Chapter 854

Tests for the Difference Between Two Linear Regression Slopes

Introduction

Linear regression is a commonly used procedure in statistical analysis. One of the main objectives in linear regression analysis is to test hypotheses about the slope and intercept of the regression equation. This module calculates power and sample size for testing whether two slopes from two groups are significantly different.

Technical Details

Suppose that the dependence of a variable *Y* on another variable *X* can be modeled using the simple linear regression equation

$$Y = \alpha + \beta X + \varepsilon$$

In this equation, α is the *Y*-intercept parameter, β is the slope parameter, *Y* is the dependent variable, *X* is the independent variable, and ε is the error. The nature of the relationship between *Y* and *X* is studied using a sample of *n* observations. Each observation consists of a data pair: the *X* value and the *Y* value. The parameters α and β are estimated using simple linear regression. We will call these estimates *a* and *b*.

Since the linear equation will not fit the observations exactly, estimated values must be used. These estimates are found using the method of least squares. Using these estimated values, each data pair may be modeled using the equation

$$Y = a + bX + e$$

Note that *a* and *b* are the estimates of the population parameters α and β. The *e* values represent the discrepancies between the estimated values ($a + bX$) and the actual values *Y*. They are called the errors or residuals.

Two Groups

Suppose there are two groups and a separate regression equation is calculated for each group. If it is assumed that these *e* values are normally distributed, a test of the hypothesis that $\beta 1 = \beta 2$ versus the alternative that they are unequal can be constructed. Dupont and Plummer (1998) state that the test statistic

$$D = \frac{(\hat{\beta}_2 - \hat{\beta}_1)\sqrt{n_2}}{S_R}$$

follows the Student's t distribution with *v* degrees of freedom where

$$v = n_1 + n_2 - 4$$

$$m = \frac{n_1}{n_2}$$

$$S_R^2 = s^2 \left[\frac{1}{m\sigma_{X_1}^2} + \frac{1}{\sigma_{X_2}^2} \right]$$

$$\sigma_{X_1}^2 = \frac{1}{n_1} \sum\nolimits_j \left(X_{1j} - \bar{X}_1 \right)^2$$

$$\sigma_{X_2}^2 = \frac{1}{n_2} \sum\nolimits_j \left(X_{2j} - \bar{X}_2 \right)^2$$

$$s^2 = \frac{1}{n_1 + n_2 - 4} \sum\nolimits_{ij} \left(Y_{ij} - \hat{Y}_{ij} \right)^2$$

$$\hat{Y}_{ij} = a_i + b_i X_{ij}$$

The power function of difference in slopes in a two-sided test is (see Dupont and Plummer, 1998) given by

$$\text{Power} = T_v\left[\delta\sqrt{n_2} - t_{v,\alpha/2}\right] + T_v\left[-\delta\sqrt{n_2} - t_{v,\alpha/2}\right]$$

where

$$N = n_1 + n_2$$

$$v = n_1 + n_2 - 4$$

$$m = \frac{n_1}{n_2}$$

$$\delta = \frac{\beta_2 - \beta_1}{\sigma_R}$$

$$\sigma_R^2 = \sigma^2 \left[\frac{1}{m\sigma_{X_1}^2} + \frac{1}{\sigma_{X_2}^2} \right]$$

$$\sigma_{X_1}^2 = \frac{1}{n_1} \sum_j \left(X_{1j} - \bar{X}_1 \right)^2$$

$$\sigma_{X_2}^2 = \frac{1}{n_2} \sum_j \left(X_{2j} - \bar{X}_2 \right)^2$$

$$\sigma^2 = \mathrm{Var}(\varepsilon)$$

Note that σ^2 is estimated by s^2.

