Economic Interpretation of Regression

Theory and Applications

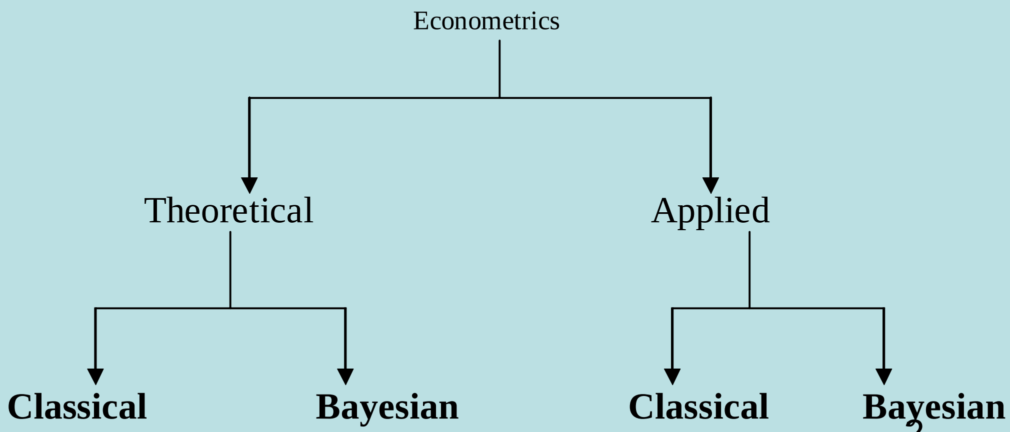
Classical and Bayesian Econometric Methods

Application of mathematical statistics to economic data for empirical support

- **Economic theory postulates a qualitative relation.**
- **Mathematical economics turns economic theory in equations.**
- **Economic statistics concerns with collecting, processing and presenting economic data.**
- **Econometricians estimate precise numerical estimates of these relations.**

- **Statement of hypothesis**
- **Specification of the mathematical model**
- **Specification of econometric model**
- **Data collection**
- **Estimation of parameters of the econometric model**
- **Hypothesis testing**
- **Forecasting or prediction**
- **Using model for control or policy analysis**

- **Mehtodology of Bayesian Econometrics**
- **Bayesian prior**
- **Sample information**
- **Posterior information**



