THE UNIVERSITY OF SYDNEY
SCHOOL OF MATHEMATICS AND STATISTICS

Solutions to Tutorial Week 8

STAT3023: Statistical Inference — Semester 2, 2023

1. Suppose X has a gamma distribution with known shape $\gamma_0 > 0$ and unknown scale parameter θ (see the computer exercise). Then X has PDF

$$f(x; \gamma_0, \theta) = \frac{x^{\gamma_0 - 1} e^{-x/\theta}}{\theta^{\gamma_0} \Gamma(\gamma_0)}$$

for $x > 0$, and 0 otherwise. Since this is an exponential family with sufficient statistic X (and X has a continuous distribution, so no randomisation required), the UMPU test of $H_0 \colon \theta = \theta_0$ against $H_1 \colon \theta \neq \theta_0$ is of the form

$$\delta(X) = \begin{cases} 1 & \text{for } X < c \text{ or } X > d \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where the constants c and d ($c < d$) are chosen so that both the following equalities hold:

$$E_{\theta_0}\left[\delta(X)\right] = \alpha\,; \tag{2}$$
$$E_{\theta_0}\left[X\delta(X)\right] = \alpha E_{\theta_0}(X)\,. \tag{3}$$

(a) Write down a formula for $E_\theta(X)$.

Solution:

$$\begin{aligned}
E_\theta(X) &= \int_{-\infty}^{\infty} x f(x; \gamma_0, \theta)\, dx \\
&= \int_0^\infty x \, \frac{x^{\gamma_0 - 1} e^{-x/\theta}}{\theta^{\gamma_0} \Gamma(\gamma_0)}\, dx \\
&= \frac{\Gamma(\gamma_0 + 1)\theta}{\Gamma(\gamma_0)} \int_0^\infty \frac{x^{(\gamma_0+1)-1} e^{-x/\theta}}{\theta^{\gamma_0+1}\Gamma(\gamma_0+1)}\, dx \\
&= \frac{\gamma_0 \Gamma(\gamma_0)\theta}{\Gamma(\gamma_0)} \\
&= \gamma_0 \theta
\end{aligned}$$

due to the well-known "factorial property" $\Gamma(\gamma_0 + 1) = \gamma_0 \Gamma(\gamma_0)$ of the gamma function and that the last integral above is of another gamma PDF with parameters $\gamma_0 + 1$ and θ, and so is 1.

(b) Show that the conditions (2) and (3) above imply

$$\int_c^d f(x; \gamma_0, \theta_0)\, dx = \int_c^d f(x; \gamma_0 + 1, \theta_0)\, dx\,. \tag{4}$$

Solution: The left-hand side of (3) is given by

$$\begin{aligned}
&\int_0^c x \, \frac{x^{\gamma_0 - 1} e^{-x/\theta_0}}{\theta_0^{\gamma_0} \Gamma(\gamma_0)}\, dx + \int_d^\infty x \, \frac{x^{\gamma_0 - 1} e^{-x/\theta_0}}{\theta_0^{\gamma_0} \Gamma(\gamma_0)}\, dx \\
&= \frac{1}{\theta_0^{\gamma_0}\Gamma(\gamma_0)} \left\{ \int_0^c x^{\gamma_0} e^{-x/\theta_0}\, dx + \int_d^\infty x^{\gamma_0} e^{-x/\theta_0}\, dx \right\} \\
&= \theta_0 \gamma_0 \left\{ \int_0^c \frac{x^{(\gamma_0+1)-1} e^{-x/\theta_0}}{\theta_0^{\gamma_0+1}\Gamma(\gamma_0+1)}\, dx + \int_d^\infty \frac{x^{(\gamma_0+1)-1} e^{-x/\theta_0}}{\theta_0^{\gamma_0+1}\Gamma(\gamma_0+1)}\, dx \right\} \\
&= E_{\theta_0}(X) \int_{-\infty}^{\infty} \delta(x) f(x; \gamma_0 + 1, \theta_0)\, dx\,,
\end{aligned}$$

where the second last line uses the factorial property of the gamma function again. Thus by (3) this last integral must equal α and so by (2) it must also be equal to

$$E_{\theta_0}\left[\delta(X)\right] = \int_{-\infty}^{\infty} \delta(x) f(x; \gamma_0, \theta_0)\, dx\,.$$

Finally note that

$$\begin{aligned}
\int_c^d f(x;\gamma_0,\theta_0)\,dx &= 1 - \int_{-\infty}^{\infty} \delta(x) f(x;\gamma_0,\theta_0)\,dx \\
&= 1 - \int_{-\infty}^{\infty} \delta(x) f(x;\gamma_0+1,\theta_0)\,dx \\
&= \int_c^d f(x;\gamma_0+1,\theta_0)\,dx\,.
\end{aligned}$$

as required.

