

Rozdział I

Repetytorium z prawdopodobieństwa

1 Zmienne losowe i ich charakterystyki

1.1 Rozkłady zmiennych losowych

W teorii prawdopodobieństwa wprowadza się *zmienną losową* X jako funkcję zdefiniowaną na przestrzeni zdarzeń elementarnych Ω . W wielu przypadkach, ani Ω ani X nie są znane. Ale co jest ważne, znany jest *rozkład* zmiennej losowej X , tj. przepis jak liczyć $F(B)$, że X przyjmie wartość z zbioru B , gdzie B jest podzbiorem (borelowskim) prostej $\mathbb{R}$.

Ze względu na matematyczną wygodę wyróżnia się rodzinę $\mathcal{F}$ podzbiorów Ω , zwaną *rodziną zdarzeń*, oraz *prawdopodobieństwo* $\mathbb{P}$ na $\mathcal{F}$ przypisujące zdarzeniu A z $\mathcal{F}$ liczbę $\mathbb{P}(A)$ z odcinka $[0, 1]$.

Rozkład można opisać przez podanie *dystrybuanty*, t.j. funkcji $F : \mathbb{R} \rightarrow [0, 1]$ takiej, że $F(x) = \mathbb{P}(X \leq x)$. Przez $\bar{F}(x) = 1 - F(x)$ będziemy oznaczać *ogon* dystrybuanty F .

W praktyce mamy dwa typy zmiennych losowych: dyskretne i absolutnie ciągłe. Mówimy, że X jest *dyskretną* jeśli istnieje przeliczalny zbiór liczb $E = \{x_0, x_1, \dots\}$ z $\mathbb{R}$ taki, że $\mathbb{P}(X \in E) = 1$. W tym przypadku definiujemy *funkcję prawdopodobieństwa* $p : E \rightarrow [0, 1]$ przez $p(x_k) = \mathbb{P}(X = x_k)$; para (E, p) podaje pełny przepis jak liczyć prawdopodobieństwa $\mathbb{P}(X \in B)$ a więc zadaje rozkład. Najważniejszą podklasą są zmienne losowe przyjmujące wartości w zbiorze E który jest podzbiorem kraty $\{hk, k \in \mathbb{Z}_+\}$ dla pewnego $h > 0$. Jeśli nie będzie powiedziane inaczej to zakładamy, że $h = 1$, tzn $E \subset \mathbb{Z}_+$, i.e. $x_k = k$. Wtedy piszemy prosto $p(x_k) = p_k$ i mówimy, że

rozkład jest *kratowy*.

Z drugiej strony mówimy, że zmienna losowa X jest *absolutnie ciągła* jeśli istnieje funkcja $f : \mathbb{R} \rightarrow \mathbb{R}_+$ taka, że $\int_{-\infty}^{\infty} f(x) dx = 1$ i $\mathbb{P}(X \in B) = \int_B f(x) dx$ dla każdego $B \subset \mathbb{R}$. Mówimy wtedy, że $f(x)$ jest *gęstością* zmiennej X .

Rozkład dyskretnej zmiennej losowej nazywamy rozkładem dyskretnym. Jeśli zmienna X jest kratowa, to jej rozkład też nazywamy kratowy. Analogicznie rozkład zmiennej losowej absolutnie ciągłej nazywamy rozkładem absolutnie ciągłym. Czasami dystrybuanta nie jest ani dyskretna ani absolutnie ciągła. Przykładem jest *mieszanka* $F = \theta F_1 + (1 - \theta) F_2$, gdzie F_1 jest dyskretna z prawdopodobieństwem θ i F_2 jest absolutnie ciągła z prawdopodobieństwem $1 - \theta$.

Aby zaznaczyć, że dystrybuanta, gęstość, funkcja prawdopodobieństwa, jest dla zmiennej losowej X , piszemy czasem $F_X, p_X, f_X, \dots$

Dla dwóch zmiennych losowych X i Y o tym samym rozkładzie piszemy $X \stackrel{d}{=} Y$. To samo stosujemy do wektorów losowych.

Mówimy, że wektor losowy $\mathbf{X} = (X_1, \dots, X_n)^T$ jest *absolutnie ciągły* jeśli istnieje nieujemna i całkowalna funkcja $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$ z

$$\int_{\mathbb{R}} \dots \int_{\mathbb{R}} f(x_1, \dots, x_n) dx_1 \dots dx_n = 1$$

taka, że $\mathbb{P}(X_1 \in B_1, \dots, X_n \in B) = \int_{B_1} \dots \int_{B_n} f(x_1, \dots, x_n) dx_n \dots dx_1$ dla $B_1, \dots, B_n \subset \mathbb{R}$. Funkcja $f(\mathbf{x})$ nazywa się *gęstością* $\mathbf{X} = (X_1, \dots, X_n)^T$. Przytoczymy teraz pożyteczny rezultat dotyczący wzoru na gęstość przekształconego wektora losowego $\mathbf{X}$, takiego, że $\mathbb{P}(\mathbf{X} \in a) = 1$. Mianowicie niech będą dane funkcje $g_i : \mathbb{R} \rightarrow \mathbb{R}$, $i = 1, \dots, n$, które spełniają następujące warunki:

- przekształcenie $\mathbf{g} = (g_1, \dots, g_n)^T$ jest wzajemnie jednoznaczne na zbiorze A ,
- przekształcenie $\mathbf{g}$ ma na A ciągłe pierwsze pochodne cząstkowe ze względu na wszystkie zmienne,
- jacobian $J\mathbf{g}(\mathbf{x}) = \frac{\partial(g_1, \dots, g_n)}{\partial(x_1, \dots, x_n)}$ jest różny od zera na A .

