UNSW Sydney
Department of Statistics
Term 1, 2020
MATH5905 - Statistical Inference
Assignment 2 Solutions

Instructions: This assignment must be submitted no later than **6 pm on Thursday, 23rd April 2020 (week 10)**. Please, declare on the first page that the assignment is your own work, except where acknowledged. State also that you have read and understood the University Rules in respect to Student Academic Misconduct. You need to submit the PDF file on **Moodle** using the assignment submission link by 6 pm on Thursday 23rd April 2020 (week 10).

Problem 1
Let $X = (X_1, X_2, \ldots, X_n)$ be sample of n i.i.d. random variables, each with a density

$$f(x,\theta) = \frac{1}{x\sqrt{2\pi\theta}} \exp\left(-\frac{1}{2\theta}\log^2(x)\right)$$

when $x > 0$ otherwise zero and where $\theta > 0$ is a parameter.

a) Find the distribution of $Y_i = \log X_i$ and hence or otherwise compute $E(\log^2 X_i)$.

b) Find the Fisher information about θ in one observation and in the sample of n observations.

c) Find the Maximum Likelihood Estimator (MLE) of θ and show that it is unbiased.

d) Does the variance of the MLE attain the Cramer Rao bound?
Note: The χ^2_k distribution has mean k and variance $2k$.

e) Determine the asymptotic distribution of the MLE of θ and also the asymptotic distribution of $h(\theta) = e^{\theta}$.

Solution:

a) Find the distribution of $Y_i = \log X_i$ and hence or otherwise compute $E(\log^2 X_i)$.
By applying the density transformation formula:

$$Y = g(X) = \log X$$

which is a one-to-one transformation from $\mathcal{X} = \{x|x > 0\}$ **into** $\mathcal{Y} = \{y|y \in R^1\}$ **with inverse** $X = g^{-1}(Y) = e^Y$ **and derivative**

$$\frac{\mathrm{d}X}{\mathrm{d}Y} = e^Y.$$

Therefore, by applying the transformation density formula, the density function of Y **becomes**

$$\begin{aligned} f_Y(y) &= f_X(g^{-1}(y))\left|\frac{\mathrm{d}X}{\mathrm{d}Y}\right| \\ &= \frac{1}{e^y\sqrt{2\pi\theta}} \exp\left(-\frac{1}{2\theta}\log^2 e^y\right)\left|e^y\right| \\ &= \frac{1}{\sqrt{2\pi\theta}} \exp\left(-\frac{1}{2\theta}y^2\right) \end{aligned}$$

Hence, this is the density of a normal distribution with mean 0 **and variance** θ**. That is** $Y_i = \log X_i \sim N(0,\theta)$ **with** $E(Y_i) = 0$ **and** $\mathbf{Var}(Y_i) = \theta$ **and so**

$$E(\log^2 X) = E(Y^2) = \mathbf{Var}(Y) + (E(Y))^2 = \theta + 0 = \theta.$$

b) Find the Fisher information about θ in one observation and in the sample of n observations.

First, compute the likelihood for one observation, for example X_1 as follows

$$f(x,\theta) = \frac{1}{x\sqrt{2\pi\theta}} \exp\left(-\frac{1}{2\theta}\log^2(x)\right),$$

and the log-likelihood function is

$$\log f(x,\theta) = -\frac{1}{2}\log 2\pi - \log x - \frac{1}{2}\log\theta - \frac{1}{2\theta}\log^2 x.$$

Take the first derivative of the log-likelihood to compute the score as

$$V(X,\theta) = \frac{\partial \log f(X,\theta)}{\partial\theta} = -\frac{1}{2\theta} + \frac{1}{2\theta^2}\log^2 x.$$

and the second derivative

$$\frac{\partial^2 \log f(X,\theta)}{\partial\theta^2} = \frac{1}{2\theta^2} - \frac{1}{\theta^3}\log^2 x.$$

The the information in a single observation is

$$\begin{aligned} I_{X_1}(\theta) &= -E\left(\frac{\partial^2 \log f(X,\theta)}{\partial\theta^2}\right) \\ &= -E\left(\frac{1}{2\theta^2} - \frac{1}{\theta^3}\log^2 x\right) \\ &= -\frac{1}{2\theta^2} + \frac{E(\log^2 X)}{\theta^3} \end{aligned}$$

Then using the fact that $E(\log^2 X) = \theta$ from part (a) we have

$$I_{X_1}(\theta) = -\frac{1}{2\theta^2} + \frac{\theta}{\theta^3} = -\frac{1}{2\theta^2} + \frac{1}{\theta^2} = \frac{1}{2\theta^2}$$

For a sample of n i.i.d. observations,

$$I_X(\theta) = nI_{X_1}(\theta) = \frac{n}{2\theta^2}$$

c) Find the Maximum Likelihood Estimator (MLE) of θ and show that it is unbiased.

