1 Parametry rozkładów

W tej serii będziemy obliczać różne parametry zmiennych losowych. Wymagana jest znajomość własności rozkładów probabilistycznych oraz umiejętne korzystanie z niezależności zmiennych losowych.

1. (Eg 48/7) Niech $X_1, X_2, ..., X_n, ...$ będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie wykładniczym o gęstości

$$f(x) = 2e^{-2x} \text{ gdy} x > 0 \text{ oraz } f(x) = 0 \text{ gdy } x \leqslant 0.$$

Niech N będzie zmienną losową, niezależną od $X_1, X_2, ..., X_n, ...$, o rozkładzie ujemnym dwumianowym $\mathbf{P}(N = n) = \frac{\Gamma(r+n)}{\Gamma(r)n!}p^r(1-p)^n$ dla $n = 0, 1, 2, ...$, gdzie $r > 0$ i $p \in (0, 1)$ są ustalonymi parametrami. Niech $Z_N = \min(X_1, X_2, ..., X_N)$, gdy $N > 0$ oraz $Z_N = 0$, gdy $N = 0$. Oblicz $\mathbf{E}(NZ_N)$ i $\mathbf{Var}(NZ_N)$.
Odp: C-> $\mathbf{E}(NZ_N) = \frac{1-p^r}{2}$ i $\mathbf{Var}(NZ_N) = \frac{1-p^{2r}}{4}$.

Rozwiązanie. Dla zmiennej Z z rozkładu wykładniczego $Exp(\lambda)$ zachodzi $\mathbf{E}Z = \frac{k!}{\lambda^k}$, $k = 0, 1, 2, \ldots$. Przypomnijmy, ze minimum z niezależnych zmiennych wykładniczych ma rozkład wykładniczy o parametrze będącym sumą parametrów dodawanych zmiennych. To znaczy rozkład Z ma postać $Exp(2n)$. W szczególności $\mathbf{E}Z_n = \frac{1}{2n}$ oraz $\mathbf{E}Z_n^2 = \frac{1}{2n^2}$ oraz $\mathbf{E}Z_0 = \mathbf{E}Z_0^2 = 0$. Obliczamy

$$\mathbf{E}(NZ_N) = \sum_{n=0}^{\infty} \mathbf{P}(N = n)n\mathbf{E}Z_n = \frac{1}{2}\sum_{n=1}^{\infty} \mathbf{P}(N = n) = \frac{1}{2}(1 - p^r),$$

$$\mathbf{E}(NZ_N)^2 = \sum_{n=0}^{\infty} \mathbf{P}(N = n)n^2\mathbf{E}Z_n^2 = \sum_{n=1}^{\infty} = \frac{1}{2}\sum_{n=1}^{\infty} \mathbf{P}(N = n) = \frac{1}{2}(1 - p^r).$$

Stąd

$$\mathbf{Var}(NZ_N) = \frac{1}{4}(2 - 2p^r - 1 + 2p^r - p^{2r}) = \frac{1}{4}(1 - p^{2r}).$$

■

2. (Eg 49/3) Niech $X_1, X_2, ..., X_n, ...$ $I_1, I_2, ..., I_n, ..., N$ będą niezależnymi zmiennymi losowymi. Zmienne $X_1, X_2, ..., X_n, ...$ mają rozkład o wartości oczekiwanej 2 i wariancji 1. Zmienne $I_1, I_2, ..., I_n, ...$ mają rozkład jednostajny na przedziale $(0, 1)$. Zmienna N ma rozkład ujemny dwumianowy $\mathbf{P}(N = n)\frac{\Gamma(2+n)}{n!}(\frac{3}{4})^2(\frac{1}{4})^n$ dla $n = 0, 1, 2, ...$ Niech $S_N = 0$, gdy $N = 0$ oraz $S_N = \sum_{i=1}^{N} I_iX_i$, gdy $N > 0$. Wtedy $\mathbf{Var}(S_N)$ jest równa: ?
Odp: C-> $\mathbf{Var}S_N = \frac{4}{3}$.

