Outline

- Wilcoxon Rank Sum test from SPSS
- The Kruskal-Wallis Test
- The Friedman Test

1

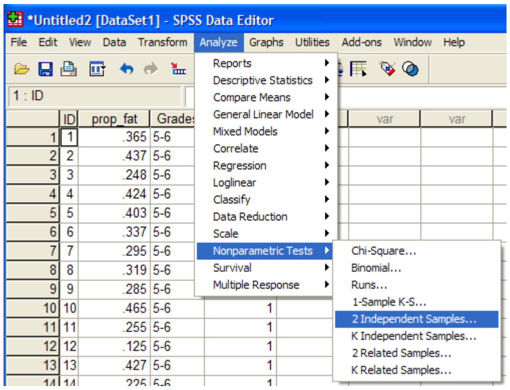
Wilcoxon Rank Sum test from SPSS

- Example 9.5 (p. 258); Excel file “Table 9.5 data (p. 258).xls”

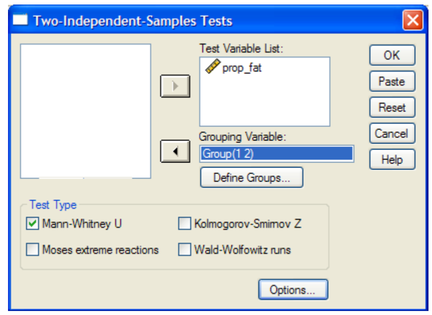
1. Data in Excel format (Groups should be in numeric codes)
2. Open file (Excel, first row variable names)
3. Request WRS test (Analyze→Nonparametric Tests→2 Independent Samples)
    - Prop_fat into “Test Variable List”
    - Group into “Grouping Variable”
    - “Define Groups” 1, 2 (known already)
    - “Test Type” Mann-Whtiney U is Wilcoxon Rank Sum
    - “Options” leave default
4. Two data tied, see bottom of Table 9.5 (How analysis change?)

2

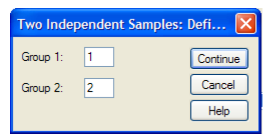
Request WRS test



Prop_fat into “Test Variable List”
Group into “Grouping Variable”



“Define Groups”
1, 2 (known already)



3

Ranks

	Group	N	Mean Rank	Sum of Ranks
prop_fat	1	14	16.04	224.50
	2	19	17.71	336.50
	Total	33		

Test Statistics[b]

	prop_fat
Mann-Whitney U	119.500
Wilcoxon W	224.500
Z	-.492
Asymp. Sig. (2-tailed)	.623
Exact Sig. [2*(1-tailed Sig.)]	.627[a]

a. Not corrected for ties.

b. Grouping Variable: Group

```
SAS

Wilcoxon Two-Sample Test
Statistic 224.5000

Normal Approximation
Z -0.4736
One-Sided Pr < Z 0.3179
Two-Sided Pr > |Z| 0.6358

t Approximation
One-Sided Pr < Z 0.3195
Two-Sided Pr > |Z| 0.6390

Z includes a continuity correction of 0.5.

Kruskal-Wallis Test
Chi-Square 0.2419
DF 1
Pr > Chi-Square 0.6229
```

Exact from Table B10, $\alpha = 0.01$
Two-sided
Critical region (168,308)
R_{WRS}=224.5
Cannot reject

4

The Kruskal-Wallis Test

- WRS for two independent populations
- Now, compare location of 2 or more independent populations
- For continuous data
- H_0: All medians are equal to one another
- H_1: At least two differ

- DATA setup (example):
  - **Homogeneous** set of subjects
  - Subjects assigned **without restrictions** to each of 3 groups
  - Change in DBP after four mos. from baseline
  - Data (n_1=8, n_2=15, n_3=16) on Table 9.8 (p.261)

