STAT 333

Nonparametric Statistics Methods

Prepared by

Abdulrahman Alfaifi — *alfaifi@ksu.edu.sa*

Kholoud Basalim — *kbasalim@ksu.edu.sa*

Munirah Alothman — *malothman3@ksu.edu.sa*

King Saud University
Department of Statistics and Operation Research

1 Mar 2026

بسم الله الرحمن الرحيم

Chapter 1

INTRODUCTION

- Some types of tests presented and their parametric counterparts.

Type of analysis	Nonparametric test	Parametric equivalent
Comparing two related samples	Wilcoxon signed ranks test and sign test	t-Test for dependent samples
Comparing two unrelated samples	Mann–Whitney U-test and Kolmogorov–Smirnov two-sample test	t-Test for independent samples
Comparing three or more related samples	Friedman test	Repeated measures, analysis of variance (ANOVA)
Comparing three or more unrelated samples	Kruskal–Wallis H-test	One-way ANOVA
Comparing categorical data	Chi square χ^2 tests and Fisher exact test	None
Comparing two rank-ordered variables	Spearman rank-order correlation	Pearson product–moment correlation
Comparing two variables when one variable is discrete dichotomous	Point-biserial correlation	Pearson product–moment correlation
Comparing two variables when one variable is continuous dichotomous	Biserial correlation	Pearson product–moment correlation
Examining a sample for randomness	Runs test	None

Parametric tests rely on six main assumptions. If these assumptions are not met, the results may be misleading, and nonparametric tests should be used instead.

1. The data are randomly drowned from a normally distributed population.
2. The populations are approximately equal variances.
3. The sample distribution is approximately normal.
4. Observations are independent of each other, expect for paired values.
5. The data are measured on an interval or ratio scale.
6. The sample size is sufficiently large.

Exercise 1:

1. Which of the following is not true of parametric statistics?

A	They are inferential tests.
B	They assume certain characteristics of population parameters.
C	They assume normality of the population.
D	They are distribution-free.

2. A collection of statistical methods that generally requires very few, if any assumptions about the population distribution is known as……………………………

A	Parametric methods	B	Nonparametric methods
C	Semiparametric	D	None of these

3. A nonparametric method for determining the differences between two populations based on two matched samples where only preference data is required is the

A	Mann-Whitney-Wilcoxon test	B	Wilcoxon signed-rank test
C	Sign test	D	Kruskal-Wallis Test

4. Parametric tests are based on some restrictive assumptions about the ………………

A	Random sample	B	Census	C	Sample	D	Population

5. Nonparametric tests for examining a sample for randomness ………………

A	Fridman test	B	Runs test	C	T - test	D	U - test

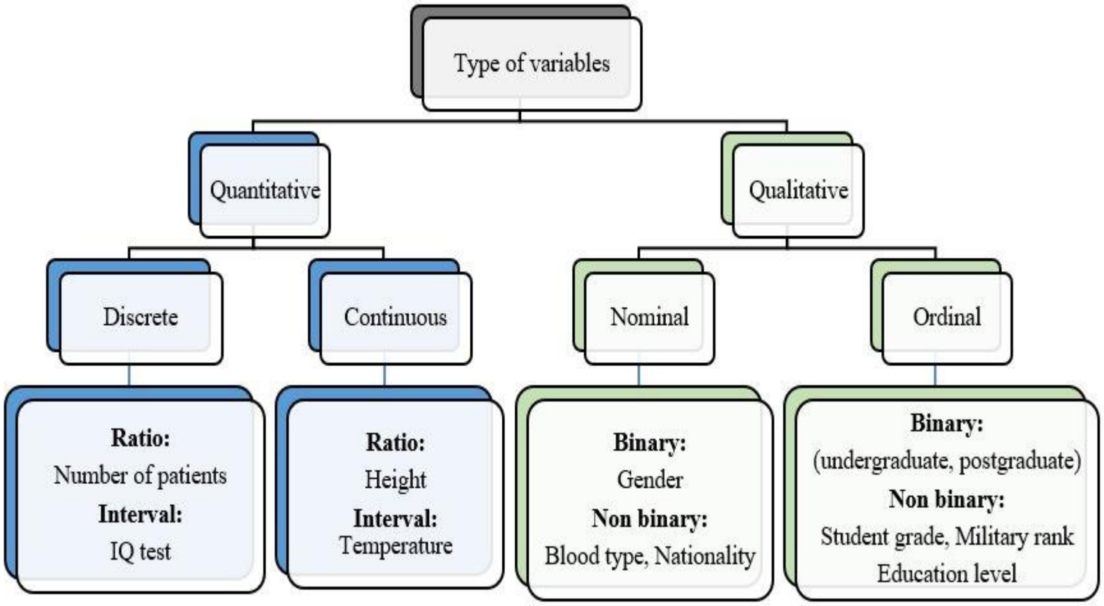
6. Point- biserial correlation is used for ………………

A	Comparing two rank-ordered variables.
B	Comparing two variables when one variable is discrete dichotomous.
C	Comparing two variables when one variable is continuous dichotomous.
D	Comparing two related samples.

7. Parametric test equivalent to Kruskal–Wallis *H*-test is: ………………

A	One-way ANOVA	B	Repeated measures
C	T-test for dependent samples	D	Fridman test

- **Measurement scales:**



Interval scale: Consider as pertinent information not only the relative order of the measurements as in the ordinal scale but also the size of the interval between measurements.
For example:

- Temperature.
- TQ-test.
- Time of the day (00:00 midnight, 14:00 afternoon)

Ratio scale: Not only the order and interval size are important, but also the ratio between two measurements is meaningful.
For example:

- Crop yields (إنتاج المحاصيل).
- Distances.
- Time to finish an exam.
- Weights.
- Heights.
- Income.

Exercise 2: Choose the correct measurement scale.

1. Gender (Male, Female) A) Nominal B) Ordinal C) Interval D) Ratio	8. Height in centimeters A) Ordinal B) Interval C) Ratio D) Nominal
2. Blood type (A, B, AB, O) A) Ordinal B) Nominal C) Ratio D) Interval	9. Saudi national ID number A) Interval B) Ratio C) Nominal D) Ordinal
3. Satisfaction level (Low, Medium, High) A) Nominal B) Ordinal C) Interval D) Ratio	10. Calendar year (2015, 2020, 2025) A) Ratio B) Ordinal C) Interval D) Nominal
4. Age in years A) Nominal B) Ordinal C) Interval D) Ratio	11. Marital status (Single, Married, Divorced) A) Ordinal B) Interval C) Ratio D) Nominal
5. Temperature in Celsius (°C) A) Nominal B) Ordinal C) Interval D) Ratio	12. Pain level (Mild, Moderate, Severe) A) Nominal B) Ordinal C) Interval D) Ratio
6. Exam grades (A, B, C, D) A) Ratio B) Interval C) Ordinal D) Nominal	13. Distance traveled (km) A) Ordinal B) Interval C) Nominal D) Ratio
7. Number of students in a class A) Nominal B) Ordinal C) Interval D) Ratio	14. Eye color A) Ratio B) Interval C) Ordinal D) Nominal

- **Ranking data:**

Example: Rank the following data:

Students who ate breakfast	Students who skipped breakfast
90	75
85	80
95	55
70	90

After ordering	Rank ignoring ties values		Rank accounting for ties values
55	1		1
70	2		2
75	3		3
80	4		4
85	5		5
90	6	$\frac{6+7}{2}=6.5$	6.5
90	7		6.5
95	8		8

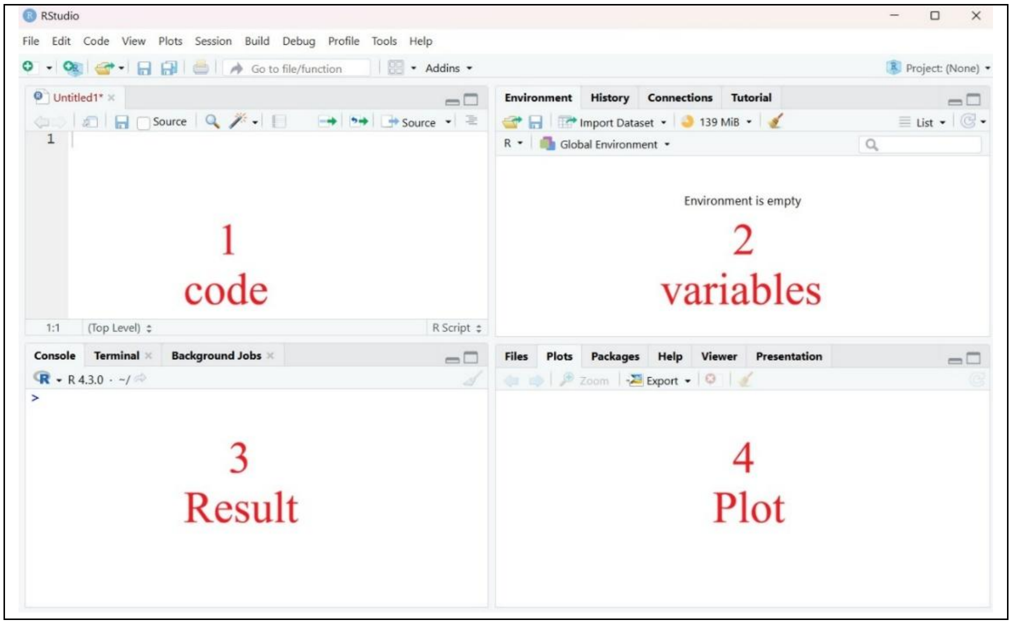
Example: the following data represent quiz score om math.

100	60	70	90	80	100	80	20	100	50

Rank the quiz score.

Quiz score	After ordering	Rank ignoring ties values		Rank accounting for ties values
100	20	1		1
60	50	2		2
70	60	3		3
90	70	4		4
80	80	5	$\frac{5+6}{2}=5.5$	5.5
100	80	6		5.5
80	90	7		7
20	100	8	$\frac{8+9+10}{3}=9$	9
100	100	9		9
50	100	10		9

Using R studio:

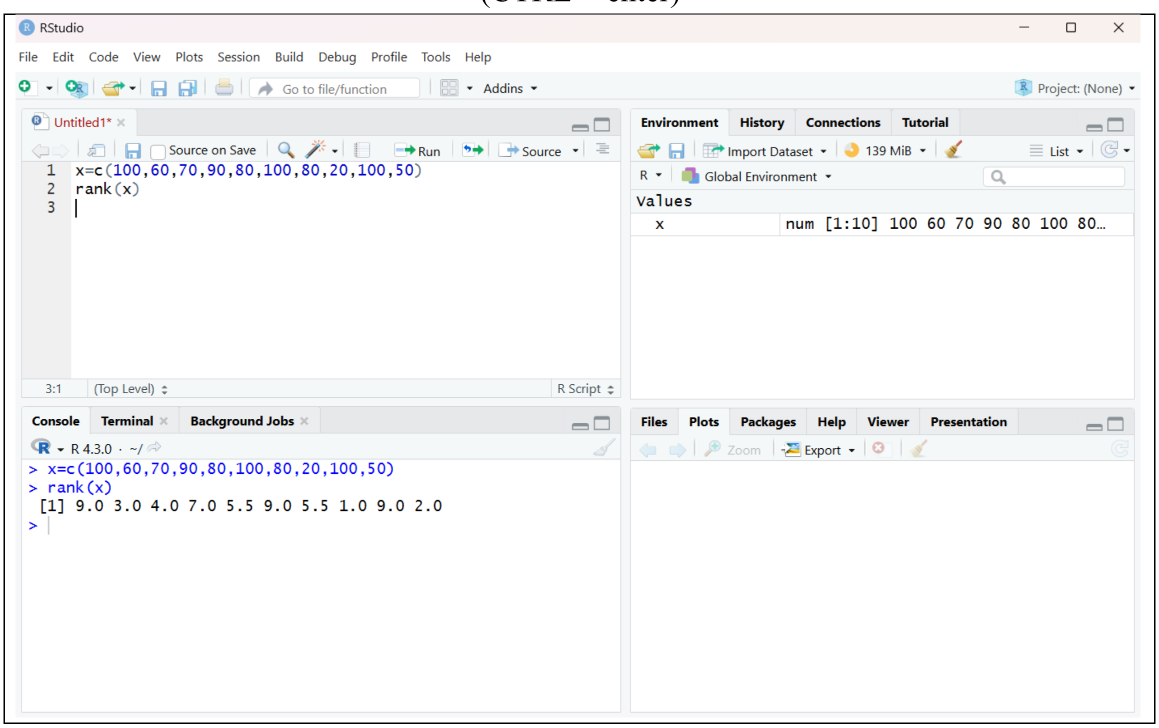


في حال ظهرت 3 شاشات إضغط (CTRL + shift + N)

R - code

```
> x=c(100,60,70,90,80,100,80,20,100,50)
> rank(x)
```

(CTRL + enter)



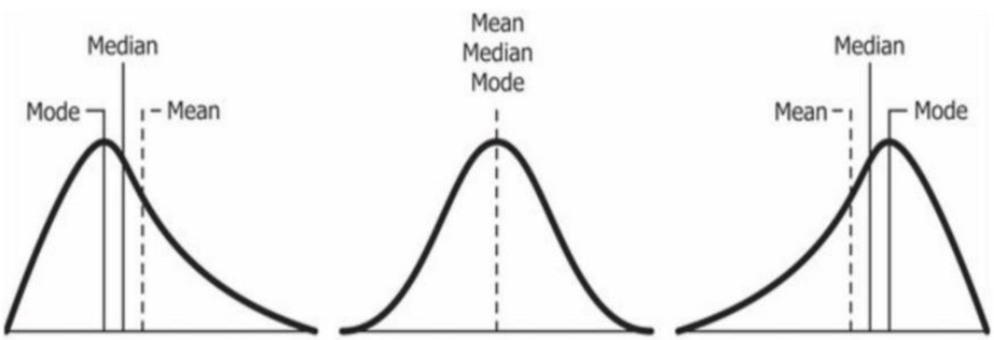
Chapter 2

TESTING DATA FOR NORMALITY

- **Skewness and Kurtosis:**

	Skewness:	Kurtosis:
Formula	$S_k = \frac{n}{(n-1)(n-2)} \sum \left(\frac{x_i - \bar{x}}{s}\right)^3$	$K = \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \left(\frac{x_i - \bar{x}}{s}\right)^4\right] - \frac{3(n-1)^2}{(n-2)(n-3)}$ $s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$
Standard error (SE)	$SE_{S_k} = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}}$	$SE_K = \sqrt{\frac{24n(n-1)^2}{(n-2)(n-3)(n+5)(n+3)}}$
Z - score	$Z_{S_k} = \frac{S_k - 0}{SE_{S_k}}$ If $Z_{S_k} \in (-1.96, 1.96)$ pass the normality assumption for $\alpha = 0.05$	$Z_k = \frac{K - 0}{SE_K}$ If $Z_k \in (-1.96, 1.96)$ pass the normality assumption for $\alpha = 0.05$

Skewness:

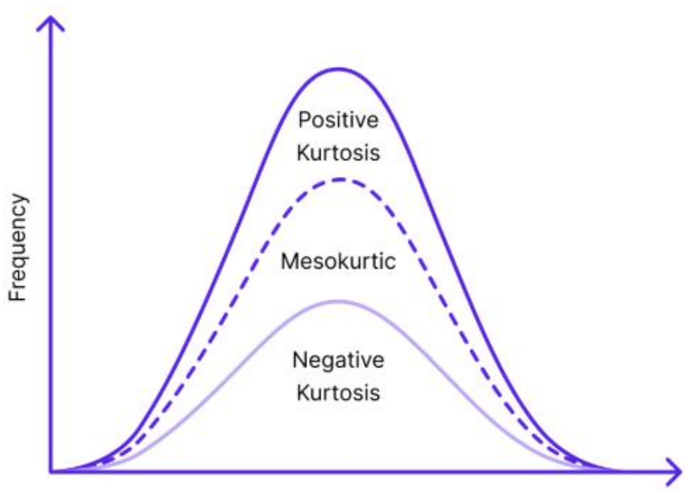


Positive skew
Right skew

No skewness

Negative skew
Left skew

Kurtosis:



Positive kurtosis

Kurtosis = 0

Negative kurtosis

Exercise 1:

The following data represent a samples of week 1 quiz score.
Calculate the skewness and kurtosis.

