Chapter 271

Confidence Intervals for One-Sample Sensitivity

Introduction

This procedure calculates the (whole table) sample size necessary for a single-sample sensitivity confidence interval, based on a specified sensitivity, interval width, confidence level, and prevalence.

Caution: This procedure assumes that the sensitivity of the future sample will be the same as the sensitivity that is specified. If the sample sensitivity is different from the one specified when running this procedure, the interval width may be narrower or wider than specified.

Sensitivity (True Positive Rate)

The sensitivity (or true positive rate) is the proportion of the individuals with a known positive condition for which the predicted condition is positive.

		Predicted Condition	
		Positive	Negative
True Condition	**Positive**	**True Positive (A)**	False Negative (C)
	Negative	False Positive (B)	True Negative (D)

Sensitivity = A / (A + C)

Prevalence

The prevalence is the overall proportion of individuals with a positive condition.

		Predicted Condition	
		Positive	Negative
True Condition	**Positive**	**True Positive (A)**	**False Negative (C)**
	Negative	False Positive (B)	True Negative (D)

Prevalence = (A + C) / (A + B + C + D)

Technical Details

In general terms, the required sample size is determined by first calculating the sample size needed for the sensitivity proportion confidence interval, followed by a prevalence adjustment. The initial sample size calculation for the sensitivity confidence interval gives the number of individuals with a positive condition that are needed. The prevalence adjustment is used to add the number of individuals with a negative condition that are needed. The resulting sample size is the total number of individuals needed to obtain a table where the number of positive condition individuals will give the needed confidence interval width for the sensitivity.

Similarly, when calculating the confidence interval width for a given sample size, the given sample size is first used to produce the number of positive condition individuals, according to the given prevalence, and then the width based on the resulting positive condition count is then calculated.

If prevalence is to be ignored, a value of 1 may be used for prevalence, or the Confidence Intervals for One Proportion procedure may be used, as the scenario has been reduced to a simple confidence interval of a single proportion.

Confidence Interval Formulas

Many methods have been devised for computing confidence intervals for a single proportion. Five of these methods are available in this procedure. The five confidence interval methods are

1. Exact (Clopper-Pearson)
2. Score (Wilson)
3. Score with continuity correction
4. Simple Asymptotic
5. Simple Asymptotic with continuity correction

For a comparison of methods, see Newcombe (1998a).

For each of the following methods, let *p* be the population sensitivity, and let *r* represent the number of true positives with *n* total positives. Let $\hat{p} = r \,/\, n$.

Exact (Clopper-Pearson)

Using a mathematical relationship (see Fleiss et al (2003), p. 25) between the *F* distribution and the cumulative binomial distribution, the lower and upper confidence limits of a 100(1-α)% exact confidence interval for the true proportion *p* are given by

$$\left[\frac{r}{r+(n-r+1)F_{1-\alpha/2;2(n-r+1),2r}}, \frac{(r+1)F_{1-\alpha/2;2(r+1),2(n-r)}}{(n-r)+(r+1)F_{1-\alpha/2;2(r+1),2(n-r)}}\right]$$

One-sided limits may be obtained by replacing $\alpha/2$ by α.

Score (Wilson)

The Wilson Score confidence interval, which is based on inverting the z-test for a single proportion, is calculated using

$$\frac{\left(2n\hat{p} + z_{1-\alpha/2}^2\right) \pm z_{1-\alpha/2}\sqrt{z_{1-\alpha/2}^2 + 4n\hat{p}(1-\hat{p})}}{2\left(n + z_{1-\alpha/2}^2\right)}$$

One-sided limits may be obtained by replacing $\alpha/2$ by α.

