Simple Linear Regression and Correlation

Shakeel Nouman

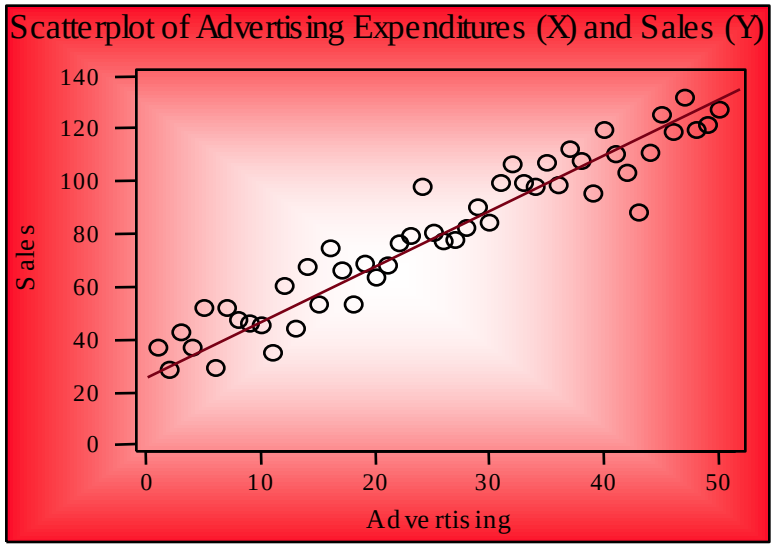
M.Phil Statistics

10 Simple Linear Regression and Correlation

- Using Statistics
- The Simple Linear Regression Model
- Estimation: The Method of Least Squares
- Error Variance and the Standard Errors of Regression Estimators
- Correlation
- Hypothesis Tests about the Regression Relationship
- How Good is the Regression?
- Analysis of Variance Table and an *F* Test of the Regression Model
- Residual Analysis and Checking for Model Inadequacies
- Use of the Regression Model for Prediction
- Summary and Review of Terms

10-1 Using Statistics

This **scatterplot** locates pairs of observations of advertising expenditures on the x-axis and sales on the y-axis. We notice that:



✓ Larger (smaller) values of sales ***tend*** to be associated with larger (smaller) values of advertising.

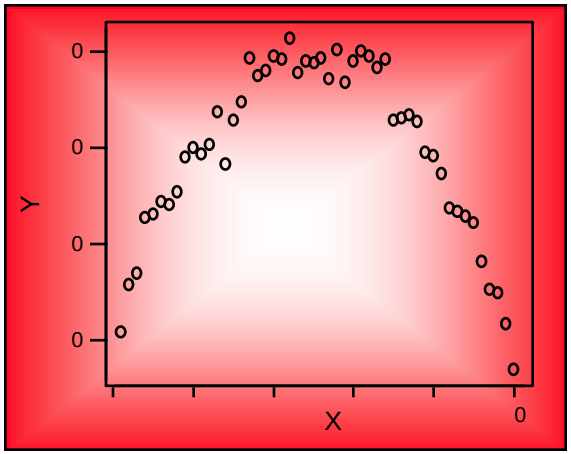
✓ The scatter of points ***tends*** to be distributed around a positively sloped straight line.

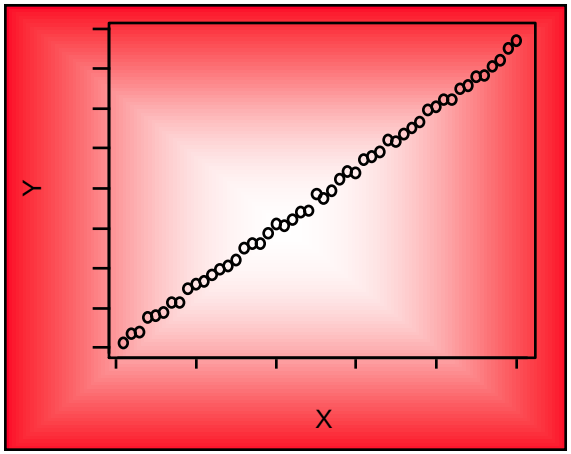
✓ The pairs of values of advertising expenditures and sales are not located exactly on a straight line.

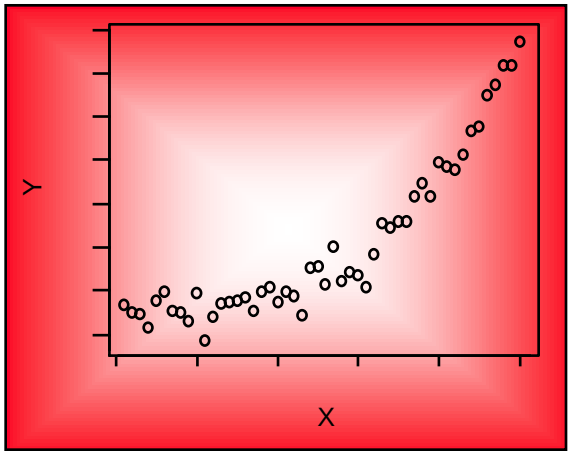
✓ The scatter plot reveals a more or less strong ***tendency*** rather than a precise ***linear relationship.***

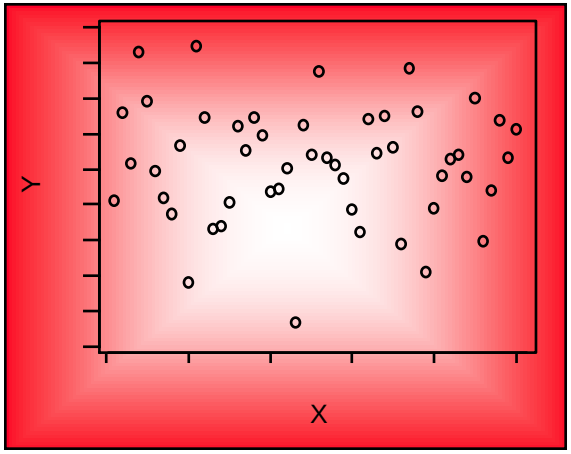
✓ The line represents the nature of the relationship on ***average***.

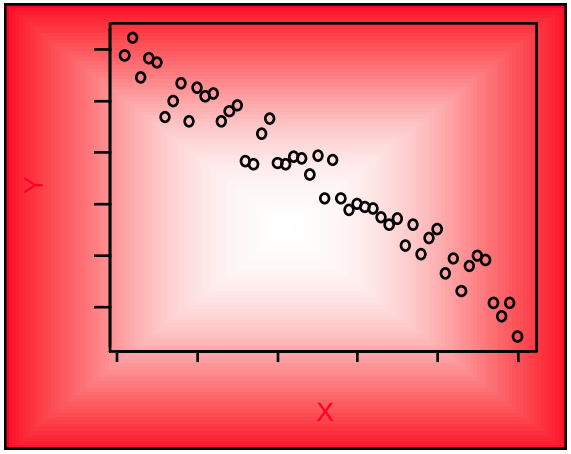
Examples of Other Scatterplots

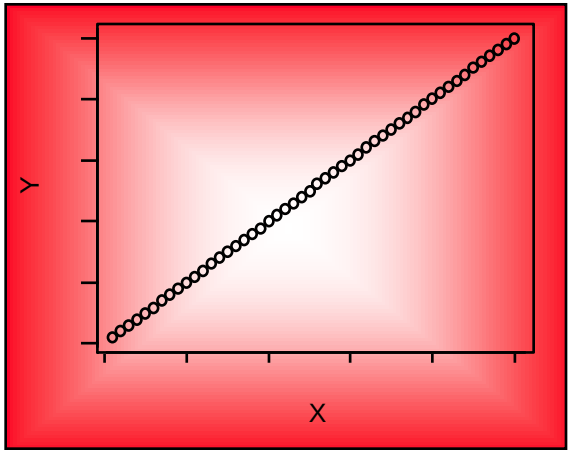








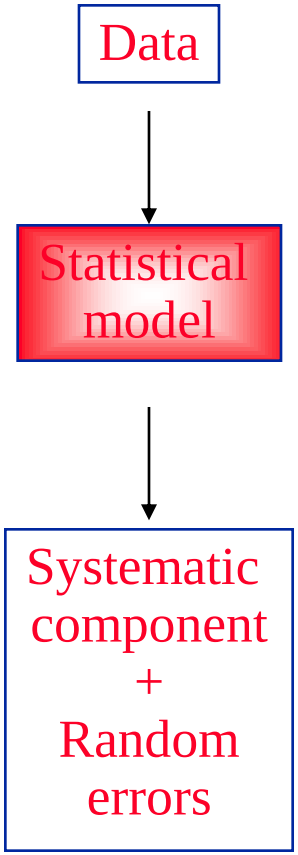




Model Building

The inexact nature of the relationship between advertising and sales suggests that a **statistical model** might be useful in analyzing the relationship.

A statistical model separates the **systematic component** of a relationship from the **random component**.



In ANOVA, the systematic component is the variation of means between samples or treatments (SSTR) and the random component is the unexplained variation (SSE).

In **regression**, the systematic component is the overall linear relationship, and the random component is the variation around the line.

10-2 The Simple Linear Regression Model

The population simple linear regression model:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Nonrandom or Systematic Component | Random Component

where

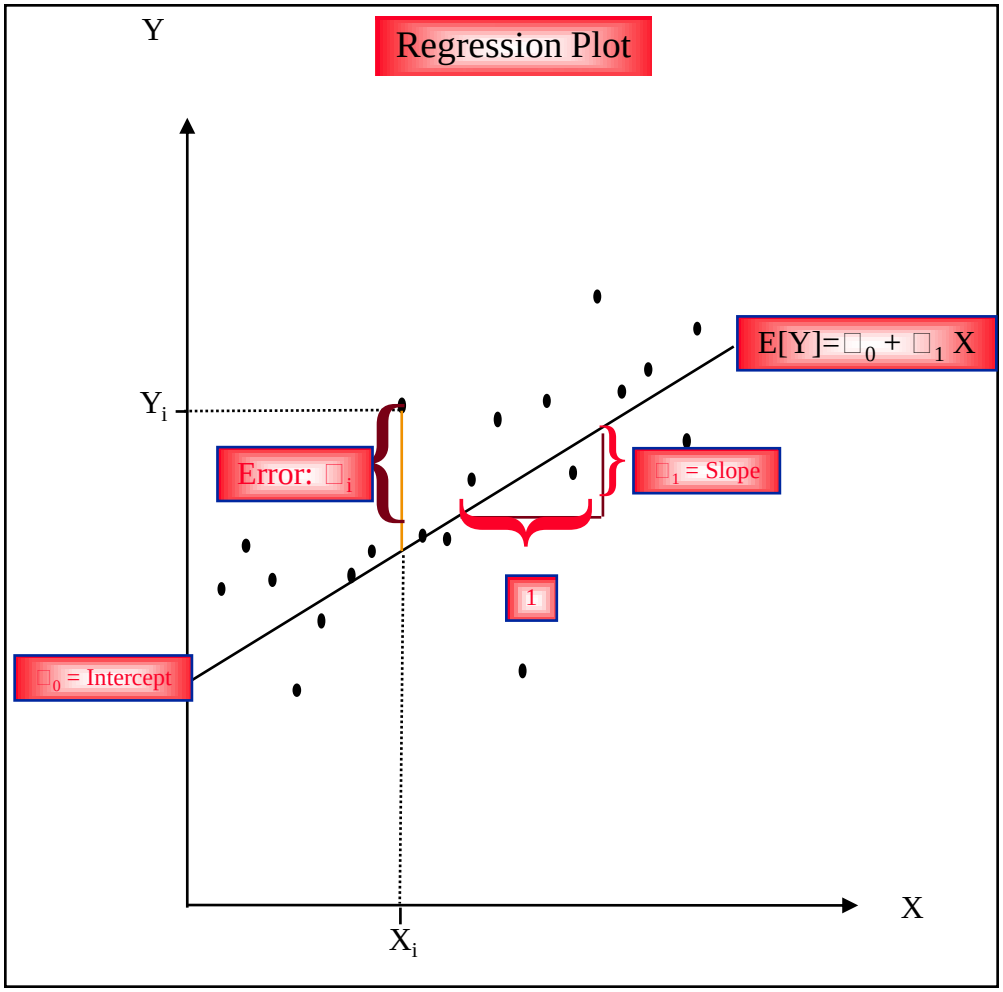
- ✓ Y is the **dependent variable**, the variable we wish to explain or predict
- ✓ X is the **independent variable**, also called the **predictor variable**
- ✓ ε is the **error term**, the only random component in the model, and thus, the only source of randomness in Y.

- ✓ β_0 is the **intercept** of the systematic component of the regression relationship.
- ✓ β_1 is the **slope** of the systematic component.

