Praca domowa #1 z SAD – przykładowe rozwiązania

Zadanie 1:
Obserwujemy dwie niezależne próby losowe $(X_1, \ldots, X_n)$, $(Y_1, \ldots, Y_m)$. Wiadomo, że $X_i \sim N(2\mu, 1)$ oraz $Y_i \sim N(\mu, 1)$.

- Wyznaczyć Metodą Największej Wiarogodności estymator parametru μ (korzystając z obydwu prób). Czy otrzymany estymator jest nieobciążony?
- Wyznaczyć ryzyko (średni błąd kwadratowy – MSE) uzyskanego estymatora.

Rozwiązanie: Gęstość zmiennej losowej X ma postać $\frac{1}{\sqrt{2\pi}} \cdot \exp(-\frac{1}{2}(x-2\mu)^2)$, a gęstość zmiennej Y: $\frac{1}{\sqrt{2\pi}} \cdot \exp(-\frac{1}{2}(y-\mu)^2)$.

Zapiszmy funkcję wiarogodności:

$$L(x_1, \ldots, x_n, y_1, \ldots, y_m, \mu) = \left(\frac{1}{\sqrt{2\pi}}\right)^{n+m} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(x_i - 2\mu)^2\right) \exp\left(-\frac{1}{2}\sum_{i=1}^{m}(y_i - \mu)^2\right)$$

Znajdźmy μ, dla którego otrzymane wyniki są najbardziej prawdopodobne. Funkcja L przyjmuje maksimum w tym samym punkcie co $\ln(L) =: l$.

$$l(x_1, \ldots, x_n, y_1, \ldots, y_m, \mu) = -\frac{n+m}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(x_i - 2\mu)^2 - \frac{1}{2}\sum_{i=1}^{m}(y_i - \mu)^2$$

Policzmy pochodną l po μ i znajdźmy μ, dla którego się ona zeruje.

$$\begin{aligned}\frac{\partial l}{\partial \mu}(x_1, \ldots, x_n, y_1, \ldots, y_m, \mu) &= -\frac{1}{2}\sum_{i=1}^{n}\big(2(x_i - 2\mu)\cdot(-2)\big) - \frac{1}{2}\sum_{i=1}^{m}\big(2(y_i - \mu)\cdot(-1)\big) = \\ &= \sum_{i=1}^{n} 2\big(x_i - 2\mu\big) + \sum_{i=1}^{m}\big(y_i - \mu\big) = \sum_{i=1}^{n} 2x_i + \sum_{i=1}^{m} y_i - \mu\cdot(4n+m)\end{aligned}$$

Zatem:

$$\frac{\partial l}{\partial \mu}(x_1, \ldots, x_n, y_1, \ldots, y_m, \mu) = 0 \iff \mu = \frac{1}{4n+m}\left(\sum_{i=1}^{n} 2x_i + \sum_{i=1}^{m} y_i\right)$$

Dostajemy więc wzór estymatora parametru μ:

$$\frac{1}{4n+m}\left(\sum_{i=1}^{n} 2x_i + \sum_{i=1}^{m} y_i\right) =: \hat{\mu}.$$

Czy l ma maksimum w $\hat{\mu}$ (czyli czy $\hat{\mu} = MLE(\mu)$)? Tak, bo:

$\frac{\partial^2 l}{\partial^2 \mu}(x_1, \ldots, x_n, y_1, \ldots, y_m, \mu) = -(4n+m) < 0$ dla każdego μ, w tym dla $\hat{\mu}$.

Czy jest on nieobciążony? Tak, bo:

$$E(\hat{\mu}) = \frac{1}{4n+m} E\left(2\sum_{i=1}^{n} X_i + \sum_{i=1}^{m} Y_i\right) = \frac{2n \cdot 2\mu + m \cdot \mu}{4n+m} = \frac{4n+m}{4n+m}\mu = \mu.$$

Wyznaczmy ryzyko estymatora $\hat{\mu}$.

$$\begin{aligned}
MSE(\hat{\mu}) &= E\left(\left(\frac{1}{4n+m}(\sum_{i=1}^{n} 2X_i + \sum_{i=1}^{m} Y_i) - \mu\right)^2\right) = \\
&= E\left(\frac{1}{(4n+m)^2}\left(\sum_{i=1}^{n} 2X_i + \sum_{i=1}^{m} Y_i\right)^2 - \frac{2\mu}{4n+m}\left(\sum_{i=1}^{n} 2X_i + \sum_{i=1}^{m} Y_i\right) + \mu^2\right) = \\
&= \mu^2 - \frac{2\mu}{4n+m} E\left(\sum_{i=1}^{n} 2X_i + \sum_{i=1}^{m} Y_i\right) + \frac{1}{(4n+m)^2} E\left(4\left(\sum_{i=1}^{n} X_i\right)^2 + 4\sum_{i=1}^{n} X_i \sum_{i=1}^{m} Y_i + \left(\sum_{i=1}^{m} Y_i\right)^2\right) = \\
&= \left\{ \begin{array}{l} X, Y \text{ niezależne } \Rightarrow \\ E(X_i Y_j) = E(X_i)E(Y_j)\ \forall_{i,j} \\ \text{podobnie dla } i \neq j \\ E(X_i X_j) = E(X_i)E(X_j) \end{array} \right\} =
\end{aligned}$$

