Zadanie 1.

O niezależnych zmiennych losowych X i Y wiadomo, że

$$\mathrm{Var}X = 1,\ \mathbb{E}X = 1,\ \mathrm{Var}Y = 2,\ \mathbb{E}Y = 2.$$

Ile wynosi $\mathrm{Var}(XY)$?

$$\mathrm{Var}(XY) = E(XY)^2 - (EXY)^2 = EX^2\, EY^2 - (EX)^2(EY)^2$$

$$EX^2 = \mathrm{Var}(X) + (EX)^2 = 1+1 = 2$$

$$EY^2 = \mathrm{Var}(Y) + (EY)^2 = 2+4 = 6$$

$$\mathrm{Var}(XY) = 2\cdot 6 - 1\cdot 4 = 12-4 = 8$$

Zadanie 2.

Zmienna losowa X ma rozkład jednostajny $\mathcal{U}(0,1000)$. Definiujemy $Y = \min(X,800)$. Ile wynosi $\mathbb{E}Y$?

$$X \sim U(0,1000)$$

$$Y = \min(X,800) = \begin{cases} X, & X < 800 \\ 800, & X > 800 \end{cases}$$

$$EY = E[\min(X,800)] = \frac{1}{1000}\left(\int_0^{800} x\,dx + \int_{800}^{1000} 800\,dx\right) =$$

$$= \frac{1}{1000}\left(\frac{x^2}{2}\Big|_0^{800} + 800x\Big|_{800}^{1000}\right) = 480$$

Zadanie 3.

Załóżmy, że niezależne obserwacje $X_1, \ldots, X_{16}$ pochodzą z rozkładu $\mathcal{N}(\mu, \sigma^2)$. Definiujemy statystyki

$$\bar{X} = \frac{1}{16}\sum_{i=1}^{16} X_i, \qquad S^2 = \frac{1}{15}\sum_{i=1}^{16}(X_i - \bar{X})^2.$$

Ile wynosi $\mathbb{P}(\bar{X} > \mu | S^2 > \sigma^2)$?

Oznaczenia:

- $F_{\chi^2_d}(x)$ to dystrybuanta rozkładu χ^2 z d stopniami swodoby w punkcie x
- $F_{t,d}(x)$ to dystrybuanta rozkładu t-Studenta z d stopniami swodoby w punkcie x
- $\Phi(x)$ to dystrybuanta standardowego rozkładu normalnego w punkcie x

<u>TW</u> (Tw. Fischera) W modelu normalnym, $\bar{X}$ i S^2 są niezależnymi zm. los.,

$\bar{X} \sim N(\mu, \frac{1}{n}\sigma^2)$

$\frac{n-1}{\sigma^2} S^2 \sim \chi^2(n-1)$

$P(\bar{X} > \mu \mid S^2 > \sigma^2) = P(\bar{X} > \mu) = \frac{1}{2}$

Zadanie 4.

Jednorodny (w czasie) łańcuch Markowa $(X_1, X_0, \ldots)$ na przestrzeni stanów $\mathcal{S} = \{1, 2\}$ ma macierz przejścia

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

Ile wynosi $\lim_{n\to\infty} \mathbb{E}\left(\frac{X_n}{X_{n+1}}\right)$?

Jaki rozkład ma $\frac{X_n}{X_{n+1}}$

Ta zm. ma możliwe 3 wartości $\frac{1}{2}, 1, 2$:

$P(\frac{X_n}{X_{n+1}} = \frac{1}{2}) = P(X_n = 1, X_{n+1} = 2) = P(X_n = 1)\, P(1 \to 2) = P(X_n = 1) \cdot \frac{2}{3}$

$P(\frac{X_n}{X_{n+1}} = 1) = P(X_n = 1, X_{n+1} = 1) + P(X_n = 2, X_{n+1} = 2) =$