The **conditional mean** of Y:

$$E[Y|X] = \beta_0 + \beta_1 X$$

Picturing the Simple Linear Regression Model



The simple linear regression model gives an exact linear relationship between the **expected** or average value of Y, the dependent variable, and X, the independent or predictor variable:

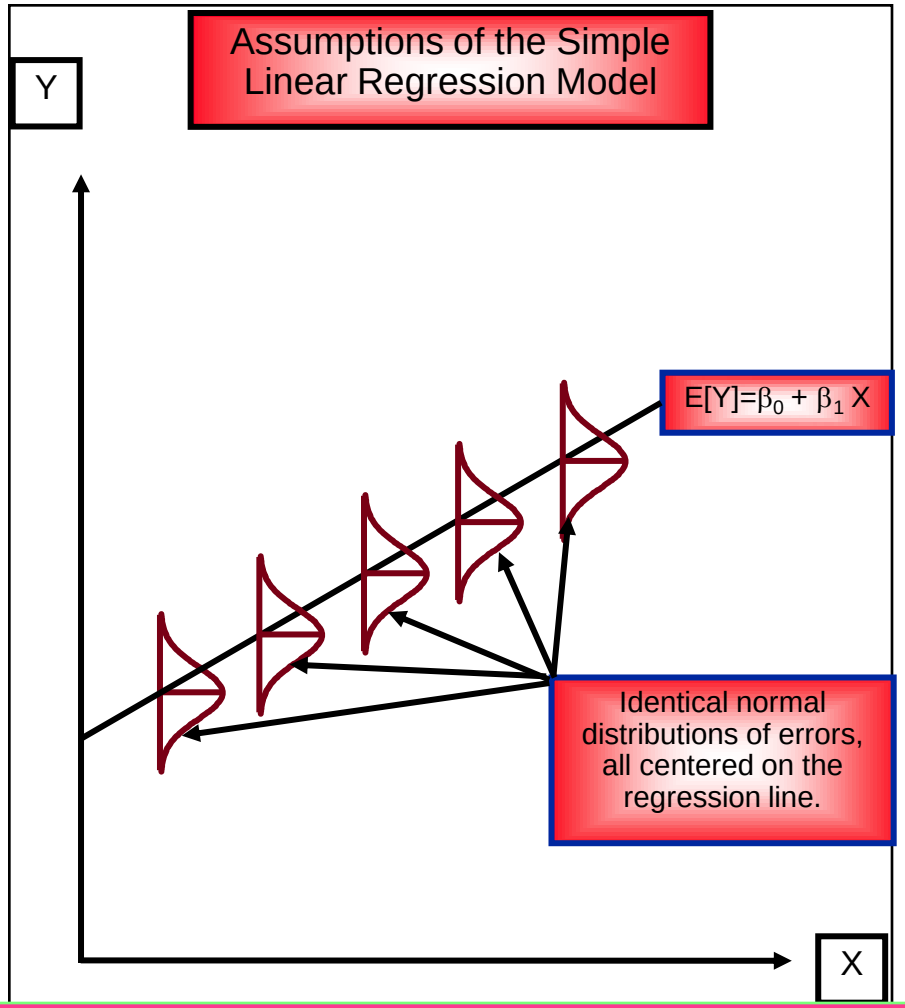
$$E[Y_i]=\beta_0 + \beta_1 X_i$$

Actual observed values of Y differ from the expected value by an unexplained or random error:

$$Y_i = E[Y_i] + \varepsilon_i$$
$$= \beta_0 + \beta_1 X_i + \varepsilon_i$$

Assumptions of the Simple Linear Regression Model

- **The relationship between X and Y is a straight-line relationship.**
- **The values of the independent variable X are assumed fixed (not random); the only randomness in the values of Y comes from the error term ε_i.**
- **The errors ε_i are normally distributed with mean 0 and variance σ^2. The errors are uncorrelated (not related) in successive observations. That is: $\varepsilon \sim N(0,\sigma^2)$**



10-3 Estimation: The Method of Least Squares

Estimation of a simple linear regression relationship involves finding estimated or predicted values of the intercept and slope of the linear regression line.

The **estimated regression equation**:

$$Y = b_0 + b_1 X + e$$

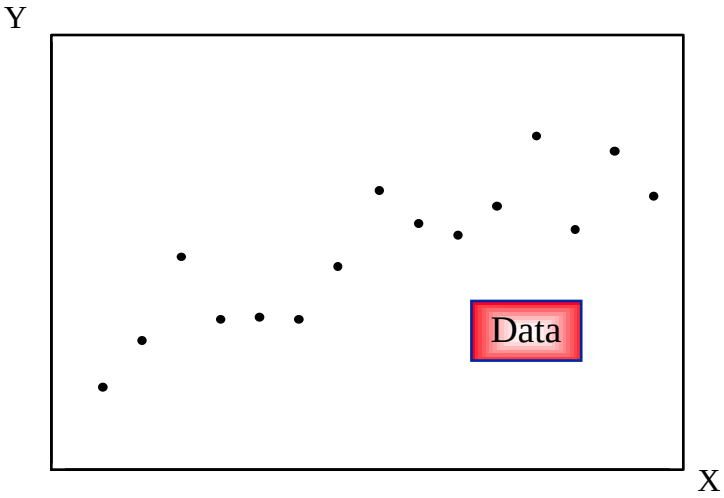
where b_0 estimates the intercept of the population regression line, β_0 ;
b_1 estimates the slope of the population regression line, β_1;
and e stands for the observed errors - the residuals from fitting the estimated regression line $b_0 + b_1 X$ to a set of n points.

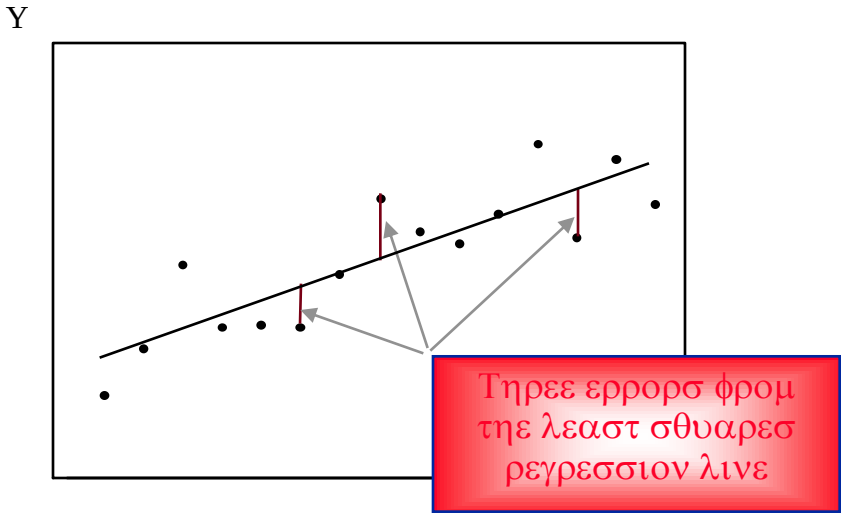
The estimated regression line:

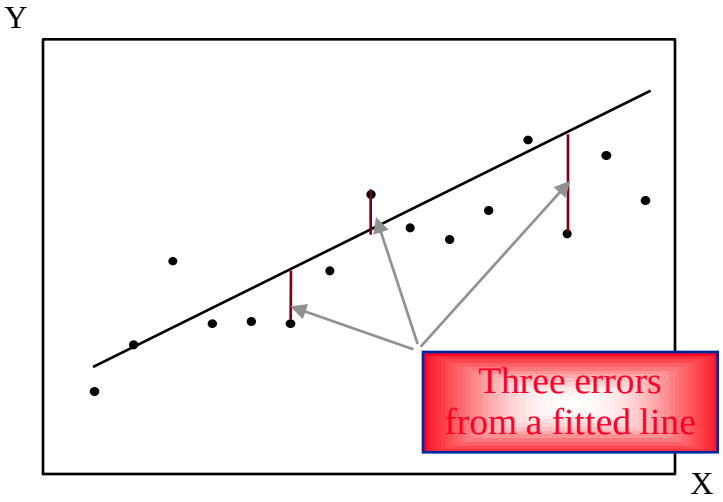
$$\hat{Y} = b_0 + b_1 X$$

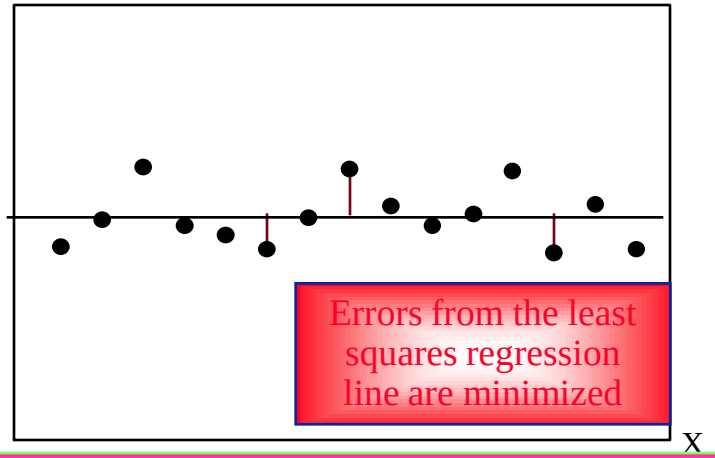
where $\hat{Y}$ (Y - hat) is the value of Y lying on the fitted regression line for a given

Fitting a Regression Line

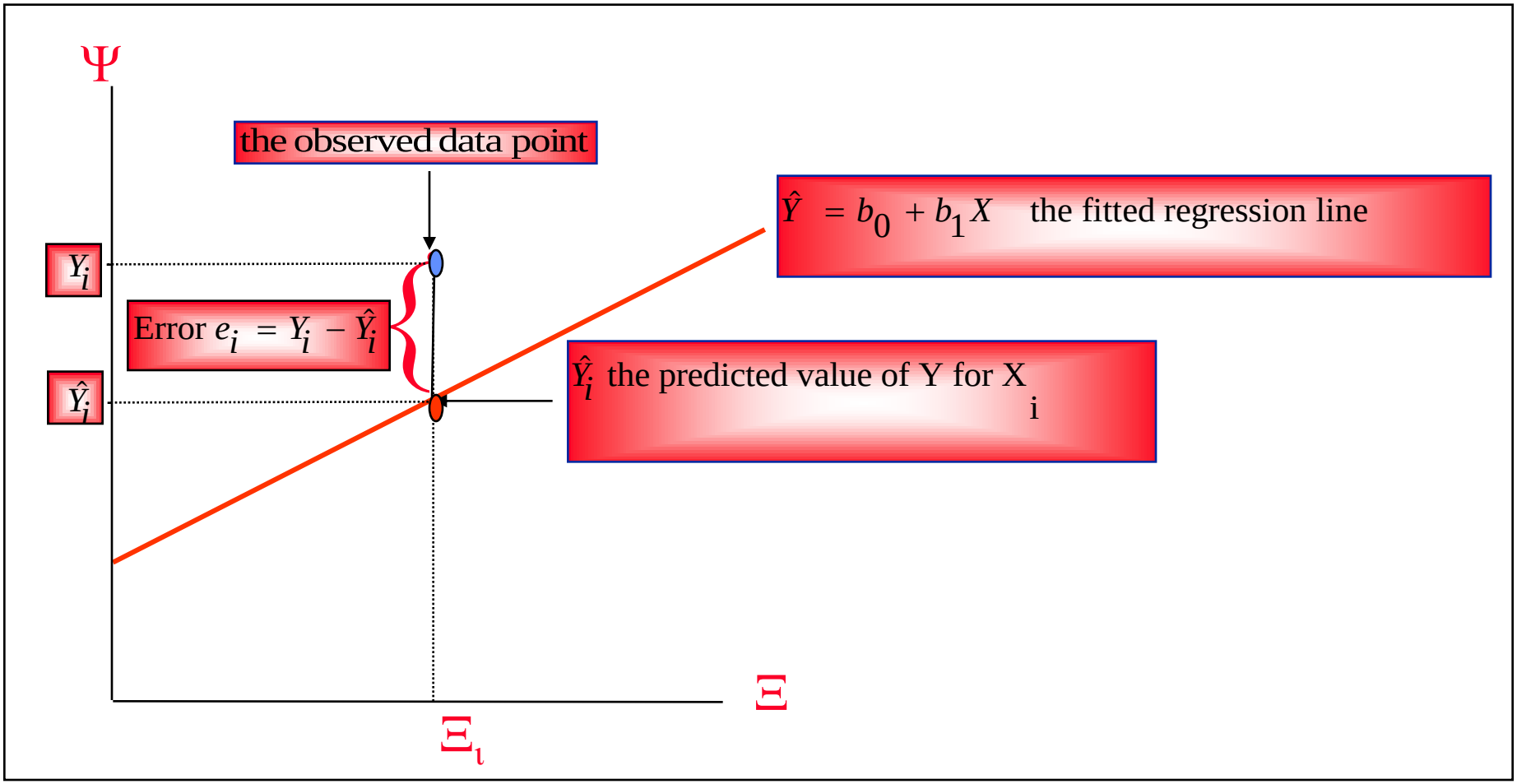








Errors in Regression



Least Squares Regression

The sum of squared errors in regression is:

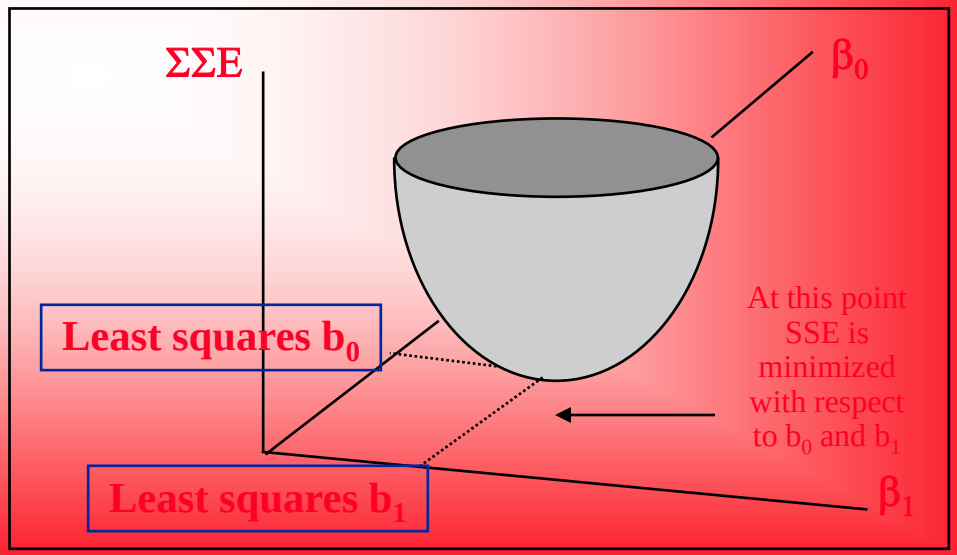
$$SSE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

The **least squares regression line** is that which *minimizes* the SSE with respect to the estimates b_0 and b_1.

