University of New South Wales
School of Mathematics and Statistics

MATH5905 Statistical Inference
Term One 2024

Assignment Two - Solutions

Given: Monday 29 March 2024 **Due date: Sunday 14 April 2024**

Problem 1
Let $\mathbf{X} = (\mathbf{X_1}, \mathbf{X_2}, \dots, \mathbf{X_n})$ be i.i.d. random variables, each with a density

$$f(x,\theta) = \begin{cases} \frac{1}{x\sqrt{2\pi\theta}} e^{\{-\frac{1}{2\theta}[\log(x)]^2\}}, x > 0 \\ 0 \text{ elsewhere} \end{cases}$$

where $\theta > 0$ is a parameter. (This is called the log-normal density.)

a) Show that $Y_i = \log X_i$ is normally distributed and determine the mean and variance of this normal distribution. Hence find $\mathbb{E}(\log(X_i)^2)$.

Note: *Density transformation formula:* For $Y = W(X)$:

$$f_Y(y) = f_X(W^{-1}(y))|\frac{d(W^{-1}(y))}{dy}| = f_X(x)|\frac{dx}{dy}|.$$

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 $h(\theta) = \theta$ and show that it is unbiased for $h(\theta)$. Is it also the UMVUE of θ? Justify your answer.

d) What is the MLE of $\tilde{h}(\theta) = \sqrt{\theta}$? Determine the asymptotic distribution of the MLE of $\tilde{h}(\theta) = \sqrt{\theta}$.

e) Prove that the family $L(\mathbf{X}, \theta)$ has a monotone likelihood ratio in $T = \sum_{i=1}^n (\log(X_i))^2$.

f) Argue that there is a uniformly most powerful (UMP) $\alpha-$size test of the hypothesis $H_0 : \theta \leq \theta_0$ against $H_1 : \theta > \theta_0$ and exhibit its structure.

g) Using f) (or otherwise), find the threshold constant in the test and hence determine completely the uniformly most powerful $\alpha-$ size test φ^* of

$$H_0 : \theta \leq \theta_0 \text{ versus } H_1 : \theta > \theta_0.$$

Solution:

a) Setting $Y_i = \log X_i = W(X_i)$ implies that $X_i = e^{Y_i} = W^{-1}(Y_i)$. Then $\frac{dW^{-1}(y)}{dy} = e^y$ and $y \in (-\infty, \infty)$. Applying the density transformation formula we get

$$f_{Y_i}(y) = f_{X_i}(e^y)e^y = \frac{e^y}{\sqrt{2\pi\theta}e^y} e^{-\frac{1}{2\theta}\left(\log e^y\right)^2} =$$

$$= \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{(y)^2}{2\theta}}, y \in (-\infty, \infty).$$

which is the density of $N(0, \theta)$ random variable.

Hence $\mathbb{E}(\log(X_i)^2) = \mathbb{E}(Y_i^2) = Var(Y) + (\mathbb{E}(y))^2 = \theta + 0 = \theta$.

b)

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

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

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

Hence, using a) we get for the information $I_{X_1}(\theta)$ about θ in one observation:

$$I_{X_1}(\theta) = -\mathbb{E}\left(\frac{\partial^2}{\partial\theta^2}\log f(x, \theta)\right) = -\frac{1}{2\theta^2} + \frac{\theta}{\theta^3} = \frac{1}{2\theta^2}.$$

Due to the additivity property, we get for the information in the sample of size n:

$$I_{\mathbf{X}}(\theta) = \frac{n}{2\theta^2}.$$

c) We substitute in $\log f(x, \theta)$ the data $x_1, x_2, \ldots, x_n$ and after summing we get

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

Hence we get for the score

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

Setting the score equal to zero (to find a stationary point) and solving the equation w.r. θ we get the MLE

$$\hat{\theta} = \frac{\sum_{i=1}^n (\log(x_i))^2}{n}.$$

This stationary point indeed delivers a maximum of the log-likelihood as calculating the second partial derivative at it we get

$$\frac{\partial^2}{\partial\theta^2}\log L(\mathbf{x}, \theta)_{|\hat{\theta}} = \frac{n}{2\theta^2} - \frac{n\hat{\theta}}{\theta^3}_{|\hat{\theta}} = -\frac{n}{2\hat{\theta}} < 0$$

(as $\hat{\theta}$ is clearly > 0).

