

nptrend — Test for trend across ordered groups[Description](#)[Quick start](#)[Menu](#)[Syntax](#)[Options](#)[Remarks and examples](#)[Stored results](#)[Methods and formulas](#)[Acknowledgments](#)[References](#)[Also see](#)

Description

`nptrend` performs a nonparametric test for trend across ordered groups.

Quick start

Nonparametric test for a trend in `v1` across ordered categories of `catvar`

```
nptrend v1, by(catvar)
```

As above, but for cases where `v2` is greater than 0

```
nptrend v1 if v2 > 0, by(catvar)
```

Menu

Statistics > Nonparametric analysis > Tests of hypotheses > Trend test across ordered groups

Syntax

```
nptrend varname [if] [in], by(groupvar) [nodetail nolabel score(scorevar) ]
```

Options

Main

by(*groupvar*) is required; it specifies the group on which the data are to be ordered.

nodetail suppresses the listing of group rank sums.

nolabel specifies that group() values be displayed instead of value labels.

score(*scorevar*) defines scores for groups. When it is not specified, the values of *groupvar* are used for the scores.

Remarks and examples

[stata.com](http://www.stata.com)

nptrend performs the nonparametric test for trend across ordered groups developed by [Cuzick \(1985\)](#), which is an extension of the Wilcoxon rank-sum test (see [\[R\] ranksum](#)). A correction for ties is incorporated into the test. nptrend is a useful adjunct to the Kruskal–Wallis test; see [\[R\] kwallis](#).

If your data are not grouped, you can test for trend with the `signtest` and `spearman` commands; see [\[R\] signrank](#) and [\[R\] spearman](#). With `signtest`, you can perform the Cox and Stuart test, a sign test applied to differences between equally spaced observations of *varname*. With `spearman`, you can perform the Daniels test, a test of zero Spearman correlation between *varname* and a time index. See [Conover \(1999, 169–175, 323\)](#) for a discussion of these tests and their asymptotic relative efficiency.

► Example 1

The following data ([Altman 1991, 217](#)) show ocular exposure to ultraviolet radiation for 32 pairs of sunglasses classified into three groups according to the amount of visible light transmitted.

| Group | Transmission of visible light | Ocular exposure to ultraviolet radiation |
|-------|-------------------------------|--|
| 1 | < 25% | 1.4 1.4 1.4 1.6 2.3 2.5 |
| 2 | 25 to 35% | 0.9 1.0 1.1 1.1 1.2 1.2 1.5 1.9 2.2 2.6 2.6 2.6 2.8 2.8 3.2 3.5 4.3 5.1 |
| 3 | > 35% | 0.8 1.7 1.7 1.7 3.4 7.1 8.9 13.5 |

Entering these data into Stata, we have

```
. use http://www.stata-press.com/data/r15/sg
. list, sep(6)
```

| | group | exposure |
|-----|-------------------------|----------|
| 1. | < 25% | 1.4 |
| 2. | < 25% | 1.4 |
| 3. | < 25% | 1.4 |
| 4. | < 25% | 1.6 |
| 5. | < 25% | 2.3 |
| 6. | < 25% | 2.5 |
| 7. | 25% to 35% | .9 |
| | <i>(output omitted)</i> | |
| 31. | > 35% | 8.9 |
| 32. | > 35% | 13.5 |

We use `nptrend` to test for a trend of (increasing) exposure across the three groups by typing

```
. nptrend exposure, by(group)
      group      score      obs  sum of ranks
    < 25%           1         6         76
    25% to 35%       2        18        290
    > 35%           3         8        162
      z = 1.52
    Prob > |z| = 0.129
```

When the groups are given any equally spaced scores (such as $-1, 0, 1$), we will obtain the same answer as above. To illustrate the effect of changing scores, an analysis of these data with scores 1, 2, and 5 (admittedly not sensible here) produces

```
. generate mysc = cond(group==3,5,group)
. nptrend exposure, by(group) score(mysc)
      group      score      obs  sum of ranks
    < 25%           1         6         76
    25% to 35%       2        18        290
    > 35%           5         8        162
      z = 1.46
    Prob > |z| = 0.143
```

This example suggests that the analysis is not all that sensitive to the scores chosen.

