Rachunek Prawdopodobieństwa 3

Rafał Latała

3 marca 2015

Poniższe notatki powstały na podstawie wykładu monograficznego z Rachunku Prawdopodobieństwa 3, prowadzonego w semestrze zimowym 2014/15. Celem wykładu było przedstawienie faktów dotyczących sum niezależnych zmiennych losowych, które znajdują się w szeroko pojętym kanonie współczesnej probabilistyki, a dla których zabrakło miejsca w kursowych wykładach z rachunku prawdopodobieństwa. Główny nacisk starano się położyć na przypadek jednowymiarowy, ale pewna część wyników dotyczy sum wektorów losowych.

Oczywiście w semestralnym wykładzie nie sposób zmieścić za wielu tematów. Ich wybór był po części kwestią gustu i zainteresowań badawczych prowadzącego, po części chęcią niepowielania materiału wykładów prowadzonych na Wydziale MIM UW w ostatnim czasie. Dlatego zabrakło tu tak ważnych pojęć jak rozkłady stabilne i nieskończenie podzielne, twierdzenia graniczne i obszary przyciągania, zasada niezmienniczości Donskera. Czytelników zainteresowanych tymi zagadnieniami oraz pogłębieniem wiedzy w tych tu zawartych pozostaje odesłać do monografii [3, 8]. Ze swej strony pragnę gorąco polecić znakomitą książkę Kallenberga [5]. Wyniki dotyczące sum wektorów losowych o wartościach w przestrzeni Banacha można znaleźć w monografii Ledoux i Talagranda [7], a wyczerpujące wprowadzenie w tematykę wielkich odchyleń w książce Dembo i Zeitouniego [2].

U słuchaczy wykładu zakładano dobrą znajomość dwusemestralnego, kursowego wykładu z rachunku prawdopodobieństwa. Wszystkie potrzebne fakty można znaleźć w doskonałych podręcznikach Jakubowskiego i Sztencla [4] oraz Billingsleya [1].

Pani Marcie Strzeleckiej i panu Krzysztofowi Ciosmakowi dziękuję za wychwycenie licznych usterek we wcześniejszych wersjach skryptu. Przepraszam za wszystkie nieścisłości i omyłki mogące nadal pojawiać się w tekście i jednocześnie zwracam się z prośbą do Czytelników, którzy zauważyli błędy lub mają jakieś inne uwagi na temat notatek o kontakt mailowy na adres `rlatala@mimuw.edu.pl` z podaniem wersji notatek (daty) do której chcą się ustosunkować.

Spis treści

1 Wielkie Odchylenia **3**
1.1 Podstawowe definicje. Nierówność Chernoffa . . . 4
1.2 Oszacowania z dołu . . . 6
1.3 Twierdzenie Cramera na $\mathbb{R}$ . . . 9
1.4 Twierdzenie Cramera na $\mathbb{R}^d$ . . . 12

2 Nierówności Wykładnicze. **17**
2.1 Nierówności dla ograniczonych przyrostów . . . 17
2.2 Nierówności typu Bernsteina . . . 18
2.3 Nierówność Bennetta . . . 23

3 Nierówności maksymalne **24**
3.1 Nierówności Levy'ego i Levy'ego-Ottavianiego . . . 25
3.2 Nierówności dla sum zmiennych o jednakowym rozkładzie . . . 28

4 Mocne prawa wielkich liczb **30**
4.1 Prawo Wielkich Liczb Kołmogorowa . . . 30
4.2 Prawo Wielkich Liczb Marcinkiewicza-Zygmunda . . . 30
4.3 Przypadek niejednakowych rozkładów. . . . 33

5 Prawo Iterowanego Logarytmu **34**
5.1 Przypadek jednowymiarowy . . . 34
5.2 Zbiór graniczny . . . 40
5.3 Przypadek wektorowy . . . 41

6 Błądzenia losowe **45**
6.1 Twierdzenie Hewitta-Savage'a. Mocna własność Markowa. . . . 46
6.2 Błądzenia chwilowe i powracające . . . 47
6.3 Punkty drabinowe. Faktoryzacja Wienera-Hopfa . . . 53

7 Elementy teorii odnowienia **58**
7.1 Stacjonarne procesy odnowienia . . . 59
7.2 Twierdzenie odnowienia . . . 61
7.3 Równanie odnowienia. Asymptotyka rozwiązania . . . 66
7.4 Rozkłady arytmetyczne . . . 67

1 Wielkie Odchylenia

W tej części, jeśli nie napiszemy inaczej będziemy zakładać, że $X, X_1, X_2, \ldots$ jest ciągiem niezależnych zmiennych losowym o jednakowym rozkładzie. Definiujemy

$$S_n := X_1 + X_2 + \ldots + X_n, \quad \bar{S}_n := \frac{S_n}{n}.$$

Ze słabego prawa wielkich liczb wiemy, że $\mathbf{P}(|\bar{S}_n - \mathbf{E}X| > \varepsilon) \to 0$ przy $n \to \infty$. Możemy zapytać jak szybka jest zbieżność. Z nierówności Czebyszewa dostajemy

$$\mathbf{P}(|\bar{S}_n - \mathbf{E}X| > \varepsilon) \leqslant \frac{\mathrm{Var}(X)}{n\varepsilon^2}.$$

Jednak oszacowanie to jest dalekie od optymalnego dla „porządnych" zmiennych X. Przykładowo, gdy $X \sim \mathcal{N}(0,1)$, to $\bar{S}_n \sim \mathcal{N}(0, 1/n) \sim X/\sqrt{n}$, zatem wykorzystując oszacowanie z poniżej sformułowanego Faktu 1.1,

$$\mathbf{P}(|\bar{S}_n| \geqslant \varepsilon) = 2\mathbf{P}(X \geqslant \varepsilon\sqrt{n}) \sim \frac{2}{\sqrt{2\pi n}\varepsilon} \exp(-n\varepsilon^2/2).$$

Fakt 1.1. *Dla dowolnego $t > 0$,*

$$(t^{-1} - t^{-3})e^{-t^2/2} \leqslant \int_t^\infty e^{-s^2/2} ds \leqslant t^{-1} e^{-t^2/2}.$$

Dowód. Mamy $(s^{-1}\exp(-s^2/2))' = -(1+s^{-2})\exp(-s^2/2)$ oraz $((s^{-1}-s^{-3})\exp(-s^2/2))' = -(1-3s^{-4})\exp(-s^2/2)$, stąd

$$\begin{aligned}(t^{-1} - t^{-3})e^{-t^2/2} = \int_t^\infty (1 - 3s^{-4}) e^{-s^2/2} ds &\leqslant \int_t^\infty e^{-s^2/2} ds \\ &\leqslant \int_t^\infty (1 + s^{-2}) e^{-s^2/2} ds = t^{-1} e^{-t^2/2}.\end{aligned}$$

□

Nasuwają się zatem następujące pytania:

1) Czy oszacowanie $\mathbf{P}(|\bar{S}_n - \mathbf{E}X| > \varepsilon)$ rzędu $\exp(-nf(\varepsilon))$ zachodzi dla szerszej klasy zmiennych losowych?

2) Czy istnieje granica $\lim_{n\to\infty} n^{-1} \log \mathbf{P}(\bar{S}_n \in A)$ dla ogólnej klasy zmiennych X i zbiorów A?

Na te pytania poszukamy odpowiedzi w kolejnych sekcjach.

1.1 Podstawowe definicje. Nierówność Chernoffa

Definicja 1.2. *Funkcję tworzącą momenty (transformatę Laplace'a)* zmiennej X definiujemy wzorem

$$M_X(t) := \mathbf{E}e^{tX}, \quad t \in \mathbb{R}.$$

Przykłady.
a) Dla $X \sim \mathcal{N}(a, \sigma^2)$, $M_X(t) = \exp(at + \sigma^2 t^2/2)$.
b) Dla $X = g^3$, gdzie $g \sim \mathcal{N}(0,1)$, $M_X(t) = \infty$ dla $t \neq 0$ i $M_X(0) = 1$.

Zachodzi oczywisty fakt.

Fakt 1.3. *Jeśli X i Y są niezależnymi zmiennymi losowymi, to $M_{X+Y} = M_X M_Y$. W szczególności, $M_{S_n} = (M_X)^n$.*

Łatwe szacowanie z nierówności Czebyszewa daje dla $s > 0$,

$$\begin{aligned}\mathbf{P}(\bar{S}_n \geqslant s) = \mathbf{P}(tS_n \geqslant stn) &\leqslant e^{-snt}\mathbf{E}e^{tS_n} = e^{-snt}(M_X(t))^n \\ &= \exp(-n(st - \log M_X(t))).\end{aligned}$$

Biorąc infimum prawej strony po wszystkich $t > 0$ dostajemy

$$\mathbf{P}(\bar{S}_n \geqslant s) \leqslant \exp\Big(- n \sup_{t \geqslant 0}(st - \log M_X(t))\Big). \tag{1}$$

Motywuje to następującą definicję

Definicja 1.4. Dla zmiennej losowej X określamy

$$\Lambda_X(t) := \log M_X(t), \quad D_{\Lambda_X} := \{t \in \mathbb{R} \colon \Lambda_X(t) < \infty\}$$

oraz

$$\Lambda_X^*(s) := \sup\{st - \Lambda_X(t) \colon t \in \mathbb{R}\}.$$

Przykłady
a) Dla $X \sim \mathcal{N}(a, \sigma^2)$, $\Lambda_X(t) = at + \sigma^2 t^2/2$, $D_{\Lambda_X} = \mathbb{R}$ oraz $\Lambda_X^*(t) = (t-a)^2/(2\sigma^2)$.
b) Dla $X = g^3$, $g \sim \mathcal{N}(0,1)$, $D_{\Lambda_X} = \{0\}$ oraz $\Lambda_X^* \equiv 0$.
c) Dla $X \sim \mathrm{Bin}(n,p)$, $D_{\Lambda_X} = \mathbb{R}$, $\Lambda_X(t) = n\log(pe^t + 1 - p)$ oraz

$$\Lambda_X^*(t) = \begin{cases} t\log(\frac{t}{np}) + (n-t)\log(\frac{n-t}{n(1-p)}) & \text{dla } t \in (0,n), \\ -n\log p & \text{dla } t = n, \\ -n\log(1-p) & \text{dla } t = 0, \\ \infty & \text{dla } t \notin [0,n]. \end{cases}$$

d) Dla $X \sim \mathrm{Poiss}(\lambda)$, $D_{\Lambda_X} = \mathbb{R}$, $\Lambda_X(t) = \lambda(e^t - 1)$ oraz

$$\Lambda_X^*(t) = \begin{cases} t\log(t/\lambda) - t + \lambda & \text{dla } t > 0, \\ \lambda & \text{dla } t = 0, \\ \infty & \text{dla } t < 0. \end{cases}$$

e) Dla $X \sim \mathrm{Exp}(\lambda)$, $D_{\Lambda_X} = (-\infty, \lambda)$, $\Lambda_X(t) = \log(\lambda/(\lambda - t))$ dla $t < \lambda$ oraz

$$\Lambda_X^*(t) = \begin{cases} \lambda t - 1 - \log(\lambda t) & \text{dla } t > 0, \\ \infty & \text{dla } t \leqslant 0. \end{cases}$$

$\Lambda_X^*(t) = \lambda t - 1 - \log(\lambda t)$.

Lemat 1.5. *Funkcje* $\Lambda_X \colon \mathbb{R} \to (-\infty, \infty]$ *oraz* $\Lambda_X^* \colon \mathbb{R} \to [0, \infty]$ *są wypukłe.*

Dowód. Wypukłość funkcji Λ_X wynika z nierówności Höldera. Funkcja Λ_X^* jest wypukła jako supremum funkcji wypukłych (liniowych). Ponadto $\Lambda_X^*(t) \geqslant 0t - \Lambda_X(0) = 0$. □

Przed sformułowaniem kolejnego lematu przypomnijmy konwencję, że wartość oczekiwana X da się określić za pomocą wzoru $\mathbf{E}X := \mathbf{E}X_+ - \mathbf{E}X_-$, jeśli tylko $\mathbf{E}X_+ < \infty$ lub $\mathbf{E}X_- < \infty$.

Lemat 1.6. *i) Jeśli* $D_{\Lambda_X} = \{0\}$, *to* $\Lambda_X^* \equiv 0$. *Jeśli* $\Lambda_X(t) < \infty$ *dla pewnego* $t > 0$, *to* $\mathbf{E}X_+ < \infty$, *a jeśli* $\Lambda_X(t) < \infty$ *dla pewnego* $t < 0$, *to* $\mathbf{E}X_- < \infty$.
ii) Jeśli $\mathbf{E}X_+ < \infty$ *lub* $\mathbf{E}X_- < \infty$, *to* $st - \Lambda_X(t) \leqslant 0$ *dla* $s \geqslant \mathbf{E}X, t \leqslant 0$ *lub* $s \leqslant \mathbf{E}X, t \geqslant 0$. *W szczególności,*

$$\Lambda_X^*(s) = \sup\{st - \Lambda_X(t) \colon t \geqslant 0\} \text{ dla } s \geqslant \mathbf{E}X \tag{2}$$

oraz

$$\Lambda_X^*(s) = \sup\{st - \Lambda_X(t) \colon t \leqslant 0\} \text{ dla } s \leqslant \mathbf{E}X.$$

iii) Jeśli $\mathbf{E}|X| < \infty$, *to* $\Lambda_X^*(\mathbf{E}X) = 0$. *Ponadto, dla dowolnego* X, $\inf_t \Lambda_X^*(t) = 0$.

Dowód. i) Dla $t > 0$ mamy $e^{tx} \geqslant tx_+$, więc skończoność $\mathbf{E}e^{tX}$ implikuje $\mathbf{E}X_+ < \infty$. Podobnie dla $t < 0$, $e^{tx} > -tx_-$, a zatem $\mathbf{E}X_- < \infty$, jeśli $\mathbf{E}e^{tX} < \infty$.

ii) Z nierówności Jensena,

$$\Lambda_X(t) = \log \mathbf{E}e^{tX} \geqslant \mathbf{E} \log e^{tX} = t\mathbf{E}X,$$

zatem $st - \Lambda_X(t) \leqslant (s - \mathbf{E}X)t \leqslant 0$ dla $s \geqslant \mathbf{E}X, t \leqslant 0$ lub $s \leqslant \mathbf{E}X, t \geqslant 0$.

iii) Jeśli $\mathbf{E}|X| < \infty$, to z udowodnionego wyżej oszacowania $\Lambda_X(t) \geqslant t\mathbf{E}X$ i

$$0 \leqslant \Lambda_X^*(\mathbf{E}X) \leqslant \sup_t (t\mathbf{E}X - t\mathbf{E}X) = 0.$$

Przypadek $\mathbf{E}|X| = \infty$ podzielimy na trzy podprzypadki:
a) $\mathbf{E}X_+ = \mathbf{E}X_- = \infty$. Wówczas $D_{\Lambda_X} = \{0\}$ i $\Lambda_X^* \equiv 0$.
b) $\mathbf{E}X_+ < \infty = \mathbf{E}X_-$, tzn. $\mathbf{E}X = -\infty$. Dla $s \in \mathbb{R}$ dostajemy (ostatnia równość wynika z (2))

$$\log \mathbf{P}(X \geqslant s) \leqslant \inf_{t \geqslant 0} \log \mathbf{E} \exp(t(X - s)) = \inf_{t \geqslant 0} (\Lambda_X(t) - ts) = -\Lambda_X^*(s),$$

czyli $0 \leqslant \Lambda_X^*(s) \leqslant -\log \mathbf{P}(X \geqslant s) \to 0$ przy $s \to -\infty$.
c) $\mathbf{E}X_+ = \infty > \mathbf{E}X_-$. Jak w b) dowodzimy, że $\Lambda_X^*(s) \to 0$ przy $s \to \infty$. □

Twierdzenie 1.7 (Nierówność Chernoffa)**.** *Jeśli* $\mathbf{E}X_+ < \infty$ *lub* $\mathbf{E}X_- < \infty$*, to*

$$\mathbf{P}(\bar{S}_n \geqslant s) \leqslant \exp(-n\Lambda_X^*(s)) \quad \text{dla } s \geqslant \mathbf{E}X$$

oraz

$$\mathbf{P}(\bar{S}_n \leqslant s) \leqslant \exp(-n\Lambda_X^*(s)) \quad \text{dla } s \leqslant \mathbf{E}X.$$

Dowód. Pierwsza nierówność wynika z (1) i Lematu 1.6 ii). Drugą nierówność dowodzimy analogicznie. □

Przykłady.
a) Dla $X \sim \mathcal{N}(a, \sigma^2)$ nierówność Chernoffa daje

$$\mathbf{P}(\bar{S}_n \geqslant t) \leqslant \exp\left(-\frac{n(t-a)^2}{2\sigma^2}\right) \qquad \text{dla } t \geqslant a$$

oraz

$$\mathbf{P}(\bar{S}_n \leqslant t) \leqslant \exp\left(-\frac{n(t-a)^2}{2\sigma^2}\right) \qquad \text{dla } t \leqslant a.$$

W rzeczywistości mamy $\bar{S}_n \sim \mathcal{N}(a, \sigma^2/n) \sim a + \sigma g/\sqrt{n}$, gdzie $g \sim \mathcal{N}(0,1)$, zatem np dla $t > a$

$$\mathbf{P}(\bar{S}_n \geqslant t) = \mathbf{P}\left(g \geqslant \frac{\sqrt{n}}{\sigma}(t-a)\right) \sim \frac{\sigma}{\sqrt{2\pi n}(t-a)} \exp\left(-\frac{n(t-a)^2}{2\sigma^2}\right).$$

b) Dla $X = g^3$, $g \sim \mathcal{N}(0,1)$, nierówność Chernoffa daje tylko trywialne oszacowanie $\mathbf{P}(\bar{S}_n \geqslant t) \leqslant 1$.
c) Jeśli $\mathbf{P}(X = 1) = p = 1 - \mathbf{P}(X = 0)$, to $S_n \sim \mathrm{Bin}(n,p)$ i Twierdzenie 1.7 implikuje klasyczne oszacowanie Chernoffa dla ogonów rozkładu dwumianowego:

$$\mathbf{P}(\mathrm{Bin}(n,p) \geqslant nt) \leqslant \left(\frac{p}{t}\right)^{nt} \left(\frac{1-p}{1-t}\right)^{n(1-t)} \qquad \text{dla } t \geqslant p.$$

1.2 Oszacowania z dołu

Widzieliśmy, że bez dodatkowych założeń nie możemy liczyć na odwrócenie nierówności Chernoffa. Daje się to jednak (w pewnym stopniu) zrobić przy założeniach o skończoności funkcji tworzącej momenty. Dlatego w tej części będziemy zazwyczaj dodatkowo zakładać, że

$$\mathbf{E}e^{tX} < \infty \text{ dla wszystkich } t \in \mathbb{R}. \tag{3}$$

Lemat 1.8. *Przy założeniu (3) funkcja* M_X *przedłuża się do funkcji analitycznej na całej płaszczyźnie zespolonej. Ponadto* $M_X^{(n)}(t) = \mathbf{E}(X^n e^{tX})$ *dla* $n = 1, 2, \ldots$.

Dowód. Funkcja $f(z) := \mathbf{E}\exp(zX)$ jest dobrze określona dla $z \in \mathbb{C}$, bo $\mathbf{E}|\exp(zX)| = \mathbf{E}\exp(\mathrm{Re}(z)X) < \infty$. Oczywiście f jest przedłużeniem M_X, nietrudno też wykazać (korzystając z twierdzenia Lebesgue'a o zbieżności zmajoryzowanej), że f jest różniczkowalna i $f'(z) = \mathbf{E}(X\exp(zX))$. Podobnie indukcyjnie dowodzimy, że $f^{(n)}(z) = \mathbf{E}(X^n\exp(zX))$. $\square$

Definicja 1.9. Dla $t \in \mathbb{R}$ definiujemy miarę probabilistyczną μ_t na $\mathbb{R}$ wzorem $d\mu_t(x) = M_X(t)^{-1}e^{tx}d\mu_X(x)$, tzn.

$$\mu_t(A) = \frac{1}{M_X(t)}\mathbf{E}(e^{tX}\mathbb{1}_{\{X\in A\}}).$$

Uwaga 1.10. Dla dowolnej funkcji f całkowalnej względem μ_t mamy

$$\mathbf{E}_{\mu_t}(f) := \int_{\mathbb{R}} f(x)\,\mathrm{d}\mu_t(x) = \frac{1}{M_X(t)}\mathbf{E}(f(X)e^{tX}).$$

Lemat 1.11. *Miara probabilistyczna μ_t spełnia warunki*

$$\mathbf{E}_{\mu_t}(x) = \int_{\mathbb{R}} x d\mu_t(x) = \Lambda'_X(t),$$

$$\mathrm{Var}_{\mu_t}(x) = \int_{\mathbb{R}} x^2 d\mu_t(x) - \left(\int_{\mathbb{R}} x d\mu_t(x)\right)^2 = \Lambda''_X(t).$$

Dowód. Zauważmy, że z Lematu 1.8, $M_X^{(n)}(t) = M_X(t)\mathbf{E}_{\mu_t}(x^n)$, więc

$$\Lambda'_X(t) = M_X(t)^{-1}M'_X(t) = \mathbf{E}_{\mu_t}(x)$$

oraz

$$\Lambda'_X(t) = M_X(t)^{-1}M''_X(t) - M_X(t)^{-2}(M'_X(t))^2 = \mathbf{E}_{\mu_t}(x^2) - (\mathbf{E}_{\mu_t}(x))^2 = \mathrm{Var}_{\mu_t}(x).$$

$\square$

Określmy

$$\alpha := \mathrm{essinf}X = \inf\{t\colon \mathbf{P}(X < t) > 0\},$$
$$\beta := \mathrm{esssup}X = \sup\{t\colon \mathbf{P}(X > t) > 0\}.$$

Oczywiście $\mathbf{P}(X \in [\alpha, \beta]) = 1$.

Lemat 1.12. *Dla dowolnej zmiennej losowej X,*

$$\Lambda^*_X(t) = \infty \quad \text{dla } t \notin [\alpha, \beta],$$
$$\mathbf{P}(X = \alpha) = \exp(-\Lambda^*_X(\alpha)), \quad \text{jeśli } \alpha > -\infty,$$
$$\mathbf{P}(X = \beta) = \exp(-\Lambda^*_X(\beta)), \quad \text{jeśli } \beta < \infty.$$

Dowód. Dla $t \leqslant 0$ mamy $\mathbf{P}(tX \leqslant \alpha t) = 1$, stąd $\Lambda_X(t) \leqslant \alpha t$ i dla $s < \alpha$,

$$\Lambda_X^*(s) \geqslant \sup\{st - \alpha t \colon t \leqslant 0\} = \infty.$$

Podobnie, dla $t \geqslant 0$, $\mathbf{P}(tX \leqslant \beta t) = 1$, zatem $\Lambda_X(t) \leqslant \beta t$ i dla $s > \beta$,

$$\Lambda_X^*(s) \geqslant \sup\{st - \beta t \colon t \geqslant 0\} = \infty.$$

Dla dowolnego $t \in \mathbb{R}$, $\Lambda_X(t) \geqslant \log(e^{t\beta}\mathbf{P}(X = \beta))$, czyli

$$\mathbf{P}(X = \beta) \leqslant \inf_t \exp(-t\beta + \Lambda_X(t)) = \exp(-\sup_t(t\beta - \Lambda_X(t)) = \exp(-\Lambda_X^*(\beta)).$$

Z drugiej strony, na mocy twierdzenia Lebesgue'a o zbieżności zmajoryzowanej,

$$\exp(-\Lambda_X^*(\beta)) \leqslant \exp(\Lambda_X(t) - t\beta) = \mathbf{E}\exp(t(X - \beta)) \to \mathbf{P}(X = \beta) \text{ dla } t \to \infty.$$

Zatem $\mathbf{P}(X = \beta) = \exp(-\Lambda_X^*(\beta))$. Tożsamość $\mathbf{P}(X = \alpha) = \exp(-\Lambda_X^*(\alpha))$ dowodzimy analogicznie. □

Lemat 1.13. *Załóżmy, że zachodzi warunek (3) oraz $\alpha < \beta$. Wówczas Λ_X' jest funkcją ściśle rosnącą z $\mathbb{R}$ na (α, β). Przekształcenie $\Theta := (\Lambda_X')^{-1} \colon (\alpha, \beta) \to \mathbb{R}$ spełnia*

$$\Lambda_X^*(s) = s\Theta(s) - \Lambda_X(\Theta(s)) \text{ dla } s \in (\alpha, \beta).$$

Dowód. Miara μ_X jest niezdegenerowana, tzn. $\mu_X(\{x\}) \neq 1$ dla $x \in \mathbb{R}$, więc miary μ_t też są niezdegenerowane. Zatem $\mathrm{Var}_{\mu_t}(x) > 0$ i z Lematu 1.11, $\Lambda_X''(t) > 0$ czyli Λ_X' jest ściśle rosnąca i funkcja $\Theta = (\Lambda_X')^{-1}$ jest dobrze określona na $\Lambda_X'(\mathbb{R})$. Oczywiście dla dowolnego t,

$$\Lambda_X'(t) = \frac{\int_\alpha^\beta xe^{tx}\,\mathrm{d}\mu_X(x)}{\int_\alpha^\beta e^{tx}\,\mathrm{d}\mu_X(x)} \in (\alpha, \beta),$$

więc $\Lambda_X'(\mathbb{R}) \subset (\alpha, \beta)$. Wykażemy, że zachodzi też odwrotna inkluzja.

Ustalmy $t \in [\mathbf{E}X, \beta)$ oraz $u \in (t, \beta)$. Dla dowolnego $s \geqslant 0$,

$$\Lambda_X(s) \geqslant \log(e^{su}\mathbf{P}(X \geqslant u)) = su + \log \mathbf{P}(X \geqslant u),$$

zatem

$$st - \Lambda_X(s) \leqslant s(t - u) - \log \mathbf{P}(X \geqslant u) < 0 \text{ dla } s > \frac{-\log \mathbf{P}(X \geqslant u)}{u - t}.$$

Z Lematu 1.6 ii), $st - \Lambda_X(s) \leqslant 0$ dla $s \leqslant 0$, zatem supremum funkcji $f(s) = st - \Lambda_X(s)$ jest osiągane w pewnym punkcie $s_0 \in [0, -(u-t)^{-1}\log \mathbf{P}(X \geqslant u)]$. W szczególności $f'(s_0) = 0$ czyli $t = \Lambda_X'(s_0)$. Wykazaliśmy więc, że $\Lambda'(\mathbb{R}) \supset [\mathbf{E}X, \beta)$ oraz $\Lambda_X^*(t) = t\Theta(t) - \Lambda_X(\Theta(t))$ dla $t \in [\mathbf{E}X, \beta)$.

Przypadek $t \in (\alpha, \mathbf{E}X]$ rozważamy analogicznie. □

Twierdzenie 1.14. *Załóżmy, że zachodzi (3). Wówczas dla $a \in (\alpha, \beta)$ oraz $\varepsilon > 0$,*

$$\mathbf{P}(|\bar{S}_n - a| < \varepsilon) \geqslant \left(1 - \frac{\Lambda''_X(\Theta(a))}{n\varepsilon^2}\right) \exp(-n(\Lambda^*_X(a) + \varepsilon|\Theta(a)|)),$$

gdzie $\Theta := (\Lambda'_X)^{-1}$.

Dowód. Zdefiniujmy $t := \Theta(a)$, wówczas z Lematu 1.13, $\Lambda^*_X(a) = at - \Lambda_X(t)$. Niech $Y, Y_1, \ldots, Y_n$ będą niezależnymi zmiennymi losowymi o rozkładzie zadanym przez miarę probabilistyczną μ_t. Na mocy Lematu 1.11, $\mathbf{E}Y = \Lambda'_X(t) = a$ i $\mathrm{Var}(Y) = \Lambda''_X(t)$. Zauważmy, że dla dowolnej nieujemnej funkcji mierzalnej f na $\mathbb{R}^n$,

$$\mathbf{E}f(Y_1, \ldots, Y_n) = (M_X(t))^{-n}\mathbf{E}\left[f(X_1, \ldots, X_n)e^{tS_n}\right].$$

Zatem

$$\begin{aligned}\mathbf{P}\left(\left|\frac{1}{n}\sum_{i=1}^n Y_i - a\right| < \varepsilon\right) &= (M_X(t))^{-n}\mathbf{E}e^{tS_n}\mathbb{1}_{\{|\bar{S}_n - a| < \varepsilon\}} \\ &\leqslant (M_X(t))^{-n}e^{n(ta+|t|\varepsilon)}\mathbf{P}(|\bar{S}_n - a| < \varepsilon).\end{aligned}$$

Z nierówności Czebyszewa

$$\mathbf{P}\left(\left|\frac{1}{n}\sum_{i=1}^n Y_i - a\right| \geqslant \varepsilon\right) \leqslant \frac{\mathrm{Var}(Y)}{n\varepsilon^2} = \frac{\Lambda''_X(t)}{n\varepsilon^2},$$

zatem

$$\begin{aligned}\mathbf{P}(|\bar{S}_n - a| < \varepsilon) &\geqslant \mathbf{P}\left(\left|\frac{1}{n}\sum_{i=1}^n Y_i - a\right| < \varepsilon\right)\exp(-n(ta - \Lambda_X(t) + |t|\varepsilon)) \\ &\geqslant \left(1 - \frac{\Lambda''_X(t)}{n\varepsilon^2}\right)\exp(-n(\Lambda^*_X(a) + |t|\varepsilon)).\end{aligned}$$

□

1.3 Twierdzenie Cramera na $\mathbb{R}$

Jesteśmy wreszcie gotowi do sformułowania twierdzenia o wielkich odchyleniach w przypadku jednowymiarowym. W odróżnieniu od poprzedniego paragrafu nie stawiamy żadnych wstępnych założeń o rozkładzie X.

