

Chapter 103

Superiority by a Margin Tests for One Proportion

Introduction

This module provides power analysis and sample size calculation for one-sample proportion tests in which the researcher is testing whether the proportion is different from a hypothesized proportion by some specified amount. Users may choose from among five popular test statistics commonly used for running the hypothesis test.

Approximate sample size formulas for tests of a single proportion are presented in Chow et al. (2008) page 85. However, only large sample (normal approximation) results are given there. This procedure has the capability for computing power using both the normal approximation and binomial enumeration for all tests. Some sample size programs use only the normal approximation to the binomial distribution for power and sample size estimates. The normal approximation is accurate for large sample sizes and for proportions between 0.2 and 0.8, roughly. When the sample sizes are small, or the proportions are extreme (i.e., less than 0.2 or greater than 0.8) the binomial calculations are much more accurate.

Example

A superiority test example will set the stage for the discussion of the terminology that follows. Suppose that the current treatment for a disease is effective 70% of the time. This treatment is inexpensive and exhibits few side-effects. A promising new treatment has been developed to the point where it can be tested, but the treatment is more expensive and may have more side-effects. One of the basic questions that must be answered is whether the new treatment is effective enough to justify its use over the treatment with the known effectiveness of 70%. In other words, do more than 70% of treated subjects respond to the new treatment?

After thoughtful discussion with several clinicians, it was decided that if a response of at least 78% was achieved, the new treatment would be adopted. The difference between these two percentages may be called the margin of superiority. The margin of superiority in this example is 8%.

The developers must design an experiment to test the hypothesis that the response rate of the new treatment is at least 0.78. The statistical hypothesis to be tested is

$$H_0: P \leq 0.78 \text{ versus } H_1: P > 0.78$$

Notice that when the null hypothesis is rejected, the conclusion is that the response rate is at least 0.78. Note that even though the response rate of the current treatment is 0.70, the hypothesis test is about a response rate of 0.78. Also notice that a rejection of the null hypothesis results in the conclusion of interest.

Binomial Model

A binomial variable should exhibit the following four properties:

1. The variable is binary --- it can take on one of two possible values.
2. The variable is observed a known number of times. Each observation or replication is called a Bernoulli trial. The number of replications is n . The number of times that the outcome of interest is observed is r . Thus, r takes on the possible values $0, 1, 2, \dots, n$.
3. The probability, P , that the outcome of interest occurs is constant for each trial.
4. The trials are independent. The outcome of one trial does not influence the outcome of the any other trial.

A binomial probability is calculated using the formula

$$b(r; n, P) = \binom{n}{r} P^r (1 - P)^{n-r}$$

where

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

The Hypergeometric Model

When samples are taken without replacement from a population of known size, N , the *hypergeometric* distribution should be used in place of the binomial distribution. The properties of a variable that is distributed according to the hypergeometric distribution are

1. The variable is binary--it can take on one of two possible values.
2. The variable is observed a known number of times. Each observation or replication is called a Bernoulli trial. The number of replications is n . The number of times that the outcome of interest is observed is r . Thus, r takes on the possible values $0, 1, 2, \dots, n$.
3. The total number of items is N . The proportion of items with the characteristic of interest is P .

The hypergeometric probability of obtaining exactly r of n items with the characteristic of interest is calculated using

$$h(r; N, n, P) = \frac{\binom{NP}{r} \binom{N-NP}{n-r}}{\binom{N}{n}}$$

Note that the quantity NP is rounded to the nearest integer.

Hypothesis Testing

Parameterizations of the Proportions

In the discussion that follows, let P represent the proportion responding as a success. That is, P is the actual probability of a *success* in a binomial experiment. Let PB represent the *baseline* proportion. In a superiority experiment, the baseline proportion is the response rate of the current treatment. The power of a test is computed at a specific value of the proportion. Let $P1$ represent the proportion at which the power is computed.

$P0$ represents the smallest value of P that still results in the conclusion that the new treatment is superior to the current treatment. The statistical hypotheses that are tested are

$$H0: P \leq P0 \quad \text{versus} \quad H1: P > P0$$

or equivalently

$$H0: P - PB \leq d0 \quad \text{versus} \quad H1: P - PB > d0$$

where $d0$ is the equal to $P0 - PB$.

There are three common methods of specifying the margin of superiority. The most direct is to simply assign values for PB and $P0$. However, it is often more meaningful to identify PB and then specify $P0$ implicitly by giving their difference, ratio, or odds ratio. Mathematically, the definitions of these parameterizations are

Parameter	Computation	Hypotheses
Difference	$d0 = P0 - PB$	$H0: P \leq PB + d0 \quad \text{vs} \quad H1: P > PB + d0$
Ratio	$r0 = P0/PB$	$H0: P \leq r0(PB) \quad \text{vs} \quad H1: P > r0(PB)$
Odds Ratio	$o0 = Odds0/OddsB$	$H0: P \leq A \quad \text{vs} \quad H1: P > A$

$$\text{where } A = \frac{(o0)(PB)}{1 - PB + (o0)(PB)}$$

Difference

The difference is perhaps the most direct method of comparison between two proportions. It is easy to interpret and communicate. It gives the absolute impact of the treatment. However, there are subtle difficulties that can arise with its interpretation.

One difficulty arises when the event of interest is rare. If a difference of 0.001 occurs when the baseline probability is 0.40, it would be dismissed as being trivial. That is, there is usually little interest in a treatment that only decreases the probability from 0.400 to 0.399. However, if the baseline probability of a disease is 0.002, a 0.001 decrease would represent a reduction of 50%. Thus, interpretation of the difference depends on the baseline probability of the event. As a rule of thumb, the difference is best suited for those cases in which $0.20 < P < 0.80$.

