Gauss Markov Proof

Gauss Markov: Given the assumptions of the classical linear regression model, the least squares estimators, in the class of unbiased linear estimators, have minimum variance, that is, they are BLUE.

Proof:

Linearity:

$$\hat{\beta}_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

Let's multiply out the Numerator:

$$\sum (X_i - \bar{X})(Y_i - \bar{Y}) = \sum (X_i - \bar{X}) Y_i - \bar{Y} \sum (X_i - \bar{X}) = \sum (X_i - \bar{X}) Y_i$$

$$\hat{\beta}_1 = \frac{\sum (X_i - \bar{X}) Y_i}{\sum (X_i - \bar{X})^2} \qquad \text{let} \quad c_i = \frac{(X_i - \bar{X})}{\sum (X_i - \bar{X})^2}$$

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

$\hat{\beta}_1$ is a weighted sum of the Y's. As $\cancel{\beta_1}$ sums are linear, $\hat{\beta}_1$ is a linear estimator.

Unbiased

To be unbiased $E[\hat{\beta}_1] = \beta_1$

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

$$= \sum c_i (B_0 + B_1 X_i + \varepsilon_i)$$

$$= B_0 \sum c_i + B_1 \sum c_i X_i + \sum c_i \varepsilon_i$$

$$\sum c_i = \frac{\sum (x_i - \bar{x})}{\sum (x_i - \bar{x})^2} = 0$$

$$\frac{\sum (x_i - \bar{x}) \varepsilon_i}{\sum (x_i - \bar{x})^2} = $$

$$= \frac{\sum x_i \varepsilon_i}{\sum (x_i - \bar{x})^2} - \frac{\bar{x} \sum \varepsilon_i}{\sum (x_i - \bar{x})^2}$$

$= 0$ by 5 $\quad$ 0 by 2

$$= B_1 \sum c_i X_i$$

$$= B_1 \sum c_i X_i + B_1 \sum c_i \bar{X} \longleftarrow \text{add 0 b/c } \sum c_i = 0$$

$$\hat{\beta}_1 = B_1 \sum c_i (X_i - \bar{X})$$

$$E[\hat{\beta}_1] = E\left[B_1 \frac{\sum (x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right]$$

$$E[\hat{\beta}_1] = B_1$$

Efficiency

Proof by contradiction:

Imagine there are some weights (a_i) that are more efficient than the OLS weights c_i. Define d_i as $d_i = a_i - c_i$

Let our imaginary estimator be called $\tilde{B}_1$

$$\tilde{B}_1 = \sum a_i Y_i = \sum (c_i + d_i) Y_i$$

$$E[\tilde{B}_1] = E\left[\sum (c_i + d_i) Y_i\right] = \underbrace{E\left[\sum c_i Y_i\right]}_{B_1} + E\left[\sum d_i Y_i\right]$$

$$= B_1 + E\left[\sum d_i (B_0 + B_1 X_i + \varepsilon_i)\right] = B_1 + \sum d_i (B_0 + B_1 X_i)$$

$$E[\tilde{B}_1] = B_1 + B_0 \sum d_i + B_1 \sum d_i X_i$$

for $\tilde{B}_1$ to be unbiased $\sum d_i = 0$ & $\sum d_i X_i = 0$

$$Var(\tilde{B}_1) = Var\left[\sum a_i Y_i\right] = Var\left[\sum (c_i + d_i) Y_i\right] =$$

Rule

Let $Z = \sum b_i X_i$ b_i are constants (weights)

Then $Var(Z) = \sum b_i^2 Var(X) \quad 2 \sum_{i \neq j} \sum b_i b_j cov(X_i, X_j)$

Proof: (put website)

$Z = \sum b_i X_i$ $\qquad$ b_i are weights

$$Var(Z) = E\left[\sum b_i X_i - E\left[\sum b_i X_i\right]\right]^2$$

$$= E\left[\sum b_i X_i - \sum b_i E\left[X_i\right]\right]^2$$

$$= E\left[b_1 X_1 + b_2 X_2 + \cdots + b_n X_n - b_1 E\left[X_1\right] - b_2 E\left[X_2\right] - b_3 E\left[X_3\right] - \cdots - b_n E\left[X_n\right]\right]^2$$

$$= E\left[b_1\left(X_1 - E\left[X_1\right]\right) + b_2\left(X_2 - E\left[X_2\right]\right) + \cdots + b_n\left(X_n - E\left[X_n\right]\right)\right]^2$$

$$= E\left[\sum_i b_i^2\left(X_i - E\left[X_i\right]\right)^2 + 2\sum_{i \neq j}\sum b_i b_j \left(X_i - E\left[X_i\right]\right)\left(X_j - E\left[X_j\right]\right)\right]$$

$$= \sum_i b_i^2 E\left(X_i - E\left[X_i\right]\right)^2 + 2\sum_{i \neq j}\sum b_i b_j E\left[\left(X_i - E\left[X_i\right]\right)\left(X_j - E\left[X_j\right]\right)\right]$$

$$= \sum_i b_i^2 Var(X) + 2\sum_{i \neq j}\sum b_i b_j Cov\left(X_i, X_j\right)$$

apply rule $\quad Var(z) = \sum b_i^2 Var(x) + 2\sum\sum b_i b_j Cov(X_i, X_j)$

to $\quad Var(\tilde{B}_1)$

$$Var(\tilde{B}_1) = \sum (c_i + d_i)^2 Var(Y_i) + 2\sum\sum_{i \neq j} (c_i + d_i)(c_j + d_j) Cov(Y_i, Y_j)$$

$Var(Y_i) = \sigma^2$ by assumption 3

$Cov(Y_i, Y_j) = 0$ by assumption 4

$$Var(\tilde{B}_1) = \sigma^2 \sum (c_i + d_i)^2$$

$$= \sigma^2 \sum c_i^2 + \sigma^2 \sum d_i^2 + 2\sigma^2 \sum c_i d_i$$

$$\frac{2\sigma^2}{\sum (x_i - \bar{x})^2} \left[\sum (x_i - \bar{x}) d_i \right]$$

$$= \frac{2\sigma^2}{\sum (x_i - \bar{x})^2} \left[\sum x_i d_i - \bar{x} \sum d_i \right]$$

$\sum d_i = 0$ & $\sum x_i d_i = 0$

$$Var(\tilde{B}_1) = \sum c_i^2 + \sigma^2 \sum d_i^2$$

This is minimized only when each $d_i = 0$