Chapter 216

Confidence Intervals for Two Proportions

Introduction
This routine calculates the group sample sizes necessary to achieve a specified interval width of the difference, ratio, or odds ratio of two independent proportions.

Caution: These procedures assume that the proportions obtained from future samples will be the same as the proportions that are specified. If the sample proportions are different from those specified when running these procedures, the interval width may be narrower or wider than specified.

Four Procedures Documented Here
There are four procedures in the menus described in this chapter. These procedures are very similar except for the type of parameterization. The parameterization can be in terms of proportions, differences in proportions, ratios of proportions, and odds ratios.

Technical Details
A background of the comparison of two proportions is given, followed by details of the confidence interval methods available in this procedure.

Comparing Two Proportions
Suppose you have two populations from which dichotomous (binary) responses will be recorded. The probability (or risk) of obtaining the event of interest in population 1 (the treatment group) is \( p_1 \) and in population 2 (the control group) is \( p_2 \). The corresponding failure proportions are given by \( q_1 = 1 - p_1 \) and \( q_2 = 1 - p_2 \).

The assumption is made that the responses from each group follow a binomial distribution. This means that the event probability \( p_i \) is the same for all subjects within a population and that the responses from one subject to the next are independent of one another.

Random samples of \( m \) and \( n \) individuals are obtained from these two populations. The data from these samples can be displayed in a 2-by-2 contingency table as follows

<table>
<thead>
<tr>
<th></th>
<th>Success</th>
<th>Failure</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population 1</td>
<td>( a )</td>
<td>( c )</td>
<td>( m )</td>
</tr>
<tr>
<td>Population 2</td>
<td>( b )</td>
<td>( d )</td>
<td>( n )</td>
</tr>
<tr>
<td>Totals</td>
<td>( s )</td>
<td>( f )</td>
<td>( N )</td>
</tr>
</tbody>
</table>
The following alternative notation is sometimes used:

<table>
<thead>
<tr>
<th></th>
<th>Success</th>
<th>Failure</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population 1</strong></td>
<td>$x_{11}$</td>
<td>$x_{12}$</td>
<td>$n_1$</td>
</tr>
<tr>
<td><strong>Population 2</strong></td>
<td>$x_{21}$</td>
<td>$x_{22}$</td>
<td>$n_2$</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td>$m_1$</td>
<td>$m_2$</td>
<td>$N$</td>
</tr>
</tbody>
</table>

The binomial proportions $p_1$ and $p_2$ are estimated from these data using the formulae

$$\hat{p}_1 = \frac{a}{m} = \frac{x_{11}}{n_1} \text{ and } \hat{p}_2 = \frac{b}{n} = \frac{x_{21}}{n_2}$$

When analyzing studies such as these, you usually want to compare the two binomial probabilities $p_1$ and $p_2$.

The most direct methods of comparing these quantities are to calculate their difference or their ratio. If the binomial probability is expressed in terms of odds rather than probability, another measure is the odds ratio. Mathematically, these comparison parameters are

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Computation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Difference</strong></td>
<td>$\delta = p_1 - p_2$</td>
</tr>
<tr>
<td><strong>Risk Ratio</strong></td>
<td>$\phi = \frac{p_1}{p_2}$</td>
</tr>
<tr>
<td><strong>Odds Ratio</strong></td>
<td>$\psi = \frac{p_1/q_1}{p_2/q_2} = \frac{p_1q_2}{p_2q_1}$</td>
</tr>
</tbody>
</table>

The choice of which of these measures is used might at seem arbitrary, but it is important. Not only is their interpretation different, but, for small sample sizes, the coverage probabilities may be different.

**Difference**

The (risk) difference $\delta = p_1 - p_2$ is perhaps the most direct method of comparison between the two event probabilities. This parameter 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 interpretation difficulty occurs when the event of interest is rare. If a difference of 0.001 were reported for an event with a baseline probability of 0.40, we would probably dismiss this as being of little importance. That is, there usually little interest in a treatment that decreases the probability from 0.40 to 0.399. However, if the baseline probability of a disease was 0.002 and 0.001 was the decrease in the disease probability, this would represent a reduction of 50%. Thus we see that interpretation depends on the baseline probability of the event.

A similar situation occurs when the amount of possible difference is considered. Consider two events, one with a baseline event rate of 0.40 and the other with a rate of 0.02. What is the maximum decrease that can occur? Obviously, the first event rate can be decreased by an absolute amount of 0.40 which the second can only be decreased by a maximum of 0.02.

So, although creating the simple difference is a useful method of comparison, care must be taken that it fits the situation.

**Ratio**

The (risk) ratio $\phi = \frac{p_1}{p_2}$ gives the relative change in the disease risk due to the application of the treatment. This parameter is also direct and easy to interpret. To compare this with the difference, consider a treatment that reduces the risk of disease for 0.1437 to 0.0793. Which single number is most enlightening, the fact that the
absolute risk of disease has been decreased by 0.0644, or the fact that risk of disease in the treatment group is only 55.18% of that in the control group? In many cases, the percentage (risk ratio) communicates the impact of the treatment better than the absolute change.

Perhaps the biggest drawback to this parameter is that it cannot be calculated in one of the most common experimental designs: the case-control study.

**Odds Ratio**

Chances are usually communicated as long-term proportions or probabilities. In betting, chances are often given as odds. For example, the odds of a horse winning a race might be set at 10-to-1 or 3-to-2. How do you translate from odds to probability? An odds of 3-to-2 means that the event will occur three out of five times. That is, an odds of 3-to-2 (1.5) translates to a probability of winning of 0.60.

The odds of an event are calculated by dividing the event risk by the non-event risk. Thus, in our case of two populations, the odds are

\[ o_1 = \frac{p_1}{1 - p_1} \quad \text{and} \quad o_2 = \frac{p_2}{1 - p_2} \]

For example, if \( p_1 \) is 0.60, the odds are 0.60/0.4 = 1.5. Rather than represent the odds as a decimal amount, it is re-scaled into whole numbers. Thus, instead of saying the odds are 1.5-to-1, we say they are 3-to-2.

Another way to compare proportions is to compute the ratio of their odds. The odds ratio of two events is

\[ \psi = \frac{o_1}{o_2} = \frac{p_1}{1 - p_1} \quad \frac{p_2}{1 - p_2} \]

Although the odds ratio is more complicated to interpret than the risk ratio, it is often the parameter of choice. Reasons for this include the fact that the odds ratio can be accurately estimated from case-control studies, while the risk ratio cannot. Also, the odds ratio is the basis of logistic regression (used to study the influence of risk factors). Furthermore, the odds ratio is the natural parameter in the conditional likelihood of the two-group, binomial-response design. Finally, when the baseline event-rates are rare, the odds ratio provides a close approximation to the risk ratio since, in this case, \( 1 - p_1 \approx 1 - p_2 \), so that

\[ \psi = \frac{1 - p_1}{p_2} \approx \frac{p_1}{p_2} = \phi \]

**Confidence Intervals for the Difference**

Many methods have been devised for computing confidence intervals for the difference between two proportions \( \delta = p_1 - p_2 \). Seven of these methods are available in the Confidence Intervals for Two Proportions [Proportions] using Proportions and Confidence Intervals for Two Proportions [Differences] procedures. The seven confidence interval methods are

1. Score (Farrington and Manning)
2. Score (Miettinen and Nurminen)
3. Score with Correction for Skewness (Gart and Nam)
4. Score (Wilson)
5. Score with Continuity Correction (Wilson)
6. Chi-Square with Continuity Correction (Yates)
7. Chi-Square (Pearson)

Newcombe (1998b) conducted a comparative evaluation of eleven confidence interval methods. He recommended that the modified Wilson score method be used instead of the Pearson Chi-Square or the Yate’s Corrected Chi-Square. Beal (1987) found that the Score methods performed very well. The lower \( L \) and upper \( U \) limits of these intervals are computed as follows. Note that, unless otherwise stated, \( z = \left| z_{\alpha/2} \right| \) is the appropriate percentile from the standard normal distribution.

**C.I. for Difference: Farrington and Manning’s Score**

Farrington and Manning (1990) proposed a test statistic for testing whether the difference is equal to a specified value \( \delta_0 \). The regular MLE’s \( \hat{p}_1 \) and \( \hat{p}_2 \) are used in the numerator of the score statistic while MLE’s \( \tilde{p}_1 \) and \( \tilde{p}_2 \) constrained so that \( \tilde{p}_1 - \tilde{p}_2 = \delta_0 \) are used in the denominator. The significance level of the test statistic is based on the asymptotic normality of the score statistic.

The test statistic formula is

\[
z_{FMD} = \frac{\hat{p}_1 - \hat{p}_2 - \delta_0}{\sqrt{\left( \frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2} \right)}}
\]

where the estimates \( \tilde{p}_1 \) and \( \tilde{p}_2 \) are computed as in the corresponding test of Miettinen and Nurminen (1985) given as

\[
\tilde{p}_1 = \tilde{p}_2 + \delta_0
\]

\[
\tilde{p}_2 = 2B \cos(A) - \frac{L_2}{3L_3}
\]

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

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

\[
C = \frac{L_3^2}{27L_3^2} - \frac{L_4L_2}{6L_3^2} + \frac{L_0}{2L_3}
\]

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

\[
L_1 = [N_2 \delta_0 - N - 2x_{21}] \delta_0 + M_1
\]

\[
L_2 = (N + N_2) \delta_0 - N - M_1
\]

\[
L_3 = N
\]
Confidence Intervals for Two Proportions

Farrington and Manning (1990) proposed inverting their score test to find the confidence interval. The lower limit is found by solving

\[ z_{FMD} = \left| z_{\alpha/2} \right| \]

and the upper limit is the solution of

\[ z_{FMD} = -\left| z_{\alpha/2} \right| \]

C.I. for Difference: Miettinen and Nurminen’s Score

Miettinen and Nurminen (1985) proposed a test statistic for testing whether the difference is equal to a specified value \( \delta_0 \). The regular MLE’s \( \hat{p}_1 \) and \( \hat{p}_2 \) are used in the numerator of the score statistic while MLE’s \( \tilde{p}_1 \) and \( \tilde{p}_2 \) constrained so that \( \tilde{p}_1 - \tilde{p}_2 = \delta_0 \) 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

\[
    z_{MND} = \frac{\hat{p}_1 - \hat{p}_2 - \delta_0}{\sqrt{\left( \frac{\tilde{p}_1 \tilde{q}_1}{n_1} + \frac{\tilde{p}_2 \tilde{q}_2}{n_2} \right) \left( \frac{N}{N-1} \right)}}
\]

where

\[
    \tilde{p}_1 = \tilde{p}_2 + \delta_0
\]

\[
    \tilde{p}_2 = 2B \cos(A) - \frac{L_2}{3L_3}
\]

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

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

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

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

\[
    L_1 = [N_2 \delta_0 - N - 2x_{21}] \delta_0 + M_1
\]

\[
    L_2 = (N + N_2) \delta_0 - N - M_1
\]

\[
    L_3 = N
\]

