

## Friedman's Two-way Analysis of Variance by Ranks -- Analysis of k-Within-Group Data with a Quantitative Response Variable

**Application:** This statistic has two applications that can appear very different, but are really just two variations of the same statistical question. In one application the same quantitative variable is measured at two or more different times from the same sample (or from two or more samples that have been matched on one or more important variables). In the other application, two or more comparable quantitative variables are measured from the same sample (usually at the same time). In both applications, Friedman's test is used to compare the distributions of the two or more quantitative variables.

Thus, it is applied in the same data situation as an ANOVA for dependent samples except that it is used when the data are either from a too-small sample, are importantly non-normally distributed, or the measurement scale of the dependent variable is ordinal (not interval or ratio). It is important to remember the null hypothesis, and to differentiate it from the null for the dependent ANOVA.

There are two specific versions of the H<sub>0</sub>, depending upon whether one characterizes the k conditions as representing a single population under two or more different circumstances (e.g., comparing treated vs. not treated or comparing different treatments -- some consider this a representation of two or more different populations) or as representing comparable variables measured from a single population (as in the example below). Here are versions of the H<sub>0</sub>: statement for each of these characterizations.

**H<sub>0</sub>:** The populations represented by the k conditions of the have the same distribution of scores.

**To reject H<sub>0</sub>:** is to say that the populations represented by the different conditions differ in some way, center, spread and/or shape. When the forms of the distributions from the two samples are similar (as is often the case -- check the size and symmetry of the IQR), then rejecting H<sub>0</sub>: is interpreted to mean that populations (or circumstances) tend to have some pattern of larger and smaller scores (different medians) among them.

**H<sub>0</sub>:** The population has the same distribution of scores on the different measures represented by the conditions.

**To reject H<sub>0</sub>:** is to say that the distributions of the variables are different in some way, center, spread and/or shape. When the forms of the distributions are similar (as is often the case -- check the size and symmetry of the IQR), then rejecting the H<sub>0</sub>: is interpreted to mean that the variables have some pattern of larger and smaller scores (medians) among them.

**The data:** In this analysis the one variable is the type of animal (fish, reptiles, or mammals), and the response variable is the number of animals on display. From our database, we use three variables **reptnum** (number of reptiles on display), **fishnum** (number of fish on display) and **mamlnum** (number of mammals on display). These scores are shown for the 12 stores below (reptnum, fishnum, ).

12, 32, 34	14, 41, 38	15, 31, 45	12, 38, 32	7, 21, 12	4, 13, 11
10, 17, 22	4, 22, 9	14, 24, 20	4, 11, 8	5, 17, 19	10, 20, 8

**Research Hypothesis:** The data come from the Pet shop database. The researcher hypothesized that stores would tend to display more fish than other types of animals, fewer reptiles, and an intermediate number of mammals.

**H<sub>0</sub>: for this analysis:** Pet stores display the same number of reptiles, fish and mammals.

**Step 1** Rearrange the data so that scores from each subject are in the appropriate columns, one for each condition.

reptnum	fishnum	mamlnum
12	32	34
14	41	38
15	31	45
12	38	32
7	21	12
4	13	11
10	17	22
4	22	9
14	24	20
4	11	8
5	17	19
10	20	8

**Step 2** Rank order the scores SEPARATELY FOR EACH SUBJECT'S DATA with the smallest score getting a value of 1. If there are ties (within the scores for a subject) each receives the average rank they would have received.

reptnum	fishnum	mamlnum	rank reptnum	rank fishnum	rank mamlnum
12	32	34	1	2	3
14	41	38	1	3	2
15	31	45	1	2	3
12	38	32	1	3	2
7	21	12	1	3	2
4	13	11	1	3	2
10	17	22	1	2	3
4	22	9	1	3	2
14	24	20	1	3	2
4	11	8	1	3	2
5	17	19	1	2	3
10	20	8	2	3	1

**Step 3** Compute the sum of the ranks for each condition.

rank reptnum	rank fishnum	rank mamlnum
1	2	3
1	3	2
1	2	3
1	3	2
1	3	2
1	3	2
1	2	3
1	3	2
1	3	2
1	3	2
1	2	3
2	3	1

$R_1 = 13$      $R_2 = 32$      $R_3 = 27$

**Step 4** Determine the number of subjects.

$$N = 12$$

**Step 5** Determine the number of conditions.

$$k = 3$$

**Step 6** Compute Friedman's F, using the following formula (you should carry at least 3 decimals in these calculations).

$$\begin{aligned}
 F &= \left[ \frac{12}{[N * k * (k+1)]} \right] * \sum R^2 - [3 * N * (k + 1)] \\
 &= \left[ \frac{12}{[12 * 3 * (3+1)]} \right] * [13^2 + 32^2 + 27^2] - [3 * 12 * (3 + 1)] \\
 &= \left[ \frac{12}{144} \right] * [169 + 1024 + 729] - 144 = [.083*1922] - 144 = 15.526
 \end{aligned}$$

**For small samples (k < 6 AND N < 14):**

**Step 7** Determine the critical value of F by looking at the table of critical values for Friedman's test

$$F(k=3, N=12, \alpha = .05) = 8.67$$

**Step 8** Compare the obtained F and the critical F values to determine whether to retain or reject the null hypothesis.