Twierdzenie 1.15 (Prawo Wielkich Odchyleń Cramera)**.**
i) Dla dowolnego zbioru domkniętego $F \subset \mathbb{R}$,

$$\limsup_{n\to\infty} \frac{1}{n}\log \mathbf{P}(\bar{S}_n \in F) \leqslant -\inf_{x\in F}\Lambda^*_X(x).$$

ii) Dla dowolnego zbioru otwartego $G \subset \mathbb{R}$,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbf{P}(\bar{S}_n \in G) \geqslant -\inf_{x\in G} \Lambda_X^*(x).$$

iii) Dla dowolnego zbioru $\Gamma \in \mathcal{B}(\mathbb{R})$,

$$-\inf_{x\in \mathrm{int}(\Gamma)} \Lambda_X^*(x) \leqslant \liminf_{n\to\infty} \frac{1}{n} \log \mathbf{P}(\bar{S}_n \in \Gamma)$$
$$\leqslant \limsup_{n\to\infty} \frac{1}{n} \log \mathbf{P}(\bar{S}_n \in \Gamma) \leqslant -\inf_{x\in \mathrm{cl}(\Gamma)} \Lambda_X^*(x).$$

Dowód. i) Niech $I_F := \inf_{x\in F} \Lambda_X^*(x)$. Jeśli $I_F = 0$, to teza jest oczywista, możemy zatem zakładać, że $I_F > 0$. Wtedy $D_{\Lambda_X} \neq \{0\}$, czyli $\mathbf{E}X_+ < \infty$ lub $\mathbf{E}X_- < \infty$. Rozważymy trzy przypadki.
a) $\mathbf{E}|X| < \infty$. Ponieważ $\Lambda_X^*(\mathbf{E}X) = 0$ więc $\mathbf{E}X \notin F$. Niech (a,b) będzie składową $\mathbb{R} \setminus F$ zawierającą $\mathbf{E}X$. Jeśli $a > -\infty$, to $a \in F$, $\Lambda_X^*(a) \geqslant I_F$, podobnie jeśli $b < \infty$, to $b \in F$, $\Lambda_X^*(b) \geqslant I_F$. Mamy zatem z nierówności Chernoffa,

$$\mathbf{P}(\bar{S}_n \in F) \leqslant \mathbf{P}(\bar{S}_n \leqslant a) + \mathbf{P}(\bar{S}_n \geqslant b) \leqslant e^{-n\Lambda_X^*(a)} + e^{-n\Lambda_X^*(b)} \leqslant 2e^{-nI_F},$$

skąd $n^{-1}\log \mathbf{P}(\bar{S}_n \in F) \leqslant -I_F + n^{-1}\log 2 \to -I_F$ przy $n \to \infty$.
b) $\mathbf{E}X = -\infty$, wówczas Λ_X^* na mocy (2) jest niemalejąca, ponadto $\Lambda_X^*(x) \to 0$ przy $x \to -\infty$. Zatem warunek $I_F > 0$ implikuje $b = \inf F > -\infty$. Z domkniętości $b \in F$, czyli $\Lambda_X^*(b) \geqslant I_F$ oraz

$$\mathbf{P}(\bar{S}_n \in F) \leqslant \mathbf{P}(\bar{S}_n \geqslant b) \leqslant e^{-n\Lambda_X^*(b)} \leqslant e^{-nI_F}.$$

c) $\mathbf{E}X = \infty$. Rozumujemy analogicznie jak w b).

ii) Musimy wykazać, że warunek $x \in G$ implikuje $\liminf n^{-1}\log \mathbf{P}(\bar{S}_n \in G) \geqslant -\Lambda_X^*(x)$. Zmienna $Y = X - x$ spełnia $\Lambda_Y(t) = \Lambda_X(t) - xt$, $\Lambda_Y^*(s) = \Lambda_X^*(s + x)$. Zatem rozpatrując $X - x$ i $G - x$ zamiast X i G widzimy, że możemy zakładać, iż $x = 0$. Z otwartości G, mamy $(-\delta, \delta) \subset G$ dla pewnego $\delta > 0$. Wystarczy zatem wykazać

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbf{P}(\bar{S}_n \in (-\delta,\delta)) \geqslant -\Lambda_X^*(0) = \inf_{t\in\mathbb{R}} \Lambda_X(t). \tag{4}$$

Dowód (4) rozbijemy na kilka przypadków.

Przypadek I. $X \geqslant 0$ p.n. Mamy wówczas z twierdzenia Lebesgue'a o zbieżności zmajoryzowanej,

$$\inf_t \Lambda_X(t) \leqslant \lim_{t\to-\infty} \Lambda_X(t) = \log \lim_{t\to-\infty} \mathbf{E}e^{tX} = \log \mathbf{E} \lim_{t\to-\infty} e^{tX} = \log \mathbf{P}(X = 0).$$

Ponadto $\mathbf{P}(\bar{S}_n \in (-\delta,\delta)) \geqslant \mathbf{P}(\bar{S}_n = 0) = \mathbf{P}(X = 0)^n$ i natychmiast dostajemy (4).

Przypadek II. $X \leqslant 0$ p.n. Dowodzimy w taki sam sposób jak w przypadku I.

Przypadek III. $\operatorname{essinf} X < 0 < \operatorname{esssup} X$ oraz $D_{\Lambda_X} = \mathbb{R}$. Z Twierdzenia 1.14 (z $a = 0$) dostajemy dla $0 < \varepsilon < \delta$,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbf{P}(\bar{S}_n \in (-\delta,\delta)) \geqslant \liminf_{n\to\infty} \frac{1}{n} \log \mathbf{P}(|\bar{S}_n| \leqslant \varepsilon) \geqslant -\Lambda_X^*(0) - \varepsilon|\Theta(0)|.$$

Biorąc $\varepsilon \to 0+$ dostajemy (4).

Przypadek IV. $\operatorname{essinf} X < 0 < \operatorname{esssup} X$ i X dowolne.

Wybierzmy $M > 0$ na tyle duże by $\mathbf{P}(X \in [-M,0)), \mathbf{P}(X \in (0,M]) > 0$. Niech Y będzie miało taki rozkład jak zmienna X pod warunkiem $|X| \leqslant M$, tzn.

$$\mathbf{P}(Y \in A) = \mathbf{P}(X \in A, |X| \leqslant M)/\mathbf{P}(|X| \leqslant M),$$

zaś $Y_1, Y_2, \ldots$ będą niezależnymi kopiami Y oraz $\bar{T}_n = (Y_1 + \ldots + Y_n)/n$. Zauważmy, że dla dowolnej funkcji mierzalnej $f\colon \mathbb{R}^n \to [0,\infty)$ mamy

$$\begin{aligned}\mathbf{E} f(X_1,\ldots,X_n) &\geqslant \mathbf{E} f(X_1,\ldots,X_n) \mathbb{1}_{\{|X_1|\leqslant M,\ldots,|X_n|\leqslant M\}} \\ &= \mathbf{P}(|X| \leqslant M)^n \mathbf{E} f(Y_1,\ldots,Y_n).\end{aligned}$$

W szczególności,

$$\mathbf{P}(\bar{S}_n \in (-\delta,\delta)) \geqslant \mathbf{P}(|X| \leqslant M)^n \mathbf{P}(\bar{T}_n \in (-\delta,\delta)).$$

Zmienna Y jest ograniczona, więc na mocy przypadku III,

$$\begin{aligned}\liminf_{n\to\infty} \frac{1}{n} \log \mathbf{P}(\bar{S}_n \in (-\delta,\delta)) &\geqslant \log \mathbf{P}(|X| \leqslant M) + \liminf_{n\to\infty} \frac{1}{n} \log \mathbf{P}(\bar{T}_n \in (-\delta,\delta)) \\ &\geqslant \log \mathbf{P}(|X| \leqslant M) + \inf_t \Lambda_Y(t) = \inf_t \Lambda^M(t),\end{aligned}$$

gdzie $\Lambda^M(t) := \log \mathbf{E} e^{tX} \mathbb{1}_{\{|X|\leqslant M\}}$. Zauważmy, że funkcja Λ^M rośnie po M, więc $\inf_t \Lambda^M(t)$ zbiega przy $M \to \infty$ do pewnego a. Wykazaliśmy, że

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbf{P}(\bar{S}_n \in (-\delta,\delta)) \geqslant a,$$

więc wystarczy wykazać, że $a \geqslant \inf_t \Lambda_X(t)$. Rozpatrzmy zbiory

$$K_M := \{t\colon \Lambda^M(t) \leqslant a\}.$$

Funkcja $\Lambda^M(t)$ zbiega do nieskończoności przy $|t| \to \infty$, więc zbiory K_M tworzą zstępującą rodzinę zbiorów zwartych, ponadto są one niepuste, bo $a \geqslant \inf \Lambda^M = \Lambda^M(t_M)$ dla pewnego t_M. Istnieje więc punkt t_0 należący do przecięcia wszystkich K_M, ale wtedy $\Lambda^M(t_0) \leqslant a$ dla wszystkich M czyli, z twierdzenia Lebesgue'a o zbieżności monotonicznej, $\Lambda_X(t_0) = \lim_{M\to\infty} \Lambda^M(t_0) \leqslant a$.

Punkt iii) wynika z i) i ii).

$\square$

Uwaga 1.16. Analiza pierwszej części dowodu pokazuje, że wykazaliśmy

$$\mathbf{P}(\bar{S}_n \in \Gamma) \leqslant 2\exp\left(-n \inf_{x\in \mathrm{cl}(\Gamma)} \Lambda_X^*(x)\right) \qquad \text{dla dowolnego } \Gamma \in \mathcal{B}(\mathbb{R}).$$

1.4 Twierdzenie Cramera na $\mathbb{R}^d$

W tej części będziemy zakładać, że $X, X_1, \ldots$ są niezależnymi d-wymiarowymi wektorami losowymi o jednakowym rozkładzie.

Definicja 1.17. Dla d-wymiarowego wektora losowego X określamy dla $t \in \mathbb{R}^d$,

$$M_X(t) := \mathbf{E}e^{\langle t,X\rangle}, \quad \Lambda_X(t) := \log M_X(t)$$

oraz dla $s \in \mathbb{R}^d$,

$$\Lambda_X^*(s) := \sup\{\langle s,t\rangle - \Lambda_X(t) \colon t \in \mathbb{R}^d\}.$$

Współrzędne wektorów X (odp. X_i) będziemy oznaczać $X^{(j)}$ (odp. $X_i^{(j)}$) i definiujemy $\bar{S}_n^{(j)} := n^{-1}\sum_{i=1}^n X_i^{(j)}$.

Będziemy zakładać, że

$$M_X(t) < \infty \text{ dla wszystkich } t \in \mathbb{R}^d. \tag{5}$$

Lemat 1.18. *Przy założeniu warunku (5),*
i) Funkcja $\Lambda_X \colon \mathbb{R}^d \to \mathbb{R}$ *jest wypukła, klasy* C^∞ *oraz*

$$\frac{\partial}{\partial t_i} M_X(t) = \mathbf{E}(X^{(i)} e^{\langle t,X\rangle}).$$

ii) Funkcja $\Lambda_X^* \colon \mathbb{R}^d \to [0,\infty]$ *jest wypukła.*
iii) Jeśli $s = \nabla\Lambda_X(t)$ *dla pewnych* $s,t \in \mathbb{R}^d$, *to* $\Lambda_X^*(s) = \langle t,s\rangle - \Lambda(t)$.

Dowód. i) Wypukłość Λ_X wynika z nierówności Höldera jak w przypadku jednowymiarowym. By udowodnić gładkość, indukcyjnie po $|\alpha| = \alpha_1 + \ldots + \alpha_d$, dowodzimy, że dla $\alpha \in \mathbb{Z}_+^d$,

$$\frac{\partial^{\alpha_1}}{\partial t_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_d}}{\partial t_d^{\alpha_d}} M_X(t) = \mathbf{E}((X^{(1)})^{\alpha_1} \cdots (X^{(d)})^{\alpha_d} e^{\langle t,X\rangle}).$$

ii) Funkcja Λ_X^* jest wypukła jako supremum funkcji liniowych. Nieujemność Λ_X^* wynika z szacowania $\Lambda_X^*(t) \geqslant \langle 0,t\rangle - \Lambda_X(0) = 0$.

iii) Ustalmy $u \in \mathbb{R}^d$ i określmy

$$g(h) := h\langle u-t,s\rangle - \Lambda_X(t+h(u-t)) + \langle t,s\rangle, \quad h \in \mathbb{R}.$$

Wypukłość Λ_X implikuje wklęsłość g, zatem

$$g(1) - g(0) \leqslant \lim_{h \to 0+} \frac{g(h) - g(0)}{h} = \langle u - t, s \rangle - \langle u - t, \nabla \Lambda_X(t) \rangle = 0.$$

Dostaliśmy więc $g(1) \leqslant g(0)$, czyli $\langle u, s \rangle - \Lambda_X(u) \leqslant \langle t, s \rangle - \Lambda_X(t)$, co po wzięciu supremum po $u \in \mathbb{R}^d$ daje tezę. □

Twierdzenie 1.19. *Załóżmy, że $M_X(t) < \infty$ dla wszystkich $t \in \mathbb{R}^d$. Wówczas*
i) Dla dowolnego zbioru domkniętego $F \subset \mathbb{R}^d$,

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbf{P}(\bar{S}_n \in F) \leqslant - \inf_{x \in F} \Lambda_X^*(x).$$

ii) Dla dowolnego zbioru otwartego $G \subset \mathbb{R}^d$,

$$\liminf_{n \to \infty} \frac{1}{n} \log \mathbf{P}(\bar{S}_n \in G) \geqslant - \inf_{x \in G} \Lambda_X^*(x).$$

iii) Dla dowolnego zbioru $\Gamma \in \mathcal{B}(\mathbb{R}^d)$,

$$\begin{aligned} - \inf_{x \in \mathrm{int}(\Gamma)} \Lambda_X^*(x) &\leqslant \liminf_{n \to \infty} \frac{1}{n} \log \mathbf{P}(\bar{S}_n \in \Gamma) \\ &\leqslant \limsup_{n \to \infty} \frac{1}{n} \log \mathbf{P}(\bar{S}_n \in \Gamma) \leqslant - \inf_{x \in \mathrm{cl}(\Gamma)} \Lambda_X^*(x). \end{aligned}$$

Dowód. i) Zauważmy, że $2\delta - \inf_{s \in F}(\Lambda_X^*(s) \wedge M) \to - \inf_{s \in F} \Lambda_X^*(s)$ przy $\delta \to 0+, M \to \infty$, więc wystarczy iż wykażemy

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbf{P}(\bar{S}_n \in F) \leqslant 2\delta - \inf_{s \in F}(\Lambda_X^*(s) \wedge M) \quad \text{dla } \delta > 0, M < \infty. \tag{6}$$

Ustalmy $\delta > 0$ i $M < \infty$, dowód (6) podzielimy na dwa kroki.

Krok I. $F \subset \mathbb{R}^d$ zwarty.

Dla $s \in F$ niech $t_s \in \mathbb{R}^d$ będzie takie, że

$$\langle s, t_s \rangle - \Lambda_X(t_s) \geqslant \Lambda_X^*(s) \wedge M - \delta$$

oraz $r_s > 0$ takie, że $r_s |t_s| \leqslant \delta$. Przez $B(s, r)$ będziemy oznaczać kulę otwartą o środku w s i promieniu r. Zauważmy, że dla dowolnego zbioru $\Gamma \in \mathcal{B}(\mathbb{R}^d)$,

$$\mathbf{P}(\bar{S}_n \in \Gamma) = \mathbf{E} \mathbb{1}_{\{\bar{S}_n \in \Gamma\}} \leqslant \mathbf{E} \exp\left(\langle t, \bar{S}_n \rangle - \inf_{x \in \Gamma} \langle t, x \rangle \right).$$

Ponadto

$$\inf_{x \in B(s, r_s)} \langle t_s, x \rangle = \langle t_s, s \rangle + \inf_{x \in B(s, r_s)} \langle t_s, x - s \rangle = \langle t_s, s \rangle - r_s |t_s| \geqslant \langle t_s, s \rangle - \delta.$$

Zatem

$$\mathbf{P}(\bar{S}_n \in B(s,r_s)) \leqslant \mathbf{E}\exp(n\langle t_s, \bar{S}_n\rangle)\exp\left(-\inf_{x\in B(s,r_s)} n\langle t_s, x\rangle\right)$$
$$\leqslant \exp(n(\Lambda_X(t_s) - \langle t_s, s\rangle + \delta)) \leqslant \exp(n(2\delta - \Lambda_X^*(s)\wedge M)).$$

Zbiór F jest zwarty więc istnieje skończony podzbiór $S \subset F$ taki, że $F \subset \bigcup_{s\in S} B(s,r_s)$ i wówczas

$$\mathbf{P}(\bar{S}_n \in F) \leqslant \sum_{s\in S}\mathbf{P}(\bar{S}_n \in B(s,r_s)) \leqslant \sum_{s\in S}\exp(n(2\delta - \Lambda_X^*(s)\wedge M)),$$

zatem

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbf{P}(\bar{S}_n \in F) \leqslant 2\delta - \min_{s\in S}(\Lambda_X^*(s)\wedge M) \leqslant 2\delta - \inf_{s\in F}(\Lambda_X^*(s)\wedge M).$$

Krok II. F dowolny domknięty.

Dla $\rho > 0$ niech $K_\rho := [-\rho,\rho]^d$. Jeśli $\rho > \max_j \mathbf{E}|X^{(j)}|$, to z nierówności Chernoffa,

$$\mathbf{P}(\bar{S}_n \notin K_\rho) \leqslant \sum_{j=1}^d \mathbf{P}(|\bar{S}_n^{(j)}| > \rho) \leqslant \sum_{j=1}^d(\exp(-n\Lambda_{X^{(j)}}^*(\rho)) + \exp(-n\Lambda_{X^{(j)}}^*(-\rho))).$$

Ponieważ $\Lambda_{X^{(j)}}^*(s) \to \infty$ przy $|s| \to \infty$, więc dla odpowiednio dużego ρ, $\mathbf{P}(\bar{S}_n \notin K_\rho) \leqslant \exp(-nM)$, czyli

$$\mathbf{P}(\bar{S}_n \in F) \leqslant \mathbf{P}(\bar{S}_n \in F\cap K_\rho) + \mathbf{P}(\bar{S}_n \notin K_\rho) \leqslant \mathbf{P}(\bar{S}_n \in F\cap K_\rho) + \exp(-nM).$$

Zbiór $F\cap K_\rho$ jest zwarty, więc na mocy Kroku I,

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbf{P}(\bar{S}_n \in F\cap K_\rho) \leqslant 2\delta - \inf_{s\in F\cap K_\rho}(\Lambda_X^*(s)\wedge M),$$

a ponieważ $2\delta - \inf_{s\in F\cap K_\rho}(\Lambda_X^*(s)\wedge M) > -M$, więc

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbf{P}(\bar{S}_n \in F) \leqslant 2\delta - \inf_{s\in F\cap K_\rho}(\Lambda_X^*(s)\wedge M) \leqslant 2\delta - \inf_{s\in F}(\Lambda_X^*(s)\wedge M).$$

ii) Wystarczy iż wykażemy, że $\liminf_{n\to\infty}\frac{1}{n}\log\mathbf{P}(\bar{S}_n \in G) \geqslant -\Lambda_X^*(s)$ dla dowolnego $s \in G$. Ponieważ $B(s,\delta) \subset G$ dla pewnego $\delta > 0$, więc musimy jedynie wykazać, że

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbf{P}(\bar{S}_n \in B(s,\delta)) \geqslant -\Lambda_X^*(s) \quad \text{dla } \delta > 0.$$

Dowód przeprowadzimy w dwóch krokach.

Krok I. $s = \nabla\Lambda_X(t)$ dla pewnego $t \in \mathbb{R}^d$.

Na mocy Lematu 1.18 mamy wówczas $\Lambda_X^*(s) = \langle t, s\rangle - \Lambda_X(t)$. Określmy miarę probabilistyczną μ_t na $\mathbb{R}^d$ wzorem

$$d\mu_t(x) := \frac{1}{M_X(t)} e^{\langle x,t\rangle} d\mu_X(x),$$

czyli $d\mu_X(x) = \exp(-\langle x,t\rangle + \Lambda_X(t))d\mu_t(x)$. Niech $Y, Y_1, Y_2, \ldots$ będą niezależnymi wektorami losowymi o rozkładzie μ_t, $T_n := (Y_1 + \ldots + Y_n)$ oraz $\bar{T}_n := T_n/n$. Dla dowolnej nieujemnej funkcji mierzalnej f na $\mathbb{R}^d$ mamy

$$\mathbf{E}f(X_1,\ldots,X_n) = \exp(n\Lambda_X(t))\mathbf{E}f(Y_1,\ldots,Y_n)\exp(-\langle t, T_n\rangle),$$

zatem

$$\begin{aligned}
\mathbf{P}(\bar{S}_n \in B(s,\delta)) &= \exp(n\Lambda_X(t))\mathbf{E}\exp(-\langle t, T_n\rangle)\mathbb{1}_{\{\bar{T}_n\in B(s,\delta)\}} \\
&= \exp(-n(\langle t,s\rangle - \Lambda_X(t)))\mathbf{E}\exp(-n\langle t, \bar{T}_n - s\rangle)\mathbb{1}_{\{|T_n - s|<\delta)\}} \\
&\geqslant \exp(-n\Lambda_X^*(s) - n\delta|t|)\mathbf{P}(|T_n - s| < \delta).
\end{aligned}$$

Mamy jednak

$$\mathbf{E}Y = \frac{1}{M_X(t)}\mathbf{E}(Xe^{\langle t,X\rangle}) = \frac{\nabla M_X(t)}{M_X(t)} = \nabla\Lambda_X(t) = s,$$

czyli ze słabego prawa wielkich liczb,

$$\begin{aligned}
\liminf_{n\to\infty}\frac{1}{n}\log\mathbf{P}(\bar{S}_n \in B(s,\delta)) &\geqslant -\Lambda_X^*(s) - \delta|t| + \liminf_{n\to\infty}\frac{1}{n}\log\mathbf{P}(|\bar{T}_n - s| < \delta)) \\
&= -\Lambda_X^*(s) - \delta|t|.
\end{aligned}$$

Z monotoniczności $\delta \mapsto \mathbf{P}(\bar{S}_n \in B(s,\delta))$ otrzymujemy

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbf{P}(\bar{S}_n \in B(s,\delta)) \geqslant \lim_{\delta'\to 0+}\liminf_{n\to\infty}\frac{1}{n}\log\mathbf{P}(\bar{S}_n \in B(s,\delta')) \geqslant -\Lambda_X^*(s).$$

Krok II. X dowolna zmienna spełniająca (5).

Niech $G, G_1, G_2, \ldots$ oznaczają niezależne kanoniczne d-wymiarowe wektory gaussowskie (tzn. współrzędne G_i są niezależnymi zmiennymi $\mathcal{N}(0,1)$), niezależne od $X, X_1, X_2, \ldots$. Ustalmy $M > 0$ i zdefinujmy

$$Y := X + \frac{1}{\sqrt{M}}G, \quad Y_i := X_i + \frac{1}{\sqrt{M}}G_i, \quad R_n := \frac{G_1 + \ldots G_n}{n\sqrt{M}}$$

Wówczas $M_Y(t) = M_X(t)M_G(t/\sqrt{M}) = M_X(t)\exp(|t|^2/(2M))$, czyli $\Lambda_Y(t) = \Lambda_X(t) + |t|^2/(2M) \geqslant \Lambda_X(t)$ oraz $\Lambda_Y^*(s) \leqslant \Lambda_X^*(s)$ dla wszystkich s. Z nierówności Jensena, $\Lambda_X(t) = \log\mathbf{E}\exp(\langle t, X\rangle) \geqslant \mathbf{E}\langle t, X\rangle$, czyli $\Lambda_Y(t) \geqslant |t|^2/(2M) + \langle t, \mathbf{E}X\rangle$ oraz

$$\langle s,t\rangle - \Lambda_Y(t) \leqslant -|t|^2/(2M) + \langle t, s - \mathbf{E}X\rangle \to -\infty \text{ przy } |t| \to \infty,$$

a więc funkcja $t \to \langle s, t\rangle - \Lambda_Y(t)$ osiąga swoje supremum. Zatem istnieje t takie, że $s = \nabla\Lambda_Y(t)$. Mamy

$$\frac{Y_1 + \ldots Y_n}{n} = \bar{S}_n + R_n,$$

na mocy przypadku I otrzymujemy

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbf{P}(\bar{S}_n + R_n \in B(s, \delta/2)) \geqslant -\Lambda_Y^*(s) \geqslant -\Lambda_X^*(s).$$

Zmienna R_n ma ten sam rozkład, co $G/\sqrt{nM}$, więc

$$\mathbf{P}\left(|R_n| \geqslant \frac{\delta}{2}\right) = \mathbf{P}\left(|G| \geqslant \sqrt{nM}\frac{\delta}{2}\right) \leqslant \sum_{j=1}^{d} \mathbf{P}\left(|G^{(j)}| \geqslant \sqrt{nM}\frac{\delta}{2d}\right) \leqslant 2d \exp\left(-\frac{nM\delta^2}{8d^2}\right),$$

gdzie ostatnie szacowanie wynika stąd, że $G^{(j)} \sim \mathcal{N}(0,1)$ dla $1 \leqslant j \leqslant d$. Jeśli przyjmiemy $M := 8\delta^{-2}d^2(\Lambda_X^*(s) + 1)$, to dostaniemy

$$\mathbf{P}\left(|R_n| \geqslant \frac{\delta}{2}\right) \leqslant 2d\exp(-n(\Lambda_X^*(s)+1)) \leqslant \frac{1}{2}\mathbf{P}\left(\bar{S}_n + R_n \in B\left(s, \frac{\delta}{2}\right)\right)$$

dla dostatecznie dużych n. Zatem dla dużych n,

$$\begin{aligned}\mathbf{P}(\bar{S}_n \in B(s,\delta)) &\geqslant \mathbf{P}\left(\bar{S}_n + R_n \in B\left(s, \frac{\delta}{2}\right)\right) - \mathbf{P}\left(|R_n| \geqslant \frac{\delta}{2}\right)\\ &\geqslant \frac{1}{2}\mathbf{P}\left(\bar{S}_n + R_n \in B\left(s, \frac{\delta}{2}\right)\right),\end{aligned}$$

czyli

$$\liminf_{n\to\infty} \frac{1}{n}\log \mathbf{P}(\bar{S}_n \in B(s,\delta)) \geqslant \liminf_{n\to\infty} \frac{1}{n}\log \mathbf{P}(\bar{S}_n + R_n \in B(s, \delta/2)) \geqslant -\Lambda_X^*(s).$$

iii) jest oczywiście natychmiastową konsekwencją i) i ii).

□

Uwaga 1.20. Twierdzenie 1.19 zachodzi przy słabszym założeniu, że M_X jest skończone w pewnym otoczeniu zera. W przeciwieństwie do przypadku jednowymiarowego, nie jest ono jednak prawdziwe dla dowolnych zmiennych X, bez dodatkowych założeń o skończoności M_X.

2 Nierówności Wykładnicze.

Nierówność Chernoffa pokazuje, że jeśli $S = \sum_{i=1}^n X_i$ oraz zmienne X_i są niezależne o jednakowym rozkładzie, to $\mathbf{P}(S \geqslant s) \leqslant \exp(-n\Lambda_X^*(s/n))$. Jednak stosowanie tej nierówności w praktyce natrafia na pewne ograniczenia:
i) Nie zawsze da się dokładnie obliczyć Λ_X, a co za tym idzie Λ_X^*;
ii) Założenie o jednakowym rozkładzie X_i bywa często niewygodne.

By te trudności obejść zauważmy, że dowód nierówności Chernoffa pokazuje, że

$$\mathbf{P}(S \geqslant s) \leqslant \exp\left(-\sup_{t>0}(st - \Lambda_S(t))\right) \qquad \text{dla } s > 0.$$

Ponadto, jeśli będziemy umieli oszacować dla wszystkich i, Λ_{X_i} z góry, to da nam to górne ograniczenie dla $\Lambda_S = \sum_{i=1}^n \Lambda_{X_i}$ i za pośrednictwem powyższej nierówności ograniczymy też $\mathbf{P}(S \geqslant s)$. Metodę tę bedziemy stosować w kolejnych paragrafach.

Szacowania $\mathbf{P}(S \geqslant s)$ z dołu są z reguły dużo trudniejsze, co już wcześniej widzieliśmy przy próbie odwrócenia nierówności Chernoffa.

2.1 Nierówności dla ograniczonych przyrostów

W tym paragrafie pokażemy jakie szacowania można wyprowadzić, jeśli potrafimy oszacować poszczególne zmienne z góry.

Lemat 2.1. *Załóżmy, że X jest ograniczoną zmienną losową o średniej zero oraz $\|X\|_\infty \leqslant d < \infty$. Wówczas*

$$M_X(t) \leqslant \exp\left(\frac{1}{2}d^2t^2\right) \qquad \textit{dla wszystkich } t.$$

Dowód. Ponieważ $\Lambda_X(t) = \Lambda_{X/d}(dt)$ możemy zakładać, że $d = 1$. Dla $|x| \leqslant 1$ mamy $\frac{1-x}{2}(-t) + \frac{1+x}{2}t = tx$, więc z wypukłości $\exp(x)$,

$$e^{tx} \leqslant \frac{1-x}{2}e^{-t} + \frac{1+x}{2}e^t = \cosh(t) + x\sinh(t) \qquad \text{dla } |x| \leqslant 1.$$

Zatem, jeśli $\|X\|_\infty \leqslant 1$, to

$$\mathbf{E}\exp(tX) \leqslant \cosh(t) + \sinh(t)\mathbf{E}X = \cosh(t) \leqslant \exp(t^2/2).$$

□

Twierdzenie 2.2. *Załóżmy, że X_i są ograniczonymi, niezależnymi zmiennymi losowymi o średniej zero oraz $\|X_i\|_\infty \leqslant d_i < \infty$. Wówczas*

$$\mathbf{E}\exp\left(t\sum_{i=1}^n X_i\right) \leqslant \exp\left(\frac{t^2}{2}\sum_{i=1}^n d_i^2\right) \qquad \textit{dla } t \in \mathbb{R}$$

oraz
$$\mathbf{P}\left(\sum_{i=1}^n X_i \geqslant s\right) \leqslant \exp\left(-\frac{s^2}{2\sum_{i=1}^n d_i^2}\right) \qquad \text{dla } s \geqslant 0.$$

Dowód. Pierwsza nierówność wynika natychmiast z Lematu 2.1. By uzyskać drugą szacujemy dla $S = \sum_{i=1}^n X_i$,

$$\mathbf{P}(S \geqslant s) \leqslant \exp\left(-\sup_{t>0}(st - \Lambda_S(t))\right) \leqslant \exp\left(-\sup_{t>0}\left(st - \frac{t^2}{2}\sum_{i=1}^n d_i^2\right)\right)$$
$$= \exp\left(-\frac{s^2}{2\sum_{i=1}^n d_i^2}\right).$$

□

Wniosek 2.3. *Załóżmy, że* (ε_i) *jest ciągiem Bernouliego, tzn. ciągiem niezależnych zmiennych losowych takim, że* $\mathbf{P}(\varepsilon_i = \pm 1) = 1/2$*. Wówczas dla dowolnego ciągu* a_i *spełniającego* $\sum_i a_i^2 < \infty$ *zachodzi*
$$\mathbf{P}\left(\sum_i a_i \varepsilon_i \geqslant s\right) \leqslant \exp\left(-\frac{s^2}{2\sum_i a_i^2}\right).$$

Dowód. Prosty argument graniczny pokazuje, że wystarczy rozważać przypadek skończony, który jest natychmiastowym wnioskiem z Twierdzenia 2.2. □

Uwaga 2.4. Można udowodnić znacznie ogólniejsze twierdzenie – *nierówność Azumy.* Dla dowolnego martyngału $(M_k, \mathcal{F}_k)_{k=1}^n$ takiego, że $\|M_k - M_{k-1}\|_\infty \leqslant d_k < \infty$ zachodzi

$$\mathbf{P}(M_n - M_0 \geqslant s) \leqslant \exp\left(-\frac{s^2}{2\sum_{i=1}^n d_i^2}\right) \quad \text{dla } s \geqslant 0.$$

2.2 Nierówności typu Bernsteina

Nierówności z poprzedniego paragrafu są mało precyzyjne – działają dobrze tylko, gdy wariancja zmiennych jest zbliżona do kwadratu ich normy supremum. Jest to związane z tym, że w założeniach występowała tylko norma L^∞. W tym paragrafie pokażemy jak można jednocześnie wykorzystywać znajomość wariancji i normy supremum.

