

9 Blocked Designs

9.1 Friedman's Test

9.1.1 Application 1: Randomized Complete Block Designs

- Assume there are k treatments of interest in an experiment. In Section 8, we considered the k -sample Extension of the Median Test and the Kruskal-Wallis Test to test for any differences in the k treatment medians.
- Suppose the experimenter is still concerned with studying the effects of a single factor on a response of interest, but variability from another factor that is not of interest is expected.
 - Suppose a researcher wants to study the effect of 4 fertilizers on the yield of cotton. The researcher also knows that the soil conditions at the 8 areas for performing an experiment are highly variable. Thus, the researcher wants to design an experiment to detect any differences among the 4 fertilizers on the cotton yield in the presence a “nuisance variable” not of interest (the 8 areas).
- Because experimental units can vary greatly with respect to physical characteristics that can also influence the response, the responses from experimental units that receive the same treatment can also vary greatly.
- If it is not controlled or accounted for in the data analysis, it can greatly inflate the experimental variability making it difficult to detect real differences among the k treatments of interest (large Type II error).
- If this source of variability can be separated from the treatment effects and the random experimental error, then the sensitivity of the experiment to detect real differences between treatments is increased (i.e., lower the Type II error).
- Therefore, the goal is to choose an experimental design in which it is possible to control the effects of a variable not of interest by bringing experimental units that are similar into a group called a “block”. The experimental units are assumed to be **homogeneous** within each block.
- Assume there are k treatments and b blocks. If we have one observation per treatment level within each block, and the order in which the treatments are run within each block is determined randomly, then we have a **randomized complete block design (RCBD)**. Because randomization occurs within each block, this is an example of **restricted randomization**.
- Question: So how would an RCBD be set up for the cotton yield experiment?

- The data from a RCBD can be summarized as:

	Treatment (i)			
Block (j)	1	2	\cdots	a
1	y_{11}	y_{21}	\cdots	y_{k1}
2	y_{12}	y_{22}	\cdots	y_{k2}
\vdots	\vdots	\vdots	\vdots	\vdots
b	y_{1b}	y_{2b}	\cdots	y_{kb}

9.1.2 Application 2: Simple Repeated Measures Designs

- Once again, assume there are k treatments of interest in an experiment.
- In experimental work in biomedical, pharmaceutical, social, behavioral sciences, and occasionally in physical sciences and engineering, the experimental units are people (or animals).
- Because subjects can vary greatly with respect to physical characteristics, health history, life experiences, training, etc., the responses of subjects who receive the same treatment can also vary greatly.
- If it is not controlled or accounted for in the data analysis, it can greatly inflate the experimental variability making it difficult to detect real differences among the k treatments of interest (large Type II error).
- If this source of variability can be accounted for and separated from the treatment effects, then the sensitivity of the experiment to detect real differences between treatments is increased (i.e., lower the Type II error).
- It may be possible to control for the subject-to-subject variability by running a design in which each subject receives all k treatments. Such a design is called a **repeated measures design (RMD)**.
- We will look at the simplest RMD in which each subject receives all k treatments in a random order.
- Assume there are k treatments and b subjects. If we have one observation per treatment level for each subject, and the order in which the treatments are run within each subject is determined randomly, then we have a **simple repeated measures design (SRMD)**.
- In a SRMD, treatment effects for subject j are measured relative to the average response of subject j across k treatments.
- In this sense, a subjects serves as its own control: responses of a subject to the treatments are measured in terms of deviations about a point which measures the average responsiveness of that subject.
- Thus, variability due to differences in the average responsiveness of the subjects to treatments is removed from the experimental error (assuming the SRMD model is appropriate).

- The SRMD data can be summarized as:

Subject (j)	Treatment (i)			
	1	2	\dots	k
1	y_{11}	y_{21}	\dots	y_{k1}
2	y_{12}	y_{22}	\dots	y_{k2}
\vdots	\vdots	\vdots	\vdots	\vdots
b	y_{1b}	y_{2b}	\dots	y_{kb}

- One potential problem is the **carry-over effect** in which the effect of receiving one treatment may influence the effect of (or carry-over to) the next treatment received by the subject.
- The goal is to prevent (or, at least minimize) any carry-over effects when designing the study. For example, in drug studies, this involves waiting a sufficient amount of time until the drug is out of the subject's system before administering the next drug.
- In our analysis, we will be assuming that there are no carry-over effects associated with a treatment. In essence, we are assuming that the subjects unique characteristics remain constant (uniform) at those times when the treatments are administered.

