# 2-nd homework

## The first part.

1. a) (4 points) Using the method of moments with the functions $g(x) = x^k, \quad k \in \mathbb{N}$, estimate the unknown parameter $\theta$, if sample is drawn from an exponential distribution $X_1, ..., X_n \sim Exp(\theta) = \theta \cdot Exp(1)$.
Explore analytically the behavior of the standard deviation of the constructed estimates. For which value of k is the estimate most efficient?
b) (4 points) Consider the same problem when the sample is drawn from $Exp(\frac{1}{\theta})$. In this case, compare the method of moments estimators with the empirical quantile $\widehat{Z_p}$ of level $p \in (0,1)$ (after transforming it into the corresponding form so that it serves as an estimator of $\theta$), and determine which of them is more efficient.
2. (2 points) Using the method of moments with the functions $g_1(x) = x, \quad g_2(x) = x^2$, estimate the vector of unknown parameters (a,b) for $X_1, ..., X_n \sim Unif[a, a+b]$.
3. (6 points (2 points per task)) a) Suppose a sample is drawn from a two-parameter exponential distribution with the density function

$$f_{\alpha,\beta}(y) = \begin{cases} \alpha^{-1} e^{-(y-\beta)/\alpha}, & \text{for } y \geq \beta, \\ 0, & \text{for } y < \beta. \end{cases}$$

Using the method of moments, estimate the scale parameter $\alpha > 0$ and the location parameter $\beta \in \mathbb{R}$.
b) Consider the same problem when the sample is drawm from Binomial distributions with parameters $m$ and $\theta$.
c) Let $P$ and $Q$ be two distributions with known expectations $a$ and $b$, respectively, where $a < b$. Let $P_\theta$ be a mixture of the distributions $P$ and $Q$:

$$P_\theta = \theta P + (1-\theta)Q, \qquad 0 \leq \theta \leq 1.$$

Using the method of moments, estimate the parameter $\theta$ based on a sample from the distribution $P_\theta$.

4*.(1 points) Why is the condition for $f(\cdot)$ of the transformation important for preserving asymptotic normality in Proposition 4.2? Give an example where asymptotic normality fails.

## The second part.

1. (8 points) Suppose a sample is drawn from the following distribution with $\theta$

| $X_i$ | 1 | 3 | 6 | 10 |
|---|---|---|---|---|
| $p_i$ | $\frac{1}{4}+\theta$ | $\frac{1}{4}-\theta$ | $\frac{1}{4}-\theta$ | $\frac{1}{4}+\theta$ |
| $\nu_i$ | 4 | 8 | 7 | 11 |

where $p_i$ are the theoretical probabilities and $\mu_i$ are the observed frequencies (counts). Estimate $\theta$ using the maximum likelihood method and using the method of moments (at least two different estimators). Compare these estimators with each other. Find the mean squared error for ML estimator and calculate the Fisher information.

2. (10 points). Suppose a sample is drawn from the following distribution with $\theta = (\theta_1, \theta_2)$:

| $X_i$ | -2 | -1 | 0 | 1 | 2 |
|---|---|---|---|---|---|
| $p_i$ | $\frac{1}{12} + \theta_1 + \theta_2$ | $\frac{1}{6} - \theta_2$ | $\frac{1}{4} - \theta_1$ | $\frac{1}{4} + \theta_2$ | $\frac{1}{4} - \theta_2$ |
| $\nu_i$ | 4 | 8 | 7 | 11 | 5 |

where $p_i$ are the theoretical probabilities and $\mu_i$ are the observed frequencies (counts). Estimate $\theta$ using the maximum likelihood method (4 points) and using the method of moments (4 points). Calculate the Fisher information (2 points).

3.(4 points) Find the mean squared error for MLE of the sample from Unif[0,$\theta$]. Calculate the Fisher Information. Explain the problem with the inequality of Rao-Cramer in this case.

4. (12 points (4 points per task )) Calculate the Fisher information and find the MLE of $\theta > 0$ and its MSE for the sample from the distribution with the following density:

a) $f(x, \theta) = \theta \cdot x^{\theta - 1}, \quad x \in [0, 1]$;

b) $f(x, \theta) = \frac{2x}{\theta^2}, \quad x \in [0, \theta]$;

c) $f(x, \theta) = \frac{e^{-|x|}}{2(1 - e^{-\theta})}, \quad |x| \leq \theta$.

5. (6 points). Suppose a sample is drawn from a two-parameter exponential distribution with the density function

$$f_{\alpha,\beta}(y) = \begin{cases} \alpha^{-1} e^{-(y-\beta)/\alpha}, & \text{for } y \geq \beta, \alpha > 0 \\ 0, & \text{for } y < \beta. \end{cases}$$

(3 points) Find the MLE for $(\alpha, \beta)$.

(3 points) Calculate the Fisher information for each parameter for the sample $(X_1, ..., X_n)$.

6* (1 point). Lets $\widehat{\theta}$ is the MLE for $\theta$ difined by the space $(X, P_\theta, \theta \in \Omega)$, where $P_\theta$. Consider $\phi$ which is the bijection between $\Theta$ and $\Xi$. Prove that in this case $\phi(\widehat{\theta})$ is MLE for some $\xi \in \Xi$.

# HW2. (deadline: Nov. 17th)

1. a) (4 points) Using the method of moments with the functions $g(x) = x^k$, $k \in \mathbb{N}$, estimate the unknown parameter $\theta$, if sample is drawn from an exponential distribution $X_1, ..., X_n \sim Exp(\theta) = \theta \cdot Exp(1)$.
Explore analytically the behavior of the standard deviation of the constructed estimates. For which value of k is the estimate most efficient?

