Mathematical Statistics

Tutorial 6

1. Let $X_1, \dots, X_n$ be an independent and identically distributed (i.i.d.) random sample from the density:
$$f(x, \theta) = \theta x^{\theta - 1} I_{(0,1)}(x) \quad \text{where } \theta \in \Theta = (0, +\infty)$$
We are interested in estimating the parameter $\lambda = 1/\theta$.
The maximum Likelihood Estimator (MLE) of λ is $\hat{\lambda}_{MV} = -\frac{\sum_{i=1}^{n} \ln x_i}{n} = -\overline{\ln X}$. and the moment estimador based on the first population moment $\hat{\lambda}_{MO} = \frac{1-\bar{X}}{\bar{X}} = \frac{1}{\bar{X}} - 1$.
   (a) Compute the Cramér-Rao Lower Bound (CRLB) and compare it with the variance of the MLE estimators.
   (b) Compare the asymptotic variances of $\hat{\lambda}_{MV}$ and $\hat{\lambda}_{MO}$ and determine which is asymptotically more efficient.

2. Let $X_1, X_2, \dots, X_n$ be an independent and identically distributed (i.i.d.) random sample from a Normal distribution $N(\mu, \rho^2)$, where $\mu \in \mathbb{R}$ and $\rho^2 \in \mathbb{R}^+$. Use the Neyman Factorization Theorem to find a sufficient statistic, $T(\mathbf{X})$, for the parameter(s) specified in each of the following cases:
   (a) Case 1: Both Parameters Unknown. Find a sufficient statistic for the parameter vector $\boldsymbol{\theta} = (\mu, \rho^2)$, where both the mean μ and the variance ρ^2 are unknown.
   (b) Case 2: Mean Unknown, Variance Known. Find a sufficient statistic for the parameter $\theta = \mu$, where the variance ρ^2 is known (a fixed constant).
   (c) Case 3: Mean Known, Variance Unknown. Find a sufficient statistic for the parameter $\theta = \rho^2$, where the mean μ is known (a fixed constant).

3. Let $X_1, X_2, \dots, X_n$ be an i.i.d. random sample from a Shifted Exponential distribution with the PDF:
$$f(x; \theta) = e^{-(x-\theta)} \mathbb{I}_{(\theta, +\infty)}(x), \quad \theta \in \mathbb{R}$$
Use the Neyman Factorization Theorem to find a sufficient statistic for the parameter θ.

Problem 1

Let $X_1, \ldots, X_n$ be an independent and identically distributed (i.i.d.) random sample from the density:

$$f(x, \theta) = \theta x^{\theta - 1} I_{(0,1)}(x) \quad \text{where } \theta \in \Theta = (0, +\infty)$$

We are interested in estimating the parameter $\lambda = 1/\theta$.

The maximum Likelihood Estimator (MLE) of λ is $\hat{\lambda}_{MV} = -\frac{\sum_{i=1}^{n} \ln x_i}{n} = -\overline{\ln X}$. and the moment estimador based on the first population moment $\hat{\lambda}_{MO} = \frac{1-\bar{X}}{\bar{X}} = \frac{1}{\bar{X}} - 1$.

1. Compute the Cramér-Rao Lower Bound (CRLB) and compare it with the variance of the MLE estimators.
2. Compare the asymptotic variances of $\hat{\lambda}_{MV}$ and $\hat{\lambda}_{MO}$ and determine which is asymptotically more efficient.

Solution

We compute the Fisher Information for the parameter λ.

We compute the Fisher Information for θ ($I_1(\theta)$): The second derivative of the log-likelihood for one observation X is $\frac{d^2 \ln f(X|\theta)}{d\theta^2} = -\frac{1}{\theta^2}$.

$$I_1(\theta) = -\mathbb{E}\left[-\frac{1}{\theta^2}\right] = \frac{1}{\theta^2}$$

then $\lambda = q(\theta) = 1/\theta$. The Cramer-Rao Bound is

$$\frac{(q'(\theta))^2}{nI_1(\theta)} = \frac{(-\frac{1}{\theta^2})^2}{n\frac{1}{\theta^2}} = \frac{1}{n\theta^2}.$$

Also we can rewrite the bound in terms of λ. Using the transformation $\theta = 1/\lambda$,

$$\frac{(q'(\theta))^2}{nI_1(\theta)} = \frac{1}{n\theta^2} = \frac{\lambda^2}{n}.$$

