The method of ordinary least squares (OLS) . Classical Normal Linear Regression Model

Presentation 3

PRF and SRF

- **Population Regression Function**

$$E(Y \mid X_i) = \beta_1 + \beta_2 X_i$$

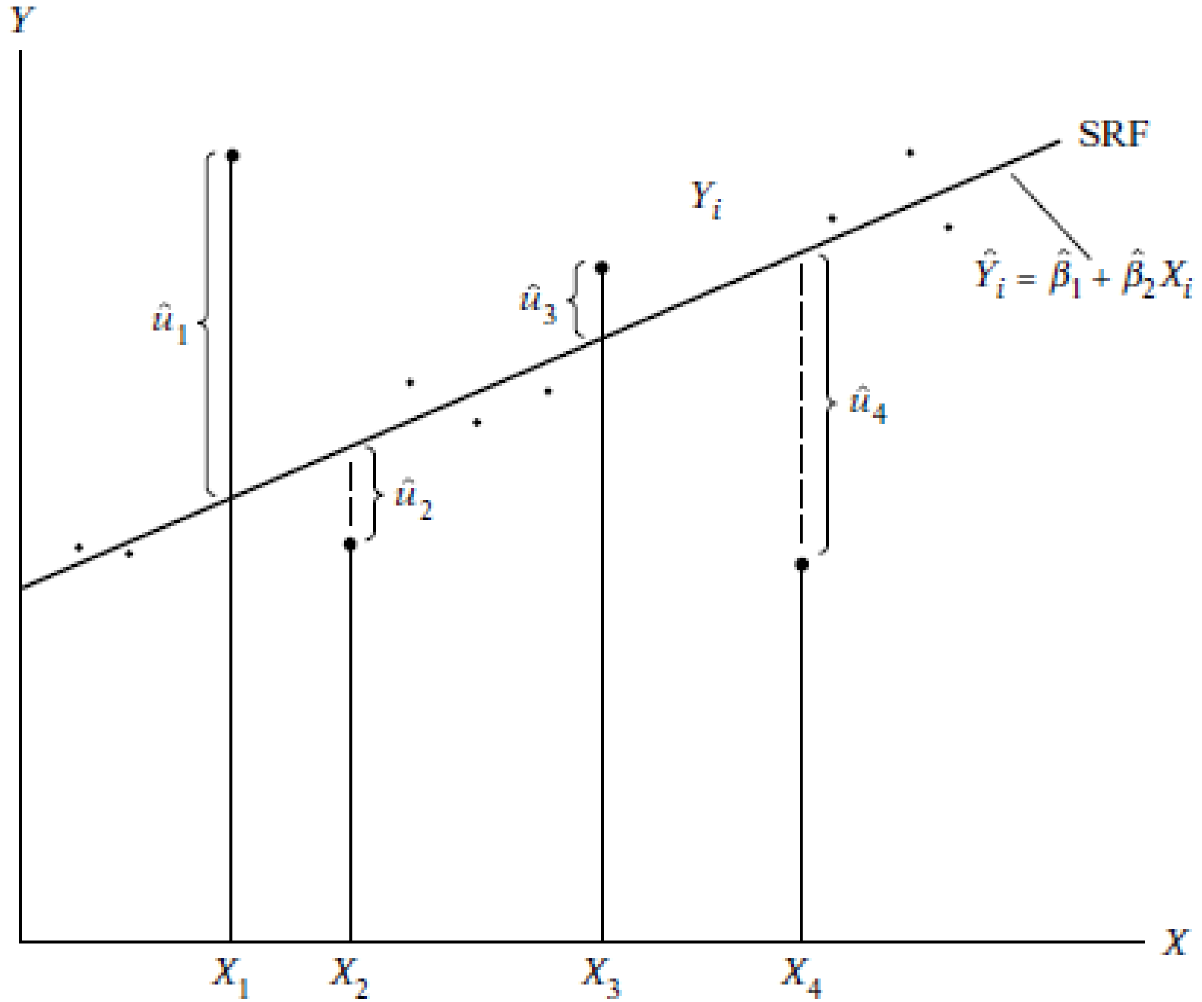
$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