Miettinen and Nurminen (1985) proposed inverting their score test to find the confidence interval. The lower limit is found by solving

\[ z_{MND} = \left| z_{\alpha/2} \right| \]

and the upper limit is the solution of

\[ z_{MND} = -\left| z_{\alpha/2} \right| \]
C.I. for Difference: Gart and Nam’s Score
Gart and Nam (1990) page 638 proposed a modification to the Farrington and Manning (1990) difference test that corrected for skewness. Let \( z_{FM}(\delta) \) stand for the Farrington and Manning difference test statistic described above. The skewness corrected test statistic \( z_{GN} \) is the appropriate solution to the quadratic equation

\[
(-\tilde{\gamma})z_{GND}^2 + (-1)z_{GND} + \left(z_{FMD}(\delta) + \tilde{\gamma}\right) = 0
\]

where

\[
\tilde{\gamma} = \frac{V^{3/2}(\delta)}{6} \left( \frac{\hat{p}_1q_1(\hat{q}_1 - \hat{p}_1)}{n_1^2} - \frac{\hat{p}_2q_2(\hat{q}_2 - \hat{p}_2)}{n_2^2} \right)
\]

Gart and Nam (1988) proposed inverting their score test to find the confidence interval. The lower limit is found by solving

\[
z_{GND} = \left| z_{\alpha/2} \right|
\]

and the upper limit is the solution of

\[
z_{GND} = -\left| z_{\alpha/2} \right|
\]

C.I. for Difference: Wilson’s Score as Modified by Newcombe (with and without Continuity Correction)
For details, see Newcombe (1998b), page 876.

\[
L = \hat{p}_1 - \hat{p}_2 - B
\]
\[
U = \hat{p}_1 - \hat{p}_2 + C
\]

where

\[
B = z \sqrt{\frac{l_1(1-l_1)}{m} + \frac{u_2(1-u_2)}{n}}
\]
\[
C = z \sqrt{\frac{u_1(1-u_1)}{m} + \frac{l_2(1-l_2)}{n}}
\]

and \( l_1 \) and \( u_1 \) are the roots of

\[
|p_1 - \hat{p}_1| - z \sqrt{\frac{p_1(1-p_1)}{m}} = 0
\]

and \( l_2 \) and \( u_2 \) are the roots of

\[
|p_2 - \hat{p}_2| - z \sqrt{\frac{p_2(1-p_2)}{n}} = 0
\]
Confidence Intervals for Two Proportions

C.I. for Difference: Yate’s Chi-Square with Continuity Correction

For details, see Newcombe (1998b), page 875.

\[
L = \hat{p}_1 - \hat{p}_2 - z \left( \frac{\hat{p}_1(1 - \hat{p}_1)}{m} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n} \right) + \frac{1}{2} \left( \frac{1}{m} + \frac{1}{n} \right)
\]

\[
U = \hat{p}_1 - \hat{p}_2 + z \left( \frac{\hat{p}_1(1 - \hat{p}_1)}{m} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n} \right) + \frac{1}{2} \left( \frac{1}{m} + \frac{1}{n} \right)
\]

C.I. for Difference: Pearson’s Chi-Square

For details, see Newcombe (1998b), page 875.

\[
L = \hat{p}_1 - \hat{p}_2 - z \sqrt{ \frac{\hat{p}_1(1 - \hat{p}_1)}{m} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n} }
\]

\[
U = \hat{p}_1 - \hat{p}_2 + z \sqrt{ \frac{\hat{p}_1(1 - \hat{p}_1)}{m} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n} }
\]

For each of the seven methods, one-sided intervals may be obtained by replacing \( \alpha/2 \) by \( \alpha \).

For two-sided intervals, the distance from the difference in sample proportions to each of the limits may be different. Thus, instead of specifying the distance to the limits we specify the width of the interval, \( W \).

The basic equation for determining sample size for a two-sided interval when \( W \) has been specified is

\[
W = U - L
\]

For one-sided intervals, the distance from the variance ratio to limit, \( D \), is specified.

The basic equation for determining sample size for a one-sided upper limit when \( D \) has been specified is

\[
D = U - (\hat{p}_1 - \hat{p}_2)
\]

The basic equation for determining sample size for a one-sided lower limit when \( D \) has been specified is

\[
D = (\hat{p}_1 - \hat{p}_2) - L
\]

Each of these equations can be solved for any of the unknown quantities in terms of the others.

Confidence Intervals for the Ratio (Relative Risk)

Many methods have been devised for computing confidence intervals for the ratio (relative risk) of two proportions \( \phi = p_1 / p_2 \). Six of these methods are available in the Confidence Intervals for Two Proportions [Ratios] procedure. The six confidence interval methods are

1. Score (Farrington and Manning)
2. Score (Miettinen and Nurminen)
3. Score with Correction for Skewness (Gart and Nam)
4. Logarithm (Katz)
5. Logarithm + 1/2 (Walter)
6. Fleiss
C.I. for Ratio: Farrington and Manning’s Score

Farrington and Manning (1990) proposed a test statistic for testing whether the ratio is equal to a specified value $\phi_0$. The regular MLE’s $\hat{p}_1$ and $\hat{p}_2$ are used in the numerator of the score statistic while MLE’s $\tilde{p}_1$ and $\tilde{p}_2$ constrained so that $\tilde{p}_1 / \tilde{p}_2 = \phi_0$ are used in the denominator. A correction factor of N/(N-1) is applied to increase the variance estimate. The significance level of the test statistic is based on the asymptotic normality of the score statistic.

Here is the formula for computing the test

$$z_{FMR} = \frac{\hat{p}_1 / \hat{p}_2 - \phi_0}{\sqrt{\left( \frac{\hat{p}_1 \hat{q}_1 + \phi_0^2 \tilde{p}_2 \tilde{q}_2}{n_1} \right)}}$$

where

$$\tilde{p}_1 = p_2 \phi_0$$

$$\tilde{p}_2 = \frac{-B - \sqrt{B^2 - 4AC}}{2A}$$

$$A = N \phi_0$$

$$B = -\left[N_1 \phi_0 + x_{11} + N_2 + x_{21} \phi_0 \right]$$

$$C = M_1$$

as in the test of Miettinen and Nurminen (1985).

Farrington and Manning (1990) proposed inverting their score test to find the confidence interval. The lower limit is found by solving

$$z_{FMR} = \left| z_{\alpha/2} \right|$$

and the upper limit is the solution of

$$z_{FMR} = -\left| z_{\alpha/2} \right|$$

C.I. for Ratio: Miettinen and Nurminen’s Score

Miettinen and Nurminen (1985) proposed a test statistic for testing whether the ratio is equal to a specified value $\phi_0$. The regular MLE’s $\hat{p}_1$ and $\hat{p}_2$ are used in the numerator of the score statistic while MLE’s $\tilde{p}_1$ and $\tilde{p}_2$ constrained so that $\tilde{p}_1 / \tilde{p}_2 = \phi_0$ 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.

Here is the formula for computing the test

$$z_{MNR} = \frac{\hat{p}_1 / \hat{p}_2 - \phi_0}{\sqrt{\left( \frac{\hat{p}_1 \hat{q}_1 + \phi_0^2 \tilde{p}_2 \tilde{q}_2}{n_1} \right)}} \cdot \left( \frac{N}{N - 1} \right)$$
where

\[
\hat{p}_1 = \hat{p}_2 \phi_0
\]

\[
\hat{p}_2 = -\frac{B - \sqrt{B^2 - 4AC}}{2A}
\]

\[
A = N \phi_0
\]

\[
B = -\left[N_1 \phi_0 + \chi_{11} + N_2 + \chi_{21} \phi_0\right]
\]

\[
C = M_1
\]

Miettinen and Nurminen (1985) proposed inverting their score test to find the confidence interval. The lower limit is found by solving

\[
z_{MNR} = \left|z_{\alpha/2}\right|
\]

and the upper limit is the solution of

\[
z_{MNR} = -\left|z_{\alpha/2}\right|
\]

**C.I. for Ratio: Gart and Nam’s Score**

Gart and Nam (1988) page 329 proposed a modification to the Farrington and Manning (1988) ratio test that corrected for skewness. Let \(z_{FM}(\phi)\) stand for the Farrington and Manning ratio test statistic described above. The skewness corrected test statistic \(z_{GN}\) is the appropriate solution to the quadratic equation

\[
(-\bar{\phi})z_{GN}^2 + (-1)z_{GN} + (z_{FMR}(\phi) + \bar{\phi}) = 0
\]

where

\[
\bar{\phi} = \frac{1}{6\bar{u}^{3/2}} \left( \frac{\bar{q}_1(\bar{q}_1 - \bar{p}_1)}{n_1 \bar{p}_1^2} - \frac{\bar{q}_2(\bar{q}_2 - \bar{p}_2)}{n_2 \bar{p}_2^2} \right)
\]

\[
\bar{u} = \frac{\bar{q}_1}{n_1 \bar{p}_1} + \frac{\bar{q}_2}{n_2 \bar{p}_2}
\]

Gart and Nam (1988) proposed inverting their score test to find the confidence interval. The lower limit is found by solving

\[
z_{GNR} = \left|z_{\alpha/2}\right|
\]

and the upper limit is the solution of

\[
z_{GNR} = -\left|z_{\alpha/2}\right|
\]
C.I. for Ratio: Logarithm (Katz)

This was one of the first methods proposed for computing confidence intervals for risk ratios.
For details, see Gart and Nam (1988), page 324.

\[
L = \hat{\phi} \exp \left( - z \left( \frac{\hat{q}_1 + \hat{q}_2}{n\hat{p}_1 - n\hat{p}_2} \right) \right)
\]
\[
U = \hat{\phi} \exp \left( z \left( \frac{\hat{q}_1 + \hat{q}_2}{n\hat{p}_1 - n\hat{p}_2} \right) \right)
\]

where
\[
\hat{\phi} = \frac{\hat{p}_1}{\hat{p}_2}
\]

C.I. for Ratio: Logarithm (Walters)

For details, see Gart and Nam (1988), page 324.

\[
L = \hat{\phi} \exp \left( - z\sqrt{u} \right)
\]
\[
U = \hat{\phi} \exp \left( z\sqrt{u} \right)
\]

where
\[
\hat{\phi} = \exp \left( \ln \left( \frac{a + \frac{1}{2}}{m + \frac{1}{2}} \right) - \ln \left( \frac{b + \frac{1}{2}}{n + \frac{1}{2}} \right) \right)
\]
\[
\hat{u} = 1 + \frac{1}{m + \frac{1}{2}} + \frac{1}{b + \frac{1}{2}} - \frac{1}{n + \frac{1}{2}}
\]
\[
\tilde{q}_2 = 1 - \tilde{p}_2
\]
\[
V = \left( \hat{\phi} \left( \frac{\tilde{q}_1 + \tilde{q}_2}{n\hat{p}_1 - n\hat{p}_2} \right) \right)^{-1}
\]
\[
\tilde{p}_1 = \hat{\phi}\tilde{p}_2
\]
\[
\tilde{q}_1 = 1 - \tilde{p}_1
\]
\[
\tilde{q}_2 = 1 - \tilde{p}_2
\]
\[
\tilde{\mu}_3 = V^{3/2} \left( \frac{\tilde{q}_1 (\tilde{q}_1 - \tilde{p}_1)}{(m\hat{p}_1)^2} - \frac{\tilde{q}_2 (\tilde{q}_2 - \tilde{p}_2)}{(n\hat{p}_2)^2} \right)
\]
\[
v = \left( \frac{\tilde{q}_1 + \tilde{q}_2}{m\hat{p}_1 - n\hat{p}_2} \right)^{-1}
\]
Confidence Intervals for Two Proportions

C.I. for Odds Ratio and Relative Risk: Iterated Method of Fleiss

Fleiss (1981) presents an improved confidence interval for the odds ratio and relative risk. This method forms the confidence interval as all those value of the odds ratio which would not be rejected by a chi-square hypothesis test. Fleiss gives the following details about how to construct this confidence interval. To compute the lower limit, do the following.

