統計學

Spring 2026

政治大學統計系余清祥
2026年05月05日
第十八章：無母數方法
http://csyue.nccu.edu.tw

Chapter Contents

CENGAGE

Introduction

Nonparametric methods can be used to make inferences about a population without requiring an assumption about the specific form of the population’s probability distribution.

- For this reason, nonparametric methods are also called **distribution-free methods**.

Nonparametric methods are often the only approach to analyze categorical (nominal or ordinal) data and draw statistical conclusions.

- Whenever the data are quantitative, we will transform the data into categorical data to conduct the nonparametric test.

In this chapter, we first show how the binomial distribution uses two categories of data to make an inference about a population median.

Then, we show how rank-ordered data are used in nonparametric tests about two or more populations.

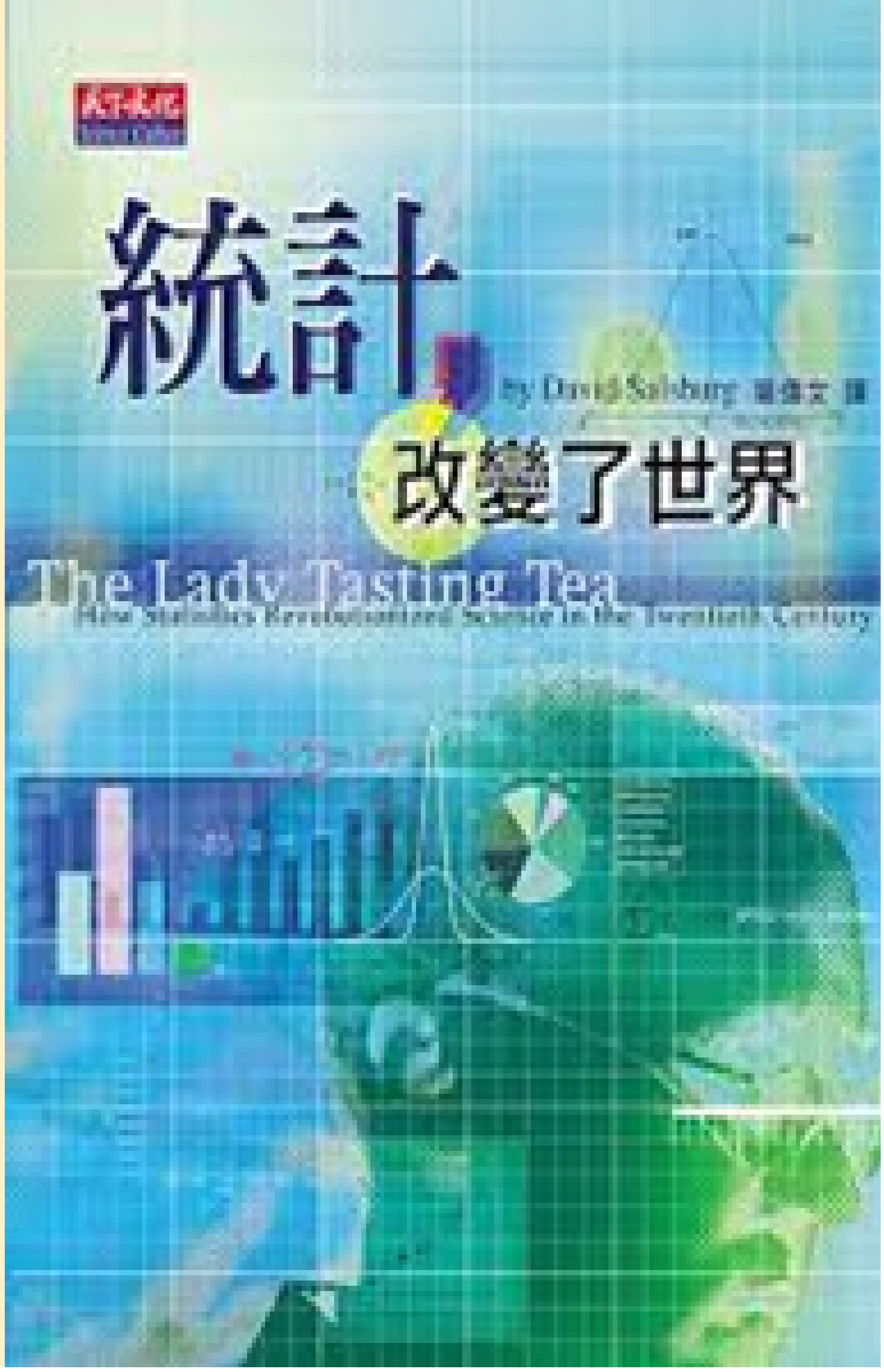
Finally, we use rank-ordered data to compute the rank correlation for two variables.