Assumptions of a Regression Model

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

$$E[e_i] = 0 \qquad \text{var}[e_i] = \sigma^2$$

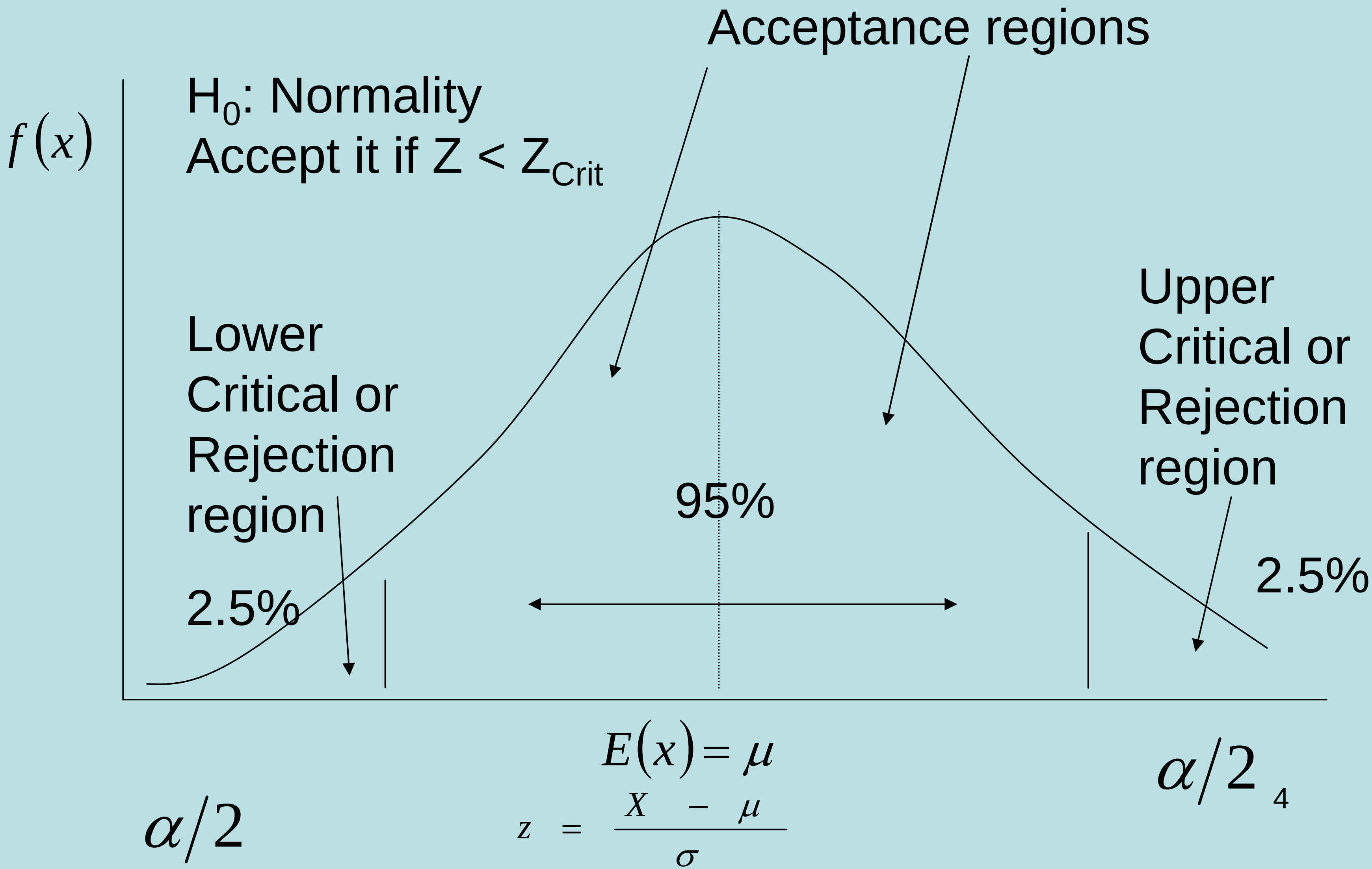
$$\text{cov}(e_i e_j) = 0 \quad \textit{for all} \quad i \neq j$$