Simple Repeated Measures Design (SRMD) Example (from *Statistical Principles in Experimental Design* by B. Winer)

- The purpose of this experiment was to study the effects of 4 drugs upon reaction time to a series of standardized tasks. The 5 subjects will be randomly selected and given extensive training on these tasks prior to the experiment.
- Each of the subjects will be observed while under the effect of each of the 4 drugs and the order of administering the drug will be randomized.
- A sufficient amount of time will be allowed between administration of the drugs to avoid a carryover effect of one drug upon the effects of subsequent drugs (which is known as a drug interaction).
- Question: So how would an SRMD be set up for this experiment?

9.1.3 Performing Friedman's Test

- The parametric approach to analyzing data from either a RCBD or SRMD is a twoway ANOVA assuming no interaction between blocks and treatments.
- Friedman's Test can be applied to data from either a RCBD or a SRMD when the parametric approach is inappropriate. For example, if the normality assumption is violated or we only have ordinal data (such as ranks).

Assumptions: Given a RCBD with k treatments and b blocks or a SRMD with k treatments and b subjects.

1. The data consist of mutually independent blocks or subjects of size k .
2. The observations are independent both within and among samples.
3. The variable of interest is continuous.
4. The measurement scale is at least ordinal within each block. That is, the observations may be ranked within each block.
5. There is no interaction between blocks and treatments.

Hypotheses:

$$H_0 : M_1 = M_2 = \dots = M_k.$$

$$H_1 : M_i \neq M_{i'} \quad \text{for at least one } i \neq i'$$

where M_1, M_2, \dots, M_k represent the median responses associated with the k treatments.

- Thus, if H_0 is true, the treatment effects are identical, while if H_0 is false, then there is at least one treatment effect that is different from at least one other treatment effect.

Rationale:

- Note that the assignment of ranks for Friedman's Test does not require combining data across groups before assigning ranks (like we did in the Kruskal-Wallis Test).
- Observations are only ranked relative to the other observations from the same block or subject. Thus, within each block or subject, there will be ranks $1, 2, \dots, k$ (unless there are ties, in which case average ranks are assigned). The following is a summary of the table of ranks such that ranks $1, 2, \dots, k$ appear in each row (block or subject).

Block or Subject (j)	Treatment (i)			
	1	2	\dots	a
1	R_{11}	R_{21}	\dots	R_{k1}
2	R_{12}	R_{22}	\dots	R_{k2}
\vdots	\vdots	\vdots	\vdots	\vdots
b	R_{1b}	R_{2b}	\dots	R_{kb}
Column Sums	R_1	R_2	\dots	R_k

- If H_0 is true and all treatments have identical effects, then the ranks that appear in any treatment column of the rank summary table occur merely by chance.
- Thus, when H_0 is true, neither the smallest ranks nor the largest ranks should tend to appear in a particular treatment column.
- In other words, the ranks in each block should be randomly distributed over the columns (i.e., the treatments) within each block or subject.
- However, if H_0 is false, then we expect a lack of randomness in the distribution of ranks in each block.
- For example, if one treatment is better than the others, then we expect to see the large ranks in that treatment column if ‘better’ implies a larger response (or expect to see the small ranks in that treatment column if ‘better’ implies a smaller response).
- A useful test would therefore be one that is sensitive to detecting departures from the pattern expected when H_0 is true. Friedman’s Test is such a test because it detects departures in the pattern of the ranks assuming H_0 is true. “Significant” departures would then lead to rejection of H_0 .
- The Friedman’s Test statistic is based on the magnitudes of the sums of the ranks associated with each treatment (i.e., sums of column ranks).
 - If H_0 is true, then we expect the column sums R_1, R_2, \dots, R_k to be relatively close to each other such that any differences can be attributed to chance.
 - If H_0 is false, then we expect at least one column sum R_i to be sufficiently different in size with at least one other column sum R_j such that the difference is unlikely to have occurred by chance alone. In this case, we would reject H_0 .

