

## NONPARAMETRIC ANALYSIS OF VARIANCE

---

## KRUSKAL-WALLIS TEST

---

## Kruskal-Wallis test

This is the nonparametric analogue of single-factor analysis of variance (completely randomized design).

In this lecture, we will discuss two nonparameteric tests.

As you will see, I have my personal reservations about these methods. But I think it is important that you know about these methods and understand their limitations.

We will start with the Kruskal-Wallis test, this is the nonparametric analogue of single-factor analysis of variance with completely randomized design.

## Assumptions

1. The  $j$ th observation in the  $i$ th group is modeled as

$$y_{ij} = \mu + \tau_i + e_{ij} (i = 1, \dots, t; j = 1, \dots, r)$$

where  $\mu$  is the overall “center”,  $\tau_i$  is the effect of treatment  $i$ , and  $e_{ij}$  is random error.

2. The  $e$ ’s are mutually independent.

3. **Each  $e$  comes from the same continuous population.** (Normality is not assumed.)

People often switch to non-parametric tests when normality assumption fails, without realizing that the non-parametric tests actually have quite strong assumptions too (more obscure or peculiar).

In particular, if the observations have unequal variances, then 3) is not satisfied.

Let’s look at the model equation assumed by the Kruskal-Wallis test. It has the same form as the one-way ANOVA model. The only difference is the assumptions on the error terms.

The K-W test still assumes that error terms to be i.i.d.—that is, independent and identically distributed: but it does not assume the error terms to be normally distributed. The K-W test assumes that each error term comes from the same continuous population.

People often switch to non-parametric tests when normality assumption fails, without realizing that the non-parametric tests actually have quite strong assumptions too (more obscure or peculiar).

In particular, if the observations have unequal variances, then 3) is not satisfied.

On the other hand, when we compare sample means among groups, the CLT says the means will be approximately normal even the individual observations are not, for reasonably large sample sizes.

Independence is the key, and is needed for both tests.

## Procedure

To test  $H_0: \tau_1 = \tau_2 = \dots = \tau_t$ ,

1. Rank all  $N$  observations jointly, from least to greatest. Let  $r_{ij}$  denote the rank of  $y_{ij}$ .

2. For  $i = 1, \dots, t$ , calculate

- $R_i = \sum_{j=1}^r r_{ij}$ , the sum of the ranks in group  $i$
- $R_{i\cdot} = R_i/r$ , the average of the ranks in group  $i$
- $R_{\cdot\cdot} = (N + 1)/2$ , the average of all the ranks

3. Compute

$$H = \frac{12}{N(N+1)} \cdot \sum_{i=1}^t r (R_{i\cdot} - R_{\cdot\cdot})^2$$

4.  $p$ -value =  $Pr(\chi_{t-1}^2 > H)$  (i.e., compare  $H$  to a chi-square distribution with  $t - 1$  d.f.) To test  $H_0: \tau_1 = \tau_2 = \dots = \tau_t$ ,

Now suppose that you have verified that all assumptions are met for the K-W test. Here is the procedure of the test:

First, we rank all  $N$  observations jointly, from least to greatest. Let  $(r_{ij})$  denote the rank of  $(y_{ij})$ .

Second, For each  $i$  in 1 through  $t$  calculate

$R_i$ , the sum of the ranks in group  $i$ ,

$R_{i\cdot}$ , the average of the ranks in group  $i$ , and

$R_{\cdot\cdot}$ , the average of all the ranks

Third, compute the sum of squared difference between  $(R_{i\cdot})$  and  $(R_{\cdot\cdot})$ , times 12 and divided by  $N$  times  $(N+1)$ .

Finally, compare this  $H$  statistic to a chi-square distribution with  $(t-1)$  degrees of freedom.

Note that strictly speaking that the null hypotheses being tests is that all groups have the same median.

## Example

Example. Data set bp0 (changes in blood pressure in 15 people treated with one of three doses of a drug). Ranks are in parentheses.

	Dose		
	0	20	50
change	14 (15)	1 (12)	-4 (8.5)
in b.p.	6 (14)	0 (11)	-6 (6)
	5 (13)	-4 (8.5)	-8 (4)
	-3 (10)	-5 (7)	-15 (2)
	-7 (5)	-13 (3)	-16 (1)
$R_{i.}$	11.4	8.3	4.3

$$R_{..} = (15 + 1)/2 = 8$$

$$H = \frac{12}{N(N+1)} \cdot \sum_{i=1}^t r(R_{i.} - R_{..})^2$$

$$= \frac{12}{15(16)} \cdot [5 \cdot (11.4 - 8)^2 + 5 \cdot (8.3 - 8)^2 + 5 \cdot (4.3 - 8)^2] = 6.34$$

$$P\text{-value} = \Pr(\chi^2_2 > 6.34) = 0.04$$

The  $P$ -value from one-way ANOVA was 0.029.

□

This page shows the how to compute K-W test by hand for the bp0 data.

