RANJAN KUMAR - 113262786
SHUBHAM AGRAWAL - 113166701
AMEYA SANKHE - 113219562
PRATIK NAGELIA - 114122014

CSE 544, Spring 2021, Probability and Statistics for Data Science

Assignment 4: Parametric Inference & Hypothesis Testing Due: 4/08, 1:15pm, via Blackboard
(8 questions, 70 points total)

I/We understand and agree to the following:
(a) Academic dishonesty will result in an 'F' grade and referral to the Academic Judiciary.
(b) Late submission, beyond the 'due' date/time, will result in a score of 0 on this assignment.

(write down the name of all collaborating students on the line below)

RANJAN KUMAR, SHUBHAM AGRAWAL, AMEYA SANKHE, PRATIK NAGELIA

1. Practice with MME **(Total 9 points)**

(a) The Gamma(x, y) distribution has mean $x \cdot y$ and variance $x \cdot y^2$. Find MME for $\hat{x}$ and $\hat{y}$. (4 points)

(b) Find MME $\hat{a}$ and $\hat{b}$ for the Uniform(a, b) distribution. Express your final answer in terms of the sample mean, $\bar{X} = (\sum X_i)/n$, and sample variance, $\overline{S^2} = ((\sum X_i^2)/n) - \bar{X}^2$. (5 points)

1. (a) Ans. Mean $= x \cdot y$, Variance $= xy^2$

Mean $= E[X] = x \cdot y$

$$\hat{E}[X] = \frac{1}{n}\sum X_i$$

As $\hat{E}[X] = E[X]$, so we have :-

$$\frac{1}{n}\sum X_i = xy$$

$$\text{or, } \hat{x} = \frac{\sum X_i}{n \cdot y} \quad \text{--- (1)}$$

$$\text{Variance} = E[X^2] - (E[X])^2$$

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

$$\text{or, } E[X^2] = xy^2 + \left(\frac{\sum X_i}{n}\right)^2$$

Since $E[X^2] = \hat{E}[X^2]$

$$\text{or, } xy^2 + \left(\frac{\sum X_i}{n}\right)^2 = \frac{1}{n}\sum X_i^2$$

$$\text{or, } y \cdot \left(\frac{\sum X_i}{n}\right) + \frac{(\sum X_i)^2}{n^2} = \frac{\sum X_i^2}{n}$$

or, $\sum X_i y = \sum X_i^2 - \frac{(\sum X_i)^2}{n}$

or, $\hat{y} = \frac{\sum X_i^2}{\sum X_i} - \frac{\sum X_i}{n}$ —— (2) {Required Answer}

Putting the value of Equation (2) in Equation (1), we get :-

$$\hat{n} = \frac{\sum X_i}{yn} = \frac{\sum X_i}{n \cdot \left(\frac{n \sum X_i^2 - \sum X_i}{n \cdot \sum X_i}\right)} = \frac{(\sum X_i)^2}{n \sum X_i^2 - \sum X_i}$$ {Required Answer}

(b) Ans. Mean $= E(X) = \frac{a+b}{2}$; $\hat{E}(X) = \frac{1}{n}\sum(X_i)$

Since $E(X) = \hat{E}(X) \Rightarrow \frac{a+b}{2} = \frac{1}{n}\sum X_i = \bar{X}$ —— (1)

Variance $[X] = E(X^2) - (E(X))^2$

or, $\frac{(b-a)^2}{12} = E(X^2) - \left(\frac{\sum X_i}{n}\right)^2 \Rightarrow E(X^2) = \frac{(b-a)^2}{12} + \left(\frac{\sum X_i}{n}\right)^2$

or, $\frac{\sum X_i^2}{n} = \frac{(b-a)^2}{12} + \left(\frac{\sum X_i}{n}\right)^2 \Rightarrow \frac{b^2 - 2ab + a^2}{12} = \frac{\sum X_i^2}{n} - \left(\frac{\sum X_i}{n}\right)^2$

or, Put $a = 2\bar{X} - b$ (from Equation (1)), we get :-

or, $\frac{b^2 + 4(\bar{x})^2 + b^2 - 4\bar{x}b - 4\bar{X}b + 2b^2}{12} = \bar{S}^2$