5

How to in SPSS

- **Example 9.6 (p. 262); Excel file "Table 9.8 data (p. 261).xls"**

1. Type-in data in Excel (Groups should be in numeric codes)
2. Open file (Excel, first row variable names)
3. Rank by diff_DBP regardless of group (**Transform → Rank Cases**)
4. Sums of ranks per group (Analyze→Report→Report Summaries in Columns)
5. Rationale similar to WRS statistic, but unpractical, rather use

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$$

n_i : sample size of ith group

$n = \sum_{i=1}^{k} n_i$

k : number of groups

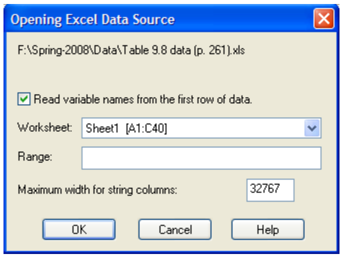
R_i : rank sum for the ith group

$H \sim \chi^2_{(k-1)}$ (H follows approximately a chi-square distribution with (k-1) degrees of freedom

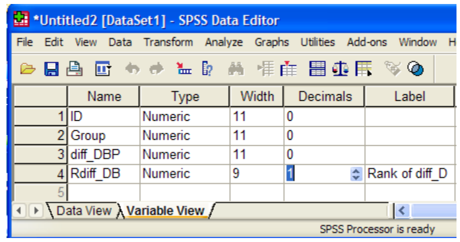
Reject H_0 if $H > \chi^2_{(k-1,1-\alpha)}$

6

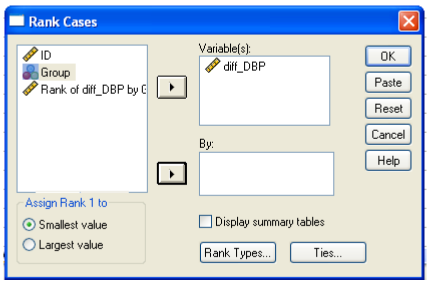
2) Open Excel file



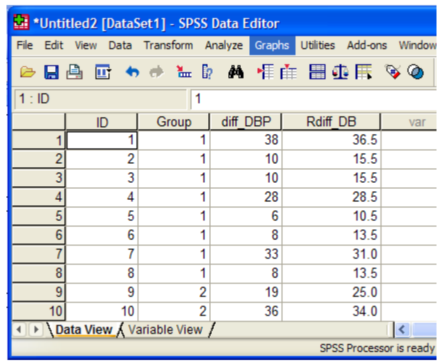
One decimal for ranks



3) Transform → Rank Cases

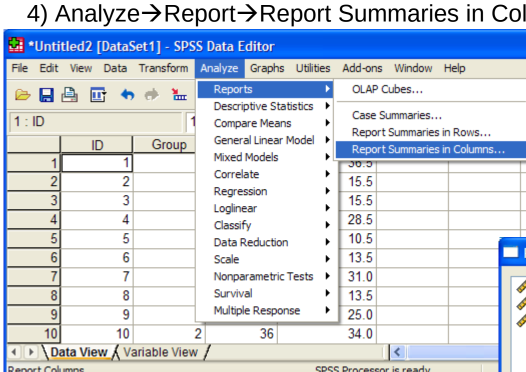


Data and Ranks

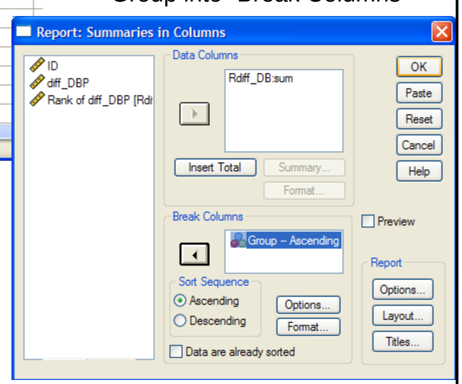


7

4) Analyze→Report→Report Summaries in Columns



Ranks into "Data Columns"
Group into "Break Columns"