Superiority using a Difference

The following example is intended to help you understand the concept of *superiority*. Suppose 60% of patients respond to the current treatment method ($PB = 0.60$). If the response rate of the new treatment is at least ten percentage points better ($d0 = 0.10$), it will be considered to be superior to the existing treatment. Substituting these figures into the statistical hypotheses gives $H0: d \leq 0.10$ versus $H1: d > 0.10$. The relationship $P0 = PB + d0$ gives $H0: P \leq 0.70$ versus $H1: P > 0.70$.

In this example, when the null hypothesis is rejected, the concluded alternative is that the response rate is at least 0.70. That is, the conclusion of superiority is that the new treatment's response rate is at least 0.10 more than that of the existing treatment.

Ratio

The ratio $r0 = PS / PB$ gives the relative change in the probability of the response. Testing non-zero null hypotheses use the same formulation $H0: r \leq r0$ versus $H1: r > r0$.

Superiority using a Ratio

The following example is intended to help you understand the concept of *superiority* as it applies to the ratio. Suppose that 60% of patients ($PB = 0.60$) respond to the current treatment method. If a new treatment increases the response rate by at least 10% ($r0 = 1.10$), it will be considered to be superior to the existing treatment. Substituting these figures into the statistical hypotheses gives $H0: r \leq 1.10$ versus $H1: r > 1.10$. The relationship $P0 = (r0)(PB)$ gives $H0: P \leq 0.66$ versus $H1: P > 0.66$.

In this example, when the null hypothesis is rejected, the concluded alternative is that the response rate is at least 66%. That is, the conclusion of superiority is that the new treatment's response rate is at least 10% more than that of the existing treatment.

Odds Ratio

The odds ratio, $o0 = (P0 / (1 - P0)) / (PB / (1 - PB))$, gives the relative change in the odds of the response. Testing superiority hypotheses uses the similar formulation $H0: o \leq o0$ versus $H1: o > o0$.

Test Statistics

Many different test statistics have been proposed for non-inferiority tests of a single proportion. Most of these were proposed before computers or hand calculators were widely available. Although these legacy methods are still presented in textbooks, their power and accuracy should be compared against modern exact methods before they are adopted for serious research. To make this comparison easy, the power and significance of several tests of a single proportion are available in this procedure.

Exact Test

The test statistic is r , the number of successes in n trials. This test should be the standard against which other test statistics are judged. The significance level and power are computed by enumerating the possible values of r , computing the probability of each value, and then computing the corresponding value of the test statistic. Hence the values that are reported in the output for these tests are exact, not approximate.

Superiority by a Margin Tests for One Proportion

Z-Tests

Several z statistics have been proposed that use the central limit theorem. This theorem states that for large sample sizes, the distribution of the z statistic is approximately normal. All of these tests take the following form:

$$z = \frac{p - PB - d0}{s} = \frac{p - P0}{s}$$

Although these z tests were developed because the distribution of z is approximately normal in large samples, the actual significance level and power can be computed exactly using the binomial distribution.

We include four z tests which are based on two methods for computing s and whether a continuity correction is applied.

Z-Test using S(P0)

This test statistic uses the value of P0 to compute s.

$$z_1 = \frac{p - PB - d0}{\sqrt{(PB + d0)(1 - (PB + d0))/n}} = \frac{p - P0}{\sqrt{P0(1 - P0)/n}}$$

Z-Test using S(P0) with Continuity Correction

This test statistic is similar to the one above except that a continuity correction is applied to make the normal distribution more closely approximate the binomial distribution.

$$z_2 = \frac{(p - PB - d0) + c}{\sqrt{(PB + d0)(1 - (PB + d0))/n}} = \frac{(p - P0) + c}{\sqrt{P0(1 - P0)/n}}$$

where

$$c = \begin{cases} \frac{-1}{2n} & \text{if } p > P0 \\ \frac{1}{2n} & \text{if } p < P0 \\ 0 & \text{if } |p - P0| < \frac{1}{2n} \end{cases}$$

Z-Test using S(Phat)

This test statistic uses the value of p to compute s.

$$z_3 = \frac{p - PB - d0}{\sqrt{p(1 - p)/n}} = \frac{p - P0}{\sqrt{p(1 - p)/n}}$$

Superiority by a Margin Tests for One Proportion

Z-Test using S(Phat) with Continuity Correction

This test statistic is similar to the one above except that a continuity correction is applied to make the normal distribution more closely approximate the binomial distribution.

$$z_4 = \frac{(p - PB - d0) + c}{\sqrt{p(1-p)/n}} = \frac{(p - P0) + c}{\sqrt{p(1-p)/n}}$$

where

$$c = \begin{cases} \frac{-1}{2n} & \text{if } p > P0 \\ \frac{1}{2n} & \text{if } p < P0 \\ 0 & \text{if } |p - P0| < \frac{1}{2n} \end{cases}$$

Power Calculation

Normal Approximation Method

Power may be calculated for one-sample proportions superiority tests using the normal approximation to the binomial distribution. This section provides the power calculation formulas for the various test statistics available in this procedure. In each case, power is presented for the lower and upper one-sided hypothesis tests. In the equations that follow, $\Phi()$ represents the standard normal cumulative distribution function, and z_α represents the value that leaves α in the upper tail of the standard normal distribution. All power values are evaluated at $P = P1$.

Exact Test

Power for the higher proportions worse hypothesis test of $H0: P \geq P0$ vs. $H1: P < P0$ is

$$Power_{ET, \text{ Lower One-Sided}} = \Phi \left(\frac{\sqrt{n}(P0 - P1) - z_\alpha \sqrt{P0(1 - P0)FPC}}{\sqrt{P1(1 - P1)FPC}} \right)$$

where $FPC = (N - n)/(N - 1)$ if the population size, N , is finite, and $FPC = 1$ if the population size is infinite.

