Rachunek Prawdopodobieństwa 3

Rafał Latała

22 maja 2007

Poniższe notatki powstają na podstawie (wybranych) wykładów z Rachunku Prawdopodobieństwa 3, prowadzonych w semestrze wiosennym 2006/07.

Przepraszam za wszystkie nieścisłości i omyłki mogące pojawić 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.

Rafał Latała

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

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

Możemy więc postawić pytania:
1) Czy oszacowanie $\mathbf{P}(|\bar{S}_n - \mathbf{E}X| > \varepsilon)$ rzędu $\exp(-nf(\varepsilon))$ zachodzi dla szerszej klasy zmiennych?
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 1.1. 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 $M_X(0) = 1$.

Fakt 1.2. *Jeśli X i Y są niezależnymi zmiennymi losowymi, to $M_{X+Y} = M_X M_Y$. W szczególności, $M_{S_n} = (M_X)^n$.*

Dowód. Oczywisty. □

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

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

Biorąc infimum lewej 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 1.2. *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 $\Lambda_X^* \equiv 0$.

Lemat 1.3. *Funkcje $\Lambda_X \colon \mathbb{R} \to (-\infty, \infty]$, $\Lambda_X^* \colon \mathbb{R} \to [0, \infty)$ są wypukłe.*

Dowód. Wypukłość funkcji Λ 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 konwencje, 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 1.4. *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,

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

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) $\mathbf{E}X_+ < \infty = \mathbf{E}X_-$. Dla $s \in \mathbb{R}$ dostajemy

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

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 $s \to -\infty$. $\square$

Twierdzenie 1.5 (Nierówność Chernoffa)**.** *Jeśli* $\mathbf{E}X_+ < \infty$ *lub* $\mathbf{E}X_- < \infty$*, to*

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

oraz

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

Dowód. Pierwsza nierówność wynika z (1) i Lematu 1.4 ii). Drugą nierówność dowodzimy analogicznie. $\square$

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

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

oraz

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

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 $t > a$

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

b) Dla $X = g^3$,$g \sim \mathcal{N}(0,1)$, nierówność Chernoffa daje tylko trywialne oszacowanie $\mathbf{P}(\bar{S}_n \geqslant t) \leqslant 1$.

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

$$\mathbf{E}e^{tX} < \infty \text{ dla wszystkich } t \in \mathbb{R}. \tag{2}$$

Lemat 1.6. *Przy założeniu (2) 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 $f^{(n)}(z) = \mathbf{E}(X^n\exp(zX))$. □

Definicja 1.3. *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.*

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

Lemat 1.7. *Miara probabilistyczna μ_t spełnia warunki*

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

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

Dowód. Zauważmy, że z Lematu 1.6, $M_X^{(n)}(t) = M_X(t)\mathbf{E}_{\mu_t}(x^n)$, więc

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

oraz

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

□

Określmy

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

Lemat 1.8. *Dla dowolnej zmiennej losowej X,*

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

Dowód. Dla $t \geqslant 0$ mamy $\Lambda_X(t) \leqslant \alpha t$, stąd dla $s > \alpha$,

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

Podobnie $\Lambda_X(t) \leqslant \beta t$ dla $t \leqslant 0$ i dla $S < \beta$,

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

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,

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

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

Lemat 1.9. *Załóżmy, że zachodzi warunek (2) 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 1.7, $\Lambda_X''(t) > 0$ czyli Λ_X' jest ściśle rosnąca i funkcja $\Theta = (\Lambda_X')^{-1}$ jest dobrze określona na $\Lambda_X'(\mathbb{R})$. Oczywiście dla dowolnego t,

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

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 1.4 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. $\square$

Twierdzenie 1.10. *Załóżmy, że zachodzi (2). Wówczas dla* $a \in (\alpha, \beta)$ *oraz* $\varepsilon > 0$,

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

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

Dowód. Zdefiniujmy $t := \Theta(a)$, wówczas z Lematu 1.9, $\Lambda_X^*(a) = at - \Lambda_X(t)$. Niech $Y, Y_1, \ldots, Y_n$ będą niezależnymi zmiennymi losowymi o rozkładzie zadanym przez miarę probabilistyczną μ_t. Na mocy Lematu 1.7, $\mathbf{E}Y = \Lambda_X'(t) = a$ i $\mathrm{Var}(Y) = \Lambda_X''(t)$. Zauważmy, że dla dowolnej nieujemnej funkcji f na $\mathbb{R}^n$,

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

Zatem

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

Z nierówności Czebyszewa

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

zatem

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

□

1.3 Twierdzenie Cramera na $\mathbb{R}$

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

Twierdzenie 1.11 (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^* jest niemalejąca, $\Lambda_X^*(x) \to 0$ przy $x\to-\infty$, a 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, 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

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbf{P}(\bar{S}_n \in (-\delta,\delta)) \geqslant -\Lambda_X^*(0) = \inf_{t\in\mathbb{R}}\Lambda_X(t). \tag{3}$$

Dowód (3) rozbijemy na kilka przypadków.

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

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

Ponadto $\mathbf{P}(\bar{S}_n \in (-\delta,\delta)) \geqslant \mathbf{P}(\bar{S}_n = 0) = \mathbf{P}(X=0)^n$ i natychmiast dostajemy (3).

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

Przypadek III. $\mathrm{essinf} X < 0 < \mathrm{esssup} X$ oraz $D_{\Lambda_X} = \mathbb{R}$. Z Twierdzenia 1.10 (z $a = 0$) dostajemy dla $0 < \varepsilon < \delta$,

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

Biorąc $\varepsilon \to 0+$ dostajemy (3).

Przypadek IV. $\mathrm{essinf}X < 0 < \mathrm{esssup}X$ 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 funkcje 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, niepustych. Istnieje więc punkt t_0 należący do ich przecięcia, ale wtedy $\Lambda_M(t_0) \leqslant a$ dla wszystkich M czyli, z twierdzenia Lebesgue'a o zbieżności monotonicznej, $\Lambda_X(t_0) = \lim_{M\to\infty} \Lambda_M(t_0) \leqslant a$. $\square$

Uwaga 1.1 Analiza pierwszej części dowodu pokazuje, że wykazaliśmy

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

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 1.4. *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

$$M_X(t) < \infty \text{ dla wszystkich } t \in \mathbb{R}^d. \tag{4}$$

Lemat 1.12. *Przy założeniu warunku (4),*
i) Funkcja $\Lambda_X \colon \mathbb{R}^d \to \mathbb{R}$ jest wypukła, klasy C^∞ oraz

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

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 $\alpha \in \mathbb{Z}_+^d$,

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

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

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

Wypukłość Λ_X implikuje wklęsłość g, zatem

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

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 1.13. *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 $\Gamma \in \mathcal{B}(\mathbb{R}^d)$,

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

Dowód. i) 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

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbf{P}(\bar{S}_n \in F) \leqslant 2\delta - \inf_{s\in F}(\Lambda_X^*(s) \wedge M) \text{ dla } \delta > 0, M < \infty. \tag{5}$$

Ustalmy $\delta > 0$ i $M < \infty$, dowód (5) 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 $G \in \mathcal{B}(\mathbb{R}^d)$,

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

Zauważmy, że

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

Zatem

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

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 zwarty. 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

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

Rozpatrzymy dwa przypadki.

