Zadanie 1.

Wektor (X,Y) ma rozkład ciągły o gęstości

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

$\mathrm{Cov}(X, \min(X,Y))$ wynosi:

$$\mathrm{Cov}(X, \min(X,Y)) = E[X \min(X,Y)] - EX\, E[\min(X,Y)]$$

$$f_X(x) = \int_0^1 dy = 1 \quad , \; x \in (0,1)$$

$$EX = \int_0^1 x\, dx = \frac{1}{2}$$

$$E[\min(X,Y)] = \int_0^1 \int_0^1 \min(x,y)\, dx\, dy = \int_0^1 \left(\int_0^y x\, dx + \int_y^1 y\, dx \right) dy = \frac{1}{3}$$

$$E[X \min(X,Y)] = \int_0^1 \int_0^1 x \min(x,y)\, dx\, dy = \int_0^1 \left(\int_0^y x^2\, dx + \int_y^1 xy\, dx \right) dy = \frac{5}{24}$$

$$\mathrm{Cov}(X, \min(X,Y)) = \frac{5}{24} - \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{24}$$

Zadanie 2.

Ubezpieczyciel zawarł kontrakt z $n \geq 1$ klientami na stały czas $T > 0$. Odpowiedzialność ubezpieczyciela kończy się z momentem zajścia szkody (po zajściu szkody wypłacane jest należne odszkodowanie). Załóżmy, że czasy do powstania szkody są niezależne dla wszystkich klientów i mają rozkład wykładniczy o średniej $\frac{1}{\theta} > 0$. Okazało się, że w trakcie czasu T wypłacono ubezpieczenie $0 < k \leq n$ klientom w chwilach $x_1, x_2, \ldots, x_k$ od zawarcia umowy (tzn. $n-k$ klientów nie spowodowało szkody w ciągu czasu T od zawarcia umowy).

Estymator największej wiarogodności dla parametru θ wynosi:

Mamy dane zagregowane: dla obserwacji X_i mamy jej dokładną wartość tylko, gdy $X_i < T$, czyli możliwe do zaobserwowania są:

$\{X_i$ dla X_i z $[0,T]\}$ - rozkład ciągły

lub „$X_i > T$" - rozkład dyskretny

Dla x_i z $[0,T]$ mamy $L(x;\theta) = \theta \exp(-\theta x)$ pierwsza część

$L(„X > T";\theta) = P(\mathcal{E}x(\theta) > T) = \exp(-\theta T)$ druga część

k klientów spowodowało szkody (w momentach $x_1, \ldots, x_k$)

Stąd pierwsza część funkcji log-wiarygodności to

$$k \log(\theta) - \theta \sum_{i=1}^{k} x_i$$

Druga część funkcji wiarygodności:

$$(n-k)(-\theta \cdot T)$$

Czyli cała f. wiarygodności to:

$$L(\theta) = (n-k)(-\theta T) + k \log(\theta) - \theta \sum_{i=1}^{k} x_i$$

$$L(\theta) = k \log(\theta) - \theta\left((n-k)T + \sum_{i=1}^{k} x_i\right)$$

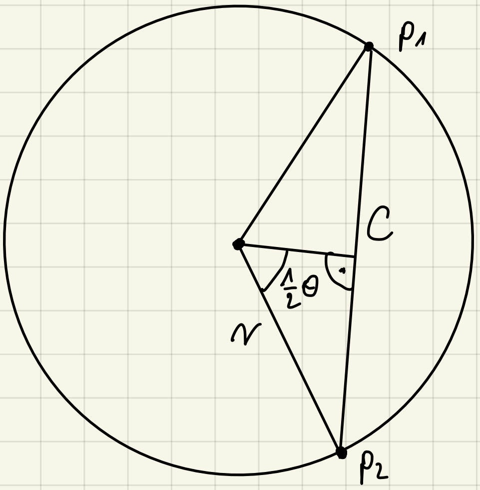
$$L'(\theta) = \frac{k}{\theta} - \left((n-k)T + \sum_{i=1}^{k} x_i\right) := 0$$

$$\theta = \frac{k}{(n-k)T + \sum_{i=1}^{k} x_i}$$

Zadanie 3.