Lemat 2.5. *Załóżmy, że* X *jest zmienną losową o średniej zero taką, że istnieją* $\sigma^2, M < \infty$ *spełniające warunek*
$$\mathbf{E}|X|^k \leqslant \frac{k!}{2}\sigma^2 M^{k-2} \qquad \text{dla } k = 2, 3, \ldots.$$
Wówczas
$$\Lambda_X(t) \leqslant \frac{\sigma^2 t^2}{2(1 - M|t|)} \qquad \text{dla } M|t| < 1.$$

Dowód. Liczymy

$$M_X(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!}\mathbf{E}X^k \leqslant 1 + \sum_{k=2}^{\infty} \frac{|t|^k}{2}\sigma^2 M^{k-2} = 1 + \frac{t^2\sigma^2}{2}\sum_{k=2}^{\infty}(|t|M)^{k-2}$$
$$= 1 + \frac{\sigma^2 t^2}{2(1-M|t|)} \leqslant \exp\left(\frac{\sigma^2 t^2}{2(1-M|t|)}\right).$$

□

Twierdzenie 2.6 (Nierówność Bernsteina)**.** *Załóżmy, że X_i są niezależnymi zmiennymi losowymi o średniej zero, zaś $\sigma_i^2, M < \infty$ są takie, że*

$$\mathbf{E}|X_i|^k \leqslant \frac{k!}{2}\sigma_i^2 M^{k-2} \qquad \text{dla } i \geqslant 1,\ k \geqslant 2. \tag{7}$$

Wówczas

$$\mathbf{E}\exp\left(t\sum_{i=1}^{n} X_i\right) \leqslant \exp\left(\frac{t^2\sum_{i=1}^n \sigma_i^2}{2(1-M|t|)}\right) \qquad \text{dla } M|t| < 1$$

oraz dla $s > 0$,

$$\mathbf{P}\left(\sum_{i=1}^{n} X_i \geqslant s\right) \leqslant \exp\left(-\frac{s^2}{2\sum_{i=1}^n \sigma_i^2 + 2Ms}\right),$$
$$\mathbf{P}\left(\left|\sum_{i=1}^{n} X_i\right| \geqslant s\right) \leqslant 2\exp\left(-\frac{s^2}{2\sum_{i=1}^n \sigma_i^2 + 2Ms}\right).$$

Dowód. Niech $S := \sum_{i=1}^n X_i$, $\sigma^2 := \sum_{i=1}^n \sigma_i^2$, wówczas $\Lambda_S = \sum_i \Lambda_{X_i}$ i pierwsze oszacowanie wynika z Lematu 2.5. Dalej szacujemy

$$\mathbf{P}(S \geqslant s) \leqslant \exp\left(-\sup_{t>0}(st - \Lambda_S(t))\right) \leqslant \exp\left(-\sup_{0<t<M^{-1}}\left(st - \frac{t^2\sigma^2}{2(1-Mt)}\right)\right)$$
$$\leqslant \exp\left(-\frac{s^2}{2\sigma^2 + 2Ms}\right),$$

gdzie ostatnią nierówność dostajemy przyjmując $t = s(\sigma^2 + Ms)^{-1}$. Ponieważ zmienne $-X_i$ spełniają te same założenia co X_i, więc dostajemy dla $s < 0$,

$$\mathbf{P}(S \leqslant s) = \mathbf{P}(-S \geqslant -s) \leqslant \exp\left(-\frac{s^2}{2\sigma^2 + 2Ms}\right)$$

i z tożsamości $\mathbf{P}(|S| \geqslant s) = \mathbf{P}(S \geqslant s) + \mathbf{P}(S \leqslant -s)$ wynika ostatnia część tezy. □

Uwaga 2.7. Na mocy centralnego twierdzenia granicznego oraz Faktu 1.1 nie możemy się spodziewać lepszego oszacowania niż $\exp(-s^2/(2\sigma^2))$. Ponadto zmienne X_i o rozkładzie symetrycznym wykładniczym z parametrem 1 (tzn. zmienne z gęstością $\exp(-|x|)/2$) spełniają $\mathbf{E}|X_i|^k = k!$, czyli dla takich zmiennych zachodzą założenia Twierdzenia 2.6 z $\sigma_i^2 = 2, M = 1$. Pokazuje to, że nie możemy uzyskać szacowania lepszego niż $\exp(-s/M)$ przy $s \to \infty$.

Wniosek 2.8. *Załóżmy, że X_i są ograniczonymi, niezależnymi zmiennymi losowymi o średniej zero, wówczas dla $s > 0$,*

$$\mathbf{P}\left(\sum_{i=1}^n X_i \geqslant s\right) \leqslant \exp\left(-\frac{s^2}{2\sigma^2 + 2as/3}\right),$$

gdzie $\sigma^2 = \mathrm{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathbf{E}X_i^2$ oraz $a = \max_i \|X_i\|_\infty$.

Dowód. Mamy dla $k \geqslant 2$,

$$\mathbf{E}|X_i|^k \leqslant a^{k-2}\mathbf{E}X_i^2 \leqslant \frac{k!}{2}\left(\frac{a}{3}\right)^{k-2}\mathbf{E}X_i^2,$$

zatem warunek (7) jest spełniony z $M = a/3$ oraz $\sigma_i = \mathbf{E}X_i^2$. □

Uwaga 2.9. Nierówność Bernsteina pokazuje, iż jeśli $s \max_i \|X_i\|_\infty \leqslant 3\varepsilon \sum_i \mathbf{E}X_i^2$, to

$$\mathbf{P}\left(\sum_{i=1}^n X_i \geqslant s\right) \leqslant \exp\left(-\frac{s^2}{2(1+\varepsilon)\sigma^2}\right).$$

Kolejny wynik pokazuje jak odwrócić szacowanie z powyższej uwagi.

Twierdzenie 2.10 (Odwrotna nierówność wykładnicza Kołmogorowa)**.** *Dla $\varepsilon > 0$ istnieją stałe $K(\varepsilon)$ i $\delta(\varepsilon)$ takie, że dla ograniczonych, niezależnych zmiennych losowych X_i o średniej zero, $s > 0$ oraz $\sigma^2 := \mathrm{Var}(\sum_{i=1}^n X_i)$ zachodzi oszacowanie*

$$\mathbf{P}\left(\sum_{i=1}^n X_i \geqslant s\right) \geqslant \exp\left(-(1+\varepsilon)\frac{s^2}{2\sigma^2}\right),$$

o ile

$$s \max_i \|X_i\|_\infty \leqslant \delta(\varepsilon)\sigma^2 \quad \textit{oraz} \quad s \geqslant K(\varepsilon)\sigma.$$

Dowód. Ustalmy $\varepsilon > 0$ i niech $u = u(\varepsilon)$ będzie taką liczbą dodatnią dla której

$$1 - \Phi(u) \geqslant 2\exp(-\sqrt{1+\varepsilon}u^2/2).$$

Wybranie u jest możliwe na podstawie np. oszacowań Faktu 1.1.

Krok I. Istnieje stała $\gamma = \gamma(\varepsilon)$ taka, że jeśli ograniczone niezależne zmienne losowe X_i o sredniej zero spełniają

$$\max_{i \leqslant k} \|X_i\|_\infty \leqslant \gamma(\varepsilon) \left(\sum_{i=1}^k \mathbf{E} X_i^2 \right)^{1/2},$$

to

$$\mathbf{P} \left(\sum_{i=1}^k X_i \geqslant u \left(\sum_{i=1}^k \mathbf{E} X_i^2 \right)^{1/2} \right) \geqslant \frac{1}{2}(1 - \Phi(u)) \geqslant \exp\left(-\sqrt{1+\varepsilon} \frac{u^2}{2} \right).$$

Oczywiście wystarczy rozpatrywać przypadek $\sum_{i=1}^k \mathbf{E} X_i^2 = 1$. Załóżmy, że takie γ nie istnieje. Wówczas dla każdego n istnieją niezależne zmienne losowe $X_{n,1}, X_{n,2}, \ldots, X_{n,k_n}$ o średniej zero takie, że $\|X_{n,i}\|_\infty \leqslant n^{-1}$, $\sum_{i=1}^{k_n} \mathbf{E} X_{n,i}^2 = 1$ oraz

$$\mathbf{P} \left(\sum_{i=1}^{k_n} X_{n,i} \geqslant u \right) < \frac{1}{2}(1 - \Phi(u)).$$

Łatwo sprawdzić, że układ trójkątny $(X_{n,i})_{n,i}$ spełnia warunek Lindeberga, więc na mocy centralnego twierdzenia granicznego,

$$\lim_{n \to \infty} \mathbf{P} \left(\sum_{i=1}^{k_n} X_{n,i} \geqslant u \right) = 1 - \Phi(u),$$

co daje sprzeczność z poprzednią nierównością. Zatem teza Kroku I została wykazana.

Krok II. Dowód twierdzenia.

Bez straty ogólności możemy zakładać, że $\sum_{i=1}^n \mathbf{E} X_i^2 = 1$. Zdefiniujmy $n_0 = 0$ oraz indukcyjnie

$$n_j := \inf \left\{ m \colon \sum_{i=n_{j-1}+1}^{m} \mathbf{E} X_i^2 \geqslant \frac{u^2}{s^2 \sqrt{1+\varepsilon}} \right\}.$$

Konstrukcję przerywamy na takim n_k dla którego

$$\sum_{i=n_k+1}^{n} \mathbf{E} X_i^2 < \frac{u^2}{s^2 \sqrt{1+\varepsilon}},$$

wtedy oczywiście $k \leqslant u^{-2} s^2 \sqrt{1+\varepsilon}$. Określmy

$$S_j := \sum_{i=n_{j-1}+1}^{n_j} X_i \text{ dla } j = 1, \ldots, k-1$$

oraz

$$S_k := \sum_{i=n_{k-1}+1}^{n} X_i.$$

Wówczas $S_1+\ldots+S_k=\sum_{i=1}^n X_i$ oraz $\mathrm{Var}(S_j)\geqslant u^2/(s^2\sqrt{1+\varepsilon})$ dla $1\leqslant j\leqslant k$. Ponadto

$$\max_i \|X_i\|_\infty \leqslant s^{-1}\delta(\varepsilon)\leqslant \gamma(\mathrm{Var}(S_j))^{1/2},$$

jeśli $\delta(\varepsilon)\leqslant u\gamma(1+\varepsilon)^{-1/4}$. Zatem na mocy kroku I,

$$\mathbf{P}\Big(S_j\geqslant \frac{u^2}{s\sqrt[4]{1+\varepsilon}}\Big)\geqslant \mathbf{P}(S_j\geqslant u(\mathrm{Var}(S_j))^{1/2})\geqslant \exp\Big(-\sqrt{1+\varepsilon}\frac{u^2}{2}\Big).$$

oraz

$$\begin{aligned}\mathbf{P}\Big(S\geqslant k\frac{u^2}{s\sqrt[4]{1+\varepsilon}}\Big)&\geqslant \prod_{j=1}^k \mathbf{P}\Big(S_j\geqslant \frac{u^2}{s\sqrt[4]{1+\varepsilon}}\Big)\geqslant \exp\Big(-k\sqrt{1+\varepsilon}\frac{u^2}{2}\Big)\\ &\geqslant \exp(-(1+\varepsilon)s^2/2).\end{aligned}$$

Oszacujmy k z dołu. Mamy dla $j=1,\ldots,k$,

$$\sum_{i=n_{j-1}+1}^{n_j}\mathbf{E}X_i^2\leqslant \frac{u^2}{s^2\sqrt{1+\varepsilon}}+\mathbf{E}X_{n_j}^2\leqslant \frac{1}{s^2}\Big(\frac{u^2}{\sqrt{1+\varepsilon}}+\delta^2(\varepsilon)\Big),$$

co daje

$$1=\sum_{i=1}^n\mathbf{E}X_i^2=\sum_{j=1}^k\sum_{i=n_{j-1}+1}^{n_j}\mathbf{E}X_i^2+\sum_{i=n_k+1}^n\mathbf{E}X_i^2\leqslant \frac{k}{s^2}\Big(\frac{u^2}{\sqrt{1+\varepsilon}}+\delta^2(\varepsilon)\Big)+\frac{u^2}{s^2\sqrt{1+\varepsilon}},$$

czyli

$$k\geqslant \frac{s^2\sqrt{1+\varepsilon}-u^2}{u^2+\sqrt{1+\varepsilon}\delta^2(\varepsilon)}\geqslant \frac{s^2}{u^2}\sqrt[4]{1+\varepsilon},$$

o ile $\delta(\varepsilon)$ jest odpowiednio małe, a $K(\varepsilon)$ odpowiednio duże. I wtedy

$$\mathbf{P}(S\geqslant s)\geqslant \mathbf{P}\left(S\geqslant k\frac{u^2}{s\sqrt[4]{1+\varepsilon}}\right)\geqslant \exp\left(-(1+\varepsilon)\frac{s^2}{2}\right).$$

$\square$

Uwaga 2.11. Nierówność z powyższego twierdzenia można też równoważnie sformułować jako istnienie stałych $\delta'(\varepsilon)>0$ i $K'(\varepsilon)<\infty$ takich, że

$$\mathbf{P}\Big(\sum_{i=1}^n X_i\geqslant s\Big)\geqslant \frac{1}{K'(\varepsilon)}\exp\Big(-(1+\varepsilon)\frac{s^2}{2\sigma^2}\Big),$$

o ile

$$\max(s,\sigma)\max_i\|X\|_\infty\leqslant \delta'(\varepsilon)\sigma^2.$$

2.3 Nierówność Bennetta

Szacowanie podane we Wniosku 2.8 jest, z uwagi na centralne twierdzenie graniczne, bliskie optymalnego dla s małych. Jednak dla s dużych można je poprawić o czynnik logarytmiczy.

Lemat 2.12. *Załóżmy, że X jest zmienną losową o średniej zero, wariancji σ^2 oraz $\|X_i\|_\infty \leqslant a$. Wówczas*

$$\Lambda_X(t) \leqslant \frac{\sigma^2}{a^2}(e^{ta} - ta - 1) \quad \text{dla } t \geqslant 0.$$

Dowód. Liczymy

$$\mathbf{E}e^{tX} = 1 + t\mathbf{E}X + \sum_{k=2}^{\infty} \frac{t^k \mathbf{E}X^k}{k!} \leqslant 1 + \sigma^2 \sum_{k=2}^{\infty} \frac{t^k a^{k-2}}{k!} = 1 + \frac{\sigma^2}{a^2}(e^{ta} - ta - 1)$$

i teza wynika natychmiast z nierówności $\ln(1+x) \leqslant x$. □

Twierdzenie 2.13 (nierówność Bennetta)**.** *Załóżmy, że X_i są ograniczonymi niezależnymi zmiennymi losowymi o średniej zero, $\sigma^2 = \mathrm{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathbf{E}X_i^2$ oraz $a \geqslant \max_i \|X_i\|_\infty$. Wówczas dla $t > 0$,*

$$\mathbf{E}\exp\left(t\sum_{i=1}^n X_i\right) \leqslant \exp\left(\frac{\sigma^2}{a^2}(e^{ta} - ta - 1)\right)$$

oraz dla $s > 0$,

$$\mathbf{P}\left(\sum_{i=1}^n X_i \geqslant s\right) \leqslant \exp\left(-\frac{\sigma^2}{a^2} h\left(\frac{sa}{\sigma^2}\right)\right) \leqslant \exp\left(-\frac{s}{2a}\ln\left(1 + \frac{sa}{\sigma^2}\right)\right),$$

gdzie

$$h(x) := (1+x)\ln(1+x) - x.$$

Dowód. Pierwsza część wynika natychmiast z Lematu 2.12. By pokazać drugą zauważamy, że dla $S = \sum_{i=1}^n X_i$,

$$\mathbf{P}(S \geqslant s) \leqslant \exp\left(-\sup_{t>0}(st - \Lambda_S(t))\right) \leqslant \exp\left(-\sup_{t>0}\left(st - \frac{\sigma^2}{a^2}(e^{ta} - ta - 1)\right)\right).$$

Prosty rachunek pokazuje, że powyższe supremum jest osiągane w punkcie

$$t = \frac{1}{a}\ln\Big(1 + \frac{as}{\sigma^2}\Big)$$

i wynosi

$$\frac{\sigma^2}{a^2}\Big[\Big(1 + \frac{sa}{\sigma^2}\Big)\ln\Big(1 + \frac{sa}{\sigma^2}\Big) - \frac{sa}{\sigma^2}\Big] = \frac{\sigma^2}{a^2} h\Big(\frac{sa}{\sigma^2}\Big) \geqslant \frac{s}{2a}\ln\Big(1 + \frac{sa}{\sigma^2}\Big),$$

gdzie ostatnie oszacowanie wynika z poniższego lematu. □

Lemat 2.14. *Dla dowolnego $x \geqslant 0$,*

$$(1+x)\ln(1+x) - x \geqslant \frac{x}{2}\ln(1+x).$$

Dowód. Niech $f(x) = (1+x)\ln(1+x) - x - (x/2)\ln(1+x) = (1+x/2)\ln(1+x) - x$. Liczymy $f'(x) = (\ln(1+x) - x(1+x)^{-1})/2$, $f''(x) = x(1+x)^{-2}$, zatem $f(0) = f'(0) = 0$ oraz $f''(x) \geqslant 0$ dla $x \geqslant 0$. $\square$

Uwaga 2.15. Jeśli $\mathbf{P}(Y_{n,i} = 1) = 1 - \mathbf{P}(Y_{n,i} = 0) = 1/n$ oraz $Y_{n,i}$ są niezależne, to rozkład $\sum_{i=1}^{n} Y_{i,n}$ zbiega do rozkładu Poissona z parametrem 1. Biorąc $X_{n,i} = Y_{n,i} - 1/n$ mamy $\sum_{i=1}^{n} \mathbf{E}X_{i,n}^2 \leqslant 1$ oraz $\max_i \|X_{i,n}\|_\infty \leqslant 1$. To pokazuje, że przy założeniach Twierdzenia 2.13 nie można uzyskać przy $s \to \infty$ oszacowania lepszego rzędu niż $s \ln s$.

Uwaga 2.16. Nierówność Bennetta ma swoją wersję martyngałową. Mianowicie, dla martyngału $(M_k, \mathcal{F}_k)_{k=0}^n$ spełniającego warunki

$$\max_k \|M_k - M_{k-1}\|_\infty \leqslant a$$

i

$$\sum_{k=1}^{n} \|\mathbf{E}((M_k - M_{k-1})^2 | \mathcal{F}_{k-1})\|_\infty \leqslant \sigma^2,$$

zachodzi nierówność

$$\mathbf{P}(M_n - M_0 \geqslant s) \leqslant \exp\left(-\frac{\sigma^2}{a^2} h\left(\frac{sa}{\sigma^2}\right)\right) \leqslant \exp\left(-\frac{s}{2a}\ln\left(1 + \frac{sa}{\sigma^2}\right)\right).$$

3 Nierówności maksymalne

Przy dowodzeniu twierdzeń granicznych przydają się nierówności pozwalające szacować ogon maksimum sum zmiennych losowych przez ogony pojedynczych sum. Nierówności takie omówimy w bieżącym rozdziale.

Będziemy zakładać, że $(F, \|\cdot\|)$ jest ośrodkową przestrzenią Banacha, a X_i są wektorami losowymi o wartościach w F. Tradycyjnie definiujemy

$$S_n := \sum_{i=1}^{n} X_i.$$

Uwaga 3.1. Założenie ośrodkowości przestrzeni Banacha F ma charakter techniczny – implikuje, że suma funkcji mierzalnych o wartościach w F jest mierzalna. Jeśli interesuje nas tylko zachowanie normy zmiennych losowych, to możemy zakładać słabszy warunek – mianowicie, że norma w F da się przedstawić jako supremum przeliczalnej liczby funkcjonałów, wtedy norma sumy funkcji mierzalnych jest mierzalna.

3.1 Nierówności Levy'ego i Levy'ego-Ottavianiego

Twierdzenie 3.2 (Nierówność Levy'ego)**.** *Załóżmy, że X_i są niezależnymi symetrycznymi wektorami losowymi w F. Wówczas dla $t \geqslant 0$,*

$$\mathbf{P}\left(\max_{1\leqslant k\leqslant n}\|S_k\| \geqslant t\right) \leqslant 2\mathbf{P}(\|S_n\| \geqslant t).$$

Dowód. Niech

$$A_k := \{\|S_1\| < t, \ldots, \|S_{k-1}\| < t, \|S_k\| \geqslant t\}, \quad 1 \leqslant k \leqslant n.$$

Zauważmy, że dla dowolnych $x, y \in F$,

$$\max\{\|x+y\|, \|x-y\|\} \geqslant \frac{1}{2}\|x+y\| + \frac{1}{2}\|x-y\| \geqslant \|x\|,$$

zatem

$$A_k \subset (A_k \cap \{\|S_k + (S_n - S_k)\| \geqslant t\}) \cup (A_k \cap \{\|S_k - (S_n - S_k)\| \geqslant t\}).$$

Łączny rozkład zmiennych $(S_1, S_2, \ldots, S_k, S_n - S_k)$ jest taki sam jak łączny rozkład zmiennych $(S_1, S_2, \ldots, S_k, -(S_n - S_k))$, więc oba zbiory po prawej stronie powyższej inkluzji mają jednakowe prawdopodobieństwo, zatem

$$\mathbf{P}(A_k) \leqslant 2\mathbf{P}(A_k \cap \{\|S_n\| \geqslant t\}),$$

czyli

$$\mathbf{P}\left(\max_{1\leqslant k\leqslant n}\|S_k\| \geqslant t\right) = \sum_{k=1}^{n}\mathbf{P}(A_k) \leqslant 2\sum_{k=1}^{n}\mathbf{P}(A_k \cap \{\|S_n\| \geqslant t\}) \leqslant 2\mathbf{P}(\|S_n\| \geqslant t).$$

□

Wniosek 3.3. *Załóżmy, że zmienne X_i są niezależne i symetryczne oraz $|a_i| \leqslant 1$. Wówczas*

$$\mathbf{P}\left(\left\|\sum_{i=1}^{n} a_i X_i\right\| \geqslant t\right) \leqslant 2\mathbf{P}(\|S_n\| \geqslant t).$$

Dowód. Możemy zakładać, że $1 \geqslant a_1 \geqslant a_2 \geqslant \ldots \geqslant a_n \geqslant 0$, wówczas

$$\sum_{i=1}^{n} a_i X_i = a_n S_n + \sum_{k=1}^{n-1}(a_k - a_{k+1})S_k,$$

stąd łatwo wynika, że $\|\sum_{i=1}^{n} a_i X_i\| \leqslant \max_{1\leqslant k\leqslant n}\|S_k\|$ i teza wniosku wynika z nierówności Levy'ego. □

Uwaga 3.4. Jeśli założenia Wniosku 3.3 są spełnione, to dla dowolnej funkcji wypukłej $f\colon F \to \mathbb{R}_+$,

$$\mathbf{E}f\left(\sum_{i=1}^n a_i X_i\right) \leqslant \mathbf{E}f\left(\sum_{i=1}^n X_i\right),$$

w szczególności dla $p \geqslant 1$,

$$\mathbf{E}\left\|\sum_{i=1}^n a_i X_i\right\|^p \leqslant \mathbf{E}\left\|\sum_{i=1}^n X_i\right\|^p.$$

Nie jest jednak prawdą, że nierówność z Wniosku 3.3 zachodzi ze stałą 1.

Przejdźmy teraz do przypadku niesymetrycznego.

Lemat 3.5. *Dla dowolnych niezależnych zmiennych losowych X_i oraz $t, u \geqslant 0$,*

$$\mathbf{P}\left(\max_{1\leqslant k\leqslant n}\|S_k\| \geqslant t+u\right) \leqslant \frac{\mathbf{P}(\|S_n\| \geqslant t)}{1-\max_{1\leqslant k\leqslant n}\mathbf{P}(\|S_n - S_k\| \geqslant u)}, \tag{8}$$

o ile $\mathbf{P}(\|S_n - S_k\| \geqslant u) < 1$ dla $1 \leqslant k \leqslant n$.

Dowód. Zdefiniujmy

$$A_k := \{\|S_1\| < t+u, \ldots, \|S_{k-1}\| < t+u, \|S_k\| \geqslant t+u\}, \quad 1 \leqslant k \leqslant n.$$

Zbiory A_k są parami rozłączne oraz

$$A := \bigcup_{k=1}^n A_k = \left\{\max_{1\leqslant k\leqslant n}\|S_k\| \geqslant t+u\right\}.$$

Zauważmy, że $\|S_n\| \geqslant \|S_k\| - \|S_n - S_k\|$, zatem wykorzystując założenie o niezależności,

$$\mathbf{P}(\|S_n - S_k\| < u)\mathbf{P}(A_k) = \mathbf{P}(A_k \cap \{\|S_n - S_k\| < u\}) \leqslant \mathbf{P}(A_k \cap \{\|S_n\| \geqslant t\}).$$

Otrzymujemy

$$\begin{aligned}\min_{1\leqslant k\leqslant n}\mathbf{P}(\|S_n - S_k\| < u)\mathbf{P}(A) &\leqslant \sum_{k=1}^n \mathbf{P}(\|S_n - S_k\| < u)\mathbf{P}(A_k)\\ &\leqslant \sum_{k=1}^n \mathbf{P}(A_k \cap \{\|S_n\| \geqslant t\}) \leqslant \mathbf{P}(\|S_n\| \geqslant t),\end{aligned}$$

co po podzieleniu stronami przez $\min_{1\leqslant k\leqslant n}\mathbf{P}(\|S_n - S_k\| < u) = 1 - \max_{1\leqslant k\leqslant n}\mathbf{P}(\|S_n - S_k\| \geqslant u)$ daje tezę lematu. □

Twierdzenie 3.6 (Nierówność Levy'ego-Ottavianiego)**.** *Załóżmy, że X_i są niezależnymi wektorami losowymi w F. Wówczas dla $t \geqslant 0$,*

$$\mathbf{P}\left(\max_{1\leqslant k\leqslant n}\|S_k\| \geqslant 3t\right) \leqslant 3\max_{1\leqslant k\leqslant n}\mathbf{P}(\|S_k\| \geqslant t). \tag{9}$$

Dowód. Szacujemy

$$\mathbf{P}(\|S_n - S_k\| \geqslant 2t) \leqslant \mathbf{P}(\|S_n\| \geqslant t) + \mathbf{P}(\|S_k\| \geqslant t) \leqslant 2 \max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_k\| \geqslant t),$$

zatem

$$\max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_n - S_k\| \geqslant 2t) \leqslant 2 \max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_k\| \geqslant t).$$

Oczywiście, by udowodnić (9) możemy założyć, że $\max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_k\| \geqslant t) < 1/3$. Kładąc $u = 2t$ w (8) dostajemy

$$\begin{aligned}\mathbf{P}\Big(\max_{1\leqslant k\leqslant n} \|S_k\| \geqslant 3t\Big) &\leqslant \frac{\mathbf{P}(\|S_n\| \geqslant t)}{1 - \max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_n - S_k\| \geqslant 2t)} \leqslant \frac{\mathbf{P}(\|S_n\| \geqslant t)}{1 - 2\max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_k\| \geqslant t)} \\ &\leqslant 3\mathbf{P}(\|S_n\| \geqslant t).\end{aligned}$$

□

Uwaga 3.7. Nierówność (9) można nieco polepszyć. Zbigniew Szewczak wykazał niedawno, że

$$\mathbf{P}\left(\max_{1\leqslant k\leqslant n} \|S_k\| \geqslant 3t\right) \leqslant 2 \max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_k\| \geqslant t).$$

Dowód. Ustalmy $t, u \geqslant 0$ i określmy dla $1 \leqslant k \leqslant n$,

$$B_k := \{\|S_n - S_{n-j}\| < t + u \text{ dla } 1 \leqslant j \leqslant k-1,\ \|S_n - S_{n-k}\| \geqslant t + u\}.$$

Wówczas zbiory B_k są rozłączne oraz

$$\bigcup_{k=1}^{n} B_k = \left\{\max_{1\leqslant k\leqslant n} \|S_n - S_{n-k}\| \geqslant t + u\right\} = \left\{\max_{0\leqslant k< n} \|S_n - S_k\| \geqslant t + u\right\}.$$

Nierówność trójkąta implikuje, że $\{\|S_n\| < t,\ \|S_n - S_{n-k}\| \geqslant t + u\} \subset \{\|S_{n-k}\| \geqslant u\}$, ponadto zdarzenie B_k jest niezależne od zmiennej S_{n-k}, stąd

$$\begin{aligned}\mathbf{P}\left(\max_{0\leqslant k< n} \|S_n - S_k\| \geqslant t + u,\ \|S_n\| < t\right) &= \sum_{k=1}^{n} \mathbf{P}(B_k \cap \{\|S_n\| < t\}) \\ &\leqslant \sum_{k=1}^{n} \mathbf{P}(B_k \cap \{\|S_{n-k}\| \geqslant u\}) = \sum_{k=1}^{n} \mathbf{P}(B_k)\mathbf{P}(\|S_{n-k}\| \geqslant u) \\ &\leqslant \max_{1\leqslant k< n} \mathbf{P}(\|S_{n-k}\| \geqslant u) \sum_{k=1}^{n} \mathbf{P}(B_k) \leqslant \max_{1\leqslant k< n} \mathbf{P}(\|S_{n-k}\| \geqslant u).\end{aligned}$$

Powyższa nierówność oraz zawieranie

$$\left\{\max_{1\leqslant k\leqslant n} \|S_k\| \geqslant 2t + u\right\} \subset \{\|S_n\| \geqslant t\} \cup \left\{\max_{0\leqslant k< n} \|S_n - S_k\| \geqslant t + u,\ \|S_n\| < t\right\}$$

implikują

$$\mathbf{P}\left(\max_{1\leqslant k\leqslant n}\|S_k\| \geqslant 2t+u\right) \leqslant \mathbf{P}(\|S_n\| \geqslant t) + \max_{1\leqslant k<n}\mathbf{P}(\|S_{n-k}\| \geqslant u).$$

Biorąc $u = t$ dostajemy tezę. □

3.2 Nierówności dla sum zmiennych o jednakowym rozkładzie

W tym paragrafie będziemy dodatkowo zakładać, że

$$\text{Niezależne zmienne } X_i \text{ o wartościach w } F \text{ mają jednakowy rozkład.} \tag{10}$$

Dla wektora losowego X o wartościach w F i $s > 0$ określmy

$$A(X,t) = \inf_{x\in F}\mathbf{P}(\|X-x\| \geqslant t).$$

Lemat 3.8. *Jeśli X i Y są niezależne, to*

$$A(X+Y,t) \geqslant A(X,t) \quad \text{dla } t \geqslant 0.$$

Dowód. Dla dowolego $x \in F$ dostajemy

$$\mathbf{P}(\|X+Y-x\| \geqslant t) = \mathbf{E}_Y\mathbf{P}_X(\|X-(x-Y)\| \geqslant t) \geqslant \mathbf{E}_Y A(X,t) = A(X,t).$$

□

Lemat 3.9. *Załóżmy, że zachodzi (10) oraz $\mathbf{P}(\|S_n\| \geqslant t) < 1/4$. Wówczas*

$$\max_{1\leqslant k\leqslant n}\mathbf{P}(\|S_k\| \geqslant 5t) \leqslant \mathbf{P}(\|S_n\| \geqslant t).$$

Dowód. Ustalmy liczbę α taką, że $1/4 > \alpha > \mathbf{P}(\|S_n\| \geqslant t)$. Wówczas oczywiście $A(S_n,t) < \alpha$, czyli na mocy Lematu 3.8, $A(S_k,t) < \alpha$ dla $k = 1,\ldots,n$. Możemy zatem wybrać $x_1,\ldots,x_{n-1} \in F$ dla których

$$\mathbf{P}(\|S_k - x_k\| \geqslant t) < \alpha, \quad k = 1,\ldots,n-1.$$

Wybierzmy $i \leqslant n-1$ takie, że $\|x_i\| = \max_{1\leqslant j\leqslant n-1}\|x_j\|$. Wiemy, że $\mathbf{P}(\|S_i - x_i\| \geqslant t) < \alpha$ oraz $\mathbf{P}(\|S_n\| \geqslant t) < \alpha$, stąd

$$\mathbf{P}(\|S_{n-i}+x_i\| \geqslant 2t) = \mathbf{P}(\|S_n-(S_i-x_i)\| \geqslant 2t) \leqslant \mathbf{P}(\|S_i-x_i\| \geqslant t) + \mathbf{P}(\|S_n\| \geqslant t) < 2\alpha.$$

Rozpatrzymy trzy przypadki.