$$E[e_i X_i] = 0$$

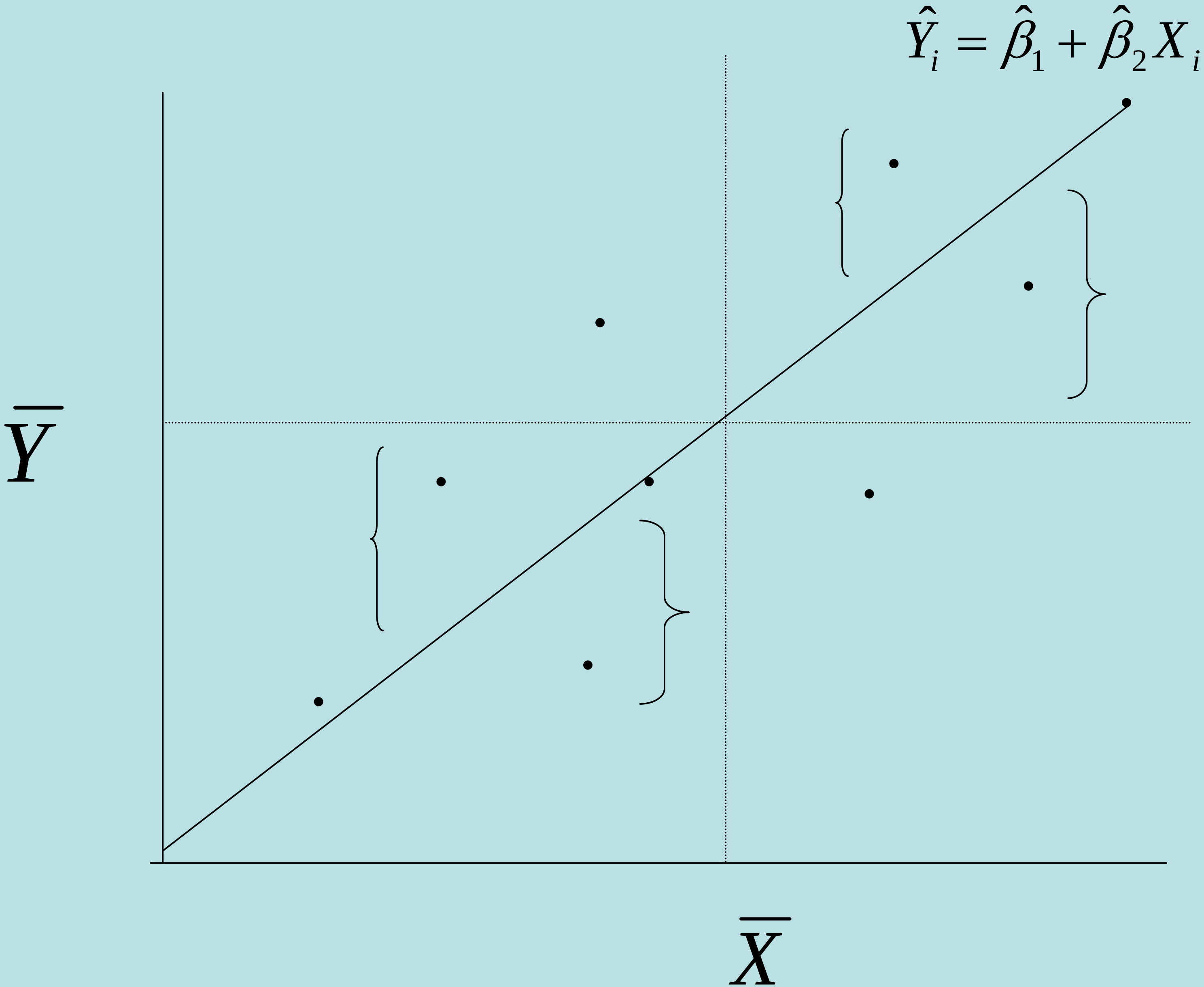
X_i is exogenous, not random

Test of Normality and Level of Significance: Two Tail Test

$$P\left(1.96 \le z \le 1.96\right) = \left(1 - \alpha\right) = 0.95$$



How Regression Can do Better than Means in Prediction ?



- The least square line is the line that best fits the data set.

- The least square line passes through the average values of variables X and Y.

- Differences between each observation and the line is represented by error terms .
  As some of them are above the line and others below the line, positive errors cancel out with the negative errors.
- Each dot in the above graph represents an observation.
- Some observations lie above the least square line and others lie below it.

These errors represent missing elements from this relationship.

Omitted variables
Measurement errors
Misspecification

Choose β_1 β_2 that Minimised the Sum of Error Square
Or best fits of the the data

$$e_i = Y_i - \beta_1 - \beta_2 X_i$$

$$S = \sum_i e_i^2 = \sum_i \left(Y_i - \beta_1 - \beta_2 X_i\right)^2$$

$$\left(-e_i\right)^2 > 0$$

$$\frac{\partial S}{\partial \beta_1} = 2\sum \left(Y_i - \beta_1 - \beta_2 X_i\right)\left(-1\right) = 0$$

$$\frac{\partial S}{\partial \beta_2} = 2\sum \left(Y_i - \beta_1 - \beta_2 X_i\right)\left(-X_i\right) = 0$$

Normal Equations and Estimators

$$\sum_i Y_i = N\beta_1 + \beta_2 \sum_i X_i$$

$$\sum_i X_i Y_i = \beta_1 \sum_i X_i + \beta_2 \sum_i X_i^2$$

$$\hat{\beta}_2 = \frac{N\sum_i X_i Y_i - \sum_i X_i \sum_i Y_i}{N\sum_i X_i^2 - \left(\sum_i X_i\right)^2}$$

$$\hat{\beta}_1 = \overline{Y} - \hat{\beta}_2 \overline{X}$$

Food expenditure and income: data and prediction

Y	X	Xy	xsquare	ysquare	Ypred	Sqpredy	prede	sqprede
4	5	20	25	16	2.866285	8.21559	1.133715	1.28531
6	8	48	64	36	5.742472	32.97598	0.257528	0.066321
7	10	70	100	49	7.65993	58.67453	-0.65993	0.435508
8	12	96	144	64	9.577388	91.72636	-1.57739	2.488153
11	14	154	196	121	11.49485	132.1315	-0.49485	0.244873
15	17	255	289	225	14.37103	206.5266	0.628967	0.395599
18	20	360	400	324	17.24722	297.4666	0.75278	0.566678
22	25	550	625	484	22.04087	485.7997	-0.04087	0.00167
Sumy	Sumx	Sumxy	sumxsq	sumysq	36.4218	Smsqpredy		smsqprede
91	111	1553	1843	1319	127.4218	1313.517	-3.9E-05	5.484111

