Wstęp do rachunku prawdopodobieństwa
Wykład 11: Nierówności probabilistyczne

Twierdzenie 1 (Nierówność Markowa)**.** *Niech X będzie zmienną losową, dla której $\mathbb{E}X < \infty$. Wówczas dla dowolnej liczby rzeczywistej $a > 0$ zachodzi:*

$$\mathbb{P}\left(|X| \geqslant a\right) \leqslant \frac{\mathbb{E}|X|}{a}.$$

Dowód. Niech $\mathbf{1}_{\{|X|\geqslant a\}}$ oznacza zmienną losową określoną na tej samej przestrzeni probabilistycznej co zmienna losowa X i taką, że

$$\mathbf{1}_{\{|X|\geqslant a\}} = \begin{cases} 0, & \text{gdy } |X| < a, \\ 1, & \text{gdy } |X| \geqslant a. \end{cases}$$

Wówczas zachodzi następujący ciąg nierówności:

$$|X| \geqslant |X| \cdot \mathbf{1}_{\{|X|\geqslant a\}} \geqslant a \cdot \mathbf{1}_{\{|X|\geqslant a\}}.$$

Rzeczywiście, gdy $|X| < a$ mamy:

$$|X| \geqslant |X| \cdot 0 = |X| \cdot \mathbf{1}_{\{|X|\geqslant a\}} = a \cdot \mathbf{1}_{\{|X|\geqslant a\}},$$

natomiast dla $|X| \geqslant a$ mamy:

$$|X| = |X| \cdot 1 = |X| \cdot \mathbf{1}_{\{|X|\geqslant a\}} \geqslant a \cdot \mathbf{1}_{\{|X|\geqslant a\}}.$$

Możemy teraz oszacować $\mathbb{E}|X|$ następująco:

$$\mathbb{E}|X| \geqslant \mathbb{E}(|X| \cdot \mathbf{1}_{\{|X|\geqslant a\}}) \geqslant \mathbb{E}(a \cdot \mathbf{1}_{\{|X|\geqslant a\}}) = a\mathbb{E}\,\mathbf{1}_{\{|X|\geqslant a\}} = a \cdot \mathbb{P}\left(|X| \geqslant a\right),$$

skąd, po podzieleniu stronami, otrzymujemy nierówność Markowa. □

Uwaga 1. W powyższym dowodzie przyjęliśmy założenie, że jeśli X i Y są zmiennymi losowymi zdefiniowanymi na tej samej przestrzeni probabilistycznej, dla których istnieje $\mathbb{E}X$ i $\mathbb{E}Y$, to jeśli $X \geqslant Y$, co należy rozumieć jako $X(\omega) \geqslant Y(\omega)$ dla każdego zdarzenia elementarnego ω, wówczas $\mathbb{E}X \geqslant \mathbb{E}Y$. Dowód tego faktu zostawiamy jako ćwiczenie.

Twierdzenie 2 (Nierówność Czebyszewa–Bienaymé)**.** *Niech X będzie zmienną losową, dla której $\mathbb{E}X < \infty$ oraz $VarX < \infty$. Wówczas dla dowolnej liczby rzeczywistej $b > 0$ zachodzi:*

$$\mathbb{P}\left(|X - \mathbb{E}X| \geqslant b\right) \leqslant \frac{VarX}{b^2}.$$

Dowód. Zastosujmy nierówność Markowa do zmiennej losowej $|X - \mathbb{E}X|^2$ oraz $a = b^2$. Wówczas mamy:

$$\mathbb{P}\left(|X - \mathbb{E}X| \geqslant b\right) = \mathbb{P}\left(|X - \mathbb{E}X|^2 \geqslant b^2\right) \leqslant \frac{\mathbb{E}(|X - \mathbb{E}X|^2)}{b^2} = \frac{\mathbb{E}((X - \mathbb{E}X)^2)}{b^2} = \frac{\mathrm{Var}X}{b^2}.$$

□

Obie powyższe nierówności są bardzo ogólne, a co za tym idzie w niektórych przypadkach mogą nie dawać sensownych oszacowań. W praktyce nierówność Czebyszewa–Bienaymé częściej okazuje się być dokładniejsza od nierówności Markowa, natomiast wymaga znajomości wariancji rozważanej zmiennej losowej. Jeśli natomiast wariancja interesującej nas zmiennej losowej nie jest znana, ale znamy jej wartość oczekiwaną, to oczywiście możemy zastosować tylko nierówność Markowa.

