

# Nonparametric Statistics

---

## 14.1 Using the Binomial Table

In this chapter, we will survey several methods of inference from Nonparametric Statistics. These methods will introduce us to several new tables designed to provide critical values. However, the first method we will discuss, called the Sign Test, will rely on a table we used in Statistics I called the binomial table. This section will acquaint us with this useful table of probabilities.

Recall that the binomial probability distribution is used to give us the probability that we have  $X$  successes out of  $n$  trials of a binomial experiment.

Where

$n$  = number of trials,

$x$  = number of successes,

$p$  = probability of a success, and

$q = 1 - p$

And the binomial experiment has the following five properties:

1. Experiment consists of  $n$  identical trials  
**For example:** Flip a coin 3 times
2. There are only two possible outcomes for each trial (success or failure)  
**For example:** Outcomes are Heads or Tails
3. The probability of a success remains the same from trial to trial  
**For example:**  $P(\text{Heads}) = .5$ ;  $P(\text{Tails}) = 1 - .5 = .5$
4. The trials are independent  
**For example:** H on flip  $i$  doesn't change  $P(H)$  on flip  $i + 1$
5. The random variable is the number of successes in  $n$  trials  
**For example:** Let  $x$  = number of heads in three flips



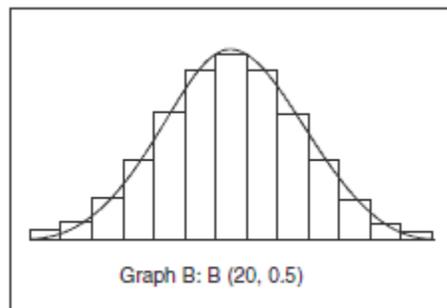
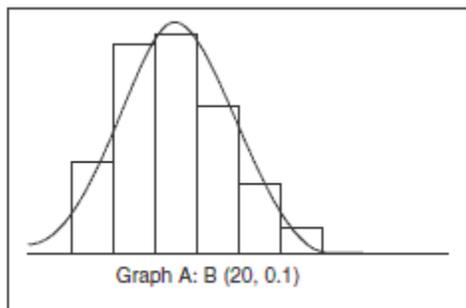
The binomial probability formula can be used without much difficulty to find the probability that we have exactly  $X$  successes out of  $n$  trials, but it is harder to find the probability that we have less than  $X$  successes, more than  $X$  successes, at least  $X$  successes, or at most  $X$  successes. Under certain conditions, we can turn to the binomial table for help.

**Example 181.5** Using the binomial table, find the probability that out of 8 flips of a fair coin we have at least 5 tails.

## 14.2 Normal as Approximation to Binomial

There are times when we are dealing with a binomial experiment consisting of a large number of trials ( $n > 25$ ). In these cases, the binomial table will not be helpful since they generally do not have probabilities for binomial experiments dealing with sample sizes greater than 25.

For binomial problems with large sample sizes, we can apply the Normal Approximation to the Binomial Distribution. The normal approximation method involves using a bell curve with a mean of  $\mu = n(p)$  and standard deviation of  $\sigma = \sqrt{npq}$  in place of the exact binomial distribution the problem involves. From the images below, you can see that when the  $p$  for the binomial random variable is 0.5 (the image on the right side) the approximation is really quite good. However, it is not as good when  $p = 0.10$  (the image on the left side).

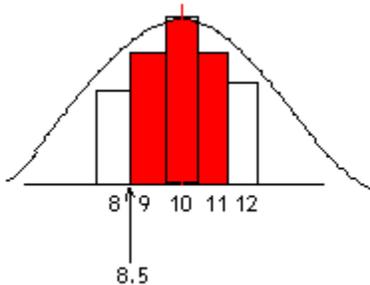


The further  $p$  is away from 0.5, the less well the approximation performs. Once  $p$  gets too far from 0.5, it is best not to use the approximation any longer. In this chapter, we will only be using this technique in cases where the probability of success is 0.5, so we will not need to worry about the approximation not being a good fit.