$$= \mu^2 - \frac{2\mu^2(4n+m)}{4n+m} + \frac{4 \cdot 2\mu^2 nm}{(4n+m)^2} + \frac{4 \cdot (2\mu)^2 n(n-1) + \mu^2 m(m-1}{(4n+m)^2} + \frac{1}{(4n+m)^2}\left(4E\left(\sum_{i=1}^{n} X_i^2\right) + E\left(\sum_{i=1}^{m} Y_i^2\right)\right) =$$

$$= \mu^2\left(1 - \frac{8n+2m}{4n+m} + \frac{8nm}{(4n+m)^2} + \frac{16n^2 - 16n + m^2 - m}{(4n+m)^2}\right) + \frac{1}{(4n+m)^2}\left(4n(Var(X) + (2\mu)^2) + m(Var(Y) + \mu^2)\right) =$$

$$= \mu^2\left(1 - \frac{8n+2m}{4n+m} + \frac{16n^2 - 16n + 8nm + m^2 - m + 16n + m}{(4n+m)^2}\right) + \frac{4n+m}{(4n+m)^2} =$$

$$= \mu^2\left(1 - \frac{8n+2m}{4n+m} + \frac{(4n+m)^2}{(4n+m)^2}\right) + \frac{4n+m}{(4n+m)^2} = \frac{1}{4n+m}$$

Zadanie 2:
Niech $(X_1, \ldots, X_n)$ będą niezależnymi zmiennymi losowymi o takim samym rozkładzie o gęstości postaci:

$$f_\lambda(x) = \frac{1}{6\lambda^4} x^3 e^{-\frac{x}{\lambda}}, x > 0, \lambda > 0$$

.

- Wyznacz estymator Metodą Największej Wiarogodności nieznanego parametru λ. Wiedząc, że wartość oczekiwana wynosi $\mathbb{E}X_i = 4\lambda$, sprawdź, czy otrzymany estymator jest estymatorem nieobciążonym.
- Wyznacz ryzyko (średni błąd kwadratowy, MSE) dla otrzymanego estymatora. Czy otrzymany estymator jest zgodny? (*Obserwacja: dla estymatorów zgodnych* $\lim_{n\to\infty} MSE(\theta) \to 0$)

Rozwiązanie:

Wyznaczenie postaci estymatora MNW:

$$L(x_1, ..., x_n, \lambda) = \prod_{i=1}^{n} \frac{1}{6\lambda^4} X_i^3 e^{-\frac{X_i}{\lambda}} = (\frac{1}{6\lambda^4})^n \prod_{i=1}^{n} X_i^3 e^{\sum_{i=1}^{n} \frac{X_i}{\lambda}}$$

$$\ln(L) = -n\ln(6) - 4n\ln(\lambda) + 3\sum_{i=1}^{n} \ln(X_i) + \frac{1}{\lambda}\sum_{i=1}^{n} X_i$$

$$\frac{\partial \ln L}{\partial \lambda} = -\frac{4n}{\lambda} - \frac{1}{\lambda^2}\sum_{i=1}^{n} = 0$$

$$\lambda_{MLE} = -\frac{\sum_{i=1}^{n} X_i}{4n}$$

Formalnie powinniśmy sprawdzić, czy znalezione rozwiązanie to faktycznie maksimum.

Sprawdzamy, czy estymator nieobciążony:

$$\mathbb{E}\,\lambda_{MLE} = \frac{1}{4n}\,\mathbb{E}\sum_{i=1}^{n} X_i = \frac{1}{4n} n 4\lambda = \lambda$$

Wartość oczekiwana estymatora jest równa szacowanemu parametrowi, czyli estymator jest nieobciążony.
Ryzyko $R = MSE(\hat{\lambda}) = b(\hat{\lambda})^2 + \mathrm{Var}(\hat{\lambda}) = 0 + \mathrm{Var}(\hat{\lambda})$

$$\mathrm{Var}(\frac{\sum_{i=0}^{n} X_i}{4n}) = \frac{1}{16n^2}\mathrm{Var}(\sum_{i=0}^{n} X_i) = \frac{1}{16n}\mathrm{Var}(X_i) = \frac{1}{16n}(\mathbb{E}X_i^2 - (\mathbb{E}X_i)^2)$$

Potrzebujemy $\mathbb{E}X_i^2$. Wyznaczamy ze wzoru:

$$\mathbb{E}X_i^2 = \int_{x=0}^{\infty} x^2 \frac{1}{6\lambda^4} x^3 e^{-\frac{x}{\lambda}} dx = \frac{1}{6}\int_{x=0}^{\infty} \frac{1}{\lambda^4} x^5 e^{-\frac{x}{\lambda}} dx = \frac{120\lambda^2}{6} = 20\lambda^2$$

Wiemy, że $(\mathbb{E}X_i)^2 = 4\lambda$. Ostatecznie: $\mathrm{Var}(X_i) = 20\lambda^2 - 16\lambda^2 = 4\lambda^2$
Ryzko wynosi $R = \frac{1}{16n} 4\lambda^2 = \frac{\lambda^2}{4n}$. Estymator jest zgodny, ponieważ $\lim_{n\to\infty} R \to 0$

Zadanie 3:

Mamy sześciościenną kostkę do gry, przy czym nie znamy prawdopodobieństwa wypadnięcia 6, oznaczonego przez p. W celu oszacowania p rzucamy kostką dopóki nie wypadnie 6 i przez Y oznaczamy liczbę wykonanych rzutów. Jednak jeśli w pierwszych k rzutach nie wypadła 6 to przerywamy eksperyment i $Y = k + 1$.

