Megoldások: MM- és ML-becslések

Elek Péter

1. feladat. *(Megoldás)*

a. The MM- and ML-estimates are $\widehat{\theta}_{MM} = 0.968$ and $\widehat{\theta}_{ML} = 0.87$ (see below for details).

b. Since $E(X) = \theta/2$, the MM-estimator is $\widehat{\theta}_{MM} = 2\overline{X}$. The likelihood function is $\frac{1}{\theta^n}$ if $0 \leq x_1, x_2, \ldots, x_n \leq \theta$, and 0 otherwise. Therefore the ML-estimator is $\widehat{\theta}_{ML} = \max(X_1, X_2, \ldots, X_n) = M_n$.

c. The MM-estimator is unbiased because it is a linear function of the first moment. The ML-estimator is not unbiased because $M_n < \theta$.

d. Since $\overline{X} \xrightarrow{p} \theta/2$ by the law of large numbers, $2\overline{X} \xrightarrow{p} \theta$ and thus the MM-estimator is consistent for θ.

   The consistency of the ML-estimator follows from a previous problem set, where it was even shown that $n(\theta - M_n)$ converges to an exponential distribution. Or showing the consistency directly: $\Pr(M_n < z) = (z/\theta)^n$ for $0 < z < \theta$ and thus for all $\varepsilon > 0$:

$$\lim_{n\to\infty} \Pr(|M_n - \theta| > \varepsilon) = \lim_{n\to\infty} \Pr(M_n < \theta - \varepsilon) = \lim_{n\to\infty} (1 - \varepsilon/\theta)^n = 0.$$

   The asymptotic distribution is as given above.

2. feladat. *(Megoldás)*

a. The likelihood function is

$$\mathcal{L}(x_1, \ldots, x_n, y_1, \ldots, y_m, \mu) =$$
$$\prod_{i=1}^{n} \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right) \times \prod_{j=1}^{m} \frac{1}{(2\pi k^2\sigma^2)^{1/2}} \exp\left(\frac{-(y_j - \mu/2)^2}{2k^2\sigma^2}\right).$$

   Since σ is assumed to be known, the log-likelihood is given by

$$\log \mathcal{L}(x_1, \ldots, x_n, y_1, \ldots, y_m, \mu) = C - \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma^2} - \sum_{j=1}^{m} \frac{(y_j - \mu/2)^2}{2k^2\sigma^2}.$$

The first-order condition is

$$\frac{\partial \log \mathcal{L}}{\partial \mu} = \sum_{i=1}^{n} \frac{x_i - \mu}{\sigma^2} + \sum_{j=1}^{m} \frac{(y_j - \mu/2)}{2k^2\sigma^2} = 0,$$

hence the ML-estimate is (after checking that it is indeed a maximum)

$$\widehat{\mu}_{ML} = \frac{n\overline{X} + \frac{m}{4k^2} 2\overline{Y}}{n + \frac{m}{4k^2}}.$$

$E(\widehat{\mu}_{ML}) = \left(n\mu + \frac{m}{4k^2}\mu\right) / \left(n + \frac{m}{4k^2}\right) = \mu$, hence it is unbiased. Its sampling variance is

$$Var(\widehat{\mu}_{ML}) = \frac{n^2\sigma^2/n + \frac{m^2}{16k^4} 4k^2\sigma^2/m}{\left(n + \frac{m}{4k^2}\right)^2} = \frac{\sigma^2}{n + \frac{m}{4k^2}}.$$

b. The (ordinary) LS-estimate minimizes

$$W(\mu) = \sum_{i=1}^{n} (x_i - \mu)^2 + \sum_{j=1}^{m} (y_j - \mu/2)^2,$$

hence the first-order condition is

$$W'(\mu) = -2\sum_{i=1}^{n} (x_i - \mu) - \sum_{j=1}^{m} (y_j - \mu/2),$$

which yields (after checking that it is indeed a minimum)

$$\widehat{\mu}_{LS} = \frac{2n\overline{X} + m\overline{Y}}{2n + m/2}.$$

This estimator is unbiased, too. But its sampling variance is

$$Var(\widehat{\mu}_{LS}) = \sigma^2 \frac{4n + mk^2}{(2n + m/2)^2}.$$

Since $\left(n + \frac{m}{4k^2}\right)\left(4n + mk^2\right) > (2n + m/2)^2$ for $k \neq 1$, the LS-estimator is less efficient than the ML-estimator for $k \neq 1$.