「淑女與下午茶」

- 在英國劍橋有位女士，聲稱把茶加到牛奶、把牛奶加到茶裡，兩種方法調出來的下午茶喝起來味道不同。在座的科學家都對她的說法嗤之以鼻，但有位來訪的瘦小紳士（RA費雪），提議要用科學的方法檢驗這位女士的假設。

→實驗設計(Experimental Design)

→如果十次測試，這位女士全部猜中茶加奶、奶加茶的順序，可能會是亂猜嗎？



18.1 Sign Test About a Population Median

The **sign test** is a versatile distribution-free method for the testing of a population median.

We conduct the nonparametric sign test about a population median by:

- recording a plus sign “+” whenever the data in the sample are above the hypothesized value of the median.
- recording a minus sign “−” whenever the data in the sample are below the hypothesized value of the median.
- discarding any observation exactly equal to the hypothesized median.

The hypotheses for a sign test about a population median are:

$$\boldsymbol{H_0: p = 0.50}$$

$$\boldsymbol{H_a: p \neq 0.50}$$

Because there are only two possible outcomes per trial, the distribution of the count of plus signs follows a binomial distribution where p denotes the probability of a plus sign.

18.1 Sign Test About a Population Median: $n \leq 20$

Lawler's Grocery Store made the decision to carry Cape May Potato Chips based on the manufacturer's estimate that the median sales should be $450 per week on a per-store basis.

One-week sales at 10 randomly selected Lawler's stores, as well as whether weekly sales were higher ("+") or lower ("−") than $450, are shown in the table.

Thus, the hypotheses for the test about a population median are:

$$H_0\text{: Median} = \$450$$

$$H_a\text{: Median} \neq \$450$$

To use the sign test, we convert to the hypotheses about the binomial probability p of a plus sign:

$$H_0: p = 0.50$$

$$H_a: p \neq 0.50$$

If we reject H_0, we can conclude that the population median is not $450.

Store Number	Weekly Sales ($)	Sign
56	485	+
19	562	+
36	415	−
128	860	+
12	426	−
63	474	+
39	662	+
84	380	−
102	515	+
44	721	+

18.1 *p*-Value Approach to a Sign Test: *n* ≤ 20

We will use an $\alpha = 0.10$ level of significance for the test.

Since the observed number of plus signs, 7, is in the upper tail of the binomial distribution, we begin by computing the probability of obtaining 7 or more plus signs out of 10 trials. Then, for a two-tailed test, we multiply the probability by two to obtain the *p*-value.

The probabilities shown in the table can be also obtained using Table 5 in Appendix B, with $n = 10$ trials or stores, and $p = 0.50$.

Thus, we have:

$$p\text{–value} = 2(0.1172 + 0.0439 + 0.0098 + 0.0010) = 2(0.1719) = 0.3438$$

Alternatively, we can use BINOM.DIST in Excel (*see notes.)

With $p\text{–value} = 0.3438 > \alpha = 0.10$, we cannot reject H_0. Thus, we cannot conclude that the population median weekly sales is not $450 per store.

Calculations about a one-tailed sign test proceed similarly.

Number of Plus Signs	**Probability**
0	0.0010
1	0.0098
2	0.0439
3	0.1172
4	0.2051
5	0.2461
6	0.2051
7	0.1172
8	0.0439
9	0.0098
10	0.0010

屏東慈天宮連擲20聖筊拚300萬 2萬人挑戰仍槓龜


https://www.ctee.com.tw/news/20260218700278-431401

18.1 Normal Approximation of a Sign Test: *n* > 20

The binomial probabilities provided in Table 5 of Appendix B are used to compute the *p*-value when the sample size is 20 or less.

