<u>Chapter 1</u>

Matrix Approach to Simple Regression Model

Overview

- Least squares estimation
- Sum of squares: SST, SSR, and SSE
- ANOVA table, F-test
- Variance of $\underline{\hat{\beta}}$
- Confidence interval for β_0 and β_1
- Confidence interval for $\mu_{Y|x}$
- Prediction interval for $Y|x$

1.1 Simple Regression Model

Consider the model

$$y_i = \beta_0 + \beta_1 x_i + \boxed{\varepsilon_i,}\; i = 1, \ldots, n \qquad (1)$$

where y_i is the dependent (or response) variable,

x_i is the independent (or predictor) variable,

β_0 and β_1 are unknown parameters,

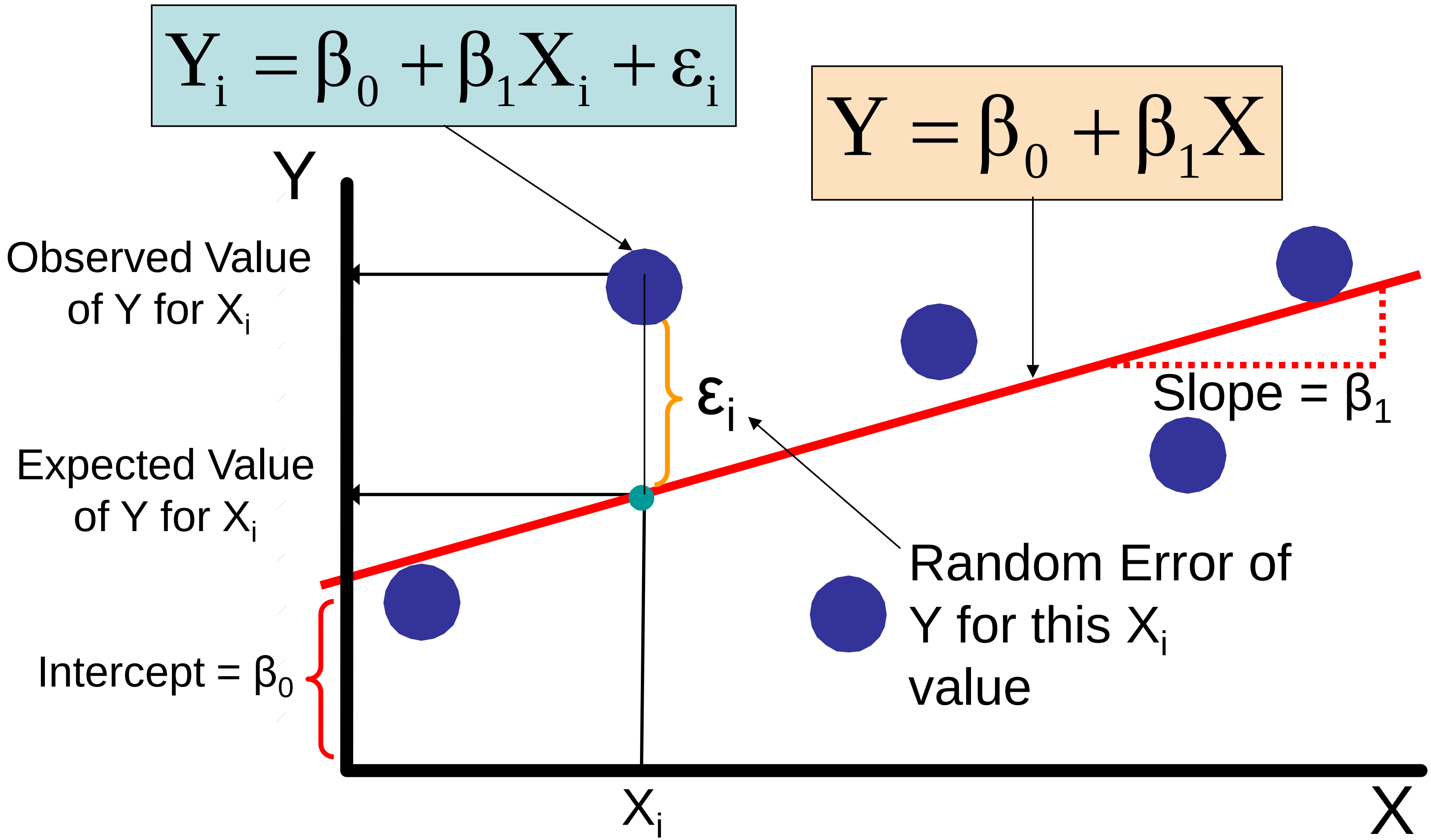
ε_i's are independent normal random variables,

with $E(\varepsilon_i) = 0$ and $Var(\varepsilon_i) = \sigma^2$ for all i.

Deterministic but unknown

Random

Simple Regression Model (Continued)



Simple Regression Model (Continued)

- Model (1) can be expressed in a matrix form as follows:

$$\underline{y} = X\,\underline{\beta} + \underline{\varepsilon}$$

where $\underline{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, $X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$, $\underline{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$, and

$\underline{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$ with $\mathrm{E}(\underline{\varepsilon}) = \underline{0}$ and $\mathrm{Var}(\underline{\varepsilon}) = \sigma^2 I_n$

1.2 Least Squares Estimation

- Least squares estimator for $\underline{\beta}$ can be obtained by minimizing $\underline{\varepsilon}'\underline{\varepsilon}$ (i.e. $\sum_{i=1}^{n} \epsilon_i^2$).

