

Probability Theory and Statistics 3, Assignment 3

Jiyeon Kim ^{*}, Younjoo Mo [†]

November 13, 2020

Exercise 1

a.

$$Y_i \mid X_i = x_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2) \quad (1)$$

Likelihood function

$$L(\beta_0, \beta_1, \sigma) = \prod_{i=1}^n f_{X,Y}(x, y) = \prod_{i=1}^n f_{Y|X}(y \mid x; \beta_0, \beta_1, \sigma) f_X(x) \quad (2)$$

$$= \left[\frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left(-\frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2}\right) \prod_{i=1}^n f_X(x) \right] \quad (3)$$

Log-likelihood function ('log' is the natural logarithm)

$$l(\beta_0, \beta_1, \sigma) = \log [L(\beta_0, \beta_1, \sigma)] \quad (4)$$

$$= -n \log(\sqrt{2\pi}\sigma) - \frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2} + \sum_{i=1}^n \log f_X(x) \quad (5)$$

^{*}Student number: 12470236

[†]Studentnumber: 12475440

b. To find the MLEs of β_0, β_1 and σ , we would check the first order condition for each of the unknown parameters as the first step. Then the term including $f_X(x)$ immediately disappears as we take the partial derivative of the log-likelihood function. Because β_0, β_1 and σ are not a component of $f_X(x)$. In other words, $f_X(x)$ is independent of those parameters. Therefore we do not need to know what $f_X(x)$ is, in order to find the MLEs of β_0, β_1 and σ .

c.

i) The first order condition for β_0

$$\frac{\partial}{\partial \beta_0} l(\beta_0, \beta_1, \sigma) \quad (6)$$

Substitute (5) and omit the terms not including β_0 (7)

$$= \frac{\partial}{\partial \beta_0} \left[-\frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2} \right] \quad (8)$$

$$= \frac{\partial}{\partial \beta_0} \left[-\frac{1}{2\sigma^2} \left(\sum_{i=1}^n y_i^2 - 2\beta_0 \sum_{i=1}^n y_i - 2\beta_1 \sum_{i=1}^n x_i y_i + n\beta_0^2 + 2\beta_0 \beta_1 \sum_{i=1}^n x_i + \beta_1^2 \sum_{i=1}^n x_i^2 \right) \right] \stackrel{\text{f.o.c}}{=} 0 \quad (9)$$

$$\Rightarrow -\sum_{i=1}^n y_i + n\beta_0 + \beta_1 \sum_{i=1}^n x_i = 0 \quad (10)$$

$$\Rightarrow \hat{\beta}_0 = \frac{\sum_{i=1}^n y_i - \beta_1 \sum_{i=1}^n x_i}{n} = \bar{Y}_n - \beta_1 \bar{X}_n \quad (11)$$

ii) The first order condition for β_1

$$\frac{\partial}{\partial \beta_1} l(\beta_0, \beta_1, \sigma) \quad (12)$$

Substitute (5) and omit the terms not including β_1 (13)

$$= \frac{\partial}{\partial \beta_1} \left[-\frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2} \right] \quad (14)$$

$$= \frac{\partial}{\partial \beta_1} \left[-\frac{1}{2\sigma^2} \left(\sum_{i=1}^n y_i^2 - 2\beta_0 \sum_{i=1}^n y_i - 2\beta_1 \sum_{i=1}^n x_i y_i + n\beta_0^2 + 2\beta_0 \beta_1 \sum_{i=1}^n x_i + \beta_1^2 \sum_{i=1}^n x_i^2 \right) \right] \stackrel{\text{f.o.c}}{=} 0 \quad (15)$$

$$\Rightarrow -\sum_{i=1}^n x_i y_i + \beta_0 \sum_{i=1}^n x_i + \beta_1 \sum_{i=1}^n x_i^2 = 0 \quad (16)$$