The approximation can be improved by using a technique called **Continuity Correction**. This technique involves adding or subtracting 0.5 to your  $x$ -value in order to correct for the difference between discrete distributions (like the binomial distribution) and continuous distributions (like the normal distribution).

The issue has to do with the fact that in discrete distributions, the probability rectangle for an  $X$  value like  $x = 9$  actually extends from 8.5 until 9.5 on the  $x$ -axis. This is not the case for continuous random variables. When dealing with probabilities like  $P(9 < x < 11)$  for continuous distributions,  $X = 9$  and  $x = 11$  are infinitesimally small points on the number line which bound our probability area. To compensate for

this difference, we will calculate the probability  $x$  is between 9 and 11 by actually calculating the probability that  $x$  is between 8.5 and 11.5 when using the normal approximation to find binomial probabilities like  $P(9 < X < 11)$ .



**Example 181.7** Use the normal approximation to find the probability that when flipping a fair coin 30 times, we have at most 12 heads.

## 14.3 Ranking Data

### Single Population Inferences Using Nonparametric Methods

The remaining sections of this chapter deal with a branch of statistics called nonparametrics. Most of the techniques we have learned thus far have been a part of parametric statistics. **Parametric tests** have requirements about the nature or shape of the populations involved. **Nonparametric tests** on the other hand do not require that samples come from populations with normal distributions or have any other particular distribution. Consequently, nonparametric tests are called distribution-free tests.

#### **There are several advantages to nonparametric tests:**

1. Nonparametric methods can be applied to a wide variety of situations because they do not have the more rigid requirements of the corresponding parametric methods. In particular, nonparametric methods do not require normally distributed populations.
2. Unlike parametric methods, nonparametric methods can often be applied to categorical data, such as the genders of survey respondents.
3. Nonparametric methods usually involve simpler computations than the corresponding parametric methods and are therefore easier to understand and apply.

**However, nonparametric tests are not perfect; they also have disadvantages:**

1. Nonparametric methods tend to waste information because exact numerical data are often reduced to a qualitative form.
2. Nonparametric tests are not as efficient (powerful) as parametric tests, so with a nonparametric test we generally need stronger evidence (such as a larger sample or greater differences) before we reject a null hypothesis.

Efficiency: Comparison of Parametric and Nonparametric Tests			
Application	Parametric Test	Nonparametric Test	Efficiency Rating of Nonparametric Test with Normal Population
Matched pairs of sample data	$t$ test or $z$ test	Sign test	0.63
		Wilcoxon signed-ranks test	0.95
Two independent samples	$t$ test or $z$ test	Wilcoxon rank-sum test	0.95
Several independent samples	Analysis of variance ( $F$ test)	Kruskal-Wallis test	0.95
Correlation	Linear correlation	Rank correlation test	0.91
Randomness	No parametric test	Runs test	No basis for comparison

There are a couple of techniques that are widely used in nonparametric tests. The following definitions will make things clearer as we move forward:

Data are **sorted** when they are arranged according to some criterion, such as smallest to the largest or best to worst.

A **rank** is a number assigned to an individual sample item according to its order in the sorted list. The first item is assigned a rank of 1, the second is assigned a rank of 2, and so on.

Assigning ranks sounds easy, and for the most part it is, but what happens if two or more values are the same?

### Handling Ties in Ranks

In the case of two or more values that are identical in the data, find the mean of the ranks involved and assign this mean rank to each of the tied items. See below:

Sorted Data	Preliminary Ranking	Rank
4	1	1
5	2	3
5	3	3
5	4	3
10	5	5
11	6	6
12	7	7.5
12	8	7.5

} Mean is 3.  
 } Mean is 7.5.

**Example 181.9** Rank the following set of points scored by the winning team in each of the last ten Super Bowls: 21, 31, 31, 27, 17, 29, 21, 24, 32, 48.