With larger sample sizes, we rely on the normal distribution approximation of the binomial distribution to compute the *p*-value; this makes the computations quicker and easier.

To conduct the normal approximation of the sampling distribution of the number of plus signs when $H_0\!: p = 0.50$, and $n > 20$, we set:

Mean: $\boldsymbol{\mu = 0.50n}$

Standard deviation: $\boldsymbol{\sigma = \sqrt{0.25n}}$

Example: a real estate firm used a sample of recent home sales to determine if the population median price of a new home is less today than it was a year ago (DATAfile: *HomeSales.)*

The real estate firm has determined that one year ago, the median price of a new home was \$336,000. We use a significance level, $\alpha = 0.05$.

18.1 Sign Test About a Population Median: *n* > 20

Because we are testing for a population median decrease from last year, the hypotheses for the one-tailed test about the population median price of a new home is

$$H_0\text{: Median} \geq \$336{,}000 \qquad H_a\text{: Median} < \$336{,}000$$

Data analysis reveals that, of the $n = 61$ home sales, 22 sold for more than $336,000, 38 for sold for less than $336,000, and one home sold for $336,000.

Thus, we delete the home sold for $336,000, and we are left with 22 "+" signs, 38 "−" signs, and a sample size, $n = 60$.

To use the sign test, we convert to the hypotheses about the probability p of a plus sign:

$$H_0\!: p \geq 0.50 \qquad H_a\!: p < 0.50$$

The normal approximation of the sampling distribution of the number of plus signs for the home sales example, when $H_0\!: p = 0.50$, and $n = 60$, is:

$$\mu = 0.50n = 0.5(60) = 30 \text{ homes and } \sigma = \sqrt{0.25n} = \sqrt{0.25(60)} = \sqrt{15} = 3.873 \text{ homes}$$

CENGAGE

18.1 *p*-Value Approach to a Sign Test: *n* > 20

Because the normal distribution is a continuous probability distribution, we need to account for the fact that the number of plus signs is a discrete variable.

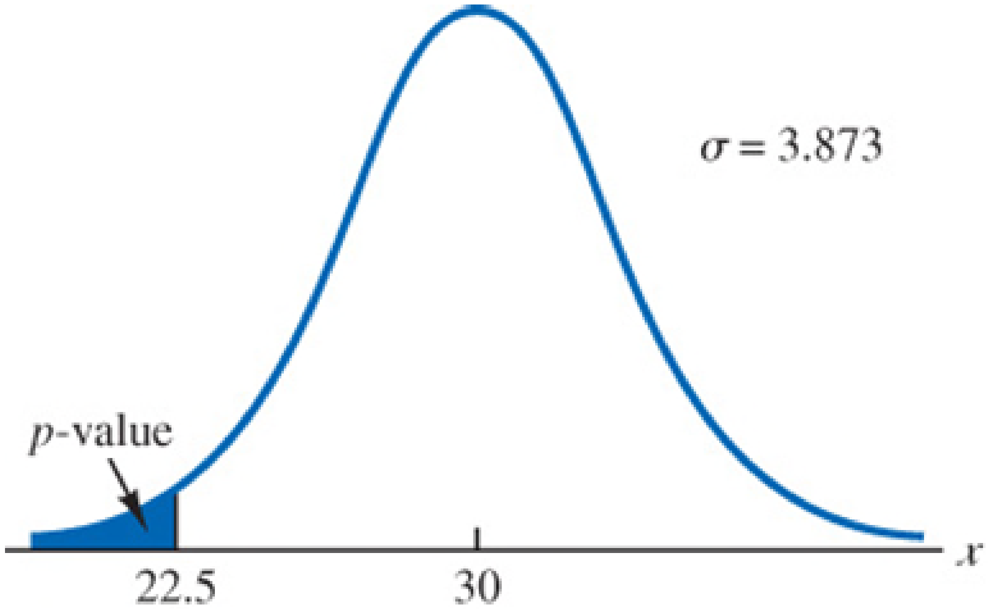
Thus, to calculate the *p*-value as the lower tail probability of having no more than 22 plus signs out of 60, we need to apply the continuity correction factor, and set the test statistic to 22.5.

