THE UNIVERSITY OF SYDNEY
SCHOOL OF MATHEMATICS AND STATISTICS

Solutions to Tutorial Week 11

STAT3023: Statistical Inference — Semester 2, 2023

1. Suppose $X_1, \ldots, X_n$ are iid $N(\theta, 1)$ random variables and $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$. If $\theta_0 < \theta < \theta_1$ and $0 < C < \infty$ show that

$$P_\theta\left\{\theta_0 + \frac{C}{\sqrt{n}} < \bar{X} < \theta_1 - \frac{C}{\sqrt{n}}\right\} \to 1$$

as $n \to \infty$.

Solution: Since $\bar{X} \sim N(\theta, \frac{1}{n})$, $\sqrt{n}(\bar{X} - \theta) \sim N(0, 1)$. Therefore, so long as n is big enough that

$$\theta_0 + \frac{C}{\sqrt{n}} \le \theta_1 - \frac{C}{\sqrt{n}}$$

(i.e. $n \ge \left\{\frac{2C}{\theta_1 - \theta_0}\right\}^2$), we have

$$\begin{aligned}
P_\theta\left\{\theta_0 + \frac{C}{\sqrt{n}} < \bar{X} < \theta_1 - \frac{C}{\sqrt{n}}\right\} &= P_\theta\left\{\sqrt{n}\left(\theta_0 + \frac{C}{\sqrt{n}} - \theta\right) < \sqrt{n}\left(\bar{X} - \theta\right) < \sqrt{n}\left(\theta_1 - \frac{C}{\sqrt{n}} - \theta\right)\right\} \\
&= P_\theta\left\{C - \sqrt{n}\,(\theta - \theta_0) < \sqrt{n}\left(\bar{X} - \theta\right) < \sqrt{n}\,(\theta_1 - \theta) - C\right\} \\
&= \Phi\left(\sqrt{n}\,(\theta_1 - \theta) - C\right) - \Phi\left(C - \sqrt{n}\,(\theta - \theta_0)\right) .
\end{aligned}$$

Since

$$\begin{aligned}
\sqrt{n}\,(\theta_1 - \theta) - C &\to +\infty \\
\Phi\left(\sqrt{n}\,(\theta_1 - \theta) - C\right) &\to 1\,;
\end{aligned}$$

also, since

$$\begin{aligned}
C - \sqrt{n}\,(\theta - \theta_0) &\to -\infty\,, \\
\Phi\left(C - \sqrt{n}\,(\theta - \theta_0)\right) &\to 0\,.
\end{aligned}$$

Therefore the difference tends to 1, as required.

2. Interval estimation of an exponential rate parameter

Suppose $\mathbf{X} = (X_1, \ldots, X_n)$ consists of iid random variables whose common distribution is exponential with *rate* $\theta \in \Theta = (0, \infty)$ unknown. Consider the formal decision problem with decision space $\mathcal{D} = \Theta = (0, \infty)$ and loss function (sequence) $L(d|\theta) = L_n(d|\theta) = 1\left\{|d - \theta| > \frac{C}{\sqrt{n}}\right\}$. This corresponds to obtaining an interval estimate of θ, with risk equal to the non-coverage probability of the interval; the *midpoint* of the interval is still regarded as an "estimator" of θ though.

Consider using the ordinary maximum likelihood estimator $\hat{\theta}_{\mathrm{ML}}$ as the estimator, giving the interval $\hat{\theta}_{\mathrm{ML}} \pm \frac{C}{\sqrt{n}}$.

(a) Write down the likelihood and derive $\hat{\theta}_{\mathrm{ML}}$ as a function of the X_i's.

Solution: The likelihood is

$$f_\theta(\mathbf{X}) = \prod_{i=1}^n \left[\theta e^{-\theta X_i}\right] = \theta^n e^{-T\theta}$$

where $T = \sum_{i=1}^n X_i$; taking logs and differentiating gives

$$\ell'(\theta; \mathbf{X}) = \frac{n}{\theta} - T\,;$$

Copyright © 2023 The University of Sydney

setting equal to zero and solving gives

$$\hat{\theta}_{\mathrm{ML}} = \frac{n}{T} = \frac{1}{\bar{X}},$$

where, as usual, $\bar{X} = T/n$ denotes the sample mean.

(b) Since $\bar{X}$ has a gamma distribution with shape n and rate $n\theta$, the product $Y_n = \theta\bar{X}$ has a gamma distribution with shape n and rate n (i.e. its distribution is free of θ). Show that the risk function

$$R(\theta|\hat{\theta}_{\mathrm{ML}}) = G_n\left(\left(1+\frac{C}{\theta\sqrt{n}}\right)^{-1}\right) + \left[1 - G_n\left(\left(1-\frac{C}{\theta\sqrt{n}}\right)^{-1}\right)\right]$$

where

$$G_n(y) = P(Y_n \le y)$$

is the CDF of Y_n.