The likelihood of the sample

$$\begin{aligned} L(X,\theta) &= \prod_{i=1}^{n} \frac{1}{x\sqrt{2\pi\theta}} \exp\left(-\frac{1}{2\theta}\log^2(x)\right) = \\ &= (2\pi\theta)^{-n/2} \prod_{i=1}^{n} \frac{1}{x_i} \exp\left(-\frac{1}{2\theta}\sum_{i=1}^{n}\log^2 x_i\right) \end{aligned}$$

with log-likelihood

$$\log L(X,\theta) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\theta - \sum_{i=1}^{n}\log\theta - \sum_{i=1}^{n}\log x_i - \frac{1}{2\theta}\sum_{i=1}^{n}\log^2 x_i.$$

Then the score function becomes

$$V(X,\theta) = \frac{\partial \log L(X,\theta)}{\partial\theta} = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n}\log^2 x_i$$

Then by setting the score to zero we can obtain the MLE

$$\hat{\theta}_{\mathbf{mle}} = \frac{1}{n}\sum_{i=1}^{n} \log^2 x_i.$$

Now consider

$$E(\hat{\theta}_{\mathbf{mle}}) = E\left(\frac{1}{n}\sum_{i=1}^{n} \log^2 x_i\right) = \frac{1}{n}nE(log^2 X_1) = \frac{1}{n}n\theta = \theta.$$

Hence the MLE is unbiased for θ.

d) Does the variance of the MLE attain the Cramer Rao bound?
Note: The χ^2_k distribution has mean k and variance $2k$.

Now we can compute the variance as follows

$$\mathbf{Var}(\hat{\theta}_{\mathbf{mle}}) = \mathbf{Var}\left(\frac{1}{n}\sum_{i=1}^{n} \log^2 x_i\right) = \frac{1}{n^2}n\mathbf{Var}(\log^2 X_i) = \frac{1}{n}\mathbf{Var}(\log^2 X_i)$$

Now recall that $Y_i = \log X_i \sim N(0,\theta)$ and so $Y_i/\sqrt{\theta} \sim N(0,1)$ and finally

$$\frac{Y_i^2}{\theta} \sim \chi^2_1$$

Hence

$$2 = \mathbf{Var}(\chi^2_1) = \mathbf{Var}\left(\frac{Y_i^2}{\theta}\right) = \frac{1}{\theta^2}\mathbf{Var}(Y_i^2)$$

which implies that

$$\mathbf{Var}(Y_i^2) = \mathbf{Var}(\log^2 X_i) = 2\theta^2.$$

Hence, the variance of the MLE is

$$\mathbf{Var}(\hat{\theta}_{\mathbf{mle}}) = \frac{1}{n}\mathbf{Var}(\log^2 X_i) = \frac{2\theta^2}{n}$$

which is the same as the CRLB since

$$nI_{X_1}(\theta) = \frac{2\theta^2}{n}$$

and hence the CRLB is attained by the variance of the MLE which makes it also the UMVUE.

e) Determine the asymptotic distribution of the MLE of θ and also the asymptotic distribution of $h(\theta) = e^{\theta}$.

The delta method sates that the asymptotic distribution of a smooth function $\tau(\theta)$ is given by

$$\sqrt{n}(\widehat{\tau(\theta)} - \tau(\theta)) \xrightarrow{d} N\left(0, \frac{(\tau'(\theta)^2)}{nI_{X_1}(\theta)}\right).$$

Then the asymptotic distribution for θ is

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N\left(0, 2\theta^2\right).$$

Then the asymptotic distribution for $h(\theta) = e^{\theta}$ can be computed by noting that $h'(\theta) = e^{\theta}$ then

$$\sqrt{n}(e^{\hat{\theta}} - e^{\theta}) \xrightarrow{d} N\left(0, 2\theta^2 e^{2\theta}\right).$$

Problem 2

Suppose $X = X_1, X_2, \ldots, X_n$ is a sample of n i.i.d. random variables from a population with a density

$$f(x;\theta) = \begin{cases} \tau\theta^\tau x^{-(\tau+1)} & \text{if} \quad x > \theta \\ 0 & \text{if} \quad \text{otherwise} \end{cases}.$$

where $\tau > 0$ is a known constant and $\theta > 0$ is an unknown parameter.

a) Show that the family $\{L(X,\theta\}, \theta > 0$ has a monotone likelihood ratio in the statistic $T = X_{(1)}$.

b) Show that the density of $T = X_{(1)}$ is

$$f_{X_{(1)}}(x) = \begin{cases} \tau n\theta^{\tau n} x^{-(\tau n+1)} & \text{if} \quad x > \theta \\ 0 & \text{if} \quad \text{otherwise} \end{cases}.$$

Hint: It might provide useful to consider the following

$$P(X_{(1)} \leq x) = 1 - P(X_1 \geq x, X_2 \geq x, \ldots, X_n \geq x).$$

c) Find the uniformly most powerful α-size test φ^* of

$$H_0 : \theta \geq \tau \qquad \text{versus} \qquad H_1 : \theta < \tau.$$

d) Calculate the power function of φ^*.

