Douczki z Rachunku Prawdopodobieństwa 2023

Dzień pierwszy

Zajęcia prowadzone są w ramach wolontariatu przez członka KPM, który jest studentem. Miej na uwadze to, że może nie być w stanie on odpowiedzieć na wszystkie zadane przez Was pytania.

## 1. Zadania podobne do zadania 1 z egzaminu

**Zadanie 1.**
Zmienne $X, Y$ są niezależne o gęstości $g(x) = \frac{e^x}{e-1} \cdot \mathbf{1}_{[0,1]}(x)$.

a) Wyznacz gęstość zmiennej $Z = X - Y$

b) Wyznacz wartość oczekiwaną i wariancję zmiennej $Z$.

**Zadanie 2.**
Zmienne losowe $X$ i $Y$ są niezależne oraz mają rozkłady geometryczne z parametrami odpowiednio $p$ i $q$. Obliczyć $P(X \leq Y)$ oraz $P(X = Y)$.

**Rozwiązanie:**
Mamy $\{X \leq Y\} = \bigcup_{k=1}^{\infty} \{X = k \leq Y\}$, przy czym zdarzenia po których sumujemy są rozłączne. Zatem

$$P(X \leq Y) = \sum_{k=1}^{\infty} P(X = k \cap k \leq Y) = \sum_{k=1}^{\infty} P(X = k) \cdot P(Y \geq k) =$$

$$= \sum_{k=1}^{\infty} p^{k-1} \cdot (1-p) \cdot \left( \sum_{i=k}^{\infty} q^{i-1}(1-q) \right) = (1-p) \cdot \sum_{k=1}^{\infty} p^{k-1} \cdot \left( (1-p) \cdot \sum_{i=k}^{\infty} q^{i-1} \right) =$$

$$= (1-p) \cdot \sum_{k=1}^{\infty} p^{k-1} \left( (1-q) \cdot q^{k-1} \cdot \frac{1}{1-q} \right) = (1-p) \cdot \sum_{k=1}^{\infty} pq = \frac{1-p}{1-pq}$$

Mamy więc

$$P(X \leq Y) = \frac{1-p}{1-pq} \qquad \text{oraz analogicznie} \qquad P(Y \leq X) = \frac{1-q}{1-pq}$$

Mamy

$$P(X \leq Y) + P(Y \leq X) + P(X = Y) = 1$$

zatem

$$P(X = Y) = \frac{1-p}{1-pq} + \frac{1-q}{1-pq} - 1 = \frac{(1-p)(1-q)}{1-pq}$$

**Zadanie 3.**
Zmienne losowe $X$ i $Y$ są niezależne oraz mają rozkłady wykładnicze z parametrami odpowiednio $\lambda$ i $\mu$. Obliczyć $P(X \leq Y)$ oraz $P(X = Y)$.

**Rozwiązanie:**
Liczymy

$$P(X \le Y) = \int_{x \le y} g_{(X,Y)}(x,y)\ d\lambda_2(x,y) = \int_{x \le y} \mathbf{1}_{(0,\infty)}(x) \cdot \mathbf{1}_{(0,\infty)}(y) \cdot \lambda e^{-\lambda x} \cdot \mu e^{-\mu y}\ d\lambda_2(x,y) =$$

$$= \int_0^\infty \mu e^{-\mu y} \cdot \left( \int_0^y \lambda e^{-\lambda x}\ dx \right)\ dy = \int_0^\infty \mu e^{-\mu y} \cdot \left(1 - e^{-\lambda y}\right)\ dy =$$

$$= \int_0^\infty \mu e^{-\mu y}\ dy - \int_0^\infty \mu e^{-(\lambda+\mu)y}\ dy = 1 - \frac{\mu}{\lambda + \mu} = \frac{\lambda}{\mu + \lambda}$$

Stąd

$$P(X = Y) = P(X \ge Y) + P(Y \ge X) - 1 = 0$$

Można też powiedzieć, że $P(X = Y) = 0$, bo $\int_{\{x=y\}} g_{(X,Y)}(x,y)\ d\lambda_2(x,y) = 0$, bo zbiór $\{x = y\}$ ma miarę zero.