Score with Continuity Correction

The Score confidence interval with continuity correction is based on inverting the z-test for a single proportion with continuity correction. The $100(1-\alpha)\%$ limits are calculated by

$$Lower\ Limit = \frac{\left(2n\hat{p} + z_{1-\alpha/2}^2 - 1\right) - z_{1-\alpha/2}\sqrt{z_{1-\alpha/2}^2 - \{2 + (1/n)\} + 4\hat{p}\{n(1-\hat{p}) + 1\}}}{2\left(n + z_{1-\alpha/2}^2\right)}$$

$$Upper\ Limit = \frac{\left(2n\hat{p} + z_{1-\alpha/2}^2 + 1\right) + z_{1-\alpha/2}\sqrt{z_{1-\alpha/2}^2 + \{2 - (1/n)\} + 4\hat{p}\{n(1-\hat{p}) - 1\}}}{2\left(n + z_{1-\alpha/2}^2\right)}$$

One-sided limits may be obtained by replacing $\alpha/2$ by α.

Simple Asymptotic

The simple asymptotic formula is based on the normal approximation to the binomial distribution. The approximation is close only for very large sample sizes. The $100(1-\alpha)\%$ confidence limits are given by

$$\hat{p} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

One-sided limits may be obtained by replacing $\alpha/2$ by α.

Simple Asymptotic with Continuity Correction

This formula is identical to the previous one, but with continuity correction. The $100(1-\alpha)\%$ confidence limits are

$$\left(\hat{p} - z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} - \frac{1}{2n}, \hat{p} + z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} + \frac{1}{2n}\right)$$

One-sided limits may be obtained by replacing $\alpha/2$ by α.

Interval Widths (One-Sided vs. Two-Sided)

For two-sided intervals, the distance from the sample sensitivity 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 sample sensitivity to limit, *D*, is specified.

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

$$D = U - \hat{p}$$

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

$$D = \hat{p} - L$$

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

Example 1 – Calculating Sample Size

Suppose a study is planned in which the researcher wishes to construct a two-sided 95% exact (Clopper-Pearson) confidence interval for the population sensitivity such that the width of the interval is no wider than 0.06. The anticipated sensitivity estimate is 0.7, but a range of values from 0.5 to 0.9 will be included to determine the effect of the sensitivity estimate on necessary sample size. Instead of examining only the interval width of 0.06, widths of 0.04, 0.08, and 0.10 will also be considered.

The goal is to determine the total sample size needed when also accounting for 30% prevalence.

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.......................... **Exact (Clopper-Pearson)**
Interval Type .... **Two-Sided**
Confidence Level **0.95**
Confidence Interval Width (Two-Sided) **0.04 0.06 0.08 0.10**
Sensitivity.. **0.5 to 0.9 by 0.05**
Prevalence.. **0.3**

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: Exact (Clopper-Pearson)
Confidence Interval Type: Two-Sided

		Confidence Interval Width			Confidence Interval Limits			
Confidence Level	Sample Size N	Target	Actual	Sensitivity	Lower	Upper	Prevalence	Number of Positives
0.95	8164	0.04	0.04	0.50	0.480	0.520	0.3	2449
0.95	8084	0.04	0.04	0.55	0.530	0.570	0.3	2425
0.95	7844	0.04	0.04	0.60	0.580	0.620	0.3	2353
0.95	7444	0.04	0.04	0.65	0.630	0.670	0.3	2233
0.95	6884	0.04	0.04	0.70	0.680	0.720	0.3	2065
0.95	6164	0.04	0.04	0.75	0.730	0.770	0.3	1849
0.95	5284	0.04	0.04	0.80	0.779	0.819	0.3	1585
0.95	4244	0.04	0.04	0.85	0.829	0.869	0.3	1273
0.95	3047	0.04	0.04	0.90	0.879	0.919	0.3	914
0.95	3660	0.06	0.06	0.50	0.470	0.530	0.3	1098
0.95	3627	0.06	0.06	0.55	0.520	0.580	0.3	1088
0.95	3520	0.06	0.06	0.60	0.570	0.630	0.3	1056
0.95	3340	0.06	0.06	0.65	0.620	0.680	0.3	1002
0.95	3094	0.06	0.06	0.70	0.669	0.729	0.3	928
0.95	2774	0.06	0.06	0.75	0.719	0.779	0.3	832
0.95	2384	0.06	0.06	0.80	0.769	0.829	0.3	715
0.95	1920	0.06	0.06	0.85	0.818	0.878	0.3	576
0.95	1390	0.06	0.06	0.90	0.867	0.927	0.3	417
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.