Rozwiązanie. Przypomnijmy własności rozkładu ujemnego dwumianowego $\mathbf{E}N = 2(\frac{3}{4})^{-1}(\frac{1}{4}) = \frac{2}{3}$ oraz $\mathbf{Var}N = 2(\frac{3}{4})^{-2}\frac{1}{4} = \frac{8}{9}$ i $\mathbf{E}N^2 = \frac{4}{3}$. Niech X ma rozkład taki jak $X_1, X_2, \ldots$ nadto I ma rozkład taki jak $I_1, I_2, \ldots$ Mamy

$$\mathbf{E}S_N = \sum_{n=0}^{\infty} \mathbf{P}(N = n)\mathbf{E}S_n = \sum_{n=0}^{\infty} \mathbf{P}(N = n)n\mathbf{E}I\mathbf{E}X = \sum_{n=1}^{\infty} \mathbf{P}(N = n)n = \mathbf{E}N.$$

Nadto

$$\mathbf{E}S_N^2 = \sum_{n=0}^{\infty} \mathbf{P}(N = n)\mathbf{E}S_n^2 = \sum_{n=0}^{\infty} \mathbf{P}(N = n)(n\mathbf{E}X^2\mathbf{E}I^2 + n(n-1)(\mathbf{E}X\mathbf{E}I)^2) =$$

$$= \sum_{n=0}^{\infty} \mathbf{P}(N = n)(n(n + \frac{2}{3})) = \mathbf{E}N(N + \frac{2}{3}).$$

Czyli

$$\mathbf{Var}S_N = \mathbf{E}N^2 + \frac{2}{3}\mathbf{E}N - (\mathbf{E}N)^2 = \frac{4}{3} + \frac{4}{9} - \frac{4}{9} = \frac{4}{3}.$$

■

3. (Eg 50/1) Niech $X_1, X_2, ..., X_n, ...$ będą niezależnymi zmiennymi losowymi z rozkładu gamma o gęstości

$$p(x) = 16xe^{-4x}, \text{ gdy } x > 0, \text{ oraz } p(x) = 0 \text{ gdy} x \leqslant 0.$$

Niech N będzie zmienną losową niezależna od zmiennych $X_1, X_2, ..., X_n, ...$ spełniającą

$$\mathbf{P}(N = 0) = \frac{1}{2}, \text{ i } \mathbf{P}(N = 1) = \mathbf{P}(N = 2) = \mathbf{P}(N = 3) = \frac{1}{6}.$$

Niech $S = 0$, gdy $N = 0$ oraz $S = \sum_{i=1}^{N} X_i$, gdy $N > 0$. Wtedy $\mathbf{E}(S - \mathbf{E}S)^3$ jest równe ?.
Odp: B-> $\frac{7}{16}$.

Rozwiązanie. Przypomnijmy momenty dla zmiennej Z z rozkładu $Gamma(\alpha, \beta)$ zachodzi

$$\mathbf{E}Z = \frac{\alpha}{\beta}, \ \mathbf{E}(Z - \mathbf{E}Z)^2 = \frac{\alpha}{\beta^2}, \ \mathbf{E}(Z - \mathbf{E}Z)^3 = \frac{2\alpha}{\beta^3}.$$

W przypadku zadania $\alpha = 2$, $\beta = 4$, stąd

$$\mathbf{E}X = \frac{1}{2}, \ \mathbf{E}(X - \mathbf{E}X)^2 = \frac{1}{8}, \ \mathbf{E}(X - \mathbf{E}X)^3 = \frac{1}{16}.$$

Podobnie wyznaczamy

$$\mathbf{E}N = 1, \ \mathbf{E}(N - \mathbf{E}N)^2 = \frac{4}{3}, \ \mathbf{E}(N - \mathbf{E}N)^3 = 1.$$

Zatem

$$\begin{aligned}
&\mathbf{E}(S - \mathbf{E}S)^3 = \mathbf{E}(S - N\mathbf{E}X + N\mathbf{E}X - \mathbf{E}S)^3 = \mathbf{E}\mathbf{E}[(S - N\mathbf{E}X + N\mathbf{E}X - \mathbf{E}S)^3|N] = \\
&= \mathbf{E}NE(X - \mathbf{E}X)^3 + 3\mathbf{E}(N - \mathbf{E}N)^2\mathbf{E}X\mathbf{E}(X - \mathbf{E}X)^2 + \mathbf{E}(N - \mathbf{E}N)^3(\mathbf{E}X)^3 = \\
&= \frac{1}{16} + 3 \cdot \frac{4}{3} \cdot \frac{1}{2} \cdot \frac{1}{8} + \frac{1}{8} = \frac{7}{16}.
\end{aligned}$$

■

4. (Eg 51/7) Załóżmy, że $X_1, X_2, ..., X_n, ...$ są niezależnymi zmiennymi losowymi o jednakowym rozkładzie wykładniczym o gęstości

$$f(x) = \exp(-x) \text{ dla } x > 0.$$

Zmienna losowa N jest niezależna od $X_1, X_2, ..., X_n, ...$ i ma rozkład Poissona o wartości oczekiwanej λ. Niech

$$Y_i = \min(X_i, 2), \ \ Z_i = X_i - Y_i,$$
$$S^{(Y)} = \sum_{i=1}^{N} Y_i, \ \ S^{(Z)} = \sum_{i=1}^{N} Z_i.$$

Oblicz $\mathbf{Cov}(S^{(Y)}, S^{(Z)})$.
Odp: A-> $\mathbf{Cov}(S^{(Y)}, S^{(Z)}) = 2\lambda e^{-2}$.