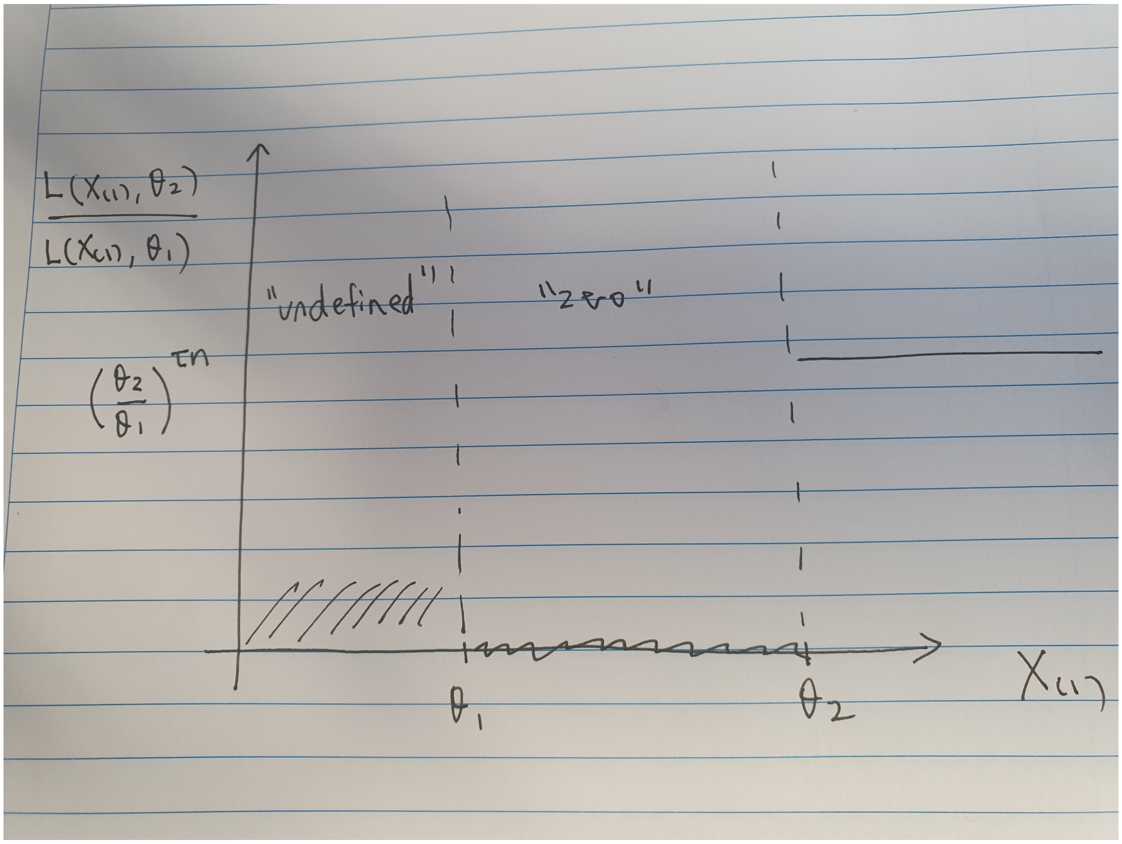
e) Compute the value of the power function at zero, τ and at the threshold constant of the test. Then, sketch a graph of the power function as precisely as possible.

Solution:

a) Show that the family $\{L(X,\theta\}, \theta > 0$ has a monotone likelihood ratio in the statistic $T = X_{(1)}$.
Let $\theta_1 < \theta_2$ and consider the ratio

$$\frac{L(X,\theta_2)}{L(X,\theta_1)} = \left(\frac{\theta_2}{\theta_1}\right)^{\tau n} \frac{I_{(0,X_{(1)})}(\theta_2)}{I_{(0,X_{(1)})}(\theta_1)} = \begin{cases} \textbf{undefined} & \textbf{if} \quad X_{(1)} < \theta \\ 0 & \textbf{if} \quad \theta_1 < X_{(1)} < \theta_2 \\ \left(\dfrac{\theta_2}{\theta_1}\right)^{\tau n} & \textbf{if} \quad X_{(1)} > \theta_2 \end{cases}.$$

A plot of this ratio as a function of $T = X_{(1)}$ can be seen below. It can be seen to be non-decreasing as $X_{(1)}$ increases (as we go along the x axis) when the ratio is properly defined, i.e. when $X_{(1)} > \theta_1$.



b) Show that the density of $T = X_{(1)}$ is

$$f_{X_{(1)}}(x) = \begin{cases} \tau n \theta^{\tau n} x^{-(\tau n+1)} & \text{if} \quad x > \theta \\ 0 & \text{if} \quad \text{otherwise} \end{cases}.$$

Hint: It might provide useful to consider the following

$$P(X_{(1)} \le x) = 1 - P(X_1 \ge x, X_2 \ge x, \dots, X_n \ge x).$$

First, the CDF is calculated as follows

$$F(x) = \int_\theta^x \tau\theta^\tau u^{-(\tau+1)} \mathrm{d}u = \tau\theta^\tau \left[-\frac{u^{-\tau}}{\tau} \right]_{u=\theta}^{u=x} = 1 - \left(\frac{\theta}{x}\right)^\tau.$$

The notice that

$$P(X > x) = 1 - F(x) = \left(\frac{\theta}{x}\right)^\tau.$$

Hence,

$$\begin{aligned} F_{X_{(1)}}(x) = P(X_{(1)} \le x) &= 1 - P(X_{(1)} > x) \\ &= 1 - P(X_1 > x, X_2 > x, \dots, X_n > x) \\ &= 1 - P(X_1 > x)^n \\ &= 1 - \left(\frac{\theta}{x}\right)^{\tau n} \end{aligned}$$

for $x \ge \theta$. Then by differentiating the CDF we obtain the density as

$$f_{X_{(1)}}(x) = \begin{cases} \tau n \theta^{\tau n} x^{-(\tau n+1)} & \textbf{if} \quad x > \theta \\ 0 & \textbf{if} \quad \textbf{otherwise} \end{cases}.$$

c) Find the uniformly most powerful α-size test φ^* of

$$H_0 : \theta \geq \tau \qquad \text{versus} \qquad H_1 : \theta < \tau.$$

Since the family $\{L(X, \theta\}, \theta > 0$ has a monotone likelihood ratio in the statistic $T = X_{(1)}$, we know the structure of the test is

$$\varphi^*(X) = \begin{cases} 1 & \textbf{if} \quad X_{(1)} < k \\ 0 & \textbf{if} \quad X_{(1)} \geq k \end{cases}.$$