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

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

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

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,

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

Otrzymujemy więc

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

Przypadek II. X dowolna zmienna spełniająca (4). Niech $G, G_1, G_2, \ldots$ 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 czyli 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

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

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

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

dla dostatecznie duzych n. Zatem

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

dla dużych n, czyli

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

□

2 Nierówności Wykładnicze.

Nierówność Chernoffa pokazuje, że jeśli $S = \sum_{i=1}^n X_i$ oraz 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

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

Ponadto jeśli będziemy umieli oszacować dla wszystkich i, Λ_{X_i} z góry, to da nam to górne ograniczenie $\Lambda_S = \sum_{i=1}^n \Lambda_{X_i}$ 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 zmienną losową o średniej zero taką, że $d = \|X\|_\infty < \infty$. Wówczas*

$$\Lambda_X(t) \leqslant \exp((dt)^2/2) \text{ dla wszystkich } t.$$

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 $\exp(x)$,

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

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 niezależnymi zmiennymi losowymi o średniej zero oraz $\|X_i\|_\infty \leqslant d_i$. Wówczas*

$$\mathbf{E}\exp\Big(t\sum_{i=1}^n X_i\Big) \leqslant \exp\Big(\sum_{i=1}^n d_i^2\frac{t^2}{2}\Big)$$

oraz

$$\mathbf{P}\Big(\sum_{i=1}^n X_i \geqslant s\Big) \leqslant \exp\Big(\frac{s^2}{2\sum_{i=1}^n d_i^2}\Big).$$

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

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

□

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*

$$\mathbf{P}\Big(\sum_i a_i\varepsilon_i \geqslant s)\Big) \leqslant \exp\Big(- \frac{s^2}{2\sum_i a_i^2}\Big).$$

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

Uwaga 2.2 Można udowodnić znacznie ogólniejsze twierdzenie - *nierówność Azumy*. Dla dowolnego martyngału $(M_k, \mathcal{F}_k)_{k=1}^n$ takiego, że $\|M_k - M_{k-1}\|_\infty \leqslant d_k$ zachodzi

$$\mathbf{P}(M_n - M_0 \geqslant s) \leqslant \exp\Big(\frac{s^2}{2\sum_{i=1}^n d_i^2}\Big).$$

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 2.4. *Załóżmy, że X jest zmienną losową o średniej zero taką, że istnieją $\sigma^2, M < \infty$ spełniające*

$$\mathbf{E}|X|^k \leqslant \frac{k!}{2}\sigma^2 M^{k-2} \text{ dla } k = 2, 3, \ldots.$$

Wówczas

$$\Lambda_X(t) \leqslant \frac{\sigma^2 t^2}{2(1 - M|t|)} \text{ dla } M|t| < 1.$$

Dowód. Liczymy

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

□

Twierdzenie 2.5 (Nierówność Bernsteina)**.** *Załóżmy, że X_i są niezależnymi zmiennymi losowymi o średniej zero, zaś $\sigma_i^2, M < \infty$ są takie, że*

$$\mathbf{E}|X_i|^k \leqslant \frac{k!}{2} \sigma_i^2 M^{k-2} \text{ dla } i \geqslant 1,\ k \geqslant 2. \tag{6}$$

Wówczas

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

oraz dla $s > 0$,

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

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 2.4. Dalej szacujemy

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

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

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

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

Uwaga 2.3 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 $\mathbf{E}|X_i|^2 = k!$, czyli dla takich zmiennych zachodzą założenia Twierdzenia 2.5 z $\sigma_i^2 = 2, M = 1$. Pokazuje to, że i nie możemy uzyskać szacowania lepszego niż $\exp(-s/M)$.

Wniosek 2.6. *Załóżmy, że X_i są ograniczonymi niezależnymi zmiennymi losowymi o średniej zero, wówczas dla $s > 0$,*

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

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 $k \geqslant 2$,

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

zatem warunek (6) jest spełniony z $M = a/3$ oraz $\sigma_i = \mathbf{E}X_i^2$. $\square$

Uwaga 2.4 Nierówność Bernsteina pokazuje, iż jeśli $s \max_i \|X_i\|_\infty \leqslant 3\varepsilon \sum_i \mathbf{E}X_i^2$, to

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

Kolejny wynik pokazuje jak odwrócić szacowanie z powyższej uwagi.

Twierdzenie 2.7 (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 oraz $\sigma^2 = \mathrm{Var}(\sum_{i=1}^n X_i)$,*

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

o ile

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

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

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

to

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

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

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

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

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

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

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

Konstrukcje 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 $\mathrm{Var}(S_j) \geqslant u^2/(s^2\sqrt{1+\varepsilon})$. 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,

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

oraz

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

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

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

co daje

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

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

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

□

Uwaga 2.5 Nierówność z powyższego twierdzenia można też równoważnie sformułować jako

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

o ile

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

2.3 Nierówność Bennetta

Szacowanie podane we Wniosku 2.6 jest z uwagi na centralne twierdzenie graniczne bliskie optymalnego dla s małych. Jednak dla s dużych można je poprawić o czynnik logarytmiczy.

Lemat 2.8. *Załóżmy, że X jest zmienną losową o średniej zero, wariancji σ^2 oraz $\|X_i\|_\infty \leqslant a$. Wówczas*

$$\Lambda_X(t) \leqslant \frac{\sigma^2}{a^2}(e^{ta} - ta - 1).$$

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 łatwo z nierówności $\ln(1+x) \leqslant x$. □

Twierdzenie 2.9 (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 $t > 0$,*

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

oraz dla $s > 0$,

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

gdzie

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

Dowód. Pierwsza część wynika natychmiast z Lematu 2.8. By pokazać drugą zauważamy, że dla $S = \sum_{i=1}^n X_i$,

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

Prosty rachunek pokazuje, że powyższe supremum jest osiągane w punkcie

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

i wynosi

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

gdzie ostatnie oszacowanie wynika z poniższego lematu. □

Lemat 2.10. *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 $x \geqslant 0$. □

Uwaga 2.6 Jeśli $\mathbf{P}(X_{n,i} = 1) = 1 - \mathbf{P}(X_{n,i}) = 1/n$ oraz $X_{n,i}$ są niezależne, to rozkład $\sum_{i=1}^n X_{i,n}$ zbiega do rozkładu Poissona z parametrem 1. Oczywiście $\sum_{i=1}^n \mathbf{E}X_{i,n}^2 = 1 = \max_i \|X_{i,n}\|_\infty$. To pokazuje, że przy założeniach Twierdzenia 2.9 nie można uzyskać przy $s \to \infty$ oszacowania lepszego rzędu niż $s\ln s$.

3 Nierówności maksymalne

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

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

Przy dowodzeniu twierdzeń granicznych przydają się nierówności pozwalające szacować $\max_k \|S_k\|$ przez pojedynczą zmienną S_k.

3.1 Nierówności Levy'ego i Levy'ego-Octavianiego

Twierdzenie 3.1 (Nierówność Levy'ego)**.** *Załóżmy, że X_i są niezależnymi symetrycznymi wektorami losowymi w F. Wówczas dla $s > 0$,*

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

Dowód. Niech

$$A_k := \{\|S_1\| < s, \ldots, \|S_{k-1}\| < s, \|S_k\| \geqslant s\}.$$

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

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

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

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

czyli

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

□

Wniosek 3.2. *Załóżmy, że zmienne X_i sa niezależne i symetryczne oraz $|a_i| \leqslant 1$. Wówczas*

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

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

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

czyli $\| \sum_{i=1}^n a_i X_i \| \leqslant \max_{1\leqslant k\leqslant n} \|S_k\|$ i teza wniosku wynika z nierówności Levy'ego. $\square$

Uwaga 3.7 Jeśli założenia Wniosku 3.2 są spełnione, to dla dowolnej funkcji wypukłej $f\colon F \to \mathbb{R}_+$,

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

w szczególności $\| \sum_{i=1}^n a_i X_i \|_p \leqslant \| \sum_{i=1}^n X_i \|_p$. Nie jest jednak prawdą, że nierówność z Wniosku 3.2 zachodzi ze stałą 1.

Przejdźmy teraz do przypadku niesymetrycznego.