Rozwiązanie. Niech Y, Z będą miały rozkład odpowiednio jak $X_1, X_2, \ldots, Y_1, Y_2, \ldots$. Wyznaczamy

$$\mathbf{E}S^{(Y)} = \mathbf{E}N\mathbf{E}Y = \lambda\mathbf{E}Y$$

oraz

$$\mathbf{E}S^{(Z)} = \mathbf{E}N(\mathbf{E}X - \mathbf{E}Y) = \lambda\mathbf{E}Z.$$

Wyznaczamy

$$\mathbf{E}S^{(Y)}S^{(Z)} = \sum_{n=0}^{\infty}\mathbf{P}(N=n)\mathbf{E}(\sum_{i=1}^{n}Y_i,\sum_{i=1}^{n}Z_i) = \sum_{n=0}^{\infty}\mathbf{P}(N=n)(n\mathbf{E}(YZ)+n(n-1)\mathbf{E}Y\mathbf{E}Z) =$$
$$= \mathbf{E}N\mathbf{E}(YZ)+\mathbf{E}(N(N-1))\mathbf{E}Y\mathbf{E}Z = \lambda\mathbf{E}(YZ)+\lambda^2\mathbf{E}Y\mathbf{E}Z.$$

Mamy

$$\mathbf{E}(YZ) = 2\mathbf{E}(X-2)_+ = 2\int_{t=0}^{\infty}\mathbf{P}(X>2+t)dt = 2e^{-2}.$$

Stąd

$$\mathbf{Cov}(S^{(Y)},S^{(Z)}) = 2\lambda e^{-2}$$

■

5. (Eg 52/8) Załóżmy, że $X_1, X_2, ..., X_n, ...$ są niezależnymi zmiennymi losowymi o tym samym rozkładzie jednostajnym na przedziale $[0,1]$, zaś N jest zmienną losową o rozkładzie ujemnym dwumianowym:

$$\mathbf{P}(N=n) = \binom{n+2}{n}p^3(1-p)^n \text{ dla } n = 0,1,2,...$$

niezależną od zmiennych losowych $X_1, X_2, ..., X_n, ...$ Niech $M_N = \max\{X_1, X_2, ..., X_N\}$, gdy $N > 0$ oraz $M_N = 0$, gdy $N = 0$. Statystyk otrzymał trzy niezależne obserwacje zmiennej losowej M_N równe

$$0, 0, \frac{1}{2}.$$

Wartość estymatora największej wiarogodności dla parametru p otrzymana na podstawie tych danych jest równa ?
Odp: B-> $0,877$.

Rozwiązanie. Zadanie należy rozwiązać przez wyznaczenie wiarygodności, skoro wśród obserwacji pojawiły się dwa zera to znaczy, że dwukrotnie pojawiło się zdarzenie o prawdopodobieństwie p^3. Z drugiej strony szansa otrzymania wartości $0 < t \leqslant 1$ wynosi

$$\sum_{n=1}^{\infty}\mathbf{P}(N=n)\mathbf{P}(\max(X_1,...,X_n)\leqslant t) = \sum_{n=1}^{\infty}\mathbf{P}(N=n)t^n = \frac{p^3}{(1-t(1-p))^3},$$

gdzie ostatnia równość bierze się ze zmiany parametru w rozkładzie ujemnym dwumianowym. Zatem warunkowo na $N > 0$ zmienna M_N ma rozkład ciągły o gęstości

$$\frac{3p^3(1-p)}{(1-t(1-p))^4}.$$

Stąd wiarygodność będzie miała postać

$$L(p,(0,0,\frac{1}{2}))) = \frac{48p^9(1-p)}{(1+p)^4}.$$

Znajdujemy maksimum $f(p) = \log L(p,(0,0,\frac{1}{2}))$,

$$0 = f'(p) = 9\frac{1}{p} - \frac{1}{1-p} - 4\frac{1}{1+p}.$$

Stąd równanie

$$6p^2+5p+9 = 0,\ \ p = \frac{-5+\sqrt{406}}{12} \simeq 0,877.$$

■

6. (Eg 53/2) Niech $X_0, X_1, ..., X_n, ...$ będą niezależnymi zmiennymi losowymi o rozkładzie jednostajnym na przedziale $(0,2)$. Niech zmienna losowa N oznacza numer pierwszej ze zmiennych losowych $X_1, ..., X_n, ...$ o wartości większej niż X_0, zatem

$$N = \inf\{n \geqslant 1 : \ X_n > X_0\}.$$

Wtedy $\mathbf{E}X_N$ jest równa?
Odp: C-> $\frac{3}{2}$.

