Zadanie 1.

Losujemy niezależnie dwie liczby X_1, X_2 z rozkładu o gęstości:

$$f(x) = \begin{cases} 2x & \text{dla } 0 \leq x \leq 1, \\ 0 & \text{w p.p.} \end{cases}$$

Dzielimy odcinek $[0,1]$ na 3 części: $[0, \min(X_1,X_2))$, $[\min(X_1,X_2), \max(X_1,X_2))$, $(\max(X_1,X_2), 1]$. Jakie jest prawdopodobieństwo, że z tych trzech części można ułożyć trójkąt?

Mamy odcinki: x, $y-x$, $1-y$

Warunki trójkąta:

$$\begin{cases} x + y - x > 1 - y \\ y - x + 1 - y > x \\ x + 1 - y > y - x \end{cases}$$

$$\begin{cases} x < 0.5 \\ y > 0.5 \\ y < x + 0.5 \end{cases}$$

Liczymy całkę

$$\int_0^{0.5} 2x \int_{0.5}^{x+0.5} 2y \, dy \, dx = \frac{11}{96}$$

Ponieważ nie wiemy czy $x > y$ czy $y > x$, dlatego wynik trzeba pomnożyć przez 2:

$$P = 2 \cdot \frac{11}{96} = \frac{11}{48}$$

Zadanie 2.

Niech X_1, X_2, X_3 będą niezależnymi zmiennymi losowymi o rozkładzie $P(X_i = k) = (1-p_i)p_i^{k-1}$ dla $k = 1, 2, \ldots$, gdzie $p_i = 2^{i-4}$. Zdefiniujmy zmienną losową

$$Y = \mathbb{1}(X_1 \leq X_2 \leq X_3) := \begin{cases} 1 & \text{jeśli} \quad X_1 \leq X_2 \leq X_3, \\ 0 & \text{w p. p.} \end{cases}$$

Ile wynosi EY ?

$$EX = \sum_{y_j \in R_Y} E[X | Y = y_j] \, P_Y(y_j)$$

$$P(X_i = k) = (1-p_i) p_i^{k-1}$$

$$F_{X_i}(k) = \sum_{j=1}^{k} (1-p_i) p_i^{j-1} = (1-p_i) \frac{1-p_i^k}{1-p_i} = 1 - p_i^k$$ dystrybuanta rozkładu geometrycznego

Teraz kolejne warunkowania :

$$EY = E[\mathbb{1}(X_1 \le X_2 \le X_3)] = \sum_{i=1}^{\infty} E[\mathbb{1}(i \le X_2 \le X_3)]P(X_1 = i) =$$

$$= \sum_{j=1}^{\infty}\sum_{i=1}^{j} E[\mathbb{1}(i \le j \le X_3)]\, P(X_1 = i)\, P(X_2 = j) = \left|\begin{array}{l} E[\mathbb{1}(i \le j \le X_3) = \sum_{k=j}^{\infty} P(X_3 = k) = \\ = P(X_3 \ge j) \end{array}\right|$$

$$= \sum_{j=1}^{\infty}\sum_{i=1}^{j} P(X_3 \ge j)\, P(X_1 = i)\, P(X_2 = j) = \sum_{j=1}^{\infty}\sum_{i=1}^{j} p_3^{j-1}(1-p_1)p_1^{i-1}(1-p_2)p_2^{j-1} =$$

$$= (1-p_1)(1-p_2)\sum_{j=1}^{\infty} p_3^{j-1} p_2^{j-1} \sum_{i=1}^{j} p_1^{i-1} =$$

$$= \cancel{(1-p_1)}(1-p_2)\sum_{j=1}^{\infty}(p_2 p_3)^{j-1} \cdot \frac{1-p_1^{j}}{\cancel{1-p_1}} =$$

$$= (1-p_2)\sum_{j=1}^{\infty}\left[(p_2 p_3)^{j-1} - p_1 p_1^{j-1}(p_2 p_3)^{j-1}\right] =$$

$$= (1-p_2)\left[\frac{1}{1-p_2 p_3} - \frac{p_1}{1-p_1 p_2 p_3}\right] =$$

$$= (1-2^{-2})\left(\frac{1}{1-2^{-2-1}} - \frac{2^{-3}}{1-2^{-3-2-1}}\right) =$$

$$= \frac{3}{4}\left(\frac{8}{7} - \frac{8}{63}\right) = \frac{16}{21}$$

Zadanie 3.