The **normal equations**:

$$\sum_{i=1}^{n} y_i = nb_0 + b_1 \sum_{i=1}^{n} x_i$$

$$\sum_{i=1}^{n} x_i y_i = b_0 \sum_{i=1}^{n} x_i + b_1 \sum_{i=1}^{n} x_i^2$$



Sums of Squares, Cross Products, and Least Squares Estimators

Sums of Squares and Cross Products:

$$SS_x = \sum (x - \bar{x})^2 = \sum x^2 - \frac{\left(\sum x\right)^2}{n}$$

$$SS_y = \sum (y - \bar{y})^2 = \sum y^2 - \frac{\left(\sum y\right)^2}{n}$$

$$SS_{xy} = \sum (x - \bar{x})(y - \bar{y}) = \sum xy - \frac{\left(\sum x\right)\left(\sum y\right)}{n}$$

Least – squares regression estimators:

$$b_1 = \frac{SS_{XY}}{SS_X}$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

Example 10-1

Μιλεσ	Δολλαρσ	Μιλεσ ^{2}	Μιλεσ*Δολλαρσ
1211	1802	1466521	2182222
1345	2405	1809025	3234725
1422	2005	2022084	2851110
1687	2511	2845969	4236057
1849	2332	3418801	4311868
2026	2305	4104676	4669930
2133	3016	4549689	6433128
2253	3385	5076009	7626405
2400	3090	5760000	7416000
2468	3694	6091024	9116792
2699	3371	7284601	9098329
2806	3998	7873636	11218388
3082	3555	9498724	10956510
3209	4692	10297681	15056628
3466	4244	12013156	14709704
3643	5298	13271449	19300614
3852	4801	14837904	18493452
4033	5147	16265089	20757852
4267	5738	18207288	24484046
4498	6420	20232004	28877160
4533	6059	20548088	27465448
4804	6426	23078416	30870504
5090	6321	25908100	32173890
5233	7026	27384288	36767056
5439	6964	29582720	37877196
79,448	106,605	293,426,946	390,185,014

$$SS_x = \sum x^2 - \frac{(\sum x)^2}{n}$$

$$= 293,426,946 - \frac{79,448^2}{25} = 40,947,557.84$$

$$SS_{xy} = \sum xy - \frac{(\sum x)(\sum y)}{n}$$

$$= 390,185,014 - \frac{(79,448)(106,605)}{25} = 51,402,852.4$$

$$b_1 = \frac{SS_{XY}}{SS_X} = \frac{51,402,852.4}{40,947,557.84} = 1.255333776 \approx 1.26$$

$$b_0 = \bar{y} - b_1\bar{x} = \frac{106,605}{25} - (1.255333776)\left(\frac{79,448}{25}\right)$$

$$= 274.85$$

Template (partial output) that can be used to carry out a Simple Regression

Simple Regression — American Express Study

	Miles X	Dollars Y	Error
1	1211	1802	6.94111
2	1345	2405	441.726
3	1422	2005	-54.9343
4	1687	2511	118.402
5	1849	2332	-263.962
6	2026	2305	-513.156
7	2133	3016	63.5234
8	2253	3385	281.883
9	2400	3090	-197.651
10	2468	3694	320.987
11	2699	3371	-291.996
12	2806	3998	200.684
13	3082	3555	-588.788
14	3209	4692	388.784
15	3466	4244	-381.837

Confidence Interval for Slope

1−α	(1-α) C.I. for β_1
95%	1.25533 + or - 0.10285

Confidence Interval for Intercept

1−α	(1-α) C.I. for β_0
95%	274.85 + or - 352.368

Prediction Interval for Y

1−α	X	(1-α) P.I. for Y given X
		+ or -

Prediction Interval for E[Y\|X]

1−α	X	(1-α) P.I. for E[Y | X]
		+ or -

r^2	0.9652	Co
r	0.9824	Co
$s(b_1)$	0.04972	Sta
$s(b_0)$	170.337	Sta
s	318.158	Sta

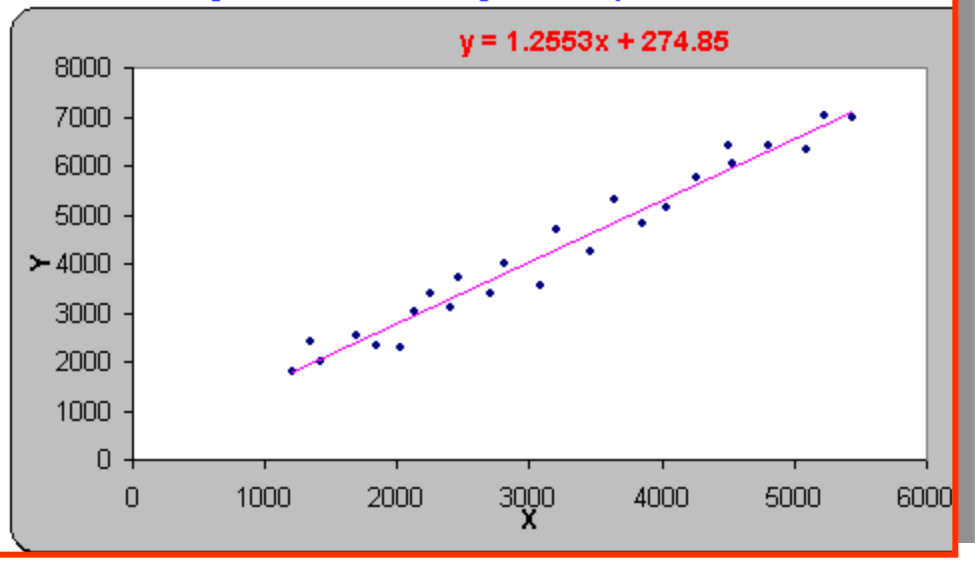
Template (continued) that can be used to carry out a Simple Regression

	A	B	C	D
19	15	3466	4244	-381.837
20	16	3643	5298	449.969
21	17	3852	4801	-309.395
22	18	4033	5147	-190.611
23	19	4267	5738	106.641
24	20	4498	6420	498.659
25	21	4533	6059	93.7223
26	22	4804	6426	120.527
27	23	5090	6321	-343.499
28	24	5233	7026	181.989
29	25	5439	6964	-138.61
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				

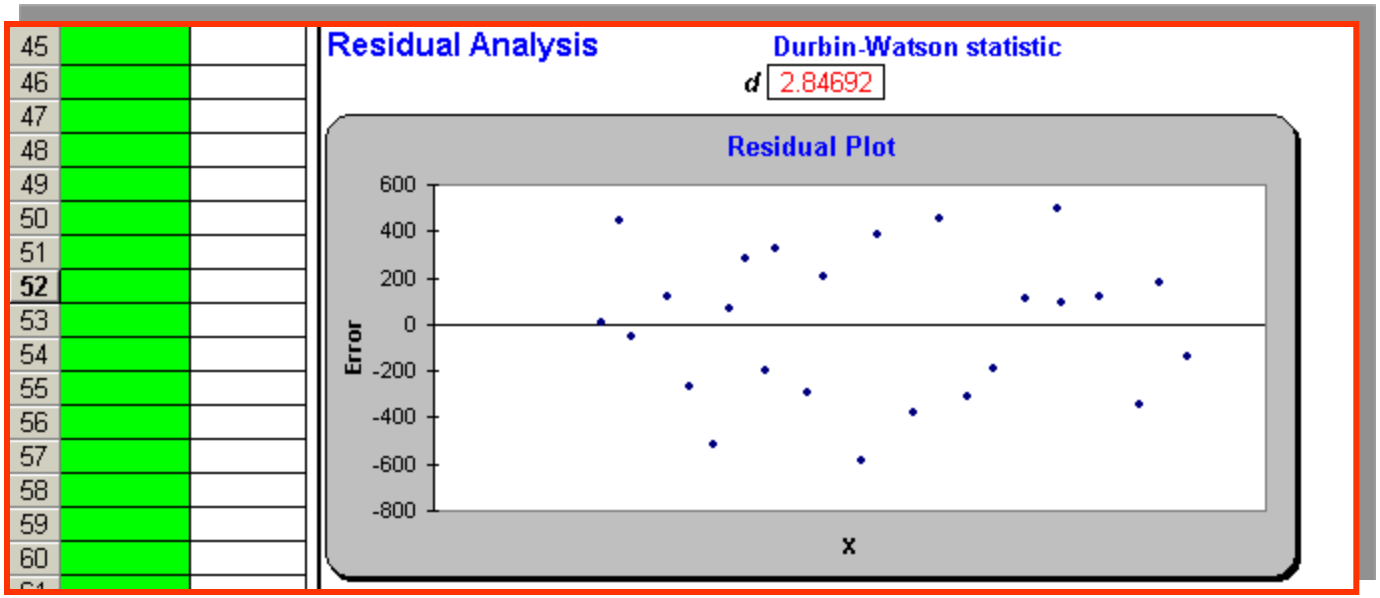
ANOVA Table

Source	SS	df	MS	F	$F_{critical}$	p-value
Regn.	6.5E+07	1	6.5E+07	637.472	4.27934	0.0000
Error	2328161	23	101224			
Total	6.7E+07	24				

Scatter Plot, Regression Line and Regression Equation

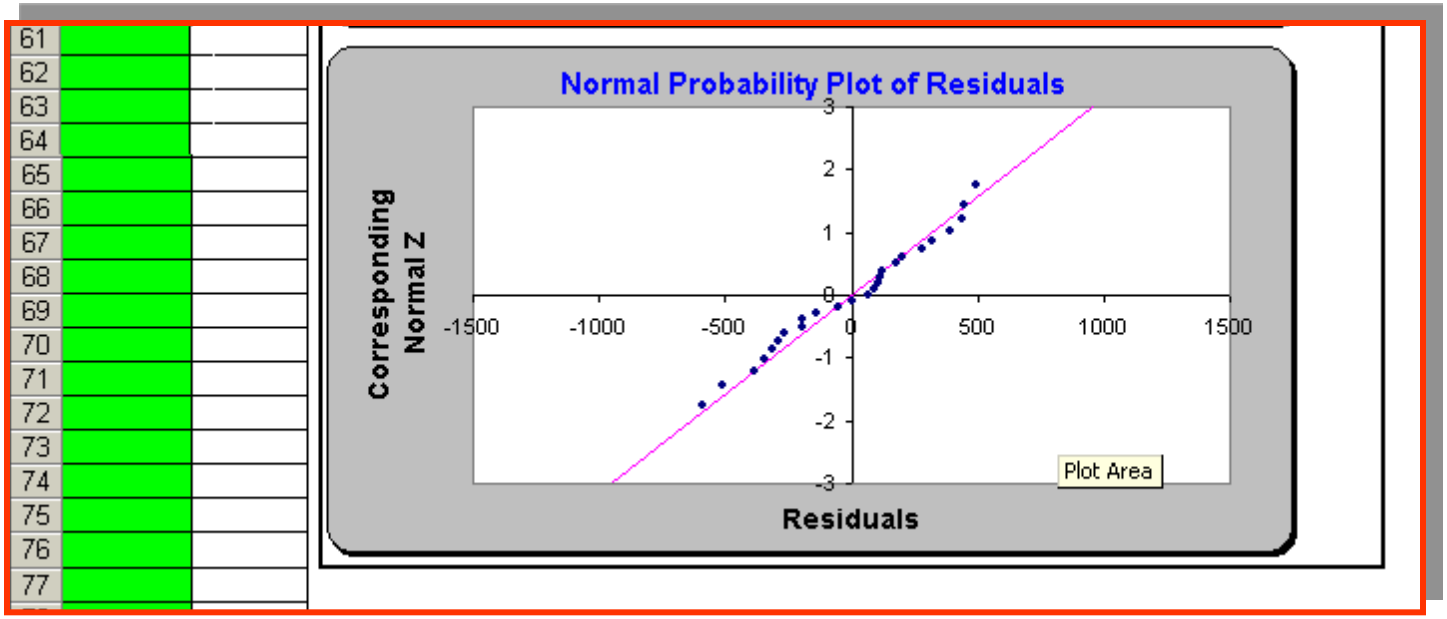


Template (continued) that can be used to carry out a Simple Regression



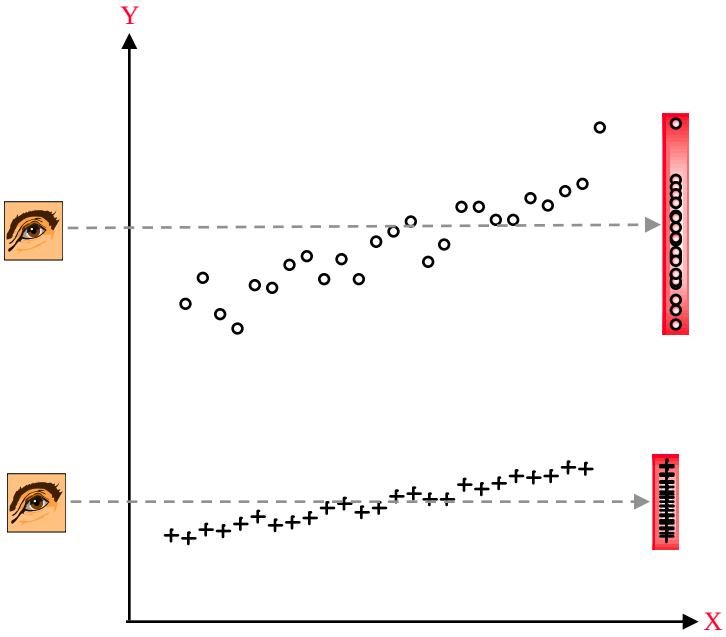
Residual Analysis. The plot shows the absence of a relationship between the residuals and the X-values (miles).

