z nierówności

$$\nu(A) = \int_A g d\mu \leq n \cdot \mu(A) \quad \text{dla } A \subset A_n$$

wynika, że miara $\nu(A)$ jest skończona i z konieczności równa 0, o ile tylko $\mu(A) < \infty$. W połączeniu z σ -skończonością miary μ oznacza to równość $\nu(A_n) = 0$ dla $n \in \mathbb{N}$, a więc także

$$\nu(\{x \in F \setminus G; g(x) < \infty\}) = 0 = \mu(\{x \in F \setminus G; g(x) < \infty\}).$$

Ostatnia z równości oznacza równość pw. funkcji g i $\tilde{f}$ na zbiorze $F \setminus G$.

Jeśli g jest funkcją skończoną, wynika stąd także równość $\mu(F \setminus G) = 0$, a to daje σ -skończoność miary $\nu = g\mu$. Odwrotnie, jeśli założyć σ -skończoność miary ν , to w przeprowadzonej wyżej konstrukcji można przyjąć $G = F$, co prowadzi do skończonej funkcji gęstości $\tilde{f}$. $\square$

Przypadek, gdy konstruowany w drugiej części dowodu zbiór F ma miarę $\nu(F) = 0$, a więc $G = F$, oznacza

Wniosek 1.3.11. *Jeśli miara μ jest σ –skończona, wówczas każda miara półskończona $\nu \ll \mu$ (def 1.1.22) jest także σ –skończona, przy czym*

$$\nu = f\mu \quad \text{dla funkcji skończonej } f \in \mathcal{M}_+(\mathfrak{T}); \quad \square$$

Przedstawione niżej twierdzenie rozszerzające równoważność sformułowaną w ćwiczeniu 1.3.1 stanowi podstawę dla charakterystyki tzw. statystyk dostatecznych (podrozdział 2.3.5) – w ramach *Probabilistyki*.

Twierdzenie 1.3.12 (O zdominowanej rodzinie miar, Halmos–Savage). *Dla każdej rodziny $\mathcal{P}$ miar σ –skończonych absolutnie ciągłych względem ustalonej σ –skończonej miary μ w przestrzeni $(T, \mathfrak{T})$ istnieje zbiór mierzalny Z taki, że*

$$\left(\forall_{\nu \in \mathcal{P}} \nu(A) = 0 \right) \iff \mu(A \cap Z) = 0, \quad (1.63)$$

dla każdego zbioru $A \in \mathfrak{T}$. Co więcej, istnieje przeliczalna podrodzina $\mathcal{P}' \subset \mathcal{P}$, dla której równoważnym warunkiem jest także $\forall_{\nu \in \mathcal{P}'} \nu(A) = 0$.

Dla przeliczalnej rodziny miar teza twierdzenia jest oczywista.

Dowód. Zgodnie z przykładem 1.3.4 możemy założyć, że miara μ jest skończona. Podążając w tym samym kierunku, rozważmy rodzinę $\tilde{\mathcal{P}}$ wszystkich miar postaci $\frac{1_S}{\nu(S)} \cdot \nu$, dla $\nu \in \mathcal{P}$ i zbiorów $S \in \mathfrak{T}$ o skończonej mierze $\nu(S) > 0$. Niech $\mathcal{Q}$ oznacza zbiór wszystkich miar postaci

$$Q = \sum_{i \in \mathbb{N}} c_i P_i, \quad \text{gdzie } P_i \in \tilde{\mathcal{P}}, \sum_i c_i = 1 \text{ i } c_i \geq 0, \text{ dla } i \in \mathbb{N}.$$

Korzystając z faktu, iż wszystkie miary $Q \in \mathcal{Q}$ są absolutnie ciągłe względem μ , wyróżniamy rodzinę zbiorów

$$\mathfrak{C} = \{C \in \mathfrak{T}; \exists_{Q \in \mathcal{Q}} Q(C) > 0, Q = q \cdot \mu \text{ i } q|_C > 0 \text{ (pw)}\}.$$

Liczba $M = \sup \{\mu(C); C \in \mathfrak{C}\} < \infty$ jest osiągalna jako granica ciągu $(\mu(C_n))_{n \in \mathbb{N}}$, przy czym dla każdego n istnieje miara $Q_n = q_n \cdot \mu \in \mathcal{Q}$ taka, że $q_n|_{C_n} > 0$ i $Q_n(C_n) > 0$. Zauważmy, że zbiór $Z = \bigcup_n C_n$ należy do rodziny $\mathfrak{C}$, a opowiadającą mu miarą jest np. $Q_o = \sum_{n \in \mathbb{N}} 2^{-n} Q_n \in \mathcal{Q}$. Istotnie, gęstością miary Q_o jest funkcja $q_o = \sum_{n \in \mathbb{N}} 2^{-n} q_n$, a jej całkowalność (i zbieżność szeregu) jest konsekwencją ograniczoności

$$\int \sum_{k \leq n} 2^{-k} q_k d\mu = \sum_{k \leq n} 2^{-k} < 1$$

i monotonicznej ciągłości całki $\int d\mu$. Oczywiście $q_o|_Z > 0$ i $\mu(Z) = M$.

Założmy teraz, że $A \in \mathfrak{T}$ jest dowolnym zbiorem mierzalnym takim, że $Q_o(A) = \int_A q_o d\mu = 0$. Wówczas $q_o|_A \stackrel{\text{pw}}{=} 0$, a więc $\mu(A \cap Z) = 0$. Dla dowolnej miary $Q = q \cdot \mu \in \mathcal{Q}$ konsekwencją równości $\mu(A \cap Z) = 0$ jest oszacowanie

$$Q(A) = Q(A \cap Z) + Q(A \setminus Z) = Q(A \setminus Z) \leq Q(\{q > 0\} \setminus Z),$$

a ponieważ zbiór $C = \{q > 0\} \setminus Z$ ma własność $(q + q_o)|_{C \cup Z} > 0$, suma $C \cup Z \in \mathfrak{C}$ nie może mieć miary większej niż $\mu(Z)$. Stąd $\mu(C) = 0$, co kończy dowód implikacji

$$Q_o(A) = 0 \implies \mu(A \cap Z) = 0 \implies \forall Q \in \mathcal{Q} \quad Q(A) = 0 \implies Q_o(A) = 0,$$

a więc wszystkie trzy warunki są równoważne. Pozostaje zauważyć, że dla dowolnej miary $\nu \in \mathcal{P}$ znikanie $\nu(A \cap K) = 0$, dla $\nu(K) < \infty$, oznacza równość $\nu(A) = 0$. $\square$

1.3.4 Całka względem σ –ciała

Udowodnione wcześniej twierdzenie 1.2.22 jest także wnioskiem z twierdzenia Radona–Nikodyma i podpowiada, że w σ –skończonej przestrzeni z miarą każda skończona funkcja mierzalna f jest *przeliczalnie całkowalna* – istnieje ciąg zbiorów mierzalnych $T_n \nearrow T$, dla których wszystkie obcięcia $f|_{T_n}$ są funkcjami całkowalnymi.

Ogólniejsze konsekwencje twierdzenia 1.3.5 zawiera

Lemat 1.3.13. *Niech $(T, \mathfrak{T}, \mu)$ będzie przestrzenią z miarą σ –skończoną, a $f \in \mathcal{M}_+(\mathfrak{T})$ nieujemną funkcją mierzalną. Dla każdego σ –ciała $\mathfrak{S} \subset \mathfrak{T}$, jeśli miara $\mu|_{\mathfrak{S}}$ jest σ –skończona, istnieje funkcja $g \in \mathcal{M}_+(\mathfrak{S})$, taka że*

$$(f \cdot \mu)|_{\mathfrak{S}} = g \cdot \mu|_{\mathfrak{S}}. \quad (1.64)$$

Ponadto, dla dowolnego zbioru $G \in \mathfrak{S}$ o mierze $\mu(G) > 0$

(i) ograniczenie funkcji g do G ma własność

$$\int_A f d\mu = \int_A g d\mu = \int_A g|_G d(\mu|_{\mathfrak{S}}), \quad \text{dla } A \in \mathfrak{S}|_G \quad (1.65)$$

i jest jednoznacznie (z dokładnością do relacji $\stackrel{\text{pw}}{=}$) wyznaczone przez ograniczenie $f|_G$ funkcji f oraz ograniczenie miary μ do podprzestrzeni $G \subset T$;

(ii) $g|_G \stackrel{\text{pw}}{=} f|_G$, jeśli funkcja $f|_G$ jest $\mathfrak{S}$ –mierzalna. $\square$

Jeżeli obcięcie $\mu|_{\mathfrak{S}}$ miary σ –skończonej w przestrzeni $(T, \mathfrak{T})$ do mniejszego σ –ciała $\mathfrak{S} \subset \mathfrak{T}$ zachowuje σ –skończoność, wszystkie funkcje z $\mathcal{M}(\mathfrak{T})$ są – w opisanym poniżej sensie – przeliczalnie μ –całkowalne względem $\mathfrak{S}$. Z drugiej strony, w dowolnej przestrzeni z miarą $(T, \mathfrak{T}, \mu)$, dla każdej funkcji $f \in \mathcal{M}(\mathfrak{T}_\mu^\sigma)$, czyli takiej, że zbiór $S = \{x \in T; f(x) \neq 0\}$ ma miarę σ –skończoną, obcięcie $f|_S \in \mathcal{M}(\mathfrak{T}|_S)$ jest w tym sensie przeliczalnie całkowalne – istnieje pokrycie $S = \bigcup_{n \in \mathbb{N}} S_n$, takie że każda z funkcji $f|_{S_n}$, $n \in \mathbb{N}$, jest całkowalna. Rozumowanie to pozostaje w mocy także w ogólniejszej sytuacji, gdy $\mathfrak{T}$ i $\mathfrak{S}$ są σ –pierścieniami, a zbiór $\{f=0\}$ nie musi być mierzalny.

Definicja 1.3.14. Dla dowolnej przestrzeni z miarą $(T, \mathfrak{T}, \mu)$ i dowolnego σ –pierścienia $\mathfrak{S} \subset \mathfrak{T}$ funkcję μ –mierzalną $f \in \mathcal{M}(T, \mu)$ będziemy nazywać przeliczalnie μ –całkowalną względem $\mathfrak{S}$, jeśli istnieją zbiory $S_n \in \mathfrak{S}_\mu^f$ takie, że $\int_{S_n} |f| d\mu < \infty$, dla $n \in \mathbb{N}$, przy czym $f \stackrel{\text{pw}}{=} 1_S f$, dla $S = \bigcup_n S_n$.

Przyjmujemy oznaczenia

$$\mathcal{I}^\sigma(\mathfrak{S}, \mu) = \{f \in \mathcal{M}(T, \mu); f \text{ jest przeliczalnie } \mu\text{–całkowalna na } \mathfrak{S}\}, \quad (1.66)$$

dla przestrzeni liniowej wszystkich funkcji przeliczalnie μ –całkowalnych względem $\mathfrak{S}$ oraz

$$L^\sigma(\mathfrak{S}, \mu) = \mathcal{I}^\sigma(\mathfrak{S}, \mu) / \stackrel{\text{pw}}{=} \subset M(T, \mu), \quad (1.67)$$

dla odpowiedniej przestrzeni klas równoważności.

Z twierdzenia Radona–Nikodyma 1.3.5(ii) otrzymujemy zatem

Stwierdzenie 1.3.15. Zachodzą inkluzje

$$\begin{aligned} \mathcal{M}(\mathfrak{S}_\mu^\sigma) &\subset \mathcal{I}^\sigma(\mathfrak{S}, \mu) \subset \mathcal{M}(\mathfrak{T}_\mu^\sigma), \quad \text{a więc także} \\ M(T, (\mu|_{\mathfrak{S}})^\sigma) &\subset L^\sigma(\mathfrak{S}, \mu) \subset M(T, \mu^\sigma). \end{aligned} \quad \square$$

Przeliczalna μ –całkowalność (względem $\mathfrak{S}$) funkcji μ –mierzalnej zależy wyłącznie od składowej σ –skończonej μ^σ miary μ i σ –pierścienia $\mathfrak{S}_\mu^\sigma$. Stąd

Twierdzenie 1.3.16 (Całka relatywna). Niech $(T, \mathfrak{T}, \mu)$ będzie dowolną przestrzenią z miarą, a $\mathfrak{S}$ dowolnym σ –pierścieniem w $\mathfrak{T}$.

(i) Istnieje dokładnie jedno odwzorowanie liniowe

$$\Pi_{\mathfrak{S}}^\mathfrak{T} \equiv \Pi_{\mathfrak{S}}: M(T, \mu) \supset L^\sigma(\mathfrak{S}, \mu) \longrightarrow M(T, \mu|_{\mathfrak{S}})$$

takie, że $\Pi_{\mathfrak{S}}[f] \geq 0$ dla $f \geq 0$, a dla klas równoważności funkcji f, g , odpowiednio, $\mathfrak{T}_\mu^\sigma$ – i $\mathfrak{S}_\mu^\sigma$ –mierzalnych, równość $\Pi_{\mathfrak{S}}[f] = [g]$, czyli $\Pi_{\mathfrak{S}}(f) \stackrel{\text{pw}}{=} g$, oznacza własność

$$\int_G f d\mu = \int_G g d(\mu|_{\mathfrak{S}}) \quad \text{dla } G \in \mathfrak{S}, \text{ takich że } \int_G |f| d\mu < \infty. \quad (1.68)$$

- (ii) Ograniczenie $\Pi_{\mathfrak{S}}$ do podprzestrzeni $M(T, (\mu|_{\mathfrak{S}})^{\sigma})$ jest identycznością, a samo odwzorowanie jest rzutowaniem na $\Pi_{\mathfrak{S}}(L^{\sigma}(\mathfrak{S}, \mu)) = M(T, (\mu|_{\mathfrak{S}})^{\sigma})$.
 (iii) Jeśli $\mathfrak{R}$ jest dowolnym σ -pierścieniem zawartym w $\mathfrak{S}$, to $\Pi_{\mathfrak{R}}^{\mathfrak{S}} \circ \Pi_{\mathfrak{S}}^{\mathfrak{R}} = \Pi_{\mathfrak{R}}^{\mathfrak{S}}$.

Z wyróżnionej w punkcie (i) nieujemności operatora $\Pi_{\mathfrak{S}}$ wynika

Wniosek 1.3.17. (i) Dla dowolnych funkcji $f, g \in \mathcal{I}^{\sigma}(\mathfrak{S}, \mu)$ nierówność $f \leq g$ implikuje $\Pi_{\mathfrak{S}}[f] \leq \Pi_{\mathfrak{S}}[g]$, czyli

$$\Pi_{\mathfrak{S}}(f) \leq \Pi_{\mathfrak{S}}(g) \text{ (pw.)}.$$

W szczególności, $|\Pi_{\mathfrak{S}}(f)| \leq \Pi_{\mathfrak{S}}(|f|)$ (pw), dla $f \in \mathcal{I}^{\sigma}(\mathfrak{S}, \mu)$.

(ii) Jeśli $(f_n)_{n \in \mathbb{N}}$ jest dowolnym niemalejącym ciągiem w $\mathcal{I}^{\sigma}(\mathfrak{S}, \mu)$ zbieżnym do funkcji $f \in \mathcal{I}^{\sigma}(\mathfrak{S}, \mu)$, to ciąg całek relatywnych $g_n \stackrel{\text{pw}}{=} \Pi_{\mathfrak{S}}(f_n)$, $n \in \mathbb{N}$, jest monotonicznie zbieżny (pw.) do $\Pi_{\mathfrak{S}}(f)$.

Dowód. Część (i) jest oczywista. Dowód (ii) możemy ograniczyć do ciągu funkcji nieujemnych $0 \leq f_n \nearrow f$. Dla każdego zbioru $G \in \mathfrak{S}$ takiego, że $\int_G f d\mu < \infty$, ciąg równości

$$\int_G g_n d\mu = \int_G f_n d\mu < \infty, \quad \text{dla } n \in \mathbb{N},$$

zapewnia zbieżność $\int_G g_n d\mu \rightarrow \int_G f d\mu$. Z monotonicznej ciągłości całki $\int d(\mu|_G)$ wynika zatem, że funkcja graniczna $g = \sup_n g_n$ jest całkowalna i

$$\int_G g d\mu = \lim_n \int_G g_n d\mu = \int_G f d\mu. \quad \square$$

Ćwiczenie 1.3.7. W przypadku, gdy miara μ na $(T, \mathfrak{T})$ jest skończona, relatywną całką $\Pi_{\{\emptyset, T\}} f$ jest liczba (funkcja stała) równa $\frac{1}{\mu(T)} \int f d\mu$. Wyznaczyć relatywną całkę $\Pi_{\mathfrak{S}}$ względem σ -ciała $\mathfrak{S}$ generowanego przez pojedynczy zbiór $C \in \mathfrak{T}$.

Dowód twierdzenia 1.3.16. (i) Niech $f \in \mathcal{I}^{\sigma}(\mathfrak{S}, \mu)$ będzie dowolną nieujemną funkcją na T przeliczalnie całkowalną względem $\mathfrak{S}$. Przyjmijmy za $S \in \mathfrak{S}_{\mu}^{\sigma}$ zbiór, na którym miara $f\mu$ jest σ -skończona, przy czym $f \stackrel{\text{pw}}{=} f \cdot 1_S$. Z twierdzenia Radona–Nikodyma (lemat 1.3.13) wynika istnienie funkcji $(\mu|_S)$ -mierzalnej g , dla której zachodzi równość (1.65). Rozszerzenie g do funkcji $f_{\mathfrak{S}} \in \mathcal{M}(\mathfrak{S})$ takiej, że $f_{\mathfrak{S}}(x) = 0$ dla $x \notin S$, spełnia warunek

$$(f \cdot \mu)(G) = (f_{\mathfrak{S}} \cdot \mu)(G), \quad \text{dla } G \in \mathfrak{S}$$

i wobec tego zależy od f jednoznacznie – z dokładnością do zbioru miary 0. Dla dowolnej funkcji $f \in \mathcal{I}^{\sigma}(\mathfrak{S}, \mu)$ rozkład $f = f^{+} - f^{-}$ nakazuje przyjąć

$\Pi_{\mathfrak{S}}(f) \stackrel{\text{pw}}{=} g - h$, gdzie $g = (f^+)_{\mathfrak{S}}$ i $h = (f^-)_{\mathfrak{S}}$, a to implikuje własność (1.68). Z liniowości całki wynika poprawność definicji oraz liniowość otrzymanego odwzorowania.

Własność (ii) wynika z faktu, że wskazane w (i) funkcje $\mathfrak{S}$ –mieralne $f_{\mathfrak{S}}$ znikają poza zbiorem σ –skończonym, przy czym, zgodnie z lematem 1.3.13 $f_{\mathfrak{S}} \stackrel{\text{pw}}{=} f$, jeśli f jest także $\mathfrak{S}$ –mieralna.

Tożsamość sformułowana w (iii) wynika z równoważności (1.68). $\square$

Szczególną postać całki relatywnej względem σ –ciała generowanego przez odwzorowanie mieralne opisuje

Stwierdzenie 1.3.18 (Lemat Dooba–Dynkina). *W przestrzeni z miarą $(T, \mathfrak{T}, \mu)$, jeśli $\mathfrak{S}$ jest σ –ciałem $\sigma(Z) \equiv Z^{-1}(\mathfrak{B}_{\mathbb{R}^m})$ generowanym przez funkcję mierzalną $Z = (Z_1, \dots, Z_m): T \rightarrow \mathbb{R}^m$, a miara $\mu_Z \equiv Z_*\mu$ jest σ –skończona, wówczas całka relatywna dowolnej funkcji $h \in \mathcal{I}^{\sigma}(\sigma(Z), \mu)$ jest postaci*

$$\Pi_{\sigma(Z)} h \stackrel{\text{pw}}{=} g(Z_1, \dots, Z_m),$$

gdzie g jest funkcją borelowską na $\mathbb{R}^m$ – określoną jednoznacznie (pw) względem miary μ_Z .

Dowód. Wystarczy ograniczyć się do przypadku $h \geq 0$, gdy miara borelowska $\nu(B) = \int_{Z^{-1}B} h d\mu$, dla $B \in B_{\mathbb{R}^m}$, jest absolutnie ciągła względem miary μ_Z . Z twierdzenia Radona–Nikodyma wynika tożsamość

$$\int_{Z^{-1}B} h d\mu = \int_B g d\mu_Z = \int_{Z^{-1}B} g \circ Z d\mu, \quad B \in B_{\mathbb{R}^m},$$

dla pewnej nieujemnej funkcji borelowskiej g . $\square$

Ćwiczenie 1.3.8. Oryginalny wynik znany jako *Lemat Dooba–Dynkina* nie korzysta z pojęcia miary. Wykazać, że każda $\sigma(Z)/\mathfrak{B}_{\mathbb{R}^p}$ –mierzalna funkcja $h: T \rightarrow \mathbb{R}^p$ jest – dokładnie, czyli wszędzie – postaci $h = g \circ Z$, gdzie g jest funkcją borelowską.

Kolejne twierdzenie włącza do rozważań klasyczną całkowalność – w nieco ogólniejszym wariancie, z p –tą potęgą. Przypominamy klasyczne

Twierdzenie 1.3.19. *Dla każdej liczby rzeczywistej $p > 0$ zbiór (klas) funkcji całkowalnych z p –tą potęgą,*

$$L^p(T, \mu) = \{[f] \in M(T, \mu); |f|^p \in L^1(T, \mu)\} \quad (1.69)$$

jest przestrzenią liniową. Jeśli $p \geq 1$, $L^p(T, \mu)$ wraz z normą

$$\|[f]\|_p \equiv \|f\|_p = \left(\int |f|^p d\mu \right)^{1/p}$$

jest przestrzenią Banacha. Każdy monotoniczny ciąg funkcji $[f_n] \in L^p(T, \mu)$, $n \in \mathbb{N}$, o ograniczonej normie jest zbieżny według normy do $f = \sup_n f_n$. $\square$

Za podstawę przyjmujemy

Lemat 1.3.20. (i) Dowolna funkcja μ -całkowalna $f \in \mathcal{I}(T, \mu)$ jest przeliczalnie całkowalna względem σ -pierścienia $\mathfrak{S}$ wtedy i tylko wtedy, gdy istnieje zbiór $S \in \mathfrak{S}$ zawierający nośnik $\{f \neq 0\}$ (z dokładnością do zbioru miary 0) i σ -skończony względem $\mu|_{\mathfrak{S}}$.

(ii) Dla każdego $p \in [1, \infty)$ przekrój $L^p(T, \mu) \cap L^\sigma(\mathfrak{S}, \mu)$ jest domkniętą podprzestrzenią w przestrzeni L^p .

(iii) $L^p(T, \mu) \subset L^\sigma(\mathfrak{S}, \mu)$, jeśli miara $\mu|_{\mathfrak{S}}$ jest σ -skończona.

Dowód. Ćwiczenie. $\square$

Twierdzenie 1.3.21 (Całka relatywna w L^p). (i) Rzut $f_{\mathfrak{S}} \stackrel{\text{pw}}{=} \Pi_{\mathfrak{S}}(f)$ funkcji całkowalnej $f \in \mathcal{I}(T, \mu) \cap \mathcal{I}^\sigma(\mathfrak{S}, \mu)$ jest całkowalny względem miary $\mu|_{\mathfrak{S}}$ i zachodzi równość

$$\int f_{\mathfrak{S}} d\mu = \int f d\mu. \quad (1.70)$$

(ii) Dla każdego $p \in [1, \infty)$ rzutowanie $\Pi_{\mathfrak{S}}$ zachowuje całkowalność funkcji z p -tą potęgą,

$$\Pi_{\mathfrak{S}}(L^p(T, \mu) \cap L^\sigma(\mathfrak{S}, \mu)) = L^p(T, \mu|_{\mathfrak{S}}) \quad \text{dla } p \geq 1,$$

i wyznacza operator liniowy ciągły o normie 1.

Dowód. (i) Własność (1.68) i przeliczalna całkowalność funkcji $f \in \mathcal{I}(T, \mu)$ implikują skończoność i równość całek $\int_T f^\pm d\mu = \int_T (f^\pm)_{\mathfrak{S}} d\mu$, a stąd wynika całkowalność funkcji $f_{\mathfrak{S}}$ i równość (1.70).

(ii) Dla $p > 1$ korzystamy z monotoniczności operatora $\Pi_{\mathfrak{S}}$ oraz z elementarnej analizy i własności

$$t^p = \sup\{at + b; (a, b) \in \Sigma_p\}, \quad \text{gdzie } \Sigma_p = \{(a, b) \in \mathbb{Q}^2, \forall x \geq 0 \ ax + b \leq x^p\},$$

prawdziwej dla $t \geq 0$ i wynikającej z wypukłości funkcji $x \mapsto x^p$. Jeżeli zatem $f \in \mathcal{I}^\sigma(\mathfrak{S}, \mu)$, to dla każdej pary współczynników $(a, b) \in \Sigma_p$ zachodzi nierówność $a|f| + b \leq |f|^p$. Jeśli dla zbioru $S \in \mathfrak{S}$ zachodzi równość $f \stackrel{\text{pw}}{=} 1_S f$, to z nierówności $a|f| + b1_S \leq |f|^p$ (w przestrzeni $\mathcal{I}^\sigma(\mathfrak{S}, \mu)$) wynika

$$a|f|_{\mathfrak{S}} + b1_S = (a|f| + b1_S)_{\mathfrak{S}} \leq (|f|^p)_{\mathfrak{S}} \stackrel{\text{pw}}{=} |f|^p, \quad (1.71)$$

gdzie indeks $\mathfrak{S}$ wskazuje (dowolne) funkcje reprezentujące klasy – wartości operatora $\Pi_{\mathfrak{S}}$. Nierówność (1.71) jest prawdziwa prawie wszędzie, czyli poza

pewnym zbiorem $Z_{(a,b)}$ miary 0. Przechodząc do kresu górnego otrzymujemy

$$(|f|_{\mathfrak{E}})^p \leqslant (|f|^p)_{\mathfrak{E}} \quad \text{poza zbiorem } Z = \bigcup_{(a,b) \in \Sigma_p} Z_{(a,b)},$$

również o mierze 0. Monotoniczność $\Pi_{\mathfrak{E}}$ (wn. 1.3.17) prowadzi do nierówności

$$|f_{\mathfrak{E}}|^p \leqslant (|f|_{\mathfrak{E}})^p \leqslant (|f|^p)_{\mathfrak{E}} \quad (\text{pw}),$$

z której wynika całkowalność funkcji $|f_{\mathfrak{E}}|^p$ i oszacowania

$$\int |f_{\mathfrak{E}}|^p d\mu \leqslant \int (|f|^p)_{\mathfrak{E}} d\mu = \int |f|^p d\mu \quad \text{ i } \quad \|f_{\mathfrak{E}}\|_p \leqslant \|f\|_p,$$

jeśli tylko funkcja $|f|^p$ jest całkowalna. Dla zakończenia dowodu pozostaje zauważyć, że ograniczenie $\Pi_{\mathfrak{E}}|_{L^p(T, \mu|_{\mathfrak{E}})}$ jest identycznością. $\square$

1.3.5 Warunkowe miary wg. Billingsleya

Zgodnie z zapowiedzią sformułowaną w uwadze 1.3.5 rozważamy ustaloną miarę μ w przestrzeni mierzalnej $(\mathbb{R}^p, \mathfrak{B}_{\mathbb{R}^p}) \times (T, \mathfrak{T})$. Brak symetrii nie przeszkadza, by sformułowane niżej twierdzenie stanowiło „wspólny mianownik” dla grupy twierdzeń wykorzystywanych w kolejnym rozdziale w kontekście probabilistycznym.

Twierdzenie 1.3.22 (O mierze warunkowej). *Niech λ będzie dowolną miarą σ –skończoną w przestrzeni $(\mathbb{R}^p \times T, \mathfrak{B}_{\mathbb{R}^p} \otimes \mathfrak{T})$, $p \in \mathbb{N}$, dla której miara brzegowa*

$$\lambda_T: \mathfrak{T} \ni B \mapsto \lambda(\mathbb{R}^p \times B) \in \overline{\mathbb{R}}_+$$

jest również σ –skończona. Istnieje funkcja

$$\lambda|_T: \mathfrak{B}_{\mathbb{R}^p} \times T \rightarrow [0, 1]$$

taka, że

- (i) *dla $\tau \in T$ funkcja $\lambda|_T(\cdot, \tau)$ jest miarą probabilistyczną na $\mathbb{R}^p$;*
- (ii) *dla każdego zbioru $A \in \mathfrak{B}_{\mathbb{R}^p}$ ewaluacja $\lambda|_T(A, \cdot)$ jest funkcją $\mathfrak{T}$ –mierzalną na T taką, że*

$$\lambda(A \times B) = \int_B \lambda|_T(A, \cdot) d\lambda_T, \quad (1.72)$$

dla dowolnych zbiorów mierzalnych $A \subset \mathbb{R}^p, B \subset T$;

- (iii) *dla każdej całkowalnej funkcji f na $\mathbb{R}^p \times T$*

$$\int f d\lambda = \iint f(x, \tau) \lambda|_T(dx, \tau) \lambda_T(d\tau). \quad (1.73)$$

Powyższa równość zachodzi także dla każdej nieujemnej funkcji mierzalnej f .

- (iv) *Dla każdej funkcji $\rho: \mathfrak{B}_{\mathbb{R}^p} \times T \rightarrow [0, 1]$ o własnościach (i)–(ii) równość miar $\rho(\cdot, \tau) = \lambda|_T(\cdot, \tau)$ zachodzi dla prawie wszystkich $\tau \in T$.*

Definicja 1.3.23. Funkcję $\lambda|_T$ będziemy nazywać warunkowym rozkładem miary λ względem miary brzegowej.

Wniosek 1.3.24. Jeśli miara $\lambda = \varphi \cdot (\nu \otimes \mu)$ jest absolutnie ciągła względem produktu σ –skończonej miary borelowskiej ν na $\mathbb{R}^p$ i σ –skończonej miary μ w przestrzeni $(T, \mathfrak{T})$, przy czym funkcje $\varphi(\cdot, \tau)$ są ν –całkowalne dla prawie wszystkich $\tau \in T$, wówczas

(i) funkcja $g = \int \varphi(x, \cdot) \nu(dx)$ jest gęstością miary miary brzegowej λ_T ; w szczególności $g(\tau) > 0$ dla p.w. $\tau \in T$ – względem miary λ_T ;

(ii) dla p.w. $\tau \in T$ funkcje $g_\tau = \varphi(\cdot, \tau)/g(\tau)$ są gęstościami miar probabilistycznych

$$\lambda|_T(A, \tau) = \int_A g_\tau d\nu, \quad \text{dla } A \in \mathfrak{B}_{\mathbb{R}^p}.$$

Naturalny i dość oczywisty rozkład gęstości $\varphi(x, \tau) = g_\tau(x)g(\tau)$ pozwala traktować równości (1.72)–(1.73) jako uogólnienie twierdzenia Fubiniego.

Dowód. Po wyznaczeniu miary brzegowej w postaci

$$\lambda_T = \left(\int \varphi(x, \cdot) \nu(dx) \right) \cdot \mu$$

potwierdzamy σ –skończoność miary $\lambda_T \equiv g \cdot \mu$ (por. tw. 1.2.22). Równość (1.72) redukuje się do wzoru Fubiniego na całkę iterowaną.

$$\lambda(A \times B) = \int_B \left(\int_A g_\tau(x) \nu(dx) \right) g(\tau) \mu(d\tau) \quad \square$$

W szczególności, jeśli w twierdzeniu 1.3.22 miara λ na produkcie $\mathbb{R}^p \times T$ jest indukowana z ustalonej σ –skończonej przestrzeni z miarą $(T, \mathfrak{T}, \mu)$ przez parę odwzorowań: $\mathfrak{T}$ –mierzalne $X: T \rightarrow \mathbb{R}^p$ oraz $\mathfrak{T}/\mathfrak{S}$ –mierzalną identyczność na T , otrzymujemy

Twierdzenie 1.3.25 (Billingsley, [3]). Załóżmy, że μ jest σ –skończoną miarą w przestrzeni mierzalnej $(T, \mathfrak{T})$, a $X: T \rightarrow \mathbb{R}^p$ dowolną funkcję $\mathfrak{T}$ –mierzalną. Jeśli obcięcie μ do σ –ciała $\mathfrak{S} \subset \mathfrak{T}$ jest miarą σ –skończoną, wówczas istnieje funkcja $\mu_{X|\mathfrak{S}}: \mathfrak{B}_{\mathbb{R}^p} \times T \rightarrow [0, 1]$ spełniająca warunki:

(i) Dla każdego $\tau \in T$ funkcja $\mu_{X|\mathfrak{S}}(\cdot, \tau)$ jest borelowską miarą probabilistyczną na $\mathbb{R}^p$.

(ii) Dla każdego $A \in \mathfrak{B}_{\mathbb{R}^p}$ funkcja $\mu_{X|\mathfrak{S}}(A, \cdot)$ jest $\mathfrak{S}$ –mierzalna, przy czym

$$\mu(X^{-1}A \cap G) = \int_G \mu_{X|\mathfrak{S}}(A, \cdot) d\mu, \quad \text{dla } G \in \mathfrak{S}. \quad (1.74)$$

Jeśli $\varphi: \mathbb{R}^p \rightarrow \mathbb{R}$ jest funkcją borelowską taką, że $\int_G |\varphi(X)| d\mu < \infty$, to także

$$\int_G \varphi(X) d\mu = \int_G \int \varphi(x) \mu_{X|\mathfrak{S}}(dx, \cdot) d\mu, \quad (1.75)$$

gdzie całka wewnętrzna jest skończona prawie wszędzie na T .

Dowód. Ćwiczenie. Dla dowodu (1.75) podstawić $f(x, \tau) = \varphi^\pm(x)1_G(\tau)$ do wzoru (1.73). $\square$

Równość (1.75) oznacza, że funkcja podcałkowa spełnia warunki charakteryzujące całkę relatywną. Stąd

Wniosek 1.3.26. *Dla każdej funkcji borelowskiej $\varphi: \mathbb{R}^p \rightarrow \mathbb{R}$ takiej, że $\varphi(X) \in \mathcal{I}^\sigma(T, \mathfrak{S})$, całka względem rodziny miar warunkowych $\mu_{f|\mathfrak{S}}(dx, \tau)$, $\tau \in T$, wyznacza relatywną całkę*

$$\Pi_{\mathfrak{S}}(\varphi(X)) \stackrel{\text{pw}}{=} \int \varphi(x) \mu_{X|\mathfrak{S}}(dx, \cdot)$$

funkcji $\varphi(X_1, \dots, X_p)$ względem σ -ciała $\mathfrak{S}$. $\square$

Konstrukcja objęta twierdzeniem 1.3.25 ma także ogólniejszy charakter i pozwala na rezygnację z założenia σ -skończoności miary μ .

Wniosek 1.3.27. *Niech $f: T \rightarrow \mathbb{R}$ będzie dowolną funkcją mierzalną względem σ -pierścienia $\mathfrak{T}$. Dla dowolnej miary μ w przestrzeni półmierzalnej $(T, \mathfrak{T})$ i σ -pierścienia $\mathfrak{S} \subset \mathfrak{T}$ takiego, że f jest przeliczalnie całkowalna względem $\mathfrak{S}$, istnieje funkcja $\mu_{f|\mathfrak{S}}: \mathfrak{B}_{\mathbb{R}} \times T \rightarrow \mathbb{R}_+$ spełniająca warunki (i)–(ii) z twierdzenia 1.3.25.*

Dowód. Po ustaleniu funkcji $f \in \mathcal{I}^\sigma(\mathfrak{S}, \mu)$ i zbioru $S \in \mathfrak{S}_\mu^\sigma$ takiego, że $f \stackrel{\text{pw}}{=} 1_S f$, teza wynika z twierdzenia 1.3.25 dla miary $\mu|_S$. $\square$

Ćwiczenie 1.3.9. Dla analogicznego rozszerzenia własności (1.75) w przypadku, gdy $\mathfrak{T}$ (i $\mathfrak{S}$) nie są σ -ciałami, mierzalność złożenia $\varphi(f)$ wymaga, by $\varphi(0) = 0$, co oznacza mierzalność $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ względem σ -pierścienia $\mathfrak{B}_{\mathbb{R} \setminus \{0\}} \subset \mathfrak{B}_{\mathbb{R}}$.

Dowód twierdzenia 1.3.22. Dla dowolnego ciągu liczb wymiernych $r \in \mathbb{Q}^p$ i zbioru $A_r = \prod_{i \leq p} (-\infty, r(i)]$ nierówność $\lambda^r(B) \equiv \lambda(A_r \times B) \leq \lambda_T(B)$, $B \in \mathfrak{T}$, implikuje absolutną ciągłość miary $\lambda^r \ll \lambda_T$. Ustalamy przeliczalną rodzinę funkcji $\mathfrak{T}$ -mierzalnych $F_r: T \rightarrow [0, 1]$ takich, że $\lambda^r = F_r \cdot \lambda_T$, dla $r \in \mathbb{Q}^p$. Dokładniej – dla dowolnego zbioru $B \in \mathfrak{T}$, dla którego $\lambda_T(B) < \infty$, ze stwierdzenia 1.3.3(i) i nierówności

$$\int_G (1 - F_r) d\lambda_T \geq 0, \quad \text{dla każdego } G \in \mathfrak{T}|B,$$

wynika własność $\lambda_T(\{1 - F_r < 0\} \cap G) = 0$. Ewentualna modyfikacja funkcji F_r na zbiorze miary 0 daje dodatkową nierówność $F_r \leq 1$. Sprawdzamy kolejno warunki sformułowane w twierdzeniu 1.1.45.

Dla dowolnych indeksów $r, s \in \mathbb{Q}^p$ takich, że $r \leq s$, czyli $r(i) \leq s(i)$, dla $i \leq p$, z σ -skończoności miary λ_T i nierówności

$$\int_B (F_s - F_r) d\lambda_T = \lambda((A_s \setminus A_r) \times B) \geq 0$$

wynika istnienie zbioru $N_{rs} \in \mathfrak{T}$ o mierze 0 takiego, że $F_r \leq F_s$ poza N_{rs} .

Każdy ciąg w $\mathbb{Q}^p$ rosnący do nieskończoności jest współkońcowy z ciągiem $r_n \equiv (n, \dots, n) \rightarrow \infty$. Przy zachowaniu monotoniczności względem r , dla dowolnego $B \in \mathfrak{T}$ o mierze $\lambda_T(B) < \infty$ zbieżność

$$\int_B F_{r_n} d\lambda_T = \lambda(A_{r_n} \times B) \longrightarrow \lambda_T(B) = \int_B 1 d\lambda_T$$

implikuje zbieżność $\lim_{n \rightarrow \infty} F_{r_n} = 1$, prawie wszędzie, a więc poza zbiorem N_∞ o mierze 0. Analogicznie, dla każdego $r \in \mathbb{Q}^p$ i każdego $j \leq p$ implikacja

$$r_n(i) = \begin{cases} r(i), & \text{dla } i \neq j \\ r(j) - n, & \text{gdy } i = j \end{cases} \implies \int_B F_{r_n} d\lambda_T = \lambda(A_{r_n} \times B) \longrightarrow 0$$

zapewnia zbieżność $\lim_{n \rightarrow \infty} F_{r_n} = 0$ – poza niezależnym od r i j zbiorem N_∞ o mierze 0.

Kolejną własność odpowiadającą półciągłości z góry w dowolnym punkcie $r \in \mathbb{Q}^p$ wykazujemy rozważając ciąg indeksów $r_n = (r(i) + \frac{1}{n})_{i \leq p}$, dla $n \in \mathbb{N}$, współkońcowy z każdym ciągiem $s_n \searrow r$. Dla dowolnego $B \in \mathfrak{T}$ o mierze $\lambda_T(B) < \infty$ z monotonicznej zbieżności

$$\lim_{n \rightarrow \infty} \int_B F_{r_n} d\lambda_T = \lambda(A_r \times B) = \int_B F_r d\lambda_T$$

wynika zbieżność $\lim_{n \rightarrow \infty} F_{r_n} = F_r$ – poza pewnym zbiorem N_r o mierze 0.

Sumując przeliczalną rodzinę zbiorów o mierze 0, znajdujemy zbiór $N \in \mathfrak{T}$ o mierze $\lambda_T(N) = 0$, poza którym spełniona jest przeliczalna rodzina warunków sformułowanych w twierdzeniu 1.1.45. Dokładniej, dla $\tau \notin N$

- (a) $F_r(\tau) \leq F_s(\tau)$, jeśli $r, s \in \mathbb{Q}^p$, $r \leq s$;
- (b) $\forall i \leq p \lim_{r_i \rightarrow -\infty} F_r(\tau) = 0$, $\lim_{r \rightarrow \infty} F_r(\tau) = 1$;
- (c) $\lim_{s \searrow r} F_s(\tau) = F_r(\tau)$;
- (d) dla każdego przedziału postaci $A = (r_1, s_1] \times \dots \times (r_p, s_p]$ o końcach wymiernych zachodzi nierówność

$$\sum_{q \in \{r_1, s_1\} \times \dots \times \{r_p, s_p\}} \text{sign}(q) F_q(\tau) \geq 0.$$

Ostatnia grupa nierówności jest konsekwencją tożsamości

$$\int_B \sum_{q \in \{r_1, s_1\} \times \dots \times \{r_p, s_p\}} \text{sign}(q) F_q d\lambda_T = \lambda(A \times B) \geq 0$$

i σ –skończoności miary λ_T . Przyjmując teraz dla dowolnego $x \in \mathbb{R}^p$

$$\tilde{F}_x(\tau) = \begin{cases} \lim_{r \searrow x} F_r(\tau), & \text{dla } \tau \notin N \\ F^o(x), & \text{dla } \tau \in N, \end{cases}$$

gdzie F^o jest np. dystrybuantą miary Diraca w $0 \in \mathbb{R}^p$, zachowujemy mierzalność względem τ i otrzymujemy dla każdego $\tau \in T$ dystrybuantę jednoznacznie określonej borelowskiej miary probabilistycznej $\lambda_{|T}(\cdot, \tau)$ w $\mathbb{R}^p$.

Dla sprawdzenia mierzalności każdej ewaluacji $\lambda_{|T}(A, \cdot)$ wystarczy zauważyć, że rodzina

$$\mathfrak{A} = \{A \in \mathfrak{B}_{\mathbb{R}^p}; \lambda_{|T}(A, \cdot) \text{ jest funkcją } \mathfrak{T}\text{–mierzalną}\}$$

jest λ –systemem i zawiera π –system przedziałów A_r , $r \in \mathbb{Q}^p$, generujących σ –ciało zbiorów borelowskich $\mathfrak{B}_{\mathbb{R}^p}$. W analogiczny sposób stwierdzamy, że dla każdego zbioru $B \in \mathfrak{T}$ o mierze $\lambda_T(B) < \infty$ rodzina

$$\mathfrak{B} = \{A \in \mathfrak{B}_{\mathbb{R}^p}; \int_B \lambda_{|T}(A, \cdot) = \lambda(A \times B)\}$$

także jest λ –systemem i również zawiera π –system przedziałów A_r , $r \in \mathbb{Q}^p$, a więc – zgodnie z tezą twierdzenia Sierpińskiego 1.1.14(i) – jest równa $\mathfrak{B}_{\mathbb{R}^p}$, a to kończy dowód własności (ii).

Uogólniony wzór Fubiniego (1.73) sprawdzamy w pierwszej kolejności dla funkcji charakterystycznych $f = 1_C$, $C \in \mathfrak{B}_{\mathbb{R}^p} \otimes \mathfrak{T}$, ale zaczynamy od tworzących π –system zbiorów postaci $C = A \times B$, gdzie $\lambda_T(B) < \infty$. Dowód powtarza zawarte w [2, str. 146–147] rozumowanie dot. dowodu twierdzenia Fubiniego. Z twierdzenia 1.1.14(iii) wynika, że rodzina zbiorów

$$\mathfrak{C} = \{C \in (\mathfrak{B}_{\mathbb{R}^p} \otimes \mathfrak{T})_{\lambda}^f; \lambda(C) = \iint 1_C(x, \tau) \lambda_{|T}(dx, \tau) \lambda_T(d\tau)\}$$

zawiera δ –pierścień generowany przez produkty $C = A \times B$, a twierdzenie strukturalne 1.1.13(iii) pozwala rozszerzyć równość (1.73) na funkcje charakterystyczne wszystkich zbiorów $C \in \mathfrak{B}_{\mathbb{R}^p} \otimes \mathfrak{T}$, a więc także na funkcje proste – kombinacje liniowe funkcji charakterystycznych zbiorów o mierze skończonej.

Dla nieujemnej funkcji mierzalnej $f: \mathbb{R}^p \times T \rightarrow \mathbb{R}_+$ równość (1.73) otrzymujemy, korzystając z dowolnej monotonicznej aproksymacji $0 \leq f_n \nearrow f$ funkcjami prostymi. W przypadku dowolnej funkcji całkowalnej f rozkład $f = f^+ - f_-$ i skończona wartość każdej z całek iterowanych oznacza, że funkcje podcałkowe są pw. skończone, a ich różnica jest określona prawie wszędzie i całkowalna.

Dowód jednoznaczności (iv) konstruowanej miary wypada rozpocząć od obserwacji, że równość lokalnie skończonych miar borelowskich w $\mathbb{R}^p$ jest równoważna równości dystrybuant, a ta – wobec półciągłości z góry – da się opisać poprzez równość na przeliczalnym zbiorze gęstym $\mathbb{Q}^p$. Implikuje to mierzalność zbioru

$$\{\tau; \rho(\cdot, \tau) = \lambda|_T(\cdot, \tau)\} = \bigcap_{r \in \mathbb{Q}^p} \{\tau; \Phi_r(\tau) = F_r(\tau)\},$$

gdzie $\Phi_r = \rho(A_r, \cdot)$, dla zdefiniowanych wyżej zbiorów $A_r, r \in \mathbb{Q}^p$. Dla zakończenia dowodu wystarczy zauważyć, że z własności (ii) wynika równość miar $\Phi_r \lambda_T = \lambda^r = F_r \lambda_T$, dla wszystkich multiindeksów r . $\square$

Jeśli w twierdzeniu 1.3.22 przyjąć, że λ jest miarą indukowaną przez parę wektorowych funkcji mierzalnych $(X, Z): T \rightarrow \mathbb{R}^p \times \mathbb{R}^m$, otrzymujemy alternatywną wersję konstrukcji Billingsleya.

Wniosek 1.3.28. *Dla dowolnej σ -skończonej przestrzeni z miarą $(T, \mathfrak{T}, \mu)$ niech $X: T \rightarrow \mathbb{R}^p, Z: T \rightarrow \mathbb{R}^m, p, m \in \mathbb{N}$, będą wektorowymi funkcjami $\mathfrak{T}$ -mierzalnymi. Jeśli miara borelowska $\mu_Z = Z_*\mu$ jest σ -skończona, to istnieje funkcja $\mu_{X|Z}: \mathfrak{B}_{\mathbb{R}^p} \times \mathbb{R}^m \rightarrow [0, 1]$ borelowska ze względu na drugi argument i taka, że*

(i) *dla każdego $z \in \mathbb{R}^m$ funkcja $\mu_{X|Z=z} \equiv \mu_{X|Z}(\cdot, z)$ jest borelowską miarą probabilistyczną, przy czym*

(ii) *dla $A \in \mathfrak{B}_{\mathbb{R}^p}$ i $B \in \mathfrak{B}_{\mathbb{R}^m}$ zachodzą równości*

$$\begin{aligned} \mu(X^{-1}A \cap Z^{-1}B) &= \int_B \mu_{X|Z}(A, \cdot) d\mu_Z \quad \text{oraz} \\ \int_{Z^{-1}B} f(X, Z) d\mu &= \int_B \left(\int f(x, z) \mu_{X|Z=z}(dx) \right) \mu_Z(dz), \end{aligned} \tag{1.76}$$

jeśli dla funkcji borelowskiej $f: \mathbb{R}^{p+m} \rightarrow \mathbb{R}$ przynajmniej jedna ze stron równości jest poprawnie określona.

(iii) *Zachodzi równość $\mu_{X|\sigma(Z)}(A, \tau) = \mu_{X|Z=Z(\tau)}(A)$, dla $A \in \mathfrak{B}_{\mathbb{R}^p}$ i prawie wszystkich $\tau \in T$.* $\square$

Definicja 1.3.29. *Podana w twierdzeniu 1.3.25 rodzina miar $\mu_{X|\mathfrak{S}}(\cdot, \tau)$, $\tau \in T$, jest warunkowym rozkładem funkcji X względem miary μ i σ -ciała $\mathfrak{S}$.*

Funkcja $\mu_{X|Z}$, czyli rodzina miar $(\mu_{X|Z=z})_{z \in \mathbb{R}^m}$ o własnościach (i)–(ii) sformułowanych w tezie wniosku 1.3.28, jest warunkowym rozkładem funkcji X względem miary μ i funkcji Z .

Łącząc wnioski 1.3.24 i 1.3.28 otrzymujemy ważną, klasyczną receptę pozwalającą na efektywne wyznaczanie rozkładów warunkowych.

Twierdzenie 1.3.30. Załóżmy, że $X: T \rightarrow \mathbb{R}^p$, $Z: T \rightarrow \mathbb{R}^m$, $p, m \in \mathbb{N}$, są funkcjami $\mathfrak{T}$ –mierzalnymi na przestrzeni z miarą σ –skończoną $(T, \mathfrak{T}, \lambda)$. Jeśli miara indukowana $\lambda_{(X,Z)} = (X, Z)_* \lambda$ jest absolutnie ciągła względem produktu σ –skończonych miar borelowskich $\nu \otimes \mu$ na $\mathbb{R}^p \times \mathbb{R}^m$, a jej gęstość g spełnia warunek $\int g(x, z) d\nu(x) < \infty$ dla μ –prawie wszystkich $z \in \mathbb{R}^m$, wówczas

- (i) funkcja $g_Z = \int g(x, \cdot) d\nu(x)$ jest gęstością miary $\lambda_Z = Z_* \lambda$ względem μ ;
- (ii) funkcje $g_{X|Z}(\cdot|z) = \frac{g(\cdot, z)}{g_Z(z)}$ określone dla pw. $z \in \mathbb{R}^m$ względem miary λ_Z , są gęstością rozkładu warunkowego $\lambda_{X|Z=z}$ względem miary ν .

Dowód. Ćwiczenie – sprawdzić założenia i uzgodnić oznaczenia. $\square$

Ćwiczenie 1.3.10. Jeśli $\mu\{Z=z\} > 0$, jednoznaczność (pw.) miary warunkowej nie pozostawia swobody i wobec tego

$$\mu_{X|Z=z}(A) = \frac{\mu\{X \in A, Z = z\}}{\mu\{Z = z\}} \quad \text{dla } A \in \mathfrak{B}_{\mathbb{R}^p}.$$

Sprawdzić.

Rozdział 2

To nieprawdopodobne

2.1 Probabilistyczny punkt widzenia

Od czasów Andrieja N. Kołmogorowa i jego aksjomatycznej definicji *przestrzeni probabilistycznej* jako przestrzeni z miarą $(\Omega, \mathfrak{F}, \mathbb{P})$ wiemy, że Ω jest przestrzenią (zbiorem) *zdarzeń elementarnych*, $\mathfrak{F}$ – rodziną *zdarzeń*, a miara $\mathbb{P}$, taka że $\mathbb{P}(\Omega) = 1$, opisuje *prawdopodobieństwo* zdarzeń. W ramach *Statystyki*, abstrakcyjne pojęcie przestrzeni zdarzeń staje się formalnym, matematycznym opisem *populacji*, czyli ogółu obiektów obserwowanych w świecie rzeczywistym – w trakcie badanego cyklu doświadczeń. W tym kontekście prawdopodobieństwo odpowiada doświadczalnie obserwowanej częstości występowania zdarzeń jako grup obiektów wyróżniających się rejestrowanymi w ramach doświadczenia cechami (por. twierdzenie 2.1.1 i wniosek 2.2.26).

2.1.1 Odświeżanie tradycji

Upředzając stosowne definicje, sygnalizujemy twierdzenie, które nadaje precyzyjny sens klasycznej, częstościowej interpretacji *prawdopodobieństwa*.

Twierdzenie 2.1.1 (Prawdopodobieństwo jako częstość). *W każdej przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$ prawdopodobieństwo $\mathbb{P}(A)$ dowolnego zdarzenia $A \in \mathfrak{F}$ wyraża się poprzez granicę ciągu średnich*

$$\mathbb{P}(A) = \lim_{n \rightarrow \infty} \frac{\#\{k \leq n; \omega_k \in A\}}{n}, \quad (2.1)$$

gdzie symbol $\#$ oznacza liczbę elementów wskazanego zbioru, a zbieżność zachodzi dla prawie wszystkich ciągów $\omega = (\omega_k)_{k \in \mathbb{N}} \in \Omega^{\mathbb{N}}$.

Dla dowolnej przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$ istnienie miary produktowej $\mathbb{P}^{\mathbb{N}}$ w przestrzeni produktowej $(\Omega, \mathfrak{F})^{\mathbb{N}}$ pozwala traktować ciągi zdarzeń elementarnych $(\omega_k)_{k \in \mathbb{N}}$ jako *abstrakcyjne próby losowe* i jednocześnie zdarzenia elementarne w $\Omega^{\mathbb{N}}$. Przypominamy zatem:

Twierdzenie 2.1.2 (O nieskończonym produkcie miar [10]). *Dla dowolnego ciągu przestrzeni probabilistycznych $(S_i, \mathfrak{S}_i, \mathbb{P}_i)$, $i \in \mathbb{N}$, w przestrzeni produktowej $(S, \mathfrak{S}) = \prod_{i \in \mathbb{N}} (S_i, \mathfrak{S}_i)$ istnieje dokładnie jedna miara probabilistyczna $\mathbb{P} = \bigotimes_{i \in \mathbb{N}} \mathbb{P}_i$, taka że*

$$(\pi_{1\dots n})_* \mathbb{P} = \mathbb{P}_1 \otimes \dots \otimes \mathbb{P}_n, \quad \text{dla } n \in \mathbb{N}, \quad (2.2)$$

gdzie $\pi_{1\dots n} = (\pi_i)_{i \leq n}$ jest rzutowaniem $S \ni (\omega_i)_{i \in \mathbb{N}} \mapsto (\omega_1, \dots, \omega_n) \in S_1 \times \dots \times S_n$.

Dowód (wg. [17]) podajemy na końcu bieżącego podrozdziału.

Wykorzystany w (2.2) skończony produkt miar jest klasycznym rozwiązaniem pojęcia produktu dla $n = 2$. Przypominamy, że produkt $\mu \otimes \nu$ dowolnych dwu skończonych miar $\mu: \mathfrak{S} \rightarrow \mathbb{R}_+$, $\nu: \mathfrak{T} \rightarrow \mathbb{R}_+$ charakteryzuje warunek

$$(\mu \otimes \nu)(A \times B) = \mu(A)\nu(B), \quad \text{dla } A \in \mathfrak{S}, B \in \mathfrak{T}, \quad (2.3)$$

a dziedziną miary produktowej jest σ -ciało $\mathfrak{S} \otimes \mathfrak{T}$ generowane przez rodzinę produktów $\{A \times B; A \in \mathfrak{S}, B \in \mathfrak{T}\}$.

Dowód twierdzenia 2.1.1. Dla dowolnego ustalonego zdarzenia $A \in \mathfrak{F}$ w przestrzeni Ω niech $Y \equiv 1_A - \mathbb{P}(A): \Omega \rightarrow \mathbb{R}$ oznacza funkcję prostą

$$Y(\omega) \mapsto \begin{cases} 1 - \mathbb{P}(A), & \text{dla } \omega \in A \\ -\mathbb{P}(A), & \text{dla } \omega \in \Omega \setminus A. \end{cases}$$

Przyjmując $Y_k = Y \circ \pi_k: \Omega^{\mathbb{N}} \rightarrow \mathbb{R}$, czyli $Y_k(\omega) = Y(\omega_k)$, dla $\omega = (\omega_i)_{i \in \mathbb{N}}$, $k \in \mathbb{N}$, zauważamy, że sumy

$$S_n \equiv \sum_{k=1}^n Y_k = \sum_{k=1}^n 1_{\pi_k^{-1}A} - n\mathbb{P}(A)$$

spełniają tożsamość $\frac{1}{n}S_n = \#\{k \leq n; \omega_k \in A\}/n - \mathbb{P}(A)$, dla $n \in \mathbb{N}$. Dowodzona teza (2.1) sprowadza się zatem do własności

$$\mathbb{P}^{\mathbb{N}}(Z) = 1, \quad \text{dla } Z = \left\{ \omega \in \Omega^{\mathbb{N}}; \lim_{n \rightarrow \infty} \frac{1}{n}S_n(\omega) = 0 \right\}. \quad (2.4)$$

Równoważność $\frac{1}{n}S_n(\omega) \not\rightarrow 0 \iff \exists_{\varepsilon > 0} \forall_m \exists_{n \geq m} |S_n(\omega)| \geq n\varepsilon$ pozwala zapisać dopełnienie zdarzenia Z w postaci

$$\Omega^{\mathbb{N}} \setminus Z = \bigcup_{\varepsilon > 0} \bigcap_{m} \bigcup_{n \geq m} \{|S_n| \geq n\varepsilon\},$$

a zatem $\mathbb{P}^{\mathbb{N}}(\Omega^{\mathbb{N}} \setminus Z) = 0$ wtedy i tylko wtedy, gdy dla każdego $\varepsilon > 0$ zachodzi zbieżność $\mathbb{P}^{\mathbb{N}}(\bigcup_{n \geq m} \{|S_n| \geq n\varepsilon\}) \rightarrow 0$, dla $m \rightarrow \infty$. Z subaddytywności

$$\mathbb{P}^{\mathbb{N}}\left(\bigcup_{n \geq m} \{|S_n| \geq n\varepsilon\}\right) \leq \sum_{n=m}^{\infty} \mathbb{P}^{\mathbb{N}}\{|S_n| \geq n\varepsilon\}$$

wynika, że wystarczy udowodnić zbieżność szeregu $\sum_{n=1}^{\infty} \mathbb{P}^{\mathbb{N}}\{|S_n| \geq n\varepsilon\}$, w którym każdy ze składników jest prawdopodobieństwem w przestrzeni Ω^n , dla $n \in \mathbb{N}$. W tym celu skorzystamy z oszacowania

$$\int X d\mu \geq \int_{\{X \geq c\}} c d\mu = c\mu\{X \geq c\}$$

prawdziwego dla dowolnej miary μ , dowolnej funkcji μ -całkowalnej $X \geq 0$ i równie dowolnego $c > 0$, zapisanego w postaci $\mu\{X \geq c\} \leq c^{-1} \int X d\mu$ lub ogólniej

$$\mu\{|X| \geq c\} \leq c^{-p} \int |X|^p d\mu, \quad \text{dla } p > 0. \quad (2.5)$$

Jest to tzw. nierówność Czebyszewa–Markowa (por. (2.9)). Dla $p = 2$ całka po prawej stronie nierówności wynosi

$$\begin{aligned} \int |S_n|^2 d\mathbb{P}^{\mathbb{N}} &= \int \left(\sum_{i \leq n} Y_i^2 + \sum_{i \neq j} Y_i Y_j \right) d\mathbb{P}^{\mathbb{N}} \\ &= n \int Y^2 d\mathbb{P} + \sum_{i \neq j} \int Y_i \left(\int Y d\mathbb{P} \right) d\mathbb{P}^{\mathbb{N}} = n \int Y^2 d\mathbb{P}, \end{aligned}$$

gdzie korzystamy także z własności $\int Y d\mathbb{P} = 0$ i twierdzenia Fubinię dla miary $\mathbb{P}^{\mathbb{N}} = \mathbb{P}^{\mathbb{N}} \otimes \mathbb{P}$. Otrzymane oszacowanie $\mathbb{P}^{\mathbb{N}}\{|S_n| \geq n\varepsilon\} \leq \frac{\sigma^2}{n\varepsilon^2}$, gdzie $\sigma^2 = \mathbb{P}(A)(1 - \mathbb{P}(A))$ nie daje wprawdzie szeregu zbieżnego, ale analogiczna tożsamość uzyskana dla $p = 4$

$$\begin{aligned} \int |S_n|^4 d\mathbb{P}^{\mathbb{N}} &= \int \left(\sum_{i \leq n} Y_i^4 + 6 \sum_{i < j \leq n} Y_i^2 Y_j^2 + 4 \sum_{i \leq n} \sum_{j \neq i} Y_i^3 Y_j \right. \\ &\quad \left. + 12 \sum_{j < k \leq n} \sum_{i \neq j, k} Y_i^2 Y_j Y_k + 24 \sum_{i < j < k < l \leq n} Y_i Y_j Y_k Y_l \right) d\mathbb{P}^{\mathbb{N}} \\ &= n \int Y^4 d\mathbb{P} + 3n(n-1) \left(\int Y^2 d\mathbb{P} \right)^2 \leq K n^2, \end{aligned}$$

w której tylko pierwsze dwie grupy składników mają niezerowe całki, prowadzi do oszacowania

$$\mathbb{P}^{\mathbb{N}}\{|S_n| \geq n\varepsilon\} \leq \frac{K\varepsilon^{-4}}{n^2},$$

gdzie K jest stałą niezależną od $n \in \mathbb{N}$. Zbieżność szeregu $\sum \frac{1}{n^2}$ kończy dowód. $\square$

Uwaga 2.1.1. Czytelnik zauważył zapewne, że w dowodzie wykorzystana jest bardzo uproszczona wersja twierdzenia Fubiniego, ograniczona do całki z iloczynu funkcji prostych, zależnych od skończonych, rozłącznych zbiorów współrzędnych. Bardziej zwarty zapis całkowy może być z powodzeniem zastąpiony przez prostą zmianę kolejności sumowania.

Istotną rolę w dowodzie twierdzenia 2.1.1 pełnią własności całkowe funkcji mierzalnych $Y: \Omega \rightarrow \mathbb{R}$ oraz $Y_n, S_n: \Omega^{\mathbb{N}} \rightarrow \mathbb{R}$, dla $n \in \mathbb{N}$. Listę niezbędnych definicji rozpoczyna zatem

Definicja 2.1.3. Zmienną losową w przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$ nazywamy każdą skończoną funkcję mierzalną $X: \Omega \rightarrow \mathbb{R}$, to znaczy taką, że zbiory

$$\{X \in A\} \equiv \{\omega \in \Omega; X(\omega) \in A\}, \quad \text{dla } A \in \mathfrak{B}_{\mathbb{R}}$$

są zdarzeniami, a więc mają określone prawdopodobieństwo $0 \leq \mathbb{P}\{X \in A\} \leq 1$. Miarę borelowską $\mathbb{P}_X = X_*\mathbb{P}$ nazywamy rozkładem prawdopodobieństwa zmiennej X .

Za uogólnioną zmienną losową będziemy uważać każdą prawie wszędzie skończoną funkcję mierzalną $X: \Omega \rightarrow \overline{\mathbb{R}}$.

... a z twierdzenia 2.1.1 wynika

Wniosek 2.1.4. Niech X będzie dowolną zmienną losową w przestrzeni $(\Omega, \mathfrak{F}, \mathbb{P})$. Dla każdego zbioru borelowskiego $B \subset \mathbb{R}$ zbieżność

$$\mathbb{P}_X(B) = \lim_{n \rightarrow \infty} \frac{1}{n} \# \{k \leq n; x_k \in B\} \quad (2.6)$$

zachodzi prawie na pewno – dla prawie wszystkich ciągów liczb rzeczywistych $(x_n)_{n \in \mathbb{N}}$ względem miary produktowej $\mathbb{P}_X^{\mathbb{N}}$ w $\mathbb{R}^{\mathbb{N}}$.

Dowód. Zbiór $C = \{\omega \in \Omega^{\mathbb{N}}; \frac{\#\{k \leq n; X(\omega_k) \in B\}}{n} \rightarrow \mathbb{P}_X(B)\}$ i zbiór borelowski $D = \{x = (x_k) \in \mathbb{R}^{\mathbb{N}}; \frac{\#\{k \leq n; x_k \in B\}}{n} \rightarrow \mathbb{P}_X(B)\}$ łączy zależność $C = \mathbb{X}^{-1}D$, gdzie $\mathbb{X}(\omega) = (X(\omega_k))_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ dla $\omega = (\omega_k) \in \Omega^{\mathbb{N}}$. Stąd

$$1 = \mathbb{P}^{\mathbb{N}}(C) = \mathbb{X}_*(\mathbb{P}^{\mathbb{N}})(D) = \mathbb{P}_X^{\mathbb{N}}(D).$$

Wykorzystana tu równość miar $\mathbb{X}_*(\mathbb{P}^{\mathbb{N}}) = \mathbb{P}_X^{\mathbb{N}}$ wynika z ich zgodności na π -systemie zbiorów borelowskich postaci $A_1 \times \dots \times A_n \times \mathbb{R}^{\mathbb{N}}$. $\square$

Elementarne definicje odwołujące się do pojęcia całki Lebesgue'a obejmują dwie podstawowe charakterystyki liczbowe zmiennych losowych.

Definicja 2.1.5. W odniesieniu do zmiennej losowej $X: \Omega \rightarrow \mathbb{R}$ całka Lebesgue'a

$$\mathcal{E}(X) := \int X d\mathbb{P} \in \overline{\mathbb{R}}$$

stanowi – jeśli istnieje – wartość średnią i nosi nazwę wartości oczekiwanej. W przypadku, gdy $\mathcal{E}(X) \in \mathbb{R}$, wielkość

$$\text{Var}(X) = \mathcal{E}(X - \mathcal{E}(X))^2 \in \overline{\mathbb{R}}_+$$

nazywamy wariancją zmiennej X . Jeśli $\text{Var}(X) < \infty$, liczbę $\sigma_X = \sqrt{\text{Var}(X)}$ nazywamy odchyleniem standardowym zmiennej X .

Ćwiczenie 2.1.2. (a) Sprawdzić oszacowanie $\mathcal{E}|X| \leq (1 + \mathcal{E}(X^2))/2$.

(b) $\text{Var}(X) = 0 \iff X \stackrel{\text{pw}}{=} \mathcal{E}(X)$ jest zmienną losową stałą.

Przykład 2.1.3. Dla dowolnego zdarzenia A w przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$ funkcja $\tau: \Omega^{\mathbb{N}} \rightarrow \mathbb{N} \cup \{\infty\}$ określona wzorem

$$\tau((\omega_i)_{i \in \mathbb{N}}) = \inf_n \{i \in \mathbb{N}; \omega_i \in A\} \quad (2.7)$$

i opisująca czas oczekiwania na sukces, czyli przypadkowe trafienie w ustalony cel, nie jest – formalnie – zmienną losową, gdyż przyjmuje wartość ∞ na produkcie $(\Omega \setminus A)^{\mathbb{N}} \subset \Omega^{\mathbb{N}}$. Funkcja τ jest uogólnioną zmienną losową wtedy i tylko wtedy, gdy $\mathbb{P}(A) > 0$. Średni czas oczekiwania na sukces jest całką równą sumie szeregu

$$\mathcal{E}(\tau) = \sum_{n \geq 1} n\mathbb{P}\{\tau = n\} = \sum_{n=1}^{\infty} n\mathbb{P}(A)(1 - \mathbb{P}(A))^{n-1} = \frac{1}{\mathbb{P}(A)}, \quad (2.8)$$

co daje alternatywną częstościową interpretację prawdopodobieństwa.

Z klasycznego zestawu nierówności dla wartości oczekiwanej wybieramy

Twierdzenie 2.1.6 (Nierówności Czebyszewa). Dla dowolnej całkownej zmiennej losowej X i dowolnego $\varepsilon > 0$,

$$\mathbb{P}\{|X| \geq \varepsilon\} \leq \frac{\mathcal{E}|X|}{\varepsilon} \quad \text{ i } \quad \mathbb{P}\{|X - \mathcal{E}(X)| \geq \varepsilon\} \leq \frac{\text{Var}(X)}{\varepsilon^2}. \quad (2.9)$$

Dowód. Pierwsza nierówność wynika z oszacowania $\int |X| d\mathbb{P} \geq \int_{\{|X| \geq \varepsilon\}} |X| d\mathbb{P}$, druga – z pierwszej, dla zmiennej $(X - \mathcal{E}(X))^2$. $\square$

Twierdzenie o produkcie miar, prawdziwe także dla nieprzeliczalnej rodziny miar probabilistycznych, pojawiło się już po cyklu prac Kołmogorowa i nie wpłynęło na ustalone w tym czasie główne kierunki rozwoju teorii. W pracach Kołmogorowa jest zresztą ogólniejszy wynik (por. tw. 2.6.6) wymagający jednak dodatkowych założeń od rodziny przestrzeni probabilistycznych. Interesujący, relatywnie prosty dowód podajemy za [17]:

Dowód twierdzenia 2.1.2 o produkcie miar probabilistycznych. Przypominamy, że produkt σ -ciała $\mathfrak{S} = \bigotimes_{i \in \mathbb{N}} \mathfrak{S}_i$ jest generowany przez zbiór *cyldrów prostych*

$$\mathfrak{S}^{oo} = \left\{ \prod_{i \in \mathbb{N}} A_i \subset S; A_i \in \mathfrak{S}_i \text{ i } \exists_{n \in \mathbb{N}} (A_i = S_i \text{ dla } i > n) \right\}.$$

Definiując $\mathbb{P}(\prod_{i \in \mathbb{N}} A_i)$ jako iloczyn skończonej liczby miar $\mathbb{P}_i(A_i)$, $i \leq n$, otrzymujemy funkcję $\mathbb{P}: \mathfrak{S}^{oo} \rightarrow [0, 1]$ będącą skończeniem addytywną miarą na półpierścieniu zbiorów. Zgodnie z twierdzeniem Hahna–Kołmogorowa (tw. 1.1.38) dla jednoznacznej rozszerzalności $\mathbb{P}$ na ciało $\mathfrak{S} = \sigma(\mathfrak{S}^{oo})$ wystarczy udowodnić σ -addytywność funkcji $\mathbb{P}$.

Lemat 2.1.7. *Jeśli półpierścień $\mathfrak{I}$ zawiera przestrzeń $T = \bigcup \mathfrak{I}$, wówczas funkcja addytywna $\mu: \mathfrak{I} \rightarrow [0, \infty)$ jest σ -addytywna, gdy dla każdego ciągu $(A_n)_{n \in \mathbb{N}}$ wzajemnie rozłącznych zbiorów w $\mathfrak{I}$*

$$\bigcup_{n \in \mathbb{N}} A_n = T \implies \sum_{n \in \mathbb{N}} \mu(A_n) = \mu(T). \quad (2.10)$$

Dowód lematu. Dla dowolnego ciągu zbiorów rozłącznych $A_n \in \mathfrak{I}$, $n \in \mathbb{N}$, o sumie $B = \bigcup_n A_n \in \mathfrak{I}$ rozważmy skończony rozkład dopełnienia $T \setminus B$ na zbiory $B_i \in \mathfrak{I}$, $i \leq m$. Z rozkładu $T = \bigcup_n A_n \cup \bigcup_{i \leq m} B_i$ wynikają równości

$$\sum_n \mu(A_n) + \sum_{i \leq m} \mu(B_i) = \mu(T) = \mu(B) + \sum_{i \leq m} \mu(B_i). \quad \square$$

Niech zatem $C_n \in \mathfrak{S}^{oo}$, $n \in \mathbb{N}$, będzie dowolnym ciągiem wzajemnie rozłącznych cylindrów, takich że $\bigcup_n C_n = S$; założmy, że $\sum_n \mathbb{P}(C_n) \neq 1$. Poszczególne cylindry możemy zapisać w postaci $C_n = \prod_{i \in \mathbb{N}} A_{ni}$, gdzie $A_{ni} \in \mathfrak{S}_i$, dla wszystkich $i \in \mathbb{N}$ natomiast $A_{ni} = S_i$, dla $i > i_n$, $n \in \mathbb{N}$. Dla $m \in \mathbb{N}$ równość

$$\sum_{n=1}^m \mathbb{P}(C_n) = \sum_{n=1}^m \prod_{i=1}^{i_n} \mathbb{P}_i(A_{ni}) = (\mathbb{P}_1 \otimes \dots \otimes \mathbb{P}_M) \left(\bigcup_{n \leq m} A_{n1} \times \dots \times A_{nM} \right),$$

gdzie $M = \max\{i_n; n \leq m\}$, implikuje oszacowanie $\sum_{n=1}^\infty \mathbb{P}(C_n) \leq 1$. Stąd

$$\sum_{n \in \mathbb{N}} \mathbb{P}(C_n) = \sum_{n \in \mathbb{N}} \prod_{i=1}^{i_n} \mathbb{P}_i(A_{ni}) = \int \sum_{n \in \mathbb{N}} 1_{A_{n1}} \prod_{i=2}^{i_n} \mathbb{P}_i(A_{ni}) d\mathbb{P}_1,$$

co oznacza $\mathbb{P}_1$ -całkowalność funkcji $\psi_1 := \sum_n 1_{A_{n1}} \prod_{i=2}^{i_n} \mathbb{P}_i(A_{ni})$. Jeśli zatem $\int \psi_1 d\mathbb{P}_1 < 1$, to $\psi_1(x_1) < 1$, dla pewnego $x_1 \in S_1$. Z analogicznej tożsamości

$$\psi_1(x_1) = \int \sum_{n \in \mathbb{N}} 1_{A_{n1}}(x_1) 1_{A_{n2}} \prod_{i=3}^{i_n} \mathbb{P}_i(A_{ni}) d\mathbb{P}_2,$$

a więc także $\mathbb{P}_2$ -całkowalności funkcji $\psi_2 := \sum_n 1_{A_{n1}}(x_1) 1_{A_{n2}} \prod_{i=3}^{i_n} \mathbb{P}_i(A_{ni})$ wnioskujemy o istnieniu $x_2 \in S_2$, takiego że $\psi_2(x_2) < 1$. Korzystając z indukcji matematycznej, możemy w ten sposób skonstruować ciąg $x = (x_n)_{n \in \mathbb{N}} \in S$, taki że

$$\sum_{n=1}^{\infty} \left(\prod_{i=1}^k 1_{A_{ni}}(x_i) \right) \prod_{i>k} \mathbb{P}_i(A_{ni}) < 1, \quad (2.11)$$

dla każdego $k \in \mathbb{N}$. Ponieważ $x \in C_m$, dla pewnego $m \in \mathbb{N}$, przyjmując $k = i_m$ otrzymujemy sprzeczność, gdyż m -ty składnik występującej w (2.11) sumy jest wówczas równy 1. $\square$

2.1.2 Zmienne losowe i ich rozkłady

Z twierdzenia 1.2.12 wynika natychmiast

Wniosek 2.1.8. *W przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$, dla każdej zmiennej losowej $X: \Omega \rightarrow \mathbb{R}$ i dowolnej funkcji borelowskiej $g: \mathbb{R} \rightarrow \mathbb{R}$, zmienna $g(X) = g \circ X$ jest całkowalna wtedy i tylko wtedy, gdy $g \in \mathcal{I}(\mathbb{R}, \mathbb{P}_X)$. Równość*

$$\int g(X) d\mu = \int g(x) \mathbb{P}_X(dx) \quad (2.12)$$

opisuje wartość oczekiwaną $\mathcal{E}(g(X))$ jako całkę względem rozkładu prawdopodobieństwa $\mathbb{P}_X = X_\mathbb{P}$ zmiennej losowej X .* $\square$

Przypominamy, że dla dowolnej zmiennej losowej $X: \Omega \rightarrow \mathbb{R}$, indukowaną miarę borelowską $\mathbb{P}_X = X_*\mathbb{P}$, $A \mapsto \mathbb{P}\{X \in A\}$, nazywamy *rozkładem prawdopodobieństwa* zmiennej, a funkcję $F_X: \mathbb{R} \ni x \mapsto \mathbb{P}\{X \leq x\} \in [0, 1]$ – jej *dystrybuantą*. Zgodnie ze stwierdzeniem 1.1.42 (prawie) każda prawostronnie ciągła niemalejąca funkcja $F: \mathbb{R} \rightarrow [0, 1]$ jest *dystrybuantą* jednoznacznie wyznaczonej miary probabilistycznej μ w przestrzeni $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$. Co więcej, każda miara probabilistyczna μ na prostej jest rozkładem prawdopodobieństwa $\mu = Q_*\ell$ szczególnej, lewostronnie ciągłej zmiennej losowej $Q: (0, 1) \rightarrow \mathbb{R}$, związanej z dystrybuantą F równoważnością

$$Q(u) \leq x \iff u \leq F(x) = \mu((-\infty, x]), \quad \text{dla } u \in (0, 1), x \in \mathbb{R}, \quad (2.13)$$

i opisanej wzorem $Q(u) = \inf\{x \in \mathbb{R}; u \leq F(x)\}$, dla $u \in (0, 1)$.

Definicja 2.1.9. *Dla rozkładu prawdopodobieństwa μ o dystrybuantcie F funkcję Q spełniającą (2.13) nazywamy funkcją kwantylową. Liczbę $Q(\frac{1}{2})$ uznajemy za (dolną) medianę rozkładu, a liczby $(Q(\frac{1}{2}) + Q(\frac{1}{2}+))/2$ oraz $Q(\frac{1}{2}+)$ – za mediany: środkową i górną. Wartości funkcji w punktach $\frac{1}{4}$ i $\frac{3}{4}$ wyznaczają, odpowiednio, pierwszy i trzeci kwartył rozkładu.*

Ćwiczenie 2.1.4. Sprawdzić równość (2.12) (wyznaczyć obie całki) w przypadku, gdy zmienna losowa X przybiera przeliczalnie wiele wartości – ma rozkład dyskretny.

Ćwiczenie 2.1.5. Rozkład zmiennej $X: \Omega \rightarrow \mathbb{R}$ jest dyskretny wtedy i tylko wtedy, gdy istnieje zmienna $X' \stackrel{\text{pw}}{=} X$, dla której obraz $X'(\Omega)$ jest zbiorem (co najwyżej) przeliczalnym. Zmienne X i X' są równoważne – wyznaczają ten sam morfizm w zdefiniowanej w [2, Ćw. 2.1.6] kategorii $\mathfrak{Q}_\mu$.

Całkowalność, wartość oczekiwana i wariancja całkowalnej zmiennej losowej zależą jedynie od jej rozkładu. W istocie przedmiotem zainteresowania *Probabilistyki* są raczej rozkłady prawdopodobieństwa, a nie – opisujące je zmienne.

Stwierdzenie 2.1.10. (i) Dla $X: \Omega \rightarrow \mathbb{R}$ zachodzą równości

$$\mathcal{E}(X) = \int t \mathbb{P}_X(dt), \quad \text{Var}(X) = \int (t - \mathcal{E}(X))^2 \mathbb{P}_X(dt). \quad (2.14)$$

(ii) Ponadto, jeśli $X \geq 0$, to dla $a \geq 1$

$$\mathcal{E}(X^a) = \int_0^\infty a t^{a-1} \mathbb{P}\{X > t\} dt. \quad (2.15)$$

Dowód. Własność (i) wynika z (2.12) i jest wnioskiem z twierdzenia o zamianie zmiennych dla zmiennej losowej. Własność (ii) wystarczy sprawdzić dla $a = 1$. Z twierdzenia Fubiniego wynika równość

$$\int_0^\infty \mathbb{P}\{X > t\} dt = \int_0^\infty \int 1_{\{X > t\}} d\mathbb{P} \otimes dt = \int \int_0^\infty 1_{\{X > t\}} dt \otimes d\mathbb{P} = \int X d\mathbb{P}. \quad \square$$

Przykład 2.1.6. Dla dowolnej całkowalnej zmiennej losowej X nierówność

$$\sum_{n \in \mathbb{N}} 1_{\{|X| > n\}} \leq |X|$$

prowadzi do obserwacji, że wskazany tu szereg jest zbieżny i zachodzi nierówność

$$\sum_{n \in \mathbb{N}} \mathbb{P}\{|X| > n\} = \mathcal{E}\left(\sum_{n \in \mathbb{N}} 1_{\{|X| > n\}}\right) \leq \mathcal{E}|X|. \quad (2.16)$$

Dla dowolnego ciągu zmiennych losowych $X_n, n \in \mathbb{N}$, o rozkładzie takim samym jak X , funkcja mierzalna $\tilde{X} = \sum_{n \in \mathbb{N}} 1_{\{|X_n| > n\}}: \Omega \rightarrow \overline{\mathbb{R}}$ nie musi być zmienną losową, gdyż może przyjmować wartość ∞ . Własność

$$\int \tilde{X} d\mathbb{P} = \sum_{n \in \mathbb{N}} \mathbb{P}\{|X_n| > n\} = \sum_{n \in \mathbb{N}} \mathbb{P}\{|X| > n\} < \infty$$

oznacza całkowalność $\tilde{X}$, a zatem $\tilde{X} < \infty$ prawie wszędzie. Stąd np.

Wniosek 2.1.11. Dla każdego ciągu $(X_n)_{n \in \mathbb{N}}$ całkowalnych zmiennych losowych o tym samym rozkładzie zachodzi równość

$$\mathbb{P}\{\omega \in \Omega; \exists_m \forall_{n \geq m} |X_n|(\omega) \leq n\} = 1. \quad \square$$

W oparciu o twierdzenie Radona–Nikodyma (tw. 1.3.5) wyróżniamy w zbiorze rozkładów ważny podzbiór.

Definicja 2.1.12. Rozkład prawdopodobieństwa μ nazywamy ciągłym, jeżeli jest absolutnie ciągły względem miary Lebesgue’a ℓ , a więc taki, że $\mu = f \cdot \ell$, czyli

$$\mu(A) = \int_A f \, d\ell \quad \text{dla } A \in \mathfrak{B}_{\mathbb{R}}, \quad (2.17)$$

dla pewnej nieujemnej funkcji borelowskiej f . Funkcję $f: \mathbb{R} \rightarrow \mathbb{R}_+$ spełniającą warunek (2.17) nazywamy gęstością rozkładu.

Z twierdzenia 1.1.29 o jednoznaczności miar wynika

Wniosek 2.1.13. Funkcja borelowska $f: \mathbb{R} \rightarrow \mathbb{R}_+$ jest gęstością rozkładu μ o dystrybuancie F wtedy i tylko wtedy, gdy

$$F(x) = \int_{-\infty}^x f(t) \, dt \quad \text{dla } x \in \mathbb{R}. \quad \square$$

Zgodnie z definicją iloczynu $f \cdot \ell$ z twierdzenia o zamianie zmiennych i wniosku 1.3.2 wynika ogólny

Wniosek 2.1.14. Jeśli zmienna losowa X ma rozkład ciągły o gęstości f , to

$$\mathcal{E}(g(X)) = \int_{-\infty}^{\infty} g(x)f(x) \, dx,$$

dla funkcji borelowskich g takich, że iloczyn gf jest ℓ –całkowalny. $\square$

W terminologii właściwej dla *Probabilistyki* zamiast równości, nierówności czy zbieżności *prawie wszędzie*, używa się równoważnego określenia *prawie na pewno* o nieco mylącym skrócie *pn.*, którego staramy się w tekście unikać. Przygotowując się do opisu pojęć i konstrukcji związanych z różnego rodzaju przejściami granicznymi, znane w teorii miary pojęcie zbieżności ciągu funkcji prawie wszędzie – czyli *pw.* – osłabiamy dla zmiennych losowych do opisanej niżej postaci.

Definicja 2.1.15. Ciąg $(X_n)_{n \in \mathbb{N}}$ zmiennych losowych w przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$ nazywamy zbieżnym według prawdopodobieństwa do zmiennej $X: \Omega \rightarrow \mathbb{R}$, co zapisujemy $X_n \xrightarrow{\mathbb{P}} X$, jeśli dla każdego $\varepsilon > 0$ zachodzi zbieżność

$$\lim_{n \rightarrow \infty} \mathbb{P}\{\omega \in \Omega; |X_n - X|(\omega) < \varepsilon\} = 1. \quad (2.18)$$

Twierdzenie 2.1.16. Niech X , oraz $X_n, n \in \mathbb{N}$, będą dowolnymi zmiennymi losowymi w przestrzeni $(\Omega, \mathfrak{F}, \mathbb{P})$.

(i) Jeśli $X_n \xrightarrow{\text{pw}} X$, to $X_n \xrightarrow{\text{p}} X$.