Rozwiązanie. Dla ustalonego $X_0 = x \in (0,2)$ rozkład $N_x = \inf\{n \geqslant : \ X_n > x\}$ jest rozkładem geometrycznym z prawdopodobieństwem sukcesu $(1-x)$. Dalej rozkład X_N dla ustalonego N i X_0 ma rozkład jednostajny na $(x,2)$ ze średnią $\frac{x+2}{2}$. Zatem

$$\mathbf{E}X_N = \mathbf{E}\mathbf{E}(X_N|X_0, N) = \mathbf{E}\frac{X_0+2}{2} = \frac{3}{2}.$$

■

7. (Eg 54/4) Zmienna losowa N ma rozkład geometryczny postaci

$$\mathbf{P}(N=k) = (\frac{1}{4})^k \frac{3}{4} \ \text{ dla } \ k = 0,1,2,...$$

Rozważamy losową liczbę zmiennych losowych $X_1, X_2, ..., X_N$ przy czym zmienne losowe $X_1, X_2, ..., X_N$ są niezależne wzajemnie i niezależne od zmiennej losowej N. Każda ze zmiennych X_i ma ten sam rozkład o parametrach

$$\mathbf{E}X = 1, \ \ \mathbf{E}(X^2) = 2, \ \ \mathbf{E}(X^3) = 3.$$

Niech $S_N = \sum_{i=1}^{N} X_i$, gdy $N > 0$ oraz $S_N = 0$, gdy $N = 0$. Współczynnik skośności $\frac{\mathbf{E}(S_N - \mathbf{E}S_N)^3}{(\mathbf{Var}S_N)^{\frac{3}{2}}}$
Odp: B-> $2,538$.

Rozwiązanie. Przypomnijmy, że

$$\begin{aligned}\mathbf{E}(S-\mathbf{E}S)^3 &= \mathbf{E}(S - N\mathbf{E}X + N\mathbf{E}X - \mathbf{E}S)^3 = \mathbf{E}\mathbf{E}[(S - N\mathbf{E}X + N\mathbf{E}X - \mathbf{E}S)^3|N] = \\ &= \mathbf{E}N\mathbf{E}(X-\mathbf{E}X)^3 + 3\mathbf{E}(N-\mathbf{E}N)^2\mathbf{E}X\mathbf{E}(X-\mathbf{E}X)^2 + \mathbf{E}(N-\mathbf{E}N)^3(\mathbf{E}X)^3.\end{aligned}$$

Dla rozkładu geometrycznego (startującego z 0) z parametrem p

$$\mathbf{E}N = \frac{1-p}{p}, \ \ \mathbf{E}(N-\mathbf{E}N)^2 = \frac{(1-p)}{p^2}, \ \ \mathbf{E}(N-\mathbf{E}N)^3 = \frac{(1-p)(2-p)}{p^3}.$$

Stad

$$\mathbf{E}N = \frac{1}{3}, \ \ \mathbf{E}(N-\mathbf{E}N)^2 = \frac{4}{9}, \ \ \mathbf{E}(N-\mathbf{E}N)^3 = \frac{20}{27}.$$

Dalej

$$\mathbf{E}X = 1, \ \ \mathbf{E}(X-\mathbf{E}X)^2 = 1, \ \ \mathbf{E}(X-\mathbf{E}X)^3 = -1.$$

Dostajemy

$$\mathbf{E}(S-\mathbf{E}S)^3 = -\frac{1}{3} + \frac{4}{3} + \frac{20}{27}$$

Podobnie

$$\mathbf{E}(S-\mathbf{E}S)^2 = \mathbf{E}N\mathbf{E}(X-\mathbf{E}X)^2 + \mathbf{E}(N-\mathbf{E}N)^2(\mathbf{E}X)^2 = \frac{1}{3} + \frac{4}{9} = \frac{7}{9}.$$

Współczynnik skośności

$$\frac{\mathbf{E}(S_N - \mathbf{E}S_N)^3}{(\mathbf{Var}S_N)^{\frac{3}{2}}} = \frac{47}{(7)^{\frac{3}{2}}} \simeq 2,538.$$