90	72	90
64	95	89
74	88	100
77	57	35
100	64	95
65	80	84
90	100	76

	data	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$\left(\frac{x_i-\bar{x}}{s}\right)^3$	$\left(\frac{x_i-\bar{x}}{s}\right)^4$
1	90	9.761905	95.29478	0.202561	0.118962
2	64	-16.2381	263.6757	-0.9323	0.91077
3	74	-6.2381	38.91383	-0.05286	0.019837
⋮	⋮	⋮	⋮	⋮	⋮
19	95	14.7619	217.9138	0.700452	0.622068
20	84	3.761905	14.15193	0.011592	0.002624
21	76	-4.2381	17.96145	-0.01658	0.004226
Total	1685		5525.81	-18.4149	69.01972

$$\bar{x} = \frac{\sum x}{n} = \frac{1685}{21} = 80.2381$$

$$S = \sqrt{\frac{\sum(x_i-\bar{x})^2}{n-1}} = \sqrt{\frac{5525.80}{20}} = 6.62199$$

$$\sum\left(\frac{x_i-\bar{x}}{s}\right)^3 = -18.4149$$

$$\sum\left(\frac{x_i-\bar{x}}{s}\right)^4 = 69.01972$$

$$S_k = \frac{n}{(n-1)(n-2)}\sum\left(\frac{x_i-\bar{x}}{s}\right)^3$$

$$= \frac{21}{(21-1)(21-2)} \times -18.4149 = -1.01766$$

$$SE_{S_k} = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}} = \sqrt{\frac{6\times 21(21-1)}{(21-2)(21+1)(21+3)}} = 0.50119$$

$$Z_{S_k} = \frac{S_k - 0}{SE_{S_k}}$$

$$= \frac{-1.01766-0}{0.50119} = -2.032$$

$$Z_{S_k} \notin (-1.96, 1.96)$$

$$K = \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)}\sum\left(\frac{x_i-\bar{x}}{s}\right)^4\right] - \frac{3(n-1)^2}{(n-2)(n-3)}$$

$$= \left[\frac{21(21+1)}{(21-1)(21-2)(21-3)} \times 69.02\right] - \frac{3(21-1)^2}{(21-2)(21-3)} = 1.153$$

$$SE_K = \sqrt{\frac{24n(n-1)^2}{(n-2)(n-3)(n+5)(n+3)}}$$

$$= \sqrt{\frac{24\times 21(21-1)^2}{(21-2)(21-3)(21+5)(21+3)}} = 0.971941$$

$$Z_k = \frac{K - 0}{SE_K}$$

$$= \frac{1.153-0}{0.971941} = 1.186$$

$$Z_k \in (-1.96, 1.96)$$

R – code

```
x=c(90,72,90,64,95,89,74,88,100,77,57,35,100,64,95,65,80,84,90,100,76)
m=mean(x)
s=sd(x)
n=length(x)

i3=sum(((x-m)/s)^3)
i4=sum(((x-m)/s)^4)

sk=n/((n-1)*(n-2))*i3
SEs=sqrt(6*n*(n-1)/(n-2)/(n+1)/(n+3))

kur=(n*(n+1)/(n-1)/(n-2)/(n-3)*i4)-(3*(n-1)^2/(n-2)/(n-3))
SEk=sqrt((24*n*(n-1)^2)/(n-2)/(n-3)/(n+5)/(n+3))
```

```
x=c(90,72,90,64,95,89,74,88,100,77,57,35,100,64,95,65,80,84,90,100,76)
m=mean(x)
s=sd(x)
```

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

```
n=length(x)
i3=sum(((x-m)/s)^3)
```

$$\sum \left(\frac{x_i - \bar{x}}{s}\right)^3$$

```
i4=sum(((x-m)/s)^4)
```

$$\sum \left(\frac{x_i - \bar{x}}{s}\right)^4$$

```
sk=n/((n-1)*(n-2))*i3
```

$$S_k = \frac{n}{(n-1)(n-2)} \sum \left(\frac{x_i - \bar{x}}{s}\right)^3$$

```
SEs=sqrt(6*n*(n-1)/(n-2)/(n+1)/(n+3))
```

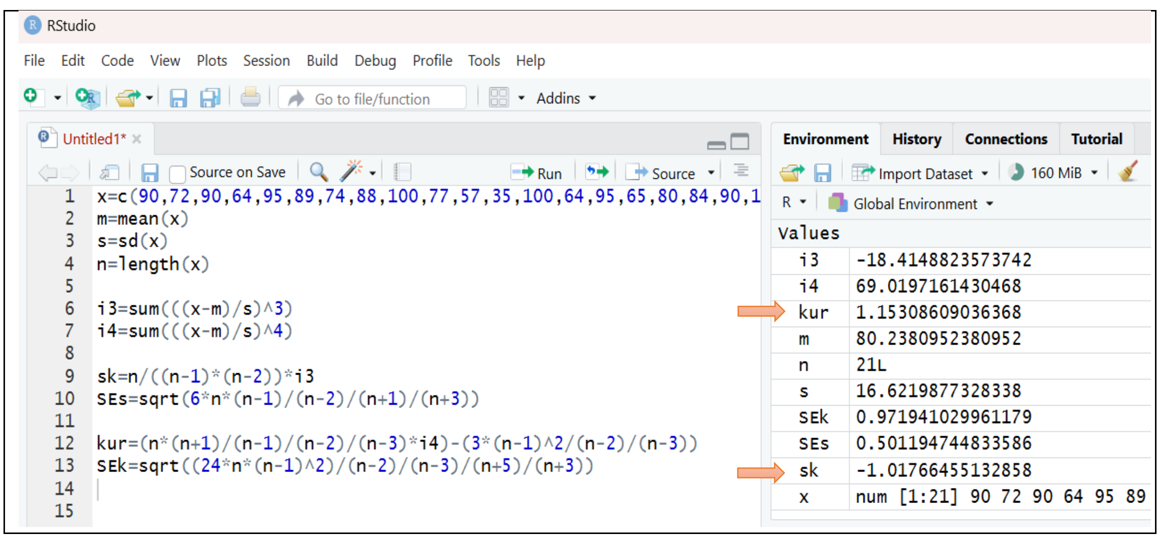
$$SE_{S_k} = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}}$$

```
kur=(n*(n+1)/(n-1)/(n-2)/(n-3)*i4)-(3*(n-1)^2/(n-2)/(n-3))
```

$$K = \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \left(\frac{x_i - \bar{x}}{s}\right)^4\right] - \frac{3(n-1)^2}{(n-2)(n-3)}$$

```
SEk=sqrt((24*n*(n-1)^2)/(n-2)/(n-3)/(n+5)/(n+3))
```

$$SE_K = \sqrt{\frac{24n(n-1)^2}{(n-2)(n-3)(n+5)(n+3)}}$$



Exercise 2:

A department store has decided to evaluate customer satisfaction. The store provides customers with a survey to rate employee friendliness. The survey uses a scale of 1–10. The survey results are:

7	3	3	6
4	4	4	5
5	5	8	9
5	5	5	7
6	8	6	2

Calculate the skewness and kurtosis.

	data	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$\left(\frac{x_i-\bar{x}}{s}\right)^3$	$\left(\frac{x_i-\bar{x}}{s}\right)^4$
1	7				
2	4				
3	5				
⋮	⋮	⋮	⋮	⋮	⋮
18	9				
19	7				
20	2				
Total	107		62.55	4.09576	45.0707

$\bar{x} =$ ....

$s =$ ....

$\sum \left(\frac{x_i-\bar{x}}{s}\right)^3 =$ ....

$\sum \left(\frac{x_i-\bar{x}}{s}\right)^4 =$ ....

$$S_k = \frac{n}{(n-1)(n-2)} \sum \left(\frac{x_i-\bar{x}}{s}\right)^3 = \text{....................................}$$

$$SE_{S_k} = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}} = \text{....................................}$$

$$Z_{S_k} = \frac{S_k - 0}{SE_{S_k}} =$$

$$K = \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \left(\frac{x_i-\bar{x}}{s}\right)^4\right] - \frac{3(n-1)^2}{(n-2)(n-3)} =$$

....

$$Z_k = \frac{K-0}{SE_K} =$$

$$SE_K = \sqrt{\frac{24n(n-1)^2}{(n-2)(n-3)(n+5)(n+3)}} = \text{....................................}$$

R – code

```
x=c(7,4,5,5,6,3,4,5,5,8,3,4,8,5,6,6,5,9,7,2)
```

Exercise 3:

Calculate the skewness and kurtosis for the following data.

	data	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$\left(\frac{x_i-\bar{x}}{S}\right)^3$	$\left(\frac{x_i-\bar{x}}{S}\right)^4$
1	25				
2	30				
3	12				
4	18				
5	20				
Total					

$\bar{x} =$ ....

$S =$ ....

$\sum \left(\frac{x_i-\bar{x}}{S}\right)^3 =$ ....

$\sum \left(\frac{x_i-\bar{x}}{S}\right)^4 =$ ....

$S_k =$ ....

....

....

$Z_{S_k} = \frac{S_k - 0}{SE_{S_k}} =$

$SE_{S_k} =$ ....

....

....

$K =$ ....

....

....

$Z_k = \frac{K - 0}{SE_K} =$

$SE_K =$ ....

....

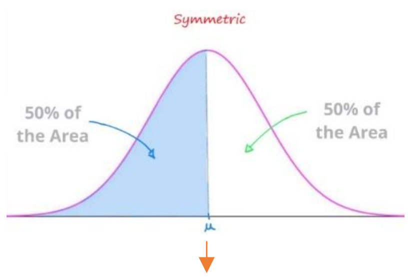
....

R – code

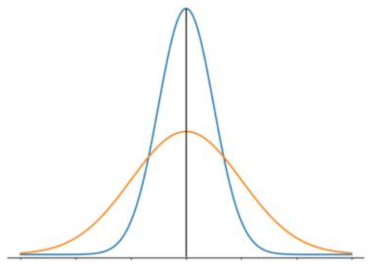
```
x=c(25,30,12,18,20)
```

The Normal Distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}; \quad -\infty < X < \infty$$



Mean = Median = Mode



μ

$\sigma > \sigma$

Normal distribution	$X \sim N(\mu, \sigma^2)$	
Standard normal	$Z \sim N(0, 1)$	$\boxed{Z = \frac{X-\mu}{\sigma} \sim N(0, 1)}$

Testing Data for Normality: Kolmogorov–Smirnov

H_0: The data approximately follow normal distribution.
H_A: The data do not follow normal distribution.

Question: Suppose that we have a random sample of size n. For α = 0.05, if the Z-score of the skewness of the sample is (−2.032), then the sample has Therefore, either the sample must be modified and rechecked or you must use a nonparametric statistical test.

A	Pass the normality assumption for kurtosis
B	Pass the normality assumption for skewness.
C	Failed the normality assumption for kurtosis
D	Failed the normality assumption for skewness.

Exercise 4:

For the following data:

8.1	8.2	8.2	8.7	8.7	8.8	8.8	8.9	8.9	8.9
9.2	9.2	9.2	9.3	9.3	9.3	9.4	9.4	9.4	9.4
9.5	9.5	9.5	9.5	9.6	9.6	9.6	9.7	9.7	9.9

(a) Find: Skewness, Standard error of the skewness, Kurtosis, Standard error of the kurtosis

Using R

```
y=c(8.1,9.2,9.5,8.2,9.2,9.5,8.2,9.2,9.5,8.7,9.3,9.5 ,8.7,9.3,9.6,8.8,9.3,9.6,8.8,9.4,9.6,8.9,9.4,9.7
        ,8.9,9.4,9.7,8.9,9.4,9.9)
m=mean(y)
s=sd(y)
n=length(y)
i3=sum(((y-m)/s)^3)
i4=sum(((y-m)/s)^4)
sk=n/((n-1)*(n-2))*i3
SEs=sqrt(6*n*(n-1)/(n-2)/(n+1)/(n+3))
kur=(n*(n+1)/(n-1)/(n-2)/(n-3)*i4)-(3*(n-1)^2/(n-2)/(n-3))
SEk=sqrt((24*n*(n-1)^2)/(n-2)/(n-3)/(n+5)/(n+3))

> sk
[1] -0.9043788
> SEs
[1] 0.4268924
> kur
[1] 0.1877582
> SEk
[1] 0.8327456
```

(b) Using a Kolmogorov–Smirnov one-sample test, is the data follow normal distribution

Using R

```
ks.test(y, "pnorm", mean = mean(y), sd = sd(y))
            Asymptotic one-sample Kolmogorov-Smirnov test
data:  y
D = 0.18377, p-value = 0.263
alternative hypothesis: two-sided
```

H_0: The data approximately follow normal distribution.
H_A: The data do not follow normal distribution.

$P - value = 0.263 > 0.05$,
We accept H_0, The data approximately follow normal distribution

Exercise 5:

A department store has decided to evaluate customer satisfaction. The store provides customers with a survey to rate employee friendliness. The survey uses a scale of 1–10. The survey results are:

7	3	3	6
4	4	4	5
5	5	8	9
5	5	5	7
6	8	6	2

Use the Kolmogorov–Smirnov one-sample test to decide if survey results approximately matching a normal distribution.

Using R

```
x=c(7,4,5,5,6,3,4,5,5,8,3,4,8,5,6,6,5,9,7,2)
ks.test(y, "pnorm", mean = mean(x), sd = sd(x))
        Asymptotic one-sample Kolmogorov-Smirnov test

data:  x
D = 0.17648, p-value = 0.5617
alternative hypothesis: two-sided
```

H_0: The data approximately follow normal distribution.
H_A: The data do not follow normal distribution.

