Math 3424 - Regression Analysis

Chapter 1: Simple & Multiple Linear Regression

1. Introduction

$$\underset{\sim}{Y}_{m\times 1}\begin{cases} m=1 & -- \quad \text{MATH3424} \\ m>1 & -- \quad \text{MATH4424} \end{cases}$$

$$\underset{\sim}{X}_{p\times 1}\begin{cases} p=1 & -- \quad \text{Simple linear regression – Chapter 2 in textbook} \\ p>1 & -- \quad \text{Multiple linear regression – Chapter 3 in textbook} \end{cases}$$

Linear – $f(x) = \alpha + \beta x = \beta_0 + \beta_1 x_1$

Linear on a parameter, not linear on X , so

$$\alpha + \beta x^2 \Rightarrow x' = x^2 \Rightarrow \alpha + \beta x' \qquad \text{(linear)}$$

$$\alpha + \beta \ln x \Rightarrow x' = \ln x \Rightarrow \alpha + \beta x' \qquad \text{(linear)}$$

$$\begin{array}{rccccc} & Y & = & \alpha\beta^x & & \\ \Rightarrow & \ln Y & = & \ln\alpha & + \quad x & \ln\beta \\ & & & \uparrow & & \uparrow \\ & & & \alpha' & & \beta' \end{array}$$

$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + e_i$

1. y_i measured with error, i.e., y_{T_i} is unobserved.
   If $y_{T_i} = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}$, then $y_i = y_{T_i} + e_{y_i}$.
2. $y_{T_i} = y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + e_i$, i.e., the relationship between y_i and $\underset{\sim}{x}_i$ is not prefect linear related.

$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + e_i \qquad\qquad i = 1, 2, 3, \ldots, n$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}_{n\times 1} = \begin{pmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1p} \\ 1 & x_{21} & x_{22} & \ldots & x_{2p} \\ \vdots & \vdots & \ddots & \ldots & \vdots \\ 1 & x_{n1} & x_{n2} & \ldots & x_{np} \end{pmatrix}_{n\times(p+1)} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}_{(p+1)} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}_{n\times 1}$$

$$\underset{\sim}{Y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{e} \qquad\qquad X : \text{design matrix}$$

Assumptions:

$$\left.\begin{array}{lll} (1) & \mathrm{E}(e_i) = 0 & \\ (2) & \mathrm{Var}(e_i) = \sigma^2 & \\ (3) & \mathrm{Cov}(e_i, e_j) = 0 & \text{(for } i \neq j) \\ (4) & e_i \sim N & \end{array}\right\} \quad e_i \underset{\sim}{iid} N(0, \sigma^2)$$

$\Rightarrow \quad \underset{\sim}{e} \sim MN(\underset{\sim}{0}, \underset{\sim}{\Sigma})$

$\Rightarrow \quad \underset{\sim}{Y} \sim MN(\underset{\sim}{X}\underset{\sim}{\beta}, \underset{\sim}{\Sigma})$

2. Estimation

2.1. Estimation of $\underset{\sim}{\beta}$

Method of Estimation: Lease Squares & Maximum likelihood

If $e_i \underset{\sim}{iid} N(0, \sigma^2)$, then estimators by methods of least squares and maximum likelihood are the same. But, it is easy to find the estimators by least squares estimation and no need to make any distribution assumption.

Least Square Estimation

Define residual

$$
\begin{aligned}
e_i &= y_i - (\beta_0 + \beta_1 x_{i1}) \qquad &\text{– unobservable} \\
\hat{e}_i &= y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1}) \qquad &\text{– observable} \\
\hat{e}_i &\text{ – estimate of } e_i = y_i - \hat{y}_i
\end{aligned}
$$

Find $\hat{\beta}_0$, $\hat{\beta}_1$ such that $\sum_{i=1}^{n} \hat{e}_i^2$ is minimized

$\sum_{i=1}^{n} \hat{e}_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$ – define: residual sum of squares RSS

For p=1

$$
\begin{cases}
\dfrac{\partial}{\partial \hat{\beta}_0} \sum_{i=1}^{n} (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1}))^2 = 2 \sum_{i=1}^{n} (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1}))(-1) = 0 \\
\dfrac{\partial}{\partial \hat{\beta}_1} \sum_{i=1}^{n} (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1}))^2 = 2 \sum_{i=1}^{n} (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1})(-x_{i1}) = 0
\end{cases}
$$

$$
\Rightarrow \begin{cases}
\sum_{i=1}^{n} y_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum_{i-1}^{n} x_{i1} = 0 \\
\sum_{i=1}^{n} x_{i1} y_i - \hat{\beta}_0 \sum_{i=1}^{n} x_{i1} - \hat{\beta}_1 \sum_{i-1}^{n} x_{i1}^2 = 0
\end{cases}
$$

$$
\Rightarrow \begin{cases}
n\hat{\beta}_0 + \sum_{i=1}^{n} x_{i1} \hat{\beta}_1 = \sum_{i=1}^{n} y_i \\
\sum_{i=1}^{n} x_{i1} \hat{\beta}_0 + \sum_{i=1}^{n} x_i^2 \hat{\beta}_1 = \sum_{i=1}^{n} x_{i1} y_i
\end{cases}
$$

$$
\Rightarrow \begin{cases}
\hat{\beta}_1 = \dfrac{\sum_{i=1}^{n} x_{i1} y_i - \sum_{i=1}^{n} x_{i1} \sum_{i=1}^{n} y_i / n}{\sum_{i=1}^{n} x_{i1}^2 - (\sum_{i=1}^{n} x_{i1})^2 / n} \\
\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \bar{x}_1
\end{cases}
$$

$$
\Rightarrow \hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_{i1} - \bar{x}_1)(y_i - \overline{y})}{\sum_{i=1}^{n} (x_{i1} - \bar{x}_1)^2} = \frac{S_{x_1 y}}{S_{x_1 x_1}}
$$

$$\sum_{i=1}^{n}(y_i - (\hat\beta_0 + \hat\beta_1 x_{i1})) = 0$$

$$\sum_{i=1}^{n} x_{i1}(y_i - (\hat\beta_0 + \hat\beta_1 x_{i1})) = 0$$

$$\Rightarrow \left.\begin{array}{r} \sum_{i=1}^{n} \hat e_i = 0 \\ \sum_{i=1}^{n} x_{i1}\hat e_i = 0 \end{array}\right\} \qquad \text{Properties of } \hat e_i$$

Example 1: Example in Simple Linear Regression

i	1	2	3	4	5	6	7	8	9
x_i	1.5	1.8	2.4	3.0	3.5	3.9	4.4	4.8	5.0
y_i	4.8	5.7	7.0	8.3	10.9	12.4	13.1	13.6	15.3

Summary statistics:

$$\begin{array}{lll} \sum_{i=1}^{9} x_i = 30.3 & \sum_{i=1}^{9} y_i = 91.1 & \sum_{i=1}^{9} x_i y_i = 345.09 \\ \sum_{i=i}^{9} x_i^2 = 115.11 & \bar x = 3.3667 & \bar y = 10.1222 \end{array}$$

$\Rightarrow \hat\beta_1 = 2.9303$ and $\hat\beta_0 = 0.2568$.

Thus, the estimated regression line (or the fitted line) is given by

$$\hat y = 0.2568 + 2.9303x$$

.

For any p

$$\text{Min } \sum_{i=1}^{n}(y_i - \hat y_i)^2$$

$$\Rightarrow \text{ Min } \sum_{i=1}^{n}(y_i - (\hat\beta_0 + \hat\beta_1 x_{i1} + \ldots + \hat\beta_p x_{ip}))^2$$

$$= \sum_{i=1}^{n} \hat e_i^2$$

$$= \begin{pmatrix} \hat e_1 & \ldots & \hat e_n \end{pmatrix} \begin{pmatrix} \hat e_1 \\ \vdots \\ \hat e_n \end{pmatrix}$$

$$= \begin{pmatrix} y_i - (\hat\beta_0 + \hat\beta_1 x_{i1} + \ldots + \hat\beta_p x_{ip}) & \ldots & y_n - (\hat\beta_0 + \hat\beta_1 x_{n1} + \ldots + \hat\beta_p x_{np}) \end{pmatrix} \begin{pmatrix} y_i - (\hat\beta_0 + \hat\beta_1 x_{i1} + \ldots + \hat\beta_p x_{ip}) \\ \vdots \\ y_n - (\hat\beta_0 + \hat\beta_1 x_{n1} + \ldots + \hat\beta_p x_{np}) \end{pmatrix}$$

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

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

Using the following formulae:

Let β be an $n \times 1$ column vector

1. A is an $n \times 1$ vector

$$\frac{\partial A^T \beta_{\ 1\times 1}}{\partial \beta_{\ n\times 1}} = \frac{\partial \beta^T A_{\ 1\times 1}}{\partial \beta_{\ n\times 1}} = A_{n\times 1}$$

2. C is an $n \times n$ matrix

$$\frac{\partial \beta^T C \beta_{\ 1\times 1}}{\partial \beta_{\ n\times 1}} = (C^T + C)\beta_{n\times 1}$$

$$\begin{aligned}
&\frac{\partial}{\partial \hat{\underset{\sim}{\beta}}}(\underset{\sim}{Y} - \underset{\sim}{X}\hat{\underset{\sim}{\beta}})^T(\underset{\sim}{Y} - \underset{\sim}{X}\hat{\underset{\sim}{\beta}}) \\
\Rightarrow \quad & -2\underset{\sim}{X}^T\underset{\sim}{Y} + 2(\underset{\sim}{X}^T\underset{\sim}{X})\hat{\underset{\sim}{\beta}} = 0 \\
\Rightarrow \quad & \hat{\underset{\sim}{\beta}} = (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{Y}
\end{aligned}$$

where

$$\underset{\sim}{X}^T\underset{\sim}{X} = \begin{pmatrix}
n & \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i2} & \dots & \sum_{i=1}^n x_{ip} \\
\sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i1}^2 & \sum_{i=1}^n x_{i1}x_{i2} & \dots & \sum_{i=1}^n x_{i1}x_{ip} \\
\sum_{i=1}^n x_{i2} & \sum_{i=1}^n x_{i2}x_{i1} & \sum_{i=1}^n x_{i2}^2 & \dots & \sum_{i=1}^n x_{i2}x_{ip} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\sum_{i=1}^n x_{ip} & \sum_{i=1}^n x_{ip}x_{i1} & \sum_{i=1}^n x_{ip}x_{i2} & \dots & \sum_{i=1}^n x_{ip}^2
\end{pmatrix}_{(p+1)\times(p+1)}$$

$$\underset{\sim}{X}^T\underset{\sim}{Y} = \begin{pmatrix}
\sum_{i=1}^n y_i \\
\sum_{i=1}^n x_{i1}y_i \\
\vdots \\
\sum_{i=1}^n x_{ip}y_i
\end{pmatrix}$$

Then,

$$\begin{aligned}
y_i &= \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + e_i \\
y_i &= \beta_0 + \beta_1(x_{i1} - \bar{x}_1) + \ldots + \beta_p(x_{ip} - \bar{x}_p) + e_i + (\beta_1\bar{x}_1 + \ldots + \beta_p\bar{x}_p) \\
y_i &= \beta_0' + \beta_1 x_{i1}' + \ldots + \beta_p x_{ip}' + e_i \qquad \text{where} \quad {\beta_0}' = \beta_0 + \beta_1\bar{x}_1 + \beta_2\bar{x}_2 + \ldots + \beta_p\bar{x}_p \\
\Rightarrow \hat{\beta}_0 &= \hat{\beta}_0' - \hat{\beta}_1\bar{x}_1 - \ldots - \hat{\beta}_p\bar{x}_p
\end{aligned}$$

Example 2: Formula in matrix form for p=1

$$\underset{\sim}{X}^T\underset{\sim}{X} = \begin{pmatrix} n & \sum_{i=1}^{n} x_{i1} \\ \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i1}^2 \end{pmatrix}_{(1+1)\times(1+1)} \quad , \quad \underset{\sim}{X}^T\underset{\sim}{Y} = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i1} y_i \end{pmatrix}$$

$$\left(\underset{\sim}{X}^T\underset{\sim}{X}\right)^{-1} = \frac{1}{n\sum_{i=1}^{n} x_{i1}^2 - \left(\sum_{i=1}^{n} x_{i1}\right)^2} \begin{pmatrix} \sum_{i=1}^{n} x_{i1}^2 & -\sum_{i=1}^{n} x_{i1} \\ -\sum_{i=1}^{n} x_{i1} & n \end{pmatrix}$$

Example 3: Centered Model

Let $x'_{ij} = x_{ij} - \bar{x}_j$ for $i = 1, \ldots, n; j = 1, \ldots, p$, then

$$\underset{\sim}{X}^T\underset{\sim}{X} = \begin{pmatrix} 1 & \ldots & 1 \\ x'_{11} & \ldots & x'_{n1} \\ \vdots & \ddots & \vdots \\ x'_{1p} & \ldots & x'_{np} \end{pmatrix}_{(p+1)\times n} \begin{pmatrix} 1 & x'_{11} & \ldots & x'_{1p} \\ 1 & x'_{21} & \ldots & x'_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x'_{n1} & \ldots & x'_{n1} \end{pmatrix}_{n\times(p+1)}$$

$$= \begin{pmatrix} n & \sum_{i=1}^{n} x'_{i1} & \sum_{i=1}^{n} x'_{i2} & \ldots & \sum_{i=1}^{n} x'_{ip} \\ \sum_{i=1}^{n} x'_{i1} & \sum_{i=1}^{n} x_{i1}^{2\,\prime} & \ldots & \ldots & \sum_{i=1}^{n} x'_{i1} x'_{ip} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} x'_{ip} & \ldots & \ldots & \ldots & \sum_{i=1}^{n} x_{ip}^{2\,\prime} \end{pmatrix}$$

$$= \begin{pmatrix} n & 0 & 0 & \ldots & 0 \\ 0 & \sum_{i=1}^{n}(x_{i1} - \bar{x}_1)^2 & \sum_{i=1}^{n}(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) & \ldots & \sum_{i=1}^{n}(x_{i1} - \bar{x}_1)(x_{ip} - \bar{x}_p) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \sum_{i=1}^{n}(x_{i1} - \bar{x}_1)(x_{ip} - \bar{x}_p) & \sum_{i=1}^{n}(x_{i2} - \bar{x}_2)(x_{ip} - \bar{x}_p) & \ldots & \sum_{i=1}^{n}(x_{ip} - \bar{x}_p)^2 \end{pmatrix}$$