■

8. (Eg 55/7) Zmienna losowa N ma rozkład geometryczny

$$\mathbf{P}(N=n)=p^n(1-p), \quad \text{dla } n=0,1,2,...,$$

gdzie $p\in(0,1)$ jest nieznanym parametrem. Rozważamy losową liczbę zmiennych losowych

$$X_1,X_2,...,X_N,$$

przy czym zmienne losowe $X_1,X_2,...,X_N$ są niezależne wzajemnie i niezależne od zmiennej losowej N. Każda ze zmiennych X_i ma rozkład jednostajny o gęstości danej wzorem: $f_\theta(x)=1/\theta$, dla $0\leqslant x\leqslant\theta$ oraz 0 w przeciwnym przypadku, gdzie $\theta>0$ jest nieznanym parametrem. Obserwujemy tylko te spośród zmiennych $X_1,X_2,...,X_N$, które są większe od 5. Nie wiemy ile jest pozostałych zmiennych ani jakie są ich wartości. Przypuśćmy, że zaobserwowaliśmy następujące wartości

$$8.5, 10, 6, 7.4, 9, 5.2.$$

Na podstawie tych danych wyznacz wartości estymatorów największej wiarogodności parametrów θ i p.
Odp: C-> $\hat{\theta}=10$ i $\hat{p}=\frac{12}{13}$.
Rozwiązanie. Niech X ma rozkład taki jak $X_1,X_2,\ldots$. Znajdujemy wiarygodność, szansa że wylosujemy 6 zmiennych większych niż 5 wynosi

$$(\mathbf{P}(X>5))^6\sum_{n=0}^{\infty}\mathbf{P}(X\leqslant 5)^n\mathbf{P}(N=n+6)=(\mathbf{P}(X>5))^6(1-p)p^6\sum_{n=0}^{\infty}\binom{n+6}{6}(\frac{5p}{\theta})^n.$$

Warunkowo względem wymagania, że 6 zmiennych przekroczyło 5 ma gęstość postaci

$$(\mathbf{P}(X>5))^{-6}(\frac{5}{\theta})^6\prod_{i=1}^{6}1_{x_i<\theta}.$$

Nadto z własności rozkładu ujemnego dwumianowego

$$\sum_{n=0}^{\infty}\binom{n+6}{6}(\frac{5p}{\theta})^n=(1-\frac{5p}{\theta})^{-7}$$

Stąd wiarygodność ma postać

$$L(\theta,p;x_1,x_2,x_3,x_4,x_5,x_6)=(1-p)(\frac{5p}{\theta})^6\prod_{i=1}^{6}1_{x_i<\theta}(1-\frac{5p}{\theta})^{-7}.$$

Najpierw zauważamy, że $\theta\geqslant 10$ (w przeciwnym razie dostajemy 0) nadto dla $\theta\geqslant 0$ funkcja L jest malejąca czyli zatem $\hat{\theta}=10$. Wówczas wystarczy znaleźć p dla którego maksimum osiąga funkcja $L(10,p;x_1,x_2,x_3,x_4,x_5,x_6)$. To oznacza, że trzeba znaleźć punkt maksimum funkcji

$$f(p)=(1-p)p^6(1-\frac{1}{2}p)^{-7}.$$

Warunek $f'(p)=0$ jest równoważny

$$(6p^5-7p^6)(1-\frac{1}{2}p)+\frac{7}{2}p^6-\frac{7}{2}p^7.$$

Stąd $\bar{p}=\frac{12}{13}$.

■

9. (Eg 56/6) Załóżmy, że $X_1, X_2, ..., X_n, ...$ są niezależnymi zmiennymi losowymi o tym samym rozkładzie jednostajnym na przedziale $[0,1]$, zaś N jest zmienną losową o rozkładzie geometrycznym,

$$\mathbf{P}(N=k) = p(1-p)^k, \quad \text{gdy} \quad k = 0,1,2,...$$

niezależną od zmiennych losowych $X_1, X_2, ..., X_n, ...$ Liczba $p \in (0,1)$ jest ustalona. Niech $Y_N = \min\{X_1, ..., X_N\}$, gdy $N > 0$ oraz $Y_N = 0$, gdy $N = 0$ i $Z_N = \max\{X_1, ..., X_N\}$, gdy $N = 0$. Obliczyć $\mathbf{P}(Z_N - Y_N > \frac{1}{2})$.
Odp: B-> $1 - \frac{4p}{(1+p)^2}$.