$P - value = 0.5617 > 0.05$,
We accept H_0, The data approximately follow normal distribution

Chapter 3

THE WILCOXON SIGNED RANK AND THE SIGN TEST

The Wilcoxon signed test and sing test: are for comparing two samples that are paired or related. $H_0: \mu_D = 0$ $H_A: \mu_D \neq 0$ Parametric equivalent: *t*-Test for dependent samples (Paired test)

The sum of the ranks with positive differences	$\sum R_+$
The sum of the ranks with negative differences	$\sum R_-$
Test statistics	$T = \min(\sum R_+, \sum R_-)$
The mean	$\bar{x}_T = \frac{n(n+1)}{4}$
Standard deviation	$S_T = \sqrt{\frac{n(n+1)(2n+1)}{24}}$
z-score	$Z = \frac{T - \bar{x}_T}{S_T}$
Effect size (ES): Determine the degree of association between the groups	$ES = \frac{\lvert Z \rvert}{\sqrt{n}}$; $0 < ES < 1$ ES closer to 0.10 small ES closer to 0.30 medium ES closer to 0.50 large

Exercise 1:

The counseling staff (فريق الإرشاد) in a school started a new program this year to reduce bullying (التنمر) in elementary schools. To see if the program worked, they compared the percentage of successful interventions (التدخلات) before the program (last year) with the percentage after the program (this year). The data were reported by 12 elementary school counselors.

(a) Is their difference in percentage of successful interventions in the two years.

Participants	Successful intervention	
	last year	this year
1	31	31
2	14	14
3	53	50
4	18	30
5	21	28
6	44	48
7	12	35
8	36	32
9	22	23
10	29	34
11	17	27
12	40	42

Participants	Successful intervention		Different	Rank	Sign
	last year	this year		without zero	
1	31	31	0	-	
2	14	14	0	-	
3	53	50	-3	3	-
4	18	30	12	9	+
5	21	28	7	7	+
6	44	48	4	4.5	+
7	12	35	23	10	+
8	36	32	-4	4.5	-
9	22	23	1	1	+
10	29	34	5	6	+
11	17	27	10	8	+
12	40	42	2	2	+

1. Using Wilcoxon test:

$H_0: \mu_D = 0$ (There is no difference in the percentages)
$H_A: \mu_D \neq 0$ (There is a difference in the percentages)

$\sum R_- = 3 + 4.5 = 7.5$
$\sum R_+ = 9 + 7 + 4.5 + 10 + 1 + 6 + 8 + 2 = 47.5$

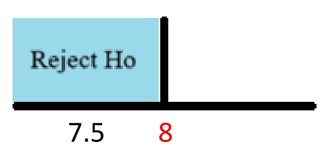
$\mathrm{T} = \min(\sum \mathrm{R}_+ , \sum \mathrm{R}_-)$
$= \min(47.5\ ,\ 7.5\ ) = 7.5$

We use (Table B.3 page 244)

TABLE B.3 Critical Values for the Wilcoxon Signed Rank Test Statistics *T*.

n	$\alpha_{\text{two-tailed}} \le 0.10$ $\alpha_{\text{one-tailed}} \le 0.05$	$\alpha_{\text{two-tailed}} \le 0.05$ $\alpha_{\text{one-tailed}} \le 0.025$	$\alpha_{\text{two-tailed}} \le 0.02$ $\alpha_{\text{one-tailed}} \le 0.01$	$\alpha_{\text{two-tailed}} \le 0.01$ $\alpha_{\text{one-tailed}} \le 0.005$
5	0			
6	2	0		
7	3	2	0	
8	5	3	1	0
9	8	5	3	1
10	10	8	5	3
11	13	10	7	5
12	17	13	9	7
13	21	17	12	9
14	25	21	15	12
15	30	25	19	15

$\mathrm{T} = 7.5 < 8$ then we reject H_0
(There is a difference in the percentages)



Using R

```
x=c(31,14,53,18,21,44,12,36,22,29,17,40)
y=c(31,14,50,30,28,48,35,32,23,34,27,42)
wilcox.test(x,y, paired = TRUE, alternative = "two.sided")
        Wilcoxon signed rank test with continuity correction

data:  x and y
V = 7.5, p-value = 0.04671 <----------- Reject H0
alternative hypothesis: true location shift is not equal to 0
```

Construct a 95% median confidence interval based on the Wilcoxon signed rank test

	-3	12	7	4	23	-4	1	5	10	2
-3	-3	4.5	2	0.5	10	-3.5	-1	1	3.5	-0.5
12		12	9.5	8	17.5	4	6.5	8.5	11	7
7			7	5.5	15	1.5	4	6	8.5	4.5
4				4	13.5	0	2.5	4.5	7	3
23					23	9.5	12	14	16.5	12.5
-4						-4	-1.5	0.5	3	-1
1							1	14	5.5	1.5
5								5	7.5	3.5
10									10	6
2										2

$$U_{ij} = \frac{D_i + D_j}{2}$$

$$1 \le i \le j \le n$$

$$\mathrm{K} = \mathrm{T} + 1 = 8 + 1 = 9$$

1	2	3	4	5	6	7	8	9	…
-4	-3.5	-3	-1.5	-1	-1	-0.5	0	0.5	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
…	47	48	49	50	51	52	53	54	55
	12	12	12.5	13.5	14	15	16.5	17.5	23

95% confidence that the difference number of activities lies between (0.5 , 12)

(c) Determine the degree of association.

$$\bar{x}_T = \frac{n(n+1)}{4} = \frac{10(10+1)}{4} = 27.5$$

$$S_T = \sqrt{\frac{n(n+1)(2n+1)}{24}} = \sqrt{\frac{10(10+1)(20+1)}{24}} = 9.81$$

$$Z = \frac{T - \bar{x}_T}{S_T} = \frac{7.5 - 27.5}{9.81} = -2.0387$$

Effect size:

$ES = \frac{|Z|}{\sqrt{n}} = \frac{|-2.0387|}{\sqrt{10}} = 0.64$ (Which indicates a strong measure of association.)

2. Using the sign test:

$H_0: P = 0.5$ (There is no difference in the percentages)
$H_A: P \neq 0.5$ (There is a difference in the percentages)

If $n < 25$ we calculate p using $P(X) = \frac{n!}{(n-X)!X!} P^X (1-P)^{n-X}$

If $n \geq 25$ we calculate p using $Z_c = \frac{max(n_p, n_n) - 0.5(n_p + n_n) - 0.5}{0.5\sqrt{n_p + n_n}}$

$n = n_p + n_n$, n_p = number of the positive differences
n_n = number of the negative differences

Participants	Successful intervention		Sign of the difference
	last year	this year	
1	31	31	
2	14	14	
3	53	50	-
4	18	30	+
5	21	28	+
6	44	48	+
7	12	35	+
8	36	32	-
9	22	23	+
10	29	34	+
11	17	27	+
12	40	42	+

$$n = n_p + n_n = 10 < 25 \text{ , then: } P(X) = \frac{n!}{(n-X)!\,X!} P^X (1-P)^{n-X}$$

$P(X=0) = \frac{10!}{(10-0)!0!} 0.5^0 (1-0.5)^{10-0} = 0.0010$

$P(X=1) = \frac{10!}{(10-1)!1!} 0.5^1 (1-0.5)^{10-1} = 0.0098$

$P(X=2) = \frac{10!}{(10-2)!2!} 0.5^2 (1-0.5)^{10-2} = 0.0439$

P-values for each tail,
we sum the probabilities for each tail until we find a probability $\geq \frac{\alpha}{2} = 0.025$

X	0	1	2	3	4	5	6	7	8	9	10
$P(X)$	0.001	0.0098	0.0439	0.1173	0.2051	0.1172	0.2051	0.1172	0.0439	0.0098	0.001

0.0547 median 0.0547

$\text{P} - \text{value} = 0.0547 + 0.0547$
$\text{P} - \text{value} = 0.1094 > 0.05$

We fail to reject H_0 (There is no difference in the percentages)

Using R

```
x=c(31,14,53,18,21,44,12,36,22,29,17,40)
y=c(31,14,50,30,28,48,35,32,23,34,27,42)
d=y-x
d_nz=d[d!=0]
np=sum(d_nz>0)
nn=sum(d_nz<0)
n=length(d_nz)
#n=np+nn
binom.test(np,n,p=0.5,alternative = "two.sided")

        Exact binomial test
data:  np and n
number of successes = 8, number of trials = 10, p-value = 0.1094 ←——— Accept H0
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
 0.4439045 0.9747893
sample estimates:
probability of success
                   0.8
#or
# you have to install BSDA package first using install.packages("BSDA")
library(BSDA)
SIGN.test(x,y,md=0,alternative = "two.sided")

Dependent-samples Sign-Test
data:  x and y
S = 2, p-value = 0.1094 ←——— Accept H0
alternative hypothesis: true median difference is not equal to 0
95 percent confidence interval:
 -9.680909  0.000000
sample estimates:
median of x-y
           -3
Achieved and Interpolated Confidence Intervals:
                  Conf.Level   L.E.pt
Lower Achieved CI     0.8540  -7.0000
Interpolated CI       0.9500  -9.6809
Upper Achieved CI     0.9614 -10.0000
                  U.E.pt
Lower Achieved CI      0
Interpolated CI        0
Upper Achieved CI      0
```

Exercise 2:

A school is trying to get more students to participate in activities that will make learning more desirable. Table below shows the number of activities that each of the 10 students in one class participated in last year before a new activity program was implemented and this year after it was implemented (تم تطبيقه). Construct a 95% median confidence interval based on the Wilcoxon signed rank test to determine whether the new activity program had a significant positive effect on the student participation.

Participants	Activities	
	last year	this year
1	18	20
2	22	28
3	10	18
4	25	23
5	16	20
6	14	21
7	21	17
8	13	18
9	28	22
10	12	21

Activities		Difference
last year X_j	this year X_i	$D_i = X_i - X_j$
18	20	2
22	28	6
10	18	8
25	23	-2
16	20	4
14	21	7
21	17	-4
13	18	5
28	22	-6
12	21	9

TABLE B.3 Critical Values for the Wilcoxon Signed Rank Test Statistics *T*.

n	$\alpha_{\text{two-tailed}} \le 0.10$ $\alpha_{\text{one-tailed}} \le 0.05$	$\boxed{\alpha_{\text{two-tailed}} \le 0.05}$ $\boxed{\alpha_{\text{one-tailed}} \le 0.025}$	$\alpha_{\text{two-tailed}} \le 0.02$ $\alpha_{\text{one-tailed}} \le 0.01$	$\alpha_{\text{two-tailed}} \le 0.01$ $\alpha_{\text{one-tailed}} \le 0.005$
5	0			
6	2	0		
7	3	2	0	
8	5	3	1	0
9	8	5	3	1
10	10	8	5	3
11	13	10	7	5
12	17	13	9	7
13	21	17	12	9
14	25	21	15	12
15	30	25	19	15

$n = 10$
$\alpha = 0.05$
$T = 8$

	2	6	8	-2	4	7	-4	5	-6	9
2	2	4	5	0	3	4.5	-1	3.5	-2	5.5
6		6	7	2	5	6.5	1	5.5	0	7.5
8			8	3	6	7.5	2	6.5	1	8.5
-2				-2	1	2.5	-3	1.5	-4	3.5
4					4	5.5	0	4.5	-1	6.5
7						7	1.5	6	0.5	8
-4							-4	0.5	-5	2.5
5								5	-0.5	7
-6									-6	4.5
9										9

$$U_{ij} = \frac{D_i + D_j}{2}$$

$$1 \le i \le j \le n$$

$$\mathrm{K} = \mathrm{T} + 1 = 8 + 1 = 9$$

1	2	3	4	5	6	7	8	9	…
-6	-5	-4	-4	-3	-2	-2	-1	-1	

…	47	48	49	50	51	52	53	54	55
	7	7	7	7.5	7.5	8	8	8.8	9

95% confidence that the difference number of activities lies between $(-1, 7)$

Exercise 3:

Twenty participants in an exercise program were measured on the number of sit-ups they could do before other physical exercise (first count) and the number they could do after they had done at least 45 min of other physical exercise (second count). Table 4 shows the results for 20 participants obtained during two separate physical exercise sessions. Determine the ES for a calculated z-score.

Participant	First count	Second count
1	18	28
2	19	18
3	20	28
4	29	20
5	15	30
6	22	25
7	21	28
8	30	18
9	22	27
10	11	30
11	20	24
12	21	27
13	21	10
14	20	40
15	18	20
16	27	14
17	24	29
18	13	30
19	10	24
20	10	36

Participant	1^{st}*count*	2^{nd} *count*	*D*	$\|D\|$	Sing	Rank $\|D\|$
1	18	28	10	10	+	11
2	19	18	-1	1	-	1
3	20	28	8	8	+	9
4	29	20	-9	9	-	10
5	15	30	15	15	+	16
6	22	25	3	3	+	3
7	21	28	7	7	+	8
8	30	18	-12	12	-	13
9	22	27	5	5	+	5.5
10	11	30	19	19	+	18
11	20	24	4	4	+	4
12	21	27	6	6	+	7
13	21	10	-11	11	-	12
14	20	40	20	20	+	19
15	18	20	2	2	+	2
16	27	14	-13	13	-	14
17	24	29	5	5	+	5.5
18	13	30	17	17	+	17
19	10	24	14	14	+	15
20	10	36	26	26	+	20

$$\sum R_{-} = 1 + 10 + 13 + 12 + 14 = 50$$
$$\sum R_{+} = 11 + 9 + \cdots + 15 + 20 = 160$$

$$\mathrm{T} = \min(\sum \mathrm{R}_{+}, \sum \mathrm{R}_{-})$$
$$= \min(\ 160\ ,\ 50\ ) = 50$$

$$\bar{x}_T = \frac{n(n+1)}{4} = \frac{20(20+1)}{4} = 105$$

$$S_T = \sqrt{\frac{n(n+1)(2n+1)}{24}} = \sqrt{\frac{20(20+1)(40+1)}{24}} = 26.786$$

$$Z = \frac{T - \bar{x}_T}{S_T} = \frac{50-105}{26.786} = -2.0533$$

Effect size:

$$ES = \frac{|Z|}{\sqrt{n}} = \frac{|-2.0533|}{\sqrt{20}} = 0.46 \text{ (Which indicates a strong measure of association.)}$$

<u>Exercise 4:</u>

Consider a clinical investigation to assess the effectiveness of a new drug designed to reduce repetitive behaviors in children affected with autism. The data are shown below.

Child	1	2	3	4	5	6	7	8	9	10
Before Treatment	30	56	48	47	43	45	36	44	44	40
After 2 Weeks of Treatment	39	46	37	44	32	39	41	40	38	46

Use a one-tailed Wilcoxon signed rank test and a one-tailed sign test to assess the effectiveness of the drug (is there differences in behavior before and after taking the drug?). Use α = 0.05.

