# 1 Zadania różne

W tym rozdziale znajdują się zadania nietypowe, często dotyczące łańcuchów Markowa oraz własności zmiennych losowych. Pojawią się także zadania z estymacji Bayesowskiej.

1. (Eg 48/3) Rozważamy łańcuch Markowa $X_1, X_2, ...,$ na przestrzeni stanów $\{1, 2, 3\}$ o macierzy przejścia
$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} \\ 0 & 1 & 0 \end{bmatrix},$$
gdzie $P_{i,j} = \mathbf{P}(X_{n+1} = j | X_n = i)$ dla $i, j = 1, 2, 3$. Załóżmy, że rozkład początkowy łańcucha jest wektorem
$$\pi = (\frac{2}{9}, \frac{4}{9}, \frac{1}{3}),$$
gdzie $\pi_i = \mathbf{P}(X_1 = i)$ dla $i = 1, 2, 3$. Oblicz $p = \mathbf{P}(X_1 = 1 | X_2 \neq 1 \vee X_3 \neq 1)$
*Odp:* B-> $\frac{1}{8}$.
**Rozwiązanie.** Najpierw zauważmy, że rozkład $\pi$ jest rozkładem stacjonarnym. Obliczamy
$$\mathbf{P}(X_1 = 1 | X_2 \neq 1 \vee X_3 \neq 1) = \frac{\mathbf{P}(X_3 \neq 1 \vee X_2 \neq 1 | X_1 = 1)\mathbf{P}(X_1 = 1)}{\mathbf{P}(X_3 \neq 1 \vee X_2 \neq 1)}.$$
Oczywiście $\mathbf{P}(X_1 = 1) = \pi_1 = \frac{2}{9}$, nadto
$$\mathbf{P}(X_3 \neq 1 \vee X_2 \neq 1 | X_1 = 1) = \mathbf{P}(X_3 = 3, X_2 = 2 | X_1 = 1) = P(3, 2)P(2, 1) = \frac{3}{4} \cdot \frac{1}{2} = \frac{3}{8}.$$
Obliczamy
$$\mathbf{P}(X_3 \neq 1 \vee X_2 \neq 1) = \mathbf{P}(X_3 = 3, X_2 = 2) + \mathbf{P}(X_3 = 2, X_2 = 3) = P(3, 2)\pi_2 + P(2, 3)\pi_3 =$$
$$= \frac{3}{4} \cdot \frac{4}{9} + 1 \cdot \frac{1}{3} = \frac{2}{3}$$
Stąd
$$\mathbf{P}(X_1 = 1 | X_2 \neq 1 \vee X_3 \neq 1) = \frac{1}{8}.$$
■

2. (Eg 50/9) Rozważamy łańcuch Markowa $X_1, X_2, ...,$ na przestrzeni stanów $\{0, 1, 2\}$ o macierzy przejścia
$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix},$$
gdzie $P_{i,j} = \mathbf{P}(X_{n+1} = j | X_n = i)$ dla $i, j = 0, 1, 2$. Niech $Z_1, Z_2, ..., Z_n, ...$ będzie ciągiem zmiennych losowych o wartościach w zbiorze $\{0, 1\}$ niezależnych od siebie nawzajem i od zmiennych $X_1, X_2, ..., X_n, ...$ o jednakowym rozkładzie prawdopodobieństwa:
$$\mathbf{P}(Z_i = 1) = \frac{3}{4} \text{ i } \mathbf{P}(Z_i = 0) = \frac{1}{4}.$$
Niech $Y_i = Z_i X_i$. Wtedy $\lim_{n\to\infty} \mathbf{P}(Y_n > Y_{n+1})$ jest równa
*Odp:* E-> $\frac{41}{144}$.
**Rozwiązanie.** Wyznaczamy rozkład stacjonarny łańcucha $X_1, X_2, ...,$ dostajemy $\pi = (\frac{10}{27}, \frac{8}{27}, \frac{9}{27})$.
Przechodząc do granicy otrzymamy
$$\lim_{n\to\infty} \mathbf{P}(Y_n > Y_{n+1}) = \mathbf{P}(Z_1 \cdot 0 > Z_2 X_2 | X_1 = 0)\pi_0 + \mathbf{P}(Z_1 \cdot 1 > Z_2 X_2 | X_1 = 1)\pi_1 +$$
$$+ \mathbf{P}(Z_1 \cdot 2 > Z_2 X_2 | X_1 = 2)\pi_2 = \mathbf{P}(Z_1 = 1)\mathbf{P}(Z_2 = 1)\mathbf{P}(X_2 = 0 | X_1 = 1)\pi_1 + \mathbf{P}(Z_1 = 1)\mathbf{P}(Z_2 = 0)\pi_1 +$$
$$+ \mathbf{P}(Z_1 = 1)\mathbf{P}(Z_2 = 1)\mathbf{P}(X_2 \in \{0, 1\} | X_1 = 2)\pi_2 + \mathbf{P}(Z_1 = 1)\mathbf{P}(Z_2 = 0)\pi_2 =$$
$$= \frac{9}{16} \cdot \frac{1}{4}$$

