Unit 6

Non-parametric Tests

Tony CHAN, Ph.D.

1

6.1 Introduction

In Units 3 and 5, we introduced some tests of hypotheses. These tests are generally concerned with such quantities as the means of a population.

- A ***mean, variance, or proportion*** is referred to as a ***parameter*** of a population. Such tests are called ***parametric tests*** of hypotheses.
- The commonly underlying ***assumption*** in testing a parameter from a ***continuous population*** is that the ***population is normally distributed***.
- Any time for a ***test statistic***, we ***assume a normal population distribution***.
- When we test more than 2 means using the ***ANOVA procedure***, we also ***assume normality of the populations.***
- Non-parametric methods are used for situations that violate the ***assumptions of the parametric procedures***.

2

6.2 Parametric Test Procedures

- Involve population parameters
  - e.g. population mean, μ
- Require interval or ratio data
  - integers or fractions
  - e.g. height in meter
- Have stringent assumptions
  - e.g. normal distribution, independence
- Examples of parametric tests
  - z-test, t-test, F-test, χ^2-test

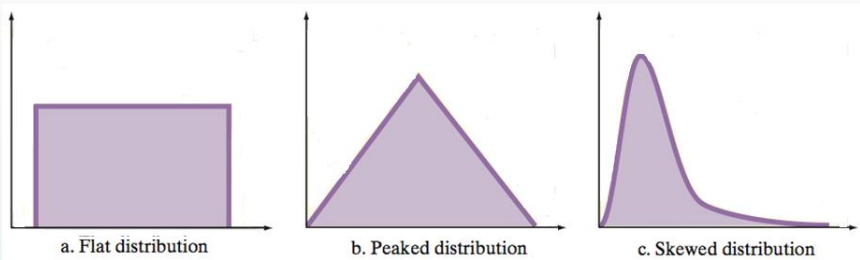
3

6.2.1 Situations where Parametric Tests Fail

- The data collected from an ***extremely skewed population*** (i.e., ***asymmetric distribution***).
- Data are ***qualitative***, say, ***nominal*** or ***ordinal***.
- Data contain ***extreme values***.

4

6.3 Non-normal Distributions



6.3.1 How to Deal with Non-normality?

Large sample size

- We may collect ***large samples***. The ***central limit theorem assures*** that the ***distribution*** of the ***sample estimators*** is ***approximately normally distributed*** regardless of the population distribution.

Small to moderate sample size

- We may use a ***non-parametric statistical procedure*** that deals with data in which the normality assumption does not hold.

6.4 Distribution-free Tests

- Distribution-free tests are statistical tests that do ***not rely on any underlying assumptions*** about the ***probability distribution of the sampled population***.
- The branch of inferential statistics devoted to distribution-free tests is called ***Non-parametric Statistics.***
- Nonparametric statistics or tests based on the ***ranks of measurements*** are called ***rank statistics*** or ***rank tests***.

7

6.5 Ranks of Observations

A common method for ***practically all non-parametric techniques*** is the use of ***ranks***. Given a set of data, we obtain a set of ranks by ***replacing each data value by its relative position***. Consider the following 8 observations:

10.8, 6.4, 11.7, 5.3, 9.5, 2.5, 15.1, 10.4

Arrange in ***ascending order of magnitudes***, we get

Position:	1	2	3	4	5	6	7	8
Value:	2.5	5.3	6.4	9.5	10.4	10.8	11.7	15.1

We say that the ***rank*** of the value 2.5 is ***1***, the rank of the value 5.3 is 2, and so forth. ***Replacing each value by its rank*** and ***maintaining the original order*** produces

6, 3, 7, 2, 4, 1, 8, 5

8

6.5 Ranks of Observations (Cont'd)

For ties or tied observations, use the ***average*** of its ***consecutive numbers*** for the rank of each of the tied data.

Example 6.1

Two small classes take a test. Rank their letter grades.

Class 1	A+	C-	B	D	A-	C+
Ranks	12	3	6.5	1.5	9.5	5
Class 2	A-	D	B	A	B+	C
Ranks	9.5	1.5	6.5	11	8	4

Grade:	D	D	C-	C	C+	B	B	B+	A-	A-	A	A+
Rank:	1.5	1.5	3	4	5	6.5	6.5	8	9.5	9.5	11	12

6.6 Ranked Data in Practice

- Ranked data for several different brands of soft drinks in a consumer taste test.
- Respondent's response selected from 'very unsatisfied', 'slightly unsatisfied' 'satisfied' or 'extremely happy' about a hotel service in a satisfaction survey.
- Time-ordered data such as the winners of a race.
- Intensity-ordered data such as the stages of cancer.
- Popularity rankings for books or music tracks.

