36-705 Intermediate Statistics Fall 2016: Homework 4.

Date: 15th July 2020.

These are my attempted solutions to the homework.

Correction status: pending.

1) Let $X_1, \dots, X_n \sim N(\mu, \Sigma)$ where $X_i \in \mathbb{R}^d$, $\mu \in \mathbb{R}^d$, and $\Sigma \in \mathbb{R}^{d \times d}$.

a) Find a minimal sufficient statistic.

b) Show that $X_1 + X_2$ is not a sufficient statistic.

We propose the following statistic:

$$T(x^n) = \left(\bar{X}_n, S^2\right) = \left(\frac{1}{n}\sum_{i=1}^{n} X_i, \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)(X_i - \bar{X}_n)^T\right)$$

Where $\bar{X}_n \in \mathbb{R}^d$ is the sample mean and $S^2 \in \mathbb{R}^{dxd}$ is the (uncorrected) sample variance.

We will check that this is sufficient, then whether it is minimal sufficient.

The Fisher-Neyman factorisation theorem states that the statistic $T(x^n)$ is sufficient for the parameter $\boldsymbol{\theta} = \{\mu, \Sigma\}$ iff the joint pdf $(x^n; \boldsymbol{\theta})$ can be factored like so:

$$p(x^n; \boldsymbol{\theta}) = h(x^n) g(t; \boldsymbol{\theta})$$

Where $h(x_1, \dots, x_n)$ is a function that does not depend on the parameter $\boldsymbol{\theta}$, and where $g(t(x_1, \dots, x_n); \boldsymbol{\theta})$ is a function depends on the parameter $\boldsymbol{\theta}$, and indirectly on the data $x_1, \dots, x_n$ through the statistic $t(x_1, \dots, x_n)$.

The data has the following joint PDF:

$$\begin{aligned}
p(X_1, \dots, X_n; \boldsymbol{\theta}) &= \frac{1}{(2\pi)^{\frac{nd}{2}} |\Sigma|^{\frac{d}{2}}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(X_i - \mu)^T \Sigma^{-1}(X_i - \mu)\right\} \\
&= \frac{1}{(2\pi)^{\frac{nd}{2}} |\Sigma|^{\frac{d}{2}}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left[(X_i - \bar{X}_n)^T \Sigma^{-1}(X_i - \bar{X}_n) + (\bar{X}_n - \mu)^T \Sigma^{-1}(\bar{X}_n - \mu)\right]\right\} \\
&= \frac{k}{(2\pi)^{\frac{nd}{2}} |\Sigma|^{\frac{d}{2}}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left[(X_i - \bar{X}_n)^T \Sigma^{-1}(X_i - \bar{X}_n)\right]\right\}
\end{aligned}$$

Where we have used relied the "trick" of expressing $\sum_{i=1}^{n}(X_i-\mu)^T\Sigma^{-1}(X_i-\mu)$ in terms of an additional sample mean $\bar{X}_n$ in the 2nd line. And where we have summed over the right-most term of the 2nd line, which is not indexed by the operator, to get $k=\exp\left\{-\frac{n}{2}(\bar{X}_n-\mu)^T\Sigma^{-1}(\bar{X}_n-\mu)\right\}$, as it is not indexed. Both k and the term within the summation in the 3rd line are of the form $\mathbf{x}^T\mathbf{A}\mathbf{x}$ where $\mathbf{x}\in\mathbb{R}^d$ and $\mathbf{A}\in\mathbb{R}^{dxd}$, that is, they are quadratic forms and hence are scalars.