- Since $\underline{\varepsilon} = \underline{y} - X\,\underline{\beta}$, therefore

$$\underline{\epsilon}'\underline{\epsilon} = \left(\underline{y} - X\underline{\beta}\right)'\left(\underline{y} - X\underline{\beta}\right)$$

$$= \underline{y}'\underline{y} - \underline{\beta}'X'\underline{y} - \underline{y}'X\underline{\beta} + \underline{\beta}'X'X\underline{\beta}$$

$$= \underline{\beta}'X'X\underline{\beta} - 2\underline{\beta}'X'\underline{y} + \underline{y}'\underline{y}$$

Least Squares Estimation (Continued)

- Assuming that $X'X$ is non-singular (i.e. $(X'X)^{-1}$ exists), then

$$\underline{\epsilon}'\underline{\epsilon} = \left[\underline{\beta} - (X'X)^{-1}X'\underline{y}\right]'(X'X)\left[\underline{\beta} - (X'X)^{-1}X'\underline{y}\right]$$

$$-\underline{y}'X(X'X)^{-1}X'\underline{y} + \underline{y}'\underline{y}$$

- If we choose $\underline{\beta} = (X'X)^{-1}X'\underline{y}$, then the first term vanishes and $\underline{\varepsilon}'\underline{\varepsilon}$ is at minimum.
- $\underline{\hat{\beta}} = (X'X)^{-1}X'\underline{y}$ is called the **least squares estimate of $\underline{\boldsymbol{\beta}}$**

1.3 Least Squares Estimation for SRM

- Refer to Model (1)

$$X'X = \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} = \begin{pmatrix} n & n\bar{x} \\ n\bar{x} & \sum_{i=1}^{n} x_i^2 \end{pmatrix}$$

$$X'\underline{y} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} n\bar{y} \\ \sum_{i=1}^{n} x_i y_i \end{pmatrix}$$

- It can be shown that

$$(X'X)^{-1} = \frac{1}{n\sum_{i=1}^{n} x_i^2 - (n\bar{x})^2} \begin{pmatrix} \sum_{i=1}^{n} x_i^2 & -n\bar{x} \\ -n\bar{x} & n \end{pmatrix}$$

1.3 Least Squares Estimation for SRM

- Hence

$$\underline{\hat{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = (X'X)^{-1}X'\underline{y}$$

$$= \frac{1}{n\sum_{i=1}^{n} x_i^2 - (n\bar{x})^2} \begin{pmatrix} \sum_{i=1}^{n} x_i^2\,(n\bar{y}) - n\bar{x}\ \sum_{i=1}^{n} x_i y_i \\ -n\bar{x}(n\bar{y}) + n\sum_{i=1}^{n} x_i y_i \end{pmatrix}$$

- Note:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2} \text{ and } \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

- The equation of the regression line is $\hat{y} = \underline{x}'\underline{\hat{\beta}}$,

  where $\underline{x}' = \begin{pmatrix} 1 \\ x \end{pmatrix}$. i.e. $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$

1.4 Unbiasedness

- $\underline{\hat{\beta}}$ is an **unbiased** estimator for $\underline{\beta}$

$$E\left(\underline{\hat{\beta}}\right) = E\left[(X'X)^{-1}X'\underline{y}\right] = (X'X)^{-1}X'E\left(\underline{y}\right)$$

$$= (X'X)^{-1}X'E\left(X\underline{\beta} + \underline{\epsilon}\right)$$

$$= (X'X)^{-1}X'\left(X\underline{\beta} + E(\underline{\epsilon})\right)$$

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

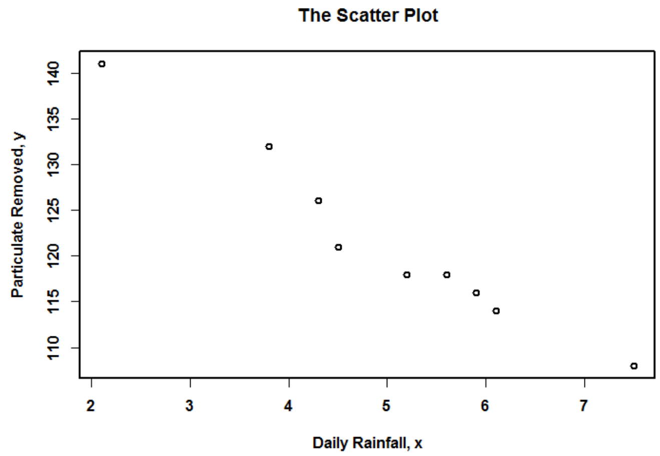
1.5 Example 1

- In a study between the amount of rainfall and the quantity of air pollutant removed, the following data were collected:

x	4.3	4.5	5.9	5.6	6.1	5.2	3.8	2.1	7.5
y	126	121	116	118	114	118	132	141	108

x: daily rainfall (0.01 cm), y: particulate removed (10^{-6} gm/m^3)

- Suppose Model (1) is appropriate. Find the Least Squares Estimate of $\underline{\beta}$.



Solution to Example 1

$$X'X = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 4.3 & 4.5 & \cdots & 7.5 \end{pmatrix} \begin{pmatrix} 1 & 4.3 \\ 1 & 4.5 \\ \vdots & \vdots \\ 1 & 7.5 \end{pmatrix} = \begin{pmatrix} 9 & 45 \\ 45 & 244.26 \end{pmatrix}$$

$$(X'X)^{-1} = \frac{1}{9(244.26) - 45^2} \begin{pmatrix} 244.26 & -45 \\ -45 & 9 \end{pmatrix}$$

$$X'\underline{y} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 4.3 & 4.5 & \cdots & 7.5 \end{pmatrix} \begin{pmatrix} 126 \\ 121 \\ \vdots \\ 108 \end{pmatrix} = \begin{pmatrix} 1094 \\ 5348.2 \end{pmatrix}$$

Solution to Example 1 (Continued)

Therefore

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

$$= \frac{1}{173.34}\begin{pmatrix} 244.26 & -45 \\ -45 & 9 \end{pmatrix}\begin{pmatrix} 1094 \\ 5348.2 \end{pmatrix} = \begin{pmatrix} 153.1755 \\ -6.3240 \end{pmatrix}$$

That is, $\hat{\beta}_0 = 153.1755$ and $\hat{\beta}_1 = -6.3240$

The regression line is $\hat{y} = 153.1755 - 6.3240x$

i.e. when the daily rainfall, x, increases by 1 unit (i.e. 0.01 cm), the amount of particulate, y, is reduced by an **average** of 6.3240 microgram per cubic meter

1.6 Sum of Squares

1.6.1 Introduction

- In the simple regression model (SRM) case, if we want to know if the predictor is useful, we test $H_0: \beta_1 = 0$ against $H_1 : \beta_1 \neq 0$.

- An alternative approach is to study how much variation that can be explained by the regression model.

