1. Assignment 6

1.1. **Problem 1.**

Suppose $X = (X_1, X_2, \ldots, X_n)$ is a vector of independent, identically distributed random variables having an exponential distribution with parameter θ_1 and $Y = (Y_1, Y_2, \ldots, Y_m)$ is a vector of independent identically distributed random variables having an exponential distribution with parameter θ_2. X and Y are independently distributed.

a) Find the value of θ_1 that maximizes the likelihood function $L(X, \theta_1)$

b) Find the value of θ_2 that maximizes the likelihood function $L(Y, \theta_2)$

c) Considering the concatenation of X and Y into a single vector of length $n + m$ given by $U = (X_1, \ldots, X_n, Y_1, \ldots, Y_m)$, what are the values of θ_1 and θ_2 that maximize the likelihood function $L(U, \theta_1, \theta_2) = L(X, \theta_1) \cdot L(Y, \theta_2)$?

d) Again considering the concatenation of X and Y into a single vector of length $n + m$ given by $U = (X_1, \ldots, X_n, Y_1, \ldots, Y_m)$, with the restriction that $\theta_1 = \theta_2 = \theta$, what is the value of θ that maximize the likelihood function $L(U, \theta)$?

1.2. **Problem 2.**

In the same setting as Problem 1, suppose

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = 13.6 \quad \text{and} \quad \overline{Y} = \frac{\sum_{i=1}^{m} Y_i}{m} = 14.1$$

with $n = m = 100$. We are interested in whether the data supports the claim that $\theta_1 = \theta_2$.

a) Find the likelihood ratio

$$\lambda = \frac{\max L(U, \theta)}{\max L(U, \theta_1, \theta_2)}$$

b) Using the approximate distribution of $-2 \ln \lambda$ (Theorem 10.2), construct an acceptance region for testing $H_0 : \theta_1 = \theta_2$ against the alternative $H_a : \theta_1 \neq \theta_2$ of size $1 - \alpha = .95$. (Hint: The number of parameters needed to fully specify the likelihood is two. The number of free parameters is two minus the number of parameter values specified).

c) Does the value of λ computed in part a) fall inside the acceptance region?

1.3. **Problem 3.** A linear model with a single continuous predictor is called a **simple regression** model and has the form

$$Y = X\beta + e$$

where

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \quad e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

a) Write an expression for $X'X$

b) Write an expression for $X'Y$

c) Recall that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Use this to write expressions for $\hat{\beta}_1$ and $\hat{\beta}_2$ where

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = (X'X)^{-1} X'Y$$

1.4. **Problem 4.** In some simple regression models, the assumption that the regression line passes through the origin can be justified. This is called a **no intercept** model and has the form

$$Y = X\beta + e$$

where

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \end{bmatrix}$$

a) Write an expression for $X'X$

b) Write an expression for $X'Y$

c) Write an expression for $\hat{\beta}_1$ where

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \end{bmatrix} = (X'X)^{-1} X'Y$$