Chapter 584

Non-Zero Null Studentized Range Tests

Introduction

This procedure computes power and sample size of tests of whether the means of two or more groups which are analyzed using a studentized range test are more than just trivially different (non-zero null). This is similar to the superiority by a margin tests, except that this is a two-sided tested.

Methodology for testing equality among three or more groups has received little attention, especially when the possibility of a non-zero null is considered. An article by Shieh (2018) gives results for two competing test procedures: The F-test and the studentized range test. Results for the F-test are available in PASS in another procedure. This procedure provides power and sample size results the studentized range test.

While the F-test is by far the most commonly used method for testing the equality of two or more means, Shieh (2018) showed that neither test is always optimal. In fact, the studentized range test is more powerful when the actual range is close to the non-zero null boundary.

Technical Details for the Studentized Range Test

Suppose G groups each have a normal distribution and with means $\mu_1, \mu_2, \dots, \mu_G$ and common variance σ^2. Let $N_1, N_2, \dots, N_G = N_i$ denote the common sample size of all groups and let N denote the total sample size. In this case of equal group sizes, $N = GN_i$. The multigroup test problem requires one to show that the means are more than trivially different. Shieh (2018) considered whether the difference between the minimum and maximum means (the range of the means) is sufficiently small so that the differences among the means can be regarded as of no practical importance.

The One-Way Model

Consider the usual one-way fixed-effects model

$$Y_{gj} = \mu_g + \varepsilon_{gj}$$

where Y_{gj} is response, μ_g are the treatment means, and ε_{gj} are the independent, normally distributed error with zero mean and common variance σ^2. Here the subscript g indexes the G groups, and the subscript j indexes the N_i subjects in each group.

Cohen (1988) showed that hypotheses about the G means may be obtained using either the variance of the means in terms of the F-test or their range in terms of the studentized range.

Equality Hypothesis with Non-Zero Null

The hypothesis of mean equality with non-zero null is

$$H_0: \frac{\delta}{\sigma} \le \frac{\delta_0}{\sigma} \ versus \ H_1: \frac{\delta}{\sigma} > \frac{\delta_0}{\sigma}$$

where $\delta = \mu_{Max} - \mu_{Nin}$ represents the range and δ_0 is the non-zero null margin.

Studentized Range Statistic

The studentized range statistic is defined as follows

$$Q = \frac{\left[\max_{g=1\ to\ G}\left(\bar{Y}_g\right) - \min_{g=1\ to\ G}\left(\bar{Y}_g\right)\right]\sqrt{N_i}}{S}$$

where $\bar{Y}_g$ are the sample means and S is the sample variance.

It turns out that the distribution of Q is a function of the pairwise mean differences $\mu_g - \mu_h$, not just the range (the maximum of these differences).

The cumulative distribution function, from which the power can be computed, is given by

$$\Theta(q) = P\{Q \le q\} = E_K\left\{\sum_{g=1}^{G} E_{Z_g}\left[\prod_{\substack{h=1\\h\neq g}}^{G}\left(\Phi\left\{Z_g + \delta_{gh}\sqrt{N_i}\right\} - \Phi\left\{Z_g + \delta_{gh}\sqrt{N_i} - q\sqrt{K/(N-G)}\right\}\right)\right]\right\}$$

where $\delta_{gh} = \mu_g - \mu_h$, K is a chi-square random variable with $N - G$ degrees of freedom, $\Phi\{z\}$ is the CDF of a standard normal distribution, Z_g are independent standard normal random variables, $E_K\{x\}$ is the expectation with respect to K, and $E_{Z_g}\{x\}$ is the expectation with respect to Z_g.

Note that the critical value is based on the set of group means. It cannot be determined from just δ_0. When only δ_0 is specified, the least favorable configuration (LFC) of the means is used. This is given by a vector with $G/2$ means equal to $-\frac{\delta_0}{2}$ and the rest equal to $\frac{\delta_0}{2}$ as follows

$$\{\mu_1, \ldots, \mu_G\} = \left\{-\frac{\delta_0}{2}, \ldots, -\frac{\delta_0}{2}, \frac{\delta_0}{2}, \ldots, \frac{\delta_0}{2}\right\}$$

If a sample size is desired, it can be determined using a standard binary search algorithm.

Example 1 – Finding Power

An experiment is being designed to test whether the range of the maximum difference among four group means is at least greater than a non-zero threshold. The hypothesis test will use studentized range test at a significance level of 0.05. Previous studies have shown a standard deviation of 2. The minimal range of the four means is 1. Power calculations assume that the actual range is 2.

To better understand the relationship between power and sample size, the researcher wants to compute the power for several group sample sizes between 20 and 120. The sample sizes will be equal across all groups.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. 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**.

Option	Value
Design Tab	
Solve For	**Power**
Alpha	**0.05**
G (Number of Groups)	**4**
Ni (Sample Size Per Group)	**20 40 60 80 100 120**
μ1i's Input Type	**Enter Range of Means|H1**
δ1 (Range of Means|H1)	**2**
μ0i's Input Type	**Enter Range of Means|H0**
δ0 (Range of Means|H0)	**1**
σ (Standard Deviation)	**2**

Annotated Output

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

Numeric Results

Numeric Results

Number of Groups 4

Power	Total Sample Size N	Sample Size Per Group Ni	H1 Range of Means δ1	H0 Range of Means δ0	Std Dev σ	Alpha
0.2673	80	20	2	1	2	0.05
0.5025	160	40	2	1	2	0.05
0.6918	240	60	2	1	2	0.05
0.8214	320	80	2	1	2	0.05
0.9018	400	100	2	1	2	0.05
0.9480	480	120	2	1	2	0.05