$$\hat{\beta}_2 = \frac{N\sum_i X_i Y_i - \sum_i X_i \sum_i Y_i}{N\sum_i X_i^2 - \left(\sum_i X_i\right)^2}$$

$$\hat{\beta}_2 = \frac{8(1553) - 111(91)}{8(1843) - (111)^2} = \frac{12424 - 10101}{14744 - 12321} = \frac{2323}{2423} = 0.95873$$

$$\hat{\beta}_1 = \overline{Y} - \hat{\beta}_2 \overline{X} = \frac{91}{8} - 0.95872\frac{111}{8} = 11.375 - 0.95872(13.875) = 11.375 - 13.30224 = -1.92724$$

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i = -1.92724 + 0.95873 X_i$$

The parameters are random variables; they vary by samples.

Data in the Matrix Form

$$X'X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 5 & 8 & 10 & 12 & 14 & 17 & 20 & 25 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ 1 & 8 \\ 1 & 10 \\ 1 & 12 \\ 1 & 14 \\ 1 & 17 \\ 1 & 20 \\ 1 & 25 \end{bmatrix}$$

$$X'X = \begin{bmatrix} 8 & 111 \\ 111 & 1843 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 5 & 8 & 10 & 12 & 14 & 17 & 20 & 25 \end{bmatrix} \begin{bmatrix} 4 \\ 6 \\ 7 \\ 8 \\ 11 \\ 15 \\ 18 \\ 22 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 91 \\ 1553 \end{bmatrix}$$

Solving the OLS Model using a Matrix

$$\beta = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} N & \sum_i x_i \\ \sum_i x_i & \sum_i x_i^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_i y_i \\ \sum_i x_i y_i \end{bmatrix}$$

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

$$\begin{bmatrix} 8 & 111 \\ 111 & 1843 \end{bmatrix}^{-1} \begin{bmatrix} 91 \\ 1553 \end{bmatrix}$$

$$\left(X'X \right)^{-1} = \frac{1}{\left| X'X \right|} Adj\left(X'X \right) = \frac{1}{2423} \begin{bmatrix} 1843 & -111 \\ -111 & 8 \end{bmatrix}$$

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 8 & 111 \\ 111 & 1843 \end{bmatrix}^{-1} \begin{bmatrix} 91 \\ 1553 \end{bmatrix} = \frac{1}{2423} \begin{bmatrix} 1843 & -111 \\ -111 & 8 \end{bmatrix} \begin{bmatrix} 91 \\ 1553 \end{bmatrix} = \frac{1}{2423} \begin{bmatrix} 1843(91) - 111(1553) \\ -111(91) + (8)1553 \end{bmatrix}$$

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \frac{1}{2423} \begin{bmatrix} 167723 - 172383 \\ -10101 + 12424 \end{bmatrix} = \begin{bmatrix} -1.92736 \\ 0.95873 \end{bmatrix}$$

Elasticity around the mean of Y and X

$$\eta = \frac{\partial Y / Y}{\partial X / X} = \frac{\partial Y}{\partial X}\frac{X}{Y} = 0.95783\frac{\bar{X}}{\bar{Y}}$$

$$= 0.95783\frac{13.875}{11.375} = 1.1683$$

[Total variation] = [Explained variation] + [Residual variation]
df = T-1 K-1 T-K-1

$$Var(Y_i) = \sum_i \left[Y_i - \bar{Y}\right]^2 = \sum_i \left[\left(Y_i - \hat{Y}_i + \hat{e}_i\right)\right]^2$$

$$= \sum_i \left(Y_i - \hat{Y}_i\right)^2 + \sum_i \hat{e}_i^2 + 2\sum_i \left(Y_i - \hat{Y}_i\right)\hat{e}_i$$

$$Var(Y_i) = \sum_i \left(\hat{Y}_i - \bar{Y}\right)^2 + \sum_i \hat{e}_i^2$$

$$\because \sum_i \left(Y_i - \hat{Y}_i\right)\hat{e}_i = 0$$

Coefficient of determination

$$1 = \frac{\sum \left(Y_i - \overline{Y}\right)^2}{\sum \left(Y_i - \overline{Y}\right)^2} = \frac{\sum \left(\hat{Y}_i - \overline{Y}\right)^2}{\sum \left(Y_i - \overline{Y}\right)^2} + \frac{\sum \hat{e}_i^2}{\sum \left(Y_i - \overline{Y}\right)^2} = R^2 + \left(1 - R^2\right)$$

$$R^2 = \frac{\sum \left(\hat{Y}_i - \overline{Y}\right)^2}{\sum \left(Y_i - \overline{Y}\right)^2}$$

$$0 \le R^2 \le 1$$

Like Parameters R-Square is also a random variable.