Oznaczmy przez $\mathbf{h} = (h_1, \dots, h_n)^T$ przekształcenie odwrotne do $\mathbf{g}$, tzn $\mathbf{x} = \mathbf{h}(\mathbf{y}) = (h_1(\mathbf{y}), \dots, h_n(\mathbf{y}))$ dla każdego $\mathbf{y} \in \mathbf{g}(A)$. Rozważmy wektor $\mathbf{Y} = \mathbf{g}(\mathbf{X})$. Wektor ten ma gęstość $f_{\mathbf{Y}}$ zadaną wzorem:

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(h_1(\mathbf{y}), \dots, h_n(\mathbf{y})) |J_{\mathbf{h}}(\mathbf{y})|.$$

Jeśli wektor losowy $(X_1, \dots, X_n)$ jest absolutnie ciągły, to $(X_1, \dots, X_{n-1})$ jest też odpowiednio z gęstościami $f_{X_1, \dots, X_n}$ i $f_{X_1, \dots, X_{n-1}}$ i jeśli $f_{X_1, \dots, X_{n-1}}(x_1, \dots, x_{n-1}) > 0$, to

$$f_{X_n | X_1, \dots, X_{n-1}}(x_n | x_1, \dots, x_{n-1}) = \frac{f_{X_1, \dots, X_n}(x_1, \dots, x_n)}{f_{X_1, \dots, X_{n-1}}(x_1, \dots, x_{n-1})}$$

nazywa się i *warunkową gęstością* X_n pod warunkiem $X_1 = x_1, \dots, X_{n-1} = x_{n-1}$.

Dla ustalonego n i wszystkich $k = 1, 2, \dots, n$ i $\omega \in \Omega$, niech $X_{(k)}(\omega)$ oznacza k -th najmniejszą wartość $X_1(\omega), \dots, X_n(\omega)$. Składowe wektora losowego $(X_{(1)}, \dots, X_{(n)})$ nazywają się *statystyką porządkową* $(X_1, \dots, X_n)$.

1.2 Podstawowe charakterystryki

Niech X będzie zmienną losową i $g : \mathbb{R} \rightarrow \mathbb{R}$ funkcją. Definiujemy $\mathbb{E}g(X)$ przez

$$\mathbb{E}g(X) = \begin{cases} \sum_k g(x_k)p(x_k) & \text{jeśli } X \text{ jest dyskretny} \\ \int_{-\infty}^{\infty} g(x)f(x)dx & \text{jeśli } X \text{ jest absolutnie ciągły} \end{cases}$$

pod warunkiem odpowiednio, że $\sum_k |g(x_k)|p(x_k) < \infty$ i $\int_{-\infty}^{\infty} |g(x)|f(x)dx < \infty$. Jeśli g jest nieujemna, to używamy symbolu $\mathbb{E}g(X)$ również gdy równa się nieskończoność. Jeśli rozkład jest mieszanką dwóch typów to definiujemy $\mathbb{E}g(X)$ przez

$$\mathbb{E}g(X) = \theta \sum_k g(x_k)p(x_k) + (1 - \theta) \int_{-\infty}^{\infty} g(x)f(x)dx. \quad (1.1)$$

Czasami wygodne jest zapisanie $\mathbb{E}g(X)$ w terminach *całki Stieltjesa* t.j.

$$\mathbb{E}g(X) = \int_{-\infty}^{\infty} g(x)dF(x),$$

która jest dobrze określona przy pewnych warunkach na g i F . W szczególności dla $g(x) = x$ wartość $\mu = \mathbb{E}X$ jest nazywana *średnią* lub *wartością oczekiwaną* lub *pierwszym momentem* X . Dla $g(x) = x^n$, $\mu^{(n)} = \mathbb{E}(X^n)$ jest nazywana n -tym *momentem*. *Wariancją* zmiennej losowej X jest $\sigma^2 = \mathbb{E}(X - \mu)^2$ a $\sigma = \sqrt{\sigma^2}$ jest *odchyleniem standardowym* lub *dyspersją*. Czasami piszemy $\text{Var } X$ zamiast σ^2 . *Współczynnik zmienności* jest zdefiniowany przez $\text{cv}_X = \sigma/\mu$, i *współczynnik asymetrii* przez $\mathbb{E}(X - \mu)^3 \sigma^{-3}$. Dla dwóch zmiennych losowych X, Y definiujemy we *kowariancję* $\text{Cov}(X, Y)$

