Assignment 10.

N1. $X_1, \dots, X_6 \sim^{iid} U[0;\theta]$ for $\theta \in [1;2]$, denote $\vec{X} = [X_1 \cdots X_6]^T$

We need to find $T_\theta(\vec{X})$: $\quad Var\, T_\theta(\vec{X}) \le \frac{1}{10}$, $\quad Bias\, T_\theta(\vec{X}) = E\, T_\theta(\vec{X}) - \theta = 0$

Since we are approximating the right boundary, it is reasonable to consider $\max\{X_1, \dots, X_6\}$

Let $T_\theta(\vec{X}) = \max \vec{X}$

$P(T_\theta(\vec{X}) < z) = P(\max\{X_1, \dots, X_6\} < z) = P(X_1 < z, X_2 < z, \dots, X_6 < z) = \quad (1)$

Since X_i are parts of a sample, they are independent, therefore:

$(1) = \prod_{i=1}^{6} P(X_i < z) = \left(\frac{z}{\theta}\right)^6$, since all $X_i \sim U[0;\theta]$ (only for $z \in [0;\theta]$)

Then $f_{T_\theta(\vec{X})}(z) = (P(T_\theta(\vec{X}) < z))'_z = \begin{cases} 0, & z \notin [0;\theta] \\ \frac{6}{\theta}\left(\frac{z}{\theta}\right)^5, & z \in [0;\theta] \end{cases}$

$$E\,T_\theta(\vec{X}) = \int_{-\infty}^{+\infty} z \cdot f_{T_\theta(\vec{X})}(z) \cdot dz = \int_0^\theta z \cdot \frac{6}{\theta} \cdot \left(\frac{z}{\theta}\right)^5 dz = \frac{6}{\theta^6} \int_0^\theta z^6 dz = \frac{6}{\theta^6}\left(\frac{z^7}{7}\Big|_{z=0}^{\theta}\right) = \frac{6}{7}\theta$$

We need $E\,T_\theta(\vec{X}) = \theta$ for $T_\theta(\vec{X})$ to be unbiased $\Rightarrow T_\theta(\vec{X}) = \max \vec{X}$ doesn't work.

Let's correct $E\,T_\theta(\vec{X})$ by setting $T'_\theta(\vec{X}) = \frac{7}{6} \max \vec{X}$

$P(T'_\theta(\vec{X}) < z) = P(\frac{7}{6} T_\theta(\vec{X}) < z) = F_{T_\theta(\vec{X})}(\frac{6}{7}z)$, $\quad f_{T'_\theta(\vec{X})}(z) = (F_{T_\theta(\vec{X})}(\frac{6}{7}z))'_z = \frac{6}{7} f_{T_\theta(\vec{X})}(\frac{6}{7}z)$

$$E\,T'_\theta(\vec{X}) = \int_{-\infty}^{+\infty} z \cdot f_{T'_\theta(\vec{X})}(z) \cdot dz = \frac{7}{6} \int_{-\infty}^{+\infty} (\tfrac{6}{7}z) \cdot f_{T_\theta(\vec{X})}(\tfrac{6}{7}z)\, d(\tfrac{6}{7}z) = \frac{7}{6} \cdot \frac{6}{7}\theta = \theta$$

Therefore, $T'_\theta(\vec{X}) = \frac{7}{6} \max \vec{X}$ is unbiased.

$$\mathrm{Var}\,T'_\theta(\vec{X}) = E\,T'^2_\theta(\vec{X}) - (E\,T'_\theta(\vec{X}))^2$$

$$E\,T'^2_\theta(\vec{X}) = \int_{-\infty}^{+\infty} z^2 f_{T'_\theta(\vec{X})}(z) dz = \left(\frac{7}{6}\right)^2 \int_{-\infty}^{+\infty} (\tfrac{6}{7}z)^2 \cdot f_{T_\theta(\vec{X})}(\tfrac{6}{7}z)\, d(\tfrac{6}{7}z) = \left(\frac{7}{6}\right)^2 \int_0^\theta z_1^2 \cdot \frac{6}{\theta}\left(\frac{z_1}{\theta}\right)^5 dz_1 = \left(\frac{7}{6}\right)^2 \cdot \frac{6}{8}\theta^2$$

$$\mathrm{Var}\,T'_\theta(\vec{X}) = E\,T'^2_\theta(\vec{X}) - (E\,T'_\theta(\vec{X}))^2 = \frac{49}{48}\theta^2 - \theta^2 = \frac{1}{48}\theta^2 \leq \frac{1}{48} \cdot 2^2 < \frac{1}{10}$$

Therefore, $T'_\theta(\vec{X}) = \frac{7}{6}\max \vec{X}$ works.

<u>N2.</u> Let $x_1, \dots, x_n \in \mathbb{R}$ be known numbers such that $\sum_{i=1}^n x_i = 0$

Let $Y_i = \alpha + \beta x_i + \varepsilon_i$ for some unknown $\alpha, \beta \in \mathbb{R}$, where $\varepsilon_i \sim N(0; \sigma^2)$ for unknown σ^2

Not mentioned in the problem statement (likely, mistakenly) is the fact that all ε_i are independent.

Therefore, $Y_i \sim N(\alpha + \beta x_i; \sigma^2)$, all Y_i are independent.

$L_\alpha(\vec{y}, \alpha) = \prod_{i=1}^n f_{Y_i}(y_i, \alpha)$, $\quad L_\beta(\vec{y}, \beta) = \prod_{i=1}^n f_{Y_i}(y_i, \beta)$, $\quad L_{\sigma^2}(\vec{y}, \beta) = \prod_{i=1}^n f_{Y_i}(y_i, \sigma^2)$

The maximum likelihood estimators for α, β, σ^2 are:

$\hat{\alpha}(\vec{Y}) = \underset{\alpha}{\mathrm{argmax}}\, L_\alpha(\vec{Y}, \alpha)$, $\quad \hat{\beta}(\vec{Y}) = \underset{\beta}{\mathrm{argmax}}\, L_\beta(\vec{Y}, \beta)$, $\quad \sigma^2(\vec{Y}) = \underset{\sigma^2}{\mathrm{argmax}}\, L_{\sigma^2}(\vec{Y}, \sigma^2)$

$$L_\alpha(\vec{y}, \alpha) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i - \alpha - \beta x_i)^2}$$

$$\mathcal{L}_\alpha(\vec{y}, \alpha) = \ln L_\alpha(\vec{y}, \alpha) = -n \ln(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^n (y_i - \alpha - \beta x_i)^2$$

$(\mathcal{L}_\alpha(\vec{y}, \alpha))'_\alpha = \frac{1}{\sigma^2}\left(\sum_{i=1}^n (y_i - \alpha - \beta x_i)\right) = \frac{1}{\sigma^2}\left(\sum_{i=1}^n y_i - n\alpha - \beta \sum_{i=1}^n x_i\right) \Rightarrow (\mathcal{L}_\alpha(\vec{y}, \alpha))'_\alpha = 0 \Leftrightarrow \alpha = \frac{\sum_{i=1}^n y_i - \beta \sum_{i=1}^n x_i}{n}$

