Chapter 159

Multi-Arm Non-Inferiority Tests for the Difference Between Treatment and Control Proportions

Introduction

This module computes power and sample size for multi-arm, noninferiority tests of the difference between treatment and control proportions. This procedure is based on the results in Machin, Campbell, Tan, and Tan (2018). In this design, there are *k* treatment groups and one control group. The groups are independent and are sampled using simple random sampling. A proportion is measured in each group. A total of *k* hypothesis tests are anticipated, each comparing a treatment group with the common control group using a noninferiority test of the difference between two proportions.

The Bonferroni multiplicity adjustment of the type I error rate may be optionally made because several tests are being constructed from the same data. Making a multiplicity adjustment is usually recommended, but not always. In fact, Saville (1990) advocates not applying it and Machin, Campbell, Tan, and Tan (2018) include omitting it as a possibility.

Whether you want to test several doses of a single treatment or several types of treatments, good research practice requires that each treatment be compared with a control. For example, a popular three-arm design consists of three groups: control, treatment A, and treatment B. Two tests are run: treatment A versus control and treatment B versus the same control. This avoids having to obtain a second control group for treatment B. Besides the obvious efficiency in subjects, it may be easier to recruit subjects if their chances of receiving a new treatment are better than 50%.

Example

Suppose that the current treatment for a disease works 70% of the time. Unfortunately, this treatment is expensive and occasionally exhibits serious side-effects. Two promising new treatments have been developed and are now ready to be tested. Hence, three groups are needed to complete this study. Two non-inferiority hypotheses need to be tested in this study: whether each new treatment is as good as the current treatment.

Because of the many benefits of the new treatment, clinicians are willing to adopt a new treatment even if it is slightly less effective than the current treatment. They must determine, however, how much less effective the new treatment can be and still be adopted. Should it be adopted if 69% respond? 68%? 65%? 60%? There is a percentage below 70% at which the difference between the two treatments is no longer considered ignorable. After thoughtful discussion with several clinicians, it was decided that if a response of at least 63% is achieved, the new treatment will be adopted. The difference between these two percentages is called the *margin of non-inferiority*. The margin of non-inferiority in this example is 7%.

The developers must design an experiment to test the hypothesis that the response rate of the new treatment is at least 0.63. The statistical hypotheses to be tested are

$$H_0: P_A \leq P_C + \delta_0 \quad \text{vs.} \quad H_1: P_A > P_C + \delta_0$$

$$H_0: P_B \leq P_C + \delta_0 \quad \text{vs.} \quad H_1: P_B > P_C + \delta_0$$

where $\delta_0 = -0.07$.

Notice that when the null hypothesis is rejected, the conclusion is that the response rate of the treatment group is at least 0.63. Note that even though the response rate of the current treatment is 0.70, the hypothesis test is about a response rate of 0.63. Also notice that a rejection of the null hypothesis results in the conclusion of interest.

Technical Details

Suppose you have *k* treatment groups with response probabilities P_i of size N_i and one control group with response probability P_C of size N_C. The total sample size is $N = N_1 + N_2 + \ldots + N_k + N_C$.

The *k* one-sided non-inferiority tests are

$$H_{0i}: P_i - P_C \leq \delta_0 \quad \text{vs.} \quad H_{1i}: P_i - P_C > \delta_0 \quad \text{for } i = 1, 2, \ldots, k$$

Note that if higher proportions are better, $\delta_0 < 0$ and if lower proportions are better, $\delta_0 > 0$.

If we define $\delta_i = P_i - P_C$, these are equivalent to

$$H_{0i}: \delta_i \leq \delta_0 \quad \text{vs.} \quad H_{1i}: \delta_i > \delta_0 \quad \text{for } i = 1, 2, \ldots, k$$

For convenience, these hypotheses are collectively referred to as

$$H_0: \delta \leq \delta_0 \quad \text{vs.} \quad H_1: \delta > \delta_0$$

Test Statistics

Several test statistics are available in this routine. These are

Z-Test (Pooled)