1. For a trial value of $\psi$, compute the quantities $X, Y, W, F, U,$ and $V$ using the formulas

   \[ X = \psi(m+s) + (n-s) \]
   \[ Y = \sqrt{X^2 - 4ms\psi(\psi - 1)} \]
   \[ A = \frac{X - Y}{2(\psi - 1)} \]
   \[ B = s - A \]
   \[ C = m - A \]
   \[ D = f - m + A \]
   \[ W = \frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D} \]
   \[ F = (a - A - \frac{1}{2})^2 W - z^2_{\alpha/2} \]
   \[ T = \frac{1}{2(\psi - 1)^2} \left( Y - n - \frac{\psi - 1}{Y} [X(m+s) - 2ms(2\psi - 1)] \right) \]
   \[ U = \frac{1}{B^2} + \frac{1}{C^2} - \frac{1}{A^2} - \frac{1}{D^2} \]
   \[ V = T \left( [a - A - \frac{1}{2}]^2 U - 2W (a - A - \frac{1}{2}) \right) \]

   Finally, use the updating equation below to calculate a new value for the odds ratio using the updating equation

   \[ \psi^{(k+1)} = \psi^{(k)} \frac{F}{V} \]

2. Continue iterating until the value of $F$ is arbitrarily close to zero.

The upper limit is found by substituting $+\frac{1}{2}$ for $-\frac{1}{2}$ in the formulas for $F$ and $V$.

Confidence limits for the relative risk can be calculated using the expected counts $A, B, C,$ and $D$ from the last iteration of the above procedure. The lower limit of the relative risk

\[ \phi_{lower} = \frac{A_{lower}n}{B_{lower}m} \]

\[ \phi_{upper} = \frac{A_{upper}n}{B_{upper}m} \]
Confidence Intervals for the Odds Ratio

Many methods have been devised for computing confidence intervals for the odds ratio of two proportions

\[ \psi = \frac{p_1}{1 - p_1} \cdot \frac{p_2}{1 - p_2} \]

Eight of these methods are available in the Confidence Intervals for Two Proportions [Odds Ratios] procedure. The eight confidence interval methods are

1. Exact (Conditional)
2. Score (Farrington and Manning)
3. Score (Miettinen and Nurminen)
4. Fleiss
5. Logarithm
6. Mantel-Haenszel
7. Simple
8. Simple + 1/2

C.I. for Odds Ratio: Conditional Exact

The conditional exact confidence interval of the odds ratio is calculated using the noncentral hypergeometric distribution as given in Sahai and Khurshid (1995). That is, a \( 100(1 - \alpha)\% \) confidence interval is found by searching for \( \psi_L \) and \( \psi_U \) such that

\[
\sum_{k=k_1}^{k_2} \binom{n_1}{k} \binom{n_2}{m_1-k} (\psi_L)^k \left( \binom{n_2}{m_1-k} (\psi_L)^k \right) = \frac{\alpha}{2}
\]

and

\[
\sum_{k=k_1}^{k_2} \binom{n_1}{k} \binom{n_2}{m_1-k} (\psi_U)^k \left( \binom{n_2}{m_1-k} (\psi_U)^k \right) = \frac{\alpha}{2}
\]

where

\[ k_1 = \max(0, m_1 - n_1) \quad \text{and} \quad k_2 = \min(n_1, m_1) \]
Confidence Intervals for Two Proportions

Farrington and Manning’s Test of the Odds Ratio

Farrington and Manning (1990) developed a test statistic similar to that of Miettinen and Nurminen but with the factor \( \frac{N}{N-1} \) removed.

The formula for computing this test statistic is

\[
z_{FMO} = \frac{(\hat{p}_1 - \hat{p}_1) - (\hat{p}_2 - \hat{p}_2)}{\frac{\hat{p}_1 \hat{q}_1}{\hat{p}_2 \hat{q}_2}} \sqrt{\frac{1}{N_2 \hat{p}_1 \hat{q}_1} + \frac{1}{N_2 \hat{p}_2 \hat{q}_2}}
\]

where the estimates \( \hat{p}_1 \) and \( \hat{p}_2 \) are computed as in the corresponding test of Miettinen and Nurminen (1985) as

\[
\hat{p}_1 = \frac{\tilde{p}_1 \psi_0}{1 + \tilde{p}_2 (\psi_0 - 1)}
\]

\[
\hat{p}_2 = -B + \sqrt{B^2 - 4AC}
\]

\[
A = N_2 (\psi_0 - 1)
\]

\[
B = N_1 \psi_0 + N_2 - M_1 (\psi_0 - 1)
\]

\[
C = -M_1
\]

Farrington and Manning (1990) proposed inverting their score test to find the confidence interval. The lower limit is found by solving

\[
z_{FMO} = z_{\alpha/2}
\]

and the upper limit is the solution of

\[
z_{FMO} = -z_{\alpha/2}
\]

C.I. for Odds Ratio: Miettinen and Nurminen

Miettinen and Nurminen (1985) proposed a test statistic for testing whether the odds ratio is equal to a specified value \( \psi_0 \). Because the approach they used with the difference and ratio does not easily extend to the odds ratio, they used a score statistic approach for the odds ratio. The regular MLE’s are \( \hat{p}_1 \) and \( \hat{p}_2 \). The constrained MLE’s are \( \tilde{p}_1 \) and \( \tilde{p}_2 \). These estimates are constrained so that \( \tilde{\psi} = \psi_0 \). A correction factor of \( \frac{N}{N-1} \) 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 the test statistic is

\[
z_{MNO} = \frac{(\hat{p}_1 - \tilde{p}_1) - (\hat{p}_2 - \tilde{p}_2)}{\frac{\hat{p}_1 \hat{q}_1}{\tilde{p}_1 \hat{q}_1} + \frac{1}{\tilde{p}_2 \hat{q}_2}} \sqrt{\frac{N}{N_2 \tilde{p}_1 \hat{q}_1} + \frac{1}{N_2 \tilde{p}_2 \hat{q}_2}}
\]

where

\[
\tilde{p}_1 = \frac{\tilde{p}_2 \psi_0}{1 + \tilde{p}_2 (\psi_0 - 1)}
\]
Confidence Intervals for Two Proportions

\[
\tilde{p}_2 = -B + \sqrt{B^2 - 4AC} \\
A = N_2(\psi_0 - 1) \\
B = N_1\psi_0 + N_2 - M_1(\psi_0 - 1) \\
C = -M_1
\]

Miettinen and Nurminen (1985) proposed inverting their score test to find the confidence interval. The lower limit is found by solving

\[
z_{MNO} = z_{a/2}
\]

and the upper limit is the solution of

\[
z_{MNO} = -z_{a/2}
\]

C.I. for Odds Ratio: Iterated Method of Fleiss

Fleiss (1981) presents an improve confidence interval for the odds ratio. This method forms the confidence interval as all those value of the odds ratio which would not be rejected by a chi-square hypothesis test. Fleiss gives the following details about how to construct this confidence interval. To compute the lower limit, do the following.

1. For a trial value of \( \psi \), compute the quantities \( X, Y, W, F, U, \) and \( V \) using the formulas

\[
X = \psi(m + s) + (n - s) \\
Y = \sqrt{X^2 - 4ms(\psi - 1)} \\
A = \frac{X - Y}{2(\psi - 1)} \\
B = s - A \\
C = m - A \\
D = f - m + A \\
W = \frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D} \\
F = (a - A - \frac{1}{2})^2W - z_{a/2}^2 \\
T = \frac{1}{2(\psi - 1)^2}\left(Y - n + \frac{\psi - 1}{Y} [X(m + s) - 2ms(2\psi - 1)]\right) \\
U = \frac{1}{B^2} + \frac{1}{C^2} - \frac{1}{A^2} - \frac{1}{D^2} \\
V = T [(a - A - \frac{1}{2})^2U - 2W(a - A - \frac{1}{2})]
\]
Confidence Intervals for Two Proportions

Finally, use the updating equation below to calculate a new value for the odds ratio using the updating equation

\[ \psi^{(k+1)} = \psi^{(k)} - \frac{F}{V} \]

2. Continue iterating until the value of \( F \) is arbitrarily close to zero.

The upper limit is found by substituting \( + \frac{1}{2} \) for \( - \frac{1}{2} \) in the formulas for \( F \) and \( V \).

Confidence limits for the relative risk can be calculated using the expected counts \( A, B, C, \) and \( D \) from the last iteration of the above procedure. The lower limit of the relative risk

\[ \phi_{lower} = \frac{A_{lower} n}{B_{lower} m} \]

\[ \phi_{upper} = \frac{A_{upper} n}{B_{upper} m} \]

C.I. for Odds Ratio: Mantel-Haenszel

The common estimate of the logarithm of the odds ratio is used to create this estimator. That is

\[ \ln(\hat{\psi}) = \ln\left(\frac{ad}{bc}\right) \]

The standard error of this estimator is estimated using the Robins, Breslow, Greenland (1986) estimator which performs well in most situations. The standard error is given by

\[ \text{se}(\ln(\hat{\psi})) = \sqrt{\frac{A}{2C} + \frac{AD + BC}{2CD} + \frac{B}{2D}} \]

where

\[ A = \frac{a + d}{N} \]
\[ B = \frac{b + c}{N} \]
\[ C = \frac{ad}{N} \]
\[ D = \frac{bc}{N} \]

The confidence limits are calculated as

\[ \hat{\psi}_{lower} = \exp(\ln(\hat{\psi}) - z_{1-\alpha/2} \text{se}(\ln(\hat{\psi}))) \]
\[ \hat{\psi}_{upper} = \exp(\ln(\hat{\psi}) + z_{1-\alpha/2} \text{se}(\ln(\hat{\psi}))) \]
C.I. for Odds Ratio: Simple, Simple + \( \frac{1}{2} \), and Logarithm

The simple estimate of the odds ratio uses the formula

\[
\hat{\psi} = \frac{\hat{p}_1\hat{q}_2}{\hat{p}_2\hat{q}_1} = \frac{ad}{bc}
\]

The standard error of this estimator is estimated by

\[
se(\hat{\psi}) = \hat{\psi}\sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}
\]

Problems occur if any one of the quantities \( a, b, c, \) or \( d \) are zero. To correct this problem, many authors recommend adding one-half to each cell count so that a zero cannot occur. Now, the formulas become

\[
\hat{\psi}' = \frac{(a + 0.5)(d + 0.5)}{(b + 0.5)(c + 0.5)}
\]

and

\[
se(\hat{\psi}') = \hat{\psi}'\sqrt{\frac{1}{a + 0.5} + \frac{1}{b + 0.5} + \frac{1}{c + 0.5} + \frac{1}{d + 0.5}}
\]

The distribution of these direct estimates of the odds ratio do not converge to normality as fast as does their logarithm, so the logarithm of the odds ratio is used to form confidence intervals. The formula for the standard error of the log odds ratio is

\[
L' = \ln(\hat{\psi}')
\]

and

\[
se(L') = \frac{1}{\sqrt{a + 0.5}} + \frac{1}{\sqrt{b + 0.5}} + \frac{1}{\sqrt{c + 0.5}} + \frac{1}{\sqrt{d + 0.5}}
\]

A \( 100(1 - \alpha)\% \) confidence interval for the log odds ratio is formed using the standard normal distribution as follows

\[
\hat{\psi}_{\text{lower}} = \exp(L' - z_{1-\alpha/2}se(L'))
\]

\[
\hat{\psi}_{\text{upper}} = \exp(L' + z_{1-\alpha/2}se(L'))
\]


Confidence Level

The confidence level, \( 1 - \alpha \), has the following interpretation. If thousands of random samples of size \( n_1 \) and \( n_2 \) are drawn from populations 1 and 2, respectively, and a confidence interval for the true difference/ratio/odds ratio of proportions is calculated for each pair of samples, the proportion of those intervals that will include the true difference/ratio/odds ratio of proportions is \( 1 - \alpha \).
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 (Common Options)

This chapter covers four procedures, each of which has different options. This section documents options that are common to all four procedures. Following this section, the unique options for each procedure (proportions, differences, ratios, and odds ratios) will be documented.

