THE LINEAR REGRESSION PROCESS

FROM GENERATION TO INTERPRETATION

PRESENTED BY

Frank Briody frankbriody@gmail.com Prospect High School Mt. Prospect, IL



2022 MMC Conference of Workshops



Contents

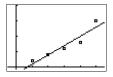
L	\mathbf{Get}	ting Least Squares Line of Best Fit	2
	1.1	From Calculator	2
	1.2	From Summary Statistics	2
	1.3	From Output	2
	1.4	Interpretation	2
		1.4.1 Slope	2
		1.4.2 V-Intercent	-

There is only one large computer program I have used in which there are to a decent approximation 0 bugs: Don Knuth's TeX

- Jaap Weel

1 Getting Least Squares Line of Best Fit

1.1 From Calculator



1.2 From Summary Statistics

Formulas (given): $\hat{y} = a + bx$ $b = r \frac{s_y}{s_x}$ $a = \bar{y} - b\bar{x}$

Descriptive Statistics: x, y

Mean SE Mean StDev Minimum Variable N N* Q1 Median Q3 Maximum 5 0 3.000 0.707 1.581 1.000 1.500 3.000 4.500 5.000 х 5 0 7.00 2.24 2.00 3.00 5.00 6.00 11.50 15.00

Correlations: x, y

Pearson correlation of x and y = 0.949

P-Value = 0.014

1.3 From Output

Regression Analysis: y versus x

The regression equation is y = -2.00 + 3.00 x

Predictor Coef SE Coef T P
Constant -2.000 1.915 -1.04 0.373
x 3.0000 0.5774 5.20 0.014

S = 1.82574 R-Sq = 90.0% R-Sq(adj) = 86.7%

Analysis of Variance

 Source
 DF
 SS
 MS
 F
 P

 Regression
 1
 90.000
 90.000
 27.00
 0.014

 Residual Error
 3
 10.000
 3.333

 Total
 4
 100.000

1.4 Interpretation

1.4.1 Slope

Slope represents the **predicted** change in response associated with each unit increase in the explanatory variable, **on average**.

1.4.2 Y-Intercept

Y-intercept is the predicted value when the explanatory (x) is 0. [Often the y-intercept is useless.]