

Regression

Data

The following data set reflects a within-subjects design with two outcome variables. The data are presented in a format suitable for entry into statistical software.

	Outcome1	Outcome2
1	.00	4.00
2	.00	7.00
3	3.00	4.00
4	5.00	9.00

Computer Output

The following tables represent typical output from statistical software. Options, labels, and layout vary from program to program.

The table of descriptive statistics can be used to assist in calculating the correlation.

	Mean	Std. Deviation	N
Outcome1	2.000	2.449	4
Outcome2	6.000	2.449	4

The table of inferential statistics shows the key elements to be calculated.

Model	R	R ²
1	0.500	0.250

Predictor	Estimate	SE	t	p	Std. Estimate
Intercept	5.000	1.785	2.801	0.107	
Outcome1	0.500	0.612	0.816	0.500	0.500

Calculations

Sum of Cross Products: The Sum of Cross Products (SCP) is not easily determined solely from the summary statistics of the output, but rather from the data.

$$SCP = \sum (X - M_X)(Y - M_Y) \\ = (0 - 2.000)(4 - 6.000) + (0 - 2.000)(7 - 6.000) + (3 - 2.000)(4 - 6.000) + (5 - 2.000)(9 - 6.000) = 9.000$$

Covariance: The Covariance is a function of the Sum of Cross Products and the sample size:

$$COV = \frac{SCP}{(N - 1)} = \frac{9.000}{(4 - 1)} = 3.000$$

Unstandardized Regression Coefficients: The Unstandardized Regression Coefficients involve Covariance, the Standard Deviations of the variables, and the Means of the variables:

$$B_1 = \frac{COV}{(SD_X)^2} = \frac{3.000}{(2.449)^2} = 0.500$$

$$B_0 = M_Y - (B_1)(M_X) = 6.000 - (0.500)(2.000) = 5.000$$

Standardized Regression Coefficients: The Standard Regression Coefficients involve the Regression Coefficient (for the predictor) and the Standard Deviations of the variables:

$$\beta_1 = B_1 \left(\frac{SD_X}{SD_Y} \right) = 0.500 \left(\frac{2.449}{2.449} \right) = 0.500$$

Multiple Correlation: In bivariate regression, the Multiple Correlation is the same as the bivariate Correlation:

$$R = \frac{COV}{(SD_X)(SD_Y)} = \frac{3.000}{(2.449)(2.449)} = 0.500$$

Proportion of Variance Accounted For: The Proportion of Variance Accounted For is a function of the Correlation:

$$R^2 = 0.500^2 = 0.250$$

APA Style

Regression provides a prediction of the outcome variable by the first variable.

For the participants ($N = 4$), the scores on Outcome 1 ($M = 2.00$, $SD = 2.45$) moderately predicted Outcome 2 ($M = 6.00$, $SD = 2.45$), $\beta = .50$, $R^2 = .25$.

Note that regression coefficients can also have inferential information associated with them (and that this information should be summarized if it is available and of interest).

For the participants ($N = 4$), the scores on Outcome 1 ($M = 2.00$, $SD = 2.45$) did not significantly predict Outcome 2 ($M = 6.00$, $SD = 2.45$), $\beta = .50$, $t = 0.82$, $p = .500$.