Rozwiązanie. Zacznijmy od prostszego zadania, rozkład (Y_n, Z_n) ma gęstość $n(n-1)(z-y)^{n-2}1_{0<y<z<1}$. Zatem

$$\mathbf{P}(Z_n - Y_n > \frac{1}{2}) = \int_{\frac{1}{2}}^1 \int_0^{z-\frac{1}{2}} n(n-1)(z-y)^{n-2} dydz = \int_{\frac{1}{2}}^1 n(z^{n-1} - (\frac{1}{2})^{n-1})dz = 1 - \frac{1}{2^n} - n\frac{1}{2^n}.$$

Stąd

$$\mathbf{P}(Z_N - Y_N > \frac{1}{2}) = \mathbf{E}\mathbf{P}(Z_N - Y_N > \frac{1}{2}|N) = \mathbf{E}(1 - \frac{1}{2^N} - N\frac{1}{2^N}) =$$
$$1 - \sum_{k=0}^{\infty}(1+k)p[(1-p)2^{-1}]^k = 1 - \frac{2p}{1+p} - \frac{2p(1-p)}{(1+p)^2} = 1\frac{4p}{(1+p)^2}.$$

■

10. (Eg 57/4) Niech $X_1, X_2, ..., X_n, ...$ $I_1, I_2, ..., I_n, ..., N$ będą niezależnymi zmiennymi losowymi. Zmienne $X_1, X_2, ..., X_n, ...$ mają rozkład wykładniczy o wartości oczekiwanej 1. Zmienne $I_1, I_2, ..., I_n, ...$ mają rozkład dwupunktowy $\mathbf{P}(I_i = 1) = 1 - \mathbf{P}(I_i = 0) = \frac{1}{2}$. Zmienna N ma rozkład ujemny dwumianowy $\mathbf{P}(N = n) = \binom{n+1}{n}(\frac{3}{4})^2(\frac{1}{4})^n$ dla $n = 0,1,2,....$ Niech $S_N = 0$, gdy $N = 0$ oraz $S_N = \sum_{i=1}^N I_i X_i$, gdy $N > 0$. Wtedy współczynnik zmienności $\frac{\sqrt{\mathbf{Var}(S_N)}}{\mathbf{E}S_N}$ jest równy: ?
Odp: C-> $\sqrt{\frac{13}{2}}$.

Rozwiązanie. Niech X, I mają rozkład taki jak odpowiednio $X_1, X_2, \ldots, I_1, I_2, \ldots$. Obliczamy $\mathbf{E}N = 2/3$, $\mathbf{Var}N = 8/9$ (rozkład ujemny dwumianowy $\mathcal{B}_-(\alpha, p)$ ma wartość oczekiwana $\alpha\frac{1-p}{p}$ i wariancję $\alpha\frac{1-p}{p^2}$), $\mathbf{E}I = \frac{1}{2}$, $\mathbf{E}X = 1$, $\mathbf{Var}X = 1$ (rozkład wykładniczy $Exp(\lambda)$ ma wartość oczekiwaną λ i wariancję $1/\lambda^2$). Zatem

$$\mathbf{E}S_N = \mathbf{E}N\mathbf{E}I\mathbf{E}X = 2(\frac{1}{4}/\frac{3}{4})\frac{1}{2} = \frac{1}{3}.$$

Z drugiej strony

$$\mathbf{Var}S_N = \mathbf{E}N\mathbf{Var}(IX) + \mathbf{Var}N(\mathbf{E}IX)^2 = \mathbf{E}N(\mathbf{E}I^2\mathbf{E}X^2 - (\mathbf{E}I\mathbf{E}X)^2) + \mathbf{Var}N(\mathbf{E}I\mathbf{E}X)^2 =$$
$$= \frac{2}{3}(\frac{1}{2}\cdot 2 - (\frac{1}{2})^2) + \frac{8}{9}(\frac{1}{2})^2 = \frac{1}{2} + \frac{2}{9} = \frac{13}{18}.$$

Czyli

$$\frac{\sqrt{\mathbf{Var}(S_N)}}{\mathbf{E}S_N} = \frac{\sqrt{13}}{\sqrt{2}}.$$

■

11. (Eg 59/8) Załóżmy, że $X_1, X_2, ..., X_n, ...$ są dodatnimi niezależnymi zmiennymi losowymi o jednakowym ciągłym rozkładzie prawdopodobieństwa. Niech $R_0 = 0$ i $R_n = \max\{X_1, ..., X_n\}$, gdy $n > 0$. Niech N i M będą niezależnymi zmiennymi losowymi o rozkładach Poissona, przy czym

$\mathbf{E}N = 1$ i $\mathbf{E}M = 2$. Wtedy $\mathbf{P}(R_{N+M} > R_N)$ jest równe?
Odp: C-> $\frac{2}{3}(1-e^{-3})$.