Template (continued) that can be used to carry out a Simple Regression

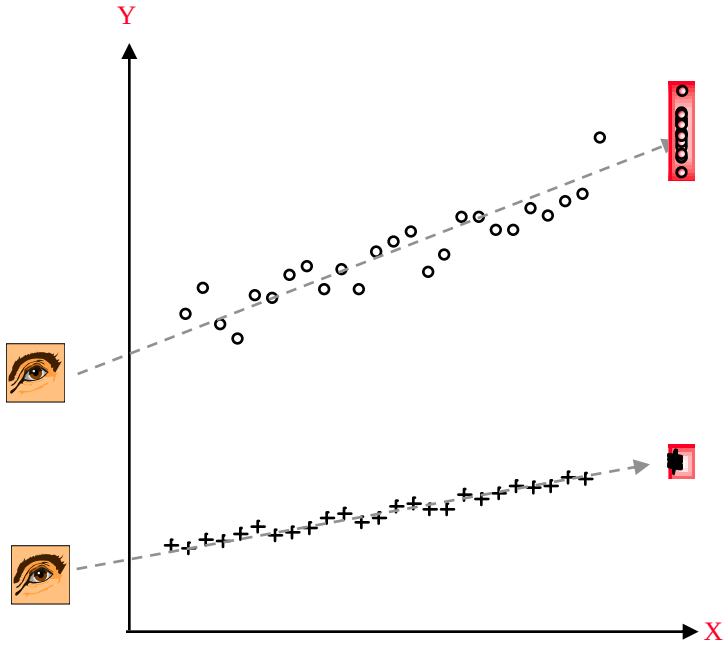


Note: The normal probability plot is approximately linear. This would indicate that the normality assumption for the errors has not been violated.

Total Variance and Error Variance



What you see when looking at the total variation of Y.



What you see when looking along the regression line at the error variance of Y.

10-4 Error Variance and the Standard Errors of Regression Estimators

Degrees of Freedom in Regression:

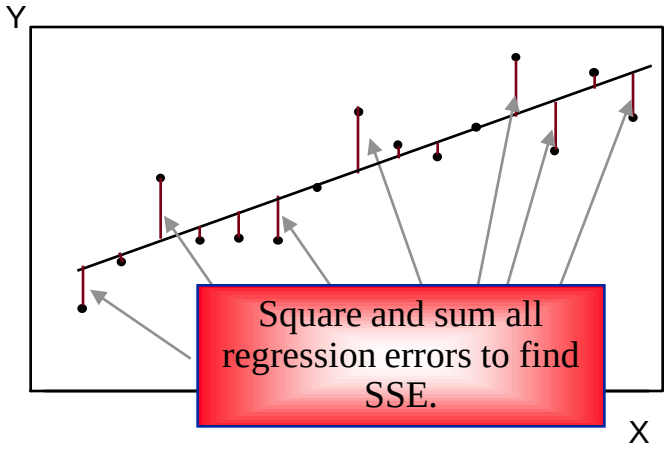
df = (n - 2) (n total observations less one degree of freedom for each parameter estimated (b_0 and b_1))

$$SSE = \sum (Y - \hat{Y})^2 = SS_Y - \frac{(SS_{XY})^2}{SS_X}$$

$$= SS_Y - b_1 SS_{XY}$$

An unbiased estimator of s^2, denoted by S^2:

$$MSE = \frac{SSE}{(n - 2)}$$



Example 10 - 1:

$$SSE = SS_Y - b_1 SS_{XY}$$

$$= 66855898 - (1.255333776)(51402852.4)$$

$$= 2328161.2$$

$$MSE = \frac{SSE}{n - 2} = \frac{2328161.2}{23}$$

$$= 101224.4$$

$$s = \sqrt{MSE} = \sqrt{101224.4} = 318.158$$

Standard Errors of Estimates in Regression

The standard error of b_0 (intercept):

$$s(b_0) = \frac{s\sqrt{\sum x^2}}{\sqrt{nSS_X}}$$

where s $= \sqrt{\text{MSE}}$

The standard error of b_1 (slope):

$$s(b_1) = \frac{s}{\sqrt{SS_X}}$$

Example 10 - 1:

$$s(b_0) = \frac{s\sqrt{\sum x^2}}{\sqrt{nSS_X}} = \frac{318.158\sqrt{293426944}}{\sqrt{(25)(4097557.84)}} = 170.338$$

$$s(b_1) = \frac{s}{\sqrt{SS_X}} = \frac{318.158}{\sqrt{40947557.84}} = 0.04972$$

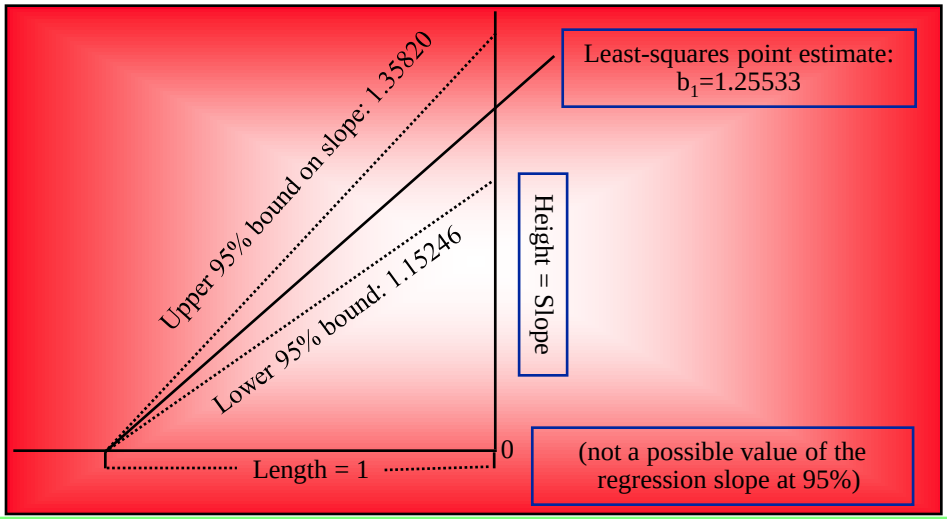
Confidence Intervals for the Regression Parameters

A $(1-\alpha)$ 100% confidence interval for b_0:

$$b_0 \pm t_{\left(\frac{\alpha}{2},(n-2)\right)} s(b_0)$$

A $(1-\alpha)$ 100% confidence interval for b_1:

$$b_1 \pm t_{\left(\frac{\alpha}{2},(n-2)\right)} s(b_1)$$



Example 10 - 1

95% Confidence Intervals:

$$b_0 \pm t_{(0.025,(25-2))} s(b_0)$$

$$= 274.85 \pm (2.069)\ (170.338)$$

$$= 274.85 \pm 352.43$$

$$= [-77.58, 627.28]$$

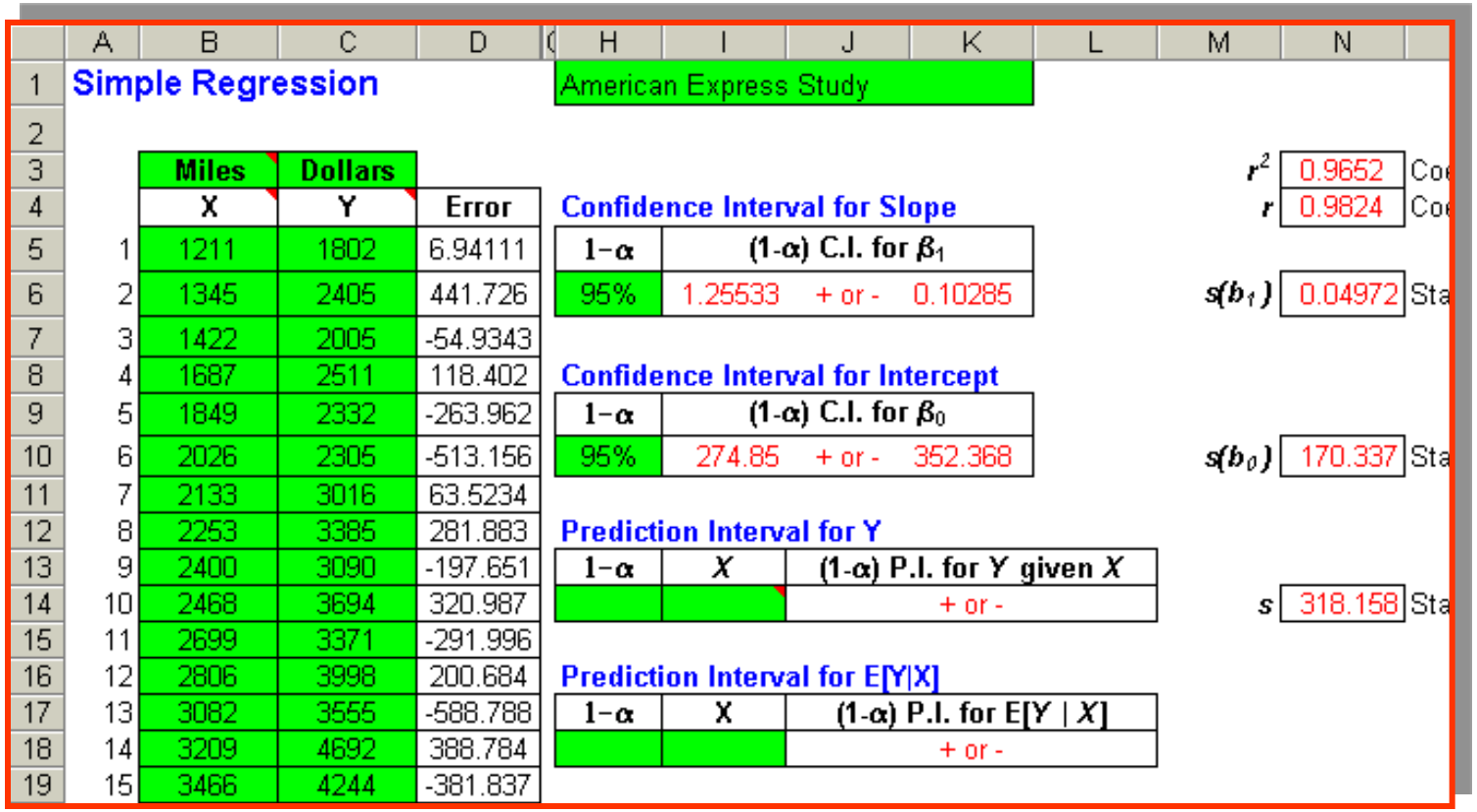
$$b_1 \pm t_{(0.025,(25-2))} s(b_1)$$

$$= 1.25533 \pm (2.069)\ (0.04972)$$

$$= 1.25533 \pm 0.10287$$

$$= [1.15246, 1.35820]$$

Template (partial output) that can be used to obtain Confidence Intervals for β_0 and β_1



Simple Regression

American Express Study

	Miles X	Dollars Y	Error
1	1211	1802	6.94111
2	1345	2405	441.726
3	1422	2005	-54.9343
4	1687	2511	118.402
5	1849	2332	-263.962
6	2026	2305	-513.156
7	2133	3016	63.5234
8	2253	3385	281.883
9	2400	3090	-197.651
10	2468	3694	320.987
11	2699	3371	-291.996
12	2806	3998	200.684
13	3082	3555	-588.788
14	3209	4692	388.784
15	3466	4244	-381.837

Confidence Interval for Slope

1−α	(1-α) C.I. for β_1
95%	1.25533 + or - 0.10285

Confidence Interval for Intercept

1−α	(1-α) C.I. for β_0
95%	274.85 + or - 352.368

Prediction Interval for Y

1−α	X	(1-α) P.I. for Y given X
		+ or -

Prediction Interval for E[Y\|X]

1−α	X	(1-α) P.I. for E[Y | X]
		+ or -

r^2	0.9652
r	0.9824
$s(b_1)$	0.04972
$s(b_0)$	170.337
s	318.158

10-5 Correlation

The **correlation** between two random variables, X and Y, is a measure of the *degree of linear association* between the two variables.

The population correlation, denoted by ρ, can take on any value from -1 to 1.

