Chapter 601

Hotelling's One-Sample T²

Introduction

This module calculates power for Hotelling's one-group T-squared (T^2) test statistic. Hotelling's One-Sample T^2 is an extension of the univariate one-sample T-test to the case where the number of response variables is greater than one.

Assumptions

The following assumptions are made when using Hotelling's T^2 to analyze one group of data.

1. The response variables are continuous.
2. The residuals follow the multivariate normal probability distribution with mean zero and constant variance-covariance matrix.
3. The subjects are independent.

Technical Details

The formulas used to perform a Hotelling's T^2 power analysis provide exact answers if the above assumptions are met. These formulas can be found in many places. We use the results in Rencher (1998). We refer you to that reference for more details.

Test Statistic

In one-group case, a set of *N* observations is available on *p* response variables. We assume that all *N* observations have the same multivariate normal distribution with mean vector μ and variance-covariance matrix Σ and that Hotelling's T^2 is used for testing the null hypothesis that $\mu = \mu_0$ versus the alternative that $\mu \neq \mu_0$ where at least one component of μ is different from the corresponding component of μ_0. Usually, μ_0 is a vector of zeros.

The value of T^2 is computed using the formula

$$T^2_{p,N-1} = N(\bar{y} - \mu_0)' S^{-1} (\bar{y} - \mu_0)$$

where $\bar{y}$ is the vector of sample means and S is the sample variance-covariance matrix.

Power Calculation

To calculate power, we need the non-centrality parameter for this distribution. With μ_1 as a vector of actual means, this non-centrality parameter is defined as follows

$$\lambda = N(\mu_1 - \mu_0)'\Sigma^{-1}(\mu_1 - \mu_0)$$

$$= N\Delta^2$$

where

$$\Delta = \sqrt{(\mu_1 - \mu_0)'\Sigma^{-1}(\mu_1 - \mu_0)}$$

We define Δ as the *effect size* because it provides an expression for the magnitude of the standardized difference between the null and alternative means.

Using this non-centrality parameter, the power of the Hotelling's T^2 may be calculated for any values of the means and standard deviations. Since there is a simple relationship between the non-central T^2 and the non-central *F*, calculations are actually based on the non-central *F* using the formula

$$\text{Power} = 1 - \Pr\left(F' < F'_{\alpha,df_1,df_2,\lambda}\right)$$

where

$$df_1 = p$$

$$df_2 = N - p$$

Variance-Covariance Matrix – Standard Deviation

The parameters in this section provide a flexible way to specify Σ, the variance-covariance matrix. Because the variance-covariance matrix is symmetric, it can be represented as

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} & \cdots & \sigma_1\sigma_p\rho_{1p} \\ \sigma_2\sigma_1\rho_{21} & \sigma_2^2 & \cdots & \sigma_2\sigma_p\rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_p\sigma_1\rho_{p1} & \sigma_p\sigma_2\rho_{p2} & \cdots & \sigma_p^2 \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_p \end{bmatrix} \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_p \end{bmatrix}.$$

where *p* is the number of response variables.

Thus, the covariance matrix can be represented with complete generality by specifying the standard deviations σ_1, σ_2, ..., σ_p and the correlation matrix

$$R = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & 1 \end{bmatrix}.$$

Example 1 – Power and Validation using Rencher (1998)

Rencher (1998) page 106 presents an example of power calculations for the one-group case in which the mean differences are both 1.88 such that

$$\mu_1 - \mu_0 = \begin{bmatrix} 1.88 \\ 1.88 \end{bmatrix}$$

and the variance-covariance matrix is

$$\Sigma = \begin{bmatrix} 56.78 & 11.98 \\ 11.98 & 29.28 \end{bmatrix}$$

When *N* is 25 and the significance level is 0.05, Rencher calculated the power to be 0.3397. To allow for a nice chart, we will calculate the power for several sample sizes and for *K* equal 1.0 and 1.5.