Using the property that the trace of scalar is a scalar $\mathrm{tr}(a)=a$, and the invariance of the operator to cyclic permutations $\mathrm{tr}(ABC)=\mathrm{tr}(CAB)=\mathrm{tr}(BCA)$ we have that

$$\begin{aligned}
p(X_1,...,X_n;\boldsymbol{\theta}) &= \frac{k}{(2\pi)^{\frac{nd}{2}}|\Sigma|^{\frac{d}{2}}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\mathrm{tr}\left[(X_i-\bar{X}_n)^T\Sigma^{-1}(X_i-\bar{X}_n)\right]\right\}\\
&= \frac{k}{(2\pi)^{\frac{nd}{2}}|\Sigma|^{\frac{d}{2}}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\mathrm{tr}\left[(X_i-\bar{X}_n)^T\Sigma^{-1}(X_i-\bar{X}_n)\right]\right\}\\
&= \frac{k}{(2\pi)^{\frac{nd}{2}}|\Sigma|^{\frac{d}{2}}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\mathrm{tr}\left[\Sigma^{-1}(X_i-\bar{X}_n)(X_i-\bar{X}_n)^T\right]\right\}\\
&= \frac{k}{(2\pi)^{\frac{nd}{2}}|\Sigma|^{\frac{d}{2}}}\exp\left\{-\frac{1}{2}\mathrm{tr}\sum_{i=1}^{n}\left[\Sigma^{-1}(X_i-\bar{X}_n)(X_i-\bar{X}_n)^T\right]\right\}\\
&= \frac{k}{(2\pi)^{\frac{nd}{2}}|\Sigma|^{\frac{d}{2}}}\exp\left\{-\frac{n}{2}\mathrm{tr}\left[\Sigma^{-1}\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n)(X_i-\bar{X}_n)^T\right]\right\}\\
&= \frac{1}{(2\pi)^{\frac{nd}{2}}|\Sigma|^{\frac{d}{2}}}\exp\left\{-\frac{n}{2}\left[\mathrm{tr}(\Sigma^{-1}S^2)+(\bar{X}_n-\mu)^T\Sigma^{-1}(\bar{X}_n-\mu)\right]\right\}
\end{aligned}$$

Setting the RHS as $g(\bar{X}_n,S^2;\mu,\Sigma)$ and $h(x_1,...,x_n)=1$, we have that $T=(\bar{X}_n,S^2)$ is a sufficient statistic for $\boldsymbol{\theta}$.

Now T is minimal sufficient iff it has the property that $R(x^n,y^n;\boldsymbol{\theta})$ does not depend on $\boldsymbol{\theta}$ if and only if $T(x^n)=T(y^n)$, where:

$$R(x^n,y^n;\boldsymbol{\theta})=\frac{p(x^n;\boldsymbol{\theta})}{p(y^n;\boldsymbol{\theta})}$$

Showing that $T(x^n)=T(y^n)\implies R(x^n,y^n;\boldsymbol{\theta})$ does not depend on $\boldsymbol{\theta}$:

Indexing S^2 with the observed data x^n and y^n, and cancelling normalisation constants, consider $R(x^n,y^n;\boldsymbol{\theta})$:

$$\begin{aligned}
R(x^n,y^n;\boldsymbol{\theta}) &= \frac{\exp\left\{-\frac{n}{2}\left[\mathrm{tr}(\Sigma^{-1}S^2_{x^n})+(\bar{X}_n-\mu)^T\Sigma^{-1}(\bar{X}_n-\mu)\right]\right\}}{\exp\left\{-\frac{n}{2}\left[\mathrm{tr}(\Sigma^{-1}S^2_{y^n})+(\bar{Y}_n-\mu)^T\Sigma^{-1}(\bar{Y}_n-\mu)\right]\right\}}\\
&= \frac{\mathrm{tr}(\Sigma^{-1}S^2_{x^n})+(\bar{X}_n-\mu)^T\Sigma^{-1}(\bar{X}_n-\mu)}{\mathrm{tr}(\Sigma^{-1}S^2_{y^n})+(\bar{Y}_n-\mu)^T\Sigma^{-1}(\bar{Y}_n-\mu)}
\end{aligned}$$

If $T(x^n) = (\bar{X}_n, S^2_{x^n}) = (\bar{Y}_n, S^2_{y^n}) = T(y^n)$, then we have that $R(x^n, y^n; \boldsymbol{\theta}) = 1$, which does not depend on $\boldsymbol{\theta}$.

