

1 Question 1

1.1 MME for \hat{x} and \hat{y}

First we know that, for Gamma distribution $Gamma(x, y)$

$$E[X] = xy \text{ and } Var(X) = xy^2$$

Equating the expectation and variance with the corresponding sample moments, we get:

$$\begin{aligned} E[X] &= \frac{1}{n} \sum_{i=1}^n X_i = \bar{X} = xy \\ E[X^2] &= \frac{1}{n} \sum_{i=1}^n X_i^2 \\ Var[X] &= E[X^2] - (E[X])^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \left(\frac{1}{n} \sum_{i=1}^n X_i\right)^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = xy^2 \end{aligned}$$

For the first equation, we have

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

Now, substituting it into the second equation, we get:

$$xy^2 = \bar{X}y = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

Now, solving for y in that last equation and we get that the MME for y ,

$$\hat{y} = \frac{1}{n\bar{X}} \sum_{i=1}^n (X_i - \bar{X})^2$$

Similarly, we could have the MME for x ,

$$\hat{x} = \frac{\bar{X}}{\hat{y}} = \frac{n\bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

1.2 MME for \hat{a} and \hat{b}

First we know that, for Uniform distribution $Uniform(a, b)$

$$E[X] = \frac{a+b}{2} \text{ and } Var(X) = \frac{(b^2 - a^2)}{12}$$

From the first moment, we get:

$$E[X] = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X} = \frac{a+b}{2}$$

For the above equation, we have

$$a = 2\bar{X} - b$$

From the second moment we get,

$$E[X^2] = \frac{1}{n} \sum_{i=1}^n X_i^2$$

Using first and second moment, variance can be obtained as follows

$$Var[X] = E[X^2] - (E[X])^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \left(\frac{1}{n} \sum_{i=1}^n X_i\right)^2 = \bar{S}^2 = \frac{(b-a)^2}{12}$$

In above statement, we can figure out the +/- based on the fact that $b > a$ in uniform distribution. Using above equation and taking square root:

$$b = \sqrt{12\bar{S}} + a$$

Now, substituting a into the above equation, we get MME for b :

$$\hat{b} = \sqrt{12\bar{S}} + 2\bar{X} - \hat{a}$$

$$2\hat{b} = 2\sqrt{3\bar{S}} + 2\bar{X}$$

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

Similarly, we could have the MME for a ,

$$\hat{a} = 2\bar{X} - \hat{b} = \bar{X} - \sqrt{3\bar{S}}$$

Q2.

Given:

$$X_1, X_2, X_3, X_4, \dots, X_n \sim \text{Exponential}\left(\frac{1}{\beta}\right)$$

Step 1: Likelihood

$$\begin{aligned} L(p) &= \prod_{i=1}^n \frac{1}{\beta} e^{-\frac{1}{\beta} X_i} \\ L(p) &= \frac{1}{\beta^n} e^{-\frac{1}{\beta} \sum_i X_i} \end{aligned}$$

Step 2: Taking Log both sides

$$\begin{aligned} \log(L(p)) &= \log\left(\frac{1}{\beta^n} e^{-\frac{1}{\beta} \sum_i X_i}\right) \\ \log(L(p)) &= -n\log(\beta) - \frac{1}{\beta} \sum_i X_i \end{aligned}$$

Step 3: Differentiate it w.r.t β and equate it to 0 to get MLE

$$\begin{aligned} \Rightarrow -\frac{n}{\beta} + \beta^{-2} \sum_i X_i &= 0 \\ \Rightarrow \hat{\beta} &= \frac{\sum_i X_i}{n} \end{aligned}$$

To show consistency:

Step 1: Bias($\hat{\beta}$) tends to 0 as n tends to infinity

$$\begin{aligned} \text{Bias}(\hat{\beta}) &= E[\hat{\beta}] - \beta \\ E[\hat{\beta}] &= E\left[\frac{\sum_i X_i}{n}\right] \end{aligned}$$

Taking constant out and applying LOE

$$\begin{aligned} E[\hat{\beta}] &= \frac{1}{n} n E[X_1] = E[X_1] \\ E[X_1] &= \beta \end{aligned}$$