To determine k we must "exhaust the α-level" as follows:

$$\alpha = P(X_{(1)} < k | \theta = \tau) = 1 - \left(\frac{\tau}{k}\right)^{\tau n},$$

or equivalently

$$k = \tau(1-\alpha)^{-\frac{1}{\tau n}}.$$

Hence, the UMP α-test is

$$\varphi^*(X) = \begin{cases} 1 & \textbf{if} \quad X_{(1)} < \tau(1-\alpha)^{-\frac{1}{\tau n}} \\ 0 & \textbf{if} \quad X_{(1)} \geq \tau(1-\alpha)^{-\frac{1}{\tau n}} \end{cases}.$$

d) Calculate the power function of φ^*.

The power can be calculated as follows,

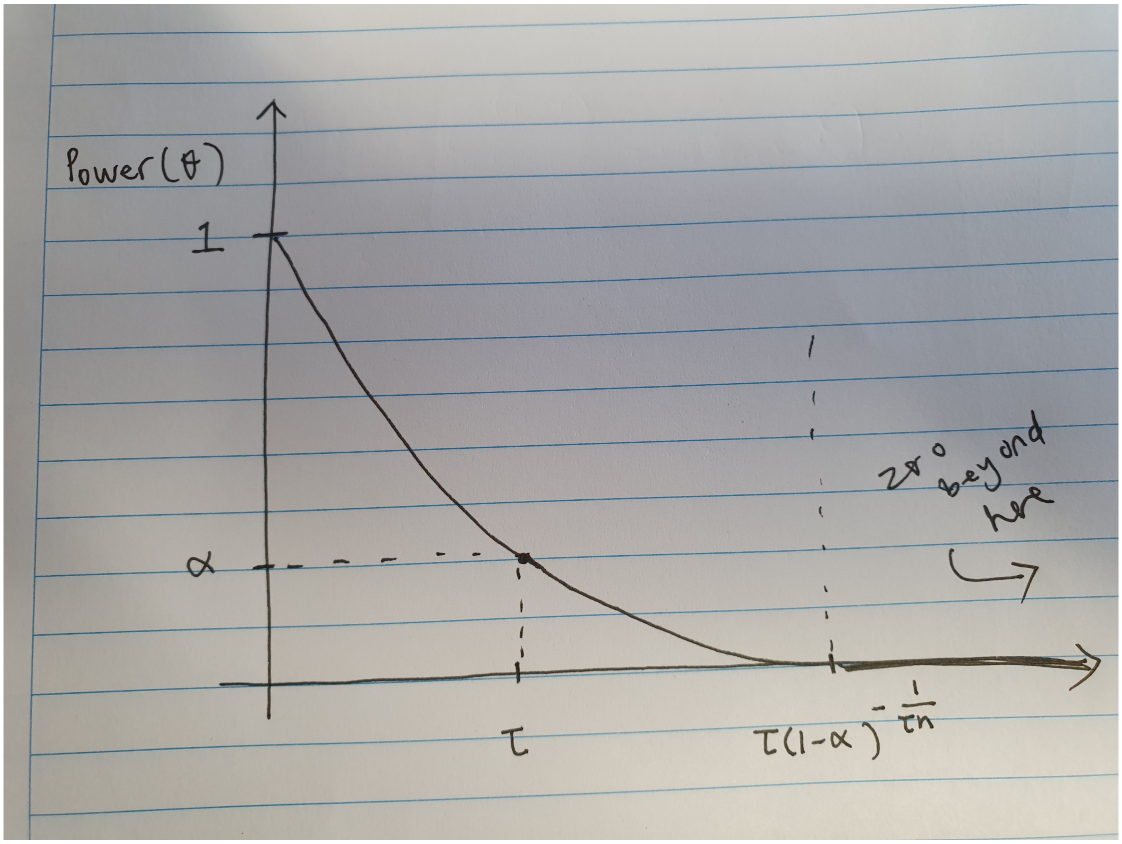
$$\begin{aligned} \textbf{Power}(\theta) = E_\theta \varphi^*(X) &= P(X_{(1)} < \tau(1-\alpha)^{-\frac{1}{\tau n}}) \\ &= 1 - \left(\frac{\theta}{\tau(1-\alpha)^{-\frac{1}{\tau n}}}\right)^{\tau n} \\ &= 1 - \left(\frac{\theta}{\tau}\right)^{\tau n} (1-\alpha) \end{aligned}$$

for $0 < \theta < k = \tau(1-\alpha)^{-\frac{1}{\tau n}}$.

e) Compute the value of the power function at zero, τ and at the threshold constant of the test. Then, sketch a graph of the power function as precisely as possible.