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

$$(\underset{\sim}{X}^T\underset{\sim}{X})^{-1} = \begin{pmatrix} \frac{1}{n} & \underset{\sim}{0} \\ \underset{\sim}{0} & (\underset{\sim}{X}_c{}^T \underset{\sim}{X}_c)^{-1} \end{pmatrix}$$

$$\underset{\sim}{X}^T\underset{\sim}{Y} = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i1}{}' y_i \\ \vdots \\ \sum_{i=1}^{n} x_{ip}{}' y_i \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} (x_{i1} - \bar{x}_1) y_i \\ \vdots \\ \sum_{i=1}^{n} (x_{i2} - \bar{x}_p) y_i \end{pmatrix}$$

<u>Example 4: Intercept is known</u>

$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + e_i, \qquad i = 1, \ldots, n, \quad \beta_0$ is known

$\Rightarrow y_i - \beta_0 = \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + e_i$

$$\underset{\sim}{Y} = \begin{pmatrix} y_1 - \beta_0 \\ y_2 - \beta_0 \\ \vdots \\ y_n - \beta_0 \end{pmatrix}, \quad \underset{\sim}{X} = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1p} \\ x_{21} & x_{22} & \ldots & x_{2p} \\ \vdots & & & \\ x_{n1} & x_{n2} & \ldots & x_{np} \end{pmatrix}, \quad \underset{\sim}{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}$$

$$\underset{\sim}{X}^T\underset{\sim}{X} = \begin{pmatrix} \sum_{i=1}^{n} x_{i1}^2 & \sum_{i=1}^{n} x_{i1}x_{i2} & \ldots & \sum_{i=1}^{n} x_{i1}x_{ip} \\ \sum_{i=1}^{n} x_{i2}x_{i1} & \sum_{i=1}^{n} x_{i2}^2 & \ldots & \sum_{i=1}^{n} x_{i1}x_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} x_{ip}x_{i1} & \sum_{i=1}^{n} x_{ip}x_{i2} & \ldots & \sum_{i=1}^{n} x_{ip}^2 \end{pmatrix}, \quad \underset{\sim}{X}^T\underset{\sim}{Y} = \begin{pmatrix} \sum_{i=1}^{n} x_{i1}(y_i - \beta_0) \\ \sum_{i=1}^{n} x_{i2}(y_i - \beta_0) \\ \vdots \\ \sum_{i=1}^{n} x_{ip}(y_i - \beta_0) \end{pmatrix}$$

For p=1,

$$\begin{aligned} \hat{\beta}_1 &= \frac{\sum_{i=1}^{n} (y_i - \beta_0) x_i}{\sum_{i=1}^{n} x_i{}^2} \\ &= \frac{\sum_{i=1}^{n} x_i y_i - \beta_0 \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i{}^2} \end{aligned}$$

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:

$$\begin{array}{lll}
\sum_{i=1}^{13} y_i = 377.5 & \sum_{i=1}^{13} y_i^2 = 11,400.15 & \sum_{i=1}^{13} x_{i1} = 59.43 \\
\sum_{i=1}^{13} x_{i2} = 81.82 & \sum_{i=1}^{13} x_{i3} = 115.40 & \sum_{i=1}^{13} x_{i1}^2 = 394.7255 \\
\sum_{i=1}^{13} x_{i2}^2 = 576.7264 & \sum_{i=1}^{13} x_{i3}^2 = 1035.96 & \sum_{i=1}^{13} x_{i1}y_i = 1877.567 \\
\sum_{i=1}^{13} x_{i2}y_i = 2246.661 & \sum_{i=1}^{13} x_{i3}y_i = 3337.78 & \sum_{i=1}^{13} x_{i1}x_{i2} = 360.6621 \\
\sum_{i=1}^{13} x_{i1}x_{i3} = 522.078 & \sum_{i=1}^{13} x_{i2}x_{i3} = 728.31 & n = 13
\end{array}$$

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

Or

$$\begin{aligned}
(\underset{\sim}{X}_c^T \underset{\sim}{X}_c)^{-1} &= \begin{pmatrix} 13 & 0 & 0 & 0 \\ 0 & 123.039 & -13.3812 & -5.4775 \\ 0 & -13.3812 & 61.7639 & 2.0002 \\ 0 & -5.4775 & 2.0002 & 11.5631 \end{pmatrix}^{-1} \\
&= \begin{pmatrix} 0.0769231 & 0 & 0 & 0 \\ 0 & 0.00847981 & 0.00171669 & 0.00371998 \\ 0 & 0.00171669 & 0.0166294 & -0.00206338 \\ 0 & 0.00371998 & -0.00206338 & 0.0886011 \end{pmatrix}
\end{aligned}$$

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

Maximum Likelihood Estimation

Maximum Likelihood Estimators of the Regression coefficient, $\hat{\beta}_0, \hat{\beta}_1, \hat{\sigma}^2$

Model:

$$y_i = \beta_0 + \beta_1 x_{i1} + e_i \qquad\qquad i = 1, \ldots, n$$

Assumption:

$$e_i \sim \mathrm{N}(0,\ \sigma^2)$$

Observations:

$$(x_{i1}, y_i) \qquad\qquad i = 1, \ldots, n$$
$$e_i = y_i - \beta_0 - \beta_1 x_{i1}$$

$$\begin{aligned}
\mathrm{L}(\beta_0, \beta_1, \sigma^2) &= f(e_1, \ldots, e_n) \\
&= \prod_{i=1}^{n} \left[\frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{1}{2\sigma^2} {e_i}^2 \right\} \right] \\
&= \prod_{i=1}^{n} \left[\frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{1}{2\sigma^2} (y_i - \beta_0 - \beta_1 x_{i1})^2 \right\} \right] \\
&= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1})^2 \right\} \\
\log \mathrm{L} &= -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1})^2
\end{aligned}$$

$$\begin{cases}
\dfrac{\partial}{\partial \beta_0} \log \mathrm{L} = -\dfrac{1}{2\sigma^2}(2) \displaystyle\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1})(-1) = 0 \\
\dfrac{\partial}{\partial \beta_1} \log \mathrm{L} = -\dfrac{1}{2\sigma^2}(2) \displaystyle\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1})(-x_{i1}) = 0 \\
\dfrac{\partial}{\partial \sigma^2} \log \mathrm{L} = -\dfrac{n}{2\sigma^2} + \dfrac{1}{2(\sigma^2)^2} \displaystyle\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1})^2 = 0
\end{cases}$$

$$\Rightarrow \begin{cases}
\displaystyle\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1}) = 0 \\
\displaystyle\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1}) x_{i1} = 0 \\
\tilde{\sigma}^2 = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1})^2
\end{cases}$$

$$\Rightarrow \begin{cases}
\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}_1 \\
\hat{\beta}_1 = \dfrac{\displaystyle\sum_{i=1}^{n} (x_{i1} - \bar{x}_1)(y_i - \bar{y})}{\displaystyle\sum_{i=1}^{n} (x_{i1} - \bar{x}_1)^2} \\
\tilde{\sigma}^2 = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1})^2
\end{cases}$$

2.2. Estimation of σ^2

Maximum likelihood estimator of σ^2 is

$$\tilde{\sigma}^2 = \frac{\displaystyle\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}$$

where $\hat{y}_i$ is the fitted value of y_i.

For p=1

$$
\begin{aligned}
RSS &= \sum_{i=1}^{n}(y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1}))^2 \\
&= \sum_{i=1}^{n}(y_i - \bar{y} - \hat{\beta}_1(x_{i1} - \bar{x}_1))^2 \\
&= \sum_{i=1}^{n}(y_i - \bar{y})^2 + \hat{\beta}_1^2 \sum_{i=1}^{n}(x_{i1} - \bar{x}_1)^2 - 2\hat{\beta}_1 \sum_{i=1}^{n}(y_i - \bar{y})(x_{i1} - \bar{x}_1) \\
&= S_{yy} + \hat{\beta}_1^2 S_{x_1x_1} - 2\hat{\beta} S_{x_1y} \\
&= S_{yy} + \hat{\beta}_1^2 S_{x_1x_1} - 2\hat{\beta}_1(\hat{\beta}_1 S_{x_1x_1}) \\
&= S_{yy} - \hat{\beta}_1^2 S_{x_1x_1}
\end{aligned}
$$

For any p

$$
\begin{aligned}
\mathrm{RSS} &= \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \\
&= \begin{pmatrix} y_1 - \hat{y}_1 & y_2 - \hat{y}_2 & \dots & y_n - \hat{y}_n \end{pmatrix} \begin{pmatrix} y_1 - \hat{y}_1 \\ y_2 - \hat{y}_2 \\ \vdots \\ y_n - \hat{y}_n \end{pmatrix} \\
&= (\underset{\sim}{Y} - \underset{\sim}{\hat{Y}})^T(\underset{\sim}{Y} - \underset{\sim}{\hat{Y}})
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{RSS} &= (\underset{\sim}{Y} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{Y})^T(\underset{\sim}{Y} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{Y}) \\
&= [(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\underset{\sim}{Y}]^T[(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\underset{\sim}{Y}] \\
&= \underset{\sim}{Y}^T(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)^T(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\underset{\sim}{Y} \\
&= \underset{\sim}{Y}^T[\underset{\sim}{I}^T - (\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)^T](\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\underset{\sim}{Y} \\
&= \underset{\sim}{Y}^T(\underset{\sim}{I} - \underset{\sim}{X}((\underset{\sim}{X}^T\underset{\sim}{X})^{-1})^T\underset{\sim}{X}^T)(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\underset{\sim}{Y} \\
&= \underset{\sim}{Y}^T(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\underset{\sim}{Y} \\
&= \underset{\sim}{Y}^T(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T + \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\underset{\sim}{Y} \\
&= \underset{\sim}{Y}^T(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\underset{\sim}{Y} \\
&= \underset{\sim}{Y}^T\underset{\sim}{Y} - \underset{\sim}{Y}^T\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{Y} \\
&= \sum_{i=1}^{n} y_i^2 - \underset{\sim}{\hat{\beta}}^T\underset{\sim}{X}^T\underset{\sim}{Y} \\
&= \sum_{i=1}^{n} y_i^2 - \begin{pmatrix} \hat{\beta}_0 & \hat{\beta}_1 & \dots & \hat{\beta}_p \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i1}y_i \\ \sum_{i=1}^{n} x_{i2}y_i \\ \vdots \\ \sum_{i=1}^{n} x_{ip}y_i \end{pmatrix} \\
&= \sum_{i=1}^{n} {y_i}^2 - \hat{\beta}_0 \sum_{i=1}^{n} y_i - \hat{\beta}_1 \sum_{i=1}^{n} x_{il}y_i - \ldots - \hat{\beta}_p \sum_{i=1}^{n} x_{ip}y_i
\end{aligned}
$$

$$\text{RSS} = \sum_{i=1}^{n} {y_i}^2 - \hat{\beta}_0 \sum_{i=1}^{n} y_i - \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} y_i - \ldots - \hat{\beta}_p \sum_{i=1}^{n} x_{ip} y_i$$

If β_0 is unknown, write $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}_1 - \ldots - \hat{\beta}_p \bar{x}_p$, then

$$\begin{aligned}
\text{RSS} &= \sum_{i=1}^{n} {y_i}^2 - [(\bar{y} - \hat{\beta}_1 x_1 - \ldots - \hat{\beta}_p \bar{x}_p) n\bar{y} + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} y_i + \ldots + \hat{\beta}_p \sum_{i=1}^{n} x_{ip} y_i] \\
&= \sum_{i=1}^{n} {y_i}^2 - n\bar{y}^2 - \hat{\beta}_1 (\sum_{i=1}^{n} x_{i1} y_i - n\bar{x}_1 \bar{y}) - \ldots - \hat{\beta}_p (\sum_{i=1}^{n} x_{ip} y_i - n\bar{x}_p \bar{y}) \\
&= \text{S}_{yy} - \hat{\beta}_1 \text{S}_{x_1 y} - \ldots - \hat{\beta}_p \text{S}_{x_p y}
\end{aligned}$$

Define $\underset{\sim}{H} = \underset{\sim}{X}(\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T$ (called Hat Matrix). Then,

$$\begin{aligned}
(\underset{\sim}{I} - \underset{\sim}{H})^T &= (\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T)^T \\
&= \underset{\sim}{I} - (\underset{\sim}{X}^T)^T ((\underset{\sim}{X}^T \underset{\sim}{X})^{-1})^T \underset{\sim}{X}^T \\
&= \underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T \\
&= \underset{\sim}{I} - \underset{\sim}{H} \qquad (\underset{\sim}{I} - \underset{\sim}{H} \text{ symmetric}) \\
(\underset{\sim}{I} - \underset{\sim}{H})(\underset{\sim}{I} - \underset{\sim}{H}) &= \underset{\sim}{I} - \underset{\sim}{H} \qquad (\text{ idempotent})
\end{aligned}$$

<u>Example 4: Intercept is known</u> (cont.)

$$\begin{aligned}
\text{RSS} &= \underset{\sim}{Y}^T (\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T) \underset{\sim}{Y} \\
&= \underset{\sim}{Y}^T \underset{\sim}{Y} - \underset{\sim}{Y}^T \underset{\sim}{X}(\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T \underset{\sim}{Y} \\
&= \sum_{i=1}^{n} (y_i - \beta_0)^2 - \hat{\underset{\sim}{\beta}}^T \underset{\sim}{X}^T \underset{\sim}{Y} \\
&= \sum_{i=1}^{n} (y_i - \beta_0)^2 - \begin{pmatrix} \hat{\beta}_1 & \ldots & \hat{\beta}_p \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{n} x_{i1}(y_i - \beta_0) \\ \sum_{i=1}^{n} x_{i2}(y_i - \beta_0) \\ \vdots \\ \sum_{i=1}^{n} x_{ip}(y_i - \beta_0) \end{pmatrix} \\
&= \sum_{i=1}^{n} (y_i - \beta_0)^2 - \hat{\beta}_1 \sum_{i=1}^{n} x_{i1}(y_i - \beta_0) - \ldots - \hat{\beta}_p \sum_{i=1}^{n} x_{ip}(y_i - \beta_0)
\end{aligned}$$

<u>Question</u> Is $\tilde{\sigma}^2$ unbiased?

