THE UNIVERSITY OF SYDNEY
SCHOOL OF MATHEMATICS AND STATISTICS

# Tutorial Week 11

STAT3023: Statistical Inference — Semester 2, 2022

Lecturers: Linh Nghiem, Michael Stewart and Rachel Wang

**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 $n \to \infty$.

**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 $\theta$, with risk equal to the non-coverage probability of the interval; the *midpoint* of the interval is still regarded as an "estimator" of $\theta$ 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.

(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 $\theta$). 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 $Y_n$.

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

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

(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 $E(Y_n) = 1$, $Var(Y_n) = \frac{1}{n}$ and $Y_n$ is asymptotically normal.

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

(a) The maximum likelihood estimator of $\theta$ 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 $[a, b]$:

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

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

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

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

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

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

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

(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 $\hat{\theta}_{\text{flat}}$.

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

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

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