- Na podstawie n niezależych powtórzeń powyższego eksperymenty wyznacz estymator Metodą Największej Wiarogodności parametru p.
- Sprawdź, czy podany estymator jest estymatorem nieobciążonym
- Wyznacz ryzyko (średni błąd kwadratowy, MSE) dla otrzymanego estymatora. Czy otrzymany estymator jest zgodny? (*Obserwacja: dla estymatorów zgodnych* $\lim_{n\to\infty} MSE(\theta) \to 0$)

Rozwiązanie: W dalszej części zadania będe używał, że $p \in (0, 1)$, tzn kiedy pisze nierówność na p mam na myśli w dziedzinie określoności.
Z treści zadania wiemy że Y ma następujący rozkład:
$P(Y = s) = p(1-p)^{s-1}$ dla $s \in \{1, ..., k\}$.
$P(Y = k+1) = 1 - \sum_{i=1}^{k} p(1-p)^i = (1-p)^k$.
Stad nasza funkcja wiarygodności ma postać:
$L(Y_1, ..., Y_n, p) = P(Y_1 = y_1, ..., Y_n = y_n) =$ (z niezależności) $= P(Y_1 = y_1) * ... * P(Y_n = y_n) = A$.
Nasze zmienne losowe mają wyszczególnione prawdopodobieństwo dla $k+1$, stąd załóżmy, że ten wynik uzyskaliśmy w (BSO $0 \leq w \leq n$) "w" ostatnich próbach. Wiedząc to podstawiamy i otrzymujemy:
$A = p(1-p)^{y_1-1} * ... * p(1-p)^{y_{n-w}-1} * ((1-p)^k)^w$.
Niech $G(p) = Ln(L(Y_1, ..., Y_n, p))$. Oczywiście, ponieważ logarytm jest funkcja ściśle rosnącą to $G(p)$ przyjmuje maksimum w $p_0 \iff$ funkcja $L(Y_1, ..., Y_n, p)$ przyjmuje maksimum w p_0.
Podstawiając do definicji funkcji G dane otrzymujemy:
$G(P) = (n-w)ln(p) + (\sum_{i=1}^{n-w}(y_i - 1))ln(1-p) + (kw)ln(1-p)$.
Teraz liczymy pochodną funkcji G.
$\frac{\partial G}{\partial p} = \frac{n-w}{p} - \frac{(\sum_{i=1}^{n-w}(y_i-1))+kw}{1-p}$
Chcemy policzyć maksimum stąd z lematu Fermata pochodna, (o ile funkcja jest różniczkowalna jak w naszym przypadku), zeruje się w punkcie przyjmowania maksimum. Po przyrównaniu do 0 dostajemy:
$\frac{\partial G}{\partial p} = 0 \iff p = \frac{n-w}{kw+\sum_{i=1}^{n-w}(y_i)}$.
Pozostaje sprawdzić czy jest to maksimum. Zauważmy, że:
$\frac{\partial G}{\partial p} > 0 \iff p < \frac{n-w}{kw+\sum_{i=1}^{n-w}(y_i)}$
$\frac{\partial G}{\partial p} < 0 \iff p > \frac{n-w}{kw+\sum_{i=1}^{n-w}(y_i)}$
Czyli pochodna zmienia znak z czego wnioskujemy że punkt $p = \frac{n-w}{kw+\sum_{i=1}^{n-w}(y_i)}$ jest maksimum. Stąd nasze wyliczone p jest dokładnie szukanym parametrem największej wiarygodnosci.

Zadanie 4:
Niech $X_1, ..., X_n$ będzie próbą prostą z rozkładu Poissona o intensywności θ
$P(X_i = x) = \frac{\theta^x}{x!}e^{-\theta}$

- Znajdź $\hat{\theta}$ estymator Metodą Największej Wiarogodności parametru θ.
- Oblicz obciążenie oraz wariancję estymatora $\hat{\theta}$, uzyskanego w poprzednim podpunkcie.
- Jak duże powinno być n, żeby błąd średniokwadratowy dla $\theta = 1$ był mniejszy niż $0,01$, gdzie $MSE(\theta) = E_\theta[(\theta - \hat{\theta})^2]$.

Rozwiązanie:

Zmiennne losowe $X_1, ..., X_n$ są niezależne bo pochodzą z próby prostej, więc

$$L(X_1, ..., X_n, \theta) = \prod_{i=1}^{n} P(X_i|\theta) = \prod_{i=1}^{n} \frac{\theta^{X_i}}{X_i!}e^{-\theta} = \frac{\theta^{\sum_{i=1}^n X_i}}{\prod_{i=1}^n X_i!}e^{-\theta n}$$

$$log(L(X_1, ..., X_n, \theta)) = log(\theta)\sum_{i=1}^{n} X_i - \theta n - \sum_{i=1}^{n} log(X_i)$$

Teraz jeszcze pochodna dla znalezienia ekstremum:

$$\frac{\partial l(X_1, ..., X_n, \theta)}{\partial \theta} = \frac{\sum_{i=1}^n X_i}{\theta} - n = 0$$

$$\Leftrightarrow \theta = \frac{\sum_{i=1}^n X_i}{n} = \overline{X}$$

Czy to maksimum?

$$\frac{\partial^2 l(X_1, ..., X_n, \theta)}{\partial \theta^2} = -\frac{\sum_{i=1}^n X_i}{\theta^2} < 0$$