przez $\text{Cov}(X, Y) = \mathbb{E}((X - \mathbb{E} X)(Y - \mathbb{E} Y))$ jeśli tylko $\mathbb{E} X^2, \mathbb{E} Y^2 < \infty$. Zauważmy, że równoważnie mamy $\text{Cov}(X, Y) = \mathbb{E} XY - \mathbb{E} X \mathbb{E} Y$. Jeśli $\text{Cov}(X, Y) > 0$, to mówimy, że X, Y are *dodatnio skorelowane*, jeśli $\text{Cov}(X, Y) < 0$ to *ujemnie skorelowane* i *nieskorelowane* jeśli $\text{Cov}(X, Y) = 0$. Przez *współczynnik korelacji* pomiędzy X, Y rozumiemy

$$\text{corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

Medianą zmiennej losowej X jest dowolna liczba $\zeta_{1/2}$ taka, że

$$\mathbb{P}(X \leq \zeta_{1/2}) \geq \frac{1}{2}, \quad \mathbb{P}(X \geq \zeta_{1/2}) \geq \frac{1}{2}.$$

Przez *indykator* $1(A) : \Omega \rightarrow \mathbb{R}$ rozumiemy zmienną losową

$$1(A, \omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{w przeciwnym razie.} \end{cases}$$

W szczególności będziemy pisać

$$1(X \in B) = l \leq \begin{cases} 1 & \text{if } X \in B, \\ 0 & \text{w przeciwnym razie.} \end{cases}$$

A więc jeśli $A \in \mathcal{F}$ to $1(A)$ jest zmienną losową i $\mathbb{P}(A) = \mathbb{E} 1(A)$. Używa się następującej konwencji, że $\mathbb{E}[X; A] = \mathbb{E}(X 1(A))$ dla zmiennej losowej X i zdarzenia $A \in \mathcal{F}$.

Jeśli chcemy zwrócić uwagę, że średnia, n -ty moment, itd. są zdefiniowane dla zmiennej X lub dystrybucyj F to piszemy $\mu_X, \mu_X^{(n)}, \sigma_X^2, \dots$ lub $\mu_F, \mu_F^{(n)}, \sigma_F^2, \dots$.

1.3 Niezależność i warunkowanie

Mówimy, że zmienne losowe $X_1, \dots, X_n$ są *niezależne* jeśli

$$\mathbb{P}(X_1 \in B_1, \dots, X_n \in B_n) = \prod_{k=1}^n \mathbb{P}(X_k \in B_k)$$

dla wszystkich $B_1, \dots, B_n$ lub, równoważnie,

$$\mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n) = \prod_{k=1}^n \mathbb{P}(X_k \leq x_k)$$

dla wszystkich $x_1, \dots, x_n \in \mathbb{R}$. Nieskończony ciąg $X_1, X_2, \dots$ jest ciągiem niezależnych zmiennych losowych, jeśli każdy skończony podciąg ma tę własność. Mówimy, że dwa ciągi $X_1, X_2, \dots$ i $Y_1, Y_2, \dots$ są *niezależne* jeśli

$$\mathbb{P}\left(\bigcap_{i=1}^n \{X_i \leq x_i\} \cap \bigcap_{i=1}^m \{Y_i \leq y_i\}\right) = \mathbb{P}\left(\bigcap_{i=1}^n \{X_i \leq x_i\}\right) \mathbb{P}\left(\bigcap_{i=1}^m \{Y_i \leq y_i\}\right)$$

dla wszystkich $n, m \in \mathbb{N}$ i $x_1, \dots, x_n, y_1, \dots, y_m \in \mathbb{R}$. Dla ciągu $X_1, X_2, \dots$ niezależnych i jednakowo rozłożonych zmiennych losowych wygodne jest wprowadzenie generycznej zmiennej losowej X , która ma taki sam rozkład. Mówimy, że zmienna losowa jest niezależna od zdarzenia, jeśli jest niezależna od indykatora $1(A)$ tego zdarzenia.

Niech $A \in \mathcal{F}$. *Warunkowym prawdopodobieństwem* zdarzenia A' pod warunkiem A jest

$$\mathbb{P}(A' | A) = \begin{cases} \mathbb{P}(A' \cap A) / \mathbb{P}(A) & \text{jeśli } \mathbb{P}(A) > 0, \\ 0 & \text{w przeciwnym razie.} \end{cases}$$

Często użyteczne jest wzór na *prawdopodobieństwo całkowite*. Niech $A, A_1, A_2, \dots \in \mathcal{F}$ będzie ciągiem zdarzeń taki, że $A_i \cap A_j = \emptyset$ dla $i \neq j$ i $\sum_{i=1}^{\infty} \mathbb{P}(A_i) = 1$. Wtedy,

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(A | A_i) \mathbb{P}(A_i). \quad (1.2)$$

Warunkową wartość oczekiwaną $\mathbb{E}(X | A)$ zmiennej losowej X pod warunkiem zdarzenia A jest warunkową wartością oczekiwaną względem rozkładu warunkowego z dystrybunatą $F_{X|A}$, gdzie $F_{X|A}(x) = \mathbb{P}(\{X \leq x\} \cap A) / \mathbb{P}(A)$.