## 14.4 The Sign Test

### Sign Test

Earlier, we learned the z-test and t-test for testing hypotheses about a population mean. If our sample size was large we were able to use the z-test; however, when our sample size was small we used the t-test and the assumption of normality. This section deals with the question, “How can we conduct a test of hypothesis when we have a small sample from a non-normal distribution?”

One possible answer is to use the **Sign Test**. The Sign Test allows us to conduct test of hypotheses about the *central tendency* of a non-normal probability distribution. We use the phrase central tendency here because the sign test provides inferences about the population median  $\eta$  rather than the population mean  $\mu$ . The only requirement that must be met before using the sign test is that the data have been randomly selected from the population.

Let’s demonstrate the Sign Test with an example:

**Example 182:** Customers at a “fast” food establishment want their food fast. To determine the median wait time at Taco Bell in the Miami area, 16 graduate students randomly selected 16 Taco Bell stores. Each student visited a Taco Bell store at 12:15 p.m. and ordered a crunchy beef taco, a burrito supreme, and a soda. The wait time to receive the meal was recorded for each restaurant. At the 5% significance level, test the claim that the median wait time is longer than 20 minutes.



Restaurant	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Wait Time	23	19	15	25	30	22	18	23	24	17	21	22	26	40	20	14

The logic will be pretty simple here. A key concept underlying this use of the **Sign Test** is that if the two sets of data have equal medians, the number of positive signs should be approximately equal to the number of negative signs. If the median is equal to 20 minutes, half of the time customers should wait more than 20 minutes. If we observe that much more than half of our sample waited longer than 20 minutes, we will be inclined to reject the idea that the median is 20 minutes. Let us see how this will work in practice:

**Step 1 Claim:**  $\eta > 20$

**Step 2 Hypotheses:**  
 $H_0 : \eta \leq 20$   
 $H_A : \eta > 20$

**Step 3 Test Stat:**  $S =$  Number of measurements greater than  $\eta_0$   $S = 10$

**Step 4 Determine n:**  $n =$  sample size minus any sample measurements which have a value equal to  $\eta_0$ .  
 $n = 15$

**Step 5 P-value:** To find the p-value go to the binomial table, use  $n$  from step 4,  $p = 0.5$ , and find  $P(X \geq S)$ .  
 $p\text{-value} = P(X \geq S) = P(X \geq 10) = 1 - 0.849 = 0.151$

**Step 6 Initial Conclusion:** Compare your p-value to alpha, if  $p < \alpha$  reject  $H_0$

In our case  $p\text{-value} = 0.151 > \alpha = 0.05$ , so we will fail to reject  $H_0$

**Step 7 Word Conclusion:** The sample data does not support the claim that the median wait time exceeds 20 minutes at the 5% significance level.

**Example 183** This is a Left – tailed example:

At  $\alpha = 0.05$ , test the claim that the length of time it takes me to find parking on campus at FIU comes from a population with a median of less than twelve minutes.

Data in minutes	11.2	9.1	13.5	10.0	11.9	10.7	11.3	11.1
-----------------	------	-----	------	------	------	------	------	------



Steps to solve:

1.  $\eta < 12$
2.  $H_0 : \eta \geq 12$  &  $H_A : \eta < 12$
3.  $S =$  number of measurements less than twelve = 7
4.  $n = 8$
5.  $p\text{-value} = P(X \geq S) = 1 - 0.965 = 0.035$
6. Reject the null
7. The data support the claim that the median is less than 12.

**Example 184** This example illustrates the two-tailed case:

At  $\alpha = 0.05$ , test the claim that the data comes from a population with a median equal to ten.

Data	7	9	3	5	11	12	8	8
------	---	---	---	---	----	----	---	---

Steps to solve:

1.  $\eta = 10$
2.  $H_0 : \eta = 10$  &  $H_A : \eta \neq 10$
3.  $S =$  Larger of  $S_S$  (# of measurements less than  $\eta_0$ ) and  $S_B$  (# of measurements more than  $\eta_0$ )  
 $S_S = 6, S_B = 2$ , so  **$S = 6$**
4.  $n = 8$
5.  $p\text{-value} = 2P(X \geq S) = 2(1 - 0.855) = 0.290$   
 (Note: the times 2 because of the two-tails)
6. Do not reject the null
7. The sample data does not warrant rejection of the claim that the median is ten.