Output

Group	Rank of diff DBP Sum
1	164.5
2	436.5
3	179.0

8

The Kruskal-Wallis test statistic

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$$

$$= \frac{12}{39(39+1)} \left[\frac{164.5^2}{8} + \frac{436.5^2}{15} + \frac{179^2}{16} \right] - 3(39+1) = 19.133$$

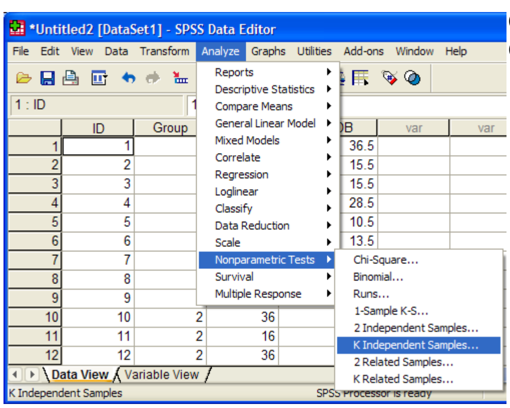
$$\alpha = 0.1$$

$H \sim \chi^2_{(2)}$ see Table B7 (p. 466)

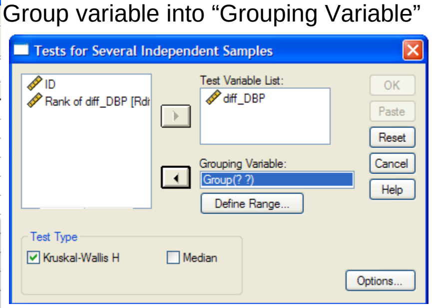
Since $H = 19.133 > 4.61 = \chi^2_{(2,0.9)}$

From Table B7 $p - value < 0.005$

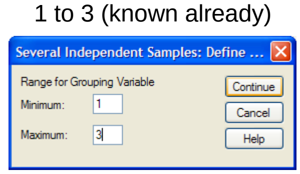
Kruskal-Wallis test directly from SPSS



Original variable into "Test Variable List"
Group variable into "Grouping Variable"



"Define Range"
1 to 3 (known already)



Ranks

	Group	N	Mean Rank
diff_DBP	1	8	20.56
	2	15	29.10
	3	16	11.19
	Total	39	

Test Statistics[a,b]

	diff_DBP
Chi-Square	19.207
df	2
Asymp. Sig.	.000

a. Kruskal Wallis Test

b. Grouping Variable: Group

The Kruskal-Wallis test from SAS

```
options nocenter formdlim='-' ;
libname aaa 'F:\Spring-2008\Data' ;
proc contents data=aaa.Table9_8 ;  run ;

PROC NPAR1WAY data=aaa.Table9_8 WILCOXON;
  CLASS GROUP;
  VAR REDUCTION;
RUN; quit;
```

```
The NPAR1WAY Procedure

         Wilcoxon Scores (Rank Sums) for Variable reduction
                    Classified by Variable group

                      Sum of      Expected       Std Dev          Mean
group        N        Scores      Under H0      Under H0         Score
---------------------------------------------------------------------
0            8        164.50         160.0     28.696466      20.56250
1           15        436.50         300.0     34.574335      29.10000
2           16        179.00         320.0     34.956383      11.18750

                 Average scores were used for ties.

   Kruskal-Wallis Test

Chi-Square          19.2065
DF                        2
Pr > Chi-Square      <.0001
```

11

The Friedman Test

- WRS for two independent populations
- K-W for 2 or more independent populations
- Now, compare location of 2 or more dependent populations

- For continuous data
- Matched subjects assigned to groups, different levels of matching
- Suitable for Randomized block design (ANOVA)
- H_0: All medians are equal to one another
- H_1: At least two differ

