Syntax

One-sample test of proportion

   prtest varname == #p [if] [in] [, level(#)]

Two-sample test of proportions using groups

   prtest varname [if] [in], by(groupvar) [level(#)]

Two-sample test of proportions using variables

   prtest varname1 == varname2 [if] [in] [, level(#)]

Immediate form of one-sample test of proportion

   prtesti #obs1 #p1 #p2 [, level(#) count]

Immediate form of two-sample test of proportions

   prtesti #obs1 #p1 #obs2 #p2 [, level(#) count]

   by is allowed with prtest; see [D] by.

Menu

   prtest
   prtesti

   Statistics > Summaries, tables, and tests > Classical tests of hypotheses > Proportion test

   Statistics > Summaries, tables, and tests > Classical tests of hypotheses > Proportion test calculator

Description

   prtest performs tests on the equality of proportions using large-sample statistics.

   In the first form, prtest tests that varname has a proportion of #p. In the second form, prtest
   tests that varname has the same proportion within the two groups defined by groupvar. In the third
   form, prtest tests that varname1 and varname2 have the same proportion.

   prtesti is the immediate form of prtest; see [U] 19 Immediate commands.
The *bitest* command is a better version of the first form of *prtest* in that it gives exact *p*-values. Researchers should use *bitest* when possible, especially for small samples; see [R] *bitest*.

**Options**

by(*groupvar*) specifies a numeric variable that contains the group information for a given observation. This variable must have only two values. Do not confuse the by(*) option with the by prefix; both may be specified.

level(#) specifies the confidence level, as a percentage, for confidence intervals. The default is level(95) or as set by set level; see [U] 20.7 Specifying the width of confidence intervals.

count specifies that integer counts instead of proportions be used in the immediate forms of *prtest*.

In the first syntax, *prtesti* expects that #obs1 and #p1 are counts—#p1 ≤ #obs1—and #p2 is a proportion. In the second syntax, *prtesti* expects that all four numbers are integer counts, that #obs1 ≥ #p1, and that #obs2 ≥ #p2.

**Remarks and examples**

The *prtest* output follows the output of *ttest* in providing a lot of information. Each proportion is presented along with a confidence interval. The appropriate one- or two-sample test is performed, and the two-sided and both one-sided results are included at the bottom of the output. For a two-sample test, the calculated difference is also presented with its confidence interval. This command may be used for both large-sample testing and large-sample interval estimation.

**Example 1: One-sample test of proportion**

In the first form, *prtest* tests whether the mean of the sample is equal to a known constant. Assume that we have a sample of 74 automobiles. We wish to test whether the proportion of automobiles that are foreign is different from 40%.

```
. use http://www.stata-press.com/data/r13/auto
(1978 Automobile Data)
. prtest foreign == .4
One-sample test of proportion foreign: Number of obs = 74
       Variable | Mean  Std. Err.   [95% Conf. Interval]
-------------+--------------------------------------
       foreign | 0.2973   .053131   .1932  .4014

   p = proportion(foreign) z = -1.8034
Ho: p = 0.4   Ha: p < 0.4   Ha: p != 0.4   Ha: p > 0.4
Pr(Z < z) = 0.0357   Pr(|Z| > |z|) = 0.0713   Pr(Z > z) = 0.9643
```

The test indicates that we cannot reject the hypothesis that the proportion of foreign automobiles is 0.40 at the 5% significance level.
Example 2: Two-sample test of proportions

We have two headache remedies that we give to patients. Each remedy’s effect is recorded as 0 for failing to relieve the headache and 1 for relieving the headache. We wish to test the equality of the proportion of people relieved by the two treatments.

```
use http://www.stata-press.com/data/r13/cure
prtest cure1 == cure2
```

Two-sample test of proportions

```
cure1: Number of obs = 50
cure2: Number of obs = 59
```

| Variable | Mean  | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|----------|-------|-----------|-------|-------|---------------------|
| cure1    | .52   | .0706541  | .3815205 | .6584795 |
| cure2    | .7118644 | .0589618  | .5963013 | .8274275 |
| diff     | -.1918644 | .0920245  | -.372229 | -.0114998 |
|          | under Ho: | .0931155 | -2.06 | 0.039 |

```
diff = prop(cure1) - prop(cure2)  z = -2.0605
Ho: diff = 0
Ha: diff < 0  Ha: diff != 0  Ha: diff > 0
Pr(Z < z) = 0.0197  Pr(|Z| < |z|) = 0.0394  Pr(Z > z) = 0.9803
```

We find that the proportions are statistically different from each other at any level greater than 3.9%.