Showing that $R(x^n, y^n; \boldsymbol{\theta})$ does not depend on $\boldsymbol{\theta} \implies T(x^n) = T(y^n)$:

Noting that the parameters are fixed and unknown, the only case where there

2) Let $X_1, X_2 \sim \text{Uniform}(0, \theta)$ where $\theta > 0$.

a) Find the distribution of (X_1, X_2) given T where $T = \max\{X_1, X_2\}$.

b) Show that $X_1 + X_2$ is not sufficient.

2a)

One way of assessing the sufficiency of a statistic T is by computing the conditional distribution $f_{X_1,X_2|T}(x_1, x_2|t;\theta)$. If it does not depend on the parameter θ, that is, if we can show that it is of the form, $f_{X_1,X_2|T}(x_1, x_2|t)$, where the θ does not appear, then T is sufficient.

Intuitively, after conditioning on a sufficient statistic, the conditional distribution of the data does not change according to the setting of the parameter.

The conditional PDF $f_{X_1,X_2|T}(x_1, x_2|t;\theta)$ in terms of joint and marginal PDFs is given by:

$$f_{X_1,X_2|T}(x_1, x_2|t;\theta) = \frac{f_{X_1,X_2,T}(x_1, x_2, t;\theta)}{f_t(t;\theta)} = \frac{f_{X_1,X_2}(x_1, x_2;\theta)}{f_T(t;\theta)}$$

Insert reason

Assuming X_1 and X_2 are independent, they have the following joint density function:

$$f_{X_1,X_2}(x_1, x_2, ;\theta) = \begin{cases} \left(\frac{1}{\theta}\right)^2 & 0 \le x_1 \le \theta, 0 \le x_2 \le \theta \\ 0 & \text{otherwise} \end{cases}$$

To compute the PDF of our statistic $T = \max\{x_1, x_2\}$, which is a transformation of the random variables X_1, X_2, we find the CDF $F_T(t)$ and differentiate to find the PDF $f_T(t)$:

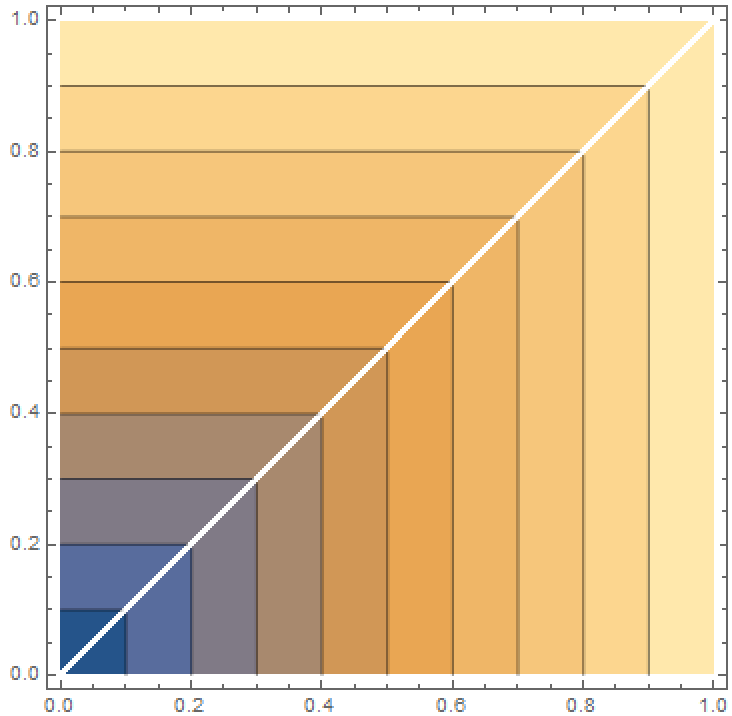
$$F_T(t) = P(T \le t) = P(\{\max\{x_1, x_2\} \le t\}) = \int\int_{A_t} f_{X_1,X_2}(x_1, x_2;\theta)dx_1dx_2$$