Przypadek I. $i = n-i$. Mamy

$$\mathbf{P}(\{\|S_i - x_i\| \geqslant t\} \cup \{\|S_{n-i}+x_i\| \geqslant 2t\}) < \alpha + 2\alpha < 1,$$

więc istnieje zdarzenie elementarne ω dla którego $\|S_i(\omega) - x_i\| < t$ oraz $\|S_i(\omega) + x_i\| = \|S_{n-i}(\omega) + x_i\| < 2t$. Na mocy nierówności trójkąta dostajemy $\|2x_i\| \leqslant 3t$.

Przypadek II. $i > n - i$. Mamy

$$\begin{aligned}\mathbf{P}(\|S_{2i-n} - 2x_i\| \geqslant 3t) &= \mathbf{P}(\|(S_i - x_i) - (S_i - S_{2i-n} + x_i)\| \geqslant 3t)\\ &\leqslant \mathbf{P}(\|S_i - x_i\| \geqslant t) + \mathbf{P}(\|S_i - S_{2i-n} + x_i\| \geqslant 2t)\\ &< \alpha + \mathbf{P}(\|S_{n-i} + x_i\| \geqslant 2t) < 3\alpha.\end{aligned}$$

Zatem

$$\mathbf{P}(\{\|S_{2i-n} - 2x_i\| \geqslant 3t\} \cup \{\|S_{2i-n} - x_{2i-n}\| \geqslant t\}) < 4\alpha < 1,$$

czyli istnieje ω takie, że $\|S_{2i-n}(\omega) - 2x_i\| < 3t$ oraz $\|S_{2i-n}(\omega) - x_{2i-n}\| < t$. Zatem $\|x_{2i-n} - 2x_i\| < 4t$. Stąd $\|x_i\| \geqslant \|x_{2i-n}\| \geqslant 2\|x_i\| - 4t$, czyli w tym przypadku $\|x_i\| \leqslant 4t$.

Przypadek III. $i < n - i$. Dostajemy

$$\begin{aligned}\mathbf{P}(\|S_{n-2i} + 2x_i\| \geqslant 3t) &= \mathbf{P}(\|(S_{n-i} + x_i) - (S_{n-i} - S_{n-2i} - x_i)\| \geqslant 3t)\\ &\leqslant \mathbf{P}(\|S_{n-i} + x_i\| \geqslant 2t) + \mathbf{P}(\|S_{n-i} - S_{n-2i} - x_i\| \geqslant t)\\ &< 2\alpha + \mathbf{P}(\|S_i - x_i\| \geqslant t) < 3\alpha.\end{aligned}$$

Tak samo jak w poprzednim przypadku pokazujemy, że $\|x_{n-2i} + 2x_i\| \leqslant 4t$ oraz $\|x_i\| \leqslant 4t$.

Udowodniliśmy zatem, że $\max_{1\leqslant j\leqslant n-1} \|x_j\| \leqslant 4t$, stąd

$$\max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_k\| \geqslant 5t) \leqslant \max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_k - x_k\| \geqslant t) < \alpha,$$

co, z dowolności $\alpha \in (\mathbf{P}(\|S_n\| \geqslant t), 1/4)$, dowodzi tezy. $\square$

Twierdzenie 3.10 (Montgomery-Smith)**.** *Przy założeniu (10),*

$$\mathbf{P}\Big(\max_{1\leqslant k\leqslant n} \|S_k\| \geqslant 6t\Big) \leqslant 4\mathbf{P}(\|S_n\| \geqslant t) \quad \textit{dla } t \geqslant 0.$$

Dowód. Oczywiście możemy zakładać, że $\mathbf{P}(\|S_n\| \geqslant t) < 1/4$. Wówczas na mocy Lematu 3.9, $\max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_k\| \geqslant 5t) \leqslant \mathbf{P}(\|S_n\| \geqslant t) < 1/4$. Z nierówności (8) (z $u = 5t$),

$$\begin{aligned}\mathbf{P}\left(\max_{1\leqslant k\leqslant n} \|S_k\| \geqslant 6t\right) &\leqslant \frac{\mathbf{P}(\|S_n\| \geqslant t)}{1 - \max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_n - S_k\| \geqslant 5t)}\\ &= \frac{\mathbf{P}(\|S_n\| \geqslant t)}{1 - \max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_{n-k}\| \geqslant 5t)} \leqslant \frac{4}{3}\mathbf{P}(\|S_n\| \geqslant t).\end{aligned}$$

$\square$

Wniosek 3.11. *Załóżmy, że zachodzi (10) oraz* $0 \leqslant a_i \leqslant 1$. *Wówczas*

$$\mathbf{P}\left(\left\|\sum_{i=1}^n a_i X_i\right\| \geqslant 6t\right) \leqslant 4\mathbf{P}(\|S_n\| \geqslant t) \quad \textit{dla } t \geqslant 0.$$

Dowód. Możemy zakładać, że $a_1 \geqslant a_2 \geqslant \ldots \geqslant a_n$ i wtedy, tak jak we Wniosku 3.3, dostajemy $\|\sum_{i=1}^n a_i X_i\| \leqslant \max_{1\leqslant k\leqslant n} \|S_k\|$ i teza łatwo wynika z Twierdzenia 3.10. $\square$

4 Mocne prawa wielkich liczb

W trakcie tego wykładu, jeśli nie powiemy inaczej, będziemy zakładali, że $X, X_1, X_2, \ldots$ są niezależnymi zmiennymi losowymi o wartościach w ośrodkowej przestrzeni Banacha $(F, \| \|)$. Tradycyjnie też $S_n = X_1 + \ldots + X_n$.

4.1 Prawo Wielkich Liczb Kołmogorowa

Okazuje się, że przeniesienie tradycyjnego sformułowania mocnego prawa wielkich na przypadek przestrzeni Banacha jest łatwe. Dzieje się tak dzięki następującemu lematowi.

Lemat 4.1. *Załóżmy, że X jest zmienną losową o wartościach w ośrodkowej przestrzeni Banacha F taką, że $\mathbf{E}\|X\| < \infty$. Wówczas dla dowolnego $\varepsilon > 0$ istnieje funkcja $\varphi\colon F \to F$, przyjmująca tylko skończenie wiele wartości, taka, że $\mathbf{E}\|X - \varphi(X)\| \leqslant \varepsilon$.*

Dowód pozostawiamy jako proste ćwiczenie.

Twierdzenie 4.2. *Jeśli $\mathbf{E}\|X\| < \infty$, to $\lim_{n\to\infty} \frac{S_n}{n} = \mathbf{E}X$ p.n..*

Dowód. Na mocy klasycznego Prawa Wielkich Liczb twierdzenie zachodzi dla $F = \mathbb{R}$, a więc również dla $F = \mathbb{R}^k$. Stąd łatwo pokazujemy tezę dla zmiennych przyjmujących wartości w pewnej skończenie wymiarowej podprzestrzeni F.

Przejdźmy do dowodu twierdzenia w ogólnym przypadku. Oczywiście możemy zakładać, że $\mathbf{E}X = 0$. Niech $\varepsilon > 0$ – wówczas, na mocy Lematu 4.1, istnieją niezależne zmienne losowe o jednakowym rozkładzie $Y_i = \varphi(X_i)$, przyjmujące tylko skończenie wiele wartości, spełniające $\mathbf{E}\|X_i - Y_i\| \leqslant \varepsilon$. Niech $T_n := Y_1 + \ldots + Y_n$, wówczas

$$\begin{aligned}\limsup_{n\to\infty} \frac{\|S_n\|}{n} &\leqslant \limsup_{n\to\infty} \frac{\|T_n\|}{n} + \limsup_{n\to\infty} \frac{\sum_{i=1}^n \|X_i - Y_i\|}{n} = \|\mathbf{E}Y\| + \mathbf{E}\|X - Y\| \\ &\leqslant \|\mathbf{E}X\| + 2\mathbf{E}\|X - Y\| \leqslant 2\varepsilon.\end{aligned}$$

Z dowolności $\varepsilon > 0$ dostajemy tezę. □

4.2 Prawo Wielkich Liczb Marcinkiewicza-Zygmunda

Okazuje się, że dla regularnych przestrzeni Banacha tradycyjne prawo wielkich liczb da się uogólnić. Zacznijmy od dwóch prostych lematów.

Lemat 4.3. *Jeśli $\mathbf{E}\|X\|^p = \infty$, to $\limsup_{n\to\infty} \frac{\|S_n\|}{n^{1/p}} = \infty$ p.n..*

Dowód. Niech

$$A := \left\{\omega\colon\ \limsup_{n\to\infty} \frac{\|S_n(\omega)\|}{n^{1/p}} < \infty\right\}.$$

Dla $\omega \in A$, dostajemy $\|X_n(\omega)\| \leqslant \|S_n(\omega)\| + \|S_{n-1}(\omega)\| \leqslant C(\omega)n^{1/p}$ dla $n \geqslant 1$ i pewnego $C(\omega) < \infty$. Zatem $A \subset \bigcup_{k=1}^{\infty} A_k$, gdzie

$$A_k := \{\|X_n\| \leqslant kn^{1/p} \text{ dla } n = 1, 2, \ldots\}.$$

Jeśli $\mathbf{P}(A_k) > 0$, to $\mathbf{P}(\limsup\{\|X_n\| > kn^{1/p}\}) < 1$, czyli na mocy Lematu Borela-Cantelliego,

$$\infty > \sum_{n=1}^{\infty} \mathbf{P}(\|X_n\| > kn^{1/p}) = \sum_{n=1}^{\infty} \mathbf{P}\left(\frac{\|X\|^p}{k^p} > n\right) \geqslant \mathbf{E}\frac{\|X\|^p}{k^p} - 1.$$

Zatem $\mathbf{P}(A_k) = 0$ dla wszystkich k, czyli $\mathbf{P}(A) = 0$. □

Lemat 4.4. *Załóżmy, że ciąg normujący liczb dodatnich* $(\gamma_n)_{n\geqslant 1}$ *spełnia warunek*

$$\gamma_{2l} \leqslant C\gamma_k \quad \text{dla } l < k \leqslant 2l.$$

Wówczas, jeśli

$$\sum_{n=1}^{\infty} \mathbf{P}(\|S_{2^n}\| \geqslant \varepsilon\gamma_{2^n}) < \infty \quad \text{dla wszystkich } \varepsilon > 0,$$

to $\lim_{n\to\infty} \frac{S_n}{\gamma_n} = 0$ *p.n.*.

Dowód. Szacujemy dla $\varepsilon > 0$,

$$\begin{aligned}\mathbf{P}\left(\sup_{m>2^{k-1}} \frac{\|S_m\|}{\gamma_m} \geqslant \varepsilon\right) &\leqslant \sum_{n=k}^{\infty} \mathbf{P}\left(\sup_{2^{n-1}<m\leqslant 2^n} \|S_m\| \geqslant \frac{\varepsilon}{C}\gamma_{2^n}\right) \\ &\leqslant \sum_{n=k}^{\infty} 4\mathbf{P}\left(\|S_{2^n}\| \geqslant \frac{\varepsilon}{6C}\gamma_{2^n}\right) \to 0 \quad \text{przy } k \to \infty,\end{aligned}$$

gdzie ostatnia nierówność wynika z Twierdzenia 3.10. □

Twierdzenie 4.5 (Prawo wielkich liczb Marcinkiewicza-Zygmunda)**.** *Załóżmy, że F jest ośrodkową przestrzenią Hilberta oraz $0 < p < 2$. Wówczas następujące warunki są równoważne.*
i) $\lim_{n\to\infty} \frac{S_n}{n^{1/p}} = 0$ *p.n.*
ii) $\mathbf{E}\|X\|^p < \infty$ *oraz jeśli* $p \geqslant 1$, *to* $\mathbf{E}X = 0$.

Dowód. ii)⇒i). Wpierw udowodnimy, że warunek ii) implikuje

$$\lim_{n\to\infty} n^{1-1/p}\mathbf{E}\left[X\mathbb{1}_{\{\|X\|\leqslant n^{1/p}\}}\right] = 0. \tag{11}$$

Dla $p \geqslant 1$, wobec $\mathbf{E}X = 0$ szacujemy

$$\begin{aligned} n^{1-1/p} \left\| \mathbf{E}\left[X \mathbb{1}_{\{\|X\| \leqslant n^{1/p}\}} \right] \right\| &= n^{1-1/p} \left\| \mathbf{E}\left[X \mathbb{1}_{\{\|X\| > n^{1/p}\}} \right] \right\| \\ &\leqslant \mathbf{E}\left[\|X\| (n^{1/p})^{p-1} \mathbb{1}_{\{\|X\| > n^{1/p}\}} \right] \\ &\leqslant \mathbf{E}\left[\|X\|^p \mathbb{1}_{\{\|X\| > n^{1/p}\}} \right] \to 0 \quad \text{przy } n \to \infty, \end{aligned}$$

na mocy skończoności $\mathbf{E}\|X\|^p$. Dla $0 < p < 1$ dostajemy dla dowolnego m,

$$\begin{aligned} n^{1-1/p} \mathbf{E}\left[\|X\| \mathbb{1}_{\{\|X\| \leqslant n^{1/p}\}} \right] &\leqslant n^{1-1/p} \mathbf{E}\left[\|X\| \mathbb{1}_{\{\|X\| \leqslant m\}} \right] + \mathbf{E}\left[\|X\| (n^{1/p})^{p-1} \mathbb{1}_{\{m < \|X\| \leqslant n^{1/p}\}} \right] \\ &\leqslant n^{1-1/p} \mathbf{E}\left[\|X\| \mathbb{1}_{\{\|X\| \leqslant m\}} \right] + \mathbf{E}\left[\|X\|^p \mathbb{1}_{\{\|X\| > m\}} \right] =: I + II. \end{aligned}$$

Składnik II można zrobić dowolnie małym dobierając odpowiednio duże m, zaś przy ustalonym m, składnik I dąży do zera przy n dążacym do nieskończoności. Czyli (11) zostało wykazane.

Na mocy Lematu 4.4, by wykazać i) wystarczy, iż udowodnimy

$$\sum_n \mathbf{P}(\|S_{2^n}\| \geqslant 2\varepsilon 2^{n/p}) < \infty \quad \text{dla } \varepsilon > 0.$$

Z (11) wynika, że $\|2^n \mathbf{E}[X \mathbb{1}_{\{\|X\| \leqslant 2^{n/p}\}}]\| \leqslant \varepsilon 2^{n/p}$ dla dużych n, więc wystarczy, iż udowodnimy, że

$$\sum_n \mathbf{P}\left(\left\| S_{2^n} - 2^n \mathbf{E}\left[X \mathbb{1}_{\{\|X\| \leqslant 2^{n/p}\}} \right] \right\| \geqslant \varepsilon 2^{n/p} \right) < \infty.$$

Określmy

$$\tilde{S}_{2^n} := \sum_{i=1}^{2^n} X_i \mathbb{1}_{\{\|X_i\| \leqslant 2^{n/p}\}},$$

wówczas

$$\begin{aligned} \mathbf{P}\left(\left\| S_{2^n} - 2^n \mathbf{E}\left[X \mathbb{1}_{\{|X| < 2^{n/p}\}} \right] \right\| \geqslant \varepsilon 2^{n/p} \right) &\leqslant \mathbf{P}(\tilde{S}_{2^n} \neq S_{2^n}) + \mathbf{P}(\|\tilde{S}_{2^n} - \mathbf{E}\tilde{S}_{2^n}\| \geqslant \varepsilon 2^{n/p}) \\ &\leqslant 2^n \mathbf{P}(\|X\| > 2^{n/p}) + \varepsilon^{-2} 2^{-2n/p} \mathbf{E}\|\tilde{S}_{2^n} - \mathbf{E}\tilde{S}_{2^n}\|^2. \end{aligned}$$

Mamy

$$\sum_{n=1}^{\infty} 2^n \mathbf{P}(\|X\| > 2^{n/p}) = \sum_{n=1}^{\infty} 2^n \mathbf{P}(\|X\|^p > 2^n) \leqslant 2\mathbf{E}\|X\|^p < \infty$$

oraz

$$\begin{aligned} \sum_{n=1}^{\infty} 2^{-2n/p} \mathbf{E}\|\tilde{S}_{2^n} - \mathbf{E}\tilde{S}_{2^n}\|^2 &\leqslant \sum_{n=1}^{\infty} 2^{n(1-2/p)} \mathbf{E}\left[\|X\|^2 \mathbb{1}_{\{\|X\| \leqslant 2^{n/p}\}} \right] \\ &= \mathbf{E}\left[\|X\|^2 \sum_{n=1}^{\infty} 2^{n(1-2/p)} \mathbb{1}_{\{\|X\|^p \leqslant 2^n\}} \right] \\ &\leqslant C_p \mathbf{E}[\|X\|^2 (\|X\|^p)^{1-2/p}] = C_p \mathbf{E}\|X\|^p < \infty. \end{aligned}$$

i)$\Rightarrow$ ii). Z Lematu 4.3, dostajemy $\mathbf{E}\|X\|^p < \infty$. W szczególności, gdy $p \geqslant 1$, to $\mathbf{E}\|X\| < \infty$ i na mocy zwykłego Mocnego Prawa Wielkich Liczb (albo udowodnionej już implikacji ii)$\Rightarrow$i) z $p = 1$ zastosowanej do $S_n - n\mathbf{E}X$), $\frac{S_n}{n} \to \mathbf{E}X$. Stąd $\mathbf{E}X = 0$. □

Uwaga 4.6. Dla $p \geqslant 2$ jeśli $\mathbf{P}(X \neq 0) > 0$, to $\limsup_{n\to\infty} \frac{S_n}{n^{1/p}} = \infty$ p.n.. Istotnie, tak jest na mocy Lematu 4.3, jeśli $\mathbf{E}X^2 = \infty$. W przypadku zaś, gdy $\mathbf{E}X^2 < \infty$, to implikacja łatwo wynika z centralnego twierdzenia granicznego.

Uwaga 4.7. Dokładna analiza dowodu Twierdzenia 4.5 pokazuje, że implikacja i)$\Rightarrow$ii) zachodzi dla dowolnych ośrodkowych przestrzeni Banacha F. Jedyną własnością F, którą potrzebujemy do dowodu implikacji ii)$\Rightarrow$i), jest nierówność

$$\mathbf{E}\left\|\sum_{i=1}^n Z_i\right\|^2 \leqslant C\sum_{i=1}^n \mathbf{E}\|Z_i\|^2$$

dla niezależnych, ograniczonych zmiennych losowych Z_i o wartościach w F i średniej 0. Przestrzenie z taką własnością nazywamy przestrzeniami typu 2 (należą do nich np. przestrzenie L^p, $2 \leqslant p < \infty$).

4.3 Przypadek niejednakowych rozkładów.

W poprzednich twierdzeniach zakładaliśmy wspólny rozkład zmiennych X_i. Jest wiele sformułowań mocnego prawa wielkich liczb, które nie zakładają tego warunku – udowodnimy jedno z nich, sformułowane dla $p = 2$ przez Kołmogorowa, a dla $p \geqslant 2$ przez Brunka.

Twierdzenie 4.8 (Mocne Prawo Wielkich Liczb Brunka)**.** *Załóżmy, że $p \geqslant 2$ oraz X_i są niezależnymi zmiennymi losowymi o średniej zero takimi, że*

$$\sum_{i=1}^\infty \frac{\mathbf{E}|X_i|^p}{i^{p/2+1}} < \infty.$$

Wówczas zachodzi mocne prawo wielkich liczb, tzn. $S_n/n \to 0$ p.n..

Dowód. Niech (X_i') będzie niezależną kopią ciągu (X_i) oraz $Y_i = X_i - X_i'$. Ponadto przez (ε_i) oznaczmy niezależny od (Y_i) ciąg Bernoulliego, tzn. ciąg niezależnych zmiennych losowych taki, że $\mathbf{P}(\varepsilon_i = \pm 1) = 1/2$. Mamy

$$\begin{aligned}
\mathbf{E}\left|\sum_{i=1}^n X_i\right|^p &\leqslant \mathbf{E}\left|\sum_{i=1}^n X_i - \sum_{i=1}^n X_i'\right|^p = \mathbf{E}\left|\sum_{i=1}^n Y_i\right|^p = \mathbf{E}_Y\mathbf{E}_\varepsilon\left|\sum_{i=1}^n \varepsilon_i Y_i\right|^p \\
&\leqslant \mathbf{E}_Y C_p\left(\sum_{i=1}^n |Y_i|^2\right)^{p/2} \leqslant C_p\mathbf{E}\left(n^{p/2-1}\sum_{i=1}^n |Y_i|^p\right) \\
&\leqslant 2^p C_p n^{p/2-1}\sum_{i=1}^n \mathbf{E}|X_i|^p,
\end{aligned}$$

gdzie kolejne nierówności wynikają z nierówności Jensena, nierówności Chinczyna, nierówności Höldera oraz oszacowania $\|Y_i\|_p \leqslant \|X_i\|_p + \|X_i'\|_p = 2\|X_i\|_p$. Stąd, na mocy Twierdzenia 3.6,

$$\begin{aligned}\mathbf{P}\Big(\max_{1\leqslant k\leqslant n}|S_k| \geqslant t\Big) &\leqslant 3\max_{1\leqslant k\leqslant n}\mathbf{P}(|S_k|\geqslant t/3) \leqslant 3\max_{1\leqslant k\leqslant n}(t/3)^{-p}\mathbf{E}|S_k|^p \\ &\leqslant C_p' t^{-p} n^{p/2-1}\sum_{i=1}^n \mathbf{E}|X_i|^p,\end{aligned}$$

gdzie $C_p' = 2^p 3^{p+1} C_p$. Zatem

$$\begin{aligned}\mathbf{P}\Big(\max_{k\geqslant 2^{l-1}}\frac{|S_k|}{k}\geqslant\varepsilon\Big) &\leqslant \sum_{j=l}^{\infty}\mathbf{P}\Big(\max_{2^{j-1}<k\leqslant 2^j}\frac{|S_k|}{k}\geqslant\varepsilon\Big) \leqslant \sum_{j=l}^{\infty}\mathbf{P}\Big(\max_{k\leqslant 2^j}|S_k|\geqslant 2^{j-1}\varepsilon\Big) \\ &\leqslant \sum_{j=l}^{\infty} C_p'(2^{j-1}\varepsilon)^{-p}(2^j)^{p/2-1}\sum_{i=1}^{2^j}\mathbf{E}|X_i|^p \\ &= C_p'(2/\varepsilon)^p\sum_{i=1}^{\infty}\mathbf{E}|X_i|^p\sum_{j\geqslant l:\,2^j\geqslant i}(2^{-j})^{p/2+1} \\ &\leqslant C_p''\varepsilon^{-p}\sum_{i=1}^{\infty}\frac{\mathbf{E}|X_i|^p}{\max\{i,2^l\}^{p/2+1}} \to 0 \quad \text{przy } l\to\infty.\end{aligned}$$

□

Uwaga 4.9. Twierdzenie Brunka (z tym samym dowodem) zachodzi też dla przestrzeni Hilberta oraz ogólniej dla przestrzeni Banacha typu 2.

5 Prawo Iterowanego Logarytmu

Podczas tego wykładu będziemy najczęściej zakładać, że $X, X_1, X_2, \ldots$ są niezależnymi zmiennymi losowymi o jednakowym rozkładzie o średniej zero i skończonej, niezerowej wariancji σ^2. Na mocy Centralnego Twierdzenia Granicznego $n^{-1/2}S_n$ zbiega wówczas według rozkładu do zmiennej gaussowskiej $\mathcal{N}(0,\sigma^2)$. Jednak, jak widzieliśmy w czasie poprzedniego wykładu, z prawdopodobieństwem jeden $\limsup_{n\to\infty}|n^{-1/2}S_n| = \infty$. Można więc zapytać jak duże musi być a_n by $\limsup_{n\to\infty}|S_n/a_n| < \infty$ p.n..

5.1 Przypadek jednowymiarowy

Twierdzenie 5.1 (Prawo Iterowanego Logarytmu Hartmana i Wintnera)**.** *Dla niezależnych zmiennych losowych o jednakowym rozkładzie ze średnią zero i wariancją σ^2 zachodzi*

$$\limsup_{n\to\infty}\frac{S_n}{\sqrt{2n\ln\ln n}} = -\liminf_{n\to\infty}\frac{S_n}{\sqrt{2n\ln\ln n}} = \sigma \qquad p.n..$$

Dowód twierdzenia podzielimy na szereg prostszych lematów. Dla uproszczenia notacji przyjmiemy $\gamma(x) := \sqrt{2x \ln\ln(x \vee 10)}$.

Lemat 5.2. *Dla dowolnego $t > 0$,*

$$\mathbf{P}\left(\max_{1\leqslant k\leqslant n} |S_k| \geqslant t + \sigma\sqrt{2n}\right) \leqslant 2\mathbf{P}(|S_n| \geqslant t).$$

Dowód. Z Lematu 3.5 dostajemy

$$\mathbf{P}\left(\max_{1\leqslant k\leqslant n} |S_k| \geqslant t + \sigma\sqrt{2n}\right) \leqslant \frac{\mathbf{P}(|S_n| \geqslant t)}{1 - \max_{1\leqslant k\leqslant n} \mathbf{P}(|S_n - S_k| \geqslant \sigma\sqrt{2n})}.$$

Na mocy nierówności Czebyszewa,

$$\mathbf{P}(|S_n - S_k| \geqslant \sigma\sqrt{2n}) \leqslant \frac{\mathrm{Var}(S_n - S_k)}{2n\sigma^2} = \frac{(n-k)\sigma^2}{2n\sigma^2} \leqslant \frac{1}{2}.$$

□

Lemat 5.3. *Załóżmy, że dla $\beta > 1, a > 0$ zachodzi warunek*

$$\sum_{n=1}^{\infty} \mathbf{P}(|S_{\lfloor \beta^n \rfloor}| \geqslant a\gamma(\beta^n)) < \infty.$$

Wówczas $\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\ln\ln n}} \leqslant a\beta^{1/2}$ *p.n..*

Dowód. Ustalmy $c > a\beta^{1/2}$ i oszacujmy,

$$\begin{aligned}\mathbf{P}\left(\sup_{n>\beta^{m-1}} \frac{|S_n|}{\sqrt{2n\ln\ln n}} > c\right) &\leqslant \sum_{k=m}^{\infty} \mathbf{P}\left(\max_{\beta^{k-1}<n\leqslant\beta^k} \frac{|S_n|}{\sqrt{2n\ln\ln n}} > c\right)\\ &\leqslant \sum_{k=m}^{\infty} \mathbf{P}\left(\max_{n\leqslant\beta^k} |S_n| > c\gamma(\beta^{k-1})\right).\end{aligned}$$

Zauważmy, że dla dostatecznie dużych k, $c\gamma(\beta^{k-1}) \geqslant a\gamma(\beta^k) + \sigma(2\beta^k)^{1/2}$, czyli na mocy Lematu 5.2,

$$\begin{aligned}\mathbf{P}\left(\max_{n\leqslant\beta^k} |S_n| > c\gamma(\beta^{k-1})\right) &\leqslant \mathbf{P}\left(\max_{n\leqslant\beta^k} |S_n| > a\gamma(\beta^k) + \sigma\sqrt{2\beta^k}\right)\\ &\leqslant 2\mathbf{P}(|S_{\lfloor\beta^k\rfloor}| \geqslant a\gamma(\beta^k)).\end{aligned}$$

Stąd,

$$\lim_{m\to\infty} \mathbf{P}\left(\sup_{n>\beta^{m-1}} \frac{|S_n|}{\sqrt{2n\ln\ln n}} > c\right) = 0$$

dla dowolnego $c > a\beta^{1/2}$ i łatwo otrzymujemy tezę. □

Lemat 5.4. *Mamy* $\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\ln\ln n}} \leqslant 4\sqrt{2}\sigma$ *p.n..*

Dowód. Załóżmy najpierw, że zmienne X_i są symetryczne i niech (ε_i) będzie ciągiem Bernoulliego niezależnym od (X_i). Wówczas ciąg (X_i) ma ten sam rozkład co $(\varepsilon_i X_i)$, zatem

$$\mathbf{P}(|S_n| \geqslant t) = \mathbf{P}\left(\left|\sum_{i=1}^{n} \varepsilon_i X_i\right| \geqslant t\right).$$

Z Wniosku 2.3 otrzymujemy dla $t, \alpha > 0$,

$$\mathbf{P}(|S_n| \geqslant t) \leqslant \mathbf{P}\left(\sum_{i=1}^{n} X_i^2 \geqslant \alpha\right) + 2\exp\left(-\frac{t^2}{2\alpha}\right).$$

Niech $T_n = \sum_{i=1}^n X_i^2$. Na mocy Mocnego Prawa Wielkich Liczb $T_n/n \to \sigma^2$ p.n., czyli również $2^{-n}(T_{2^{n+1}} - T_{2^n}) \to \sigma^2$ p.n.. Zatem na mocy Lematu Borela-Cantelliego,

$$\infty > \sum_{n=1}^{\infty} \mathbf{P}(2^{-n}(T_{2^{n+1}} - T_{2^n}) \geqslant 2\sigma^2) = \sum_{n=1}^{\infty} \mathbf{P}\left(\sum_{i=1}^{2^n} X_i^2 \geqslant 2^{n+1}\sigma^2\right).$$

Szacujemy,

$$\begin{aligned}\mathbf{P}(|S_{2^n}| \geqslant a\gamma(2^n)\sigma) &\leqslant \mathbf{P}\left(\sum_{i=1}^{2^n} X_i^2 \geqslant 2^{n+1}\sigma^2\right) + 2\exp\left(-\frac{a^2\sigma^2 2^{n+1}\ln\ln 2^n}{2^{n+2}\sigma^2}\right)\\ &\leqslant \mathbf{P}\left(\sum_{i=1}^{2^n} X_i^2 \geqslant 2^{n+1}\sigma^2\right) + 2(n\ln 2)^{-a^2/2}.\end{aligned}$$