1.4 Transformaty

Niech $I = \{s \in \mathbb{R} : \mathbb{E}e^{sX} < \infty\}$. Zauważmy, że I musi być odcinkiem (właściwym), prostą, półprostą lub nawet punktem $\{0\}$. *Funkcję tworzącą momenty* $\hat{m} : I \rightarrow \mathbb{R}$ zmiennej losowej X jest zdefiniowana przez $\hat{m}(s) = \mathbb{E}e^{sX}$. Zauważmy różnicę pomiędzy funkcją tworzącą momenty a *trasformatą Laplace-Stieltjesa* $\hat{l}(s) = \mathbb{E}e^{-sX} = \int_{-\infty}^{\infty} e^{-sx} dF(x)$ zmiennej losowej X lub dystrybucji F zmiennej X .

Dla zmiennych kratowych z funkcją prawdopodobieństwa $\{p_k, k \in \mathbb{N}\}$ wprowadza się *funkcję tworzącą* $\hat{g} : [-1, 1] \rightarrow \mathbb{R}$ zdefiniowaną przez $\hat{g}(s) = \sum_{k=0}^{\infty} p_k s^k$.

Niech teraz X będzie dowolną zmienną losową z dystrybunatą F . *Funkcją charakterystyczną* $\hat{\varphi} : \mathbb{R} \rightarrow \mathbb{C}$ zmiennej X jest zdefiniowana przez

$$\hat{\varphi}(s) = \mathbb{E}e^{isX}. \quad (1.3)$$

W szczególności jeśli F jest absolutnie ciągła z gęstością f , to

$$\hat{\varphi}(s) = \int_{-\infty}^{\infty} e^{isx} f(x) dx = \int_{-\infty}^{\infty} \cos(sx) f(x) dx + i \int_{-\infty}^{\infty} \sin(sx) f(x) dx.$$

Jeśli X jest kratową zmienną losową z funkcją prawdopodobieństwa $\{p_k\}$, to

$$\hat{\varphi}(s) = \sum_{k=0}^{\infty} e^{isk} p_k. \quad (1.4)$$

Jeśli chcemy podkreślić, że transformata jest zmiennej X z dystrybuantą F , to piszemy $\hat{m}_X, \dots, \hat{\varphi}_X$ lub $\hat{m}_F, \dots, \hat{\varphi}_F$.

Zauważmy, też formalne związki $\hat{g}(e^s) = \hat{l}(-s) = \hat{m}(s)$ lub $\hat{\varphi}(s) = \hat{g}(e^{is})$. Jednakże delikatnym problemem jest zdecydowanie, gdzie dana transformata jest określona. Na przykład jeśli X jest nieujemna, to $\hat{m}(s)$ jest zawsze dobrze zdefiniowana przynajmniej dla $s \leq 0$ podczas gdy $\hat{l}(s)$ jest dobrze zdefiniowana przynajmniej dla $s \geq 0$. Są też przykłady pokazujące, że $\hat{m}(s)$ może być ∞ dla wszystkich $s > 0$, podczas gdy w innych przypadkach, że $\hat{m}(s)$ jest skończone $(-\infty, a)$ dla pewnych $a > 0$.

Ogólnie jest wzajemna jednoznaczność pomiędzy rozkładem a odpowiadającą jemu transformatą, jednakże w każdym wypadku sformułowanie twierdzenia wymaga odpowiedniego wysłowienia. Jeśli n -ty moment zmiennej losowej $|X|$ jest skończony, to n -ta pochodna $\hat{m}^{(n)}(0), \hat{l}^{(n)}(0)$ istnieje jeśli tylko $\hat{m}(s), \hat{l}(s)$ są dobrze zdefiniowane w pewnym otoczeniu $s = 0$. Wtedy mamy

$$\mathbb{E} X^n = \hat{m}^{(n)}(0) = (-1)^n \hat{l}^{(n)}(0) = (-i)^n \hat{\varphi}^{(n)}(0). \quad (1.5)$$

Jeśli $\hat{m}(s)$ i $\hat{l}(s)$ są odpowiednio dobrze zdefiniowane na $(-\infty, 0]$ i $[0, \infty)$, to trzeba pochodne w (1.5) zastąpić pochodnymi jednostronnymi, t.j.

