Mathematical Statistics Assignment 2

Liam Carroll - 830916

17th June 2020

Q1. Loss functions

Let $X_1, \dots, X_n \overset{i.i.d.}{\sim} \text{Uniform}(0, \theta),\ \theta \in \Theta = (0, \infty)$. Consider estimators of θ of the form $T_b = bX_{(n)}$, where $X_{(n)} = \max(X_1, \dots, X_n)$.

We first note the standard fact that the maximum order statistic is distributed as

$$F_{X_{(n)}}(x) = \prod_{i=1}^{n} P(X_i \leq x) = F_{X_1}(x)^n = \begin{cases} 0 & \text{if } x < 0 \\ \left(\frac{x}{\theta}\right)^n & \text{if } 0 \leq x \leq \theta \\ 1 & \text{if } x > \theta \end{cases} \tag{1.1}$$

$$\text{with pdf} \quad f_{X_{(n)}}(x) = \frac{nx^{n-1}}{\theta^n} \mathbb{1}(0 \leq x \leq \theta)\,. \tag{1.2}$$

Hence we can calculate

$$\mathbb{E}[X_{(n)}] = \int_0^\theta x \frac{n}{\theta^n} x^{n-1} dx = \frac{n}{n+1}\theta\,, \tag{1.3}$$

$$\text{and} \quad \mathbb{E}[X_{(n)}^2] = \int_0^\theta x^2 \frac{n}{\theta^n} x^{n-1} dx = \frac{n}{n+2}\theta^2\,. \tag{1.4}$$

Part a)

We will use the loss function $L(\theta, t) = (t - \theta)^2$ to calculate the risk function $R(\theta, T_b)$. We can calculate

$$\begin{aligned} R(\theta, T_b) = \mathbb{E}[L(\theta, T_b)] &= \mathbb{E}[(bX_{(n)} - \theta)^2] \\ &= b^2 \mathbb{E}[X_{(n)}^2] - 2b\theta \mathbb{E}[X_{(n)}] + \theta^2 \\ &= \theta^2 \underbrace{\left(\frac{n}{n+2} b^2 - \frac{2n}{n+1} b + 1 \right)}_{f(b)}\,. \end{aligned} \tag{1.5}$$

We can then solve $\partial R / \partial b = 0$ (i.e. $f'(b) = 0$) for all values of $\theta \in \Theta$ by noting that it is a simple quadratic, hence

$$\tilde{b} = -\frac{1}{2} \left(\frac{-2n}{n+1} \right) \left(\frac{n+2}{n} \right) = \frac{n+2}{n+1} \tag{1.6}$$

is the value of b that minimises the risk function.

Part b)

Now instead consider the loss function $L(\theta, t) = t/\theta - 1 - \log(t/\theta)$ and calculate its corresponding risk:

$$\begin{aligned} R(\theta, T_b) &= \mathbb{E}[bX_{(n)}/\theta - 1 - \log(bX_{(n)}/\theta)] \\ &= \underbrace{\frac{b}{\theta}\frac{n\theta}{n+1} - 1 - \log(b)}_{g(b)} - \mathbb{E}[\log(X_{(n)}/\theta)] \,. \end{aligned} \tag{1.7}$$

To optimise R we can calculate the minimum of $g(b)$ to find

$$g'(b) = 0 = \frac{n}{n+1} - \frac{1}{b}\,, \quad \text{so} \quad \tilde{b} = \frac{n+1}{n} \tag{1.8}$$

is the value of b that minimises the risk $R(\theta, T_b)$.

Q2. Bayesian approach

Let $X_1, \ldots, X_n$ be a random sample from a population with pdf

$$f(x|\theta) = \sqrt{\frac{2\theta}{\pi}} e^{-\theta x^2} \mathbb{1}(x \geq 0)\,, \quad \text{so} \quad f(\vec{\mathbf{x}}|\theta) = \left(\frac{2\theta}{\pi}\right)^{n/2} e^{-\theta \sum_{i=1}^n x_i^2} \mathbb{1}(\vec{\mathbf{x}} \geq 0) \tag{2.1}$$

where $\theta > 0$ is unknown.

Part a)

Define a prior $\pi(\theta)$ as Gamma(a, b) with $a, b > 0$ being known constants, that is,

$$\pi(\theta|a, b) = \frac{1}{\Gamma(a)b^a} \theta^{a-1} e^{-\theta/b} \mathbb{1}(\theta > 0)\,. \tag{2.2}$$

Then we can first calculate the marginal distribution of $\vec{\mathbf{x}}$ for $f(\vec{\mathbf{x}}, \theta) = f(\vec{\mathbf{x}}|\theta)\pi(\theta|a, b)$, where we set $K = \sum_{i=1}^n x_i^2$ for notational simplicity:

$$\begin{aligned}
m(\vec{\mathbf{x}}) &= \int_\Theta f(\vec{\mathbf{x}}, \theta) d\theta \\
&= \int_0^\infty \left(\frac{2\theta}{\pi}\right)^{n/2} e^{-K\theta} \frac{1}{\Gamma(a)b^a} \theta^{a-1} e^{-\theta/b} d\theta \\
&= \left(\frac{2}{\pi}\right)^{n/2} \frac{1}{\Gamma(a)b^a} \int_0^\infty \theta^{(a+n/2)-1} e^{-(K+1/b)\theta} d\theta \\
&= \left(\frac{2}{\pi}\right)^{n/2} \frac{1}{\Gamma(a)b^a} \int_0^\infty \left(\frac{\alpha}{K + 1/b}\right)^{(a+n/2)-1} e^{-\alpha} \frac{d\alpha}{K + 1/b} \\
&= \left(\frac{2}{\pi}\right)^{n/2} \frac{1}{\Gamma(a)b^a} \left(\frac{1}{K + 1/b}\right)^{a+n/2} \int_0^\infty \alpha^{(a+n/2)-1} e^{-\alpha} d\alpha \\
&= \left(\frac{2}{\pi}\right)^{n/2} \frac{1}{\Gamma(a)b^a} \left(\frac{1}{K + 1/b}\right)^{a+n/2} \Gamma(a + n/2)\,.
\end{aligned} \tag{2.3}$$

Then the posterior distribution is

$$\begin{aligned}
f_{\theta|\vec{\mathbf{x}}}(\theta|\vec{\mathbf{x}}) &= \frac{f(\vec{\mathbf{x}}|\theta)\pi(\theta)}{m(\vec{\mathbf{x}})} \\
&= \left(\left(\frac{2}{\pi}\right)^{n/2} \frac{\Gamma(a + n/2)}{\Gamma(a)b^a} \left(\frac{1}{K + 1/b}\right)^{a+n/2}\right)^{-1} \left(\frac{2}{\pi}\right)^{n/2} \frac{1}{\Gamma(a)b^a} \theta^{(a+n/2)-1} e^{-(K+1/b)\theta} \\
&= \frac{(\sum_{i=1}^n x_i^2 + 1/b)^{a+n/2}}{\Gamma(a + n/2)} \theta^{(a+n/2)-1} e^{-(\sum_{i=1}^n x_i^2+1/b)\theta} \\
&= \pi\left(\theta \;\middle|\; a' = a + n/2,\ b' = \frac{1}{\sum_{i=1}^n x_i^2 + 1/b}\right),
\end{aligned} \tag{2.4}$$

for $\theta > 0$. Hence, since the posterior distribution is also a Gamma distribution, i.e. $f_{\theta|\vec{\mathbf{x}}}(\theta|\vec{\mathbf{x}}) \in \Pi = \{\text{Gamma}(a, b) :\ a, b > 0\}$ for all $\pi \in \Pi$, for all $f \in \mathcal{F}$ (specified by (2.2)) and for all $x \in \mathbb{R}$, we conclude that the Gamma prior is a conjugate prior for θ. $\square$