Another posibility could be rewrite the density as a function of λ and compute the information number, $I_1(\lambda)$. We can see that the

$$I_1(\lambda) = I_1(\theta)\left(\frac{d\theta}{d\lambda}\right)^2 = \frac{1}{(1/\lambda)^2}\left(-\frac{1}{\lambda^2}\right)^2 = \lambda^2 \cdot \frac{1}{\lambda^4} = \frac{1}{\lambda^2}$$

then

$$\text{CRLB}(\lambda) = \frac{1}{nI_1(\lambda)} = \frac{1}{n/\lambda^2} = \frac{\lambda^2}{n}$$

Comparison with Estimator Variances of MLE ($\hat{\lambda}_{MV}$)

It is easy to see that $Y_i = -\ln X_i \sim \text{Exponential}(\theta)$, so $\hat{\lambda}_{MV} = \bar{Y}$. The variance of $\hat{\lambda}_{MV}$ is $\text{Var}(\hat{\lambda}_{MV}) = \frac{\text{Var}(Y_i)}{n}$. Since $\text{Var}(Y_i) = 1/\theta^2 = \lambda^2$:

$$\text{Var}(\hat{\lambda}_{MV}) = \frac{\lambda^2}{n}$$

The MLE is efficient for finite samples because $\text{Var}(\hat{\lambda}_{MV}) = \text{CRLB}(\lambda)$.

Asymptotic Distribution of $\hat{\lambda}_{MV}$ (MLE)

The asymptotic distribution is:

$$\sqrt{n}(\hat{\lambda}_{MV} - q(\theta)) \xrightarrow{D} N(0, (q'(\theta))^2/(I_1(\theta)))$$

then the asymprtotic variance is

$$\text{AVar}(\hat{\lambda}_{MV}) = \frac{1}{\theta^2} = \lambda^2$$

Asymptotic Distribution of $\hat{\lambda}_{MO}$ (MME)

We can compute the mean and the variance of X

- Mean: $\mu_1 = \mathbb{E}[X] = \frac{\theta}{\theta+1}$.
- Variance: $\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$.

$$\text{Var}(X) = \frac{\theta}{\theta+2} - \left(\frac{\theta}{\theta+1}\right)^2 = \frac{\theta(\theta+1)^2 - \theta^2(\theta+2)}{(\theta+2)(\theta+1)^2} = \frac{\theta}{(\theta+2)(\theta+1)^2}$$

Since the MME is $\hat{\lambda}_{MO} = \frac{1}{\bar{X}} - 1$, we define the function $g(m)$

$$g(m) = \frac{1}{m} - 1$$

Then $\hat{\lambda}_{MO} = \frac{1}{\bar{X}} - 1 = g(\bar{X})$ and $g(\frac{\theta}{\theta+1}) = \frac{1}{\theta}$ The using the Central Limit Theorem and applying the Delta Method, we have that

$$\sqrt{n}(\hat{\lambda}_{MO} - q(\theta)) \xrightarrow{D} N(0, \text{AVar}(\hat{\lambda}_{MO}))$$

where the asymptotic variance of $\hat{\lambda}_{MO}$ is given by:

$$\text{AVar}(\hat{\lambda}_{MO}) = \text{Var}(X) \cdot \left[g'(\mathbb{E}[X])\right]^2$$

First, we compute the derivative of $g(m)$ with respect to m:

$$g'(m) = \frac{d}{dm}\left(\frac{1}{m} - 1\right) = -\frac{1}{m^2}$$

Now we evaluate this derivative at the population mean $\mathbb{E}[X] = \frac{\theta}{\theta+1}$:

$$g'(\mathbb{E}[X]) = -\frac{1}{\left(\frac{\theta}{\theta+1}\right)^2} = -\frac{(\theta+1)^2}{\theta^2}$$

Substituting $\text{Var}(X)$ and $g'(\mathbb{E}[X])$ into the Delta Method formula:

$$\begin{aligned}
\text{AVar}(\hat{\lambda}_{MO}) &= \left(\frac{\theta}{(\theta+2)(\theta+1)^2}\right) \cdot \left(-\frac{(\theta+1)^2}{\theta^2}\right)^2 \\
&= \left(\frac{\theta}{(\theta+2)(\theta+1)^2}\right) \cdot \left(\frac{(\theta+1)^4}{\theta^4}\right) \\
&= \frac{(\theta+1)^2}{\theta^3(\theta+2)}
\end{aligned}$$

