LECTURE - 32

Agenda:

1) Properties of least squares estimators

We have two characteristics Y and X in the population or phenomenon that is under consideration. The linear model says that

$$Y = \beta_0 + \beta_1 X + \varepsilon,$$

where ε is assumed to be a random error with $E[\varepsilon] = 0$ and $V(\varepsilon) = \sigma^2$. Typically, we have n independent data points $(X_1, Y_1), (X_2, Y_2), ---, (X_n, Y_n)$ and hence,

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i,$$

where $\varepsilon_1, \varepsilon_2, ---, \varepsilon_n$ are independent and identically distributed with mean 0 and variance σ^2. The "LEAST SQUARES" estimates of β_0 and β_1 are defined as

$$(\hat{\beta}_0, \hat{\beta}_1) = \underset{\beta_0, \beta_1}{\arg\min} \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_i)^2.$$

In the previous lecture, we derived that

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

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

Today, we will study some properties of $\hat{\beta}_0$ and $\hat{\beta}_1$.

$\boxed{\text{PROPERTY 1:}}$ Both $\hat{\beta}_0$ and $\hat{\beta}_1$ are unbiased estimators of β_0 and β_1 respectively, i.e.,

$$E[\hat{\beta}_0] = \beta_0, \quad E[\hat{\beta}_1] = \beta_1.$$

PROOF (sketch): Note first that,

$$\sum_{i=1}^{n}(x_i-\bar{x})(Y_i-\bar{Y}) = \sum_{i=1}^{n}(x_i-\bar{x})Y_i \quad \text{(Why?)}$$

Hence,
$$E[\hat{\beta}_1] = E\left[\frac{\sum_{i=1}^{n}(x_i-\bar{x})Y_i}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right]$$

$$= \frac{1}{\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)} E\left[\sum_{i=1}^{n}(x_i-\bar{x})Y_i\right]$$

PROPERTY 2: $\quad V(\hat{\beta}_1) = \sigma^2 / \sum_{i=1}^{n} (x_i - \bar{x})^2$

and $V(\hat{\beta}_0) = \dfrac{\sigma^2}{n} + \dfrac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$

PROOF (sketch):

$$V(\hat{\beta}_1) = V\left(\frac{\sum_{i=1}^{n} (x_i - \bar{x}) Y_i}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \right)$$

$$= \frac{1}{\left(\sum_{i=1}^{n} (x_i - \bar{x})^2\right)^2} V\left(\sum_{i=1}^{n} (x_i - \bar{x}) Y_i \right)$$

$$= \frac{1}{\left(\sum_{i=1}^{n} (x_i - \bar{x})^2\right)^2} \sum_{i=1}^{n} (x_i - \bar{x})^2 \, V(Y_i)$$

($\because$ By independence)

$$= \frac{1}{\left(\sum_{i=1}^{n} (x_i - \bar{x})^2\right)^2} \sum_{i=1}^{n} (x_i - \bar{x})^2 \, V(\beta_0 + \beta_1 x_i + \varepsilon_i)$$

$$= \frac{1}{\left(\sum_{i=1}^{n} (x_i - \bar{x})^2\right)^2} \sum_{i=1}^{n} (x_i - \bar{x})^2 \, V(\varepsilon_i)$$

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

$$= \frac{\sigma^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}.$$

$$= \frac{1}{\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)} \sum_{i=1}^{n}(x_i-\bar{x})\, E[Y_i]$$

$$= \frac{1}{\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)} \sum_{i=1}^{n}(x_i-\bar{x})(\beta_0+\beta_1 x_i)$$

Note that

$$\sum_{i=1}^{n}(x_i-\bar{x})(\beta_0+\beta_1 x_i) = \beta_0 \sum_{i=1}^{n}(x_i-\bar{x}) + \beta_1 \sum_{i=1}^{n}(x_i-\bar{x})x_i$$

$$= \beta_0\left(\sum_{i=1}^{n} x_i - n\bar{x}\right) + \beta_1\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right) \quad \text{(why?)}$$

$$= 0 + \beta_1\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)$$

$$= \beta_1\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right).$$

Hence, it follows that,

$$E[\hat{\beta}_1] = \frac{1}{\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)} \beta_1\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)$$

$$= \beta_1 .$$

Similarly, one can prove that $E[\hat{\beta}_0] = \beta_0$.

Similarly, one can prove that

$$V(\hat{\beta}_0) = \frac{\sigma^2}{n} + \frac{\sigma^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}.$$

Note that σ^2 is also an unknown parameter in the model. What is a good estimate of σ^2?

Note that the "ERROR SUM OF SQUARES" is defined as

$$\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)^2.$$

The "ESTIMATED ERROR SUM OF SQUARES" is defined as

$$SSE = \sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2$$

PROPERTY 3: $\hat{\sigma}^2 = \frac{1}{n-2}\sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2$ is an unbiased estimate of σ^2.