Metody Probabilistyczne i Statystyka — zadania

dr Małgorzata Kuchta, Agnieszka Plucińska, Edmund Pluciński

opracował: Konrad Kleczkowski i inni

Lista 8

Zadanie 1. Rzucamy dwa razy kostką symetryczną o k ścianach z numerammi od 1 do k i otrzymujemy wartości X_1 i X_2. Ile wynosi $\mathbb{E}[\max(X_1, X_2)]$? Ile wynosi $\mathbb{E}[\min(X_1, X_2)]$? Ile wynosi $\mathbb{E}[\max(X_1, X_2) + \min(X_1, X_2)]$?

Rozwiązanie.
By obliczyć $\mathbb{E}[\min(X_1, X_2)]$, należy zliczyć ile jest takich X_1 oraz X_2, że $\min(X_1, X_2) = i$, gdzie $i \in \{1, \ldots, k\}$.

Ustalmy i. Zauważmy, że równanie jest spełnione przez pary (j, i), (i, i) oraz (i, j), gdzie $k \geqslant j > i \geqslant 1$. Stąd mamy, że $j \in \{i+1, \ldots, k\}$. Wobec tego otrzymujemy, że $|\{(X_1, X_2) : \min(X_1, X_2) = i\}| = 2\,(k - (i+1) + 1) + 1 = 2(k-i) + 1$, czyli $\mathbb{P}[\min(X_1, X_2) = i] = \frac{2(k-i)+1}{k^2}$. Więc:

$$
\begin{aligned}
\mathbb{E}[\min(X_1, X_2)] &= \sum_{i=1}^{k} i \frac{2(k-i)+1}{k^2} \\
&= \sum_{i=1}^{k} \frac{2ki - 2i^2 + i}{k^2} \\
&= \sum_{i=1}^{k} \frac{(2k+1)i - 2i^2}{k^2} \\
&= \sum_{i=1}^{k} \frac{(2k+1)i}{k^2} - \sum_{i=1}^{k} \frac{2i^2}{k^2} \\
&= \frac{2k+1}{k^2} \left(\sum_{i=1}^{k} i \right) - \frac{2}{k^2} \left(\sum_{i=1}^{k} i^2 \right) \\
&= \frac{2k+1}{k^2} \cdot \frac{k(k+1)}{2} - \frac{2}{k^2} \cdot \frac{k(k+1)(2k+1)}{6} \\
&= \frac{k(k+1)(2k+1)}{k^2} \left(\frac{1}{2} - \frac{2}{6} \right) \\
&= \frac{(k+1)(2k+1)}{6k}
\end{aligned}
$$

By obliczyć wartość oczekiwaną $\mathbb{E}[\max(X_1, X_2) + \min(X_1, X_2)]$ należy zauważyć, że:

$$
\begin{aligned}
\mathbb{E}[\max(X_1, X_2) + \min(X_1, X_2)] &= \mathbb{E}[X_1 + X_2] \\
&= 2\mathbb{E}[X_1] \\
&= 2 \sum_{i=1}^{k} \frac{i}{k} \\
&= k + 1
\end{aligned}
$$

Korzystając z liniowości w. o., mamy, że:

$$
\mathbb{E}[\max(X_1, X_2)] = k + 1 - \frac{(k+1)(2k+1)}{6k}
$$

Zadanie 2. Wybieramy losowo i niezależnie dwie liczby X i Y ze zbioru $\Omega = \{1, \dots, n\}$, to znaczy, dla każdej pary $(i,j) \in \Omega$ mamy $\mathbb{P}[X = i, Y = j] = \frac{1}{n^2}$. Wyznacz wartość oczekiwaną zmiennej losowej $Z = |X - Y|$.

Rozwiązanie.
By obliczyć $\mathbb{E}|X-Y|$ musimy obliczyć moc zbioru $\{(X,Y) : |X-Y| = i\}$ dla pewnego i. Ustalmy $i \in \{0, 1, \dots, n-1\}$. Zauważmy, że jeśli $|X - Y| = i$, to $Y = X - i$ lub $Y = X + i$. Stąd $|\{(X,Y) : |X - Y| = i\}| = 2(n-i)$, czyli $\mathbb{P}[|X - Y| = i] = \frac{2(n-i)}{n^2}$. Liczymy więc wartość oczekiwaną:

$$\begin{aligned}
\mathbb{E}|X - Y| &= \sum_{i=0}^{n-1} i\frac{2(n-i)}{n^2} \\
&= \sum_{i=0}^{n-1} \frac{2ni - 2i^2}{n^2} \\
&= \frac{2n}{n^2}\left(\sum_{i=0}^{n-1} i\right) - \frac{2}{n^2}\left(\sum_{i=0}^{n-1} i^2\right) \\
&= \frac{2}{n} \cdot \frac{n(n-1)}{2} - \frac{2}{n^2} \cdot \frac{(n-1)n(2n-1)}{6} \\
&= n - 1 - \frac{(n-1)(2n-1)}{3n}
\end{aligned}$$

Zadanie 3. Zmienna losowa X ma rozkład logarytmiczno-normalny, jeśli $\ln X \sim \mathcal{N}(\mu, \sigma)$. Wyznacz gęstość rozkładu logarytmiczno-normalnego.

Rozwiązanie.
Niech $Y = \ln X$. Przypomnijmy, że

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

Teraz zauważmy, że $F_Y(y) = \mathbb{P}[Y < y] = \mathbb{P}[\ln X < y] = \mathbb{P}[X < e^y] = F_X(e^y)$. Podstawiając $x = e^y$ otrzymujemy równanie $F_Y(\ln x) = F_X(x)$, czyli $f_X(x) = \frac{d}{dx}[F_Y(\ln x)] = \frac{f_y(\ln x)}{x}$. Zatem $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}$.

Zadanie 4. Do n urn wrzucamy losowo k kul. Niech X oznacza liczbę urn niepustych. Niech $Y = \frac{X}{n}$. Oblicz $\lim_{n\to\infty} \mathbb{E}Y$, gdzie $k = n$.