Child	Repetitive behaviors		Different	Rank	sign
	Before treatment	After treatment		without zero	
1	30	39	9	7	+
2	56	46	-10	8	-
3	48	37	-11	9.5	-
4	47	44	-3	1	-
5	43	32	-11	9.5	-
6	45	39	-6	5	-
7	36	41	5	3	+
8	44	40	-4	2	-
9	44	38	-6	5	-
10	40	46	6	5	+

1. Using Wilcoxon test:
   $H_0: \mu_D \geq 0$ (The effectiveness of a new drug does not reduce repetitive behaviors)
   $H_A: \mu_D < 0$ (The effectiveness of a new drug reduced repetitive behaviors)

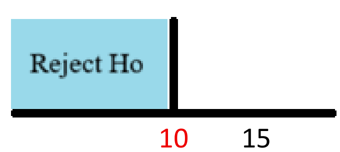
$\sum R_- = 8 + 9.5 + 1 + 9.5 + 5 + 2 + 5 = 40$
$\sum R_+ = 7 + 3 + 5 = 15$

$\mathrm{T} = \min(\sum \mathrm{R}_+, \sum \mathrm{R}_-) = \min(40, 15) = 15$

TABLE B.3 Critical Values for the Wilcoxon Signed Rank Test Statistics *T*.

n	$\alpha_{\text{two-tailed}} \leq 0.10$ $\alpha_{\text{one-tailed}} \leq 0.05$	$\alpha_{\text{two-tailed}} \leq 0.05$ $\alpha_{\text{one-tailed}} \leq 0.025$	$\alpha_{\text{two-tailed}} \leq 0.02$ $\alpha_{\text{one-tailed}} \leq 0.01$	$\alpha_{\text{two-tailed}} \leq 0.01$ $\alpha_{\text{one-tailed}} \leq 0.005$
5	0			
6	2	0		
7	3	2	0	
8	5	3	1	0
9	8	5	3	1
10	10	8	5	3
11	13	10	7	5
12	17	13	9	7
13	21	17	12	9
14	25	21	15	12
15	30	25	19	15

$\mathrm{T} = 15 > 10$ then we fail to reject H_0
(The effectiveness of a new drug does not reduce repetitive behaviors)



Using R

```
before=c(30,56,48,47,43,45,36,44,44,40)
after  =c(39,46,37,44,32,39,41,40,38,46)
wilcox.test(before,after, paired = TRUE, alternative = "greater")
            Wilcoxon signed rank test with continuity correction
data:  before and after
V = 40, p-value = 0.1099 <------ Accept H0
alternative hypothesis: true location shift is greater than 0
```

2. Using the sign test:

$H_0: P \geq 0.5$ (The effectiveness of a new drug does not reduce repetitive behaviors)

$H_A: P < 0.5$ (The effectiveness of a new drug reduced repetitive behaviors)

$$n = n_p + n_n = 10 < 25 \;,\; \text{then: } P(X) = \frac{n!}{(n-X)!\,X!} P^X (1-P)^{n-X}$$

$$P(X=0) = \frac{10!}{(10-0)!0!} 0.5^0 (1-0.5)^{10-0} = 0.0010$$

$$P(X=1) = \frac{10!}{(10-1)!1!} 0.5^1 (1-0.5)^{10-1} = 0.0098$$

$$P(X=2) = \frac{10!}{(10-2)!2!} 0.5^2 (1-0.5)^{10-2} = 0.0439$$

P-values for left tail, $P - \text{value} = P(X \leq 3)$

X	0	1	2	3	4	5	6	7	8	9	10
$P(X)$	0.001	0.0098	0.0439	0.1173	0.2051	0.1172	0.2051	0.1172	0.0439	0.0098	0.001

0.1719 median

$P - \text{value} = 0.1719 > 0.05$ We fail to reject H_0

(The effectiveness of a new drug does not reduce repetitive behaviors)

Using R

```
before=c(30,56,48,47,43,45,36,44,44,40)
after  =c(39,46,37,44,32,39,41,40,38,46)
d = before-after
d_nz=d[d!=0]
np=sum(d_nz>0)
nn=sum(d_nz<0)
#n=np+nn
n=length(d_nz)
binom.test(np,n,p=0.5,alternative = "greater")
#d = after-before
#binom.test(np,n,p=0.5,alternative = "less")

            Exact binomial test
data:  np and n
number of successes = 7, number of trials = 10, p-value = 0.1719 <------ Accept H0
alternative hypothesis: true probability of success is greater than 0.5
95 percent confidence interval:
 0.3933758 1.0000000
sample estimates:
probability of success
                0.7
```

Chapter 4

THE MANN−WHITNEY U-TEST AND THE KOLMOGOROV−SMIRNOV TWO-SAMPLE TEST

The Mann – Whitney U – test and The Kolmogorov – Smirnov two sample test:
For comparing two samples that are independent, or not related.

H_0: There is no significant different between the two methods.
H_A: There is a significant different between the two methods.

Parametric equivalent: *t*-Test for independent samples.

The sum of the ranks for group i	$\sum R_i$
The number of values from the i^{th} sample.	n_i
Test statistics is the smallest of	$U_i = n_1 n_2 + \frac{n_i(n_i+1)}{2} - \sum R_i$
The mean	$\bar{x}_U = \frac{n_1 n_2}{2}$
Standard deviation	$S_U = \sqrt{\frac{n_1 n_2 (n_1+n_2+1)}{12}}$
z-score	$Z = \frac{U_i - \bar{x}_U}{S_U}$
Effect size (ES): Determine the degree of association between the groups	$ES = \frac{\lvert Z \rvert}{\sqrt{n}}$; $0 < ES < 1$ ES closer to 0.10 small ES closer to 0.30 medium ES closer to 0.50 large

Critical value:
If $n \le 20$, Use Table B.4 (page 245).
If $n > 20$, Compute a *z*-score and use a table with the normal distribution.

Exercise 1:

The following data were collected from a study comparing two methods being used to teach reading recovery in the 4th grade. Method 1 was a pull-out program in which the children were taken out of the classroom for 30 min a day, 4 days a week. Method 2 was a small group program in which children were taught in groups of four or five for 45 min a day in the classroom, 4 days a week. The students were tested using a reading comprehension test after 4 weeks of the program. The test results are shown in the table below.

Method 1	48	40	39	50	41	38	53
Method 2	14	18	20	10	12	102	17

1. Using Mann – Whitney U – test:

H_0: There is no significant different between the two methods.
H_A: There is a significant different between the two methods.

Rank	Score	Sample
1	10	Method 2
2	12	Method 2
3	14	Method 2
4	17	Method 2
5	18	Method 2
6	20	Method 2
7	38	Method 1
8	39	Method 1
9	40	Method 1
10	41	Method 1
11	48	Method 1
12	50	Method 1
13	53	Method 1
14	102	Method 2

For method 1:

$n_1 = 7$

$\sum R_1 = 7 + 8 + 9 + 10 + 11 + 12 + 13 = 70$

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - \sum R_1$$
$$= 7 \times 7 + \frac{7(7+1)}{2} - 70 = 7$$

For method 2:

$n_2 = 7$

$\sum R_2 = 1 + 2 + 3 + 4 + 5 + 14 = 35$

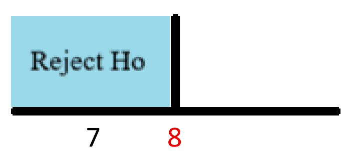
$$U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - \sum R_2$$
$$= 7 \times 7 + \frac{7(7+1)}{2} - 35 = 42$$

$$\boxed{U = \min(U_1, U_2) = \min(7, 42) = 7}$$

To find the critical value for rejection, we use Table B.4 (p.245)

α	m	n = 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0.025	1																				
	2																				
	3																				
	4				0																
	5			0	1	2															
	6			1	2	3	5														
	7			1	3	5	6	8													
	8		0	2	4	6	8	10	13												
	9		0	2	4	7	10	12	15	17											
	10		0	3	5	8	11	14	17	20	23										
	11		0	3	6	9	13	16	19	23	26	30									
	12		1	4	7	11	14	18	22	26	29	33	37								
	13		1	4	8	12	16	20	24	28	33	37	41	45							
	14		1	5	9	13	17	22	26	31	36	40	45	50	55						
	15		1	5	10	14	19	24	29	34	39	44	49	54	59	64					
	16		1	6	11	15	21	26	31	37	42	47	53	59	64	70	75				
	17		2	6	11	17	22	28	34	39	45	51	57	63	69	75	81	87			
	18		2	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99		
	19		2	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	
	20		2	8	14	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127

$U = 7 < 8$, Reject H_0
(There is a significant different between the two methods).



Using R

```
method1 = c(48, 40, 39, 50, 41, 38, 53)
method2 = c(14, 18, 20, 10, 12, 102, 17)
wilcox.test(method1,method2, alternative = "two.sided")
                    Wilcoxon rank sum exact test

data:  method1 and method2
W = 42, p-value = 0.02622   ←——Reject H0
alternative hypothesis: true location shift is not equal to 0
```

Finding 95% confidence interval for the difference between two location parameters

				1st Sample (X_i)			
2nd Sample (Y_j)	38	39	40	41	48	50	53
10	28	29	30	31	38	40	43
12	26	27	28	29	36	38	41
14	24	25	26	27	34	36	39
17	21	22	23	24	31	33	36
18	20	21	22	23	30	32	35
20	18	19	20	21	28	30	33
102	-64	-63	-62	-61	-54	-52	-49

$D_{ij} = X_i - Y_j$

Lower limit is the 9th value from the bottom is 19
Upper limit is the 9th value from the top is 36

95% confidence interval is $(19\,,36)$

Using R

```
method1 = c(48, 40, 39, 50, 41, 38, 53)
method2 = c(14, 18, 20, 10, 12, 102, 17)
wilcox.test(method1,method2, alternative = "two.sided",conf.int = TRUE,conf.level = 0.95)
                    Wilcoxon rank sum exact test

data:  method1 and method2
W = 42, p-value = 0.02622
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 19 36  <------------------------------ lower and upper limit
sample estimates:
difference in location
              27
```

2. Using Kolmogorov – Smirnov two sample test

H_0: There is no significant different between the two methods.
[F(t) = G(t), for every t]
H_A: There is a significant different between the two methods.
[F(t) ≠ G(t) for at least one value of t]

$$F_m(t) = \frac{\text{number of observed X's} \le t}{m}$$
$$G_m(t) = \frac{\text{number of observed Y's} \le t}{n}$$

	Score Z_i	Sample	$F_7(Z_i)$	$G_7(Z_i)$	$\lvert F_7(Z_i) - G_7(Z_i)\rvert$
1	10	Method 2	0/7	1/7	1/7
2	12	Method 2	0/7	2/7	2/7
3	14	Method 2	0/7	3/7	3/7
4	17	Method 2	0/7	4/7	4/7
5	18	Method 2	0/7	5/7	5/7
6	20	Method 2	0/7	6/7	6/7
7	38	Method 1	1/7	6/7	5/7
8	39	Method 1	2/7	6/7	4/7
9	40	Method 1	3/7	6/7	3/7
10	41	Method 1	4/7	6/7	2/7
11	48	Method 1	5/7	6/7	1/7
12	50	Method 1	6/7	6/7	0/7
13	53	Method 1	7/7	6/7	1/7
14	102	Method 2	7/7	7/7	0/7

$$D_{max}(\text{largest divergence}) = 6/7$$

$$Z = D_{max}\sqrt{\frac{mn}{m+n}} = (0.86)\sqrt{\frac{(7)(7)}{7+7}} = 0.86 \times 1.87 = 1.604$$

Since $1 \le Z < 3.2$, we use the p-value formula:

$$Q = e^{-2Z^2} = e^{-2(1.604)^2} = e^{-5.146} = 0.0058$$

$$p = 2(Q - Q^4 + Q^9 - Q^{16})$$
$$= 2(0.0058 - 0.0058^4 + 0.0058^9 - 0.0058^{16}) = 0.012$$

$\text{p} - \text{value} = 0.012 < \alpha = 0.05$

Then we reject H_0 (There is a difference in the percentages)

Using R

```
method1 = c(48, 40, 39, 50, 41, 38, 53)
method2 = c(14, 18, 20, 10, 12, 102, 17)
ks.test(method1, method2, alternative = "two.sided")
            Exact two-sample Kolmogorov-Smirnov test
data:  method1 and method2
D = 0.85714, p-value = 0.008159  <----------Reject H0
alternative hypothesis: two-sided
```

Exercise 2:

Table below shows assessment scores of two different classes who are being taught computer skills using two different methods.

Method 1	Method 2
53	91
41	18
17	14
45	21
44	23
12	99
49	16
50	10

Use two-tailed Mann–Whitney U and Kolmogorov–Smirnov two-sample tests to determine which method was better for teaching reading. Set α = 0.05. Report your findings.

1. Using Mann – Whitney U – test:

H_0: There is no significant difference between the methods.
H_1: There is significant difference between the methods.

Rank	Score	Sample
1	10	Method 2
2	12	Method 1
3	14	Method 2
4	16	Method 2
5	17	Method 1
6	18	Method 2
7	21	Method 2
8	23	Method 2
9	41	Method 1
10	44	Method 1
11	45	Method 1
12	49	Method 1
13	50	Method 1
14	53	Method 1
15	91	Method 2
16	99	Method 2

From sample 1:

$n_1 = 8$

$\sum R_1 = 76$

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - \sum R_1$$
$$= 8 \times 8 + \frac{8(8+1)}{2} - 76 = 20$$

From sample 2:

$n_2 = 8$

$\sum R_2 = 60$

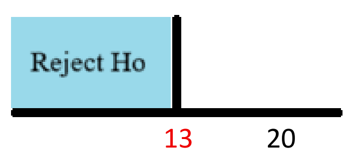
$$U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - \sum R_2$$
$$= 8 \times 8 + \frac{8(8+1)}{2} - 60 = 36$$

$$\boxed{U = \min(U_1, U_2) = \min(20, 36) = 20}$$

To find the critical value for rejection, we use Table B.4 (p.245)

		n																			
α	m	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0.025	1																				
	2																				
	3																				
	4				0																
	5			0	1	2															
	6			1	2	3	5														
	7			1	3	5	6	8													
	8		0	2	4	6	8	10	13												
	9		0	2	4	7	10	12	15	17											
	10		0	3	5	8	11	14	17	20	23										
	11		0	3	6	9	13	16	19	23	26	30									
	12		1	4	7	11	14	18	22	26	29	33	37								
	13		1	4	8	12	16	20	24	28	33	37	41	45							
	14		1	5	9	13	17	22	26	31	36	40	45	50	55						
	15		1	5	10	14	19	24	29	34	39	44	49	54	59	64					
	16		1	6	11	15	21	26	31	37	42	47	53	59	64	70	75				
	17		2	6	11	17	22	28	34	39	45	51	57	63	69	75	81	87			
	18		2	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99		
	19		2	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	
	20		2	8	14	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127

$U = 20 > 13$, Fail to reject H_0
(There is no significant difference between the methods).