(c) By integrating the right-hand side of (4) by parts, show that c and d in (1) satisfy

$$c^{\gamma_0} e^{-c/\theta_0} = d^{\gamma_0} e^{-d/\theta_0}\,. \tag{5}$$

Solution: The right-hand side of (4) may be written as

$$\begin{aligned}
\int_c^d f(x;\gamma_0+1,\theta_0)\,dx &= \frac{1}{\theta_0^{\gamma_0+1}\Gamma(\gamma_0+1)} \int_c^d x^{\gamma_0} e^{-x/\theta_0}\,dx \\
&= \frac{1}{\theta_0^{\gamma_0+1}\Gamma(\gamma_0+1)} \left\{ \left[x^{\gamma_0}\left(-\theta_0 e^{-x/\theta_0}\right)\right]_c^d + \int_c^d \gamma_0 x^{\gamma_0-1}\theta_0 e^{-x/\theta_0}\,dx \right\} \\
&= \frac{1}{\theta_0^{\gamma_0+1}\Gamma(\gamma_0+1)} \left\{ \left[x^{\gamma_0}\left(-\theta_0 e^{-x/\theta_0}\right)\right]_c^d \right\} + \frac{1}{\theta_0^{\gamma_0}\Gamma(\gamma_0)} \int_c^d x^{\gamma_0-1} e^{-x/\theta_0}\,dx \\
&= \frac{1}{\theta_0^{\gamma_0+1}\Gamma(\gamma_0+1)} \left\{ \left[x^{\gamma_0}\left(-\theta_0 e^{-x/\theta_0}\right)\right]_c^d \right\} + \int_c^d f(x;\gamma_0,\theta_0)\,dx \\
&= \int_c^d f(x;\gamma_0,\theta_0)\,dx - \frac{1}{\theta_0^{\gamma_0}\Gamma(\gamma_0+1)} \left[x^{\gamma_0} e^{-x/\theta_0}\right]_c^d\,.
\end{aligned}$$

Therefore by condition (4) this last term must be zero, so that

$$c^{\gamma_0} e^{-c/\theta_0} = d^{\gamma_0} e^{-d/\theta_0}\,.$$

(d) Explain why the UMPU test of $H_0\colon \theta = \theta_0$ against $H_1\colon \theta \neq \theta_0$ rejects for large values of

$$\frac{X}{\mu_0} - \log\left\{\frac{X}{\mu_0}\right\}$$

where $\mu_0 = \gamma_0\theta_0 = E_{\theta_0}(X)$, the expected value of X under H_0.

Solution: The condition (5) is equivalent to

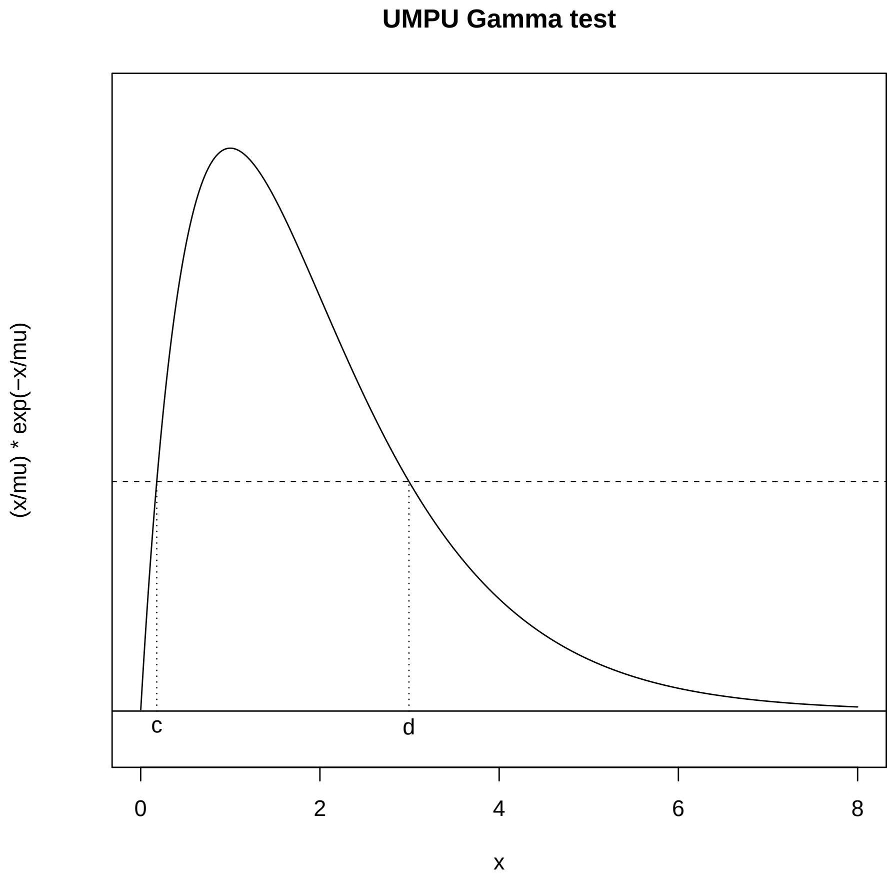
$$c e^{-c/(\theta_0\gamma_0)} = d e^{-d/(\theta_0\gamma_0)}$$

