

Chapter 601

Hotelling's One-Sample T^2

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.

One-Group Technical Details

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 $\mu = \mu_A$ where at least one component of μ_A is different from the corresponding component of μ_0 . Usually, the vector μ_0 is a vector of zeros.

The value of T^2 is computed using the formula

$$T_{p,N-1}^2 = N(\bar{y} - \mu_0)' S^{-1}(\bar{y} - \mu_0)$$

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

To calculate power we need the non-centrality parameter for this distribution. This non-centrality parameter is defined as follows

Hotelling's One-Sample T²

$$\begin{aligned}\lambda &= N(\mu_A - \mu_0)' \Sigma^{-1}(\mu_A - \mu_0) \\ &= N\Delta^2\end{aligned}$$

where

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

We define Δ as *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 T² may be calculated for any value of the means and standard deviations. Since there is a simple relationship between the non-central T² and the non-central F , calculations are actually based on the non-central F using the formula

$$\beta = \Pr(F' < F'_{\alpha, df1, df2, \lambda})$$

where

$$df1 = p$$

$$df2 = N - p$$

Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design and Covariance tabs. For more information about the options of other tabs, go to the Procedure Window chapter.

Design Tab

The Design tab contains many of the options that you will be primarily concerned with.

Solve For

Solve For

This option specifies the parameter to be solved for.

When you choose to solve for *Sample Size*, the program searches for the lowest sample size that meets the alpha and power criterion you have specified.

Power and Alpha

Power

This option specifies one or more values for power. Power is the probability of rejecting a false null hypothesis, and is equal to one minus Beta. Beta is the probability of a type-II error, which occurs when a false null hypothesis is not rejected. In this procedure, a type-II error occurs when you fail to reject the null hypothesis of equal means when in fact the means are different.

Values must be between zero and one. Historically, the value of 0.80 (Beta = 0.20) was used for power. Now, 0.90 (Beta = 0.10) is also commonly used.

A single value may be entered here or a range of values such as *0.8 to 0.95 by 0.05* may be entered.

Hotelling's One-Sample T²

Alpha

This option specifies one or more values for the probability of a type-I error. A type-I error occurs when a true null hypothesis is rejected. In this procedure, a type-I error occurs when you reject the null hypothesis of equal means when in fact the means are equal.

Values must be between zero and one. Historically, the value of 0.05 has been used for alpha. This means that about one test in twenty will falsely reject the null hypothesis. You should pick a value for alpha that represents the risk of a type-I error you are willing to take in your experimental situation.

You may enter a range of values such as *0.01 0.05 0.10* or *0.01 to 0.10 by 0.01*.

Sample Size

Sample Size

This is the number of individuals measures (multiple times).

Effect Size – Response Variables

Number of Response Variables

Enter the number of response (dependent or Y) variables. For a true multivariate test, this value will be greater than one.

The number of mean differences entered in the Mean Differences box or in the Means column must equal this value. If you read-in the covariance matrix from the spreadsheet, the number of columns specified must equal this value.

Effect Size – Mean Differences

Mean Differences (= # of Response Vars)

Enter a list of values representing the mean differences under the alternative hypothesis. Under the null hypothesis, these values are all zero. The values entered here represent the differences that you want the experiment (study) to be able to detect.

Note that the number of values must match the number of Response Variables.

If you like, you can enter these values in a column on the spreadsheet. This column is specified using the 'Means Column' option. When that option is specified, any values entered here are ignored.

Means Differences Column

Use this option to specify the spreadsheet column containing the hypothesized mean differences. The response variables are represented down the rows. The number of rows with data must equal the number of response variables. When this option is used, the 'Mean Differences' box is ignored.

You can obtain the spreadsheet by selecting 'Window', then 'Data', from the menus.

Effect Size – Mean Multiplier

K (Means Multipliers)

These values are multiplied times the mean differences to give you various effect sizes. A separate power calculation is generated for each value of K. If you want to ignore this setting, enter '1'.

Covariance Tab

This tab specifies the covariance matrix.

Covariance Matrix Specification

Specify Which Covariance Matrix Input Method to Use

This option specifies which method will be used to define the covariance matrix.

- **Standard Deviation and Correlation**

This option generates a covariance matrix based on the settings for the standard deviation (SD) and the pattern of correlations as specified in the Correlation Pattern and R options.