Power for the higher proportions better hypothesis test of $H0: P \leq P0$ vs. $H1: P > P0$ is

$$Power_{ET, \text{ Upper One-Sided}} = 1 - \Phi \left(\frac{\sqrt{n}(P0 - P1) + z_\alpha \sqrt{P0(1 - P0)FPC}}{\sqrt{P1(1 - P1)FPC}} \right)$$

Superiority by a Margin Tests for One Proportion

Z Test using S(P0)

Power for the higher proportions worse hypothesis test of $H_0: P \geq P_0$ vs. $H_1: P < P_0$ is

$$Power_{ZS(P_0), \text{ Lower One-Sided}} = \Phi \left(\frac{\sqrt{n}(P_0 - P_1) - z_\alpha \sqrt{P_0(1 - P_0)}}{\sqrt{P_1(1 - P_1)}} \right)$$

Power for the higher proportions better hypothesis test of $H_0: P \leq P_0$ vs. $H_1: P > P_0$ is

$$Power_{ZS(P_0), \text{ Upper One-Sided}} = 1 - \Phi \left(\frac{\sqrt{n}(P_0 - P_1) + z_\alpha \sqrt{P_0(1 - P_0)}}{\sqrt{P_1(1 - P_1)}} \right)$$

Z Test using S(P0) with Continuity Correction

Power for the higher proportions worse hypothesis test of $H_0: P \geq P_0$ vs. $H_1: P < P_0$ is

$$Power_{ZS(P_0)CC, \text{ Lower One-Sided}} = \Phi \left(\frac{\sqrt{n}(P_0 - P_1) - z_\alpha \sqrt{P_0(1 - P_0)} - c}{\sqrt{P_1(1 - P_1)}} \right)$$

Power for the higher proportions better hypothesis test of $H_0: P \leq P_0$ vs. $H_1: P > P_0$ is

$$Power_{ZS(P_0)CC, \text{ Upper One-Sided}} = 1 - \Phi \left(\frac{\sqrt{n}(P_0 - P_1) + z_\alpha \sqrt{P_0(1 - P_0)} + c}{\sqrt{P_1(1 - P_1)}} \right)$$

where $c = 1/2\sqrt{n}$ if $|P_1 - P_0| < 1/2n$ otherwise $c = 0$.

Z Test using S(Phat)

Power for the higher proportions worse hypothesis test of $H_0: P \geq P_0$ vs. $H_1: P < P_0$ is

$$Power_{ZS(P_1), \text{ Lower One-Sided}} = \Phi \left(\frac{\sqrt{n}(P_0 - P_1) - z_\alpha \sqrt{P_1(1 - P_1)}}{\sqrt{P_1(1 - P_1)}} \right)$$

Power for the higher proportions better hypothesis test of $H_0: P \leq P_0$ vs. $H_1: P > P_0$ is

$$Power_{ZS(P_1), \text{ Upper One-Sided}} = 1 - \Phi \left(\frac{\sqrt{n}(P_0 - P_1) + z_\alpha \sqrt{P_1(1 - P_1)}}{\sqrt{P_1(1 - P_1)}} \right)$$

Z Test using S(Phat) with Continuity Correction

Power for the higher proportions worse hypothesis test of $H_0: P \geq P_0$ vs. $H_1: P < P_0$ is

$$Power_{ZS(P_1)CC, \text{ Lower One-Sided}} = \Phi \left(\frac{\sqrt{n}(P_0 - P_1) - z_\alpha \sqrt{P_1(1 - P_1)} - c}{\sqrt{P_1(1 - P_1)}} \right)$$

Power for the higher proportions better hypothesis test of $H_0: P \leq P_0$ vs. $H_1: P > P_0$ is

$$Power_{ZS(P_1)CC, \text{ Upper One-Sided}} = 1 - \Phi \left(\frac{\sqrt{n}(P_0 - P_1) + z_\alpha \sqrt{P_1(1 - P_1)} + c}{\sqrt{P_1(1 - P_1)}} \right)$$

where $c = 1/2\sqrt{n}$ if $|P_1 - P_0| < 1/2n$ otherwise $c = 0$.

Steps to Calculate Power using Binomial Enumeration of All Possible Outcomes

Historically, power and sample size calculations for a one-sample proportion test have been based on normal approximations to the binomial. However, with the speed of modern computers using the normal approximation is unnecessary, especially for small samples. Rather, the significance level and power can be computed using complete enumeration of all possible values of x , the number of successes in a sample of size n .

This is done as follows.

1. The critical value of the test is computed using standard techniques.
2. For each possible value of x , the value of the test statistic (z-test or exact test) is computed along with its associated probability of occurrence.
3. The significance level and power are computed by summing the probabilities of occurrence for all values of the test statistic that are greater than (or less than) the critical value. Each probability of occurrence is calculated using P_0 for the significance level and P_1 for the power.

Other variables such as the sample size are then found using an efficient search algorithm. Although this method is not as elegant as a closed-form solution, it is completely accurate.

Example 1 – Finding the Power

Suppose 50% of patients with a certain type of cancer survive two years using the current treatment. The current treatment is inexpensive and has few severe side effects. A new treatment has more side effects and is more expensive. A superiority trial is to be conducted to show that the two-year survival rate of the new treatment is better than the current treatment. After serious consideration, the margin of superiority is set at 5%. What power will be achieved by sample sizes of 50, 100, 200, 300, 500, or 800 and a significance level of 0.05? For comparative purposes, also calculate the power for a margin of superiority of 10%. Assume that the true survival rate of the new treatment is 62%.