$$\mathbb{E} X^n = \hat{m}^{(n)}(0-) = (-1)^n \hat{l}^{(n)}(0+), \quad (1.6)$$

Jeśli X przyjmuje wartość w podzbiorze $\mathbb{Z}_+$ i jeśli $\mathbb{E} X^n < \infty$, to

$$\mathbb{E} (X(X-1) \dots (X-n+1)) = \hat{g}^{(n)}(1-). \quad (1.7)$$

Inną ważną własnością funkcji tworzącej momenty $\hat{m}(s)$ (i również dla transformaty Laplace'a $\hat{l}(s)$ i funkcji charakterystycznej $\hat{\varphi}(s)$) jest własność, że jeśli $X_1, \dots, X_n$ są niezależne to

$$\hat{m}_{X_1+\dots+X_n}(s) = \prod_{k=1}^n \hat{m}_{X_k}(s). \quad (1.8)$$

Podobnie, jeśli $X_1, \dots, X_n$ są niezależnymi krataowymi zmiennymi losowymi z $\mathbb{Z}_+$ to

$$\hat{g}_{X_1+\dots+X_n}(s) = \prod_{k=1}^n \hat{g}_{X_k}(s). \quad (1.9)$$

Ogólnie dla dowolnych funkcji $g_1, \dots, g_n : \mathbb{R} \rightarrow \mathbb{R}$ mamy

$$\mathbb{E} \left(\prod_{k=1}^n g_k(X_k) \right) = \prod_{k=1}^n \mathbb{E} g_k(X_k) \quad (1.10)$$

jeśli tylko $X_1, \dots, X_n$ są niezależne.

Niech $X, X_1, X_2, \dots$ będą zmiennymi losowymi. Mówimy, że ciąg $\{X_n\}$ *zbiega słabo* (lub *zbiega według rozkładu*) do X jeśli $F_{X_n}(x) \rightarrow F_X(x)$ dla wszystkich punktów ciągłości F_X . Piszemy wtedy $X_n \xrightarrow{d} X$.

1.5 Operacje na rozkładach

Mieszanki Dla ciągu dystrybuant $F_1, F_2, \dots$ i funkcji prawdopodobieństwa $\{p_k, n = 1, 2, \dots\}$ definiujemy dystrybuantę $F = \sum_{k=1}^{\infty} p_k F_k$ nazywaną *mieszką*. Dla rodziny rozkładów F_θ zparametryzowanej przez θ , gdzie θ przebiega po zbiorze Θ z $\mathbb{R}$ względem dystrybuanty G skoncentrowanej na Θ . *Mieszką* $F(x)$ rodziny $\{F_\theta, \theta \in \Theta\}$ względem *dystrybuanty mieszającej* G nazywamy $F(x) = \int_{\Theta} F_\theta(x) dG(\theta)$, $x \in \mathbb{R}$.

Przesunięcie Przez *przesunięcie* dystrybuanty $F(x)$ o $a \in \mathbb{R}$, otrzymujemy nową dystrybuantę $G(x) = F(x - a)$. To oznacza, że jeśli X ma dystrybuantę F , to $X + a$ ma dystrybuantę G .

Skalowanie Przez *skalowanie* argumentu dystrybuanty $F(x)$ przez $a > 0$, dostajemy nową dystrybuantę $G(x) = F(x/a)$. To oznacza, że jeśli X ma dystrybuantę F , to aX ma dystrybuantę G .

1.6 Kilka funkcji specjalnych

Podamy teraz kilka funkcji specjalnych i ich własności.

- *Funkcja gamma*

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt, \quad x > 0.$$

Mamy $\Gamma(1) = 1$ i $\Gamma(1/2) = \sqrt{\pi}$. Całkowanie przez części pokazuje, że $\Gamma(x) = (x-1)\Gamma(x-1)$ i stąd $\Gamma(n) = (n-1)!$.

- *Funkcja beta*

$$B(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt = \int_0^\infty \frac{t^{a-1}}{(1+t)^{a+b}} dt, \quad a, b > 0.$$

Zauważmy, że $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b) = \Gamma(b, a)$.

- *Zmodyfikowana funkcja Bessela*

$$I_\nu(x) = \sum_{k=0}^{\infty} \frac{(x/2)^{2k+\nu}}{k!\Gamma(\nu+k+1)}, \quad x \in \mathbb{R}, \nu \in \mathbb{R}. \quad (1.11)$$

1.7 Rozkłady z lekkimi i ciężkimi ogonami

Mówimy, że rozkład z dystrybuntą $F(x)$ ma *wykładniczo ograniczony ogon* jeśli dla pewnego $a, b > 0$ mamy $\bar{F}(x) \leq ae^{-bx}$ dla wszystkich $x > 0$. Wtedy mówimy też że rozkład ma *lekki ogon*. Zauważmy, że wtedy $\hat{m}_F(s) < \infty$ dla $0 < s < s_0$, gdzie $s_0 > 0$. W przeciwnym razie mówimy, że rozkład ma *ciężki ogon* a więc wtedy $\hat{m}_F(s) = \infty$ dla wszystkich $s > 0$.