Stąd dla symetrycznych zmiennych losowych,

$$\sum_n \mathbf{P}(|S_{2^n}| \geqslant 2\gamma(2^n)\sigma) < \infty. \tag{12}$$

W przypadku gdy X_i są dowolne oznaczmy przez (X_i') niezależną kopię (X_i) oraz niech $S_n' := \sum_{i=1}^n X_i'$. Na mocy nierówności Czebyszewa, $\mathbf{P}(|S_n'| > \sqrt{2n}\sigma) \leqslant 1/2$, więc z niezależności,

$$\mathbf{P}(|S_n| \geqslant t) \leqslant 2\mathbf{P}(|S_n| \geqslant t, |S_n'| \leqslant \sqrt{2n}\sigma) \leqslant 2\mathbf{P}(|S_n - S_n'| \geqslant t - \sqrt{2n}\sigma).$$

Zatem dla dużych n,

$$\mathbf{P}(|S_n| \geqslant 4\gamma(n)\sigma) \leqslant 2\mathbf{P}(|S_n - S_n'| \geqslant 2\sqrt{2}\gamma(n)\sigma).$$

Zauważmy, że $S_n - S_n' = \sum_{i=1}^n Y_i$, gdzie $Y_i = X_i - X_i'$ są niezależnymi, symetrycznymi zmiennymi losowymi o jednakowym rozkładzie z wariancją $2\sigma^2$, czyli na mocy (12),

$$\sum_{n>n_0} \mathbf{P}(|S_{2^n}| \geqslant 4\gamma(2^n)\sigma) \leqslant 2\sum_{n>n_0} \mathbf{P}(|S_{2^n} - S_{2^n}'| \geqslant 2\gamma(2^n)\sigma_Y) < \infty.$$

Możemy więc stosować Lemat 5.3 z $a = 4\sigma$ i $\beta = 2$. $\square$

Lemat 5.5. *Jeśli zmienne X_i są ograniczone, to* $\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\ln\ln n}} \leqslant \sigma$ *p.n..*

Dowód. Załóżmy, że $\|X\|_\infty \leqslant a$, wówczas na mocy nierówności Bernsteina (Wniosek 2.8),

$$\mathbf{P}(|S_n| \geqslant t) \leqslant 2\exp\left(-\frac{t^2}{2n\sigma^2 + 2at/3}\right).$$

Zatem dla $\beta > 1$ i $\varepsilon > 0$ oraz dostatecznie duzych n,

$$\begin{aligned}\mathbf{P}(|S_{\lfloor\beta^n\rfloor}| \geqslant (1+\varepsilon)\sigma\gamma(\beta^n)) &\leqslant 2\exp\left(-\frac{(1+\varepsilon)^2\sigma^2 2\beta^n \ln\ln\beta^n}{2\lfloor\beta^n\rfloor\sigma^2 + 2a(1+\varepsilon)\sigma\sqrt{2\beta^n\ln\ln\beta^n}/3}\right)\\ &\leqslant 2\exp(-(1+\varepsilon)\ln\ln\beta^n) = 2(n\ln\beta)^{-1+\varepsilon}.\end{aligned}$$

Możemy więc stosować Lemat 5.3 (z $a = (1+\varepsilon)\sigma$) i dostajemy

$$\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\ln\ln n}} \leqslant (1+\varepsilon)\sqrt{\beta}\sigma \quad \text{p.n. .}$$

Z dowolności $\beta > 1$ i $\varepsilon > 0$ wynika teza. □

Wniosek 5.6. *Dla dowolnych zmiennych o średniej 0 i wariancji σ^2,*

$$\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\ln\ln n}} \leqslant \sigma \quad \textit{p.n..}$$

Dowód. Ustalmy $M > 0$ i rozłóżmy $X_i = Y_i + Z_i$, gdzie

$$Y_i := X_i\mathbb{1}_{\{|X_i|\leqslant M\}} - \mathbf{E}X_i\mathbb{1}_{\{|X_i|\leqslant M\}}, \quad Z_i := X_i\mathbb{1}_{\{|X_i|>M\}} - \mathbf{E}X_i\mathbb{1}_{\{|X_i|>M\}}.$$

Określmy $T_n := Y_1 + \ldots + Y_n$, $R_n := Z_1 + \ldots + Z_n$. Na mocy Lematów 5.4 i 5.5,

$$\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\ln\ln n}} \leqslant \limsup_{n\to\infty} \frac{|T_n|}{\sqrt{2n\ln\ln n}} + \limsup_{n\to\infty} \frac{|R_n|}{\sqrt{2n\ln\ln n}} \leqslant \sigma_Y + 4\sqrt{2}\sigma_Z \quad \text{p.n..}$$

Mamy jednak $\sigma_Y^2 \leqslant \mathbf{E}X^2\mathbb{1}_{\{|X|\leqslant M\}} \leqslant \mathbf{E}X^2 = \sigma^2$ oraz $\sigma_Z^2 \leqslant \mathbf{E}X^2\mathbb{1}_{\{|X|>M\}} \to 0$ przy $M \to \infty$. □

Lemat 5.7. *Załóżmy, że dla liczby naturalnej $C > 1$ oraz $a > 0$ zachodzi*

$$\sum_n \mathbf{P}(S_{C^n - C^{n-1}} \geqslant a\gamma(C^n - C^{n-1})) = \infty.$$

Wówczas $\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} \geqslant a(1-C^{-1})^{1/2} - \sigma C^{-1/2}$ *p.n.*

Dowód. Zdarzenia $\{S_{C^n} - S_{C^{n-1}} \geqslant a\gamma(C^n - C^{n-1})\}$, $n = 1, 2, \ldots$ są niezależne, zatem na mocy Lematu Borela-Cantelliego,

$$1 = \mathbf{P}(\limsup\{S_{C^n} - S_{C^{n-1}} \geqslant a\gamma(C^n - C^{n-1})\}) \leqslant \mathbf{P}\left(\limsup_{n\to\infty} \frac{S_{C^n} - S_{C^{n-1}}}{\gamma(C^n - C^{n-1})} \geqslant a\right).$$

Stąd na podstawie Wniosku 5.6,

$$\begin{aligned}\limsup_{n\to\infty} \frac{S_{C^n}}{\gamma(C^n)} \geqslant & \limsup_{n\to\infty} \frac{S_{C^n} - S_{C^{n-1}}}{\gamma(C^n - C^{n-1})} \lim_{n\to\infty} \frac{\gamma(C^n - C^{n-1})}{\gamma(C^n)} \\ & - \limsup_{n\to\infty} \frac{-S_{C^{n-1}}}{\gamma(C^{n-1})} \lim_{n\to\infty} \frac{\gamma(C^{n-1})}{\gamma(C^n)} \\ \geqslant & a(1 - C^{-1})^{1/2} - \sigma C^{-1/2} \quad \text{p.n..}\end{aligned}$$

□

Lemat 5.8. *Jeśli zmienne X_i są ograniczone, to* $\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n \ln\ln n}} \geqslant \sigma$ *p.n..*

Dowód. Oszacujemy $\mathbf{P}(S_n \geqslant (1+\varepsilon)^{-1/2}\gamma(n)\sigma)$ stosując odwrotną nierówność wykładniczą Kołmogorowa (Twierdzenie 2.10) z $s = (1+\varepsilon)^{-1/2}\gamma(n)\sigma$ i $n\sigma^2$ zamiast σ^2. Zauważmy, że dla dużych $n \geqslant n(\varepsilon)$ jego założenia są spełnione, bo

$$s \max_i \|X_i\|_\infty \leqslant \|X\|_\infty \gamma(n)\sigma \leqslant \delta(\varepsilon) n\sigma^2$$

oraz

$$s = (1+\varepsilon)^{-1/2}\gamma(n)\sigma \geqslant K(\varepsilon)\sqrt{n}\sigma.$$

Zatem dla $n \geqslant n(\varepsilon)$,

$$\mathbf{P}\left(S_n \geqslant (1+\varepsilon)^{-1/2}\gamma(n)\sigma\right) \geqslant \exp\left(-\frac{(1+\varepsilon)(1+\varepsilon)^{-1}\gamma^2(n)\sigma^2}{2n\sigma^2}\right) = (\ln n)^{-1}.$$

Zatem

$$\sum_n \mathbf{P}\left(S_{C^n - C^{n-1}} \geqslant (1+\varepsilon)^{-1/2}\sigma\gamma(C^n - C^{n-1})\right) \geqslant \sum_{n \geqslant n(\varepsilon)} (\ln((C-1)C^{n-1}))^{-1} = \infty.$$

Czyli, na mocy Lematu 5.7,

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} \geqslant (1+\varepsilon)^{-1/2}\sigma(1 - C^{-1})^{1/2} - \sigma C^{-1/2} \quad \text{p.n..}$$

Biorąc $C \to \infty$ oraz $\varepsilon \to 0+$ dostajemy tezę. □

Wniosek 5.9. *Dla dowolnych zmiennych o średniej 0 i wariancji σ^2,*

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} \geqslant \sigma \quad \textit{p.n.}.$$

Dowód. Jak w dowodzie Wniosku 5.6 dla $M > 0$ rozkładamy $X_i = Y_i + Z_i$ oraz $S_n = R_n + T_n$ i szacujemy używając Lematu 5.8 w Wniosku 5.6,

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} \geqslant \limsup_{n\to\infty} \frac{T_n}{\sqrt{2n\ln\ln n}} - \limsup_{n\to\infty} \frac{-R_n}{\sqrt{2n\ln\ln n}} \geqslant \sigma_Y - \sigma_Z \quad \text{p.n.}.$$

By zakończyć dowód wystarczy zauważyć, że $\sigma_Y \to \sigma$ oraz $\sigma_Z \to 0$ przy $M \to \infty$. □

Dowód Twierdzenia 5.1. . Wnioski 5.6 i 5.9 implikują

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} = \sigma \quad \text{p.n.}.$$

Ponieważ $-S_n = \sum_{i=1}^n (-X_i)$, więc

$$\liminf_{n\to\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} = -\limsup_{n\to\infty} \frac{-S_n}{\sqrt{2n\ln\ln n}} = -\sigma_{-X} = -\sigma \quad \text{p.n.}.$$

□

Okazuje się, że jeśli spełnione jest prawo iterowanego logarytmu, to zmienna musi mieć średnią zero i skończoną wariancję.

Twierdzenie 5.10. *Załóżmy, że*

$$\mathbf{P}\Big(\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} < \infty\Big) > 0,$$

wówczas $\mathbf{E}X = 0$ *oraz* $\mathbf{E}X^2 < \infty$.

Dowód. Na mocy założeń istnieje $M < \infty$ takie, że

$$\mathbf{P}\left(\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\ln\ln n}} \leqslant M\right) > 0,$$

ale na mocy prawa 0-1, powyższe prawdopodobieństwo wynosi 0 lub 1. Stąd

$$\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\ln\ln n}} \leqslant M \quad \text{p.n.}.$$

Załóżmy najpierw dodatkowo, że zmienne X_i są symetryczne. Dla $t > 0$ zdefinujmy

$$X_n^t := X_n \mathbb{1}_{\{|X_n| \leqslant t\}} - X_n \mathbb{1}_{\{|X_n| > t\}}$$

oraz $S_n^t := \sum_{i=1}^n X_i^t$. Zmienne (X_n^t) mają ten sam rozkład co zmienne X_n, zatem

$$\limsup_{n\to\infty} \frac{|S_n^t|}{\sqrt{2n\ln\ln n}} \leqslant M \quad \text{p.n.},$$

czyli

$$\limsup_{n\to\infty} \frac{|S_n + S_n^t|}{\sqrt{2n\ln\ln n}} \leqslant 2M \quad \text{p.n.}.$$

Ale $S_n + S_n^t = \sum_{i=1}^n 2X_i \mathbb{1}_{\{|X_i|\leqslant t\}}$. Zmienne $X_i \mathbb{1}_{\{|X_i|\leqslant t\}}$ mają średnią zero i skończoną wariancję, więc na mocy Twierdzenia 5.1,

$$\limsup_{n\to\infty} \frac{|S_n + S_n^t|}{\sqrt{2n\ln\ln n}} = (\mathrm{Var}(2X\mathbb{1}_{\{|X|\leqslant t\}}))^{1/2} = (2\mathbf{E}X^2\mathbb{1}_{\{|X|\leqslant t\}})^{1/2} \quad \text{p.n.}$$

Stąd $\mathbf{E}X^2 I_{\{|X|\leqslant t\}} \leqslant 2M^2$ i z dowolności $t > 0$ dostajemy $\mathbf{E}X^2 < \infty$.

W przypadku, gdy rozkład X_i nie jest symetryczny, niech (X_n') będzie niezależną kopią (X_n), $S_n' = \sum_{i=1}^n X_i'$. Wówczas

$$\limsup_{n\to\infty} \frac{|S_n - S_n'|}{\sqrt{2n\ln\ln n}} \leqslant 2M \quad \text{p.n.}.$$

Mamy jednak $S_n - S_n' = \sum_{i=1}^n (X_i - X_i')$, stąd na mocy rozważanego wcześniej przypadku, $\mathbf{E}(X - X')^2 < \infty$, a więc również $\mathbf{E}X^2 < \infty$.

By zakończyć dowód zauważmy, że na mocy mocnego prawa wielkich liczb

$$\mathbf{E}X = \lim_{n\to\infty} \frac{S_n}{n} = \limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} \lim_{n\to\infty} \frac{\sqrt{2n\ln\ln n}}{n} = 0.$$

□

5.2 Zbiór graniczny

Okazuje się, że nie tylko jesteśmy w stanie podać granicę górną i dolną ciągu $\frac{S_n}{\sqrt{2n\ln\ln n}}$, ale również określić zbiór wszystkich możliwych granic jego podciągów. Wprowadźmy najpierw definicję.

Definicja 5.11. Dla ciągu liczbowego (lub ogólniej ciągu o wartościach w przestrzeni metrycznej) a_n przez $C(a_n)$ oznaczamy zbiór wszystkich punktów skupienia ciągu a_n, czyli

$$x \in C(a_n) \ \Leftrightarrow\ \exists_{n_k}\ a_{n_k} \to x.$$

Twierdzenie 5.12. *Załóżmy, że* $\mathbf{E}X = 0$ *oraz* $\mathbf{E}X^2 = \sigma^2 < \infty$. *Wówczas*
i) $\mathrm{dist}(\frac{S_n}{\sqrt{2n\ln\ln n}}, [-\sigma,\sigma]) \to 0$ *p.n.,*
ii) $C(\frac{S_n}{\sqrt{2n\ln\ln n}}) = [-\sigma,\sigma]$ *p.n..*

Dowód opiera się na następującym prostym lemacie.

Lemat 5.13. *Załóżmy, że* a_n *jest ciągiem liczbowym takim, że* $b = \liminf_{n\to\infty} a_n > -\infty$, $c = \limsup_{n\to\infty} a_n < \infty$ *oraz* $\lim_{n\to\infty}(a_{n+1} - a_n) = 0$. *Wówczas*
i) $\mathrm{dist}(a_n, [b,c]) \to 0$,
ii) $C(a_n) = [b,c]$.

Dowód. i) Z definicji granicy dolnej i górnej wynika, że dla $\varepsilon > 0$ dla dostatecznie dużych n, $b - \varepsilon \leqslant x_n \leqslant c + \varepsilon$, co pociąga $\mathrm{dist}(a_n, [b,c]) < \varepsilon$.

ii) Oczywiście $\{b,c\} \subset C(a_n)$, ustalmy $d \in (b,c)$. Możemy skonstruować podciąg n_k taki, że $a_{n_{2k-1}} < d < a_{n_{2k}}$. Zdefiniujmy

$$m_k := \sup\{n \leqslant n_{2k} \colon a_n \leqslant d\},$$

wówczas $a_{m_k} \leqslant d < a_{m_k+1}$, zatem $|d - a_{m_k}| \leqslant |a_{m_k+1} - a_{m_k}| \to 0$, czyli $d \in C(a_n)$. □

Dowód Twierdzenia 5.12.. Na mocy Lematu 5.13 oraz Twierdzenia 5.1 wystarczy wykazać, że

$$\lim_{n\to\infty}\left(\frac{S_n}{\gamma(n)} - \frac{S_{n-1}}{\gamma(n-1)}\right) = 0 \quad \text{p.n.},$$

gdzie $\gamma(n) := \sqrt{2n \ln\ln n}$. Zauważmy, że $\gamma(n)^{-1} - \gamma(n-1)^{-1} = o(\gamma(n-1)^{-1})$ oraz $\limsup |\frac{S_{n-1}}{\gamma(n-1)}| = \sigma < \infty$ p.n., więc wystarczy udowodnić, że

$$\lim_{n\to\infty} \frac{X_n}{\gamma(n)} = 0 \quad \text{p.n..}$$

Szacujemy,

$$\begin{aligned}\sum_n \mathbf{P}(|X_n| \geqslant \varepsilon\gamma(n)) &= \mathbf{E}\sum_n \mathbb{1}_{\{|X|\geqslant\varepsilon\gamma(n)\}} \leqslant \mathbf{E}\sum_n \mathbb{1}_{\{|X|\geqslant\varepsilon\sqrt{2n}\}} \\ &\leqslant \mathbf{E}\sum_n \mathbb{1}_{\{2n\leqslant|X|^2\varepsilon^{-2}\}} \leqslant \mathbf{E}\frac{|X|^2}{2\varepsilon^2} < \infty,\end{aligned}$$

zatem na podstawie Lematu Borela-Cantelliego, $\limsup \frac{|X_n|}{\gamma(n)} \leqslant \varepsilon$ p.n.. □

5.3 Przypadek wektorowy

W tej części będziemy rozważać Prawo Iterowanego Logarytmu dla wektorów losowych.

Dla $x \in \mathbb{R}^d$ przez $|x|$ będziemy oznaczać euklidesową długość wektora x.

Twierdzenie 5.14. *Załóżmy, że* $X, X_1, \ldots$ *są niezależnymi wektorami losowymi o jednakowym rozkładzie o wartościach w* $\mathbb{R}^d$. *Jeśli* $\mathbf{E}X = 0$ *oraz* $\mathbf{E}|X|^2 < \infty$, *to*

$$\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\ln\ln n}} = \sup_{|t|\leqslant 1}(\mathbf{E}\langle X,t\rangle^2)^{1/2} = \sup_{|t|\leqslant 1}(\langle Ct,t\rangle)^{1/2} \quad \textit{p.n.},$$

gdzie C *oznacza macierz kowariancji* X.

Dowód. Ustalmy $t \in \mathbb{R}^d$ z $|t| \leqslant 1$, wówczas $|x| \geqslant \langle x, t\rangle$ dla dowolnego $x \in \mathbb{R}^d$ i

$$\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\ln\ln n}} \geqslant \limsup_{n\to\infty} \frac{\langle S_n, t\rangle}{\sqrt{2n\ln\ln n}} = (\mathbf{E}\langle X, t\rangle^2)^{1/2} \text{ p.n.},$$

stąd dostajemy

$$\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\ln\ln n}} \geqslant \sup_{|t|\leqslant 1} (\mathbf{E}\langle X, t\rangle^2)^{1/2} \qquad \text{p.n..}$$

By udowodnić nierówność w drugą stronę, ustalmy $\varepsilon > 0$ i wybierzmy skończony podzbiór T kuli domknietej $\overline{B(0,1)}$ taki, że

$$\overline{B(0,1)} \subset \bigcup_{t\in T} B(t, \varepsilon).$$

Dla $x \in \mathbb{R}^d$, istnieje $|s| = 1$ takie, że $\langle s, x\rangle = |x|$, wybierając $t \in T$ takie, że $|s - t| < \varepsilon$ dostajemy $\langle t, x\rangle = \langle s, x\rangle - \langle s - t, x\rangle \geqslant |x| - \varepsilon|x|$. Zatem

$$\forall_{x\in\mathbb{R}^n} \ \max_{t\in T} \langle t, x\rangle \geqslant (1 - \varepsilon)|x|,$$

w szczególności

$$\frac{|S_n|}{\sqrt{2n\ln\ln n}} \leqslant \frac{1}{1-\varepsilon} \max_{t\in T} \frac{\langle S_n, t\rangle}{\sqrt{2n\ln\ln n}}.$$

Zatem ze skończoności zbioru T dostajemy

$$\begin{aligned}\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\ln\ln n}} &\leqslant \frac{1}{1-\varepsilon} \max_{t\in T} \limsup_{n\to\infty} \frac{\langle S_n, t\rangle}{\sqrt{2n\ln\ln n}} \\ &= \frac{1}{1-\varepsilon} \max_{t\in T} (\mathbf{E}\langle X, t\rangle^2)^{1/2} \qquad \text{p.n..}\end{aligned}$$

Z dowolności $\varepsilon > 0$,

$$\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\ln\ln n}} \leqslant \sup_{|t|\leqslant 1} (\mathbf{E}\langle X, t\rangle^2)^{1/2} \qquad \text{p.n..}$$

□

Podobnie jak w przypadku skalarnym, również dla d-wymiarowych wektorów losowych warunki $\mathbf{E}X = 0$ oraz $\mathbf{E}|X|^2 < \infty$ są konieczne dla zachodzenia Prawa Iterowanego Logaytmu.

Twierdzenie 5.15. *Załóżmy, że $X, X_1, \ldots$ są niezależnymi wektorami losowymi o jednakowym rozkładzie o wartościach w $\mathbb{R}^d$. Jeśli*

$$\mathbf{P}\Big(\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\ln\ln n}} < \infty\Big) > 0,$$

to $\mathbf{E}X = 0$ oraz $\mathbf{E}|X|^2 < \infty$.

Dowód. Szacowanie $\langle x,t\rangle \leqslant |x||t|$ oraz Twierdzenie 5.10 implikują, że dla dowolnego $t\in\mathbb{R}^d$, $\mathbf{E}\langle X,t\rangle = 0$ oraz $\mathbf{E}\langle X,t\rangle^2<\infty$. □

Dla wektora losowego X o wartościach w $\mathbb{R}^d$ takiego, że $\mathbf{E}|X|^2<\infty$ definiujemy

$$K_X := \{\mathbf{E}(\xi X)\colon\ \xi\in L^2(\Omega),\ \mathbf{E}\xi^2\leqslant 1\}$$

Jeśli $\mathrm{Cov}(X) = C = UDU^T$, gdzie $U^T = U^{-1}$ oraz D jest macierzą diagonalną, to kładziemy $C^{1/2} := UD^{1/2}U^T$.

Lemat 5.16. *Mamy*

$$K_X = \{Ct\colon\ \langle Ct,t\rangle\leqslant 1\} = \{C^{1/2}s\colon\ |s|\leqslant 1\},$$

czyli K_X jest elipsoidą.

Dowód. Weźmy $\xi\in L^2(\Omega)$ wówczas $\xi = \sum_{i=1}^d t_iX(i)+\eta$, gdzie $\mathbf{E}(\eta X(i)) = 0$ dla $i = 1,\ldots,d$. Mamy

$$\mathbf{E}\xi^2 = \mathbf{E}\left(\sum_{i=1}^d t_iX(i)\right)^2 + \mathbf{E}\eta^2 = \langle Ct,t\rangle + \mathbf{E}\eta^2$$

oraz

$$\mathbf{E}(\xi X(j)) = \sum_{i=1}^d t_i\mathbf{E}X(i)X(j) = \sum_{i=1}^d t_ic_{ji} = (Ct)_j,$$

skąd wynika pierwsza równość. Druga jest konsekwencją tożsamości $\langle Ct,t\rangle = |C^{1/2}t|^2$ i $Ct = C^{1/2}C^{1/2}t$. □

Twierdzenie 5.17. *Załóżmy, że spełnione są założenia Twierdzenia 5.14. Wówczas*
a) $\mathrm{dist}(\frac{S_n}{\sqrt{2n\ln\ln n}}, K_X)\to 0$ *p.n.,*
b) $C(\frac{S_n}{\sqrt{2n\ln\ln n}}) = K_X$ *p.n..*

Dowód. Zacznijmy od rozpatrzenia przypadku $C = \mathrm{Cov}(X) = \mathrm{Id}$. Wówczas $K_X = \overline{B(0,1)}$ i punkt a) wynika stąd, że $\limsup_{n\to\infty}\frac{|S_n|}{\gamma(n)} = 1$ p.n. (Twierdzenie 5.14).

By udowodnić b) dla $C = \mathrm{Id}$ weźmy najpierw $t\in S^{d-1}$, wówczas $\mathbf{E}\langle t,X\rangle^2 = 1$ i z Twierdzenia 5.1 z prawdopodobieństwem 1 istnieje podciąg n_k taki, że $\frac{\langle S_{n_k},t\rangle}{\gamma(n_k)}\to 1$, ale wtedy

$$\limsup_{k\to\infty}\left|\frac{S_{n_k}}{\gamma(n_k)} - t\right|^2 = \limsup_{k\to\infty}\Big(\frac{|S_{n_k}|}{\gamma(n_k)}\Big)^2 + |t|^2 - 2\lim_{k\to\infty}\frac{\langle S_{n_k},t\rangle}{\gamma(n_k)} = 0.$$

Zatem dla $|t| = 1$, $\mathbf{P}(t\in C(\frac{S_n}{\gamma(n)})) = 1$.

Jeśli $|t|<1$, to rozpatrzmy (ε_i) ciąg Bernoulliego niezależny od ciągu (X_i) i połóżmy $Y_i := X_i + \varepsilon_ie_{d+1}$. Wówczas (Y_i) jest ciągiem niezależnych wektorów losowych w $\mathbb{R}^{d+1}$

o średniej zero i macierzy kowariancji Id. Jeśli położymy $\tilde{S}_n := \sum_{i\leqslant n} Y_i$, to na mocy poprzednich rozważań,

$$t + (1-|t|^2)^{1/2}e_{d+1} \in C\left(\frac{\tilde{S}_n}{\gamma(n)}\right) \quad \text{p.n.},$$

co w szczególności implikuje $t \in C(\frac{S_n}{\gamma(n)})$ p.n.. A zatem dla dowolnego $t \in \overline{B(0,1)}$, $\mathbf{P}(t \in C(\frac{S_n}{\gamma(n)})) = 1$, skąd łatwo wynika (z domkniętości $C(a_n)$ i ośrodkowości kuli) $\mathbf{P}(\overline{B(0,1)} \subset C(\frac{S_n}{\gamma(n)})) = 1$.

Jeśli X dowolne takie, że $C \neq 0$, to istnieje $k \leqslant d$ oraz przekształcenia liniowe $A\colon \mathbb{R}^d \to \mathbb{R}^k$ i $B\colon \mathbb{R}^k \to \mathbb{R}^d$ takie, że $Y := AX$ jest zmienną o średniej zero i macierzy kowariancji Id oraz $X = BY$ p.n. Wówczas $K_X = BK_Y = B\overline{B_k(0,1)}$, $S_n = BT_n$ p.n., gdzie $T_n = \sum_{i\leqslant n} AX_i$, czyli

$$\operatorname{dist}\left(\frac{S_n}{\gamma(n)}, K_X\right) = \operatorname{dist}\left(B\frac{T_n}{\gamma(n)}, BK_Y\right) \leqslant \|B\|\operatorname{dist}\left(\frac{T_n}{\gamma(n)}, K_Y\right) \to 0 \quad \text{p.n.}$$

oraz

$$C\left(\frac{S_n}{\gamma(n)}\right) = BC\left(\frac{T_n}{\gamma(n)}\right) = BK_Y = K_X \quad \text{p.n.}.$$

Dla $C = 0$, $X = 0$ p.n. oraz $K_X = \{0\}$ i teza twierdzenia jest oczywista. $\square$

W przypadku wektorów losowych o wartościach w przestrzeni Banacha warunki konieczne i dostateczne dla zachodzenia prawa iterowanego logarytmu są bardziej skomplikowane. Ich znalezienie zajęło długi okres czasu, ostateczny wynik podali Ledoux i Talagrand w 1988 roku, wykorzystując subtelne wyniki Talagranda dotyczące koncentracji miary.

Zauważmy, że dla dowolnego funkcjonału φ o normie 1,

$$\limsup_{n\to\infty} \frac{\|S_n\|}{\sqrt{2n\ln\ln n}} \geqslant \limsup_{n\to\infty} \frac{|\varphi(S_n)|}{\sqrt{2n\ln\ln n}}.$$

W szczególności, jeśli $\mathbf{P}(X = 0) < 1$, to $\limsup_{n\to\infty} \|S_n\|/\sqrt{2n\ln\ln n} > 0$ p.n.. Zanim sformułujemy twierdzenie podające warunki równoważne skończoności tej granicy, przypomnimy definicję stochastycznej ograniczoności.

Definicja 5.18. Mówimy, że rodzina wektorów losowych $(Y_i)_{i\in I}$ o wartościach w przestrzeni Banacha $(F, \|\ \|)$ jest *stochastycznie ograniczona*, jeśli

$$\forall_{\varepsilon>0}\exists_{M<\infty}\forall_{i\in I}\ \mathbf{P}(\|Y_i\| \geqslant M) \leqslant \varepsilon.$$

W przypadku zmiennych o wartościach w $\mathbb{R}^d$ stochastyczna ograniczoność jest równoważna ciasności, dla przestrzeni Banacha nieskończonego wymiaru tak już nie jest.

Dla uproszczenia notacji będziemy używać oznaczenia $LLt := \log\log(\max\{t, 10\})$ dla $t \geqslant 0$.

Twierdzenie 5.19 (Ledoux-Talagrand)**.** *Załóżmy, że $X, X_1, \ldots$ są niezależnymi wektorami losowymi o jednakowym rozkładzie o wartościach w ośrodkowej przestrzeni Banacha $(F, \| \ \|)$. Wówczas spełnione jest Prawo Iterowanego Logarytmu, tzn.*

$$\limsup_{n\to\infty} \frac{\|S_n\|}{\sqrt{2n \ln\ln n}} < \infty \quad p.n.$$

wtedy i tylko wtedy, gdy spełnione są następujące trzy warunki:
i) $\mathbf{E}(\|X\|^2/LL\|X\|) < \infty$*,*
ii) dla dowolnego funkcjonału $\varphi \in F^*$*,* $\mathbf{E}\varphi(X) = 0$ *oraz* $\mathbf{E}\varphi(X)^2 < \infty$*,*
iii) ciąg $(\|S_n\|/\sqrt{nLLn})$ *jest stochastycznie ograniczony.*

Konieczność warunków i)-iii) jest dość łatwa do wykazania, jednak udowodnienie ich dostateczności jest trudne. Zainteresowanego czytelnika odsyłamy do rozdziału 8 monografii [7].

Okazuje się, że w przestrzeniach Hilberta (ogólniej w przestrzeniach typu 2) można się pozbyć warunku iii).

Fakt 5.20. *Załóżmy, że spełnione są założenia Twierdzenia 5.19 oraz przestrzeń Banacha F ma typ 2. Wówczas, jeśli* $\mathbf{E}(\|X\|^2/LL\|X\|) < \infty$ *oraz* $\mathbf{E}X = 0$*, to* $\|S_n\|/\sqrt{n\ln\ln n}$ *zbiega do* 0 *według prawdopodobieństwa.*

Uwaga 5.21*.* Z prawa 0-1 łatwo wynika, że jeśli zachodzą warunki i)-iii) Twierdzenia 5.19, to istnieje stała $c = c(X) > 0$ (zależna tylko od rozkładu zmiennej X) taka, że $\limsup_{n\to\infty} \|S_n\|/\sqrt{2n\ln\ln n} = c$ p.n.. Ogólnie identyfikacja stałej c jest problemem otwartym. Wiadomo jednak, że jeśli warunek iii) zastąpimy silniejszym warunkiem zbieżności według prawdopodobieństwa $\|S_n\|/\sqrt{nLLn}$ do zera, to

$$\limsup_{n\to\infty} \frac{\|S_n\|}{\sqrt{2n\ln\ln n}} = \sup_{\varphi\in F^*, \|\varphi\|\leqslant 1} \sqrt{\mathbf{E}\varphi(X)^2}.$$

W szczególności umiemy zidentyfikować granicę w Prawie Iterowanego Logarytmu dla przestrzeni Banacha typu 2.