Wybierzmy losowo dwa punkty p_1, p_2 na obwodzie okręgu o promieniu 2 (wybierając jednostajnie, wzajemnie niezależnie, dwa kąty $\theta_1, \theta_2 \in [0, 2\pi)$). Niech C będzie długością odcinka łączącego p_1 i p_2.
$Pr(C > 2\sqrt{3})$ wynosi:

Obliczę jaki minimalny kąt spełnia założenie $C > 2\sqrt{3}$. Maksymalny to 180°.



$$\sin\left(\tfrac{1}{2}\theta\right) = \frac{\frac{1}{2}C}{r} = \frac{C}{4}$$

$$\sin\left(\tfrac{1}{2}\theta\right) = \frac{2\sqrt{3}}{4} = \frac{\sqrt{3}}{2}$$

$\theta = 120^\circ$

Czyli możliwy zakres kątów to $180^\circ - 120^\circ = 60^\circ$

$$Pr(C > 2\sqrt{3}) = \frac{60^\circ}{180^\circ} = \frac{1}{3}$$

Zadanie 4.

Niech $X_1, \ldots, X_n$ jest będzie próbką z rozkładu wykładniczego o średniej $1/\lambda > 0$. Obserwujemy próbkę $Y_1, \ldots, Y_n$, gdzie

$$Y_i = \left\lfloor \frac{X_i}{a} \right\rfloor,$$

$\lfloor x \rfloor$ jest częścią całkowitą z x (tzn. $\lfloor x \rfloor = \max\{k \in \mathbb{Z} : k \leq x\}$) oraz $a > 0$ jest znanym parametrem. Niech $\hat{S} = \frac{1}{n}\sum_{i=1}^{n} Y_i$. Przez $\lceil x \rceil$ oznaczamy tzw. *sufit* z liczby x tzn. $\lceil x \rceil = \min\{k \in \mathbb{Z} : k \geq x\}$.

Estymator *największej wiarogodności* $\hat{\lambda}$ nieznanego parametru λ wyraża się wzorem:

$$Y=0: \; P\left(0 < \frac{X}{a} < 1\right) = P(0 < X < a)$$

$$Y=1: \; P\left(1 < \frac{X}{a} < 2\right) = P(a < X < 2a)$$

$$Y=k: \; P\left(k < \frac{X}{a} < k+1\right) = P(ka < X < a(k+1)) =$$

$$= \int_{ka}^{a(k+1)} \lambda e^{-\lambda x}\,dx = 1 - e^{-\lambda a(k+1)} - 1 + e^{-\lambda a k} =$$

$$= e^{-\lambda a k} - e^{-\lambda a(k+1)} = \underbrace{\left(e^{-\lambda a}\right)}_{=q}{}^{k} \underbrace{\left(1 - e^{-\lambda a}\right)}_{=p}$$

$$L(y_i; \lambda) = \prod_{i=1}^{n} q^{y_i} p = q^{\sum_{i=1}^{n} y_i} p^n$$

$$l(y_i; \lambda) = \sum_{i=1}^{n} y_i \cdot \ln(q) + n \ln(p) = \sum_{i=1}^{n} y_i \cdot (-\lambda a) + n \ln(1 - e^{-\lambda a})$$

$$l' = -a \sum_{i=1}^{n} y_i + \frac{n a e^{-\lambda a}}{1 - e^{-\lambda a}} \; := 0$$

$$\frac{n \cancel{a} q}{p} = \cancel{a} \sum_{i=1}^{n} y_i$$

$$\frac{q}{p} = \frac{1}{n} \sum_{i=1}^{n} y_i = \hat{S}$$

$$q = p \hat{S}$$

$$e^{-\lambda a} = (1 - e^{-\lambda a})\hat{s}$$

$$e^{-\lambda a} = \hat{s} - \hat{s} e^{-\lambda a}$$

$$e^{-\lambda a} + \hat{s} e^{-\lambda a} = \hat{s}$$

$$e^{-\lambda a}(1 + \hat{s}) = \hat{s}$$

$$e^{-\lambda a} = \frac{\hat{s}}{1 + \hat{s}}$$

$$-\lambda a = \ln\left(\frac{\hat{s}}{1+\hat{s}}\right)$$

$$\lambda = -\frac{1}{a} \ln\left(\frac{\hat{s}}{1+\hat{s}}\right)$$

$$\hat{\lambda} = \ln \sqrt[a]{1 + \frac{1}{\hat{s}}}$$

Zadanie 5.

