CSE 544

PROBABILITY AND STATISTICS FOR DATA SCIENCE

ASSIGNMENT 4

Team members:

Name	**SBU ID**
Venkata Satwik Chilukuri	113263262
Akshay Somayaji	113322316
Shardul Churi	114403283

1)

a) Given,

$(x_1, x_2, \ldots x_n) \approx \text{Gamma}(x, y)$

Mean $(\bar{x}) = x \cdot y$, variance $(\bar{s}^2) = x \cdot y^2$

To find,

$\hat{x}_{MME}$, $\hat{y}_{MME}$.

No of parameters to find 2

so,

$$\hat{\alpha}_1 = \hat{E}[x_i] = \frac{1}{n} \sum_{i=1}^{n} (x_i)$$

(plug in estimate)

$$\hat{\alpha}_2 = \hat{E}[x_i^2] = \frac{1}{n} \sum_{i=1}^{n} (x_i^2)$$

Now,

for the Moments,

1^{st} Moment $\alpha_1 = E[\text{Gamma}(x, y)]$

$= x \cdot y$ (Given)

$$\alpha_2 = \left[E\left[(\text{Gamma}(x,y))^2\right\}\right.$$

Now, we know $E(x^2) = Var(x) + (E(x))^2$

$$\Rightarrow \alpha_2 = var(\text{Gamma}(x,y)) + (E(\text{Gamma}(x,y)))^2$$

$$= x \cdot y^2 + (x \cdot y)^2$$

i.e

$$\alpha_1 = x \cdot y, \quad \alpha_2 = x \cdot y^2 + (x \cdot y)^2$$

$$\hat{\alpha}_1 = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i, \quad \hat{\alpha}_2 = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i^2$$

Now,

$$\hat{\alpha}_1 = \alpha_1 \qquad \hat{\alpha}_2 = \alpha_2$$

$$\Rightarrow \frac{1}{n} \cdot \sum_{i=1}^{n} x_i = x \cdot y \quad — (1) \qquad \Rightarrow \frac{1}{n} \cdot \sum_{i=1}^{n} x_i^2 = x \cdot y^2 + (x \cdot y)^2 \quad — (2)$$

substituting (1) in (2)

$$\Rightarrow \frac{1}{n} \cdot \sum_{i=1}^{n} x_i^2 = \hat{y}_{MME} \cdot \frac{1}{n} \sum_{i=1}^{n} x_i + \left(\frac{\sum x_i}{n}\right)^2$$

$$\Rightarrow \hat{y}_{MME} = \frac{\frac{1}{n}\cdot\sum_{i=1}^{n} x_i^2 - \left(\frac{\sum_{i=1}^{n} x_i}{n}\right)^2}{\frac{1}{n}\cdot\sum x_i}$$

let $\frac{\sum_{i=1}^{n} x_i}{n} = \bar{x} \rightarrow$ mean.

$$\Rightarrow \hat{y}_{MME} = \frac{\frac{1}{n}\cdot\sum_{i=1}^{n} x_i^2 - (\bar{x})^2}{\bar{x}}$$

$$= \frac{\frac{1}{n}\left[\sum_{i=1}^{n} x_i^2 - n\cdot\bar{x}^2\right]}{\bar{x}}$$

$$\Rightarrow \hat{y}_{MME} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n\bar{x}}$$

we know Mean $= x\cdot y$

$$\Rightarrow x = \frac{\bar{x}}{y}$$

$$\Rightarrow \hat{x}_{MME} = \frac{n\cdot\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

So,

$$\hat{y}_{MME} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n\cdot\bar{x}}, \quad \hat{x}_{MME} = \frac{n\cdot\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

where, $\bar{x}$ is sample Mean

(b)

Given,

$(x_1, x_2, \ldots x_n) \sim$ uniform (a, b)

Sample Mean $(\bar{x}) = \frac{\sum x_i}{n}$ — ①, variance $(s^2) = \left(\frac{\sum x_i^2}{n}\right) - \bar{x}^2$ — ②

To find,

$\hat{a}_{MME}$, $\hat{b}_{MME}$

No of parameters to find 2,

So, plug-in values

$\hat{\alpha}_1 = E[x_i] = \sum_{i=1}^{n} x_i \times \frac{1}{n}$

$\hat{\alpha}_2 = E[x_i^2] = \sum_{i=1}^{n} x_i^2 \times \frac{1}{n}$

Now for the Moments,

1st Moment is Mean

$\alpha_1 = E[\text{unif}(a, b)]$

we know that for a Unif (a, b),

its Mean is $\frac{a+b}{2}$

variance is $\frac{(b-a)^2}{12}$

So,

$$\alpha_1 = \frac{a+b}{2}$$

$$\alpha_2 = E\left[\text{unif}(a,b)^2\right]$$

$$= \text{var}\left(\text{unif}(a,b)\right) + \left(E\left(\text{unif}(a,b)\right)\right)^2$$

$$\left(\because E(x^2) = \text{var}(x) + (E(x))^2\right)$$

$$= \frac{(b-a)^2}{12} + (\alpha_1)^2$$

$$\Rightarrow \hat{\alpha}_1 = \frac{\sum_{i=1}^{n} x_i}{n}, \quad \alpha_1 = \frac{a+b}{2}$$

$$\Rightarrow \hat{\alpha}_2 = \frac{\sum_{i=1}^{n} x_i^2}{n}, \quad \alpha_2 = \frac{(b-a)^2}{12} + \alpha_1^2$$