Thus Bias ($\hat{\beta}$) = 0

Step 2: SE($\hat{\beta}$) tends to 0 as n tends to infinity

$$\begin{aligned} se(\hat{\beta}) &= \sqrt{Var(\hat{\beta})} = \sqrt{Var\left(\frac{\sum_i X_i}{n}\right)} \\ &\Rightarrow \sqrt{\frac{1}{n^2} Var(\sum_i X_i)} \end{aligned}$$

As they are iid

$$\Rightarrow \sqrt{\frac{1}{n^2} n Var(X_1)} = \sqrt{\frac{Var(X_1)}{n}} = \frac{\beta}{\sqrt{n}}$$

Thus,

$$se(\hat{\beta}) = \frac{\beta}{\sqrt{n}}$$

tends to 0 as n tends to infinity

Q3

a

First we know that the probability for Poisson distribution is

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

For X_1, \dots, X_n iid Poisson random variables will have a joint frequency function that is a product of the marginal frequency functions. Thus, the likelihood will be:

$$L(\lambda) = \prod_i \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = \frac{\lambda^{\sum x_i} e^{-n\lambda}}{\prod_i x_i!}$$

The log likelihood will then be

$$l(\lambda) = \log \lambda \cdot \sum_i x_i - n\lambda - \sum_i \log x_i!$$

We need to find the maximum by finding the derivative:

$$l'(\lambda) = \frac{1}{\lambda} \sum_i x_i - n = 0$$

which implies that the MLE should be

$$\hat{\lambda} = \frac{\sum_i x_i}{n} = \bar{x}$$

b) $x_1, \dots, x_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$

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

$$P_{\mathcal{D}}(\mathcal{D}) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2\sigma^2} \sum_i (x_i - \mu)^2}$$

$$\ln(P_{\mathcal{D}}(\mathcal{D})) = -\ln(\sqrt{2\pi}\sigma)^n - \frac{1}{2\sigma^2} \sum (x_i - \mu)^2$$

$$\frac{\partial}{\partial \mu} (\cdot) = 0 - \frac{1}{2\sigma^2} \sum (x_i - \mu) (-1) = 0$$

$$\sum x_i - n\mu = 0 \Rightarrow \hat{\mu}_{MLE} = \frac{\sum x_i}{n} \rightarrow \text{same as sample mean}$$

$$\frac{\partial}{\partial \sigma} (\cdot) = -\frac{n}{\sigma} - \frac{(-2)}{2\sigma^3} \sum (x_i - \mu)^2 = 0$$

$$\Rightarrow \frac{n}{\sigma} = \frac{1}{\sigma^3} \sum (x_i - \mu)^2$$

$$\sigma^2 = \frac{1}{n} \sum (x_i - \mu)^2 \rightarrow \text{same as uncorrected sample variance.}$$

c

Let $X_1, \dots, X_n \sim N(\theta, 1)$. The MLE for θ is $\hat{\theta} = \bar{X}$. Let $\delta = E[I_{X_1 > 0}]$. Thus,

$$\begin{aligned}\delta &= E[I_{X_1 > 0}] \\&= P(X_1 > 0) \\&= 1 - P(X_1 \leq 0) \\&= 1 - F_{X_1}(0) \\&= 1 - \Phi\left(\frac{0 - \mu}{\sigma}\right) \\&= \Phi\left(\frac{\mu}{\sigma}\right) \\&= \Phi\left(\frac{\theta}{1}\right) \\&= \Phi(\theta)\end{aligned}$$

So the MLE for δ is $\Phi(\bar{X}) = \Phi\left(\frac{\sum_i X_i}{n}\right)$

Q 4

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

① $K=1$

② $\hat{x}_1 = \frac{1}{n} \sum x_i$

③ $x_1(\theta) = E[X(\theta)] = \sum_{x \in D} x \cdot p(x)$

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

④ $\frac{1}{n} \sum x_i = \hat{x}$
 $\hat{x} = 3 - \frac{\sum x_i}{n} = 3 - \frac{(2+3+2)}{3} = \frac{2}{3}$

$$b) \text{ se}(\hat{\theta}_{MME}) = ? , \quad \hat{\theta}_{MME} = 3 - \bar{x}$$

$$\hookrightarrow = \sqrt{\text{Var}(3 - \bar{x})} = \sqrt{\text{Var}(\bar{x})} = \sqrt{\text{Var}\left(\frac{1}{n} \sum x_i\right)}$$