Method: For a given α

- Assign ranks $1, 2, \dots, k$ to the observations within each block or subject. Let R_{ij} be the rank of the i^{th} treatment observation within block or subject j .
- Let $R_i =$ the sum of the ranks for the i^{th} treatment (the i^{th} column sum). The **Friedman Test statistic is:**

$$\chi_F^2 = \frac{12}{bk(k+1)} \sum_{i=1}^k \left[R_i - \frac{b(k+1)}{2} \right]^2$$

where $\frac{b(k+1)}{2} = E_0(R_i)$ is the mean of the R_i ’s when H_0 is true.

- Asymptotically, $\chi_F^2 \sim \chi_{(k-1)}^2$ (χ^2 with $k - 1$ degrees of freedom).
- Note that large differences between the R_i ’s and $E_0(R_i)$ inflate the χ_F^2 statistic, and a sufficiently large χ_F^2 will lead to rejecting H_0 .
- An equivalent computational form for the test statistic is

$$\chi_F^2 = \frac{12}{bk(k+1)} \sum_{i=1}^k R_i^2 - 3b(k+1)$$

- Some texts will use the following scaled version of χ_F^2 :

$$W = \frac{12 \sum_{i=1}^k R_i^2 - 3b^2k(k+1)^2}{b^2k(k^2-1)}$$

That is, $W = \chi_F/[b(k-1)]$. A table of critical values for W is given in the notes.

Decision Rule

- Using W : If the value of P in the table is $\leq \alpha$ for the value of test statistic W , then reject H_0 .
- Using χ_F^2 : Reject H_0 if χ_F^2 is $\geq \chi_\alpha^2$ where χ_α^2 is the value from a χ^2 distribution with $k-1$ degrees of freedom with right-tail probability α .
- Otherwise, **Fail to Reject H_0** .

Adjustment for Ties:

- When there are ties within a block or subject, we assign average ranks.
- The formula for W or χ_F^2 can be adjusted for ties to improve the estimated p -value. Replace the denominator of W with

$$b^2k(k^2-1) - b \sum_{j=1}^b (t^3 - t)$$

where t is the number of observations with tied ranks within each block or subject.

Example of Friedman's Test for a RCBD (from *Applied Nonparametric Statistics* by W. Daniel).

Researchers compared three methods of determining serum amylase values in patients with pancreatitis. The results are given in the following table. Do these data indicate a difference among the three methods? Thus, test

$$H_0 : M_1 = M_2 = M_3 \quad \text{against} \quad H_1 : M_i \neq M_{i'} \quad \text{for at least one } i \neq i'$$

Specimen	Method A	Method B	Method C
1	4000	3210	6120
2	1600	1040	2410
3	1600	647	2210
4	1200	540	2060
5	840	445	1400
6	352	156	249
7	224	155	224
8	200	99	208
9	184	70	227

R code for Friedman's Test for RCBD Example

```
# Friedman's Test of Serum Specimen Data

y <- c(4000,3210,6120,1600,1040,2410,1600,647,2210,
1200,570,2060,840,445,1400,352,156,249,
224,155,224,200,99,208,184,70,227)

specimen <- c(1,1,1,2,2,2,3,3,3,4,4,4,5,5,5,6,6,6,7,7,7,8,8,8,9,9,9)
method <- c(rep(c("A","B","C"),9))

friedman.test(y,method,specimen)
```

R output for Friedman's Test for RCBD Example

```
> # Friedman's Test of Serum Specimen Data

Friedman rank sum test

data: y, method and specimen

Friedman chi-squared = 15.9429, df = 2, p-value = 0.0003452
```


