

nptrend — Tests 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 four different nonparametric tests for trend: the Cochran–Armitage test, the Jonckheere–Terpstra test, the linear-by-linear trend test, and a test using ranks developed by Cuzick.

Quick start

Cochran–Armitage test for the association of a binary outcome in `y` and the ordered values of the categories of `x`

```
nptrend y, group(x) carmitage
```

As above, but report an exact p -value calculated using Monte Carlo permutations

```
nptrend y, group(x) carmitage exact
```

As above, but perform 100,000 Monte Carlo permutations rather than the default of 10,000, and set the random-number seed for reproducibility

```
nptrend y, group(x) carmitage exact(montecarlo, reps(100000) rseed(1234))
```

Rather than testing the trend in `y` by the values of `x`, which are 1, 2, 3, test the trend by 1, 2, 4

```
nptrend y, group(x) carmitage scoregroup(1 2 4)
```

Jonckheere–Terpstra test for trend in `y` across ordered categories of `catvar`

```
nptrend y, group(catvar) jterpstra
```

As above, but perform the linear-by-linear trend test

```
nptrend y, group(catvar) linear
```

As above, but perform Cuzick’s test using ranks

```
nptrend y, group(catvar) cuzick
```

Menu

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

Syntax

```
nptrend varname [if] [in] [weight] , group(groupvar) testopt [options]
```

<i>options</i>	Description
Main	
* <u>g</u> roup(<i>groupvar</i>)	ordered group variable
<u>s</u> coreresponse(<i>numlist</i>)	scores for the response <i>varname</i>
<u>s</u> coregroup(<i>numlist</i>)	scores for the groups specified by <i>groupvar</i>
<u>e</u> xact [<i>(exact_specs)</i>]	report an exact <i>p</i> -value
<u>n</u> olabel	display numerical values of <i>groupvar</i> rather than value labels
<u>n</u> otable	do not display the table of mean response scores by group
* <i>testopt</i>	
Description	
<u>c</u> armitage	Cochran–Armitage test for trend
<u>j</u> terpstra	Jonckheere–Terpstra test for trend
<u>l</u> inear	linear-by-linear test for trend
<u>c</u> uzick	use ranks for the response scores in the test due to Cuzick

*group(*groupvar*) and one of the choices for *testopt* are required.

collect is allowed; see [U] 11.1.10 Prefix commands.

fweights are allowed, except when exact is specified; see [U] 11.1.6 weight.

Options

Main

group(*groupvar*) defines the ordered groups across which differences in the response *varname* are to be tested. group() is required.

When option scoregroup() is not specified, the values of *groupvar* are the group scores. When scoregroup() is specified, *groupvar* defines the ordered categories, and the scores are the values specified in scoregroup(). When scoregroup() is specified, *groupvar* must be a positive integer (with gaps allowed).

testopt specifies the test that is reported. It is required and must be one of carmitage, jterpstra, linear, or cuzick.

carmitage specifies that the Cochran–Armitage test for trend be reported. Typically, the response *varname* is a 0/1 variable, and the values of *groupvar* represent a set of ordered categories. It gives a test of the trend of the proportions of positive responses in the groups, with the hypothesized trend given by the values of *groupvar* (or the values of scoregroup() if specified). If the response *varname* is not 0/1, it can be valued 1/2 and mapped to 0/1 or 1/0 using scoreresponse(0 1) or scoreresponse(1 0), respectively.

Response and group specifications for carmitage can be flipped. The *groupvar* variable can be a 0/1 predictor, and the response *varname* can contain ordered outcomes. If *groupvar* is not 0/1, it can be valued 1/2 and mapped to 0/1 or 1/0 using scoregroup(0 1) or scoregroup(1 0), respectively.

`jterpstra` specifies that the Jonckheere–Terpstra test for trend be reported. The response *varname* contains outcomes, which can be ordered categories or continuous values, and the *groupvar* variable is an ordered group indicator that is believed to be a predictor of outcome. The Jonckheere–Terpstra statistic is dependent on which is the response variable and which is the group variable. If response and group variables are interchanged, the value of the statistic will be different.

Only the orderings given by the response *varname* and *groupvar* are used to calculate the Jonckheere–Terpstra statistic. Different numerical values that give the same orderings produce the same statistic. Hence, this statistic is used when a test for trend dependent only on order is desired. Typically, the options `scorerresponse()` and `scoregroup()` are not specified. However, either or both of these options can be used to change the orderings given by *varname* and *groupvar*.

`linear` specifies that the linear-by-linear test for trend be reported. The linear-by-linear statistic is symmetric in the response variable *varname* and the group indicator *groupvar*. Interchanging response and group variables produces the same result. Despite the name “linear”, nonlinear trends can be tested. The trend, linear or nonlinear, is based on the numerical values of the response and group variables. `scorerresponse()` and `scoregroup()` can be specified to test different trends.

`cuzick` specifies that ranks be used for the responses and that the test for trend across ordered groups developed by Cuzick (1985) be reported. Because ranks are used, only the ordering given by response variable *varname* matters. The numerical values of *groupvar* are used in calculation, and different values, even those that give the same ordering, will result in different values of the statistic for Cuzick’s test. `scorerresponse()` can be specified to change the ordering of the responses. `scoregroup()` can be specified to test different trends by group.

`scorerresponse(numlist)` specifies scores for the responses in *varname*. When specified, *varname* must contain only positive integers (with gaps allowed), and the response score for *varname* = *i* is the *i*th number in *numlist*. When `scorerresponse()` is not specified, the response scores are the values of the response *varname*.