1.6.2 Total Sum of Squares

An Example

- Consider the following exam marks for 10 students,

$$\underset{\sim}{y} = (50, 53, 54, 59, 60, 64, 66, 65, 67, 69)'$$

  - Mean of the 10 exam marks is 60.7
  - Variation of the 10 exam marks about the mean is given by

$$(50 - 60.7)^2 + \cdots + (69 - 60.7)^2 = 388.1$$

  - It is called the **Total Sum of Squares** (SST)
  - Sometimes it is referred as the corrected total sum of squares

1.6.3 Regression Sum of Squares

- Suppose we also have the continuous assessment mark for each of the10 students.

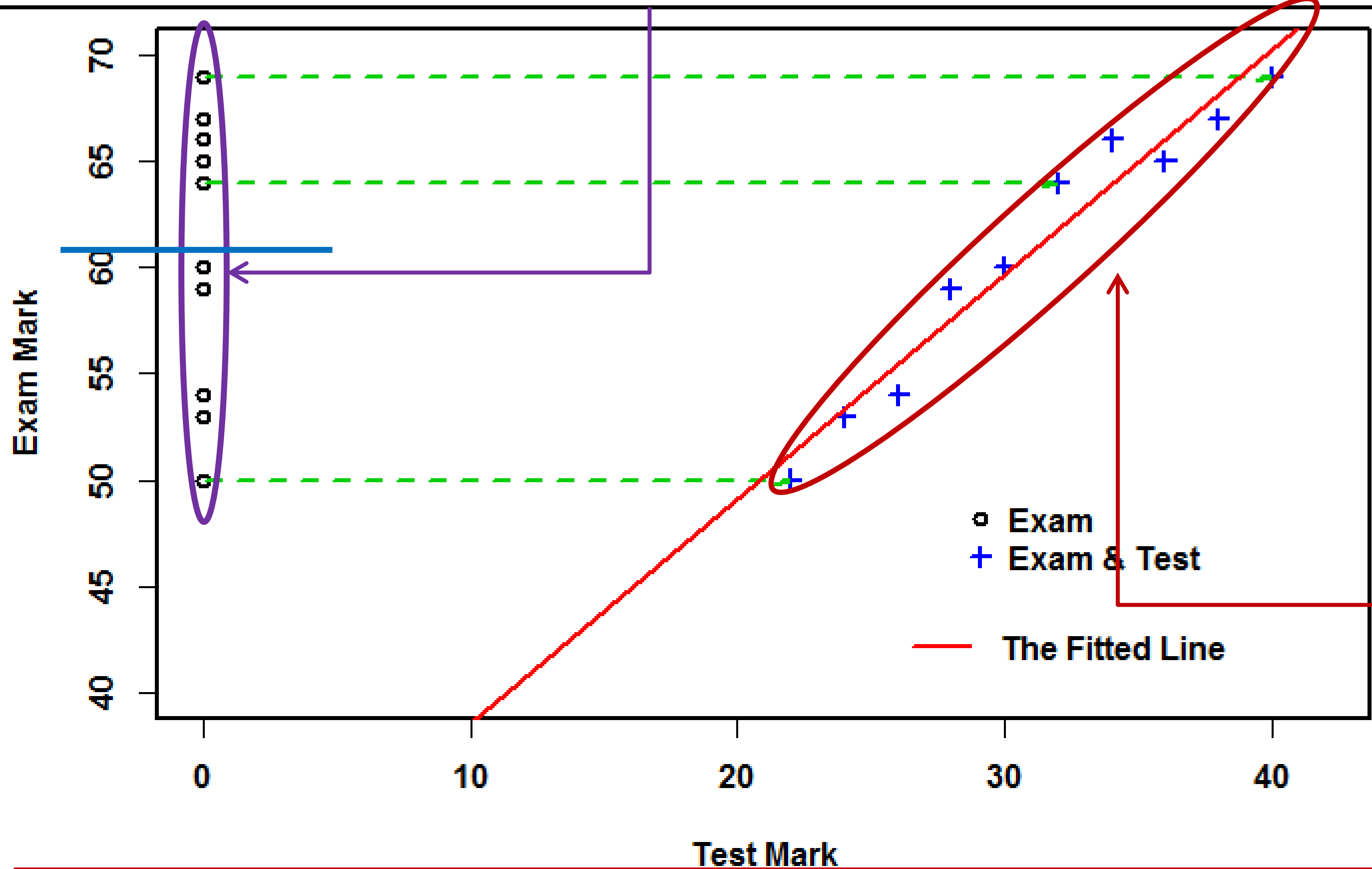
- They are

$$\underline{x} = (22, 24, 26, 28, 30, 32, 34, 36, 38, 40)'$$

- Would the knowledge of the CA marks (x) helps to explain some of the variation in the exam marks (y)?

1.6.3 Regression Sum of Squares

Without the info on test marks, it seems that the observed exam marks deviate a lot from the expected exam mark (the blue line)



With the info on test marks, it seems that the observed exam marks are close to the expected exam mark (estimated by the red fitted line)

1.6.3 Regression Sum of Squares

- Without the test marks, E(y) is estimated by the sample mean of the exam marks, 60.7.
- With the test marks and the assumption of a linear relationship between the test mark and the exam mark, E(y) for a student with x test mark is estimated by

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

- If the model is good, then we expect the observed exam mark should be close to the estimated exam mark. (i.e. $y - \hat{y}$ should be small)

1.6.3 Regression Sum of Squares

- Hence if the model is good, then $(y_i - \bar{y})^2$ should be close to $(\hat{y}_i - \bar{y})^2$
- Define the Regression Sum of Squares (SSR) as

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

- In the above example, SSR = 369.09
- SSR is close to SST (= 388.1)
- SSR is considered as the variation of y explained by the model $y = \beta_0 + \beta_1 x + \varepsilon$

1.6.4 Error Sum of Squares

- The variation of y which is not explained by the model is given by

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

- It is called Error Sum of Squares (SSE)
- In the above example, SSE = 19.01
- If the model is good, then SSE should be small.