■

3. (Eg 51/6) Załóżmy, że $X_1, X_2, ...X_n$ są niezależnymi zmiennymi losowymi o jednakowym, ciągłym rozkładzie prawdopodobieństwa, mającymi momenty rzędu 1, 2 i 3. Znamy

$$\mu = \mathbf{E}(X_i) \ \text{ i } \ \sigma^2 = \mathbf{Var}(X_i).$$

Niech $f(x)$ oznacza gęstość rozkładu pojedynczej zmiennej $X_i$. Wiemy, że rozkład jest symetryczny w tym sensie, że $f(\mu + x) = f(\mu - x)$ dla każdego $x$. Oblicz trzeci moment sumy $\mathbf{E}(S_n^3)$, gdzie $S_n = X_1 + ... + X_n$.
*Odp:* C-> $n^2\mu(n\mu^2 + 3\sigma^2)$.

**Rozwiązanie.** Z faktu symetrii wynika, że

$$\mathbf{E}X_i = \mu, \ \ \mathbf{E}(X_i - \mu)^2 = \sigma^2, \ \ \mathbf{E}(X_i - \mu)^3 = 0.$$

Stąd również $\mathbf{E}(S_n - n\mu) = 0$, $\mathbf{E}(S_n - n\mu)^2 = n\sigma^2$, $\mathbf{E}(S_n - n\mu)^3 = 0$. Pozostaje obliczyć

$$\mathbf{E}(S_n^3) = \mathbf{E}(S_n - n\mu + n\mu)^3 = 3n^2\mu\sigma^2 + n^3\mu^3 = n^2\mu(n\mu^2 + 3\sigma^2).$$

■

4. (Eg 52/3) Niech $X_1, X_2, ..., X_n$ będą niezależnymi zmiennymi losowymi z rozkładu o gęstości

$$f_\theta(x) = \exp(-2|x - \theta|).$$

Niech $T_n = X_{[0,5n]:n}$, gdzie $[x]$ oznacza część całkowitą liczby $x$. Które z poniższych stwierdzeń jest prawdziwe?
*Odp:* A-> $\lim_{n\to\infty} \mathbf{P}(((T_n - \theta)\sqrt{n} > 1) = 0,023$.

**Rozwiązanie.** Zmienne $X_i$ mają rozkład o dystrybuancie

$$F(\theta + t) = \begin{cases} \frac{1}{2}(2 - e^{-2(t-\theta)}) & t \geqslant \theta \\ \frac{1}{2}e^{-2(\theta - t)} & t < \theta \end{cases}$$

Przypomnijmy, że $\mathbf{P}(T_n \leqslant t)$ jest takie same jak to, że zmienna $S_n(t)$ z rozkładu Bernoulliego $\mathcal{B}(n, F(t))$ będzie miała co najmniej $[0,5n]$ sukcesów. Obliczamy dla zmiennej $S_n = S_n(\theta + \frac{1}{\sqrt{n}})$

$$\mathbf{P}(((T_n - \theta)\sqrt{n} > 1) = \mathbf{P}(T_n > \theta + \frac{1}{\sqrt{n}}) = \mathbf{P}(S_n < [0,5n]) =$$

$$= \mathbf{P}(\frac{S_n - nF(\theta + \frac{1}{\sqrt{n}})}{\sqrt{nF(\theta + \frac{1}{\sqrt{n}})(1 - F(\theta + \frac{1}{\sqrt{n}}))}} < \frac{[0,5n] - \frac{n}{2}(2 - e^{-\frac{2}{\sqrt{n}}})}{\sqrt{\frac{n}{2}(1 - \frac{1}{2}e^{-\frac{2}{\sqrt{n}}})e^{-\frac{2}{\sqrt{n}}}}}).$$