or, $\frac{4b^2 + 4\bar{x}^2 - 8\bar{X}b}{12} = \bar{S}^2$, or, $\frac{b^2 + \bar{X}^2 - 2\bar{X}b}{3} = \bar{S}^2$

or, $b^2 - 2\bar{X}b + \bar{X}^2 - 3\bar{S}^2 = 0$, or, $b = \frac{2\bar{X} \pm \sqrt{4\bar{x}^2 - 4\bar{x}^2 + 12\bar{S}^2}}{2}$

or, $\hat{b} = \frac{2\bar{X} \pm 2\sqrt{3}\,\bar{S}}{2} = \bar{X} \pm \sqrt{3}\,\bar{S}$ (~~Required Answer~~)

Therefore, $a = 2\bar{X} - b$ or, $\hat{a} = 2\bar{X} - \bar{X} \mp \sqrt{3}\bar{S}$

$= \bar{X} \mp \sqrt{3} \cdot \bar{S}$

(~~Required Answer~~)

In uniform distribution, as $a < b$

Therefore $\hat{a} = \overline{X} - \sqrt{3} \cdot \overline{S}$

$$\hat{b} = \overline{X} + \sqrt{3} \cdot \overline{S}$$

Q2) Pdf of exponential distribution is defined as :-

$$f(x, \lambda) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

Likehood function $L(\lambda, x_1, x_2 \ldots x_n) = \prod_{i=1}^{n} f(x_i, \lambda)$

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

For MLE, $\frac{dL}{d\lambda} = 0 \Rightarrow \frac{d \ln(L(\lambda, x_1, x_2 \ldots x_n))}{d\lambda} = 0$

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

$$\Rightarrow \frac{d \ln\left(n \ln(\lambda) - \lambda \sum_{i=1}^{n} x_i\right)}{d\lambda} = 0$$

$$\Rightarrow \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0$$

$$\Rightarrow \frac{n}{\lambda} = \sum_{i=1}^{n} x_i$$

In the question, $\lambda = \frac{1}{\beta}$, putting it in the above equation, we get :- $n\beta = \sum_{i=1}^{n} x_i$

$$\Rightarrow MLE(\hat{\beta}) = \frac{\sum_{i=1}^{n} x_i}{n}$$; As $n \to \infty$, $MLE(\hat{\beta}) \to 0$

The consistency ... sequence of random variable contd.

2. Consistency of MLE **(Total 6 points)**

Let $X_1, X_2, \ldots, X_n$ be distributed as Exponential(1/β), all i.i.d. Show that the MLE($\hat{\beta}$) will converge to the unknown parameter β. Prove this by showing that bias($\hat{\beta}$) and se($\hat{\beta}$) tends to 0 as n tends to ∞. You can use the fact that the mean and variance of Exponential(λ) are 1/λ and 1/λ², respectively.

Ans. Let $D = \{X_1, X_2, X_3 \ldots X_n\} \overset{iid}{\sim} Exp\left(\frac{1}{\beta}\right)$

Mean: $E(X) = \beta$

$\hat{E}(X) = \frac{1}{n} \sum X_i$

Since $\hat{E}(X) = E(X)$

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

or, $\hat{\beta}_{MLE} = \frac{1}{n} \sum X_i$

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

$= E\left[\frac{1}{n} \sum X_i\right] - \beta = \frac{1}{n} \cdot \sum E(X_i) - \beta$

$= \frac{1}{n} \times n \cdot \beta - \beta = 0$ (Proved)

$se(\hat{\beta}) = \sqrt{Var(\hat{\beta})} = \sqrt{Var\left(\frac{1}{n} \cdot \sum X_i\right)}$

$= \sqrt{\frac{1}{n^2} \cdot \sum Var(X_i)} = \sqrt{\frac{1}{n^2} \times n \times Var(X)}$

$= \sqrt{\frac{Var(X) \times n}{n^2}} = \sqrt{\frac{\beta^2}{n}}$

Ans As $bias(\hat{\beta}_{MLE})$ and $se(\beta)$ tends to 0 as n tends to ∞. Therefore $\hat{\beta}_{MLE}$ converges to β.

$\therefore$ As $n \to \infty$, $se(\hat{\beta}) \to 0$. (Proved).

