Chapter 102

Confidence Intervals for the Difference Between Two Correlated Proportions

Introduction

This routine calculates the sample sizes needed to achieve a specified confidence interval width for the difference between two correlated proportions. Correlated proportions occur when two binary variables are obtained on each subject.

Several formulas have been proposed to calculate confidence intervals for the difference between two correlated proportions. This procedure provides results for five methods.

Technical Details

This section provides the technical details of how the sample sizes are determined.

Comparing Two Proportions

Suppose two dichotomous responses (Yt and Ys) are obtained on each of n subjects. A common example is when Ys represents a pretest binary response and Yt represents a posttest response.

Suppose the probability of obtaining the event of interest in variable Yt is Pt. That is, $P(Yt = 1) = Pt$. Similarly, the probability of obtaining the event of interest in variable Ys is Ps. That is, $P(Ys = 1) = Ps$. The corresponding failure probabilities are given by $Qt = P(Yt = 0)$ and $Qs = P(Ys = 0)$. The purpose of the study is to compare Pt and Ps by evaluating their difference.

The following table gives both the marginal (total) probabilities and the joint (or cell) probabilities.

	Ys = 1	Ys = 0	Totals
Yt = 1	P_{11}	P_{10}	P_t
Yt = 0	P_{01}	P_{00}	Q_t
Totals	P_s	Q_s	1

Note that $P_t = P_{11} + P_{10}$ and $P_s = P_{11} + P_{01}$, so the quantity P_{11} is common to both terms. This shows that P_t and P_s are correlated.

An assumption is made that the responses for each variable follow binomial distributions. Since both variables are measured on the same individual, the variables are assumed to be correlated.

A random sample of n individuals (or pairs) are obtained, and the two variables are measured on each individual. The data from these samples can be displayed in a 2-by-2 contingency table as follows

	Ys = 1	Ys = 0	Totals
Yt = 1	f_{11}	f_{10}	$f_{1\bullet}$
Yt = 0	f_{01}	f_{00}	$f_{0\bullet}$
Totals	$f_{\bullet 1}$	$f_{\bullet 0}$	n

The binomial proportions P_t and P_s are estimated using the formulae

$$\hat{P}_t = p_t = \frac{f_{1\bullet}}{n} \text{ and } \hat{P}_s = p_s = \frac{f_{\bullet 1}}{n}$$

The (risk) difference $\delta = P_t - P_s$ is perhaps the most direct method of comparison between these 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, it would probability be dismissed as being of little importance. That is, there is usually little interest in a treatment that decreases the probability from 0.400 to 0.399. However, if the baseline probably 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 of the size of the difference depends on the baseline probability of the event. So, although using the simple difference is a useful method of comparison, care must be taken that it fits the situation.

Confidence Intervals for the Difference (Pt – Ps)

Many formulas have been proposed for computing confidence intervals for the difference between two paired proportions. Five of these methods are available in this procedure.

It is difficult to recommend one procedure as the best. However, all comparisons in the literature that we have reviewed have found the *Wald* and the *Wald with continuity correction* to be lacking in several aspects. Therefore, we recommend avoiding these methods. They are provided because they often appear in elementary statistics texts. But, in real work, they should be avoided.

These confidence interval methods are as follows.

Wald

Fleiss *et al.* (2003, p. 378) and Newcombe (1998c) mention a confidence interval for the difference in proportions given by

$$(p_1 - p_2) \pm z_{1-\alpha/2} se_{p_1-p_2}$$

where

$$se_{p_1-p_2} = \frac{\sqrt{n(f_{10} + f_{01}) - (f_{10} - f_{01})^2}}{n\sqrt{n}}$$

Wald with Continuity Correction

Fleiss *et al.* (2003, p. 378) and Newcombe (1998c) mention a confidence interval for the difference in proportions given by

$$(p_1 - p_2) \pm \left(z_{1-\alpha/2} se_{p_1-p_2} + 1/n \right)$$

where

$$se_{p_1-p_2} = \frac{\sqrt{n(f_{10} + f_{01}) - (f_{10} - f_{01})^2}}{n\sqrt{n}}$$

Agresti and Min's Wald+2

Agresti and Min (2005) provide an improvement to the Wald Method which is made by adding two observations to the 2x2 table before calculating the interval. This modification improves the overall coverage probability and allows results to be calculated even though one or more marginal frequency totals are zero.

