

# EE3211 Modelling Techniques

# Lecture 7

## Nonparametric Methods

# Overview

## **Parametric statistical methods:**

- For estimation and hypothesis testing
- We assume that the parametric form of the distribution is known

## **Nonparametric statistical methods**

- If no assumptions about the shape of the distribution
- Cannot apply central-limit theorem due to small sample size ( $n$ )
- Make fewer assumptions about the distributional shape

# Types of Data

- **Categorical variables:** two or more categories that do not have any ordering  
e.g. race and ethnicity
- **Dichotomous variables:** two possible values  
e.g. gender, death, disease status
- **Ordinal variables:** more than two ranked or ordered values  
e.g., amount of current smoking: none, <10/day, 10-20/day, 21-30/day, >30/day

# Ordinal data

- Measure relative ordering of different categories

E.g. two people: A and B

→ Whether score for A is  $>$ ,  $<$ , or  $=$  to the score for B  
~~relative magnitude of the differences~~



# Example on Dermatology

- Objective: compare the effectiveness of two ointments (A, B) in reducing excessive redness in people who cannot otherwise be exposed to sunlight
- Ointment A: randomly applied to either the left or right arm
- Ointment B: applied to the corresponding area on the other arm.
- Person is then exposed to 1 hour of sunlight
- Compare the degrees of redness of the two arms
- Make qualitative assessments:
  1. Arm A is not as red as arm B
  2. Arm B is not as red as arm A
  3. Both arms are equally red
- Of 45 people tested with the condition, 22 are better off on arm A, 18 are better off on arm B, 5 are equally well off on both arms.
- Q: How can we decide whether this evidence is enough to conclude that ointment A is better than ointment B?

# Large-sample Method

- Degree of redness can be measured on a quantitative scale (higher number indicates more redness)
- Let  $x_i$  = degree of redness on arm A for the  $i$ th person  
 $y_i$  = degree of redness on arm B for the  $i$ th person  
 $d_i = x_i - y_i$  = difference in redness between the A and B arms for the  $i$ th participant
- $H_0: \Delta = 0$  vs.  $H_1: \Delta \neq 0$  ( $\Delta$  : the population median of the  $d_i$  / the 50<sup>th</sup> percentile of the underlying distribution of the  $d_i$ )
  - (i) if  $\Delta = 0$  : ointments are equally effective
  - (ii) if  $\Delta < 0$  : ointment A is better (arm A is less red than arm B)
  - (iii) if  $\Delta > 0$  : ointment B is better (arm A is redder than arm B)
- $d_i$  cannot be observed, can only observe whether:
  - (i)  $d_i > 0$
  - (ii)  $d_i < 0$
  - (iii)  $d_i = 0$

# The Sign Test

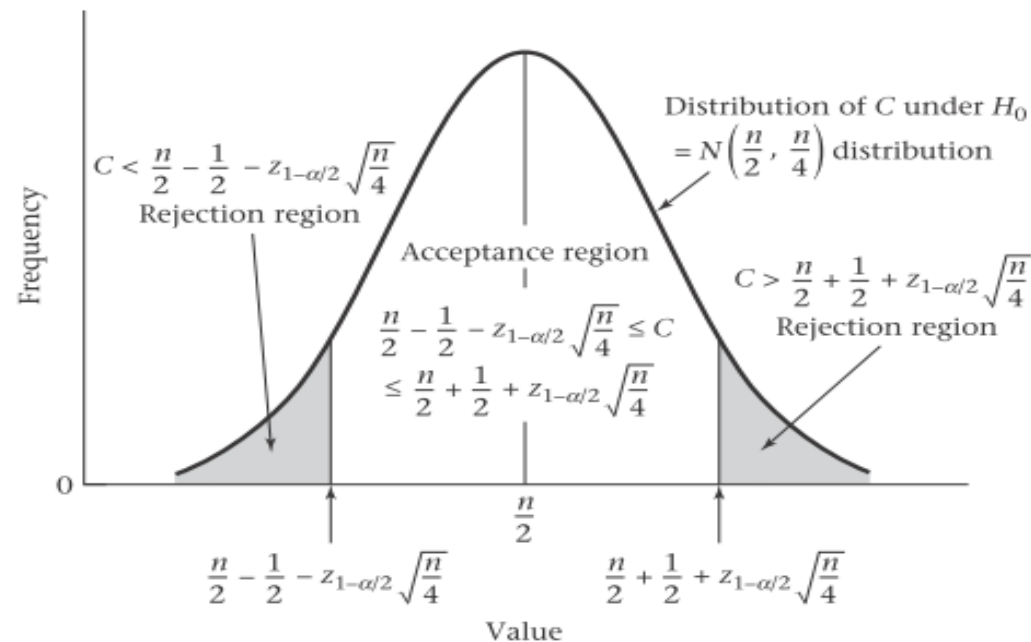
- Depends only on the sign of the difference, not their actual magnitude
- To test the hypothesis  $H_0: \Delta = 0$  vs.  $H_1: \Delta \neq 0$   
\*number of nonzero  $d_i$ 's =  $n \geq 20$  and  $C$  = the number of  $d_i$ 's where  $d_i > 0$ ,

if

$$C > c_2 = \frac{n}{2} + \frac{1}{2} + z_{1-\alpha/2} \sqrt{n/4} \quad \text{or} \quad C < c_1 = \frac{n}{2} - \frac{1}{2} - z_{1-\alpha/2} \sqrt{n/4}$$

→ Reject  $H_0$ ; otherwise accept  $H_0$

Figure 9.1 Acceptance and rejection regions for the sign test





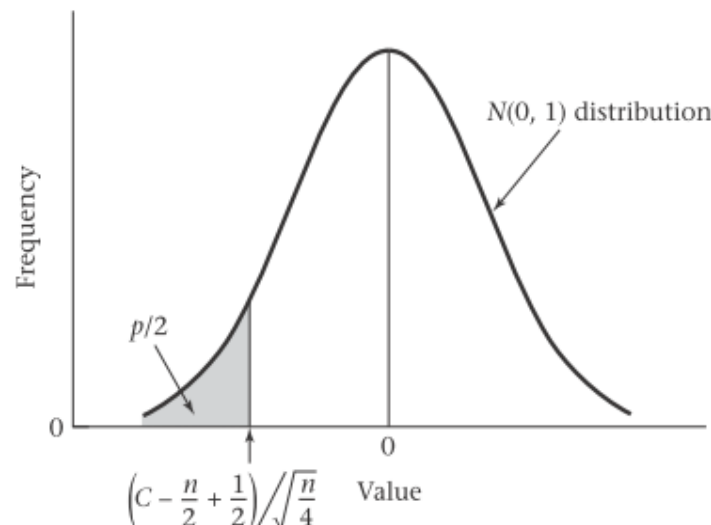
# Computation of the $p$ -Value for the Sign Test (Normal-Theory Method)

$$p = 2 \times \left[ 1 - \Phi \left( \frac{C - \frac{n}{2} - .5}{\sqrt{n/4}} \right) \right] \quad \text{if } \underline{C > \frac{n}{2}}$$