Sol: $f_i(x) = \theta e^{-\theta x}$ $x > 0$. $\mathbb{E}[X_i] = \frac{1}{\theta}$.

$$\mathbb{E}[X^k] = \int_0^{+\infty} x^k \theta e^{-\theta x} dx \xlongequal[dx=\frac{1}{\theta}dt]{\theta x = t} \int_0^{+\infty} t^k \theta^{-k} \cdot \theta^{-1} \cdot \theta e^{-t} dt = \theta^{-k}\Gamma(k+1) = \theta^{-k} \cdot k!$$

i.e. $h(\theta) = \theta^{-k} \cdot k!$

the moment $m_k = \frac{1}{n}\sum_{i=1}^n X_i^k$, estimator $\hat{\theta}_k = h^{-1}(m_k) = \left(\frac{k!}{m_k}\right)^{1/k}$

$$\mathbb{D}g(X) = Var(X^k) = \mathbb{E}[X^{2k}] - (\mathbb{E}[X^k])^2 = \theta^{-2k}((2k)! - (k!)^2)$$

we need $((h^{-1})'(h(\theta)))^2$ $(h^{-1})'(y) = -\frac{(k!)^{1/k}}{k} y^{-\frac{1}{k}-1}$

$$h^{-1\prime}(h(\theta)) = -\frac{(k!)^{1/k}}{k}(\theta^{-k} \cdot k!)^{-\frac{k+1}{k}} = -\frac{(k!)^{1/k}}{k} \cdot \theta^{k+1} \cdot (k!)^{-\frac{k+1}{k}} = -\frac{(k!)^{-k}}{k}\theta^{k+1}$$

$$\mathbb{D}g(X) \cdot ((h^{-1})' h(\theta))^2 = \theta^{-2k}((2k)! - (k!)^2)\left(\frac{\theta^{1+k}}{k \cdot k!}\right)^2 = \frac{\theta^2}{k^2}\left(\frac{(2k)!}{(k!)^2} - 1\right)$$

by the asymptotic normality. $\sqrt{n} \cdot (\hat{\theta}_k - \theta) \xrightarrow{d} N\left(0, \frac{\theta^2}{k^2}\left(\frac{(2k)!}{(k!)^2} - 1\right)\right)$

$$\Rightarrow \quad \hat{\theta}_k \xrightarrow{d} N\left(\theta, \frac{\theta^2}{nk^2}\left(\frac{(2k)!}{(k!)^2} - 1\right)\right)$$

$$SD(\hat{\theta}_k) = \sqrt{Var\,\theta_k} = \frac{1}{\sqrt{n}} \cdot \frac{\theta}{k} \cdot \sqrt{\frac{(2k)!}{(k!)^2} - 1}$$

$\frac{1}{k}\sqrt{\frac{(2k)!}{(k!)^2} - 1}$ will increase w.r.t. k. thus the most efficient estimator is $k=1$.

b) (4 points) Consider the same problem when the sample is drawn from $Exp(\frac{1}{\theta})$. In this case, compare the method of moments estimators with the empirical quantile $\widehat{Z_p}$ of level $p \in (0, 1)$ (after transforming it into the corresponding form so that it serves as an estimator of $\theta$), and determine which of them is more efficient.

Sol: $f_\theta(x) = \frac{1}{\theta} e^{-\frac{1}{\theta}x}$

$$\mathbb{E}[X^k] = \int_0^{+\infty} x^k \cdot \frac{1}{\theta} e^{-\frac{x}{\theta}} dx \xlongequal[dx = \theta dt]{t = \frac{x}{\theta}} \int_0^{+\infty} \theta^k \cdot t^k \cdot \frac{1}{\theta} e^{-t} \cdot \theta dt = \theta^k \int_0^{+\infty} t^k e^{-t} dt = \theta^k \Gamma(k+1) = \theta^k k!$$

by a). we have the conclusion that $k=1$. the moment estimator is the most efficient

i.e. $\bar{X} = \theta \Rightarrow \hat{\theta}_1 = \frac{1}{n}\sum X_i$ $Var(\hat{\theta}_1) \approx \frac{\theta^2}{n}$ (by a). same computation).

the population quartile $\xi$ s.t. $F(\xi_p) = p$. i.e. $1 - e^{-\frac{1}{\theta}\xi_p} = p \Rightarrow \ln(1-p) = -\frac{1}{\theta}\xi_p \Rightarrow \theta = -\frac{\xi_p}{\ln(1-p)}$

by thm 2.1., $\sqrt{n}(\widehat{Z}_p - \xi_p) \xrightarrow{d} N\left(0, \frac{p(1-p)}{f^2(\xi_p)}\right)$. $Var(\widehat{Z}_p) = \frac{\theta^2 p(1-p)}{n(1-p)^2} = \frac{\theta^2}{n}\frac{p}{1-p}$.

$$Var(\hat{\theta}_p) = Var\left(-\frac{1}{\ln(1-p)}\widehat{Z}_p\right) = \frac{\theta^2}{n} \cdot \frac{p}{(1-p)\ln^2(1-p)}$$

consider func. $f(p) = \frac{p}{(1-p)\ln^2(1-p)}$ $p \in (0,1)$. $f(p)$ always $> 1$. $p \in (0,1)$

thus. $Var(\hat{\theta}_p) > Var(\hat{\theta}_1)$. the moment estimator of $k=1$ always more efficient.

2. (2 points) Using the method of moments with the functions $g_1(x) = x$, $g_2(x) = x^2$, estimate the vector of unknown parameters (a,b) for $X_1, ..., X_n \sim Unif[a, a+b]$.