6 Błądzenia losowe

W tym wykładzie będziemy analizować zachowanie błądzeń losowych, tzn. procesów postaci

$$S_n = X_1 + \ldots + X_n, \quad n \geqslant 0,$$

gdzie $X, X_1, \ldots$ są niezależnymi zmiennymi losowymi (jedno bądź wielowymiarowymi).

6.1 Twierdzenie Hewitta-Savage'a. Mocna własność Markowa.

Przy badaniu błądzeń losowych często pojawiają się zdarzenia postaci $\limsup\{S_n \in B\} = \{S_n \in B \text{ dla nieskończenie wielu } n\}$. Okazuje się, że prawdopodobieństwo takich zdarzeń jest równe 0 albo 1, wynika to z ogólniejszego wyniku – prawa 0-1 Hewitta-Savage'a. Zanim je sformułujemy i udowodnimy będziemy potrzebowali kilku definicji.

Definicja 6.1. Mówimy, że przekształcenie π na liczbach całkowitych dodatnich jest *skończoną permutacją*, jeśli $\pi(n) = n$ dla dużych n.

Dla przestrzeni mierzalnej $(S, \mathcal{S})$ przez $(S^\infty, \mathcal{S}^\infty)$ oznaczamy jej nieskończony produkt. Każda permutacja skónczona π indukuje przekształcenie T_π z S^∞ w S^∞, dane wzorem $T_\pi((x_i)_{i\geqslant 1}) := (x_{\pi(i)})_{i\geqslant 1}$ oraz przekształcenie T_π^{-1} z $\mathcal{S}^\infty$ w $\mathcal{S}^\infty$ określone jako

$$T_\pi^{-1}(A) := \{x \in S^\infty \colon\ T_\pi(x) = (x_{\pi(i)})_{i\geqslant 1} \in A\}.$$

Mówimy, że zbiór $A \in \mathcal{S}^\infty$ jest *permutowalny* bądź *niezmienniczy na skończone permutacje*, jeśli $T_\pi^{-1}(A) = A$ dla dowolnej skończonej permutacji π.

Twierdzenie 6.2 (Prawo 0-1 Hewitta-Savage'a)**.** *Załóżmy, że X_i są niezależnymi zmiennymi losowymi o jednakowym rozkładzie i wartościach w przestrzeni $(S, \mathcal{S})$, zaś $A \in \mathcal{S}^\infty$ jest zdarzeniem permutowalnym. Wówczas*

$$\mathbf{P}((X_1, X_2, \ldots) \in A) \in \{0, 1\}.$$

Dowód będzie korzystał z prostego lematu o σ-ciałach.

Lemat 6.3. *Załóżmy, że $(\mathcal{F}_n)_{n\geqslant 1}$ jest wstępującym ciągiem σ-ciał oraz $\mathcal{F} := \sigma(\bigcup_{n\geqslant 1} \mathcal{F}_n)$. Wówczas dla dowolnego $\varepsilon > 0$ i $A \in \mathcal{F}$ istnieją n oraz $A_n \in \mathcal{F}_n$ takie, że $\mathbf{P}(A \triangle A_n) \leqslant \varepsilon$.*

Dowód. Nietrudno sprawdzić, że zbiór $\mathcal{A}$ zdarzeń A takich, że dla każdego $\varepsilon > 0$ istnieją n i $A_n \in \mathcal{F}_n$ spełniające $\mathbf{P}(A \triangle A_n) \leqslant \varepsilon$ tworzy σ-ciało. Oczywiście wszystkie $\mathcal{F}_n$ są zawarte w $\mathcal{A}$, więc $\mathcal{F} \subset \mathcal{A}$. □

Dowód Twierdzenia 6.2. Niech $\mathcal{F}_n = \mathcal{S}^n \times S^\infty$, czyli $\mathcal{F}_n$ to σ-ciało mierzalnych podzbiorów S^∞ zależnych tylko od pierwszych n współrzędnych. Wówczas, oczywiście $\mathcal{S}^\infty = \sigma(\bigcup_{n\geqslant 1} \mathcal{F}_n)$. Niech μ oznacza miarę probabilistyczna na $(S^\infty, \mathcal{S}^\infty)$, która jest rozkładem zmiennej $(X_1, X_2 \ldots)$.

Niech A będzie permutowalnym zdarzeniem z $\mathcal{S}^\infty$. Na mocy Lematu 6.3 istnieją $A_n \in \mathcal{F}_n$ takie, że $\mu(A \triangle A_n) \to 0$. W szczególności $\mu(A_n) \to \mu(A)$.

Niech π_n oznacza permutację skończoną przeprowadzającą i na $n + i$ dla $i = 1, \ldots, n$ oraz $A_n' := T_{\pi_n}^{-1}(A_n)$. Wówczas, na mocy permutowalności, $\mu(A \triangle A_n') = \mu(A \triangle A_n)$. Mamy więc $\mu(A \triangle (A_n \cap A_n')) \leqslant \mu(A \triangle A_n') + \mu(A \triangle A_n) \to 0$.

Zdarzenie A_n' zależy tylko od współrzędnych $n + 1, \ldots, 2n$, więc względem μ jest niezależne od A_n. Stąd

$$\mu(A) = \lim \mu(A_n \cap A_n') = \lim \mu(A_n)\mu(A_n') = \mu(A)^2,$$

czyli $\mathbf{P}((X_1, X_2, \ldots) \in A) = \mu(A) \in \{0, 1\}$.

$\square$

Wniosek 6.4. *Załóżmy, że S_n jest błądzeniem losowym, wówczas dla dowolnego zbioru borelowskiego B, $\mathbf{P}(\limsup\{S_n \in B\})$ jest równe 0 lub 1.*

Dowód. Zauważmy, że jeśli π jest taką permutację, że $\pi(k) = k$ dla $k \geqslant n_0$, oraz $A_n := \{\sum_{i=1}^n x_i \in B\}$, to $T_\pi^{-1}(A_n) = A_n$ dla $n \geqslant n_0$. Stąd zdarzenie $\limsup A_n$ jest permutowalne. $\square$

Będziemy też często korzystali z tego, że błądzenie losowe ma mocną własność Markowa, tzn. jeśli zatrzymamy je w losowym momencie, a następnie znowu wystartujemy, to będzie się zachowywać tak jak przesunięte błądzenie losowe.

Twierdzenie 6.5. *Załóżmy, że $S_n = X_1 + \ldots + X_n$ jest błądzeniem losowym oraz $\mathcal{F}_n := \sigma(X_1, \ldots, X_n)$. Wówczas dla dowolnego momentu zatrzymania τ takiego, że $\mathbf{P}(\tau < \infty) > 0$, ciąg $(S_{\tau+n} - S_\tau)_{n\geqslant 0}$, warunkowany zdarzeniem $\tau < \infty$, jest błądzeniem losowym, niezależnym od $\mathcal{F}_\tau$, o takim samy rozkładzie co (S_n).*

Dowód. Ustalmy $A \in \mathcal{F}_\tau$, wówczas $A \cap \{\tau = k\} \in \mathcal{F}_k$ i jest niezależne od $(X_n)_{n>k}$. Stąd dla dowolnego zbioru $B \in (\mathcal{B}(\mathbb{R}^d))^\infty$

$$\begin{aligned}
\mathbf{P}(\{(S_{\tau+n} - S_\tau)_{n\geqslant 0} \in B\} \cap A \cap \{\tau = k\}) &= \mathbf{P}(\{(S_{k+n} - S_k)_{n\geqslant 0} \in B\} \cap A \cap \{\tau = k\}) \\
&= \mathbf{P}((S_{k+n} - S_k)_{n\geqslant 0} \in B)\mathbf{P}(A \cap \{\tau = k\}) \\
&= \mathbf{P}((S_n)_{n\geqslant 0} \in B)\mathbf{P}(A \cap \{\tau = k\}).
\end{aligned}$$

Sumując powyższą tożsamość po k otrzymujemy

$$\mathbf{P}(\{(S_{\tau+n} - S_\tau)_{n\geqslant 0} \in B\} \cap A \cap \{\tau < \infty\}) = \mathbf{P}((S_n)_{n\geqslant 0} \in B)\mathbf{P}(A \cap \{\tau < \infty\}).$$

$\square$

6.2 Błądzenia chwilowe i powracające

Dla bładzenia losowego S_n, o wartościach w $\mathbb{R}^d$, określmy miarę η liczącą ile razy błądzenie losowe odwiedza dany zbiór wzorem

$$\eta(B) := \sum_{n\geqslant 0} \mathbb{1}_{\{S_n \in B\}}, \quad B \in \mathcal{B}(\mathbb{R}^d).$$

Zauważmy, że

$$\mathbf{E}\eta(B) = \sum_{n\geqslant 0} \mathbf{P}(S_n \in B), \quad B \in \mathcal{B}(\mathbb{R}^d).$$

Określmy też zbiory

$$A := \{x \in \mathbb{R}^d\colon\ \mathbf{E}\eta(B(x,r)) > 0 \text{ dla wszystkich } r > 0\},$$

$$M := \{x \in \mathbb{R}^d\colon\ \mathbf{E}\eta(B(x,r)) = \infty \text{ dla wszystkich } r > 0\}$$

oraz

$$R := \{x \in \mathbb{R}^d\colon\ \eta(B(x,r)) = \infty \text{ p.n. dla wszystkich } r > 0\}.$$

Twierdzenie 6.6. *Dla dowolnego błądzenia losowego S_n o wartościach w $\mathbb{R}^d$ zachodzi dokładnie jeden z dwóch warunków*
i) $\mathbf{P}(|S_n| \to \infty) = 0$, $A = M = R$ *oraz zbiór ten jest domknietą podgrupą* $\mathbb{R}^d$*;*
ii) $M = R = \emptyset$ *oraz* $|S_n| \to \infty$ *p.n..*

Błądzenie spełniające warunek i) nazywamy *powracającym*, zaś warunek ii) *chwilowym.*

Dowód. Zauważy, że zawsze $R \subset M \subset A$. Łatwo też widać, że A jest domkniętą półgrupą. Rozpatrzymy dwa przypadki.

Przypadek I. $\mathbf{P}(|S_n| \to \infty) < 1$.

Wówczas oczywiście istnieje $V < \infty$ takie, że $\mathbf{P}(\limsup\{|S_n| < V\}) > 0$. Ustalmy $r > 0$, ponieważ kulę $B(0, V)$ można pokryć skończoną liczbą kul $B(x_i, r/2)$, $i = 1, \ldots, N$, więc dla pewnego k, $\mathbf{P}(\limsup\{S_n \in B(x_k, r/2)\}) > 0$. Na mocy twierdzenia Hewitta-Savage'a, $\mathbf{P}(\limsup\{S_n \in B(x_k, r/2)\}) = 1$, czyli $\tau := \inf\{n\colon\ S_n \in B(x_k, r/2)\} < \infty$ p.n.. Na mocy mocnej własności Markowa dostajemy

$$1 = \mathbf{P}(\limsup\{S_n \in B(x_k, r/2)\}) \leqslant \mathbf{P}(\limsup\{|S_{\tau+n} - S_\tau| < r\}) = \mathbf{P}(\limsup\{|S_n| < r\}),$$

stąd dostajemy, że $0 \in R$.

Pokażemy teraz, że $A \subset R$, w tym celu ustalmy $x \in A$ oraz $r > 0$. Niech $\sigma := \inf\{n\colon\ |S_n - x| < r/2\}$. Mamy, na mocy mocnej własności Markowa,

$$\begin{aligned}\mathbf{P}(\limsup\{|S_n - x| < r\}) &\geqslant \mathbf{P}(\sigma < \infty, \limsup\{|S_{\sigma+n} - S_\sigma| < r/2\})\\ &= \mathbf{P}(\sigma < \infty)\mathbf{P}(\limsup\{|S_n| < r/2\}) > 0.\end{aligned}$$

Stąd z prawa 0-1 Hewitta-Savage'a i dowolności r dostajemy $x \in R$.

By zakończyć dowód w tym przypadku wystarczy wykazać $-x \in A$. Wiemy, że $\sigma < \infty$ p.n., więc na podstawie mocnej własności Markowa,

$$\begin{aligned}\mathbf{P}(\limsup\{|S_n + x| < r\}) &= \mathbf{P}(\limsup\{|S_{\sigma+n} - S_\sigma + x| < r\})\\ &\geqslant \mathbf{P}(\limsup\{|S_{\sigma+n}| < r/2\}) = \mathbf{P}(\limsup\{|S_n| < r/2\}).\end{aligned}$$

Przypadek II. $|S_n| \to \infty$ p.n..

Ustalmy liczby naturalne m, k, wtedy

$$\mathbf{P}\left(|S_m| < r, \inf_{n\geqslant k} |S_{m+n}| \geqslant r\right) \geqslant \mathbf{P}\left(|S_m| < r, \inf_{n\geqslant k} |S_{m+n} - S_m| \geqslant 2r\right)$$
$$= \mathbf{P}(|S_m| < r)\mathbf{P}\left(\inf_{n\geqslant k} |S_n| \geqslant 2r\right).$$

Zdarzenie $\{|S_m| < r, \inf_{n\geqslant k} |S_{m+n}| \geqslant r\}$ może się powtórzyć tylko k razy, więc

$$\mathbf{P}\left(\inf_{n\geqslant k} |S_n| \geqslant 2r\right) \sum_{m=0}^{\infty} \mathbf{P}(|S_m| < r) \leqslant k < \infty.$$

Ponieważ $\mathbf{P}(\inf_{n\geqslant k} |S_n| \geqslant 2r)$ dąży do 1 przy $k \to \infty$, to $\sum_{m=0}^{\infty} \mathbf{P}(|S_m| < r) < \infty$. Stąd otrzymujemy, że $M \cap B(0, r) = \emptyset$ i z dowolności $r > 0$, $M = \emptyset$, co implikuje też, że $R = \emptyset$. $\square$

Wykażemy teraz użyteczny Lemat.

Lemat 6.7. *Dla dowolnego błądzenia losowego (S_n) w $\mathbb{R}^d$,*

$$\sum_{n=0}^{\infty} \mathbf{P}(|S_n| < r\varepsilon) \leqslant (5r)^d \sum_{n=0}^{\infty} \mathbf{P}(|S_n| < \varepsilon) \quad r \geqslant 1, \varepsilon > 0.$$

Dowód. Niech $x_1, \ldots, x_m$ oznacza maksymalny zbiór punktów w kuli $B(0, r\varepsilon)$ z których każde dwa mają odległość przynajmniej $\varepsilon/2$. Wtedy kule $B(x_i, \varepsilon/4)$ są rozłączne i są zawarte w kuli $B(0, \varepsilon(r+1/4))$. Stąd z porównania objętości wynika, że $m(r\varepsilon/4)^d \leqslant (\varepsilon(r+1/4))^d$, czyli $m \leqslant (4r+1)^d \leqslant (5r)^d$. Zatem kulę $B(0, r\varepsilon)$ można pokryć kulami otwartymi $B_1, \ldots, B_m$ o promieniu $\varepsilon/2$, przy czym $m \leqslant (5r)^d$.

Niech $\tau_k := \inf\{n\colon S_n \in B_k\}$, na podstawie mocnej własności Markowa dostajemy

$$\sum_{n=0}^{\infty} \mathbf{P}(|S_n| < r\varepsilon) \leqslant \sum_{k=1}^{m} \sum_{n=0}^{\infty} \mathbf{P}(S_n \in B_k) \leqslant \sum_{k=1}^{m} \sum_{n=0}^{\infty} \mathbf{P}(|S_{n+\tau_k} - S_{\tau_k}| < \varepsilon, \tau_k < \infty)$$
$$= \sum_{k=1}^{m} \mathbf{P}(\tau_k < \infty) \sum_{n=0}^{\infty} \mathbf{P}(|S_n| < \varepsilon) \leqslant (5r)^d \sum_{n=0}^{\infty} \mathbf{P}(|S_n| < \varepsilon).$$

$\square$

Wniosek 6.8. *Błądzenie losowe (S_n) jest powracalne wtedy i tylko wtedy, gdy dla pewnego $\varepsilon > 0$ (równoważnie dowolnego $\varepsilon > 0$) $\sum_{n=0}^{\infty} \mathbf{P}(|S_n| < \varepsilon) = \infty$.*

Dowód. Jeśli błądzenie jest powracalne, to z Twierdzenia 6.6 wiemy, że $0 \in M$, czyli $\sum_{n=0}^{\infty} \mathbf{P}(|S_n| < \varepsilon) = \infty$ dla każdego $\varepsilon > 0$.

Jeśli $\sum_{n=0}^{\infty} \mathbf{P}(|S_n| < \varepsilon) = \infty$ dla pewnego $\varepsilon > 0$, to z Lematu 6.7 wynika, że tak jest dla wszystkich $\varepsilon > 0$, a zatem $0 \in M$. $\square$

Twierdzenie 6.9. *Następujące warunki są wystarczające na to, by błądzenie losowe S_n było powracające:*
i) $d = 1$ oraz $\frac{1}{n}S_n \to 0$ według prawdopodobieństwa;
ii) $d \leqslant 2$, $\mathbf{E}X = 0$ oraz $\mathbf{E}|X|^2 < \infty$.

Dowód. i) Dla $\varepsilon > 0$ i $r \geqslant 1$ mamy z Lematu 6.7,

$$\sum_{n=0}^{\infty} \mathbf{P}(|S_n| < \varepsilon) \geqslant \frac{1}{5r} \sum_{n=0}^{\infty} \mathbf{P}(|S_n| < r\varepsilon) = \frac{1}{5} \int_0^{\infty} \mathbf{P}(|S_{\lfloor rt \rfloor}| < r\varepsilon) \mathrm{d}t.$$

Na mocy Lematu Fatou

$$\sum_{n=0}^{\infty} \mathbf{P}(|S_n| < \varepsilon) \geqslant \frac{1}{5} \int_0^{\infty} \liminf_{r\to\infty} \mathbf{P}(|S_{\lfloor rt \rfloor}| < r\varepsilon) \mathrm{d}t = \int_0^{\infty} 1 \mathrm{d}t = \infty.$$

ii) Analogicznie jak w i) dostajemy

$$\sum_{n=0}^{\infty} \mathbf{P}(|S_n| < \varepsilon) \geqslant \frac{1}{25r^2} \sum_{n=0}^{\infty} \mathbf{P}(|S_n| < r\varepsilon) = \frac{1}{25} \int_0^{\infty} \mathbf{P}(|S_{\lfloor r^2 t \rfloor}| < r\varepsilon) \mathrm{d}t.$$

Na mocy CTG $n^{-1/2}S_n$ zbiega według rozkładu do zmiennej G o rozkładzie normalnym (być może zdegenerowanym). Łatwo widać, że istnieje $c, \delta > 0$ takie, że $\mathbf{P}(|G| \leqslant t) \geqslant ct^2$ dla $0 < t \leqslant \delta$. Stąd z Lematu Fatou,

$$\begin{aligned}\sum_{n=0}^{\infty} \mathbf{P}(|S_n| < \varepsilon) &\geqslant \frac{1}{25} \int_0^{\infty} \liminf_{r\to\infty} \mathbf{P}(|S_{\lfloor r^2 t \rfloor}| < r\varepsilon) \mathrm{d}t = \frac{1}{25} \int_0^{\infty} \mathbf{P}(|G| < \varepsilon t^{-1/2}) \mathrm{d}t \\ &\geqslant \frac{c\varepsilon^2}{25} \int_{(\varepsilon/\delta)^2}^{\infty} t^{-1} \mathrm{d}t = \infty.\end{aligned}$$

□

Twierdzenie 6.10 (Chung-Fuchs)**.** *Załóżmy, że φ jest funkcją charakterystyczną zmiennej X. Wówczas błądzenie S_n jest powracające wtedy i tylko wtedy, gdy dla pewnego (równoważnie dla każdego) $\varepsilon > 0$,*

$$\sup_{0<r<1} \int_{|t|\leqslant\varepsilon} \mathrm{Re}\left(\frac{1}{1 - r\varphi(t)}\right) \mathrm{d}t = \infty.$$

Lemat 6.11. *Dla dowolnych miar probabilistycznych μ, ν na $\mathbb{R}^d$,*

$$\int_{\mathbb{R}^d} \varphi_\mu \mathrm{d}\nu = \int_{\mathbb{R}^d} \varphi_\nu \mathrm{d}\mu.$$

Dowód. Na mocy Twierdzenia Fubiniego

$$\int_{\mathbb{R}^d} \varphi_\mu(t)\mathrm{d}\nu(t) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{i\langle t,s\rangle}\mathrm{d}\mu(s)\mathrm{d}\nu(t) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{i\langle t,s\rangle}\mathrm{d}\nu(t)\mathrm{d}\mu(s) = \int_{\mathbb{R}^d} \varphi_\nu(s)\mathrm{d}\mu(s).$$

□

Dowód Twierdzenia 6.10. . Miara ν o gęstości $g(x) = (1-|x|)_+$ ma funkcję charakterystyczną $\hat{g}(t) = \varphi_\nu(t) = 2t^{-2}(1-\cos t)$. Zatem miara produktowa o gęstości $a^d g^{\otimes d}(ax) = a^d \prod_{k=1}^d g(ax_k)$ ma funkcję charakterystyczną $\hat{g}^{\otimes d}(t/a) = \prod_{k=1}^d \hat{g}(t_k/a)$. Z Lematu 6.11 otrzymujemy dla $a > 0$,

$$\int_{\mathbb{R}^d} \hat{g}^{\otimes d}(x/a)\mathrm{d}\mu_{S_n}(x) = a^d \int_{\mathbb{R}^d} \varphi(t)^n g^{\otimes d}(at)\mathrm{d}t.$$

Twierdzenie Fubiniego implikuje, że dla dowolnego $r \in (0,1)$,

$$\int_{\mathbb{R}^d} \hat{g}^{\otimes d}(x/a)\sum_{n=0}^\infty r^n \mathrm{d}\mu_{S_n}(x) = a^d \int_{\mathbb{R}^d}\sum_{n=0}^\infty r^n\varphi(t)^n g^{\otimes d}(at)\mathrm{d}t = a^d\int_{\mathbb{R}^d}\frac{g^{\otimes d}(at)}{1-r\varphi(t)}\mathrm{d}t. \tag{13}$$

Ustalmy $\varepsilon > 0$ i załóżmy, że $\sup_{0<r<1}\int_{|t|\leqslant\varepsilon}\mathrm{Re}(\frac{1}{1-r\varphi(t)})\mathrm{d}t < \infty$. Zauważmy, że $\hat{g}$ jest ciągłe i nie zeruje się na $[-1,1]$ zatem $\hat{g}(t) \geqslant c > 0$ dla $|t| \leqslant 1$, stąd

$$\begin{aligned}\sum_{n\geqslant 0}\mathbf{P}(|S_n|<\delta) &= \sup_{0<r<1}\sum_{n\geqslant 0} r^n \mu_{S_n}(B(0,\delta)) \leqslant \sup_{0<r<1} c^{-d}\int_{\mathbb{R}^d}\hat{g}^{\otimes d}(x/\delta)\sum_{n=0}^\infty r^n\mathrm{d}\mu_{S_n}(x)\\ &= (\delta/c)^d \sup_{0<r<1}\int_{\mathbb{R}^d}\frac{g^{\otimes d}(\delta t)}{1-r\varphi(t)}\mathrm{d}t.\end{aligned}$$

Zauważmy, że $g^{\otimes d}(\delta t)$ się zeruje poza kostką $[-\delta^{-1},\delta^{-1}]^n \subset B(0, d^{1/2}\delta^{-1})$. Stąd przyjmując $\delta = d^{1/2}\varepsilon^{-1}$ dostajemy

$$\sum_{n\geqslant 0}\mathbf{P}(|S_n|<\delta) \leqslant C(d,\varepsilon)\sup_{0<r<1}\int_{B(0,\varepsilon)}\frac{1}{1-r\varphi(t)}\mathrm{d}t < \infty,$$

czyli błądzenie losowe jest chwilowe na mocy Wniosku 6.8.

By udowodnić przeciwną implikację zauważmy, że (na mocy twierdzenia o odwracaniu transformaty Fouriera) miara z gęstością $(2\pi)^{-d}\hat{g}^{\otimes d}(x)$ ma funkcję charakterystyczną $g^{\otimes d}(t)$, więc analogicznie jak (13) dowodzimy, że

$$\int_{\mathbb{R}^d} g^{\otimes d}(x/a)\sum_{n=0}^\infty r^n\mathrm{d}\mu_{S_n}(x) = (a/(2\pi))^d\int_{\mathbb{R}^d}\frac{\hat{g}^{\otimes d}(at)}{1-r\varphi(t)}\mathrm{d}t.$$

Stąd, jeśli S_n jest chwilowy, to

$$\begin{aligned}\sup_{0<r<1}\int_{B(0,\varepsilon)}\frac{1}{1-r\varphi(t)}\mathrm{d}t &\leqslant c^{-d}\sup_{0<r<1}\int_{\mathbb{R}^d}\frac{\hat{g}^{\otimes d}(t/\varepsilon)}{1-r\varphi(t)}\mathrm{d}t\\ &= (2\pi\varepsilon/c)^d\sup_{0<r<1}\int_{\mathbb{R}^d}g^{\otimes d}(\varepsilon x)\sum_{n=0}^{\infty}r^n\mathrm{d}\mu_{S_n}(x)\\ &= (2\pi\varepsilon/c)^d\sup_{0<r<1}\int_{|x|<d^{1/2}/\varepsilon}g^{\otimes d}(\varepsilon x)\sum_{n=0}^{\infty}r^n\mathrm{d}\mu_{S_n}(x)\\ &\leqslant (2\pi\varepsilon/c)^d\sum_{n=0}^{\infty}\mu_{S_n}(B(0,d^{1/2}\varepsilon^{-1})<\infty.\end{aligned}$$

□

Wniosek 6.12. *Ustalmy $\varepsilon > 0$. Wówczas błądzenie (S_n) jest powracalne, jeśli*

$$\int_{|t|\leqslant\varepsilon}\mathrm{Re}\left(\frac{1}{1-\varphi(t)}\right)\mathrm{d}t=\infty$$

oraz chwilowe, jeśli

$$\int_{|t|\leqslant\varepsilon}\frac{1}{1-\mathrm{Re}(\varphi(t))}\mathrm{d}t<\infty.$$

W szczególności błądzenie symetryczne jest powracalne wtedy i tylko wtedy, gdy

$$\int_{|t|\leqslant\varepsilon}\frac{1}{1-\varphi(t)}\mathrm{d}t=\infty.$$

Wniosek 6.13. *Niech (S_n') oznacza niezależną kopię błądzenia losowego (S_n). Wówczas, jeśli (S_n) jest powracające, to $(S_n - S_n')$ też jest powracające.*

Dowody obu wniosków pozostawiamy jako proste ćwiczenie.

Twierdzenie 6.14. *Jeśli nośnik zmiennej X jest więcej niż dwuwymiarowy, to błądzenie S_n jest chwilowe.*

Dowód. Załóżmy wpierw, że X jest symetryczny i ma nośnik wymiaru $k > 2$. Ewentualnie obcinając błądzenie do odpowiedniej podprzestrzeni liniowej wymiaru k, możemy zakładać, że $k = d$. Wówczas dla odpowiednio dużego M zmienna ograniczona $X_M := X\mathbb{1}_{|X|\leqslant M}$ też ma nośnik pełnowymiarowy, czyli ma niezdegenerowaną macierz kowariancji C_M. Zatem przy $t \to 0$

$$1-\varphi_X(t)=\mathbf{E}(1-\cos(\langle t,X\rangle))\geqslant\mathbf{E}(1-\cos(\langle t,X_M\rangle))=1-\varphi_{X_M}(t)=\frac{1}{2}\langle C_M t,t\rangle+o(|t|),$$

stąd istnieje $\delta > 0$ takie, że $1 - \varphi_X(t) \geqslant \delta|t|^2$ dla t odpowiednio małych. Dostajemy

$$\sup_{0<r<1} \int_{|t|\leqslant\varepsilon} \frac{1}{1 - r\varphi(t)} \mathrm{d}t \leqslant \int_{|t|\leqslant\varepsilon} \frac{1}{1 - \varphi(t)} \mathrm{d}t \leqslant \int_{|t|\leqslant\varepsilon} \frac{1}{\delta|t|^2} \mathrm{d}t < \infty,$$

a zatem błądzenie jest chwilowe.

W przypadku ogólnym załóżmy, że (S_n) jest powracające, wówczas $(S_n - S'_n)$ też jest powracające, zatem $X - X'$ ma nośnik wymiaru conajwyżej dwa. Wynika stąd, że X zawiera się w pewnej przestrzeni afinicznej wymiaru 2. Jeśli ta przestrzeń nie zawiera 0, to istnieje $t \in \mathbb{R}^d$ takie, że $\langle t, X \rangle = 1$, co łatwo przeczy powracalności (S_n). □

6.3 Punkty drabinowe. Faktoryzacja Wienera-Hopfa

W tej części będziemy rozpatrywali błądzenia jednowymiarowe. Zacznijmy od użytecznej definicji punktów drabinowych.

Definicja 6.15. Załóżmy, że (S_n) jest jednowymiarowym błądzeniem losowym. Określamy indukcyjnie ciąg $(\tau_k)_{k\geqslant 0}$ wzorem $\tau_0 = 0$ oraz

$$\tau_k := \inf\{n \geqslant \tau_{k-1} + 1 \colon \ S_n > S_{\tau_{k-1}}\}. \tag{14}$$

Moment zatrzymania τ_k nazywamy *k-tym (wstępującym) momentem drabinowym* Zmienną S_{τ_k} nazywamy *k-tą (wstępującą) wysokością drabinową*, a parę (τ_k, S_{τ_k}) – *k-tym (wstępującym) punktem drabinowym.* Zamieniając w (14) nierówność $>$ na $<$ otrzymujemy *k–te zstępujące momenty drabinowe* τ_k^- i odpowiadające im *zstępujące wysokości drabinowe* $S_{\tau_k^-}$ i *zstępujące punkty drabinowe* $(\tau_k^-, S_{\tau_k^-})$. Podobnie, zamieniając w (14) $>$ na $\geqslant$ (odp. $\leqslant$) dostajemy *słabe wstępujące momenty drabinowe* σ_n, *słabe wstępujące wysokości drabinowe* S_{σ_n} oraz *słabe wstępujące punkty drabinowe* (σ_n, S_{σ_n}) (odp. *słabe zstępujące momenty drabinowe* σ_n^-, *słabe zstępujące wysokości drabinowe* $S_{\sigma_n^-}$ oraz *słabe zstępujące punkty drabinowe* $(\sigma_n^-, S_{\sigma_n^-})$).

Uwaga 6.16. Mamy $\mathbf{P}(\tau_n < \infty) = \mathbf{P}(\tau_1 < \infty)^n$. Ponadto, jeśli $\tau_1 < \infty$ p.n., to na podstawie mocnej własności Markowa wstępujące punkty drabinowe (τ_n, S_{τ_n}) tworzą dwuwymiarowe błądzenie losowe. Podobne fakty zachodzą dla punktów zstępujących i słabych punktów wstępujących/zstępujących.