which may be rewritten as

$$\frac{c}{\mu_0} e^{-\frac{c}{\mu_0}} = \frac{d}{\mu_0} e^{-\frac{d}{\mu_0}}\,.$$

Thus the interval (c, d) may be viewed as the "level set" obtained by drawing a horizontal line that cuts the graph of the function given by

$$g(x) = \frac{x}{\mu_0} e^{-x/\mu_0}\,.$$



Thus we reject for all X such that the random variable

$$\frac{X}{\mu_0} e^{-X/\mu_0}$$

is **less than** some threshold (chosen so the test has level α). Equivalently, we reject for **large** values of

$$\frac{X}{\mu_0} - \log\left(\frac{X}{\mu_0}\right) .$$

2. Suppose $X_1, \ldots, X_n$ (for $n \geq 2$) are iid $N(\mu, \sigma^2)$ and we are interesting in testing

$$H_0\colon \sigma^2 = \sigma_0^2 \quad \text{against} \quad H_1\colon \sigma^2 \neq \sigma_0^2 \tag{6}$$

(a) The statistic $Y = \frac{1}{2}\sum_{i=1}^n (X_i - \bar{X})^2 \sim \frac{\sigma^2}{2}\chi^2_{n-1}$ which is the same as a gamma random variable with (known) shape $\frac{n-1}{2}$ and (unknown) scale parameter σ^2. It turns out that the UMPU test of (6) is the same as the test from question 1 applied to the statistic Y (which makes sense since Y is sufficient for the scale parameter σ^2). Show that this test rejects for large values of

$$\frac{S^2}{\sigma_0^2} - \log\left(\frac{S^2}{\sigma_0^2}\right) \tag{7}$$

where $S^2 = \frac{2Y}{n-1} = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$ is the sample variance.

Solution: Applying the test from question 1, we reject for large values of

$$\frac{Y}{\mu_0} - \log\left(\frac{Y}{\mu_0}\right) ,$$

where $\mu_0 = E_{H_0}(Y) = \frac{\sigma_0^2(n-1)}{2}$ is the expected value of Y under H_0. But note that

$$\frac{Y}{\mu_0} = \frac{2Y}{\sigma_0^2(n-1)} = \frac{S^2}{\sigma_0^2}$$

as required.

(b) Is the GLRT of the hypotheses (6) above equivalent to the test rejecting for large values of (7)?

Solution: The log-likelihood is

$$\ell(\mu, \sigma^2; \mathbf{X}) = -n \log \sigma - n \log \sqrt{2\pi} - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (X_i - \mu)^2 .$$

The "unrestricted" maximum likelihood estimators are $\hat{\mu} = \bar{X}$ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$. Under the restriction of $H_0 \colon \sigma^2 = \sigma_0^2$, the maximum likelihood estimator of μ remains $\bar{X}$. Thus the GLRT rejects for large values of the statistic

$$\begin{aligned}
& - n \log \hat{\sigma} - \frac{1}{2\hat{\sigma}^2} \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2 - \{-n \log \sigma_0\} + \frac{1}{2\sigma_0^2} \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2 \\
&= -n \log \frac{\hat{\sigma}}{\sigma_0} - \frac{n}{2} + \frac{n\hat{\sigma}^2}{2\sigma_0^2} \\
&= -\frac{n}{2} \log \left(\frac{\hat{\sigma}}{\sigma_0}\right)^2 - \frac{n}{2} + \frac{n\hat{\sigma}^2}{2\sigma_0^2} \\
&= \frac{n}{2} \left\{ \frac{\hat{\sigma}^2}{\sigma_0^2} - \log\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right) - 1 \right\}
\end{aligned}$$

which is an increasing function of

$$\frac{\hat{\sigma}^2}{\sigma_0^2} - \log\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right) .$$