3. Properties of $\underset{\sim}{\hat{\beta}}$ and Res.S.S.

3.1. Properties of $\underset{\sim}{\hat{\beta}}$

For p=1

Theorem 3.1: Let $Y_1, \ldots, Y_n$ be uncorrelated random variables with $Var(Y_i) = \sigma^2$ for all $i = 1, \ldots, n$. Let $c_1, \ldots, c_n$ and $d_1, \ldots, d_n$ be two sets of constants. Then

$$Cov\left(\sum_{i=1}^{n} c_i Y_i, \sum_{i=1}^{n} d_i Y_i\right) = \left(\sum_{i=1}^{n} c_i d_i\right)\sigma^2$$

$$\begin{aligned}
\mathrm{E}(\hat{\beta}_1) &= \frac{\mathrm{E}\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)(y_i-\bar{y})}{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2} \\
&= \frac{\mathrm{E}(\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)y_i) - (\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)\bar{y})}{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2} \\
&= \frac{\mathrm{E}(\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)y_i)}{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2} \\
&= \frac{\mathrm{E}(\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)(\beta_0+\beta_1 x_{i1}+e_i))}{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2} \\
&= \frac{\mathrm{E}(\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)\beta_0 + \sum_{i=1}^{n}(x_{i1}-\bar{x}_1)\beta_1 x_{i1} + \sum_{i=1}^{n}(x_{i1}-\bar{x}_1)e_i)}{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2} \\
&= \frac{\beta_1\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)x_{i1}}{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2} \\
&= \frac{\beta_1\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)(x_{i1}-\bar{x}_1)}{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2} \\
&= \beta_1 \quad \text{— unbiased}
\end{aligned}$$

$$
\begin{aligned}
\mathrm{Var}(\hat{\beta}_1) &= \mathrm{Var}(\frac{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)(y_i-\bar{y})}{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2}) \\
&= \frac{1}{(\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2)^2}\mathrm{Var}(\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)y_i) \\
&= \frac{1}{(\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2)^2}\mathrm{Var}((x_{11}-\bar{x}_1)y_1+(x_{21}-\bar{x}_1)y_2+\ldots+(x_{n1}-\bar{x}_1)y_n) \\
&= \frac{1}{(\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2)^2}\{(x_{11}-\bar{x}_1)^2\mathrm{Var}(y_1)+(x_{21}-\bar{x}_1)^2\mathrm{Var}(y_2)+\ldots+(x_{n1}-\bar{x}_1)^2\mathrm{Var}(y_n)+0\} \\
&= \frac{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2\mathrm{Var}(y_i)}{(\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2)^2} \\
&= \frac{\sigma^2\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2}{(\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2)^2} \\
&= \frac{\sigma^2}{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2}
\end{aligned}
$$

$$
e_i \sim \mathrm{N}(0,\ \sigma^2) \Rightarrow y_i \sim \mathrm{N}\left(\beta_0+\beta_1 x_{i1},\ \sigma^2\right)
$$

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

$$
\begin{aligned}
\mathrm{E}(\hat{\beta}_0) &= \mathrm{E}(\bar{y} - \hat{\beta}_1 \bar{x}_1) \\
&= \mathrm{E}(\bar{y}) - \mathrm{E}(\hat{\beta}_1 \bar{x}_1) \\
&= \mathrm{E}(\frac{1}{n}\sum_{i=1}^{n} y_i) - \mathrm{E}(\hat{\beta}_1 \bar{x}_1) \\
&= \frac{1}{n}\mathrm{E}(\sum_{i=1}^{n} y_i) - \beta_1 \bar{x}_1 \\
&= \frac{1}{n}\mathrm{E}(\sum_{i=1}^{n} (\beta_0 + \beta_1 x_{i1} + e_i)) - \beta \bar{x}_1 \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}(\beta_0 + \beta_1 x_{i1} + e_i) - \beta_1 \bar{x}_1 \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}(\beta_0 + \beta_1 x_{i1}) - \beta_1 \bar{x}_1 \\
&= \frac{1}{n}(n\beta_0 + \beta_1 \sum_{i=1}^{n} x_{i1}) - \beta_1 \bar{x}_1 \\
&= \beta_0 + \beta_1 \bar{x}_1 - \beta_1 \bar{x}_1 \\
&= \beta_0 \qquad \text{— unbiased}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}(\hat{\beta}_0) &= \mathrm{Var}(\bar{y} - \hat{\beta}_1 \bar{x}_1) \\
&= \mathrm{Var}(\bar{y}) + \mathrm{Var}(-\hat{\beta}_1 \bar{x}_1) + 2\mathrm{Cov}(\bar{y}, -\hat{\beta}_1 \bar{x}_1) \\
&= \frac{\sigma^2}{n} + \frac{\bar{x}_1^2 \sigma^2}{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)^2} + 0 \\
&= \frac{\sigma^2}{n} + \frac{\bar{x}_1^2 \sigma^2}{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)^2}
\end{aligned}
$$

$$
e_i \sim \mathrm{N}(0,\ \sigma^2) \Rightarrow y_i \sim \mathrm{N}\left(\beta_0 + \beta_1 x_{i1},\ \sigma^2\right)
$$

$$
\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}_1 \sim \mathrm{N}\left(\beta_0, \quad \frac{\sigma^2}{n} + \frac{\bar{x}_1^2 \sigma^2}{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)^2}\right)
$$

$$
\begin{aligned}
\mathrm{Cov}(\bar{y},\ \hat{\beta}_1) &= \mathrm{Cov}(\frac{1}{n}\sum_{i=1}^{n} y_i,\quad \frac{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)y_i}{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2}) \\
&= \frac{1}{n\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2}\mathrm{Cov}(\sum_{i=1}^{n} y_i,\quad \sum_{i=1}^{n}(x_{i1}-\bar{x}_1)y_i) \\
&= \frac{1}{n\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2}\mathrm{Cov}(y_1+\ldots+y_n,\quad (x_{11}-\bar{x}_1)y_1+\ldots+(x_{n1}-\bar{x}_1)y_n) \\
&= \frac{1}{n\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2}\{(x_{11}-\bar{x}_1)\mathrm{Var}(y_1)+\ldots+(x_{n1}-\bar{x}_1)\mathrm{Var}(y_n)\} \\
&= \frac{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)\mathrm{Var}(y_i)}{n\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2} \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Cov}(\hat{\beta}_0,\ \hat{\beta}_1) &= \mathrm{Cov}(\bar{y}-\hat{\beta}_1\bar{x}_1,\ \hat{\beta}_1) \\
&= \mathrm{Cov}(\bar{y},\ \hat{\beta}_1)-\mathrm{Cov}(\hat{\beta}_1\bar{x}_1,\ \hat{\beta}_1) \\
&= -\frac{\bar{x}_1\sigma^2}{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2}
\end{aligned}
$$

$$
\begin{pmatrix}\hat{\beta}_0 \\ \hat{\beta}_1\end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix}\beta_0 \\ \beta_1\end{pmatrix},\quad \begin{pmatrix}\sigma^2(\frac{1}{n}+\frac{\bar{x}_1^2}{S_{x_1x_1}}) & -\frac{\bar{x}_1\sigma^2}{S_{x_1x_1}} \\ -\frac{\bar{x}_1\sigma^2}{S_{x_1x_1}} & \frac{\sigma^2}{S_{x_1x_1}}\end{pmatrix}\right)
$$

For any p

Theorem 3.2:

$$
\begin{aligned}
\mathrm{E}(\underset{\sim}{c}\,\underset{\sim}{Z}+\underset{\sim}{d}) &= \underset{\sim}{c}\,\mathrm{E}(\underset{\sim}{Z})\ +\ \underset{\sim}{d} \\
\mathrm{Var}(\underset{\sim}{c}\,\underset{\sim}{Z}+\underset{\sim}{d}) &= \underset{\sim}{c}\,\mathrm{Var}(\underset{\sim}{Z})\,\underset{\sim}{c}^T && c\ ,\ d \text{ are constants} \\
\mathrm{Cov}(\underset{\sim}{c}\,\underset{\sim}{Z},\underset{\sim}{d}\,\underset{\sim}{Z}) &= \underset{\sim}{c}\,\mathrm{Var}(\underset{\sim}{Z})\,\underset{\sim}{d}^T && c\ ,\ d \text{ are constants}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}(\hat{\underset{\sim}{\beta}}) &= \mathrm{E}((\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{Y}) \\
&= (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\mathrm{E}(\underset{\sim}{Y}) \\
&= (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{X}\,\underset{\sim}{\beta} \\
&= \underset{\sim}{I}\,\underset{\sim}{\beta} \\
&= \underset{\sim}{\beta}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}(\hat{\underset{\sim}{\beta}}) &= \mathrm{Var}((\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{Y}) \\
&= (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\ \mathrm{Var}(\underset{\sim}{Y})\ ((\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)^T \\
&= (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\ \mathrm{Var}(\underset{\sim}{Y})\ \underset{\sim}{X}((\underset{\sim}{X}^T\underset{\sim}{X})^{-1})^T \\
&= (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\ \mathrm{Var}(\underset{\sim}{Y})\ \underset{\sim}{X}((\underset{\sim}{X}^T\underset{\sim}{X})^T)^{-1} \\
&= (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\ \mathrm{Var}(\underset{\sim}{Y})\ \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1} \\
&= (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\sigma^2\underset{\sim}{I}\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1} \\
&= \sigma^2(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{I}\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1} \\
&= \sigma^2(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1} \\
&= \sigma^2\underset{\sim}{I}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1} \\
&= \sigma^2(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}
\end{aligned}
$$

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

<u>Example 2: Formula in matrix form for p=1</u> (cont.)

Consider $y_i = \beta_0 + \beta_1 x_{i1} + e_i \qquad i = 1, \ldots, n$

$$
\underset{\sim}{X} = \begin{pmatrix} 1 & x_{11} \\ \vdots & \vdots \\ 1 & X_{n1} \end{pmatrix} \qquad \underset{\sim}{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \qquad \underset{\sim}{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}
$$

Then

$$
\begin{aligned}
\underset{\sim}{X}^T\underset{\sim}{X} &= \begin{pmatrix} 1 & \ldots & 1 \\ x_{11} & \ldots & x_{n1} \end{pmatrix} \begin{pmatrix} 1 & x_{11} \\ \vdots & \vdots \\ 1 & x_{n1} \end{pmatrix} \\
&= \begin{pmatrix} n & \sum_{i=1}^n x_{i1} \\ \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i1}^2 \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
(\underset{\sim}{X}^T\underset{\sim}{X})^{-1} &= \frac{1}{n\sum_{i=1}^n x_{i1}^2 - \left(\sum_{i=1}^n x_{i1}\right)^2} \begin{pmatrix} \sum_{i=1}^n x_{i1}^2 & -\sum_{i=1}^n x_{i1} \\ -\sum_{i=1}^n x_{i1} & n \end{pmatrix} \\
&= \frac{1}{n\left(\sum_{i=1}^n x_{i1}^2 - \frac{\left(\sum_{i=1}^n x_{i1}\right)^2}{n}\right)} \begin{pmatrix} \sum_{i=1}^n x_{i1}^2 & -\sum_{i=1}^n x_{i1} \\ -\sum_{i=1}^n x_{i1} & n \end{pmatrix} \\
&= \frac{1}{nS_{x_1x_1}} \begin{pmatrix} \sum_{i=1}^n x_{i1}^2 & -\sum_{i=1}^n x_{i1} \\ -\sum_{i=1}^n x_{i1} & n \end{pmatrix}
\end{aligned}
$$

$$\mathrm{Var}(\hat{\underset{\sim}{\beta}}) = \begin{pmatrix} \dfrac{\sigma^2 \sum_{i=1}^{n} x_{i1}^2}{nS_{x_1x_1}} & -\dfrac{\sigma^2 \sum_{i=1}^{n} x_{i1}}{nS_{x_1x_1}} \\ -\dfrac{\sigma^2 \sum_{i=1}^{n} x_{i1}}{nS_{x_1x_1}} & \dfrac{\sigma^2}{S_{x_1x_1}} \end{pmatrix}$$

Example 4: Intercept is known (cont.)

$$\begin{aligned} \mathrm{Var}(\hat{\beta}_1) &= \frac{\sum_{i=1}^{n} x_{i1}{}^2 \mathrm{Var}(y_i)}{(\sum_{i=1}^{n} x_{i1}{}^2)^2} \\ &= \frac{\sigma^2}{\sum_{i=1}^{n} x_{i1}{}^2} \end{aligned}$$

3.2. Properties of Res.S.S.

3.2.1. Unbiased estimator of σ^2

For p=1

$$\begin{aligned} \mathrm{E}(S_{yy}) &= (n-1)\sigma^2 + \beta_1^2 Sx_1x_1 \quad \text{and} \\ \mathrm{E}(\hat{\beta}_1^2) &= \mathrm{Var}(\hat{\beta}_1) + (\mathrm{E}(\hat{\beta}_1))^2 \end{aligned}$$

$$\begin{aligned} \mathrm{E}(RSS) &= (n-2)\sigma^2 \\ \Rightarrow \quad \hat{\sigma}^2 &= \frac{RSS}{n-2} \end{aligned}$$

$$\frac{(n-2)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-2} \Rightarrow \frac{RSS}{\sigma^2} \sim \chi^2_{n-2}$$

For any p

Theorem 3.3: Let $\underset{\sim}{Y}$ be a n random vector and let $\mathrm{E}(\underset{\sim}{Y}) = \underset{\sim}{\mu}$, $\mathrm{Cov}(\underset{\sim}{Y}) = \underset{\sim}{\Sigma}$. Then $\mathrm{E}[\underset{\sim}{Y}^T \underset{\sim}{A}\underset{\sim}{Y}] = \mathrm{trace}(\underset{\sim}{A}\underset{\sim}{\Sigma}) + \underset{\sim}{\mu}^T \underset{\sim}{A}\underset{\sim}{\mu}$

$$\begin{aligned} \mathrm{E(RSS)} &= \mathrm{E}(\underset{\sim}{Y}^T(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\underset{\sim}{Y}) \\ &= \mathrm{trace}((\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\underset{\sim}{\Sigma}\ ) + \underset{\sim}{\mu}^T(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\underset{\sim}{\mu} \\ &= \mathrm{trace}(\underset{\sim}{I}\underset{\sim}{\Sigma}\ ) - \mathrm{trace}(\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{\Sigma}\ ) + \underset{\sim}{\beta}^T\underset{\sim}{X}^T(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\underset{\sim}{X}\underset{\sim}{\beta} \\ &= n\sigma^2 - \mathrm{trace}(\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\sigma^2\underset{\sim}{I}) + \underset{\sim}{\beta}^T(\underset{\sim}{X}^T\underset{\sim}{X} - \underset{\sim}{X}^T\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{X})\underset{\sim}{\beta} \\ &= n\sigma^2 - \sigma^2\mathrm{trace}(\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T) + 0 \\ &= n\sigma^2 - \sigma^2\mathrm{trace}((\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{X}) \\ &= n\sigma^2 - \sigma^2\mathrm{trace}(\underset{\sim}{I}_{p'}) \\ &= (n-p')\sigma^2 \end{aligned}$$

$$\hat{\sigma}^2 \text{ (unbiased estimator)} = \frac{RSS}{n-p'}$$

where p' – no. of unknown parameters in the model. If the model has an intercept, then $p' = p+1$ where p is no. of independent variables.