**Zadanie 4 PRACA DOMOWA.**
Zmienne losowe $X$ i $Y$ są niezależne oraz mają rozkłady geometryczne z parametrami odpowiednio $p$ i $q$. Wyznaczyć rozkłady zmiennych $X + Y$ oraz $X - Y$.

**Zadanie 5.**
Niech $X_1, X_2$ są niezależnymi zmiennymi losowymi o rozkładzie jednostajnym na $[0,1]$, czyli $g_1(x) = g_2(x) = 1_{[0,1]}(x)$. Oblicz gęstość $X_1 + X_2$.

**Rozwiązanie:**

$$g_1 * g_2(x) = \int_{\mathbb{R}} 1_{[0,1]}(x-y) \cdot 1_{[0,1]}(y)\, dy = \int_{\mathbb{R}} 1_{[x-1,x]}(y) \cdot 1_{[0,1]}(y)\, dy =$$

$$= |[x-1,x] \cap [0,1]| = \begin{cases} 0 & \text{dla } x < 0, \\ x & \text{dla } 0 \le x < 1, \\ 2 - x & \text{dla } 1 \le x < 2, \\ 0 & \text{dla } x \ge 2. \end{cases}$$

**Zadanie 6.**
Niech $X_1, X_2$ są niezależnymi zmiennymi losowymi o rozkładach $N(m_1, \sigma_1^2)$, $N(m_2, \sigma_2^2)$, czyli

$$g_i(x) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left( -\frac{(x - m_i)^2}{2\sigma_i^2} \right), \quad i = 1, 2.$$

Oblicz gęstość $X_1 + X_2$.

**Rozwiązanie:**
Zmienna $X_1 + X_2$ ma gęstość

$$g_1 * g_2(x) = \frac{1}{2\pi\sigma_1\sigma_2} \int_{\mathbb{R}} \exp\left( -\frac{(x - y - m_1)^2}{2\sigma_1^2} - \frac{(y - m_2)^2}{2\sigma_2^2} \right)\ dy =$$

$$= \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}} \exp\left(-\frac{(x-m_1-m_2)^2}{2(\sigma_1^2+\sigma_2^2)}\right).$$

Dowód ostatniej równości pozostawiamy jako ćwiczenie. Zatem $X_1 + X_2 \sim N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$. I ogólniej, przez indukcję: jeśli $X_1, X_2, \ldots, X_n$ są niezależnymi zmiennymi losowymi o rozkładach $N(m_1, \sigma_1^2)$, $N(m_2, \sigma_2^2)$, $\ldots$, $N(m_n, \sigma_n^2)$, to $X_1 + X_2 + \ldots + X_n$ ma rozkład $N(m_1 + m_2 + \ldots + m_n, \sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2)$.

**Zadanie 7 PRACA DOMOWA.**
Zmienne losowe $X, Y$ są niezależne. $X$ ma gęstość $g(x) = 2x \cdot \mathbf{1}_{[0,1]}(x)$ a $Y$ ma rozkład jednostajny na przedziale $[0,1]$.

a) Wyznacz gęstość zmiennej $Z = X + Y$.

b) Wyznacz wartość oczekiwaną i wariancję zmiennej $Z$.

## 2. Zadania podobne do zadania 2 z egzaminu

**Zadanie 1.**
W $n$ rozróżnialnych urnach rozmieszczono $k$ rozróżnialnych kul. Niech $X$ oznacza liczbę pustych urn. Wyznacz $\mathbb{E}(X)$ oraz $Var(X)$.