6.7 Wilcoxon Signed Rank Test for Single Sample

Conditions:

Data must be ***randomly selected*** and have a ***symmetric distribution***.

- The symmetric assumption does not guarantee normality, simply that there seems to be roughly the same number of values above and below the median.
- It is a test for the central location (mean, median or mode) of a single population.
- This test takes both ***direction*** and ***magnitude*** into account.
- A sample of size n , $(x_1, x_2, \ldots, x_n)$ from an ***unknown*** distribution
- The test applies to the case of ***symmetric*** and ***continuous*** distributions
- Under this assumption, mean = median

Null Hypothesis	Alternative Hypothesis	Type of Test
H_0: $\mu = \mu_0$	H_1: $\mu > \mu_0$	Right-tailed
H_0: $\mu = \mu_0$	H_1: $\mu < \mu_0$	Left-tailed
H_0: $\mu = \mu_0$	H_1: $\mu \neq \mu_0$	Two-tailed

11

6.7.1 Procedures of Wilcoxon Signed Rank Test for Single Sample

- For ***Wilcoxon signed rank test***, we assign ranks to the absolute values of $(x_1 - \mu_0, x_2 - \mu_0, \ldots, x_n - \mu_0)$, where μ_0 is a certain hypothesized or assumed value to be tested in H_0: $\mu = \mu_0$.
- A rank of ***1*** to the value of $x_i - \mu_0$ which is ***smallest in absolute value.***
- A rank of ***n*** to the value of $x_i - \mu_0$ which is ***largest in absolute value.***
- W^+ = the ***sum of the ranks*** associated with ***positive values*** of $x_i - \mu_0$.
- W^- = the ***sum of the ranks*** associated with ***negative values*** of $x_i - \mu_0$.

12

6.7.1 Procedures of Wilcoxon Signed Rank Test for Single Sample (Cont'd)

- $W^+ + W^- = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$.
- Left-tailed test $(H_0: \mu = \mu_0 \leftrightarrow H_1: \mu - \mu_0 < 0)$
  Reject H_0 in favor of $H_1: \mu < \mu_0$ only if W^+ is ***SMALL*** and W^- is large.
- Right-tailed test $(H_0: \mu = \mu_0 \leftrightarrow H_1: \mu - \mu_0 > 0)$
  Reject H_0 in favor of $H_1: \mu > \mu_0$ only if W^- is ***SMALL*** and W^+ is large.
- Two-tailed test ($H_0: \mu = \mu_0 \leftrightarrow H_1: \mu \neq \mu_0 \Leftrightarrow (\mu < \mu_0$ **or** $\mu > \mu_0)$)
  Reject H_0 in favor of H_1 if ***either W^+ or W^-***, and hence,
  $W = min(W^+, W^-)$ is sufficiently small.
- Decision Rule:
  Reject H_0 if $\boxed{w^+, w^- \text{ or } w = min(w^+, w^-) \leq \text{tabled value}}$

13

6.7.1 Procedures of Wilcoxon Signed Rank Test for Single Sample (Cont'd)

H_0	H_1	Test Statistic to be Computed
$\mu = \mu_0$	$\mu < \mu_0$	W^+
	$\mu > \mu_0$	W^-
	$\mu \neq \mu_0$	$W = min(W^+, W^-)$

Or more simply, the test statistic for the above 3 tests (left-tailed, right-tailed, and 2-tailed) is

$$W = min(W^+, W^-)$$

14

6.7.2 Wilcoxon Signed Rank Test Statistic for Single Sample

- Ranks the observations (obs) according to magnitude (***absolute value***) from the smallest obs (rank = 1) to the largest (rank = *n*) obs.
- Test statistic, $W = min(W^+, W^-)$
- Specify a significance level α, say, 5% or 1%.
- Find the critical value corresponding to α. Refer to Table A.16 on the next slide.