This is *almost* the same as the UMPU statistic, but the unbiased variance estimator $S^2 = \frac{2Y}{n-1}$ has been replaced with the *biased* MLE $\hat{\sigma}^2 = \frac{2Y}{n}$.
To see that these are different tests, note that if we express the two statistics in terms of $Y = \frac{1}{2}\sum_{i=1}^n (X_i - \bar{X})^2$, the UMPU statistic is an increasing function of

$$\frac{2Y}{(n-1)\sigma_0^2} - \log Y$$

while the GLRT is an increasing function of

$$\frac{2Y}{n\sigma_0^2} - \log Y .$$

3. Suppose $X_1, \ldots, X_5$ are iid Poisson(θ) random variables. Determine the UMP test of $H_0 \colon \theta = 1$ against $H_1 \colon \theta < 1$ at level 0.05. **Hint:** you will need to calculate a few Poisson probabilities!

Solution: The likelihood is

$$\prod_{i=1}^{5} \frac{e^{-\theta}\theta^{X_i}}{X_i!} = e^{\log(\theta)\sum_{i=1}^5 X_i - 5\theta - \sum_{i=1}^5 \log(X_i!)} .$$

This is a 1-parameter **exponential** family with sufficient statistic $T = \sum_{i=1}^5 X_i$ (indeed it is monotone likelihood ratio in this sum) so the UMP test rejects for small values of the sum, indeed it is of the form

$$\delta(\mathbf{X}) = 1\{T < c\} + \gamma 1\{T = c\}$$

where c and γ are chosen so that

$$E_0[\delta(\mathbf{X})] = \alpha .$$

Under H_0 the sufficient statistic $T \sim$ Poisson(5). The first few values of its distribution are

t	0	1	2	$\cdots$
$P(T=t)$	e^{-5}	$5e^{-5}$	$\frac{25e^{-5}}{2}$	$\cdots$
$P(T\le t)$	$e^{-5}\approx 0.00674$	$6e^{-5}\approx 0.04043$	$18.5\, e^{-5}\approx 0.12465$	$\cdots$
$\delta(t)$	1	1	γ	0

Thus the test is $\delta(t) = 1\{t \le 1\} + \gamma 1\{t = 2\}$ where γ is such that

$$0.05 = E_0[\delta(T)] = e^{-5}(6 + 12.5\gamma)$$
$$\gamma = \frac{0.05e^5 - 6}{12.5} \approx 0.1137\,.$$

4. Suppose that $X_1, \ldots, X_n$ are iid $U(0, \theta)$, with common density

$$f_\theta(x) = \frac{1\{0 \le x \le \theta\}}{\theta}\,.$$

(a) Show that the sample maximum $X_{(n)}$ is a sufficient statistic for θ.

Solution: Writing $X_{(1)}$ and $X_{(n)}$ for the sample minimum and maximum (respectively), the likelihood is

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

and so by the Factorisation Theorem, the sample maximum $X_{(n)}$ is sufficient for θ.

(b) Derive the maximum likelihood estimator $\hat\theta_{\mathrm{ML}}$ and show that it is biased.

Solution: The likelihood, viewed as a function of θ is **zero** for $0 < \theta < X_{(n)}$ and then for $\theta \ge X_{(n)}$ it is a positive decreasing function of θ (i.e. θ^{-n}). Thus it is maximised at $\theta = X_{(n)}$. See part (c) for bias.

(c) For what value of the multiplier c_n is the estimator $c_n\hat\theta_{\mathrm{ML}}$

(i) unbiased;

Solution: The CDF of $X_{(n)}$ is given by

$$\begin{aligned}
P_\theta(X_{(n)} \le x) &= P_\theta(X_1 \le x, \ldots, X_n \le x)\\
&= P_\theta(X_1 \le x) \cdots P_\theta(X_n \le x)\\
&= P_\theta(X_1 \le x)^n\\
&= \begin{cases} 0 & \text{for } x < 0\\ \left(\frac{x}{\theta}\right)^n & \text{for } 0 \le x \le \theta\\ 1 & \text{for } x > \theta. \end{cases}
\end{aligned}$$

The corresponding PDF is then

$$f_n(x) = \begin{cases} \frac{nx^{n-1}}{\theta^n} & \text{for } 0 \le x \le \theta,\\ 0 & \text{otherwise.} \end{cases}$$

Thus for positive integer k,

$$E_\theta\left[X_{(n)}^k\right] = \int_0^\theta x^k f_n(x)\,dx = \frac{n}{\theta^n}\int_0^\theta x^{k+n-1}\,dx = \frac{n}{\theta^n}\left[\frac{x^{k+n}}{k+n}\right]_0^\theta = \frac{n\theta^k}{n+k}\,. \tag{8}$$

Setting $k = 1$ gives

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

Thus choosing $c_n = \frac{n+1}{n}$ gives us that

$$E_\theta\left[c_n X_{(n)}\right] = \theta\,,$$

so $c_n X_{(n)}$ is unbiased.