(ii) Jeśli $X_n \xrightarrow{\text{p}} X$, to rozkłady zmiennych tworzą ciąg słabo zbieżny

$$\mathbb{P}_{X_n} \xrightarrow{\text{w}} \mathbb{P}_X.$$

(iii) Słaba zbieżność $\mathbb{P}_{X_n} \xrightarrow{\text{w}} \mathbb{P}_X$ ciągu rozkładów w $\mathbb{R}$ jest równoważna zbieżności prawie na pewno ciągu zmiennych kwantylowych $Q_n: (0, 1) \rightarrow \mathbb{R}$,

$$Q_n \xrightarrow{\text{pw}} Q,$$

do funkcji kwantylowej Q rozkładu zmiennej X , przy czym zbieżność zachodzi poza co najwyżej przeliczalnym zbiorem punktów nieciągłości funkcji Q .

Przypominamy, że słabą zbieżność ciągu miar opisuje definicja 1.1.48 i stwierdzenie 1.1.50.

Dowód. (i) Zbieżność $|X_n - X| \xrightarrow{\text{pw}} 0$ oznacza, że przekrój po $\varepsilon > 0$ zdarzeń postaci $\bigcup_m \{|X_n - X| < \varepsilon \text{ dla } n \geq m\}$ ma miarę 1. Tym samym każde z wyróżnionych zdarzeń ma miarę 1, czyli

$$\lim_{m \rightarrow \infty} \mathbb{P}\{\forall n \geq m |X_n - X| < \varepsilon\} = 1$$

dla wszystkich $\varepsilon > 0$. Jest to własność istotnie mocniejsza, niż (2.18).

(ii) Dla dowolnego podzbioru domkniętego $F \subset \mathbb{R}$ i dowolnego $\varepsilon > 0$ implikacja $(X_n(\omega) \in F) \Rightarrow (X(\omega) \in F_\varepsilon \text{ lub } |X_n - X|(\omega) \geq \varepsilon)$ prowadzi do nierówności

$$\mathbb{P}\{X_n \in F\} \leq \mathbb{P}\{X \in F_\varepsilon\} + \mathbb{P}\{|X_n - X| \geq \varepsilon\},$$

czyli $\mathbb{P}_{X_n}(F) \leq \mathbb{P}_X(F_\varepsilon) + \mathbb{P}\{|X_n - X| \geq \varepsilon\}$, gdzie zbiór F_ε oznacza domknięte ε -otoczenie zbioru F . W granicy przy $n \rightarrow \infty$ otrzymujemy

$$\limsup_{n \rightarrow \infty} \mathbb{P}_{X_n}(F) \leq \mathbb{P}_X(F_\varepsilon), \quad \text{dla } \varepsilon > 0,$$

a więc $\limsup_n \mathbb{P}_{X_n}(F) \leq \mathbb{P}_X(F)$, co oznacza słabą zbieżność ciągu miar.

(iii) Własności (i)-(ii) pozwalają ograniczyć dowód do implikacji „ $\Rightarrow$ ”. Niech zatem $F, F_n, n \in \mathbb{N}$, oznaczają dystrybuanty badanych rozkładów. Z definicji słabej zbieżności ciągu rozkładów wynika zbieżność ciągu dystrybuant – poza punktami nieciągłości funkcji F . Ponieważ zbiór punktów ciągłości dystrybuanty F jest gęsty, dla dowolnych $u \in (0, 1)$ i $\varepsilon > 0$ przedział $[Q(u) - \varepsilon, Q(u)]$ zawiera punkt ciągłości x . Z równoważności (2.13)

wynika nierówność $F(x) < u$, czyli także $F_n(x) < u$ – dla pewnego $m \in \mathbb{N}$ i wszystkich $n \geq m$. Otrzymujemy stąd

$$Q(u) - \varepsilon \leq x < Q_n(u), \quad \text{dla } n \geq m,$$

co daje $Q(u) \leq \liminf_{n \rightarrow \infty} Q_n(u)$. Dla wykazania przeciwnej nierówności rozważmy dodatkowo dowolne $v \in (u, 1)$, a w przedziale $[Q(v), Q(v) + \varepsilon)$ ustalmy punkt ciągłości x' . Z nierówności $x' \geq Q(v)$ wynika $F(x') \geq v > u$, a zatem $F_n(x') > u$ dla pewnego $m \in \mathbb{N}$ i $n \geq m$. Stąd

$$Q_n(u) \leq x' < Q(v) + \varepsilon, \quad \text{dla } n \geq m$$

i $\limsup_{n \rightarrow \infty} Q_n(u) \leq Q(v) + \varepsilon$. Wobec dowolności $\varepsilon > 0$ i $v > u$ stwierdzamy

$$Q(u) \leq \liminf_{n \rightarrow \infty} Q_n(u) \leq \limsup_{n \rightarrow \infty} Q_n(u) \leq Q(u+). \quad (2.19)$$

Poza przeliczalnym zbiorem punktów nieciągłości funkcji Q oznacza to zbieżność $\lim_{n \rightarrow \infty} Q_n(u) = Q(u)$. $\square$

Uwaga 2.1.7. Zgodnie z definicją 1.3.29 badanie rozkładu zmiennej losowej X w kontekście drugiej zmiennej Y – na tej samej przestrzeni Ω prowadzi do wyróżnienia odpowiednich warunkowych miar probabilistycznych, którymi zajmujemy się dokładniej w podrozdziale 2.1.4. W najprostszym przypadku rozkład warunkowy $\mathbb{P}_{X|Y}$ jest określony (niemal) jednoznacznie poprzez warunek

$$\mathbb{P}\{X \in A, Y \in B\} = \int_B \mathbb{P}_{X|Y=y}(A) \mathbb{P}_Y(dy),$$

dla dowolnych $A, B \in \mathfrak{B}_{\mathbb{R}}$. Dla danego zbioru borelowskiego A funkcja mierzalna $\mathbb{P}_{X|Y=y}(A)$ nie zależy od zmiennej $y \in \mathbb{R}$, jest pw. stała (i równa $\mathbb{P}_X(A)$) wtedy i tylko wtedy, gdy zachodzi równość

$$\mathbb{P}\{X \in A, Y \in B\} = \mathbb{P}_X(A) \cdot \mathbb{P}_Y(B)$$

dla $B \in \mathfrak{B}_{\mathbb{R}}$. Stąd już krok do rozważanego niżej pojęcia niezależności zmiennych losowych.

2.1.3 Najważniejsza jest niezależność

To, co nie pozwala uważać rachunku prawdopodobieństwa za szczególną dziedzinę teorii miary, to nie specyficzne, powszechnie przyjęte oznaczenia i nazwy, ale pojęcie niezależności: zdarzeń, σ -ciał, zmiennych losowych itp. Jest to precyzyjna, matematyczna postać oczekiwań formułowanych w świecie rzeczywistym – by wyniki kolejnych eksperymentów (w tym rejestrowane błędy) nie wpływały na siebie.

Niech zatem $(\Omega, \mathfrak{F}, \mathbb{P})$ będzie dowolną ustaloną przestrzenią probabilistyczną.

Definicja 2.1.17. Rodzina zdarzeń $\{A_a\}_{a \in \mathcal{A}}$ w $\mathfrak{F}$ jest niezależna, jeśli dla każdego skończonego zbioru różnych indeksów $\{a_1, \dots, a_n\} \subset \mathcal{A}$, $n \in \mathbb{N}$, zachodzi równość

$$\mathbb{P}(A_{a_1} \cap \dots \cap A_{a_n}) = \mathbb{P}(A_{a_1}) \cdot \dots \cdot \mathbb{P}(A_{a_n}). \quad (2.20)$$

Rodzinę podzbiorów $\mathfrak{C}_a \subset \mathfrak{F}$, $a \in \mathcal{A}$, nazywamy niezależną, jeśli każdy układ zdarzeń $A_a \in \mathfrak{C}_a$, $a \in \mathcal{A}$, jest niezależny.

Upraszczając terminologię, w opisanej wyżej sytuacji przyjmuje się, że zbiory zdarzeń $\mathfrak{C}_a$, $a \in \mathcal{A}$, i same zdarzenia A_a , $a \in \mathcal{A}$, są niezależne.

Ćwiczenie 2.1.8. Przekrój dowolnej przeliczalnej rodziny zbiorów o miarze 1 również ma miarę 1 – zdarzenia takie są zatem niezależne. Ogólnie, zdarzenie $A \in \mathfrak{F}$ jest niezależne od siebie (tzn. ciąg zdarzeń (A, A) jest niezależny) wtedy i tylko wtedy, gdy $\mathbb{P}(A) = 1$ lub $\mathbb{P}(A) = 0$. Proponujemy czytelnikowi sprawdzenie, że każde zdarzenie jest niezależne od dowolnej rodziny zdarzeń o prawdopodobieństwie równym 0 lub 1.

W uzupełnieniu proponujemy także – jako ćwiczenie

Stwierdzenie 2.1.18. (i) Dowolna rodzina zdarzeń o prawdopodobieństwie 0 lub 1 generuje σ -ciało zbiorów o tej samej własności.

(ii) Funkcja (zmienna losowa) $X: \Omega \rightarrow \mathbb{R}$ mierzalna względem σ -ciała zdarzeń o prawdopodobieństwie 0 lub 1 jest prawie na pewno stała. $\square$

Porządkując znane definicje, przypominamy

Stwierdzenie 2.1.19. (i) Niezależność zdarzeń $A, B \in \mathfrak{F}$ oznacza niezależność σ -ciał $\mathcal{A} = \{\emptyset, A, \Omega \setminus A, \Omega\}$ i $\mathcal{B} = \{\emptyset, B, \Omega \setminus B, \Omega\}$.

(ii) Dla $n \in \mathbb{N}$ zbiory zdarzeń $\mathfrak{C}_i \subset \mathfrak{F}$, $i \leq n$, tworzą ciąg niezależny wtedy i tylko wtedy, gdy po ewentualnym rozszerzeniu do $\mathfrak{C}'_i = \mathfrak{C}_i \cup \{\Omega\}$ spełniają warunek

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1) \cdot \dots \cdot \mathbb{P}(A_n) \quad (2.21)$$

dla dowolnych zdarzeń $A_i \in \mathfrak{C}'_i$, $i \leq n$. $\square$

Pojęcie niezależności rodzin zbiorów dotyczy w istocie σ -ciał. Niezależność dowolnej rodziny σ -ciał zdarzeń jest przy tym własnością skończonego typu, tzn. zachodzi wtedy i tylko wtedy, gdy ma ją każdy skończony podciąg (wyznaczony przez skończony podzbiór zbioru indeksów).

Twierdzenie 2.1.20 (O niezależności σ -ciał). *Niech $\mathfrak{C}_1, \dots, \mathfrak{C}_n$ ($n > 1$) będzie skończonym ciągiem niezależnych π -systemów w przestrzeni $(\Omega, \mathfrak{F}, \mathbb{P})$. Wówczas σ -ciała $\sigma_\Omega(\mathfrak{C}_i)$, $i \leq n$, są również niezależne.*

Dowód. Zgodnie z wcześniejszą obserwacją możemy założyć, że każdy z rozważanych π -systemów zawiera $\{\Omega\}$. Z założenia niezależności wynika, że dla dowolnych zbiorów $A_2 \in \mathfrak{C}_2, \dots, A_n \in \mathfrak{C}_n$ rodzina

$$\mathfrak{B}_1 = \left\{ B \in \mathfrak{F}; \mathbb{P}(B \cap A_2 \cap \dots \cap A_n) = \mathbb{P}(B)\mathbb{P}(A_2) \cdots \mathbb{P}(A_n) \right\}$$

zawiera $\mathfrak{C}_1$ i jest λ -systemem (por. Def. 1.1.9). Z twierdzenia 1.1.14 wynika zatem inkluzja $\sigma(\mathfrak{C}_1) \subset \mathfrak{B}_1$, przy czym założenie niezależności przenosi się na ciąg π -systemów $\sigma(\mathfrak{C}_1), \mathfrak{C}_2, \dots, \mathfrak{C}_n$. Ogólnie, założmy, że dla pewnego k takiego, że $2 \leq k \leq n$, π -systemy

$$\sigma(\mathfrak{C}_1), \dots, \sigma(\mathfrak{C}_{k-1}), \mathfrak{C}_k, \dots, \mathfrak{C}_n$$

są niezależne. Wówczas dla dowolnych zbiorów $B_i \in \sigma(\mathfrak{C}_i)$, $i < k$, oraz $A_j \in \mathfrak{C}_j$, $k < j \leq n$, rodzina

$$\mathfrak{B}_k = \left\{ B \in \mathfrak{F}; \mathbb{P}\left(\bigcap_{i < k} B_i \cap B \cap \bigcap_{j > k} A_j\right) = \prod_{i < k} \mathbb{P}(B_i) \cdot \mathbb{P}(B) \cdot \prod_{j > k} \mathbb{P}(A_j) \right\}$$

zawiera $\mathfrak{C}_k$ i jest λ -systemem, a zatem własność niezależności rozciąga się na π -systemy $\sigma(\mathfrak{C}_1), \dots, \sigma(\mathfrak{C}_k), \mathfrak{C}_{k+1}, \dots, \mathfrak{C}_n$. Indukcja kończy dowód. $\square$

Wniosek 2.1.21. (i) *Jeśli $\mathfrak{C}_1, \dots, \mathfrak{C}_n$ ($n > 1$) jest skończonym ciągiem niezależnych π -systemów, to dla dowolnego $k \leq n$ σ -ciała $\sigma_\Omega(\mathfrak{C}_i)$, $i < k$, oraz $\sigma_\Omega(\mathfrak{C}_k \cup \dots \cup \mathfrak{C}_n)$ są niezależne.*

(ii) *Jeśli $(\mathfrak{C}_n)_{n \in \mathbb{N}}$ jest nieskończonym ciągiem niezależnych π -systemów, to dla dowolnego $k > 1$ σ -ciała $\sigma_\Omega(\mathfrak{C}_i)$, $i < k$, oraz $\sigma_\Omega(\bigcup_{n \geq k} \mathfrak{C}_n)$ są niezależne.*

Dowód. Bez zmniejszenia ogólności zakładamy, że każdy z π -systemów zawiera zbiór $\{\Omega\}$. Własność (i) wynika z twierdzenia 2.1.20 dla dodatkowego π -systemu

$$\mathfrak{C}'_k = \{A_k \cap \dots \cap A_n; A_i \in \mathfrak{C}_i \text{ dla } i \geq k\}.$$

Dla dowodu (ii) wystarczy teraz rozważyć π -system $\mathfrak{C}''_k = \bigcup_{n > k} \sigma\left(\bigcup_{i=k}^n \mathfrak{C}_i\right)$. $\square$

Definicja 2.1.22. *Dla dowolnego ciągu σ -ciał $(\mathfrak{F}_n)_{n \in \mathbb{N}}$ w przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$ przekrój*

$$\mathfrak{F}_\infty = \bigcap_{m \in \mathbb{N}} \sigma(\mathfrak{F}_{m+1} \cup \mathfrak{F}_{m+2} \cup \dots) \quad (2.22)$$

nazywamy σ -ciałem resztowym lub ogonem ciągu. Zdarzenia $A \in \mathfrak{F}_\infty$ będziemy nazywać zbiorami resztowymi.

Przykład 2.1.9. Dla dowolnego ciągu zmiennych losowych $(X_n)_{n \in \mathbb{N}}$ pytanie o ewentualną zbieżność ciągu prowadzi do wyróżnienia w $\sigma(X_1, X_2, \dots)$ zdarzeń (zbiorów) złożonych z tych $\omega \in \Omega$, dla których ciąg ma granicę skończoną,

$$A = \{\omega; \forall_{\varepsilon > 0} \exists_{n_0 \in \mathbb{N}} \forall_{n \geq n_0} |X_n(\omega) - X_{n_0}(\omega)| < \varepsilon\},$$

lub nieskończoną $A^+ = \{\omega; \forall_{M > 0} \exists_{n_0 \in \mathbb{N}} \forall_{n \geq n_0} X_n(\omega) > M\}$ i odpowiednio A^- dla rozbieżności do $-\infty$. Z kolei zdarzenia

$$A^u = \{\omega; \exists_M \forall_{n \in \mathbb{N}} X_n(\omega) \leq M\} \quad \text{ i } \quad A^l = \{\omega; \exists_M \forall_{n \in \mathbb{N}} X_n(\omega) \geq -M\}$$

opisują ograniczoność badanego ciągu (z góry lub z dołu). Dla dowolnego $m \in \mathbb{N}$ każdy ze zbiorów można przedstawić w postaci np.

$$A = X^{-1}\{x \in \mathbb{R}^{\mathbb{N}}; \forall_k \exists_{n_0 > m} \forall_{n \geq n_0} |x_n - x_{n_0}| < \frac{1}{k}\} = X^{-1}(\mathbb{R}^m \times B_m) \text{ lub}$$

$$A^u = X^{-1}\{x \in \mathbb{R}^{\mathbb{N}}; \exists_M \exists_{n_0 > m} \forall_{n \geq n_0} x_n \leq M\} = X^{-1}(\mathbb{R}^m \times B_m^u),$$

gdzie $X := (X_n): \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$, a $B_m, B_m^u \in \mathfrak{B}_{\mathbb{R}^{\mathbb{N}}} = \bigotimes_{n \in \mathbb{N}} \mathfrak{B}_{\mathbb{R}}$ (prz. 1.1.2). Opisane wyżej zdarzenia należą do resztowego σ -ciała $\bigcap_m \sigma(X_{m+1}, X_{m+2}, \dots)$.

Wykorzystana w przykładzie terminologia wiąże zmienne losowe i odpowiadające im (najmniejsze – [2, par. 2.2.1]) σ -ciała.

Definicja 2.1.23. Dla dowolnego ciągu zmiennych losowych $X_k, k \in \mathbb{N}$ lub $k \leq n, n \in \mathbb{N}$, rodzinę zdarzeń

$$\sigma(X_1, X_2, \dots) := \sigma\left(\bigcup_k X_k^{-1} \mathfrak{B}_{\mathbb{R}}\right)$$

nazywamy σ -ciałem generowanym przez zmienne losowe. Dla skończonego ciągu oznacza to równość $\sigma(X_1, \dots, X_n) = X^{-1} \mathfrak{B}_{\mathbb{R}^n}$.

Prezentowane w kolejnych podrozdziałach raczej zaskakujące (dla niewtajemniczonych) własności graniczne zmiennych losowych i ich rozkładów mają często u swych źródeł stwierdzenie 2.1.18 oraz

Twierdzenie 2.1.24 (Prawo 0/1 Kołmogorowa). Jeśli σ -ciała $\mathfrak{F}_n, n \in \mathbb{N}$, są niezależne, to

$$\mathbb{P}(A) = 0 \quad \text{ albo } \quad \mathbb{P}(A) = 1$$

dla każdego zdarzenia resztowego $A \in \mathfrak{F}_\infty$.

Dowód. Dla każdego $m \in \mathbb{N}$ z wniosku 2.1.21(ii) wynika niezależność ciągu $\mathfrak{F}_1, \dots, \mathfrak{F}_m$ oraz σ -ciała $\sigma(\bigcup_{n>m} \mathfrak{F}_n)$, a więc także ciągu $\mathfrak{F}_\infty, \mathfrak{F}_1, \dots, \mathfrak{F}_m$. Oznacza to niezależność ciągu nieskończonego i – w konsekwencji – niezależność σ -ciał $\mathfrak{F}_\infty$ i $\sigma(\bigcup_{n \in \mathbb{N}} \mathfrak{F}_n) \supset \mathfrak{F}_\infty$. Tym samym każde zdarzenie resztowe $A \in \mathfrak{F}_\infty$ jest niezależne ze sobą, czyli $\mathbb{P}(A \cap A) = \mathbb{P}(A) \cdot \mathbb{P}(A)$. $\square$

Przypominamy podstawowe pojęcia dotyczące niezależności zmiennych. Zgodnie z twierdzeniem 2.1.27, poniżej, niezależność zmiennych oznacza w istocie niezależność σ -ciał.

Definicja 2.1.25. *Zmienne losowe $X_1, \dots, X_n: \Omega \rightarrow \mathbb{R}$ określone na tej samej przestrzeni probabilistycznej uważamy za niezależne, jeśli dla dowolnych zbiorów borelowskich $A_1, \dots, A_n \in \mathfrak{B}_\mathbb{R}$, zdarzenia $\{X_1 \in A_1\}, \dots, \{X_n \in A_n\}$ są niezależne.*

Nieskończony ciąg zmiennych losowych $X_i: \Omega \rightarrow \mathbb{R}, i \in \mathbb{N}$, jest niezależny, a tworzące go zmienne uznajemy za niezależne, jeśli dla każdego $n \in \mathbb{N}$ ciąg $X_1, \dots, X_n$ jest niezależny.

Ze stwierdzenia 2.1.18 oraz twierdzenia 2.1.24 wynika zatem

Wniosek 2.1.26. *Dla dowolnego ciągu niezależnych zmiennych losowych $(X_n)_{n \in \mathbb{N}}$ własności takie jak ograniczoność, skończoność granicy dolnej lub górnej oraz zbieżność zachodzą prawie na pewno, jeśli tylko mają prawdopodobieństwo niezerowe. Granice ciągu – skończone lub nieskończone – są funkcjami stałymi (pw.).*

Dowód. Zdefiniowane w przykładzie 2.1.9 zbiory A, A^u, A^l mają prawdopodobieństwo 0 lub 1. To znaczy – ciąg jest ograniczony prawie wszędzie lub prawie nigdzie, a jeśli któraś ze zbieżności ma prawdopodobieństwo niezerowe, to zachodzi prawie na pewno. W analogiczny sposób sprawdzamy, że każda z funkcji $\limsup_n X_n, \liminf_n X_n$ – jeśli jest (prawie wszędzie) skończona, to jest mierzalna względem σ -ciała resztowego, a więc – stała. $\square$

Przykład 2.1.10. Ćwiczenie [2, ćw. 2.3.7] dotyczy równości

$$\ell(\varepsilon_{i_1}^{-1}\{\delta_1\} \cap \dots \cap \varepsilon_{i_n}^{-1}\{\delta_n\}) = \prod_{k \leq n} \ell(\varepsilon_{i_k}^{-1}\{\delta_k\}) = 2^{-n},$$

dla dowolnego n -elementowego ciągu indeksów $1 \leq i_1 < \dots < i_n$ i ciągu cyfr $\delta_k \in \{0, 1\}$, $k \leq n \in \mathbb{N}$, co oznacza niezależność ciągu cyfr binarnych – zmiennych losowych $\varepsilon_n: [0, 1] \rightarrow \{0, 1\}$, dla $n \in \mathbb{N}$. Każda ze zmiennych ma ten sam rozkład, przy czym

$$\mathbb{P}_{\varepsilon_n}\{0\} = \mathbb{P}_{\varepsilon_n}\{1\} = \frac{1}{2} \quad \text{dla } n \in \mathbb{N}.$$

Przykład 2.1.11. W dowolnej przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$ każdy skończony ciąg zdarzeń elementarnych $\omega = (\omega_1, \dots, \omega_n) \in \Omega^n$ nazywamy *próbką n -elementową*. Wartości ustalonej zmiennej losowej dla danej próbki możemy interpretować jako wartości zmiennych losowych

$$X_i(\omega) = X(\omega_i), \quad \text{czyli } X_i = X \circ \pi_i: \Omega^n \rightarrow \mathbb{R}, \quad i \leq n,$$

gdzie przestrzeń probabilistyczną Ω^n rozważamy z miarą produktową. Powstałe w ten sposób zmienne losowe $X_i, i \leq n$, są niezależne i mają ten sam rozkład. Ciąg niezależnych zmiennych losowych o zadanych z góry rozkładach jest zatem możliwy do zrealizowania np. w przestrzeni produktowej.

Twierdzenie 2.1.27 (Niezależność zmiennych losowych). *Poniższe warunki są równoważne niezależności zmiennych losowych $X_1, \dots, X_n: \Omega \rightarrow \mathbb{R}$.*

- (i) σ -ciała $\sigma(X_i) := (X_i)^{-1}\mathfrak{B}_{\mathbb{R}} \subset \mathfrak{F}, i \leq n$, są niezależne.
- (ii) Dla dowolnych zbiorów borelowskich $A_1, \dots, A_n \subset \mathbb{R}$, zachodzi równość

$$\mathbb{P}\{X_1 \in A_1, \dots, X_n \in A_n\} = \mathbb{P}\{X_1 \in A_1\} \cdot \dots \cdot \mathbb{P}\{X_n \in A_n\}. \quad (2.23)$$

- (iii) Dystrybuanta $F_X: \mathbb{R}^n \rightarrow [0, 1]$ wektora $X = (X_1, \dots, X_n)$ określona wzorem

$$F_X(a_1, \dots, a_n) = \mathbb{P}\{\omega \in \Omega; X_1(\omega) \leq a_1, \dots, X_n(\omega) \leq a_n\}, \quad (2.24)$$

spełnia warunek $F_X(a_1, \dots, a_n) = \prod_{i \leq n} F_{X_i}(a_i)$ dla dowolnych $a_1, \dots, a_n \in \mathbb{R}$.

- (iv) Rozkład $\mathbb{P}_X = X_*\mathbb{P}$ wektora losowego $X = (X_1, \dots, X_n): \Omega \rightarrow \mathbb{R}^n$ jest produktem miar

$$\mathbb{P}_X = \mathbb{P}_{X_1} \otimes \dots \otimes \mathbb{P}_{X_n} \quad (2.25)$$

w borelowskiej przestrzeni zdarzeń $(\mathbb{R}^n, \mathfrak{B}_{\mathbb{R}^n})$.

Zgodnie z twierdzeniem (1.1.45) zdefiniowana w (2.24) dystrybuanta F_X dowolnego wektora losowego $X: \Omega \rightarrow \mathbb{R}^p$ wyznacza jednoznacznie rozkład $\mathbb{P}_X$.

Z zawartej w twierdzeniu 2.1.2 charakteryzacji produktu miar probabilistycznych wynika zatem

Wniosek 2.1.28. *Ciąg zmiennych losowych $X_n: \Omega \rightarrow \mathbb{R}, n \in \mathbb{N}$, jest niezależny wtedy i tylko wtedy, gdy rozkład $\mathbb{P}_X = X_*\mathbb{P}$ wektora losowego $X = (X_n)_{n \in \mathbb{N}}: \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$ jest produktem miar*

$$\mathbb{P}_X = \bigotimes_{n \in \mathbb{N}} \mathbb{P}_{X_n} \quad (2.26)$$

w przestrzeni zdarzeń $(\mathbb{R}^{\mathbb{N}}, \mathfrak{B}_{\mathbb{R}^{\mathbb{N}}})$. □

Ćwiczenie 2.1.12. Sprawdzić, że dowolna borelowska miara probabilistyczna w $\mathbb{R}^{\mathbb{N}}$ jest produktem postaci (2.26) wtedy i tylko wtedy, gdy współrzędne (projekcje) $\pi_n: \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}$ są niezależnymi zmiennymi losowymi.

Dowód twierdzenia 2.1.27. Własność (i) jest równoważnym z definicją opisem niezależności. Stwierdzenie 2.1.19 implikuje równoważność (i) $\Leftrightarrow$ (ii).

(ii) $\Rightarrow$ (iii) Własność sformułowana w (iii) dla dystrybuant jest szczególnym przypadkiem równości (2.23) – dla zbiorów $A_i = (-\infty, a_i]$, $i \leq n$.

(iii) $\Rightarrow$ (iv) Jeśli prawdziwa jest podana w twierdzeniu zależność dla dystrybuant, to miara borelowska $\mathbb{P}_X$ oraz produkt $\mathbb{P}_{X_1} \otimes \dots \otimes \mathbb{P}_{X_n}$ są równe na tworzących π -system zbiorach postaci $(-\infty, a_1] \times \dots \times (-\infty, a_n]$. Równość (2.25) wynika zatem z twierdzenia 1.1.29 (o jednoznaczności miar) oraz z faktu, że wyróżnione tu nieskończone przedziały generują $\mathfrak{B}_{\mathbb{R}^n}$.

(iv) $\Rightarrow$ (ii) Zgodnie z definicją produktu miar zachodzi równość

$$\mathbb{P}_X(A_1 \times \dots \times A_n) = \mathbb{P}_{X_1}(A_1) \cdot \dots \cdot \mathbb{P}_{X_n}(A_n)$$

dla dowolnych zbiorów $A_1, \dots, A_n \in \mathfrak{B}_{\mathbb{R}}$. Jest to własność równoważna (2.23). $\square$

Ze względu na własność (i), istnienie niezależnych zmiennych losowych $(X_k)_{k \in \mathbb{N}}$ w przestrzeni $(\Omega, \mathfrak{F}, \mathbb{P})$ pociąga za sobą istnienie odpowiedniej rodziny niezależnych σ -ciał w $\mathfrak{F}$.

Stwierdzenie 2.1.29 (Niezależność grup zmiennych). *Założmy, że zmienne $X_1, \dots, X_n$ są niezależne. Wówczas*

(i) *dla dowolnych liczb naturalnych $0 = k_0 < k_1 < \dots < k_m = n$ każdy ciąg funkcji borelowskich $f_i: \mathbb{R}^{k_i - k_{i-1}} \rightarrow \mathbb{R}$, $i \leq m$, wyznacza niezależne zmienne losowe $f_i(X_{k_{i-1}+1}, \dots, X_{k_i})$, $i \leq m$;*

(ii) *dla dowolnej funkcji borelowskiej $f: \mathbb{R}^n \rightarrow \mathbb{R}$ rozkład prawdopodobieństwa zmiennej $f(X_1, \dots, X_n)$ jest miarą*

$$f(X_1, \dots, X_n)_* \mathbb{P} = f_*(\mathbb{P}_{X_1} \otimes \dots \otimes \mathbb{P}_{X_n})$$

i zależy tylko od f i rozkładów poszczególnych zmiennych.

Dowód. Ćwiczenie. $\square$

Stwierdzenie 2.1.30. *Dla dowolnego zbioru borelowskiego $B \subset \mathbb{R}^2$ i dowolnej pary niezależnych zmiennych losowych $X, Y: \Omega \rightarrow \mathbb{R}$*

$$\mathbb{P}\{\omega \in \Omega; (X(\omega), Y(\omega)) \in B\} = \int \mathbb{P}\{\omega \in \Omega; (x, Y(\omega)) \in B\} \mathbb{P}_X(dx).$$

W szczególności

$$\mathbb{P}_{X+Y}(A) = \int \mathbb{P}_Y(A-x) \mathbb{P}_X(dx) \quad \text{dla } A \in \mathfrak{B}_{\mathbb{R}}, \quad (2.27)$$

gdzie zbiór $A-x = \{a-x; a \in A\}$ powstaje z A w wyniku translacji, oraz

$$F_{X+Y}(x) = \int F_Y(t-x) \mathbb{P}_X(dt) \quad \text{dla } x \in \mathbb{R}.$$

Dowód. Skorzystać z twierdzenia Fubiniego (wn. 1.2.25). $\square$

Wniosek 2.1.31. Jeśli zmienne niezależne $X, Y: \Omega \rightarrow \mathbb{R}$ mają rozkład ciągły o gęstości, odpowiednio f i g , to ich splot, czyli funkcja

$$f * g(x) = \int_{\mathbb{R}} f(t)g(x-t) dt \quad \text{dla } x \in \mathbb{R}$$

jest gęstością rozkładu sumy $X + Y$. $\square$

Uwaga 2.1.13. Dla dowolnych dwu miar borelowskich μ i ν na prostej miarę $\mu * \nu$ określoną wzorem

$$(\mu * \nu)(A) = \int \nu(A-x) \mu(dx) \quad \text{dla } A \in \mathfrak{B}_{\mathbb{R}}, \quad (2.28)$$

nazywamy *splotem*. Innymi słowy, rozkład sumy niezależnych zmiennych losowych jest splotem rozkładów poszczególnych zmiennych.

Ćwiczenie 2.1.14. Wykazać przemienność i łączność operacji splotu.

Ćwiczenie 2.1.15. Sprawdzić, że dla dowolnej zmiennej losowej X i liczby $m \in \mathbb{R}$ rozkład zmiennej $\max(X, m)$ zależy tylko od liczby m i rozkładu $\mathbb{P}_X$. Wyznaczyć dystrybuanty rozkładu zmiennych $\min(X, Y)$ i $\max(X, Y)$ dla dowolnych dwu niezależnych zmiennych losowych X i Y .

Podstawowe warunki konieczne niezależności pary zmiennych zawiera

Twierdzenie 2.1.32. Dla dowolnych całkowalnych niezależnych zmiennych losowych X, Y iloczyn XY jest również całkowalny i zachodzą równości

$$\mathcal{E}(X \cdot Y) = \mathcal{E}(X)\mathcal{E}(Y), \quad (2.29)$$

$$\text{Var}(aX + bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y) \quad \text{dla } a, b \in \mathbb{R}, \quad (2.30)$$

gdzie druga z równości wymaga, by przynajmniej jedna ze zmiennych miała skończoną wariancję.

Dowód. Aby skorzystać z twierdzenia o zamianie zmiennych (wn. 1.38), oznaczmy przez $h: \mathbb{R}^2 \rightarrow \mathbb{R}$ mnożenie $(x, y) \mapsto |xy|$. Dla dowolnych całkowalnych zmiennych losowych X, Y w przestrzeni $(\Omega, \mathfrak{F}, \mathbb{P})$ całka z funkcji $|XY| = h \circ (X, Y)$ wyraża się wzorem

$$\int |XY| d\mathbb{P} = \int h d\mathbb{P}_{(X,Y)} = \int |xy| d\mathbb{P}_{(X,Y)}(x, y).$$

Jeśli zmienne są niezależne, z własności (2.25) i twierdzenia Tonellego otrzymujemy

$$\int |XY| d\mathbb{P} = \int |y| \left(\int |x| d\mathbb{P}_X(x) \right) d\mathbb{P}_Y(y) = \mathcal{E}|X| \cdot \mathcal{E}|Y| < \infty.$$

Oznacza to całkowalność zmiennej XY . Równość (2.29) wynika z twierdzenia Fubiniego. Własność (2.30) wyprowadzamy z tożsamości

$$(X+Y - \mathcal{E}(X+Y))^2 = (X - \mathcal{E}(X))^2 + 2(X - \mathcal{E}(X))(Y - \mathcal{E}(Y)) + (Y - \mathcal{E}(Y))^2,$$

gdzie niezależne zmienne $X - \mathcal{E}(X), Y - \mathcal{E}(Y)$ mają wartość oczekiwaną 0. $\square$

2.1.4 Warunkowe rozkłady i wartość oczekiwana

Precyzując oznaczenia, przypominamy to, co wprowadziła definicja 1.3.29, a co pojawiło się jako obiekty (miary) probabilistyczne w ogólniejszym kontekście miarowym – jako wniosek z twierdzenia Radona–Nikodyma. Ograniczamy kontekst do ustalonej przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$.

Definicja 2.1.33. Rozkładem warunkowym $\mathbb{P}_{X|\mathfrak{G}}$ dowolnego wektora losowego $X = (X_1, \dots, X_p): \Omega \rightarrow \mathbb{R}^p$ względem σ -ciała $\mathfrak{G} \subset \mathfrak{F}$ nazywamy każdą $\mathfrak{G}$ -mierzalną rodzinę borelowskich miar probabilistycznych $\mathbb{P}_{X|\mathfrak{G}}(\cdot, \omega)$ w $\mathbb{R}^p$, dla $\omega \in \Omega$, spełniającą warunek (por. twierdzenie 1.3.25)

$$\mathbb{P}\{X^{-1}A \cap G\} = \int_G \mathbb{P}_{X|\mathfrak{G}}(A, \cdot) d\mathbb{P}, \quad \text{dla } A \in \mathfrak{B}_{\mathbb{R}^p}, G \in \mathfrak{G}. \quad (2.31)$$

Jeśli zmienna X jest przeliczalnie całkowalna na $\mathfrak{G}$, jej warunkową wartość oczekiwaną $\mathcal{E}(X|\mathfrak{G})$ względem σ -ciała $\mathfrak{G}$ nazywamy $\mathfrak{G}$ -mierzalną całką relatywną $\Pi_{\mathfrak{G}}(X) = (\Pi_{\mathfrak{G}}(X_i))_{i \leq p}$ scharakteryzowaną przez warunek

$$\int_G \mathcal{E}(X|\mathfrak{G}) d(\mathbb{P}|_{\mathfrak{G}}) = \int_G X d\mathbb{P}$$

dla $G \in \mathfrak{G}$ takich, że $\int_G \|X\| d\mathbb{P} < \infty$.

Pozorną niezgodność przyjętej tu definicji z ogólną regułą, w której wartość oczekiwana jest całką względem rozkładu, niweluje wniosek 1.3.26 sformułowany jako

Twierdzenie 2.1.34. *Zachodzi równość*

$$\mathcal{E}(\varphi(X)|\mathfrak{G}) = \int \varphi(x) \mathbb{P}_{X|\mathfrak{G}}(d\omega, \cdot), \quad (2.32)$$

jeśli φ jest funkcją borelowską na $\mathbb{R}^p$ taką, że złożenie $\varphi(X)$ jest przeliczalnie całkowalne na $\mathfrak{G}$. $\square$

W przypadku, gdy σ -ciało $\mathfrak{G}$ jest jawnie opisane jako generowane przez skończony ciąg zmiennych losowych, przyjmujemy bardziej szczegółową terminologię i oznaczenia.

Definicja 2.1.35. Rozkładem warunkowym $\mathbb{P}_{X|Z}$ dowolnej zmiennej losowej $X: \Omega \rightarrow \mathbb{R}^p$ względem zmiennej wektorowej $Z = (Z_1, \dots, Z_m)$ nazywamy każdą borelowską rodzinę miar probabilistycznych $(\mathbb{P}_{X|Z=z})_{z \in \mathbb{R}^m}$, spełniającą warunek (por. wniosek 1.3.28)

$$\mathbb{P}\{X \in A, Z \in B\} = \int_B \mathbb{P}_{X|Z=z}(A) \mathbb{P}_Z(dz), \quad \text{dla } A, B \in \mathfrak{B}_{\mathbb{R}^p} \times \mathfrak{B}_{\mathbb{R}^m}. \quad (2.33)$$

Jeśli X jest przeliczalnie całkowalna na σ -ciele $\sigma(Z)$, warunkową wartością oczekiwaną $\mathcal{E}(X|Z)$ zmiennej X względem zmiennej Z nazywamy borelowską zmienną losową na $\mathbb{R}^p$ taką, że

$$\mathcal{E}(X|Z) \circ Z = \mathcal{E}(X|\sigma(Z)): \Omega \longrightarrow \mathbb{R}^p. \quad (2.34)$$

Funkcja $\mathcal{E}(X|Z)$ jest wyznaczona jednoznacznie – pw. z dokładnością do zbioru miary 0 względem $\mathbb{P}_Z$. Podstawowy schemat obliczeniowy dla rozkładu warunkowego zapewnia twierdzenie 1.3.30.

Uwaga 2.1.16. Tradycyjnie przyjęte oznaczenie $\mathcal{E}(X|Z=z)$ dla wartości funkcji $\mathcal{E}(X|Z)$ w punktach $z \in \mathbb{R}^m$ ma *warunkowy* sens ograniczony do kontekstu, w którym wyróżniono jedną z funkcji równych pw.

Jeśli $\mathbb{P}\{Z=z\} > 0$, to

$$\mathcal{E}(X|Z=z) = \frac{1}{\mathbb{P}\{Z=z\}} \int X|_{\{Z=z\}} d\mathbb{P}$$

jest wartością średnią zmiennej X obciętej do zbioru $\{Z=z\}$. W tym przypadku rozkład warunkowy ma jednoznacznie określoną wartość

$$\mathbb{P}_{X|Z=z}(A) = \frac{\mathbb{P}(X^{-1}A \cap Z^{-1}\{z\})}{\mathbb{P}\{Z=z\}} \equiv \mathbb{P}(\{X \in A\}|Z=z) \quad \text{dla } A \in \mathfrak{B}_{\mathbb{R}^p}$$

i jest klasycznym *prawdopodobieństwem warunkowym*. Z drugiej strony, za paradoks przyjętej terminologii można uznać fakt, iż zależność (2.34) dotyczy wartości funkcji borelowskiej $\mathcal{E}(X|Z)$ na zbiorze $Z(\Omega)$, który borelowski

być nie musi. Możliwa jest zatem niezerowa wartość $\mathcal{E}(X|Z=z)$ także wtedy, gdy zbiór $\{Z=z\}$ jest pusty. Za kolejne wyjaśnienie konwencji zapisu *warunku* w postaci „zmienna = wartość” możemy przyjąć niemal oczywistą własność, którą formułujemy w nieco szerszym kontekście.

Stwierdzenie 2.1.36. *Załóżmy, że w przestrzeni $(\Omega, \mathfrak{F})$ dana jest dowolna σ -skończona miara P . Jeśli zmienna warunkująca jest złożeniem $Z = TX$ zmiennej $X: \Omega \rightarrow \mathbb{R}^p$ i funkcji borelowskiej $T: \mathbb{R}^p \rightarrow \mathbb{R}^m$, przy czym miara indukowana $P_{TX} = Z_*P$ jest σ -skończona, wówczas dla P_{TX} -prawie wszystkich $y \in \mathbb{R}^m$ funkcja T jest prawie na pewno stała względem miary warunkowej $P_{X|TX=y}$, tzn.*

$$P_{TX}\{y \in \mathbb{R}^m; P_{X|TX=y}\{T \neq y\} > 0\} = 0. \quad (2.35)$$

Dowód. Przyjęte założenia implikują istnienie rozkładu warunkowego. Kładąc we wzorze (1.76) $f(x, y) \equiv \|T(x) - y\|$, z tożsamości

$$0 = \int \|TX - TX\| dP = \iint \|T(x) - y\| dP_{X|TX=y}(x) dP_{TX}(y)$$

otrzymujemy tezę. $\square$

Zależność warunkowej wartości oczekiwanej od rozkładu warunkowego precyzuje

Twierdzenie 2.1.37. *Zachodzi równość*

$$\mathcal{E}(f(X, Z)|Z=z) = \int f(x, z) \mathbb{P}_{X|Z=z}(dx) \quad \text{dla } z \in \mathbb{R}^m \text{ (pw)}, \quad (2.36)$$

jeśli $f: \mathbb{R}^{p+m} \rightarrow \mathbb{R}$ jest funkcją borelowską taką, że złożenie $f(X, Z)$ jest przeliczalnie całkowalne względem $\sigma(Z)$.

Dowód. Zawarta w tezie wniosku 1.3.28 własność (1.76) oznacza równość

$$\int_G f(X, Z) d\mathbb{P} = \int_G \left(\int f(x, Z(\omega)) \mathbb{P}_{X|Z=Z(\omega)}(dx) \right) \mathbb{P}(d\omega), \quad \text{dla } G \in \sigma(Z),$$

a zatem prawa strona dowodzonej równości (2.36) spełnia warunek (2.34) charakteryzujący warunkową wartość oczekiwaną $\mathcal{E}(f(X, Z)|Z)$. $\square$

Wniosek 2.1.38. *Każda całkowalna zmienna losowa X ma warunkową wartość oczekiwaną względem dowolnej zmiennej losowej Z , oraz – względem dowolnego σ -ciała zdarzeń $\mathfrak{G}$ przy czym*

$$\mathcal{E}(X) = \int \mathcal{E}(X|Z=z) d\mathbb{P}_Z(z) = \int \mathcal{E}(X|\mathfrak{G}) d\mathbb{P}. \quad \square$$

Przykład 2.1.17. Dla *każdej* wektorowej zmiennej losowej X i dla $Z \equiv X$ warunek (2.34) w postaci

$$\mathbb{P}\{X \in A \cap B\} = \int_B \mathbb{P}_{X|X=x}(A) \mathbb{P}_X(dx)$$

spełnia ta sama rodzina miar Diraca $\mathbb{P}_{X|X=x} = \delta_x$, $x \in \mathbb{R}^p$. Warunkową wartością oczekiwaną jest zatem np. funkcja tożsamościowa $\mathcal{E}(X|X=x) = x$. W przypadku, gdy zmienne $X_i, i \leq p$, są niezależne i mają ten sam rozkład, z symetrii wynika wzór

$$\mathcal{E}(X_i|X_1 + \dots + X_p = x) = \frac{x}{p} \quad \text{dla } i \leq p, x \in \mathbb{R}.$$

Przykład 2.1.18. Dla $n \in \mathbb{N}$ w przestrzeni ciągów $\{0, 1\}^n$ wzór

$$\mathbb{P}_z^n\{x\} = z^y(1-z)^{n-y}, \quad \text{gdzie } y = \sum_{i \leq n} x_i, \text{ dla } x = (x_i),$$

gdzie przyjmujemy $0^0 = 1$, daje rozkład prawdopodobieństwa y sukcesów dla schematu n prób *Bernoulliego* z prawdopodobieństwem sukcesu $z \in [0, 1]$. Ten sam wzór definiuje gęstość

$$\varphi(x, z) = z^{\Sigma x}(1-z)^{n-\Sigma x}, \quad \text{dla } (x, z) \in \Omega = \{0, 1\}^n \times [0, 1],$$

miary probabilistycznej $\mathbb{P} \ll \rho \otimes \ell|_{[0,1]}$ bezwzględnie ciągłej względem produktu miary zliczającej ρ i miary Lebesgue'a. Jeśli zmienne $X = (X_1, \dots, X_n)$ i Z oznaczają odpowiednie projekcje, korzystamy z twierdzenia 1.3.30 i otrzymujemy rozkład warunkowy $\mathbb{P}_{X|Z=z} = \mathbb{P}_z^n$, przy czym $\mathbb{P}_Z = \ell|_{[0,1]}$ jest rozkładem brzegowym.

Zależność gęstości od zmiennej x wyraża się poprzez sumę $\sum x_i$, co kieruje naszą uwagę na przestrzeń $\tilde{\Omega} = \{0, 1, \dots, n\} \times [0, 1]$ z miarą $\tilde{\mathbb{P}}$ indukowaną przez odwzorowanie $(\Sigma, id): (x, z) \mapsto (\sum_{i \leq n} x_i, z)$. Z równości

$$\tilde{\mathbb{P}}(\{y\} \times B) = \mathbb{P}(\Sigma^{-1}(y) \times B) = \binom{n}{y} \cdot \int_B z^y(1-z)^{n-y} dz$$

wynika, że miara $\tilde{\mathbb{P}}$ ma gęstość $\tilde{\varphi}(y, z) = \binom{n}{y} z^y(1-z)^{n-y}$ względem produktu $\rho_{\{0,1,\dots,n\}} \otimes \ell|_{[0,1]}$. Dla projekcji (Y, Z) z $\tilde{\Omega}$ na poszczególne osie – dyskretną i ciągłą – obie miary brzegowe są jednostajne

$$\tilde{\mathbb{P}}_Y = \frac{1}{n+1} \cdot \rho \quad \text{ i } \quad \tilde{\mathbb{P}}_Z = \ell|_{[0,1]},$$

a rozkład warunkowy tworzy rodzina *rozkładów Bernoulliego*

$$\tilde{\mathbb{P}}_{Y|Z=z}\{y\} = \binom{n}{y} z^y(1-z)^{n-y}, \quad \text{dla } z \in [0, 1], 0 \leq y \leq n, \quad (2.37)$$

zliczających liczbę sukcesów w n -krotnym schemacie Bernoulliego. Z drugiej strony rodzina miar

$$\tilde{\mathbb{P}}_{Z|Y=y}(B) = (n+1) \int_B \binom{n}{y} z^y (1-z)^{n-y} dz, \quad \text{dla } B \subset [0, 1], 0 \leq y \leq n, \quad (2.38)$$

stanowi przykład tzw. *rozkładów Beta* na przedziale $(0, 1)$. Dokładniej

$$\tilde{\mathbb{P}}_{Z|Y=y} = \text{Beta}(y+1, n-y+1) = \mathbb{P}_{Z|\Sigma X_i=y}.$$

Przypominamy, że dla dowolnych $\alpha, \beta > 0$ rozkład $\text{Beta}(\alpha, \beta)$ na przedziale $(0, 1)$ charakteryzuje gęstość

$$f(x|\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad (2.39)$$

przy czym *funkcja gamma* Γ ma wartości $\Gamma(\gamma) = (\gamma-1)!$, dla liczb całkowitych $\gamma \in \mathbb{N}$.

2.1.5 Charakterystyki wektorów losowych

Definicja 2.1.39. Kowariancją dowolnej pary $X, Y: \Omega \rightarrow \mathbb{R}$ zmiennych losowych nazywamy liczbę

$$\text{Cov}(X, Y) = \mathcal{E}((X - \mathcal{E}(X))(Y - \mathcal{E}(Y))), \quad (2.40)$$

jeśli wskazana wartość oczekiwana istnieje.

Oczywiście $\text{Cov}(X, X) = \text{Var}(X)$.

Wniosek 2.1.40. (i) Kowariancja $\text{Cov}(X, Y)$ istnieje, przy czym

$$|\text{Cov}(X, Y)| \leq \sqrt{\text{Var}(X) \text{Var}(Y)},$$

jeśli obie zmienne mają skończoną wariancję.

(ii) $\text{Cov}(X, Y) = 0$, jeśli zmienne X, Y są niezależne.

Dowód. Własność (i) jest klasyczną nierównością Schwarza. Własność (ii) wynika z (2.29). $\square$

Definicja 2.1.41. Zmienne losowe $X, Y: \Omega \rightarrow \mathbb{R}$ o własności $\text{Cov}(X, Y) = 0$ nazywamy nieskorelowanymi. Dla zmiennych o niezerowej wariancji liczba

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}} \in [-1, 1]$$

jest współczynnikiem korelacji zmiennych.

W zależności od znaku liczby $\rho(X, Y)$ mówimy także o *dodatniej* lub *ujemnej korelacji* zmiennych losowych, co można interpretować jako zgodny lub przeciwny przeciętny kierunek wzrostu.

Naturalne rozszerzenie pojęcia wartości oczekiwanej na zmienne losowe o wartościach wektorowych (i macierzowych) – całkowanie *po współrzędnych* – prowadzi do rozszerzenia definicji 2.1.39 i wzoru (2.40). Zwracamy uwagę na klasyczną konwencję (notację), w której ciąg liczb interpretowany jako wektor staje się macierzą 1-kolumnową, co jednoznacznie podkreśla użycie operatora transpozycji.

Definicja 2.1.42. Kowariancją wektora losowego $X = (X_1, \dots, X_n)^t$ zmiennych losowych o skończonej wariancji nazywamy macierz symetryczną

$$\text{Cov}(X) = [\text{Cov}(X_i, X_j)]_{i,j \leq n} = \mathcal{E}((X - \mathcal{E}(X))(X - \mathcal{E}(X))^t). \quad (2.41)$$

Kowariancja pozwala na włączenia do rozważań pojęcia *liniowej zależności* zmiennych tworzących wektor losowy X i prowadzi do następującej charakterystyki

Stwierdzenie 2.1.43. Niech $X = (X_1, \dots, X_n)^t$ będzie dowolnym wektorem losowym złożonym ze zmiennych o skończonej wariancji.

(i) Dla dowolnego wektora $a \in \mathbb{R}^n$ zachodzą równości

$$\mathcal{E}(a^t X) = a^t \mathcal{E}(X) \quad \text{ i } \quad \text{Var}(a^t X) = a^t \text{Cov}(X) a, \quad (2.42)$$

przy czym $\text{Cov}(AX) = A \text{Cov}(X) A^t$ dla dowolnej macierzy $A \in \mathbb{R}^{m \times n}$.

(ii) Macierz kowariancji $\text{Cov}(X)$ jest nieujemnie określona, a jest macierzą osobliwą wtedy i tylko wtedy, gdy zmienne X_i , $i \leq n$, oraz stała $X_0 \equiv 1$ są liniowo zależne (pw).

Wniosek 2.1.44. Zmienne losowe X, Y o skrajnej wartości współczynnika korelacji $\rho(X, Y) = \pm 1$ wiążą liniową zależność

$$X - \rho(X, Y)Y = \text{const} \quad (\text{pw}) \quad \square$$

Dowód. Wzory (2.42) korzystają z notacji macierzowej (dla iloczynu macierzy i wektorów) i wynikają z liniowości wartości oczekiwanej oraz wzoru (2.41). Nieujemność wariancji implikuje zatem określoność macierzy $\text{Cov}(X)$. Dla $0 \neq a \in \mathbb{R}^n$ zależność $\text{Cov}(X)a = 0$ jest równoważna zanikaniu wariancji $\text{Var}(a^t X) = 0$, a to oznacza równości

$$\begin{aligned} \mathcal{E}(a_1 X_1 + \dots + a_n X_n - c)^2 &= 0 \\ a_1 X_1 + \dots + a_n X_n - c &= 0 \quad (\text{pw}). \end{aligned}$$

Druga z równości, jeśli zachodzi, pozwala wyznaczyć $c = \mathcal{E}(a^t X)$. $\square$

Bardziej szczegółowa analiza wektora losowego wykorzystuje diagonalizację dowolnej nieujemnie określonej macierzy symetrycznej $W \in \mathbb{R}^{n \times n}$ – wg wartości i wektorów własnych. Przypominamy, że w $\mathbb{R}^n$ istnieje baza ortonormalna złożona z wektorów własnych $v_1, \dots, v_n$ o wartościach własnych $\lambda_i = v_i^t W v_i \geq 0$ dla $i \leq n$. Równości $W v_i = \lambda_i v_i$ dla $i \leq n$ implikują rozkład

$$W = V \cdot \Delta \cdot V^{-1}, \quad (2.43)$$

gdzie macierz $V = [v_1 \dots v_n]$ utworzona z wektorów własnych (kolumn) spełnia warunek ortogonalności $V^t = V^{-1}$, a w macierzy diagonalnej Δ wartości własne wypełniają główną przekątną. Zakładając porządek

$$\lambda_1 \geq \dots \geq \lambda_n \geq 0 \quad (2.44)$$

i wynikającą stąd numerację wektorów własnych stwierdzamy, że ostatnią niezerową wartością własną jest λ_r , gdzie $r = \text{rank}(W) \leq n$ oznacza rząd badanej macierzy.

Twierdzenie 2.1.45. *Niech $X = (X_1, \dots, X_n)^t$ będzie dowolnym wektorem losowym o wartości oczekiwanej $m = \mathcal{E}(X)$ i macierzy kowariancji $\text{Cov}(X)$ rzędu r . Istnieje r -wymiarowy wektor losowy $Z = (Z_1, \dots, Z_r)^t$ o nieskorelowanych współrzędnych oraz macierz P wymiaru $n \times r$, takie że*

$$X = m + P \cdot Z \text{ (pw)} \quad i \quad P^t P = I_r. \quad (2.45)$$

Dowód. Korzystając z rozkładu (2.43)–(2.44) macierzy $W = \text{Cov}(X)$, wystarczy przyjąć za P macierz utworzoną z pierwszych r kolumn macierzy V , a więc wektorów własnych $v_i, i \leq r$. Ewentualne zerowe wiersze macierzy Δ , o indeksach powyżej r , redukują wzór (2.43) do postaci $W = P \cdot \Delta' \cdot P^t$, gdzie

$$\Delta' = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_r \end{bmatrix},$$

jest macierzą kowariancji wektora losowego $Z := P^t(X - m)$. Istotnie,

$$\text{Cov}(Z) = P^t \text{Cov}(X) P = P^t P \Delta' P^t P = \Delta'. \quad \square$$

Definicja 2.1.46. *Zmienne losowe $Z_i, i \leq \text{rank Cov}(X)$, nazywamy składowymi głównymi, a kolumny macierzy P – kierunkami głównymi wektora X , jeśli mają własności wyróżnione w twierdzeniu 2.1.45.*

Składowe i kierunki główne wektora losowego zamieniają częściową korelację oryginalnych współrzędnych na wzajemnie ortogonalne, dominujące składniki objaśniające częściową zmienność czyli *wariancję wektora*.

Definicja 2.1.47. Ślad macierzy kowariancji wektora losowego X ,

$$\text{Var}(X) := \text{tr Cov}(X) = \text{Var}(X_1) + \dots + \text{Var}(X_n) \quad (2.46)$$

nazywamy całkowitą wariancją wektora X .

Wniosek 2.1.48. Zachodzi równość $\text{Var}(X) = \text{Var}(Z) = \lambda_1 + \dots + \lambda_r$, gdzie $\lambda_i, i \leq r$, są wartościami własnym macierzy kowariancji. $\square$

Uwaga 2.1.19. W przypadku danych empirycznych, gdy wartości zmiennych losowych X_i są obciążone błędem pomiaru, z reguły wszystkie wartości własne macierzy kowariancji są niezerowe. Dodatkowy warunek postaci $\lambda_i < \varepsilon$ lub $\sum_{i \geq k} \lambda_i < \varepsilon \text{Var}(X)$ pozwala wówczas odrzucić *nieistotne* kierunki główne i oceniać *empiryczną liniową zależność* badanych zmiennych losowych. Podejście empiryczne, w ujęciu właściwym dla *Statystyki* – a więc teoretycznym, ale ukierunkowanym na zastosowania – rozwijamy dalej w podrozdziale 2.2.3.

2.2 Prawo wielkich liczb

Tytuł bieżącego podrozdziału kryje (dla niewtajemniczonych) rodzinę twierdzeń o granicznym zachowaniu ciągów i szeregów wyznaczonych z reguły przez ciągi niezależnych zmiennych losowych.

2.2.1 Dokąd prowadzi błędzenie losowe

Dla dowolnego ciągu niezależnych zmiennych losowych $(X_n)_{n \in \mathbb{N}}$ analiza zbieżności szeregu $S_n = X_1 + \dots + X_n, n \in \mathbb{N}$, opisującego tzw. *błędzenie losowe* (ang. random walking) lub, ogólniej, zbieżności ciągów postaci $c_n S_n$ dla $c_n > 0, n \in \mathbb{N}$, jest bardziej interesująca niż badana wyżej zbieżność samego ciągu.

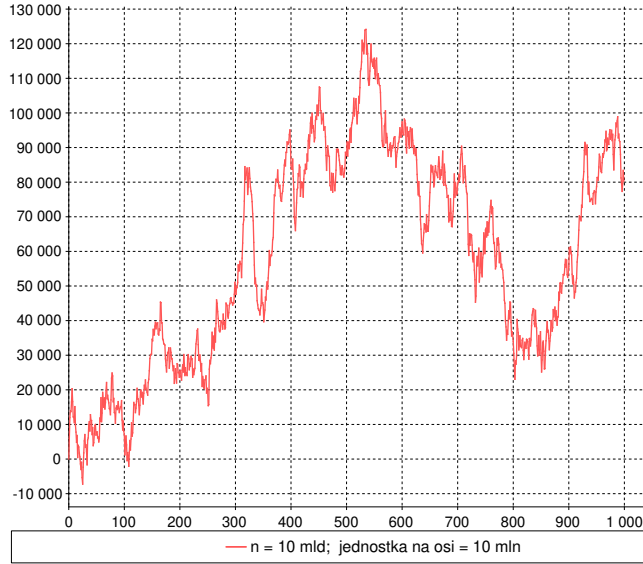
Ćwiczenie 2.2.1. Sprawdzić, że funkcja $\limsup_n c_n S_n: \Omega \rightarrow \overline{\mathbb{R}}$ jest prawie na pewno stała i równa $\pm\infty$ lub jest skończona prawie wszędzie. To samo dotyczy granicy dolnej $\liminf_n c_n S_n$.

W przypadku zmiennych o tym samym rozkładzie mamy alternatywę

Twierdzenie 2.2.1 (O błędzeniu losowym: Alternatywa I). *Jeśli $(X_n)_{n \in \mathbb{N}}$ jest dowolnym ciągiem niezależnych zmiennych losowych o tym samym rozkładzie, to zachodzi dokładnie jedna z wymienionych poniżej możliwości:*

- (i) $S_n \xrightarrow{\text{pw}} 0$ dla $n \in \mathbb{N}$, albo
- (ii) $S_n \longrightarrow +\infty$ prawie na pewno, albo

- (iii) $S_n \longrightarrow -\infty$ *prawie na pewno, albo*
 (iv) $\liminf_n S_n = -\infty$ i $\limsup_n S_n = +\infty$ *prawie na pewno.*



Rysunek 2.1: Spacer wg. rozkładu normalnego

Dowód. Przypuśćmy, że granica górna $f(\omega) = \limsup_n S_n(\omega) \in \mathbb{R}$ jest skończona dla prawie wszystkich $\omega \in \Omega$, a więc jest zmienną losową. Ponieważ dla $c \in \mathbb{R}$

$$f(\omega) < c \iff \exists_{c' < c} \exists_{n_0} (S_n(\omega) < c' \text{ dla } n \geq n_0),$$

σ -ciało $\sigma(f)$ generują zależne od parametru c zbiory $A = X^{-1}B$, gdzie

$$A = \{\exists_{n_0} \forall_{n \geq n_0} S_n < c\} \quad \text{ i } \quad B = \{x \in \mathbb{R}^\infty; \exists_{n_0} \forall_{n \geq n_0} x_1 + \dots + x_n < c\}.$$

Nie są to zbiory resztowe, niemniej jednak mają szczególną własność – podobnie jak funkcja f oraz indukowana w $(\mathbb{R}^\mathbb{N}, \mathfrak{B}_{\mathbb{R}^\mathbb{N}})$ miara $\mathbb{P}_X$ nie zależą od kolejności wyrazów ciągu (por. twierdzenie 2.1.2). W szczególności, dla każdego $m \in \mathbb{N}$ permutacja współrzędnych $\tau_m: \mathbb{R}^\mathbb{N} \rightarrow \mathbb{R}^\mathbb{N}$ postaci

$$\tau_m(x', x'', x''') = (x'', x', x''') \quad \text{ dla } (x', x'', x''') \in \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^\mathbb{N} \equiv \mathbb{R}^\mathbb{N} \quad (2.47)$$

jest funkcją borelowską taką, że $\tau_m(B) = B$. Stąd $\mathbb{P}_X(B) \in \{0, 1\}$. Istotnie, korzystając z twierdzenia *O aproksymacji miar* [2, stw. 1.5.14], znajdujemy ciąg zbiorów borelowskich $B_k \subset \mathbb{R}^{n_k}$, $k \in \mathbb{N}$, takich że

$$\mathbb{P}_X(B \triangle (B_k \times \mathbb{R}^\mathbb{N})) = \varepsilon_k \rightarrow 0.$$

Dla dowolnego ustalonego $k \in \mathbb{N}$ zbiory $C_k = B_k \times \mathbb{R}^{n_k}$ i $D_k = \mathbb{R}^{n_k} \times B_k$ wyznaczają w $\mathbb{R}^{\mathbb{N}}$ dwa niezależne zdarzenia

$$\tilde{C}_k := C_k \times \mathbb{R}^{\mathbb{N}} = B_k \times \mathbb{R}^{\mathbb{N}} \quad \text{ i } \quad \tilde{D}_k := D_k \times \mathbb{R}^{\mathbb{N}} = \tau_{n_k}(B_k \times \mathbb{R}^{\mathbb{N}}),$$

a ponieważ miara $\mathbb{P}_X$ jest τ_{n_k} -niezmiennicza, $\mathbb{P}_X(B \triangle \tilde{C}_k) = \mathbb{P}_X(B \triangle \tilde{D}_k) = \varepsilon_k$ i w konsekwencji $\mathbb{P}_X(\tilde{C}_k \triangle \tilde{D}_k) \leq 2\varepsilon_{n_k}$.

Z tożsamości $\tilde{C}_k \triangle \tilde{D}_k = (\tilde{C}_k \cup \tilde{D}_k) \setminus (\tilde{C}_k \cap \tilde{D}_k)$ otrzymujemy nierówność $\mathbb{P}_X(\tilde{C}_k \cap \tilde{D}_k) \geq \mathbb{P}_X(\tilde{C}_k) - \mathbb{P}_X(\tilde{C}_k \triangle \tilde{D}_k)$ i oszacowanie

$$\mathbb{P}_X(\tilde{C}_k) \geq \mathbb{P}_X(\tilde{C}_k)^2 \geq \mathbb{P}_X(\tilde{C}_k) - 2\varepsilon_{n_k} \quad \text{ dla } k \in \mathbb{N},$$

co w granicy daje równość $\mathbb{P}_X(B) = \mathbb{P}_X(B)^2$, $\mathbb{P}(A) = \mathbb{P}_X(B) \in \{0, 1\}$.

W podanym wyżej rozumowaniu wykorzystaliśmy następującą własność.

Definicja 2.2.2. Podzbiór borelowski $B \subset \mathbb{R}^{\mathbb{N}}$ nazywamy permutowalnym, jeśli jest niezmienniczy względem operatorów (2.47), tzn. $\tau_m(B) = B$ dla każdego $m \in \mathbb{N}$.

Łatwo sprawdzić, że permutowalność jest równoważna niezmienniczości ze względu na permutacje współrzędnych – dowolne, ale ograniczone do skończonej ich liczby. Badając miarę zbioru B udowodniliśmy przy okazji

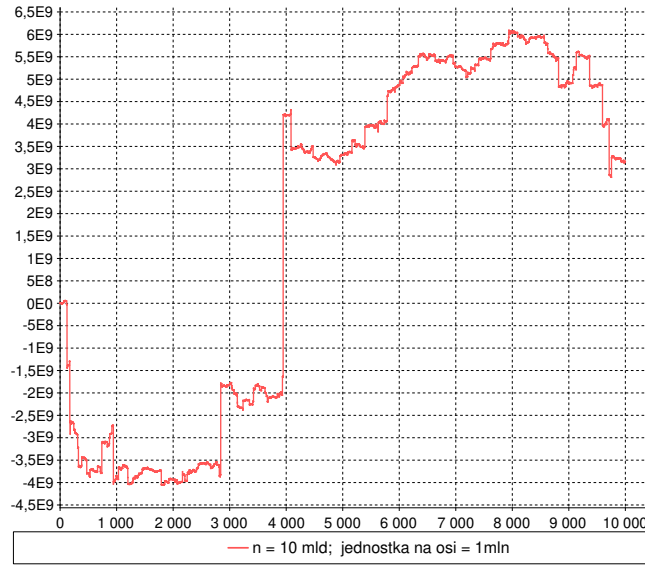
Twierdzenie 2.2.3 (Prawo 0/1 Hewitta–Savage’a). *Dla dowolnego rozkładu prawdopodobieństwa $\mathbb{P}$ na prostej, miara każdego zbioru permutowalnego w przestrzeni $(\mathbb{R}^{\mathbb{N}}, \mathfrak{B}_{\mathbb{R}^{\mathbb{N}}}, \mathbb{P}^{\mathbb{N}})$ jest równa 0 lub 1.* $\square$

Kontynuując dowód twierdzenia 2.2.1, z udowodnionej już własności wnioskujemy, że na mocy stwierdzenia 2.1.18 funkcja $f = \limsup_n S_n$ jest (prawie na pewno) stała. Ponieważ ta sama własność przysługuje granicy

$$\limsup_n S_{n+1} - X_1 = f - X_1,$$

stała (pw.) jest także funkcja X_1 . Taka sama konkluzja wynika z przypuszczenia, że granica dolna $\liminf_n S_n$ jest prawie wszędzie skończona. Z założenia skończoności jednej z granic wynika, że $X_1 \stackrel{\text{pw}}{=} 0$, co odpowiada przypadkowi (i).

Zgodnie z wcześniejszą obserwacją (ćw. 2.2.1) pozostają do rozważenia przypadki, gdy każda z badanych granic jest nieskończona, równa (pw.) $\pm\infty$. W zależności od znaku odpowiada to wymienionym w tezie twierdzenia przypadkom (ii)–(iv). $\square$



Rysunek 2.2: Spacer wg. rozkładu Cauchy'ego

Przykład 2.2.2. Rys. 2.2 przedstawia symulację błędzenia losowego dla ciągu zmiennych o rozkładzie *Cauchy'ego* i gęstości równej $g(x) = \frac{1}{\pi(1+x^2)}$ dla $x \in \mathbb{R}$. Graniczne zachowanie ciągu sum S_k wyjaśnia w tym przypadku

Stwierdzenie 2.2.4. *Jeśli wspólny rozkład każdej z niezerowych niezależnych zmiennych X_k jest symetryczny – taki sam, jak zmiennych $\tilde{X}_k = -X_k$, $k \in \mathbb{N}$, to zachodzi przypadek (iv) i obie granice ciągu sum są nieskończone.*

Dowód. Istotnie, zbiór

$$Z = \{\omega; \forall_M \exists_n S_n(\omega) > M\} = X^{-1}\{x; \forall_M \exists_n x_1 + \dots + x_n > M\}$$

ma taką samą miarę jak zbiór

$$\tilde{Z} = \tilde{X}^{-1}\{x; \forall_M \exists_n x_1 + \dots + x_n > M\} = \{\omega; \forall_M \exists_n S_n(\omega) < -M\},$$

a to oznacza, że z równości $\limsup_n S_n \stackrel{\text{pw}}{=} \infty$ wynika $\liminf_n S_n \stackrel{\text{pw}}{=} -\infty$. $\square$

Przykład 2.2.3. Translacja $x \mapsto x - \frac{1}{2}$,

$$[0, 1] \ni \sum_{n=1}^{\infty} \frac{a_n}{2^n} \mapsto \sum_{n=1}^{\infty} \frac{2a_n - 1}{2^{n+1}} \in \left[-\frac{1}{2}, \frac{1}{2}\right] \quad (2.48)$$

pozwala na alternatywny binarny zapis liczb z przedziału $\Omega = [-\frac{1}{2}, \frac{1}{2}]$ poprzez ciągi liczb ± 1 , przy czym poza zbiorem przeliczalnym rozkład ten jest

jednoznaczny. W przestrzeni probabilistycznej Ω z miarą Lebesgue'a funkcje $X_n = 2a_n - 1: [-\frac{1}{2}, \frac{1}{2}] \rightarrow \{-1, 1\}, n \in \mathbb{N}$, są niezależnymi zmiennymi losowymi o rozkładzie dyskretnym

$$\mathbb{P}\{X_n = 1\} = \mathbb{P}\{X_n = -1\} = \frac{1}{2}, \quad (2.49)$$

a zatem także $\mathcal{E}(X_n) = 0$ i $\text{Var}(X_n) = 1$ dla $n \in \mathbb{N}$. Ponieważ zmienne X_n są symetryczne, z poprzedniego przykładu wynika rozbieżność obu granic

$$\liminf_n S_n \stackrel{\text{pw}}{=} -\infty \quad \text{i} \quad \limsup_n S_n \stackrel{\text{pw}}{=} \infty$$

dla sum $S_n = X_1 + \dots + X_n, n \in \mathbb{N}$, a średnia częstość wyrazów ± 1 w rozkładzie binarnym (2.48) liczb z przedziału $[-\frac{1}{2}, \frac{1}{2}]$ jest dla prawie wszystkich liczb taka sama.

Teza twierdzenia 2.2.1 przybiera szczególnie prostą, czytelną postać w sytuacji, gdy dla rozważanych zmiennych losowych istnieje (wspólna) wartość oczekiwana $\mu = \mathcal{E}(X_1) \in \overline{\mathbb{R}}$. Z rozważań wykluczamy trywialny przypadek $X_n \stackrel{\text{pw}}{=} 0$ dla $n \in \mathbb{N}$.

Twierdzenie 2.2.5 (O błędzeniu losowym II). *Niech $X_n: \Omega \rightarrow \mathbb{R}, n \in \mathbb{N}$, będzie dowolnym ciągiem nietrywialnych niezależnych zmiennych losowych o tym samym rozkładzie i wartości oczekiwanej $\mu = \mathcal{E}(X_1) \in \overline{\mathbb{R}}$. Błędzenie losowe $(S_n)_{n \in \mathbb{N}}$ podlega jednej z reguł (ii)–(iv) twierdzenia 2.2.1 zależnie od tego, czy $\mu > 0$, $\mu < 0$, czy też, odpowiednio, $\mu = 0$.*

Podanego tu założenia nie spełnia np. prezentowany w przykładzie 2.2.2 rozkład Cauchy'ego – a jego zachowanie daje się uzasadnić symetrią rozkładu.

Dla dowodu wykażemy jeszcze dodatkowe ogólne kryterium rozstrzygające, która z potencjalnych możliwości (i)–(iv) jest w przypadku danego ciągu $(X_n)_{n \in \mathbb{N}}$ właściwa.

Stwierdzenie 2.2.6. *Alternatywę sformułowaną w twierdzeniu 2.2.1 jednoznacznie charakteryzują prawdopodobieństwa zdarzeń $\Omega^+ = \{\omega; \forall_n S_n(\omega) \geq 0\}$ i $\Omega^- = \{\omega; \forall_n S_n(\omega) \leq 0\}$:*

- (i) $(\mathbb{P}(\Omega^+) > 0 \text{ i } \mathbb{P}(\Omega^-) > 0) \iff \forall_n S_n = 0 \text{ (pw)}$;
- (ii) $(\mathbb{P}(\Omega^+) > 0 \text{ i } \mathbb{P}(\Omega^-) = 0) \iff S_n \rightarrow \infty \text{ (pw)}$;
- (iii) $(\mathbb{P}(\Omega^+) = 0 \text{ i } \mathbb{P}(\Omega^-) > 0) \iff S_n \rightarrow -\infty \text{ (pw)}$;
- (iv) $\mathbb{P}(\Omega^+) = 0 = \mathbb{P}(\Omega^-) \iff (\limsup_n S_n \stackrel{\text{pw}}{=} \infty \text{ i } \liminf_n S_n \stackrel{\text{pw}}{=} -\infty)$.

Dowód. Z twierdzenia 2.2.1 wynika, że naturalne implikacje dotyczące zbieżności i ograniczoności ciągów,

$$(i') \forall_n S_n = 0 \implies (\sup_n S_n < \infty \text{ i } \inf_n S_n > -\infty),$$

$$(ii') S_n \rightarrow \infty \implies (\sup_n S_n = \infty \text{ i } \inf_n S_n > -\infty),$$

$$(iii') S_n \rightarrow -\infty \implies (\sup_n S_n < \infty \text{ i } \inf_n S_n = -\infty) \text{ oraz}$$

$$(iv') (\limsup_n S_n = \infty \text{ i } \liminf_n S_n = -\infty) \implies (\sup_n S_n = \infty \text{ i } \inf_n S_n = -\infty),$$

prawdziwe także w wariancie (pw.), są równoważnościami – skoro podane własności kresów wzajemnie się wykluczają.

Aby zastosować ten sam schemat rozumowania i zakończyć dowód, zauważmy prawdziwość implikacji

$$(i'') (\mathbb{P}(\Omega^+) > 0 \text{ i } \mathbb{P}(\Omega^-) > 0) \implies (\sup_n S_n < \infty \text{ i } \inf_n S_n > -\infty \text{ (pw)}),$$

$$(ii'') (\mathbb{P}(\Omega^+) > 0 \text{ i } \mathbb{P}(\Omega^-) = 0) \implies (\sup_n S_n = \infty \text{ i } \inf_n S_n > -\infty \text{ (pw)}),$$

$$(iii'') (\mathbb{P}(\Omega^+) = 0 \text{ i } \mathbb{P}(\Omega^-) > 0) \implies (\sup_n S_n < \infty \text{ i } \inf_n S_n = -\infty \text{ (pw)}),$$

$$(iv'') (\mathbb{P}(\Omega^+) = 0 \text{ i } \mathbb{P}(\Omega^-) = 0) \implies (\sup_n S_n = \infty \text{ i } \inf_n S_n = -\infty \text{ (pw)}),$$

będących kombinacją własności

$$\mathbb{P}(\Omega^-) > 0 \implies \mathbb{P}\{\sup_n S_n < \infty\} = 1 \text{ i} \quad (2.50)$$

$$\mathbb{P}(\Omega^-) = 0 \implies \mathbb{P}\{\sup_n S_n = \infty\} = 1, \quad (2.51)$$

oraz analogicznych – dla miary $\mathbb{P}(\Omega^+)$ i kresu dolnego $\inf_n S_n$.

Istotnie, jeśli $\mathbb{P}(\Omega^-) > 0$, to zbiór resztowy

$$\{\limsup_n S_n < \infty\} = \{\omega; \sup_n S_n(\omega) < \infty\} \supset \Omega^-,$$

ma miarę 1. Z drugiej strony, jeśli $\mathbb{P}\{\exists_n X_1 + \dots + X_n > 0\} = 1$, czyli $\mathbb{P}(\Omega^-) = 0$, to ze względu na ten sam rozkład zmiennych (i ich niezależność), miarę 1 ma także każdy ze zbiorów

$$B_m = \{\omega; \exists_{n_m > m} X_{m+1}(\omega) + \dots + X_{n_m}(\omega) > 0\} \quad \text{dla } m \in \mathbb{N},$$

oraz ich przekrój $B = \bigcap_{m \in \mathbb{N}} B_m$. Tym samym dla każdego $\omega \in B$ ciąg $(S_n(\omega))_{n \in \mathbb{N}}$ zawiera dodatni podciąg rosnący, a więc $\limsup_n S_n(\omega) > 0$. Wyklucza to przypadki (i) oraz (iii) i implikuje nieograniczoność ciągu sum z góry. Dowód analogicznych własności

$$\mathbb{P}(\Omega^+) > 0 \implies \mathbb{P}\{\inf_n S_n > -\infty\} = 1, \quad \mathbb{P}(\Omega^+) = 0 \implies \mathbb{P}\{\inf_n S_n = -\infty\} = 1,$$

pozostawiamy czytelnikowi.

Ponieważ poprzedniki udowodnionych wyżej implikacji (i'')–(iv'') obejmują wszystkie możliwe przypadki i wzajemnie się wykluczają, implikacje są równoważnościami. $\square$

Rozważmy funkcje mierzalne $\alpha, \bar{\beta}: \Omega \rightarrow \mathbb{N} \cup \{\infty\}$, takie że

$$\alpha(\omega) = \inf\{n; S_n(\omega) > 0\} \quad \text{ i } \quad \bar{\beta}(\omega) = \inf\{n; S_n(\omega) \leq 0\} \quad \text{ dla } \omega \in \Omega,$$

a więc, w szczególności, $\Omega^- = \{\alpha = \infty\}$ (oraz $\Omega^+ = \{\beta = \infty\}$ dla odpowiednio zdefiniowanej funkcji $\beta \geq \bar{\beta}$) zgodnie z konwencją, że brak stosownego indeksu oznacza kres dolny $\inf \emptyset = \infty$.

Lemat 2.2.7. *Dla każdego ciągu niezależnych zmiennych losowych o tym samym rozkładzie $X_n, n \in \mathbb{N}$, zachodzą równości*

$$\sum_{m=0}^n \mathbb{P}\{\alpha > m\} \cdot \mathbb{P}\{\bar{\beta} > n - m\} = 1 \quad \text{ dla } n \in \mathbb{N}, \quad (2.52)$$

$$\mathbb{P}\{\alpha = \infty\} \cdot \mathbb{P}\{\bar{\beta} = \infty\} = 0. \quad (2.53)$$

Ze względu na symetrię założeń, analogiczne równości zachodzą także dla odpowiednio zdefiniowanych funkcji $\bar{\alpha}$ i β .

Dowód. Przyjmijmy $S_0 = 0$. Dla $n \in \mathbb{N}$ i dowolnego $\omega \in \Omega$ niech $m = m(\omega) \leq n$ oznacza ostatni indeks, dla którego $S_m(\omega) = \min\{S_i(\omega); 0 \leq i \leq n\}$, tzn.

$$\omega \in A_m^n := \{0 = S_0, \dots, S_{m-1} \geq S_m < S_{m+1}, \dots, S_n\}.$$

Zachodzą równości: $\Omega = \bigcup_{m=0}^n A_m^n$ i $A_m^n = B_m \cap C_m^n$, gdzie

$$B_m = \{X_1 + \dots + X_m \leq 0, X_2 + \dots + X_m \leq 0, \dots, X_m \leq 0\} \in \sigma(X_1, \dots, X_m),$$

$$C_m^n = \{X_{m+1} > 0, X_{m+1} + X_{m+2} > 0, \dots, X_{m+1} + \dots + X_n > 0\} \in \sigma(X_{m+1}, \dots, X_n),$$

przy czym miary powyższych zbiorów zależą od rozkładów, a nie od opisu-
jących je zmiennych. Stąd

$$\mathbb{P}(B_m) = \mathbb{P}\{X_1 \leq 0, X_1 + X_2 \leq 0, \dots, X_1 + \dots + X_m \leq 0\} = \mathbb{P}\{\alpha > m\},$$

$$\mathbb{P}(C_m^n) = \mathbb{P}\{X_1 > 0, X_1 + X_2 > 0, \dots, X_1 + \dots + X_{n-m} > 0\} = \mathbb{P}\{\bar{\beta} > n - m\},$$

a zatem $\mathbb{P}(A_m^n) = \mathbb{P}\{\alpha > m\} \cdot \mathbb{P}\{\bar{\beta} > n - m\}$ dla $m \in \{0, \dots, n\}$, co kończy dowód (2.52). Kolejna równość jest konsekwencją oszacowania

$$(n+1) \cdot \mathbb{P}\{\alpha = \infty\} \cdot \mathbb{P}\{\bar{\beta} = \infty\} \leq 1 \quad \text{ dla } n \in \mathbb{N}. \quad \square$$

Dowód twierdzenia 2.2.5. Wykażemy w pierwszej kolejności, że w przypadku zerowej wartości oczekiwanej zmiennych X_n przynajmniej drugi z czynników w (2.53) jest równy 0. W tym celu założmy, że $\mathbb{P}\{\bar{\beta} = \infty\} = b > 0$, co

pociąga za sobą równość $\mathbb{P}\{\alpha = \infty\} = 0$. Funkcja α jest w tej sytuacji (pw. skończoną) zmienną losową, a wzór (2.15) i wynikające z (2.52) nierówności

$$\left(\sum_{m=0}^n \mathbb{P}\{\alpha > m\} \right) \cdot \mathbb{P}\{\bar{\beta} = \infty\} \leq 1 \quad \text{dla } n \in \mathbb{N},$$

implikują istnienie wartości oczekiwanej $\mathcal{E}(\alpha) = \int_0^\infty \mathbb{P}\{\alpha > t\} dt \leq \frac{1}{b}$. Prawie wszędzie skończona suma

$$S_\alpha := \sum_{n=1}^{\infty} 1_{\{\alpha \geq n\}} X_n, \quad \text{czyli} \\ S_\alpha: \omega \mapsto X_1(\omega) + \dots + X_{\alpha(\omega)}(\omega) > 0, \quad (2.54)$$

wskazuje pierwszą dodatnią wartość w ciągu $(S_n(\omega))_{n \in \mathbb{N}}$ i jest zmienną losową określoną dla $\omega \notin \Omega^-$. Zauważmy, że dla każdego $n \in \mathbb{N}$ zmienne X_n (i $|X_n|$) oraz $1_{\{\alpha \geq n\}} = 1 - 1_{\{\alpha < n\}}$ są niezależne. Przy założeniu całkowalności zmiennych X_n , skończona wartość całki

$$\int \sum_{n=1}^{\infty} 1_{\{\alpha \geq n\}} |X_n| d\mathbb{P} = \sum_{n=1}^{\infty} \int 1_{\{\alpha \geq n\}} d\mathbb{P} \int |X_n| d\mathbb{P} = \mathcal{E}(\alpha) \cdot \mathcal{E}(|X_1|)$$

oraz klasyczne twierdzenie Lebesgue'a o zmajoryzowanej zbieżności (por. twierdzenie 3.1.19) implikują całkowalność zmiennej S_α i równość

$$\mathcal{E}(S_\alpha) = \sum_{n=1}^{\infty} \int 1_{\{\alpha \geq n\}} X_n d\mathbb{P} = \mathcal{E}(\alpha) \cdot \mathcal{E}(X_1). \quad (2.55)$$

Ponieważ z (2.54) wynika nierówność $\mathcal{E}(S_\alpha) > 0$, przyjęte wcześniej założenie o mierze zbioru $\{\bar{\beta} = \infty\}$ i otrzymana wyżej tożsamość (2.55) wykluczają przypadek $\mu := \mathcal{E}(X_1) = 0$. Zgodnie z zapowiedzią wykazaliśmy tym samym prawdziwość implikacji $\mathcal{E}(X_1) = 0 \implies \mathbb{P}\{\bar{\beta} = \infty\} = 0$.

