Wstęp do rachunku prawdopodobieństwa
6. Funkcje zmiennej losowej. Zmienne losowe wielowymiarowe. Wartość oczekiwana, wariancja i kowariancja.

Definicja. Wartością oczekiwaną *zmiennej losowej dyskretnej X o wartościach (atomach) w zbiorze A nazywamy liczbę*

$$\mathbb{E}X = \sum_{a \in A} a\mathbb{P}(X = a).$$

Uwaga: Jeśli $X : \Omega \to \mathbb{R}$ jest zmienną losową, a $g : \mathbb{R} \to \mathbb{R}$ jest funkcją (mierzalną), to funkcja $g(X) : \Omega \to \mathbb{R}$ będąca złożeniem funkcji X i g jest również zmienną losową.

Twierdzenie (Prawo leniwego statystyka)**.** *Niech X będzie zmienną losową dyskretną o wartościach (atomach) w zbiorze A i $f : \mathbb{R} \to \mathbb{R}$ będzie funkcją (mierzalną), wtedy*

$$\mathbb{E}(f(X)) = \sum_{a \in A} f(a)\mathbb{P}(X = a).$$

Definicja. Wariancją *zmiennej losowej X nazywamy liczbę*

$$VarX = \mathbb{E}((X - \mathbb{E}X)^2).$$

Uwaga: Przy liczeniu wariancji praktyczniejsze zastosowanie ma wzór:

$$\mathrm{Var}X = \mathbb{E}X^2 - (\mathbb{E}X)^2.$$

Uwaga: Istnieją zmienne losowe X, dla których $\mathbb{E}X$ nie jest określona (bo definiujący ją szereg nie jest bezwzględnie zbieżny). W takim przypadku $\mathrm{Var}X$ również jest niezdefiniowana. Są również zmienne losowe X, dla których istnieje $\mathbb{E}X$, a nie istnieje $\mathrm{Var}X$.

Definicja. *Funkcję (mierzalną)*

$$(X, Y) : \Omega \to \mathbb{R}^2 : \omega \to (X(\omega), Y(\omega))$$

nazywamy **dwuwymiarową zmienną losową** *lub czasami* **dwuwymiarowym wektorem losowym***. Współrzędne tej funkcji, X i Y, są „zwykłymi" zmiennymi losowymi określonymi na tej samej przestrzeni probabilitycznej $(\Omega, \mathcal{F}, \mathbb{P})$.*

We wszystkich definicjach poniżej ograniczymy się do przypadku, gdy zmienne losowe X i Y są zmiennymi losowymi dyskretnymi o atomach odpowiednio $\{x_1, x_2, \ldots\}$ i $\{y_1, y_2, \ldots\}$. Wtedy (X, Y) nazywamy **dwuwymiarową zmienną losową dyskretną** o atomach $\{(x_i, y_j) : i, j = 1, 2, \ldots\}$.

Definicja. Rozkład (łączny)*, lub* **funkcję masy prawdopodobieństwa***, zmiennej losowej (X, Y) definiujemy, podając wszystkie wartości*

$$p_{i,j} = \mathbb{P}(X = x_i, Y = y_j).$$

Wtedy oczywiście

$$\sum_{i,j} p_{i,j} = \sum_i \sum_j p_{i,j} = 1.$$

Definicja. Rozkłady brzegowe *zmiennych losowych X i Y wektora losowego (X, Y) obliczamy za pomocą wzorów*

$$\mathbb{P}(X = x_i) = \sum_j \mathbb{P}(X = x_i, Y = y_j)$$

oraz

$$\mathbb{P}(Y = y_j) = \sum_i \mathbb{P}(X = x_i, Y = y_j).$$

Jeśli dla **każdego** *atomu (x_i, y_j) zachodzi*

$$\mathbb{P}(X = x_i, Y = y_j) = \mathbb{P}(X = x_i)\mathbb{P}(Y = y_j),$$

to mówimy, że zmienne losowe X i Y są **niezależne***.*

Twierdzenie (Prawo leniwego statystyka). *Niech (X,Y) będzie zmienną losową dyskretną o wartościach (atomach) $\{(x_i,y_j): i,j=1,2,\dots\}$, a $f:\mathbb{R}^2\to\mathbb{R}$ będzie funkcją (mierzalną). Wtedy*

$$\mathbb{E}(f(X,Y)) = \sum_{i,j} f(x_i,y_j)\mathbb{P}\left(X=x_i, Y=y_j\right).$$