Example 4: Intercept is known (cont.)

$$\text{RSS} = \sum_{i=1}^{n}(y_i - \beta_0)^2 - \hat{\beta}_1 \sum_{i=1}^{n} x_{i1}(y_i - \beta_0) - \ldots - \hat{\beta}_p \sum_{i=1}^{n} x_{ip}(y_i - \beta_0)$$

and

$$\hat{\sigma}^2 = \frac{RSS}{n-p}$$

Example 5: Example in Multiple Linear Regression (cont.)

$$\begin{aligned}
\text{RSS} &= \sum_{i=1}^{n} y_i^2 - \sum_{j=0}^{3} \hat{\beta}_j \sum_{i=1}^{13} x_{ij} y_i \\
&= 11400.15 - (39.1574)(377.5) + (1.0161)(1877.567) + (-1.8616)(2246.6610) + (-0.3433)(3337.78) \\
&= 11400.15 - 11361.47 \\
&= 38.68
\end{aligned}$$

$$\begin{aligned}
\hat{\sigma}^2 &= \frac{38.68}{9} \\
&= 4.298
\end{aligned}$$

3.2.2. Distribution of Res.S.S.

Theorem 3.4: Let the n random vector $\underset{\sim}{Y}$ be distributed $MN(\underset{\sim}{\mu}, \underset{\sim}{I})$. The quadratic form $\underset{\sim}{Y}^T \underset{\sim}{A} \ \underset{\sim}{Y}$ has a non-central chi-square distribution with k d.f. and $\lambda = \underset{\sim}{\mu}^T \underset{\sim}{A} \ \underset{\sim}{\mu}$ (defined as $\chi^2(k, \lambda)$) iff $\underset{\sim}{A}$ is a symmetric idempotent matrix of rank k.

m.g.f of $\chi^2(k, \lambda) = \dfrac{\exp\left(\dfrac{\lambda t}{1-2t}\right)}{(1-2t)^{k/2}}$

mean $= k + \lambda$

variance $= 2(k + 2\lambda)$

Let $\underset{\sim}{Y}^* = \dfrac{\underset{\sim}{Y}}{\sigma} \sim MN(\dfrac{\underset{\sim}{X}\underset{\sim}{\beta}}{\sigma}, \underset{\sim}{I})$

Residual Sum of Squares

$$\underset{\sim}{Y} - \hat{\underset{\sim}{Y}} = (\underset{\sim}{I} - \underset{\sim}{H}) \underset{\sim}{Y}$$

$$\frac{(\underset{\sim}{Y} - \hat{\underset{\sim}{Y}})^T (\underset{\sim}{Y} - \hat{\underset{\sim}{Y}})}{\sigma^2} = \underset{\sim}{Y}^{*T} (\underset{\sim}{I} - \underset{\sim}{H}) \underset{\sim}{Y}^*$$

$$\begin{aligned}
\text{d.f.} &= \text{rank of } \underset{\sim}{I} - \underset{\sim}{H} \\
&= \text{trace}(\underset{\sim}{I} - \underset{\sim}{H}) \\
&= n - p'
\end{aligned}$$

$$\begin{aligned}
\lambda &= \underset{\sim}{\mu}^T \ (\underset{\sim}{I} - \underset{\sim}{H}) \ \underset{\sim}{\mu} \\
&= \frac{1}{\sigma^2} \underset{\sim}{\beta}^T \underset{\sim}{X}^T \ (\underset{\sim}{I} - \underset{\sim}{H}) \ \underset{\sim}{X} \, \underset{\sim}{\beta} \\
&= \frac{1}{\sigma^2} \underset{\sim}{\beta}^T (\underset{\sim}{X}^T \underset{\sim}{X} \ - \ \underset{\sim}{X}^T \underset{\sim}{X} (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T \underset{\sim}{X}) \underset{\sim}{\beta} \\
&= 0
\end{aligned}$$

$$\Rightarrow \frac{\text{Res. S. S.}}{\sigma^2} \ \sim \ \chi^2(n - p')$$

3.3. Independence of $\hat{\underset{\sim}{\beta}}$ and $\hat{\sigma}^2$

For $p = 1$

Write

$$\begin{aligned}
\hat{e}_i &= y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} \\
&= y_i - \sum_{j=1}^{n} (c_j + d_j x_{i1}) y_j \\
&= \sum_{j=1}^{n} \left(\delta_{ij} - (c_j + d_j x_{i1})\right) y_j
\end{aligned}$$

where $c_j = \frac{1}{n} - \frac{(x_{j1} - \bar{x}_1)\bar{x}_1}{S_{x_1 x_1}}$, $d_j = \frac{x_{j1} - \bar{x}_1}{S_{x_1 x_1}}$ and $\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$

$$\begin{aligned}
Cov(\hat{e}_i, \hat{\beta}_0) &= \sigma^2 \sum_{j=1}^{n} \left(\delta_{ij} - (c_j + d_j x_{i1})\right) c_j \\
&= \sigma^2 \left\{ c_i - \sum_{j=1}^{n} c_j \left(c_j + d_j x_{i1}\right) \right\} \\
&= \sigma^2 \left\{ c_i - \sum_{j=1}^{n} c_j^2 - x_{i1} \sum_{j=1}^{n} c_j d_j \right\} \\
&= \sigma^2 \left\{ c_i - \sum_{j=1}^{n} \left(\frac{1}{n} - \frac{(x_{j1} - \bar{x}_1)\bar{x}_1}{S_{x_1 x_1}} \right)^2 - x_{i1} \sum_{j=1}^{n} \left(\frac{1}{n} - \frac{(x_{j1} - \bar{x}_1)\bar{x}_1}{S_{x_1 x_1}} \right) \left(\frac{x_{j1} - \bar{x}_1}{S_{x_1 x_1}} \right) \right\} \\
&= \sigma^2 \left\{ c_i - \left(\frac{1}{n} + \frac{\bar{x}_1^2}{S_{x_1 x_1}} \right) + \frac{x_{i1} \bar{x}_1}{S_{x_1 x_1}} \right\} \\
&= 0
\end{aligned}$$

$$
\begin{aligned}
Cov(\hat{e}_i, \hat{\beta}_1) &= \sigma^2 \sum_{j=1}^{n} (\delta_{ij} - (c_j + d_j x_{i1}))\, d_j \\
&= \sigma^2 \left\{ d_i - \sum_{j=1}^{n} d_j \,(c_j + d_j x_{i1}) \right\} \\
&= \sigma^2 \left\{ d_i - \sum_{j=1}^{n} c_j d_j - x_{i1} \sum_{j=1}^{n} d_j^2 \right\} \\
&= \sigma^2 \left\{ d_i - \sum_{j=1}^{n} \left(\frac{1}{n} - \frac{(x_{j1} - \bar{x}_1)\bar{x}_1}{S_{x_1 x_1}} \right) \left(\frac{x_{j1} - \bar{x}_1}{S_{x_1 x_1}} \right) - x_{j1} \sum_{j=1}^{n} \left(\frac{x_{j1} - \bar{x}_1}{S_{x_1 x_1}} \right)^2 \right\} \\
&= \sigma^2 \left\{ d_i + \frac{\bar{x}_1}{S_{x_1 x_1}} - \frac{x_{i1}}{S_{x_1 x_1}} \right\} \\
&= 0
\end{aligned}
$$

$\Rightarrow$ $\hat{\sigma}^2$ and $(\hat{\beta}_0, \hat{\beta}_1)$ are independent.

4. Confidence Interval & Hypothesis Testing

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

$$\frac{(n-p')\hat{\sigma}^2}{\sigma^2} = \frac{RSS}{\sigma^2} \sim \chi_{(n-p')}$$

Independence of sample mean and sample variance

$\Rightarrow RSS$ and $\hat{\underset{\sim}{\beta}}$ are independent.

$\Rightarrow \hat{\sigma}^2$ and $\hat{\underset{\sim}{\beta}}$ are independent.

4.1. *T* test

For p=1

$$\hat{\beta}_1 \sim \mathrm{N}(\beta_1, \quad \frac{\sigma^2}{S_{x_1x_1}})$$

$$\Rightarrow \frac{\hat{\beta}_1 - \beta_1}{\sigma/\sqrt{S_{x_1x_1}}} \sim \mathrm{N}(0,1)$$

$$\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}/\sqrt{S_{x_1x_1}}} \sim t_{(n-2)}$$

$H_0 : \beta_1 = \beta_{10}$

$$\hat{\beta}_1 \sim \mathrm{N}(\beta_{10}, \frac{\sigma^2}{S_{x_1x_1}}), \qquad \frac{\hat{\beta}_1 - \beta_{10}}{\hat{\sigma}/\sqrt{S_{x_1x_1}}} \sim t_{(n-2)}$$

$$\begin{aligned}
&Pr(-t_{\alpha/2,(n-2)} \leq T \leq t_{\alpha/2,(n-2)}) = 1-\alpha \\
&\Rightarrow Pr(-t_{\alpha/2,(n-2)} \leq \frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}/\sqrt{S_{x_1x_1}}} \leq t_{\alpha/2,(n-2)}) = 1-\alpha \\
&\Rightarrow Pr(-\hat{\beta}_1 - t_{\alpha/2,(n-2)}\frac{\hat{\sigma}}{\sqrt{S_{x_1x_1}}} \leq -\beta_1 \leq -\hat{\beta}_1 + t_{\alpha/2,(n-2)}\frac{\hat{\sigma}}{\sqrt{S_{x_1x_1}}}) = 1-\alpha \\
&\Rightarrow Pr(\hat{\beta}_1 - t_{\alpha/2,(n-2)}\frac{\hat{\sigma}}{\sqrt{S_{x_1x_1}}} \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2(n-2)}\frac{\hat{\sigma}}{\sqrt{S_{x_1x_1}}}) = 1-\alpha
\end{aligned}$$

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

$$\hat{\beta}_0 \sim \mathrm{N}(\beta_{00}, \sigma^2(\frac{1}{n} + \frac{\bar{x}_1^2}{S_{x_1x_1}}))$$

$$\frac{\hat{\beta}_0 - \beta_{00}}{\sigma\sqrt{\frac{1}{n} + \frac{\bar{x}_1^2}{S_{x_1x_1}}}} \sim \mathrm{N}(0,1) \qquad \frac{\hat{\beta}_0 - \beta_{00}}{\hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{x}_1^2}{S_{x_1x_1}}}} \sim t_{n-2}$$

C.I. of β_0 :

$$(\hat{\beta}_0 - t_{\alpha/2,(n-2)}\hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{x}_1^2}{S_{x_1x_1}}}, \quad \hat{\beta}_0 + t_{\alpha/2,(n-2)}\hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{x}_1^2}{S_{x_1x_1}}})$$

Example 1: Example in Simple Linear Regression (cont.)

Summary statistics:

$$\begin{aligned}
\sum_{i=1}^{9} y_i^2 &= 1036.65 \\
S_{xx} &= 115.11 - \frac{(30.3)^2}{9} = 13.10 \\
S_{yy} &= 1036.65 - \frac{(91.1)^2}{9} = 114.52 \\
S_{xy} &= 345.09 - \frac{(30.3)(91.9)}{9} = 38.39 \\
\hat{\beta}_1 &= 2.9303 \\
\hat{\sigma}^2 &= \frac{S_{yy} - \hat{\beta}_1 S_{xy}}{n-2} \\
&= \frac{114.52 - (2.9303)(38.39)}{n-2} = 0.2894 \\
\Rightarrow \quad \hat{\sigma} &= 0.538 \\
t_{0.05/2,7} &= 2.365
\end{aligned}$$

$H_0 : \beta_1 = 2.5, \qquad H_1 : \beta_1 > 2.5$

$$\begin{aligned}
t &= \frac{2.9303 - 2.5}{0.538/\sqrt{13.10}} \\
&= 2.8945 \\
&< 2.998 = t_{0.01,7}
\end{aligned}$$

Can't reject H_0 $\quad \Rightarrow \beta_1$ does not significantly differ from 2.5

95% C.I. of β_1:

$$(2.9305 - \frac{2.365 * 0.538}{\sqrt{13.10}}, \quad 2.9305 - \frac{2.365 * 0.538}{\sqrt{13.10}})$$

$$\Rightarrow 2.579 < \beta_1 < 3.282$$

$H_0 : \beta_0 = 0$

$$\frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}\sqrt{\dfrac{1}{n} + \dfrac{\bar{x}_1^2}{S_{x_1 x_1}}}} = \frac{0.2568 - 0}{0.538\sqrt{\frac{1}{9} + \frac{3.3667^2}{13.10}}} = 0.4831 < 1 \text{ (can't reject } H_0)$$

95% C.I. of β_0 :

$$(0.2568 - (2.365)0.538\sqrt{\frac{1}{9} + \frac{3.3667^2}{13.10}}, \quad 0.2568 + (2.365)(0.538)\sqrt{\frac{1}{9} + \frac{3.3667^2}{13.10}})$$

$$\Rightarrow -1.0005 < \beta_0 < 1.514$$

Example 4: Intercept is known (cont.)