Rozwiązanie.
Wyznaczmy funkcję masy prawdopodobieństwa $\mathbb{P}[X = i]$. Wszystkich możliwych losowań kul do urn jest $\binom{k+n-1}{n-1}$. Teraz policzmy ile mamy sprzyjających losowań. Wybieramy i urn spośród n urn, które będą niepuste, na $\binom{n}{i}$ sposobów. Następnie do każdej z nich wrzucamy po jednej kuli i losujemy resztę kul na $\binom{(n-i)+i-1}{i-1} = \binom{n-1}{i-1}$ sposobów. Wobec tego mamy, że:

$$\mathbb{P}[X = i] = \frac{\binom{n}{i}\binom{n-1}{i-1}}{\binom{k+n-1}{n-1}}$$

Połóżmy $k = n$. Policzmy granicę:

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}\left[\frac{X}{n}\right] &= \lim_{n\to\infty} \frac{1}{n}\mathbb{E}X \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} i\frac{\binom{n}{i}\binom{n-1}{i-1}}{\binom{2n-1}{n-1}} \\
&= \lim_{n\to\infty} \frac{1}{\binom{2n-1}{n-1}}\sum_{i=1}^{n} \frac{i}{n}\binom{n}{i}\binom{n-1}{i-1} \\
&= \lim_{n\to\infty} \frac{1}{\binom{2n-1}{n-1}}\sum_{i=1}^{n} \binom{n-1}{i-1}^2 \\
&= \lim_{n\to\infty} \frac{1}{\binom{2n-1}{n-1}}\binom{2n-2}{n-1} \\
&= \lim_{n\to\infty} \frac{n}{2n-1} = \frac{1}{2}
\end{aligned}$$

Zadanie 5. Niech X będzie zmienną losową. Pokaż, że liczba $m = \mathbb{E}X$ minimalizuje funkcję $f(m) = \mathbb{E}(X-m)^2$.

Rozwiązanie.
Przekształcając:

$$\begin{aligned}
f(m) &= \mathbb{E}(X-m)^2 \\
&= \mathbb{E}(m^2 - 2Xm + X^2) \\
&= m^2 - 2\mathbb{E}[X]m + \mathbb{E}\left[X^2\right]
\end{aligned}$$

ponieważ $f(m)$ jest trójmianem kwadratowym

$$m_{\min} = -\frac{-2\mathbb{E}X}{2} = \mathbb{E}X \quad \square$$

Zadanie 6. Wyznacz wartość oczekiwaną i wariancję zmiennych losowych o rozkładzie:

a) hipergeometrycznym z parametrami N, M, n,

b) ujemnym rozkładzie dwumianowym z parametrami m, p,

c) z zadania 2 z listy 6.

Metoda (a).
Dana jest populacja $\mathcal{U}$ oraz podzbiór elementów populacji o interesującej własności $\mathcal{A} \subseteq \mathcal{U}$. Losujemy podzbiór $\mathcal{P} \subseteq \mathcal{U}$. Niech X będzie zmienną losową:

$$X : \mathcal{P} \mapsto |\mathcal{P} \cap \mathcal{A}|$$

Tak zdefiniowana zmienna X ma rozkład hipergeometryczny. Parametry określają następująco: $|\mathcal{U}| = N$, $|\mathcal{A}| = M$ oraz $|\mathcal{P}| = n$. Oznaczamy to jako $X \sim \mathcal{HG}(N, M, n)$.

Rozkład również można interpretować intuicyjnie w następujący sposób. Spośród grupy N studentów możemy wyróżnić M, które nie zdadzą Języków Formalnych i Teorii Translacji. Losujemy n studentów. Zdarzenie $\{X = i\}$ opisuje, że w próbce znajduje się i studentów, którzy nie zdadzą Języków Formalnych i Teorii Translacji.

Pomocniczo obliczmy $\mathbb{P}[X = i]$. Wszystkich możliwych próbek jest $\binom{N}{n}$. Obliczmy ile jest sprzyjających próbek. Wybieramy najpierw i spośród M elementów na $\binom{M}{i}$ sposobów, a następnie wybieramy $n - i$ spośród $N - M$ na $\binom{N-M}{n-i}$ sposobów. Wobec tego:

$$\mathbb{P}[X = i] = \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}}$$

Rozwiązanie (a).
Obliczmy wartość oczekiwaną:

$$\begin{aligned}\mathbb{E}X &= \sum_{i=0}^{n} i \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}} \\ &= \frac{1}{\binom{N}{n}} \sum_{i=0}^{n} i \binom{M}{i}\binom{N-M}{n-i}\end{aligned}$$

Suma $\sum_{i=0}^{n} i\binom{M}{i}\binom{N-M}{n-i}$ jest równa $M\binom{N-1}{n-1}$. Aby udowodnić tę zależność, wystarczy wyjść z równości

$$x[(1+x)^M]'(1+x)^{N-M} = Mx(1+x)^{N-1}$$

rozwinąć obustronnie w szereg i wykonać mnożenie szeregów formalnych. Mamy stąd:

$$\mathbb{E}X = \frac{M\binom{N-1}{n-1}}{\binom{N}{n}} = \frac{Mn}{N}$$

Obliczmy pomocniczo $\mathbb{E}X^2$.

$$\begin{aligned}\mathbb{E}X^2 &= \sum_{i=0}^{n} i^2 \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}} \\ &= \frac{1}{\binom{N}{n}} \sum_{i=0}^{n} i^2 \binom{M}{i}\binom{N-M}{n-i}\end{aligned}$$

By obliczyć sumę $\sum_{i=0}^{n} i^2\binom{M}{i}\binom{N-M}{n-i}$, zauważmy, że $x\left[x\left[(1+x)^M\right]'\right]' = \sum_{m=0}^{M} m^2\binom{M}{m}x^m$. Wtedy korzystamy z tożsamości

$$x\left[x\left[(1+x)^M\right]'\right]' \cdot (1+x)^{N-M} = \left[M^2x^2 + Mx\right](1+x)^{N-2}$$

i otrzymujemy, że suma jest równa $M^2\binom{N-2}{n-2} + M\binom{N-2}{n-1}$. Podstawiając do wzoru na wariancję, otrzymujemy:

$$\mathrm{Var}X = \mathbb{E}X^2 - \mathbb{E}^2X = \frac{M^2\binom{N-2}{n-2} + M\binom{N-2}{n-1}}{\binom{N}{n}} - \left(\frac{Mn}{N}\right)^2$$

Metoda (b).
Zanim rozpoczniemy jakiekolwiek rozważania, przyda się parę definicji.

Definicja (Rozkład zero-jedynkowy). Ustalmy $p \in [0,1]$. Niech X będzie zmienną losową, która przyjmuje wartości ze zbioru $\{0,1\}$ taką, że $\mathbb{P}[X=1]=p$. Taka zmienna ma rozkład zerojedynkowy, oznaczamy jako $X \sim \mathcal{O}(p)$.

Bardziej Biegły Czytelnik zauważy, że $\mathbb{E}X = p$ oraz $\mathrm{Var}X = p(1-p)$, jednakże wyprowadzenie tego zostawiam jako krótkie ćwiczenie w liczeniu wartości oczekiwanej i wariancji dla takiego rozkładu[1].

Niech $\{X_i\}_{i\in\mathbb{N}}$ będą ciągiem niezależnych zmiennych losowych o rozkładzie zero-jedynkowym z prawdopodobieństwem sukcesu równym p (zapisujemy to jako $X_i \sim \mathcal{O}(p)$).