Lemat 3.3. *Dla dowolnych niezależnych zmiennych losowych X_i oraz $u, t > 0$,*

$$\mathbf{P}\Big(\max_{1\leqslant k\leqslant n} \|S_k\| \geqslant t + u \Big) \leqslant \frac{\mathbf{P}(\|S_n\| \geqslant u)}{1 - \max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_n - S_k\| \geqslant t)}, \tag{7}$$

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

Dowód. Zdefiniujmy

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

zbiory A_k są parami rozłączne oraz

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

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

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

Otrzymujemy

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

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

Twierdzenie 3.4 (Nierówność Levy'ego-Octavianiego)**.** . *Załóżmy, że X_i są niezależnymi wektorami losowymi w F. Wówczas dla $s > 0$,*

$$\mathbf{P}\Big(\max_{1\leqslant k\leqslant n} \|S_k\| \geqslant s\Big) \leqslant 3 \max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_k\| \geqslant s/3). \tag{8}$$

Dowód. Szacujemy

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

zatem

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

Oczywiście by udowodnić (8) możemy założyć, że $\max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_k\| \geqslant s/3) < 1/3$. Kładąc $u = s/3$ oraz $t = 2s/3$ w (7) dostajemy

$$\mathbf{P}\Big(\max_{1\leqslant k\leqslant n} \|S_k\| \geqslant s\Big) \leqslant \frac{\mathbf{P}(\|S_n\| \geqslant s/3)}{1 - 2\max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_k\| \geqslant s/3)} \leqslant 3\mathbf{P}(\|S_n\| \geqslant s/3).$$

□

3.2 Nierówności dla sum zmiennych o jednakowym rozkładzie

W tym paragrafie będziemy dodatkowo zakładać, że

Niezależne zmienne X_i o wartościach w F mają jednakowy rozkład. (9)

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

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

Lemat 3.5. *Jeśli X i Y są niezależne, to*

$$A(X + Y, s) \geqslant A(X, s) \ \textit{dla}\ s > 0.$$

Dowód. Dla dowolego $x \in F$ dostajemy

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

□

Lemat 3.6. *Załóżmy, że zachodzi (9) oraz $\mathbf{P}(\|S_n\| > s/5) < 1/4$. Wówczas*

$$\max_{1\leqslant k\leqslant n} \mathbf{P}(\|S_k\| > s) \leqslant \mathbf{P}(\|S_n\| > s/5).$$

Dowód. Ustalmy liczbę α taką, że $1/4 > \alpha > \mathbf{P}(\|S_n\| > s/5)$. Wówczas oczywiście $A(S_n, s/5) < \alpha$, czyli na mocy Lematu 3.5, $A(S_k, s/5) < \alpha$ dla $k = 1, \ldots, n$. Możemy zatem wybrać $x_1, \ldots, x_{n-1} \in F$ dla których

$$\mathbf{P}(\|S_k - x_k\| > s/5) \leqslant \alpha, \ k = 1, \ldots, n-1.$$

Wybierzmy $i \leqslant n-1$ takie, że $\|x_i\| = \max_{1 \leqslant j \leqslant n-1} \|x_j\|$. Wiemy, że $\mathbf{P}(\|S_i - x_i\| > s/5) \leqslant \alpha$ oraz $\mathbf{P}(\|S_n\| > s/5) \leqslant \alpha)$, stąd

$$\mathbf{P}(\|S_{n-i} + x_i\| > 2s/5) = \mathbf{P}(\|S_n - (S_i - x_i)\| > 2s/5) \leqslant 2\alpha.$$

Rozpatrzymy trzy przypadki.

i) $i = n - i$. Mamy

$$\mathbf{P}(\{\|S_i - x_i\| > s/5\} \cup \{\|S_{n-i} + x_i\| > 2s/5\}) < 3\alpha < 1,$$

więc istnieje zdarzenie elementarne ω dla którego $\|S_i(w) - x_i\| \leqslant s/5$ oraz $\|S_{n-i}(w) + x_i\| = \|S_i(w) + x_i\| \leqslant 2s/5$. Na mocy nierówności trójkąta dostajemy $\|2x_i\| \leqslant 3s/5$.

ii) $i > n - i$. Mamy

$$\begin{aligned}\mathbf{P}(\|S_{2i-n} - 2x_i\| > 3s/5) &= \mathbf{P}(\|(S_i - x_i) - (S_i - S_{2i-n} + x_i)\| > 3s/5) \\ &\leqslant \mathbf{P}(\|S_i - x_i\| > s/5) + \mathbf{P}(\|S_i - S_{2i-n} + x_i\| > 2s/5) \\ &\leqslant \alpha + \mathbf{P}(\|S_{n-i} + x_i\| > 2s/5) \leqslant 3\alpha.\end{aligned}$$

Zatem

$$\mathbf{P}(\{\|S_{2i-n} - 2x_i\| > 3s/5\} \cup \{\|S_{2i-n} - x_{2i-n}\| > 2s/5\}) < 4\alpha < 1,$$

czyli istnieje ω takie, że $\|S_{2i-n}(w) - 2x_i\| \leqslant 3s/5$ and $\|S_{2i-n}(w) - x_{2i-n}\| \leqslant s/5$. Zatem $\|x_{2i-n} - 2x_i\| \leqslant 4s/5$. Stąd $\|x_i\| \geqslant \|x_{2i-n}\| \geqslant 2\|x_i\| - 4s/5$, czyli w tym przypadku $\|x_i\| \leqslant 4s/5$.

iii) $i < n - i$. Dostajemy

$$\begin{aligned}\mathbf{P}(\|S_{n-2i} + 2x_i\| > 3s/5) &= \mathbf{P}(\|(S_{n-i} + x_i) - (S_{n-i} - S_{n-2i} - x_i)\| > 3s/5) \\ &\leqslant \mathbf{P}(\|S_{n-i} + x_i\| > 2s/5) + \mathbf{P}(\|S_{n-i} - S_{n-2i} - x_i\| > s/5) \\ &\leqslant 2\alpha + \mathbf{P}(\|S_i - x_i\| > s/5) \leqslant 3\alpha.\end{aligned}$$

Tak samo jak w poprzednim przypadku pokazujemy, że $\|x_{n-2i} + 2x_i\| \leqslant 4s/5$ oraz $\|x_i\| \leqslant 4s/5$.

Udowodniliśmy zatem, że $\max_{1 \leqslant j \leqslant n-1} \|x_j\| \leqslant 4s/5$, stąd

$$\max_{1 \leqslant k \leqslant n} \mathbf{P}(\|S_k\| > s) \leqslant \max_{1 \leqslant k \leqslant n} \mathbf{P}(\|S_k - x_k\| > s/5) \leqslant \alpha,$$

co z dowolności $\alpha \in (\mathbf{P}(\|S_n\| > s/5), 1/4)$ dowodzi tezy. □

Twierdzenie 3.7. *Przy założeniu (9),*

$$\mathbf{P}\Big(\max_{1 \leqslant k \leqslant n} \|S_k\| > s\Big) \leqslant 4\mathbf{P}(\|S_n\| > s/6).$$

Dowód. Oczywiście możemy zakładać, że $\mathbf{P}(\|S_n\| > s/6) < 1/4$. Wówczas na mocy Lematu 3.6, $\max_{1 \leqslant k \leqslant n} \mathbf{P}(\|S_k\| > 5s/6) \leqslant \mathbf{P}(\|S_n\| > s/6) < 1/4$. Z nierówności (7) (z $t = 5s/6, u = s/6$),

$$\mathbf{P}\Big(\max_{1 \leqslant k \leqslant n} \|S_k\| > s\Big) \leqslant \frac{\mathbf{P}(\|S_n\| > s/6)}{1 - \max_{1 \leqslant k \leqslant n} \mathbf{P}(\|S_n - S_k\| > 5s/6)}$$
$$= \frac{\mathbf{P}(\|S_n\| > s/6)}{1 - \max_{1 \leqslant k \leqslant n} \mathbf{P}(\|S_{n-k}\| > 5s/6)} \leqslant \frac{4}{3}\mathbf{P}(\|S_n\| > s/6).$$

□

Wniosek 3.8. *Załóżmy, że zachodzi (9) oraz* $0 \leqslant a_i \leqslant 1$. *Wówczas*

$$\mathbf{P}\Big(\Big\|\sum_{i=1}^{n} a_i X_i\Big\| > s\Big) \leqslant 4\mathbf{P}(\|S_n\| > s/6).$$

Dowód. Możemy zakładać, że $a_1 \geqslant a_2 \geqslant \ldots \geqslant a_n$ i wtedy tak jak we Wniosku 3.2 dostajemy $\|\sum_{i=1}^{n} a_i X_i\| \leqslant \max_{1 \leqslant k \leqslant n} \|S_k\|$. □

4 Mocne prawa wielkich liczb

Będziemy zakładali, że $X, X_1, X_2, \ldots$ są niezależnymi zmiennymi losowymi o wartościach w ośrodkowej przestrzeni Banacha $(F, \| \ \|)$.

