CGS698C

Assignment-01

SAHIL TOMAR
210898

Part 1: Probability

1.1

(a) sample space (Ω) = $\{HH, HT, TH, TT\}$

(b) Event space ($\mathcal{F}$) = $\mathcal{P}(\Omega) = \{\{HH\}, \{HT\}, \{TH\}, \{TT\},$
$\{HH, HT\}, \{HH, TH\}, \{HH, TT\}, \{HT, TH\}, \{HT, TT\}, \{TH, TT\},$
$\{HT, TH, TT\}, \{HH, HT, TT\}, \{HH, HT, TH\}, \{HH, TT, TH\},$
$\phi, \Omega\}$

(c) Using the three axioms of probability:

(1) E be any event, i.e. $E \in \mathcal{F}$, then $P(E) \geq 0$

(2) w_i be any elementary events, i.e. $w_i \in \Omega \ \forall i$,
s.t. $\bigcup_i w_i = \Omega$, the $P(\bigcup_i w_i) = P(\Omega) = 1$

(3) for $w_1, w_2 .. w_k$ mutually exclusive events,
$P(w_1 \cup w_2 .. w_k) = \sum_{i=1}^{k} P(w_i)$

Also, it is known that all elementary events are mutually exclusive to each other.

Now $P(\Omega) = 1$

$\Rightarrow P(\{HH\} \cup \{HT\} \cup \{TH\} \cup \{TT\}) = 1$

$\Rightarrow P(\{HH\}) + .. P(\{TT\}) = 1$

now, given that all elementary events are equiprobable,
let $P(w_i) = p \ \ \forall w_i \in \Omega$.

$\Rightarrow p + p + p + p = 1 \Rightarrow (p = \frac{1}{4})$

So, probability of occurence of each elemaryevent is $\frac{1}{4}$ or 0.25.

(c) (ii) let event E_1 : atleast one heads occurs

$\Rightarrow \quad E_1 : \{\{HH\} \cup \{HT\} \cup \{TH\}\}$

$P(E_1) = P(\{HH\} \cup \{HT\} \cup \{TH\})$

$= P(HH) + P(HT) + P(TH) = 3p = \frac{3}{4}$

So, $P(E_1) = 0.75$

(iii) let event E_2 : exactly one head appears

$E_2 = \{HT\} \cup \{TH\}$

So, $P(E_2) = P(\{HT\} \cup \{TH\}) = 2p = \frac{1}{2}$

So, $P(E_2) = \frac{1}{2}$ or 0.5.

Part 2 : Discrete random variables

2.1

Given: probability assignment function

* $f(k, n, p) = \frac{n!}{k!(n-k)!} p^k (1-p)^{(n-k)}$

* k = # correctly recognised words = 45

n = total # words = 50

p = probability of guessing correctly = 0.9

so, P(correctly recognising 45 out of 50 words)