Solve For

Solve For

This option specifies the parameter to be solved for from the other parameters.

One-Sided or Two-Sided Interval

Interval Type

Specify whether the interval to be used will be a two-sided confidence interval, an interval that has only an upper limit, or an interval that has only a lower limit.

Confidence

Confidence Level

The confidence level, $1 - \alpha$, has the following interpretation. If thousands of random samples of size $n_1$ and $n_2$ are drawn from populations 1 and 2, respectively, and a confidence interval for the true difference/ratio/odds ratio of proportions is calculated for each pair of samples, the proportion of those intervals that will include the true difference/ratio/odds ratio of proportions is $1 - \alpha$.

Often, the values 0.95 or 0.99 are used. You can enter single values or a range of values such as $0.90$, $0.95$ or $0.90$ to $0.99$ by $0.01$.

Sample Size (When Solving for Sample Size)

Group Allocation

Select the option that describes the constraints on $N_1$ or $N_2$ or both.

The options are

- **Equal ($N_1 = N_2$)**
  
  This selection is used when you wish to have equal sample sizes in each group. Since you are solving for both sample sizes at once, no additional sample size parameters need to be entered.

- **Enter $N_1$, solve for $N_2$**
  
  Select this option when you wish to fix $N_1$ at some value (or values), and then solve only for $N_2$. Please note that for some values of $N_1$, there may not be a value of $N_2$ that is large enough to obtain the desired power.

- **Enter $N_2$, solve for $N_1$**
  
  Select this option when you wish to fix $N_2$ at some value (or values), and then solve only for $N_1$. Please note that for some values of $N_2$, there may not be a value of $N_1$ that is large enough to obtain the desired power.
Confidence Intervals for Two Proportions

- Enter $R = \frac{N_2}{N_1}$, solve for $N_1$ and $N_2$
  
  For this choice, you set a value for the ratio of $N_2$ to $N_1$, and then PASS determines the needed $N_1$ and $N_2$, with this ratio, to obtain the desired power. An equivalent representation of the ratio, $R$, is
  
  $$N_2 = R \times N_1.$$  

- Enter percentage in Group 1, solve for $N_1$ and $N_2$
  
  For this choice, you set a value for the percentage of the total sample size that is in Group 1, and then PASS determines the needed $N_1$ and $N_2$ with this percentage to obtain the desired power.

**N1 (Sample Size, Group 1)**

*this option is displayed if Group Allocation = “Enter $N_1$, solve for $N_2$”*

$N_1$ is the number of items or individuals sampled from the Group 1 population.

$N_1$ must be $\geq 2$. You can enter a single value or a series of values.

**N2 (Sample Size, Group 2)**

*this option is displayed if Group Allocation = “Enter $N_2$, solve for $N_1$”*

$N_2$ is the number of items or individuals sampled from the Group 2 population.

$N_2$ must be $\geq 2$. You can enter a single value or a series of values.

**R (Group Sample Size Ratio)**

*this option is displayed only if Group Allocation = “Enter $R = \frac{N_2}{N_1}$, solve for $N_1$ and $N_2$.”*

$R$ is the ratio of $N_2$ to $N_1$. That is,

$$R = \frac{N_2}{N_1}.$$  

Use this value to fix the ratio of $N_2$ to $N_1$ while solving for $N_1$ and $N_2$. Only sample size combinations with this ratio are considered.

$N_2$ is related to $N_1$ by the formula:

$$N_2 = \lceil R \times N_1 \rceil,$$

where the value $\lceil Y \rceil$ is the next integer $\geq Y$.

For example, setting $R = 2.0$ results in a Group 2 sample size that is double the sample size in Group 1 (e.g., $N_1 = 10$ and $N_2 = 20$, or $N_1 = 50$ and $N_2 = 100$).

$R$ must be greater than 0. If $R < 1$, then $N_2$ will be less than $N_1$; if $R > 1$, then $N_2$ will be greater than $N_1$. You can enter a single or a series of values.

**Percent in Group 1**

*this option is displayed only if Group Allocation = “Enter percentage in Group 1, solve for $N_1$ and $N_2$.”*

Use this value to fix the percentage of the total sample size allocated to Group 1 while solving for $N_1$ and $N_2$. Only sample size combinations with this Group 1 percentage are considered. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single or a series of values.
Sample Size (When Not Solving for Sample Size)

Group Allocation
Select the option that describes how individuals in the study will be allocated to Group 1 and to Group 2.

The options are:

- **Equal (N1 = N2)**
  This selection is used when you wish to have equal sample sizes in each group. A single per group sample size will be entered.

- **Enter N1 and N2 individually**
  This choice permits you to enter different values for N1 and N2.

- **Enter N1 and R, where N2 = R * N1**
  Choose this option to specify a value (or values) for N1, and obtain N2 as a ratio (multiple) of N1.

- **Enter total sample size and percentage in Group 1**
  Choose this option to specify a value (or values) for the total sample size (N), obtain N1 as a percentage of N, and then N2 as N - N1.

Sample Size Per Group
This option is displayed only if Group Allocation = “Equal (N1 = N2).”

The Sample Size Per Group is the number of items or individuals sampled from each of the Group 1 and Group 2 populations. Since the sample sizes are the same in each group, this value is the value for N1, and also the value for N2.

The Sample Size Per Group must be ≥ 2. You can enter a single value or a series of values.

**N1 (Sample Size, Group 1)**
This option is displayed if Group Allocation = “Enter N1 and N2 individually” or “Enter N1 and R, where N2 = R * N1.”

N1 is the number of items or individuals sampled from the Group 1 population.

N1 must be ≥ 2. You can enter a single value or a series of values.

**N2 (Sample Size, Group 2)**
This option is displayed only if Group Allocation = “Enter N1 and N2 individually.”

N2 is the number of items or individuals sampled from the Group 2 population.

N2 must be ≥ 2. You can enter a single value or a series of values.

**R (Group Sample Size Ratio)**
This option is displayed only if Group Allocation = “Enter N1 and R, where N2 = R * N1.”

R is the ratio of N2 to N1. That is,

\[ R = \frac{N2}{N1} \]

Use this value to obtain N2 as a multiple (or proportion) of N1.

N2 is calculated from N1 using the formula:

\[ N2 = \lceil R \times N1 \rceil, \]

where the value \( \lceil Y \rceil \) is the next integer ≥ Y.
For example, setting $R = 2.0$ results in a Group 2 sample size that is double the sample size in Group 1. $R$ must be greater than 0. If $R < 1$, then $N2$ will be less than $N1$; if $R > 1$, then $N2$ will be greater than $N1$. You can enter a single value or a series of values.

**Total Sample Size (N)**
This option is displayed only if Group Allocation = “Enter total sample size and percentage in Group 1.”
This is the total sample size, or the sum of the two group sample sizes. This value, along with the percentage of the total sample size in Group 1, implicitly defines $N1$ and $N2$.
The total sample size must be greater than one, but practically, must be greater than 3, since each group sample size needs to be at least 2.
You can enter a single value or a series of values.

**Percent in Group 1**
This option is displayed only if Group Allocation = “Enter total sample size and percentage in Group 1.”
This value fixes the percentage of the total sample size allocated to Group 1. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.
The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single value or a series of values.

---

**Design Tab (Proportions)**
This section documents options that are used when the parameterization is in terms of the values of the two sample proportions, $P1$ and $P2$. The corresponding procedure is Confidence Intervals for the Difference between Two Proportions using Proportions.

---

**Confidence Interval Method**

**Confidence Interval Formula**
Specify the formula to be in used in calculation of confidence intervals.

- **Score (Farrington & Manning)**
  This formula is based on inverting Farrington and Manning's score test.

- **Score (Miettinen & Nurminen)**
  This formula is based on inverting Miettinen and Nurminen's score test.

- **Score w/ Skewness (Gart & Nam)**
  This formula is based on inverting Gart and Nam's score test, with a correction for skewness.

- **Score (Wilson)**
  This formula is based on the Wilson score method for a single proportion, without continuity correction.

- **Score (Wilson C.C.)**
  This formula is based on the Wilson score method for a single proportion, with continuity correction.

- **Chi-Square C.C. (Yates)**
  This is the commonly used simple asymptotic method, with continuity correction.
Chi-Square (Pearson)
This is the commonly used simple asymptotic method, without continuity correction.

**Precision**

**Confidence Interval Width (Two-Sided)**
This is the distance from the lower confidence limit to the upper confidence limit.
You can enter a single value or a list of values. The value(s) must be greater than zero.

**Distance from Diff to Limit (One-Sided)**
This is the distance from the difference in sample proportions to the lower or upper limit of the confidence interval, depending on whether the Interval Type is set to Lower Limit or Upper Limit.
You can enter a single value or a list of values. The value(s) must be greater than zero.

**Proportions (Difference = P1 – P2)**

**P1 (Proportion Group 1)**
Enter an estimate of the proportion for group 1. The sample size and width calculations assume that the value entered here is the proportion estimate that is obtained from the sample. If the sample proportion is different from the one specified here, the width may be narrower or wider than specified.
The value(s) must be between 0.0001 and 0.9999.
You can enter a range of values such as .1 .2 .3 or .1 to .5 by .1.

**P2 (Proportion Group 2)**
Enter an estimate of the proportion for group 2. The sample size and width calculations assume that the value entered here is the proportion estimate that is obtained from the sample. If the sample proportion is different from the one specified here, the width may be narrower or wider than specified.
The value(s) must be between 0.0001 and 0.9999.
You can enter a range of values such as .1 .2 .3 or .1 to .5 by .1.

**Design Tab (Differences)**
This section documents options that are used when the parameterization is in terms of the difference in sample proportions and the value of the second sample proportion, P2. The corresponding procedure is Confidence Intervals for the Difference between Two Proportions using Differences.

**Confidence Interval Method**

**Confidence Interval Formula**
Specify the formula to be in used in calculation of confidence intervals.

- **Score (Farrington & Manning)**
  This formula is based on inverting Farrington and Manning's score test.