This test was first proposed by Karl Pearson in 1900. Although this test is usually expressed directly as a chi-square statistic, it is expressed here as a *z* statistic so that it can be more easily used for one-sided hypothesis testing. The proportions are pooled (averaged) in computing the standard error. The formula for the test statistic is

$$z_t = \frac{\hat{p}_i - \hat{p}_C - \delta_0}{\hat{\sigma}_1}$$

where

$$\hat{\sigma}_1 = \sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_i}+\frac{1}{n_C}\right)}$$

$$\bar{p} = \frac{n_i\hat{p}_i + n_C\hat{p}_C}{n_i + n_C}$$

Z-Test (Unpooled)

This test statistic does not pool the two proportions in computing the standard error.

$$z_t = \frac{\hat{p}_i - \hat{p}_C - \delta_0}{\hat{\sigma}_2}$$

where

$$\hat{\sigma}_2 = \sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{n_i} + \frac{\hat{p}_C(1-\hat{p}_C)}{n_C}}$$

Z-Test with Continuity Correction (Pooled)

This test is the same as Z-Test (Pooled), except that a continuity correction is used. Remember that in the null case, the continuity correction makes the results closer to those of Fisher's Exact test.

$$z_t = \frac{\hat{p}_i - \hat{p}_C - \delta_0 + \frac{F}{2}\left(\frac{1}{n_i}+\frac{1}{n_C}\right)}{\hat{\sigma}_1}$$

where

$$\hat{\sigma}_1 = \sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_i}+\frac{1}{n_C}\right)}$$

$$\bar{p} = \frac{n_i\hat{p}_i + n_C\hat{p}_C}{n_i + n_C}$$

where *F* is -1 for lower-tailed hypotheses and 1 for upper-tailed hypotheses.

Z-Test with Continuity Correction (Unpooled)

This test is the same as the Z-Test (Unpooled), except that a continuity correction is used. Remember that in the null case, the continuity correction makes the results closer to those of Fisher's Exact test.

$$z_t = \frac{\hat{p}_i - \hat{p}_C - \delta_0 - \frac{F}{2}\left(\frac{1}{n_i} + \frac{1}{n_C}\right)}{\hat{\sigma}_2}$$

where

$$\hat{\sigma}_2 = \sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{n_i} + \frac{\hat{p}_C(1-\hat{p}_C)}{n_C}}$$

where *F* is -1 for lower-tailed hypotheses and 1 for upper-tailed hypotheses.

Miettinen and Nurminen's Likelihood Score Test

Miettinen and Nurminen (1985) proposed a test statistic for testing whether the difference is equal to a specified, non-zero, value, δ_0. The regular MLE's, $\hat{p}_i$ and $\hat{p}_C$, are used in the numerator of the score statistic while MLE's $\tilde{p}_i$ and $\tilde{p}_C$, constrained so that $\tilde{p}_i - \tilde{p}_C = \delta_0$, are used in the denominator. A correction factor of *N*/(*N*-1) is applied to make the variance estimate less biased. The significance level of the test statistic is based on the asymptotic normality of the score statistic. The formula for computing this test statistic is

$$z_{MND} = \frac{\hat{p}_i - \hat{p}_C - \delta_0}{\hat{\sigma}_{MND}}$$

where

$$\hat{\sigma}_{MND} = \sqrt{\left(\frac{\tilde{p}_i\tilde{q}_i}{n_i} + \frac{\tilde{p}_C\tilde{q}_C}{n_C}\right)\left(\frac{N}{N-1}\right)}$$

$$\tilde{p}_i = \tilde{p}_C + \delta_0$$

$$\tilde{p}_C = 2B\cos(A) - \frac{L_2}{3L_3}$$

$$A = \frac{1}{3}\left[\pi + \cos^{-1}\left(\frac{C}{B^3}\right)\right]$$

$$B = \text{sign}(C)\sqrt{\frac{L_2^2}{9L_3^2} - \frac{L_1}{3L_3}}$$

$$C = \frac{L_2^3}{27L_3^3} - \frac{L_1L_2}{6L_3^2} + \frac{L_0}{2L_3}$$

$$L_0 = x_{21}\delta_0(1-\delta_0)$$

$$L_1 = [n_C\delta_0 - N - 2x_{21}]\delta_0 + m_1$$

$$L_2 = (N+n_C)\delta_0 - N - m_1$$

$$L_3 = N$$

Farrington and Manning's Likelihood Score Test

Farrington and Manning (1990) proposed a test statistic for testing whether the difference is equal to a specified value δ_0. The regular MLE's, $\hat{p}_i$ and $\hat{p}_C$, are used in the numerator of the score statistic while MLE's $\tilde{p}_i$ and $\tilde{p}_C$, constrained so that $\tilde{p}_i - \tilde{p}_C = \delta_0$, are used in the denominator. The significance level of the test statistic is based on the asymptotic normality of the score statistic.