Definicja (Rozkład Pascala). Ustalmy liczbę całkowitą $r \geqslant 1$. Dana jest zmienna losowa T

$$T = \min\left\{n \geqslant 1 : \sum_{i=1}^{n} X_i = r\right\}$$

jako czas oczekiwania na pierwsze r sukcesów w schemacie Bernoulliego. Tak zdefiniowana zmienna ma rozkład Pascala (zwany też ujemnym rozkładem dwumianowym). Oznaczamy to jako $T \sim \mathcal{NB}(r,p)$.

[1] To jest naprawdę proste.

Powyższy rozkład można interpretować następująco. Student nie zdaje kursu[2] z prawdopodobieństwem p. Zdarzenie $\{T = t\}$ opisuje, że student w minimalnym czasie t nie zdaje kursu r razy.

By móc wyprowadzić ważne twierdzenie i funkcję wagi dla rozkładu Pascala, musimy wprowadzić rozkład geometryczny.

Definicja (Rozkład geometryczny). Dana jest zmienna losowa T

$$T = \min\{n \geqslant 1 : X_n = 1\}$$

jako czas oczekiwania na pierwszy sukces w schemacie Bernoulliego. Tak zdefiniowana zmienna ma rozkład geometryczny. Oznaczamy to jako $T \sim \mathcal{G}(p)$.

Obliczmy funkcję masy prawdopodobieństwa.

$$\begin{aligned}
\mathbb{P}[T = t] &= \mathbb{P}\left[\{X_1 = 0, X_2 = 0, \ldots, X_t = 1\}\right] \\
&= \prod_{i=1}^{t-1} \mathbb{P}[X_i = 0] \cdot \mathbb{P}[X_t = 1] && \text{z niezależności zmiennych} \\
&= (1-p)^{t-1} p
\end{aligned}$$

Zachodzi przydatne twierdzenie:

Twierdzenie. Niech $\{T_i\}_{i \in \mathbb{N}}$ będzie ciągiem niezależnych zmiennych losowych takich, że $T_i \sim \mathcal{G}(p)$. Niech $T = \sum_{i=1}^{r} T_i$. Wtedy $T \sim \mathcal{NB}(r, p)$.

Dowód. Interpretacja tego jest dość prosta. Bierzemy nasze niezależne zmienne i każda z tych zmiennych niezależnie zlicza każdy pierwszy występujący sukces. Dodane zmienne losowe dają nam totalny czas oczekiwania na r sukcesów.

Bardziej rygorystyczny dowód wygląda następująco. Niech $\{X_i\}_{i \in \mathbb{N}}$ będą ciągiem niezależnych zmiennych losowych o rozkładzie zero-jedynkowym z prawdopodobieństwem sukcesu równym p. Niech $T_i = \min\{n : X_n = 1\} \sim \mathcal{G}(p)$ oraz $T = \sum_{i=1}^{r} T_i$. Obliczmy zdarzenie $\{T = t\}$. Mamy, że $\{T_1 = k_1, T_2 = k_2, \ldots, T_r = k_r\}$, gdzie $k_1 + k_2 + \ldots + k_r = t$. Bez straty ogólności możemy założyć, że $1 \leqslant k_1 \leqslant k_2 \leqslant \ldots \leqslant k_r \leqslant r$. Ponieważ zmienne T_i są niezależne, to nasze zdarzenie jest równe $\bigcap_{i=1}^{r}\{T_i = k_i\}$. Stąd mamy, że $\bigcap_{i=1}^{r}\{X_{k_i} = 1\}$. Czyli z niezależności $\{X_{k_1} = 1, X_{k_2} = 1, \ldots, X_{k_r} = 1\}$. Wobec tego $X_i = 0$ dla $i \notin \{k_1, k_2, \ldots, k_r\}$. Czyli mamy zdarzenie równoważne:

$$\{X_{k_1} = 1, X_{k_2} = 1, \ldots, X_{k_r} = 1\} = \{X_{k_1} + X_{k_2} + \ldots + X_{k_r} = r\} = \left\{\sum_{i=1}^{n} X_i = r\right\}$$

gdzie $n = \max\{k_1, k_2, \ldots, k_r\}$ i w końcu

$$T = \sum_{i=1}^{r} T_i = \min\left\{n \geqslant 1 : \sum_{i=1}^{n} X_i = r\right\}$$

$\square$

To twierdzenie pozwala nam na wyznaczenie funkcji masy prawdopodobieństwa rozkładu Pascala w bardzo szybki sposób. Ustalmy r i t, przy czym $t \geqslant r$. Wtedy mamy:

$$\mathbb{P}[T = t] = \mathbb{P}\left[\sum_{i=1}^{r} T_i = t\right]$$

[2] Sukces w prawdopodobieństwie interpretuje się jako zdarzenie, które według nas ma ciekawe lub pożądane własności. Niezdanie kursu również bywa obiektem badań probabilistyki.

Równanie $\sum_{i=1}^{r} T_i = t$ można policzyć na $\binom{t+r-1}{r-1}$ sposobów, stąd

$$= \binom{t+r-1}{r-1} \mathbb{P}[T_1 = k_1, T_2 = k_2, \ldots, T_r = k_r]$$

dla ustalonych $k_1, k_2, \ldots, k_r$ takich, że $k_1 + k_2 + \ldots + k_r = t$. Z niezależności mamy

$$\begin{aligned}
&= \binom{t+r-1}{r-1} \prod_{i=1}^{r} \mathbb{P}[T_i = k_i] \\
&= \binom{t+r-1}{r-1} \prod_{i=1}^{r} \left((1-p)^{k_i - 1} p\right) \\
&= \binom{t+r-1}{r-1} (1-p)^{\sum_{i=1}^{r} k_i - r} p^r \\
&= \binom{t+r-1}{r-1} (1-p)^{t-r} p^r
\end{aligned}$$

Rozwiązanie (b).

Niech $T \sim \mathcal{NB}(m,p)$. Ponieważ zmienna T jest sumą niezależnych zmiennych takich, że $T_i \sim \mathcal{G}(p)$ dla każdego $i \in \{1, 2, \ldots, m\}$, dostajemy:

$$\mathbb{E}T = \mathbb{E}\left[\sum_{i=1}^{m} T_i\right] = \sum_{i=1}^{m} \mathbb{E}T_i = m\frac{1}{p}$$

Podobnie wariancja:

$$\mathrm{Var}T = \mathrm{Var}\left[\sum_{i=1}^{m} T_i\right] = \sum_{i=1}^{m} \mathrm{Var}T_i = m\frac{1-p}{p^2}$$

Metoda.

Generalnie należy skorzystać ze znanej sumy:

$$\frac{1}{1-x} = \sum_{i=0}^{\infty} x^i$$

gdzie $|x| < 1$. Na tej sumie można wykonywać pochodne i przesuwać szereg przez domnożenie do niego x. W taki sposób zostały wyprowadzone „magiczne” równości w rozwiązaniu.