Pozostaje obliczyć

$$\lim_{n\to\infty} \frac{[0,5n] - \frac{n}{2}(2 - e^{-\frac{2}{\sqrt{n}}})}{\sqrt{\frac{n}{2}(1 - \frac{1}{2}e^{-\frac{2}{\sqrt{n}}})e^{-\frac{2}{\sqrt{n}}}}} = -2.$$

Z drugiej strony z CTG wynika, że

$$\frac{S_n - nF(\theta + \frac{1}{\sqrt{n}})}{\sqrt{nF(\theta + \frac{1}{\sqrt{n}})(1 - F(\theta + \frac{1}{\sqrt{n}}))}} \to Z,$$

gdzie $Z$ ma rozkład $\mathcal{N}(0,1)$. Otrzymujemy wynik

$$\mathbf{P}(Z < -2) \simeq 0,23.$$

■

5. (Eg 53/7) Niech $X_1, X_2, ..., X_{13}$ będą niezależnymi zmiennymi losowymi o tym samym rozkładzie normalnym $\mathcal{N}(m, 1)$. Parametr $m$ jest nieznany i jest realizacją zmiennej losowej o rozkładzie normalnym $\mathcal{N}(1, 3)$. Wyznaczamy estymator bayesowski parametru $m$ przy funkcji straty LINEX danej wzorem

$$L(m, a) = e^{m-a} - (m - a) - 1,$$

gdzie $a$ oznacza wartość estymatora. Załóżmy, że w wyniku doświadczenia uzyskano próbkę losową taką, że $\sum_{i=1}^{13} X_i = 15$. Wtedy estymator bayesowski przyjmuje wartość
*Odp:* E-> $\frac{19}{16}$.

**Rozwiązanie.** W teorii decyzji statystycznej mamy do czynienia z regułami decyzyjnymi $\delta : \mathcal{X} \to \mathcal{A}$, dalej z funkcjami starty $L : \Theta \times \mathcal{A} \to \mathbb{R}$ oraz funkcjami ryzyka

$$R(\theta, \delta) = \mathbf{E}_\theta(\theta, \delta(X)).$$

W przypadku reguł bayesowskich mamy zadany rozkład a priori $\nu$ na przestrzeni parametrów $\Theta$. Dzięki temu można zdefiniować

$$r(\nu, \delta) = \int R(\theta, \delta)\nu(d\theta).$$

Dysponując powyższym funkcjonałem definiujemy optymalną regułę bayesowską $\delta_\mu$ jako argument minimum funkcji $r(\mu, \delta)$. Wyznaczenie optymalnej reguły bayesowskiej polega na skorzystaniu ze wzoru Fubiniego

$$\int_\Theta \int_\mathcal{X} L(\theta, \delta(x))\mu_\theta(dx)\nu(d\theta) = \int_\mathcal{X} \int_\Theta L(\theta, \delta)\nu_x(d\theta)\mu(dx),$$

gdzie miara $\mu_\theta$ jest rozkładem $X$ na $\mathcal{X}$ przy prawdopodobieństwie $\mathbf{P}_\theta$, nadto miary $\mu$ i $\nu_x(\theta)$ wyznacza się ze wzoru

$$\mu_\theta(dx)\nu(d\theta) = \nu_x(d\theta)\mu(dx).$$

Rozkład $\mu_x$ nazywa się rozkładem a posteriori. Dla każdego $x \in \mathcal{X}$ wybieramy wartość $\delta_\mu(x)$ jako argument minimum funkcji $f : \mathcal{A} \to \mathbb{R}$

$$f(\delta) = \int_\Theta L(\theta, \delta)\nu_x(d\theta).$$

Dla $L(\theta, a) = (\theta - a)^2$ estymatorem bayesowski jest wartość oczekiwana względem $\nu_x$, nadto dla $L(\theta, a) = |\theta - a|$ tym estymatorem jest mediana $\nu_x$. W przypadku funkcji LINEX $L(m, a) = e^{m-a} - (m - a) - 1$ obliczamy

$$f'(\delta) = \int_\Theta e^{m-\delta}\nu_x(dm) - 1$$

Czyli

$$\delta(x) = \log \int_\mathbb{R} e^m \nu_x(dm).$$

Należy zatem wyznaczyć rozkład a posteriori $\nu_x$. Mamy

$$\mu_m(dx) = (2\pi)^{-\frac{n}{2}} \exp(-\frac{1}{2}\sum_{i=1}^{n}(x_i - m)^2)dx_1...dx_n$$