The MLE is an unbiased for θ. Indeed, from a) we know that $\mathbb{E}(\log(X_i)^2) = \theta$ holds, so

$$\mathbb{E}(\hat{\theta}) = \frac{n\theta}{n} = \theta.$$

It is also an UMVUE. Indeed, the density $f(x, \theta)$ belongs to a one-parameter exponential family with $a(\theta) = \frac{1}{\sqrt{2\pi\theta}}, b(x) = \frac{1}{x}, c(\theta) = -\frac{1}{2\theta}, d(x) = (\log(x))^2$ hence $\hat{\theta}$ is a function of the complete and sufficient statistic $\sum_{i=1}^n d(X_i)$ and was shown to be unbiased. Hence the conclusion follows from the Lehmann-Scheffe theorem.

d)The MLE (due to the translation-invariance property) is $\hat{\tilde{h}} = \sqrt{\frac{\sum_{i=1}^n (\log(X_i))^2}{n}}$. We now apply the delta method to find the asymptotic distribution. Note that $\tilde{h}'(\theta) = \frac{1}{2\sqrt{\theta}}$. We also have by b):

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, I_{X_1}(\theta)^{-1}) = N(0, 2\theta^2)$$

Hence, by the delta method:

$$\sqrt{n}(\hat{\tilde{h}} - \tilde{h}) \xrightarrow{d} N(0, \theta/2).$$

e) The easiest way to show this is by observing that $f(x, \theta)$ belongs to a one-parameter exponential family (which we already did in c). Note that the function $c(\theta) = -\frac{1}{2\theta}$ is strictly monotonically increasing in θ. Hence we have the MLR property in the statistic $T = \sum_{i=1}^n d(X_i) = \sum_{i=1}^n (\log(X_i))^2$.

f)As we have shown the MLR property in the statistic T, we know from the Blackwell-Girshick Theorem that a UMP α-size test ϕ^* of $H_0 : \theta \leq \theta_0$ against $H_1 : \theta > \theta_0$ exists and its structure is

$$\phi^*(\mathbf{X}) = \begin{cases} 1 \text{ if } T > k \\ 0 \text{ if } T \leq k \end{cases}$$

g) To determine the threshold constant, we observe that we need to "exhaust " the level at the borderline value θ_0 hence $P_{\theta_0}(T > k) = \alpha$ must hold. First we note that from a) we have $Y_i/\sqrt{\theta} \sim N(0, 1)$ which implies that $\sum_{i=1}^n Y_i^2/\theta \sim \chi_n^2$. Now we have to "exhaust the level":

$$P_{\theta_0}(\sum_{i=1}^n (\log(X_i))^2 > k) = P_{\theta_0}(\sum_{i=1}^n {Y_i}^2/\theta_0 > k/\theta_0) = P(\chi_n^2 > k/\theta_0) = \alpha$$

Hence $k/\theta_0 = \chi^2_{n,\alpha}$ where $\chi^2_{n,\alpha}$ is the upper $\alpha \times 100\%$ point of the χ^2_n distribution. The threshold constant (and in this way) the test is completely determined: $k = \theta_0 \times \chi^2_{n,\alpha}$.

Problem 2

Let $X = (X_1, X_2, \ldots, X_n)$ be a sample of n observations each with a uniform in $[0, \theta)$ density

$$f(x, \theta) = \begin{cases} \frac{1}{\theta} \text{ if } 0 < x < \theta \\ 0 \text{ else} \end{cases}$$

where $\theta > 0$ is an unknown parameter. Denote the joint density by $L(X, \theta)$.