$H_0 : \beta_1 = \beta_{10}$

$$\hat{\beta}_1 \sim \mathrm{N}(\beta_1, \frac{\sigma^2}{\sum_{i=1}^{n} x_i^2})$$

$$\Rightarrow \frac{\hat{\beta}_1 - \beta_{10}}{\sigma \Big/ \sqrt{\sum_{i=1}^{n} {x_i}^2}} \quad \sim \quad \mathrm{N}(0,1)$$

$$\Rightarrow \frac{\hat{\beta}_1 - \beta_{10}}{\hat{\sigma} \Big/ \sqrt{\sum_{i=1}^{n} {x_i}^2}} \quad \sim \quad t_{(n-1)}$$

For any p

$H_0 : \beta_j = \beta_{j0} \qquad$ for $j = 0, \ldots, p$

$$t = \frac{\hat{\beta}_j - \beta_{j0}}{\text{s.e. of } (\hat{\beta}_j)} = \frac{\hat{\beta}_j - \beta_{j0}}{\hat{\sigma}\sqrt{c^{jj}}} \sim t_{(n-p')}$$

Reject H_0 if $|t| > t_{\alpha/2,(n-p')}$

Example 5: Example in Multiple Linear Regression (cont.)

$H_0 : \beta_2 = -2.5, \qquad H_1 : \beta_2 > -2.5$

$$\begin{aligned} t &= \frac{\hat{\beta}_2 - \beta_{20}}{\hat{\sigma}\sqrt{c^{22}}} \\ &= \frac{(-1.8616) - (-2.5)}{2.073\sqrt{0.0166}} \\ &= 2.391 \\ &> 1.833 \end{aligned}$$

Reject H_0.

Example 1: Example in Simple Linear Regression (cont.)

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

Point estimate of : $5\beta_0 + \beta_1 = 5\hat{\beta}_0 + \hat{\beta}_1$

$$\begin{aligned} \mathrm{Var}(5\hat{\beta}_0 + \hat{\beta}_1) &= 25\mathrm{Var}(\hat{\beta}_0) + \mathrm{Var}(\hat{\beta}_1) + 2 * 5\mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_1) \\ &= \hat{\sigma}^2 \left\{ 25 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{x_1 x_1}} \right) + \frac{1}{S_{x_1 x_1}} - \frac{10\bar{x}}{S_{x_1 x_1}} \right\} \\ &= \hat{\sigma}^2 \left(\frac{25}{n} + \frac{(5\bar{x} - 1)^2}{S_{x_1 x_1}} \right) \end{aligned}$$

$$\frac{(5\hat{\beta}_0 + \hat{\beta}_1) - (5\beta_0 + \beta_1)}{\hat{\sigma}\sqrt{\frac{25}{n} + \frac{(5\bar{x}-1)^2}{S_{x_1 x_1}}}} \sim t_{n-2}$$

$$t_{obs} = \frac{(5 * 0.2568 + 2.9303) - 2}{0.538\sqrt{\frac{25}{9} + \frac{(5 * 3.3667 - 1)^2}{13.10}}}$$

- For both one-sided and two-sided alternatives
- For ONE linear combination of regression coefficients (including intercept) only

Simultaneous C.I.

Let $Pr(A)$ is the Prob. that the confident interval on α does not cover the true parameter

Let $Pr(B)$ is the Prob. that the confident interval on β does not cover the true parameter

$Pr(\bar{A} \cap \bar{B})$ is the Prob. that the two confident intervals simultaneously cover their respective true parameter

$$Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B) \text{ and } Pr(\bar{A} \cap \bar{B}) = 1 - Pr(A \cup B)$$

$$\begin{aligned} Pr(\bar{A} \cap \bar{B}) &= 1 - [Pr(A) + Pr(B) - Pr(A \cap B)] \\ &= 1 - [Pr(A) + Pr(B)] + Pr(A \cap B) \\ &\geq 1 - [Pr(A) + Pr(B)] \\ &\geq 1 - 2\alpha \qquad \text{(Bonferroni inequality)} \end{aligned}$$

$$1 - 2\alpha' = 1 - \alpha \Rightarrow \alpha' = \frac{\alpha}{2}$$

So we have at least $1 - \alpha$ confident that

$$\hat{\beta}_1 - t_{\frac{\alpha}{4}} \frac{\hat{\sigma}}{S_{x_1 x_1}} < \beta_1 < \hat{\beta}_1 + t_{\frac{\alpha}{4}} \frac{\hat{\sigma}}{S_{x_1 x_1}}$$

and

$$\hat{\beta}_0 - t_{\frac{\alpha}{4}} \hat{\sigma} \sqrt{\frac{1}{n} + \frac{\bar{x}_1^2}{S_{x_1 x_1}}} < \beta_0 < \hat{\beta}_0 + t_{\frac{\alpha}{4}} \hat{\sigma} \sqrt{\frac{1}{n} + \frac{\bar{x}_1^2}{S_{x_1 x_1}}}$$

m simultaneous C.I.: We have at least $(1 - \alpha)$ confident that ......

$\Rightarrow$ we choose $t_{\frac{\alpha}{2m}}$

4.2. *F* test

4.2.1. All regression coefficient equal to zero

A. Partitioning total variability

For the model of $y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + e_i$ for $i = 1, \ldots, p$,

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

For

$$\begin{aligned}
\sum_{i=1}^{n}(y_i-\hat{y}_i)(\hat{y}_i-\bar{y}) &= \sum_{i=1}^{n}(y_i-\hat{y}_i)\hat{y}_i - \sum_{i=1}^{n}(y_i-\hat{y}_i)\bar{y} \\
&= \sum_{i=1}^{n}(y_i-\hat{y}_i)(\hat{\beta}_0+\hat{\beta}_1 x_i) \\
&= \hat{\beta}_0\sum_{i=1}^{n}(y_i-\hat{y}_i) + \hat{\beta}_1\sum_{i=1}^{n}(y_i-\hat{y}_i)x_{i1} \\
&= \hat{\beta}_1\sum_{i=1}^{n}[y_i-(\bar{y}+\hat{\beta}_1(x_{i1}-\bar{x}_1))](x_{i1}-\bar{x}_1) \\
&= \hat{\beta}_1\sum_{i=1}^{n}[(y_i-\bar{y})(x_{i1}-\bar{x}_1)-\hat{\beta}_1(x_{i1}-\bar{x}_1)^2] \\
&= 0
\end{aligned}$$

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

$$\begin{array}{ccccc}
\begin{pmatrix}\text{Total variability}\\ \text{in response}\end{pmatrix} & = & \begin{pmatrix}\text{Variability}\\ \text{explained by model}\end{pmatrix} & + & \begin{pmatrix}\text{Unexplained}\\ \text{variability}\end{pmatrix} \\
\textit{Total S.S.} & = & \textit{Reg. S.S.} & + & \textit{Residual S.S.} \\
\sum_{i=1}^{n}(y_i-\bar{y})^2 & = & \sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2 & + & \sum_{i=1}^{n}(y_i-\hat{y}_i)^2
\end{array}$$

B. Distributions

Residual Sum of Squares

$$\underset{\sim}{Y}-\hat{\underset{\sim}{Y}} = (\underset{\sim}{I}-\underset{\sim}{H})\,\underset{\sim}{Y}$$

$$\frac{(\underset{\sim}{Y}-\hat{\underset{\sim}{Y}})^T\,(\underset{\sim}{Y}-\hat{\underset{\sim}{Y}})}{\sigma^2} = \underset{\sim}{Y}^{*T}\,(\underset{\sim}{I}-\underset{\sim}{H})\,\underset{\sim}{Y}^{*}$$

$$\Rightarrow \frac{\text{Res. S. S.}}{\sigma^2} \sim \chi^2(n-p')$$

Total Sum of Squares

$$\underset{\sim}{Y}-\bar{\underset{\sim}{Y}} = (\underset{\sim}{I}-\frac{1}{n}\underset{\sim}{J})\,\underset{\sim}{Y}$$

$$\begin{aligned}
\frac{(\underset{\sim}{Y}-\bar{\underset{\sim}{Y}})^T\,(\underset{\sim}{Y}-\bar{\underset{\sim}{Y}})}{\sigma^2} &= \underset{\sim}{Y}^{*T}\,(\underset{\sim}{I}-\frac{1}{n}\underset{\sim}{J})^T\,(\underset{\sim}{I}-\frac{1}{n}\underset{\sim}{J})\,\underset{\sim}{Y}^{*} \\
&= \underset{\sim}{Y}^{*T}\,(\underset{\sim}{I}-\frac{1}{n}\underset{\sim}{J})\,\underset{\sim}{Y}^{*}
\end{aligned}$$

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

Regression Sum of Squares

$$\hat{\underset{\sim}{Y}} = \underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T\,\underset{\sim}{Y} \quad \text{where} \quad \underset{\sim}{X} = \begin{pmatrix} 1 & x'_{11} & \dots & x'_{1p} \\ 1 & x'_{21} & \dots & x'_{2p} \\ \vdots & \vdots & \dots & \vdots \\ 1 & x'_{n1} & \dots & x'_{np} \end{pmatrix}$$

$$\bar{\underset{\sim}{Y}} = \frac{1}{n}\underset{\sim}{J}\,\underset{\sim}{Y}$$

$$\begin{aligned} \text{Reg S.S.} &= (\hat{\underset{\sim}{Y}}-\bar{\underset{\sim}{Y}})^T\,(\hat{\underset{\sim}{Y}}-\bar{\underset{\sim}{Y}}) \\ &= \underset{\sim}{Y}^T\,(\underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T - \frac{1}{n}\underset{\sim}{J})^T\,(\underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T - \frac{1}{n}\underset{\sim}{J})\,\underset{\sim}{Y} \end{aligned}$$

$$\begin{aligned} & (\underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T - \frac{1}{n}\underset{\sim}{J})^T\,(\underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T - \frac{1}{n}\underset{\sim}{J}) \\ = {} & (\underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T - \frac{1}{n}\underset{\sim}{J})\,(\underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T - \frac{1}{n}\underset{\sim}{J}) \\ = {} & \underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T - \frac{1}{n}\underset{\sim}{J}\,\underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T - \frac{1}{n}\underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T\,\underset{\sim}{J} + \frac{1}{n^2}\underset{\sim}{J}\underset{\sim}{J} \\ = {} & \underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T - \frac{1}{n}\underset{\sim}{J} \end{aligned}$$

$$\Rightarrow \frac{\text{Reg. S. S.}}{\sigma^2} \sim \chi^2\left(p,\ \frac{1}{\sigma^2}\sum_{i=1}^{p}\sum_{j=1}^{p}\beta_i\beta_j S_{x_i,x_j}\right)$$

Theorem 4.1: Let the n random vector $\underset{\sim}{Y}$ be distributed $MN(\underset{\sim}{\mu}, \underset{\sim}{\Sigma})$, where $\underset{\sim}{\Sigma}$ has rank n. If $\underset{\sim}{A}^T\,\underset{\sim}{\Sigma}\,\underset{\sim}{B}=0$, then two quadratic forms of $\underset{\sim}{Y}^T\,\underset{\sim}{A}\underset{\sim}{Y}$ and $\underset{\sim}{Y}^T\,\underset{\sim}{B}\underset{\sim}{Y}$ are independent.

Theorem 4.2: Let the two independent random variables U_1 and U_2 be distributed as $U_1 \sim \chi^2(n_1,\lambda_1)$ and $U_2 \sim \chi^2(n_2,\lambda_2)$, respectively, then $U = U_1 + U_2 \sim \chi^2(n,\lambda)$ where $n = n_1 + n_2$ and $\lambda = \lambda_1 + \lambda_2$.

$$\begin{aligned} \text{Reg.S.S.} &= (\hat{\underset{\sim}{Y}}-\bar{\underset{\sim}{Y}})^T\,(\hat{\underset{\sim}{Y}}-\bar{\underset{\sim}{Y}}) \\ &= \underset{\sim}{Y}^T\,(\underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T - \frac{1}{n}\underset{\sim}{J})\,\underset{\sim}{Y} \\ \text{Res.S.S.} &= (\underset{\sim}{Y}-\hat{\underset{\sim}{Y}})^T\,(\underset{\sim}{Y}-\hat{\underset{\sim}{Y}}) \\ &= \underset{\sim}{Y}^T\,(\underset{\sim}{I} - \underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T)\,\underset{\sim}{Y}^* \end{aligned}$$

$$\begin{aligned} & (\underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T - \frac{1}{n}\underset{\sim}{J})\,(\underset{\sim}{I} - \underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T) \\ = {} & \underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T - \frac{1}{n}\underset{\sim}{J} - \underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T\underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}^T + \frac{1}{n}\underset{\sim}{J}\underset{\sim}{X}\,(\underset{\sim}{X}^T\,\underset{\sim}{X})^{-1}\underset{\sim}{X}^T \\ = {} & \underset{\sim}{0} \end{aligned}$$

$$\begin{array}{ccccc} \mathit{Total\,S.S.} & = & \mathit{Reg.\,S.S.} & + & \mathit{Res.\,S.S.} \\ \displaystyle\sum_{i=1}^{n}(y_i-\bar{y})^2 & = & \displaystyle\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2 & + & \displaystyle\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \\ \sim \sigma^2\chi^2(n-1,\lambda) & & & & \sim \sigma^2\chi^2(n-p') \end{array}$$

$Reg.\,S.S.$ and $Res.\,S.S.$ are independent $\Rightarrow Reg.\,S.S. \sim \sigma^2\chi^2(p,\lambda)$, where $\lambda = \frac{1}{\sigma^2}\sum_{i=1}^{p}\sum_{j=1}^{p}\beta_i\beta_j S_{x_i,x_j}$.

C. Test statistic

Under $H_0 : \beta_1 = \ldots = \beta_p = 0$, the test statistic

$$\mathrm{F} = \frac{RegS.S./p}{RSS/(n-(p+1))} = \frac{Reg\,M.S.}{\hat{\sigma}^2} \sim \mathrm{F}(p,\, n-p')$$

Under alternative hypothesis,

$$\begin{aligned}
\mathrm{E}(SS_{reg}) &= p\,\sigma^2 \;+\; \sum_{i=1}^{p}\sum_{j=1}^{p}\beta_i\beta_j S_{x_i,x_j} \\
\mathrm{E}(MS_{reg}) &= \sigma^2 \;+\; \frac{1}{p}\sum_{i=1}^{p}\sum_{j=1}^{p}\beta_i\beta_j S_{x_i,x_j} \\
\mathrm{E}(\hat{\sigma}^2) &= \sigma^2 \\
\mathrm{E}(\mathrm{F}) &\approx \frac{\mathrm{E}(MS_{Reg})}{\mathrm{E}(\hat{\sigma}^2)} \\
&= 1 + \frac{1}{p\sigma^2}\sum_{i=1}^{p}\sum_{j=1}^{p}\beta_i\beta_j S_{x_i,x_j}
\end{aligned}$$

For $p = 1$

$$\mathrm{E}(\mathrm{F}) \;\approx\; 1 + \frac{\beta_1^2 S_{x_1x_1}}{\sigma^2} > 1$$

ANOVA table for $H_0 : \beta_1 = \beta_2 = \ldots = \beta_p = 0$

Source	Sum of squares (S.S.)	d.f.	Mean Squares (M.S.)	F
Regression	Reg.S.S $= \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$	p	SS_{reg}/p	$F = \frac{SS_{reg}/p}{SS_{res}/(n-p')} = \frac{MS_{reg}}{\hat{\sigma}^2}$
Residual	Res. S.S. $= \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$	$n-p'$	$SS_{res}/(n-p')$	
Total	Total S.S. $= \sum_{i=1}^{n}(y_i - \bar{y})^2$	$n-1$		

Example 1: Example in Simple Linear Regression (cont.)