The formula for computing the test statistic is

$$z_{FMD} = \frac{\hat{p}_i - \hat{p}_C - \delta_0}{\sqrt{\left(\frac{\tilde{p}_i\tilde{q}_i}{n_i} + \frac{\tilde{p}_C\tilde{q}_C}{n_C}\right)}}$$

where the estimates $\tilde{p}_i$ and $\tilde{p}_C$ are computed as in the corresponding test of Miettinen and Nurminen (1985) given above.

Gart and Nam's Likelihood Score Test

Gart and Nam (1990), page 638, proposed a modification to the Farrington and Manning (1988) difference test that corrects for skewness. Let $z_{FMD}(\delta)$ stand for the Farrington and Manning difference test statistic described above. The skewness corrected test statistic, z_{GND}, is the appropriate solution to the quadratic equation

$$(-\tilde{\gamma})z_{GND}^2 + (-1)z_{GND} + (z_{FMD}(\delta) + \tilde{\gamma}) = 0$$

where

$$\tilde{\gamma} = \frac{\tilde{V}^{3/2}(\delta)}{6}\left(\frac{\tilde{p}_i\tilde{q}_i(\tilde{q}_i - \tilde{p}_i)}{n_i^2} - \frac{\tilde{p}_C\tilde{q}_C(\tilde{q}_C - \tilde{p}_C)}{n_C^2}\right)$$

Asymptotic Approximation to Power

A large sample approximation is used to compute power. The large sample approximation is made by replacing the values of $\hat{p}_i$ and $\hat{p}_C$ in the *z* statistic with the corresponding values of P_i and P_C, and then computing the results based on the normal distribution. Note that in large samples, the Farrington and Manning statistic is substituted for the Gart and Nam statistic.

Multiplicity Adjustment

Because *k* z-tests between treatment groups and the control group are run when analyzing the results of this study, many statisticians recommend that the Bonferroni adjustment be applied. This adjustment is easy to apply: the value of alpha that is used in the test is found by dividing the original alpha by the number of tests. For example, if the original alpha is set at 0.05 and the number of treatment (not including the control) groups is five, the individual tests will be conducted using an alpha of 0.01.

The main criticism of this procedure is that if there are many tests, the value of alpha becomes very small. To mitigate against this complaint, some statisticians recommend separating the treatment groups into those that are of primary interest and those that are of secondary interest. The Bonferroni adjustment is made by the using the number of primary treatments rather than the total number of treatments.

There are some who advocate ignoring the adjustment entirely in the case of randomized clinical trials. See for example Saville (1990) and the discussion in chapter 14 of Machin, Campbell, Tan, and Tan (2018).

Size of the Control Group

Because the control group is used over and over, some advocate increasing the number of subjects in this group. The standard adjustment is to include $\sqrt{k}$ subjects in the control group for each subject in one of the treatment groups. See Machin, Campbell, Tan, and Tan (2018, pages 231-232). Note that often, the treatment groups all have the same size.

Example 1 – Finding the Sample Size

A parallel-group, clinical trial is being designed to compare three doses of a test compound against the standard therapy using three non-inferiority tests. Suppose the standard therapy has a response rate of 60%. The investigators would like a sample size large enough to find statistical significance at an overall 0.05 level and an individual-test power of 0.80. The response rates of group 1 are 62%, 65%, or 68%. The response rate of group 2 is 70%. The response rate of group 3 is 75%. The non-inferiority difference is -0.06 (10% of the standard therapy response rate).

Following common practice, the control-group sample-size multiplier will be set to $\sqrt{k} = \sqrt{3} = 1.732$ since there are three treatment groups in this design.

