UPPSALA UNIVERSITY
Dept of Mathematics
Rolf Larsson

Matematical statistics
Inference II 1MS037
2023–11–28

Hand-in assignment 2, solutions

1. Suppose we have a random sample $(x_1, x_2, x_3) = (1, 11, 12)$ from a distribution with expectation θ. Estimate θ by using

   (a) the least squares estimate. (0.5p)

   Solution: The least squares estimate is the mean (example 4.11), i.e. $(1 + 11 + 12)/3 = 8$.

   (b) the least absolute value. (0.5p)

   Solution: The least absolute value estimate is the median (example 4.12), i.e. 11.

   (c) the trimmed mean with $k = 3$ (see example 4.13). (2p)

   Solution: Let

   $$g(\theta) = \sum_{|x_i-\theta|\leq 3} (x_i - \theta), \quad h(\theta) = \sum_{|x_i-\theta|\leq 3} (x_i - \theta)^2 + 3^2 \sharp\{|x_i - \theta| > 3\},$$

   where $\sharp\{|x_i - \theta| > 3\}$ is the number of x_i such that $|x_i - \theta| > 3$.
   Among the θ values such that $g(\theta) = 0$, we want to find the one that minimizes $h(\theta)$.
   For all $\theta > 15$ or $\theta < -2$, we have $g(\theta) = 0$, so all of these are potential solutions to our problem. Here, $h(\theta) = 3^2 \cdot 3 = 27$.
   For $15 \geq \theta > 14$, we have that $0 = g(\theta) = 12 - \theta$ yields the solution $\theta = 12$. This contradicts $15 \geq \theta > 14$.
   For $14 \geq \theta \geq 9$, $0 = g(\theta) = (12 - \theta) + (11 - \theta) = 23 - 2\theta$ gives $\theta = 11.5$. This corresponds to $h(\theta) = (12 - 11.5)^2 + (11 - 11.5)^2 + 3^2 \cdot 1 = 9.5$.
   For $9 > \theta \geq 8$, we get $0 = g(\theta) = 11 - \theta$, i.e. $\theta = 11$, but this contradicts $9 > \theta \geq 8$.
   For $8 > \theta > 4$, we have $g(\theta) = 0$, so all of these are potential solutions, and here $h(\theta) = 3^2 \cdot 3 = 27$.
   Finally, for $4 \geq \theta \geq -2$, $0 = g(\theta) = 1 - \theta$ gives $\theta = 1$, corresponding to $h(\theta) = (1 - 1)^2 + 3^2 \cdot 2 = 18$.
   To sum up, among the potential θ that solve $g(\theta) = 0$, it is $\theta = 11.5$ that gives the smallest value of $h(\theta)$. Hence, the trimmed mean estimate is 11.5.

(d) the Winzorized mean with $k = 3$ (see example 4.14). (2p)

Solution: Let

$$g(\theta) = \sum_{|x_i - \theta| \leq 3} (x_i - \theta) - 3\sharp\{x_i - \theta < -3\} + 3\sharp\{x_i - \theta > 3\}.$$

(It turns out that we need not specify an h function as in (b).)
For all $\theta > 15$ we have $g(\theta) = 0 - 3 \cdot 3 = -9$, and for all $\theta < -2$ we have $g(\theta) = 0 + 3 \cdot 3 = 9$, hence no solutions here.
For $15 \geq \theta > 14$, we have that $0 = g(\theta) = 12 - \theta - 3 \cdot 2 = 6 - \theta$, which yields the solution $\theta = 6$, a contradiction.
For $14 \geq \theta \geq 9$, $0 = g(\theta) = (12 - \theta) + (11 - \theta) - 3 \cdot 1 = 20 - 2\theta$ gives $\theta = 10$, which is a permitted solution.
For $9 > \theta \geq 8$, we get $0 = g(\theta) = 11 - \theta - 3 \cdot 1 + 3 \cdot 1 = 11 - \theta$, i.e. $\theta = 11$, a contradiction.
For $8 > \theta > 4$, we have $g(\theta) = 0 - 3 \cdot 1 + 3 \cdot 2 = 3$, hence no solutions.
Finally, for $4 \geq \theta \geq -2$, $g(\theta) = 1 - \theta + 3 \cdot 2 = 7 - \theta$, giving $\theta = 7$, a contradiction.
Hence, the only solution to $g(\theta) = 0$ is $\theta = 10$, so this is the Winzorized mean estimate. Observe that this estimate lies between the median and the mean.