Załóżmy, że losujemy jednostajnie, niezależnie, ze zwracaniem $n \geq 7$ liter ze zbioru $\{\text{A, C, G, T}\}$, oznaczmy przez X_i wynik i-tego losowania. Innymi słowy mamy n niezależnych zmiennych losowych o rozkładzie $P(X_i = \text{A}) = P(X_i = \text{C}) = P(X_i = \text{G}) = P(X_i = \text{T}) = 1/4, i = 1, \ldots, n$. Niech Y_n będzie zmienną losową oznaczającą liczbę wystąpień wzorca $(A, ?, A)$ (gdzie '?' oznacza dowolną z liter A, C, G, T) w ciągu $(X_1, \ldots, X_n)$.

Uwaga: np. dla $n = 7$ w ciągu (A,C,A,T,A,T,G) wzorzec (A,?,A) występuje $Y_7 = 2$ razy (wystąpienia się nakładają, pierwszy wzorzec $(X_1, X_2, X_3) =$ (A,C,A), drugi wzorzec $(X_3, X_4, X_5) =$ (A,T,A)).

Ile wynosi EY_n?

Miejsca w jakich mogą rozpocząć się trójki:

x x x x x x _ _

A zatem tych trójek jest $n-2$

Prawdopodobieństwo, że w jednej trójce wystąpi wzorzec: $\frac{1\cdot 4\cdot 1}{4\cdot 4\cdot 4} = \frac{1}{16}$

Czyli $EY_n = \frac{n-2}{16}$

Zadanie 4.

Załóżmy, że X_1, X_2, X_3, X_4 są niezależnymi zmiennymi losowymi o jednakowym rozkładzie $P(X_i = 1) = p = 1 - P(X_i = 0), i = 1, \ldots, 4$, gdzie $p \in (0, 1)$. Testujemy hipotezę

$$H_0 : p = \frac{1}{2} \qquad vs. \qquad H_1 : p = \frac{1}{4}$$

na poziomie istotności $\alpha = \frac{11}{16}$. Oznaczmy $S = \sum_{i=1}^{4} X_i$. Jednostajnie najmocniejszy test odrzuca hipotezę H_0, gdy

$\alpha = P(S < \tilde{k} \mid \theta = \frac{1}{2})$

$S \sim Bin(4, \frac{1}{2})$

$\frac{11}{16} = \left(\frac{1}{2}\right)^4 \left[\binom{4}{0} + \binom{4}{1} + \binom{4}{2}\right]$

$S < 3 \ \vee \ S \leq 2$

Zadanie 5.

Mamy dane dwa rozkłady prawdopodobieństwa $\{p_i\}, \{q_i\}, i = 1, \ldots, 10$:

$i:$	1	2	3	4	5	6	7	8	9	10
$p_i:$	0.1	0.1	0.1	0.1	0.2	0.0	0.0	0.0	0.3	0.1
$q_i:$	0.2	0.2	0.2	0.1	0.1	0.1	0.0	0.0	0.0	0.1

Oznaczmy $E := \{1, \ldots, 10\}$. Rozpatrzmy zbiór dwuwymiarowych rozkładów, takich, że rozkładami brzegowymi są $\{p_i\}, \{q_i\}, i \in E$,

$$\mathscr{W} := \{w: \text{ rozkład na } E^2 \text{ taki, że } \forall(i \in E) \sum_{j=1}^{10} w(i,j) = p_i \text{ oraz } \forall(j \in E) \sum_{i=1}^{10} w(i,j) = q_j\}.$$

$(X,Y) \sim w$ oznacza, iż wektor (X,Y) ma rozkład w, tzn. $P(X = i, Y = j) = w(i,j)$.
Ile wynosi

$$\inf_{\substack{(X,Y)\sim w \\ w \in \mathscr{W}}} P(X \neq Y)?$$

Wystarczy posumować minima z każdej pary dwóch rozkładów:

$\inf P = 1 - 0.6 = 0.4 = \frac{2}{5}$

Zadanie 6.

