Confidence interval & hypothesis testing

24 Sept

For $p = 1$

$$\hat{\beta}_0 \sim N\left(\beta_0, \sigma^2\left(\frac{1}{n} + \frac{\bar{X}_1^2}{S_{X_1X_1}}\right)\right)$$

$$\Rightarrow \boxed{\frac{\dfrac{\hat{\beta}_0 - \beta_0}{\sigma\sqrt{\frac{1}{n} + \frac{\bar{X}_1^2}{S_{X_1X_1}}}}}{\sqrt{\dfrac{Res.S.S.}{\sigma^2}/(n-2)}}} \sim N(0,1) \quad \sim t(n-2)$$

$$\Rightarrow \frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{X}_1^2}{S_{X_1X_1}}}} \sim t(n-2)$$

$$\Rightarrow \text{C.I. of } \beta_0 = \hat{\beta}_0 \pm t_{\alpha/2, n-2}\, \hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{X}_1^2}{S_{X_1X_1}}}$$

& $H_0 = \beta_0 = \beta_{00}$

$$t = \frac{\hat{\beta}_0 - \boxed{\beta_{00}}}{\hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{X}_1^2}{S_{X_1X_1}}}}$$

value under H_0

Page 21

e.g. p.21 $\quad H_0 = \beta_0 = 0$

p.3

$$t = \frac{0.2568 - 0}{0.538\sqrt{\frac{1}{9} + \frac{3.3667^2}{13.10}}} = 0.4831 < 1 \quad (\text{Can't reject } H_0)$$

e.g. $\underline{\beta_0 \text{ is known}}$ Model: $\underbrace{y_i - \beta_0}_{y_i'} = \beta_1 x_{i1} + e_i$

$$\hat{\beta}_1 = (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T \underset{\sim}{Y}$$

$$= \frac{\sum_{i=1}^{n} x_{i1} y_i'}{\sum_{i=1}^{n} x_{i1}^2} \quad y_i' = y_i - \beta_0 \qquad \underset{\sim}{Y} = \begin{pmatrix} y_1 - \beta_0 \\ \vdots \\ y_n - \beta_0 \end{pmatrix} \quad \underset{\sim}{\beta} = \beta_1 \quad \underset{\sim}{X} = \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix}$$

$$\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma^2}{\sum_{i=1}^{n} x_{i1}^2}\right)$$

$$\Rightarrow \frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}\sqrt{\frac{1}{\sum_{i=1}^{n} x_{i1}^2}}} \sim t_{(n-1)}$$

<u>For any p</u>

$$\hat{\underset{\sim}{\beta}} = (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T \underset{\sim}{Y} \sim N(\underset{\sim}{\beta}, \sigma^2 (\underset{\sim}{X}^T \underset{\sim}{X})^{-1})$$

$$\mathrm{Var}(\hat{\underset{\sim}{\beta}}) = \sigma^2 (\underset{\sim}{X}^T \underset{\sim}{X})^{-1}$$

$$\hat{\sigma}^2 = \frac{\text{Res S.S}}{n - p'}$$

Example 5: Example in Multiple Linear Regression

The percent survival of a certain type of animal semen after storage was measured at various combinations of concentrations of three materials used to increase chance of survival. The data are as follows:

y (% survival)	x_1 (weight %)	x_2 (weight %)	x_3 (weight %)
25.5	1.74	5.30	10.80
31.2	6.32	5.42	9.40
25.9	6.22	8.41	7.20
38.4	10.52	4.63	8.50
18.4	1.19	11.60	9.40
26.7	1.22	5.85	9.90
26.4	4.10	6.62	8.00
25.9	6.32	8.72	9.10
32.0	4.08	4.42	8.70
25.2	4.15	7.60	9.20
39.7	10.15	4.83	9.40
35.7	1.72	3.12	7.60
26.5	1.70	5.30	8.20

Summary statistics:

$H_0: \beta_2 = -2.5$

$H_1: \beta_2 \neq -2.5$

$t = \dfrac{\hat{\beta}_2 - (-2.5)}{\text{s.e. of } \hat{\beta}_2}$

$\sum_{i=1}^{13} y_i = 377.5$ $\quad \sum_{i=1}^{13} y_i^2 = 11,400.15$ $\quad \sum_{i=1}^{13} x_{i1} = 59.43$

$\sum_{i=1}^{13} x_{i2} = 81.82$ $\quad \sum_{i=1}^{13} x_{i3} = 115.40$ $\quad \sum_{i=1}^{13} x_{i1}^2 = 394.7255$

$\sum_{i=1}^{13} x_{i2}^2 = 576.7264$ $\quad \sum_{i=1}^{13} x_{i3}^2 = 1035.96$ $\quad \sum_{i=1}^{13} x_{i1}y_i = 1877.567$

$\sum_{i=1}^{13} x_{i2}y_i = 2246.661$ $\quad \sum_{i=1}^{13} x_{i3}y_i = 3337.78$ $\quad \sum_{i=1}^{13} x_{i1}x_{i2} = 360.6621$

$\sum_{i=1}^{13} x_{i1}x_{i3} = 522.078$ $\quad \sum_{i=1}^{13} x_{i2}x_{i3} = 728.31$ $\quad n = 13$

Original

$$\begin{pmatrix} 13 & 59.43 & 81.82 & 115.40 \\ 59.43 & 394.7255 & 360.6621 & 522.078 \\ 81.82 & 360.6621 & 576.7264 & 728.31 \\ 115.40 & 522.078 & 728.31 & 1035.96 \end{pmatrix}^{-1} = \begin{pmatrix} 8.06479 & -0.0825927 & -0.0941951 & -0.790527 \\ -0.0825927 & 0.00847982 & 0.00171669 & 0.00372002 \\ -0.0941951 & 0.00171669 & \boxed{0.0166294} & -0.00206331 \\ -0.790527 & 0.00372002 & -0.00206331 & 0.0886013 \end{pmatrix}$$

Or

$\text{s.e. of } \hat{\beta}_2 = \sqrt{\hat{\sigma}^2 (0.0166294)}$

$= \hat{\sigma}\sqrt{0.0166294}$

$= \sqrt{4.298}\sqrt{0.0166294}$

p.17

Centered

$$(\underset{\sim}{X}_c^T \underset{\sim}{X}_c)^{-1} = \begin{pmatrix} 13 & 0 & 0 & 0 \\ 0 & 123.039 & -13.3812 & -5.4775 \\ 0 & -13.3812 & 61.7639 & 2.0002 \\ 0 & -5.4775 & 2.0002 & 11.5631 \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} 0.0769231 & 0 & 0 & 0 \\ 0 & 0.00847981 & 0.00171669 & 0.00371998 \\ 0 & 0.00171669 & \boxed{0.0166294} & -0.00206338 \\ 0 & 0.00371998 & -0.00206338 & 0.0886011 \end{pmatrix}$$

$\Rightarrow \hat{\beta}_0 = 39.1574,\ \hat{\beta}_1 = 1.0161,\ \hat{\beta}_2 = -1.8616,\ \hat{\beta}_3 = -0.3433.$

p.22

$t = \dfrac{(-1.8016) - (-2.5)}{2.075\sqrt{0.0166294}}$

$= 2.391$

$\text{C.V.} = t_{\alpha/2,\, 13-4} = t_{0.025,9}$

$= 1.833$

3

e.g. Data in simple linear regression p. 3

$H_0 : 5\beta_0 + \beta_1 = 2$

vs $H_1 : 5\beta_0 + \beta_1 \neq 2$

$5\hat{\beta}_0 + \hat{\beta}_1 \sim N(5\beta_0 + \beta_1, \boxed{})$

$$Var(5\hat{\beta}_0 + \hat{\beta}_1)$$

$$= 25\, Var(\hat{\beta}_0) + Var(\hat{\beta}_1) + 2 * 5 * cov(\hat{\beta}_0, \hat{\beta}_1)$$