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

b) Using a) (or otherwise), completely determine find the uniformly most powerful α-size test of $H_0 : \theta \leq 3$ versus $H_1 : \theta > 3$. Justify all steps in your argument.

c) Find the power function of the test in b) and sketch the graph of $E_\theta \varphi^*$ as accurately as possible.

d) Let $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ be the order statistics from this distribution. Show that $X_{(1)}/X_{(n)}$ and $X_{(n)}$ are independent random variables.

Hint: Find the joint density of $X_{(1)}$ and $X_{(n)}$ first, then use the density transformation formula for vector of two variables.

Solution:

a) We have $L(\mathbf{X},\theta)=\frac{1}{\theta^n}I_{(0,\theta)}(x_{(n)})$. Choose $0<\theta'<\theta''$ and consider the Likelihood ratio

$$\frac{L(\mathbf{X},\theta'')}{L(\mathbf{X},\theta')}=\left(\frac{\theta'}{\theta''}\right)^n\frac{I_{(0,\theta'')}(x_{(n)})}{I_{(0,\theta')}(x_{(n)})}.$$

Analyzing this ratio as a function of $X_{(n)}\in(0,\infty)$ we see that it is:

- is equal to $\left(\frac{\theta'}{\theta''}\right)^n>0$ when $0<X_{(n)}<\theta'$
- is equal to ∞ when $X_{(n)}\in[\theta',\theta'')$
- not defined when $X_{(n)}\geq\theta''$

Hence, in the area where it is defined ($(0,\theta'')$) this ratio is non-decreasing as a function of $X_{(n)}$ and we have a MLR property in $X_{(n)}$.

b) In a) we have shown MLR ratio in the statistic $X_{(n)}$. Therefore the UMP α-size test φ^* exists due to the Blackwell Girshick Theorems and its structure is given by

$$\varphi^*(\mathbf{X})=\begin{cases}1 \text{ if } X_{(n)}>k\\ 0 \text{ if } X_{(n)}\leq k\end{cases}$$

To determine k we must "exhaust" the level α at the borderline value θ_0. This means

$$\alpha=E_{\theta_0}\varphi^*=P_{\theta_0}(X_{(n)}>k)=1-P_{\theta_0}(X_{(n)}\leq k)=1-\Big(P_{\theta_0}(X_1<k)\Big)^n=1-(k/\theta_0)^n$$

From this relation we find that $k=\theta_0(1-\alpha)^{1/n}$. As $\theta_0=3$ in this example, we get $k=3(1-\alpha)^{1/n}$, and the UMP α-size test is completely determined.

c) For any fixed $\theta>0$ the power function is given by

$$E_\theta\varphi^*=P_\theta(X_{(n)}>3(1-\alpha)^{1/n})=\int_{3(1-\alpha)^{1/n}}^{\theta}\frac{nx^{n-1}}{\theta^n}dx=\left(\frac{x}{\theta}\right)^n|_{3(1-\alpha)^{1/n}}^{\theta}=1-(1-\alpha)(\frac{3}{\theta})^n$$

