Chapter 513

One Proportion – Equivalence Tests

Introduction

This procedure computes confidence limits and equivalence hypothesis tests for a single proportion. For example, you might want confidence limits for the proportion of individuals with the common cold who took ascorbic acid (vitamin C) and recovered within twenty-four hours. You might want to test the equivalence hypothesis that the percentage of individuals with the common cold who recovered immediately after taking ascorbic acid is equivalent to the standard proportion of 50%.

Exact results, based on the binomial distribution, are calculated. Approximate results based on the normal approximation to the binomial distribution are also given. In the case of a single proportion, the exact results are preferable to the approximate results and should always be used. The approximate results are available in the software because they are commonly presented in elementary statistical texts.

This procedure accepts data entered as summary counts (number of "successes" and sample size) but can also tabulate data from columns in the database.

The Binomial Model

Binomial data must exhibit the following four conditions:

1. The response can take on only one of two possible values. This is a binary response variable.
2. The response is observed a known number of times. Each replication is called a Bernoulli trial. The number of replications is labeled *n*. The number of responses out of the *n* total that exhibit the outcome of interest is labeled *X*. Thus, *X* takes on the possible values 0, 1, 2, …, n.
3. The probability that a particular outcome (a success) occurs is constant for each trial. This probability is labeled P.
4. The trials are independent. The outcome of one trial does not influence the outcome of the any other trial.

The binomial probability, $b(X; n, P)$, is calculated using:

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

where

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

The estimate of P from a sample is labeled p and is estimated using:

$$p = \frac{X}{n}$$

The label $\hat{p}$ is often used in place of p in practice.

Confidence Limits

Using a mathematical relationship (see Ostle(1988), page 110) between the *F* distribution and the cumulative binomial distribution, the lower and upper confidence limits of a $100(1-\alpha)\%$ confidence interval are given by:

$$LCL = \frac{XF_{[\alpha/2],[2X,2(n-X+1)]}}{(n-X+1) + XF_{[\alpha/2],[2X,2(n-X+1)]}}$$

$$UCL = \frac{(X+1)F_{[1-\alpha/2],[2(X+1),2(n-X)]}}{(n-X) + (X+1)F_{[1-\alpha/2],[2(X+1),2(n-X)]}}$$

Note that although these limits are based on direct calculation of the binomial distribution, they are only "exact" for a few values of alpha. Otherwise, these limits are conservative (wider than necessary).
These limits may be approximated using the normal approximation to the binomial as

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

If a correction for continuity is added, the above formula becomes

$$CI_{cc} = p \pm \left(z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} + \frac{1}{2n} \right)$$

Although these two approximate confidence intervals are found in many elementary statistics books, they are not recommended in general. For example, Newcombe (1998) made a comparative study of seven confidence interval techniques and these methods came in last. Instead, Newcombe (1998) recommended the Wilson Score confidence interval method because of its performance. The Wilson Score confidence interval is calculated using

$$CI_{Wilson\ Score} = \frac{\left(2np + z_{\alpha/2}^2\right) \pm z_{\alpha/2}\sqrt{z_{\alpha/2}^2 + 4np(1-p)}}{2\left(n + z_{\alpha/2}^2\right)}$$

Equivalence Hypothesis Tests using TOST (Two One-Sided Tests)

Schuirmann's (1987) two one-sided tests (TOST) approach is used to test equivalence. The equivalence test essentially reverses the roles of the null and alternative hypothesis. Assume that P represents the population proportion of the response, S is a standard reference proportion, and *M* is the so-called *margin of equivalence*. The null and alternative hypotheses are

$$H_0: P < S - M \quad \text{or} \quad P > S + M$$

$$H_1: S - M < P < S + M$$

The null hypothesis is made up of two simple one-sided hypotheses:

$$H_{01}: P < S - M$$

$$H_{02}: P > S + M$$

If both of these one-sided tests are rejected, we conclude H_1 that the response is equivalent to the standard proportion (their difference is confined within a small margin). Schuirmann showed that if we want the alpha level of the equivalence test to be α, then each of the one-sided tests should be α as well (not $\alpha/2$ as you might expect). The probability level (*p*-value) of the equivalence test is equal to the maximum of the probability levels of the two one-sided tests. These tests are conducted using the standard formulas for the one-sample proportions test.