Niech $\alpha_F = \limsup_{x \rightarrow \infty} M(x)$, gdzie $M(x) = -\log \bar{F}(x)$ jest *funkcją hazardową* dla F . Poniższe rezultaty sformułujemy dla rozkładów na $\mathbb{R}_+$, t.j. spełniających $F(0-) = 0$.

Theorem 1.1 *Jeśli $\alpha_F = 0$, to F ma ciężki ogon.*

Proof Załóżmy, że $\alpha_F = 0$. Wtedy $\lim_{x \rightarrow \infty} M(x) = 0$. A więc, dla każdego $\varepsilon > 0$ istnieje $x' > 0$ taki, że $M(x) \leq \varepsilon x$ dla wszystkich $x \geq x'$. Dlatego dla pewnego $c > 0$ mamy $\bar{F}(x) \geq ce^{-\varepsilon x}$ dla wszystkich $x \geq 0$ i stąd

$$\int_0^\infty e^{sx} \bar{F}(x) dx = \infty, \quad (1.12)$$

dla wszystkich $s \geq \varepsilon$. Ponieważ $\varepsilon > 0$ jest dowolny, stąd (1.12) jest spełnione dla wszystkich $s > 0$, co oznacza, że F ma ciężki ogon. $\square$

Uwaga Dla dystrybunaty z ciężkim ogonem F mamy

$$\lim_{x \rightarrow \infty} e^{sx} \bar{F}(x) = \infty \quad (1.13)$$

dla wszystkich $s > 0$.

2 Rodziny rozkładów

Podamy teraz kilka klas rozkładów dyskretnych i absolutnie ciągłych. Wszystkie są scharakteryzowane przez skończony zbiór parametrów. Podstawowe charakterystyki podane są w tablicy rozkładów w rozdziale III.3.1.

W zbiorze rozkładów absolutnie ciągłych wyróżniamy dwie podklasy: rozkłady z lekkimi ogonami i ciężkimi ogonami. Rozróżnienie to jest bardzo ważne w praktyce i w teorii Monte Carlo.

2.1 Rozkłady dyskretnie

Wśród dyskretnych rozkładów mamy:

- *Rozkład zdegenerowany* δ_a skoncentrowany w $a \in \mathbb{R}$ z $p(x) = 1$ jeśli $x = a$ i $p(x) = 0$ w przeciwnym razie.
- *Rozkład Bernoulliego* $\text{Ber}(p)$ z $p_k = p^k(1-p)^{1-k}$ dla $k = 0, 1$; $0 < p < 1$: rozkład przyjmujący tylko wartości 0 i 1.
- *Rozkład dwumianowy* $B(n, p)$ z $p_k = \binom{n}{k} p^k (1-p)^{n-k}$ dla $k = 0, 1, \dots, n$; $n \in \mathbb{Z}_+$, $0 < p < 1$: rozkład sumy n niezależnych i jednakowo rozłożonych zmiennych losowych o rozkładzie $\text{Ber}(p)$. A więc, $B(n_1, p) * B(n_2, p) = B(n_1 + n_2, p)$.
- *Rozkład Poissona* $\text{Poi}(\lambda)$ z $p_k = e^{-\lambda} \lambda^k / k!$ dla $k = 0, 1, \dots$; $0 < \lambda < \infty$: jedna z podstawowych cegiełek na których zbudowany jest rachunek prawdopodobieństwa. Historycznie pojawił się jako graniczny dla rozkładów dwumianowych $B(n, p)$ gdy $np \rightarrow \lambda$ dla $n \rightarrow \infty$. Ważną własnością jest zamkniętość ze względu na splot, tzn. suma niezależnych zmiennych Poissonowskich ma znowu rozkład Poissona. Formalnie piszemy, że $\text{Poi}(\lambda_1) * \text{Poi}(\lambda_2) = \text{Poi}(\lambda_1 + \lambda_2)$.
- *Rozkład geometryczny* $\text{Geo}(p)$ z $p_k = (1-p)p^k$ dla $k = 0, 1, \dots$; $0 < p < 1$: rozkład mający dyskretny brak pamięci. To jest $\mathbb{P}(X \geq i+j \mid X \geq j) = \mathbb{P}(X \geq i)$ dla wszystkich $i, j \in \mathbb{N}$ wtedy i tylko wtedy gdy X jest geometrycznie rozłożony.
- *Rozkład obcięty geometryczny* $\text{TG}(p)$ z $p_k = (1-p)p^{k-1}$ dla $k = 1, 2, \dots$; $0 < p < 1$: jeśli $X \sim \text{Geo}(p)$, to $X|X > 0 \sim \text{TG}(p)$, jest to rozkład liczb prób Bernoulliego do pierwszego sukcesu, jeśli prawdopodobieństwo sukcesu wynosi $p' = 1-p$. A więc jeśli $Y \sim \text{TG}(p')$, to $\mathbb{E}Y = 1/p'$