$$\stackrel{\text{Def}}{=} \sqrt{\frac{1}{n^2} (n) \text{Var}(x)} = \sqrt{\frac{\text{Var}(x)}{n}}$$

$$\therefore E(x^2) = 4(\theta) + 9(1-\theta) = 9-5\theta$$

$$E(x) = 3-\theta$$

$$\text{Var}(x) = 9-5\theta - (3-\theta)^2 = 9-5\theta - (9+ \theta^2 - 6\theta) = \theta - \theta^2 = \theta(1-\theta)$$

$$\text{se}(\hat{\theta}_{MME}) = \sqrt{\frac{\theta(1-\theta)}{n}} ; \quad \text{se}(\hat{\theta}_{MME}) = \sqrt{\frac{\hat{\theta}_{MME}(1-\hat{\theta}_{MME})}{n}}$$

$$\hat{\text{se}}(\hat{\theta}_{MME}) = \sqrt{\frac{2/3 \times 1/3}{3}} = \frac{1}{3\sqrt{3}} = 0.2721$$

$\hat{\theta}_{MME}$ is AN
Estimate \downarrow this applies
Normal CI
of 95%ile = $2/3 \pm 1.96(0.2721) = [0.1332, 1.2034]$
CI-Normal

(4) (c).

Let X_1, \dots, X_n be i.i.d from the given distribution.

$$P(X_1, \dots, X_n | \theta) = \prod_{i=1}^n P(X_i | \theta).$$

$$\therefore L(\theta) = \prod_{i=1}^n P(X_i | \theta)$$

$$\text{We have } P(X_i | \theta) = \theta^{(3-X_i)} (1-\theta)^{(X_i-2)}$$

$$\therefore l(\theta) = \log L(\theta)$$

$$= \sum_{i=1}^n \log P(X_i | \theta)$$

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

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

$$= \log \theta \sum_i (3-X_i) + \log (1-\theta) \sum_i (X_i-2)$$

$$\begin{aligned} \therefore \frac{\partial l}{\partial \theta} &= 0 \\ \Rightarrow \frac{1}{\theta} \sum_i (3 - X_i) - \frac{1}{1-\theta} \sum_i (X_i - 2) &= 0 \\ (1-\theta) \sum_i (3 - X_i) - \theta \sum_i (X_i - 2) &= 0 \\ \sum_i (3 - X_i) - \theta \left[\sum_i (3 - X_i) + \sum_i (X_i - 2) \right] &= 0 \\ \Rightarrow \hat{\theta} &= \frac{\sum_i 3 - X_i}{\sum_i 3 - X_i + \sum_i (X_i - 2)} = \frac{\sum_i 3 - X_i}{n}. \\ \text{One can verify } \left. \frac{\partial^2 l}{\partial \theta^2} \right|_{\theta=\hat{\theta}} &< 0 \end{aligned}$$

$$\begin{aligned} \text{For } D &= \{2, 3, 2\} \\ \hat{\theta} &= \frac{(3-2)+(3-3)+(3-2)}{3} \\ &= 2/3 \end{aligned}$$

Q5

MME of $\text{Exp}(\lambda)$, $\hat{\lambda}_{\text{MME}}$

Step 0: $K = 1$

$$\text{Step 1: } \hat{\alpha}_1 = \frac{1}{n} \sum_{j=1}^n x_j$$

$$\text{Step 2: } \alpha_1(\lambda) = E[\text{exp}(\lambda)] = \frac{1}{\lambda}$$

$$\text{Step 3: } \alpha_1(\hat{\lambda}) = \hat{\alpha}_1$$

$$\frac{1}{\hat{\lambda}} = \frac{1}{n} \sum_{j=1}^n x_j$$

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

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

Scanned with
CamScanner

b)