$Var(X+Y) = Var(X) + Var(Y) + 2\,cov(X,Y)$

$$= 25\, \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}_1^2}{S_{x_1x_1}}\right) + \frac{\sigma^2}{S_{x_1x_1}} - 10\, \frac{\sigma^2 \bar{x}_1}{S_{x_1x_1}}$$

$$= \sigma^2 \left\{ \frac{25}{n} + \frac{25\bar{x}_1^2 - 10\bar{x}_1 + 1}{S_{x_1x_1}} \right\}$$

$$= \sigma^2 \left\{ \frac{25}{n} + \frac{(5\bar{x}_1 - 1)^2}{S_{x_1x_1}} \right\}$$

$$\frac{\dfrac{(5\hat{\beta}_0 + \hat{\beta}_1) - 2}{\cancel{\sigma}\sqrt{\dfrac{25}{n} + \dfrac{(5\bar{x}_1-1)^2}{S_{x_1x_1}}}} \sim N(0,1)}{\sqrt{\dfrac{Res.S.S.}{\cancel{\sigma^2}} \Big/ (n-2)}} \sim t$$

$\dfrac{Res.S.S.}{(n-2)} = \hat{\sigma}^2$

$$t = \frac{(5\hat{\beta}_0 + \hat{\beta}_1) - \text{value under } H_0}{\text{s.e. of } (5\hat{\beta}_0 + \hat{\beta}_1)}$$

$$= \frac{(5\hat{\beta}_0 + \hat{\beta}_1) - 2}{\hat{\sigma}\sqrt{\dfrac{25}{n} + \dfrac{(5\bar{x}_1 - 1)^2}{S_{x_1x_1}}}}$$

$$= \frac{(5 * 0.2568 + 2.9303) - 2}{0.538\sqrt{\dfrac{25}{9} + \dfrac{(5 * 3.667 - 1)^2}{13.10}}}$$

p. 22

Critical value $= t_{\alpha/2,\, (n-2)}$ ← d.f. of Res S.S.

Example 5: Example in Multiple Linear Regression

The percent survival of a certain type of animal semen after storage was measured at various combinations of concentrations of three materials used to increase chance of survival. The data are as follows:

y (% survival)	x_1 (weight %)	x_2 (weight %)	x_3 (weight %)
25.5	1.74	5.30	10.80
31.2	6.32	5.42	9.40
25.9	6.22	8.41	7.20
38.4	10.52	4.63	8.50
18.4	1.19	11.60	9.40
26.7	1.22	5.85	9.90
26.4	4.10	6.62	8.00
25.9	6.32	8.72	9.10
32.0	4.08	4.42	8.70
25.2	4.15	7.60	9.20
39.7	10.15	4.83	9.40
35.7	1.72	3.12	7.60
26.5	1.70	5.30	8.20

Summary statistics:

$H_0 : \beta_0 = 40$

$H_1 : \beta_0 \neq 40$

$\sum_{i=1}^{13} y_i = 377.5$ $\quad \sum_{i=1}^{13} y_i^2 = 11,400.15$ $\quad \sum_{i=1}^{13} x_{i1} = 59.43$

$\sum_{i=1}^{13} x_{i2} = 81.82$ $\quad \sum_{i=1}^{13} x_{i3} = 115.40$ $\quad \sum_{i=1}^{13} x_{i1}^2 = 394.7255$

$\sum_{i=1}^{13} x_{i2}^2 = 576.7264$ $\quad \sum_{i=1}^{13} x_{i3}^2 = 1035.96$ $\quad \sum_{i=1}^{13} x_{i1}y_i = 1877.567$

$\sum_{i=1}^{13} x_{i2}y_i = 2246.661$ $\quad \sum_{i=1}^{13} x_{i3}y_i = 3337.78$ $\quad \sum_{i=1}^{13} x_{i1}x_{i2} = 360.6621$