Więc $\hat{\theta} = \overline{X}$

Zmiennne losowe $X_1, ..., X_n$ mają rozkład Poissona z parametrem θ więc $\mathbb{E}(X_i) = \theta$ i $Var(X_i) = \theta$ dla każdego i

$$\mathbb{E}(\hat{\theta}) = \mathbb{E}(\frac{\sum_{i=1}^n X_i}{n}) = \frac{1}{n}\mathbb{E}(\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i) = \frac{1}{n}n\mathbb{E}(X_1) = \theta$$

Czyli estymator $\hat{\theta}$ jest nieobciążony, więc jego obciążenie wynosi 0.

$$Var(\hat{\theta}) = Var(\frac{\sum_{i=1}^n X_i}{n}) = \frac{1}{n^2}Var(\sum_{i=1}^{n} X_i) = \frac{1}{n^2}nVar(X_1) = \frac{\theta}{n}$$

$$\mathbb{E}[(\theta - \hat{\theta})^2] = \mathbb{E}[\theta^2 - 2\theta\hat{\theta} + \theta^2] = \mathbb{E}[\theta^2] - 2\mathbb{E}[\theta\hat{\theta}] + \mathbb{E}[\hat{\theta}^2] =$$

$$= \mathbb{E}[1] - 2\mathbb{E}[\hat{\theta}] + Var[\hat{\theta}] + (\mathbb{E}[\hat{\theta}])^2 = 1 - 2 + \frac{\theta}{n} + 1^2 = \frac{\theta}{n} = \frac{1}{n}$$

$$\frac{1}{n} < 0.01 \Leftrightarrow n > 100$$

Zadanie 5:
Niech $X_1, ..., X_n$ będzie próbą prostą z rozkładu Pareto o parametrach $a > 0, \theta > 0$ o gęstości $f_{\theta,a} = \frac{\theta a^\theta}{x^{\theta+1}} 1(x > a)$

- Znajdź $\hat{\theta}$ oraz $\hat{a}$ estymatory Metodą Największej Wiarogodności parametrów θ oraz a.
- Sprawdź, czy znalezione estymatory są nieobciążone.
- Wyznacz ryzyko (średni błąd kwadratowy, MSE) dla jednego z estymatorów.

Rozwiązanie: Wyznaczamy najpierw funkcję wiarygodności (standardowo jako iloczyn gęstości):

$$L(a, \theta; X_1, \ldots, X_n) = \prod_{i=1}^{n} \frac{\theta a^\theta}{X_i{}^{\theta+1}} 1\{X_i > a\} = \theta^n a^{n\theta} \prod_{i=1}^{n} \left(\frac{1}{X_i{}^{\theta+1}} \right) 1\{\min_{1 \le i \le n} X_i > a\}.$$

Dla wygody przechodzimy do logarytmicznej funkcji wiarygodności:

$$\begin{aligned} l(a, \theta; X_1, \ldots, X_n) &= \ln L(a, \theta; X_1, \ldots, X_n) = \ln \left(\theta^n a^{n\theta} \prod_{i=1}^{n} \frac{1}{X_i{}^{\theta+1}} \right) \\ &= n \ln \theta + n\theta \ln a - (\theta + 1) \sum_{i=1}^{n} \ln X_i, \; x_i > a \end{aligned}$$

Zauważmy, że jedynym składnikiem zależnym od a jest $n\theta \ln a$; logarytm jest funkcją ściśle rosnącą - więc zwiększanie parametru a będzie zwiększało funkcję wiarygodności. Zatem estymatorem największej wiarygodności dla parametru a będzie największa wartość a, jaką może przyjąć ten parametr. Ponieważ mamy $X_1, \ldots, X_n \in (a, +\infty)$, to największą dopuszczalną wartością a będzie $\min_{1 \le i \le n} X_i$. Wobec tego estymatorem największej wiarygodności dla parametru a jest

$$\hat{a} = \min_{1 \le i \le n} X_i.$$

Estymator parametru θ wyznaczymy bardziej standardowo, różniczkując logarytmiczną funkcję wiarygodności. Mamy

$$\begin{aligned} \frac{\partial}{\partial \theta} l(a, \theta; X_1, \ldots, X_n) &= \frac{\partial}{\partial \theta} \left(n \ln \theta + n\theta \ln a - (\theta + 1) \sum_{i=1}^{n} \ln X_i \right) \\ &= \frac{n}{\theta} + n \ln a - \sum_{i=1}^{n} \ln X_i. \end{aligned}$$

Przyrównując obliczoną pochodną do zera, otrzymujemy

$$\begin{aligned} &\frac{\partial}{\partial \theta} l(a, \theta; X_1, \ldots, X_n) = 0 \\ &\frac{n}{\theta} + n \ln a - \sum_{i=1}^{n} \ln X_i = 0 \\ &\theta \left(n \ln a - \sum_{i=1}^{n} \ln X_i \right) + n = 0 \\ &\theta = \frac{n}{\sum_{i=1}^{n} \ln X_i - n \ln a}. \end{aligned}$$

Druga pochodna to

$$\frac{\partial^2}{\partial \theta^2 s} l(a, \theta; X_1, \ldots, X_n) = \frac{\partial}{\partial \theta} \left(\frac{n}{\theta} + n \ln a - \sum_{i=1}^{n} \ln X_i \right) = -\frac{n}{\theta^2} < 0.$$