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**
Higher Proportions Are	**Better (H1: δ > δ0)**
Test Type	**Likelihood Score (Gart & Nam)**
Power of Each Test	**0.8**
Overall Alpha	**0.05**
Bonferroni Adjustment	**Standard Bonferroni**
Group Allocation	**Enter Group Allocation Pattern, solve for group sample sizes**
δ0 (Non-Inferiority Difference)	**-0.06**
Control Proportion	**0.6**
Control Sample Size Allocation	**1.732**
Set A Number of Groups	**1**
Set A Proportion	**0.62 0.65 0.68**
Set A Sample Size Allocation	**1**
Set B Number of Groups	**1**
Set B Proportion	**0.7**
Set B Sample Size Allocation	**1**
Set C Number of Groups	**1**
Set C Proportion	**0.75**
Set C Sample Size Allocation	**1**
Set D Number of Groups	**0**
More	**Unchecked**

Output

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

Numeric Reports

Numeric Results

Solve For:	Sample Size
Group Allocation:	Enter Group Allocation Pattern, solve for group sample sizes
Test Type:	Gart & Nam Likelihood Score Test
Higher Proportions Are:	Better
Hypotheses:	H0: δ ≤ δ0 vs. H1: δ > δ0
Number of Groups:	4
Bonferroni Adjustment:	Standard Bonferroni (Divisor = 3)

	Power		Sample Size		Proportion		Difference		Alpha	
Comparison	Target	Actual	Ni	Allocation	Pi|H0 Pi.0	Pi|H1 Pi.1	Non-Inferiority δ0	Actual δi	Overall	Bonferroni-Adjusted
Control			904	1.732	0.60	0.60				
vs A	0.8	0.80039	522	1.000	0.54	0.62	-0.06	0.02	0.05	0.016667
vs B	0.8	0.99997	522	1.000	0.54	0.70	-0.06	0.10	0.05	0.016667
vs C	0.8	1.00000	522	1.000	0.54	0.75	-0.06	0.15	0.05	0.016667
Total			2470							
Control			473	1.732	0.60	0.60				
vs A	0.8	0.80083	273	1.000	0.54	0.65	-0.06	0.05	0.05	0.016667
vs B	0.8	0.98877	273	1.000	0.54	0.70	-0.06	0.10	0.05	0.016667
vs C	0.8	0.99994	273	1.000	0.54	0.75	-0.06	0.15	0.05	0.016667
Total			1292							
Control			288	1.732	0.60	0.60				
vs A	0.8	0.80069	166	1.000	0.54	0.68	-0.06	0.08	0.05	0.016667
vs B	0.8	0.90272	166	1.000	0.54	0.70	-0.06	0.10	0.05	0.016667
vs C	0.8	0.99363	166	1.000	0.54	0.75	-0.06	0.15	0.05	0.016667
Total			786							

Comparison	The group that is involved in the comparison between the treatment and control displayed on this report line. The comparison is made using the difference.
Target Power	The power desired. Power is probability of rejecting a false null hypothesis for this comparison. This power is of the comparison shown on this line only.
Actual Power	The power actually achieved.
Ni	The number of subjects in the ith group. The total sample size shown below the groups is equal to the sum of all individual group sample sizes.
Allocation	The group sample size allocation ratio of the ith group. The value on each row represents the relative number of subjects assigned to the group.
Pi.0	The response proportion in the ith group assumed by the null hypothesis, H0. Note that Pi.0 = Pc + δ0, where Pc is the control group proportion.
Pi.1	The response proportion in the ith group at which the power is calculated.
δ0	The non-inferiority difference in proportions is the boundary that separates a non-inferior result from an inferior result.
δi	The difference between the ith group proportion (Pi.1) and the control group proportion (Pc) at which the power is calculated. The formula is δi = Pi.1 - Pc.
Overall Alpha	The probability of rejecting at least one of the comparisons in this experiment when each null hypothesis is true.
Bonferroni Alpha	The adjusted significance level at which each individual comparison is made.