**Large sample case:**

Most binomial tables do not exceed  $n = 25$ , so when we encounter problems that have a sample size greater than 25 we can use a large sample normal approximation. The test stat will become:

$$z = \frac{(S_{\min} + 0.5) - \frac{n}{2}}{\frac{\sqrt{n}}{2}}, \text{ Where } S_{\min} \text{ is the number of times the less frequent sign occurs. This test will}$$

always be left-tailed regardless of the claim involved in the specific problem.

Let's look at an example:

**Example 185:** Use the sign test to test the claim that the median is less than 98.6°F. There are 68 subjects with temperatures below 98.6°F, 23 subjects with temperatures above 98.6°F, and 15 subjects with temperatures equal to 98.6°F.



**Solution:**

**Step 1 and 2:** Since the claim is that the median is less than 98.6°F, the test involves only the left tail.

$H_0$ : Median is equal to 98.6°F.

$H_1$ : Median is less than 98.6°F.

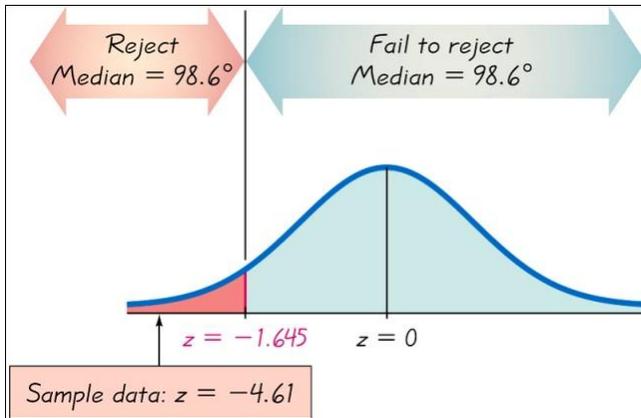
**Step 3:** Discard the 15 values that are tied with 98.6. Use ( - ) to denote the 68 temperatures below 98.6°F, and use ( + ) to denote the 23 temperatures above 98.6°F.

So  $n = 91$  and  $S_{\min} = 23$

**Step 4:** Calculate our test stat  $z = \frac{(23 + 0.5) - \frac{91}{2}}{\frac{\sqrt{91}}{2}} = -4.61$

**Step 5:** We get our critical value using the Z Table to get the critical z value of  $-1.645$ .

**Step 6:** The test statistic of  $z = -4.61$  falls into the critical region.



**Step 7:** Conclusion: We support the claim that the median body temperature of healthy adults is less than  $98.6^\circ\text{F}$ .

## 14.5 Wilcoxon Ranked Sum Test for Independent Samples

### Comparing Two Populations: Independent Samples

In earlier sections, we learned how to make comparisons between the means of two independent populations using either the z-test or t-test. However, if the t-test is not appropriate to compare the two populations that you wish to compare because we cannot meet the normality assumption, we can use the nonparametric test: Wilcoxon Rank Sum Test.

The Wilcoxon Rank Sum Test can be used to test the hypothesis that the probability distributions (and as such the medians) associated with the two populations are equivalent. The procedure uses the ranks of the sample data from two independent populations to test the null hypothesis which will contain the condition that the two independent samples come from populations with equal medians (as in previous tests, the null hypothesis will use either the  $\leq$ ,  $\geq$ , or  $=$  symbol).

The logic of the test is simple. We first rank all of the data from the lowest value ( $=1$ ) to the highest value. If the two populations have the same probability distribution, then the ranks should be randomly mixed between the two populations, and the sum of the ranks (the rank sums) for each sample should be approximately equal. If one distribution is shifted to the right of the other it would have a larger rank sum than the other distribution.