(11) can be substituted for β_0 of (16) (17)

$$\Rightarrow -\sum_{i=1}^n x_i y_i + (\bar{Y}_n - \beta_1 \bar{X}_n) \sum_{i=1}^n x_i + \beta_1 \sum_{i=1}^n x_i^2 = 0 \quad (18)$$

$$\Rightarrow \hat{\beta}_1 = \frac{-\sum_{i=1}^n x_i y_i + \bar{Y}_n \sum_{i=1}^n x_i}{\bar{X}_n \sum_{i=1}^n x_i - \sum_{i=1}^n x_i^2} \quad (19)$$

iii) The first order condition for σ

$$\frac{\partial}{\partial \sigma} l(\beta_0, \beta_1, \sigma) \quad (20)$$

Substitute (5) and omit the terms not including σ (21)

$$= \frac{\partial}{\partial \sigma} \left[-n \log(\sqrt{2\pi}\sigma) - \frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2} \right] \stackrel{\text{f.o.c}}{=} 0 \quad (22)$$

$$\Rightarrow -\frac{n}{\sigma} + \frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2}{\sigma^3} = 0 \quad (23)$$

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2}{n} \quad (24)$$

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2}{n}} \quad (25)$$

d.

```
<code>
```

```
df<-read.table ("PTS3_assignment3_data.txt")
```

```
x<-df [,1]
```

```
y<-df [,2]
```

```
beta1<- (-sum(x*y)+mean(y)*sum(x))/(mean(x)*sum(x)-sum(x^2))
```

```
beta1
```

```
beta0<- mean(y)-beta1*mean(x)
```

```
beta0
```

```
sigma<-sqrt(sum((y-beta0-beta1*x)^2)/50)
```

```
sigma
```

If we run the code, we get the following result.

$$\beta_1 = 1.546367$$

$$\beta_0 = -1.47139$$

$$\sigma = 1.191652$$

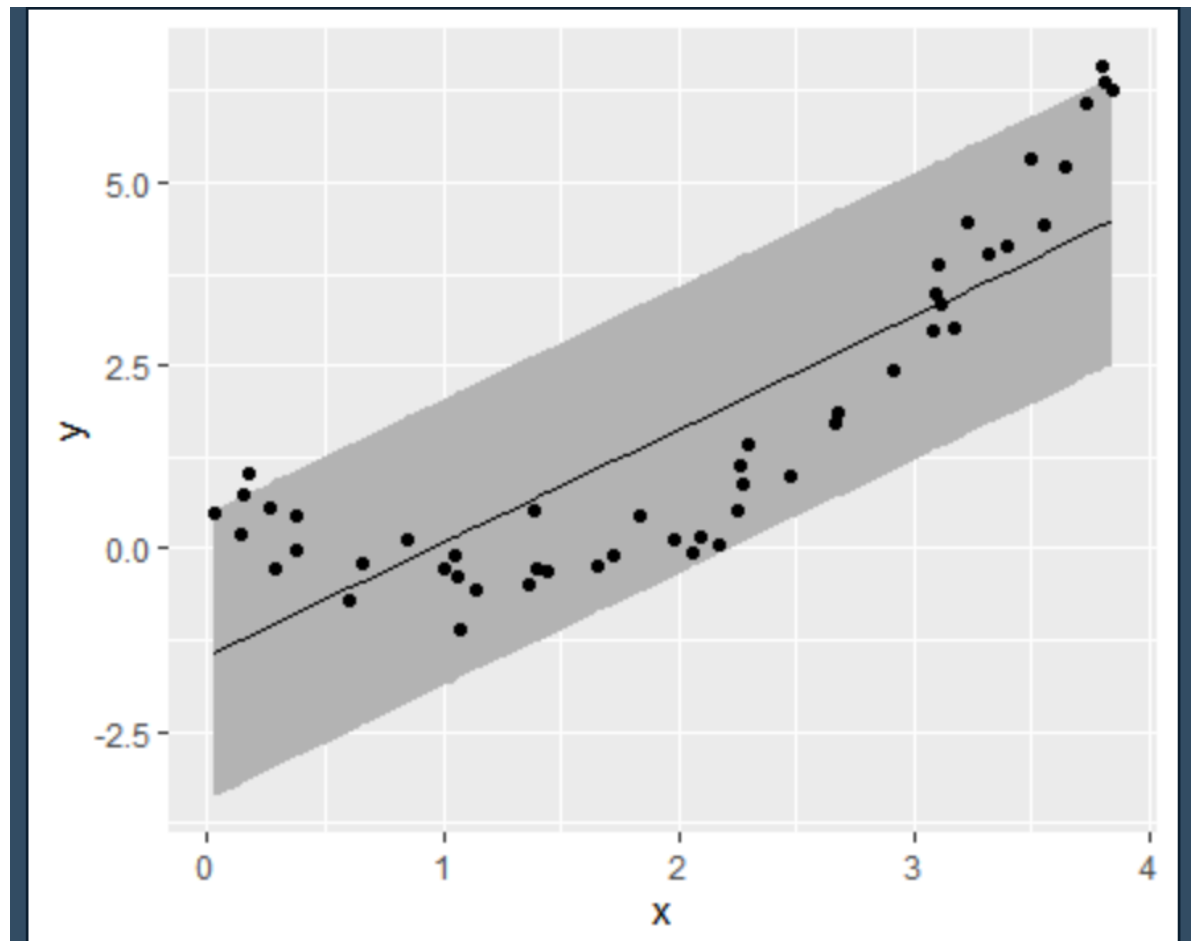
e. In order to compute 95-percent confidence interval for Y, we used $\sigma \approx \hat{\sigma} = 1.191652$