**Rozwiązanie:**
Niech

$$X_i = \begin{cases} 1 & \text{gdy w } i\text{-tej próbie nic nie ma} \\ 0 & \text{w przeciwnym przypadku} \end{cases}$$

wówczas $X = X_1 + \ldots + X_n$. Mamy więc

$$\mathbb{E}(X) = \mathbb{E}(X_1 + \ldots + X_n) = \mathbb{E}(X_1) + \ldots + \mathbb{E}(X_n) = n \cdot \mathbb{E}(X_1)$$

gdyż prawdopodobieństwo że $i$-ta urna będzie pusta jest takie samo jak prawdopodobieństwo że $j$-ta urna będzie pusta, co pociąga za sobą, że $\mathbb{E}(X_i) = \mathbb{E}(X_j)$. Policzmy więc $\mathbb{E}(X_1)$. Mamy $\mathbb{E}(X_1) = 1 \cdot P(X_1 = 1) + 0 \cdot P(X_1 = 0) = \left(\frac{n-1}{n}\right)^k$, gdyż wszystkie $k$ kul muszą trafić do pozostałych $n-1$ urn. Stąd $\mathbb{E}X = n \cdot \left(\frac{n-1}{n}\right)^k$. Dalej liczymy $VarX$. Wiemy, że

$$Var(X) = Var(X_1 + \ldots + X_n) = \sum_{i=1}^{n} Var(X_i) + 2 \cdot \sum_{1 \le i < j \le n} Cov(X_i, X_j)$$

Musimy więc policzyć $Var(X_i)$ oraz $Cov(X_i, X_j)$ dla $i \neq j$. Mamy

$$Var(X_i) = \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2 = \mathbb{E}(X_i) - (\mathbb{E}(X_i))^2 = \left(\frac{n-1}{n}\right)^k - \left(\frac{n-1}{n}\right)^{2k}$$

oraz mamy

$$Cov(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i) \cdot \mathbb{E}(X_j) = \mathbb{E}(X_i X_j) - (\mathbb{E}(X_i))^2 =$$

$$= 1 \cdot P(X_i X_j = 1) + 0 \cdot P(X_i X_j = 0) - \left(\frac{n-1}{n}\right)^{2k} =$$

$$= P(X_i = 1, X_j = 1) - \left(\frac{n-1}{n}\right)^{2k} = \left(\frac{n-2}{n}\right)^{k} - \left(\frac{n-1}{n}\right)^{2k}$$

Zatem

$$Var(X) = n \cdot Var(X_1) + \binom{n}{2} \cdot 2 \cdot Cov(X_1, X_2)$$

**Zadanie 2.**
Do punktu badania jakości wody dostarczono 100 próbek. Prawdopodobieństwo tego, że w danej próbce znajdują się bakterie E. coli wynosi $0,1$. W celu zaoszczędzenia na badaniach, próbki podzielono na 10 grup po dziesięć w każdej. Następnie, w obrębie każdej grupy zmieszano wodę i przeprowadzono badanie. Jeśli wynik jest negatywny, wszystkie dziesięć próbek jest wolnych od bakterii. Jeśli natomiast wynik jest pozytywny, przeprowadza się dodatkowe 10 badań, dla każdej próbki z osobna. Niech $X$ oznacza liczbę przeprowadzonych badań. Oblicz $\mathbb{E}X$.

**Rozwiązanie:**
Niech

$$X_i = \begin{cases} 1 & \text{gdy w danej próbie wszystkie próbki są nieskażone} \\ 11 & \text{w przeciwnym przypadku} \end{cases}$$

Wówczas $X = X_1 + \ldots + X_{10}$ oraz $\mathbb{E}X_i = \mathbb{E}X_j$ dla $i \neq j$, więc $\mathbb{E}X = \mathbb{E}X_1 + \ldots + \mathbb{E}X_{10} = 10 \cdot \mathbb{E}X_1$. Mamy $\mathbb{E}X_1 = 1 \cdot P(X_1 = 1) + 11 \cdot P(X_1 = 11) = \left(\frac{9}{10}\right)^{10} + 11 \cdot \left(1 - (\frac{9}{10})^{10}\right) = 11 - 10 \cdot \left(\frac{9}{10}\right)^{10}$.