$= P(X_n = 1)\, P(1 \to 1) + P(X_n = 2)\, P(2 \to 2) =$

$= P(X_n = 1) \cdot \frac{1}{3} + 0 = P(X_n = 1) \cdot \frac{1}{3}$

$P(\frac{X_n}{X_{n+1}} = 2) = P(X_n = 2, X_{n+1} = 1) = P(X_n = 2)\, P(2 \to 1) = P(X_n = 2)$

$$E\left(\frac{X_n}{X_{n+1}}\right) = \frac{1}{2}\cdot\frac{2}{3}P(X_n=1) + 1\cdot\frac{1}{3}P(X_n=1) + 2P(X_n=2) =$$

$$= \frac{2}{3}P(X_n=1) + 2P(X_n=2)$$

$$\pi_j = \lim_{n\to\infty} P(X_n = j)$$

Teraz należy skorzystać z tw. ergodycznego:

$$\lim_{n\to\infty} E\left(\frac{X_n}{X_{n+1}}\right) = \lim_{n\to\infty} \frac{2}{3}P(X_n=1) + 2P(X_n=2) = [\text{tw. ergodyczne}] =$$

$= \frac{2}{3}\pi_1 + 2\pi_2$, gdzie $[\pi_1, \pi_2]$ jest rozkładem stacjonarnym

Równania na rozkład stacjonarny wyglądają następująco:

$$[\pi_1 \ \ \pi_2] = [\pi_1 \ \ \pi_2]P = [\pi_1 \ \ \pi_2]\begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ 1 & 0 \end{bmatrix} = \left[\frac{1}{3}\pi_1 + \pi_2 \quad \frac{2}{3}\pi_1\right]$$

Mamy też: $\pi_1 + \pi_2 = 1$

$$\frac{2}{3}\pi_1 = \pi_2$$

$$\pi_1 = \frac{3}{5}, \quad \pi_2 = \frac{2}{5}$$

Wstawiamy do wzoru

$$\lim_{n\to\infty} E\left(\frac{X_n}{X_{n+1}}\right) = \frac{2}{3}\pi_1 + 2\pi_2 = \frac{2}{3}\cdot\frac{3}{5} + 2\cdot\frac{2}{5} = \frac{6}{5}$$

Zadanie 5.

X_1, X_2, X_3 są niezależnymi zmiennymi losowymi o rozkładach wykładniczych z wartościami oczekiwanymi $\mathbb{E}X_i = \frac{1}{i}, i = 1,2,3.$

Ile wynosi $\mathbb{P}(X_1 = \min(X_1, X_2, X_3))$?

$X_1, \ldots, X_n$ są zm. los. wykładniczymi niezależnymi z parametrami $\lambda_1, \ldots, \lambda_n$, to:

- $\min(X_1, \ldots, X_n)$ ma rozkład $Exp(\lambda_1 + \ldots + \lambda_n)$
- $P(X_i = \min(X_1, \ldots, X_n)) = \frac{\lambda_i}{\lambda_1 + \ldots + \lambda_n}$

$P(X_1 > X_i, X_2 > X_i, \ldots, X_n > X_i) =$

$= \int_0^\infty P(X_1 > x, \ldots [\text{bez } X_i] \ldots, X_n > x)\, p(X_i = x)\, dx =$

$= \int_0^\infty e^{-\lambda_1 x} \cdot \ldots [\text{bez } X_i] \ldots \cdot e^{-\lambda_n x} \cdot \lambda_i e^{-\lambda_i x}\, dx =$

$= \frac{\lambda_i}{\lambda_1 + \ldots + \lambda_n}$

$X_1 \sim Exp(1)$

$\min(X_1, X_2, X_3) \sim Exp(6)$

$P(X_1 = \min(X_1, X_2, X_3)) = \frac{1}{6}$

Zadanie 6.

