

## Chapter 256

# Multiple Comparisons of Proportions for Treatments vs. a Control

---

### Introduction

This module computes power and sample size for multiple comparisons of treatment proportions versus a control proportion based on the results in Chow, Shao, and Wang (2008). In this design, there are  $k$  treatment groups and one control group. 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 simple z-test of the difference between two proportions. The type I error rate is controlled using a Bonferroni adjustment.

### Background

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, placebo, and treatment. Two tests are run: placebo versus control and treatment versus control. Of course, you could design an experiment with four groups by having a separate control group for treatment and for placebo. But it appears to be more efficient to use just one control group. Besides efficiency, it may be easier to recruit subjects if their chances of receiving the new treatment are better than 50-50.

---

### Technical Details

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

The hypotheses for two-sided tests are

$$H_{0i}: P_i = P_C \text{ versus } H_{a}: P_i \neq P_C$$

For one-sided tests the hypotheses are

$$H_{0i}: P_i \leq P_C \text{ versus } H_{a}: P_i > P_C$$

## Multiple Comparisons of Proportions for Treatments vs. a Control

### Test Statistic

Chow, Shao, and Wang (2008) give the test procedure for each of the  $k$  two-sided comparisons as rejecting the null hypothesis at the  $\alpha$  level of significance if

$$\left| \frac{\hat{p}_i - \hat{p}_c}{\sqrt{\frac{\hat{p}_i(1 - \hat{p}_i)}{N_i} + \frac{\hat{p}_c(1 - \hat{p}_c)}{N_c}}} \right| > |z_{\alpha/(2k)}| \quad i = 1, 2, \dots, k$$

where  $z$  is the standard normal deviate. Dividing  $\alpha$  by 2 in the above expression occurs because there are two-sided tests. Dividing  $\alpha$  by  $k$  in the above expression is used to make the Bonferroni adjustment so that the overall significance level is  $\alpha$ .

### Power

The power of each test is

$$\text{Power} = \Phi \left( \frac{P_i - P_c}{\sqrt{\frac{P_i(1 - P_i)}{N_i} + \frac{P_c(1 - P_c)}{N_c}}} \right)$$

where  $\Phi(x)$  is the standard normal CDF.

This formula can be rearranged to provide formulas for the sample size as well.

---

## Procedure Options

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

---

### Design Tab

The Design tab contains most of the parameters and options that you will be concerned with.

---

#### Solve For

##### Solve For

This option specifies the parameter to be solved for from the other parameters. The parameters that may be selected are *Power* or *Sample Size*.

---

#### Comparisons

##### Alternative Hypothesis

Specify whether a one-sided or two-sided test will be used.

## Multiple Comparisons of Proportions for Treatments vs. a Control

---

### Power and Alpha

#### Power of Each Test

Enter the target power to be used for each comparison test. All comparisons use this same power. Power is the probability of rejecting a false null hypothesis.

The valid range is 0 to 1. Usually, a power between 0.8 and 0.99 is desired. Different disciplines have different standards for setting power. Common values are 0.80 and 0.90.

You can enter a single value such as 0.90 or a series of values such as 0.70 0.80 0.90 or 0.70 to 0.90 by 0.1.

When a series of values is entered, PASS will generate a separate calculation result for each value of the series.

#### Overall Alpha

Overall Alpha is the probability of rejecting at least one of the comparison hypothesis tests when the actual proportion differences are all zero. Thus it is a combined alpha for the whole set of tests. The individual comparison alphas are found using the Bonferroni adjustment.

Since Alpha is a probability, it is bounded by 0 and 1. Commonly, it is between 0.001 and 0.10. Alpha is often set to 0.05.

You can enter a single value such as 0.05 or a series of values such as 0.05 0.10 0.15 or 0.05 to 0.15 by 0.01.

When a series of values is entered, PASS will generate a separate calculation result for each value of the series.

---

### Sample Size (if solving for power)

#### Group Allocation

Select the method used to allocate subjects to groups. The choices are

- **Equal**  
All group sample sizes are the same. The common value of  $N_i$  is entered in the  $N_i$  (Group Sample Size) box that appears immediately below this box.
- **Enter base group sample size and multipliers**  
The group sample sizes are found by multiplying the corresponding Sample Size Multiplier value by the Base Group Sample Size that appears immediately below this box.
- **Enter N (total sample size) and group percentages of N**  
The group sample sizes are found by taking the corresponding percentage of the N (Total Sample Size) that will appear immediately below this box.  
Note that the Percent of N values are adjusted so they sum to 100%.
- **Enter individual group sample sizes**  
Enter each group's sample size directly in the corresponding Sample Size box below.