Druga pochodna jest, szczęśliwie, ujemna ($n, \theta > 0$); zatem rzeczywiście estymatorem największej wiarygodności dla parametru θ będzie (pamiętając, że wyznaczyliśmy już MLE dla a):

$$\hat{\theta} = \frac{n}{\sum_{i=1}^{n} \ln X_i - n \ln \hat{a}} = \frac{n}{\sum_{i=1}^{n} \ln \frac{X_i}{\hat{a}}}.$$

Sprawdzimy teraz, czy otrzymane estymatory są obciążone. Polecenie nie nakazuje wyznaczać obciążenia, więc dla $\hat{a}$ podamy argument dowodzący tylko, że $Bias(\hat{a}, a) \neq 0$, czyli $\hat{a}$ jest obciążony. Zauważmy, że $X_1, \ldots, X_n > a$ dla dowolnego n (tak jest zdefiniowany w poleceniu rozkład Pareto), więc dla każdego skończonego rozmiaru próby n mamy

$$\hat{a} = \min_{1 \leq i \leq n} X_i > a$$

(tj. estymator przeszacowuje a), zatem dla skończonych n jest $Bias(\hat{a}, a) = \mathbb{E}[\hat{\theta}] - \theta \neq 0$ (estymator będzie nieobciążony asymptotycznie).

Dla $\hat{\theta}$ już standardowo policzymy obciążenie. Zauważmy najpierw, że

$$\mathbb{E}[\hat{\theta}] = \mathbb{E}\left[\frac{n}{\sum_{i=1}^{n} \ln \frac{X_i}{a}}\right] = n\mathbb{E}\left[\frac{1}{\sum_{i=1}^{n} \ln \frac{X_i}{a}}\right].$$

Przypomnijmy, że dystrybuanta rozkładu (a, θ) zadana jest wzorem (zakładam, że można z tego typu rzeczy korzystać):

$$F(t) = 1 - \left(\frac{a}{t}\right)^{\theta}.$$

Zatem

$$\mathbb{P}\left(\ln \frac{X_i}{a} \leq t\right) = \mathbb{P}\left(\frac{X_i}{a} \leq \exp(t)\right) = \mathbb{P}(X_i \leq a \exp(t)) = 1 - \left(\frac{a}{a \exp(t)}\right)^{\theta} = 1 - \exp -\theta t.$$

Rozpoznajemy dystrybuantę rozkładu wykładniczego $\mathrm{Exp}(\theta)$ (o wartości oczekiwanej $\frac{1}{\theta}$). Stosując fakt z RP II stwierdzający, że suma n niezależnych zmiennych o rozkładzie $\mathrm{Exp}(\lambda)$ ma rozkład (n, λ), stwierdzamy, że

$$\sum_{i=1}^{n} \ln \frac{X_i}{a} \sim (n, \theta),$$

gdzie gęstość rozkładu (n, θ) to

$$f(x) = \frac{\theta^n}{\Gamma(n)} x^{n-1} \exp(-\theta x), \quad x > 0.$$

Szukamy zatem $\mathbb{E}[\hat{a}] = \mathbb{E}[\frac{n}{T}]$, gdzie $T \sim (n, \theta)$. Najpierw rachunkowo przekonujemy się (po drodze korzystając z własności wartości oczekiwanej rozkładu ciągłego o znanej gęstości oraz z zależności między funkcją Γ a silnią), że

$$\begin{aligned}
\mathbb{E}[\frac{1}{T}] &= \int_0^{\infty} \frac{1}{x} \cdot \frac{\theta^n}{\Gamma(n)} x^{n-1} \exp(-\theta x)\, dx = \int_0^{\infty} \frac{\theta^n}{\Gamma(n)} x^{n-2} \exp(-\theta x)\, dx = \\
&= \int_0^{\infty} \frac{\theta^n}{(n-1)!} x^{n-2} \exp(-\theta x)\, dx = \frac{\theta}{n-1} \int_0^{\infty} \frac{\theta^{n-1}}{(n-2)!} x^{n-2} \exp(-\theta x)\, dx = \\
&= \frac{\theta}{n-1} \int_0^{\infty} f_{(n-1,\theta)}(x)\, dx = \frac{\theta}{n-1} \cdot 1 = \frac{\theta}{n-1},
\end{aligned}$$

przy czym ostatnia całka jest równa 1, ponieważ okazuje się być całką z gęstości rozkładu $(n-1, \theta)$ po całym nośniku.

Wobec tego możemy wreszcie obliczyć

$$\mathbb{E}[\hat{a}] = \mathbb{E}[\frac{n}{T}] = \sum_{i=1}^{n} \mathbb{E}[\frac{1}{T}] = n\mathbb{E}[\frac{1}{T}] = \frac{n}{n-1}\theta.$$

Widzimy więc, że estymator $\hat{\theta}$ jest obciążony, konkretnie

$$Bias(\hat{\theta}, \theta) = \mathbb{E}[\hat{\theta}] - \theta = \frac{n}{n-1}\theta - \theta = \frac{1}{n-1}\theta$$

(ponownie: asymptotycznie dla $n \to \infty$ dostaniemy równość).