**Zadanie 3.**
Jest $N$ listów i $N$ zaadresowanych kopert z różnymi adresami. Każdy list odpowiada dokładnie jednemu adresowi i na odwrót. Włożono listy do kopert na chybił trafił, po jednym liście do każdej koperty. Niech $X$ oznacza liczbę listów, które trafiły do właściwych kopert. Obliczyć wartość oczekiwaną i wariancję zmiennej $X$.

**Rozwiązanie:**
Od liczby wszystkich możliwych permutacji odejmujemy te, gdzie co najmniej jeden jest dobry, dodajemy te gdzie co najmniej dwa są dobre, itd. Z zasady włączeń-wyłączeń mamy więc

$$|A_n| = n! - \binom{n}{1} \cdot (n-1)! + \binom{n}{2} \cdot (n-2)! - \ldots = \sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i)!$$

Skąd prawdopodobieństwo wynosi

$$P(A_n) = \frac{1}{n!} \cdot \sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i)!$$

**Zadanie 4.**
Z urny zawierającej 15 kul białych i 5 kul czarnych losujemy 4 kule bez zwracania. Obliczyć wartość oczekiwaną liczby wylosowanych kul białych.

**Rozwiązanie:**
Mamy $X \in \{0,1,2,3,4\}$. Można policzyć że $P(X=i) = \frac{\binom{15}{i}\cdot\binom{5}{4-i}}{\binom{20}{4}}$ i potem wartość oczekiwaną ze wzoru $\mathbb{E}X = \sum_{i=0}^{4} i \cdot P(X=i)$. Ale można też inaczej. Niech

$$X_i = \begin{cases} 1 & \text{gdy } i\text{-ta kula jest biała} \\ 0 & \text{w przeciwnym przypadku} \end{cases}$$

Wówczas jako że $P(X_i = 1)$ oznacza prawdopodobieństwo na to, że $i$-ta kula jest biała, czyli z symetrii również na to, że pierwsza kula jest biała, to zmienne losowe $X_1, X_2, X_3, X_4$ mają taki sam rozkład prawdopodobieństwa, czyli $\mathbb{E}X = \mathbb{E}(X_1+X_2+X_3+X_4) = \mathbb{E}X_1+\mathbb{E}X_2+\mathbb{E}X_3+\mathbb{E}X_4 = 4\cdot\mathbb{E}X_1 = 4\cdot\frac{15}{20} = 3$.

**Zadanie 5.**
Każdy bok i każdą przekątną sześciokąta foremnego malujemy losowo na jeden z trzech kolorów. Wybór każdego koloru jest jednakowo prawdopodobny, kolorowania różnych odcinków są niezależne. Niech $X$ oznacza liczbę jednobarwnych trójkątów o wierzchołkach będących wierzchołkami sześciokąta. Obliczyć wartość oczekiwaną zmiennej $X$.

**Rozwiązanie:**
Wszystkich trójkątów jest łącznie $\binom{6}{3} = \frac{6\cdot5\cdot4}{3!} = 20$. Niech

$$X_i = \begin{cases} 1 & \text{gdy } i\text{-ty trójkąt jest jednobarwny} \\ 0 & \text{w przeciwnym przypadku} \end{cases}$$

Wówczas $X = X_1 + \ldots + X_{20}$. Jako, że $\mathbb{E}X_i = \mathbb{E}X_j$ dla $i \neq j$, to $\mathbb{E}X = \mathbb{E}(X_1+\ldots+X_n) = \mathbb{E}X_1 + \ldots + \mathbb{E}X_{20} = 20\cdot\mathbb{E}X_1 = 20\cdot P(X_1=1) = 20\cdot\frac{3\cdot1\cdot1}{3^3}$.

