ASSIGNMENT

Piyush Soni [SBU ID # 113273375]

Abhishek Bhattad [SBUID# 113277131]

Aditya Nandan Bhide [SBU ID# 113216565]

Niket Shah [SBU ID# 113274949]

1)

a) let $Z \sim \text{Gamma}(x, y)$ To find $\hat{x}_{MME}, \hat{y}_{MME}$

Given : $\mu = xy$

$\sigma^2 = xy^2$

$E[Z] = xy$

Number of unknowns = 2.

We find $\hat{x}_{MME}, \hat{y}_{MME}$ by equating $\hat{E[Z^i]} = E[Z^i]$ for $i = \{1, 2\}$

$$\hat{E[Z^i]} = \frac{\sum_{j=1}^{n}(Z_j^i)}{n} \quad \text{for } i = \{1, 2\}$$

$$\therefore \hat{E[Z^1]} = \frac{1}{n}\sum_{j=1}^{n} z_j = E[Z]$$

$$\therefore \frac{1}{n}\sum_{j=1}^{n} Z_j = \mu = xy. \quad -①$$

Similarly

$$\hat{E[Z^2]} = E[Z^2]$$

$$\frac{1}{n}\sum_{j=1}^{n} Z_j^2 = E[Z^2]$$

But we know that $E[Z^2] = \sigma^2 + \mu^2$

$$\therefore \frac{1}{n}\sum_{j=1}^{n} z_j^2 = xy^2 \left(\frac{1}{n}\sum_{j=1}^{n} Z_j\right)^2.$$

$$\therefore xy^2 = \frac{1}{n}\sum_{j=1}^{n} z_j^2 - \left(\frac{1}{n}\sum_{j=1}^{n} z_j\right)^2 \quad -②$$

Dividing ② by ①

$$\hat{y}_{MME} = \frac{\frac{1}{n}\sum_{j=1}^{n} z_j^2 - \left(\frac{1}{n}\sum_{j=1}^{n} z_j\right)^2}{\frac{1}{n}\sum z_j}$$

$$\boxed{\hat{y}_{MME} = \frac{\sum_{j=1}^{n} z_j^2 - \frac{1}{n}\left(\sum_{j=1}^{n} z_j\right)^2}{\sum_{j=1}^{n} z_j}}$$

from (1)

$$xy = \frac{1}{n}\sum_{j=1}^{n} z_j$$

$$\boxed{\hat{x}_{MME} = \frac{\frac{1}{n}\left(\sum_{j=1}^{n} z_j\right)^2}{\sum_{j=1}^{n} z_j^2 - \frac{1}{n}\left(\sum_{j=1}^{n} z_j\right)^2}}$$

1 b) $X \sim Unif(a, b)$ Given $\bar{X} = \frac{\sum X_i}{n}$

find $\hat{a}_{MME}$, $\hat{b}_{MME}$ $\bar{S}^2 = \left(\left(\frac{\sum X_i^2}{n}\right)\right) - \bar{X}^2$

Number of unknowns = 2

$\therefore$ lets Equate $\cancel{E[\hat{X}^j] = E[X}$

$$E[\hat{X}^j] = E[X^j] \quad \text{for } j = \{1, 2\}$$

$E[X] = \frac{a+b}{2}$ — for uniform distro.

$$\therefore \frac{1}{n}\sum X_i = \frac{a+b}{2} \quad (1)$$

$$\therefore \bar{X} = \frac{a+b}{2} \qquad b = 2\bar{X} - a \quad \text{—(1)}$$

$$E[X^2] = \int_a^b x^2 \cdot f(x)\, dx$$

$$= \int_a^b x^2 \cdot \frac{1}{b-a}\, dx = \left.\frac{x^3}{3}\right|_a^b \frac{1}{b-a}$$

$$= \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3}$$

We know that $E[X^2] = \bar{S}^2 + (\bar{X})^2$

$$\frac{a^2+b^2+ab}{3} = \bar{S}^2 + (\bar{X})^2$$

$$\therefore a^2+b^2+ab = 3\left(\bar{S}^2 + (\bar{X})^2\right) \quad -(2)$$

Substituting (1) in (2)

$$a^2 + (2\bar{X} - a)^2 + a(2\bar{x} - a) = 3(\bar{S}^2 + \bar{x}^2)$$

$$a^2 + 4\bar{X}^2 + a^2 - 4a\bar{X} + 2\bar{x}a - a^2 = 3(\bar{S}^2 + \bar{x}^2)$$

$$\therefore a^2 - 2\bar{x}a + 4\bar{x}^2 = 3\bar{S}^2 + 3\bar{x}^2$$

$$\therefore a^2 - 2\bar{x}a + \bar{x}^2 - 3\bar{S}^2 = 0$$

Solving for a

$\because$ Soln of quad. eqn $= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

$$a = \frac{2\bar{X} \pm \sqrt{4\bar{x}^2 - 4(\bar{x}^2 - 3\bar{S}^2)}}{2}$$

$$= \frac{2\bar{x} \pm \sqrt{(3\bar{S}^2)4}}{2}$$

$$= \bar{x} \pm \sqrt{3\bar{S}^2} \quad — (3)$$

Substituting (3) in (1)

$$b = 2\bar{X} - a$$
$$= 2\bar{X} - (\bar{X} \pm \sqrt{3\bar{S}^2})$$
$$= \bar{X} \mp \sqrt{3\bar{S}^2}$$