The modification is simply to add $\frac{1}{2}$ to each of the cell frequencies before making any calculations and then to calculate the confidence interval using the Wald method. They call this the "Wald + 2" method.

Bonett and Price's Adjusted Wald

Bonett and Price (2012) provide an improvement to the Wald method which is similar to the method of Agresti and Min method. The adjustment is made by adding two observations to the 2x2 table before calculating the interval. This modification improves the overall coverage probability and allows results to be calculated even though one or more marginal frequency totals are zero.

The modification is simply to add one to each of the non-diagonal cell frequencies before making any calculations and then calculate the confidence interval using the Wald method. Bonett and Price call this the "Adjusted Wald" method. They show that it performs well in most situations.

Newcombe's Score

Newcombe (1998c, p. 2639, method 10), Altman *et al.* (2000, p. 52), and Machin *et al.* (2018, p. 141-142) provide a formula for computing a confidence interval of the proportion difference. Through extensive analysis, Newcombe shows that this method is a very good choice.

Newcombe describes this method as using score intervals for the marginal proportions but adding a continuity correction for the estimated correlation coefficient.

The lower limit is given by

$$D_L = (p_t - p_s) - W_L$$

where

$$W_L = \sqrt{(p_t - L_t)^2 - 2\rho(p_t - L_t)(U_s - p_s) + (U_s - p_s)^2}$$

and the upper limit is given by

$$D_U = (p_t - p_s) + W_U$$

where

$$W_U = \sqrt{(p_s - L_s)^2 - 2\rho(p_s - L_s)(U_t - p_t) + (U_t - p_t)^2}$$

$$p_t = \frac{f_{11} + f_{10}}{n}$$

$$p_s = \frac{f_{11} + f_{01}}{n}$$

$$L_t = \frac{A_t - B_t}{C_t}$$

$$U_t = \frac{A_t + B_t}{C_t}$$

$$A_t = 2np_t + z_{1-\alpha/2}^2$$

$$B_t = z_{1-\alpha/2}\sqrt{z_{1-\alpha/2}^2 + 4np_t(1 - p_t)}$$

$$C_t = 2\left(n + z_{1-\alpha/2}^2\right)$$

$$L_s = \frac{A_s - B_s}{C_s}$$

$$U_s = \frac{A_s + B_s}{C_s}$$

$$A_s = 2np_s + z_{1-\alpha/2}^2$$

$$B_s = z_{1-\alpha/2}\sqrt{z_{1-\alpha/2}^2 + 4np_s(1 - p_s)}$$

$$C_s = 2\left(n + z_{1-\alpha/2}^2\right)$$

The within subject correlation coefficient with continuity correction, ρ_{cc}, is calculated as follows.

If $f_{11} + f_{10}$, $f_{01} + f_{00}$, $f_{11} + f_{01}$, or $f_{10} + f_{00}$ then set $\rho_{cc} = 0$.

Otherwise, calculate

$$A = (f_{11} + f_{10})(f_{01} + f_{00})(f_{11} + f_{01})(f_{10} + f_{00})$$

$$B = (f_{11}f_{00}) - (f_{10}f_{01})$$

$$C = \begin{cases} B - \dfrac{n}{2} & \text{if } B > \dfrac{n}{2} \\ 0 & \text{if } 0 \le B \le \dfrac{n}{2} \\ B & \text{if } B < 0 \end{cases}$$

and then calculate

$$\rho_{cc} = C/\sqrt{A}$$

The width of this interval is $D_U - D_L$.

Correlation of Yt and Ys (Pearson's Phi Coefficient)

The (uncorrected) correlation of the two binary variables, *Yt* and *Ys*, measured on each subject is sometimes called Pearson's phi coefficient, φ. It is be calculated as follows:

$$\rho = \varphi = \frac{P_{11}P_{00} - P_{10}P_{01}}{\sqrt{P_t P_s Q_t Q_s}}$$

This correlation may be used with the marginal probabilities to obtain the joint (cell) probabilities using the following relationships:

$$P_{11} = P_t P_s + \rho \sqrt{P_t P_s Q_t Q_s}$$

$$P_{10} = P_t Q_s - \rho \sqrt{P_t P_s Q_t Q_s}$$

$$P_{01} = Q_t P_s - \rho \sqrt{P_t P_s Q_t Q_s}$$

$$P_{00} = Q_t Q_s + \rho \sqrt{P_t P_s Q_t Q_s}$$

The lower and upper limits of this correlation is given by

$$\rho_L = \max\left\{ -\sqrt{\frac{P_s P_t}{(1 - P_s)(1 - P_t)}}, -\sqrt{\frac{(1 - P_s)(1 - P_t)}{P_s P_t}} \right\}$$

$$\rho_U = \min\left\{ \sqrt{\frac{P_s(1 - P_t)}{P_t(1 - P_s)}}, \sqrt{\frac{P_t(1 - P_s)}{P_s(1 - P_t)}} \right\}$$

Continuity Corrected Correlation

Newcombe's score method uses a continuity corrected correlation (phi) in the calculation of interval width. This correction, shown above, can be summarized as follows.

$$\rho_{cc} = \rho - \frac{1}{2n\sqrt{P_t P_s Q_t Q_s}}$$

To keep all input values consistent, always enter the uncorrected estimate of ρ. The program will make necessary adjustments when needed with Newcombe's method.

Alternative Ways of Setting the Parameter Values

The main parameter of interest is $\delta = P_t - P_s$. Most of the formulas for the confidence intervals of δ involve the joint probabilities. When the study is being planned it is usually very difficult to find estimates of the joint probabilities $(P_{11}, P_{10}, P_{01}, P_{00})$ from previous studies or from experts. However, estimates of the marginal probabilities (P_t, Q_t, P_s, Q_s) and the within-subject correlation (ρ) are usually much easier to obtain. Using the relationships given above, the values of the joint probabilities may be calculated from the marginal probabilities.

In order to make comparisons with other results easier, **PASS** allows for both input types: marginal probabilities or joint probabilities.

Example 1 – Calculating Sample Size

Suppose a study is planned in which the researcher wishes to compare the binary (yes/no) response to a certain treatment. The responses of each subject are measured before (Ys) and after (Yt) the treatment is administered.

The researchers want to estimate the sample size needed to construct a 95% confidence interval for the difference in response proportions such that the width of the interval is no wider than 0.1. Newcombe's score method will be used to construct the interval.

Previous studies provide a planning value for P_s of 0.41 and the (uncorrected) within-subject correlation will be between 0.51 and 0.62.

Planning values for P_t will be 0.5, 0.6, and 0.7.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For	**Sample Size**
Confidence Interval Formula..........................	**Newcombe's Score**
Confidence Level (1 - Alpha)	**0.95**
W_D (Width of Confidence Interval)	**0.1**
Probability Input Type	**Enter Marginal Probabilities**
Pt (Probability (Yt = 1))	**0.5 0.6 0.7**
Ps (Probability (Ys = 1))...................................	**0.41**
ρ (Correlation of Yt and Ys)	**0.51 0.62**

Output

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

Numeric Reports

Numeric Results
──
Solve For: Sample Size
Confidence Interval Formula: Newcombe's Score
──

Width of Confidence Interval W_D	Number of Pairs N	Marginal Probabilities		Difference $P_t - P_s$	Correlation of Yt and Ys ρ	Confidence Interval Segment Widths		Confidence Level
		Pt	Ps			Lower W_L	Upper W_U	
0.09998	371	0.5	0.41	0.09	0.51	0.05044	0.04954	0.95
0.09999	289	0.5	0.41	0.09	0.62	0.05050	0.04949	0.95
0.09991	364	0.6	0.41	0.19	0.51	0.05094	0.04897	0.95
0.09984	284	0.6	0.41	0.19	0.62	0.05104	0.04881	0.95
0.09990	343	0.7	0.41	0.29	0.51	0.05149	0.04842	0.95
		0.7	0.41	0.29	0.62	0.05149	0.04842	0.95

──
Warning: One or more input parameter combinations resulted in an undefined table for which no calculations are possible.

W_D	The width of the confidence interval for the difference between the two proportions.
N	The number of pairs in the study.
Pt	The marginal probability that the response variable Yt = 1.
Ps	The marginal probability that the response variable Ys = 1.
Pt - Ps	The difference between the two proportions, Pt and Ps. It is the parameter of interest.
ρ	The within-subject correlation between Yt and Ys. It is also called the phi (ϕ) coefficient.
W_L	The distance from the difference to the lower C.I. limit. Note that $W_D = W_L + W_U$.
W_U	The distance from the difference to the upper C.I. limit. Note that $W_D = W_L + W_U$.
Confidence Level	The confidence level of the confidence interval.