Niech $X_1, X_2, \ldots, X_n, n \geq 3$ będą niezależnymi zmiennymi losowymi o rozkładzie $\mathcal{U}(a,b)$, tj. o rozkładzie jednostajnym na odcinku (a,b), gdzie $a < b$. Dla jakiego α estymator

$$\hat{\theta} = \alpha\left[\max(X_1,\ldots,X_n) - \min(X_1,\ldots,X_n)\right]$$

jest nieobciążonym estymatorem parametru $\theta = b - a$?

$X \sim U(a,b) \qquad F_X(x) = \frac{x-a}{b-a}$

$\hat{\theta} = \alpha(Z - Y)$

$F_Z(x) = \left(\frac{x-a}{b-a}\right)^n$

$F_Y(x) = 1 - \left(1 - \frac{x-a}{b-a}\right)^n = 1 - \left(\frac{b-a-x+a}{b-a}\right)^n = 1 - \left(\frac{b-x}{b-a}\right)^n$

$EZ = \int_a^b 1 - \left(\frac{x-a}{b-a}\right)^n dx = x - \frac{b-a}{n+1}\left(\frac{x-a}{b-a}\right)^{n+1}\Big|_a^b =$

$= b - \frac{b-a}{n+1} - a = (b-a)\left(1 - \frac{1}{n+1}\right) = (b-a)\left(\frac{n}{n+1}\right)$

$EY = \int_a^b 1 - 1 + \left(\frac{b-x}{b-a}\right)^n dx = -\frac{b-a}{n+1}\left(\frac{b-x}{b-a}\right)^{n+1}\Big|_a^b =$

$= \frac{b-a}{n+1}$

$\alpha\left[(b-a)\left(\frac{n}{n+1}\right) - \left(\frac{b-a}{n+1}\right)\right] = b - a$

$\alpha\left(\frac{n-1}{n+1}\right) = 1$

$\alpha = \frac{n+1}{n-1}$

Zadanie 7.

Ustalmy $n \geq 4$. Niech $X_1, X_2, \ldots, X_n$ będą niezależnymi zmiennymi losowymi o rozkładzie $\mathcal{N}(\mu, \sigma_1^2)$, a $Y_1, \ldots, Y_n$ niech będą niezależnymi (od siebie i od ciągu $X_1, \ldots, X_n$) zmiennymi losowymi o rozkładzie $\mathcal{N}(\mu, \sigma_2^2)$. Wówczas dla dowolnego $\alpha \in (0,1)$ estymator

$$\hat{\mu} = \alpha \bar{X} + (1-\alpha)\bar{Y},$$

gdzie

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i,$$

jest nieobciążonym estymatorem parametru μ. Dla jakiego α średniokwadratowy błąd estymatora, tj. $\mathbf{E}(\hat{\mu}-\mu)^2$, jest najmniejszy?

$$E(\hat{\mu}-\mu)^2 = E(\hat{\mu}^2 - 2\hat{\mu}\mu + \mu^2) = E\hat{\mu}^2 - 2\mu E\hat{\mu} + \mu^2$$

$$E\hat{\mu} = E(\alpha \bar{X} + (1-\alpha)\bar{Y}) = \alpha E\bar{X} + (1-\alpha)E\bar{Y}$$

$$E\hat{\mu}^2 = E(\alpha\bar{X} + (1-\alpha)\bar{Y})^2 = E(\alpha^2\bar{X}^2 + 2\alpha(1-\alpha)\bar{X}\bar{Y} + (1-\alpha)^2\bar{Y}^2) =$$

$$= \alpha^2 E\bar{X}^2 + 2\alpha(1-\alpha)E\bar{X}\,E\bar{Y} + (1-\alpha)^2 E\bar{Y}^2$$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\mu, \tfrac{1}{n}\sigma_1^2)$$