- **Sample Regression Function**

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$$

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

Model Estimation



Model Estimation

$$\sum \hat{u}_i^2 = \sum (Y_i - \hat{Y}_i)^2$$
$$= \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2$$

$$\sum \hat{u}_i^2 = f(\hat{\beta}_1, \hat{\beta}_2)$$

$$\min\left(\sum \hat{u}_i^2\right) = \min\left(f\left(\hat{\beta}_1, \hat{\beta}_2\right)\right) = ?$$

Example 1

Model Estimation

$$\frac{\partial\left(\sum \hat{u}_i^2\right)}{\partial \hat{\beta}_1} = 0$$

$$\frac{\partial\left(\sum \hat{u}_i^2\right)}{\partial \hat{\beta}_2} = 0$$

$$\frac{\partial\left(\sum \hat{u}_i^2\right)}{\partial \hat{\beta}_1} = -2\sum(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i) = -2\sum \hat{u}_i$$

$$\frac{\partial\left(\sum \hat{u}_i^2\right)}{\partial \hat{\beta}_2} = -2\sum(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)X_i = -2\sum \hat{u}_i X_i$$

Model Estimation

$$\sum Y_i = n\hat{\beta}_1 + \hat{\beta}_2 \sum X_i$$

$$\sum Y_i X_i = \hat{\beta}_1 \sum X_i + \hat{\beta}_2 \sum X_i^2$$

$$\begin{aligned}\hat{\beta}_2 &= \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - \left(\sum X_i\right)^2} \\ &= \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} \\ &= \frac{\sum x_i y_i}{\sum x_i^2}\end{aligned}$$

$$\begin{aligned}\hat{\beta}_1 &= \frac{\sum X_i^2 \sum Y_i - \sum X_i \sum X_i Y_i}{n \sum X_i^2 - \left(\sum X_i\right)^2} \\ &= \bar{Y} - \hat{\beta}_2 \bar{X}\end{aligned}$$

Numerical properties of estimators

1. The regression line passes through the sample means of *Y* and *X*
2. The mean value of the estimated model is equal to the mean value of the actual *Y*
3. The mean value of the estimated residuals is zero

Numerical properties of estimators

4. The estimated residuals are uncorrelated with the predicted Y
5. The estimated residuals are uncorrelated with X_i

Example 2

Assumptions of the Classical Linear Regression Model (CLRM)

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

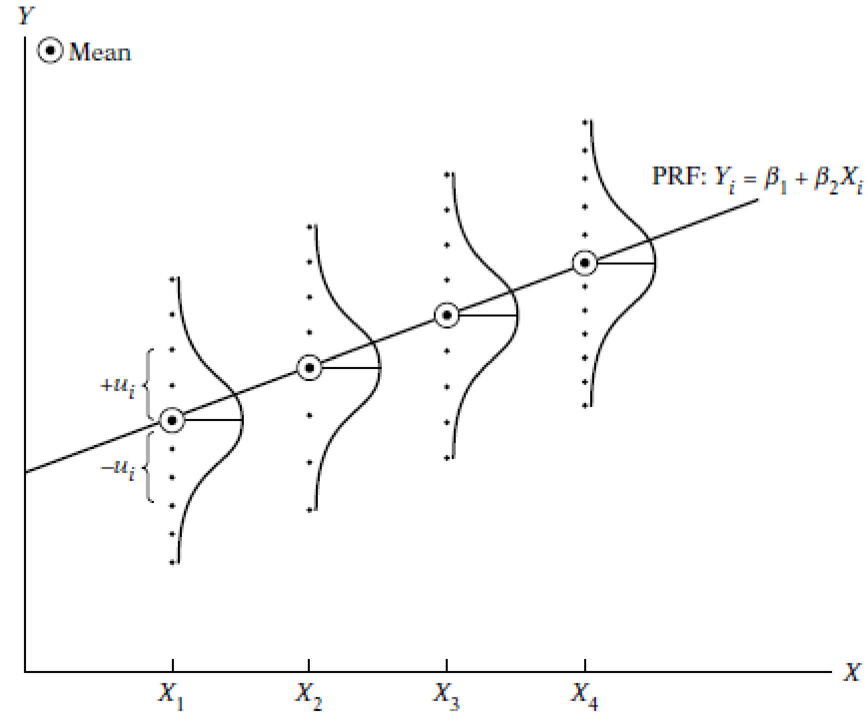
- The PRF depends on *X* and *u* and their properties determine the characteristics of $\beta_1$and β_2.