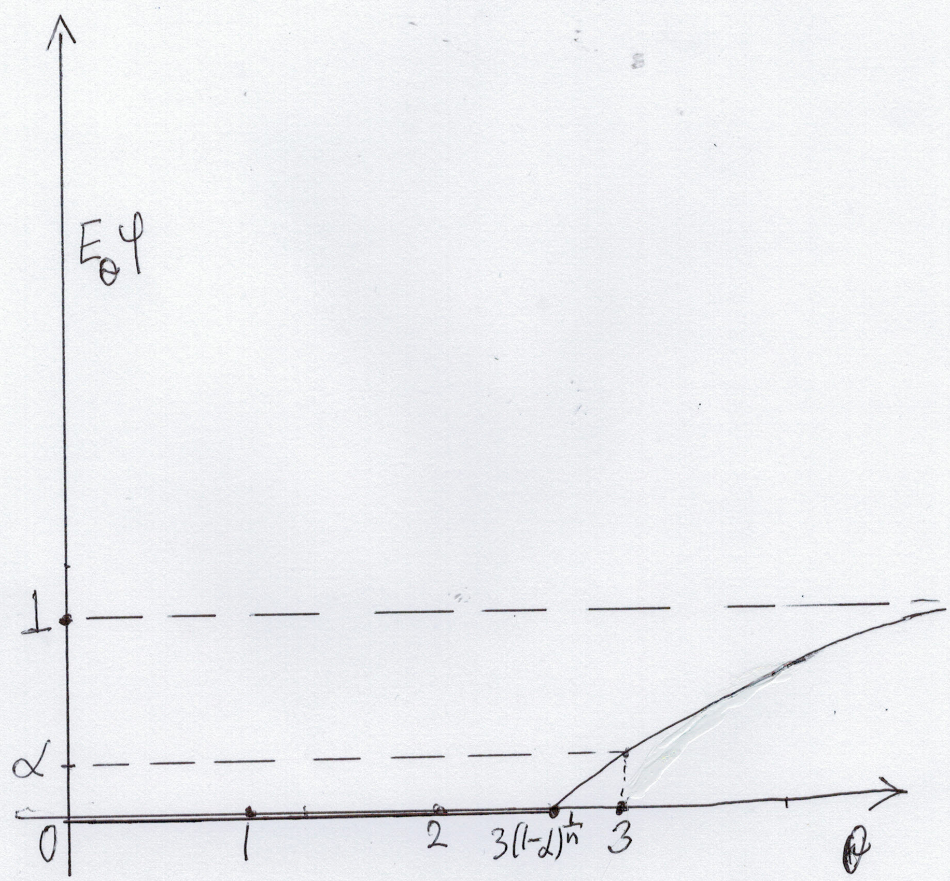
when $\theta>k=3(1-\alpha)^{1/n}$. When $0<\theta\leq 3(1-\alpha)^{1/n}$, the power function is equal to zero. The function is sketched below.

d) We need to find the joint density first. Then we will be using a density transformation formula where one of the transformations is $X_{(1)}/X_{(n)}$ and the other one is kept as $X_{(n)}$. Note that for the uniform in $(0,\theta)$ random variable we have the density $f_X(x)=\frac{1}{\theta}$ when $x\in(0,\theta)$ (and zero else). The cdf is $F_X(x)=\frac{x}{\theta}$ when $x\in(0,\theta)$ (and it is a constant 0 when $x\leq 0$ and a constant 1 when $x\geq\theta$). The joint density of $X_{(1)},X_{(n)}$ is given by

$$f_{X_{(1)},X_{(n)}}(x,y)=\frac{n!}{0!(n-2)!0!}\frac{1}{\theta}\frac{1}{\theta}\left[\frac{y-x}{\theta}\right]^{n-2}=\frac{(n-1)n}{\theta^n}(y-x)^{n-2}, 0<x<y<\theta.$$

Now introduce $U=X_{(1)}/X_{(n)},V=X_{(n)}$. Then $X_{(1)}=UV,x_{(n)}=V$ holds. The constraints $0<X_{(1)}<X_{(n)}<\theta$ transform into $0<U<1<\frac{\theta}{V}$. The Jacobian of the transformation $X_{(1)}=UV,X_{(n)}=V$ is equal to V. Hence the joint density $f_{U,V}(u,v)$ is given by

$$f_{U,V}(u,v)=f_{X_{(1)},X_{(n)}}(uv,v)\times v=\frac{(n-1)n}{\theta^n}(v-uv)^{n-2}v=\frac{(n-1)n}{\theta^n}v^{n-1}(1-u)^{n-2}, 0<u<1, 0<v<\theta.$$