Rozwiązanie (c).

Rozkład opisany na tamtej liście był rozkładem geometrycznym (por. definicję zmiennej losowej i jej funkcję masy prawdopodobieństwa). Obliczmy, i tak już wcześniej używaną, wartość oczekiwaną oraz wariancję dla tego rozkładu. Wobec tego niech $\{X_i\}_{i \in \mathbb{N}}$ będzie schematem Bernoulliego, gdzie $X_i \sim \mathcal{O}(p)$ dla każdego $i \in \mathbb{N}$. Niech $T = \min\{n : X_n = 1\} \sim \mathcal{G}(p)$. Przypomnijmy, że

$$\mathbb{P}[T = t] = (1-p)^{t-1} p$$

Obliczmy wartość oczekiwaną i drugi moment zmiennej losowej X:

$$\begin{aligned}
\mathbb{E}X &= \sum_{i=1}^{\infty} i(1-p)^{i-1} p \\
&= p \sum_{i=1}^{\infty} i(1-p)^{i-1}
\end{aligned}$$

Korzystając z równości $\frac{1}{(1-x)^2} = \sum_{i=1}^{\infty} i x^{i-1}$:

$$\begin{aligned}
&= p \frac{1}{(1-(1-p))^2} \\
&= \frac{1}{p}
\end{aligned}$$

Teraz drugi moment zmiennej:

$$
\begin{aligned}
\mathbb{E}X^2 &= \sum_{i=1}^{\infty} i^2(1-p)^{i-1}p \\
&= p\sum_{i=1}^{\infty} i^2(1-p)^{i-1}
\end{aligned}
$$

Korzystając z równości $-\frac{x+1}{(x-1)^3} = \sum_{i=1}^{\infty} i^2x^{i-1}$:

$$
= -p\frac{1-p+1}{(1-p-1)^3} = p\frac{2-p}{p^3}
$$

Stąd łatwo policzyć wariancję:

$$
\begin{aligned}
\mathrm{Var}X &= \mathbb{E}X^2 - \mathbb{E}^2X \\
&= p\frac{2-p}{p^3} - \frac{1}{p^2} \\
&= \frac{2p-p^2-p}{p^3} \\
&= \frac{p-p^2}{p^3} = \frac{1-p}{p^2}
\end{aligned}
$$

Zadanie 7 (Próba ethernetowa)**.**

a) Rozważamy n urządzeń, które starają się o dostęp do wspólnego zasobu. W tym celu każde urządzenie niezależnie od pozostałych losuje liczbę 0 lub 1, przy czym prawdopodobieństwo wylosowania 1 wynosi p. Dostęp do zasobu uzyska to urządzenie, które jako jedyne wylosuje liczbę 1. Wyznacz taką wartość p, aby prawdopodobieństwo sukcesu w jednej próbie było największe.

b) Powtarzamy próbę ethernetową z optymalnym parametrem p aż do osiągnięcia pierwszego sukcesu. Wyznacz wartość oczekiwaną i wariancję liczby powtórzeń.

Metoda.
Warto znać taką definicję zmiennej losowej o rozkładzie dwumianowym.

Definicja (Rozkład dwumianowy). Ustalmy $n \geqslant 1$. Niech $\{X_i\}_{i\in\mathbb{N}}$ będą ciągiem niezależnych zmiennych losowych o tym samym rozkładzie zero-jedynkowym z prawdopodobieństwem sukcesu równym p (oznaczamy $X_i \sim \mathcal{O}(p)$). Definiujemy zmienną

$$
X = \sum_{i=1}^{n} X_i
$$

która opisuje ilość sukcesów w takim ciągu. Taka zmienna jest rozkładu dwumianowego. Oznaczamy ją jako $X \sim \mathcal{B}(n,p)$.

Wartość oczekiwana i wariancja (równe kolejno $\mathbb{E}X = np$ oraz $\mathrm{Var}X = np(1-p)$), przy takiej definicji zmiennej są bajecznie proste do wyprowadzenia, pozostawiam to jako łatwe ćwiczenie. Głównie podczas wyprowadzania należy pamiętać o tym, że wartości oczekiwane i wariancje dla zmiennych losowych niezależnych są liniowe (tzn. $\mathbb{E}(X+Y) = \mathbb{E}X + \mathbb{E}Y$ oraz $\mathrm{Var}(X+Y) = \mathrm{Var}X + \mathrm{Var}Y$).

Rozwiązanie (a).
Na próbę ethernetową można spojrzeć jak na schemat Bernoulliego $\{K_i\}_{i\in\mathbb{N}}$, gdzie $K_i \sim \mathcal{O}(p)$ dla każdego $i \in \mathbb{N}$, gdzie $\{K_i = 1\}$ oznacza, że i-te urządzenie wylosowało 1. Niech $X = \sum_{i=1}^{n} K_i$ będzie zmienną losową, która

przypisuje liczbę urządzeń, które wylosowały 1. Jak widać $X \sim \mathcal{B}(n,p)$. Interesującym nas zdarzeniem jest $\{X = 1\}$, czyli tylko jedno urządzenie wylosowało 1, a to jest równoważne z uzyskaniem dostępu. Wobec tego

$$\mathbb{P}[X = 1] = \binom{n}{1} p^1 (1-p)^{n-1} = np(1-p)^{n-1}$$

Niech $f(p) = np(1-p)^{n-1}$. Obliczmy punkty stacjonarne tej funkcji:

$$\begin{aligned} f'(p) &= n\left((1-p)^{n-1} - p(n-1)(1-p)^{n-2}\right) \\ &= n(1-p)^{n-2}\,(1-np) = 0 \end{aligned}$$

Mamy stąd $p = 1$ lub $p = \frac{1}{n}$. Podstawiając $p = 1$ otrzymujemy zdarzenie niemożliwe. Wobec tego $p = \frac{1}{n}$ jest punktem stacjonarnym dla którego $f''(p) \leqslant 0$, czyli jest to maksimum $f(p)$.

Rozwiązanie (b).
Niech $\{X_i\}_{i\in\mathbb{N}}$ będzie ciągiem niezależnych zmiennych losowych o tym samym rozkładzie $X_i \sim \mathcal{O}\left(\frac{1}{n}\right)$. Powyższe zmienne stanowią niezależne próby ethernetowe z optymalnym prawdopodobieństwem równym $\frac{1}{n}$. Niech $T = \min\{n \geqslant 1 : X_n = 1\}$ będzie zmienną losową czasu oczekiwania na pierwszy sukces próby ethernetowej. Oczywiście $T \sim \mathcal{G}\left(\frac{1}{n}\right)$. Wobec tego mamy:

$$\begin{aligned} \mathbb{E}T &= \frac{1}{\frac{1}{n}} = n \\ \mathrm{Var}T &= \frac{1-\frac{1}{n}}{\frac{1}{n^2}} = n^2 - n \end{aligned}$$

Zadanie 8. Niech X będzie losowym punktem z koła jednostkowego. Pokaż, że $\mathbb{E}|X| = \frac{2}{3}$. Wyznacz wariancję zmiennej losowej $|X|$.