Since $b > a$

$$\boxed{\begin{aligned} \hat{a}_{MME} &= \bar{X} - \sqrt{3\bar{S}^2} \\ \hat{b}_{MME} &= \bar{X} + \sqrt{3\bar{S}^2} \end{aligned}}$$

2 $X_1, X_2 \ldots X_n \overset{iid}{\sim} Exp(1/\beta)$ To find $\hat{\beta}_{MLE}$

We know for $Exp(\lambda)$ $f_X(x) = \lambda e^{-\lambda x}$

let $\frac{1}{\beta} = \lambda$

$$l(\cancel{\lambda}) = \prod_{i=1}^{n} f_X(x_i)$$

where $l(\beta)$ is the function we want to maximize according to MLE definition

$$l(\cancel{\lambda}) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i}$$

$$= \lambda^n \cdot e^{-\lambda \sum x_i}$$

$$l(\beta) = \left(\frac{1}{\beta}\right)^n e^{-1/\beta \sum X_i}$$

$$\ln(l(\beta)) = -n \log \beta + \left(-\frac{1}{\beta} \sum_{i=1}^{n} X_i \log e \right)$$

$$= -n \log \beta - \frac{1}{\beta} \sum_{i=1}^{n} X_i$$

To maximize, $\frac{d(\ln(l(\beta)))}{d\beta} = 0$

$$\therefore \frac{d(\ln(l(\beta)))}{d\beta} = -\frac{n}{\beta} + \frac{1}{\beta^2} \cdot \sum_{i=1}^{n} X_i = 0$$

$$\therefore \frac{1}{\beta^2} \sum_{i=1}^{n} X_i = \frac{n}{\beta}$$

$$\therefore \boxed{\hat{\beta}_{MLE} = \frac{\sum_{i=1}^{n} X_i}{n}} \longrightarrow \text{sample mean}$$

By def^n. of bias

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

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

we

$$= \frac{n}{n} E[X_i] - \mu$$

$$= \frac{n\mu}{n} - \mu = \underline{\underline{0}}$$

$$\therefore \text{Bias}(\hat{\beta}_{MLE}) = 0 \quad \text{— (1)}$$

~~Se~~ By def^n of se

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

$$= \sqrt{var\left(\frac{\sum X_i}{n}\right)}$$

cov

$$= \sqrt{\frac{n \, Var(X_i)}{n^2}} \qquad \text{let } Var(X_i) = \sigma^2$$

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

$$se(\hat{\beta}_{MLE}) = \frac{\sigma}{\sqrt{n}} \quad \text{— (2)}$$

We can see that $se(\hat{\beta}_{MLE}) \to 0$ as $n \to \infty$
& $\text{Bias}(\hat{\beta}_{MLE}) = 0$

Thus $\hat{\beta}_{MLE}$ is a consistent estimator, which will converge to β (true value)

Q 3 a] $X_1, X_2, X_3, \dots X_2 \overset{iid}{\sim} \text{Poisson}(\lambda)$

We have to find λ_{MLE}.

p.d.f of Poisson is

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

$$L(\lambda) = \prod_{i=1}^{n} f_x(X_i)$$

$$= \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{X_i}}{X_i!}$$

$$= \frac{e^{-n\lambda} \lambda^{\Sigma X_i}}{\prod_{i=1}^{n} (X_i)!}$$

Taking log on both sides to get log likelihood

$$l(\lambda) = \ln \left[\frac{e^{-n\lambda} \lambda^{\Sigma X_i}}{\prod_{i=1}^{n} (X_i!)} \right]$$

$$= \ln(e^{-n\lambda}) + \ln(\lambda^{\Sigma X_i}) - \ln\left(\prod_{i=1}^{n} (X_i!)\right)$$

$$= -n\lambda \ln(e) + \sum_{i=1}^{n} X_i \ln(\lambda) - \sum_{i=1}^{n} \ln(X_i!)$$

For getting maximum $\frac{dl(\lambda)}{d\lambda} = 0$

$\therefore$ Differentiating on both sides.