Założmy, że $\mu = 0$. Własność $\mathbb{P}(\bar{\beta} = \infty) = 0$ dla ciągu $(X_n)_{n \geq m}$ oznacza, że każdy ze zbiorów

$$C_m = \{\omega; \exists n \geq m \ X_m(\omega) + \dots + X_n(\omega) \leq 0\} \quad \text{dla } m \in \mathbb{N}$$

ma miarę 1, co przenosi się także na przekrój $C = \bigcap_{m \in \mathbb{N}} C_m$. Wynika stąd, że dla $\omega \in C$ ciąg $(S_n(\omega))_{n \in \mathbb{N}}$ zawiera podciąg nierosnący, a zatem $S_n(\omega) \not\rightarrow \infty$. Ponieważ przyjęte założenia spełnia także ciąg $(-X_n)_{n \in \mathbb{N}}$, otrzymana wyżej własność oznacza, że dla prawie wszystkich $\omega \in \Omega$ również ciąg $(-S_n(\omega))_{n \in \mathbb{N}}$ zawiera podciąg nierosnący – i $S_n(\omega) \not\rightarrow -\infty$. Po wykluczeniu przypadków (ii)-(iii) stwierdzamy, że znikanie wartości oczekiwanej μ implikuje ostatni z przypadków wyróżnionych w twierdzeniu 2.2.1.

Zachowanie ciągu $(S_n)_{n \in \mathbb{N}}$, tj. rozbieżność do ∞ lub $-\infty$ przy założeniu, że wartość oczekiwana $\mu = \mathcal{E}(X_1)$ istnieje i jest niezerowa (skończona lub nieskończona), jest prostym wnioskiem z udowodnionego niżej *Prawa wielkich liczb*, czyli twierdzeń 2.2.8 i 2.2.14. $\square$

Uwaga 2.2.4. Symulacje komputerowe wykonane w tym i następnym podrozdziale korzystają z generatora liczb losowych `Random` w standardowej bibliotece Java JDK – dla obiektu losującego utworzonego z parametrem `seed = 12345` (przykład 2.3.3). Rozkład Cauchy’ego powstaje z rozkładu jednostajnego (funkcja `nextDouble()`) przekształconego za pomocą funkcji kwantylowej $Q(t) = \operatorname{tg}(\pi(t - \frac{1}{2}))$ – zgodnie z regułą opisaną w podrozdziale 2.3.2, natomiast rozkład normalny symuluje funkcja `nextGaussian()` i jest on zgodny z algorytmem Marsaglii (podr. 2.4.4).

2.2.2 Mocne prawo wielkich liczb

Tytułowe dla bieżącego podrozdziału *Prawo*, czyli twierdzenie, formułujemy w tzw. wersji mocnej (ang. strong), w odniesieniu do zbieżności *prawie na pewno* badanego ciągu średnich zmiennych losowych. Samo twierdzenie nie jest skomplikowane, istotne *konstrukcje matematyczne* pojawiają się jednak w klasycznej sekwencji obserwacji i faktów prowadzących do dowodu oraz – w klasycznych empirycznych zastosowaniach PWL. Alternatywną i nieco silniejszą wersję PWL zawiera także twierdzenie 2.2.15.

Twierdzenie 2.2.8 (Mocne prawo wielkich liczb). *W dowolnej przestrzeni probabilistycznej Ω , dla dowolnego ciągu niezależnych zmiennych losowych $X_i, i \in \mathbb{N}$, o tym samym rozkładzie i wartości oczekiwanej μ zachodzi równość*

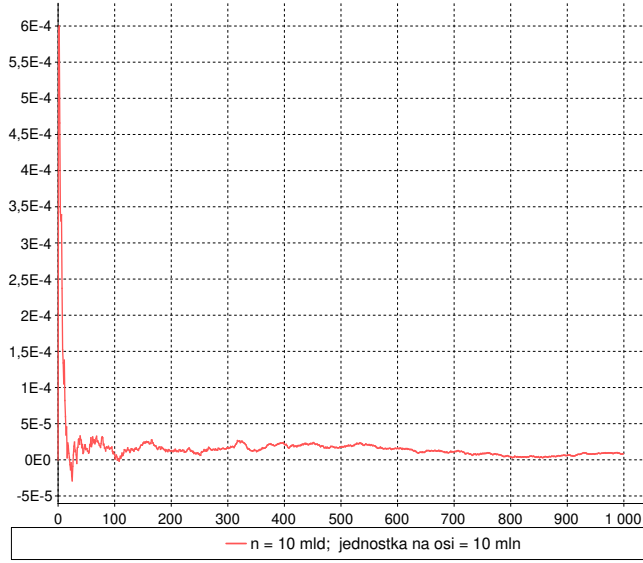
$$\mathbb{P}\left\{\omega \in \Omega; \lim_{n \rightarrow \infty} \frac{X_1(\omega) + \dots + X_n(\omega)}{n} = \mu\right\} = 1, \quad (2.56)$$

czyli ciąg średnich jest prawie na pewno zbieżny do wartości oczekiwanej.

Krok I. Zastępując zmienne X_i przez $X_i - \mu$ dla $i \in \mathbb{N}$, możemy założyć, że wspólną wartością oczekiwaną jest $\mu = 0$ – podobnie jak w Przykładzie 2.2.3. Sformułowany niżej lemat sprowadza badanie miary zbioru rozważanego w (2.56) do podzbioru tych $\omega \in \Omega$, w których szereg $\sum_n \frac{X_n(\omega)}{n}$ jest zbieżny – niekoniecznie bezwzględnie.

Lemat 2.2.9 (Kronecker). *Dla dowolnego ciągu liczbowego $(x_n)_{n \in \mathbb{N}}$ oraz rosnącego do ∞ ciągu liczb dodatnich $(c_n)_{n \in \mathbb{N}}$ prawdziwa jest implikacja*

$$\text{szereg } \sum_n \frac{x_n}{c_n} \text{ jest zbieżny} \implies \lim_{n \rightarrow \infty} \frac{1}{c_n} \sum_{k=1}^n x_k = 0. \quad (2.57)$$



Rysunek 2.3: PWL - symulacja dla rozkładu normalnego

Dowód. Oznaczmy $b_n = \sum_{k \leq n} \frac{x_k}{c_k}$, a więc $x_n = (b_n - b_{n-1})c_n$ dla $n \in \mathbb{N}$. Stąd

$$\frac{1}{c_n} \sum_{k \leq n} x_k = b_n - \frac{1}{c_n} \sum_{k < n} (c_{k+1} - c_k) b_k. \quad (2.58)$$

Niech $g = \lim_n b_n \in \mathbb{R}$ oznacza sumę szeregu $\sum_{n=1}^{\infty} \frac{x_n}{c_n}$. Dla dowolnego $\varepsilon > 0$ ustalmy indeks n_ε taki, że $|b_n - g| < \varepsilon$, jeśli tylko $n \geq n_\varepsilon$. Oszacowanie

$$\left| \frac{1}{c_n} \sum_{k < n} (c_{k+1} - c_k) b_k - g \right| \leq \frac{1}{c_n} \left(\sum_{k < n_\varepsilon} (c_{k+1} - c_k) |b_k| + c_{n_\varepsilon} |g| \right) + \frac{c_n - c_{n_\varepsilon}}{c_n} \varepsilon$$

dowodzi wówczas, że także drugi ze składników różnicy (2.58) dąży do g . $\square$

Krok II dotyczy wyjaśnienia pojęcia *zbieżności prawie wszędzie* lub – w ramach terminologii obowiązującej w probabilistyce – *zbieżności prawie na pewno*. Dla rozpoznania co to znaczy, że prawdopodobieństwo zdarzenia (czyli miara zbioru) złożonego z tych $\omega \in \Omega$, dla których ciąg wartości zmiennych losowych jest zbieżny, wynosi 1, udowodnimy

Lemat 2.2.10 (Zbieżność prawie na pewno). *Dowolny ciąg zmiennych losowych $Y_n: \Omega \rightarrow \mathbb{R}$, $n \in \mathbb{N}$, jest zbieżny prawie na pewno wtedy i tylko wtedy, gdy dla każdego $\varepsilon > 0$ zachodzi równość*

$$\lim_m \mathbb{P}\{\omega \in \Omega; \sup_{k > h \geq m} |Y_k(\omega) - Y_h(\omega)| > \varepsilon\} = 0. \quad (2.59)$$

Dowód. Korzystając z warunku Cauchy'ego charakteryzującego zbieżność ciągu $Y_n(\omega)$,

$$\forall_{n \in \mathbb{N}} \exists_{m \in \mathbb{N}} \forall_{k, h \geq m} |Y_k(\omega) - Y_h(\omega)| \leq \frac{1}{n},$$

stwierdzamy równość zbiorów

$$\{\omega; (Y_n(\omega))_{n \in \mathbb{N}} \text{ zbieżny}\} = \bigcap_{n \in \mathbb{N}} \bigcup_{m \in \mathbb{N}} \left\{ \sup_{k, h \geq m} |Y_k - Y_h| \leq \frac{1}{n} \right\}.$$

Miara przekroju jest równa 1 wtedy i tylko wtedy, gdy każdy z czynników ma miarę 1, co daje równoważny układ warunków

$$\lim_m \mathbb{P} \left\{ \sup_{k, h \geq m} |Y_k - Y_h| \leq \frac{1}{n} \right\} = \mathbb{P} \left(\bigcup_{m \in \mathbb{N}} \left\{ \sup_{k, h \geq m} |Y_k - Y_h| \leq \frac{1}{n} \right\} \right) = 1$$

dla każdego $n \in \mathbb{N}$, gdzie korzystamy także z ciągłości miary. Zastępując ułamek $\frac{1}{n}$ przez dowolną liczbę $\varepsilon > 0$ i przechodząc do dopełnień otrzymujemy jako równoważny warunek (2.59). $\square$

Krok III korzysta z faktu, iż w badanym szeregu składniki $\frac{X_n}{n}, n \in \mathbb{N}$, są niezależnymi zmiennymi losowymi o zerowej wartości oczekiwanej, ale wymaga *dotatkowego założenia* skończonej i szybko zbieżnej do 0 wariancji poszczególnych składników. Dla takiego szeregu prawdziwe jest

Twierdzenie 2.2.11 (Kryterium Kołmogorowa zbieżności). *Jeśli niezależne zmienne losowe $Z_n, n \in \mathbb{N}$, spełniają warunki*

$$\mathcal{E}(Z_n) = 0 \text{ dla } n \in \mathbb{N} \quad \text{oraz} \quad \sum_{n=1}^{\infty} \text{Var}(Z_n) < \infty, \quad (2.60)$$

to szereg $\sum_n Z_n$ jest zbieżny prawie na pewno.

Lemat 2.2.12 (Nierówność Kołmogorowa). *Niech $Z_1, \dots, Z_n$ będą niezależnymi zmiennymi losowymi o wartości oczekiwanej równej 0 i skończonej wariancji $\text{Var}(Z_k) < \infty$ dla wszystkich k . Dla $x > 0$ zachodzi nierówność*

$$\mathbb{P} \left\{ \max_{k \leq n} |Z_1 + \dots + Z_k| > x \right\} \leq \frac{1}{x^2} \sum_{k \leq n} \text{Var}(Z_k). \quad (2.61)$$

Dowód. Oznaczmy $S_k = Z_1 + \dots + Z_k$ dla $k \leq n$. Dla ustalonego $x > 0$ rozkład na zbiory wzajemnie rozłączne

$$\left\{ \max_{k \leq n} |S_k| > x \right\} = \bigcup_{k \leq n} A_k, \quad A_k = \left\{ \max_{j < k} |S_j| \leq x, |S_k| > x \right\} \text{ dla } k \leq n,$$

prowadzi do nierówności

$$\sum_{k \leq n} \text{Var}(Z_k) = \mathcal{E}(S_n^2) \geq \sum_{k \leq n} \int_{A_k} S_n^2 d\mathbb{P}.$$

Korzystając z niezależności zmiennych $S_k 1_{A_k}$, $S_n - S_k$ i oszacowania

$$\begin{aligned} \int_{A_k} S_n^2 d\mathbb{P} &= \int_{A_k} (S_k^2 + 2S_k(S_n - S_k) + (S_n - S_k)^2) d\mathbb{P} \\ &\geq \int_{A_k} S_k^2 d\mathbb{P} + 2 \int (S_k 1_{A_k})(S_n - S_k) d\mathbb{P} \\ &= \int_{A_k} S_k^2 d\mathbb{P} + 0 \geq x^2 \cdot \mathbb{P}(A_k) \quad \text{dla } k \leq n, \end{aligned}$$

otrzymujemy tezę $\frac{1}{x^2} \sum_{k \leq n} \text{Var}(Z_k) \geq \sum_{k \leq n} \mathbb{P}(A_k) = \mathbb{P}(\bigcup_{k \leq n} A_k)$. $\square$

Dowód twierdzenia 2.2.11. Teza o zbieżności prawie na pewno badanego szeregu jest prostym wnioskiem z lematu 2.2.10. W istocie, z nierówności Kołmogorowa (2.61) wynikają dla $n \geq m$ oszacowania

$$\begin{aligned} \mathbb{P}\left\{\max_{m \leq k \leq n} |Z_m + \dots + Z_k| > \varepsilon\right\} &\leq \frac{1}{\varepsilon^2} \sum_{k=m}^n \text{Var}(Z_k), \quad \text{a więc} \\ \mathbb{P}\left\{\sup_{k \geq m} |Z_m + \dots + Z_k| > \varepsilon\right\} &\leq \frac{1}{\varepsilon^2} \sum_{k=m}^{\infty} \text{Var}(Z_k), \end{aligned}$$

gdzie reszta szeregu zbieżnego dąży do 0 dla $m \rightarrow \infty$. $\square$

Dowód twierdzenia 2.2.8. Przedstawione wyżej kroki dowodu sprowadzają tezę do problemu zbieżności (prawie na pewno) szeregu o wyrazach będących niezależnymi zmiennymi losowymi postaci $\frac{X_n}{n}$, $n \in \mathbb{N}$. W przypadku, gdy rozważane zmienne X_n mają skończoną wariancję, teza wynika już z kryterium Kołmogorowa (2.60).

W przypadku, gdy wariancja zmiennych jest nieskończona, korzystamy z wniosku 2.1.11 i rozważamy niezależne zmienne losowe $\tilde{X}_n = X_n \cdot 1_{\{|X_n| \leq n\}}$ dla $n \in \mathbb{N}$.

Lemat 2.2.13. *Zachodzi nierówność*

$$\sum_{n \in \mathbb{N}} \frac{\text{Var}(\tilde{X}_n)}{n^2} \leq 4\mathcal{E}|X_1| < \infty.$$

Dowód. Z tożsamości (2.15) otrzymujemy dla $n \in \mathbb{N}$

$$\begin{aligned} \text{Var}(\tilde{X}_n) = \mathcal{E}(\tilde{X}_n^2) &= 2 \int_0^n t \mathbb{P}\{|\tilde{X}_n| > t\} dt \\ &\leq 2 \int_0^n t \mathbb{P}\{|X_n| > t\} dt = 2 \int_0^\infty 1_{(0,n)} t \mathbb{P}\{|X_1| > t\} dt, \end{aligned}$$

a zatem

$$\sum_{n \in \mathbb{N}} \frac{\text{Var}(\tilde{X}_n)}{n^2} \leq 2 \int_0^\infty t \left(\sum_{n>t} \frac{1}{n^2} \right) \mathbb{P}\{|X_1| > t\} dt. \quad (2.62)$$

Dla $t \in (0, 1)$ zachodzi nierówność $t \left(\sum_{n>t} \frac{1}{n^2} \right) \leq \sum_{n \in \mathbb{N}} \frac{1}{n^2} = \frac{\pi^2}{6}$, natomiast dla $t \geq 1$ sumowanie po $n \geq \lfloor t \rfloor + 1$ daje

$$t \sum_{n>t} \frac{1}{n^2} \leq t \int_{\lfloor t \rfloor}^\infty \frac{dx}{x^2} = \frac{t}{\lfloor t \rfloor} < \frac{\lfloor t \rfloor + 1}{\lfloor t \rfloor} \leq 2.$$

Stąd – oszacowania wielkości (2.62) przez $4 \int_0^\infty \mathbb{P}\{|X_1| > t\} dt = 4\mathcal{E}|X_1|$. $\square$

Wracając do dowodu prawa wielkich liczb, z kryterium Kołmogorowa wnioskujemy zbieżność szeregu $\sum_{n \in \mathbb{N}} \frac{\tilde{X}_n(\omega)}{n}$, a więc także – zgodnie z wnioskiem 2.1.11 – szeregu $\sum_{n \in \mathbb{N}} \frac{X_n(\omega)}{n}$, dla prawie wszystkich $\omega \in \Omega$. $\square$

Ćwiczenie 2.2.5. Korzystając z kolejnych etapów przedstawionego wyżej dowodu, wykazać, że dla każdego $a \in (\frac{1}{2}, 1]$ funkcja

$$(-\frac{1}{2}, \frac{1}{2}) \ni x \mapsto \sum_{n=1}^\infty \frac{X_n(x)}{n^a}$$

jest określona prawie wszędzie i nieograniczona na każdym przedziale otwartym. Można pokazać (patrz np. [8]), że dla $a = \frac{1}{2}$ badany losowy szereg naprzemienny jest prawie wszędzie (czyli prawie na pewno) rozbieżny.

Ćwiczenie 2.2.6. Skorzystać z PWL i wykazać, że jeśli zmienne losowe $X, X_n: \Omega \rightarrow \mathbb{R}, n \in \mathbb{N}$, są niezależne i mają ten sam rozkład, natomiast $A \subset \mathbb{R}$ jest dowolnym zbiorem borelowskim, to dla prawie wszystkich $\omega \in \Omega$ zachodzi równość

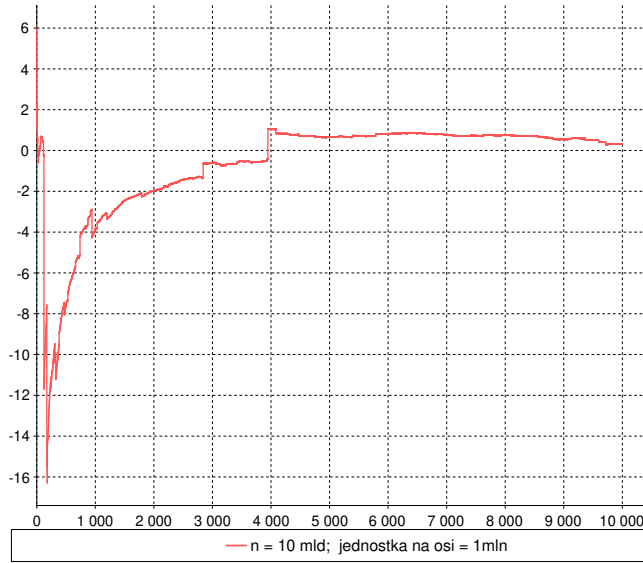
$$\mathbb{P}\{X \in A\} = \lim_{n \rightarrow \infty} \frac{\#\{k \leq n; X_k(\omega) \in A\}}{n},$$

gdzie symbol $\#$ oznacza liczbę elementów wskazanego zbioru.

Mocne prawo wielkich liczb nie ogranicza się do zmiennych o skończonej wartości oczekiwanej, ale „nie działa” – gdy wartość oczekiwana nie istnieje, co ilustruje rozważany wcześniej rozkład Cauchy’ego (rys. 2.4).

Twierdzenie 2.2.14 (Prawo wielkich liczb cd.). *W przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$, jeśli ciąg niezależnych zmiennych losowych $X_i, i \in \mathbb{N}$, o tym samym rozkładzie ma wartość oczekiwaną $+\infty$, wówczas*

$$\mathbb{P}\left\{\omega \in \Omega; \lim_n \frac{X_1(\omega) + \dots + X_n(\omega)}{n} = \infty\right\} = 1. \quad (2.63)$$



Rysunek 2.4: Średnia dla rozkładu Cauchy'ego

Dowód. Dla $m \in \mathbb{N}$ zmienne losowe $X_{im} = \min\{X_i, m\}$ są niezależne i mają ten sam rozkład, a przy tym $\mathcal{E}(X_{im}) = \mathcal{E}(X_{1m}) \leq m$. Zgodnie z twierdzeniem 2.2.8 zbiór

$$\Omega_m = \left\{ \omega \in \Omega; \lim_{n \rightarrow \infty} \frac{X_{1m}(\omega) + \dots + X_{nm}(\omega)}{n} = \mathcal{E}(X_{1m}) \right\}$$

ma miarę 1, przy czym

$$\liminf_{n \rightarrow \infty} \frac{X_1(\omega) + \dots + X_n(\omega)}{n} \geq \int X_{1m} d\mathbb{P}$$

dla $\omega \in \Omega_m$, $m \in \mathbb{N}$. Z twierdzenia 1.2.17(ii) – dla części dodatnich funkcji X_{1m} – wynika możliwość przejścia z granicą pod znak całki. Stąd ograniczenie granicy dolnej ciągu średnich (z dołu) przez całkę $\int X_1 d\mathbb{P}$, a więc

$$\lim_{n \rightarrow \infty} \frac{X_1(\omega) + \dots + X_n(\omega)}{n} = \infty$$

dla $\omega \in \bigcap_{m \in \mathbb{N}} \Omega_m$. Wskazany zbiór ma miarę 1. $\square$

Wiemy już, że ciąg sum $Y_n = \sum_{i \leq n} \frac{X_i}{i}$, $n \in \mathbb{N}$, utworzony ze zmiennych niezależnych o tym samym rozkładzie i wartości oczekiwanej $\mathcal{E}(X_n) = 0$ jest zbieżny prawie na pewno. Bardziej precyzyjne wykorzystanie zależności (2.57) prowadzi do nieco mocniejszego wyniku.

Twierdzenie 2.2.15 (Mocne PWL – wersja 2). *Niech $X_n: \Omega \rightarrow \mathbb{R}$, $n \in \mathbb{N}$, będzie dowolnym ciągiem niezależnych zmiennych losowych o tym samym rozkładzie i skończonej wartości oczekiwanej μ . Dla dowolnego $\varepsilon > 0$ zachodzi równość*

$$\lim_{m \rightarrow \infty} \mathbb{P}\left\{\omega; \sup_{n \in \mathbb{N}} \left| \frac{X_1 + \dots + X_n}{m+n}(\omega) - \mu \right| \leq \varepsilon\right\} = 1. \quad (2.64)$$

Dowód. Ograniczamy się do przypadku $\mu \equiv \mathcal{E}(X_n) = 0$. Własność (2.59) charakteryzująca zbieżność *prawie na pewno* ciągu sum $Y_n = \sum_{i \leq n} \frac{X_i}{i}$ oznacza, że dla każdej pary liczb $\varepsilon, \varepsilon' > 0$ istnieje $m_o \in \mathbb{N}$, takie że

$$\mathbb{P}\left\{\omega; \sup_{n > i \geq m_o} |Y_n(\omega) - Y_i(\omega)| \leq \varepsilon\right\} > 1 - \varepsilon'.$$

Niech A oznacza wyróżniony wyżej zbiór (zdarzenie) o mierze $\mathbb{P}(A) > 1 - \varepsilon'$. Z tożsamości (2.58) i wynikających z niej równości

$$\sum_{i \leq n} X_i = nY_n - \sum_{i < n} Y_i \quad \text{dla } n \in \mathbb{N}$$

wyprowadzamy, dla $n > m_o$, oszacowanie

$$\begin{aligned} |X_{m_o+1} + \dots + X_n| &= \left| (nY_n - \sum_{i < n} Y_i) - (m_o Y_{m_o} - \sum_{i < m_o} Y_i) \right| \\ &= \left| m_o(Y_n - Y_{m_o}) + \sum_{i=m_o}^{n-1} (Y_n - Y_i) \right| \\ &\leq n \max \{|Y_n - Y_i|; m_o \leq i < n\}. \end{aligned}$$

Jeśli zatem $\omega \in A$, to dla każdego $n > m_o$

$$\frac{1}{n} |X_{m_o+1}(\omega) + \dots + X_n(\omega)| \leq \varepsilon,$$

co oznacza inkluzję $A \subset B := \{\forall n > m_o \frac{1}{n} |X_{m_o+1} + \dots + X_n| \leq \varepsilon\}$ i pociąga za sobą nierówność $\mathbb{P}(B) > 1 - \varepsilon'$. Ze względu na ten sam rozkład i niezależność zmiennych X_i , zdarzenie $B' = \{\forall n > m_o \frac{1}{n} |X_1 + \dots + X_{n-m_o}| \leq \varepsilon\}$ ma takie samo prawdopodobieństwo, a zatem dla każdego $m \geq m_o$

$$\mathbb{P}\left\{\forall n \in \mathbb{N} \frac{1}{n+m} |X_1 + \dots + X_n| \leq \varepsilon\right\} \geq \mathbb{P}(A) > 1 - \varepsilon',$$

co było do okazania. □

Uwaga 2.2.7. Z nierówności (2.64) wynika oszacowanie

$$\mathbb{P}\left\{\omega; \limsup_n \left| \frac{X_1(\omega) + \dots + X_n(\omega)}{n} - \mu \right| \leq \varepsilon\right\} > 1 - \varepsilon'$$

dla dowolnych $\varepsilon, \varepsilon' > 0$, a to jest klasyczna wersja PWL.

2.2.3 Podejście empiryczne

Precyzując i formalizując terminologię, przyjmujemy

Definicja 2.2.16. Dla dowolnej zmiennej losowej $X: \Omega \rightarrow \mathbb{R}$ w przestrzeni $(\Omega, \mathfrak{F}, \mathbb{P})$ zbiór $\mathbb{X} = X(\Omega) \subset \mathbb{R}$ stanowi przestrzeń próbek.

Dla dowolnego $n \in \mathbb{N}$ każdy ciąg liczb $(x_1, \dots, x_n) \in \mathbb{X}^n$ nazywamy n -elementową próbką losową zmiennej X .

Każda n -elementowa próbka zmiennej X jest obrazem ciągu zdarzeń elementarnych $\omega = (\omega_1, \dots, \omega_n) \in \Omega^n$, który możemy nazywać *abstrakcyjną* próbką losową. Dopuszczamy możliwość powtórzeń w próbkach, a interpretację „losowości” poszczególnych próbek pozostawiamy do dalszej oceny w kontekście indukowanej borelowskiej miary probabilistycznej $\mathbb{P}_X^n$ w $\mathbb{R}^n$ lub – ogólniej – miary produktowej $\mathbb{P}^n$ w Ω^n .

Definicja 2.2.17. Dla dowolnej zmiennej losowej X w przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$ każdy nieskończony ciąg zdarzeń $\omega = (\omega_n)_{n \in \mathbb{N}} \in \Omega^{\mathbb{N}}$ wyznacza próbę losową $X^{\mathbb{N}}(\omega) := (X(\omega_n))_{n \in \mathbb{N}} \in \mathbb{X}^{\mathbb{N}} \subset \mathbb{R}^{\mathbb{N}}$.

Nazwę „próba losowa” rezerwujemy dla nieskończonych ciągów danych oraz dla ciągów zmiennych losowych. W szczególności, dotyczy to ciągów zmiennych postaci $X_k = X \circ \pi_k: (\omega_i)_{i \leq n} \mapsto X(\omega_k)$, $k \leq n$, określonych na przestrzeni $\Omega^{\mathbb{N}}$ abstrakcyjnych prób losowych, niezależnych i mających rozkład $\mathbb{P}_X$.

W *modelowej* przestrzeni borelowskiej $(\mathbb{R}^{\mathbb{N}}, \mathfrak{B}_{\mathbb{R}}^{\mathbb{N}}, \mathbb{P}_X^{\mathbb{N}})$ projekcje na poszczególne współrzędne $\pi_k: \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}$, $k \in \mathbb{N}$, traktowane jako zmienne losowe tworzą *standardową próbę losową*, uniwersalną dla dowolnej zmiennej losowej X .

Uwaga 2.2.8. Możliwość ograniczenia prób losowych do podprzestrzeni $\mathbb{X}^{\mathbb{N}} \subset \mathbb{R}^{\mathbb{N}}$ ma znaczenie raczej teoretyczne w sytuacji, gdy rozważane w ramach *Statystyki* zmienne losowe są związane z niekoniecznie dokładnym pomiarem wybranych cech badanych obiektów. Korzystanie z przybliżonych wartości liczbowych i wnioskowanie dotyczące rozkładów prawdopodobieństwa umożliwia regularność miar borelowskich w $\mathbb{R}^{\mathbb{N}}$. W szczególności, pierwsza z własności (1.26) pozwala wyznaczać miarę (prawdopodobieństwo) dowolnego zdarzenia A w oparciu o miary dostatecznie bliskich zbiorów otwartych $U \supset A$.

Definicja 2.2.18. Przestrzenią prób losowych zmiennej X nazywamy borelowską przestrzeń probabilistyczną $(\mathbb{X}^{\mathbb{N}}, \mathfrak{B}_{\mathbb{X}^{\mathbb{N}}}, \mathbb{P}_{\mathbb{X}})$, w której miarę $\mathbb{P}_{\mathbb{X}}$ charakteryzuje

Stwierdzenie 2.2.19. *Dla dowolnej zmiennej losowej $X: \Omega \rightarrow \mathbb{R}$ w podprzestrzeni $\mathbb{X}^{\mathbb{N}} \subset \mathbb{R}^{\mathbb{N}}$ istnieje dokładnie jedna miara probabilistyczną $\mathbb{P}_X$ indukującą w $\mathbb{R}^{\mathbb{N}}$ miarę produktową $\mathbb{P}_X^{\mathbb{N}}$, a więc taka, że*

$$\mathbb{P}_X^{\mathbb{N}}(A) = \mathbb{P}_X(A \cap \mathbb{X}^{\mathbb{N}}) \quad \text{dla } A \in \mathfrak{B}_{\mathbb{R}}^{\mathbb{N}}. \quad \square$$

Ćwiczenie 2.2.9. W przestrzeni $\mathbb{R}^{\mathbb{N}}$ miara $\mathbb{P}_X^{\mathbb{N}}$ jest indukowana z $\Omega^{\mathbb{N}}$, a więc dla każdego zdarzenia $A \in \mathfrak{B}_{\mathbb{R}}^{\mathbb{N}}$, jeśli $\mathbb{P}_X^{\mathbb{N}}(A) > 0$, to $A \cap \mathbb{X}^{\mathbb{N}} \neq \emptyset$. Z drugiej strony, jest to miara regularna, a zatem liczba $\mathbb{P}_X^{\mathbb{N}}(A)$ jest kresem górnym miar $\mathbb{P}_X^{\mathbb{N}}(K)$, gdzie $K \subset A$ przebiega rodzinę zwartych podzbiorów. Korzystając z faktu, iż dla każdego $n \in \mathbb{N}$ rzut $K_n = \pi_{1..n}(K) \subset \mathbb{R}^n$ jest zbiorem zwartym, wykazać równość $K = \bigcap_{n \in \mathbb{N}} \pi_{1..n}^{-1} K_n$ oraz

$$\begin{aligned} \mathbb{P}_X^{\mathbb{N}}(K) &= \inf_n \mathbb{P}_X^n(\pi_{1..n} K) \\ &= \inf_n \mathbb{P}^n\{(\omega_i) \in \Omega^n; \exists x \in \mathbb{R}^{\mathbb{N}} (X(\omega_1), \dots, X(\omega_n), x_1, x_2, \dots) \in K\}. \end{aligned} \quad (2.65)$$

Powyższa równość urealnia przybliżanie miary $\mathbb{P}_X^{\mathbb{N}}$ i sprowadza je do operacji na skończonych próbkach losowych – a tylko takie występują w świecie rzeczywistym.

Podstawą powszechnie uznanego formalizmu upraszczającego pojęcie próby losowej jest

Definicja 2.2.20. *Dla dowolnej borelowskiej miary probabilistycznej μ w $\mathbb{R}$ próbą losową o rozkładzie μ nazywamy każdy niezależny ciąg zmiennych losowych $X_1, X_2, \dots: \Omega \rightarrow \mathbb{R}$ o tym samym rozkładzie $\mathbb{P}_{X_i} = \mu$.*

Przyjęte w definicji uproszczenie sprawia, że zmienne X_i mają ten sam rozkład, ale równość miar $\mathbb{P}_{X_i}$ nie zapewnia równości obrazów $X_i(\Omega)$. Pojęcie „przestrzeni próbek” traci tym samym realny sens – nie jest już zbiorem wszystkich możliwych pomiarów, wyników powtarzalnego *doświadczenia* X , którego rozkład μ badamy. Przestrzeń Ω jest tu jedynie tworem formalnym, niezbędnym dla zdefiniowania niezależności ciągu zmiennych.

W odniesieniu do prób losowych z PWL wynika

Wniosek 2.2.21. *Niech $X: \Omega \rightarrow \mathbb{R}$ będzie dowolną zmienną losową. Dla prawie wszystkich prób losowych $(X(\omega_n))_{n \in \mathbb{N}}$ zachodzą równości*

(i) $\lim_{n \rightarrow \infty} \frac{1}{n} (X(\omega_1) + \dots + X(\omega_n)) = \mathcal{E}(X)$, *jeśli całka $\mathcal{E}(X)$ istnieje (skończona lub nieskończona), oraz*

(ii) $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i \leq n} (X(\omega_i) - \bar{X}_n(\omega))^2 = \text{Var}(X)$, *jeśli zmienna X jest całkowalna, a $\bar{X}_n(\omega) = \frac{1}{n} (X(\omega_1) + \dots + X(\omega_n))$ jest występującą w (i) estymacją, czyli zmienną będącą przybliżeniem wartości oczekiwanej.*

Dowód. Jeśli zmienna X jest całkowalna i $\mu = \mathcal{E}(X)$, z twierdzenia 2.2.8 zastosowanego dla standardowej próby losowej $(\pi_n)_{n \in \mathbb{N}}$ wynika równość

$$\mathbb{P}_X^{\mathbb{N}}\{(x_n)_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}; \lim_{n \rightarrow \infty} \frac{x_1 + \dots + x_n}{n} = \mu\} = 1.$$

Własność (i) otrzymujemy po ograniczeniu ciągów do przestrzeni prób $X(\Omega)^{\mathbb{N}}$. Przypadek $\mathcal{E}(X) = \infty$ wynika z twierdzenia 2.2.14.

Dla dowodu własności (ii) korzystamy z tożsamości

$$\begin{aligned} \frac{1}{n} \sum_{i \leq n} \left(x_i - \frac{1}{n} \sum_{j \leq n} x_j\right)^2 &= \frac{1}{n} \sum_{i \leq n} \left((x_i - \mu) - \frac{1}{n} \sum_{j \leq n} (x_j - \mu)\right)^2 \\ &= \frac{1}{n} \sum_{i \leq n} (x_i - \mu)^2 - \left(\frac{1}{n} \sum_{i \leq n} x_i - \mu\right)^2, \end{aligned} \quad (2.66)$$

dla $(x_i)_{i \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$, gdzie μ jest skończoną wartością oczekiwaną zmiennej X .

Ponieważ zmienne $(\pi_i - \mu)^2: \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}, i \in \mathbb{N}$, pozostają niezależne i mają ten sam rozkład co $(X - \mu)^2$, pierwszy ze składników zbiega (prawie na pewno) do wielkości $\mathcal{E}(X - \mu)^2 = \text{Var}(X)$. Drugi ze składników dąży do 0. $\square$

Przybliżającą wariancję $\text{Var}(X)$ funkcja $Z_n = \frac{1}{n} \sum_{i \leq n} (X_i - \bar{X}_n)^2$, gdzie $X_i \equiv X \circ \pi_i$, jest zmienną losową na przestrzeni Ω^n – dla każdego $n \in \mathbb{N}$. Naturalną drogę poprawienia dokładności przybliżenia wskazuje

Lemat 2.2.22. *Jeśli zmienna losowa $X: \Omega \rightarrow \mathbb{R}$ ma skończoną wariancję, wówczas dla każdego $n \in \mathbb{N}$ zachodzi równość*

$$\mathcal{E}\left(\frac{1}{n} \sum_{i \leq n} (X_i - \bar{X}_n)^2\right) = \left(1 - \frac{1}{n}\right) \text{Var}(X).$$

Dowód. Korzystając z (2.66) dla rozważanej zmiennej Z_n , otrzymujemy

$$\begin{aligned} \mathcal{E}Z_n &= \frac{1}{n} \sum_{i \leq n} \mathcal{E}(X_i - \mu)^2 - \frac{1}{n^2} \left(\sum_{i=j \leq n} + \sum_{i \neq j} \right) \mathcal{E}(X_i - \mu)(X_j - \mu) \\ &= \text{Var}(X) - \frac{1}{n} \text{Var}(X), \end{aligned}$$

gdzie znikanie trzeciej sumy wynika z niezależności czynników dla $i \neq j$. $\square$

Ponieważ zmienne $\frac{n}{n-1}Z_n$ mają dla $n > 1$ wartość oczekiwaną dokładnie równą wariancji $\text{Var}(X)$, możemy ustalić n , a następnie – rezygnując z przejścia granicznego – do przybliżania wariancji wykorzystać średnią wartość otrzymanego w ten sposób poprawionego, a więc *nieobciążonego* estymatora.

Definicja 2.2.23. Wariancją empiryczną zmiennej losowej $X: \Omega \rightarrow \mathbb{R}$ nazywamy określony dla dowolnej próby losowej ciąg $(\text{Var}_n X)_{n \in \mathbb{N}}$ wariancji n -elementowych próbek

$$\text{Var}_n X = \frac{1}{n-1} \sum_{i \leq n} (X_i - \bar{X}_n)^2: \Omega^n \rightarrow \mathbb{R}. \quad (2.67)$$

W każdej przestrzeni borelowskiej $\mathbb{R}^n$ z miarą $\mathbb{P}_X^n$ wariancją empiryczną jest funkcja

$$\text{Var}_n: (x_i)_{i \leq n} \mapsto \frac{1}{n-1} \sum_{i \leq n} \left(x_i - \frac{1}{n} \sum_{j \leq n} x_j \right)^2, \quad \text{dla } n \geq 2. \quad (2.68)$$

Uwaga 2.2.10. Nieco więcej przygotowań wymaga analogiczne twierdzenie o estymacji kowariancji dowolnych dwu zmiennych losowych X, Y w tej samej przestrzeni probabilistycznej Ω . Jeśli $\mu_X, \mu_Y \in \mathbb{R}$ oznaczają wartości oczekiwane, to dla ciągu niezależnych zmiennych losowych

$$(X_n - \mu_X)(Y_n - \mu_Y): \Omega^{\mathbb{N}} \rightarrow \mathbb{R}, \quad n \in \mathbb{N},$$

prawo wielkich liczb oznacza równość

$$\begin{aligned} 1 &= \mathbb{P}^{\mathbb{N}} \left\{ \omega \in \Omega^{\mathbb{N}}; \frac{1}{n} \sum_{i \leq n} (X(\omega_i) - \mu_X)(Y(\omega_i) - \mu_Y) \rightarrow \text{Cov}(X, Y) \right\} \\ &= \mathbb{P}_{(X, Y)}^{\mathbb{N}} \left\{ ((x_i, y_i))_{i \in \mathbb{N}}; \frac{1}{n} \sum_{i \leq n} (x_i - \mu_X)(y_i - \mu_Y) \rightarrow \text{Cov}(X, Y) \right\}, \end{aligned}$$

gdzie druga z miar jest przeliczalnym produktem (potęgą) rozkładu wektora (X, Y) – w przestrzeni borelowskiej $(\mathbb{R}^2)^{\mathbb{N}}$. Tożsamość

$$\begin{aligned} &\frac{1}{n} \sum_{i \leq n} \left(x_i - \frac{1}{n} \sum_{j \leq n} x_j \right) \left(y_i - \frac{1}{n} \sum_{j \leq n} y_j \right) \\ &= \frac{1}{n} \sum_{i \leq n} (x_i - \mu_X)(y_i - \mu_Y) - \left(\frac{1}{n} \sum_{i \leq n} x_i - \mu_X \right) \left(\frac{1}{n} \sum_{i \leq n} y_i - \mu_Y \right), \end{aligned} \quad (2.69)$$

pozwała wywnioskować, że średnie (2.69) dążą do kowariancji $\text{Cov}(X, Y)$ dla pw. *wektorowych* prób losowych $((x_n, y_n))_{n \in \mathbb{N}} = ((X, Y)(\omega_n))_{n \in \mathbb{N}}$. Stosując to rozumowanie do współrzędnych wektorów losowych oraz uwzględniając korektę wartości oczekiwanej (ćwiczenie!), przymujemy

Definicja 2.2.24. Kowariancją empiryczną *wektorowej zmiennej losowej* X nazywamy ciąg macierzy

$$\text{Cov}_n X(\omega) = \frac{1}{n-1} \sum_{i \leq n} \left(X(\omega_i) - \frac{1}{n} \sum_{j \leq n} X(\omega_j) \right) \left(X(\omega_i) - \frac{1}{n} \sum_{j \leq n} X(\omega_j) \right)^t$$

określony dla dowolnej próby losowej $(X(\omega_n))_{n \in \mathbb{N}}$.

Wniosek 2.2.25. Niech $X = (X_1, \dots, X_p)^t: \Omega \rightarrow \mathbb{R}^p$, $p \in \mathbb{N}$, będzie dowolnym skończonym wymiarowym wektorem losowym zmiennych o skończonej wariancji. W przestrzeni macierzy $\mathbb{R}^{p \times p}$ zachodzi równość

$$\lim_{n \rightarrow \infty} \text{Cov}_n X(\omega) = \text{Cov}(X)$$

dla prawie wszystkich $\omega = (\omega_i)_{i \in \mathbb{N}} \in \Omega^{\mathbb{N}}$ lub, równoważnie, dla p.w. wektorowych prób losowych $(X(\omega_i))_{i \in \mathbb{N}} \in (\mathbb{R}^p)^{\mathbb{N}}$ – względem miary produktowej $\mathbb{P}_X^{\mathbb{N}}$. $\square$

2.2.4 Twierdzenie Gliwienki–Cantelliego

Nieco abstrakcyjne twierdzenie 2.1.1 pozwala na przybliżone wyznaczania rozkładu prawdopodobieństwa dowolnej zmiennej losowej $X: \Omega \rightarrow \mathbb{R}$ z wykorzystaniem formuły $\mathbb{P}_X(A) = \lim_{n \rightarrow \infty} \#\{k \leq n; X(\omega_k) \in A\}/n$ prawdziwej dla prawie wszystkich prób losowych $(X(\omega_n))_{n \in \mathbb{N}}$ w $\mathbb{R}^{\mathbb{N}}$. Dokładniej – dla każdego zbioru $A \in \mathfrak{B}_{\mathbb{R}}$ każda losowo wybrana próba pozwala na wyznaczenie z dowolną dokładnością prawdopodobieństwa $\mathbb{P}_X(A)$. Zauważmy jednak, że – o ile zmienna losowa X nie jest dyskretna – żaden ustalony, najbardziej losowo wybrany ciąg $(X(\omega_n))_{n \in \mathbb{N}}$ nie jest uniwersalny, a podany we wniosku wzór nie zachodzi np. dla zbioru $A = \{X(\omega_n); n \in \mathbb{N}\}$, dla którego podejście częstościowe daje wartość 1. Dotyczy to także każdego zbioru otwartego $U \supset A$ o mierze $\mathbb{P}_X(U) < 1$. Ten pozorny paradoks wynika z faktu, iż losowość wyboru próby zależy od kontekstu i trudno uznać za losowy wybór ciągu $X^{\mathbb{N}}(\omega)$, którego wszystkie wyrazy należą do znanego wcześniej zdarzenia A o prawdopodobieństwie $\mathbb{P}\{X \in A\} < 1$. Z zawartego w sformułowanym niżej twierdzeniu 2.2.29 dobrego, empirycznego przybliżenia dystrybucyjnego badanego rozkładu, wynika jednak

Wniosek 2.2.26. Niech X będzie dowolną zmienną losową. Prawie na pewno, czyli dla prawie wszystkich prób losowych $(X(\omega_n))_{n \in \mathbb{N}}$, zachodzi równość

$$\mathbb{P}_X(J) = \lim_{n \rightarrow \infty} \frac{\#\{k \leq n; X(\omega_k) \in J\}}{n} \quad (2.70)$$

dla każdego przedziału $J \subset \mathbb{R}$ jednocześnie. $\square$

Definicja 2.2.27. Dla dowolnego ciągu $X_n: \Omega \rightarrow \mathbb{R}, n \in \mathbb{N}$, niezależnych zmiennych losowych o tym samym rozkładzie, rodzinę zmiennych losowych

$$F_n(x, \cdot) = \frac{1}{n} \sum_{k \leq n} 1_{(-\infty, x]} \circ X_k \quad \text{dla } x \in \mathbb{R}, n \in \mathbb{N}, \quad (2.71)$$

nazywamy empiryczną dystrybucyjną rozkładu prawdopodobieństwa dowolnej ze zmiennych X_n .

Każda z funkcji $F_n^\omega := F_n(\cdot, \omega)$ jest dystrybuantą miary dyskretnej μ_n^ω na prostej, skupionej na skończonym zbiorze $\{X_k(\omega); k \leq n\} \subset \mathbb{R}$, dla $n \in \mathbb{N}$ i $\omega \in \Omega$. Zgodnie ze stwierdzeniem 1.1.42, miary μ_n^ω są indukowane z miary Lebesgue'a $\ell|_{(0,1)}$ przez funkcje kwantylowe $Q_n^\omega := Q^{F_n(\cdot, \omega)}$ takie, że

$$Q_n^\omega(u) = \inf\{x \in \mathbb{R}; u \leq F_n^\omega(x)\} \quad (2.72)$$

$$= X_{k_i}(\omega) \quad \text{dla } \frac{i-1}{n} < u \leq \frac{i}{n}, \quad (2.73)$$

jeśli $X_{k_1}(\omega) \leq \dots \leq X_{k_n}(\omega)$ jest uporządkowanym ciągiem wartości $X_i(\omega)$ dla $i \leq n, \omega \in \Omega$.

Ćwiczenie 2.2.11. Wyprowadzić równość (2.73) z własności (2.72).

Definicja 2.2.28. Rodzinę funkcji Q_n^ω , dla $n \in \mathbb{N}$ i $\omega \in \Omega$, będziemy nazywać empiryczną funkcją kwantylową wspólnego rozkładu zmiennych X_n .

Twierdzenie 2.2.29 (Gliwienki–Cantelliego). Niech $X_n, n \in \mathbb{N}$, będzie dowolnym ciągiem niezależnych zmiennych losowych o tym samym rozkładzie. Dla prawie wszystkich $\omega \in \Omega$ (czyli – prawie na pewno)

(i) dystrybuanty F_n^ω dążą jednostajnie do funkcji $F: \mathbb{R} \rightarrow [0, 1]$ będącej wspólną dystrybuantą zmiennych X_n , oraz

(ii) funkcje kwantylowe Q_n^ω dążą do funkcji Q^F w sensie

$$\liminf_n Q_n^\omega(u) = Q^F(u-) \quad \text{i} \quad \limsup_n Q_n^\omega(u) = Q^F(u+) \quad \text{dla } u \in (0, 1). \quad (2.74)$$

Równości (2.74) obejmują w szczególności zawartą w twierdzeniu 2.1.16 zbieżność $\lim_{n \rightarrow \infty} Q_n^\omega(u) = Q^F(u)$ w każdym punkcie ciągłości granicznej funkcji kwantylowej, czyli dla $u \in (0, 1)$ takich, że przedział $F^{-1}(u)$ redukuje się do punktu lub jest zbiorem pustym.

W kontekście rozważanej wcześniej słabej zbieżności miar i ich dystrybuant jednostajna zbieżność ciągu dystrybuant empirycznych uzyskana w twierdzeniu Gliwienki–Cantelliego jest własnością raczej nietypową, niezależną od ewentualnej nieciągłości wspólnej dystrybuanty rozkładu.

Uwaga 2.2.12. Zbiór $\{\omega \in \Omega; F_n^\omega \rightrightarrows F\}$ jest zdarzeniem (zbiorem mierzalnym), gdyż badana zbieżność jednostajna jest równoważna zbieżności do zera ciągu zmiennych losowych,

$$\sup_{x \in \mathbb{Q}} |F_n^\omega(x) - F(x)| = \sup_{x \in \mathbb{R}} |F_n^\omega(x) - F(x)| \longrightarrow 0 \quad \text{dla } n \rightarrow \infty,$$

czego sprawdzenie pozostawiamy czytelnikowi.

Dowód twierdzenia 2.2.29. (i) Dla każdego $x \in \mathbb{R}$ z twierdzenia 2.1.1 wynika zbieżność prawie wszędzie

$$F_n(x, \cdot) \longrightarrow F(x), \quad \text{oraz}$$

$$F_n(x-, \cdot) = \frac{1}{n} \sum_{k \leq n} 1_{(-\infty, x)} \circ X_k \longrightarrow F(x-)$$

dla $n \rightarrow \infty$, gdzie wykorzystana notacja odnosi się do lewostronnych granic rozważanych dystrybuant. Aby skonstruować jeden wspólny zbiór $\Omega_0 \subset \Omega$ o mierze $\mathbb{P}(\Omega_0) = 1$, na którym zbieżność $F_n(x, \cdot) \rightarrow F(x)$ zachodzi jednostajnie dla każdego $x \in \mathbb{R}$, wyróżniamy kwantyle

$$x_{mi} = Q^F\left(\frac{i}{m}\right) = \inf \left\{ x; F(x) \geq \frac{i}{m} \right\} \quad \text{dla } 0 < i < m, m \in \mathbb{N}.$$

Wskazane punkty oraz $x_{m0} = -\infty$ i $x_{mm} = \infty$ wyznaczają podział prostej $\mathbb{R}$ w ten sposób, że

$$F(x_{m,i-1}) \leq F(x_{mi}-) \leq F(x_{m,i-1}) + \frac{1}{m}, \quad \text{gdy } x_{m,i-1} < x_{mi}, \quad (2.75)$$

dla $i \leq m, m \in \mathbb{N}$. Istotnie, negacja nieoczywistej drugiej nierówności oznacza istnienie liczby $x \in (x_{m,i-1}, x_{mi})$, dla której $F(x) > F(x_{m,i-1}) + \frac{1}{m} \geq \frac{i}{m}$, co jest sprzeczne z nierównością $x < x_{mi}$.

Za Ω_0 przyjmujemy przekrój przeliczalnej rodziny zdarzeń o prawdopodobieństwie 1 złożony z tych $\omega \in \Omega$, dla których ma miejsce zbieżność

$$F_n(x_{mi}, \cdot) \longrightarrow F(x_{mi}) \quad \text{i} \quad F_n(x_{mi}-, \cdot) \longrightarrow F(x_{mi}-)$$

jednocześnie dla wszystkich wyróżnionych punktów; oczywiście $\mathbb{P}(\Omega_0) = 1$. Zauważmy, że dla $\omega \in \Omega_0$ ciąg funkcji $F_n^\omega, n \in \mathbb{N}$ jest jednostajnie zbieżny do F . Istotnie, dla $\omega \in \Omega_0$ i dowolnego $m \in \mathbb{N}$ istnieje indeks N_m^ω taki, że

$$|F_n(x_{mi}, \omega) - F(x_{mi})| < \frac{1}{m} \quad \text{i} \quad |F_n(x_{mi}-, \omega) - F(x_{mi}-)| < \frac{1}{m} \quad \text{dla } i \leq m,$$

jeśli tylko $n \geq N_m^\omega$. Dla $x \in [x_{m,i-1}, x_{mi})$ z (2.75) i oszacowania

$$F_n(x, \omega) \leq F_n(x_{mi}-, \omega) < F(x_{mi}-) + \frac{1}{m} \leq F(x_{m,i-1}) + \frac{2}{m} \leq F(x) + \frac{2}{m}, \quad \text{oraz}$$

$$F_n(x, \omega) \geq F_n(x_{m,i-1}, \omega) > F(x_{m,i-1}) - \frac{1}{m} \geq F(x_{mi}-) - \frac{2}{m} \geq F(x) - \frac{2}{m}$$

wynika nierówność $|F_n(x, \omega) - F(x)| < \frac{2}{m}$, a zatem

$$\sup_{x \in \mathbb{R}} |F_n(x, \omega) - F(x)| \leq \frac{2}{m} \quad \text{dla } n \geq N_m^\omega,$$

co dowodzi jednostajnej zbieżności rozważanego ciągu funkcyjnego.

(ii) Dla dowodu specyficznej, opisanej w (2.74) zbieżności dla empirycznych funkcji kwantylowych, ograniczamy się do punktów nieciągłości funkcji Q^F tworzących co najwyżej przeliczalny zbiór

$$P = \{p \in (0, 1); \text{int } F^{-1}\{p\} \neq \emptyset\}.$$

Dla $p \in P$ i liczb $a_p = Q^F(p-) = Q^F(p)$ i $b_p = Q^F(p+) > a_p$ zachodzą równości $F(a_p) = p = F(b_p-)$, a zatem każde zdarzenie $\{X_n \in (a_p, b_p)\}$, $n \in \mathbb{N}$, ma prawdopodobieństwo 0. Oznacza to, że zbiór

$$\Omega'_0 = \Omega_0 \setminus \bigcup_{n \in \mathbb{N}} \bigcup_{p \in P} X_n^{-1}(a_p, b_p)$$

ma miarę $\mathbb{P}(\Omega'_0) = 1$. Ustalmy dowolny punkt nieciągłości $p \in P$ i oznaczmy $(a, b) = (a_p, b_p)$, $q = 1 - p$. Dla $\omega \in \Omega'_0$ i dowolnego $n \in \mathbb{N}$ w próbie $X_1(\omega), \dots, X_n(\omega)$ nie ma liczb z przedziału (a, b) , a po uporządkowaniu zachodzą nierówności

$$X_{k_1}(\omega) \leq \dots \leq X_{k_s}(\omega) \leq a < b \leq X_{k_{s+1}}(\omega) \leq \dots \leq X_{k_n}(\omega) \quad (2.76)$$

dla jednoznacznie określonego indeksu $s \in \{0, \dots, n\}$. Ponieważ zmienne losowe

$$Z_n = q \cdot 1_{\{X_n \leq a\}} - p \cdot 1_{\{X_n > a\}} \quad \text{dla } n \in \mathbb{N}$$

są niezależne, mają ten sam rozkład i wartość oczekiwaną $\mathcal{E}(Z_1) = 0$, z twierdzenia 2.2.5 o błędzeniu losowym II wynika, że dla prawie wszystkich $\omega \in \Omega'_0$ w ciągu sum

$$S_n^Z(\omega) := (Z_1 + \dots + Z_n)(\omega) = s \cdot q - (n - s) \cdot p = s - np, \quad n \in \mathbb{N},$$

zarówno dodatnie jak i ujemne wartości występują nieskończenie wiele razy.

Zgodnie z definicją, wartość funkcji kwantylowej $Q_n^\omega(p)$ zależy od numeru i_n przedziału $(\frac{i-1}{n}, \frac{i}{n}]$ zawierającego p , czyli takiego, że $i_n - 1 < pn \leq i_n$. Sumy dodatnie $S_n^Z(\omega) > 0$ oznaczają zatem $s > pn > i_n - 1$, przy czym

$$i_n \leq s \implies k_{i_n} \leq k_s \implies Q_n^\omega(p) \leq X_{k_s}(\omega) \leq a.$$

co w granicy prowadzi do nierówności $\liminf_n Q_n^\omega(p) \leq a$. Z drugiej strony, sumy ujemne oznaczają nierówności $s < pn \leq i_n$, a implikacje

$$s + 1 \leq i_n \implies k_{s+1} \leq k_{i_n} \implies b \leq X_{k_{s+1}} \leq Q_n^\omega(p)$$

prowadzą w granicy do nierówności $b \leq \limsup_n Q_n^\omega(p)$. Nierówności przeciwnie zawarte są w udowodnionej wcześniej zależności (2.19), a wykazane w ten sposób równości zachodzą dla prawie wszystkich $\omega \in \Omega'_0$ – na zbiorze, który możemy oznaczyć Ω_0^p .

Dla zakończenia dowodu ograniczamy się do przekroju $\Omega_{00} = \bigcap_{p \in P} \Omega_0^p$ przeliczalnej rodziny zdarzeń o prawdopodobieństwie $\mathbb{P}(\Omega_{00}) = 1$. $\square$

2.2.5 Nieskończone próby losowe

Empiryczny charakter twierdzenia Gliwienki–Cantelliego podkreśla odwołujący się do pojedynczej zmiennej losowej

Wniosek 2.2.30. *Niech $X: \Omega \rightarrow \mathbb{R}$ będzie dowolną zmienną losową. Dla prawie wszystkich prób losowych $(X(\omega_n))_{n \in \mathbb{N}}$ w $\mathbb{R}^{\mathbb{N}}$, czyli prawie na pewno*

(i) *funkcje*

$$F_n(x) = \frac{1}{n} \cdot \#\{k \leq n; X(\omega_k) \leq x\} \quad \text{dla } x \in \mathbb{R}, n \in \mathbb{N},$$

dążą jednostajnie do dystrybuanty F_X zmiennej X , oraz

(ii) *ciąg funkcji kwantylowych $Q_n: (0, 1) \rightarrow \mathbb{R}$, takich że*

$$Q_n(u) = X(\omega_{k_i}) \quad \text{dla } \frac{i-1}{n} < u \leq \frac{i}{n},$$

gdzie $X(\omega_{k_1}) \leq \dots \leq X(\omega_{k_n})$ jest uporządkowanym ciągiem wartości $X(\omega_i)$ dla $i \leq n$, dąży do funkcji Q^F w sensie równości (2.74) $\square$

Opis i pomiar wyników ciągu powtarzalnych doświadczeń w świecie rzeczywistym generuje (skończony) ciąg liczb $(x_k)_{k \leq n}$, na podstawie którego chcemy wyznaczać przybliżony rozkład zmiennej losowej X , za jaką uważamy wynik pomiaru. Zakładając *praktyczną* niezależność kolejnych pomiarów, możemy uznać średnią $F(x) = \frac{1}{n} \#\{k \leq n; x_k \leq x\}$, $x \in \mathbb{R}$, za empiryczną dystrybuantę badanej zmiennej.

Wyznaczona w ten sposób funkcja nie zależy oczywiście od kolejności wyrazów ciągu – i kolejności dokonywanych pomiarów. Jednak, jeśli kolejne pomiary stanowią np. ciąg rosnący, możemy mieć uzasadnione wątpliwości co do niezależności pomiarów – a więc także co do wiarygodności uzyskanych wyników. Skorzystajmy zatem z uogólnienia, w którym próby losowe grupowane są w ciągi stałej długości: $(x_1, \dots, x_p), (x_{a+1}, \dots, x_{a+p})$, itd.

Twierdzenie 2.2.31 (O regularnej różnorodności prób losowych). *Niech $\mathbb{P}$ będzie dowolnym rozkładem prawdopodobieństwa na prostej $\mathbb{R}$. Dla każdego $p \in \mathbb{N}$, przesunięcia $a \leq p$ i dowolnych przedziałów $J_i \subset \mathbb{R}$, $i \leq p$, zbiór Θ tych prób losowych $\omega = (x_k)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$, dla których zachodzi zbieżność*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \#\{0 \leq k < n; x_{ak+i} \in J_i \text{ dla } i = 1, \dots, p\} = \prod_{i \leq p} \mathbb{P}(J_i) \quad (2.77)$$

ma prawdopodobieństwo $\mathbb{P}^{\mathbb{N}}(\Theta) = 1$.

Udowodnimy najpierw wielowymiarowy wariant twierdzenia Gliwienki–Cantelliego ograniczony do zbieżności dystrybuant.

Lemat 2.2.32. Niech $X_n: \Omega \rightarrow \mathbb{R}, n \in \mathbb{N}$, będzie ciągiem niezależnych zmiennych losowych o tym samym rozkładzie. Dla ustalonej liczby naturalnej p niech $Y_k = (X_{p(k-1)+1}, \dots, X_{pk}), k \in \mathbb{N}$, oznacza wyznaczony przez (X_n) ciąg zmiennych wektorowych. Wówczas dla prawie wszystkich $\omega \in \Omega$ (czyli – prawie na pewno) ciąg dystrybuant empirycznych

$$F_n^\omega(t) = \frac{1}{n} \# \{k \leq n; Y_k(\omega) \leq t\} \quad \text{dla } t = (t_i)_{i \leq p} \in \mathbb{R}^p, n \in \mathbb{N},$$

dąży jednostajnie do funkcji $F_Y: \mathbb{R}^p \rightarrow [0, 1]$ będącej wspólną dystrybuantą zmiennych wektorowych $Y_k, k \in \mathbb{N}$.

Zgodnie z powszechnie przyjętą konwencją zdania postaci $Y_k(\omega) \leq t (< t)$ rozumiemy jako koniunkcję nierówności $X_{p(k-1)+i} \leq t_i (< t_i)$ dla $i \leq p$.

Dowód. Schemat dowodu powtarza I część dowodu twierdzenia 2.2.29 i korzysta z analogicznych oznaczeń. Niech F będzie wspólną dystrybuantą zmiennych X_n , a więc $F_Y(t) \equiv \prod_{i \leq p} F(t_i)$, a dla dowolnego ustalonego $m \in \mathbb{N}$ rozważamy kwantyle $x_{mi}, i \leq m$, spełniające warunki (2.75).

Dla każdego $t \in \mathbb{R}^p$ z twierdzenia 2.1.1 wynika zbieżność prawie wszędzie

$$F_n^\omega(t) \longrightarrow F_Y(t), \quad \text{oraz} \\ F_n^\omega(t-) \equiv \frac{1}{n} \# \{k \leq n; Y_k(\omega) < t\} \longrightarrow F_Y(t-),$$

dla $n \rightarrow \infty$, gdzie wykorzystana notacja odnosi się do lewostronnych granic rozważanych dystrybuant. Aby w pełni wykorzystać zbieżność dla punktów t o współrzędnych kwantylowych, zauważamy, że w przypadkach wykraczających poza $\mathbb{R}^p$, gdy poszczególne współrzędne są nieskończone, zbieżność także ma miejsce – dla odpowiednich rozkładów brzegowych (gdy $t_i = \infty$) lub dla ciągów stałych równych 0 (gdy $t_i = -\infty$). Tak określony skończony (dla danego m) zbiór punktów pozwala znaleźć indeks $N_m \in \mathbb{N}$ i zależny od m wspólny zbiór $\Omega_m \subset \Omega$ o mierze $\mathbb{P}(\Omega_m) = 1$, takie że

$$|F_Y(t) - F_n^\omega(t)| < \frac{1}{m} \quad \text{i} \quad |F_Y(t-) - F_n^\omega(t-)| < \frac{1}{m} \quad \text{dla } n \geq N_m, \omega \in \Omega_m \quad (2.78)$$

i dla wszystkich $t = (x_{m\alpha_i})_{i \leq p} \in Q_m^p$, gdzie $Q_m = \{x_{mi}; 0 \leq i \leq m\}$. Współrzędne dowolnego punktu $t \in \mathbb{R}^p$ spełniają nierówności $x_{m, \alpha_i - 1} \leq t_i < x_{m, \alpha_i}$ dla $i \leq p$ i pewnych $\alpha_i \leq m$. Stąd, z (2.75) i z (2.78) otrzymujemy

$$\begin{aligned} F_Y(t) &\leq \prod_{i \leq p} F(x_{m, \alpha_i} -) \leq \prod_{i \leq p} (F(x_{m, \alpha_i - 1}) + \frac{1}{m}) \\ &\leq F_Y((x_{m, \alpha_i - 1})_{i \leq p}) + \sum_{i=1}^p \binom{p}{i} \frac{1}{m^p} < F_n^\omega(t) + \frac{2^p}{m} \quad \text{oraz} \\ F_n^\omega(t) &\leq F_n^\omega((x_{m, \alpha_i})_{i \leq p} -) < \prod_{i \leq p} F(x_{m, \alpha_i} -) + \frac{1}{m} \\ &\leq \prod_{i \leq p} (F(x_{m, \alpha_i - 1}) + \frac{1}{m}) + \frac{1}{m} \\ &\leq \prod_{i \leq p} F(x_{m, \alpha_i - 1}) + \frac{2^p}{m} \leq F_Y(t) + \frac{2^p}{m}. \end{aligned}$$

Udowodnione oszacowanie $|F_Y(t) - F_n^\omega(t)| < \frac{2^p}{n}$, dla $t \in \mathbb{R}^p$, $\omega \in \Omega_m$ i $n \geq N_m$ implikuje jednostajną zbieżność badanego ciągu empirycznych dystrybuant dla wszystkich $\omega \in \bigcap_{m \in \mathbb{N}} \Omega_m$. $\square$

Ćwiczenie 2.2.13. Zbieżność jednostajna $F_n^\omega \Rightarrow F_Y$ przenosi się na granice lewostronne funkcji (por. ćwiczenie 1.1.14).

Dowód twierdzenia 2.2.31. Dla dowolnego punktu $t = (t_i)_{i \leq p} \in \mathbb{R}^p$ duża (równa $2^p - 1$) liczba lewostronnych, częściowych granic dystrybuant w t sugeruje wprowadzenie adekwatnych oznaczeń. Dla dowolnego rozkładu zbioru indeksów na dwa rozłączne podzbiory $\{1, \dots, p\} = A \cup B$ punkt t możemy zapisać jako parę (formalnie – sumę mnogościową) $t = (t_A, t_B)$, gdzie $t_A: A \rightarrow \mathbb{R}$ i $t_B: B \rightarrow \mathbb{R}$ są obcięciami funkcji $t: \{1, \dots, p\} \rightarrow \mathbb{R}$. Pozwala to na wprowadzenie oznaczeń $F_Y(t_{A-}, t_B)$ i $F_n^\omega(t_{A-}, t_B)$ dla lewostronnych granic rozważanych dystrybuant w stosunku do współrzędnych o indeksach ze zbioru A . Korzystamy z lematu 2.2.32 w przypadku, gdy $X_n: \mathbb{R}^N \rightarrow \mathbb{R}$ oznaczają kolejne współrzędne, a współrzędnymi zmiennych wektorowych Y_k są $Y_{k,i} \equiv X_{p(k-1)+i}$, dla $i \leq p$. Z udowodnionej w lemacie jednostajnej zbieżności wynika zbieżność ciągu $F_n^\omega(t_{A-}, t_B)$ dla p.w. $\omega \in \mathbb{R}^N$,

$$\frac{1}{n} \# \{k \leq n; Y_{k\alpha} < t_\alpha, Y_{k\beta} \leq t_\beta \text{ dla } \alpha \in A, \beta \in B\} \rightarrow \prod_{\alpha \in A} F(t_{\alpha-}) \prod_{\beta \in B} F(t_\beta).$$

Ze wzoru 1.23 i analogicznej formuły dla produktu dowolnych przedziałów (otwartych, domkniętych itd. – por. ćwiczenie 1.1.14) wnioskujemy o zbieżności

$$\frac{1}{n} \# \{k \leq n; Y_{k,i}(\omega) \in J_i, \text{ dla } i \leq p\} \rightarrow \prod_{i \leq p} \mathbb{P}(J_i) \quad (2.79)$$

dla prób losowych $\omega = (x_k)_{k \in \mathbb{N}}$ z tego samego zbioru Ω_o o prawdopodobieństwie $\mathbb{P}^N(\Omega_o) = 1$.

W ostatniej części dowodu ustalamy dowolne przesunięcie $a \leq p$. W badanym ciągu wektory $(x_1, \dots, x_p)$ i $(x_{ak+1}, \dots, x_{ak+p})$ mają rozłączne zbiory indeksów – jeśli $ak \geq p$, a więc dla $k \geq \lceil p/a \rceil$. Oznacza to, że rozważany ciąg p -wektorów można rozdzielić na $k_o = \lceil p/a \rceil$ podciągów, z których każdy spełnia założenia lematu 2.2.32 – i ma własność (2.79). Wystarczy przeprowadzić podział zbioru $Z_a(n) = \{0 \leq k < n; x_{ak+i} \in J_i \text{ dla } 1 \leq i \leq p\}$ według reszt $j = k \bmod k_o$. Dokładniej, dla $n = k_o n' + h$, gdzie $0 \leq h < k_o$ mamy

$$\begin{aligned} Z_a(n) = & \bigcup_{j=0}^{h-1} \{0 \leq k' \leq n'; x_{(ak_o)k'+j+i} \in J_i \text{ dla } i \leq p\} \\ & \cup \bigcup_{j=h}^{k_o-1} \{0 \leq k' < n'; x_{(ak_o)k'+j+i} \in J_i \text{ dla } i \leq p\}, \end{aligned}$$

co zapisujemy jako rozkład $Z_a(n) = \bigcup_{j < h} Z_{a_{k_o, j}}(n' + 1) \cup \bigcup_{j \geq h} Z_{a_{k_o, j}}(n')$. Możemy ograniczyć próby losowe ω do przekroju Θ zbiorów, dla których każdy z ciągów postaci $\frac{1}{m} \# Z_{a_{k_o, j}}(m)$ jest zbieżny – zgodnie z (2.79) – dla wszystkich p -wymiarowych przedziałów w $\mathbb{R}^p$. Dla ustalonego przedziału o mierze $V \equiv \prod_{i \leq p} \mathbb{P}(J_i)$ i dowolnego $\varepsilon > 0$ ustalmy liczbę M_ε taką, że wyrazy każdego z ciągów różnią się od granicy o mniej niż $\varepsilon > 0$, jeśli $m \geq M_\varepsilon$. Dla każdego n postaci $n = k_o n' + h$, gdzie $n' \geq M_\varepsilon, h < k_o$, stwierdzamy nierówność

$$\begin{aligned} \left| \frac{1}{n} \# Z_a(n) - V \right| &\leq \frac{1}{n} ((n' + 1) \sum_{j < h} \left| \frac{1}{n' + 1} \# Z_{a_{k_o, j}}(n' + 1) - V \right| \\ &\quad + n' \sum_{j \geq h} \left| \frac{1}{n'} \# Z_{a_{k_o, j}}(n') - V \right|) < \varepsilon. \end{aligned}$$

Otrzymana zbieżność kończy dowód twierdzenia. $\square$

2.3 Modelowanie rzeczywistości czyli nieco statystyki

Tytuł bieżącego podrozdziału ma sugerować odwrócenie ról, które pojawia się wtedy, gdy punktem wyjścia do analizy nie jest rozkład prawdopodobieństwa lecz ciąg danych empirycznych powstałych w wyniku sekwencji pomiarów (obserwacji) lub – jako efekt symulacji komputerowej.

2.3.1 Jednostajna losowość

Teza twierdzenia 2.2.31 o *regularnej różnorodności prób losowych* wyróżnia w $\mathbb{R}^N$ ciągi liczb, które potencjalnie mogłyby być próbami losowymi dla nieznanego rozkładu. W przypadku rozkładu jednostajnego na odcinku $[0, 1)$, a więc dla miary Lebesgue'a ℓ , przyjmujemy:

Definicja 2.3.1. *Ciąg liczb rzeczywistych $(x_n)_{n \in \mathbb{N}}$ z przedziału jednostkowego jest p -jednostajny, jeśli dla każdego niepustego p -wymiarowego przedziału $[a_1, b_1) \times \dots \times [a_p, b_p) \subset [0, 1)^p$ zachodzi zbieżność*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \# \{0 \leq k < n; \forall i \leq p \ a_i \leq x_{k+i} < b_i\} = \prod_{i \leq p} (b_i - a_i).$$

Ciąg $(x_n)_{n \in \mathbb{N}}$ nazywamy ∞ -jednostajnym, jeśli jest p -jednostajny dla każdego $p \in \mathbb{N}$.

Warunek p -jednostajności ciągu jest słabszy niż uwzględniająca także przesunięcia rodzina warunków odpowiadająca (2.77) i oznaczająca tzw. (p, a) -jednostajność. W książce Donalda Knutha [14, str. 156-8] podany jest dowód twierdzenia, że ∞ -jednostajny ciąg jest także (p, a) -jednostajny dla dowolnych $p, a \in \mathbb{N}$. Zachęcając czytelnika do przestudiowania istotnych

uwag i alternatywnych propozycji sformułowanych tamże, możemy traktować ∞ –jednostajność za podstawowe kryterium intuicyjnie pojmowanej *losowości ciągu* – czyli zgodności z teoretyczną próbą losową rozkładu jednostajnego. W pracy [7] znajdujemy m.in. dowód następującego faktu

Twierdzenie 2.3.2. *Ciąg liczb $U_n^\theta = \theta^n \bmod 1$, $n \in \mathbb{N}$, jest ∞ –jednostajny dla ℓ –prawie wszystkich liczb rzeczywistych $\theta > 1$.*

Uwaga 2.3.1. Jeśli $1 + a_1\theta + \dots + a_p\theta^p = 0$, dla pewnych $a_i \in \mathbb{Q}$, $i \leq p$, a więc θ jest liczbą algebraiczną stopnia $\leq p$, to ciąg U_n^θ nie jest $(p+1)$ –jednostajny, a wszystkie ciągi postaci $(U_n, U_{n+1}, \dots, U_{n+p}) \in (0, 1)^{p+1}$ leżą na skończonej liczbie równoległych hiperpłaszczyzn $x_0 + a_1x_1 + \dots + a_px_p = k \in \mathbb{Z}$.

Przeprowadzane w latach 60-tych ubiegłego wieku intensywne badania ciągu dla $\theta = \pi$ nie zakończyły się pozytywnym wynikiem. W istocie nie jest znany algorytm, który dla danej liczby przestępnej θ pozwalałby potwierdzić ∞ –jednostajność ciągu U_n^θ . Dla odmiany sam Knuth w jednej ze swoich prac skonstruował (jawnie) ∞ –jednostajny ciąg złożony z liczb wymiernych.

Analogiczne, równoległe podejście do ∞ –jednostajności od strony liczb całkowitych umożliwiając

Definicja 2.3.3. *Dla dowolnej liczby naturalnej $b \geq 2$ ciąg liczb całkowitych $(x_n)_{n \in \mathbb{N}}$ ze zbioru $\mathbb{Z}_b \equiv \{0, 1, \dots, b-1\}$ będziemy nazywać ciągiem b –arnym. Ciąg b –arny $(x_n)_{n \in \mathbb{N}}$ jest p –jednostajny, jeśli dla każdego punktu $a \in \mathbb{Z}_b^p$*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \# \{0 \leq k < n; (x_{k+1}, \dots, x_{k+p}) = a\} = 1/b^p.$$

Ciąg nazywamy ∞ –jednostajnym, jeśli jest p –jednostajny dla każdego p .

... i – jako ćwiczenie – łatwe do udowodnienia

Twierdzenie 2.3.4. (i) *Jeśli ciąg $(x_n)_{n \in \mathbb{N}}$ liczb z przedziału $[0, 1)$ jest ∞ –jednostajny, to dla każdej liczby naturalnej $b \geq 2$ części całkowite $\lfloor bx_n \rfloor$, $n \in \mathbb{N}$, tworzą ciąg b –arny ∞ –jednostajny.*

(ii) *Jeśli dla ciągu $x = (x_n)_{n \in \mathbb{N}}$ liczb z przedziału $[0, 1)$ istnieje ciąg liczb naturalnych $1 < b_1 < b_2 < \dots$ taki, że każdy z ciągów b_i –arnych $(\lfloor b_i x_n \rfloor)_{n \in \mathbb{N}}$ jest ∞ –jednostajny, to x jest ∞ –jednostajny. $\square$*

Stąd *praktyczny wniosek*: Ciąg liczb rzeczywistych $(x_n)_{n \in \mathbb{N}}$ jest ∞ –jednostajny, wtedy i tylko wtedy, gdy dla dowolnej podstawy $b \geq 2$ kolejne cyfry $x_n^{(i)}$ rozwinięcia b –arnego liczb x_n tworzą ciągi $(x_n^{(i)})_{n \in \mathbb{N}}$ ∞ –jednostajne.

Podstawową wadą sugerowanej wyżej definicji losowości jest jej graniczny, *nieskończony* charakter. Ciąg 1000-jednostajny musi np. zawierać podciągi złożone z 1000 kolejnych wyrazów $x_i < \frac{1}{10}$, by w granicy nie pominąć zbioru o mierze $10^{-1000} > 0$.

2.3.2 Skończone ciągi pseudolosowe

Wyjątkowo proste rekurencyjne formuły postaci

$$X_{n+1} = (aX_n + c) \bmod m \quad \text{lub po prostu} \quad X_{n+1} = aX_n \bmod m, \quad (2.80)$$

dla $n \geq 0$, nazywane są *metodą liniowej kongruencji* i od połowy ubiegłego wieku służą za podstawę generowania *pseudolosowych* ciągów liczb całkowitych z przedziału $0 \leq X_n \leq m-1$. Ciąg ilorazów $X'_n = X_n/m \in [0, 1)$ jest wówczas przybliżeniem próby losowej z rozkładu jednostajnego. Praktyka pokazała, że liczne bardziej skomplikowane i w pewnej mierze losowo dobierane wzory generują rozpoznawane po dłuższym czasie trudne do lokalizacji anomalie, natomiast raczej nie zwiększają postulowanej *losowości ciągu*.

Formuła (2.80) zależy od parametrów (X_0, a, c, m) , gdzie $a, m > 1$, a pierwsze 3 parametry można zredukować do reszt $\bmod m$. Liczba m , czyli *moduł* ogranicza zakres możliwych wartości i wymusza okresowość ciągu. Oczywiście – długość okresu równa m oznacza, że pierwsze (i każde kolejne) m wyrazów ciągu jest permutacją liczb $0, \dots, m-1$. Możliwie duża długość okresu stanowi skończony odpowiednik 1-jednostajności ciągu. Za [14, par. 3.2.1.2] przytaczamy krótki zestaw istotnych faktów.

Twierdzenie 2.3.5. *Ciąg (2.80) ma okres m wtedy i tylko wtedy, gdy*

- *liczby m, c są względnie pierwsze oraz*
- *$a \bmod p = 1$, dla każdej liczby pierwszej p takiej, że $p|m$, przy czym $a \bmod 4 = 1$, jeśli $4|m$.*

W ogólnej sytuacji działa (co podajemy jako proste ćwiczenie)

Stwierdzenie 2.3.6. *Jeśli $m = p_1^{e_1} \dots p_t^{e_t}$ jest rozkładem modułu m na czynniki pierwsze, a λ_j oznacza długość okresu ciągu (2.80) z parametrami $(X_0, a, c, p_j^{e_j})$, dla $j \leq t$, to najmniejsza wspólna wielokrotna*

$$\lambda = \text{NWW}(\lambda_1, \dots, \lambda_t)$$

jest długością okresu ciągu (2.80). □

W przypadku $c = 0$ przy naturalnym założeniu, że m i X_0 (a więc także wszystkie kolejne wyrazy ciągu) są względnie pierwsze, największą długość okresu wskazuje pochodzące sprzed 2 wieków

Twierdzenie 2.3.7 (Gauss 1801). *Jeśli moduł ciągu jest postaci $m = p^e$, a wartość początkowa X_0 jest niepodzielna przez liczbę pierwszą p , wówczas ciąg $X_n = a^n X_0 \bmod m$ ma okres o maksymalnej długości $\lambda(m)$, gdzie*

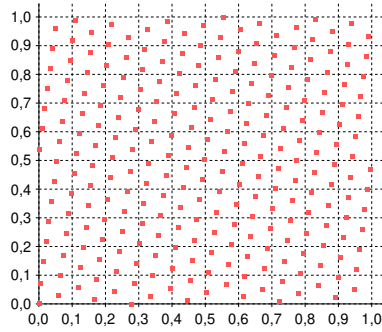
$$\lambda(2^e) = 2^{e-2} \quad \text{dla } e \geq 3, \quad \lambda(p^e) = p^{e-1}(p-1), \quad \text{gdy } p > 2 \text{ lub } e \leq 2,$$

wtedy i tylko wtedy, gdy czynnik a jest elementem pierwotnym modulo m . Z wyłączeniem wartości $m \in \{2, 4, 8\}$ okres ciągu ma długość $\lambda(m)$, gdy

- (i) $p = 2, e \geq 4$ i $a \bmod 8 \in \{3, 5\}$;
- (ii) $p > 2, e = 1$, $a \bmod p \neq 0$ i $a^{(p-1)/q} \bmod p \neq 1$ dla każdego dzielnika pierwszego q liczby $p-1$;
- (iii) $p > 2, e > 1$, a spełnia warunek (ii) i dodatkowo $a^{p-1} \bmod p^2 \neq 1$.

Jak stąd wynika, jeśli $m = p$ jest liczbą pierwszą, okres ciągu jest mniejszy niż m , ale może być równy $\varphi(m) = m-1$ w opisanym wyżej przypadku (ii).

Przypominamy, że dla $p > 2$ wartość $\lambda(p^e)$ jest równa funkcji Eulera $\varphi(p^e)$. To, że czynnik a jest *pierwotny* względem m , oznacza, że $n = \varphi(m)$ jest najmniejszym wykładnikiem, dla którego $a^n \bmod m = 1$.



Rysunek 2.5: Pary (y_n, y_{n+1}) ciągu pseudolosowego

Przykład 2.3.2. Parametry $(0, 137, 1, 2^8)$ wyznaczają ciąg kongruencyjny (x_n) o okresie równym $m = 256$. Ciąg ułamków $y_n = x_n/m$, $n < m$, regularnie wypełnia przedział $(0, 1]$ – na ile to możliwe – natomiast rys. 2.5 podaje rozkład par kolejnych wyrazów ciągu w kwadracie $[0, 1)^2$. Widoczna na ilustracji udana symulacja 2–jednostajności ciągu nie jest przypadkiem lecz jednym z dostępnych dla $m = 256$ optymalnych wyborów.

Obiektywnym miernikiem pozwalającym na ocenę symulowanej 2–jednostajności ciągu są liczby wyznaczone w tzw. *teście spektralnym*, odpowiadające największej odległości pomiędzy sąsiednimi równoległymi prostymi

zawierającymi zbiór par. Dane widoczne na rys. 2.5 leżą na każdej z trzech rodzin prostych równoległych

$$\begin{cases} 2y_{n+1} - 18y_n - \frac{2}{256} = k \\ 13y_{n+1} + 11y_n - \frac{13}{256} = k \\ 15y_{n+1} - 7y_n - \frac{15}{256} = k \end{cases} \text{ dla } k \in \mathbb{Z}, \text{ czyli } \begin{cases} 2x_{n+1} - 18x_n \equiv 2 \pmod{m} \\ 13x_{n+1} + 11x_n \equiv 13 \pmod{m} \\ 15x_{n+1} - 7x_n \equiv 15 \pmod{m} \end{cases}$$

W ramach każdej z rodzin o odległości prostych decyduje norma wektora współczynników. Jej kwadrat wynosi odpowiednio 328, 290 i 274, wyznaczona w teście wielkość jest równa $\nu_2^2 = 274$, a największą odległością jest $1/\nu_2 \approx 0.06$. Dla porównania – czynnik $a = 133$ prowadzi do ciągu spełniającego zależność $2x_{n+1} - 10x_n \equiv 2 \pmod{256}$, a czynnik $a = 141$ do równania liniowego $9x_{n+1} + 11x_n \equiv 9 \pmod{256}$. Odpowiednie wartości testu wynoszą 104 i 202 i są gorsze od osiągniętej wcześniej wielkości 274.

Podany w [14, sec. 3.3.4] algorytm pozwala efektywnie – zależnie od szybkości dostępnych komputerów – wyznaczać analogiczne wartości ν_k^2 ściśle związane z największą odległością kolejnych równoległych $(k-1)$ -wymiarowych hiperpłaszczyzn zawierających punkty $(y_{n+1}, \dots, y_{n+k}) \in [0, 1]^k$, dla $k > 2$ (por. uwaga 2.3.1). W przypadku badanego ciągu zależność

$$x_{n+1} = 137x_n + 1 \pmod{256} \implies x_{n+2} + 5x_{n+1} + 2x_n \equiv 143 \pmod{256}$$

pozwała wyznaczyć $\nu_3^2 = 30$ i maksymalną odległość sąsiednich płaszczyzn równą $\frac{1}{\sqrt{30}} \approx 0.316$. Tak duże „luki” w $[0, 1]^3$ i związany z tym brak 3-jednostajności jest także konsekwencją bardzo małej długości ciągu.