Using R

```
method1=c(53,41,17,45,44,12,49,50)
method2=c(91,18,14,21,23,99,19,10)
wilcox.test(method1,method2, alternative = "two.sided")
                Wilcoxon rank sum exact test

data:  method1 and method2
W = 40, p-value = 0.4418  <———— Accept H0
alternative hypothesis: true location shift is not equal to 0
```

2. Using Kolmogorov – Smirnov two sample test

H_0: There is no significant different between the two methods.
[F(t) = G(t), for every t]
H_A: There is a significant different between the two methods.
[F(t) ≠ G(t) for at least one value of t]

$$F_m(t) = \frac{\text{number of observed } X's \le t}{m}$$
$$G_m(t) = \frac{\text{number of observed } Y's \le t}{n}$$

	Score Z_i	Sample	$F_8(Z_i)$	$G_8(Z_i)$	$\lvert F_8(Z_i) - G_8(Z_i)\rvert$
1	10	Method 2	0/8	1/8	1/8
2	12	Method 1	1/8	1/8	0/8
3	14	Method 2	1/8	2/8	1/8
4	16	Method 2	1/8	3/8	2/8
5	17	Method 1	2/8	3/8	1/8
6	18	Method 2	2/8	4/8	2/8
7	21	Method 2	2/8	5/8	3/8
8	23	Method 2	2/8	6/8	4/8
9	41	Method 1	3/8	6/8	3/8
10	44	Method 1	4/8	6/8	2/8
11	45	Method 1	5/8	6/8	1/8
12	49	Method 1	6/8	6/8	0/8
13	50	Method 1	7/8	6/8	1/8
14	53	Method 1	8/8	6/8	2/8
15	91	Method 2	8/8	7/8	1/8
16	99	Method 2	8/8	8/8	0/8

$$D_{max}(\text{largest divergence}) = 4/8$$

$$Z = D_{max}\sqrt{\frac{mn}{m+n}} = (0.5)\sqrt{\frac{(8)(8)}{8+8}} = 0.5 \times 2 = 1$$

Since $1 < Z < 3.2$, we use the p-value formula:

$$Q = e^{-2Z^2} = e^{-2(1)^2} = e^{-2} = 0.135335$$
$$p = 2(Q - Q^4 + Q^9 - Q^{16})$$
$$= 2(0.1353 - 0.1353^4 + 0.1353^9 - 0.1353^{16}) = 0.2699$$

$p - \text{value} = 0.2699 > \alpha = 0.05$ then we fail to reject H_0

(There is no significant different between the two methods.)

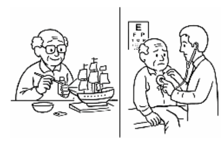
Using R

```
method1 = c(48, 40, 39, 50, 41, 38, 53)
method2 = c(14, 18, 20, 10, 12, 102, 17)
ks.test(method1, method2, alternative = "two.sided")
Exact two-sample Kolmogorov-Smirnov test
data:  method1 and method2
D = 0.5, p-value = 0.2827 ← Accept H0
alternative hypothesis: two-sided
```

Exercise 3:

A research study was conducted to see if an active involvement in a hobby (هواية) had a positive effect on the health of a person who retires after age 65. The data in Table below describe the health (number of doctor visits in 1 year) for participants who are involved in a hobby almost daily and those who are not.

No hobby	Hobby
12	9
15	5
8	10
11	3
9	4
17	2



Use one-tailed Mann–Whitney U and Kolmogorov–Smirnov two-sample tests to determine whether the hobby tends to reduce the need for doctor visits. Set $\alpha = 0.05$. Report your findings.

1. Using Mann – Whitney U – test:

H_0: Having a hobby does not significantly reduce doctor visits. (No hoppy $\leq$ hobby)
H_1: Having a hobby reduces significantly doctor visits. (No hoppy > hobby)

Using R

```
nh=c(12,15,8,11,9,17)
h=c(9,5,10,3,4,2)
wilcox.test(nh,h, alternative = "greater")

        Wilcoxon rank sum test with continuity correction

data:  nh and h
W = 32.5, p-value = 0.01236 <----------- Reject H0

alternative hypothesis: true location shift is greater than 0
```

The results from the Mann-Whitney U test suggest that active involvement in a hobby significantly reduce the number of doctor visits.

2. Using Kolmogorov – Smirnov two sample test:

H_0: Having a hobby does not significantly reduce doctor visits. (No hoppy $\leq$ hobby)
H_1: Having a hobby reduces significantly doctor visits. (No hoppy > hobby)

Using R

```
No_hobby=c(12,15,8,11,9,17)
Hobby=c(9,5,10,3,4,2)
ks.test(No_hobby,Hobby, alternative = "greater")

        Exact two-sample Kolmogorov-Smirnov test

data:  No_hobby and Hobby
D^+ = 0, p-value = 1        Accept H0
alternative hypothesis: the CDF of x lies above that of y
```

The results from the Kolmogorov–Smirnov two-sample test suggest that active involvement in a hobby does not significantly reduce the number of doctor visits.

Exercise *:

Method 1

48	40	39	50	41
38	71	30	15	33
47	51	60	59	58
42	11	46	36	27
93	72	57	45	53

Method 2

14	18	20	10	12
102	21	19	100	23
16	82	13	25	24
97	28	9	34	52
70	22	26	8	17

For each of the following questions (1-4), determine which would be the simplest type of statistical analysis that would be appropriate to use. Use each type of analysis only once.

(A) Paired t test	(B) Two sample t-test	(C) ANOVA
(D) Kruskal-Wallis	(E) Wilcoxon Rank-Sum Test	

Compare the average number of hours per week spent on Facebook for Freshmen, Sophomore, Juniors and Seniors at UF, based on a random sample of 100 students.	(C) ANOVA
Compare the average number of hours per week spent on Facebook during the first week in April and the first week in May (finals week) for random students at UF, measured on the same 100 students.	(A) Paired t test
Compare the distribution of the number of hours per week spent on Facebook for male and female students at UF, based on a random sample of 10 students. There was an outlier in one of the groups.	(E) Wilcoxon Rank-Sum Test
Compare the average number of hours per week spent on Facebook for male and female students at UF, based on a random sample of 100 students.	(B) Two sample t-test

Exercise *:

The following data were obtained from a reading-level test for 1st-grade children. Compare the performance gains of the two different methods for teaching reading. Two different classes being taught a basic mathematics skills using two different methods.

Gain score (Method 1)	16	13	16	16	13	9	12	12	20	17
Gain score (Method 1)	11	2	10	4	9	8	5	6	4	16

Use two-tailed Mann–Whitney U and Kolmogorov–Smirnov two-sample tests to determine which method was better for teaching reading. Set $\alpha = 0.05$.

[1] The hypothesis associated with this test ……………………………
[2] The calculated value of the test statistic is ………………………...
[3] The critical value ……………...

The Mann-Whitney U test is preferred to a t-test when ……………………………

A	Data are paired	B	Sample sizes are small
C	Sample are dependent	D	The assumption of normality is not met

Chapter 5

THE FRIEDMAN TEST

The Friedman test:

For comparing more than two samples that are dependent, or related.

H_0: There is no significant different between the all groups.
H_A: At least one of the groups is different.

Parametric equivalent: repeated measured ANOVA.

Test statistics For data without ties	$F_r = \left[\frac{12}{nk(k+1)} \sum_{i=1}^{k} R_i^2\right] - 3n(k+1)$
Test statistics For data with ties	$F_r = \frac{n(k-1)\left[\sum_{i=1}^{k} \frac{R_i^2}{n} - C_F\right]}{\sum r_{ij}^2 - C_F}$ Where, $C_F = \frac{1}{4} nk(k+1)^2$
Degree of freedom	$df = k - 1$
Adjusted level of risk	$\alpha_B = \frac{\alpha}{k}$

Critical value:

If n and k exist in Friedman critical value Table , Use Table B.5 (pages 247 and 248).

If not, Use χ^2 Table B.2 (page 243).

<u>Exercise 1 (small data without ties):</u>

A manager has seven employees who are often late to work. She wants to reduce their tardiness (التأخر), so she tries two different strategies. In the first month, she deducts \$10 from an employee's paycheck for each day the employee is late. In the second month, she increases the deduction to \$20 for each day the employee is late.
The table shows how many times each employee was late:

- Before any strategy was used (baseline),
- After the \$10 deduction month,
- After the \$20 deduction month.

The goal is to compare the results and see which strategy was more effective in reducing tardiness.

Employee	Monthly tardiness		
	Baseline	Month 1	Month 2
1	16	13	12
2	10	5	2
3	7	8	9
4	13	11	5
5	17	2	6
6	10	7	9
7	11	6	7

Use the Friedman test to test whether the paycheck deductions reduced tardiness

$H_0: \theta_B = \theta_{M1} = \theta_{M2}$
H_A: One or both strategies will reduce employee tardiness.

The rank table (by row):

Employee	Rank of monthly tardiness		
	Baseline	Month 1	Month 2
1	3	2	1
2	3	2	1
3	1	2	3
4	3	2	1
5	3	1	2
6	3	1	2
7	3	1	2
Total	$R_1 = 19$	$R_2 = 11$	$R_3 = 12$

$$F_r = \left[\frac{12}{nk(k+1)}\sum_{i=1}^{k} R_i^2\right] - 3n(k+1)$$
$$= \left[\frac{12}{7(3)(3+1)}(626)\right] - 3(7)(3+1)$$
$$= 5.429$$

$$\sum_{I=1}^{k} R_i^2 = 19^2 + 11^2 + 12^2$$
$$= 626$$

Table B.5 (pages 247 and 248).

k = 3

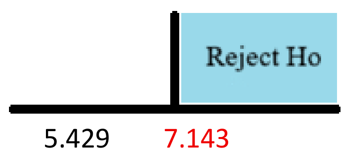
n = 7

TABLE B.5 Critical Values for the Friedman Test Statistic F_r.

k	N	$\alpha \le 0.10$	$\alpha \le 0.05$	$\alpha \le 0.025$	$\alpha \le 0.01$
3	3	6.000	6.000		
	4	6.000	6.500	8.000	8.000
	5	5.200	6.400	7.600	8.400
	6	5.333	7.000	8.333	9.000
	7	5.429	7.143	7.714	8.857
	8	5.250	6.250	7.750	9.000
	9	5.556	6.222	8.000	8.667
	10	5.000	6.200	7.800	9.600
	11	4.909	6.545	7.818	9.455
	12	5.167	6.500	8.000	9.500
	13	4.769	6.000	7.538	9.385
	14	5.143	6.143	7.429	9.000
	15	4.933	6.400	7.600	8.933
4	2	6.000	6.000		

5.429 < 7.143 Fail to reject H_0

(The paycheck deductions did not significantly reduced tardiness)



Using R

```
baseline = c(16,10,7,13,17,10,11)
month1   = c(13,5,8,11,2,7,6)
month2   = c(12,2,9,5,6,9,7)
data = cbind(baseline, month1, month2)
friedman.test(data)

        Friedman rank sum test

data:  data
Friedman chi-squared = 5.4286, df = 2    p-value = 0.06625
```

F_r degree of freedom k − 1

Accept H_0

Exercise 2 (small data with ties):

The manager was not successful in reducing tardiness by deducting money from employees' paychecks. So, she decided to try a different method. This time, she rewarded employees for arriving on time.
In the first month, employees received a \$10 bonus for each day they arrived on time. In the second month, employees received a \$20 bonus for each day they arrived on time. She wanted to see if giving bonuses would improve employee timeliness.

Employee	Monthly tardiness		
	Baseline	Month 1	Month 2
1	16	17	11
2	10	5	2
3	7	8	0
4	13	9	5
5	17	2	2
6	10	10	9
7	11	6	5

The table shows how many times each employee was late during a month. The baseline shows their tardiness before using any strategy. Month 1 shows their tardiness after receiving a \$10 bonus for arriving on time. Month 2 shows their tardiness after receiving a \$20 bonus for arriving on time.
We want to find out if either bonus strategy reduced employee tardiness.

$H_0: \theta_B = \theta_{M1} = \theta_{M2}$
H_A: H_A: One or both strategies will reduce employee tardiness.

The rank table (by row):

Employee	Ranks of monthly tardiness		
	Baseline	Month 1	Month 2
1	2	3	1
2	3	2	1
3	2	3	1
4	3	2	1
5	3	1.5	1.5
6	2.5	2.5	1
7	3	2	1
Total	18.5	16	7.5

Ranks^2 of monthly tardiness		
Baseline	Month 1	Month 2
4	9	1
9	4	1
4	9	1
9	4	1
9	2.25	2.25
6.25	6.25	1
9	4	1
50.25	38.50	8.25

$$F_r = \frac{n(k-1)\left[\sum_{i=1}^{k}\frac{R_i^2}{n} - C_F\right]}{\sum r_{ij}^2 - C_F}$$

$$= \frac{7(3-1)(93.5-84)}{97-84} = 10.23$$

$$\sum r_{ij}^2 = 50.25 + 38.50 + 8.25 = 97$$

$$\sum_{i=1}^{k}\frac{R_i^2}{n} = \frac{18.5^2}{7} + \frac{16^2}{7} + \frac{7.5^2}{7} = 93.5$$

$$C_F = \frac{1}{4}nk(k+1)^2$$

$$= \frac{1}{4}(7)(3)(3+1)^2 = 84$$

Table B.5 (pages 247 and 248).

k = 3

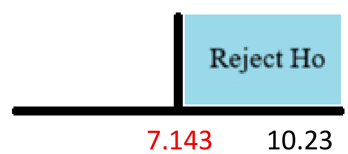
n = 7

TABLE B.5 Critical Values for the Friedman Test Statistic F_r.

k	N	$\alpha \le 0.10$	$\alpha \le 0.05$	$\alpha \le 0.025$	$\alpha \le 0.01$
3	3	6.000	6.000		
	4	6.000	6.500	8.000	8.000
	5	5.200	6.400	7.600	8.400
	6	5.333	7.000	8.333	9.000
	7	5.429	$\boxed{7.143}$	7.714	8.857
	8	5.250	6.250	7.750	9.000
	9	5.556	6.222	8.000	8.667
	10	5.000	6.200	7.800	9.600
	11	4.909	6.545	7.818	9.455
	12	5.167	6.500	8.000	9.500
	13	4.769	6.000	7.538	9.385
	14	5.143	6.143	7.429	9.000
	15	4.933	6.400	7.600	8.933
4	2	6.000	6.000		

$10.23 > 7.143$ Reject H_0

(The bonus strategies led to a statistically significant reduction in tardiness)



Using R

```
baseline = c(16,10,7,13,17,10,11)
Month1 = c(17,5,8,9,2,10,6)
Month2 = c(11,2,0,5,2,9,5)
data = cbind(baseline,Month1,Month2)
friedman.test(data)

        Friedman rank sum test
                                   F_r      degree of freedom k − 1
data:  data
Friedman chi-squared = 10.231, df = 2, p-value = 0.006004  ←— Accept H_0
```

For identifying the difference between groups:
Sample Contrasts, or Post Hoc Tests.
Using Wilcoxon signed rank test for each pair, with type I error = $\alpha_B = \frac{\alpha}{k} = \frac{0.05}{2} = 0.025$

*Note: Usually we should have $\frac{3(3-1)}{2} = 3$ comparisons, but since we will not compare Month 1 with Month 2 we will have only 2 comparisons.