Zatem np.

$$\mathbb{E}XY = \sum_{i,j} x_i y_j \mathbb{P}\left(X=x_i, Y=y_j\right) .$$

Twierdzenie. *Dla dowolnych liczb rzeczywistych $a,b,c\in\mathbb{R}$ i funkcji (mierzalnych) $f,g:\mathbb{R}^2\to\mathbb{R}$ mamy*

$$\mathbb{E}(af(X,Y)+bg(X,Y)+c) = a\mathbb{E}f(X,Y)+b\mathbb{E}g(X,Y)+c.$$

Definicja. Kowariancją *zmiennej losowej (X,Y) nazywamy liczbę*

$$Cov(X,Y) = \mathbb{E}(X-\mathbb{E}X)(Y-\mathbb{E}Y) = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y,$$

a **współczynnik korelacji** $\rho(X,Y)$ *zdefiniowany jest wzorem*

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{VarX}\sqrt{VarY}}.$$

Uwaga: Istnieją zmienne losowe (X,Y), dla których $\mathrm{Cov}(X,Y)$ i $\rho(X,Y)$ nie są określone (bo definiujący je szereg nie jest zbieżny)! Jeśli jednak $\rho(X,Y)$ istnieje, to $|\rho(X,Y)|\leqslant 1$.

Twierdzenie. *Jeśli zmienne losowe X i Y są niezależne, to $\mathbb{E}XY=\mathbb{E}X\mathbb{E}Y$, a zatem*

$$Cov(X,Y)=\rho(X,Y)=0.$$

Uwaga: Twierdzenie odwrotne do powyższego nie jest prawdziwe.

Dodatek A. Zadania na ćwiczenia

Zadanie A.1. Zmienna losowa X posiada rozkład podany w poniższej tabeli:

k	0	1	2	3
$\mathbb{P}\left(X=k\right)$	0,4	0,3	0,2	0,1

a) Oblicz wartość oczekiwaną oraz wariancję zmiennej losowej X.
b) Znajdź rozkład zmiennej losowej $Y=|2-X|$ i używając go, znajdź jej wartość oczekiwaną.
c) Oblicz $\mathbb{E}|2-X|$, korzystając z „prawa leniwego statystyka”.

Zadanie A.2. Zmienna losowa X posiada rozkład podany w poniższej tabeli:

i	-2	-1	0	1	2	7
$\mathbb{P}(X=i)$	0,18	0,16	0,28	0,23	0,14	0,01

Znajdź rozkład zmiennej losowej $Y=sgn(X)$ i oblicz $\mathbb{E}Y$ oraz $\mathrm{Var}Y$.

Zadanie A.3. Zmienna losowa (X,Y) ma następujący rozkład:

$$\mathbb{P}\left(X=0,Y=-1\right)=1/2,\ \mathbb{P}\left(X=1,Y=1\right)=\mathbb{P}\left(X=-1,Y=1\right)=1/4.$$

Wyznacz rozkłady brzegowe obu zmiennych losowych i znajdź $\rho(X,Y)$. Czy zmienne te są niezależne?

Zadanie A.4. Rzucamy symetryczną kostką n razy. Dla każdego $n\geqslant 1$ rozstrzygnij, czy liczba wszystkich wyrzuconych jedynek X_n i liczba wszystkich wyrzuconych dwójek Y_n są niezależnymi zmiennymi losowymi. (Spróbuj najpierw zgadnąć odpowiedź i uzasadnić ją bez dokonywania obliczeń, a następnie przeprowadź formalny dowód, żeby potwierdzić swoją tezę.)