#### $N_i$ (Group Sample Size)

This is the group sample size of all groups. One or more integer values, separated by blanks or commas, may be entered. A separate analysis is performed for each value listed here.

You can enter integers greater than 5.

## Multiple Comparisons of Proportions for Treatments vs. a Control

### Base Group Sample Size

This is the base sample size per group. One or more positive numbers, separated by blanks or commas, may be entered. A separate analysis is performed for each value listed here.

The individual group samples sizes are determined by multiplying this value times the corresponding *Sample Size Multiplier* value entered below in the *Effect Size and Design* section.

If the individual group *Sample Size Multiplier* numbers are represented by  $m_1, m_2, m_3, \dots$  and this value is represented by  $N_i$ , the group sample sizes  $N_1, N_2, N_3, \dots$  are calculated as

$$N_1 = [N_i(m_1)]$$

$$N_2 = [N_i(m_2)]$$

$$N_3 = [N_i(m_3)]$$

where the operator,  $[X]$  means the next integer after  $X$ , e.g.  $[3.1]=4$ .

For example, suppose there are three groups and the multipliers are set to **1, 2, and 3**. If  $N_i$  is **5**, the resulting group sample sizes will be **5, 10, and 15**.

### N (Total Sample Size)

This is the total sample size. The smallest value of  $N$  should be large enough to guarantee that all group sample sizes are greater than five. The individual group samples sizes are determined by multiplying this value times the corresponding *Percentage of N* value entered in the *Effect Size and Design* section below.

One or more values, separated by blanks or commas, may be entered. A separate analysis is performed for each value listed here.

If the group percentage values are represented by  $Pct_1, Pct_2, Pct_3, \dots$  and this value is represented by  $N$ , the group sample sizes are calculated as follows:

$$N_1 = [N(Pct_1)/100]$$

$$N_2 = [N(Pct_2)/100]$$

$$N_3 = [N(Pct_3)/100]$$

where the operator,  $[X]$  means the next integer after  $X$ , e.g.  $[3.1]=4$ .

For example, suppose there are three groups and the percentages are set to **25, 25, and 50**. If  $N$  is **36**, the resulting group sample sizes will be **9, 9, and 18**.

---

## Sample Size (if solving for sample size)

### Group Allocation Method

Select the method used to allocate subjects to groups during the search for the optimum sample size. The choices are

- **Equal**  
The group sample sizes are constrained so that all group sample sizes are equal.
- **Equal (Fixed control group size with equal treatment group sizes)**  
The sample size of the control group is fixed by the user and the minimum sample size of the treatment groups is found so that all target powers are met.
- **Unequal (Use sample size multipliers)**  
The minimum total sample size is found that uses the multipliers for each group.

## Multiple Comparisons of Proportions for Treatments vs. a Control

- **Unequal (Fixed control group size with flexible treatment group sizes)**  
The control group sample size is fixed by the user. The sample size for each treatment group is then found such that its target power is just met.
- **Unequal (Find control size, then minimize treatment sizes)**  
First find the minimum control-group sample size that meets all target powers and uses the sample size multipliers. Next, recalculate each treatment-group sample size using the minimum control-group size individually, so that the target power is barely achieved.
- **Exploratory (Control group sizes can vary for each treatment group)**  
Exploratory. Does not provide an actual design configuration. Calculate the required sample size for each control-treatment pair, using the sample size multipliers.

### Control Group Size

Enter a fixed sample size for the control group. One or more positive numbers, separated by blanks or commas, may be entered. A separate analysis is performed for each value listed here.

---

## Effect Size and Design

### Groups

Specify the number of groups given on this line. Usually, you will enter a "1" to specify a single group, or you will enter a "0" to indicate that this line should be ignored. This option lets you easily specify many identical groups with a single set of parameters.

The total number of groups is equal to the sum of these values.

### Control Proportion

Specify one or more values of  $P_{\text{control}}$ , the proportion of successes in the control group.

This parameter is used in the calculation of power as the actual value of this proportion and in the calculation of the difference between the proportions to specify the effect size.

You can enter a single value such as 0.3 or a series of values such as 0.3 0.35 0.39 or 0.3 to 0.39 by 0.01.

Proportions must be between 0 and 1. They cannot take on the values 0 or 1. Also,  $P_{\text{control}}$  cannot be equal to any of the  $P_i$ .

### Treatment Proportion

Specify one or more values of  $P_i$ , the proportion of successes on this treatment group.

This parameter is used in the calculation of power as the actual value of this proportion and in the calculation of the difference between the treatment and control proportions to specify the effect size.