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. *Jeśli X jest zmienną losową o wartościach w ośrodkowej przestrzeni Banacha 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. Proste ćwiczenie. □

Twierdzenie 4.2. *Załóżmy, że F jest dowolną ośrodkową przestrzenią Banacha oraz $\mathbf{E}\|X\| < \infty$. Wówczas $\lim_{n\to\infty} \frac{S_n}{n} = \mathbf{E}X$ p.n..*

Dowód. Zauważmy, że 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.

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

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

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

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

Dla $\omega \in A$, dostajemy $\|X_n(\omega)\| \leqslant \|S_n(\omega)\| + \|S_{n-1}(\omega)\| \leqslant C(\omega)n^{1/p}$ dla 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,

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

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

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

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

Wówczas, jeśli

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

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

Dowód. Szacujemy dla $\varepsilon > 0$,

$$\mathbf{P}\Big(\sup_{m>2^{k-1}} \frac{\|S_m\|}{\gamma_m} \geqslant \varepsilon\Big) \leqslant \sum_{n=k}^\infty \mathbf{P}\Big(\sup_{2^{n-1}<m\leqslant 2^n} \|S_m\| \geqslant C^{-1}\varepsilon\gamma_{2^n}\Big)$$
$$\leqslant \sum_{n=k}^\infty 4\mathbf{P}\big(\|S_{2^n}\| \geqslant C^{-1}\varepsilon\gamma_{2^n}/6\big) \to 0 \text{ przy } k \to \infty,$$

gdzie ostatnia nierówność wynika z Twierdzenia 3.7. □

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)$\Rightarrow$i). Wpierw udowodnimy, że warunek ii) implikuje

$$\lim_{n\to\infty} n^{1-1/p}\mathbf{E}[XI_{\{\|X\|\leqslant n^{1/p}\}}] = 0. \tag{10}$$

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

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

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

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

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 (10) zostało wykazane.

Na mocy Lematu 4.4 by wykazać i) wystarczy, iż wykażemy, że dla $\varepsilon > 0$, $\sum_n \mathbf{P}(\|S_{2^n}\| \geqslant 2\varepsilon 2^{n/p}) < \infty$. Z (10) wynika, że $2^n\mathbf{E}[XI_{\{\|X\|\leqslant 2^{n/p}\}}] \leqslant \varepsilon 2^{n/p}$ dla dużych n, więc wystarczy, iż udowodnimy, że

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

Określmy

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

wówczas

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

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

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

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 $p = 1$), $\frac{S_n}{n} \to \mathbf{E}X$. Stąd $\mathbf{E}X = 0$. □

Uwaga 4.8 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 4.9 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

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

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 4.6 (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

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

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 3.4,

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

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

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

□

Uwaga 4.10 Twierdzenie Brunka (z tym samym dowodem) zachodzi też dla przestrzeni Hilberta oraz ogólniej dla przestrzeni F typu 2.

5 Prawo Iterowanego Logarytmu

W tej części, jeśli nie zaznaczymy inaczej, będziemy zakładać, że $X, X_1, X_2, \ldots$ są niezależnymi zmiennymi losowymi o jednakowym rozkładzie o średniej zero i wariancji σ^2. Jak zwykle też $S_k = X_1 + \ldots X_k$.

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

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

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 $t > 0$,*

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

Dowód. Z lematu 3.3,

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

Mamy jednak na mocy nierówności Czebyszewa,

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

□

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,

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

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

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

Stąd,

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

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

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

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

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

Niech $T_n = \sum_{i=1}^n X_i^2$. Na mocy Silnego 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,

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

Szacujemy,

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

Stąd dla symetrycznych zmiennych losowych,

$$\sum_n \mathbf{P}(|S_{2^n}| \geqslant 2\gamma(2^n)\sigma) < \infty. \tag{11}$$

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 (11),

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

Możemy więc stosować Lemat 5.3 z $a = 4$ 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 2.6),

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

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

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

Możemy więc stosować Lemat 5.3 (z $a = \sigma\sqrt{1+2\varepsilon}$) i dostajemy

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

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

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

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

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

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

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

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

Mamy jednak $\sigma_Y^2 \leqslant \mathbf{E}X^2 I_{\{|X|\leqslant M\}} \leqslant \mathbf{E}X^2 = \sigma^2$ oraz $\sigma_Z^2 \leqslant \mathbf{E}X^2 I_{\{|X|>M\}} \to 0$ 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. Zmienne $S_{C^n} - S_{C^{n-1}}$, $n = 1, 2, \ldots$ są niezależne oraz

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

Zatem na mocy Lematu Borela-Cantelliego,

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

Stąd na podstawie Wniosku 5.6,

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

□

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 $\mathbf{P}(S_n\geqslant(1+\varepsilon)^{-2}\gamma(n)\sigma)$ stosując odwrotną nierówność wykładniczą Kołmogorowa (Twierdzenie 2.7) z $s=(1+\varepsilon)^{-2}\gamma(n)\sigma$ i $n\sigma^2$ zamiast σ^2. Zauważmy, że dla dużych $n\geqslant n(\varepsilon)$ jego założenia są spełnione, bo

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

oraz

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

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

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

Zatem

$$\begin{aligned}\sum_n\mathbf{P}(S_{C^n-C^{n-1}})\geqslant(1+\varepsilon)^{-2}\sigma\gamma(C^n-C^{n-1}))&\\\geqslant\sum_{n\geqslant n(\varepsilon)}(\log((C-1)C^{n-1}))^{-1/(1+\varepsilon)}=\infty.\end{aligned}$$

Czyli, na mocy Lematu 5.7,

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

Biorąc $C\to\infty$ oraz $\varepsilon\to0+$ dostajemy tezę. □

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

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

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,

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

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

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

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

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

□

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

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

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

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

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

$$X_n^t := X_n I_{\{|X_n| \leqslant t\}} - X_n I_{\{|X_n| > t\}}$$

oraz $S_n^t := \sum_{i=1}^n X_i^t$. Zmienne (X_n^t) mają ten sam rozkład co zmienne X_n, zatem $\limsup_{n\to\infty} \frac{|S_n^t|}{\sqrt{2n \ln \ln n}} \leqslant M$ p.n., czyli

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

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

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

Stąd $\mathbf{E}X^2 I_{\{|X|\leqslant t\}} \leqslant M^2$ 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

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

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

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

$\square$

5.1 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 określić zbiór wszystkich możliwych granic jego podciągów. Wprowadźmy najpierw definicję.

Definicja 5.1. *Dla ciągu liczbowego (lub ogólniej o wartościach w przestrzeni metrycznej) a_n określamy $C(a_n)$ zbiór wszystkich punktów skupienia ciągu a_n, czyli*

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

Twierdzenie 5.11. *Załóżmy, że $\mathbf{E}X = 0$ oraz $\mathbf{E}X^2 = \sigma^2 < \infty$. Wówczas*
i) $\operatorname{dist}(\frac{S_n}{\sqrt{2n\ln\ln n}}, [-\sigma,\sigma]) \to 0$ p.n.,
ii) $C(\frac{S_n}{\sqrt{2n\ln\ln n}}) = [-\sigma,\sigma]$ p.n..

Dowód opiera się na następującym prostym lemacie.

Lemat 5.12. *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) $\operatorname{dist}(a_n, [b,c]) \to 0$,
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 $\operatorname{dist}(a_n, [b,c]) < \varepsilon$.