$H_0 : \beta_1 = 0$

Reg S.S. $= \hat{\beta}_1^2 S_{x_1x_1} = 2.9303^2 \times 13.1$

Or, Reg S.S. $= \hat{\beta}_1 S_{x_1y} = 2.9303 \times 38.39$

$S_{yy} = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 = 114.52$

Source	Sum of squares (S.S.)	d.f.	Mean Squares (M.S.)	F
Regression	112.48	1	$112.48/1 = 112.48$	$F = \frac{112.48/1}{2.03188/7} = 387.52$
Residual	2.03188	9 - 2	$2.03189/7 = 0.29027$	
Total	114.52	9 - 1		

Critic value $= F_{\alpha,1,7}$

Reject H_0, $\Rightarrow \ \beta_1 \neq 0$

Example 5: Example in Multiple Linear Regression (cont.)

$H_0 : \beta_1 = \beta_2 = \beta_3 = 0$

$$\begin{aligned} \text{Res.S.S.} &= \mathrm{S}_{yy} - \hat{\beta}_1 \mathrm{S}_{x_1 y} - \ldots - \hat{\beta}_p \mathrm{S}_{x_p y} \\ \Rightarrow \text{Reg.S.S.} &= \hat{\beta}_1 \mathrm{S}_{x_1 y} + \ldots + \hat{\beta}_p \mathrm{S}_{x_p y} \end{aligned}$$

Or

$$\begin{aligned} \text{Res.S.S.} &= \sum_{i=1}^{n} {y_i}^2 - \hat{\beta}_0 \sum_{i=1}^{n} y_i - \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} y_i - \ldots - \hat{\beta}_p \sum_{i=1}^{n} x_{ip} y_i \\ \Rightarrow \text{Reg.S.S.} &= \hat{\beta}_0 \sum_{i=1}^{n} y_i + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} y_i + \ldots + \hat{\beta}_p \sum_{i=1}^{n} x_{ip} y_i - \frac{(\sum_{i=1}^{n} y_i)^2}{n} \end{aligned}$$

Source	Sum of squares (S.S.)	d.f.	Mean Squares (M.S.)	F
Regression	399.45437	12-9=3	399.45437/3 = 133.15146	$F = \frac{133.15146}{4.29738} = 30.98$
Residual	38.6764	9	38.6764/9 = 4.29738	
Total	438.13	12		

Another view

Under $H_0 : \beta_1 = \ldots = \beta_p = 0$, find β_0 s.t. $\sum_{i=1}^{n} (y_i - \beta_0)^2$ is minimized.

$$\begin{aligned} \frac{\partial \sum_{i=1}^{n} (y_i - \beta_0)^2}{\partial \beta_0} &= 2 \sum_{i=1}^{n} (y_i - \beta_0)(-1) \\ &= 0 \\ &\Rightarrow \hat{\beta}_0 = \bar{y} \\ \text{fitted value} &= \hat{\beta}_0 = \bar{y} \qquad \forall\, i \end{aligned}$$

$$\Rightarrow \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} (y_i - \hat{y}|_{\mathrm{H_0}})^2 = \text{Res S.S. under } \mathrm{H_0}$$

$$\Rightarrow \mathit{Reg.\, S.S.} = \mathit{Res\, S.S.}|_{\mathrm{H_0}} - \mathit{Res\, S.S.}$$

Example 4: Intercept is known (cont.)

For $p = 1$

Under the model :

$$y_i' = \beta_1 x_{i1} + e_i \qquad \text{where } y_i' = y_i - \beta_0 \quad \text{and} \quad \hat{y}_i' = \hat{y}_i - \beta_0$$

RSS for the model :

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

Regression S.S. for the model :

$$\sum_{i=1}^{n}(\hat{y}_i' - \widetilde{y}_i')^2 = \sum_{i=1}^{n}(\hat{y}_i')^2 = \sum_{i=1}^{n}(\hat{y}_i - \beta_0)^2$$

Under $H_0 : \beta_1 = 0$,

$$\begin{aligned} \Rightarrow y_i' &= e_i & i &= 1, \ldots, n \\ \widetilde{y}_i' &= 0 & i &= 1, \ldots, n \end{aligned}$$

Residual S.S under $H_0 : \beta_1 = 0$

$$\sum_{i=1}^{n}(y_i' - \widetilde{y}_i')^2 = \sum_{i=1}^{n}(y_i - \beta_0)^2$$

RSS for the model of $y_i = \beta_0 + e_i$		Reg.S.S. for the model		RSS for the model
$\sum_{i=1}^{n}(y_i - \beta_0)^2$	$=$	$\sum_{i=1}^{n}(\hat{y}_i - \beta_0)^2$	$+$	$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

<u>For any p</u>

ANOVA table for $H_0 : \beta_1 = \beta_2 = \ldots = \beta_p = 0$

Source	Sum of squares (S.S.)	d.f.	Mean Squares (M.S.)	F
Regression	$\hat{\underset{\sim}{\beta}}^T \underset{\sim}{X}^T \underset{\sim}{Y}$	p	$\hat{\underset{\sim}{\beta}}^T \underset{\sim}{X}^T \underset{\sim}{Y}/p$	$F = \dfrac{\hat{\underset{\sim}{\beta}}^T \underset{\sim}{X}^T \underset{\sim}{Y}/p}{\underset{\sim}{Y}^T \underset{\sim}{Y} - \hat{\underset{\sim}{\beta}}^T \underset{\sim}{X}^T \underset{\sim}{Y}/(n-p)}$
Residual	$\underset{\sim}{Y}^T \underset{\sim}{Y} - \hat{\underset{\sim}{\beta}}^T \underset{\sim}{X}^T \underset{\sim}{Y}$	$n-p$	$\dfrac{\underset{\sim}{Y}^T \underset{\sim}{Y} - \hat{\underset{\sim}{\beta}}^T \underset{\sim}{X}^T \underset{\sim}{Y}}{n-p}$	
Total	$\underset{\sim}{Y}^T \underset{\sim}{Y}$	n		

where

$$\underset{\sim}{Y} = \begin{pmatrix} y_1 - \beta_0 \\ y_2 - \beta_0 \\ \vdots \\ y_n - \beta_0 \end{pmatrix}$$

<u>4.2.2. Subset of regression coefficients</u>

<u>Explanation</u>

$$\begin{array}{ccccc} Reg\,S.S.|_F & = & Reg.\,S.S.|_R & + & Increase\ in\ Reg.S.S. \\ \sim \sigma^2\chi^2(p, \lambda) & & \sim \sigma^2\chi^2(p-r, \lambda_1) & & \end{array}$$

where $\lambda = \frac{1}{\sigma^2}\underset{\sim}{\beta}_1^T \underset{\sim}{X}_1^T \underset{\sim}{X}_1 \underset{\sim}{\beta}_1$ and $\lambda_1 = \frac{1}{\sigma^2}\underset{\sim}{\beta}_s^T \underset{\sim}{X}_s^T \underset{\sim}{X}_s \underset{\sim}{\beta}_s$.

By Theorem 4.2, $Reg.\ S.S.|_R$ and Increase in $Reg.S.S.$ are independent $\Rightarrow$ $Increase\ in\ Reg.S.S. \sim \sigma^2\chi^2(r, \lambda_2)$, where $\lambda_2 = \frac{1}{\sigma^2}\left(\underset{\sim}{\beta}_1^T \underset{\sim}{X}_1^T \underset{\sim}{X}_1 \underset{\sim}{\beta}_1 \ - \ \underset{\sim}{\beta}_s^T \underset{\sim}{X}_s^T \underset{\sim}{X}_s \underset{\sim}{\beta}_s\right)$

Examples

1. $p = 4$ $\qquad \beta_0,\ \beta_1,\ \beta_2,\ \beta_3,\ \beta_4$

   $H_0 : \beta_3 = \beta_4 = 0$

   Under H_0 :

   Intercept $+ x_1 + x_2$ $\qquad$ $Reg\, S.S.$(reduced model) $= \mathrm{R}(\beta_1, \beta_2|\beta_0)$

   Full model :

   Intercept $+ x_1 + x_2 + x_3 + x_4$ $\qquad$ $Reg\, S.S.$(full model) $= \mathrm{R}(\beta_1, \beta_2, \beta_3, \beta_4|\beta_0)$

$$\begin{aligned}
\text{Increasing in } Reg\, S.S. &= \mathrm{R}(\beta_3, \beta_4|\beta_2, \beta_1, \beta_0) \\
&= \mathrm{R}(\beta_1, \beta_2, \beta_3, \beta_4|\beta_0) - \mathrm{R}(\beta_1, \beta_2|\beta_0) \\
&= Reg\, S.S.(\text{full}) - Reg\, S.S.(\text{reduced}) \\
&= Res\, S.S.(\text{reduced}) - Res\, S.S.(\text{full}) \\
&= d.f._{\text{reduced}} \times \hat{\sigma}^2_{\text{reduced}} - d.f._{\text{full}} \times \hat{\sigma}^2_{\text{full}} \\
&= \underset{\sim}{Y}^T(\underset{\sim}{I} - \underset{\sim}{X}_1(\underset{\sim}{X}_1{}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}_1{}^T)\underset{\sim}{Y} - \underset{\sim}{Y}^T(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\underset{\sim}{Y} \\
&= \underset{\sim}{Y}^T(\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T - \underset{\sim}{X}_1(\underset{\sim}{X}_1{}^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1{}^T)\underset{\sim}{Y} \\
&\quad X_1 : \text{ reduced design matrix}
\end{aligned}$$

   Reject H_0 if $\mathrm{R}(\beta_3, \beta_4|\beta_2, \beta_1, \beta_0)$ is significantly large.

$$\mathrm{F} = \frac{\mathrm{R}(\beta_3, \beta_4|\beta_2, \beta_1, \beta_0)/2}{\hat{\sigma}^2}$$

   Reject H_0 if $\mathrm{F} > \mathrm{F}_\alpha(2, n - p')$

2. $p = 4$ $\qquad \beta_0,\ \beta_1,\ \beta_2,\ \beta_3,\ \beta_4$

   $H_0 : \beta_3 = \beta_4$

   Full model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_{i4} x_{i4} + e_i$$

   Reduced model (under H_0) :

$$\begin{aligned}
y_i &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + e_i \\
&= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 (x_{i3} + x_{i4}) + e_i
\end{aligned}$$

$$\mathrm{F} = \frac{[Res\, S.S._{(\text{reduced})} - Res\, S.S._{(\text{full})}]/[d.f._{(\text{reduced})} - d.f._{(\text{full})}]}{\hat{\sigma}^2}$$

   Reject H_0 if $\mathrm{F} > \mathrm{F}_\alpha(1, n - p')$

3. <u>Example 5: Example in Multiple Linear Regression</u> (cont.)

$H_0 : \beta_3 = 0$

$$
\begin{aligned}
t &= \frac{\hat{\beta}_3 - 0}{\hat{\sigma}\sqrt{c_{33}}} \\
&= \frac{-0.3433}{2.073\sqrt{0.0886}} \\
&= -0.556
\end{aligned}
$$

$$
\begin{pmatrix} 13 & 59.43 & 81.82 \\ 59.43 & 394.7255 & 360.6621 \\ 81.82 & 360.6621 & 576.7264 \end{pmatrix} \begin{pmatrix} \tilde{\beta}_0 \\ \tilde{\beta}_1 \\ \tilde{\beta}_2 \end{pmatrix} = \begin{pmatrix} 377.5 \\ 1877.576 \\ 2246.661 \end{pmatrix}
$$

$\Rightarrow \tilde{\beta}_0 = 36.094,\ \tilde{\beta}_1 = 1.031,\ \hat{\beta}_2 = -1.870.$

$$
\begin{aligned}
Reg\ S.S.(\beta_1, \beta_2|\beta_0) &= \sum_{j=0}^{2} \tilde{\beta}_j \sum_{i=1}^{13} x_{ji} y_i - \frac{(\sum_{i=1}^{13} y_i)^2}{13} \\
&= 398.12 \\
Reg\ S.S.(\beta_1, \beta_2, \beta_3|\beta_0) &= 399.45
\end{aligned}
$$

$$
\begin{aligned}
\text{Increase in } Reg\ S.S. &= \mathrm{R}(\beta_3|\beta_1, \beta_2, \beta_0) \\
&= Reg\ S.S.(\beta_1, \beta_2, \beta_3|\beta_0) - Reg\ S.S.(\beta_1, \beta_2|\beta_0) \\
&= 399.45 - 398.12 = 1.33 \\
F &= \frac{Reg\ S.S.(\beta_3|\beta_1, \beta_2, \beta_0)}{\hat{\sigma}^2} = \frac{1.33}{4.298} = 0.309
\end{aligned}
$$

4. $p = 4$ $\qquad \beta_0,\ \beta_1,\ \beta_2,\ \beta_3,\ \beta_4$

$H_0 : \beta_3 = 2$

For t-test,

$$
\begin{aligned}
&\underset{\sim}{\beta} \sim \mathrm{N}(\underset{\sim}{\beta}, \quad (\underset{\sim}{X}^T \underset{\sim}{X})^{-1}\sigma^2) \\
\Rightarrow \quad & t = \frac{\hat{\beta}_j - \beta_{j0}}{\hat{\sigma}\sqrt{c_{jj}}}, \qquad j = 3
\end{aligned}
$$

For F test, Full :

$$
\begin{aligned}
y_i &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + e_i \\
Total\,SS(\text{full}) &= \sum_{i=1}^{n} (y_i - \bar{y})^2
\end{aligned}
$$

Reduced :

$$
\begin{aligned}
\underbrace{y_i - 2x_{i3}}_{y_i'} &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_4 x_{i4} + e_i \\
Total\,S.S.(\text{reduced}) &= \sum_{i=1}^{n} (y_i' - \bar{y}')^2
\end{aligned}
$$

- For two-sided alternative

- Total S.S. for full model must be equal to Total S.S. for reduced model (i.e., the model under H_0)

4.2.3. General linear hypothesis

Theorem 4.3: Let $\underset{\sim}{X}$ be distributed $MN_r(\underset{\sim}{\mu}, \underset{\sim}{\Sigma})$ with $|\underset{\sim}{\Sigma}| > 0$. Then $(\underset{\sim}{X} - \underset{\sim}{\mu})' \underset{\sim}{\Sigma}^{-1} (\underset{\sim}{X} - \underset{\sim}{\mu})$ is distributed as χ^2_r, where χ^2_r denotes the chi-square distribution with r degrees of freedom.