Niech $X_1, \ldots, X_n, n \geq 2$ będą niezależnymi zmiennymi losowymi, każda o rozkładzie Poissona z parametrem $\lambda > 0$. Niech $k \geq 2$, rozpatrzmy wektor losowy $(Z_1, \ldots, Z_n)$ o rozkładzie

$$Pr(Z_1 = x_1, \ldots, Z_n = x_n) = Pr(X_1 = x_1, \ldots, X_n = x_n | X_1 + \ldots + X_n = k)$$

Korelacja $\mathrm{Corr}(Z_1, Z_2)$ wynosi:

(A) $-\frac{1}{n-2}$

(B) $\frac{2}{n}$

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

(D) $-\frac{1}{2n-1}$

(E) $\frac{1}{n}$

Jeśli w odpowiedziach pojawiają się zmienne, to aby uniknąć rachunków warto wziąć ustaloną wartość zmiennej, dla której łatwo się liczy i sprawdzić, która odpowiedź pasuje. Jeśli tylko jedna to jest to wynik zadania.

Niech $n = 2$, wtedy $Z_2 = k - Z_1$ oraz:

$$\mathrm{Corr}(Z_1, Z_2) = \mathrm{Corr}(Z_1, k - Z_1) = \mathrm{Corr}(Z_1, -Z_1) = -1$$

Podstawiając $n = 2$ do odpowiedzi otrzymujemy wynik -1 tylko dla odpowiedzi C.

Zadanie 6.

Zmienna losowa (X,Y) ma rozkład prawdopodobieństwa o gęstości

$$f(x,y) = \begin{cases} \frac{3}{4}x^2y & \text{dla } (x,y) \in (0,2) \times (0,1) \\ 0 & \text{w przeciwnym przypadku} \end{cases}$$

Wówczas $E(2X - Y | X - Y = 1)$ wynosi

Nie pozostaje nic innego jak obliczenie gęstości rozkładu łącznego dla następujących zm.:

$$\begin{cases} M = 2X - Y \\ V = X - Y \end{cases}$$

$$\begin{cases} X = M - V \\ Y = M - 2V \end{cases}$$

Jakobian (trik: jeżeli w macierzy są same liczby to nie trzeba dalej liczyć bo w ostatecznym równaniu i tak się skróci):

$$J = \begin{vmatrix} 1 & -1 \\ 1 & -2 \end{vmatrix} = -2 + 1 = 1$$

$$f(m, v) = \frac{3}{4}(m - v)^2(m - 2v)$$

Warunki brzegowe wystarczy obliczyć dla m przy ustalonym $v = 1$:

$$\begin{cases} 0 < x < 2 \\ 0 < y < 1 \end{cases}$$

$$\begin{cases} 0 < m - 1 < 2 \\ 0 < m - 2 < 1 \end{cases}$$

$$\begin{cases} 1 < m < 3 \\ 2 < m < 3 \end{cases}$$

$$2 < m < 3$$

Ostatecznie :

$$E[2X - Y \mid X - Y = 1] = \frac{\int_2^3 m \frac{3}{4}(m-v)^2 (m-2v)\, dm}{\int_2^3 \frac{3}{4}(m-v)^2 (m-2v)\, dm} =$$

$$= \frac{\int_2^3 m(m-1)^2 (m-2)\, dm}{\int_2^3 (m-1)^2 (m-2)\, dm} =$$

$$= \frac{58}{15} \cdot \frac{12}{17} = \frac{232}{85}$$

Zadanie 7.

Niech $X_1, X_2, \ldots, X_{10}$ będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie o ciągłej, ściśle rosnącej dystrybuancie F. Niech $X_{1:10}, \ldots, X_{10:10}$ oznaczają odpowiednie statystyki pozycyjne.
Niech m_X oznacza medianę zmiennej losowej o rozkładzie F_X.
$Pr(X_{3:10} \leq m_X \leq X_{8:10})$ wynosi

Elementarna obserwacja: jeśli przyłożymy do obu stron nierówności $X \leq Y$ funkcję rosnącą F, to $X \leq Y$ wtedy i tylko wtedy gdy $F(X) \leq F(Y)$

Pytają o $P(X_{3:10} \leq m_X \leq X_{8:10})$

Zdefiniujmy sobie zm.:

$Y_i = \mathbb{1}(X_i < m_X)$

Wtedy

$P(X_{3:10} < m_X < X_{8:10}) = P(Y_{3:10} = 0, Y_{8:10} = 1)$ (bo jeśli 3 z kolei X był $< m_X$, to trzeci z kolei Y jest równy 0, i analogicznie dla 8:10)

