Q(1) Consider a dataset $(x_1, y_1), (x_2, y_2) \ldots, (x_n, y_n)$ where x_i is the predictor and y_i is the response variable for $i = 1, 2, \ldots, n$. We assume the simple linear regression model:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

where $\varepsilon_i \sim N(0, \sigma^2)$ are independent and identically distributed normal random variables.

(a) Likelihood function $L(\beta_0, \beta_1, \sigma^2)$

The likelihood function gives us the probability of observing our data given the parameter β_0, β_1 and σ^2. Since the errors (ε_i) are normally distributed, we can say that y_i also follows a normal distribution.

For this normal distribution to find the probability density function (PDF):

$$P(X = x) = \frac{1}{\sqrt{2\pi\sigma^2}} * e\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$$

where, we know that the response y_i is normally distributed with mean $\beta_0 + \beta_1 x_i$ and the variance is σ^2.

where the PDF can be written as:

$$P(y_i \mid \beta_0, \beta_1, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2}\right)$$

from this we can say that, the observations are independent, where the joint likelihood function for the complete dataset is

$$L(\beta_0, \beta_1, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2}\right)$$

Thus:

$$L(\beta_0, \beta_1, \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(\sum_{i=1}^{n} \frac{(y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2}\right)$$

(b) Log-Likelihood Function $l(\beta_0, \beta_1, \sigma^2)$

Taking the natural logarithm of the likelihood function helps to simplify our calculations. The log-likelihood function $l(\beta_0, \beta_1, \sigma^2)$ is natural log of likelihood function, which means

$$l(\beta_0, \beta_1, \sigma^2) = \log L(\beta_0, \beta_1, \sigma^2)$$

we can write the logarithm of the $L(\beta_0, \beta_1, \sigma^2)$

$$l(\beta_0, \beta_1, \sigma^2) = \log\left[\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(\sum_{i=1}^{n} \frac{(y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2}\right)\right]$$

using the properties of logarithm we can write as:

$$l(\beta_0, \beta_1, \sigma^2) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

The log-likelihood is easier to work with, especially when finding the maximum likelihood estimates (MLEs).

(c) Maximum Likelihood Estimates (MLEs)

To find the MLEs for $\beta_0, \beta_1, \sigma^2$ we need to differentiate the log-likelihood function wrt each parameter and set these derivatives to zero.

We know the parameters are $\beta_0, \beta_1, \sigma^2$.

(i) Differentiating w.r.t to β_0:

The log-likelihood function contain a terms

$$\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 \quad \text{w.r.t } \beta_0.$$

$$\frac{\partial}{\partial \beta_0} l(\beta_0, \beta_1, \sigma^2) = \frac{1}{\sigma^2} \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)$$

where we are equal to zero gives us the first normal equation:

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

which can be written as:

$$n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

Therefore: $\boxed{\beta_0 = \bar{y} - \beta_1 \bar{x}}$ $\left(\because \bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} ; \bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}\right)$

(ii) Differentiate w.r.t to β_1:

$$\frac{\partial l}{\partial \beta_1} = \frac{1}{\sigma^2} \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i) x_i = 0$$

Similar to (i), equating to zero we get the second normal equation

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

we know the $y = \beta_0 + \beta_1 x \Rightarrow \beta_0 = \bar{y} - \beta_1 \bar{x}$

substituteng the β_0 in the equation

$$\sum_{i=1}^{n} (y_i - \bar{y} - \beta_1 \bar{x} - \beta_1 x_i) x_i = 0$$

$$\sum_{i=1}^{n} (y_i - \bar{y} - \beta_1 (\bar{x} - x_i)) x_i = 0$$

$$\beta_1 \sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

where the MLE for β_1 is

$$\boxed{\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}}$$

(iii) Differentiating wrt to σ^2

$$\frac{\partial l}{\partial \sigma^2} = \frac{-n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

similar to above, equating to zero the third normal equation is

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

therfore MLE is σ^2 is

$$\boxed{\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2}$$

(d) Least squares Estimates

The MLEs for β_0 and β_1 are identical to the least squares estimates.

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

from part c, we know the MLE for β_0, β_1

These are exactly the formula for the least square estimates of β_0 & β_1

so the MLE for β_0 & β_1 is the corresponding to the least square estimates.

(e) The MLE of σ^2 is

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2.$$

from the part c, we the MLE - for σ^2

where this formula for the MSE in simple linear regression and

this is the standard result for the MLE of σ^2 in σ^2 in

simple linear regression model.