

Zadanie 1.

Niech $X_1, \dots, X_n, n \geq 4$ będą niezależnymi zmiennymi losowymi o rozkładzie $N(\theta, 1)$ (normalnym o średniej θ i wariancji 1). Dla każdego $i \in \{1, \dots, n\}$ zaobserwowano jedynie czy obserwacja była mniejsza od 0 czy nie, a dokładniej:

$$Y_i = \begin{cases} 1 & \text{jeśli } X_i < 0, \\ 0 & \text{jeśli } X_i \geq 0. \end{cases}$$

Oznaczmy $T = \sum_{i=1}^n Y_i$, przez Φ oznaczamy z kolei dystrybucję rozkładu standardowego normalnego $N(0, 1)$. Na podstawie tak otrzymanych danych estymator parametru θ otrzymany metodą największej wiarygodności wyraża się wzorem:

$$P(Y_i = 0) = P(X \geq 0) = 1 - P(X < 0) = 1 - P(Z < -\theta) = 1 - \Phi(-\theta)$$

$$P(Y_i = 1) = P(X < 0) = P(X < -\theta) = \Phi(-\theta)$$

Funkcja wiarygodności

$$L(\theta) = \prod_{i=1}^n P(Y_i = y_i)$$

$$L(\theta) = \ln\left(\prod_{i=1}^n P(Y_i = y_i)\right) = \sum_{i=1}^n \ln(P(Y_i = y_i))$$

$$\frac{\partial \ln(\theta)}{\partial \theta} = \sum_{i=1}^n \frac{P'(Y_i = y_i)}{P(Y_i = y_i)} = 0$$

$$\frac{\partial \ln(\theta)}{\partial \theta} = \sum_{i=1}^n \left[\frac{f_X(-\theta)}{P(Y_i = 0)} + \frac{-f_X(-\theta)}{P(Y_i = 1)} \right] = 0 \quad / : f_X(-\theta)$$

$$\sum_{i=1}^n \left[\frac{1}{P(Y_i = 0)} + \frac{-1}{P(Y_i = 1)} \right] = 0$$

$$\frac{\left[\sum_i \text{takie, że liczba } Y_i = 0 \right]}{1 - \Phi(-\theta)} + \frac{\left[-\sum_i \text{takie, że liczba } Y_i = 1 \right]}{\Phi(-\theta)} = 0$$

$$\frac{n - T}{1 - \Phi(-\theta)} + \frac{-T}{\Phi(-\theta)} = 0$$

$$\frac{n - T}{1 - \Phi(-\theta)} = \frac{T}{\Phi(-\theta)}$$

$$n \Phi(-\theta) - T \Phi(-Q) = T - T \Phi(-Q)$$

$$\Phi(-\theta) = \frac{T}{n}$$

$$-\theta = \Phi^{-1}\left(\frac{T}{n}\right)$$

$$\theta = -\Phi^{-1}\left(\frac{T}{n}\right)$$

Zadanie 2.

Wektor losowy (X, Y) ma rozkład ciągły o gęstości

$$f(x, y) = \begin{cases} 2 & \text{dla } \{(x, y) : (x, y) \in (0, 1)^2, y > x\}, \\ 0 & \text{w przeciwnym przypadku.} \end{cases}$$

Jakie jest prawdopodobieństwo, że $\left\lfloor \frac{Y}{X} \right\rfloor$ jest liczbą nieparzystą? Przez $[z]$ oznaczamy liczbę całkowitą najbliższą liczbie z (przyjmujemy $[n + 1/2] = n$ dla liczby całkowitej n).