We now need to find the set $A_t = \{(x_1, x_2) : \max\{x_1, x_2\} \le t\}$ for relevant values of t. Noting that $X_1, X_2 \sim \text{Uniform}(0, \theta)$; the minimum value of $t = 0$ occurs when either $X_1 = 0$ or $X_2 = 0$ or both, and applying similar arguments to the maximum value of t, we have that the relevant interval is $0 \le t \le \theta$.

We can visualise the set of points contained by contour plots, with each iso-contour corresponding to the set where $t = \max\{x_1, x_2\}$ and find an expression for this area in terms of t. These were plotted by hand, but for presentation purposes were then plotted in the Wolfram Mathematica symbolic computing package:

```
In[1]:= ContourPlot[Max[x1, x2], {x1, 0, 1}, {x2, 0, 1}]
```

Out[1]=



We find that the set A_t corresponds to a shaded square of length t, area t^2 inside a square of area θ^2. This gives the following CDF $F_T(t)$:

$$F_T(t) = \begin{cases} 0 & t < 0 \\ \left(\frac{t}{\theta}\right)^2 & 0 \leq t \leq \theta \\ 1 & t > 0 \end{cases}$$

Differentiating yields the following PDF:

$$f_T(t) = \begin{cases} \left(\frac{2t}{\theta^2}\right) & 0 \leq t \leq \theta \\ 0 & \text{otherwise} \end{cases}$$

Hence the conditional distribution of X_1, X_2 given $T = \max\{x_1, x_2\}$ is:

$$f_{X_1,X_2|T;\theta}(x_1, x_2|t;\theta) = \frac{\frac{1}{\theta^2} \cdot \mathbb{I}(0 \leq x_1, x_2 \leq \theta)}{\frac{2t}{\theta^2} \cdot \mathbb{I}(0 \leq t \leq \theta)} = \frac{1 \cdot \mathbb{I}(x_{(2)} \leq \theta)}{2t \cdot \mathbb{I}(t \leq \theta)} = \frac{1}{2t}$$

Where we have used dropped the lower bound of 0 in the indicator functions as they are redundant - both X_1, X_2 are greater than 0 by assumption, meaning that T is also greater than 0. And where we have because $X_1, X_2 \leq \theta \implies \max\{x_1, x_2\} \leq \theta \implies t \leq \theta$.

The conditional distribution does not depend on θ and hence T is sufficient.

2b)

Following the same arguments as above, we compute conditional distribution of

3) Let $X_1, X_2, \ldots, X_n \sim \text{Uniform}(-\theta, 2\theta)$ where $\theta > 0$. Find the likelihood function.

As the data is IID, denoting the PMF of X_i parametrised by θ as $f_{X_i}(x_i;\theta)$, the likelihood function $L(\theta)$ is given by:

$$L(\theta) = \prod_{i=1}^{n} f_{X_i}(x_i;\theta)$$

The $L(\theta) = 0$ if any of the PMFs $f_{X_i}(X_i;\theta) = 0$.

That means we only need to consider the 1st and nth order statistics of the data, that is $X_{(1)} = \min\{X_1, \ldots, X_n\}$ and $X_{(n)} = \max\{X_1, \ldots, X_n\}$, and their values relative to the parameters.

