Slide 77

$$RSS = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

$$\begin{cases} \dfrac{\partial RSS}{\partial \beta_0} = -2\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = -2\left[\sum y_i - n\beta_0 - \beta_1 \sum x_i\right] \\ \dfrac{\partial RSS}{\partial \beta_1} = \sum_{i=1}^{n} -2x_i (y_i - \beta_0 - \beta_1 x_i) = -2\left[\sum_{i=1} x_i y_i - \beta_0 \sum x_i - \beta_1 \sum x_i^2\right] \end{cases}$$

$$\left.\begin{array}{l} \dfrac{\partial RSS}{\partial \beta_0} = 0 \Rightarrow \overset{\sum x_i \times}{\sum y_i - n\beta_0 - \beta_1 \sum x_i = 0} \\ \dfrac{\partial RSS}{\partial \beta_1} = 0 \Rightarrow \overset{n \times}{\sum x_i y_i - \beta_0 \sum x_i - \beta_1 \sum x_i^2 = 0} \end{array}\right\} \text{normal equation}$$

$$\sum x_i \sum y_i - n\beta_0 \sum x_i - \beta_1 \left(\sum x_i\right)^2 = 0$$

$$-1 \times \quad \underset{-}{n\sum x_i y_i} \underset{+}{-\, n\beta_0 \sum x_i} \underset{+}{-\, \beta_1 n \sum x_i^2} = 0$$

$$\sum x_i \sum y_i - n\sum x_i y_i - \beta_1 \left(\sum x_i\right)^2 - \beta_1 n \sum x_i^2 = 0$$

$$\hat{\beta}_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n(\bar{x})^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

unhaisedness of $\hat{\beta}_0$ and $\hat{\beta}_1$

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

$$= \frac{1}{\sum (x_i - \bar{x})^2} \sum (x_i - \bar{x}) \, E(Y_i - \bar{Y})$$

$$= \frac{1}{\sum (x_i - \bar{x})^2} \sum (x_i - \bar{x}) \left[E(Y_i) - E(\bar{Y}) \right]$$

$$= \frac{1}{\sum (x_i - \bar{x})^2} \sum (x_i - \bar{x}) \left[\cancel{\beta_0} + \beta_1 x_i - (\cancel{\beta_0} + \beta_1 \bar{x}) \right]$$

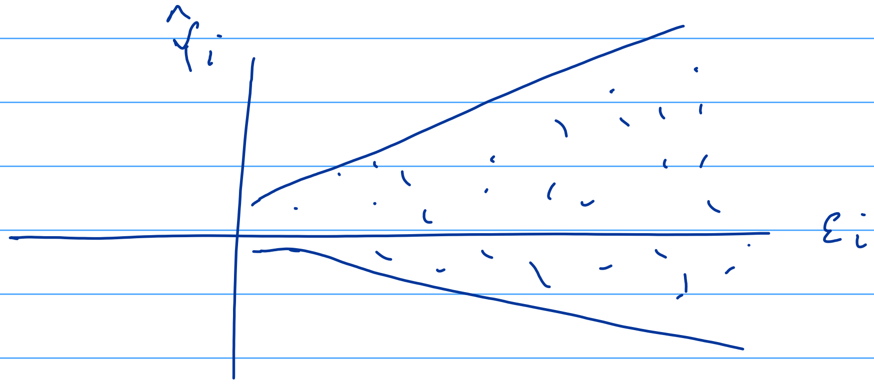
$$\beta_1 (x_i - \bar{x})$$

$$= \beta_1$$

$$\begin{cases} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{cases}$$

$$t = \frac{\hat{\beta}_1 - 0}{SE(\hat{\beta}_1)} \sim t_{n-2}$$

F-statistic for SLR $\equiv t^2$



Slide 91

- Y_i belongs to exponential family
  - distribution is in canonical form $a(Y_i) = Y_i$
  - natural parameter $b(\theta_i$

$$f_{Y_i}(y_i) = \exp\left(y_i b(\theta_i) + c(\theta_i) + d(y_i) \right)$$

we know that

$$E(Y_i) = \mu_i = -\frac{c'(\theta_i)}{b'(\theta_i)} \quad (*)$$

$$Var(Y_i) = \frac{b''(\theta_i) c'(\theta_i) - c''(\theta_i) b'(\theta_i)}{(b'(\theta_i))^3} \quad (**)$$

$$g(\mu_i) = \eta_i = x_i^T(\beta)$$

To find MLE we should minmize $l(\theta)$

$$l(\theta; Y_1, \ldots, Y_N) = \sum_{i=1}^{N} l_i$$

$$= \sum_{i=1}^{N} Y_i b(\theta_i) + \sum_{i=1}^{N} c(\theta_i) + \sum_{i=1}^{N} d(Y_i)$$

score function

$$U_j = \frac{\partial l}{\partial \beta_j} = \underbrace{\frac{\partial l}{\partial \theta_i}}_{(1)} \times \underbrace{\frac{\partial \theta_i}{\partial \mu_i}}_{(2)} \times \underbrace{\frac{\partial \mu_i}{\partial \beta_j}}_{(3)}$$

use chain rule

① $$\frac{\partial l}{\partial \theta_i} = \sum_{i=1}^{N} Y_i b'(\theta_i) + \sum_{i=1}^{N} c'(\theta_i)$$