`scoregroup(numlist)` specifies scores for the groups specified by *groupvar*. When specified, *groupvar* must contain only positive integers (with gaps allowed), and the group score for *groupvar* = *i* is the *i*th number in *numlist*. When `scoregroup()` is not specified, the group scores are the values of *groupvar*.

`exact` and `exact(exact_specs)` specify that an exact *p*-value be reported.

`exact` specifies that an exact *p*-value from a Monte Carlo permutation test be reported. `exact` is a synonym for `exact(montecarlo)`.

`exact(montecarlo|enumerate[, options])` specifies that an exact *p*-value be reported in addition to the approximate or asymptotic *p*-value. Specifying `exact(montecarlo)` does a Monte Carlo permutation test. Specifying `exact(enumerate)` does an enumeration of all possible permutations. Because the number of all possible permutations is typically extremely large, enumeration is feasible only for very small datasets. The number of permutations will be displayed, and you can click on *Break* to stop the computation. The exact *p*-value is computed by `permute`.

`exact(montecarlo[, options])` allows *options* `show`, `reps(#)`, `rseed(#)`, `saving(filename[, sav_options])`, `level(#)`, `dots(#)`, `nodots`, and `eps(#)`. The `show` option specifies that the table produced by `permute` also be displayed. By default, 10,000 Monte Carlo permutations are done. That is, the default is the same as specifying `exact(montecarlo, reps(10000))`. The default for `dots()` is `dots(100)` when `reps()` is $\geq 10,000$; otherwise, it is `dots(1)`. See *Options* in [R] [permute](#).

`exact(enumerate[, options])` allows *options* `show`, `saving(filename[, sav_options])`, `dots(#)`, `nodots`, and `eps(#)`. The `show` option specifies that the table produced by `permute` also be displayed. The default for `dots()` is `dots(100)`. See *Options* in [R] [permute](#).

`noLabel` causes numerical values of *groupvar* to be displayed in the table of mean response scores rather than value labels.

`notable` suppresses the display of the table of mean response scores by group.

Remarks and examples

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

Remarks are presented under the following headings:

- [Introduction](#)
- [Cochran–Armitage test](#)
- [Jonckheere–Terpstra test](#)
- [Linear-by-linear trend test](#)
- [Cuzick’s test](#)

Introduction

`nptrend` performs four different nonparametric tests for trend: the Cochran–Armitage test, the Jonckheere–Terpstra test, the linear-by-linear trend test, and a test using ranks developed by Cuzick.

Data for these tests consist of two variables. The first variable, specified immediately after `nptrend`, contains responses. The second variable, specified in `group()`, identifies groups.

For the Cochran–Armitage test, the response is typically a 0/1 variable, and the values of the group variable are ordered categories. It tests the trend of the proportions of positive responses in the groups. The numerical values of the hypothesized trend are called “scores”. For this form of the test, we call the scores that specify the hypothesized trend “group scores” because it is the trend by group being tested. For example, group scores of 1, 2, 3 would test for a linear trend. Group scores of 1, 4, 9 would test for a quadratic trend.

The specification of the Cochran–Armitage test can be flipped. That is, the response variable can identify ordered outcome categories, and the group variable can be 0/1. In this case, we call the scores that specify the hypothesized trend “response scores”.

The Cochran–Armitage test requires either a 0/1 response variable or a 0/1 group variable. It cannot test responses with more than two levels when there are more than two groups.

The Jonckheere–Terpstra test is typically used when there are more than two groups and more than two response levels. The test assesses whether there is an association between the response scores and the group scores, but only the relative orders of the responses and groups matter. Because there is no need to specify a hypothesized trend, this test is in this sense “more nonparametric” than the other tests for trend reported by `nptrend`.

The linear-by-linear trend test uses response scores to specify the trend being tested. So it is typically used when the trend of the response scores, rather than just their ordering, is meaningful. The linear-by-linear statistic is asymptotically equivalent to the Pearson correlation coefficient (see [R] [correlate](#)).

The statistics for the Cochran–Armitage test (when the response is 0/1), the linear-by-linear trend test, and Cuzick’s test are based on the numerical values of the group scores. The linear-by-linear test requires numerical response scores as well. Different scores, even those that have the same ordering, produce different values of the statistic. For instance, group scores 1, 2, 3 would give a different result than 1, 2, 4. For the linear-by-linear test, the same holds for response scores.

The options `scoreresponse()` and `scoregroup()` are provided as convenience tools to test different response scores and group scores without having to modify the underlying variables. They do, however, require that the underlying variables be valued 1, 2, 3, . . .

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.

Cochran–Armitage test

The Cochran–Armitage test tests the trend in ordered groups of the probability of a positive response. As we stated earlier, response and group specifications can be flipped, and it can assess whether a trend based on ordered responses is associated with membership in one of two groups. In the discussion that follows we will consider only the former case: ordered groups with 0/1 responses.

The Cochran–Armitage test is based on a linear probability model for θ_i , the probability of a positive response for subjects in the i th group,

$$\theta_i = \alpha + \beta(g_i - \bar{g})$$

where α is the intercept, β is the slope, g_i are the group scores, and \bar{g} is the mean of the group scores.

The estimate for β is the ordinary least-squares (OLS) estimate $\hat{\beta}$, and this is the statistic for the Cochran–Armitage test.

However, the standard error of the statistic $\hat{\beta}$ for the Cochran–Armitage test has a different formulation than that of the standard error from OLS regression—although they are the same asymptotically. The standard error $se(\hat{\beta})$ for the Cochran–Armitage statistic is based on decomposing the Pearson statistic for independence.