Metoda.
Zazwyczaj, mając do czynienia z losowaniem punktów, wprowadza się tzw. wektor losowy, czyli wielowymiarową funkcję mierzalną w danej przestrzeni probabilistycznej. Prawdopodobne, że autor tego zadania chciał zataić ten fakt przed nami, by nie siać w paniki wśród Czytelników.

Ogólnie podejście do takich zmiennych sprowadza się do całek wielokrotnych lub sum, w zależności od rodzaju zmiennej. Wartość oczekiwaną, odpowiednie momenty zmiennych, wariancje liczy się z przekształcenia zmiennych losowych $g(X,Y)$, która przeprowadza wektor losowy w liczby rzeczywiste.

Na to zadanie można spojrzeć bardziej formalnie, myśląc o losowaniu punktu jako wektorze losowym (X,Y), które to zmienne są rozkładu jednostajnego ciągłego, które zwykło, przynajmniej mi, oznaczać jako $(X,Y) \sim \mathcal{U}(D)$, gdzie D to interesujący nas zbiór mierzalny (uogólnienie przedziału dla przypadku jednowymiarowego). Natomiast liczymy wartości oczekiwane dla przekształceń $g(X,Y) = \sqrt{X^2+Y^2}$ oraz dla $h(X,Y) = X^2 + Y^2$.

Również to zadanie można wyliczyć bez takich [ucięte — przyp. red.] narzędzi jak wyżej przytoczone. Można wtedy posilić się interpretacją naszego modelu, ułożyć dystrybuantę dla zmiennej $|X|$ a następnie znaleźć funkcję gęstości prawdopodobieństwa dla tej zmiennej.

Rozwiązanie.
Niech X będzie zmienną losową o rozkładzie jednostajnym ciągłym na kole jednostkowym K. Mamy funkcję gęstości dwóch zmiennych:

$$f_X(x,y) = \begin{cases} 0 & x^2+y^2 > 1 \\ \frac{1}{\pi} & x^2+y^2 \leqslant 1 \end{cases}$$

Obliczmy wartość oczekiwaną $\mathbb{E}|X|$ oraz wariancję $\mathrm{Var}|X|$:

$$\begin{aligned} \mathbb{E}|X| &= \iint_K \sqrt{x^2+y^2}\, f_X(x,y)\,dxdy \\ &= \frac{1}{\pi}\iint_K \sqrt{x^2+y^2}\,dxdy \end{aligned}$$

Wprowadzając współrzędne biegunowe:

$$\begin{aligned}
&= \frac{1}{\pi}\int_0^{2\pi} d\varphi \int_0^1 r^2 dr \\
&= \frac{1}{\pi}\cdot 2\pi \cdot \left[\frac{r^3}{3}\right]_0^1 \\
&= \frac{1}{\pi}\cdot 2\pi \cdot \frac{1}{3} \\
&= \frac{2}{3}
\end{aligned}$$

Obliczmy pomocniczo $\mathbb{E}|X|^2$:

$$\begin{aligned}
\mathbb{E}|X|^2 &= \iint_K (x^2+y^2) f_X(x,y)dxdy \\
&= \frac{1}{\pi}\iint_K (x^2+y^2)dxdy \\
&= \frac{1}{\pi}\int_0^{2\pi} d\varphi \int_0^1 r^3 dr \\
&= \frac{1}{\pi}\cdot 2\pi \cdot \left[\frac{r^4}{4}\right]_0^1 \\
&= \frac{1}{\pi}\cdot 2\pi \cdot \frac{1}{4} \\
&= \frac{1}{2}
\end{aligned}$$

Czyli:

$$\begin{aligned}
\mathrm{Var}|X| &= \mathbb{E}|X|^2 - \mathbb{E}^2|X| \\
&= \frac{1}{2} - \frac{4}{9} \\
&= \frac{1}{18}
\end{aligned}$$

Zadanie 9. Niech $A \subseteq \Omega$. Niech $X = \mathbf{1}_A$. Pokaż, że $\mathrm{Var}X = \mathbb{P}[A](1-\mathbb{P}[A])$. Jaka może być maksymalna wartość wariancji dla takiej zmiennej losowej?

Rozwiązanie.
Zmienna $X = \mathbf{1}_A$ jest rozkładu zero-jedynkowego z prawdopodobieństwem sukcesu $\mathbb{P}[A]$. Wartość oczekiwana wynosi $\mathbb{E}X = \mathbb{P}[A]$ oraz $\mathbb{E}X^2 = \mathbb{P}[A]$ Wobec tego wariancja wynosi $\mathrm{Var}X = \mathbb{P}[A] - \mathbb{P}^2[A] = \mathbb{P}[A](1-\mathbb{P}[A])$.

Zmaksymalizujmy sobie funkcję $f(x) = x(1-x)$. Mamy $f'(x) = 1-2x$, czyli punktem stacjonarnym tej funkcji jest $x = \frac{1}{2}$. Ponieważ $f''\left(\frac{1}{2}\right) < 0$, jest to maksimum tej funkcji.

Zadanie 10. Rozważmy ciąg $(A_n)_{n\geqslant 1}$ niezależnych zdarzeń takich, że $\mathbb{P}[A_n] = p$ dla każdego n. Niech r będzie ustalone. Definiujemy

$$L = \min\left\{k : \sum_{i=1}^{k} \mathbf{1}_{A_i} = r\right\}$$

Jaki rozkład ma zmienna L? Wyznacz wartość oczekiwaną i wariancję zmiennej L.

Rozwiązanie.
Niech $X_i = \mathbf{1}_{A_i}$. Zmienna X_i jest rozkładu zero-jedynkowego z prawdopodobieństwem sukcesu $\mathbb{P}[A_i] = p$. Wobec tego zmienna L jest rozkładu Pascala o parametrach r i p. Wyznaczanie wartości oczekiwanej i wariancji pozostawiam jako ćwiczenie.

Zadanie 11. Niech X i Y będą niezależnymi zmiennymi losowymi. Pokaż, że $\mathrm{Var}[X-Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$.

Metoda.
Przypomnijmy sobie, czym jest niezależność zmiennych losowych. Równoważnie ten fakt formułuje się tak:

1. Zmienne losowe X i Y są niezależne
2. $\mathbb{P}[X \leqslant x \wedge Y \leqslant y] = \mathbb{P}[X \leqslant x] \cdot \mathbb{P}[Y \leqslant y]$
3. $F_{(X,Y)}(x,y) = F_X(x) \cdot F_Y(y)$
4. $f_{(X,Y)}(x,y) = f_X(x) \cdot f_Y(y)$

Z tego twierdzenia (i definicji) skorzystamy w rozwiązaniu.