Sol: for $Unif\ [a, a+b]$ $\mu = \frac{2a+b}{2}$ $\sigma^2 = \frac{b^2}{12}$

$$\mathbb{E}[X^2] = \int_a^{a+b} x^2 \cdot \frac{1}{b}\, dx = \frac{(a+b)^3 - a^3}{3b} = a^2 + ab + \frac{b^2}{3}$$

denote $m_1 = \bar{X}$ $m_2 = \frac{1}{n}\sum X_i^2$

$$\begin{cases} 2a+b = 2m_1 \\ a^2+ab+\frac{b^2}{3} = m_2 \end{cases} \Rightarrow \begin{cases} \hat{a} = m_1 - \sqrt{3}\cdot\sqrt{m_2 - m_1^2} \\ \hat{b} = 2\sqrt{3}\cdot\sqrt{m_2 - m_1^2} \end{cases}$$

3. (6 points (2 points per task)) a) Suppose a sample is drawn from a two-parameter exponential distribution with the density function

$$f_{\alpha,\beta}(y) = \begin{cases} \alpha^{-1}e^{-(y-\beta)/\alpha}, & \text{for } y \geq \beta, \\ 0, & \text{for } y < \beta. \end{cases}$$

Using the method of moments, estimate the scale parameter $\alpha > 0$ and the location parameter $\beta \in \mathbb{R}$.
b) Consider the same problem when the sample is drawm from Binomial distributions with parameters $m$ and $\theta$.
c) Let $P$ and $Q$ be two distributions with known expectations $a$ and $b$, respectively, where $a < b$. Let $P_\theta$ be a mixture of the distributions $P$ and $Q$:

$$P_\theta = \theta P + (1-\theta)Q, \qquad 0 \leq \theta \leq 1.$$

Using the method of moments, estimate the parameter $\theta$ based on a sample from the distribution $P_\theta$.

a) $\mathbb{E}[Y] = \int_\beta^{+\infty} y \cdot \alpha^{-1} e^{-\frac{y-\beta}{\alpha}} dy = -ye^{-\frac{y-\beta}{\alpha}}\Big|_\beta^{+\infty} + \int_\beta^{+\infty} e^{-\frac{y-\beta}{\alpha}} dy = \beta + -\alpha e^{-\frac{y-\beta}{\alpha}}\Big|_\beta^{+\infty} = \alpha+\beta.$

$\mathbb{E}[Y^2] = \int_\beta^{+\infty} y^2 \cdot \alpha^{-1} e^{-\frac{y-\beta}{\alpha}} dy = -y^2 e^{-\frac{y-\beta}{\alpha}}\Big|_\beta^{+\infty} + \int_\beta^{+\infty} 2y\, e^{-\frac{y-\beta}{\alpha}} dy = \beta^2 + 2\alpha(\alpha+\beta) = (\beta+\alpha)^2 + \alpha^2$

denote $m_1 = \frac{1}{n}\sum Y_i$. $m_2 = \frac{1}{n}\sum Y_i^2$

$$\Rightarrow \begin{cases} \alpha+\beta = m_1 \\ (\alpha+\beta)^2 + \alpha^2 = m_2 \end{cases} \Rightarrow \begin{cases} \hat{\alpha} = \sqrt{m_2 - m_1^2} \\ \hat{\beta} = m_1 - \sqrt{m_2 - m_1^2} \end{cases}$$

b). $X_i \sim Bin(m, \theta)$. $\mathbb{E}X = m\theta$. $Var(X) = m\theta(1-\theta)$.

$\mathbb{E}X^2 = m\theta(1-\theta) + m^2\theta^2$

denote $m_1 = \frac{1}{n}\sum X_i$ $m_2 = \frac{1}{n}\sum X_i^2$

$$\Rightarrow \begin{cases} m_1 = m\theta \\ m_2 = m\theta(1-\theta) + m^2\theta^2 \end{cases} \Rightarrow \begin{cases} \hat{m} = \frac{m_1^2}{m_1^2 + m_1 - m_2} \\ \hat{\theta} = 1 - \frac{m_2 - m_1^2}{m_1}. \end{cases}$$

c). $\mathbb{E}_{P_\theta}[X] = \theta a + (1-\theta)b = \theta(a-b) + b.$

denote $\bar{X} = \frac{1}{n}\sum X_i$ $\Rightarrow \hat{\theta} = \frac{\bar{X}-b}{a-b} = \frac{b-\bar{X}}{b-a}$ $(a<b)$.

4*.(1 points) Why is the condition for $f(\cdot)$ of the transformation important for preserving asymptotic normality in Proposition 4.2? Give an example where asymptotic normality fails.

We need $f$ cont. and $f'(a) \neq 0$ to make Taylor expansion

$f(T_n) = f(a) + (T_n - a) f'(a) + o(T_n - a)$.

example. $T_n = \bar{X}_n = \frac{1}{n}\sum X_i \quad X_i \sim N(0,1) \quad f(x) = x^2. \quad f(T_n) = \bar{X}_n^2$

$\sigma = 1$, $a = 0$ $\quad \sqrt{n}\,\bar{X}_n \xrightarrow{d} N(0,1)$.

but $f'(a) = 0$ $\quad (f'(0) = 2 \cdot 0 = 0)$. which makes Asymptotic normality fails.

consider $n f(T_n)$, we have $n f(T_n) \sim \chi_1^2$, which means. $\sqrt{n}(f(T_n))$ is not normal distribution.

# Second Part.

1. (8 points) Suppose a sample is drawn from the following distribution with $\theta$

| $X_i$ | 1 | 3 | 6 | 10 |
|---|---|---|---|---|
| $p_i$ | $\frac{1}{4}+\theta$ | $\frac{1}{4}-\theta$ | $\frac{1}{4}-\theta$ | $\frac{1}{4}+\theta$ |
| $\nu_i$ | 4 | 8 | 7 | 11 |

where $p_i$ are the theoretical probabilities and $\mu_i$ are the observed frequencies (counts). Estimate $\theta$ using the maximum likelihood method and using the method of moments (at least two different estimators). Compare these estimators with each other. Find the mean squared error for ML estimator and calculate the Fisher information.