- Notice that it can be shown that

$$\text{SST} = \text{SSR} + \text{SSE}$$

1.6.5 Sum of Squares Identity

- The **sum of squares** identity is

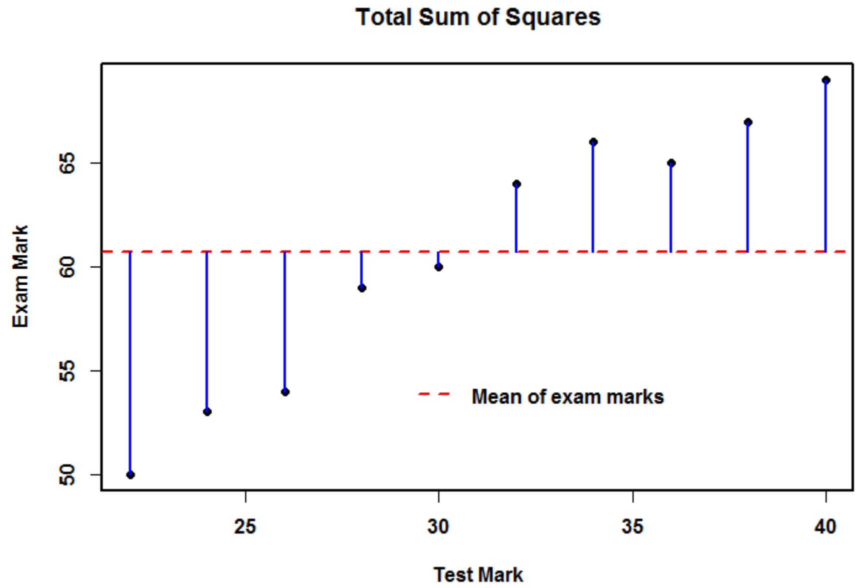
$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

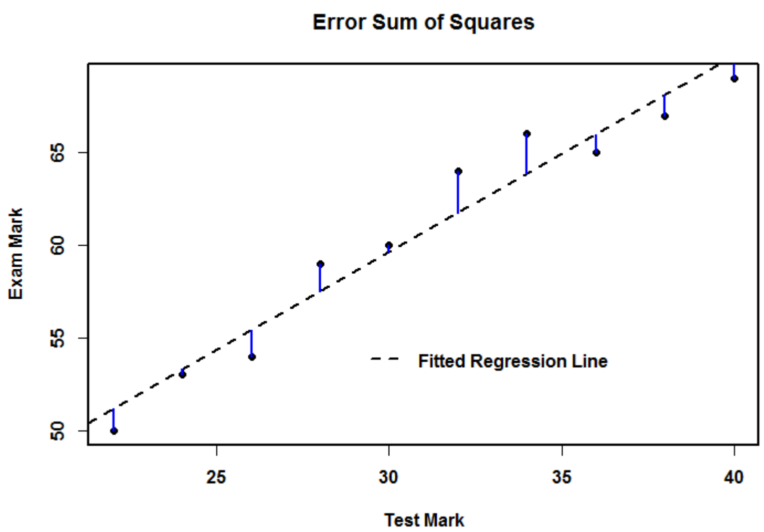
- Symbolically,

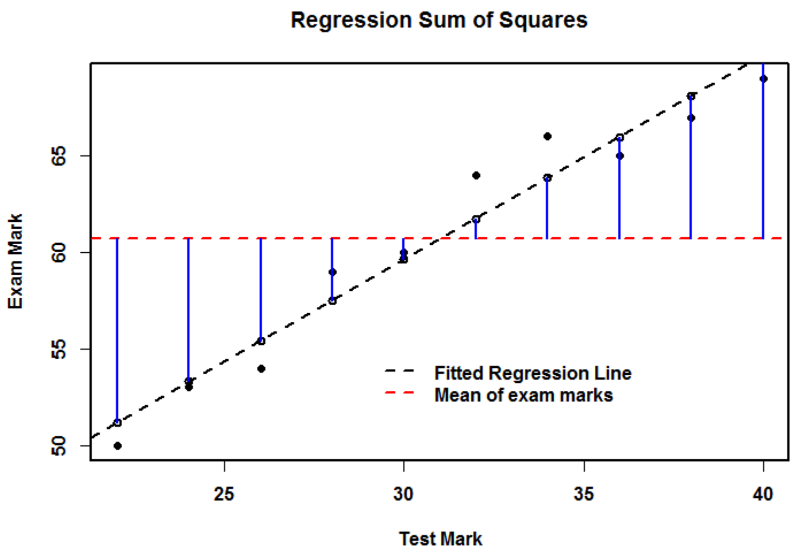
$$SST = SSR + SSE$$

- *SST:* Total Sum of Squares with n – 1 df
- *SSR:* Regression Sum of Squares (Variation explained by the model) with 1 df (for SRM)
- *SSE:* Sum of Squares Error (Variation unexplained by the model) with n – 2 df (for SRM)

Sum of squares (Continued)







1.6.6 Total Sum of Squares (Matrix)

Total variation around the mean is measured by the corrected total sum of squares.

$$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 = \underline{y}'\underline{y} - n\bar{y}^2$$

with n – 1 d.f.

Note: $\sum_{i=1}^{n} a_i^2 = (a_1 \quad \cdots \quad a_n)\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \underline{a}'\underline{a}$

1.6.7 Regression Sum of Squares (Matrix)

Variation explained by the regression model (due to β_1) is measured by the regression sum of squares.

$$SSR = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 = \left(\underline{\hat{y}} - \bar{y}\underline{1}_n\right)'\left(\underline{\hat{y}} - \bar{y}\underline{1}_n\right)$$

$$= \underline{\hat{y}}'\underline{\hat{y}} - \bar{y}\underline{1}_n'\underline{\hat{y}} - \bar{y}\underline{\hat{y}}'\underline{1}_n + \bar{y}^2\underline{1}_n'\underline{1}_n$$

$$= \underline{\hat{y}}'\underline{\hat{y}} - 2\bar{y}\underline{1}_n'\underline{\hat{y}} + n\bar{y}^2 = \underline{\hat{y}}'\underline{\hat{y}} - n\bar{y}^2$$

Note: $\bar{y}\underline{1}_n'\underline{\hat{y}} = \bar{y}\underline{1}_n'X\underline{\hat{\beta}} = \bar{y}(n \quad n\bar{x})\begin{pmatrix}\hat{\beta}_0 \\ \hat{\beta}_1\end{pmatrix}$

$$= n\bar{y}\left(\hat{\beta}_0 + \hat{\beta}_1\bar{x}\right) = n\bar{y}^2$$

Regression Sum of Squares (Matrix) (Continued)

Also $\underline{\hat{y}}'\underline{\hat{y}} = \underline{\hat{\beta}}'X'X\underline{\hat{\beta}} = \underline{\hat{\beta}}'X'\underline{y}$