1. The regression model is linear in the parameters
2. X values are fixed in repeated sampling

Assumptions of the Classical Linear Regression Model (CLRM)

3. Given the value of X, the mean, or expected, value of the random disturbance term ui is zero

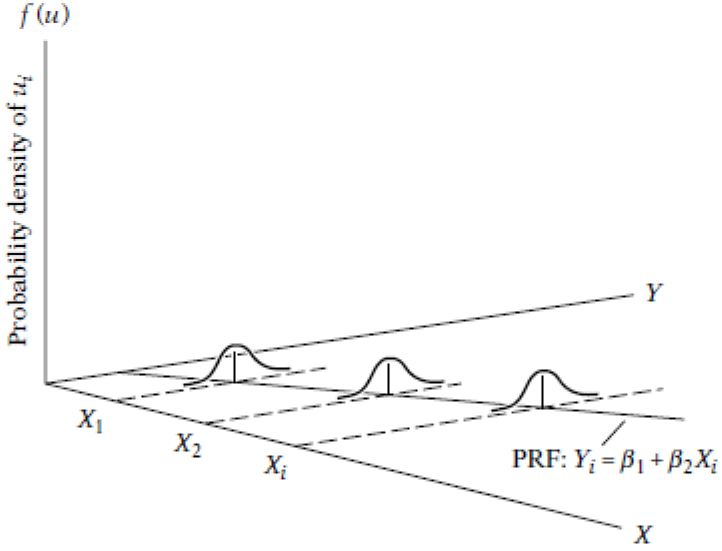
$$E(u_i \mid X_i) = 0$$

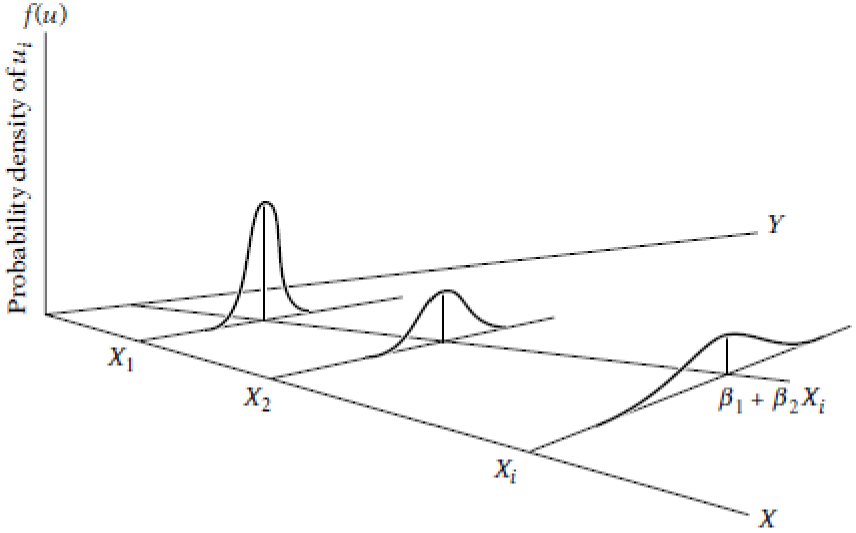


Assumptions of the Classical Linear Regression Model (CLRM)