Wartość $c = 1$ prowadzi do ciągów o wyrażnie „nielosowej” formule

$$Y_n \equiv a^{n-1} + \dots + a + 1 = \frac{a^n - 1}{a - 1} \pmod{m}, \quad \text{dla } n \in \mathbb{N},$$

jeśli przyjąć $Y_0 = 0$. Wówczas ogólny wzór dla ciągu (X_0, a, c, m) z $c > 1$ podaje

Stwierdzenie 2.3.8. *Zachodzi równość*

$$X_n = (\tilde{a}Y_n + X_0) \pmod{m}, \quad \text{dla } n \in \mathbb{N},$$

gdzie współczynnik $\tilde{a} = (a-1)X_0 + c$ może być również zredukowany $\pmod{m}$.

Dowód. Sprawdzamy tożsamość $X_n = (a^n X_0 + (a^{n-1} + \dots + a + 1)c) \pmod{m}$, po czym postawiamy $a^n = (a-1)Y_n + 1 \pmod{m}$. $\square$

Wybór przesunięcia $c \geq 2$ oznacza zatem dodatkową permutację w ciągu $(X_0, a, 1, m)$ – na tyle szczególną, że każdej nietrywialnej liniowej zależności postaci

$$\alpha_0 Y_n + \alpha_1 Y_{n+1} + \dots + \alpha_k Y_{n+k} = \text{const} \pmod{m}, \quad \text{dla } n \in \mathbb{N},$$

gdzie $\alpha_i \in \mathbb{Z}$ dla $i \leq k$, odpowiada liniowa zależność dla ciągu $(X_n)_{n \in \mathbb{N}}$ – z tymi samymi współczynnikami $\alpha_i, i \leq k$, ale z inną stałą. Niezależną od m interpretację wskaźników k –jednostajności

$$\nu_k^2 = \min \sum_{i \leq k} \alpha_i^2, \quad \text{dla } k \geq 2,$$

ułatwia bliski związek odległości $d_k \equiv \nu_k^{-1}$ i maksymalnej średnicy kuli w $[0, 1]^k$ niezawierającej punktów $(X_n, X_{n+1}, \dots, X_{n+k})$, $n \in \mathbb{N}$. Ze względu na prostotę obliczeń proponujemy porównanie d_k z długością krawędzi przeciętnej kostki w $[0, 1]^k$ o mierze $\frac{1}{m}$. Stąd *miara k –niejednostajności*

$$\gamma_k = m^{1/k} / \nu_k, \quad k = 2, 3, 4, \dots, \quad (2.81)$$

dla której $d_k = \gamma_k \sqrt[k]{\frac{1}{m}}$, tzn. im mniejsza wartość $\gamma_k \approx 1$ tym większy poziom k –jednostajności ciągu. Dla poprzedniego przykładu wyznaczamy $\gamma_2 \approx 0.9666$, $\gamma_3 \approx 1.159$.

Przykład 2.3.3. Współczesne systemy komputerowe oferują zrównoważone i przetestowane generatory liczb pseudolosowych (całkowitych i rzeczywistych). W standardowej bibliotece języka Java za pseudolosową symulację rozkładu jednostajnego odpowiada klasa `java.util.Random`. Podstawą zapisanego tam algorytmu są parametry: moduł $m = 2^{48} = 281474976710656$, czynnik $a = 25214903917$ i przesunięcie $c = 11$. Generowany (wewnątrz funkcji `next`) w myśl reguły

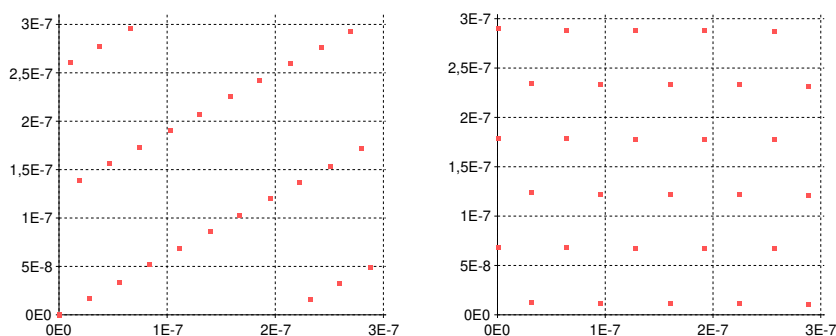
$$\begin{aligned} \text{oldseed} &= \text{seed} \\ \text{nextseed} &= a \cdot \text{oldseed} + c \pmod{m} \\ \text{seed} &= \text{nextseed} \end{aligned}$$

ciąg ma zgodnie z twierdzeniem 2.3.5 maksymalną długość okresu przekraczającą $2.8 \cdot 10^{15}$. Ciąg przebiega zatem w ustalonej w ten sposób kolejności wszystkie liczby całkowite z zakresu $0 \dots m-1$. Obiekt losujący (egzemplarz klasy `Random`) przechowuje zawsze aktualny, ostatnio wyznaczony wyraz ciągu jako tzw. *ziarno* (ang. *seed*). Jako wartość początkową Java jednorazowo wykorzystuje przetworzony parametr s podany w chwili tworzenia obiektu losującego: $\text{seed} = s \text{ xor } a \pmod{m}$. Wskaźniki 2– i 3–jednostajności ciągu wydają się akceptowalne i wynoszą

$$\begin{aligned} \nu_2^2 &= 84862060372330 = 7847617^2 + 4824621^2, & \gamma_2 &\approx 1.82 \\ \nu_3^2 &= 3489362614 = 15843^2 + 36586^2 + 43587^2, & \gamma_3 &\approx 1.11 \\ \nu_4^2 &= 4788790 = 448^2 + 1082^2 + 1529^2 + 1039^2, & \gamma_4 &\approx 1.87 \end{aligned}$$

Najważniejsza funkcja $\text{next}(\text{nbits})$ odrzuca zawsze przynajmniej 12 dolnych bitów liczby nextseed , a jako wynik $\text{nextseed} \ggg (48 - \text{nbits})$ podaje dokładnie górne nbits bitów (cyfr w rozwinięciu binarnym). Zapotrzebowanie na kolejną liczbę losową realizowane jest poprzez jedno lub dwa wywołania funkcji $\text{next}(\text{nbits})$

$$\begin{aligned}\text{nextInt}() &= \text{next}(32) \\ \text{nextLong}() &= 2^{32} \cdot \text{next}(32) + \text{next}(32) \\ \text{nextFloat}() &= \text{next}(24)/2^{24} \\ \text{nextDouble}() &= (2^{27} \cdot \text{next}(26) + \text{next}(27))/2^{53}\end{aligned}$$



Rysunek 2.6: Rozkład par: ciągi pseudolosowe dla przykładów 2.3.3 i 2.3.5

Uwaga 2.3.4. Odrzucanie dolnych bitów generowanych liczb całkowitych uzasadnia następująca obserwacja. Jeśli liczby pseudolosowe $X_n, n \geq 0$, są związane rekurencją postaci $X_{n+1} = aX_n + c \bmod 2^e$, to dla każdego wykładnika $f < e$ ciąg $Z_n \equiv X_n \bmod 2^f$ spełnia zależność $Z_{n+1} = aZ_n + c \bmod 2^f$. Implikuje to skrócenie do 2^f długości okresu i istotne obniżenie losowości dolnych cyfr rozwinięcia binarnego liczb X_n .

Przykład 2.3.5. Zgodnie z [14, str. 106-7] M. Lavaux and F. Janssens znaleźli czynnik $a = 31167285$, dla którego ciąg pseudolosowy o parametrach $(X_0, a, c, 2^{48})$ ma istotnie lepsze parametry spektralne (rys. 2.6):

$$\begin{aligned}\nu_2^2 &= 322492826755072 = 9031104^2 + 15521984^2, & \gamma_2 &\approx 0.9342 \\ \nu_3^2 &= 4111841446 = 53901^2 + 17554^2 + 29973^2, & \gamma_3 &\approx 1.0220 \\ \nu_4^2 &= 17341510 = 496^2 + 890^2 + 3515^2 + 1987^2, & \gamma_4 &\approx 0.9836\end{aligned}$$

2.3.3 Nieparametryczne testy losowości

Skończone, realnie wykorzystywane ciągi liczb *pseudolosowych* odpowiadają na zapotrzebowanie, by sekwencje *nierealistyczne* choć teoretycznie możliwe

nie pojawiały się w „skończonym” – dostępnym obserwacji czasie. W przypadku ciągu binarnego za naturalne, oczywiste zakłócenie losowości gotowi jesteśmy uznać zarówno zbyt długie sekwencje powtórzonych wartości (ciągi zer lub ciągi jedynek) jak i długie ciągi takich samych zmian, np. 0101...0101. W ogólnym przypadku możemy badać nierealistyczne, nie-zrównoważone powtarzalne zależności opisywalne np. poprzez relacje $\leq$ lub $<$, a więc wynikające z porównywania elementów badanego ciągu.

Przykład 2.3.6. W schemacie Bernoulliego dla danego zbioru n wyników możemy traktować względny udział liczby jedynek (sukcesów) w stosunku do liczności próby jako przybliżoną wartość prawdopodobieństwa sukcesu $p = \mathbb{P}\{1\} \in (0, 1)$. Dla wykluczenia skrajnych, trywialnych wartości p wystarczy, by zarówno sukces jak i porażka zdarzyły się co najmniej raz. Jeśli nie znamy p , w ciągu n próbek losowych żadna (nawet bardzo mała lub bardzo duża) liczba sukcesów $n_1 < n$ nie wskazuje na brak losowości próbki. Jeżeli jednak n jest relatywnie duże, a w próbce $(a_i)_{i \leq n} \in \{0, 1\}^n$ wystąpiły np. jedynie $n_1 = 3$ jedynki, to o możliwym braku wiarygodności (losowości) próby może świadczyć co najwyżej ich wzajemne położenie w ciągu obserwacji.

Dokładniej – z danych wynika, że $p > 0$, przy czym każda sekwencja $n+3$ liczb zawierająca zera i dokładnie 3 jedynki jest tak samo prawdopodobna. Zaobserwowano jedną z $\binom{n+3}{3}$ sekwencji – prawdopodobieństwo każdej wynosi $\binom{n+3}{3}^{-1}$. Wśród nich są

- 2 wyjątkowe – postaci 1110...0 i 0...0111
- $(n-1)+2 = n+1$ ciągów postaci ...01110... oraz 11...01 i 10...11
- $4(n-1)$ ciągów 11...010..., 10...110... oraz ...010...11 i ...011...01
- $(n-1)^2$ ciągów postaci ...011...010..., ...010...110... oraz 1...010...1
- $(n-1)(n-2)$ ciągów postaci 10...10...10... i ...01...01...01
- $\binom{n-1}{3}$ ciągów ...01...010...10...

Uważny i cierpliwy czytelnik zauważył zapewne, że kolejne przypadki zależą od łącznej liczby R przebiegów, czyli sekwencji (podciągów) złożonych na przemian z samych sukcesów lub samych porażek. Możliwe wartości to liczby 2,...,7. Dla dużych wartości n oznacza to rozkład prawdopodobieństwa zmiennej losowej R postaci

$$\mathbb{P}\{R = 2\} \approx \frac{12}{n^3}, \quad \mathbb{P}\{R \leq 4\} \approx \frac{30}{n^2}, \quad \text{ i } \quad \mathbb{P}\{R \leq 6\} \approx \frac{12}{n}.$$

Wynika stąd, że w przypadkach, gdy w ciągu $n = 250$ prób pojawiają się dokładnie 3 sukcesy, maksymalna liczba przebiegów $R = 7$ występuje przeciętnie w 95% obserwacji, a wartość $R \geq 5$ w 99.5% obserwacji.

Analogicznie, dla $n = 2500$ wartość $R = 7$ jest wspólna dla ponad 99.5% n -elementowych prób losowych (z liczbą sukcesów równą 3). Dla pseudolosowych ciągów liczb 0, 1 prawdopodobieństwo $\mathbb{P}\{R \leq 4\} \approx 5/10^6$ nie jest jednak nierealistycznie małe, a próbki (pseudo-)losowe długości 2500 powinny przeciętnie co 200 tys. zawierać $R = 3, 4$ przebiegi.

Dla $n \geq 2500$ próbki z liczbą przebiegów $R=2$ (wszystkie sukcesy na początku lub – wszystkie na końcu) należałoby odrzucać jako nierealistyczne. Równie *skrajnie nielosowe* są także próbki, dla których $R \leq 4$ i $n \geq 200000$.

Opisana w przykładzie sytuacja (b. duża liczba prób i tylko 3 sukcesy) sama w sobie mogłaby być uznana za skrajnie nierealistyczną, gdyby nie to, że dotyczy np. wzajemnego położenia w próbie 3 największych (lub 3 najmniejszych) wartości, albo – dwu wartości skrajnych i np. mediany. Jest to oczywiście związane z zamianą opisu wielkości (liczb) tworzących próbkę z ilościowego na jakościowy, gdzie rozróżnione zostają np. wyrazy „przeciętne” od „skrajnych” lub – dla odmiany – „ujemne” od „nieujemnych”. Naturalne dwa podejścia pozwalają wyróżnić dwie komplementarne koncepcje losowości ciągu danych liczbowych.

Definicja 2.3.9. *Dla skończonego ciągu liczb rzeczywistych $(x_k)_{k \leq n}$ przebiegiem monotonicznym (rosnącym – up lub malejącym – down) nazywamy każdy maksymalny monotoniczny podciąg*

$$x_i \leq x_{i+1} \leq \dots \leq x_{i+h}.$$

Dla ciągu liczb $(x_k)_{k \leq n}$ ograniczonych do relatywnie niewielkiego i niezależnego od n zbioru wartości przebiegiem nazywamy także każdy maksymalny podciąg stały $(x_{i+1}, \dots, x_{i+h})$ utworzony z kolejnych wyrazów rozważanego ciągu.

Liczbę $h \geq 1$ – zależną od kontekstu – nazywamy długością przebiegu.

W przypadku, gdy wyrazy rozważanego ciągu ograniczają się do dwu wartości całkowitych – zazwyczaj $\{-1, 1\}$ lub $\{0, 1\}$ – łączna liczba następujących po sobie przebiegów jest elementarną *statystyką*, cechą ciągu, którą możemy powiązać z jego hipotetyczną losowością.

Jako przestrzeń probabilistyczną rozważamy zbiór $B_{n_1 n_2} \subset \{0, 1\}^n$ wszystkich n -elementowych ciągów zawierających dokładnie n_1 zer (obiektów pierwszego typu) i n_2 jedynek (drugiego typu). Niezależnie od nieznanego prawdopodobieństwa sukcesu w schemacie Bernoulliego w $\{0, 1\}^n$ wszystkie ciągi w $B_{n_1 n_2}$ są tak samo prawdopodobne.

Ćwiczenie 2.3.7. Niech R_1 i R_2 oznaczają odpowiednio liczbę przebiegów n_1 obiektów pierwszego oraz n_2 obiektów drugiego typu – w próbach losowych rozmiaru $n = n_1 + n_2$. Sprawdzić, że łączny rozkład prawdopodobieństwa pary (R_1, R_2) opisuje wzór

$$\mathbb{P}\{R_1 = r_1, R_2 = r_2\} = c \frac{\binom{n_1-1}{r_1-1} \binom{n_2-1}{r_2-1}}{\binom{n_1+n_2}{n_1}} \quad \text{dla } r_1 \leq n_1, r_2 \leq n_1,$$

gdzie $c = 1$, gdy $r_1 - r_2 = \pm 1$, $c = 2$, gdy $r_1 = r_2$, i $c = 0$ (dlaczego?) w pozostałych przypadkach.

Podstawą dla oceny losowości badanego ciągu liczb może być

Twierdzenie 2.3.10 (Test runs). *Rozkład łącznej liczby $R = R_1 + R_2$ przebiegów w przestrzeni $B_{n_1 n_2}$ prób losowych złożonych z n_1 obiektów typu pierwszego i n_2 typu drugiego podaje wzór*

$$\mathbb{P}\{R = r\} = \begin{cases} 2 \binom{n_1-1}{r'-1} \binom{n_2-1}{r'-1} / \binom{n_1+n_2}{n_1}, & \text{gdy } r = 2r', \\ \left(\binom{n_1-1}{r'} \binom{n_2-1}{r'-1} + \binom{n_1-1}{r'-1} \binom{n_2-1}{r'} \right) / \binom{n_1+n_2}{n_1}, & \text{gdy } r = 2r' + 1 \end{cases}$$

dla $2 \leq r \leq n_1 + n_2$. Wartość oczekiwana i wariancja zmiennej R wynoszą

$$\mathcal{E}(R) = 1 + \frac{2n_1 n_2}{n} \quad \text{Var}(R) = \frac{2n_1 n_2 (2n_1 n_2 - n)}{n^2 (n - 1)}. \quad (2.82)$$

Definicja 2.3.11. *Opisany powyżej rozkład $WW(n_1, n_2)$ zmiennej R nazywamy rozkładem Walda–Wolfowitza.*

Dowód. Podane w tezie wzory na prawdopodobieństwo zdarzenia $R = r$ wynikają bezpośrednio z ćwiczenia 2.3.7 i z rozkładu każdego z przypadków na dwa zdarzenia zależne od wartości pierwszej próby.

Łączna liczba przebiegów $R = 1 + I_2 + \dots + I_n$ jest sumą jedynki oznaczającej stały początek pierwszego przebiegu oraz zmiennych losowych

$$I_k(x) = \begin{cases} 1, & \text{gdy } x_{k-1} \neq x_k \\ 0 & \text{gdy } x_{k-1} = x_k \end{cases} \quad \text{dla } 2 \leq k \leq n$$

sygnalizujących ewentualny początek kolejnego przebiegu. Z prostej zależności zmiennych I_k od dwu sąsiednich współrzędnych wynikają równości

$$\mathbb{P}\{I_k = 1\} = \frac{2n_1 n_2}{n(n-1)} = \mathcal{E}(I_k).$$

Stąd – podany wyżej wzór na $\mathcal{E}(R)$. Ponieważ pomocnicze zmienne są zależne, wariancję wyznaczamy z tożsamości $\text{Var}(R) = \mathcal{E}((\sum_k I_k)^2) - (\mathcal{E}(\sum_k I_k))^2$, tzn.

$$\text{Var}(R) = (n-1) \frac{2n_1 n_2}{n(n-1)} + 2 \sum_{2 \leq j < k \leq n} \mathcal{E}(I_j I_k) - \left(\frac{2n_1 n_2}{n} \right)^2.$$

Rozdzielając sumę na $n-2$ składniki, w których $k = j+1$, oraz $\binom{n-2}{2}$ składniki, w których $k-j \geq 2$, wyliczamy

$$\mathcal{E}(I_{k-1} I_k) = \frac{n_1(n_1-1)n_2}{n(n-1)(n-2)} + \frac{n_1 n_2(n_2-1)}{n(n-1)(n-2)} = \frac{n_1 n_2}{n(n-1)},$$

a dla $k \geq j+2$,

$$\mathcal{E}(I_j I_k) = 4 \cdot \frac{n_1(n_1-1)n_2(n_2-1)}{n(n-1)(n-2)(n-3)}.$$

Po wykonaniu prostych obliczeń we wzorze na $\text{Var}(R)$ otrzymujemy tezę. $\square$

Wniosek 2.3.12. *Jeśli długość n próbki rośnie do nieskończoności w ten sposób, że $n_1/n \rightarrow \lambda$ i $n_2/n \rightarrow 1-\lambda$, to*

$$\lim_n \mathcal{E}(R)/n = 2\lambda(1-\lambda), \quad \lim_n \text{Var}(R)/n = 4\lambda^2(1-\lambda)^2. \quad \square$$

Uwaga 2.3.8 (Test runs Walda–Wolfowitza). Graniczne wartości charakterystyk $\mathcal{E}(R)$ i $\text{Var}(R)$ oraz odpowiadająca im zbieżność ciągu odpowiednio znormalizowanych rozkładów prawdopodobieństwa do rozkładu normalnego (co jest stopniowo formalizowane w kolejnych podrozdziałach – por. twierdzenie 2.4.4), pozwala na ocenę losowości badanego ciągu prób. Dokładniej, dla dostatecznie dużych n i dowolnej n -elementowej próby – pomiarów procesu o nieznanym rozkładzie – *test runs* wyznacza prawdopodobieństwo podziału danych na przebiegi *poniżej* i *powyżej* dowolnego kwantyla $\lambda \in (0, 1)$, gdy $n_1 \approx \lambda n$, a $n_2 \approx (1-\lambda)n$.

Test runs jest w istocie rodziną testów – zmiennych losowych $R_{n_1 n_2}$, $n_1, n_2 \in \mathbb{N}$, określonych na przestrzeni borelowskiej $\mathbb{R}^n$, $n = n_1 + n_2$, z produktową miarą probabilistyczną $\mathbb{P} = P^n$, gdzie P oznacza dowolny rozkład na prostej $\mathbb{R}$ taki, że $P\{x\} = 0$ dla wszystkich zdarzeń elementarnych $x \in \mathbb{R}$. Uniwersalność testu potwierdza

Twierdzenie 2.3.13. *Niech $n_1, n_2 \in \mathbb{N}$ będą dowolnymi liczbami naturalnymi. Dla każdej przestrzeni probabilistycznej $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}}, P)$, w której zdarzenia elementarne mają miarę 0, łączna liczba przebiegów w zbiorze dolnych n_1 elementów i n_2 górnych elementów ciągu $n = n_1 + n_2$ elementowego jest zmienną losową $R_{n_1 n_2} : \mathbb{R}^n \rightarrow \mathbb{N}$ określoną prawie wszędzie względem miary $\mathbb{P} = P^n$. Rozkład indukowany $\mathbb{P}_R \equiv R_{n_1 n_2 *} \mathbb{P}$ nie zależy od miary P i jest rozkładem Walda–Wolfowitza $\text{WW}(n_1, n_2)$.*

Dowód. Założenie dotyczące rozkładu P oznacza, że funkcja kwantylowa Q rozkładu jest ściśle monotoniczna, oraz że $\mathbb{P}$ –prawie wszystkie ciągi w $\mathbb{R}^n$ mieszczą się w podprzestrzeni $\mathbb{R}^{(n)} \subset \mathbb{R}^n$ ciągów różnowartościowych. Przypominamy, że zgodnie z twierdzeniem 1.1.42 zachodzą równości $P = Q_*\ell_{[0,1]}$ i $\mathbb{P} = Q_*^n\ell^n$. Podział wyrazów ciągów różnowartościowych na „dolne” i „górne” wykonuje funkcja klasyfikująca $\Lambda_{n_1n_2}: \mathbb{R}^{(n)} \rightarrow B_{n_1n_2} \subset \{0,1\}^n$ określona wzorem

$$\Lambda_{n_1n_2}(x) = (b_i)_{i \leq n}, \quad \text{dla } b_i = \begin{cases} 1, & \text{gdy } \#\{j \leq n; x_j < x_i\} \geq n_1 \\ 0, & \text{gdy } \#\{j \leq n; x_j > x_i\} \geq n_2 \end{cases},$$

przy czym $R_{n_1n_2} = R \circ \Lambda_{n_1n_2}$ jest złożeniem funkcji klasyfikującej i rozważanej wyżej zmiennej R .

Z monotoniczności funkcji Q wynika równość $\Lambda_{n_1n_2} \circ Q^n = \Lambda_{n_1n_2}$, a więc $R_{n_1n_2} \circ Q^n = R_{n_1n_2}$ i $R_{n_1n_2*}\mathbb{P} = R_*\lambda_{n_1n_2*}\ell^n|_{[0,1]^n}$, co dowodzi niezależności miary indukowanej od P i sprowadza problem do miary Lebesgue’a na przedziale jednostkowym. Pozostaje zauważyć, że ze względu na symetrię względem współrzędnych funkcja klasyfikująca $\Lambda_{n_1n_2}$ indukuje z miary Lebesgue’a na kostce $[0,1]^n$ jednostajną miarę zliczającą na $B_{n_1n_2}$. $\square$

Porównanie obserwowanej liczby przebiegów z ilością teoretyczną, o akceptowalnym poziomie ufności, pozwala na odrzucenie (lub nie) hipotezy o losowości próby – w zależności od tego, czy ciąg danych $x = (x_i)_{i \leq n}$ wpada do któregoś ze zbiorów krytycznych, określonych przez skrajne wartości $R_{n_1n_2}$ bliskie 2 lub $2 \min(n_1, n_2) + 1$. Efektywną implementację testu Walda–Wolfowitza zawierają np. funkcje `randtest::runs.test()` i `TSA::runs()` dostępne w środowisku R-Studio.

Przebiegi monotoniczne prowadzą do analogicznego testu korzystającego z liczby *punktów zwrotnych*, dla których $x_{i-1} < x_i > x_{i+1}$ lub $x_{i-1} > x_i < x_{i+1}$, a więc z łącznej liczby przebiegów `up` i `down`. Test ten jest zaimplementowany np. w funkcji `randtest::turning.point.test()`.

Ćwiczenie 2.3.9. Wyznaczyć rozkład zmiennej losowej zliczającej liczbę punktów zwrotnych w n –elementowym losowym ciągu liczbowym. Obliczyć wartość oczekiwaną i wariancję.

Uwaga 2.3.10. Udowodnione wyżej twierdzenie wykorzystuje fakt, że każdy borelowski rozkład probabilistyczny na prostej może być otrzymany jako miara indukowana $Q_*\ell|_{[0,1]}$ z rozkładu jednostajnego na odcinku jednostkowym. Jeśli Q jest (dowolną) funkcją kwantylową rozkładu $\mathbb{P}$, to dla dowolnego ciągu pseudolosowego $(x_i)_{i \leq m}$ o rozkładzie jednostajnym na $[0,1]$ ciąg

$Q(x_i)_{i \leq m}$ stanowi symulację rozkładu $\mathbb{P}$. Skorzystaliśmy już z tej własności, przygotowując rysunki do podrozdziału 2.2.

2.3.4 Jaką informację wnosi skończony zbiór obserwacji?

Założmy, że dysponujemy (skończonym) ciągiem danych liczbowych będących wynikiem przybliżonego pomiaru wielkości opisujących zjawisko obserwowane w świecie realnym (a nie – w matematyce). Upewniwszy się, że niejednoznaczność wyniku, a więc *błąd pomiaru* ma charakter losowy – co potwierdzają np. opisane w poprzednim podrozdziale testy – możemy podjąć się wyznaczenia empirycznego, przybliżonego rozkładu obserwowanej zmiennej. Podstawy teoretyczne z reguły podpowiadają rodzinę rozkładów prawdopodobieństwa (jedno- lub wieloparametrową), w ramach której wskazanie „właściwego rozkładu” oznacza wyznaczenie najbardziej prawdopodobnych, oczekiwanych wartości parametrów.

Przykład 2.3.11. Jeśli ciąg zer i jedynek $x_1, \dots, x_n \in \{0, 1\}$ jest wartością próby losowej $X = (X_1, \dots, X_n)$ o rozkładzie $\mathbb{P}_\theta\{x\} = \theta^x(1-\theta)^{1-x}$, dla $x = 0, 1$, z nieznanym prawdopodobieństwem *sukcesu* θ , wówczas zgodnie z przykładem 2.1.18 miara $\mathbb{P}$ w produkcie $\Omega \equiv \{0, 1\}^n \times [0, 1]$ o gęstości

$$g_n(x, \theta) = \theta^{\sum x} (1 - \theta)^{n - \sum x},$$

wprowadza jednolity kontekst dla danych i parametrów rozkładu. Warunkowy rozkład prawdopodobieństwa zmiennej $\Theta: (x, \theta) \mapsto \theta$,

$$\mathbb{P}_{\Theta|X=x} = \mathbb{P}_{\Theta|\sum X_i=y} = \text{Beta}(y+1, n-y+1) \quad (2.83)$$

o gęstości $g_n(\theta|x) = (n+1) \binom{n}{y} \theta^y (1-\theta)^{n-y}$, gdzie $y = \sum_i x_i$, pozwala oceniać *wiarygodność* (ang. likelihood) dopasowania parametru $\theta \in [0, 1]$ do zaobserwowanych danych. Klasyczna terminologia wyróżnia dwa rywalizujące *optymalne* rozwiązania. Wartość θ_{mle} maksymalizuje wiarygodność, czyli gęstość rozkładu (2.83), zaś konkurencję stanowi wartość oczekiwana $\theta_e = \mathcal{E}(\Theta|X=x)$. W tym przypadku

$$\theta_{mle} = \frac{y}{n}, \quad \theta_e = \frac{y+1}{n+2}.$$

Liczba θ_e jest tu równa wartości maksymalizującej wiarygodność dla ciągu danych rozszerzonego o dwie dodatkowe „próby”: 0, 1. Można to traktować jako *wiedzę ekspercką* nakazującą umiarkowany sceptycyzm co do akceptacji skrajnych wartości parametru θ (prawdopodobieństwa sukcesu).

Przykład 2.3.12. Rozważmy ciągi liczb $x_1, \dots, x_n \geq 0$ będące próbkami losowymi z rozkładu jednostajnego na przedziale $[0, \theta]$ o nieznanej długości $\theta > 0$. Dla każdego θ miara o gęstości $g_\theta = \frac{1}{\theta} 1_{[0, \theta]}$ wyznacza w przestrzeni próbek $\mathbb{R}_+^n$ miarę probabilistyczną $\mathbb{P}_\theta^n$ o gęstości $g_n(x|\theta) = g_\theta(x_1) \cdots g_\theta(x_n)$ względem miary Lebesgue'a ℓ^n . Ten sam wzór określa gęstość miary borelowskiej $\mathbb{P}$ w przestrzeni $\Omega \equiv \mathbb{R}_+^n \times \mathbb{R}_+$,

$$g_n(x_1, \dots, x_n, \theta) = \theta^{-n} \cdot 1_{[0, \theta]}(\max_i x_i).$$

Miara $\mathbb{P}$ i projekcje X, Θ na $\mathbb{R}_+^n, \mathbb{R}_+$ spełniają założenia twierdzenia 1.3.30, przy czym miarą brzegową $\mathbb{P}_\Theta$ jest (jednostajna) miara Lebesgue'a, co interpretuje się jako „brak informacji” a priori dotyczących parametru θ . Drugi z rozkładów brzegowych $\mathbb{P}_X$ ma gęstość

$$g_n(x) = \int_{\max x}^{\infty} \theta^{-n} d\theta = \frac{1}{n-1} (\max x)^{-n+1},$$

dla $n > 1$, a zatem rozkład warunkowy $\mathbb{P}_{\Theta|X=x}$ o gęstości

$$g_n(\theta|x) = \begin{cases} (n-1)(\max x)^{n-1} \theta^{-n}, & \text{dla } \theta \geq \max x \\ 0, & \text{gdy } \theta < \max x, \end{cases}$$

jest tzw. *rozkładem Pareta*

$$\mathbb{P}_{\Theta|X=x} = \text{Pareto}(y, n-1) = \mathbb{P}_{\Theta|\max X=y}, \quad (2.84)$$

gdzie $y = \max_i x_i$. Zgodnie z przyjętą w poprzednim przykładzie notacją optymalne dopasowanie do próbki losowej wyróżnia wartości

$$\theta_{mle} = y \quad \text{ i } \quad \theta_e = \frac{n-1}{n-2} y > y, \quad \text{ dla } n > 2.$$

Jeśli rozszerzymy ciąg x o dodatkową „pozorną” obserwację $x_0 = 0$, otrzymamy gęstość $\tilde{g}_n(\theta|x) = n(\max x)^n / \theta^{n+1}$ rozkładu $\text{Pareto}(y, n)$ o wartości oczekiwanej $\tilde{\theta}_e = \frac{n}{n-1} y < \theta_e$. Wprowadzona modyfikacja odpowiada *wiedzy eksperckiej* dotyczącej możliwych wartości parametru θ i zapisanej w postaci wzoru

$$\tilde{g}_n(x, \theta) = g_n(x|\theta) \cdot \theta^{-1}$$

modyfikującego wyjściową miarę w przestrzeni Ω .

Ćwiczenie 2.3.13. Sprawdzić, że jeśli uzyskany wyżej rozkład warunkowy $\mathbb{P}_{\Theta|X=\xi}$ wykorzystać jako dostępną – dla danej próbki $\xi = (\xi_1, \dots, \xi_n)$ – informację na temat rozkładu parametru θ , wówczas miara o gęstości

$$\tilde{g}_m(x, \theta) = g_m(x|\theta) \cdot g_n(\theta|\xi) \quad \text{ w przestrzeni } \mathbb{R}_+^m \times \mathbb{R}_+$$

łączy informację z obu ciągów i prowadzi do rozkładu $\mathbb{P}_{\Theta|X=(\xi, x)}$ wyznaczonego przez próbkę o łącznej długości $n+m$.

Przykład 2.3.14. Jak wiadomo (ćwiczenie – jaka to zbieżność?) graniczną formą ciągu rozkładów dwumianowych $\text{Bin}(n, p_n)$ o stabilnej wartości oczekiwanej $np_n \rightarrow \lambda$ jest *rozkład Poissona* $\text{Pois}(\lambda)$. Ponieważ rozkłady dyskretne są absolutnie ciągłe względem miary zliczającej ρ , w tym przypadku na zbiorze $\mathbb{N}_o \equiv \{0\} \cup \mathbb{N}$, możemy mówić o gęstości równej $g_\lambda(i) = \frac{\lambda^i}{i!} e^{-\lambda}$, dla $i \in \mathbb{N}_o$. Jeśli zatem rozważamy ciągi liczb $x_1, \dots, x_n \in \mathbb{N}_o$ będące próbkami losowymi z rozkładu $\text{Pois}(\theta)$ o nieznannej wartości oczekiwanej $\theta > 0$, konfiguracją próbek i parametrów w przestrzeni $\Omega = \mathbb{N}_o^n \times \mathbb{R}_+$ zarządza miara borelowska $\mathbb{P}$ o gęstości

$$g_n(x_1, \dots, x_n, \theta) = \frac{\theta^{x_1 + \dots + x_n}}{x_1! \cdots x_n!} e^{-n\theta}$$

względem $\rho \otimes \ell$, przy czym $\mathbb{P}_\Theta = \ell_{\mathbb{R}_+}$ jest brzegowym rozkładem a’priori dla θ . Gęstość rozkładu brzegowego $\mathbb{P}_X$ wyznaczamy jako całkę

$$g_n(x) = \int_0^\infty g_n(x, \theta) d\theta = \frac{1}{x_1! \cdots x_n!} \int_0^\infty \left(\frac{t}{n}\right)^y e^{-t} \frac{dt}{n} = \frac{y! n^{-(y+1)}}{x_1! \cdots x_n!},$$

gdzie $y = \sum_i x_i$. Z twierdzenia 1.3.30 wynika zatem, że rozkład warunkowy $\mathbb{P}_{\Theta|X=x}$ ma gęstość

$$g_n(\theta|x) = \frac{(n\theta)^{x_1 + \dots + x_n}}{n(x_1 + \dots + x_n)!} e^{-n\theta}$$

i jest tzw. rozkładem *gamma*

$$\mathbb{P}_{\Theta|X=x} = \text{Gamma}(y+1, n) = \mathbb{P}_{\Theta|\Sigma X_i=y}, \quad (2.85)$$

gdzie $y = \sum_i x_i$. W tym przypadku

$$\theta_{mle} = \frac{y}{n}, \quad \theta_e = \frac{y+1}{n}.$$

Podobnie, jak w poprzednim przykładzie, wprowadzenie do gęstości $g_n(x, \theta)$ czynnika a’priori postaci $c\theta^\alpha e^{-m\theta}$ pozwala na kumulowanie informacji pobieranych z kolejnych próbek losowych.

Ćwiczenie 2.3.15. Rozkład *wykładniczy* $\text{Exp}(\theta) \equiv \text{Gamma}(1, \theta)$ ma gęstość $g_\theta(x) = \theta e^{-x\theta}$, dla $x \geq 0$. Sprawdzić, że dla próbek $x \in \mathbb{R}_+^n$ rozkład warunkowy $\mathbb{P}_{\Theta|X=x}$ jest jednym z rozkładów **Gamma** i jest równy $\mathbb{P}_{\Theta|T=T(x)}$ dla $T(x) \equiv \sum x$. Wyznaczyć optymalne wartości θ_{mle} i θ_e . Wykazać, że warunkowy rozkład prawdopodobieństwa $\mathbb{P}_{X|T=y}$ jest przeskalowaną miarą Lebesgue’a na $(n-1)$ -wymiarowych sympleksach $\{x \in \mathbb{R}_+^n; \sum_i x_i = y\}$, dla $y > 0$.

Wspólną cechą prezentowanych wyżej przykładów jest wyodrębnienie funkcji $T(X)$ nałożonej na próbę losową X , której wartości $T(x_1, \dots, x_n)$ są ściśle związane z przybliżoną (estymowaną) optymalną wartością nieznanego parametru θ . W kontekście parametru θ zmienna losowa T jest nazywana *statystyką*. W przykładzie 2.3.12 jest to maksimum $T(X) \equiv \max_i X_i$, a w pozostałych dwu (i w ćwiczeniu) – suma $T(X) \equiv \sum_i X_i$. W każdym z przykładów zależny od parametru rozkład n –elementowej próby losowej ma gęstość postaci $g(x|\theta) = g_\theta(T(x))h(x)$ względem miary jednostajnej (zliczającej lub Lebesgue’a), a więc gęstość $g_\theta \circ T$ względem pewnej niezależnej od θ miary σ –skończonej.

2.3.5 Statystyki dostateczne czyli dobre

Precyzyjny matematyczny opis tego, co i jak można wyznaczać w oparciu o dostępne dane, ma zaskakująco abstrakcyjną formę, a jej kulminację stanowi pojęcie *dostatecznej statystyki* i związana z nim grupa twierdzeń. Klasyczne twierdzenie (poniżej) formalizuje specyfikę statystyki T w postaci niezależnej od jakiegokolwiek miary w zbiorze parametrów.

Twierdzenie 2.3.14 (O faktoryzacji). *Niech $\mathcal{P} = \{P_\theta; \theta \in \Theta\}$ będzie rodziną borelowskich miar probabilistycznych na $\mathbb{R}^p$ ($p \in \mathbb{N}$) absolutnie ciągłych względem σ –skończonej miary μ ; niech X i T oznaczają, odpowiednio, identyczność na przestrzeni $(\mathbb{R}^p, \mathfrak{B}_{\mathbb{R}^p}, P_\theta)$ i zmienną losową $T: \mathbb{R}^p \rightarrow \mathbb{R}^m$ niezależną od wartości θ . Następujące warunki są równoważne:*

(i) *Rozkład P_X jest postaci*

$$P_\theta = (g_\theta \circ T) h \cdot \mu, \quad (2.86)$$

gdzie $h: \mathbb{R}^p \rightarrow \mathbb{R}_+$ i $g_\theta: \mathbb{R}^m \rightarrow \mathbb{R}_+$, dla $\theta \in \Theta$, są funkcjami borelowskimi.

(ii) *Rozkład warunkowy $P_{X|T}$ nie zależy od wartości parametru θ .*

Dowód podajemy w dalszej części bieżącego podrozdziału. Naturalną konsekwencją twierdzenia jest

Definicja 2.3.15. *Przy oznaczeniach twierdzenia 2.3.14 funkcję T nazywamy statystyką dostateczną dla rodziny miar $\mathcal{P} = \{P_\theta; \theta \in \Theta\}$, jeśli rozkład warunkowy $P_{X|T}$ nie zależy od wartości parametru θ .*

Uogólniając zależności opisane w poprzednim podrozdziale, proponujemy w tym miejscu pomocnicze, alternatywne twierdzenie, w którym przestrzeń parametrów jest $\mathbb{R}^q$ wraz z ustaloną miarą borelowską, co obejmuje także rodziny parametrów ograniczone do podzbioru w $\mathbb{R}^q$, dla pewnego

$q \in \mathbb{N}$. Symbol Θ rezerwujemy w tym przypadku dla projekcji $(x, \theta) \mapsto \theta$ określonej na produkcie $\mathbb{R}^p \times \mathbb{R}^q$.

Stwierdzenie 2.3.16 (Bayesowskie statystyki dostateczne). *Niech μ oraz λ będą dowolnymi σ -skończonymi miarami borelowskimi odpowiednio w $\mathbb{R}^p$ i $\mathbb{R}^q$, $p, q \in \mathbb{N}$. Załóżmy, że $\mathcal{P} = \{P_\theta; \theta \in \mathbb{R}^q\}$ jest rodziną borelowskich miar probabilistycznych na $\mathbb{R}^p$ taką, że wzór*

$$\mathbb{P}(A \times C) = \int_C P_\theta(A) d\lambda(\theta), \quad \text{dla } A \in \mathfrak{B}_{\mathbb{R}^p}, C \in \mathfrak{B}_{\mathbb{R}^q}, \quad (2.87)$$

określa σ -skończoną borelowską miarę na $\mathbb{R}^{p+q}$ absolutnie ciągłą względem miary produktowej $\mu \otimes \lambda$. Niech wreszcie X, Θ i T oznaczają, odpowiednio, projekcje na poszczególne czynniki $\mathbb{R}^{p+q} = \mathbb{R}^p \times \mathbb{R}^q$ oraz zmienną losową $T: \mathbb{R}^p \rightarrow \mathbb{R}^m$ niezależną od wartości $\theta \in \mathbb{R}^q$. Jeśli miara brzegowa

$$\mathbb{P}_{TX}(B) = \int P_\theta(T^{-1}B) d\lambda(\theta), \quad \text{dla } B \in \mathfrak{B}_{\mathbb{R}^m},$$

jest σ -skończona, wówczas T jest statystyką dostateczną dla rodziny $\mathcal{P}$ wtedy i tylko wtedy, gdy zachodzi równość rozkładów warunkowych

$$\mathbb{P}_{\Theta|X=x} = \mathbb{P}_{\Theta|TX=T(x)} \quad \text{dla } \mathbb{P}_X - (\text{pw}) x \in \mathbb{R}^p. \quad (2.88)$$

Dowód. Równość $\mathbb{P}_\Theta(C) = \int_C P_\theta(\mathbb{R}^p) d\lambda(\theta) = \lambda(C)$ implikuje σ -skończoność miary brzegowej $\mathbb{P}_\Theta = \lambda$. Jeśli przyjmiemy, że $\mathbb{P} = h \cdot (\mu \otimes \lambda)$ dla skończonej funkcji borelowskiej h na $\mathbb{R}^p \times \mathbb{R}^q$, wówczas rodzina $\mathcal{P}$ jest rozkładem warunkowym $\mathbb{P}_{X|\Theta}$, przy czym

$$P_\theta = \mathbb{P}_{X|\Theta=\theta} = h(\cdot, \theta)\mu, \quad \text{dla } (\text{pw}) \theta \in \mathbb{R}^q,$$

co wynika z tożsamości $\int_C P_\theta(A) d\lambda(\theta) = \int_C \int_A h(x, \theta) d\mu(x) d\lambda(\theta)$.

Analogicznie, równość $\mathbb{P}(A \times \mathbb{R}^q) = \int_A (\int h(x, \cdot) d\lambda) d\mu(x)$, dla $A \in \mathfrak{B}_{\mathbb{R}^p}$, wyznacza miarę brzegową

$$\mathbb{P}_X = H \cdot \mu, \quad \text{gdzie } H(x) = \int h(x, \cdot) d\lambda, \quad \text{dla } x \in \mathbb{R}^p,$$

a więc, zgodnie z twierdzeniem 1.3.30, $\mathbb{P}_{\Theta|X=x} = \frac{h(x, \cdot)}{H(x)} \cdot \lambda$, dla $\mathbb{P}_X - (\text{pw}) x \in \mathbb{R}^p$. Dla wyznaczenia rozkładu warunkowego względem statystyki T korzystając z lematu Dooba–Dynkina (stw. 1.3.18) otrzymujemy tożsamość

$$\mathbb{P}_{TX}(B) = \int_{T^{-1}B} H(x) d\mu(x) = \int_B G(y) d\mu_T(y), \quad \text{a zatem } \mathbb{P}_{TX} = G \cdot \mu_T,$$

dla pewnej funkcji borelowskiej $G: \mathbb{R}^m \rightarrow \mathbb{R}_+$, i analogicznie

$$\mathbb{P}(T^{-1}B \times C) = \int_{T^{-1}B \times C} h(x, \theta) d(\mu \otimes \lambda)(x, \theta) = \int_{B \times C} g(y, \theta) d(\mu_T \otimes \lambda)(y, \theta),$$

dla funkcji borelowskiej $g: \mathbb{R}^m \times \mathbb{R}^q \rightarrow \mathbb{R}_+$. Wynikający stąd opis rozkładu warunkowego $\mathbb{P}_{\Theta|TX=y} = \frac{g(y, \cdot)}{G(y)} \cdot \lambda$, dla $\mathbb{P}_{TX} - (\text{pw}) y \in \mathbb{R}^m$, pozwala zauważyć, że konsekwencją równości miar $\mathbb{P}_{\Theta|X=x} = \mathbb{P}_{\Theta|TX=T(x)}$ jest równość wyrażeń $h(x, \theta)/H(x) = g(T(x), \theta)/G(T(x))$ dla $\mu - (\text{pw}) x \in \mathbb{R}^p$ i $\lambda - (\text{pw}) \theta \in \mathbb{R}^q$, czyli

$$h(x, \theta) = g(T(x), \theta) \frac{H(x)}{G(T(x))}, \quad \text{dla } \mu \otimes \lambda - (\text{pw}) (x, \theta) \in \mathbb{R}^{p+q},$$

a to oznacza dostateczność statystyki T .

Dla dowodu implikacji odwrotnej założymy, że rodzina $\mathcal{P}$ spełnia warunek faktoryzacji (2.86) w postaci

$$P_\theta = (g_\theta \circ T) \cdot \tilde{\mu}, \quad \text{dla } \tilde{\mu} \equiv \tilde{h} \cdot \mu,$$

przy czym $\tilde{h}$ i $g_\theta, \theta \in \mathbb{R}^q$, są nieujemnymi funkcjami borelowskimi. Założymy dodatkowo, że wzór $\tilde{g}(y, \theta) \equiv g_\theta(y)$ definiuje funkcję borelowską na $\mathbb{R}^p \times \mathbb{R}^q$. W tym przypadku z równości (2.87) wynika tożsamość

$$\begin{aligned} \mathbb{P}(A \times C) &= \int_C \left(\int_A g_\theta(T(x)) d\tilde{\mu}(x) \right) d\lambda(\theta) \\ &= \int_C \int_A \tilde{g}(T(x), \theta) d\tilde{\mu}(x) d\lambda(\theta) = \int_{A \times C} h d(\tilde{\mu} \otimes \lambda), \end{aligned}$$

a więc funkcja $h(x, \theta) \equiv \tilde{g}(T(x), \theta)$ jest gęstością miary $\mathbb{P} \ll \tilde{\mu} \otimes \lambda$. Rozważania ogólne – z pierwszej części dowodu – prowadzą w tym przypadku do tożsamości

$$\mathbb{P}(T^{-1}B \times C) = \int_{B \times C} \tilde{g}(y, \theta) d(\tilde{\mu}_T \otimes \lambda)(y, \theta) \quad \text{oraz}$$

$$H(x) = \int \tilde{g}(T(x), \cdot) d\lambda = \tilde{G}(T(x)), \quad \text{gdzie } \tilde{G}(y) = \int \tilde{g}(y, \cdot) d\lambda,$$

i wobec tego $\mathbb{P}_{\Theta|X=x} = \frac{\tilde{g}(T(x))}{\tilde{G}(T(x))} \cdot \lambda = \mathbb{P}_{\Theta|TX=T(x)}$, dla $x \in \mathbb{R}^p$.

Aby usunąć potencjalny brak mierzalności g_θ względem θ , zauważamy, iż równość (2.87) kryje w sobie założenie, że dla każdego zbioru borelowskiego $A \subset \mathbb{R}^p$ funkcja $\theta \mapsto P_\theta(A)$ jest mierzalna (borelowska). Korzystając z monotonicznej ciągłości całki i stosując twierdzenie Sierpińskiego o monotonicznych klasach zbiorów (tw. 1.1.14), rozszerzamy (2.87) do postaci

$$\mathbb{P}(M) = \int P_\theta(M^\theta) d\lambda(\theta), \quad \text{dla } M \in \mathfrak{B}_{\mathbb{R}^{p+q}},$$

gdzie $M^\theta = \{x; (x, \theta) \in M\}$.

Wykażemy w pierwszej kolejności absolutną ciągłość $\mathbb{P} \ll \tilde{\mu} \otimes \lambda$. Istotnie, jeśli M jest dowolnym zbiorem, dla którego

$$(\tilde{\mu} \otimes \lambda)(M) = \int \tilde{\mu}(M^\theta) d\lambda(\theta) = 0,$$

wówczas zbiór $O = \{\theta \in \mathbb{R}^q; \tilde{\mu}(M^\theta) = 0\}$ ma dopełnienie miary 0, a dla $o \in O$ i dla każdego θ (w szczególności dla $\theta = o$) $P_\theta(M^o) = 0$. Stąd

$$\mathbb{P}(M) = \int_O P_o(M^o) d\lambda(o) = 0,$$

a zatem $\mathbb{P} = h \cdot (\tilde{\mu} \otimes \lambda)$, dla pewnej funkcji borelowskiej $h: \mathbb{R}^p \times \mathbb{R}^q \rightarrow \mathbb{R}_+$. Przy oznaczeniach ze wstępnej części dowodu mamy zatem tożsamości

$$\mathbb{P}(T^{-1}B \times C) = \int_C \int_B g(y, \theta) d\tilde{\mu}_T(y) d\lambda(\theta),$$

dla funkcji borelowskiej g na $\mathbb{R}^m \times \mathbb{R}^q$ i jednocześnie

$$\mathbb{P}(T^{-1}B \times C) = \int_C \left(\int_{T^{-1}B} g_\theta(T(x)) d\tilde{\mu}(x) \right) d\lambda(\theta),$$

dla dowolnych zbiorów $B \in \mathfrak{B}_{\mathbb{R}^m}$ i $C \in \mathfrak{B}_{\mathbb{R}^q}$. Wynika stąd zależność

$$\int_B g(y, \theta) d\tilde{\mu}_T(y) = \int_{T^{-1}B} g_\theta(T(x)) d\tilde{\mu}(x) = \int_B g_\theta(y) d\tilde{\mu}_T(y),$$

dla λ -(pw) $\theta \in \mathbb{R}^q$, czyli także $g(y, \theta) = g_\theta(y)$, dla $\tilde{\mu}_T$ -(pw) $y \in \mathbb{R}^m$ i w końcu

$$g(T(x), \theta) = g_\theta(T(x)), \quad \text{dla } \tilde{\mu}\text{-(pw) } x \in \mathbb{R}^p.$$

Ostatnia z równości implikuje równość miar $P_\theta = g(\cdot, \theta) \cdot \tilde{\mu}$, dla prawie wszystkich wartości parametru θ . $\square$

Dowód twierdzenia 2.3.14. Bez zmniejszenia ogólności możemy założyć (por. przykład 1.3.4), że μ jest miarą skończoną. Spełnione są wówczas założenia sformułowane we wniosku 1.3.28, a wzór (1.76) zapewnia równość

$$\int_{T^{-1}B} f(x, T(x)) \mu(dx) = \int_B \left(\int f(x, y) \mu_{X|T=y}(dx) \right) \mu_T(dy), \quad \text{dla } B \in \mathfrak{B}_{\mathbb{R}^m},$$

dla każdej nieujemnej funkcji borelowskiej f na $\mathbb{R}^{p+m}$.

(i) $\Rightarrow$ (ii) Biorąc za f gęstość miary P_θ , otrzymujemy

$$\int_{T^{-1}B} g_\theta(T(x)) h(x) \mu(dx) = \int_B \left(\int g_\theta(y) h(x) \mu_{X|T=y}(dx) \right) \mu_T(dy), \quad \text{czyli}$$

$$P_T(B) \equiv P_\theta(T^{-1}B) = \int_B g_\theta(y) w(y) \mu_T(dy),$$

gdzie $w(y) = \int h d\mu_{X|T=y}$ jest funkcją borelowską pw. skończoną względem miary μ_T . Zatem $P_T = (g_\theta w) \cdot \mu_T$, a rozkład warunkowy $P_{X|T}$ definiuje warunek (2.33) w postaci

$$P_\theta\{X \in A, T \in B\} = \int_B P_{X|T=y}(A) P_T(dy) = \int_B g_\theta(y) w(y) P_{X|T=y}(A) \mu_T(dy),$$

dla borelowskich $A \times B \subset \mathbb{R}^p \times \mathbb{R}^m$, przy czym wartość miary P_θ wynosi

$$\int_{T^{-1}B} g_\theta(T(x)) (1_A h)(x) \mu(dx) = \int_B g_\theta(y) \left(\int_A h d\mu_{X|T=y} \right) \mu_T(dy).$$

Rozkładem warunkowym $P_{X|T=y}$ jest wobec tego każda borelowska rodzina miar na $\mathfrak{B}_{\mathbb{R}^p} \times \mathbb{R}^m$, dla której

$$w(y) P_{X|T=y} = h \cdot \mu_{X|T=y}, \quad \text{dla } y \in \mathbb{R}^m.$$

Jeśli zatem $w(y) > 0$, miara probabilistyczna $P_{X|T=y}$ jest efektem unormowania miary $h \cdot \mu_{X|T=y}$. Przypadek $w(y) = 0$ jest nieistotny, gdyż $P_T\{w=0\} = 0$.

(ii) $\Rightarrow$ (i) Istnienie rozkładu warunkowego $P_{X|T}$ nie wymaga dodatkowych założeń – w przestrzeni probabilistycznej $(\mathbb{R}^p, \mathfrak{B}_{\mathbb{R}^p}, P_\theta)$. Z twierdzenia 1.3.12 o zdominowanej rodzinie miar wynika istnienie zbioru borelowskiego $Z \subset \mathbb{R}^p$, dla którego zachodzi równoważność

$$(\forall \theta \in \Theta \ P_\theta(N) = 0) \iff \tilde{\mu}(N) \equiv \mu(N \cap Z) = 0, \quad \text{dla } N \in \mathfrak{B}_{\mathbb{R}^p}, \quad (2.89)$$

gdzie $\tilde{\mu} = 1_Z \cdot \mu$ jest miarą dominującą równoważną z rodziną $P_\theta, \theta \in \Theta$. Z absolutnej ciągłości $P_\theta \ll \tilde{\mu}$ wynika analogiczna zależność dla miar brzegowych, a więc $P_T = g_\theta \cdot \tilde{\mu}_T$, dla zależnej od θ funkcji skończonej $g_\theta \geq 0$. Całka iterowana

$$\hat{P}(A) = \int \left(\int 1_A(x) dP_{X|T=y}(x) \right) d\tilde{\mu}_T(y), \quad \text{dla } A \in \mathfrak{B}_{\mathbb{R}^p},$$

wyznacza niezależną od $\theta \in \Theta$ skończoną miarę borelowską na $\mathbb{R}^p$ taką, że

$$\int f(x) d\hat{P} = \int \left(\int f(x) dP_{X|T=y}(x) \right) d\mu_T(y),$$

dla dowolnej nieujemnej funkcji borelowskiej f na $\mathbb{R}^p$, co sprawdzamy korzystając z monotonicznej ciągłości całki Lebesgue'a. W szczególności,

$$\int_A g_\theta(T(x)) d\hat{P}(x) = \int \int_A g_\theta(T(x)) dP_{X|T=y}(x) d\tilde{\mu}_T(y), \quad \text{dla } A \in \mathfrak{B}_{\mathbb{R}^p}.$$

Z własności (2.35) dla niezależnego od θ zbioru $M = \{y; P_{X|T=y}\{T \neq y\} > 0\}$ wynikają równości $P_\theta\{T \in M\} = P_T(M) = 0$ i $\tilde{\mu}_T(M) = \tilde{\mu}\{T \in M\} = 0$. Dla rozważanej wyżej całki oznacza to równość

$$\begin{aligned}\int_A g_\theta(T(x)) d\hat{P}(x) &= \int \int_A g_\theta(y) dP_{X|T=y}(x) d\tilde{\mu}_T(y) \\ &= \int P_{X|T=y}(A) dP_T(y) = P_\theta(A), \quad \text{dla } A \in \mathfrak{B}_{\mathbb{R}^p},\end{aligned}$$

czyli $P_\theta = (g_\theta \circ T) \cdot \hat{P}$. Pozostaje sprawdzić absolutną ciągłość $\hat{P} \ll \mu$. Niech zatem $N \in \mathfrak{B}_{\mathbb{R}^p}$ będzie dowolnym zbiorem o mierze $\mu(N) = 0$, co implikuje równość

$$0 = P_\theta(N) = \int P_{X|T=y}(N) dP_T(y), \quad \text{dla } \theta \in \Theta,$$

a więc także $P_T(\hat{N}) = 0$ dla zbioru $\hat{N} = \{y; P_{X|T=y}(N) > 0\}$ niezależnego od θ . Z własności (2.89) otrzymujemy $\tilde{\mu}_T(\hat{N}) = 0$ i $\hat{P}(N) = 0$. $\square$

Przykład 2.3.16. Dla rozważanych w przykładzie 2.3.12 próbek z rozkładu jednostajnego rozkład warunkowy $P_{X|T=y}$, gdzie $T(x) \equiv \max_i x_i$, redukuje się do miary na powierzchni

$$\{x; T(x)=y\} = \bigcup_{k \leq n} \{x; x_k = \max x = y\}, \quad \text{dla } 0 < y \leq \theta,$$

a naturalnym kandydatem na miarę probabilistyczną jest odpowiednio unormowana miara Lebesgue'a w wymiarze $n-1$. Przyjmując na każdym ze składników sumy

$$\bigcup_k \{k\} \times [0, y]^{n-1} \ni (k, x) \xrightarrow{\iota_k} (\dots, x_{k-1}, y, x_k, \dots) \in \{x; T(x)=y\}$$

miarę $\frac{1}{ny^{n-1}} \ell^{n-1}$, sprawdzamy równość (2.33) dla miary warunkowej

$$P_{X|T=y}\{X \leq t\} = \sum_{k \leq n} \frac{1}{ny^n} \ell^{n-1}(\iota_k^{-1}\{X \leq t\}) \quad (2.90)$$

oraz miary brzegowej P_T o dystrybuancie $P_\theta^n\{T \leq y\} = \min((y/\theta)^n, 1)$ i gęstości ny^{n-1}/θ^n , dla $0 \leq y \leq \theta$. Dla $t \in [0, \theta]^n$ i $z \leq \theta$ porządkujemy indeksy w ten sposób, by

$$0 \leq t_{i_1} \leq \dots \leq t_{i_k} < z \leq t_{i_{k+1}} \leq \dots \leq \theta,$$

co pozwala na wyznaczenie poszczególnych składników sumy

$$\begin{aligned}\theta^n \left(\int_0^{t_{i_1}} + \dots + \int_{t_{i_k}}^z \right) P_{X|T=y}\{X \leq t\} P_T(dy) \\ &= \int_0^{t_{i_1}} ny^{n-1} dy + \int_{t_{i_1}}^{t_{i_2}} (n-1)t_{i_1}y^{n-2} dy + \int_{t_{i_2}}^{t_{i_3}} (n-2)t_{i_1}t_{i_2}y^{n-3} dy \\ &\quad + \int_{t_{i_k}}^z (n-k)t_{i_1} \dots t_{i_k}y^{n-k-1} dy \\ &= t_{i_1}^n + [t_{i_1}y^{n-1}]_{t_{i_1}}^{t_{i_2}} + [t_{i_1}t_{i_2}y^{n-2}]_{t_{i_2}}^{t_{i_3}} + \dots + [t_{i_1} \dots t_{i_k}y^{n-k}]_{t_{i_k}}^z \\ &= t_{i_1} \dots t_{i_k} z^{n-k} = \theta^n P_\theta^n\{X \leq t, T \leq z\}.\end{aligned}$$

Ćwiczenie 2.3.17. Sprawdzić, że w przykładzie 2.3.11 zachodzi równość

$$P_{X|T=y}\{x\} = \begin{cases} \binom{n}{y}^{-1}, & \text{gdy } T(x)=y, \\ 0, & \text{gdy } T(x) \neq y. \end{cases}$$

Ta sama statystyka $T(x) \equiv \sum_i x_i$ prowadzi w przykładzie 2.3.14 do rozkładu

$$P_{X|T=y}\{x\} = \begin{cases} \frac{y!n^{-y}}{x_1! \cdots x_n!}, & \text{gdy } T(x)=y, \\ 0, & \text{gdy } T(x) \neq y. \end{cases}$$

W każdym z przykładów wyznaczyć wartość oczekiwaną $\mathcal{E}_\theta(TX)$. Wielkość ta (zależna tylko od θ) może być przybliżana za pomocą danych empirycznych – próbek losowych zmiennej TX – to kolejna, trzecia już metoda estymacji nieznanej wartości parametru.

Uwaga 2.3.18. Niezależny od parametru θ znany rozkład warunkowy $\mathbb{P}_{X|T}$ możemy symulować poprzez odpowiednią modyfikację ciągu liczb pseudolosowych (podr. 2.3.2). W połączeniu z obserwowanym empirycznym rozkładem *statystyki dostatecznej* T można zatem wygenerować (pseudo-)próbki losowe o rozkładzie zgodnym z rozkładem zmiennej $X = (X_1, \dots, X_n)$. Mimo nieznajomości faktycznej wartości parametru utworzone w ten sposób „fałszywe dane” niosą tę samą informację o θ co dane oryginalne.

2.4 Rozkład normalny

2.4.1 Skąd TO się wzięło?

Znany wszystkim(?) *rozkład normalny* nie pojawił się do tej pory w prezentowanym tekście – został jedynie wspomniany w uwadze 2.3.8 jako graniczna postać opisującego łączną liczbę przebiegów rozkładu Walda–Wolfowitza.

Przy oznaczeniach przyjętych w podrozdziale 2.2, dla niezależnych zmiennych losowych $X_n, n \in \mathbb{N}$, o rozkładzie dyskretnym (2.49) uwzględnienie normalizującego czynnika $\frac{1}{\sqrt{n}}$ w równości $\text{Var}(X_1 + \dots + X_n) = n$ prowadzi do ciągu zmiennych o wartości oczekiwanej 0 i stałej wariancji równej 1.

Twierdzenie 2.4.1 (de Moivre’a – Laplace’a). *Ciąg dystrybuant $(F_n)_{n \in \mathbb{N}}$ zmiennych losowych $Y_n = \frac{1}{\sqrt{n}}(X_1 + \dots + X_n)$, dla $n \in \mathbb{N}$, jest jednostajnie zbieżny do funkcji*

$$\Phi: \mathbb{R} \ni t \mapsto \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{x^2}{2}} dx \in (0, 1). \quad (2.91)$$

Definicja 2.4.2. *Miarę probabilistyczną na prostej $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$, o dystrybucie opisanej wzorem całkowym (2.91), nazywamy standardowym rozkładem normalnym lub rozkładem Gaussa – i oznaczamy symbolem $\mathcal{N}(0, 1)$.*

Uwaga 2.4.1. Zbieżność ciągu zmiennych losowych czy dystrybuant nie gwarantuje zbieżności ciągu wartości oczekiwanych, czy też wariancji rozkładów. Bezpośredni rachunek pozwala sprawdzić, że graniczny rozkład normalny ma wartość oczekiwaną 0 i wariancję 1. Jeśli zmienna losowa N ma rozkład $\mathcal{N}(0, 1)$, to dla dowolnych $m \in \mathbb{R}$ i $\sigma \neq 0$ rozkład $\mathcal{N}(m, \sigma^2)$ zmiennej $\sigma N + m$ nazywamy ogólnym 1-wymiarowym *rozkładem normalnym o wartości oczekiwanej m i wariancji σ^2* . Miara $\mathcal{N}(m, \sigma^2)$ jest rozkładem ciągłym o gęstości $\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$ dla $x \in \mathbb{R}$. Funkcje stałe będziemy uważać za zmienne losowe o *dyskretnym* rozkładzie normalnym i zerowej wariancji.

Dowód twierdzenia 2.4.1. Suma parzystej liczby zmiennych $X_i, i \leq 2n$, przyjmuje wartości całkowite będące liczbami parzystymi, przy czym

$$\mathbb{P}\{X_1 + \dots + X_{2n} = 2k\} = \binom{2n}{n+k} \cdot 2^{-2n}, \quad \text{dla } |k| \leq n, \quad (2.92)$$

a zatem

$$F_{2n}(t) = \mathbb{P}\{Y_{2n} \leq t\} = \sum_{k=-n}^{\lfloor t\sqrt{n/2} \rfloor} \binom{2n}{n+k} \cdot 2^{-2n},$$

gdzie sumujemy składniki spełniające nierówność $\frac{2k}{\sqrt{2n}} \leq t$. Niech G_n będzie funkcją ciągłą zgodną z F_{2n} w punktach $t_k^n = \frac{2k}{\sqrt{2n}}$ dla $k = -n-1, \dots, n$, liniową w każdym z przedziałów $[t_{k-1}^n, t_k^n]$, $k = -n, \dots, n$, i taką że

$$G_n(t) = \begin{cases} 0 & \text{dla } t \leq t_{-n-1}^n, \\ 1 & \text{dla } t \geq t_n^n. \end{cases}$$

Zachodzą równości

$$G'_n|_{(t_{k-1}^n, t_k^n)} = \frac{G(t_k^n) - G(t_{k-1}^n)}{t_k^n - t_{k-1}^n} = \binom{2n}{n-k} 2^{-2n} \sqrt{\frac{n}{2}}$$

dla $|k| \leq n$. Aby uzyskać symetrię i wykorzystać ją do uproszczenia obliczeń, rozważmy funkcje ciągłe $H_n(x) = G_n(x - \frac{1}{\sqrt{2n}})$ dla $x \in \mathbb{R}$ i $n \in \mathbb{N}$, o własności

$$H'_n(x) = \begin{cases} \binom{2n}{n-k} 2^{-2n} \sqrt{\frac{n}{2}} & \text{dla } x \in (\frac{2k-1}{\sqrt{2n}}, \frac{2k+1}{\sqrt{2n}}), k = 0, \pm 1, \dots, \pm n, \\ 0 & \text{dla } x < -\frac{2n+1}{\sqrt{2n}} \text{ i dla } x > \frac{2n+1}{\sqrt{2n}}, \end{cases} \quad (2.93)$$

a więc także $H_n(x) + H_n(-x) = \text{const} = 1$ i $H_n(0) = \frac{1}{2}$ dla $n \in \mathbb{N}$. Sprawdzamy, że ciąg pochodnych $(H'_n)_{n \in \mathbb{N}}$, określony poza przeliczalnym zbiorem

$N = \{\pm \frac{2k+1}{\sqrt{2n}}; n \in \mathbb{N}, k=0, \dots, n\}$, spełnia lokalnie założenia twierdzenia Lebesgue'a o zmajoryzowanej zbieżności i jest zbieżny do funkcji podcałkowej we wzorze (2.91). Istotnie, dla dowolnego ustalonego $x \in \mathbb{R} \setminus N$ i dla $n > \frac{x^2}{2}$ istnieje liczba całkowita k_n taka, że $|k_n| \leq n$ i $k_n - \frac{1}{2} < x\sqrt{\frac{n}{2}} < k_n + \frac{1}{2}$, co w szczególności implikuje zbieżność do x ciągu $x_n = \frac{2k_n}{\sqrt{2n}}$. Wyznaczone na tej podstawie wartości pochodnej $H'_n(x)$ dla $n > \frac{x^2}{2}$, możemy przekształcić, korzystając ze wzoru Stirlinga (por. [1])

$$n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n} \cdot e^{\theta_n/12n}, \quad \theta_n \in (0, 1). \quad (2.94)$$

Jeśli $|k_n| < n$, a więc dla $x \in (-\frac{2n-1}{\sqrt{2n}}, \frac{2n-1}{\sqrt{2n}})$, otrzymujemy

$$\begin{aligned} H'_n(x) &= \binom{2n}{n-k_n} 2^{-2n} \sqrt{\frac{n}{2}} = \frac{(2n)!}{(n-k_n)!(n+k_n)! 2^{2n}} \sqrt{\frac{n}{2}} \\ &= \left(\frac{2n}{e}\right)^{2n} \left(\frac{n-k_n}{e}\right)^{-(n-k_n)} \left(\frac{n+k_n}{e}\right)^{-(n+k_n)} \frac{\sqrt{2\pi \cdot 2n} \sqrt{n/2}}{\sqrt{2\pi(n-k_n)} \sqrt{2\pi(n+k_n)}} \frac{c_n}{2^{2n}}, \end{aligned}$$

gdzie c_n jest iloczynem potęg liczby e o wykładnikach zawierających θ_n, θ_{n-k_n} i θ_{n+k_n} . Dla $n \rightarrow \infty$, przy ustalonym x , z rozbieżności każdego z ciągów $n \pm k_n$ do ∞ wynika zbieżność $c_n \rightarrow 1$. Odpowiednie pogrupowanie czynników prowadzi do równości

$$\begin{aligned} H'_n(x) &= \frac{c_n}{\sqrt{2\pi}} \cdot \left(1 - \frac{k_n^2}{n^2}\right)^{-n-\frac{1}{2}} \left(1 - \frac{k_n}{n}\right)^{k_n} \left(1 + \frac{k_n}{n}\right)^{-k_n} \\ &= \frac{c_n}{\sqrt{2\pi}} \cdot \left(1 - \frac{x_n^2}{2n}\right)^{-n-\frac{1}{2}} \left(1 - \frac{x_n^2}{2k_n}\right)^{k_n} \left(1 + \frac{x_n^2}{2k_n}\right)^{-k_n}, \end{aligned}$$

gdzie $x_n \rightarrow x$. Zatem

$$\lim_{n \rightarrow \infty} H'_n(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{x^2}{2}} e^{-\frac{x^2}{2}} e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

przy czym zbieżność każdego z czynników wynika ze sformułowanego niżej lematu 2.4.3(ii). Dla oszacowania z góry, w postaci

$$\begin{aligned} \sqrt{\left(1 - \frac{k_n^2}{n^2}\right)} H'_n(x) &= \frac{c_n}{\sqrt{2\pi}} \cdot \left(1 - \frac{k_n}{n}\right)^{-n+k_n} \left(1 + \frac{k_n}{n}\right)^{-n-k_n} \\ &= \frac{c_n}{\sqrt{2\pi}} \cdot \left(1 + \frac{k_n}{n-k_n}\right)^{n-k_n} \left(1 - \frac{k_n}{n+k_n}\right)^{n+k_n} \\ &\leq \frac{c_n}{\sqrt{2\pi}} \cdot e^{k_n} e^{-k_n} = \frac{c_n}{\sqrt{2\pi}} < c := 1.09, \end{aligned}$$

korzystamy z lematu 2.4.3(i) i nierówności $c_n < e^{1/12}$. Stąd

$$0 \leq H'_n(x) \leq c \left(1 - \frac{k_n^2}{n^2}\right)^{-\frac{1}{2}} \leq c \left(1 - \left(\frac{|x|}{\sqrt{2n}} + \frac{1}{2n}\right)^2\right)^{-\frac{1}{2}} \quad (2.95)$$

dla $x \in (-\frac{2n-1}{\sqrt{2n}}, \frac{2n-1}{\sqrt{2n}})$. Ponieważ funkcje występujące po prawej stronie nierówności (2.95) są całkowalne i tworzą ciąg malejący o rozszerzających się dziedzinach, w każdym z przedziałów $(-\frac{2m-1}{\sqrt{2m}}, \frac{2m-1}{\sqrt{2m}})$, $m \in \mathbb{N}$, otrzymaliśmy oszacowanie

$$0 \leq H'_n(x) \leq c \left(1 - \left(\frac{|x|}{\sqrt{2m}} + \frac{1}{2m}\right)^2\right)^{-\frac{1}{2}} \quad \text{dla } n \geq m.$$

Niech $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, dla $x \in \mathbb{R}$, oraz $\Phi(t) = \int_{-\infty}^t \varphi(x) dx$, dla $t \in \mathbb{R}$, oznaczają odpowiednio gęstość i dystrybuantę rozkładu normalnego. Dla dowolnego $t \in \mathbb{R}$, z twierdzenia Lebesgue'a wynika zbieżność ciągu całek

$$\int_0^t H'_n(x) dx \rightarrow \int_0^t \varphi(x) dx.$$

Ponieważ $\int_{-\infty}^0 H'_n(x) dx = \frac{1}{2} = \Phi(0)$ dla $n \in \mathbb{N}$, wnioskujemy stąd równość

$$\lim_{n \rightarrow \infty} H_n(t) = \Phi(t).$$

Dla dowodu, że jest to zbieżność jednostajna, ustalmy dowolne $\varepsilon \in (0, \frac{1}{2})$; niech $a < 0$ i $b = -a > 0$ będą liczbami, dla których $\Phi(a) = \varepsilon = 1 - \Phi(b)$. Jeśli przyjmiemy za $N \in \mathbb{N}$ liczbę taką, że dla $n \geq N$ zachodzą nierówności $|H_n(a) - \Phi(a)| < \varepsilon$ i $\int_a^b |H'_n(x) - \varphi(x)| dx < \varepsilon$, to z monotoniczności całek otrzymujemy

$$-\varepsilon < -\Phi(t) \leq H_n(t) - \Phi(t) < H_n(t) \leq H_n(a) < 2\varepsilon,$$

czyli $|H_n(t) - \Phi(t)| < 2\varepsilon$ dla $t < a$ (i dla $t > b$), oraz

$$\left| \int_a^t H'_n(x) dx - \int_a^t \varphi(x) dx \right| \leq \int_a^t |H'_n(x) - \varphi(x)| dx < \varepsilon,$$

a zatem $|H_n(t) - \Phi(t)| < |H_n(a) - \Phi(a)| + \varepsilon < 2\varepsilon$ dla $t \in [a, b]$.

Pozostaje wykazać, że jednostajna zbieżność $H_n \rightrightarrows \Phi$ przenosi się na ciąg $(F_n)_{n \in \mathbb{N}}$. Bezpośrednio z definicji obu funkcji wynikają nierówności

$$F_{2n}(t_{k-1}^n) \leq H_n(t) \leq F_{2n}(t_k^n) \quad \text{dla } t \in \left(\frac{2k-1}{\sqrt{2n}}, \frac{2k+1}{\sqrt{2n}}\right), \quad -n \leq k \leq n,$$

przy czym

$$F_{2n}(t) = \begin{cases} F_{2n}(t_{k-1}^n) & \text{dla } t \in \left(\frac{2k-1}{\sqrt{2n}}, t_k^n\right) \\ F_{2n}(t_k^n) & \text{dla } t \in [t_k^n, \frac{2k+1}{\sqrt{2n}}). \end{cases}$$

Stąd

$$|H_n(t) - F_{2n}(t)| \leq F_{2n}(t_k^n) - F_{2n}(t_{k-1}^n) \leq \binom{2n}{n} 2^{-2n} = \sqrt{\frac{2}{n}} H'_n(0),$$

co implikuje jednostajną zbieżność $F_{2n} \Rightarrow \Phi$. Zbieżność wyrazów ciągu o indeksach nieparzystych wyprowadzamy z tożsamości

$$\sqrt{\frac{n+1}{n}}Y_{n+1} = Y_n + \frac{1}{\sqrt{n}}X_{n+1},$$

z której wynikają oszacowania

$$F_{2n}\left(\sqrt{\frac{2n+1}{2n}}t - \frac{1}{\sqrt{2n}}\right) \leq F_{2n+1}(t) \leq F_{2n}\left(\sqrt{\frac{2n+1}{2n}}t + \frac{1}{\sqrt{2n}}\right).$$

Ponieważ Φ jest funkcją jednostajnie ciągłą, otrzymujemy $F_{2n+1} \Rightarrow \Phi$. $\square$

Wykorzystany w dowodzie lemat odwołuje się do elementarnych metod Analizy 1 (por. [1]):

Lemat 2.4.3. (i) $(1 + \frac{x}{a})^a \leq e^x$ dla $x > -a$, $a > 0$.

(ii) Jeśli $a_n \rightarrow \infty$, $b_n \rightarrow 0$ i $a_n b_n \rightarrow x$, to $(1 + b_n)^{a_n} \rightarrow e^x$. $\square$

Możemy w tym miejscu wrócić do prezentowanego w podrozdziale 2.3.3 rozkładu $WW(n_1, n_2)$ i uzupełnić badaną rodzinę testów losowości o bardziej zaawansowaną, graniczną postać.

Twierdzenie 2.4.4 (Wald–Wolfowitz, 1940). *Jeśli ciągi liczb naturalnych $(n_1, n_2, n = n_1 + n_2)$ dążą do nieskończoności w ten sposób, że*

$$\left(\frac{n_1}{n}, \frac{n_2}{n}\right) \rightarrow (\lambda, 1 - \lambda), \quad \text{dla pewnego } \lambda \in (0, 1),$$

wówczas dla dowolnej liczby rzeczywistej t zachodzi zbieżność

$$\lim_{n \rightarrow \infty} \mathbb{P}\left\{R \leq 2n\lambda(1-\lambda)\left(1 + \frac{t}{\sqrt{n}}\right)\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-x^2/2} dx,$$

przy czym zbieżność jest jednostajna względem $t \in \mathbb{R}$.

Dowód. Rozkład $\mathbb{P}_R$ zmiennej $R = R_{n_1 n_2}$ opisują wzory

$$\mathbb{P}_R\{2k\} = \frac{2k^2(n_1!)^2(n_2!)^2}{n_1 n_2 (k!)^2 (n_1 - k)! (n_2 - k)! n!}, \quad \mathbb{P}_R\{2k+1\} = \frac{n-2k}{2k} \mathbb{P}_R\{2k\},$$

dla $k \in \mathbb{N}$. Po zastosowaniu wzoru Stirlinga (2.94) otrzymujemy

$$\begin{aligned} \mathbb{P}_R\{2k\} &= \sqrt{\frac{2}{\pi n}} \left(\frac{(n_1 - k)(n_2 - k)}{k^2}\right)^{k-\frac{1}{2}} \left(\frac{n_1^2}{(n_1 - k)n}\right)^{n_1} \left(\frac{n_2^2}{(n_2 - k)n}\right)^{n_2} c_n \\ &= \sqrt{\frac{2}{\pi n}} \left(1 + \frac{n_1 n_2 - kn}{k^2}\right)^{k-\frac{1}{2}} \left(1 + \frac{n_1 n_2 - kn}{n_1^2}\right)^{-n_1} \left(1 + \frac{n_1 n_2 - kn}{n_2^2}\right)^{-n_2} c_n, \end{aligned}$$

gdzie $c_n \rightarrow 0$. Skalowanie zmiennej R uwzględniające parametry $\mathcal{E}(R) \approx \frac{2n_1n_2}{n}$ i $\text{Var}(R) \approx \frac{4n_1^2n_2^2}{n^3}$ wprowadza zmienną $Z_n = \sqrt{n}(\frac{n}{2n_1n_2}R - 1)$ oraz ciąg węzłów $t_k \in \mathbb{R}$ zdefiniowanych poprzez równości

$$\frac{t_k}{\sqrt{n}} = \frac{kn - n_1n_2}{n_1n_2}, \quad \text{dla } k \in \mathbb{N},$$

przy czym $\mathbb{P}\{Z \leq t_k\} = \mathbb{P}\{R \leq 2k\}$. Jeśli przyjąć, jak w dowodzie tw. 2.4.1, że G jest kawałkami liniową aproksymacją dystrybucji F_Z o wartościach w węzłach

$$G(t_k) = \sum_{r \leq 2k+1} \mathbb{P}_R(r) = F_Z(t_k) + \mathbb{P}_R\{2k+1\},$$

wówczas pochodna G' jest stała pomiędzy kolejnymi węzłami i równa

$$G'|_{(t_{k-1}, t_k)} = \frac{n_1n_2}{n\sqrt{n}}(\mathbb{P}_R\{2k\} + \mathbb{P}_R\{2k+1\}) = \frac{n_1n_2}{2k\sqrt{n}}\mathbb{P}_R\{2k\}.$$

Dla ustalonego $x \in \mathbb{R}$ niebędącego węzłem wskazujemy $k = \lceil \frac{n_1n_2}{n}(1 + \frac{x}{\sqrt{n}}) \rceil$, dla którego $x \in (t_{k-1}, t_k)$, co implikuje zbieżność $t_k \rightarrow x$ oraz $k/n \rightarrow \lambda(1-\lambda)$ – niezależnie od x . Jednostajna zbieżność ciągu funkcji G wynika teraz ze zbieżności ciągu pochodnych

$$G'(x) = \frac{1}{\sqrt{2\pi}} \frac{n_1n_2}{kn} \left(1 - \frac{t_k}{\sqrt{n}} \frac{n_1n_2}{k^2}\right)^{k-\frac{1}{2}} \left(1 - \frac{t_k}{\sqrt{n}} \frac{n_2}{n_1}\right)^{-n_1} \left(1 - \frac{t_k}{\sqrt{n}} \frac{n_1}{n_2}\right)^{-n_2} \cdot c_n.$$

Istotnie, wiemy już, że $\frac{n_1n_2}{kn} \left(1 - \frac{t_k}{\sqrt{n}} \frac{n_1n_2}{k^2}\right)^{-\frac{1}{2}} \rightarrow 1$, dla $n \rightarrow \infty$. Pozostały iloczyn logarytmujemy i korzystamy z rozwinięcia funkcji $\ln(1-z)$, co sprowadza go do postaci

$$\begin{aligned} & -k \left(\frac{t_k}{\sqrt{n}} \frac{n_1n_2}{k^2} + \frac{t_k^2}{2n} \frac{n_1^2n_2^2}{k^4} + \dots \right) + n_1 \left(\frac{t_k}{\sqrt{n}} \frac{n_2}{n_1} + \frac{t_k^2}{2n} \frac{n_2^2}{n_1^2} + \dots \right) + n_2 \left(\frac{t_k}{\sqrt{n}} \frac{n_1}{n_2} + \frac{t_k^2}{2n} \frac{n_1^2}{n_2^2} + \dots \right) \\ & = \frac{t_k}{\sqrt{n}} \left(-\frac{n_1n_2}{k} + n_2 + n_1 \right) + \frac{t_k^2}{2n} \left(-\frac{n_1^2n_2^2}{k^3} + \frac{n_2^2}{n_1} + \frac{n_1^2}{n_2} \right) + \dots, \end{aligned}$$

gdzie pomijamy dążące do zera składniki zawierające $t_k^i, i \geq 3$. Ponieważ czynnik przy $\frac{t_k}{\sqrt{n}}$ jest równy $\frac{kn - n_1n_2}{k} = \frac{t_k n_1n_2}{k\sqrt{n}}$, do wyznaczenia pozostaje granica

$$\lim_{n \rightarrow \infty} \frac{t_k^2}{2n} \left(\frac{2n_1n_2}{k} - \frac{n_1^2n_2^2}{k^3} + \frac{n_1^3 + n_2^3}{n_1n_2} \right) = \frac{x^2}{2} \left(2 + \frac{-1 + (\lambda^3 + (1-\lambda)^3)}{\lambda(1-\lambda)} \right) = \frac{-x^2}{2}.$$

Zachęcamy czytelnika do uzupełnienia brakujących szczegółów – zgodnie ze schematem zrealizowanym w dowodzie tw. 2.4.1. $\square$

2.4.2 ...i od strony rozkładów ciągłych

Podobnie jak rozkład opisany wzorem (2.49) jest najprostszym nietrywialnym rozkładem dyskretnym, miara Lebesgue'a, czyli rozkład jednostajny na przedziale jednostkowym $I = [0, 1]$ może być uważana za najprostszy rozkład ciągły. Dla każdej $(2n+1)$ -elementowej próbki losowej $x = (x_i)_{i \leq 2n+1} \in I^{2n+1}$ porządek liniowy

$$x_{i_0} \leq x_{i_1} \leq \dots \leq x_{i_{2n}}$$

pozwala wybrać *element centralny*, środkowy, $V_n(x) = x_{i_n}$. Powstałą w ten sposób zmienną losową (medianę) nazywamy *centralną statystyką pozycyjną* rozkładu jednostajnego. Aby wykazać, że również ta zmienna ma w granicy ścisły związek z rozkładem normalnym, wyznaczmy rozkład $\mathbb{P}_{V_n}$.

Twierdzenie 2.4.5. *Dla każdego $n \in \mathbb{N}$ miara $\mathbb{P}_{V_n}$ jest rozkładem ciągłym o gęstości*

$$f_n(x) = \begin{cases} (2n+1) \binom{2n}{n} x^n (1-x)^n & \text{dla } x \in [0, 1] \\ 0 & \text{dla } x < 0 \text{ lub } x > 1, \end{cases}$$

przy czym $\mathcal{E}(V_n) = \frac{1}{2}$ i $\text{Var}(V_n) = \frac{1}{4(2n+3)}$.