$$\frac{dl(\lambda)}{d\lambda} = -n + \sum_{i=1}^{n} X_i \times \frac{1}{\lambda} = 0$$

$$\therefore \quad 0 = -n + \frac{\sum_{i=1}^{n} x_i}{\lambda}$$

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

b] $X_1, X_2, X_3 \dots X_n \overset{iid}{\sim} \text{Normal}(\mu, \sigma^2)$

$$\therefore \quad f_X(x_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}}$$

We have to find μ_{MLE} & σ^2_{MLE}

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

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

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

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

Taking log on both sides to get log likelihood.

$$\ell(P) = \ln\left[\frac{1}{(\sqrt{2\pi}\sigma)^n} e^{\left(\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)}\right]$$

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

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

$$= -n \ln(\sqrt{2\pi}) - n \ln \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (X_i - \mu)^2 \quad \dots \text{(i)}$$

Now we need to take partial derivative w.r.t. μ & σ

$\therefore$ For max likelihood $\frac{\partial l(P)}{\partial \mu} = 0$ & $\frac{\partial l(P)}{\partial \sigma} = 0$.

$\therefore$ Taking partial derivative w.r.t. μ for (i)

$$\frac{\partial l(P)}{\partial \mu} = -0 - 0 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} 2(X_i - \mu)$$

$$\therefore 0 = -\frac{2}{2\sigma^2} \sum_{i=1}^{n} (X_i - \mu)$$

$$\therefore \sum_{i=1}^{n} (X_i - \mu) = 0$$

$$\therefore \sum_{i=1}^{n} X_i - n\mu = 0$$

$$\boxed{\therefore \hat{\mu}_{MLE} = \sum_{i=1}^{n} \frac{X_i}{n} = \text{Sample mean}}$$

$\therefore$ Taking partial derivative w.r.t. σ for (i)

$$\frac{\partial l(P)}{\partial \sigma} = -0 - \frac{n}{\sigma} + \frac{2}{2\sigma^3} \sum_{i=1}^{n} (X_i - \mu)^2$$

$$0 = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{n} (X_i - \mu)^2$$

$$\therefore \sigma^2 n = \sum_{i=1}^{n} (X_i - \mu)^2$$

$$\boxed{\therefore \sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2 = \text{uncorrected Sample variance}}$$

c] $X_1, X_2, X_3 \dots X_n \sim \text{Normal}(\theta, 1)$

$\delta = E[I_{X_1 > 0}]$

By equivariance $\hat{\theta}_{MLE}$ is equivariant. then

$\therefore \hat{f}(\theta)_{MLE} = f(\hat{\theta}_{MLE})$

$\therefore \delta = P(X_i > 0)$

$\therefore \quad = 1 - F_x(0)$

Now to get values for standard Normal.

$Z \sim \text{Normal}(0, 1)$

$\theta a + b = 0$ ... compare mean

$1^2 \times a = 0$ ... compare varience

$\therefore a = 0$

$\therefore b = -\theta$

$\therefore z = x - \theta$

$\therefore$ The value of δ will be

$\therefore \delta = 1 - F_x(\theta)$

$= 1 - \Phi(-\theta)$

$= 1 - Pr(z < -\theta)$

$= Pr(z > -\theta)$

$= Pr(z < \theta)$ ... (By symmetry in Standard normal.)

$= \Phi(\theta)$

Now by equivariance $\theta \rightarrow \hat{\theta}_{MLE}$

$\therefore \delta = \Phi(\hat{\theta}_{MLE})$

Now from 3b]

$\hat{\theta}_{MLE}$ = MLE of mean of normal = sample variance

$\therefore \hat{\theta}_{MLE} = \hat{\sum_{i=1}} \frac{X_i}{n}$

$\therefore \delta = \cancel{\Phi}\ \Phi\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$

Ans) Q.4)

Given,

$$X = \begin{cases} 2 & \theta \\ 3 & \text{otherwise } (1-\theta). \end{cases}$$

$D = \{2, 3, 2\}.$

A) $\hat{\theta}_{MME} = ?.$

we find first moment of X.

$$\therefore E[X] = \sum x P_X(x)$$

$$= 2\theta + 3(1-\theta).$$

$$= 2\theta + 3 - 3\theta$$

$$\therefore E[X] = 3 - \theta.$$

$\therefore$ For $\hat{\theta}_{MME}$, $E[X] = \frac{1}{n} \sum X_i$

$$\therefore 3 - \hat{\theta}_{MME} = \overline{X}$$

$$\therefore \hat{\theta}_{MME} = 3 - \overline{X} = 3 - \frac{(2+3+2)}{3}$$

$$= 3 - \frac{7}{3}$$

$$\therefore \boxed{\hat{\theta}_{MME} = \frac{2}{3}}$$

B) $se(\hat{\theta}_{MME}) = ?$

Now, $se(\hat{\theta}_{MME}) = \sqrt{Var(3-\overline{X})}$

$$\overset{cov}{=} \sqrt{Var(3) + Var(\overline{X})}$$

$$= \sqrt{Var\left(\frac{X_1 + X_2 + \cdots X_n}{n}\right)}$$

$$= \sqrt{\frac{1}{m^2} \text{Var}(\bar{X})}$$

$$\overset{\text{iid}}{=} \sqrt{\frac{1}{m^2} \text{Var}(\Sigma X_i)}$$

$$\overset{\text{LOV}}{=} \sqrt{\frac{1}{m^2} \times m \text{Var}(X_1)}$$

$$= \sqrt{\frac{\text{Var}(X)}{m}} \quad \because X_1 \text{ is distributed as } X. \quad —(i)$$