Również będziemy korzystać ze wzorów $\mathrm{Var}(aX+b) = a^2\mathrm{Var}X$ oraz $\mathrm{Var}(X \pm Y) = \mathrm{Var}X + \mathrm{Var}Y \pm 2\mathrm{Cov}(X,Y)$, gdzie $\mathrm{Cov}(X,Y) = \mathbb{E}XY - (\mathbb{E}X)(\mathbb{E}Y)$ to współczynnik kowariancji. Jeśli zmienne są niezależne, to właśnie $\mathrm{Cov}(X,Y) = 0$ i stąd mamy magiczny wzór dla niezależnych zmiennych $\mathrm{Var}(X \pm Y) = \mathrm{Var}X + \mathrm{Var}Y$.

Rozwiązanie.
Udowodnijmy mocniejsze twierdzenie, to znaczy, jeśli X i Y są zmiennymi losowymi niezależnymi, to $\mathbb{E}XY = \mathbb{E}X \cdot \mathbb{E}Y$. Jednakże dowód nie jest obowiązkowy do znania w całości. Dla mniej dociekliwego Czytelnika wystarczy pominąć następny akapit.

Załóżmy, że X i Y są zmiennymi niezależnymi, to znaczy dla wektora losowego (X,Y) zachodzi $F_{(X,Y)}(x,y) = F_X(x)F_Y(y)$. Wektorowi losowemu (X,Y) określamy funkcję gęstości $f_{(X,Y)}(x,y)$ i wtedy z niezależności mamy $f_{(X,Y)}(x,y) = f_X(x)f_Y(y)$. Teraz rozpatrujemy dwa przypadki: dyskretny i ciągły. Dla zmiennej dyskretnej:

$$\mathbb{E}XY = \sum_x \sum_y xy f_{(X,Y)}(x,y) = \sum_x \sum_y xy f_X(x) f_Y(y) = \left(\sum_x x f_X(x)\right)\left(\sum_y y f_Y(Y)\right) = \mathbb{E}X \cdot \mathbb{E}Y$$

Dowód jest analogiczny dla zmiennej ciągłej z twierdzenia Fubiniego (przez znaki sumy zastępujemy znakami całki podwójnej po $\mathbb{R}^2$).

Wobec tego współczynnik kowariancji jest równy zero, dlatego wariancja jest liniowa. Ostatecznie, $\mathrm{Var}[X-Y] = \mathrm{Var}[X] + \mathrm{Var}[-Y] = \mathrm{Var}[X] + (-1)^2\mathrm{Var}[Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$. $\square$

Druga lista „powtórkowa"

Zadanie 12. Niech X, Y będą niezależnymi zmiennymi losowymi o rozkładzie jednostajnym na odcinku $(0,2)$. Oblicz $\mathbb{E}(X+Y)$, $\mathbb{E}(X-Y)$, $\mathrm{Var}(X+Y)$, $\mathrm{Var}(X-Y)$. Które z wyników mogą się zmienić, gdy zrezygnujemy z założenia o niezależności?

Metoda.
Generalnie lista i takie zadania sprawdzają, czy Czytelnik potrafi zapamiętać wybrane rozkłady prawdopodobieństwa. Z przyjemnością mogę wyprowadzić rozkład jednostajny ciągły jednakże podaruję tę przyjemność Czytelnikowi. Natomiast warto zapamiętać poniższy wzór na funkcję gęstości prawdopodobieństwa zmiennej o rozkładzie jednostajnym ciągłym na odcinku $[a,b]$:

$$f_X(x) = \begin{cases} 0 & x \notin [a,b] \\ \frac{1}{b-a} & x \in [a,b] \end{cases}$$

Powyższa konstrukcja również przez analogię do „objętości" przedziału uogólnia się na wektory losowe (por. lista 8., zadanie 8.).

Wyprowadźmy natomiast wartość oczekiwaną i wariancję takiej zmiennej. Obliczmy więc wartość oczekiwaną i drugi moment zmiennej X:

$$\begin{aligned}
\mathbb{E}X &= \int_a^b x\frac{1}{b-a}dx \\
&= \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b \\
&= \frac{1}{b-a}\cdot\frac{b^2-a^2}{2} \\
&= \frac{b^2-a^2}{2(b-a)} \\
&= \frac{b+a}{2}
\end{aligned}$$

Zaś drugi moment:

$$\begin{aligned}
\mathbb{E}X^2 &= \int_a^b x^2\frac{1}{b-a}dx \\
&= \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b \\
&= \frac{b^3-a^3}{3(b-a)} \\
&= \frac{b^2+ba+a^2}{3}
\end{aligned}$$

Stąd można policzyć wariancję:

$$\begin{aligned}
\mathrm{Var}X &= \mathbb{E}X^2 - \mathbb{E}^2X \\
&= \frac{b^2+ba+a^2}{3} - \frac{(b+a)^2}{4} \\
&= \frac{4(b^2+ba+a^2)-3(b+a)^2}{12} \\
&= \frac{4b^2+4ba+4a^2-3b^2-6ba-3a^2}{12} \\
&= \frac{b^2-2ba+a^2}{12} \\
&= \frac{(b-a)^2}{12}
\end{aligned}$$

Co do niezależności zmiennych: dla zmiennych niezależnych wartość oczekiwana i wariancja są liniowe, tzn. wartość oczekiwana sumy zmiennych jest sumą wartości oczekiwanych zmiennych, analogicznie dla wariancji.

Rozwiązanie.
Liczymy:

$$\begin{aligned}
\mathbb{E}(X+Y) &= \mathbb{E}X + \mathbb{E}Y && \text{ponieważ zmienne są niezależne} \\
&= 1+1=2 \\
\mathbb{E}(X-Y) &= 0 && \text{ponieważ mają ten sam rozkład} \\
\mathrm{Var}(X+Y) &= \mathrm{Var}X + \mathrm{Var}Y && \text{ponieważ zmienne są niezależne} \\
&= \frac{2^2}{12} + \frac{2^2}{12} = \frac{2}{3} \\
\mathrm{Var}(X-Y) &= \mathrm{Var}X + \mathrm{Var}(-Y) \\
&= \mathrm{Var}X + (-1)^2\mathrm{Var}(Y) && \text{korzystając z } \mathrm{Var}(aX+b) = a^2\mathrm{Var}X \\
&= \frac{2}{3}
\end{aligned}$$

Wartość wariancji może ulec zmianie, ponieważ kowariancja jest niezerowa i stąd wariancja jest nieliniowa przy założeniu, że zmienne są od siebie zależne.