Condition Comparison	Wilcoxon T statistic	Rank sum difference	One-tailed significance
Baseline-Month 1	3.0	18.0-3.0=15.0	0.057
Baseline-Month2	0.0	28.0-0.0	0.009

1- Testing Baseline with Month 1: p-value = $0.057 \nless \alpha_B = 0.025$
(fail to reject H_o, there is no significant difference).

2- Testing Baseline with Month 2: p-value $= 0.009 < \alpha_B = 0.025$
(reject H_o, there is a significant difference).

Therefore, the data indicate that the $20 bonus reduces tardiness while the $10 bonus does not.

<u>Exercise 3:</u>

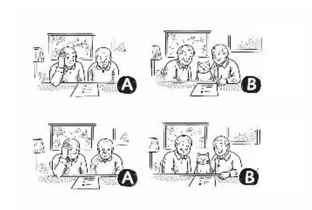
A graduate student conducted a small pilot study to examine whether having an animal affects the quality of life of elderly men in a nursing home (دار المسنين). Ten male participants took part in the study.

The study followed an ABAB design:

- Week A: no cat present
- Week B: cat present

At the end of each week, participants completed a 20-point survey measuring life satisfaction. The table shows the survey scores for four weeks.

Participants	Week 1	Week 2	Week 3	Week 4
1	7	6	8	9
2	9	8	10	7
3	15	18	16	17
4	7	6	8	9
5	7	8	10	11
6	10	14	13	11
7	12	19	11	13
8	7	4	2	5
9	8	7	9	5
10	12	16	14	15



Use the Friedman test to determine whether there are significant differences among the weeks. Since this is a pilot study, use $\boldsymbol{\alpha = 0.10}$. If a significant difference is found, use Wilcoxon signed-rank tests to identify which weeks differ. Apply the Bonferroni correction to control for Type I error.

$H_0: \theta_{W1} = \theta_{W2} = \theta_{W3} = \theta_{W4}$

H_A: At least one group are singificantly different.

The rank table (by row):

Participants	Week 1	Week 2	Week 3	Week 4
1	2	1	3	4
2	3	2	4	1
3	1	4	2	3
4	2	1	3	4
5	1	2	3	4
6	1	4	3	2
7	2	4	1	3
8	4	2	1	3
9	3	2	4	1
10	1	4	2	3
Total	$R_1 = 20$	$R_2 = 26$	$R_3 = 26$	$R_4 = 28$

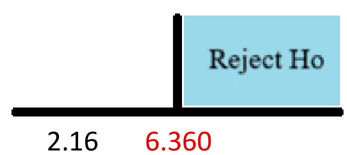
$$F_r = \left[\frac{12}{nk(k+1)}\sum_{i=1}^{k} R_i^2\right] - 3n(k+1)$$
$$= \left[\frac{12}{10(4)(4+1)}(2536)\right] - 3(7)(3+1)$$
$$= 2.16$$

$$\sum_{I=1}^{k} R_i^2 = 20^2 + 26^2 + 26^2 + 28^2 = 2536$$

Table B.5 (pages 247 and 248).

	k	N	$\alpha \le 0.10$	$\alpha \le 0.05$	$\alpha \le 0.025$	$\alpha \le 0.01$
		⋮	⋮	⋮	⋮	⋮
$k = 4$	4	2	6.000	6.000		
		3	6.600	7.400	8.200	9.000
		4	6.300	7.800	8.400	9.600
		5	6.360	7.800	8.760	9.960
		6	6.400	7.600	8.800	10.200
		7	6.429	7.800	9.000	10.371
		8	6.300	7.650	9.000	10.500
		9	6.467	7.800	9.133	10.867
$n = 10$		10	$\boxed{6.360}$	7.800	9.120	10.800
		11	6.382	7.909	9.327	11.073
		12	6.400	7.900	9.200	11.100
		13	6.415	7.985	7.369	11.123
		14	6.343	7.886	9.343	11.143
		15	6.440	8.040	9.400	11.240

$2.16 < 6.360$, Fail to reject H_0
(No significant different between groups)



Using R

```
w1 = c(7,9,15,7,7,10,12,7,8,12)
w2 = c(6,8,18,6,8,14,19,4,7,16)
w3 = c(8,10,16,8,10,13,11,2,9,14)
w4 = c(9,7,17,9,11,11,13,5,5,15)

data = cbind(w1,w2,w3,w4)
friedman.test(data)

        Friedman rank sum test
                                   F_r    degree of freedom k − 1
data:  data
Friedman chi-squared = 2.16, df = 3, p-value = 0.5399  ←  Accept H0
```

Exercise 4:

A physical education teacher wanted to see if a strength program improved performance. He tested 12 male students and measured how many curl-ups each could do in one minute. He recorded their performance:

- Before the program (baseline),
- After one month,
- After two months.

Participant	Number of curl ups in one minute		
	Baseline	Month 1	Month 2
1	66	67	69
2	49	50	56
3	51	52	49
4	65	65	69
5	42	43	46
6	38	39	40
7	33	31	39
8	41	41	44
9	46	47	48
10	45	46	46
11	36	33	34
12	51	55	67

Use the Friedman test with $\alpha = 0.05$ to check if there are differences across the three time periods. Since improvement is expected, if the result is significant, use Wilcoxon signed-rank tests to find which time periods differ. Apply the Bonferroni correction to control Type I error. Report your results.

$H_0: \theta_B = \theta_{M1} = \theta_{M2}$
H_A: There are significant differences across the three times in the number of curl-ups.

The rank table (by row):

Participant	Ranks of curl ups in one minute		
	Baseline	Month 1	Month 2
1	1	2	3
2	1	2	3
3	2	3	1
4	2	1	3
5	1	3	2
6	1	2	3
7	2	1	3
8	1.5	1.5	3
9	1	2	3
10	1	2.5	2.5
11	3	1	2
12	1	2	3
Total	17.5	23	31.5

$\Rightarrow$

Ranks of curl ups in one minute		
Baseline	Month 1	Month 2
1	4	9
1	4	9
4	9	1
4	1	9
1	9	4
1	4	9
4	1	9
2.25	2.25	9
1	4	9
1	6.25	6.25
9	1	4
1	4	9
30.25	49.5	87.25

$$F_r = \frac{n(k-1)\left[\sum_{i=1}^{k}\frac{R_i^2}{n} - C_F\right]}{\sum r_{ij}^2 - C_F} = \frac{12(3-1)(152.29-144)}{167-144} = 8.6522$$

$$\sum r_{ij}^2 = 30.25 + 49.5 + 87.25 = 167$$

$$\sum_{i=1}^{k}\frac{R_i^2}{n} = \frac{17.5^2}{12} + \frac{23^2}{12} + \frac{31.5^2}{12} = 152.29$$

$$C_F = \frac{1}{4}nk(k+1)^2 = \frac{1}{4}(12)(3)(3+1)^2 = 144$$

Table B.5 (pages 247 and 248).

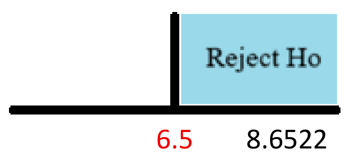
TABLE B.5 Critical Values for the Friedman Test Statistic F_r.

$k = 3$, $n = 12$

k	N	$\alpha \le 0.10$	$\alpha \le 0.05$	$\alpha \le 0.025$	$\alpha \le 0.01$
3	3	6.000	6.000		
	4	6.000	6.500	8.000	8.000
	5	5.200	6.400	7.600	8.400
	6	5.333	7.000	8.333	9.000
	7	5.429	7.143	7.714	8.857
	8	5.250	6.250	7.750	9.000
	9	5.556	6.222	8.000	8.667
	10	5.000	6.200	7.800	9.600
	11	4.909	6.545	7.818	9.455
	12	5.167	6.500	8.000	9.500
	13	4.769	6.000	7.538	9.385
	14	5.143	6.143	7.429	9.000
	15	4.933	6.400	7.600	8.933
4	2	6.000	6.000		

$8.6522 > 6.5$, Reject H_0

(There are significant differences across the three times in the number of curl-ups)



Using R

```
#rm(list=ls())   #for clear the environment

b    = c(66,49,51,65,42,38,33,41,46,45,36,51)
m1  = c(67,50,52,43,65,39,31,41,47,46,33,55)
m2  = c(69,56,49,69,46,40,39,44,48,46,34,67)

data = cbind(b, m1, m2)
friedman.test(data)

          Friedman rank sum test

data:  data
Friedman chi-squared = 8.6522, df = 2, p-value = 0.01322
```

F_r degree of freedom $k-1$

Reject H_0

For identifying the difference between groups:
Sample Contrasts, or Post Hoc Tests.
Using Wilcoxon signed rank test for each pair, with type I error

$$\alpha_B = \frac{\alpha}{k} = \frac{0.05}{3} = 0.0167$$

Using R

```
wilcox.test(b, m1, paired = TRUE, alternative = "two.sided")
          Wilcoxon signed rank test with
          continuity correction

data:  b and m1
V = 25, p-value = 0.4973 <---- Accept Ho
alternative hypothesis: true location shift is not equal to 0

wilcox.test(b, m2, paired = TRUE, alternative = "two.sided")
          Wilcoxon signed rank test with
          continuity correction

data:  b and m2
V = 7, p-value = 0.01304 <---- Reject Ho
alternative hypothesis: true location shift is not equal to 0

wilcox.test(m1, m2, paired = TRUE, alternative = "two.sided")
          Wilcoxon signed rank test with
          continuity correction

data:  m1 and m2
V = 15.5, p-value = 0.1297 <---- Accept Ho
alternative hypothesis: true location shift is not equal to 0
```

1- Testing Baseline with Month 1:
   p-value $= 0.4973 > 0.0167$
   (Fail to reject H_0, there is no significant difference)

2- Testing Baseline with Month 2:
   p-value $= 0.01304 < 0.0167$
   (Reject H_0, there is a significant difference).

3- Testing Month 1 with Month 2:
   p-value $= 0.1297 > 0.0167$
   (Fail to reject H_0, there is no significant difference).

Chapter 6

THE KRUSKAL – WALLIS H - TEST

The Kruskal-Wallis H-test:

For comparing more than two independent samples.

H_0: There is no significant different between the all groups.

H_A: At least one of the groups is different.

Parametric equivalent: One-way ANOVA.

Test statistics For small data $(n < 20)$	$H = \frac{12}{N(N+1)} \sum \frac{R_i^2}{n_i} - 3\,(N+1)$
Test statistics For small data $(n < 20)$ with ties	In case of ties: Find (Corrected H) because of ties: $\text{Corrected H} = \frac{\text{Original H}}{C_H}$ $C_H = 1 - \frac{\sum(T^3 - T)}{N^3 - N}$; $T = df = k - 1$
Degree of freedom	$df = k - 1$
Adjusted level of risk	$\alpha_B = \frac{\alpha}{k}$

Critical value:

If n and k exist in ‹Kruskal − Wallis critical value Table , Use Table B.6 (page 248).

Exercise 1:

Researchers wanted to examine whether social interaction (التفاعل الإجتماعي) is related to self-confidence (الثقة بالنفس). They classified 17 adults into three groups based on their level of social interaction:

- High: very social and talks to many people
- Medium: interacts with others but sometimes isolates
- Low: mostly isolated and interacts very little

After grouping them, participants completed a 25-point self-confidence scale. The table shows their scores, where 25 indicates high self-confidence.

	High	Medium	Low
1	21	19	7
2	23	5	8
3	18	10	15
4	12	11	3
5	19	9	6
6	20	-	4

Use Kruskal-Wallis H-test to determine if there is a difference between any of the three groups.

$H_0: \theta_H = \theta_M = \theta_L$ (No significant different between the three groups)
H_A: There are significant different between the three groups.

	High	Rank H	Medium	Rank M	Low	Rank L
1	21	**16**	19	**13.5**	7	**5**
2	23	**17**	5	**3**	8	**6**
3	18	**12**	10	**8**	15	**11**
4	12	**10**	11	**9**	3	**1**
5	19	**13.5**	9	**7**	6	**4**
6	20	**15**	-	**-**	4	**2**
Total		**83.5**		**40.5**		**29**

$n_H = 6$
$n_M = 5$
$n_L = 6$

$N = 17$

$$H = \frac{12}{N(N+1)} \sum \frac{R_i^2}{n_i} - 3\,(N+1)$$
$$= \frac{12}{17(17+1)}(1630.26) - 3\,(17+1)$$
$$= 9.93$$

$$\sum \frac{R_i^2}{n_i} = \frac{83.5^2}{6} + \frac{40.5^2}{5} + \frac{29^2}{6} = 1630.26$$

Finding (Corrected H) because of ties:

$$\text{Corrected H} = \frac{\text{Original H}}{C_H}$$
$$= \frac{9.93}{0.9988} = 9.94$$

$$C_H = 1 - \frac{\sum(T^3 - T)}{N^3 - N}$$
$$= 1 - \frac{(2^3 - 2)}{17^3 - 17} = 0.9988$$

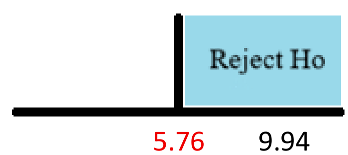
Table B.6 pages 248

(The Critical Values for the Kruskal–Wallis H-Test Statistic, $k = 3$).

n_1	n_2	n_3	$\alpha \le 0.10$	$\alpha \le 0.05$	$\alpha \le 0.01$
6	2	1	4.200000	4.822222	–
6	2	2	4.436364	5.345455	6.654545
6	3	1	3.909091	4.854545	6.581818
6	3	2	4.681818	5.348485	6.969697
6	3	3	4.538462	5.615385	7.192308
6	4	1	4.037879	4.946970	7.083333
6	4	2	4.493590	5.262821	7.339744
6	4	3	4.604396	5.604396	7.467033
6	4	4	4.523810	5.666667	7.795238
6	5	1	4.128205	4.989744	7.182051
6	5	2	4.595604	5.318681	7.375824
6	5	3	4.535238	5.601905	7.590476
6	5	4	4.522500	5.660833	7.935833
6	5	5	4.547059	5.698529	8.027941
6	6	1	4.000000	4.857143	7.065934
6	6	2	4.438095	5.409524	7.466667
6	6	3	4.558333	5.625000	7.725000
6	6	4	4.547794	5.724265	8.000000
6	6	5	4.542484	$\boxed{5.764706}$	8.118954
6	6	6	4.538012	5.719298	8.222222
7	1	1	4.266667	–	–

$9.94 > 5.76$, Reject H_0

(There are significant different between the three groups)



Using R

```
h = c(21,23,18,12,19,20)
m = c(19,5,10,11,9)
l = c(7,8,15,3,6,4)

v = c(h, m, l)
g = factor(c(rep("High"     , length(h)),
             rep("Medium",length(m)),
             rep("Low"      , length(l))))
kruskal.test(v ~ g)

        Kruskal-Wallis rank sum test

data:  v by g
Kruskal-Wallis chi-squared = 9.9439, df = 2, p-value = 0.00693
```

H degree of freedom $k-1$

Reject H_0

For statistically identifying difference between groups

Sample Contrasts, or Post Hoc Tests.

