Outline

- The Wilcoxon Signed Rank Test
- The Wilcoxon Rank Sum Test

1

The Wilcoxon Signed Rank Test

- For continuous data
- Regression to the mean Example (pp. 253-255)
  - Data on 14 boys with extreme “day 1” values
  - Change in direction of mean (**+**)
  - Change away from mean (**-**)
  - If same value, no sign assigned (excluded from analysis)
  - Rank absolute differences (all, ascending)
  - Sum ranks separately (**+/-**)
  - Use $\mathbf{R_{WSR}}$: signed rank sum for positive differences (those changing toward the mean) as test statistic

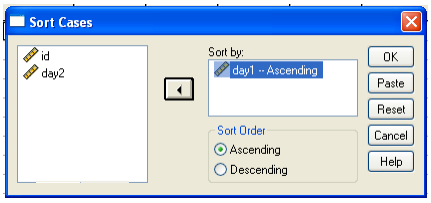
2

How to in SPSS

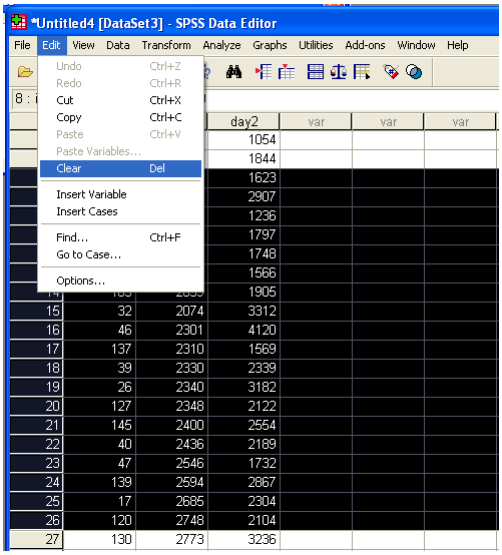
1. Open file (Excel, first row variable names)
2. Sort by day1 (Data → Sort Cases…)
3. Clear rows 8-26 (Hold-down left click on row 8, drag down to row 26 →Edit→Clear)
4. Rank by day1 (Transform → Rank Cases)
5. Sort by id (Data→Sort Cases)
6. Create “toward” (Transform→Compute→’Target Variable’ toward→”If” ::: →’Numeric Expression’)
7. Create ”away” (ibid but :::)
8. Create rtm (same as “toward” and “away” but on same column)
9. Find ranks by rtm (Transform → Rank Cases)
10. Create +/- groups (Transform→Compute)
11. Sums from group ranks (Analyze→Reports→Report Summaries in Columns)