**Example 186:** Researchers measured the bursting strength of two different bottle designs (old vs. new). At the 5% significance level, test the claim that the probability distributions associated with the two bottle designs are equivalent.

Old Design	210	212	211	211	190	213	212	211	164	209
New Design	216	214	162	137	219	218	179	153	152	217



**Solution:**

**Step 1** Claim:  $D_1$  (distribution for pop 1) and  $D_2$  (distribution for pop 2) are identical.

**Step 2**  $H_0$  :  $D_1$  and  $D_2$  are identical,

$H_A$  :  $D_1$  is shifted either to the left or to the right of  $D_2$ .

**Step 3** Rank the data (If ties exist give the tied values the average of the ranks they would have gotten if they were in successive order)

Old	210 9	212 13.5	211 11	211 11	190 7	213 15	212 13.5	211 11	164 5	209 8
New	216 17	214 16	162 4	137 1	219 20	218 19	179 6	153 3	152 2	217 18

**Step 4** Calculate  $T_1$  = sum of the ranks for pop 1 and  $T_2$  = sum of the ranks for pop 2

(As a check make sure  $T_1 + T_2 = \frac{n(n+1)}{2}$ ) If  $n_1 < n_2$ ,  $T_1 = T$  = Test stat, If  $n_1 > n_2$ ,  $T_2 = T$  = Test stat (if both sample sizes are equal use either sum as your test stat).

$T_1 = 104$

**Step 5** The rejection region is determined by looking up alpha in the Wilcoxon Rank-Sum table.

(Reject the null if  $T \leq T_L = 79$  or  $T \geq T_U = 131$ )

**Step 6** Form your initial conclusion

Since  $T_1 = 104$ , neither of the conditions for rejection are met, so we will not reject the null.

**Step 7** Word final conclusion

The sample data does not warrant rejection of the claim that the two distributions are identical.

**Summary of the Wilcoxon Rank Sum Test**

<b>One-tailed test</b>	<b>Two-tailed test</b>
$H_0$ : $D_1$ and $D_2$ are identical	$H_0$ : $D_1$ and $D_2$ are identical
$H_a$ : $D_1$ is shifted right of $D_2$ <b>or</b> [ $H_a$ : $D_1$ is shifted left of $D_2$ ]	$H_a$ : $D_1$ is shifted right or left of $D_2$
<b>Test statistic:</b>	<b>Test statistic:</b>
$T_1$ , if $n_1 < n_2$	$T_1$ , if $n_1 < n_2$
$T_2$ , if $n_1 > n_2$	$T_2$ , if $n_1 > n_2$
Either if $n_1 = n_2$	Either if $n_1 = n_2$
<b>Rejection region:</b>	<b>Rejection region:</b>
$T_1$ : $T_1 \geq T_U$ or $[T_1 \leq T_L]$	$T \leq T_L$ or $T \geq T_U$
$T_2$ : $T_2 \leq T_L$ or $[T_2 \geq T_U]$	

For the Wilcoxon rank sum test there are two assumptions we should note:

1. The two samples are random and independent
2. The data is continuous (this ensures the probability of ties = 0)

Finally, as was the case with the sign test, when we have a large sample for both populations (large here means either  $n_1 > 10$  or  $n_2 > 10$ ) we will use a slightly different approach.

**Wilcoxon Rank Sum Test for Large Samples**

- |  |  |
|--|--|
| <ul style="list-style-type: none"> <li>■ One-tailed test</li> <li><math>H_0</math>: <math>D_1</math> and <math>D_2</math> are identical</li> <li><math>H_a</math>: <math>D_1</math> is shifted right of <math>D_2</math> <b>or</b> <math>H_a</math>:<br/><math>D_1</math> is shifted left of <math>D_2</math></li> <li><b>Rejection region:</b></li> <li><math> z  &gt; z_\alpha</math></li> </ul> | <ul style="list-style-type: none"> <li>■ Two-tailed test</li> <li><math>H_0</math>: <math>D_1</math> and <math>D_2</math> are identical</li> <li><math>H_a</math>: <math>D_1</math> is shifted either right<br/>or left of <math>D_2</math></li> <li><b>Rejection region:</b></li> <li><math> z  &gt; z_{\alpha/2}</math></li> </ul> |
|--|--|