Sol: 1) MLE.

$\ln \ell(\theta) = \sum \nu_i \ln p_i = 15 \ln(\frac{1}{4}+\theta) + 15 \ln(\frac{1}{4}-\theta)$

$\frac{\partial \ln \ell(\theta)}{\partial \theta} = 15 \cdot \left(\frac{1}{\frac{1}{4}+\theta} - \frac{1}{\frac{1}{4}-\theta}\right) = 15 \frac{-2\theta}{(\frac{1}{4}+\theta)(\frac{1}{4}-\theta)} \qquad \frac{\partial \ln L}{\partial p} = 0 \quad \Rightarrow \hat{\theta}_L = 0$.

2) Moment

i) $m_1 = \frac{1}{n}\sum X_i \qquad E[X] = (\frac{1}{4}+\theta) + 3(\frac{1}{4}-\theta) + 6(\frac{1}{4}-\theta) + 10(\frac{1}{4}+\theta) = \frac{1+3+6+10}{4} + (1+10-3-6)\theta = 5+2\theta$.

$\bar{X} = \frac{1\times4 + 3\times8 + 6\times7 + 10\times11}{30} = \frac{4+24+42+110}{30} = 6$

$5 + \hat{\theta}_1 = 6 \quad \Rightarrow \quad \hat{\theta}_1 = 0.5$

ii) $m_2 = \frac{1}{n}\sum X_i^2$

$E[X^2] = 1^2 \cdot (\frac{1}{4}+\theta) + 3^2(\frac{1}{4}-\theta) + 6^2(\frac{1}{4}-\theta) + 10^2(\frac{1}{4}+\theta) = \frac{1+9+36+100}{4} + (1-9-36+100)\theta = 36.5 + 56\theta$.

$\bar{X}^2 = \frac{4 + 9\times8 + 36\times7 + 100\times11}{30} = 47.6. \quad \Rightarrow \hat{\theta}_2 = \frac{111}{560}. \approx 0.198$

$E[X^4] = 1^4(\frac{1}{4}+\theta) + 3^4(\frac{1}{4}-\theta) + 6^4(\frac{1}{4}-\theta) + 10^4(\frac{1}{4}+\theta) = \frac{1^4+3^4+6^4+10^4}{4} + (10^4+1-3^4-6^4)\theta. = 2844.5 + 8624\theta$.

$Var(\hat{\theta}_2) = Var\left(\frac{m_2 - 36.5}{56}\right) = \frac{1}{n} \cdot \frac{1}{56^2} Var(X^2) = \frac{1}{56^2 n} [EX^4 - (EX^2)^2] = \frac{1512.25 + 4536\theta - 3136\theta^2}{3136\, n}$

$$I_n(\theta) = n I_1(\theta) = n \mathbb{E}_\theta\left[\left(\frac{\partial \ln f}{\partial \theta}\right)^2\right] = n \sum_{i=1}^{4}\left[\left(\frac{\partial p_i}{\partial \theta}\right)^2\right] \cdot p_i(\theta) = n \cdot \frac{[p_i'(\theta)]^2}{p_i(\theta)}$$

$$= n\left(\frac{2}{\theta+\frac{1}{4}} + \frac{2}{\frac{1}{4}-\theta}\right) = \frac{n}{\frac{1}{16}-\theta^2}$$

when $n=30$. Fisher information $I(\theta) = \frac{30}{\frac{1}{16}-\theta^2}$

by thm 4.3. $Var(\hat{\theta}_L) = \frac{1}{I_n(\theta)} = \frac{1-16\theta^2}{16n}$

$Var(\hat{\theta}_2) = \frac{1512 \cdot 25 + 4536\theta - 3136\theta^2}{3136 n^2}$ $\quad Var(\hat{\theta}_L) = \frac{196 - 3136\theta^2}{3136 n}$ $\quad Var(\hat{\theta}_L) > Var(\hat{\theta}_2)$.

$\hat{\theta}_1$ is not proper, since $\hat{\theta}_1 = 0.5$ will lead $p_i < 0$.

$\hat{\theta}_L$ is more efficient than $\hat{\theta}_2$

$$MSE(\hat{\theta}_L) = Var(\hat{\theta}_L) + (Bias(\hat{\theta}))^2$$

$$Bias(\hat{\theta}) = \mathbb{E}[\hat{\theta}] - \theta = \frac{1}{4}\sum p_i - \theta = 0.$$

thus $MSE(\hat{\theta}_L) = Var(\hat{\theta}_L) = \frac{1}{I_n(\theta)} = \frac{1-16\theta^2}{480}$

2. (10 points). Suppose a sample is drawn from the following distribution with $\theta = (\theta_1, \theta_2)$:

| $X_i$ | -2 | -1 | 0 | 1 | 2 |
|---|---|---|---|---|---|
| $p_i$ | $\frac{1}{12}+\theta_1+\theta_2$ | $\frac{1}{6}-\theta_2$ | $\frac{1}{4}-\theta_1$ | $\frac{1}{4}+\theta_2$ | $\frac{1}{4}-\theta_2$ |
| $\nu_i$ | 4 | 8 | 7 | 11 | 5 |

where $p_i$ are the theoretical probabilities and $\mu_i$ are the observed frequencies (counts). Estimate $\theta$ using the maximum likelihood method (4 points) and using the method of moments (4 points). Calculate the Fisher information (2 points).

i) MLE.