Now, $\text{Var}(X) = E[X^2] - (E[X])^2$

$$= \Sigma x^2 P_X(x) - (3-\theta)^2$$

$$= 4\theta + 9(1-\theta) - (9 - 6\theta + \theta^2)$$

$$= 4\theta + \cancel{9} - 9\theta - \cancel{9} + 6\theta - \theta^2.$$

$$= \theta - \theta^2$$

$$= \theta(1-\theta) \quad -(ii).$$

Substituting Eq (ii) in Eq (i),

$$\therefore se(\bar{\theta}_{MME}) = \sqrt{\frac{\theta(1-\theta)}{m}}$$

Now estimator of se is $\hat{se}$ and MME is a good estimator, we get

$$\therefore \hat{se}(\bar{\theta}_{MME}) = \sqrt{\frac{\hat{\theta}_{MME}(1-\bar{\theta}_{MME})}{m}}$$

$$= \sqrt{\frac{\frac{2}{3}\left(1-\frac{2}{3}\right)}{3}}$$

$$= \sqrt{\frac{\frac{2}{3}\times\frac{1}{3}}{3}}$$

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

Since, $\hat{\theta}_{MME}$ is AN.

$(1-\alpha)$ CI will be

$$\hat{\theta}_{MME} \pm Z_{\alpha/2} \cdot \hat{se}(\hat{\theta}_{MME}).$$

$\therefore$ 95% CI $\Rightarrow$ $(1-0.05)$ CI $\Rightarrow$

$$\hat{\theta}_{MME} \pm Z_{0.025} \cdot \hat{se}(\hat{\theta}_{MME}).$$

$$= \frac{2}{3} \pm 1.96 \times \frac{1}{3}\sqrt{\frac{2}{3}}$$

i.e. $\boxed{\frac{2}{3} - 1.96 \times \frac{1}{3}\sqrt{\frac{2}{3}},\ \frac{2}{3} + \frac{1.96}{3}\sqrt{\frac{2}{3}}}$

i.e. $\boxed{0.1332, 1.2001}$

9 $\qquad \vec{\theta}_{MLE} = ?$ $\qquad X = \begin{cases} 2 & \theta \\ 3 & 1-\theta \end{cases}$

Writing in terms of

$$P_X(x) = \theta^{\alpha} (1-\theta)^{\beta}, \quad \beta = 1-\alpha.$$

Let $\alpha = ax + b$.

For $x = 2$, $\alpha = 1$.

$\therefore 2a + b = 1$ — (i).

For $x = 3$, $\alpha = 0$.

$\therefore 3a + b = 0$ — (ii).

$\therefore$ Eq (ii) − Eq (i) → Substraction

$\therefore a = -1$.

$\therefore b = 3$.

$$\boxed{\begin{aligned} &\alpha = 3 - x \\ &\beta = 1-\alpha = 1-(3-x) = x-2 \end{aligned}}$$

$$\therefore P_X(x) = \theta^{(3-x)} (1-\theta)^{(x-2)}.$$

Likelihood, $\mathcal{L}(P) = \prod_{i=1}^{m} \theta^{(3-x)} (1-\theta)^{(x-2)}$

$$= \theta^{3m - \sum X_i} (1-\theta)^{\sum X_i - 2m}$$

Taking log-likelihood,

$$l(P) = (3m - \sum X_i) \ln(\theta) + (\sum X_i - 2m) \ln(1-\theta$$

For max-likelihood

$$\frac{dl}{d\theta} = 0.$$

$$\therefore \quad 3m - \sum X_i \times \frac{1}{\theta} - \frac{1}{(1-\theta)} \left(\sum X_i - 2m\right) = 0.$$

$$\therefore \quad (1-\theta)\left(3m - \sum X_i\right) = \theta\left(\sum X_i - 2m\right).$$

$$\therefore \quad 3m - \sum X_i - 3m\theta + \cancel{\theta \sum X_i} = \cancel{\theta \sum X_i} - 2m\theta.$$

$$\therefore \quad m\theta = 3m - \sum X_i$$

$$\therefore \quad \theta = 3 - \frac{\sum X_i}{m}.$$

Now, replacing θ with $\hat{\theta}_{MLE}$ and $D = \{2, 3, 2\}$.