You can enter a single value such as 0.3 or a series of values such as 0.3 0.35 0.39 or 0.3 to 0.39 by 0.01.

Proportions must be between 0 and 1. They cannot take on the values 0 or 1. Also,  $P_i$  and  $P_{\text{control}}$  cannot be equal.

## Multiple Comparisons of Proportions for Treatments vs. a Control

### Sample Size Multiplier

The definitions depends on whether you are solving for power or sample.

- **If solving for Power**

The sample size for this group is equal to the product of this multiplier and the Base Group Sample Size entered above. That is, the sample size for group  $i$  is given by  $N_i = (\text{Base Group Sample Size}) \times (\text{Sample Size Multiplier})$ . Hence, only the product of these two numbers is used.

#### Usage

There are a couple of ways to use these parameters.

First, you might enter as standard amount for the Base value such as '30', and then enter a multiplier such as '1' or '2'.

Second, you might enter the standard amount as the Multiplier and then enter the ratio as the Base amount. For example, you might enter '50' for the Multiplier and '1 1.5 2' as the Base values.

- **If solving for Sample Size**

The sample size for this group is equal to the product of this multiplier and an implicit base group sample size that is found during the search. That is, the sample size for group  $i$  is given by  $N_i = \text{Base} \times \text{Multiplier}$ . Hence, only the product of these two numbers matters.

### Percentage of N

This is the percentage of the total sample size that is allocated to this group. The individual group samples sizes are determined by multiplying this value times the  $N$  (*Total Sample Size*) value entered above and then dividing by 100. If the percentages do not sum to 100, they are rescaled so that they do. For example, the values 20, 20, 40 become 25, 25, 50 after rescaling.

For example, if these percentage values are represented by Pct1, Pct2, Pct3, ..., the group sample sizes are calculated as follows:

$$N_1 = [N(\text{Pct1})/100]$$

$$N_2 = [N(\text{Pct2})/100]$$

$$N_3 = [N(\text{Pct3})/100]$$

where the operator,  $[X]$  means the next integer after  $X$ , e.g.  $[3.1]=4$ .

For example, suppose there are three groups and these percentages are set to 25, 25, and 50. If  $N$  is 36, the resulting group sample sizes will be 9, 9, and 18.

## Multiple Comparisons of Proportions for Treatments vs. a Control

## 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 in patients with a specific type of disease. Suppose the standard therapy has a response rate of 60%. The investigators would like a sample size large enough to find statistical significance at the 0.05 level if the actual response rate is at least 70% and the power is 0.80 in each test. They also want to see the impact on sample size of response rates of 75% and 80%. The tests will be two-sided.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Multiple Comparisons of Proportions for Treatments vs. a Control** 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**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
Alternative Hypothesis .....	<b>Two-Sided</b>
Power of Each Test .....	<b>0.8</b>
Overall Alpha .....	<b>0.05</b>
Group Allocation Method .....	<b>Equal</b>
Control Proportion.....	<b>0.6</b>
Set A Number of Groups .....	<b>3</b>
Set A Proportion .....	<b>0.7 0.75 0.8</b>
Set B Number of Groups .....	<b>0</b>

### Annotated Output

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

#### Numeric Results

Numeric Results							
Hypothesis Type: Two-Sided							
Number of Treatment Groups: 3							
Group Allocation Search Algorithm: All groups have equal sample size							
Comparison	Target Power	Actual Power	Sample Size Ni	Proportion Pi	Proportion Difference Pi - Pcontrol	Overall Alpha	Bonferroni Adjusted Alpha
Control			472	0.6000			
vs A1	0.8000	0.80085	472	0.7000	0.1000	0.0500	0.016667
vs A2	0.8000	0.80085	472	0.7000	0.1000	0.0500	0.016667
vs A3	0.8000	0.80085	472	0.7000	0.1000	0.0500	0.016667
Total			1888				
Control			199	0.6000			
vs A1	0.8000	0.80020	199	0.7500	0.1500	0.0500	0.016667
vs A2	0.8000	0.80020	199	0.7500	0.1500	0.0500	0.016667
vs A3	0.8000	0.80020	199	0.7500	0.1500	0.0500	0.016667
Total			796				
Control			105	0.6000			
vs A1	0.8000	0.80133	105	0.8000	0.2000	0.0500	0.016667
vs A2	0.8000	0.80133	105	0.8000	0.2000	0.0500	0.016667
vs A3	0.8000	0.80133	105	0.8000	0.2000	0.0500	0.016667
Total			420				

## Multiple Comparisons of Proportions for Treatments vs. a Control

### References

Chow, S.C.; Shao, J.; Wang, H. 2008. Sample Size Calculations in Clinical Research, 2nd Edition. Chapman & Hall/CRC. Boca Raton, FL. Pages 99-100.