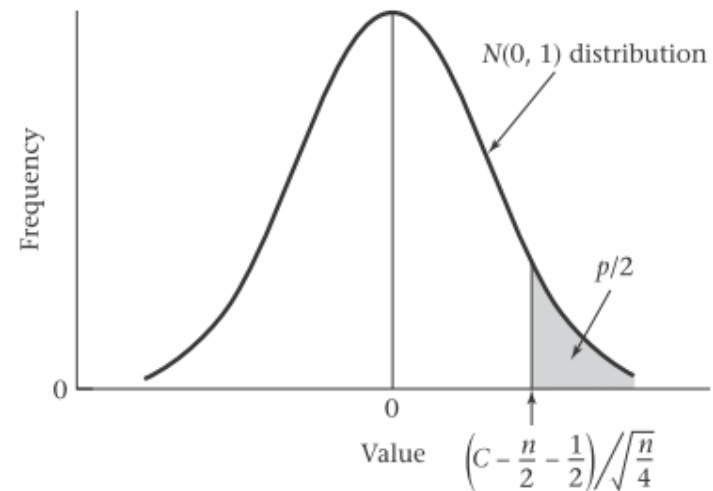
$$\underline{p = 1.0} \quad \text{if } \underline{C = \frac{n}{2}}$$

$$p = 2 \times \Phi \left( \frac{C - \frac{n}{2} + .5}{\sqrt{n/4}} \right) \quad \text{if } \underline{C < \frac{n}{2}}$$

Figure 9.2 Computation of the  $p$ -value for the sign test



If  $C < n/2$ , then  $p = 2 \times$  area to the left of  $\left( C - \frac{n}{2} + \frac{1}{2} \right) / \sqrt{\frac{n}{4}}$  under an  $N(0, 1)$  distribution



If  $C > n/2$ , then  $p = 2 \times$  area to the right of  $\left( C - \frac{n}{2} - \frac{1}{2} \right) / \sqrt{\frac{n}{4}}$  under an  $N(0, 1)$  distribution



# R commands to perform the sign test

#x=number of subjects with  $d_i > 0$

#n=total number of subjects with  $d_i$  not equal to 0

#codes for one-sample binomial test with  $p_0 = 0.5$  with continuity correction

```
>prop.test(x,n,p=0.5, alternative="two.sided", correct=TRUE)
```

# Computation of the $p$ -Value for the Sign Test (Exact Method)

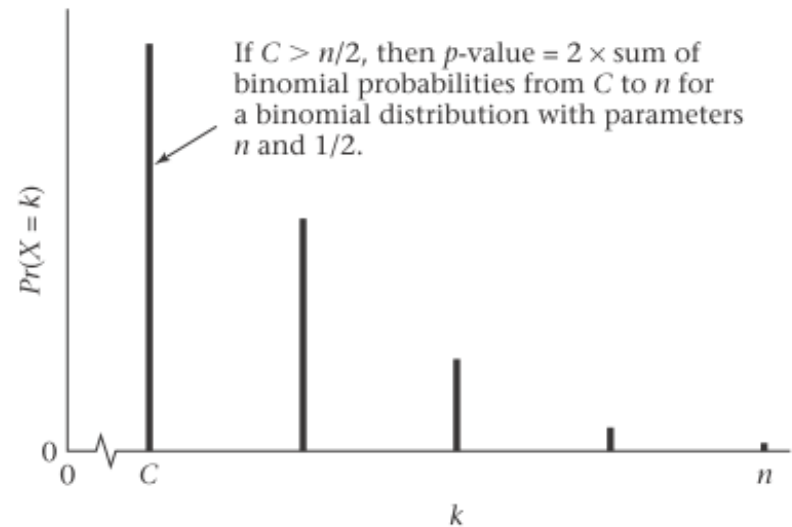
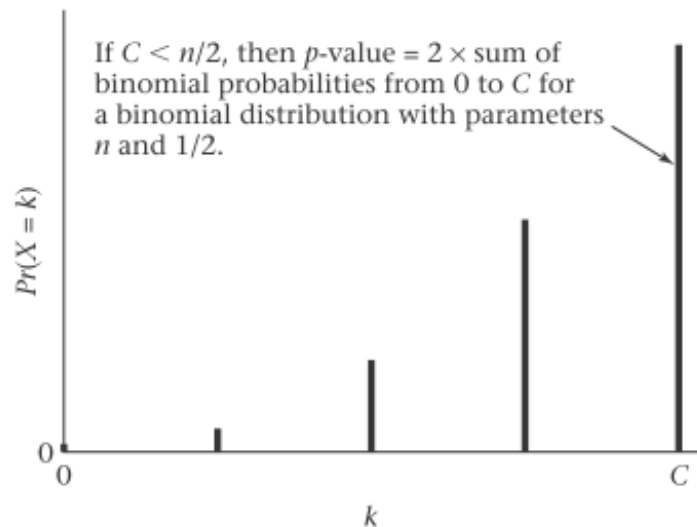
- $n < 20$ : exact binomial probabilities to compute the  $p$ -value
- If  $C$  is very large or very small: reject  $H_0$

$$\text{If } C > n/2, \quad p = 2 \times \sum_{k=C}^n \binom{n}{k} \left(\frac{1}{2}\right)^n$$

$$\text{If } C < n/2, \quad p = 2 \times \sum_{k=0}^C \binom{n}{k} \left(\frac{1}{2}\right)^n$$

$$\text{If } C = n/2, \quad p = 1.0$$

**Figure 9.3** Computation of the  $p$ -value for the sign test (exact test)



# Example on Sign Test - Ophthalmology

- Suppose we wish to compare two different types of eye drops (A, B) that are intended to prevent redness in people with hay fever.
- Drug A is randomly administered to one eye and drug B to the other eye.
- The redness is noted at baseline and after 10 minutes by an observer who is unaware of which drug has been administered to which eye.
- We find that for 15 people with an equal amount of redness in each eye at baseline, after 10 minutes the drug A eye is less red than the drug B eye for 2 people ( $d_i < 0$ ); the drug B eye is less red than the drug A eye for 8 people ( $d_i > 0$ ); and the eyes are equally red for 5 people ( $d_i = 0$ ).
- **Q: Assess the statistical significance of the results.**



# Example on Sign Test - Ophthalmology

## Solution:

The test is based on the 10 people who had a differential response to the two types of eye drops.