- *Rozkład ujemny dwumianowy* lub inaczej *rozkład Pascala* $NB(\alpha, p)$ z $p_k = \Gamma(\alpha + k)/(\Gamma(\alpha)\Gamma(k + 1))(1 - p)^\alpha p^k$ dla $k = 0, 1, \dots$; $\alpha > 0, 0 < p < 1$. Powyżej piszemy

$$\binom{x}{0} = 1, \quad \binom{x}{k} = \frac{x(x-1)\dots(x-k+1)}{k!},$$

dla $x \in \mathbb{R}$, $k = 1, 2, \dots$. Mamy

$$p_k = \binom{\alpha + k - 1}{k} (1 - p)^\alpha p^k = \binom{-\alpha}{k} (1 - p)^\alpha (-p)^k.$$

Co więcej, dla $\alpha = 1, 2, \dots$, $NB(\alpha, p)$ jest rozkładem sumy α niezależnych i jednakowo rozłożonych $Geo(p)$ zmiennych losowych. To oznacza na przykład, że podklasa ujemnych rozkładów dwumianowych $\{NB(\alpha, p), \alpha = 1, 2, \dots\}$ jest zamknięta ze względu na spłot t.j. $NB(\alpha_1, p) * NB(\alpha_2, p) = NB(\alpha_1 + \alpha_2, p)$ jeśli $\alpha_1, \alpha_2 = 1, 2, \dots$. Co więcej poprzedni wzór jest spełniony dla $\alpha_1, \alpha_2 > 0$ and $0 < p < 1$.

- *Rozkład logarytmiczny* $Log(p)$ z $p_k = p^k / (-k \log(1 - p))$ dla $k = 1, 2, \dots$; $0 < p < 1$: pojawia się jako granica obciętych rozkładów dwumianowych.
- *Rozkład (dyskretny) jednostajny* $UD(n)$ z $p_k = n^{-1}$ dla $k = 1, \dots, n$; $n = 1, 2, \dots$

2.2 Rozkłady absolutnie ciągłe

Wśród rozkładów absolutnie ciągłych mamy:

- *Rozkład normalny* $N(\mu, \sigma^2)$ z $f(x) = (2\pi\sigma^2)^{-1/2} e^{-(x-\mu)^2/(2\sigma^2)}$ dla wszystkich $x \in \mathbb{R}$; $\mu \in \mathbb{R}, \sigma^2 > 0$: obok rozkładu Poissona druga cegiełka na której opiera się rachunek prawdopodobieństwa. Pojawia się jako granica sum niezależnych asymptotycznie zaniedbywalnych zmiennych losowych. Rozkłady normalne są zamknięte na branie spłotów, t.j. $N(\mu_1, \sigma_1^2) * N(\mu_2, \sigma_2^2) = N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.
- *Rozkład wykładniczy* $Exp(\lambda)$ z $f(x) = \lambda e^{-\lambda x}$ dla $x > 0$; $\lambda > 0$: podstawowy rozkład w teorii procesów Markowa ze względu na własność braku pamięci, t.j. $\mathbb{P}(X > t + s | X > s) = \mathbb{P}(X > t)$ dla wszystkich $s, t \geq 0$ jeśli X jest rozkładem wykładniczym.
- *Rozkład Erlanga* $Erl(n, \lambda)$ z $f(x) = \lambda^n x^{n-1} e^{-\lambda x} / (n-1)!$; $n = 1, 2, \dots$: rozkład sumy n niezależnych i jednakowo rozłożonych $Exp(\lambda)$ zmiennych losowych. Tak więc, $Erl(n_1, \lambda) * Erl(n_2, \lambda) = Erl(n_1 + n_2, \lambda)$.

- *Rozkład chi kwadrat* $\chi^2(n)$ z $f(x) = (2^{n/2}\Gamma(n/2))^{-1}x^{(n/2)-1}e^{-x/2}$ jeśli $x > 0$ and $f(x) = 0$ if $x \leq 0$; $n = 1, 2, \dots$: Rozkład sumy n niezależnych i jednakowo rozłożonych zmiennych losowych, które są kwadratami $N(0, 1)$ zmiennych losowych. To oznacza, że $\chi^2(n_1) * \chi^2(n_2) = \chi^2(n_1 + n_2)$.
- *Rozkład gamma* $\Gamma(a, \lambda)$ z $f(x) = \lambda^a x^{a-1} e^{-\lambda x} / \Gamma(a)$ for $x \geq 0$; z parametrem kształtu $a > 0$ i parametrem skali $\lambda > 0$: jeśli $a = n \in \mathbb{N}$, to $\Gamma(n, \lambda) = \text{Erl}(n, \lambda)$. Jeśli $\lambda = 1/2$ i $a = n/2$, to $\Gamma(n/2, 1/2) = \chi^2(n)$.
- *Rozkład jednostajny* $U(a, b)$ z $f(x) = (b - a)^{-1}$ dla $a < x < b$; $-\infty < a < b < \infty$.
- *Rozkład beta* $\text{Beta}(a, b, \eta)$ z

$$f(x) = \frac{x^{a-1}(\eta - x)^{b-1}}{B(a, b)\eta^{a+b-1}}$$

dla $0 < x < \eta$, gdzie $B(a, b)$ oznacza funkcję beta (patrz 1.6); $a, b, \eta > 0$; jeśli $a = b = 1$, to $\text{Beta}(1, 1, \eta) = U(0, \eta)$.