$$\therefore \quad \hat{\theta}_{MLE} = 3 - \frac{\sum X_i}{m}$$

$$= 3 - \frac{7}{3}$$

$$\therefore \boxed{\hat{\theta}_{MLE} = \frac{2}{3}}$$

(5) $Acc \sim Nor(\mu, \sigma^2)$
$mod \sim Unif(a, b)$
$mpg \sim exp(\lambda)$

(a) $mpg \sim exp(\lambda)$

To find λ_{MME}, we use first moment -

$$\Rightarrow E[X] = \frac{1}{\lambda} \quad \{\text{from class}\}$$

$$\Rightarrow \frac{\sum X_i}{n} = \frac{1}{\lambda}$$

$$\Rightarrow \boxed{\hat{\lambda}_{MME} = \frac{n}{\sum X_i}}$$

(b) $mpg \sim exp(\lambda) \Rightarrow f_X(x) = \lambda e^{-\lambda x}$

In order to find $\hat{\lambda}_{MLE}$, we maximize the likelihood -

$$\Rightarrow \text{likelihood}, \mathcal{L}(p) = \prod_{i=1}^{n} f_X(x)$$

$$= \prod_{i=1}^{n} \lambda e^{-\lambda x_i}$$

$$\mathcal{L}(p) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i}$$

$$= \lambda^n e^{-\lambda \Sigma x_i}$$

$\Rightarrow$ log-likelihood –

$$l(p) = \ln\left(\lambda^n e^{-\lambda \Sigma x_i}\right)$$

$$= \ln(\lambda^n) + \ln\left(e^{-\lambda \Sigma x_i}\right)$$

$$= n \ln(\lambda) - \lambda \Sigma x_i$$

$\Rightarrow$ for finding, $\hat{\lambda}_{MLE}$, maximizing $l(p)$

$$\therefore \frac{d\,l(p)}{d\lambda} = 0$$

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

$$\Rightarrow \quad \boxed{\hat{\lambda}_{MLE} = \frac{n}{\Sigma x_i}}$$

(c) For Accel^n dataset -

$$ACC \sim Nor(\mu, \sigma^2)$$

$$\Rightarrow \sigma^2_{MME} \stackrel{class}{=} \frac{1}{n}\sum x_j^2 - \left(\frac{1}{n}\sum x_j\right)^2$$

from dataset - $n = 398$

$$\Rightarrow \frac{1}{n}\sum x_i = 15.56809$$

$$\Rightarrow \frac{1}{n}\sum x_j^2 = 249.9512$$

$\Rightarrow$ Thus, $\sigma^2_{MME} = (249.9512) - (15.56809)^2$

$$= 7.58577$$

$$\boxed{\sigma^2_{MME} \approx 7.586}$$

$$\Rightarrow \hat{\mu}_{MME} \stackrel{class}{=} \frac{\sum X_i}{n}$$

$$\boxed{\hat{\mu}_{MME} = 15.56809 \approx 15.568}$$

For model year dataset -

$$\text{mod} \sim \text{Unif}(a, b)$$

$$\Rightarrow \hat{a}_{MME} \overset{1(b)}{=} \bar{x} - \sqrt{3\bar{s}^2}$$

$$= \frac{\sum x_i}{n} - \sqrt{3\left(\frac{1}{n}\sum x_i^2 - \left(\frac{1}{n}\sum x\right)^2\right)}$$

from dataset,

$$n = 398$$

$$\frac{\sum x_i}{n} = 76.01005$$

$$\frac{\sum x_i^2}{n} = 5791.66$$

Thus,

$$\hat{a}_{MME} = 76.01005 - \sqrt{3 \times \left(5791.166 - (76.01005)^2\right)}$$

$$= 76.01005 - \sqrt{3 \times 13.638298}$$

$$\boxed{\hat{a}_{MME} = 69.614}$$

$$\hat{b}_{MME} \overset{1(b)}{=} \bar{x} + \sqrt{3 \bar{s}^2}$$

$$= \frac{\sum x_i}{n} + \sqrt{3\left(\frac{\sum x_i^2}{n} - \left(\frac{1}{n}\sum x_i\right)^2\right)}$$

$$= 76.01005 + \sqrt{3(5791.166 - (76.01005)^2)}$$

$$= 76.01005 + \sqrt{3 \times 13.63828}$$

$$\boxed{\hat{b}_{MME} \approx 82.407}$$

- for mpg dataset –

$$mp \sim \exp(\lambda)$$

From dataset

$$\frac{\sum x_i}{n} = 23.51457$$

$$\hat{\lambda}_{MME} \overset{5(a)}{=} \frac{n}{\sum x_i} = \frac{1}{23.51457}$$