Zadanie A.5. W pierwszej urnie znajduje się 5 kul białych i 3 czarne, a w drugiej – 2 białe i 2 czarne. Z pierwszej urny losujemy jedną kulę, a z drugiej również jedną kulę. Niech X będzie zmienną losową oznaczającą liczbę kul białych wśród dwóch wylosowanych kul, a Y oznacza liczbę wylosowanych kul czarnych.

a) Znajdź rozkład łączny wektora losowego (X,Y).

b) Czy zmienne losowe X i Y są niezależne?
c) Oblicz $\mathbb{E}X$ i $\mathbb{E}Y$, korzystając z rozkładów brzegowych zmiennych losowych X i Y.
d) Oblicz $\mathbb{E}X$, przedstawiając X w postaci sumy $X = X_1 + X_2$, gdzie X_i oznacza liczbę kul białych wyciągniętych z i-tej urny.
e) Co można powiedzieć o rozkładzie zmiennej losowej $Z = X + Y$?
f) Znajdź współczynnik korelacji $\rho(X,Y)$. Czy można to zrobić prościej, korzystając z podpunktu e) i nie używając rozkładu łącznego zmiennej losowej (X,Y)?

Dodatek B. Zadania domowe

Zadanie B.1. Wiedząc, że $\mathbb{E}X = 1$ i $\mathrm{Var}X = 5$, znajdź $\mathbb{E}(2+X)^2$ oraz $\mathrm{Var}(4+3X)$.

Zadanie B.2. Zmienna losowa X ma rozkład
$$\mathbb{P}(X=i) = \frac{1}{2^i}, \text{ dla } i = 1,2,3, \quad \text{oraz} \quad \mathbb{P}(X=8) = \frac{1}{8}.$$
Wyznacz rozkład zmiennych losowych $Y = (-2)^X$ oraz $Z = 4X+1$. Oblicz wartości oczekiwane i wariancje zmiennych losowych Y oraz Z na dwa sposoby: korzystając z rozkładu i korzystając z „prawa leniwego statystyka”.

Zadanie B.3. Wektor losowy (X,Y) posiada rozkład podany w poniższej tabeli:

$X \setminus Y$	-1	0	1
1	1/9	0	1/3
2	0	1/9	0
3	1/3	0	1/9

a) Znajdź rozkłady brzegowe zmiennych losowych X i Y.
b) Czy zmienne losowe X i Y są niezależne?
c) Oblicz $\mathbb{E}X$ i $\mathbb{E}Y$, korzystając z rozkładów brzegowych zmiennych losowych X i Y.
d) Wyznacz rozkład zmiennej losowej $Z = X+Y$, następnie oblicz $\mathbb{E}Z$ na dwa sposoby – korzystając z „prawa leniwego statstyka” i korzystając z rozkładu zmiennej losowej Z.
e) Znajdź współczynnik korelacji $\rho(X,Y)$.

Zadanie B.4. Wyznacz rozkład zmiennej losowej $Y = \cos(\pi X)$, jeżeli zmienna losowa X ma rozkład:
$$\mathbb{P}(X=k) = \frac{1}{2^k}, \quad k = 1,2,\ldots.$$
Następnie oblicz $\mathbb{E}Y$ i $\mathrm{Var}Y$.
Wskazówka: Na początku ustal jaki jest zbiór atomów zmiennej losowej Y.

Zadanie B.5. Rzucamy kostką tak długo, dopóki nie wyrzucimy wyniku **różnego** od jedynki. Niech X oznacza liczbę wykonanych rzutów.
a) Znajdź wartość oczekiwaną i wariancję zmiennej losowej X.
b) Znajdź wartość oczekiwaną i wariancję zmiennej losowej $Y = 6^X$.
c) Znajdź wartość oczekiwaną i wariancję zmiennej losowej $Y = 3^X$.
Wskazówka: Należy wykorzystać wzory (prawdziwe dla każdego $x \in (-1,1)$):
$$\sum_{k=1}^{\infty} kx^k = \frac{x}{(1-x)^2},$$
$$\sum_{k=1}^{\infty} k^2x^k = \frac{x(1+x)}{(1-x)^3}.$$

Zadanie B.6. Zmienna losowa X posiada rozkład podany w poniższej tabeli

k	-1	0	1	3
$\Pr(X=k)$	$\frac{2}{10}$	$\frac{5}{10}$	$\frac{3}{20}$	$\frac{3}{20}$

a) Znajdź rozkład zmiennej losowej $Y = |X|$, jej wartość oczekiwaną oraz wariancję.
b) Znajdź rozkład zmiennej losowej $Z = \sin(\pi X/2)$, jej wartość oczekiwaną oraz wariancję.