3. Practice with MLE **(Total 10 points)**

(a) Let $X_1, X_2, \ldots, X_n$ be distributed i.i.d. as Poisson(λ). Find the MLE of λ. (3 points)

(b) Let $X_1, X_2, \ldots, X_n$ be distributed i.i.d. as Normal(μ, σ²). Show that the MLE of μ and σ² is the same as the sample mean and (uncorrected) sample variance, respectively. (4 points)

(c) Let $X_1, X_2, \ldots, X_n \sim$ Normal(θ, 1). Let $\delta = E[I_{X_1>0}]$. Use the Equivariance property to show that the MLE of δ is $\Phi(\frac{1}{n}\sum_{i=1}^{n} X_i)$, where $\Phi()$ is the CDF of the standard Normal. You can use the MLE of the Normal as provided in 3(b). (3 points)

a) $D = \{X_1, X_2 \ldots\ldots, X_n\} \overset{iid}{\sim} X \sim \text{Poisson}(\lambda)$

We know P.d.f of Poisson is

$$P_X(x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad \text{where } x = 0, 1, 2, \ldots\ldots$$

Now let liklihood be $L(\lambda)$

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

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

Taking log on both sides

$$\log \mathcal{L}(\lambda) = \log \prod_{i=1}^{n} \left(\frac{\lambda^{X_i} e^{-\lambda}}{X_i!} \right)$$

$$= \sum_{i=1}^{n} \log \left(\frac{\lambda^{X_i} e^{-\lambda}}{X_i!} \right)$$

$$l(\lambda) = \log L(\lambda) = \sum_{i=1}^{n} \left(\log \lambda^{x_i} + \log e^{-\lambda} - \log x_i! \right)$$

$\ldots \log\left(\frac{ab}{c}\right) = \log a + \log b - \log c$

$$= \sum_{i=1}^{n} \left(x_i \log \lambda - \lambda - \log x_i! \right)$$

$\ldots \log_a b^c = c \log_a b$

$\log e^a = a$

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

$$l(\lambda) = (\log \lambda) \sum_{i=1}^{n} x_i - n\lambda - \sum_{i=1}^{n} \log x_i!$$

Now differentiate $l(\lambda)$ w.r.t λ and equate it to zero

$$\frac{d\, l(\lambda)}{d\lambda} = \frac{\sum_{i=1}^{n} x_i}{\lambda} - n = 0$$

$$\therefore n = \frac{\sum_{i=1}^{n} x_i}{\lambda}$$

$$\therefore \lambda = \frac{\sum_{i=1}^{n} x_i}{n}$$

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

b) $D = \{X_1, X_2, \dots, X_n\} \overset{iid}{\sim} X \sim Normal(\mu, \sigma^2)$

We know Pdf of Normal is

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

Now let liklihood be $L(\lambda)$

$$L_{\mu,\sigma^2}(\cancel{\lambda}) = \prod_{i=1}^{n} P_X(X_i)$$

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

Taking log on both sides

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

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

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

$\dots \log\left(\frac{ab}{c}\right) = \log a + \log b - \log c$

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

$\dots \log e^b = b$

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

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

$$= l(\mu,\sigma^2)$$

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

partially differentiate $l(\mu,\sigma^2)$ w.r.t μ and equate it to zero

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

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

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

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

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

$$\mu = \frac{\sum_{i=1}^{n} X_i}{n}$$

$\therefore \hat{\mu}_{MLE} = \frac{\sum_{i=1}^{n} X_i}{n}$	$\therefore \hat{\mu}_{MLE}$ is same as sample Mean

Now Partially differentiate $l(\mu,\sigma^2)$ w.r.t σ and equate it to 0

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

$\ldots \frac{\partial \log x}{\partial x} = \frac{1}{x}$

$\frac{\partial x^n}{\partial x} = nx^{n-1}$

when $n=-2$

$= -2x^{-3}$

$$\frac{n}{\sigma} = \frac{\cancel{2}}{\cancel{2}\sigma^{\cancel{3}2}}\sum_{i=1}^{n}(x_i-\mu)^2$$

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

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

$\therefore$ MLE of σ^2 ($\hat{\sigma}^2_{MLE}$) is uncorrected Sample Variance.