The exact *p*-values for each of these situations may be computed as follows:

1. $P(|\tilde{p} - P_0| \geq |p - P_0|)$, where $\tilde{p}$ represents all possible values of p. This probability is calculated using the binomial distribution.

2. $\sum_{r=0}^{X} b(r; n, p)$

3. $\sum_{r=X}^{n} b(r; n, p)$

The simple large-sample *z*-test is based on the test statistic

$$z = \frac{p - P_0}{\sqrt{\frac{P_0(1 - P_0)}{n}}}$$

Two approximations to the exact p-values are also available. One uses P_0 and the other uses p in the calculation of the standard error. The first approximation uses P_0 in the calculation of the standard error:

$$z_c = \frac{X + 0.5 - nP_0}{\sqrt{nP_0(1 - P_0)}}$$

if $X < nP_0$, or

$$z_c = \frac{X - 0.5 - nP_0}{\sqrt{nP_0(1 - P_0)}}$$

if $X > nP_0$.

The second approximation uses p in the calculation of the standard error:

$$z_c = \frac{X + 0.5 - nP_0}{\sqrt{np(1 - p)}}$$

if $X < nP_0$, or

$$z_c = \frac{X - 0.5 - nP_0}{\sqrt{np(1 - p)}}$$

if $X > nP_0$.

These z-values are used to calculate probabilities using the standard normal probability distribution.

Data Structure

The procedure accepts data entered as summary counts (number of "successes" (*X*) and sample size (*n*)), but can also tabulate data from columns in the database. A separate report is generated for each outcome variable entered.

Example 1 – Equivalence Test of One Proportion using Summary Data

This section presents an example of how to run an equivalence test of summary data in which *n* is 100, *X* is 55, the standard proportion, *S*, is 0.50, and the equivalence margin, *M*, is set to 0.10.

Setup

To run this example, complete the following steps:

1 Specify the One Proportion – Equivalence Tests procedure options

- Find and open the **One Proportion – Equivalence Tests** procedure using the menus or the Procedure Navigator.
- The settings 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.

Data Tab

Type of Data Input ....**Enter the Number of "Successes" (X) and the Sample Size (n)**
Number of "Successes" (X) **55**
Sample Size (n) **100**
Equivalence Bounds ....**Symmetric**
Standard Proportion..**0.50**
Equivalence Margin **0.10**

Reports Tab

Exact (Binomial)...**Checked**
Simple Z ....**Checked**
Simple Z with Continuity Correction................**Checked**
Wilson Score..**Checked**
Exact (Binomial) Test.....................................**Checked**
Simple Z-Test ....**Checked**
Z Approximation Test using P0.......................**Checked**
Z Approximation Test using p ....**Checked**

2 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Data Report

Data Report

Number of "Successes" (X)	Sample Size (n)	Sample Proportion (p)
55	100	0.55

This report documents the values that were entered.

Confidence Limits Section

Exact Confidence Interval of P

				95% C. I. of P	
Number of "Successes" (X)	Sample Size (n)	Sample Proportion (p)	Z*	Lower Limit	Upper Limit
55	100	0.55	1.95996	0.4472802	0.6496798

Simple Z Confidence Interval of P

				95% C. I. of P	
Number of "Successes" (X)	Sample Size (n)	Sample Proportion (p)	Z*	Lower Limit	Upper Limit
55	100	0.55	1.95996	0.452493	0.647507

Simple Z Confidence Interval of P with Continuity Correction

				95% C. I. of P	
Number of "Successes" (X)	Sample Size (n)	Sample Proportion (p)	Z*	Lower Limit	Upper Limit
55	100	0.55	1.95996	0.447493	0.6525069

Wilson Score Confidence Interval of P

				95% C. I. of P	
Number of "Successes" (X)	Sample Size (n)	Sample Proportion (p)	Z*	Lower Limit	Upper Limit
55	100	0.55	1.95996	0.452446	0.6438546

The Wilson score interval is the interval recommended in Newcombe (1998). The limits are based on the formulas that were presented earlier.