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
Power Calculation Method	Normal Approximation
Higher Proportions Are	Better
Test Type	Z-test using S(P0)
Alpha	0.05
n (Sample Size)	50 100 200 300 500 800
Input Type	Differences
PB (Baseline Proportion)	0.50
d0 (Superiority Difference)	0.05 0.10
d1 (Actual Difference)	0.12

Superiority by a Margin Tests for One Proportion

Output

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

Numeric Reports

Numeric Results

Solve For: Power
 Higher Proportions Are: Better
 Alternative Hypothesis: One-Sided (H0: $P \leq P_0$ vs. H1: $P > P_0$)
 Test Type: Z-Test with S(P0)

Power*	Sample Size n	Baseline Proportion PB	Difference		Alpha	Reject H0 If Z >
			Superiority d0	Actual d1		
0.25266	50	0.5	0.05	0.12	0.05	1.6449
0.08553	50	0.5	0.10	0.12	0.05	1.6449
0.40372	100	0.5	0.05	0.12	0.05	1.6449
0.10600	100	0.5	0.10	0.12	0.05	1.6449
0.63819	200	0.5	0.05	0.12	0.05	1.6449
0.14065	200	0.5	0.10	0.12	0.05	1.6449
0.79160	300	0.5	0.05	0.12	0.05	1.6449
0.17196	300	0.5	0.10	0.12	0.05	1.6449
0.93808	500	0.5	0.05	0.12	0.05	1.6449
0.23002	500	0.5	0.10	0.12	0.05	1.6449
0.99165	800	0.5	0.05	0.12	0.05	1.6449
0.31040	800	0.5	0.10	0.12	0.05	1.6449

* Power was computed using the normal approximation method.

Power The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
 n The size of the sample drawn from the population.
 PB The baseline or standard value of the proportion.
 d0 The distance above PB that is considered superior. $d_0 = P_0 - PB$.
 d1 The value of the difference at which the power is calculated. $d_1 = P_1 - PB$.
 Alpha The probability of rejecting a true null hypothesis.
 Reject H0 If... Gives the critical value(s) for the test.

Summary Statements

A single-group design will be used to test whether the proportion is superior by a margin, with a superiority difference ($P_0 - PB$) of 0.05 (H0: $P \leq P_0$ versus H1: $P > P_0$). The comparison will be made using a one-sided, one-sample Z-test that uses S(P0) to estimate the standard deviation, with a Type I error rate (α) of 0.05. The baseline proportion (PB) is assumed to be 0.5. To detect a difference ($P_1 - PB$) of 0.12 with a sample size of 50, the power is 0.25266.

Superiority by a Margin Tests for One Proportion

Dropout-Inflated Sample Size

Dropout Rate	Sample Size n	Dropout- Inflated Enrollment Sample Size n'	Expected Number of Dropouts D
20%	50	63	13
20%	100	125	25
20%	200	250	50
20%	300	375	75
20%	500	625	125
20%	800	1000	200

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, 63 subjects should be enrolled to obtain a final sample size of 50 subjects.

References

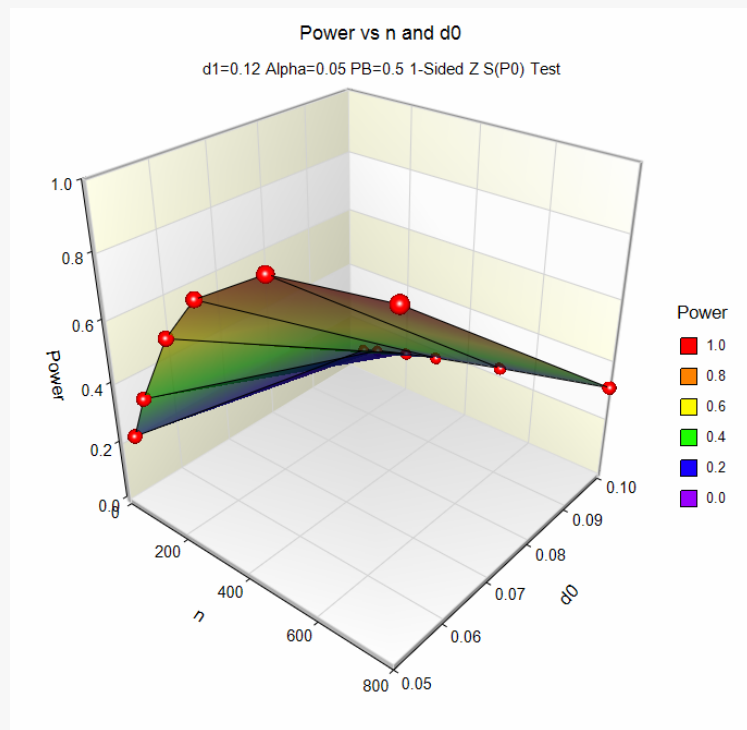
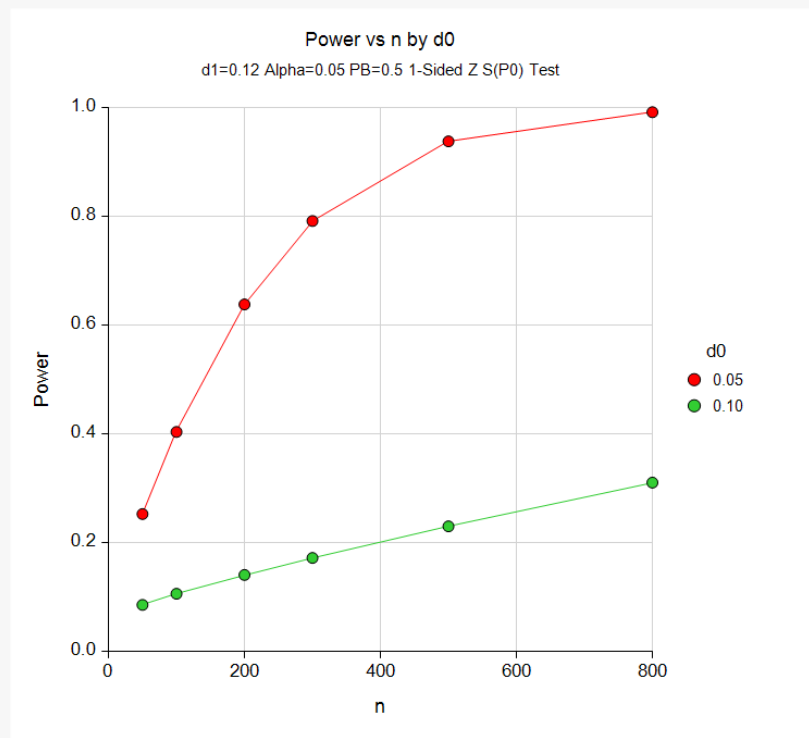
- Blackwelder, W.C. 1998. 'Equivalence Trials.' In Encyclopedia of Biostatistics, John Wiley and Sons. New York. Volume 2, 1367-1372.
- Chow, S. C., Shao, J., and Wang, H. 2008. Sample Size Calculations in Clinical Research, Second Edition. Chapman & Hall/CRC. Boca Raton, Florida.
- Fleiss, J. L., Levin, B., and Paik, M.C. 2003. Statistical Methods for Rates and Proportions. Third Edition. John Wiley & Sons. New York.