- **Covariance Matrix Variables**

When this option is selected, the covariance matrix is read in from the columns of the spreadsheet. This is the most flexible method, but specifying a covariance matrix is tedious. You will usually only use this method when a specific covariance is given to you.

Note that the spreadsheet is shown by selecting the menus: 'Window' and then 'Data'.

Covariance Matrix Specification- Input Method = 'Standard Deviation and Correlation'

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

$$\begin{aligned} \Sigma &= \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1p} & \sigma_{2p} & \cdots & \sigma_{pp} \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} & \cdots & \sigma_1\sigma_p\rho_{1p} \\ \sigma_1\sigma_2\rho_{12} & \sigma_2^2 & \cdots & \sigma_2\sigma_p\rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_1\sigma_p\rho_{1p} & \sigma_2\sigma_p\rho_{2p} & \cdots & \sigma_p^2 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_p \end{bmatrix} \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{12} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1p} & \rho_{2p} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_p \end{bmatrix} \end{aligned}$$

where p is the number of response variables.

Thus, the covariance matrix can be represented with complete generality by specifying the standard deviations $\sigma_1, \sigma_2, \dots, \sigma_p$ and the correlation matrix

Hotelling's One-Sample T²

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

SD (Common Standard Deviation)

This value is used to generate the covariance matrix. This option specifies a single standard deviation to be used for all response variables. The square of this value becomes the diagonal elements of the covariance matrix. Since this is a standard deviation, it must be greater than zero.

This option is only used when the first Covariance Matrix Input Method is selected.

R (Correlation)

Specify a correlation to be used in calculating the off-diagonal elements of the covariance matrix. Since this is a correlation, it must be between -1 and 1. This option is only used when the first Covariance Matrix Input Method is selected.

Specify Correlation Pattern

This option specifies the pattern of the correlations in the variance-covariance matrix. Two options are available:

- **Constant**

The value of R is used as the constant correlation. For example, if $R = 0.6$ and $p = 6$, the correlation matrix would appear as

$$R = \begin{bmatrix} 1 & 0.600 & 0.600 & 0.600 & 0.600 & 0.600 \\ 0.600 & 1 & 0.600 & 0.600 & 0.600 & 0.600 \\ 0.600 & 0.600 & 1 & 0.600 & 0.600 & 0.600 \\ 0.600 & 0.600 & 0.600 & 1 & 0.600 & 0.600 \\ 0.600 & 0.600 & 0.600 & 0.600 & 1 & 0.600 \\ 0.600 & 0.600 & 0.600 & 0.600 & 0.600 & 1 \end{bmatrix}$$

- **1st-Order Autocorrelation**

The value of R is used as the base autocorrelation in a first-order, serial correlation pattern. For example, $R = 0.6$ and $p = 6$, the correlation matrix would appear as

$$R = \begin{bmatrix} 1 & 0.600 & 0.360 & 0.216 & 0.130 & 0.078 \\ 0.600 & 1 & 0.600 & 0.360 & 0.216 & 0.130 \\ 0.360 & 0.600 & 1 & 0.600 & 0.360 & 0.216 \\ 0.216 & 0.360 & 0.600 & 1 & 0.600 & 0.360 \\ 0.130 & 0.216 & 0.360 & 0.600 & 1 & 0.600 \\ 0.078 & 0.130 & 0.216 & 0.360 & 0.600 & 1 \end{bmatrix}$$

This pattern is often chosen as the most realistic when little is known about the correlation pattern and the responses variables are measured across time.

Hotelling's One-Sample T²**Covariance Matrix Specification- Input Method =
'Covariance Matrix Variables'**

This option instructs the program to read the covariance matrix from the spreadsheet.

Spreadsheet Columns Containing the Covariance Matrix

This option designates the columns on the current spreadsheet holding the covariance matrix. It is used when the 'Specify Which Covariance Matrix Input Method to Use' option is set to *Covariance Matrix Variables*. The number of columns and number of rows must match the number of response variable at which the subjects are measured.