- Generalization of the sign test for more than 2 groups

12

Procedure

- Rank data separately for each block (matching level)
- Find sum of ranks for each of the comparison groups
- Use statistic

$$T = \frac{12}{bk(k+1)} \sum_{i=1}^{k} R_i^2 - 3b(k+1) \qquad \text{(similar to K-W } H \text{ statistic)}$$

b: number of blocks

k: number of comparison groups

R_i: rank sum for the ith group

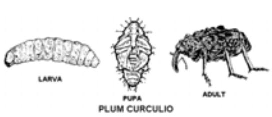
$T \sim \chi^2_{(k-1)}$ (T follows approximately a chi-square distribution with (k-1) degrees of freedom

Reject H_0 if $T > \chi^2_{(k-1,1-\alpha)}$

13

Example 9.7 (p. 263)

- Insecticide effectiveness
- Four blocks
- Data: number of living adult plum curculios, 5 insecticides+control
- NOTE: counts in (0,217); normality, equality of variances doubtful
- Table 9.9 (p. 263) shows conts (ranks), sum of ranks for comparison groups



14

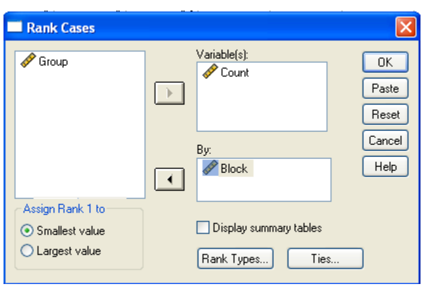
How to in SPSS

- **Example 9.7 (p. 263); Excel file "Table 9.9 data (p. 263).xls"**

1. Type-in data in Excel (Groups should be in numeric codes)
2. Open file (Excel, first row variable names)
3. Rank by diff_DBP within block (Transform → Rank Cases)
    - Count into "Variable(s)"
    - Group into "Block"
    - Uncheck "Display summary tables"
4. Sums of ranks per group (Analyze→Report→Report Summaries in Columns)
    - Rcount into "Data Columns"
    - Group into "Break Columns"
    - Rest as is
5. Get *T* statistic

15

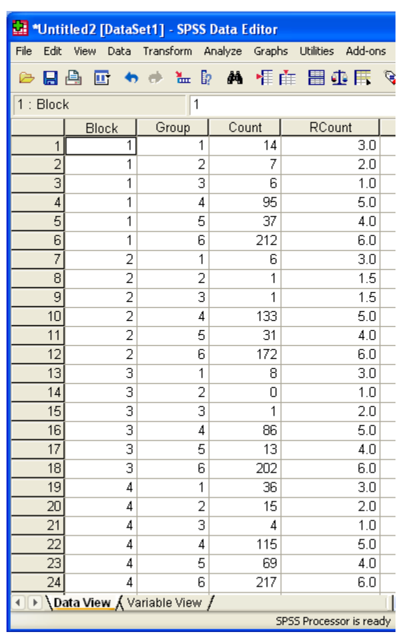
3) Count into "Variable(s)"
Group into "Block"



Group	Rank of Count by Block Sum
1	12.0
2	6.5
3	5.5
4	20.0
5	16.0
6	24.0

Analyze→Report→Report Summaries in Columns

Data after setting ranks to 1 decimal



	Block	Group	Count	RCount
1	1	1	14	3.0
2	1	2	7	2.0
3	1	3	6	1.0
4	1	4	95	5.0
5	1	5	37	4.0
6	1	6	212	6.0
7	2	1	6	3.0
8	2	2	1	1.5
9	2	3	1	1.5
10	2	4	133	5.0
11	2	5	31	4.0
12	2	6	172	6.0
13	3	1	8	3.0
14	3	2	0	1.0
15	3	3	1	2.0
16	3	4	86	5.0
17	3	5	13	4.0
18	3	6	202	6.0
19	4	1	36	3.0
20	4	2	15	2.0
21	4	3	4	1.0
22	4	4	115	5.0
23	4	5	69	4.0
24	4	6	217	6.0