SAS code for Friedman's Test

```
DM 'LOG;CLEAR;OUT;CLEAR;';
OPTION PS=60 LS=64 NODATE NONUMBER;

*****;
*** Friedman's Test Example ***;
*****;

*** Analysis by entering ranks within blocks ***;

DATA IN;
  DO SPECIMEN = 1 TO 9;
  DO METHOD = 'A' , 'B' , 'C';
    INPUT SERUM @@; OUTPUT;
  END; END;
CARDS;
  2 1 3   2 1 3   2 1 3
  2 1 3   2 1 3   3 1 2
  2.5 1 2.5   2 1 3   2 1 3
;

PROC FREQ DATA=IN;
  TABLE SPECIMEN*METHOD*SERUM / NOPRINT CMH;
  TITLE 'FRIEDMAN TEST (INPUTS ARE RANKS WITHIN BLOCKS)';
RUN;

*** Analysis by entering original responses ***;

DATA IN;
  DO specimen = 1 TO 9;
  DO method = 'A' , 'B' , 'C';
    INPUT serum @@; OUTPUT;
  END; END;
CARDS;
4000 3210 6120   1600 1040 2410   1600 647 2210
1200  570 2060   840  445 1400   352 156  249
 224  155  224   200   99  208   184  70  227
;

*** Proc Rank converts the y-values to ranks ***;

PROC SORT DATA=IN; BY ;
PROC RANK DATA=IN OUT=rankset; BY specimen;
  VAR serum;
  RANKS srmrank;
PROC PRINT DATA=rankset;
TITLE 'FRIEDMAN TEST (INPUTS ARE RAW DATA)';

PROC FREQ DATA=rankset;
  TABLE specimen*method*srmrank / NOPRINT CMH;
RUN;
```

SAS output for Friedman's Test

FRIEDMAN TEST (INPUTS ARE RANKS WITHIN BLOCKS)

The FREQ Procedure

Summary Statistics for METHOD by SERUM
Controlling for SPECIMEN

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	2.0571	0.1515
2	Row Mean Scores Differ	2	15.9429	0.0003 <-- pvalue
3	General Association	6	27.0000	0.0001

Total Sample Size = 27

FRIEDMAN TEST (INPUTS ARE RAW DATA)

Obs	specimen	method	serum	srmrank
1	1	A	4000	2.0
2	1	B	3210	1.0
3	1	C	6120	3.0
4	2	A	1600	2.0
5	2	B	1040	1.0
6	2	C	2410	3.0
7	3	A	1600	2.0
8	3	B	647	1.0
9	3	C	2210	3.0
10	4	A	1200	2.0
11	4	B	570	1.0
12	4	C	2060	3.0
13	5	A	840	2.0
14	5	B	445	1.0
15	5	C	1400	3.0
16	6	A	352	3.0
17	6	B	156	1.0
18	6	C	249	2.0
19	7	A	224	2.5
20	7	B	155	1.0
21	7	C	224	2.5
22	8	A	200	2.0
23	8	B	99	1.0
24	8	C	208	3.0
25	9	A	184	2.0
26	9	B	70	1.0
27	9	C	227	3.0

The FREQ Procedure

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	2.0571	0.1515
2	Row Mean Scores Differ	2	15.9429	0.0003 <-- p-value
3	General Association	6	27.0000	0.0001

Total Sample Size = 27

Simple Repeated Measures Design (SRMD) Example (from *Statistical Principles in Experimental Design* by B. Winer)

- The purpose of this experiment was to study the effects of 4 drugs upon reaction time to a series of standardized tasks.
- The 5 subjects were randomly selected and had been given extensive training on these tasks prior to the experiment.
- Each of the subjects was observed under each of the 4 drugs and the order of administering the drug was randomized.
- A sufficient amount of time was allowed between administration of the drugs to avoid a carryover effect of one drug upon the effects of subsequent drugs (which is known as a drug interaction). The following table summarizes the reaction times.