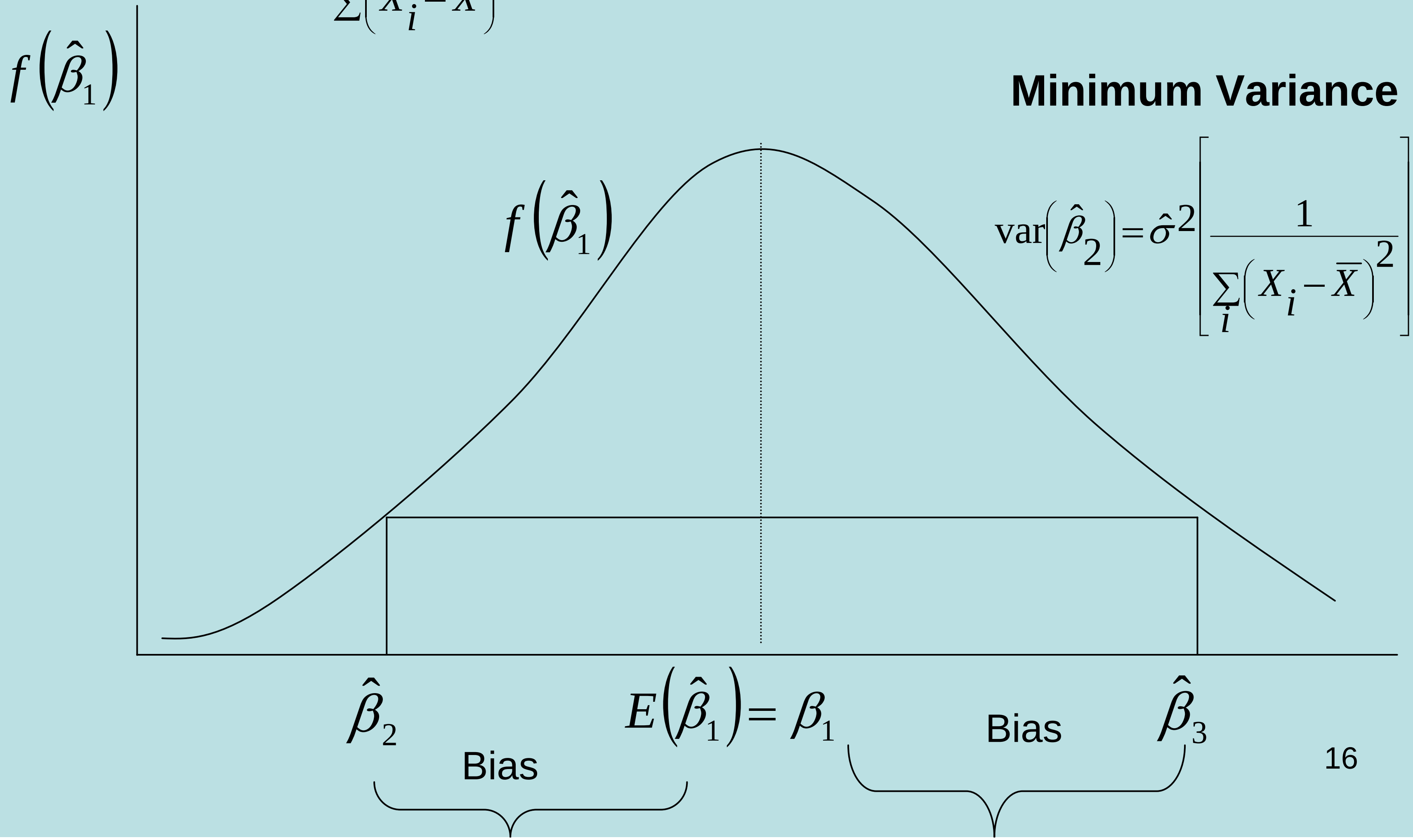
Linear, Unbiasedness and Minimum Variance Properties of an Estimator (BLUE Property)