2. Suppose we have one observation of the continuous random variable X, with density function

$$f(x) = \beta^{-2} x \exp\left(-\frac{x}{\beta}\right),$$

for $x \geq 0$ and 0 otherwise, with $\beta > 0$. Consider the estimator $T(X) = X^2/6$ of the parameter $\theta = \beta^2$.

Hint: Without proof, you may use that $E(X^k) = (k+1)!\beta^k$ for $k = 1, 2, ...$

(a) Show that $T(X)$ is unbiased for θ. (1p)

Solution: From the hint, we have

$$E\{T(X)\} = \frac{1}{6}E(X^2) = \frac{1}{6} \cdot 3!\beta^2 = \beta^2 = \theta,$$

showing unbiasedness.

(b) Is $T(X)$ efficient for θ? Motivate your answer. (4p)

Solution: Efficiency means that the variance of $T(X)$ attains the Cramér-Rao lower bound $1/I_X(\theta)$, where $I_X(\theta)$ is the Fisher information.
As in (a), we have

$$E\{T(X)^2\} = \left(\frac{1}{6}\right)^2 E(X^4) = \frac{1}{6^2} \cdot 5!\beta^4 = \frac{10}{3}\theta^2,$$

which yields the variance

$$\mathrm{Var}\{T(X)\} = E\{T(X)^2\} - [E\{T(X)\}]^2 = \frac{10}{3}\theta^2 - \theta^2 = \frac{7}{3}\theta^2.$$

Inserting $\beta = \theta^{1/2}$, the likelihood is

$$L(\theta) = \theta^{-1} x \exp\left(-x\theta^{-1/2}\right),$$

which yields the log likelihood and its first two derivatives as

$$\begin{aligned} l(\theta) &= \log x - \log\theta - x\theta^{-1/2}, \\ l'(\theta) &= -\frac{1}{\theta} + \frac{1}{2}x\theta^{-3/2}, \\ l''(\theta) &= \frac{1}{\theta^2} - \frac{3}{4}x\theta^{-5/2}, \end{aligned}$$

and it seems to be easiest to calculate the Fisher information as (note that $E(X) = 2!\beta = 2\theta^{1/2}$)

$$\begin{aligned} I_X(\theta) = -E\{l''(\theta; X)\} &= -\frac{1}{\theta^2} + \frac{3}{4}E(X)\theta^{-5/2} \\ &= -\frac{1}{\theta^2} + \frac{3}{4}2\theta^{1/2} \cdot \theta^{-5/2} = \frac{1}{2}\theta^{-2}, \end{aligned}$$

giving the Cramér-Rao bound $1/I_X(\theta) = 2\theta^2$.
Hence, $\mathrm{Var}\{T(X)\}$ does not attain the Cramér-Rao bound. The estimator is not efficient.

3. Suppose that $X_1, ..., X_n$ are independent Bernoulli variables with parameter p, i.e. $P(X_i = 1) = p = 1 - P(X_i = 0)$ for $i = 1, ..., n$. Suppose we want to estimate

$$\gamma(p) = P(\cap_{i=1}^{n-1}\{X_i > X_n\} = 1).$$

Hence, $\gamma(p)$ is the probability that all $X_1, ..., X_{n-1}$ are strictly greater than X_n.

(a) Show that $U = I\{\cap_{i=1}^{n-1}\{X_i > X_n\}\}$, where $I\{A\} = 1$ if A is true and 0 otherwise, is an unbiased estimator of $\gamma(p)$. (1p)

Solution: Because U is an indicator variable, it follows easily that

$$E(U) = P(\cap_{i=1}^{n-1} I\{X_i > X_n\} = 1) = \gamma(p),$$

showing unbiasedness.

(b) Use the Rao-Blackwell theorem to construct an unbiased estimator of $\gamma(p)$ with smaller variance than U. (3p)

Solution: At first, note that in fact

$$\gamma(p) = P(X_1 = 1, ..., X_{n-1} = 1, X_n = 0) = p^{n-1}(1-p).$$

Suppose that we have a sample $x_1, ..., x_n$. Let $t = \sum_{i=1}^n x_i$. Since t is the number of ones in the sample, we get the likelihood

$$L(p) = p^t(1-p)^{n-t} = (1-p)^n \left(\frac{p}{1-p}\right)^t,$$

from which we see by the factorization theorem that $T = \sum_{i=1}^n X_i$ is sufficient for p.