Now,

$$\hat{\alpha}_1 = \alpha_1, \quad \hat{\alpha}_2 = \alpha_2$$

$$\Rightarrow \sum_{i=1}^{n} x_i \times \frac{1}{n} = \frac{a+b}{2}, \quad \frac{\sum_{i=1}^{n} x_i^2}{n} = \frac{(b-a)^2}{12} + \alpha_1^2$$

Given

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

so,

$$\frac{a+b}{2} = \bar{x} \quad , \quad \frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2 = \frac{(b-a)^2}{12}$$

$$\frac{a+b}{2} = \bar{x} \quad , \quad \overline{s^2} = \frac{(b-a)^2}{12} \quad \text{(Given variance)}$$

$$\Rightarrow 12 \cdot \overline{s^2} = (b-a)^2$$

$$\Rightarrow b-a = \sqrt{12 \cdot \overline{s^2}}$$

$$\Rightarrow b-a = 2\sqrt{3 \cdot \overline{s^2}}$$

$$\Rightarrow a+b = 2\bar{x} \quad , \quad b-a = 2\sqrt{3 \cdot \overline{s^2}}$$

By solving,

we set $\quad \hat{a}_{MME} = \bar{x} - \sqrt{3 \cdot \overline{s^2}}$

$$\hat{b}_{MME} = \bar{x} + \sqrt{3 \cdot \overline{s^2}}$$

Given,

$$(x_1, x_2, \ldots, x_n) \overset{iid}{\sim} \text{Exponential}(1/\beta)$$

pdf of an Exponential (λ) distribution is

$$f(x) = \lambda e^{-\lambda x}$$

so Here,

$$\text{pdf } f(x_i) = \lambda \cdot e^{-\lambda x_i} \quad — ①$$

Likelihood of the data,

$$L\left(\frac{1}{\beta}\right) = \prod_{i=1}^{n} f_x(x_i)$$

$$= \prod_{i=1}^{n} \frac{1}{\beta} \cdot e^{-\frac{1}{\beta} \cdot x_i} \quad (\text{from } ①)$$

$$= \left(\frac{1}{\beta} \cdot e^{-\frac{1}{\beta} \cdot x_1}\right) \times \left(\frac{1}{\beta} \cdot e^{-\frac{1}{\beta} \cdot x_2}\right) \ldots \text{ n times}$$

$$= \frac{1}{\beta^n} \cdot e^{-\frac{1}{\beta} \cdot \sum_{i=1}^{n} x_i}$$

To find the value of β which Maximises the value $\frac{d}{d\beta}\left(L\left(\frac{1}{\beta}\right)\right)$ should be 0

So,

$$\frac{d}{d\beta}\left(L\left(\frac{1}{\beta}\right)\right) = 0$$

$$\Rightarrow \frac{d}{d\beta}\left[\frac{1}{\beta^n} \cdot e^{-\frac{1}{\beta}\cdot\sum_{i=1}^{n} x_i}\right] = 0$$

using the law of partial derivatives

$$\Rightarrow \frac{d}{d\beta}\left(\frac{1}{\beta^n}\right) \cdot e^{-\frac{1}{\beta}\cdot\sum_{i=1}^{n} x_i} + \frac{1}{\beta^n} \cdot \frac{d}{d\beta}\left(e^{-\frac{1}{\beta}\cdot\sum_{i=1}^{n} x_i}\right) = 0$$

$$\Rightarrow -n \times \beta^{-n-1} \times e^{-\frac{1}{\beta}\cdot\sum_{i=1}^{n} x_i} + \frac{1}{\beta^n} \cdot e^{-\frac{1}{\beta}\cdot\sum_{i=1}^{n} x_i} \times \frac{d}{d\beta}\left(-\frac{1}{\beta}\cdot\sum_{i=1}^{n} x_i\right) = 0$$

$$\left(\because \frac{d}{dx}\left(\frac{1}{x^n}\right) = -n \times \frac{1}{x^{n+1}},\ \frac{d}{dx} e^{g(x)} = e^{g(x)} \cdot \frac{d}{dx}(g(x))\right)$$

$$\Rightarrow -n \times \frac{1}{\beta^{n+1}} \times \frac{1}{e^{\frac{1}{\beta}\cdot\sum_{i=1}^{n} x_i}} + \frac{1}{\beta^n} \cdot \frac{1}{e^{\frac{1}{\beta}\sum_{i=1}^{n} x_i}}\left[\sum x_i \times \frac{1}{\beta^2}\right] = 0$$

$$\Rightarrow \frac{\sum x_i}{\beta^n \times \beta^2 \times \cancel{e^{\frac{1}{\beta}\cdot\sum_{i=1}^{n} x_i}}} = \frac{+n}{\beta^{n+1} \times \cancel{e^{\frac{1}{\beta}\sum_{i=1}^{n} x_i}}}$$

$$\Rightarrow \frac{\sum_{i=1}^{n} x_i}{\beta} = n \qquad \boxed{\therefore \hat{\beta}_{MLE} = \frac{\sum_{i=1}^{n} x_i}{n}}$$

Now,

to prove MLE ($\hat{\beta}$) will converge to the unknown parameter β we should show that bias($\hat{\beta}$) or se($\hat{\beta}$) tends to 0 as $n \to \infty$