Example 1 – Power and Validation

Rencher (1998) page 106 presents an example of power calculations for the one-group case in which the mean differences are both 1.88 and the 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 samples sizes and for K equal 1.0 and 1.5.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Hotelling's One-Sample T²** procedure window by expanding **Means**, then clicking on **Multivariate Means**, and then clicking on **Hotelling's One-Sample T²**. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**. You can see that the values have been loaded into the spreadsheet by clicking on the spreadsheet button.

| <u>Option</u> | <u>Value</u> |
|---|----------------------------------|
| Design Tab | |
| Solve For | Power |
| Alpha..... | 0.05 |
| Sample Size..... | 5 15 25 35 50 75 100 150 |
| Number of Response Variables..... | 2 |
| Mean Differences..... | 1.88 1.88 |
| Mean Differences Column | <i>blank</i> |
| K (Means Multiplier)..... | 1.0 1.5 |
| Covariance Tab | |
| Specify Covariance Method..... | Covariance Matrix Columns |
| Spreadsheet Columns | VC1-VC2 |
| Press the Spreadsheet button to enter the following values into the spreadsheet for columns VC1 and VC2 | |
| Row 1 | 56.78 11.98 |
| Row 2..... | 11.98 29.28 |
| Reports Tab | |
| Show Numeric Reports..... | Checked |
| Show Means Matrix | Checked |
| Show Covariance Matrix..... | Checked |

Hotelling's One-Sample T²

Annotated Output

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

Numeric Report

| Power | N | Multiply Means By | Alpha | Beta | Effect Size | Number of Y's (DF1) | DF2 |
|--------|-----|-------------------------|--------|--------|----------------|---------------------------|-----|
| 0.0737 | 5 | 1.0000 | 0.0500 | 0.9263 | 0.38 | 2 | 3 |
| 0.1996 | 15 | 1.0000 | 0.0500 | 0.8004 | 0.38 | 2 | 13 |
| 0.3379 | 25 | 1.0000 | 0.0500 | 0.6621 | 0.38 | 2 | 23 |
| 0.4707 | 35 | 1.0000 | 0.0500 | 0.5293 | 0.38 | 2 | 33 |
| 0.6409 | 50 | 1.0000 | 0.0500 | 0.3591 | 0.38 | 2 | 48 |
| 0.8311 | 75 | 1.0000 | 0.0500 | 0.1689 | 0.38 | 2 | 73 |
| 0.9282 | 100 | 1.0000 | 0.0500 | 0.0718 | 0.38 | 2 | 98 |
| 0.9895 | 150 | 1.0000 | 0.0500 | 0.0105 | 0.38 | 2 | 148 |
| 0.1040 | 5 | 1.5000 | 0.0500 | 0.8960 | 0.38 | 2 | 3 |
| 0.4033 | 15 | 1.5000 | 0.0500 | 0.5967 | 0.38 | 2 | 13 |
| 0.6635 | 25 | 1.5000 | 0.0500 | 0.3365 | 0.38 | 2 | 23 |
| 0.8302 | 35 | 1.5000 | 0.0500 | 0.1698 | 0.38 | 2 | 33 |
| 0.9475 | 50 | 1.5000 | 0.0500 | 0.0525 | 0.38 | 2 | 48 |
| 0.9943 | 75 | 1.5000 | 0.0500 | 0.0057 | 0.38 | 2 | 73 |
| 0.9995 | 100 | 1.5000 | 0.0500 | 0.0005 | 0.38 | 2 | 98 |
| 1.0000 | 150 | 1.5000 | 0.0500 | 0.0000 | 0.38 | 2 | 148 |

Report Definitions

Power is the probability of rejecting a false null hypothesis. Note that Power = 1 - Beta.

N is the sample size, the number of subjects in the experiment or study.

K is a constant by which all means are multiplied.

Alpha is the probability of rejecting a true null hypothesis.

Beta is the probability of accepting a false null hypothesis. Note that Beta = 1 - Power.

Effect Size is a standardized version of T2 under the alternative hypothesis.

DF1 is the first degrees of freedom of T2. It is the number of response variables.

DF2 is the second degrees of freedom of T2.

Summary Statements

A sample size of 5 achieves 7% power to detect an effect size of 0.38 which represents the differences between the null and alternative means of the 2 response variables, adjusted by the variance-covariance matrix. The one-sample Hotelling's T-squared test statistic is used with a significance level of 0.0500.

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 Section

| Means Section | |
|---------------|--------|
| Name | Mean |
| Y1 | 1.8800 |
| Y2 | 1.8800 |

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.