dalej

$$\nu(dm) = (2\pi 3)^{-\frac{1}{2}} \exp(-\frac{1}{6}(m - 1)^2).$$

Stąd $\mu_m(dx)\nu(dm)$ jest rozkładem Gaussowskim, a zatem również $\nu_x$ma rozkład Gaussowski, co natychmiast pozwala wyznaczyć jego postać $\mathcal{N}(\frac{1+3\sum_{i=1}^n x_i}{1+3n}, \frac{3}{1+3n})$. Rozkład $\mu$ też jest Gaussowski i można wyznaczyć jego postać, nie ma to jednak znaczenia dla tego zadania. Pozostaje wyznaczyć

$$\int_{\mathbb{R}} e^m \nu_x(dm) = \exp(\frac{1+3\sum_{i=1}^n x_i}{1+3n} + \frac{3}{2(1+3n)}).$$

Stąd

$$\delta(x) = \frac{1+3\sum_{i=1}^n x_i}{1+3n} + \frac{3}{2(1+3n)} = \frac{5+6\sum_{i=1}^n x_i}{2(1+3n)}.$$

Podstawiając $n = 13$ oraz $\sum_{i=1}^{13} X_i = 13$ dostajemy

$$\delta(X) = \frac{19}{16}.$$

■

6. (Eg 54/6) O zmiennych $X_1, X_2, ..., X_n$ o tej samej wartości oczekiwanej równej $\mu$ oraz tej samej wariancji równej $\mu$ oraz tej samej wariancji równej $\sigma^2$ zakładamy, iż:

$$\mathbf{Cov}(X_i, X_j) = \rho\sigma^2 \text{ dla } i \neq j.$$

Zmienne losowe $\varepsilon_1, \varepsilon_2, ..., \varepsilon_n$ są nawzajem niezależne oraz niezależne od zmiennych losowych $X_1, X_2, ..., X_n$ i mają rozkłady prawdopodobieństwa postaci:

$$\mathbf{P}(\varepsilon_1 = 1) = \mathbf{P}(\varepsilon_i = \frac{1}{2}) = \mathbf{P}(\varepsilon_i = 0) = \frac{1}{3}.$$

Wariancja zmiennej losowej $S = \sum_{i=1}^n \varepsilon_i X_i$ jest równa.
*Odp:* A-> $\frac{n}{12}(5\sigma^2 + 2\mu^2 + 3(n-1)\rho\sigma^2)$.

**Rozwiązanie.** Obliczamy wariancję

$$\begin{aligned}\mathbf{Var}(S) &= \sum_{i=1}^n \mathbf{Var}(\varepsilon_i X_i) + \sum_{i\neq j} \mathbf{Cov}(\varepsilon_i X_i, \varepsilon_j X_j) = \\ &= n[\frac{5}{12}(\mu^2+\sigma^2) - \frac{1}{4}\mu^2] + n(n-1)[\frac{1}{4}\rho\sigma^2] = \\ &= \frac{n}{12}(5\sigma^2 + 2\mu^2 + 3(n-1)\rho\sigma^2).\end{aligned}$$

■

7. (Eg 55/1) Niech $X$ i $Y$ będą niezależnymi zmiennymi losowymi o rozkładach wykładniczych, przy czym $\mathbf{E}X = 4$, $\mathbf{E}Y = 6$. Rozważmy zmienną $Z = \frac{Y}{X+Y}$. Wtedy
*Odp:* B-> mediana rozkładu $Z$ jest równa $0,4$.

**Rozwiązanie.** Poszukujemy $C$ takiego, że

$$\mathbf{P}(\frac{Y}{X+Y} > C) = \frac{1}{2}.$$

stąd

$$\frac{1}{2} = \mathbf{P}((1-C)Y > CX) = \mathbf{EP}((1-C)Y > CX|X)) = \int_0^\infty \frac{1}{4} e^{-\frac{Cx}{6(1-C)}} e^{-\frac{x}{4}} dx = \frac{1}{1+\frac{2C}{3(1-C)}}.$$

Czyli

$$3(1-C) = 2C, \text{ zatem } C = 0,6.$$

■

8. (Eg 56/8) Cyfry $1, 2, 3, ..., 9$ ustawiamy losowo na miejscach o numerach $1, 2, 3, ..., 9$. Niech $X$będzie zmienną losową równą liczbie cyfr stojących na miejscach o numerach równych cyfrom. Wariancja zmiennej $X$ jest równa
*Odp:* B-> 1.

