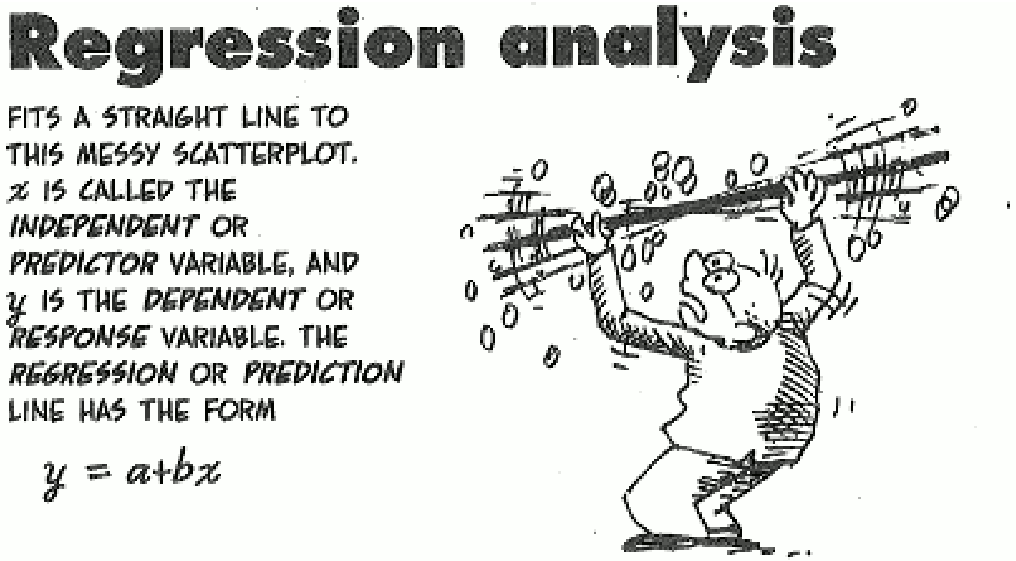
Welcome To….

Regression Analysis

(Part-2)



Learning Outcomes

When you will complete this chapter, you would be able to-

- Development of Mathematical Equation for modeling the relationship of the variables.
- Use of the Equation for the purpose of prediction.

Outlines

From this lecture, you are going to learn…

- Definition of Regression with Examples
- Define Independent and Dependent variables
- Types of Regression with definition (Simple & multiple)
- Fit/ Estimate Simple linear regression Line/model
- Estimation of regression coefficients with interpretation
- Interpret Coefficient of determination

(Part-1)

(Part-2)

Example of Estimation of regression coefficients with interpretation

Problem: Suppose a company's owner wants to forecast sales on the basis of advertising expenses. The owner would like to review the relationship between sales and the amount spent on advertising. Below is the information on sales and advertising expense for the last four months:

Month	Advertising expense(x) ($ million)	Sales revenue(y) ($ million)
July	2	7
August	1	3
September	3	8
October	4	10