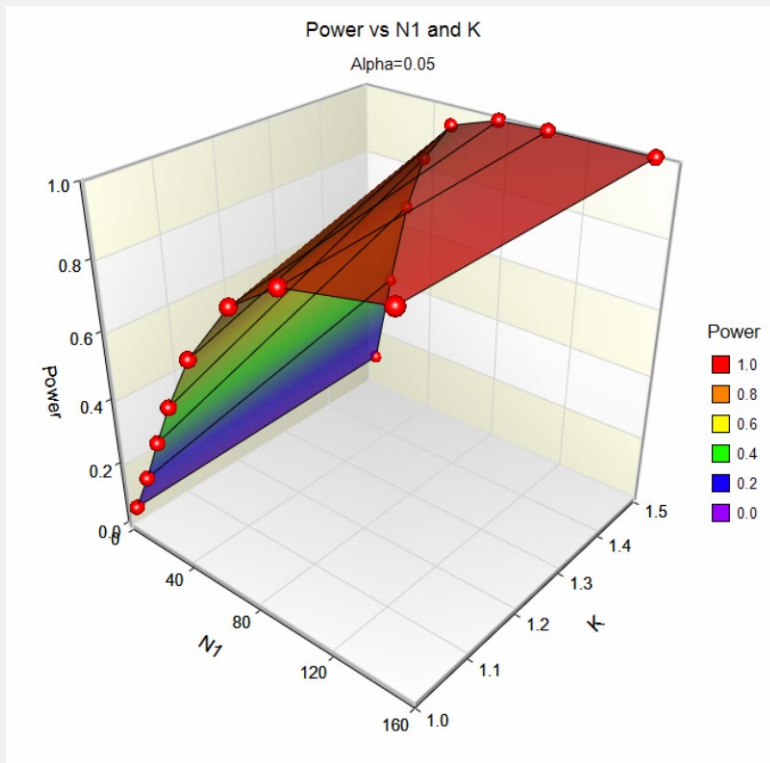
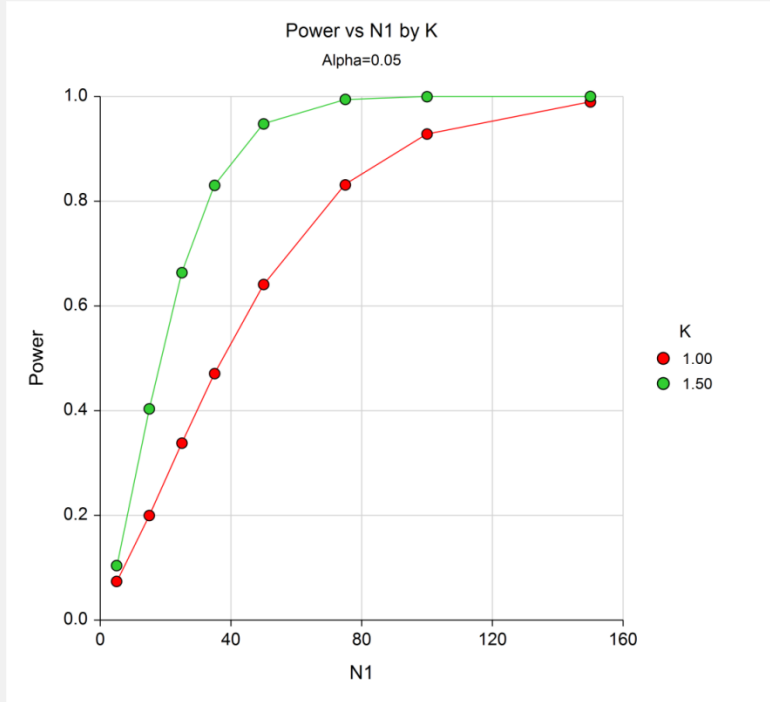
Variance-Covariance Matrix Section

| Variance-Covariance Matrix Section | | |
|------------------------------------|--------|--------|
| Response | Y1 | Y2 |
| Y1 | 7.5353 | 0.2938 |
| Y2 | 0.2938 | 5.4111 |

Hotelling's One-Sample T²

This report shows the variance-covariance matrix that was read in from the spreadsheet or generated by the settings of on the Covariance tab. The standard deviations are given on the diagonal and the correlations are given off the diagonal.

Chart Section



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