$$P(-1.645 < \frac{Y | (X = x) - (\beta_0 + \beta_1 x_i)}{\sigma} < 1.645) \quad (26)$$

$$\Rightarrow P(\beta_0 + \beta_1 x_i - 1.645\sigma < Y | (X = x) < \beta_0 + \beta_1 x_i + 1.645\sigma) \quad (27)$$

$$\text{Let } \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_i \quad (28)$$

$$\Rightarrow P(\hat{y} - 1.645\hat{\sigma} < Y | (X = x) < \hat{y} + 1.645\hat{\sigma}) \quad (29)$$



f. Likelihood function

$$Y_i | X_i = x_i \sim N(\beta_0 + \beta_1 x_i + \beta_2 x_i^2, \sigma^2) \quad (30)$$

$$L(\beta_0, \beta_1, \beta_2, \sigma) = \prod_{i=1}^n f_{X,Y}(x, y; \beta_0, \beta_1, \beta_2, \sigma) = \prod_{i=1}^n f_{Y|X}(y | x; \beta_0, \beta_1, \beta_2, \sigma) f_X(x) \quad (31)$$

$$= \left[\frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left(-\frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2}{2\sigma^2}\right) \prod_{i=1}^n f_X(x) \right] \quad (32)$$

Log-likelihood function ('log' is the natural logarithm)

$$l(\beta_0, \beta_1, \beta_2, \sigma) = \log [L(\beta_0, \beta_1, \beta_2, \sigma)] \quad (33)$$

$$= -n \log(\sqrt{2\pi}\sigma) - \frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2}{2\sigma^2} + \sum_{i=1}^n \log f_X(x) \quad (34)$$

i) The first order condition for β_0

$$\frac{\partial}{\partial \beta_0} l(\beta_0, \beta_1, \beta_2, \sigma) \quad (35)$$

$$= \frac{\partial}{\partial \beta_0} \left[-\frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2}{2\sigma^2} \right] \quad (36)$$

$$= \frac{\partial}{\partial \beta_0} \left[-\frac{1}{2\sigma^2} \left(\sum_{i=1}^n y_i^2 - 2\beta_0 \sum_{i=1}^n y_i - 2\beta_1 \sum_{i=1}^n x_i y_i - 2\beta_2 \sum_{i=1}^n x_i^2 y_i \right. \right. \quad (37)$$

$$\left. + n\beta_0^2 + \beta_1^2 \sum_{i=1}^n x_i^2 + \beta_2^2 \sum_{i=1}^n x_i^4 + 2\beta_0 \beta_2 \sum_{i=1}^n x_i^2 + 2\beta_1 \beta_2 \sum_{i=1}^n x_i^3 + 2\beta_0 \beta_1 \sum_{i=1}^n x_i \right) \right] \stackrel{\text{f.o.c}}{=} 0 \quad (38)$$

