ST117 Lab 9 Workbook

11 March, 2024

1. Linear Regression Model

Consider the regression model:

$$y_i = \alpha + \beta x_i + \epsilon_i$$

where $\epsilon_i \sim \mathrm{N}(0, \sigma^2)$ for $i = 1, \ldots, n$. Here, y represents heights(in cm) and x represents weights(in kg).

Given a sample of heights and weights,

```
x<-c(82.62954, 66.73767, 83.29799, 82.72429, 74.14641, 54.60050, 60.71433,
     67.05280, 69.94233, 94.04653)
y<-c(202.5511,166.1328,194.1981,197.4564,181.9675,146.2233,160.5469, 166.2354,
     178.0746, 213.0961)
```

(a) Find the maximum likelihood estimates for α and β based on the given sample.

```
#TODO: write your codes here
```

(b) Plot a scatter plot of the sample and include the estimated regression line using the MLE of α and β. Providing the exact regression equation $y = 50 + 1.8 * x + \epsilon$, where $\epsilon \sim N(0, 5^2)$, add this line into the plot.

```
#TODO: write your codes here
```

(c) Generate 100 simulations and calculate 100 values of $\hat{\alpha}$ and $\hat{\beta}$ using the regression $y = 50 + 1.8 * x + \epsilon$, where $\epsilon \sim N(0, 5^2)$ and $x \sim N(70, 10^2)$. Calculate the variance of the MLE $(\hat{\alpha}, \hat{\beta})$ using R and the formulas provided in Tuesday-1 Week 8 Lecture notes:

$$\mathrm{Var}(\hat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$\mathrm{Var}(\hat{\alpha}) = \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \right)$$

```
#TODO: write your codes here
```

(d) Find the estimators' variances as the sample size n increases, considering $n = 100, 1000, 10000$.

```
#TODO: write your codes here
```

2. Diagnostic Check

(a) Consider `women` dataset which contains average heights and weights for American Women. Create a scatter plot of this data to examine the relationship between heights and weights. Is there evidence of a linear relationship?

```
 #access the dataset
data(women) #access the dataset
```

```r
head(women)
```

```
##   height weight
## 1     58    115
## 2     59    117
## 3     60    120
## 4     61    123
## 5     62    126
## 6     63    129
```

```r
#TODO: write your codes here
```

(b) Fit a linear model to the data and perform a regression diagnostics check, including standard linear model diagnostics.

```r
#TODO: write your codes here
```