Dowód. Dla dowolnego ustalonego przedziału $[a, b] \subset I$ rozważamy zbiór $A = V_n^{-1}[a, b] \subset I^{2n+1}$ będący sumą $(2n+1) \binom{2n}{n} = \frac{(2n+1)!}{(n!)^2}$ zbiorów

$$A_{j,J} = \{(x_i)_{i \leq 2n+1}; x_j \in [a, b], x_i \leq x_j \leq x_{i'} \text{ dla } i \in J, i' \notin J\},$$

gdzie $j \leq 2n+1$, a $J \not\ni j$ przebiega rodzinę n -elementowych podzbiorów pomniejszonego zbioru indeksów. Dla każdej pary (j, J) zachodzą inkluzje $A_{j,J}^- \subset A_{j,J} \subset A_{j,J}^+$, gdzie

$$A_{j,J}^- = \{x \in I^{2n+1}; x_j \in [a, b], x_i \leq a \text{ dla } i \in J, x_{i'} \geq b \text{ dla } i' \notin J \cup \{j\}\},$$

$$A_{j,J}^+ = \{x \in I^{2n+1}; x_j \in [a, b], x_i \leq b \text{ dla } i \in J, x_{i'} \geq a \text{ dla } i' \notin J \cup \{j\}\}.$$

Ponieważ zbiory $A_{j,J}^-$ są wzajemnie rozłączne, wynikają stąd nierówności

$$\frac{(2n+1)!}{(n!)^2} (b-a) a^n (1-b)^n \leq \ell^{2n+1}(A) \leq \frac{(2n+1)!}{(n!)^2} (b-a) b^n (1-a)^n.$$

Dla zakończenia dowodu wystarczy stwierdzić różniczkowalność dystrybucyjności F_{V_n} i wyznaczyć jej pochodną w przedziale $[0, 1]$. Dla $x \in I$ i $h \neq 0$, takich że $x+h \in I$, mamy

$$\frac{(2n+1)!}{(n!)^2} x^n (1-x-h)^n \leq \frac{F_{V_n}(x+h) - F_{V_n}(x)}{h} \leq \frac{(2n+1)!}{(n!)^2} (x+h)^n (1-x)^n,$$

jeśli $h > 0$ i, analogicznie,

$$\frac{(2n+1)!}{(n!)^2}(x+h)^n(1-x)^n \leq \frac{F_{V_n}(x+h) - F_{V_n}(x)}{h} \leq \frac{(2n+1)!}{(n!)^2}x^n(1-x-h)^n,$$

dla $h < 0$. Wspólna wartość graniczna dla $h \rightarrow 0$ jest szukaną gęstością rozkładu. Z udowodnionego wzoru wynika także równość $\int_{\mathbb{R}} f_n(x) dx = 1$.

Wartość oczekiwaną $\int_{\mathbb{R}} x f_n(x) dx$ i wariancję $\int_{\mathbb{R}} (x - \frac{1}{2})^2 f_n(x) dx$ wyznaczamy z tożsamości

$$\int_0^1 x^{n+1}(1-x)^n dx = \int_0^1 (1-x)^{n+1} x^n dx = \int_0^1 ((1-x)^n x^n - (1-x)^n x^{n+1}) dx,$$

oraz

$$\begin{aligned} \int_0^1 (x - \frac{1}{2})^2 x^n (1-x)^n dx &= - \int_0^1 x^{n+1} (1-x)^{n+1} dx + \frac{1}{4} \int_0^1 x^n (1-x)^n dx \\ &= \frac{-(n+1)!^2}{(2n+3)!} + \frac{(n!)^2}{4(2n+1)!} = \frac{(n!)^2}{4(2n+1)!(2n+3)}. \quad \square \end{aligned}$$

Wniosek 2.4.6. Ciąg dystrybuant zmiennych losowych

$$Y_n = 2\sqrt{2n+3}(V_n - \frac{1}{2}), \quad \text{dla } n \in \mathbb{N},$$

jest zbieżny do dystrybuanty standardowego rozkładu normalnego.

Dowód. Dla uproszczenia obliczeń zajmijmy się zmienną $Z_n = 2\sqrt{2n}(V_n - \frac{1}{2})$. Tożsamość

$$\ell^n\{Z_n \leq t\} = \int_{-\infty}^{\frac{1}{2} + \frac{t}{2\sqrt{2n}}} f_n(x) dx,$$

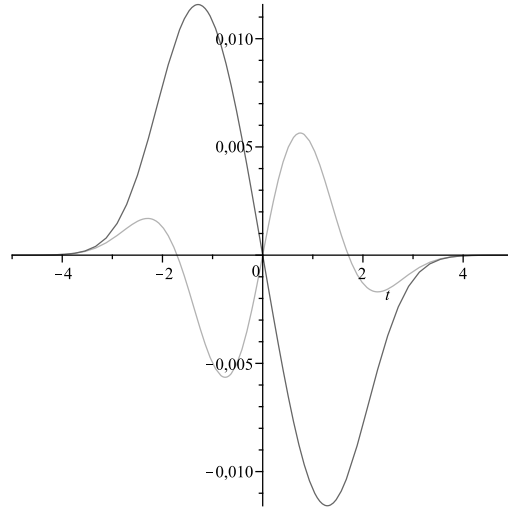
po podstawieniu $\xi = (2x - 1)\sqrt{2n}$, prowadzi do wzoru

$$\begin{aligned} F_{Z_n}(t) &= \frac{(2n+1)!}{2(n!)^2\sqrt{2n}} \int_{-\sqrt{2n}}^t \left(\frac{1}{2} + \frac{\xi}{2\sqrt{2n}}\right)^n \left(\frac{1}{2} - \frac{\xi}{2\sqrt{2n}}\right)^n d\xi \\ &= \binom{2n}{n} 2^{-2n} \sqrt{2n} \int_{-\sqrt{2n}}^t \left(1 - \frac{\xi^2}{2n}\right)^n d\xi \quad \text{dla } |t| \leq \sqrt{2n}. \end{aligned}$$

Zbieżność (jednostajna) ciągu dystrybuant do funkcji $\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{\xi^2}{2}} d\xi$, dla $t \in \mathbb{R}$, wynika z istnienia granic

$$\binom{2n}{n} 2^{-2n} \sqrt{2n} \longrightarrow \frac{1}{\sqrt{2\pi}} \quad \text{i} \quad \left(1 - \frac{\xi^2}{2n}\right)^n \nearrow e^{-\frac{\xi^2}{2}}$$

dla $n \rightarrow \infty$. □



Rysunek 2.7: Błędy przybliżenia rozkład zmiennych Y_{11} (jasnoszary) i Z_{11}

Uwaga 2.4.2. Dla $-\sqrt{2n+3} \leq a < b \leq \sqrt{2n+3}$, sformułowana wyżej własność prowadzi do asymptotycznej formuły

$$\ell^n \left\{ \frac{a}{2\sqrt{2n+3}} \leq V_n - \frac{1}{2} \leq \frac{b}{2\sqrt{2n+3}} \right\} \xrightarrow{\approx} \Phi(b) - \Phi(a),$$

pozwalającej na wyznaczanie przybliżonych wartości rozkładu. Błąd przybliżenia dla $n = 11$ jest zbliżony do 0,006.

Na rysunku 2.7 podany jest także wykres błędu (kol. ciemnoszary) dla rozważanej w dowodzie zmiennej Z_n ze zredukowanym czynnikiem normalizującym.

2.4.3 Rozkład wielowymiarowy

Analizę (nie-)zależności ciągu zmiennych losowych o rozkładzie normalnym rozpoczyna

Lemat 2.4.7. *Suma lub różnica $X \pm Y$ dwu niezależnych zmiennych losowych o rozkładzie normalnym, odpowiednio, $\mathcal{N}(\bar{x}, a)$ i $\mathcal{N}(\bar{y}, b)$ dla $a, b > 0$, ma rozkład normalny $\mathcal{N}(\bar{x} \pm \bar{y}, a+b)$.*

Dowód. Gęstość f sumy $X + Y$ wyznaczamy ze wzoru

$$f(z) = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi a}} e^{-\frac{(t-\bar{x})^2}{2a}} \frac{1}{\sqrt{2\pi b}} e^{-\frac{(z-t-\bar{y})^2}{2b}} dt \quad \text{dla } z \in \mathbb{R},$$

czyli, po zamianie zmiennej t na $t + \bar{x}$,

$$\begin{aligned}
2\pi\sqrt{ab}f(z + (\bar{x} + \bar{y})) &= \int_{\mathbb{R}} \exp\left(-\frac{t^2}{2a} - \frac{(z-t)^2}{2b}\right) dt \\
&= \int_{\mathbb{R}} \exp\left(-\frac{a+b}{2ab}\left(t^2 - \frac{2a}{a+b}tz + \frac{a}{a+b}z^2\right)\right) dt \\
&= \exp\left(-\frac{z^2}{2(a+b)}\right) \int_{\mathbb{R}} \exp\left(-\frac{a+b}{2ab}\left(t - \frac{a}{a+b}z\right)^2\right) dt \\
&= \exp\left(-\frac{z^2}{2(a+b)}\right) \sqrt{\frac{2\pi ab}{a+b}}.
\end{aligned}$$

dla $z \in \mathbb{R}$. Przypadek różnicy wynika z faktu, że zmienna $-Y$ ma rozkład $\mathcal{N}(-\bar{y}, b)$. $\square$

Wniosek 2.4.8. *Jeśli współrzędne wektora losowego $Z = (Z_1, \dots, Z_p)^t$ są dowolnymi zmiennymi niezależnymi o rozkładzie normalnym, wówczas*

- (i) *dla każdego $a \in \mathbb{R}^p$ zmienna $a^t Z$ ma rozkład normalny, oraz*
- (ii) *dla dowolnej nieosobliwej macierzy P wymiaru $p \times p$ rozkład wektora $X = P \cdot Z$ ma postać*

$$\mathbb{P}_X(A) = \frac{1}{(2\pi)^{p/2} \sqrt{\det(W)}} \int_A \exp\left(-\frac{(x-m)^t W^{-1} (x-m)}{2}\right) dx \quad (2.96)$$

dla $A \in \mathfrak{B}_{\mathbb{R}^p}^p$, gdzie $m = \mathcal{E}(X)$, $W = \text{Cov}(X) = P \cdot \Delta \cdot P^t$, a $\Delta = \text{Cov}(Z)$ jest macierzą diagonalną.

Dowód. Własność (i) jest prostym uogólnieniem lematu 2.4.7 na kombinację liniową dowolnej liczby zmiennych niezależnych.

Dla dowodu (ii) zauważmy, że jeśli współrzędne Z_i mają rozkład normalny $\mathcal{N}(m_i, \sigma_i^2)$ dla $i \leq p$, to $\mathbb{P}_Z = \bigotimes_{i \leq p} \mathbb{P}_{Z_i}$ (tw. 2.1.27), a więc

$$\begin{aligned}
\mathbb{P}_Z(A) &= \frac{1}{(2\pi)^{p/2} \sigma_1 \dots \sigma_p} \int_A e^{-\frac{(z_1-m_1)^2}{2\sigma_1^2} - \dots - \frac{(z_p-m_p)^2}{2\sigma_p^2}} dz_1 \dots dz_p \quad (2.97) \\
&= \frac{1}{(2\pi)^{p/2} \sqrt{\det(\Delta)}} \int_A e^{-\frac{1}{2}(z-m_Z)^t \Delta^{-1} (z-m_Z)} dz
\end{aligned}$$

dla $A \in \mathfrak{B}_{\mathbb{R}^p}^p$, gdzie $m_Z = \mathcal{E}(Z)$, $\Delta = \text{Cov}(Z)$. Jeśli zatem zmienna wektorowa $X = PZ: \Omega \rightarrow \mathbb{R}^p$ jest złożeniem Z i przekształcenia liniowego $L_P: \mathbb{R}^p \rightarrow \mathbb{R}^p$ o macierzy P , to $\mathbb{P}_X(A) = \mathbb{P}_Z(L_P^{-1}(A))$, a po zamianie zmiennych otrzymujemy – zgodnie z wnioskiem 1.2.16

$$\mathbb{P}_X(A) = \frac{|\det(P^{-1})|}{(2\pi)^{p/2} \sqrt{\det(\Delta)}} \int_A e^{-\frac{1}{2}(P^{-1}x-m_Z)^t \Delta^{-1} (P^{-1}x-m_Z)} dx,$$

co dla $m = P \cdot m_Z$ oznacza dokładnie równość (2.96). Liniowe zależności pomiędzy (wektorowymi) wartościami oczekiwanymi i macierzami kowariancji wektorów X i Z wynikają ze stwierdzenia 2.1.43(i). $\square$

Definicja 2.4.9. Dla $p \in \mathbb{N}$ i dowolnej symetrycznej dodatnio określonej macierzy $W \in \mathbb{R}^{p \times p}$ p -wymiarowym rozkładem normalnym $\mathcal{N}_p(m, W)$ o wartości oczekiwanej $m \in \mathbb{R}^p$ i macierzy kowariancji W nazywamy miarę probabilistyczną w $\mathbb{R}^p$, której gęstość f wyraża się wzorem

$$f(x) = \frac{1}{(2\pi)^{p/2} \sqrt{\det(W)}} \exp\left(-\frac{(x-m)^t W^{-1}(x-m)}{2}\right) \quad \text{dla } x \in \mathbb{R}^p. \quad (2.98)$$

Możliwość rozkładu względem wartości własnych $W = P \cdot \Delta \cdot P^t$ oznacza, zgodnie z wnioskiem 2.4.8, że dla każdej dodatnio określonej macierzy symetrycznej W wzór (2.98) jest gęstością miary probabilistycznej.

Uogólniając własność (i) z wniosku 2.4.8, przyjmujemy

Definicja 2.4.10. Wektor losowy $X = (X_1, \dots, X_p)^t$ nazywamy p -wymiarową zmienną typu Gaussa (ang. Gaussian vector), jeśli dla każdego $a \in \mathbb{R}^p$ kombinacja liniowa $a^t X = a_1 X_1 + \dots + a_p X_p$ jest zmienną o rozkładzie normalnym. Zmienne $X_1, \dots, X_p$ nazywamy w takim przypadku zmiennymi typu Gaussa.

Każda ze zmiennych typu Gaussa ma rozkład normalny. Zgodnie ze stwierdzeniem 2.1.43(i), znamy rozkłady $\mathbb{P}_{a^t X} = \mathcal{N}(a^t \mathcal{E}(X), a^t \text{Cov}(X) a)$ dla $a \in \mathbb{R}^p$.

Twierdzenie 2.4.11 (Charakteryzacja wektora typu Gaussa). Wektor losowy $X = (X_1, \dots, X_n)^t$ o nieosobliwej macierzy kowariancji $W = \text{Cov}(X)$ jest typu Gaussa wtedy i tylko wtedy, gdy rozkład prawdopodobieństwa $\mathbb{P}_X$ jest p -wymiarowym rozkładem normalnym $\mathcal{N}_p(\mathcal{E}(X), \text{Cov}(X))$.

Stąd niebanalny, ale w tym kontekście oczywisty

Wniosek 2.4.12. Zmienne losowe $X_1, \dots, X_n$ typu Gaussa są niezależne wtedy i tylko wtedy, gdy $\text{Cov}(X_i, X_j) = 0$ dla wszystkich $i \neq j$. $\square$

Dowód twierdzenia 2.4.11. „ $\Leftarrow$ ” Korzystamy z rozkładu macierzy kowariancji $W = P^t \cdot \Delta \cdot P$, w którym Δ jest macierzą diagonalną. Przyjmując $Z = P^{-1}X$, odwracamy rozumowanie wykorzystane w dowodzie wniosku 2.4.8. Istotnie, dla dowolnego zbioru borelowskiego $A \subset \mathbb{R}^p$ zachodzi równość $\mathbb{P}_Z(A) = \mathbb{P}_X(L_P(A))$, a więc, po zamianie zmiennych

$$\mathbb{P}_Z(A) = \frac{|\det(P)|}{(2\pi)^{p/2} \sqrt{\det(W)}} \int_A e^{-\frac{1}{2}(Pz-m)^t W^{-1}(Pz-m)} dz,$$

co po uproszczeniu przybiera postać (2.97). Wynika stąd, że zmienne $Z_i, i \leq p$, tworzące wektor Z są niezależne. Na mocy wniosku 2.4.8(ii) dowolne ich

kombinacje liniowe – w tym kombinacje liniowe zmiennych $X_1, \dots, X_p$ – mają rozkład normalny.

Implikacja „ $\Rightarrow$ ” wynika z faktu, iż – jak wykażemy w kolejnym podrozdziale (wniosek 2.5.18) – rozkład dowolnej p -wymiarowej zmiennej wektorowej X jest jednoznacznie wyznaczony przez rozkłady zmiennych jednowymiarowych $a^t X$, dla $a \in \mathbb{R}^p$. Dowód cytowanego wniosku korzysta z pojęcia *funkcji charakterystycznej* rozkładu. $\square$

2.4.4 Jak „wprostować” rozkład normalny?

1-wymiarowy rozkład normalny $\mathcal{N}(0, 1)$ jest – podobnie jak każdy rozkład prawdopodobieństwa na prostej – indukowany z rozkładu jednostajnego $\ell|_I$ przez swoją funkcję kwantylową $Q: (0, 1) \rightarrow \mathbb{R}$; jest *jej rozkładem*, jako zmiennej losowej na $I = [0, 1]$. W tym przypadku $Q = \Phi^{-1}$ jest funkcją odwrotną do dystrybuanty (2.91).

Formalnie rzecz biorąc, jeśli ciąg liczb $x_i \in (0, 1)$, $i \leq n$, jest próbka losową z rozkładu jednostajnego, to ciąg $y_i = \Phi^{-1}(x_i)$, $i \leq n$, jest próbka z rozkładu normalnego. Praktyczne wykorzystanie tej własności jest jednak niewielkie, a wynika to z braku dostatecznie dokładnych i szybkich algorytmów wyznaczania Φ^{-1} .

Problem indukowania rozkładu normalnego z miary Lebesgue’a, a więc z rozkładu jednostajnego, upraszcza się w wymiarze 2, a dla miary $\mathcal{N}_2(0, I_2)$ znane są przynajmniej dwa rozwiązania (Box–Muller 1958 i Marsaglia 1964). Oba korzystają z oczywistej redukcji całki

$$\iint e^{-(x^2+y^2)/2} dx dy = \int_0^\infty \int_0^{2\pi} r e^{-r^2/2} d\varphi dr$$

powstałej przy zmianie współrzędnych na biegunowe $(x, y) = (r \cos \varphi, r \sin \varphi)$.

Definicja 2.4.13. Transformację Boxa–Mullera nazywamy *dyfeomorfizm* $G: (0, 1)^2 \rightarrow \mathbb{R}^2$, określony wzorem

$$G: (u, v) \mapsto \sqrt{-2 \ln(u)} (\cos(2\pi v), \sin(2\pi v)). \quad (2.99)$$

Transformację Marsaglii nazywamy *dyfeomorfizm* $H: \mathbb{R}^2 \supset B(0, 1) \setminus \{0\} \rightarrow \mathbb{R}^2$ postaci

$$H: (u, v) \mapsto (u \sqrt{-2 \ln(s)/s}, v \sqrt{-2 \ln(s)/s}), \quad s = u^2 + v^2. \quad (2.100)$$

Twierdzenie 2.4.14. Każda ze zdefiniowanych wyżej transformacji wyznacza izomorfizm borelowskich przestrzeni z miarą

$$\begin{aligned} G: (I^2, \mathfrak{B}_{I^2}, \ell^2) &\xrightarrow{\cong} (\mathbb{R}^2, \mathfrak{B}_{\mathbb{R}^2}, \mathcal{N}_2(0, I_2)) \quad \text{oraz} \\ H: (B(0, 1), \mathfrak{B}_{B(0, 1)}, \tfrac{1}{\pi} \ell^2) &\xrightarrow{\cong} (\mathbb{R}^2, \mathfrak{B}_{\mathbb{R}^2}, \mathcal{N}_2(0, I_2)), \end{aligned}$$

co oznacza równość miar indukowanych $G_*\ell^2|_{I^2} = \mathcal{N}_2(0, I_2) = H_*(\frac{1}{\pi}\ell^2|_{B(0,1)})$ i tożsamości

$$\iint_A \frac{1}{2\pi} e^{-(x^2+y^2)/2} dx dy = \int_{G^{-1}A} d\ell^2 = \frac{1}{\pi} \int_{H^{-1}A} d\ell^2,$$

dla $A \in \mathfrak{B}_{\mathbb{R}^2}$.

Uzyskany w twierdzeniu brak funkcji podcałkowej w dwu ostatnich całkach wyjaśnia sformułowanie użyte w tytule bieżącego podrozdziału.

Dowód. Przekształcenie (2.99) wykorzystuje – zamiast parametru r – funkcję odwrotną do dystrybuanty $\int_0^\rho r e^{-r^2/2} dr = 1 - e^{-\rho^2/2}$, dla $\rho > 0$, a jej własności są prostą konsekwencją złożenia wskazanych transformacji.

Zamiast bezpośrednich i niezbyt przyjaznych rachunków pokażemy, że odkryte kilka lat później przekształcenie (2.100) jest wynikiem poszukiwań dyfeomorfizmu postaci

$$H(u, v) = r(s)(\cos \varphi(u, v), \sin \varphi(u, v)),$$

gdzie $s = u^2 + v^2$, dla (u, v) przebiegających nieznany jeszcze podzbiór w $\mathbb{R}^2$. Jakobian transformaty H wynosi $J(H) = 2(rr')(s)(u\varphi'_v - v\varphi'_u)$, a zatem

$$\iint_A \frac{1}{2\pi} e^{-(x^2+y^2)/2} dx dy = \frac{1}{\pi} \iint_{H^{-1}A} r e^{-\frac{s}{2}} |r'| |u\varphi'_v - v\varphi'_u| du dv.$$

Czynnik zależny od s jest równy 1, dla $r(s) = \sqrt{-2\ln(s)}$, a to ogranicza dziedzinę szukanego dyfeomorfizmu do koła $0 < u^2 + v^2 < 1$ o polu π . Jeśli przyjmiemy, że $\varphi(u, v) = \lambda(t)$ jest lokalnie funkcją ilorazu $t \equiv v/u$, to pod całką pozostaje wyrażenie $|(1+t^2)\lambda'(t)|$ równe 1 wtedy, gdy

$$\lambda(t) = \pm \arctan(t), \quad \text{czyli} \quad t = \pm \tan \varphi.$$

Z tożsamości trygonometrycznej $\tan^2 \varphi = \frac{1}{\cos^2 \varphi} - 1$ wyznaczamy $\cos^2 \varphi = \frac{u^2}{u^2+v^2}$ i przyjmujemy

$$\cos \varphi = \frac{u}{\sqrt{s}}, \quad \sin \varphi = \frac{v}{\sqrt{s}}, \quad \text{dla } 0 < s < 1,$$

co nadaje szukanej transformacji postać (2.100). Opisane wyżej przekształcenia stanowią dowód równości $H_*\mu = \mathcal{N}_2(0, I_2)$, w której μ oznacza miarę probabilistyczną $\frac{1}{\pi}\ell^2|_{B(0,1)}$ na kole jednostkowym. $\square$

Dysponując opisanymi w podrozdziale 2.2 wiarygodnymi generatorami ciągów liczb pseudolosowych, proponujemy zatem:

Algorytm Boxa–Mullera

Algorytm B-M polega na bezpośrednim zastosowaniu transformacji (2.99) do pary kolejnych liczb pseudolosowych $(u, v) = (X_n, X_{n+1})$ symulujących rozkład jednostajny w przedziale $[0, 1)$.

Algorytm Marsaglii

Transformata Marsaglii prowadzi do klasycznego nieco prostszego obliczeniowo algorytmu zamieniającego 2-elementową próbkę (u, v) z rozkładu jednostajnego na odcinku $[-1, 1]$ na 2-elementową próbkę (ξ, η) z rozkładu normalnego $\mathcal{N}(0, 1)$ – jeśli tylko $(u, v) \in B(0, 1)$. Przykładowa implementacja obejmuje kroki:

- pobrać $(u, v) = (X_n, X_{n+1})$ jako kolejne dwie liczby pseudolosowe z odcinka $[0, 1)$
- $(U, V) = (2u - 1, 2v - 1)$; jeśli $S \equiv U^2 + V^2 \geq 1$, wrócić do poprzedz. punktu
- liczby
$$\xi = U \sqrt{\frac{-2 \ln(S)}{S}}, \quad \eta = V \sqrt{\frac{-2 \ln(S)}{S}}$$

są kolejnymi dwiema próbkami zgodnymi z rozkładem normalnym $\mathcal{N}(0, 1)$.

2.5 Przestrzenie rozkładów prawdopodobieństwa

Niech $\mathcal{P}(\mathbb{R}^p)$ oznacza zbiór wszystkich *rozkładów prawdopodobieństwa*, czyli miar probabilistycznych w przestrzeni $(\mathbb{R}^p, \mathfrak{B}_{\mathbb{R}^p})$ dla ustalonego $p \in \mathbb{N}$. Jak wiemy, każda miara $\mu \in \mathcal{P}(\mathbb{R}^p)$ może być jednoznacznie opisana poprzez swą dystrybuantę $F_\mu: \mathbb{R}^p \rightarrow [0, 1]$.

2.5.1 Transformata Fouriera

Wprowadzone w pierwszych latach XIX w. koncepcje *transformaty*, pochodzące od Jeana-Baptiste J. Fouriera oraz Pierre’a S. de Laplace’a i intensywnie wykorzystywane m.in. w teorii równań różniczkowych, podpowiadają, by badanie miary jako funkcjonału na $\mathcal{C}_b(\mathbb{R})$ rozszerzyć do $\mathcal{C}_b(\mathbb{R}, \mathbb{C})$, a następnie ograniczyć do szczególnej 1-parametrowej rodziny funkcji. Korzystamy z rozszerzenia całki Lebesgue’a względem μ jako dodatniej miary Radona (podrozdział 3.2.2), a podstawą jest

Lemat 2.5.1. *Dla dowolnej skończonej miary borelowskiej μ na prostej $\mathbb{R}$ funkcjonal*

$$\int d\mu: \mathcal{C}_b(\mathbb{R}, \mathbb{C}) \ni f \mapsto \int \Re f d\mu + i \int \Im f d\mu \in \mathbb{C}$$

jest $\mathbb{C}$ -liniowy, przy czym zachodzi nierówność

$$\left| \int f d\mu \right| \leq \int |f| d\mu \quad \text{dla } f \in \mathcal{C}_b(\mathbb{R}, \mathbb{C}). \quad (2.101)$$

Dowód. Liniowość rozszerzenia jest oczywista. Dla dowodu (2.101) ustalamy funkcję $f \in \mathcal{C}_b(\mathbb{R}, \mathbb{C})$, a za $z \in \mathbb{C}$ weźmy liczbę o module 1, dla której

$$\left| \int f d\mu \right| = z \int f d\mu = \int zf d\mu.$$

Wówczas $\left| \int f d\mu \right| = \int \Re(zf) d\mu \leq \int |zf| d\mu$. □

Definicja 2.5.2. Funkcją charakterystyczną $\varphi_\mu: \mathbb{R} \rightarrow \mathbb{C}$ rozkładu prawdopodobieństwa μ na prostej nazywamy funkcję zespoloną zmiennej rzeczywistej określoną wzorem

$$\varphi_\mu(t) = \int e^{itx} \mu(dx) = \int \cos(tx) \mu(dx) + i \int \sin(tx) \mu(dx) \quad \text{dla } t \in \mathbb{R}. \quad (2.102)$$

Stwierdzenie 2.5.3 (Własności funkcji charakterystycznej). *Dla dowolnego rozkładu μ*

- (i) $|\varphi_\mu(t)| \leq \varphi_\mu(0) = 1$,
- (ii) $\varphi_\mu(-t) = \overline{\varphi_\mu(t)}$, oraz
- (iii) $|\varphi_\mu(t+h) - \varphi_\mu(t)| \leq \int |e^{ihx} - 1| \mu(dx)$

dla $t, h \in \mathbb{R}$. W szczególności, funkcja φ_μ jest jednostajnie ciągła.

Dowód. Ćwiczenie. Jednostajna ciągłość wynika z (iii) i z twierdzenia Lebesgue'a o zmajoryzowanej zbieżności. □

Uwaga 2.5.1. Jeśli μ jest rozkładem ciągłym o gęstości $f: \mathbb{R} \rightarrow \mathbb{R}_+$, to równość

$$\varphi_\mu(t) = \int_{\mathbb{R}} e^{itx} f(x) dx \quad \text{dla } t \in \mathbb{R}$$

oznacza, że funkcja φ_μ jest *transformatą Fouriera* funkcji $f \in L^1(\mathbb{R})$. Klasyczny wzór na transformatę odwrotną,

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_\mu(t) dt \quad \text{dla } x \in \mathbb{R}, \quad (2.103)$$

jeśli $\int |\varphi_\mu| d\ell < \infty$, ustala w tym przypadku właściwą dla transformacji wzajemnie jednoznaczność $\mu \mapsto \varphi_\mu$.

W przypadku dowolnej miary probabilistycznej μ w $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$, przy rezygnacji z założenia całkowalności funkcji charakterystycznej φ_μ , formalne zastąpienie części wyrażenia podcałkowego w (2.103) przez całkę

$$\int_a^b e^{-itx} dx = \frac{e^{-ita} - e^{-itb}}{it} \quad \text{dla } a, b, t \in \mathbb{R},$$

nie prowadzi do funkcji całkowalnej. Pomimo to, zachodzi

Twierdzenie 2.5.4 (Wzór na transformatę odwrotną). *Miary w przestrzeni $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ są równe wtedy i tylko wtedy, gdy równe są ich funkcje charakterystyczne. Dokładniej, dla dowolnego rozkładu prawdopodobieństwa μ na prostej i dowolnych liczb $a < b$,*

$$\lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \varphi_{\mu}(t) dt = \mu((a, b)) + \frac{1}{2} \mu(\{a, b\}). \quad (2.104)$$

Dowód. Dla $T > 0$ przyjmijmy

$$\begin{aligned} I_T &= \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \varphi_{\mu}(t) dt = \int_{-T}^T \int \frac{e^{-ita} - e^{-itb}}{it} e^{itx} \mu(dx) dt \\ &= \int \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} e^{itx} dt \mu(dx), \end{aligned}$$

gdzie ciągłość i ograniczoność funkcji podcałkowej zapewniają całkowalność względem miary produktowej $\mu \otimes \ell|_{[-T, T]}$ i pozwalają na skorzystanie z twierdzenia Fubiniego. Po zamianie $t \mapsto -t$ w przedziale $[-T, 0]$ i drobnych przekształceniach otrzymujemy całkę w postaci rzeczywistej

$$\begin{aligned} I_T &= 2 \int \int_0^T \frac{\sin(t(x-a)) - \sin(t(x-b))}{t} dt \mu(dx) \\ &= 2 \int (S(x-a, T) - S(x-b, T)) \mu(dx), \end{aligned} \quad (2.105)$$

gdzie przyjęliśmy oznaczenie

$$S(\alpha, T) = \int_0^T \frac{\sin(\alpha t)}{t} dt = \text{sign}(\alpha) \int_0^{|\alpha|T} \frac{\sin(t)}{t} dt \quad \text{dla } \alpha \in \mathbb{R},$$

a zatem $\lim_{T \rightarrow \infty} S(\alpha, T) \rightarrow \pm \frac{\pi}{2}$, dla $\alpha \neq 0$ (por. [2, prz. 3.1.8]). Wspólne ograniczenie

$$|S(x-a, T) - S(x-b, T)| \leq 2 \sup\left\{ \left| \int_0^a \frac{\sin(t)}{t} dt \right|; a > 0 \right\} < \infty$$

i twierdzenie o zmajoryzowanej zbieżności umożliwiają przejście w równości (2.105) do granicy pod znakiem całki. Ponieważ dla $T \rightarrow \infty$

$$S(x-a, T) - S(x-b, T) \rightarrow \begin{cases} \pi, & \text{jeśli } a < x < b \\ \frac{\pi}{2} & \text{dla } x = a \text{ lub } x = b \\ 0 & \text{dla } x < a \text{ lub } x > b, \end{cases}$$

otrzymujemy w granicy równość (2.104). Pozostawiamy czytelnikowi modyfikację wzoru pozwalającą na wyznaczenie dystrybucyj miary μ . $\square$

Udowodniony wzór na transformatę odwrotną pozwala na niezależny dowód i wzmocnienie formuły (2.103).

Wniosek 2.5.5. *Jeśli funkcja charakterystyczna φ_μ jest całkowalna, to μ jest rozkładem ciągłym, o ciągłej gęstości f danej wzorem (2.103).*

Dowód. Przy podanych założeniach, dla dowolnych $a < b$ nierówność

$$\int_{-T}^T \left| \frac{e^{-ita} - e^{-itb}}{it} \varphi_\mu(t) \right| dt \leq |b - a| \int_{\mathbb{R}} |\varphi_\mu(t)| dt \quad (2.106)$$

implikuje istnienie całki po całej prostej $\mathbb{R}$, a równość (2.104) przyjmuje postać

$$\mu((a, b)) + \frac{1}{2}\mu(\{a, b\}) = \int_{-\infty}^{\infty} \int_a^b e^{-itx} dx \varphi_\mu(t) dt.$$

Ponieważ ograniczenie wskazane w (2.106) dąży do 0 dla $a \nearrow b$, graniczną wartością lewej strony jest $\mu\{b\} = 0$ dla $b \in \mathbb{R}$; z twierdzenia Fubiniego otrzymujemy zatem

$$\mu((a, b)) = \int_a^b \int_{-\infty}^{\infty} e^{-itx} \varphi_\mu(t) dt dx \quad \text{dla dowolnych } a < b. \quad (2.107)$$

Sprawdzenie, że funkcja $f(x) = \int_{\mathbb{R}} e^{-itx} \varphi_\mu(t) dt$, dla $x \in \mathbb{R}$, jest rzeczywista i ciągła pozostawiamy czytelnikowi. Z tożsamości (2.107) wynika, że jest także nieujemna i całkowalna, przy czym $\mu((\infty, b]) = \int_{-\infty}^b f(x) dx$. $\square$

2.5.2 Ciągłość

Dla zmiennych losowych przyjmujemy klasyczną terminologię odnoszącą się w istocie do zbieżności ciągu miar.

Definicja 2.5.6. *Ciąg p -wymiarowych zmiennych losowych $(X_n)_{n \in \mathbb{N}}$ nazywamy zbieżnym według rozkładu do borelowskiej miary probabilistycznej μ w $\mathbb{R}^p$, co zapisujemy jako $X_n \xrightarrow{d} \mu$, jeśli ciąg rozkładów zmiennych X_n jest słabo zbieżny do μ . Jeśli graniczna miara μ jest rozkładem zmiennej X , korzystamy z uproszczonego zapisu $X_n \xrightarrow{d} X$ i mówimy o zbieżności wg. rozkładu do zmiennej X .*

Wbrew nazwie – słaba zbieżność ciągu miar jest własnością raczej wymagającą, wpisującą się w klasyczne ramy topologii przestrzeni metrycznych, o czym przekonamy się w podrozdziale 2.5.5.

Znaczenie funkcji charakterystycznych potwierdza ich ścisły związek ze słabą zbieżnością rozkładów, będący kolejną równoważnością – poczynawszy od definicji 1.1.48, poprzez twierdzenia 1.2.11 i 1.1.54.

Twierdzenie 2.5.7 (Lévy'ego–Craméra, O ciągłości). (i) *Dowolny ciąg rozkładów prawdopodobieństwa $(\mu_n)_{n \in \mathbb{N}}$ jest słabo zbieżny do rozkładu μ wtedy i tylko wtedy, gdy ciąg funkcji charakterystycznych $\varphi_n := \varphi_{\mu_n}$ jest punktowo zbieżny do φ_μ ,*

$$\mu_n \xrightarrow{w} \mu \iff \varphi_n \rightarrow \varphi_\mu.$$

(ii) *Jeśli ciąg funkcji charakterystycznych $(\varphi_n)_{n \in \mathbb{N}}$ jest zbieżny do funkcji $\varphi: \mathbb{R} \rightarrow \mathbb{C}$ ciągłej w zerze, to rozważany ciąg miar jest słabo zbieżny, a $\varphi = \varphi_\mu$ jest funkcją charakterystyczną rozkładu μ takiego, że $\mu_n \xrightarrow{w} \mu$.*

Dowód. (i) Implikacja „ $\Rightarrow$ ” wynika wprost z twierdzenia 1.2.11 i wzoru definiującego funkcje charakterystyczne.

Dla dowodu własności (ii), obejmującej jako przypadek szczególny implikację „ $\Leftarrow$ ”, założmy zbieżność $\varphi_n \rightarrow \varphi$. Jeśli funkcja φ jest ciągła w $u = 0$, to dla dowolnego $\varepsilon > 0$ istnieje $u > 0$ takie, że

$$|1 - \varphi(t)| < \frac{\varepsilon}{2} \quad \text{dla } |t| < u,$$

a zatem $\frac{1}{u} \left| \int_{-u}^u (1 - \varphi(t)) dt \right| < \varepsilon$ i

$$\frac{1}{u} \left| \int_{-u}^u (1 - \varphi_n(t)) dt \right| < \varepsilon$$

dla dostatecznie dużych $n \in \mathbb{N}$. Przetłumaczenie uzyskanego ograniczenia na zbieżność ciągu miar umożliwia

Lemat 2.5.8. *Dowolny rozkład prawdopodobieństwa μ i jego funkcja charakterystyczna φ_μ spełniają nierówność*

$$\mu([-2/u, 2/u]) \geq 1 - \frac{1}{u} \int_{-u}^u (1 - \varphi_\mu(t)) dt \quad \text{dla } u > 0,$$

co w szczególności oznacza, że wskazana całka jest liczbą rzeczywistą.

Dowód lematu. Ponieważ

$$\frac{1}{u} \int_{-u}^u (1 - \varphi_\mu(t)) dt = \frac{1}{u} \int \int_{-u}^u (1 - e^{itx}) dt \mu(dx) = 2 \int \left(1 - \frac{\sin(ux)}{ux}\right) \mu(dx),$$

a wyrażenie podcałkowe jest nieujemne, ograniczenie całki do podzbioru daje

$$\frac{1}{u} \int_{-u}^u (1 - \varphi_\mu(t)) dt \geq 2 \int_{\{x; |xu| > 2\}} \left(1 - \frac{\sin(ux)}{ux}\right) \mu(dx),$$

czyli $\frac{1}{u} \int_{-u}^u (1 - \varphi_\mu(t)) dt \geq \mu\{x; |xu| > 2\} = 1 - \mu([-2/u, 2/u])$. $\square$

Lemat 2.5.9. Ciąg miar $(\mu_n)_{n \in \mathbb{N}}$ w $\mathcal{P}(\mathbb{R})$ o dystrybuantach $F_n = F_{\mu_n}$ zawiera podciąg zbieżny, jeśli dla każdego $\varepsilon > 0$ istnieje liczba $M > 0$, taka że

$$\limsup_{n \rightarrow \infty} (1 - F_n(M) + F_n(-M)) \leq \varepsilon \quad (2.108)$$

lub, równoważnie, $\liminf_{n \rightarrow \infty} \mu_n((-M, M]) \geq 1 - \varepsilon$.

Dowód wg. [6]. Niech $(q_n)_{n \in \mathbb{N}}$ oznacza ustalony ciąg zawierający wszystkie liczby wymierne. Zaczynając od ciągu $m_0 = (j)_{j \in \mathbb{N}}$ i korzystając z indukcji oraz zwartości przedziału $[0, 1]$, możemy dla każdego $k \in \mathbb{N}$ znaleźć ciąg m_k będący podciągiem m_{k-1} i taki, że wartości dystrybuant $F_{m_k(j)}(q_k)$, $j \in \mathbb{N}$, tworzą ciąg zbieżny; przyjmijmy

$$G(q_k) = \lim_{j \rightarrow \infty} F_{m_k(j)}(q_k) \in [0, 1].$$

Ciąg diagonalny o wyrazach $n_k = m_k(k)$, $k \in \mathbb{N}$, jest – począwszy od k -go wyrazu – podciągiem ciągu m_k , dla $k \in \mathbb{N}$, a zatem

$$\lim_{k \rightarrow \infty} F_{n_k}(q) = G(q) \quad \text{dla wszystkich } q \in \mathbb{Q}.$$

Funkcja $F(x) := \inf\{G(q); x < q \in \mathbb{Q}\}$, dla $x \in \mathbb{R}$, jest niemalejąca i prawostronnie ciągła (sprawdzić!). Jeśli $x \in \mathbb{R}$ jest punktem ciągłości funkcji F , to dla dowolnego $\varepsilon > 0$ wybieramy liczby $r', r, s \in \mathbb{Q}$, takie że $r' < r < x < s$ i

$$F(x) - \varepsilon < F(r') \leq F(r) \leq F(x) \leq F(s) < F(x) + \varepsilon.$$

Porównanie granic $F_{n_k}(r) \rightarrow G(r) \geq F(r')$ i $F_{n_k}(s) \rightarrow G(s) \leq F(s)$ prowadzi do nierówności

$$F(x) - \varepsilon < F_{n_k}(r) \leq F_{n_k}(x) \leq F_{n_k}(s) < F(x) + \varepsilon$$

dla dostatecznie dużych $k \in \mathbb{N}$. Zatem $\lim_{k \rightarrow \infty} F_{n_k}(x) = F(x)$.

Powyższa konstrukcja nie korzysta z założenia (2.108) i stanowi treść twierdzenia *Helly'ego*. Dla dowolnego $\varepsilon > 0$, jeśli ciąg dystrybuant i liczba $M = M_\varepsilon$ spełniają warunek (2.108), to dla dowolnych punktów ciągłości $s < -M_\varepsilon$ i $t > M_\varepsilon$

$$0 \leq 1 - F(s) + F(t) = \lim_{k \rightarrow \infty} (1 - F_{n_k}(s) + F_{n_k}(t)) \leq \varepsilon.$$

Z monotoniczności F wynikają nierówności

$$0 \leq \lim_{s \rightarrow -\infty} (1 - F(s)) + \lim_{t \rightarrow \infty} F(t) \leq \varepsilon \quad \text{dla } \varepsilon > 0,$$

a zatem funkcja graniczna F jest dystrybuantą i $F_{n_k} \xrightarrow{w} F$. □

Wracając do dowodu części (ii) twierdzenia 2.5.7 stwierdzamy, że dla dowolnego $\varepsilon > 0$ istnieją $u > 0$ i $m \in \mathbb{N}$, takie że

$$\mu_n([-2/u, 2/u]) > 1 - \varepsilon \quad \text{dla } n \geq m. \quad (2.109)$$

Z lematu 2.5.9 wynika zatem, iż nie tylko ciąg $(\mu_n)_{n \in \mathbb{N}}$, ale też każdy jego podciąg zawiera słabo zbieżny podciąg miar. Granicą każdego ze zbieżnych podciągów jest ta sama miara μ – jedyna miara, której funkcją charakterystyczną jest granica $\varphi = \lim_{k \rightarrow \infty} \varphi_n$. Dla dowolnej funkcji $f \in \mathcal{C}_b(\mathbb{R})$, sformułowana w poprzednim zdaniu własność implikuje zbieżność ciągu liczbowego

$$\lim_{n \rightarrow \infty} \int f d\mu_n = \int f d\mu,$$

co, na mocy twierdzenia 1.2.11, oznacza słabą zbieżność ciągu miar. $\square$

Jeśli miara $\mu = X_*\mathbb{P}$ jest rozkładem dowolnej zmiennej losowej X w przestrzeni $(\Omega, \mathfrak{F}, \mathbb{P})$, to funkcję $\varphi_X \equiv \varphi_\mu$, a więc

$$\varphi_X(t) := \int e^{itX} d\mathbb{P} = \mathcal{E}(e^{itX}) \quad \text{dla } t \in \mathbb{R}, \quad (2.110)$$

po odpowiednim rozszerzeniu pojęcia wartości oczekiwanej, nazywamy *funkcją charakterystyczną zmiennej losowej X* . Z twierdzenia 2.1.32 – i wzorów na sinus i kosinus sumy argumentów – otrzymujemy ważny

Wniosek 2.5.10. *Dla dowolnej zmiennej losowej X zachodzą równości*

- (i) $\varphi_{aX+b}(t) = e^{itb} \varphi_X(at)$ dla $a, b \in \mathbb{R}$, oraz
- (ii) $\varphi_{X+Y}(t) = \varphi_X(t) \varphi_Y(t)$, jeśli zmienne X i Y są niezależne. $\square$

W odniesieniu do miar równość (ii) jest szczególnym przypadkiem ogólnej własności: *transformacja Fouriera* zamienia spłot obiektów na iloczyn ich transformat.

Uwaga 2.5.2. Jeśli wartość oczekiwana $\mathcal{E}(X)$ zmiennej losowej $X: \Omega \rightarrow \mathbb{R}$ istnieje i jest skończona, to funkcja charakterystyczna φ_X jest różniczkowalna w punkcie $t = 0$, a istnienie granicy

$$\lim_{t \rightarrow 0} \frac{\varphi_X(t) - 1}{t} = i \lim_{t \rightarrow 0} \int \frac{e^{ixt} - 1}{it} \mathbb{P}_X(dx) = i \int x \mathbb{P}_X(dx)$$

wynika wówczas z twierdzenia Lebesgue’a o ograniczonej zbieżności. Zatem

$$\varphi'_X(0) = i\mathcal{E}(X), \quad \text{jeśli } \mathcal{E}|X| < \infty. \quad (2.111)$$

2.5.3 Czas na Centralne twierdzenie graniczne

Dalsze własności funkcji charakterystycznych i wynikające z nich wnioski, w tym słynne *Centralne twierdzenie graniczne*, wymagają pomocniczych oszacowań w dziedzinie liczb zespolonych. Podstawową rolę pełni tu

Stwierdzenie 2.5.11. *Jeśli $\mathcal{E}(X^2) < \infty$, to*

$$\varphi_X(t) = 1 + it\mathcal{E}(X) - \frac{t^2}{2}\mathcal{E}(X^2) + t^2r(t), \quad (2.112)$$

gdzie $\lim_{t \rightarrow 0} r(t) = 0$.

Dowód. Klasyczne wzory Taylora z resztą w postaci całkowej

$$e^{ix} = 1 + ix + i^2 \int_0^x (x-t)e^{it} dt = 1 + ix + \frac{i^2 x^2}{2} + \frac{i^3}{2} \int_0^x (x-t)^2 e^{it} dt$$

prowadzą do oszacowania

$$\begin{aligned} \left| e^{ix} - \left(1 + ix - \frac{x^2}{2}\right) \right| &\leq \left| \int_0^x |x-t| dt \right| + \frac{x^2}{2} = x^2, \quad \text{oraz} \\ \left| e^{ix} - \left(1 + ix - \frac{x^2}{2}\right) \right| &\leq \frac{1}{2} \left| \int_0^x |x-t|^2 dt \right| = \frac{|x|^3}{6}, \end{aligned}$$

dla $x \in \mathbb{R}$. Stąd

$$\begin{aligned} \frac{1}{t^2} \left| \varphi_X(t) - \left(1 + it\mathcal{E}(X) - \frac{t^2}{2}\mathcal{E}(X^2)\right) \right| &\leq \frac{1}{t^2} \int \left| e^{itx} - \left(1 + itx - \frac{t^2 x^2}{2}\right) \right| \mathbb{P}_X(dx) \\ &\leq \frac{1}{t^2} \int \min(|tx|^2, \frac{|tx|^3}{6}) \mathbb{P}_X(dx) \\ &= \int x^2 \min\left(1, \frac{|tx|}{6}\right) \mathbb{P}_X(dx) \end{aligned}$$

dla $t \in \mathbb{R} \setminus \{0\}$, gdzie ostatnia całka jest ograniczona przez $\mathcal{E}(X^2) < \infty$ i dąży do 0 dla $t \rightarrow 0$. $\square$

Rozszerzamy także własność znaną w wariancie rzeczywistym jako lemat 2.4.3(ii).

Stwierdzenie 2.5.12. *Jeśli $c_n \rightarrow c \in \mathbb{C}$, to $(1 + c_n/n)^n \rightarrow e^c$.*

Dowód. Niech $C := \sup_n |c_n| < \infty$. Dla liczb zespolonych $z = e^{c_n/n}$ oraz $w = 1 + \frac{c_n}{n}$ korzystamy z elementarnej nierówności

$$|z^n - w^n| \leq n\theta^{n-1}|z - w|, \quad \text{gdzie } \theta = \max\{|z|, |w|\},$$

a ponieważ $|z| = e^{\Re c_n/n} \leq e^{C/n}$ i $|w| \leq 1 + \frac{C}{n} \leq e^{C/n}$, za wspólne oszacowanie modułów możemy przyjąć $\theta = e^{C/n}$. Stąd dla $n > C$

$$\left| e^{c_n} - \left(1 + \frac{c_n}{n}\right)^n \right| \leq n e^{\frac{C(n-1)}{n}} \left| e^{\frac{c_n}{n}} - \left(1 + \frac{c_n}{n}\right) \right| \leq n e^C \left(\frac{C}{n}\right)^2 \longrightarrow 0,$$

gdzie w ostatniej nierówności wykorzystujemy rozwinięcie w szereg,

$$|e^b - (1+b)| = \left| \sum_{k \geq 2} \frac{b^k}{k!} \right| \leq \frac{|b|^2}{2} \left(1 + \frac{1}{2} + \frac{1}{2^2} + \dots\right) = |b|^2,$$

o ile $|b| \leq 1$. □

Podsumowując, otrzymujemy stąd tytułowe dla bieżącego podrozdziału

Twierdzenie 2.5.13 (Centralne twierdzenie graniczne). *Dla dowolnego ciągu $(X_n)_{n \in \mathbb{N}}$ niezależnych zmiennych losowych o tym samym rozkładzie, wartości oczekiwanej $\mathcal{E}(X_1) = 0$ i skończonej wariancji $\sigma^2 = \text{Var}(X_1)$, rozkłady prawdopodobieństwa zmiennych $\frac{1}{\sqrt{n}}S_n$ dążą do rozkładu normalnego,*

$$\frac{1}{\sqrt{n}}(X_1 + \dots + X_n) \xrightarrow{d} \mathcal{N}(0, \sigma^2) \quad \text{dla } n \rightarrow \infty.$$

Podana wyżej zbieżność wg. rozkładu zakłada, że rolę zmiennej granicznej pełni dowolna zmienna losowa o wskazanym rozkładzie normalnym – niekoniecznie istniejąca w oryginalnej przestrzeni Ω . Równoważny, bardziej konkretny zapis dla dystrybuant oznacza własność

$$\lim_{n \rightarrow \infty} \mathbb{P}\left\{a \leq \frac{1}{\sqrt{n}}(X_1 + \dots + X_n) \leq b\right\} = \frac{1}{\sqrt{2\pi}\sigma} \int_a^b e^{-\frac{x^2}{2\sigma^2}} dx \quad (2.113)$$

dla dowolnych liczb $a \leq b$.

Dowód. Dla zmiennych $\tilde{S}_n = \frac{1}{\sqrt{n}}S_n, n \in \mathbb{N}$, z wniosku 2.5.10 i stwierdzenia 2.5.11 wynikają równości

$$\begin{aligned} \varphi_{\tilde{S}_n}(t) &= \prod_{k \leq n} \varphi_{X_k}\left(\frac{t}{\sqrt{n}}\right) = \varphi_{X_1}\left(\frac{t}{\sqrt{n}}\right)^n \\ &= \left(1 + \frac{t^2}{n} \left(\frac{-\sigma^2}{2} + r\left(\frac{t}{\sqrt{n}}\right)\right)\right)^n \longrightarrow e^{-\frac{t^2 \sigma^2}{2}} \quad \text{dla } t \in \mathbb{R}, \end{aligned}$$

przy czym zbieżność do funkcji wykładniczej jest konsekwencją stwierdzenia 2.5.12. W szczególnym przypadku, dla niezależnych zmiennych losowych o rozkładzie (2.49), gdy $\tilde{S}_n \xrightarrow{d} \mathcal{N}(0, 1)$, z twierdzenia 2.4.1 (de Moivre’a – Laplace’a) i twierdzenia 2.5.7 (o ciągłości) otrzymujemy równość

$$\varphi_{\mathcal{N}(0,1)}(t) = e^{-\frac{t^2}{2}} \quad \text{dla } t \in \mathbb{R}.$$

Ponieważ $\mathcal{N}(0, \sigma^2)$ jest wynikiem skalowania standardowego rozkładu normalnego, funkcja charakterystyczna $\varphi_{\mathcal{N}(0, \sigma^2)}$ jest równa otrzymanej wyżej funkcji granicznej – w ogólnym przypadku, gdy $\mathcal{E}(X_1^2) = \text{Var}(X_1) = \sigma^2$. $\square$

Wniosek 2.5.14. *Funkcją charakterystyczną rozkładu $\mathcal{N}(x, a)$ jest*

$$\varphi(t) = e^{itx - \frac{t^2 a}{2}} \quad \text{dla } t \in \mathbb{R}. \quad (2.114)$$

Dowód. Ćwiczenie. Skorzystać z wniosku 2.5.10(i). $\square$

Ćwiczenie 2.5.3. Wyznaczyć funkcję charakterystyczną (2.114) rozkładu normalnego bezpośrednio z definicji, jako całkę.

Przypominając, że przy przyjętych wyżej założeniach dotyczących ciągu $X_n, n \in \mathbb{N}$, nierówność Kołmogorowa (2.61) przyjmuje postać

$$\mathbb{P}\left\{\max_{k \leq n} |X_1 + \dots + X_k| > \lambda \sigma \sqrt{n}\right\} \leq n \frac{\sigma^2}{(\lambda \sigma \sqrt{n})^2} = \frac{1}{\lambda^2},$$

możemy z CTG wywnioskować mocniejszą własność

Stwierdzenie 2.5.15. *Dla dowolnego ciągu $X_n, n \in \mathbb{N}$, niezależnych zmiennych losowych o tym samym rozkładzie, wartości oczekiwanej 0 i wariancji $\text{Var}(X_k) = \sigma^2 < \infty$, oraz dla dowolnego $m \in \mathbb{N}$ zachodzi równość*

$$\lim_{\lambda \rightarrow \infty} \limsup_{n \rightarrow \infty} \lambda^m \mathbb{P}\left\{\max_{k \leq n} |X_1 + \dots + X_k| > \lambda \sigma \sqrt{n}\right\} = 0. \quad (2.115)$$

Dowód. Odrzucając trywialny przypadek $\sigma = 0$ i zastępując X_n przez X_n/σ , dla $n \in \mathbb{N}$, możemy założyć, że $\sigma = 1$. Ustalmy $n > 1$ oraz dowolne $\lambda > 0$. Podobnie jak w dowodzie lematu 2.2.12 korzystamy z rozkładu

$$\left\{\max_{k \leq n} |S_k| > x\right\} = \bigcup_{k \leq n} A_k, \quad \text{dla } A_k = \left\{\max_{j < k} |S_j| \leq x, |S_k| > x\right\}, \quad k \leq n,$$

gdzie $x \equiv \lambda \sqrt{n}$. Dla dowolnego $c \in (0, \lambda)$ korzystamy z nierówności

$$\begin{aligned} \mathbb{P}\left\{\max_{k \leq n} |S_k| > \lambda \sqrt{n}\right\} &\leq \mathbb{P}\{|S_n| > (\lambda - c) \sqrt{n}\} \\ &\quad + \sum_{k < n} \mathbb{P}\{\omega \in A_k; |S_n(\omega)| \leq (\lambda - c) \sqrt{n}\} \\ &\leq \mathbb{P}\{|S_n| > (\lambda - c) \sqrt{n}\} \\ &\quad + \sum_{k < n} \mathbb{P}\{\omega \in A_k; |S_n(\omega) - S_k(\omega)| > c \sqrt{n}\}. \end{aligned}$$

Ostatnia grupa składników pozwala na oszacowania

$$\begin{aligned}\mathbb{P}\{\omega \in A_k; |S_n(\omega) - S_k(\omega)| > c\sqrt{n}\} &\leq \frac{1}{c^2 n} \int 1_{A_k} |S_n - S_k|^2 d\mathbb{P} \\ &= \frac{1}{c^2 n} \mathbb{P}(A_k) \sum_{i=k+1}^n \mathcal{E}(X_i^2) \leq \frac{1}{c^2} \mathbb{P}(A_k),\end{aligned}$$

dla $k < n$, a zatem

$$\mathbb{P}\{\max_{k \leq n} |S_k| > \lambda\sqrt{n}\} \leq \mathbb{P}\{|S_n| > (\lambda - c)\sqrt{n}\} + \frac{1}{c^2} \mathbb{P}\{\max_{k < n} |S_k| > \lambda\sqrt{n}\}, \quad (2.116)$$

co w szczególności, dla $c = 2 < \lambda$, daje nierówność

$$\mathbb{P}\{\max_{k \leq n} |S_k| > \lambda\sqrt{n}\} \leq \frac{4}{3} \mathbb{P}\{|S_n| > (\lambda - 2)\sqrt{n}\} \quad \text{dla } n \in \mathbb{N}.$$

W granicy otrzymujemy stąd – i z twierdzenia 2.5.13

$$\limsup_{n \rightarrow \infty} \mathbb{P}\{\max_{k \leq n} |S_k| > \lambda\sqrt{n}\} \leq \frac{4}{3} \lim_{n \rightarrow \infty} \mathbb{P}\{|\frac{S_n}{\sqrt{n}}| > \lambda - 2\} = \frac{4}{3} \mathbb{P}\{|N| > \lambda - 2\}$$

dla dowolnej zmiennej N o rozkładzie normalnym $\mathcal{N}(0, 1)$. Teza wynika z oszacowania $\mathbb{P}\{|N| > \lambda - 2\} \leq \mathcal{E}(|N|^{m+1})/(\lambda - 2)^{m+1}$. $\square$

2.5.4 ...i wariant wektorowy

Ograniczając się do wektorów losowych $X = (X_1, \dots, X_p): \Omega \rightarrow \mathbb{R}^p$, $p \in \mathbb{N}$, przyjmujemy

Definicja 2.5.16. Funkcję charakterystyczną $\varphi_X: \mathbb{R}^p \rightarrow \mathbb{C}$ dowolnej wektorowej zmiennej losowej $X: \Omega \rightarrow \mathbb{R}^p$ określa równość

$$\varphi_X(t) = \mathcal{E}(e^{i\langle t, X \rangle}) = \int e^{i\langle t, x \rangle} \mathbb{P}_X(dx) \quad (2.117)$$

dla $t \in \mathbb{R}^p$, gdzie $\langle \cdot, \cdot \rangle$ oznacza euklidesowy iloczyn skalarny w $\mathbb{R}^p$, a więc $\langle t, X \rangle = t_1 X_1 + \dots + t_p X_p$.

Technika sprawdzona w podrozdziale 2.5.1 pozwala wykazać

Twierdzenie 2.5.17 (Transformata odwrotna II). Funkcja charakterystyczna φ_X wyznacza jednoznacznie rozkład $\mathbb{P}_X$ zmiennej losowej X . Dokładniej, dla dowolnego przedziału $A = [a_1, b_1] \times \dots \times [a_p, b_p]$ takiego, że $\mathbb{P}_X(\partial A) = 0$ zachodzi równość

$$\mathbb{P}_X(A) = \lim_{T \rightarrow \infty} \frac{1}{(2\pi)^p} \int_{[-T, T]^p} \prod_{j \leq p} \frac{e^{-it_j a_j} - e^{-it_j b_j}}{it_j} \varphi_X(t) dt. \quad (2.118)$$

Dowód. Z twierdzenia Fubiniego wynika, że całka po prawej stronie (2.118) jest równa

$$\begin{aligned} & \int_{[-T,T]^p} \int \prod_{j \leq p} \frac{e^{-it_j a_j} - e^{-it_j b_j}}{it_j} e^{i\langle t, x \rangle} \mathbb{P}_X(dx) dt \\ &= \int \prod_{j \leq p} \int_{-T}^T \frac{e^{-it_j a_j} - e^{-it_j b_j}}{it_j} e^{it_j x_j} dt_j \mathbb{P}_X(dx) = \int \prod_{j \leq p} \psi_j(x_j, T) \mathbb{P}_X(dx), \end{aligned}$$

gdzie funkcje $\psi_j(x_j, T)$, $j \leq p$, są wspólnie ograniczone i zgodnie z (2.104)

$$\begin{aligned} \lim_{T \rightarrow \infty} \prod_{j \leq p} \psi_j(x_j, T) &= \pi^p \prod_{j \leq p} (1_{(a_j, b_j)}(x_j) + 1_{[a_j, b_j]}(x_j)) \\ &= (2\pi)^p 1_A(x) \end{aligned}$$

dla $\mathbb{P}_X$ –prawie wszystkich $x \in \mathbb{R}^p$. Równość (2.118) wynika z twierdzenia Lebesgue’a o ograniczonej zbieżności. Dla zakończenia dowodu wystarczy zauważyć, że warunek $\mathbb{P}_X(\partial A) = 0$ wyklucza z rozważań co najwyżej przeliczalną ilość wartości a_j, b_j dla każdej z osi $j \leq p$, a *dopuszczalne* przedziały stanowią π –system generujący σ –ciało zbiorów borelowskich. $\square$

Z tożsamości $\varphi_X(t) \equiv \varphi_{\langle t, X \rangle}(1)$ wynika zatem

Wniosek 2.5.18. *Zmienne losowe $X, Y: \Omega \rightarrow \mathbb{R}^p$ mają ten sam rozkład $\mathbb{P}_X = \mathbb{P}_Y$ wtedy i tylko wtedy, gdy dla każdego $a \in \mathbb{R}^p$ zachodzi równość $\mathbb{P}_{a^t X} = \mathbb{P}_{a^t Y}$ odpowiednich 1–wymiarowych rozkładów brzegowych.* $\square$

Jak sygnalizowaliśmy wcześniej, wniosek ten uzupełnia dowód twierdzenia 2.4.11, a jego niebanalność pokazuje, że teoria rozkładów p –wymiarowych jest istotnie bogatsza od wariantu 1–wymiarowego. Stosowne twierdzenia graniczne – analogiczne do sformułowanych w poprzednich podrozdziałach – wymagają niewielkich modyfikacji wcześniejszych dowodów.

Naturalne p –wymiarowe uogólnienie twierdzenia Lévy’ego-Craméra (por. tw. 2.5.7) formułujemy w postaci

Twierdzenie 2.5.19 (O ciągłości – w $\mathbb{R}^p$). *Następujące własności ciągu zmiennych wektorowych $X, X_n: \Omega \rightarrow \mathbb{R}^p$, $n \in \mathbb{N}$, są równoważne;*

- (i) $X_n \xrightarrow{d} X$,
- (ii) $a^t X_n \xrightarrow{d} a^t X$ dla każdego $a \in \mathbb{R}^p$, oraz
- (iii) $\varphi_{X_n}(t) \rightarrow \varphi_X(t)$ dla każdego $t \in \mathbb{R}^p$.

Dowód. Implikacja (i) $\Rightarrow$ (ii) wynika z tożsamości $\int g d\mathbb{P}_{a^t X_n} = \int g \circ L_a d\mathbb{P}_{X_n}$ dla $g \in \mathcal{C}_b(\mathbb{R})$, gdzie $L_a: \mathbb{R}^p \rightarrow \mathbb{R}$ jest iloczynem przez wektor a (por. stwierdzenie 1.2.12). Z kolei twierdzenie 2.5.7 pozwala z (ii) wywnioskować zbieżność $\varphi_{\langle t, X_n \rangle}(1) \rightarrow \varphi_{\langle t, X \rangle}(1)$ dla każdego $t \in \mathbb{R}^p$, co oznacza własność (iii).

(iii) $\Rightarrow$ (i) Załóżmy, że ciąg funkcji charakterystycznych φ_{X_n} , $n \in \mathbb{N}$, jest punktowo zbieżny do φ_X . Dla dowolnego $\varepsilon > 0$, niech $u > 0$ określa przedział $[-u, u]^p$ na którym $|1 - \varphi_X| < \frac{\varepsilon}{2}$, a więc także $\int_{[-u, u]^p} |1 - \varphi_X| d\ell^p < (2u)^p \frac{\varepsilon}{2}$ oraz, co wynika z twierdzenia Lebesgue'a o ograniczonej zbieżności,

$$\int_{[-u, u]^p} |1 - \varphi_{X_n}(x)| dx < (2u)^p \frac{\varepsilon}{2} \quad \text{dla dostatecznie dużych } n \in \mathbb{N}.$$

Przekształcenia analogiczne do wykorzystanych w dowodzie lematu 2.5.8 prowadzą do oszacowania – jak się okazuje – liczby rzeczywistej

$$\begin{aligned} \int_{[-u, u]^p} (1 - \varphi_{X_n}(t)) dt &= (2u)^p \int \left(1 - \prod_{i \leq p} \frac{\sin(ux_i)}{ux_i}\right) \mathbb{P}_{X_n}(dx) \\ &\geq (2u)^p \int_{B_u} \left(1 - \frac{1}{2}\right) d\mathbb{P}_{X_n} = \frac{(2u)^p}{2} \mathbb{P}_{X_n}(B_u), \end{aligned}$$

gdzie dla dowolnego $u > 0$ wszystkie czynniki są ograniczone przez 1, a na zbiorze $B_u = \mathbb{R}^p \setminus [-\frac{2}{u}, \frac{2}{u}]^p$ przynajmniej jeden jest $\leq \frac{1}{2}$. Istnieje zatem liczba $m \in \mathbb{N}$, taka że

$$\mathbb{P}_{X_n}([- \frac{2}{u}, \frac{2}{u}]^p) > 1 - \varepsilon \quad \text{dla } n \geq m. \quad (2.119)$$

Udowodniona własność, analogiczna do (2.109), pozwala na powtórzenie konstrukcji wykorzystanej w dowodzie lematu 2.5.9. Po stosownych modyfikacjach stwierdzamy, że ciąg dystrybuant $(F_{X_n})_{n \in \mathbb{N}}$ – i każdy jego podciąg – zawiera podciąg słabo zbieżny do funkcji $F: \mathbb{R}^p \rightarrow [0, 1]$ spełniającej warunki podane w twierdzeniu 1.1.45, a więc będącej dystrybuantą pewnej miary probabilistycznej $\mathbb{P}$. W szczególności, własność (ii) normalizująca graniczną funkcję F wynika z (2.119) (sprawdzić!).

Ponieważ wskazany podciąg ciągu funkcji charakterystycznych pozostaje zbieżny do φ_X , z twierdzenia 2.5.17 wynika równość $\mathbb{P} = \mathbb{P}_X$. Dla dowolnej funkcji $g \in \mathcal{C}_b(\mathbb{R}^p)$ zbieżność $\mathcal{E}(g \circ X_n) \rightarrow \mathcal{E}(g \circ X)$, jest konsekwencją udowodnionej własności ciągu liczbowego $(\mathcal{E}(g \circ X_n))_{n \in \mathbb{N}}$, którego każdy podciąg zawiera podciąg zbieżny do $\mathcal{E}(g \circ X)$. $\square$

Cykl rozważań dotyczących rozkładów p -wymiarowych kończy

Twierdzenie 2.5.20 (Centralne twierdzenie graniczne w $\mathbb{R}^p$). *Dla każdego ciągu niezależnych p -wymiarowych wektorów losowych $(X_n)_{n \in \mathbb{N}}$ o tym samym rozkładzie, wartości oczekiwanej $\mu = \mathcal{E}(X_n) \in \mathbb{R}^p$ i nieosobliwej macierzy kowariancji $W = \text{Cov}(X_n) \in \mathbb{R}^{p \times p}$ zachodzi zbieżność*

$$\sqrt{n} \left(\frac{X_1 + \dots + X_n}{n} - \mu \right) \xrightarrow{d} \mathcal{N}_p(0, W). \quad (2.120)$$

W odniesieniu do rozkładów (miar) własność (2.120) oznacza równość

$$\lim_{n \rightarrow \infty} \mathbb{P} \left\{ \sqrt{n} \left(\frac{X_1 + \dots + X_n}{n} - \mu \right) \in A \right\} = \frac{1}{(2\pi)^{p/2} \sqrt{\det(W)}} \int_A e^{-\frac{x^t W^{-1} x}{2}} dx$$

dla każdego zbioru borelowskiego $A \subset \mathbb{R}^p$.

Dowód. Jeśli X oznacza dowolną wektorową zmienną losową o rozkładzie $\mathcal{N}_p(0, W)$, to dla dowolnego wektora $a \in \mathbb{R}^p$ zmienna $a^t X$ ma rozkład Gaussa $\mathcal{N}(0, a^t W a)$ i z twierdzenia 2.5.13 dla ciągu $a^t(X_n - \mu)$, $n \in \mathbb{N}$, wynika zbieżność wg. rozkładu

$$a^t \frac{X_1 + \dots + X_n - n\mu}{\sqrt{n}} \xrightarrow{d} a^t X.$$

Teza wynika z twierdzenia 2.5.19 o ciągłości w $\mathbb{R}^p$. $\square$

Ćwiczenie 2.5.4. Wywnioskować wzór na funkcję charakterystyczną rozkładu $\mathcal{N}_p(\mu, W)$. Sprawdzić, że CTG w $\mathbb{R}^p$ obejmuje także przypadek, gdy wspólna macierz kowariancji $\text{Cov}(X_n)$ jest osobliwa, a wektory są *pw.* linio-wo zależne.

2.5.5 Metryzowalność

Nich (M, ρ) oznacza dowolną *ośrodkową* przestrzeń metryczną. Jeśli chcemy, by odwzorowania mierzalne $(\Omega, \mathfrak{F}) \rightarrow (M, \mathfrak{B}_M)$ tworzyły przestrzeń liniową, możemy dodatkowo założyć, że (M, ρ) ma strukturę *topologicznej przestrzeni liniowej*, co oznacza ciągłość działań: mnożenia $\mathbb{R} \times M \rightarrow M$ i dodawania $M \times M \rightarrow M$. Założenia te obejmują oczywiście p -wymiarowe zmienne losowe dla $M = \mathbb{R}^p$, $p \in \mathbb{N}$, a także ogólniejsze przestrzenie funkcji ciągłych (por. ćwiczenie 3.2.2), oraz rozważaną w podrozdziale 2.6.4 przestrzeń $\mathcal{D}$.

Definicja 2.5.21. M -wartościową zmienną losową w przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$ nazywamy każde odwzorowanie $X: \Omega \rightarrow M$ mierzalne względem σ -ciała zbiorów borelowskich w metryzowalnej liniowej przestrzeni topologicznej M .

Indukowaną miarę borelowską $\mathbb{P}_X = X_*\mathbb{P}$ na M nazywamy rozkładem prawdopodobieństwa zmiennej X .

Z założenia ośrodkowości przestrzeni M wynika

Stwierdzenie 2.5.22. (i) Odległość $d(X, Y): \Omega \rightarrow \mathbb{R}$ dowolnych zmiennych losowych $X, Y: \Omega \rightarrow M$ jest zmienną losową.

(ii) Jeśli M jest także liniową przestrzenią topologiczną, to kombinacja liniowa $aX + bY$ dowolnych M -wartościowych zmiennych losowych jest zmienną losową (jest mierzalna).

Dowód. Rozważane funkcje są złożeniami odwzorowania $(X, Y): \Omega \rightarrow M^2$ i funkcji ciągłych określonych na produkcie. Równość $(M^2, \mathfrak{B}_{M^2}) = (M, \mathfrak{B}_M)^2$ prawdziwa dla przestrzeni ośrodkowych [2, wn. 2.2.8] kończy dowód. $\square$

Rozszerzając definicję 2.5.6 przyjmujemy

Definicja 2.5.23. Ciąg M -wartościowych zmiennych losowych $(X_n)_{n \in \mathbb{N}}$ dąży według rozkładu do zmiennej X , jeśli ciąg rozkładów zmiennych X_n jest słabo zbieżny do rozkładu $\mathbb{P}_X$ zmiennej X . Zbieżność wg. rozkładu zapisujemy w postaci $X_n \xrightarrow{d} X$ lub $X_n \xrightarrow{d} \mathbb{P}_X$.

Wykażemy, że zbiór $\mathcal{P}(M)$ wszystkich borelowskich miar probabilistycznych na M nosi naturalną strukturę metryzowalnej przestrzeni topologicznej, w której słaba zbieżność jest tożsama ze zbieżnością wg. metryki.

Definicja 2.5.24. Odległością Prochorowa dowolnych miar probabilistycznych $\lambda, \mu \in \mathcal{P}(M)$ nazywamy liczbę

$$d_\rho(\lambda, \mu) = \inf\{\varepsilon > 0; \lambda(A) \leq \mu(A^\varepsilon) + \varepsilon \text{ dla } A \in \mathfrak{B}_M\}, \quad (2.121)$$

gdzie $A^\varepsilon = \{x \in M; \rho(x, A) < \varepsilon\}$ jest ε -otoczeniem zbioru $A \in \mathfrak{B}_M$.

Funkcję d_ρ nazywamy metryką Prochorowa.

Sprawdzenia wymaga oczywiście

Lemat 2.5.25. Funkcja opisana wzorem (2.121) jest metryką w $\mathcal{P}(M)$.

Dowód. Dla dowolnych miar $\lambda, \mu \in \mathcal{P}(M)$ symetrię zapewnia równoważność $A \cap B^\varepsilon = \emptyset \Leftrightarrow A^\varepsilon \cap B = \emptyset$. Jeśli zatem dla każdego zbioru borelowskiego $B \subset M$ zachodzi nierówność $\lambda(B) \leq \mu(B^\varepsilon) + \varepsilon$, to przyjmując $B = M \setminus A^\varepsilon$ dla dowolnego $A \in \mathfrak{B}_M$, otrzymujemy

$$\lambda(A^\varepsilon) = 1 - \lambda(B) \geq 1 - \mu(B^\varepsilon) - \varepsilon = \mu(M \setminus B^\varepsilon) - \varepsilon \geq \mu(A) - \varepsilon.$$

Jeśli miary $\lambda, \mu \in \mathcal{P}(M)$ mają odległość $d_\rho(\lambda, \mu) = 0$, to dla każdego podzbioru domkniętego $F \subset M$ z nierówności $\lambda(F) \leq \mu(F^\varepsilon) + \varepsilon$ dla $\varepsilon \rightarrow 0$ wynika nierówność $\lambda(F) \leq \mu(F)$, a więc (z symetrii) równość $\lambda(F) = \mu(F)$. Równość miar $\lambda = \mu$ jest konsekwencją ich regularności (tw. 1.1.47).

Dowód nierówności trójkąta dla d_ρ pozostawiamy czytelnikowi. $\square$

Ćwiczenie 2.5.5. Na prostej $\mathbb{R}$ równoważność rozkładów prawdopodobieństwa i dystrybuant pozwala na zdefiniowanie metryki Levy'ego, takiej że

$$d_L(F, G) = \inf\{\varepsilon; F(x - \varepsilon) - \varepsilon \leq G(x) \leq F(x + \varepsilon) + \varepsilon \text{ dla } x \in \mathbb{R}\} \quad (2.122)$$

dla dowolnych dystrybuant F i G . Sprawdzić równoważność metryk – Levy'ego i Prochorowa.

Ćwiczenie 2.5.6. Wykazać (por. [3, Ćw. 26.15-16]), że funkcja

$$d_\varphi(\mu, \nu) = \sup \left\{ \frac{|\varphi_\mu(t) - \varphi_\nu(t)|}{1 + |t|}; t \in \mathbb{R} \right\} \quad \text{dla } \mu, \nu \in \mathcal{P}(\mathbb{R})$$

jest metryką zupełną równoważną z metryką Lévy'ego.

Twierdzenie 2.5.26. W ośrodkowej przestrzeni metrycznej (M, ρ) , dla dowolnych miar probabilistycznych $\mu, \mu_n \in \mathcal{P}(M)$, $n \in \mathbb{N}$,

$$d_\rho(\mu_n, \mu) \rightarrow 0 \iff \mu_n \xrightarrow{w} \mu. \quad (2.123)$$

Dowód. „ $\Rightarrow$ ” Jeśli $d_\rho(\mu_n, \mu) < \varepsilon_n$ i $\varepsilon_n \rightarrow 0$, to dla dowolnego zbioru domkniętego $F \subset M$ własność

$$\limsup_{n \rightarrow \infty} \mu_n(F) \leq \lim_{n \rightarrow \infty} (\mu(F^{\varepsilon_n}) + \varepsilon_n) = \mu(F)$$

oznacza słabą zbieżność ciągu miar.

„ $\Leftarrow$ ” Niech $\varepsilon > 0$; ustalamy rozkład $M = \bigcup_{n \in \mathbb{N}} A_n$ na przeliczalną rodzinę rozłącznych zbiorów borelowskich o średnicy mniejszej niż ε i liczbę $k \in \mathbb{N}$, taką że $\mu(\bigcup_{i > k} A_i) < \varepsilon$. Dla dowolnego zbioru $A \in \mathfrak{B}_M$ zachodzi inkluzja

$$A \subset A_0 \cup \bigcup_{i > k} A_i, \quad \text{gdzie } A_0 = \bigcup \{A_i; i \leq k, A_i \cap A \neq \emptyset\} \subset A^\varepsilon.$$

Skończona rodzina ε -otoczeń $\mathcal{G} = \{(A_{i_1} \cup \dots \cup A_{i_m})^\varepsilon; 1 \leq i_1 \dots i_m \leq k\}$ pozwala na wybór indeksu $n_0 \in \mathbb{N}$, takiego że słaba zbieżność $\mu_n \xrightarrow{w} \mu$ implikuje nierówność

$$\mu_n(U) > \mu(U) - \varepsilon \quad \text{dla każdego } U \in \mathcal{G} \text{ i } n \geq n_0.$$

Dla $n \geq n_0$ otrzymujemy zatem

$$\mu(A) \leq \mu(A_0^\varepsilon) + \varepsilon < \mu_n(A_0^\varepsilon) + 2\varepsilon \leq \mu_n(A^{2\varepsilon}) + 2\varepsilon,$$

czyli $d_\rho(\mu_n, \mu) \leq 2\varepsilon$. □

Ponieważ słaba zbieżność ciągu miar borelowskich jest pojęciem *topologicznym* – nie zależy od wyboru konkretnej metryki w metryzowalnej przestrzeni M , ta sama własność przysługuje wyznaczonej w $\mathcal{P}(M)$ topologii.

Ćwiczenie 2.5.7. Wykazać, że metryzowalna przestrzeń topologiczna $\mathcal{P}(M)$ jest także ośrodkowa.

Twierdzenie 2.5.27. Dla dowolnej przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$ i ośrodkowej przestrzeni metrycznej (M, ρ) niech $X_n, Y_n: (\Omega, \mathfrak{F}) \rightarrow (M, \mathfrak{B}_M)$, $n \in \mathbb{N}$, będą dwoma ciągami funkcji mierzalnych. Jeśli

$$X_n \xrightarrow{d} \mu \quad i \quad \rho(X_n, Y_n) \xrightarrow{p} 0$$

dla pewnej miary borelowskiej μ na M , to także $Y_n \xrightarrow{d} \mu$.

W szczególności, jeśli ciąg zmiennych $(X_n)_{n \in \mathbb{N}}$ jest zbieżny wg. prawdopodobieństwa do $X: \Omega \rightarrow \mathcal{C}(T)$, czyli

$$\lim_{n \rightarrow \infty} \mathbb{P}\{\omega \in \Omega; \rho(X_n(\omega), X(\omega)) \geq \varepsilon\} = 0 \quad \text{dla } \varepsilon > 0,$$

to $X_n \xrightarrow{d} X$.

Dowód. Dla $\varepsilon > 0$ i domkniętego ε -otoczenia F_ε dowolnego zbioru domkniętego $F \subset M$, z implikacji

$$Y_n(\omega) \in F \implies (X_n(\omega) \in F_\varepsilon \text{ lub } \rho(X_n(\omega), Y_n(\omega)) > \varepsilon)$$

wynikają nierówności $\mathbb{P}\{y_n \in F\} \leq \mathbb{P}\{X_n \in F_\varepsilon\} + \mathbb{P}\{\rho(X_n(\omega), Y_n(\omega)) > \varepsilon\}$, a więc

$$\limsup_{n \rightarrow \infty} \mathbb{P}\{Y_n \in F\} \leq \limsup_{n \rightarrow \infty} \mathbb{P}\{X_n \in F_\varepsilon\} \leq \mathbb{P}\{X \in F_\varepsilon\},$$

a to oznacza zbieżność wg. rozkładu $Y_n \xrightarrow{d} X$. □

2.6 Krok w stronę procesów stochastycznych

2.6.1 Definicja – czyli twierdzenie Kołmogorowa

Definicja 2.6.1. Dla dowolnego zbioru parametrów $T \neq \emptyset$, interpretowanego jako czas, procesem stochastycznym na przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$ nazywamy dowolną rodzinę zmiennych losowych $X_t: \Omega \rightarrow \mathbb{R}$, $t \in T$. Procesy stochastyczne $(X_t)_{t \in T}$ oraz $(X'_t)_{t \in T}$ na tej samej przestrzeni będziemy uważać za równoważne, jeśli $X_t \stackrel{\text{pw}}{=} X'_t$ dla każdego $t \in T$.

W przypadku procesów równoważnych mówi się także, że jeden jest modyfikacją drugiego. Jedną z klasycznych interpretacji procesu stochastycznego $(X_t)_{t \in T}$ jest odwzorowanie mierzalne $X: (\Omega, \mathfrak{F}) \rightarrow (\mathbb{R}, \mathfrak{B}_\mathbb{R})^T$.

Ćwiczenie 2.6.1. Sprawdzić, że $\mathfrak{F}/\mathfrak{B}_\mathbb{R}^T$ -mierzalność odwzorowania

$$\Omega \ni \omega \mapsto (X_t(\omega))_{t \in T} \in \mathbb{R}^T$$

jest równoważna $\mathfrak{F}/\mathfrak{B}_\mathbb{R}$ -mierzalności funkcji X_t , $t \in T$.

Definicja 2.6.2. *Miarę indukowaną $\mathbb{P}_X = X_*\mathbb{P}$ na przestrzeni $(\mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T)$ nazywamy rozkładem prawdopodobieństwa procesu $X = (X_t)_{t \in T}$.*

Stwierdzenie 2.6.3. *Równoważne procesy stochastyczne mają ten sam rozkład prawdopodobieństwa.*

Dowód. Ćwiczenie. □

Jako odwzorowania mierzalne, procesy równoważne nie są na ogół równe prawie wszędzie, mają jednak ten sam rozkład prawdopodobieństwa. W podstawowych zastosowaniach czas T jest półprostą $\mathbb{R}_+ = [0, \infty)$ lub jej przeliczalnym podzbiorem – dla tzw. procesów dyskretnych. Analogicznie jak dla 1-wymiarowych zmiennych losowych i ich rozkładów zachodzi równie oczywiste, ważne

Stwierdzenie 2.6.4. *Dla każdego procesu stochastycznego $(X_t)_{t \in T}$, w przestrzeni probabilistycznej $(\mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T, \mathbb{P}_X)$ rodzina rzutowań $\pi_t: \mathbb{R}^T \rightarrow \mathbb{R}$, $t \in T$, jest procesem stochastycznym o rozkładzie $\mathbb{P}_X$.* □

Pochodząca od Kołmogorowa i stanowiąca treść twierdzenia 2.6.6 konstrukcja pozwala na charakteryzację rozkładów prawdopodobieństwa procesów stochastycznych, czyli – właściwie – dowolnych miar prawdopodobieństwa w produktach $\mathbb{R}^T$, poprzez układy wzajemnie powiązanych miar w skończone wymiarowych przestrzeniach $\mathbb{R}^S$, dla $S = \{t_1, \dots, t_n\} \subset T$, $n \in \mathbb{N}$. W tym sensie identyczność na $\mathbb{R}^T$ jest *uniwersalnym* procesem stochastycznym $(\pi_t)_{t \in T}$ dla każdej miary prawdopodobieństwa w produkcie.