$$\Rightarrow \boxed{\hat{\lambda}_{MME} = 0.0425}$$

(d) • For ACC. dataset -

$$ACC \sim Nor(\mu, \sigma^2)$$

$$\Rightarrow \boxed{\hat{\mu}_{MLE} \overset{3b}{=} \frac{\sum x_i}{n} = 15.568}$$

$$\hat{\sigma}^2_{MLE} \overset{3b}{=} \frac{\sum (x-\mu)^2}{n} \longrightarrow \text{from dataset}$$

$$= 7.585741$$

$$\boxed{\hat{\sigma}^2_{MLE} = 7.586}$$

• for mpg dataset -

$$mpg \sim exp(\lambda)$$

$$\Rightarrow \hat{\lambda}_{MLE} \overset{5(b)}{=} \frac{n}{\sum x_i} = \frac{1}{23.51457} = 0.0425$$

$$\boxed{\hat{\lambda}_{MLE} = 0.0425}$$

- For model year dataset

$$mo \sim \text{unif}(a, b)$$

$$\Rightarrow \hat{a}_{MLE} \overset{\text{class}}{=} \min(D)$$

$$= 70$$

$$\Rightarrow \boxed{\hat{a}_{MLE} = 70}$$

$$\Rightarrow \hat{b}_{MLE} \overset{\text{class}}{=} \max(D)$$

$$= 82$$

$$\Rightarrow \boxed{\hat{b}_{MLE} = 82}$$

Q6 100 health 100 sick.

let null hypothesis be patient is healthy

$\therefore H_0$ = heathy H_1 = not healthy = sick.

The truth table will be

Test output →	Patient heathy	Patient Sick.
H_0 true	98	2
H_0 false	1	99

ground truth

$\therefore$ True Negative (TN) = 98

False Negative (FN) = 1

False Positive (FP) = 2

True positive (TP) = 99

a) Precision = $\frac{TP}{TP+FP} = \frac{99}{99+2} = \frac{99}{101} = 0.9802$

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

c) Type I error = False +ve = 2

d) Type II error = False -ve = 1.

(7) (a) $H_0: \theta = \theta_0$ true value, $\theta = \theta^*$

Now,

Pr(Type-II error) =

$Pr(\text{Accept } H_0 / H_0 \text{ is false})$

Since, H_0 is false $\rightarrow \theta \neq \theta_0$

thus, $\hat{\theta}$ will not be a Standard Normal

Assuming, $\hat{\theta}$ be asymptotic Normal, centered around true value (θ^*)

$\Rightarrow$ For H_0 is Accepted in wald's test, if

$$\Rightarrow |W| \leq z_{\alpha/2}$$

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

se = σ

true deviation

$$\Rightarrow \hat{\theta} \sim Nor(\theta_*, \sigma^2)$$

$$\Rightarrow \frac{\hat{\theta} - \theta_0}{se(\hat{\theta})} \sim Nor\left(\frac{\theta_* - \theta_0}{se(\hat{\theta})}, \frac{\cancel{\sigma^2}}{\cancel{(se(\hat{\theta}))^2}} \right) = Nor\left(\frac{\theta_* - \theta_0}{se(\hat{\theta})}, 1 \right)$$

$$\frac{\hat{\theta} - \theta_0}{se(\hat{\theta})} \sim \text{St Nor}(0,1) + \frac{\theta^* - \theta_0}{se(\hat{\theta})}$$

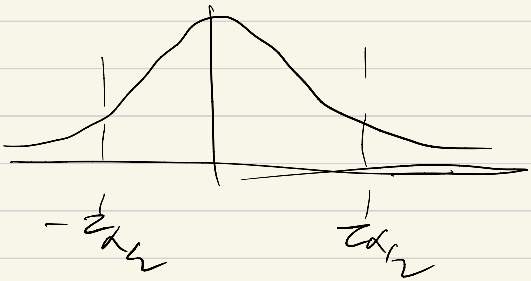
By equivariance, we can say that -

$$\frac{\hat{\theta} - \theta_0}{\hat{se}(\hat{\theta})} \sim Z + \frac{\theta^* - \theta_0}{\hat{se}(\hat{\theta})}$$