References

Shieh, G. 2018. 'On Detecting a Minimal Important Difference among Standardized Means'. Current Psychology, Vol 37, Pages 640-647. Doi: 10.1007/s12144-016-9549-5
Cohen, Jacob. 1988. Statistical Power Analysis for the Behavioral Sciences. Lawrence Erlbaum Associates. Hillsdale, New Jersey

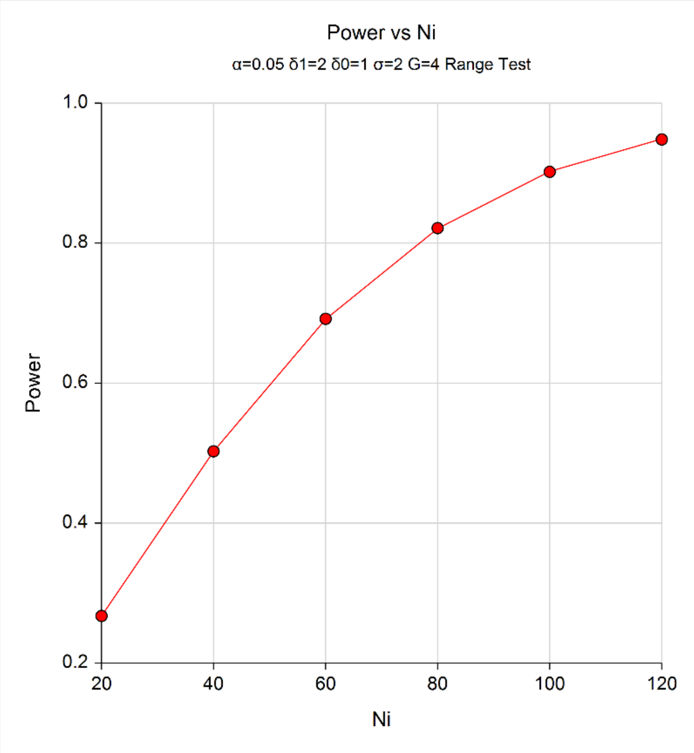
Report Definitions
Power is the probability of rejecting a false null hypothesis in favor of the alternative hypothesis.
Total Sample Size N is the total number of subjects in the study.
Sample Size Per Group Ni is the number of subjects sampled per group.
H1 Range of Means δ1 is the range of the group means assumed by the alternative hypothesis. It is the value at which the power is computed. Note that you must have 0 ≤ δ0 < δ1.
H0 Range of Means δ0 is the range of the group means assumed by the null hypothesis. This value is the equivalence limit (bound). Note that you must have 0 ≤ δ0 < δ1.
Std Dev σ is the standard deviation of the responses for all groups.
Alpha is the significance level of the test: the probability of rejecting the null hypothesis when it is actually true.

Summary Statements
In a one-way study with a non-zero null hypothesis, a sample of 80 subjects, divided among 4 groups, achieves a power of 27%. This power assumes a studentized range test of a non-zero null margin at a significance level of 0.05. The group subject counts are 20. The non-zero margin under the null hypothesis is 1. The range at which the power is computed is 2. The standard deviation of all groups is 2.

This report shows the numeric results of this power study.

Chart Section

Chart Section



This plot gives a visual presentation of the results in the Numeric Report.

Example 2 – Finding the Sample Size Necessary to Reject

Continuing with the last example, we will determine how large the sample size would need to have been for alpha = 0.05 and power = 0.80 or 0.9.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

Option	Value
Design Tab	
Solve For	**Sample Size**
Power	**0.8 0.9**
Alpha	**0.05**
G (Number of Groups)	**4**
μ1i's Input Type	**Enter Range of Means|H1**
δ1 (Range of Means|H1)	**2**
μ0i's Input Type	**Enter Range of Means|H0**
δ0 (Range of Means|H0)	**1**
σ (Standard Deviation)	**2**

Output

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

Numeric Results

Numeric Results
Number of Groups 4

Power	Total Sample Size N	Sample Size Per Group Ni	H1 Range of Means δ1	H0 Range of Means δ0	Std Dev σ	Alpha
0.8055	308	77	2	1	2	0.05
0.9018	400	100	2	1	2	0.05

This report shows the necessary sample sizes for achieving powers of 0.8 and 0.9.

Example 3 – Validation using Shieh (2018)

Shieh (2018) page 644 presents an example in which alpha = 0.05, G = 3, σ = 3.189, means under alternative hypothesis are {7.77, 9.77, 6.68}, δ0 = 0.2(σ) = 0.637809, and power = 0.8. The resulting sample size was 27 per group for a total of 81.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. You may then make the appropriate entries as listed below, or open **Example 3** by going to the **File** menu and choosing **Open Example Template**.

Option	Value
Design Tab	
Solve For	**Sample Size**
Power	**0.8**
Alpha	**0.05**
G (Number of Groups)	**3**
μ1i's Input Type	**Enter μ11, μ12, ..., μ1G**
μ11, μ12, ..., μ1G	**7.77 9.77 6.68**
μ0i's Input Type	**Enter Range of Means|H0**
δ0 (Range of Means|H0)	**0.637809**
σ (Standard Deviation)	**3.189**

Output

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

Numeric Results

Numeric Results
Number of Groups 3

Power	Total Sample Size N	Sample Size Per Group Ni	H1 Group Means Set μ1i	H1 Range of Means δ1	H0 Range of Means δ0	Std Dev σ	Alpha
0.8088	81	27	μ1i(1)	3.09	0.638	3.189	0.05

Name	Values
μ1i(1)	7.77, 9.77, 6.68

PASS also found Ni = 27 and N = 81.