15

Table A.16 Critical Values for the Signed-Rank Test

n	One-Sided $\alpha = 0.01$ Two-Sided $\alpha = 0.02$	One-Sided $\alpha = 0.025$ Two-Sided $\alpha = 0.05$	One-Sided $\alpha = 0.05$ Two-Sided $\alpha = 0.1$
5			1
6		1	2
7	0	2	4
8	2	4	6
9	3	6	8
10	5	8	11
11	7	11	14
12	10	14	17
13	13	17	21
14	16	21	26
15	20	25	30
16	24	30	36
17	28	35	41
18	33	40	47
19	38	46	54
20	43	52	60
21	49	59	68
22	56	66	75
23	62	73	83
24	69	81	92
25	77	90	101
26	85	98	110
27	93	107	120
28	102	117	130
29	111	127	141
30	120	137	152

Reproduced from F. Wilcoxon and R. A. Wilcox, *Some Rapid Approximate Statistical Procedures, American Cyanamid Company*, Pearl River, N.Y., 1964, by permission of the American Cyanamid Company.

16

Example 6.2

The following data represent the number of hours that a rechargeable hedge trimmer operates before a recharge is required:

1.5, 2.2, 0.9, 1.3, 2.0, 1.6, 1.5, 2.0, 1.2, 1.7

Use the Wilcoxon signed rank test to test the hypothesis, at the 5% level of significance, that this particular trimmer operates with a median of 1.8 hours before requiring a recharge.

17

Solution

1. Hypotheses to be tested
   H_0: $\tilde{\mu} = \widetilde{\mu_0} = 1.8$ $\leftrightarrow$ H_1: $\tilde{\mu} \neq 1.8$, where $\tilde{\mu}$ = the mean waiting time.

2. Critical value
   $n = 10$, $\alpha = 5\%$. From Table A.16 (Slide 16), we have the critical value $w \leq 8$.

3. Value of test statistic, w

$d_i = x_i - \mu_0$	-0.3	0.4	-0.9	-0.5	0.2	-0.2	-0.3	0.2	-0.6	-0.1
Signed Rank	-5.5	+7	-10	-8	+3	-3	-5.5	+3	-9	-1

$w^+ = 7 + 3 + 3 = 13,$
$w^- = 5.5 + 10 + 8 + 3 + 5.5 + 9 + 1 = 42$
$w = min(w^+, w^-) = min(13,42) = 13$

4. Decision
   $\because$ w = 13 does not lie within the critical region $w \leq 8$,
   $\therefore$ do not reject H_0 at the 5% level.

5. Conclusion
   The median operating time is not significantly different from 1.8 hours.

18

Exercise 6.1

1. The following data represent the time, in minutes, that a patient has to wait during 12 visits to a doctor's office before being seen by the doctor:

17, 15, 20, 20, 32, 28, 12, 26, 25, 25, 35, 24

Use the Wilcoxon signed rank test to test, at the 0.05 level of significance, the doctor's claim that the median waiting time for her patients is not more than 20 minutes.

Solution

i. Hypotheses to be tested

H_0: $\tilde{\mu} = \widetilde{\mu_0} = 20 \quad \leftrightarrow \quad H_1$: $\tilde{\mu} > 20$, where $\tilde{\mu}$ = the median waiting time.

ii. Critical value

$n = 10$ ***(discard the two "20" as they are the same as $\widetilde{\mu_0} = 20$ to be tested.)***, $\alpha = 5\%$. From Table A.16 (Slide 16), we have the critical value $w \leq 11$.

iii. Value of test statistic, w

$d_i = x_i - \mu_0$	-3	-5	12	8	-8	6	5	5	15	4
Signed Rank	-1	-4	+9	+7.5	-7.5	+6	+4	+4	+10	+2

$w^+ = 9 + 7.5 + 6 + 4 + 4 + 10 + 2 = 42.5,$

$w^- = 1 + 4 + 7.5 = 12.5$

$w = min(w^+, w^-) = min(42.5, 12.5) = 12.5$

iv. Decision

$\because$ w = 12.5 does not lie within the critical region $w \leq 11$,

$\therefore$ do not reject H_0 at the 5% level.

v. Conclusion

The median waiting time for her patients is not more than 20 minutes.

6.8 Wilcoxon Signed Rank Test for Paired Samples

- Two ***continuous*** and ***symmetric*** populations with $\mu_1 = \mu_2$ for the paired sample case are sampled to test H_0.
- The ***differences***, $d_i's$, of the paired observations are ***ranked without regard to sign*** and proceed as with the single sample case. The test procedures are summarized below:

	H_0	H_1	Test Statistic to be Computed
Single Sample	$\mu = \mu_0$	$\mu < \mu_0$	W^+
		$\mu > \mu_0$	W^-
		$\mu \neq \mu_0$	$W = min(W^+, W^-)$
Paired Samples	$\mu_1 = \mu_2$	$\mu_1 < \mu_2$	W^+
		$\mu_1 > \mu_2$	W^-
		$\mu_1 \neq \mu_2$	$W = min(W^+, W^-)$

