GAUSS-MARKOV THEOREM

WITHIN A CLASS OF LINEAR UNBIASED ESTIMATORS THE LEAST SQUARES ESTIMATORS HAVE THE MINIMUM VARIANCE.

THIS THEOREM ESTABLISHES THAT

(A) OLS ESTIMATORS ARE LINEAR

(B) OLS ESTIMATORS ARE UNBIASED

(C) THERE IS NO OTHER LINEAR UNBIASED ESTIMATOR THAT HAS LOWER VARIANCE THAN THE OLS ESTIMATORS.

UNBIASEDNESS

CONSIDER SOME POPULATION PARAMETER θ & SOME SAMPLE STATISTIC $\tilde{\theta}$ WHICH IS AN ESTIMATOR OF θ. NOTE THAT θ IS A CONSTANT BUT $\tilde{\theta}$ TAKES DIFFERENT VALUES IN DIFFERENT SAMPLES. THE DISTRIBUTION OF $\tilde{\theta}$ ACROSS ALL POSSIBLE SAMPLES OF A GIVEN SIZE IS CALLED THE SAMPLING DISTRIBUTION OF THE ESTIMATOR $\tilde{\theta}$. IT IS AN <u>UNBIASED ESTIMATOR</u> IF THE SAMPLING DISTRIBUTION HAS A MEAN EQUAL TO THE PARAMETER θ.

<u>EXAMPLE</u> CONSIDER THE POPULATION MEAN (μ) AND THE SAMPLE MEAN $\bar{x} = \frac{1}{n}\sum x_i$

THEN $E(\bar{x}) = E\left[\frac{1}{n}\sum x_i\right] = \frac{1}{n}\sum E(x_i)$

BUT $E(x_i) = \mu$. HENCE $E(\bar{x}) = \mu$

& THE SAMPLE MEAN IS AN UNBIASED ESTIMATOR OF THE POPULATION MEAN.

(SAMPLING) VARIANCE OF AN ESTIMATOR IS THE VARIANCE OF THE SAMPLING DISTRIBUTION.

EXAMPLE: $\text{VAR}(\bar{x}) = \text{Var}\left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right)$

$$= \sum \text{Var}\left(\frac{x_i}{n}\right) = \sum \frac{\sigma^2}{n^2} = \sigma^2/n.$$

WHY IS A SMALL VARIANCE DESIRABLE?

CHEBYCHEV'S INEQUALITY SHOWS THAT FOR ANY RANDOM VARIABLE x WITH MEAN μ & VARIANCE σ^2

$$\Pr\left[\mu - r\sigma \le x \le \mu + r\sigma\right] \ge 1 - \frac{1}{r^2}$$

THUS A SMALLER VALUE OF σ^2 (& HENCE σ) IMPLIES THAT A RANDOMLY DRAWN VALUE OF x LIES WITHIN A SMALLER INTERVAL μ EXCEEDS $1 - \frac{1}{r^2}$.

FOR A KNOWN DISTRIBUTION, WE CAN BE MORE PRECISE.

IF $x \sim N(\mu, \sigma^2)$,

$$\Pr\left[\mu - 1.96\sigma \le x \le \mu + 1.96\sigma\right] = 0.95$$

SUPPOSE $\mu = 10$ & $\sigma^2 = 4$

THEN $\Pr\left[6.08 \le x \le 13.92\right] = 0.95$

BUT IS $\sigma^2 = 1$

$$\Pr\left[8.04 \le x \le 11.96\right] = 0.95$$

THUS A SMALLER VARIANCE IMPLIES THAT THE DISTRIBUTION IS MORE CONCENTRATED AROUND THE MEAN.

PROOF OF THE GAUSS-MARKOV THEOREM

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

DEFINE $A \equiv (X'X)^{-1}X'$. THEN $\hat{\beta} = Ay$

$$A = \begin{bmatrix} a_{11} & a_{12} \cdots & a_{1n} \\ a_{21} & a_{22} & a_{2n} \\ a_{k1} & a_{k2} & a_{kn} \end{bmatrix} \qquad y = \begin{bmatrix} y_1 \\ y_2 \\ y_n \end{bmatrix}$$

(A) THEN $\hat{\beta} = Ay \Rightarrow \hat{\beta}_i = \sum_j a_{ij} y_j \quad (i = 1, 2, \ldots, k)$

THIS SHOWS THAT $\hat{\beta}$ IS LINEAR IN y.

(B)

$$y = X\beta + u \Rightarrow \hat{\beta} = (X'X)^{-1}X'y$$

$$\hat{\beta} = (X'X)^{-1}X'(X\beta + u)$$

$$\hat{\beta} = (X'X)^{-1}X'X\beta + (X'X)^{-1}X'u = \beta + (X'X)^{-1}X'u$$

$$E(\hat{\beta}) = \beta + E\left((X'X)^{-1}X'u\right) = \beta + (X'X)^{-1}X'E(u)$$

$$= \beta \quad (\text{b/c } E(u) = 0)$$

THIS SHOWS UNBIASEDNESS.

(C)

$$V(\hat{\beta}) = E\left[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'\right]$$

$$= E\left[(X'X)^{-1}X'uu'X(X'X)^{-1}\right]$$

$$= (X'X)^{-1}X'E(uu')X(X'X)^{-1}$$

$$= (X'X)^{-1}X'(\sigma^2 I)X(X'X)^{-1} = \sigma^2 (X'X)^{-1}(X'X)(X'X)^{-1}$$

$$= \sigma^2 (X'X)^{-1}.$$

CONSIDER SOME OTHER LINEAR & UNBIASED ESTIMATOR $\tilde{\beta}$.

BECAUSE $\tilde{\beta}$ IS LINEAR IN y IT CAN BE WRITTEN AS $\tilde{\beta} = Cy$ (4-4)

THEN $\tilde{\beta} = C(X\beta + u) = CX\beta + Cu$

$$E(\tilde{\beta}) = CX\beta + CE(u) = CX\beta.$$

BUT $\tilde{\beta}$ IS UNBIASED. HENCE

$E(\tilde{\beta}) = \beta \Rightarrow CX = I$.

RECALL THAT $\hat{\beta} = (X'X)^{-1}X'y \equiv Ay$

DEFINE $D = C - A$.

NOW, $\tilde{\beta} = \beta + Cu$

$\Rightarrow \tilde{\beta} - \beta = Cu$

$V(\tilde{\beta}) = E[Cuu'C'] = C(\sigma^2 I)C' = \sigma^2 CC'$.

BUT $C = A + D$

$$= D + (X'X)^{-1}X'$$

$$C' = D' + X(X'X)^{-1}$$

$$CC' = [D + (X'X)^{-1}X'][D' + X(X'X)^{-1}]$$

$$= DD' + (X'X)^{-1}X'D' + DX(X'X)^{-1} + (X'X)^{-1}X'X(X'X)^{-1}$$

NOW, $D = C - (X'X)^{-1}X'$

$$DX = CX - (X'X)^{-1}X'X = I - I = \underline{0}$$

$\Rightarrow X'D' = \underline{0}'$

HENCE $CC' = DD' + (X'X)^{-1}$.

$$V(\tilde{\beta}) - V(\hat{\beta})$$
$$= \sigma^2 DD'$$

IS A POSITIVE DEFINITE MATRIX.

$$V(\tilde{\beta}) = V(\hat{\beta}) + \sigma^2 DD'$$