Ćwiczenie 2.6.2. Dla dowolnej miary μ w przestrzeni $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})^T$, każdy skończony ciąg $(t_1, \dots, t_n)$ w T wyznacza rzutowanie $(\pi_{t_1}, \dots, \pi_{t_n}): \mathbb{R}^T \rightarrow \mathbb{R}^n$, a więc indukuje miarę borelowską $\mu_{t_1 \dots t_n} = (\pi_{t_1}, \dots, \pi_{t_n})_* \mu$ w $\mathbb{R}^n$. Jeśli miara μ jest rozkładem prawdopodobieństwa procesu X na $(\Omega, \mathfrak{F}, \mathbb{P})$, to $\mu_{t_1 \dots t_n}$ jest rozkładem zmiennej losowej $(X_{t_1}, \dots, X_{t_n}): \Omega \rightarrow \mathbb{R}^n$. Ograniczając się do ciągów różnowartościowych sprawdzić, że otrzymana w ten sposób rodzina miar spełnia następujące warunki zgodności – dla każdego $n > 1$:

- (i) $\pi'_* \mu_{t_1 \dots t_n} = \mu_{t_1 \dots t_{n-1}}$, dla $\pi': \mathbb{R}^n \ni (x_1, \dots, x_n) \mapsto (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$.
- (ii) $\tilde{\alpha}_* \mu_{t_1 \dots t_n} = \mu_{t_{\alpha(1)} \dots t_{\alpha(n)}}$, dla dowolnej permutacji α zbioru $\{1, \dots, n\}$ i odpowiadającej jej permutacji $\tilde{\alpha}: (x_1, \dots, x_n) \mapsto (x_{\alpha(1)}, \dots, x_{\alpha(n)})$ współrzędnych w $\mathbb{R}^n$.

Definicja 2.6.5. *Rodzina borelowskich miar prawdopodobieństwa μ_S na $\mathbb{R}^n$ określona dla każdego n -elementowego podzbioru $S = \{t_1, \dots, t_n\} \subset T$,*

$n \in \mathbb{N}$, ustalonego zbioru T jest zgodna, jeśli dla każdej pary podzbiorów S, S' , takich że $S' \subset S \subset T$, zachodzi równość $(\pi_{S'}^S)_* \mu_S = \mu_{S'}$, czyli

$$\mu_{S'}(A) = \mu_S((\pi_{S'}^S)^{-1}A) \quad \text{dla} \quad A \in \mathfrak{B}_{\mathbb{R}^{S'}}, \quad (2.124)$$

gdzie $\pi_{S'}^S: \mathbb{R}^S \ni x \mapsto x|_{S'} \in \mathbb{R}^{S'}$ jest odwzorowaniem obcięcia.

Ćwiczenie 2.6.3. Wykazać, że warunek (2.124) jest równoważny własnościom (i)–(ii) sformułowanym w ćwiczeniu 2.6.2. Dokładniej – każdej mierze postaci μ_S odpowiada rodzina miar $\mu_{t_1 \dots t_n}$ dla $S = \{t_1, \dots, t_n\}$, gdzie pozorną zależność od uporządkowania zbioru S rekompensuje własność (ii), Sprawdzenie warunku (2.124) wystarczy ograniczyć do par, w których $S \setminus S'$ jest zbiorem jednoelementowym.

Twierdzenie 2.6.6 (Daniell–Kołmogorow). *Dla dowolnego zbioru parametrów T każda zgodna rodzina miar prawdopodobieństwa μ_S , dla skończonych podzbiorów $S = \{t_1, \dots, t_n\} \subset T$, $n \in \mathbb{N}$, pochodzi od dokładnie jednej miary μ w przestrzeni produktowej $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})^T$.*

Podane tu twierdzenia rozszerza wcześniejszą, przeliczalną konstrukcję pochodzącą od Daniella. Przypominamy, że szczególny, opisany w twierdzeniu 2.1.2, przypadek dowolnego produktu miar probabilistycznych nie ogranicza się do $\mathbb{R}$ i nie wymaga żadnych założeń na przestrzenie mierzalne.

Dowód. Przyjmijmy oznaczenie $\pi_S: \mathbb{R}^T \ni x \mapsto x|_S \in \mathbb{R}^S$, dla dowolnego podzbioru $S \subset T$. Konstrukcja miary μ polega na rozszerzeniu danej zgodnej rodziny miar w pierwszej kolejności na ciało zbiorów cylindrycznych

$$\mathfrak{B}^o = \{\pi_S^{-1}C; C \in \mathfrak{B}_{\mathbb{R}^S}, S \subset T, \#S < \infty\}, \quad (2.125)$$

a następnie na skorzystaniu z twierdzenia 1.1.38 i faktu, iż ciało $\mathfrak{B}^o$ generuje σ –ciało $\mathfrak{B}_{\mathbb{R}}^T$ zbiorów mierzalnych w $\mathbb{R}^T$. Pozostawiając czytelnikowi sprawdzenie, że warunek zgodności (2.124) gwarantuje poprawność i skończoną addytywność rozszerzenia

$$\mu^o(A) := \mu_S(C), \quad \text{jeśli} \quad A = \pi_S^{-1}C,$$

ograniczymy się do wykazania ciągłości (por. uwaga 1.1.10) utworzonej w ten sposób miary na ciele. Niech zatem zbiory $A_n \in \mathfrak{B}^o$, $n \in \mathbb{N}$, tworzą dowolny ciąg zstępujący. Zakładając nierówności $\mu^o(A_n) \geq \varepsilon$, dla pewnego $\varepsilon > 0$ i $n \in \mathbb{N}$, wystarczy wykazać, że przekrój $\bigcap_n A_n$ jest niepusty. Każdy ze zbiorów rozważanego ciągu jest postaci podanej w (2.125), przy czym niejednoznaczność opisu pozwala założyć, że dla każdego n zbiór S_{n+1} indeksów charakteryzujących A_{n+1} jest nadzbiorem analogicznego zbioru S_n

dla A_n . Możemy zatem wskazać ciąg zbiorów borelowskich $C_n \in \mathfrak{B}_{\mathbb{R}}^{S_n}$ dla $n \in \mathbb{N}$ taki, że

$$A_n = \pi_{S_n}^{-1} C_n, \quad \text{ i } \quad \mu_{S_n}(C_n) \geq \varepsilon$$

dla $n \in \mathbb{N}$. Zgodnie z twierdzeniem 1.1.47(ii), dla każdego $n \in \mathbb{N}$ istnieje zwarty podzbiór $K_n \subset C_n$ taki, że $\mu_{S_n}(C_n \setminus K_n) \leq \frac{\varepsilon}{2^n}$. Oznacza to nierówności

$$\mu^o(A_n) - \mu^o(H_n) \leq \frac{\varepsilon}{2^n}, \quad \text{ gdzie } H_n := \pi_{S_n}^{-1} K_n \subset A_n,$$

dla $n \in \mathbb{N}$. Aby ze zbiorów H_n utworzyć zstępujący ciąg zbiorów domkniętych, przyjmujemy $B_n = H_1 \cap \dots \cap H_n$, dla $n \in \mathbb{N}$. Nierówności

$$\mu^o(A_n) - \mu^o(B_n) = \mu^o\left(\bigcup_{i \leq n} (A_i \setminus H_i)\right) \leq \sum_{i \leq n} \mu^o(A_i \setminus H_i) < \varepsilon,$$

a więc $\mu^o(B_n) > \mu^o(A_n) - \varepsilon \geq 0$, pozwalają stwierdzić, że każdy ze zbiorów $B_n, n \in \mathbb{N}$, jest niepusty. Z definicji zbiorów B_n wynika istnienie jednoznacznie określonych domkniętych podzbiorów $K'_n \subset K_n$ takich, że $B_n = \pi_{S_n}^{-1} K'_n$, przy czym zależności $B_{n+1} = H_{n+1} \cap B_n$ implikują tożsamości

$$\begin{aligned} K'_{n+1} &= K_{n+1} \cap (\pi_{S_n}^{S_{n+1}})^{-1}(K'_n), \quad \text{ czyli} \\ x \in K'_{n+1} &\iff (x \in K_{n+1} \text{ i } x|_{S_n} \in K'_n) \quad \text{ dla } x \in \mathbb{R}^{S_{n+1}}, n \in \mathbb{N}. \end{aligned}$$

Zatem

$$x \in K'_{n+1} \implies (x|_{S_{n+1} \setminus S_n} \in J_n \text{ i } x|_{S_n} \in K'_n)$$

gdzie $J_n = \pi_{S_{n+1} \setminus S_n}^{S_{n+1}}(K_{n+1})$ jest zbiorem zwartym, dla $n \in \mathbb{N}$. W przestrzeni $\mathbb{R}^S$, gdzie $S = \bigcup_{n \in \mathbb{N}} S_n$, zbiory zwarte $D_n \equiv K'_n \times J_n \times J_{n+1} \times \dots$, takie że

$$D_n = \left\{ x \in \mathbb{R}^S; \ x|_{S_n} \in K'_n, \forall m \geq n \ x|_{S_{m+1} \setminus S_m} \in J_m \right\}, \quad n \in \mathbb{N},$$

tworzą ciąg zstępujący o niepustym przekroju. Inkluzje $\pi_S^{-1} D_n \subset B_n \subset A_n$, dla $n \in \mathbb{N}$, pociągają za sobą własność $\bigcap_{n \in \mathbb{N}} A_n \neq \emptyset$. $\square$

2.6.2 Trajektorie i ośrodkowość

Zgodnie ze stwierdzeniem 2.6.4 przestrzeń mierzalna $(\mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T) = (\mathbb{R}, \mathfrak{B}_{\mathbb{R}})^T$ jest uważana za uniwersalną, modelową dla procesów stochastycznych o przestrzeni czasu T . Jest to jednak tylko *jeden z modeli* – nie najważniejszy i nie zawsze najprostszy. Brak jednoznaczności (izomorficzności), wymaganej dla produktu w sensie kategorijskim, pozwala szukać modeli o dodatkowych własnościach – przy zachowaniu tego samego rozkładu prawdopodobieństwa. Względy praktyczne, czyli zastosowania, kierują naszą uwagę na własności *trajektorii procesów* zwanych też realizacjami lub ścieżkami.

Definicja 2.6.7. Trajektorią procesu stochastycznego $X: (\Omega, \mathfrak{F}) \rightarrow (\mathbb{R}, \mathfrak{B}_{\mathbb{R}})^T$ nazywamy każdą z wartości $X(\omega)$, dla $\omega \in \Omega$, a więc specyficzną dla procesu funkcję $X(\omega): T \rightarrow \mathbb{R}$.

Trajektoriami uniwersalnego procesu stochastycznego, jakim jest identyczność na przestrzeni produktowej $\mathbb{R}^T$, są *wszystkie* funkcje $T \rightarrow \mathbb{R}$. Okazuje się, że w ważnym, szczególnym przypadku, gdy przestrzeń czasu T jest przedziałem prostej rzeczywistej, modelowa przestrzeń $(\mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T)$ jest zbyt duża. Podstawową wadą jest *niemierzalność* podzbioru $\mathcal{C}(T) \subset \mathbb{R}^T$ złożonego z funkcji ciągłych, tzn. niemożność scharakteryzowania ciągłości funkcji (czyli ciągłości trajektorii procesu) poprzez ich wartości na przeliczalnym zbiorze punktów. Przedstawiamy poniżej rozwiązanie tego problemu pochodzące od J.L. Dooba. W dalszej części bieżącego podrozdziału przyjmujemy ogólniejsze założenie, że przestrzeń czasu T jest ośrodkową przestrzenią metryczną (lub metryzowalną) – por. [2, tw. 2.4.4; Tichonowa–Urysohna].

Definicja 2.6.8. Dla dowolnego przeliczalnego gęstego podzbioru $Q \subset T$, funkcję $x = (x_t)_{t \in T} \in \mathbb{R}^T$ nazwiemy Q –ośrodkową (ang. Q separable), jeśli dla każdego punktu $\tau \in T$ istnieje w Q ciąg $(q_n)_{n \in \mathbb{N}}$ zbieżny do τ i taki, że $x_{q_n} \rightarrow x_\tau$ lub $|x_{q_n}| \rightarrow \infty$. Zbiór Q będziemy też nazywać ośrodkiem dla funkcji x .

Zbiór $\mathcal{S}^Q(T) \subset \mathbb{R}^T$ wszystkich funkcji Q –ośrodkowych na T jest wprowadzie na ogół niemierzalny, ale – zgodnie z definicją 1.1.16 – wraz z σ –ciałem $\mathfrak{S}_T^Q = \mathfrak{B}_{\mathbb{R}}^T | \mathcal{S}^Q(T)$ ma naturalną strukturę podprzestrzeni.

Definicja 2.6.9. Przestrzenią funkcji ośrodkowych na T będziemy nazywać podprzestrzeń $(\mathcal{S}(T), \mathfrak{S}_T)$ w $(\mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T)$ złożoną z sumy $\mathcal{S}(T) = \bigcup_Q \mathcal{S}^Q(T)$ po przeliczalnych gęstych podzbiórach $Q \subset T$ oraz z σ –ciała $\mathfrak{S}_T = \mathfrak{B}_{\mathbb{R}}^T | \mathcal{S}(T)$.

Lemat 2.6.10 (Doob). (i) Dla każdego procesu stochastycznego $X = (X_t)_{t \in T}$ na dowolnej przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$, jeśli przestrzeń czasu T jest ośrodkową przestrzenią metryczną, to istnieje ośrodek $Q \subset T$ oraz równoważny proces o trajektoriach ośrodkowych

$$X' = (X'_t)_{t \in T}: (\Omega, \mathfrak{F}) \rightarrow (\mathcal{S}^Q(T), \mathfrak{S}_T^Q).$$

(ii) Każda miara prawdopodobieństwa μ w $(\mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T)$ jest redukowalna do podprzestrzeni $\mathcal{S}^Q(T) \subset \mathbb{R}^T$, dla pewnego ośrodka $Q \subset T$. Miara μ redukuje się do dokładnie jednej miary w przestrzeni $(\mathcal{S}(T), \mathfrak{S}_T)$.

Wniosek 2.6.11. Dla dowolnej ośrodkowej metrycznej przestrzeni parametrów T , każda zgodna rodzina miar prawdopodobieństwa $\mu_{t_1 \dots t_n}$, w $\mathbb{R}^n$, $t_1, \dots, t_n \in T$, $n \in \mathbb{N}$, pochodzi od dokładnie jednej miary ς w przestrzeni $(\mathcal{S}(T), \mathfrak{S}_T)$. $\square$

Uwaga 2.6.4. Dla dowolnego procesu stochastycznego X na $(\Omega, \mathfrak{F}, \mathbb{P})$, jedyną miarę ς w przestrzeni funkcji ośrodkowych $(\mathcal{S}(T), \mathfrak{S}_T)$, do której redukuje się miara $\mathbb{P}_X = X_*\mathbb{P}$ możemy uważać za zredukowany rozkład prawdopodobieństwa procesu. Miara ς nie może być określona jako ograniczenie miary $\mathbb{P}_X$ do podprzestrzeni, gdyż funkcje ośrodkowe nie stanowią mierzalnego podzbioru w $\mathbb{R}^T$.

Dowód lematu 2.6.10 wg. [3]. Wykażemy w pierwszej kolejności drugą część lematu. Ustalmy dowolną miarę prawdopodobieństwa μ w $(\mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T)$. Znajdziemy ośrodek Q w T i skonstruujemy mierzalne odwzorowanie rzutowania $\lambda: \mathbb{R}^T \rightarrow \mathcal{S}^Q(T)$, specyficzne dla μ i takie, że miara $\varsigma := \lambda_*\mu$ indukowana w przestrzeni $(\mathcal{S}^Q(T), \mathfrak{S}_T^Q)$ funkcji Q -ośrodkowych spełnia warunek $i_*\varsigma = \mu$, gdzie i jest włożeniem $\mathcal{S}^Q(T) \hookrightarrow \mathbb{R}^T$. W tym celu dla dowolnej pary $(U, J) \neq (\emptyset, \emptyset)$ złożonej z podzbioru otwartego $U \subset T$ i otwartego przedziału $J \subset \mathbb{R}$ i dla dowolnego zbioru przeliczalnego $C \subset U$ rozważamy wielkość

$$\rho(C) = \mu\{x; \forall_{s \in C} x_s \notin J\},$$

którą można interpretować jako lokalny współczynnik niezależności nośnika miary μ i zbioru C . Ustalając ciąg podzbiorów $(C_n)_{n \in \mathbb{N}}$ w U taki, że $\lim_n \rho(C_n) = \inf\{\rho(C); C \subset U\}$, przyjmujemy $C(U, J) = \bigcup_n C_n$, co oznacza, że $\rho(C(U, J))$ jest wartością minimalną. Dzięki temu, dla każdego $t \in U$ ma miejsce równość

$$\mu\{x; x_t \in J \text{ i } \forall_{s \in C(U, J)} x_s \notin J\} = 0. \quad (2.126)$$

Oznaczmy przez $\mathcal{Q}$ przeliczalną rodzinę wszystkich par (U, J) , w których U przebiega ustaloną przeliczalną bazę topologii w T , a J – wszystkie przedziały o końcach wymiernych. Przyjmijmy

$$\begin{aligned} Q &= \bigcup_{(U, J) \in \mathcal{Q}} C(U, J) \\ N(t) &= \bigcup_{(U, J) \in \mathcal{Q}_t} \left\{ x; x_t \in J \text{ i } \forall_{s \in C(U, J)} x_s \notin J \right\}, \end{aligned} \quad (2.127)$$

gdzie skorzystaliśmy z oznaczenia $\mathcal{Q}_t = \{(U, J) \in \mathcal{Q}; t \in U\}$ dla $t \in T$. Z własności (2.126) wynika, że każdy ze zbiorów $N(t)$, $t \in T$, ma miarę $\mu(N(t)) = 0$.

Ponieważ zbiory $C(U, J) \subset U$ są niepuste, suma $Q \subset T$ jest zbiorem gęstym. Dla każdego $t \in T \setminus Q$ ustalmy w Q ciąg $(q_n^t)_{n \in \mathbb{N}}$ zbieżny do t . Dla dowolnego $x \in \mathbb{R}^T$ przyjmujemy

$$\lambda(x)_t = x_t, \quad \text{jeśli } t \in Q \text{ lub } x \notin N(t), \quad (2.128)$$

natomiast w przypadku $x \in N(t)$ i $t \notin Q$,

$$\lambda(x)_t = \begin{cases} \limsup_n x_{q_n^t}, & \text{gdy ciąg } (x(q_n^t))_{n \in \mathbb{N}} \text{ jest ograniczony} \\ 0, & \text{jeśli ciąg } (x(q_n^t))_{n \in \mathbb{N}} \text{ jest nieograniczony.} \end{cases} \quad (2.129)$$

Dla ustalonego $t \in T \setminus Q$ założymy, że funkcja $x \in \mathbb{R}^T$ nie należy do $N(t)$. Rozważając ciąg par $(U_n, J_n) \in \mathcal{Q}$ zbiorów o średnicach mniejszych niż $\frac{1}{n}$ i takich, że $t \in U_n$ oraz $x_t \in J_n$, dla $n \in \mathbb{N}$, wnioskujemy z (2.127), że dla każdego $n \in \mathbb{N}$ istnieje punkt $q_n \in Q \cap U_n$ taki, że $x_{q_n} \in J_n$. Zatem $q_n \rightarrow t$ i $x_{q_n} \rightarrow x_t$, a więc także $\lambda(x)_{q_n} \rightarrow \lambda(x)_t$. Jeśli natomiast $x \in N(t)$ i ciąg $\lambda(x)(q_n^t) = x(q_n^t)$ jest ograniczony, to – zgodnie z (2.129) – istnieje podciąg (q'_n) w ciągu $(q_n^t)_{n \in \mathbb{N}}$ taki, że $\lambda(x)_{q'_n} \rightarrow \lambda(x)_t$. Ponieważ definicja 2.6.8 nie nakłada żadnych ograniczeń na wartości funkcji ośrodkowej w punktach ośrodka Q oraz w punktach, w których otoczeniu funkcja jest nieograniczona na Q , stwierdzamy tym samym ośrodkowość każdej z funkcji $\lambda(x)$, tzn. $\lambda(\mathbb{R}^T) \subset \mathcal{S}(T)$.

Zgodnie z definicją $\mathcal{S}(T)$ jako podprzestrzeni w $\mathbb{R}^T$, dla dowodu $\mathfrak{B}_{\mathbb{R}}^T / \mathfrak{S}_T$ -mierzalności odwzorowania λ wystarczy sprawdzić mierzalność poszczególnych współrzędnych $\lambda_t = \pi_t \circ \lambda$, dla $t \in T$, a więc mierzalność złożenia $i \circ \lambda: \mathbb{R}^T \rightarrow \mathbb{R}^T$. Zauważmy przy tym, że brak zupełności miary μ nie pozwala wywnioskować mierzalności funkcji λ_t wprost z równości (2.128). Korzystając z (2.129) stwierdzamy jednak, że dla $t \notin Q$, funkcje λ_t i π_t różnią się co najwyżej na zbiorze mierzalnym $N(t)$, przy czym $\lambda_t = 0$ na zbiorze $N_o(t) := \{x \in N(t); \forall m \in \mathbb{N} \exists n |x(q_n^t)| > m\}$, podczas gdy na dopełnieniu $N(t) \setminus N_o(t)$ funkcja λ_t jest mierzalną granicą górną. Przy okazji wykazaliśmy także równoważność procesów

$$\lambda_t \stackrel{\text{pw}}{=} \pi_t \quad \text{dla } t \in T, \quad (2.130)$$

co oznacza sprawdzenie własności sformułowanej w pierwszej części twierdzenia dla kanonicznego procesu stochastycznego na $(\mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T, \mu)$.

Przyjmijmy $\varsigma := \lambda_* \mu$. Równość $\mu = \iota_* \varsigma = (i \circ \lambda)_* \mu$ wynika ze stwierdzenia 2.6.3. Z kolei jednoznaczność miary ς jest konsekwencją definicji zbiorów mierzalnych w podprzestrzeni.

Dla dowodu (i) rozważmy dowolny proces stochastyczny $X = (X_t)_{t \in T}$ na przestrzeni $(\Omega, \mathfrak{F}, \mathbb{P})$. Dla rozkładu prawdopodobieństwa $\mathbb{P}_X = X_* \mathbb{P}$ istnieje

skonstruowane wyżej odwzorowanie mierzalne $\lambda: \mathbb{R}^T \rightarrow \mathcal{S}(T)$ spełniające (2.130). Równości $\mathbb{P}\{\lambda_t \neq \pi_t\} = 0$ oznaczają $\mathbb{P}\{\lambda_t \circ X \neq X_t\} = 0$, czyli $\lambda_t \circ X = X_t$ (pw), dla każdego $t \in T$. Wystarczy przyjąć $X' = \lambda \circ X$. $\square$

2.6.3 Ciągłość trajektorii jako problem

Podstawowym celem konstrukcji zawartej w lemacie Dooba 2.6.10 jest stworzenie probabilistycznego kontekstu, w którym problem posiadania przez proces stochastyczny ciągłych trajektorii ma czytelny opis i rozwiązanie. Klasycznym przykładem nietrywialnego procesu stochastycznego o ciągłych trajektoriach są badane w podrozdziale 2.7 *ruchy Browna*.

Ponieważ ciągłe trajektorie leżą w podprzestrzeni $\mathcal{C}(T) \subset \mathcal{S}(T)$, należy jeszcze zidentyfikować σ -ciało $\mathfrak{S}_T|\mathcal{C}(T) = \mathfrak{B}_{\mathbb{R}}^T|\mathcal{C}(T)$. Wykażemy

Lemat 2.6.12. *Założmy, że ośrodkowa przestrzeń metryczna T jest także lokalnie zwarta. Wówczas*

- (i) *dla dowolnego ośrodka Q w T przestrzeń funkcji ciągłych $\mathcal{C}(T)$ stanowi mierzalny podzbiór w przestrzeni $(\mathcal{S}^Q(T), \mathfrak{S}_T^Q)$, oraz*
- (ii) *zachodzi równość*

$$\mathfrak{B}_{\mathcal{C}(T)} = \mathfrak{B}_{\mathbb{R}}^T|\mathcal{C}(T),$$

jeśli przestrzeń $\mathcal{C}(T)$ rozważać z topologią zbieżności niemal jednostajnej.

Wyróżnionym w tym miejscu pojęciem topologii zbieżności zajmujemy się dokładniej w podrozdziale 3.2.1.

Dowód. Dla dowolnego ośrodka $Q \subset T$ i względnie zwartego podzbioru otwartego $U \subset T$, ciągłość funkcji $\bar{U} \rightarrow \mathbb{R}$ oznacza ciągłość jednostajną. Po ograniczeniu $x \in \mathcal{S}^Q(T)$ do zbioru $Q \cap U$, stwierdzamy, że ciągłość funkcji na $\bar{U}$ jest równoważna przynależności do zbioru

$$\left\{ x \in \mathcal{S}^Q(T); \forall_{m \in \mathbb{N}} \exists_{n \in \mathbb{N}} \forall_{\substack{p, q \in Q \cap U \\ d(p, q) < 1/n}} |x_p - x_q| < \frac{1}{m} \right\} \in \mathfrak{S}_T^Q.$$

Dla zakończenia dowodu własności (i) pozostaje ustalić dowolne przeliczalne otwarte pokrycie $T = \bigcup_k U_k$ zbiorami względnie zwartymi.

Bazę topologii zbieżności niemal jednostajnej w $\mathcal{C}(T)$ tworzą *kule otwarte* postaci $B_K(g, r) = \{f \in \mathcal{C}(T); \sup_K |f - g| < r\}$, dla $g \in \mathcal{C}(T)$, zbiorów zwartych $K \subset T$ i liczb $r > 0$. W szczególności – otwarte w $\mathcal{C}(T)$ są zbiory

$$B_{\{t\}}(g, r) = B_t(g(t), r) \cap \mathcal{C}(T), \quad \text{dla}$$

$$B_t(c, r) = \{f \in \mathbb{R}^T; |f(t) - c| < r\}, \quad \text{gdzie } t \in T, c \in \mathbb{R},$$

przy czym kule $B_t(c, r)$ generują σ -ciało $\mathfrak{B}_{\mathbb{R}}^T$. Z drugiej strony, jeśli ograniczyć się do zbiorów zwartych postaci $K = \bar{V}$, to równość

$$B_K(g, r) = \{f \in \mathcal{C}(T); \sup\{|f(p) - g(p)|; p \in Q \cap U\} < r\} \quad \text{dla } g \in \mathcal{C}(T)$$

oznacza, że każda taka kula otwarta może być przedstawiona w postaci

$$B_K(g, r) = \mathcal{C}(T) \cap \bigcup_{n \in \mathbb{N}} \bigcap_{p \in Q \cap U} B_p(f(p), r - \frac{1}{n}),$$

a więc należy do σ -ciała indukowanego z $\mathfrak{B}_{\mathbb{R}}^T$. $\square$

Twierdzenie 2.6.13 (O procesach ciągłych). *Niech $(\Omega, \mathfrak{F}, \mathbb{P})$ będzie dowolną przestrzenią probabilistyczną, a $X = (X_t)_{t \in T}$ – dowolnym procesem stochastycznym na Ω , dla którego przestrzeń czasu T jest lokalnie zwartą ośrodkową przestrzenią metryczną. Następujące warunki są równoważne:*

(i) *Rozkład prawdopodobieństwa $\mathbb{P}_X = X_*\mathbb{P}$ procesu X jest redukowalny do miary borelowskiej w $\mathcal{C}(T)$.*

(ii) *Istnieje równoważny z X proces o trajektoriach ciągłych*

$$\tilde{X}: (\Omega, \mathfrak{F}) \rightarrow (\mathcal{C}(T), \mathfrak{B}_{\mathcal{C}(T)}).$$

(iii) *Każdy równoważny z X proces $X': (\Omega, \mathfrak{F}) \rightarrow (\mathcal{S}^Q(T), \mathfrak{S}_T^Q)$ o trajektoriach ośrodkowych względem dowolnego ustalonego ośrodka $Q \subset T$ ma $\mathbb{P}$ -prawie wszystkie trajektorie ciągłe.*

(iv) *Istnieje równoważny z X proces $X': (\Omega, \mathfrak{F}) \rightarrow (\mathcal{S}^Q(T), \mathfrak{S}_T^Q)$, taki że*

$$(X'_*\mathbb{P})(\mathcal{C}(T)) = 1.$$

Dowód. Dla ustalonego ośrodka Q wprowadźmy oznaczenia dla wykorzystanych w dowodzie odwzorowań przestrzeni mierzalnych.

$$\begin{array}{ccccc} & & (\Omega, \mathfrak{F}) & & \\ & \swarrow X & \downarrow X' & \searrow \tilde{X} & \\ (\mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T) & \xleftarrow{i} & (\mathcal{S}^Q(T), \mathfrak{S}_T^Q) & \xleftarrow{j} & (\mathcal{C}(T), \mathfrak{B}_{\mathcal{C}(T)}) \end{array} \quad (2.131)$$

(i) $\Rightarrow$ (ii) Załóżmy, że w przestrzeni $\mathcal{C}(T)$ istnieje miara ν spełniająca warunek $\iota_*\nu = X_*\mathbb{P}$, dla $\iota = i \circ j$. Własność $\nu(\mathcal{C}(T)) = (\iota_*\nu)(\mathbb{R}^T) = 1$ oznacza, że ν jest miarą probabilistyczną. Zgodnie z lematem Dooba, dla X istnieje równoważny proces ośrodkowy $X': \Omega \rightarrow \mathcal{S}^Q(T)$, przy czym równość $i_*(j_*\nu) = i_*(X'_*\mathbb{P})$ implikuje zgodność miar $X'_*\mathbb{P} = j_*\nu$ na przestrzeni

$(\mathcal{S}^Q(T), \mathfrak{S}_T^Q)$. W szczególności, na zbiorze mierzalnym funkcji ciągłych mamy równość $(X'_*\mathbb{P})(\mathcal{C}(T)) = 1$. Szukany proces $\tilde{X}$ o ciągłych trajektoriach otrzymamy, przyjmując np.

$$\tilde{X}(\omega) = \begin{cases} X'(\omega) & \text{dla } \omega \in X'^{-1}\mathcal{C}(T) \\ 0 & \text{dla } \omega \in \Omega \setminus X'^{-1}\mathcal{C}(T). \end{cases}$$

(ii) $\Rightarrow$ (iii) Jeśli procesy stochastyczne $X': \Omega \rightarrow \mathcal{S}^Q(T)$ i $\tilde{X}: \Omega \rightarrow \mathcal{C}(T)$ są równoważne z X , to mają ten sam rozkład $(i \circ X')_*\mathbb{P} = (\iota \circ \tilde{X})_*\mathbb{P}$ redukowalny do $X'_*\mathbb{P} = j_*\tilde{X}_*\mathbb{P}$. Wspólną wartością obu miar na zbiorze funkcji ciągłych jest zatem $\mathbb{P}(X'^{-1}\mathcal{C}(T)) = (\tilde{X}_*\mathbb{P})(\mathcal{C}(T)) = 1$.

Implikacja (iii) $\Rightarrow$ (iv) jest oczywista.

(iv) $\Rightarrow$ (i) Mierzalność podzbioru $\mathcal{C}(T) \subset \mathcal{S}^Q(T)$ pozwala zdefiniować miarę borelowską $\nu = X'_*\mathbb{P}|_{\mathcal{C}(T)}$ taką, że $j_*\nu = X'_*\mathbb{P}$. Wówczas $\iota_*\nu = i_*X'_*\mathbb{P} = \mathbb{P}_X$, gdyż proces $i \circ X' = (X'_t)_{t \in T}$ jest równoważny z X . $\square$

Ważne, praktyczne kryterium redukowalności procesu stochastycznego do przestrzeni funkcji ciągłych pochodzi oczywiście od Kołmogorowa (1956).

Twierdzenie 2.6.14 (Kołmogorow–Czencow). *W dowolnej przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$ proces stochastyczny $(X_t)_{0 \leq t \leq T}$ dopuszcza ciągłą modyfikację $\tilde{X}$, jeśli istnieją stałe $\alpha, \beta, C > 0$, takie że*

$$\mathcal{E}|X_t - X_s|^\alpha \leq C|t - s|^{1+\beta}, \quad \text{dla } 0 \leq s, t \leq T. \quad (2.132)$$

Dla dowolnej liczby dodatniej $\gamma < \frac{\beta}{\alpha}$, prawie wszystkich $\omega \in \Omega$ i dostatecznie bliskich s i t zachodzi nierówność

$$\forall_{0 \leq s, t \leq T} |\tilde{X}_t(\omega) - \tilde{X}_s(\omega)| \leq K|t - s|^\gamma, \quad (2.133)$$

przy czym stała $K = K_\gamma$ nie zależy od ω .

Dowód. Łącząc założenie (2.132) z nierówności Czebyszewa (2.9) w postaci $\mathbb{P}(|X| \geq \varepsilon) \leq \frac{\mathcal{E}|X|^\alpha}{\varepsilon^\alpha}$ otrzymujemy oszacowanie

$$\mathbb{P}(|X_t - X_s| \geq |t - s|^\gamma) \leq C \frac{|t - s|^{1+\beta}}{(|t - s|^\gamma)^\alpha} \leq C|t - s|^{1+\beta-\alpha\gamma}$$

dla dowolnego $\gamma > 0$; ustalmy $\gamma \in (0, \frac{\beta}{\alpha})$.

Wykażemy, że na zbiorze liczb binarnych $D = \{\frac{i}{2^n}; 0 \leq i \leq \lfloor T2^n \rfloor, n \in \mathbb{N}\}$ prawie wszystkie trajektorie procesu są jednostajnie ciągłe. Istotnie, dla dowolnego ustalonego $n \in \mathbb{N}$ i ciągu liczb $t_i = \frac{i}{2^n}, i \leq \lfloor T2^n \rfloor$, prawdopodobieństwo zdarzenia

$$\Omega_n = \{\omega \in \Omega; \max_{i < \lfloor T2^n \rfloor} |X_{t_{i+1}}(\omega) - X_{t_i}(\omega)| \geq 2^{-n\gamma}\}$$

jest ograniczone z góry przez $C\lfloor T2^n \rfloor 2^{-n(1+\beta-\alpha\gamma)} \leq CT2^{-n(\beta-\alpha\gamma)}$. Stąd

$$\mathbb{P}\left(\bigcup_{n \geq m} \Omega_n\right) \leq CT \frac{2^{-m(\beta-\alpha\gamma)}}{1 - 2^{-m(\beta-\alpha\gamma)}} = \frac{CT}{2^{m(\beta-\alpha\gamma)} - 1} \longrightarrow 0 \quad \text{dla } m \rightarrow \infty,$$

a zdarzenie graniczne $\Omega_o = \bigcap_{m \in \mathbb{N}} \bigcup_{n \geq m} \Omega_n$ ma prawdopodobieństwo 0. Dla prawie wszystkich $\omega \in \Omega$ (dokł. dla $\omega \notin \Omega_o$) istnieje zatem $m = m(\omega)$, takie że

$$|X_{\frac{i+1}{2^n}}(\omega) - X_{\frac{i}{2^n}}(\omega)| < 2^{-n\gamma} \quad \text{dla } i < \lfloor T2^n \rfloor, n \geq m.$$

Postulowana jednostajna ciągłość trajektorii $X_t(\omega)$ dla $t \in D$ wymaga porównania wartości w chwilach $s, t \in D$, takich że $|t - s| < \frac{1}{2^n}$ dla odpowiednio dużego $n \in \mathbb{N}$. Rozważmy zatem dowolne $n \geq m(\omega)$ i liczbę $s \in D \cap [\frac{i}{2^n}, \frac{i+1}{2^n})$ dla pewnego $i \leq \lfloor T2^n \rfloor$. Rozwinięcie dwójkowe

$$s = \frac{i}{2^n} + \frac{a_1}{2^{n+1}} + \dots + \frac{a_k}{2^{n+k}} \leq T,$$

gdzie $a_j \in \{0, 1\}$ dla $j \leq k$, pozwala porównywać wartości trajektorii $X(\omega)$ w chwilach

$$s_j = \frac{i}{2^n} + \sum_{h=1}^j \frac{a_h}{2^{n+h}} = \frac{2^h i + \dots + 2a_{j-1} + a_j}{2^{n+h}}, \quad j = 0, \dots, k,$$

co prowadzi do oszacowania

$$\begin{aligned} |X_s(\omega) - X_{\frac{i}{2^n}}(\omega)| &\leq |X_{s_k}(\omega) - X_{s_{k-1}}(\omega)| + \dots + |X_{s_1}(\omega) - X_{s_0}(\omega)| \\ &< 2^{-(n+k)\gamma} + \dots + 2^{-(n+1)\gamma} < \frac{1}{2^\gamma - 1} 2^{-n\gamma}. \end{aligned}$$

Wraz z analogicznym oszacowaniem dla liczby $t \in D$ odległej od s o mniej niż 2^{-n} otrzymujemy dla $s, t \in D$

$$|X_t(\omega) - X_s(\omega)| < (1 + 2 \frac{1}{2^\gamma - 1}) 2^{-n\gamma} = \frac{2^\gamma + 1}{2^\gamma - 1} 2^{-n\gamma}, \quad \text{jeśli } |t - s| < \frac{1}{2^n}, \quad (2.134)$$

gdzie skorzystaliśmy także z faktu, że liczba t leży w tym samym lub w jednym z sąsiednich przedziałów o długości 2^{-n} . Otrzymana nierówność zachodzi dla wszystkich $n \geq m(\omega)$, a to oznacza jednostajną ciągłość $X(\omega)|_D$. Dla każdego $\omega \notin \Omega_o$ niech $\tilde{X}(\omega) \in \mathcal{C}[0, T]$ będzie zatem jednoznaczny ciągłym rozszerzeniem funkcji $X(\omega)|_D$; dla $\omega \in \Omega_o$ przyjmujemy $\tilde{X}(\omega) \equiv 0$.

Wartość wykorzystywanej w dowodzie liczby $\gamma < \frac{\beta}{\alpha}$ nie była do tej pory szczególnie istotna. Jeśli jednak dla $\omega \notin \Omega_o$ zastosować własność (2.134) dla różnych $s, t \in D$, takich że $2^{-(n+1)} \leq |t - s| < 2^{-n}$, otrzymujemy nierówność

$$|X_t(\omega) - X_s(\omega)| \leq \frac{2^\gamma + 1}{2^\gamma - 1} (2|t - s|)^\gamma = K_\gamma |t - s|^\gamma,$$

prawdziwą dla wszystkich $s, t \in D$, jeśli tylko $|t - s| < 2^{-m(\omega)}$. Rozszerzenie do nierówności (2.133) wynika z ciągłości trajektorii. $\square$

2.6.4 Trajektorie jednostronnie ciągłe

Czytelnik zauważył zapewne, że równoważności analogiczne do opisanych w twierdzeniu 2.6.13 zachodzą dla dowolnej przestrzeni funkcji $\mathcal{C}$ stanowiącej mierzalny podzbiór w każdej przestrzeni $\mathcal{S}^Q(T)$. W szczególności, dla dowolnego procesu stochastycznego $(X_t)_{t \in T}$ istnienie równoważnego procesu o wartościach w $\mathcal{C}$ jest wówczas równoważne redukowalności (do $\mathcal{C}$) rozkładu prawdopodobieństwa procesu X .

Przykład 2.6.5. Przestrzeń $\mathcal{D}(\mathbb{R}_+)$. Mówimy, że funkcja rzeczywista f na półprostej $\mathbb{R}_+$ jest klasy $\mathcal{D}$, jeśli posiada w każdym punkcie t skończoną granicę prawostronną $f(t+)$ oraz lewostronną $f(t-)$ (dla $t > 0$), a wartość $f(t)$ jest jedną z powyższych granic. Półciągłość funkcji f w każdym punkcie implikuje Q -ośrodkowość, dla każdego przeliczalnego podzbioru gęstego $Q \subset \mathbb{R}_+$. Niech zatem $\mathcal{D}(\mathbb{R}_+) \subset \mathcal{S}(\mathbb{R}_+)$ oznacza zbiór wszystkich funkcji klasy $\mathcal{D}$.

Ćwiczenie 2.6.6. Sprawdzić, że dla $\varepsilon > 0$ dowolna funkcja $f: \mathbb{R}_+ \rightarrow \mathbb{R}$ klasy $\mathcal{D}$ ma w każdym przedziale $[0, T]$, $T > 0$, skończoną liczbę punktów nieciągłości t ze skokiem $|f(t+) - f(t-)| > \varepsilon$.

Lemat 2.6.12 zapewnia mierzalność zbioru funkcji ciągłych – jednostajnie ciągłych na zbiorach zwartych. Potrzebna nam w tym miejscu analogiczna przeliczalna charakteryzacja funkcji klasy $\mathcal{D}$ opiera się na ograniczeniu dozwolonej liczby oscylacji.

Lemat 2.6.15. (i) Dowolna funkcja $f \in \mathcal{D}(\mathbb{R}_+)$ ma dla $\varepsilon > 0$ na każdym przedziale $[0, T]$, $T > 0$, skończoną, ograniczoną z góry liczbę oscylacji przekraczających ε , tzn. prawdziwa jest implikacja

$$\exists_{n \in \mathbb{N}} \forall_{t_1, \dots, t_n \geq 0} (t_1 < \dots < t_n \leq T \implies \min_{i < n} |f(t_i) - f(t_{i+1})| \leq \varepsilon) \quad (2.135)$$

dla dowolnych $\varepsilon, T > 0$.

(ii) Jeśli Q jest dowolnym ośrodkiem w $\mathbb{R}_+$, to każda funkcja $f \in \mathcal{S}^Q(\mathbb{R}_+)$ półciągała we wszystkich punktach ośrodka i mająca własność (2.135) ograniczoną do punktów $t_1, \dots, t_n \in Q$ – dla dowolnych $\varepsilon, T > 0$ – jest klasy $\mathcal{D}$.

Dowód. Przypuśćmy, że funkcja $f \in \mathcal{D}(\mathbb{R}_+)$ nie spełnia (2.135) na przedziale $[0, T]$ dla pewnego $\varepsilon > 0$. Wychodząc on $n = 2$ i dowolnego 2-elementowego zbioru $W_1 = \{\tau_1, \tau_2\}$, takiego, że $|f(\tau_1) - f(\tau_2)| > \frac{\varepsilon}{2}$, skonstruujemy wstępujący ciąg zbiorów skończonych $(W_m)_{m \in \mathbb{N}}$ zaprzeczających własności (2.135) dla oscylacji przekraczających $\frac{\varepsilon}{2}$ i coraz większych $n \in \mathbb{N}$. Dokładniej, założmy, że dla pewnego $m \in \mathbb{N}$ jest już znany zbiór W_m mający n elementów

$0 \leq \tau_1 < \dots < \tau_n \leq T$ takich, że $|f(\tau_i) - f(\tau_{i+1})| > \frac{\varepsilon}{2}$ dla $i < n$. Z przyjętego dla f założenia wynika istnienie rosnącego ciągu $(3n+1)$ -elementowego $t_1, \dots, t_{3n+1} \in [0, T]$ takiego, że $|f(t_i) - f(t_{i+1})| > \varepsilon$ dla $i \leq 3n$. Odrzucając ewentualnie wyrazy, które już należą do W_m , możemy – na mocy zasady szufladkowej Dirichleta – wskazać w ciągu $(t_i)_{i \leq 3n+1}$ dwa początkowe wyrazy mniejsze niż τ_1 lub dwa końcowe wyrazy większe niż τ_n lub trzy kolejne wyrazy należące do jednego z przedziałów otwartych (τ_i, τ_{i+1}) . Czytelnik z łatwością sprawdzi, że w każdym przypadku pozwala to na rozszerzenie W_m o przynajmniej 1 ze wskazanych punktów w ten sposób, by otrzymany zbiór W_{m+1} spełniał podane założenie indukcyjne.

Mając zbudowany ciąg $(W_m)_{m \in \mathbb{N}}$, oznaczmy przez t_0 dowolny punkt skupienia zbioru $W = \bigcup_m W_m$. Liczba t_0 jest granicą pewnego monotonicznego ciągu $(\tau_n)_{n \in \mathbb{N}}$ zawartego w W – dla ustalenia uwagi przyjmujemy, że jest to ciąg rosnący, $\tau_n \nearrow t_0$. Dla każdego n punkty τ_n i τ_{n+1} należą do pewnego W_m – niech $t_n \leq \tau_{n+1}$ oznacza następny po τ_n wyraz w W_m . Otrzymany w ten sposób ciąg $(t_n)_{n \in \mathbb{N}}$ również rośnie do t_0 , podczas gdy nierówności $|f(\tau_n) - f(t_n)| > \frac{\varepsilon}{2}$, dla $n \in \mathbb{N}$, wykluczają zbieżność obu ciągów do tej samej granicy $f(t_0-)$. Sprzeczność kończy dowód części (i).

Dla dowodu własności (ii) ustalmy przeliczalny gęsty podzbiór $Q \subset \mathbb{R}_+$ i załóżmy, że funkcja $f \in \mathcal{S}^Q(\mathbb{R}_+)$ spełnia ograniczony do punktów z Q warunek (2.135). Wykażemy, że f posiada w każdym punkcie granice prawo- i lewostronne (z oczywistym ograniczeniem dla końców przedziału). Przypuśćmy przeciwnie, że w pewnym punkcie $t_0 \in [0, T]$ nie istnieje np. granica lewostronna $f(t_0-)$. Dla pewnego $\varepsilon > 0$ mamy zatem

$$\forall \delta > 0 \exists t, t' \in (t_0 - \delta, t_0) \cap Q |f(t) - f(t')| > \varepsilon,$$

gdzie możliwość ograniczenia $t, t' \in Q$ wynika z ośrodkowości funkcji. Korzystając z indukcji, łatwo teraz zbudować w Q ciąg rosnący $(q_n)_{n \in \mathbb{N}}$, dążący do t_0 i taki, że $|f(q_n) - f(q_{n+1})| > \frac{\varepsilon}{2}$ dla wszystkich n . Otrzymana sprzeczność z (2.135) potwierdza istnienie wszystkich jednostronnych granic dla funkcji f . Półciągłość w punktach dopełnienia $\mathbb{R}_+ \setminus Q$ wynika z ośrodkowości f , natomiast analogiczną własność w punktach zbioru Q dołączamy do przyjętych w (ii) założeń. $\square$

Wniosek 2.6.16. Dla dowolnego ośrodka $Q \subset \mathbb{R}_+$ zbiór $\mathcal{D}(\mathbb{R}_+)$ jest mierzalnym podzbiorem w $\mathcal{S}^Q(\mathbb{R}_+)$. $\square$

Funkcje klasy $\mathcal{D}$ *prawostronnie ciągle* stanowią naturalne trajektorie dla tzw. procesów Levy'ego. Niech zatem podprzestrzeń $\mathcal{D}_r(\mathbb{R}_+) \subset \mathcal{D}(\mathbb{R}_+)$ oznacza przestrzeń liniową funkcji prawostronnie ciągłych. W obu przestrzeniach

rozważamy σ -ciała zbiorów mierzalnych $\mathfrak{D}_r$ i $\mathfrak{D}$ indukowane z $\mathcal{S}(\mathbb{R}_+)$, a więc generowane przez ograniczenia projekcji $\pi_t: \mathbb{R}^{\mathbb{R}_+} \rightarrow \mathbb{R}$, $t \in \mathbb{R}_+$. Można pokazać, że w przestrzeni $\mathcal{D}_r(\mathbb{R}_+)$ istnieje *metryka Skorochoda*, w której $\mathfrak{D}_r$ jest dokładnie σ -ciałem zbiorów borelowskich, natomiast $\mathcal{D}_r(\mathbb{R}_+)$ staje się przestrzenią ośrodkową. Wprawdzie zbiór $\mathcal{D}_r(\mathbb{R}_+)$ nie jest podzbiorem mierzalnym w żadnej z przestrzeni $\mathcal{S}^Q(\mathbb{R}_+)$, niemniej jednak prawdziwa jest następująca charakteryzacja procesów prawostronnie ciągłych.

Twierdzenie 2.6.17. *Niech $(\Omega, \mathfrak{F}, \mathbb{P})$ będzie dowolną przestrzenią probabilistyczną, a $X = (X_t)_{t \in \mathbb{R}_+}$ – dowolnym procesem stochastycznym na Ω , dla którego przestrzeń czasu jest półprosta $\mathbb{R}_+$. Następujące warunki są równoważne:*

- (i) *Rozkład prawdopodobieństwa $\mathbb{P}_X = X_*\mathbb{P}$ procesu X jest redukowalny do miary borelowskiej w $\mathcal{D}_r(\mathbb{R}_+)$.*
- (ii) *Istnieje równoważny z X proces $\tilde{X}: (\Omega, \mathfrak{F}) \rightarrow (\mathcal{D}_r(\mathbb{R}_+), \mathfrak{D}_r)$.*
- (iii) *Każdy równoważny z X proces $X': (\Omega, \mathfrak{F}) \rightarrow (\mathcal{S}^Q(\mathbb{R}_+), \mathfrak{S}_{\mathbb{R}_+}^Q)$ ma $\mathbb{P}$ -prawie wszystkie trajektorie klasy $\mathcal{D}$, przy czym dla każdego $t \in \mathbb{R}_+$ $\mathbb{P}$ -prawie wszystkie trajektorie procesu X' są prawostronnie ciągłe w t .*
- (iv) *Istnieje równoważny z X proces $X': (\Omega, \mathfrak{F}) \rightarrow (\mathcal{S}^Q(\mathbb{R}_+), \mathfrak{S}_{\mathbb{R}_+}^Q)$, którego zredukowany rozkład prawdopodobieństwa $\varsigma = X'_*\mathbb{P}$ ma własność*

$$\forall t \in \mathbb{R}_+ \quad \varsigma\{f \in \mathcal{D}(\mathbb{R}_+); f(t) = f(t+)\} = 1.$$

Dowód. Ćwiczenie – zmodyfikować schemat dowodu twierdzenia 2.6.13. $\square$

2.7 Ruchy Browna

2.7.1 ...czyli miara Wienera

Standardowy rozkład normalny Gaussa zdefiniowany i skonstruowany w poprzednim rozdziale 2.4 i pełniący centralną rolę w *Centralnym twierdzeniu granicznym* (tw. 2.5.13) jest ciągły, a jego gęstość jest funkcją gładką – dowolnie różniczkowalną. Skrajnie odmienny, losowy i nieróżniczkowalny charakter przejawia ściśle związany z rozkładem normalnym fundamentalny proces stochastyczny, skonstruowany przez Norberta Wienera (1923), a formalizujący obserwacje botanika Roberta Browna (1827).

Definicja 2.7.1. *Ruchem Browna lub procesem Wienera nazywamy każdy proces stochastyczny $\{B_t\}_{t \geq 0}$ w przestrzeni $(\Omega, \mathfrak{F}, \mathbb{P})$, którego trajektorie $\mathbb{R}_+ \ni t \mapsto B_t(\omega) \in \mathbb{R}$, $\omega \in \Omega$, są ciągle, przy czym*

- (i) $B_0 = 0$, a dla każdego $t > 0$ zmienna B_t ma rozkład Gaussa $\mathcal{N}(0, t)$;

(ii) dla każdego ciągu liczb $0 < t_1 < t_2 < \dots < t_n, n \in \mathbb{N}$, przyrosty

$$B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}$$

są niezależnymi zmiennymi losowymi.

Z poniższego lematu wynika, że każdy z przyrostów $B_t - B_s$ ruchu Browna – o ile takowy istnieje – ma także rozkład normalny.

Lemat 2.7.2. *Dla dowolnej zmiennej losowej X o rozkładzie normalnym i dowolnej zmiennej Y niezależnej od X (w tej samej przestrzeni) suma $X' = X + Y$ ma rozkład normalny o wariancji $\text{Var}(X') > \text{Var}(X)$ wtedy i tylko wtedy, gdy Y ma rozkład normalny.*

Dowód. Z niezależności X i Y wynika równość $\varphi_{X'} = \varphi_X \cdot \varphi_Y$ dla funkcji charakterystycznych rozważanych zmiennych. Niech X ma rozkład $\mathcal{N}(x, a)$. Jeśli X' ma rozkład $\mathcal{N}(\xi, b)$ i $b > a$, to z wniosku 2.5.14 i równości

$$\varphi_Y(t) = \frac{\varphi_{X'}(t)}{\varphi_X(t)} = e^{(\xi-x)t - \frac{t^2(b-a)}{2}} \quad \text{dla } t \in \mathbb{R}$$

wynika, że iloraz jest funkcją charakterystyczną rozkładu normalnego. Teza wynika z twierdzenia 2.5.4.

Implikacja przeciwna może być potwierdzona w analogiczny sposób, aczkolwiek została już wykazana w lemacie 2.4.7. $\square$

Wniosek 2.7.3. *Proces stochastyczny $\{B_t\}_{t \geq 0}$, którego prawie wszystkie trajektorie są ciągłe, jest ruchem Browna wtedy i tylko wtedy, gdy dla każdego ciągu różnych liczb $\tau = (t_1, \dots, t_n)$ w $\mathbb{R}_+$ zmienna wektorowa $(B_{t_1}, \dots, B_{t_n})$ ma rozkład Gaussa $\mathcal{N}_n(0, W_\tau)$ o macierzy kowariancji $W_\tau = [\min(t_i, t_j)]_{i,j \leq n}$.*

Dowód. Podana własność ruchu Browna jest konsekwencją przyjętej definicji, lematu 2.7.2 oraz wniosku 2.4.8. Z kolei z wniosku 2.4.12 wynika, że kształt macierzy kowariancji decyduje o niezależności kolejnych różnic $B_{t_{i+1}} - B_{t_i}$. Pozostaje sprawdzić, że dla $s < t$

$$\text{Cov}(B_s, B_t) = \text{Cov}(B_s, B_t - B_s) + \text{Cov}(B_s, B_s) = 0 + \text{Var}(B_s). \quad \square$$

Rezygnując czasowo z ciągłości trajektorii, wykażemy jednoznaczność:

Twierdzenie 2.7.4. *W przestrzeni $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})^{\mathbb{R}_+}$ istnieje dokładnie jedna miara $\mathbb{P}_B$, będąca rozkładem każdego procesu stochastycznego $(B_t)_{t \geq 0}$ spełniającego warunki (i)–(ii) sformułowane w definicji 2.7.1.*

Dowód. Twierdzenie Kołmogorowa 2.6.6 oraz wniosek 2.7.3 redukują tezę do zgodności rodziny miar postaci $\mathcal{N}_n(0, W_\tau)$ dla $\tau \in \mathbb{R}_+^n, n \in \mathbb{N}$.

Aby potwierdzić zgodność badanej rodziny rozkładów normalnych ustalmy dowolny podzbiór $S \subset \mathbb{R}_+$ postaci $S = \{t_1, \dots, t_n\}$, gdzie $t_1 < \dots < t_n$, $n \geq 2$, oraz liczbę $t_k \in S$. Dla $S' = S \setminus \{t_k\}$ warunek zgodności (2.124) sprowadza się do równości

$$\pi_{\hat{k}*} \mathcal{N}_n(0, W_\tau) = \mathcal{N}_{n-1}(0, W_{\tau_{\hat{k}}}), \quad (2.136)$$

gdzie $\tau = (t_1, \dots, t_n)$, $\tau_{\hat{k}} = \pi_{\hat{k}}(\tau)$, a rzutowanie

$$\pi_{\hat{k}}: \mathbb{R}^n \ni (x_1, \dots, x_n) \mapsto (x_1, \dots, \hat{x}_k, \dots, x_n) \in \mathbb{R}^{n-1} \quad \text{dla } k \leq n$$

pomija k -tą współrzędną. W przestrzeni $\mathbb{R}^n$ z miarą $\mathcal{N}(0, W_\tau)$ współrzędne $\pi_1, \dots, \pi_n$ tworzą wektor typu Gaussa o macierzy kowariancji W_τ . Pomińnięcie k -tej współrzędnej tworzy w $\mathbb{R}^n$ $(n-1)$ -wymiarowy wektor typu Gaussa o rozkładzie normalnym $\mathcal{N}_{n-1}(0, W')$, którego macierz kowariancji W' powstaje z W_τ przez usunięcie k -tego wiersza i k -tej kolumny (por. twierdzenie 2.4.11). $\square$

Do zakończenia dowodu poprawności definicji *ruchu Browna* skorzystamy z elementarnej własności rozkładu normalnego.

Lemat 2.7.5. *Dla dowolnej zmiennej losowej X o rozkładzie normalnym $\mathcal{N}(0, a)$ i dowolnego $n \in \mathbb{N}$*

$$\mathcal{E}(X^{2n}) = \frac{(2n)!}{2^n n!} a^n.$$

Dowód. Ćwiczenie. Wyrażenie $I_n = \int_{\mathbb{R}} t^{2n} \exp(-\frac{t^2}{2a}) dt$ wyznaczyć z rekurencyjnej zależności $I_n = (2n-1)aI_{n-1}$. $\square$

Jako wniosek otrzymujemy

Twierdzenie 2.7.6. *W przestrzeni $\mathcal{C}(\mathbb{R}_+)$ istnieje dokładnie jedna miara borelowska $\mathbb{P}_W$, dla której ewaluacje $\pi_t: \varphi \mapsto \varphi(t)$, $t \geq 0$, stanowią opisany w def. 2.7.1 standardowy ruch Browna.*

Definicja 2.7.7. *Miarę borelowską $\mathbb{P}_W$ w przestrzeni funkcji ciągłych $\mathcal{C}(\mathbb{R}_+)$ będącą rozkładem dowolnego ruchu Browna nazywamy miarą Wienera.*

Dowód. Z lematu 2.7.5 wynika nierówność

$$\mathcal{E}|\pi_t - \pi_s|^4 = 3|t - s|^2 \quad \text{dla } s, t \geq 0,$$

a zatem proces stochastyczny $\pi_t: \mathbb{R}^{\mathbb{R}_+} \rightarrow \mathbb{R}, t \geq 0$, spełnia warunek dostateczny ciągłości (2.132) względem miary $\mathbb{P}_B$. Z twierdzenia Kołmogorowa–Czencowa 2.6.14 wynika istnienie ciągłej modyfikacji procesu, a z twierdzenia 2.6.13 – jednoznaczna redukowalność miary $\mathbb{P}_B$ do miary borelowskiej w przestrzeni funkcji ciągłych na $\mathbb{R}_+$. $\square$

Zgodnie z miarą Wienera prawie wszystkie funkcje ciągłe $f \in \mathcal{C}(\mathbb{R}_+)$ są nigdzie nieróżniczkowalne – nigdzie, w żadnym punkcie(!). Prosty dowód tego faktu czytelnik znajdzie np. w [12, par. 13.2].

2.7.2 Konstrukcja procesu Wienera wg. Ciesielskiego

Istnienie ruchu Browna $(B_t)_{t \geq 0}$ w przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$ wymaga, by przestrzeń Ω dopuszczała ciąg niezależnych zmiennych losowych o tym samym rozkładzie normalnym $\mathcal{N}(0, 1)$, np. $X_n = B_n - B_{n-1}$, dla $n \in \mathbb{N}$. Oznacza to w szczególności, że przestrzeń Hilberta funkcji całkowalnych z kwadratem $L^2(\Omega, \mathbb{P})$ zawiera układ ortonormalny $(X_n)_{n \in \mathbb{N}}$ – będący bazą (zbiorem zupełnym) w podprzestrzeni $L^2 X = \overline{\text{span}\{X_n; n \in \mathbb{N}\}}$. Niezależnie od pochodzenia układu ortonormalnego $(X_n)_{n \in \mathbb{N}}$ w $L^2(\Omega, \mathbb{P})$ poszukiwanie w $L^2 X$ zmiennych B_s tworzących proces Wienera sprowadza się do wskazania ciągów współczynników Fouriera $c(s) = (c_n(s))_{n \in \mathbb{N}}$ takich, że

$$B_s = \sum_{n \in \mathbb{N}} c_n(s) X_n, \quad \text{gdzie } c_n(s) = \langle B_s, X_n \rangle \quad \text{dla } n \in \mathbb{N}, s \geq 0.$$

Ciągłość trajektorii procesu zależy od ciągłości funkcji c_n i np. od ewentualnej jednostajnej zbieżności wskazanego wyżej szeregu, a właściwa postać macierzy kowariancji wymaga spełnienia równości

$$\langle B_s, B_t \rangle = \sum_{n \in \mathbb{N}} c_n(s) c_n(t) = \min(s, t), \quad \text{dla } s, t \geq 0. \quad (2.137)$$

Zaproponowana przez Zbigniewa Ciesielskiego (1961) konstrukcja ogranicza przestrzeń czasu do odcinka jednostkowego, a w przestrzeni Hilberta funkcji całkowalnych z kwadratem $L^2[0, 1]$ – względem miary Lebesgue’a – korzysta z własności

$$\min(s, t) = \int 1_{[0, \min(s, t)]} d\ell = \langle 1_{[0, s]}, 1_{[0, t]} \rangle, \quad \text{dla } s, t \geq 0. \quad (2.138)$$

W tym kontekście właściwym wyborem ortonormalnego układu zupełnego w $L^2[0, 1]$ okazuje się dwuparametrowa, przeliczalna rodzina funkcji Haara

$H = (\chi_{kn})_{(k,n) \in N}$ takich, że $\chi_{10} \equiv 1$ i

$$\chi_{kn}(\tau) = \begin{cases} 2^{(n-1)/2}, & \text{gdy } \frac{k'}{2^{n-1}} < \tau \leq \frac{k}{2^n} \\ -2^{(n-1)/2}, & \text{gdy } \frac{k}{2^n} < \tau \leq \frac{k'+1}{2^{n-1}} \\ 0 & \text{w pozostałych przypadkach,} \end{cases} \quad (2.139)$$

określone dla nieparzystych $k = 2k' + 1 \leq 2^n$, $n \in \mathbb{N}$.

Ćwiczenie 2.7.1. Sprawdzić ortonormalność układu $(\chi_{kn})_{(k,n) \in N}$. Wykazać, że dla dowolnej funkcji $f \in L^2[0, 1]$ równości $\int f \chi_{k,n} d\ell = 0$ implikują znikanie ciągłej funkcji pierwotnej F (dla której $F' \stackrel{\text{pw}}{=} f$) we wszystkich punktach postaci $k/2^n \in [0, 1]$. Stąd – zupełność układu.

Dla każdej pary indeksów $(k, n) \in N$ niech $c_{kn}(s) = \int_{[0,s]} \chi_{kn} d\ell$ będzie współczynnikiem Fouriera rozwinięcia funkcji $1_{[0,s]}$ w bazie G , dla $s \in [0, 1]$. Załóżmy, że w przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$ istnieje ciąg niezależnych zmiennych losowych X_{kn} o tym samym rozkładzie normalnym $\mathcal{N}(0, 1)$; z oczywistego w tym momencie powodu wybrane zmienne indeksujemy za pomocą par $(k, n) \in N$.

Twierdzenie 2.7.8 (Ciesielski). *Proces stochastyczny $(B_t)_{t \in [0,1]}$ utworzony przez sumy w przestrzeni $L^2(\Omega, \mathbb{P})$*

$$B_t = \sum_{n \geq 0} \sum_{(k,n) \in N} c_{kn}(t) X_{kn}$$

jest ograniczonym do przedziału $[0, 1]$ ruchem Browna.

Oznaczmy $Y_n(t, \cdot) = \sum_{(k,n) \in N} c_{kn}(t) X_{kn}$, dla $n \geq 0, t \in [0, 1]$.

Lemat 2.7.9 (Losowe kryterium Weierstrassa). *Jeśli zachodzą nierówności*

$$\mathbb{P}\{\sup_t |Y_n(t, \cdot)| > a_n\} < b_n, \quad \text{dla } n \in \mathbb{N}$$

i oba szeregi $\sum a_n, \sum b_n$ są zbieżne, wówczas dla prawie wszystkich $\omega \in \Omega$ szereg $\sum_{n \geq 0} Y_n(\cdot, \omega)$ jest jednostajnie zbieżny.

Dowód. Dla dowolnego indeksu $k \in \mathbb{N}$ zachodzą nierówności

$$\mathbb{P}\left\{\bigcup_{i \geq k} \sup_t |Y_i(t, \cdot)| > a_i\right\} \leq \sum_{i \geq k} \mathbb{P}(\sup_t |Y_i(t, \cdot)| > a_i) \leq \sum_{i \geq k} b_i \xrightarrow{k \rightarrow \infty} 0.$$

Wynika stąd, że będący dopełnieniem przekroju zbiór

$$\Omega^* = \Omega \setminus \bigcap_k \bigcup_{i \geq k} \left\{ \sup_t |Y_i(t, \cdot)| > a_i \right\} = \bigcup_k \bigcap_{i \geq k} \left\{ \sup_t |Y_i(t, \cdot)| \leq a_i \right\}$$

ma miarę 1, a dla każdego $\omega \in \Omega^*$ istnieje $k \in \mathbb{N}$ takie, że $|Y_i(t, \omega)| \leq a_i$, dla wszystkich $i \geq k$ i $t \in [0, 1]$. Oszacowania

$$\sup_t \left| \sum_{i=n}^m Y_i(t, \omega) \right| \leq \sum_{i \geq n} a_i, \quad \text{dla } m \geq n \geq k$$

kończą dowód jednostajnej zbieżności szeregu $\sum_{n \geq 0} Y_n(\cdot, \omega)$. $\square$

Pierwsza część dowodu jest oczywiście zastosowaniem (jedynym w tej książce) klasycznego, elementarnego *Lematu Borela–Cantelliego*.

Dowód twierdzenia 2.7.8. Dla znalezienia zgodnego z lematem ciągu oszacowań zauważmy, że dla każdego $n \in \mathbb{N}$ zachodzą równości

$$\sup_t |Y_n(t, \omega)| = \max_k \left| \sup \chi_{kn} \left(\left(\frac{k-1}{2^n}, \frac{k+1}{2^n} \right) \right) X_{kn}(\omega) \right| = 2^{-(n+1)/2} \max_k |X_{kn}(\omega)|.$$

Stąd, niezależnie od nieznanych jeszcze wartości $a_n > 0$, szacujemy prawdopodobieństwa

$$\mathbb{P} \left\{ \sup_t |Y_n(t, \cdot)| > a_n \right\} \leq \mathbb{P} \left(\bigcup_k \{ |X_{kn}| > 2^{\frac{n+1}{2}} a_n \} \right) \leq 2^{n-1} \mathbb{P} \{ |X_{1n}| > 2^{\frac{n+1}{2}} a_n \},$$

przy czym ostatnią z miar wyznaczamy z rozkładu normalnego, co daje

$$\mathbb{P} \{ |X_{1n}| > \alpha_n \} = \frac{2}{\sqrt{2\pi}} \int_{\alpha_n}^{\infty} e^{-\frac{x^2}{2}} dx \leq \frac{2}{\sqrt{2\pi}} \int_{\alpha_n}^{\infty} \frac{x}{\alpha_n} e^{-\frac{x^2}{2}} dx = \frac{2}{\sqrt{2\pi} \alpha_n} e^{-\frac{\alpha_n^2}{2}},$$

dla $\alpha_n \equiv 2^{(n+1)/2} a_n$. Wynika stąd oszacowanie

$$\mathbb{P} \left\{ \sup_t |Y_n(t, \cdot)| > a_n \right\} \leq b_n \equiv \frac{2^{n/2}}{\sqrt{4\pi} a_n} e^{-2^n a_n^2}, \quad \text{dla } n \in \mathbb{N}.$$

Jednoczesną zbieżność obu szeregów zapewnia ciąg $a_n = \sqrt{n/2^n}$, $n \in \mathbb{N}$. Ewentualny brak definicji (lub ciągłości) na zbiorze $\Omega \setminus \Omega^*$ korygujemy przyjmując $B_t(\omega) = 0$ dla $\omega \notin \Omega^*$. $\square$

Podsumowanie i zakończenie konstrukcji stanowi

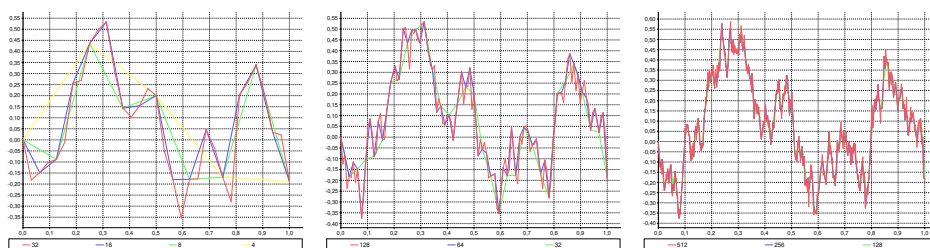
Stwierdzenie 2.7.10. *Dla dowolnej przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$, jeśli przestrzeń funkcji całkowalnych z kwadratem $L^2(\Omega, \mathbb{P})$ zawiera nieskończony ciąg niezależnych zmiennych o rozkładzie normalnym $\mathcal{N}(0, 1)$, wówczas na Ω istnieje proces Wienera $B_t: \Omega \rightarrow \mathbb{R}$, dla $t \geq 0$.*

Dowód. Grupując odpowiednio istniejące z założenia niezależne zmienne o rozkładzie normalnym, tworzymy przeliczalną rodzinę wzajemnie niezależnych nieskończonych ciągów zmiennych o takim samym rozkładzie $\mathcal{N}(0, 1)$.

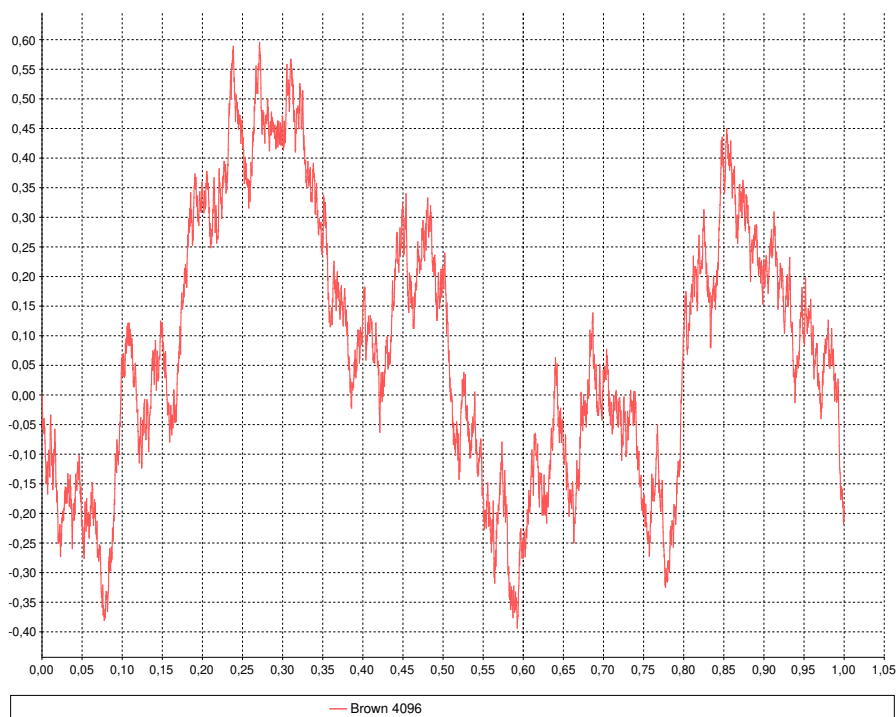
Z twierdzenie 2.7.8 wynika istnienie przeliczalnej rodziny ruchów Browna $B_t^{(n)}: \Omega \rightarrow \mathbb{R}$, $n \in \mathbb{N}$, wzajemnie niezależnych i określonych dla $t \in [0, 1]$. Ustawiając procesy „jeden za drugim”, przyjmujemy $B_n = B_1^{(1)} + \dots + B_1^{(n)}$, dla $n \in \mathbb{N}$, i ogólnie

$$B_t = \sum_{i=1}^{\lfloor t \rfloor} B_1^{(i)} + B_{t-\lfloor t \rfloor}^{[t]}, \quad \text{dla } t \geq 0.$$

Sprawdzenie warunków wymaganych od procesu Wienera pozostawiamy czytelnikowi. $\square$



Rysunek 2.8: Kolejne przybliżenia trajektorii – zależne od liczności próbek



Rysunek 2.9: Symulacja ruchu Browna

2.7.3 Twierdzenie graniczne Donskera

Punktem wyjścia dla zapowiedzianej *konstrukcji granicznej* jest określony na $(\Omega, \mathfrak{F}, \mathbb{P})$ dowolny (ustalony) ciąg $(X_n)_{n \in \mathbb{N}}$ niezależnych zmiennych losowych o tym samym rozkładzie, zerowej wartości oczekiwanej $\mathcal{E}(X_1) = 0$ i skończonej wariancji $\text{Var}(X_1) = \sigma^2 > 0$. Ciąg sum częściowych

$$S_0 = 0, \quad S_n = X_1 + \dots + X_n \quad \text{dla } n \in \mathbb{N},$$

rozszerzamy do ciągłego procesu stochastycznego, przyjmując dla $0 < \tau < 1$ $S_{n+\tau} := (1 - \tau)S_n + \tau S_{n+1}$, czyli

$$S_t = S_{[t]} + (t - [t])X_{[t]} \quad \text{dla } t \geq 0, \quad (2.140)$$

gdzie symbol $[\cdot]$ oznacza część całkowitą. Ostatni krok obejmuje skalowanie

$$X_t^{(n)} := \frac{1}{\sigma\sqrt{n}} S_{nt} \quad \text{dla } t \geq 0, n \in \mathbb{N}, \quad (2.141)$$

co daje ciąg $\mathcal{C}(\mathbb{R}_+)$ –wartościowych zmiennych losowych $X^{(n)}, n \in \mathbb{N}$, na Ω . Udowodnimy słabą zbieżność ciągu rozkładów do miary Wienera,

$$\mathbb{P}_n \equiv (X^{(n)})_* \mathbb{P} \xrightarrow{w} \mathbb{P}_W. \quad (2.142)$$

Twierdzenie 2.7.11 (Zasada niezmienniczości Donskera). *Niech $(X_n)_{n \in \mathbb{N}}$ będzie dowolnym ciągiem niezależnych zmiennych losowych w przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$, o tym samym rozkładzie, wartości oczekiwanej 0 i niezerowej skończonej wariancji σ^2 . Zdefiniowane w (2.140)–(2.141) procesy stochastyczne $X^{(n)}: \Omega \rightarrow \mathcal{C}(\mathbb{R}_+)$, $n \in \mathbb{N}$, tworzą ciąg zbieżny wg. rozkładu do miary Wienera $\mathbb{P}_W$, dla której kanoniczny proces stochastyczny złożony z ewaluacji*

$$\pi_t: \mathcal{C}(\mathbb{R}_+) \ni \phi \mapsto \phi(t) \in \mathbb{R}, \quad t \geq 0,$$

jest ruchem Browna.

Z twierdzenia 1.1.52, dla odwzorowania rzutowania $h = \pi_{t_1 \dots t_p}: \mathcal{C}(\mathbb{R}_+) \rightarrow \mathbb{R}^p$, wynika że słaba zbieżność ciągu skończenie wymiarowych rozkładów wyznaczonych przez $X^{(n)}$ jest warunkiem koniecznym postulowanej zbieżności miar $\mathbb{P}_n, n \in \mathbb{N}$. W pierwszej kolejności wykazujemy zatem

Lemat 2.7.12. *Dla każdego $p \in \mathbb{N}$ i dowolnego $\tau = (t_1, \dots, t_p) \in \mathbb{R}_+^p$, takiego że $0 \leq t_1 < \dots < t_p$, zachodzi zbieżność*

$$\begin{aligned} \text{(i)} \quad & \frac{1}{\sigma\sqrt{n}} (S_{[nt_1]}, \dots, S_{[nt_p]}) \xrightarrow{d} \mathcal{N}_p(0, W_\tau), \quad \text{oraz} \\ \text{(ii)} \quad & (X_{t_1}^{(n)}, \dots, X_{t_p}^{(n)}) \xrightarrow{d} \mathcal{N}_p(0, W_\tau), \end{aligned}$$

gdzie $W_\tau = [\min(t_i, t_j)]_{i,j \leq p}$ jest macierzą kowariancji opisującą p –wymiarowy rozkład ruchu Browna (por. wniosek 2.7.3).

Dowód. Zastępując zmienne X_n przez X_n/σ , możemy założyć, bez zmniejszenia ogólności, że $\sigma = 1$.

(i) Zgodnie z twierdzeniem 2.5.19 wystarczy dla dowolnego $(a_k)_{k \leq p} \in \mathbb{R}^p$ wykazać zbieżność ciągu zmiennych

$$Z_n \equiv \frac{1}{\sqrt{n}} \sum_{k \leq p} a_k S_{[nt_k]} = \frac{1}{\sqrt{n}} \sum_{k \leq p} b_k (S_{[nt_k]} - S_{[nt_{k-1}]}) ,$$

gdzie przyjęliśmy $t_0 = 0$ i $b_k = a_k + \dots + a_p$ dla $k \leq p$. Z niezależności składników otrzymujemy

$$\begin{aligned} \varphi_{Z_n}(1) &= \prod_{k \leq p} \prod_{\{nt_{k-1} < j \leq nt_k\}} \varphi_{X_j}\left(\frac{b_k}{\sqrt{n}}\right) \\ &= \prod_{k \leq p} \varphi_{X_1}\left(\frac{b_k}{\sqrt{n}}\right)^{[nt_k] - [nt_{k-1}]} \\ &= \prod_{k \leq p} \left(1 - \frac{b_k^2}{2n} + \dots\right)^{\frac{[nt_k] - [nt_{k-1}]}{n}} \longrightarrow \prod_{k \leq p} e^{-\frac{b_k^2(t_k - t_{k-1})}{2}} \end{aligned}$$

dla $n \rightarrow \infty$, gdzie łącznym wykładnikiem jest wyrażenie

$$-\frac{1}{2} \left(\sum_{k < p} (b_k^2 - b_{k+1}^2) t_k + b_p^2 t_p \right) = -\frac{1}{2} \left(\sum_k t_k a_k^2 + 2 \sum_{j < k} t_j a_j a_k \right),$$

będące formą kwadratową o macierzy współczynników równej $-\frac{1}{2}W_\tau$.

(ii) Dla każdego $t \geq 0$ z równości (2.140)–(2.141) definiujących zmienne $X_t^{(n)}$ wynikają oszacowania

$$|X_t^{(n)} - \frac{1}{\sqrt{n}} S_{[nt]}| \leq \frac{1}{\sqrt{n}} |X_{[nt]+1}|.$$

Przyjmując oznaczenie $Y_t^{(n)} \equiv \frac{1}{\sqrt{n}} S_{[nt]}$ i korzystając z nierówności Czebyszewa (2.9) otrzymujemy zatem

$$\mathbb{P}\{|X_t^{(n)} - Y_t^{(n)}| > \varepsilon\} \leq \mathbb{P}\{|X_{[nt]+1}| > \sqrt{n}\varepsilon\} \leq \frac{1}{n\varepsilon^2}$$

dla $t \geq 0$, $n \in \mathbb{N}$ i dowolnego $\varepsilon > 0$. Stąd dla $\tau = (t_1, \dots, t_p) \in \mathbb{R}_+^p$ i wektorów losowych $X_\tau^{(n)} = (X_{t_k}^{(n)})_{k \leq p}$, $Y_\tau^{(n)} = (Y_{t_k}^{(n)})_{k \leq p}$, oraz normy *supremum* $\|\cdot\|$ w $\mathbb{R}^p$, wynikają nierówności

$$\mathbb{P}\{\|X_\tau^{(n)} - Y_\tau^{(n)}\| > \varepsilon\} \leq \sum_{k \leq p} \mathbb{P}\{|X_{t_k}^{(n)} - Y_{t_k}^{(n)}| > \varepsilon\} \leq \frac{p}{n\varepsilon^2}$$

oznaczające zbieżność według prawdopodobieństwa $\|X_\tau^{(n)} - Y_\tau^{(n)}\| \xrightarrow{p} 0$ dla $n \rightarrow \infty$. Na mocy twierdzenia 2.5.27 własność (ii) jest zatem konsekwencją udowodnionej już własności (i). $\square$

W drugiej części dowodu twierdzenia Donskera wykazujemy zbieżność wg. rozkładu na każdym skończonym przedziale postaci $[0, T]$ dla dowolnego ustalonego $T \in \mathbb{N}$. Korzystamy z własności dotyczącej *modułów ciągłości* rozważanych zmiennych losowych – bliskiej pojęciu względnej zwartości w przestrzeni funkcji ciągłych. Niech π^T oznacza obcięcie $\mathcal{C}(\mathbb{R}_+) \rightarrow \mathcal{C}[0, T]$.