Rozwiązanie. Zauważmy, że dla $n+m>0$

$$\mathbf{P}(R_{n+m} > R_n) = \frac{m}{n+m},$$

bo szanse, że któraś ze zmiennych jest największa są równe a szansa, że dwie zmienne przyjmują tę samą wartość jest zerowa. Zatem

$$\mathbf{P}(R_{N+M} > R_N) = \mathbf{E}\mathbf{P}(R_{N+M} > R_N|N,M) = \mathbf{E}\frac{M}{N+M}1_{N+M>0}.$$

Przypomnijmy, że rozkład warunkowy M pod warunkiem $N+M$ ma postać $\mathcal{B}(\frac{\mathbf{E}M}{\mathbf{E}N+\mathbf{E}M}, N+M)$ nadto $N+M$ ma rozkład $Poiss(\mathbf{E}N+\mathbf{E}M)$. Czyli

$$\mathbf{E}\frac{M}{N+M}1_{N+M>0} = \mathbf{E}\frac{1}{N+M}1_{N+M>0}\mathbf{E}(M|N+M) = \frac{2}{3}\mathbf{E}1_{N+M>0} = \frac{2}{2}(1-e^{-3}).$$

■

12. (Eg 60/2) Zmienne losowe $X_1, X_2, ..., X_n$ są warunkowo niezależne przy znanej wartości zmiennej losowej θ i mają rozkłady o wartości oczekiwanej $\mathbf{E}(X_i|\theta) = 10\theta$ i wariancji $\mathbf{Var}(X_i|\theta) = 100\theta^2$. Niech N będzie zmienną losową warunkowo niezależną od $X_1, X_2, ..., X_n$ przy znanym θ i o warunkowym rozkładzie

$$\mathbf{P}(N = n|\theta) = n(1-\theta)^{n-1}\theta^2, \quad \text{dla} \ \ n = 1,2,3,...$$

Zmienna losowa θ ma rozkład Beta o gęstości $p(\theta) = 6\theta(1-\theta)$, gdy $\theta \in (0,1)$. Niech $S_N = \sum_{i=1}^{N} X_i$. Wtedy wariancja $\mathbf{Var}(\frac{S_N}{N})$ jest równa?
Odp: C-> 25.

Rozwiązanie. Zauważmy, że $\mathbf{E}(N^{-1}|\theta) = \theta^2\sum_{n=1}^{\infty}(1-\theta)^{n-1} = \theta$, zatem

$$\begin{aligned}\mathbf{E}(\frac{S_N}{N} - \mathbf{E}\frac{S_N}{N})^2 &= \mathbf{E}\mathbf{E}[(\frac{S_N - N\mathbf{E}(X|\theta)}{N} + \mathbf{E}(X|\theta) - \mathbf{E}\frac{S_N}{N})^2|\theta] = \\ &= \mathbf{E}[\mathbf{E}(N^{-1}|\theta)\mathbf{Var}(X|\theta)] + \mathbf{Var}(\mathbf{E}(X|\theta)) = 100\mathbf{E}\theta^3 + 100\mathbf{Var}(\theta) =\end{aligned}$$

Dalej $\mathbf{E}\theta^3 = 6\int_0^1 \theta^4(1-\theta)d\theta = \frac{1}{5}$, $\mathbf{E}\theta^2 = 6\int_0^1\theta^3(1-\theta)d\theta = \frac{3}{10}$, $\mathbf{E}\theta = 6\int_0^1\theta^2(1-\theta) = \frac{1}{2}$, stąd $\mathbf{Var}X = \frac{1}{20}$. Ogólnie można pamiętać wzór że dla rozkładu $Beta(\alpha,\beta)$, $\mathbf{E}X = \frac{\alpha}{\alpha+\beta}$, $\mathbf{Var}X = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$, w zadaniu $\alpha = 2, \beta = 2$. Czyli

$$\mathbf{Var}\frac{S_N}{N} = 20 + 5 = 25$$

■

13. (Eg 61/10) Niech $X_1, X_2, ..., X_n, ...$ będą niezależnymi zmiennymi losowymi o rozkładzie wykładniczym o wartości oczekiwanej 4. Niech N będzie zmienną losową o rozkładzie geometrycznym o funkcji p-stwa

$$\mathbf{P}(N = k) = \frac{4}{5}\cdot(\frac{1}{5})^k \quad \text{dla} \ \ k = 0,1,2,...,$$

niezależną od zmiennych $X_1, X_2, ..., X_n, ...$ Niech $S_n = \sum_{i=1}^{N} X_i$, gdy $N > 0$ oraz $S_N = 0$ gdy $N = 0$. Wtedy $\mathbf{P}(S_N < 5)$ jest równe: ?
Odp: C-> $1 - 0,2e^{-1}$.