C) Now we know

$\delta = E[I_{x_1 > 0}]$ where I is Indicator RV

as $X_1, X_2 \dots X_n \overset{iid}{\sim} \text{Normal}(\theta, 1) \sim X$

$\therefore$ We can Write $\delta = E[I_{x > 0}]$

We ~~kno~~ know $E[I_{x>0}]$ is $P(X>0)$

$$\therefore \delta = E[I_{x>0}]$$
$$= P(X>0) \quad \text{——— (1)}$$

To Convert X to Standard Normal

$Z = X - \theta$ as $X \sim \text{Normal}(\theta, 1)$

where Z is Standard Normal

$$\therefore X = Z + \theta$$

Substituting the above value in (1)

We get

$$\delta = P((Z+\theta) > 0)$$

$$\delta = 1 - P((Z+\theta) \leq 0) \quad \because P(Y>0) + P(Y \leq 0) = 1$$

$\therefore \delta = 1 - P(Z \leq -\theta)$

We know $\Phi()$ is CDF of standard Normal

$\therefore P(\cancel{Y} Z \leq x) = \Phi(x)$

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

we know, $\Phi(\theta) + \Phi(-\theta) = 1$

$\therefore \delta = \Phi(\theta)$

$\therefore$ Let $g(\theta) = \delta = \Phi(\theta)$

We know $\hat{\theta}_{MLE} = \text{MLE of } \mu$

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

$\therefore$ ~~By equi~~ As $\hat{\theta}_{MLE}$ is equivariant

We know

$$g(\hat{\theta}_{MLE}) = \text{MLE of } g(\theta)$$

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

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

$$\boxed{\begin{aligned} \therefore \text{MLE of } \delta &= \hat{\delta} \\ &= \Phi\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) \end{aligned}}$$

4. Parametric Inference with Data Samples **(Total 10 points)**

Let $X = \begin{cases} 2 & \text{with prob } \theta \\ 3 & \text{otherwise} \end{cases}$, where θ is unknown. Let D = {2, 3, 2} be drawn i.i.d. from X.

(a) Derive $\hat{\theta}_{MME}$ using D as the sample data. Clearly show all your steps. (3 points)

(b) Provide a numerical estimate of the 95%ile confidence intervals for $\hat{\theta}_{MME}$. Start by deriving $\widehat{se}(\hat{\theta}_{MME})$: first derive $se(\hat{\theta}_{MME})$ in terms of θ, and then estimate $\widehat{se}(\hat{\theta}_{MME})$, as in class. Show all your steps. Your final answer should be a numerical range. (4 points)

(c) Derive $\hat{\theta}_{MLE}$ using D as the sample data. Clearly show all your steps. (3 points)

4(a) $X = \begin{cases} 2 & \text{with prob}\,\theta \\ 3 & \text{otherwise} \end{cases}$

$= \begin{cases} 2 & \text{with } P\theta \\ 3 & \text{with } P(1-\theta) \end{cases}$

Expectation of the above distribution:

$E[X] = \sum_{i=1}^{n} x\, P_X(x)$

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

$\hat{\theta}_{mmE} = \frac{\sum X_i}{n}$ — ②

Putting values in ① and ②

$3 - \hat{\theta}_{mmE} = \frac{2+3+2}{3}$

$\therefore \hat{\theta}_{mmE} = -\frac{7}{3} + 3 = \frac{2}{3}$ <u>Ans</u>.

(b) $se(\hat{\theta}_{mmE}) = \sqrt{Var(3-\bar{X})}$ where $\bar{X} = \frac{\sum X_i}{n}$

$= \sqrt{Var(3) + Var(\bar{X})}$

$= \sqrt{Var\left(\frac{\sum X_i}{n}\right)} \overset{LOV}{=} \sqrt{\frac{1}{n^2} Var(\sum X_i)}$