$\sum_{i=1}^{13} x_{i1}x_{i3} = 522.078$ $\quad \sum_{i=1}^{13} x_{i2}x_{i3} = 728.31$ $\quad n = 13$

$$t = \frac{\hat{\beta}_0 - 40}{\text{S.e. of } \hat{\beta}_0}$$

$$\text{S.e. of } \hat{\beta}_0 = \sqrt{\hat{\sigma}^2 (\quad)}$$

Original

$$\begin{pmatrix} 13 & 59.43 & 81.82 & 115.40 \\ 59.43 & 394.7255 & 360.6621 & 522.078 \\ 81.82 & 360.6621 & 576.7264 & 728.31 \\ 115.40 & 522.078 & 728.31 & 1035.96 \end{pmatrix}^{-1} = \begin{pmatrix} 8.06479 & -0.0825927 & -0.0941951 & -0.790527 \\ -0.0825927 & 0.00847982 & 0.00171669 & 0.00372002 \\ -0.0941951 & 0.00171669 & 0.0166294 & -0.00206331 \\ -0.790527 & 0.00372002 & -0.00206331 & 0.0886013 \end{pmatrix}$$

Or

Centred

$$(\underset{\sim}{X}_c^T \underset{\sim}{X}_c)^{-1} = \begin{pmatrix} 13 & 0 & 0 & 0 \\ 0 & 123.039 & -13.3812 & -5.4775 \\ 0 & -13.3812 & 61.7639 & 2.0002 \\ 0 & -5.4775 & 2.0002 & 11.5631 \end{pmatrix}^{-1}$$

Given

$$= \begin{pmatrix} 0.0769231 & 0 & 0 & 0 \\ 0 & 0.00847981 & 0.00171669 & 0.00371998 \\ 0 & 0.00171669 & 0.0166294 & -0.00206338 \\ 0 & 0.00371998 & -0.00206338 & 0.0886011 \end{pmatrix}$$

$\Rightarrow \hat{\beta}_0 = 39.1574,\ \hat{\beta}_1 = 1.0161,\ \hat{\beta}_2 = -1.8616,\ \hat{\beta}_3 = -0.3433.$

$H_0: \beta_0 = 40$ intercept in centered model

$$\hat\beta_0 = \hat\beta_0' - \hat\beta_1 \bar X_1 - \hat\beta_2 \bar X_2 - \hat\beta_3 \bar X_3$$

$$= \bar y - \hat\beta_1 \bar X_1 - \hat\beta_2 \bar X_2 - \hat\beta_3 \bar X_3 \rightarrow (1, -\bar X_1, -\bar X_2, -\bar X_3)\begin{pmatrix}\bar y\\ \hat\beta_1\\ \hat\beta_2\\ \hat\beta_3\end{pmatrix}$$

$$Var(\hat\beta_0) = Var\Big(\underbrace{(1, -\bar X_1, -\bar X_2, -\bar X_3)}_{\underset{\sim}{c}}\underbrace{\begin{pmatrix}\bar y\\ \hat\beta_1\\ \hat\beta_2\\ \hat\beta_3\end{pmatrix}}_{\underset{\sim}{z}}\Big)$$

By Theorem 3.2 in p.14 $\Rightarrow Var(\underset{\sim}{c}\underset{\sim}{z}) = \underset{\sim}{c}\, Var(\underset{\sim}{z})\, \underset{\sim}{c}^T$

$$\Rightarrow Var(\hat\beta_0) = (1, -\bar X_1, -\bar X_2, -\bar X_3)\, Var\begin{pmatrix}\bar y\\ \hat\beta_1\\ \hat\beta_2\\ \hat\beta_3\end{pmatrix}\begin{pmatrix}1\\ -\bar X_1\\ -\bar X_2\\ -\bar X_3\end{pmatrix}$$

$Var(y_i) = \sigma^2$

$Var(\bar y) = \frac{\sigma^2}{n}$

0 0

$$\begin{pmatrix} Var(\bar y) & Cov(\bar y, \hat\beta_1) & Cov(\bar y, \hat\beta_2) & Cov(\bar y, \hat\beta_3) \\ & & \boxed{\text{given}} & \end{pmatrix}$$

0

$$t = \frac{\hat\beta_0 - 40}{\text{s.e. of } \hat\beta_0} = -0.1431$$ Can't reject H_0

t-test

- one-sided alternative / two-sided alternative
- For one linear combination of reg. coeff & intercept.

