CHAPTER 7
NONPARAMETRIC TESTS

7.1 Introduction

The statistical tests in chapters 3,4and 6 used the assumption that the population(or populations) is normally distributed or at least approximately normally distributed and also, assumed the sample size small or large. These tests called parametric or classical statistical methods(tests). Methods such as the Chi-square goodness-of-fit test in chapter 5, that do not assume any specific distribution or the sample size, are known as nonparametric methods.

7.2 sign test (quantitative or qualitative variable)

We use this test if the variable is quantitative or ordinal qualitative where no information about the population that the sample taken from. The test will be for the median denoted by η.

In this test we have to account the number of values that is less than η_o and denoted by (T-) , also, the number of values that is greater than η_o and denoted by (T+). We will ignore the values that equals η_o and then the sample size become smaller than n.

7.2.1 Test steps:

1) Data: η_o , n , x_1 , x_2 ,…, x_n , α

2) statistical hypothesis:

H_o : $\eta = \eta_o$

H_1 : $\eta > \eta_o$ or $\eta \neq \eta_o$ or $\eta < \eta_o$

3)the test statistic:

As before, it depends on H_1

1 – H_1: $\eta < \eta_o$ (the statistic = T +)

2 – H_1: $\eta > \eta_o$ (the statistic = T -)

3 – H_1: $\eta \neq \eta_o$ (the statistic = T = min (T^- , T^+))

4) Rejection region:

1 –if H_1 : $\eta < \eta_o$, we reject H_o <u>if p-value = Bin(T^+ ,n ,0.5)< α</u>

2 –if H_1 : $\eta > \eta_o$, we reject H_o <u>if p-value = Bin(T^- ,n ,0.5)< α</u>

3 –if H_a : $\eta \neq \eta_o$, we reject H_o <u>if p-value = 2Bin(T ,n ,0.5)< α</u>

Where:

P-value can be obtained from computer results,

<u>Bin</u> means the binomial distribution ,

<u>Bin(T ,n ,0.5)</u> means that the table value for binomial distribution at 0.5, since n is the new sample size after deleted the values equals to η_o and T can be T^+ or T- or T = min (T^- , T^+)

	T
n	Bin. Valu

Ex(1)

A random sample of grades in 105 statistic course taught in KSU for girls who took the course, we have:

D+ B+ F A A C B+ A C+ B C+ D A D+ B

Test whether the median grade is different from C. use $\alpha = 0.1$.

Solu.

1) Data: $n = 15$, $\alpha = 0.1$, $\eta_o = C$

When we delete the elements that $\eta_o = C$ then, n=14

2) The hypothesis:

$H_o: \eta = C$

$H_1: \eta \neq C$ $T^- = 4$ $T^+ = 10$

3) The statistic is: $T = \min(T^-, T^+)$

4) reject H_o, if $2Bin(4, 14, 0.5) < \alpha$

$$2\,Bin\,(4{,}14{,}0.5) = 2(0.09) = 0.18 > \alpha = 0.1$$

5) Decision:

Accept H_o, that is the median grades is C

7.3 The test for independent samples

When we tested two populations means, we assuming normality or having large samples. Also, we had two cases ; one for independent samples and the other for dependent(paired) samples.

Similarly, when we do not assume normality, we must test these two cases. In this section we will consider independent samples without the assumptions of normality or large samples.

7.3.1 The Rank sum test

We assume independent samples with size $n_1 and\ n_2$ from two populations with a variable measured on at least an ordinal scale. An appropriate test for this case is known as **rank sum** (also called **Mann-Whitney** rank sum test).

To find the rank sum , we combine the data from the both samples and rank them $1, 2, \ldots, (n_1 + n_2)$ in an increasing order and if we have repeated values, we calculate the mean rank of these values.

7.3.2 Test steps:

1) Data: $n_1,\ n_2, \alpha$

2) Statistical hypothesis:

H$_o$: $\eta_1 = \eta_2$

H$_1$: $\eta_1 > \eta_2$ or $\eta_1 < \eta_2$ or $\eta_1 \neq \eta_2$

3)The test statistic:

$$W_s = W_1 - \frac{n_1(n_1+1)}{2}$$

Where $W_1 = the\ sum\ of\ the\ ranks\ for\ the\ first\ sample.$

Note that: the first sample is the sample with smallest size.

4) Rejection region:

As before, it depends on H$_1$

1 –if H$_1$: $\eta_1 < \eta_2$, we reject H$_o$ if

2 –if H$_1$: $\eta_1 > \eta_2$, we reject H$_o$ if

$$W_s > W_{1-\alpha, n_1, n_2}$$

3 –if H_a : $\eta_1 \neq \eta_2$, we reject H_o if

$$W_s > W_{1-\frac{\alpha}{2},n_1,n_2} \quad or \quad W_s < W_{\frac{\alpha}{2},n_1,n_2}$$

Note that

The gives us the tabulated values for small , thus:

$$W_{1-\alpha,n_1,n_2} = n_1 n_2 - W_{\alpha,n_1,n_2}$$

And

$$W_{1-\frac{\alpha}{2},n_1,n_2} = n_1 n_2 - W_{\frac{\alpha}{2},n_1,n_2}$$

Ex(2)

If we have two independent samples A and B with the following data:

A	7	5.5	2.5	6	3	4.5	7.5	8.5
B	3	7	5	1.5	2	3	6.5	9

Can we conclude that the median of sample A is greater than the median of sample B at $\alpha = 0.05$.

Solu.

1) Data: $n_1 = 8,\ n_2 = 8, \alpha = 0.05$
2) Statistical hypothesis:

H_o : $\eta_1 = \eta_2$ against H_1 : $\eta_1 > \eta_2$

we will rank the two samples as one group, to find W_1.

A	7	5.5	2.5	6	3	4.5	7.5	8.5
rank	12.5	9	3	10	5	7	14	15
B	3	7	5	1.5	2	3	6.5	9
rank	5	12.5	8	1	2	5	11	16

$$W_1 = 12.5 + 9 + 3 + 10 + 5 + 7 + 14 + 15 = 75.5$$

3)The test statistic:

$$W_s = W_1 - \frac{n_1(n_1+1)}{2} = 75.5 - \frac{8(9)}{2} = 39.5$$

4) Rejection region:

Since, $H_1 : \eta_1 > \eta_2$,then, we reject H_o if

$$W_s > W_{1-\alpha,n_1,n_2}$$

$$W_{1-\alpha,n_1,n_2} = W_{0.95,8,8} = 8(8) - W_{0.05,8,8} = 64 - 16 = 48$$

Thus , $39.5 \ngtr 48 \Longrightarrow we\ accept\ H_0$

Which means that the median of the sample A equal the median of the sample B.