Statistics (II) Quiz (2)

Version A Solution

(I) Multiple Choice (7 pts each)

1. B
2. C
3. A
4. C
5. B
6. D
7. B
8. A
9. C
10. A
11. C
12. B
13. C
14. A
15. C

(II) Bonus Question (+15 pts)

Let $(x_i, y_i), i = 1, \ldots, n$, be a sample of paired observations. Consider the simple linear regression model, use the least squares method to estimate the regression coefficients.
The simple linear regression model is:

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

1. Objective Function

Minimize the sum of squared error (SSE):

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

2. Partial Differentiation and Derivation

To minimize SSE, we take partial derivatives with respect to β_0 and β_1. Using the chain rule:

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

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

3. Deriving Normal Equations

Let (b_0, b_1) denote the particular value of (β_0, β_1) that minimizes Q.

$$\begin{cases} \sum_{i=1}^{n} 2(y_i - b_0 - b_1 x_i) \cdot (-1) = 0 \\ \sum_{i=1}^{n} 2(y_i - b_0 - b_1 x_i) \cdot (-x_i) = 0 \end{cases}$$

Divide by -2:

$$\begin{cases} \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i) = 0 \\ \sum_{i=1}^{n} (x_i y_i - b_0 x_i - b_1 x_i^2) = 0 \end{cases}$$

Distribute the summation:

$$\begin{cases} \sum y_i - \sum b_0 - \sum b_1 x_i = 0 \\ \sum x_i y_i - b_0 \sum x_i - b_1 \sum x_i^2 = 0 \end{cases}$$

Since $\sum_{i=1}^{n} b_0 = n b_0$:

$$\begin{cases} \sum y_i = n b_0 + b_1 \sum x_i \quad \text{(Equation 1)} \\ \sum x_i y_i = b_0 \sum x_i + b_1 \sum x_i^2 \quad \text{(Equation 2)} \end{cases}$$

4. Solving for Coefficients

From (1):

$$b_0 = \frac{\sum y_i - b_1 \sum x_i}{n} = \bar{y} - b_1 \bar{x}$$

Substituting b_0 into (2) and solving for b_1:

$$b_1 = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$