$\overset{iid}{=} \sqrt{\frac{1}{n^2} \times n\, Var(X_i)} = \sqrt{\frac{Var(X_i)}{n}}$

$$\therefore \hat{se}(\hat{\theta}_{mme}) = \sqrt{\frac{\hat{Var}(x)}{n}}$$

$$\text{Now } Var(x) = E[x^2] - (E[x])^2 \quad \text{--- (3)}$$

$$E[x^2] = \sum x^2 P(x)$$
$$= 2^2 \times \theta + 3^2(1-\theta) = 4\theta + 9 - 9\theta = 9 - 5\theta$$

$$(E[x])^2 = (3-\theta)^2 = 9 + \theta^2 - 6\theta \quad [\text{From (1)}]$$

$$\cancel{\hat{Var}(x) = \hat{\theta}_{mme} - \hat{\theta}^2_{mme}}$$

$$Var(x) = 9 - 5\theta - (9 + \theta^2 - 6\theta)$$
$$= \theta - \theta^2$$

$$\therefore \hat{Var}(x) = \hat{\theta}_{mme} - \hat{\theta}^2_{mme} = \hat{\theta}_{mme}(1 - \hat{\theta}_{mme})$$

$$\therefore \hat{se}(\hat{\theta}_{mme}) = \sqrt{\frac{\hat{Var}(x)}{n}} = \sqrt{\frac{\hat{\theta}_{mme}(1-\hat{\theta}_{mme})}{n}}$$

Putting the value of $\hat{\theta}_{mme} = 2/3$ from 4(a)

$$\hat{se}(\hat{\theta}_{mme}) = \sqrt{\frac{2/3(1-2/3)}{3}} = \sqrt{\frac{2/3 \times 1/3}{3}} = \sqrt{\frac{2}{27}}$$

Numerical estimate of 95% ile confidence intervals

$$\Rightarrow \hat{\theta}_{mme} \pm 1.96 \times \hat{se}_{mme} \qquad [Z_{\alpha/2} = 1.96]$$

$$\Rightarrow \frac{2}{3} \pm 1.96 \times \sqrt{\frac{2}{27}} \qquad [\text{Putting the values of } \hat{\theta}_{mme} \text{ \& } \hat{se}_{mme}]$$

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

$$\Rightarrow \frac{2}{3} \pm 0.65\sqrt{\frac{2}{3}} \qquad \text{Ans.}$$

$$\Rightarrow (0.1332,\ 1.2001)$$

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

Now, finding the closed form:

$$P_X(x) = \theta^{ax+b} (1-\theta)^{1-(ax+b)}$$

When $x = 2$;

$2a + b = 1$ — Eqn. (1)

When $x = 3$;

$3a + b = 0$ — Eqn. (2)

Solving eqn. (1) and (2)

$$\begin{aligned} 2a + \cancel{b} &= 1 \\ 3a + b &= 0 \\ - \quad - \quad &\quad - \\ \hline -a &= 1 \end{aligned}$$

$\therefore \quad a = -1$

$\therefore \quad b = 1 - (2 \times -1) = 1 + 2 = 3$

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

Likelihood, $L(\hat{\theta})_{MLE} = \prod_{i=1}^{n} P_X(x_i)$

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

Taking log on both sides

$$\log L(\theta) = \sum_{i=1}^{n} (3-x_i) \log \theta + \sum_{i=1}^{n} (x_i - 2) \log(1-\theta)$$

$$\frac{d(L(\theta))}{d\theta} = 3n - \Sigma X_i \times \frac{1}{\theta} + (\Sigma X_i - 2n) \times \frac{1}{1-\theta} \times -1$$

$$= \frac{(3n - \Sigma X_i)}{\theta} - \frac{(\Sigma X_i - 2n)}{1-\theta} \quad \text{— Eq}^n \ (3)$$

In order to find MLE: $\frac{d(L(\theta))}{d\theta} = 0$

$\therefore$ Equating eq^n (3) to 0

$$\frac{(\Sigma X_i - 2n)}{1-\theta} = \frac{(3n - \Sigma X_i)}{\theta}$$

$$\cancel{\theta \times \Sigma X_i} - 2n\theta = 3n - \Sigma X_i - 3n\theta + \cancel{\theta \times \Sigma X_i}$$

$$n\theta = 3n - \Sigma X_i$$

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

$$= 3 - \frac{\Sigma X_i}{n}$$

$$= 3 - \frac{2+3+2}{3}$$ [putting sample data]

$$\hat{\theta}_{MLE} = 3 - \frac{7}{3} = \frac{2}{3}$$ Ans.