Warunki brzegowe w tym zadaniu to :

$$n - \frac{1}{2} < \frac{y}{x} < n + \frac{1}{2}, \text{ gdzie } n \text{ to liczba nieparzysta}$$

Wypiszmy kilka przykładów

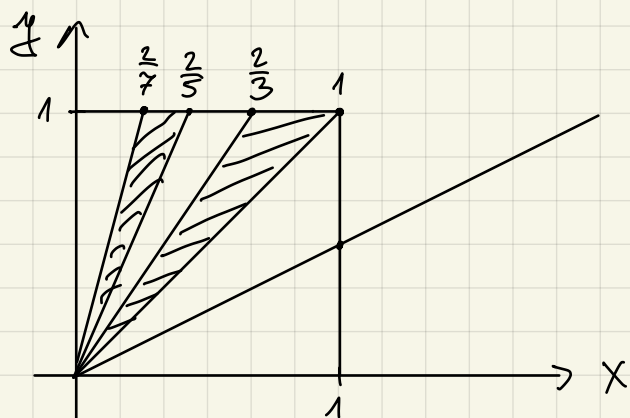
$$n=1: \frac{1}{2}x < y < \frac{3}{2}x$$

$$n=3: \frac{5}{2}x < y < \frac{7}{2}x$$

$$n=5: \frac{9}{2}x < y < \frac{11}{2}x$$

$$n=7: \frac{13}{2}x < y < \frac{15}{2}x$$

Pozostałe policzyć pole dla poszczególnych n i pomnożyć przez gęstość rozkładu, kilka pól jest zamalowane na poniższym rysunku:



$$p_1 = 2 \cdot \frac{1}{2} \cdot 1 \cdot \left(1 - \frac{2}{3}\right) = \frac{1}{3}$$

$$p_3 = \frac{2}{5} - \frac{2}{7} = \frac{4}{35}$$

$$p_5 = \frac{2}{9} - \frac{2}{11} = \frac{4}{99}$$

$$p_7 = \frac{2}{13} - \frac{2}{15} = \frac{4}{195}$$

Cyfli :

$$p = \frac{1}{3} + \frac{4}{35} + \frac{4}{99} + \frac{4}{195} + \dots$$

Trzeba skorzystać z wziora na ciag potegowy :

$$\arctg(x) = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7 + \dots$$

$$\frac{\pi}{4} = \arctg(1) = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots = \frac{2}{3} + \frac{2}{35} + \frac{2}{99} + \frac{2}{195} + \dots$$

$$\frac{\pi}{2} = \frac{4}{3} + \frac{4}{35} + \frac{2}{99} + \frac{2}{195} + \dots$$

Wstawiaj do naszego wziora :

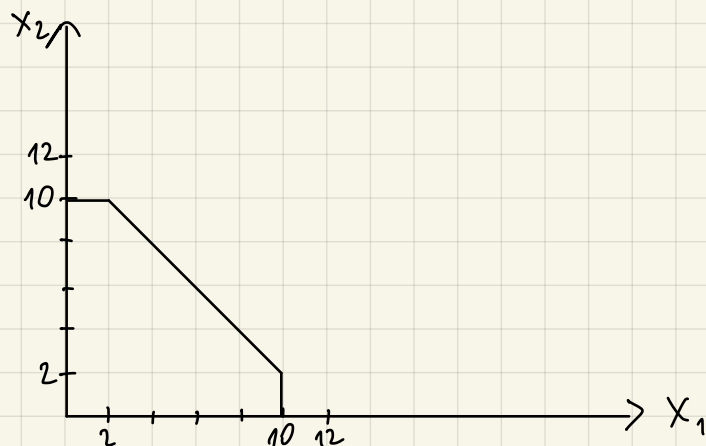
$$p = \frac{1}{3} - \frac{4}{3} + \frac{4}{3} + \frac{4}{35} + \frac{4}{99} + \frac{4}{195} + \dots = \frac{1}{3} - \frac{4}{3} + \frac{\pi}{2} = \frac{\pi}{2} - 1$$

Zadanie 3.

Niech X_1 oznacza czas, który upływa od wypadku samochodowego do zgłoszenia szkody, natomiast X_2 oznacza czas, który upływa od zgłoszenia szkody do wypłaty odszkodowania. Oznaczmy $T = X_1 + X_2$ (czas od wypadku do wypłaty odszkodowania). Wektor losowy (X_1, X_2) ma rozkład łączny:

$$f(x_1, x_2) = \begin{cases} \frac{1}{c} & \text{dla } 0 < x_1 < 10, 0 < x_2 < 10, x_1 + x_2 \leq 12, \\ 0 & \text{w przeciwnym przypadku,} \end{cases}$$

gdzie $c > 0$ jest stałą. Ile wynosi ET ?