Therefore $SSR = \underline{\hat{\beta}}'X'\underline{y} - n\bar{y}^2$ with p d.f., where p is the number of predictors in the model

Note: The degrees of freedom for SSR is 1 for the simple regression model (SRM)

1.6.8 Error Sum of Squares (Matrix)

- Unexplained variation is measured by error sum of squares

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

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

$$= \underline{y}'\underline{y} - \underline{\hat{\beta}}'X'\underline{y}$$

with $n - p - 1$ df

- Note: $\underline{y}'\underline{\hat{y}} = \underline{y}'X\underline{\hat{\beta}} = \underline{\hat{\beta}}'X'\underline{y}$ and $\underline{\hat{\beta}}'\left(X'X\underline{\hat{\beta}}\right) = \underline{\hat{\beta}}'X'\underline{y}$
- Remark: The degrees of freedom for the simple regression model ($y = \beta_0 + \beta_1 x + \varepsilon$) is $n - 2$.

1.6.9 F-test

- It can also be shown that if the model is not significant, then

  SSR/σ^2 and SSE/σ^2 are independent

  and follow $\chi^2(1)$ and $\chi^2(n-2)$ distributions respectively (for the simple regression model).

- Therefore

$$F = \frac{\frac{SSR/\sigma^2}{1}}{\frac{SSE/\sigma^2}{n-2}} = \frac{MSR}{MSE} \sim F(1, n-2)$$

F-test (Continued)

- To test whether the model is significant or not, we **reject** the null hypothesis of no significant model if

$$F_{\text{observed}} > F_{\alpha}(1, n-2).$$

- Note that for the **simple** regression model, test of significant model is equivalent to testing H_0: $\beta_1 = 0$ against H_1: $\beta_1 \neq 0$.

1.6.10 ANOVA table

- We can summarize the above results in an ANOVA table (for the simple regression model)

Source	SS	Df	MS	F-ratio	p-value
Regression	$SSR = \hat{\underline{\beta}}'X'\underline{y} - n\bar{y}^2$	1	$\frac{SSR}{1}$	$F_{obs} = \frac{MSR}{MSE}$	$\Pr\left(F(1, n-2) > F_{obs}\right)$
Error	$SSE = \underline{y}'\underline{y} - \hat{\underline{\beta}}'X'\underline{y}$	$n-2$	$\frac{SSE}{n-2}$		
Total	$SST = \underline{y}'\underline{y} - n\bar{y}^2$	$n-1$			

Example 1 (Continued)

$$SST = \underline{y}'\underline{y} - n\bar{y}^2$$

$$= (126 \quad \cdots \quad 108)\begin{pmatrix} 126 \\ \vdots \\ 108 \end{pmatrix} - 9(121.556)^2$$

$$= 133786 - 132981.778 = 804.222$$

$$SSR = \underline{\hat{\beta}}'X'\underline{y} - n\bar{y}^2$$

$$= (153.1755 \quad -6.3240)\begin{pmatrix} 1094 \\ 5348.2 \end{pmatrix} - 9(121.556)^2$$

$$= 133751.980 - 132981.778 = 770.202$$

$$SSE = SST - SSR = 804.222 - 770.202 = 34.020$$

Example 1 (Continued)

- ANOVA Table

Source	SS	Df	MS	F-ratio	p-value
Regression	770.202	1	770.202	158.48	$4.607(10)^{-6}$
Error	34.020	7	4.860		
Total	804.222	8			

- Since $F_{\text{obs}} = 158.48 > F_{0.05}(1, 7) = 5.59$ (or p-value < 0.0001), therefore we reject the null hypothesis that there is no significant relationship between daily rainfall and the amount particulate removed at the 5% level of significance.

1.8 Variance-Covariance Matrix of $\hat{\underline{\beta}}$

$$Var\left(\hat{\underline{\beta}}\right) = Var\left(A\underline{y}\right), \qquad \text{where } A = (X'X)^{-1}X'$$

$$= A\,Var\left(\underline{y}\right)A'$$

But

$$Var\left(\underline{y}\right) = Var\left(X\underline{\beta} + \underline{\epsilon}\right) = Var(\underline{\epsilon}) = \sigma^2 I_n$$

Therefore

$$Var\left(\hat{\underline{\beta}}\right) = (X'X)^{-1}X'Var\left(\underline{y}\right)\,X(X'X)^{-1}$$

$$= (X'X)^{-1}X'\sigma^2 I_n\,X(X'X)^{-1}$$

$$= \sigma^2(X'X)^{-1}X'X(X'X)^{-1} = \sigma^2(X'X)^{-1}$$

Variance-Covariance Matrix of $\hat{\beta}$

For the simple regression model,

$$(X'X)^{-1} = \frac{1}{n\sum_{i=1}^{n} x_i^2 - (n\bar{x})^2} \begin{pmatrix} \sum_{i=1}^{n} x_i^2 & -n\bar{x} \\ -n\bar{x} & n \end{pmatrix}$$

therefore

$$Var(\hat{\beta}_0) = \sigma^2 \frac{\sum_{i=1}^{n} x_i^2}{n\left(\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2\right)}$$

$$Var(\hat{\beta}_1) = \sigma^2 \frac{1}{\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2}$$

$$Cov(\hat{\beta}_0, \hat{\beta}_1) = -\sigma^2 \frac{\bar{x}}{\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2}$$

1.9 Estimate of $V(\underline{\hat{\beta}})$

Recall: σ^2 can be estimated by the MSE.