Now we can compute the power function at the following to get some idea how to plot it:

$$\begin{aligned} \textbf{Power}(0) &= 1 - \left(\frac{0}{\tau}\right)^{\tau n}(1-\alpha) = 1 - 0 = 1 \\ \textbf{Power}(\tau) &= 1 - \left(\frac{\tau}{\tau}\right)^{\tau n}(1-\alpha) = 1 - (1-\alpha) = \alpha \\ \textbf{Power}(\tau(1-\alpha)^{-\frac{1}{\tau n}}) &= 1 - \left(\frac{\tau(1-\alpha)^{-\frac{1}{\tau n}}}{\tau}\right)^{\tau n}(1-\alpha) = 1 - 1 = 0 \end{aligned}$$


Power(θ)
1
α
τ
$\tau(1-\alpha)^{-\frac{1}{\tau n}}$
zero beyond here

Problem 3
Suppose that X is uniform on $[0,\theta]$ with density $f(x,\theta)=1/\theta,\ 0<x<\theta$, zero elsewhere.

a) Let $X=(X_1,\dots,X_n)$ be a sample of n i.i.d. observations from this distribution. Similar to the first question it can be shown that the family $\{L(X,\theta)\}$ has a monotone likelihood ratio in the statistic $T=X_{(n)}$.

   i) Compute the density and distribution function for $T=X_{(n)}$.

   ii) Determine the uniformly most powerful α-size test of

$$H_0:\theta\geq\frac{1}{2}\qquad\text{versus}\qquad H_1:\theta<\frac{1}{2}.$$

   iii) Show that the random variable

$$W_n=n\left(1-\frac{X_{(n)}}{\theta}\right)$$

   converges in distribution to the exponential distribution with mean one as $n\to\infty$.

   iv) Hence or otherwise justify that $X_{(n)}$ is a consistent estimator of θ.

b) Now suppose that $X_{(1)},X_{(2)},X_{(3)},X_{(4)},X_{(5)}$ are the order statistics of a random sample of size five from this distribution. Let the observed value of $X_{(5)}$ be $x_{(5)}$. The test rejects $H_0:\theta=2$ and accepts $H_1:\theta\neq 2$ when either $x_{(5)}\leq 1$ or $x_{(5)}>2$.

   i) Find the power function $\gamma(\theta)$ for $0<\theta$ of this particular test.

   ii) Plot the power function $\gamma(\theta)$ for all values $\theta>0$.

Solution:

a) Let $X=(X_1,\dots,X_n)$ be a sample of n i.i.d. observations from this distribution. Similar to the first question it can be shown that the family $\{L(X,\theta)\}$ has a monotone likelihood ratio in the statistic $T=X_{(1)}$.

   i) Compute the density and distribution function for $T=X_{(n)}$.
   The density can be directly computed from the formula in the lecture slides by noting that the distribution function for each X_i is $F(x)=x/\theta$ for $0<x<\theta$

$$f_{X_{(n)}}(x)=n[F(x)]^{n-1}f(x)=n\left(\frac{x}{\theta}\right)^{n-1}\frac{1}{\theta}=\frac{nx^{n-1}}{\theta^n}$$

   for $0<x<\theta$ and zero elsewhere. Then the distribution function can be computed as follows

$$F_{X_{(n)}}(x)=\int_0^x f_{X_{(n)}}(t)\mathrm{d}t=\int_0^x\frac{nt^{n-1}}{\theta^n}\mathrm{d}t=\left[\frac{t^n}{\theta^n}\right]_{t=0}^{t=x}=\left(\frac{x}{\theta}\right)^n$$

   for $0<x<\theta$ and when $x<0$ we have $F(x,\theta)=0$ and when $x>\theta$ we have $F(x,\theta)=1$.

   ii) Determine the uniformly most powerful α-size test of

$$H_0:\theta\geq\frac{1}{2}\qquad\text{versus}\qquad H_1:\theta<\frac{1}{2}.$$

   From part i) we know that the structure of the UMP level α test for testing $H_0:\theta\geq\frac{1}{2}$ against $H_1:\theta<\frac{1}{2}$ must take the form

$$\varphi^*=\begin{cases}1 & \textbf{if}\quad X_{(n)}<k\\ 0 & \textbf{if}\quad X_{(n)}\geq k\end{cases}.$$

Then using part b) we are able to "exhaust the α level" as follows

$$E_{\theta_0}\varphi^* = \alpha = P_{\theta_0}(X_{(n)} < k) = \left(\frac{k}{\theta_0}\right)^n = (2k)^n$$

This implies that

$$k = \frac{1}{2}\alpha^{\frac{1}{n}}.$$

Hence, the UMP α size test is

$$\varphi^*(X) = \begin{cases} 1 & \textbf{if} \quad X_{(n)} < \frac{1}{2}\alpha^{\frac{1}{n}} \\ 0 & \textbf{if} \quad X_{(n)} \geq \frac{1}{2}\alpha^{\frac{1}{n}} \end{cases}.$$

iii) Show that the random variable

$$W_n = n\left(1 - \frac{X_{(n)}}{\theta}\right)$$

converges in distribution to the exponential distribution with mean one as $n \to \infty$.
First, from part (b) we have

$$f_{X_{(n)}}(x) = \frac{nx^{n-1}}{\theta^n}, \quad \textbf{for} \quad 0 < x < \theta, \quad \textbf{with}$$

$$F_{X_{(n)}}(x) = \begin{cases} 0 & \textbf{if} \quad x < 0 \\ \left(\frac{x}{\theta}\right)^n & \textbf{if} \quad 0 \leq x < \theta \\ 1 & \textbf{if} \quad x \geq \theta \end{cases}$$

