Friedman Test

• A nonparametric test used to compare several treatments (populations) when all treatments are applied (in random order) to each subject in the sample. (If there is no blocking, use the Kruskal-Wallis test if a nonparametric procedure is called for.)

Set Up. Each of k treatments is independently applied to b blocks (or subjects) and each outcome X_{ij} (where $i = 1, ..., b$ and $j = 1, ..., k$) is recorded. (We assume that the results within one block do not influence the results within the other blocks.)

	Treatments			
Blocks	1	2	...	k
1	X_{11}	X_{12}	...	X_{1k}
2	X_{21}	X_{22}	...	X_{2k}
$\ddots$	$\ddots$		$\ddots$	
b	X_{b1}	X_{b2}	...	X_{bk}

Hypotheses.

H_0: treatments are similarly effective ($p_{i1} = p_{i2} = ... = p_{ic}$, see below)
H_1: treatments differ in effectiveness

Obtaining Test Statistic

First, for each block replace X_{ij} with its rank $R(X_{ij})$ within that block. Use average ranks in case of ties. Let R_j denote the sum of ranks for treatment j.

	Treatments			
Blocks	1	2	...	k
1	$R(X_{11})$	$R(X_{12})$	...	$R(X_{1k})$
2	$R(X_{21})$	$R(X_{22})$	...	$R(X_{2k})$
$\ddots$	$\ddots$		$\ddots$	
b	$R(X_{b1})$	$R(X_{b2})$	...	$R(X_{bk})$
Treatment rank sums:	R_1	R_2	...	R_k

Let $A = \sum_{i=1}^{b}\sum_{j=1}^{k}[R(X_{ij})]^2$ and $C = \frac{bk(k+1)^2}{4}$. Then Friedman's **test statistic** (adjusted for possible ties) is given by

$$T = \frac{(k-1)\sum_{j=1}^{k}\left(R_j - \frac{b(k+1)}{2}\right)^2}{A-C},$$

which has an approximate ${\chi^2}_{(k-1)}$ distribution under the null hypothesis.

If there are no ties, T simplifies to $T = \frac{12}{bk(k+1)}\sum_{j=1}^{k}\left(R_j - \frac{b(k+1)}{2}\right)^2$.

Another route that gives a better approximation of the true distribution uses

$$F^* = \frac{(b-1)T}{b(k-1)-T},$$

which has an approximate $F_{[k-1,\,(b-1)(k-1)]}$ distribution under the null hypothesis.

The test statistic F^* can be obtained by two-way ANOVA on the ranks $R(X_{ij})$.

Decision Rule. Reject H_0 if the test statitistic is too large or if the p-value is too small.

Multiple Comparisons. If the null hypothesis is rejected, it is appropriate to ascertain which pairs of treatments differ. Treatments i and j are considered different if

$$|R_i - R_j| > t_{1-\alpha/2}\sqrt{\frac{2(bA - \sum R_j^2)}{(b-1)(k-1)}}\ .$$

Example.(from Biomed Central Ltd). Pain scores of five patients receiving four separate treatments.

	Treatment				
Patient (block)	A	B	C	D	
1	6	9	10	16	
2	9	16	16	32	
3	14	14	22	67	
4	10	14	40	19	
5	11	16	17	60	

Table of ranks

	Treatment				
Patient (block)	A	B	C	D	
1	1	2	3	4	
2	1	2.5	2.5	4	
3	1.5	1.5	3	4	
4	1	2	4	3	
5	1	2	3	4	
Sum(Rj)	5.5	10	15.5	19	

H_o: All four treatments are equally effective in treating pain.
H_1: Not all treatments are equally effective in treating pain.

$$A = \sum_{i=1}^{b}\sum_{j=1}^{k}[R(X_{ij})]^2 = 1^2 + 2^2 + ... + 4^2 = 149$$

and $C = \frac{bk(k+1)^2}{4} = \frac{5(4)(5^2)}{4} = 125$, thus

$$T = \frac{(k-1)\sum_{j=1}^{k}\left(R_j - \frac{b(k+1)}{2}\right)^2}{A-C} = \frac{3(7^2+2.5^2+3^3+6.5^2)}{149-125} = 13.3125.$$

with p-value $= .004$ [from TI-83: χ^2cdf(13.3125,10^99,3)].

Or use $F^* = \frac{(b-1)T}{b(k-1)-T} = \frac{4(13.3125)}{5(3)-13.3125} = 31.56$ with p-value $= .0000056$.

Since the null hypothesis is rejected, for multiple comparisons, with $\alpha = .05$, we have

$$t_{1-\alpha/2}\sqrt{\frac{2(bA-\sum R_j^2)}{(b-1)(k-1)}} = 2.1788\sqrt{\frac{2(5*149-5.5^2-10^2-15.5^2-19^2)}{3(4)}} = 3.2682.$$

Hence if $|R_i - R_j| > 3.2682$, we can say that treatments i and j differ.
Here, we see that all the treatments have a statistically significant difference.

Reference. *Practical Nonparametric Statistics*, 3rd ed, W.J. Conover, John Wiley & Sons, 1999