STA310 HW3

Olivia Fu

2025-02-07

```r
library(ggplot2)
```

Exercise 1

The probability mass function of a Poisson random variable is given by:

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, ...$$

The likelihood function is:

$$\begin{aligned} L(\lambda) &= \prod_{i=1}^{n} P(X_i = x_i) \\ &= \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \\ &= \lambda^{\sum_{i=1}^{n} x_i} \cdot e^{-n\lambda} \cdot \prod_{i=1}^{n} \frac{1}{x_i!} \end{aligned}$$

Exercise 2

The log-likelihood function is:

$$\begin{aligned}
\ell(\lambda) &= \log L(\lambda) \\
&= \log\left(\lambda^{\sum_{i=1}^{n} x_i} \cdot e^{-n\lambda} \cdot \prod_{i=1}^{n} \frac{1}{x_i!}\right) \\
&= \log(\lambda^{\sum_{i=1}^{n} x_i}) + \log(e^{-n\lambda}) + \log(\prod_{i=1}^{n} \frac{1}{x_i!}) \\
&= \sum_{i=1}^{n} x_i \cdot \log \lambda - n\lambda + \sum_{i=1}^{n} \log \frac{1}{x_i!}
\end{aligned}$$

The first derivative of the log-likelihood function is:

$$\frac{\partial \ell(\lambda)}{\partial \lambda} = \frac{\sum_{i=1}^{n} x_i}{\lambda} - n$$

We then set the derivative equal to zero to find the MLE:

$$\begin{aligned}
\frac{\sum_{i=1}^{n} x_i}{\hat{\lambda}} - n &= 0 \\
\hat{\lambda} &= \frac{\sum_{i=1}^{n} x_i}{n}
\end{aligned}$$

The second derivative of the log-likelihood function is:

$$\frac{\partial^2 \ell(\lambda)}{\partial \lambda^2} = -\frac{\sum_{i=1}^{n} x_i}{\lambda^2}$$

Since $\lambda > 0$ and $x_i \geq 0$, the second derivative is negative (exclude the trivial case where all $x_i = 0$), confirming that the critical point corresponds to a maximum.