4. Homoscedasticity or equal variance of ui

$$\operatorname{var}(u_i \mid X_i) = \sigma^2 = const$$



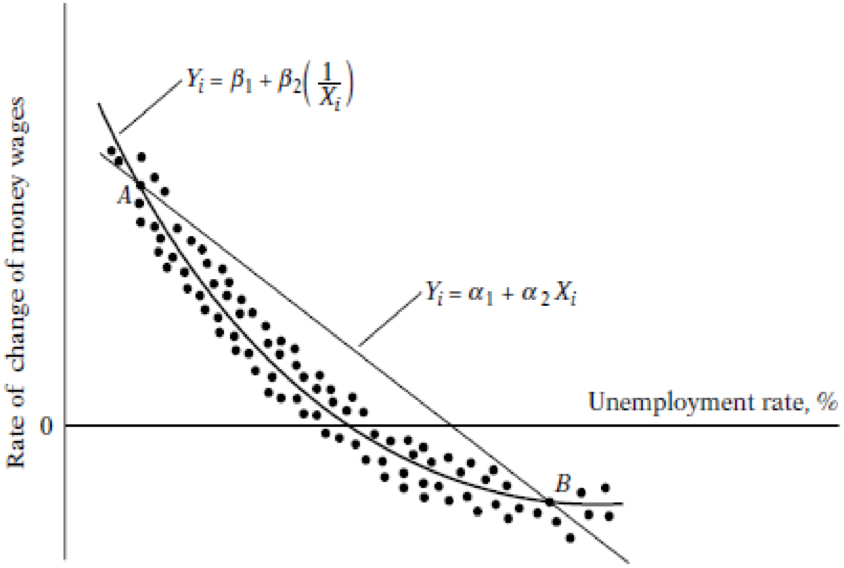


Assumptions of the Classical Linear Regression Model (CLRM)

5. No autocorrelation between the disturbances $$\operatorname{cov}\left(u_i, u_j \mid X_i, X_j\right) = 0$$
6. Zero covariance between u_i and X_i $$\operatorname{cov}\left(u_i, X_i\right) = 0$$
7. The number of observations n must be greater than the number of parameters to be estimated
8. Variability in X values $$\operatorname{var}\left(X\right) = \sigma^2 > 0$$

Assumptions of the Classical Linear Regression Model (CLRM)

9. The regression model is correctly specified



10. There is no perfect multicollinearity

Precision of the estimators

$$\operatorname{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2}$$

$$\operatorname{se}(\hat{\beta}_2) = \frac{\sigma}{\sqrt{\sum x_i^2}}$$

$$\operatorname{var}(\hat{\beta}_1) = \frac{\sum X_i^2}{n \sum x_i^2}\sigma^2$$

$$\operatorname{se}(\hat{\beta}_1) = \sqrt{\frac{\sum X_i^2}{n \sum x_i^2}}\sigma$$

Precision of the estimators

$$\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n-2}$$

$$\begin{aligned}\operatorname{cov}(\hat{\beta}_1, \hat{\beta}_2) &= -\bar{X} \operatorname{var}(\hat{\beta}_2) \\ &= -\bar{X}\left(\frac{\sigma^2}{\sum x_i^2}\right)\end{aligned}$$

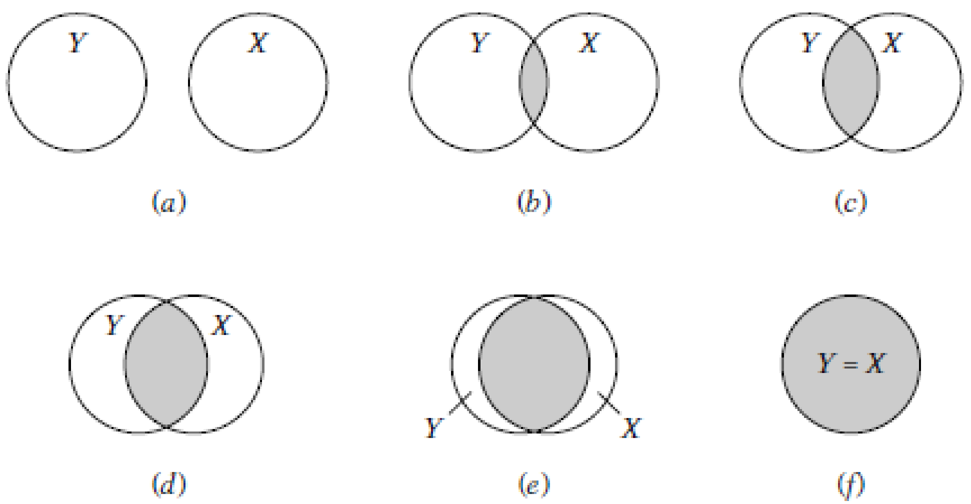
Example 3

Gauss–Markov Theorem

- Given the assumptions of the classical linear regression model, the least-squares estimators, in the class of unbiased linear estimators, have minimum variance, that is, they are BLUE (Best Linear Unbiased Estimator).

Coefficient of determination r^2

- The coefficient of determination is a summary measure that tells how well the sample regression line fits the data



The Ballentine view of r^2: (a) $r^2 = 0$; (f) $r^2 = 1$.

