

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

$$H_0: \frac{\delta}{\sigma} \geq \frac{\delta_0}{\sigma} \quad \text{versus} \quad H_1: \frac{\delta}{\sigma} < \frac{\delta_0}{\sigma}$$

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

Studentized Range Statistic

The studentized range statistic is defined as follows

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

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

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

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

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

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

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
G (Number of Groups)	4
Ni (Sample Size Per Group)	20 40 60 80 100 120
μ_{1i} 's Input Type	Enter Range of Means H1
δ_1 (Range of Means H1).....	2
μ_{0i} 's Input Type	Enter Range of Means H0
δ_0 (Range of Means H0).....	1
σ (Standard Deviation).....	2

Non-Zero Null Studentized Range Tests

Output

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

Numeric Reports

Numeric Results

Solve For: **Power**
 Number of Groups: 4

Power	Sample Size		Range of Means		Standard Deviation σ	Alpha
	Total N	Group Ni	H1 δ_1	H0 (Non-Zero Null Margin) δ_0		
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

Power The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
 N The total number of subjects in the study.
 Ni The number of subjects sampled per group.
 δ_1 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 $0 \leq \delta_0 < \delta_1$.
 δ_0 The range of the group means assumed by the null hypothesis. This value is the equivalence limit (bound). Note that you must have $0 \leq \delta_0 < \delta_1$.
 σ The standard deviation of the responses for all groups.
 Alpha The probability of rejecting a true null hypothesis.

Summary Statements

A one-way design with 4 groups will be used to test whether the range (maximum difference in means) is different from the non-zero margin 1. The comparison will be made using a studentized range test (with a non-zero null margin) with a Type I error rate (α) of 0.05. The common within-group standard deviation of responses for all groups is assumed to be 2. To detect a range of group means of 2, with group sample sizes of 20 subjects per group (for a total of 80 subjects), the power is 0.2673.

Non-Zero Null Studentized Range Tests

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout- Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	80	100	20
20%	160	200	40
20%	240	300	60
20%	320	400	80
20%	400	500	100
20%	480	600	120

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 Lohknygina, Y. (2018) pages 32-33.)
D	The expected number of dropouts. $D = N' - N$.

Dropout Summary Statements

Anticipating a 20% dropout rate, 100 subjects should be enrolled to obtain a final sample size of 80 subjects.