Rozwiązanie. Przypomnijmy, że S_n, $n \geqslant 1$ ma rozkład $Gamma(n, \frac{1}{4})$ Obliczamy

$$\mathbf{P}(S_N \geqslant 5) = \mathbf{EP}(S_N \geqslant 5|N) = \sum_{n=1}^{\infty} \mathbf{P}(S_n \geqslant 5)\mathbf{P}(N = n) = \sum_{n=1}^{\infty} \int_5^{\infty} \frac{1}{(n-1)!4^n} x^{n-1} e^{-\frac{1}{4}x} dx =$$

$$= \frac{4}{5} \int_5^{\infty} (\sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!(20)^n}) e^{-\frac{1}{4}x} dx = \frac{1}{25} \int_{n=5}^{\infty} e^{(\frac{1}{20}-\frac{1}{4})x} dx = \frac{1}{5} e^{-1}.$$

Stąd $\mathbf{P}(S_N < 5) = 1 - 0,2e^{-1}$.

■

14. (Eg 62/9) Niech $X_1, ..., X_n, ...$ będą niezależnymi zmiennymi losowymi z rozkładu jednostajnego na przedziale $[0, 1]$. Niech N będzie zmienną losową o rozkładzie ujemnym dwumianowym niezależną od zmiennych $X_1, ..., X_n, ...$ o funkcji p-stwa

$$\mathbf{P}(N = n) = \frac{(n+2)(n+1)}{2} p^3 (1-p)^n, \quad \text{dla } n = 0, 1, 2, ...$$

Niech $Y_N = \min\{X_1, ..., X_N\}$, gdy $N > 0$ oraz $Y_N = 0$, gdy $N = 0$ i $Z_N = \max\{X_1, ..., X_N\}$, oraz $Z_N = 0$ gdy $N = 0$. Wyznacz $\mathbf{E}(Y_N Z_N)$.
Odp: E-> $\frac{p(1-p^2)}{2}$.
Rozwiązanie. Przypomnijmy, że (Y_n, Z_n) ma rozkład o gęstości $n(n-1)(z-y)^{n-2} 1_{0 \leqslant y < z \leqslant 1}$. Nadto z rozkładu beta wiemy, że $\int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ Zatem podstawiając $y = zw$

$$\mathbf{E}Y_n Z_n = \int_0^1 \int_0^z n(n-1) yz(z-y)^{n-2} dy dz = n(n-1) \int_0^1 z^{n+1} dz \int_0^1 w(1-w)^{n-2} dw = \frac{1}{n+2}.$$

Stąd

$$\mathbf{E}(Y_N Z_N) = \mathbf{EE}(Y_N Z_N|N) = \mathbf{E}\frac{1}{N+2} 1_{N>0} = \sum_{n=1}^{\infty} \frac{(n+1)}{2} p^3 (1-p)^n.$$

Z ujemnego rozkładu dwumianowego wiemy, że $\sum_{n=0}^{\infty} (n+1)p^2(1-p)^n = 1$. Zatem

$$\mathbf{E}(Y_N Z_N) = -\frac{p^3}{2} + \frac{p}{2} = \frac{p(1-p^2)}{2}.$$

■

15. (Eg 63/6) Niech X_1 będzie zmienną losową o rozkładzie jednostajnym na przedziale $(0, 1)$, X_2 zmienną losową o rozkładzie jednostajnym na przedziale $(0, X_1)$, X_3 zmienną losową o rozkładzie jednostajnym na przedziale $(0, X_2)$ i tak dalej. Niech N oznacza zmienną losową, taką że

$$\mathbf{P}(N = n) = \frac{\lambda^n}{n!(e^\lambda - 1)}, \quad \text{gdy } n = 1, 2, 3, ...,$$

gdzie $\lambda > 0$ jest ustaloną liczbą. Zmienna N jest niezależna od zmiennych $X_1, X_2, X_3, ...$ Wtedy $\mathbf{E}(N! X_1 \cdot X_2 \cdot ... \cdot X_N)$ jest równa
Odp: D-> $\frac{e^\lambda - \lambda - 1}{\lambda(e^\lambda - 1)}$.
Rozwiązanie. Zaważmy, że

$$\mathbf{E}(X_1...X_n) = \mathbf{E}(X_1...X_{n-1})\mathbf{E}(X_n|X_1, ..., X_{n-1}) = \mathbf{E}(X_1...X_{n-2})\frac{X_{n-1}^2}{2} =$$

$$= \mathbf{E}(X_1...X_{n-2})\mathbf{E}(\frac{X_{n-1}^2}{2}|X_1, ..., X_{n-2}) = ... = \frac{1}{(n+1)!}.$$

Zatem

$$\mathbf{E}N!(X_1...X_N) = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!(e^\lambda - 1)} \frac{1}{n+1} =$$
$$= \sum_{n=1}^{\infty} \frac{\lambda^n}{(n+1)!(e^\lambda - 1)} = \frac{e^\lambda - \lambda - 1}{\lambda(e^\lambda - 1)}.$$

$\blacksquare$