Coefficient of determination r^2

- The formula for r^2 computation can be derived with the help of deviation form of the model

$$y_i = \hat{y}_i + \hat{u}_i$$
$$\hat{y}_i = \hat{\beta}_2 x_i$$

$$\begin{aligned}\sum y_i^2 &= \sum \hat{y}_i^2 + \sum \hat{u}_i^2 + 2\sum \hat{y}_i \hat{u}_i \\ &= \sum \hat{y}_i^2 + \sum \hat{u}_i^2 \\ &= \hat{\beta}_2^2 \sum x_i^2 + \sum \hat{u}_i^2\end{aligned}$$

Coefficient of determination r^2

- Total sum of squares(TSS)
- Explained sum of squares(ESS)
- Residual sum of squares (RSS)

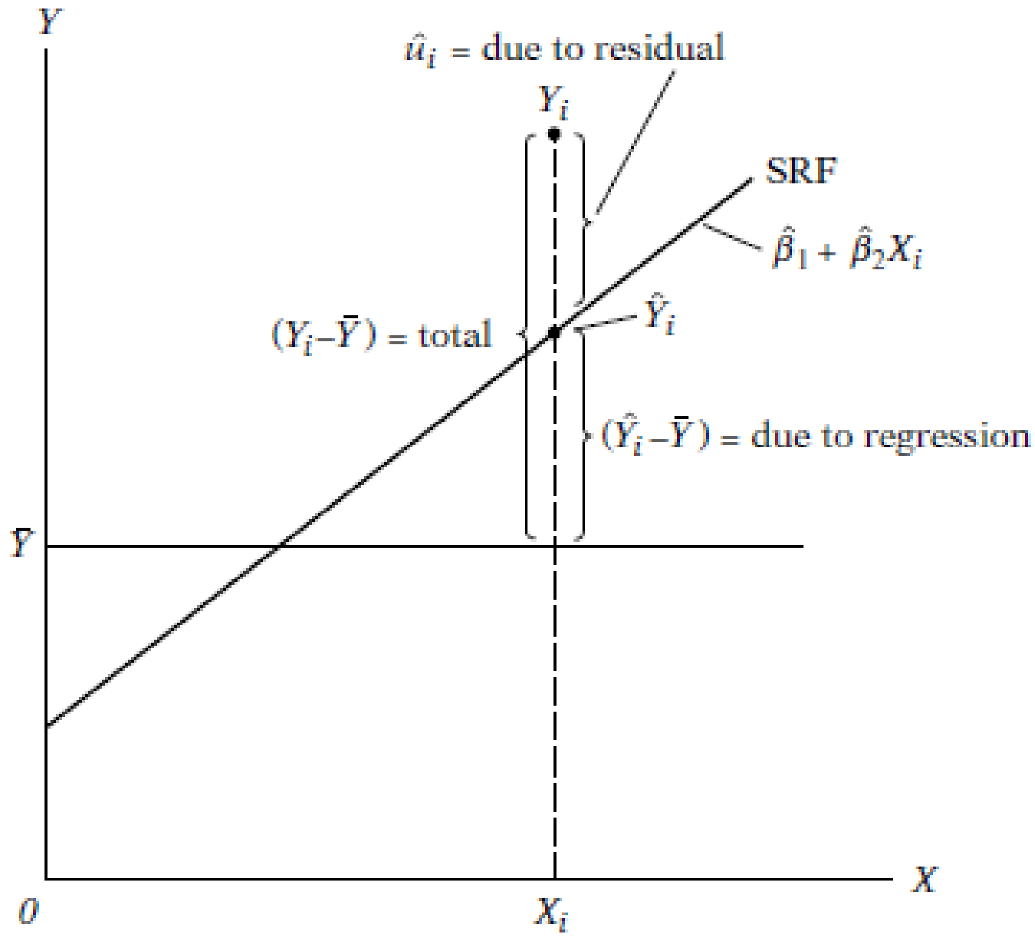
$$\sum y_i^2 = \sum (Y_i - \bar{Y})^2$$

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

$$\sum \hat{u}_i^2$$

$$\text{TSS} = \text{ESS} + \text{RSS}$$

Coefficient of determination r^2



Coefficient of determination r^2

$$1 = \frac{\text{ESS}}{\text{TSS}} + \frac{\text{RSS}}{\text{TSS}}$$

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

$$r^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = \frac{\text{ESS}}{\text{TSS}}$$