Następny lemat o dualności mówi, że miary czasu przebywania dla procesów $(S_{\tau_n})_{n\geqslant 0}$ (odp. $(S_{\sigma_n})_{n\geqslant 0}$) mają takie same wartości oczekiwane jak miary czasu przebywania dla procesu $(S_n)_{0\leqslant n<\sigma_1^-}$ (odp. $(S_n)_{0\leqslant n<\tau_1^-}$).

Lemat 6.17. *Dla dowolnego zbioru borelowskiego B*

$$\sum_{n=0}^{\infty} \mathbf{P}(S_{\tau_n} \in B,\ \tau_n < \infty) = \sum_{n=0}^{\infty} \mathbf{P}(S_n \in B,\ n < \sigma_1^-)$$

oraz

$$\sum_{n=0}^{\infty}\mathbf{P}(S_{\sigma_n}\in B,\ \sigma_n<\infty)=\sum_{n=0}^{\infty}\mathbf{P}(S_n\in B,\ n<\tau_1^-).$$

Dowód. Udowodnimy pierwszy ze wzorów, drugi dowodzi się analogicznie. Na mocy definicji momentów drabinowych wystarczy go oczywiście sprawdzić dla $B\in\mathcal{B}(0,\infty)$. Zauważmy, że

$$(S_1,\ldots,S_n) \text{ ma ten sam rozkład co } (S_n-S_{n-1},\ldots,S_n-S_0), \tag{15}$$

zatem dla $B\in\mathcal{B}(0,\infty)$,

$$\begin{aligned}\mathbf{P}(\min\{S_1,\ldots,S_{n-1}\}>0,S_n\in B)&=\mathbf{P}(\max\{S_1,\ldots,S_{n-1}\}<S_n,S_n\in B)\\&=\sum_{k\geqslant 0}\mathbf{P}(\tau_k=n,S_{\tau_k}\in B).\end{aligned}\tag{16}$$

Sumując po $n\geqslant 1$ otrzymujemy pierwszy ze wzorów. □

Fakt 6.18. *Dla niezdegenerowanego jednowymiarowego błądzenia losowego* (S_n) *zachodzi dokładnie jeden z trzech warunków:*
i) $S_n\to\infty$ *p.n.,* $\mathbf{E}\sigma_1\leqslant\mathbf{E}\tau_1<\infty$ *oraz* $\mathbf{E}\sigma_1^-=\mathbf{E}\tau_1^-=\infty$*,*
ii) $S_n\to-\infty$ *p.n.,* $\mathbf{E}\sigma_1=\mathbf{E}\tau_1=\infty$ *oraz* $\mathbf{E}\sigma_1^-\leqslant\mathbf{E}\tau_1^-<\infty$*,*
iii) $\limsup\pm S_n=\infty$ *p.n. oraz* $\mathbf{E}\sigma_1=\mathbf{E}\tau_1=\mathbf{E}\sigma_1^-=\mathbf{E}\tau_1^-=\infty$.

Dowód. Istnieje takie $A>-\infty$, że $\mathbf{P}(\limsup\{S_n<A\})<1$ lub istnieje takie $A<\infty$, że $\mathbf{P}(\limsup\{S_n>A\})<1$ lub $\limsup\pm S_n=\infty$ p.n.. W pierwszym przypadku, $R\subset[A,\infty)$, stąd błądzenie jest chwilowe, czyli $|S_n|\to\infty$ p.n.. Ponieważ, z prawa 0-1, $\mathbf{P}(\limsup\{S_n<A\})=0$, więc $S_n\to\infty$ p.n.. Analogicznie, w drugim przypadku $S_n\to-\infty$ p.n..

Załóżmy, że zachodzi drugi przypadek, czyli $S_n\to-\infty$ p.n.. Niech

$$Z=\inf\{n:\ \sigma_n=\infty\}.$$

Oczywiście $Z<\infty$ p.n., ponadto Z ma rozkład geometryczny z parametrem $p=\mathbf{P}(\sigma_1<\infty)<1$. Stosując drugą równość Lematu 6.17 do $B=\mathbb{R}$ otrzymujemy $\infty>\mathbf{E}Z=\mathbf{E}\tau_1^-$.

Jeśli $\limsup\pm S_n=\infty$ p.n., to $\tau_n<\infty$ p.n. i $\tau_n^-<\infty$ p.n. dla wszystkich n, więc stosując Lemat 6.17 do $B=\mathbb{R}$ dostajemy $\mathbf{E}\sigma_1=\mathbf{E}\sigma_1^-=\infty$. □

W kolejnym fakcie używamy konwencji, że $\mathbf{E}X=\mathbf{E}X_+-\mathbf{E}X_-$ można określić, jeśli przynajmniej jedna z wielkości $\mathbf{E}X_+$, $\mathbf{E}X_-$ jest skończona

Fakt 6.19. *Niech* (S_n) *będzie niezdegenerowanym, jednowymiarowym błądzeniem losowym generowanym przez zmienną* X.
i) jeśli $\mathbf{E}X=0$*, to* $\limsup\pm S_n=\infty$ *p.n.;*
ii) jeśli $0<\mathbf{E}X\leqslant\infty$*, to* $S_n\to\infty$ *p.n. oraz* $\mathbf{E}S_{\tau_1}=\mathbf{E}\tau_1\mathbf{E}X$*;*
iii) jeśli $\mathbf{E}X_+=\mathbf{E}X_-=\infty$*, to* $\mathbf{E}S_{\tau_1}=-\mathbf{E}S_{\tau_1^-}=\infty$.

Dowód. i) Z symetrii i Faktu 6.18 możemy zakładać, że $\limsup S_n = \infty$. Gdyby $\mathbf{E}\tau_1 < \infty$, to na mocy tożsamości Walda $\mathbf{E}S_{\tau_1} = \mathbf{E}X\mathbf{E}\tau_1 = 0$, co przeczy temu, że $S_{\tau_1} > 0$ p.n.. Zatem $\mathbf{E}\tau_1 = \infty$, co na mocy Faktu 6.18 implikuje, że $\liminf S_n = -\infty$.

ii) $S_n \to \infty$ na mocy MPWL, a druga równość wynika z tożsamości Walda.

iii) Wystarczy zauważyć, że $S_{\tau_1} \geqslant (X_1)_+$ oraz $S_{\tau_1^-} \leqslant -(X_1)_-$. □

Zanim sformułujemy bardzo ważne twierdzenie Wienera-Hopfa, będziemy potrzebowali kolejnej definicji.

Definicja 6.20. Przez χ^+ (odp. χ^-, ψ^+, ψ^-) oznaczamy (być może zdegenerowany) rozkład punktu drabinowego (τ_1, S_{τ_1}) (odp. $(\tau_1^-, S_{\tau_1^-})$, (σ_1, S_{σ_1}), $(\sigma_1^-, S_{\sigma_1^-})$), tzn. dla $B \in \mathcal{B}(\mathbb{N} \times \mathbb{R})$,

$$\begin{aligned}
\chi^+(B) &= \mathbf{P}(\tau_1 < \infty,\ (\tau_1, S_{\tau_1}) \in B),\\
\chi^-(B) &= \mathbf{P}(\tau_1^- < \infty,\ (\tau_1^-, S_{\tau_1^-}) \in B),\\
\psi^+(B) &= \mathbf{P}(\sigma_1 < \infty,\ (\sigma_1, S_{\sigma_1}) \in B),\\
\psi^-(B) &= \mathbf{P}(\sigma_1^- < \infty,\ (\sigma_1^-, S_{\sigma_1^-}) \in B).
\end{aligned}$$

Definiujemy też miary $\chi_n^\pm$ i $\psi_n^\pm$ wzorem

$$\chi_n^\pm(A) = \chi^\pm(\{n\} \times A), \quad \psi_n^\pm(A) = \psi^\pm(\{n\} \times A), \quad n \geqslant 1,\ A \in \mathcal{B}(\mathbb{R}).$$

Wreszcie określamy miarę χ^0 na $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$ jako

$$\chi^0(\{n\}) := \mathbf{P}(\min\{S_1, \ldots, S_{n-1}) > 0, S_n = 0) = \mathbf{P}(\max\{S_1, \ldots, S_{n-1}) < 0, S_n = 0),$$

gdzie druga równość wynika z (15). Miarę χ^0 będziemy też traktować jako miarę na $\mathbb{N} \times \mathbb{R}$ (skoncentrowaną na $\mathbb{N}^* \times \{0\}$), zaś przez χ_n^0 będziemy oznaczać zarówno miarę na $\mathbb{R}$ równą $\chi^0(\{n\})\delta_0$ jak i liczbę $\chi^0(\{n\})$.

Uwaga 6.21. Z mocnej własności Markowa łatwo wynika, że k-ty wstępujący punkt drabinowy (τ_k, S_{τ_k}) ma (być może zdegenerowany) rozkład χ_+^{*k}. Podobne fakty zachodzą dla pozostałych rodzajów punktów drabinowych.

Twierdzenie 6.22 (Faktoryzacja Wienera-Hopfa)**.** *Dla dowolnego jednowymiarowego błądzenia losowego zachodzą następujące tożsamości dotyczące miar na* $\mathbb{N} \times \mathbb{R}$*:*

$$\delta_0 - \delta_1 \otimes \mu = (\delta_0 - \chi^+) * (\delta_0 - \psi^-) = (\delta_0 - \psi^+) * (\delta_0 - \chi^-), \tag{17}$$

$$\delta_0 - \psi^\pm = (\delta_0 - \chi^\pm) * (\delta_0 - \chi^0). \tag{18}$$

Dowód. Określmy miarę $\rho = (\rho_n)_{n \geqslant 0}$ na $\mathbb{N} \times \mathbb{R}$ wzorem $\rho_0 = \delta_0$ oraz

$$\rho_n(B) = \mathbf{P}(\min\{S_1, \ldots, S_n\} > 0,\ S_n \in B).$$

Zauważmy, że ρ_n zeruje się na $(-\infty, 0]$, ponadto dla $n \geqslant 1$,

$$(\rho_n + \psi_n^-)(B) = \mathbf{P}(\min\{S_1, \ldots, S_{n-1}\} > 0,\ S_{n-1} + X_n \in B) = \rho_{n-1} * \mu(B).$$

Stąd

$$\rho_n + \psi_n^- = \rho_{n-1} * \mu = (\rho * (\delta_1 \otimes \mu))_n, \quad n \geqslant 1,$$

zatem (po rozpatrzeniu co się dzieje dla $n = 0$)

$$\rho + \psi^- = \delta_0 + \rho * (\delta_1 \otimes \mu), \text{ czyli } \quad \rho * (\delta_0 - \delta_1 \otimes \mu) = \delta_0 - \psi^-.$$

Zauważmy, że ze wzoru (16) dostajemy dla $n \geqslant 1$

$$\begin{aligned}\rho_n &= \sum_{k \geqslant 1} ((\chi^+)^{*k})_n = \sum_{k \geqslant 1} \sum_{l=1}^{n} \chi_l^+ * ((\chi^+)^{*(k-1)})_{n-l} = \sum_{l=1}^{n} \chi_l^+ * \sum_{k \geqslant 1} ((\chi^+)^{*(k-1)})_{n-l} \\ &= \sum_{l=1}^{n} \chi_l^+ * \rho_{n-l} = (\chi^+ * \rho)_n.\end{aligned}$$

Rozpatrując dodatkowo $n = 0$ dostajemy

$$\rho = \delta_0 + \chi^+ * \rho, \text{ czyli } \quad (\delta_0 - \chi^+) * \rho = \delta_0.$$

Mamy zatem

$$\delta_0 - \delta_1 \otimes \mu = \delta_0 * (\delta_0 - \delta_1 \otimes \mu) = (\delta_0 - \chi^+) * \rho * (\delta_0 - \delta_1 \otimes \mu) = (\delta_0 - \chi^+) * (\delta_0 - \psi^-).$$

Druga część równości (17) wynika z pierwszej oraz symetrii.

By udowodnić (18) zauważmy, że dla $B \in \mathcal{B}(0, \infty)$ i $n \geqslant 1$,

$$\begin{aligned}(\chi_n^+ - \psi_n^+ + \chi_n^0)(B) &= \mathbf{P}((\{\tau_1 = n\} \setminus \{\sigma_1 = n\}) \cap \{S_n \in B\}) = \mathbf{P}\left(\max_{1 \leqslant k < n} S_k = 0,\ S_n \in B\right) \\ &= \sum_{l=1}^{n-1} \mathbf{P}\left(\max_{1 \leqslant k < l} S_k < 0 = S_l,\ \max_{l \leqslant k < n} (S_k - S_l) \leqslant 0,\ S_n - S_l \in B\right) \\ &= \sum_{l=1}^{n-1} \mathbf{P}\left(\max_{1 \leqslant k < l} S_k < 0 = S_l\right) \mathbf{P}\left(\max_{k < n-l} S_k \leqslant 0,\ S_{n-l} \in B\right) \\ &= \sum_{l=1}^{n-1} \chi_l^0 \chi_{n-l}^+(B) = (\chi^0 * \chi^+)_n(B).\end{aligned}$$

Łatwo sprawdzić, że dla $B \subset (-\infty, 0]$ $(\chi_n^+ - \psi_n^+ + \chi_n^0)(B) = 0 = (\chi^0 * \chi^+)_n(B)$. Stąd $\chi^+ - \psi^+ + \chi^0 = \chi^0 * \chi^+$, co dowodzi „plusowej” części (18). Część „minusową” dowodzimy podobnie. $\square$

Twierdzenie 6.23. *Jeśli (S_n) jest niezdegenerowanym, jednowymiarowym błądzeniem losowym, to dla $|s| < 1$ i $u \geqslant 0$*

$$\mathbf{E}\left[s^{\tau_1}\exp(-uS_{\tau_1})\mathbb{1}_{\{\tau_1<\infty\}}\right] = 1 - \exp\left(-\sum_{n=1}^{\infty}\frac{s^n}{n}\mathbf{E}\left[e^{-uS_n}\mathbb{1}_{\{S_n>0\}}\right]\right).$$

Analogiczny wzór zachodzi dla słabego punktu drabinowego (σ_1, S_{σ_1}) – zdarzenie $S_n > 0$ należy zastąpić zdarzeniem $S_n \geqslant 0$.

Dowód. Określmy

$$F(s,t) := \mathbf{E}\left[s^{\tau_1}\exp(itS_{\tau_1})\mathbb{1}_{\{\tau_1<\infty\}}\right], \quad G(s,t) := \mathbf{E}\left[s^{\sigma_1^-}\exp(itS_{\sigma_1^-})\mathbb{1}_{\{\sigma_1^-<\infty\}}\right].$$

Całkując funkcję $\mathbb{N}\times\mathbb{R} \ni (n,x) \mapsto s^n\exp(itx)$ względem obu stron (17) dostajemy

$$1 - s\varphi_\mu(t) = (1-F(s,t))(1-G(s,t)), \qquad \text{dla } |s|<1,\ t\in\mathbb{R}.$$

Logarytmując obie strony i rozwijając w szereg Taylora otrzymujemy

$$\sum_{n\geqslant 1}\frac{1}{n}(s\varphi_\mu(t))^n = \sum_{n\geqslant 1}\frac{1}{n}F(s,t)^n + \sum_{n\geqslant 1}\frac{1}{n}G(s,t)^n.$$

Ustalmy $s\in(-1,1)$, wówczas zauważamy, że obie strony powyższej równości są transformatami Fouriera miar (ze znakiem dla $s<0$) na $\mathbb{R}$, stąd te miary muszą się pokrywać, tzn.

$$\sum_{n\geqslant 1}\frac{1}{n}s^n\mu^{*n} = \sum_{n\geqslant 1}\frac{1}{n}\left(\sum_{k\geqslant 1}s^k\chi_k^+\right)^{*n} + \sum_{n\geqslant 1}\frac{1}{n}\left(\sum_{k\geqslant 1}s^k\psi_k^-\right)^{*n}.$$

Zauważmy, że pierwsza z miar po prawej stronie równości ma nośnik w $(0,\infty)$ a druga w $(-\infty,0]$. Ustalając $u\geqslant 0$ i całkując funckję $\mathbb{R}\ni t\to\exp(-ut)\mathbb{1}_{t>0}$ względem obu stron dostajemy

$$\begin{aligned}\sum_{n\geqslant 1}\frac{s^n}{n}\mathbf{E}\left[e^{-uS_n}\mathbb{1}_{\{S_n>0\}}\right] &= \sum_{n\geqslant 1}\frac{1}{n}\left(\mathbf{E}s^{\tau_1}\exp(-uS_{\tau_1})\mathbb{1}_{\{\tau_1<\infty\}}\right)^n \\ &= -\ln\left(1-\mathbf{E}s^{\tau_1}\exp(-uS_{\tau_1})\mathbb{1}_{\{\tau_1<\infty\}}\right).\end{aligned}$$

Wzór dla słabego punktu drabinowego dowodzimy analogicznie. □

Wniosek 6.24. *Dla dowolnego jednowymiarowego błądzenia losowego (S_n),*

$$\mathbf{P}(\tau_1=\infty) = \frac{1}{\mathbf{E}\sigma_1^-} = \exp\left(-\sum_{n\geqslant 1}\frac{1}{n}\mathbf{P}(S_n>0)\right),$$

$$\mathbf{P}(\sigma_1 = \infty) = \frac{1}{\mathbf{E}\tau_1^-} = \exp\left(-\sum_{n\geqslant 1} \frac{1}{n}\mathbf{P}(S_n \geqslant 0) \right),$$

Ponadto, każdy z następujących warunków jest równoważny temu, że $S_n \to -\infty$ p.n.:

$$\sum_{n\geqslant 1} \frac{1}{n}\mathbf{P}(S_n > 0) < \infty, \quad \sum_{n\geqslant 1} \frac{1}{n}\mathbf{P}(S_n \geqslant 0) < \infty.$$

Dowód. Biorąc w Lemacie 6.17 $B = \mathbb{R}$ i korzystając z tego, że $\mathbf{P}(\tau_n < \infty) = \mathbf{P}(\tau_1 < \infty)^n$ oraz $\mathbf{P}(\sigma_n < \infty) = \mathbf{P}(\sigma_1 < \infty)^n$ dostajemy pierwsze z równości z obu wzorów.

Przyjmując w Twierdzeniu 6.23 $u = 0$ i biorąc $s \to 1-$ dostajemy drugie części wzorów na $\mathbf{P}(\tau_1 = \infty)$ i $\mathbf{P}(\sigma_1 = \infty)$. Stąd też mamy, że zbieżność szeregów z ostatniej części Twierdzenia jest równoważna $\mathbf{P}(\tau_1 = \infty) > 0$ bądź $\mathbf{P}(\sigma_1 = \infty) > 0$, a oba te fakty są równoważne temu, że $S_n \to -\infty$ p.n.. □

7 Elementy teorii odnowienia

W tej części będziemy rozważać jednowymiarowe błądzenia losowe z niekoniecznie zerowym warunkiem początkowym, tzn. błądzenia postaci

$$S_n = S_0 + X_1 + \ldots + X_n, \quad n \geqslant 0,$$

gdzie $S_0, X_1, \ldots$ są niezależnymi zmiennymi losowymi, przy czym zmienne X_i mają jednakowy rozkład. Rozkład $(S_n)_{n\geqslant 0}$ zależy wtedy od dwóch parametrów - rozkładu poczatkowego zmiennej S_0 i rozkładu zmiennych X_i, które będziemy dalej oznaczać przez ν i μ.

Przez η oznaczamy *proces czasu przebywania*, tzn. miarę losową na $\mathbb{R}$ daną jako

$$\eta = \sum_{n\geqslant 0} \delta_{S_n}.$$

W przypadku, gdy błądzenie jest nieujemne, η nazywa się również *procesem odnowienia.* Miarę (deterministyczną) $\mathbf{E}\eta$ nazywamy *intensywnością procesu czasu przebywania/miarą czasu przebywania* (*miarą odnowienia* w przypadku nieujemnym). Mamy

$$\mathbf{E}\eta(B) = \sum_{n\geqslant 0} \mathbf{P}(S_n \in B) \quad \text{oraz} \quad \mathbf{E}\eta = \nu * \sum_{n\geqslant 0} \mu^{*n}.$$

Będziemy rozważali chwilowe błądzenia losowe. Wówczas, jak wiemy z poprzednich rozważań, $|S_n| \to \infty$ p.n., η jest z prawdopodobieństwem 1 lokalnie skończoną miarą na $\mathbb{R}$, zaś $\mathbf{E}\eta$ lokalnie skończoną miarą na $\mathbb{R}$.

7.1 Stacjonarne procesy odnowienia

Definicja 7.1. Dla miary α na $\mathbb{R}$ i $t \geqslant 0$ przez $\theta_t\alpha$ będziemy oznaczali miarę α przesuniętą o t i obciętą do $\mathbb{R}_+$, tzn.

$$\theta_t\alpha(B) := \alpha(B+t) \quad \text{dla } B \in \mathcal{B}(\mathbb{R}_+).$$

Powiemy, że proces czasu przebywania η jest *stacjonarny* na $\mathbb{R}_+$, jeśli $\theta_t\eta$ ma ten sam rozkład, co $\theta_0\eta$ dla dowolnego $t \geqslant 0$.

Uwaga 7.2. i) Zauważmy, że $\theta_0\alpha$ to obcięcie α do $\mathbb{R}_+$.
ii) Dla $t \geqslant 0$ oznaczmy $\sigma(t) := \inf\{n\colon\ S_n \geqslant t\}$. Wówczas $\theta_t\eta$ ma ten sam rozkład co $\theta_0\eta_t$, gdzie η_t jest miarą czasu przebywania dla błądzenia losowego $(S_{\sigma(t)+n} - t)$.

Szczególnie istotna jest druga część powyższej uwagi. Wiąże się ona z nazwą teorii odnowienia – wielokrotnie będziemy bowiem rozważali momenty zatrzymania τ i nowe (odnowione) błądzenia losowe $(S_{\tau+n})_{n\geqslant 0}$, które, na mocy mocnej własności Markowa, też są błądzeniami losowymi, o takim samym rozkładzie przyrostów co (S_n).

Przez λ będziemy oznaczać miarę Lebesgue'a na $\mathbb{R}_+$. Łatwo udowodnić, że jeśli α jest lokalnie skończoną miarą na $\mathbb{R}_+$, niezmienniczą na przesunięcia z $\mathbb{R}_+$, to $\alpha = a\lambda$ dla pewnego $a \geqslant 0$.

Fakt 7.3. *Załóżmy, że błądzenie losowe ma przyrosty nieujemne o rozkładzie μ na $\mathbb{R}_+$ i średniej c. Wówczas istnieje rozkład początkowy ν na $\mathbb{R}_+$ taki, że proces odnowienia η jest stacjonarny wtedy i tylko wtedy, gdy $0 < c < \infty$. Ponadto, gdy η jest stacjonarny, to $\mathbf{E}\eta = c^{-1}\lambda$ oraz $\nu = c^{-1}(\delta_0 - \mu) * \lambda$, czyli*

$$\nu[0,t] = c^{-1}\int_0^t \mu(s,\infty)ds, \quad t \geqslant 0. \tag{19}$$

Dowód. Mamy

$$\mathbf{E}\eta = \sum_{n\geqslant 0} \nu * \mu^{*n} = \nu + \mu * \sum_{n\geqslant 0} \nu * \mu^{*n} = \nu + \mu * \mathbf{E}\eta,$$

stąd

$$\nu = (\delta_0 - \mu) * \mathbf{E}\eta. \tag{20}$$

Jeśli η jest stacjonarny, to $\mathbf{E}\eta$ jest niezmienniczą na przesunięcia, lokalnie skończoną miarą na $\mathbb{R}_+$, zatem $\mathbf{E}\eta = a\lambda$ dla pewnego $a \geqslant 0$. Zatem $\nu = a(\delta_0 - \mu) * \lambda$, skąd wynika, że zachodzi równość (19) z c^{-1} zastąpionym przez a. Biorąc $t \to \infty$ dostajemy $1 = ac$, czyli $a = c^{-1}$, w szczególności $c < \infty$.

Na odwrót załóżmy, że $0 < c < \infty$ i $\nu = c^{-1}(\delta_0 - \mu) * \lambda$ (tzn. zachodzi (19)). Mamy

$$\mathbf{E}\eta = \nu * \sum_{n\geqslant 0} \mu^{*n} = c^{-1}(\delta_0 - \mu) * \lambda * \sum_{n\geqslant 0} \mu^{*n} = c^{-1} * \lambda * \left(\sum_{n\geqslant 0} \mu^{*n} - \sum_{n\geqslant 1} \mu^{*n}\right) = c^{-1}\lambda.$$

W szczególności $\theta_t \mathbf{E}\eta = \mathbf{E}\eta$. Ustalmy $t > 0$, niech $\sigma(t) := \inf\{n\colon S_n \geqslant t\}$. Na mocy mocnej własności Markowa, $(S_{\sigma(t)+n} - t)_{n\geqslant 0}$ jest błądzeniem losowym o rozkładzie przyrostów μ i pewnym rozkładzie początkowym ν_t, ponadto proces odnowienia dla tego błądzenia to $\theta_t\eta$. Z wzoru (20) dostajemy

$$\nu_t = (\delta_0 - \mu) * \mathbf{E}\theta_t\eta = (\delta_0 - \mu) * \theta_t\mathbf{E}\eta = (\delta_0 - \mu) * \mathbf{E}\eta = \nu,$$

co oznacza, że $(S_{\sigma(t)+n} - t)_{n\geqslant 0}$ ma ten sam rozkład, co $(S_n)_{n\geqslant 0}$, zatem proces odnowienia η jest niezmienniczy. □

Kolejny wynik jest uogólnieniem drugiej części poprzedniego faktu na przypadek błądzeń dla których μ ma dodatnią wartość oczekiwaną, ale nie musi mieć nośnika w $\mathbb{R}_+$.

Fakt 7.4. *Załóżmy, że miara probabilistyczna μ na $\mathbb{R}$ ma średnią $c \in (0, \infty)$. Niech $\tilde{\mu}$ będzie rozkładem pierwszej wysokości drabinowej dla błądzenia startującego z zera o przyrostach z rozkładem μ, zaś $\tilde{c}$ oznacza średnią $\tilde{\mu}$. Określmy miarę probabilistyczną ν na $\mathbb{R}_+$ wzorem*

$$\tilde{\nu}[0, t] = \tilde{c}^{-1} \int_0^t \tilde{\mu}(s, \infty) ds, \quad t \geqslant 0.$$

Wówczas błądzenie losowe (S_n) o rozkładzie początkowym $\tilde{\nu}$ i rozkładzie przyrostów μ ma stacjonarny proces czasu przebywania η, ponadto $\mathbf{E}\eta = c^{-1}\lambda$.

Dowód. Na mocy MPWL $S_n^0 := S_n - S_0 \to \infty$ p.n., stąd na mocy Faktów 6.18 i 6.19 czasy drabinowe τ_n oraz wysokości drabinowe $S^0_{\tau_n} = S_{\tau_n} - S_0$ dla błądzenia (S_n^0) mają skończone wartości oczekiwane. Niech $H_n = S_{\tau_n}$, wówczas H_n jest błądzeniem losowym o nieujemnych przyrostach i rozkładzie początkowym $\tilde{\nu}$ – zatem z Faktu 7.3 wynika, że proces odnowy dla tego błądzenia $\xi = \sum_n \delta_{H_n}$ jest stacjonarny. Ustalmy $t \geqslant 0$ i niech $\sigma(t) := \inf\{n\colon\ S_n \geqslant t\}$ oraz $\tau(t) := \inf\{n\colon\ H_n \geqslant t\}$. Wówczas ze stacjonarności i wzoru (20) wynika, że $S_{\sigma(t)} - t = H_{\tau(t)} - t$ ma też rozkład $\tilde{\nu}$, a stąd i z mocnej własności Markowa dostajemy, że błądzenie $(S_{\sigma(t)+n} - t)_n$ ma ten sam rozkład co (S_n). Stąd otrzymujemy $\theta_t\eta$ ma ten sam rozkład co $\theta_0\eta$, czyli postulowaną stacjonarność η.

By zidentyfikować $\mathbf{E}\eta$, niech $\eta_n := \sum_{k=\tau_n}^{\tau_{n+1}-1} \delta_{S_k - H_n}$ będzie procesem czasu przebywania dla ciągu $(S_k - H_n)_{\tau_n \leqslant k < \tau_{n+1}}$. Wówczas, na mocy mocnej własności Markowa, zmienna H_n jest niezależna od procesu η_n i wszystkie η_n mają ten sam rozkład. Zatem

$$\mathbf{E}\eta = \mathbf{E}\sum_n \eta_n * \delta_{H_n} = \sum_n \mathbf{E}\eta_n * \mathbf{E}\delta_{H_n} = \mathbf{E}\eta_0 * \mathbf{E}\sum_n \delta_{H_n} = \mathbf{E}\eta_0 * \mathbf{E}\xi.$$

Na mocy Faktu 7.3 $\mathbf{E}\xi = \tilde{c}^{-1}\lambda$. Ponadto $\mathbf{E}\eta_0(0, \infty) = 0$ i $\tilde{c} = c\mathbf{E}\tau_1$. Stąd dla dowolnego $B \in \mathcal{B}(\mathbb{R}_+)$,

$$\mathbf{E}\eta(B) = \int_{(-\infty,0]} \tilde{c}^{-1}\lambda(B - t) d\mathbf{E}\eta_0(t) = \tilde{c}^{-1}\lambda(B)\mathbf{E}\eta_0(\mathbb{R}_-) = \tilde{c}^{-1}\lambda(B)\mathbf{E}\tau_1 = c^{-1}\lambda(B).$$

□

7.2 Twierdzenie odnowienia

Przejdziemy teraz do analizy zachowania asymptotycznego $\theta_t\eta$ i $\theta_t\mathbf{E}\eta$ przy $t \to \infty$. By móc sformułować stosowne twierdzenie będziemy musieli wprowadzić odpowiednie pojęcia zbieżności miar lokalnie skończonych (losowych bądź deterministycznych).

Definicja 7.5. Załóżmy, że $\nu, \nu_1, \nu_2, \ldots$ są lokalnie skończonymi miarami na $\mathbb{R}_+$. Powiemy, że ν_n zbiega *słabo* do ν (ozn. $\nu_n \xrightarrow{v} \nu$), jeśli $\int f d\nu_n$ zbiega do $\int f d\nu$ dla dowolnej funkcji ciągłej f na $\mathbb{R}_+$ o nośniku zwartym.

Podobnie, jeśli $\eta, \eta_1, \eta_2, \ldots$ są losowymi lokalnie skończonymi miarami na $\mathbb{R}_+$, to mówimy, że η_n *zbiega według rozkładu do* η (ozn. $\eta_n \xrightarrow{d} \eta$) , jeśli $\int f d\eta_n$ zbiega według rozkładu do $\int f d\eta$ dla dowolnej funkcji ciągłej f na $\mathbb{R}_+$ o nośniku zwartym.

Uwaga 7.6. Dla miar probabilistycznych słabą zbieżność definiuje się jako zbieżność całek funkcji ciągłych ograniczonych, bez założenia o zwartości nośnika (taką zbieżność będziemy oznaczać przez $\xrightarrow{w}$). Jeśli ν_n i ν są miarami probabilistycznymi, to $\nu_n \xrightarrow{v} \nu$ wtedy i tylko wtedy, gdy $\nu_n \xrightarrow{w} \nu$, co upoważnia nas do użycia tej samej nazwy (w języku angielskim używa się czasem pojęć *vague* oraz *weak convergence*). Zauważmy, że w sensie wprowadzonej powyżej definicji granica miar probabilistycznych nie musi być probabilistyczna, np. $\delta_n \xrightarrow{v} 0$.