The Asymptotic Variance of the MME in terms of θ is:

$$\text{AVar}(\hat{\lambda}_{MO}) = \frac{(\theta+1)^2}{\theta^3(\theta+2)}$$

The asymptotic distribution is:

$$\sqrt{n}(\hat{\lambda}_{MO} - q(\theta)) \xrightarrow{D} N\left(0, \frac{(\theta+1)^2}{\theta^3(\theta+2)}\right)$$

Comparison of Asymptotic Variances

We compare $\text{AVar}(\hat{\lambda}_{MV}) = 1/\theta^2$ with $\text{AVar}(\hat{\lambda}_{MO}) = \frac{(\theta+1)^2}{\theta^3(\theta+2)}$.

$$\frac{\frac{(\theta+1)^2}{\theta^3(\theta+2)}}{1/\theta^2} = \frac{(\theta+1)^2}{\theta(\theta+2)}$$

Since $\theta > 0$, the ratio is greater than 1. Therefore, $\text{AVar}(\hat{\lambda}_{MV}) < \text{AVar}(\hat{\lambda}_{MO})$.

The Maximum Likelihood Estimator ($\hat{\lambda}_{MV}$) is **asymptotically more efficient** than the Method of Moments Estimator ($\hat{\lambda}_{MO}$), as it achieves the minimum possible asymptotic variance (λ^2).

Problem 2

Let $X_1, X_2, \ldots, X_n$ be an independent and identically distributed (i.i.d.) random sample from a Normal distribution $N(\mu, \rho^2)$, where $\mu \in \mathbb{R}$ and $\rho^2 \in \mathbb{R}^+$. Use the Neyman Factorization Theorem to find a sufficient statistic, $T(\mathbf{X})$, for the parameter(s) specified in each of the following cases:

1. Case 1: Both Parameters Unknown. Find a sufficient statistic for the parameter vector $\boldsymbol{\theta} = (\mu, \rho^2)$, where both the mean μ and the variance ρ^2 are unknown.

2. Case 2: Mean Unknown, Variance Known. Find a sufficient statistic for the parameter $\theta = \mu$, where the variance ρ^2 is known (a fixed constant).

3. Case 3: Mean Known, Variance Unknown. Find a sufficient statistic for the parameter $\theta = \rho^2$, where the mean μ is known (a fixed constant).

Solution

The probability density function (PDF) for a single observation X_i is:

$$f(x_i; \mu, \rho^2) = \frac{1}{\sqrt{2\pi\rho^2}} \exp\left(-\frac{(x_i - \mu)^2}{2\rho^2}\right)$$

The likelihood function for the i.i.d. sample $\mathbf{x} = (x_1, \ldots, x_n)$ is:

$$L(\mathbf{x}; \mu, \rho^2) = \prod_{i=1}^{n} f(x_i; \mu, \rho^2) = \left(\frac{1}{2\pi\rho^2}\right)^{n/2} \exp\left(-\frac{1}{2\rho^2}\sum_{i=1}^{n}(x_i - \mu)^2\right)$$

We expand the sum of squares in the exponent: $\sum(x_i - \mu)^2 = \sum x_i^2 - 2\mu\sum x_i + n\mu^2$.

$$L(\mathbf{x}; \mu, \rho^2) = \left(\frac{1}{2\pi\rho^2}\right)^{n/2} \exp\left(-\frac{1}{2\rho^2}\left(\sum x_i^2 - 2\mu\sum x_i + n\mu^2\right)\right)$$

According to the **Neyman Factorization Theorem**, $T(\mathbf{X})$ is a sufficient statistic for $\boldsymbol{\theta}$ if $L(\mathbf{x}; \boldsymbol{\theta})$ can be factored as $g(T(\mathbf{x}); \boldsymbol{\theta})h(\mathbf{x})$, where $h(\mathbf{x})$ does not depend on $\boldsymbol{\theta}$.

1. Case 1: Both Parameters Unknown ($\boldsymbol{\theta} = (\mu, \rho^2)$)

Since $h(\mathbf{x})$ must not depend on μ or ρ^2, we can set $h(\mathbf{x}) = 1$. The entire likelihood function must then be the function g.

$$L(\mathbf{x}; \mu, \rho^2) = \underbrace{\left(\frac{1}{2\pi\rho^2}\right)^{n/2} \exp\left(-\frac{1}{2\rho^2}\left(n\mu^2 - 2\mu\sum x_i + \sum x_i^2\right)\right)}_{g(T_1(\mathbf{x});\mu,\rho^2)} \cdot \underbrace{1}_{h(\mathbf{x})}$$

The function g depends on the sample $\mathbf{x}$ only through the values of $\sum x_i$ and $\sum x_i^2$.