### Report Definitions

Comparison shows the test that is displayed on this line. Each treatment group is compared to the control group. The null hypothesis,  $H_0$ , is that the treatment group proportion is equal to the control group proportion.

The alternative hypothesis,  $H_a$ , is that the treatment group proportion is different from the control group proportion.

Power is the probability of rejecting a false null hypothesis for this comparison.

Target Power is the power desired.

Actual Power is the power actually achieved.

Sample Size,  $N_i$ , is the number of subjects in the this group. The total sample size,  $N$ , is shown as the last row of the set.

Proportion,  $P_i$ , is the proportion of successes in this group. These proportions are used in calculations involving the alternative hypothesis.

Proportion Difference,  $P_i - P_{\text{control}}$ , is the difference in the group proportions. When the actual proportion difference is greater than this amount, the null hypothesis of equality will be rejected with probability equal to the power.

Overall Alpha is the probability of rejecting at least one of the comparisons when the actual proportion differences are all zero.

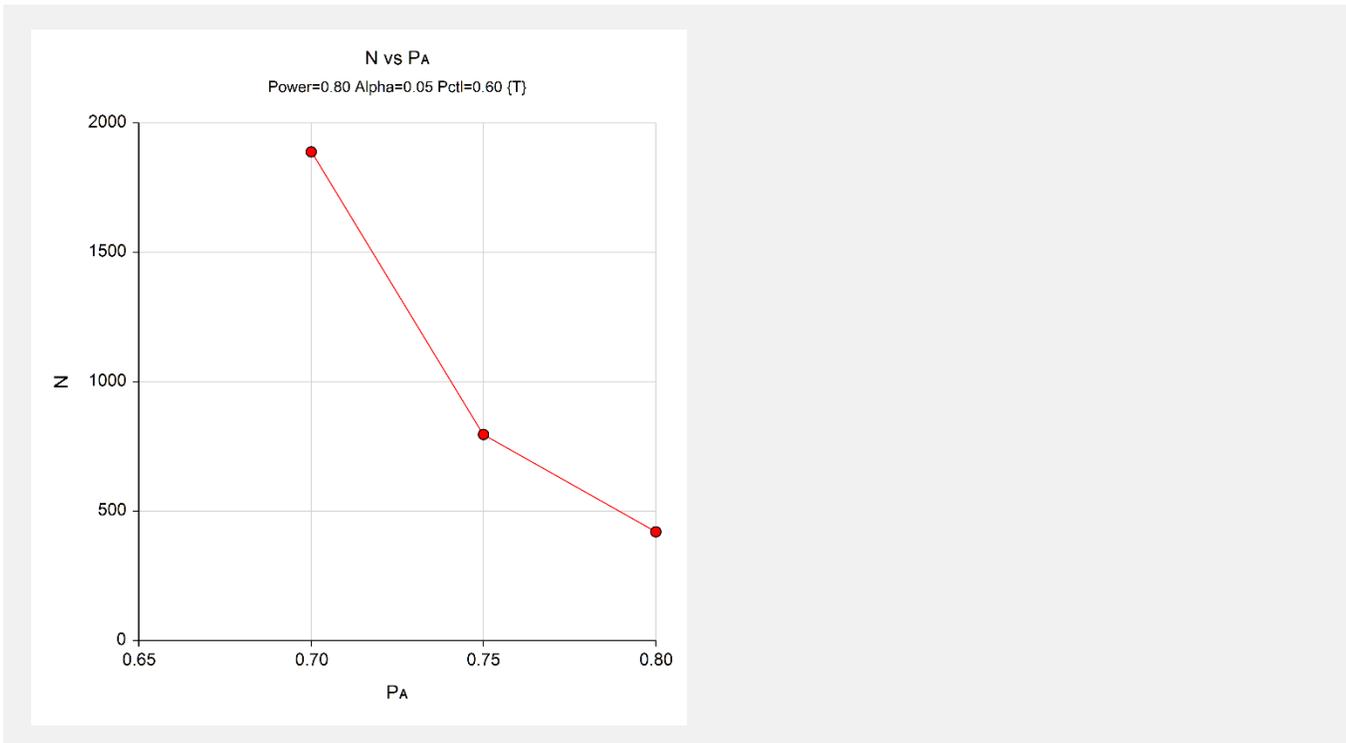
Bonferroni Adjusted Alpha is the significance level at which each individual comparison is made. It is calculated by dividing the overall alpha by the number of treatment groups.