Do obliczenia stałej c wystarczy obliczyć pole figury wyznaczonej przez granice x_1 i x_2 :

$$P_F = 10 \cdot 2 + 2 \cdot 2 + 0.5 \cdot 2 \cdot 2 = 62$$

Stąd:

$$c = 62$$

$$f(x_1, x_2) = \begin{cases} \frac{1}{62} & \text{dla } 0 < x_1 < 10, 0 < x_2 < 10, x_1 + x_2 \leq 12 \\ 0 & \text{p.p.} \end{cases}$$

Rozkład margowy X_1 :

$$f_{X_1}(x_1) = \frac{1}{62} \int_0^{10} dx_2 = \frac{10}{62} \quad \text{dla } 0 < x_1 < 2$$

$$f_{X_1}(x_1) = \frac{1}{62} \int_0^{12-x_1} dx_2 = \frac{12-x_1}{62} \quad \text{dla } 2 < x_1 < 10$$

Wartość oczekiwana X_1 :

$$EX_1 = \frac{10}{62} \int_0^2 x_1 dx_1 + \frac{1}{62} \int_2^{10} (12 - x_1) x_1 dx_1 = \frac{5}{17} + \frac{124}{51} = \frac{199}{51}$$

$$EX_2 = EX_1 = \frac{199}{51}$$

$$ET = EX_1 + EX_2 = \frac{398}{51}$$

Zadanie 4.

Niech E będzie dyskretną przestrzenią stanów. Dla dwóch rozkładów na tej przestrzeni: $p \equiv \{p(e)\}, q \equiv \{q(e)\}, e \in E$ definiujemy normę całkowitego wahania:

$$d_{tv}(p, q) = \frac{1}{2} \sum_{e \in E} |p(e) - q(e)|.$$

Ustalmy $n \geq 3$. Rozważamy zbiór permutacji S_n zbioru n -elementowego. Niech π oznacza rozkład jednostajny na tych permutacjach, tzn. $\forall \sigma \in S_n$ mamy $\pi(\sigma) = \frac{1}{n!}$.

Pomyślmy o tych permutacjach jako o układach n kart. Rozważmy następującą procedurę tasowania kart, jej jeden krok to:

- Bierzemy kartę, która jest obecnie na samej górze i wkładamy ją losowo (tzn. z prawdopodobieństwem $1/n$) na każdą z możliwych pozycji.

(Dla przykładu, gdy $n = 4$ i startujemy z permutacji $(1, 2, 3, 4)$, to możemy w 1 kroku dojść z prawdopodobieństwem $1/4$ do każdej z permutacji $(1, 2, 3, 4), (2, 1, 3, 4), (2, 3, 1, 4), (2, 3, 4, 1)$).

Założmy, że startujemy z permutacji identycznościowej $(1, 2, \dots, n)$.

Niech $\rho_2(\cdot)$ oznacza rozkład na permutacjach po wykonaniu dwóch kroków powyższego tasowania kart. Norma całkowitego wahania $d_{tv}(\rho_2, \pi)$ wynosi:

Obserwacja: wartość ρ_2 dla danej permutacji e to $\rho_2(e) = 0$ albo $\rho_2(e) > \frac{1}{n!}$.