- **Score (Miettinen & Nurminen)**
  This formula is based on inverting Miettinen and Nurminen's score test.
Confidence Intervals for Two Proportions

- **Score w/Skewness (Gart & Nam)**
  This formula is based on inverting Gart and Nam's score test, with a correction for skewness.

- **Score (Wilson)**
  This formula is based on the Wilson score method for a single proportion, without continuity correction.

- **Score (Wilson C.C.)**
  This formula is based on the Wilson score method for a single proportion, with continuity correction.

- **Chi-Square C.C. (Yates)**
  This is the commonly used simple asymptotic method, with continuity correction.

- **Chi-Square (Pearson)**
  This is the commonly used simple asymptotic method, without continuity correction.

---

**Precision**

**Confidence Interval Width (Two-Sided)**
This is the distance from the lower confidence limit to the upper confidence limit.
You can enter a single value or a list of values. The value(s) must be greater than zero.

**Distance from Diff to Limit (One-Sided)**
This is the distance from the difference in sample proportions to the lower or upper limit of the confidence interval, depending on whether the Interval Type is set to Lower Limit or Upper Limit.
You can enter a single value or a list of values. The value(s) must be greater than zero.

---

**Proportions (Difference = P1 – P2)**

**Difference in Sample Proportions**
Enter an estimate of the difference between sample proportion 1 and sample proportion 2. The sample size and width calculations assume that the value entered here is the difference estimate that is obtained from the sample. If the sample difference is different from the one specified here, the width may be narrower or wider than specified.

The value(s) must be between -1 and 1, and such that P1 = Difference + P2 is between 0.0001 and 0.9999.
You can enter a range of values such as .1 .2 .3 or .1 to .5 by .1.

**P2 (Proportion Group 2)**
Enter an estimate of the proportion for group 2. The sample size and width calculations assume that the value entered here is the proportion estimate that is obtained from the sample. If the sample proportion is different from the one specified here, the width may be narrower or wider than specified.

The value(s) must be between 0.0001 and 0.9999.
You can enter a range of values such as .1 .2 .3 or .1 to .5 by .1.

---

**Design Tab (Ratios)**
This section documents options that are used when the parameterization is in terms of the ratio of sample proportions and the value of the second sample proportion, P2. The corresponding procedure is Confidence Intervals for the Difference between Two Proportions using Ratios.
Confidence Interval Method

Confidence Interval Formula
Specify the formula to be used in calculation of confidence intervals.

- **Score (Farrington & Manning)**
  This formula is based on inverting Farrington and Manning's score test.

- **Score (Miettinen & Nurminen)**
  This formula is based on inverting Miettinen and Nurminen's score test.

- **Score w/ Skewness (Gart & Nam)**
  This formula is based on inverting Gart and Nam's score test, with a correction for skewness.

- **Logarithm (Katz)**
  This formula is based on the asymptotic normality of log(P1/P2).

- **Logarithm + 1/2 (Walter)**
  This formula is based on the asymptotic normality of log(P1/P2), but 1/2 is used as an adjustment.

- **Fleiss**
  This is an iterative method that was developed for the odds ratio and adapted to the proportion ratio.

Precision

Confidence Interval Width (Two-Sided)
This is the distance from the lower confidence limit to the upper confidence limit.
You can enter a single value or a list of values. The value(s) must be greater than zero.

Distance from Ratio to Limit (One-Sided)
This is the distance from the ratio of sample proportions to the lower or upper limit of the confidence interval, depending on whether the Interval Type is set to Lower Limit or Upper Limit.
You can enter a single value or a list of values. The value(s) must be greater than zero.

Proportions (Ratio = P1/P2)

Ratio of Sample Proportions
Enter an estimate of the ratio of sample proportion 1 to sample proportion 2. The sample size and width calculations assume that the value entered here is the ratio estimate that is obtained from the samples. If the sample ratio is different from the one specified here, the width may be narrower or wider than specified.
The value(s) must be greater than 0, and such that P1 = Ratio * P2 is between 0.0001 and 0.9999.
You can enter a range of values such as .7 .8 .9 or .5 to .9 by .1.

P2 (Proportion Group 2)
Enter an estimate of the proportion for group 2. The sample size and width calculations assume that the value entered here is the proportion estimate that is obtained from the sample. If the sample proportion is different from the one specified here, the width may be narrower or wider than specified.
The value(s) must be between 0.0001 and 0.9999.
Confidence intervals for two proportions

You can enter a range of values such as .1 .2 .3 or .1 to .5 by .1.

Design Tab (Odds Ratios)

This section documents options that are used when the parameterization is in terms of the odds ratio and the value of the second sample proportion, P2. The corresponding procedure is Confidence Intervals for the Difference between Two Proportions using Odds Ratios.

Confidence Interval Method

Confidence Interval Formula

Specify the formula to be used in calculation of confidence intervals.

- **Exact (Conditional)**
  This conditional exact confidence interval formula is calculated using the non-central hypergeometric distribution.

- **Score (Farrington & Manning)**
  This formula is based on inverting Farrington and Manning's score test.

- **Score (Miettinen & Nurminen)**
  This formula is based on inverting Miettinen and Nurminen's score test.

- **Fleiss**
  This iterative method forms the confidence interval as all those value of the odds ratio which would not be rejected by a chi-square hypothesis test.

- **Logarithm**
  This formula is similar to SIMPLE + 1/2, but with the logarithm of the odds ratio.

- **Mantel-Haenszel**
  This formula is based on the Mantel-Haenszel formula for the odds ratio.

- **Simple**
  This uses the simple odds ratio formula and large sample standard error estimate.

- **Simple + 1/2**
  This uses the simple odds ratio formula and large sample standard error estimate, but with 1/2 added to frequencies as a bias reduction device.

Precision

Confidence Interval Width (Two-Sided)

This is the distance from the lower confidence limit to the upper confidence limit.

You can enter a single value or a list of values. The value(s) must be greater than zero.

Distance from OR to Limit (One-Sided)

This is the distance from the odds ratio to the lower or upper limit of the confidence interval, depending on whether the Interval Type is set to Lower Limit or Upper Limit.

You can enter a single value or a list of values. The value(s) must be greater than zero.
Proportions (OR = O1/O2)

Odds Ratio
Enter an estimate of the sample odds ratio (O1/O2). The sample size and width calculations assume that the value entered here is the odds ratio estimate that is obtained from the samples. If the sample odds ratio is different from the one specified here, the width may be narrower or wider than specified. The value(s) must be greater than 0.

You can enter a range of values such as .7 .8 .9 or .5 to .9 by .1.

P2 (Proportion Group 2)
Enter an estimate of the proportion for group 2. The sample size and width calculations assume that the value entered here is the proportion estimate that is obtained from the sample. If the sample proportion is different from the one specified here, the width may be narrower or wider than specified. The value(s) must be between 0.0001 and 0.9999.

You can enter a range of values such as .1 .2 .3 or .1 to .5 by .1.

Example 1 – Calculating Sample Size using Proportions
Suppose a study is planned in which the researcher wishes to construct a two-sided 95% confidence interval for the difference in proportions such that the width of the interval is no wider than 0.1. The confidence interval method to be used is the Yates chi-square simple asymptotic method with continuity correction. The confidence level is set at 0.95, but 0.99 is included for comparative purposes. The proportion estimates to be used are 0.6 for Group 1, and 0.4 for Group 2. Instead of examining only the interval width of 0.1, a series of widths from 0.05 to 0.3 will also be considered.

The goal is to determine the necessary sample size.

Setup
This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Confidence Intervals for Two Proportions using Proportions procedure window by expanding Proportions, then Two Independent Proportions, then clicking on Confidence Interval, and then clicking on Confidence Intervals for Two Proportions using Proportions. 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.

<table>
<thead>
<tr>
<th>Option</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design Tab</td>
<td></td>
</tr>
<tr>
<td>Solve For</td>
<td>Sample Size</td>
</tr>
<tr>
<td>Confidence Interval Formula</td>
<td>Chi-Square C.C. (Yates)</td>
</tr>
<tr>
<td>Interval Type</td>
<td>Two-Sided</td>
</tr>
<tr>
<td>Confidence Level</td>
<td>0.95 0.99</td>
</tr>
<tr>
<td>Group Allocation</td>
<td>Equal (N1 = N2)</td>
</tr>
<tr>
<td>Confidence Interval Width</td>
<td>0.05 to 0.30 by 0.05</td>
</tr>
<tr>
<td>P1</td>
<td>0.6</td>
</tr>
<tr>
<td>P2</td>
<td>0.4</td>
</tr>
</tbody>
</table>
**Annotated Output**

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

**Numeric Results**

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Target Width</th>
<th>Actual Width</th>
<th>P1</th>
<th>P2</th>
<th>P1 - P2</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950</td>
<td>3030</td>
<td>3030</td>
<td>6060</td>
<td>0.050</td>
<td>0.050</td>
<td>0.60</td>
<td>0.40</td>
<td>0.20</td>
<td>0.18</td>
<td>0.22</td>
</tr>
<tr>
<td>0.950</td>
<td>778</td>
<td>778</td>
<td>1556</td>
<td>0.100</td>
<td>0.100</td>
<td>0.60</td>
<td>0.40</td>
<td>0.20</td>
<td>0.15</td>
<td>0.25</td>
</tr>
<tr>
<td>0.950</td>
<td>354</td>
<td>354</td>
<td>708</td>
<td>0.150</td>
<td>0.150</td>
<td>0.60</td>
<td>0.40</td>
<td>0.20</td>
<td>0.13</td>
<td>0.27</td>
</tr>
<tr>
<td>0.950</td>
<td>204</td>
<td>204</td>
<td>408</td>
<td>0.200</td>
<td>0.200</td>
<td>0.60</td>
<td>0.40</td>
<td>0.20</td>
<td>0.10</td>
<td>0.30</td>
</tr>
<tr>
<td>0.950</td>
<td>134</td>
<td>134</td>
<td>268</td>
<td>0.250</td>
<td>0.250</td>
<td>0.60</td>
<td>0.40</td>
<td>0.20</td>
<td>0.08</td>
<td>0.22</td>
</tr>
<tr>
<td>0.950</td>
<td>95</td>
<td>95</td>
<td>190</td>
<td>0.300</td>
<td>0.300</td>
<td>0.60</td>
<td>0.40</td>
<td>0.20</td>
<td>0.05</td>
<td>0.35</td>
</tr>
<tr>
<td>0.990</td>
<td>5176</td>
<td>5176</td>
<td>10352</td>
<td>0.050</td>
<td>0.050</td>
<td>0.60</td>
<td>0.40</td>
<td>0.20</td>
<td>0.18</td>
<td>0.22</td>
</tr>
<tr>
<td>0.990</td>
<td>1314</td>
<td>1314</td>
<td>2628</td>
<td>0.100</td>
<td>0.100</td>
<td>0.60</td>
<td>0.40</td>
<td>0.20</td>
<td>0.15</td>
<td>0.25</td>
</tr>
<tr>
<td>0.990</td>
<td>593</td>
<td>593</td>
<td>1186</td>
<td>0.150</td>
<td>0.150</td>
<td>0.60</td>
<td>0.40</td>
<td>0.20</td>
<td>0.13</td>
<td>0.27</td>
</tr>
<tr>
<td>0.990</td>
<td>339</td>
<td>339</td>
<td>678</td>
<td>0.200</td>
<td>0.200</td>
<td>0.60</td>
<td>0.40</td>
<td>0.20</td>
<td>0.10</td>
<td>0.30</td>
</tr>
<tr>
<td>0.990</td>
<td>220</td>
<td>220</td>
<td>440</td>
<td>0.250</td>
<td>0.250</td>
<td>0.60</td>
<td>0.40</td>
<td>0.20</td>
<td>0.08</td>
<td>0.22</td>
</tr>
<tr>
<td>0.990</td>
<td>155</td>
<td>155</td>
<td>310</td>
<td>0.300</td>
<td>0.300</td>
<td>0.60</td>
<td>0.40</td>
<td>0.20</td>
<td>0.05</td>
<td>0.35</td>
</tr>
</tbody>
</table>

**References**


**Report Definitions**

Confidence level is the proportion of confidence intervals (constructed with this same confidence level, sample size, etc.) that would contain the true difference in population proportions.