Pozostaje wyznaczyć błąd średniokwadratowy dla $\hat{\theta}$. Ogólnie mamy

$$MSE(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \theta)^2] = \mathrm{Var}(\hat{\theta}) + \mathbb{E}[\hat{\theta} - \theta]^2 = \mathrm{Var}(\hat{\theta}) + (Bias(\hat{\theta}, \theta))^2,$$

przy czym (zgodnie ze wcześniejszymi rozważaniami)

$$\mathrm{Var}(\hat{\theta}) = \mathrm{Var}\left(\frac{n}{T}\right) = n^2 \mathrm{Var}\left(\frac{1}{T}\right),$$

gdzie $T \sim (n, \theta)$. Moglibyśmy w tym miejscu zacząć znów z uporem całkować (daje się to policzyć), ale lepiej przypomnieć sobie (RP II), że $\frac{1}{(n,\theta)}$ ma tzw. odwrotny rozkład gamma $^{-1}(n, \theta)$ o znanych, stablicowanych parametrach, w szczególności o wariancji $\frac{\theta^2}{(n-1)^2(n-2)}$, $n > 2$. Stąd wyliczamy

$$\mathrm{Var}(\hat{\theta}) = n^2 \mathrm{Var}\left(\frac{1}{T}\right) = \frac{(n\theta)^2}{(n-1)^2(n-2)}.$$

ostatecznie

$$MSE(\hat{\theta}) = \mathrm{Var}(\hat{\theta}) + (Bias(\hat{\theta}, \theta))^2 = \frac{(n\theta)^2}{(n-1)^2(n-2)} + \frac{\theta^2}{(n-1)^2} = \theta^2 \frac{(n-1)(n+2)}{(n-2)}.$$

Zadanie 6:
Niech $X_1, ..., X_n$ będzie próba prostą z rozkładu Log-normalnego o parametrach $\mu, \sigma^2 > 0$, o gęstości

$$f_{\mu,\sigma^2} = \frac{1}{x\sqrt{2\pi}\sigma} \exp(-\frac{(\ln(x) - \mu)^2}{2\sigma^2})$$

- Znajdź $\hat{\mu}, \hat{\sigma}^2$ Estymatory Największej Wiarogodności parametrów μ, σ^2,
- Oblicz obciążenie oraz wariancje estymatora $\hat{\mu}$ uzyskanego w poprzednim podpunkcie,
- Jak duże powinno być n, żeby błąd średniokwadratowy dla $\mu = 0$, $MSE(0)$, był mniejszy niż 0.01, gdzie $MSE(\mu) = \mathbb{E}_\theta(\mu - \hat{\mu})^2$.

Rozwiązanie: Liczymy funkcję wiarygodności

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{x_i\sqrt{2\pi}\sigma} \exp(-\frac{(\ln x_i - \mu)^2}{2\sigma^2}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp(-\frac{\sum_{i=1}^{n}(\ln x_i - \mu)^2}{2\sigma^2}) \prod_{i=1}^{n} \frac{1}{x_i}$$

Policzmy jej logarytm

$$\ln L(\mu, \sigma^2) = -n \ln(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (\ln x_i - \mu)^2 - \sum_{i=1}^{n} \ln(x_i)$$

Policzmy pochodną po μ i przyrównajmy do zera

$$\ln L\mu = \frac{1}{\sigma^2} \sum_{i=1}^{n} \ln(x_i) - \frac{\mu}{\sigma^2} n = 0 \Rightarrow \mu = \frac{\sum_{i=1}^{n} \ln(x_i)}{n}$$

Druga pochodna

$$-\frac{n}{\sigma^2} < 0$$

zatem

$$\hat{\mu} = \frac{\sum_{i=1}^{n} \ln(x_i)}{n}.$$

Teraz policzmy pochodną po σ i przyrównajmy ją do zera

$$\ln L\sigma = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{n} (\ln x_i - \mu)^2 = 0 \Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n} (\ln x_i - \mu)^2}{n}$$

Korzystając z tego, że $\mathbb{E}(\ln x_i) = \mu$ oraz $Var(\ln x_i) = n\sigma^2$, policzmy obciążenie i wariancję estymatora $\hat{\mu}$

$$b(\hat{\mu}) = \mathbb{E}(\hat{\mu}) - \mu = \frac{\sum_{i=1}^{n} \mathbb{E}(\ln(x_i))}{n} - \mu = \frac{n\mu}{n} - \mu = 0$$

$$Var(\hat{\mu}) = Var(\frac{\sum_{i=1}^{n} \ln(x_i)}{n}) = \frac{1}{n^2} \sum_{i=1}^{n} Var(\ln(x_i)) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

MSE dla $\mu = 0$

$$MSE(\hat{\mu}) = Var(\hat{\mu}) - b(\hat{\mu}) = \frac{\sigma^2}{n} = \frac{\sum_{i=1}^{n} \ln^2(x_i)}{n^2} < 0.01$$

$$n > 10\sqrt{\sum_{i=1}^{n} \ln^2(x_i)}$$

Zadanie 7:
Niech $X_1, ..., X_n$ będzie próba prostą z rozkładu normalnego o parametrach $\mu, \sigma^2 > 0$.

- Znajdź $\hat{\mu}, \hat{\sigma}^2$ Estymatory Największej Wiarogodności parametrów μ, σ^2,
- Oblicz obciążenie oraz wariancję estymatora $\hat{\mu}$ uzyskanego w poprzednim podpunkcie,
- Jak duże powinno być n, żeby błąd średniokwadratowy dla $\mu = 0$, $MSE(0)$, był mniejszy niż 0.01, gdzie $MSE(\mu) = \mathbb{E}_\theta(\mu - \hat{\mu})^2$.