We can define the likelihood function piecewise like so:

$$L(\theta) = \begin{cases} 0 & \text{if} \quad -\theta > X_{(1)} \\ 0 & \text{if} \quad 2\theta < X_{(n)} \\ 0 & \text{if} \quad 2\theta < X_{(n)} \text{and} -\theta > X_{(1)} \\ \left(\frac{1}{3\theta}\right)^n & \text{if} \quad 2\theta > X_{(n)} \text{and} -\theta < X_{(1)} \end{cases}$$

Hence the likelihood function is given by:

$$L(\theta) = \mathbb{I}(2\theta > X_{(n)})\mathbb{I}(-\theta < X_{(1)}) \left(\frac{1}{3\theta}\right)^n$$

4) Chapter 6 problem 1.

Let $X_1, \ldots, X_n \sim \text{Poisson}(\lambda)$, and let $\widehat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Find the bias, standard error, and mean squared error of this estimator.

We have that the mean squared error is the sum of the squared bias and variance of an estimator:

$$\text{MSE} = B^2 + V$$

$$\mathbb{E}[(\lambda - \widehat{\lambda})^2] = \left(\mathbb{E}[\widehat{\lambda}] - \lambda\right)^2 + \mathbb{E}\left[(\widehat{\lambda} - \mathbb{E}[\widehat{\lambda}])^2\right]$$

where the expectation, $\mathbb{E}[\cdot]$, is with respect to a joint probability distribution parameterised by a value of the parameter at its true value, λ.

As $\widehat{\lambda}$, being a sample mean, is an unbiased estimator of λ, we have that the bias is equal to 0, $B = 0$, and the mean squared error is equal to the variance, $\text{MSE} = V$.

As $\widehat{\lambda}$ is a sample mean, and the data is IID, the mean and variance of $\widehat{\lambda}$ are given by:

$$\mathbb{E}[\widehat{\lambda}] = \lambda, \quad \text{Var}[\widehat{\lambda}] = \frac{\lambda}{n}$$

Hence MSE $= V = \text{Var}[\widehat{\lambda}] = \frac{\lambda}{n}$.

And the standard error is given by se $= \sigma_{\widehat{\lambda}} = \sqrt{\frac{\lambda}{n}}$

5) Chapter 6, problem 3.

Let $X_1, \ldots, X_n \sim \text{Uniform}(0, \theta)$, and let $\widehat{\theta} = 2\bar{X}_n$.

Find the bias, standard error, and mean squared error of this estimator.

IID random variables with uniformly distributed $X_i \sim \text{Unif}(a, b)$ have mean and variance given by:

$$\mathbb{E}[X_i] = \frac{a+b}{2} = \frac{\theta}{2}$$

$$\text{Var}[X_i] = \frac{(b-a)^2}{12} = \frac{\theta^2}{12}$$

Using the same formulae for mean squared error as above, we have that:

$$\text{MSE} = B^2 + V$$

$$\text{MSE} = \mathbb{E}_\theta[(\theta - \widehat{\theta})^2] = \left(\mathbb{E}_\theta[\widehat{\theta}] - \theta\right)^2 + \mathbb{E}_\theta\left[(\widehat{\theta} - \mathbb{E}_\theta[\widehat{\theta}])^2\right]$$

Computing the mean of the estimator, we have that:

$$\mathbb{E}_\theta[\widehat{\theta}] = \mathbb{E}[2\bar{X}_n] = 2\mathbb{E}[\bar{X}_n] = 2\mathbb{E}[X_i] = 2\left(\frac{\theta}{2}\right) = \theta$$

Hence we have that the estimator $\widehat{\theta}$ is unbiased, and that $B = 0$.

