Properties of Least Squares Estimators

Simple Linear Regression

Model: $Y = \beta_0 + \beta_1 x + \epsilon$

- ϵ is the random error so Y is a random variable too.

Sample:

$$(x_1, Y_1), (x_2, Y_2), \ldots, (x_n, Y_n)$$

Each (x_i, Y_i) satisfies $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$

Least Squares Estimators:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2}\,, \qquad \hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{x}$$

Properties of Least Squares Estimators

Assumptions on the Random Error ϵ

- $E[\epsilon_i] = 0$
- $V(\epsilon_i) = \sigma^2$

Implications for the Response Variable Y

- $E[Y_i] = \beta_0 + \beta_1 x_i$
- $V(Y_i) = \sigma^2$

What can be said about:

- $E[\hat{\beta}_1]$?
- $V(\hat{\beta}_1)$?
- $E[\hat{\beta}_0]$?
- $V(\hat{\beta}_0)$?
- $Cov(\hat{\beta}_0, \hat{\beta}_1)$?

Properties of Least Squares Estimators

Proposition: The estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ are unbiased; that is,

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

Proof:

$$\begin{aligned}
\hat{\beta}_1 &= \frac{\sum_{i=1}^{n}(x_i - \overline{x})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \\
&= \frac{\sum_{i=1}^{n}(x_i - \overline{x})Y_i - \overline{Y}\sum_{i=1}^{n}(x_i - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \\
&= \frac{\sum_{i=1}^{n}(x_i - \overline{x})Y_i}{\sum_{i=1}^{n}(x_i - \overline{x})^2}
\end{aligned}$$

Thus

$$
\begin{aligned}
E[\hat{\beta}_1] &= \sum_{i=1}^{n} \frac{(x_i - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \, E[Y_i] \\
&= \sum_{i=1}^{n} \frac{(x_i - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \, (\beta_0 + \beta_1 x_i) \\
&= \frac{\sum_{i=1}^{n}(x_i - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \, \beta_0 + \beta_1 \frac{\sum_{i=1}^{n}(x_i - \overline{x}) x_i}{\sum_{i=1}^{n}(x_i - \overline{x})^2}
\end{aligned}
$$

$$
\begin{aligned}
E[\hat{\beta}_0] &= E[\overline{Y} - \hat{\beta}_1 \overline{x}] \\
&= \frac{1}{n} \sum_{i=1}^{n} E[Y_i] - E[\hat{\beta}_1] \overline{x} \\
&= \frac{1}{n} \sum_{i=1}^{n} (\beta_0 + \beta_1 x_i) - \beta_1 \frac{1}{n} \sum_{i=1}^{n} x_i
\end{aligned}
$$

Properties of Least Squares Estimators

Proposition: The variances of $\hat{\beta}_0$ and $\hat{\beta}_1$ are:

$$V(\hat{\beta}_0) = \frac{\sigma^2 \sum_{i=1}^n x_i^2}{\sum_{i=1}^n (x_i - \overline{x})^2} = \frac{\sigma^2 \sum_{i=1}^n x_i^2}{S_{xx}}$$

and

$$V(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \overline{x})^2} = \frac{\sigma^2}{S_{xx}} \; .$$

Proof:

$$\begin{aligned}
V(\hat{\beta}_1) &= V\left(\frac{\sum_{i=1}^n (x_i - \overline{x}) Y_i}{S_{xx}}\right) \\
&= \left(\frac{1}{S_{xx}}\right)^2 \sum_{i=1}^n (x_i - \overline{x})^2 V(Y_i) \\
&= \left(\frac{1}{S_{xx}}\right)^2 \left(\sum_{i=1}^n (x_i - \overline{x})^2\right) \sigma^2 \\
&= \left(\frac{1}{S_{xx}}\right) \sigma^2
\end{aligned}$$

$$
\begin{aligned}
V(\hat{\beta}_0) &= V(\overline{Y} - \hat{\beta}_1 \overline{x}) \\
&= V(\overline{Y}) + V(-\hat{\beta}_1 \overline{x}) + 2Cov(\overline{Y}, -\overline{x}\hat{\beta}_1) \\
&= V(\overline{Y}) + \overline{x}^2 V(\hat{\beta}_1) - 2\overline{x} Cov(\overline{Y}, \hat{\beta}_1) \\
&= \frac{\sigma^2}{n} + \overline{x}^2 \left(\frac{\sigma^2}{S_{xx}} \right) - 2\overline{x} Cov(\overline{Y}, \hat{\beta}_1)
\end{aligned}
$$