Solution: The risk function is

$$\begin{aligned}
R(\theta|\hat{\theta}_{\mathrm{ML}}) &= P_\theta\left\{\theta < \hat{\theta}_{\mathrm{ML}} - \frac{C}{\sqrt{n}}\right\} + P_\theta\left\{\theta > \hat{\theta}_{\mathrm{ML}} + \frac{C}{\sqrt{n}}\right\} \\
&= P_\theta\left\{\bar{X} < \frac{1}{\theta + \frac{C}{\sqrt{n}}}\right\} + P_\theta\left\{\bar{X} > \frac{1}{\theta - \frac{C}{\sqrt{n}}}\right\} \\
&= P_\theta\left\{\theta\bar{X} < \frac{\theta}{\theta + \frac{C}{\sqrt{n}}}\right\} + P_\theta\left\{\theta\bar{X} > \frac{\theta}{\theta - \frac{C}{\sqrt{n}}}\right\} \\
&= P\left\{Y_n < \frac{1}{1 + \frac{C}{\theta\sqrt{n}}}\right\} + P\left\{Y_n > \frac{1}{1 - \frac{C}{\theta\sqrt{n}}}\right\} \\
&= G_n\left(\left(1+\frac{C}{\theta\sqrt{n}}\right)^{-1}\right) + \left[1 - G_n\left(\left(1-\frac{C}{\theta\sqrt{n}}\right)^{-1}\right)\right].
\end{aligned}$$

(c) Determine, for any $0 < a < b < \infty$, the maximum risk

$$\max_{a\le\theta\le b} R(\theta|\hat{\theta}_{\mathrm{ML}}).$$

Solution: Each of the two terms making up the risk is an *increasing* function of θ and so the sum of them is also an increasing function of θ. Thus

$$\max_{a\le\theta\le b} R(\theta|\hat{\theta}_{\mathrm{ML}}) = R(b|\hat{\theta}_{\mathrm{ML}}) = G_n\left(\left(1+\frac{C}{b\sqrt{n}}\right)^{-1}\right) + 1 - G_n\left(\left(1-\frac{C}{b\sqrt{n}}\right)^{-1}\right).$$

(d) Determine, for any $0 < a < b < \infty$, the *limiting* maximum risk

$$\lim_{n\to\infty}\max_{a\le\theta\le b} R(\theta|\hat{\theta}_{\mathrm{ML}}).$$

You may use the facts that $E(Y_n) = 1$, $\mathrm{Var}(Y_n) = \frac{1}{n}$, and Y_n is asymptotically normal.

Solution: Note firstly that for any sequence $z_n \to z$, since $\sqrt{n}(Y_n - 1) \xrightarrow{d} N(0,1)$ we have

$$P\left\{\sqrt{n}(Y_n - 1) \le z_n\right\} \to \Phi(z)$$

where $\Phi(\cdot)$ is the $N(0,1)$ CDF. Consider the first term making up the maximum risk:

$$\begin{aligned}
G_n\left(\left(1+\frac{C}{b\sqrt{n}}\right)^{-1}\right) &= P\left\{Y_n \le \frac{1}{1 + \frac{C}{b\sqrt{n}}}\right\} \\
&= P\left\{\sqrt{n}(Y_n - 1) \le \sqrt{n}\left(\frac{1}{1 + \frac{C}{b\sqrt{n}}} - 1\right)\right\} \\
&= P\left\{\sqrt{n}(Y_n - 1) \le \left(\frac{-\frac{C}{b}}{1 + \frac{C}{b\sqrt{n}}}\right)\right\}.
\end{aligned}$$

Since

$$\left(\frac{-\frac{C}{b}}{1+\frac{C}{b\sqrt{n}}}\right) \to -\frac{C}{b},$$

this tends to

$$\Phi\left(-\frac{C}{b}\right) = 1 - \Phi\left(\frac{C}{b}\right).$$

In a similar way,

$$\begin{aligned} 1 - G_n\left(\left(1-\frac{C}{b\sqrt{n}}\right)^{-1}\right) &= P\left\{\sqrt{n}(Y_n - 1) > \sqrt{n}\left(\frac{1}{1-\frac{C}{b\sqrt{n}}} - 1\right)\right\} \\ &= P\left\{\sqrt{n}(Y_n - 1) > \left(\frac{\frac{C}{b}}{1-\frac{C}{b\sqrt{n}}}\right)\right\} \\ &\to 1 - \Phi\left(\frac{C}{b}\right). \end{aligned}$$

Therefore the limiting maximum risk is

$$\lim_{n\to\infty} \max_{a\le\theta\le b} R(\theta|\hat{\theta}_{\mathrm{ML}}) = 2\left[1 - \Phi\left(\frac{C}{b}\right)\right].$$

3. Suppose $X_1, \ldots, X_n$ are iid $U[0,\theta]$ random variables and that we wish to estimate θ using squared error loss (so the risk is the mean-squared error). Assume that $n \ge 3$.

