**Andrzej Sierociński**

# RACHUNEK PRAWDOPODOBIEŃSTWA

## Zmienne losowe
## Wykład 4

# 8 Funkcje zmiennych losowych

Niech $X$ będzie zmienną losową zdefiniowaną na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, P)$ a $g$ pewną funkcją rzeczywistą. Zdefiniujmy nową zm. los.

$$Y(\omega) = g(X(\omega)), \quad \omega \in \Omega.$$

Powstaje pytanie: *Jak uzyskać rozkład prawdopodobieństwa zmiennej losowej $Y$ jeżeli znamy rozkład prawdopodobieństwa zmiennej losowej $X$?*

## 8.1 Przypadek, gdy $X$ jest typu dyskretnego

Nietrudno zauważyć, że skoro nośnik zm.los. $X$ jest co najwyżej przeliczalny, to także co najwyżej przeliczalny jest nośnik zm. los. $Y$, ponieważ

$$\mathcal{Y} = g(\mathcal{X}),$$

gdzie $\mathcal{X}$ jest nośnikiem zm. los. $X$. Zatem $Y$ jest również zm. los. typu dyskretnego.

Załóżmy, że $\mathcal{Y} = \{y_1, y_2, \ldots\}$. Wówczas funkcja prawdopodobieństwa zm. los. $Y$ zdefiniowana jest następująco:

$$\begin{aligned} P(Y = y_i) &= P(\{\omega \in \Omega : Y(g(X(\omega)) = y_i\}) = \\ &= \sum_{\{x_j \in \mathcal{X} : g(x_j) = y_i\}} P(X = x_j). \end{aligned}$$

**Przykład.** *Wiadomo, że płyty DVD-RW produkowane przez pewną firmę są wadliwe z prawdopodobieństwem 0.01 oraz że przypadki wadliwych płyt są wzajemnie niezależne. Firma sprzedaje płyty w paczkach po 10 sztuk oraz prowadzi politykę zwrotu gotówki polegającą na tym, że w przypadku uznanej reklamacji za jedną wadliwą płytę w paczce zwraca koszt tej płyty, a w przypadku, gdy w paczce jest więcej niż jedna wadliwa płyta - zwraca koszt całej paczki. Cena jednej płyty wynosi $c$, a całej paczki $10c$.*
*Niech $Y$ oznacza zm. los. równą dodatkowym kosztom związanym ze stosowaniem tej polityki przy sprzedaży jednej paczki. Znaleźć funkcję prawdopodobieństwa zm. los. $Y$.*

***Rozwiązanie.***
Oznaczmy przez $X$ zm. los. równą liczbie wadliwych płyt w jednej paczce. Nietrudno zauważyć, że liczba wadliwych płyt jest liczbą sukcesów w schemacie dwumianowym $B(10, 0.01)$, 10 niezależnych doświadczeń Bernoulliego z prawdopodobieństwem sukcesu $p$. Wówczas funkcja prawdopodobieństwa zm. los. $X$ dana jest równością

$$P(X = k) = b(k; 10, 0.01) = \binom{10}{k} \cdot 0.01^k (1 - 0.01)^{10-k},$$

gdzie $k = 0, 1, 2, \ldots, 10$.
Zdefiniujmy funkcję

$$g(k) = \begin{cases} 0 & \text{dla} \quad k = 0 \\ c & \text{dla} \quad k = 1 \\ 10c & \text{dla} \quad k \geqslant 2 \end{cases}$$

Zatem $Y = g(X)$, a nośnikiem jest zbiór $\mathcal{Y} = \{0, c, 10c\}$.

Oczywiście

$$Y = 0 \quad \text{wtw} \quad X = 0,$$

$$Y = c \quad \text{wtw} \quad X = 1,$$

oraz

$$Y = 10c \quad \text{wtw} \quad X \geqslant 2.$$

Ostatecznie otrzymujemy funkcję prawdopodobieństwa zm. los. $Y$, ($p(y) = P(Y = y)$)

| $y$ | 0 | c | 10c |
|---|---|---|---|
| $p(y)$ | 0.904 | 0.091 | 0.005 |

## 8.2 Przypadek, gdy $X$ jest typu ciągłego

W tym przypadku zm. los. $Y$ może być typu ciągłego, dyskretnego lub mieszanego. Jak znaleźć rozkład $Y$ pokażemy na przykładach.

Załóżmy, że dla zm. los. $X$ dane są dystrybuanta $F_X$ oraz gęstość $f_X$. Jako pierwszy rozważmy przypadek funkcji liniowej

$$Y = aX + b, \quad \text{gdzie} \quad a \neq 0.$$

W przypadku, gdy $a > 0$ mamy

$$\begin{aligned} F_Y(y) &= P(Y \leqslant y) = P(aX + b \leqslant y) = P\left(X \leqslant \frac{y-b}{a}\right) = \\ &= F_X\left(\frac{y-b}{a}\right) \end{aligned}$$

i co za tym idzie $f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X\left(\frac{y-b}{a}\right) = \frac{1}{a}f_X\left(\frac{y-b}{a}\right)$.