This report shows the values of each of the parameters, one scenario per row. The values from this table are plotted in the chart below.

Superiority by a Margin Tests for One Proportion

Plots Section

Plots



These plots show the relationship between power, sample size, and the minimum difference.

Example 2 – Finding the Sample Size

Continuing from Example 1, suppose you want to find the exact sample size necessary to achieve 90% power when the superiority difference is 0.05.

In this example we'll show you how to setup the calculation by inputting proportions, differences, ratios, and odds ratios. In all cases, you'll see that the sample sizes are exactly the same. The only difference is in the way the non-inferiority margin and effect size are specified.

Setup (Proportions)

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 2a** 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 Calculation Method	Normal Approximation
Higher Proportions Are	Better
Test Type	Z-test using S(P0)
Power	0.90
Alpha	0.05
Input Type	Proportions
P0 (Superiority Proportion)	0.45
P1 (Actual Proportion)	0.62

Output (Proportions)

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

Numeric Results					
Solve For:	Sample Size				
Higher Proportions Are:	Better				
Alternative Hypothesis:	One-Sided (H0: P ≤ P0 vs. H1: P > P0)				
Test Type:	Z-Test with S(P0)				
Power*	Sample Size n	Proportion		Alpha	Reject H0 If Z >
		Superiority P0	Actual P1		
0.90037	424	0.55	0.62	0.05	1.6449

* Power was computed using the normal approximation method.

This report shows that a sample size of 424 will be necessary to achieve the design requirements.

Setup (Differences)

To input $P_0 = 0.55$ and $P_1 = 0.62$ using differences, recall the relationships between P_0 , P_1 , PB , d_0 and d_1 :

$$d_0 = P_0 - PB = 0.55 - 0.5 = 0.05$$

$$d_1 = P_1 - PB = 0.62 - 0.5 = 0.12$$

The setup for differences is exactly the same as that for proportions except for the following inputs on the Design tab. You may then make the appropriate entries as listed below, or open **Example 2b** by going to the **File** menu and choosing **Open Example Template**.

Design Tab	
Input Type.....	Differences
PB (Baseline Proportion)	0.50
d0 (Superiority Difference)	0.05
d1 (Actual Difference)	0.12

Output (Differences)

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

Numeric Results						
Solve For:	Sample Size					
Higher Proportions Are:	Better					
Alternative Hypothesis:	One-Sided (H0: $P \leq P_0$ vs. H1: $P > P_0$)					
Test Type:	Z-Test with S(P0)					
Power*	Sample Size n	Baseline Proportion PB	Difference		Alpha	Reject H0 If Z >
			Superiority d0	Actual d1		
0.90037	424	0.5	0.05	0.12	0.05	1.6449

* Power was computed using the normal approximation method.

This report shows the same sample size when the difference $d_0 = 0.05$ is entered as the superiority difference and $d_1 = 0.12$ is entered as the actual difference.

Setup (Ratios)

To input $P_0 = 0.55$ and $P_1 = 0.62$ using ratios, recall the relationships between P_0 , P_1 , PB , r_0 and r_1 :

$$r_0 = P_0/PB = 0.55/0.5 = 1.1$$

$$r_1 = P_1/PB = 0.62/0.5 = 1.24$$

The setup for ratios is exactly the same as that for proportions except for the following inputs on the Design tab. You may then make the appropriate entries as listed below, or open **Example 2c** by going to the **File** menu and choosing **Open Example Template**.

Design Tab	
Input Type.....	Ratios
PB (Baseline Proportion)	0.50
r0 (Superiority Ratio).....	1.1
r1 (Actual Ratio).....	1.24

Output (Ratios)

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

Numeric Results						
Solve For:	Sample Size					
Higher Proportions Are:	Better					
Alternative Hypothesis:	One-Sided (H0: $P \leq P_0$ vs. H1: $P > P_0$)					
Test Type:	Z-Test with S(P0)					
Power*	Sample Size n	Baseline Proportion PB	Ratio		Alpha	Reject H0 If Z >
			Superiority r0	Actual r1		
0.90037	424	0.5	1.1	1.24	0.05	1.6449

* Power was computed using the normal approximation method.