Definicja 2.7.13. Modułem ciągłości funkcji $x \in \mathcal{C}[0, T]$ nazywamy funkcję

$$w(x, \cdot): \delta \mapsto \max \{|x(t) - x(s)|; 0 \leq s, t \leq T, |s - t| \leq \delta\}. \quad (2.143)$$

Lemat 2.7.14. Dla każdego $T > 0$ i dowolnego $\varepsilon > 0$ zachodzi równość

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}\{\omega \in \Omega; w(\pi^T X^{(n)}(\omega), \delta) > \varepsilon\} = 0. \quad (2.144)$$

Dowód. Dla dowolnej funkcji $x \in \mathcal{C}[0, T]$ i dowolnego $\delta > 0$, jeśli liczby $0 = t_0 < t_1 < \dots < t_v = T$ dzielą przedział $[0, T]$ w ten sposób, że wszystkie przedziały wewnętrzne mają długość $t_{i+1} - t_i > \delta$ dla $0 < i < v - 1$, to każda para liczb s, t występująca w (2.143) zawiera się w pojedynczym przedziale lub sumie dwu sąsiednich. Stąd – łatwa do wyprowadzenia nierówność

$$w(x, \delta) \leq 3 \max_{0 \leq i < v} \sup_{t_i \leq s \leq t_{i+1}} |x(s) - x(t_i)|$$

prawdziwa dla każdej funkcji $x \in \mathcal{C}[0, T]$. Uwzględniając definicję funkcji postaci $x = X^{(n)}(\omega)$, dla każdego $n \in \mathbb{N}$ i dowolnego ustalonego $\delta > 0$ niech $m = \lceil n\delta \rceil$ będzie liczbą naturalną taką, że $m - 1 < n\delta \leq m$. Rozważmy *niemal równomierny* podział przedziału $[0, T]$ liczbami wymiernymi postaci $t_i = \frac{m_i}{n}$, gdzie $m_i = mi$ dla $i < v$, a $m_v = nT$, przy czym wybór liczb $v = \lceil \frac{nT}{m} \rceil$ zapewnia, że ostatni z przedziałów ma długość $t_v - t_{v-1} \leq \frac{m}{n}$. Pozwala to na oszacowanie modułu ciągłości każdej z kawałkami liniowych realizacji $X^{(n)}(\omega)$, $\omega \in \Omega$, poprzez potrójne maksimum liczb

$$\sup_{t_i \leq s \leq t_{i+1}} \left| X_s^{(n)}(\omega) - \frac{S_{m_i}(\omega)}{\sqrt{n}} \right| = \max_{m_i \leq k \leq m_{i+1}} \frac{|S_k - S_{m_i}|(\omega)}{\sqrt{n}}, \quad i < v.$$

Stąd, dla każdego $\varepsilon > 0$,

$$\begin{aligned} \mathbb{P}\{w(\pi^T X^{(n)}, \delta) > 3\varepsilon\} &\leq \sum_{i < v} \mathbb{P}\left\{ \max_{m_i \leq k \leq m_{i+1}} |S_k - S_{m_i}| > \sqrt{n}\varepsilon \right\} \\ &\leq v \mathbb{P}\left\{ \max_{k \leq m} |S_k| > \sqrt{n}\varepsilon \right\}, \end{aligned}$$

gdzie korzystamy z niezależności zmiennych i wspólnego ograniczenia dla różnic $m_{i+1} - m_i \leq m$. Aby skorzystać ze stwierdzenia 2.5.15, zauważmy,

że przy ustalonych $\varepsilon, \delta > 0$ zachodzi zbieżność $\frac{m}{n} \rightarrow \delta$ i $v \rightarrow \frac{T}{\delta}$ dla $n \rightarrow \infty$.
Zatem dla pewnego n_δ i wszystkich $n \geq n_\delta$ spełnione są nierówności

$$\sqrt{n}\varepsilon > \lambda\sqrt{m} \quad \text{ i } \quad v \leq \frac{2T}{\delta} = \frac{4T\lambda^2}{\varepsilon^2},$$

w których przyjęliśmy $\lambda = \varepsilon/\sqrt{2\delta}$. Uzyskane dla $n \geq n_\delta$ oszacowanie

$$\mathbb{P}\{w(\pi^T X^{(n)}, \delta) > 3\varepsilon\} \leq \frac{4T}{\varepsilon^2} \lambda^2 \mathbb{P}\left\{\max_{k \leq m} |S_k| > \lambda\sqrt{m}\right\}$$

pozwała natychmiast wywnioskować własność (2.144) z (2.115). $\square$

Dowód twierdzenia 2.7.11 wg. [4]. Aby połączyć lematy 2.7.12 i 2.7.14, dla ustalonego $T \in \mathbb{N}$, definiujemy ciąg odwzorowań $M_u: \mathcal{C}[0, T] \rightarrow \mathcal{C}[0, T]$, takich że

$$M_u x(t) = \begin{cases} x(\frac{i}{u}) & \text{dla } t = \frac{i}{u}, 0 \leq i \leq uT \\ (1-\tau)x(\frac{i}{u}) + \tau x(\frac{i+1}{u}) & \text{dla } t = \frac{i+\tau}{u}, \tau \in (0, 1). \end{cases}$$

Zauważmy, że z lematu 2.7.12(ii) wynika zbieżność wg. dystrybuant ujęta w górnym wierszu diagramu

$$\begin{array}{ccc} M_u(\pi^T X^{(n)}) & \xrightarrow{d} & M_{u*} \pi_*^T \mathbb{P}_W \\ \downarrow \wr & & \downarrow w \\ \pi^T X^{(n)} & \xrightarrow{d} & \pi_*^T \mathbb{P}_W. \end{array} \quad (2.145)$$

Istotnie, stosując lemat do ciągu $t_i = \frac{i}{u}$, $0 \leq i \leq uT$, stwierdzamy zbieżność wg. rozkładu

$$\pi_{t_0 \dots t_{uT}} \circ X^{(n)} \xrightarrow{d} (\pi_{t_0 \dots t_{uT}})_* \mathbb{P}_W. \quad (2.146)$$

Dla każdego $u \in \mathbb{N}$ zachodzi równość $M_u \circ \pi^T = L_u \circ \pi_{t_0 \dots t_{uT}}$, co możemy zapisać jako przemienność diagramu

$$\begin{array}{ccc} \mathcal{C}(\mathbb{R}_+) & \xrightarrow{\pi_{t_0 \dots t_{uT}}} & \mathbb{R}^{1+uT} \\ \downarrow \pi^T & & \downarrow L_u \\ \mathcal{C}[0, T] & \xrightarrow{M_u} & \mathcal{C}[0, T], \end{array}$$

w którym L_u jest odwzorowaniem ciągłym, przypisującym dowolnemu ciągowi liczb $\alpha = (\alpha_0, \dots, \alpha_{uT})$ funkcję $L_u(\alpha) \in \mathcal{C}[0, T]$ liniową na każdym przedziale $[\frac{i-1}{u}, \frac{i}{u}]$ i taką że $L_u(\alpha)(\frac{i}{u}) = \alpha_i$ dla $i \leq uT$. Twierdzenie o odwzorowaniu pozwala przekształcić (2.146) i – po zastosowaniu L_u – uzyskać zbieżność

$$M_u(\pi^T X^{(n)}) \xrightarrow{d} M_{u*} \pi_*^T \mathbb{P}_W \quad \text{ dla } n \rightarrow \infty.$$

Łatwe do sprawdzenia oszacowanie normy $\|M_u x - x\| \leq 2w(x, \frac{1}{u})$, korzystające z modułu ciągłości, implikuje jednostajną zbieżność $M_u x \rightarrow x$ dla $x \in \mathcal{C}[0, T]$, czyli $M_u \rightarrow \text{id}_{\mathcal{C}[0, T]}$ w przestrzeni metrycznej $\mathcal{C}[0, T]$, dla $u \rightarrow \infty$. Dla prawej części diagramu (2.145) wynika stąd słaba zbieżność ciągu miar indukowanych z $\mathbb{P}_W$, natomiast w odniesieniu do zmiennych $X_T^n := \pi^T X^{(n)}$ implikuje to także nierówności $\|M_u X_T^n - X_T^n\| \leq 2w(X_T^n, \frac{1}{u})$ oraz

$$\limsup_{n \rightarrow \infty} \mathbb{P}\{\|M_u X_T^n - X_T^n\| > \varepsilon\} \leq \limsup_{n \rightarrow \infty} \mathbb{P}\{w(X_T^n, \frac{1}{u}) > \frac{\varepsilon}{2}\} \rightarrow 0, \quad (2.147)$$

dla $u \rightarrow \infty$ i dowolnego $\varepsilon > 0$, na mocy lematu 2.7.14.

Dla dowodu zbieżności wg. rozkładu $X_T^n \xrightarrow{d} \pi_*^T \mathbb{P}_W$ rozważmy teraz dowolny podzbiór domknięty $F \subset \mathcal{C}[0, T]$ i dowolną liczbę $\varepsilon > 0$. Niech F_ε oznacza domknięte ε -otoczenie zbioru F . Z oczywistej implikacji

$$X_T^n(\omega) \in F \implies (M_u(X_T^n(\omega))) \in F_\varepsilon \text{ lub } \|M_u(X_T^n(\omega)) - X_T^n(\omega)\| > \varepsilon$$

wynikają nierówności

$$\mathbb{P}\{X_T^n \in F\} \leq \mathbb{P}\{M_u X_T^n \in F_\varepsilon\} + \mathbb{P}\{\|M_u X_T^n - X_T^n\| > \varepsilon\}$$

dla $n, u \in \mathbb{N}$. Interesuje nas granica górna ciągu miar $\mathbb{P}\{X_T^n \in F\}$, co w połączeniu ze zbieżnością $M_u(\pi^T X^{(n)}) \xrightarrow{d} M_{u*} \pi_*^T \mathbb{P}_W$ dla $n \rightarrow \infty$ daje nierówność

$$\limsup_{n \rightarrow \infty} \mathbb{P}\{X_T^n \in F\} \leq M_{u*} \pi_*^T \mathbb{P}_W(F_\varepsilon) + \limsup_{n \rightarrow \infty} \mathbb{P}\{\|M_u X_T^n - X_T^n\| > \varepsilon\},$$

a więc, po uwzględnieniu (2.147),

$$\limsup_{n \rightarrow \infty} \mathbb{P}\{\pi^T X^{(n)} \in F\} \leq \pi_*^T \mathbb{P}_W(F_\varepsilon).$$

Wobec dowolności ε i zbioru domkniętego F oznacza to zbieżność wg. rozkładu $\pi^T X^{(n)} \xrightarrow{d} \pi_*^T \mathbb{P}_W$ na każdym przedziale $[0, T]$, $T \in \mathbb{N}$.

Dla zakończenia dowodu przypominamy (tw. 1.1.51(i)), że dla każdego $T \in \mathbb{N}$ przestrzeń metryczna $\mathcal{C}[0, T]$ zawiera przeliczalną bazę $\mathfrak{V}_T$ złożoną z kul będących zbiorami ciągłości miary $\pi_*^T \mathbb{P}_W$. Wszystkie przeciwobrazy $(\pi^T)^{-1}V$, $V \in \mathfrak{V}_T$, są zbiorami ciągłości miary $\mathbb{P}_W$, a udowodniona wyżej zbieżność

$$\mathbb{P}\{X^{(n)} \in (\pi^T)^{-1}V\} \longrightarrow \mathbb{P}_W\{(\pi^T)^{-1}V\} \quad \text{dla } V \in \mathfrak{V}_T, T \in \mathbb{N},$$

zachodzi także dla skończonych przekrojów utworzonej w ten sposób bazy przestrzeni $\mathcal{C}(\mathbb{R}_+)$. Z Twierdzenia 1.1.51(ii) otrzymujemy ostatecznie zbieżność według rozkładu

$$X^{(n)} \xrightarrow{d} \mathbb{P}_W, \quad \text{dla } n \rightarrow \infty. \quad \square$$

Rozdział 3

Całki jako miary

Pierwszą dojrzałą koncepcję całki przedstawił Henri Lebesgue ok. roku 1902 – jako uogólnienie jeszcze wcześniejszej całki (Georga) Riemanna z 1854 r.

Celem bieżącego rozdziału jest próba zamiany ról i przywołanie tych konstrukcji, w których specyficznie pojmowana całka będąca *miarą funkcji* logicznie poprzedza pojęcie *miary zbioru*. Prezentowane w kolejnych podrozdziałach abstrakcyjne podejście do pojęcia całki (i miary) zaproponował ok. 1917 r. John Daniell. Dalsze podrozdziały dotyczą także tzw. miary i całki (Johanna) Radona.

3.1 Abstrakcyjna konstrukcja całki Daniella–Stone’a

3.1.1 Od funkcjonału do miary zewnętrznej

Definicja 3.1.1. Dla dowolnego niepustego zbioru T i będącej kratą przestrzeni liniowej $\mathcal{A} \subset \mathbb{R}^T$ funkcjonał liniowy $\chi: \mathcal{A} \rightarrow \mathbb{R}$ będziemy nazywać miarą abstrakcyjną, jeśli jest nieujemny i monotonicznie ciągły, tzn. spełnia warunki

$$\begin{aligned} \chi(f) &\geq 0, & \text{dla } f = f^+ \in \mathcal{A}, \\ \lim_{n \rightarrow \infty} \chi(f_n) &= \chi(f), & \text{jeśli } f_n \nearrow f \in \mathcal{A} \text{ i } f_n \in \mathcal{A}, \text{ dla } n \in \mathbb{N}. \end{aligned} \quad (3.1)$$

Stwierdzenie 3.1.2. Jeśli χ jest dowolną miarą abstrakcyjną określoną na kracie $\mathcal{A}$, to

$$|\chi(f)| \leq \chi(|f|) \quad \text{dla } f \in \mathcal{A}. \quad (3.2)$$

Dowód. Ćwiczenie. □

Zgodnie z uwagą 1.2.1 konstrukcja całki Lebesgue’a polega na stopniowym rozszerzaniu funkcjonału $\tilde{\mu}: \mathcal{P}(\mathcal{I}, \mu) \rightarrow \mathbb{R}$ określonego na przestrzeni funkcji prostych i będącego ważnym przykładem miary abstrakcyjnej (por. przykład 3.1.2). Inne abstrakcyjne miary są wykorzystywane w podrozdziale 3.2.

Zasadniczy schemat konstrukcji całki, rozwinięty przez Daniella i powiązany przez Stone’a z całką Lebesgue’a, składa się z dwu części (por. [5], [15]), z których pierwszą opisuje

Lemat 3.1.3. *Niech $\chi: \mathcal{A} \rightarrow \mathbb{R}$ będzie dowolną abstrakcyjną miarą w zbiorze T . Istnieje dokładnie jedna funkcja $\chi^*: \overline{\mathbb{R}}^T \rightarrow \overline{\mathbb{R}}$, będąca rozszerzeniem χ i taka, że*

(i) *na zbiorze $\mathcal{A}^* = \{f \in \overline{\mathbb{R}}^T; \exists (g_n)_{n \in \mathbb{N}} (\forall_n g_n \in \mathcal{A}) \text{ i } g_n \nearrow f\}$ funkcji osiągalnych z $\mathcal{A}$ (z dołu), χ^* jest addytywna, dodatnio jednorodna i ma własność monotonicznej ciągłości*

$$\lim_{n \rightarrow \infty} \chi^*(f_n) = \chi^*\left(\lim_{n \rightarrow \infty} f_n\right), \text{ jeśli } f_n \in \mathcal{A}^* \text{ i } f_n \leq f_{n+1} \text{ dla } n \in \mathbb{N}; \quad (3.3)$$

(ii) *dla $f \in \overline{\mathbb{R}}^T$ zachodzi równość*

$$\chi^*(f) = \begin{cases} \inf\{\chi^*(g); g \in \mathcal{A}^* \text{ i } f \leq g\} \\ +\infty, \text{ gdy } \{g \in \mathcal{A}^*; f \leq g\} = \emptyset. \end{cases} \quad (3.4)$$

Funkcja χ^ jest niemalejąca, subaddytywna (z ograniczeniem do funkcji, dla których $\chi^* > -\infty$) i dodatnio jednorodna. Dla dowolnego niemalejącego ciągu funkcji $(f_n)_{n \in \mathbb{N}}$, takiego że $\chi^*(f_1) > -\infty$, zachodzi równość*

$$\chi^*\left(\lim_{n \rightarrow \infty} f_n\right) = \lim_{n \rightarrow \infty} \chi^*(f_n). \quad (3.5)$$

Dowód. Postulowana własność (3.3) wymaga, by przyjąć

$$\chi^*(f) = \sup\{\chi(u); u \in \mathcal{A} \text{ i } u \leq f\} \quad \text{dla } f \in \mathcal{A}^*. \quad (3.6)$$

Traktując (3.6) jako definicję, rozważmy w $\mathcal{A}^*$ dowolny ciąg niemalejący $(f_n)_{n \in \mathbb{N}}$. Skoro każda z funkcji f_n jest granicą niemalejącego ciągu $(g_{mn})_{m \in \mathbb{N}}$ funkcji z $\mathcal{A}$, dla $n \in \mathbb{N}$, funkcje

$$h_n = \max_{p, q \leq n} g_{pq} \in \mathcal{A}, \quad \text{dla } n \in \mathbb{N}, \quad (3.7)$$

dążą monotonicznie do f . Wynika stąd, że funkcja graniczna $f = \lim_n f_n$ również jest osiągalna z dołu z $\mathcal{A}$, a więc $f \in \mathcal{A}^*$.

Ponieważ nierówności $\chi^*(f_n) \leq \chi^*(f)$ implikują $\lim_n \chi^*(f_n) \leq \chi^*(f)$, pozostaje wykazać nierówność przeciwną. W tym celu ustalmy dowolną funkcję

$u \leq f$, $u \in \mathcal{A}$. Korzystając z ciągu (3.7), stwierdzamy monotoniczną zbieżność funkcji $h'_n = \min(u, h_n) \in \mathcal{A}$ do u , przy zachowaniu nierówności $h'_n \leq f_n$ dla $n \in \mathbb{N}$. Własność (3.1) prowadzi zatem do nierówności

$$\chi(u) = \lim_n \chi(h'_n) \leq \lim_n \chi^*(f_n),$$

z której wynika (3.3). Jest to zarazem dowód jednoznaczności rozszerzenia χ na funkcje osiągalne z dołu. Dla dowolnych dwu funkcji $f, g \in \mathcal{A}^*$ i niemalejących ciągów funkcji w $\mathcal{A}$, $f_n \nearrow f$ i $g_n \nearrow g$, zbieżność $f_n + g_n \nearrow f + g$ i wynikające z niej równości

$$\chi^*(f + g) = \lim_n \chi^*(f_n + g_n) = \chi^*(f) + \chi^*(g)$$

dowodzą addytywności χ^* na $\mathcal{A}^*$.

Możemy teraz przyjąć wzór (3.4) za definicję rozszerzenia χ^* na dowolne funkcje z $\overline{\mathbb{R}}^T$. Łatwo sprawdzić, że otrzymaliśmy w ten sposób funkcję niemalejącą, zgodną na $\mathcal{A}^*$ z poprzednią definicją i, podobnie, dodatnio jednorodną. Dla dowodu subaddytywności zauważmy, że jeśli wartości χ^* na funkcjach f_1, f_2 są skończone, to dla dowolnych liczb $a_i > \chi^*(f_i)$ możemy znaleźć $g_i \in \mathcal{A}^*$ takie, że $f_i \leq g_i$ oraz $\chi^*(g_i) < a_i$, dla $i = 1, 2$. Wówczas zachodzą nierówności

$$\chi^*(f_1 + f_2) \leq \chi^*(g_1 + g_2) = \chi^*(g_1) + \chi^*(g_2) < a_1 + a_2.$$

Wobec dowolności a_1 i a_2 otrzymujemy stąd $\chi^*(f_1 + f_2) \leq \chi^*(f_1) + \chi^*(f_2)$.

Ponieważ dla dowolnego niemalejącego ciągu funkcji $(f_n)_{n \in \mathbb{N}}$ nierówność $\chi^*(\lim_n f_n) \geq \lim_n \chi^*(f_n)$ jest oczywista, przyjmijmy, że prawa strona nierówności jest skończona (a więc także $\chi^*(f_n) < \infty$ dla wszystkich n) i wykażemy nierówność przeciwną. Z założenia skończoności wynika, że wielkości $\chi^*(f_n)$ są osiągalne z góry przez $\chi^*(g_n)$, dla pewnych funkcji $g_n \in \mathcal{A}^*$, $n \in \mathbb{N}$. Dokładniej, dla dowolnej ustalonej liczby $\varepsilon > 0$ skonstruujemy w $\mathcal{A}^*$ *niemalejący* ciąg $(g_n)_{n \in \mathbb{N}}$ taki, że

$$f_n \leq g_n \quad \text{oraz} \quad \chi^*(g_n) \leq \chi^*(f_n) + (1 - 2^{-n})\varepsilon \quad \text{dla } n \in \mathbb{N}. \quad (3.8)$$

W tym celu dla każdego $n \in \mathbb{N}$ wybierzmy w $\mathcal{A}^*$ funkcję $h_n \geq f_n$ taką, że $\chi^*(f_n) \leq \chi^*(h_n) \leq \chi^*(f_n) + 2^{-n}\varepsilon$ i przyjmijmy $g_n = \max\{h_1, \dots, h_n\} \in \mathcal{A}^*$. Zauważmy w tym miejscu, że zbiór funkcji osiągalnych z dołu z $\mathcal{A}$ dziedziczy własność kraty – jest zamknięty ze względu na operacje maksimum i minimum skończonej liczby argumentów. Dla $n = 1$ nierówność (3.8) jest oczywista. Indukcyjny dowód korzysta z tożsamości

$$\max(g_n, h_{n+1}) + \min(g_n, h_{n+1}) = g_n + h_{n+1} \quad \text{dla } n \in \mathbb{N}, \quad (3.9)$$

prawdziwej dla dowolnej pary funkcji rzeczywistych o wspólnej dziedzinie. Obie strony równości (3.9) wiążą funkcje z $\mathcal{A}^*$, przy czym pierwsza z nich jest równa $\max(g_n, h_{n+1}) = g_{n+1}$, natomiast druga spełnia nierówność $\min(g_n, h_{n+1}) \geq \min(h_n, f_{n+1}) \geq f_n$. Addytywność i monotoniczność χ^* na $\mathcal{A}^*$ dają kolejno

$$\begin{aligned}\chi^*(g_{n+1}) + \chi^*(f_n) &\leq \chi^*(g_n) + \chi^*(h_{n+1}) \\ \chi^*(g_{n+1}) - \chi^*(f_{n+1}) &\leq \chi^*(g_n) - \chi^*(f_n) + \chi^*(h_{n+1}) - \chi^*(f_{n+1}).\end{aligned}$$

Stąd, jeśli nierówność (3.8) jest prawdziwa dla n , to

$$\chi^*(g_{n+1}) - \chi^*(f_{n+1}) \leq (1 - 2^{-n})\varepsilon + 2^{-(n+1)}\varepsilon = (1 - 2^{-(n+1)})\varepsilon.$$

Nierówności $f_n \leq g_n, n \in \mathbb{N}$, implikują nierówność $f \leq g$ dla funkcji granicznych obu ciągów, przy czym $g \in \mathcal{A}^*$ i prawdziwe jest oszacowanie

$$\chi^*(f) \leq \chi^*(g) = \lim_n \chi^*(g_n) \leq \lim_n \chi^*(f_n) + \varepsilon.$$

Wobec dowolności ε oznacza to własność (3.5) i koniec dowodu lematu. $\square$

Definicja 3.1.4. Funkcję $\chi^*: \overline{\mathbb{R}}^T \rightarrow \overline{\mathbb{R}}$ będziemy nazywać miarą zewnętrzną Daniella (lub całką zewnętrzną) wyznaczoną przez abstrakcyjną miarę χ .

Wniosek 3.1.5. W przestrzeni liniowej

$$\mathcal{M}(T, \chi) = \{f \in \mathbb{R}^T; \chi^*(|f|) < \infty\}. \quad (3.10)$$

miara zewnętrzna wyznacza półnormę $N_1(f) = \chi^*(|f|)$ dla $f \in \mathcal{M}(T, \chi)$. $\square$

Definicja 3.1.6. Dla dowolnej miary abstrakcyjnej $\chi: \mathcal{A} \rightarrow \mathbb{R}$ w zbiorze T , całką Daniella–Stone’a względem χ nazywamy funkcjonal liniowy

$$\int d\chi: \mathcal{I}(T, \chi) \longrightarrow \mathbb{R}$$

będący jedynym ciągłym rozszerzeniem miary χ na domknięcie $\mathcal{I}(T, \chi) = \overline{\mathcal{A}}$ podprzestrzeni $\mathcal{A} \subset \mathcal{M}(T, \chi)$ względem półnormy N_1 . Funkcje należące do dziedziny całki nazywamy χ -całkowalnymi.

W kolejnym podrozdziale wykażemy (twierdzenie 3.1.12), że całka $\int d\chi$ jest ograniczeniem $\chi^*|_{\overline{\mathcal{A}}}$ miary zewnętrznej, a jej dziedzina $\overline{\mathcal{A}}$ jest największą podprzestrzenią liniową w $\mathcal{M}(T, \chi)$, na której miara zewnętrzna jest funkcjonalem liniowym.

Przykład 3.1.1. Całka Riemanna (por. [1]) na ograniczonym przedziale domkniętym jest równa całce Lebesgue’a na prostej (względem miary Lebesgue’a ℓ). Przypominamy:

Podziałem przedziału $[a, b] \subset \mathbb{R}$ nazywamy dowolny skończony podzbiór $P \subset [a, b]$ zawierający końce a i b przedziału. Przyjmując rosnący porządek elementów zbioru $P = \{x_i; i \leq n\}$ w ten sposób, że $a = x_0 < x_1 < \dots < x_n = b$, $n \in \mathbb{N}$, dowolnej ograniczonej funkcji rzeczywistej $f: [a, b] \rightarrow \mathbb{R}$ możemy przypisać odpowiednio *dolną* i *górną sumę Riemanna*,

$$L(f, P) = \sum_{i=1}^n (x_i - x_{i-1}) \inf \{f(x); x_{i-1} < x < x_i\}, \text{ oraz} \\ U(f, P) = \sum_{i=1}^n (x_i - x_{i-1}) \sup \{f(x); x_{i-1} < x < x_i\}.$$

Dolną całkę Riemanna funkcji $f: [a, b] \rightarrow \mathbb{R}$ nazywamy liczbę

$$L_a^b(f) = \sup \{L(f, P); P \subset [a, b] \text{ jest podziałem}\}.$$

Górną całkę Riemanna funkcji f nazywamy kres dolny

$$U_a^b(f) = \inf \{U(f, P); P \subset [a, b] \text{ jest podziałem}\}.$$

Funkcję f uważamy za *całkowalną w sensie Riemanna*, jeśli zachodzi równość $L_a^b(f) = U_a^b(f)$. Wspólną wartość $\int_a^b f(x) dx$ obu całek nazywamy wówczas *całką Riemanna*.

Sumy górne $U(f, P)$ i dolne $L(f, P)$ są całkami Lebesgue’a specyficznych funkcji prostych ograniczających f z góry i z dołu, a twierdzenie Lebesgue’a (p. [1]) charakteryzuje funkcje całkowalne w sensie Riemanna jako te funkcje ograniczone, dla których zbiór punktów nieciągłości ma miarę Lebesgue’a 0.

Pozostawiamy czytelnikowi sprawdzenie, że każda funkcja f całkowalna w sensie Riemanna jest ℓ -całkowalna i wówczas

$$\int_a^b f(x) dx = \int_{[a,b]} f d\ell.$$

Stosowne konstrukcje i twierdzenia dla całki Riemanna w $\mathbb{R}^m$ czytelnik znajdzie np. w [18].

Przykład 3.1.2 (Konstrukcja Daniella dla miary Lebesgue’a). W przypadku dowolnej przestrzeni z miarą $(T, \mathfrak{T}, \mu)$ przedstawiona wyżej konstrukcja przypomina rozszerzanie funkcjonału $\tilde{\mu}: \mathcal{P}(\mathfrak{T}, \mu) \rightarrow \mathbb{R}$ do całki Lebesgue’a (por. uwaga 1.2.1). Przestrzeń liniowa $\mathcal{P} = \mathcal{P}(\mathfrak{T}, \mu)$ jest krata, funkcjonał $\tilde{\mu}$ jest całką abstrakcyjną, a przestrzeń funkcji $\tilde{\mu}$ -całkowalnych zawiera przestrzeń $\mathcal{I}(T, \mu)$.

Istotnie, dla dowodu inkluzji $\mathcal{I}(T, \mu) \subset \overline{\mathcal{P}}$, wystarczy rozważyć dowolną funkcję μ -całkowalną f . Składowe $f^+, f^- \in \mathcal{I}(T, \mu)$ można przybliżać funkcjami prostymi $0 \leq g \leq f^+$ i $0 \leq h \leq f^-$ w ten sposób, że $\tilde{\mu}(g) > \tilde{\mu}^*(f^+) - \varepsilon$ i $\tilde{\mu}(h) > \tilde{\mu}^*(f^-) - \varepsilon$ dla danego $\varepsilon > 0$. Stąd oszacowanie

$$\tilde{\mu}^*(|f - (g - h)|) \leq \tilde{\mu}^*(f^+ - g) + \tilde{\mu}^*(f^- - h) < 2\varepsilon,$$

w którym $g - h \in \mathcal{P}$. Wobec dowolności $\varepsilon > 0$ oznacza to przynależność f do domknięcia zbioru $\mathcal{P}$. W podrozdziale 3.1.5 wykazujemy, że całka Daniella–Stone’a jest w tym przypadku całką Lebesgue’a, ale – względem miary stanowiącej uzupełnienie miary μ .

3.1.2 Co można pominąć?

Podobnie jak w przypadku całki Lebesgue’a ogólna konstrukcja całki powinna nie tylko wyróżniać przestrzeń liniową funkcji całkowalnych, ale także relację równoważności (pw) pozwalającą traktować rozszerzone funkcje całkowalne jako skończone funkcje mierzalne. Badanie całkowalności rozszerzonych funkcji rzeczywistych zaczynamy zatem od analizy zbiorów i funkcji, których konstrukcja opisana w lemacie 3.1.3 nie rozróżnia. Stąd

Definicja 3.1.7. Funkcję $f: T \rightarrow \overline{\mathbb{R}}$ uznajemy za χ -pomijalną, jeśli zachodzi równość $\chi^*(|f|) = 0$. Zbiór $N \subset T$ będziemy nazywać χ -pomijalnym, jeśli pomijalna jest funkcja charakterystyczna 1_N zbioru N .

W dalszym ciągu dla ustalonego funkcjonału χ będziemy mówić po prostu o funkcjach i zbiorach pomijalnych. Z monotoniczności χ^* wynika, że każdy podzbiór zbioru pomijalnego jest także pomijalny.

Stwierdzenie 3.1.8. Własności zbiorów pomijalnych:

- (i) $\chi^*(|f|) = 0 \iff$ zbiór $\{f \neq 0\}$ jest pomijalny.
- (ii) $\chi^*(f) < \infty \implies$ zbiór $\{f = \infty\}$ jest pomijalny.
- (iii) $\forall_{f, g: T \rightarrow \overline{\mathbb{R}}} \text{zbiór } \{f \neq g\} \text{ jest pomijalny} \implies \chi^*(f) = \chi^*(g)$.
- (iv) Suma przeliczalnej rodziny zbiorów pomijalnych jest zbiorem pomijalnym.

Dowód. Dla dowolnej funkcji $f: T \rightarrow \overline{\mathbb{R}}$, oznaczmy $N = \{f \neq 0\}$. Zależności $1_N \leq \sup_n(n|f|)$ oraz $|f| \leq \sup_n(n \cdot 1_N)$ pozwalają wywnioskować równoważność (i) z własności (3.5).

Jeśli $\chi^*(f) < \infty$, to istnieje funkcja $g \in \mathcal{A}^*$ taka, że $\chi^*(g) < \infty$ i $f \leq g$, a więc dla dowodu (ii) wystarczy wykazać pomijalność zbioru $N' = \{g = \infty\}$. W tym celu ustalmy w $\mathcal{A}$ dowolną (skończoną) funkcję $h \leq g$. Funkcja

$g - h \in \mathcal{A}^*$ jest nieujemna i nieskończona na zbiorze N' , a więc nierówności $n\chi^*(1_{N'}) \leq \chi^*(g - h) < \infty$ dla $n \in \mathbb{N}$ implikują $\chi^*(1_{N'}) = 0$.

Jeśli funkcje f i g są równe poza pomijalnym zbiorem N , to porównując każdą z nich z $\max(f, g)$ możemy sprowadzić implikację (iii) do przypadku $f \leq g$, co daje od razu nierówność $\chi^*(f) \leq \chi^*(g)$. Aby wykazać nierówność przeciwną, założymy, że $\chi^*(f) \neq \infty$ i weźmy w $\mathcal{A}^*$ dowolną funkcję $h \geq f$. Dla funkcji $u \geq 0$ nieskończonej na N i równej 0 na $T \setminus N$, zachodzi nierówność $g \leq h + u$, a zatem

$$\chi^*(g) \leq \chi^*(h) + \chi^*(u) = \chi^*(h).$$

Wobec dowolności $h \geq f$ oznacza to brakującą nierówność $\chi^*(g) \leq \chi^*(f)$.

Pozostaje sprawdzić pomijalność sumy $N = \bigcup_n N_n$ ciągu zbiorów pomijalnych $N_n, n \in \mathbb{N}$. Z własności miary zewnętrznej wynika oszacowanie

$$\chi^*(1_N) = \lim_{n \rightarrow \infty} \chi^*(1_{N_1 \cup \dots \cup N_n}) \leq \lim_{n \rightarrow \infty} \sum_{i \leq n} \chi^*(1_{N_i}) = 0,$$

co kończy dowód. $\square$

Będziemy korzystać z klasycznej konwencji i oznaczeń dotyczących własności zależnych od $x \in T$ prawdziwych dla *prawie wszystkich* x czyli χ -*prawie wszędzie* – traktując zbiory pomijalne jako *zbiory miary zero*. W tym sensie w stwierdzeniu 3.1.8 własność (ii) dotyczy funkcji pw. skończonych, a pozostałe własności – równości pw.

Definicja 3.1.9. Funkcje $f, g: T \rightarrow \overline{\mathbb{R}}$ nazywamy równymi χ -prawie wszędzie, co zapisujemy w skrócie $f \stackrel{\text{pw}}{=} g$, jeśli zbiór $\{x \in T; f(x) \neq g(x)\}$ jest χ -pomijalny.

Ćwiczenie 3.1.3. Sprawdzić, że relacja $\stackrel{\text{pw}}{=}$ jest równoważnością w zbiorze $\overline{\mathbb{R}}^T$. Dla funkcji skończonych, $f \stackrel{\text{pw}}{=} g$ wtedy i tylko wtedy, gdy różnica $f - g$ jest funkcją pomijalną.

Wprost z definicji wynika, iż skończone funkcje pomijalne tworzą przestrzeń liniową $[0]_\chi \subset \mathcal{I}(T, \chi)$, a relacja równoważności $\stackrel{\text{pw}}{=}$ prowadzi do przestrzeni liniowej ilorazowej $L(T, \chi) := \mathcal{M}(T, \chi)/[0]_\chi$, na której N_1 indukuje normę

$$\|[f]\| := \chi^*(|f|), \quad \text{dla } [f] \in L(T, \chi), \quad (3.11)$$

gdzie poprzez $[f] = f + [0]_\chi$ oznaczamy klasę równoważności (warstwę) dowolnej funkcji $f \in \mathcal{M}(T, \chi)$.

Wniosek 3.1.10. *Przestrzeń ilorazowa $L^1(T, \chi) = \mathcal{I}(T, \chi)/[0]_\chi$ jest podprzestrzenią w $L(T, \chi)$. Całka Daniella–Stone’a wyznacza funkcjonal liniowy i ciągły (oznaczany tym samym symbolem)*

$$\int d\chi: L^1(T, \chi) \longrightarrow \mathbb{R}, \quad [f] \mapsto \int f d\chi,$$

którego wartości nie zależą od wyboru funkcji całkowlanej w ramach warstwy względem podprzestrzeni $[0]_\chi$. $\square$

3.1.3 Funkcje całkowlane wg Daniella i Stone’a

Dla uzyskania jawnej charakteryzacji χ –całkowalności rozszerzamy klasę funkcji całkowlanych i przyjmujemy:

Definicja 3.1.11. *Funkcję $f: T \rightarrow \overline{\mathbb{R}}$ nazywamy χ –całkowalną, gdy jest równoważna funkcji całkowlanej o wartościach skończonych. Zbiór rozszerzonych funkcji χ –całkowlanych będziemy oznaczać $\overline{\mathcal{I}}(T, \chi)$, tzn.*

$$\overline{\mathcal{I}}(T, \chi) = \{f \in \overline{\mathbb{R}}^T; \exists g \in \mathcal{I}(T, \chi) f \stackrel{\text{pw}}{=} g\}.$$

Niech $\mathcal{A}_* \subset \overline{\mathbb{R}}^T$ oznacza – dla symetrii – zbiór funkcji osiągalnych z podprzestrzeni $\mathcal{A}$ z góry. Zachodzi oczywista równoważność $f \in \mathcal{A}_* \iff -f \in \mathcal{A}^*$.

Twierdzenie 3.1.12. *Na przestrzeni $\mathcal{I}(T, \chi)$ całka Daniella–Stone’a $\int d\chi$ jest równa mierze zewnętrznej χ^* . Miara zewnętrzna χ^* jest jedynym rozszerzeniem całki na zbiór $\overline{\mathcal{I}}(T, \chi)$ zgodnym z relacją równoważności $\stackrel{\text{pw}}{=}$. Ponadto, funkcje χ –całkowlane mają następujące własności:*

(i) *Funkcja $f: T \rightarrow \overline{\mathbb{R}}$ jest całkowlana $\iff$ zachodzi jeden z równoważnych warunków:*

$$\chi^*(f) + \chi^*(-f) = 0 \quad (3.12)$$

$$\forall \varepsilon > 0 \exists g \in \mathcal{A}^* \exists h \in \mathcal{A}_* h \leq f \leq g \text{ i } \chi^*(g - h) < \varepsilon \quad (3.13)$$

$$\forall \varepsilon > 0 \exists u \in \mathcal{A} \chi^*(|f - u|) < \varepsilon. \quad (3.14)$$

(ii) *Jeśli f jest funkcją całkowlaną, to funkcje $|f|$, f^+ i f^- są także całkowlane. Dla dowolnych $f, g \in \overline{\mathcal{I}}(T, \chi)$, funkcje $\max(f, g)$ i $\min(f, g)$ są również całkowlane.*

(iii) *Funkcja $f \in \mathcal{A}^*$ osiągalna z $\mathcal{A}$ z dołu jest całkowlana $\iff \chi^*(f) < \infty$.*

Naturalnym rozszerzeniem definicji 3.1.6 jest:

Definicja 3.1.13. Całką Daniella–Stone’a względem miary abstrakcyjnej $\chi: \mathcal{A} \rightarrow \mathbb{R}$ w przestrzeni T nazywamy funkcję

$$\int d\chi: \overline{\mathcal{I}}(T, \chi) \longrightarrow \mathbb{R}$$

będącą ograniczeniem miary zewnętrzną χ^* do zbioru funkcji całkowalnych.

Ćwiczenie 3.1.4. Przestrzeń $\overline{\mathcal{I}}(T, \chi)$ skończonych funkcji całkowalnych jest kratą, a całka $\int d\chi$ – miarą abstrakcyjną w sensie definicji 3.1.1. Sprawdzić, że wyznaczona przez całkę miara zewnętrzna jest identyczna z χ^* .

Wniosek 3.1.14. Dla dowolnego monotonicznego ciągu funkcji χ –całkowalnych $f_n \in \overline{\mathcal{I}}(T, \chi)$, $n \in \mathbb{N}$, granica $f = \lim_{n \rightarrow \infty} f_n$ jest χ –całkowalna wtedy i tylko wtedy, gdy całki $\int f_n d\chi$, $n \in \mathbb{N}$, są wspólnie ograniczone. Zachodzi wówczas równość

$$\int f d\chi = \lim_{n \rightarrow \infty} \int f_n d\chi.$$

Dowód. Rozważmy dowolny niemalejący ciąg funkcji całkowalnych $f_n \nearrow f$. Wynikająca z (3.5) równość $\sup_n \chi^*(f_n) = \chi^*(f)$ oznacza, że warunek wspólnej ograniczoności miar $\chi^*(f_n)$ jest konieczny dla całkowalności funkcji granicznej f . Załóżmy zatem, że jest spełniony i dla dowolnego ustalonego $\varepsilon > 0$ wybierzmy $n \in \mathbb{N}$ takie, by $\chi^*(f) \leq \chi^*(f_n) + \frac{\varepsilon}{2}$. Skoro całkowalność f_n implikuje nierówność $\chi^*(|f_n - u|) < \frac{\varepsilon}{2}$ dla pewnego $u \in \mathcal{A}$, podobne oszacowanie ma miejsce dla odległości f i u ,

$$\begin{aligned} \chi^*(|f - u|) &\leq \chi^*(f - f_n) + \chi^*(|f_n - u|) \\ &< \chi^*(f) + \chi^*(-f_n) + \frac{\varepsilon}{2} \leq \varepsilon. \end{aligned}$$

Wobec (3.14) oznacza to całkowalność f . Przypadek ciągu nierosnącego $f_n \searrow f$ sprowadzamy do poprzedniego poprzez zmianę znaku. $\square$

Dowód twierdzenia 3.1.12. W pierwszej kolejności wykażemy równoważności sformułowane w punkcie (i). Ustalmy dowolne $\varepsilon > 0$. Jeśli funkcja $f \in \overline{\mathbb{R}}^T$ ma własność (3.12), to obie miary zewnętrzne $\chi^*(f)$ i $\chi^*(-f)$ są skończone, a zatem istnieją funkcje $g, h' \in \mathcal{A}^*$ takie, że $f \leq g$ i $-f \leq h'$, a także $\chi^*(g) < \chi^*(f) + \frac{\varepsilon}{2}$, $\chi^*(h') < \chi^*(-f) + \frac{\varepsilon}{2}$. Przyjmując $h = -h' \in \mathcal{A}_*$, otrzymujemy $h \leq f \leq g$ i

$$\chi^*(g - h) \leq \chi^*(g) + \chi^*(h') < \varepsilon,$$

co oznacza własność (3.13).

Dla dowodu implikacji (3.13) $\Rightarrow$ (3.14) zauważmy, że w przypadku nierówności $h \leq f \leq g$ dla pewnych $g, -h \in \mathcal{A}^*$, osiągalność funkcji g z dołu

pozwała wskazać w $\mathcal{A}$ dowolnie bliskie $u \leq g$, przy czym ma miejsce oszacowanie

$$|f - u| \leq \max(g - h, g - u) \leq (g - h) + (g - u).$$

Jeśli zatem $\chi^*(g - h) = \varepsilon' < \varepsilon$, to dla $u \in \mathcal{A}$ takiego, że $\chi(u) > \chi^*(g) - (\varepsilon - \varepsilon')$, otrzymujemy

$$\chi^*(|f - u|) \leq \chi^*(g - h) + \chi^*(g - u) < \varepsilon.$$

Jeśli funkcje f oraz $u \in \mathcal{A}$ spełniają nierówność $\chi^*(|f - u|) < \varepsilon$, to zgodnie z (3.4) istnieje osiągalna z dołu funkcja $g \geq |f - u|$ taka, że $\chi^*(g) < \varepsilon$. Wynikające stąd nierówności $\pm f \leq g \pm u$ pozwalają na oszacowanie

$$\begin{aligned} 0 \leq \chi^*(f) + \chi^*(-f) &\leq \chi^*(g + u) + \chi^*(g - u) \\ &= \chi^*(g) + \chi(u) + \chi^*(g) - \chi(u) < 2\varepsilon, \end{aligned}$$

co wobec dowolności $\varepsilon > 0$ oznacza implikację (3.14) $\Rightarrow$ (3.12).

Pozostaje zauważyć, że ostatnia z badanych własności implikuje skończoność miary zewnętrznej $\chi^*(|f|)$, przy czym dla równoważnej z f funkcji skończonej $g \stackrel{\text{pw}}{=} f$ własność (3.14) oznacza dokładnie przynależność do domknięcia podprzestrzeni $\mathcal{A} \subset \mathcal{M}(T, \chi)$ – czyli całkowalność g i f .

Nierówność $||f| - u| \leq |f - u|$ oraz własność $|u| \in \mathcal{A}$ dla $u \in \mathcal{A}$ oznaczają, że całkowalność funkcji f pociąga za sobą całkowalność modułu $|f|$ – zgodnie z warunkiem (3.14). Stąd także całkowalność funkcji $f^\pm = (|f| \pm f)/2$. Druga część własności (ii) wynika bezpośrednio z pierwszej.

W przypadku własności (iii) wystarczy pokazać, że dla dowolnej funkcji $f \in \mathcal{A}^*$ osiągalnej z $\mathcal{A}$ warunek $\chi^*(f) < \infty$ zapewnia całkowalność. Skończoność miary $\chi^*(f)$ oznacza, że dla dowolnego $\varepsilon > 0$ istnieje funkcja $u \in \mathcal{A}$ taka, że $u \leq f$ i $\chi^*(f) < \chi(u) + \varepsilon$. Stąd – nierówność $\chi^*(|f - u|) < \varepsilon$ i całkowalność f .

Możemy teraz powrócić do tezy stanowiącej pierwszą część twierdzenia. Zauważmy, że udowodniona już własność (3.12) zapewnia liniowość i ograniczoność miary zewnętrznej χ^* na przestrzeni $\mathcal{I}(T, \chi)$. Istotną w tym przypadku addytywność dają oszacowania $\chi^*(f + g) \leq \chi^*(f) + \chi^*(g)$ oraz $\chi^*(-f - g) \leq \chi^*(-f) + \chi^*(-g)$, z których dla f i g całkowalnych wynika

$$0 \leq \chi^*(f + g) + \chi^*(-f - g) \leq 0.$$

Ograniczoność względem półnormy N_1 ,

$$|\chi^*(f)| = |\chi^*(f^+) - \chi^*(-f^-)| \leq \chi^*(f^+) + \chi^*(f^-) = \chi^*(|f|)$$

dla $f \in \mathcal{I}(T, \chi)$, oznacza, że to miara zewnętrzna χ^* jest (jedynym) ciągłym rozszerzeniem miary χ na przestrzeń funkcji całkowalnych. $\square$

Wniosek 3.1.15. *Dla każdej funkcji χ -całkowalnej $f \in \overline{\mathcal{I}}(T, \chi)$ istnieją funkcje całkowalne g i h równe χ -prawie wszędzie i takie, że $h \leq f \leq g$, przy czym g jest granicą nierosnącego ciągu funkcji całkowalnych $g_n \in \mathcal{A}^*$, a h – granicą niemalejącego ciągu funkcji całkowalnych $h_n \in \mathcal{A}_*$. $\square$*

Uwaga 3.1.5. Wskazane we wniosku funkcje g i h są mieralne względem σ -pierścienia $\mathcal{A}^\sigma$ generowanego przez $\mathcal{A}$. Każda klasa równoważności $[f] \in L^1(T, \chi)$ zawiera funkcje całkowalne $\mathcal{A}^\sigma$ -mieralne, a sama relacja równoważności i własność χ -pomijalności (funkcji i zbiorów) także dają się zdefiniować poprzez χ -pomijalne funkcje $\mathcal{A}^\sigma$ -mieralne.

3.1.4 Klasyczne twierdzenia graniczne dla całek

W odniesieniu do klas równoważnych (pw. równych) funkcji całkowalnych zachodzi następująca własność, stanowiąca uzupełnienie wniosku 3.1.14 i rozszerzenie wniosku 1.2.21.

Wniosek 3.1.16 (Twierdzenie o monotonicznej zbieżności). *Dla dowolnej abstrakcyjnej miary χ w przestrzeni T , całka $\int d\chi: L^1(T, \chi) \rightarrow \mathbb{R}$ ma własność monotonicznej ciągłości,*

$$\lim_{n \rightarrow \infty} \int f_n d\chi = \int \lim_{n \rightarrow \infty} f_n d\chi, \quad (3.15)$$

dla dowolnego monotonicznego ciągu $f_n \in L^1(T, \chi)$, $n \in \mathbb{N}$, takiego, że całki $\int f_n d\chi$, $n \in \mathbb{N}$, są wspólnie ograniczone. $\square$

Własność monotonicznej ciągłości, kluczowa dla konstrukcji całki Daniella–Stone’a i ważna w zastosowaniach, ma naturalne konsekwencje dla „zwykłych” ciągów funkcyjnych, których ewentualna zbieżność nie musi być w żadnej mierze monotoniczna. Wykorzystanie ciągów monotonicznych ułatwia następujący

Lemat 3.1.17. *Dla dowolnego ciągu funkcji χ -całkowalnych $(f_n)_{n \in \mathbb{N}}$, ograniczonych z góry przez funkcję χ -całkowalną, kres górny $f = \sup_n f_n$ jest funkcją χ -całkowalną. Kres dolny $\inf_n f_n$ jest całkowalny, jeśli funkcje f_n , $n \in \mathbb{N}$, są wspólnie ograniczone z dołu przez funkcję całkowalną.*

Dowód. Zgodnie z twierdzeniem 3.1.12(ii) funkcje $g_n = \max(f_1, \dots, f_n)$, $n \in \mathbb{N}$, są całkowalne i tworzą ciąg niemalejący zbieżny do $f = \sup_n f_n$. Wspólna ograniczoność (z góry) przez funkcję całkowalną g zapewnia oszacowanie $\sup_n \int g_n d\chi \leq \int g d\chi$, co implikuje całkowalność funkcji f . Druga część tezy wynika z równości $\inf_n f_n = -\sup_n (-f_n)$. $\square$

Twierdzenie 3.1.18 (Lemat Fatou). *Niech $f_n \in \overline{\mathbb{R}}^T$, $n \in \mathbb{N}$, będzie dowolnym ciągiem funkcji ograniczonym z dołu przez funkcję χ –całkowalną. Wówczas*

$$\chi^*(\liminf_{n \rightarrow \infty} f_n) \leq \liminf_{n \rightarrow \infty} \chi^*(f_n). \quad (3.16)$$

Jeśli funkcje f_n są całkowne i spełniają nierówność $|f_n| \leq g$, dla $n \in \mathbb{N}$ i pewnej funkcji całkownej g , to obie granice: dolna $\liminf_{n \rightarrow \infty} f_n$ i górna $\limsup_{n \rightarrow \infty} f_n$ są funkcjami całkowalnymi i zachodzą nierówności

$$\int \liminf_{n \rightarrow \infty} f_n d\chi \leq \liminf_{n \rightarrow \infty} \int f_n d\chi \leq \limsup_{n \rightarrow \infty} \int f_n d\chi \leq \int \limsup_{n \rightarrow \infty} f_n d\chi. \quad (3.17)$$

Dowód. Ćwiczenie. Dlaczego obie granice są funkcjami mierzalnymi? $\square$

Wniosek 3.1.19 (Twierdzenie Lebesgue’a o zmajoryzowanej zbieżności). *Dla dowolnego ciągu funkcji χ –całkowalnych $(f_n)_{n \in \mathbb{N}}$ zbieżnego (pw.) do funkcji f i zmajoryzowanego przez funkcję całkowną, tzn. takiego, że $|f_n| \leq g$ (pw), dla wszystkich $n \in \mathbb{N}$ i pewnej funkcji $g \in \mathcal{I}(T, \chi)$, funkcja graniczna f jest χ –całkowalna i zachodzą równości*

$$\lim_{n \rightarrow \infty} \int f_n d\chi = \int f d\chi, \quad \text{oraz} \quad (3.18)$$

$$\lim_{n \rightarrow \infty} \int |f - f_n| d\chi = 0. \quad (3.19)$$

Dowód. Całkowalność funkcji f wynika bezpośrednio z twierdzenia 3.1.18, a równość (3.18) – z nierówności (3.17). Równość (3.19) wynika ze zmajoryzowanej zbieżności $|f - f_n| \xrightarrow{\text{pw}} 0$. $\square$

3.1.5 Powrót do całki Lebesgue’a

Nazwisko Stone związane jest z opisaną wyżej konstrukcją całki ze względu na autorstwo twierdzenia, w którym całka Daniella–Stone’a okazuje się być – przy dość naturalnych założeniach – tożsama z całką Lebesgue’a. Sformułowanie i dowód twierdzenia Stone’a poprzedzimy wstępną charakterystyką dwu ważnych dla konstrukcji σ –pierścieni oraz miary wyznaczonej przez całkę $\int d\chi$. Motywacją dla poniższej definicji jest własność sformułowana w lemacie 1.2.5.

Definicja 3.1.20. *Podprzestrzeń stanowiąca kratę $\mathcal{A} \subset \mathbb{R}^T$ jest kratą Stone’a, jeśli prawdziwa jest implikacja*

$$f \in \mathcal{A} \implies \min(f, 1) \in \mathcal{A}, \quad (3.20)$$

nazywana także warunkiem Stone’a.

Warunek Stone’a implikuje zamkniętość przestrzeni $\mathcal{A}$ względem operacji $\min(f, c)$ oraz $\max(f, -c)$ dla $c > 0$ i dowolnej funkcji $f \in \mathcal{A}$, i jest automatycznie spełniony w przypadku kraty zawierającej funkcje stałe.

Ćwiczenie 3.1.6. Sprawdzić, że warunek Stone’a dla dziedziny całki abstrakcyjnej χ przenosi się na kratę $\mathcal{I}(T, \chi)$ funkcji χ -całkowalnych, tzn.

$$f \in \mathcal{I}(T, \chi) \implies \min(f, 1) \in \mathcal{I}(T, \chi). \quad (3.21)$$

W tym celu zauważyć, że dla całkowalnej funkcji f i liczby $\varepsilon > 0$ dowolna osiągalna z dołu funkcja $g \in \mathcal{A}^*$ spełniająca warunki $g \geq f$ i $\chi^*(g - f) < \varepsilon$, wyznacza całkowalną funkcję $g' = \min(g, 1)$ równie bliską badanej funkcji $f' = \min(f, 1)$.

Jak wiemy – zgodnie z przykładem 1.1.7 – niezależnie od miary abstrakcyjnej $\chi: \mathcal{A} \rightarrow \mathbb{R}$ i wyznaczonej przez nią całki krata $\mathcal{A}$ generuje σ -pierścień $\mathcal{A}^\sigma$ i σ -ciało $\sigma\mathcal{A}$, względem których wszystkie funkcje przestrzeni liniowej $\mathcal{A}$ są mierzalne. Z drugiej strony, całka Daniella–Stone’a pozwala na wyróżnienie w T zbiorów związanych z całkowalnością. Wychodząc od rodziny $\mathfrak{D}_\chi^f = \{A \subset T; 1_A \in \mathcal{I}(T, \chi)\}$, przyjmijmy oznaczenia

$$\mathfrak{D}_\chi^\sigma = \sigma(\mathfrak{D}_\chi^f) \quad \text{oraz} \quad \mathfrak{D}_\chi = \sigma_T(\mathfrak{D}_\chi^f) \quad (3.22)$$

dla σ -pierścienia i σ -ciała generowanych przez rodzinę wszystkich zbiorów o całkowalnych funkcjach charakterystycznych.

Twierdzenie 3.1.21 (Stone). Niech χ będzie dowolną miarą abstrakcyjną w przestrzeni mierzalnej T , określoną na kracie Stone’a $\mathcal{A} \subset \mathbb{R}^T$.

(i) Przestrzeń $(T, \mathfrak{D}_\chi^\sigma, \varsigma_\chi)$ z miarą

$$\varsigma_\chi: \mathfrak{D}_\chi^\sigma \ni A \mapsto \chi^*(1_A) \in \overline{\mathbb{R}}_+ \quad (3.23)$$

jest uzupełnieniem jedynej miary ς na σ -pierścieniu $\mathcal{A}^\sigma$, dla której $\chi = \int d\varsigma$ na przestrzeni $\mathcal{A}$. Dla zbiorów $A \in \mathfrak{D}_\chi \setminus \mathfrak{D}_\chi^\sigma$ zachodzi równość $\chi^*(1_A) = \infty$.

(ii) Całka Daniella–Stone’a $\int d\chi$ jest równa całce Lebesgue’a $\int d\varsigma_\chi$. Obie całki mają te same przestrzenie funkcji całkowalnych

$$\mathcal{I}(T, \chi) = \mathcal{I}(T, \varsigma_\chi) \quad \text{i} \quad L^1(T, \chi) = L^1(T, \varsigma_\chi).$$

Definicja 3.1.22. Miarę ς_χ w półprzestrzeni mierzalnej $(T, \mathfrak{D}_\chi^\sigma)$ oraz jej trywialne rozszerzenie na przestrzeń $(T, \mathfrak{D}_\chi)$ będziemy nazywać miarą Daniella–Stone’a wyznaczoną przez miarę abstrakcyjną χ .

Dowód. Z lematu 1.2.5 warunku Stone’a wynika, że każda z funkcji charakterystycznych zbiorów $\{f > c\}$, dla $f \in \mathcal{A}$ i $c > 0$, jest osiągalna z dołu przez ciąg funkcji $f_n = \min(n(f - \min(f, c)), 1)$. Jeśli zatem ς jest miarą na $\mathcal{A}^\sigma$ i $\chi = \int d\chi$ na $\mathcal{A}$, to z równości

$$\varsigma\{f > c\} = \lim_n \chi(f_n) \leq c^{-1} \chi(f) \quad \text{ i } \quad \varsigma\{f \geq 1\} = \lim_n \varsigma\{f > 1 - \frac{1}{n}\}$$

wynika jednoznaczność ς oraz inkluzje $\mathcal{A}^f \subset \mathfrak{D}_\chi^\sigma$ i $\mathcal{A}^\sigma \subset \mathfrak{D}_\chi^\sigma$.

Dla dowodu faktu, że wzór (3.23) definiuje miarę, zauważmy, że skoro każdy zbiór $A \in \mathfrak{D}_\chi^\sigma$ jest sumą wstępującego ciągu zbiorów z $\mathfrak{D}_\chi^f$, prawdziwa jest równoważność $\chi^*(1_A) < \infty \iff A \in \mathfrak{D}_\chi^f$, dla $A \in \mathfrak{D}_\chi^\sigma$. Dzięki temu skończona addytywność funkcji ς_χ na $\mathfrak{D}_\chi^\sigma$ wynika z addytywności (liniowości) całki na przestrzeni $\mathcal{I}(T, \chi)$ i subaddytywności χ^* , natomiast ciągłość z dołu, równoważna σ -addytywności jest konsekwencją własności (3.5). Pozostaje sprawdzić równość $\chi^*(1_B) = \infty$ dla $B \in \mathfrak{D}_\chi \setminus \mathfrak{D}_\chi^\sigma$. Przypuśćmy przeciwnie, że dopełnienie rodziny $\mathfrak{D}_\chi^\sigma$ zawiera zbiór $B \in \mathfrak{D}_\chi$, dla którego $\chi^*(1_B) < \infty$. Wówczas istnieje funkcja całkowalna $g \in \mathcal{A}^*$ taka, że $1_B \leq g$, a zatem B jest podzbiorem zbioru $\{g \geq 1\} \in \mathcal{A}^\sigma$, co oznacza sprzeczność z tezą twierdzenia 1.1.13(i).

Dla zakończenia dowodu własności (i) pozostaje wykazać związek pomiędzy σ -pierścieniami $\mathcal{A}^\sigma$ i $\mathfrak{D}_\chi^\sigma$. Rozważmy zatem dowolny zbiór $A \in \mathfrak{D}_\chi^f$, a więc taki, że funkcja 1_A jest χ -całkowalna. Zgodnie z wnioskiem 3.1.15 istnieje funkcja całkowalna $g \geq 1_A$ będąca granicą nierosnącego ciągu $(g_n)_{n \in \mathbb{N}}$ funkcji całkowalnych i taka, że $\int g d\chi = \int 1_A d\chi$. Dzięki temu nierówności $1_A \leq 1_G \leq g$, dla zbioru $G := \{g \geq 1\} \in \mathcal{A}^\sigma$, prowadzą do zależności

$$\int 1_G d\chi = \int 1_A d\chi \quad \text{ i } \quad \int 1_{G \setminus A} d\chi = 0.$$

Wynika stąd, że zbiór $N = G \setminus A$ jest χ -pomijalny. Udowodniliśmy tym samym, że każdy zbiór $A \in \mathfrak{D}_\chi^f$ jest różnicą zbioru $G \in \mathcal{A}^\sigma$ i zbioru pomijalnego. Ponieważ rodzina $\mathfrak{D}_\chi^f$ jest δ -pierścieniem, własność ta przenosi się na σ -pierścień $\mathfrak{D}_\chi^\sigma$ (por. twierdzenie 1.1.13).

Przystępując do dowodu własności (ii) zauważmy, że dla dowolnej nieujemnej χ -całkowalnej funkcji $g \in \mathcal{I}(T, \chi)$ i liczby $c > 0$, z warunku Stone’a i lematu 1.2.5 wynika całkowalność funkcji charakterystycznej $1_{\{g > c\}}$, a zatem dla każdego $c > 0$ zbiór $\{g > c\}$ należy do $\mathfrak{D}_\chi^f$. Pozostając przy funkcjach nieujemnych, dowolną funkcję $f \geq 0$ możemy przedstawić jako granicę niemalejącego ciągu

$$f_n = \sum_{m=0}^{n \cdot 2^n - 1} \frac{m}{2^n} \cdot 1_{\{\frac{m}{2^n} < f \leq \frac{m+1}{2^n}\}} + n \cdot 1_{\{n < f\}}, \quad n \in \mathbb{N}. \quad (3.24)$$

Dodatkowe założenie χ -całkowalności f pociąga za sobą całkowalność każdej z funkcji f_n zarówno względem χ , jak i względem miary ς_χ – oraz równość całek $\int f_n d\chi = \int f_n d\varsigma_\chi$. Z drugiej strony, założenie ς_χ -całkowalności funkcji f implikuje ograniczenie dla miary

$$\varsigma_\chi(\{f > c\}) \leq c^{-1} \int f d\varsigma_\chi < \infty \quad \text{dla } c > 0,$$

co pociąga za sobą przynależność zbioru $\{f > c\}$ do $\mathfrak{D}_\chi^f$ oraz χ -całkowalność funkcji charakterystycznej, dla każdego $c > 0$. Tym samym także w tym przypadku każda z funkcji f_n jest całkowalna zarówno względem χ , jak i względem miary ς_χ , a odpowiednie całki są równe. Z monotonicznej ciągłości każdej z całek wynika zatem inkluzja

$$\mathcal{I}_+(T, \chi) \cup \mathcal{I}_+(T, \varsigma_\chi) \subset \mathcal{I}(T, \chi) \cap \mathcal{I}(T, \varsigma_\chi),$$

a więc równość zarówno przestrzeni funkcji całkowalnych jak i samych całek. Klasy abstrakcji funkcji równych prawie wszędzie są równe, co wynika z faktu, iż zbiory χ -pomijalne są tożsame ze zbiorami miary 0. $\square$

Przykład 3.1.7. *Całka względem miary określonej na półpierzścieniu.* Dla dowolnego półpierzścienia $\mathfrak{I}$ w przestrzeni T przestrzeń liniowa

$$\mathcal{P}_\mathfrak{I} = \left\{ \sum_{i \leq n} c_i \cdot 1_{A_i}; c_i \in \mathbb{R}, A_i \in \mathfrak{I}, \text{ dla } i \leq n, n \in \mathbb{N} \right\}$$

jest kratą Stone’a. Istotnie, z definicji półpierzścienia łatwo wywnioskować, że funkcje z $\mathcal{P}_\mathfrak{I}$ są *proste* – można je przedstawić jako sumy $\sum_{i \leq n} c_i \cdot 1_{A_i}$, w których zbiory $A_i \in \mathfrak{I}$ są wzajemnie rozłączne. Korzystając z [2] i powtarzając pierwszą część dowodu twierdzenia 1.2.3 – dla dowolnej skończonej miary μ na półpierzścieniu $\mathfrak{I}$ – stwierdzamy, iż funkcjonal

$$\chi: \mathcal{P}_\mathfrak{I} \ni \sum_{i \leq n} c_i \cdot 1_{A_i} \mapsto \sum_{i \leq n} c_i \mu(A_i) \in \mathbb{R}$$

jest poprawnie określoną abstrakcyjną miarą. Pozostawiamy czytelnikowi sprawdzenie, że całka Daniella–Stone’a $\int d\chi$ wyznacza na σ -ciele $\sigma_T(\mathfrak{I})$ miarę stanowiącą *minimalne* rozszerzenie miary μ . Z drugiej strony, miara Daniella–Stone’a ς_χ w przestrzeni $(T, \mathfrak{D}_\chi)$ jest minimalnym rozszerzeniem *zupełnym* obu miar, zgodnym z miarą Carathéodory’ego w zakresie składowych σ -skończonych – porównaj [2, par. 1.5.3].

3.2 Miary Radona

W bieżącym podrozdziale prezentujemy kolejne zastosowanie abstrakcyjnej konstrukcji Daniella–Stone’a rozwijanej w podrozdziale 3.1. Klasyczna konstrukcja całki Radona jest wcześniejsza, wychodzi – zamiast od przestrzeni

mierzalnej – od lokalnie zwartej przestrzeni topologicznej i oznacza zgodny z ogólnym abstrakcyjnym schematem proces rozszerzania funkcjonału liniowego określonego początkowo na przestrzeni funkcji ciągłych o zwartych nośnikach. Dodatkową cechą miary Radona i związanej z nią całki jest włączenie do dziedziny funkcji zespolonych – naturalne np. dla znanych z analizy zespolonej całek krzywoliniowych, oraz regularność – opisana dokładniej w kolejnym podrozdziale.

W odniesieniu do przestrzeni topologicznych lokalnie zwartych przywołujemy za [2, podr. 1.6]

Twierdzenie 3.2.1 (O rozkładzie jedynki). *W przestrzeni lokalnie zwartej T , dla dowolnej skończonej rodziny zbiorów otwartych $\{V_i; i \leq m\}$ i zbioru zwartego $K \subset \bigcup_{i \leq m} V_i$, istnieją nieujemne funkcje ciągłe $f_1, \dots, f_m \in \mathcal{C}_c(T)$ takie, że*

$$(i) \sum_{i \leq m} f_i(x) = 1, \text{ dla } x \in K \text{ oraz}$$

$$(ii) \forall_{i \leq m} \operatorname{supp}(f_i) \subset V_i. \quad \square$$

3.2.1 Topologie zbieżności (niemal) jednostajnej

W przestrzeni lokalnie zwartej (T, τ) niech $\mathfrak{K} = \mathfrak{K}_T$ będzie rodziną wszystkich zwartych podzbiorów. Dla $\mathbb{F}$ oznaczającego ciało liczb rzeczywistych $\mathbb{R}$ lub ciało liczb zespolonych $\mathbb{C}$, niech $\mathcal{C}(T, \mathbb{F})$ oznacza przestrzeń liniową funkcji ciągłych na T o wartościach w $\mathbb{F}$, natomiast

$$\mathcal{C}_c(T, \mathbb{F}) = \{f \in \mathcal{C}(T, \mathbb{F}); \operatorname{supp}(f) \in \mathfrak{K}\} \quad (3.25)$$

– podprzestrzeń funkcji o zwartym nośniku. Przypominamy, że nośnikiem funkcji ciągłej f na T nazywamy domknięcie zbioru otwartego na którym $f \neq 0$, tzn. zbiór domknięty

$$\operatorname{supp}(f) = \overline{f^{-1}(\mathbb{F} \setminus \{0\})}. \quad (3.26)$$

Dla każdego zbioru zwartego $K \subset T$ przestrzeń $\mathcal{C}(K, \mathbb{F})$ ma naturalną normę i topologię zbieżności jednostajnej. Z drugiej strony, dla dowolnego zwartego podzbioru $K \subset T$, norma jednostajna $\|f\| := \sup |f|$ wyznacza topologię w przestrzeni $\mathcal{C}_K(T, \mathbb{F}) \subset \mathcal{C}_c(T, \mathbb{F})$ funkcji ciągłych o nośniku zawartym w K . Z definicji (3.26) wynika charakteryzacja

$$\mathcal{C}_K(T, \mathbb{F}) = \{f \in \mathcal{C}_c(T, \mathbb{F}); f|_{T \setminus K} = 0\}, \text{ dla } K \in \mathfrak{K}_T. \quad (3.27)$$

Definicja 3.2.2. Topologią zbieżności niemal jednostajnej w $\mathcal{C}(T, \mathbb{F})$ nazywamy najmniejszą topologię $\tau_{\mathcal{C}}$ zapewniającą ciągłość odwzorowań

$$\text{pr}_K: \mathcal{C}(T, \mathbb{F}) \ni f \mapsto f|_K \in \mathcal{C}(K, \mathbb{F}), \quad \text{dla } K \in \mathfrak{K}_T.$$

Topologią zbieżności jednostajnej w $\mathcal{C}_c(T, \mathbb{F})$ nazywamy największą topologię τ_c , w której odwzorowania inkluzji

$$\iota_K: \mathcal{C}_K(T, \mathbb{F}) \hookrightarrow \mathcal{C}_c(T, \mathbb{F}), \quad \text{dla } K \in \mathfrak{K}_T$$

są włożeniami (homeomorfizmami) na podprzestrzenie domknięte.

Podstawowe własności topologii zbieżności jednostajnej gromadzi

Twierdzenie 3.2.3. (i) Topologia $\tau_{\mathcal{C}}$ jest generowana przez podbazę złożoną z przeciwobrazów – przez odwzorowanie obcięcia – wszystkich zbiorów otwartych w przestrzeniach $\mathcal{C}(K, \mathbb{F})$. Bazę topologii $\tau_{\mathcal{C}}$ stanowią zbiory (relatywne kule) postaci

$$B(g, K, \varepsilon) = \{f \in \mathcal{C}(T, \mathbb{F}); \sup_{x \in K} |f(x) - g(x)| < \varepsilon\} = g + B(0, K, \varepsilon),$$

gdzie $g \in \mathcal{C}(T, \mathbb{F})$, $K \in \mathfrak{K}_T$ i $\varepsilon > 0$. Topologia $\tau_{\mathcal{C}}$ jest zgodna ze strukturą przestrzeni liniowej nad $\mathbb{F}$.

Dowolny funkcjonal liniowy $\lambda: \mathcal{C}(T, \mathbb{F}) \rightarrow \mathbb{F}$ jest ciągły w topologii niemal jednostajnej zbieżności $\tau_{\mathcal{C}}$ wtedy i tylko wtedy, gdy istnieją zbiór zwarty $K \subset T$ oraz stała $C \geq 0$, takie że

$$|\lambda(f)| \leq C \sup_{x \in K} |f(x)|, \quad \text{dla } f \in \mathcal{C}(T, \mathbb{F}). \quad (3.28)$$

(ii) W topologii τ_c dowolny podzbiór $U \subset \mathcal{C}_c(T, \mathbb{F})$ jest otwarty wtedy i tylko wtedy, gdy każdy przekrój $U \cap \mathcal{C}_K(T, \mathbb{F})$ jest otwartym podzbiorem w $\mathcal{C}_K(T, \mathbb{F})$, dla wszystkich zbiorów zwartych $K \subset T$. Każda translacja $\mathcal{C}_c(T, \mathbb{F}) \ni f \mapsto f + g \in \mathcal{C}_c(T, \mathbb{F})$ o funkcję g o nośniku zwartym jest odwzorowaniem ciągłym.

Funkcjonał (norma) $\sup |\cdot|$ jest funkcją ciągłą. Dowolny funkcjonal liniowy $\lambda: \mathcal{C}_c(T, \mathbb{F}) \rightarrow \mathbb{F}$ jest ciągły w topologii τ_c wtedy i tylko wtedy, gdy dla każdego zbioru zwartego $K \subset T$ istnieje stała $C_K \geq 0$, taka że

$$|\lambda(f)| \leq C_K \sup_{x \in K} |f(x)|, \quad \text{jeśli } \text{supp}(f) \subset K. \quad (3.29)$$

Obie topologie $\tau_{\mathcal{C}}$ i τ_c są klasy T_2 .

Dowód. (i) Jawny opis zbiorów otwartych w τ_C wynika bezpośrednio z definicji i zapewnia, że $\mathcal{C}(T, \mathbb{F})$ jest przestrzenią liniowo topologiczną. W szczególności każda relatywna kula $B(g, K, \varepsilon) = \text{pr}_K^{-1}B(g|_K, \varepsilon)$ jest otwarta jako przeciwobraz kuli otwartej w $\mathcal{C}(K, \mathbb{F})$. Aby sprawdzić własność bazy, rozważmy dowolne dwie relatywne kule o niepustym przekroju $B(g_i, K_i, \varepsilon_i)$, $i = 1, 2$, oraz dowolną funkcję

$$g \in B(g_1, K_1, \varepsilon_1) \cap B(g_2, K_2, \varepsilon_2),$$

co oznacza nierówności $\sup_{K_i} |g - g_i| < \varepsilon_i$, dla $i = 1, 2$. Przyjmując

$$\varepsilon = \min\{\varepsilon_i - \sup_{x \in K_i} |g(x) - g_i(x)|; i = 1, 2\} > 0,$$

sprawdzamy inkluzję $B(g, K_1 \cup K_2, \varepsilon) \subset B(g_1, K_1, \varepsilon_1) \cap B(g_2, K_2, \varepsilon_2)$. Istotnie, dla każdej funkcji $f \in B(g, K_1 \cup K_2, \varepsilon)$ z własności

$$|f(x) - g(x)| < \varepsilon_i - \sup_{K_i} |g - g_i|, \quad \text{dla } x \in K_i$$

wynikają nierówności

$$\sup_{x \in K_i} |f(x) - g_i(x)| \leq \sup_{K_i} |f - g| + \sup_{K_i} |g - g_i| < \varepsilon_i,$$

a więc $f \in B(g_i, K_i, \varepsilon_i)$, dla $i = 1, 2$.

Dowolny funkcjonal liniowy λ jest ciągły w topologii τ_C wtedy i tylko wtedy, gdy przeciwobraz $\lambda^{-1}(-1, 1)$ zawiera zbiór otwarty postaci $B(0, K, \varepsilon)$. Oznacza to implikację

$$\sup_{x \in K} |f(x)| < \varepsilon \implies |\lambda(f)| < 1 \quad \text{dla } f \in \mathcal{C}(T, \mathbb{F})$$

równoważną własności (3.28).

(ii) Zauważmy, że podana charakteryzacja rodziny zbiorów otwartych w topologii τ_C jest ograniczeniem z góry – jakikolwiek zbiór „otwarty” bez tej własności wyklucza ciągłość przynajmniej jednego z włożeń ι_K . Pozostaje w tym punkcie sprawdzić (co pozostawiamy czytelnikowi), że wyróżniona rodzina $\mathcal{U}$ jest topologią. Ponieważ dowolna funkcja h określona na $\mathcal{C}_c(T, \mathbb{F})$ jest ciągła względem $\mathcal{U}$ wtedy i tylko wtedy, gdy ciągłe są wszystkie obcięcia $h \circ \iota_K$, stwierdzamy ciągłość ewaluacji $e_x: f \mapsto f(x)$, $x \in T$, a stąd wynika domkniętość podprzestrzeni

$$\mathcal{C}_K(T, \mathbb{F}) = \bigcap_{x \notin K} e_x^{-1}(0) \subset \mathcal{C}_c(T, \mathbb{F}), \quad \text{dla } K \in \mathfrak{K}_T.$$

Rodzina $\mathcal{U}$ jest zatem największą topologią, w której ciągłe są wszystkie ewaluacje. Aby udowodnić, że ι_K jest homeomorfizmem na obraz, ustalmy

teraz dowolny zbiór zwarty $K \in \mathfrak{K}_T$ i podzbiór otwarty $U_o \subset \mathcal{C}_K(T, \mathbb{F})$, i przyjmijmy

$$U = U_o \cup \{f \in \mathcal{C}_c(T, \mathbb{F}); \exists_{x \notin K} f(x) \neq 0\}.$$

Zauważmy, że $\mathcal{C}_K(T, \mathbb{F}) \cap U = U_o$, a dla dowolnego innego zbioru zwartego $L \in \mathfrak{K}_T$ przekrój

$$U_L = (\mathcal{C}_L(T, \mathbb{F}) \cap U_o) \cup \bigcup_{x \in L \setminus K} \{f \in \mathcal{C}_L(T, \mathbb{F}); f(x) \neq 0\}$$

jest otwarty w $\mathcal{C}_L(T, \mathbb{F})$. Istotnie, drugi składnik jest zbiorem otwartym, a dla dowolnego $g \in \mathcal{C}_L(T, \mathbb{F}) \cap U_o$ i liczby $\varepsilon > 0$, takiej że

$$\{f \in \mathcal{C}_K(T, \mathbb{F}); \sup |f - g| < \varepsilon\} \subset U_o,$$

kula otwarta $\{f \in \mathcal{C}_L(T, \mathbb{F}); \sup |f - g| < \varepsilon\}$ zawiera się w U_L . Wykazaliśmy tym samym, że $\tau_c = \mathcal{U}$ jest jednoznacznie wyróżnioną w definicji 3.2.2 topologią zbieżności jednostajnej w $\mathcal{C}_c(T, \mathbb{F})$.

Dla dowolnej funkcji $g \in \mathcal{C}_c(T, \mathbb{F})$ o nośniku L ciągle translacje

$$t_g: \mathcal{C}_K(T, \mathbb{F}) \ni f \longmapsto f + g \in \mathcal{C}_{K \cup L}(T, \mathbb{F}), \quad K \in \mathfrak{K}_T,$$

wyznaczają w diagramie

$$\begin{array}{ccc} \mathcal{C}_K(T, \mathbb{F}) & \xrightarrow{t_g} & \mathcal{C}_{K \cup L}(T, \mathbb{F}) \\ \downarrow \iota_K & \searrow & \downarrow \iota_{K \cup L} \\ \mathcal{C}_c(T, \mathbb{F}) & \xrightarrow{t_g} & \mathcal{C}_c(T, \mathbb{F}) \end{array}$$

ciągle odwzorowania diagonalne, a to oznacza ciągłość translacji w $\mathcal{C}_c(T, \mathbb{F})$.

Dla zakończenia dowodu wystarczy zauważyć, że warunek (3.29) oznacza dokładnie ciągłość względem normy w (każdej) przestrzeni $\mathcal{C}_K(T, \mathbb{F})$. Własność T_2 jest prostą konsekwencją ciągłości normy sup. $\square$

Uwaga 3.2.1. Przestrzeń topologiczna $(\mathcal{C}_c(T, \mathbb{F}), \tau_c)$ jest *granicą prostą* tzw. *systemu injektywnego*

$$\begin{array}{ccc} \mathcal{C}_K(T, \mathbb{F}) & & \text{dla } K, L \in \mathfrak{K}_T, \\ & \searrow \subset & \\ \mathcal{C}_L(T, \mathbb{F}) & \xrightarrow{\subset} & \mathcal{C}_{K \cup L}(T, \mathbb{F}) \xrightarrow{\iota_{K \cup L}} \mathcal{C}_c(T, \mathbb{F}) \end{array}$$

co oznacza zgodną rodzinę odwzorowań (tu: homeomorficznych włożeń) $\mathcal{C}_K \hookrightarrow \mathcal{C}_{K'}$, dla par indeksów związanych relacją częściowego porządku $K \subset K'$.

Analogicznie (lub – dla odmiany) przestrzeń topologiczna $(\mathcal{C}(T, \mathbb{F}), \tau_{\mathcal{C}})$ jest *granicą odwrotną* tzw. *systemu projektywnego*

$$\begin{array}{ccc} \mathcal{C}(T, \mathbb{F}) & \xleftarrow{\text{pr}_{K \cup L}} & \mathcal{C}(K \cup L, \mathbb{F}) \xrightarrow{\text{pr}_K} \mathcal{C}(K, \mathbb{F}) \quad \text{dla } K, L \in \mathfrak{K}_T. \\ & \searrow \text{pr}_L & \\ & & \mathcal{C}(L, \mathbb{F}) \end{array}$$

Granica prosta stanowi uogólnienie pojęcia sumy rozłącznej przestrzeni, granica odwrotna uogólnia pojęcie produktu. Proponujemy czytelnikowi – jako ćwiczenie – szczegółowe opracowanie pojęć ww. granic (prostej i odwrotnej) w języku teorii kategorii (por. [2, rozdz. 2.1]).

Przestrzenie $\mathcal{C}(T, \mathbb{F})$ i $\mathcal{C}_c(T, \mathbb{F})$ nie są na ogół metryczne.

Ćwiczenie 3.2.2. Jeśli przestrzeń lokalnie zwarta T spełnia II ap. i – zgodnie z [2, tw. 1.6.21] – jest metryzowalna, to T jest sumą wstępującego ciągu zbiorów zwartych postaci $K_n = \overline{U_n}$, $n \in \mathbb{N}$, przy czym każdy zwarty podzbiór $K \in \mathfrak{K}_T$ zawiera się w dostatecznie dużym zbiorze K_n . Sprawdzić, że metryka

$$\rho(f, g) = \sum_{n \in \mathbb{N}} 2^{-n} \min \left(\sup_{x \in K_n} |f - g|(x), 1 \right) \quad \text{dla } f, g \in \mathcal{C}(T, \mathbb{F})$$

wyznacza topologię $\tau_{\mathcal{C}}$ niemal jednostajnej zbieżności. Wykazać ośrodkowość przestrzeni $\mathcal{C}(T, \mathbb{F})$. Podać jawny opis ośrodka w $\mathcal{C}(\mathbb{R})$.

W opisanej sytuacji przestrzeń $(\mathcal{C}(T, \mathbb{F}), \tau_{\mathcal{C}})$ jest tzw. *przestrzenią Frecheta*. Odpowiednia granica prosta $(\mathcal{C}_c(T, \mathbb{F}), \tau_c)$ również daje się opisać poprzez wstępujący ciąg podprzestrzeni i jest tzw. *LF-przestrzenią* (por. [19]).

3.2.2 Mierzalność oparta na ciągłości

Zgodnie z udowodnionym wyżej twierdzeniem 3.2.3(ii) przyjmujemy

Definicja 3.2.4. Miarą Radona na przestrzeni lokalnie zwartej T – zespoloną lub rzeczywistą, zależnie od ciała $\mathbb{F} \in \{\mathbb{C}, \mathbb{R}\}$ – nazywamy funkcjonal liniowy $\lambda: \mathcal{C}_c(T, \mathbb{F}) \rightarrow \mathbb{F}$ spełniający następujący warunek ciągłości

$$\forall K \in \mathfrak{K} \exists C_K > 0 |\lambda(f)| \leq C_K \sup_{x \in K} |f(x)|, \quad \text{jeśli } \text{supp}(f) \subset K. \quad (3.30)$$

Rzeczywistą miarę Radona λ nazywamy dodatnią, jeśli spełnia warunek

$$\lambda(f) \geq 0 \quad \text{dla } f \geq 0. \quad (3.31)$$

Miarami Radona związanymi z ciągłymi funkcjonalami liniowymi na przestrzeni $\mathcal{C}(T, \mathbb{F})$ zajmiemy się w dalszej kolejności (tw. 3.3.12).

Dla miar rzeczywistych posługujemy się skróconą notacją $\mathcal{C}(T) = \mathcal{C}(T, \mathbb{R})$ i $\mathcal{C}_c(T) = \mathcal{C}_c(T, \mathbb{R})$, zgodną z konwencją stosowaną w poprzednich rozdziałach. Związki pomiędzy trzema wyróżnionymi rodzajami miar Radona stanowią treść kilku kolejnych ćwiczeń i wniosku 3.2.12.

Ćwiczenie 3.2.3. Wykazać, że każdy funkcjonal liniowy $\lambda: \mathcal{C}_c(T) \rightarrow \mathbb{R}$ spełniający warunek (3.31) jest ciągły w sensie (3.30), a więc jest dodatnią miarą Radona. Skorzystać z nierówności $|f| \leq \|f\|_\varphi$ – dla dowolnej nieujemnej funkcji $\varphi \in \mathcal{C}_c(T)$ równą 1 na $\text{supp}(f)$.

Klasyczne twierdzenie Diniego o jednostajnej zbieżności monotonicznego ciągu zbieżnego na zbiorze zwartym (por. [2, tw. 1.6.7]) wiąże dodatnie miary Radona z rozwijaną w podrozdziale 3.1 abstrakcyjną teorią całki.

Lemat 3.2.5. *Każda dodatnia miara Radona $\lambda: \mathcal{C}_c(T) \rightarrow \mathbb{R}$ w lokalnie zwartej przestrzeni T jest miarą abstrakcyjną w T . Przestrzeń $\mathcal{C}_c(T)$ jest kratą Stone’a.*

Dowód. Własność kraty i warunek Stone’a (3.20) są oczywiste. Sprawdzenia wymaga jedynie druga z własności (3.1) oznaczająca monotoniczną ciągłość λ . Wychodząc zatem od dowolnego ciągu niemalejącego $f_n \nearrow f$ zbieżnego w $\mathcal{C}_c(T)$ i zastępując go ciągiem funkcji nieujemnych $f_n - f_1 \nearrow f - f_1$ o zwartych nośnikach zawartych w $K = \text{supp}(f - f_1)$, stwierdzamy, na mocy twierdzenia Diniego, jednostajną zbieżność $f - f_n \rightrightarrows 0$. Ponieważ zachowany jest także warunek $\text{supp}(f - f_n) \subset K$ dla $n \in \mathbb{N}$, własność (3.30) miary Radona implikuje równość $\lim_n \lambda(f - f_n) = 0$. $\square$

Możliwość skorzystania z twierdzenia Stone’a (tw. 3.1.21) kieruje naszą uwagę na σ –pierścień $\mathcal{C}_c(T)^\sigma$. Przyjmujemy

Definicja 3.2.6. *W przestrzeni lokalnie zwartej T rodzinę zbiorów*

$$\mathfrak{B}_{\text{aire}}^\sigma(T) = \mathcal{C}_c(T)^\sigma = \sigma\{f^{-1}[1, \infty); f \in \mathcal{C}_c(T)\} \quad (3.32)$$

nazywamy σ –pierścieniem Baire’a. Miarą Baire’a w przestrzeni T nazywać będziemy każdą miarę na σ –pierścieniu $\varsigma: \mathfrak{B}_{\text{aire}}^\sigma(T) \rightarrow [0, \infty]$ przyjmującą skończone wartości na zbiorach zwartych.

Z ogólnych własności całki Lebesgue’a wynika

Wniosek 3.2.7. *Zachodzi inkluzja $\mathcal{C}_c(T) \subset \mathcal{I}(T, \varsigma)$, a więc funkcje ciągłe o zwartym nośniku są całkowne względem każdej miary Baire’a.* $\square$

Stwierdzenie 3.2.8 (por. [2, lem. 1.6.28]). *W przestrzeni lokalnie zwartej T , następujące własności zbioru zwartego K są równoważne:*

- (i) $K = \{f \geq 1\}$ dla pewnej funkcji $f \in \mathcal{C}_c(T)$;
- (ii) K jest zbiorem typu G_δ . □

Zbiory Baire’a stanowią najmniejszy σ –pierścień, dla którego wszystkie funkcje rzeczywiste ciągle o zwartych nośnikach na T są $\mathfrak{B}_{\text{ave}}^\sigma(T)/\mathfrak{B}_\mathbb{R}$ –mieralne. Równoważność (i) $\Leftrightarrow$ (ii) oznacza dodatkowo, że σ –pierścień $\mathfrak{B}_{\text{ave}}^\sigma(T)$ jest generowany przez rodzinę $\mathfrak{K}_T^G$ zbiorów zwartych typu G_δ .