Przykład 1. Rzucamy 10000 razy symetryczną monetą. Chcielibyśmy oszacować z góry prawdopodobieństwo, że liczba wyrzuconych orłów wynosi co najmniej 6000.

Zauważmy, że zmienna losowa X równa liczbie wyrzuconych orłów w trakcie 10000 rzutów monetą ma rozkład dwumianowy z parametrami 10000 i 1/2, tzn. $X \sim Bin(10000, 1/2)$. A zatem $\mathbb{E}X = 10000 \cdot 1/2 = 5000$ oraz $\mathrm{Var}X = 10000 \cdot 1/2 \cdot 1/2 = 2500$. Ponadto zmienna losowa X przyjmuje tylko wartości nieujemne, zatem $X = |X|$.

Oszacujmy na początek szukane prawdopodobieństwo wykorzystując nierówność Markowa. Mamy:

$$\mathbb{P}\left(X \geqslant 6000\right) \leqslant \frac{\mathbb{E}X}{6000} = \frac{5}{6}.$$

Nie jest to specjalnie dobre oszacowanie, bo wydaje się, że szanse na zajście tego zdarzenia powinny być dużo niższe. Spróbujmy zatem oszacować to prawdopodobieństwo jeszcze raz, tym razem w oparciu o nierówność Czebyszewa–Bienaymé:

$$\mathbb{P}\left(X \geqslant 6000\right) = \mathbb{P}\left(X - \mathbb{E}X \geqslant 1000\right) \leqslant \mathbb{P}\left(|X - \mathbb{E}X| \geqslant 1000\right) \leqslant \frac{\mathrm{Var}X}{1000^2} = \frac{2500}{1000000} = \frac{1}{400}.$$

Jak widać to oszacowanie jest o wiele mocniejsze! Ile natomiast wynosi to prawdopodobieństwo? Pamiętając, że zmienna losowa X ma rozkład dwumianowy $X \sim Bin(10000, 1/2)$ otrzymujemy

$$\mathbb{P}\left(X \geqslant 6000\right) = \sum_{k=6000}^{10000} \binom{10000}{k} \left(\frac{1}{2}\right)^{10000} \approx 8{,}702158 \cdot 10^{-90}.$$

Jak nietrudno się domyślić, do obliczenia powyższego prawdopodobieństwa musimy posłużyć się komputerem. Jedną z możliwości jest użycie programu R, gdzie powyższe prawdopodobieństwo policzymy za pomocą komendy

sum(dbinom(6000:10000, 10000, 0.5)).

Tabelka najczęściej spotykanych rozkładów zmiennych losowych **dyskretnych**.

Rozkład	Parametry	$\mathbb{P}\left(X = k\right)$	dla $k =$	$\mathbb{E}(X)$	Var(X)	Uwagi
dwumianowy	n, p	$\binom{n}{k} p^k (1-p)^{n-k}$	$0, 1, \ldots, n$	np	$np(1-p)$	Liczba sukcesów w n próbach Bernoulliego.
Poissona	λ	$\frac{\lambda^k}{k!} e^{-\lambda}$	$0, 1, 2, \ldots$	λ	λ	Rozkład zdarzeń „rzadkich”.
geometryczny	p	$(1-p)^{k-1} p$	$1, 2, \ldots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	Liczba doświadczeń do pierwszego sukcesu.
Pascala	p, r	$\binom{k-1}{r-1} p^r (1-p)^{k-r}$	$r, r+1, \ldots$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	Liczba doświadczeń do r-tego sukcesu.
hipergeometryczny	N, m, n	$\frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}$	$0, 1, 2, \ldots, n$	$\frac{nm}{N}$	$\frac{nm(N-n)(N-m)}{N^2(N-1)}$	Liczba wylosowanych kul typu A, jeśli losujemy jednocześnie n kul z urny, w której jest N kul, w tym m kul typu A.