Niech X będzię zmienną losową o rozkładzie Poissona z parametrem 2, a Y zmienną losową Poissona z parametrem 8, niezależną od X.
(Zmienna losowa Z jest zmienną losową o rozkładzie Poissona z parametrem $\lambda > 0$ jeśli $P(Z = k) = \frac{\lambda^k e^{-\lambda}}{k!}, k = 0, 1, \ldots$).
Niech T oznacza rozkład zmiennej losowej X pod warunkiem $X + Y = 100$, tzn.

$$P(T = k) = P(X = k | X + Y = 100), \quad k = 0, 1, \ldots, 100$$

Ile wynosi $VarT$?

Jeśli $X \sim P(2)$, $Y \sim P(8)$ to $X | X+Y = 100 \sim B\left(100, \frac{2}{2+8}\right)$

T ma rozkład dwumianowy czyli:
$Var(T) = 100 \cdot \frac{2}{10} \cdot \frac{8}{10} = 16$

Zadanie 7.

Niech $X_1,\dots,X_n, n \geq 1$ oznaczają wyniki kolejnych niezależnych rzutów kostką, tzn. są to niezależne zmienne losowe o rozkładzie $P(X_i = k) = \frac{1}{6}, k = 1,\dots,6$.

Niech Y_n oznacza liczbę różnych wyników wśród $X_1, X_2, \dots, X_n$ (np. gdy $X_1 = X_2 = X_3 = 4$, to $Y_3 = 1$, natomiast, gdy $X_1 = X_2 = 3, X_3 = 6, X_4 = 5, X_5 = 3$, to $Y_4 = 3$). Przyjmujemy $Y_0 = 0$.

Ile średnio trzeba wykonać rzutów, by każdy z wyników się pojawił? Innymi słowy, ile wynosi ET, gdzie

$$T = \min\{n \geq 0 : Y_n = 6\}?$$

$T = \min\{n \geq 0 : Y_n = 6\}$

S - uzyskanie w k-tym rzucie innej liczby niż we wszystkich poprzednich rzutach

$Y_1 \neq 6$

$Y_2 \neq 6$

$\vdots$

$Y_{n-1} \neq 6$

$Y_n = 6$

$P(T=n) = P(Y_1 \neq 6) P(Y_2 \neq 6) \cdot \dots \cdot$

$\cdot P(Y_{n-1} \neq 6) P(Y_n = 6)$

$T_k = \min\{n \geq 0 : Y_n = k\}$

$P(S=1) = 1, \quad ET_1 = 1$

$P(S=2) = \frac{5}{6}, \quad ET_2 = \frac{6}{5} = 1{,}2$

$P(S=3) = \frac{4}{6}, \quad ET_3 = \frac{6}{4} = 1{,}5$

$P(S=4) = \frac{3}{6}, \quad ET_4 = \frac{6}{3} = 2$

$P(S=5) = \frac{2}{6}, \quad ET_5 = \frac{6}{2} = 3$

$P(S=6) = \frac{1}{6}, \quad ET_6 = \frac{6}{1} = 6$

$ET = 1 + 1{,}2 + 1{,}5 + 2 + 3 + 6 = 14{,}7$

Zadanie 8.

Niech $X_1, X_2, \dots, X_n, n \geq 7$ będą niezależnymi zmiennymi losowymi o rozkładzie Erlanga z parametrem kształtu $k > 0$ (liczba całkowita) oraz częstością $\lambda > 0$ (liczba rzeczywista), tj. o gęstości

$$f(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!}, \quad \text{dla } x \geq 0.$$

Oznaczmy $\hat{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Ustalmy $s \in \{1,\dots,n\}$. Ile wynosi $Cov(\hat{X}, X_s - \hat{X})$?