$= f(45, 50, 0.9) = \frac{50!}{45!\,5!} (0.9)^{45} (0.1)^5$

$\simeq 0.185$

2.2 Given: PMF: $f(k,\lambda) = \dfrac{\lambda^k e^{-\lambda}}{k!}$

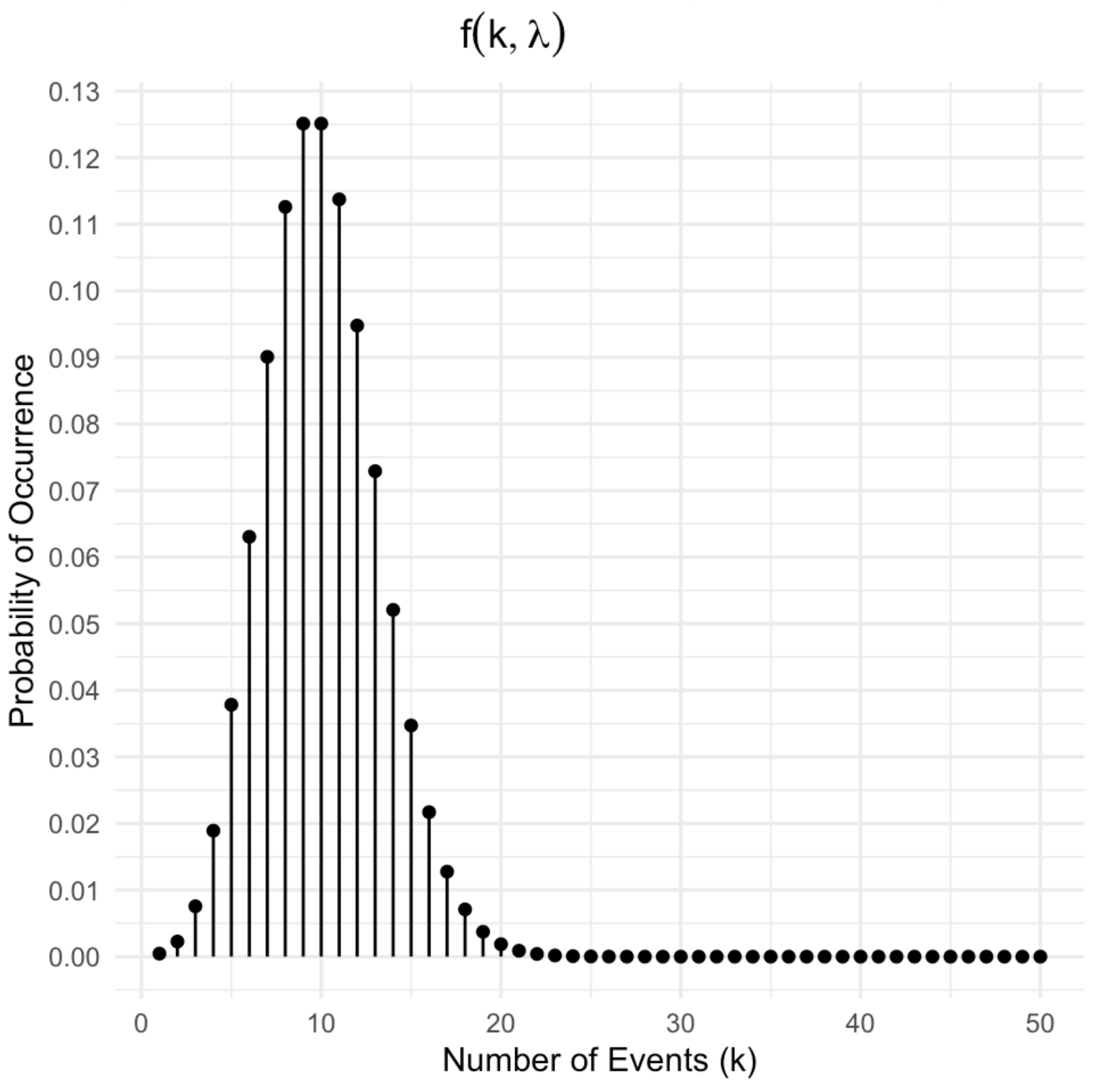
and $\lambda = 10$

(a) P(zero road accidents) $= P(k=0) = \dfrac{10^0 e^{-10}}{0!} = e^{-10} \approx 4.54 \times 10^{-5}$

(b) P(more than 7 but less than 10) $= P(\{k=8\} \cup \{k=9\})$

$= P(k=8) + P(k=9) = \dfrac{10^8 \cdot e^{-10}}{8!} + \dfrac{10^9 e^{-10}}{9!} \approx 0.24$

(c) We calculate the values of $f(k,\lambda)$ for $k = 1, 2 \ldots 50$ and plot the graph of PMF below. Graph is drawn using ggplot in R language (code added to the end of the PDF file).



Part 3 : Continuous Random Variables

3.1

given : pdf : $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

(a) $f(X=0;\ \mu=1, \sigma=1) = \frac{1}{(1)\sqrt{2\pi}} e^{-\frac{(0-1)^2}{2\cdot(1)}} = \frac{e^{-0.5}}{\sqrt{2\pi}} \approx$

(b) $f(x=1;\ \mu=0, \sigma=1) = \frac{1}{(1)\sqrt{2\pi}} e^{-\frac{(1-0)^2}{2\cdot(1)}} = \frac{e^{-0.5}}{\sqrt{2\pi}} \approx$

(c) we know that

$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x) \cdot dx$ where $f(x)$ is the p.d.f. of random var. X.

so, given $P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x) dx = 0.3$ ... ①