② $$\frac{\partial \theta_i}{\partial \mu_i} = \left(\frac{\partial \mu_i}{\partial \theta_i}\right)^{-1} \overset{*}{=} \left[\frac{\partial}{\partial \theta_i}\left(\frac{-c'(\theta_i)}{b'(\theta_i)}\right)\right]^{-1}$$

$$= \left[-\frac{c''(\theta_i)}{b'(\theta_i)} + \frac{c'(\theta_i)\, b''(\theta_i)}{(b'(\theta_i))^2}\right]^{-1}$$

$$\overset{**}{=} \left[b'(\theta_i)\, Var(Y_i)\right]^{-1}$$

$$\frac{\partial \mu_i}{\partial \beta_j} = \frac{\partial \mu_i}{\partial \eta_i} \times \underbrace{\frac{\partial \eta_i}{\partial \beta_j}}_{} = \frac{\partial \mu_i}{\partial \eta_i} x_{ij}$$

$$\eta_i = x_i^T \beta = x_{i1}\beta_1 + x_{i2}\beta_2 + \cdots + x_{ip}\beta_p$$

$$\longrightarrow \quad U_j = \sum_{i=1}^{N}\left[\frac{(y_i - \mu_i)}{Var(Y_i)} x_{ij}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)\right]$$

$$I_{jk} = E(U_j U_k)$$

$$= E\left[\sum_{i=1}^{N} \frac{Y_i - \mu_i}{Var(Y_i)} X_{ij}\left(\frac{\partial \mu_i}{\partial \eta_i}\right) \times \sum_{i'=1}^{N} \frac{Y_{i'} - \mu_{i'}}{Var(Y_{i'})} X_{i'k}\left(\frac{\partial \mu_{i'}}{\partial \eta_{i'}}\right)\right]$$

Y_i are indep $\Rightarrow 0$ except $i = i'$

$$= \sum_i \sum_{i'} \frac{d\mu_i}{d\eta_i} \frac{\partial \mu_{i'}}{d\eta_{i'}} \frac{x_{ij} x_{i'k}}{Var(Y_i)\, Var(Y_{i'})} \underbrace{E(Y_i - \mu_i)(Y_{i'} - \mu_{i'})}_{\text{if } i = i' \quad Var(Y_i)}$$

$$= \sum_i \frac{x_{ij} x_{ik}}{Var(Y_i)} \left(\frac{d\mu_i}{d\eta_i} \right)^2$$

---

in normal dist since $\frac{d\mu_i}{d\eta_i} = 1$

$$I_{jk} = \sum_{i=1}^{N} \frac{x_{ij} x_{ik}}{Var\, Y_i}$$

Rules:

$$\frac{\partial x^T A}{\partial x} = A$$

$$\frac{\partial A x}{\partial x} = A^T$$

$$\frac{\partial x^T A x}{\partial x} = (A + A^T) x$$

Therefore if A is symmetric then $A = A^T$

$$\Rightarrow \frac{\partial x^T A x}{\partial x} = 2 A x$$

$$Y \sim N(X\beta, \sigma^2 I)$$

$$(I-H)Y \sim N($$

$$E((I-H)Y) = (I-H)X\beta$$

$$= X\beta - HX\beta$$

$$= X\beta - X(\underbrace{(X^TX)^{-1}X^TX}_{I}\beta)$$

$$= X\beta - X\beta = 0$$

$$Var((I-H)Y) = (I-H)\overbrace{Var(Y)}^{\sigma^2 I}(I-H)$$

$$= \frac{1}{\sigma^2}(I-H)(I-H) = \frac{1}{\sigma^2}(I-H)$$

when H is a projection matrix the $I-H$ is projection matrix $\Rightarrow (I-H)^2 = (I-H)$

* if $X \sim N(\mu, \sigma^2 I)$ then

$$\frac{x^TAX}{\sigma^2} \sim \chi^2(\text{trace}(A), \mu^TA\mu)$$

minimal model $\qquad Y_i = \beta_0 + \varepsilon_i$

$$X = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \qquad y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \qquad \beta = \beta_0$$

$$\widehat{RSS} = \sum (y_i - \bar{y})^2 = TSS$$

$$\hat{\beta} = (\underbrace{X^T X}_{N})^{-1} \underbrace{X^T y}_{\sum y_i} = \frac{1}{N} \sum y_i = \bar{y}$$

maximal model

μ_i for each observation $Y_i \Rightarrow \beta$ has N elements

$$X_{N \times N} = I_{N \times N}$$

$$\Rightarrow \hat{\beta} = (X^T X)^{-1} X^T Y = Y$$

$$\Rightarrow \hat{\mu}_i = y_i$$

$$RSS = \sum (y_i - \hat{y}_i)^2 = \sum 0 = 0$$