$$\Rightarrow -\sum_{i=1}^n y_i + n\beta_0 + \beta_1 \sum_{i=1}^n x_i + \beta_2 \sum_{i=1}^n x_i^2 = 0 \quad (39)$$

$$\Rightarrow \hat{\beta}_0 = \frac{\sum_{i=1}^n y_i - \beta_1 \sum_{i=1}^n x_i - \beta_2 \sum_{i=1}^n x_i^2}{n} = \bar{y} - \beta_1 \bar{x} - \beta_2 \bar{x}^2 \quad (40)$$

$$\therefore \Rightarrow \bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x} + \hat{\beta}_2 \bar{x}^2 \quad (41)$$

ii) The first order condition for β_1

$$\frac{\partial}{\partial \beta_1} l(\beta_0, \beta_1, \beta_2, \sigma) \quad (42)$$

$$= \frac{\partial}{\partial \beta_1} \left[- \frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2}{2\sigma^2} \right] \quad (43)$$

$$= \frac{\partial}{\partial \beta_1} \left[- \frac{1}{2\sigma^2} \left(\sum_{i=1}^n y_i^2 - 2\beta_0 \sum_{i=1}^n y_i - 2\beta_1 \sum_{i=1}^n x_i y_i - 2\beta_2 \sum_{i=1}^n x_i^2 y_i \right. \right. \quad (44)$$

$$\left. + n\beta_0^2 + \beta_1^2 \sum_{i=1}^n x_i^2 + \beta_2^2 \sum_{i=1}^n x_i^4 + 2\beta_0 \beta_2 \sum_{i=1}^n x_i^2 + 2\beta_1 \beta_2 \sum_{i=1}^n x_i^3 + 2\beta_0 \beta_1 \sum_{i=1}^n x_i \right) \Big] \stackrel{\text{f.o.c}}{=} 0 \quad (45)$$

$$\Rightarrow - \sum_{i=1}^n x_i y_i + \beta_1 \sum_{i=1}^n x_i^2 + \beta_0 \sum_{i=1}^n x_i + \beta_2 \sum_{i=1}^n x_i^3 = 0 \quad (46)$$

$$\Rightarrow \hat{\beta}_1 = \frac{\sum_{i=1}^n x_i y_i - \beta_0 \sum_{i=1}^n x_i - \beta_2 \sum_{i=1}^n x_i^3}{\sum_{i=1}^n x_i^2} = \frac{\bar{xy} - \beta_0 \bar{x} - \beta_2 \bar{x}^3}{\bar{x}^2} \quad (47)$$

$$\therefore \Rightarrow \bar{xy} = \hat{\beta}_0 \bar{x} + \hat{\beta}_1 \bar{x}^2 + \hat{\beta}_2 \bar{x}^3 \quad (48)$$

iii) The first order condition for β_2

$$\frac{\partial}{\partial \beta_1} l(\beta_0, \beta_1, \beta_2, \sigma) \quad (49)$$

$$= \frac{\partial}{\partial \beta_1} \left[-\frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2}{2\sigma^2} \right] \quad (50)$$

$$= \frac{\partial}{\partial \beta_1} \left[-\frac{1}{2\sigma^2} \left(\sum_{i=1}^n y_i^2 - 2\beta_0 \sum_{i=1}^n y_i - 2\beta_1 \sum_{i=1}^n x_i y_i - 2\beta_2 \sum_{i=1}^n x_i^2 y_i \right. \right. \quad (51)$$

$$\left. + n\beta_0^2 + \beta_1^2 \sum_{i=1}^n x_i^2 + \beta_2^2 \sum_{i=1}^n x_i^4 + 2\beta_0 \beta_2 \sum_{i=1}^n x_i^2 + 2\beta_1 \beta_2 \sum_{i=1}^n x_i^3 + 2\beta_0 \beta_1 \sum_{i=1}^n x_i \right) \Big] \stackrel{\text{f.o.c}}{=} 0 \quad (52)$$

$$\Rightarrow -\sum_{i=1}^n x_i^2 y_i + \beta_2 \sum_{i=1}^n x_i^4 + \beta_1 \sum_{i=1}^n x_i^3 + \beta_0 \sum_{i=1}^n x_i^2 = 0 \quad (53)$$