Rozwiązanie:

(a) Każda z obserwacji pochodzi z rozkładu normalnego o parametrach μ oraz $\sigma^2 > 0$. Jest to rozkład ciągły więc opisywany jest funkcją gestości. Ma ona postać:

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} exp(-\frac{(X-\mu)^2}{2\sigma^2})$$

(b) Funkcja wiarogodności. Przemnażamy wartości prawdopodobieństw dla każdej obserwacji z próby:

$$L(X_1, X_2, ..., X_n, p) = \frac{1}{\sigma \cdot \sqrt{2\pi}} exp(-\frac{(X_1-\mu)^2}{2\sigma^2}) \cdot \frac{1}{\sigma \cdot \sqrt{2\pi}} exp(-\frac{(X_2-\mu)^2}{2\sigma^2}) ... \frac{1}{\sigma \cdot \sqrt{2\pi}} exp(-\frac{(X_n-\mu)^2}{2\sigma^2})$$

$$= (\frac{1}{\sigma \cdot \sqrt{2\pi}})^n \cdot \prod_{i=1}^{n} exp(-\frac{(X_n-\mu)^2}{2\sigma^2})$$

(c) Zlogarytmowanie funkcji wiarogodności i obliczenie pochodnych cząstkowych po parametrach:

$$l(X_1, X_2, ..., X_n, p) = \ln L = -n \ln \sigma - \frac{n}{2} \ln 2\pi - \frac{1}{2\sigma^2} (\sum_{i=1}^{n} (X_i - \mu)^2)$$

$$\frac{\partial l}{\partial \mu} = -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (2\mu - 2X_i)$$

$$\frac{\partial l}{\partial \sigma^2} = \frac{1}{2\sigma^4} (\sum_{i=1}^{n} (X_i - \mu)^2) - \frac{n}{2\sigma^2}$$

bo

$$\frac{\partial}{\partial x} \ln \sqrt{x} = \frac{1}{2x}$$

(d) Przyrównanie pochodnych cząstkowych do zera (zakładamy $p \in (0, 1)$ bo oba parametry rozkładu normalnego są >0) :

$$\frac{\partial l}{\partial \mu} = 0 \Leftrightarrow \sum_{i=1}^{n} (2\mu - 2X_i) = 0$$

więc

$$\widetilde{\mu} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$$

$$\frac{\partial l}{\partial \sigma^2} = 0 \Leftrightarrow \frac{1}{\sigma^2}(\sum_{i=1}^{n}(X_i - \mu)^2) = n$$

więc

$$\widetilde{\sigma^2} = \frac{1}{n} \cdot \sum_{i=1}^{n}(X_i - \mu)^2$$

.

(e) Sprawdzenie, czy faktycznie uzyskaliśmy maksimum w punkcie $(\widetilde{\mu}, \widetilde{\sigma^2})$. W tym przypadku jest to oczywiste, bo badana funkcja logarytmu funkcji wiarogodności jest wklęsla, a więc znalezione ekstremum to maksimum.

(f) Obliczenie obciążenia estymatora $\widetilde{\mu}$

Najpierw liczymy $E[\widetilde{\mu}]$

$$E[\widetilde{\mu}] = E[\frac{1}{n} \cdot \sum_{i=1}^{n} X_i] = \frac{1}{n}E[\sum_{i=1}^{n} X_i] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n} \cdot n\mu = \mu$$

bo wartością oczekiwaną średniej dla rozkładu normalnego jest parametru μ stąd obciążenie estymatora $\widetilde{\mu}$ to:

$$b(\widetilde{\mu}) = \mu - \frac{1}{n} \cdot \sum_{i=1}^{n} X_i = 0$$

Stąd $(\widetilde{\mu})$ to estymator nieobciążony.

(g) Obliczenie wariancji estymatora $\widetilde{\mu}$

$$D^2[\widetilde{\mu}] = D^2[\frac{1}{n} \cdot \sum_{i=1}^{n} X_i] = \frac{1}{n^2}D^2[\sum_{i=1}^{n} X_i] = \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n}$$

(h) Oszacowanie błędu średniokwadratowego:

Błąd średniokwadratowy estymatora można obliczyć jako sumę wariancji estymatora i jego obciążenia podniesionego do kwadratu. W tym przypadku będzie on wynosił zatem

$$MSE(\widetilde{\mu}) = \frac{\sigma^2}{n} + 0^2 = \frac{\sigma^2}{n}$$

wówczas $MSE < 0.01$ wtw, gdy $n > \frac{\sigma^2}{0.01}$

Zadanie 8:
Liczba wypadków samochodowych zgłoszonych do towarzystwa ubezpieczeniowego w k-tym miesiącu jest zmienną losową W_k o rozkładzie Poissona z parametrem λz_k , gdzie z_k jest liczbą samochodów zgłoszonych do ubezpieczenia w tym miesiącu, zaś λ jest nieznanym parametrem. Zmienne losowe W_k są niezależne.

- Wyznaczyć estymator Metodą Największej Wiarogodności parametru λ na podstawie próby $W_1, \ldots, W_{12}$.
- Sprawdzić, czy ten estymator jest nieobciążony.
- Wyznaczyć ryzyko (średni błąd kwadratowy, MSE) dla uzyskanego estymatora.

Rozwiązanie: Wiemy, że

$$P(W_k = i) = \frac{(\lambda z_k)^i}{i!} e^{-\lambda z_k}$$

oraz W_k są niezależne.