Now,

for H_0 to be accepted -

$|w| \le z_{\alpha/2}$

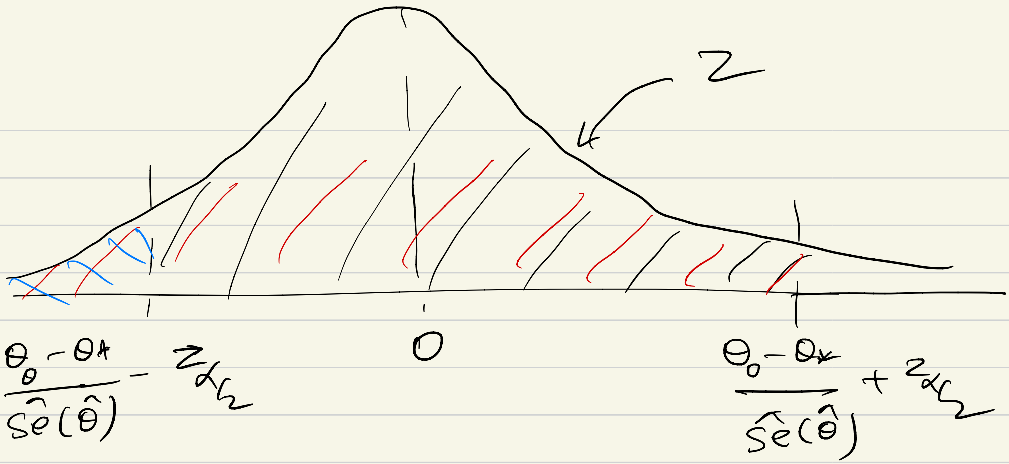


$$\left| \frac{\hat{\theta} - \theta_0}{\hat{se}(\hat{\theta})} \right| \le z_{\alpha/2}$$

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

$$\Rightarrow -z_{\alpha/2} \le Z + \frac{\theta^* - \theta_0}{\hat{se}(\hat{\theta})} \le z_{\alpha/2}$$

$$\Rightarrow \frac{\theta_0 - \theta^*}{\hat{se}(\hat{\theta})} - z_{\alpha/2} \le Z \le z_{\alpha/2} + \frac{\theta_0 - \theta^*}{\hat{se}(\hat{\theta})}$$



Thus, $Pr(\text{Type-II error})$ (Red Area − Blue Area)

$$= Pr\left(\frac{\theta_0 - \theta_*}{\hat{se}(\hat{\theta})} - z_{\alpha/2} \le Z \le \frac{\theta_0 - \theta_*}{\hat{se}(\hat{\theta})} + z_{\alpha/2} \right)$$

$$= Pr\left(Z \le \frac{\theta_0 - \theta_*}{\hat{se}(\hat{\theta})} + z_{\alpha/2} \right) - Pr\left(Z \le \frac{\theta_0 - \theta_*}{\hat{se}(\hat{\theta})} - z_{\alpha/2} \right)$$

$$= \Phi\left(\frac{\theta_0 - \theta_*}{\hat{se}(\hat{\theta})} + z_{\alpha/2} \right) - \Phi\left(\frac{\theta_0 - \theta_*}{\hat{se}(\hat{\theta})} - z_{\alpha/2} \right)$$

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

(b) H_0 : coin unbiased, $p = 0.5$ $\qquad$ $\underline{\underline{P_0 = 0.5}}$

H_1 : coin biased, $p \neq 0.5$

→ 46 successes in 100 trials

$$\hat{P}_{MLE} = \frac{\Sigma x_i}{n} = \frac{46}{100} = 0.46$$

∵ MLE estimators are AN, thus Wald's test can be applied.

$$w = \frac{\hat{P} - P_0}{\hat{se}(\hat{p})}$$

⇒ Numerator, $= \hat{P}_{MLE} - P_0$

$$= 0.46 - 0.5 = \underline{\underline{-0.04}}$$

⇒ Deno, $= \hat{se}(\hat{P}_{MLE})$

$$se(\hat{P}_{MLE}) = \sqrt{Var(\hat{P})}$$

$$= \sqrt{Var\left(\frac{\Sigma x_i}{n}\right)}$$

$$se(\hat{p}_{MLE}) = \sqrt{var\left(\frac{\sum x_i}{n}\right)}$$

$$\overset{iov}{=} \sqrt{\frac{\sum var(x_i)}{n^2}} = \frac{1}{n}\sqrt{\sum var(x_i)}$$

$$= \sqrt{\frac{n\, var(x_1)}{n^2}}$$

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

$$= \sqrt{\frac{p(1-p)}{n}} \Leftarrow \text{var of Bern coin}$$

By equivariance

$p \to \hat{p}$

$$\Rightarrow \quad \hat{se}(\hat{p}_{MLE}) = \sqrt{\frac{\hat{p}_{MLE}(1-\hat{p}_{MLE})}{n}} = \sqrt{\frac{0.46(1-0.46)}{100}}$$

$$= \sqrt{\frac{0.46 \times 0.54}{100}}$$

$$= \underline{0.0498}$$

$$w = \frac{-0.04}{0.0498} = -0.803$$

$$|w| = 0.803$$

$z_{\alpha/2}$ for $\alpha = 0.05$

$$z_{\alpha/2} = 1.96$$

Since, $\boxed{|w| < z_{\alpha/2} \quad 0.803 < 1.96}$

Thus, H_0 is true according to wald's test

i.e. coin is unbiased

from prev part

If, $H_0 : p = 0.7$, $H_1 : p \neq 0.7$, $\hat{se}(\hat{p}_{MLE}) = 0.0498$

$$|w| = \left| \frac{0.46 - 0.7}{0.0498} \right| = \left| \frac{-0.24}{0.0498} \right| = 4.81$$

$$|w| > z_{\alpha/2} \Rightarrow 4.81 > 1.96$$

Thus, if $P_0 = 0.7$, then H_0 will be rejected

thus, coin is biased.