Therefore, $\hat{\alpha}(\vec{Y}) = \underset{\alpha}{\mathrm{argmax}}\, L_\alpha(\vec{Y}, \alpha) = \underset{\alpha}{\mathrm{argmax}}\, \mathcal{L}_\alpha(\vec{Y}, \alpha) = \overline{Y} - \frac{\beta}{n}\sum_{i=1}^n x_i = \overline{Y}$

Since all Y_i are independent, $\vec{Y}$ is multivariate normal, therefore:

$$\overline{Y} \sim N\left(\begin{bmatrix}\frac{1}{n} & \cdots & \frac{1}{n}\end{bmatrix} \times \begin{bmatrix}\alpha + \beta x_1 \\ \vdots \\ \alpha + \beta x_n\end{bmatrix}; \begin{bmatrix}\frac{1}{n} & \cdots & \frac{1}{n}\end{bmatrix} \times \begin{bmatrix}\sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2\end{bmatrix} \times \begin{bmatrix}\frac{1}{n} \\ \vdots \\ \frac{1}{n}\end{bmatrix}\right) = N\left(\alpha + \frac{\beta}{n}\sum_{i=1}^n x_i; \frac{\sigma^2}{n}\right) \Rightarrow \hat{\alpha}(\vec{Y}) \sim N\left(\alpha; \frac{\sigma^2}{n}\right)$$

$$L_\beta(\vec{y},\beta) = L_\alpha(\vec{y},\alpha) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\alpha-\beta x_i)^2}$$

$$l_\beta(\vec{y},\beta) = l_\alpha(\vec{y},\alpha) = -n\ln(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\alpha-\beta x_i)^2$$

$$(l_\beta(\vec{y},\beta))'_\beta = -\frac{1}{2\sigma^2}\cdot\sum_{i=1}^n 2(y_i-\alpha-\beta x_i)\cdot(-x_i) = \frac{1}{\sigma^2}\left(\sum_{i=1}^n y_i x_i - \alpha\sum_{i=1}^n x_i - \beta\sum_{i=1}^n x_i^2\right)$$

$$(l_\beta(\vec{y},\beta))'_\beta = 0 \Leftrightarrow \beta = \frac{\sum_{i=1}^n y_i x_i - \alpha\sum_{i=1}^n x_i}{\sum_{i=1}^n x_i^2}$$

$$\hat{\beta}(\vec{Y}) = \underset{\beta}{\arg\max}\, L_\beta(\vec{Y},\beta) = \frac{\sum_{i=1}^n Y_i x_i - \alpha\sum_{i=1}^n x_i}{\sum_{i=1}^n x_i^2} = \frac{\sum_{i=1}^n Y_i x_i}{\sum_{i=1}^n x_i^2}$$

$$L_{\sigma^2}(\vec{y},\sigma^2) = L_\alpha(\vec{y},\alpha) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\alpha-\beta x_i)^2}$$

$$L_{\sigma^2}(\vec{y},\sigma^2) = -n\ln(\sqrt{2\pi}) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\alpha-\beta x_i)^2$$

$$(L_{\sigma^2}(\vec{y},\sigma^2))'_{\sigma^2} = -\frac{n}{2\sigma^2} - \frac{1}{2}\sum_{i=1}^n (y_i-\alpha-\beta x_i)^2\cdot\left(-\frac{1}{(\sigma^2)^2}\right)$$

$$(L_{\sigma^2}(\vec{y},\sigma^2))'_{\sigma^2} = 0 \Leftrightarrow \frac{1}{\sigma^2}\left(\frac{1}{\sigma^2}\sum_{i=1}^n (y_i-\alpha-\beta x_i)^2 - n\right) = 0 \qquad (\sigma^2 \neq 0)$$

$$\Leftrightarrow \frac{1}{n}\sum_{i=1}^n (y_i-\alpha-\beta x_i)^2 = \sigma^2$$

Therefore, $\hat{\sigma^2}(\vec{Y}) = \frac{1}{n}\sum_{i=1}^n (Y_i-\alpha-\beta x_i)^2$

Since $\vec{Y}$ is multivariate normal, $\sum_{i=1}^n x_i Y_i \sim N\left([x_1 \cdots x_n]\times\begin{bmatrix}\alpha+\beta x_1\\ \vdots \\ \alpha+\beta x_n\end{bmatrix}; [x_1\cdots x_n]\times\begin{bmatrix}\sigma^2 & & 0\\ & \ddots & \\ 0 & & \sigma^2\end{bmatrix}\times\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\right)$

$$\sum_{i=1}^n x_i Y_i \sim N\left(\alpha\sum_{i=1}^n x_i + \beta\sum_{i=1}^n x_i^2; \sum_{i=1}^n \sigma^2 x_i^2\right)$$

$$\hat{\beta}(\vec{Y}) = \frac{1}{\sum_{i=1}^n x_i^2}\sum_{i=1}^n x_i Y_i - \frac{\alpha\sum_{i=1}^n x_i}{\sum_{i=1}^n x_i^2} \Rightarrow \hat{\beta}(\vec{Y}) \sim N\left(\beta; \frac{\sigma^2}{\sum_{i=1}^n x_i^2}\right)$$

$$\hat{\sigma^2}(\vec{Y}) = \frac{1}{n}\sum_{i=1}^n (Y_i-\alpha-\beta x_i)^2 = \frac{1}{n}\sum_{i=1}^n \varepsilon_i^2 = \frac{1}{n}\sum_{i=1}^n \sigma^2\eta_i^2, \text{ where } \eta_i = \frac{\varepsilon_i}{\sigma},\ \eta_i \sim N(0;1)$$

Therefore, $\hat{\sigma^2}(\vec{Y}) = \frac{\sigma^2}{n}\sum_{i=1}^n \eta_i^2 \Rightarrow \hat{\sigma^2}(\vec{Y}) \sim \chi^2(n)\cdot\frac{\sigma^2}{n}$

To fit the model, we will compute $\hat{\alpha}, \hat{\beta}, \hat{\sigma^2}$ of the actual realization of $\vec{Y}$:

$\vec{x} = [-3\ -2\ -1\ 0\ 1\ 2\ 3]$, $\vec{y} = [-5\ 0\ 3\ 4\ 3\ 0\ -5]$, $n=7$

$\sum_{i=1}^n x_i = 0$, $\sum_{i=1}^n y_i = 0$, $\sum_{i=1}^n x_i y_i = 0$, $\sum_{i=1}^n x_i^2 = 24$

$\hat{\alpha}(\vec{y}) = 0$, $\hat{\beta}(\vec{y}) = 0$, $\hat{\sigma^2}(\vec{y}) = \frac{1}{n}\sum_{i=1}^n (y_i - \hat{\alpha} - \hat{\beta}x_i)^2 = \frac{1}{n}\sum_{i=1}^n y_i^2 = \frac{84}{7} = 12$

The fact that $\hat{\sigma^2}$ is so large means that the linear model is inappropriate for this data.