$$r^2 = 1 - \frac{\sum \hat{u}_i^2}{\sum(Y_i - \bar{Y})^2}$$

$$= 1 - \frac{\text{RSS}}{\text{TSS}}$$

$$r^2 = \frac{\text{ESS}}{\text{TSS}}$$

$$= \frac{\sum \hat{y}_i^2}{\sum y_i^2}$$

$$= \frac{\hat{\beta}_2^2 \sum x_i^2}{\sum y_i^2}$$

$$= \hat{\beta}_2^2 \left(\frac{\sum x_i^2}{\sum y_i^2} \right)$$

Coefficient of determination r^2

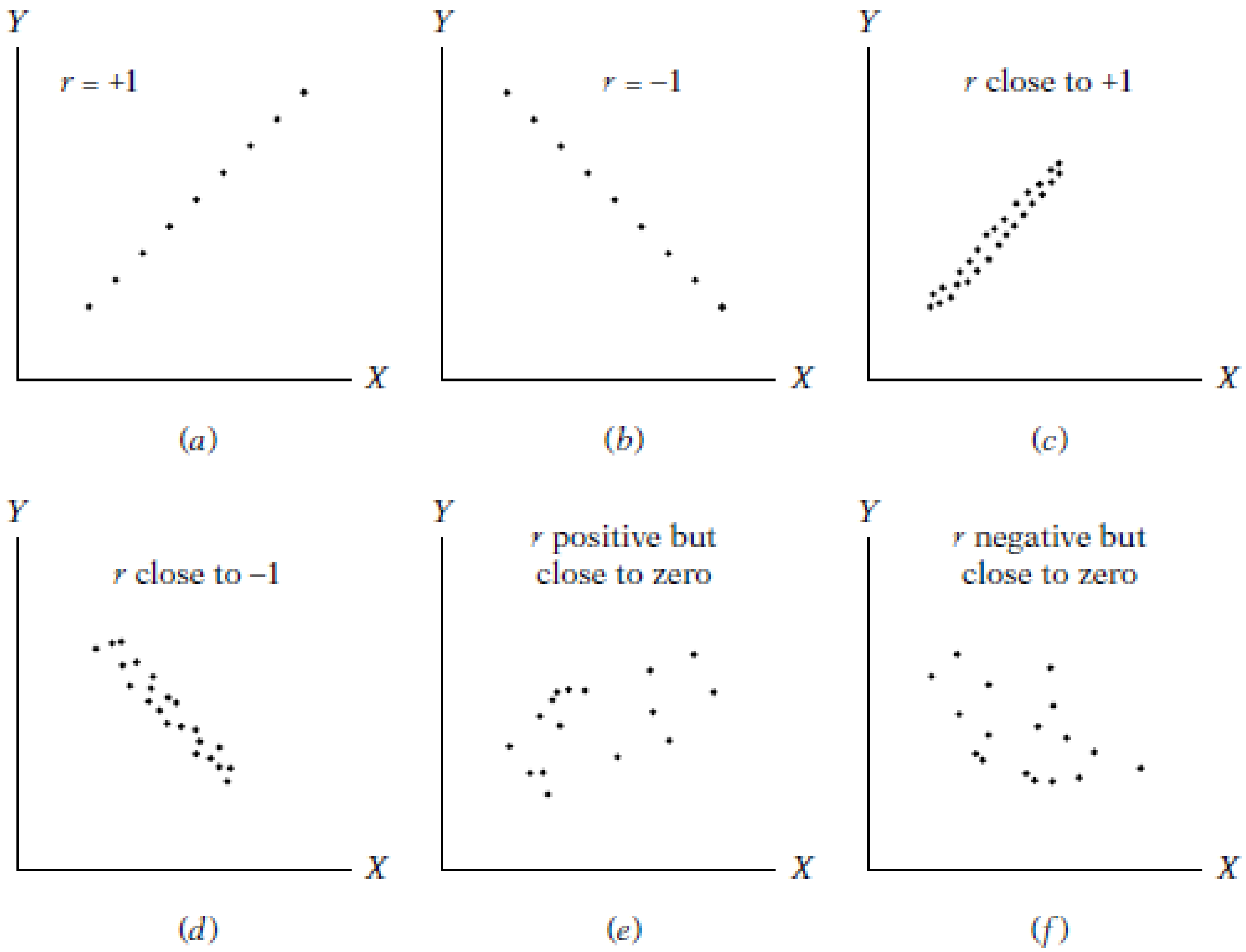
- The Coefficient of determination r^2 as a squared correlation between Y_i and the model