$\ln L(\theta) = \sum \nu_i \ln p_i(\theta)$

$$\frac{\partial \ln L(\theta)}{\partial \theta_1} = \sum \nu_i \frac{p_i'(\theta)}{p_i(\theta)} = 4 \cdot \frac{1}{\frac{1}{12}+\theta_1+\theta_2} + 7 \cdot \frac{-1}{\frac{1}{4}-\theta_1}$$

$$\frac{\partial \ln L(\theta)}{\partial \theta_2} = \sum \nu_i \frac{p_i'(\theta)}{p_i(\theta)} = 4 \cdot \frac{1}{\frac{1}{12}+\theta_1+\theta_2} + 8 \cdot \frac{-1}{\frac{1}{6}-\theta_2} + 11 \cdot \frac{1}{\frac{1}{4}+\theta_2} + 5 \cdot \frac{-1}{\frac{1}{4}-\theta_2}$$

Solve:

$$\begin{cases} 4 \cdot \frac{1}{\frac{1}{12}+\theta_1+\theta_2} + 7 \cdot \frac{-1}{\frac{1}{4}-\theta_1} = 0 \\ 4 \cdot \frac{1}{\frac{1}{12}+\theta_1+\theta_2} + 8 \cdot \frac{-1}{\frac{1}{6}-\theta_2} + 11 \cdot \frac{1}{\frac{1}{4}+\theta_2} + 5 \cdot \frac{-1}{\frac{1}{4}-\theta_2} = 0 \end{cases} \Rightarrow \begin{cases} \hat{\theta}_{L1} \approx 0.0292 \\ \hat{\theta}_{L2} \approx 0.0136 \end{cases}$$

(numerical sol. of this non-linear system).

$$4\left(\frac{1}{4}-\theta_1\right) + 7\left(\frac{1}{12}+\theta_1+\theta_2\right) = 0$$

ii) $\mathbb{E}(X) = \sum X_i p_i(\theta) = -2\left(\frac{1}{12}+\theta_1+\theta_2\right) + (-1)\left(\frac{1}{6}-\theta_2\right) + (1)\cdot\left(\frac{1}{4}+\theta_2\right) + 2\left(\frac{1}{4}-\theta_2\right) = \frac{5}{12} - 2(\theta_1+\theta_2)$

$\mathbb{E}(X^2) = \sum X_i^2 p_i(\theta) = 4\left(\frac{1}{12}+\theta_1+\theta_2\right) + (1)\left(\frac{1}{6}-\theta_2\right) + (1)\cdot\left(\frac{1}{4}+\theta_2\right) + 4\left(\frac{1}{4}-\theta_2\right) = \frac{7}{4} + 4\theta_1$

$\bar{X} = \frac{1}{n}\sum \nu_i X_i = \frac{1}{35}(-8-8+10+11) = \frac{1}{7}$

$\bar{X}^2 = \frac{1}{n}\sum \nu_i X_i^2 = \frac{1}{35}(16+8+11+20) = \frac{11}{7}$

$$\Rightarrow \begin{cases} 4\theta_1 + \frac{7}{4} = \frac{11}{4} \\ 2\theta_1 + 2\theta_2 + \frac{1}{7} = \frac{5}{12} \end{cases} \Rightarrow \begin{cases} \hat{\theta}_1 = -\frac{5}{112} \\ \hat{\theta}_2 = \frac{61}{336}. \end{cases}$$

$$I(\theta) = \sum p_i \left( \frac{\partial \ln p_i}{\partial \theta_j} \right)^2 = \sum_i p_i \frac{1}{p_i^2} \left( \frac{\partial p_i}{\partial \theta_j} \right)^2 = \sum \frac{1}{p_i} \left( \frac{\partial p_i}{\partial \theta_j} \right)^2$$

| $\frac{\partial p_i}{\partial \theta_j}$ | $-2$ | $-1$ | $0$ | $1$ | $2$ |
|---|---|---|---|---|---|
| $\theta_1$ | 1 | 0 | $-1$ | 0 | 0 |
| $\theta_2$ | 1 | $-1$ | 0 | 1 | $-1$ |

$$I_{11} = \frac{1}{\frac{1}{12}+\theta_1+\theta_2} + \frac{1}{\frac{1}{4}+\theta_2} \qquad I_{21} = I_{12} = \frac{1\cdot 1}{\frac{1}{12}+\theta_1+\theta_2} = \frac{1}{\frac{1}{12}+\theta_1+\theta_2}$$

$$I_{22} = \frac{1}{\frac{1}{12}+\theta_1+\theta_2} + \frac{1}{\frac{1}{6}-\theta_2} + \frac{1}{\frac{1}{4}+\theta_2} + \frac{1}{\frac{1}{4}-\theta_2} \qquad I(\theta) = n I_1(\theta) = 35 \begin{pmatrix} I_{11} & I_{21} \\ I_{12} & I_{22} \end{pmatrix}$$

3.(4 points) Find the mean squared error for MLE of the sample from Unif[0,$\theta$]. Calculate the Fisher Information. Explain the problem with the inequality of Rao-Cramer in this case.