Part b)

We will calculate the Bayes estimator T_B such that $BR(T_B) = \min_T BR(T)$ under the loss function $L(\theta, t) = (t-\theta)^2$, where $BR(T) = \int_\Theta R(\theta, T)\pi(\theta)d\theta$. From the theorem in class, this is equivalent to an estimator that minimises the posterior expected loss $\mathbb{E}_{\theta|\vec{\mathbf{x}}}[L(\theta, T(\vec{\mathbf{x}}))]$ over all estimators, for each fixed $\vec{\mathbf{x}} \in S$. Then

$$\begin{aligned} \mathbb{E}_{\theta|\vec{\mathbf{x}}}[L(\theta, T(\vec{\mathbf{x}}))] = \mathbb{E}_{\theta|\vec{\mathbf{x}}}[(t-\theta)^2] = t^2 - 2\mathbb{E}_{\theta|\vec{\mathbf{x}}}[\theta]t + \mathbb{E}_{\theta|\vec{\mathbf{x}}}[\theta^2]\,, \\ \text{which is minimised at} \quad t = \mathbb{E}_{\theta|\vec{\mathbf{x}}}[\theta]\,. \end{aligned} \tag{2.5}$$

We then appeal to the fact that for $G \sim \pi(\theta|\alpha, \beta)$ we have $\mathbb{E}[G] = \alpha\beta$, so using (2.4) we have the Bayesian estimator of θ

$$T_B = \mathbb{E}_{\theta|\vec{\mathbf{x}}}[\theta] = \frac{a + n/2}{\sum_{i=1}^n x_i^2 + 1/b}\,. \tag{2.6}$$

Part c)

Using all of the proceeding theorems, the Bayes estimator of $g(\theta) = \sqrt{2/\pi}\theta^{1/2}$ under square error loss is the posterior expected value of $g(\theta)$, hence we can calculate (where C refers to the horrendous constants in the distribution $\text{Gamma}(a', b')$ that we substitute in afterwards),

$$\begin{aligned} \mathbb{E}_{\theta|\vec{\mathbf{x}}}[g(\theta)] &= \int_0^\infty \sqrt{\frac{2}{\pi}}\theta^{1/2} C\ \theta^{a+n/2-1} e^{-(\sum_{i=1}^n x_i^2 + 1/b)\theta} d\theta \\ &= \sqrt{\frac{2}{\pi}} \frac{\Gamma(a + n/2 + 1/2)(\sum_{i=1}^n x_i^2 + 1/b)^{-(a+n/2+1/2)}}{\Gamma(a+n/2)(\sum_{i=1}^n x_i^2 + 1/b)^{-(a+n/2)}} \\ &= \sqrt{\frac{2}{\pi}} \frac{\Gamma(a + (n+1)/2)}{\Gamma(a+n/2)\sqrt{\sum_{i=1}^n x_i^2 + 1/b}}\,. \end{aligned} \tag{2.7}$$

Q3. Moment estimator and asymptotic distributions

Let $X_1, \ldots, X_n$ be a random sample from the following discrete distribution:

$$P(X_1 = 1) = \frac{2(1-\theta)}{2-\theta}, \quad P(X_1 = 2) = \frac{\theta}{2-\theta}, \tag{3.1}$$

where $\theta \in (0,1)$ is unknown. We can first obtain a moment estimator of θ by equating means,

$$m_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \mathbb{E}[X_1] = (1)\frac{2(1-\theta)}{2-\theta} + (2)\frac{\theta}{2-\theta} = \frac{2}{2-\theta},$$
$$\text{so } \tilde{\theta} = 2 - \frac{2}{\overline{X}_n} \tag{3.2}$$

is our method of moments estimator for this distribution. We can then use the delta method to find its asymptotic distribution. The central limit theorem tells us that

$$\sqrt{n}(\overline{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2) \qquad \text{where } \mu = \frac{2}{2-\theta} \text{ and } \sigma^2 = \frac{2\theta(1-\theta)}{(2-\theta)^2}. \tag{3.3}$$

We can then set $\tilde{\theta} = g(\overline{X}_n)$ where

$$g(y) = 2 - \frac{2}{y}, \qquad \text{so} \quad g'(y) = \frac{2}{y^2}. \tag{3.4}$$

Then the delta method tells us

$$\sqrt{n}\left\{g(\overline{X}_n) - g(\mu)\right\} \xrightarrow{d} N(0, \sigma^2 g'(\mu)^2) = N\left(0, \frac{4\sigma^2}{\mu^4}\right) = N\left(0, \frac{\theta(1-\theta)(2-\theta)^2}{2}\right). \tag{3.5}$$

Hence, noting the easy calculation that $g(\mu) = \theta$, we arrive at the asymptotic distribution of $\tilde{\theta}$,

$$\tilde{\theta} = g(\overline{X}_n) \xrightarrow{d} N\left(\theta, \frac{\theta(1-\theta)(2-\theta)^2}{2n}\right). \tag{3.6}$$

□

Q4. Test functions for Gamma

Let $X_1, \ldots, X_n \overset{i.i.d.}{\sim} \text{Gamma}(r, \lambda)$ where $r > 0$ is known and $\lambda > 0$ is unknown. We use the shape-scale parametrisation with pdf

$$f(x|r,\lambda) = \frac{1}{\Gamma(r)\lambda^r} x^{r-1} e^{-x/\lambda} \mathbb{1}(x > 0)\,. \tag{4.1}$$

We can easily calculate $L(\lambda) = f(\vec{\mathbf{x}}|\lambda, r)$ as

$$L(\lambda) = \left(\frac{1}{\Gamma(r)\lambda^r}\right)^n \left(\prod_{i=1}^n x_i\right)^{r-1} \exp\left[-\frac{1}{\lambda}\sum_{i=1}^n x_i\right] \tag{4.2}$$

Part a)