N1 and N2 are the number of items sampled from each population.

N is the total sample size, N1 + N2.

Target Width is the value of the width that is entered into the procedure.

Actual Width is the value of the width that is obtained from the procedure.

P1 and P2 are the assumed sample proportions for sample size calculations.

P1 - P2 is the difference between sample proportions at which sample size calculations are made.

Lower Limit and Upper Limit are the lower and upper limits of the confidence interval for the true difference in proportions (Population Proportion 1 - Population Proportion 2).

**Summary Statements**

Group sample sizes of 3030 and 3030 produce a two-sided 95% confidence interval for the difference in population proportions with a width that is equal to 0.050 when the estimated sample proportion 1 is 0.60 and the estimated sample proportion 2 is 0.40.

This report shows the calculated sample sizes for each of the scenarios.
These plots show the group sample size versus the confidence interval width for the two confidence levels.
Example 2 – Validation (Proportions and Differences) using Newcombe

Newcombe (1998b) page 877 gives an example of a calculation for a confidence interval for the difference in proportions when the confidence level is 95%, the sample proportions are 0.9 and 0.3, and the interval width is 0.6790 for the Chi-Square (Pearson) method, 0.8395 for the Chi-Square C.C. (Yates) method, 0.67064 for the Score (Miettinen and Nurminen) method, 0.6385 for the Score (Wilson) method, and 0.7374 for the Score C.C. (Wilson) method. The necessary sample size in each case is 10 per group.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Confidence Intervals for Two Proportions using Proportions procedure window by expanding Proportions, then Two Independent Proportions, then clicking on Confidence Interval, and then clicking on Confidence Intervals for Two Proportions using Proportions. You may then make the appropriate entries as listed below, or open Example 2 by going to the File menu and choosing Open Example Template.

Option

<table>
<thead>
<tr>
<th>Option</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design Tab</td>
<td></td>
</tr>
<tr>
<td>Solve For</td>
<td>Sample Size</td>
</tr>
<tr>
<td>Confidence Interval Formula</td>
<td>Varies [Chi-Square (Pearson), Chi-Square C.C. (Yates), Score</td>
</tr>
<tr>
<td></td>
<td>(Miettinen &amp; Nurminen), Score (Wilson), Score C.C. (Wilson)]</td>
</tr>
<tr>
<td>Interval Type</td>
<td>Two-Sided</td>
</tr>
<tr>
<td>Confidence Level</td>
<td>0.95</td>
</tr>
<tr>
<td>Group Allocation</td>
<td>Equal (N1 = N2)</td>
</tr>
<tr>
<td>Confidence Interval Width</td>
<td>Varies (0.6790, 0.8395, 0.67064, 0.6385, 0.7374)</td>
</tr>
<tr>
<td>P1</td>
<td>0.9</td>
</tr>
<tr>
<td>P2</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Output

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

Chi-Square (Pearson)

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Target Width</th>
<th>Actual Width</th>
<th>P1</th>
<th>P2</th>
<th>P1 - P2</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>0.6790</td>
<td>0.6790</td>
<td>0.9000</td>
<td>0.3000</td>
<td>0.6000</td>
<td>0.2605</td>
<td>0.9395</td>
</tr>
</tbody>
</table>

PASS also calculates the necessary sample size to be 10 per group.

Chi-Square C.C. (Yates)

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Target Width</th>
<th>Actual Width</th>
<th>P1</th>
<th>P2</th>
<th>P1 - P2</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>0.8395</td>
<td>0.8395</td>
<td>0.9000</td>
<td>0.3000</td>
<td>0.6000</td>
<td>0.1605</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

PASS also calculates the necessary sample size to be 10 per group.
### Confidence Intervals for Two Proportions

#### Score (Miettinen & Nurminen)

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Target Width</th>
<th>Actual Width</th>
<th>P1</th>
<th>P2</th>
<th>P1 - P2</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>0.6706</td>
<td>0.6706</td>
<td>0.9000</td>
<td>0.3000</td>
<td>0.6000</td>
<td>0.1700</td>
<td>0.8406</td>
</tr>
</tbody>
</table>

PASS also calculates the necessary sample size to be 10 per group.

#### Score (Wilson)

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Target Width</th>
<th>Actual Width</th>
<th>P1</th>
<th>P2</th>
<th>P1 - P2</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>0.6385</td>
<td>0.6385</td>
<td>0.9000</td>
<td>0.3000</td>
<td>0.6000</td>
<td>0.1705</td>
<td>0.8090</td>
</tr>
</tbody>
</table>

PASS also calculates the necessary sample size to be 10 per group.

#### Score C.C. (Wilson)

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Target Width</th>
<th>Actual Width</th>
<th>P1</th>
<th>P2</th>
<th>P1 - P2</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>0.7374</td>
<td>0.7374</td>
<td>0.9000</td>
<td>0.3000</td>
<td>0.6000</td>
<td>0.1013</td>
<td>0.8387</td>
</tr>
</tbody>
</table>

PASS also calculates the necessary sample size to be 10 per group.
Example 3 – Validation (Proportions and Differences) using Gart and Nam

Gart and Nam (1990) page 640 give an example of a calculation for a confidence interval for the difference in proportions when the confidence level is 95%, the sample proportions are 0.28 and 0.08, and the interval width is 0.4281 for the Score (Gart and Nam) method. The necessary sample size in each case is 25 per group.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Confidence Intervals for Two Proportions using Proportions procedure window by expanding Proportions, then Two Independent Proportions, then clicking on Confidence Interval, and then clicking on Confidence Intervals for Two Proportions using Proportions. You may then make the appropriate entries as listed below, or open Example 3 by going to the File menu and choosing Open Example Template.

Option | Value
--- | ---
Design Tab | Sample Size
Solve For | Sample Size
Confidence Interval Formula | Score w/Skewness (Gart & Nam)
Interval Type | Two-Sided
Confidence Level | 0.95
Group Allocation | Equal (N1 = N2)
Confidence Interval Width (Two-Sided) | 0.4281
P1 | 0.28
P2 | 0.08

Output

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

**Numeric Results**

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Target Width</th>
<th>Actual Width</th>
<th>P1</th>
<th>P2</th>
<th>P1 - P2</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>0.4281</td>
<td>0.4281</td>
<td>0.2800</td>
<td>0.0800</td>
<td>0.2000</td>
<td>-0.0143</td>
<td>0.4137</td>
</tr>
</tbody>
</table>

PASS also calculates the necessary sample size to be 25 per group.
Example 4 – Calculating Sample Size using Differences

Suppose a study is planned in which the researcher wishes to construct a two-sided 95% confidence interval for the difference in proportions such that the width of the interval is no wider than 0.1. The confidence interval method to be used is the Yates chi-square simple asymptotic method with continuity correction. The confidence level is set at 0.95, but 0.99 is included for comparative purposes. The difference estimate to be used is 0.05, and the estimate for proportion 2 is 0.3. Instead of examining only the interval width of 0.1, a series of widths from 0.05 to 0.3 will also be considered.

The goal is to determine the necessary sample size.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Confidence Intervals for Two Proportions using Differences procedure window by expanding Proportions, then Two Independent Proportions, then clicking on Confidence Interval, and then clicking on Confidence Intervals for Two Proportions using Differences. You may then make the appropriate entries as listed below, or open Example 4 by going to the File menu and choosing Open Example Template.

Option Value
---
Design Tab
Solve For ................................................ Sample Size
Confidence Interval Formula .................. Chi-Square C.C. (Yates)
Interval Type ........................................... Two-Sided
Confidence Level ................................. 0.95 0.99
Group Allocation ................................. Equal (N1 = N2)
Confidence Interval Width (Two-Sided) .. 0.05 to 0.30 by 0.05
Difference in Sample Proportions........... 0.05
P2........................................................... 0.3

Annotated Output

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

Numeric Results

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Target Width</th>
<th>Actual Width</th>
<th>P1</th>
<th>P2</th>
<th>P1 - P2</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950</td>
<td>2769</td>
<td>2769</td>
<td>5538</td>
<td>0.050</td>
<td>0.050</td>
<td>0.35</td>
<td>0.30</td>
<td>0.05</td>
<td>0.03</td>
<td>0.07</td>
</tr>
<tr>
<td>0.950</td>
<td>712</td>
<td>712</td>
<td>1424</td>
<td>0.100</td>
<td>0.100</td>
<td>0.35</td>
<td>0.30</td>
<td>0.05</td>
<td>0.00</td>
<td>0.10</td>
</tr>
<tr>
<td>0.950</td>
<td>325</td>
<td>325</td>
<td>650</td>
<td>0.150</td>
<td>0.150</td>
<td>0.35</td>
<td>0.30</td>
<td>0.05</td>
<td>-0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>0.950</td>
<td>188</td>
<td>188</td>
<td>376</td>
<td>0.200</td>
<td>0.200</td>
<td>0.35</td>
<td>0.30</td>
<td>0.05</td>
<td>-0.05</td>
<td>0.15</td>
</tr>
<tr>
<td>0.950</td>
<td>124</td>
<td>124</td>
<td>248</td>
<td>0.250</td>
<td>0.249</td>
<td>0.35</td>
<td>0.30</td>
<td>0.05</td>
<td>-0.07</td>
<td>0.17</td>
</tr>
<tr>
<td>0.950</td>
<td>88</td>
<td>88</td>
<td>176</td>
<td>0.300</td>
<td>0.299</td>
<td>0.35</td>
<td>0.30</td>
<td>0.05</td>
<td>-0.10</td>
<td>0.20</td>
</tr>
<tr>
<td>0.950</td>
<td>4725</td>
<td>4725</td>
<td>9450</td>
<td>0.050</td>
<td>0.050</td>
<td>0.35</td>
<td>0.30</td>
<td>0.05</td>
<td>0.03</td>
<td>0.07</td>
</tr>
<tr>
<td>0.950</td>
<td>1201</td>
<td>1201</td>
<td>2402</td>
<td>0.100</td>
<td>0.100</td>
<td>0.35</td>
<td>0.30</td>
<td>0.05</td>
<td>0.00</td>
<td>0.10</td>
</tr>
<tr>
<td>0.950</td>
<td>543</td>
<td>543</td>
<td>1086</td>
<td>0.150</td>
<td>0.150</td>
<td>0.35</td>
<td>0.30</td>
<td>0.05</td>
<td>-0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>0.990</td>
<td>310</td>
<td>310</td>
<td>620</td>
<td>0.200</td>
<td>0.200</td>
<td>0.35</td>
<td>0.30</td>
<td>0.05</td>
<td>-0.05</td>
<td>0.15</td>
</tr>
<tr>
<td>0.990</td>
<td>202</td>
<td>202</td>
<td>404</td>
<td>0.250</td>
<td>0.250</td>
<td>0.35</td>
<td>0.30</td>
<td>0.05</td>
<td>-0.07</td>
<td>0.17</td>
</tr>
<tr>
<td>0.990</td>
<td>143</td>
<td>143</td>
<td>286</td>
<td>0.300</td>
<td>0.299</td>
<td>0.35</td>
<td>0.30</td>
<td>0.05</td>
<td>-0.10</td>
<td>0.20</td>
</tr>
</tbody>
</table>
Confidence Intervals for Two Proportions

References

Report Definitions
Confidence level is the proportion of confidence intervals (constructed with this same confidence level, sample size, etc.) that would contain the true difference in population proportions.
N1 and N2 are the number of items sampled from each population.
N is the total sample size, N1 + N2.
Target Width is the value of the width that is entered into the procedure.
Actual Width is the value of the width that is obtained from the procedure.
P1 and P2 are the assumed sample proportions for sample size calculations.
P1 - P2 is the difference between sample proportions at which sample size calculations are made.
Lower Limit and Upper Limit are the lower and upper limits of the confidence interval for the true difference in proportions (Population Proportion 1 - Population Proportion 2).