The Pearson statistic X_{Ind}^2 for independence can be written as

$$X_{\text{Ind}}^2 = z^2 + X_{\text{DL}}^2$$

where $z = \hat{\beta}/se(\hat{\beta})$ is the z score for the Cochran–Armitage test and X_{DL}^2 is a statistic for departure from the linear trend. Asymptotically, z has a standard normal distribution, and X_{DL}^2 has a χ^2 distribution with $I - 2$ degrees of freedom, where I is the number of ordered categories. See [Cochran–Armitage test for trend](#) in *Methods and formulas* below.

This is one of the advantages of performing the Cochran–Armitage test. We get not only a test for trend but also a test for departure from this trend. Let’s be clear about what we mean when we say “departure from this trend”. The Pearson statistic for independence is decomposed into two pieces. Either piece being large means that the test for independence is rejected. The linear piece z^2 being large means there is a linear trend that rejects independence. The departure piece X_{DL}^2 being large means there are differences other than the linear trend that reject independence.

We should also be clear about what we mean when we say “linear trend”. Linear trend means whatever trend is given by the group scores g_i , which do not have to be 1, 2, 3, They could be, for instance, 1, 3, 10, 12, It is linear in the scores but not necessarily linear by the ordering of the groups.

When the `exact` option is specified, a p -value from the permutation test of the null hypothesis $\beta = 0$ is reported.

▷ Example 1

We have fictional data from a clinical trial of a new drug for treating migraines. The variable `dose` contains the dose of the drug given to a subject. The variable `relief` is 0/1 with 0 indicating no relief of the migraine and 1 indicating partial or total relief. Here is a tabulation of the data:

```
. use https://www.stata-press.com/data/r17/migraine
(Fictional migraine drug data)
. tabulate dose relief, row nokey
```

Mycureit dose in mg	Relief of migraine after 2 hours		Total
	0	1	
10	80 40.00	120 60.00	200 100.00
20	92 46.00	108 54.00	200 100.00
30	83 41.50	117 58.50	200 100.00
40	63 31.50	137 68.50	200 100.00
Total	318 39.75	482 60.25	800 100.00

There appears to be a trend of more subjects reporting relief at the higher doses. Let's calculate the Cochran–Armitage test for trend:

```
. nptrend relief, group(dose) carmitage
Cochran–Armitage test for trend
  Number of observations =      800
    Number of groups =         4
Number of response levels =         2
```

Group	Group score	Mean response score	Number of obs
dose			
10	10	.6	200
20	20	.54	200
30	30	.585	200
40	40	.685	200

```
Statistic = .003
Std. err. = .0015476
z = 1.939
Prob > |z| = 0.0526
```

```
Test of departure from trend:
  chi2(2) = 5.45
Prob > chi2 = 0.0656
```

The asymptotic p -value for the test of a linear trend is 0.0526. The p -value for the test of departure from linearity is 0.0656.

The p -value for the test of a linear trend is close to 0.05. Rather than report the asymptotic p -value, we may want to report an exact p -value. We run `nptrend` again, this time specifying the `exact` option.

```
. nptrend relief, group(dose) carmitage exact
Permutations (10,000): .....1,000.....2,000.....3,000.....
> 4,000.....5,000.....6,000.....7,000.....8,000.....
> 9,000.....10,000 done
```

Cochran-Armitage test for trend

```
Number of observations = 800
Number of groups = 4
Number of response levels = 2
```

Group	Group score	Mean response score	Number of obs
dose			
10	10	.6	200
20	20	.54	200
30	30	.585	200
40	40	.685	200

```
Statistic = .003
Std. err. = .0015476
z = 1.939
Prob > |z| = 0.0526
Exact prob = 0.0576 (10,000 Monte Carlo permutations)
```

Test of departure from trend:

```
chi2(2) = 5.45
Prob > chi2 = 0.0656
```

The exact p -value reported is 0.0576, larger than the asymptotic p -value of 0.0526. The exact p -value comes from a permutation test performed by computing 10,000 random permutations. This means that there is some random error associated with the reported exact p -value.

To see the Monte Carlo error in the exact p -value, we can use the `show` suboption of `exact`, which displays the full output of `permute`. The option we specify is

```
exact(montecarlo, show reps(100000) dots(1000) rseed(1234))
```

We increase the number of random permutations to 100,000. `dots(1000)` displays a dot after every 1,000th permutation rather than the default of every 100th. `rseed(1234)` sets the random-number seed so we can reproduce our result. We did not set the random-number seed when we ran `nptrend ... exact` earlier. If we were to run that again, we would get a different value for the exact p -value. Here are the results with `exact()` specified this new way:


```
. nptrend relief, group(dose) carmitage
> exact(montecarlo, show reps(100000) dots(1000) rseed(1234))
Permutations (100,000): .....10,000.....20,000.....30,000.....
> ...40,000.....50,000.....60,000.....70,000.....80,000....
> .....90,000.....100,000 done
Monte Carlo permutation results                Number of observations =    800
Permutation variable: relief                  Number of permutations = 100,000
```

T	T(obs)	Test	c	n	p	Monte Carlo error		
						SE(p)	[95% CI(p)]	
_pm_1	.003	lower	97435	100000	.9744	.0005	.9734	.9753
		upper	2959	100000	.0296	.0005	.0285	.0307
		two-sided			.0592	.0007	.0577	.0606

Notes: For lower one-sided test, $c = \#\{T \leq T(\text{obs})\}$ and $p = p_{\text{lower}} = c/n$.
 For upper one-sided test, $c = \#\{T \geq T(\text{obs})\}$ and $p = p_{\text{upper}} = c/n$.
 For two-sided test, $p = 2 \cdot \min(p_{\text{lower}}, p_{\text{upper}})$; SE and CI approximate.

