SMIII lectures

Week 2

Estimable functions and best linear unbiased estimates

Consider the linear model formulation

$$\boldsymbol{Y} = X\beta + \boldsymbol{E}$$

- Theorems 1.2 and 1.3 (last lecture) describe the sampling distribution of the least squares estimates.
- The least squares estimate $\hat{\boldsymbol{\eta}} = X\hat{\beta}$ is also the best linear unbiased estimate (BLUE) of $\boldsymbol{\eta} = E[\boldsymbol{Y}]$.

Definition 1.4

A quantity of the form $\boldsymbol{\lambda}^T \boldsymbol{\eta}$ where $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)^T$ is a fixed vector is said to be an **estimable function**.

$\square$

Examples

Consider the multiple regression model $\boldsymbol{Y} = X\boldsymbol{\beta} + \boldsymbol{\mathcal{E}}$ where the columns of X are assumed to be linearly independent.

- Each individual regression coefficient β_j is an estimable function. To see this, take

$$\boldsymbol{\lambda}^T = (0, \ldots, 0, 1, 0, \ldots, 0)(X^T X)^{-1} X^T.$$

- The point prediction, $\beta_0 + \beta_1 x_1 + \ldots + \beta_r x_r$ is an estimable function for any fixed values $x_1, x_2, \ldots, x_r$. To see this, take

$$\boldsymbol{\lambda}^T = (1, x_1, x_2, \ldots, x_r)(X^T X)^{-1} X^T.$$

The Gauss-Markov Theorem

Theorem 1.4

Suppose $E(\boldsymbol{Y}) = \boldsymbol{\eta} = X\beta$ and $\text{Var}(\boldsymbol{Y}) = \sigma^2 I$.

If $\boldsymbol{a}^T\boldsymbol{Y}$ is an unbiased linear estimator for $\boldsymbol{\lambda}^T\boldsymbol{\eta}$, then

$$\text{var}(\boldsymbol{a}^T\boldsymbol{Y}) \geq \text{var}(\boldsymbol{\lambda}^T\hat{\boldsymbol{\eta}})$$

with equality if and only if

$$\boldsymbol{a} = X(X^TX)^{-1}X^T\boldsymbol{\lambda}.$$

$\square$

Inference for multiple regression: regression coefficients

- Theorem 1.3 (last lecture) allows for inference concerning individual regression coefficients.
- Suppose the multiple regression model

$$M : \boldsymbol{Y} = X\beta + \boldsymbol{\mathcal{E}}$$

with

$$\boldsymbol{\mathcal{E}} \sim N_n(\mathbf{0}, \sigma^2 I)$$

holds.

Pivotal quantity

A **pivotal quantity** for β_j is given by

$$\frac{\hat{\beta}_j - \beta_j}{s_e\sqrt{c_{jj}}} \sim t_{n-p}$$

where c_{jj} is the jth diagonal element of $(X^TX)^{-1}$.

Confidence interval for β_j

The random interval

$$\hat{\beta}_j \pm t_{n-p}(\alpha/2) s_e \sqrt{c_{jj}}$$

is a $100(1-\alpha)\%$ confidence interval for β_j.

Hypothesis testing of β_j

To test $H_0 : \beta_j = 0$ we use the test statistic

$$t = \frac{\hat{\beta}_j - 0}{s_e \sqrt{c_{jj}}}$$

and the rule,

$$\text{Reject } H_0 \text{ for } |t| \geq t_{n-p}(\alpha/2),$$

to define a two-sided test with significance level α.

The P-value can be calculated in the usual way.

Testing multiple coefficients

Hypotheses of the form

$$H_0 : \beta_p = \beta_{p-1} = \ldots = \beta_{p_0+1} = 0$$

may be tested using an F-test. The ANOVA table is

Source	SS	DF	MS	F
H_0 vs M	$Q(\hat{\beta}_0) - Q(\hat{\beta})$	$p - p_0$	$MS_H = \frac{Q(\hat{\beta}_0) - Q(\hat{\beta})}{p - p_0}$	$F = \frac{MS_H}{MS_E}$
Error	$Q(\hat{\beta})$	$n - p$	$MS_E = \frac{Q(\hat{\beta})}{n-p}$	
Total	$Q(\hat{\beta}_0)$	$n - p_0$		

- If H_0 is true then F has the F-distribution, $F \sim F_{p-p_0,n-p}$ so we reject H_0, with significance level α, if $F \geq F_{p-p_0,n-p}(\alpha)$.

Prediction in multiple regression

Consider the multiple regression model

$$\boldsymbol{Y} = X\beta + \boldsymbol{\mathcal{E}} \text{ with } \boldsymbol{\mathcal{E}} \sim N_n(\mathbf{0}, \sigma^2 I).$$

- Suppose we have a new observation $(Y_0, \boldsymbol{x}_0)$
- with $Y_0 \sim N(\boldsymbol{x}_0^T \beta, \sigma^2)$ (independently from $\boldsymbol{Y}$).

A $100(1-\alpha)\%$ **confidence interval** for $\eta_0 = \boldsymbol{x}_0^T \boldsymbol{\beta}$

The $100(1-\alpha)\%$ confidence interval for η_0 is therefore

$$\hat{\eta}_0 \pm t_{n-p}(\alpha/2) s_e \sqrt{\boldsymbol{x}_0^T (X^T X)^{-1} \boldsymbol{x}_0}$$

A $100(1-\alpha)\%$ **prediction interval** for Y_0

The $100(1-\alpha)\%$ prediction interval for Y_0 is therefore

$$\hat{\eta}_0 \pm t_{n-p}(\alpha/2) s_e \sqrt{1 + \boldsymbol{x}_0^T (X^T X)^{-1} \boldsymbol{x}_0}$$

Case study

Rubber example