Because  $n = 10 < 20 \rightarrow$  cannot use normal-theory method  
 $\rightarrow$  exact method due to:

$$C = 8 > \frac{10}{2} = 5, p = 2 \times \sum_{k=8}^{10} \binom{10}{k} (1/2)^{10}$$

- binomial tables using

$n = 10, p = 0.5$ , and note that  $\Pr(X = 8) = .0439$ ,

$\Pr(X = 9) = .0098, \Pr(X = 10) = .001$

$$p = 2 \times \Pr(X \geq 8) = 2(.0439 + 0.0098 + 0.0010)$$

$$= 2 \times 0.0547 = 0.109 \text{ (not statistically significant)}$$

- accept  $H_0$ : two types of eye drops are equally effective in reducing redness in people with hay fever



TABLE 1

Exact binomial probabilities  $\Pr(X = k) = \binom{n}{k} p^k q^{n-k}$  (continued)

n	k	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
9	4	.0004	.0048	.0185	.0459	.0865	.1361	.1875	.2322	.2627	.2734
	5	.0000	.0004	.0026	.0092	.0231	.0467	.0808	.1239	.1719	.2188
	6	.0000	.0000	.0002	.0011	.0038	.0100	.0217	.0413	.0703	.1094
	7	.0000	.0000	.0000	.0001	.0004	.0012	.0033	.0079	.0164	.0313
	8	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0007	.0017	.0039
	0	.6302	.3874	.2316	.1342	.0751	.0404	.0207	.0101	.0046	.0020
	1	.2985	.3874	.3679	.3020	.2253	.1556	.1004	.0605	.0339	.0176
	2	.0629	.1722	.2597	.3020	.3003	.2668	.2162	.1612	.1110	.0703
	3	.0077	.0446	.1089	.1762	.2336	.2668	.2716	.2508	.2119	.1641
10	4	.0006	.0074	.0283	.0661	.1168	.1715	.2194	.2508	.2600	.2461
	5	.0000	.0008	.0050	.0165	.0389	.0735	.1181	.1672	.2128	.2461
	6	.0000	.0001	.0006	.0028	.0087	.0210	.0424	.0743	.1160	.1641
	7	.0000	.0000	.0000	.0003	.0012	.0039	.0098	.0212	.0407	.0703
	8	.0000	.0000	.0000	.0000	.0001	.0004	.0013	.0035	.0083	.0176
	9	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0008	.0020	.0046
	0	.5987	.3487	.1969	.1074	.0563	.0282	.0135	.0060	.0025	.0010
	1	.3151	.3874	.3474	.2884	.1877	.1211	.0725	.0403	.0207	.0098
	2	.0746	.1937	.2759	.3020	.2816	.2335	.1757	.1209	.0763	.0439
11	3	.0105	.0574	.1298	.2013	.2503	.2668	.2522	.2150	.1665	.1172
	4	.0010	.0112	.0401	.0881	.1460	.2001	.2377	.2508	.2384	.2051
	5	.0001	.0015	.0085	.0264	.0584	.1029	.1536	.2007	.2340	.2461
	6	.0000	.0001	.0012	.0055	.0162	.0368	.0689	.1115	.1596	.2051
	7	.0000	.0000	.0001	.0008	.0031	.0090	.0212	.0425	.0746	.1172
	8	.0000	.0000	.0000	.0001	.0004	.0014	.0043	.0106	.0229	.0439
	9	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0016	.0042	.0098
	10	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010
	0	.5688	.3138	.1673	.0859	.0422	.0198	.0088	.0036	.0014	.0005
12	1	.3293	.3835	.3248	.2362	.1549	.0932	.0518	.0266	.0125	.0054
	2	.0867	.2131	.2866	.2953	.2581	.1998	.1395	.0887	.0513	.0269
	3	.0137	.0710	.1517	.2215	.2581	.2568	.2254	.1774	.1259	.0806
	4	.0014	.0158	.0536	.1107	.1721	.2201	.2428	.2365	.2060	.1611
	5	.0001	.0025	.0132	.0388	.0803	.1321	.1830	.2207	.2360	.2256
	6	.0000	.0003	.0023	.0097	.0268	.0566	.0985	.1471	.1931	.2256
	7	.0000	.0000	.0003	.0017	.0064	.0173	.0379	.0701	.1128	.1611
	8	.0000	.0000	.0000	.0002	.0011	.0037	.0102	.0234	.0462	.0806
	9	.0000	.0000	.0000	.0000	.0001	.0005	.0018	.0052	.0126	.0269
13	10	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0007	.0021	.0054
	11	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0005
	0	.5404	.2824	.1422	.0687	.0317	.0138	.0057	.0022	.0008	.0002
	1	.3413	.3766	.3012	.2062	.1267	.0712	.0368	.0174	.0075	.0029
	2	.0988	.2301	.2924	.2835	.2323	.1678	.1088	.0639	.0339	.0161
	3	.0173	.0852	.1720	.2362	.2581	.2397	.1954	.1419	.0923	.0537
	4	.0021	.0213	.0683	.1329	.1936	.2311	.2367	.2128	.1700	.1208
	5	.0002	.0038	.0193	.0532	.1032	.1585	.2039	.2270	.2225	.1934
	6	.0000	.0005	.0040	.0155	.0401	.0792	.1281	.1766	.2124	.2256
14	7	.0000	.0000	.0006	.0033	.0115	.0291	.0591	.1009	.1489	.1934
	8	.0000	.0000	.0001	.0005	.0024	.0078	.0199	.0420	.0762	.1208
	9	.0000	.0000	.0000	.0001	.0004	.0015	.0048	.0125	.0277	.0537
	10	.0000	.0000	.0000	.0000	.0000	.0002	.0008	.0025	.0068	.0161
	11	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	.0029
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002
	0	.5133	.2542	.1209	.0550	.0238	.0097	.0037	.0013	.0004	.0001
	1	.3512	.3672	.2774	.1787	.1029	.0540	.0259	.0113	.0045	.0016
	2	.1109	.2448	.2937	.2680	.2059	.1388	.0836	.0453	.0220	.0095
	3	.0214	.0997	.1900	.2457	.2517	.2181	.1651	.1107	.0660	.0349
15	4	.0028	.0277	.0638	.1335	.2097	.2337	.2222	.1845	.1350	.0873
	5	.0003	.0055	.0266	.0691	.1258	.1803	.2154	.2214	.1989	.1571

# Modified Scenario on Dermatology

- New assumption: degree of burn can be quantified on a 10-point scale (10: worst burn; 1 no burn)
- Compute  $d_i = x_i - y_i$  and  $x_i = \text{degree of burn}$  for ointment A and  $y_i = \text{degree of burn}$  for ointment B
  - If  $d_i > 0$  : ointment B is doing better than ointment A
  - If  $d_i < 0$  : ointment A is doing better than ointment B
  - E.g.  $d_i = +5$  : degree of redness is 5 units > on the ointment A arm than on the ointment B arm
  - $d_i = -3$  : degree of redness is 3 units < on the ointment A arm than on the ointment B arm
- **Q: how can this additional information be used to test if the ointments are equally effectively?** **Wilcoxon Signed-Rank**

## Wilcoxon Signed-Rank test's ranking procedure:

1. Arrange  $d_i$  (the differences) in order of *absolute value*
2. Count the number of differences with the same absolute value
3. Rank  $d_i$  from 1 (observation with the lowest absolute value) to  $n$  (highest absolute value)
4. Group of several observations with the same absolute value
  - find the lowest rank in the range =  $1+R$  and the highest rank in the range =  $G + R$
  - $R$  = highest rank used prior to considering this group and  $G$  = the number of differences in the *range of ranks* for the group
  - Assign the *average rank* = (lowest rank in the range + highest rank in the range)/2 as the rank for each difference in the group



# The Wilcoxon Signed-Rank Test

$H_0: \Delta = 0$  vs.  $H_1: \Delta \neq 0$

- $\Delta$  = median score difference between ointment A and B
- If  $\Delta < 0$ , then ointment A is better
- If  $\Delta > 0$ , then ointment B is better