①

we know that for a plug in value $\hat{\theta}$

$$\text{bias}(\hat{\theta}) = E[\hat{\theta}] - \theta$$

so,

$$\text{bias}(\hat{\beta}) = E[\hat{\beta}] - \beta$$

$$= E\left[\frac{\sum_{i=1}^{n} x_i}{n}\right] - \beta$$

$$= \frac{1}{n} \cdot E\left[\sum_{i=1}^{n} x_i\right] - \beta$$

using linearity of expectation

$$= \frac{1}{n} \cdot \left[E[x_1] + E[x_2] + \ldots + E[x_n]\right] - \beta$$

as $x_1, x_2, \ldots x_n$ are iid their expectations are equal

$$= \frac{1}{n} \, n \cdot E[x_1] - \beta$$

$$= E[x_1] - \beta$$

as $x_1 \sim \text{Exponential}(1/\beta)$, Given Mean of

$$\text{Exponential}(\lambda) = \frac{1}{\lambda}$$

$$= \frac{1}{\left(\frac{1}{\beta}\right)} - \beta$$

$$= 0$$

$$\therefore \text{bias}(\hat{\beta}) = 0$$

② we know that for a plug-in value $\hat{\theta}$

$$se(\hat{\theta}) = \sqrt{\text{var}(\hat{\theta})}$$

so,

$$se(\hat{\beta}) = \sqrt{\text{var}(\hat{\beta})}$$

$$= \sqrt{\text{var}\left(\frac{1}{n} \sum_{i=1}^{n} x_i\right)}$$

$$= \sqrt{\frac{1}{n^2} \cdot \text{var}\left(\sum_{i=1}^{n} x_i\right)}$$

using linearity of variance

$$= \sqrt{\frac{1}{n^2} \left[\text{var}(x_1) + \text{var}(x_2) + \dots \text{var}(x_n)\right]}$$

as $x_1, x_2, \dots x_n$ are iid, their variances are equal

$$= \sqrt{\frac{1}{n^2} \times n \cdot var(x_1)}$$

$$= \sqrt{\frac{1}{n} \cdot var(x_1)}$$

as $x_1 \sim$ exponential $\left(\frac{1}{\beta}\right)$, Given variance of

exponential $(\lambda) = \frac{1}{\lambda^2}$

$$= \sqrt{\frac{1}{n} \cdot \frac{1}{\frac{1}{\beta^2}}}$$

$$= \sqrt{\frac{\beta^2}{n}}$$

So,

as $n \to \infty$, $se(\hat{\beta}) = 0$

Both, bias$(\hat{\beta})$ and $se(\hat{\beta})$ tends to 0 as $n \to \infty$

So we can say that MLE $(\hat{\beta})$ will converge to the unknown parameter β.

3a) Given, $\{x_1, x_2 \dots x_n\} \sim$ Poisson (λ). Find $\hat{\lambda}_{MLE}$

Now, pmf for Poisson distⁿ is given by

$$P(X=x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$\Rightarrow L(\lambda) = \prod_{i=1}^{n} p_X(x_i) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!}$$

$$= \frac{e^{-n\lambda} \cdot \lambda^{\Sigma x_i}}{\prod_{i=1}^{n} x_i!}$$

Taking ln on both sides,

$$\ln(L(\lambda)) = \ln(e^{-n\lambda} \cdot \lambda^{\Sigma x_i}) - \ln(\prod_{i=1} x_i!)$$

$$= \Sigma x_i \cdot \ln(\lambda) - n\lambda - \sum_{i=1}^{n} \ln(x_i!)$$

$(\because \ln(a \cdot b) = \ln(a) + \ln(b))$

$$\therefore \ell(\lambda) = \Sigma x_i \cdot \ln(\lambda) - (n\lambda + \Sigma \ln(x_i!))$$

→ Now, to find maxima, $\frac{d}{d\lambda}(\ell(\lambda)) = 0$

$$\Rightarrow \frac{d}{d\lambda}(\ell(\lambda)) = 0 = \frac{d}{d\lambda}\left(\Sigma x_i \cdot \ln(\lambda) - n\lambda - \Sigma \ln(x_i!)\right)$$

$$\Rightarrow \frac{\Sigma x_i}{\lambda} - n - \underbrace{\frac{d}{d\lambda}(\Sigma \ln(x_i!))}_{\text{constant}} = 0$$

$$\Rightarrow \frac{\Sigma x_i}{\lambda} - n - 0 = 0$$

$$\Rightarrow \boxed{\hat{\lambda}_{MLE} = \frac{\sum_{i=1}^{n} x_i}{n}}$$

3(b)

Given,

$$x_1, x_2, x_3, \ldots, x_n \overset{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$

we know that,

$$\text{Sample variance uncorrected} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

where,

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

likelihood of the data;

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} f_x(x_i)$$

$$= \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \times e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}} \times e^{-\frac{(x_1-\mu)^2}{2\sigma^2}}\right) \times \left(\frac{1}{\sigma\sqrt{2\pi}} \times e^{-\frac{(x_2-\mu)^2}{2\sigma^2}}\right)$$

$$\times \cdots \left(\frac{1}{\sigma\sqrt{2\pi}} \times e^{-\frac{(x_n-\mu)^2}{2\sigma^2}}\right)$$

$$= \frac{1}{(\sigma\sqrt{2\pi})^n} \times e^{-\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}}$$