We will first find a most powerful test (MPT) of size α for testing

$$H_0 : \lambda = \lambda_0 \quad \text{versus} \quad H_1 : \lambda = \lambda_1\,, \tag{4.3}$$

where λ_0 and λ_1 are fixed real numbers satisfying $0 < \lambda_0 < \lambda_1$. We know from Neymann-Pearson's lemma that for a continuous jpdf we have a MPT of the form

$$\phi(\vec{\mathbf{x}}_n) = \begin{cases} 1 & \text{if } f(\vec{\mathbf{x}}|\lambda_1) > cf(\vec{\mathbf{x}}|\lambda_0) \\ 0 & \text{if } f(\vec{\mathbf{x}}|\lambda_1) \leq cf(\vec{\mathbf{x}}|\lambda_0) \end{cases} \tag{4.4}$$

for some $c \geq 0$ which we aim to calculate. We consider the first inequality and calculate

$$\begin{aligned}
\left(\frac{1}{\Gamma(r)\lambda_1^r}\right)^n \left(\prod_{i=1}^n x_i\right)^{r-1} \exp\left[-\frac{1}{\lambda_1}\sum_{i=1}^n x_i\right] &> c\left(\frac{1}{\Gamma(r)\lambda_0^r}\right)^n \left(\prod_{i=1}^n x_i\right)^{r-1} \exp\left[-\frac{1}{\lambda_0}\sum_{i=1}^n x_i\right], \\
\implies \qquad \lambda_1^{-nr} \exp\left[-\frac{1}{\lambda_1}\sum_{i=1}^n x_i\right] &> c\lambda_0^{-nr} \exp\left[-\frac{1}{\lambda_0}\sum_{i=1}^n x_i\right], \\
\implies \qquad -nr\log\lambda_1 - \frac{1}{\lambda_1}\sum_{i=1}^n x_i &> \log c - nr\log\lambda_0 - \frac{1}{\lambda_0}\sum_{i=1}^n x_i\,, \\
\implies \qquad \left(\frac{1}{\lambda_0} - \frac{1}{\lambda_1}\right)\sum_{i=1}^n x_i &> \log c + nr\log\frac{\lambda_1}{\lambda_0}\,, \\
\implies \qquad \sum_{i=1}^n x_i &> \frac{\lambda_0\lambda_1}{\lambda_1 - \lambda_0}\left(\log c + nr\log\frac{\lambda_1}{\lambda_0}\right) = c_1\,.
\end{aligned} \tag{4.5}$$

In the last step we used the fact that $\lambda_0 < \lambda_1$, so we didn't have to flip the inequality. So our condition now becomes

$$\mathbb{E}_{\lambda_0}[\phi(\vec{\mathbf{x}}_n)] = P_{\lambda_0}\left(\sum_{i=1}^n X_i > c_1\right) = \alpha\,. \tag{4.6}$$

Then we know from elementary properties of the Gamma function that

$$\sum_{i=1}^{n} X_i \sim \text{Gamma}\left(\sum_{i=1}^{n} r_i, \lambda\right) = \text{Gamma}(nr, \lambda)\,, \tag{4.7}$$

which then allows us to write, under H_0,

$$\frac{2}{\lambda_0}\sum_{i=1}^{n} X_i \sim \text{Gamma}(nr, 2) \sim \chi^2_{2nr}\,. \tag{4.8}$$

So we can then write, where $F(x; 2nr)$ is the CDF of χ^2_{2nr},

$$\begin{aligned}\alpha = P_{\lambda_0}\left(\sum_{i=1}^{n} X_i > c_1\right) &= P_{\lambda_0}\left(\frac{2}{\lambda_0}\sum_{i=1}^{n} X_i > \frac{2c_1}{\lambda_0}\right) \\ &= 1 - P_{\lambda_0}\left(\frac{2}{\lambda_0}\sum_{i=1}^{n} X_i \le \frac{2c_1}{\lambda_0}\right) \\ &= 1 - F\left(\frac{2c_1}{\lambda_0}; 2nr\right)\,. \end{aligned} \tag{4.9}$$

Denoting $\chi^2_{2nr}(k)$ as the kth quantile of a χ^2_{2nr} distribution, we rearrange the above to see

$$c_1 = \frac{\lambda_0}{2}\chi^2_{2nr}(1-\alpha)\,. \tag{4.10}$$

Therefore we can write the most powerful test for this hypothesis test as

$$\phi(\vec{\mathbf{x}}_n) = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} X_i > \frac{\lambda_0}{2}\chi^2_{2nr}(1-\alpha) \\ 0 & \text{if } \sum_{i=1}^{n} X_i \le \frac{\lambda_0}{2}\chi^2_{2nr}(1-\alpha) \end{cases}\,. \tag{4.11}$$

Part b)

We now want to find a uniformly most powerful (UMP) test of size α for testing

$$H_0 : \lambda \le \lambda_0 \quad \text{versus} \quad H_1 : \lambda > \lambda_0\,, \tag{4.12}$$

where $\lambda_0 \in \mathbb{R}^+$ is fixed. Letting $(0, \lambda_0] = \Theta_0 \subset \Theta = (0, \infty)$, we see that we can apply the theorem from lectures. It is clear that the MPT in (4.11) is not dependent on $\lambda_1 \notin \Theta_0$, so we just need to check that $\max_{\lambda\in\Theta_0} \mathbb{E}_\lambda[\phi(\vec{\mathbf{x}}_n)] = \alpha$. We see that

$$\begin{aligned}\mathbb{E}_\lambda[\phi(\vec{\mathbf{x}}_n)] &= P_\lambda\left(\sum_{i=1}^{n} X_i > \frac{\lambda_0}{2}\chi^2_{2nr}(1-\alpha)\right) \\ &= P_\lambda\left(\frac{2}{\lambda}\sum_{i=1}^{n} X_i > \frac{\lambda_0}{\lambda}\chi^2_{2nr}(1-\alpha)\right) \\ &= P_\lambda\left(\chi^2_{2nr} > \frac{\lambda_0}{\lambda}\chi^2_{2nr}(1-\alpha)\right) \end{aligned} \tag{4.13}$$

which, viewed as a function of λ is increasing, meaning that the maximum occurs at the boundary, $\lambda = \lambda_0$. Hence,

$$\max_{\lambda\in\Theta_0} \mathbb{E}_\lambda[\phi(\vec{\mathbf{x}}_n)] = \mathbb{E}_{\lambda_0}[\phi(\vec{\mathbf{x}}_n)] = \alpha\,. \tag{4.14}$$

Therefore, by this theorem we have that (4.11) is a UMP test for this hypothesis.

Part c)

We will now find a likelihood ratio test of size α for testing

$$H_0 : \lambda = \lambda_0 \quad \text{versus} \quad H_1 : \lambda \neq \lambda_0 \,, \tag{4.15}$$

where $\lambda_0 > 0$ is a fixed real number. Under H_0, the MLE of λ is λ_0. For the domain Θ, we can calculate the MLE (where r is known, hence a fixed constant):

$$\begin{aligned}
L(\lambda|\vec{\mathbf{x}}_n) &= \left(\frac{1}{\Gamma(r)\lambda^r}\right)^n \left(\prod_{i=1}^n x_i\right)^{r-1} e^{-\sum_{i=1}^n x_i/\lambda} \mathbb{1}(x_{(1)} > 0) \,, \\
\text{so} \quad \log L &= -n \log \Gamma(r) - rn \log \lambda + (r-1)\left(\sum_{i=1}^n \log x_i\right) - \frac{1}{\lambda}\sum_{i=1}^n x_i \,, \\
\text{so setting} \quad \frac{\partial \log L}{\partial \lambda} &= -\frac{rn}{\lambda} + \frac{1}{\lambda^2}\sum_{i=1}^n x_i = 0 \\
\text{we have} \quad \hat{\lambda} &= \frac{1}{rn}\sum_{i=1}^n x_i = \frac{1}{r}\overline{X}_n \,.
\end{aligned} \tag{4.16}$$