Then define the random variable $W_n = n\left(1 - \frac{X_{(n)}}{\theta}\right)$ which then has distribution function:

$$\begin{aligned} F_{W_n}(w) = P(W_n \leq w) &= P\left(n\big(1 - \frac{X_{(n)}}{\theta}\big) \leq w\right) \\ &= P\left(-\frac{X_{(n)}}{\theta} \leq \frac{w}{n} - 1\right) \\ &= P\left(X_{(n)} \geq \theta\big(1 - \frac{w}{n}\big)\right) \\ &= 1 - P\left(X_{(n)} < \theta\big(1 - \frac{w}{n}\big)\right) \end{aligned}$$

Then by evoking the independence of each of the X_i's we get the following

$$\begin{aligned} F_{W_n}(w) &= 1 - P\left(X_1 < \theta\big(1 - \frac{w}{n}\big)\right) \times P\left(X_2 < \theta\big(1 - \frac{w}{n}\big)\right) \ldots P\left(X_n < \theta\big(1 - \frac{w}{n}\big)\right) \\ &= 1 - \left[F_X\left(\theta\big(1 - \frac{w}{n}\big)\right)\right]^n \end{aligned}$$

Hence we have the following

$$F_{W_n}(w) = \begin{cases} 0 & \textbf{if} \quad w < 0 \\ 1 - \left(\frac{\theta(1-\frac{w}{n})}{\theta}\right)^n = 1 - \big(1 - \frac{w}{n}\big)^n & \textbf{if} \quad 0 \leq x < n \\ 1 & \textbf{if} \quad w \geq n \end{cases}$$

Then using the fact that

$$\lim_{n\to\infty}\left(1-\frac{w}{n}\right)^n = e^{-w}$$

we see that

$$\lim_{n\to\infty} F_{W_n}(w) = F_W(w) = \begin{cases} 0 & \textbf{if} \quad w < 0 \\ 1-e^{-w} & \textbf{if} \quad w \geq 0 \end{cases}$$

Here, $F_W(w)$ is the distribution function of an exponential random variable with mean one. Therefore, W_n converges in distribution to an exponential random variable with mean one as $n \to \infty$.

iv) Hence or otherwise justify that $X_{(n)}$ is a consistent estimator of θ.

We consider evaluating $P(|X_{(n)}-\theta| < \epsilon)$ directly by noting that $X_{(n)}$ cannot possibly be greater than θ. Hence

$$P(|X_{(n)}-\theta| < \epsilon) = P(X_{(n)} > \theta-\epsilon) = 1 - P(X_{(n)} \leq \theta - \epsilon)$$

Now, the maximum $X_{(n)}$ is less than some constant if and only if each of the random variables $X_1, \ldots, X_n$ is less than that constant. Therefore, since the X_i are i.i.d.,

$$P(X_{(n)} \leq \theta-\epsilon) = \left[P(X_1 \leq \theta - \epsilon)\right] = \begin{cases} \left(1-\frac{\epsilon}{\theta}\right)^n & \textbf{if} \quad 0 \leq \epsilon < \theta \\ 0 & \textbf{if} \quad \epsilon <\geq \theta \end{cases}$$

since $1-\frac{\epsilon}{\theta}$ is strictly less than one, we conclude no matter what positive value ϵ takes,

$$P(X_{(n)} \leq \theta - \epsilon) \to 0$$

as desired.

b) Now suppose that $X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}, X_{(5)}$ are the order statistics of a random sample of size five from this distribution. Let the observed value of $X_{(5)}$ be $x_{(5)}$. The test rejects $H_0 : \theta = 2$ and accepts $H_1 : \theta \neq 2$ when either $x_{(5)} \leq 1$ or $x_{(5)} > 2$.

i) Find the power function $\gamma(\theta)$ for $0 < \theta$ of this particular test.

We first notice that the distribution function of X is $F(x,\theta) = x/\theta$ for $0 < x < \theta$ and hence the density of $X_{(5)}$ is

$$f_{X_{(5)}}(x) = 5\left(\frac{x}{\theta}\right)^{5-1}\frac{1}{\theta} = \frac{5x^4}{\theta^5} \qquad 0 < x < \theta.$$