Applying log on both sides

$$\Rightarrow \ln(L(\mu, \sigma^2)) = \ln\left(\frac{1}{(\sigma\sqrt{2\pi})^n} \times e^{-\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}}\right)$$

$$\Rightarrow l(\mu, \sigma^2) = \ln\left((\sigma\sqrt{2\pi})^{-n}\right) + \ln\left(e^{\frac{-1}{2\sigma^2} \cdot \sum_{i=1}^{n}(x_i-\mu)^2}\right)$$

$$\Rightarrow l(\mu, \sigma^2) = \ln\left((2\pi\sigma^2)^{\frac{-n}{2}}\right) + \left(\frac{-1}{2\sigma^2} \cdot \sum_{i=1}^{n}(x_i-\mu)^2\right)$$

$$\left(\because \ln(e^x) = x\right)$$

$$\Rightarrow l(\mu, \sigma^2) = \frac{-n}{2}\ln(2\pi\sigma^2) - \left(\frac{1}{2\sigma^2} \cdot \sum_{i=1}^{n}(x_i-\mu)^2\right)$$

$$= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2} \cdot \sum_{i=1}^{n}(x_i-\mu)^2$$

conditions for Maximum value are,

$$\frac{d}{d\mu}\left(l(\mu, \sigma^2)\right) = 0 \quad - (1)$$

$$\frac{d}{d\sigma^2}\left(l(\mu, \sigma^2)\right) = 0 \quad - (2)$$

lets solve (1), i.e

partial derivative of log likelihood with respect

* Mean

$$\Rightarrow \frac{d}{d\mu}\left(\underbrace{\frac{-n}{2}\ln(2\pi)}_{\text{const}} - \underbrace{\frac{n}{2}\ln(\sigma^2)}_{\text{const}} - \frac{1}{2\sigma^2} \cdot \sum_{i=1}^{n}(x_i-\mu)^2\right) = 0$$

$$\Rightarrow \frac{d}{d\mu}\left(\frac{-1}{2\sigma^2} \cdot \sum_{i=1}^{n} (x_i^2 + \mu^2 - 2x_i\mu)\right) = 0$$

$$\Rightarrow \frac{-1}{2\sigma^2} \cdot \sum_{i=1}^{n} \left(0 + 2\mu - 2x_i \right) = 0$$

$$\Rightarrow \sum_{i=1}^{n} (2\mu - 2x_i) = 0$$

$$\Rightarrow 2n\mu - 2\sum x_i = 0$$

$$\boxed{\Rightarrow \hat{\mu}_{MLE} = \frac{\sum_{i=1}^{n} x_i}{n}}$$

Hence,

$\hat{\mu}_{MLE}$ is equal to the Sample Mean $\bar{x}$

Lets solve ②

$$\Rightarrow \frac{d}{d\sigma^2}\left(\underbrace{\frac{-n}{2}\ln(2\pi)}_{const} - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2} \cdot \sum_{i=1}^{n}(x_i - \mu)^2 \right) = 0$$

$$\Rightarrow \frac{d}{d\sigma^2}\left(-\frac{n}{2}\ln(\sigma^2)\right) - \frac{d}{d\sigma^2}\left[\frac{1}{2\sigma^2} \cdot \sum_{i=1}^{n}(x_i - \mu)^2 \right] = 0$$

$$\Rightarrow -\frac{n}{2} \times \frac{1}{\sigma^2} - \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2} \left[\frac{d}{d\sigma^2}\left(\frac{1}{\sigma^2}\right)\right] = 0$$

$$\Rightarrow \frac{-n}{2\sigma^2} - \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2} \left(\frac{-1}{(\sigma^2)^2} \right) = 0$$

$$\Rightarrow \frac{1}{2\sigma^2}\left[-n + \frac{1}{\sigma^2}\left(\sum_{i=1}^{n}(x_i - \mu)^2 \right)\right] = 0$$

$$\Rightarrow \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 = n$$

$$\boxed{\Rightarrow \hat{\sigma}^2_{MLE} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu}_{MLE})^2}$$ where,

$$\hat{\mu}_{MLE} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Hence,

$\hat{\sigma}^2_{MLE}$ is same as sample variance uncorrected

3c) Given $\{X_1, X_2 \ldots X_n\} \sim Normal(\theta, 1)$. $\delta = E[I_{X_1 > 0}]$.
T.P - $\hat{\delta}_{MLE} = \phi(\frac{1}{n} \sum_{i=1}^{n} X_i)$, where $\phi()$ is the CDF of $Z(0,1)$.

→ From assignment 1, we know $E[I_E] = Pr(E)$

$\Rightarrow \delta = E[I_{X_1 > 0}] = Pr(X_1 > 0)$
$= 1 - Pr(X_1 \le 0) = 1 - F_{X_1}(0)$
$= 1 - F_X(0)$ $(\because X_1$ is from $X \sim Nor(\theta, 1))$

Now, if we have $X \sim Normal(\mu, \sigma^2)$, we can transform into standard Normal $Z \sim Normal(0,1)$ by the property

$$Z = \frac{X - \mu}{\sigma} = \frac{X - \theta}{1} \quad (i) \quad \left(\because \mu = \theta, \sigma^2 = 1\right)$$

$\Rightarrow \delta = 1 - F_X(0) = 1 - F_Z(0 - \theta) = 1 - F_Z(-\theta)$
$= 1 - \phi(-\theta)$ $(\because \phi()$ is CDF of $Z(0,1))$
$= 1 - (1 - \phi(\theta))$ $(\because \phi(a) + \phi(-a) = 1)$
$\therefore \delta = \phi(\theta)$

→ By equivariance property, we have that if $\hat{\theta}_{MLE}$ is the MLE of θ, then $g(\hat{\theta}_{MLE})$ is the MLE of $g(\theta)$.
For normal (μ, σ^2), we have $\hat{\mu}_{MLE} = \frac{\sum X_i}{n}$
$\therefore \hat{\mu}_{MLE} = \hat{\theta}_{MLE} = \sum_{i=1}^{n} X_i / n$
We also have $g(\theta) = \delta = \phi(\theta)$
$\therefore g(\hat{\theta}_{MLE}) = \hat{\delta}_{MLE} = \phi(\hat{\theta}_{MLE}) = \phi\left(\frac{1}{n} \cdot \sum_{i=1}^{n} X_i\right)$