Or more simply, the test statistic for all 3 different tests is

$$W = min(W^+, W^-)$$

Example 6.3

It is claimed that a new diet will reduce a person's weight by 4.5 kg on average in a two-week period. The weights of 10 women who followed this diet were recorded before and after the period, yielding the following data:

Women	Weight Before	Weight After
1	58.5	60.0
2	60.3	54.9
3	61.7	58.1
4	69.0	62.1
5	64.0	58.5
6	62.6	59.9
7	56.7	54.4
8	63.6	60.2
9	68.2	62.3
10	59.4	58.7

Use the Wilcoxon signed rank test, at the 5% level of significance, to test the following null hypothesis that joining the diet has no effect on median weight versus the alternative hypothesis that median weight decreases after joining the diet.

Solution

i. Hypotheses to be tested

H_0: $\widetilde{\mu_1} = \widetilde{\mu_2}$ $\leftrightarrow$ H_1: $\widetilde{\mu_1} > \widetilde{\mu_2}$, where $\widetilde{\mu_1}$ and $\widetilde{\mu_2}$ are respectively the median weights before and after taking the diet.

ii. Critical value

$n = 10$, $\alpha = 5\%$. From Table A.16 (Slide 16), we have the critical value $w \leq 11$.

iii. Value of test statistic, w

$d_i = x_{1i} - x_{2i}$	-1.5	5.4	3.6	6.9	5.5	2.7	2.3	3.4	5.9	0.7
Signed Rank	-2	+7	+6	+10	+8	+4	+3	+5	+9	+1

$w^+ = 7 + 6 + 10 + 8 + 4 + 3 + 5 + 9 + 1 = 53$,
$w^- = 2$
$w = min(w^+, w^-) = min(53, 2) = 2$

iv. Decision

$\because$ $w = 2$ lies within the critical region $w \leq 11$,
$\therefore$ reject H_0 at the 5% level.

v. Conclusion

The median weight decreases after joining the diet.

6.8 Wilcoxon Signed Rank Test for Paired Samples (Cont'd)

- The signed rank test can also be used to test

$$H_0:\ \mu_1 - \mu_2 = d_0,$$

where d_0 is a given number, $d_0 \neq 0$.

- In this case, the 2 populations ***need not be symmetric***.
- Subtract d_0 from each difference.
- Rank the ***adjusted differences without regard to sign***.
- Apply the same procedure as above.

Example 6.4

It is claimed that a college senior can increase his/her score in the major field area of the GRE by at least 50 points if he/she is provided with sample problems in advance. To test this claim, 20 college seniors are divided into 10 pairs such that the students in each matched pair have almost the same overall GPA's for their first 3 years in college. Sample problems and answers are provided at random to one member of each pair 1 week prior to the examination. The examination scores are given below:

	Pair									
	1	2	3	4	5	6	7	8	9	10
With Sample Problems	531	621	663	579	451	660	591	719	543	575
Without Sample Problems	509	540	688	502	424	683	568	748	530	524

Use the Wilcoxon signed rank test to test the null hypothesis at the 5% level of significance that sample problems increase the scores by 50 points against the alternative hypothesis that the increase is less than 50 points.

Solution

Let μ_1 and μ_2 be the mean score of all students taking the test in question with and without sample problems, respectively.

1. Hypotheses to be tested

   H_0: $\mu_1 - \mu_2 = d_0 = 50 \quad \leftrightarrow \quad H_1$: $\mu_1 - \mu_2 < 50$.

2. Critical value

   $n = 10$, $\alpha = 5\%$. From Table A.16 (Slide 16), we have the critical value $w \leq 11$.

3. Value of test statistic, w

$d_i = x_{1i} - x_{2i}$	22	81	-25	77	27	-23	23	-29	13	51
$d_i - d_0$	-28	31	-75	27	-23	-73	-27	-79	-37	1
Signed Rank	-5	+6	-9	+3.5	-2	-8	-3.5	-10	-7	+1

$w^+ = 6 + 3.5 + 1 = 10.5,\ w^- = 5 + 9 + 2 + 8 + 3.5 + 10 + 7 = 44.5$

$w = min(w^+, w^-) = min(10.5, 44.5) = 10.5$

4. Decision: $\because 10.5 < 11$, $\therefore$ reject H_0 at the 5% level.

5. Conclusion: Sample problems do not, on average, increase one's graduate record score by as much as 50 points.

Exercise 6.2

1. The weights (kg) of 5 people before and 5 weeks after stopping smoking are shown below:

Status	Person's Weight				
	1	2	3	4	5
Before	66	80	69	52	75
After	71	82	68	56	73

Use the Wilcoxon signed rank test to test the claim, at the 5% level of significance, that the median weight increases if people quit smoking.