5. MME versus MLE using real data **(Total 10 points)**

For this question, we will use the acceleration, model, and mpg data from the Auto-mpg dataset (https://www.kaggle.com/uciml/autompg-dataset). Please use the data files on the class website. We will assume that acceleration is Normal(μ, σ^2) distributed, model year is Uniform(a, b) distributed, and mpg is Exponential(λ) distributed. You are to find the MME and MLE estimates of the parameters of the distributions for all 3 datasets. For the Normal MME and Uniform MLE, you can directly use the results from class. For the Normal MLE, use the result from Q3(b); for Uniform MME, use the result from Q1(b). For the Exponential, we will first derive the estimates.

(a) For the Exp(λ) distribution, find the $\hat{\lambda}_{MME}$. (2 points)

(b) For the Exp(λ) distribution, find the $\hat{\lambda}_{MLE}$. (2 points)

(c) For the 3 datasets, find the MME estimates. That is, find the MME for μ and σ^2 for the acceleration dataset, a and b for the model dataset, and λ for the mpg dataset. Provide your answer as a number with 3 significant digits. (3 points)

(d) Same as part (c), but this time find the MLE estimates. (3 points)

5(a) For the exponential distribution:

$$E[Exp(\lambda)] = \frac{1}{\lambda} \quad \text{— (1)}$$

Sample Mean of the dataset $= \frac{\sum X_i}{n}$ — (2)

Equating (1) and (2) to get $\hat{\lambda}_{mme}$

$$\frac{\sum X_i}{n} = \frac{1}{\hat{\lambda}_{mme}}$$

$$\boxed{\therefore \hat{\lambda}_{mme} = \frac{n}{\sum X_i}}$$ Ans.

(b) Likelihood function

$$L(\lambda, x_1, x_2 \ldots x_n) = \prod_{i=1}^{n} f(x_i, \lambda)$$

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

For MLE, $\frac{dl}{d\lambda} = 0$

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

$$\Rightarrow \frac{d \ln\left(n \ln \lambda - \lambda \sum_{i=1}^{n} x_i\right)}{d\lambda} = 0$$

$$\Rightarrow \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0$$

$$\Rightarrow \frac{n}{\lambda} = \sum_{i=1}^{n} x_i$$

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

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

(c) $\hat{\mu}_{MME} = 15.6$

$\hat{\sigma}^2_{MME} = \cancel{250}\ 7.57$

$\hat{a}_{MME} = \cancel{75.7}\ 69.6$

$\hat{b}_{MME} = \cancel{76.3}\ 82.4$

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

(d) $\hat{\mu}_{MLE} = 15.6$

$\hat{\sigma}^2_{MLE} = \cancel{250}\ 7.57$

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

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

$\hat{\lambda}_{MLE} = \cancel{23.5}\ 0.0425$

6. Clinical Testing **(Total 4 points)**

Consider the sick patient example from class. In a clinical trial of a new disease detection test, there were 100 healthy patients and 100 sick patients. The test correctly identified 98 out of the 100 healthy patients as healthy. The test also correctly identified 99 of the 100 sick patients as sick. The remaining patients were incorrectly classified.

(a) What is the precision of the test? (1 point)

(b) What is the recall of the test? (1 point)

(c) What is the Type I error of the test? (1 point)

(d) What is the Type II error of the test? (1 point)

Ans:- Let Ho means the hypothesis that Patient is healthy.

As per definition,

True Positives

= Test Rejects Ho | patient is sick

= 99

Test Output →

Ground truth ↓	Accept	Reject
Ho true	98 True -ve	2 False +ve
Ho false.	1 False -ve	99 True +ve

False Positives

= Test Reject Ho | Ho is true

= Test Reject Ho | Patient is Healthy = 2

False Negative = Test Accepts Ho | Ho false = Test Says Healthy | patient is sick

= 100 - 99 = 1 Ans.

Hence, we draw the truth table as above.

(1) Precision = $\frac{\text{True Positive}}{\text{True Positive + False Positive}} = \frac{99}{99+2} = \frac{99}{101} = 0.98$

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

(3) Type I error = $2/100 = 0.02$ = False Positive.

(4) Type II error = $\frac{1}{100} = 0.01$ = False Negative.

7. Wald's test **(Total 11 points)**

(a) Suppose the null hypothesis is H₀: θ = θ₀, but the true value of θ is θ$_*$. Show that, under Wald's test, the probability of a Type II error is $\Phi(\frac{\theta_0-\theta_*}{\widehat{se}} + z_{\alpha/2}) - \Phi(\frac{\theta_0-\theta_*}{\widehat{se}} - z_{\alpha/2})$.