e.g. $H_0: 5\beta_0 + \beta_1 = 2$

$H_0: \beta_0 = 0$

But, e.g. $H_0 = \beta_0 = 0$, $\beta_1 = 1$

$H_0 = \beta_0 = 0 \leftarrow$ t-test $\updownarrow$ NO

$H_1 = \beta_1 = 1 \leftarrow$ t-test

$$\widehat{Var}(\hat\beta_0) = \hat\sigma^2 \left(\frac{1}{n} + \frac{\bar{x}_1^2}{S_{x_1x_1}}\right)$$

$$\widehat{Var}(\hat\beta_1) = \frac{\hat\sigma^2}{S_{x_1x_1}}$$

4.2 F-test

- two-sided alternative $H_0 = \beta_0 = 0$ vs $H_1 = \beta_0 \neq 0$

- one or more one linear ~~to~~ ~~bet~~ ~~to~~ combinations of $\underline{\beta}$

A Partitioning total variability

$$= \sum_{i=1}^n (y_i - \boxed{\bar{y}})^2 = \text{total } \cancel{\text{S.S}} \text{ sum of squares} = \text{corrected total S.S.}$$

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (y_i - \hat{y}_i + \hat{y}_i - \bar{y})^2$$

$$= \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + 2 \underbrace{\sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{y})}$$

$\hat{y}_i = \boxed{\hat\beta_0} + \hat\beta_1 x_{i1} + \dots + \hat\beta_p x_{ip}$

$$\text{Total S.S.} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$$

$$\sum_{i=1}^n (y_i - \hat{y}_i)\hat{y}_i - \bar{y}\sum_{i=1}^n (y_i - \hat{y}_i)$$

$$= \sum_{i=1}^n \hat{e}_i(\hat\beta_0 + \hat\beta_1 x_{i1} + \dots + \hat\beta_p x_{ip}), \qquad \sum_{i=1}^n \hat{e}_i = 0$$

$$= \hat\beta_0 \boxed{\sum_{i=1}^n \hat{e}_i} + \hat\beta_1 \boxed{\sum_{i=1}^n \hat{e}_i x_{i1}} + \dots + \hat\beta_p \boxed{\sum_{i=1}^n \hat{e}_i x_{ip}} = 0$$

Total S.S. ↑ total variability

= Res S.S. ↑ unexplained variability

= $\boxed{\text{Reg S.S.}}$ = $R(\underline{\beta} \mid \beta_0)$ ↑ Variability explained by the model

$$R^2 = \frac{\text{Reg S.S}}{\text{total S.S}}$$

B. Distributions

① Distribution of Total S.S.

Theorem 3.4 in p.17 $\quad \underset{\sim}{Y} \sim MN(\underset{\sim}{\mu}, \underset{\sim}{I})$

then $\underset{\sim}{Y}^T \underset{\sim}{A} \underset{\sim}{Y} \sim \chi^2(k, \lambda)$

k: rank of $\underset{\sim}{A}$ $\qquad \lambda$: $\underset{\sim}{\mu}^T \underset{\sim}{A} \underset{\sim}{\mu}$

Total S.S. $\sum_{i=1}^{n} (y_i - \bar{y})^2$

$$= \underbrace{(\underset{\sim}{Y} - \bar{\underset{\sim}{Y}})}_{\parallel \; (\underset{\sim}{I} - \frac{1}{n}\underset{\sim}{J})\underset{\sim}{Y}}{}^T (\underset{\sim}{Y} - \bar{\underset{\sim}{Y}})$$

$$\underset{n\times 1}{\bar{\underset{\sim}{Y}}} = \begin{pmatrix} \bar{y} \\ \vdots \\ \bar{y} \end{pmatrix} \qquad \bar{y} = \sum_{i=1}^{n} y_i / n$$

$$= \begin{pmatrix} \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & & \vdots \\ \frac{1}{n} & \cdots & \frac{1}{n} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$= \frac{1}{n} \underbrace{\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix}}_{J} \underbrace{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}}_{\underset{\sim}{Y}}$$