$Cov(\hat{X}, X_s - \hat{X}) = Cov(\hat{X}, X_s) - Var(\hat{X})$

$Var(\hat{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2} \cdot n \cdot \frac{k}{\lambda^2} = \frac{1}{n} \cdot \frac{k}{\lambda^2}$

$Cov(\hat{X}, X_s) = Cov\left(\frac{1}{n}\sum_{i=1}^{n} X_i, X_s\right) = \frac{1}{n}\sum_{i=1}^{n} Cov(X_i, X_s) = \left| \begin{matrix} 0 & \text{gdy } i \neq 1 \\ Var(X) & \text{gdy } i = 1 \end{matrix} \right| =$

$= \frac{1}{n} \frac{k}{\lambda^2}$

$Cov(\hat{X}, X_s - \hat{X}) = \frac{1}{n}\frac{k}{\lambda^2} - \frac{1}{n}\frac{k}{\lambda^2} = 0$

Zadanie 9.

Załóżmy, że X_1, X_2, X_3, X_4 są niezależnymi zmiennymi losowymi o rozkładzie wykładniczym z parametrem 3, tzn. o gęstości

$$f_X(x) = \begin{cases} 3e^{-3x} & \text{jeśli} \quad x \geq 0, \\ 0 & \text{w p. p.} \end{cases}$$

Oznaczmy $S = \frac{1}{4}\sum_{i=1}^{4} X_i$. Ile wynosi $P(X_1 > S)$?

$Y = X_2 + X_3 + X_4 \sim \text{gamma}(3,3)$

$P(X_1 > S) = P(X_1 > \frac{1}{4}(X_1+X_2+X_3+X_4)) = P(4X_1 > X_1 + Y) = P(3X_1 > Y) =$

$= \int_0^\infty P(3X_1 > Y \mid Y = y) f_Y(y)\,dy = \int_0^\infty P(X_1 > \frac{y}{3}) f_Y(y)\,dy =$

$= \int_0^\infty [1 - P(X_1 < \frac{y}{3})] f_Y(y)\,dy = \int_0^\infty (1 - 1 + e^{-3\cdot\frac{y}{3}}) f_Y(y)\,dy =$

$= \int_0^\infty e^{-y} f_Y(y)\,dy = E[e^{-Y}] = |\text{FGM roz. gamma}| = \left(\frac{3}{3+1}\right)^3 = \frac{27}{64}$

Zadanie 10.

Zmienna losowa X ma rozkład standardowy normalny $N(0,1)$, tj. ma gęstość: $\phi(x) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{x^2}{2}\right), x \in \mathbb{R}$. Ustalmy $r > 0$ i zdefiniujmy

$$Y := \begin{cases} X & \text{jeśli} \quad |X| < r, \\ -X & \text{jeśli} \quad |X| \geq r. \end{cases}$$

Kowariancja $Cov(X,Y)$ wynosi 0 dla parametru r spełniającego równość:

$cov(X,Y) = EXY - \overbrace{EXEY}^{0} = EXY = E X X \cdot \mathbb{1}_{(|X|<r)} + EX(-X)\cdot \mathbb{1}_{(|X|\geq r)} =$

$= EX^2 \mathbb{1}_{(|X|<r)} - EX^2 \mathbb{1}_{(|X|\geq r)} = EX^2 \mathbb{1}_{(|X|<r)} - EX^2(1 - \mathbb{1}_{(|X|<r)}) =$

$= EX^2 \mathbb{1}_{(|X|<r)} - EX^2 + EX^2 \mathbb{1}_{(|X|<r)} = 0$

$2 EX^2 \mathbb{1}_{(|X|<r)} = 1$

$EX^2 \mathbb{1}_{(|X|<r)} = \frac{1}{2}$

$2\int_0^r x^2 \varphi(x)\,dx = \frac{1}{2}$ $\longrightarrow$ rozkład normalny symetryczny względem OY

$\int_0^r x^2 \varphi(x)\,dx = \frac{1}{4}$