Subject	Drug Treatment			
	1	2	3	4
1	30	28	16	34
2	14	18	10	22
3	24	20	18	30
4	38	34	20	44
5	26	28	14	30

R output for Friedman's Test for SRMD Example

```
> # Friedman's Test of Simple Repeated Measures Data
```

```
   y subject treatment
1  30      1         A
2  28      1         B
3  16      1         C
4  34      1         D
5  14      2         A
6  18      2         B
7  10      2         C
8  22      2         D
9  24      3         A
10 20      3         B
11 18      3         C
12 30      3         D
13 38      4         A
14 34      4         B
15 20      4         C
16 44      4         D
17 26      5         A
18 28      5         B
19 14      5         C
20 30      5         D
```

```
Friedman rank sum test
```

```
data: y, treatment and subject
Friedman chi-squared = 13.56, df = 3, p-value = 0.00357 <-- Reject H0
```

R code for Friedman's Test for SRMD Example

```
# Friedman's Test of Simple Repeated Measures Data
y <- c(30,28,16,34,14,18,10,22,24,20,18,30,38,34,20,44,26,28,14,30)
subject <- c(1,1,1,1,2,2,2,2,3,3,3,3,4,4,4,4,5,5,5,5)
treatment <- c(rep(c("A","B","C","D"),5))

# Verify data was entered correctly
viewdata <- data.frame(y,subject,treatment)
viewdata

friedman.test(y,treatment,subject)
```

9.1.4 Multiple Comparison Procedure for use with Friedman's Test

Investigators are usually not satisfied to know simply that their data allow them to conclude that not all sampled populations or all treatment effects are identical. For example, when the application of the Friedman test leads us to reject H_0 , we are usually interested in exactly where the differences are located. What we need, then, is a multiple-comparison procedure to use after the Friedman test.

When we compare all possible differences between pairs of samples, when the experimentwise error rate is α , and when the number of blocks is large, then we declare R_j and $R_{j'}$ significantly different if

$$|R_j - R_{j'}| \geq z \sqrt{\frac{bk(k+1)}{6}} \quad (7.4)$$

where R_j and $R_{j'}$ are the j th and j' th treatment rank totals, and z is a value from Table A.2 corresponding to $\alpha/k(k-1)$. For an alternative multiple-comparisons formula, see Hollander and Wolfe (T38).

9.1.5 Example of the exact distribution of the W statistic for Friedman's Test

- Consider the case with $k = 3$ treatments and $b = 3$ blocks.
- For each block there are $3 \times 2 \times 1 = 6$ randomizations of the 3 ranks.
- Thus, there are $6 \times 6 \times 6 = 216$ randomizations of the 3 ranks across the 3 blocks.
- The following table contains a subset of these 216 randomizations.
- A relative frequency table and reverse relative frequency table are then constructed for the set of possible W values. This provides the exact right and left tail probabilities for the W -statistic.
- Note that these match the values in Table A.14.