$$r^2 = \frac{\left[\sum (Y_i - \bar{Y})(\hat{Y}_i - \bar{Y})\right]^2}{\sum (Y_i - \bar{Y})^2 \sum (\hat{Y}_i - \bar{Y})^2}$$

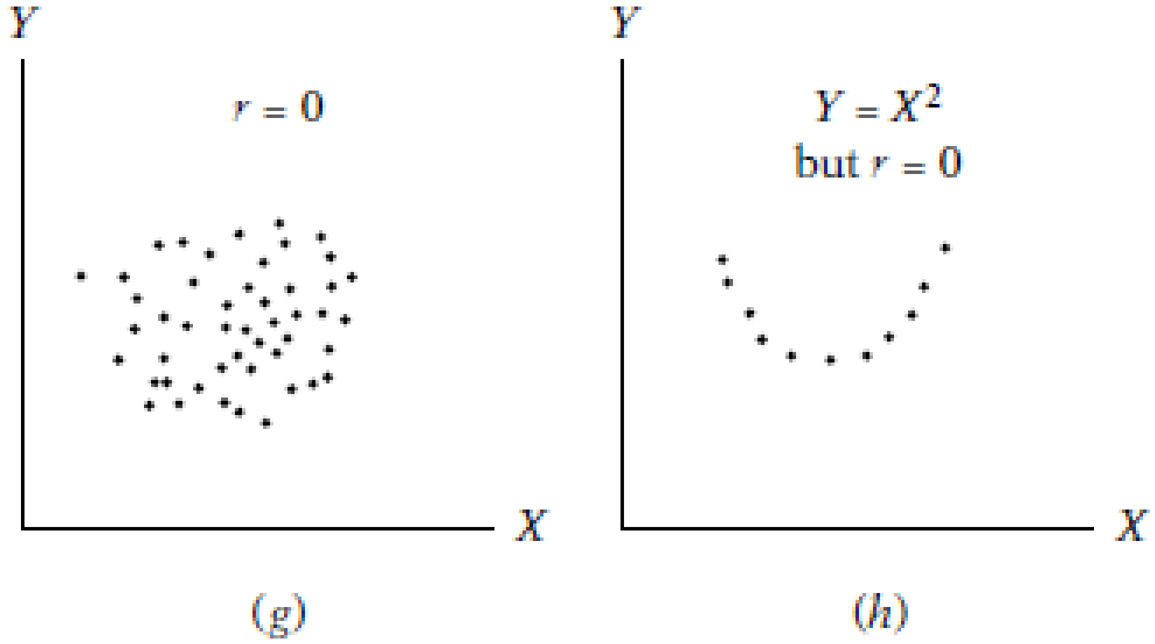
Example 4

$$r^2 = \frac{\left(\sum y_i \hat{y}_i\right)^2}{\left(\sum y_i^2\right)\left(\sum \hat{y}_i^2\right)}$$

Coefficient of determination r^2



Coefficient of determination r^2



Classical Normal Linear Regression Model

$$\hat{\beta}_2 = \sum k_i Y_i \qquad k_i = x_i / \sum x_i^2$$

$$\hat{\beta}_2 = \sum k_i(\beta_1 + \beta_2 X_i + u_i)$$

The estimator of β_2 is a linear function of the residuals u_i

Classical Normal Linear Regression Model

- According the Classical Normal Linear Regression Model the residuals u_i are with Normal distribution and:

$$E(u_i) = 0$$

$$E[u_i - E(u_i)]^2 = E(u_i^2) = \sigma^2$$

$$E\{[(u_i - E(u_i)][u_j - E(u_j)]\} = E(u_i\, u_j) = 0 \quad i \neq j$$

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

Classical Normal Linear Regression Model

- The grounds for the assumption that the residuals are normally distributed:
  - The residuals reflect the summary influence of multitude incidental factors
  - The Normal Distribution is linear, well studied and permits the usage of the connected distributions