This report shows the same sample size when the ratio $r_0 = 1.1$ is entered as the superiority ratio and $r_1 = 1.24$ is entered as the actual ratio.

Setup (Odds Ratios)

To input $P_0 = 0.55$ and $P_1 = 0.62$ using odds ratios, recall the relationships between P_0 , P_1 , P_B , o_0 and o_1 :

$$o_0 = \text{odds}_0 / \text{odds}_B = [0.55 / (1 - 0.55)] / [0.5 / (1 - 0.5)] = 1.222222222$$

$$o_1 = \text{odds}_1 / \text{odds}_B = [0.62 / (1 - 0.62)] / [0.5 / (1 - 0.5)] = 1.631578947$$

The setup for odds ratios is exactly the same as that for proportions except for the following inputs on the Design tab. You may then make the appropriate entries as listed below, or open **Example 2d** by going to the **File** menu and choosing **Open Example Template**.

Design Tab	
Input Type.....	Odds Ratios
PB (Baseline Proportion)	0.50
o0 (Superiority Odds Ratio)	1.222222222
o1 (Actual Odds Ratio).....	1.631578947

Output (Odds Ratios)

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

Numeric Results						
Solve For:	Sample Size					
Higher Proportions Are:	Better					
Alternative Hypothesis:	One-Sided (H0: $P \leq P_0$ vs. H1: $P > P_0$)					
Test Type:	Z-Test with S(P0)					
Power*	Sample Size n	Baseline Proportion PB	Odds Ratio		Alpha	Reject H0 If Z >
			Superiority o0	Actual o1		
0.90037	424	0.5	1.2222	1.6316	0.05	1.6449

* Power was computed using the normal approximation method.

This report shows the same sample size when the ratio $o_0 = 1.222222222$ is entered as the superiority odds ratio and $o_1 = 1.631578947$ is entered as the actual odds ratio.

Example 3 – Comparing Test Statistics

One important decision that must be made before conducting the experiment is to decide which of the available test statistics to use. This procedure makes it easy to make this comparison. The differences in the characteristics of test statistics are most noticeable in small samples. Hence, the investigation done here is for sample sizes of 20 to 200 in steps of 20. The trivial difference will be set to 0.05. We'll calculate exact power using binomial enumeration when the baseline proportion (PB) is 0.5, the actual difference (d1) is 0.12, and $\alpha = 0.05$.

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	Power
Power Calculation Method.....	Binomial Enumeration
Max n for Binomial Enumeration.....	10000
Higher Proportions Are	Better
Test Type.....	Z-test using S(P0)
Alpha.....	0.05
n (Sample Size)	20 to 200 by 20
Input Type.....	Differences
PB (Baseline Proportion)	0.50
d0 (Superiority Difference)	0.05
d1 (Actual Difference)	0.12

Reports Tab

Show Comparative Reports	Checked
--------------------------------	----------------

Comparative Plots Tab

Show Comparative Plots.....	Checked
-----------------------------	----------------

Superiority by a Margin Tests for One Proportion

Output

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

Power Comparison of Five Different Superiority Tests for One Proportion

Higher Proportions Are: Better
 Alternative Hypothesis: One-Sided ($H_0: P \leq P_0$ vs. $H_1: P > P_0$)

Sample Size n	Baseline Proportion PB	Difference			Power				
		Superiority d0	Actual d1	Target Alpha	Exact Test	Z-Test S(P0)	Z-Test S(P0)C	Z-Test S(P)	Z-Test S(P)C
20	0.5	0.05	0.12	0.05	0.07261	0.16707	0.07261	0.16707	0.16707
40	0.5	0.05	0.12	0.05	0.19049	0.19049	0.19049	0.29333	0.19049
60	0.5	0.05	0.12	0.05	0.27278	0.27278	0.27278	0.27278	0.27278
80	0.5	0.05	0.12	0.05	0.33369	0.33369	0.33369	0.33369	0.33369
100	0.5	0.05	0.12	0.05	0.38160	0.38160	0.38160	0.46227	0.38160
120	0.5	0.05	0.12	0.05	0.42094	0.49551	0.42094	0.49551	0.42094
140	0.5	0.05	0.12	0.05	0.45425	0.52359	0.45425	0.52359	0.52359
160	0.5	0.05	0.12	0.05	0.54792	0.54792	0.54792	0.54792	0.54792
180	0.5	0.05	0.12	0.05	0.56939	0.62849	0.56939	0.62849	0.56939
200	0.5	0.05	0.12	0.05	0.58862	0.64398	0.58862	0.64398	0.64398

Note: Power was computed using binomial enumeration of all possible outcomes.

Actual Alpha Comparison of Five Different Superiority Tests for One Proportion

Higher Proportions Are: Better
 Alternative Hypothesis: One-Sided ($H_0: P \leq P_0$ vs. $H_1: P > P_0$)

Sample Size n	Baseline Proportion PB	Difference			Alpha				
		Superiority d0	Actual d1	Target	Exact Test	Z-Test S(P0)	Z-Test S(P0)C	Z-Test S(P)	Z-Test S(P)C
20	0.5	0.05	0.12	0.05	0.0189	0.0553	0.0189	0.0553	0.0553
40	0.5	0.05	0.12	0.05	0.0386	0.0386	0.0386	0.0751	0.0386
60	0.5	0.05	0.12	0.05	0.0446	0.0446	0.0446	0.0446	0.0446
80	0.5	0.05	0.12	0.05	0.0449	0.0449	0.0449	0.0449	0.0449
100	0.5	0.05	0.12	0.05	0.0429	0.0429	0.0429	0.0651	0.0429
120	0.5	0.05	0.12	0.05	0.0399	0.0587	0.0399	0.0587	0.0399
140	0.5	0.05	0.12	0.05	0.0365	0.0526	0.0365	0.0526	0.0526
160	0.5	0.05	0.12	0.05	0.0470	0.0470	0.0470	0.0470	0.0470
180	0.5	0.05	0.12	0.05	0.0419	0.0573	0.0419	0.0573	0.0419
200	0.5	0.05	0.12	0.05	0.0373	0.0505	0.0373	0.0505	0.0505