Hypothesis Test Section

Exact Test for Equivalence using TOST (Two One-Sided Tests)

Equivalence Hypothesis: 0.5 - 0.1 < P < 0.5 + 0.1

Test	Alternative Hypothesis	Sample Proportion	Prob Level	Reject H0 at α = 0.05?
Lower Boundary	P > 0.4	0.55	0.00171	Yes
Upper Boundary	P < 0.6	0.55	0.17890	No
Equivalence	0.4 < P < 0.6		0.17890	No

Simple Z-Test for Equivalence using TOST (Two One-Sided Tests)

Equivalence Hypothesis: 0.5 - 0.1 < P < 0.5 + 0.1

Test	Alternative Hypothesis	Sample Proportion	Z-Statistic	Prob Level	Reject H0 at α = 0.05?
Lower Boundary	P > 0.4	0.55	3.06190	0.00110	Yes
Upper Boundary	P < 0.6	0.55	-1.02060	0.15372	No
Equivalence	0.4 < P < 0.6			0.15372	No

Z Approximation Test using P0 for Equivalence using TOST (Two One-Sided Tests)

Equivalence Hypothesis: 0.5 - 0.1 < P < 0.5 + 0.1

Test	Alternative Hypothesis	Sample Proportion	Z-Statistic	Prob Level	Reject H0 at α = 0.05?
Lower Boundary	P > 0.4	0.55	2.95980	0.00154	Yes
Upper Boundary	P < 0.6	0.55	-0.91860	0.17916	No
Equivalence	0.4 < P < 0.6			0.17916	No

Z Approximation Test using p for Equivalence using TOST (Two One-Sided Tests)

Equivalence Hypothesis: 0.5 - 0.1 < P < 0.5 + 0.1

Test	Alternative Hypothesis	Sample Proportion	Z-Statistic	Prob Level	Reject H0 at α = 0.05?
Lower Boundary	P > 0.4	0.55	2.91460	0.00178	Yes
Upper Boundary	P < 0.6	0.55	-0.90450	0.18286	No
Equivalence	0.4 < P < 0.6			0.18286	No

The formulas for these tests were shown earlier. The test fail to conclude equivalence.

Example 2 – Equivalence Test of One Proportion using Raw Data

This section presents an example of how to run an analysis of raw data that may have been used to generate the summary data in example 1. The data for this example are found in the **Prop1** dataset.

Setup

To run this example, complete the following steps:

1 Open the Prop1 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Prop1** and click **OK**.

2 Specify the One Proportion – Equivalence Tests procedure options

- Find and open the **One Proportion – Equivalence Tests** procedure using the menus or the Procedure Navigator.
- The settings 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.

Data Tab

Type of Data Input	**Tabulate "Successes" from a Categorical Outcome Variable in the Database**
Outcome Variable(s)	**Response**
"Success" Value	**1**
Equivalence Bounds	**Symmetric**
Standard Proportion	**0.50**
Equivalence Margin	**0.10**

Reports Tab

Exact (Binomial)	**Checked**
Simple Z	**Checked**
Simple Z with Continuity Correction	**Checked**
Wilson Score	**Checked**
Exact (Binomial) Test	**Checked**
Simple Z-Test	**Checked**
Z Approximation Test using P0	**Checked**
Z Approximation Test using p	**Checked**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

Data Report for Response

Number of "Successes" (X)	Sample Size (n)	Sample Proportion (p)
55	100	0.55

The results are exactly the same as they were in Example 1.