Sol: $f_i(x,\theta) = \begin{cases} \frac{1}{\theta}, & 0 \le x \le \theta. \\ 0 . & \text{otherwise.} \end{cases} \quad (\theta > 0)$

$$L(\theta) = \Pi f_i(x,\theta) = \prod_{i=1}^{n} \left( \frac{1}{\theta} \cdot \mathbb{1}(0 \le X_i \le \theta) \right) = \begin{cases} \frac{1}{\theta^n} . & \theta \ge X_{(n)} \\ 0 , & \theta < X_{(n)} \end{cases}$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{-n \ln \theta}{\partial \theta} = -\frac{n}{\theta}. < 0., \text{ thus. } \hat{\theta}_L = \theta_{min} = X_{(n)}$$

for the order statistic. $g_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} [F(x)]^{k-1} [1-F(x)]^{n-k} f(x)$

thus here $g_{X_{(n)}}(x) = n \cdot \left(\frac{x}{\theta}\right)^{n-1} \cdot \frac{1}{\theta} = \frac{n x^{n-1}}{\theta^n} \qquad 0 \le x \le \theta.$

$$\mathbb{E}[X_{(n)}] = \int_0^\theta x g(x) = \frac{n}{\theta^n} \int_0^\theta x^n dx = \frac{n}{\theta^n} \cdot \frac{1}{n+1} x^{n+1} \Big|_0^\theta = \theta \cdot \frac{n}{n+1}$$

$$\mathbb{E}[X_{(n)}^2] = \int_0^\theta x^2 g(x) = \frac{n}{\theta^n} \int_0^\theta x^{n+1} dx = \frac{n}{\theta^n} \cdot \frac{1}{n+2} x^{n+2} \Big|_0^\theta = \frac{n}{n+2} \theta^2$$

$$MSE(\hat{\theta}_L) = \mathbb{E}[(\hat{\theta}_L - \theta)^2] = \mathbb{E}[X_{(n)}^2] - 2\theta \mathbb{E}[X_{(n)}] + \theta^2 = \left(\frac{n}{n+2}\theta^2\right) - 2\theta\left(\frac{n}{n+1}\theta\right) + \theta^2 = \theta^2 \left( \frac{n(n+1) - 2n(n+2) + (n+1)(n+2)}{(n+2)(n+1)} \right)$$

$$= \frac{2\theta^2}{(n+1)(n+2)}$$

$$I(\theta) = \int \left( \frac{\partial \ln L(\theta)}{\partial \theta} \right)^2 f_\theta(x) \mu(dx) = \int_0^\theta \left(-\frac{n}{\theta}\right)^2 \frac{1}{\theta} dx = \frac{n^2}{\theta^2}$$

check regularity $\frac{d}{d\theta} \int T(x) f(x,\theta) d\mu = \int T(x) f_\theta'(x,\theta) d\mu.$ for any $T$.

$$T(x) = 1. \qquad \frac{d}{d\theta} \int_0^\theta \frac{1}{\theta} dx = \frac{d}{d\theta} \left( \frac{1}{\theta} \cdot \theta \right) = 0 \qquad \int f_\theta'(x,\theta) d\mu = \int_0^\theta -\frac{1}{\theta^2} dx = -\frac{1}{\theta}$$

LHS $\neq$ RHS. the premise of the Rao-Cramer inequality doesn't holds.

4. (12 points (4 points per task )) Calculate the Fisher information and find the MLE of $\theta > 0$ and its MSE for the sample from the distribution with the following density:
a) $f(x, \theta) = \theta \cdot x^{\theta-1}, \quad x \in [0, 1]$;
b) $f(x, \theta) = \frac{2x}{\theta^2}, \quad x \in [0, \theta]$;
c) $f(x, \theta) = \frac{e^{-|x|}}{2(1-e^{-\theta})}, \quad |x| \leq \theta$.

a). $L(x,\theta) = \theta^n \left(\prod X_i\right)^{\theta-1} \quad x \in [0,1]$

$$\frac{\partial \ln L(x,\theta)}{\partial \theta} = \frac{\partial(n \ln\theta + (\theta-1)\ln\prod X_i)}{\partial\theta} = n\cdot\frac{1}{\theta} + \sum_{i=1}^{n} \ln X_i$$

$$\frac{\partial L(x,\theta)}{\partial\theta} = 0 \quad \Rightarrow \quad \hat{\theta}_L = \frac{-n}{\sum_{i=1}^{n} \ln X_i}$$

$$I_1(\theta) = \mathbb{E}_\theta\left(\frac{\partial L(x,\theta)}{\partial\theta}\right)^2 = \mathbb{E}_\theta\left(\frac{1}{\theta^2} + \frac{2}{\theta}\ln x + (\ln x)^2\right) = \int_0^1 \frac{1}{\theta} x^{\theta-1} dx + 2\int_0^1 x^{\theta-1}\ln x\, dx + \int_0^1 \theta x^{\theta-1}(\ln x)^2 dx$$

$$= \frac{1}{\theta}\cdot\frac{x^\theta}{\theta}\Big|_0^1 + 2\,\frac{(\ln x)\,x^\theta}{\theta}\Big|_0^1 - \frac{2}{\theta}\cdot\frac{x^\theta}{\theta}\Big|_0^1 + \theta\cdot\frac{(\ln x)^2 x^\theta}{\theta}\Big|_0^1 - \theta\int_0^1 \frac{x^\theta}{\theta}\cdot\frac{2\ln x}{x}dx$$

$$= \frac{1}{\theta^2} - \frac{2}{\theta^2} - \left(-\frac{2}{\theta^2}\right) = \frac{1}{\theta^2}$$

$$I_n(\theta) = n I_1(\theta) = \frac{n}{\theta^2}$$

denote $Y_i = -\ln X_i \qquad X_i \sim f(x) = \theta x^{\theta-1} \qquad F_X(x) = \int_0^x \theta t^{\theta-1}dt = x^\theta$

$F_Y(y) = P(Y \leq y) = P(-\ln X \leq y) = P(X \geq e^{-y}) = 1 - F_X(e^{-y}) = 1 - e^{-\theta y}$ which is $Exp(\frac{1}{\theta})$

denote $S = \sum Y_i$. $Y_i \sim Exp(\frac{1}{\theta})$. which means $S \sim Gamma(n, \theta)$.