Cochran-Armitage test for trend

```
Number of observations =    800
Number of groups =      4
Number of response levels =  2
```

Group	Group score	Mean response score	Number of obs
dose			
10	10	.6	200
20	20	.54	200
30	30	.585	200
40	40	.685	200

```
Statistic =    .003
Std. err. = .0015476
z =    1.939
Prob > |z| =    0.0526
Exact prob =    0.0592 (100,000 Monte Carlo permutations)
```

Test of departure from trend:

```
chi2(2) =    5.45
Prob > chi2 =    0.0656
```

The reported exact p -value from this specification is 0.0592, and we see from the table reported by `permute` that the Monte Carlo 95% confidence interval for it is [0.0577, 0.0606]. Clearly, the asymptotic p -value 0.0526 is a bit anticonservative. If we were going to publish the result, we might want to run the command again with even more permutations, perhaps 1,000,000 or more, to narrow the width of the confidence level to less than 0.001. With only 800 observations in these data, the Monte Carlo permutations are computed quickly.

▷ Example 2

We continue with the [previous example](#) to illustrate the use of the `scoregroup()` option. When we tested the linear trend of dose, we got a p -value of 0.0656 for the departure from linearity. So maybe the trend is not linear. Let's test a quadratic trend.

To use the `scoregroup()` option, the variable in `group()` must be 1, 2, ... The variable dose is 10, 20, 30, 40. We can convert it to 1, 2, 3, 4 in several ways. Here we use the `egen group()` function because this works in general.

```
. egen idose = group(dose)
. tabulate idose
```

group(dose)	Freq.	Percent	Cum.
1	200	25.00	25.00
2	200	25.00	50.00
3	200	25.00	75.00
4	200	25.00	100.00
Total	800	100.00	

Now, we specify `scoregroup(1 4 9 16)` to test a quadratic trend.

```
. nptrend relief, group(idose) carmitage scoregroup(1 4 9 16)
```

Cochran-Armitage test for trend

```
Number of observations = 800
Number of groups = 4
Number of response levels = 2
```

Group	Group score	Mean response score	Number of obs
idose			
1	1	.6	200
2	4	.54	200
3	9	.585	200
4	16	.685	200

```
Statistic = .0070543
Std. err. = .0030468
z = 2.315
Prob > |z| = 0.0206
```

Test of departure from trend:

```
chi2(2) = 3.85
Prob > chi2 = 0.1462
```

Of course, we could have just typed `generate dose2 = dose^2` and used this variable for `group()` to test the quadratic trend.

The p -value for the quadratic trend is 0.0206, and the departure p -value is 0.1462. The p -value for the linear trend was 0.0526. So it appears the relationship between dose and relief is closer to quadratic than linear. But should we be using `nptrend` to search for the trend with the smallest p -value? Surely, `logit` is the tool for that!

```

. generate dose2 = dose^2
. logit relief dose dose2
Iteration 0:   log likelihood = -537.58798
Iteration 1:   log likelihood = -532.96973
Iteration 2:   log likelihood = -532.96346
Iteration 3:   log likelihood = -532.96346

Logistic regression                                Number of obs =    800
LR chi2(2)    =    9.25
Prob > chi2   = 0.0098
Pseudo R2    = 0.0086

Log likelihood = -532.96346

```

relief	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
dose	-.0720094	.0367396	-1.96	0.050	-.1440177	-1.19e-06
dose2	.0017005	.0007276	2.34	0.019	.0002744	.0031266
_cons	.946503	.4021702	2.35	0.019	.1582638	1.734742

To be clear, we are not recommending a model such as this for these data. There are only four doses, and this logit model fits three of the four degrees of freedom. The purpose of the Cochran–Armitage test is to test a single hypothesized trend and report an honest p -value for it. If we want to test for a trend but do not have any idea what the trend might be, we may be better off using the Jonckheere–Terpstra test.

◀

Jonckheere–Terpstra test

The Jonckheere–Terpstra test is useful when it is not clear what the trend might be and we simply want to test for any trend. That is, it tests whether the ordering of the groups is associated with the ordering of the responses. The test is typically used when there are many response levels and any number of groups.

Suppose there are $J \geq 2$ ordered groups. Let y_{jk} be the k th response in the j th group, where $k = 1, 2, \dots, n_j$. The responses y_{jk} are given by *varname* (the variable specified immediately after `nptrend`) or by *varname* mapped to the values of `scorerresponse()`.

For groups j and j' , where $j < j'$, consider the $n_j n_{j'}$ pairs

$$(y_{jk}, y_{j'k'})$$

Let $C_{jj'}$ be the number of pairs such that $y_{jk} < y_{j'k'}$ and $D_{jj'}$ be the number of pairs such that $y_{jk} > y_{j'k'}$. The Jonckheere–Terpstra statistic is

$$T = \sum_{j < j'} (C_{jj'} - D_{jj'})$$

In other words, count the number of response pairs, $C_{jj'}$, where the group j response is less than the group j' response. That is, pairs in which the ordering of the responses is concordant with the ordering of the groups. Subtract from this the number of discordant pairs, $D_{jj'}$. Then, sum over the $J(J-1)/2$ combinations of selecting two of the J groups, where group j is ordered before group j' .

Clearly, only the ordering of responses matters relative to the ordering of the groups. There is no need to hypothesize what the trend is.

Kendall's τ (see [R] [spearman](#)) also uses only relative differences, and the Jonckheere–Terpstra test can be viewed as a generalization of Kendall's τ for groups of unequal sizes. See *Jonckheere–Terpstra test for trend* in *Methods and formulas* below.