Using Table 2 of Appendix B, or the NORM.S.DIST function in Excel, we have:

$$p\text{–value} = P(x \leq 22.5) = P\left(z \leq \frac{22.5 - 30}{3.873}\right)$$

$$= P(z \leq -1.94) = 0.0262$$

With $p\text{–value} = 0.0262 \leq \alpha = 0.05$, we reject H_0 and conclude that the median price of a new home is less than the $336,000 median price a year ago.



18.1 Sign Test With Matched Samples

A common application of the sign test involves using a sample of customers to determine whether there is a difference in preference between the two brands being compared.

We record the preference data, using a plus sign (”+”) if the individual prefers one brand and a minus sign (“–”) if the individual prefers the other brand.

Because the data are recorded as plus and minus signs, we use the nonparametric sign test to analyze the matched-sample data.

Consider the following application of matched sample data.

Sun Coast Farms conducted a consumer preference study between its orange juice, Citrus Valley, and the primary competition, Tropical Orange.

The table shows the study results for the brand of orange juice preferred by 14 randomly selected individuals.

Individual	Preference	Sign
1	Tropical Orange	–
2	Tropical Orange	–
3	Citrus Valley	+
4	Tropical Orange	–
5	Tropical Orange	–
6	No Preference	
7	Tropical Orange	–
8	Tropical Orange	–
9	Tropical Orange	–
10	No Preference	
11	Tropical Orange	–
12	Citrus Valley	+
13	Tropical Orange	–
14	Tropical Orange	–

CENGAGE

18.1 Hypothesis Test With Matched Samples

Deleting the two individuals who could not express a preference for either brand, we are left with a sign test having 2 plus signs and 10 minus signs for the $n = 12$ individuals who could express a preference for one of the two brands.

Letting p indicate the proportion of the population of customers who prefer Citrus Valley orange juice, we want to test the hypotheses that there is no difference between the preferences for the two brands as follows:

$$H_0 : p = 0.50$$

$$H_a : p \neq 0.50$$

We will use an $\alpha = 0.05$ level of significance for the test.

Since the observed number of plus signs, 2, is in the lower tail of the binomial distribution, we will calculate the *p*-value by multiplying by two the probability of obtaining 2 or less plus signs out of 12 individuals.

CENGAGE

18.1 *p*-Value Approach to a Matched Samples Sign Test

The probabilities shown in the table can also be obtained using Table 5 in Appendix B, with $n = 10$ trials or stores, and $p = 0.50$.

Number of Plus Signs	Probability
0	0.0002
1	0.0029
2	0.0161
3	0.0537
4	0.1208
5	0.1934
6	0.2256
7	0.1934
8	0.1208
9	0.0537
10	0.0161
11	0.0029
12	0.0002

Thus, we have:

$$p\text{−value} = 2(0.0002 + 0.0029 + 0.0161) = 2(0.0192) = 0.0384$$

Alternatively, we can use Excel to calculate the *p*-value:

$$p\text{−value}{=}2{*}\text{BINOM.DIST}(2{,}12{,}0.5{,}\text{TRUE}){=}0.0384$$

With $p\text{−value} = 0.0384 \leq \alpha = 0.05$, we reject H_0.

The taste test provides significant evidence that consumer preference differs significantly for the two brands of orange juice.

We would advise Sun Coast Farms of this result and conclude that the competitor’s Tropical Orange product is the preferred one.

Sun Coast Farms can then pursue a strategy to address this issue.

18.2 Wilcoxon Signed-Rank Test

The **Wilcoxon signed-rank test** analyzes quantitative data from a matched samples experiment, but only requires the assumption that the differences have a symmetric distribution.