**Zadanie 6 PRACA DOMOWA.**
W turnieju gry w bierki bierze udział $n$ osób. Każdy rozgrywa po jednej grze z innym z graczy. Każda z gier może skończyć się wygraną jednego z graczy lub remisem, przy czym każda z tych trzech możliwości jest jednakowo prawdopodobna. Wyniki poszczególnych gier są niezależne. Niech $N$ oznacza liczbę graczy którzy nie przegrali (czyli wygrali bądź zremisowali). Oblicz wartość oczekiwaną i wariancję zmiennej $N$.

## 3. Zadania podobne do zadania 3. z egzaminu

**Zadanie 1.**
W urnie jest $N \geq 2$ kul. Dwie z nich są białe a reszta niebieskie. gracz losuje kule z urny bez zwracania do momentu wyciągnięcia kuli białej. Niech $T$ oznacza liczbę przeprowadzonych losowań.

a) Wyznacz rozkład zmiennej $T$.

b) Wyznacz wartość oczekiwaną zmiennej $\frac{1}{N-T}$

**Rozwiązanie:**

a) Mamy

$$P(T=k) = \frac{N-2}{N} \cdot \frac{N-3}{N-1} \cdot \ldots \cdot \frac{N-(k-1)-1}{N-(k-1)+1} \cdot \frac{2}{N-(k-1)} =$$
$$= \frac{N-2}{N} \cdot \frac{N-3}{N-1} \cdot \ldots \cdot \frac{N-k}{N-k+2} \cdot \frac{2}{N-k+1} = \frac{2(N-k)}{N(N-1)}$$

Mamy tu doczynienie z iloczynem teleskopowym i wszystkie wyrazy poza pierwszymi dwoma mianownikami oraz ostatnimi dwoma licznikami się skracają.

**Zadanie 2.**
Kij złamano w losowym miejscu. Niech $X$ oznacza stosunek długości lewego kawałka do długości prawego kawałka. Wyznacz rozkład zmiennej $X$.

**Rozwiązanie:**
Mamy $X \in [0, \infty)$ oraz $X$ jest rozkładem ciągłym. Zachodzi $X = \frac{x}{1-x}$ oraz $\frac{x}{1-x} \leq t \Leftrightarrow x \leq \frac{t}{t+1}$, zatem

$$F_\lambda(t) = P(X \leq t) = \begin{cases} 0 & t < 0 \\ \frac{t}{t+1} & t \geq 0 \end{cases}$$

Wiadomo też, że dystrybuanta wyznacza rozkład.

**Zadanie 3.**
Wybieramy losowo niepusty podzbiór zbioru $\{1, 2, \ldots, n\}$. Niech $X$ oznacza moc wylosowanego podzbioru. Obliczyć $\mathbb{E}X$.

**Rozwiązanie:**
Mamy $2^n - 1$ niepustych podzbiorów zbioru $\{1, \ldots, n\}$. Zbiorów o mocy $k$ jest $\binom{n}{k}$, zatem $P(X = k) = \frac{\binom{n}{k}}{2^n-1}$. Stąd $\mathbb{E}X = \sum\limits_{k=0}^{\infty} k \cdot P(X = k) = \frac{1}{2^n-1} \cdot \sum\limits_{k=0}^{\infty} k \cdot \binom{n}{k}$. Mamy $\sum\limits_{k=0}^{\infty} k \cdot \binom{n}{k} = n \cdot 2^{n-1}$, co udowodnimy przez interpretację kombinatoryczną. Z jednej strony, wybieramy z $n$ osób $k$ osób i wśród tych osób wybieramy lidera otrzymując wszystkie podziały $n$ osób na dwie grupy (z liderem lub bez). Z drugiej strony wybieramy lidera na $n$ sposobów, a następnie pozostałe $n-1$ osób dzielimy na dwie grupy (z liderem lub bez). Stąd $\mathbb{E}X = \frac{2^{n-1} \cdot n}{2^n-1}$.