N3. Let $X_1 \sim Po(e^\alpha)$, $X_2 \sim Po(e^{\alpha+\beta})$ be two independent RVs.

Denote $\vec{X} = \begin{bmatrix}X_1\\ X_2\end{bmatrix}$ the sample, then the likelihood:

$$L(\vec{x},\alpha,\beta) = f_{X_1}(x_1,\alpha,\beta)\cdot f_{X_2}(x_2,\alpha,\beta) = e^{-e^\alpha}\cdot\frac{e^{\alpha x_1}}{x_1!}\cdot e^{-e^{\alpha+\beta}}\cdot\frac{e^{(\alpha+\beta)x_2}}{x_2!}$$

$$l(\vec{x},\alpha,\beta) = \ln L(\vec{x},\alpha,\beta) = -e^\alpha + \alpha x_1 - \ln(x_1!) - e^{\alpha+\beta} + (\alpha+\beta)x_2 - \ln(x_2!)$$

a) $\frac{\partial l(\vec{x},\alpha,\beta)}{\partial\alpha} = -e^\alpha + x_1 - e^{\alpha+\beta} + x_2$, $\frac{\partial l(\vec{x},\alpha,\beta)}{\partial\beta} = -e^{\alpha+\beta} + x_2$

$\frac{\partial^2 l(\vec{x},\alpha,\beta)}{(\partial\alpha)^2} = -e^\alpha - e^{\alpha+\beta}$, $\frac{\partial^2 l(\vec{x},\alpha,\beta)}{\partial\alpha\partial\beta} = -e^{\alpha+\beta}$, $\frac{\partial^2 l(\vec{x},\alpha,\beta)}{(\partial\beta)^2} = -e^{\alpha+\beta}$

Our unknown parameters form a vector $\Rightarrow$ to get the estimators, we set the gradient to 0.

$$\begin{bmatrix} \frac{\partial L(\vec{x}, \alpha, \beta)}{\partial \alpha} \\ \frac{\partial L(\vec{x}, \alpha, \beta)}{\partial \beta} \end{bmatrix} = 0 \Leftrightarrow \begin{cases} -e^{\alpha} + x_1 - e^{\alpha+\beta} + x_2 = 0 \\ -e^{\alpha+\beta} + x_2 = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha = \ln x_1 \\ \beta = \ln x_2 - \ln x_1 \end{cases}$$

Therefore, $\hat{\beta}(\vec{X}) = \underset{\beta}{\operatorname{argmax}} L(\vec{X}, \alpha, \beta) = \ln \frac{X_2}{X_1}$.

<u>N4</u>. Let $x_1, \dots, x_n$ be known numbers such that $\sum_{i=1}^{n} x_i = 0$

Let $Y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i$ for some unknown $\beta_0, \beta_1, \beta_2$, where $\varepsilon_i \sim N(0; \sigma^2)$, σ^2 unknown, all ε_i are independent.

Thus, Y_i are also independent, $Y_i \sim N(\beta_0 + \beta_1 x_i + \beta_2 x_i^2; \sigma^2)$

$$L(\vec{y}, \beta_0, \beta_1, \beta_2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2}{2\sigma^2}} = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2}$$

$$\mathcal{L}(\vec{y}, \beta_0, \beta_1, \beta_2) = \ln L(\vec{y}, \beta_0, \beta_1, \beta_2) = -n \ln(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2$$

To get equations for estimators, we set the gradient to zero:

$$\begin{bmatrix} \frac{\partial \mathcal{L}(\vec{y}, \beta_0, \beta_1, \beta_2)}{\partial \beta_0} \\ \frac{\partial \mathcal{L}(\vec{y}, \beta_0, \beta_1, \beta_2)}{\partial \beta_1} \\ \frac{\partial \mathcal{L}(\vec{y}, \beta_0, \beta_1, \beta_2)}{\partial \beta_2} \end{bmatrix} = 0 \Leftrightarrow \begin{cases} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2) = 0 \\ \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2) x_i = 0 \\ \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2) x_i^2 = 0 \end{cases} \Leftrightarrow \begin{cases} \beta_0 = \bar{y} - \frac{\beta_2}{n} \sum_{i=1}^{n} x_i^2 \\ \beta_1 = \frac{\sum_{i=1}^{n} y_i x_i - \beta_2 \sum_{i=1}^{n} x_i^3}{\sum_{i=1}^{n} x_i^2} \\ \beta_2 = \frac{\sum_{i=1}^{n} y_i x_i^2 - \beta_0 \sum_{i=1}^{n} x_i^2 - \beta_1 \sum_{i=1}^{n} x_i^3}{\sum_{i=1}^{n} x_i^4} \end{cases}$$

To obtain estimators $\hat{\beta}_0(\vec{Y}), \hat{\beta}_1(\vec{Y}), \hat{\beta}_2(\vec{Y})$, one solves the system above and replaces y_i for Y_i.

б) the same problem was solved in N2, the estimator $\hat{\beta}(\vec{Y})$.

<u>N5</u>. Let $X_1, \dots, X_n \sim^{iid} N(\mu; \sigma^2)$, μ and σ^2 are unknown.

Let $\vec{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$, then $f_{\vec{X}}(x_1, \dots, x_n) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{1}{2\sigma^2}(\sum_{i=1}^{n} (x_i - \mu)^2)}$

а) $\mathcal{L}(\vec{x}, \mu, \sigma^2) = \ln L(\vec{x}, \mu, \sigma^2) = -n \ln(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2}(\sum_{i=1}^{n} (x_i - \mu)^2)$

б) We need to find a pair of statistics $\vec{T}(\vec{X}) = \begin{bmatrix} T_1(\vec{X}) \\ T_2(\vec{X}) \end{bmatrix}$ that is jointly sufficient for $\begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix}$

According to the factorization criterion for several parameters, $\vec{T}(\vec{X})$ is jointly sufficient for $\begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix}$ if:

- $f_{\vec{X}}(\vec{x}) = g(T_1(\vec{x}), T_2(\vec{x}), \mu, \sigma^2) \cdot h(\vec{x})$
- $g(T_1(\vec{x}), T_2(\vec{x}), \mu, \sigma^2)$ doesn't depend on $\vec{x}$ directly, only through the statistics
- $h(\vec{x})$ doesn't depend on μ, σ^2

$$f_{\vec{X}}(\vec{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}} \cdot \frac{1}{\sigma^n} \cdot e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} x_i^2} \cdot e^{-\frac{1}{2\sigma^2} \cdot 2\mu \sum_{i=1}^{n} x_i} \cdot e^{-\frac{1}{2\sigma^2} \cdot n\mu^2}$$

Set $h(\vec{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}}$ (the only factor that doesn't depend on μ, σ^2)