(a) The maximum likelihood estimator of θ is the *sample maximum* $X_{(n)} = \max_{i=1,\ldots,n} X_i$. In last week's tutorial, using the fact that $E_\theta(X_{(n)}) = \frac{n\theta}{n+1}$ and $\mathrm{Var}_\theta\left(X_{(n)}\right) = \frac{n\theta^2}{(n+1)^2(n+2)}$, we showed that the exact risk of this estimator is

$$E_\theta\left\{\left[X_{(n)} - \theta\right]^2\right\} = \frac{2\theta^2}{(n+2)(n+1)}.$$

Determine, for $0 \le a < b < \infty$, the limiting maximum (rescaled) risk over $[a, b]$:

$$\lim_{n\to\infty} \max_{a\le\theta\le b} n^2 E_\theta\left\{\left[X_{(n)} - \theta\right]^2\right\}.$$

Solution:

$$\begin{aligned} \lim_{n\to\infty} \max_{a\le\theta\le b} n^2 E_\theta\left\{\left[X_{(n)} - \theta\right]^2\right\} &= \lim_{n\to\infty} \max_{a\le\theta\le b} \theta^2 \left\{\frac{2n^2}{(n+2)(n+1)}\right\} \\ &= \lim_{n\to\infty} b^2 \left\{\frac{2n^2}{n^2\left(1+\frac{2}{n}\right)\left(1+\frac{1}{n}\right)}\right\} \\ &= \lim_{n\to\infty} b^2 \left\{\frac{2}{\left(1+\frac{2}{n}\right)\left(1+\frac{1}{n}\right)}\right\} \\ &= 2b^2. \end{aligned}$$

(b) In last week's tutorial we also showed that the unbiased estimator $\hat{\theta}_{\mathrm{unb}} = \left(\frac{n+1}{n}\right) X_{(n)}$ has exact risk

$$E_\theta\left\{\left[\hat{\theta}_{\mathrm{unb}} - \theta\right]^2\right\} = \frac{\theta^2}{n(n+2)}.$$

For $0 \le a < b < \infty$ find

$$\lim_{n\to\infty} \max_{a\le\theta\le b} n^2 E_\theta\left\{\left[\hat{\theta}_{\mathrm{unb}} - \theta\right]^2\right\}.$$

Solution:

$$\begin{aligned}
\lim_{n\to\infty} \max_{a\le\theta\le b} n^2 E_\theta \left\{\left[\hat{\theta}_{\text{unb}} - \theta\right]^2\right\} &= \lim_{n\to\infty} \max_{a\le\theta\le b} \theta^2 \left\{\frac{n^2}{n(n+2)}\right\} \\
&= \lim_{n\to\infty} b^2 \left\{\frac{n^2}{n^2\left(1+\frac{2}{n}\right)}\right\} \\
&= \lim_{n\to\infty} b^2 \left\{\frac{1}{\left(1+\frac{2}{n}\right)}\right\} \\
&= b^2 \,.
\end{aligned}$$

(c) Show that the Bayes procedure using the "flat prior" weight function $w(\theta) \equiv 1$ is given by

$$\hat{\theta}_{\text{flat}}(\mathbf{X}) = \left(\frac{n-1}{n-2}\right) X_{(n)} \,.$$

Solution: The likelihood is

$$f_\theta(\boldsymbol{X}) = \prod_{i=1}^{n} \left[\frac{1\{0 \le X_i \le \theta\}}{\theta}\right] = \frac{1\left\{X_{(1)} \ge 0\right\} 1\left\{X_{(n)} \le \theta\right\}}{\theta^n}$$

(where $X_{(1)}$ is the sample *minimum*). Assuming $X_{(1)} \ge 0$, when viewed this is a multiple of a Pareto density (with shape $n-1$ and scale $X_{(n)}$). (Recall that if X has a Pareto density with shape α and scale m, then X has pdf $f(x) = \frac{\alpha m^\alpha}{x^{\alpha+1}}$ for $x \ge m$.) Since the weight function is just $w(\theta) \equiv 1$ this Pareto distribution is also the posterior distribution. Thus the Bayes procedure/estimator is the posterior mean, which is

$$\frac{\text{shape} \times \text{scale}}{\text{shape} - 1} = \frac{(n-1)X_{(n)}}{(n-2)} \,.$$

(d) Using the expressions for the expectation and variance of $X_{(n)}$ given in part (a) above, determine the variance, bias and thus exact risk of $\hat{\theta}_{\text{flat}}$.