3

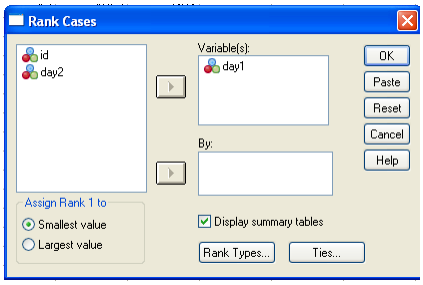
2. Sort cases by day1



3. Clear rows 8-26

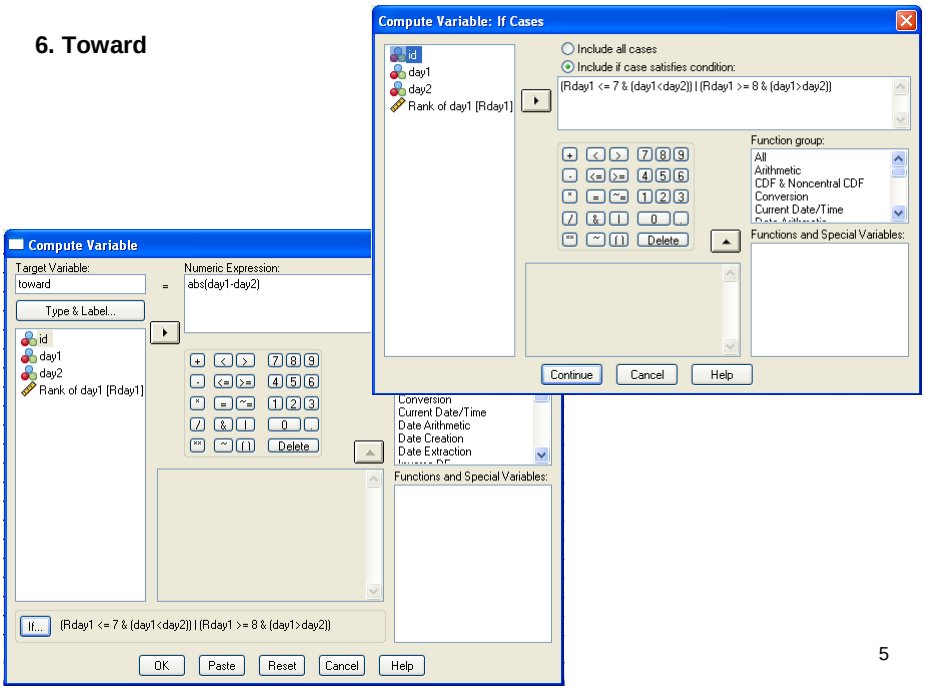


4. Rank data by day1



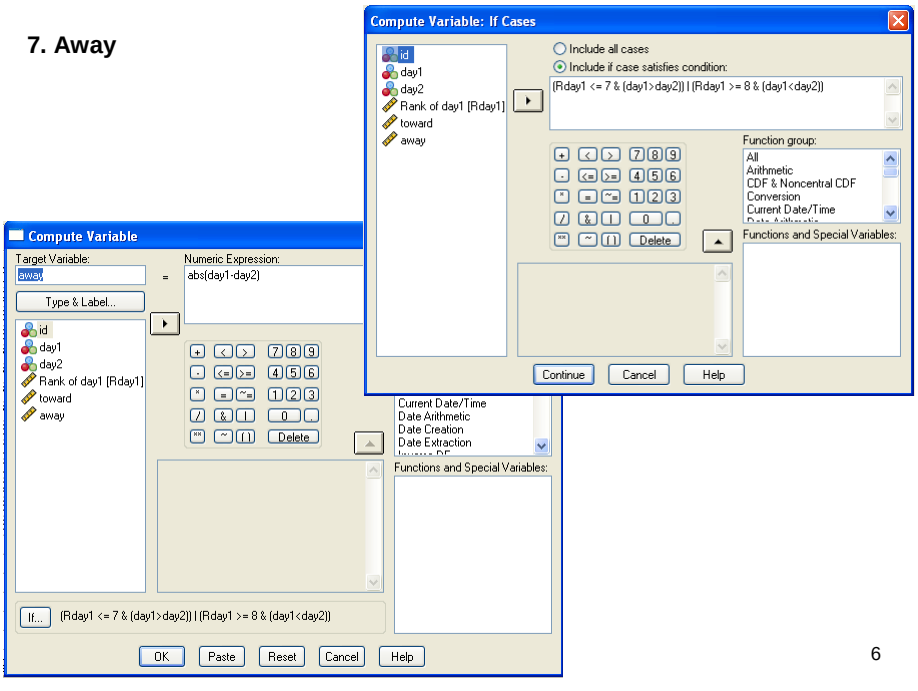
4

6. Toward



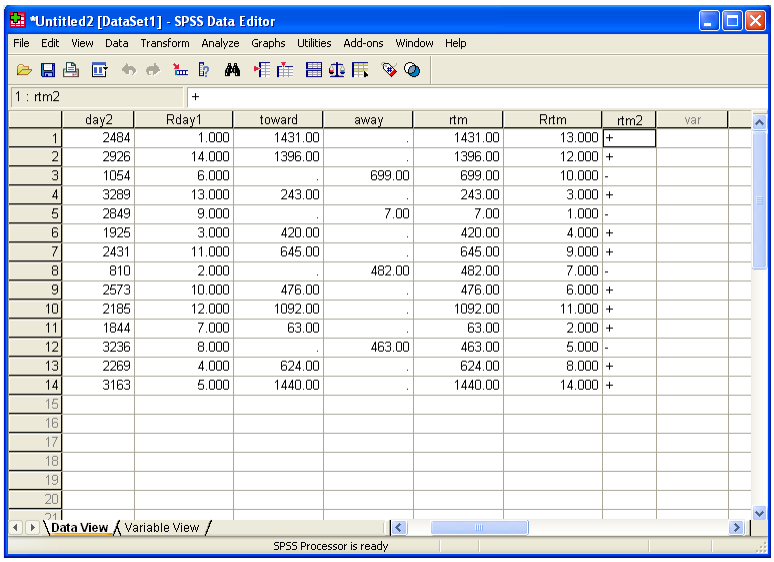
5

7. Away

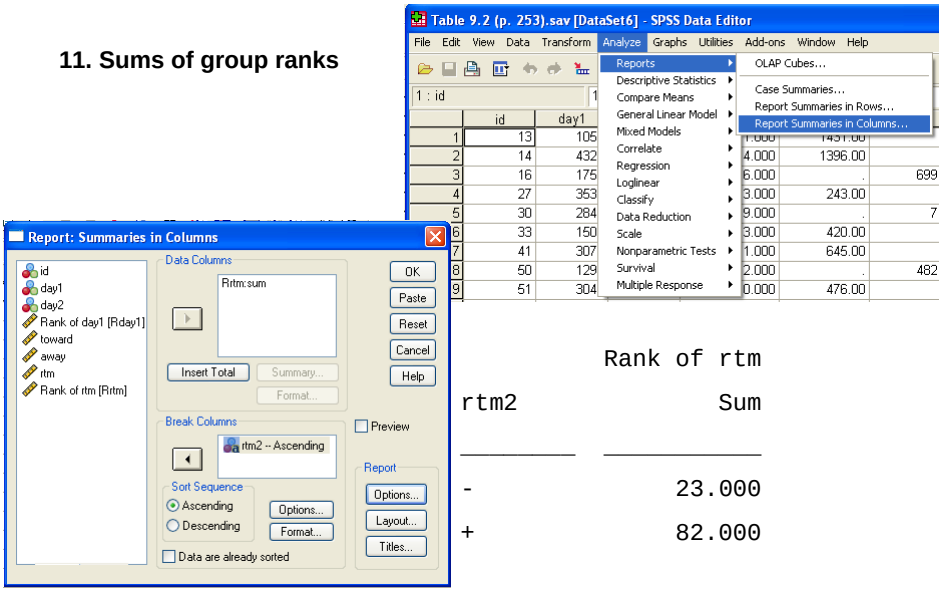


6

Data after step 10



11. Sums of group ranks



rtm2	Rank of rtm Sum
-	23.000
+	82.000

Rationale of test

- For n observations, ranks are *1, …, n*
- Sum of ranks is $n(n+1)/2$, and average rank is $(n+1)/2$
- Null hypothesis
- If the null is true, distribution of differences is symmetric
  - Half positive, half negative
- If null hypothesis is true, the sum of the ranks should be $n/2$ times the average rank, i.e. $(n/2)(n+1)/2 = n(n+1)/4$
- Tables 9.3 and 9.4 exemplify the construction of boundaries of the rejection region as shown in Table B9 for n<30 (469):
  - N on first column, then
  - Two sided
  - One sided

9

Example 9.3 (p. 255)

- N=14, $\alpha = 0.05$ $\alpha = 0.05$
- Null hypothesis $H_0 : M_d = 0$ (median of differences equals zero, no regression toward the mean)
- Alternative $H_0 : M_d > 0$ (median differences larger than zero, regression toward the mean)
- Critical values $(25, 80)$ (if R_{WSR} is larger than or equal to 80, reject)
- Since $R_{WSR} = 82$ we can reject the null