$$\text{Total S.S.} = \underset{\sim}{Y}^T \boxed{(\underset{\sim}{I} - \frac{1}{n}\underset{\sim}{J})^T (\underset{\sim}{I} - \frac{1}{n}\underset{\sim}{J})} \underset{\sim}{Y}$$

$$\boxed{(\underset{\sim}{I} - \frac{1}{n}\underset{\sim}{J})^T (\underset{\sim}{I} - \frac{1}{n}\underset{\sim}{J})} = (\underset{\sim}{I} - \frac{1}{n}\underset{\sim}{J})(\underset{\sim}{I} - \frac{1}{n}\underset{\sim}{J})$$

$$= \underset{\sim}{I} - \frac{1}{n}\underset{\sim}{J} - \frac{1}{n}\underset{\sim}{J} + \frac{1}{n^2}\boxed{\underset{\sim}{J}\underset{\sim}{J}}$$

$$\underset{\sim}{J}\underset{\sim}{J} = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} n & \cdots & n \\ \vdots & & \vdots \\ n & \cdots & n \end{pmatrix} = n\underset{\sim}{J}$$

$$= \underset{\sim}{Y}^T \boxed{(\underset{\sim}{I} - \frac{1}{n}\underset{\sim}{J})} \underset{\sim}{Y}$$

symmetric & idempotent

By Theorem 3.4 $\quad \dfrac{\text{Total S.S.}}{\sigma^2} = \underset{\sim}{Y}^{*T} (\underset{\sim}{I} - \frac{1}{n}\underset{\sim}{J}) \underset{\sim}{Y}^*$ where $\underset{\sim}{Y}^* = \dfrac{\underset{\sim}{Y}}{\sigma}$

$$\sim \chi^2(k, \lambda)$$

$$k = \text{rank of } \Big(\underset{n\times n}{\underset{\sim}{I}} - \frac{1}{n}\underset{n\times n}{\underset{\sim}{J}}\Big) \qquad \underset{\sim}{J} = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$$

$$= \text{trace of } \Big(\underset{\sim}{I} - \frac{1}{n}J\Big) \text{ as } \underset{\sim}{I} - \frac{1}{n}J \text{ is idempotent matrix}$$

$$= n-1$$

$$\lambda = \underset{\sim}{\mu}^{*T}\Big(\underset{\sim}{I} - \frac{1}{n}\underset{\sim}{J}\Big)\underset{\sim}{\mu}^{*} \qquad \underset{\sim}{\mu}^{*} = \frac{\underset{\sim}{X}\underset{\sim}{\beta}}{\sigma}$$

$$= \frac{1}{\sigma^2}\underset{\sim}{\beta}^T\underset{\sim}{X}^T\Big(\underset{\sim}{I} - \frac{1}{n}\underset{\sim}{J}\Big)\underset{\sim}{X}\underset{\sim}{\beta}$$

$$= \frac{1}{\sigma^2}\underset{\sim}{\beta}^T\Big(\boxed{\underset{\sim}{X}^T\underset{\sim}{X}} - \frac{1}{n}\boxed{\underset{\sim}{X}^T\underset{\sim}{J}\underset{\sim}{X}}\Big)\underset{\sim}{\beta} \qquad \boxed{\begin{matrix} x'_{ij} = x_{ij} - \bar{x}_j \\ \sum_{i=1}^{n} x'_{ij} = 0 \end{matrix}}$$

$$\begin{pmatrix} n & 0 \\ 0 & \underset{\sim}{X}_c^T\underset{\sim}{X}_c \end{pmatrix}$$

$$\underset{\sim}{X}^T\underset{\sim}{J} = \begin{pmatrix} 1 & \cdots & 1 \\ x'_{11} & \cdots & x'_{n1} \\ \vdots & & \vdots \\ x'_{1p} & \cdots & x'_{np} \end{pmatrix}\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$$