Uwaga 2. Jeśli X ma rozkład dwumianowy z parametrami n i p, to $\mathbb{P}\left(X = k\right)$ najpierw rośnie, przyjmuje największą wartość dla $k = \lfloor (n+1)p \rfloor$ (w przypadku, gdy $(n+1)p$ jest całkowite przyjmuje największą wartość również dla $k = \lfloor (n+1)p \rfloor - 1$), a następnie maleje.

Przykład 2. Niech X będzie zmienną losową równą liczbie szóstek wyrzuconych w 100 rzutach symetryczną kostką. Wówczas X ma rozkład dwumianowy $Bin(100, 1/6)$. Jeśli chcielibyśmy ustalić najbardziej prawdopodobną wartość zmiennej losowej X, to będzie ona równa $\lfloor(100+1)\cdot 1/6\rfloor = 16$. Z kolei najmniej prawdopodobne są skrajne wyniki, czyli w tym przypadku $X = 0$ oraz $X = 100$. Prawdopodobieństwa obydwu tych zdarzeń wynoszą odpowiednio $(5/6)^{100}$ oraz $(1/6)^{100}$, a zatem najmniej prawdopodobną wartością zmiennej losowej X jest 100.

Tabelka najczęściej spotykanych rozkładów zmiennych losowych **ciągłych**.

Rozkład	Parametry	Zbiór wartości	Gęstość	$\mathbb{E}(X)$	Var(X)
jednostajny	$a < b$	$[a, b]$	$g(x) = \frac{1}{b-a}$ dla $x \in [a, b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
wykładniczy	λ	$(0, \infty)$	$g(x) = \lambda e^{-\lambda x}$ dla $x > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
normalny $\mathcal{N}(\mu, \sigma^2)$	μ, σ^2	$\mathbb{R}$	$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}$	μ	σ^2
standardowy normalny $\mathcal{N}(0, 1)$	$0, 1$	$\mathbb{R}$	$g(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$	0	1

Uwaga 3. Parametry μ i σ^2 w rozkładzie normalnym $\mathcal{N}(\mu, \sigma^2)$ to odpowiednio wartość oczekiwana i wariancja zmiennej losowej o tym rozkładzie. Jeżeli X jest zmienną losową o rozkładzie normalnym z parametrami μ i σ^2, tzn. $X \sim \mathcal{N}(\mu, \sigma^2)$, to zmienna losowa

$$\frac{X - \mu}{\sigma}$$

ma również rozkład normalny, ale z parametrami 0 i 1, tzn. $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$. Powyższe przekształcenie nazywamy **standaryzacją** zmiennej losowej X.

Szkic dowodu. Z liniowości wartości oczekiwanej otrzymujemy:

$$\mathbb{E}\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma}\left(\mathbb{E}X - \mu\right) = \frac{1}{\sigma}(\mu - \mu) = 0,$$

natomiast z własności wariancji dostajemy

$$\mathrm{Var}\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma^2} \cdot \mathrm{Var}(X - \mu) = \frac{1}{\sigma^2} \cdot \mathrm{Var}X = \frac{1}{\sigma^2} \cdot \sigma^2 = 1.$$

A zatem zmienna losowa $\frac{X-\mu}{\sigma}$ ma wartość oczekiwaną równą 0 oraz wariancję równą 1. Aby pokazać, że ma ona rozkład normalny, należałoby obliczyć jeszcze jej gęstość i porównać z gęstością w tabelce, ale ponieważ wymaga to zbyt skomplikowanych rachunków pominiemy ten krok. □

Uwaga 4. Rozważając zmienne losowe o rozkładzie normalnym $\mathcal{N}(0, 1)$ w praktyce nie będziemy posługiwać się gęstością tego rozkładu, tylko będziemy odczytywać przybliżone wartości dystrybunaty tego rozkładu z gotowych tablic. Z kolei rozważając zmienną losową o rozkładzie normalnym $\mathcal{N}(\mu, \sigma)$ możemy dokonać najpierw jej standaryzacji, żeby sprowadzić ją do postaci zmiennej losowej o rozkładzie normalnym $\mathcal{N}(0, 1)$ i wtedy odczytać jej rozkład prawdopodobieństwa z tablic.