**Rozwiązanie.** Warto zapamiętać, że graniczna liczba koincydencji jest zmienną Poissona z parametrem 1, a więc i wariancję równą 1. W przypadku skończonym tego zadania korzystamy ze zmiennych włączeniowych $X = X_1 + ... + X_9$, gdzie $X_i$ przyjmuje wartość 1 jeśli $i$-ta cyfra stoi na swoim miejscu i 0 w przeciwnym przypadku. Jest jasne, że $\mathbf{P}(X_i = 1) = \frac{1}{9}$ oraz $\mathbf{P}(X_i = 1, X_j = 1) = \frac{1}{9} \cdot \frac{1}{8}$ zatem

$$\mathbf{Var}X_i = \frac{8}{9 \cdot 9}, \quad \mathbf{Cov}(X_i, X_j) = \frac{1}{8 \cdot 9 \cdot 9}.$$

Stąd

$$\mathbf{Var}(X) = 9\mathbf{Var}X_1 + 9 \cdot 8\mathbf{Cov}(X_1, X_2) = \frac{8}{9} + \frac{1}{9} = 1.$$

■

9. (Eg 57/2) Niech $X_1, X_2, ..., X_n$ będą niezależnymi zmiennymi losowymi z rozkładu jednostajnego na przedziale $(0, \theta)$, gdzie $\theta > 0$ jest nieznanym parametrem. Rozważamy estymator parametru $\theta$ postaci

$$T_n = (n + 1) \min\{X_1, X_2..., X_n\}.$$

Jeśli $\theta = 1$, to dla każdego $\varepsilon \in (0, 1)$ granica $\lim_{n\to\infty} \mathbf{P}(|T_n - 1| > \varepsilon)$ jest równa
*Odp:* B-> $1 - e^{\varepsilon-1} - e^{-\varepsilon-1}$.

**Rozwiązanie.** Najpierw wyznaczamy rozkład $T_n$, dla $0 < t < n + 1$

$$\mathbf{P}(T_n > t) = (1 - \frac{t}{n+1})^n \to e^{-t}.$$

Stąd dla $T$ o rozkładzie wykładniczym

$$\lim_{n\to\infty} \mathbf{P}(|T_n - 1| > \varepsilon) = \mathbf{P}(|T - 1| > \varepsilon) = \mathbf{P}(T > 1 + \varepsilon) + \mathbf{P}(T < 1 - \varepsilon) =$$
$$= e^{-\varepsilon-1} + 1 - e^{\varepsilon-1}.$$

■

10. (Eg 57/9) Zmienne losowe $X_1, X_2, ..., X_n, ...$ są niezależnymi o jednakowym rozkładzie

$$\mathbf{P}(X_n = 1) = \mathbf{P}(X_n = 2) = \mathbf{P}(X_n = 3) = \mathbf{P}(X_n = 4) = \frac{1}{4}.$$

Niech $Y_0 = 2$ oraz niech dla $n = 1, 2, 3, ...$ zachodzi

$$Y_n = \begin{cases} 4 & \text{gdy} \quad X_n = 4 \\ \min(Y_{n-1}, X_n) & \text{gdy} \quad X_n < 4 \end{cases}$$

Oblicz $\lim_{n\to\infty} \mathbf{P}(Y_n \leqslant 3)$
*Odp:* B-> $\frac{3}{4}$.

**Rozwiązanie.** Nietrudno zauważyć, że $(Y_n)_{n=0}^{\infty}$ jest łańcuchem Markowa o macierzy przejścia

$$P = \begin{bmatrix} \frac{3}{4} & 0 & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{2}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

Obliczamy rozkład graniczny ze wzoru $\pi = \pi P$, zachodzi $\pi = (\frac{1}{2}, \frac{1}{6}, \frac{1}{12}, \frac{1}{4})$. Zatem

$$\lim_{n\to\infty} \mathbf{P}(Y_n \leqslant 3) = \pi + 1 + \pi + \pi_3 = \frac{3}{4}.$$

■

11. (Eg 58/1) Niech $X_1, X_2, ..., X_n, ...$ będą zmiennymi losowymi o tym samym rozkładzie ujemnym dwumianowym

$$\mathbf{P}_\theta(X_i = k) = (k+1)\theta^2(1-\theta)^k, \quad k = 0, 1, 2, ..., \quad i = 1, 2, ..., n+1,$$

gdzie $\theta \in (0,1)$ jest nieznanym parametrem. Zmienne $X_1, X_2, ..., X_n, X_{n+1}$ są warunkowo niezależne przy danym $\theta$. Załóżmy, że rozkład a priori parametru $\theta$ jest rozkładem o gęstości