In general, write $H_0 : \underset{\sim}{C}\underset{\sim}{\beta} = \underset{\sim}{d}$

For $H_0 : \beta_3 = \beta_4 = 0$

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

For $H_0 : \beta_3 = \beta_4$

$$\begin{pmatrix} 0 & 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix} = 0$$

For $H_0 : \beta_3 = 2$

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix} = 2$$

$H_0 : \underset{\sim}{C}\underset{\sim}{\beta} = \underset{\sim}{d}$

$$\begin{gathered} \hat{\underset{\sim}{\beta}} \sim \mathrm{N}(\underset{\sim}{\beta}, (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \sigma^2) \\ \underset{\sim}{C}\hat{\underset{\sim}{\beta}} \sim \mathrm{N}(\underset{\sim}{C}\underset{\sim}{\beta}, \sigma^2 \underset{\sim}{C}(\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{C}^T) \end{gathered}$$

So, under H_0,

$$\begin{aligned} & \underset{\sim}{C}\hat{\underset{\sim}{\beta}} \sim \mathrm{N}(\underset{\sim}{d}, \sigma^2 \underset{\sim}{C}(\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{C}^T) \\ & \Rightarrow \underset{\sim}{C}\hat{\underset{\sim}{\beta}} - \underset{\sim}{d} \sim \mathrm{N}(\underset{\sim}{0}, \sigma^2 \underset{\sim}{C}(\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{C}^T) \\ & \Rightarrow \frac{(\underset{\sim}{C}\hat{\underset{\sim}{\beta}} - \underset{\sim}{d})^T [\underset{\sim}{C}(\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{C}^T]^{-1} (\underset{\sim}{C}\hat{\underset{\sim}{\beta}} - \underset{\sim}{d})}{\sigma^2} \sim \chi^2_{(r)} \\ & \Rightarrow \frac{(\underset{\sim}{C}\hat{\underset{\sim}{\beta}} - \underset{\sim}{d})^T [\underset{\sim}{C}(\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{C}^T]^{-1} (\underset{\sim}{C}\hat{\underset{\sim}{\beta}} - \underset{\sim}{d})}{r\hat{\sigma}^2} \sim \mathrm{F}_{(r, n-p')} \end{aligned}$$

- For two-sided alternative
- Handle one or more than one linear combinations of regression coefficients including intercept

Sequential sum of squares

$\mathrm{R}(\beta_1, \beta_2, \ldots, \beta_p | \beta_0) = \mathrm{R}(\beta_1 | \beta_0) + \mathrm{R}(\beta_2 | \beta_1, \beta_0) + \mathrm{R}(\beta_3 | \beta_2, \beta_1, \beta_0) + \ldots + \mathrm{R}(\beta_p | \beta_{p-1}, \beta_{p-2}, \ldots, \beta_1, \beta_0)$

Partial sum of squares

$$
\begin{array}{l}
\mathrm{R}(\beta_1|\beta_p,\ldots,\beta_2,\beta_0)\\
\mathrm{R}(\beta_2|\beta_p,\ldots,\beta_3,\beta_1,\beta_0)\\
\vdots\\
\mathrm{R}(\beta_p|\beta_{p-1},\ldots,\beta_2,\beta_1,\beta_0)
\end{array}
$$

5. Prediction

For p=1

$$\begin{aligned} y &= \beta_0 + \beta_1 x_1 + e \\ \mathrm{E}(y_0) &= \beta_0 + \beta_1 x_{01} \\ \widehat{\mathrm{E}(y_0)} &= \hat{\beta}_0 + \hat{\beta}_1 x_{01} \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(\widehat{\mathrm{E}(y_0)}) &= \mathrm{Var}(\hat{\beta}_0 + \hat{\beta}_1 x_{01}) \\ &= \mathrm{Var}(\hat{\beta}_0) + x_{01}^2 \mathrm{Var}(\hat{\beta}_1) + 2x_{01}\mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_1) \\ &= \sigma^2(\frac{1}{n} + \frac{\bar{x}_1^2}{S_{x_1x_1}}) + x_{01}^2 \frac{\sigma^2}{S_{x_1x_1}} - 2x_{01}\sigma^2 \frac{\bar{x}_1}{S_{x_1x_1}} \\ &= \sigma^2(\frac{1}{n} + \frac{(x_{01} - \bar{x}_1)^2}{S_{x_1x_1}}) \end{aligned}$$

$(1-\alpha)$ C.I. for mean value of y at x_{01} ($\mu_{y|x_{01}}$):

$$(\hat{\beta}_0 + \hat{\beta}_1 x_{01}) - t_{\alpha/2,(n-2)}\hat{\sigma}\sqrt{\frac{1}{n} + \frac{(x_{01} - \bar{x}_1)^2}{S_{x_1x_1}}} < \mu_{y|x_{01}} < (\hat{\beta}_0 + \hat{\beta}_1 x_{01}) + t_{\alpha/2,(n-2)}\hat{\sigma}\sqrt{\frac{1}{n} + \frac{(x_{01} - \bar{x}_1)^2}{S_{x_1x_1}}}$$

Individual value of y at x_0 (y_0): $y_0 = \beta_0 + \beta_1 x_{01} + e_0$

Point estimation: $\hat{y_0} = \hat{\beta}_0 + \hat{\beta}_1 x_{01}$ (take e_0 equal to zero)

$$\begin{aligned} E(\hat{y_0} - y_0) &= E[\hat{\beta}_0 + \hat{\beta}_1 x_{01} - (\beta_0 + \beta_1 x_{01} + e_0)] \\ &= 0 \end{aligned}$$

Since $\widehat{\mathrm{E}(y_0)}$ & e_0 are independent (i.e., $\mathrm{Cov}(\widehat{\mathrm{E}(y_0)}, e_0) = 0$)

$$\begin{aligned} \mathrm{Var}(\hat{y_0} - y_0) &= Var(\hat{\beta}_0 + \hat{\beta}_1 x_{01} - e_0) \\ &= \sigma^2\left(\frac{1}{n} + \frac{(x_{01} - \bar{x}_1)^2}{S_{x_1x_1}}\right) + \sigma^2 \end{aligned}$$

$(1-\alpha)$ C.I. for y_0 at x_0 :

$$(\hat{\beta}_0 + \hat{\beta}_1 x_{01}) - t_{\alpha/2,(n-2)}\hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(x_{01} - \bar{x}_1)^2}{S_{x_1x_1}}} < y_0 < (\hat{\beta}_0 + \hat{\beta}_1 x_{01}) + t_{\alpha/2,(n-2)}\hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(x_{01} - \bar{x}_1)^2}{S_{x_1x_1}}}$$

Example 1: Example in Simple Linear Regression (cont.)

Construct 95% confidence limits for the mean response $\mu_{Y|x}$ at $x_0 = 2$

$$\hat{y_0} = 0.2568 + (2.9303)(2) = 6.1174$$

Therefore, a 95% confidence interval for $\mu_{Y|x_0=2}$ is given by

$$6.1174 - (2.365)(0.538)\sqrt{\frac{1}{9} + \frac{(2-3.3667)^2}{13.10}} < \mu_{Y|x_0=2} < 6.1174 + (2.365)(0.538)\sqrt{\frac{1}{9} + \frac{(2-3.3667)^2}{13.10}}$$

$\Rightarrow 5.4765 < \mu_{Y|x_0=2} < 6.7583$

A 95% confidence interval for y_0 is given by

$$6.1174-(2.365)(0.538)\sqrt{1+\frac{1}{9}+\frac{(2-3.3667)^2}{13.10}} < y_0 < 6.1174+(2.365)(0.538)\sqrt{1+\frac{1}{9}+\frac{(2-3.3667)^2}{13.10}}$$

$\Rightarrow 4.6927 < y_0 < 7.5421$

For any p

New observation: $\underset{\sim}{x}_0^T = (1, x_{01}, \ldots, x_{0p})$

$$\begin{aligned}
\hat{\mu}_{y_0|\underset{\sim}{x}_0} &= \hat{\beta}_0 + \hat{\beta}_1 x_{01} + \hat{\beta}_2 x_{02} + \ldots + \hat{\beta}_p x_{0p} \\
&= \begin{pmatrix} 1 & x_{01} & \ldots & x_{0p} \end{pmatrix} \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_p \end{pmatrix} \\
&= \underset{\sim}{x}_0^T \hat{\underset{\sim}{\beta}}
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}(\hat{\mu}_{y_0|\underset{\sim}{x}_0}) &= \underset{\sim}{x}_0^T \mathrm{Var}(\hat{\underset{\sim}{\beta}})(\underset{\sim}{x}_0)^T \\
&= \underset{\sim}{x}_0^T \mathrm{Var}(\hat{\underset{\sim}{\beta}})\underset{\sim}{x}_0 \\
&= \sigma^2 \underset{\sim}{x}_0{}^T (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{x}_0
\end{aligned}$$

C.I. for mean value of y at x_0 ($\mu_{y_0|\underset{\sim}{x}_0}$):

$$\underset{\sim}{x}_0^T \hat{\underset{\sim}{\beta}} - t_{\alpha/2,(n-p')}\hat{\sigma}\sqrt{\underset{\sim}{x}_0^T (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{x}_0} < \mu_{y_0|\underset{\sim}{x}_0} < \underset{\sim}{x}_0^T \hat{\underset{\sim}{\beta}} + t_{\alpha/2,(n-p')}\hat{\sigma}\sqrt{\underset{\sim}{x}_0^T (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{x}_0}$$

C.I. for individual value of y at x_0 (prediction interval) :

$$\underset{\sim}{x}_0^T \hat{\underset{\sim}{\beta}} - t_{\alpha/2,(n-p')}\hat{\sigma}\sqrt{1+\underset{\sim}{x}_0^T (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{x}_0} < y_0 < \underset{\sim}{x}_0^T \hat{\underset{\sim}{\beta}} + t_{\alpha/2,(n-p')}\hat{\sigma}\sqrt{1+\underset{\sim}{x}_0^T (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{x}_0}$$

Example 5: Example in Multiple Linear Regression (cont.)

Construct 95% confidence limits for the mean response $\mu_{Y|x}$ when $x_1 = 3$, $x_2 = 8$ and $x_3 = 9$.

$$\hat{y}_0 = 39.1574 + (1.0161)(3) + (-1.8616)(8) + (-0.3433)(9) = 24.2232$$

$$\underset{\sim}{x}_0^T (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{x}_0 = 0.1267, \quad t_{0.025}(9) = 2.262$$

Therefore, a 95% confidence interval for $\mu_{Y|x_0=2}$ is given by

$$24.2232 - (2.262)(2.073)\sqrt{0.1267} < \mu_{Y|\underset{\sim}{x}_0} < 24.2232 + (2.262)(2.073)\sqrt{0.1267}$$

$\Rightarrow 22.5541 < \mu_{Y|\underset{\sim}{x}_0} < 25.8923$

A 95% confidence interval for y_0 is given by

$$24.2232 - (2.262)(2.073)\sqrt{1.1267} < y_0 < 24.2232 + (2.262)(2.073)\sqrt{1.1267}$$

$\Rightarrow 19.2459 < y_0 < 29.2$

Coefficient of determination

$$
\begin{aligned}
R^2 &= \frac{Reg\ S.S.}{Total\ S.S.} \qquad \text{(coefficient of determination)} \\
&= \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \\
&= \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2 - \sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \\
&= 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \\
&= 1 - \frac{Res\ S.S.}{Total\ S.S.}
\end{aligned}
$$

$$0 \leq R^2 \leq 1$$

For p=1

$$
\begin{aligned}
R^2 &= \frac{Reg\ S.S.}{Total\ S.S.} \\
&= \frac{\hat{\beta}^2 S_{xx}}{S_{yy}} \\
&= \frac{(\frac{S_{xy}}{S_{xx}})^2 S_{xx}}{S_{yy}} \\
&= \frac{S_{xy}^2}{S_{xx}S_{yy}} \\
&= (\frac{S_{xy}}{\sqrt{S_{xx}}\sqrt{S_{yy}}})^2
\end{aligned}
$$

$$\sqrt{R^2}\begin{cases} \text{simple linear reg. - } r \text{ simply correlation coeff. (linear relationship between } y \text{ and } x) \\ \text{multple linear reg. - multiple correlation coeff. (linear relationship between } y \text{ and } \underset{\sim}{x}) \end{cases}$$

Example 1: Example in Simple Linear Regression (cont.)

$$R^2 = \frac{112.4852}{114.511} = 0.9823 \quad \Rightarrow \quad r = 0.9911$$

Example 5: Example in Multiple Linear Regression (cont.)

$$R^2 = \frac{399.45}{438.17} = 0.9117$$

91.17% of the variation in Y has been explained by the linear regression model.