- This occurs whenever the shapes of the two populations are the same and the focus is on determining whether there is a difference between the two populations' medians.
- The Wilcoxon signed-rank test does not require the assumption that the differences between the paired observations are normally distributed.

To perform a Wilcoxon signed-rank test:

1. Compute the difference for each matched pair in the sample.
2. Delete any difference equal to zero.
3. Rank the remaining differences in ascending order, with tied values being assigned the average of the ranks.

When H_0 is true, the medians of the two populations are equal, and the sum of the negative signed ranks, T^-, and the sum of the positive signed ranks, T^+, are approximately the same.

18.2 Wilcoxon Signed-Rank Test Statistic

The Wilcoxon signed-rank test uses the sum of the positive signed ranks, T^+, as a test statistic.

T^+ is a random variable that follows a normally-distributed sampling distribution with mean and standard deviation defined as:

$$\boldsymbol{\mu_{T^+} = \frac{n(n+1)}{4}} \qquad \boldsymbol{\sigma_{T^+} = \sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

so long as the sample size $n \geq 10$.

As an application of the Wilcoxon signed-rank test, consider a manufacturing firm that wants to determine whether two production methods differ in terms of task completion time.

Using a matched-samples experimental design, 11 randomly selected workers completed the production task two times, once using method A and once using method B, and the completion times for the two methods is recorded.

The production method that the worker used first was randomly selected.

18.2 Application of the Wilcoxon Signed-Rank Test

Because we are testing for the difference between the median completion times for the two production methods, the hypotheses are

$$H_0\text{: Median Method A} - \text{Median Method B} = 0$$

$$H_a\text{: Median Method A} - \text{Median Method B} \neq 0$$

We delete worker 8 from the data set because it had a difference of zero,

Then, we compute the absolute difference in completion time for the $n = 10$ remaining workers and rank them in ascending order.

The sum of the positive signed ranks is

$$T^+ = 8 + 3.5 + 5.5 + 10 + 1 + 7 + 5.5 + 9$$
$$= 49.5$$

				Signed Ranks	
Worker	**Difference**	**Absolute Difference**	**Rank**	**Negative**	**Positive**
1	0.7	0.7	8		8
2	-0.2	0.2	2	-2	
3	0.4	0.4	3.5		3.5
4	0.5	0.5	5.5		5.5
5	-0.4	0.4	3.5	-3.5	
6	0.9	0.9	10		10
7	0.1	0.1	1		1
8	0.0	0.0			
9	0.6	0.6	7		7
10	0.5	0.5	5.5		5.5
11	0.8	0.8	9		9
		Sum of positive signed ranks			$T^+ = 49.5$

18.2 *p*-Value Approach to a Wilcoxon Signed-Rank Test

With $n = 10$, the sampling distribution is normal with mean and standard deviation equal to

$$\mu_{T^+} = \frac{n(n+1)}{4} = \frac{10(10+1)}{4} = \frac{110}{4} = 27.5$$

$$\sigma_{T^+} = \sqrt{\frac{n(n+1)(2n+1)}{24}} = \sqrt{\frac{10(10+1)(20+1)}{24}} = \sqrt{\frac{2310}{24}} = 9.8107$$

Because $T^+ = 49.5$ is a discrete value located in the upper tail of a contintuous normal distribution, we apply the continuity correction factor and set $T^+ = 49$.

Thus, for a two-tailed test, we calculate the *p*-value as

$$p\text{−value} = 2P(T^+ \geq 49) = 2P\left(z \geq \frac{49 - 27.5}{9.8107}\right) = 2P(z \geq 2.19) = 2[1 - P(z \leq 2.19)] = 0.0286$$

Because $p\text{−value} = 0.0286 \leq \alpha = 0.05$, we reject H_0 and conclude that the median completion times for the two production methods are not equal.

18.2 Sign Test vs. Wilcoxon Signed-Rank Test

In Section 18.1, we learned that the sign test can be used for both a hypothesis test about a population median and a hypothesis test with matched samples.

Besides a hypothesis test with matched samples, as shown in this section, the Wilcoxon signed-rank test can also be used for a nonparametric test about a population median.