Then $g(\vec{x}) = \frac{1}{\sigma^n} \cdot e^{-\frac{n\mu^2}{2\sigma^2}} \cdot e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} x_i^2} \cdot e^{-\frac{\mu}{\sigma^2} \sum_{i=1}^{n} x_i}$, we can take $T_1(\vec{X}) = \sum_{i=1}^{n} X_i^2$, $T_2(\vec{X}) = \sum_{i=1}^{n} X_i$

c) $\frac{\partial L(\vec{x}, \mu, \sigma^2)}{\partial \mu} = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(x_i-\mu)\cdot(-1) = \frac{\sum_{i=1}^{n} x_i - n\mu}{\sigma^2}$ $(\sigma^2 \neq 0)$

$$\frac{\partial L(\vec{x}, \mu, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\cdot\left(-\frac{1}{(\sigma^2)^2}\right) = \frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_i-\mu)^2\cdot\frac{1}{\sigma^2} - n\right)$$

$$\begin{cases} \frac{\partial L(\vec{x}, \mu, \sigma^2)}{\partial \mu} = 0 \\ \frac{\partial L(\vec{x}, \mu, \sigma^2)}{\partial \sigma^2} = 0 \end{cases} \Leftrightarrow \begin{matrix} \mu = \bar{x} \\ \sum_{i=1}^{n}(x_i-\mu)^2\cdot\frac{1}{\sigma^2} = n \end{matrix} \Leftrightarrow \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \sigma^2$$

$\hat{\mu}(\vec{X}) = \underset{\mu}{\operatorname{argmax}}\, L(\vec{X}, \mu, \sigma^2) = \bar{X}$, $\quad \hat{\sigma}^2(\vec{X}) = \underset{\sigma^2}{\operatorname{argmax}}\, L(\vec{X}, \mu, \sigma^2) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})$

Since $X_1, \dots, X_n$ are independent, $\vec{X}$ is multivariate normal.

Therefore, $\hat{\mu}(\vec{X}) = \bar{X} \sim N\left([\frac{1}{n} \cdots \frac{1}{n}] \times \begin{bmatrix}\mu \\ \vdots \\ \mu\end{bmatrix}; [\frac{1}{n} \cdots \frac{1}{n}] \times \begin{bmatrix}\sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2\end{bmatrix} \times \begin{bmatrix}\frac{1}{n} \\ \vdots \\ \frac{1}{n}\end{bmatrix}\right) = N(\mu; \frac{\sigma^2}{n})$

Let $\eta \sim N(0;1)$ such that $\frac{\sigma}{\sqrt{n}}\eta + \mu = \hat{\mu}(\vec{X})$

We need to find ε such that $P(|\hat{\mu}(\vec{X}) - \mu| < \varepsilon) = 0,95$

$P(|\hat{\mu}(\vec{X}) - \mu| < \varepsilon) = P(\mu - \varepsilon < \frac{\sigma}{\sqrt{n}}\eta + \mu < \mu + \varepsilon) = P(-\frac{\varepsilon\sqrt{n}}{\sigma} < \eta < \frac{\varepsilon\sqrt{n}}{\sigma}) = 2\Phi(\frac{\varepsilon\sqrt{n}}{\sigma}) - 1$

Therefore, $2\Phi(\frac{\varepsilon\sqrt{n}}{\sigma}) - 1 = 0,95 \Leftrightarrow \frac{\varepsilon\sqrt{n}}{\sigma} = \Phi^{-1}(0,975) \Leftrightarrow \varepsilon = \frac{1,96\sigma}{\sqrt{n}}$

<u>N6.</u> $X_1, \dots, X_n \sim U[-\theta; 2\theta]$ for unknown $\theta > 0$

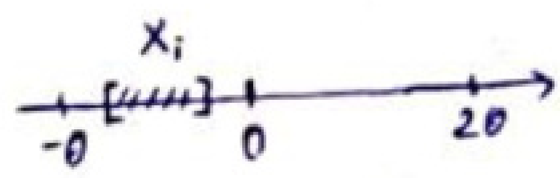
$L(\vec{X}, \theta) = \prod_{i=1}^{n} \frac{1}{3\theta}\cdot I_{x_i \in [-\theta; 2\theta]} = \frac{1}{(3\theta)^n}\cdot I_{\forall i:\, x_i \in [-\theta; 2\theta]} = \frac{1}{(3\theta)^n}\cdot I_{-\theta \leq \min\vec{x}} \cdot I_{\max\vec{x} \leq 2\theta} = \frac{1}{(3\theta)^n}\cdot I_{\theta \geq \max\{-\min\vec{x}, \frac{\max\vec{x}}{2}\}}$

$$\underset{\theta > 0}{\operatorname{argmax}}\, L(\vec{X}, \theta) = \underset{\substack{\theta \geq \max\{0, -\min\vec{X}, \frac{\max\vec{X}}{2}\} \\ \theta \neq 0}}{\operatorname{argmax}} \frac{1}{(3\theta)^n}$$

Since $X_i \sim U[-\theta; 2\theta]$, they can only take values in range $[-\theta; 2\theta]$

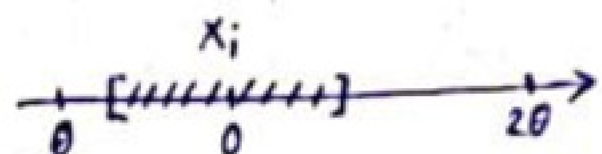
Let's consider the following cases:

I. $-\theta \leq \min\vec{X} \leq \max\vec{X} \leq 0$



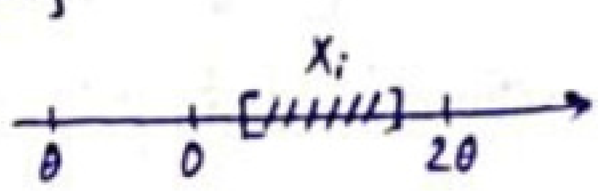
Then $\max\{0, -\min\vec{X}, \frac{\max\vec{X}}{2}\} = -\min\vec{X}$

II. $-\theta \leq \min\vec{X} \leq 0 \leq \max\vec{X} \leq 2\theta$



Then $\max\{0, -\min\vec{X}, \frac{\max\vec{X}}{2}\} = \max\{-\min\vec{X}, \frac{\max\vec{X}}{2}\}$

III. $0 \leq \min\vec{X} \leq \max\vec{X} \leq 2\theta$



Then $\max\{0, -\min\vec{X}, \frac{\max\vec{X}}{2}\} = \frac{\max\vec{X}}{2}$

Therefore, $\hat{\theta}(\vec{X}) = \begin{cases} \max\{-\min\vec{X}, \frac{\max\vec{X}}{2}\}, & \text{if } \sum_{i=1}^{n}|X_i| \neq 0 \\ \text{d.n.e.}, & \text{if } X_i = 0 \;\; \forall i \end{cases}$ (maximum likelihood estimator, other estimators still exist)

<u>N7.</u> Let X_1, X_2, X_3, X_4 be independent RVs denoting measurements of angles.