Hence proved, $\boxed{\hat{\delta}_{MLE} = \phi\left(\frac{1}{n} \sum_{i=1}^{n} X_i\right)}$

4) $X = \begin{cases} 2 & \text{with prob } \theta \\ 3 & \text{with prob } 1-\theta \end{cases}$. $D = \{2, 3, 2\} \overset{iid}{\sim} X(\theta)$

a) To find - $\hat{\theta}_{MME}$ using D. Steps:

Step 0 : $k = 1$, since only 1 unknown parameter θ

Step 1 : For $i = 1$, $\hat{\alpha}_1 = \sum_{j=1}^{n} x_j / n = \bar{x}$

Step 2 : For $i = 1$, $\alpha_1 = E[X(\theta)] = \sum_{x \in \Omega} x \cdot p_X(x)$

$\Rightarrow \alpha_1 = 2\theta + 3(1-\theta) = 3 - \theta$

Step 3 : For $i = 1$, we have $\hat{\alpha}_1 = \alpha_1(\theta)$

$$\Rightarrow \frac{\sum_{j=1}^{n} x_j}{n} = 3 - \theta$$

$$\Rightarrow \theta = 3 - \frac{\sum_{j=1}^{n} x_j}{n} = 3 - \frac{(2+3+2)}{3}$$

$$\boxed{\therefore \hat{\theta}_{MME} = \frac{9-7}{3} = 2/3}$$

b) Now, $\hat{\theta}_{MME} = 3 - \frac{\sum_{j=1}^{\hat{n}} x_j}{n} = 3 - \bar{x}$.

$$se(\hat{\theta}_{MME}) = \sqrt{Var(\hat{\theta}_{MME})} = \sqrt{Var(3-\bar{x})}$$

$$\overset{LOV}{=} \sqrt{Var(3) + Var(-\bar{x})} = \sqrt{Var(\bar{x})} \quad (\because Var(3) = 0)$$

$$= \sqrt{Var\left(\frac{\sum x_j}{n}\right)} = \sqrt{\frac{1}{n^2} \cdot Var(\sum x_j)}$$

$$\overset{LOV, iid}{=} \sqrt{\frac{1}{n^{\cancel{2}}} \cdot \cancel{n} \cdot Var(x_1)} = \sqrt{\frac{Var(x)}{n}} \quad (\because x_1 \sim x)$$

Now, $Var(x) = E[x^2] - (E[x])^2$ (i)

$E[x^2] = \sum_{x \in \Omega} x^2 \cdot p_X(x) = 2^2 \cdot \theta + 3^2(1-\theta) = 9 - 5\theta$

Substituting in (i), $Var(x) = (9-5\theta) - (3-\theta)^2$ (from Step 2)

$\Rightarrow Var(x) = \cancel{9} - 5\theta - \cancel{9} - \theta^2 + 6\theta = \theta - \theta^2 = \theta(1-\theta)$

$\therefore se(\hat{\theta}_{MME}) = \sqrt{\frac{\theta(1-\theta)}{n}}$. Since θ is true & unknown, we can take estimate $\hat{se}(\hat{\theta}_{MME})$ by plugging $\theta \to \hat{\theta}_{MME}$ above. $\Rightarrow \hat{se}(\hat{\theta}_{MME}) = \sqrt{\frac{\hat{\theta}_{MME}(1-\hat{\theta}_{MME})}{n}}$

(ii)

We know $\hat{\theta}_{MME} = 2/3$ from part a).

Replacing $\hat{\theta}_{MME}$ in ii), $\hat{se}(\hat{\theta}_{MME}) = \sqrt{\frac{\frac{2}{3}\cdot(1-\frac{2}{3})}{3}}$

$\therefore \hat{se}(\hat{\theta}_{MME}) = \frac{1}{3}\sqrt{\frac{2}{3}}$

Now, $(1-\alpha)\,CI = \hat{\theta}_{MME} \pm z_{\alpha/2} \cdot \hat{se}(\hat{\theta}_{MME})$

For 95% CI, $\alpha = 1 - 0.95 = 0.05$

$\therefore$ 95% CI $= \hat{\theta}_{MME} \pm z_{0.025} \cdot \hat{se}(\hat{\theta}_{MME})$

$= \frac{2}{3} \pm 1.96 \cdot \left(\frac{1}{3}\sqrt{\frac{2}{3}}\right) = \boxed{\frac{2}{3} \pm 0.533}$

$\therefore$ the 95% CI lies in the range of $\boxed{0.133 \text{ to } 1.20}$

c) To find: $\hat{\theta}_{MLE}$ using D

now, $\mathcal{L}(\theta) = \prod_{i=1}^{n} p_X(x_i)$. In order to calculate $p_X(x_i)$, we need to consider this dist^n to be similar to Bern(p), whose closed form pmf is given by $p^x \cdot (1-p)^{1-x}$.