**Test Statistic:** 
$$z = \frac{T_1 - \frac{n_1(n_1 + n_2 + 1)}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

**Example 187:** Use the data below with the Wilcoxon rank-sum test and a 0.05 significance level to test the claim that the median BMI of men is greater than the median BMI of women.

BMI Measurements	
Men	Women
23.8 (11.5)	19.6 (2.5)
23.2 (9)	23.8 (11.5)
24.6 (14)	19.6 (2.5)
26.2 (17)	29.1 (22)
23.5 (10)	25.2 (15.5)
24.5 (13)	21.4 (5)
21.5 (6)	22.0 (7)
31.4 (24)	27.5 (19)
26.4 (18)	33.5 (25)
22.7 (8)	20.6 (4)
27.8 (20)	29.9 (23)
28.1 (21)	17.7 (1)
25.2 (15.5)	
$n_1 = 13$	$n_2 = 12$
$R_1 = 187$	$R_2 = 138$

## 14.6 Wilcoxon Signed-Ranks Test for Paired Difference Experiments

### Paired Difference Experiment

We can use nonparametric methods to analyze data from a paired difference experiment also. The test we will demonstrate is called the Wilcoxon Sign-Rank Test. The Wilcoxon signed-ranks test uses ranks of sample data consisting of matched pairs. This test is used with a null hypothesis that the population of differences from the matched pairs has a median equal to zero. The hypotheses we will use in this test are as follows (note, the one-tailed procedures are given later):

$$H_0 : \eta_d = 0$$

$$H_A : \eta_d \neq 0$$

Consider the example below:

**Example 188:** Use the Wilcoxon Sign-Rank Test to test the claim at the 5% significance level that there is no difference between the yields from regular and kiln dried seed.

### Yields of Corn from Different Seeds

Regular	1903	1935	1910	2496	2108	1963	2060	1444	1612	1316	1511
Kiln	2009	1915	2011	2496	2180	1925	2122	1482	1542	1443	1535

**Step 1:** State your claim.

There is no difference between the yields from regular and kiln dried seed.

**Step 2:** List your hypotheses.

$$H_0 : \eta_d = 0$$

$$H_A : \eta_d \neq 0$$

**Step 3:** Get your ranks—a.) For each pair of data, find the difference (d) by subtracting the second value from the first. Discard any pairs for which d = 0. b.) Take the absolute values of d and rank them from low (= 1) to high. If any differences are tied give the tied differences the average of the ranks they would have if they were in successive order.

Regular	1903	1935	1910	2496	2108	1963	2060	1444	1612	1316	1511
Kiln Dried	2009	1915	2011	2496	2180	1925	2122	1482	1542	1443	1535
<b>Differences</b>	-106	20	-101	0	-72	38	-62	-38	70	-127	-24
<b>Absolute d</b>	106	20	101	X	72	38	62	38	70	127	24
<b>Rank</b>	9	1	8	X	7	3.5	5	3.5	6	10	2

**Step 4:** Add up all the ranks of the negative differences ( $T_-$ ) and do the same for the ranks of the positive differences ( $T_+$ ). Let  $T = \min(T_+, T_-)$

$$T_- = 9 + 8 + 7 + 5 + 3.5 + 10 + 2 = 44.5,$$

$$T_+ = 1 + 3.5 + 6 = 10.5,$$

$$T = 10.5$$

**Step 5:** Let n = number of pairs that have a nonzero difference, then

find your Critical Value  $T_0$  from the Wilcoxon Sign-Rank Test table.

$$N = 10, T_0 = 8$$

**Step 6:** Form your initial conclusion

We reject if our test stat is below our critical value (i.e. – Reject when  $T \leq T_0$ )

$T = 10.5 > T_0 = 8$ , so do not reject the null

**Step 7:** Word your final conclusion

The sample data does not allow rejection of the claim that there is no difference between the yields from regular and kiln dried seed.