Using Mann - Whitney U-test for each pair, with type I error

$$\alpha_B = \frac{\alpha}{k} = \frac{0.05}{3} = 0.0167$$

Using R

```
wilcox.test(h,m, alternative = "two.sided")
        Wilcoxon rank sum test with continuity correction
data:  h and m
W = 27.5, p-value = 0.0281  ←— Accept Ho
alternative hypothesis: true location shift is not equal to 0

wilcox.test(h,l, alternative = "two.sided")
        Wilcoxon rank sum exact test
data:  h and l
W = 35, p-value = 0.004329  ←— Reject Ho
alternative hypothesis: true location shift is not equal to 0

wilcox.test(m,l, alternative = "two.sided")
        Wilcoxon rank sum exact test
data:  m and l
W = 23, p-value = 0.1775  ←— Accept Ho
alternative hypothesis: true location shift is not equal to 0
```

1- Testing “High” with “Medium”:
   p-value $= 0.0281 > 0.0167$
   (Fail to reject H_0, there is no significant difference)

2- Testing “High” with “Low”:
   p-value $= 0.004329 < 0.0167$
   (Reject H_0, there is a significant difference).

3- Testing “Medium” with “Low”:
   p-value $= 0.1775 > 0.0167$
   (Fail to reject H_0, there is no significant difference).

Exercise 2:

A researcher studied 15 participants to see if exercise increased strength. The participants were divided into three groups, and each group received a different treatment. Their strength gains were measured and ranked. The table shows these rankings.

	I	II	III
1	7	13	12
2	2	1	5
3	4	7	16
4	11	8	9
5	15	3	14

Use Kruskal-Wallis H-test to determine if there is a difference between any of the three groups. If a significant difference exists, use two tailed Mann–Whitney *U*-tests or two-sample Kolmogorov–Smirnov tests to identify which groups are significantly different. Use the Bonferroni procedure to limit the type I error rate.

H_0: $\theta_I = \theta_{II} = \theta_{III}$ (No significant different between the three treatments)
H_A: There are significant different between the three treatments.

	I	Rank I	II	Rank II	III	Rank III
1	7	**6.5**	13	**12**	12	**11**
2	2	**2**	1	**1**	5	**5**
3	4	**4**	7	**6.5**	16	**15**
4	11	**10**	8	**8**	9	**9**
5	15	**14**	3	**3**	14	**13**
Total		**36.5**		**30.5**		**53**

$n_I = 5$
$n_{II} = 5$
$n_{III} = 5$
$N = 15$

$$H = \frac{12}{N(N+1)} \sum \frac{R_i^2}{n_i} - 3\,(N+1)$$
$$= \frac{12}{15(15+1)}(1014.3) - 3\,(15+1)$$
$$= 2.715$$

$$\sum \frac{R_i^2}{n_i} = \frac{36.5^2}{5} + \frac{30.5^2}{5} + \frac{53^2}{5} = 1014.3$$

Finding (Corrected H) because of ties:

$$\text{Corrected H} = \frac{\text{Original H}}{C_H}$$
$$= \frac{2.715}{0.9982} = 2.72$$

$$C_H = 1 - \frac{\sum(T^3 - T)}{N^3 - N}$$
$$= 1 - \frac{(2^3 - 2)}{15^3 - 15} = 0.9982$$

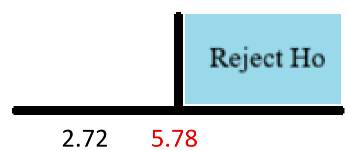
Table B.6 pages 248

(The Critical Values for the Kruskal–Wallis H-Test Statistic, $k = 3$).

n_1	n_2	n_3	$\alpha \le 0.10$	$\alpha \le 0.05$	$\alpha \le 0.01$
⋮	⋮	⋮	⋮	⋮	⋮
4	4	3	4.545455	5.598485	7.143939
4	4	4	4.653846	5.692308	7.653846
5	2	1	4.200000	5.000000	–
5	2	2	4.373333	5.160000	6.533333
5	3	1	4.017778	4.871111	–
5	3	2	4.650909	5.250909	6.821818
5	3	3	4.533333	5.648485	7.078788
5	4	1	3.987273	4.985455	6.954545
5	4	2	4.540909	5.272727	7.204545
5	4	3	4.548718	5.656410	7.444872
5	4	4	4.668132	5.657143	7.760440
5	5	1	4.109091	5.127273	7.309091
5	5	2	4.623077	5.338462	7.338462
5	5	3	4.545055	5.626374	7.578022
5	5	4	4.522857	5.665714	7.791429
5	5	5	4.560000	$\boxed{5.780000}$	8.000000

2.72 > 5.78 , Accept H_0

(There are no significant different between the three treatments)



Using R

```
t1= c(7,2,4,11,15)
t2= c(13,1,7,8,3)
t3= c(12,5,16,9,14)

v = c(t1, t2, t3)
g = factor(c(rep("t1", length(t1)),
             rep("t2", length(t2)),
             rep("t3", length(t3))))
kruskal.test(v ~ g)
          Kruskal-Wallis rank sum test
                                      H          degree of freedom k − 1
data:  v by g
Kruskal-Wallis chi-squared = 2.7199, df = 2, p-value = 0.2567  ←  Reject H0
```

Exercise 3:

A researcher investigated how physical attraction influences the perception among others of a person's effectiveness with difficult tasks. The photographs of 24 people were shown to a focus group. The group was asked to classify the photos into three groups: very attractive, average, and very unattractive. Then, the group ranked the photographs according to their impression of how capable they were of solving difficult problems. Table below shows the classification and rankings of the people in the photos (1 = most effective, 24 = least effective).

	Very attractive	Average	Very unattractive
1	1	3	11
2	2	4	15
3	5	8	16
4	6	9	18
5	7	12	20
6	10	14	21
7	12	19	23
8	17	22	24

Use Kruskal-Wallis H-test to determine if there is a difference between any of the three groups. If a significant difference exists, use two tailed Mann–Whitney *U*-tests or two-sample Kolmogorov–Smirnov tests to identify which groups are significantly different. Use the Bonferroni procedure to limit the type I error rate.

Very attractive: VA
Average: A
Very unattractive: VU

$H_0: \theta_{VA} = \theta_A = \theta_{VU}$ (No significant different between the three groups)
H_A: There are significant different between the three groups.

	VA	Rank VA	A	Rank A	VU	Rank VU
1	1	**1**	3	**3**	11	**11**
2	2	**2**	4	**4**	15	**15**
3	5	**5**	8	**8**	16	**16**
4	6	**6**	9	**9**	18	**18**
5	7	**7**	13	**13**	20	**20**
6	10	**10**	14	**14**	21	**21**
7	12	**12**	19	**19**	23	**23**
8	17	**17**	22	**22**	24	**24**
Total		**60**		**92**		**148**

$n_{VA} = 8$
$n_A = 8$
$n_{VU} = 8$

$N = 24$

$$H = \frac{12}{N(N+1)} \sum \frac{R_i^2}{n_i} - 3\,(N+1)$$

$$= \frac{12}{24(24+1)}(4246) - 3\,(24+1)$$

$$= 9.92$$

$$\sum \frac{R_i^2}{n_i} = \frac{60^2}{8} + \frac{92^2}{8} + \frac{148^2}{8} = 4246$$

No need to calculate C_H (there is no ties)

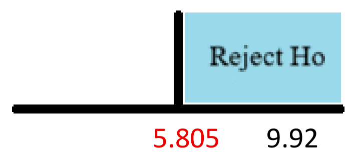
Table B.6 pages 248

(The Critical Values for the Kruskal–Wallis H-Test Statistic, $k = 3$).

n_1	n_2	n_3	$\alpha \le 0.10$	$\alpha \le 0.05$	$\alpha \le 0.01$
8	8	6	4.572134	5.778656	8.366601
8	8	7	4.570652	5.791149	8.418866
8	8	8	4.595000	$\boxed{5.805000}$	8.465000
9	1	1	4.545455	–	–
9	2	1	3.905983	4.841880	6.346154
9	2	2	4.483516	5.260073	6.897436

$9.92 > 5.805$, Reject H_0

(There are significant different between the three groups)



Using R

```
va= c(1,2,5,6,7,10,12,17)
  a= c(3,4,8,9,13,14,19,22)
vu= c(11,15,16,18,20,21,23,24)

  v = c(va,a,vu)
  g = factor(c(rep("va", length(va)),
               rep("a", length(a)),
              rep("vu", length(vu))))

kruskal.test(v ~ g)
          Kruskal-Wallis rank sum test

data:  v by g
Kruskal-Wallis chi-squared = 9.92, df = 2,  p-value = 0.007013
```

H — degree of freedom $k-1$ — Reject H_0

For statistically identifying difference between groups

Sample Contrasts, or Post Hoc Tests.

Using Mann - Whitney U-test for each pair, with type I error

$$\alpha_B = \frac{\alpha}{k} = \frac{0.05}{3} = 0.0167$$

Using R

```
wilcox.test(va,a , alternative = "two.sided")

            Wilcoxon rank sum exact test
data:  va and a
W = 20, p-value = 0.2345  <----  Accept Ho
alternative hypothesis: true location shift is not equal to 0

wilcox.test(va,vu, alternative = "two.sided")

            Wilcoxon rank sum exact test
data:  va and vu
W = 4, p-value = 0.001865  <----  Reject Ho
alternative hypothesis: true location shift is not equal to 0

wilcox.test(a ,vu, alternative = "two.sided")

            Wilcoxon rank sum exact test
data:  a and vu
W = 12, p-value = 0.03792  <----  Accept Ho
alternative hypothesis: true location shift is not equal to 0
```

1- Testing “Very attractive” with “Average”:
   p-value $= 0.2345 > 0.0167$
   (Fail to reject H_0, there is no significant difference)

2- Testing “Very attractive “Very unattractive”:
   p-value $= 0.001865 < 0.0167$
   (Reject H_0, there is a significant difference).

3- Testing “Average” with “Very unattractive”:
   p-value $= 0.13792 > 0.0167$
   (Fail to reject H_0, there is no significant difference).

Chapter 7

SPEARMAN RANK – ORDER
POINT BISERIAL
AND BISERIAL CORRELATIONS

Spearman rank-order correlation:

For comparing two rank-ordered variables

$H_0: \rho_s = 0$ (No significant correlation between the two groups)
$H_A: \rho_s \neq 0$ (There is significant correlation between the two groups)

Parametric equivalent: Pearson product–moment correlation

Test statistics For small data $(n < 20)$	$r_s = 1 - \frac{6\sum D_i^2}{n(n^2-1)}$
Test statistics For small data $(n < 20)$ with ties	$r_s = \frac{(n^3-n)-6\sum D_i^2-\frac{(T_x+T_y)}{2}}{\sqrt{(n^3-n)^2-(T_x+T_y)(n^3-n)+T_xT_y}}$ $T_x = \sum(t_i^3 - t_i)$ for group 1 $T_y = \sum(t_i^3 - t_i)$ for group 2
Degree of freedom	$df = n - 1$

Critical value:
If n exist in ‹Spearman test critical value Table , Use Table B.7 (page 255).

Exercise 1: (Without ties)

Eight men took part in a study to see if going to the gym more often affects resting heart rate (معدل ضربات القلب اثناء الراحة). The idea is that people who visit the gym more frequently will have a lower heart rate. The table shows how many times each man went to the gym during the month. It also shows the average heart rate measured at the end of each week during the last three weeks of the month.

Participant	Number of visits	Mean heart rate
1	5	100
2	12	89
3	7	78
4	14	66
5	2	77
6	8	103
7	15	67
8	17	63

$H_0: \rho_s = 0$ (No significant correlation between number of visits and mean heart rate)
$H_A: \rho_s \neq 0$ (There is significant correlation between number of visits and mean heart rate)

Participant	Number of visits		Mean heart rate		D
	Data	Rank	Data	Rank	$\text{Rank}_V - \text{Rank}_H$
1	5	2	100	7	-5
2	12	5	89	6	-1
3	7	3	78	5	-2
4	14	6	66	2	4
5	2	1	77	4	-3
6	8	4	103	8	-4
7	15	7	67	3	4
8	17	8	63	1	7

$n = 8$
$df = n - 1 = 6$

$$r_s = 1 - \frac{6\sum D_i^2}{n(n^2-1)}$$

$$= 1 - \frac{6\,(136)}{8(8^2-1)} = -0.619$$

$$\sum D_i^2 = (-5)^2 + (-1)^2 + (-2)^2 + (4)^2 + (-3)^2 + (-4)^2 + (4)^2 + (7)^2 = 136$$

Table B.7 pages 255

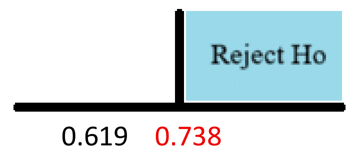
n = 8

n	$\alpha_{\text{two-tailed}} \le 0.10$ $\alpha_{\text{one-tailed}} \le 0.05$	$\alpha_{\text{two-tailed}} \le 0.05$ $\alpha_{\text{one-tailed}} \le 0.025$	$\alpha_{\text{two-tailed}} \le 0.02$ $\alpha_{\text{one-tailed}} \le 0.01$	$\alpha_{\text{two-tailed}} \le 0.01$ $\alpha_{\text{one-tailed}} \le 0.005$
4	1.000			
5	0.900	1.000	1.000	
6	0.829	0.886	0.943	1.000
7	0.714	0.786	0.893	0.929
8	0.643	0.738	0.833	0.881
9	0.600	0.700	0.783	0.833
10	0.564	0.648	0.745	0.794

$$r_s = -0.619$$
$$|r_s| = 0.619$$

$0.619 < 0.738$, Accept H_0

(No significant correlation between number of visits and mean heart rate)



Using R

```
v =c(5,12,7,14,2,8,15,17)
hr=c(100,89,78,66,77,103,67,63)
cor.test(v, hr, method = "spearman")

            Spearman's rank correlation rho
data:  v and hr
S = 136, p-value = 0.115 <-------- Accept Ho
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
-0.6190476   <-------- r_s
```

Exercise 2: (With ties)

Eight men took part in a study to see if going to the gym more often affects resting heart rate (معدل ضربات القلب اثناء الراحة). The idea is that people who visit the gym more frequently will have a lower heart rate. The table shows how many times each man went to the gym during the month. It also shows the average heart rate measured at the end of each week during the last three weeks of the month.