Now let's evaluate the covariance term:

$$
\begin{aligned}
Cov(\overline{Y}, \hat{\beta}_1) &= Cov\left(\sum_{i=1}^{n} \frac{1}{n} Y_i, \sum_{j=1}^{n} \frac{x_j - \overline{x}}{S_{xx}} \, Y_j \right) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{x_j - \overline{x}}{n S_{xx}} \, Cov(Y_i, Y_j) \\
&= \sum_{i=1}^{n} \frac{x_i - \overline{x}}{n S_{xx}} \, \sigma^2 \qquad + \qquad 0 \\
&= 0
\end{aligned}
$$

Thus

$$
\begin{aligned}
V(\hat{\beta}_0) &= \frac{\sigma^2}{n} + \overline{x}^2 \left(\frac{\sigma^2}{S_{xx}} \right) \\
&= \sigma^2 \, \frac{S_{xx} + n\overline{x}^2}{nS_{xx}} \\
&= \sigma^2 \, \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2 + n\overline{x}^2}{nS_{xx}} \\
&= \sigma^2 \, \frac{\sum_{i=1}^{n}(x_i^2 - 2\overline{x}x_i + \overline{x}^2) + n\overline{x}^2}{nS_{xx}} \\
&= \sigma^2 \, \frac{\sum_{i=1}^{n} x_i^2}{nS_{xx}}
\end{aligned}
$$

Properties of Least Squares Estimators

A Practical Matter

The variance σ^2 of the random error ϵ is usually not known. So it is necessary to estimate it.

Proposition: The estimator

$$S^2 = \frac{1}{n-2} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 = \frac{1}{n-2} \, SSE$$

is an unbiased estimator of σ^2.

Properties of Least Squares Estimators

Assumptions on the Random Error ϵ

- $E[\epsilon_i] = 0$
- $V(\epsilon_i) = \sigma^2$
- **Each ϵ_i is normally distributed.**

Implications for the Estimators

- $\hat{\beta}_1$ is normally distributed with mean β_1 and variance $\frac{\sigma^2}{S_{xx}}$;
- $\hat{\beta}_0$ is normally distributed with mean β_0 and variance $\sigma^2 \frac{\sum_{i=1}^n x_i^2}{nS_{xx}}$;
- $\frac{(n-2)S^2}{\sigma^2}$ has a χ^2 distribution with $n-2$ degrees of freedom;
- S^2 is independent of $\hat{\beta}_0$ and $\hat{\beta}_1$.

Properties of Least Squares Estimators

Multiple Linear Regression

Model: $$Y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \epsilon$$

Sample:

$$\begin{array}{c} (x_{11}, x_{12}, \ldots, x_{1k}, Y_1) \\ (x_{21}, x_{22}, \ldots, x_{2k}, Y_2) \\ \vdots \\ (x_{n1}, x_{n2}, \ldots, x_{nk}, Y_n) \end{array}$$

Each (x_i, Y_i) satisfies $Y_i = \beta_0 + \beta_1 x_i + \cdots + \beta_k x_k + \epsilon_i$

Least Squares Estimators:

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

Properties of Least Squares Estimators

- Each $\hat{\beta}_i$ is an unbiased estimator of β_i: $E[\hat{\beta}_i] = \beta_i$;

- $V(\hat{\beta}_i) = c_{ii}\sigma^2$, where c_{ii} is the element in the ith row and ith column of $(\mathbf{X}'\mathbf{X})^{-1}$;

- $Cov(\hat{\beta}_i, \hat{\beta}_i) = c_{ij}\sigma^2$;

- The estimator

$$S^2 = \frac{SSE}{n-(k+1)} = \frac{\mathbf{Y}'\mathbf{Y} - \hat{\beta}'\mathbf{X}'\mathbf{Y}}{n-(k+1)}$$

  is an unbiased estimator of σ^2.

Properties of Least Squares Estimators

When ϵ is normally distributed,

- Each $\hat{\beta}_i$ is normally distributed;

- The random variable
$$\frac{(n-(k+1))S^2}{\sigma^2}$$
has a χ^2 distribution with $n-(k+1)$ degrees of freeedom;

- The statistics S^2 and $\hat{\beta}_i$, $i = 0, 1, \ldots, k$, are independent.