Definicja 7.7. Mówimy, że miara μ na $\mathbb{R}$ jest niearytmetyczna, jeśli grupa addytywna generowana przez $\mathrm{supp}(\mu)$ jest gęsta w $\mathbb{R}$.

Używamy też konwencji, że zmienna X (bądź jej rozkład) ma średnią $c \in [-\infty, \infty]$, jeśli przynajmniej jedna z wielkości $\mathbf{E}X_+, \mathbf{E}X_-$ jest skończona oraz $c = \mathbf{E}X_+ - \mathbf{E}X_-$.

Twierdzenie 7.8 (Dwustronne twierdzenie odnowienia - Blackwell, Feller, Orey)**.** *Załóżmy, że η jest procesem czasu przebywania dla jednowymiarowego błądzenia losowego o rozkładzie początkowym ν i rozkładzie przyrostów μ oraz, że μ jest niearytmetyczny o średniej $c \in [-\infty, \infty] \setminus \{0\}$. Dla $0 < c < \infty$ oznaczmy przez $\tilde{\eta}$ stacjonarny proces przebywania dla błądzenia z rozkładem przyrostów μ i rozkładem początkowym $\tilde{\nu}$ zadanym przez Fakt 7.4, zaś dla innych c przyjmijmy, że $\tilde{\eta} = 0$. Wówczas przy $t \to \infty$,*
i) $\theta_t\eta \xrightarrow{d} \tilde{\eta}$,
ii) $\theta_t\mathbf{E}\eta \xrightarrow{v} \mathbf{E}\tilde{\eta} = \max\{c^{-1}, 0\}\lambda$.

Dowód będzie oparty na dwóch lematach. Pierwszy jest prostym ogólnym faktem dla błądzeń chwilowych.

Lemat 7.9. *Niech η będzie miarą czasu przebywania dla chwilowego błądzenia losowego (S_n) na $\mathbb{R}^d$ z dowolnym rozkładem początkowym, zaś B będzie ograniczonym zbiorem borelowskim w $\mathbb{R}^d$. Wówczas rodzina zmiennych losowych $(\eta(B + x))_{x \in \mathbb{R}^d}$ jest jednostajnie całkowalna.*

Dowód. Ustalmy $x \in \mathbb{R}^d$ i określmy $\tau := \inf\{t \geqslant 0\colon\ S_n \in B+x\}$. Niech η_0 będzie procesem czasu przebywania dla niezależnego błądzenia losowego o tym samym rozkładzie przyrostów co S_n, ale startującego z zera. Z mocnej własności Markowa wynika, że $\eta(B+x)$ ma ten sam rozkład co $\eta_0(B+x-S_\tau)\mathbb{1}_{\{\tau<\infty\}}$, a ta ostatnia zmienna się szacuje przez $\eta_0(B-B)$. Z chwilowości bładzenia i ograniczoności $B-B$ wynika, że $\mathbf{E}\eta_0(B-B) < \infty$ (Wniosek 6.8), co łatwo implikuje jednostajną całkowalność rozważanej rodziny. □

Zanim sformułujemy drugi lemat, przywołajmy pojawiający się w dowodzie stacjonarności procesu czasu przebywania rozkład ν_t zmiennej $S_{\sigma(t)}-t$, gdzie $\sigma(t) := \inf\{n\colon\ S_n \geqslant t\}$. Okazuje się, że, przy odpowiednich założeniach, rozkład ten zbiega słabo do rozkładu $\tilde{\nu}$ określonego w Fakcie 7.4.

Lemat 7.10. *Jeśli miara μ przyrostów błądzenia losowego jest niearytmetyczna i ma średnią $c \in (0,\infty)$, to $\nu_t \xrightarrow{w} \tilde{\nu}$ przy $t \to \infty$.*

Dowód lematu opiera się na idei parowania (ang. coupling) błądzeń losowych, a poza tym wymaga pewnych dodatkowych pomysłów, które mogą nieco utrudnić jego zrozumienie. Dlatego najpierw wykażemy go przy nieco mocniejszym założeniu o niearytmetyczności rozkładu $X_1 - X_2$.

Dowód lematu 7.10 przy założeniu niearytmetyczności rozkładu $X_1 - X_2$. Wybierzmy niezależne zmienne $S_0, S_0', X_1, X_1', X_2, X_2', \ldots$ takie, że X_i oraz X_i' mają rozkład μ, S_0 rozkład ν, a S_0' rozkład $\tilde{\nu}$. Połóżmy

$$S_n = S_0 + \sum_{i=1}^{n} X_i, \quad S_n' = S_0' + \sum_{i=1}^{n} X_i', \quad \text{oraz} \quad \tilde{S}_n := S_n' - S_n$$

Wówczas przyrosty błądzenia $(\tilde{S}_n)$ mają średnią zero oraz (zgodnie z naszym założeniem) rozkład niearytmetyczny. Na mocy Twierdzeń 6.6 i 6.9i) ciąg $\tilde{S}_n$ jest zatem z prawdopodobieństwem 1 gęsty w $\mathbb{R}$. Ustalmy $\varepsilon > 0$ i określmy moment zatrzymania

$$\tau := \inf\{n\colon\ \tilde{S}_n \in [0,\varepsilon]\},$$

który jest skończony p.n.. Na mocy mocnej własności Markowa błądzenia $(S_{\tau+n} - S_\tau)$ i $(S_{\tau+n}' - S_\tau')$ mają ten sam rozkład, ponadto oba są niezależne od $(S_k'\mathbb{1}_{\{k\leqslant\tau\}})$. Zatem, jeśli określimy

$$S_n'' := S_n'\mathbb{1}_{\{n<\tau\}} + [S_\tau' + (S_n - S_\tau)]\mathbb{1}_{\{n\geqslant\tau\}},$$

to otrzymamy błądzenie (S_n'') o tym samym rozkładzie co (S_n'). Ponadto

$$(S_n'' - S_n)\mathbb{1}_{\{n\geqslant\tau\}} = (S_\tau' - S_\tau)\mathbb{1}_{\{n\geqslant\tau\}} \in [0,\varepsilon], \quad \text{czyli} \quad S_n \leqslant S_n'' \leqslant S_n + \varepsilon \text{ dla } n \geqslant \tau.$$

Miara ν_t jest rozkładem $S_{\sigma(t)} - t$, zaś ze względu na stacjonarność $\tilde{\nu} = \tilde{\nu}_t$ jest rozkładem dla $S''_{\sigma''(t)} - t$, gdzie $\sigma''(t) := \inf\{n\colon S''_n \geqslant t\}$. Niech $Z := \max_{k\leqslant\tau} \max\{S_k, S'_k\}$, czyli $\{Z < t\} = \{\sigma(t), \sigma''(t) > \tau\}$. Mamy zatem dla $x > 0$,

$$\begin{aligned}\nu_t[0,x] - \mathbf{P}(Z \geqslant t) &\leqslant \mathbf{P}(S_{\sigma(t)} - t \in [0,x], Z < t\} \\ &\leqslant \mathbf{P}(S_{\sigma(t)} - t \in [0,x], S''_k \leqslant t + \varepsilon \text{ dla } k < \sigma(t), S''_{\sigma(t)} \in [0, x+\varepsilon]) \\ &\leqslant \mathbf{P}(S''_{\sigma''(t)} \in [0, x+\varepsilon]) = \tilde{\nu}[0, x+\varepsilon]\end{aligned}$$

oraz

$$\begin{aligned}\tilde{\nu}[\varepsilon,x] - \mathbf{P}(Z \geqslant t) &\leqslant \mathbf{P}(S''_{\sigma''(t)} - t \in [\varepsilon,x], Z < t\} \\ &\leqslant \mathbf{P}(S''_{\sigma''(t)} - t \in [\varepsilon,x], S_k < t \text{ dla } k < \sigma''(t), S_{\sigma''(t)} \in [0,x]) \\ &\leqslant \mathbf{P}(S_{\sigma(t)} \in [0,x]) = \nu_t[0,x].\end{aligned}$$

Stąd

$$\tilde{\nu}[\varepsilon,x] - \mathbf{P}(Z \geqslant t) \leqslant \nu_t[0,x] \leqslant \tilde{\nu}[0,x+\varepsilon] + \mathbf{P}(Z \geqslant t).$$

Zatem

$$\tilde{\nu}[\varepsilon,x] \leqslant \liminf_{t\to\infty} \nu_t[0,x] \leqslant \limsup_{t\to\infty} \nu_t[0,x] \leqslant \tilde{\nu}[0,x+\varepsilon].$$

Biorąc $\varepsilon \to 0$ i korzystając z tego, że $\tilde{\nu}\{0\} = 0$ dostajemy $\nu_t[0,x] \to \tilde{\nu}[0,x]$. □

Dowód Lematu 7.10 w pełnej ogólności. Niech $S_0, S'_0, X_1, \varepsilon_1, X_2, \varepsilon_2, \ldots$ będą niezależnymi zmiennymi losowymi takimi, że S_0 ma rozkład ν, S'_0 rozkład $\tilde{\nu}$, zmienne X_k rozkład μ oraz $\mathbf{P}(\varepsilon_k = \pm 1) = 1/2$. Rozważmy błądzenie losowe

$$\tilde{S}_n = S'_0 - S_0 - \sum_{k=1}^{n} \varepsilon_k X_k, \quad n = 0, 1, \ldots.$$

Wówczas przyrost błądzenia $(\tilde{S}_n)$ ma rozkład niearytmetyczny o średniej zero, stąd na mocy Twierdzeń 6.6 i 6.9i) ciąg $\tilde{S}_n$ jest z prawdopodobieństwem 1 gęsty w $\mathbb{R}$. W szczególności dla dowolnego $\varepsilon > 0$ moment zatrzymania $\tau := \inf\{n\colon\ \tilde{S}_n \in [0,\varepsilon]\}$ jest p.n. skończony.

Określmy $\varepsilon'_k := (-1)^{\mathbb{1}_{\{k\leqslant\tau\}}}\varepsilon_k$. Sprawdzamy, że ciąg (ε'_k) jest tak jak (ε_k) ciągiem niezależnych symetrycznych zmiennych losowych o wartościach ± 1, niezależnych od S_0, S'_0 i X_k. Niech $\kappa_1 < \kappa_2 < \ldots$ będą kolejnymi wartościami k dla których $\varepsilon_k = 1$, podobnie zdefiniujmy ciąg $\kappa'_1 < \kappa'_2 < \ldots$ poprzez zmienne ε'_k. Łatwo zauważyć, że ciągi

$$S_n = S_0 + \sum_{j\leqslant n} X_{\kappa_j}, \quad S'_n = S'_0 + \sum_{j\leqslant n} X_{\kappa'_j}$$

są błądzeniami losowymi o rozkładzie przyrostów μ i rozkładzie początkowym równym odpowiednio ν i $\tilde{\nu}$. Określmy

$$\tau_{\pm} = \sum_{k\leqslant\tau} \mathbb{1}_{\{\varepsilon_k = 1\}}.$$

Jak łatwo sprawdzić
$$S'_{\tau_-+n} - S_{\tau_++n} = \tilde{S}_\tau \in [0,\varepsilon].$$
Ponieważ ν_t jest rozkładem pierwszej nieujemnej wartości $S_n - t$, zaś ze względu na stacjonarność $\tilde{\nu} = \tilde{\nu}_t$ jest podobnym rozkładem dla $S'_n - t$, więc, podobnie jak w poprzednim dowodzie,
$$\tilde{\nu}[\varepsilon, x] - \mathbf{P}(Z \geqslant t) \leqslant \nu_t[0,x] \leqslant \tilde{\nu}[0, x+\varepsilon] + \mathbf{P}(Z \geqslant t)$$
gdzie $Z := \max\{\max_{k\leqslant\tau_-} S'_k, \max_{k\leqslant\tau_+} S_k\}$. Biorąc $\varepsilon \to 0$ i korzystając z tego, że $\tilde{\nu}\{0\} = 0$ dostajemy $\nu_t[0,x] \to \tilde{\nu}[0,x]$. □

Dowód Twierdzenia 7.8. Przypadek I. $c < \infty$. Zauważmy, że z Lematu 7.9 wynika, że zmienne $X_t(f) = \int f d\nu_t$ są jednostajnie całkowalne dla dowolnej funkcji ciągłej o nośniku zwartym. Stąd zbieżność $X_t(f)$ według rozkładu implikuje zbieżność wartości oczekiwanych, zatem wystarczy udowodnić i).

Jeśli $c < 0$, to $S_n \to -\infty$ p.n., czyli $\sup S_n < \infty$ p.n., ponadto $\theta_t\eta = 0$ dla $t > \sup S_n$, więc i) jest w tym przypadku oczywista. Jeśli $0 < c < \infty$, to z Lematu 7.10 $\nu_t \xrightarrow{w} \tilde{\nu}$, możemy zatem skonstruować zmienne Y_t i Y o rozkładzie równym odpowiednio ν_t i $\tilde{\nu}$, takie, że Y_t zbiega do Y punktowo. Niech η_0 będzie procesem czasu przebywania dla, niezależnego od Y i Y_t, błądzenia losowego (S'_n) o przyrostach μ startującego z 0.

Ustalmy funkcję ciągłą f na $\mathbb{R}_+$ o nośniku zwartym i przedłużmy ją na $\mathbb{R}$ kładąc $f(x) = 0$ dla $x < 0$. Zauważmy, że $f(Y_t + x) \to f(Y + x)$ poza zbiorem $\{x = -Y\}$. Ponadto $\mathbf{E}\eta_0(\{-Y\}) = \sum_{n\geqslant 0} \mathbf{P}(Y = -S'_n) = 0$, gdyż rozkład Y jest bezatomowy, zaś S'_n jest niezależne od Y. Stąd $\eta_0(\{-Y\}) = 0$ p.n.. Z mocnej własności Markowa dostajemy
$$\int f \mathrm{d}\theta_t\eta \stackrel{d}{=} \int f(Y_t + x)\mathrm{d}\eta_0(x) \to \int f(Y + x)\mathrm{d}\eta_0(x) \stackrel{d}{=} \int f \mathrm{d}\tilde{\eta},$$
gdzie zbieżność na zbiorze $\eta_0(\{-Y\}) = 0$ wynika z twierdzenia Lebesgue'a o zbieżności zmajoryzowanej.

Przypadek II. $c = \infty$. Tu wystarczy udowodnić ii) (zbieżność średnich zmiennych nieujemnych do zera implikuje zbieżność tych zmiennych do 0 wg prawdopodobieństwa, a więc i wg rozkładu).

Załóżmy wpierw, że miara μ jest nieujemna i $\nu = \delta_0$, czyli błądzenie staruje z zera. Ustalmy $u > 0$ i połóżmy $I = [0,u]$. Na mocy Lematu 7.9, $a := \sup \mathbf{E}\eta(I + t) < \infty$. Określmy $b := \limsup_{t\to\infty} \mathbf{E}\eta(I + t)$ i wybierzmy $t_k \to \infty$ takie, że $\mathbf{E}\eta(I + t_k) \to b$. Zauważmy, że dla ustalonego $m \geqslant 0$, $\mu^{*m}(I + t_k) \to 0$ przy $k \to \infty$, więc
$$\mu^{*m} * \mathbf{E}\eta(I + t_k) = \mathbf{E}\eta(I + t_k) - \sum_{l=0}^{m} \mu^{*l}(I + t_k) \to b \quad \text{ dla } m = 0, 1, \ldots.$$

Dla dowolnego $B \in \mathcal{B}(\mathbb{R}_+)$ mamy

$$\begin{aligned}
\liminf_{k\to\infty} \mathbf{E}\eta(I - B + t_k)\mu^{*m}(B) &\geqslant \liminf_{k\to\infty} \int_B \mathbf{E}\eta(I - x + t_k)\mathrm{d}\mu^{*m}(x) \\
&= \liminf_{k\to\infty} \left(\mu^{*m} * \mathbf{E}\eta(I + t_k) - \int_{B^c} \mathbf{E}\eta(I - x + t_k)\mathrm{d}\mu^{*m}(x)\right) \\
&= b - \limsup_{k\to\infty} \int_{B^c} \mathbf{E}\eta(I - x + t_k)\mathrm{d}\mu^{*m}(x) \\
&\geqslant b - \int_{B^c} \limsup_{k\to\infty} \mathbf{E}\eta(I - x + t_k)\mathrm{d}\mu^{*m}(x) \\
&\geqslant b - \int_{B^c} b\mathrm{d}\mu^{*m}(x) = b\mu^{*m}(B).
\end{aligned}$$

gdzie przedostatnia nierówność wynika z Lematu Fatou (zastosowanego do fukcji nieujemnych $a - \mathbf{E}\eta(I - x + t_k)$). Dodając stronami po wszystkich $m \geqslant 0$ dostajemy

$$\liminf_{k\to\infty} \mathbf{E}\eta(I - B + t_k) \geqslant b, \text{ jeśli } \mathbf{E}\eta(B) > 0. \tag{21}$$

Ustalmy teraz $h_1 > h_2 > 0$ takie, że $\mu(h_2, h_2) > 0$, wtedy dla dowolnego $r > 0$,

$$\mathbf{E}\eta(r, r + h_1] = \sum_{n\geqslant 1} \mathbf{P}(S_n \in (r, r + h_1]) \geqslant \mathbf{P}(X_k \in (h_2, h_1) \text{ dla } 1 \leqslant k \leqslant 1 + \lceil r/h_2 \rceil) > 0.$$

Niech $J = [0, s)$ oraz $s = u + h_1$, stosując (21) dla $B = (r - h_1, r]$ dostajemy dla $r > h_1$,

$$\liminf_{k\to\infty} \mathbf{E}\eta(J + t_k - r) \geqslant b.$$

Z tożsamości $\delta_0 = (\delta_0 - \mu) * \mathbf{E}\eta$ wynika

$$\begin{aligned}
1 = \delta_0[0, t_k] &= \int_{[0,t_k]} \mu(t_k - x, \infty) d\mathbf{E}\eta(x) = \sum_{n\geqslant 1} \int_{[t_k - ns, t_k - (n-1)s)} \mu(t_k - x, \infty) d\mathbf{E}\eta(x) \\
&\geqslant \sum_{n\geqslant 1} \mu(ns, \infty)\mathbf{E}\eta(J + t_k - ns).
\end{aligned}$$

Biorąc $k \to \infty$ dostajemy $1 \geqslant b \sum_{n\geqslant 1} \mu(ns, \infty)$. Z założenia o nieskończonej średniej μ dostajemy $b = 0$, czyli $\theta_t \mathbf{E}\eta([0, u]) \to 0$, co oznacza, że $\theta_t \mathbf{E}\eta \xrightarrow{v} 0$.

W przypadku, gdy ν jest dowolne, a μ niekoniecznie jest skoncentrowane na $\mathbb{R}_+$, na mocy mocnej własności Markowa mamy $\mathbf{E}\eta = \nu * \mathbf{E}\xi * \mathbf{E}\kappa$, gdzie κ jest procesem przebywania dla wysokości drabinowych błądzenia $(S_n - S_0)$, a ξ procesem przebywania dla błądzenia $(S_n - S_0)$ przed pierwszą wysokością drabinową τ_1. Zauważmy, że pierwsza wysokość drabinowa $(S_n - S_0)$ ma nieskończoną wartość oczekiwaną, stąd z rozpatrzonego poprzednio przypadku $\theta_t \mathbf{E}\kappa \to 0$. Ponadto $\mathbf{E}\xi(\mathbb{R}) = \mathbf{E}\tau_1 < \infty$, więc $\nu * \mathbf{E}\xi$ jest miarą skończoną i dostajemy łatwo $\theta_t \mathbf{E}\eta \xrightarrow{v} 0$.

$\square$

7.3 Równanie odnowienia. Asymptotyka rozwiązania

Definicja 7.11. Funkcję f na $\mathbb{R}_+$ nazywamy *regularną funkcją schodkową*, jeśli ma postać

$$f = \sum_{j=1}^{\infty} a_j \mathbb{1}_{[(j-1)h, jh)}$$

dla pewnego $h > 0$ i $a_1, a_2, \dots \in \mathbb{R}$. Mówimy, że funkcja mierzalna f na $\mathbb{R}_+$ jest *bezpośrednio całkowalna w sensie Riemanna*, jeśli $\int_0^\infty |f(x)|dx < \infty$ oraz istnieją ciągi $(f_n^\pm)$ regularnych funkcji schodkowych takie, że

$$f_n^- \leqslant f \leqslant f_n^+ \quad \text{oraz} \quad \int_0^\infty (f_n^+(x) - f_n^-(x))dx \to 0.$$

Uwaga 7.12. i) Funkcje ciągłe o zwartym nośniku, funkcje monotoniczne z $L_1(\mathbb{R}_+)$ są bezpośrednio całkowalne w sensie Riemanna.
ii) Funkcje o nośniku zawartym w przedziale skończonym $[0, x]$ są bezpośrednio całkowalne w sensie Riemanna wtedy i tylko wtedy, gdy są całkowalne w sensie Riemanna na $[0, x]$.

Wniosek 7.13. *Załóżmy, że $\mu \neq \delta_0$ jest miarą probabilistyczną na $\mathbb{R}_+$, $\bar{\mu} := \sum_{n \geqslant 0} \mu^{*n}$, zaś f jest lokalnie ograniczoną mierzalną funkcją na $\mathbb{R}_+$. Wówczas równanie $F = f + F * \mu$ ma jednoznaczne lokalnie ograniczone rozwiązanie dane wzorem $F = f * \bar{\mu}$. Ponadto, jeśli f jest bezpośrednio całkowalna w sensie Riemanna oraz μ jest niearytmetyczne ze średnią $c \in (0, \infty]$, to $F(t) \to c^{-1} \int_0^\infty f d\lambda$ przy $t \to \infty$.*

Dowód. Iterując równanie odnowienia $F = f + F * \mu$ dostajemy

$$F = \sum_{k=0}^{n} f * \mu^{*k} + F * \mu^{*n}, \quad n = 1, 2, \dots. \tag{22}$$

Ze słabego prawa wielkich liczb wynika, że dla $t \geqslant 0$, $\mu^{*n}[0, t] \to 0$ przy $n \to \infty$. Stąd dla lokalnie ograniczonej funkcji F dostajemy, że $F * \mu^{*n}$ zbiega punktowo do 0. Jeśli F i f są lokalnie ograniczone, to przechodząc w (22) z $n \to \infty$ i korzystając z twierdzenia Fubiniego dostajemy

$$F = \sum_{k=0}^{\infty} f * \mu^{*k} = f * \sum_{k=0}^{\infty} \mu^{*k} = f * \bar{\mu}.$$

Na odwrót zauważmy, że jeśli f jest lokalnie ograniczone, to $f * \bar{\mu}$ też takie jest oraz

$$f + f * \bar{\mu} * \mu = f + f * \sum_{k=1}^{\infty} \mu^{*k} = f * \bar{\mu},$$

czyli $F = f * \bar{\mu}$ jest rozwiązaniem równania odnowy.

Załóżmy teraz, że rozkład μ jest niearytmetyczny. Jeśli $f = \sum_{j=1}^{\infty} a_j \mathbb{1}_{[(j-1)h,jh)}$ jest regularną, całkowalną funkcją schodkową, to z Twierdzenia 7.8 i twierdzenia Lebesgue'a o zbieżności zmajoryzowanej dostajemy

$$F(t) = \int_0^t f(t-s)d\bar{\mu}(s) = \sum_{j\geqslant 1} a_j \bar{\mu}((0,h]+t-jh) \to \sum_{j\geqslant 1} a_j c^{-1} h = c^{-1}\int f d\lambda.$$

W przypadku, gdy f jest bezpośrednio całkowalna w sensie Riemanna to na mocy definicji istnieją regularne, całkowalne funkcje schodkowe $f_n^{\pm}$ takie, że $f_n^- \leqslant f \leqslant f_n^+$ oraz $\int (f_n^+(x) - f_n^-(x))d\lambda \to 0$. Mamy oczywiście

$$(f_n^- * \bar{\mu})(t) \leqslant F(t) \leqslant (f_n^+ * \bar{\mu})(t), \quad t \geqslant 0, n = 1,2,\ldots.$$

Biorąc wpierw $t \to \infty$ a potem $n \to \infty$ dostajemy $F(t) \to c^{-1}\int_0^\infty f d\lambda$. □

7.4 Rozkłady arytmetyczne

Formułując powyżej twierdzenie odnowienia oraz twierdzenie o asymptotyce rozwiązania równania odnowienia zakładaliśmy niearytmetyczność rozkładów. Oba twierdzenia można przeformułować na przypadek rozkładów niearytmetycznych, dowody są nawet w pewnym sensie prostsze - gdyż rozpatrywane błądzenia losowe są wówczas łańcuchami Markowa z dyskretną przestrzenią stanów.

Jeśli rozkład μ jest arytmetyczny, to grupa addytywna generowana przez jego nośnik jest dyskretna, zatem musi mieć postać $h\mathbb{Z}$ dla pewnego $h > 0$.

Definicja 7.14. Mówimy, że rozkład μ jest *arytmetyczny ze skokiem* $h > 0$, jeśli grupa addytywna generowana przez nośnik μ jest równa $h\mathbb{Z}$.

Dla błądzeń losowych na $h\mathbb{Z}$ nie ma sensu mówić o stacjonarności na $\mathbb{R}_+$, ale o stacjonarności na $h\mathbb{N}$.

Definicja 7.15. Powiemy, że proces czasu przebywania η jest *stacjonarny* na $h\mathbb{N}$, jeśli $\theta_{kh}\eta$ ma ten sam rozkład, co $\theta_0\eta$ dla $h = 0,1,\ldots$.

Przez $\lambda_{h\mathbb{N}}$ będziemy oznaczali miarę liczącą na $h\mathbb{N}$. Z dokładnością do stałej to jedyna miara na $h\mathbb{N}$ niezmiennicza na przesunięcia.

Poniżej formułujemy tylko odpowiednie wyniki dla błądzeń o arytmetycznych przyrostach - ich dowody są analogiczne jak w przypadku rozkładów niearytmetycznych.

Fakt 7.16. *Załóżmy, że błądzenie losowe ma przyrosty nieujemne o rozkładzie arytmetycznym μ o skoku h i średniej c. Wówczas istnieje rozkład początkowy ν na $h\mathbb{N}$ taki, że proces odnowienia η jest stacjonarny na $h\mathbb{N}$ wtedy i tylko wtedy, gdy $0 < c < \infty$. Ponadto, gdy η jest stacjonarny na $h\mathbb{N}$, to $\mathbf{E}\eta = \frac{h}{c}\lambda_{h\mathbb{N}}$ oraz $\nu = \frac{h}{c}(\delta_0 - \mu) * \lambda_{h\mathbb{N}}$, czyli*

$$\nu\{kh\} = \frac{h}{c}\mu(kh,\infty), \quad k = 0,1,\ldots.$$

Fakt 7.17. *Załóżmy, że rozkład arytmetyczny μ o skoku h ma średnią $c \in (0, \infty)$. Niech $\tilde{\mu}$ będzie rozkładem pierwszej wysokości drabinowej dla błądzenia startującego z zera o przyrostach z rozkładem μ, zaś $\tilde{c}$ oznacza średnią $\tilde{\mu}$. Określmy miarę probabilistyczną ν na $h\mathbb{N}$ wzorem*

$$\tilde{\nu}\{kh\} = \frac{h}{c}\tilde{\mu}(kh, \infty)ds, \quad k = 0, 1, 2, \ldots.$$

Wówczas błądzenie losowe (S_n) o rozkładzie początkowym $\tilde{\nu}$ i rozkładzie przyrostów μ ma stacjonarny na $h\mathbb{N}$ proces czasu przebywania η, ponadto $\mathbf{E}\eta = \frac{h}{c}\lambda_{h\mathbb{N}}$.

Twierdzenie 7.18 (Dwustronne twierdzenie odnowienia dla rozkładów arytmetycznych)**.** *Załóżmy, że η jest procesem czasu przebywania dla błądzenia losowego na $h\mathbb{Z}$ o rozkładzie początkowym ν i rozkładzie przyrostów μ oraz, że μ jest arytmetyczny o skoku h i średniej $c \in [-\infty, \infty] \setminus \{0\}$. Dla $0 < c < \infty$ oznaczmy przez $\tilde{\eta}$ stacjonarny proces przebywania dla błądzenia z rozkładem przyrostów μ i rozkładem początkowym $\tilde{\nu}$ zadanym przez Fakt 7.17, zaś dla innych c przyjmijmy, że $\tilde{\eta} = 0$. Wówczas przy $k \to \infty$,*
i) $\theta_{kh}\eta \xrightarrow{d} \tilde{\eta}$,
ii) $\theta_{kh}\mathbf{E}\eta \xrightarrow{v} \mathbf{E}\tilde{\eta} = \max\{hc^{-1}, 0\}\lambda$.

Uwaga 7.19. Punkt ii) dla mówi po prostu, że $\mathbf{E}\eta\{kh\} \to \max\{hc^{-1}, 0\}$ przy $k \to \infty$.

Zauważmy, że każda funkcja na $h\mathbb{Z}$ jest lokalnie ograniczona oraz można ją przedstawić jako przeliczalną sumę delt Diraca. Zatem dla rozkładów arytmetycznych na $h\mathbb{Z}$ równanie $F = f + F * \mu$ ma zawsze rozwiązanie, a jego asymptotyka jest dana przez następujący wniosek.

Wniosek 7.20. *Załóżmy, że $h > 0$, $\mu \neq \delta_0$ jest miarą probabilistyczną na $h\mathbb{N}$ ze średnią $c \in (0, \infty)$ oraz grupa generowana przez nośnik μ jest równa $h\mathbb{Z}$. Niech f będzie funkcją na $h\mathbb{N}$ taką, że $\sum_{k \geqslant 0} |f(kh)| < \infty$. Wówczas rozwiązanie równania $F = f + F * \mu$ spełnia*

$$F(kh) \to \frac{h}{c}\sum_{n=0}^{\infty} f(nh) \quad \textit{przy } k \to \infty.$$

Literatura

[1] P. Billingsley *Prawdopodobieństwo i miara*, wyd. II, PWN, Warszawa 2009.

[2] A. Dembo, O. Zeitouni, *Large deviations techniques and applications*, Springer, New York, 1998.

[3] W. Feller, *Wstęp do rachunku prawdopodobieństwa*, tom II, wyd. IV, PWN, Warszawa 2009.

[4] J. Jakubowski, R. Sztencel, *Wstęp do teorii prawdopodobieństwa*, wyd. IV, Script, Warszawa 2010.

[5] O. Kallenberg, *Foundations of modern probability*, 2nd edition, Springer, New York 2001.

[6] S. Kwapień, W. A. Woyczyński, *Random Series and Stochastic Integrals: Simple and Multiple*, Birkhauser, New York 1992.

[7] M. Ledoux, M. Talagrand, *Probability in Banach spaces*, Springer, Berlin, 1991.

[8] M. Loéve, *Probability theory I,II*, 4th edition, Springer, New York-Heidelberg, 1977,1978.

[9] D. W. Stroock, *Probability theory, an analytic view*, Cambridge University Press, Cambridge, 1993.