Similarly, for dist^n $X(\theta)$, we can determine the pdf $p_X(x_i)$ in the form $\theta^{f(x)} \cdot (1-\theta)^{1-f(x)}$,

where $f(x) = ax + b = 1$ when $x = 2$

and $f(x) = ax + b = 0$ when $x = 3$

$\therefore$ we have $\Rightarrow$ $2a + b = 1$ $\Rightarrow$ $2a - 3a = 1$

$-3a + b = 0$ $\Rightarrow \boxed{a = -1,\ b = 3}$

$\Rightarrow p_X(x_i) = \theta^{-x+3} \cdot (1-\theta)^{(1-(-x+3))}$

$p_X(x_i) = \theta^{3-x_i} \cdot (1-\theta)^{x_i - 2}$ (iii)

$\therefore \mathcal{L}(\theta) = \prod_{i=1}^{n} \theta^{3-x_i} \cdot (1-\theta)^{x_i-2}$

$= \theta^{3n - \Sigma x_i} \cdot (1-\theta)^{\Sigma x_i - 2n}$

taking ln on both sides, we get

$\ln(\mathcal{L}(\theta)) = \ln\left(\theta^{3n-\Sigma x_i} \cdot (1-\theta)^{\Sigma x_i - 2n}\right)$

log-likelihood $\Rightarrow \ell(\theta) = (3n - \Sigma x_i)\ln(\theta) + (\Sigma x_i - 2n)\ln(1-\theta)$

Since we are trying to find $\hat{\theta}_{MLE}$, we need to find the value of θ that maximizes $l(\theta)$, that is, its maxima. This can be found by equating $\frac{d(l\theta)}{d\theta} = 0$

$$\Rightarrow 0 = (3n - \Sigma x_i)\frac{d}{d\theta}(\ln(\theta)) + (\Sigma x_i - 2n)\frac{d}{d\theta}(\ln(1-\theta))$$

$$\Rightarrow 0 = \frac{3n - \Sigma x_i}{\theta} - \frac{\Sigma x_i - 2n}{1-\theta}$$

$$\Rightarrow (1-\theta)(3n - \Sigma x_i) = \theta(\Sigma x_i - 2n)$$

$$\Rightarrow 3n - \Sigma x_i - 3n\theta + \cancel{\theta.\Sigma x_i} = \cancel{\theta.\Sigma x_i} - 2n\theta$$

$$\therefore \boxed{\hat{\theta}_{MLE} = \frac{3n - \Sigma x_i}{n}}$$

using D

$$\Rightarrow \boxed{\hat{\theta}_{MLE} = \frac{3.3 - (2+3+2)}{3} = 2/3}$$

Scanned with CamScanner

5a) To find - $\hat{\lambda}_{MME}$ for $X \sim Exp(\lambda)$

Step 0 : $k=1$ since only one unknown parameter λ

Step 1 : For $i=1$, $\hat{\alpha}_1 = \frac{\sum_{j=1}^{n} x_j}{n} = \bar{x}$

Step 2 : For $i=1$, $\alpha_1(\lambda) = E[Exp(\lambda)] = 1/\lambda$

Step 3 : For $i=1$, we have $\hat{\alpha}_1 = \alpha_1(\lambda)$

$\Rightarrow \frac{\sum_{j=1}^{n} x_j}{n} = \frac{1}{\lambda} \Rightarrow \boxed{\hat{\lambda}_{MME} = \frac{n}{\sum_{j=1}^{n} x_j}}$

b) To find - $\hat{\lambda}_{MLE}$ for $X \sim Exp(\lambda)$

now, $L(\lambda) = \prod_{i=1}^{n} p_X(x_i) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n . e^{-\lambda \sum_{i=1}^{n} x_i}$

$\Rightarrow \ln(L(\lambda)) = \ln(\lambda^n . e^{-\lambda \sum x_i}) = n\ln(\lambda) - \lambda \sum x_i$

Now to find maxima, $\frac{d(\ln(L(\lambda)))}{d\lambda} = \frac{d(l(\lambda))}{d\lambda} = 0$

$\Rightarrow n\frac{d(\ln(\lambda))}{d\lambda} - \sum x_i \frac{d\lambda}{d\lambda} = 0 \Rightarrow \frac{n}{\lambda} = \sum x_i \Rightarrow \boxed{\hat{\lambda}_{MLE} = \frac{n}{\sum_{i=1}^{n} x_i}}$

5c) – For the Normally distributed Acceleration dataset,

$\hat{\mu}_{MME} = 15.568$, $\hat{\sigma}^2_{MME} = 7.586$

– For the Uniformly distributed Model dataset,

$\hat{a}_{MME} = 69.614$, $\hat{b}_{MME} = 82.406$

– For the Exponentially distributed MPG dataset,

$\hat{\lambda}_{MME} = 0.043$

5d) – For the acceleration dataset,

$\hat{\mu}_{MLE} = 15.568$, $\hat{\sigma}^2_{MLE} = 7.586$

– For the model dataset,

$\hat{a}_{MLE} = 70.000$, $\hat{b}_{MLE} = 82.000$

– For the MPG dataset,

$\hat{\lambda}_{MLE} = 0.043$

Scanned with CamScanner

Question 6 : Clinical Testing

Total Healthy Patients = 100
Total Sick Patients = 100

Results of Disease Detection Test ⇒
98/100 correctly identified as healthy
99/100 correctly identified as sick

Let null hypothesis H_0 : patient is healthy
∴ H_1 : patient is sick.