Jako wniosek z twierdzenia 3.1.21 otrzymujemy

Twierdzenie 3.2.9 (O dualności dla miar Radona). *Dla dowolnej dodatniej miary Radona $\lambda: \mathcal{C}_c(T) \rightarrow \mathbb{R}$ w przestrzeni lokalnie zwartej T istnieje dokładnie jedna miara Baire’a $\varsigma_\lambda: \mathfrak{B}_{\text{ave}}^\sigma(T) \rightarrow [0, \infty]$, dla której $\lambda = \int d\varsigma_\lambda$.*

Odwrotnie – dla każdej miary Baire’a ς w T funkcjonal

$$\lambda_\varsigma: \mathcal{C}_c(T) \ni f \mapsto \int f d\varsigma \in \mathbb{R}$$

jest dodatnią miarą Radona.

Dowód. Z twierdzenia Stone’a wiemy, że miara Daniella–Stone’a ς_λ jest jednoznaczным uzupełnieniem miary określonej na $\mathfrak{B}_{\text{ave}}^\sigma(T)$ i spełnia postulowaną równość $\int d\varsigma_\lambda = \lambda$. Aby wykazać, że oznaczane tym samym symbolem ograniczenie $\varsigma_\lambda|_{\mathfrak{B}_{\text{ave}}^\sigma(T)}$ jest miarą Baire’a, rozważmy dowolny zbiór zwarty $K \in \mathfrak{B}_{\text{ave}}^\sigma(T)$. Z twierdzenia 3.2.1 wynika istnienie funkcji $f \in \mathcal{C}_c(T)$ takiej, że $1_K \leq f$, a więc $\varsigma_\lambda(K) \leq \lambda(f) < \infty$. Druga część tezy jest oczywista. □

W kolejnych podrozdziałach twierdzenie 3.2.9 rozszerzamy do *twierdzenia Riesz’a o reprezentacji* – tw. 3.3.3.

Przykład 3.2.4. Opisana w przykładzie 3.1.1 całka Riemanna wyznacza nieujemny funkcjonal liniowy na kracie $\mathcal{C}([a, b])$, spełniający warunek

$$\left| \int_a^b f(x) dx \right| \leq (b - a) \sup_{x \in [a, b]} |f(x)|$$

dla $f \in \mathcal{C}([a, b])$. Konstrukcja Daniella prowadzi w tym przypadku do całki Lebesgue’a na prostej i w konsekwencji do miary Lebesgue’a.

3.2.3 Struktura rzeczywistych i zespolonych całek Radona

Miarę zespoloną możemy uważać za rzeczywistą, jeśli na funkcjach rzeczywistych ma wartości rzeczywiste. Dla każdej miary zespolonej λ funkcjonal $\bar{\lambda}$ określony wzorem $\bar{\lambda}(f) := \overline{\lambda(f)}$ jest także zespoloną miarą Radona.

Ćwiczenie 3.2.5. Miara λ jest rzeczywista wtedy i tylko wtedy, gdy $\bar{\lambda} = \lambda$. Dla dowolnej zespolonej miary Radona λ na T , funkcjonały wyznaczone przez części rzeczywistą i urojoną, $\lambda_1(f) = \Re \lambda(f)$ i $\lambda_2(f) = \Im \lambda(f)$ dla $f \in \mathcal{C}_c(T)$, są rzeczywistymi miarami Radona. Z drugiej strony, każda rzeczywista miara Radona $\lambda: \mathcal{C}_c(T) \rightarrow \mathbb{R}$ rozszerza się jednoznacznie do miary zespolonej $\lambda(f_1 + if_2) := \lambda(f_1) + i\lambda(f_2)$, dla $f_1, f_2 \in \mathcal{C}_c(T)$.

Lemat 3.2.10. Dla dowolnej zespolonej miary Radona λ w przestrzeni T istnieje rzeczywista miara Radona $|\lambda|: \mathcal{C}_c(T) \rightarrow \mathbb{R}$ spełniająca warunek

$$|\lambda|(f) = \sup\{|\lambda(g)|; g \in \mathcal{C}_c(T, \mathbb{C}), |g| \leq f\} \quad \text{dla } f \geq 0. \quad (3.33)$$

Definicja 3.2.11. Jedyną dodatnią miarę Radona $|\lambda|$ spełniającą warunek (3.33) nazywamy wartością bezwzględną lub modulem miary λ .

Ćwiczenie 3.2.6. Sprawdzić, że miara $|\lambda|$ jest najmniejszą miarą dodatnią (zdefiniować naturalną relację częściowego porządku!), taką że

$$|\lambda(f)| \leq |\lambda|(|f|) \quad \text{dla } f \in \mathcal{C}_c(T, \mathbb{C}). \quad (3.34)$$

Dowód lematu 3.2.10. Z warunku ciągłości (3.30) miary Radona λ wynika ograniczenie $|\lambda|(f) \leq C_K \sup_{x \in K} |f(x)|$ dla $K = \text{supp } f$, $f \in \mathcal{C}_c(T)$, co oznacza poprawność definicji i ewentualną ciągłość $|\lambda|$ – po sprawdzeniu liniowości.

Dla dowodu addytywności, $|\lambda|(f_1 + f_2) = |\lambda|(f_1) + |\lambda|(f_2)$ dla $f_1, f_2 \geq 0$, weźmy dowolną liczbę $\varepsilon > 0$ i funkcję $g \in \mathcal{C}_c(T, \mathbb{C})$ taką, że $|g| \leq f_1 + f_2$ oraz $|\lambda|(f_1 + f_2) \leq |\lambda(g)| + \varepsilon$. Aby przedstawić g jako sumę $g = g_1 + g_2$, w której $|g_i| \leq f_i$ dla $i = 1, 2$, wystarczy przyjąć

$$g_i(x) = \begin{cases} g(x) \frac{f_i(x)}{f_1(x) + f_2(x)}, & \text{dla } f_1(x) + f_2(x) > 0 \\ 0, & \text{jeśli } f_1(x) = f_2(x) = 0, \end{cases}$$

dla $i = 1, 2$. Ciągłość każdej z funkcji w punktach $x \in T$, w których suma $f_1(x) + f_2(x)$ jest równa 0, a więc także $g(x) = 0$, wynika z nierówności $|g_i| \leq |g|$. Wobec dowolności ε , nierówności

$$|\lambda(g)| \leq |\lambda(g_1)| + |\lambda(g_2)| \leq |\lambda|(f_1) + |\lambda|(f_2)$$

prowadzą do $|\lambda|(f_1 + f_2) \leq |\lambda|(f_1) + |\lambda|(f_2)$. Nierówność przeciwna wynika wprost z definicji (3.33).

Pozostawiamy czytelnikowi sprawdzenie, że udowodniona wyżej addytywność $|\lambda|$ i oczywista dodatnia jednorodność – na funkcjach nieujemnych – implikują liniowość rozszerzenia $|\lambda|(f) := |\lambda|(f^+) - |\lambda|(f^-)$, dla dowolnych $f \in \mathcal{C}_c(T)$. $\square$

Wniosek 3.2.12. Dla dowolnej zespolonej miary Radona λ miary

$$\lambda_{\Re} = \frac{1}{2}(\lambda + \bar{\lambda}) \quad i \quad \lambda_{\Im} = \frac{1}{2i}(\lambda - \bar{\lambda})$$

są rzeczywiste, przy czym $\lambda = \lambda_{\Re} + i\lambda_{\Im}$.

Każda rzeczywista miara Radona λ jest różnicą $\lambda = \lambda^+ - \lambda^-$ miar dodatnich $\lambda^+ = \frac{1}{2}(\lambda + |\lambda|)$ i $\lambda^- = \frac{1}{2}(|\lambda| - \lambda)$. $\square$

Ćwiczenie 3.2.7. Dla dowolnej miary rzeczywistej λ i funkcji $f \in \mathcal{C}_c(T, \mathbb{C})$ zachodzi równość $|\lambda(f)| = \lambda(\zeta f)$, gdzie $\zeta \in \mathbb{C}$ jest odpowiednio dobraną liczbą o module równym 1. Wywnioskować stąd, że we wzorze (3.33) można się ograniczyć do funkcji rzeczywistych $g \in \mathcal{C}_c(T)$ takich, że $|g| \leq f$. W szczególności wykazać, że każda miara dodatnia jest równa swej wartości bezwzględnej.

Za całkę Radona względem dodatniej miary Radona λ w przestrzeni lokalnie zwartej T przyjmujemy wstępnie stosowną całkę Daniella–Stone’a, będącą także całką Lebesgue’a

$$\int d\lambda: \mathcal{I}(T, \lambda) \rightarrow \mathbb{R}.$$

Czeka nas jeszcze proces jednoznacznego rozszerzania miary Radona – w dwu kierunkach – do całki Radona względem dowolnej zespolonej miary Radona, oraz do stosownej całki i miary borelowskiej.

Definicja 3.2.13. Dla dowolnej rzeczywistej miary Radona $\lambda: \mathcal{C}_c(T) \rightarrow \mathbb{R}$, całką Radona nazywamy każdy z funkcyjałów $\mathbb{F}$ –liniowych

$$\int d\lambda: \mathbb{F} \otimes \mathcal{I}(T, |\lambda|) \rightarrow \mathbb{F}, \quad \mathbb{F} = \mathbb{R}, \mathbb{C},$$

określonych za pomocą wzorów

$$\begin{aligned} \int f d\lambda &= \int f d\lambda^+ - \int f d\lambda^- \quad \text{dla } f \in \mathcal{I}(T, |\lambda|) \\ \int (f_1 + if_2) d\lambda &= \int f_1 d\lambda + i \int f_2 d\lambda \quad \text{dla } f_1, f_2 \in \mathcal{I}(T, |\lambda|). \end{aligned} \quad (3.35)$$

Dla dowolnej zespolonej miary Radona $\lambda: \mathcal{C}_c(T, \mathbb{C}) \rightarrow \mathbb{C}$, całką Radona względem λ nazywamy funkcyjał $\int d\lambda: \mathbb{C} \otimes \mathcal{I}(T, |\lambda|) \rightarrow \mathbb{C}$ taki, że

$$\int f d\lambda = \int f d\lambda_{\Re} + i \int f d\lambda_{\Im} \quad \text{dla } f \in \mathbb{C} \otimes \mathcal{I}(T, |\lambda|). \quad (3.36)$$

Całkowalność funkcji podcałkowych występujących w (3.35)–(3.36) wymaga dowodu. Stąd

Stwierdzenie 3.2.14. *Zachodzą równości*

$$\mathcal{I}(T, \lambda^+) \cap \mathcal{I}(T, \lambda^-) = \mathcal{I}(T, |\lambda|),$$

dla dowolnej rzeczywistej miary Radona $\lambda: \mathcal{C}_c(T) \rightarrow \mathbb{R}$, oraz

$$\mathcal{I}(T, |\lambda_{\Re}|) \cap \mathcal{I}(T, |\lambda_{\Im}|) = \mathcal{I}(T, |\lambda|),$$

dla miary zespolonej $\lambda: \mathcal{C}_c(T, \mathbb{C}) \rightarrow \mathbb{C}$. Dla dowolnej funkcji $f \in \mathcal{I}(T, |\lambda|)$ prawdziwa jest nierówność

$$|\int f d\lambda| \leq \int |f| d|\lambda|.$$

Dowód. Ćwiczenie. □

Ćwiczenie 3.2.8. Sprawdzić, że znana z analizy zespolonej całka krzywoliniowa jest zespoloną miarą Radona. Opisać przestrzeń funkcji całkowlanych.

W przypadku miar i funkcji o wartościach zespolonych rolę podstawowego twierdzenia dotyczącego *przechodzenia do granicy pod znakiem całki* przejmuje

Twierdzenie 3.2.15 (Lebesgue’a o zbieżności jednostajnej). *Jeśli funkcje stałe są całkowlane względem $|\lambda|$, czyli – jeśli przestrzeń T ma skończoną miarę, to dla dowolnego jednostajnie zbieżnego ciągu funkcji $|\lambda|$ -całkowalnych $(f_n)_{n \in \mathbb{N}}$ granica $f = \lim_{n \rightarrow \infty} f_n$ jest funkcją całkowlaną i zachodzi równość*

$$\int f d\lambda = \lim_{n \rightarrow \infty} \int f_n d\lambda.$$

Dowód. Z jednostajnej zbieżności ciągu wynika ograniczenie $|f_n - f_{n_0}| \leq 1$ dla pewnego $n_0 \in \mathbb{N}$ i wszystkich $n > n_0$. Ciąg ograniczony $|f_n| \leq 1 + |f_{n_0}|$, $n > n_0$, spełnia założenia twierdzenia Lebesgue’a o zmajoryzowanej zbieżności (wniosek 3.1.19). □

3.3 Regularne miary i całki

3.3.1 Miary wyznaczone poprzez zbiory zwarte

Możliwość korzystania z całki Radona, w postaci sformułowanej w poprzednim podrozdziale, czy też – równoważnie – z całki względem miary Baire’a, ograniczona jest wymaganiem mierzalności badanych funkcji względem σ -ciała zbiorów Baire’a. W ogólnej sytuacji, gdy $\mathfrak{B}_{\text{aire}}(T) \subsetneq \mathfrak{B}_T$, oznacza to np. brak stosowalności całki do dowolnych funkcji ciągłych, czy –

ogólniej – borelowskich. Klasycznym sposobem na usunięcie tego ograniczenia jest jednoznaczne rozszerzenie miary Radona do *regularnej* miary borelowskiej (twierdzenie 3.3.4).

Związki miar borelowskich z topologią przestrzeni prowadzą do wyróżnienia miar mniej lub bardziej regularnych, przy czym należy zaznaczyć, że w literaturze przedmiotu istnieje pewna niejednoznaczność terminologii. Przypominamy także, że zgodnie z twierdzeniem 1.1.47 zbliżona do zdefiniowanej poniżej *regularność* jest naturalną cechą miar w przestrzeniach metrycznych.

Definicja 3.3.1. *Miara borelowska μ w przestrzeni Hausdorffa T jest regularna, jeśli spełnia następujące warunki.*

(i) *Dla każdego zbioru borelowskiego $A \subset T$,*

$$\mu(A) = \inf \{ \mu(U); A \subset U, U \text{ jest zbiorem otwartym} \}.$$

(ii) *Zbiory zwarte mają miarę skończoną, natomiast dla każdego zbioru otwartego $U \subset T$ zachodzi równość*

$$\mu(U) = \sup \{ \mu(K); K \subset U, \text{zbiór } K \text{ jest zwarty} \}.$$

Własność (i) charakteryzuje miary regularne z góry (upper-regular).

Stwierdzenie 3.3.2. *Składowa σ –skończona dowolnej miary regularnej μ opisana jest poprzez tożsamość*

$$\mu(A) = \sup \{ \mu(K); K \subset A, \text{zbiór } K \text{ jest zwarty} \}, \quad (3.37)$$

dla dowolnego σ –skończonego zbioru borelowskiego A , a zatem jest bezpośrednio wyznaczona poprzez wartości na zbiorach zwartych.

Dowód. Jeśli $\mu(A) < \infty$, to istnieje zbiór otwarty $U \supset A$ o skończonej mierze. Dla dowolnego $\varepsilon > 0$, wybierzmy podzbiór zwarty $K \subset U$ taki, że $\mu(U \setminus K) < \varepsilon$ i zawarty w U zbiór otwarty $V \supset U \setminus A$, taki że $\mu(V \setminus (U \setminus A)) < \varepsilon$. Otrzymujemy w ten sposób konfigurację, w której zbiór zwarty $K \setminus V$ jest zawarty w A , natomiast miara różnicy

$$A \setminus (K \setminus V) \subset (U \setminus K) \cup (A \cap V)$$

jest ograniczona przez 2ε .

W przypadku $\mu(A) = \infty$ wystarczy rozważyć wstępujący ciąg zbiorów borelowskich $A_n \nearrow A$, $n \in \mathbb{N}$, o mierze skończonej, i każdy z nich „przybliżyć” zbiorem zwartym $K_n \subset A_n$ takim, by $\mu(K_n) \geq \mu(A_n) - 1$. $\square$

Ćwiczenie 3.3.1. Sprawdzić, że uzupełnienie regularnej miary borelowskiej spełnia oba warunki regularności sformułowane w definicji 3.3.1.

Twierdzenie 3.3.3 (Riesza o reprezentacji). *Niech T będzie dowolną przestrzenią lokalnie zwartą. Regularne miary borelowskie i dodatnie miary Radona na T pozostają w naturalnej wzajemnie jednoznacznej odpowiedniości przedstawionej poniżej.*

(i) *Dla każdej regularnej miary borelowskiej μ w T funkcje ciągłe o zwartych nośnikach są całkowalne, a funkcjonal $\lambda_\mu: C_c(T) \ni f \mapsto \int f d\mu \in \mathbb{R}$ jest dodatnią miarą Radona.*

(ii) *Dla każdej dodatniej miary Radona $\lambda: C_c(T) \rightarrow \mathbb{R}$ istnieje dokładnie jedna regularna miara borelowska ϱ_λ w T taka, że zachodzi równość*

$$\int f d\varrho_\lambda = \lambda(f) \quad \text{dla } f \in C_c(T).$$

Teza jest rozwinięciem twierdzenia 3.2.9.

Szkic dowodu. Z twierdzenia Stone’a 3.1.21 wynika istnienie i jednoznaczność zgodnej z λ miary Baire’a ς_λ , takiej że

$$\lambda(f) = \int f d\varsigma_\lambda \quad \text{dla } f \in C_c(T).$$

W kolejnym podrozdziale (twierdzenie 3.3.4) wykażemy, że miara ς_λ rozszerza się jednoznacznie do regularnej miary borelowskiej $\rho_\lambda = \tilde{\varsigma}_\lambda$ w przestrzeni T , co prowadzi do rozszerzenia samej całki i przestrzeni funkcji całkowalnych. Wniosek 1.2.8 zapewnia wymaganą równość całek. $\square$

3.3.2 Charakteryzacje miar regularnych

Pobieżna analiza własności miary borelowskiej μ w przestrzeni lokalnie zwartej T , składających się na ewentualną regularność miary, prowadzi do dwu obserwacji:

- Wartości miary regularnej są wyznaczone przez ograniczenie μ do pewnej rodziny $\mathfrak{K}$ zbiorów zwartych – przynajmniej takiej, że dla każdej pary (K, U) złożonej ze zbioru zwartego K i zawierającego go zbioru otwartego U istnieje zbiór zwarty $F \in \mathfrak{K}$ taki, że $K \subset F \subset U$.
- Równość $\mu(F) = \inf\{\mu(U); F \subset U\}$ oznacza, że zbiory zwarte z rodziny $\mathfrak{K}$ pozostają w szczególnej relacji ze swoimi otwartymi otoczeniami.

Twierdzenie 3.3.4 (O miarach regularnych). *W dowolnej przestrzeni lokalnie zwartej (T, τ) ograniczenie do $\mathfrak{B}_{\text{ave}}^\sigma(T) \subset \mathfrak{B}_T$ wyznacza naturalną, wzajemnie jednoznaczność odpowiedniość pomiędzy:*

- (i) regularnymi miarami borelowskimi $\mathfrak{B}_T \rightarrow [0, \infty]$ i
- (ii) miarami Baire'a $\mathfrak{B}_{\text{ave}}^\sigma(T) \rightarrow [0, \infty]$.

Ponadto, dla dowolnych miar Baire'a μ i ν w przestrzeniach S i T zachodzi równość

$$(\mu \otimes \nu)^\sim(A) = \tilde{\mu} \otimes \tilde{\nu}(A) \quad \text{dla } A \in (\mathfrak{B}_S \otimes \mathfrak{B}_T)_{\tilde{\mu} \otimes \tilde{\nu}}^\sigma, \quad (3.38)$$

gdzie $\tilde{\mu}$ i $\tilde{\nu}$ oraz $(\mu \otimes \nu)^\sim$ oznaczają jedyne regularne rozszerzenia miar μ , ν i produktu $\mu \otimes \nu: \mathfrak{B}_{\text{ave}}^\sigma(S \times T) \rightarrow [0, \infty]$.

Korzystając z twierdzenia 3.2.1, wykażemy podstawowy

Lemat 3.3.5. Dla dowolnej miary Baire'a $\nu: \mathfrak{B}_{\text{ave}}^\sigma(T) \rightarrow [0, \infty]$ w przestrzeni lokalnie zwartej (T, τ) , funkcja $\tilde{\nu}$ określona wzorem

$$\tilde{\nu}(U) = \sup \{ \nu(K); K \subset U, K \in \mathfrak{K}^G \} \quad \text{dla } U \in \tau, \quad (3.39)$$

jest subaddytywna i spełnia warunek

$$\nu(K) + \tilde{\nu}(U \setminus K) \leq \tilde{\nu}(U) \quad (3.40)$$

dla dowolnej pary zbiorów $U \in \tau$ i $K \in \mathfrak{K}^G$ takich, że $K \subset U$.

Przypominamy, że $\mathfrak{K}^G = \mathfrak{K}_T^G$ oznacza rodzinę zbiorów zwartych typu G_δ .

Dowód. Dla dowolnego ciągu zbiorów otwartych $(V_i)_{i \in \mathbb{N}}$ sumującego się do $V \in \tau$ niech K będzie dowolnym zbiorem zwartym z $\mathfrak{K}^G$ zawartym w V . Ustalając ciągły rozkład jedynek $\{f_i\}_{i \leq m}$ dla K , wpisany w skończone pokrycie $K \subset \bigcup_{i \leq m} V_i$, znajdujemy zbiory zwarte $K_i = \{f_i \geq \frac{1}{m}\} \subset V_i$, $i \leq m$, pokrywające K i należące do $\mathfrak{K}^G$. Z subaddytywności miary ν otrzymujemy zatem

$$\nu(K) \leq \sum_{i \leq m} \nu(K_i) \leq \sum_{i \leq m} \tilde{\nu}(V_i),$$

co, wobec dowolności K , oznacza subaddytywność $\tilde{\nu}(V) \leq \sum_{i \in \mathbb{N}} \tilde{\nu}(V_i)$.

Dowód (3.40) wynika z faktu, iż dla $L \in \mathfrak{K}^G$ takiego, że $L \subset U \setminus K$, z addytywności miary ν wynika nierówność $\nu(K) + \nu(L) = \nu(K \cup L) \leq \tilde{\nu}(U)$. $\square$

Dowód twierdzenia 3.3.4. Zgodnie z uwagą 1.1.12 i twierdzeniem Carathéodory'ego, funkcja $\tilde{\nu}$ zdefiniowana w (3.39) rozszerza się, co uzasadnia zachowanie oznaczenia, do miary Carathéodory'ego $\tilde{\nu}: \mathfrak{C}_{\tilde{\nu}} \rightarrow [0, \infty]$, przy czym

$$\tilde{\nu}(A) = \inf \{ \tilde{\nu}(U); A \subset U, U \in \tau \} \quad (3.41)$$

dla dowolnych mierzalnych $A \subset T$. Zauważmy, że dla dowolnych zbiorów otwartych U, V w T i zbioru zwartego $K \subset U \cap V$, z własności (3.40) otrzymujemy $\nu(K) + \tilde{\nu}(U \setminus V) \leq \tilde{\nu}(U)$ – jeśli tylko $K \in \mathfrak{K}^G$. Wynikająca stąd nierówność

$$\tilde{\nu}(U \cap V) + \tilde{\nu}(U \setminus V) \leq \tilde{\nu}(U),$$

oznacza – w myśl twierdzenia 1.1.35 – mierzalność zbiorów otwartych i w konsekwencji wszystkich zbiorów borelowskich. Aby teraz zakończyć konstrukcję i stwierdzić, że miara Carathéodory’ego $\tilde{\nu}$ jest regularną miarą borelowską, pozostaje sprawdzić równość $\tilde{\nu}(K) = \nu(K)$ dla $K \in \mathfrak{K}^G$. W tym celu wykazemy najpierw zgodność obu miar na zbiorach otwartych postaci

$$\{f > c\} \in \mathfrak{B}_{\text{are}}^\sigma(T), \quad \text{dla } c > 0 \text{ i } f \in \mathcal{C}_c(T).$$

Istotnie, z tożsamości $\{x; f(x) > c\} = \bigcup_{n \in \mathbb{N}} \{x; f(x) \geq c + \frac{1}{n}\}$ oraz z ciągłości miary ν wynika nierówność

$$\tilde{\nu}(\{f > c\}) \geq \sup_{c' > c} \nu(\{f \geq c'\}) = \nu(\{f > c\}). \quad (3.42)$$

Ponieważ każdy zbiór zwarty $K \subset \{f > c\}$ ma własność $K \subset \{f \geq c'\}$ dla $c' = \inf(f|_K) > c$, zachodzi także nierówność $\tilde{\nu}(\{f > c\}) \leq \nu(\{f > c\})$, przeciwna do (3.42). W analogiczny sposób, korzystając z wykazanej już zgodności miar na wyróżnionych zbiorach otwartych, stwierdzamy nierówność

$$\tilde{\nu}(\{f \geq c\}) \leq \inf_{c' < c} \tilde{\nu}(\{f > c'\}) = \inf_{c' < c} \nu(\{f > c'\}) = \nu(\{f \geq c\}).$$

Aby uzyskać nierówność przeciwną, dla dowolnego $\varepsilon > 0$ rozważmy zbiór otwarty $U \supset \{f \geq c\}$ taki, że $\tilde{\nu}(\{f \geq c\}) + \varepsilon \geq \tilde{\nu}(U)$. Skoro $f(x) < c$ dla $x \notin U$, wystarczy przyjąć $c' = \sup(f|_{T \setminus U}) < c$ – ze względu na zwarty nośnik funkcji f – by uzyskać inkluzję $U \supset \{f > c'\}$. Stąd nierówności

$$\tilde{\nu}(\{f \geq c\}) + \varepsilon \geq \nu(\{f > c'\}) \geq \nu(\{f \geq c\}),$$

potwierdzające zgodność $\tilde{\nu}(K) = \nu(K)$ dla dowolnego $K \in \mathfrak{K}^G$. Z twierdzenia o jednoznaczności miar wynika równość $\tilde{\nu}(A) = \nu(A)$ dla $A \in \mathfrak{B}_{\text{are}}^\sigma(T)$. Wzajemnie jednoznaczna odpowiedniość miar ν i $\tilde{\nu}$ jest konsekwencją jednoznaczności rozszerzenia i wynika z wykorzystania w konstrukcji własności (3.39) i (3.41).

Dla dowodu własności (3.38) przypominamy, że zgodnie z definicją produktu miar zachodzi równość

$$(\mathfrak{B}_S \otimes \mathfrak{B}_T)_{\tilde{\mu} \otimes \tilde{\nu}}^\sigma = \mathfrak{B}_S^\sigma \otimes \mathfrak{B}_T^\sigma,$$

gdzie korzystamy z uproszczonych oznaczeń dla σ -pierścieni $(\mathfrak{B}_S)_\mu^\sigma$ i $(\mathfrak{B}_T)_\nu^\sigma$ zbiorów σ -skończonych względem poszczególnych miar regularnych. Rozważmy zatem dowolny zbiór postaci $A = B \times C$, dla $B \in \mathfrak{B}_S^\sigma$ i $C \in \mathfrak{B}_T^\sigma$. We wzorze (3.37)

$$(\mu \otimes \nu)^\sim(A) = \sup \{(\mu \otimes \nu)(L); L \subset B \times C \text{ i zbiór } L \text{ jest zwarty}\},$$

możemy ograniczyć się do zbiorów zwartych postaci $L = K \times K'$, gdzie $K \subset B$ i $K' \subset C$ są rzutami na odp. współrzędne. Otrzymujemy stąd

$$\begin{aligned} (\mu \otimes \nu)^\sim(A) &= \sup \{(\mu \otimes \nu)(K \times K'); K \times K' \subset A, K \subset B \text{ i } K' \subset C - \text{zwarte}\} \\ &= \sup \{\mu(K); K \subset B\} \cdot \sup \{\nu(K'); K' \subset C\} \\ &= \tilde{\mu}(B) \cdot \tilde{\nu}(C) = (\tilde{\mu} \otimes \tilde{\nu})(A). \end{aligned}$$

Dowodzona równość wynika z twierdzenia 1.1.29 o jednoznaczności miar zastosowanego do π -systemu generującego produkt $\mathfrak{B}_S^\sigma \otimes \mathfrak{B}_T^\sigma$. $\square$

Wybrane fragmenty powyższego dowodu pozwalają na wyróżnienie własności, które można uznać za *regularność* miar Baire'a.

Stwierdzenie 3.3.6. *Dla dowolnej miary Baire'a ν w przestrzeni lokalnie zwartej (T, τ) zachodzi równość*

$$\nu(A) = \begin{cases} \inf \{\nu(U); A \subset U \in \mathfrak{B}_{\text{ave}}^\sigma(T), U \text{ otwarty}\} \\ \sup \{\nu(K); A \supset K \in \mathfrak{B}_{\text{ave}}^\sigma(T), K \text{ zwarty}\} \end{cases} \quad (3.43)$$

dla $A \in \mathfrak{B}_{\text{ave}}^\sigma(T)$.

Wniosek 3.3.7. *Jeśli $\mathfrak{B}_{\text{ave}}^\sigma(T) = \mathfrak{B}_T$, to każda lokalnie skończona miara borelowska jest regularna (por. twierdzenie 1.1.47).* $\square$

Dowód. Funkcja $\mu = \nu|_{\tau \cap \mathfrak{B}_{\text{ave}}^\sigma(T)}$ rozszerza się do miary zewnętrznej i wyznacza miarę Carathéodory'ego $\tilde{\mu}: \mathfrak{C}_\mu \rightarrow [0, \infty]$ taką, że

$$\tilde{\mu}(A) = \inf \{\nu(U); A \subset U \in \tau \cap \mathfrak{B}_{\text{ave}}^\sigma(T)\} \quad \text{dla } A \in \mathfrak{C}_\mu.$$

Z udowodnionego już twierdzenia wiemy, że dla dowolnego otwartego zbioru Baire'a $U \in \tau \cap \mathfrak{B}_{\text{ave}}^\sigma(T)$ zachodzi równość

$$\nu(U) = \tilde{\nu}(U) = \sup \{\nu(K); K \subset U \text{ i } K \in \mathfrak{K}^G\}.$$

Tym samym nierówność $\nu(K) + \mu^*(U \setminus V) \leq \nu(K) + \mu(U \setminus K) = \nu(U)$, dla $U, V \in \tau \cap \mathfrak{B}_{\text{ave}}^\sigma(T)$ i dowolnego $K \in \mathfrak{K}^G$ takiego, że $K \subset U \cap V$, prowadzi do równości $\nu(U \cap V) + \mu^*(U \setminus V) = \nu(U)$ oznaczającej mierzalność otwartych

zbiorów Baire’a i inkluzję $\mathfrak{B}_{\text{aire}}^\sigma(T) \subset \mathfrak{C}_\mu$. Ponieważ miara Carathéodory’ego $\tilde{\mu}$ pokrywa się z miarą Baire’a ν na π -systemie $\tau \cap \mathfrak{B}_{\text{aire}}^\sigma(T)$, z twierdzenia 1.1.29 wynika równość na $\mathfrak{B}_{\text{aire}}^\sigma(T)$.

Dla zakończenia dowodu pozostaje jeszcze wykazać drugą z równości (3.43) – dla zbioru $A \in \mathfrak{B}_{\text{aire}}^\sigma(T)$ o mierze skończonej. Skorzystamy z wykazanej już częściowej regularności miary ν i głównej linii dowodu stwierdzenia 3.3.2. Dla dowolnego $\varepsilon > 0$, ustalając otwarty zbiór Baire’a $U \supset A$ o skończonej mierze, znajdujemy zwarty zbiór Baire’a $K \subset U$ taki, że $\nu(U \setminus K) < \varepsilon$, oraz otwarty zbiór Baire’a $V \supset U \setminus A$ taki, że $\nu(V \setminus (U \setminus A)) < \varepsilon$, przy czym $V \subset U$. Wynikająca stąd inkluzja $K \setminus V \subset A$ oraz nierówność $\nu(A \setminus (K \setminus V)) \leq 2\varepsilon$ oznaczają możliwość dowolnie bliskiego przybliżania A zwartymi zbiorami Baire’a. $\square$

Przykład 3.3.2. Niech $T = \bigcup_{x \in I} \{x\} \times \mathbb{R}$ będzie sumą rozłączną nieprzeliczalnej liczby przestrzeni $\mathbb{R}$ (podprzestrzeni otwartych). Miara μ równa mierze Lebesgue’a na podprzestrzeni $\{0\} \times \mathbb{R}$, a równa 0 na każdym z pozostałych składników jest miarą lokalnie skończoną na σ -pierścieniu Baire’a $\mathfrak{B}_{\text{aire}}^\sigma(T)$. Generowana przez μ miara regularna znika na dopełnieniu $T \setminus (\{0\} \times \mathbb{R})$, podczas gdy druga z miar (Carathéodory’ego) przyjmuje wartość ∞ na każdej sumie $B = \bigcup_{x \in C} \{x\} \times \mathbb{R}$, o ile tylko zbiór indeksów $C \subset I$ jest nieprzeliczalny. Zbiór borelowski B należy do σ -ciała $\mathfrak{B}_{\text{aire}}^\sigma(T)$ wtedy i tylko wtedy, gdy zbiór indeksów C jest przeliczalny lub ma przeliczalne dopełnienie.

3.3.3 Nośnik miary regularnej

Kontynuując badanie regularnych miar borelowskich przyjmujemy

Definicja 3.3.8. Nośnikiem $\text{supp}(\mu)$ regularnej miary borelowskiej μ w przestrzeni topologicznej T nazywamy najmniejszy zbiór domknięty $F \subset T$, dla którego $\mu(T \setminus F) = 0$.

Dopełnienie nośnika miary składa się z tych punktów, w których otoczeniu miara znika. Własność ta zapewnia jednocześnie poprawność definicji.

Stwierdzenie 3.3.9. Dla dowolnej miary regularnej μ w przestrzeni (T, τ) zachodzi równość

$$T \setminus \text{supp}(\mu) = \{x \in T; \exists U \in \tau \ x \in U \text{ i } \mu(U) = 0\}. \quad (3.44)$$

Dowód. Niech $V \in \tau$ oznacza zbiór występujący po prawej stronie równości (3.44). Dla dowodu równości $\mu(V) = 0$ wystarczy sprawdzić, że $\mu(K) = 0$,

dla dowolnego zbioru zwanego $K \subset V$. Ponieważ taki zbiór K można pokryć skończonym ciągiem zbiorów otwartych $U_i, i \leq n$, dla których $\mu(U_i) = 0$, wnioskujemy stąd, że $\mu(K) \leq \mu(U_1 \cup \dots \cup U_n) = 0$. $\square$

Definicja 3.3.10. Dla dowolnej miary Radona λ w przestrzeni lokalnie zwartej (T, τ) , nośnikiem miary nazywamy zbiór domknięty

$$\text{supp}(\lambda) = T \setminus \bigcup \{U \in \tau; \forall f \in \mathcal{C}_c(T) \text{ supp}(f) \subset U \Rightarrow \lambda(f) = 0\}. \quad (3.45)$$

Ćwiczenie 3.3.3. Sprawdzić, że $\text{supp}(\lambda) = \text{supp}(|\lambda|)$.

Stwierdzenie 3.3.11 (O nośniku miary). Dla dowolnej regularnej miary borelowskiej μ w przestrzeni lokalnie zwartej T zachodzi równość

$$\text{supp}(\mu) = \text{supp}(\lambda),$$

gdzie $\lambda: \mathcal{C}_c(T) \rightarrow \mathbb{R}$ oznacza związaną z μ dodatnią miarę Radona. Ponadto:

- (i) Dla każdej funkcji $f \in \mathcal{C}_c(T)$, jeśli $f|_{\text{supp}(\lambda)} = 0$, to $\lambda(f) = 0$.
- (ii) Dowolna funkcja ciągła $f: T \rightarrow \mathbb{R}$ jest całkowalna względem miary μ , jeśli przekrój $\text{supp}(f) \cap \text{supp}(\mu)$ jest zbiorem zwartym. Wówczas

$$\int f d\mu = \lambda(\varphi \cdot f), \quad (3.46)$$

dla każdej funkcji $\varphi \in \mathcal{C}_c(T)$ równej 1 na zbiorze $\text{supp}(f) \cap \text{supp}(\lambda)$.

- (iii) Nośnik miary jest zawarty w zbiorze zwartym $K \subset T$ (i zwarty) wtedy i tylko wtedy, gdy istnieje stała $C \geq 0$ taka, że

$$|\lambda(f)| \leq C \sup_{x \in K} |f(x)| \quad \text{dla } f \in \mathcal{C}_c(T). \quad (3.47)$$

Dowód. Jeśli miara Radona λ znika na zbiorze otwartym U , to dla każdego $x \in U$ i dowolnego względnie zwanego otoczenia V o domknięciu zawartym w U istnieje funkcja nieujemna $f \in \mathcal{C}_c(T)$, taka że $f|_V = 1$ i $\text{supp}(f) \subset U$. Nierówność

$$\mu(V) \leq \int f d\mu = \lambda(f) = 0$$

oznacza, że $\mu(V) = 0$, a więc $x \in T \setminus \text{supp}(\mu)$ i $\text{supp}(\mu) \subset \text{supp}(\lambda)$. Inkluzja przeciwna jest oczywista.

Własność (i) odwołuje się bezpośrednio do definicji, a postulowana w (ii) równość $\int (1 - \varphi)f d\mu = 0$ także wynika z faktu, że na nośniku miary funkcja podcałkowa jest równa 0. Istotnie, pierwszy z czynników znika na przekroju $\text{supp}(f) \cap \text{supp}(\mu)$, a drugi – na różnicy $\text{supp}(\mu) \setminus \text{supp}(f)$.

(iii) Jeśli $\text{supp}(\mu) \subset K$, to dla dowolnej funkcji $f \in \mathcal{C}_c(T)$

$$|\lambda(f)| \leq \int |f| d\mu = \int_K |f| d\mu \leq \mu(K) \sup_{x \in K} |f(x)|.$$

Odwrotnie, z nierówności (3.47) wynika, że $\lambda(f) = 0$ dla każdej funkcji f o nośniku rozłącznym z K . Stąd $T \setminus K \subset T \setminus \text{supp}(\lambda)$. $\square$

Miary Radona o zwartych nośnikach rozszerzają się zatem do ciągłych funkcjonałów na $\mathcal{C}(T, \mathbb{R})$. Jako wniosek otrzymujemy

Twierdzenie 3.3.12. *W przestrzeni lokalnie zwartej T operacje obcięcia*

$$\mu \leftrightarrow \lambda = (\int d\mu)|_{\mathcal{C}_c(T)} \leftrightarrow \tilde{\lambda} = (\int d\mu)|_{\mathcal{C}(T)}$$

ustalają naturalną bijekcję pomiędzy

- (i) *miarami regularnymi μ w przestrzeni $(T, \mathfrak{B}_T)$, dla których nośnik $\text{supp}(\mu)$ jest zbiorem zwartym,*
- (ii) *dodatnimi miarami Radona $\lambda: \mathcal{C}_c(T) \rightarrow \mathbb{R}$ o zwartym nośniku $\text{supp}(\lambda)$,*
- (iii) *ciągłymi funkcjonałami liniowymi $\tilde{\lambda}: \mathcal{C}(T) \rightarrow \mathbb{R}$ spełniającymi warunek nieujemności $\tilde{\lambda}(f) \geq 0$ dla $f \geq 0$.*

Dowód. Ćwiczenie. $\square$

3.3.4 Całka względem miary regularnej

Zgodnie z twierdzeniem Riesza 3.3.3, istnieje naturalna droga prowadząca od dowolnej dodatniej miary Radona $\lambda: \mathcal{C}_c(T) \rightarrow \mathbb{R}$, poprzez regularną miarę borelowską ϱ_λ , do całki Lebesgue'a $\int d\varrho_\lambda$ będącej rozszerzeniem funkcjonału λ . W dalszym ciągu skupimy się na pytaniu: *Jaka własność całki Lebesgue'a charakteryzuje w przestrzeni lokalnie zwartej regularne miary borelowskie?* Skorzystamy w tym miejscu z opisanej w [11] konstrukcji odwołującej się do rodziny funkcji półciągłych z dołu i stanowiącej w pewnym sensie rozszerzenie całki Daniella–Stone'a.

W dowolnej przestrzeni lokalnie zwartej T niech $\mathcal{C}_\uparrow(T)$ oznacza zbiór funkcji rozszerzonych półciągłych z dołu i ograniczonych z dołu przez funkcje o zwartym nośniku (czyli – niezerowe poza, ewentualnie, zbiorem zwartym). Analogicznie, niech $\mathcal{C}_\downarrow(T) = \{-f; f \in \mathcal{C}_\uparrow(T)\}$ oznacza zbiór funkcji półciągłych z góry. Przypominamy także oznaczenie *zbioru pod wykresem* $W_g = \{(x, t); 0 < t < g(x)\}$ dla dowolnej funkcji nieujemnej g .

Twierdzenie 3.3.13 (Całka względem miary regularnej). *Dla dowolnej regularnej miary borelowskiej μ w przestrzeni lokalnie zwartej T*

- (i) $\int g d\mu = \sup \{ \int \varphi d\mu; \varphi \leq g, \varphi \in \mathcal{C}_c(T) \}$ dla $g \in \mathcal{C}_+(T)$, oraz
- (ii) $\int f d\mu = \inf \{ \int g d\mu; f \leq g, g \in \mathcal{C}_+(T) \}$ dla $f \in \overline{\mathcal{M}}_+(\mathfrak{B}_T)$.
- (iii) Całka z dowolnej funkcji $f \in \overline{\mathcal{I}}(T, \mu)$ wynosi

$$\int f d\mu = \begin{cases} \inf \{ \int g d\mu; f \leq g, g \in \mathcal{C}_+(T) \} \\ \sup \{ \int h d\mu; h \leq f, h \in \mathcal{C}_-(T) \}, \end{cases} \quad (3.48)$$

przy czym dla każdego $\varepsilon > 0$ istnieje funkcja $\varphi \in \mathcal{C}_c(T)$ taka, że $\int |f - \varphi| d\mu < \varepsilon$.

Każda lokalnie skończona miara borelowska μ w T spełniająca warunki (i)–(ii) jest miarą regularną.

Ćwiczenie 3.3.4. Sprawdzić, że równości (3.48) przenoszą się na funkcje całkowalne $f \in \overline{\mathcal{I}}(T, \mu^c)$ względem uzupełnienia μ^c miary regularnej μ .

Dla dowodu twierdzenia skorzystamy ze stwierdzenia 1.2.26.

Lemat 3.3.14. (i) *Dla dowolnego podzbioru otwartego $U \subset T \times \mathbb{R}$ funkcja*

$$g_U : T \ni x \mapsto \int 1_U(x, t) \ell(dt) \in [0, \infty]$$

jest półciągła z dołu. Jeśli $U \subset W_f$ dla funkcji $f \geq 0$, to $g_U \leq f$.

(ii) *Dla dowolnej nieujemnej funkcji półciągłej z dołu $g \in \mathcal{C}_+(T)$ i dowolnego zbioru zwartego $L \subset W_g$ istnieje funkcja $\varphi \in \mathcal{C}_c(T)$ taka, że $0 \leq \varphi \leq g$ i $L \subset W_\varphi$.*

Dowód. Dla dowodu (i) rozważmy otwarty podzbiór U produktu $T \times \mathbb{R}$, oraz zbiór $U^x = \{t \in \mathbb{R}; (x, t) \in U\}$, dla dowolnego ustalonego punktu $x \in T$. Ze względu na regularność miary Lebesgue'a zachodzi równość

$$g_U(x) = \ell(U^x) = \sup \{ \ell(K); K \subset U^x, K \text{ – zwarty} \}.$$

Jeśli liczba c spełnia warunek $c < g_U(x)$, a zbiór zwarty $K \subset U^x$ ma miarę $\ell(K) > c$, to istnieje otoczenie $V \ni x$ takie, że $V \times K \subset U$. Wówczas $g_U(y) > c$ dla $y \in V$.

Jeśli U zawiera się w zbiorze W_f dla pewnej funkcji f , to z inkluzji $U^x \subset (0, f(x))$ wynika nierówność $g_U(x) \leq f(x)$, dla $x \in T$.

(ii) Załóżmy, że $g \in \mathcal{C}_+(T)$, a $L \subset W_g$ jest dowolnym podzbiorem zwartym. Zauważmy, że dla każdego punktu $x \in K := \pi_1(L) \subset T$ istnieje względnie zwarte otoczenie U_x oraz liczba $r_x > 0$ taka, że

$$L \cap (U_x \times \mathbb{R}) \subset U_x \times (0, r_x) \quad \text{ i } \quad \overline{U_x} \times (0, r_x] \subset W_g. \quad (3.49)$$

Istotnie, przy ustalonym x zbiór zwarty $L \cap (\{x\} \times \mathbb{R})$ posiada otoczenie postaci $U' \times (0, r_x)$, którego domknięcie zawiera się w W_g . Z drugiej strony, dla liczby $M = \sup(\pi_2(L))$ zbiór zwarty $\{x\} \times [r_x, M]$ jest rozłączny z L , a zatem istnieje otoczenie $U'' \ni x$ takie, że $L \cap (U'' \times [r_x, M]) = \emptyset$. Wystarczy przyjąć $U_x = U' \cap U''$.

Wybierając spośród zbiorów $U_x, x \in K$, skończone pokrycie, otrzymujemy ciąg zbiorów $V_i = U_{x_i}, i \leq m$, o własności (3.49), dla pewnych liczb $c_i = r_{x_i}, i \leq m$. Niech $\psi_1, \dots, \psi_m \in \mathcal{C}_c(T)$ będzie ciągłym rozkładem jedynki wpisanym w pokrycie. Wykażemy, że funkcja $\varphi = \sum_{i \leq m} c_i \psi_i \in \mathcal{C}_c(T)$ spełnia tezę lematu. Istotnie, z (3.49) wynika prawdziwość implikacji

$$(x, t) \in L \text{ i } x \in V_i \implies t < c_i \text{ dla } i \leq m.$$

Jeśli zatem dla dowolnego punktu $x \in \text{supp } \varphi$ rozważyć zbiór indeksów $\{i_1, \dots, i_r\} = \{i \leq m; \psi_i(x) > 0\}$, to dla $(x, t) \in L$ otrzymujemy

$$t < \min(c_{i_1}, \dots, c_{i_r}) \leq \sum_{k \leq r} c_{i_k} \psi_{i_k}(x) = \varphi(x),$$

czyli $(x, t) \in W_\varphi$. Z drugiej strony, z inkluzji $\{x\} \times (0, c_{i_k}] \subset W_g$, dla $k \leq r$, wynikają nierówności

$$\varphi(x) = \sum_{k \leq r} c_{i_k} \psi_{i_k}(x) \leq \max(c_{i_1}, \dots, c_{i_r}) < g(x),$$

co kończy dowód. $\square$

Dowód twierdzenia 3.3.13. Niech μ będzie regularną miarą borelowską w T . Zauważmy, że dla każdej nieujemnej funkcji $g: T \rightarrow [0, \infty]$ półciągłej z dołu zachodzi równość

$$\int g d\mu = (\mu \otimes \ell)(W_g) = \sup \{(\mu \otimes \ell)(L); L \subset W_g \text{ i } L \in \mathfrak{K}_{T \times \mathbb{R}}\}. \quad (3.50)$$

Istotnie, w przypadku, gdy $W_g \in \mathfrak{B}_T^\sigma \otimes \mathfrak{B}_\mathbb{R}$, równość (3.50) wynika bezpośrednio z częściowej regularności (3.38) produktu miar. Załóżmy zatem, że zbiór W_g nie jest σ -skończony. Z rozkładu (1.53) wynika, że tę samą własność ma przynajmniej jeden ze zbiorów $\{g > c\} \times (0, c)$, a więc $\{g > c\} \notin \mathfrak{B}_T^\sigma$, dla pewnego $c > 0$. Regularność miary μ implikuje związek

$$\mu(\{g > c\}) = \sup \{\mu(K); K \subset \{g > c\} \text{ i } K \in \mathfrak{K}_T\},$$

a zatem miary zbiorów zwartych postaci $K \times [a, b] \subset \{g > c\} \times (0, c)$ mają kres górny równy ∞ . Oznacza to równość (3.50) także i w tym przypadku. Korzystając z lematu 3.3.14(ii) otrzymujemy pierwszą część tezy.

Przypadek ogólny, funkcji półciągłej $g \geq \phi \in \mathcal{C}_c(T)$, sprowadzamy do poprzedniego – dla funkcji $g - \phi$.

Dla dowodu (ii) założmy na początek, że $f \geq 0$ jest funkcją $\mathfrak{B}_T^\sigma$ -mierzalną, co zgodnie ze stwierdzeniem 1.2.26 jest równoważne $\mathfrak{B}_T^\sigma \otimes \mathfrak{B}_\mathbb{R}$ -mierzalności zbioru W_f . Przy takim założeniu, z częściowej regularności (3.38) produktu miar wynika równość

$$(\mu \otimes \ell)(W_f) = \inf \{ (\mu \otimes \ell)(U); W_f \subset U \in \mathfrak{B}_T^\sigma \otimes \mathfrak{B}_\mathbb{R}, U \text{ otwarty} \},$$

to znaczy

$$\int f d\mu = \inf \left\{ \int g_U d\mu; W_f \subset U, U - \text{otwarty i } \sigma\text{-skończony} \right\}.$$

W tym przypadku tożsamość (ii) wynika z lematu 3.3.14(i).

Jeśli $\int f d\mu = \infty$ i zbiór W_f nie jest σ -skończony, wówczas ta sama własność przysługuje przynajmniej jednemu ze zbiorów $\{f > c\} \times (0, c)$ dla pewnego $c > 0$. Wnioskujemy stąd, że zbiór borelowski $\{f > c\} \notin \mathfrak{B}_T^\sigma$ ma miarę nieskończoną, a zatem dla każdej funkcji $g \in \mathcal{C}_\uparrow(T)$ takiej, że $g \geq f$, inkluzja $\{f > c\} \subset \{g > c\}$ pociąga za sobą nieskończoność całki $\int g d\mu = \infty = \int f d\mu$.

(iii) Z udowodnionych już własności (i)–(ii) wynika, że dla nieujemnej funkcji borelowskiej μ -całkowalnej i liczby $\varepsilon > 0$ istnieją funkcje $g \in \mathcal{C}_\uparrow(T)$ i $\phi \in \mathcal{C}_c(T)$ takie, że $f, \phi \leq g$ oraz $\int (g - f) d\mu < \frac{\varepsilon}{2}$ i $\int (g - \phi) d\mu < \frac{\varepsilon}{2}$. Stąd

$$\int |f - \phi| d\mu \leq \int (|f - g| + |g - \phi|) d\mu < \varepsilon.$$

Analogiczne nierówności dla składowych f^+, f^- dowolnej funkcji $f \in \overline{\mathcal{I}}(T, \mu)$ prowadzą do stanowiącej część tezy nierówności $\int |f - \varphi| d\mu < \varepsilon$ dla pewnej funkcji ciągłej φ o zwartym nośniku. Zgodnie z (ii), dla każdego $\varepsilon > 0$ istnieją zatem funkcje $\varphi \in \mathcal{C}_c(T)$ i $g' \in \mathcal{C}_\uparrow(T)$ takie, że $|f - \varphi| \leq g'$ i $\int g' d\mu < \varepsilon$. Funkcje $h := -g' + \varphi \in \mathcal{C}_\downarrow(T)$ i $g := g' + \varphi \in \mathcal{C}_\uparrow(T)$ spełniają warunki

$$h \leq f \leq g \quad \text{ i } \quad \max \left\{ \int (f - h) d\mu, \int (g - f) d\mu \right\} < \varepsilon,$$

co kończy dowód. $\square$

Uwaga 3.3.5. Z twierdzenia 3.3.13 wynika *bezpośrednia* konstrukcja rozszerzenia dowolnej dodatniej miary Radona $\lambda: \mathcal{C}_c(T) \rightarrow \mathbb{R}$ do całki względem regularnej miary borelowskiej ϱ_λ . Jeśli przyjąć

$$\lambda_\uparrow(g) = \sup \{ \lambda(\varphi); \varphi \leq g, \varphi \in \mathcal{C}_c(T) \}$$

dla $g \in \mathcal{C}_\uparrow(T)$, oraz

$$\lambda_\uparrow(f) = \inf \{ \lambda_\uparrow(g); f \leq g, g \in \mathcal{C}_\uparrow(T) \}$$

dla $f \in \overline{\mathcal{M}}(\mathfrak{B}_T)$, to – analogicznie jak w ogólnej konstrukcji Daniella – otrzymujemy półnormę $N(f) = \lambda_\uparrow(|f|)$ w podprzestrzeni przestrzeni liniowej

$\mathcal{M}(\mathfrak{B}_T)$ i ciągły funkcjonal λ o jednoznacznym rozszerzeniu na domknięcie podprzestrzeni $\mathcal{C}_c(T)$. Pozostawiamy czytelnikowi sprawdzenie, że owym rozszerzeniem jest całka Lebesgue’a $\int d\varrho_\lambda$ względem odpowiedniej regularnej miary borelowskiej ϱ_λ .

Ćwiczenie 3.3.6. W odróżnieniu od konstrukcji z podrozdziału 3.1, funkcje półciągłe z dołu tworzące $\mathcal{C}_+(T)$ nie muszą być osiągalne z dołu. Mimo to pierwsze z rozszerzeń, $\lambda_+ : \mathcal{C}_+(T) \rightarrow (-\infty, \infty]$, jest funkcją addytywną (i dodatnio jednorodną), co wynika z twierdzenia 3.3.13 i odpowiedniej własności całki. Wykazać addytywność bezpośrednio z definicji λ_+ .

Definicja 3.3.15. Borelowską całką Radona *względem dodatniej miary Radona w przestrzeni lokalnie zwartej T nazywamy każdą z całek Lebesgue’a względem miary regularnej λ^r odpowiadającej mierze Radona λ ,*

$$\int d\lambda^r : \mathcal{I}(T, \lambda^r) \rightarrow \mathbb{R}, \quad \text{oraz} \quad \int d\lambda^r : L^1(T, \lambda^r) \rightarrow \mathbb{R}.$$

Całki Radona $\int d\lambda^r$ stanowią rozszerzenie całki Daniella–Stone’a z przestrzeni liniowej skończonych funkcji λ –całkowalnych $\mathcal{I}(T, \lambda)$, oraz, odpowiednio, z przestrzeni unormowanej klas równoważności funkcji λ –całkowalnych.

Spis symboli i skrótów

Alfabet wykorzystywany w tekście matematycznym – klasyczny (roman), fraktura (gotyk) i kaligrafia (script):

$A, \mathfrak{A}, \mathcal{A}$	$B, \mathfrak{B}, \mathcal{B}$	$C, \mathfrak{C}, \mathcal{C}$	$D, \mathfrak{D}, \mathcal{D}$	$E, \mathfrak{E}, \mathcal{E}$	$F, \mathfrak{F}, \mathcal{F}$	$G, \mathfrak{G}, \mathcal{G}$
$H, \mathfrak{H}, \mathcal{H}$	$I, \mathfrak{I}, \mathcal{I}$	$J, \mathfrak{J}, \mathcal{J}$	$K, \mathfrak{K}, \mathcal{K}$	$L, \mathfrak{L}, \mathcal{L}$	$M, \mathfrak{M}, \mathcal{M}$	$N, \mathfrak{N}, \mathcal{N}$
$O, \mathfrak{O}, \mathcal{O}$	$P, \mathfrak{P}, \mathcal{P}$	$Q, \mathfrak{Q}, \mathcal{Q}$	$R, \mathfrak{R}, \mathcal{R}$	$S, \mathfrak{S}, \mathcal{S}$	$T, \mathfrak{T}, \mathcal{T}$	$U, \mathfrak{U}, \mathcal{U}$
	$V, \mathfrak{V}, \mathcal{V}$	$W, \mathfrak{W}, \mathcal{W}$	$X, \mathfrak{X}, \mathcal{X}$	$Y, \mathfrak{Y}, \mathcal{Y}$	$Z, \mathfrak{Z}, \mathcal{Z}$	

Symbol	Opis	Definicja
$\mathfrak{P}(T)$	rodzina wszystkich podzbiorów zbioru T	str. 5
$\sigma_T(\mathfrak{C})$	σ –ciało generowane	str. 5
$\mathfrak{B}_\tau, \mathfrak{B}_T$	σ –ciało zbiorów borelowskich w przestrzeni topologicznej (T, τ)	str. 6
$\mathfrak{C} A$	ograniczenie rodziny zbiorów	str. 6
$\mathcal{M}(\mathfrak{T}), \overline{\mathcal{M}}(\mathfrak{T})$	przestrzenie funkcji mierzalnych względem σ –ciała	str. 6
$\prod_a(S_a, \mathfrak{S}_a), \pi_\alpha$	produkt przestrzeni mierzalnych, projekcje	str. 7
$\mathfrak{B}_\mathbb{R}^T$	produktowe σ –ciało zbiorów w $\mathbb{R}^T$	str. 7
$\bigotimes_a \mathfrak{S}_a$	produktowe σ –ciało	str. 7
$\sigma(\mathfrak{C}), \delta(\mathfrak{C})$	σ –pierścień/ δ –pierścień generowany przez rodzinę zbiorów	str. 8
$\mathcal{M}(\mathfrak{F}), \overline{\mathcal{M}}(\mathfrak{F})$	przestrzenie funkcji mierzalnych względem σ –pierścienia	str. 10
$\mathcal{A}^\sigma, \sigma\mathcal{A}$	σ –pierścień/ σ –ciało generowane przez przestrzeń funkcji $\mathcal{A}$	str. 11
δ_x	miara Diraca w punkcie x	str. 11
ρ_S	miara zliczająca względem zbioru S	str. 11
$\mathfrak{T}_\mu^i$	δ –pierścień zbiorów o mierze μ skończonej	str. 12
$\mathfrak{T}_\mu^\sigma, \mathfrak{T}_\mu^{loc}$	σ –pierścień zbiorów σ –skończonych/ pół-skończonych względem miary μ	str. 12

Symbol	Opis	Definicja
μ^σ, μ^{loc}	składowa σ –skończona/półskończona miary μ	str. 12
$\mathcal{M}(T, \mu)$	zbiór funkcji μ –mierzalnych	str. 12
$\mathcal{M}(T, \mu)$	przestrzeń <i>skończonych funkcji</i> μ –mierzalnych.	str. 12
$g_*\mu, g^*$	miara indukowana/homomorfizm indukowany przez odwzorowanie mierzarne	str. 13
$\mu _S$	miara ograniczona (obcięta) do podprzestrzeni	str. 13
ℓ	miara Lebesgue’a na prostej $\mathbb{R}$	str. 14
ν^*	miara zewnętrzna wyznaczona przez ν	str. 15
$\mu_{\mathfrak{C}}: \mathfrak{C}_\mu \rightarrow \mathbb{R}_+$	miara Carathéodory’ego	str. 15
$\mu_1 \otimes \dots \otimes \mu_n$	produkt miar	str. 16
F_μ	dystrybuanta miary borelowskiej	str. 17
Q^F	funkcja kwantylowa miary borelowskiej w $\mathbb{R}$	str. 17
$\Delta_A F$	zależność miary przedziału w $\mathbb{R}^n$ od dystrybuanty	str. 18
$\mu_n \xrightarrow{w} \mu$	słaba zbieżność ciągu miar borelowskich	str. 20
∂C	brzeg zbioru	str. 21
$F_n \xrightarrow{w} F$	słaba zbieżność ciągu dystrybuant	str. 22
f^+, f^-	części – dodatnia i ujemna funkcji mierzarnej	str. 23
$\int d\mu$	całka Lebesgue’a względem miary μ	str. 23
$\mathcal{I}(T, \mu)$	przestrzeń skończonych funkcji μ –całkowalnych (w sensie Lebesgue’a)	str. 23
$\mathcal{P}(T, \mu)$	przestrzeń funkcji prostych	str. 25
$\int_S d\mu$	całka po zbiorze mierzalnym S	str. 26
$\mathcal{C}_b(M)$	przestrzeń ograniczonych funkcji ciągłych	str. 26
$\int dF$	całka Lebesgue’a–Stieltjesa	str. 28
$\overline{\mathcal{I}}(T, \mu)$	przestrzeń funkcji μ –całkowalnych	str. 29
$\mathcal{D}(T, \mu)$	formalna, maksymalna dziedzina całki Lebesgue’a $\int d\mu$	str. 29
$L^1(T, \mu)$	przestrzeń unormowana klas równoważności funkcji μ –całkowalnych	str. 30
W_f	zbiór pod wykresem	str. 32
$f\mu$	iloczyn funkcji mierzarnej $f \geq 0$ i miary	str. 34
$\nu \ll \mu$	miara ν absolutnie ciągła względem miary μ	str. 35
$\frac{d\nu}{d\mu}$	gęstość, pochodna Radona–Nikodyma miar	str. 35
μ_T	miara brzegowa	str. 36
$\lambda \perp \mu$	miary wzajemnie singularne	str. 37
$\mathcal{I}^\sigma(\mathfrak{S}, \mu), L^\sigma(\mathfrak{S}, \mu)$	przestrzeń funkcji przeliczalnie całkowalnych	str. 45
$\Pi_{\mathfrak{S}}^\pi, \Pi_{\mathfrak{S}}$	całka względem σ –ciała, operator projekcji	str. 45
$(L^p(T, \mu), \ \cdot\ _p)$	przestrzeń Banacha klas funkcji μ –całkowalnych z p –tą potęgą	str. 47

Symbol	Opis	Definicja
$\lambda _T$	rozkład warunkowy miary (względem projekcji)	str. 49
$\mu_{X \mathfrak{G}}$	rozkład warunkowy funkcji względem σ -ciała	str. 50
$\mu_{X Z}, \mu_{X Z=z}$	rozkład warunkowy funkcji względem miary i funkcji	str. 54
$\mathbb{P}^{\mathbb{N}}$	miara probabilistyczna w przestrzeni abstrakcyjnych prób losowych, $\Omega^{\mathbb{N}}$	str. 57
$\mathbb{P}_X$	rozkład zmiennej losowej X	str. 59
$\mathcal{E}(X)$	wartość oczekiwana zmiennej X	str. 60
$\text{Var}(X), \sigma_X$	wariancja, odchylenie standardowe zmiennej	str. 60
F_X	dystrybuenta zmiennej losowej X	str. 62
$X_n \xrightarrow{p} X$,	zbieżność według prawdopodobieństwa	str. 64
$\mathfrak{F}_{\infty}$	σ -ciało resztowe, ogon ciągu σ -ciał	str. 68
$\sigma(X_1, X_2, \dots)$	σ -ciało generowane przez ciąg zmiennych losowych	str. 69
$\mu * \nu$	splot miar borelowskich	str. 73
$\mathbb{P}_{X \mathfrak{G}}$	warunkowy rozkład zmiennej X względem σ -ciała	str. 74
$\mathcal{E}(X \mathfrak{G})$	warunkowa wartość oczekiwana zmiennej X względem σ -ciała	str. 74
$\mathbb{P}_{X Z}, \mathbb{P}_{X Z=z}$	warunkowy rozkład zmiennej X względem zmiennej Z	str. 75
$\mathcal{E}(X Z), \mathcal{E}(X Z=z)$	warunkowa wartość oczekiwana zmiennej X względem zmiennej Z	str. 75
$\text{Beta}(\alpha, \beta)$	rozkład beta	str. 78
$\text{Cov}(X, Y)$	kovariancja pary zmiennych losowych	str. 78
$\rho(X, Y)$	współczynnik korelacji zmiennych	str. 78
$\text{Cov}(X)$	macierz kovariancji zmiennej wektorowej	str. 79
$\text{Var}(X)$	wariancja całkowita zmiennej wektorowej	str. 81
$\mathbb{X} = X(\Omega)$	przestrzeń próbek zmiennej losowej	str. 96
$X^{\mathbb{N}}(\omega)$	próba losowa zmiennej losowej X	str. 96
$\mathbb{P}_{\mathbb{X}}$	miara w przestrzeni prób zmiennej	str. 97
$\bar{X}_n(\omega)$	estymator wartości oczekiwanej zmiennej X	str. 97
Var_n	wariancja empiryczna zmiennej/ próbki	str. 99
$\text{Cov}_n X$	kovariancja empiryczna zmiennej wektorowej	str. 99
F_n^{ω}	empiryczna dystrybuenta	str. 101
Q_n^{ω}	empiryczna funkcja kwantylowa	str. 101
(X_0, a, c, m)	parametry kongruencyjnego ciągu pseudolosowego	str. 109
ν_k^2	wynik testu spektralnego dla k -jednostajności ciągu	str. 111
γ_k	miara k -niejednostajności	str. 112

Symbol	Opis	Definicja
R	zmienna losowa – łączna liczba przebiegów w próbkce	str. 114
$R_{n_1 n_2}$	łączna liczba przebiegów w $\mathbb{R}^n$ – wśród n_1 dolnych i n_2 górnych wyrazów ciągu	str. 117
$WW(n_1, n_2)$	rozkład Walda–Wolfowitza	str. 117
$\text{Pareto}(y, n)$	rozkład Pareta	str. 120
$\text{Pois}(\lambda)$	rozkład Poissona	str. 121
$\text{Gamma}(y + 1, n)$	rozkład Gamma	str. 121
$\text{Exp}(\theta)$	rozkład wykładniczy	str. 121
Φ	dystybuanta rozkładu normalnego Gaussa	str. 128
$\mathcal{N}(0, 1), \mathcal{N}(m, \sigma^2)$	standardowy/ogólny rozkład normalny Gaussa	str. 129
$\mathcal{N}_p(m, W)$	p –wymiarowy rozkład normalny Gaussa	str. 138
$X_n \xrightarrow{d} \mu, X_n \xrightarrow{d} X$	zbieżność ciągu zmiennych losowych wg. rozkładu	str. 144
φ_μ	funkcja charakterystyczna rozkładu μ	str. 142
φ_X	funkcja charakterystyczna zmiennej losowej	str. 147
$\mathcal{P}(M)$	przestrzeń miar probabilistycznych na $(M, \mathfrak{B}_M)$	str. 155
d_ρ	metryka Prochorowa	str. 155
d_L	metryka Levy’ego w $\mathcal{P}(\mathbb{R})$	str. 155
$(\mathcal{S}^Q(T), \mathfrak{S}_T^Q)$	przestrzeń funkcji Q –ośrodkowych w $(\mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T)$	str. 161
$(\mathcal{S}(T), \mathfrak{S}_T)$	przestrzeń funkcji ośrodkowych w $(\mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T)$	str. 161
$\mathcal{D}(\mathbb{R}_+)$	zbiór funkcji (półciągłych) klasy $\mathcal{D}$	str. 168
$\mathcal{D}_r(\mathbb{R}_+)$	przestrzeń funkcji prawostronnie ciągłych	str. 169
$\{B_t\}_{t \geq 0}$	ruch Browna – proces stochastyczny	str. 169
$\mathbb{P}_W$	miara Wienera w $\mathcal{C}(\mathbb{R}_+)$	str. 172
$X^{(n)}$	ciąg procesów zbieżnych do miary Wienera	str. 177
$w(x, \delta)$	moduł ciągłości funkcji x	str. 179
$\chi^*: \overline{\mathcal{M}}(\mathfrak{T}) \rightarrow \overline{\mathbb{R}}$	miara zewnętrzna wyznaczona przez miarę abstrakcyjną χ	str. 183
$\mathcal{A}^*$	zbór funkcji osiągalnych (z dołu) z kraty $\mathcal{A}$	str. 183
$(\mathcal{M}(\mathfrak{T}, \chi), N_1)$	przestrzeń unormowana skończonych funkcji mierzalnych o skończonej mierze zewnętrznej	str. 185
$\int d\chi$	całka Daniella–Stone’a względem miary abstrakcyjnej χ	str. 185
$\mathcal{I}(\mathfrak{T}, \chi), \mathcal{I}(T, \chi)$	przestrzeń skończonych funkcji χ –całkowalnych	str. 185
$f \stackrel{\text{pw}}{=} g$	równość funkcji χ –prawie wszędzie	str. 188
$[0]_\chi$	podprzestrzeń funkcji χ –pomijalnych	str. 188
$(L^1(T, \chi), \ \cdot\)$	przestrzeń unormowana klas równoważności funkcji χ –całkowalnych	str. 189

Symbol	Opis	Definicja
$\overline{\mathcal{I}}(T, \chi)$	zbiór funkcji χ -całkowalnych	str. 189
$\mathcal{A}_*$	zbiór funkcji osiągalnych z $\mathcal{A}$ z góry	str. 189
$\mathfrak{D}_\chi^f$	δ -pierścień zbiorów o χ -całkowalnych funkcjach charakterystycznych	str. 194
$\mathfrak{D}_\chi^\sigma, \mathfrak{D}_\chi$	σ -pierścień i σ -ciało – dziedzina miary wyznaczonej przez całkę Daniella-Stone’a	str. 194
ς_χ	miara Daniella-Stone’a wyznaczona przez χ	str. 194
$(\mathcal{C}(T, \mathbb{F}), \tau_{\mathcal{C}})$	przestrzeń funkcji ciągłych o wartościach w ciele $\mathbb{F} = \mathbb{R}, \mathbb{C}$	str. 197
$(\mathcal{C}_c(T, \mathbb{F}), \tau_c)$	podprzestrzeń funkcji o zwartych nośnikach	str. 197
$\mathcal{C}_K(T, \mathbb{F})$	przestrzeń funkcji ciągłych o nośnikach zawartych w $K \subset T$	str. 197
$\mathfrak{B}_{\text{aire}}^\sigma(T)$	σ -pierścień zbiorów Baire’a	str. 202
$\mathfrak{R}_T^G$	rodzina zbiorów zwartych typu G_δ	str. 203
$\bar{\lambda}, \lambda $	sprzężenie, wartość bezwzględna miary Radona λ	str. 204
$\lambda_{\Re}, \lambda_{\Im}$	część rzeczywista, urojona miary Radona λ	str. 205
λ^+, λ^-	część dodatnia, część ujemna rzeczywistej miary Radona λ	str. 205
$\int d\lambda$	całka Radona względem miary Radona λ	str. 205
$\text{supp}(\mu)$	nośnik miary regularnej	str. 212
$\text{supp}(\lambda)$	nośnik miary Radona	str. 213
$\mathcal{C}_\uparrow(T), \mathcal{C}_\downarrow(T)$	zbiór funkcji półciągłych z dołu, z góry	str. 214
$\lambda_\uparrow$	bezpośrednie rozszerzenie miary Radona λ do całki regularnej	str. 217

Skorowidz

algebra zbiorów, 5

błądzenie losowe, 81

całka

Daniella–Stone’a, 185, 190

Lebesgue’a, 23

po zbiorze mierzalnym, 26

rozszerzenie, 28

względem σ –ciała, 45

Lebesgue’a–Stieletjesa, 28

Radona, 205

borelowska, 218

Riemanna na przedziale, 186

względem miary

na półpierzścieniu, 196

centralne twierdzenie graniczne, 149

w $\mathbb{R}^p$, 153

ciało zbiorów, 5

ciąg

liczbowy

∞ –jednostajny, 107

b –arny, 108

p –jednostajny, 107

miar, słabo zbieżny, 20

cylinder prosty, 61

δ –pierścień, 8

δ –system, 8

dystribuanta

empiryczna, 100

miary borelowskiej

w $\mathbb{R}$, 17

w $\mathbb{R}^p$, 18

wektora losowego, 71

estymator nieobciążony, 98

funkcja

χ –całkowalna, 185

charakteryzacja, 189

χ –całkowalna, rozszerzona, 189

χ –pomijalna, 187

kwantylowa, 62

empiryczna, 101

μ –całkowalna, 24, 29

μ –mierzalna, 12

mierzalna

μ –całkowalna, 29

względem σ –pierścienia, 10

prosta, 25

przeliczalnie całkowalna, 45

Q –ośrodkowa, 161

$\mathfrak{S}/\mathfrak{T}$ –mierzalna, 5

funkcja charakterystyczna

rozkładu normalnego, 150

rozkładu prawdopodobieństwa, 142

transformata odwrotna, 143

transformata odwrotna

w $\mathbb{R}^p$, 151

zmiennej losowej, 147

wektorowej, 151

funkcje

mierzalne, 6

osiągalne

z dołu, 183

z góry, 189

przeliczalnie całkowalne, 30

homomorfizm indukowany, 13

- iloczyn miary i funkcji, 34
- kierunki główne wektor losowego, 80
- kowariancja
 - empiryczna, 99
 - pary zmiennych losowych, 78
 - wektora losowego, 79
- krata, 23
 - Stone'a, 193
- kryterium Kołmogorowa, 91
- λ -system, 8
- lemat
 - Dooba, 161
 - Fatou, 193
 - Kroneckera, 89
 - o zbieżności prawie na pewno, 90
- mediana, 62
- metoda liniowej kongruencji, 109
- metryka
 - Levy'ego, 155
 - Prochorowa, 155
- miara
 - k -niejednostajności, 112
 - absolutnie ciągła, 35
 - gęstość, 35
 - abstrakcyjna, 182
 - Baire'a, 202
 - regularność, 211
 - borelowska, 16
 - dystrybuanta, 17
 - lokalnie skończona, 16
 - regularna, 207
 - Carathéodory'ego, 16
 - Daniella–Stone'a, 194
 - delta Diraca, 11
 - dyskretna, 35
 - indukowana przez odwzorowanie, 13
 - Lebesgue'a na prostej, 14
 - Lebesgue'a w $\mathbb{R}^n$, 16
 - obcięcie do podprzestrzeni, 13
 - półskończona, 12
 - produktowa, 16
 - Radona, 201
 - dodatnia, 201
 - zespólna, 204
 - σ -skończona, 11
 - singularna, 37
 - składowa σ -skończona, 12
 - składowa półskończona, 12
 - skończona, 11
 - uzupełnienie, 14
 - w przestrzeni mierzalnej, 11
 - Wienera, 172
 - zewnętrzna, 15
 - Daniella, 185
 - zliczająca, 11
 - zupełna, 14
- moduł
 - ciągłości funkcji, 179
 - ciągu kongruencyjnego, 109
 - miary Radona, 204
- monotoniczna ciągłość
 - całki Lebesgue'a, 23
 - miary, 13
- nierówność Kołmogorowa, 91
- nierówności Czebyszewa, 60
- nieujemny funkcjonal liniowy, 23
- niezależność
 - ciągu zmiennych losowych, 70
 - rodziny zdarzeń, 67
- nośnik
 - miary Radona, 213
 - regularnej miary borelowskiej, 212
- odchylenie standardowe, 60
- odległość Prochorowa, 155
- odwzorowanie przestrzeni mierzalnych, 5
- ośrodek, 161
- π -system, 7
- pierścień zbiorów, 8
- podprzestrzeń
 - przestrzeni mierzalnych, 6
 - przestrzeni półmierzalnej, 10
- podział przedziału, 186
- półpierścień zbiorów, 16
- populacja, 56
- próba losowa, 96
 - abstrakcyjna, 57
 - standardowa, 96
 - wektorowa, 99
- próbka losowa, 96
- próbka zdarzeń, n -elementowa, 71
- prawo 0/1
 - Hewitta–Savage'a, 83

- Kołmogorowa, 69
- prawo wielkich liczb, mocne, 89, 93
 - dowód, 92
- proces stochastyczny, 157
 - rozkład prawdopodobieństwa, 158
 - uniwersalny, 158
 - Wienera, 170
- produkt
 - miar, 16
 - przeźrzeni mierzalnych, 7
- przebieg
 - długość, 115
 - rosnący/malejący, 115
- przeliczalna całkowalność, 44
- przeźrzeni
 - funkcji
 - ośrodkowych, 161
 - prostych, 25
 - klas równoważności funkcji
 - całkowalnych, 30
 - całkowalnych z p -tą potęgą, 48
 - mierzalna, 5
 - półmierzalna, 9
 - prób losowych, 96
 - probabilistyczna, 56
 - topologiczna
 - liniowa, 154
 - z miarą, 11
 - uzupełnienie, 14
 - zdarzeń elementarnych, 56
- PWL, *zob.* prawo wielkich liczb, mocne
- relacja równości
 - χ -prawie wszędzie, 188
- równoważność
 - procesów stochastycznych, 157
- rozkład
 - beta, 78
 - gamma, 121
 - Gaussa, 129
 - Hahna, 38
 - normalny, 129
 - p -wymiarowy, 138
 - standardowy, 129
 - Pareta, 120
 - Poissona, 121
 - prawdopodobieństwa, 62
 - ciągły, 64
 - procesu stochastycznego, 158, 162
 - warunkowy
 - miary, 50
 - warunkowy funkcji, 54
 - wektora losowego, 71
 - wykładniczy, 121
- ruchy Browna, 170
- schemat Bernoulliego, 77
- σ -ciało, 5
 - generowane przez
 - przeźrzeni funkcji, 11
 - zmienne losowe, 69
 - resztowe, 69
 - zbiorów borelowskich, 6
- σ -pierścień, 8
 - Baire'a, 202
 - generowany przez
 - przeźrzeni funkcji, 11
- składowe główne wektora losowego, 80
- splot
 - funkcji gęstości, 73
 - miar borelowskich w $\mathbb{R}$, 73
- standardowy rozkład normalny, 129
- statystyka dostateczna, 122
- suma Riemanna, 186
- test
 - runs, Walda–Wolfowitza, 117
 - spektralny, 110
- topologia
 - zbieżności jednostajnej, 198
 - zbieżności niemal jednostajnej, 198
 - metryzowalność, 201
- trajektoria procesu stochastycznego, 161
- trajektorie
 - ciągłe, 164
 - jednostronnie ciągłe, 168
- transformata Fouriera, 142
- twierdzenie
 - Carathéodory'ego, 15
 - Daniella–Kołmogorowa, 159
 - de Moivre'a – Laplace'a, 128
 - Donskera, 177
 - dowód, 180
 - Fubini'ego, 32
 - Gliwienki–Cantelliego, 101
 - Hahna–Kołmogorowa, 16

- Kołmogorowa
 - prawo 0/1, 69
- Kołmogorowa–Czencowa, 166
- Lebesgue’a
 - o rozkładzie miary, 37
 - o zbieżności jednostajnej, 206
 - o zmajoryzowanej zbieżności, 193
- Lebesgue’a–Radona–Nikodyma, 37
- Lévy’ego–Craméra, o ciągłości, 145
- o błędzeniu losowym I, 81
- o błędzeniu losowym II, 85
 - dowód, 87
- o całce Lebesgue’a, 23
- o całce relatywnej, 45
 - w L^p , 48
- o całce względem miary regularnej, 215
 - dowód, 216
- o charakteryzacji
 - funkcji χ –całkowalnych, 189
 - wektora losowego typu Gaussa, 138
- o ciągłości, 145
 - w $\mathbb{R}^p$, 152
- o dualności dla miar Radona, 203
- o geometrycznej interpretacji całki, 32
- o jednoznaczności miar, 14
- o miarach regularnych, 208
 - dowód, 209
- o mierze warunkowej, 49
- o monotonicznej zbieżności
 - dla całki Daniella–Stone’a, 192
- o monotonicznych klasach zbiorów, 9
- o nieskończonym produkcie miar, 57
 - dowód, 61
- o niezależności
 - grup zmiennych losowych, 72
 - σ –ciał, 68
 - zmiennych losowych, 71
- o nośniku miary, 213
- o odwzorowaniu dla miar, 21
- o prawdopodobieństwie jako częstości, 56
- o procesach ciągłych, 165
- o produkcie miar
 - nieskończonym, 57
 - o przeliczalnej całkowalności, 30
 - o równoważności
 - zbieżności miar i dystrybuant, 22
 - o regularności miar borelowskich, 20
 - o rozkładzie jedyńki, 197
 - o transformacie odwrotnej II, 151
 - o uniwersalności testu runs, 117
 - o własnościach
 - funkcji charakterystycznej, 142
 - o zależności całki od dziedziny, 34
 - o zamianie zmiennych, 27
 - w $\mathbb{R}^n$, 28
- Radona–Nikodyma, 36
 - dowód, cz. I, 39
 - dowód, cz. II, 40
- Riesza o reprezentacji, 208
- Stone’a, 194
- strukturalne dla σ –ciał, 9
- Tonellego, 32
- uzupełnienie przestrzeni z miarą, 14
- wariancja, 60
 - całkowita, wektora losowego, 81
 - empiryczna, 99
- wartość oczekiwana, 60
 - warunkowa, 74, 75
- warunek
 - Carathéodory’ego, 15
 - Stone’a, 23, 193
- wektor losowy, 79
 - typu Gaussa, 138
- własność lokalizacji miary, 12
- współczynnik korelacji, 78
- wzór na całkowanie przez części, 33
- wzór Stirlinga, 130
- zasada niezmienniczości Donskera, 177
- zbieżność
 - ciągu dystrybuant, słaba, 22
 - ciągu miar borelowskich, 20
 - prawie na pewno, 64
 - według rozkładu, 144, 155
 - wg. prawdopodobieństwa, 64
- zbiór
 - χ –pomijalny, 187
 - ciągłości miary borelowskiej, 21
 - mierzalny, 5
 - nieosiągalny, 13

- półskończony, 12
- σ -skończony, 11
- typu $0/\infty$, 11
- w sensie Carathéodory'ego, 15
- permutowalny, 83
- pod wykresem, 32
- zwarty typu G_δ , 203
- zdarzenia niezależne, 67
- zdarzenie, 56
 - resztowe, 69
- zgodna rodzina miar borelowskich, 158
- zmienna losowa, 59
 - M -wartościowa, 154
 - o rozkładzie dyskretnym, 63
 - rozkład, 59
 - uogólniona, 59
 - wektorowa, 79
 - typu Gaussa, 138
- zmienne losowe
 - nieskorelowane, 78
 - niezależne, 70

Bibliografia

- [1] Andrzejczak G., (2019), *Analiza I, Materiał do wykładu*, Politechnika Łódzka, Łódź.
- [2] Andrzejczak G., (2020), *Konstrukcje I. Wybrane konstrukcje matematyki teoretycznej: Topologia, teoria miary i całka Lebesgue'a*, Politechnika Łódzka, Łódź.
- [3] Billingsley P., (2009), *Prawdopodobieństwo i miara*, PWN, Warszawa (wydanie oryginalne: Billingsley P., (1995), *Probability and Measure*, Wiley-Interscience, New York-Chichester-Brisbane-Toronto-Singapore).
- [4] Billingsley P., (1999), *Convergence of Probability Measures*, Second Ed., Wiley-Interscience, New York-Chichester-Brisbane-Toronto-Singapore.
- [5] Dieudonné J., (1976), *Treatise On Analysis*, Vol. II, Academic Press, Manchester (England).
- [6] Durrett R., (2010), *Probability. Theory and Examples*, Fourth Ed., Cambridge.
- [7] Franklin J.N., *Deterministic Simulation of Random Processes*, "Mathematics of Computation" 1963, vol. 17.
- [8] Gut A., (2005), *Probability: A Graduate Course*, Springer, Uppsala.
- [9] Halmos P.R., (1974, repr. z 1950), *Measure Theory*, Springer-Verlag, New York-Heidelberg-Berlin.

- [10] Hewitt E., Stromberg K., (1975, repr. z 1965), *Real and Abstract Analysis*, Springer-Verlag, New York–Heidelberg–Berlin.
- [11] Hewitt E., Ross K.A., (1979, wyd. I – 1963), *Abstract Harmonic Analysis*, Vol. I, Springer-Verlag, New York–Heidelberg–Berlin.
- [12] Jakubowski J., Sztencel R., (2001), *Wstęp do teorii prawdopodobieństwa*, wyd. II, SCRIPT, Warszawa.
- [13] eener R.W., (2010), *Theoretical Statistics, Topics for a Core Course*, Springer, New York–Dordrecht–Heilberg–Londyn.
- [14] Knuth D.E., (1998), *The Art of Computer Programming, Vol. 2, Seminumerical Algorithms*, wyd. 3, Addison–Wesley.
- [15] Maurin K., (2010), *Analiza I*, PWN, Warszawa.
- [16] Parthasarathy K.R., (1967), *Probability Measures on Metric Spaces*, Academic Press, New York.
- [17] Saeki S., *A Proof of the Existence of Infinite Product Probability Measures*, "The Amerian Mathematical Monthly" 1996, vol. 103, no. 8., October.
- [18] Stroock D.W., (2011), *Essentials of Integration Theory for Analysis*, Springer, New York–Dordrecht–Heidelberg–London.
- [19] Treves F., (1967), *Topological Vector Spaces, Distributions and Kernels*, Academic Press, New York.