**Table 9.1** Difference in degree of redness between ointment A and ointment B arms after 10 minutes of exposure to sunlight

$ d_i $	Negative		Positive		Number of people with same absolute value	Range of ranks	Average rank
	$d_i$	$f_i$	$d_i$	$f_i$			
10	-10	0	10	0		—	—
9	-9	0	9	0	0	—	—
8	-8	1	8	0	1	40	40.0
7	-7	3	7	0	3	37–39	38.0
6	-6	2	6	0	2	35–36	35.5
5	-5	2	5	0	2	33–34	33.5
4	-4	1	4	0	1	32	32.0
3	-3	5	3	2	7	25–31	28.0
2	-2	4	2	6	10	15–24	19.5
1	-1	4	1	10	14	1–14	7.5
		22		18			
0	0	5					

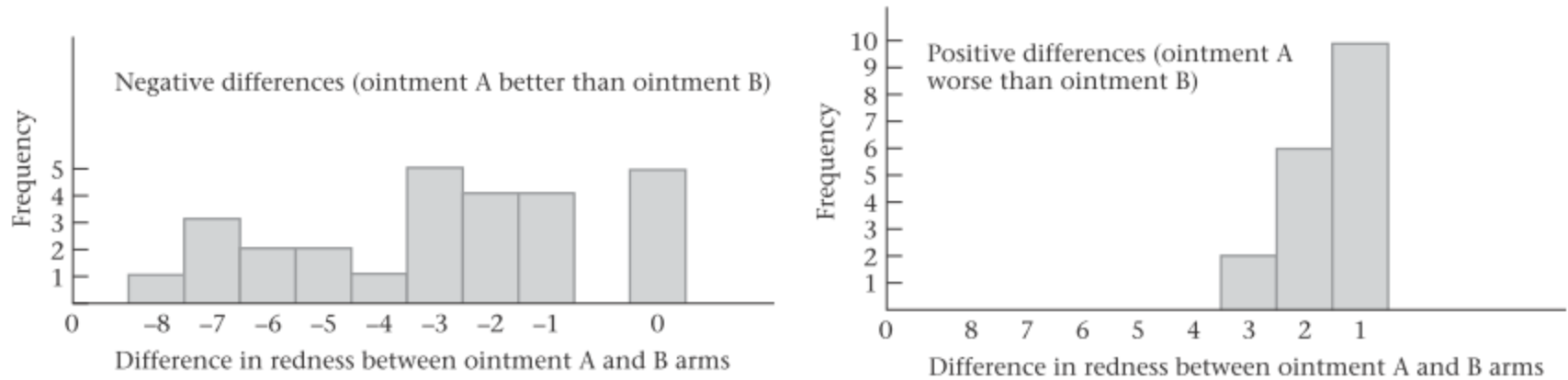


**Question: Compute the ranks for the skin-ointment data in Table 9.1.**

**Solution:**

- First collect the differences with the same absolute value
- Fourteen people have absolute value 1; this group has a rank range from 1 to 14 and an average rank of  $(1 + 14)/2 = 7.5$ .
- The group of 10 people with absolute value 2 has a rank range from  $(1 + 14)$  to  $(10 + 14) = 15$  to 24 and an average rank =  $(15 + 24)/2 = 19.5$ , . . . , and so on. T
- The column 'Average rank' in table 9.1 includes the ranks for each row of absolute value.

Bar graph of the differences in redness between the ointment A and ointment B arms for the data in Example 9.10



## Wilcoxon Signed-Rank test:

- nonparametric test that is analogous to the paired  $t$  test
- based on the ***rank***s of the observations rather than on their actual values

# Wilcoxon Signed-Rank Test (Normal Approximation Method for Two-Sided Level $\alpha$ Test)

\*number of nonzero di's  $\geq 16$ \* (normal approximation  $\rightarrow$  sampling distribution of  $R_1$ )

1. Rank the differences

2. Compute the rank sum  $R_1$  of the positive differences

3. If  $R_1 \neq [n(n+1)]/4$  and there are no ties (no groups of differences with the same absolute value), then

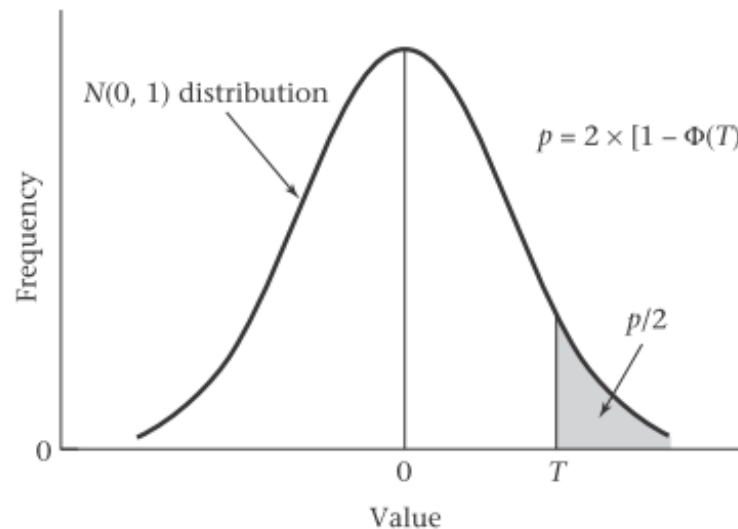
$$T = \left[ \left| R_1 - \frac{n(n+1)}{4} \right| - \frac{1}{2} \right] / \sqrt{n(n+1)(2n+1)/24}$$

4. If  $R_1 \neq [n(n+1)]/4$  and there are ties, where  $t_i$  refers to the number of differences with the same absolute value in the  $i$ th tied group and  $g$  is the number of tied groups, then

$$T = \left[ \left| R_1 - \frac{n(n+1)}{4} \right| - \frac{1}{2} \right] / \sqrt{n(n+1)(2n+1) / 24 - \sum_{i=1}^g (t_i^3 - t_i) / 48}$$

5. If  $T > z_{1-\alpha/2}$  then reject  $H_0$ . Otherwise, accept  $H_0$ .
6. The  $p$ -value for the test is given by  $p = 2 \times [1 - \Phi(T)]$
7. This test should be used only if the number of nonzero differences is  $\geq 16$  and if the difference scores have an underlying continuous symmetric distribution.

**Figure 9.5** Computation of the  $p$ -value for the Wilcoxon signed-rank test





# R Commands to perform the Wilcoxon Signed Rank Test

#For large sample method with one set of difference scores in variable x

```
>wilcox.test(x, y=NULL, alternative="two.sided", mu=0,  
paired=FALSE, exact=NULL, correct=TRUE, conf.int=FALSE)
```

#Two sets of paired scores in variables x and y

```
>wilcox.test(x, y, alternative="two.sided", mu=0, paired=TRUE,  
exact=NULL, correct=TRUE, conf.int=FALSE)
```

# Example on Wilcoxon Signed-rank test - Dermatology

- Question: Perform the Wilcoxon signed-rank test for the data below.