(Hints: (i) might help to draw a figure; (ii) think about the distribution of the estimate.) (6 points)

(b) You observe 46 successes in 100 trials of a coin. If the null hypothesis is that the coin is unbiased, use the Wald's test with the MLE or MME with α = 0.05 to Reject/Accept the null. What if the null hypothesis is that the coin has p=0.7? (5 points)

We know,

a) $$W = \frac{\hat{\theta} - \theta_0}{\hat{se}(\hat{\theta})}$$

We know True Value of $\theta = \theta_*$

$\therefore \theta \neq \theta_0$ and the Null hypothesis is rejected

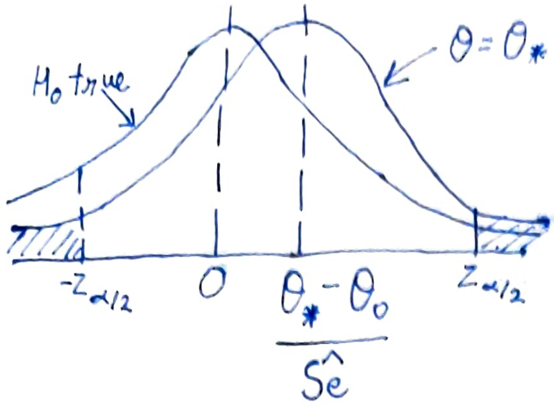
Now $\hat{\theta}$ is Asymptotically Normal

$\therefore \hat{\theta} \sim \text{Normal}(\theta_*, se^2)$

$$\therefore \frac{\hat{\theta} - \theta_0}{se(\hat{\theta})} \sim \text{Normal}\left(\frac{\theta_* - \theta_0}{se(\hat{\theta})}, 1\right)$$

$\hat{se}(\hat{\theta})$ is a Consistent estimator of $se(\hat{\theta})$

$$\therefore \frac{\hat{\theta} - \theta_0}{\hat{se}(\hat{\theta})} \sim \text{Normal}\left(\frac{\theta_* - \theta_0}{\hat{se}(\hat{\theta})}, 1\right)$$



$$
\begin{aligned}
P(\text{Type II error}) &= P(\text{Accept } H_0 / H_0 \text{ is False}) \\
&= P(|W| \le Z_{\alpha/2} / \theta \ne \theta_0) \\
&= P(-Z_{\alpha/2} \le W \le Z_{\alpha/2} / \theta \ne \theta_0) \\
&= P(-Z_{\alpha/2} \le W \le Z_{\alpha/2} / \theta = \theta_*) \\
&= P\left(-Z_{\alpha/2} \le \frac{\hat{\theta} - \theta_0}{\hat{se}} \le Z_{\alpha/2}\right) \\
&= P\left(\frac{\hat{\theta} - \theta_0}{\hat{se}} \le Z_{\alpha/2}\right) - P\left(\frac{\hat{\theta} - \theta_0}{\hat{se}} \le -Z_{\alpha/2}\right)
\end{aligned}
$$

To Convert $\dfrac{\hat{\theta} - \theta_0}{\hat{se}}$ to standard Normal

$$Z = \frac{\hat{\theta} - \theta_0}{\hat{se}} - \left(\frac{\theta_* - \theta_0}{\hat{se}}\right)$$ where Z is Standard Normal

as $\dfrac{\hat{\theta} - \theta_0}{\hat{se}} \sim \text{Normal}\left(\dfrac{\theta_* - \theta_0}{\hat{se}}, 1\right)$

Now, Putting the Values

$$= P\left(Z + \left(\frac{\theta_* - \theta_0}{\hat{Se}}\right) \le Z_{\alpha/2}\right) - P\left(Z + \left(\frac{\theta_* - \theta_0}{\hat{Se}}\right) \le -Z_{\alpha/2}\right)$$

$$= P\left(Z \le Z_{\alpha/2} - \left(\frac{\theta_* - \theta_0}{\hat{Se}}\right)\right) - P\left(Z \le -Z_{\alpha/2} - \left(\frac{\theta_* - \theta_0}{\hat{Se}}\right)\right)$$

$$= P\left(Z \le Z_{\alpha/2} + \left(\frac{\theta_0 - \theta_*}{\hat{Se}}\right)\right) - P\left(Z \le -Z_{\alpha/2} + \left(\frac{\theta_0 - \theta_*}{\hat{Se}}\right)\right)$$