$E_\theta \varphi$
1
α
0
1
2
$3(1-\alpha)^{\frac{1}{n}}$
3
θ

Now, the marginal density of $V = X_{(n)}$ is of course $f_V(v) = \frac{d}{dv}F_V(v) = \frac{d}{dv}(v/\theta)^n = \frac{nv^{n-1}}{\theta^n}$. Hence we have that the conditional density

$$f_{U|V}(u \mid v) = f_{U,V}(u,v)/f_V(v) = (n-1)(1-u)^{n-2}, 0 < u < 1$$

does not depend on the condition (only depends on u) which implies that U and V are independent. Alternatively, you could have calculated the marginal density of U, as well. For this you integrate out the unwanted variable from the joint:

$$f_U(u) = \int_0^\theta \frac{(n-1)n}{\theta^n} v^{n-1}(1-u)^{n-2}dv = (1-u)^{n-2}\frac{(n-1)n}{\theta^n}\int_0^\theta v^{n-1}dv = (n-1)(1-u)^{n-2},$$

which implies that the joint density $f_{U,V}(u,v)$ is a product of the two marginals:

$$f_{U,V}(u,v) = \frac{nv^{n-1}}{\theta^n} \times (n-1)(1-u)^{n-2}.$$

Hence U and V are independent.

Problem 3

Assume $\mathbf{X} = (X_1, X_2, \dots, X_n)$ is a random sample of size n form the density

$$f(x,\theta) = (1+\theta)x^\theta, x \in (0,1),$$

where $\theta > -1$ is an unknown parameter.

1. Why is $T = \sum_{i=1}^n \log(X_i)$ complete and minimal sufficient for θ?
2. If $h(\theta) = \frac{1}{1+\theta}$ argue that the MLE $\hat{h}$ of $h(\theta)$ is unbiased for $h(\theta)$.
3. State the asymptotic distribution of $\sqrt{n}(\hat{h} - h(\theta))$.
4. Suppose that besides the sample $\mathbf{X} = (X_1, X_2, \dots, X_n)$ from the above distribution with parameter $\theta = \theta_1$, another independent sample $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$ is available with parameter θ_2. Show that for the MLE $\hat{g}$ of $g(\theta_1, \theta_2) = \frac{\theta_1+1}{\theta_2+1}$, it holds

$$\sqrt{n}(\hat{g} - g(\theta_1,\theta_2)) \xrightarrow{d} N(0, \frac{2(\theta_1+1)^2}{(\theta_2+1)^2}).$$

Solution: 1.) The density $f(x,\theta)$ belongs to the one-parameter exponential family with

$$a(\theta) = 1+\theta, b(x) = 1, c(\theta) = \theta, d(x) = \log(x).$$

Hence T is complete and minimal sufficient (as it equals $\sum_{i=1}^n d(X_i)$).

2.) We first find the MLE. We have:

$$L(\mathbf{X},\theta) = (1+\theta)^n(\prod_{i=1}^n X_i)^\theta$$

$$\log L(\mathbf{X},\theta) = n\log(1+\theta) + \theta\sum_{i=1}^n \log(X_i)$$

$$V(\mathbf{X},\theta)=\frac{n}{1+\theta}+\sum_{i=1}^{n}\log(X_i)$$

Solving the equation of $V(\mathbf{X},\theta)=0$ gives the MLE as $\hat{\theta}=-1-\frac{n}{\sum_{i=1}^{n}\log(X_i)}$. As $h(\theta)=\frac{1}{1+\theta}$ we can get the MLE $\hat{h}$ by substituting $\hat{\theta}$ for θ in the formula for h (due to the invariance property of the MLE). We get then $\hat{h}=-\frac{\sum_{i=1}^{n}\log(X_i)}{n}$. We can argue about unbiasedness as follows. First we have $\mathbb{E}(\log X_1)=\int_0^1\log(x)(1+\theta)x^{\theta}dx=\int_0^1\log(x)dx^{\theta+1}$. We can integrate by parts by noticing that

$$\lim_{x\to 0_+}\log(x)x^{\theta+1}=0$$

when $\theta>-1$. The integration gives $\mathbb{E}(\log X_1)=-\int_0^1 x^{\theta+1}d\log x=-\int_0^1 x^{\theta}dx=-\frac{1}{1+\theta}$. Hence $\mathbb{E}\hat{h}=n\frac{1}{n(1+\theta)}=\frac{1}{1+\theta}$.