### Summary Statements

A one-way study is planned to test whether each of 3 treatment-group proportions are significantly different from the control-group proportion at the 0.0500 significance level. The sample size of the control group is 472 and the sample sizes of the individual treatment groups are 472, 472, and 472. Hence the total sample size is 1888. The corresponding powers of each of the tests are 0.801, 0.801, and 0.801. These powers are computed assuming that the actual control group proportion is 0.600 and the treatment group proportions are 0.700, 0.700, and 0.700.

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



This plot gives a visual presentation to 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 Chow, Shao, and Wang (2008)

Chow, Shao, and Wang (2008) page 100 gives an example in which the control response rate is 20%, treatment 1 response rate is 40%, and treatment 2 response rate is 50%. They use two-sided tests and set the overall significance level at 0.05. They set the target power at 0.80. They calculate the sample sizes for treatment 1 as about 95 and treatment 2 as about 45. Thus the required sample size is 95.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Multiple Comparisons of Proportions for Treatments vs. a Control** procedure window. You may then make the appropriate entries as listed below, or open **Example 2 (a, b, or c)** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
Alternative Hypothesis .....	<b>Two-Sided</b>
Power of Each Test .....	<b>0.8</b>
Overall Alpha .....	<b>0.05</b>
Group Allocation Method .....	<b>Exploratory (Control group sizes can vary for each treatment group)</b>
Control Proportion.....	<b>0.2</b>
Control Multiplier.....	<b>1</b>
Set A Number of Groups .....	<b>1</b>
Set A Proportion .....	<b>0.4</b>
Set A Multiplier.....	<b>1</b>
Set B Number of Groups .....	<b>1</b>
Set B Proportion .....	<b>0.5</b>
Set B Multiplier.....	<b>1</b>

### Annotated Output

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

#### Numeric Results

<b>Numeric Results</b>									
Hypothesis Type: Two-Sided									
Number of Treatment Groups: 3									
Group Allocation Search Algorithm: EXPLORATORY, sample size of each comparison figured individually									
Comparison	Target Power	Actual Power	Sample Size Ni/Nctrl	Sample Size Multiplier Mi	Proportion Pi	Proportion Difference Pi - Pcontrol	Overall Alpha	Bonferroni Adjusted Alpha	
Control				1.00	0.2000				
vs A	0.8000	0.80427	96/96	1.00	0.4000	0.2000	0.0500	0.025000	
vs B	0.8000	0.80687	44/44	1.00	0.5000	0.3000	0.0500	0.025000	

PASS obtained 96 and 44 which matches Chow, Shao, and Wang's results to within rounding. The final design would be 96 subjects in each of the three groups.

## Multiple Comparisons of Proportions for Treatments vs. a Control

If you change the *Group Allocation Method* to *Unequal (Use sample size multipliers)*, you will obtain the following

**Numeric Results**

Hypothesis Type: Two-Sided

Number of Treatment Groups: 3

Group Allocation Search Algorithm: Unequal group sizes using sample size multipliers

Comparison	Target Power	Actual Power	Sample Size		Proportion Pi	Proportion Difference Pi - Pcontrol	Overall Alpha	Bonferroni Adjusted Alpha
			Size Ni	Multiplier Mi				
Control			96	1.00	0.2000			
vs A	0.8000	0.80427	96	1.00	0.4000	0.2000	0.0500	0.025000
vs B	0.8000	0.99059	96	1.00	0.5000	0.3000	0.0500	0.025000
Total			288					

This algorithm first found the required group size of 96 and then provided the results when 96 is the common group size. Notice that the actual power for the B vs Control comparison is 0.99059 which is much higher than the target power of 0.80.

If you change the *Group Allocation Method* to *Unequal (Find control size, then minimize treatment sizes)*, you will obtain the following

**Numeric Results**

Hypothesis Type: Two-Sided

Number of Treatment Groups: 3

Group Allocation Search Algorithm: Unequal group sizes, first find minimum control-group size using multipliers

Comparison	Target Power	Actual Power	Sample Size		Proportion Pi	Proportion Difference Pi - Pcontrol	Overall Alpha	Bonferroni Adjusted Alpha
			Size Ni	Multiplier Mi				
Control			96	1.00	0.2000			
vs A	0.8000	0.80157	95	1.00	0.4000	0.2000	0.0500	0.025000
vs B	0.8000	0.81033	33	1.00	0.5000	0.3000	0.0500	0.025000
Total			224					

This algorithm first found the required control group size of 96 and then provides the results when the target powers are just met: 95 in treatment A and 33 in treatment B.

Notice that the actual powers are now close to the target powers. Also note that the total sample size has been reduced from 288 to 224, a reduction of 22%.