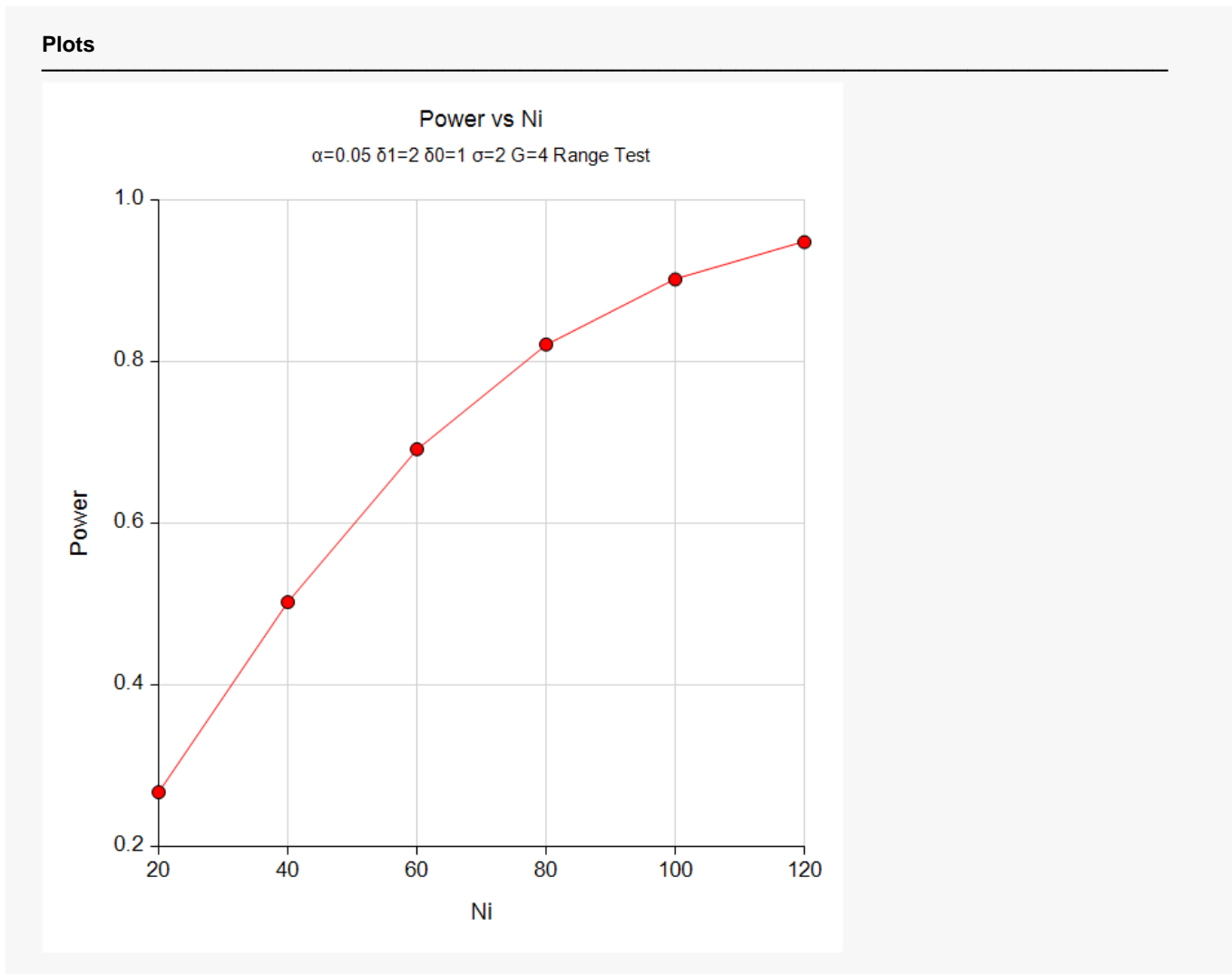
References

- 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.

This report shows the numeric results of this power study.

Non-Zero Null Studentized Range Tests

Plots 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

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**
 Power..... **0.8 0.9**
 Alpha..... **0.05**
 G (Number of Groups) **4**
 μ_1 's Input Type **Enter Range of Means|H1**
 δ_1 (Range of Means|H1)..... **2**
 μ_0 's Input Type **Enter Range of Means|H0**
 δ_0 (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

Solve For: [Sample Size](#)
 Number of Groups: 4

Power	Sample Size		Range of Means		Standard Deviation σ	Alpha
	Total N	Group Ni	H1 δ_1	H0 (Non-Zero Null Margin) δ_0		
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 $\alpha = 0.05$, $G = 3$, $\sigma = 3.189$, means under alternative hypothesis are $\{7.77, 9.77, 6.68\}$, $\delta_0 = 0.2(\sigma) = 0.637809$, and power = 0.8. The resulting sample size was 27 per group for a total of 81.

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 3** 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**
 Power..... **0.8**
 Alpha..... **0.05**
 G (Number of Groups) **3**
 μ_{1i} 's Input Type **Enter $\mu_{11}, \mu_{12}, \dots, \mu_{1G}$**
 $\mu_{11}, \mu_{12}, \dots, \mu_{1G}$ **7.77 9.77 6.68**
 μ_{0i} 's Input Type **Enter Range of Means|H0**
 δ_0 (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

Solve For: [Sample Size](#)
 Number of Groups: 3

Power	Sample Size		H1		H0 (Null Margin)	Standard Deviation σ	Alpha
	Total N	Group Ni	Means μ_{1i}	Range δ_1	Range δ_0		
	0.8088	81	27	$\mu_{1i(1)}$	3.09		

Item	Values
$\mu_{1i(1)}$	7.77, 9.77, 6.68

PASS also found $N_i = 27$ and $N = 81$.