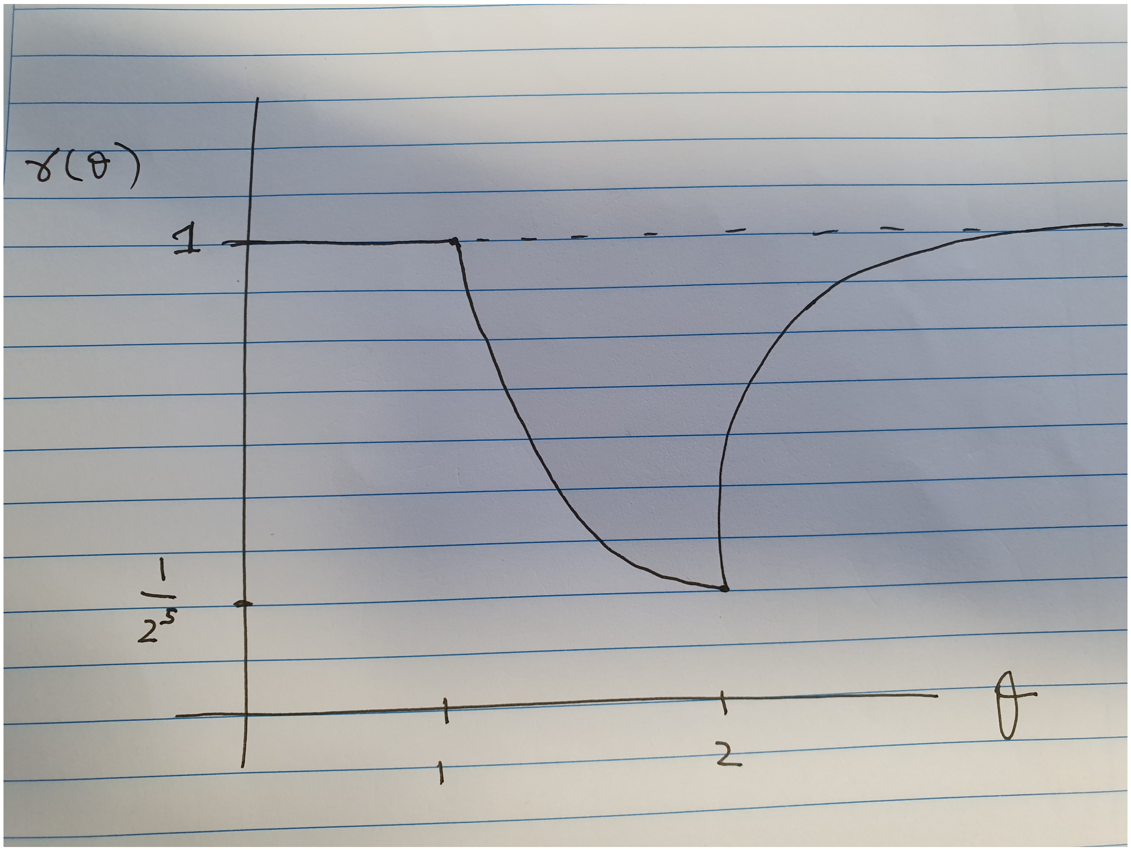
Clearly if $\theta < 1$ then $\gamma(\theta) = 1$ since all values of X must be less than $\theta = 1$ and so must $X_{(5)}$. Then when $1 < \theta \leq 2$ the power function $\gamma(\theta)$ is

$$P(X_{(5)} \leq 1 \quad \textbf{or} \quad X_{(5)} \geq 2) = \int_0^1 \frac{5t^4}{\theta^5}\mathrm{d}t = \frac{1}{\theta^5}$$

Now when $\theta > 2$ the power function $\gamma(\theta)$ is

$$\begin{aligned} P(X_{(5)} \leq 1 \quad \textbf{or} \quad X_{(5)} \geq 2) &= \int_0^1 \frac{5t^4}{\theta^5}\mathrm{d}t + \int_2^\theta \frac{5t^4}{\theta^5} \\ &= \frac{1}{\theta^5} + 1 - \left(\frac{2}{\theta}\right)^5 \\ &= 1 - \frac{31}{\theta^5}. \end{aligned}$$

ii) Plot the power function $\gamma(\theta)$ for all values $\theta > 0$.



Problem 4
Suppose $X_1, \ldots, X_n$ are an i.i.d. sample with density

$$f(x, , \theta) = \theta(1-x)^{\theta-1} \qquad 0 < x < 1,$$

zero elsewhere and $\theta > 0$.

a) Find the form (you do not have to find the threshold constant) of the uniformly most powerful test of

$$H_0 : \theta = 1 \qquad \text{against} \qquad H_1 : \theta > 1.$$

b) Find the likelihood ratio for testing

$$H_0 : \theta = 1 \qquad \text{against} \qquad H_1 : \theta \neq 1.$$

Solution:

a) Find the form (you do not have to find the threshold constant) of the uniformly most powerful test of

$$H_0 : \theta = 1 \qquad \text{against} \qquad H_1 : \theta > 1.$$

The likelihood function for this problem is

$$L(X, \theta) = \theta^n \left[\prod_{i=1}^{n} (1 - x_i) \right]^{\theta - 1}$$

Now for $\theta' < \theta''$, the ratio of the likelihoods is

$$\frac{L(X, \theta'')}{L(X, \theta')} = \left(\frac{\theta''}{\theta'} \right)^n \left[\prod_{i=1}^{n} (1 - x_i) \right]^{\theta'' - \theta'}$$

Since $\theta'' - \theta' > 0$ the above likelihood ratio is monotonically increasing in the statistic

$$T = \prod_{i=1}^{n} (1 - x_i).$$

Hence the structure of the UMP test is given by

$$\varphi^*(X) = \begin{cases} 1 & \textbf{if} \quad \prod_{i=1}^{n}(1 - x_i) > k \\ 0 & \textbf{if} \quad \prod_{i=1}^{n}(1 - x_i) < k \end{cases}.$$

b) Find the likelihood ratio for testing

$$H_0 : \theta = 1 \qquad \text{against} \qquad H_1 : \theta \neq 1.$$

By calculating the partial derivative of the log of the likelihood, the MLE can be obtained:

$$\hat{\theta}_{\mathbf{mle}} = \frac{n}{-\log \prod_{i=1}^{n} (1 - x_i)}.$$

Hence, the likelihood ratio test statistic is

$$\Lambda = \frac{1}{\hat{\theta}_{\mathbf{mle}} \left(\prod_{i=1}^{n} (1 - x_i) \right)^{\hat{\theta}_{\mathbf{mle}} - 1}}.$$

Problem 5
Suppose $X_{(1)} < X_{(2)} < X_{(3)} < X_{(4)} < X_{(5)}$ are the order statistics based on a random sample of size five from the standard exponential density $f(x) = e^{-x}, x > 0$.

a) Find $E(X_{(4)})$.
Hint: Use a computer package to calculate the integral.

b) Find the density of the midrange $A = \frac{1}{2}(X_{(1)} + X_{(5)})$.
Hint: Use the following expansion

$$(e^{-v} - e^{-2u+v})^3 = -3e^{-2u-v} + 3e^{v-4u} - e^{3v-6u} + e^{-3v}.$$

c) Using this result (or otherwise), find $P(A > 2)$.
Hint: Use a computer package to calculate the integral.