Note: this exercise can be regarded as a simple example for regression with heteroscedasticity. The ML-estimator takes account of heteroscedasticity by giving more weight to the terms with smaller variance, while the LS-estimator does not.

3. feladat. *(Megoldás)*

a. The likelihood of observing x_i is $f(x_i;\theta) = \frac{1}{\theta}\exp\left(-\frac{x_i}{\theta}\right)$ (for $x_i > 0$), so the likelihood function and the log-likelihood are (for $x_1,\ldots,x_n > 0$):

$$\begin{aligned}\mathcal{L}(x_1,x_2,\ldots,x_n;\theta) &= \frac{1}{\theta^n}\exp\left(-\frac{\sum_{i=1}^n x_i}{\theta}\right)\\ \log\mathcal{L}(x_1,x_2,\ldots,x_n;\theta) &= -n\log\theta - \frac{\sum_{i=1}^n x_i}{\theta}.\end{aligned}$$

The first-order condition is

$$\frac{\partial\log\mathcal{L}}{\partial\theta} = -\frac{n}{\theta} + \frac{\sum_{i=1}^n x_i}{\theta^2} = 0,$$

which gives $\widehat{\theta}_{ML} = \frac{\sum_{i=1}^n X_i}{n} = \overline{X}$. (It can be shown easily that this is indeed a maximum.)

b. We know that $E(X_i) = \theta$ for the exponential distributon, hence $\widehat{\theta}_{ML} = \overline{X}$ is unbiased for θ. The second derivative of the log-likelihood is

$$\frac{\partial^2\log\mathcal{L}}{\partial\theta^2} = \frac{n}{\theta^2} - 2\frac{\sum_{i=1}^n x_i}{\theta^3}.$$

Since the sample elements are random variables, this has expected value

$$E\left(\frac{\partial^2\log\mathcal{L}}{\partial\theta^2}\right) = \frac{n}{\theta^2} - 2\frac{\sum_{i=1}^n E(X_i)}{\theta^3} = \frac{n}{\theta^2} - 2\frac{n\theta}{\theta^3} = -\frac{n}{\theta^2}.$$

Then the CRLB is given by $-\left[E\frac{\partial^2\log\mathcal{L}}{\partial\theta^2}\right]^{-1} = \frac{\theta^2}{n}$. But this is equal to $Var\left(\overline{X}\right) = \frac{Var(X)}{n} = \frac{\theta^2}{n}$. Therefore $\widehat{\theta}_{ML} = \overline{X}$ is the best unbiased estimator of θ.

4. feladat. *(Megoldás)*

In this case we have $E(X) = \lambda$, so $\widehat{\lambda}_{MM} = \overline{X}$.
$\Pr(X = x_i) = \frac{\lambda^{x_i}}{x_i!}e^{-\lambda}$ for $x_i = 0, 1, \ldots$, so the likelihood function is

$$\mathcal{L}(x_1,x_2,\ldots,x_n;\lambda) = \prod_{i=1}^n \Pr(X = x_i) = \frac{\lambda^{\sum x_i}}{\prod x_i!}e^{-n\lambda}.$$

Again, it is more convenient to maximize the log-likelihood:

$$\log\mathcal{L}(x_1,x_2,\ldots,x_n;\lambda) = \left(\sum_{i=1}^n x_i\right)\log\lambda - \sum_{i=1}^n \log(x_i!) - n\lambda.$$

Derivative with respect to λ is

$$\frac{\sum_{i=1}^n x_i}{\lambda} - n = 0,$$

which gives $\widehat{\lambda}_{ML} = \frac{\sum_{i=1}^{n} X_i}{n} = \overline{X}$. (It can be checked that it is indeed a maximum.)

The second derivative of the log-likelihood is $\frac{\partial^2 \log \mathcal{L}}{\partial \lambda^2} = -\frac{\sum_{i=1}^{n} x_i}{\lambda^2}$, the expected value of this is $-\frac{n\lambda}{\lambda^2} = -\frac{n}{\lambda}$, so the CRLB is $-\left[E\frac{\partial^2 \log \mathcal{L}}{\partial \lambda^2}\right]^{-1} = \frac{\lambda}{n}$, which is equal to $Var\left(\overline{X}\right) = \frac{Var(X)}{n} = \frac{\lambda}{n}$. Therefore $\overline{X}$ is the best unbiased estimator of λ.