$$d(\rho_2, \text{jednostajny}) = \frac{1}{2} \sum_{e - \text{wszystkie permutacje}} \left| \rho_2(e) - \frac{1}{n!} \right| =$$

niech A będzie zbiorem permutacji n -elementowych, dla których

$$\rho_2(*) > 0$$

$$= \frac{1}{2} \sum_{e \in A} (\delta_2(e) - \frac{1}{n!}) + \frac{1}{2} \sum_{e \notin A} \frac{1}{n!} = \frac{1}{2} \sum_{e \in A} (\delta_2(e) - \frac{1}{n!}) + \frac{1}{2} - \frac{1}{2} \sum_{e \in A} \frac{1}{n!} =$$

$$= \frac{1}{2} \sum_{e \in A} \delta_2(e) - \frac{1}{2} \sum_{e \in A} \frac{1}{n!} + \frac{1}{2} - \frac{1}{2} \sum_{e \in A} \frac{1}{n!} = \frac{1}{2} - \sum_{e \in A} \frac{1}{n!} + \frac{1}{2} = 1 - \frac{|A|}{n!}$$

Ile elementów ma zbiór A ? Rozważmy permutacje tego zbioru. Elementy $\{1, 2\}$ mogą stać na dowolnym miejscu, natomiast $\{3, 4, \dots, n\}$ muszą być w dokładnie tej kolejności.

Element 2 możemy wstawić do $(n-1)$ możliwych „luk”, a element 1 do n możliwych „luk”:

$$(n-1)n = |A|$$

$$\text{Wynik} = 1 - \frac{1}{(n-2)!}$$

Zadanie 5.

Niech X i Y będą dwoma niezależnymi zmiennymi losowymi o standardowym rozkładzie normalnym $N(0, 1)$. Zdefiniujmy $M = X + Y, N = X - Y$.

Ile wynosi $E(M^{17}N^{18})$?

Jeśli dane są niezależne zm. los. X, Y o rozkładzie normalnym to dla $\forall \alpha, \beta$ zm. $\alpha X + \beta Y$ ma rozkład normalny. Ponadto wektor złożony z takich kombinacji liniowych ma rozkład 2-czny normalny.

Jeśli wektor (M, N) ma rozkład 2-czny normalny oraz $\text{cov}(M, N) = 0$, to zm. M, N są niezależne:

$$\text{cov}(M, N) = \text{cov}(X+Y, X-Y) = \text{cov}(X, X) - \text{cov}(X, Y) + \text{cov}(Y, X) - \text{cov}(Y, Y) =$$

$$= \text{Var}(X) - \text{Var}(Y) = 1 - 1 = 0$$

Funkcje od zm. niezależnych są zm. niezależnymi. W szczególności z niezależności M i N wynika niezależność M^{17} i N^{18} . Mamy więc:

$$E(M^{17} N^{18}) = E(M^{17}) E(N^{18})$$

$$E(M) = E(X+Y) = 0$$

Ponieważ N ma rozkład normalny o średniej 0, to z symetrii tego rozkładu względem zera wynika, że wszystkie nieparzyste momenty są równe. Ostatecznie:

$$E(M^{17} N^{18}) = 0 \cdot E(N^{18}) = 0.$$

Zadanie 6.

Pewne urządzenie przestaje działać tylko wtedy, gdy przestają działać jego dwa komponenty. Czasy działania tych komponentów T_1 oraz T_2 są niezależnymi zmiennymi losowymi o rozkładzie wykładniczym z parametrem 1. Koszt działania urządzenia wyliczony został jako $X = 2T_1 + T_2$. Ile wynosi $\text{Var}(X)$?

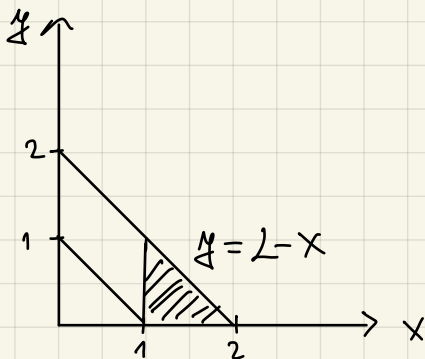
$$\text{Var}(X) = \text{Var}(2T_1 + T_2) = 4\text{Var}(T_1) + \text{Var}(T_2) = 4 + 1 = 5$$

Zadanie 7.