16

The Friedman test statistic

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

$$= \frac{12}{4(6)(6+1)} \left[12^2 + 6.5^2 + 5.5^2 + 20^2 + 16^2 + 24^2 \right] - 3(4)(6+1) = 19.5$$

$$\alpha = 0.05$$

$H \sim \chi^2_{(5)}$ see Table B7 (p. 466)

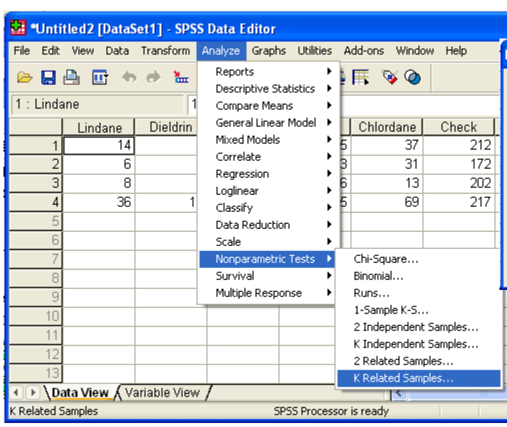
Since $H = 19.5 > 11.07 = \chi^2_{(5, 0.95)}$

From Table B7 $p - value < 0.005$

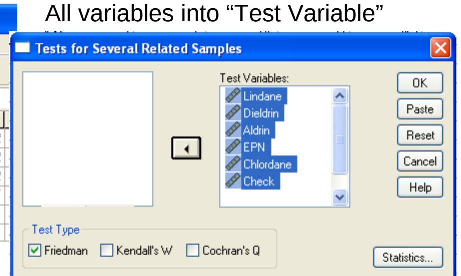
17

Friedman test directly from SPSS

Data: Groups in columns, Blocks in rows



All variables into "Test Variable"



Ranks

	Mean Rank
Lindane	3.00
Dieldrin	1.63
Aldrin	1.38
EPN	5.00
Chlordane	4.00
Check	6.00

Test Statistics[a]

N	4
Chi-Square	19.604
df	5
Asymp. Sig.	.001

a. Friedman Test

18

The Friedman test from SAS

```
options nocenter formdlim='-' ;
data insect;
input block insecticide number;
datalines;
1 1  14
1 2   7
1 3   6
1 4  95
1 5  37
1 6 212
2 1   6
2 2   1
2 3   1
2 4 133
2 5  31
2 6 172
3 1   8
3 2   0
3 3   1
3 4  86
3 5  13
3 6 202
4 1  36
4 2  15
4 3   4
4 4 115
4 5  69
4 6 217
;
proc freq;
  tables block*insecticide*number/cmh2 scores=rank noprint;
run;
```

The FREQ Procedure

Summary Statistics for insecticide by number
Controlling for block

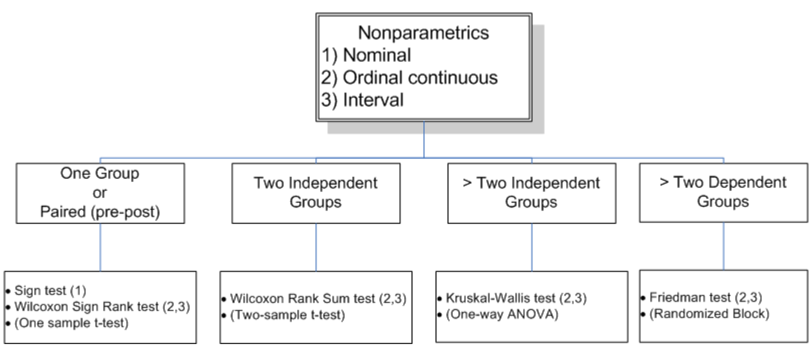
Cochran-Mantel-Haenszel Statistics (Based on Rank Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	10.9034	0.0010
2	Row Mean Scores Differ	5	19.6043	0.0015

Total Sample Size = 24

19

Nonparametrics roadmap



20