Hence, we have our LRTS,

$$\begin{aligned}
\lambda(\vec{\mathbf{x}}_n) = \frac{L(\lambda_0|\vec{\mathbf{x}}_n)}{L(\hat{\lambda}|\vec{\mathbf{x}}_n)} &= \frac{\left(\frac{1}{\Gamma(r)\lambda_0^r}\right)^n \left(\prod_{i=1}^n x_i\right)^{r-1} e^{-\sum_{i=1}^n x_i/\lambda_0} \mathbb{1}(x_{(1)} > 0)}{\left(\frac{1}{\Gamma(r)\hat{\lambda}^r}\right)^n \left(\prod_{i=1}^n x_i\right)^{r-1} e^{-\sum_{i=1}^n x_i/\hat{\lambda}} \mathbb{1}(x_{(1)} > 0)} \\
&= \left(\frac{\lambda_0}{\hat{\lambda}}\right)^{-nr} e^{nr} e^{-\frac{n}{\lambda_0}\overline{X}_n} \,.
\end{aligned} \tag{4.17}$$

Our likelihood ratio test is then defined as

$$\phi(\vec{\mathbf{x}}_n) = \begin{cases} 1 & \text{if } \lambda(\vec{\mathbf{x}}_n) < c \\ 0 & \text{if } \lambda(\vec{\mathbf{x}}_n) \geq c \end{cases} \,, \tag{4.18}$$

which satisfies $\mathbb{E}_{\lambda_0}[\phi(\vec{\mathbf{x}}_n)] = P_{\lambda_0}(\lambda(\vec{\mathbf{x}}_n) < c) = \alpha$. We first find a better condition on our LRTS (where $c_1 > 0$ is another constant)

$$\begin{aligned}
& \lambda(\vec{\mathbf{x}}_n) < c \\
\Longrightarrow \quad & \left(\frac{\lambda_0}{\hat{\lambda}}\right)^{-nr} e^{nr} e^{-\frac{n}{\lambda_0}\overline{X}_n} < c \\
\Longrightarrow \quad & \left(\frac{\lambda_0}{\hat{\lambda}}\right)^{-nr} e^{-\frac{n}{\lambda_0}\overline{X}_n} < c_1 \\
\Longrightarrow \quad & \left(\frac{\sum_{i=1}^n x_i}{nr\lambda_0}\right)^{nr} e^{-\frac{1}{\lambda_0}\sum_{i=1}^n x_i} < c_1 \,.
\end{aligned} \tag{4.19}$$

As in part a), we can then define

$$Y = \frac{2}{\lambda_0}\sum_{i=1}^n X_i \sim \chi^2_{2nr} \tag{4.20}$$

where our inequality (4.19) now becomes

$$\left(\frac{Y}{2nr}\right)^{nr} e^{-Y/2} < c_1$$
$$\implies \qquad Y^{nr} e^{-Y/2} < c_2 \,, \tag{4.21}$$

for some constant $c_2 > 0$, meaning we can now define

$$g(y) = y^{nr} e^{-y/2} \tag{4.22}$$

which satisfies $P_{\lambda_0}(g(Y) < c_2) = \alpha$.

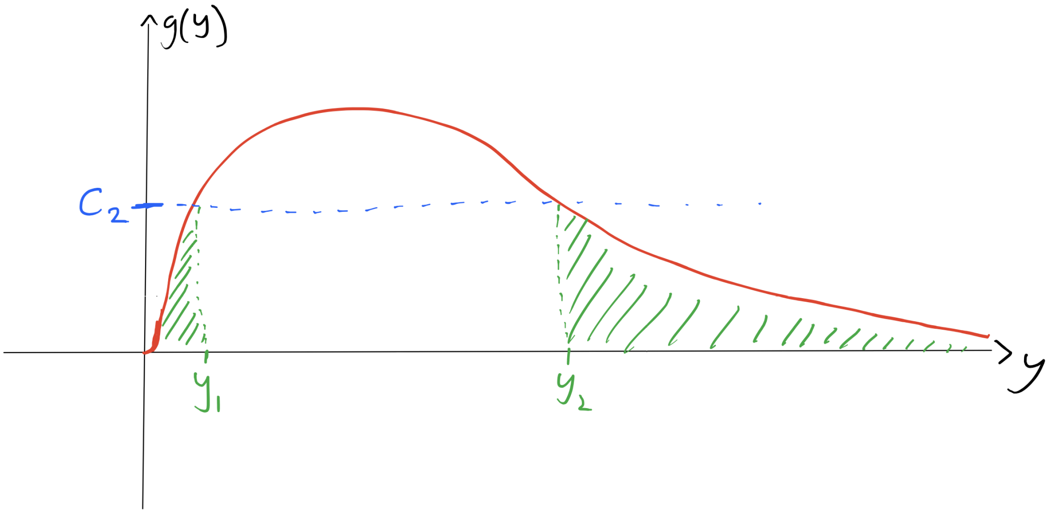

Figure 4.1: Plot of $g(y)$ displaying values for which inequality holds.

Using Figure 4.1 as a guide, we can translate our α condition into

$$P_{\lambda_0}(0 < Y < y_1) + P_{\lambda_0}(Y > y_2) = \alpha \,, \quad \text{where } g(y_1) = g(y_2) = c_2 \,. \tag{4.23}$$

Then since $Y \sim \chi^2_{2nr}$, we can define quantiles $q_1, q_2 > 0$ where

$$P_{\lambda_0}(0 < Y < y_1) = q_1 \;\text{ and }\; P_{\lambda_0}(Y \leq y_2) = q_2 \quad \text{such that } 1 + q_1 - q_2 = \alpha \,. \tag{4.24}$$

Hence we can now write

$$y_1 = \chi^2_{2nr}(q_1) \quad \text{and} \quad y_2 = \chi^2_{2nr}(1 - \alpha + q_1) \,. \tag{4.25}$$

Therefore, after much effort, our acceptance and rejection regions are

$$A_\phi(\lambda_0) = \left[\frac{\lambda_0}{2}\chi^2_{2nr}(q_1), \frac{\lambda_0}{2}\chi^2_{2nr}(1 - \alpha + q_1)\right] \tag{4.26}$$

$$R_\phi(\lambda_0) = A(\lambda_0)^c = \left(0, \frac{\lambda_0}{2}\chi^2_{2nr}(q_1)\right) \cup \left(\frac{\lambda_0}{2}\chi^2_{2nr}(1 - \alpha + q_1), \infty\right) \,, \tag{4.27}$$

meaning our LRT of size α is

$$\phi(\vec{\mathbf{x}}_n) = \begin{cases} 1 & \text{if } \sum_{i=1}^n X_i \in R(\lambda_0) \\ 0 & \text{if } \sum_{i=1}^n X_i \in A(\lambda_0) \end{cases} \,. \tag{4.28}$$