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

$$m_k := \sup\{n \leqslant n_k \colon a_n \leqslant d\},$$

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 $d \in C(a_n)$. $\square$

Dowód Twierdzenia 5.11. Na mocy Lematu 5.12 oraz Twierdzenia 5.1 wystarczy wykazać, że

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

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

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

Szacujemy,

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

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

5.2 Przypadek wektorowy

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

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

Twierdzenie 5.13. *Jeśli* $\mathbf{E}X = 0$ *oraz* $\mathbf{E}|X|^2 < \infty$*, to*

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

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

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

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

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

Dla $x \in \mathbb{R}$, 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

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

Z dowolności $\varepsilon > 0$,

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

□

Twierdzenie 5.14. *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. Szacowanie $\langle x, t\rangle \leqslant |x||t|$ oraz Twierdzenie 5.10 implikują $\mathbf{E}\langle X, t\rangle = 0$ oraz $\mathbf{E}\langle X, t\rangle^2 < \infty$ dla dowolnego $t \in \mathbb{R}^d$. □

Dla wektora X 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 5.15. *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 $\xi = \sum_{i=1}^d X_i + \eta$, gdzie $\mathbf{E}\eta X_i = 0$ dla $i = 1, \ldots, d$. Mamy

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

oraz

$$\mathbf{E}(\xi X_j) = \sum_{i=1}^{d} t_i \mathbf{E} X_i X_j = \sum_{i=1}^{d} t_i c_{ji} = Ct,$$

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

Twierdzenie 5.16. *Załóżmy, że* $\mathbf{E}X = 0$ *oraz* $\mathbf{E}|X|^2 < \infty$. *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. Przyjmijmy $\gamma(n) := \sqrt{2n\ln\ln n}$ i zacznijmy od rozpatrzenia przypadku $C = \mathrm{Cov}(X) = Id$. 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 5.13).

By udowodnić b) dla $C = Id$ 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

$$\limsup_{k\to\infty} \Big|\frac{S_{n_k}}{\gamma(n_k)} - 1\Big|^2 \limsup_{k\to\infty} \Big(\frac{|S_{n_k}|}{\gamma(n_k)}\Big)^2 + |t|^2 - 2\lim_{k\to\infty} \frac{\langle S_{n_k}, t\rangle}{\gamma(n_k)} = 0.$$

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 $Y_i = (X_i, \varepsilon_i)$. Wówczas (Y_i) jest ciągiem niezależnych wektorów losowych w $\mathbb{R}^{d+1}$ o sredniej zero i macierzy kowariancji Id. Jeśli położymy $\tilde{S}_n := \sum_{i\leqslant n} Y_i$, to na mocy poprzednich rozważań,

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

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

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

oraz

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

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

6 Rozkłady nieskończenie podzielne

Definicja 6.1. *Rozkład probabilistyczny μ na $\mathbb{R}$ nazywamy* nieskończenie podzielnym*, jeśli dla każdego $n = 1, 2, \ldots$ istnieje rozkład $\mu_{1/n}$ taki, że $\mu = \mu_{1/n}^{*n}$.*

Mówimy, że zmienna losowa X jest nieskończenie podzielna*, jeśli rozkład X jest nieskończenie podzielny, czyli dla każdego $n = 1, 2, \ldots$ istnieją zmienne $X_{n,1}, \ldots, X_{n,n}$ o jednakowym rozkładzie takie, że $X \sim X_{n,1} + \ldots + X_{n,1}$.*

Klasę wszystkich rozkładów nieskończenie podzielnych będziemy oznaczać przez ID.

Przykłady.
i) $\mu = \delta_a$, $\mu_{1/n} = \delta_{a/n}$
ii) $\mu = \mathcal{N}(a, \sigma^2)$, $\mu_{1/n} = \mathcal{N}(a/n, \sigma^2/n)$,
iii) $\mu = \text{Poiss}(\lambda)$, $\mu_{1/n} = \text{Poiss}(\lambda/n)$,
iv) $\mu = \Gamma(\alpha)$, $\mu_{1/n} = \Gamma(\alpha/n)$.

Definicja 6.2. *Dla $\lambda \geqslant 0$ oraz ν miary probabilistycznej na $\mathbb{R}$ definiujemy* złożony rozkład Poissona $\pi_{\lambda\nu}$ *wzorem*

$$\pi_{\lambda\nu} := e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \mu^{*n}.$$

Przez CP będziemy oznaczać klasę wszystkich złożonych rozkładów Poissona, czyli

$$CP := \{\pi_{\lambda\nu} \colon \lambda \geqslant 0, \nu \text{ miara probabilistyczna na } \mathbb{R}\}.$$

Domknięcie CP w topologii zbieżności według rozkładu będziemy oznaczać przez $\overline{CP}$ czyli

$$\mu \in \overline{CP} \Leftrightarrow \exists_{\mu_n \in CP}\ \mu_n \Rightarrow \mu.$$

Dla uproszczenia i ujednolicenia notacji wprowadźmy następującą definicję.

Definicja 6.3. *Dla miar skończonych ν na $\mathbb{R}$ definiujemy* transformatę Fouriera *ν wzorem*

$$\widehat{\nu}(t) := \int_{\mathbb{R}} e^{itx} d\nu(x).$$

Podobnie dla funkcji $f \in L^1(\mathbb{R})$ określamy

$$\widehat{f}(t) := \int_{\mathbb{R}} e^{itx} f(x) dx.$$

Oczywiście gdy ν jest miarą probabilistyczną, to $\widehat{\nu} = \varphi_\nu$ jest funkcją charakterystyczną ν.

Twierdzenie 6.1. *Dowolny rozkład nieskończenie podzielny jest granicą złożonych rozkładów Poissona, czyli $ID \subset \overline{CP}$.*

Dowód. Weźmy $\mu \in ID$ oraz zdefiniujmy

$$\pi_n := e^{-n} \sum_{k=0}^{\infty} \frac{n^k}{k!} \mu_{1/n}^{*k},$$

pokażemy, że $\pi_n \Rightarrow \mu$. W tym celu wystarczy, że wykażemy iż

$$\int \varphi d\pi_n \to \int \varphi d\mu \text{ dla dowolnej ograniczonej funkcji Lipschitzowskiej } \varphi.$$

Ustalmy więc ograniczoną funkcję Lipschitzowską φ ze stałą Lipschitza L i dla $R > 0$ określmy

$$\Delta_n(R) := e^{-n} \sum_{|k-n| \leqslant \sqrt{n}R} \frac{n^k}{k!} \Big| \int \varphi d\mu_{1/n}^{*k} - \int \varphi d\mu \Big|.$$

Wówczas,

$$\begin{aligned} \Big| \int \varphi d\pi_n - \int \varphi d\mu \Big| &\leqslant e^{-n} \sum_{k \geqslant 0} \frac{n^k}{k!} \Big| \int \varphi d\mu_{1/n}^{*k} - \int \varphi d\mu \Big| \\ &\leqslant \Delta_n(R) + 2\|\varphi\|_\infty \varepsilon_n(R), \end{aligned}$$

gdzie

$$\begin{aligned} \varepsilon_n(R) := e^{-n} \sum_{|k-n| > \sqrt{n}R} \frac{n^k}{k!} &\leqslant \frac{1}{nR^2} e^{-n} \sum_{k \geqslant 0} \frac{n^k}{k!} (n-k)^2 \\ = \frac{1}{nR^2} \mathrm{Var}(\mathrm{Poiss}(n)) = \frac{1}{R^2}. \end{aligned}$$

Wykazaliśmy więc, że

$$\Big| \int \varphi d\pi_n - \int \varphi d\mu \Big| \leqslant \Delta_n(R) + 2R^{-2}\|\varphi\|_\infty$$

i by zakończyć dowód wystarczy pokazać, że

$$\forall_{R>0} \lim_{n \to \infty} \Delta_n(R) = 0. \tag{12}$$

Szacujemy

$$\begin{aligned} \Delta_n(R) &\leqslant e^{-n} \sum_{|k-n| \leqslant \sqrt{n}R} \frac{n^k}{k!} \sup_{|k-n| \leqslant \sqrt{n}R} \Big| \int \varphi d\mu_{1/n}^{*k} - \int \varphi d\mu_{1/n}^{*n} \Big| \\ &\leqslant \sup_{0 \leqslant l \leqslant \sqrt{n}R, m \geqslant 0} \Big| \int \varphi d\mu_{1/n}^{*(m+l)} - \int \varphi d\mu_{1/n}^{*m} \Big|. \end{aligned}$$

