Multiple linear regression

$(k-1)$ regressor variables

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_{k-1} x_{i,k-1} + \varepsilon_i$$

Assumption $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$

$$Y = X\beta + \varepsilon$$

LSE determine the parameters by minimizing

$$SS_{Res} = \sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2$$

$$= e'e = (Y - \hat{Y})'(Y - \hat{Y}) = (Y - X\hat{\beta})(Y - X\hat{\beta})'$$

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

$$= Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}X'X\hat{\beta}.$$

$$SS_{Res} = \sum (Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \cdots - \hat{\beta}_{k-1} x_{ik-1})^2$$

Normal equations

$$\sum (Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} \cdots - \hat{\beta}_{k-1} x_{ik-1}) = 0$$

$$\sum e_i x_{i1} = 0$$

$$\vdots$$

$$\sum e_i x_{ik-1} = 0$$

$$\frac{\partial SS_{Res}}{\partial \hat{\beta}} = 0 \Rightarrow \quad -2X'Y + 2X'X\hat{\beta} = 0$$

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

Statistical properties of LSE

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

$$V(\hat{\beta}) = (X'X)^{-1}X'\sigma^2 I \; X(X'X)^{-1} = \sigma^2(X'X)^{-1}$$

$$SS_{Res} = Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'(X'X)\hat{\beta}$$

$$= Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'X'X(X'X)^{-1}X'Y$$

$$= Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'X'Y$$

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

SS_{Res} has $(n-k)$ d.f.

$$MS_{Res} = \frac{SS_{Res}}{n-k}, \quad E(MS_{Res}) = \sigma^2$$

$$\frac{SS_{Res}}{\sigma^2} \sim \chi^2_{n-k}$$

other way to express SS_{Res}

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

$$= Y - X(X'X)^{-1}X'Y = (I - X(X'X)^{-1}X')Y$$

$$= (I - H)Y = Y - HY.$$

H is $n \times n$ and is called hat matrix.

$$SS_{Res} = e'e = Y'(I-H)'(I-H)Y$$

$$= Y'(I-H)Y$$

$$SS_T = \Sigma(Y_i - \bar{Y})^2 = \Sigma Y_i^2 - n\bar{Y}^2 \quad \text{has d.f. } n-1$$

$$SS_{Reg} = SS_T - SS_{Res}$$

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

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

$$SS_T = SS_{Reg} + SS_{Res}$$

$n-1$ $\quad$ $n-K$ $\quad$ $K-1$

Test for Significance of Reg model

If there is a linear relationship between the response and any of the regressor variables $x_1, \ldots x_{k-1}$.

$H_0: \beta_1 = \beta_2 = \cdots = \beta_{k-1} = 0$

vs. $H_1: \beta_j \neq 0$ for atleast one j.

$$\frac{SS_{Reg}}{\sigma^2} \sim \chi^2_{k-1}$$

$$\frac{SS_{Res}}{\sigma^2} \sim \chi^2_{n-k}$$

$$F = \frac{\frac{SS_{Reg}}{k-1}}{\frac{SS_{Res}}{n-k}} \sim F_{k-1, n-k}$$

$E(MS_{Res}) = \sigma^2$; $E(MS_{Reg}) = \sigma^2 + \dfrac{\beta^{*'} X_c' X_c \beta^*}{(k-1)\sigma^2}$

$$\beta^* = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{k-1} \end{pmatrix} \qquad X_c = \begin{pmatrix} x_{11}-\bar{x} & \cdots & x_{1k-1}-\bar{x} \\ \\ x_{n1}-\bar{x} & & x_{nk-1}-\bar{x} \end{pmatrix}$$

$$F = \frac{MS_{Reg}}{MS_{Res}}$$

atleast one $\beta_j \neq 0$

we reject H_0 $\beta_1 = \beta_2 = \cdots = \beta_{K-1} = 0$ if

$$F > F_{\alpha,\, K-1,\, n-K}.$$

ANOVA TABLE

Source of variation	DF	SS	MS	F
Reg	$K-1$	SS_{Reg}	$MS_{Reg} = \frac{SS_{Reg}}{K-1}$	$F = \frac{MS_{Reg}}{MS_{Res}}$
Res	$n-K$	SS_{Res}	$MS_{Res} = \frac{SS_{Res}}{n-K}$	
Total	$n-1$	SS_T	MS_T	

$$F \sim F_{K-1,\, n-K}$$

we reject H_0 if $F > F_{\alpha,\, K-1,\, n-K}$

Test on individual regression Coefficient

(partial / Marginal Test)

$H_0 : \beta_j = 0$ ag.

$H_1 : \beta_j \neq 0$.

Test the significance of x_j in the presence of other regressors in the model.

6

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

$$\hat{\beta} \sim N(\beta, \sigma^2 (X'X)^{-1})$$

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\sigma^2 (X'X)^{-1}_{jj}}} \sim N(0,1)$$

Test statistic

$$t = \frac{\hat{\beta}_j}{\sqrt{MS_{Res} (X'X)^{-1}_{jj}}} \sim t_{n-k} \quad \text{under } H_0.$$

$H_0: \beta_j = 0$ is rejected if

$$|t| < t_{\alpha/2, n-k}$$

Extra Sum of Squares Method

Test for several parameters being zero.