□

Q5. Another UMP Test

Let $X_1, \dots, X_n$ be a random sample from a population with pdf

$$f(x|\theta) = \frac{x^{1/\theta - 1}}{\theta} \mathbb{1}(0 < x < 1)\,, \tag{5.1}$$

where $\theta \in \Theta = (0, \infty)$. We want to find a UMP test for testing

$$H_0 : \lambda \leq \lambda_0 \quad \text{versus} \quad H_1 : \lambda > \lambda_0\,, \tag{5.2}$$

where $\theta_0 \in \Theta$ is fixed. We start by finding a MPT for

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta = \theta_1\,, \tag{5.3}$$

where $\theta_1 > \theta_0$. We note that the joint pdf of f is in an exponential family since we can write

$$\begin{aligned} f(\vec{\mathbf{x}}|\theta) &= \left(\frac{1}{\theta}\right)^n \left(\prod_{i=1}^n x_i\right)^{1/\theta - 1} \prod_{i=1}^n \mathbb{1}(0 < x_i < 1) \\ &= \underbrace{\theta^{-n}}_{c(\theta)} \underbrace{\prod_{i=1}^n \mathbb{1}(0 < x_i < 1)}_{h(\vec{\mathbf{x}}_n)} \exp\left[\underbrace{(1 - 1/\theta)}_{w(\theta)} \underbrace{\left(-\sum_{i=1}^n \log x_i\right)}_{t(\vec{\mathbf{x}}_n)}\right]. \end{aligned} \tag{5.4}$$

Since $w(\theta)$ is non decreasing in θ on Θ, we see that this family of pdf's has a monotone likelihood ratio in $t(\vec{\mathbf{x}}_n)$ as labelled above. By the theorem in lectures, this tells us we have a UMP test of size α as

$$\phi(\vec{\mathbf{x}}_n) = \begin{cases} 1 & \text{if } -\sum_{i=1}^n \log x_i > c \\ 0 & \text{if } -\sum_{i=1}^n \log x_i \leq c \end{cases}. \tag{5.5}$$

We then will need to find the distribution of $t(\vec{\mathbf{x}}_n)$, so we start by finding the distribution of $Y = -\log X$:

$$\begin{aligned} F_Y(y) &= P(-\log X \leq y) \\ &= P(X > e^{-y}) \\ &= 1 - \int_0^{e^{-y}} \frac{1}{\theta} t^{1/\theta - 1} dt \\ &= 1 - [t^{1/\theta}]_0^{e^{-y}} \\ &= 1 - e^{-y/\theta}\,, \end{aligned} \tag{5.6}$$

which tells us that $Y \sim \text{Exp}(1/\theta)$, hence we have

$$t(\vec{\mathbf{x}}_n) = -\sum_{i=1}^n \log x_i \sim \text{Gamma}(n, \theta)\,, \tag{5.7}$$

where Gamma has the shape-scale distribution as in Q4, hence we can use the same facts about chi-square in our calculations. So to determine c, we set

$$\begin{aligned}\alpha = \mathbb{E}_{\theta_0}[\phi(\vec{\mathbf{x}}_n)] &= P_{\theta_0}(t(\vec{\mathbf{x}}_n) > c)\\ &= P_{\theta_0}\left(\frac{2}{\theta_0}t(\vec{\mathbf{x}}_n) > \frac{2}{\theta_0}c\right)\\ &= P_{\theta_0}\left(\chi^2_{2n} > \frac{2}{\theta_0}c\right).\end{aligned} \tag{5.8}$$

Using the exact same arguments and notation as in Q4, we arrive at our UMP test for this hypothesis test,

$$\phi(\vec{\mathbf{x}}_n) = \begin{cases} 1 & \text{if } \sum_{i=1}^n X_i > \frac{\theta_0}{2}\chi^2_{2n}(1-\alpha) \\ 0 & \text{if } \sum_{i=1}^n X_i \leq \frac{\theta_0}{2}\chi^2_{2n}(1-\alpha) \end{cases}. \tag{5.9}$$

□

Q8. Confidence intervals

Part a)

To find a $1-\alpha$ confidence set for λ we can invert the likelihood ratio test established in question 3. We had an acceptance region of

$$\begin{aligned} A_\phi(\lambda_0) &= \left\{ \vec{\mathbf{x}}_n : \frac{\lambda_0}{2}\chi^2_{2nr}(q_1) \leq \sum_{i=1}^n x_i \leq \frac{\lambda_0}{2}\chi^2_{2nr}(1-\alpha+q_1) \right\} \\ &= \left\{ \vec{\mathbf{x}}_n : \frac{2}{\chi^2_{2nr}(1-\alpha+q_1)} \sum_{i=1}^n x_i \leq \lambda_0 \leq \frac{2}{\chi^2_{2nr}(q_1)} \sum_{i=1}^n x_i \right\} . \end{aligned} \tag{8.1}$$

Hence our $1-\alpha$ confidence region for λ is

$$C(\vec{\mathbf{x}}_n) = \left\{ \lambda : \frac{2}{\chi^2_{2nr}(1-\alpha+q_1)} \sum_{i=1}^n x_i \leq \lambda \leq \frac{2}{\chi^2_{2nr}(q_1)} \sum_{i=1}^n x_i \right\} . \tag{8.2}$$

Part b)

Throughout this question we have met the location-scale based statistic

$$Q(\vec{\mathbf{x}}_n, \lambda) = \frac{2}{\lambda_0} \sum_{i=1}^n X_i \sim \chi^2_{2nr} , \tag{8.3}$$

and so since Q does *not* depend on λ, we see that this is a well defined pivotal quantity. Hence we can define $c_1, c_2 > 0$ such that

$$P_\lambda(c_1 \leq Q \leq c_2) = 1-\alpha . \tag{8.4}$$

Setting it to be an equi-tail confidence region then gives us

$$P_\lambda(Q \leq c_1) = P_\lambda(Q \geq c_2) = \alpha/2 , \tag{8.5}$$

hence meaning we have

$$c_1 = \chi^2_{2nr}(\alpha/2) \quad \text{and} \quad c_2 = \chi^2_{2nr}(1-\alpha/2) . \tag{8.6}$$

Therefore our $1-\alpha$ confidence region for λ based on the pivot Q is

$$C(\vec{\mathbf{x}}_n) = \left\{ \lambda : \frac{2}{\chi^2_{2nr}(\alpha/2)} \sum_{i=1}^n x_i \leq \lambda \leq \frac{2}{\chi^2_{2nr}(1-\alpha/2)} \sum_{i=1}^n x_i \right\} . \tag{8.7}$$

Q9. More likelihood ratio tests

Let $X_1, \ldots, X_n \overset{i.i.d}{\sim} N(\mu, \sigma^2)$, where $\mu \in \mathbb{R}$ and $\sigma^2 > 0$ are both unknown.

