Agenda:

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

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

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

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

where $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ are independent and identically distributed with mean 0 and variance $\sigma^2$. The "LEAST SQUARES" estimates of $\beta_0$ and $\beta_1$ are defined as

$$(\hat{\beta}_0, \hat{\beta}_1) = \arg \min_{\beta_0, \beta_1} \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_i)^2.$$
In the previous lecture, we derived that

\[ \hat{\beta}_2 = \frac{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2} \]

and \[ \hat{\beta}_0 = \bar{y} - \hat{\beta}_2 \bar{x} \).

Today, we will study some properties of \( \hat{\beta}_0 \) and \( \hat{\beta}_2 \).

**PROPERTY 1:** Both \( \hat{\beta}_0 \) and \( \hat{\beta}_2 \) are unbiased estimators of \( \beta_0 \) and \( \beta_2 \) respectively, i.e.,

\[ E[\hat{\beta}_0] = \beta_0, \quad E[\hat{\beta}_2] = \beta_2. \]

**PROOF (sketch):** Note first that,

\[ \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}) y_i \quad (\text{Why?}) \]

Hence,

\[ E[\hat{\beta}_2] = E \left[ \frac{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}) y_i}{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2} \right] \]

\[ = \frac{1}{\left(\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2\right)} \quad E \left[ \frac{\sum_{i=1}^{n} (x_i - \bar{x}) y_i}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \right] \]
**Property 2:**

\[ V(\hat{\beta}_2) = \sigma^2 / \sum_{i=1}^{n} (x_i - \bar{x})^2 \]

and \[ V(\hat{\beta}_0) = \frac{\sigma^2}{\sum_{i=1}^{n} y_i} + \frac{\sigma^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \]

**Proof (sketch):**

\[ V(\hat{\beta}_2) = V\left( \frac{n}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \left( \sum_{i=1}^{n} \frac{(x_i - \bar{x}) y_i}{(x_i - \bar{x})^2} \right) \right) \]

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

\[ = \frac{1}{\left( \sum_{i=1}^{n} (x_i - \bar{x})^2 \right)^2} \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2 V(y_i)}{(x_i - \bar{x})^2} \quad \text{(by independence)} \]

\[ = \frac{1}{\left( \sum_{i=1}^{n} (x_i - \bar{x})^2 \right)^2} \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2 V(f_0 + \hat{\beta}_2 x_i + \epsilon_i)}{(x_i - \bar{x})^2} \]

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

\[ = \sigma^2 / \sum_{i=1}^{n} (x_i - \bar{x})^2 \]

\[ = \sigma^2 / \sum_{i=1}^{n} (x_i - \bar{x})^2. \]
\[
\frac{1}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sum_{i=1}^{n} (x_i - \bar{x}) E[\beta_i] = \frac{1}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sum_{i=1}^{n} (x_i - \bar{x}) (\beta_0 + \beta_1 x_i)
\]

Note that
\[
\sum_{i=1}^{n} (x_i - \bar{x}) (\beta_0 + \beta_1 x_i) = \beta_0 \sum_{i=1}^{n} (x_i - \bar{x}) + \beta_1 \sum_{i=1}^{n} (x_i - \bar{x}) x_i
\]

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

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

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

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

\[
= \beta_2.
\]

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 \( \sigma^2 \) is also an unknown parameter in the model. What is a good estimate of \( \sigma^2 \)?

Note that the "ERROR SUM OF SQUARES" \( SSE \) is defined as

\[ \frac{1}{n} \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_2 x_i)^2 \]

The "ESTIMATED ERROR SUM OF SQUARES" \( SSE \) is defined as

\[ SSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_2 x_i)^2 \]

**Property 3:** \( \frac{\hat{\sigma}^2}{n-2} \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_2 x_i)^2 \) is an unbiased estimate of \( \sigma^2 \).