Now,

$$\begin{aligned}
E(U|T=t) &= E\left(I\{\cap_{i=1}^{n-1}\{X_i > X_n\}\}|T=t\right) \\
&= P\left(\cap_{i=1}^{n-1} I\{X_i > X_n\} = 1|T=t\right) \\
&= P(X_1 = 1, ..., X_{n-1} = 1, X_n = 0|T=t) \\
&= \frac{P(X_1 = 1, ..., X_{n-1} = 1, X_n = 0, \sum_{i=1}^n X_i = t)}{P\left(\sum_{i=1}^n X_i = t\right)} \\
&= \begin{cases} \frac{p^{n-1}(1-p)}{\binom{n}{n-1}p^{n-1}(1-p)} = \frac{1}{n} & \text{if } t = n-1, \\ 0 & \text{otherwise} \end{cases} \\
&= \frac{1}{n} I\{t = n-1\}.
\end{aligned}$$

Hence, by the Rao-Blackwell theorem, an unbiased estimator of $\gamma(p)$ with smaller variance than U is

$$Y = \frac{1}{n} I\{T = n-1\}.$$

(c) Give the variances of U and of the estimator i (b) explicitly, to verify that the latter estimator has the lowest variance among the two. (1p)

Solution: If Z is an indicator variable, $E(Z^2) = E(Z)$ which gives $\mathrm{Var}(Z) = E(Z) - E(Z)^2 = E(Z)\{1 - E(Z)\}$. Utilizing this fact, we find

$$\mathrm{Var}(U) = \gamma(p)\{1 - \gamma(p)\} = p^{n-1}(1-p)\{1 - p^{n-1}(1-p)\}.$$

Moreover,

$$\begin{aligned} E(Y^2) &= \left(\frac{1}{n}\right)^2 E[I\{T = n-1\}] = \frac{1}{n^2} P(T = n-1) \\ &= \frac{1}{n^2}\binom{n}{n-1} p^{n-1}(1-p) = \frac{1}{n} p^{n-1}(1-p), \end{aligned}$$

and since Y is unbiased by the Rao-Blackwell theorem (and this is also easy to check), we get

$$\begin{aligned} \mathrm{Var}(Y) = E(Y^2) - \{E(Y)\}^2 &= \frac{1}{n} p^{n-1}(1-p) - \{p^{n-1}(1-p)\}^2 \\ &= p^{n-1}(1-p)\left\{\frac{1}{n} - p^{n-1}(1-p)\right\}, \end{aligned}$$

which is seen to be smaller than the variance of U, as was to be verified.

4. Consider a random sample $\mathbf{X} = (X_1, ..., X_n)$ where the X_i are Exponentially distributed with intensity β, i.e. with density function

$$f(x) = \beta \exp(-\beta x), \quad x > 0,$$

and 0 otherwise, with $\beta > 0$. The goal is to estimate $\mu = E(X_i) = 1/\beta$.

(a) Let $T = \sum_{i=1}^n X_i$. Show that T is complete and sufficient. (3p)

Solution: With observations $x_1, ..., x_n$, the likelihood is

$$L(\mu) = \prod_{i=1}^n \frac{1}{\mu} \exp\left(-\frac{1}{\mu} x_i\right) = \mu^{-n} \exp\left(-\frac{1}{\mu} \sum_{i=1}^n x_i\right)$$

We note that this is a one parameter exponential family with sufficient statistic $T = \sum_{i=1}^n X_i$.
Because $\beta > 0$, we have for the natural parameter $1/\mu \in (0, \infty)$.
Thus, the natural parameter space contains a nonempty interval. Hence, by theorem 4.6, T is complete and sufficient.
Alternatively, use corollary 4.1.

(b) Make use of the fact that $U = X_1$ is unbiased for μ to construct the best unbiased estimator (BUE) of μ. (2p)

Solution: Since T is complete and sufficient, then according to the Lehmann-Sheffé theorem, the BUE is constructed by Rao-Blackwellization. Now, as in remark 4.5,

$$E(U|T = t) = E\left(X_1 \,\middle|\, \sum_{i=1}^n X_i = t\right) = \frac{t}{n},$$

and so, the BUE is given by $T/n = \sum_{i=1}^n X_i/n$.