THE UNIVERSITY OF SYDNEY
SCHOOL OF MATHEMATICS AND STATISTICS

Solutions to Tutorial Week 10

STAT3023: Statistical Inference — Semester 2, 2022

Lecturers: Linh Nghiem, Michael Stewart and Rachel Wang

1. If Y has a gamma distribution with shape α and **rate**† λ its PDF is

$$f_Y(y) = \frac{y^{\alpha-1}e^{-\lambda y}\lambda^\alpha}{\Gamma(\alpha)},$$

for $y > 0$.

(a) Determine a formula for $E\left\{Y^{-k}\right\}$ which is valid for all (positive) integers k such that $0 < k < \alpha$.

Solution:

$$\begin{aligned}
E\left(Y^{-k}\right) &= \int_0^\infty y^{-k}\,\frac{y^{\alpha-1}e^{-\lambda y}\lambda^\alpha}{\Gamma(\alpha)}\,dy \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty y^{\alpha-k-1}e^{-\lambda y}\,dy \\
&= \frac{\lambda^k}{\Gamma(\alpha)}\int_0^\infty z^{\alpha-k-1}e^{-z}\,dz \quad \text{(changing to } z = \lambda y) \\
&= \frac{\lambda^k\Gamma(\alpha-k)}{\Gamma(\alpha)}.
\end{aligned}$$

Now, for all $\alpha > 0$, $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$. So we also have

$$\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1) = (\alpha-1)(\alpha-2)\Gamma(\alpha-2) = \cdots = (\alpha-1)(\alpha-2)\cdots(\alpha-k)\Gamma(\alpha-k)$$

so long as $\alpha - k > 0$. Substituting this into the denominator above we get

$$E\left\{Y^{-k}\right\} = \frac{\lambda^k}{(\alpha-1)(\alpha-2)\cdots(\alpha-k)}.$$

(b) The random variable $X = Y^{-1}$ is said to have an *inverse Gamma* distribution. Use the previous part to determine the mean and variance of X (for $\alpha > 2$).

Solution:

$$\begin{aligned}
E(X) = E\left(Y^{-1}\right) &= \frac{\lambda}{\alpha-1}; \\
E(X^2) = E\left(Y^{-2}\right) &= \frac{\lambda^2}{(\alpha-1)(\alpha-2)}.
\end{aligned}$$

Therefore

$$\begin{aligned}
\text{Var}(X) &= E(X^2) - [E(X)]^2 \\
&= \frac{\lambda^2}{(\alpha-1)(\alpha-2)} - \left[\frac{\lambda}{\alpha-1}\right]^2 \\
&= \frac{\lambda^2\left\{(\alpha-1)-(\alpha-2)\right\}}{(\alpha-1)^2(\alpha-2)} \\
&= \frac{\lambda^2}{(\alpha-1)^2(\alpha-2)}.
\end{aligned}$$

†Note that the gamma *rate* parameter is the reciprocal of the gamma *scale* parameter.

(c) Use the CDF method to derive the PDF of X.

Solution: Writing $F_Y(\cdot)$ and $f_Y(\cdot)$ for the CDF and PDF (respectively) of Y, the CDF of X is

$$\begin{aligned} F_X(x) = P(X \leq x) = P\left(Y^{-1} \leq x\right) = P\left(Y \geq x^{-1}\right) &= 1 - P\left(Y < x^{-1}\right) \\ &= 1 - P\left(Y \leq x^{-1}\right) \\ &= 1 - F_Y\left(x^{-1}\right) \end{aligned}$$

and thus the PDF of X is

$$f_X(x) = \frac{d}{dx} F_X(x) = -\frac{d}{dx}\left[F_Y\left(x^{-1}\right)\right] = -f_Y\left(x^{-1}\right)\frac{d}{dx}\left(x^{-1}\right) = x^{-2} f_Y\left(x^{-1}\right)$$

using the Chain Rule. That is,

$$f_X(x) = \frac{x^{-2}\left(x^{-1}\right)^{\alpha-1} e^{-\lambda\left(x^{-1}\right)}\lambda^{\alpha}}{\Gamma(\alpha)} = \frac{\lambda^{\alpha} e^{-\lambda/x}}{x^{\alpha+1}\Gamma(\alpha)},$$

for $x > 0$.