$$\mathbb{E}[\hat{\theta}_L] = \int_0^{+\infty}\left(\frac{n}{S}\right)\cdot\frac{\theta^n}{\Gamma(n)} s^{n-1}e^{-\theta s}ds = \frac{n\theta^n}{(n-1)!}\cdot\frac{\Gamma(n-1)}{\theta^{n-1}} = \frac{n\theta^n}{(n-1)!}\cdot\frac{(n-2)!}{\theta^{n-1}} = \frac{n\theta}{n-1}$$

$$\mathbb{E}[\hat{\theta}_L^2] = \int_0^{+\infty}\frac{n^2}{s^2}\frac{\theta^n}{\Gamma(n)}s^{n-1}e^{-\theta s}ds = \frac{n^2\theta^n}{(n-1)!}\int_0^{+\infty}s^{n-3}e^{-\theta s}ds = \frac{n^2\theta^n}{(n-1)!}\frac{\Gamma(n-2)}{\theta^{n-2}} = \frac{n^2\theta^2}{(n-1)(n-2)}$$

$$MSE(\hat{\theta}_L) = \mathbb{E}[\hat{\theta}_L^2] - 2\theta\,\mathbb{E}[\hat{\theta}_L] + \theta^2 = \theta^2\left[\frac{n^2}{(n-1)(n-2)} - \frac{2n}{n-1} + 1\right] = \theta^2\left[\frac{n^2 - 2n(n-2) + (n-1)(n-2)}{(n-1)(n-2)}\right]$$

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

b). $L(x,\theta) = 2^n\cdot\theta^{-2n}\left(\prod_{i=1}^{n} X_i\cdot \mathbb{1}(0 \leq X_i \leq \theta)\right)$.

$$\frac{\partial\ln L(x,\theta)}{\partial\theta} = \frac{\partial(n\ln 2 - 2n\ln\theta + \sum\ln X_i\cdot \mathbb{1})}{\partial\theta} = -\frac{2n}{\theta}\,\mathbb{1}(X_{(n)} \leq \theta).$$

$\frac{\partial\ln L}{\partial\theta} < 0 \quad (\theta > 0)$. $\max L(\theta) = L(\theta = X_{(n)}) \quad \Rightarrow \quad \hat{\theta}_L = X_{(n)}$

$$I(\theta) = \mathbb{E}_\theta\left[\left(\frac{\partial\ln L}{\partial\theta}\right)^2\right] = \mathbb{E}_\theta\left[\frac{4n^2}{\theta^2}\right] = \frac{4n^2}{\theta^2}$$

by property of order statistics.

$$g_{X_{(n)}}(x) = n\left[\frac{x^2}{\theta^2}\right]^{n-1}\left(\frac{2x}{\theta^2}\right) = n\cdot\frac{x^{2n-2}}{\theta^{2n-2}}\cdot\frac{2x}{\theta^2} = \frac{2nx^{2n-1}}{\theta^{2n}}$$

$$\mathbb{E}[X_{(n)}] = \frac{2n}{\theta^{2n}}\int_0^\theta x^{2n}dx = \frac{2n}{\theta^{2n}}\cdot\frac{1}{2n+1}x^{2n+1}\Big|_0^\theta = \frac{2n}{2n+1}\theta.$$

$$\mathbb{E}[X_{(n)}^2] = \frac{2n}{\theta^{2n}}\int_0^\theta x^{2n+1}dx = \frac{2n}{\theta^{2n}}\cdot\frac{1}{2n+2}x^{2n+2}\Big|_0^\theta = \frac{n}{n+1}\theta^2$$

$$MSE[\hat{\theta}_L] = \mathbb{E}[X_{(n)}^2] - 2\theta\,\mathbb{E}[X_{(n)}] + \theta^2 = \theta^2\left[\frac{n}{n+1} - 4\frac{n}{2n+1} + 1\right] = \frac{\theta^2}{(n+1)(2n+1)}$$

c). $L(x,\theta) = \dfrac{e^{-\sum|X_i|} \cdot \mathbb{1}(|X|_{(n)} \le \theta)}{2^n (1-e^{-\theta})^n}$

$1-e^{-\theta}$ increase. $\dfrac{1}{(1-e^{-\theta})^n}$ decrease. thus $\max L(\theta) = L(\theta_{min})$, i.e. $\hat{\theta}_L = |X|_{(n)}$

$$\frac{\partial \ln L}{\partial\theta} = \frac{\partial(-\sum|X_i| - n\ln 2 - n\ln(1-e^{-\theta}))}{\partial\theta} = -n \cdot \frac{e^{-\theta}}{1-e^{-\theta}}$$

$$\mathbb{E}_\theta\left(\frac{\partial \ln L}{\partial\theta}\right)^2 = \mathbb{E}_\theta\left(\frac{n^2 e^{-2\theta}}{(1-e^{-\theta})^2}\right) = \frac{n^2 e^{-2\theta}}{(1-e^{-\theta})^2}$$

let $Y = |X|$. $F_Y(y) = \int_{-y}^{y} \dfrac{e^{-|x|}}{2(1-e^{-\theta})} dx = \dfrac{1}{1-e^{-\theta}} \int_0^y e^{-x} dx = \dfrac{1-e^{-y}}{1-e^{-\theta}}$. $\quad f_Y(y) = \dfrac{e^{-y}}{1-e^{-\theta}}$.