$$\Rightarrow \hat{\beta}_2 = \frac{\sum_{i=1}^n x_i^2 y_i - \beta_1 \sum_{i=1}^n x_i^3 - \beta_0 \sum_{i=1}^n x_i^2}{\sum_{i=1}^n x_i^4} = \frac{\bar{x}^2 y - \beta_1 \bar{x}^3 - \beta_0 \bar{x}^2}{\bar{x}^4} \quad (54)$$

$$\therefore \Rightarrow \overline{x^2 y} = \hat{\beta}_0 \bar{x}^2 + \hat{\beta}_1 \bar{x}^3 + \hat{\beta}_2 \bar{x}^4 \quad (55)$$

we get the matrix shown in the question.

Next, we are going to prove $\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2}{n}}$

iv) The first order condition for $\hat{\sigma}$

$$\frac{\partial}{\partial \sigma} l(\beta_0, \beta_1, \beta_2, \sigma) \quad (56)$$

$$\text{Substitute (34) and omit the terms not including } \sigma \quad (57)$$

$$= \frac{\partial}{\partial \sigma} \left[-n \log(\sqrt{2\pi}\sigma) - \frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2}{2\sigma^2} \right] \stackrel{\text{f.o.c}}{=} 0 \quad (58)$$

$$\Rightarrow -\frac{n}{\sigma} + \frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2}{\sigma^3} = 0 \quad (59)$$

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2}{n} \quad (60)$$

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2}{n}} \quad (61)$$

The second part of Subquestion 1-f : Compute $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \sigma$

```
<code>
```

```
df<-read.table("PTS3_assignment3_data.txt")
```

```
x<-df[,1]
```

```
y<-df[,2]
```

```
# Use the first given expression to get the beta hats
```

```
# Create the matrix at the left side
```

```
mean.of.x <- c(1, mean(x), mean(x^2),  
               mean(x), mean(x^2), mean(x^3),  
               mean(x^2), mean(x^3), mean(x^4))
```

```
matA <- matrix(mean.of.x, nrow=3)
```

```
# Create the matrix at the right side
```

```
mean.of.xy <- c(mean(y), mean(x*y), mean(y*x^2))
```

```
matB <- matrix(mean.of.xy, nrow=3)
```

```
# Beta hats
```

```
beta <- solve(matA) %*% matB
```

```
beta
```

```
# Use the second given expression to get the sigma hat
```

```
# Sigma hat
```

```
sigma <- sqrt(sum((y - beta[1] - beta[2]*x - beta[3]*x^2)^2)/50)
```

If we run the code, we get the following result.