2. Suppose now $X_1, \ldots, X_n$ are iid exponential random variables with **mean** θ, so the common PDF is, for $x, \theta > 0$, given by

$$f_\theta(x) = \frac{1}{\theta} e^{-x/\theta}.$$

(a) Determine the maximum likelihood estimator $\hat{\theta}_{\text{ML}}(\mathbf{X})$.

Solution: The likelihood is

$$\prod_{i=1}^{n} f_\theta(X_i) = \prod_{i=1}^{n}\left[\frac{1}{\theta} e^{-X_i/\theta}\right] = \frac{1}{\theta^n} e^{-\left(\sum_{i=1}^{n} X_i\right)/\theta} \tag{1}$$

Taking logs, the log-likelihood is

$$-n \log\theta - \frac{1}{\theta}\sum_{i=1}^{n} X_i,$$

differentiating and setting equal to zero gives the maximum-likelihood estimator as

$$\hat{\theta}_{\text{ML}}(\mathbf{X}) = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{X}.$$

(b) Determine the Bayes estimator $\hat{\theta}_{\text{flat}}(\mathbf{X})$ under squared-error loss using the weight function $w(\theta) \equiv 1$ (the "flat prior").

Solution: Using the likelihood in (1),

$$w(\theta)\prod_{i=1}^{n} f_\theta(X_i) = \text{const.}\frac{\left(\sum_{i=1}^{n} X_i\right)^{n-1} e^{-\left(\sum_{i=1}^{n} X_i\right)/\theta}}{\theta^{(n-1)+1}\Gamma(n-1)},$$

where the constant part does not depend on θ. This is an inverse gamma density with shape $n-1$ and rate $T = \sum_{i=1}^{n} X_i$, which is the posterior density of θ.
Since we have squared-error loss, the Bayes estimator is the posterior *mean* which, according to question 1(b), equals

$$\hat{\theta}_{\text{flat}}(\mathbf{X}) = \frac{T}{n-2}.$$

(c) Determine the Bayes estimator $\hat{\theta}_{\text{conj}}(\mathbf{X})$ under squared-error loss using the conjugate prior

$$w(\theta) = \frac{\lambda_0^{\alpha_0} e^{-\lambda_0/\theta}}{\theta^{\alpha_0+1}\Gamma(\alpha_0)},$$

for $\theta > 0$.

Solution: The product of the weight function and the likelihood in (1) is of the form

$$\text{const.} \frac{e^{-(T+\lambda_0)/\theta}}{\theta^{n+\alpha_0+1}}$$

where (as usual) "const." is a factor depending on everything *except* θ. This in turn is a multiple of the inverse gamma density with shape $n + \alpha_0$ and rate $T + \lambda_0$; that density is the posterior. Again, the estimator is the posterior *mean*, which in this case is

$$\frac{T + \lambda_0}{n + \alpha_0 - 1}.$$

(d) Determine the risk $R(\theta|d)$ of the estimator

$$d(\mathbf{X}) = \frac{\ell + \sum_{i=1}^n X_i}{n + k}$$

and hence also determine the limiting (rescaled) risk $\lim_{n\to\infty} nR(\theta|d)$.

Solution: We know that, as the sum of n iid exponentials with *mean* θ, $T = \sum_{i=1}^n X_i$ has a gamma distribution with shape n and *scale* parameter θ (the reciprocal of the rate!). Thus we have $E(T) = n\theta$ and $Var(T) = n\theta^2$.

