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:

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

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

$$Var[X] = E[X^2] - (E[X])^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\frac{1}{n}\sum_{i=1}^{n} X_i)^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = xy^2$$

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

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

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:

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

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

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

Similarly, we could have the MME for a,

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

Q2.

Given:

$$X_1, X_2, X_3, X_4....X_n \sim Exponential(\frac{1}{\beta})$$

Step 1: Likelihood

$$L(p) = \prod_{i=1}^{n} \frac{1}{\beta} e^{-\frac{1}{\beta} X_i}$$

$$L(p) = \frac{1}{\beta^n} e^{-\frac{1}{\beta} \sum_i X_i}$$

Step 2: Taking Log both sides

$$log(L(p)) = log(\frac{1}{\beta^n} e^{-\frac{1}{\beta} \sum_i X_i})$$

$$log(L(p)) = -nlog(\beta) - \frac{1}{\beta} \sum_i X_i$$

Step3 : Differntiate it w.r.t β and equate it to 0 to get MLE

$$\implies -\frac{n}{\beta} + \beta^{-2} \sum_i X_i = 0$$

$$\implies \hat{\beta} = \frac{\sum_i X_i}{n}$$

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

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

$$E[\hat{\beta}] = E[\frac{\sum_i X_i}{n}]$$

Taking constant out and applying LOE

$$E[\hat{\beta}] = \frac{1}{n} nE[X_1] = E[X_1]$$

$$E[X_1] = \beta$$

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

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

$$\implies \sqrt{\frac{1}{n^2} Var(\sum_i X_i)}$$

As they are iid

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

Thus,

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

tends to 0 as n tends ∞

Q3

a

First we know that the probability for Poisson distribution is

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

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

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

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.) $X_1, \ldots, X_n \overset{iid}{\sim} \text{Normal}(\mu, \sigma^2)$

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

$$Pr(D) = \frac{1}{(\sqrt{2\pi}\,\sigma)^n} e^{-\frac{1}{2\sigma^2} \sum_1^n (x_i-\mu)^2}$$

$$\ln(Pr(D)) = -\ln(\sqrt{2\pi}\,\sigma)^n - \frac{1}{2\sigma^2} \Sigma (x_i-\mu)^2$$

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

$$\Sigma x_i - n\mu = 0 \Rightarrow \hat{\mu}_{MLE} = \frac{\Sigma x_i}{n}$$ → same as sample mean

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

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

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

c

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

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

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

Q4

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

① $K = 1$

② $\hat{\alpha}_1 = \frac{1}{n} \sum_i X_i$

③ $\alpha_1(\theta) = E[X(\theta)] = \sum_{x \in \Omega} x \cdot p(x)$

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

④ $\frac{1}{n} \sum_i X_i = 3 - \hat{\theta}$

$\hat{\theta} = 3 - \frac{\sum_i X_i}{n} = 3 - \frac{(2+3+2)}{3} = \frac{2}{3}$

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

$$\hookrightarrow = \sqrt{Var(3-\bar{X})} = \sqrt{Var(\bar{X})} = \sqrt{Var\left(\frac{1}{n}\sum X_i\right)}$$

$$\overset{LOV}{=} \sqrt{\frac{1}{n^2}(n)Var(X)} = \sqrt{\frac{Var(X)}{n}}$$

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

$E(X) = 3 - \theta$

$Var(X) = 9 - 5\theta - (3-\theta)^2 = 9 - 5\theta - (9 + \theta^2 - 6\theta) = \theta - \theta^2 = \theta(1-\theta)$

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

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

$\boxed{\hat{\theta}_{MME} \text{ is AN}}$ This applies Normal CI

Estimate of 95%ile CI-Normal $= \quad 2/3 \pm 1.96(0.2721) = [0.1332, 1.2034]$

(4) (c).