$P(x_1 \leq X \leq x_3) = \int_{x_1}^{x_3} f(x) dx = 0.45$ ... ②

subtract ① from ②

$\Rightarrow \int_{x_1}^{x_3} f(x) dx - \int_{x_1}^{x_2} f(x) dx = 0.15$

$\Rightarrow \int_{x_2}^{x_3} f(x) dx + \int_{x_1}^{x_2} f(x) dx - \int_{x_1}^{x_2} f(x) dx = 0.15$

$\Rightarrow \int_{x_2}^{x_3} f(x) dx = 0.15 = P(x_2 \leq X \leq x_3)$

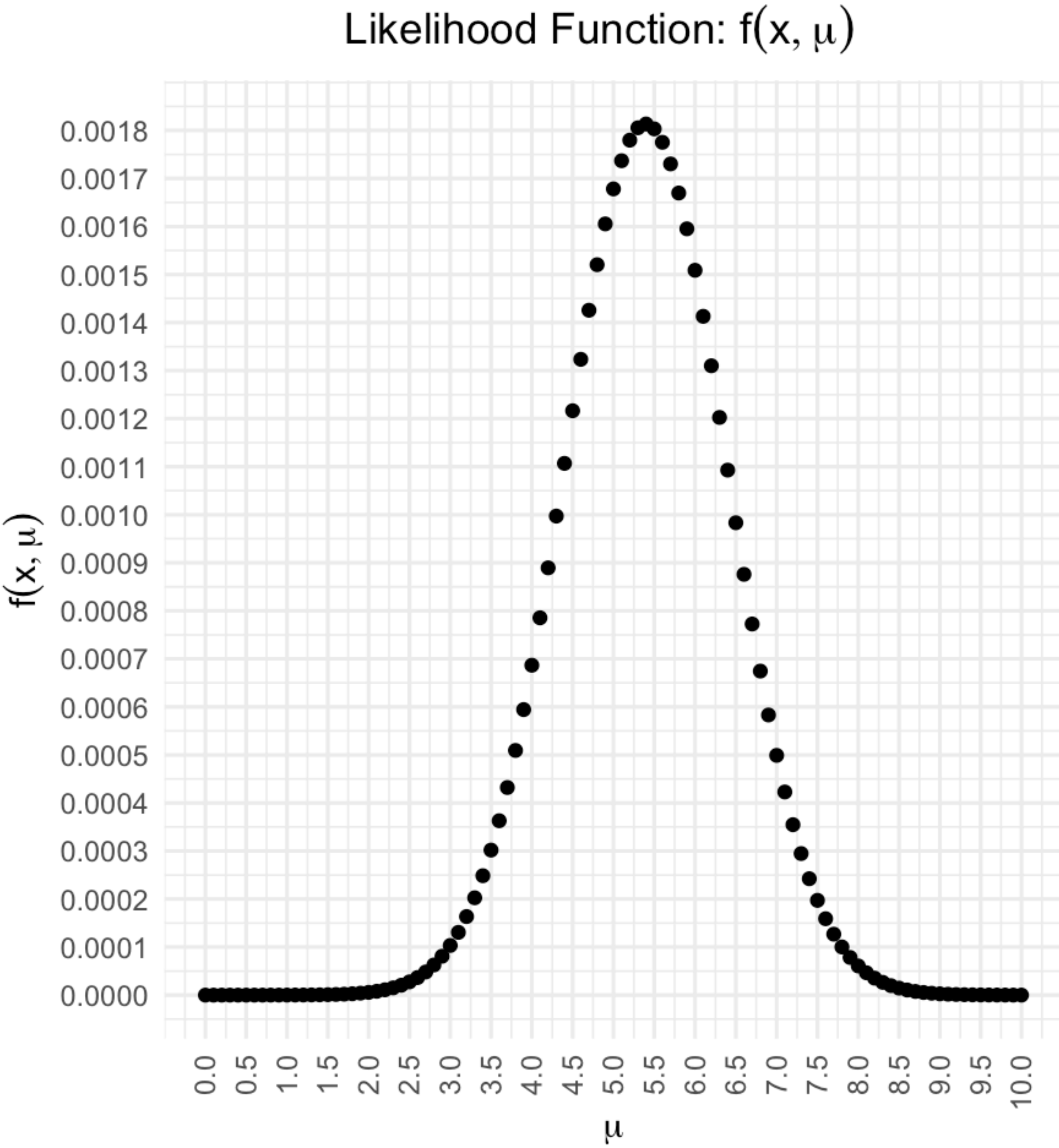
Part 4 : The Likelihood function

4.1:

given : $f(x,\mu) = \dfrac{1}{x\sqrt{2\pi}} e^{-\frac{(\log x - \mu)^2}{2}}$

(a) when you fix x to 220, $f(220,\mu) = \dfrac{1}{220\sqrt{2\pi}} e^{-\frac{(\log(220) - \mu)^2}{2}}$

Again using this function and R-language, the graph of the likelihood function is plotted below and code is attached at the end of the pdf.