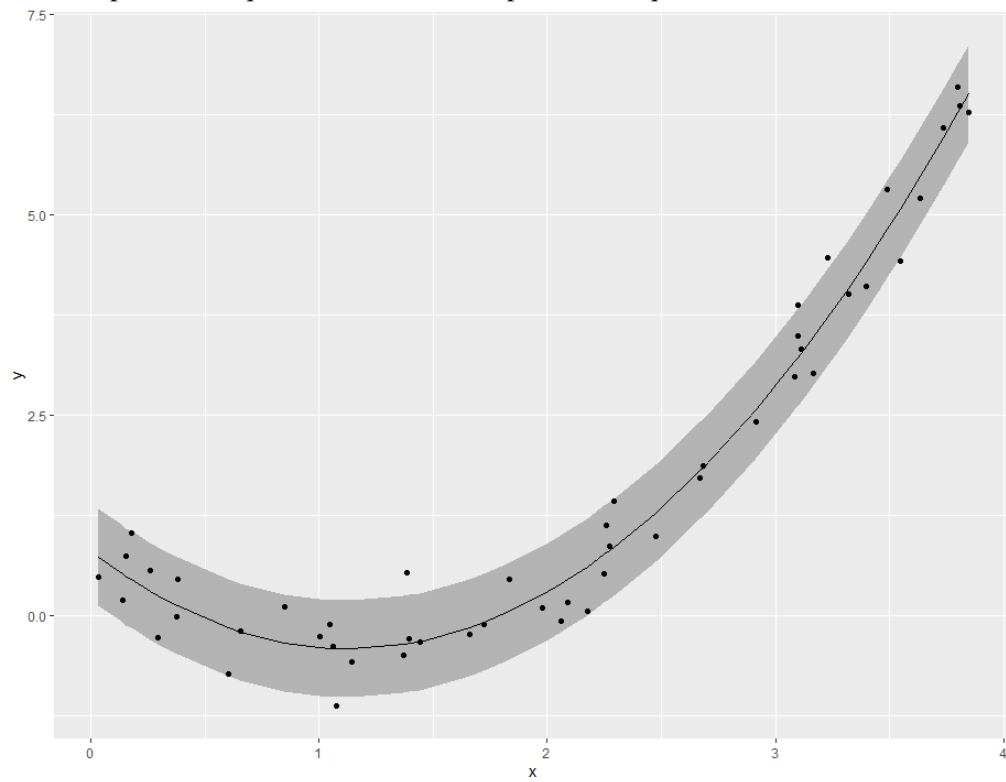
$\hat{\beta}_0 = 0.7963816$

$$\beta_1 = -2.1290201$$

$$\beta_2 = 0.9402489$$

$$\sigma = 0.3672581$$

The Last part of Subquestion 1-f : Create a plot for the quadratic model



Exercise 2

a.

$$X_i \sim POI(\mu) \quad (62)$$

$$E[X] = Var[X] = \mu \quad (63)$$

$$\text{The MLE of } \mu = \bar{X}_n \quad (64)$$

$$X_i \sim POI(\mu) \approx POI(\bar{X}_n) \quad (65)$$

$$\hat{p} = \frac{Y_n}{n} = e^{-\bar{X}_n} + \bar{X}_n \cdot e^{-\bar{X}_n} \quad (66)$$

$$E[\hat{p}] = E[e^{-\bar{X}_n} + \bar{X}_n \cdot e^{-\bar{X}_n}] \quad (67)$$

$$= E[e^{-\bar{X}_n}] + E[\bar{X}_n] \cdot E[e^{-\bar{X}_n}] \quad (68)$$

$$\text{By CLT} \quad (69)$$

$$\sqrt{n} \cdot \frac{\bar{X}_n - \mu}{\sigma} \xrightarrow{d} Z \sim N(0, 1) \quad (70)$$

$$\Rightarrow \text{MGF of } Z \text{ is } M_Z(t) = E[e^{Zt}] = e^{\frac{1}{2}t^2} \quad (71)$$

$$\Rightarrow \text{When } n \rightarrow \infty, \quad E\left[e^{\sqrt{n} \cdot \frac{\bar{X}_n - \mu}{\sigma} \cdot t}\right] = e^{\frac{1}{2}t^2} \quad (72)$$

$$\text{Substitute } t = -\frac{\sigma}{\sqrt{n}} \text{ to (72)} \quad (73)$$

$$\Rightarrow E[e^{-\bar{X}_n + \mu}] = E[e^{-\bar{X}_n}] \cdot E[e^{\mu}] = E[e^{-\bar{X}_n}] \cdot e^{\mu} = e^{\frac{\sigma^2}{2n}} \quad (74)$$

$$\Rightarrow E[e^{\bar{X}_n}] = e^{\frac{\sigma^2}{2n} - \mu} \quad (75)$$

$$\frac{\sigma^2}{2n} \rightarrow 0 \text{ as } n \rightarrow \infty \quad (76)$$

$$\therefore E[e^{\bar{X}_n}] = e^{-\mu} \quad (77)$$

Substitute (77) to (68) (78)

$$E[\hat{p}] = e^{-\mu} + \mu \cdot e^{-\mu} = e^{-\mu}(1 + \mu) = p \quad (79)$$

$\therefore \hat{p}$ is an unbiased estimator of p . (80)

b.