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

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

i.e., $$L(\theta) = \prod_{i=1}^{n} P(x_i | \theta)$$

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

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

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

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

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

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

$$\therefore \frac{\partial l}{\partial \theta} = 0$$

$$\Rightarrow \frac{1}{\theta} \sum_i (3 - X_i) - \frac{1}{1-\theta} \sum_i (X_i - 2) = 0$$

$$(1-\theta) \sum_i (3 - X_i) - \theta \sum_i (X_i - 2) = 0$$

$$\sum_i (3 - X_i) - \theta \left[\sum_i (3 - X_i) + \sum_i (X_i - 2) \right] = 0$$

$$\Rightarrow \hat{\theta} = \frac{\sum_i 3 - X_i}{\sum_i 3 - X_i + \sum_i (X_i - 2)} = \frac{\sum_i 3 - X_i}{n}.$$

One can verify $\left. \frac{\partial l^2}{\partial \theta^2} \right|_{\theta = \hat{\theta}} < 0$

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

$$\hat{\theta} = \frac{(3-2) + (3-3) + (3-2)}{3}$$

$$= 2/3$$

Q5

a)

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

Step 0: $k = 1$

Step 1: $\hat{\alpha}_1 = \frac{1}{n}\sum_{i=1}^{n} x_i$

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

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

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

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

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

b)

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

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

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

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

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

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

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

5)

c) mme(normal):

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

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

mme(uniform):

$a = 69.614\ (69.61)$

$b = 82.406\ (82.41)$

mme(exp):

$\hat{\lambda} = 0.0425$

d) mle(normal):

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

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

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.

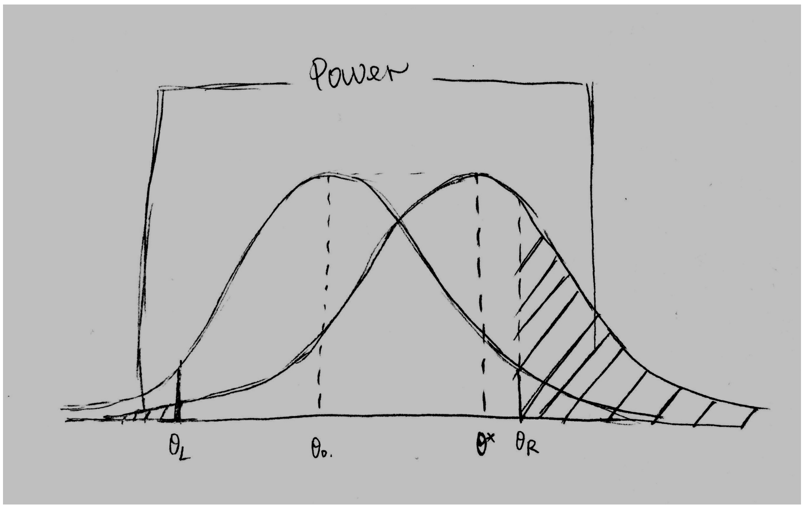
(a)

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



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

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

Thus,

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

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

$$\Rightarrow \quad \hat{p} = \bar{X} = \frac{\sum X_i}{n} = \frac{46}{100} = 0.46$$

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

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

$$|W| = 0.803$$

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

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

p-value $= 2(1 - \Phi(|W|)) = 2\Phi(-|W|) = 2\Phi(-0.803)$
$= 0.42371$

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 <u>reject</u> the hypothesis

p-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) $\hat{\theta} = 0.541$

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

$W = \frac{0.041}{0.00326} = 12.56 \geq z_{\alpha/2} = 2.326$

Reject null, H_0.

b) $W = \left| \frac{\hat{\theta} - 0}{\hat{se}(\hat{\theta})} \right|$

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

$$= \left| \frac{\bar{X} - \bar{Y}}{\hat{se}(\bar{X} - \bar{Y})} \right| = \frac{|\bar{X} - \bar{Y}|}{\sqrt{\frac{Var(X_i)}{750} + \frac{Var(Y_i)}{750}}}$$

$$= \frac{|5.0048 - 5.846|}{\sqrt{\frac{2.36}{750} + \frac{6.47}{750}}} = \frac{0.842}{\sqrt{0.0118}}$$

$$= \frac{0.842}{0.109} = 7.725$$

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