(b) let ~~g(x,u)~~ $f(x_0,\mu)$ be the likelihood fxn at fixed $x = x_0$

then likelihood fn. for observing $x_1, x_2, \ldots, x_k$ is

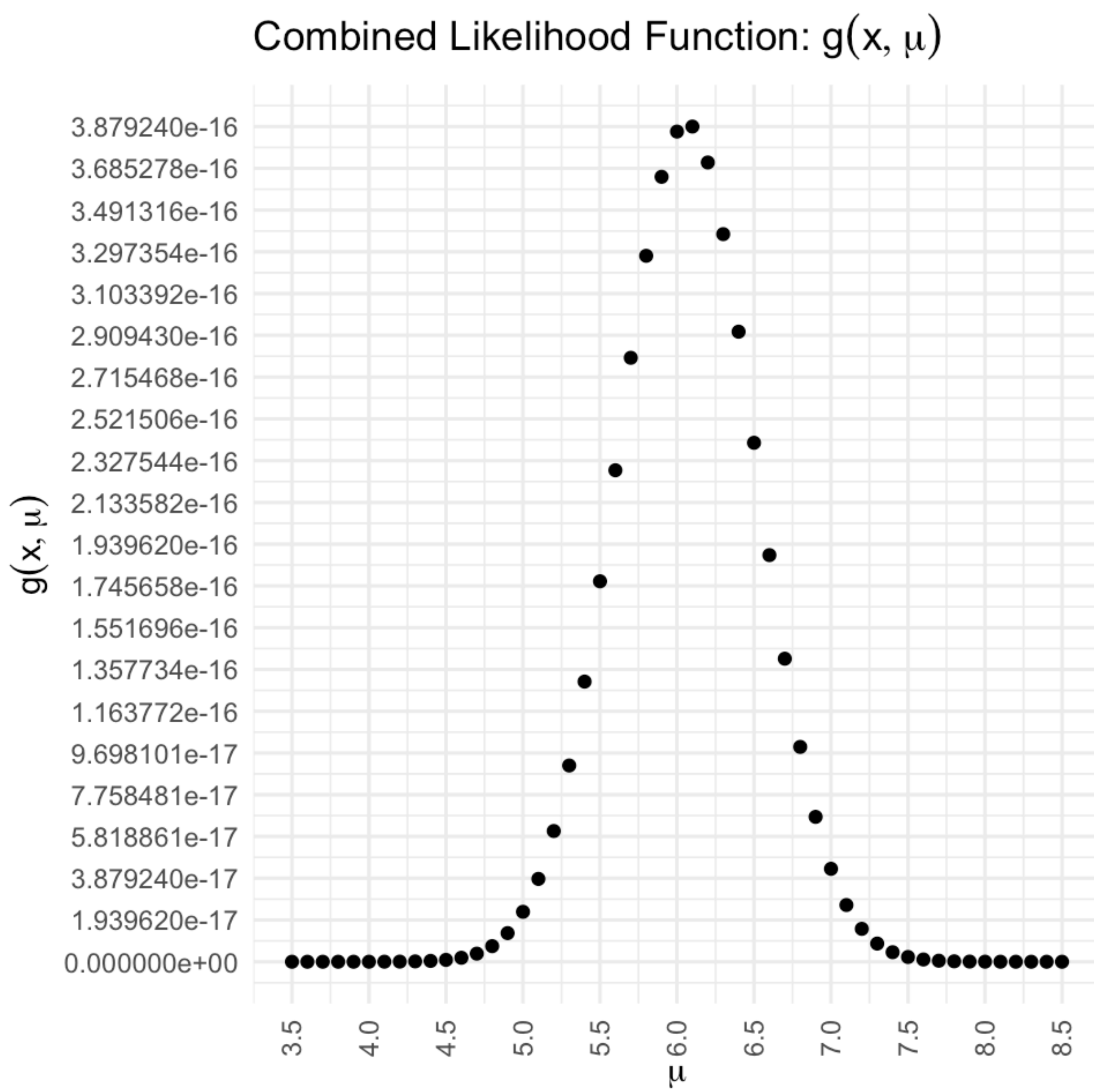
$$g(\cancel{x},\mu) = f(x_1,\mu)\cdot f(x_2,\mu)\ldots f(x_k,\mu)$$