Confidence Level The proportion of confidence intervals (constructed with this same confidence level, sample size, etc.) that would contain the population sensitivity.
N The size of the sample drawn from the population.
Confidence Interval Width The distance from the lower limit to the upper limit.
Target Width The value of the width that is entered into the procedure.
Actual Width The value of the width that is obtained from the procedure.
Sensitivity The assumed sample sensitivity, or true positive rate.
Confidence Interval Limits The lower and upper limits of the confidence interval.
Prevalence The assumed overall proportion of individuals with a positive condition.
Number of Positives The count upon which the confidence interval width calculation is based. Number of Positives = Sample Size × Prevalence, with appropriate rounding.

Summary Statements

A single-group diagnostic test design will be used to obtain a two-sided 95% confidence interval for the sensitivity. The Exact (Clopper-Pearson) formula will be used to calculate the confidence interval. The sample sensitivity is assumed to be 0.5 and the prevalence is assumed to be 0.3. To produce a confidence interval with a width of no more than 0.04, 8164 subjects will be needed.

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	8164	10205	2041
20%	8084	10105	2021
20%	7844	9805	1961
20%	7444	9305	1861
20%	6884	8605	1721
20%	6164	7705	1541
20%	5284	6605	1321
20%	4244	5305	1061
20%	3047	3809	762
20%	3660	4575	915
20%	3627	4534	907
20%	3520	4400	880
20%	3340	4175	835
20%	3094	3868	774
20%	2774	3468	694
20%	2384	2980	596
20%	1920	2400	480
20%	1390	1738	348
.	.	.	.
.	.	.	.
.	.	.	.

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 the confidence interval is computed. If N subjects are evaluated out of the N' subjects that are enrolled in the study, the design will achieve the stated confidence interval.
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, 10205 subjects should be enrolled to obtain a final sample size of 8164 subjects.