$$\pi(\theta) = 12\theta^2(1-\theta), \quad \text{gdy} \quad \theta \in (0,1).$$

Na podstawie próby losowej $X_1, X_2, ..., X_n$ wyznaczamy predyktor bayesowski. zmiennej $X_{n+1}$ przy kwadratowej funkcji straty. Wariancja tego predykatora jest równa
*Odp:* D-> $\frac{8n}{n+1}$.
**Rozwiązanie.** Podstawową wiedzą z teorii warunkowych wartości oczekiwanych jest, że przy kwadratowej funkcji straty najlepszym estymatorem $X_{n+1}$ jest $\mathbf{E}(X_{n+1}|X_1, ..., X_n)$, gdzie wartość oczekiwana oznacza całkowanie względem miary $\mathbf{P}_\theta\mu(d\theta)$. Oczywiście z warunkowej niezależności $X_1, .., X_n$ pod warunkiem $\Theta = \theta$ dostajemy

$$\begin{aligned} &\mathbf{E}(X_{n+1}|X_1, ...., X_n) = \mathbf{E}(\mathbf{E}(X_{n+1}|\Theta, X_1, ..., X_n)|X_1, ..., X_n) = \\ &= \mathbf{E}(2\frac{1-\Theta}{\Theta}|X_1, ..., X_n). \end{aligned}$$

Musimy wyznaczyć rozkład warunkowy $\Theta$ pod warunkiem $X_1 = k_1, ..., X_n = k_n$. Najpierw wyprowadzamy gęstość rozkładu łącznego

$$f(\theta, k) = 12\theta^{2(n+1)}(1-\theta)^{1+\sum_{i=1}^n k_i} \prod_{i=1}^n (k_i + 1).$$

Pozostaje wyznaczyć, gęstość $f(\theta|k)$. Nietrudno zauważyć, że jest rozkład $Beta(2n+3, 2+\sum_{i=1}^n k_i)$. Dla zmiennej $Z$ z rozkładu $Beta(\alpha, \beta)$ oraz $\alpha > 1$ wartość $\mathbf{E}Z^{-1} = \frac{\alpha+\beta-1}{\alpha-1}$, $\mathbf{E}Z^{-2} = \frac{(\alpha+\beta-1)(\alpha+\beta-2)}{(\alpha-1)(\alpha-2)}$.
Stąd

$$\mathbf{E}(2\frac{1-\Theta}{\Theta}|X_1, ..., X_n) = 2 \cdot \frac{2 + \sum_{i=1}^n X_i}{2(n+1)} =: T.$$

Nadto $X_1 + ... + X_n$ pod warunkiem $\Theta = \theta$ ma rozkład ujemny dwumianowy $\mathcal{B}_-(2n, \theta)$. Zatem

$$\mathbf{Var}T = \mathbf{E}\mathbf{Var}(T|\Theta) + \mathbf{Var}\mathbf{E}(T|\Theta).$$

Z własności rozkładu ujemnego dwumianowego

$$\mathbf{E}(T|\Theta) = \frac{2n}{n+1}\frac{1}{\Theta} - \frac{2(n-1)}{n+1}$$

oraz

$$\mathbf{Var}(T|\Theta) = \frac{8n}{4(n+1)^2} \cdot \frac{1-\Theta}{\Theta^2}.$$

Korzystając z własności rozkładu $Beta(3,2)$ obliczamy

$$\mathbf{Var}\mathbf{E}(T|\Theta) = 8\frac{n^2}{(n+1)^2}$$

nadto

$$\mathbf{EVar}(T|\Theta) = 8\frac{n}{(n+1)^2}$$

Podsumowując $\mathbf{Var}T = \frac{8n}{n+1}$.

$\blacksquare$

12. (Eg 58/8) Załóżmy, że $W_1, W_2, ..., W_n, ...$ jest ciągiem zmiennych losowych takim, że

   - zmienna $W_1$ ma rozkład jednostajny na przedziale $(0,1)$,
   - dla każdej liczby naturalnej $n$ zmienna losowa $W_{n+1}$ warunkowo przy danych $W_1, W_2, ..., W_n$ ma gęstość

$$f(w_{n+1}|w_1, ..., w_n) = \begin{cases} 1 & \text{gdy} \quad w_n \leqslant 0,5 \\ 3x^2 & \text{gdy} w_n > 0,5 \end{cases} \quad \text{dla } w_{n+1} \in (0,1).$$

   Wtedy $\lim_{n\to\infty} \mathbf{P}(W_n > 0,25)$ jest równa
   *Odp:* B-> $\frac{15}{16}$.