Rozpiszmy funkcję wiarogodności.

$$L(w_1, ...w_{12}, \lambda) = \frac{(\lambda z_1)^{w_1}}{w_1!} e^{-\lambda z_1} \cdot ... \cdot \frac{(\lambda z_{12})^{w_{12}}}{w_{12}!} e^{-\lambda z_{12}} = \exp\Big(- \lambda \sum_{k=1}^{12} z_k \Big) \cdot \prod_{k=1}^{12} \frac{(\lambda z_k)^{w_k}}{w_k!}$$

L i $\ln(L) =: l$ przyjmują maksimum dla tego samego λ.

$$l(w_1, ...w_{12}, \lambda) = -\lambda \sum_{k=1}^{12} z_k + \sum_{k=1}^{12} w_k \ln(\lambda z_k) - \sum_{k=1}^{12} \ln(w_k)$$

$$\frac{\partial l}{\partial \lambda}(w_1, ...w_{12}, \lambda) = -\sum_{k=1}^{12} z_k + \sum_{k=1}^{12} \frac{w_k \cdot z_k}{\lambda z_k} = \sum_{k=1}^{12} \frac{w_k}{\lambda} - \sum_{k=1}^{12} z_k$$

Zatem:

$$\frac{\partial l}{\partial \lambda}(w_1, ...w_{12}, \lambda) = 0 \iff \lambda = \frac{\sum_{k=1}^{12} w_k}{\sum_{k=1}^{12} z_k}$$

Sprawdźmy, czy l ma maksimum w wyznaczonym

$$\frac{\sum_{k=1}^{12} w_k}{\sum_{k=1}^{12} z_k} =: \lambda_0.$$

$$\frac{\partial^2 l}{\partial^2 \lambda}(w_1, ...w_{12}, \lambda) = -\frac{1}{\lambda^2} \sum_{k=1}^{12} w_k < 0 \text{ dla każdego } \lambda, \text{ w tym dla } \lambda_0.$$

Wobec tego l rzeczywiście przyjmuje tam maksimum.

Sprawdźmy, czy estymator λ_0 jest nieobciążony.

$$E(\lambda_0) = E\left(\frac{\sum_{k=1}^{12} W_k}{\sum_{k=1}^{12} z_k}\right) = \frac{\lambda \sum_{k=1}^{12} z_k}{\sum_{k=1}^{12} z_k} = \lambda$$

Stąd jest to estymator nieobciążony.

Policzmy jego średni błąd kwadratowy (ryzyko).

$$MSE(\lambda_0) = E\left(\left(\frac{\sum_{k=1}^{12} W_k}{\sum_{k=1}^{12} z_k} - \lambda\right)^2\right) = \frac{1}{\left(\sum_{k=1}^{12} z_k\right)^2} E\left(\left(\sum_{k=1}^{12} W_k\right)^2 - 2\lambda \sum_{k=1}^{12} W_k \sum_{k=1}^{12} z_k + \lambda^2 \left(\sum_{k=1}^{12} z_k\right)^2\right) =$$

$$= \frac{1}{\left(\sum_{k=1}^{12} z_k\right)^2} E\left(\left(\sum_{k=1}^{12} W_k\right)^2\right) - \frac{1}{\left(\sum_{k=1}^{12} z_k\right)^2} \cdot 2\lambda^2 \left(\sum_{k=1}^{12} z_k\right)^2 + \frac{1}{\left(\sum_{k=1}^{12} z_k\right)^2} \cdot \lambda^2 \left(\sum_{k=1}^{12} z_k\right)^2 = \frac{E\left((\sum_{k=1}^{12} W_k)^2\right)}{\left(\sum_{k=1}^{12} z_k\right)^2} -$$

$$= \left\{ \begin{array}{l} W_k \text{ - zmienne niezależne } \Rightarrow \\ E(W_k W_j) = E(W_k)E(W_j) \text{ dla } i \neq j \\ \text{oraz} \\ E(W_k^2) = Var(W_k) + (E(W_k))^2 \\ Var(W_k) = \lambda z_k \end{array} \right\} = \frac{\sum_{k=1}^{12} \left((\lambda z_k)^2 + \lambda z_k\right) + 2\lambda^2 \sum_{1 \leq j < k \leq 12} z_j z_k}{\left(\sum_{k=1}^{12} z_k\right)^2} - \lambda^2 =$$

$$= \frac{\sum_{k=1}^{12} \lambda z_k + \lambda^2 \left(\sum_{k=1}^{12} z_k\right)^2}{\left(\sum_{k=1}^{12} z_k\right)^2} - \lambda^2 = \frac{\lambda}{\sum_{k=1}^{12} z_k}$$

Zadanie 9:
W jeziorze pływa pewna nieznana liczba θ ryb. Aby oszacować tę liczbę postępujemy następująco: odławiamy m ryb, znaczymy je, a następnie wpuszczamy do jeziora. Czekamy, aż ryby wymieszają się, łowimy n ryb i zliczamy liczbę ryb znaczonych.

- Wyznaczyć estymator Metodą Największej Wiarogodności parametru θ. Sprawdzić, czy ten estymator jest nieobciążony.
- Wyznaczyć ryzyko (średni błąd kwadratowy, MSE) dla uzyskanego estymatora.

Rozwiązanie: *Zadanie jest bardzo znanym przykładem wykorzystania Metody Największej Wiarogodności – daje podstawę stosowanej w terenowych badaniach liczebności Metodzie Wielokrotnych Złowień – Mark and Recapture.*

Przykładowy fragment rozwiązania:

`https://www.math.drexel.edu/~tolya/Rice%20%5B%20capture-recapture%20%5D.pdf`