- *Rozkład odwrotny gaussowski (inverse Gaussian distribution)* $\text{IG}(\mu, \lambda)$ z

$$f(x) = (\lambda/(2\pi x^3))^{1/2} \exp(-\lambda(x - \mu)^2/(2\mu^2 x))$$

dla wszystkich $x > 0$; $\mu \in \mathbb{R}, \lambda > 0$.

- *Rozkład statystyki ekstremalnej* $\text{EV}(\gamma)$ z $F(x) = \exp(-(1 + \gamma x)_+^{-1/\gamma})$ dla wszystkich $\gamma \in \mathbb{R}$, gdzie wielkość $-(1 + \gamma x)_+^{-1/\gamma}$ jest zdefiniowana jako e^{-x} kiedy $\gamma = 0$. W tym ostatnim przypadku rozkład jest znany jako *rozkład Gumbela*. Dla $\gamma > 0$ otrzymujemy *rozkład Fréchéta*; rozkład ten jest skoncentrowany na półprostej $(-1/\gamma, +\infty)$. W końcu, dla $\gamma < 0$ otrzymujemy *rozkład ekstremalny Weibulla*, skoncentrowany na półprostej $(-\infty, -1/\gamma)$.

2.3 Rozkłady z ciężkimi ogonami

Przykładami absolutnie ciągłych rozkładów, które mają ciężki ogon są podane niżej.

- *Rozkład logarytmiczno normalny* $\text{LN}(a, b)$ z $f(x) = (xb\sqrt{2\pi})^{-1} \exp(-(\log x - a)^2/(2b^2))$ dla $x > 0$; $a \in \mathbb{R}, b > 0$; jeśli X jest o rozkładzie $N(a, b)$ to e^X ma rozkład $\text{LN}(a, b)$.

- *Rozkład Weibulla* $W(r, c)$ z $f(x) = rcx^{r-1} \exp(-cx^r)$ dla $x > 0$ gdzie parametr kształtu $r > 0$ i parametr skali $c > 0$; $\overline{F}(x) = \exp(-cx^r)$; jeśli $r \geq 1$, to $W(r, c)$ ma lekki ogon, w przeciwnym razie $W(r, c)$ ma ciężki.
- *Rozkład Pareto* $\text{Par}(\alpha)$ z $f(x) = \alpha(1/(1+x))^{\alpha+1}$ for $x > 0$; gdzie eksponent $\alpha > 0$; $\overline{F}(x) = (1/(1+x))^\alpha$.
- *Rozkład loggamma* $\text{LG}(a, \lambda)$ z $f(x) = \lambda^a / \Gamma(a) (\log x)^{a-1} x^{-\lambda-1}$ dla $x > 1$; $\lambda, a > 0$; jeśli X jest $\Gamma(a, \lambda)$, to e^X ma rozkład $\text{LG}(a, \lambda)$.

2.4 Rozkłady wektorów losowych

- *Rozkład jednostajny na* $B \subset \mathbb{R}^d$, $U(B)$, gdzie $|B| < \infty$ z $f(\mathbf{x}) = 1/|B|$ dla $\mathbf{x} \in B$ oraz $f(\mathbf{x}) = 0$ dla $\mathbf{x} \notin B$.
- *Standardowy d wymiarowy rozkład normalny* $N_d(\mathbf{0}, \mathbf{I})$ jest rozkładem zmiennych losowych $X_1, \dots, X_d$ niezależnych o jednakowym rozkładzie $N(0,1)$; ma gęstość $f(\mathbf{x}) = 1/(2\pi)^{d/2} \exp(-\mathbf{x}^T \mathbf{x}/2)$, $\mathbf{x} \in \mathbb{R}^d$.
- *d wymiarowy rozkład normalny* $N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ jest rozkładem wektora losowego $\mathbf{X} = (X_1, \dots, X_d)^T$ z wektorem średniej $\boldsymbol{\mu}$ i macierzą kowariancji $\boldsymbol{\Sigma}$, który dla dowolnego $\mathbf{a} \in \mathbb{R}^d$ ma własność, że zmienna losowa $\mathbf{a}^T \mathbf{X}$ ma rozkład normalny; jeśli $\mathbf{Y} \sim N(\mathbf{0}, \mathbf{I})$ oraz $\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^T$ to $\mathbf{X} = \mathbf{L}\mathbf{Y} + \boldsymbol{\mu}$ ma rozkład $N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$; jeśli $\boldsymbol{\Sigma}$ jest niesingularna, to $\mathbf{X}$ ma gęstość $f(\mathbf{x}) = 1/(2\pi)^{d/2} (\det(\boldsymbol{\Sigma}))^{-1/2} \exp((\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})/2)$.