多元

$$Y_i = \beta_1 + \beta_2 X_i + \mu_i$$

$$\Rightarrow Y_i = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \mu_i$$

一般表达：

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_2 X_{3i} + \cdots \beta_k X_{ki} + \mu_i$$

$$E(Y|X_{2i}, X_{3i}, \cdots, X_{ki}) = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_k X_{ki}$$

$$Y_i = \hat{Y}_i + e_i$$

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \cdots + \hat{\beta}_k X_{ki} + e_i$$

① 矩阵形式：

$$\begin{cases} Y = X\beta + U \\ E(Y) = X\beta \end{cases}$$ 总体

$$\begin{cases} Y = X\hat{\beta} + e \\ \hat{Y} = X\hat{\beta} \end{cases}$$ 样本

复习：古典假定、基本假定

1. 零均值

$$E(u_i)=0$$

$$E(U)=0$$

2. 无自相关

$$Cov(u_i,u_k)=E[u_i-E(u_i)]-E[u_k-E(u_k)]$$
$$=E(u_iu_k)=\begin{cases}\sigma^2 & i=k\\ 0 & i\neq k\end{cases}$$

3. 同方差

$$Var(U)=E[(U-EU)(U-EU)']$$
$$=E(UU')=\sigma^2 I_n.$$

4. 随机干扰项与解释变量不相关

$$Cov(X_{ji},u_i)=0.$$

5. 无多重共线性

$Rank(X)=k$

$Rank(X'X)=k$

$(X'X)$存在

6. 正态性

$u_i \sim N(0, \sigma^2)$

最小二乘估计

正规方程组：

$$X'Y = X'X\hat{\beta}$$

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

1) 线性性质

$\hat{\beta}$ 与 $(X'X)^{-1}X'$ 线性

2) 无偏性

$$\begin{aligned}\hat{\beta} = (X'X)^{-1}X'Y &= (X'X)^{-1}X'(X\beta+U) \\ &= (X'X)^{-1}X'X\beta + (X'X)^{-1}X'U \\ &= \beta + (X'X)^{-1}X'U\end{aligned}$$

$$E(\hat{\beta}) = \beta + (X'X)^{-1}X'\underbrace{[E(U)]}_{=0} = \beta$$

3) 最小方差性

L 最佳线性无偏估计量

标准误差

$$SE(\hat{\beta}_j) = \sigma\sqrt{c_{jj}}$$

$$\hat{\beta}_j \sim N[\beta_j, \sigma^2 c_{jj}]$$

$$e = Y - \hat{Y} = Y - X\hat{\beta}$$

$$\sum e_i^2 = e'e$$

无偏估计

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-k}$$

参数估计量方差

$$\widehat{Var}(\hat{\beta}_2) = \hat{\sigma}^2 c_{jj} = \left(\frac{\sum e_i^2}{n-k}\right) c_{jj}$$

拟合优度检验

$$TSS = ESS + RSS$$

$$\sum(Y_i - \bar{Y})^2 = \sum(\hat{Y}_i - \bar{Y})^2 + \sum(Y_i - \hat{Y})^2$$

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum e_i^2}{\sum(Y_i - \bar{Y})^2}$$

多重

$$TSS = \sum(Y_i - \bar{Y})^2 = \sum Y_i^2 - n\bar{Y}^2 = Y'Y - n\bar{Y}^2$$

$$ESS = \hat{\beta}'X'Y - n\bar{Y}^2$$

$$\therefore R^2 = \frac{ESS}{TSS} = \frac{\hat{\beta}'X'Y - n\bar{Y}^2}{Y'Y - n\bar{Y}^2}$$

修正的可决系数

$$\bar{R}^2 = 1 - \frac{\sum e_i^2/(n-k)}{\sum(Y_i-\bar{Y})^2/n-1} = 1 - \frac{n-1}{n-k} \cdot \frac{\sum e_i^2}{\sum(Y_i-\bar{Y})^2}$$

$$\bar{R}^2 = 1 - (1-R^2)\frac{n-1}{n-k}$$

F检验：

1° $H_0: \beta_2 = \beta_3 = \cdots = \beta_k = 0$

$H_1: \beta_j\ (j=2,3,\cdots,k)$ 不全为0.

2° $$F = \frac{ESS/(k-1)}{RSS/(n-k)} \sim F(k-1, n-k)$$

3° 找临界值 $F_\alpha(k-1, n-k)$ 与F比较.

4° $F > F_\alpha$，拒绝原假设

$F < F_\alpha$，不能拒，不显著

F与可决系数的关系

$$F = \frac{n-k}{k-1} \cdot \frac{R^2}{1-R^2}$$ $R^2 = 1$，$F \to +\infty$

t检验

$$t = \frac{\hat{\beta}_j - \beta}{\sqrt{\hat{\sigma}^2 c_{jj}}} = \frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}\sqrt{c_{jj}}} \sim t(n-k)$$

$$t^* = \frac{\hat{\beta}_j}{\hat{\sigma}\sqrt{c_{jj}}}$$

然后找 $t_{\frac{\alpha}{2}}(n-k)$

1° $|t^*| \geq t_{\frac{\alpha}{2}}(n-k)$，拒绝 H_0

2° $|t^*| < t_{\frac{\alpha}{2}}(n-k)$，不能拒绝 H_0.

点预测

$$Y = \beta X + U$$

$$\hat{Y} = \hat{\beta} X$$

$$\hat{Y}_f = \hat{\beta}_1 + \hat{\beta}_2 X_{2f} + \hat{\beta}_3 X_{3f} + \cdots + \hat{\beta}_k X_{kf}$$

$$E(\hat{Y}_f) = E(Y_f)$$

案例分析

3.5 ①

TSS的自由度 $= n-1 = 19 \quad n=20$

$ESS = 3-1 = 2$

$RSS = n-k = 20-3 = 17.$

②.

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

$$= \frac{\Sigma(\hat{Y}_i - \bar{Y})^2}{\Sigma(Y_i - \bar{Y})^2} = 1 - \frac{\Sigma(e_i^2)}{\Sigma(Y_i - \bar{Y})^2} = 0.8417.$$

$$\Sigma(Y_i - \bar{Y})^2 = \Sigma(Y_i - \hat{Y}_i)^2 - \Sigma(\hat{Y} - \bar{Y})^2$$

$$= 447962.19$$

$$\bar{R}^2 = 1 - \frac{n-1}{n-k} \frac{\Sigma(e_i^2)}{\Sigma(Y_i - \bar{Y})^2} = 1 - \frac{19}{17}(1-R^2) = 0.823$$

③.

$$F = \frac{ESS/(k-1)}{RSS/(n-k)} = \frac{377067.19/2}{470895.00/17}.$$

3.4.

A:

$$\bar{R}^2 = 1-(1-R)\frac{n-1}{n-k} = 1-(1-0.994202)\times\frac{38}{35} = \bigcirc.$$

$$\sqrt{\frac{RSS}{(n-k)}}$$

$$F = \frac{n-k}{n-1} \cdot \frac{R^2}{1-R^2}$$

3.1

看 R^2 0.6，$\bar{R}^2$ 0.6 拟合数据

$F = 17.95 > F(3, 27)$