Part a)

We start by finding a likelihood ratio test of size α for testing

$$H_0 : \mu = \mu_0 \quad \text{versus} \quad H_1 : \mu \neq \mu_0$$

where $\mu_0 \in \mathbb{R}$ is fixed. Hence we define

$$\begin{aligned} \Theta_0 &= \{(\mu, \sigma^2) : \mu = \mu_0,\ \sigma^2 > 0\} \\ \Theta &= \{(\mu, \sigma^2) : \mu \in \mathbb{R},\ \sigma^2 > 0\}\,. \end{aligned} \tag{9.1}$$

We begin by calculating the MLE of $\theta = (\mu, \sigma^2)$ over the two sets to determine the LRTS. The likelihood function for a normal distribution is

$$L(\mu, \sigma^2|\vec{\mathbf{x}}_n) = (2\pi)^{-\frac{n}{2}}(\sigma^2)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \mu)^2\right], \tag{9.2}$$

and then we can differentiate $\log L$ to see

$$\frac{\partial \log L}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^n (x_i - \mu)\,, \qquad \frac{\partial \log L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^n (x_i - \mu)^2\,. \tag{9.3}$$

Setting both derivatives to 0 we see that the MLE for θ over Θ is, where $\hat{\theta} = (\hat{\mu}, \hat{\sigma}^2)$,

$$\hat{\mu} = \overline{x}_n \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \overline{x}_n)^2\,. \tag{9.4}$$

By contrast, over the set Θ_0 we have $\hat{\theta}_0 = (\hat{\mu}_0, \hat{\sigma}_0^2)$ where

$$\hat{\mu}_0 = \mu_0 \qquad \hat{\sigma}_0^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \mu_0)^2\,. \tag{9.5}$$

Therefore we calculate our likelihood ratio test statistic as

$$\begin{aligned} \lambda(\vec{\mathbf{x}}_n) = \frac{L(\hat{\theta}_0|\vec{\mathbf{x}}_n)}{L(\hat{\theta}|\vec{\mathbf{x}}_n)} &= \frac{(2\pi)^{-\frac{n}{2}}\left(\frac{1}{n}\sum_{i=1}^n (x_i - \mu_0)^2\right)^{-\frac{n}{2}} \exp\left[-\frac{\sum_{i=1}^n (x_i-\mu_0)^2}{2\frac{1}{n}\sum_{i=1}^n (x_i-\mu_0)^2}\right]}{(2\pi)^{-\frac{n}{2}}\left(\frac{1}{n}\sum_{i=1}^n (x_i - \overline{x}_n)^2\right)^{-\frac{n}{2}} \exp\left[-\frac{\sum_{i=1}^n (x_i-\overline{x}_n)^2}{2\frac{1}{n}\sum_{i=1}^n (x_i-\overline{x}_n)^2}\right]} \\ &= \left(\frac{\sum_{i=1}^n (x_i - \mu_0)^2}{\sum_{i=1}^n (x_i - \overline{x}_n)^2}\right)^{-\frac{n}{2}} \end{aligned} \tag{9.6}$$

So, the LRT of size α is

$$\phi(\vec{\mathbf{x}}_n) = \begin{cases} 1 & \text{if } \lambda(\vec{\mathbf{x}}_n) < c \\ 0 & \text{if } \lambda(\vec{\mathbf{x}}_n) \geq c \end{cases}, \tag{9.7}$$

where c satisfies $P_{\mu_0}(\lambda(\vec{\mathbf{x}}_n) < c) = \alpha$ and since σ^2 is also unknown, this has to hold for all $\sigma^2 \in \Theta$ too. Noting the following identity derived in the first assignment,

$$\sum_{i=1}^{n}(x_i - \mu_0)^2 = \sum_{i=1}^{n}(x_i - \overline{x}_n)^2 + \sum_{i=1}^{n}(\overline{x}_n - \mu_0)^2 = \sum_{i=1}^{n}(x_i - \overline{x}_n)^2 + n(\overline{x}_n - \mu_0)^2\,, \tag{9.8}$$

we can then calculate, for constants $c_1, c_2 > 0$

$$\begin{aligned} \lambda(\vec{\mathbf{x}}_n) = \left(\frac{\sum_{i=1}^{n}(x_i - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \overline{x}_n)^2}\right)^{-\frac{n}{2}} &< c \\ \implies \qquad \frac{\sum_{i=1}^{n}(x_i - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \overline{x}_n)^2} &> c_1 \\ \implies \qquad \frac{\sum_{i=1}^{n}(x_i - \overline{x}_n)^2 + n(\overline{x}_n - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \overline{x}_n)^2} &> c_1 \\ \implies \qquad \frac{n(\overline{x}_n - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \overline{x}_n)^2} &> c_2\,. \end{aligned} \tag{9.9}$$

We then note that the denominator term looks very close to the sample variance, hence implying that we should multiply by $(n-1)$ and then take the square root to get a familiar distribution. Hence, we have for $c_3 > 0$

$$\implies \qquad \frac{\sqrt{n}|\overline{x}_n - \mu_0|}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x}_n)^2}} > c_3\,. \tag{9.10}$$

We can then define the new statistic under the null hypothesis, where S_n^2 is the sample variance and t_{n-1} is the Student's t-distribution with $n-1$ degrees of freedom,

$$T = \frac{\sqrt{n}(\overline{X}_n - \mu_0)}{S_n} \sim t_{n-1}\,, \tag{9.11}$$

and we see that our condition on the LRTS becomes

$$P_{\mu_0}(|T| > c_3) = 1 - P_{\mu_0}(|T| \leq c_3) = \alpha\,, \tag{9.12}$$

hence indicating that we should choose $c_3 = t_{n-1}(1-\alpha/2)$, the $(1-\alpha/2)$-quantile of t_{n-1}. Therefore, the LRT of size α for this hypothesis testing scenario is

$$\phi(\vec{\mathbf{x}}_n) = \begin{cases} 1 & \text{if } |T| > t_{n-1}(1-\alpha/2) \\ 0 & \text{if } |T| \leq t_{n-1}(1-\alpha/2) \end{cases}. \tag{9.13}$$

□

Part b)

From part a), our acceptance region for this LRT is

$$\begin{aligned}
A_\phi(\theta_0) &= \left\{ \vec{\mathbf{x}}_n : \left| \frac{\sqrt{n}(\overline{X}_n - \mu_0)}{S_n} \right| \leq t_{n-1}(1-\alpha/2) \right\} \\
&= \left\{ \vec{\mathbf{x}}_n : \left|\overline{X}_n - \mu_0\right| \leq \frac{S_n t_{n-1}(1-\alpha/2)}{\sqrt{n}} \right\} \\
&= \left\{ \vec{\mathbf{x}}_n : -\frac{S_n t_{n-1}(1-\alpha/2)}{\sqrt{n}} \leq \overline{X}_n - \mu_0 \leq \frac{S_n t_{n-1}(1-\alpha/2)}{\sqrt{n}} \right\} \\
&= \left\{ \vec{\mathbf{x}}_n : \overline{X}_n - \frac{S_n t_{n-1}(1-\alpha/2)}{\sqrt{n}} \leq \mu_0 \leq \overline{X}_n + \frac{S_n t_{n-1}(1-\alpha/2)}{\sqrt{n}} \right\} . \qquad (9.14)
\end{aligned}$$

Therefore, our $1-\alpha$ confidence set for μ is

$$C(\boldsymbol{X}_n) = \left[\overline{X}_n - \frac{S_n t_{n-1}(1-\alpha/2)}{\sqrt{n}}, \overline{X}_n + \frac{S_n t_{n-1}(1-\alpha/2)}{\sqrt{n}} \right] . \qquad (9.15)$$

We note that this confidence set is indeed an interval.