Definicja 1. *Jeśli zmienna losowa X ma wariancję $VarX$, to pierwiastek z wariancji nazywamy* **odchyleniem standardowym** *zmiennej losowej X i oznaczamy symbolem $\sigma_X = \sqrt{VarX}$.*

Przykład 3 (Reguła trzech sigm)**.** Jeśli zmienna losowa X ma wariancję $\mathrm{Var}X = \sigma^2$, to wówczas (korzystając z nierówności Czebyszewa–Bienaymé) prawdopodobieństwo, że zmienna ta odchyla się od swojej wartości oczekiwanej o więcej niż trzy odchylenia standardowe σ można oszacować z góry przez:

$$\mathbb{P}\left(|X - \mathbb{E}X| \geqslant 3\sigma\right) \leqslant \frac{\mathrm{Var}X}{9\sigma^2} = \frac{1}{9}.$$

W przypadku standardowego rozkładu normalnego $\mathcal{N}(0, 1)$, to prawdopodobieństwo nie przekracza 0,003. Oznacza to, że zmienna losowa o tym rozkładzie jest stosunkowo dobrze skoncentrowana wokół swojej wartości oczekiwanej, czyli wokół zera.

Przykład 4. Średnia liczba pożarów, które wybuchają w ciągu roku w Biebrzańskim Parku Narodowym równa jest 5. Ile wynosi prawdopodobieństwo, że po pierwszym wiosennym pożarze w tym roku więcej pożarów nie nawiedzi parku?

Ponieważ mamy tutaj do czynienia ze zdarzeniem rzadkim, możemy przyjąć, że liczba pożarów X ma rozkład Poissona ze średnią 5, czyli rozważamy zmienną losową X o rozkładzie $Po(5)$. Pytamy się zatem, ile wynosi prawdopodobieństwo, że $X = 1$. Z jednej strony możemy policzyć to prawdopodobieństwo wprost:

$$\mathbb{P}\left(X = 1\right) = 5e^{-5} \approx 0{,}034.$$

Możemy też oszacować z góry to prawdopodobieństwo z nierówności Czebyszewa-Bienaymé pamiętając, że dla zmiennej losowej $X \sim Po(5)$ mamy $\mathbb{E}X = \mathrm{Var}X = 5$. Zatem:

$$\mathbb{P}\left(X = 1\right) \leqslant \mathbb{P}\left(X \leqslant 1\right) = \mathbb{P}\left(X - \mathbb{E}X \leqslant -4\right) \leqslant \mathbb{P}\left(|X - \mathbb{E}X| \geqslant 4\right) \leqslant \frac{\mathrm{Var}X}{4^2} = \frac{5}{16} \approx 0{,}31.$$

Jak widzimy, w tym przypadku oszacowanie z nierówności Czebyszewa-Bienaymé jest niemalże 10-krotnie gorsze od właściwego wyniku.

Gdybyśmy natomiast chcieli oszacować prawdopodobieństwo, że w tym roku wybuchnie co najmniej 10 pożarów, wówczas aby policzyć to wprost musielibyśmy obliczyć:

$$\mathbb{P}\left(X \geqslant 10\right) = \sum_{k=10}^{\infty} \frac{5^k}{k!} e^{-5} = 1 - \sum_{k=0}^{9} \frac{5^k}{k!} e^{-5} \approx 0{,}03.$$

Natomiast z nierówności Czebyszewa-Bienaymé otrzymujemy oszacowanie:

$$\mathbb{P}\left(X \geqslant 10\right) = \mathbb{P}\left(X - \mathbb{E}X \geqslant 5\right) \leqslant \mathbb{P}\left(|X - \mathbb{E}X| \geqslant 5\right) \leqslant \frac{\mathrm{Var}X}{5^2} = 0{,}2.$$

Powyższe oszacowanie jest ponownie o wiele słabsze od właściwego wyniku, natomiast obliczenia są o tyle prostsze, że jesteśmy w stanie wykonać je w pamięci bez użycia kalkulatora.