Solution:

$$E_\theta\left[\hat{\theta}_{\text{flat}}(\mathbf{X})\right] = \left(\frac{n-1}{n-2}\right) E_\theta\left(X_{(n)}\right) = \left(\frac{n-1}{n-2}\right) \frac{n\theta}{n+1} \,.$$

Thus the bias is

$$\begin{aligned}
\text{Bias}_\theta\left[\hat{\theta}_{\text{flat}}\right] = \left(\frac{n-1}{n-2}\right) E_\theta\left(X_{(n)}\right) - \theta &= \frac{(n-1)n\theta - \theta(n-2)(n+1)}{(n-2)(n+1)} \\
&= \theta\left(\frac{n^2 - n - [n^2 - n - 2]}{(n-2)(n+1)}\right) \\
&= \frac{2\theta}{(n-2)(n+1)} \,.
\end{aligned}$$

The variance is given by

$$\begin{aligned}
\text{Var}_\theta\left(\hat{\theta}_{\text{flat}}\right) &= \left(\frac{n-1}{n-2}\right)^2 \text{Var}_\theta\left(X_{(n)}\right) \\
&= \left(\frac{n-1}{n-2}\right)^2 \frac{n\theta^2}{(n+1)^2(n+2)} .
\end{aligned}$$

Thus the exact risk is

$$\begin{aligned}
E_\theta\left\{\left[\hat{\theta}_{\text{flat}} - \theta\right]^2\right\} &= \text{Var}_\theta\left(\hat{\theta}_{\text{flat}}\right) + \left\{\text{Bias}_\theta\left[\hat{\theta}_{\text{flat}}\right]\right\}^2 \\
&= \frac{\theta^2}{(n-2)^2(n+1)^2}\left[\frac{n(n-1)^2}{n+2} + 4\right] \\
&= \frac{\theta^2(n-1)^2}{(n-2)^2(n+1)^2}\left[\frac{n}{n+2} + \frac{4}{(n-1)^2}\right] .
\end{aligned}$$

(e) Determine, for $0 \le a < b < \infty$,

$$\lim_{n\to\infty} \max_{a\le\theta\le b} n^2 E_\theta \left\{ \left[\hat{\theta}_{\text{flat}} - \theta\right]^2 \right\} .$$

Solution:

$$\begin{aligned}
\lim_{n\to\infty} \max_{a\le\theta\le b} n^2 E_\theta \left\{ \left[\hat{\theta}_{\text{flat}} - \theta\right]^2 \right\} &= \lim_{n\to\infty} \max_{a\le\theta\le b} \theta^2 \left\{ \frac{n^2(n-1)^2}{(n-2)^2(n+1)^2} \left[\frac{n}{n+2} + \frac{4}{(n-1)^2} \right] \right\} \\
&= \lim_{n\to\infty} b^2 \left\{ \frac{n^2(n-1)^2}{(n-2)^2(n+1)^2} \left[\frac{n}{n+2} + \frac{4}{(n-1)^2} \right] \right\} \\
&= b^2
\end{aligned}$$

since

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

and

$$\left[\frac{n}{n+2} + \frac{4}{(n-1)^2} \right] = \frac{n}{n\left(1+\frac{2}{n}\right)} + \frac{4}{(n-1)^2} = \frac{1}{\left(1+\frac{2}{n}\right)} + \frac{4}{(n-1)^2} \to 1\,.$$

(f) Comment on what is interesting about the 3 estimators compared in the previous parts.

Solution: The maximum likelihood estimator (the sample maximum) has a bias such that the squared bias and the variance are “of the same order” (i.e. like $\frac{1}{n^2}$) and thus *both* contribute to the limiting (maximum, rescaled) risk. The unbiased version $\hat{\theta}_{\text{unb}}$ however removes the bias and only increases the variance slightly, so that in the limit, the (maximum, rescaled) risk is halved. Perhaps most interestingly, the Bayes estimator is “automatically (approximately) bias-corrected”; it is not unbiased, but it is “bias-corrected” enough so that the squared bias is of *smaller order* than the variance and thus does **not** contribute to the limiting (maximum, rescaled) risk.