$\rho = -1$ indicates a perfect negative linear relationship

$-1 < \rho < 0$ indicates a negative linear relationship

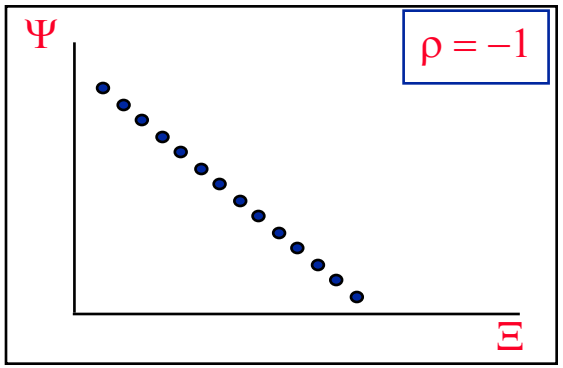
$\rho = 0$ indicates no linear relationship

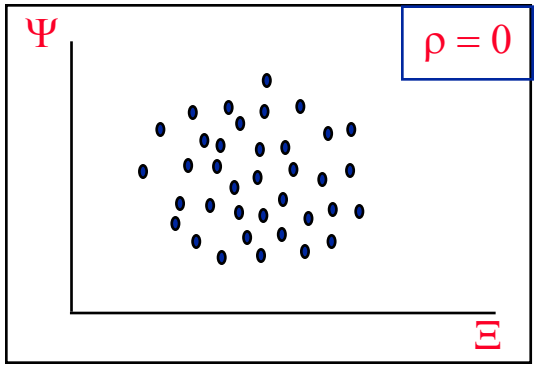
$0 < \rho < 1$ indicates a positive linear relationship

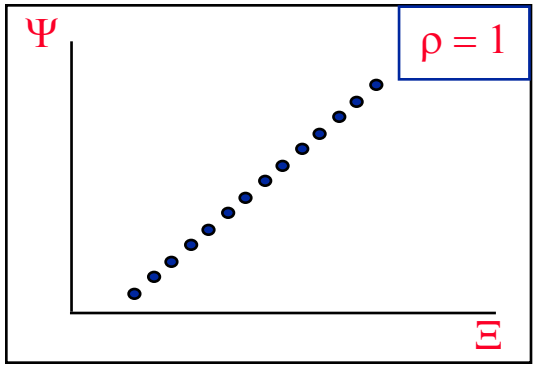
$\rho = 1$ indicates a perfect positive linear relationship

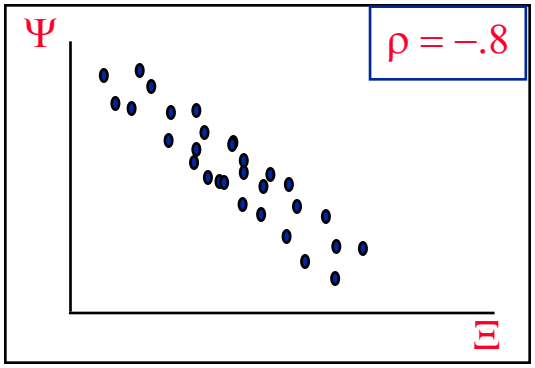
The absolute value of ρ indicates the strength or exactness of the relationship.

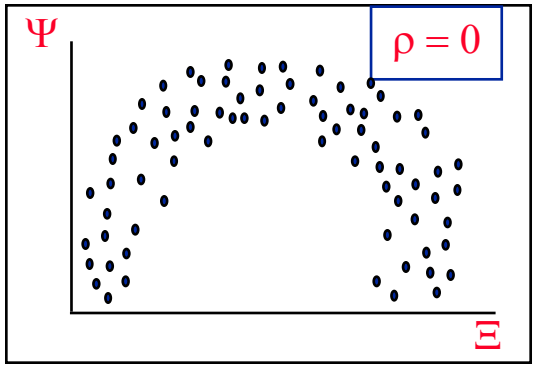
Illustrations of Correlation

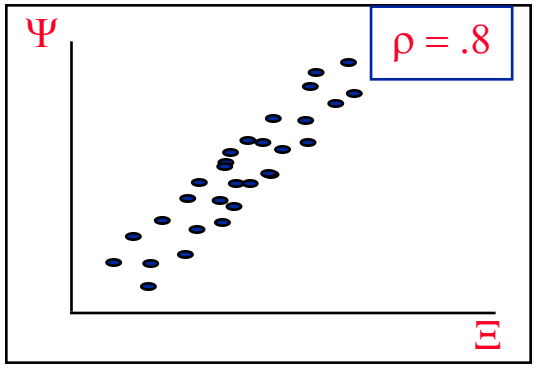












Covariance and Correlation

The covariance of two random variables X and Y:

$$Cov(X,Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

where μ_X and μ_Y are the population means of X and Y respectively.

The population correlation coefficient:

$$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

The sample correlation coefficient*:

$$r = \frac{SS_{XY}}{\sqrt{SS_X SS_Y}}$$

Εξαμπλε 10−1:

$$\rho = \frac{\Sigma\Sigma_{\Xi\Psi}}{\sqrt{\Sigma\Sigma_{\Xi}\Sigma\Sigma_{\Psi}}}$$

$$= \frac{514028524}{\sqrt{(4094755784)(668558983)}}$$

$$= \frac{514028524}{523219432.9} = .9824$$

Hypothesis Tests for the Correlation Coefficient

H_0: $\rho = 0$ (No linear relationship)
H_1: $\rho \neq 0$ (Some linear relationship)

Test Statistic:

$$t_{(n-2)} = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

Example 10-1:

$$t_{(n-2)} = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

$$= \frac{0.9824}{\sqrt{\frac{1-0.9651}{25-2}}}$$

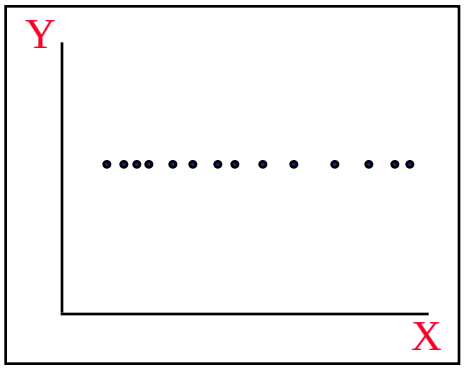
$$= \frac{0.9824}{0.0389} = 25.25$$

$t_{0.005} = 2.807 < 25.25$

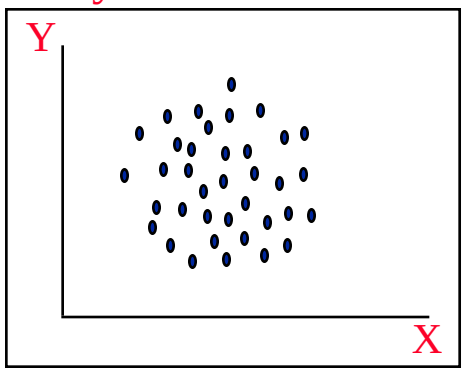
H_0 rejected at 1% level

10-6 Hypothesis Tests about the Regression Relationship

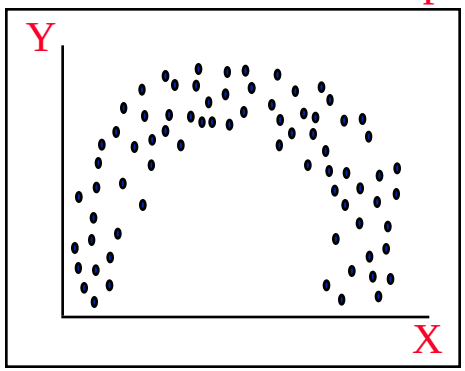
Constant Y



Unsystematic Variation



Nonlinear Relationship



A hypothesis test for the existence of a linear relationship between X and Y:

$$H_0 : b_1 = 0$$

$$H_1 : b_1 \neq 0$$

Test statistic for the existence of a linear relationship between X and Y:

$$t_{(n-2)} = \frac{b_1}{s(b_1)}$$

where b_1 is the least-squares estimate of the regression slope and $s(b_1)$ is the standard error of b_1.

When the null hypothesis is true, the statistic has a t distribution with $n - 2$ degrees of freedom.

Hypothesis Tests for the Regression Slope

Example 10 - 1:

$H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

$$t_{(n-2)} = \frac{b_1}{s(b_1)}$$

$$= \frac{1.25533}{0.04972} = 25.25$$

$$t_{(0.005,23)} = 2.807 < 25.25$$

H_0 is rejected at the 1% level and we may conclude that there is a relationship between charges and miles traveled.

Example 10 - 4:

$H_0: \beta_1 = 1$

$H_1: \beta_1 \neq 1$

$$t_{(n-2)} = \frac{b_1 - 1}{s(b_1)}$$

$$= \frac{1.24 - 1}{0.21} = 1.14$$

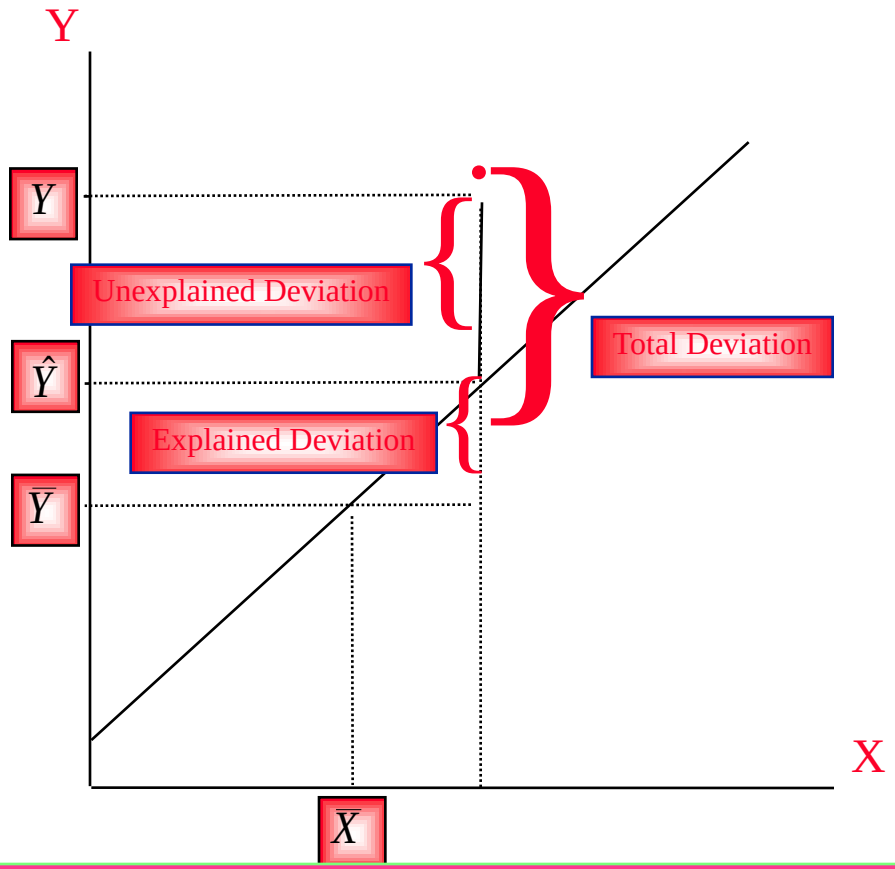
$$t_{(0.05,58)} = 1.671 > 1.14$$

H_0 is not rejected at the 10% level.

We may not conclude that the beta coefficient is different from 1.

10-7 How Good is the Regression?

The **coefficient of determination, r^2**, is a descriptive measure of the strength of the regression relationship, a measure of how well the regression line fits the data.



$$(y-\bar{y}) = (y-\hat{y}) + (\hat{y}-\bar{y})$$

Total Deviation = Unexplained Deviation (Error) + Explained Deviation (Regression)

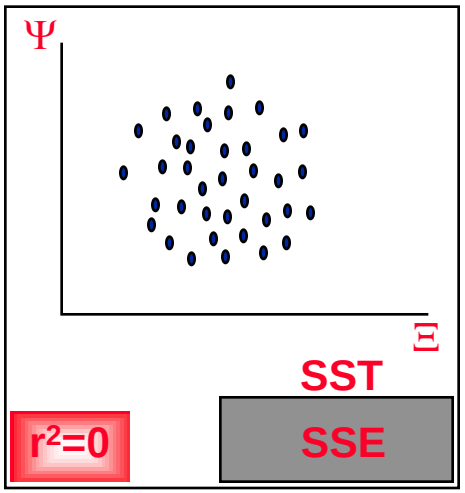
$$\sum(y-\bar{y})^2 = \sum(y-\hat{y})^2 + \sum(\hat{y}-\bar{y})^2$$

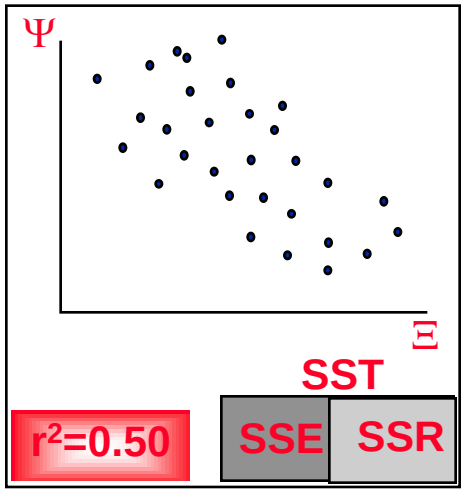
SST = SSE + SSR

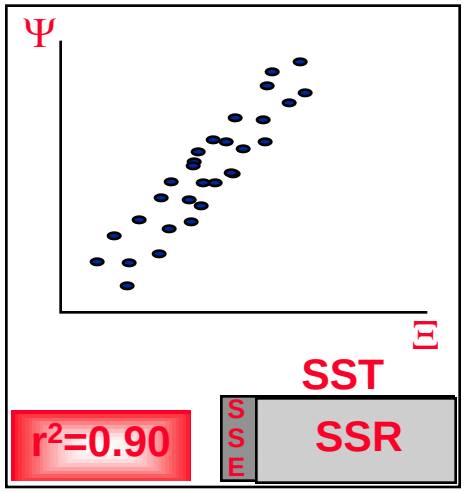
$$r^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

Percentage of total variation explained by the regression.