Zauważmy, że dla miar probabilistycznych ν_1 i ν_2,

$$\Big|\int \varphi d\nu_1 * \nu_2 - \int \varphi d\nu_1\Big| = \Big|\int\int (\varphi(x+y) - \varphi(x))d\nu_2(y)d\nu_1(x)\Big|$$
$$\leqslant \sup_x \int |\varphi(x+y) - \varphi(x)|d\nu_2(y),$$

ponadto dla dowolnego x na mocy Lipschitzowskości φ,

$$\int |\varphi(x+y) - \varphi(x)|d\nu_2(y) = \Big(\int_{|y|\leqslant r} + \int_{|y|>r}\Big)|\varphi(x+y) - \varphi(x)|d\nu_2(y)$$
$$\leqslant rL + 2\|\varphi\|_\infty \nu_2(\mathbb{R} \setminus [-r,r]).$$

Zatem by wykazać (12) pozostaje udowodnić, że

$$\forall R, r > 0 \ \lim_{n\to\infty} \sup_{0\leqslant l\leqslant\sqrt{n}R} \mu_{1/n}^{*l}(\mathbb{R} \setminus [-r,r]) = 0. \tag{13}$$

Udowodnijmy wpierw następujące oszacowanie.

Lemat 6.2. *Dla* $|v| \leqslant 1/2$ *oraz* $\alpha \in [0,1]$, $|e^{\alpha\log(1+v)} - 1| \leqslant 2\alpha|v|$, *gdzie* $\log(z)$ *oznacza kanoniczną gałąź logarytmu w kole* $|z-1| \leqslant 1/2$.

Dowód. Mamy

$$e^{\alpha\log(1+v)} - 1 = \int_0^1 \frac{d}{dt}e^{\alpha\log(1+tv)}dt = \int_0^1 \alpha v(1+tv)^{\alpha-1}dt,$$

zatem

$$|e^{\alpha\log(1+v)} - 1| \leqslant \alpha|v| \int_0^1 |1+tv|^{\alpha-1} \leqslant 2\alpha|v|,$$

gdzie ostatnia nierówność wynika stąd, że $|1+tv| \geqslant 1 - t|v| \geqslant 1/2$ oraz $0 \leqslant \alpha - 1 \leqslant -1$. □

Ustalmy $\delta > 0$ takie, że $|\widehat{\mu}(t) - 1| \leqslant 1/2$ dla $|t| \leqslant \delta$. Ponieważ $\widehat{\mu_{1/n}}^n = \widehat{\mu}$, więc

$$\widehat{\mu_{1/n}^{*l}}(t) = \text{"}\widehat{\mu}^{l/n}(t)\text{"} = e^{\frac{l}{n}\log\widehat{\mu}(t)} \text{ dla } |t| \leqslant \delta,$$

gdzie log oznacza standardową gałąź logarytmu w kole $|z-1| \leqslant 1/2$. Na podstawie Lematu 6.2,

$$\forall_{|t|\leqslant\delta} \ \big|\widehat{\mu_{1/n}^{*l}}(t) - 1\big| \leqslant 2\frac{l}{n}|\widehat{\mu}(t) - 1| \leqslant \frac{l}{n}.$$

Dla miary ν liczymy

$$\frac{1}{\delta}\int_0^\delta (1 - \mathrm{Re}(\widehat{\nu}(t))dt = \frac{1}{\delta}\int_0^\delta \int_{\mathbb{R}} (1 - \cos(tx))d\nu(x)dt$$
$$= \int_{\mathbb{R}} \frac{1}{\delta}\int_0^\delta (1 - \cos(tx))dtd\nu(x) = \int_{\mathbb{R}} (1 - \frac{\sin(x\delta)}{x\delta})d\nu(x)$$
$$\geqslant m(r\delta)\nu(\mathbb{R} \setminus [-r,r]),$$

gdzie

$$m(s) := \inf\{1 - \frac{\sin(y)}{y} : |y| \geqslant s\} > 0 \text{ dla } s > 0.$$

Zatem

$$\mu_{1/n}^{*l}(\mathbb{R} \setminus [-r, r]) \leqslant \frac{1}{m(r\delta)} \frac{1}{\delta} \int_0^\delta \left|1 - \widehat{\mu_{1/n}^{*l}}(t)\right| dt \leqslant \frac{l}{nm(r\delta)},$$

co łatwo implikuje (13). □

Uwaga. Można udowodnić bezpośrednio, że słaba granica rozkładów nieskończenie podzielnych jest nieskończenie podzielna, skąd łatwo wynika, że $ID = \overline{CP}$. My jednak pójdziemy bardziej skomplikowaną drogą, gdyż zależy nam na dobrym opisaniu klasy ID.

Lemat 6.3. *Mamy*

$$\widehat{\pi_{\lambda\nu}}(t) = \exp\Big(\int (e^{itx} - 1) d\lambda\nu(x)\Big) =: f_{\lambda\nu}(t).$$

Dowód. Liczymy

$$\widehat{\pi_{\lambda\nu}}(t) = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \widehat{\nu}^n(t) = \exp(-\lambda + \lambda\widehat{\nu}(t)) = \exp\Big(\int (e^{itx} - 1) d\lambda\nu(x)\Big).$$

□

By podać klasyfikację rozkładów nieskończenie podzielnych wprowadzimy następującą definicję.

Definicja 6.4. *Niech*

$$\mathcal{M} := \{\textit{miary nieujemne } \nu \textit{ na } \mathbb{R} \setminus \{0\} \colon \int 1 \wedge x^2 d\nu(x) < \infty\}.$$

Elementy zbioru $\mathcal{M}$ nazywamy miarami Levy'ego. *Klasę $\mathcal{L}$* układów Levy'ego *określamy jako*

$$\mathcal{L} := \mathbb{R} \times [0, \infty) \times \mathcal{M} = \{(a, \sigma, \nu) \colon a \in \mathbb{R}, \sigma \geqslant 0, \nu \in \mathcal{M}\}.$$

Dla $(a, \sigma, \nu) \in \mathcal{L}$ definujemy $f_{a,\sigma^2,\nu} := \exp(l_{a,\sigma^2,\nu})$, gdzie

$$l_{a,\sigma^2,\nu} := ita - \frac{\sigma^2}{2} t^2 + \int_{\mathbb{R}\setminus\{0\}} \Big(e^{itx} - 1 - \frac{itx}{1 + x^2}\Big) d\nu(x).$$

Twierdzenie 6.4. *Dla $(a, \sigma, \nu) \in \mathcal{L}$ istnieje miara probabilistyczna $\pi_{a,\sigma^2,\nu} \in ID$ taka, że $\widehat{\pi_{a,\sigma^2,\nu}} = f_{a,\sigma^2,\nu}$.*

Dowód. Przeprowadzimy konstrukcję $\pi_{a,\sigma^2,\nu}$ w trzech krokach.

Krok I. Dla $a \in \mathbb{R}$ oraz miary skończonej ν na $\mathbb{R}$ istnieje rozkład $\pi_{a,\nu}$ taki, że

$$\widehat{\pi_{a,\nu}}(t) = f_{a,\nu}(t) = \exp\Big(ita + \int (e^{itx} - 1) d\nu(x)\Big).$$

Mamy $f_{a,\nu}(t) = e^{iat} f_\nu(t)$, więc za $\pi_{a,\nu}$ wystarczy wziąć przesunięcie π_ν o a i skorzystać z Lematu 6.3.