Summary Statements
Group sample sizes of 2769 and 2769 produce a two-sided 95% confidence interval for the difference in population proportions with a width that is equal to 0.050 when the estimated sample proportion 1 is 0.35, the estimated sample proportion 2 is 0.30, and the difference in sample proportions is 0.05.

This report shows the calculated sample sizes for each of the scenarios.

Plots Section

These plots show the group sample size versus the confidence interval width for the two confidence levels.

Validation using Differences

The validation for the procedure Confidence Intervals for the Difference between Two Proportions using Differences is shown in Examples 2 and 3, which is the validation for the proportion specification.
Example 5 – Calculating Sample Size using Ratios

Suppose a study is planned in which the researcher wishes to construct a two-sided 95% confidence interval for the ratio of proportions such that the width of the interval is no wider than 0.2. The confidence interval method to be used is the Logarithm (Katz) method. The confidence level is set at 0.95, but 0.99 is included for comparative purposes. The ratio estimate to be used is 1.2, and the estimate for proportion 2 is 0.6. Instead of examining only the interval width of 0.2, a series of widths from 0.1 to 0.3 will also be considered.

The goal is to determine the necessary sample size.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Confidence Intervals for Two Proportions using Ratios procedure window by expanding Proportions, then Two Independent Proportions, then clicking on Confidence Interval, and then clicking on Confidence Intervals for Two Proportions using Ratios. You may then make the appropriate entries as listed below, or open Example 5 by going to the File menu and choosing Open Example Template.

<table>
<thead>
<tr>
<th>Option</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design Tab</td>
<td>Solve For: Sample Size</td>
</tr>
<tr>
<td>Confidence Interval Formula</td>
<td>Logarithm (Katz)</td>
</tr>
<tr>
<td>Interval Type</td>
<td>Two-Sided</td>
</tr>
<tr>
<td>Confidence Level</td>
<td>0.95 0.99</td>
</tr>
<tr>
<td>Group Allocation</td>
<td>Equal (N1 = N2)</td>
</tr>
<tr>
<td>Confidence Interval Width (Two-Sided)</td>
<td>0.10 to 0.30 by 0.05</td>
</tr>
<tr>
<td>Ratio of Sample Proportions</td>
<td>1.2</td>
</tr>
<tr>
<td>P2</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Annotated Output

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

Numeric Results

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Target Width</th>
<th>Actual Width</th>
<th>P1</th>
<th>P2</th>
<th>P1/P2</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950</td>
<td>2337</td>
<td>2337</td>
<td>4674</td>
<td>0.100</td>
<td>0.100</td>
<td>0.72</td>
<td>0.60</td>
<td>1.20</td>
<td>1.15</td>
<td>1.25</td>
</tr>
<tr>
<td>0.950</td>
<td>1040</td>
<td>1040</td>
<td>2080</td>
<td>0.150</td>
<td>0.150</td>
<td>0.72</td>
<td>0.60</td>
<td>1.20</td>
<td>1.13</td>
<td>1.28</td>
</tr>
<tr>
<td>0.950</td>
<td>586</td>
<td>586</td>
<td>1172</td>
<td>0.200</td>
<td>0.200</td>
<td>0.72</td>
<td>0.60</td>
<td>1.20</td>
<td>1.10</td>
<td>1.30</td>
</tr>
<tr>
<td>0.950</td>
<td>376</td>
<td>376</td>
<td>752</td>
<td>0.250</td>
<td>0.250</td>
<td>0.72</td>
<td>0.60</td>
<td>1.20</td>
<td>1.08</td>
<td>1.33</td>
</tr>
<tr>
<td>0.950</td>
<td>261</td>
<td>261</td>
<td>522</td>
<td>0.300</td>
<td>0.300</td>
<td>0.72</td>
<td>0.60</td>
<td>1.20</td>
<td>1.06</td>
<td>1.36</td>
</tr>
<tr>
<td>0.990</td>
<td>4037</td>
<td>4037</td>
<td>8074</td>
<td>0.100</td>
<td>0.100</td>
<td>0.72</td>
<td>0.60</td>
<td>1.20</td>
<td>1.15</td>
<td>1.25</td>
</tr>
<tr>
<td>0.990</td>
<td>1796</td>
<td>1796</td>
<td>3592</td>
<td>0.150</td>
<td>0.150</td>
<td>0.72</td>
<td>0.60</td>
<td>1.20</td>
<td>1.13</td>
<td>1.28</td>
</tr>
<tr>
<td>0.990</td>
<td>1011</td>
<td>1011</td>
<td>2022</td>
<td>0.200</td>
<td>0.200</td>
<td>0.72</td>
<td>0.60</td>
<td>1.20</td>
<td>1.10</td>
<td>1.30</td>
</tr>
<tr>
<td>0.990</td>
<td>648</td>
<td>648</td>
<td>1296</td>
<td>0.250</td>
<td>0.250</td>
<td>0.72</td>
<td>0.60</td>
<td>1.20</td>
<td>1.08</td>
<td>1.33</td>
</tr>
<tr>
<td>0.990</td>
<td>451</td>
<td>451</td>
<td>902</td>
<td>0.300</td>
<td>0.300</td>
<td>0.72</td>
<td>0.60</td>
<td>1.20</td>
<td>1.06</td>
<td>1.36</td>
</tr>
</tbody>
</table>
Confidence Intervals for Two Proportions

**References**

**Report Definitions**
Confidence level is the proportion of confidence intervals (constructed with this same confidence level, sample size, etc.) that would contain the true ratio of population proportions.
N1 and N2 are the number of items sampled from each population.
N is the total sample size, N1 + N2.
Target Width is the value of the width that is entered into the procedure.
Actual Width is the value of the width that is obtained from the procedure.
P1 and P2 are the assumed sample proportions for sample size calculations.
P1/P2 is the ratio of sample proportions at which sample size calculations are made.
Lower Limit and Upper Limit are the lower and upper limits of the confidence interval for the true ratio of proportions (Population Proportion 1 / Population Proportion 2).

**Summary Statements**
Group sample sizes of 2337 and 2337 produce a two-sided 95% confidence interval for the ratio of population proportions with a width that is equal to 0.100 when the estimated sample proportion 1 is 0.72, the estimated sample proportion 2 is 0.60, and the ratio of the sample proportions is 1.20.

This report shows the calculated sample sizes for each of the scenarios.

**Plots Section**

These plots show the group sample size versus the confidence interval width for the two confidence levels.
Example 6 – Validation (Ratios) using Gart and Nam

Gart and Nam (1988) page 331 give an example (Example 2) of a calculation for a confidence interval for the ratio of proportions when the confidence level is 95%, the sample proportion ratio is 2 and the sample proportion 2 is 0.3, the sample size for group 2 is 20, and the interval width is 3.437 for the Logarithm + 1/2 (Walter) method, 3.751 for the Score (Farrington and Manning) method, and 4.133 for the Score w/Skewness (Gart and Nam) method. The necessary sample size for group 1 in each case is 10.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Confidence Intervals for Two Proportions using Ratios procedure window by expanding Proportions, then Two Independent Proportions, then clicking on Confidence Interval, and then clicking on Confidence Intervals for Two Proportions using Ratios. You may then make the appropriate entries as listed below, or open Example 6 by going to the File menu and choosing Open Example Template.

Option | Value
--- | ---
Design Tab | Sample Size
Solve For | Sample Size
Confidence Interval Formula | Varies [Logarithm + 1/2 (Walter), Score (Farrington and Manning), Score w/Skewness (Gart and Nam)]
Interval Type | Two-Sided
Confidence Level | 0.95
Group Allocation | Enter N2, solve for N1
N2 | 20
Confidence Interval Width (Two-Sided) | Varies (3.437, 3.751, 4.133)
Ratio of Sample Proportions | 2
P2 | 0.3

Output

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

**Logarithm + 1/2 (Walter)**

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Target Width</th>
<th>Actual Width</th>
<th>P1</th>
<th>P2</th>
<th>P1/P2</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>3.437</td>
<td>3.431</td>
<td>0.60</td>
<td>0.30</td>
<td>2.00</td>
<td>0.88</td>
<td>4.31</td>
</tr>
</tbody>
</table>

PASS also calculates the necessary sample size for Group 1 to be 10.

**Score (Farrington and Manning)**

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Target Width</th>
<th>Actual Width</th>
<th>P1</th>
<th>P2</th>
<th>P1/P2</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>3.751</td>
<td>3.751</td>
<td>0.60</td>
<td>0.30</td>
<td>2.00</td>
<td>0.84</td>
<td>4.59</td>
</tr>
</tbody>
</table>

PASS also calculates the necessary sample size for Group 1 to be 10.
Score w/Skewness (Gart and Nam)

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Target Width</th>
<th>Actual Width</th>
<th>P1</th>
<th>P2</th>
<th>P1/P2</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>4.133</td>
<td>4.132</td>
<td>0.60</td>
<td>0.30</td>
<td>2.00</td>
<td>0.82</td>
<td>4.95</td>
</tr>
</tbody>
</table>

PASS also calculates the necessary sample size for Group 1 to be 10.

Example 7 – Validation (Ratios) using Katz et al

Katz et al (1978) pages 472-473 give an example of a calculation for a lower limit confidence interval for the ratio of proportions when the confidence level is 97.5%, the sample proportion ratio is 1.596078 and the sample proportion 2 is 0.153153, the sample size for group 2 is 111, and the distance from the ratio to the limit is 0.6223 for the Logarithm (Katz) method. The necessary sample size for group 1 is 225.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Confidence Intervals for Two Proportions using Ratios procedure window by expanding Proportions, then Two Independent Proportions, then clicking on Confidence Interval, and then clicking on Confidence Intervals for Two Proportions using Ratios. You may then make the appropriate entries as listed below, or open Example 7 by going to the File menu and choosing Open Example Template.