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 .... **Power**
Alpha... **0.05**
Sample Size.. **5 15 25 35 50 75 100 150**
Number of Response Variables **2**
Mean Difference Input Type........................... **List of Mean Differences**
List of Mean Differences **1.88 1.88**
K (Means Multiplier)....................................... **1 1.5**

Covariance Tab

Variance-Covariance Matrix Input Type......... **Variance-Covariance Matrix in Spreadsheet**
Columns Containing the V-C Matrix............... **VC1-VC2**

Input Spreadsheet Data

Row	Diffs	VC1	VC2
1	1.88	56.78	11.98
2	1.88	11.98	29.28

Output

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

Numeric Reports

Numeric Results

Solve For: Power
Mean Differences: 1.88, 1.88
Variance-Covariance Matrix: Input Columns = VC1-VC2

			Degrees of Freedom			
Power	Sample Size N	Means Multiplier K	Number of Y's DF1	DF2	Effect Size	Alpha
0.07372	5	1.0	2	3	0.38	0.05
0.19962	15	1.0	2	13	0.38	0.05
0.33787	25	1.0	2	23	0.38	0.05
0.47067	35	1.0	2	33	0.38	0.05
0.64085	50	1.0	2	48	0.38	0.05
0.83108	75	1.0	2	73	0.38	0.05
0.92822	100	1.0	2	98	0.38	0.05
0.98954	150	1.0	2	148	0.38	0.05
0.10403	5	1.5	2	3	0.57	0.05
0.40330	15	1.5	2	13	0.57	0.05
0.66345	25	1.5	2	23	0.57	0.05
0.83020	35	1.5	2	33	0.57	0.05
0.94754	50	1.5	2	48	0.57	0.05
0.99435	75	1.5	2	73	0.57	0.05
0.99952	100	1.5	2	98	0.57	0.05
1.00000	150	1.5	2	148	0.57	0.05

Power The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N The sample size, the number of subjects in the experiment or study.
K A constant by which all means are multiplied.
DF1 The first degrees of freedom for T^2. It is equal to the number of response variables (Y's).
DF2 The second degrees of freedom for T^2.
Effect Size A standardized version of T^2 under the alternative hypothesis.
Alpha The probability of rejecting a true null hypothesis.

Summary Statements

A single-group design with 2 response variables will be used to test whether the multivariate mean is different from the null multivariate mean. The comparison will be made using a one-sample Hotelling's T^2 test with a Type I error rate (α) of 0.05. The variance-covariance matrix is assumed to be defined as 'Input Columns = VC1-VC2'. To detect an effect size of 0.38 (derived from the mean differences '1.88, 1.88' and the variance-covariance matrix), with a sample size of 5, the power is 0.07372.

.
.
.

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	5	7	2
20%	15	19	4
20%	25	32	7
20%	35	44	9
20%	50	63	13
20%	75	94	19
20%	100	125	25
20%	150	188	38

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.
N The evaluable sample size at which power is computed (as entered by the user). If N subjects are evaluated out of the N' subjects that are enrolled in the study, the design will achieve the stated power.
N' The total number of subjects that should be enrolled in the study in order to obtain N evaluable subjects, based on the assumed dropout rate. N' is calculated by inflating N using the formula N' = N / (1 - DR), with N' 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.)
D The expected number of dropouts. D = N' - N.

Dropout Summary Statements

Anticipating a 20% dropout rate, 7 subjects should be enrolled to obtain a final sample size of 5 subjects.

References

Rencher, Alvin C. 1998. Multivariate Statistical Inference and Applications. John Wiley. New York, New York.

This report gives the power for each value of *N* and *K*. Notice that the power for *K* = 1 and *N* = 25 is 0.3379. This is slightly different than the 0.3397 obtained by interpolation by Rencher.

Means Matrix

Means Matrix

Name	Mean
Y1	1.88
Y2	1.88

This report shows the mean differences that were read in. When a Means Multiplier, *K*, is used, each value of *K* is multiplied times each of these values.

