

Biostatistics

Lecture 12

Prepared by

Baneen Ahmed

Simple linear Regression:

The purpose of using the simple linear regression analysis method is to study and analyze the effect of a quantitative variable on another quantitative variable.

In simple regression analysis, we find that the researcher is interested in studying the effect of one of the variables, which is called the independent variable or the predicted one on the second variable, and it is called the dependent or predicted variable. Then the linear regression model can be presented in the form of a linear equation of the first degree that reflects the dependent variable as a function of the independent variable as follows:

$$y = \beta_0 + \beta_1 x + e$$

y : is the dependent variable (which is affected)

x : is the independent variable (which influences)

β_0 : It is the truncated part of the vertical y-axis and it reflects the value of the dependent variable in the

absence of the value of the independent variable x ,
 $x=0$.

β_1 : The slope of a straight line ($\beta_0 + \beta_1 x$) reflects the change in y if x changes by one unit.

e : It is the random error, which expresses the difference between the actual value y and the estimated value.

Simple Linear Regression Model Estimation:

The regression coefficients can be estimated using the method of least squares, and this estimate is what makes the sum of the squares of random errors as less as possible. This estimate is calculated by the following equation:

$$\hat{\beta}_1 = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$

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

The estimated simple regression equation is as follows:

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

Example: The following are data on the daily amount of protein in grams that the child needs, and the amount of increase in the weight of the calf in kilograms, for a sample of children of size 10.

Protein amount	10	11	14	15	20	25	46	50	59	70
weight gain	10	10	12	12	13	13	19	15	16	20

1-How much is the weight regression equation for the amount of protein?

2-What is the amount of weight gain when the child is given 50 grams of protein? What is the value of the random error?

Solution:

Protein amount(x)	weight gain (y)	xy	X ²
10	10	100	100
11	10	110	121
14	12	168	196
15	12	180	225

20	13	260	400
25	13	325	625
46	19	874	2116
50	15	750	2500
59	16	944	3481
70	20	1400	4900
320	140	5111	14664
$\bar{x} = 32$	$\bar{y} = 14$		

1-

$$\hat{\beta}_1 = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$

$$= \frac{10(5111) - (320)(140)}{10(14664) - (320)^2} = 0.1426$$

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

$$= 14 - (0.1426)(32) = 9.4368$$

Then $\hat{y} = 9.44 + 0.143x$

Interpretation of the equation:

$\hat{\beta}_0 = 9.44$ It indicates that if protein is not used for food, the weight will increase by 9.44 kg.

$\hat{\beta}_1=0.143$ It indicates that the higher the amount of protein by one gram, an increase in the child weight occurred by 0.143kg, that is an increase of 143 grams.

$$2- \hat{y} = 9.44 + 0.143(50)=16.59$$

The twelve lecture has ended

I wish you all the best.