Sufficient Statistic:

$$\mathbf{T}_1(\mathbf{X}) = \left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2\right)$$

2. Case 2: μ Unknown ($\theta = \mu$), ρ^2 Known

Since ρ^2 is known, terms involving only ρ^2 can be placed in $h(\mathbf{x})$. We rearrange the exponent to separate terms containing μ from terms that do not:

$$L(\mathbf{x}; \mu) = \underbrace{\exp\left(\frac{\mu}{\rho^2}\sum x_i - \frac{n\mu^2}{2\rho^2}\right)}_{g(T_2(\mathbf{x});\mu)} \cdot \underbrace{\left(\frac{1}{2\pi\rho^2}\right)^{n/2} \exp\left(-\frac{1}{2\rho^2}\sum x_i^2\right)}_{h(\mathbf{x})}$$

The function g depends on $\mathbf{x}$ only through $\sum x_i$ and depends on the unknown parameter μ. The function h depends on $\mathbf{x}$ but is independent of μ.

Sufficient Statistic:

$$T_2(\mathbf{X}) = \sum_{i=1}^{n} X_i \quad \text{or equivalently } \bar{X}$$

3. Case 3: ρ^2 Unknown ($\theta = \rho^2$), μ Known

Since μ is known, we use the initial unexpanded form of the exponent $\sum(x_i - \mu)^2$.

$$L(\mathbf{x}; \rho^2) = \underbrace{\left(\frac{1}{2\pi\rho^2}\right)^{n/2} \exp\left(-\frac{1}{2\rho^2}\sum_{i=1}^{n}(x_i - \mu)^2\right)}_{g(T_3(\mathbf{x});\rho^2)} \cdot \underbrace{1}_{h(\mathbf{x})}$$

The function g depends on $\mathbf{x}$ only through the quantity $\sum_{i=1}^{n}(x_i - \mu)^2$ and depends on the unknown parameter ρ^2.

Sufficient Statistic:

$$T_3(\mathbf{X}) = \sum_{i=1}^{n}(X_i - \mu)^2$$

Problem 3

Let $X_1, X_2, \ldots, X_n$ be an i.i.d. random sample from a Shifted Exponential distribution with the PDF:

$$f(x;\theta) = e^{-(x-\theta)}\mathbb{I}_{(\theta,+\infty)}(x), \quad \theta \in \mathbb{R}$$

Use the Neyman Factorization Theorem to find a sufficient statistic for the parameter θ.

Solution

The joint PDF for the sample $\mathbf{X}$ is:

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x};\theta) &= \prod_{i=1}^{n} f(x_i;\theta) \\ &= \prod_{i=1}^{n} e^{-(x_i-\theta)}\mathbb{I}_{(\theta,+\infty)}(x_i) \\ &= e^{-\sum_{i=1}^{n}(x_i-\theta)} \prod_{i=1}^{n} \mathbb{I}_{(\theta,+\infty)}(x_i) \\ &= e^{n\theta} e^{-\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} \mathbb{I}_{(\theta,+\infty)}(x_i) \end{aligned}$$

The product of indicator functions is non-zero (equal to 1) if and only if $x_i > \theta$ for all $i = 1, \ldots, n$, which is equivalent to requiring that the smallest observation is greater than θ: $\min\{x_i\} > \theta$.

$$\prod_{i=1}^{n} \mathbb{I}_{(\theta,+\infty)}(x_i) = \mathbb{I}_{(\theta,+\infty)}(\min\{x_i\})$$

Substituting this back, we factor the joint PDF:

$$f_{\mathbf{X}}(\mathbf{x};\theta) = \underbrace{e^{n\theta}\mathbb{I}_{(\theta,+\infty)}(\min\{x_i\})}_{k_1(t(\mathbf{x}),\theta)} \underbrace{e^{-\sum_{i=1}^{n} x_i}}_{k_2(\mathbf{x})}$$

The function k_1 depends on $\mathbf{x}$ only through $t(\mathbf{x}) = \min\{x_i\}$ and the parameter θ. The function k_2 depends on $\mathbf{x}$ but not on θ.

By the Neyman Factorization Theorem, the statistic $T = \min\{X_i\}$ is sufficient for θ.