Participant	Number of visits	Mean heart rate
1	5	96
2	12	63
3	7	78
4	14	66
5	3	79
6	8	95
7	15	67
8	12	64
9	2	99
10	16	62
11	12	65
12	7	76
13	17	61

$H_0: \rho_s = 0$ (No significant correlation between number of visits and mean heart rate)
$H_A: \rho_s \neq 0$ (There is significant correlation between number of visits and mean heart rate)

Participant	Number of visits		Mean heart rate		D
	Data	Rank	Data	Rank	$\text{Rank}_V - \text{Rank}_H$
1	5	3	96	12	-9
2	12	**8**	63	3	5
3	7	**4.5**	78	9	-4.5
4	14	10	66	6	4
5	3	2	79	10	-8
6	8	6	95	11	-5
7	15	11	67	7	4
8	12	8	64	4	4
9	2	1	99	13	-12
10	16	12	62	2	10
11	12	**8**	65	5	3
12	7	**4.5**	76	**8**	-3.5
13	17	13	61	1	12

$n = 13$
$df = n - 1 = 11$

$$r_s = \frac{(n^3-n)-6\sum D_i^2-\frac{(T_x+T_y)}{2}}{\sqrt{(n^3-n)^2-(T_x+T_y)(n^3-n)+T_xT_y}}$$

$$= \frac{(13^3-13)-6\,(672.5)-\frac{(30+0)}{2}}{\sqrt{(13^3-13)^2-(30+0)(13^3-13)+(30)(0)}} = -0.86$$

$$\sum D_i^2 = (-9)^2+(5)^2+(-4.5)^2+(4)^2+(-8)^2+(-5)^2+(4)^2+(4)^2+(-12)^2+(10)^2+(3)^2+(-3.5)^2+(12)^2 = \mathbf{672.5}$$

$$T_x = \sum(t_i^3-t_i) = (2^3-2)+(3^3-3) = \mathbf{30}$$
$$T_y = \mathbf{0}$$

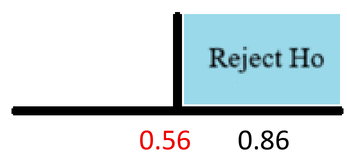
Table B.7 pages 255

	n	$\alpha_{\text{two-tailed}} \le 0.10$ $\alpha_{\text{one-tailed}} \le 0.05$	$\alpha_{\text{two-tailed}} \le 0.05$ $\alpha_{\text{one-tailed}} \le 0.025$	$\alpha_{\text{two-tailed}} \le 0.02$ $\alpha_{\text{one-tailed}} \le 0.01$	$\alpha_{\text{two-tailed}} \le 0.01$ $\alpha_{\text{one-tailed}} \le 0.005$
	4	1.000			
	5	0.900	1.000	1.000	
	6	0.829	0.886	0.943	1.000
	7	0.714	0.786	0.893	0.929
	8	0.643	0.738	0.833	0.881
	9	0.600	0.700	0.783	0.833
	10	0.564	0.648	0.745	0.794
	11	0.536	0.618	0.709	0.755
	12	0.503	0.587	0.671	0.727
n = 13	13	0.484	$\boxed{0.560}$	0.648	0.703
	14	0.464	0.538	0.622	0.675
	15	0.443	0.521	0.604	0.654

$$r_s \;= -0.86$$
$$|r_s| = 0.86$$

$0.86 > 0.56$, Reject H_0

(There is significant correlation between number of visits and mean heart rate)



Using R

```
v=c(5,12,7,14,3,8,15,12,2,16,12,7,17)
 hr=c(96,63,78,66,79,95,67,64,99,62,65,76,61)
cor.test(v, hr, method = "spearman")

            Spearman's rank correlation rho
data:  v and hr
S = 677.16, p-value = 0.0001609   <———— Reject Ho
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
-0.8603249  <———— r_s
```

Point - Biserial correlation:
For comparing two variables when one variable is discrete dichotomous

$H_0: \rho_{pb} = 0$ (No significant correlation between the two groups)
$H_A: \rho_{pb} \neq 0$ (There is significant correlation between the two groups)

Parametric equivalent: Pearson product–moment correlation

Test statistics For small data $(n < 20)$	$r_{pb} = \frac{\bar{x}_p - \bar{x}_q}{s}\sqrt{P_p P_q}$
Degree of freedom	$df = n - 1$

Critical value:
If n exist in ‹Person product − Momemt Table , Use Table B.8 (page 257).

Exercise 1:

A psychology researcher (باحث في علم النفس) wanted to see if males and females differ in their ability to notice and remember visual details (قوة الملاحظة البصرية).17 participants were placed alone in a room with different objects for 10 minutes. After that, they completed a 30-question test about details in the room. The table shows each participant's gender and test score. The researcher wants to know whether there is a relationship between gender and test score.

Participant	Gender	Score
1	M	7
2	M	19
3	M	8
4	M	10
5	M	7
6	M	15
7	M	6
8	M	13
9	F	14
10	F	11
11	F	18
12	F	23
13	F	17
14	F	20
15	F	14
16	F	24
17	F	22

Since gender has two categories (male/female) and test scores are numerical,
a point-biserial correlation will be used.

H_0: $\rho_{pb} = 0$ (No significant correlation between gender and test score).
H_A: $\rho_{pb} \neq 0$ (There is significant correlation between gender and test score).

	Sample size	Summation	Mean	Proportion
Male	$n_p = 8$	$\sum x_p = 85$	$\bar{x}_p = \frac{85}{8} = 10.625$	$P_p = \frac{n_p}{n} = \frac{8}{17} = 0.47$
Female	$n_q = 9$	$\sum x_q = 163$	$\bar{x}_q = \frac{163}{9} = 18.11$	$P_q = \frac{n_q}{n} = \frac{9}{17} = 0.53$
All	$n = 17$	$\sum x_i = 248$	$\bar{x} = \frac{248}{17} = 14.588$	-

The standard deviation of test score: $S = \sqrt{\frac{\sum(x_i-\bar{x})^2}{n-1}} = \sqrt{\frac{\sum(x_i-14.588)^2}{17-1}} = 5.86$

$$r_{pb} = \frac{\bar{x}_p-\bar{x}_q}{S}\sqrt{P_pP_q}$$

$$= \frac{10.625-18.11}{5.86}\sqrt{(0.47)(0.53)} = -0.637$$

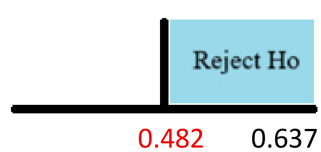
Table B.8 pages 257

$$\begin{aligned} df &= n-2 \\ &= 17-2 \\ &= 15 \end{aligned}$$

df	$\alpha_{\text{two-tailed}} \leq 0.10$ $\alpha_{\text{one-tailed}} \leq 0.05$	$\alpha_{\text{two-tailed}} \leq 0.05$ $\alpha_{\text{one-tailed}} \leq 0.025$	$\alpha_{\text{two-tailed}} \leq 0.025$ $\alpha_{\text{one-tailed}} \leq 0.0125$	$\alpha_{\text{two-tailed}} \leq 0.01$ $\alpha_{\text{one-tailed}} \leq 0.005$
⋮	⋮	⋮	⋮	⋮
12	0.458	0.532	0.594	0.661
13	0.441	0.514	0.575	0.641
14	0.426	0.497	0.557	0.623
15	0.412	$\boxed{0.482}$	0.541	0.606
16	0.400	0.468	0.526	0.590

$r_{pb} = -0.637$
$|r_{pb}| = 0.637$

$0.637 > 0.482$, Reject H_0
(There is significant correlation between gender and test score)



Exercise 2:

A researcher wanted to see if poverty (الفقر) is related to self-esteem (الثقة بالنفس). Participants were divided into two groups based on income: below poverty and above poverty. Each participant completed a 20-question self-esteem survey. The table shows their scores. Is there a relationship between poverty and test self-esteem.

Participant	Poverty	Score
1	Above	15
2	Above	19
3	Above	15
4	Above	20
5	Above	7
6	Above	12
7	Above	3
8	Above	15
9	Below	9
10	Below	5
11	Below	13
12	Below	13
13	Below	11
14	Below	10
15	Below	8
16	Below	9
17	Below	10
18	Below	17

Since gender has two categories (Poverty) and test scores are numerical,
a point-biserial correlation will be used.

$H_0: \rho_{pb} = 0$ (No significant correlation between poverty and test self-esteem.).
$H_A: \rho_{pb} \neq 0$ (There is significant correlation between poverty and test self-esteem.).

	Sample size	Summation	Mean	Proportion
Above	$n_p = 8$	$\sum x_p = 106$	$\bar{x}_p = \frac{106}{8} = 13.25$	$P_p = \frac{n_p}{n} = \frac{8}{18} = 0.44$
Below	$n_q = 10$	$\sum x_q = 105$	$\bar{x}_q = \frac{105}{10} = 10.5$	$P_q = \frac{n_q}{n} = \frac{10}{18} = 0.56$
All	$n = 18$	$\sum x_i = 211$	$\bar{x} = \frac{211}{18} = 11.72$	-

The standard deviation of test score: $S = \sqrt{\frac{\sum(x_i-\bar{x})^2}{n-1}} = \sqrt{\frac{\sum(x_i-11.72)^2}{18-1}} = 4.625$

$$r_{pb} = \frac{\bar{x}_p - \bar{x}_q}{S}\sqrt{P_p P_q}$$

$$= \frac{13.25-10.5}{4.625}\sqrt{(0.44)(0.56)} = 0.295$$

Table B.8 pages 257

df	$\alpha_{\text{two-tailed}} \leq 0.10$ $\alpha_{\text{one-tailed}} \leq 0.05$	$\alpha_{\text{two-tailed}} \leq 0.05$ $\alpha_{\text{one-tailed}} \leq 0.025$	$\alpha_{\text{two-tailed}} \leq 0.025$ $\alpha_{\text{one-tailed}} \leq 0.0125$	$\alpha_{\text{two-tailed}} \leq 0.01$ $\alpha_{\text{one-tailed}} \leq 0.005$
⋮	⋮	⋮	⋮	⋮
12	0.458	0.532	0.594	0.661
13	0.441	0.514	0.575	0.641
14	0.426	0.497	0.557	0.623
15	0.412	0.482	0.541	0.606
16	0.400	$\boxed{0.468}$	0.526	0.590

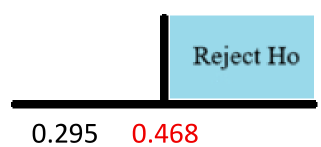
$$\begin{aligned} df &= n - 2 \\ &= 18 - 2 \\ &= 16 \end{aligned}$$

$r_{pb} = 0.295$

$|r_{pb}| = 0.295$

$0.295 > 0.468$, Accept H_0

(No significant correlation between poverty and test self-esteem.)



Biserial correlation:

For comparing two variables when one variable is continuous dichotomous

$H_0: \rho_b = 0$
$H_A: \rho_b \neq 0$

Parametric equivalent: Pearson product–moment correlation

Test statistics For small data ($n < 20$)	$r_{pb} = \frac{\bar{x}_p - \bar{x}_q}{S}\sqrt{P_p P_q}$ $r_b = \left[\frac{\bar{x}_p - \bar{x}_q}{S_x}\right]\frac{P_p P_q}{y}$ $r_b = r_{pb}\frac{\sqrt{P_p P_q}}{y}$
Degree of freedom	$df = n - 1$

Exercise 1:

A graduate department at a university wanted to examine whether students' GPAs could predict their performance on the comprehensive exam (الإختبار الشامل) required for graduation. The exam is graded on a pass/fail basis. Last year, students took the exam. The table shows each student's GPA and exam result.

Participant	Exam performance	GPA
1	F	3.5
2	F	3.4
3	F	3.3
4	F	3.2
5	F	3.6
6	P	4.0
7	P	3.6
8	P	4.0
9	P	4.0
10	P	3.8
11	P	3.9
12	P	3.9
13	P	4.0
14	P	3.8
15	P	3.5
16	P	3.6

The goal is to determine whether there is a relationship between GPA and exam outcome, and whether GPA can be used to predict whether a student will pass or fail.

Exam performance is a continuous dichotomous variable and GPA is an interval scale variable. Therefore, we will use a **biserial correlation**.

	Sample size	Summation	Mean	Proportion
Fail	$n_p = 5$	$\sum x_p = 17$	$\bar{x}_p = \frac{17}{5} = 3.4$	$P_p = \frac{n_p}{n} = \frac{5}{16} = 0.3125$
Pass	$n_q = 11$	$\sum x_q = 42.1$	$\bar{x}_q = \frac{42.1}{11} = 3.83$	$P_q = \frac{n_q}{n} = \frac{11}{16} = 0.6875$
All	$n = 16$	$\sum x_i = 211$	$\bar{x} = \frac{59.1}{16} = 3.69$	-

The standard deviation of test score: $S = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{\sum(x_i - 3.69)^2}{16-1}} = 0.267$

$$y = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

$$r_{pb} = \frac{\bar{x}_p - \bar{x}_q}{S} \sqrt{P_p P_q}$$

$$= \frac{13.25 - 10.5}{4.625} \sqrt{(0.44)(0.56)} = 0.295$$

Table B.8 pages 257

df	$\alpha_{\text{two-tailed}} \le 0.10$ $\alpha_{\text{one-tailed}} \le 0.05$	$\alpha_{\text{two-tailed}} \le 0.05$ $\alpha_{\text{one-tailed}} \le 0.025$	$\alpha_{\text{two-tailed}} \le 0.025$ $\alpha_{\text{one-tailed}} \le 0.0125$	$\alpha_{\text{two-tailed}} \le 0.01$ $\alpha_{\text{one-tailed}} \le 0.005$
⋮	⋮	⋮	⋮	⋮
12	0.458	0.532	0.594	0.661
13	0.441	0.514	0.575	0.641
14	0.426	0.497	0.557	0.623
15	0.412	0.482	0.541	0.606
16	0.400	$\boxed{0.468}$	0.526	0.590

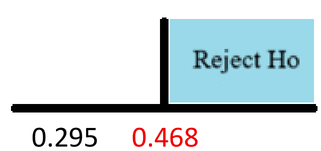
$$\begin{aligned} df &= n - 2 \\ &= 18 - 2 \\ &= 16 \end{aligned}$$

$r_{pb} = 0.295$

$|r_{pb}| = 0.295$

$0.295 > 0.468$, Accept H_0

(No significant correlation between poverty and test self-esteem.)



Chapter 8

CHI – SQUARE
AND FISHER EAXCT TEST

Chapter 9

THE RUN TEST