Solution

i. Hypotheses to be tested

H_0: $\widetilde{\mu_1} = \widetilde{\mu_2}$ $\leftrightarrow$ H_1: $\widetilde{\mu_1} < \widetilde{\mu_2}$, where $\widetilde{\mu_1}$ and $\widetilde{\mu_2}$ are respectively the median weights before and after stopping smoking.

ii. Critical value

$n = 5$, $\alpha = 5\%$. From Table A.16 (Slide 16), we have the critical value $w \leq 1$.

iii. Value of test statistic, w

$d_i = x_{1i} - x_{2i}$	-5	-2	1	-4	2
Signed Rank	-5	-2.5	+1	-4	+2.5

$w^+ = 1 + 2.5 = 3.5$,
$w^- = 5 + 2.5 + 4 = 11.5$,
$w = min(w^+, w^-) = min(3.5, 11.5) = 3.5$

iv. Decision

$\because w = 3.5 > 1$,
$\therefore$ do not reject H_0 at the 5% level.

v. Conclusion

The median weight does not increase if people quit smoking.

2. The following figures give the systolic blood pressure of 16 joggers before and after an 8-kilometer run:

Jogger	Systolic Blood Pressure	
	Before (x_1)	After (x_2)
1	158	164
2	149	139
3	160	163
4	155	160
5	164	172
6	138	129
7	163	167
8	159	169
9	165	173
10	145	147

Use the Wilcoxon signed rank test, at the 5% level, to test the null hypothesis that jogging 8 km increases in the median systolic blood pressure by 8 points against the alternative hypothesis that the increase in the median is less than 8 points.

29

Solution

Let $\widetilde{\mu_1}$ and $\widetilde{\mu_2}$ be the median systolic blood pressure before and after the run, respectively.

i. Hypotheses to be tested

H_0: $\widetilde{\mu_2} - \widetilde{\mu_1} = d_0 = 8 \;\leftrightarrow\; H_1: \widetilde{\mu_2} - \widetilde{\mu_1} < 8.$

ii. Critical value

$n = 10$, $\alpha = 5\%$. From Table A.16 (Slide 16), we have the critical value $w \leq 11$.

iii. Value of test statistic, w

$d_i = x_{1i} - x_{2i}$	-6	10	-3	-5	-8	9	-4	-10	-8	-2
$d_i - d_0$	-14	2	-11	-13	-16	1	-12	-18	-16	-10
Signed Rank	-7	+2	-4	-6	-8.5	+1	-5	-10	-8.5	-3

$w^+ = 2 + 1 = 3,\; w^- = 7 + 4 + 6 + 8.5 + 5 + 10 + 8.5 + 3 = 52$

$w = min(w^+, w^-) = min(3, 52) = 3$

iv. Decision: $\because 3 < 11$, $\therefore$ reject H_0 at the 5% level.

v. Conclusion: The increase in the median systolic blood pressure is less than 8 points.

30

6.9 Wilcoxon Rank Sum Test for Independent Samples

- When we are interested in testing equality of means of 2 ***continuous distributions*** that are obviously ***non-normal***, and ***samples are independent*** (i.e., there is ***no pairing of observations***), the Wilcoxon rank-sum test is an appropriate ***alternative*** to the ***two-sample unpaired t-test***.

- First, select a ***random sample*** from each of the populations.
  Let n_1 be the number of observations in the ***smaller*** sample;
  n_2 be the number of observations in the ***larger*** sample.
  When $\boldsymbol{n_1 = n_2}$, the ***smaller*** sample can be ***randomly*** assigned.

- Arrange the $n_1 + n_2$ observations of the ***combined sample*** in ***ascending order.***

- Assign a rank of $1, 2, \cdots, (n_1 + n_2)$ for each observation.

- In the case of ties (identical observations), replace the observations by the mean of the ranks.

31

6.9 Wilcoxon Rank Sum Test for Indep. Samples (Cont'd)

Let

w_1 = ***sum of ranks*** corresponding to the $\boldsymbol{n_1}$ observations belonging to the ***smaller*** sample in the ***combined sample*** of size, $n_1 + n_2$;

w_2 = ***sum of ranks*** corresponding to the $\boldsymbol{n_2}$ observations belonging to the ***larger*** sample in the ***combined sample*** of size, $n_1 + n_2$.