Let $\theta_1, \theta_2, \theta_3, \theta_4 : \sum_{i=1}^{4}\theta_i = 2\pi$ be actual values of angles, unknown, in radians to avoid units

$X_i = \theta_i + \varepsilon_i$, where $\varepsilon_i \sim N(0; \sigma^2)$, therefore, $X_i \sim N(\theta_i; \sigma^2)$

Let $S(\vec{X}, \vec{\theta}) = \sum_{i=1}^{4}(X_i - \theta_i)^2$, then $\hat{\vec{\theta}}(\vec{X}) = \underset{\theta}{\operatorname{argmin}}\, S(\vec{X}, \vec{\theta})$ - least squares estimators

Since θ_i are angles of a quadrilateral, we will minimize $S(\vec{X}, \vec{\theta})$ under constraint $\sum_{i=1}^{4}\theta_i = 2\pi$

Using Lagrange multipliers:

$$L(\vec{X}, \vec{\theta}, \lambda) = \sum_{i=1}^{4}(X_i-\theta_i)^2 - \lambda\left(\sum_{i=1}^{4}\theta_i - 2\pi\right)$$ // Lagrange, not likelihood

$$\frac{\partial L(\vec{X}, \vec{\theta}, \lambda)}{\partial \theta_i} = -2(X_i-\theta_i) - \lambda$$

$$\begin{cases} \frac{\partial L(\vec{X}, \vec{\theta}, \lambda)}{\partial \theta_i} = 0 \\ \sum_{i=1}^{4}\theta_i = 2\pi \end{cases} \Leftrightarrow \begin{cases} \theta_i - X_i = \frac{\lambda}{2} \\ \sum_{i=1}^{4}\theta_i = 2\pi \end{cases} \Leftrightarrow \begin{cases} \theta_i = \frac{\lambda}{2} + X_i \\ 2\pi = 2\lambda + \sum_{i=1}^{4} X_i \end{cases} \Leftrightarrow \begin{cases} \theta_i = \frac{\pi}{2} - \overline{X} + X_i \\ \lambda = \pi - \frac{\sum_{i=1}^{4} X_i}{2} \end{cases}$$

Therefore, the least squares estimators are $\hat{\theta}_i(\vec{X}) = \frac{\pi}{2} - \overline{X} + X_i$

b) We need to find $\hat{\sigma}^2(\vec{X})$: $E\hat{\sigma}^2(\vec{X}) = \sigma^2$

$$f_{\vec{X}}(\vec{x}) = \prod_{i=1}^{4} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x_i-\theta_i)^2} = \frac{1}{(\sqrt{2\pi}\sigma)^4} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{4}(x_i-\theta_i)^2}$$

$$L(\vec{x}, \sigma^2) = \ln f_{\vec{X}}(\vec{x}) = -2\ln(2\pi) - 2\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{4}(x_i-\theta_i)^2$$

$$\frac{\partial L(\vec{x}, \sigma^2)}{\partial \sigma^2} = -\frac{2}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{4}(x_i-\theta_i)^2$$

$$\frac{\partial L(\vec{x}, \sigma^2)}{\partial \sigma^2} = 0 \Leftrightarrow \frac{1}{\sigma^2}\left(-2 + \frac{1}{2\sigma^2}\sum_{i=1}^{4}(x_i-\theta_i)^2\right) = 0 \qquad (\sigma^2 \neq 0)$$

$$\Leftrightarrow \frac{1}{\sigma^2}\sum_{i=1}^{4}(x_i-\theta_i)^2 = 4 \Leftrightarrow \sigma^2 = \frac{1}{4}\left(\sum_{i=1}^{4}(x_i-\theta_i)^2\right)$$

Using the least squares estimators from (a), we may construct a MLE for σ^2:

$$\hat{\sigma}^2(\vec{X}) = \arg\max_{\sigma^2} L(\vec{X}, \sigma^2) = \frac{1}{4}\sum_{i=1}^{4}(X_i - \hat{\theta}_i(\vec{X}))^2 = \frac{1}{4}\sum_{i=1}^{4}\left(\overline{X} - \frac{\pi}{2}\right)^2 = \left(\overline{X} - \frac{\pi}{2}\right)^2$$

$X_1, \dots, X_4$ are independent $\Rightarrow \vec{X}$ is multivariate normal $\Rightarrow \overline{X}$ is normally distributed

$$\overline{X} \sim N\left(\left[\tfrac{1}{4} \cdots \tfrac{1}{4}\right] \times \begin{bmatrix}\theta_1 \\ \vdots \\ \theta_4\end{bmatrix}; \left[\tfrac{1}{4} \cdots \tfrac{1}{4}\right] \times \begin{bmatrix}\sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2\end{bmatrix} \times \begin{bmatrix}\frac{1}{4} \\ \vdots \\ \frac{1}{4}\end{bmatrix}\right) = N\left(\frac{\pi}{2}; \frac{\sigma^2}{4}\right)$$

$$E(\hat{\sigma}^2(\vec{X})) = E\left(\left(\overline{X} - \tfrac{\pi}{2}\right)^2\right) = E\overline{X}^2 - \pi E\overline{X} + \frac{\pi^2}{4} = \operatorname{Var}\overline{X} + (E\overline{X})^2 - \pi E\overline{X} + \frac{\pi^2}{4} = \frac{\sigma^2}{4}$$

To make $\hat{\sigma}^2(\vec{X})$ unbiased, we may multiply it by 4:

$$\hat{\sigma}^2(\vec{X}) = 4\left(\overline{X} - \tfrac{\pi}{2}\right)^2$$

c) In addition to the constraint $\sum_{i=1}^{4}\theta_i = 2\pi$ we have $\theta_1 = \theta_3$ and $\theta_2 = \theta_4$

Using Lagrange multipliers:

$$L(\vec{X}, \vec{\theta}, \lambda, \mu, \eta) = \sum_{i=1}^{4}(X_i-\theta_i)^2 - \lambda\left(\sum_{i=1}^{4}\theta_i - 2\pi\right) - \mu(\theta_1 - \theta_3) - \eta(\theta_2 - \theta_4)$$

$$\frac{\partial L(\vec{X}, \vec{\theta}, \lambda, \mu, \eta)}{\partial \theta_1} = -2(X_1-\theta_1) - \lambda - \mu \qquad \frac{\partial L(\vec{X}, \vec{\theta}, \lambda, \mu, \eta)}{\partial \theta_2} = -2(X_2-\theta_2) - \lambda - \eta$$

$$\frac{\partial L(\vec{X}, \vec{\theta}, \lambda, \mu, \eta)}{\partial \theta_3} = -2(X_3-\theta_3) - \lambda + \mu \qquad \frac{\partial L(\vec{X}, \vec{\theta}, \lambda, \mu, \eta)}{\partial \theta_4} = -2(X_4-\theta_4) - \lambda + \eta$$