Wektor losowy (X, Y) ma rozkład łączny o gęstości

$$f(x, y) = \begin{cases} c \cdot x & \text{dla } 0 \leq x \leq 2, \max(0, 1-x) \leq y \leq 2-x, \\ 0 & \text{w przeciwnym przypadku,} \end{cases}$$

gdzie $c > 0$ jest stałą. Wariancja warunkowej zmiennej losowej Y pod warunkiem $X = x$, tj. $\text{Var}(Y|X = x)$ dla $x \in (1, 2)$ wynosi:



$$f_{Y|X}(y|x) = \frac{f_{XY}(x, y)}{f_X(x)}$$

$$f_X(x) = \int_0^{2-x} c x dy = c x y \Big|_0^{2-x} = c x (2-x), \quad 1 < x < 2$$

$$f_{Y|X}(y|x) = \frac{c \cdot x}{c \cdot x(2-x)} = \frac{1}{2-x} \quad 0 < y < 2-x$$

$$E[Y|X=x] = \int_0^{2-x} \frac{y}{2-x} dy = \frac{2-x}{2}$$

$$E[Y^2|X=x] = \int_0^{2-x} \frac{y^2}{2-x} dy = \frac{(2-x)^2}{3}$$

$$\text{Var}(Y|X=x) = E[Y^2|X] - (E[Y|X])^2 = \frac{(2-x)^2}{12}$$

Można też zauważyć, że $f_{Y|X}(y|x)$ to gęstość rozkładu jednostajnego, wtedy:

$$\text{Var}(Y|X) = \frac{(b-a)^2}{12} = \frac{(2-x)^2}{12}$$

Zadanie 8.

Dyskretny wektor losowy (X, Y) ma rozkład łączny $p(x, y)$ dla $x = 0, 1$ oraz $y = 0, 1, 2$. Wiadomo, iż $p(1, 2) = 2p(1, 1)$ i wartość $p(1, 1)$ jest taka, że wariancja zmiennej losowej $Z = XY$ jest największa (zmaksymalizowana).

Ile wynosi prawdopodobieństwo tego, że X lub Y przyjmuje wartość 0?

$X \backslash Y$	0	1	2
0			
1		$p(1,1)$	$2p(1,1)$

$$R_Z = \{0, 0, 0, 0, 1, 2\} = \{0, 1, 2\}$$

$$P(Z=1) = p(1,1)$$

$$P(Z=2) = 2p(1,1)$$

$$EZ = 1 \cdot p(1,1) + 4 \cdot p(1,1) = 5p(1,1)$$

$$EZ^2 = 1p(1,1) + 4p(1,1) = 5p(1,1)$$

$$p = p(1,1)$$

$$\text{Var}(Z) = 5p - 25p^2$$

$$\text{Var}'(Z) = 5 - 50p = 0$$

$$p = 0.1$$

$$P(X=0 \vee Y=0) = 1 - 3p = 0.46$$

Zadanie 9.

Niech X i Y będą niezależnymi zmiennymi losowymi o rozkładach wykładniczych takich, że $EX = EY = \frac{1}{\lambda} > 0$. Wiadomo również, że $\lambda \neq 1$.

Rozważmy zmienną losową $Z = \frac{X}{X+Y}$. Ile wynosi $Var(Z)$?

$$X \sim \text{Exp}(\lambda) \Rightarrow X \sim P(1, \lambda)$$

$$Y \sim \text{Exp}(\lambda) \Rightarrow Y \sim P(1, \lambda)$$

$$X \sim P(1, \lambda), Y \sim P(1, \lambda) \Rightarrow \frac{X}{X+Y} \sim \text{Beta}(1, 1) \sim U(0, 1)$$

Czyli:

$$Z \sim U(0, 1)$$

$$Var(Z) = \frac{1}{12}$$

Zadanie 10.

Zmienne losowe X_1, X_2 są niezależnymi zmiennymi losowymi o rozkładzie $U(\theta, \theta+1)$, tj. o rozkładzie jednostajnym na przedziale $(\theta, \theta+1)$, gdzie $\theta > 0$.