We know Φ is CDF of Standard Normal

$$P(Z \le x) = \Phi(x)$$

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

$\therefore$ P(Type II error)

$$\boxed{= \Phi\left(\frac{\theta_0 - \theta_*}{\hat{Se}} + Z_{\alpha/2}\right) - \Phi\left(\frac{\theta_0 - \theta_*}{\hat{Se}} - Z_{\alpha/2}\right)}$$

7(b) Given: 46 success in 100 trials of a coin toss

$$\hat{P}_{MLE} = \frac{46}{100} = 0.46 \quad \text{— (1)}$$

(i) Null Hypothesis: $H_0 : p = 0.5$, vs $H_1 : p \neq 0.5$

Since, MLE is Asymptotically Normal, Wald's test applies

According to Wald's test $|W| > z_{\alpha/2}$, reject H_0, null hypothesis

$z_{\alpha/2} = 1.96$ [Given $\alpha = 0.05$]

$$W = \frac{\hat{\theta} - \theta_0}{\hat{se}(\hat{\theta})} = \frac{\hat{\theta}_{MLE} - \theta_0}{\hat{se}(\hat{\theta}_{MLE})}$$

Numerator: $\hat{\theta}_{MLE} - \theta_0 = 0.46 - 0.5 = -0.04$ [From 1] — (2)

Denominator: $\hat{se}(\hat{\theta}_{MLE})$

$$se(\hat{p}) = \sqrt{Var(\hat{p})} = \sqrt{Var\left(\frac{\sum X_i}{n}\right)} \underset{IID}{\overset{Lov}{=}} \sqrt{\frac{Var(X_i)}{n}}$$

$$= \sqrt{\frac{p(1-p)}{n}} \quad \text{[Variance of Bernoulli]}$$

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

$$= 0.0498 \quad \text{— (3)}$$

$$\therefore W = \frac{-0.04}{0.0498} = -0.8 \quad \text{[Dividing (2) and (3)]}$$

$\therefore |W| < z_{\alpha/2}$, we accept the Hypothesis

(ii) Now, when Null hypothesis, $H_0 : p = 0.7$ vs $H_1\ p \neq 0.7$

$$W = \frac{0.46 - 0.7}{0.0498}$$ [Taking value ~~from~~ for denominator from (3)]

$$\therefore W = \frac{-0.24}{0.0498} = -4.82$$

$\therefore |W| > Z_{\alpha/2}$

$\therefore$ We reject the Hypothesis : $H_0 p \neq 0.7$

When null hypothesis is $p = 0.5$, we accept the hypothesis

When null hypothesis is $p = 0.7$, we reject the hypothesis

8. More on Wald's test **(Total 10 points)**

(a) Use q8_a.csv dataset and assume it is distributed as Normal(θ, σ^2). Apply the Wald's test with α = 0.02 to check whether the true mean is $\theta_0 = 0.5$. Use sample mean to obtain $\hat{\theta}$ and corrected sample variance estimator for obtaining $\widehat{\sigma^2}$. (4 points)

(b) Use q8_b_X.csv and q8_b_Y.csv available at the class website for this question. Each contains 750 samples for X and Y drawn from two independent Normal distributions. Without worrying about the applicability of the test, use Wald's 2-population test with α = 0.05 to test whether the population means of X and Y are same (null) or not (alternative). Is this test applicable here? (6 points)

8. (a) Hypothesis H_0 is that θ_0 is equal to 0.5.

$|W| = 12.5476$

$Z_{\alpha/2} = 2.32$ $\because |W| > Z_{\alpha/2}$

$\therefore$ Hypothesis is rejected.

(b) Hypothesis H_0 is mean of X ~~and~~ and Y dist. is same.

$|W| = 7.748$

$Z_{\alpha/2} = 1.96.$

$\therefore$ Hypothesis is rejected.

~~The test is applicable here, since by we assume that.~~

~~X and Y are~~

Yes, The test is applicable here since We assume the distributions are Asymptotic Normal, ~~and~~ the estimators are consistent, and X and Y are normal Distribution.