   **Rozwiązanie.** Nietrudno zauważyć, że mamy do czynienia z jednorodnym łańcuchem Markowa zdanym przez funkcję przejścia

$$P(x, A) = |A| 1_{x \leqslant \frac{1}{2}} + \int_A 3y^2 dy 1_{x > \frac{1}{2}}.$$

   Poszukujemy rozkładu stacjonarnego $\pi$ na $[0,1]$ takiego, że

$$\pi(A) = |A| \pi([0, \frac{1}{2}]) + \int_A 3y^2 dy \pi((\frac{1}{2}, 1])$$

   Stąd natychmiast wynika, że $\pi$ jest absolutnie ciągłą względem miary Lebsegue'a której gęstość $f$ spełnia warunek

$$f(x) = \pi([0, \frac{1}{2}]) 1_{0 \leqslant x \leqslant 1} + \pi(\frac{1}{2}, 1]) 3x^2 1_{0 \leqslant x \leqslant 1}.$$

   Współczynniki $a = \pi([0, \frac{1}{2}])$, $b = \pi((\frac{1}{2}, 1])$ wyznaczamy ze wzorów

$$\begin{cases} a &= a\frac{1}{2} + b\frac{1}{8} \\ 1 &= a + b \end{cases}$$

   Stąd $4a = b$ oraz $a = \frac{1}{5}$, $b = \frac{4}{5}$. Obliczamy

$$\lim_{n\to\infty} \mathbf{P}(W_n > \frac{1}{4}) = \pi((\frac{1}{4}, 1]) = \\ = \frac{1}{5}\frac{3}{4} + \frac{4}{5}(1 - \frac{1}{64}) = \frac{3}{20} + \frac{63}{80} = \frac{75}{80} = \frac{15}{16}.$$

$\blacksquare$

13. (Eg 59/4) Dysponujemy dwiema urnami. W urnie I mamy dwie kule białe i jedną czarną, w urnie II mamy trzy kule białe i trzy czarne. Powtarzamy $n$ razy eksperyment polegający na tym, że losujemy jedną kulę z urny I, nie oglądając jej wkładamy ją do urny II, następnie losujemy jedną kulę z urny II i nie oglądając jej wkładamy ją do urny I. Niech $X_n$ oznacza zmienną losową równą liczbie kul białych w urnie I po $n$ doświadczeniach. Wtedy $\lim_{n\to\infty} \mathbf{E}(X_n X_{n+1})$ jest równa
   *Odp:* C-> $\frac{65}{21}$.

   **Rozwiązanie.** Ponownie korzystamy z teorii łańcuchów Markowa. Pod długim czasie rozkład kul będzie się stabilizował, aby wyznaczyć rozkład graniczny piszemy macierz przejścia dla liczby kul w I urnie

$$P = \begin{bmatrix} S & 0 & 1 & 2 & 3 \\ 0 & \frac{2}{7} & \frac{5}{7} & 0 & 0 \\ 1 & \frac{2}{21} & \frac{11}{21} & \frac{8}{21} & 0 \\ 2 & 0 & \frac{6}{21} & \frac{12}{21} & \frac{3}{21} \\ 3 & 0 & 0 & \frac{4}{7} & \frac{3}{7} \end{bmatrix}$$

Rozwiązujemy układ równań

$$\begin{cases} \pi_0 & = & \frac{2}{7}\pi_0 + \frac{2}{21}\pi_1 \\ \pi_1 & = & \frac{5}{7}\pi_0 + \frac{11}{21}\pi_1 + \frac{6}{21}\pi_2 \\ \pi_3 & = & \frac{3}{21}\pi_2 + \frac{3}{7}\pi_3 \\ 1 & = & \pi_0 + \pi_1 + \pi_2 + \pi_3 \end{cases}$$

którego rozwiązaniem jest $\pi_0 = \frac{1}{21}, \pi_1 = \frac{15}{42}$, $\pi_2 = \frac{10}{21}$, $\pi_3 = \frac{5}{42}$. Zatem

$$\lim_{n\to\infty} \mathbf{E}(X_n X_{n+1}) = \mathbf{E}_\pi(X_0 X_1) = \sum_{k=0}^{3}\sum_{l=0}^{3} klP(k,l)\pi_k.$$