Krok II. Dla $a \in \mathbb{R}$ oraz miary Levy'ego $\nu \in \mathcal{M}$ istnieje rozkład $\pi_{a,0,\nu}$ taki, że

$$\widehat{\pi_{a,0,\nu}}(t) = f_{a,0,\nu}(t) = \exp\Big(ita + \int (e^{itx} - 1 - \frac{itx}{1+x^2}) d\nu(x)\Big).$$

Zauważmy, że $|e^{itx} - 1 - itx| \leqslant t^2 x^2/2$, więc

$$\Big|e^{itx} - 1 - \frac{itx}{1+x^2}\Big| \leqslant \min\Big(2 + |t|, \frac{1}{2}|t|^2 x^2 + \frac{|t||x|^3}{1+x^2}\Big) \leqslant \Big(2 + |t| + \frac{t^2}{2}\Big)\min(1, x^2),$$

skąd wynika, że funkcja $f_{a,0,\nu}$ jest dobrze określona i ciągła w zerze. Określmy miarę skończoną ν_ε wzorem $\nu_\varepsilon(A) := \nu(A \setminus (-\varepsilon, \varepsilon))$, wówczas

$$f_{a,0,\nu_\varepsilon}(t) = \exp\Big(it(a - \int_{|x|\geqslant\varepsilon} \frac{x}{1+x^2} d\nu) + \int (e^{itx} - 1) d\nu_\varepsilon(x)\Big) = \widehat{\pi_{a_\varepsilon,\nu_\varepsilon}}$$

dla $a_\varepsilon := a - \int_{|x|\geqslant\varepsilon} \frac{x}{1+x^2} d\nu$. Ponieważ $f_{a,0,\nu_\varepsilon}$ zbiega punktowo do $f_{a,0,\nu}$, więc na mocy twierdzenia Levy'ego $\pi_{a_\varepsilon,\nu_\varepsilon}$ zbiega słabo do szukanego rozkładu $\pi_{a,0,\nu}$.

Krok III. Dla $(a, \sigma, \nu) \in \mathcal{L}$ istnieje miara probabilistyczna $\pi_{a,\sigma^2,\nu} \in ID$ taka, że $\widehat{\pi_{a,\sigma^2,\nu}} = f_{a,\sigma^2,\nu}$.

Za $\pi_{a,\sigma^2,\nu}$ bierzemy rozkład $X + Y$, gdzie X i Y są niezależne, $X \sim \pi_{a,0,\nu}$, a $Y \sim \mathcal{N}(0, \sigma^2)$.

Z postaci funkcji charakterystycznej natychmiast dostajemy, że $\pi_{a,\sigma^2,\nu} = \pi^{*n}_{a/n,\sigma^2/n,\nu/n}$, czyli $\pi_{a,\sigma^2,\nu} \in ID$. □

Fakt 6.5. *Jeśli $f_{a_1,\sigma_1,\nu_1} = f_{a_2,\sigma_2,\nu_2}$ dla pewnych $(a_i, \sigma_i, \nu_i) \in \mathcal{L}$, to $a_1 = a_2, \sigma_1 = \sigma_2$ oraz $\nu_1 = \nu_2$.*

Fakt 6.6. *Załóżmy, że $\pi_{a_n,\sigma_n,\nu_n} \Rightarrow \mu$ dla pewnego ciągu układów Levy'ego $(a_n, \sigma_n, \nu_n) \in \mathcal{L}$. Wówczas istnieje $(a, \sigma, \nu) \in \mathcal{L}$ taki, że $\mu = \pi_{a,\sigma,\nu}$.*

Twierdzenie 6.7. *Jeśli μ jest rozkładem nieskończenie podzielnym to istnieje dokładnie jeden układ Levy'ego $(a, \sigma, \nu) \in \mathcal{L}$ taki, że $\widehat{\mu} = f_{a,\sigma,\nu}$.*

6.1 Rozkłady nieskończenie podzielne jako granice sum

Przypomnijmy, że macierz zmiennych losowych $(X_{n,k})_{k\leqslant k_n, n=1,2,\ldots}$ nazywamy *układem trójkątnym*, jeśli dla każdego n zmienne $X_{n,k}$, $1 \leqslant k \leqslant k_n$ są niezależne. Połóżmy $S_n := \sum_{k=1}^{k_n} X_{n,k}$.

Centralne twierdzenie graniczne mówi, że jeśli $\mathbf{E}S_n \to a$, $\mathrm{Var}(S_n) \to \sigma^2$ oraz zachodzi warunek Lindeberga, to S_n zbiega według rozkładu do zmiennej $\mathcal{N}(a, \sigma^2)$.

Twierdzenie Poissona stwierdza, że jeśli $\mathbf{P}(X_{n,k} = 1) = p_n = 1 - \mathbf{P}(X_{n,k} = 0)$, $k_n \to \infty$ oraz $k_n p_n \to \lambda$, to S_n zbiega według rozkładu do zmiennej Poiss(λ).

Można się zapytać jakie inne rozkłady można uzyskać jako granice sum układów trójkątnych? Oczywiście trzeba coś założyć, inaczej wystarczy przyjąć $X_{n,1} \sim X$ oraz wziąć $X_{n,2} + \ldots + X_{n,k_n}$ zbieżne do zera by otrzymać X jako granicę S_n. Naturalnym założeniem jest

$$\forall_{\varepsilon>0} \lim_{n\to\infty} \sup_{k\leqslant k_n} \mathbf{P}(|X_{n,k}| \geqslant \varepsilon) = 0 \tag{14}$$

mówiące, że wkład pojedynczej zmiennej w sumę S_n jest granicznie zaniedbywalny.

Celem tego paragrafu jest udowodnienie następującego twierdzenia.

Twierdzenie 6.8. *Załóżmy, że $(X_{n,k})_{k\leqslant k_n, n\geqslant 1}$ jest układem trójkątnym spełniającym (14) oraz $a_n \in \mathbb{R}$. Wówczas jeśli ciąg zmiennych losowych*

$$S_n := \sum_{k=1}^{k_n} X_{n,k} + a_n$$

zbiega według rozkładu do zmiennej S, to S ma rozkład nieskończenie podzielny.

Uwaga. Na odwrót - jeśli S ma rozkład nieskończenie podzielny μ, to biorąc $a_n = 0$ oraz $X_{k,n} \sim \mu_{1/k_n}$ dostajemy $S_n \sim S$ czyli w szczególności S_n zbiega według rozkładu do S oraz jak łatwo sprawdzić warunek (14) jest spełniony, jeśli $k_n \to \infty$.

Lemat 6.9 (nierówność Paleya-Zygmunda)**.** *Załóżmy, że Z jest nieujemną całkowalną zmienną losową. Wówczas*

$$\forall_{\lambda\in(0,1)} \mathbf{P}(Z \geqslant \lambda \mathbf{E}Z) \geqslant (1-\lambda)^2 \frac{(\mathbf{E}Z)^2}{\mathbf{E}Z^2}.$$

Dowód. Na podstawie nierówności Höldera,

$$\begin{aligned}(\mathbf{E}Z^2)^{1/2}(\mathbf{P}(Z \geqslant \lambda\mathbf{E}Z))^{1/2} &\geqslant \mathbf{E}ZI_{\{Z\geqslant\lambda\mathbf{E}Z\}} = \mathbf{E}Z - \mathbf{E}ZI_{\{Z<\lambda\mathbf{E}Z\}} \\ &\geqslant (1-\lambda)\mathbf{E}Z.\end{aligned}$$

Podnosząc obie strony nierówności do kwadratu dostajemy tezę. □

Wniosek 6.10. *Niech (ε_i) będzie ciągiem Bernoulliego, wówczas dla dowolnych liczb a_i,*

$$\mathbf{P}\left(\Big|\sum_{i=1}^n a_i\varepsilon_i\Big| \geqslant \frac{1}{2}\sqrt{\sum_{i=1}^n a_i^2}\right) \geqslant \frac{3}{16}.$$

Dowód. Niech $S := \sum_{i=1}^n a_i\varepsilon_i$, łatwo sprawdzić, że

$$\mathbf{E}S^4 = \sum_i a_i^4 + 6\sum_{i<j} a_i^2 a_j^2 \leqslant 3(\sum_i a_i^2)^2 = 3(\mathbf{E}S^2)^2,$$

stąd na podstawie nierówności Paleya-Zygmunda,

$$\mathbf{P}\Big(|S| \geqslant \frac{1}{2}\Big(\sum_{i=1}^{n} a_i^2\Big)^{1/2}\Big) = \mathbf{P}(|S|^2 \geqslant \frac{1}{4}\mathbf{E}S^2) \geqslant (1-\frac{1}{4})^2\frac{(\mathbf{E}S^2)^2}{\mathbf{E}S^4} \geqslant \frac{3}{16}.$$