<table>
<thead>
<tr>
<th>Option</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design Tab</td>
<td></td>
</tr>
<tr>
<td>Solve For</td>
<td>Sample Size</td>
</tr>
<tr>
<td>Confidence Interval</td>
<td>Logarithm (Katz)</td>
</tr>
<tr>
<td>Interval Type</td>
<td>Lower Limit</td>
</tr>
<tr>
<td>Confidence Level</td>
<td>0.975</td>
</tr>
<tr>
<td>Group Allocation</td>
<td>Enter N2, solve for N1</td>
</tr>
<tr>
<td>N2</td>
<td>111</td>
</tr>
<tr>
<td>Distance to from Ratio</td>
<td>0.6223</td>
</tr>
<tr>
<td>to Limit</td>
<td></td>
</tr>
<tr>
<td>Ratio of Sample</td>
<td>1.596078</td>
</tr>
<tr>
<td>Proportions</td>
<td>0.153153</td>
</tr>
</tbody>
</table>

Output

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

Logarithm (Katz)

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Target Dist from Ratio to Limit</th>
<th>Actual Dist from Ratio to Limit</th>
<th>P1</th>
<th>P2</th>
<th>P1/P2</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.975</td>
<td>225</td>
<td>111</td>
<td>336</td>
<td>0.622</td>
<td>0.24</td>
<td>0.15</td>
<td>1.60</td>
<td>0.97</td>
<td>Inf</td>
<td></td>
</tr>
</tbody>
</table>

PASS also calculates the necessary sample size for group 1 to be 225.
Example 8 – Calculating Sample Size using Odds Ratios

Suppose a study is planned in which the researcher wishes to construct a two-sided 95% confidence interval for the odds ratio such that the width of the interval is no wider than 0.5. The confidence interval method to be used is the Logarithm method. The confidence level is set at 0.95, but 0.99 is included for comparative purposes. The odds ratio estimate to be used is 1.5, and the estimate for proportion 2 is 0.4. Instead of examining only the interval width of 0.5, a series of widths from 0.1 to 1.0 will also be considered.

The goal is to determine the necessary sample size.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Confidence Intervals for Two Proportions using Odds Ratios procedure window by expanding Proportions, then Two Independent Proportions, then clicking on Confidence Interval, and then clicking on Confidence Intervals for Two Proportions using Odds Ratios. You may then make the appropriate entries as listed below, or open Example 8 by going to the File menu and choosing Open Example Template.

Option | Value
--- | ---
Design Tab | 
Solve For | Sample Size
Confidence Interval Formula | Logarithm
Interval Type | Two-Sided
Confidence Level | 0.95 0.99
Group Allocation | Equal (N1 = N2)
Confidence Interval Width (Two-Sided) | 0.1 to 1.0 by 0.1
Odds Ratio | 1.5
P2 | 0.4

Annotated Output

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

Numeric Results

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Target Width</th>
<th>Actual Width</th>
<th>P1</th>
<th>P2</th>
<th>Odds Ratio O1/O2</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950</td>
<td>28244</td>
<td>28244</td>
<td>56488</td>
<td>0.100</td>
<td>0.100</td>
<td>0.50</td>
<td>0.40</td>
<td>1.50</td>
<td>1.45</td>
<td>1.55</td>
</tr>
<tr>
<td>0.950</td>
<td>7068</td>
<td>7068</td>
<td>14136</td>
<td>0.200</td>
<td>0.200</td>
<td>0.50</td>
<td>0.40</td>
<td>1.50</td>
<td>1.40</td>
<td>1.60</td>
</tr>
<tr>
<td>0.950</td>
<td>3146</td>
<td>3146</td>
<td>6292</td>
<td>0.300</td>
<td>0.300</td>
<td>0.50</td>
<td>0.40</td>
<td>1.50</td>
<td>1.36</td>
<td>1.66</td>
</tr>
<tr>
<td>0.950</td>
<td>1774</td>
<td>1774</td>
<td>3548</td>
<td>0.400</td>
<td>0.400</td>
<td>0.50</td>
<td>0.40</td>
<td>1.50</td>
<td>1.31</td>
<td>1.71</td>
</tr>
<tr>
<td>0.950</td>
<td>1138</td>
<td>1138</td>
<td>2276</td>
<td>0.500</td>
<td>0.500</td>
<td>0.50</td>
<td>0.40</td>
<td>1.50</td>
<td>1.27</td>
<td>1.77</td>
</tr>
<tr>
<td>0.950</td>
<td>793</td>
<td>793</td>
<td>1586</td>
<td>0.600</td>
<td>0.600</td>
<td>0.50</td>
<td>0.40</td>
<td>1.50</td>
<td>1.23</td>
<td>1.83</td>
</tr>
<tr>
<td>0.950</td>
<td>585</td>
<td>585</td>
<td>1170</td>
<td>0.700</td>
<td>0.700</td>
<td>0.50</td>
<td>0.40</td>
<td>1.50</td>
<td>1.19</td>
<td>1.89</td>
</tr>
<tr>
<td>0.950</td>
<td>450</td>
<td>450</td>
<td>900</td>
<td>0.800</td>
<td>0.800</td>
<td>0.50</td>
<td>0.40</td>
<td>1.50</td>
<td>1.15</td>
<td>1.95</td>
</tr>
<tr>
<td>0.950</td>
<td>358</td>
<td>358</td>
<td>716</td>
<td>0.900</td>
<td>0.900</td>
<td>0.50</td>
<td>0.40</td>
<td>1.50</td>
<td>1.11</td>
<td>2.01</td>
</tr>
<tr>
<td>0.950</td>
<td>291</td>
<td>291</td>
<td>582</td>
<td>1.000</td>
<td>1.000</td>
<td>0.50</td>
<td>0.40</td>
<td>1.50</td>
<td>1.08</td>
<td>2.08</td>
</tr>
<tr>
<td>0.990</td>
<td>48783</td>
<td>48783</td>
<td>97566</td>
<td>0.100</td>
<td>0.100</td>
<td>0.50</td>
<td>0.40</td>
<td>1.50</td>
<td>1.45</td>
<td>1.55</td>
</tr>
<tr>
<td>0.990</td>
<td>12208</td>
<td>12208</td>
<td>24416</td>
<td>0.200</td>
<td>0.200</td>
<td>0.50</td>
<td>0.40</td>
<td>1.50</td>
<td>1.40</td>
<td>1.60</td>
</tr>
<tr>
<td>0.990</td>
<td>5435</td>
<td>5435</td>
<td>10870</td>
<td>0.300</td>
<td>0.300</td>
<td>0.50</td>
<td>0.40</td>
<td>1.50</td>
<td>1.36</td>
<td>1.66</td>
</tr>
<tr>
<td>0.990</td>
<td>3065</td>
<td>3065</td>
<td>6130</td>
<td>0.400</td>
<td>0.400</td>
<td>0.50</td>
<td>0.40</td>
<td>1.50</td>
<td>1.31</td>
<td>1.71</td>
</tr>
<tr>
<td>0.990</td>
<td>1967</td>
<td>1967</td>
<td>3934</td>
<td>0.500</td>
<td>0.500</td>
<td>0.50</td>
<td>0.40</td>
<td>1.50</td>
<td>1.27</td>
<td>1.77</td>
</tr>
<tr>
<td>0.990</td>
<td>1371</td>
<td>1371</td>
<td>2742</td>
<td>0.600</td>
<td>0.600</td>
<td>0.50</td>
<td>0.40</td>
<td>1.50</td>
<td>1.23</td>
<td>1.83</td>
</tr>
</tbody>
</table>
Confidence Intervals for Two Proportions

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>Total Sample Size</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.990</td>
<td>1012</td>
<td>2024</td>
<td>3036</td>
<td>0.700</td>
<td>1.50</td>
</tr>
<tr>
<td>0.990</td>
<td>778</td>
<td>1556</td>
<td>2334</td>
<td>0.800</td>
<td>1.50</td>
</tr>
<tr>
<td>0.990</td>
<td>618</td>
<td>1236</td>
<td>1854</td>
<td>0.900</td>
<td>1.50</td>
</tr>
<tr>
<td>0.990</td>
<td>504</td>
<td>1008</td>
<td>1512</td>
<td>0.999</td>
<td>1.50</td>
</tr>
</tbody>
</table>

References

Report Definitions
Confidence level is the proportion of confidence intervals (constructed with this same confidence level, sample size, etc.) that would contain the true odds ratio of population proportions. N1 and N2 are the number of items sampled from each population. N is the total sample size, N1 + N2. Target Width is the value of the width that is entered into the procedure. Actual Width is the value of the width that is obtained from the procedure. P1 and P2 are the assumed sample proportions for sample size calculations. Odds Ratio O1/O2 is the sample odds ratio at which sample size calculations are made. Lower Limit and Upper Limit are the lower and upper limits of the confidence interval for the true odds ratio of proportions (Population Odds 1 / Population Odds 2).

Summary Statements
Group sample sizes of 28244 and 28244 produce a two-sided 95% confidence interval for the population odds ratio with a width that is equal to 0.100 when the estimated sample proportion 1 is 0.50, the estimated sample proportion 2 is 0.40, and the sample odds ratio is 1.50. This report shows the calculated sample sizes for each of the scenarios.

Plots Section

These plots show the group sample size versus the confidence interval width for the two confidence levels.
Example 9 – Validation (Odds Ratios) using Fleiss et al

Fleiss et al (2003) pages 117, 119 give an example of a calculation for a confidence interval for the odds ratio when the confidence level is 95%, the sample odds ratio is 2.25 and the sample proportion 2 is 0.1, the sample size for group 2 is 150, and the interval width is 4.387 for the Logarithm method, and 4.980 for the Fleiss method. The necessary sample size for group 1 in each case is 50.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Confidence Intervals for Two Proportions using Odds Ratios procedure window by expanding Proportions, then Two Independent Proportions, then clicking on Confidence Interval, and then clicking on Confidence Intervals for Two Proportions using Odds Ratios. You may then make the appropriate entries as listed below, or open Example 9 by going to the File menu and choosing Open Example Template.

Option Value
Design Tab
Solve For .................................................. Sample Size
Confidence Interval Formula .................. Varies [Logarithm, Fleiss]
Interval Type ........................................... Two-Sided
Confidence Level .................................... 0.95
Group Allocation ................................. Enter N2, solve for N1
N2 ........................................................... 150
Confidence Interval Width (Two-Sided) .. Varies (4.387, 4.980)
Odds Ratio .............................................. 2.25
P2 ............................................................ 0.1

Output

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

Logarithm

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Target Width</th>
<th>Actual Width</th>
<th>P1</th>
<th>P2</th>
<th>Odds Ratio O1/O2</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950</td>
<td>50</td>
<td>150</td>
<td>200</td>
<td>4.387</td>
<td>4.387</td>
<td>0.20</td>
<td>0.10</td>
<td>2.25</td>
<td>0.96</td>
<td>5.35</td>
</tr>
</tbody>
</table>

PASS also calculates the necessary sample size for Group 1 to be 50.

Fleiss

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Target Width</th>
<th>Actual Width</th>
<th>P1</th>
<th>P2</th>
<th>Odds Ratio O1/O2</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950</td>
<td>50</td>
<td>150</td>
<td>200</td>
<td>4.980</td>
<td>4.980</td>
<td>0.20</td>
<td>0.10</td>
<td>2.25</td>
<td>0.86</td>
<td>5.84</td>
</tr>
</tbody>
</table>

PASS also calculates the necessary sample size for Group 1 to be 50.