$$\Rightarrow g(\cancel{x}\,\mu) = \frac{1}{\left(\prod_i^n x_i\right)\sqrt{2\pi}}\, e^{-\left(\frac{\sum_i^n (\log x - \mu)^2}{2}\right)}$$

using this function and R language for

$x_1 = 303.5$, $x_2 = 443$, $x_3 = 220$, $x_4 = 560$, $x_5 = 880$

the following plot is generated, code is attached at the end of the pdf file.



(c) $g(\mu)$ = likelihood fn for 303, 443, 220, 560, 880

$$g(\mu) = \frac{e^{-\frac{\sum(\log x_i - \mu)^2}{2}}}{\prod x_i \left(\sqrt{2\pi}\right)}$$ is greatest when

e^{-y} is greatest i.e y is lowest $\Rightarrow y = 0$

$\Rightarrow$ $g(\mu)$ is greatest when $\sum_i (\log x_i - \mu)^2 = 0$

$\Rightarrow 5\mu^2 - 2\mu \sum_i \log(x_i) + \sum_i \log^2(x_i) = 0 \Rightarrow 5\mu^2 - 26.325\mu + 34.87 = 0$

$\Rightarrow \boxed{\mu \approx 6.062}$

Code for 2.2 (c)

```r
library(ggplot2)

poisson_pmf <- function(k, lambda) {
  (lambda^k * exp(-lambda)) / factorial(k)
}

lambda <- 10
k_values <- 1:50

# Compute PMF values for each k
pmf_values <- sapply(k_values, poisson_pmf, lambda = lambda)

# Create a data frame for plotting
data <- data.frame(k = k_values, PMF = pmf_values)

# Plot the PMF using ggplot2 with vertical lines and detailed axis labels
ggplot(data, aes(x = k, y = PMF)) +
  geom_point() +
  geom_segment(aes(x = k, xend = k, y = 0, yend = PMF)) +
  labs(title = expression(paste("Poisson PMF: ", f(k, lambda))),
       x = "Number of Events (k)",
       y = "Probability of Occurrence") +
  scale_y_continuous(breaks = seq(0, 0.5, by = 0.01)) +
  theme_minimal()
```

Code for 4.1 (a)

```r
library(ggplot2)

f <- function(x, mu) {
  return (1/(x * sqrt(2*pi))) * exp(-((log(x) - mu)^2)/2)
}

x_fixed <- 220
mu_values <- seq(3.5, 8.5, length.out = 100)  # Range of mu values

data <- data.frame(mu = mu_values, fx_mu = sapply(mu_values, function(mu) f(x_fixed, mu)))

# Plot the function using ggplot2
ggplot(data, aes(x = mu, y = fx_mu)) +
  geom_line() +
  labs(title = "Plot of f(x, mu) for x = 220",
       x = "Value of mu",
       y = "f(x, mu)") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5))
```

Code for 4.1 (b)

```r
library(ggplot2)

f <- function(x, mu) {
  return (1/(x * sqrt(2*pi))) * exp(-((log(x) - mu)^2)/2)
}

x_values <- c(303.5, 443, 220, 560, 880)
mu_values <- seq(3.5, 8.5, length.out = 100)  # Range of mu values

# Calculate likelihood for each x and mu
likelihoods <- sapply(x_values, function(x) sapply(mu_values, function(mu) f(x, mu)))

# Calculate the combined likelihood (product of likelihoods for each x)
combined_likelihood <- apply(likelihoods, 2, prod)

# Create data frame for plotting
data <- data.frame(mu = mu_values, combined_likelihood = combined_likelihood)

# Plot the combined likelihood using ggplot2
ggplot(data, aes(x = mu, y = combined_likelihood)) +
  geom_line() +
  labs(title = "Combined Likelihood for Fixed Values of x",
       x = "Value of mu",
       y = "Combined Likelihood") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5))
```