Example 3: Immediate form of one-sample test of proportion

```
prtesti
```

is like `prtest`, except that you specify summary statistics rather than variables as arguments. For instance, we are reading an article that reports the proportion of registered voters among 50 randomly selected eligible voters as 0.52. We wish to test whether the proportion is 0.7:

```
prtesti 50 .52 .70
```

One-sample test of proportion

```
x: Number of obs = 50
```

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>.52</td>
<td>.0706541</td>
<td>.3815205</td>
</tr>
<tr>
<td></td>
<td>p = proportion(x)</td>
<td>z = -2.7775</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ho: p = 0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ha: p &lt; 0.7</td>
<td>Ha: p != 0.7</td>
<td>Ha: p &gt; 0.7</td>
</tr>
<tr>
<td></td>
<td>Pr(Z &lt; z) = 0.0027</td>
<td>Pr(</td>
<td>Z</td>
</tr>
</tbody>
</table>

Example 4: Immediate form of two-sample test of proportions

To judge teacher effectiveness, we wish to test whether the same proportion of people from two classes will answer an advanced question correctly. In the first classroom of 30 students, 40% answered the question correctly, whereas in the second classroom of 45 students, 67% answered the question correctly.
. prtesti 30 .4 45 .67
Two-sample test of proportions
x: Number of obs = 30
y: Number of obs = 45

| Variable | Mean  | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|----------|-------|-----------|-------|------|---------------------|
| x        | .4    | .0894427  | .2246955 | .5753045 | [.5753045] |
| y        | .67   | .0700952  | .532616 | .807384  | [.807384] |
| diff     | -.27  | .1136368  | -.4927241 | .0472759 | [.0472759] |

under Ho: .1169416 -2.31 0.021

diff = prop(x) - prop(y)  
z = -2.3088
Ho: diff = 0
Ha: diff < 0
Ha: diff != 0
Ha: diff > 0
Pr(Z < z) = 0.0105  Pr(|Z| < |z|) = 0.0210  Pr(Z > z) = 0.9895

Stored results
prtest and prtesti store the following in r():
Scalars
r(z)  z statistic
r(P_#) proportion for variable #
r(N_#) number of observations for variable #

Methods and formulas
See Acock (2014, 155–161) for additional examples of tests of proportions using Stata.
A large-sample 100(1 − α)% confidence interval for a proportion \(\hat{p}\) is

\[
\hat{p} \pm z_{1-\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}}
\]

and a 100(1 − α)% confidence interval for the difference of two proportions is given by

\[
(\hat{p}_1 - \hat{p}_2) \pm z_{1-\alpha/2} \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}
\]

where \(\hat{q} = 1 - \hat{p}\) and \(z\) is calculated from the inverse cumulative standard normal distribution.

The one-tailed and two-tailed tests of a population proportion use a normally distributed test statistic calculated as

\[
z = \frac{\hat{p} - p_0}{\sqrt{p_0q_0/n}}
\]

where \(p_0\) is the hypothesized proportion. A test of the difference of two proportions also uses a normally distributed test statistic calculated as
\[ z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_p q_p \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \]

where

\[ \hat{p}_p = \frac{x_1 + x_2}{n_1 + n_2} \]

and \( x_1 \) and \( x_2 \) are the total number of successes in the two populations.

References


Also see

[R] bitest — Binomial probability test
[R] proportion — Estimate proportions
[R] ttest — t tests (mean-comparison tests)
[MV] hotelling — Hotelling’s T-squared generalized means test