$Y = X\beta + \epsilon$

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon$

$H_0: \begin{pmatrix} \beta_3 \\ \beta_4 \end{pmatrix} = 0$ vs. $H_1: \begin{pmatrix} \beta_3 \\ \beta_4 \end{pmatrix} \neq 0$

In general, $Y = X\beta + \varepsilon$ $\quad \underset{k\times 1}{\beta} = \begin{pmatrix} \beta_1^{\,k-r\times 1} \\ \beta_2^{\,r\times 1} \end{pmatrix}$

$X = (X_1 \vdots X_2)$

$H_0: Y = X_1\beta_1 + \varepsilon$ $\quad \Leftrightarrow \quad$ $H_0: \beta_2 = 0$

$H_1: Y = X_1\beta_1 + X_2\beta_2 + \varepsilon$ $\quad$ $H_1: \beta_2 \neq 0$

We compute SS_{Reg} for both the full and restricted model.

$SS_{Reg}^{Full} = \hat{\beta}' X' Y - n\bar{Y}^2$ has d.f. $k-1$

$SS_{Res}^{Full} = Y'Y - \hat{\beta}' X' Y$ has d.f. $n-k$

$SS_{Reg}^{Restricted} = \hat{\beta}_1' X_1' Y - n\bar{Y}^2$ has d.f. $k-r-1$

$SS_{Reg}^{Full} - SS_{Reg}^{Res.}$ = extra sum of squares due to β_2 given that β_1 is already in the model.

$$\left.\begin{array}{l} \dfrac{SS_{Reg}^{Full} - SS_{Reg}^{Restr}}{\sigma^2} \sim \chi^2_r \\ \dfrac{SS_{Res}^{Full}}{\sigma^2} \sim \chi^2_{\cancel{k}\, n-K} \end{array}\right\} \text{ ind.}$$

$$F = \frac{\left(SS_{Reg}^{Full} - SS_{Reg}^{Restr.}\right)/r}{SS_{Res}^{Full}/n-K} \sim F_{r,\, n-K}$$

under H_0

If $F > F_{\alpha, r, n-K}$, we reject H_0.

atleast one of the regressors in β_2 is significant.

<u>Extra Sum of Squares</u>

$H_0: \beta_2 = 0$ as.

$H_1: \beta_2 \neq 0$

$$F = \frac{\left(SS_{Reg}^{Full} - SS_{Reg}^{Restr.}\right)/1}{\dfrac{SS_{Res}}{n-K}} \sim F_{1,\, n-K}$$

If $F > F_{\alpha, 1, n-K}$, we reject H_0.

$\boxed{t^2 = F}$

Confidence interval on regression coefficients

$$\hat{\beta} = (X'X)^{-1} X'Y \qquad V(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

$$\hat{\beta}_i \sim N\left(\beta_i, \sigma^2 (X'X)^{-1}_{ii}\right)$$

$$\frac{\hat{\beta}_i - \beta_i}{\sqrt{MS_{Res} (X'X)^{-1}_{ii}}} \sim t_{n-K}$$

$$P_r \left\{ \left| \frac{\hat{\beta}_i - \beta_i}{\sqrt{MS_{Res} (X'X)^{-1}_{ii}}} \right| < t_{\alpha/2, n-K} \right\} = 1-\alpha$$

A $100(1-\alpha)\%$ CI for the parameter β_i is

$$\hat{\beta}_i \pm t_{\alpha/2, n-K} \sqrt{MS_{Res} (X'X)^{-1}_{ii}}$$

Confidence Interval on Mean response at a particular point, say, $x_0 = (1, x_{01}, x_{02}, \ldots, x_{0,K-1})$

$$E(Y|x_0) = x_0 \beta$$

An unbiased estimator of $E(Y/x_0)$ is $\hat{y}_0 = x_0 \hat{\beta}$

$$E(x_0 \hat{\beta}) = x_0 E(\hat{\beta}) = x_0 \beta.$$

$$V(\hat{y}_0) = V(x_0 \hat{\beta}) = x_0 \, V(\hat{\beta}) \, x_0'$$
$$= x_0 \, \sigma^2 (X'X)^{-1} x_0'$$

$$\frac{\hat{y}_0 - x_0 \beta}{\sqrt{x_0 \, MS_{Res} \, (X'X)^{-1} x_0'}} \sim t_{n-K}$$

A $100(1-\alpha)\%$ CI on mean response at the point x_0 is

$$x_0 \hat{\beta} \pm t_{\alpha/2, n-K} \sqrt{x_0 \, MS_{Res} \, (X'X)^{-1} x_0'}$$

Prediction of New observation

$$x_0 = (1, x_{01}, \ldots x_{0,K-1})$$

$$y_0 = x_0 \beta + \varepsilon$$

A point estimator of y_0 ~~at x_0~~ is

$$\hat{y}_0 = x_0 \hat{\beta}$$

$$\psi = \hat{y}_0 - y_0 , \quad E(\psi) = 0$$

$$V(\psi) = V(\hat{y}_0 - y_0) = \sigma^2 (1 + x_0 (X'X)^{-1} x_0')$$

$$\frac{\hat{y}_0 - y_0}{\sqrt{MS_{Res} (1 + x_0 (X'X)^{-1} x_0')}} \sim t_{n-K}$$

A (100) (1-α) % PI for y_0 is

$$x\hat{\beta}_0 \pm t_{\alpha/2, n-k} \sqrt{MS_{Res} \left(1 + x_0 (X'X)^{-1} x_0' \right)}$$