Summary Statements

A parallel, 4-group design (with one control group and 3 treatment groups) will be used to test whether the proportion for each treatment group is non-inferior to the control group proportion, with a non-inferiority difference of -0.06 (H0: $\delta \leq -0.06$ versus H1: $\delta > -0.06$, $\delta = Pi - Pc$). In this study, higher proportions are considered to be better. The non-inferiority hypotheses will be evaluated using 3 one-sided, two-sample, Bonferroni-adjusted Gart & Nam Likelihood Score tests, with an overall (experiment-wise) Type I error rate (α) of 0.05. The control group proportion is assumed to be 0.6. To detect the treatment proportions 0.62, 0.7, and 0.75 with at least 80% power for each test, the control group sample size needed will be 904 and the number of needed subjects for the treatment groups will be 522, 522, and 522 (totaling 2470 subjects overall).

Dropout-Inflated Sample Size

Group	Dropout Rate	Sample Size Ni	Dropout-Inflated Enrollment Sample Size Ni'	Expected Number of Dropouts Di
1	20%	904	1130	226
2	20%	522	653	131
3	20%	522	653	131
4	20%	522	653	131
Total		2470	3089	619
1	20%	473	592	119
2	20%	273	342	69
3	20%	273	342	69
4	20%	273	342	69
Total		1292	1618	326
1	20%	288	360	72
2	20%	166	208	42
3	20%	166	208	42
4	20%	166	208	42
Total		786	984	198

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

Dropout Summary Statements

Anticipating a 20% dropout rate, group sizes of 1130, 653, 653, and 653 subjects should be enrolled to obtain final group sample sizes of 904, 522, 522, and 522 subjects.