Czyli

$$\lim_{n\to\infty} \mathbf{E}(X_n X_{n+1}) = 1\cdot 1\cdot \frac{11}{21}\cdot\frac{15}{42} + +1\cdot 2\frac{8}{21}\cdot\frac{15}{42} + 2\cdot 1\cdot\frac{6}{21}\cdot\frac{10}{21} +$$
$$+ 2\cdot 2\cdot\frac{12}{21}\cdot\frac{10}{21} + 2\cdot 3\cdot\frac{3}{21}\cdot\frac{10}{21} + 3\cdot 2\cdot\frac{4}{7}\cdot\frac{5}{42} + 3\cdot 3\cdot\frac{3}{7}\frac{5}{42} = \frac{65}{21}.$$

■

14. (Eg 60/10) Niech $X_1, X_2, ..., X_n$, $n > 2$, będą niezależnymi zmiennymi losowymi z rozkładu Pareto o gęstości

$$f(x) = \frac{3}{(1+x)^4} 1_{x>0}.$$

Niech $U = \min\{X_0, X_1, X_2, ..., X_n\}$. Wtedy $\mathbf{Cov}(U, X_0)$ jest równa
*Odp:* C-> $\frac{3}{2(3n+1)(3n+2)}$.

**Rozwiązanie.** Niech $X$ ma rozkład taki jak $X_0, X_1, X_2, ..., X_n$. Mamy

$$\mathbf{P}(X > t) = \frac{1}{(1+t)^3}, \quad t \geqslant 0.$$

Stąd $\mathbf{E}X_0 = \int_0^\infty \mathbf{P}(X > t)dt = \frac{1}{2}$. Nadto

$$\mathbf{P}(U > t) = \frac{1}{(1+t)^{3n}}, \quad t \geqslant 0.$$

Zatem $\mathbf{E}U = \int_0^\infty \mathbf{P}(X > t)dt = \frac{1}{3n+2}$. Pozostaje obliczyć

$$\mathbf{E}UX_0 = \mathbf{E}X_0\mathbf{E}(U|X_0) = \mathbf{E}X_0 \int_0^{X_0} \frac{1}{(1+t)^{3n}}dt =$$
$$= \frac{1}{3n-1}\mathbf{E}X_0(1 - \frac{1}{(1+X_0)^{3n-1}}) = \frac{1}{2(3n-1)} - \frac{1}{3n-1}\int_0^\infty \frac{3t}{(1+t)^{3n+3}}dt.$$

Do policzenia występujących powyżej wartości oczekiwanych najprościej użyć podstawienia $x = \frac{1}{1+t}$

$$\int_0^\infty \frac{t}{(1+t)^{3n+3}}dt = \int_0^1 x^{3n}(1-x)dx = \frac{\Gamma(3n+1)\Gamma(2)}{\Gamma(3n+3)} =$$
$$\frac{1}{(3n+1)(3n+2)}.$$

Zatem

$$\mathbf{Cov}(U, X_0) = \frac{3}{3n-1}(\frac{1}{2} - \frac{1}{(3n+1)(3n+2)}) - \frac{1}{2(3n+2)} =$$
$$= \frac{9n-3}{2(3n+2)(9n^2-1)} = \frac{3}{2(3n+1)(3n+2)}.$$

■

15. (Eg 61/2) Niech zmienna losowa $S_n$ będzie liczbą sukcesów w $n$ ($n > 1$) próbach Bernoulliego z prawdopodobieństwem sukcesu $p$. O zdarzeniu losowym A wiemy, że

$$\mathbf{P}(A|S_n = k) = a\frac{k}{n} \text{ dla } k = 0, 1, 2, ..., n,$$

gdzie $a$ jest znaną liczbą $0 < a \leqslant 1$. Oblicz $\mathbf{E}(S_n|A)$.
*Odp:* A-> $pn + 1 - p$.

**Rozwiązanie.** Obliczamy

$$\mathbf{P}(A) = \sum_{k=0}^{n} \mathbf{P}(A|S_n = k)\mathbf{P}(S_n = k) = \sum_{k=0}^{n} a\frac{k}{n} \cdot \binom{n}{k} p^k (1-p)^{n-k} = ap.$$

Zatem korzystając z własności rozkładu Bernoulliego

$$\mathbf{E}(S_n|A) = \mathbf{P}(A)^{-1} \sum_{k=0}^{n} k\mathbf{P}(\{S_n = k\} \cap A) = (ap)^{-1} \sum_{k=0}^{n} k\mathbf{P}(A|S_n = k)\mathbf{P}(S_n = k) =$$
$$= (ap)^{-1} \sum_{k=0}^{n} a\frac{k^2}{n} \binom{n}{k} p^k (1-p)^{n-k} = (1-p) + np.$$

■