FRIEDMAN TEST -- DISTRIBUTION OF W FOR K=3 AND B=3

Obs	A	A	A	B	B	B	C	C	C	R	R	R	W
	1	2	3	1	2	3	1	2	3	1	2	3	
1	1	2	3	1	2	3	1	2	3	3	6	9	1.000
2	1	2	3	1	2	3	1	3	2	3	7	8	0.778
3	1	2	3	1	2	3	2	1	3	4	5	9	0.778
4	1	2	3	1	2	3	2	3	1	4	7	7	0.333
5	1	2	3	1	2	3	3	1	2	5	5	8	0.333
6	1	2	3	1	2	3	3	2	1	5	6	7	0.111
7	1	2	3	1	3	2	1	2	3	3	7	8	0.778
8	1	2	3	1	3	2	1	3	2	3	8	7	0.778
9	1	2	3	1	3	2	2	1	3	4	6	8	0.444
10	1	2	3	1	3	2	2	3	1	4	8	6	0.444
11	1	2	3	1	3	2	3	1	2	5	6	7	0.111
12	1	2	3	1	3	2	3	2	1	5	7	6	0.111
13	1	2	3	2	1	3	1	2	3	4	5	9	0.778
14	1	2	3	2	1	3	1	3	2	4	6	8	0.444
15	1	2	3	2	1	3	2	1	3	5	4	9	0.778
16	1	2	3	2	1	3	2	3	1	5	6	7	0.111
17	1	2	3	2	1	3	3	1	2	6	4	8	0.444
18	1	2	3	2	1	3	3	2	1	6	5	7	0.111
19	1	2	3	2	3	1	1	2	3	4	7	7	0.333
20	1	2	3	2	3	1	1	3	2	4	8	6	0.444
21	1	2	3	2	3	1	2	1	3	5	6	7	0.111
22	1	2	3	2	3	1	2	3	1	5	8	5	0.333
23	1	2	3	2	3	1	3	1	2	6	6	6	0.000
24	1	2	3	2	3	1	3	2	1	6	7	5	0.111
25	1	2	3	3	1	2	1	2	3	5	5	8	0.333
26	1	2	3	3	1	2	1	3	2	5	6	7	0.111
27	1	2	3	3	1	2	2	1	3	6	4	8	0.444
28	1	2	3	3	1	2	2	3	1	6	6	6	0.000
29	1	2	3	3	1	2	3	1	2	7	4	7	0.333
30	1	2	3	3	1	2	3	2	1	7	5	6	0.111
31	1	2	3	3	2	1	1	2	3	5	6	7	0.111
32	1	2	3	3	2	1	1	3	2	5	7	6	0.111
33	1	2	3	3	2	1	2	1	3	6	5	7	0.111
34	1	2	3	3	2	1	2	3	1	6	7	5	0.111
35	1	2	3	3	2	1	3	1	2	7	5	6	0.111
36	1	2	3	3	2	1	3	2	1	7	6	5	0.111
37	1	3	2	1	2	3	1	2	3	3	7	8	0.778
38	1	3	2	1	2	3	1	3	2	3	8	7	0.778
39	1	3	2	1	2	3	2	1	3	4	6	8	0.444
40	1	3	2	1	2	3	2	3	1	4	8	6	0.444
41	1	3	2	1	2	3	3	1	2	5	6	7	0.111
42	1	3	2	1	2	3	3	2	1	5	7	6	0.111
43	1	3	2	1	3	2	1	2	3	3	8	7	0.778
44	1	3	2	1	3	2	1	3	2	3	9	6	1.000
45	1	3	2	1	3	2	2	1	3	4	7	7	0.333
46	1	3	2	1	3	2	2	3	1	4	9	5	0.778
47	1	3	2	1	3	2	3	1	2	5	7	6	0.111
48	1	3	2	1	3	2	3	2	1	5	8	5	0.333
49	1	3	2	2	1	3	1	2	3	4	6	8	0.444
50	1	3	2	2	1	3	1	3	2	4	7	7	0.333
51	1	3	2	2	1	3	2	1	3	5	5	8	0.333
52	1	3	2	2	1	3	2	3	1	5	7	6	0.111
53	1	3	2	2	1	3	3	1	2	6	5	7	0.111
54	1	3	2	2	1	3	3	2	1	6	6	6	0.000
:	:	:	:	:	:	:	:	:	:	:	:	:	:
73	2	1	3	1	2	3	1	2	3	4	5	9	0.778
74	2	1	3	1	2	3	1	3	2	4	6	8	0.444
75	2	1	3	1	2	3	2	1	3	5	4	9	0.778
:	:	:	:	:	:	:	:	:	:	:	:	:	:

Obs	A	A	A	B	B	B	C	C	C	R	R	R	W
	1	2	3	1	2	3	1	2	3	1	2	3	
181	3	2	1	1	2	3	1	2	3	5	6	7	0.111
182	3	2	1	1	2	3	1	3	2	5	7	6	0.111
183	3	2	1	1	2	3	2	1	3	6	5	7	0.111
184	3	2	1	1	2	3	2	3	1	6	7	5	0.111
185	3	2	1	1	2	3	3	1	2	7	5	6	0.111
186	3	2	1	1	2	3	3	2	1	7	6	5	0.111
187	3	2	1	1	3	2	1	2	3	5	7	6	0.111
188	3	2	1	1	3	2	1	3	2	5	8	5	0.333
189	3	2	1	1	3	2	2	1	3	6	6	6	0.000
190	3	2	1	1	3	2	2	3	1	6	8	4	0.444
191	3	2	1	1	3	2	3	1	2	7	6	5	0.111
192	3	2	1	1	3	2	3	2	1	7	7	4	0.333
193	3	2	1	2	1	3	1	2	3	6	5	7	0.111
194	3	2	1	2	1	3	1	3	2	6	6	6	0.000
195	3	2	1	2	1	3	2	1	3	7	4	7	0.333
196	3	2	1	2	1	3	2	3	1	7	6	5	0.111
197	3	2	1	2	1	3	3	1	2	8	4	6	0.444
198	3	2	1	2	1	3	3	2	1	8	5	5	0.333
199	3	2	1	2	3	1	1	2	3	6	7	5	0.111
200	3	2	1	2	3	1	1	3	2	6	8	4	0.444
201	3	2	1	2	3	1	2	1	3	7	6	5	0.111
202	3	2	1	2	3	1	2	3	1	7	8	3	0.778
203	3	2	1	2	3	1	3	1	2	8	6	4	0.444
204	3	2	1	2	3	1	3	2	1	8	7	3	0.778
205	3	2	1	3	1	2	1	2	3	7	5	6	0.111
206	3	2	1	3	1	2	1	3	2	7	6	5	0.111
207	3	2	1	3	1	2	2	1	3	8	4	6	0.444
208	3	2	1	3	1	2	2	3	1	8	6	4	0.444
209	3	2	1	3	1	2	3	1	2	9	4	5	0.778
210	3	2	1	3	1	2	3	2	1	9	5	4	0.778
211	3	2	1	3	2	1	1	2	3	7	6	5	0.111
212	3	2	1	3	2	1	1	3	2	7	7	4	0.333
213	3	2	1	3	2	1	2	1	3	8	5	5	0.333
214	3	2	1	3	2	1	2	3	1	8	7	3	0.778
215	3	2	1	3	2	1	3	1	2	9	5	4	0.778
216	3	2	1	3	2	1	3	2	1	9	6	3	1.000

W	Freq	Percent	Cumulative Percent
0	12	5.56	5.56
0.1111111111	90	41.67	47.22
0.3333333333	36	16.67	63.89
0.4444444444	36	16.67	80.56
0.7777777778	36	16.67	97.22
1	6	2.78	100.00

W	Freq	Percent	Cumulative Percent
1	6	2.78	2.78
0.7777777778	36	16.67	19.44
0.4444444444	36	16.67	36.11
0.3333333333	36	16.67	52.78
0.1111111111	90	41.67	94.44
0	12	5.56	100.00

216

9.2 Quade's Test

Application: Studying k treatments ($k \geq 3$). The assumptions are the same as those for Friedman's Test (k treatments studied in b blocks or b subjects).

	Treatment			
Block	1	2	\dots	k
1	x_{11}	x_{12}	\dots	x_{1k}
2	x_{21}	x_{22}	\dots	x_{2k}
3	x_{31}	x_{32}	\dots	x_{3k}
\dots	\dots	\dots	\dots	\dots
b	x_{b1}	x_{b2}	\dots	x_{bk}

- Quade's Test is a more powerful test than Friedman's Test when there are $k = 3$ treatments. That is, you are more likely to detect significant differences in treatment effects when differences really exist. Or, you are more likely to reject the null hypothesis when it is false.
- Quade's Test and Friedman's Test have similar power for $k = 4$ or $k = 5$ treatments.
- For $k \geq 6$ treatments, Friedman's Test is typically more powerful than Quade's Test.

Rationale:

- Recall that Friedman's Test is based entirely on within-block ranks. Thus, each block has equal weight in the calculation of the test statistic.
- However, if block effects differ in magnitude, we would like to use information about blocks to weight individual observations.
- Quade's Test was developed for such an application.

Hypotheses:

$$H_0 : M_1 = M_2 = \dots = M_k.$$

$$H_1 : M_i \neq M_{i'} \quad \text{for at least one } i \neq i'$$

where M_1, M_2, \dots, M_k represent the median responses associated with the k treatments.

- Thus, if H_0 is true, the treatment effects are identical, while if H_0 is false, then there is at least one treatment effect that is different from at least one other treatment effect.