MLE of $\text{Exp}(\lambda)$, $\hat{\lambda}_{\text{MLE}}$

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

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

$$l(\theta) = \log [\lambda^n e^{-\lambda \sum x_i}]$$

$$= n \log \lambda - \lambda \sum x_i$$

$$\text{Now, } \frac{d l(\theta)}{d \theta} = 0$$

$$\therefore n \cdot \frac{1}{\lambda} - \sum x_i = 0$$

$$\Rightarrow \lambda = \frac{n}{\sum x_i}$$

$$\therefore \hat{\lambda}_{\text{MLE}} = \frac{n}{\sum x_i}$$

5)

c)

mme(normal):

$$\hat{\mu} = 15.568 \quad (15.57)$$

$$\hat{\sigma}^2 = 7.585 \quad (7.59)$$

mme(uniform):

$$a = 69.614 \quad (69.61)$$

$$b = 82.406 \quad (82.4)$$

mme(exp):

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

d)

mle(Normal):

$$\hat{\mu} = 15.568 \quad (15.57)$$

$$\hat{\sigma}^2 = 7.585 \quad (7.59)$$

mle(uniform):

$$a = 70$$

$$b = 82$$

mle(exp):

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

Q6.

	Healthy	Sick
Healthy	98 (TN)	2 (FP)
Sick	1 (FN)	99 (TP)

(a) Precision

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{99}{99 + 2} = \frac{99}{101}$$

(b) Recall

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

(c) Type 1 Error

$$\text{Type 1 Error} = \frac{FP}{FP + TN} = \frac{2}{98 + 2} = \frac{2}{100}$$

(d) Type 2 Error

$$\text{Type 2 Error} = \frac{FN}{FN + TP} = \frac{1}{99 + 1} = \frac{1}{100}$$

Q7.

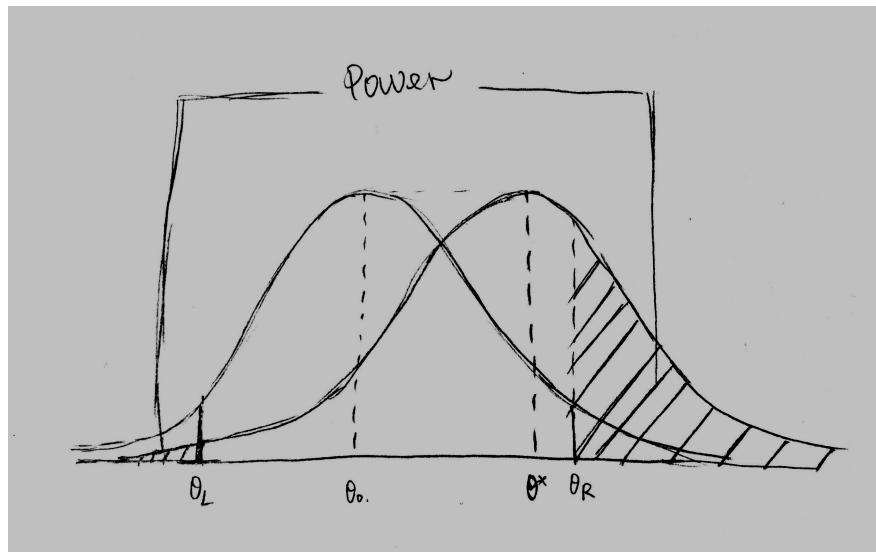
(a)

Based on the hypothesis, we know that this is a 2-sided test and the type II Error is failing to reject H_0 even when H_1 is true. Given

- $H_0 : \theta = \theta_0$
- $H_1 : \theta \neq \theta_0$

we know that we reject H_0 when $|W| > z_{\alpha/2}$, which means that

- Case 1: $W > z_{\alpha/2} \Rightarrow \frac{\theta - \theta_0}{\hat{s}_e} > z_{\alpha/2} \Rightarrow \theta > \theta_0 + \hat{s}_e \cdot z_{\alpha/2} = \theta_R$
- Case 2: $W < -z_{\alpha/2} \Rightarrow \frac{\theta - \theta_0}{\hat{s}_e} < -z_{\alpha/2} \Rightarrow \theta < \theta_0 - \hat{s}_e \cdot z_{\alpha/2} = \theta_L$