Linearity: $\hat{\beta}_2 = \dfrac{\sum\left(X_i - \overline{X}\right)Y_i}{\sum\left(X_i - \overline{X}\right)^2} = \sum w_i Y_i$

Unbiasedness $E\left(\hat{\beta}_1\right) = \beta_1$

Minimum Variance

$$\operatorname{var}\left(\hat{\beta}_2\right) = \hat{\sigma}^2 \left[\frac{1}{\sum_i \left(X_i - \overline{X}\right)^2}\right]$$



Interval Estimation and the level of significance

$P\left(t_c \le \dfrac{\hat{\beta}_k - \beta}{\sqrt{\operatorname{var}\left(\hat{\beta}_k\right)}} \le t_c \right) = 1 - \alpha$ $P\left(t \ge t_c\right) = \alpha/2$	$P\left(\hat{\beta}_k - t_c SE\left(\hat{\beta}_k\right) \le \beta_k \le \hat{\beta}_k + t_c SE\left(\hat{\beta}_k\right)\right) = 1 - \alpha$ $P\left[\hat{\beta}_k - t_c SE\left(\hat{\beta}_k\right), \beta_k + t_c SE\left(\hat{\beta}_k\right)\right]$

Hypothesis Testing About the Mean

One-tailed hypothesis test	Two-tailed hypothesis test
$H_0: \beta_k = 0$	$H_0: \beta_k = 0$
$H_A: \beta_k \geq 0$	$H_A: \beta_k \neq 0$
Get the value of $\hat{\beta}_k$	Get the value of $\hat{\beta}_k$
Get the standard error of $\hat{\beta}_k$	Get the standard error of $\hat{\beta}_k$
Compute t-ratio	Compute t-ratio
$t = \frac{\hat{\beta}_k - \beta}{SE\left(\hat{\beta}_k\right)} \sim t_{(T-K)}$	$t = \frac{\hat{\beta}_k - \beta}{SE\left(\hat{\beta}_k\right)} \sim t_{(T-K)}$
Compare it with the critical value of t from the t-distribution table.	Compare it with the critical value of t from the t-distribution table.
• reject H_0 if $t \leq t_c$	• reject H_0 if the computed t-value is greater than or equal to t_c, or less than or equal to $-t_c$.

Testing Restrictions in a Multiple Regression Model: F-test

Model $Y_t = \beta_0 + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + e_t$

Test

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

Alternative hypothesis $H_1 : \beta_1$ or β_2 or β_3

or any two of them or all are nonzero.

At least one of them is significant.

F-test for overall significance of the model

$$F = \frac{V_1/m_1}{V_2/m_2} \sim F_{(m_1, m_2)}$$

V_1 = sum of variation due to explanatory variables and

V_2 = sum of variation not explained (squared residuals)

m1 = degrees of freedom of k explanatory variables (K-1)

m2 = degrees of freedom of for residual (N-K)

Derivation of Normal Equations in a Multiple Regression Analysis

$$\frac{\partial S}{\partial \beta_1} = 2\sum\left(Y_i - \beta_1 - \beta_2 X_{1,i} - \beta_3 X_{2,i}\right)(-1) = 0$$ and

$$\frac{\partial S}{\partial \beta_2} = 2\sum\left(Y_i - \beta_1 - \beta_2 X_{1,i} - \beta_3 X_{2,i}\right)\left(-X_{1,i}\right) = 0$$

$$\frac{\partial S}{\partial \beta_3} = 2\sum\left(Y_i - \beta_1 - \beta_2 X_{1,i} - \beta_3 X_{2,i}\right)\left(-X_{2,i}\right) = 0$$

Thus normal equations are

$$\sum_i Y_i = N\beta_1 + \beta_2 \sum_i X_{1,i} + \beta_3 \sum_i X_{2,i} \qquad (2)$$

$$\sum_i X_{1,i} Y_i = \beta_1 \sum_i X_{1,i} + \beta_2 \sum_i X^2_{1,i} + \beta_3 \sum_i X_{1,i} X_{2,i} \qquad (3)$$

$$\sum_i X_{2,i} Y_i = \beta_1 \sum_i X_{2,i} + \beta_2 \sum_i X_{1,i} X_{2,i} + \beta_3 \sum_i X^2_{2,i} \qquad (4)$$