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

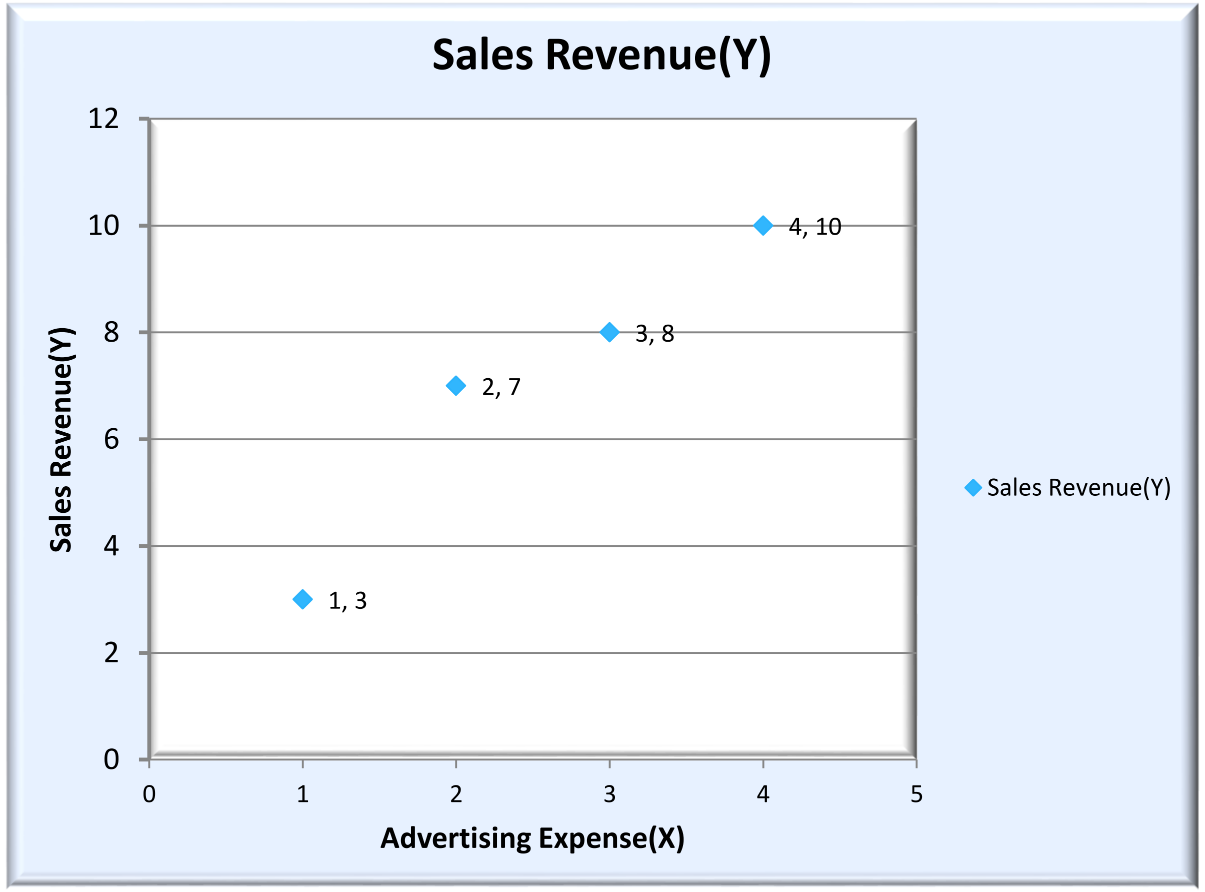
$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x}$$

a) Draw scatter plot.
b) Determine the estimated regression model.
c) Interpret the value $\boldsymbol{\beta_0}$ and $\boldsymbol{\beta_1}$.
d) Estimate sales when $9 million is spent on advertising.

Example of Estimation of regression coefficients with interpretation

Solution:

a) If we plot the data and draw scatter plot we get as below plot



Example of Estimation of regression coefficients with interpretation

Solution:

b) Let, Estimated Simple linear regression Model,

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \varepsilon_i$$

we know that,

$$\bar{x} = \frac{\sum_{i=1}^{4} x_i}{n} = \frac{10}{4} = 2.5 \quad \text{and}$$

$$\bar{y} = \frac{\sum_{i=1}^{4} y_i}{n} = \frac{28}{4} = 7$$

Advertising expense(x)	Sales revenue(y)	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
2	7	-0.5	0	0.25	0
1	3	-1.5	-4	2.25	6
3	8	0.5	1	0.25	0.5
4	10	1.5	3	2.25	4.5
				$\sum_{i=1}^{4} (x_i - \bar{x})^2$ =5	$\sum_{i=1}^{4} (x_i - \bar{x})(y_i - \bar{y})$ =11

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

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Example of Estimation of regression coefficients with interpretation