- This result differs with that from Sign test (didn't reject):
  - Wilcoxon Sum Rank test has larger power
  - WSR uses more information, not only counts of changes toward the mean
  - Average rank away=5.75 vs. average rank toward=7.5
  - P-value from sig test is 0.0898, not so far from 0.05

10

TIES

When applying WSR, two types of ties can occur:

- The difference is zero (e.g. regression to the mean data) or the observed value equals the hypothesized value (e.g. Example 9.4 next).
  - REMOVE observation(s) from analysis;
  - FEW ties of this type are allowed
- Two or more differences have the same nonzero value.
  - CONVENTION: use average rank (e.g. 2,3→2.5; 2,3,4→3).
  - Few ties of this type: still can use WSR statistic (but results are approximate)
  - Many ties of this type: adjustment for ties needed or next Chapter (contingency table)

$$R_{WSR} \quad H_0 : M_d = 0 \qquad R_{WSR} = 82$$
$$(25,80) \quad H_0 : M_d > 0$$

11

Example 9.4

DATA: Lead concentration in one lab (n=13)

- Are measurements significantly different from 41?

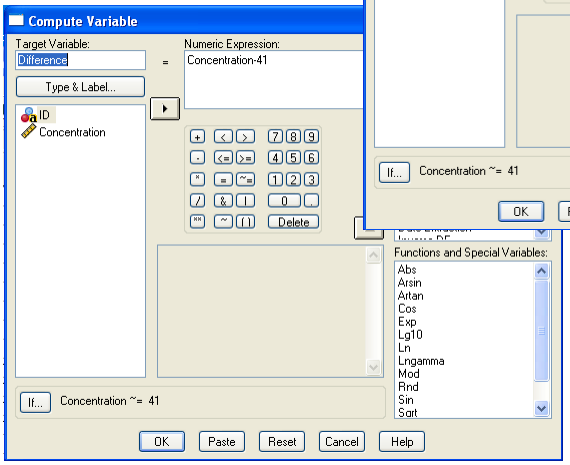
$$H_0 : M_d = 41 \quad \text{vs.} \quad H_0 : M_d \neq 41 \qquad \alpha = 0.05$$

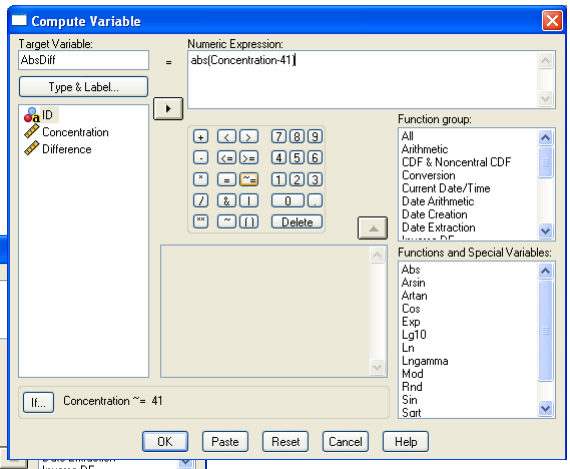
- Eleventh datum dropped, new n=12
- Test statistic $R_{WSR} = 58.5$

- From Table B9 (p. 469): n=12, $\alpha \leq 0.05$, two-sided the boundaries are (13,65)
- 58.5<65 → cannot reject
- WSR decision agrees with sign test