Classical Normal Linear Regression Model

- If we assume u_i to be Normally distributed then LS Estimators have the following properties:

$$E(\hat{\beta}_1) = \beta_1 \qquad E(\hat{\beta}_2) = \beta_2$$

$$\sigma^2_{\hat{\beta}_1} = \frac{\sum X_i^2}{n\sum x_i^2}\sigma^2 \qquad \sigma^2_{\hat{\beta}_2} = \frac{\sigma^2}{\sum x_i^2}$$

$$\hat{\beta}_1 \sim N(\beta_1, \sigma^2_{\hat{\beta}_1}) \qquad \hat{\beta}_2 \sim N(\beta_2, \sigma^2_{\hat{\beta}_2})$$

Classical Normal Linear Regression Model

- If we assume u_i to be Normally distributed then LS Estimators have the following properties:

$(n-2)(\hat{\sigma}^2/\sigma^2)$ is distributed as χ^2

and $(\hat{\beta}_1, \hat{\beta}_2)$ are independently distributed from $\hat{\sigma}^2$

Maximum Likelihood Estimation

In the accordance with CNLRM

$$E(Y_i) = \beta_1 + \beta_2 X_i$$

$$\mathrm{var}(Y_i) = \sigma^2$$

or $$Y_i \sim N(\beta_1 + \beta_2 X_i, \sigma^2)$$

Maximum Likelihood Estimation

- The joint density function of $Y_1, Y_2, \ldots, Y_n$ is:

$$f(Y_1, Y_2, \ldots, Y_n \mid \beta_1 + \beta_2 X_i, \sigma^2)$$
$$= f(Y_1 \mid \beta_1 + \beta_2 X_i, \sigma^2) f(Y_2 \mid \beta_1 + \beta_2 X_i, \sigma^2) \cdots f(Y_n \mid \beta_1 + \beta_2 X_i, \sigma^2)$$

where $$f(Y_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\frac{(Y_i - \beta_1 - \beta_2 X_i)^2}{\sigma^2}\right\}$$

Maximum Likelihood Estimation

- After substitution we obtain the following likelihood function:

$$\mathrm{LF}(\beta_1, \beta_2, \sigma^2) = \frac{1}{\sigma^n\left(\sqrt{2\pi}\right)^n} \exp\left\{-\frac{1}{2}\sum \frac{(Y_i - \beta_1 - \beta_2 X_i)^2}{\sigma^2}\right\}$$

According the maximum likelihood principle we search for the maximum of this function

Maximum Likelihood Estimation

- Usually we take a logarithm om LF

$$\ln \mathrm{LF} = -n \ln \sigma - \frac{n}{2} \ln (2\pi) - \frac{1}{2} \sum \frac{(Y_i - \beta_1 - \beta_2 X_i)^2}{\sigma^2}$$

$$= -\frac{n}{2} \ln \sigma^2 - \frac{n}{2} \ln (2\pi) - \frac{1}{2} \sum \frac{(Y_i - \beta_1 - \beta_2 X_i)^2}{\sigma^2}$$

Maximum Likelihood Estimation

- Necessary conditions are:

$$\frac{\partial \ln \mathrm{LF}}{\partial \beta_1} = -\frac{1}{\sigma^2} \sum (Y_i - \beta_1 - \beta_2 X_i)(-1)$$

$$\frac{\partial \ln \mathrm{LF}}{\partial \beta_2} = -\frac{1}{\sigma^2} \sum (Y_i - \beta_1 - \beta_2 X_i)(-X_i)$$

$$\frac{\partial \ln \mathrm{LF}}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum (Y_i - \beta_1 - \beta_2 X_i)^2$$

To be set equal to zero

Maximum Likelihood Estimation

- Finally we get the known set of equations:

$$\sum Y_i = n\tilde{\beta}_1 + \tilde{\beta}_2 \sum X_i$$

$$\sum Y_i X_i = \tilde{\beta}_1 \sum X_i + \tilde{\beta}_2 \sum X_i^2$$

and estimator of the variance:

$$\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} \hat{u}_i^2$$

Example 5