- The total $w_1 + w_2$ depends only on the number of observations in the 2 samples and is ***not affected*** by the results of the experiment.
  e.g. If $n_1 = 3, n_2 = 4$, then $w_1 + w_2 = 1 + 2 + \cdots + 7 = 28$, ***regardless*** of the ***numerical values of the observations***. In general,

$$w_1 + w_2 \equiv \frac{1}{2}(n_1 + n_2)(n_1 + n_2 + 1)$$

In fact, the above sum is the sum of the integers: $1, 2, \cdots, (n_1 + n_2)$.

- Once we have determined w_1, w_2 can be obtained by subtraction:

$$w_2 \equiv \frac{1}{2}(n_1 + n_2)(n_1 + n_2 + 1) - w_1$$

32

6.9.1 Test Statistic for Wilcoxon Rank Sum Test

- In practice, we do not use W_1 or W_2 as the test statistic. Rather, we usually base our decision on the ***value*** of U_1, U_2 or U:

$$u_1 = w_1 - \frac{1}{2}n_1(n_1+1),$$

or

$$u_2 = w_2 - \frac{1}{2}n_2(n_2+1),$$

or

$$u = min(u_1, u_2).$$

H_0	H_1	Test Statistic to be Computed
$\mu_1 = \mu_2$	$\mu_1 < \mu_2$	u_1
	$\mu_1 > \mu_2$	u_2
	$\mu_1 \neq \mu_2$	$u = min(u_1, u_2)$

These statistics simplify the construction of tables of critical values.

- Decision Rule: H_0 will be rejected if

$$\boxed{U_1,\ U_2 \text{ or } U\ \leq \text{tabled critical value}}$$

The critical value can be found in Table A.17 in the Slides 35-38.

33

Table A.17 Critical Values for the Wilcoxon Rank-Sum Test

One-Tailed Test at $\alpha = 0.001$ or Two-Tailed Test at $\alpha = 0.002$

	n_2														
n_1	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1															
2															
3												0	0	0	0
4					0	0	0	1	1	1	2	2	3	3	3
5		0	0	1	1	2	2	3	3	4	5	5	6	7	7
6	0	1	2	2	3	4	4	5	6	7	8	9	10	11	12
7		2	3	3	5	6	7	8	9	10	11	13	14	15	16
8			5	5	6	8	9	11	12	14	15	17	18	20	21
9				7	8	10	12	14	15	17	19	21	23	25	26
10					10	12	14	17	19	21	23	25	27	29	32
11						15	17	20	22	24	27	29	32	34	37
12							20	23	25	28	31	34	37	40	42
13								26	29	32	35	38	42	45	48
14									32	36	39	43	46	50	54
15										40	43	47	51	55	59
16											48	52	56	60	65
17												57	61	66	70
18													66	71	76
19														77	82
20															88

34

One-Tailed Test at $\alpha = 0.01$ or Two-Tailed Test at $\alpha = 0.02$

n_1 \ n_2	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1																
2									0	0	0	0	0	0	1	1
3			0	0	1	1	1	2	2	2	3	3	4	4	4	5
4	0	1	1	2	3	3	4	5	5	6	7	7	8	9	9	10
5	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
6		3	4	6	7	8	9	11	12	13	15	16	18	19	20	22
7			6	8	9	11	12	14	16	17	19	21	23	24	26	28
8				10	11	13	15	17	20	22	24	26	28	30	32	34
9					14	16	18	21	23	26	28	31	33	36	38	40
10						19	22	24	27	30	33	36	38	41	44	47
11							25	28	31	34	37	41	44	47	50	53
12								31	35	38	42	46	49	53	56	60
13									39	43	47	51	55	59	63	67
14										47	51	56	60	65	69	73
15											56	61	66	70	75	80
16												66	71	76	82	87
17													77	82	88	93
18														88	94	100
19															101	107
20																114

Based in part on Tables 1, 3, 5, and 7 of D. Auble, "Extended Tables for the Mann-Whitney Statistic," *Bulletin of the Institute of Educational Research at Indiana University*, **1**, No. 2, 1953, by permission of the director.