- If the population distribution is symmetric, the Wilcoxon signed-rank test is the preferred nonparametric test for a population median (* see notes.)
- If the population distribution is skewed, the sign test is preferred.

To perform a Wilcoxon signed-rank test on a population median, the differences between the observations and the hypothesized value of the population median are used instead of the differences between the matched-pair observations.

Otherwise, the calculations are exactly as shown in this section.

18.3 Mann-Whitney-Wilcoxon Test

The **Mann-Whitney-Wilcoxon (MWW) test** is a nonparametric method that uses two independent sample to determine whether there is a difference between two populations.

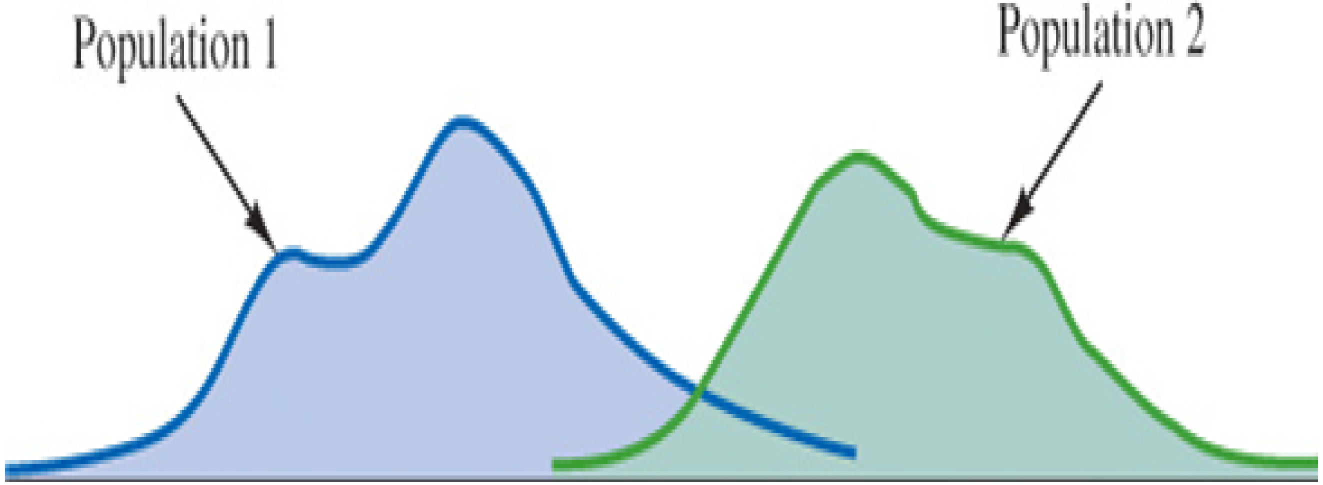
Advantages of this procedure are:

- It can be used with either ordinal data or quantitative data.
- It does not require the assumption that the populations have a normal distribution.

The MWW test determines whether the two populations are identical:

H_0: The two populations are identical

H_a: The two populations are not identical



18.3 Mann-Whitney-Wilcoxon Test Statistic

The steps to conduct at Mann-Whitney-Wilcoxon test are:

1. Combine the data from the two independent samples into one.
2. Rank the combined sample data in ascending order, with tied values being assigned the average of the ranks.
3. Compute W, the sum of the ranks for the first sample.
4. Compare the observed value of W to the sampling distribution of W for identical populations.
    - For small sample sizes, the exact sampling distribution of W has to be computed.
    - For larger sample sizes such that $n_1 \geq 7$ and $n_2 \geq 7$, the sampling distribution of W can be approximated by a normal distribution with mean and standard deviation:

$$\boldsymbol{\mu_W = \frac{n_1(n_1 + n_2 + 1)}{2}} \qquad \boldsymbol{\sigma_W = \sqrt{\frac{n_1 n_2(n_1 + n_2 + 1)}{12}}}$$

18.3 Mann-Whitney-Wilcoxon Test: Small Samples Case

A manager at Showtime Cinemas, rated all 35 employees from best (rating 1) to worst (rating 35) in the theater's annual report.