The Coefficient of Determination

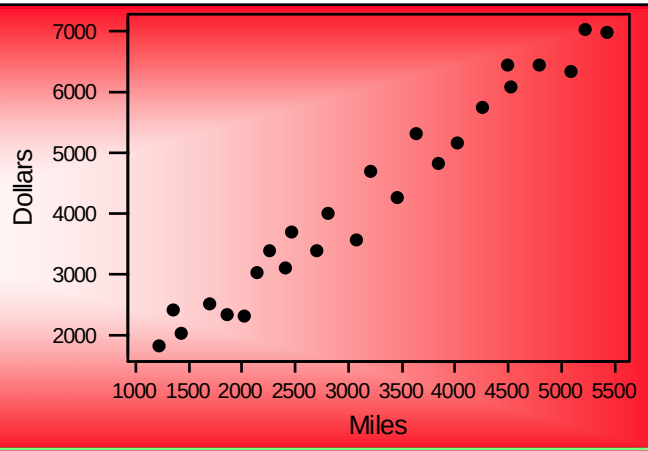






Example 10-1:

$$r^2 = \frac{SSR}{SST} = \frac{64527736.8}{66855898} = 0.96518$$



10-8 Analysis of Variance and an F Test of the Regression Model

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio
Regression	SSR	(1)	MSR	$\frac{MSR}{MSE}$
Error	SSE	(n-2)	MSE	
Total	SST	(n-1)	MST	

Example 10-1

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	p Value
Regression	64527736.8	1	64527736.8	637.47	0.000
Error	2328161.2	23	101224.4		
Total	66855898.0	24			

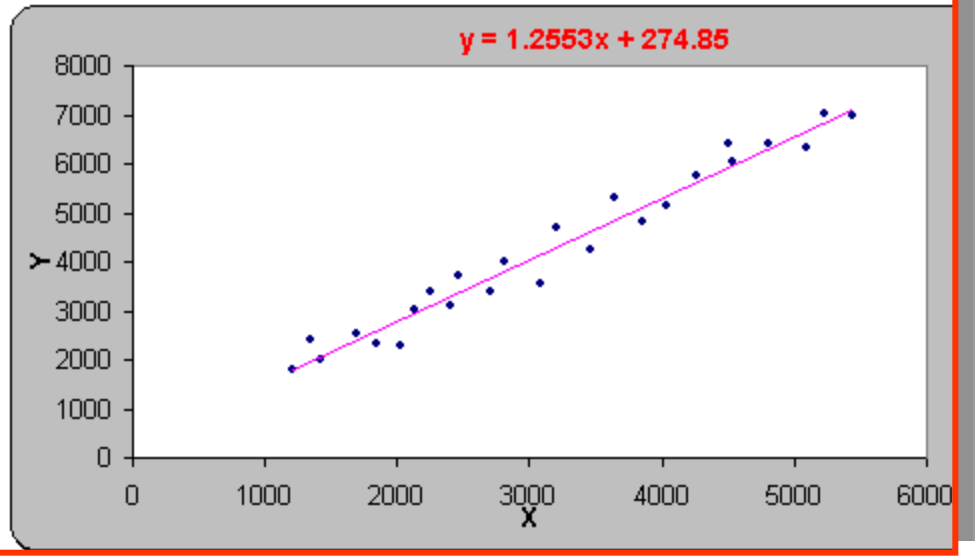
Template (partial output) that displays Analysis of Variance and an F Test of the Regression Model

	A	B	C	D
19	15	3466	4244	-381.837
20	16	3643	5298	449.969
21	17	3852	4801	-309.395
22	18	4033	5147	-190.611
23	19	4267	5738	106.641
24	20	4498	6420	498.659
25	21	4533	6059	93.7223
26	22	4804	6426	120.527
27	23	5090	6321	-343.499
28	24	5233	7026	181.989
29	25	5439	6964	-138.61
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				

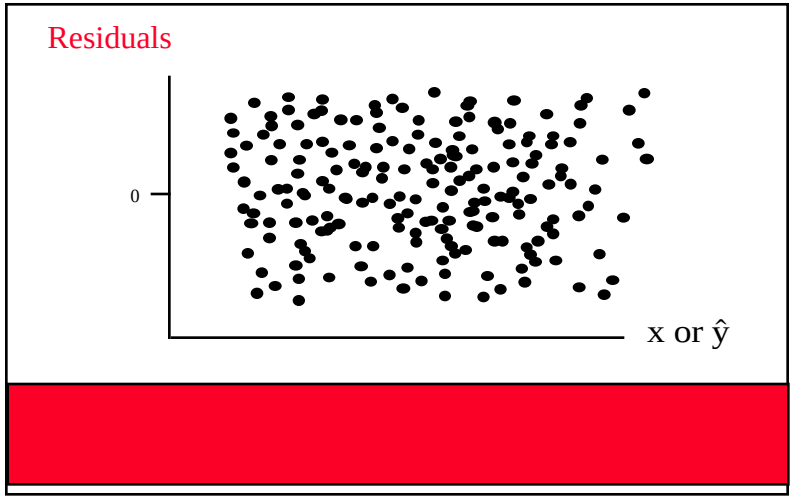
ANOVA Table

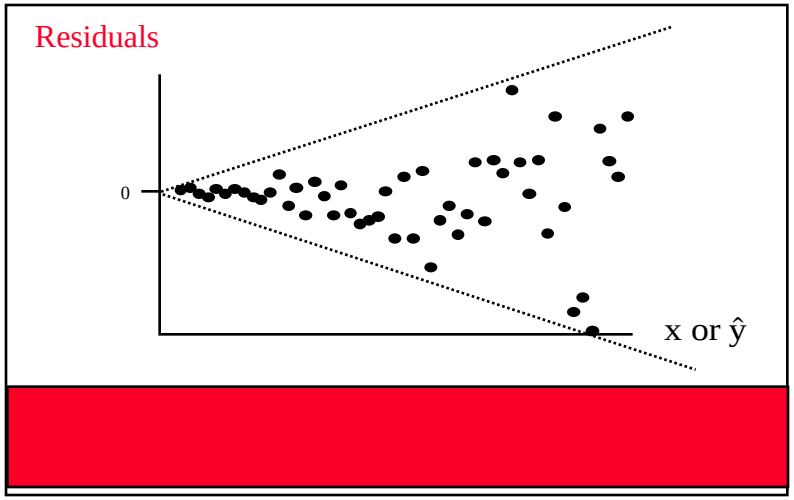
Source	SS	df	MS	*F*	$F_{critical}$	*p*-value
Regn.	6.5E+07	1	6.5E+07	637.472	4.27934	0.0000
Error	2328161	23	101224			
Total	6.7E+07	24				

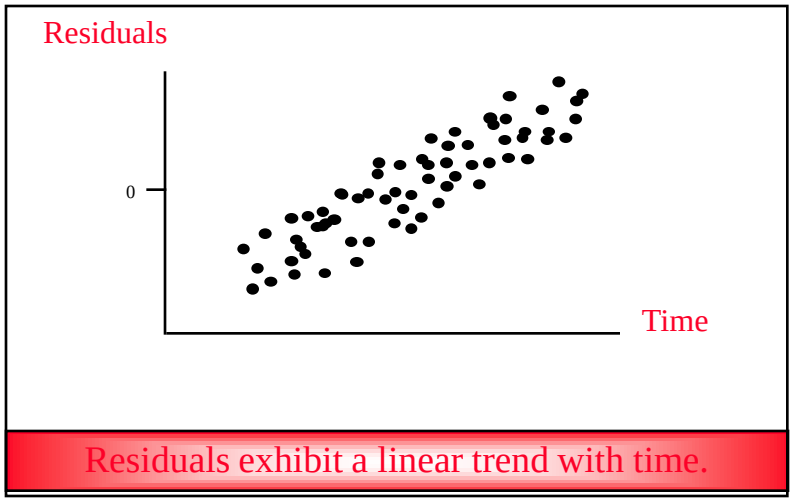
Scatter Plot, Regression Line and Regression Equation

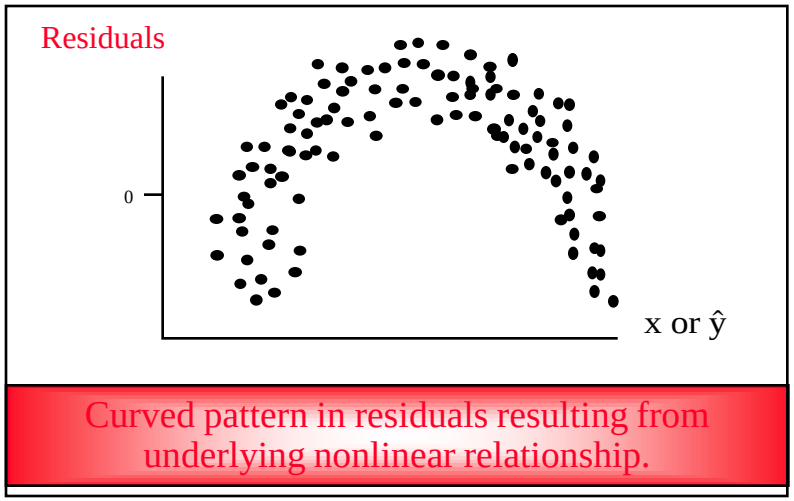


10-9 Residual Analysis and Checking for Model Inadequacies



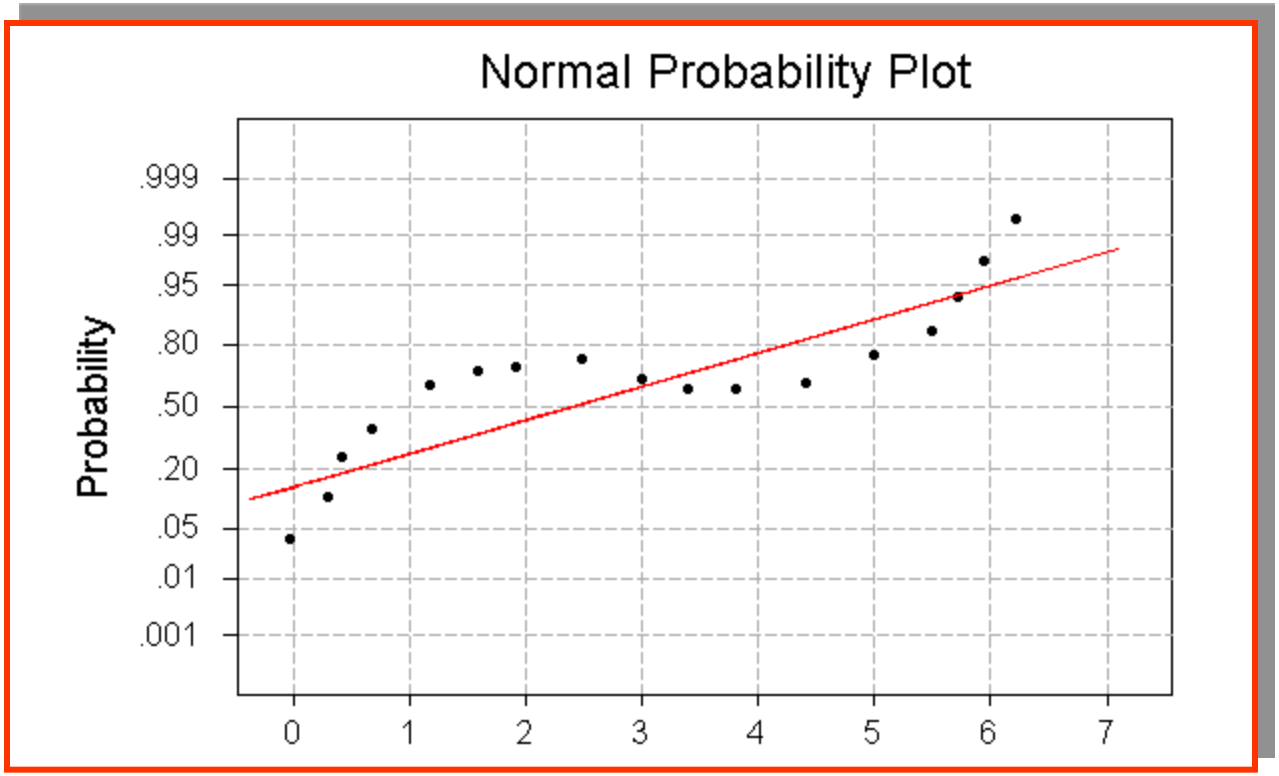






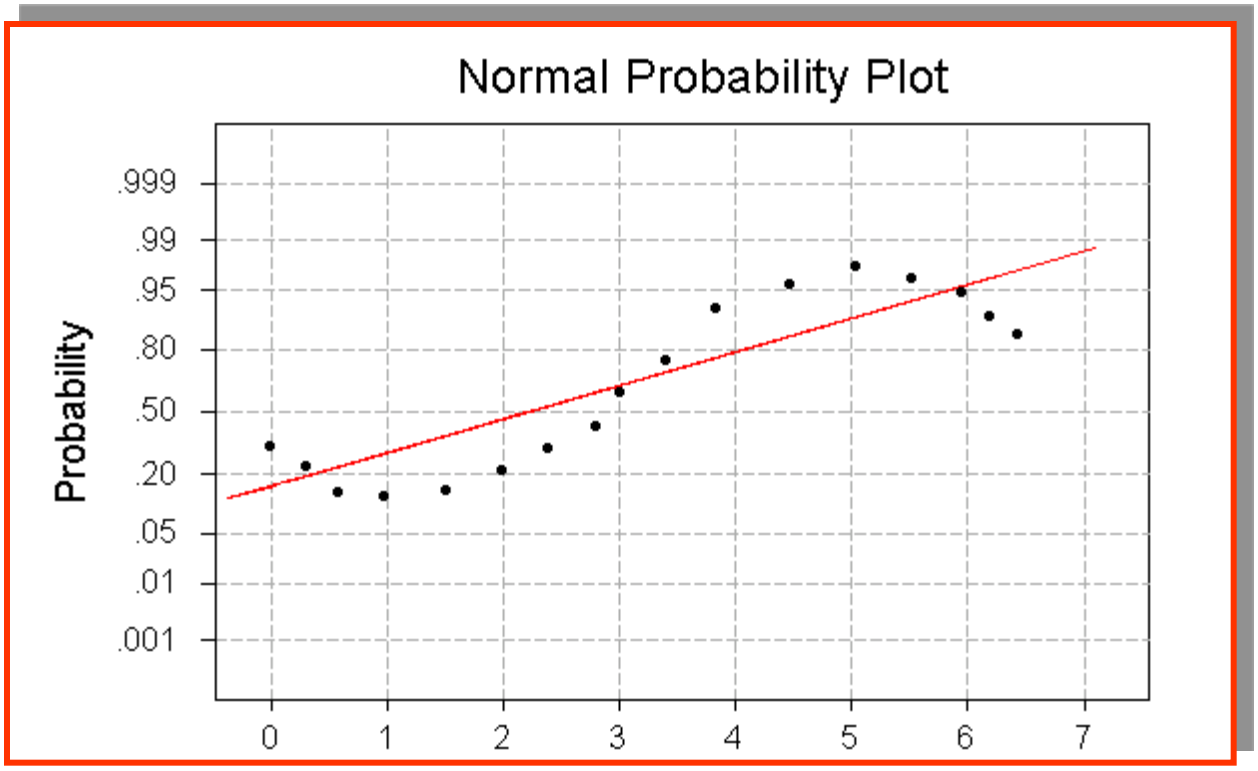
Normal Probability Plot of the Residuals