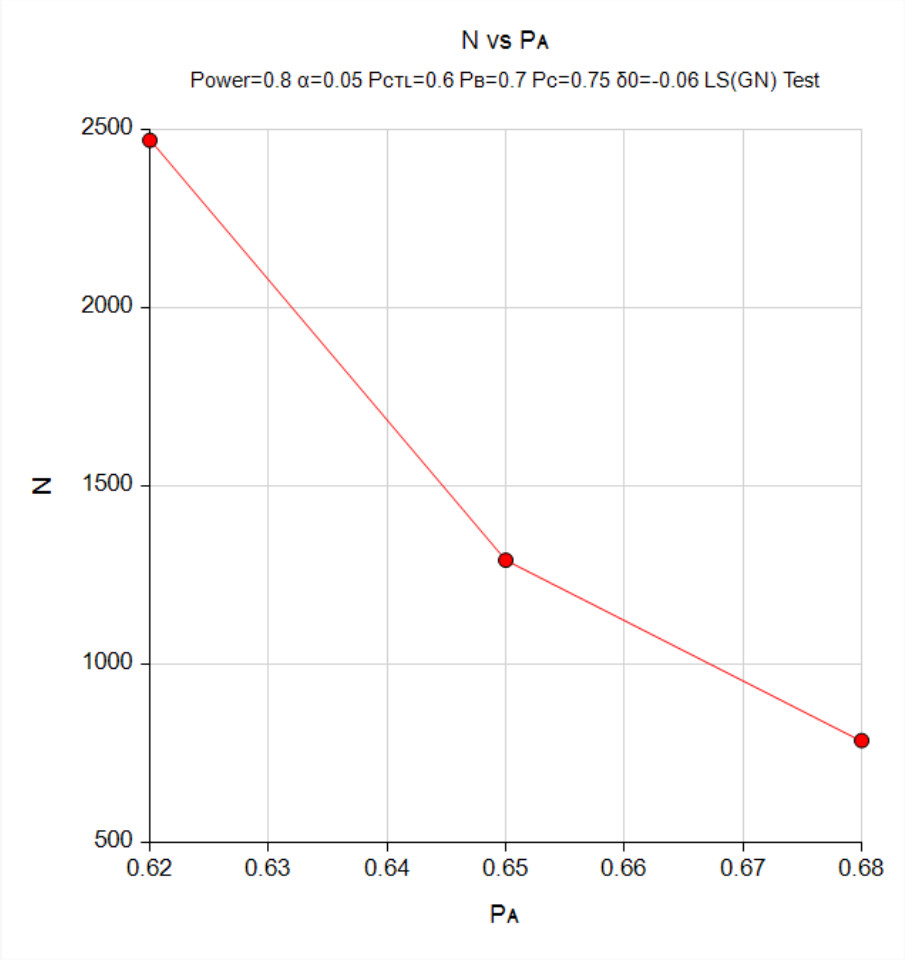
References

Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. 2018. Sample Size Calculations in Clinical Research, 3rd Edition. Chapman & Hall/CRC. Boca Raton, FL. Pages 86-88.
D'Agostino, R.B., Chase, W., and Belanger, A. 1988. 'The Appropriateness of Some Common Procedures for Testing the Equality of Two Independent Binomial Populations', The American Statistician, August 1988, Volume 42 Number 3, pages 198-202.
Farrington, C. P. and Manning, G. 1990. 'Test Statistics and Sample Size Formulae for Comparative Binomial Trials with Null Hypothesis of Non-Zero Risk Difference or Non-Unity Relative Risk.' Statistics in Medicine, Vol. 9, pages 1447-1454.
Fleiss, J. L., Levin, B., Paik, M.C. 2003. Statistical Methods for Rates and Proportions. Third Edition. John Wiley & Sons. New York.
Gart, John J. and Nam, Jun-mo. 1988. 'Approximate Interval Estimation of the Ratio in Binomial Parameters: A Review and Corrections for Skewness.' Biometrics, Volume 44, Issue 2, 323-338.
Gart, John J. and Nam, Jun-mo. 1990. 'Approximate Interval Estimation of the Difference in Binomial Parameters: Correction for Skewness and Extension to Multiple Tables.' Biometrics, Volume 46, Issue 3, 637-643.
Julious, S. A. and Campbell, M. J. 2012. 'Tutorial in biostatistics: sample sizes for parallel group clinical trials with binary data.' Statistics in Medicine, 31:2904-2936.
Lachin, J.M. 2000. Biostatistical Methods. John Wiley & Sons. New York.
Machin, D., Campbell, M.J., Tan, S.B, and Tan, S.H. 2018. Sample Sizes for Clinical, Laboratory, and Epidemiology Studies, 4th Edition. Wiley Blackwell.
Miettinen, O.S. and Nurminen, M. 1985. 'Comparative analysis of two rates.' Statistics in Medicine 4: 213-226.

This report shows the numeric results of this power study. Notice that the results are shown in blocks of three rows at a time. Each block represents a single design.

Plots Section

Plots



This plot gives a visual presentation of the results in the Numeric Report. We can quickly see the impact on the sample size of decreasing the difference between the treatment and control proportions.

Example 2 – Validation using a Previously Validated Procedure

We could not find a validation result in the statistical literature, so we will use a previously validated **PASS** procedure (**Non-Inferiority Tests for the Difference Between Two Proportions**) to produce the results for the following example.

Suppose a parallel-group, clinical trial is being designed to compare three doses of a test compound against the standard therapy using three non-inferiority tests. Suppose the standard therapy has a response rate of 60%. The investigators would like a sample size large enough to find statistical significance at an overall 0.05 level and an individual-test power of 0.80. The response rates of groups 1, 2, and 3 are 65%, 70%, and 75%, respectively. The non-inferiority difference is -0.06 (10% of the standard therapy response rate). The analysis will use the continuity-corrected unpooled Z-Test.

Following common practice, the control-group sample-size multiplier will be set to $\sqrt{k} = \sqrt{3} = 1.732$ since there are three treatment groups in this design.

The **Non-Inferiority Tests for the Difference Between Two Proportions** procedure is set up as follows.

Design Tab

Solve For	**Sample Size**
Power Calculation Method	**Normal Approximation**
Higher Proportions Are	**Better (H1: P1 - P2 > δ0)**
Test Type	**Z-Test C.C. (UnPooled)**
Power	**0.80**
Alpha	**0.016667** (which is Alpha / k)
Group Allocation	**Enter R = N2/N1, solve for N1 and N2**
R	**1.732** (which is Control SS Multiplier / Treatment SS Multiplier)
Input Type	**Mixture (δ0 and P1.1)**
δ0 (Non-Inferiority Difference)	**-0.06**
P1.1 (Actual Proportion)	**0.65 0.7 0.75**
P2 (Group 2 Proportion)	**0.6**

This set of options generates the following report.