Algebraic Method

It is easier to solve this system in a deviation form defining deviation from the mean $\sum_i x_{i,i} = \sum_i \left(X_{1,i} - \bar{X}_1\right) = 0$; $\sum_i x_{2,i} = \sum_i \left(X_{2,i} - \bar{X}_2\right) = 0$; $\sum_i y_i = \sum_i \left(Y_i - \bar{Y}_i\right) = 0$

$$\sum_i x_{1,i} y_i = \beta_2 \sum_i x_{1,i}^2 + \beta_3 \sum_i x_{1,i} x_{2,i} \qquad (3')$$

$$\sum_i x_{2,i} y_i = \beta_2 \sum_i x_{1,i} x_{2,i} + \beta_3 \sum_i x_{2,i}^2 \qquad (4')$$

In order get value of β_3 eliminate β_2 by multiplying the (3') by $\sum x_{1,i} x_{2,i}$ and (4') by $\sum x_{1,i}^2$

$$\hat{\beta}_3 = \frac{\sum_i x_{1,i} x_{2,i} \sum_i x_{1,i} y_i - \sum_i x_{2,i} y_i \sum_i x_{1,i}^2}{\left(\sum_i x_{1,i} x_{2,i}\right)^2 - \sum_i x_{2,i}^2 \sum_i x_{1,i}^2}$$

Algebraic Derivation of Parameters

Use value of the $\hat{\beta}_3$ in equation (3') to get value of $\hat{\beta}_2$

$$\sum_i x_{1,i} y_i = \hat{\beta}_2 \sum_i x_{1,i}^2 + \hat{\beta}_3 \sum_i x_{1,i} x_{2,i} \qquad (3')$$

$$\Rightarrow \frac{\sum_i x_{1,i} y_i}{\sum_i x_{1,i}^2} = \hat{\beta}_2 + \hat{\beta}_3 \frac{\sum_i x_{1,i} x_{2,i}}{\sum_i x_{1,i}^2}$$

$$\hat{\beta}_2 = \frac{\sum_i x_{1,i} y_i}{\sum_i x_{1,i}^2} - \left[\frac{\sum_i x_{1,i} x_{2,i} \sum_i x_{1,i} y_i - \sum_i x_{2,i} y_i \sum_i x_{1,i}^2}{\left(\sum_i x_{1,i} x_{2,i} \right)^2 - \sum_i x_{2,i}^2 \sum_i x_{1,i}^2} \right] \frac{\sum_i x_{1,i} x_{2,i}}{\sum_i x_{1,i}^2}$$

$$\hat{\beta}_2 = \frac{\sum_i x_{1,i} x_{2,i} \sum_i x_{2,i} y_i - \sum_i x_{1,i} y_i \sum_i x_{2,i}^2}{\left(\sum_i x_{1,i} x_{2,i} \right)^2 - \sum_i x_{2,i}^2 \sum_i x_{1,i}^2}$$

The values of $\hat{\beta}_3$ and $\hat{\beta}_2$ can be used to find the value of $\hat{\beta}_1$.

$$\hat{\beta}_1 = \overline{Y} - \hat{\beta}_2 \overline{X}_1 - \hat{\beta}_3 \overline{X}_2$$

Matrix Approach

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_2 \end{bmatrix} = \underbrace{\begin{bmatrix} N & \sum_i X_{1,i} & \sum_i X_{21,i} \\ \sum_i X_{1,i} & \sum_i X_{1,i}^2 & \sum_i X_{1,i} X_{2,i} \\ \sum_i X_{21,i} & \sum_i X_{1,i} X_{2,i} & \sum_i X_{1,i}^2 \end{bmatrix}^{-1}}_{\text{X'X}} \underbrace{\begin{bmatrix} \sum_i Y_i \\ \sum_i Y_i X_{1,i} \\ \sum_i Y_i X_{2,i} \end{bmatrix}}_{\text{X'Y}}$$

$$= \left(X'X \right)^{-1} X'Y$$

Using Cramer's Rule for Deriving the Slope Parameters

$$\begin{bmatrix} \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} \sum_i x_{1,i}^2 & \sum_i x_{1,i} x_{2,i} \\ \sum_i x_{1,i} x_{2,i} & \sum_i x_{2,i}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_i y_i x_{1,i} \\ \sum_i y_i x_{2,i} \end{bmatrix} \qquad (18)$$