Flatter than Normal



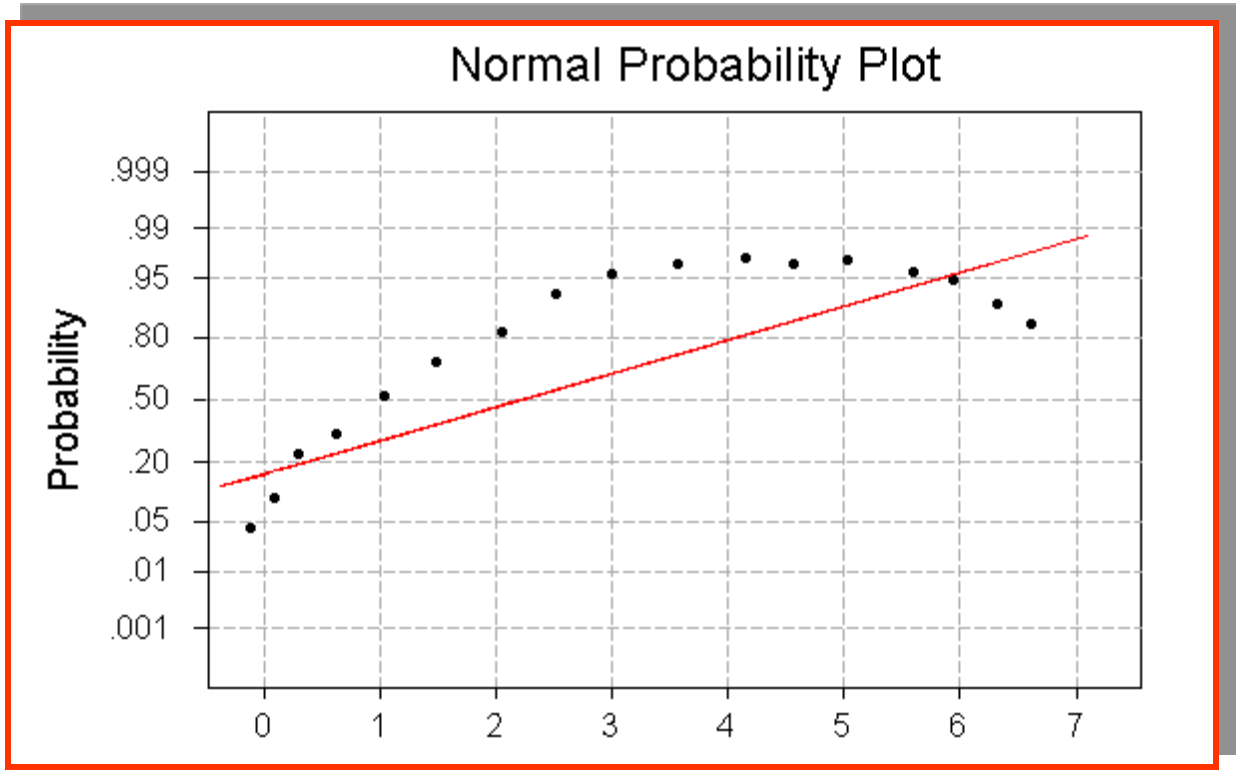
Normal Probability Plot of the Residuals

More Peaked than Normal



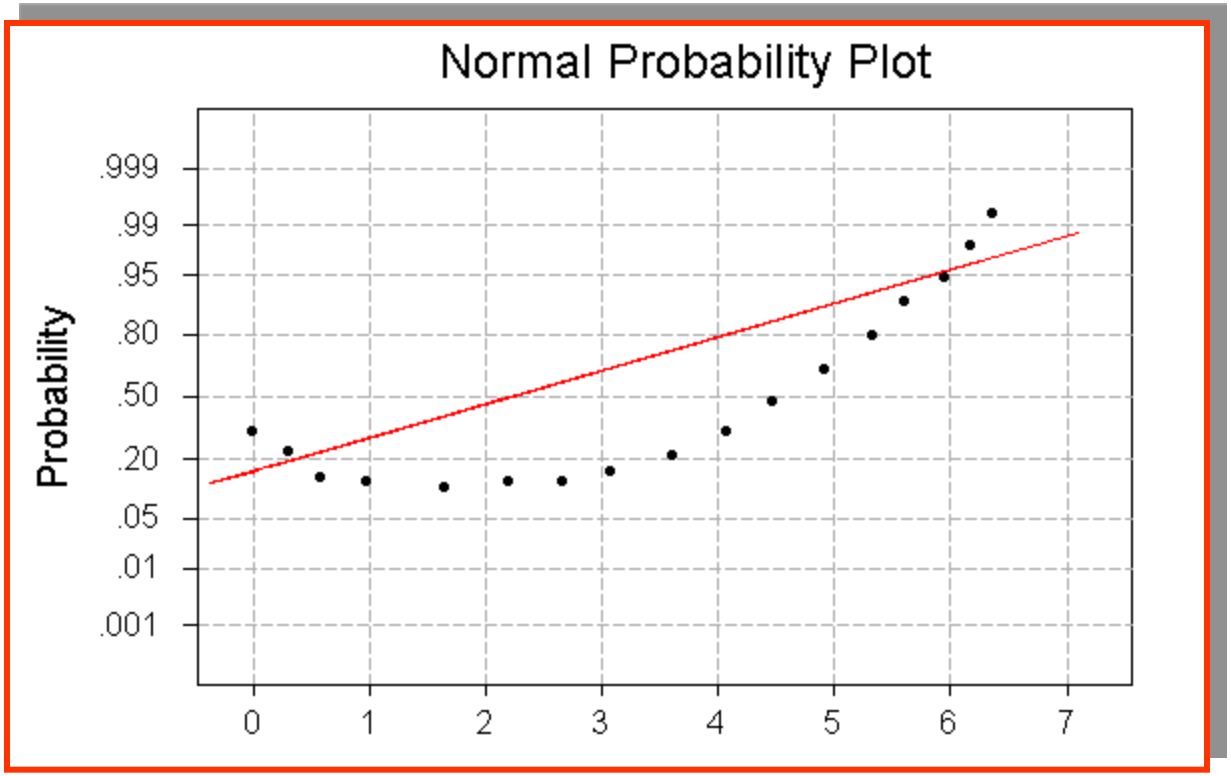
Normal Probability Plot of the Residuals

More Positively Skewed than Normal



Normal Probability Plot of the Residuals

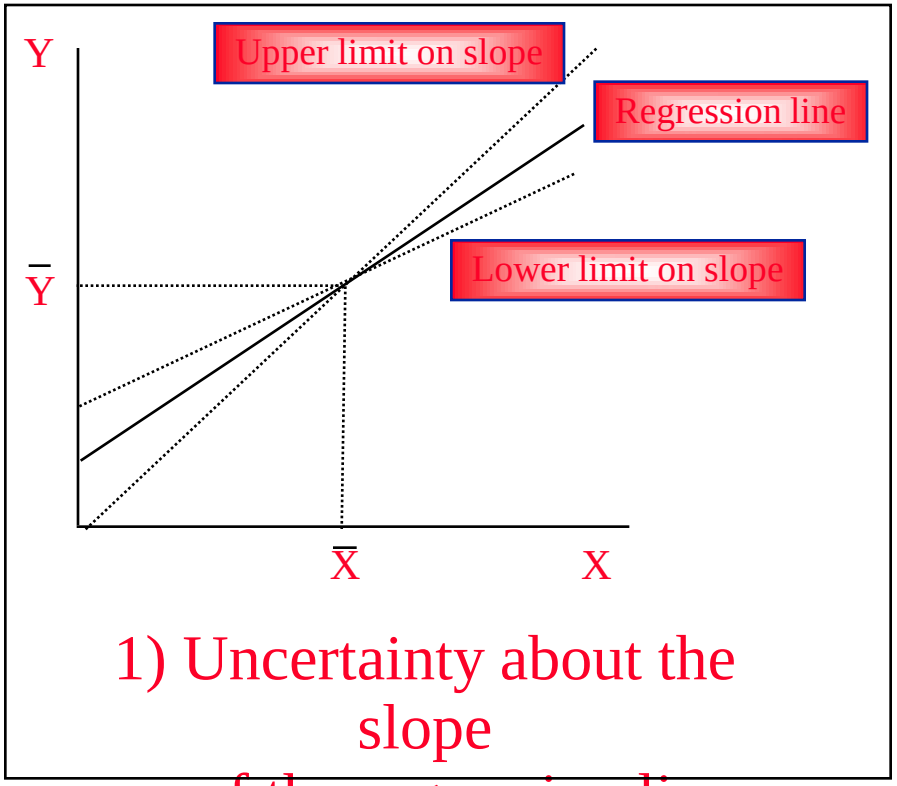
More Negatively Skewed than Normal



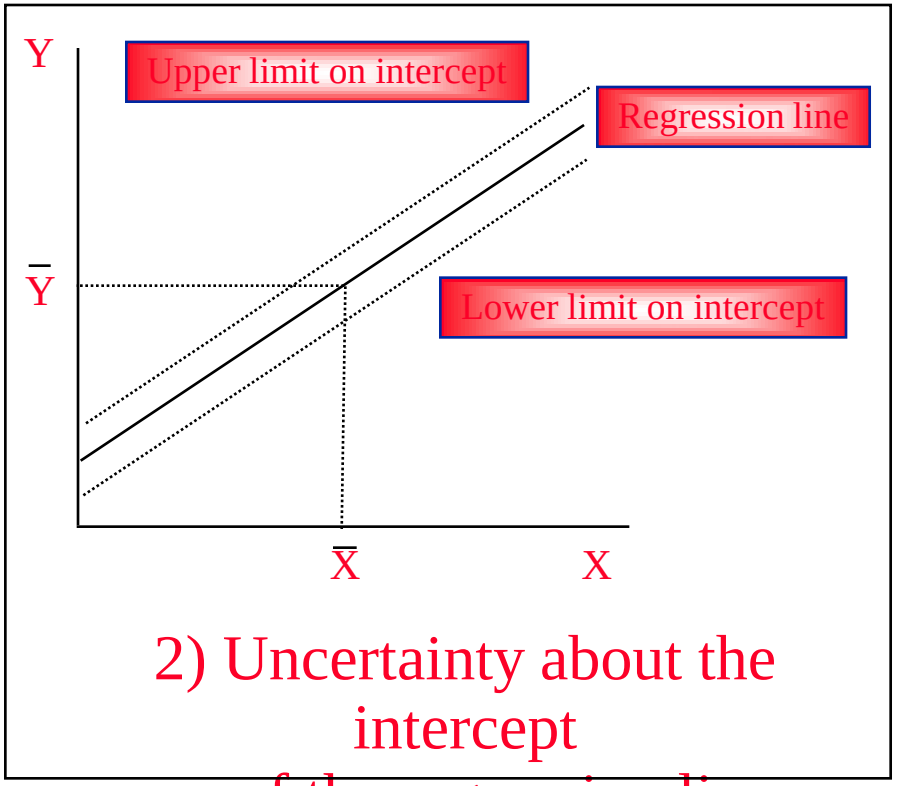
10-10 Use of the Regression Model for Prediction

- **Point Prediction**
  - ✓ A single-valued estimate of Y for a given value of X obtained by inserting the value of X in the estimated regression equation.
- **Prediction Interval**
  - ✓ For a value of Y given a value of X
    - » Variation in regression line estimate
    - » Variation of points around regression line
  - ✓ For an average value of Y given a value of X
    - » Variation in regression line estimate

Errors in Predicting E[Y|X]