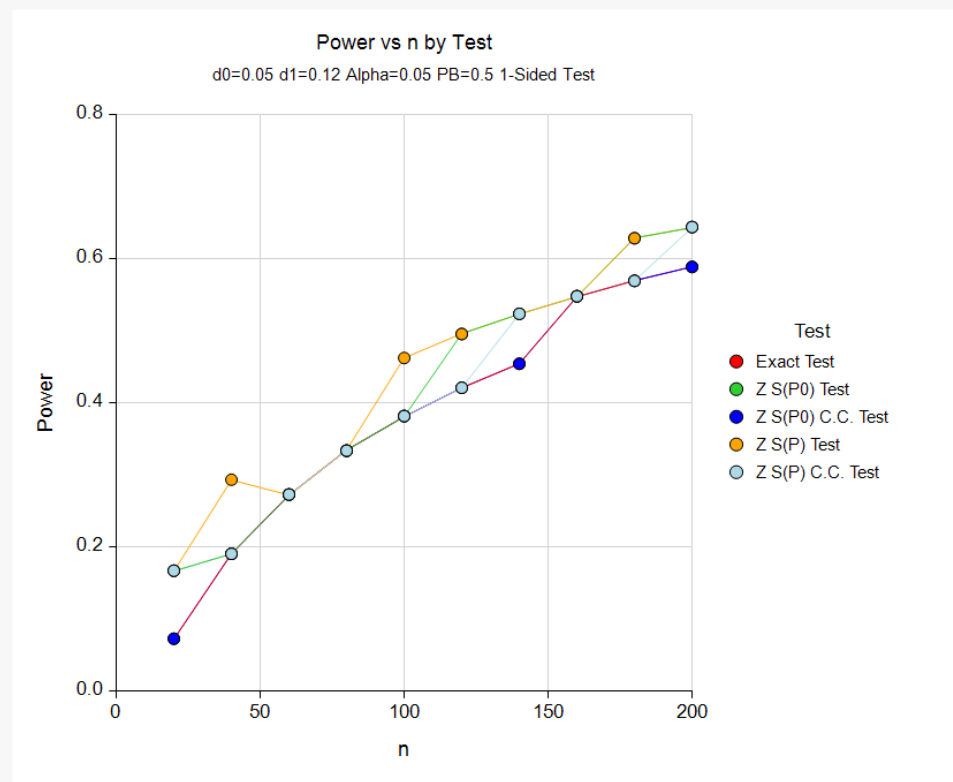
Note: Actual alpha was computed using binomial enumeration of all possible outcomes.

References

- Blackwelder, W.C. 1998. 'Equivalence Trials.' In Encyclopedia of Biostatistics, John Wiley and Sons. New York. Volume 2, 1367-1372.
- Chow, S. C., Shao, J., and Wang, H. 2008. Sample Size Calculations in Clinical Research, Second Edition. Chapman & Hall/CRC. Boca Raton, Florida.
- Fleiss, J. L., Levin, B., and Paik, M.C. 2003. Statistical Methods for Rates and Proportions. Third Edition. John Wiley & Sons. New York.

Superiority by a Margin Tests for One Proportion

Plots



The first numeric report shows the power for each test statistic. The second shows the actual alpha achieved by the design.

An examination of the first report shows that the power is often different for at least one of the tests. Also notice that the exact test always has the minimum power in each row. This would lead us to discard this test statistic. However, consider the second report which shows the actual alpha level (the target was 0.05) for each test. By inspecting corresponding entries in both tables, we see that whenever a test statistic achieves a better power than the exact test, it also yields an actual alpha level larger than the target alpha.

For example, look at the powers for $n = 20$. The z-test using \hat{s} has an unusually large power = 0.16707. This is a much larger power than the exact test's value of 0.07261. However, note that the actual alpha for this test is 0.0553 which is larger than the target alpha of 0.05.

We conclude that indeed, the exact test is consistently the best test since it always achieves a significance level that is less than the target value.

Example 4 – Comparing Power Calculation Methods

Continuing with Example 3, let's see how the results compare if we were to use approximate power calculations instead of power calculations based on binomial enumeration.

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 4** 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
Power Calculation Method	Binomial Enumeration
Max n for Binomial Enumeration	10000
Higher Proportions Are	Better
Test Type	Z-test using S(P0)
Alpha	0.05
n (Sample Size)	20 to 200 by 20
Input Type	Differences
PB (Baseline Proportion)	0.50
d0 (Superiority Difference)	0.05
d1 (Actual Difference)	0.12

Reports Tab

Show Power Detail Report	Checked
--------------------------------	----------------

Superiority by a Margin Tests for One Proportion

Output

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

Power Detail Report for Testing Superiority of One Proportion using the Z-Test with S(P0)

Higher Proportions Are: Better
 Alternative Hypothesis: One-Sided (H0: $P \leq P_0$ vs. H1: $P > P_0$)

Sample Size n	Baseline Proportion PB	Difference		Normal Approximation		Binomial Enumeration	
		Superiority d0	Actual d1	Power	Alpha	Power	Alpha
20	0.5	0.05	0.12	0.14895	0.05	0.16707	0.0553
40	0.5	0.05	0.12	0.21953	0.05	0.19049	0.0386
60	0.5	0.05	0.12	0.28475	0.05	0.27278	0.0446
80	0.5	0.05	0.12	0.34606	0.05	0.33369	0.0449
100	0.5	0.05	0.12	0.40372	0.05	0.38160	0.0429
120	0.5	0.05	0.12	0.45776	0.05	0.49551	0.0587
140	0.5	0.05	0.12	0.50817	0.05	0.52359	0.0526
160	0.5	0.05	0.12	0.55500	0.05	0.54792	0.0470
180	0.5	0.05	0.12	0.59831	0.05	0.62849	0.0573
200	0.5	0.05	0.12	0.63819	0.05	0.64398	0.0505

Notice that the power values are fairly consistent for both methods. For some sample sizes, however, the actual alpha level computed using binomial enumeration is greater than the target of 0.05.