Ground Truth \ Testing		Healthy (Accept)	Sick (Reject)
Healthy	H_0 true	98 TN	2 FP
Sick	H_0 false	1 FN	99 TP

True +ve = 99
True −ve = 98
False +ve = 100 − 98 = 2
False −ve = 100 − 99 = 1

(a) Precision $= \frac{TP}{TP+FP} = \frac{99}{99+2} = \frac{99}{101} = \underline{\underline{0.98}}$

(b) Recall $= \frac{TP}{TP+FN} = \frac{99}{99+1} = \frac{99}{100} = \underline{\underline{0.99}}$

(c) Type 1 Error = FP = $\underline{\underline{2}}$

(d) Type 2 Error = FN = $\underline{\underline{1}}$

7a) Given, Null hypothesis $H_0 : \theta = \theta_0$, so $H_1 : \theta \neq \theta_0$

Also, true value of $\theta = \theta_*$. We have to show that

→ $Pr(\text{Type II error}) = \phi\left(\frac{\theta_0 - \theta_*}{\hat{se}} + z_{\alpha/2}\right) - \phi\left(\frac{\theta_0 - \theta_*}{\hat{se}} - z_{\alpha/2}\right)$

→ Now, Type II error is same as False ~~Positive~~ Negative. FN

∴ $Pr(\text{Type II error}) = Pr(FN) = Pr(\text{accept } H_0 \mid H_0 \text{ is false})$

– In Wald's test, we reject H_0 if $|W| > z_{\alpha/2}$

∴ to accept H_0, we need $|W| = \left|\frac{\hat{\theta} - \theta_0}{\hat{se}(\hat{\theta})}\right| \leq z_{\alpha/2}$

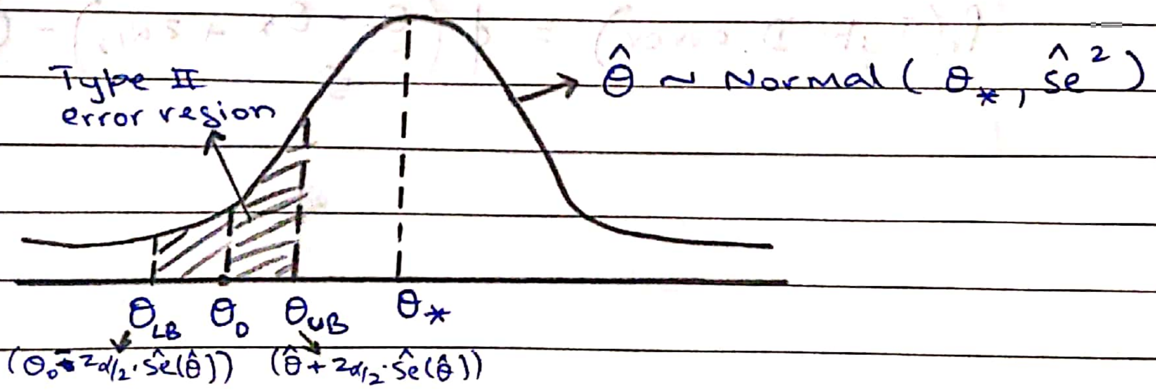
$$\Rightarrow -z_{\alpha/2} \leq \frac{\hat{\theta} - \theta_0}{\hat{se}(\hat{\theta})} \leq z_{\alpha/2}$$

$$\Rightarrow \theta_0 - z_{\alpha/2} \cdot \hat{se}(\hat{\theta}) \leq \hat{\theta} \leq \theta_0 + z_{\alpha/2} \cdot \hat{se}(\hat{\theta}),$$

which is the expected region where we can find Type II errors.

→ Let us denote the lower bound $\theta_0 - z_{\alpha/2} \cdot \hat{se}(\hat{\theta})$ as θ_{LB} and the upper bound $\theta_0 + z_{\alpha/2} \cdot \hat{se}(\hat{\theta})$ as θ_{UB}

→ Now, we need to plot the distribution of estimate $\hat{\theta}$. It will be distributed normally, and centered around true mean θ_*. θ_0 can be on either side of θ_*, but let us assume $\theta_0 < \theta_*$. We can draw the curve for $\hat{\theta} \sim \text{Normal}(\theta_*, \hat{se}^2)$ as below



From above, it is clear that we need to find area of shaded region, i.e, $Pr(\text{Type II error}) = Pr(\theta_{LB} \leq \hat{\theta} \leq \theta_{UB})$

$$\Rightarrow Pr(\theta_{LB} \le \hat{\theta} \le \theta_{UB}) = Pr(\hat{\theta} \le \theta_{UB}) - Pr(\hat{\theta} \le \theta_{LB})$$
$$= F_{\hat{\theta}}(\theta_{UB}) - F_{\hat{\theta}}(\theta_{LB}) \quad (i)$$

Now, as we did in 3c), we can transform $\hat{\theta} \sim Normal(\theta_*, \hat{se}^2)$ to $Z \sim Normal(0,1)$ using

$$Z = \frac{X - \mu}{se} \Rightarrow \frac{\theta_{UB} - \theta_*}{\hat{se}} \text{ for } \hat{\theta} = \theta_{UB},$$

$$\frac{\theta_{LB} - \theta_*}{\hat{se}} \text{ for } \hat{\theta} = \theta_{LB}$$

$\Rightarrow$ i) becomes $F_Z\left(\frac{\theta_{UB} - \theta_*}{\hat{se}}\right) - F_Z\left(\frac{\theta_{LB} - \theta_*}{\hat{se}}\right)$

$$\rightarrow F_Z\left(\frac{\theta_{UB} - \theta_*}{\hat{se}}\right) = \phi\left(\frac{(\theta_0 + z_{\alpha/2} \cdot \hat{se}) - \theta_*}{\hat{se}}\right)$$
$$= \phi\left(\frac{\theta_0 - \theta_*}{\hat{se}} + z_{\alpha/2}\right) \quad (ii)$$

$$\rightarrow F_Z\left(\frac{\theta_{LB} - \theta_*}{\hat{se}}\right) = \phi\left(\frac{(\theta_0 - z_{\alpha/2} \cdot \hat{se}) - \theta_*}{\hat{se}}\right)$$
$$= \phi\left(\frac{\theta_0 - \theta_*}{\hat{se}} - z_{\alpha/2}\right) \quad (iii)$$

Substituting (ii) & (iii) in (i), we get

$$\boxed{Pr(\text{Type II error}) = \phi\left(\frac{\theta_0 - \theta_*}{\hat{se}} + z_{\alpha/2}\right) - \phi\left(\frac{\theta_0 - \theta_*}{\hat{se}} - z_{\alpha/2}\right)}$$

Scanned with CamScanner

7b) Given, we observe 46 successes out of 100 trials.