The variance of the estimator, in terms of its 2nd moment and mean is given by:

$$\begin{aligned}
\text{Var}[\widehat{\theta}] &= \mathbb{E}[\widehat{\theta}^2] - \mathbb{E}[\widehat{\theta}]^2 \\
&= \mathbb{E}[(2\bar{X}_n)^2] - \mathbb{E}[2\bar{X}_n]^2 \\
&= 4\mathbb{E}[\bar{X}_n^2] - \theta^2
\end{aligned}$$

Computing the 2nd moment of the sample mean yields:

$$\begin{aligned}
\mathbb{E}[\bar{X}_n^2] &= \text{Var}[\bar{X}_n] + \mathbb{E}[\bar{X}_n]^2 \\
&= \frac{1}{n}\text{Var}[X_i] + \theta^2 \\
&= \frac{\theta^2}{12n} + \theta^2
\end{aligned}$$

We can now compute the variance of the estimator:

$$\text{Var}[\widehat{\theta}] = 4\left(\frac{\theta^2}{12n} + \theta^2\right) - \theta^2 = \frac{\theta^2}{3n} + 3\theta^2$$

The standard error is given by:

$$\begin{aligned}
\text{se} = \sigma_{\hat{\theta}} = \sqrt{\text{Var}_\theta(\widehat{\theta})} &= \sqrt{\left(\frac{\theta^2}{3n} + 3\theta^2\right)} \\
&= \sqrt{\frac{9n\theta^2 + \theta^2}{3n}} \\
&= \sqrt{\frac{\theta^2(9n+1)}{3n}} \\
&= \theta\sqrt{\frac{9n+1}{3n}}
\end{aligned}$$

Hence we have that the estimator $\widehat{\theta} = 2\bar{X}_n$ is an unbiased of θ, i.e. $B = 0$.

As the estimator is unbiased, the mean squared error is equal to the variance $\text{MSE} = V = \frac{\theta^2}{3n} + 3\theta^2$.

And the standard error is given by $\text{se} = \sigma_{\hat{\theta}} = \theta\sqrt{(9n+1)/3n)}$

6) Chapter 9, problem 2 (a,b,c).

Let $X_1, \ldots, X_n \sim \text{Unif}(a, b)$ where a and b are unknown parameters with $a < b$.

a) Find the method of moments estimator for a and b.

b) Find the maximum likelihood estimator, $\widehat{a}$ and $\widehat{b}$.

c) Let $\tau = \int x dF(x)$, and find the maximum likelihood estimator of τ.

6a) The mean and variance of a random variable $X_i \sim \text{Unif}(a, b)$ where $a < b$ is given by:

$$\mathbb{E}[X_i] = \frac{a+b}{2}$$

$$\text{Var}[X_i] = \frac{(b-a)^2}{12}$$

To find the method of moments estimators $\widehat{a}_{MoM}$ and $\widehat{b}_{MoM}$, we equate the kth sample moment $m_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$ with the kth theoretical moment $\mathbb{E}[X_i^k]$ for $k = 1, 2$; i.e. the means and variances:

Equating the sample mean and population mean (1st moments):

$$m_1 = \mathbb{E}[X_i] \implies \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{a+b}{2}$$

Computing the 2nd theoretical moment:

$$\begin{aligned}
\mathbb{E}[X_i^2] &= \mathrm{Var}[X_i] + \mathbb{E}[X_i]^2 \\
&= \frac{(b-a)^2}{12} + \left(\frac{a+b}{2}\right)^2 \\
&= \frac{(b-a)^2}{12} + \frac{3(a+b)^2}{12} \\
&= \frac{3(a^2+2ab+b^2) + (b^2-2ab+a^2)}{12} \\
&= \frac{4a^2+4ab+4b^2}{12} \\
&= \frac{a^2+ab+b^2}{3}
\end{aligned}$$

Equating the 2nd sample and theoretical moment, we have that:

$$m_2 = \mathbb{E}[X_i^2] \implies \frac{1}{n}\sum_{i=1}^{n} X_i^2 = \frac{a^2+ab+b^2}{3}$$