In the simple regression model,

$$MSE = \frac{1}{n-2}\left(\underline{y}'\underline{y} - \underline{\hat{\beta}}'X'\underline{y}\right)$$

Let

$$\hat{\sigma}^2 = MSE = \frac{1}{n-2}\left(\underline{y}'\underline{y} - \underline{\hat{\beta}}'X'\underline{y}\right)$$

Estimate of $Var\left(\underline{\hat{\beta}}\right) = \hat{\sigma}^2(X'X)^{-1}$

Estimate of $V(\underline{\hat{\beta}})$

Therefore

$$\widehat{Var}(\hat{\beta}_0) = \hat{\sigma}^2 \frac{\sum x_i^2}{n\left(\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2\right)}$$

$$\widehat{Var}(\hat{\beta}_1) = \hat{\sigma}^2 \frac{1}{\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2}$$

$$\widehat{Cov}(\hat{\beta}_0, \hat{\beta}_1) = -\hat{\sigma}^2 \frac{\bar{x}}{\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2}$$

1.10 Example 1 (Continued)

$$\hat{\sigma}^2 = MSE = 4.86$$

and

$$(X'X)^{-1} = \frac{1}{173.34}\begin{pmatrix} 244.26 & -45 \\ -45 & 9 \end{pmatrix} = \begin{pmatrix} 1.4901 & -0.2596 \\ -0.2596 & 0.0519 \end{pmatrix}$$

Therefore

$$\widehat{Var}\left(\underline{\hat{\beta}}\right) = \hat{\sigma}^2 (X'X)^{-1} = 4.86\begin{pmatrix} 1.4901 & -0.2596 \\ -0.2596 & 0.0519 \end{pmatrix}$$

Note: $\sqrt{\widehat{Var}(\hat{\beta}_1)}$ denotes the standard error of the estimator of β_1. Hence $\text{s.e.}(\hat{\beta}_1) = \sqrt{4.86(0.0519)} = 0.5022$

1.11 Confidence Intervals for β_0 and β_1

We know that $\hat{\beta}_0 \sim N\left(\beta_0, Var\left(\hat{\beta}_0\right)\right)$

Therefore

$$\frac{\hat{\beta}_0 - \beta_0}{\sqrt{Var\left(\hat{\beta}_0\right)}} \sim N(0,1)$$

Hence

$$\frac{\hat{\beta}_0 - \beta_0}{\sqrt{\widehat{Var}\left(\hat{\beta}_0\right)}} \sim t(n-2) \text{ for SRM}$$

Confidence Intervals for β_0 and β_1 (Continued)

Note: For the simple regression model, the degrees of freedom is n – 2 for the distribution of $\hat{\beta}_i$, i = 0, 1.

Therefore a 100(1 – α)% confidence interval for β_0 is given by

$$\hat{\beta}_0 \pm t_{\alpha/2}(n-2)\sqrt{\widehat{Var}(\hat{\beta}_0)}$$

Similarly, a 100(1 – α)% confidence interval for β_1 is given by

$$\hat{\beta}_1 \pm t_{\alpha/2}(n-2)\sqrt{\widehat{Var}(\hat{\beta}_1)}$$

1.12 Confidence region for β

It can be shown that for simple regression model

$$\frac{1}{\sigma^2}\left(\underline{\beta}-\underline{\hat{\beta}}\right)' X'X\left(\underline{\beta}-\underline{\hat{\beta}}\right) \sim \chi^2(2)$$

We also know that $\frac{(n-2)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-2)$

Therefore

$$\frac{\left(\underline{\beta}-\underline{\hat{\beta}}\right)' X'X\left(\underline{\beta}-\underline{\hat{\beta}}\right)}{2\hat{\sigma}^2} \sim F(2, n-2)$$

[Recall: $\frac{\chi^2(n)/n}{\chi^2(m)/m} \sim F(n,m)$]

Confidence region for $\underline{\beta}$ (Continued)

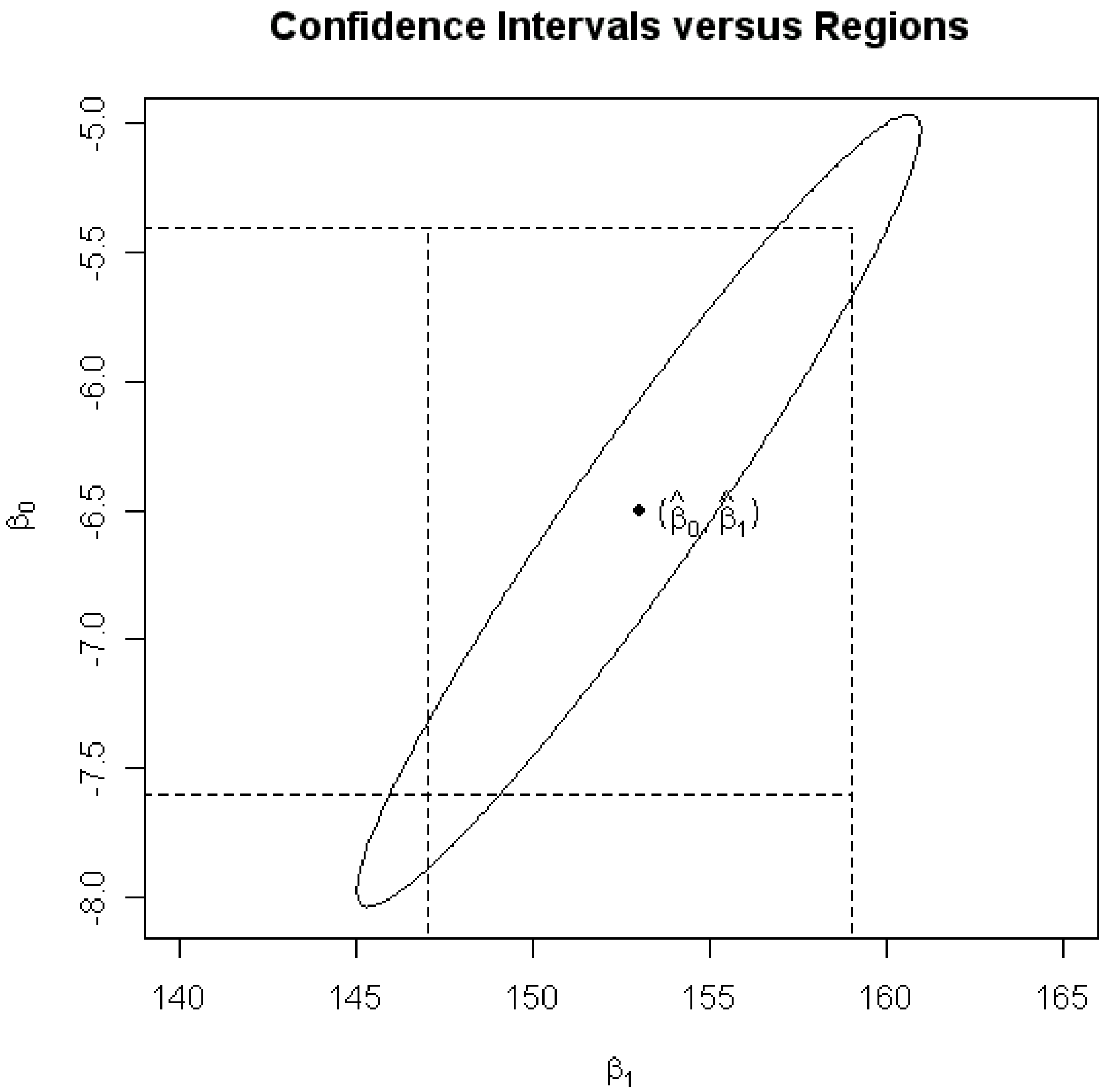
Therefore a 100(1 – α)% confidence region for $\underline{\beta}$ is given by