Knowing that the part-time employees were primarily college and high school students, the district manager asked if there was evidence of a significant difference in their performance.

Among the 35 employees, four were college students and five were high school students, for a combined sample of nine students.

The hypothesis test, with $\alpha = 0.05$, was set as follows:

H_0: College and high school student populations are identical

H_a: College and high school student populations are not identical

The ratings were sorted, ranked in ascending order, and summed.

The sum of ranks for the first sample (the college students) will be the test statistic, $W = 4 + 1 + 7 + 2 = 14$.

College Student	Manager's Performance Rating	Rank
1	15	4
2	3	1
3	23	7
4	8	2
	Sum of ranks	14

High School Student	Manager's Performance Rating	Rank
1	18	5
2	20	6
3	32	9
4	9	3
5	25	8
	Sum of ranks	31

CENGAGE

18.3 *p*-Value Approach to MWW Test: Small Samples Case

To build the sampling distribution, we observe that the possible values of W, for 4 college students out of 9, range between $W = 1 + 2 + 3 + 4 = 10$ and $W = 6 + 7 + 8 + 9 = 30$.

Because there is a total of 126 possible ranking combinations, the sampling distribution of W can be computed as shown in the table (*see notes.)

When H_0 is true, the two populations are identical, and we would expect a mix in the ordering of the ranks so that W is near the average of the two extremes, $(10 + 30)/2 = 20$.

Thus, for a two-tailed test, the *p*-value will be twice the probability that W is less than or equal to 14.

$$p\text{–value} = 2P(W \leq 14)$$
$$= 2(0.0079 + 0.0079 + 0.0159 + 0.0238 + 0.0397) = 0.1904$$

Because $p\text{–value} > \alpha$, we cannot reject H_0 and conclude that the populations of college and high school students are not identical.

W	Probability	W	Probability
10	0.0079	20	0.0952
11	0.0079	21	0.0873
12	0.0159	22	0.0873
13	0.0238	23	0.0714
14	0.0397	24	0.0635
15	0.0476	25	0.0476
16	0.0635	26	0.0397
17	0.0714	27	0.0238
18	0.0873	28	0.0159
19	0.0873	29	0.0079
		30	0.0079

CENGAGE

18.3 Mann-Whitney-Wilcoxon Test: Large Samples Case

A bank manager wants to determine if the populations of checking account balances at two Third National Bank branches are identical.

Two independent samples of checking accounts are taken with sample sizes $n_1 = 12$ at branch 1 and $n_2 = 10$ at branch 2.

As it was for the small samples case, the balances for the 22 combined accounts were sorted and ranked in ascending order, with tied values being assigned the average of the ranks.

The test statistic W is the sum of the rankings for sample 1.

$$W = 3 + 6 + 9 + 11 + 12.5 + 14 + 15 + 17 + 19 + 20 + 21 + 22 = 169.5$$

Because $n_1 \geq 7$ and $n_2 \geq 7$, we can approximate the sampling distribution of W with a normal distribution, having mean and standard deviation:

$$\mu_W = \frac{12(12 + 10 + 1)}{2} = 138 \quad \sigma_W = \sqrt{\frac{(12)(10)(12 + 10 + 1)}{12}} = 15.1658$$

Branch	Account	Balance	Rank
2	6	750	1
2	5	800	2
1	7	805	3
2	2	850	4
2	7	865	5
1	9	875	6
2	1	885	7
2	3	915	8
1	5	925	9
2	10	935	10
1	8	945	11
1	6	950	12.5
2	4	950	12.5
1	2	955	14
1	12	975	15
2	8	1000	16
1	11	1025	17
2	9	1050	18
1	10	1055	19
1	1	1095	20
1	4	1195	21
1	3	1200	22

CENGAGE

18.3 *p*-Value Approach to MWW Test: Large Samples Case

Because the normal distribution is a continuous probability distribution, we need to account for the fact that the test statistic W is a discrete variable.

In this case, to calculate the *p*-value as the upper tail probability of having $W \geq 169.5$, we need to apply the continuity correction factor, and set the test statistic to $W = 169$.