▷ Example 3

The following data (Altman 1991, 217) show ocular exposure to ultraviolet radiation for 32 pairs of sunglasses classified into 3 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
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 https://www.stata-press.com/data/r17/sg, clear
(Ultraviolet radiation exposure with sunglasses)
. list, separator(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
<hr/>		
7.	25% to 35%	.9
8.	25% to 35%	1
<i>(output omitted)</i>		
31.	> 35%	8.9
32.	> 35%	13.5

We can now use `nptrend` to report the Jonckheere–Terpstra test results.

```
. nptrend exposure, group(group) jterpstra
Jonckheere–Terpstra test for trend
Number of observations = 32
Number of groups = 3
Number of response levels = 23
```

Group	Group score	Mean response score	Number of obs
group < 25%	1	1.766667	6
25% to 35%	2	2.311111	18
> 35%	3	4.85	8

```
Statistic = 82
Std. err. = 54.80056
z = 1.496
Prob > |z| = 0.1346
```

The approximate p -value is 0.1346. This p -value is an approximation to the permutation test p -value. That is, it comes from a z statistic that uses the exact variance of the permutation distribution. With only 32 observations, it seems better to simply compute the exact p -value using Monte Carlo permutations.

```
. set seed 1234
. nptrend exposure, group(group) jterpstra exact
Permutations (10,000): .....1,000.....2,000.....3,000.....
> 4,000.....5,000.....6,000.....7,000.....8,000.....
> 9,000.....10,000 done
Jonckheere-Terpstra test for trend
  Number of observations =      32
    Number of groups     =       3
Number of response levels =      23
```

Group	Group score	Mean response score	Number of obs
group < 25%	1	1.766667	6
25% to 35%	2	2.311111	18
> 35%	3	4.85	8

```
Statistic =      82
Std. err. = 54.80056
      z =    1.496
Prob > |z| =    0.1346
Exact prob =    0.1378 (10,000 Monte Carlo permutations)
```

Despite the small number of observations, the Monte Carlo exact p -value 0.1378 is close to the approximate p -value 0.1346. Because of the small number of observations, we might be tempted to try the `exact(enumerate)` option. We do so.

```
. nptrend exposure, group(group) jterpstra exact(enumerate, dots(1000))
(enumerating all 1.42e+12 possible permutations)
Permutations (1,415,721,106,800): .....10,000.....—Break—
r(1);
```

`nptrend` calls `permute` to do the enumeration, and `permute` tells us there are 1.42×10^{12} possible permutations. Because we do not want to wait years for the result, we pressed `Break`.

Let's randomly cut the dataset in half, from 32 observations to 16.

```
. set seed 25
. generate r = runiform()
. sort r
. keep in 1/16
(16 observations deleted)
. tabulate group
```

Transmission of visible light	Freq.	Percent	Cum.
< 25%	2	12.50	12.50
25% to 35%	9	56.25	68.75
> 35%	5	31.25	100.00
Total	16	100.00	

Then, let's do the enumeration for these data:

```
. nptrend exposure, group(group) jterpstra exact(enumerate, dots(1000))
(enumerating all 240,240 possible permutations)
Permutations (240,240): .....10,000.....20,000.....30,000.....
> ...40,000.....50,000.....60,000.....70,000.....80,000....
> .....90,000.....100,000.....110,000.....120,000.....130,
> 000.....140,000.....150,000.....160,000.....170,000.....
> .....180,000.....190,000.....200,000.....210,000.....220,0
> 00.....230,000.....240,000.240,240 done
Jonckheere–Terpstra test for trend
      Number of observations =      16
      Number of groups      =       3
Number of response levels =      15
```

Group	Group score	Mean response score	Number of obs
group < 25%	1	1.5	2
25% to 35%	2	2.377778	9
> 35%	3	6.92	5

```
Statistic =      49
Std. err. = 19.57188
      z =      2.504
Prob > |z| =      0.0123
Exact prob =      0.0103 (enumerated all 240,240 permutations)
```

There are only 240,240 possible permutations now, and the computation takes little time. Despite there being only 16 observations, the exact p -value (without any error) 0.0103 is close to the approximate p -value 0.0123. Enumeration works quite well for tiny datasets such as this.

◀

Linear-by-linear trend test

The linear-by-linear trend test is an alternative to the Jonckheere–Terpstra test. The difference is that the linear-by-linear trend test uses response scores to specify the trend being tested. How the trend is hypothesized to vary across groups is specified by the group scores. The linear-by-linear trend test could be described as a parametric specification for trend that is tested nonparametrically!

If scores are not specified in the options, then the values of the response *varname* are used as response scores, and the values of *groupvar* are used as group scores.

The linear-by-linear statistic is

$$T = \frac{1}{N} \sum_{i=1}^I \sum_{j=1}^J n_{ij} (g_j - \bar{g})(r_i - \bar{r})$$

where N is the total number of observations, g_j is the group score for the j th group, \bar{g} is the mean group score over all observations, r_i is the response score for the i th ordered response, and \bar{r} is the mean response score over all observations.

The linear-by-linear statistic is equivalent to the Pearson correlation coefficient (see [R] [correlate](#)), the difference being that the Pearson correlation coefficient is standardized by the standard deviations of the scores. The p -values are slightly different because the p -value for the linear-by-linear test is based on its permutation distribution while the p -value for the Pearson correlation coefficient assumes normality.

Clearly, the statistic depends on the numerical values of both the response scores and the group scores. Different scores, even ones that give the same ordering, will produce different values of the statistic.