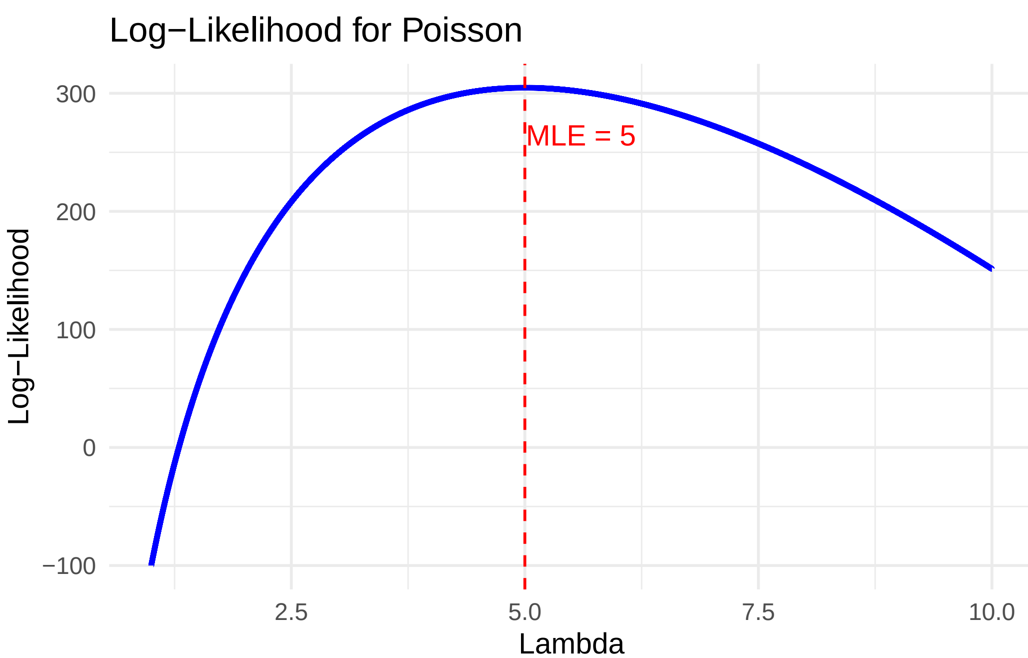
Exercise 3

```r
# Given data
n <- 100
sum_x <- 500

# Log-likelihood function for Poisson distribution
log_likelihood_function <- function(lambda) {
  if (lambda <= 0) {
    return(-Inf)
  } else {
    # The constant term can be ignored as
    # it does not affect the maximization of the log-likelihood
    return(sum_x * log(lambda) - n * lambda)
  }
}

# Set up a sequence of lambda values
lambda_values <- seq(1, 10, length.out = 100000)

# Calculate the log-likelihood for each lambda
log_likelihood_values <- sapply(lambda_values, log_likelihood_function)

# Find the lambda that maximizes the log-likelihood
mle_lambda <- lambda_values[which.max(log_likelihood_values)]

# Print the MLE for lambda
cat("MLE for lambda:", mle_lambda, "\n")
```

```
MLE for lambda: 5
```

```r
# Create a dataframe
data <- data.frame(lambda = lambda_values, log_likelihood = log_likelihood_values)

# Plot
ggplot(data, aes(x = lambda, y = log_likelihood)) +
  geom_line(color = "blue", size = 1) +
  geom_vline(xintercept = mle_lambda, color = "red", linetype = "dashed") +
  annotate("text",
           x = mle_lambda + 0.6,
           y = max(log_likelihood_values) - 40,
```

```r
           label = paste("MLE =", round(mle_lambda, 3)),
           color = "red") +
  labs(title = "Log-Likelihood for Poisson",
       x = "Lambda",
       y = "Log-Likelihood") +
  theme_minimal()
```



Therefore, the approximate MLE is 5, which matches the formula we derived in Exercise 2.

$$\hat{\lambda} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{500}{100} = 5$$

Exercise 4

(a)

Game	First five shots	Likelihood (no hot hand)	Likelihood (hot hand)
1	BMMBB	$p_B^3(1-p_B)^2$	$(p_B)(1-p_{B\|B})(1-p_B)(p_B)(p_{B\|B}) = (p_B)^2(1-p_{B\|B})(1-p_B)(p_{B\|B})$
2	MBMBM	$p_B^2(1-p_B)^3$	$(1-p_B)(p_B)(1-p_{B\|B})(p_B)(1-p_{B\|B}) = (p_B)^2(1-p_{B\|B})^2(1-p_B)$
3	MMBBB	$p_B^3(1-p_B)^2$	$(1-p_B)(1-p_B)(p_B)(p_{B\|B})(p_{B\|B}) = (p_B)(1-p_B)^2(p_{B\|B})^2$
4	BMMMB	$p_B^2(1-p_B)^3$	$(p_B)(1-p_{B\|B})(1-p_B)(1-p_B)(p_B) = (p_B)^2(1-p_{B\|B})(1-p_B)^2$
5	MMMMM	$(1-p_B)^5$	$(1-p_B)(1-p_B)(1-p_B)(1-p_B)(1-p_B) = (1-p_B)^5$
Total		$p_B^{10}(1-p_B)^{15}$	$(p_B)^7(1-p_B)^{11}(p_{B\|B})^3(1-p_{B\|B})^4$

(b)

```
likelihood <- function(pb){
  pb^(10) * (1 - pb)^(15)
}
likelihood(0.4)
```

```
[1] 4.930247e-08
```

```
likelihood(0.3)
```

```
[1] 2.803388e-08
```

When we substitute 0.4 and 0.3 for p_B in the likelihood function, we find that $p_B = 0.4$ produces a higher likelihood than $p_B = 0.3$ (4.93×10^{-8} vs. 2.80×10^{-8}). This means that the data is more consistent with $p_B = 0.4$.