The variance of this estimator is given by

$$Var_\theta\left[d(\mathbf{X})\right] = \left(\frac{1}{n+k}\right)^2 Var_\theta\left(T\right) = \frac{n\theta^2}{(n+k)^2}.$$

The expected value of the estimator is

$$E_\theta\left[d(\mathbf{X})\right] = \frac{\ell + E(T)}{n+k} = \frac{\ell + n\theta}{n+k}.$$

Thus the bias is given by

$$Bias_\theta\left[d(\mathbf{X})\right] = E_\theta\left[d(\mathbf{X})\right] - \theta = \frac{\ell + n\theta - (n+k)\theta}{n+k} = \frac{\ell - k\theta}{n+k}.$$

Therefore the risk is

$$\begin{aligned} R(\theta|d) &= E_\theta\left\{\left[d(\mathbf{X}) - \theta\right]^2\right\} \\ &= Var_\theta\left[d(\mathbf{X})\right] + \left\{Bias_\theta\left[d(\mathbf{X})\right]\right\}^2 \\ &= \frac{n\theta^2 + (\ell - k\theta)^2}{(n+k)^2}. \end{aligned}$$

The limiting (rescaled) risk is

$$\begin{aligned} \lim_{n\to\infty} nR(\theta|d) &= \lim_{n\to\infty} n\left(\frac{n\left(\theta^2 + \frac{(\ell-k\theta)^2}{n}\right)}{n^2\left(1+\frac{k}{n}\right)^2}\right) \\ &= \lim_{n\to\infty} \frac{\left(\theta^2 + \frac{(\ell-k\theta)^2}{n}\right)}{\left(1+\frac{k}{n}\right)^2} \\ &= \theta^2, \end{aligned}$$

the same for all (fixed) ℓ and k.

(e) Determine the risk $R(\theta|d)$ and limiting (rescaled) risk $\lim_{n\to\infty} nR(\theta|d)$ where d is replaced by each of the 3 estimators in the questions (a)–(c) above.

Solution: The three estimators $\hat{\theta}_{\text{ML}}$, $\hat{\theta}_{\text{flat}}$ and $\hat{\theta}_{\text{conj}}$ are all special cases of $d(\mathbf{X})$, for certain choices of ℓ and k:

- $\hat{\theta}_{\text{ML}}$ corresponds to $\ell = 0$, $k = 0$;
- $\hat{\theta}_{\text{flat}}$ corresponds to $\ell = 0$, $k = -2$;
- $\hat{\theta}_{\text{conj}}$ corresponds to $\ell = \lambda_0$, $k = \alpha_0 - 1$.

Therefore we have

$$\begin{aligned}
R\left(\theta|\hat{\theta}_{\text{ML}}\right) &= \frac{\theta^2}{n}\,;\\
R\left(\theta|\hat{\theta}_{\text{flat}}\right) &= \frac{(n+4)\theta^2}{(n-2)^2}\,;\\
R\left(\theta|\hat{\theta}_{\text{conj}}\right) &= \frac{n\theta^2 + (\lambda_0 - (\alpha_0+1)\theta)^2}{(n+\alpha_0+1)^2}\,.
\end{aligned}$$

and

$$nR\left(\theta|\hat{\theta}_{\text{ML}}\right) = \lim_{n\to\infty} nR\left(\theta|\hat{\theta}_{\text{flat}}\right) = \lim_{n\to\infty} nR\left(\theta|\hat{\theta}_{\text{conj}}\right) = \theta^2\,.$$

3. Suppose $X_1, \ldots, X_n$ are iid $U[0,\theta]$ and it is desired to estimate θ using squared-error loss.

(a) Write down the CDF $F_\theta(x) = P_\theta\{X_1 \le x\}$.

Solution:

$$F_\theta(x) = \begin{cases} 0 & \text{for } x < 0,\\ \frac{x}{\theta} & \text{for } 0 \le x \le \theta,\\ 1 & \text{for } x > \theta. \end{cases}$$

(b) For any n iid random variables $Y_1, \ldots, Y_n$ the CDF of the maximum $Y_{(n)} = \max_{i=1,\ldots,n} Y_i$ is given by

$$P(Y_{(n)} \le y) = P(Y_1 \le y, \ldots, Y_n \le y) = P(Y_1 \le y)\cdots P(Y_n \le y) \;\; \text{(by independence)}\,.$$

Use this to derive the CDF of $X_{(n)}$ above (the $U[0,\theta]$ sample maximum).