Example 1 – Finding Sample Size

Suppose a sample size needs to be found for a study to compare two slopes. The parameters of the study are two-sided alpha of 0.05, power of 0.90, equal subject allocation to both groups, δ of 1, σ of 2, 3, or 4, and σX1 and σX2 of 2.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For	**Sample Size**
Alternative Hypothesis	**δ ≠ 0**
Power	**0.90**
Alpha	**0.05**
Group Allocation	**Equal (N1 = N2)**
δ (β1-β2, Slope Difference)	**1**
σ (SD of Residuals)	**2 3 4**
σX1 (SD of X in Group 1)	**2**
σX2 (SD of X in Group 2)	**2**

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Reports

Numeric Results

Solve For: Sample Size
Alternative Hypothesis: Ha: δ = β1 - β2 ≠ 0

Power		Sample Size			Slope Difference	Standard Deviation Residuals	Standard Deviation X Values		
Target	Actual	N1	N2	N	δ	σ	σX1	σX2	Alpha
0.9	0.91149	23	23	46	1	2	2	2	0.05
0.9	0.90403	49	49	98	1	3	2	2	0.05
0.9	0.90308	86	86	172	1	4	2	2	0.05

Target Power: The desired power value (or values) entered in the procedure. Power is the probability of rejecting a false null hypothesis.
Actual Power: The power obtained in this scenario. Because N1 and N2 are discrete, this value is often (slightly) larger than the target power.
N1 and N2: The number of items sampled from each group.
N: The total sample size. N = N1 + N2.
δ: The difference between population slopes at which power and sample size calculations are made. δ = β1 - β2.
σ: The standard deviation of the residuals.
σX1 and σX2: The assumed population standard deviations of X for groups 1 and 2, respectively. Note that the divisor is n, not n - 1.
Alpha: The probability of rejecting a true null hypothesis.

Summary Statements

A two-group simple linear regression (Y versus X) design will be used to test whether the Group 1 slope (β1) is different from the Group 2 slope (β2) (H0: δ = 0 versus H1: δ ≠ 0, δ = β1 - β2). The comparison will be made using a two-sided slope-difference t-test, with a Type I error rate (α) of 0.05. The standard deviation of X for Group 1 is assumed to be 2 and the standard deviation of X for Group 2 is assumed to be 2. The common standard deviation of residuals for both groups is assumed to be 2. To detect a slope difference (δ = β1 - β2) of 1 with 90% power, the number of needed subjects will be 23 in Group 1 and 23 in Group 2.

Dropout-Inflated Sample Size

	Sample Size			Dropout-Inflated Enrollment Sample Size			Expected Number of Dropouts		
Dropout Rate	**N1**	**N2**	**N**	**N1'**	**N2'**	**N'**	**D1**	**D2**	**D**
20%	23	23	46	29	29	58	6	6	12
20%	49	49	98	62	62	124	13	13	26
20%	86	86	172	108	108	216	22	22	44

Dropout Rate The percentage of subjects (or items) that are expected to be lost at random during the course of the study and for whom no response data will be collected (i.e., will be treated as "missing"). Abbreviated as DR.
N1, N2, and N The evaluable sample sizes at which power is computed. If N1 and N2 subjects are evaluated out of the N1' and N2' subjects that are enrolled in the study, the design will achieve the stated power.
N1', N2', and N' The number of subjects that should be enrolled in the study in order to obtain N1, N2, and N evaluable subjects, based on the assumed dropout rate. After solving for N1 and N2, N1' and N2' are calculated by inflating N1 and N2 using the formulas N1' = N1 / (1 - DR) and N2' = N2 / (1 - DR), with N1' and N2' always rounded up. (See Julious, S.A. (2010) pages 52-53, or Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. (2018) pages 32-33.)
D1, D2, and D The expected number of dropouts. D1 = N1' - N1, D2 = N2' - N2, and D = D1 + D2.

Dropout Summary Statements

Anticipating a 20% dropout rate, 29 subjects should be enrolled in Group 1, and 29 in Group 2, to obtain final group sample sizes of 23 and 23, respectively.