$$\hat{\beta}_2 = \frac{\begin{vmatrix} \sum_i y_i x_{1,i} & \sum_i x_{1,i} x_{2,i} \\ \sum_i y_i x_{2,i} & \sum_i x_{2,i}^2 \end{vmatrix}}{\begin{vmatrix} \sum_i x_{1,i}^2 & \sum_i x_{1,i} x_{2,i} \\ \sum_i x_{1,i} x_{2,i} & \sum_i x_{2,i}^2 \end{vmatrix}} = \frac{\sum_i y_i x_{1,i} \sum_i x_{2,i}^2 - \sum_i x_{1,i} x_{2,i} \sum_i y_i x_{2,i}}{\sum_i x_{1,i}^2 \sum_i x_{2,i}^2 - \left(\sum_i x_{1,i} x_{2,i} \right)^2}$$

(19)

$$\hat{\beta}_3 = \frac{\begin{vmatrix} \sum_i x_{1,i}^2 & \sum_i y_i x_{1,i} \\ \sum_i x_{1,i} x_{2,i} & \sum_i y_i x_{2,i} \end{vmatrix}}{\begin{vmatrix} \sum_i x_{1,i}^2 & \sum_i x_{1,i} x_{2,i} \\ \sum_i x_{1,i} x_{2,i} & \sum_i x_{1,i}^2 \end{vmatrix}} = \frac{\sum_i y_i x_{2,i} \sum_i x_{1,i}^2 - \sum_i x_{1,i} x_{2,i} \sum_i y_i x_{1,i}}{\sum_i x_{1,i}^2 \sum_i x_{2,i}^2 - \left(\sum_i x_{1,i} x_{2,i} \right)^2}$$

(20)

Mutlicollinearity

$$\hat{\beta}_3 = \frac{\sum_i y_i x_{2,i} \sum_i x_{1,i}^2 - \sum_i x_{1,i} x_{2,i} \sum_i y_i x_{1,i}}{\sum_i x_{1,i}^2 \sum_i x_{2,i}^2 - \left(\sum_i x_{1,i} x_{2,i} \right)^2} \qquad (7.11)$$

- consequences:

Here x1 is constant, x2 and x3 are explanatory variables. Further assume that x2 and x3 are perfectly correlated: $x_{2,i} = \lambda x_{1,i}$. Then (7.10) and (7.11) become as following:

$$\hat{\beta}_2 = \frac{\lambda^2 \sum_i y_i x_{1,i} \sum_i x_{1,i}^2 - \lambda^2 \sum_i x_{1,i}^2 \sum_i y_i x_{1,i}}{\lambda^2 \left(\sum_i x_{1,i}^2 \right)^2 - \lambda^2 \left(\sum_i x_{1,i}^2 \right)^2} = \frac{0}{0} = \infty \qquad (7.12)$$

$$\hat{\beta}_3 = \frac{\lambda \sum_i y_i x_{1,i} \sum_i x_{1,i}^2 - \lambda \sum_i x_{1,i}^2 \sum_i y_i x_{1,i}}{\lambda^2 \left(\sum_i x_{1,i}^2 \right)^2 - \lambda^2 \left(\sum_i x_{1,i}^2 \right)^2} = \frac{0}{0} = \infty \qquad (7.13)$$

This is the proof of the fact that when two variables are exactly correlated to each other the least square procedure completely breaks down.

What is hetrosckedasticity?

- LS assumption: variance of e_i is constant $\text{var}\left[e_i\right] = \sigma^2$ for every ith observation, $\text{var}\left(\hat{\beta}_2\right) = \hat{\sigma}^2 \left[\frac{1}{\sum_i \left(x_i - \bar{x}\right)^2}\right]$ but it is possible that

$$\sigma_i^2 = \sigma^2 x_i$$

Causes: Learning, growth, improved data collection, outliers, omitted variables;

What is autocorrelation

Assumption behind the OLS

$$\operatorname{cov}(e_i e_j) = 0 \quad for\ all \quad i \neq j$$

Autocorrelation exists when

$$\operatorname{cov}(e_i e_j) \neq 0 \quad for\ all \quad i \neq j$$

$$e_i = \rho e_{i-1} + v_i \qquad v_i \sim N\left(0, \sigma^2\right)$$

ρ **correlation coefficient between –1 and 1**