**Table 9.1** Difference in degree of redness between ointment A and ointment B arms after 10 minutes of exposure to sunlight

$ d_j $	Negative		Positive		Number of people with same absolute value	Range of ranks	Average rank
	$d_j$	$f_j$	$d_j$	$f_j$			
10	-10	0	10	0		—	—
9	-9	0	9	0	0	—	—
8	-8	1	8	0	1	40	40.0
7	-7	3	7	0	3	37-39	38.0
6	-6	2	6	0	2	35-36	35.5
5	-5	2	5	0	2	33-34	33.5
4	-4	1	4	0	1	32	32.0
3	-3	5	3	2	7	25-31	28.0
2	-2	4	2	6	10	15-24	19.5
1	-1	4	1	10	14	1-14	7.5
		22		18			
0	0	5					

# Example on Wilcoxon Signed-rank test - Dermatology

**Solution:** Because the number of nonzero differences ( $22 + 18 = 40$ )  $\geq 16$ , the normal approximation method can be used.

Compute the rank sum for the people with positive  $d_i$ —that is, where ointment B performs better than ointment A, as follows:

$$R_1 = 10(7.5) + 6(19.5) + 2(28.0) = 75 + 117 + 56 = 248$$

The expected rank sum is given by

$$E(R_1) = 40(41)/4 = 410$$

$$[n(n+1)/4]$$

The variance of the rank sum corrected for ties is given by

$$\begin{aligned} \text{Var}(R_1) &= 40(41)(81)/24 - [(14^3 - 14) + (10^3 - 10) + (7^3 - 7) + (1^3 - 1) + (2^3 - 2) + (2^3 - 2) + (3^3 - 3) + (1^3 - 1)]/48 = 5535 - \\ &= (2730 + 990 + 336 + 0 + 6 + 6 + 24 + 0)/48 = 5535 - 4092/48 = 5449.75 \end{aligned}$$

$$T = \left[ R_1 - \frac{n(n+1)}{4} \right] - \frac{1}{2} \bigg/ \sqrt{ \frac{n(n+1)(2n+1)}{24} - \sum_{i=1}^g (t_i^3 - t_i)/48 }$$



# Example on Wilcoxon Signed-rank test - Dermatology

Thus,  $sd(R_1) = \sqrt{5449.75} = 73.82$ .

Therefore, the test statistic  $T$  is given by

$$T = \left[ R_1 - \frac{n(n+1)}{4} - \frac{1}{2} \right] / \sqrt{n(n+1)(2n+1) / 24 - \sum_{i=1}^g (t_i^3 - t_i) / 48}$$

$$T = \frac{|248 - 410| - 0.5}{73.82} = \frac{161.5}{73.82} = 2.19$$

The  $p$ -value of the test is given by  $p = 2[1 - \Phi(2.19)] = 2 \times (1 - 0.9857) = 0.029$   $p = 2 \times [1 - \Phi(T)]$

Conclusion: there is a significant difference between ointments, with ointment A doing better than ointment B because the observed rank sum (248) is smaller than the expected rank sum (410).



TABLE 3 The normal distribution

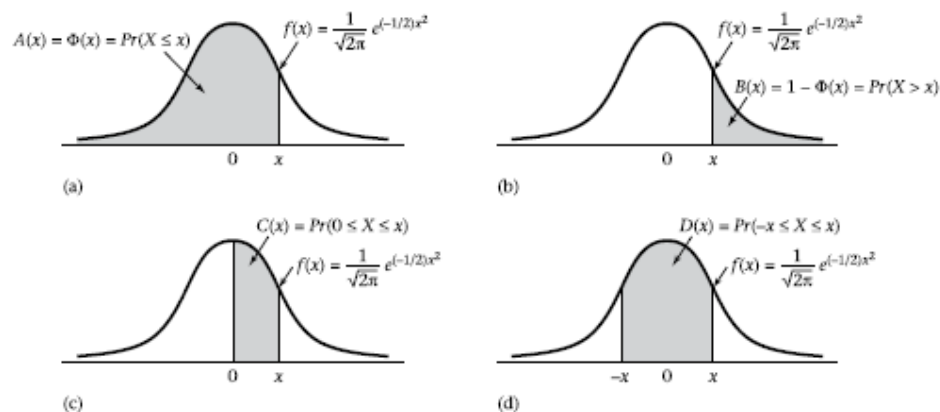


TABLE 3 The normal distribution (continued)

$x$	$A^a$	$B^b$	$C^c$	$D^d$
1.82	.9656	.0344	.4656	.9312
1.83	.9664	.0336	.4664	.9327
1.84	.9671	.0329	.4671	.9342
1.85	.9678	.0322	.4678	.9357
1.86	.9686	.0314	.4686	.9371
1.87	.9693	.0307	.4693	.9385
1.88	.9699	.0301	.4699	.9399
1.89	.9706	.0294	.4706	.9412
1.90	.9713	.0287	.4713	.9426
1.91	.9719	.0281	.4719	.9439
1.92	.9726	.0274	.4726	.9451
1.93	.9732	.0268	.4732	.9464
1.94	.9738	.0262	.4738	.9476
1.95	.9744	.0256	.4744	.9488
1.96	.9750	.0250	.4750	.9500
1.97	.9756	.0244	.4756	.9512
1.98	.9761	.0239	.4761	.9523
1.99	.9767	.0233	.4767	.9534
2.00	.9772	.0228	.4772	.9545
2.01	.9778	.0222	.4778	.9556
2.02	.9783	.0217	.4783	.9566
2.03	.9788	.0212	.4788	.9576
2.04	.9793	.0207	.4793	.9586
2.05	.9798	.0202	.4798	.9596
2.06	.9803	.0197	.4803	.9606
2.07	.9808	.0192	.4808	.9615
2.08	.9812	.0188	.4812	.9625
2.09	.9817	.0183	.4817	.9634
2.10	.9821	.0179	.4821	.9643
2.11	.9826	.0174	.4826	.9651
2.12	.9830	.0170	.4830	.9660
2.13	.9834	.0166	.4834	.9668
2.14	.9838	.0162	.4838	.9676
2.15	.9842	.0158	.4842	.9684
2.16	.9846	.0154	.4846	.9692
2.17	.9850	.0150	.4850	.9700
2.18	.9854	.0146	.4854	.9707
2.19	.9857	.0143	.4857	.9715
2.20	.9861	.0139	.4861	.9722
2.21	.9864	.0136	.4864	.9729
2.22	.9868	.0132	.4868	.9736
2.23	.9871	.0129	.4871	.9743
2.24	.9875	.0125	.4875	.9749
2.25	.9878	.0122	.4878	.9756
2.26	.9881	.0119	.4881	.9762
2.27	.9884	.0116	.4884	.9768
2.28	.9887	.0113	.4887	.9774
2.29	.9890	.0110	.4890	.9780
2.30	.9893	.0107	.4893	.9786
2.31	.9896	.0104	.4896	.9791
2.32	.9898	.0102	.4898	.9797
2.33	.9901	.0099	.4901	.9802
2.34	.9904	.0096	.4904	.9807
2.35	.9906	.0094	.4906	.9812
2.36	.9909	.0091	.4909	.9817
2.37	.9911	.0089	.4911	.9822
2.38	.9913	.0087	.4913	.9827