⇒ The given experiment ~ Bernoulli (p).

Also given, null hypothesis is that coin is unbiased.

i) ⇒ $H_0: p = 0.5$, $H_1: p \neq 0.5$ (p_0 here is 0.5)

→ Let us take $\hat{p}_{MLE}$ as the estimator, since it is AN & consistent. We can hence do Wald's test using $\hat{p}_{MLE}$.

$$\hat{p}_{MLE} = \frac{\Sigma X_i}{n} = \frac{46}{100} = 0.46$$

→ By Wald's test, $W = \left| \frac{\hat{p} - p_0}{\hat{se}(\hat{p})} \right|$

now, $\hat{se}(p) = \sqrt{\frac{Var(p)}{n}}$ (∵ as shown in class)

$$= \sqrt{\frac{p(1-p)}{n}}$$

$$\therefore \hat{se}(\hat{p}) = \sqrt{\frac{\hat{p}_{MLE}(1-\hat{p}_{MLE})}{n}} \quad \text{(By equivariance)}$$

$$= \sqrt{\frac{0.46(1-0.46)}{100}} = \frac{0.498}{10} = 0.0498$$

$$\therefore W = \left| \frac{0.46 - 0.50}{0.0498} \right| = |-0.803| = 0.803$$

→ For $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$

since $|W| \leq z_{\alpha/2}$, we fail to reject H_0, that is, accept that the coin is unbiased

ii) Now, if $H_0: p = 0.7$, $H_1: p \neq 0.7$ (p_0 is 0.7 here),

~~then~~ $W = \left| \frac{0.46 - 0.7}{0.0498} \right| = |-4.815| = 4.815$

For same $\alpha = 0.05$, we now have $|W| > z_{\alpha/2}$

Hence, we reject H_0, i.e, we say that the coin is biased

8a) Given, a dataset distributed as Normal (θ, σ^2)
H_0: True mean $\theta = \theta_0$, $H_1: \theta \neq \theta_0$, where $\theta_0 = 0.5$
We also need to use $\alpha = 0.02$.
$\Rightarrow z_{\alpha/2} = z_{0.01} \simeq 2.3263$

Now, by wald's test, we reject H_0 if

$$W = \left| \frac{\hat{\theta} - \theta_0}{\hat{se}(\hat{\theta})} \right| > z_{\alpha/2} \qquad \boxed{\begin{array}{l} \hat{\theta} = 0.541 \\ \hat{se}(\hat{\theta}) = 0.1032 \end{array}}$$

From our experiment through the program, we calculated $W \simeq 0.397$
$\therefore$ since $W = 0.397 \leq z_{\alpha/2}$ (2.3263), we can say that we failed to reject the null hypothesis that true mean, $\theta = 0.5$

b) Given, datasets X & Y are drawn from 2 $\perp$ Normal distributions. Using $\alpha = 0.05$, we need to use Wald's 2-population test to determine if the sample means of X & Y are same.

$\rightarrow$ Let us denote sample mean estimators of X and Y as μ_X & μ_Y.
Thus, we have $\delta = \mu_X - \mu_Y$
$\Rightarrow H_0: \delta = 0$, $H_1: \delta \neq 0$

$\rightarrow$ For wald's 2-population test, we have

$$W = \left| \frac{\hat{\delta}}{\hat{se}(\hat{\delta})} \right|,$$ where $\hat{\delta} = \hat{\mu}_X - \hat{\mu}_Y$, where
$\hat{\mu}_X$ - sample mean of X
μ_Y - sample mean of Y

$$\Rightarrow W = \left| \frac{\hat{\delta}}{\sqrt{\frac{\hat{se}_X^2}{n} + \frac{\hat{se}_Y^2}{m}}} \right| = \left| \frac{\hat{\mu}_X - \hat{\mu}_Y}{\sqrt{\frac{\hat{se}_X^2}{n} + \frac{\hat{se}_Y^2}{m}}} \right| \quad (i)$$

$(m = n = 750)$

now, $\hat{u}_x = \frac{\sum x_i}{n} \simeq 5.005$, $\hat{u}_y = \frac{\sum y_i}{m} = 5.846$

$$\hat{se}_x^2 = \frac{1}{n} \sum (x_i - \hat{u}_x)^2 = 2.3615$$

$$se_y^2 = \frac{1}{m} \sum (y_i - \hat{u}_y)^2 = 6.4724$$

$\therefore$ From (i), $W = \left| \frac{5.005 - 5.846}{\sqrt{\frac{2.3615}{750} + \frac{6.4724}{750}}} \right|$

$$\Rightarrow \boxed{W = 7.748}$$

→ Since $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$

$\therefore$, we have $W > z_{\alpha/2}$ and can reject the Null hypothesis (H_0) that the sample means are equal, i.e, sample means of X & Y are not same.

→ Wald's 2-population test is applicable here, because it is given that X and Y are independant and also because of the estimators we are using, which is the sample mean u_x / u_y. By central limit theory, the sample mean estimator is Asymptotically Normal, around its true mean μ, and we can hence perform this test on it.