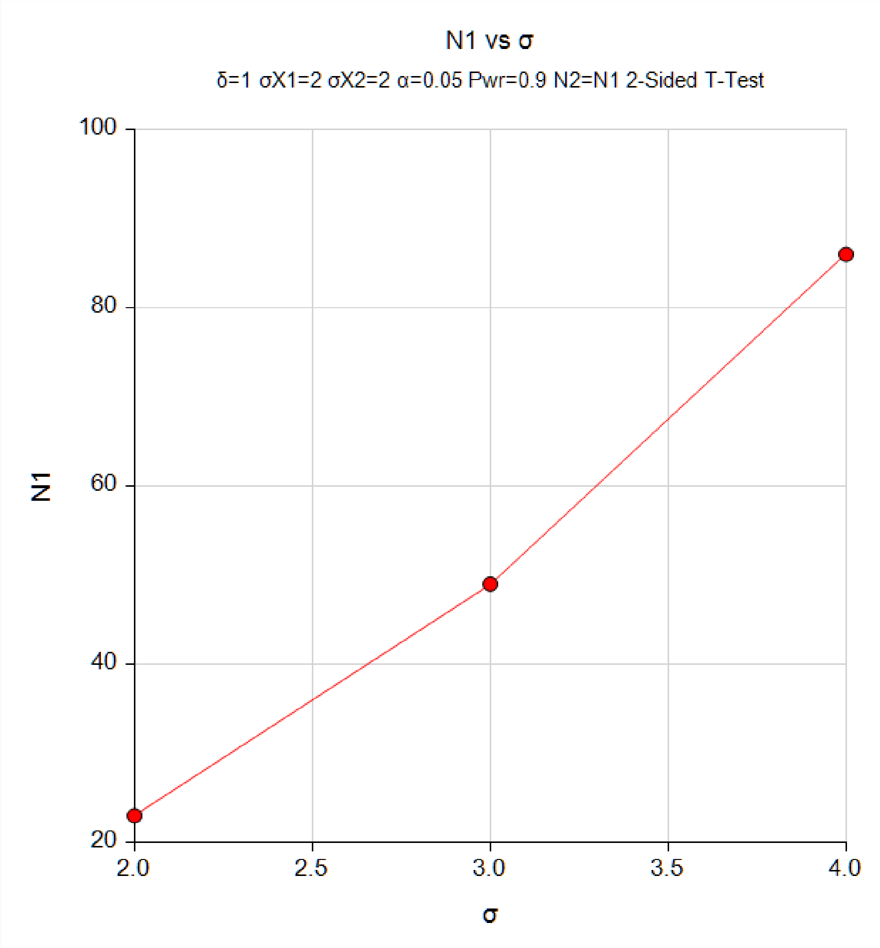
References

Dupont, W.D. and Plummer, W.D. Jr. 1998. Power and Sample Size Calculations for Studies Involving Linear Regression. Controlled Clinical Trials. Vol 19. Pages 589-601.

This report shows the calculated sample size for each of the scenarios.

Plots Section

Plots



This plot shows the sample size required for each value of σ.

Example 2 – Validation using Dupont and Plummer (1998)

Dupont and Plummer (1998, page 594-595) provide a worked example that we can use to validate this procedure. The parameters of the study are estimated to be a two-sided alpha of 0.05, power of 0.80, R = N2/N1 = 28/44 = 0.636, δ of -0.159, σ of 0.574, σX1 of 12.0, and σX2 of 9.19. They obtained N1 = 261 and N2 = 166.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For .... **Sample Size**
Alternative Hypothesis .... **δ ≠ 0**
Power.. **0.80**
Alpha.. **0.05**
Group Allocation .... **Enter R = N2/N1, solve for N1 and N2**
R **0.636**
δ (β1-β2, Slope Difference)........................... **-0.0159**
σ (SD of Residuals) **0.574**
σX1 (SD of X in Group 1).............................. **12**
σX2 (SD of X in Group 2).............................. **9.19**

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Solve For: Sample Size
Alternative Hypothesis: Ha: δ = β1 - β2 ≠ 0

Power		Sample Size			R (Allocation Ratio)		Slope Difference	Standard Deviation Residuals	Standard Deviation X Values		
Target	Actual	N1	N2	N	Target	Actual	δ	σ	σX1	σX2	Alpha
0.8	0.80003	263	167	430	0.636	0.635	-0.016	0.574	12	9.19	0.05

PASS obtained N1 = 263 and N2 = 167. This matches Dupont and Plummer within rounding. We checked the sample sizes that they found and obtained a power of 0.79748, which is slightly less than the desired power of 0.80. This is why **PASS**'s sample size is slightly larger in this case. If you set the desired power to 0.797, you will obtain the same results as they did.