$g_{|X|_{(n)}}(y) = n \cdot [F_Y(y)]^{n-1} \cdot f_Y(y) = \dfrac{n(1-e^{-y})^{n-1} e^{-y}}{(1-e^{-\theta})^n}$ $\quad (y \in [0,\theta])$

$$\mathbb{E}[\hat{\theta}] = \frac{n}{(1-e^{-\theta})^n} \int_0^\theta y e^{-y} (1-e^{-y})^{n-1} dy = \frac{n}{(1-e^{-\theta})^n} \int_0^\theta y e^{-y} \left(\sum_{k=0}^{k-1} \binom{n-1}{k} (-1)^k e^{-ky}\right) dy$$

$$= \frac{n}{(1-e^{-\theta})^n} \sum \binom{n-1}{k} (-1)^k \int_0^\theta y e^{-(k+1)y} dy$$

$$\mathbb{E}[\hat{\theta}_L^2] = \frac{n}{(1-e^{-\theta})^n} \int_0^\theta y^2 e^{-y} (1-e^{-y})^{n-1} dy$$

$$MSE[\hat{\theta}_L] = \mathbb{E}[\hat{\theta}_L] - 2\theta \mathbb{E}[\hat{\theta}_L] + \theta^2$$

5. (6 points). Suppose a sample is drawn from a two-parameter exponential distribution with the density function

$$f_{\alpha,\beta}(y) = \begin{cases} \alpha^{-1} e^{-(y-\beta)/\alpha}, & \text{for } y \ge \beta, \alpha > 0 \\ 0, & \text{for } y < \beta. \end{cases}$$

(3 points) a) Find the MLE for $(\alpha, \beta)$.
(3 points) b) Calculate the Fisher information for each parameter for the sample $(X_1, ..., X_n)$.

Sol: a). $\ln L(\alpha,\beta,y) = \left(-n\ln\alpha + \left(-\frac{1}{\alpha}\sum y_i\right) + \frac{n\beta}{\alpha}\right) \cdot \mathbb{1}(Y_{(1)} \ge \beta)$.

$$\frac{\partial \ln L(\alpha,\beta,y)}{\partial\alpha} = -\frac{n}{\alpha} + \frac{1}{\alpha^2}\sum y_i - \frac{n\beta}{\alpha^2}$$

$\dfrac{\partial \ln L(\alpha,\beta,y)}{\partial\beta} = \dfrac{n}{\alpha} > 0$. fixed $\alpha$. $L(\beta)$ is increasing, thus $L_{max}$ when $\hat{\beta} = \beta_{min} = Y_{(1)}$

$\Rightarrow -\frac{n}{\alpha} + \frac{1}{\alpha^2} \cdot \sum(Y_i - Y_{(1)}) \Rightarrow \hat{\alpha} = \frac{1}{n}\sum(Y_i - Y_{(1)})$.

b). $I_n(\alpha,\beta) = n I_1(\alpha,\beta)$

$$\mathbb{E}\left[\left(\frac{\partial \ln f}{\partial\alpha}\right)^2\right] = \mathbb{E}\left[\left(\frac{y-\beta-\alpha}{\alpha^2}\right)^2\right] = \frac{1}{\alpha^4}\mathbb{E}[((y-\beta)-\alpha)^2]$$

we have $\mathbb{E}(y-\beta) = \int_\beta^{+\infty} (y-\beta)(\alpha^{-1} e^{-(y-\beta)/\alpha}) dy \overset{\lambda = y-\beta}{=\!=\!=} \int_0^{+\infty} \lambda \alpha^{-1} \cdot e^{-\frac{\lambda}{\alpha}} d\lambda = \alpha$.

thus $\mathbb{E}[(y-\beta-\mathbb{E}(y-\beta))^2] = Var(y-\beta) = \alpha^2$

$$\mathbb{E}\left[\left(\frac{\partial \ln f}{\partial\alpha}\right)^2\right] = \frac{1}{\alpha^2}$$

$$\mathbb{E}\left[\left(\frac{\partial \ln f}{\partial \alpha}\right)\left(\frac{\partial \ln f}{\partial \beta}\right)\right] = \mathbb{E}\left(\frac{y-\beta-\alpha}{\alpha^2} \cdot \frac{1}{\alpha}\right) = \frac{1}{\alpha^3}\mathbb{E}\left(y-\beta-\alpha\right) = \frac{1}{\alpha^3}\left(\mathbb{E}[y-\beta] + \mathbb{E}[-\alpha]\right) = \frac{1}{\alpha^3}(\alpha-\alpha) = 0$$

$$\mathbb{E}\left[\left(\frac{\partial \ln f}{\partial \beta}\right)^2\right] = \mathbb{E}\left[\frac{1}{\alpha^2}\right] = \frac{1}{\alpha^2}$$

$$I_n(\alpha, \beta) = \begin{pmatrix} \frac{n}{\alpha^2} & 0 \\ 0 & \frac{n}{\alpha^2} \end{pmatrix}$$

6* (1 point). Lets $\widehat{\theta}$ is the MLE for $\theta$ difined by the space $(X, P_\theta, \theta \in \Omega)$, where $P_\theta$. Consider $\phi$ which is the bijection between $\Theta$ and $\Xi$. Prove that in this case $\phi(\widehat{\theta})$ is MLE for some $\xi \in \Xi$.

Pf: by def of MLE. $\forall \theta \in \Omega$. $L(X, \hat{\theta}) \geq L(X, \theta)$.

denote $\xi = \phi(\theta)$. $\phi$ is bijection between $\Theta$, $\Xi$

define a new likelihood func. $L_1$, s.t. $L_1(\xi) = L(X, \phi^{-1}(\xi))$.

since they are obtained by the same sampling.

$L_1(\phi(\hat{\theta})) = L(X, \phi^{-1}(\phi(\hat{\theta}))) = L(X, \hat{\theta})$.

thus $L_1(\phi(\hat{\theta})) = L(X, \hat{\theta}) \geq L(X; \phi^{-1}(\xi)) = L_1(\xi) \quad \forall \xi \in \Xi$

thus. $\phi(\hat{\theta})$ is MLE of $\xi$.