Example 5 – Finding Power after an Experiment

The proportion of successes of the current treatment is known to be 0.64 based on years of treatment use. Researchers have developed a new method of treatment which costs much more than the current treatment price. Before the new treatment can be approved it must be shown that the success of the proposed treatment is superior to that of the current treatment. It is determined that use of the new treatment is justifiable if it is shown that it is effective more than 74% of the time. Sixty individuals are randomly selected to receive the new method of treatment. Forty-six (76.67%) of the 60 individuals responded positively to the treatment. The p-value for the test based on exact binomial probabilities is 0.3814. Because the researchers were unable to show the new treatment is superior, they desire to know the power of the test if the actual difference is 0.13.

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 5** 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
Power Calculation Method	Binomial Enumeration
Max n for Binomial Enumeration	10000
Higher Proportions Are	Better
Test Type	Exact Test
N (Population Size)	Infinite
Alpha	0.05
n (Sample Size)	60
Input Type	Differences
PB (Baseline Proportion)	0.64
d0 (Superiority Difference)	0.10
d1 (Actual Difference)	0.13

Superiority by a Margin Tests for One Proportion

Output

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

Numeric Results

Solve For: Power
 Higher Proportions Are: Better
 Alternative Hypothesis: One-Sided (H0: $P \leq P_0$ vs. H1: $P > P_0$)
 Test Type: Exact Test

Power*	Sample Size n	Baseline Proportion PB	Difference		Alpha		Reject H0 If $R \geq$
			Superiority d0	Actual d1	Target	Actual*	
0.08932	60	0.64	0.1	0.13	0.05	0.0312	51

* Power and actual alpha were computed using binomial enumeration of all possible outcomes.

This report shows that the power for the test was only 0.08932.

Example 6 – Finding Sample Size based on the Odds Ratio

The odds for success of the current treatment are known to be 3.87. A new treatment is developed to compete with the current treatment. It must be shown to be superior to the current treatment to be adopted. The researchers want to determine the sample size necessary to achieve 80% power in this test of superiority. The researchers determine that the new treatment will be considered superior if the odds for success are greater than 1.2 times the odds for success of the current treatment. The baseline proportion is calculated as $PB = \text{odds}/(1+\text{odds}) = 3.87/(1+3.87) = 0.7947$. They want to calculate sample sizes for actual odds ratios between 1.3 and 1.5.

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 6** 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 Calculation Method **Normal Approximation**
 Higher Proportions Are **Better**
 Test Type **Z-test using S(P0)**
 N (Population Size) **Infinite**
 Power **0.80**
 Alpha **0.05**
 Input Type **Odds Ratios**
 PB (Baseline Proportion) **0.7947**
 o0 (Superiority Odds Ratio) **1.2**
 o1 (Actual Odds Ratio) **1.3 1.4 1.5**

Superiority by a Margin Tests for One Proportion

Output

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

Numeric Results

Solve For: [Sample Size](#)
 Higher Proportions Are: Better
 Alternative Hypothesis: One-Sided (H0: $P \leq P_0$ vs. H1: $P > P_0$)
 Test Type: Z-Test with S(P0)

Power*	Sample Size n	Baseline Proportion PB	Odds Ratio		Alpha	Reject H0 If Z >
			Superiority o0	Actual o1		
0.80005	6853	0.7947	1.2	1.3	0.05	1.6449
0.80019	1909	0.7947	1.2	1.4	0.05	1.6449
0.80016	939	0.7947	1.2	1.5	0.05	1.6449

The sample size range between 6853 and 939 to show superiority under these conditions based on the actual odds ratio.

Example 7 – Validation using Chow, Shao, and Wang (2008)

Chow, Shao, and Wang (2008) page 87 gives the result of a sample size calculation for the z-test with S(Phat). Although the test given there is for non-inferiority, the form of the test is the same so we can use those results to validate the calculations of this procedure. They calculate a sample size of 18 when alpha = 0.05, power = 0.80, PB = 0.30, non-inferiority difference = -0.10 ($P_0 = 0.30 - 0.10 = 0.2$), and actual difference = 0.20 (such that actual proportion = 0.50).

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 7** 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 Calculation Method	Normal Approximation
Higher Proportions Are	Better
Test Type	Z-test using S(Phat)
Power	0.80
Alpha	0.05
Input Type	Proportions
P0 (Superiority Proportion)	0.2
P1 (Actual Proportion)	0.5

Output

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

Numeric Results					
Solve For:	Sample Size				
Higher Proportions Are:	Better				
Alternative Hypothesis:	One-Sided (H0: $P \leq P_0$ vs. H1: $P > P_0$)				
Test Type:	Z-Test with S(Phat)				
Power*	Sample Size n	Proportion		Alpha	Reject H0 If Z >
		Superiority P0	Actual P1		
0.81613	18	0.2	0.5	0.05	1.6449

* Power was computed using the normal approximation method.

PASS also calculated a sample size of 18.