Now, if we know that the true value of θ is θ^* Therefore,

$$\begin{aligned}
Power &= P[\text{case 1}|\text{true mean is } \theta^*] + P[\text{case 2}|\text{true mean is } \theta^*] \\
&= P(\theta > \theta_R | \theta = \theta^*) + P(\theta < \theta_L | \theta = \theta^*) \\
&= P(\theta > \theta_0 + \hat{s}e \cdot z_{\alpha/2} | \theta = \theta^*) + P(\theta < \theta_0 - \hat{s}e \cdot z_{\alpha/2} | \theta = \theta^*) \\
&= P(W > \frac{\theta_0 + \hat{s}e \cdot z_{\alpha/2} - \theta^*}{\hat{s}e}) + P(W < \frac{\theta_0 - \hat{s}e \cdot z_{\alpha/2} - \theta^*}{\hat{s}e}) \\
&= P(W > \frac{\theta_0 - \theta^*}{\hat{s}e} + z_{\alpha/2}) + P(W < \frac{\theta_0 - \theta^*}{\hat{s}e} - z_{\alpha/2}) \\
&= 1 - P(W < \frac{\theta_0 - \theta^*}{\hat{s}e} + z_{\alpha/2}) + P(W < \frac{\theta_0 - \theta^*}{\hat{s}e} - z_{\alpha/2}) \\
&= 1 - \Phi(\frac{\theta_0 - \theta^*}{\hat{s}e} + z_{\alpha/2}) + \Phi(\frac{\theta_0 - \theta^*}{\hat{s}e} - z_{\alpha/2})
\end{aligned}$$

Thus,

$$P[\text{Type II error}] = 1 - Power = \Phi(\frac{\theta_0 - \theta^*}{\hat{s}e} + z_{\alpha/2}) - \Phi(\frac{\theta_0 - \theta^*}{\hat{s}e} - z_{\alpha/2})$$

Given experiment is Bernoulli Distribution, then $p_0 = 0.5$

(b) Given - $H_0 : p = p_0$, $H_1 : p \neq p_0$

We have 46 success

$$\Rightarrow \hat{p} = \bar{x} = \frac{\sum x_i}{n} = \frac{46}{100} = 0.46$$

Wald's test $\rightarrow W = \frac{\hat{p} - p_0}{\hat{s}_e} = \frac{\hat{p} - p_0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$

$$W = \frac{0.46 - 0.5}{\sqrt{\frac{0.46(1-0.46)}{100}}} = -0.803$$

$$|W| = 0.803$$

$$\alpha = 0.05 \quad Z_{\alpha/2} = Z_{0.025} = 1.96$$

$|W| < Z_{\alpha/2} \Rightarrow$ we accept the hypothesis.

$$\begin{aligned} p\text{-value} &= 2(1 - \Phi(|W|)) = 2\Phi(-|W|) = 2\Phi(-0.803) \\ &= 0.42371 \end{aligned}$$

Now, if $p = 0.7$, so $p_0 = 0.7$

$$W = \frac{0.46 - 0.7}{\sqrt{\frac{0.46(1-0.46)}{100}}} = -4.819$$

$$\therefore |W| = 4.819$$

Now, $|W| > Z_{\alpha/2} \Rightarrow$ we reject the hypothesis

$$p\text{-value} = 2\Phi(-|W|) = 2\Phi(-4.819)$$

< 0.00001 , i.e. we get a very small value from the p-value table.

8)

a) $\hat{\theta} = 0.541$

$$\hat{s}_\theta(\hat{\theta}) = \sqrt{\frac{0.0106}{1000}} = 0.00326$$

$$z = \frac{0.041}{0.00326} \approx 12.56 > z_{\alpha/2} = 2.326$$

Reject null, H_0 .

$$b) W = \left| \frac{\hat{\theta} - \theta}{\text{se}(\hat{\theta})} \right|$$

$\hat{\theta} = \bar{x} - \bar{y}$, where X and Y data samples are normally distributed and are independent as well making the test applicable.

$$= \left| \frac{\bar{x} - \bar{y}}{\text{se}(\bar{x} - \bar{y})} \right| = \frac{|\bar{x} - \bar{y}|}{\sqrt{\frac{\text{var}(x_i)}{750} + \frac{\text{var}(y_i)}{750}}}$$

$$= \frac{|5.0048 - 5.846|}{\sqrt{\frac{2.34}{750} + \frac{6.47}{750}}} = \frac{0.842}{\sqrt{0.0118}} = \frac{0.842}{0.109} = 7.725$$

$|W| > Z_{\alpha/2}$ for $\alpha = 0.05$, hence we reject the null hypothesis, H_0 .