Note: We could have used the sign test here, but the Wilcoxon test is a more powerful option, what does that mean? Why does it have more power? (Answer: Having more power means that it is better at detecting differences between two populations. The reason why it has more power is because it considers not just the sign of the difference, but also the rank (magnitude) of the difference between the matched pairs).

**Assumptions:**

1. The sample of differences was randomly selected from the population of differences.
2. The probability distribution for the paired differences is continuous.
3. The population differences have a distribution that is symmetric.

**Example 189** The following data shows the scores given to two different products (A and B) by ten different judges. Use the Wilcoxon-Signed Ranks test to determine if product A has a higher given median score by the judges.

Judge	Product A	Product B	A - B	A - B	Rank of  A - B
1	6	4	2	2	5
2	8	5	3	3	7.5
3	4	5	-1	1	2
4	9	8	1	1	2
5	4	1	3	3	7.5
6	7	9	-2	2	5
7	6	2	4	4	9
8	5	3	2	2	5
9	6	7	-1	1	2
10	8	2	6	6	10

## 14.7 Kruskal-Wallis H-Test: Completely Randomized Design

### Nonparametric Analysis of Data from a CRD

In the earlier sections, we used an ANOVA F-test to analyze the data from a CRD to compare the  $k$  population means of the treatments. The **Kruskal-Wallis H-test** is a nonparametric option for comparing  $k$  populations. It is used to test the null hypothesis that the independent samples come from populations with the equal medians. When conducting a Kruskal-Wallis H-test, we compute the test statistic  $H$ , which has a distribution that can be approximated by the chi-square ( $\chi^2$ ) distribution as long as each sample has at least 5 observations. When we use the *chi*-square distribution in this context, the number of degrees of freedom is  $k - 1$ , where  $k$  is the number of samples.

The hypotheses we use when running the KW test are as follows:

$$H_0 : \eta_1 = \eta_2 = \cdots = \eta_k$$

$H_A$  : At least two medians differ from each other.

We are about to work an example to demonstrate the technique involved in the KW test; however, before we begin we should introduce some notation:

- $N$  = total number of observations in all observations combined
- $k$  = number of samples
- $R_1$  = sum of ranks for Sample 1
- $n_1$  = number of observations in Sample 1
- For Sample 2, the sum of ranks is  $R_2$  and the number of observations is  $n_2$ , and similar notation is used for the other samples.

Consider the example below:

**Example 190** Test the claim at the 5% significance level that all four of the treatments produce the same median weight for poplar trees grown from seedlings.

Effects of treatments on Poplar tree weights

Treatments			
None	Fertilizer	Irrigation	Both
0.15	1.34	0.23	2.03
0.02	0.14	0.04	0.27
0.16	0.02	0.34	0.92
0.37	0.08	0.16	1.07
0.22	0.08	0.05	2.38



**Step 1 Claim**

All four of the treatments produce the same median weight for poplar trees.

**Step 2 Hypotheses**

$$H_0 : \eta_1 = \eta_2 = \eta_3 = \eta_4$$

$H_A$  : At least two medians differ from each other.

**Step 3** Temporarily view the entire data set as a whole and rank the data values (Average the ranks in case of a tie as usual). Then for each treatment add up its ranks.