$$\left\{\underline{\beta}: \left(\underline{\beta} - \underline{\hat{\beta}}\right)' X'X \left(\underline{\beta} - \underline{\hat{\beta}}\right) \le 2\hat{\sigma}^2 F_\alpha(2, n-2)\right\}$$

where $F_\alpha(2, n-2)$ satisfies

$$\Pr(F(2, n-2) > F_\alpha(2, n-2)) = \alpha.$$

1.13 Confidence Intervals versus Regions



1.14 Example 1 (Continued)

(i) $\hat{\beta}_0 = 153.1755,\ \widehat{Var}(\hat{\beta}_0) = 6.8482$ and $t_{0.025}(7) = 2.365$

Therefore a 95% confidence interval for β_0 is

$$\hat{\beta}_0 \pm t_{0.025}(7)\sqrt{\widehat{Var}(\hat{\beta}_0)}$$

$$= 153.1755 \pm 2.365(6.8482)^{0.5}$$

$$= 153.1755 \pm 6.1890$$

$$= (146.9865, 159.3645).$$

Example 1 (Continued)

(ii) Similarly, a 95% confidence interval for β_1 is

$$-6.3240 \pm 2.365(0.2522)^{0.5}$$

$$= -6.3240 \pm 1.1877$$

$$= (-7.5117, -5.1363)$$

(Note: The 95% confidence interval for β_1 consists only negative values, hence we can reject the hypothesis H_0: $\beta_1 = 0$ at the 5% significance level.)

Example 1 (Continued)

(iii) $F_{0.05}(2,7) = 4.74.$

A 95% confidence region for $\underline{\beta}$ is given by (β_0, β_1) satisfying

$$(\beta_0 - 153.1755 \quad \beta_1 - (-6.4324)) \begin{pmatrix} 9 & 45 \\ 45 & 244.25 \end{pmatrix} \begin{pmatrix} \beta_0 - 153.1755 \\ \beta_1 - (-6.4324) \end{pmatrix} \leq 2(4.86)(4.74)$$

Or

$$9(\beta_0 - 153.1755)^2 - 2(45)(\beta_0 - 153.1755)(\beta_1 + 6.4324) + 244.26(\beta_1 + 6.4324)^2 \leq 46.0728$$

1.15 Confidence Interval for $\mu_{Y|x0}$

Let $\underline{x}_0 = \begin{pmatrix} 1 \\ x_0 \end{pmatrix}$ and $\mu_{Y|\underline{x}_0} = \underline{x}_0{}'\underline{\beta}$

An estimate of $\mu_{Y|\underline{x}_0}$ is given by

$$\hat{\mu}_{Y|\underline{x}_0} = \underline{x}_0{}'\hat{\underline{\beta}} = \hat{\beta}_0 + \hat{\beta}_1 x_0$$

and

$$Var\left(\hat{\mu}_{Y|\underline{x}_0}\right) = Var\left(\underline{x}_0{}'\hat{\underline{\beta}}\right) = \underline{x}_0{}'Var\left(\hat{\underline{\beta}}\right)\underline{x}_0$$
$$= \sigma^2 \underline{x}_0{}'(X'X)^{-1}\underline{x}_0$$

Confidence Interval for $\mu_{Y|x0}$ (Continued)

It can be shown that

$$\hat{\mu}_{Y|\underline{x}_0} \sim N\left(\underline{x}_0{}'\underline{\beta},\ \sigma^2\underline{x}_0{}'(X'X)^{-1}\underline{x}_0\right)$$

Therefore

$$\frac{\hat{\mu}_{Y|\underline{x}_0} - \underline{x}_0{}'\underline{\beta}}{\sqrt{\hat{\sigma}^2\underline{x}_0{}'(X'X)^{-1}\underline{x}_0}} \sim t(n-2)$$

(for Simple Regression Model)

Confidence Interval for $\mu_{Y|x0}$ (Continued)

Hence a 100(1 – α)% confidence interval for $\mu_{Y|x_0}$ is given by

$$\underline{x}_0{}'\underline{\hat{\beta}} \pm t_{\alpha/2}(n-2)\sqrt{\hat{\sigma}^2\underline{x}_0'(X'X)^{-1}\underline{x}_0}$$

1.16 Prediction Interval for $Y|x_0$

Let $\hat{y}_0 = \underline{x}_0{}'\hat{\underline{\beta}}$ and $y_0 = \underline{x}_0{}'\underline{\beta} + \epsilon$

We know that

$$\hat{y}_0 \sim N\left(\underline{x}_0{}'\underline{\beta}, \sigma^2 \underline{x}_0{}'(X'X)^{-1}\underline{x}_0\right)$$

$$y_0 \sim N(\underline{x}_0{}'\underline{\beta}, \sigma^2)$$

and they are independent.