Solution: The CDF is

$$G_n(x;\theta) = P_\theta\left\{X_{(n)} \le x\right\} = F_\theta(x)^n = \begin{cases} 0 & \text{for } x < 0,\\ \left(\frac{x}{\theta}\right)^n & \text{for } 0 \le x \le \theta,\\ 1 & \text{for } x > \theta. \end{cases}$$

(c) Derive the PDF of $X_{(n)}$ and hence a formula for $E\left\{X_{(n)}^k\right\}$, for $k = 1, 2, \ldots$.

Solution: The PDF is

$$g_n(x;\theta) = \frac{\partial}{\partial x} G_n(x;\theta) = \begin{cases} \frac{nx^{n-1}}{\theta^n} & \text{for } 0 \le x \le \theta,\\ 0 & \text{for } x < 0 \text{ or } x > \theta. \end{cases}$$

Thus

$$E_\theta\left\{X_{(n)}^k\right\} = \int_{-\infty}^{\infty} x^k g_n(x;\theta)\,dx = \frac{n}{\theta^n}\int_0^\theta x^{k+n-1}\,dx = \frac{n}{\theta^n}\left[\frac{x^{k+n}}{k+n}\right]_0^\theta = \frac{n\theta^k}{n+k}\,.$$

(d) Determine the bias, variance and thus mean-squared error (risk) of the maximum likelihood estimator $X_{(n)}$.

Solution: The expectation of the sample maximum is

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

so the bias is

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

The mean-square of the sample maximum is

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

so the variance is

$$\begin{aligned}
\text{Var}_\theta\left[X_{(n)}\right] &= E_\theta\left[X_{(n)}^2\right] - \left\{E_\theta\left[X_{(n)}\right]\right\}^2 \\
&= \frac{n\theta^2}{n+2} - \left(\frac{n\theta}{n+1}\right)^2 \\
&= \theta^2\left\{\frac{n(n+1)^2 - n^2(n+2)}{(n+2)(n+1)^2}\right\} \\
&= \theta^2\left\{\frac{n^3+2n^2+n-(n^3+2n^2)}{(n+2)(n+1)^2}\right\} \\
&= \frac{n\theta^2}{(n+2)(n+1)^2}\,.
\end{aligned}$$

Therefore the mean-squared error

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

(e) Determine the limiting (rescaled) risk $\lim_{n\to\infty} n^2 E_\theta\left\{\left[X_{(n)} - \theta\right]^2\right\}$.

Solution:

$$\begin{aligned}
n^2 E_\theta\left\{\left[X_{(n)} - \theta\right]^2\right\} &= n^2\frac{2\theta^2}{(n+2)(n+1)} \\
&= 2\theta^2\left\{\frac{n^2}{(n+2)(n+1)}\right\} \\
&= 2\theta^2\left\{\frac{n^2}{n\left(1+\frac{2}{n}\right)n\left(1+\frac{1}{n}\right)}\right\} \\
&= \frac{2\theta^2}{\left(1+\frac{2}{n}\right)\left(1+\frac{1}{n}\right)} \\
&\to 2\theta^2
\end{aligned}$$

as $n \to \infty$.

(f) Defining the unbiased estimator $\hat{\theta}_{\text{unb}}(\boldsymbol{X}) = \left(\frac{n+1}{n}\right) X_{(n)}$, determine

$$\lim_{n\to\infty} n^2 E_\theta \left\{ \left[\hat{\theta}_{\text{unb}}(\boldsymbol{X}) - \theta \right]^2 \right\} .$$

Solution: Firstly, since this estimator is unbiased, the mean-squared error is simply the variance, which is

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

Therefore

$$n^2 E_\theta \left\{ \left[\hat{\theta}_{\text{unb}}(\boldsymbol{X}) - \theta \right]^2 \right\} = \theta^2 \left(\frac{n}{n+2} \right) = \theta^2 \left(\frac{1}{1+\frac{2}{n}} \right) \to \theta^2 .$$