Part c)

We see that along the way we have already found our pivot quantity, namely T, whose distribution does not depend on μ. For a $1-\alpha$ equi-tail confidence set, in setting $P_\mu(c_1 \leq T \leq c_2) = 1-\alpha$ we have the same calculation as in (8.4) and (8.5), hence giving us

$$c_1 = t_{n-1}(\alpha/2) \qquad \text{and} \qquad c_2 = t_{n-1}(1-\alpha/2) . \qquad (9.16)$$

Therefore our $1-\alpha$ confidence region for μ based on the pivot T is

$$C(\boldsymbol{X}_n) = \left[\overline{X}_n - \frac{S_n t_{n-1}(1-\alpha/2)}{\sqrt{n}}, \overline{X}_n - \frac{S_n t_{n-1}(\alpha/2)}{\sqrt{n}} \right] . \qquad (9.17)$$

Q10. Pivoting on a CDF

Let $X_1, \ldots, X_n$ be a random sample from a population with pdf

$$f(x|\theta) = \frac{3}{\theta^3} x^2 \;\; \mathbb{1}(0 < x < \theta)\,, \tag{10.1}$$

where $\theta > 0$ unknown. An elementary calculation shows that the cdf is

$$F_X(x|\theta) = \begin{cases} 0 & \text{if } \; x < 0 \\ \left(\frac{x}{\theta}\right)^3 & \text{if } \; 0 \leq x \leq \theta \\ 1 & \text{if } \; x > \theta \end{cases} . \tag{10.2}$$

Part a)

We will first find a $1 - \alpha$ confidence interval for θ by pivoting the cdf of $X_{(n)} = \max\{X_1, \ldots, X_n\}$. We first calculate the cdf of $X_{(n)}$:

$$\begin{aligned} F_{X_{(n)}}(x|\theta) = P(X_{(n)} \leq x) &= P(\max\{X_1, \ldots, X_n\} \leq x) \\ &= P(X_1 \leq y, \ldots, X_n \leq x) \\ &= \prod_{i=1}^{n} P(X_i \leq x) = \prod_{i=1}^{n} F_{X_i}(x|\theta) \\ &= \begin{cases} 0 & \text{if } \; x < 0 \\ \left(\frac{x}{\theta}\right)^{3n} & \text{if } \; 0 \leq x \leq \theta \\ 1 & \text{if } \; x > \theta \end{cases} . \end{aligned} \tag{10.3}$$

Therefore, in defining the random variable $F_{X_{(n)}} \sim \text{Unif}(0,1)$, we have a pivotal quantity. We then note that for any fixed value of x, $F_{X_{(n)}}(x|\theta)$ is a decreasing function of θ. Hence by the theorem in class we define $C(\boldsymbol{X}_n) = [\theta_L(x), \theta_U(x)]$ by

$$F_{X_{(n)}}(x|\theta_U(x)) = \alpha_1 \qquad \text{and} \qquad F_{X_{(n)}}(x|\theta_L(x)) = 1 - \alpha_2\,, \tag{10.4}$$

where $\alpha_1, \alpha_2 < 1$ satisfy $\alpha_1 + \alpha_2 = \alpha$. We will then assume an equi-tail confidence interval for simplicity, setting $\alpha_1 = \alpha_2 = \alpha/2$. Then we solve (where we note $0 < \alpha/2 < 1$ when solving),

$$\left(\frac{x}{\theta_U(x)}\right)^{3n} = \frac{\alpha}{2}\,, \qquad \text{so} \quad \theta_U(x) = \left(\frac{2}{\alpha}\right)^{1/3n} x\,, \tag{10.5}$$

$$\text{and similarly} \quad \left(\frac{x}{\theta_L(x)}\right)^{3n} = \frac{2-\alpha}{2}\,, \qquad \text{so} \quad \theta_L(x) = \left(\frac{2}{2-\alpha}\right)^{1/3n} x\,. \tag{10.6}$$

Therefore, our $1 - \alpha$ confidence interval for θ is

$$C(X_{(n)}) = \left\{\theta : \left(\frac{2}{2-\alpha}\right)^{1/3n} X_{(n)} \leq \theta \leq \left(\frac{2}{\alpha}\right)^{1/3n} X_{(n)}\right\} . \tag{10.7}$$

Part b)

This time we construct a confidence interval based on a pivotal quantity. We have already seen that $X_{(n)}$ has a favourable distribution, so appealing to the fact that we can create pivots from a location-scale family, we can define a new pivotal quantity $Y = X_{(n)}/\theta$. We verify that it is indeed a pivot:

$$P(Y \le y) = P(X_{(n)}/\theta \le y) = P(X_{(n)} \le \theta y) = \begin{cases} 0 & \text{if } x < 0 \\ \left(\frac{\theta y}{\theta}\right)^{3n} & \text{if } 0 \le \theta y \le \theta \\ 1 & \text{if } \theta y > \theta \end{cases}$$

$$= \begin{cases} 0 & \text{if } x < 0 \\ y^{3n} & \text{if } 0 \le y \le 1 \\ 1 & \text{if } y > 1 \end{cases} . \tag{10.8}$$

We can hence clearly see that the distribution of Y is *independent of* θ, meaning it is a well defined pivot. We then define $c_1, c_2 > 0$ such that

$$P_\theta(c_1 \le Y \le c_2) = 1 - \alpha\,,$$

and once again using an equi-tail confidence region we set

$$P_\theta(Y \le c_1) = P_\theta(Y \ge c_2) = \alpha/2\,.$$

Respectively, this yields

$$c_1 = \left(\frac{\alpha}{2}\right)^{1/3n} \qquad \text{and} \qquad c_2 = \left(\frac{2-\alpha}{2}\right)^{1/3n} . \tag{10.9}$$

So we can now write our confidence interval as

$$C(Y) = \left\{\theta : \left(\frac{\alpha}{2}\right)^{1/3n} \le \frac{X_{(n)}}{\theta} \le \left(\frac{2-\alpha}{2}\right)^{1/3n}\right\}$$

$$= \left\{\theta : \left(\frac{2}{2-\alpha}\right)^{1/3n} X_{(n)} \le \theta \le \left(\frac{\alpha}{2}\right)^{1/3n} X_{(n)}\right\} \tag{10.10}$$

As anticipated, this is the same interval that we arrived at in part a). Hallelujah!

Q11. Evaluation of confidence intervals

Let $X_1, \ldots, X_n$ be a random sample from a population with pdf

$$f(x|\theta) = \theta x^{\theta-1} \, \mathbb{1}(0 < x < 1) \,, \tag{11.1}$$

where $\theta \in \Theta = (0, \infty)$, with cdf

$$F(x|\theta) = \begin{cases} 0 & \text{if } \; x < 0 \\ x^\theta & \text{if } \; 0 \leq y \leq 1 \\ 1 & \text{if } \; y > 1 \end{cases} . \tag{11.2}$$

Part a)

We will find a $1-\alpha$ confidence interval for θ based on the statistic $T(\boldsymbol{X}_n) = -\sum_{i=1}^n \log X_i$. By boxing smart, we notice that this is actually the same distribution as in Q5, but with $\theta_{Q11} = 1/\theta_{Q5}$. Hence we can use the exact same calculation as in (5.6) and (5.7) to get

$$T(\boldsymbol{X}_n) \sim \text{Gamma}(n, 1/\theta) \,. \tag{11.3}$$

Hence we can scale this statistic to produce our pivot,

$$T' = 2\theta T(\boldsymbol{X}_n) = \text{Gamma}(n, 2) \sim \chi^2_{2n} \,. \tag{11.4}$$