12

Difference from 41

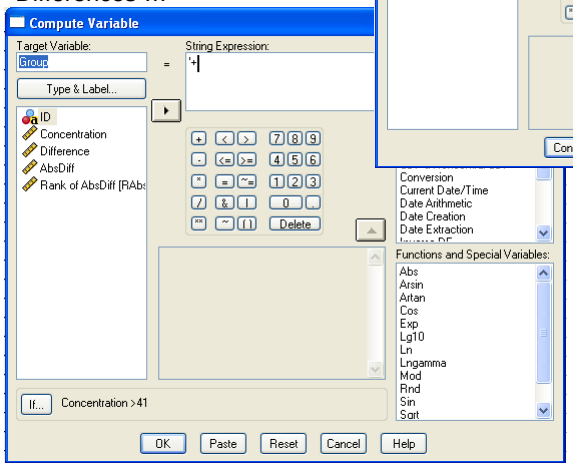


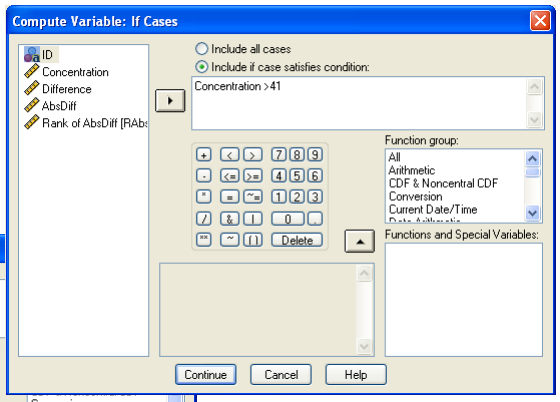


Absolute difference from 41

13

Define group with positive Differences …



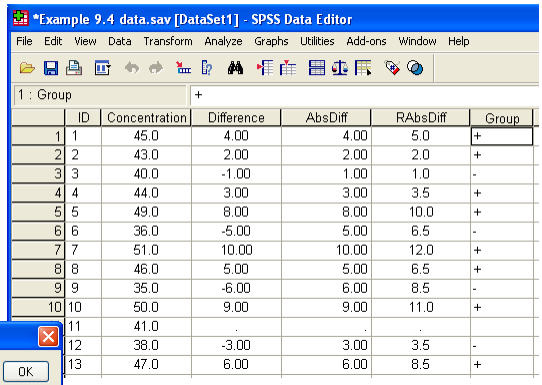


… for cases with concentration larger than 41

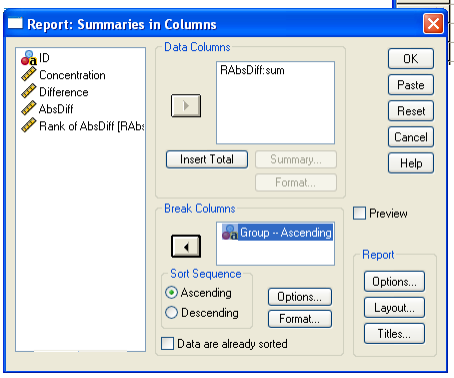
Repeat for concentrations smaller than 41, assign '-'

14

Data



Getting sum of ranks by group



Group	Rank of AbsDiff Sum
-	19.5
+	58.5

15

Normal Approximation to WSR

- For $n \geq 16$, R_{WSR} follows approx. a normal distribution
- Can use

$$\frac{\left|R_{WSR} - \overbrace{[n(n+1)/4]}^{\text{Expected value}}\right| - 0.5}{\underbrace{\sqrt{n(n+1)(2n+1)/24}}_{\text{Variance}}} \quad \text{Continuity correction}$$

- For Example 9.3 (recall one-sided)

$$\frac{\left|82 - [14(14+1)/4]\right| - 0.5}{\sqrt{14(14+1)(2(14)+1)/24}} = \frac{\left|82 - 52.5\right| - 0.5}{15.93} = 1.82$$

- From Table B4, p-value=P[Z>1.82]=0.0344
- This p-value is close to the exact (0.0338). Approximation is good for this dataset