3.)Using the delta method we have asymptotic normality with a zero mean and an asymptotic variance $h'(\theta)^2/I_{X_1}(\theta)$. Clearly $h'(\theta)^2=\frac{1}{(1+\theta)^4}$. Also: $\frac{\partial}{\partial\theta}V(\mathbf{X},\theta)=-\frac{n}{(1+\theta)^2}=-I_{\mathbf{X}}(\theta)=-nI_{X_1}(\theta)$ hence $I_{X_1}(\theta)=\frac{1}{(1+\theta)^2}$ holds. Therefore

$$\sqrt{n}(\hat{h}-h(\theta))\xrightarrow{d}N(0,1/(1+\theta)^2).$$

4.) From the lecture notes we have

$$\sqrt{n}(\hat{g}-g(\theta_1,\theta_2))\xrightarrow{d}N\left\{0,[\frac{\partial g(\theta_1,\theta_2)}{\partial\theta_1},\frac{\partial g(\theta_1,\theta_2)}{\partial\theta_2}][I_{X_1,Y_1}(\theta_1,\theta_2)]^{-1}[\frac{\partial g(\theta_1,\theta_2)}{\partial\theta_1},\frac{\partial g(\theta_1,\theta_2)}{\partial\theta_2}]'\right\}.$$

Now we have:

$$\frac{\partial g(\theta_1,\theta_2)}{\partial\theta_1}=\frac{1}{1+\theta_2}$$
$$\frac{\partial g(\theta_1,\theta_2)}{\partial\theta_2}=-\frac{\theta_1+1}{(\theta_2+1)^2}$$

The 2×2 information matrix is diagonal as $\mathbf{X}$ and $\mathbf{Y}$ are independent and the expected value of the score is equal to zero. In the same way like in 3.) we have $I_{X_1}(\theta_1,\theta_2)=\frac{1}{(1+\theta_1)^2}, I_{Y_1}(\theta_1,\theta_2)=\frac{1}{(1+\theta_2)^2}$. The information matrix is

$$I_{X_1,Y_1}(\theta_1,\theta_2)=\begin{pmatrix}\frac{1}{(1+\theta_1)^2} & 0\\ 0 & \frac{1}{(1+\theta_2)^2}\end{pmatrix}$$

with an inverse

$$[I_{X_1,Y_1}(\theta_1,\theta_2)]^{-1}=\begin{pmatrix}(1+\theta_1)^2 & 0\\ 0 & (1+\theta_2)^2\end{pmatrix}$$

Now substituting in the formula for the asymptotic variance we get

$$[\frac{\partial g(\theta_1,\theta_2)}{\partial\theta_1},\frac{\partial g(\theta_1,\theta_2)}{\partial\theta_2}][I_{X_1,Y_1}(\theta_1,\theta_2)]^{-1}[\frac{\partial g(\theta_1,\theta_2)}{\partial\theta_1},\frac{\partial g(\theta_1,\theta_2)}{\partial\theta_2}]'=\frac{2(\theta_1+1)^2}{(\theta_2+1)^2}.$$

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

1. Find the numerical value of $\mathbb{E}(X_{(2)})$.

2. Find the density of the midrange $M=\frac{1}{2}(X_{(1)}+X_{(4)})$.

3. Find $P(M > 1/2)$.

   Hint: To completely solve this problem, you might need to use a computer package to approximate some integrals. The `integrate` function in `R` might be handy.