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \sim N(\mu, \tfrac{1}{n}\sigma_2^2)$$

$$E\hat{\mu} = \alpha\mu + (1-\alpha)\mu$$

$$E\hat{\mu}^2 = \alpha^2(\mu^2 + \tfrac{1}{n}\sigma_1^2) + (2\alpha - 2\alpha^2)\mu^2 + (1-2\alpha+\alpha^2)(\mu^2 + \tfrac{1}{n}\sigma_2^2)$$

$$E(\hat{\mu}-\mu)^2 = \alpha^2(\mu^2 + \tfrac{1}{n}\sigma_1^2) + (2\alpha - 2\alpha^2)\mu^2 + (1-2\alpha+\alpha^2)(\mu^2 + \tfrac{1}{n}\sigma_2^2) -$$
$$- 2\alpha\mu^2 - 2(1-\alpha)\mu^2 + \mu^2$$

$$E(\hat{\mu}-\mu)^2 = 2\alpha(\mu^2 + \tfrac{1}{n}\sigma_1^2) + (2-4\alpha)\mu^2 + (2\alpha-2)(\mu^2 + \tfrac{1}{n}\sigma_2^2) -$$
$$- \cancel{2\mu^2} + \cancel{2\mu^2} =$$
$$= \cancel{2\alpha\mu^2} + 2\alpha\sigma_1^2\tfrac{1}{n} + \cancel{2\mu^2} - \cancel{4\alpha\mu^2} + \cancel{2\alpha\mu^2} + 2\alpha\sigma_2^2\tfrac{1}{n} - \cancel{2\mu^2} -$$
$$- 2\sigma_2^2\tfrac{1}{n}$$

$$\cancel{2}\alpha\sigma_1^2\cancel{\tfrac{1}{n}} + \cancel{2}\alpha\sigma_2^2\cancel{\tfrac{1}{n}} = \cancel{2}\sigma_2^2\cancel{\tfrac{1}{n}}$$

$$\alpha(\sigma_1^2 + \sigma_2^2) = \sigma_2^2$$

$$\alpha = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

Druga metoda

$$Var(\hat{\mu}) = Var(\alpha\bar{X} + (1-\alpha)\bar{Y}) = \alpha^2 Var(\bar{X}) + (1-\alpha)^2 Var(\bar{Y}) =$$

$$= \alpha^2 \frac{1}{n}\sigma_1^2 + (1 - 2\alpha + \alpha^2)\frac{1}{n}\sigma_2^2$$

$$Var'(\hat{\mu}) = 2\alpha\frac{1}{n}\sigma_1^2 + (-2 + 2\alpha)\frac{1}{n}\sigma_2^2 = 0$$

$$2\alpha\sigma_1^2 - 2\sigma_2^2 + 2\alpha\sigma_2^2 = 0$$

$$\alpha(\sigma_1^2 + \sigma_2^2) = \sigma_2^2$$

$$\alpha = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

Zadanie 8.

Zmienne losowe $X_1, \ldots, X_n$ przy danej wartości parametru $\Theta = \theta \in (0,1)$ są warunkowo niezależne i mają rozkład

$$\begin{aligned} P(X_i = 1|\Theta = \theta) &= \theta, \\ P(X_i = 0|\Theta = \theta) &= 1 - \theta. \end{aligned}$$

Z kolei zmienna losowa Θ ma rozkład o gestości

$$f_\Theta(\theta) = 3(1-\theta)^2, \quad \theta \in (0,1).$$

Oznaczmy $S_n = \sum_{i=1}^{n} X_i$. Ile wynosi $\mathbb{P}(S_4 > 0|S_2 = 0)$?

$$P(S_4 > 0 | S_2 = 0) = P(X_1 + X_2 + X_3 + X_4 > 0 \mid X_1 + X_2 = 0) =$$

$$= P(X_3 + X_4 > 0 \mid X_1 + X_2 = 0) =$$

$$= P(X_3 + X_4 = 2 \mid X_1 + X_2 = 0) + P(X_3 + X_4 = 1 \mid X_1 + X_2 = 0)$$

Rozwiązanie opiera się na fakcie, że do liczenia p-stwa warunkowego $P(A|B)$ dla warunkowo niezależnych zdarzeń / zmiennych A, B używamy rozkładu aposteriori (czyli zaktualizowanego) parametru $\Theta|B$.