- ***Study last paragraph of Section 9.3 (p. 257)***

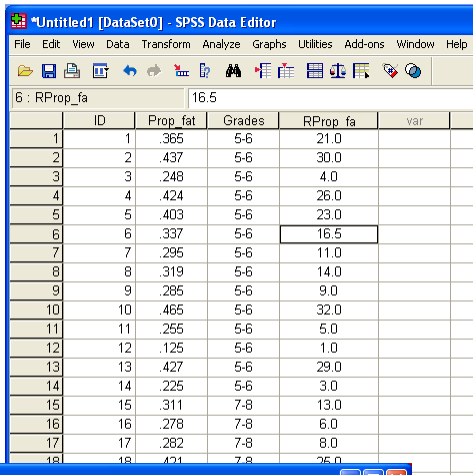
16

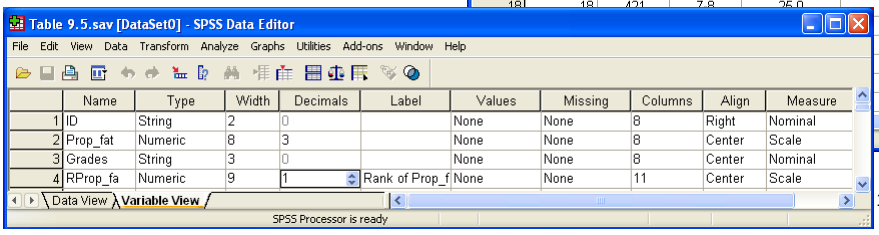
The Wilcoxon Rank Sum Test

- *Sign and WSR test for one group or paired data*
- Now assume we have two groups
- WRS is also known as Mann-Whitney (unequal sample sizes)
- Data: proportion of calories from fat in two groups

17

1. Type in data (ID, Prop_fat, Grades), for one case
2. Change variable Name
3. Change data Type
4. Change Alignment
5. Rank all data
6. Get sum of ranks per group (Analize→Reports→Report Summaries in Columns)





18

• Data for two cases were rounded to allow a tie

•Test statistic is the sum of ranks of smaller sample

$$R_{WRS} = 224.5, \quad n_1 = 14, \quad n = 33$$

Grades	Rank of Prop fat Sum
5-6	**224.5**
7-8	336.5

• RATIONALE:

• If there were no differences between groups, then the rank sum for the smaller group is the product $n_1(n+1)/2 = 14(33+1)/2 = 238$

• A value of R_{WRS} far away from the expected under the null are grounds for rejecting the hypothesis of no difference

• Use critical values on Table B10 (p. 470-473)

- N_1 smaller group, N_2 larger group
- two-sided or one-sided
- Critical regions are determined similarly to WSR test
- An example for n=8, n_1=4 is described on page 258 and Tables 9.6 and 9.7 show relevant counts and probabilities

19

• Example 9.5 (p. 260)

• Null: no difference in proportion of calories from fat in these two groups

• Alternative: there is a difference

- Two-sided, let's do $\alpha \le 0.01$
- Will reject if extremely large or small values of R_{WRS}

• Observed $R_{WRS} = 224.5, \quad n_1 = 14, \quad n_2 = 19$

• From Table B10 (p. 473) critical values are (168,308)

• Reject the null if observed R_{WRS} is smaller than 168 or larger than 308

• $R_{WRS} = 224.5$ is not in the rejection region. Cannot reject the null.

• Conclusion: based on the WRS test there is no evidence that grade groups differ in the proportion of calories from fat

N_1

20

Normal Approximation to WRS

- If Table B10 cannot accommodate data or $n_1 > 8$ and $n_2 > 8$
- Can use $$\frac{\left|R_{WRS} - \overbrace{[n_1(n+1)/2]}^{\text{Expected value}}\right| - 0.5}{\sqrt{\underbrace{n_1 n_2 (n+1)/12}_{\text{Variance}}}}$$ Continuity correction
- For Example 9.3 (recall two-sided) $R_{WRS} = 224.5, \quad n_1 = 14, \quad n_2 = 19$

$$\frac{\left|224.5 - [14(33+1)/2]\right| - 0.5}{\sqrt{14(19)(33+1)/12}} = \frac{\left|224.5 - 238\right| - 0.5}{27.453} = 0.4735$$

- From Table B4, p-value=2P[Z>0.4735]=2(0.3174)=0.63 (interpolation)
- For many ties use adjustment or next Chapter approach

21