$$= \begin{pmatrix} n & n & \cdots & n \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\underset{\sim}{X}^T\underset{\sim}{J}\underset{\sim}{X} = \begin{pmatrix} n & n & \cdots & n \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}\begin{pmatrix} 1 & x'_{11} & \cdots & x'_{1p} \\ \vdots & \vdots & & \vdots \\ 1 & x'_{n1} & \cdots & x'_{np} \end{pmatrix}$$

$$= \begin{pmatrix} \boxed{n^2} & 0 & \cdots & 0 \\ 0 & 0 & & \underset{\sim}{0} \\ \vdots & & \ddots & \\ 0 & \underset{\sim}{0} & & 0 \end{pmatrix}$$

$$= \frac{1}{\sigma^2}\underset{\sim}{\beta}^T\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \boxed{\underset{\sim}{X}_c^T\underset{\sim}{X}_c} & \\ 0 & & \end{pmatrix}\underset{\sim}{\beta}$$

$(i,j)^{th}$ element in $\underset{\sim}{X}_c^T\underset{\sim}{X}_c$

$= S_{x_ix_j}$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{p}\sum_{j=1}^{p}\beta_i\beta_j S_{x_ix_j}$$

$$\frac{\text{Total S.S}}{\sigma^2} \sim \chi^2\Big(n-1,\ \frac{1}{\sigma^2}\sum_{i=1}^{p}\sum_{j=1}^{p}\beta_i\beta_j S_{x_ix_j}\Big)$$

For $p=1$, $E(\text{Total S.S.}) = E\Big(\sum_{i=1}^{n}(y_i - \bar{y})^2\Big)$

$$= (n-1)\sigma^2 + \beta_1^2 S_{x_1x_1}$$

$$\frac{\text{Total S.S}}{\sigma^2} = \frac{\text{Res S.S.}}{\sigma^2} + \frac{\text{Reg S.S.}}{\sigma^2}$$

$$\Downarrow \qquad\qquad \Downarrow \qquad\qquad \Downarrow$$

$$\sim \chi^2(n-1,\lambda) \qquad \underbrace{\sim \chi^2(n-p')}_{\text{indep}} \qquad \sim ?$$

where $\lambda = \frac{1}{\sigma^2}\sum_{i=1}^{p}\sum_{j=1}^{p}\beta_i\beta_j S_{x_ix_j}$

MATH3423

$$\underset{\sim \chi^2(r)}{W} = \underset{\sim \chi^2(r_1)}{U} + V \Rightarrow V \sim \chi^2(r-r_1)$$

Theorem 4.2 (p.25)

$W \sim \chi^2(r,\lambda)$, $U \sim \chi^2(r_1,\lambda_1)$

U & V are indep.

$\Rightarrow V \sim \chi^2(r-r_1, \lambda-\lambda_1)$

$\frac{\text{Total S.S.}}{\sigma^2} \sim \chi^2(n-1,\lambda)$, $\frac{\text{Res S.S.}}{\sigma^2} \sim \chi^2(n-\boxed{p'})$

$p' = 1+p$

intercept

If Res S.S. & Reg S.S. are indep.

then $\frac{\text{Reg S.S.}}{\sigma^2} \sim \chi^2(p,\lambda)$

Theorem 4.1 (p.25)

$\underset{\sim}{Y} \sim MN(\underset{\sim}{\mu}, \underset{\sim}{\Sigma})$ with rank n

If $\underset{\sim}{A}^T\underset{\sim}{\Sigma}\underset{\sim}{B} = 0$, then two quadratic forms of

$\underset{\sim}{Y}^T\underset{\sim}{A}\underset{\sim}{Y}$ & $\underset{\sim}{Y}^T\underset{\sim}{B}\underset{\sim}{Y}$ are independent.

In our model, $\underset{\sim}{\Sigma} = \sigma^2\underset{\sim}{I}$

Res S.S. $= \underset{\sim}{Y}^T\underset{\sim}{A}\underset{\sim}{Y}$, Reg S.S. $= \underset{\sim}{Y}^T\underset{\sim}{B}\underset{\sim}{Y}$

Show: $\underset{\sim}{A}^T\underset{\sim}{B} = \underset{\sim}{0}$