The formula of CRLB is

$$\frac{[\tau'(\mu)]^2}{-nE\left[\frac{\partial^2 \log f_X(x; \mu)}{\partial \mu^2}\right]} \quad (81)$$

\therefore When certain differentiability requirements holds,

$$E\left[\left(\frac{\partial}{\partial \theta} \log f_X(x; \mu)\right)^2\right] = -E\left[\frac{\partial^2}{\partial \theta^2} \log f_X(x; \mu)\right] \quad (82)$$

quantity equation is as follows

$$[\tau(\mu)] = P[X \leq 1] = e^{-\mu}(1 + \mu) \quad (83)$$

Take the partial derivative respect to μ

$$[\tau'(\mu)] = -e^{-\mu}(1 + \mu) + e^{-\mu} = -\mu e^{-\mu} \quad (84)$$

We get the discrete pdf $f_X(X; \mu)$ from the table B.1

$$f_X(X; \mu) = \frac{e^{-\mu} \mu^X}{X!} \quad (85)$$

$$\log f_X(X; \mu) = -\mu + X \log \mu - \log X! \quad (86)$$

Take the partial derivative respect to μ

$$\frac{\partial}{\partial \mu} \log f_X(X; \mu) = -1 + \frac{X}{\mu} \quad (87)$$

The second order condition for μ

$$\frac{\partial^2}{\partial \mu^2} \log f_X(X; \mu) = -\frac{X}{\mu^2} \quad (88)$$

$$E\left[\frac{\partial^2}{\partial \mu^2} \log f_X(X; \mu)\right] = \frac{-E[X]}{\mu^2} = -\frac{1}{\mu} \quad (89)$$

(8) $\rightarrow \cdot E(X) = \mu$ in poisson distribution

$$Var(T) \geq \frac{[\tau'(\mu)]^2}{-nE\left[\frac{\partial^2}{\partial \theta^2} \log f_X(X; \mu)\right]} = \frac{\mu^3 \exp^{-2\mu}}{n} \quad (90)$$

c.

From the previous sub question, we got the following formulas

$$f_X(X; \mu) = \frac{e^{-\mu} \mu^X}{X!} \quad (91)$$

'Log' is natural logarithm.

$$\log f_X(X; \mu) = -\mu + X \log \mu - \log X! \quad (92)$$

$$\frac{\partial}{\partial \mu} \log f_X(X; \mu) = -1 + \frac{X}{\mu} \quad (93)$$

We are going to use definition that If unbiased estimator attains CRLB, there should be a linear function of $\sum_{i=1}^n \frac{\partial}{\partial \mu} \log f_X(X; \mu)$

$$\sum_{i=1}^n \frac{\partial}{\partial \mu} \log f_X(x; \mu) = (-1 + \frac{X_1}{\mu}) + (-1 + \frac{X_1}{\mu}) + \dots + (-1 + \frac{X_n}{\mu}) \quad (94)$$

$$= -n + \frac{X_1 + X_2 + \dots + X_n}{\mu} \quad (95)$$

$$[\tau(\mu)] = P[X \leq 1] = e^{-\mu}(1 + \mu) \quad (96)$$

Substitute \hat{p} in equation (19).

$$[\tau(\hat{p})] = \exp^{-\hat{p}}(1 + \hat{p}) = \exp^{-\bar{X}_n}(1 + \bar{X}_n) \quad (97)$$

$$\neq c \left[-n + \frac{X_1 + X_2 + \dots + X_n}{\mu} \right], c \in \mathbb{R} \setminus \{0\} \quad (98)$$

There is no coefficient c that makes equation (20) and (21) linear

$\therefore \hat{p}$ doesnt attain CRLB for unbiased estimators of p.

d.

From the sub-question (a), \hat{p} is an unbiased estimator of p . In addition to this, if $Var(\hat{p})$ is equal to CRLB, then we can tell \hat{p} is an UMVUE of p . It happens only when \hat{p} is a linear function of $\sum_{i=1}^n \frac{\partial}{\partial \mu} \log f_X(x; \mu)$. However, we learnt that they do not have a linear relationship, from the sub-question (c).

We can't find UMVUE by using the CRLB-method. However, we don't know whether \hat{p} is UMVUE of p only based on the previous sub-questions and findings. We can't be certain that \hat{p} is not an UMVUE. There are possibilities that if we use other methods from the next lecture or more, probably we could be able to find this UMVUE.