(ii) have the smallest possible MSE, i.e. for which $E_\theta\left[\left(c_n\hat{\theta}_{ML}-\theta\right)^2\right]$ is minimised?

Solution: Setting $k=2$ in (8) above gives

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

So for any choice of c_n, the MSE is

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

Differentiating with respect to c_n gives

$$2n\theta^2\left[\frac{c_n}{n+2}-\frac{1}{n+1}\right]$$

and setting this equal to zero gives

$$c_n=\frac{n+2}{n+1}\,.$$

The resulting estimator

$$\left(\frac{n+2}{n+1}\right)X_{(n)}$$

is *biased*, but still has a smaller MSE than both $X_{(n)}$ itself and the *unbiased* estimator

$$\left(\frac{n+1}{n}\right)X_{(n)}$$

found in the previous part.

(d) Determine the form of the GLRT for testing $H_0\colon\theta=\theta_0$ against $H_1\colon\theta\neq\theta_0$ at level α.

Solution: For any $\theta_1\neq\theta_0$ the likelihood ratio is (assuming $X_{(1)}>0$)

$$L(\theta_1)=\prod_{i=1}^n\left[\frac{f_{\theta_1}(X_i)}{f_{\theta_0}(X_i)}\right]=\frac{1\left\{X_{(n)}\le\theta_1\right\}}{1\left\{X_{(n)}\le\theta_0\right\}}\left(\frac{\theta_0}{\theta_1}\right)^n=\begin{cases}+\infty & \text{if } \theta_0<X_{(n)}\le\theta_1\\ 0 & \text{if } \theta_1<X_{(n)}\le\theta_0\\ \left(\frac{\theta_0}{\theta_1}\right)^n & \text{if } X_{(n)}\le\theta_1 \text{ and } X_{(n)}\le\theta_0.\end{cases}$$

The case that $X_{(n)}>\theta_1$ and $X_{(n)}>\theta_0$ is not possible under either θ_0 or θ_1. Now, the maximum likelihood estimator of θ is $X_{(n)}$ itself. Thus the GLRT rejects for large values of

$$L(X_{(n)})=\begin{cases}+\infty & \text{if } X_{(n)}>\theta_0,\\ \left(\frac{\theta_0}{X_{(n)}}\right)^n & \text{if } X_{(n)}\le\theta_0.\end{cases}$$

Equivalently it rejects for *small* values of

$$\frac{1}{L(X_{(n)})}=\begin{cases}0 & \text{if } X_{(n)}>\theta_0,\\ \left(\frac{X_{(n)}}{\theta_0}\right)^n & \text{if } X_{(n)}\le\theta_0.\end{cases}$$

For any positive constant c,

$$\left\{\frac{1}{L(X_{(n)})}<c\right\}=\left\{X_{(n)}>\theta_0\right\}\cup\left\{X_{(n)}<\theta_0c^{\frac{1}{n}}\right\}\,.$$

We need to choose the critical value $c>0$ so that $P_{\theta_0}\left\{\frac{1}{L(X_{(n)})}<c\right\}=\alpha$. But note that

$$\begin{aligned}P_{\theta_0}\left\{\frac{1}{L(X_{(n)})}<c\right\}&=P_{\theta_0}\left\{X_{(n)}>\theta_0\right\}+P_{\theta_0}\left\{X_{(n)}<\theta_0c^{\frac{1}{n}}\right\}\\&=0+P_{\theta_0}\left\{X_{(n)}<\theta_0c^{\frac{1}{n}}\right\}\,;\end{aligned}$$

for this last probability to be ≤ 1 we need $0 < c \leq 1$; in that case it reduces to

$$P_{\theta_0}\left\{X_1 \leq \theta_0 c^{\frac{1}{n}}\right\}^n = c\,.$$

so we should choose $c = \alpha$. Thus the GLRT is given by

$$\delta(\mathbf{X}) = \begin{cases} 1 & \text{for } X_{(n)} > \theta_0 \\ 0 & \text{for } \theta_0 \alpha^{\frac{1}{n}} \leq X_{(n)} \leq \theta_0 \\ 1 & \text{for } X_{(n)} < \theta_0 \alpha^{\frac{1}{n}}\,. \end{cases}$$