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

$$= \frac{11}{5} = 2.2$$

$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1\overline{x}$$

=7-2.2*2.5

Now if we put the values the Estimated Simple linear regression Model,

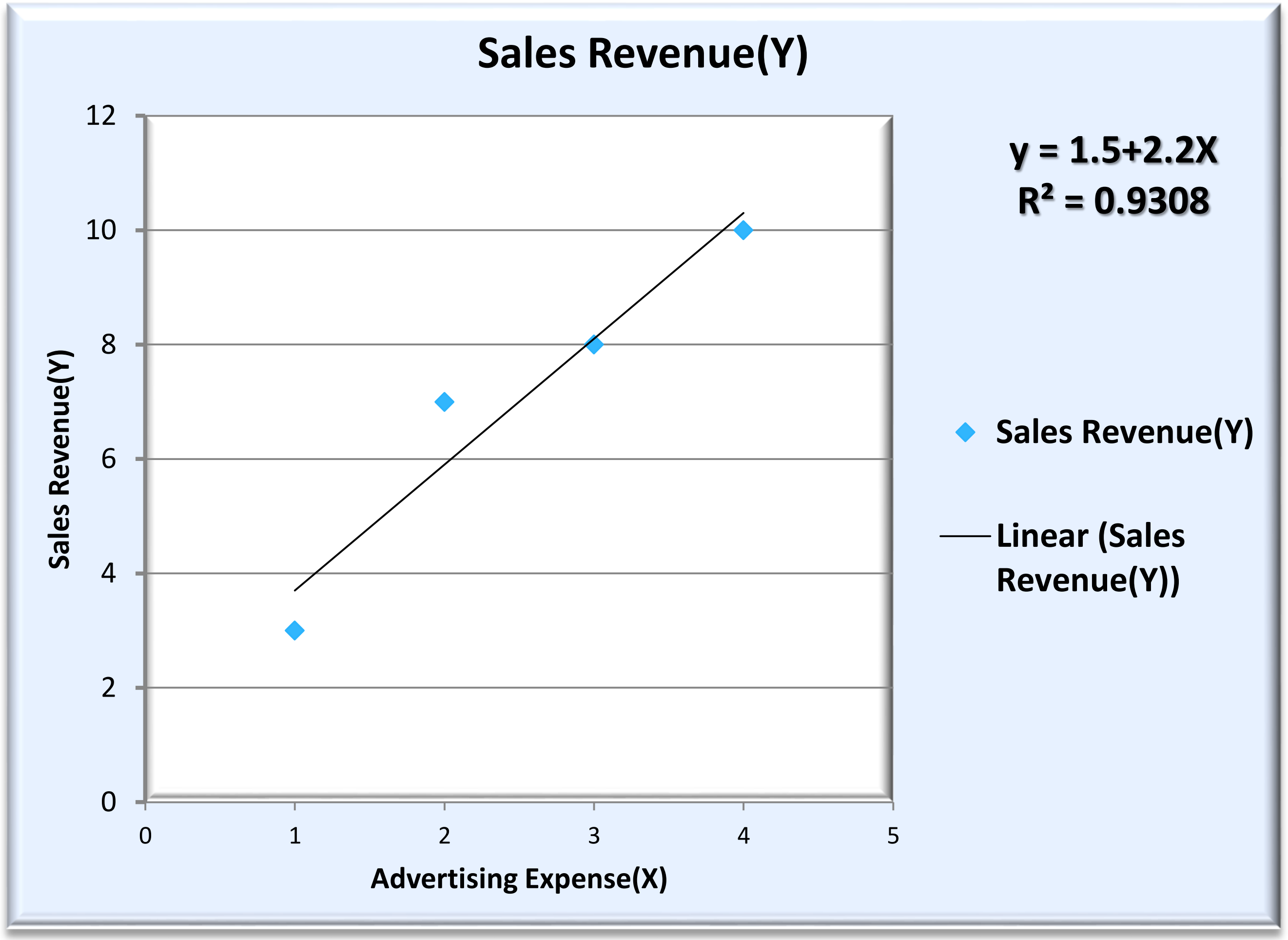
$$\hat{Y}_i = 1.5 + 2.2X_i + \varepsilon_i,$$

Example of Estimation of regression coefficients with interpretation

c) Interpretation:

$\hat{\beta}_1$=2.2 means that an increase of $1million in advertising cost, the sales revenue will increase $2.2 million.

$\hat{\beta}_0$=1.5 means that, if there is no advertisement cost, then sales revenue would be $1.5 million.



Prediction of Dependent variable in terms of independent variable and calculation of error

d) Now if x= 9, then

$\hat{\mathbf{Y}}$ =1.5+2.2*9=21.3.So, when advertisement cost is \$9 million, the expected sales revenue would be \$21.3 million.

Calculation of error for a single value:

In the given data set , when X= 3, Y=8. But using the Estimated regression equation the value is,

$\hat{Y}$ *=1.5+2.2*3=8.1.*

So the amount of error is, $Y_i - \hat{Y}_i$ =8-8.1=-0.1.

Coefficient of Determination(r^2)

Coefficient of Determination (r^2): The coefficient of determination tells the percent of the variation in the dependent variable that is explained (determined) by the model and the explanatory variable.

Interpretation of r^2**:** Suppose r^2=92.7%.
Interpretation: Almost 93% of the variability of the dependent variables explained by the independent variables


Thank You