Therefore

$$\hat{y}_0 - y_0 \sim N\left(0, \sigma^2\left(1 + \underline{x}_0{}'(X'X)^{-1}\underline{x}_0\right)\right)$$

Prediction Interval for $Y|x_0$ (Continued)

Therefore

$$\frac{\hat{y}_0 - y_0}{\sqrt{\sigma^2\left(1 + \underline{x}_0'(X'X)^{-1}\underline{x}_0\right)}} \sim N(0,1)$$

Hence

$$\frac{\hat{y}_0 - y_0}{\sqrt{\hat{\sigma}^2\left(1 + \underline{x}_0'(X'X)^{-1}\underline{x}_0\right)}} \sim t(n-2)$$

Prediction Interval for $Y|x_0$ (Continued)

Therefore a 100(1 – α)% **prediction interval** for an individual response y_0 at $\underline{x}_0 = \begin{pmatrix} 1 \\ x_0 \end{pmatrix}$ is given by

$$\underline{x}_0{}'\underline{\hat{\beta}} \pm t_{\alpha/2}(n-2)\sqrt{\hat{\sigma}^2\left(1 + \underline{x}_0{}'(X'X)^{-1}\underline{x}_0\right)}$$

1.17 Example 1 (Continued)

Let $\underline{x}_0 = \begin{pmatrix} 1 \\ 5.5 \end{pmatrix}$. Then

$$\underline{x}_0'(X'X)^{-1}\underline{x}_0 = (1 \quad 5.5)\begin{pmatrix} 244.26 & -45 \\ -45 & 9 \end{pmatrix}\begin{pmatrix} 1 \\ 5.5 \end{pmatrix} = \frac{21.51}{173.34}$$

A 95% **confidence interval** for the mean response when $x = 5.5$ is

$$(1 \quad 5.5)\begin{pmatrix} 153.1755 \\ -6.3240 \end{pmatrix} \pm 2.365\sqrt{4.860\,\frac{21.51}{173.34}}$$

$$= 118.3935 \pm 2.365\sqrt{0.6031}$$

$$= 118.3935 \pm 1.836 = (116.5569, 120.2301)$$

Example 1 (Continued)

A 95% **prediction interval** for a response when $x = 5.5$ is

$$(1 \quad 5.5)\begin{pmatrix}153.1755\\-6.3240\end{pmatrix} \pm 2.365\sqrt{4.860\left(1+\frac{21.51}{173.34}\right)}$$

$$= 118.3935 \pm 2.365\sqrt{5.4631}$$

$$= 118.3935 \pm 5.5278 = (112.9304, 123.8566)$$

1.18 Example 2

An experiment to measure the microscopic magnetic relaxation time in crystals (μ sec) as a function of the strength of the external biasing magnetic field (KG) yielded the following data.

x	11.0	12.5	15.2	17.2	19.0	20.8
y	187	225	305	318	367	365

x	22.0	24.2	25.3	27.0	29.0
y	400	435	450	506	558

The summary statistics are $\Sigma\, x_i = 223.2$, $\Sigma\, y_i = 4116$, $\Sigma\, x_i^2 = 4877.5$, $\Sigma\, x_i\, y_i = 90096.1$ and $\Sigma\, y_i^2 = 1666782$.

Example 2 (Continued)

The following simple linear regression model is considered to be appropriate

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \;\; i = 1, \ldots, 11,$$

where ε_i's are independent $N(0, \sigma^2)$ random variables.

(i) Obtain the least squares estimates of β_0, β_1 and their standard errors using the matrix approach.

(ii) Test for the significance of the simple relationship by using

(a) a t-test for the slope and

(b) the ANOVA approach (i.e. F-test).

Solutions to Example 2

(i)

$$X'X = \begin{pmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{pmatrix} = \begin{pmatrix} 11 & 223.2 \\ 223.2 & 4877.5 \end{pmatrix}$$

$$(X'X)^{-1} = \frac{1}{11(4877.5) - 223.3^2} \begin{pmatrix} 4877.5 & -223.2 \\ -223.2 & 11 \end{pmatrix}$$

$$X'\underline{y} = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{pmatrix} = \begin{pmatrix} 4116 \\ 90096.1 \end{pmatrix}$$

Solutions to Example 2 (Continued)

Therefore

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

$$= \frac{1}{3834.26}\begin{pmatrix} 4877.5 & -223.2 \\ -223.2 & 11 \end{pmatrix}\begin{pmatrix} 4116 \\ 90096.1 \end{pmatrix}$$

$$= \begin{pmatrix} -8.7786 \\ 18.8735 \end{pmatrix}$$

$$SSE = \underline{y}'\underline{y} - \underline{\hat{\beta}}'X'\underline{y}$$

$$= 166782 - \begin{pmatrix} -8.7786 & 18.8735 \end{pmatrix}\begin{pmatrix} 4116 \\ 90096.1 \end{pmatrix}$$

$$= 2486.196$$

Solutions to Example 2 (Continued)

Hence $\hat{\sigma}^2 = MSE = \frac{2456.196}{11-2} = 276.244$

Since estimate of $Var\left(\hat{\underline{\beta}}\right) = \hat{\sigma}^2(X'X)^{-1}$ therefore

$$s.e.(\hat{\beta}_0) = \sqrt{276.244\left(\frac{4877.50}{3834.25}\right)} = 18.7458$$

and

$$s.e.(\hat{\beta}_1) = \sqrt{276.244\left(\frac{11}{3834.25}\right)} = 0.8902$$

Solutions to Example 2 (Continued)

(ii) Test H_0: $\beta_1 = 0$ against H_1: $\beta_1 \neq 0$.

(a) The t-test for the slope is to reject H_0 if

$$\left|\frac{\hat{\beta}_1}{s.e.(\hat{\beta}_1)}\right| > t_{0.025}(9) = 2.26$$

From the data, we have

$$\left|\frac{\hat{\beta}_1}{s.e.(\hat{\beta}_1)}\right| = \frac{18.8735}{0.8902} = 21.20$$

Solutions to Example 2 (Continued)

Since the observed t-value = 21.20 which is greater than the critical value $t_{0.025}(9) = 2.262$, hence we reject H_0 at the 5% significance level.

Alternatively, the p-value = 2 $\Pr(t(9) > 21.20) = 5.42665(10)^{-9}$, which is small than 5%, hence we reject H_0.

Solutions to Example 2 (Continued)

$$SST = \underline{y}'\underline{y} - n\bar{y}^2 = 166782 - 11\left(\frac{4116}{11}\right)^2 = 126649.64$$

$$SSR = \underline{\hat{\beta}}'X'\underline{y} - n\bar{y}^2 = 124163.4271$$

$$SSE = SST - SSR = 2486.2092$$

ANOVA table

Source	SS	Df	MS	F-ratio	p-value
Regression	124163.43	1	124163.43	449.47	$5.42(10)^{-9}$
Error	2486.21	9	276.25		
Total	126649.64	10			

Since F_{obs} = 449.47 > $F_{0.05}(1,9)$ = 5.12, (or p-value = $5.42(10)^{-9}$ < 0.05) we conclude that the overall regression is significant.