Variance-Covariance Matrix

Variance-Covariance Matrix

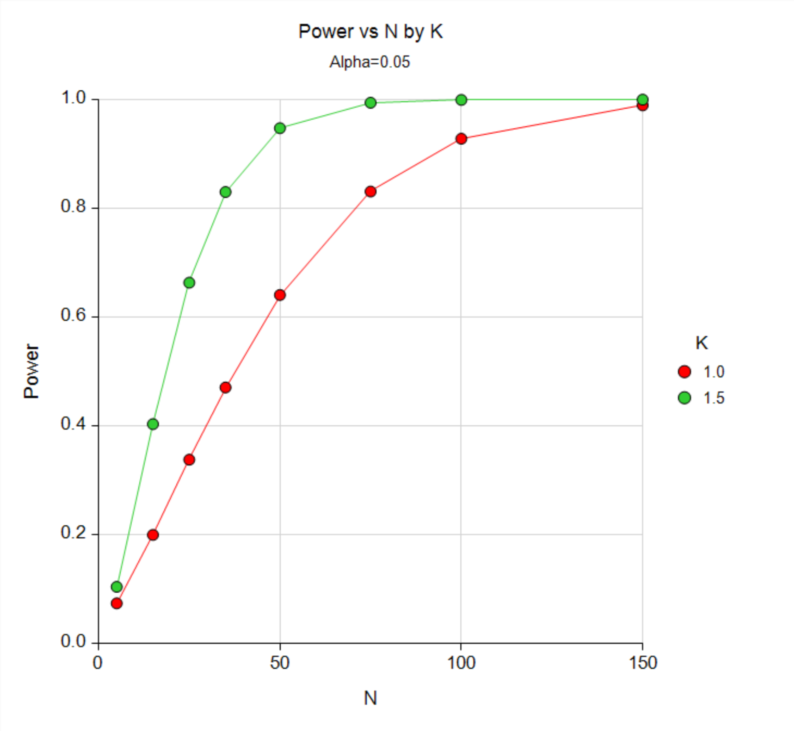
Response	Y1	Y2
Y1	7.535	0.294
Y2	0.294	5.411

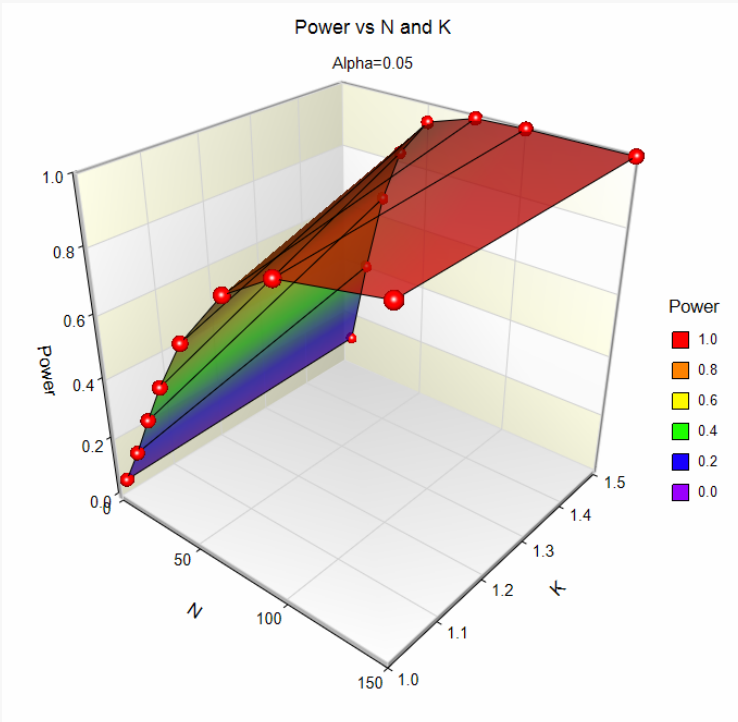
σ's on the diagonal, ρ's on the off diagonal(s)

This report shows a modified variance-covariance matrix with σ's on the diagonal, ρ's on the off diagonal(s). This matrix is created from the variance-covariance matrix that was read in from the spreadsheet or generated by the σ and ρ settings on the Covariance tab.

Plots Section

Plots





These plots show the relationship between power and *N* for each value of *K*.

The very same report can be generated by inputting the mean differences from the spreadsheet (**Example 1b**):

Design Tab

Mean Difference Input Type.......................... **Mean Differences in Spreadsheet**
Column Containing the Mean Differences ..... **DIFFS**

Input Spreadsheet Data

Row	Diffs	VC1	VC2
1	1.88	56.78	11.98
2	1.88	11.98	29.28