35

Table A.17 (continued) Critical Values for the Wilcoxon Rank-Sum Test

One-Tailed Test at $\alpha = 0.025$ or Two-Tailed Test at $\alpha = 0.05$

n_1 \ n_2	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1																	
2					0	0	0	0	1	1	1	1	1	2	2	2	2
3		0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	13
5		2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6			5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7				8	10	12	14	16	18	20	22	24	26	28	30	32	34
8					13	15	17	19	22	24	26	29	31	34	36	38	41
9						17	20	23	26	28	31	34	37	39	42	45	48
10							23	26	29	33	36	39	42	45	48	52	55
11								30	33	37	40	44	47	51	55	58	62
12									37	41	45	49	53	57	61	65	69
13										45	50	54	59	63	67	72	76
14											55	59	64	67	74	78	83
15												64	70	75	80	85	90
16													75	81	86	92	98
17														87	93	99	105
18															99	106	112
19																113	119
20																	127

36

One-Tailed Test at $\alpha = 0.05$ or Two-Tailed Test at $\alpha = 0.1$

n_1 \ n_2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1																	0	0
2			0	0	0	1	1	1	1	2	2	3	3	3	3	4	4	4
3	0	0	1	2	2	3	4	4	5	5	6	7	7	8	9	9	10	11
4		1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18
5			4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
6				7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
7					11	13	15	17	19	21	24	26	28	30	33	35	37	39
8						15	18	20	23	26	28	31	33	36	39	41	44	47
9							21	24	27	30	33	36	39	42	45	48	51	54
10								27	31	34	37	41	44	48	51	55	58	62
11									34	38	42	46	50	54	57	61	65	69
12										42	47	51	55	60	64	68	72	77
13											51	56	61	65	70	75	80	84
14												61	66	71	77	82	87	92
15													72	77	83	88	94	100
16														83	89	95	101	107
17															96	102	109	115
18																109	116	123
19																	123	130
20																		138

37

Example 6.5

The nicotine content of two brands of cigarettes, measured in milligrams, was found to be as follows:

Brand A	2.1	4.0	6.3	5.4	4.8	3.7	6.1	3.3	-	-
Brand B	4.1	0.6	3.1	2.5	4.0	6.2	1.6	2.2	1.9	5.4

Test the hypothesis, at the 5% level of significance, that the median nicotine contents of the two brands are equal against the alternative that they are unequal.

38

Solution (Cont'd)

1. Hypotheses to be tested

H_0: $\widetilde{\mu_1} = \widetilde{\mu_2}$ $\leftrightarrow$ H_1: $\widetilde{\mu_1} \neq \widetilde{\mu_2}$,

where $\widetilde{\mu_1}$ and $\widetilde{\mu_2}$ are respectively the median nicotine contents of the 2 brands.

2. Critical value
$n_1 = 8, n_2 = 10,\ \alpha = 5\%$.
From Table A.17 (Slide 36),
we have the critical value $u \leq 17$.

3. Value of test statistic, u

The ranks of the obs belonging to sample A, the smaller sample are marked with asterisks.

Original Data	Ranks	Original Data	Ranks
0.6	1	4.0	10.5*
1.6	2	4.0	10.5
1.9	3	4.1	12
2.1	4*	4.8	13*
2.2	5	5.4	14.5*
2.5	6	5.4	14.5
3.1	7	6.1	16*
3.3	8*	6.2	17
3.7	9*	6.3	18*

*The ranks marked with an asterisk belong to sample A.

$w_1 = 4 + 8 + 9 + 10.5 + 13 + 14.5 + 18 = 93$

$w_2 = \frac{1}{2}(n_1 + n_2)(n_1 + n_2 + 1) - w_1 = \frac{1}{2}(8 + 10)(8 + 10 + 1) - 93 = 78$

39

Solution

3. Value of test statistic (Cont'd)

$$u_1 = w_1 - \frac{1}{2}n_1(n_1 + 1) = 93 - \frac{1}{2}8 \times 9 = 57$$

$$u_2 = w_2 - \frac{1}{2}n_2(n_2 + 1) = 78 - \frac{1}{2}10 \times 11 = 23$$

$$u = min(u_1, u_2) = min(57,23) = 23$$

4. Decision
$\because u = 23 > 17$,
$\therefore$ do not reject H_0 at the 5% level.

5. Conclusion
There is no significant difference in the median nicotine contents of the 2 brands of cigarettes.

40

Exercise 6.3

1. A cigarette manufacturer claims that the median tar content of brand B cigarettes is lower than that of brand A cigarettes. To test this claim, the following determinations of tar content, in milligrams, were recorded:

Brand A	1	12	9	13	11	14
Brand B	8	10	7			

Use the Wilcoxon rank sum test with α=0.05 to test whether or not the claim is valid.