6. Lack of fit

Let y_{ij} represent the jth response at the ith experimental combination, $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n_i$

Example:

Table 3.1 Breadwrapper Stock Data

y(g/in.)	x_1 (°F)	x_2 (°F)	x_3 (weight in %)
6.6	225	46	0.5
6.9	285	46	0.5
7.9	225	64	0.5
6.1	285	64	0.5
9.2	225	46	1.7
6.8	285	46	1.7
10.4	225	64	1.7
7.3	285	64	1.7
9.8	204.5	55	1.1
5.0	305.5	55	1.1
6.9	255	39.9	1.1
6.3	255	70.1	1.1
4.0	255	55	0.09
8.6	255	55	2.11
10.1	255	55	1.1
9.9	255	55	1.1
12.2	255	55	1.1
9.7	255	55	1.1
9.7	255	55	1.1
9.6	255	55	1.1

We fit a model of x_1, x_2, x_3, x_1^2, x_2^2, x_3^2, x_1x_2, x_1x_3, $x_2x_3 \Rightarrow p' = 10$.

$$m = 15, n_1 = 1, n_2 = 1, \ldots, n_{14} = 1, n_{15} = 6$$

$$\textstyle\sum_{j=1}^{m} n_i = n(= 20)$$

$$m > p'(= 10)$$

$$\begin{aligned}
\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_i)^2 &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i + \bar{y}_i - \hat{y}_i)^2 \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2 + \sum_{i=1}^{m}\sum_{j=1}^{n_i}(\bar{y}_i - \hat{y}_i)^2 + 2\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)(\bar{y}_i - \hat{y}_i)
\end{aligned}$$

Then

$$\underbrace{\textstyle\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_i)^2}_{\text{Res S.S.}} = \underbrace{\textstyle\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2}_{\text{Pure error S.S.}} + \underbrace{\textstyle\sum_{i=1}^{m} n_i(\bar{y}_i - \hat{y}_i)^2}_{\text{Lack of fit S.S.}}$$

$$\begin{aligned}
\mathrm{E}(\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_i)^2) &= \sum_{i=1}^{m}\mathrm{E}(\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_i)^2) \\
&= \sum_{i=1}^{m}(n_i-1)\sigma^2 \\
&= \sigma^2\sum_{i=1}^{m}(n_i-1) \\
&= \sigma^2(n-m)
\end{aligned}$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_i)^2}{n-m} \qquad - \text{unbiased estimator}$$

$$\mathrm{E}(\underset{\sim}{Y}^T A\underset{\sim}{Y}) = \mathrm{trace}(\underset{\sim}{\Sigma}\underset{\sim}{A}) + \underset{\sim}{\mu}^T\underset{\sim}{A}\underset{\sim}{\mu} = \sigma^2\mathrm{trace}(\underset{\sim}{A}) + \underset{\sim}{\mu}^T\underset{\sim}{A}\underset{\sim}{\mu}$$

true model :

$$\underset{\sim}{Y} = \underset{\sim}{X}_1\underset{\sim}{\beta}_1 + \underset{\sim}{X}_2\underset{\sim}{\beta}_2 + \underset{\sim}{e}$$

short model :

$$\underset{\sim}{Y} = \underset{\sim}{X}_1\underset{\sim}{\beta}_1 + \underset{\sim}{e}^*$$

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

$$\begin{aligned}
\mathrm{trace}(\underset{\sim}{A}) &= \mathrm{trace}(\underset{\sim}{I} - \underset{\sim}{X}_1(\underset{\sim}{X}_1^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1^T) \\
&= n - \mathrm{trace}(\underset{\sim}{X}_1(\underset{\sim}{X}_1^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1^T) \\
&= n - \mathrm{trace}((\underset{\sim}{X}_1^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1^T\underset{\sim}{X}_1) \\
&= n - \mathrm{trace}(\underset{\sim}{I}) \\
&= n - p'
\end{aligned}$$

$$\begin{aligned}
\underset{\sim}{\mu}^T\underset{\sim}{A}\underset{\sim}{\mu} &= (\underset{\sim}{X}_1\underset{\sim}{\beta}_1 + \underset{\sim}{X}_2\underset{\sim}{\beta}_2)^T(\underset{\sim}{I} - \underset{\sim}{X}_1(\underset{\sim}{X}_1^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1^T)(\underset{\sim}{X}_1\underset{\sim}{\beta}_1 + \underset{\sim}{X}_2\underset{\sim}{\beta}_2) \\
&= (\underset{\sim}{\beta}_1^T\underset{\sim}{X}_1^T + \underset{\sim}{\beta}_2^T\underset{\sim}{X}_2^T)(\underset{\sim}{I} - \underset{\sim}{X}_1(\underset{\sim}{X}_1^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1^T)(\underset{\sim}{X}_1\underset{\sim}{\beta}_1 + \underset{\sim}{X}_2\underset{\sim}{\beta}_2) \\
&= \underset{\sim}{\beta}_1^T\underset{\sim}{X}_1^T(\underset{\sim}{I} - \underset{\sim}{X}_1(\underset{\sim}{X}_1^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1^T)\underset{\sim}{X}_1\underset{\sim}{\beta}_1 + \underset{\sim}{\beta}_2^T\underset{\sim}{X}_2^T(\underset{\sim}{I} - \underset{\sim}{X}_1(\underset{\sim}{X}_1^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1^T)\underset{\sim}{X}_2\underset{\sim}{\beta}_2 \\
&\quad + \underset{\sim}{\beta}_1^T\underset{\sim}{X}_1^T(\underset{\sim}{I} - \underset{\sim}{X}_1(\underset{\sim}{X}_1^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1^T)\underset{\sim}{X}_2\underset{\sim}{\beta}_2 + \underset{\sim}{\beta}_2^T\underset{\sim}{X}_2^T(\underset{\sim}{I} - \underset{\sim}{X}_1(\underset{\sim}{X}_1^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1^T)\underset{\sim}{X}_1\underset{\sim}{\beta}_1 \\
&= \underset{\sim}{\beta}_2^T\underset{\sim}{X}_2^T\underset{\sim}{X}_2\underset{\sim}{\beta}_2 - \underset{\sim}{\beta}_2^T\underset{\sim}{X}_2^T\underset{\sim}{X}_1(\underset{\sim}{X}_1^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1^T\underset{\sim}{X}_2\underset{\sim}{\beta}_2 \\
&= \underset{\sim}{\beta}_2^T(\underset{\sim}{X}_2^T\underset{\sim}{X}_2 - \underset{\sim}{X}_2^T\underset{\sim}{X}_1(\underset{\sim}{X}_1^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1^T\underset{\sim}{X}_2)\underset{\sim}{\beta}_2
\end{aligned}$$

$$\mathrm{E}(RSS) = (n-p')\sigma^2 + \underset{\sim}{\beta}_2^T(\underset{\sim}{X}_2^T\underset{\sim}{X}_2 - \underset{\sim}{X}_2^T\underset{\sim}{X}_1(\underset{\sim}{X}_1^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1^T\underset{\sim}{X}_2)\underset{\sim}{\beta}_2$$

$$\begin{aligned}
\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij}-\hat{y}_i)^2 &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_i)^2 + \sum_{i=1}^{m}n_i(\bar{y}_i-\hat{y}_i)^2 \\
\mathrm{E}(\sum_{i=1}^{m}n_i(\bar{y}_i-\hat{y}_i)^2) &= (m-p')\sigma^2 + \underset{\sim}{\beta}_2^T(\underset{\sim}{X}_2^T\underset{\sim}{X}_2 - \underset{\sim}{X}_2^T\underset{\sim}{X}_1(\underset{\sim}{X}_1^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1^T\underset{\sim}{X}_2)\underset{\sim}{\beta}_2 \\
\Rightarrow \mathrm{E}(\frac{\sum_{i=1}^{m}n_i(\bar{y}_i-\hat{y}_i)^2}{m-p'}) &= \sigma^2 + \frac{\underset{\sim}{\beta}_2^T(\underset{\sim}{X}_2^T\underset{\sim}{X}_2 - \underset{\sim}{X}_2^T\underset{\sim}{X}_1(\underset{\sim}{X}_1^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1^T\underset{\sim}{X}_2)\underset{\sim}{\beta}_2}{m-p'}
\end{aligned}$$

$$\mathrm{F} = \frac{\displaystyle\sum_{i=1}^{m} n_i(\bar{y}_i - \hat{y}_i)^2/(m-p')}{\displaystyle\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2/(n-m)} \sim \mathrm{F}(m-p', n-m)$$

$$\text{Reject} H_0 \text{ if } \mathrm{F} > \mathrm{F}_\alpha(m-p', n-m)$$

$$\begin{aligned}
\mathrm{E}(\mathrm{F}) &\approx \frac{\mathrm{E}(\displaystyle\sum_{i=1}^{m} n_i(\bar{y}_i - \hat{y}_i)^2/(m-p'))}{\mathrm{E}(\displaystyle\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2/(n-m))} \\
&= \frac{\sigma^2 + \dfrac{\underset{\sim}{\beta}_2^T(\underset{\sim}{X}_2^T\underset{\sim}{X}_2 - \underset{\sim}{X}_2^T\underset{\sim}{X}_1(\underset{\sim}{X}_1^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1^T\underset{\sim}{X}_2)\underset{\sim}{\beta}_2}{m-p'}}{\sigma^2} \\
&= 1 + \frac{\underset{\sim}{\beta}_2^T(\underset{\sim}{X}_2^T\underset{\sim}{X}_2 - \underset{\sim}{X}_2^T\underset{\sim}{X}_1(\underset{\sim}{X}_1^T\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1^T\underset{\sim}{X}_2)\underset{\sim}{\beta}_2}{\sigma^2(m-p')}
\end{aligned}$$

Table 3.1

Source	S.S.	d.f.	M.S.	F
Regression	70.302	9	7.8113	
Error	11.8678	10	1.18678	
Lack of fit	6.9078	5	1.3816	1.39
Pure Error	4.96	5	0.9920	
Total	82.17	19		

7. Added variable plot

For $p = 2$, $\quad y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + e_i \quad i = 1, \ldots, n$

1. Fit y_i on x_{i2}

$$\begin{aligned}\Rightarrow \hat{e}_{Y(x_1)} &= y_i - (\hat{\delta}_0 + \hat{\delta}_1 x_{i2}) \\ &= y_i - [\bar{y} + \hat{\delta}_1(x_{i2} - \bar{x}_2)] \\ &= y_i - \bar{y} - \hat{\delta}_1(x_{i2} - \bar{x}_2)\end{aligned}$$

2. Fit x_{i1} on x_{i2}

$$\begin{aligned}\Rightarrow \hat{e}_1 &= x_{i1} - (\hat{\gamma}_0 + \hat{\gamma}_1 x_{i2}) \\ &= x_{i1} - (\bar{x}_1 - \hat{\gamma}_1 \bar{x}_2 + \hat{\gamma}_1 x_{i2}) \\ &= x_{i1} - \bar{x}_1 - \hat{\gamma}_1(x_{i2} - \bar{x}_2)\end{aligned}$$

3. Plot $\hat{e}_{Y(x_1)}$ versus $\hat{e}_1$,

$$\begin{aligned}\text{slope} &= \frac{\sum_{i=1}^{n}((y_i - \bar{y}) - \hat{\delta}_1(x_{i2} - \bar{x}_2))((x_{i1} - \bar{x}_1) - \hat{\gamma}_1(x_{i2} - \bar{x}_2))}{\sum_{i=1}^{n}((x_{i1} - \bar{x}_1) - \hat{\gamma}_1(x_{i2} - \bar{x}_2))^2} \\ &= \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_{i1} - \bar{x}_1) - \hat{\delta}_1\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) - \hat{\gamma}_1\sum_{i=1}^{n}(y_i - \bar{y})(x_{i2} - \bar{x}_2) + \hat{\delta}_1\hat{\gamma}_1\sum_{i=1}^{n}(x_{i2} - \bar{x}_2)^2}{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)^2 - 2\hat{\gamma}_1\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) + \hat{\gamma}_1^2\sum_{i=1}^{n}(x_{i2} - \bar{x}_2)^2} \\ &= \frac{S_{y1} - \hat{\delta}_1 S_{12} - \hat{\gamma}_1 S_{y2} + \hat{\delta}_1\hat{\gamma}_1 S_{22}}{S_{11} - 2\hat{\gamma}_1 S_{12} + \hat{\gamma}_1^2 S_{22}} \\ &= \frac{S_{y1}S_{22} - S_{12}S_{y2}}{S_{11}S_{22} - S_{12}^2} \\ &\left[\text{Since } \ \hat{\delta}_1 = \frac{S_{y2}}{S_{22}} \ \text{ (Regress } y \text{ on } x_2\text{)}, \ \ \hat{\gamma}_1 = \frac{S_{12}}{S_{22}} \ \text{ (Regress } x_1 \text{ on } x_2\text{)}\right] \\ &= \hat{\beta}_1\end{aligned}$$

For $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \ldots + \beta_p x_{ip} + e_i$

1. Fit y_i on all the x's except x_k

$$\begin{aligned}\text{Residual} &= y_i - \text{ fitted value of } y_i \\ &= \hat{e}_{Y(x_k)}\end{aligned}$$

2. Fit x_k on the other x's

$$\begin{aligned}\text{Residual} &= x_{ik} - \text{ fitted value of } x_{ik} \\ &= \hat{e}_k\end{aligned}$$

3. Plot $\hat{e}_{Y(x_k)}$ versus $\hat{e}_k$. Then, slope $= \hat{\beta}_k$

Partial Correlation Coefficient

1. $\hat{e}_{Y(x_k)}$: residual from the regression of y on all variance except x_k
2. $\hat{e}_k$: residual from the regression of x_k on all other x variables
3. slope of the regression of $\hat{e}_{Y(x_k)}$ on $\hat{e}_k = \hat{\beta}_k$
4. simple corr. coeff. of $\hat{e}_{Y(x_k)}$ and $\hat{e}_k$
   = partial corr. coeff. of y and x_k on $x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_p$

Thus, partial corr. coeff. of y and x_1 on x_2

$$
\begin{aligned}
&= \frac{\sum_{i=1}^{n}((y_i-\bar{y})-\hat{\delta}_1(x_{i2}-\bar{x}_2))((x_{i1}-\bar{x}_1)-\hat{\gamma}_1(x_{i2}-\bar{x}_2))}{\sqrt{\sum_{i=1}^{n}((y_i-\bar{y})-\hat{\delta}_1(x_{i2}-\bar{x}_2))^2}\sqrt{\sum_{i=1}^{n}((x_{i1}-\bar{x}_1)-\hat{\gamma}_1(x_{i2}-\bar{x}_2))^2}} \\
&= \frac{S_{y1}S_{22}-S_{12}S_{y2}}{\sqrt{S_{yy}S_{22}-S_{y2}^2}\sqrt{S_{11}S_{22}-S_{12}^2}} \\
&= \frac{S_{y1}-\frac{S_{12}S_{y2}}{S_{22}}}{\sqrt{S_{yy}-\frac{S_{y2}^2}{S_{22}}}\sqrt{S_{11}-\frac{S_{12}^2}{S_{22}}}} \\
&= \frac{r_{y1}-r_{12}r_{y2}}{\sqrt{1-r_{y2}^2}\sqrt{1-r_{12}^2}}
\end{aligned}
$$