Summary Statements
──
A paired design will be used to obtain a two-sided 95% confidence interval for the difference between two correlated proportions. The Newcombe's Score method will be used to compute the confidence interval. The response proportions of the treatment and standard are assumed to be 0.5 and 0.41, respectively. The within-subject correlation is assumed to be 0.51. To produce a confidence interval with a width of no more than 0.1, 371 subject pairs will be needed.
──

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	371	464	93
20%	289	362	73
20%	364	455	91
20%	284	355	71
20%	343	429	86
20%			

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. 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. After solving for N, 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 Lokhnygina, Y. (2018) pages 32-33.)
D	The expected number of dropouts. D = N' - N.

Dropout Summary Statements

Anticipating a 20% dropout rate, 464 subjects should be enrolled to obtain a final sample size of 371 subjects.

References

Newcombe, R.G. 1998. 'Improved Confidence Intervals for the Difference Between Binomial Proportions Based on Paired Data.' Statistics in Medicine. Volume 17. Pages 2635-2650.

Altman, D.G, Machin, D., Bryant, T.N., Gardner, M.J. 2000. Statistics with Confidence, 2nd Edition. BMJ Publishing.

Agresti, A. and Min, Y. 2005. 'Simple improved confidence intervals for comparing matched proportions.' Statistics in Medicine. Volume 24. Pages 729-740. DOI: 10.1002/sim.1781.

Bonett, D.G. and Price, R.M. 2012. 'Adjusted Wald Confidence Interval for a Difference of Binomial Proportions Based on Paired Data.' Journal of Educational and Behavioral Statistics. Volume 17, Number 4, Pages 479-488. DOI:10.3102/1076998611411915.

Machin, D., Campbell, M., Tan, S.B., and Tan, S.H. 2018. Sample Size Tables for Clinical, Laboratory and Epidemiology Studies, 4th Edition. John Wiley & Sons. Hoboken, NJ.

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

Plots Section

Plots
_____
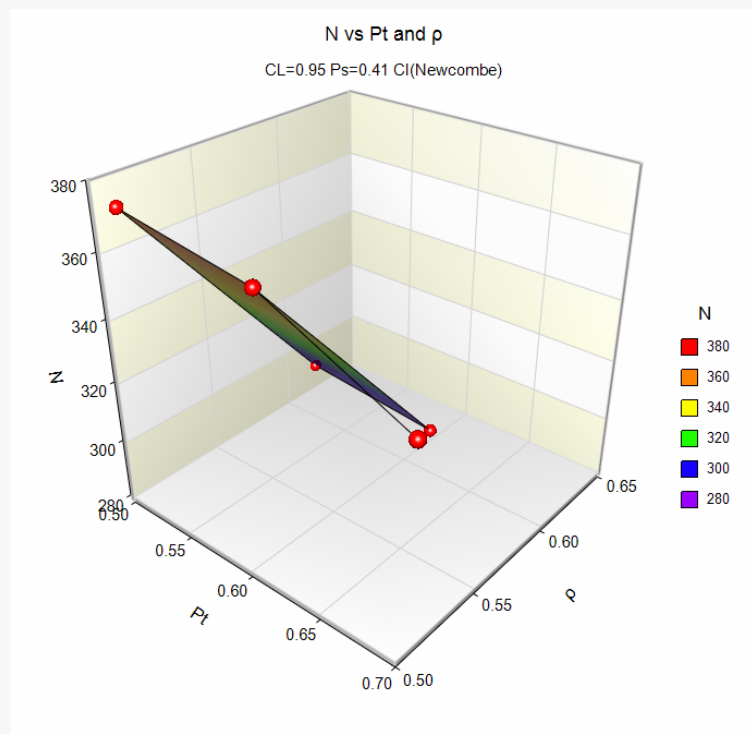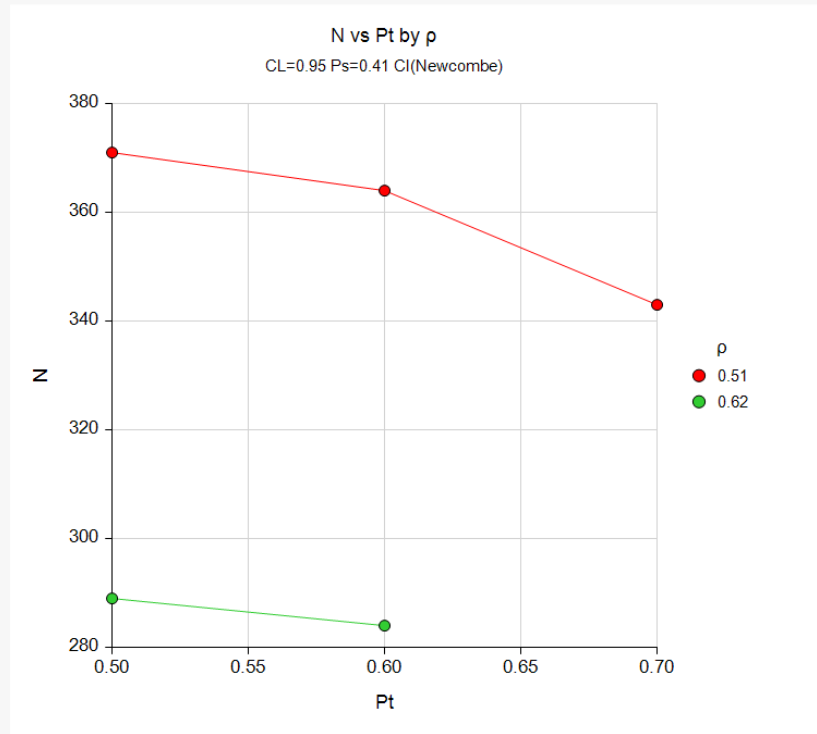
These plots show the group sample sizes for each of the scenarios. It appears that the value of the within-subject correlation has the largest impact on sample size in this study.