Solution

i. Hypotheses to be tested

$$H_0: \widetilde{\mu_B} = \widetilde{\mu_A} \;\; \leftrightarrow \;\; H_1: \widetilde{\mu_B} < \widetilde{\mu_A},$$

where $\widetilde{\mu_B}$ and $\widetilde{\mu_A}$ are respectively the median tar contents of the brands B and A.

ii. Critical value

$n_B = 3, n_A = 6, \; \alpha = 5\%.$

From Table A.17 (Slide 37), we have the critical value $u_B \leq 2$.

iii. Value of test statistic, u_B

The ranks of the obs belonging to brand B, the smaller sample are marked with asterisks.

Original Data	1	7	8	9	10	11	12	13	14
Rank	1	2*	3*	4	5*	6	7	8	9

Solution

iii. Value of test statistic, u_B (Cont'd)

$w_B = 2 + 3 + 5 = 10$ (The sum of the ranks marked asterisks in the above table)

$u_B = w_1 - \frac{1}{2}n_B(n_B + 1) = 10 - \frac{1}{2}3 \times 4 = 4$

iv. Decision

$\because u_B = 4 > 2$,

$\therefore$ do not reject H_0 at the 5% level.

v. Conclusion

The claim that the tar content of brand B cigarettes is lower than that of brand A is not statistically supported.

2. The following data represent the number of hours that 2 different types of electric motor cars can travel before a recharge is required.

Car A	5.5	5.6	6.3	4.6	5.3	5.0	6.2	5.8	5.1
Car B	3.8	4.8	4.3	4.2	4.0	4.9	4.5	5.2	4.5

Use the Wilcoxon rank sum test with $\alpha = 1\%$ to determine if the median travelling time of the electric motor A travels longer than that of the electric motor B on a full battery charge.

Solution

i. Hypotheses to be tested

H_0: $\widetilde{\mu_A} = \widetilde{\mu_B}$ $\leftrightarrow$ H_1: $\widetilde{\mu_A} > \widetilde{\mu_B}$,

where $\widetilde{\mu_A}$ and $\widetilde{\mu_B}$ are respectively the median travelling time of the cars A and B.

ii. Critical value

$n_A = 9, n_B = 9, \ \alpha = 1\%$.

From Table A.17 (Slide 35), we have the critical value $u_B \leq 14$.

45

Solution

iii. Value of test statistic, u_B

The ranks of the obs belonging to car B, regarded as the smaller sample, are marked with asterisks.

Original Data	3.8	4.0	4.2	4.3	4.5	4.5	4.6	4.8	4.9
Rank	1*	2*	3*	4*	5.5*	5.5*	7	8*	9*
Original Data	5.0	5.1	5.2	5.3	5.5	5.6	5.8	6.2	6.3
Rank	10	11	12*	13	14	15	16	17	18

$w_B = 1 + 2 + 3 + 4 + 5.5 + 5.5 + 8 + 9 + 12 = 50$

$u_B = w_B - \frac{1}{2} n_B(n_B + 1) = 50 - \frac{1}{2} 9 \times 10 = 5$

iv. Decision

$\because u_B = 5 < 14$,

$\therefore$ reject H_0 at the 1% level.

v. Conclusion: Car A can travel longer on a full battery charge.

46

6.10 Merits & Demerits of Nonparametric Methods

Merits

- No normality assumption is required.
- They are easier to do and to be understood.
  Most nonparametric tests do not demand laborious computations:
  A nonparametric test only requires us to replace numerical values by ranks.

Demerits

- Replace actual data values with ranks. Thus, the actual data values are wasted.
- Difficult to compute by hand for large samples
- They do ***not*** use ***all the information*** provided by the sample. Consequently, to achieve the same power of a test, ***a non-parametric test will require a larger sample size than will the corresponding parametric test***.
- When the sample size is ***small***, the ***sensitivity of non-parametric test is low***.

47

Remark

If a parametric and a non-parametric test are both applicable to the same set of data, we should carry out the ***more efficient parametric test***.

48

6.11 Parametric Tests vs. Non-parametric Tests

Situation	Non-parametric Method	Parametric Method
One sample	Wilcoxon signed rank test	Unpaired Z test / unpaired t test
Two dependent/paired samples	Wilcoxon signed rank test	Paired Z test or paired t test
Two independent samples	Wilcoxon rank sum test	Z test or t test for 2 independent samples