# Scenario on Ophthalmology

- Different genetic types of the disease retinitis pigmentosa (RP) are thought to have different rates of progression
  - Dominant form of the disease: slowest progression
  - Recessive form: next slowest
  - Gender-linked form: fastest progression
- Test the hypothesis by comparing the visual acuity of people who have different genetic types of RP
- Scenario: 25 people aged 10-19 with dominant disease and 30 people with gender-linked disease
- The best-corrected visual acuities (i.e. with appropriate glasses) in the better eye of these people are shown in the next slide.
- **Q: how can these data be used to test if the distribution of visual acuity is different in the two groups?**

# Wilcoxon Rank Sum Test

- Nonparametric analog to the  $t$  test for two independent samples
- Hypothesis  $H_0: F_D = F_{SL}$  vs.  $H_1: F_D(x) = F_{SL}(x - \Delta)$  where  $\Delta \neq 0$
- $F_D$  = cdf of visual acuity for dominant grp
- $F_{SL}$  = cdf of visual acuity for sex-linked grp
- $\Delta$  = location shift of cdf for the sex linked grp relative to dominant group

**Table 9.3** Comparison of visual acuity in people ages 10–19 with dominant and sex-linked RP

Visual acuity	Dominant	Sex-linked	Combined sample	Range of ranks	Average rank
20–20	5	1	6	1–6	3.5
20–25	9	5	14	7–20	13.5
20–30	6	4	10	21–30	25.5
20–40	3	4	7	31–37	34.0
20–50	2	8	10	38–47	42.5
20–60	0	5	5	48–52	50.0
20–70	0	2	2	53–54	53.5
20–80	<u>0</u>	<u>1</u>	<u>1</u>	55	55.0
	25	30	55		



# Ranking Procedure for the Wilcoxon Rank-Sum Test

1. Combine data from the two groups, order the values smallest to largest (or e.g. visual acuity from best [20-20] to worst [20-80])
2. Assign ranks to the individual values  
e.g. lowest rank: with the best visual acuity (20-20)  
highest rank: with worst visual acuity (20-80)  
or vice versa
3. Group of observations has the same value
  - compute the range of ranks for the group (similar to signed-rank test)
  - assign the average rank for each observation in the group

# Wilcoxon Rank-Sum Test (Normal Approximation Method for Two-Sided Level $\alpha$ Test)

1. Rank the observations
2. Compute the rank sum  $R_1$  in the first sample (the choice of sample is arbitrary)
3. If (i)  $R_1 \neq n_1(n_1+n_2+1)/2$  and there are no ties, then compute

$$T = \left[ R_1 - \frac{n_1(n_1+n_2+1)}{2} \right] - \frac{1}{2} \bigg/ \sqrt{\left( \frac{n_1 n_2}{12} \right) (n_1 + n_2 + 1)}$$

- 
- 
- (ii)  $R_1 \neq n_1(n_1+n_2+1)/2$  and there are ties, then compute

$$T = \left[ R_1 - \frac{n_1(n_1+n_2+1)}{2} \right] - \frac{1}{2} \bigg/ \sqrt{\left( \frac{n_1 n_2}{12} \right) \left[ n_1 + n_2 + 1 - \frac{\sum_{i=1}^g t_i(t_i^2 - 1)}{(n_1 + n_2)(n_1 + n_2 - 1)} \right]}$$

where  $t_i$  refers to the number of observations with the same value in the  $i$ th tied group and  $g$  is the number of tied groups

(iii) If  $R1 = n1(n1+n2+1)/2$ , then  $T = 0$

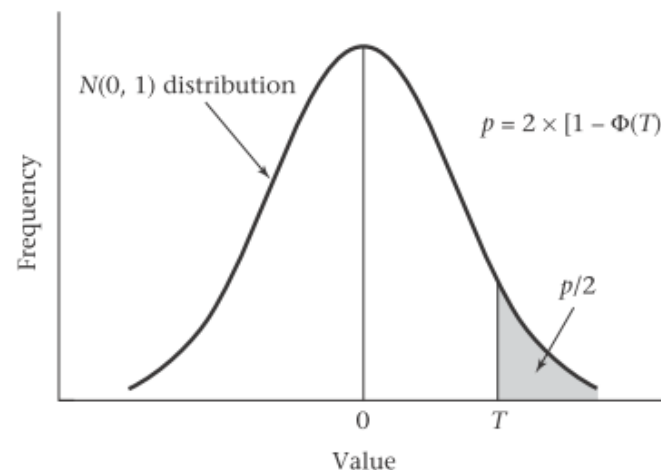
4. If  $T > z_{1-\alpha/2}$  then reject  $H_0$ . Otherwise, accept  $H_0$ .

5. Compute the exact p-value by  $p = 2 \times [1 - \Phi(T)]$

5. Conditions:

- both  $n1$  and  $n2$  are  $\geq 10$
- there is an underlying continuous distribution

Figure 9.6 Computation of the p-value for the Wilcoxon rank-sum test





# Example on Wilcoxon Rank-sum Test - Ophthalmology

- Question: Compute the ranks for the visual-acuity data in Table 9.3.

**Table 9.3** Comparison of visual acuity in people ages 10–19 with dominant and sex-linked RP

Visual acuity	Dominant	Sex-linked	Combined sample	Range of ranks	Average rank
20–20	5	1	6	1–6	3.5
20–25	9	5	14	7–20	13.5
20–30	6	4	10	21–30	25.5
20–40	3	4	7	31–37	34.0
20–50	2	8	10	38–47	42.5
20–60	0	5	5	48–52	50.0
20–70	0	2	2	53–54	53.5
20–80	0	1	1	55	55.0
	25	30	55		

# Example on Wilcoxon Rank-sum Test - Ophthalmology

## Solution:

- First collect all people with the same visual acuity over the two groups, as shown in Table 9.3.
- There are 6 people with visual acuity 20–20 who have a rank range of 1–6 and are assigned an average rank of  $(1 + 6)/2 = 3.5$ .
- There are 14 people for the two groups combined with visual acuity 20–25. The rank range for this group is from  $(1 + 6)$  to  $(14 + 6) = 7$  to 20.
- All people in this group are assigned the average rank  $= (7 + 20)/2 = 13.5$ , and similarly for the other groups. The column 'Average rank' in table 9.3 depicts the ranks as required.

# Example on Wilcoxon Rank-sum Test - Ophthalmology

**Question:** Perform the Wilcoxon rank-sum test for the data in Table 9.3.

**Solution:**

Because the minimum sample size in the two samples is  $25 \geq 10$ , the normal approximation can be used.