Zadanie 13. Niech X będzie zmienną losową o rozkładzie jednostajnym na odcinku $[0,1]$. Wyznacz dystrybuantę oraz gęstość zmiennej losowej $Y=\sqrt{X}$.

Metoda.
Z wykładu wiemy, że:

Twierdzenie. Jeśli X jest zmienną losową ciągłą o gęstości, która koncentruje się na przedziale (a,b), oraz $g(x)$ jest funkcją ściśle monotoniczną, różniczkowalną o pochodnej $g'(x)\neq 0$ dla $x\in(a,b)$ oraz dla funkcji g istnieje funkcja odwrotna, to gęstość zmiennej $g(X)$ zadana jest wzorem:

$$f_{g(X)}(y) = \begin{cases} f_X(g^{-1}(y))\cdot\left|(g^{-1}(y))'\right| & y\in(g(a^+),g(b^-)) \\ 0 & y\notin(g(a^+),g(b^-)) \end{cases}$$

gdzie $g(x_0^+) = \lim_{x\to x_0^+} g(x)$ oraz $g(x_0^-) = \lim_{x\to x_0^-} g(x)$.

To twierdzenie jest bardzo przydatne przy przekształceniu zmiennej losowej. Analogiczne twierdzenie uzyskujemy dla wektorów losowych, tylko zamiast pochodnej funkcji odwrotnej mamy jakobian przekształcenia g^{-1} w punkcie y.

Rozwiązanie.
Niech $X\sim\mathcal{U}[0,1]$. Wobec tego zmienna X ma następującą funkcję gęstości prawdopodobieństwa:

$$f_X(x) = \begin{cases} 0 & x\notin[0,1] \\ 1 & x\in[0,1] \end{cases}$$

Niech $g(x)=\sqrt{x}$. Ze wzoru wyznaczamy funkcję gęstości:

$$f_Y(y) = \begin{cases} f_X(y^2)\cdot|2y| = 2y & y\in[0,1] \\ 0 & y\notin[0,1] \end{cases}$$

Zadanie 14. Wybieramy losowo punkt $P\in\langle 0,1\rangle^2$. Zakładamy, że dla dowolnego (mierzalnego) podzbioru $A\subseteq\langle 0,1\rangle^2$ mamy $\mathbb{P}[P\in A]=\mathrm{vol}(A)$. Dla $\varepsilon>0$ kładziemy $A_\varepsilon=\{(x,y)\in\langle 0,1\rangle^2 : |x-y|<\varepsilon\}$. Wyznacz $\mathbb{P}[P\in A_\varepsilon]$.

Rozwiązanie (S. Wróbel, J. Nigiel).
Liczymy pole A_ε. Kreśląc nierówności $y<x+\varepsilon$ oraz $y>x-\varepsilon$, które ograniczone są kwadratem jednostkowym $\langle 0,1\rangle^2$, otrzymujemy, że pole jest równe $1-|A_\varepsilon^c|$, czyli $1-2\cdot\frac{(1-\varepsilon)^2}{2}=\varepsilon(2-\varepsilon)$.

Zadanie 15. Gęstość pewnej zmiennej losowej X zadana jest wzorem:

$$f_X(x) = \begin{cases} 3x^2 & x \in (0,1) \\ 0 & x \notin (0,1) \end{cases}$$

Wyznacz dystrybuantę F_X zmiennej losowej X. Oblicz i zaznacz na wykresach gęstości i dystrybuanty $\mathbb{P}\left[X \geqslant \frac{1}{2}\right]$.

Metoda.
Dystrybuanta dla zmiennej ciągłej to:

$$F_X(x) = \mathbb{P}\left[X < x\right] = \int_{-\infty}^{x} f_X(t)dt$$

Interpretacja takiego zdarzenia głownie wynika na wzorze wyżej podczas rysowania. Generalnie rozwiązania nie będą zawierać wykresów, ponieważ samo sporządzenie wykresu i zaznaczenie jest dość trywialne.

Rozwiązanie.
Dystrybuanta dla $x \leqslant 0$ będzie przyjmować wartość $F_X(x) = 0$, ponieważ $x \notin (0,1)$. Policzmy dystrybuantę dla $0 < x < 1$:

$$F_X(x) = \int_0^x 3t^2 dt = \left[t^3\right]_0^x = x^3$$

Natomiast dla $x \geqslant 1$ mamy $F_X(x) = 1$, ponieważ:

$$F_X(x) = \int_0^1 3t^2 dt + \int_1^x 0 dt = 1$$

Mamy $\mathbb{P}\left[X \geqslant \frac{1}{2}\right] = 1 - \mathbb{P}\left[X < \frac{1}{2}\right] = 1 - F_X\left(\frac{1}{2}\right) = \int_0^1 3t^2dt - \int_0^{\frac{1}{2}} 3t^2dt = \int_{\frac{1}{2}}^1 3t^2dt.$

Zadanie 16. Dystrybuanta pewnej zmiennej losowej X zadana jest wzorem:

$$F_X(x) = \begin{cases} 0 & x \leqslant 2 \\ (x-2)^2 & x \in (2,3) \\ 1 & x \geqslant 3 \end{cases}$$

Wyznacz gęstość f_X zmiennej losowej X. Oblicz i zaznacz na wykresach gęstości i dystrybuanty $\mathbb{P}\left[X \geqslant \frac{5}{2}\right]$.

Rozwiązanie.
Należy policzyć pochodną.

$$f_X(x) = \begin{cases} 0 & x \leqslant 2 \wedge x \geqslant 3 \\ 2(x-2) & x \in (2,3) \end{cases}$$

Mamy $\mathbb{P}\left[X \geqslant \frac{5}{2}\right] = 1 - \mathbb{P}\left[X < \frac{5}{2}\right] = 1 - F_X\left(\frac{5}{2}\right) = \int_{\frac{5}{2}}^3 2(t-2)dt.$

Zadanie 17. Niech X będzie zmienną losową przyjmująca wartości ze zbioru $\mathbb{N} \setminus \{0\}$ taką, że $\mathbb{P}[X = k] = \frac{c}{3^k}$.

a) Wyznacz stałą c.

b) Jaka jest wartość oczekiwana zmiennej losowej X?

Rozwiązanie.
Suma funkcji wag prawdopdobieństwa po wszystkich $k \in \mathbb{N} \setminus \{0\}$ musi dać nam 1, czyli:

$$\sum_{i=1}^{\infty} \frac{c}{3^k} = c\frac{1}{1-\frac{1}{3}} = 1$$

Stąd $c = \frac{2}{3}$. Niech $Y \sim \mathcal{G}\left(\frac{2}{3}\right)$. Mamy $\mathbb{E}X = \frac{1}{3}\mathbb{E}Y$, ponieważ $\mathbb{P}[X = k] = \frac{c}{3^k} = \frac{2}{3}\left(\frac{1}{3}\right)^k = \frac{1}{3}\left(\frac{2}{3}\left(\frac{1}{3}\right)^{k-1}\right) = \frac{1}{3}\mathbb{P}[Y = k]$.
Stąd wartość oczekiwana wynosi $\mathbb{E}X = \frac{1}{3}\mathbb{E}Y = \frac{1}{3} \cdot \frac{3}{2} = \frac{1}{2}$.