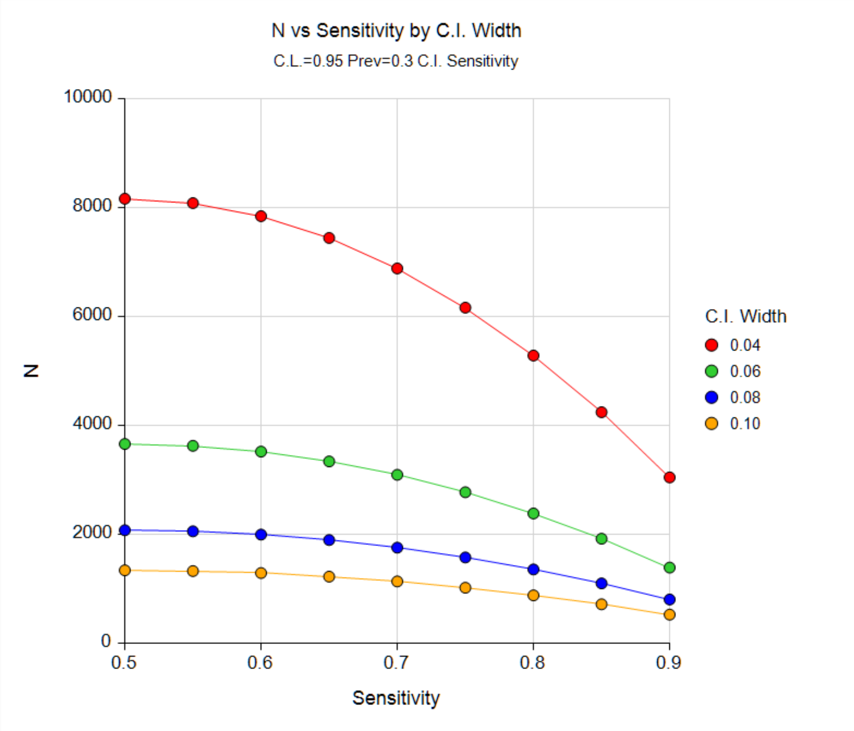
References

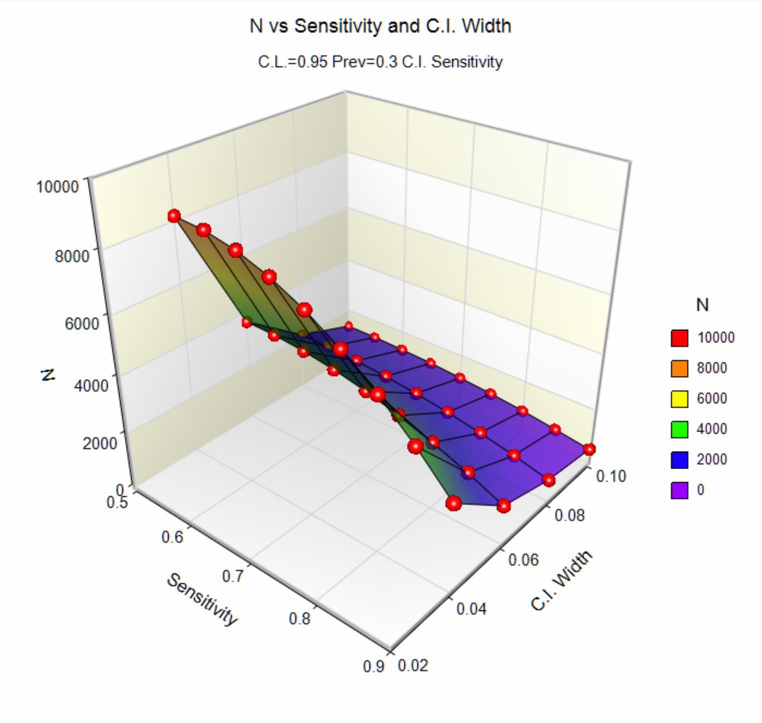
Hajian-Tilaki, K. 2014. 'Sample size estimation in diagnostic test studies of biomedical informatics.' Journal of Biomedical Informatics, 48, pp. 193-204.
Fleiss, J. L., Levin, B., Paik, M.C. 2003. Statistical Methods for Rates and Proportions. Third Edition. John Wiley & Sons. New York.
Newcombe, R. G. 1998. 'Two-Sided Confidence Intervals for the Single Proportion: Comparison of Seven Methods.' Statistics in Medicine, 17, pp. 857-872.

These reports show the calculated sample size for each of the scenarios.

Plots Section

Plots





These plots show the sample size versus the sample sensitivity for the four confidence interval widths.

Example 2 – Validation using Hajian-Tilaki (2014)

Hajian-Tilaki (2014), page 195, gives an example of a calculation for a simple asymptotic two-sided confidence interval for a single sensitivity when the confidence level is 95%, the sensitivity is 0.8, the prevalence is 0.1, and the margin of error is 7% (With a margin of error (precision) of 7%, the width is 0.14). The necessary sample size is calculated to be 1254.

Setup

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

Design Tab

Solve For .... **Sample Size**
Confidence Interval Formula.......................... **Simple Asymptotic**
Interval Type .... **Two-Sided**
Confidence Level **0.95**
Confidence Interval Width (Two-Sided) **0.14**
Sensitivity.. **0.8**
Prevalence.. **0.1**

Output

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

Numeric Results

Solve For: Sample Size
Confidence Interval Formula: Simple Asymptotic
Confidence Interval Type: Two-Sided

Confidence Level	Sample Size N	Confidence Interval Width Target	Confidence Interval Width Actual	Sensitivity	Confidence Interval Limits Lower	Confidence Interval Limits Upper	Prevalence	Number of Positives
0.95	1260	0.14	0.14	0.8	0.73	0.87	0.1	126

PASS calculates the necessary sample size to be 1260. The sample size calculated in **PASS** is slightly different from the article. In the article the sample sizes are calculated directly, while **PASS** calculates the sample size needed before prevalence is taken into account, and then adjusts for the prevalence. With a sample size of 1254, the number of positives would be 125.4, which should be rounded up to 126. Adjusting 126 for prevalence gives 1260.