Interesuje nas testowanie hipotezy $H_0: \theta = 0$ przy alternatywie $H_1: \theta > 0$. Rozważmy 2 testy:

$\phi_1(X_1)$: Odrzucamy H_0 jeśli $X_1 > 0.95$,

$\phi_2(X_1, X_2)$: Odrzucamy H_0 jeśli $X_1 + X_2 > c$.

Ile wynosi stała c jeśli wiadomo, że test ϕ_2 ma taki sam rozmiar jak test ϕ_1 ?

$$P(X_1 > 0.95) = P(X_1 + X_2 > c)$$

$$\frac{1}{20} = P(X_1 + X_2 > c)$$

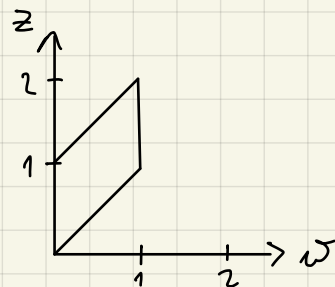
$$Z = X_1 + X_2$$

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X_1}(w) f_{X_2}(z-w) dw$$

$$\begin{aligned} 0 &< w < 1 \\ 0 &< z-w < 1 \\ w &< z < 1+w \end{aligned}$$

Dla $0 < z < 1$:

$$f_Z(z) = \int_0^z dw = z$$



Dla $1 < z < 2$:

$$f_z(z) = \int_{z-1}^1 d\omega = 2 - z$$

$$\frac{1}{20} = P(Z > c) = \int_c^2 2 - z \, dz = 2 - 2c + \frac{c^2}{2}$$

$$c^2 - 4c + \frac{39}{10} = 0$$

$$\Delta = \frac{2}{5}$$

$$c = 2 - \frac{\sqrt{10}}{10}$$

$$Z = X + Y$$

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P(X + Y \leq z) = \int_{-\infty}^{\infty} P(X + Y \leq z \mid Y = \omega) f_Y(\omega) \, d\omega = \\ &= \int_{-\infty}^{\infty} P(X + \omega \leq z \mid Y = \omega) f_Y(\omega) \, d\omega = \int_{-\infty}^{\infty} P(X \leq z - \omega) f_Y(\omega) \, d\omega = \\ &= \int_{-\infty}^{\infty} F_X(z - \omega) f_Y(\omega) \, d\omega \end{aligned}$$

DRUGA METODA

$$0 < X_1 + X_2 < 2$$

$$P(X_1 + X_2 > c) = \int_{-\infty}^{\infty} P(X_1 + X_2 > c \mid X_2 = x_2) f_{X_2}(x_2) \, dx_2 =$$

$$= \int_{-\infty}^{\infty} 1 - P(X_1 \leq c - x_2) \, dx_2 = \left| \begin{array}{l} 0 < c - x_2 < 1 \quad | \cdot (-1) \\ 0 > x_2 - c > -1 \quad | + c \\ c > x_2 > c - 1 \\ 1 > x_2 > c - 1 \end{array} \right. \begin{array}{l} \text{podnosz na rozwiązanie} \\ P(X_1 + X_2 > c) = \frac{1}{10} \text{ stąd } c > 1 \\ \text{bo musi być blisko górnej} \\ \text{granicy} \end{array}$$

$$= \int_{c-1}^1 1 - c + x_2 \, dx_2 = x_2 - cx_2 + \frac{x_2^2}{2} \Big|_{c-1}^1 =$$

$$= 1 - c + \frac{1}{2} - (c-1) + c(c-1) - \frac{c^2 - 2c + 1}{2} =$$

$$= 1 - c + \cancel{\frac{1}{2}} - c + 1 + c^2 - \cancel{c} - \frac{c^2}{2} + \cancel{c} - \cancel{\frac{1}{2}} =$$

$$= 2 - 2c + \frac{c^2}{2} \quad := \frac{1}{20}$$

$$c^2 - 4c + \frac{39}{10} = 0$$

$$\Delta = \frac{2}{5}$$

$$c = 2 - \frac{\sqrt{10}}{10}$$