## In R

Use

```
kruskal.test(y = change, group = dose)
```

or

```
kruskal.test(change ~ dose)
```

You can find a section on this test in [script1.html](#).

In real applications, there is an R function for this test.

You can find a section on this test in [script1.html](#).

## FRIEDMAN'S TEST

---



### Friedman's test

This is the nonparametric analogue of the randomized complete block analysis (or two-way ANOVA with one observation per treatment combination).

Friedman's test is the nonparametric analogue of the randomized complete block analysis (or two-way ANOVA with one observation per treatment combination).

## Assumptions

1. The model is

$$y_{ij} = \mu + \tau_i + \rho_j + e_{ij} (i = 1, \dots, t; j = 1, \dots, r)$$

where

- $\mu$  = overall “center”
- $\tau_i$  = effect of  $i$ th treatment
- $\rho_j$  = effect of  $j$ th block
- $e_{ij}$  = random error

2. The  $e$ 's are mutually independent.

3. Each  $e$  comes from the same continuous (not necessarily normal) population.

The model equation assumed for the Friedman's test looks the same as the model equation for the RCBD model.

Again, the difference is in the model assumption on the random errors. In the ANOVA model, it is assumed that all errors are iid normal with mean 0.

For Friedman's test, we still need to assume that all error terms are iid, but they don't have to be normal.

## Procedure

To test  $H_0: \tau_1 = \tau_2 = \dots = \tau_t$ ,

1. Within each block, rank the  $t$  observations from least to greatest. Let  $r_{ij}$  be the rank of  $y_{ij}$  in the joint ranking of  $y_{1j}, y_{2j}, \dots, y_{tj}$ .

2. For  $i = 1, \dots, t$ , calculate

- $R_i = \sum_{j=1}^r r_{ij}$ , the sum of the ranks in treatment  $i$
- $R_{i\cdot} = R_i/r$ , the average of the ranks in treatment  $i$
- $R_{\cdot\cdot} = (t + 1)/2$ , the average rank expected under  $H_0$

3. Compute

$$S = \frac{12r}{t(t+1)} \sum_{i=1}^t (R_{i\cdot} - R_{\cdot\cdot})^2.$$

4.  $p\text{-value} = \Pr(\chi_{t-1}^2 > S)$ . (Compare  $S$  to a chi-square distribution with  $t - 1$  d.f.)

This slide summarizes the procedure for performing a Friedman's test.

The form of the test statistic looks similar to the K-W test. The key difference is that the ranking of the observations is done within each block separately.

## Example

Example. Nitrogen content of wheat as a function of six fertilization methods, applied in four blocks (data set WHEAT).

Block	Fert. meth.	Nitrogen content	Rank	Block	Fert. meth.	Nitrogen content	Rank
1	1	35	2	3	1	36.9	1
1	2	40.9	5	3	2	46.6	5
1	3	42.1	6	3	3	52.7	6
1	4	37.2	3	3	4	40.2	3
1	5	38	4	3	5	37.6	2
1	6	34.9	1	3	6	44.6	4
2	1	41.2	1	4	1	40	2
2	2	46.7	4	4	2	41.9	4
2	3	49.4	5	4	3	42.9	5
2	4	45.9	3	4	4	39.2	1
2	5	42	2	4	5	40.5	3
2	6	50.1	6	4	6	43.3	6

$$R_{1.} = \frac{2+1+1+2}{4} = 1.5 \quad \dots \quad R_{6.} = \frac{1+6+4+6}{4} = 4.25$$

$$R_{..} = (6+1)/2 = 3.5$$

$$\begin{aligned}
 S &= \frac{12r}{t(t+1)} \sum_{i=1}^t (R_{i.} - R_{..})^2 \\
 &= \frac{12 \cdot 4}{6 \cdot 7} [(1.5 - 3.5)^2 + (4.5 - 3.5)^2 + (5.5 - 3.5)^2 + (2.5 - 3.5)^2 + \\
 &\quad (2.75 - 3.5)^2 + (4.25 - 3.5)^2] \\
 &= 12.71
 \end{aligned}$$

The  $P$ -value is  $\Pr(\chi^2_3 > 12.71) = 0.026$ . (The  $P$ -value from the RCB analysis was 0.004.)  $\square$

Here we show how to compute the Friedman's test by hand using the nitrogen content of wheat data.

## In R

Use

```
friedman.test(y = nitro, groups = fert, blocks = block)
```

See the R code in  
[example8.1.html](#)

In practice, we can use the R function “`friedman.test`” to perform this non-parametric test.

## Summary

Kruskal-Wallis: nonparametric analogue of one-way ANOVA for completely randomized design

Friedman's test: nonparametric analogue of two-way ANOVA for RCBD

They are both based on ranks.

Use with caution: nonparametric tests are not assumption free!

Kruskal-Wallis is a nonparametric analogue of one-way ANOVA for completely randomized design

Friedman's test is nonparametric analogue of two-way ANOVA for RCBD

They are both based on ranks.

Use them with caution: nonparametric tests are not assumption free!