**Zadanie 4.**
Rzucamy kostką aż do momentu wypadnięcia wszystkich możliwych liczb oczek. Obliczyć wartość oczekiwaną liczby rzutów.

**Rozwiązanie:**
Niech $X_i$ to liczba rzutów pomiędzy pojawieniem się $i-1$-szego nowego wyniku a pojawieniem się $i$-tego nowego wyniku wliczając nową wyrzuconą wartość. Zmienne $X_i$ to zmienne o rozkładzie geometrycznym z prawdopodopieństwem sukcesu porażki $\frac{i-1}{6}$ oraz prawdopodobieństwem sukcesu $\frac{6-i+1}{6} = \frac{7-i}{6}$. Mamy więc $\mathbb{E}X_i = \sum\limits_{k=0}^{\infty} k \cdot P(X_i = k) = \sum\limits_{k=0}^{\infty} k \cdot \left(\frac{i-1}{6}\right)^{k-1} \cdot \left(\frac{7-i}{6}\right) = \frac{7-i}{6} \cdot \frac{1}{\left(1-\frac{i-1}{6}\right)^2} = \frac{6}{7-i}$.
Mamy $X = X_1 + \ldots + X_6$, skąd wartość oczekiwana wynosi $\mathbb{E}X = \mathbb{E}X_1 + \ldots + \mathbb{E}X_6 = 1 + \frac{6}{5} + \frac{6}{4} + \frac{6}{3} + \frac{6}{2} + \frac{6}{1} = \frac{147}{10}$.

**Zadanie 5.**
Liczby $1, 2, \ldots, 100$ ustawiono losowo w ciąg $(a_1, a_2, \ldots, a_{100})$. Niech $N$ będzie największą liczbą taką, że $a_1 < a_2 < \ldots < a_N$. Wyznaczyć rozkład zmiennej $N$ oraz $\mathbb{E}N$.

**Rozwiązanie:**
Mamy $N \in \{1, \ldots, 100\}$. Mamy $P(N \geq n) = \frac{\binom{100}{n} \cdot 1 \cdot (100-n)!}{100!} = \frac{1}{n!}$ skąd otrzymujemy $P(N = n) = P(X \geq n) - P(X \geq n+1) = \frac{1}{n!} - \frac{1}{(n+1)!}$, więc $\mathbb{E}N = \sum_{n=1}^{100} n \cdot \left(\frac{1}{n!} - \frac{1}{(n+1)!}\right) = \sum_{n=1}^{100} \frac{n^2}{(n+1)!}$.

**Zadanie 6.**
Pięć uczennic i pięciu uczniów bierze udział w teście. Następnie tworzą listę rankingową. Zakładamy, że żadne dwa wyniki nie mogą się powtórzyć. Niech $X$ oznacza najwyższe miejsce osiągnięte przez uczennicę. Wyznacz rozkład $X$.

**Rozwiązanie:**
Oznaczmy uczennice liczbami od 1 do 5 oraz uczniów od 6 do 10. Mamy $X \in \{1, 2, 3, 4, 5, 6\}$. Niech $\Omega = \{(x_1, x_2, \ldots, x_{10}) \mid x_i \in \{1, \ldots, 10\},\ x_i \neq x_j \text{ dla } i \neq j\}$, $\mathcal{F} = 2^\Omega$ oraz niech $P$ to prawdopodobieństwo klasyczne. Wówczas $P(X = j) = \frac{5! \cdot 5 \cdot (10-j)!}{10! \cdot (5-j+1)!}$, bo najpierw siadamy $i - 1$ chłopców na $\frac{5!}{(5-i+1)!}$ sposobów, następnie wybieramy najzdolniejszą dziewczynkę na 5 sposobów oraz pozostałe 4 dziewczynki i $5 - i + 1$ chłopców ustawiamy na końcu na $(10 - i)!$ sposobów. Całość dzielimy przez $10!$, bo tyle jest permutacji 10 osób.