▷ Example 4

We again use the sunglasses data from [example 3](#), where we did the Jonckheere–Terpstra test.

```
. use https://www.stata-press.com/data/r17/sg, clear
(Ultraviolet radiation exposure with sunglasses)
```

We calculate the linear-by-linear trend test for these data.

```
. nptrend exposure, group(group) linear
Linear-by-linear test for trend
Number of observations = 32
Number of groups = 3
Number of response levels = 23
```

Group	Group score	Mean response score	Number of obs
group < 25%	1	1.766667	6
25% to 35%	2	2.311111	18
> 35%	3	4.85	8

```
Statistic = .7035156
Std. err. = .3063377
z = 2.297
Prob > |z| = 0.0216
```

The approximate p -value is 0.0216, which is much smaller than the Jonckheere–Terpstra test approximate p -value 0.1346. Such a difference is not surprising. The linear-by-linear trend test uses the actual values of the response in the statistic. The values of the response seem to follow a linear trend by group, and when this is true, one generally expects the p -value from the linear-by-linear trend test to be smaller than the p -value from the Jonckheere–Terpstra test.

Again, we might want to compute the exact p -value using Monte Carlo permutations.

```
. nptrend exposure, group(group) linear
> exact(montecarlo, reps(100000) dots(1000) rseed(1234))
Permutations (100,000): .....10,000.....20,000.....30,000.....
> ...40,000.....50,000.....60,000.....70,000.....80,000....
> .....90,000.....100,000 done
Linear-by-linear test for trend
  Number of observations =      32
    Number of groups     =       3
Number of response levels =      23
```

Group	Group score	Mean response score	Number of obs
group < 25%	1	1.766667	6
25% to 35%	2	2.311111	18
> 35%	3	4.85	8

```

Statistic = .7035156
Std. err. = .3063377
      z = 2.297
Prob > |z| = 0.0216
Exact prob = 0.0146 (100,000 Monte Carlo permutations)
```

In this case, the exact p -value 0.0146 is smaller than the approximate p -value 0.0216.

◀

Cuzick's test

Cuzick's test (Cuzick 1985) is an extension of the Kruskal–Wallis test (see [R] [kwallis](#)) for ordered groups. The response scores are ranked, the sum of the ranks is calculated for each group, and then these sums of ranks are weighted by the deviation of the group scores from their mean.

The statistic for Cuzick's test is

$$T = \frac{1}{N} \sum_{j=1}^J (g_j - \bar{g}) R_j$$

where N is the total number of observations, g_j is the group score for the j th group, \bar{g} is the mean group score over all observations, and R_j is the sum of the ranks of the response scores in the j th group.

The form of the statistic is similar to that of the statistic for the linear-by-linear test, but instead of using response scores being used directly, the ranks of the response scores are used. So only the ordering of the response scores matters. The numerical values of the group scores, however, are used in the calculation, so different group scores will give different values of the statistic.

▷ Example 5

We continue with the sunglasses data from [example 4](#). We now calculate Cuzick's test, calculating the exact p -value as well.

```
. nptrend exposure, group(group) cuzick
> exact(montecarlo, reps(100000) dots(1000) rseed(1234))
Permutations (100,000): .....10,000.....20,000.....30,000.....
> ...40,000.....50,000.....60,000.....70,000.....80,000....
> .....90,000.....100,000 done
Cuzick's test with rank scores
  Number of observations =      32
  Number of groups      =       3
Number of response levels =      23
```

Group	Group score	Mean response score	Number of obs
group < 25%	1	1.766667	6
25% to 35%	2	2.311111	18
> 35%	3	4.85	8

```
Statistic = 1.65625
Std. err. = 1.090461
z = 1.519
Prob > |z| = 0.1288
Exact prob = 0.1322 (100,000 Monte Carlo permutations)
```

The approximate p -value 0.1288 is close to the exact p -value 0.1322. These results are also close to the results from the Jonckheere–Terpstra test, where the approximate p -value was 0.1346. This is not surprising because using ranks of the responses in Cuzick's test is likely to be similar to the comparisons of orderings in the Jonckheere–Terpstra test. The latter test is better known than Cuzick's test, so for this reason alone, it may be preferable to use the Jonckheere–Terpstra test when reporting results.

◀

Stored results

nptrend stores the following in `r()`:

Scalars

```
r(N)          number of observations
r(n_groups)   number of groups
r(n_rsp_levels) number of response levels
r(beta)       slope from Cochran–Armitage linear probability model
r(T)          statistic
r(se)         standard error of the statistic
r(z)          z statistic
r(p)          two-sided  $p$ -value from  $z$  statistic
r(p_l)        lower one-sided  $p$ -value from  $z$  statistic
r(p_u)        upper one-sided  $p$ -value from  $z$  statistic
r(chi2_depart)  $\chi^2$  for departure from Cochran–Armitage linear probability model
r(df_depart)  degrees of freedom for departure  $\chi^2$ 
r(p_depart)    $p$ -value for departure  $\chi^2$ 
r(p_exact)    two-sided exact  $p$ -value
```

<code>r(p_l_exact)</code>	lower one-sided exact p -value
<code>r(p_u_exact)</code>	upper one-sided exact p -value
<code>r(n_perm)</code>	number of permutations performed

Macros

<code>r(test)</code>	"carmitage", "jterpstra", "linear", or "cuzick"
<code>r(group)</code>	group variable
<code>r(exact)</code>	"montecarlo" or "enumerate"
<code>r(rngstate)</code>	random-number state used for Monte Carlo permutations

Matrices

<code>r(table)</code>	table of mean response scores by group
-----------------------	--

If `exact(..., show)` is specified, the stored results from `permute` are returned as well; see [Stored results](#) in [\[R\] permute](#).