□

Lemat 6.11. *Załóżmy, że X_i są nieujemnymi, niezależnymi zmiennymi losowymi. Wówczas*

$$\mathbf{P}\Big(\sum_i X_i \geqslant \frac{1}{4}\sum \mathbf{E}(X_i \wedge 1)\Big) \geqslant \frac{9}{32}\min\Big(1, \sum \mathbf{E}(X_i \wedge 1)\Big).$$

Dowód. Niech $S := \sum_i X_i \wedge 1$, wówczas

$$\begin{aligned}\mathbf{E}S^2 &= \mathbf{E}\sum_{i,j}(X_i \wedge 1)(X_j \wedge 1) = \sum_i \mathbf{E}(X_i \wedge 1)^2 + 2\sum_{i<j}\mathbf{E}(X_i \wedge 1)\mathbf{E}(X_j \wedge 1) \\ &\leqslant \sum_i \mathbf{E}(X_i \wedge 1) + (\sum_i \mathbf{E}(X_i \wedge 1))^2 = \mathbf{E}S + (\mathbf{E}S)^2.\end{aligned}$$

Stąd na mocy nierówności Paleya-Zygmunda,

$$\mathbf{P}\Big(\sum_i X_i \geqslant \frac{1}{4}\sum \mathbf{E}(X_i \wedge 1)\Big) \geqslant \mathbf{P}(S \geqslant \frac{1}{4}\mathbf{E}S) \geqslant \frac{9}{16}\frac{\mathbf{E}S}{\mathbf{E}S + (\mathbf{E}S)^2} \geqslant \frac{9}{32}\min(1, \mathbf{E}S).$$

□

Wniosek 6.12. *Załóżmy, że X_i są symetrycznymi, niezależnymi zmiennymi losowymi. Wówczas*

$$\mathbf{P}\Big(\Big|\sum_i X_i\Big| \geqslant \frac{1}{4}\Big(\sum \mathbf{E}(X_i^2 \wedge 1)\Big)^{1/2}\Big) \geqslant \frac{1}{20}\min\Big(1, \sum \mathbf{E}(X_i^2 \wedge 1)\Big).$$

Dowód. Niech (ε_i) będzie ciągiem Bernoulliego niezależnym od (X_i), wówczas $\sum_i X_i$ ma ten sam rozkład co $\sum_i \varepsilon_i X_i$. Na mocy Wniosku 6.10 dostajemy $\mathbf{P}(|\sum_i \varepsilon_i x_i| \geqslant (\sum_i x_i^2)^{1/2}/2) \geqslant 3/16)$, zatem przyjmując $a := \sum_i \mathbf{E}(X_i^2 \wedge 1)$ mamy

$$\begin{aligned}\mathbf{P}\Big(\Big|\sum_i X_i\Big| \geqslant \frac{1}{4}\sqrt{a}\Big) &= \mathbf{P}\Big(\Big|\sum_i \varepsilon_i X_i\Big| \geqslant \frac{1}{4}\sqrt{a}\Big) \geqslant \frac{3}{16}\mathbf{P}\Big(\frac{1}{2}(\sum_i X_i^2)^{1/2} \geqslant \frac{1}{4}\sqrt{a}\Big) \\ &= \frac{3}{16}\mathbf{P}\Big(\sum_i X_i^2 \geqslant \frac{1}{4}a\Big) \geqslant \frac{3}{16}\cdot\frac{9}{32}\min(1, a),\end{aligned}$$

gdzie ostatnia nierówność wynika z Lematu 6.11. □

7 Rozkłady stabilne

Definicja 7.1. *Mówimy, że niezdegenerowana zmienna losowa X ma* rozkład stabilny*, jeśli dla każdego $n \geqslant 1$ istnieją liczby $a_n > 0$ oraz $b_n \in \mathbb{R}$ takie, że*

$$X_1 + \ldots + X_n \sim a_n X + b_n,$$

gdzie $X_1, X_2, \ldots$ są niezależnymi kopiami X.

Lemat 7.1. *Załóżmy, że X_n jest ciągiem zmiennych losowych, $a_n, c_n > 0$, $b_n, d_n \in \mathbb{R}$ są takie, że*

$$a_n X_n + b_n \Rightarrow X, \quad c_n X_n + d_n \Rightarrow Y$$

dla pewnych niezdegenerowanych zmiennych losowych X i Y. Wówczas istnieją granice

$$\lim_{n\to\infty} \frac{a_n}{c_n} = A > 0, \quad \lim_{n\to\infty} \Big(d_n - \frac{c_n b_n}{a_n}\Big) = B$$

oraz $Y \sim AX + B$.

Twierdzenie 7.2. *Załóżmy, że $X_1, X_2, \ldots$ są niezależnymi zmiennymi losowymi o jednakowym rozkładzie, $a_n > 0$ oraz $b_n \in \mathbb{R}$ są takie, że*

$$\frac{1}{a_n} \sum_{i=1}^{n} X_i \Rightarrow Y.$$

Wówczas Y jest zdegenerowane lub Y ma rozkład stabilny.

Lemat 7.3. *Załóżmy, że spełnione są założenia Twierdzenia 7.2 oraz Y jest niezdegenerowane. Wówczas $a_n \to \infty$, $\frac{a_{n+1}}{a_n} \to 1$.*

Lemat 7.4. *Załóżmy, że $g\colon (0,\infty) \to [0,\infty)$ niemalejąca oraz*

$$\forall_{x>0} \forall_n \ g(a_n x) = \lambda_n g(x)$$

dla pewnych ciągów liczb dodatnich (a_n) i (λ_n) takich, że $a_n \to \infty$, $\frac{\lambda_{n+1}}{\lambda_n} \to 1$. Wówczas $g(x) \equiv 0$ lub $g(x) = Cx^\rho$ dla pewnych $C > 0$, $\rho \geqslant 0$.

Twierdzenie 7.5. *Załóżmy, że X ma rozkład stabilny. Wówczas X ma rozkład nieskończenie podzielny. Ponadto $a_n = n^{1/\alpha}$ dla pewnego $0 < \alpha \leqslant 2$ oraz*
i) jeśli $\alpha = 2$, to $X \sim \mathcal{N}(a, \sigma^2)$ dla pewnych $a \in \mathbb{R}$ oraz $\sigma > 0$,
ii) jeśli $0 < \alpha < 2$, to $X \sim \pi_{a,0,\nu_{\alpha,C_1,C_2}}$ dla pewnych $a \in \mathbb{R}$, $C_1, C_2 \geqslant 0$, $C_1 + C_2 > 0$, gdzie

$$d\nu_{\alpha,C_1,C_2}(x) = (C_1 x^{-1-\alpha} I_{\{x>0\}} + C_2 (-x)^{-1-\alpha} I_{\{x<0\}})dx.$$

Na odwrót, zmienne $\mathcal{N}(a, \sigma^2)$ są stabilne z $a_n = n^{1/2}$ a zmienne o rozkładzie $\pi_{a,0,\nu_{\alpha,C_1,C_2}}$ są stabilne z $a_n = n^{1/\alpha}$.

Literatura

[1] A. Dembo, O. Zeitouni, *Large deviations techniques and applications*, Springer-Verlag, New York, 1998.

[2] W. Feller, *Wstęp do rachunku prawdopodobieństwa*, tom II, wyd. III, PWN, Warszawa 1981.

[3] J. Jakubowski, R. Sztencel, *Wstęp do teorii prawdopodobieństwa*, wyd. II, Script, Warszawa 2001.

[4] S. Kwapień, W. A. Woyczyński, *Random Series and Stochastic Integrals: Simple and Multiple*, Birkhauser, New York 1992.

[5] M. Ledoux, M. Talagrand, *Probability in Banach spaces*, Springer-Verlag, Berlin, 1991.

[6] V. Petrov, *Limit theorems of probability theory. Sequences of independent random variables.*, Oxford University Press, New York, 1995.

[7] D. W. Stroock, *Probability theory, an analytic view*, Cambridge University Press, Cambridge, 1993.