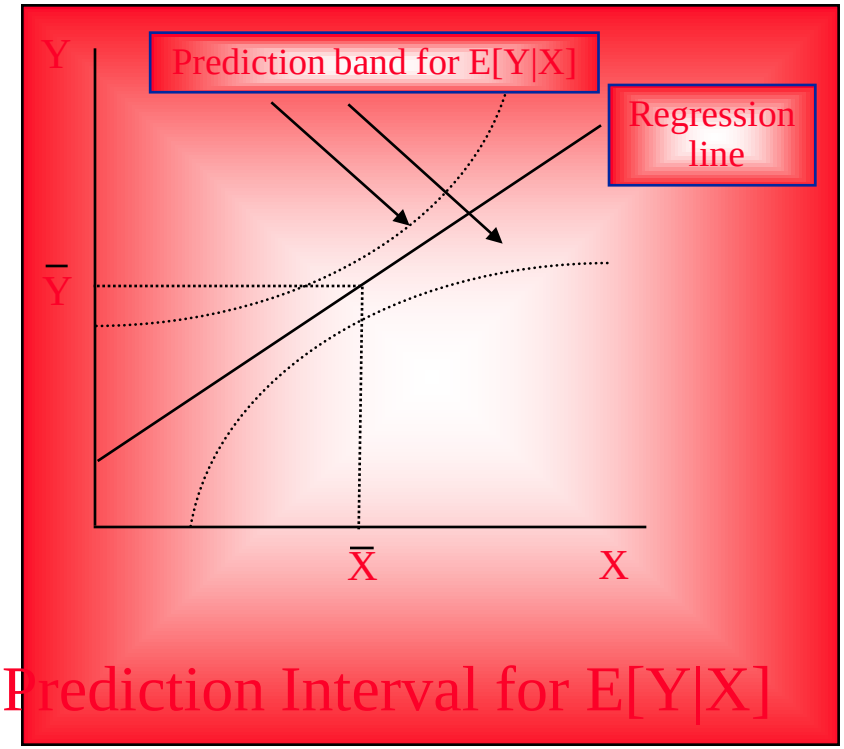


1) Uncertainty about the slope of the regression line



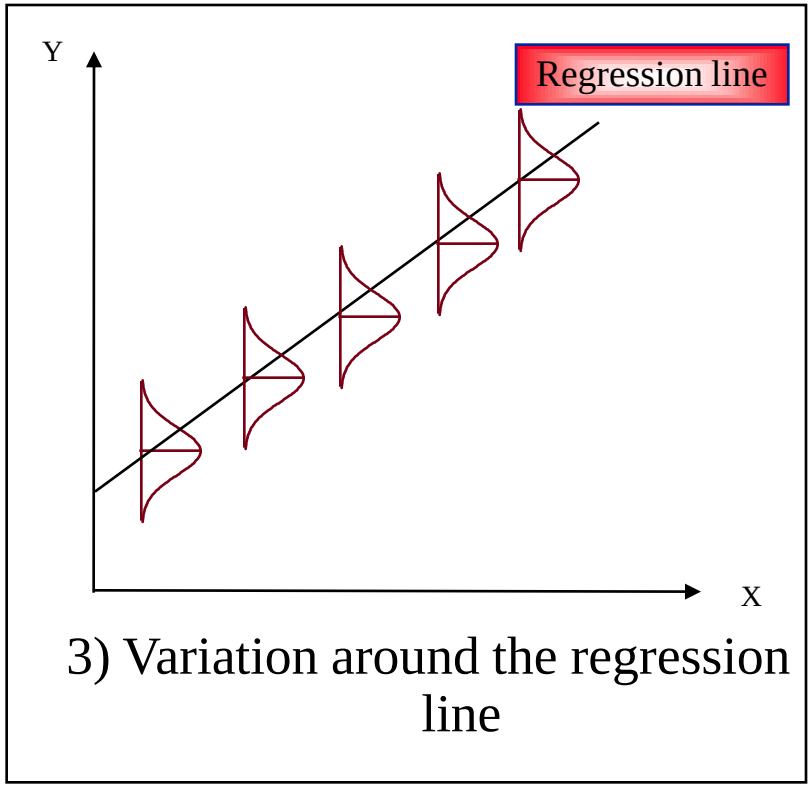
2) Uncertainty about the intercept of the regression line

Prediction Interval for E[Y|X]

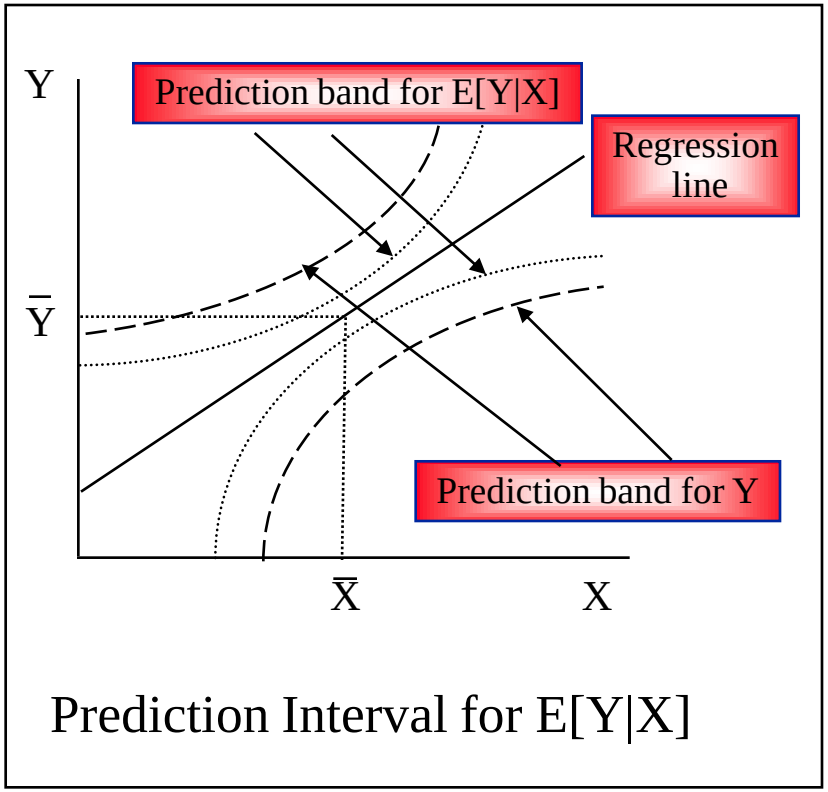


- **The prediction band for E[Y|X] is narrowest at the mean value of X.**
- **The prediction band widens as the distance from the mean of X increases.**
- **Predictions become very unreliable when we extrapolate beyond the range of the sample itself.**

Additional Error in Predicting Individual Value of Y



3) Variation around the regression line



Prediction Interval for E[Y|X]

Prediction Interval for a Value of Y

A $(1-\alpha)100\%$ prediction interval for Y :

$$\hat{y} \pm t_{\frac{\alpha}{2}} \times s\sqrt{1+\frac{1}{n}+\frac{(x-\bar{x})^2}{SS_x}}$$

Example 10 -1 (X = 4,000) :

$$\{274.85+(1.2553)(4{,}000)\} \pm 2.069 \times 318.16\sqrt{1+\frac{1}{25}+\frac{(4{,}000-3{,}177.92)^2}{40{,}947{,}557.84}}$$

$$= 5296.05 \pm 676.62 = [4619.43, 5972.67]$$

Prediction Interval for the Average Value of Y

A $(1-\alpha)100\%$ prediction interval for the $E[Y|X]$:

$$\hat{y} \pm t_{\frac{\alpha}{2}} \times s\sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{SS_x}}$$

Example 10-1 (X = 4,000):

$$\{274.85 + (1.2553)(4{,}000)\} \pm 2.069 \times 318.16\sqrt{\frac{1}{25} + \frac{(4{,}000 - 3{,}177.92)^2}{40{,}947{,}557.84}}$$

$$= 5{,}296.05 \pm 156.48 = [5139.57, 5452.53]$$

Template Output with Prediction Intervals

Simple Regression

American Express Study

	Miles X	Dollars Y	Error
1	1211	1802	6.94111
2	1345	2405	441.726
3	1422	2005	-54.9343
4	1687	2511	118.402
5	1849	2332	-263.962
6	2026	2305	-513.156
7	2133	3016	63.5234
8	2253	3385	281.883
9	2400	3090	-197.651
10	2468	3694	320.987
11	2699	3371	-291.996
12	2806	3998	200.684
13	3082	3555	-588.788
14	3209	4692	388.784
15	3466	4244	-381.837

Confidence Interval for Slope

1−α	(1-α) C.I. for β_1
95%	1.25533 + or - 0.10285

Confidence Interval for Intercept

1−α	(1-α) C.I. for β_0
95%	274.85 + or - 352.368

Prediction Interval for Y

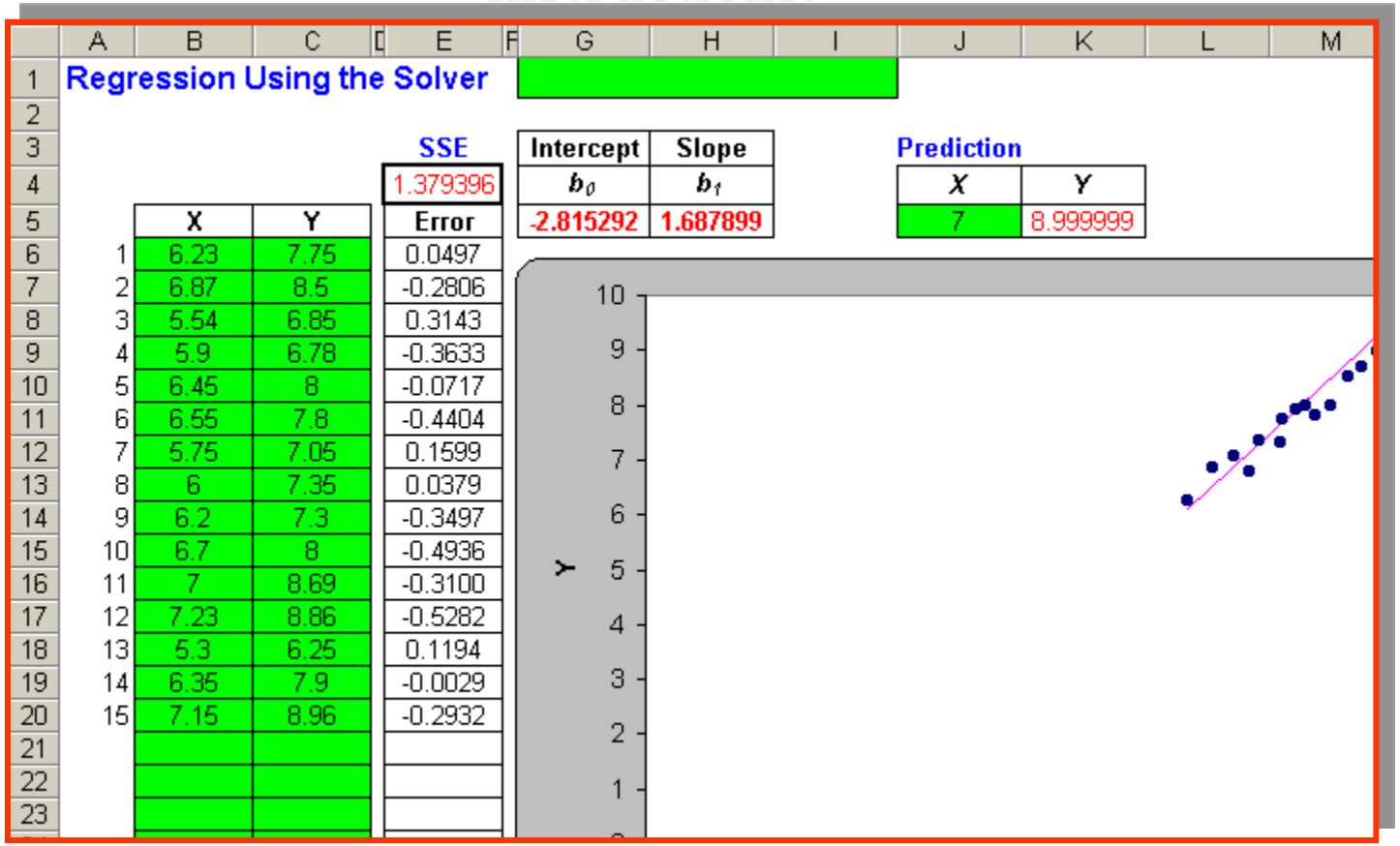
1−α	X	(1-α) P.I. for Y given X
95%	4000	5296.18 + or - 676.498

Prediction Interval for E[Y|X]

1−α	X	(1-α) P.I. for E[Y | X]
95%	4000	5296.18 + or - 156.449

10-11 The Solver Method for Regression

The solver macro available in EXCEL can also be used to conduct a simple linear regression. **See the text for instructions.**



Regression Using the Solver

	SSE	Intercept	Slope
	1.379396	b_0	b_1
		-2.815292	1.687899

Prediction

X	Y
7	8.999999

	X	Y	Error
1	6.23	7.75	0.0497
2	6.87	8.5	-0.2806
3	5.54	6.85	0.3143
4	5.9	6.78	-0.3633
5	6.45	8	-0.0717
6	6.55	7.8	-0.4404
7	5.75	7.05	0.1599
8	6	7.35	0.0379
9	6.2	7.3	-0.3497
10	6.7	8	-0.4936
11	7	8.69	-0.3100
12	7.23	8.86	-0.5282
13	5.3	6.25	0.1194
14	6.35	7.9	-0.0029
15	7.15	8.96	-0.2932

Name	Shakeel Nouman
Religion	Christian
Domicile	Punjab (Lahore)
Contact #	0332-4462527. 0321-9898767
E.Mail	sn_gcu@yahoo.com sn_gcu@hotmail.com

M.Phil (Statistics)	GC University, . (Degree awarded by GC University)
M.Sc (Statistics)	GC University, . (Degree awarded by GC University)
Statitical Officer (BS-17) (Economics & Marketing Division)	Livestock Production Research Institute Bahadurnagar (Okara), Livestock & Dairy Development Department, Govt. of Punjab