Methods and formulas

Methods and formulas are presented under the following headings:

- [Overview](#)
- [Cochran–Armitage test for trend](#)
- [Jonckheere–Terpstra test for trend](#)
- [Linear-by-linear test for trend](#)
- [Cuzick’s test with rank scores](#)

Overview

We will use the same notation for all the tests. Let the set of distinct responses be indexed by $i = 1, 2, \dots, I$. Let the groups be indexed by $j = 1, 2, \dots, J$.

We call the scores for the responses “response scores” and denote them by r_i .

For the Cochran–Armitage test (when the response is 0/1), linear-by-linear trend test, and Cuzick’s test the groups have scores as well. We call these scores “group scores”, and denote them by g_j .

All the tests for trend reported by `nptrend` are based on tests of independence between the ordering of the responses (or their scores) and the ordering of the groups (or their scores). Because these are tests of independence, there is no need to hypothesize a direction of causal inference. Responses could be outcomes predicted by group membership, or “groups” could be outcomes predicted by what we are calling “responses”.

Let n_{ij} be the number of observations in the i th response category and j th group. Let $N = \sum_{ij} n_{ij}$ be the total number of observations. Let $n_{i+} = \sum_j n_{ij}$ be the row margins, and let $n_{+j} = \sum_i n_{ij}$ be the column margins.

Cochran–Armitage test for trend

When `carmitage` is specified, the Cochran–Armitage test is performed (Cochran 1954; Armitage 1955; Agresti 2013). It assesses whether the probability of a positive response is associated with a trend based on ordered groups. Or the specification can be flipped. It can be used to assess whether a trend based on ordered responses is associated with membership in one of two groups.

In these methods, we assume that the response is 0/1—that is, the response is indexed by $i = 0, 1$ —and that there are J ordered groups with scores g_j . When the specification is flipped, response and group are interchanged in the equations below.

The Cochran–Armitage test is based on a linear probability model for θ_j , the probability of having response $i = 1$ for the subjects in the j th group,

$$\theta_j = \alpha + \beta (g_j - \bar{g})$$

where α is the intercept, β is the slope, g_j are the group scores, and \bar{g} is the mean of the group scores.

The equation is solved using OLS,

$$\hat{\theta}_j = p + \hat{\beta} (g_j - \bar{g})$$

where $p = (\sum_j n_{1j})/N$ and $\bar{g} = (\sum_j n_{+j}g_j)/N$.

The OLS solution for $\hat{\beta}$ is

$$\hat{\beta} = \frac{\sum_j n_{+j}(p_j - p)(g_j - \bar{g})}{\sum_j n_{+j}(g_j - \bar{g})^2}$$

where $p_j = n_{1j}/n_{+j}$.

The test for independence between groups and responses can be separated into a test for the linear trend and a test for departure from the linear trend. The Pearson statistic for independence

$$X_{\text{Ind}}^2 = \frac{1}{p(1-p)} \sum_{j=1}^J n_{+j}(p_j - p)^2$$

is decomposed as

$$X_{\text{Ind}}^2 = z^2 + X_{\text{DL}}^2$$

where

$$z^2 = \frac{\hat{\beta}^2}{p(1-p)} \sum_{j=1}^J n_{+j}(g_j - \bar{g})^2$$

and

$$X_{\text{DL}}^2 = \frac{1}{p(1-p)} \sum_{j=1}^J n_{+j}(p_j - \hat{\theta}_j)^2$$

z is the z statistic for the linear trend, and X_{DL}^2 is the statistic for departure from a linear trend. Asymptotically, z has a standard normal distribution, and X_{DL}^2 has a χ^2 distribution with $J - 2$ degrees of freedom.

Jonckheere–Terpstra test for trend

When `jterpstra` is specified, the Jonckheere–Terpstra test is performed (Terpstra 1952; Jonckheere 1954). The Jonckheere–Terpstra test is a test of whether the hypothesized ordering of the responses is associated with the hypothesized ordering of the groups. Typically, this test is used when there are only a few groups but many response categories or even continuous responses.

Only the ordering of the responses and groups matter. The numerical values of the scores do not matter, but scores can be specified to change the ordering.

Let y_{jk} be the k th response in the j th group, where $k = 1, 2, \dots, n_{+j}$. If the response falls in the i th response category, then $y_{jk} = r_i$, the response score for the i th category.

For groups j and j' , where $j < j'$, consider the pairs

$$(y_{jk}, y_{j'k'})$$

where $k = 1, 2, \dots, n_{+j}$ and $k' = 1, 2, \dots, n_{+j'}$. A total of $n_{+j}n_{+j'}$ pairs are formed for these two groups.

Let $C_{jj'}$ be the number of pairs such that $y_{jk} < y_{j'k'}$, that is, pairs in which the response ordering is concordant with the group ordering. Let $D_{jj'}$ be the number of discordant pairs with $y_{jk} > y_{j'k'}$. The Jonckheere–Terpstra statistic is

$$T = \sum_{j=1}^{J-1} \sum_{j'=j+1}^J (C_{jj'} - D_{jj'})$$

The Jonckheere–Terpstra statistic is sometimes expressed as sums over $C_{jj'} + E_{jj'}/2$, where $E_{jj'}$ is the number of tied pairs with $y_{jk} = y_{j'k'}$. Because $C_{jj'} + D_{jj'} + E_{jj'} = n_{+j}n_{+j'}$, the statistic T is twice the statistic based on $C_{jj'} + E_{jj'}/2$ minus a constant. The advantage of the formulation using $C_{jj'} - D_{jj'}$ is that under the null hypothesis of independence between groups and responses, its expectation is zero. When T is positive, the observed trend is a positive trend between the group ordering and the response ordering, and when T is negative, it is a negative trend.