Podobnie, dla $a < 0$, mamy

$$\begin{aligned} F_Y(y) &= P(Y \leqslant y) = P(aX + b \leqslant y) = P\left(X \geqslant \frac{y-b}{a}\right) = \\ &= 1 - F_X\left(\frac{y-b}{a}\right), \end{aligned}$$

oraz

$$f_Y(y) = \frac{d}{dy}[1 - F_Y(y)] = -\frac{d}{dy}F_X\left(\frac{y-b}{a}\right) = -\frac{1}{a}f_X\left(\frac{y-b}{a}\right).$$

Ostatecznie

**Gęstość liniowej funkcji zm.los. typu ciągłego**

$$\forall a \neq 0,\ b \in R \quad f_{aX+b}(y) = \frac{1}{|a|}f_X\left(\frac{y-b}{a}\right).$$

Zauważmy, że liniowa funkcja $f(x) = ax + b$, dla $a \neq 0$ jest funkcją różnowartościową (wzajemnie jednoznaczną).

Problem się komplikuje, jeżeli funkcja $g(x)$ nie jest wzajemnie jednoznaczna.

Rozważmy funkcję kwadratową, tzn. $Y = X^2$, gdzie $X$ jest typu ciągłego o dystrybuancie $F_X$ i gęstości $f_X$. Wówczas dla wszystkich $y \geqslant 0$ mamy

$$\begin{aligned} F_Y(y) &= P(Y \leqslant y) = P(X^2 \leqslant y) = P\left(|X| \leqslant \sqrt{y}\right) = \\ &= F_X\left(\sqrt{y}\right) - F_X\left(-\sqrt{y}\right). \end{aligned}$$

Natomiast gęstość dla wszystkich $y > 0$ ma postać

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_Y(y) = \frac{d}{dy}[F_X\left(\sqrt{y}\right) - F_X\left(-\sqrt{y}\right)] = \\ &= \frac{1}{2\sqrt{y}}[f_X\left(\sqrt{y}\right) + f_X\left(-\sqrt{y}\right)]. \end{aligned}$$

### Gęstość kwadratu zm. los. $X$

Dla dowolnej zm. los. $X$ o gęstości $f_X$

$$f_{X^2}(y) = \begin{cases} 0 & \text{dla} \quad x \leqslant 0 \\ \frac{1}{2\sqrt{y}}\left[f_X\left(\sqrt{y}\right) + f_X\left(-\sqrt{y}\right)\right] & \text{dla} \quad x > 0. \end{cases}$$

**Przykład.** *Niech $X$ będzie zm. los. typu ciągłego o gęstości*

$$f(x) = \frac{1}{2} e^{-|x|}.$$

*Wyznaczyć gęstości zmiennych losowych $Y = 2X + 1$, $Z = X^2$ oraz $T = |X|$.*

**Rozwiązanie.**
Wykorzystując wcześniej wyznaczoną formułę na gęstość funkcji liniowej ($a = 2$ and $b = 1$) mamy

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right) = \frac{1}{4} \exp\left\{\frac{|y-1|}{2}\right\} \text{ dla all } y \in R.$$

Podobnie, dla $z > 0$ mamy

$$f_Z(z) = \frac{1}{2\sqrt{z}}[f_X\left(\sqrt{z}\right) + f_X\left(-\sqrt{z}\right)] = \frac{1}{2\sqrt{z}} e^{-\sqrt{z}}.$$

Ostatecznie

$$f_Z(z) = \begin{cases} 0 & \text{dla} \quad z \leqslant 0 \\ \frac{1}{2\sqrt{z}} e^{-\sqrt{z}} & \text{dla} \quad z > 0. \end{cases}$$

Zmienna losowa $T = |X|$ przyjmuje tylko nieujemne wartości. Niech $t \geqslant 0$.

$$F_T(t) = P(T \leqslant t) = P\left(|X| \leqslant t\right) = F_X\left(t\right) - F_X\left(-t\right),$$

oraz

$$f_T(t) = f_X(t) + f_X(-t) = e^{-t}.$$

Zatem

$$f_T(t) = \begin{cases} 0 & \text{dla} \quad t \leqslant 0 \\ e^{-t} & \text{dla} \quad t > 0. \end{cases}$$

**Twierdzenie** Jeżeli $g$ jest funkcją ściśle monotoniczną, różniczkowalną a $X$ zm. los. typu ciągłego o gęstości $f_X$, to zm. los. $Y = g(X)$ jest również typu ciągłego o gęstości

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \frac{1}{|g'(g^{-1}(y))|}.$$

**Dowód.**
Załóżmy, że $g$ jest ściśle rosnąca.
Istnieje zatem funkcja odwrotna $g^{-1}$, która także jest funkcją ściśle rosnącą .
Stąd otrzymujemy

$$\begin{aligned} F_Y(y) &= P(Y \leqslant y) = P(g(X) \leqslant y) = P(X \leqslant g^{-1}(y)) = \\ &= F_X(g^{-1}(y)). \end{aligned}$$

Różniczkując ostatnie wyrażenie względem $y$, otrzymujemy

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X(g^{-1}(y)) = f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y) = \\ &= f_X(g^{-1}(y)) \cdot \frac{1}{g'(g^{-1}(y))}. \end{aligned}$$

Podobnie, w przypadku, gdy $g$ jest ściśle malejąca, wówczas $g^{-1}$ jest również ściśle malejąca, i

$$F_Y(y) = P(g(X) \leqslant y) = P(X \geqslant g^{-1}(y)) = 1 - F_X(g^{-1}(y)).$$

Różniczkując dystrybuantę $F_Y(y)$ względem $y$, otrzymujemy, że

$$f_Y(y) = -f_X(g^{-1}(y)) \cdot \frac{1}{g'(g^{-1}(y))}.$$

$\blacksquare$