Method: For a given α

- Assign ranks $1, 2, \dots, b$ to the observations within each block. For tied values, assign average ranks. Let R_{ij} be the rank assigned to block i and treatment j .
- Calculate the sample range within each block

$$\text{sample range in block } i = \max x_{ij} - \min x_{ij}$$

- Assign a second set of ranks to the blocks based on the sample ranges. Assign 1 to the block with the smallest range, 2 to the block with the next smallest range, and so on. For tied values, assign average ranks. Let Q_1, Q_2, \dots, Q_b be the ranks to blocks $1, 2, \dots, b$, respectively.
- Calculate the weighted observations

$$S_{ij} = Q_i \left[R(x_{ij}) - \frac{k+1}{2} \right].$$

Note that $(k+1)/2$ is the average rank within a block. Thus, S_{ij} is a weighted quantity based on the relative importance of the block.

- Calculate the sum of the weighted observations for each treatment

$$s_j = \sum_{i=1}^b S_{ij} \quad \text{for } j = 1, 2, \dots, k$$

- Calculate $A = \sum_{i=1}^b \sum_{j=1}^k S_{ij}^2$ and $B = \frac{1}{b} \sum_{j=1}^k S_j^2$.

If there are no ties, this reduces to $A = b(b+1)(2b+1)k(k+1)(k-1)/72$.

Test Statistic: The test statistic is $T = \frac{(b-1)B}{A-B}$.

Decision Rule

- If H_0 is true, then T is approximately distributed with the $F(k-1, (b-1)(k-1))$ distribution.
- Let F^* be the α critical value from the right tail of the $F(k-1, (b-1)(k-1))$ distribution.
- If $T > F^*$, then **Reject** H_0 . Otherwise, **Fail to Reject** H_0 .

EXAMPLE: Seven stores are selected for a marketing survey. Five different brands of hand lotion are placed side-by-side in identical sampling displays at each store. At the end of the week, the number of bottles of lotion sold for each brand is recorded. The following table contains the data:

Store	Brand					Store	Ranks				
	A	B	C	D	E		A	B	C	D	E
1	5	4	7	10	12	1					
2	1	3	1	0	2	2					
3	16	12	22	22	35	3					
4	5	4	3	5	4	4					
5	10	9	7	13	10	5					
6	19	18	28	37	58	6					
7	10	7	6	8	7	7					

Store	sample range	block rank Q_i	S_{ij} for Brand				
			A	B	C	D	E
1							
2							
3							
4							
5							
6							
7							

$S_j =$

$$A = \sum_{i=1}^b \sum_{j=1}^k S_{ij}^2 =$$

$$B = \frac{1}{b} \sum_{j=1}^k S_j^2 =$$

Test statistic $T =$

R output for Quade's Test

```
> ## Quade's Test
> ## Numbers of five brands of a new hand lotion sold
> ## in seven stores during one week.
```

```
      Brand
Store  A  B  C  D  E
  1    5  4  7 10 12
  2    1  3  1  0  2
  3   16 12 22 22 35
  4    5  4  3  5  4
  5   10  9  7 13 10
  6   19 18 28 37 58
  7   10  7  6  8  7
```

```
> quade.test(y)
```

```
      Quade test
```

```
Quade F = 3.8293, num df = 4, denom df = 24, p-value = 0.01519
```


R code for Quade's Test

```
## Quade's Test
## Numbers of five brands of a new hand lotion sold
## in seven stores during one week.

y <- matrix(c( 5,  4,  7, 10, 12,
              1,  3,  1,  0,  2,
              16, 12, 22, 22, 35,
              5,  4,  3,  5,  4,
              10,  9,  7, 13, 10,
              19, 18, 28, 37, 58,
              10,  7,  6,  8,  7),
            nr = 7, byrow = TRUE,
            dimnames = list(Store = as.character(1:7), Brand = LETTERS[1:5]))
y
quade.test(y)
```

9.2.1 Multiple Comparisons When Using Quade's Test

- If the null hypothesis is rejected, then multiple comparisons can be made.
- Treatments i and j are considered different if the following inequality is satisfied:

$$|S_i - S_j| > t^* \sqrt{\frac{2b(A - B)}{(b - 1)(k - 1)}}$$

where t^* is the $\alpha/2$ critical value from a t -distribution with $(b - 1)(k - 1)$ degrees of freedom.