The variance of T can be computed exactly based on the permutation distribution of responses where the table margins n_{i+} and n_{+j} are fixed and each permutation is considered equally likely under the null hypothesis. The formula, however, is complicated when there are ties.

Let M be the number of distinct values in the responses. Let e_m be the multiplicity of the m th distinct response value. That is, the N observations y_{jk} consist of e_1 observations with value z_1 , e_2 observations with value z_2 , etc., with $\sum_m e_m = N$. The variance of T under the null permutation distribution is exactly

$$\begin{aligned} \text{Var}(T) = & \frac{1}{18} \left[N(N-1)(2N+5) - \sum_{j=1}^J n_{+j}(n_{+j}-1)(2n_{+j}+5) \right. \\ & \left. - \sum_{m=1}^M e_m(e_m-1)(2e_m+5) \right] \\ & + \frac{1}{2N(N-1)} \left\{ \sum_{j=1}^J n_{+j}(n_{+j}-1) \right\} \left\{ \sum_{m=1}^M e_m(e_m-1) \right\} \\ & + \frac{1}{9N(N-1)(N-2)} \left\{ \sum_{j=1}^J n_{+j}(n_{+j}-1)(n_{+j}-2) \right\} \\ & \quad \times \left\{ \sum_{m=1}^M e_m(e_m-1)(e_m-2) \right\} \end{aligned}$$

The approximate p -value is based on a normal approximation using the exact variance of T .

Linear-by-linear test for trend

When `linear` is specified, the linear-by-linear trend test is calculated based on the statistic

$$T = \frac{1}{N} \sum_{i=1}^I \sum_{j=1}^J n_{ij} (g_j - \bar{g}) (r_i - \bar{r})$$

where g_j is the group score for the j th group, $\bar{g} = (\sum_j n_{+j} g_j) / N$ is the mean group score over all observations, r_i is the response score for the i th ordered response, and $\bar{r} = (\sum_i n_{i+} r_i) / N$ is the mean response score over all observations.

Under the null hypothesis of independence between groups and responses, the expectation of T is zero. The variance of T is based on the permutation distribution of responses where the table margins n_{i+} and n_{+j} are fixed and each permutation is considered equally likely under the null hypothesis. The variance of T is exactly

$$\text{Var}(T) = \frac{1}{N^2(N-1)} \left\{ \sum_{j=1}^J n_{+j} (g_j - \bar{g})^2 \right\} \left\{ \sum_{i=1}^I n_{i+} (r_i - \bar{r})^2 \right\}$$

The approximate p -value is based on a normal approximation using the exact variance of T .

Note the symmetry between the group and response scores in these equations. In the linear-by-linear test for trend, it is arbitrary as to what is a group with a group score and what is a response with a response score. The two are interchangeable.

Cuzick's test with rank scores

When `cuzick` is specified, the test for trend is based on a method in [Cuzick \(1985\)](#), which is described in [Altman \(1991, 215–217\)](#).

Rank all N responses, using average ranks for ties. Let R_j be the sum of the ranks in the j th group. The statistic for Cuzick's test T is

$$T = \frac{1}{N} \sum_{j=1}^J (g_j - \bar{g}) R_j$$

where g_j is the group score for the j th group and $\bar{g} = (\sum_j n_{+j} g_j) / N$.

Under the null hypothesis of independence between groups and responses, the expectation of T is zero. When there are no ties, the variance of T is

$$\text{Var}(T) = \frac{N+1}{12N} \sum_{j=1}^J n_{+j} (g_j - \bar{g})^2$$

When there are ties, we again let e_m be the multiplicity of the m th distinct response value, $m = 1, 2, \dots, M$. Define

$$a = \frac{1}{N(N^2 - 1)} \sum_{m=1}^M e_m (e_m^2 - 1)$$

The corrected variance of T when there are ties is

$$\widetilde{\text{Var}}(T) = (1 - a) \text{Var}(T)$$

Exact p-values

Exact p -values are computed by `permut`. The variable permuted is that specified by `group()`, except for `carmitage` when the response variable is 0/1; in this case, the response variable is permuted. For details on the permutation computation, see [R] [permut](#).

Acknowledgments

The original version of `nptrend` was written by K. A. Stepniewska, Worldwide Antimalarial Resistance Network, University of Oxford, UK, and D. G. Altman (1948–2018).

References

- Agresti, A. 2013. *Categorical Data Analysis*. 3rd ed. Hoboken, NJ: Wiley.
- Altman, D. G. 1991. *Practical Statistics for Medical Research*. London: Chapman & Hall/CRC.
- Armitage, P. 1955. Tests for linear trends in proportions and frequencies. *Biometrics* 11: 375–386. <https://doi.org/10.2307/3001775>.
- Cochran, W. G. 1954. Some methods for strengthening the common chi-squared tests. *Biometrics* 10: 417–451. <https://doi.org/10.2307/3001616>.
- 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. <https://doi.org/10.1002/sim.4780040112>.
- Jonckheere, A. R. 1954. A distribution-free k-sample test against ordered alternatives. *Biometrika* 41: 133–145. <https://doi.org/10.2307/2333011>.
- Terpstra, T. J. 1952. The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking. *Indagationes Mathematicae* 14: 327–333.

Also see

- [R] [Epitab](#) — Tables for epidemiologists
- [R] [kwallis](#) — Kruskal–Wallis equality-of-populations rank test
- [R] [permut](#) — Monte Carlo permutation tests
- [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