Gauss Markov Theorem

Dr. Frank Wood
Digression : Gauss-Markov Theorem

In a regression model where \( E\{\epsilon_i\} = 0 \) and variance \( \sigma^2\{\epsilon_i\} = \sigma^2 < \infty \) and \( \epsilon_i \) and \( \epsilon_j \) are uncorrelated for all \( i \) and \( j \) the least squares estimators \( b_0 \) and \( b_1 \) are unbiased and have minimum variance among all unbiased linear estimators.

Remember

\[
\begin{align*}
    b_1 &= \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} = \sum k_i Y_i, \quad k_i = \frac{(X_i - \bar{X})}{\sum (X_i - \bar{X})^2} \\
    b_0 &= \bar{Y} - b_1 \bar{X}
\end{align*}
\]

\[
\begin{align*}
    \sigma^2\{b_1\} &= \sigma^2\{\sum k_i Y_i\} = \sum k_i^2 \sigma^2\{Y_i\} \\
    &= \sigma^2 \frac{1}{\sum(X_i - \bar{X})^2}
\end{align*}
\]
Gauss-Markov Theorem

- The theorem states that $b_1$ has minimum variance among all unbiased linear estimators of the form

$$
\hat{\beta}_1 = \sum c_i Y_i
$$

- As this estimator must be unbiased we have

$$
E\{\hat{\beta}_1\} = \sum c_i E\{Y_i\} = \beta_1
$$

$$
= \sum c_i(\beta_0 + \beta_1 X_i) = \beta_0 \sum c_i + \beta_1 \sum c_i X_i = \beta_1
$$

- This imposes some restrictions on the $c_i$'s.
Proof

- Given these constraints

\[ \beta_0 \sum c_i + \beta_1 \sum c_i X_i = \beta_1 \]

clearly it must be the case that \( \sum c_i = 0 \) and \( \sum c_i X_i = 1 \)

- The variance of this estimator is

\[ \sigma^2 \{\hat{\beta}_1\} = \sum c_i^2 \sigma^2 \{Y_i\} = \sigma^2 \sum c_i^2 \]

- This also places a kind of constraint on the \( c_i \)'s
Proof cont.

Now define $c_i = k_i + d_i$ where the $k_i$ are the constants we already defined and the $d_i$ are arbitrary constants. Let’s look at the variance of the estimator

$$
\sigma^2 \{ \hat{\beta}_1 \} = \sum c_i^2 \sigma^2 \{ Y_i \} = \sigma^2 \sum (k_i + d_i)^2 \\
= \sigma^2 (\sum k_i^2 + \sum d_i^2 + 2 \sum k_i d_i)
$$

Note we just demonstrated that

$$\sigma^2 \sum k_i^2 = \sigma^2 \{ b_1 \}$$

So $\sigma^2 \{ \hat{\beta}_1 \}$ is related to $\sigma^2 \{ b_1 \}$ plus some extra stuff.
Proof cont.

Now by showing that \( \sum k_i d_i = 0 \) we’re almost done

\[
\sum k_i d_i = \sum k_i (c_i - k_i)
\]

\[
= \sum k_i (c_i - k_i)
\]

\[
= \sum k_i c_i - \sum k_i^2
\]

\[
= \sum c_i \left( \frac{X_i - \bar{X}}{\sum (X_i - \bar{X})^2} \right) - \frac{1}{\sum (X_i - \bar{X})^2}
\]

\[
= \frac{\sum c_i X_i - \bar{X} \sum c_i}{\sum (X_i - \bar{X})^2} - \frac{1}{\sum (X_i - \bar{X})^2} = 0
\]
Proof end

So we are left with

\[ \sigma^2 \{ \hat{\beta}_1 \} = \sigma^2 ( \sum k_i^2 + \sum d_i^2 ) = \sigma^2 (b_1) + \sigma^2 (\sum d_i^2) \]

which is minimized when the \( d_i = 0 \ \forall \ i \).

If \( d_i = 0 \) then \( c_i = k_i \).

This means that the least squares estimator \( b_1 \) has minimum variance among all unbiased linear estimators.