Ans) 8|A| From Wald's test, we have

$H0 : \theta = \theta_0$ $H1 : \theta \neq \theta_0$,

If $W = \frac{\tilde{\theta} - \theta_0}{\tilde{se}(\tilde{\theta})}$,

$|W| > Z_{\alpha/2}$, Reject H0.

Here $\theta_0 = 0.5$.

$\alpha = 0.02$ ∴ $Z_{\alpha/2} = Z_{0.01} = 2.326$.

From data we have,

$n = 1000$. Total $= 540.948$.

Here, $\tilde{\theta}$ = Sample Mean $= \frac{540.948}{1000} = 0.540948$.

$$\therefore \tilde{se}(\tilde{\theta}) = \sqrt{Var(\tilde{\theta})}$$

$$= \sqrt{Var\left(\frac{\Sigma x_i}{n}\right)}$$

$$\tilde{se}(\tilde{\theta}) \overset{cov, \, iid}{=} \sqrt{\frac{1}{n} Var(X)}$$

Now, $Var(X) = \frac{1}{(n-1)} \Sigma(x_i - \bar{x})^2$ (corrected Sample Variance)

$$= \frac{10.6390873}{999}$$

$$\therefore Var(X) = 0.010649737.$$

$\therefore \hat{se}(\hat{\theta}) = \sqrt{\dfrac{0.010649737}{1000}}$

$= \sqrt{0.000010649737}$

$\therefore \hat{se}(\hat{\theta}) = 0.0032634031.$

Now, $W = \dfrac{0.540948 - 0.5}{0.0032634031}$

$\therefore W = 12.5476377711.$

Now, By Wald's test,

$|W| = 12.548 > Z_{\alpha/2}(2.326).$

$\therefore$ Reject H_0 i.e. $\theta \neq \theta_0$.

$\boxed{b)}$ Given, $X \sim Nor$, $Y \sim Nor$, $\alpha = 0.05$.

To test,

$H_0 : \theta_1 = \theta_2$ vs $H_1 : \theta_1 \neq \theta_2$

Population mean

i.e. $\delta = \mu_X - \mu_Y$

$H_0 : \delta = 0$ vs $H_1 : \delta \neq 0$.

From Wald's test for 2 population,

$$W = \frac{\hat{\delta}}{\hat{se}(\hat{\delta})}$$

$$= \frac{\bar{X} - \bar{Y}}{\sqrt{Var(\bar{X} - \bar{Y})}}$$

$$\tilde{se}(\tilde{\delta}) = \sqrt{Var(\bar{X} - \bar{Y})}$$

$$\overset{LOV}{=} \sqrt{Var(\bar{X}) + Var(\bar{Y})}$$

$$\overset{iid, LOV}{=} \sqrt{\frac{Var(X)}{m} + \frac{Var(Y)}{m}}, \quad \text{~~Hence~~} (m = n) \to 750$$

From Q.3 b), we have,

$$\bar{X} = \frac{\sum x_i}{m}, \quad Var(X) = \frac{1}{n} \sum (x_i - \bar{x})^2 .$$

$\to$ uncorrected.

∴ Substituting the values from data,

$$W = \frac{5.0048 - 5.8456}{\sqrt{\dfrac{\frac{1771.0946}{750} + \frac{4854.3183}{750}}{750}}}$$

$$= \frac{-0.8408}{\sqrt{\dfrac{\frac{1771.0946}{750}}{750} + \dfrac{\frac{4854.3183}{750}}{750}}}$$

$$= \frac{-0.8408}{\sqrt{0.011769}} = \frac{-0.8408}{0.10849}$$

$$\therefore \boxed{W = 7.7500}$$

→ Kindly note, this value might differ in decimals because I have rounded values for calculation.

$\therefore \quad |W| = 7.7500 > Z_{\alpha/2} \, (Z_{0.025} = 1.96)$.

$\therefore$ Ho is rejected and θ_1 (Sample Mean X) and θ_2 (Sample Mean Y) are not same.

Wald's test is applicable here because :-

i) $\hat{\theta}_{1_{MLE}}$ (Sample Mean X) and $\hat{\theta}_{2_{MLE}}$ (Sample Mean Y) are asymtotically normal by CLT.

ii) X and Y are independent Normal Distributions.