◀

□ Technical note

The grouping variable may be either a string variable or a numeric variable. If it is a string variable and no score variable is specified, the natural numbers 1, 2, 3, ... are assigned to the groups in the sort order of the string variable. This may not always be what you expect. For example, the sort order of the strings “one”, “two”, “three” is “one”, “three”, “two”.

□

Stored results

`nptrend` stores the following in `r()`:

Scalars

| | | | |
|-------------------|------------------------|-------------------|----------------|
| <code>r(N)</code> | number of observations | <code>r(z)</code> | z statistic |
| <code>r(p)</code> | two-sided p -value | <code>r(T)</code> | test statistic |

Methods and formulas

`nptrend` is based on a method in [Cuzick \(1985\)](#). The following description of the statistic is from [Altman \(1991, 215–217\)](#). We have k groups of sample sizes n_i ($i = 1, \dots, k$). The groups are given scores, l_i , which reflect their ordering, such as 1, 2, and 3. The scores do not have to be equally spaced, but they usually are. $N = \sum n_i$ observations are ranked from 1 to N , and the sums of the ranks in each group, R_i , are obtained. L , the weighted sum of all the group scores, is

$$L = \sum_{i=1}^k l_i n_i$$

The statistic T is calculated as

$$T = \sum_{i=1}^k l_i R_i$$

Under the null hypothesis, the expected value of T is $E(T) = 0.5(N+1)L$, and its standard error is

$$\text{se}(T) = \sqrt{\frac{N+1}{12} \left(N \sum_{i=1}^k l_i^2 n_i - L^2 \right)}$$

so that the test statistic, z , is given by $z = \{T - E(T)\} / \text{se}(T)$, which has an approximately standard normal distribution when the null hypothesis of no trend is true.

The correction for ties affects the standard error of T . Let \tilde{N} be the number of unique values of the variable being tested ($\tilde{N} \leq N$), and let t_j be the number of times the j th unique value of the variable appears in the data. Define

$$a = \frac{\sum_{j=1}^{\tilde{N}} t_j (t_j^2 - 1)}{N(N^2 - 1)}$$

The corrected standard error of T is $\tilde{\text{se}}(T) = \sqrt{1 - a} \text{se}(T)$.

Acknowledgments

nptrend was written by K. A. Stepniewska, Worldwide Antimalarial Resistance Network, University of Oxford, UK, and D. G. Altman, Botnar Research Center, University of Oxford, UK, (1992).

References

- Altman, D. G. 1991. *Practical Statistics for Medical Research*. London: Chapman & Hall/CRC.
- Conover, W. J. 1999. *Practical Nonparametric Statistics*. 3rd ed. New York: Wiley.
- Cuzick, J. 1985. A Wilcoxon-type test for trend. *Statistics in Medicine* 4: 87–90.
- Sasieni, P. D. 1996. [snp12: Stratified test for trend across ordered groups](#). *Stata Technical Bulletin* 33: 24–27. Reprinted in *Stata Technical Bulletin Reprints*, vol. 6, pp. 196–200. College Station, TX: Stata Press.
- Sasieni, P. D., K. A. Stepniewska, and D. G. Altman. 1996. [snp11: Test for trend across ordered groups revisited](#). *Stata Technical Bulletin* 32: 27–29. Reprinted in *Stata Technical Bulletin Reprints*, vol. 6, pp. 193–196. College Station, TX: Stata Press.
- Stepniewska, K. A., and D. G. Altman. 1992. [snp4: Non-parametric test for trend across ordered groups](#). *Stata Technical Bulletin* 9: 21–22. Reprinted in *Stata Technical Bulletin Reprints*, vol. 2, p. 169. College Station, TX: Stata Press.

Also see

- [R] [epitab](#) — Tables for epidemiologists
- [R] [kwallis](#) — Kruskal–Wallis equality-of-populations rank test
- [R] [signrank](#) — Equality tests on matched data
- [R] [spearman](#) — Spearman’s and Kendall’s correlations
- [R] [symmetry](#) — Symmetry and marginal homogeneity tests
- [ST] [strate](#) — Tabulate failure rates and rate ratios