$$\begin{cases} \frac{\partial L(\vec{X}, \vec{\theta}, \lambda, \mu, \eta)}{\partial \theta_1} = 0 \\ \frac{\partial L(\vec{X}, \vec{\theta}, \lambda, \mu, \eta)}{\partial \theta_2} = 0 \\ \frac{\partial L(\vec{X}, \vec{\theta}, \lambda, \mu, \eta)}{\partial \theta_3} = 0 \\ \frac{\partial L(\vec{X}, \vec{\theta}, \lambda, \mu, \eta)}{\partial \theta_4} = 0 \\ \sum_{i=1}^{4} \theta_i = 2\pi \\ \theta_1 = \theta_3 \\ \theta_2 = \theta_4 \end{cases} \Leftrightarrow \begin{cases} -2(X_1 - \theta_1) - \lambda - \mu = 0 \\ -2(X_2 - \theta_2) - \lambda - \eta = 0 \\ -2(X_3 - \theta_3) - \lambda + \mu = 0 \\ -2(X_4 - \theta_4) - \lambda + \eta = 0 \\ \theta_1 + \theta_2 + \theta_3 + \theta_4 = 2\pi \\ \theta_1 = \theta_3 \\ \theta_2 = \theta_4 \end{cases} \Leftrightarrow \begin{cases} \theta_1 = \frac{\lambda+\mu}{2} + X_1 \\ \theta_2 = \frac{\lambda+\eta}{2} + X_2 \\ \theta_3 = \frac{\lambda-\mu}{2} + X_3 \\ \theta_4 = \frac{\lambda-\eta}{2} + X_4 \\ \lambda = \pi - \frac{\sum_{i=1}^{4} X_i}{2} \\ \mu = X_3 - X_1 \\ \eta = X_4 - X_2 \end{cases}$$

Therefore, the estimators of θ_i are:

$$\hat{\theta}_1(\vec{X}) = \frac{\pi}{2} - \bar{X} + \frac{X_1 + X_3}{2} = \hat{\theta}_3(\vec{X}), \qquad \hat{\theta}_2(\vec{X}) = \frac{\pi}{2} - \bar{X} + \frac{X_2 + X_4}{2} = \hat{\theta}_4(\vec{X})$$

Recall from (B) that $\widehat{\sigma^2}(\vec{X}) = \frac{1}{4}\left(\sum_{i=1}^{4}(x_i - \hat{\theta}_i(\vec{X}))^2\right)$

Using the new estimators:

$$\widehat{\sigma^2}(\vec{X}) = \frac{1}{4}\left(\left(\frac{X_1 - X_3}{2} - \frac{\pi}{2} + \bar{X}\right)^2 + \left(\frac{X_2 - X_4}{2} - \frac{\pi}{2} + \bar{X}\right)^2 + \left(\frac{X_3 - X_1}{2} - \frac{\pi}{2} + \bar{X}\right)^2 + \left(\frac{X_4 - X_2}{2} - \frac{\pi}{2} + \bar{X}\right)^2\right) =$$

$$= \frac{1}{4}\left(\frac{(X_1 - X_3)^2}{4} + 2\,\frac{X_1 - X_3}{2}\left(\bar{X} - \frac{\pi}{2}\right) + \left(\bar{X} - \frac{\pi}{2}\right)^2 + \ldots\right) =$$

$$= \frac{1}{4}\left(\frac{X_1^2 - 2X_1X_3 + X_3^2 + X_2^2 - 2X_2X_4 + X_4^2}{2} + 4\left(\bar{X} - \frac{\pi}{2}\right)^2\right) =$$

$$= \frac{1}{8}\sum_{i=1}^{4} X_i^2 - \frac{X_1X_3}{4} - \frac{X_2X_4}{4} + \bar{X}^2 - \pi\bar{X} + \frac{\pi^2}{4}$$

$$E(\widehat{\sigma^2}(\vec{X})) = \frac{1}{8}\sum_{i=1}^{4} EX_i^2 - \frac{1}{4}EX_1 \cdot EX_3 - \frac{1}{4}EX_2 \cdot EX_4 + E\bar{X}^2 - \pi E\bar{X} + \frac{\pi^2}{4} =$$

$$= \frac{1}{8}\sum_{i=1}^{4}(\sigma^2 + \theta_i^2) - \frac{1}{4}\theta_1\theta_3 - \frac{1}{4}\theta_2\theta_4 + \left(\frac{\sigma^2}{4} + \frac{\pi^2}{4}\right) - \frac{\pi^2}{2} + \frac{\pi^2}{4} = \frac{3\sigma^2}{4}$$

To make $\widehat{\sigma^2}(\vec{X})$ unbiased, we may multiply it by $\frac{4}{3}$:

$$\widehat{\sigma^2}(\vec{X}) = \frac{4}{3}\left(\frac{1}{8}\sum_{i=1}^{4} X_i^2 - \frac{X_1X_3}{4} - \frac{X_2X_4}{4} + \bar{X}^2 - \pi\bar{X} + \frac{\pi^2}{4}\right)$$

N8. $X_1, \ldots, X_n \sim^{i.i.d.} N(\mu; \sigma^2)$, $\quad \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$, $\quad V = \sum_{i=1}^{n}(X_i - \bar{X})^2$

Since X_i are independent, $\vec{X}$ is multivariate normal.

Let C_i be a $2 \times n$ matrix such that: // $\begin{array}{c} \quad\quad\quad\quad\ \ i\text{-th} \\ \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & \frac{n-1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \end{bmatrix} \end{array}$

$C_i[1][j] = 1$ for any j, $\qquad C_i[2][i] = \frac{n-1}{n}$, $\quad C_i[2][j] = -\frac{1}{n}$ for $j \neq i$

Therefore, $\begin{bmatrix} \sum X_i \\ X_i - \bar{X} \end{bmatrix} = C_i \times \vec{X}$ for each i, $\begin{bmatrix} \sum X_i \\ X_i - \bar{X} \end{bmatrix}$ is also multivariate normal

$Cov\left(\sum_{i=1}^{n} X_i, X_i - \bar{X}\right) = 0$ (obtained from the covariance matrix $C_i \times \begin{bmatrix} \sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2 \end{bmatrix} \times C_i^T$)

For components of a multivariate normal vector, $Cov = 0 \Leftrightarrow$ independence $\Rightarrow$

$\Rightarrow \sum_{i=1}^{n} X_i$ and $X_i - \bar{X}$ are independent $\forall i$.

Let $f(\vec{z}) = \frac{1}{n} z_1$, $\quad g(\vec{z}) = \sum_{i=1}^{n} z_i^2$

Let $\vec{u} = \begin{bmatrix} \Sigma x_i \\ \vdots \\ \Sigma x_i \end{bmatrix}$, $\vec{v} = \begin{bmatrix} x_1 - \bar{x} \\ \vdots \\ x_n - \bar{x} \end{bmatrix}$, then $\bar{x} = f(\vec{u})$, $\sum_{i=1}^{n} (x_i - \bar{x})^2 = g(\vec{v})$

Components of $\vec{u}$ and $\vec{v}$ are independent $\Rightarrow$ the vectors themselves are independent

According to the property that functions of independent random variables or vectors are independent, we have that $\bar{x}$ and $\sum_{i=1}^{n} (X_i - \bar{x})^2$ are independent.