Example 2 – Validation of Newcombe's Score Method using Newcombe (1998c)

Newcombe (1998c, page 2641) presents a table from which we will use the second row as our validation example. In this row, $P_{11} = 0.4, P_{10} = 0.24, P_{01} = 0.04, P_{00} = 0.32,$ and $N = 50.$ The resulting 95% confidence interval for the difference is 0.0562 to 0.3292 which gives a width of 0.273.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For ....**C.I. Width**
Confidence Interval Formula..........................**Newcombe's Score**
Confidence Level (1 - Alpha) **0.95**
N (Number of Pairs)......................................**50**
Probability Input Type ....**Enter Joint (Cell) Probabilities**
P10 (Probability (Yt = 1, Ys = 0)) **0.24**
P01 (Probability (Yt = 0, Ys = 1)) **0.04**
P11 (Probability (Yt = 1, Ys = 1)) **0.4**

Output

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

Numeric Results

Solve For: C.I. Width
Confidence Interval Formula: Newcombe's Score

Width of Confidence Interval W_D	Number of Pairs N	Joint (Cell) Probabilities			Difference $P_t - P_s$	Confidence Interval Segment Widths		Confidence Level
		P10	P01	P11		Lower W_L	Upper W_U	
0.27305	50	0.24	0.04	0.4	0.2	0.14384	0.12921	0.95

This report gives the width as 0.27305 which validates this procedure.

Example 3 – Validation of the Wald Method using Newcombe (1998c)

Newcombe (1998c, page 2640) presents a table from which we will use the first row as our validation example. In this row, $P_{11} = 0.36, P_{10} = 0.24, P_{01} = 0.04, P_{00} = 0.36,$ and $N = 50$. The resulting 95% confidence interval for the difference is 0.0642 to 0.3358 which gives a width of 0.2716.

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 ....**C.I. Width**
Confidence Interval Formula.........................**Wald**
Confidence Level (1 - Alpha) **0.95**
N (Number of Pairs).....................................**50**
Probability Input Type ....**Enter Joint (Cell) Probabilities**
P10 (Probability (Yt = 1, Ys = 0)) **0.24**
P01 (Probability (Yt = 0, Ys = 1)) **0.04**

Output

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

Numeric Results
───
Solve For: C.I. Width
Confidence Interval Formula: Wald
───

Width of Confidence Interval W_D	Number of Pairs N	Joint (Cell) Probabilities P10	P01	Difference Pt - Ps	Standard Error of the Difference SE	Confidence Level
0.27158	50	0.24	0.04	0.2	0.06928	0.95

This report shows gives the width as 0.27158 which validates this procedure since it rounds to 0.2716.