-- if the obtained F value (from Step 6) is larger than the critical value of F, then reject H0:

-- if the obtained F value is less than or equal to the critical value of F, then reject H0:

For the example data, we would decide to reject the null hypothesis, because the obtained value of F (15.526) is greater than the larger critical F value (8.67).

**For large samples (k > 5 OR N > 13):**

**Step 9** Determine the critical value of F by looking at the table of critical values for the Chi-Square test (df = k - 1).  
(As an example, here is how you would apply this version of the significance test to these data.)

$$X^2 (df = 2, p = .05) = 5.99$$

**Step 10** Compare the obtained F and the critical X<sup>2</sup> values to determine whether to retain or reject the null hypothesis.

-- if the obtained F value (from Step 6) is less than or equal the critical value of X<sup>2</sup>, then retain H0:

-- if the obtained F value is larger than the critical value of X<sup>2</sup>, then reject H0:

For the example data, we would decide to reject the null hypothesis, because the obtained value of F (15.526) is greater than the larger critical X<sup>2</sup> value (5.99).

**Step 11** IF you reject the null hypothesis, determine whether the pattern of the data completely supports, partially supports, or does not support the research hypothesis.

- IF you reject the null hypothesis, AND if the pattern of data agrees exactly with the research hypothesis, then the research hypothesis is completely supported.
- IF you reject the null hypothesis, AND if part of the pattern of the data agrees with the research hypothesis, BUT part of the pattern of the data does not, then the research hypothesis is partially supported.
- IF you retain the null hypothesis, OR you reject the null BUT NO PART of the pattern of the data agrees with the research hypothesis, then the research hypothesis is not at all supported.

**By the way:** To properly determine whether the hypothesized pattern of differences was found, one should perform pairwise comparisons (using Friedman's test); the report of the results given below makes use of these follow-up tests (although the computations are not shown).

**By the way:** Usually the researcher hypothesizes that there is a difference between the conditions. Sometimes, however, the research hypothesis is that there is NO difference between the conditions. If so, the research hypothesis and  $H_0$  are the same! When this is the case, retaining  $H_0$  provides support for the research hypothesis, whereas rejecting  $H_0$  provides evidence that research hypothesis is incorrect.

For the example data, we would decide that the research hypothesis is partially supported, because the null hypothesis was rejected, and because, as hypothesized, there were fewer reptiles displayed than mammals or fish. However, there was not a significant difference between the number of fish and mammals displayed at these stores.

## Step 12 Reporting the results

You will want to compute medians and IQR values to help describe the data before reporting the results of the significance test. With multiple-group designs it is often easier to present these data in a table. As for the other statistical tests, the report includes the "wordy" part and the statistical values based upon which you based your statistical decision. If you reject  $H_0$ , be sure to describe how the groups differed, rather than just reporting that there was "a difference".

Table 1 summarizes the data for the numbers of animals displayed at the stores. There was a significant difference among the distributions of the three types of animals (based on Friedman's test,  $\chi^2(2) = 15.526$ ,  $p = .0003$ . Pairwise Friedman's tests ( $p < .05$ ) revealed that, as hypothesized, fewer reptiles were displayed than either fish or mammals. However, contrary to the research hypothesis, there was not a difference in the mean numbers of fish and mammals displayed.

Table 1  
Summary of the number of animals of each type displayed in the pet stores.

---

	Type of Animal		
	Fish	Mammals	Reptiles
Mdn	21.50	19.50	10.00
Q1	17.00	9.50	4.25
Q3	31.75	33.50	13.50

---

## $F_r$ Critical values for Friedman's two-way analysis of Variance by ranks

$k$	$N$	$\alpha < .10$	$\alpha \leq .05$	$\alpha < .01$
3	3	6.00	6.00	---
	4	6.00	6.50	8.00
	5	5.20	6.40	8.40
	6	5.33	7.00	9.00
	7	5.43	7.14	8.86
	8	5.25	6.25	9.00
	9	5.56	6.22	8.67
	10	5.00	6.20	9.60
	11	4.91	6.54	8.91
	12	5.17	6.17	8.67
	13	4.77	6.00	9.39
	¥	4.61	5.99	9.21
	4	2	6.00	6.00
3		6.60	7.40	8.60
4		6.30	7.80	9.60
5		6.36	7.80	9.96
6		6.40	7.60	10.00
7		6.26	7.80	10.37
8		6.30	7.50	10.35
¥		6.25	7.82	11.34
5	3	7.47	8.53	10.13
	4	7.60	8.80	11.00
	5	7.68	8.96	11.52
	¥	7.78	9.49	13.28