Treatments							
None		Fertilizer		Irrigation		Both	
0.15	8	1.34	18	0.23	12	2.03	19
0.02	1.5	0.14	7	0.04	3	0.27	13
0.16	9.5	0.02	1.5	0.34	14	0.92	16
0.37	15	0.08	5.5	0.16	9.5	1.07	17
0.22	11	0.08	5.5	0.05	4	2.38	20
$n_1 = 5$	$R_1 = 45$	$n_2 = 5$	$R_2 = 37.5$	$n_3 = 5$	$R_3 = 42.5$	$n_4 = 5$	$R_4 = 85$

**Step 4** Calculate the test statistic  $H = \frac{12}{N(N+1)} \left( \frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \dots + \frac{R_k^2}{n_k} \right) - 3(N+1)$

$$H = \frac{12}{20(20+1)} \left( \frac{45^2}{5} + \frac{37.5^2}{5} + \frac{42.5^2}{5} + \frac{85^2}{5} \right) - 3(20+1) = 8.214$$

**Step 5** Get Critical Value: The H statistic can be approximated by a Chi-Squared distribution with  $k - 1$  degrees of freedom, so we will look up alpha on the chi-square critical value table. The test is always a right tailed test.

$$\chi_{.05,3}^2 = 7.81473$$

**Step 6** State your initial conclusion—We will reject the null when  $H > \chi_{\alpha, k-1}^2$

$$H = 8.214 > \chi_{.05,3}^2 = 7.81473$$

**Step 7** Word final conclusion

There is sufficient evidence to warrant rejection of the claim that all four of the treatments produce the same median weight for poplar trees.

**Assumptions:**

1. All the samples are independent and randomly selected.
2. Each sample has at least 5 observations.
3. The  $k$  probability distributions are continuous.

## 14.8 Friedman Fr-Test: Randomized Block Design

### Nonparametric Analysis of Data from RBD

The nonparametric test we will use to analyze data from a randomized block design experiment is called Friedman  $F_r$ -Test .

Consider the example below:

**Example 191** Test the claim at the 5% significance level that all three drugs have the same probability distribution.

Reaction Time for Three Drugs

Subject	Drug A	Rank	Drug B	Rank	Drug C	Rank
1	1.21		1.48		1.56	
2	1.63		1.85		2.01	
3	1.42		2.06		1.70	
4	2.43		1.98		2.64	
5	1.16		1.27		1.48	
6	1.94		2.44		2.81	
		$R_1 =$		$R_2 =$		$R_3 =$

**Step 1 Claim**

All three drugs have the same probability distribution.

**Step 2 Hypotheses**

$H_0$  : The probability distributions for the 3 treatments are identical.

$H_A$  : At least two of the distributions differ in location.

**Step 3** Rank the data values in each block (Average the ranks in case of a tie as usual). Then for each treatment add up its ranks.

Subject	Drug A	Rank	Drug B	Rank	Drug C	Rank
1	1.21	1	1.48	2	1.56	3
2	1.63	1	1.85	2	2.01	3
3	1.42	1	2.06	3	1.70	2
4	2.43	2	1.98	1	2.64	3
5	1.16	1	1.27	2	1.48	3

6	1.94	1	2.44	2	2.81	3
		$R_1 = 7$		$R_2 = 12$		$R_3 = 17$

**Step 4** Calculate the test statistic  $F_r = \frac{12}{bk(k+1)} \sum R_j^2 - 3b(k+1)$

$$F_r = \frac{12}{6 * 3(3+1)} (7^2 + 12^2 + 17^2) - 3(6)(3+1) = 8.33$$

**Step 5** Get Critical Value: The  $F_r$  statistic can be approximated by a Chi-Squared distribution with  $k - 1$  degrees of freedom, so we will look up alpha on our chi-squared table. The test is always a right tailed test.

$$\chi_{.05,2}^2 = 5.99147$$

**Step 6** State your initial conclusion—We will reject the null when  $F_r > \chi_{\alpha,k-1}^2$

$$F_r = 8.33 > \chi_{.05,2}^2 = 5.99147$$

**Step 7** Word final conclusion

There is sufficient evidence to warrant rejection of the claim that all three drugs have the same probability distribution.

**Assumptions:**

1. The treatments are randomly assigned to the blocks.
2. The probability distributions from which the samples within each block are drawn are continuous.
3. There is no interaction between blocks and treatments.
4. The observations in each block may be ranked.