The rank sum in the dominant group is given by

$$R_1 = 5(3.5) + 9(13.5) + 6(25.5) + 3(34) + 2(42.5) \\ = 17.5 + 121.5 + 153 + 102 + 85 = 479$$

Furthermore,  $E(R_1) = 25(56)/2 = \frac{1400}{2} = 700$   $n_1(n_1+n_2+1)/2$

and  $Var(R_1)$  corrected for ties is given by

$$[25(30)/12]\{56 - [6(6^2 - 1) + 14(14^2 - 1) + 10(10^2 - 1) + 7(7^2 - 1) \\ + 10(10^2 - 1) + 5(5^2 - 1) + 2(2^2 - 1) + 1(1^2 - 1)]/[55(54)]\} \\ = 62.5(56 - 5382 / 2970) = 3386.74$$

$$T = \left[ R_1 - \frac{n_1(n_1+n_2+1)}{2} \right] \sqrt{\frac{1}{n_1 n_2}} \left[ \frac{\sum_{i=1}^n t_i(t_i^2 - 1)}{(n_1+n_2)(n_1+n_2-1)} \right]$$



# Wilcoxon Rank-sum Test - Ophthalmology

Therefore, the test statistic  $T$  is given by

$$T = \left[ R_1 - \frac{n_1(n_1 + n_2 + 1)}{2} \right] - \frac{1}{2} \bigg/ \sqrt{\left( \frac{n_1 n_2}{12} \right) \left[ n_1 + n_2 + 1 - \frac{\sum_{i=1}^g t_i(t_i^2 - 1)}{(n_1 + n_2)(n_1 + n_2 - 1)} \right]}$$

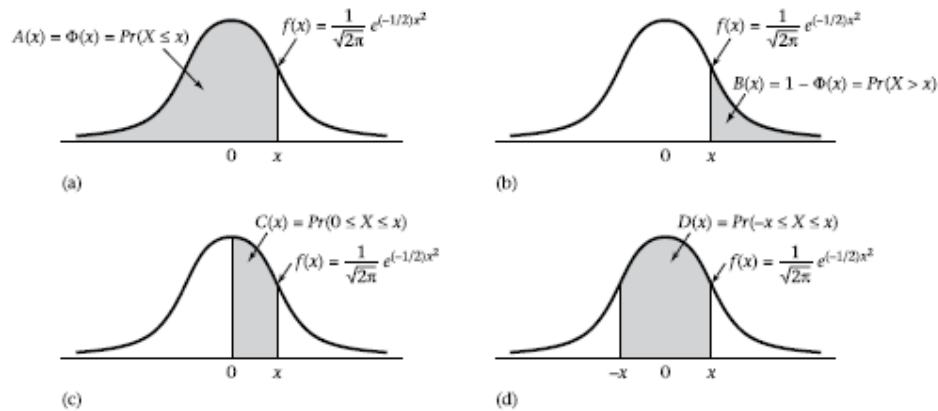
$$T = \frac{|479 - 700| - 0.5}{\sqrt{3386.74}} = \frac{220.5}{58.2} = 3.79$$

which follows an  $N(0, 1)$  distribution under  $H_0$ .

The  $p$ -value of the test is  $p = 2[1 - \Phi(3.79)] < 0.001$

Conclusion: the visual acuities of the two groups are significantly different. Because the observed rank sum in the dominant group (479) is lower than the expected rank sum (700), the dominant group has better visual acuity than the sex-linked group.

TABLE 3 The normal distribution



$x$	$A$	$B$	$C$	$D$
3.49	.9998	.0002	.4998	.9995
3.50	.9998	.0002	.4998	.9995
3.51	.9998	.0002	.4998	.9996
3.52	.9998	.0002	.4998	.9996
3.53	.9998	.0002	.4998	.9996
3.54	.9998	.0002	.4998	.9996
3.55	.9998	.0002	.4998	.9996
3.56	.9998	.0002	.4998	.9996
3.57	.9998	.0002	.4998	.9996
3.58	.9998	.0002	.4998	.9997
3.59	.9998	.0002	.4998	.9997
3.60	.9998	.0002	.4998	.9997
3.61	.9998	.0002	.4998	.9997
3.62	.9999	.0001	.4999	.9997
3.63	.9999	.0001	.4999	.9997
3.64	.9999	.0001	.4999	.9997
3.65	.9999	.0001	.4999	.9997
3.66	.9999	.0001	.4999	.9997
3.67	.9999	.0001	.4999	.9998
3.68	.9999	.0001	.4999	.9998
3.69	.9999	.0001	.4999	.9998
3.70	.9999	.0001	.4999	.9998
3.71	.9999	.0001	.4999	.9998
3.72	.9999	.0001	.4999	.9998
3.73	.9999	.0001	.4999	.9998
3.74	.9999	.0001	.4999	.9998
3.75	.9999	.0001	.4999	.9998
3.76	.9999	.0001	.4999	.9998
3.77	.9999	.0001	.4999	.9998
3.78	.9999	.0001	.4999	.9998
3.79	.9999	.0001	.4999	.9998
3.80	.9999	.0001	.4999	.9999
3.81	.9999	.0001	.4999	.9999
3.82	.9999	.0001	.4999	.9999
3.83	.9999	.0001	.4999	.9999
3.84	.9999	.0001	.4999	.9999
3.85	.9999	.0001	.4999	.9999
3.86	.9999	.0001	.4999	.9999
3.87	.9999	.0001	.4999	.9999
3.88	.9999	.0001	.4999	.9999
3.89	.9999	.0001	.4999	.9999
3.90	1.0000	.0000	.5000	.9999
3.91	1.0000	.0000	.5000	.9999
3.92	1.0000	.0000	.5000	.9999
3.93	1.0000	.0000	.5000	.9999
3.94	1.0000	.0000	.5000	.9999
3.95	1.0000	.0000	.5000	.9999
3.96	1.0000	.0000	.5000	.9999
3.97	1.0000	.0000	.5000	.9999
3.98	1.0000	.0000	.5000	.9999
3.99	1.0000	.0000	.5000	.9999