We then find $c_1, c_2 > 0$ such that $P_\theta(c_1 \leq 2\theta T(\boldsymbol{X}_n) \leq c_2) = 1 - \alpha$. As in question 8b), setting an equi-tail once again, we have

$$c_1 = \chi^2_{2n}(\alpha/2) \qquad \text{and} \qquad c_2 = \chi^2_{2n}(1 - \alpha/2) \,. \tag{11.5}$$

Therefore, our $1-\alpha$ confidence interval is

$$C(\boldsymbol{X}_n) = \left\{ \theta : \frac{\chi^2_{2n}(\alpha/2)}{2(-\sum_{i=1}^n \log X_i)} \leq \theta \leq \frac{\chi^2_{2n}(1-\alpha/2)}{2(-\sum_{i=1}^n \log X_i)} \right\} . \tag{11.6}$$

Part b)

We now want to find the shortest $1-\alpha$ interval for θ of the form $[a/T, b/T]$, with T as before and $a \leq b$ are real numbers. We can calculate the confidence coefficient as follows:

$$\begin{aligned} P_\theta\left(\frac{a}{T} \leq \theta \leq \frac{b}{T}\right) &= P_\theta\left(2a \leq 2\theta T \leq 2b\right) \\ &= P\left(2\theta T \leq 2b\right) - P\left(2\theta T \leq 2a\right) \\ &= F_{T'}(2b) - F_{T'}(2a) \,. \end{aligned} \tag{11.7}$$

Noting that we have $\mathbb{E}_\theta[b/T - a/T] = (b-a)\mathbb{E}_\theta[1/T]$, this suggests we want to minimise $b - a$ subject to

$$F_{T'}(2b) - F_{T'}(2a) = 1 - \alpha \,, \tag{11.8}$$

hence we can rearrange to find

$$a = \frac{1}{2} F_{T'}^{-1} \left[F_{T'}(2b) - (1-\alpha) \right] . \tag{11.9}$$

We then note for an arbitrary bijective function $g(x) : \mathbb{R} \to [0,1]$, we have

$$\frac{dg^{-1}(x)}{dx} = \frac{1}{g'(g^{-1}(x))} . \tag{11.10}$$

We see that $F_{T'}^{-1}$ satisfies these requirements, hence we can set

$$h(b) = b - \frac{1}{2} F_{T'}^{-1} \left[F_{T'}(2b) - (1-\alpha) \right] , \tag{11.11}$$

we can then calculate the derivative as follows:

$$\frac{dh}{db} = 1 - \frac{f_{T'}(2b)}{f_{T'}(F_{T'}^{-1} \left[F_{T'}(2b) - (1-\alpha) \right])} . \tag{11.12}$$

Hence, the value of b that minimises h satisfies

$$f_{T'}(2b) = f_{T'}(F_{T'}^{-1} \left[F_{T'}(2b) - (1-\alpha) \right]) . \tag{11.13}$$

Unfortunately, $f_{T'}(t)$ is not actually a bijection, meaning it is difficult to progress further from here.

Whilst the mathematics of this calculation are quite awful to look at, there is a relatively simple intuitive explanation for what we seek. We know from lectures that for a unimodal pdf $f(x)$, if we can find an interval $[a,b]$ such that i) $\int_a^b f(x)dx = 1-\alpha$, ii) $f(a) = f(b) > 0$ and iii) a and b fall either side of the mode of f, then $[a,b]$ is the shortest interval that we seek. Clearly this theorem is telling us that the shortest interval occurs around the region of highest 'mass', being the mode.

Drawing a visual picture, we can imagine a line $y = k$ that begins tangential to the mode on f. As we slowly reduce the value of k (move the line down), hence yielding intercepts of $f(a) = f(b)$ on either side of the mode, the total enclosed integral will be some value M. Our shortest interval is then found by finding the particular value of k such that $M = 1-\alpha$. With suitable numerical calculation, this can be easily determined using such constraints.

Part c)

Suppose θ has the prior $\pi(\theta|r,\lambda)$ as Gamma(r,λ) with the same pdf as in (2.2), where both r and λ are known. We want to find a $1-\alpha$ Bayes highest posterior density (HPD) credible set for θ. We have the posterior distribution as:

$$\begin{aligned} f_{\theta|\vec{\mathbf{x}}_n}(\theta|\vec{\mathbf{x}}_n) &\propto f(\vec{\mathbf{x}}_n|\theta)\pi(\theta|r,\lambda) \\ &\propto \theta^n \left(\prod_{i=1}^n x_i \right)^{\theta-1} \frac{1}{\Gamma(r)\lambda^r} \theta^{r-1} e^{-\theta/\lambda} \mathbb{1}(\theta > 0) \\ &\propto \theta^{n+r-1} \exp\left[-\theta \left(\frac{1}{\lambda} - \sum_{i=1}^n \log x_i \right) \right] \mathbb{1}(\theta > 0) , \end{aligned} \tag{11.14}$$

meaning we can write

$$f_{\theta|\vec{\mathbf{x}}_n}(\theta|\vec{\mathbf{x}}_n) \sim \text{Gamma}\left(n+r,\ \left[\frac{1}{\lambda} - \sum_{i=1}^{n} \log x_i\right]^{-1}\right) . \tag{11.15}$$

Then, we know from lectures that a $1-\alpha$ Bayes HPD credible set for θ has the form

$$C(\vec{\mathbf{x}}_n) = \left\{\theta > 0 : f_{\theta|\vec{\mathbf{x}}_n}(\theta|\vec{\mathbf{x}}_n) \geq k\right\} , \tag{11.16}$$

for some $k > 0$ such that

$$P(\theta \in C(\vec{\mathbf{X}}_n)|\ \vec{\mathbf{X}}_n = \vec{\mathbf{x}}_n) = 1 - \alpha\,. \tag{11.17}$$

Since Gamma is a unimodal distribution, we know that this credible set will take the form of an interval,

$$C(\vec{\mathbf{X}}_n) = \left[\theta_L(\vec{\mathbf{X}}_n), \theta_U(\vec{\mathbf{X}}_n)\right] , \tag{11.18}$$

with the additional constraint from (11.16) giving us

$$\begin{aligned} f_{\theta|\vec{\mathbf{x}}_n}(\theta_L(\vec{\mathbf{X}}_n)|\vec{\mathbf{x}}_n) &= f_{\theta|\vec{\mathbf{x}}_n}(\theta_U(\vec{\mathbf{X}}_n)|\vec{\mathbf{x}}_n) = k \\ \text{so} \quad \theta_L^{n+r-1} \exp\left[-\theta_L\left(\frac{1}{\lambda} - \sum_{i=1}^{n} \log x_i\right)\right] &= \theta_U^{n+r-1} \exp\left[-\theta_U\left(\frac{1}{\lambda} - \sum_{i=1}^{n} \log x_i\right)\right] . \end{aligned} \tag{11.19}$$

As long as all of these constraints are satisfied, we have found our HPD credible set of level $1 - \alpha$ for θ - in order to gain more specific results we would need the assistance of numerical calculations.

□