Numeric Results

Solve For: Sample Size
Groups: 1 = Treatment, 2 = Reference
Test Statistic: Continuity Corrected Z-Test with Unpooled Variance
Hypotheses: H0: P1 - P2 ≤ δ0 vs. H1: P1 - P2 > δ0

Power		Sample Size			R (Allocation Ratio)		Proportions			Difference		
Target	Actual*	N1	N2	N	Target	Actual	Non-Inferiority P1.0	Actual P1.1	Reference P2	Non-Inferiority δ0	Actual δ1	Alpha
0.8	0.800168	281	487	768	1.732	1.7331	0.54	0.65	0.6	-0.06	0.05	0.0167
0.8	0.800742	130	225	355	1.732	1.7308	0.54	0.70	0.6	-0.06	0.10	0.0167
0.8	0.802355	73	126	199	1.732	1.7260	0.54	0.75	0.6	-0.06	0.15	0.0167

* Power was computed using the normal approximation method.

In order to maintain a power of 80% for all three groups, it is apparent that the treatment groups will all need to have a sample size of 281 and the control group should be 487. We then calculate the powers of the three groups using these sample sizes. The results are displayed in the following table.

Numeric Results

Solve For: Power
Groups: 1 = Treatment, 2 = Reference
Test Statistic: Continuity Corrected Z-Test with Unpooled Variance
Hypotheses: H0: P1 - P2 ≤ δ0 vs. H1: P1 - P2 > δ0

	Sample Size			Proportions			Difference		
Power*	N1	N2	N	Non-Inferiority P1.0	Actual P1.1	Reference P2	Non-Inferiority δ0	Actual δ1	Alpha
0.800168	281	487	768	0.54	0.65	0.6	-0.06	0.05	0.0167
0.990247	281	487	768	0.54	0.70	0.6	-0.06	0.10	0.0167
0.999962	281	487	768	0.54	0.75	0.6	-0.06	0.15	0.0167

* Power was computed using the normal approximation method.

This table contains the validation values. We will now run these values through the current procedure and compare the results with these values.

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**
Higher Proportions Are	**Better (H1: δ > δ0)**
Test Type	**Z-Test C.C. (UnPooled)**
Power of Each Test	**0.8**
Overall Alpha	**0.05**
Bonferroni Adjustment	**Standard Bonferroni**
Group Allocation	**Enter Group Allocation Pattern, solve for group sample sizes**
δ0 (Non-Inferiority Difference)	**-0.06**
Control Proportion	**0.6**
Control Sample Size Allocation	**1.732**
Set A Number of Groups	**1**
Set A Proportion	**0.65**
Set A Sample Size Allocation	**1**
Set B Number of Groups	**1**
Set B Proportion	**0.7**
Set B Sample Size Allocation	**1**
Set C Number of Groups	**1**
Set C Proportion	**0.75**
Set C Sample Size Allocation	**1**
Set D Number of Groups	**0**
More	**Unchecked**

Output

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

Numeric Results

Solve For:	Sample Size
Group Allocation:	Enter Group Allocation Pattern, solve for group sample sizes
Test Type:	Continuity Corrected Z-Test with Unpooled Variance
Higher Proportions Are:	Better
Hypotheses:	H0: δ ≤ δ0 vs. H1: δ > δ0
Number of Groups:	4
Bonferroni Adjustment:	Standard Bonferroni (Divisor = 3)

	Power		Sample Size		Proportion		Difference		Alpha	
Comparison	Target	Actual	Ni	Allocation	Pi|H0 Pi.0	Pi|H1 Pi.1	Non-Inferiority δ0	Actual δi	Overall	Bonferroni-Adjusted
Control			487	1.732	0.60	0.60				
vs A	0.8	0.800166	281	1.000	0.54	0.65	-0.06	0.05	0.05	0.016667
vs B	0.8	0.990247	281	1.000	0.54	0.70	-0.06	0.10	0.05	0.016667
vs C	0.8	0.999962	281	1.000	0.54	0.75	-0.06	0.15	0.05	0.016667
Total			1330							

As you can see, the sample sizes and powers match, thus validating this procedure.