Setting $b = 2\bar{X}_n - a$, and substituting into the 2nd moment equation one line above, we have that:

$$\begin{aligned}
3\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) &= a^2 + a(2\bar{X}_n - a) + (2\bar{X}_n - a)^2 \\
&= a^2 + 2a\bar{X}_n - a^2 + 4\bar{X}_n^2 - 4a\bar{X}_n + a^2 \\
&= a^2 - 2a\bar{X}_n + 4\bar{X}_n^2 \\
&= (a - \bar{X}_n)^2 + 3\bar{X}_n^2 \\
\implies \widehat{a}_{MoM} &= \bar{X}_n + \sqrt{3\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) - 3\left(\bar{X}_n\right)^2} \\
&= m_1 + \sqrt{3(m_2 - m_1^2)}
\end{aligned}$$

And we also have that:

$$\begin{aligned}
\widehat{b}_{MoM} &= 2\bar{X}_n - \widehat{a}_{MoM} \\
&= \bar{X}_n - \sqrt{3\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) - 3\left(\bar{X}_n\right)^2} \\
&= m_1 - \sqrt{3(m_2 - m_1^2)}
\end{aligned}$$

6b)

The PMF of $X_i \sim \text{Unif}(a, b)$ is given by:

$$f_{X_i}(x_i; a, b) = \begin{cases} \frac{1}{b-a} & \text{if} \quad x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$$

We first need to derive the likelihood function. Much of the reasoning for deriving the likelihood function under a uniform distribution is similar to question 3) above, so we lift from there.

As the data is IID, denoting the PMF of X_i parametrised by a, b as $f_{X_i}(x_i; a, b)$, the likelihood function $L(a, b)$ is given by:

$$L(a, b) = \prod_{i=1}^{n} f_{X_i}(x_i; a, b)$$

Then $L(a, b) = 0$ if any of the PMFs $f_{X_i}(X_i; a, b) = 0$.

That means we only need to consider the 1st and nth order statistics of the data, that is $X_{(1)} = \min\{X_1, \ldots, X_n\}$ and $X_{(n)} = \max\{X_1, \ldots, X_n\}$, and their values relative to the parameters.

$$L(a, b) = \begin{cases} 0 & \text{if} \quad a > X_{(1)} \\ 0 & \text{if} \quad b < X_{(n)} \\ 0 & \text{if} \quad b < X_{(n)} \text{and} a > X_{(1)} \\ \left(\frac{1}{b-a}\right)^n & \text{if} \quad b \geq X_{(n)} \text{and} a \leq X_{(1)} \end{cases}$$

Hence the likelihood function is given by:

$$L(a, b) = \mathbb{I}(b \geq X_{(n)})\mathbb{I}(a \leq X_{(1)}) \left(\frac{1}{b-a}\right)^n$$

The maximum likelihood estimators $\widehat{a}_{\text{ML}}$ and $\widehat{b}_{\text{ML}}$ are set so as to maximise the likelihood function $L(a, b)$.

In order to maximise $L(a, b)$, we need to minimise $(b - a)$ in such a way that $L(a, b) \neq 0$, i.e. with constraints that $a \leq X_{(1)}$ and $b \geq X_{(n)}$.

The smallest we can make $(b - a)$ with the above constraints satisfied is by setting:

$$\widehat{a}_{\text{MLE}} = X_{(1)}$$

$$\widehat{b}_{\text{MLE}} = X_{(n)}$$

6c)

τ, which is the population mean, is expressed as a function of the parameters a, b of the Uniform distribution:

$$\tau = \mathbb{E}[X_i] = \frac{a+b}{2} = g(a,b)$$

As the maximum likelihood estimator is equivariant, we have that $\widehat{\tau}_{\text{MLE}} = g(\widehat{a}_{\text{MLE}}, \widehat{b}_{\text{MLE}})$.

Hence we have that the maximum likelihood estimator of τ is given by:

$$\widehat{\tau}_{\text{MLE}} = \frac{\widehat{a}_{\text{MLE}} + \widehat{b}_{\text{MLE}}}{2} = \frac{X_{(1)} + X_{(n)}}{2}$$