Zadanie 18. Niech X będzie zmienną losową przyjmującą wartości ze zbioru $\{1, 2, \ldots, 2n\}$ taką, że $\mathbb{P}[X = k] = \frac{1}{2n}$, gdzie $k = 1, 2, \ldots, 2n$.

a) Oblicz $\mathbb{E}X$.

b) Oblicz $\mathbb{P}[X > n]$.

Rozwiązanie.
Liczymy wartość oczekiwaną:

$$\begin{aligned}\mathbb{E}X &= \sum_{i=1}^{2n} \frac{i}{2n} \\ &= \frac{1}{2n} \cdot \frac{(2n+1)2n}{2} \\ &= \frac{2n+1}{2}\end{aligned}$$

Policzmy $\mathbb{P}[X > n]$:

$$\begin{aligned}\mathbb{P}[X > n] &= 1 - \mathbb{P}[X \leqslant n] \\ &= 1 - \sum_{i=1}^{n} \frac{1}{2n} \\ &= \frac{1}{2}\end{aligned}$$

Zadanie 19. Zmienna losowa X ma następujący rozkład prawdopodobieństwa:

$$\begin{aligned}\mathbb{P}[X = 0] = \mathbb{P}\left[X = \frac{\pi}{2}\right] &= \frac{1}{4} \\ \mathbb{P}[X = 2\pi] &= \frac{1}{2}\end{aligned}$$

Znajdź rozkład prawdopodobieństwa zmiennej losowej $Y = \cos X$.

Metoda.
Metoda sprowadza się do obliczenia odpowiednich zdarzeń po nałożeniu funkcji cos i policzenia dystrybuanty. Ponieważ funkcja cos nie jest różnowartościowa, trzeba będzie wliczyć sklejone wartości do jednego zdarzenia.

Rozwiązanie.
Widzimy, że:

$$\begin{aligned}\mathbb{P}[\cos X = 1] = \mathbb{P}\left[X = 2\pi\right] + \mathbb{P}[X = 0] &= \frac{3}{4} \\ \mathbb{P}[\cos X = 0] = \mathbb{P}\left[X = \frac{\pi}{2}\right] &= \frac{1}{2}\end{aligned}$$

Czyli nasza dystrybuanta wygląda tak:

$$F_{\cos X}(y) = \begin{cases} 0 & x < 0 \\ \frac{3}{4} & 0 \leqslant x < 1 \\ 1 & x \geqslant 1 \end{cases}$$

Zadanie 20. Jaś rzuca niesymetryczną monetą (prawdopodobiestwo wypadnięcia orła wynosi $\frac{1}{3}$) aż do momentu wypadnięcia pierwszego orła. Następnie Małgosia rzuca tą samą monetą do momentu wypadnięcia pierwszej reszki. Niech Z będzie łączną liczbą wszystkich rzutów. Wyznacz $\mathbb{E}Z$ oraz $\mathrm{Var}Z$.

Rozwiązanie.
Niech $T_J \sim \mathcal{G}\left(\frac{1}{3}\right)$ będzie zmienną losową opisującą czas oczekiwania Jasia na wypadnięcie pierwszego orła. Analogicznie dla Małgosi definiujemy $T_M \sim \mathcal{G}\left(\frac{2}{3}\right)$. Oczywiście T_J i T_M są niezależne — Jaś nie wpływa na rzuty Małgosi i odwrotnie (chyba że są zakochani). Wtedy $Z = T_J + T_M$ i wyznaczamy wartości:

$$\begin{aligned}
\mathbb{E}Z &= \mathbb{E}T_J + \mathbb{E}T_m \\
&= 3 + \frac{3}{2} \\
&= \frac{9}{2} \\
\mathrm{Var}Z &= \mathrm{Var}T_J + \mathrm{Var}T_M \\
&= \frac{1-\frac{1}{3}}{\frac{1}{9}} + \frac{1-\frac{2}{3}}{\frac{4}{9}} \\
&= 6 + \frac{3}{4} \\
&= \frac{27}{4}
\end{aligned}$$

Zadanie 21. Załóżmy, że $X_k \sim \mathcal{B}\left(k, \frac{1}{k^2}\right)$ dla $k = 1, 2, 3, \ldots$. Niech $Y_n = X_1 + X_2 + \ldots + X_n$. Wyznacz granicę:

$$\lim_{n\to\infty} \mathbb{E}\left(\frac{Y_n}{n}\right)$$

Metoda.
Przyda nam się twierdzenie z analizy matematycznej (bez dowodu), znane jako twierdzenie Stolza, twierdzenie Stolza-Cesàro, czy *dyskretna reguła de l'Hôpitala.*

Twierdzenie (Stolz, Cesàro). *Niech (a_n) i (b_n) będą ciągami rzeczywistymi, przy czym (b_n) jest ciągiem rosnącym i $b_n \xrightarrow{n\to\infty} \infty$. Wtedy zachodzi następujący fakt:*

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{a_n - a_{n-1}}{b_n - b_{n-1}}$$

o ile granica $\lim_{n\to\infty} \frac{a_n - a_{n-1}}{b_n - b_{n-1}}$ istnieje, to znaczy, może być rozbieżna albo skończona.

Rozwiązanie.
Liczymy granicę:

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}\left(\frac{Y_n}{n}\right) &= \lim_{n\to\infty} \frac{1}{n}\mathbb{E}Y_n \\
&= \lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left(\sum_{k=1}^{n} X_k\right) \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}X_k \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} k\frac{1}{k^2} \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \frac{1}{k} \\
&= \lim_{n\to\infty} \frac{H_n}{n}
\end{aligned}$$

Korzystając z twierdzenia Stolza

$$
\begin{aligned}
&= \lim_{n\to\infty} \frac{H_n - H_{n-1}}{n-(n-1)} \\
&= \lim_{n\to\infty} H_n - H_{n-1} \\
&= \lim_{n\to\infty} \frac{1}{n} = 0
\end{aligned}
$$

Zadanie 22. Niech X będzie zmienną losową o rozkładzie jednostajnym na odcinku $[0,1]$ oraz Y będzie zmienną losową o rozkładzie jednostajnym na odcinku $[2,4]$. Oblicz $\mathbb{E}(X+Y)$.

Rozwiązanie.

$$\mathbb{E}(X+Y) = \mathbb{E}X + \mathbb{E}Y = \frac{1}{2} + \frac{6}{2} = \frac{7}{2}$$