Solution:

1) We use the formula in the lecture notes with $f_X(x) = 2e^{-2x}$, $x > 0$ and $F_X(x) = 1 - e^{-2x}$ with $n = 4$ and $r = 2$:

$$f_{X_{(r)}}(y_r) = \frac{n!}{(r-1)!(n-r)!}\Big[F_X(y_r)\Big]^{r-1}\Big[1 - F_X(y_r)\Big]^{n-r} f_X(y_r)$$

to obtain

$$\begin{aligned} f_{X_{(2)}}(y_2) &= \frac{4!}{(2-1)!(4-2)!}\Big[1 - e^{-2y_2}\Big]^{2-1}\Big[1 - (1 - e^{-2y_2})\Big]^{4-2} 2e^{-2y_2} \\ &= 24(1 - e^{-2y_2})e^{-6y_2} \end{aligned}$$

Therefore,

$$\mathbb{E}(X_{(2)}) = \int_0^\infty 24 y_2 e^{-6y_2} \mathrm{d}y_2 - \int_0^\infty 24 y_2 e^{-8y_2} \mathrm{d}y_2$$

Substituting in the first integral $6y_2 = z$ and in the second integral $8y_2 = z$ leads to

$$\mathbb{E}(X_{(2)}) = [\frac{24}{36} - \frac{24}{64}] \int_0^\infty z e^{-z} dz = 2/3 - 3/8 = 7/24 = 0.2917.$$

You could also obtained the latter value via numerical integration in `R Studio`.

2) The joint density of $X_{(i)}$ and $X_{(j)}$ is

$$f_{X_{(i)},X_{(j)}}(x,y) = \frac{n!}{(i-1)!(j-1-i)!(n-j)!} f_X(x) f_X(y)\Big[F_X(x)\Big]^{i-1}\Big[F_X(y) - F_X(x)\Big]^{j-1-i}\Big[1 - F_X(y)\Big]^{n-j}$$

where here we have $i = 1$, $j = 4$, $n = 4$ which gives

$$\begin{aligned} f_{X_{(1)},X_{(4)}}(x,y) &= \frac{4!}{0!2!0!} 2e^{-2x} 2e^{-2y}\Big[1 - e^{-2y} - (1 - e^{-2x})\Big]^{4-1-1} \\ &= 48e^{-2x}e^{-2y}(e^{-2x} - e^{-2y})^2 \\ &= 48e^{-6x-2y} + 48e^{-2x-6y} - 96e^{-4x-4y} \end{aligned}$$

for $0 < x < y < \infty$. Now we apply the density transformation formula with

$$U = \frac{1}{2}(X_{(1)} + X_{(4)}) \qquad \text{and} \qquad V = X_{(1)}.$$

This can equivalently be written as

$$X_{(1)} = V \qquad \text{and} \qquad X_{(4)} = 2U - V.$$

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:

$$
\begin{aligned}
f_{U,V}(u,v) &= f_{X_{(1)},X_{(4)}}(x_{(1)}(u,v), x_{(4)}(u,v))|J(u,v)| \\
&= f_{X_{(1)},X_{(4)}}(v, 2u-v)\cdot 2 \\
&= 96[e^{-4v-4u} + e^{4v-12u} - 2e^{-8u}]
\end{aligned}
$$

for $0 < v < u < \infty$ since $0 < X_{(1)} < \frac{1}{2}(X_{(1)} + X_{(4)}) < \infty$ which implies $2X_{(1)} < X_{(1)} + X_{(4)}$ or $X_{(1)} < X_{(4)}$. Then if $M = \frac{1}{2}(X_{(1)} + X_{(4)})$, then the marginal density can be computed using integration as follows

$$
\begin{aligned}
f_M(u) &= 96[e^{-4u}\int_0^u e^{-4v}dv + e^{-12u}\int_0^u e^{4v}dv - 2e^{-8u}\int_0^u 1dv] \\
&= 24[-e^{-4u}(e^{-4u}-1) + e^{-12u}(e^{4u}-1) - 8ue^{-8u}] \\
&= 24[e^{-4u} - e^{-12u} - 8ue^{-8u}]
\end{aligned}
$$

for $0 < u < \infty$.

3)

$$P(M > 0.5) = \int_{0.5}^{\infty} 24[e^{-4u} - e^{-12u} - 8ue^{-8u}]\mathrm{d}u \approx 0.53232$$

(using the `integrate` function in `R`).