Solution:

a) Find $E(X_{(4)})$.
Hint: Use a computer package to calculate the integral.

Using the formula in the lecture notes with $f(x) = e^{-x}, x > 0$ and hence $F(x) = 1 - e^{-x}$, and also with $n = 5$ and $r = 4$,we get

$$f_{X_{(4)}}(x) = \frac{5!}{3!\,1!}(1 - e^{-x})^3(1 - (1 - e^{-x}))^1 e^{-x} = 20(1 - e^{-x})^3 e^{-2x}$$

Then

$$E(X_{(4)}) = \int_0^\infty 20x(1 - e^{-x})^3 e^{-2x} \mathrm{d}x = 1.283$$

This requires numerical integration using a computer program such as `R Studio`.

b) Find the density of the midrange $A = \frac{1}{2}(X_{(1)} + X_{(5)})$.
Hint: Use the following expansion

$$(e^{-v} - e^{-2u+v})^3 = -3e^{-2u-v} + 3e^{v-4u} - e^{3v-6u} + e^{-3v}.$$

The joint density of $X_{(1)}$ and $X_{(5)}$ is

$$\begin{aligned} f_{X_{(1)},X_{(5)}}(x, y) &= \frac{5!}{0!\,3!\,0!} e^{-x}e^{-y}(1 - e^{-x})^0(1 - e^{-y} - (1 - e^{-x}))^3(1 - (1 - e^{-v}))^0 \\ &= 20e^{-x-y}(e^{-x} - e^{-y})^3, \end{aligned}$$

for $0 < x < y < \infty$. If we apply the transformation

$$U = \frac{1}{2}(X_{(1)} + X_{(5)}) \qquad \textbf{and} \qquad V = X_{(1)},$$

then we can equivalently write this as

$$X_{(1)} = V \qquad \textbf{and} \qquad X_{(5)} = 2U - V.$$

The value of the Jacobian of this transformation is equal to -2 since:

$$J(u, v) = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 2 & -1 \end{vmatrix} = 0 \times -1 - 2 \times 1 = -2$$

Therefore the absolute value of the Jacobian is equal to two. Using the density transformation formula we have:

$$\begin{aligned} f_{U,V}(u,v) &= f_{X_{(1)},X_{(n)}}(x(u,v),y(u,v))|J(u,v)| \\ &= f_{X_{(1)},X_{(n)}}(v,2u-v)\cdot 2 \\ &= 20e^{-v-(2u-v)}(e^{-v}-e^{-(2u-v)})^3\cdot 2 \\ &= 40e^{-2u}(e^{-v}-e^{-2u+v})^3, \end{aligned}$$

for $0<v<u<\infty$ **since** $0<X_{(1)}<\frac{1}{2}(X_{(1)}+X_{(5)})<\infty$**.**

Then if $A=\frac{1}{2}(X_{(1)}+X_{(5)})$**, then the marginal density can be computed using integration as follows**

$$\begin{aligned} f_A(u) &= \frac{5}{3}\int_0^u e^{-2u}(e^{-v}-e^{-2u+v})^3\mathrm{d}v \\ &= \frac{5}{3}e^{-2u}\int_0^u -3e^{-2u-v}+3e^{v-4u}-e^{3v-6u}+e^{-3v}\mathrm{d}v \end{aligned}$$

where here we have applied the hint with $x=v$ **and** $y=u$**.**

Then

$$\begin{aligned} f_A(u) &= 40e^{-2u}\left[3e^{-2u-v}+3e^{v-4u}-\frac{1}{3}e^{3v-6u}-\frac{1}{3}e^{-3v}\right]_{v=0}^{v=u} \\ &= 40e^{-2u}\left[\left(3e^{-3u}+3e^{-3u}-\frac{1}{3}e^{-3u}-\frac{1}{3}e^{-3u}\right)-\left(3e^{-2u}+3e^{-4u}-\frac{1}{3}e^{-6u}-\frac{1}{3}\right)\right] \\ &= 40e^{-2u}\left(\frac{1}{3}-3e^{-2u}+\frac{16}{3}e^{-3u}-3e^{-4u}+\frac{1}{3}e^{-6u}\right) \end{aligned}$$

for $0<u<\infty$**.**

c) Using this result (or otherwise), find $P(A>2)$.
Hint: Use a computer package to calculate the integral.

$$P(A>2)=\int_2^\infty 40e^{-2u}\left(\frac{1}{3}-3e^{-2u}+\frac{16}{3}e^{-3u}-3e^{-4u}+\frac{1}{3}e^{-6u}\right)\mathrm{d}u\approx 0.1138547.$$