Since $\vec{X}$ is multivariate normal, $\bar{X} \sim N([\frac{1}{n} \cdots \frac{1}{n}] \times \begin{bmatrix} \mu \\ \vdots \\ \mu \end{bmatrix}; [\frac{1}{n} \cdots \frac{1}{n}] \begin{bmatrix} \sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2 \end{bmatrix} \begin{bmatrix} \frac{1}{n} \\ \vdots \\ \frac{1}{n} \end{bmatrix}) = N(\mu; \frac{\sigma^2}{n})$

For the same reason, $X_i - \bar{X} \sim N(C_{2i} \times \begin{bmatrix} \mu \\ \vdots \\ \mu \end{bmatrix}; C_{2i} \times \begin{bmatrix} \sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2 \end{bmatrix} \times C_{2i}^T) = N(0; \frac{n-1}{n}\sigma^2)$,

where C_{2i} is the second row of the aforementioned matrix C_i

Let $\eta_i \sim N(0;1)$ such that $X_i - \bar{X} = \sqrt{\frac{n-1}{n}}\sigma \cdot \eta_i$, then $(X_i - \bar{X})^2 = \frac{n-1}{n}\sigma^2 \cdot \eta_i^2$

Therefore, $\sum_{i=1}^{n} (X_i - \bar{X})^2 = \frac{n-1}{n}\sigma^2 \cdot \sum_{i=1}^{n} \eta_i^2 \sim \frac{n-1}{n}\sigma^2 \cdot \chi_n^2$

N9 skipped: we didn't cover confidence intervals.

<u>N10.</u> Let $Y_{ij} = Ax_i + B + \varepsilon_{ij}$, where A, B are unknowns, $\varepsilon_{ij} \overset{iid}{\sim} N(0;1)$

Let $N = \sum_{i=1}^{k} n_i$, $\bar{x} = \frac{1}{N}\sum_{i=1}^{k} n_i x_i$, $u_i = x_i - \bar{x}$

Then we can use u_i instead of x_i, and we now have $\sum_{i=1}^{k} n_i u_i = 0 = \sum_{i,j} u_i$

$Y_{ij} = A(u_i + \bar{x}) + B + \varepsilon_{ij} = Au_i + (\bar{x} + B) + \varepsilon_{ij}$, denote $(\bar{x} + B)$ as B^*

$Y_{ij} \sim N(Au_i + B^*; 1)$, Y_{ij} are independent as are $\varepsilon_{ij} \Rightarrow \vec{Y}$ is multivariate normal,

where $\vec{Y}$ is a vector of Y_{ij} for $i \in [1;k]$ for $j \in [1; n_i]$

$$L(\vec{y}, A, B^*) = \frac{1}{(\sqrt{2\pi})^N} e^{-\frac{1}{2}\sum_{ij}(y_{ij} - Au_i - B^*)^2}$$

$$l(\vec{y}, A, B^*) = \ln L(\vec{y}, A, B^*) = -N\ln(\sqrt{2\pi}) - \frac{1}{2}\sum_{i,j}(y_{ij} - Au_i - B^*)^2$$

MLE: $\hat{A}(\vec{Y}) = \underset{A}{\operatorname{argmax}}\, l(\vec{Y}, A, B^*)$, $\hat{B}(\vec{Y}) = \underset{B^*}{\operatorname{argmax}}\, l(\vec{Y}, A, B^*) - \bar{x}$

$$\frac{\partial l(\vec{y}, A, B^*)}{\partial A} = -\frac{1}{2}\sum_{i,j} 2(y_{ij} - Au_i - B^*) \cdot (-u_i), \qquad \frac{\partial l(\vec{y}, A, B^*)}{\partial B} = -\frac{1}{2}\sum_{i,j} 2(y_{ij} - Au_i - B^*) \cdot (-1)$$

$$\begin{cases} \frac{\partial l(\vec{y}, A, B^*)}{\partial A} = 0 \\ \frac{\partial l(\vec{y}, A, B^*)}{\partial B} = 0 \end{cases} \Leftrightarrow \begin{cases} \sum_{ij} y_{ij} u_i - A\sum_{i,j} u_i^2 - B^* \cancelto{0}{\sum_{ij} u_i} = 0 \\ \sum_{i,j} y_{ij} - A \cancelto{0}{\sum_{i,j} u_i} - NB^* = 0 \end{cases} \Leftrightarrow \begin{cases} A = \frac{\sum_{ij} y_{ij} u_i}{\sum_{ij} u_i^2} \\ B^* = \frac{\sum_{ij} y_{ij}}{N} \end{cases}$$

$$S(\vec{y}, A, B^*) = \sum_{i,j} (y_{ij} - Au_i - B^*)^2$$

LSE: $\hat{A}(\vec{Y}) = \underset{A}{\operatorname{argmin}}\, S(\vec{Y}, A, B^*)$, $\hat{B}(\vec{Y}) = \underset{B^*}{\operatorname{argmin}}\, S(\vec{Y}, A, B^*) - \bar{x}$

$$\frac{\partial S(\vec{y}, A, B^*)}{\partial A} = \sum_{i,j} 2(y_{ij} - Au_i - B^*) \cdot (-u_i), \qquad \frac{\partial S(\vec{y}, A, B)}{\partial B} = \sum_{i,j} 2(y_{ij} - Au_i - B^*) \cdot (-1)$$

One can observe that the gradients of l and S are in the same point $\Rightarrow$ MLE = LSE

$$\hat{A}(\vec{Y}) = \frac{\sum_{ij} Y_{ij} \cdot u_i}{\sum_{ij} u_i^2} = \sum_{i,j} \frac{u_i}{\sum_{ij} u_i^2} \cdot Y_{ij}, \qquad \hat{B}^*(\vec{Y}) = \sum_{ij} \frac{1}{N} Y_{ij}, \qquad \hat{B}(\vec{Y}) = \hat{B}^*(\vec{Y}) - \bar{x}$$

Since $\vec{Y}$ is multivariate normal and $\hat{A}(\vec{Y})$ and $\hat{B}^*(\vec{Y})$ are linear combinations of its components,

$$\begin{bmatrix} \hat{A}(\vec{Y}) \\ \hat{B}^*(\vec{Y}) \end{bmatrix} \sim N\left(C \times \begin{bmatrix} Au_i + B^* \end{bmatrix}; \ C \times I \times C^T\right), \text{ where } C = \begin{bmatrix} \frac{u_i}{\sum_{i,j} u_i^2} & \cdots \\ \frac{1}{N} & \cdots \end{bmatrix} \text{ for } i \in [1;k] \text{ for } j \in [1;n_i]$$

$$= N\left(\begin{bmatrix} A \\ B^* \end{bmatrix}; \begin{bmatrix} \frac{1}{\sum_{i,j} u_i^2} & 0 \\ 0 & \frac{1}{N} \end{bmatrix}\right) \Rightarrow \hat{A}(\vec{Y}) \text{ and } \hat{B}^*(\vec{Y}) \text{ are independent.}$$

$$P(\hat{B}(\vec{Y}) < z) = P(\hat{B}^*(\vec{Y}) < z + \bar{x}) \Rightarrow f_{\hat{B}(\vec{Y})}(z) = f_{\hat{B}^*(\vec{Y})}(z + \bar{x})$$