TABLE 11 Two-tailed critical values for the Wilcoxon rank-sum test

$n_2^b$	$\alpha = .10$ $n_1^a$							$\alpha = .05$ $n_1$						
	4	5	6	7	8	9		4	5	6	7	8	9	
	$T_1^L$ $T_1^U$	$T_1^L$ $T_1^U$	$T_1^L$ $T_1^U$	$T_1^L$ $T_1^U$	$T_1^L$ $T_1^U$	$T_1^L$ $T_1^U$		$T_1^L$ $T_1^U$	$T_1^L$ $T_1^U$	$T_1^L$ $T_1^U$	$T_1^L$ $T_1^U$	$T_1^L$ $T_1^U$	$T_1^L$ $T_1^U$	
4	11-25	17-33	24-42	32-52	41-63	51-75		10-26	16-34	23-43	31-53	40-64	49-77	
5	12-28	19-36	26-46	34-57	44-68	54-81		11-29	17-38	24-48	33-58	42-70	52-83	
6	13-31	20-40	28-50	36-62	46-74	57-87		12-32	18-42	26-52	34-64	44-76	55-89	
7	14-34	21-44	29-55	39-66	49-79	60-93		13-35	20-45	27-57	36-69	46-82	57-96	
8	15-37	23-47	31-59	41-71	51-85	63-99		14-38	21-49	29-61	38-74	49-87	60-102	
9	16-40	24-51	33-63	43-76	54-90	66-105		14-42	22-53	31-65	40-79	51-93	62-109	
10	17-43	26-54	35-67	45-81	56-96	69-111		15-45	23-57	32-70	42-84	53-99	65-115	
11	18-46	27-58	37-71	47-86	59-101	72-117		16-48	24-61	34-74	44-89	55-105	68-121	
12	19-49	28-62	38-76	49-91	62-106	75-123		17-51	26-64	35-79	46-94	58-110	71-127	
13	20-52	30-65	40-80	52-95	64-112	78-129		18-54	27-68	37-83	48-99	60-116	73-134	
14	21-55	31-69	42-84	54-100	67-117	81-135		19-57	28-72	38-88	50-104	62-122	76-140	
15	22-58	33-72	44-88	56-105	69-123	84-141		20-60	29-76	40-92	52-109	65-127	79-146	
16	24-60	34-76	46-92	58-110	72-128	87-147		21-63	30-80	42-96	54-114	67-133	82-152	
17	25-63	35-80	47-97	61-114	75-133	90-153		21-67	32-83	43-101	56-119	70-138	84-159	
18	26-66	37-83	49-101	63-119	77-139	93-159		22-70	33-87	45-105	58-124	72-144	87-165	
19	27-69	38-87	51-105	65-124	80-144	96-165		23-73	34-91	46-110	60-129	74-150	90-171	
20	28-72	40-90	53-109	67-129	83-149	99-171		24-76	35-95	48-114	62-134	77-155	93-177	
21	29-75	41-94	55-113	69-134	85-155	102-177		25-79	37-98	50-118	64-139	79-161	95-184	
22	30-78	43-97	57-117	72-138	88-160	105-183		26-82	38-102	51-123	66-144	81-167	98-190	
23	31-81	44-101	58-122	74-143	90-166	108-189		27-85	39-106	53-127	68-149	84-172	101-196	
24	32-84	45-105	60-126	76-148	93-171	111-195		27-89	40-110	54-132	70-154	86-178	104-202	
25	33-87	47-108	62-130	78-153	96-176	114-201		28-92	42-113	56-136	72-159	89-183	107-208	
26	34-90	48-112	64-134	81-157	98-182	117-207		29-95	43-117	58-140	74-164	91-189	109-215	
27	35-93	50-115	66-138	83-162	101-187	120-213		30-98	44-121	59-145	76-169	93-195	112-221	
28	36-96	51-119	67-143	85-167	103-193	123-219		31-101	45-125	61-149	78-174	96-200	115-227	
29	37-99	53-122	69-147	87-172	106-198	126-225		32-104	47-128	63-153	80-179	98-206	118-233	
30	38-102	54-126	71-151	89-177	109-203	129-231		33-107	48-132	64-158	82-184	101-211	121-239	
31	39-105	55-130	73-155	92-181	111-209	132-237		34-110	49-136	66-162	84-189	103-217	123-246	
32	40-108	57-133	75-159	94-186	114-214	135-243		34-114	50-140	67-167	86-194	106-222	126-252	
33	41-111	58-137	77-163	96-191	117-219	138-249		35-117	52-143	69-171	88-199	108-228	129-258	
34	42-114	60-140	78-168	98-196	119-225	141-255		36-120	53-147	71-175	90-204	110-234	132-264	
35	43-117	61-144	80-172	100-201	122-230	144-261		37-123	54-151	72-180	92-209	113-239	135-270	
36	44-120	62-148	82-176	102-206	124-236	148-266		38-126	55-155	74-184	94-214	115-245	137-277	
37	45-123	64-151	84-180	105-210	127-241	151-272		39-129	57-158	76-188	96-219	117-251	140-283	
38	46-126	65-155	85-185	107-215	130-246	154-278		40-132	58-162	77-193	98-224	120-256	143-289	
39	47-129	67-158	87-189	109-220	132-252	157-284		41-135	59-166	79-197	100-229	122-262	146-295	
40	48-132	68-162	89-193	111-225	135-257	160-290		41-139	60-170	80-202	102-234	125-267	149-301	
41	49-135	69-166	91-197	114-229	138-262	163-296		42-142	61-174	82-206	104-239	127-273	151-308	
42	50-138	71-169	93-201	116-234	140-268	166-302		43-145	63-177	84-210	106-244	129-279	154-314	
43	51-141	72-173	95-205	118-239	143-273	169-308		44-148	64-181	85-215	108-249	132-284	157-320	
44	52-144	74-176	96-210	120-244	146-278	172-314		45-151	65-185	87-219	110-254	134-290	160-326	
45	53-147	75-180	98-214	123-248	148-284	175-320		46-154	66-189	88-224	112-259	137-295	163-332	
46	55-149	77-183	100-218	125-253	151-289	178-326		47-157	68-192	90-228	114-264	139-301	165-339	
47	56-152	78-187	102-222	127-258	154-294	181-332		48-160	69-196	92-232	116-269	141-307	168-345	
48	57-155	79-191	104-226	129-263	156-300	184-338		48-164	70-200	93-237	118-274	144-312	171-351	
49	58-158	81-194	106-230	132-267	159-305	187-344		49-167	71-204	95-241	120-279	146-318	174-357	
50	59-161	82-198	107-235	134-272	162-310	190-350		50-170	73-207	97-245	122-284	149-323	177-363	

<sup>a</sup> $n_1$  = minimum of the two sample sizes.<sup>b</sup> $n_2$  = maximum of the two sample sizes.<sup>c</sup> $T_1^L$  = lower critical value for the rank sum in the first sample.<sup>d</sup> $T_1^U$  = upper critical value for the rank sum in the first sample.



# **R commands to perform the Wilcoxon Rank Sum Test**

#scores for the two groups

#x=values for one group and y=values for the other group

```
>wilcox.test(x,y,alternative="two.sided", mu=0, paired=FALSE, exact=NULL,  
Correct=TRUE, conf.int=FALSE)
```

# Summary

1. Advantage of nonparametric methods: assumption of normality can be relaxed when such assumptions are unreasonable
2. Disadvantage of nonparameteric procedures: some power is lost relative to using a parametric procedure (such as a t test) if the data truly follow a normal distribution (or central-limit theorem is applicable)
3. The sign test and the signed-rank test are nonparametric analogs to the **paired t test**.
  - Sign test: to determine whether one member of a matched pair has a higher or lower score than the other member of the pair
  - For the signed-rank test: the magnitude of the absolute value of the difference score and its sign is used in performing the significance test.
4. The Wilcoxon rank-sum test is an analog to the **two-sample t test for independent samples** in which the actual values are replaced by rank scores.

# Parametric tests and analogous nonparametric procedures

Analysis Type	Example	Parametric Procedure	Nonparametric Procedure
Compare means between two distinct/independent groups	Is the mean systolic blood pressure (at baseline) for patients assigned to placebo different from the mean for patients assigned to the treatment group?	Two-sample t-test	Wilcoxon rank-sum test
Compare two quantitative measurements taken from the same individual	Was there a significant change in systolic blood pressure between baseline and the six-month follow-up measurement in the treatment group?	Paired t-test	Wilcoxon signed-rank test
Estimate the degree of association between two quantitative variables	Is systolic blood pressure associated with the patient's age?	Pearson coefficient of correlation	Spearman's rank correlation