Zadanie B.7. Rzucamy raz symetryczną kostką. Pierwszy gracz wygrywa 2, gdy liczba oczek jest parzysta, a przegrywa 1 (tj. wygrywa -1), gdy liczba oczek jest nieparzysta. Drugi gracz wygrywa 2, gdy liczba oczek wynosi co najmniej 4, a przegrywa 1 w przeciwnym przypadku. Niech X_i oznacza wygraną i-tego gracza.

a) Znajdź rozkład wektora losowego (X_1, X_2).
b) Znajdź rozkłady brzegowe wektora losowego (X_1, X_2) i sprawdź, czy zmienne losowe X_1 i X_2 są niezależne.
c) Oblicz współczynnik korelacji $\rho(X_1, X_2)$.
d) Znajdź $\mathbb{E}(X_1^2 X_2^3)$.

Odpowiedzi do zadań domowych

B.1 $\mathbb{E}((2+X)^2) = 14$, $\mathrm{Var}(4+3X) = 45$

B.2

k	-8	-2	4	256
$\mathbb{P}(Y=k)$	1/8	1/2	1/4	1/8

$\mathbb{E}Y = 31$, $\mathrm{Var}Y = 7245$

k	5	9	13	33
$\mathbb{P}(Z=k)$	1/2	1/4	1/8	1/8

$\mathbb{E}Z = 10,5$, $\mathrm{Var}Z = 79,75$

B.3

a)

k	1	2	3
$\mathbb{P}(X=k)$	$\frac{4}{9}$	$\frac{1}{9}$	$\frac{4}{9}$

k	-1	0	1
$\mathbb{P}(Y=k)$	$\frac{4}{9}$	$\frac{1}{9}$	$\frac{4}{9}$

b) Nie

c) $\mathbb{E}X = 2$, $\mathbb{E}Y = 0$

d)

k	0	2	4
$\mathbb{P}(X=k)$	$\frac{1}{9}$	$\frac{7}{9}$	$\frac{1}{9}$

$\mathbb{E}Z = 2$

e) $\rho(X,Y) = -\frac{1}{2}$

B.4

$\mathbb{P}(Y=-1) = \frac{2}{3}$, $\mathbb{P}(Y=1) = \frac{1}{3}$, $\mathbb{E}Y = -\frac{1}{3}$, $\mathrm{Var}Y = \frac{8}{9}$

B.5

a) $\mathbb{E}X = \frac{6}{5}$, $\mathrm{Var}X = \frac{6}{25}$
b) Zmienna losowa Y nie ma wartości oczekiwanej ani wariancji.
c) $\mathbb{E}Y = 2,5$; wariancja nie istnieje.

B.6

a)

k	0	1	3
$\mathbb{P}(Y=k)$	$\frac{1}{2}$	$\frac{7}{20}$	$\frac{3}{20}$

$\mathbb{E}Y = 0,8$, $\mathrm{Var}Y = 1,06$

b)

k	-1	0	1
$\mathbb{P}(Z=k)$	$\frac{7}{20}$	$\frac{1}{2}$	$\frac{3}{20}$

$\mathbb{E}Z = -0,2$, $\mathrm{Var}Z = 0,46$

B.7

a)

$X \setminus Y$	-1	2
-1	$\frac{1}{3}$	$\frac{1}{6}$
2	$\frac{1}{6}$	$\frac{1}{3}$

b)

k	-1	2
$\mathbb{P}(X_1=k)$	$\frac{1}{2}$	$\frac{1}{2}$

k	-1	2
$\mathbb{P}(X_2=k)$	$\frac{1}{2}$	$\frac{1}{2}$

Nie są niezależne.

c) $\rho(X_1, X_2) = \frac{1}{3}$

d) $\mathbb{E}X_1^2 X_2^3 = 11$