Wzór na warunkową niezależność:

$$P(A, B \mid C) = P(A \mid C) P(B \mid C)$$

$$P(A \mid B) = \frac{P(A,B)}{P(B)} = \frac{1}{P(B)} P(A,B) =$$

$$= \frac{1}{P(B)} \cdot \int_0^1 P(A, B \mid \Theta = \theta) f_\Theta(\theta)\, d\theta = \left| \begin{matrix} \text{warunkowa} \\ \text{niezależność} \end{matrix} \right| =$$

$$= \frac{1}{P(B)} \int_0^1 P(A \mid \Theta = \theta) P(B \mid \Theta = \theta) f_\Theta(\theta)\, d\theta =$$

$$= \int_0^1 \frac{P(A \mid \Theta = \theta) P(B, \Theta = \theta)}{P(B)}\, d\theta =$$

$$= \int_0^1 P(A \mid \Theta = \theta) f_{\Theta \mid B}(\theta)\, d\theta$$

Rozkład sumy $X_i + X_j$:

$X_i + X_j$	0	1	2
p_i	$(1-\theta)^2$	$2\theta(1-\theta)$	θ^2

Wracamy do zadania. Rozkład aposteriori w naszym przypadku wynosi (ze wzoru Bayesa):

$$f_{\Theta \mid X_1 + X_2 = 0} = c \cdot P(X_1 + X_2 = 0 \mid \Theta = \theta) f_\Theta(\theta) =$$

$$= c(1-\theta)^2 \cdot 3(1-\theta)^2 = c(1-\theta)^4 = \left| \text{stała } c \text{ dobierany tak aby gęstość całkowała się do } 1 \right| =$$

$$= 5(1-\theta)^4$$

Mamy:

$$P(X_3 + X_4 = 2 \mid X_1 + X_2 = 0) = \int_0^1 \theta^2 \cdot 5(1-\theta)^4\, d\theta = \frac{1}{7} \cdot \frac{1}{3}$$

Oraz

$$P(X_3 + X_4 = 1 \mid X_1 + X_2 = 0) = \int_0^1 2\theta(1-\theta) \cdot 5(1-\theta)^4\, d\theta = \frac{1}{7} \cdot \frac{5}{3}$$

Ostatecznie:

$$p = \frac{1}{7} \cdot \frac{1}{3} + \frac{1}{7} \cdot \frac{5}{3} = \frac{2}{7}$$

Zadanie 9.

Zaobserwowano $x_1, x_2, \ldots, x_n$, niezależną próbkę z rozkładu Poissona z nieznaną średnią $\lambda > 0$. Oznaczmy $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. Niech $z_{1-\alpha/2}$ oznacza kwantyl rzędu $1-\alpha/2$ standardowego rozkładu normalnego, tj.

$$z_{1-\alpha/2} = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right),$$

gdzie $\Phi(\cdot)$ to dystrybuanta standardowego rozkładu normalnego $\mathcal{N}(0,1)$.

Asymptotyczny przedział ufności dla estymatora λ wyznaczonego metodą największej wiarogodności z wykorzystaniem informacji Fishera jest postaci:

Szukany przedział to $\hat{\theta} \pm c\left(nI(\hat{\theta})\right)^{-\frac{1}{2}}$

MLE:

$$f_X(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

$$L = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = \prod_{i=1}^{n} \frac{1}{x_i!} \cdot \lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda}$$

$$l = \ln\left(\prod_{i=1}^{n} \frac{1}{x_i!}\right) + \sum_{i=1}^{n} x_i \cdot \ln(\lambda) - n\lambda$$

$$l' = \frac{\sum_{i=1}^{n} x_i}{\lambda} - n$$

$$\hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\hat{\lambda} = \bar{x}$$