Since subtracting a constant doesn't break independence,

$$f_{\hat{A},\hat{B}}(w, z) = f_{\hat{A}}(w) \cdot f_{\hat{B}}(z) = f_{\hat{A}}(w) \cdot f_{\hat{B}^*}(z + \bar{x}) \Rightarrow \begin{bmatrix} \hat{A}(\vec{Y}) \\ \hat{B}(\vec{Y}) \end{bmatrix} \sim N\left(\begin{bmatrix} A \\ B \end{bmatrix}; \begin{bmatrix} \frac{1}{\sum_{i,j} u_i^2} & 0 \\ 0 & \frac{1}{N} \end{bmatrix}\right)$$

<u>N11.</u>a) Let $f(x;\theta) = I_{x \geq \theta} e^{\theta - x}$ be the PDF of a sample X_1, X_2, X_3

Let $t(\vec{X}) = \min \vec{X}$ be a statistics.

According to the factorization criterion, $t(\vec{X})$ is sufficient for θ if $\exists g, h$:

$$f_{\vec{X}}(\vec{x};\theta) = g(t(\vec{x}), \theta) \cdot h(\vec{x})$$

Since the sample consists of independent RVs by definition,

$$f_{\vec{X}}(\vec{x};\theta) = \prod_{i=1}^{3} f(x_i;\theta) = I_{x_1 \geq 0} \cdot I_{x_2 \geq 0} \cdot I_{x_3 \geq 0} \cdot e^{3\theta - x_1 - x_2 - x_3} = I_{\min \vec{x} \geq 0} \cdot e^{3\theta - x_1 - x_2 - x_3}$$

Then we set $g(t(\vec{x}), \theta) = I_{\min \vec{x} \geq 0} \cdot e^{3\theta}$, $h(\vec{x}) = e^{-x_1 - x_2 - x_3} \Rightarrow t(\vec{x})$ is sufficient for θ

B) Let $Y_1 = \min\{X_1, X_2, X_3\}$, $\xi = Y_1 - \theta$

$$F_\xi(z) = P(Y_1 < z + \theta) = 1 - P(\min\{X_1, X_2, X_3\} > z + \theta) = 1 - P(X_1 > z + \theta, X_2 > z + \theta, X_3 > z + \theta)$$

Since X_i are independent:

$$F_\xi(z) = 1 - P(X_1 > z + \theta) P(X_2 > z + \theta) P(X_3 > z + \theta)$$

Since X_i are identically distributed:

$$F_\xi(z) = 1 - (1 - P(X_1 < z + \theta))^3 = 3F_{X_1}(z + \theta) - 3(F_{X_1}(z + \theta))^2 + (F_{X_1}(z + \theta))^3$$

$$f_\xi(z) = (F_\xi(z))'_z = (F_\xi(z))'_{z+\theta} = 3f_{X_1}(z + \theta) - 6F_{X_1}(z + \theta) \cdot f_{X_1}(z + \theta) + 3(F_{X_1}(z + \theta))^2 \cdot f_{X_1}(z + \theta) =$$

$$= f_{X_1}(z + \theta) \cdot (3 - 6F_{X_1}(z + \theta) + 3(F_{X_1}(z + \theta))^2) = I_{z > 0} \cdot e^{-z}(3 - 6F_{X_1}(z + \theta) + 3(F_{X_1}(z + \theta))^2)$$

$$F_{X_1}(z) = \int_{-\infty}^{z} f_{X_1}(t)dt = I_{z > \theta} \cdot \int_{\theta}^{z} e^{\theta - t}dt = I_{z > \theta} \cdot e^{\theta} \cdot (-1) \cdot \int_{\theta}^{z} -e^{-t}dt = I_{z > \theta} \cdot e^{\theta} \cdot (e^{-\theta} - e^{-z}) = I_{z > \theta} \cdot (1 - e^{\theta - z})$$

Then $f_\xi(z) = I_{z > 0} \cdot e^{-z}(3 - 6 \cdot (I_{z > 0} \cdot (1 - e^{-z})) + 3 \cdot (I_{z > 0}^2 \cdot (1 - 2e^{-z} + e^{-2z}))) =$

$$= I_{z > 0} \cdot e^{-z}(3 - 6 + 6e^{-z} + 3 - 6e^{-z} + 3e^{-2z}) = I_{z > 0} \cdot 3e^{-3z}$$

Thus we've shown that $\xi \sim Exp(3)$

Let $\hat{\theta}_1(\vec{X}) = X_3 - 1$, $\hat{\theta}_2(\vec{X}) = \min \vec{X}$, $\hat{\theta}_3(\vec{X}) = \bar{X} - 1$ be estimators for θ.

$$EX_i = \int_{-\infty}^{+\infty} x \cdot I_{x \geq \theta} e^{\theta - x}dx = e^{\theta} \cdot \int_{\theta}^{+\infty} xe^{-x}dx = -e^{\theta} \cdot \left(\int_{\theta}^{+\infty} x \cdot (-e^{-x})dx\right) = -e^{\theta} \cdot \left(xe^{-x}\Big|_{x=\theta}^{+\infty} - \int_{\theta}^{+\infty} e^{-x}dx\right) = \theta + 1$$

$$EX_i^2 = \int_{-\infty}^{+\infty} x^2 \cdot I_{x \geq \theta} e^{\theta - x}dx = -e^{\theta} \cdot \int_{\theta}^{+\infty} x^2 \cdot (-e^{-x})dx = -e^{\theta} \cdot \left(x^2e^{-x}\Big|_{x=\theta}^{+\infty} - 2\int_{\theta}^{+\infty} xe^{-x}dx\right) = \theta^2 + 2\theta + 2$$

$E\hat{\theta}_1(\vec{X}) = EX_3 - 1 = \theta \quad \Rightarrow \hat{\theta}_1(\vec{X})$ is unbiased

$E\hat{\theta}_2(\vec{X}) = E(Y_1 - \theta + \theta) = E(Y_1 - \theta) + \theta = \frac{1}{3} + \theta \quad \Rightarrow \hat{\theta}_2(\vec{X})$ is biased, unsuitable

$E\hat{\theta}_3(\vec{X}) = \sum_{i=1}^{3} \frac{1}{3} EX_i - 1 = \theta \quad \Rightarrow \hat{\theta}_3(\vec{X})$ is unbiased

$Var\ \hat{\theta}_1(\vec{X}) = Var\ X_3 = EX_3^2 - (EX_3)^2 = 1$

$Var\ \hat{\theta}_3(\vec{X}) = \sum_{i=1}^{3} \frac{1}{9} Var\ X_i = \frac{1}{3}$

$\Rightarrow \hat{\theta}_3(\vec{X})$ has a lesser mean square error

$\hat{\theta}_3(\vec{X})$ is the best candidate among the three.

N12. (a) done in N2, (b) skipped.