Example 4 – Validation of the Wald C.C. Method using Newcombe (1998c)

Newcombe (1998c, page 2640) presents a table from which we will use the second row as our validation example. In this row, $P_{11} = 0.36, P_{10} = 0.24, P_{01} = 0.04, P_{00} = 0.36,$ and $N = 50.$ The resulting 95% confidence interval for the difference is 0.0442 to 0.3558 which gives a width of 0.3116.

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 ....**C.I. Width**
Confidence Interval Formula.........................**Wald with Continuity Correction**
Confidence Level (1 - Alpha) **0.95**
N (Number of Pairs).......................................**50**
Probability Input Type ....**Enter Joint (Cell) Probabilities**
P10 (Probability (Yt = 1, Ys = 0)) **0.24**
P01 (Probability (Yt = 0, Ys = 1)) **0.04**

Output

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

Numeric Results
_____

Solve For: C.I. Width
Confidence Interval Formula: Wald with Continuity Correction
_____

Width of Confidence Interval W_D	Number of Pairs N	Joint (Cell) Probabilities P10	P01	Difference Pt - Ps	Standard Error of the Difference SE	Confidence Level
0.31158	50	0.24	0.04	0.2	0.06928	0.95
_____

This report shows gives the width as 0.31158 which validates this procedure since it rounds to 0.3116.

Example 5 – Validation of the Agresti and Min Method using Agresti and Min (2005)

Agresti and Min (2005, page 736) present a Table V from which we will use the second row as our validation example. In this row, $P_{10} = \frac{8}{86} = 0.093023, P_{01} = \frac{16}{86} = 0.186047,$ and $N = 86$. The resulting 95% confidence interval for the difference is -0.019 to 0.201 which gives a rounded width of 0.22.

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 ....**C.I. Width**
Confidence Interval Formula..........................**Agresti and Min's Wald+2**
Confidence Level (1 - Alpha) **0.95**
N (Number of Pairs)......................................**86**
Probability Input Type ....**Enter Joint (Cell) Probabilities**
P10 (Probability (Yt = 1, Ys = 0)) **0.093023**
P01 (Probability (Yt = 0, Ys = 1)) **0.186047**

Output

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

Numeric Results

Solve For: C.I. Width
Confidence Interval Formula: Agresti and Min's Wald + 2

Width of Confidence Interval W_D	Number of Pairs N	Joint (Cell) Probabilities		Difference Pt - Ps	Standard Error of the Difference SE	Confidence Level
		P10	P01			
0.21946	86	0.09302	0.18605	-0.09302	0.05599	0.95

This report gives the width as 0.21946 which validates this procedure since it is within rounding of 0.22.

Example 6 – Validation of the Bonett and Price Method using Bonett and Price (2012)

Bonett and Price (2012, page 484) present an example which we will use as our validation example. In this example, $P_{10} = \frac{50}{200} = 0.25, P_{01} = \frac{22}{200} = 0.11,$ and $N = 200$. The resulting 95% confidence interval for the difference is 0.057 to 0.220 which gives a width of 0.163.

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 ....**C.I. Width**
Confidence Interval Formula..........................**Bonett and Price's Adjusted Wald**
Confidence Level (1 - Alpha) **0.95**
N (Number of Pairs)..**200**
Probability Input Type ....**Enter Joint (Cell) Probabilities**
P10 (Probability (Yt = 1, Ys = 0)) **0.25**
P01 (Probability (Yt = 0, Ys = 1)) **0.11**

Output

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

Numeric Results

Solve For: C.I. Width
Confidence Interval Formula: Bonett and Price's Adjusted Wald

Width of Confidence Interval W_D	Number of Pairs N	Joint (Cell) Probabilities		Difference Pt - Ps	Standard Error of the Difference SE	Confidence Level
		P10	P01			
0.1625	200	0.25	0.11	0.14	0.04145	0.95

This report gives the width as 0.1625 which validates this procedure since it is within rounding of 0.163.