Informacja Fishera

$$\frac{\partial \ln f(x;\lambda)}{\partial \lambda} = \frac{\partial}{\partial \lambda}\left(-\ln(x!) + x\ln(\lambda) - \lambda\right) = \frac{x}{\lambda} - 1 = \frac{x-\lambda}{\lambda}$$

$$E\left[\left(\frac{\partial \ln f(x;\lambda)}{\partial \lambda}\right)^2\right] = \frac{E(x-\lambda)^2}{\lambda^2} = \frac{\sigma^2}{\lambda^2} = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

Przedział:

$$\left(nI(\hat{\lambda})\right)^{-\frac{1}{2}} = \sqrt{\frac{\bar{x}}{n}}$$

$$\left(\bar{X} - z_{1-\alpha/2}\sqrt{\frac{\bar{X}}{n}}\ ,\ \bar{X} + z_{1-\alpha/2}\sqrt{\frac{\bar{X}}{n}}\right)$$

Zadanie 10.

Rzucamy niezależnie 2 razy symetryczną n-ścienną kostką do gry, $n \geq 6$ (kostki przyjmują z jednakowym prawdopodobieństwem wartości ze zbioru $\{1,2,\ldots,n\}$). Oznaczmy wyniki przez X_1, X_2 oraz zdefiniujmy $Y = \max(X_1, X_2)$.
Ile wynosi $\mathbb{E}Y$?

(A) $\frac{2(n+1)}{3}$

(B) $\frac{(n+1)(4n-1)}{6n}$

(C) $\frac{4n-1}{6}$

(D) $\frac{(n+1)(3n-1)}{6n}$

(E) Żadne z powyższych

Dla $n = 3$ zm. Y przedstawiona jest w tabeli:

$X_1 \backslash X_2$	1	2	3
1	1	2	3
2	2	2	3
3	3	3	3

$$EY = \frac{1}{9}(1+2+2+2+3+3+3+3+3) = \frac{1}{9}\cdot 22 = \frac{22}{9}$$

Wstawiając $n=3$ kolejno do odpowiedzi otrzymuję odp. B.

DRUGA METODA

Dla ułatwienia w pisaniu $Z = \max(X, Y)$

$$\max(X,Y) = \begin{cases} X, & X \geqslant Y \\ Y, & X < Y \end{cases}$$

$$EZ = E[E[Z|X=x]]$$

$$E[Z|X=x] = \sum_{k=1}^{x} x \cdot \frac{1}{n} + \sum_{k=x+1}^{n} k \cdot \frac{1}{n} = \frac{x^2}{n} + \frac{1}{n}\left[\frac{n(n+1)}{2} - \frac{x(x+1)}{2}\right] =$$

$$= \frac{x^2}{n} + \frac{1}{2n}\left(n^2 + n - x^2 - x\right) = \frac{1}{2n}\left(2x^2 + n^2 + n - x^2 - x\right) = \frac{1}{2n}\left(x^2 - x + n^2 + n\right)$$

$$E\left[\frac{1}{n}\left(X^2 - X + n^2 + n\right)\right] = \frac{1}{2n}\left(EX^2 - EX + n^2 + n\right)$$

$$EX^2 = \sum_{k=1}^{n} k^2 \cdot \frac{1}{n} = \frac{1}{\cancel{n}} \frac{\cancel{n}(n+1)(2n+1)}{6} = \frac{2n^2 + 3n + 1}{6}$$

$$EX = \sum_{k=1}^{n} k \cdot \frac{1}{n} = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}$$

$$EZ = \frac{1}{2n}\left(\frac{2n^2 + 3n + 1}{6} - \frac{3n+3}{6} + \frac{6n^2}{6} + \frac{6n}{6}\right) =$$

$$= \frac{1}{12n}\left(2n^2 + \cancel{3n} + 1 - \cancel{3n} - 3 + 6n^2 + 6n\right) =$$

$$= \frac{1}{12n}\left(8n^2 + 6n - 2\right) = \frac{2}{12n}\left(4n^2 + 3n - 1\right) =$$

$$= \frac{1}{6n}(n+1)(4n-1)$$