Policzmy $P(Y_{3:10} = 0, Y_{8:10} = 1)$ przez zdarzenie dopełniające:

$P(Y_{3:10} = 0, Y_{8:10} = 1) = 1 - P(Y_{3:10} = 1 \vee Y_{8:10} = 0) =$ | zasada włączeń i wyłączeń, przy czym przecięcie zdarzeń jest puste | $= 1 - P(Y_{3:10} = 1) - P(Y_{8:10} = 0)$

Zauważmy, $P(Y_i = 1) = P(Y_i = 0) = \frac{1}{2}$ (definicja mediany rozkładu m_X)

Więc

$P(Y_{3:10} = 1) = P($w schemacie 10 rzutów symetryczną monetą wyrzucimy 0, 1 lub 2 orły$) = \left(\frac{1}{2}\right)^{10}(1 + 10 + 45) = \frac{56}{1024}$

$P(Y_{8:10} = 0)$ – zupełnie analogicznie $= \frac{56}{1024}$

Czyli

$= 1 - P(Y_{3:10} = 1) - P(Y_{8:10} = 0) = (1024 - 2 \cdot 56)/1024 = \frac{912}{1024}$

Zadanie 8.

Załóżmy, że $U_1, U_2, \dots, U_n, \dots$ są niezależnymi zmiennymi losowymi o takim samym rozkładzie jednostajnym na $(0,1)$. Niech

$$N = \min\left\{n \geq 0 : \sum_{i=1}^{n} U_i > \frac{1}{2}\right\}.$$

Wartość oczekiwana EN wynosi:

Ogólny wzór : dla zm. X o wartościach naturalnych :

$$EX = \sum_{i=1}^{\infty} P(X \geq i)$$

Czyli

$EN = \sum_{n=1}^{\infty} P(N \geq n) = \sum_{n=1}^{\infty} P($minimum z indeksów na których ciąg sum częściowych przekracza $\frac{1}{2} \geq n) = \sum_{n=1}^{\infty} P($ciąg sum częściowych na indeksie $n-1$ jest $\frac{1}{2}) = \sum_{n=1}^{\infty} P(U_1 + U_2 + \dots U_{n-1} < \frac{1}{2}) = |$ można zauważyć, że zbiór $x_1 + x_2 + \dots + x_{n-1} < \frac{1}{2}$ to sympleks $(n-1)$-wymiarowy o krawędzi $\frac{1}{2}$ więc ma objętość $\frac{1}{(n-1)!} \cdot \left(\frac{1}{2}\right)^{n-1}$ $| =$

$$= \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \cdot \left(\frac{1}{2}\right)^{n-1} = \sum_{n=0}^{\infty} \frac{1}{n!} \cdot \left(\frac{1}{2}\right)^{n} = e^{\frac{1}{2}}$$

Zadanie 10.

Niech $X = \{X_n\}_{n\geq 0}$ będzie łańcuchem Markowa na przestrzeni stanów $\mathbb{E} = \{1,2\}$ z następującą macierzą przejść:

$$\mathbf{P} = \begin{pmatrix} \frac{5}{8} & \frac{3}{8} \\ \frac{1}{8} & \frac{7}{8} \end{pmatrix}.$$

Wiadomo, że $Pr(X_{10} = 1) = \dfrac{1025}{4096}$. Ile wynosi $Pr(X_0 = 1)$?

Wzorek: $\pi^{(n)} = \pi^{(0)} P^{n}$

$$\pi^{(10)} = \left(\frac{1025}{4096}, 1 - \frac{1025}{4095}\right) = \left(\frac{1025}{4096}, \frac{3071}{4096}\right)$$

Polecam potęgować macierz na kalkulatorze.

$$\left(\frac{1025}{4096}, \frac{3071}{4096}\right) = (p_1, p_2)\begin{pmatrix} \frac{5}{8} & \frac{3}{8} \\ \frac{1}{8} & \frac{7}{8} \end{pmatrix}^{10} = (p_1, p_2)\begin{pmatrix} 0.2507 & 0.7492 \\ 0.2497 & 0.7502 \end{pmatrix}$$

$$\begin{cases} 0.2507\,p_1 + 0.2497\,p_2 = \frac{1025}{4096} \\ 0.7492\,p_1 + 0.7502\,p_2 = \frac{3071}{4096} \end{cases}$$

$$p_1 = 0.51 \approx 0.5$$