Using Table 2 of Appendix B, we have:

$$p\text{-value} = 2P(W \geq 169) = 2P\left(z \geq \frac{169 - 138}{15.1658}\right) = 2P(z \geq 2.04) = 2[1 - P(z \leq 2.04)] = 0.0286$$

With $p\text{-value} = 0.0286 \leq \alpha = 0.05$, we reject the null hypothesis and conclude that the two populations of account balances are not identical.

The upper tail value for test statistic W indicates that the population of account balances at branch 1 tends to be larger.

CENGAGE

18.3 Mann-Whitney-Wilcoxon Test: Final Remarks

If the two populations have identical shapes, then the hypothesis test for the Mann-Whitney-Wilcoxon test may be stated in terms of the difference between the two population medians.

Thus, any difference between the medians can be interpreted as the shift in location of one population compared to the other.

In such case, the three forms of the Mann-Whitney-Wilcoxon test of hypotheses about the medians of the two populations are as follows:

Lower Tail Test	Upper Tail Test	Two-Tailed Test
$H_0\text{: Median}_1 \geq \text{Median}_2$	$H_0\text{: Median}_1 \leq \text{Median}_2$	$H_0\text{: Median}_1 = \text{Median}_2$
$H_a\text{: Median}_1 < \text{Median}_2$	$H_a\text{: Median}_1 > \text{Median}_2$	$H_a\text{: Median}_1 \neq \text{Median}_2$

18.4 Kruskal-Wallis Test

The **Kruskal-Wallis test** is based on the analysis of independent random samples from k populations. It can be used with either ordinal data or quantitative data and does not require the assumption that the populations have normal distributions.

The general form of the null and alternative hypotheses is as follows (*see notes):

$$\boldsymbol{H_0}\text{: All } \boldsymbol{k} \text{ populations are identical}$$

$$\boldsymbol{H_a}\text{: Not all } \boldsymbol{k} \text{ populations are identical}$$

The Kruskal-Wallis test statistic, H, is calculated using the sum of the ranks, R_i, for k samples:

$$\boldsymbol{H = \left[\frac{12}{n_T(n_T+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i}\right] - 3(n_T+1)} \qquad \text{where } \boldsymbol{n_T = \sum_{i=1}^{k} n_i}$$

When H_0 is true, the sampling distribution of H can be approximated by a chi-square distribution with $k-1$ degrees of freedom, so long as $n_i \geq 5$ for each of the k samples.

18.4 An Application of the Kruskal-Wallis Test

The personnel director at Williams Manufacturing Company reviewed the annual performance reports for the management staff to determine whether there are differences in the performance ratings among the managers who graduated from either college A, B, or C.

$n_T = 20$ performance ratings, recorded on a scale from 0 to 100, are available for independent samples of $n_1 = 7$ managers who graduated from college A, $n_2 = 6$ managers who graduated from college B, and $n_3 = 7$ managers who graduated from college C.

To test whether the three populations of managers have identical performance ratings, the director:

1. ranked the combined 20 ratings
2. assigned the average of ranks to tied ratings
3. summed the ranks for each sample

A 0.05 significance level was used for this test.

College A	Rank	College B	Rank	College C	Rank
25	3	60	9	50	7
70	12	20	2	70	12
60	9	30	4	60	9
85	17	15	1	80	15.5
95	20	40	6	90	18.5
90	18.5	35	5	70	12
80	15.5			75	14
	95		27		88

CENGAGE

18.4 *p*-Value Approach to the Kruskal-Wallis Test

The Kruskal-Wallis test statistic for the three rating samples is computed as follows:

$$H = \left[\frac{12}{n_T(n_T+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i}\right] - 3(n_T+1) = \frac{12}{20(20+1)}\left[\frac{95^2}{7}+\frac{27^2}{6}+\frac{88^2}{7}\right] - 3(20+1) = 8.92$$

Using the chi-square distribution table (Table 3 of Appendix B) for $k = 3 - 1 = 2$ degrees of freedom, we find $P(\mathrm{X}^2 \geq 0.025) = 7.378$ and $P(\mathrm{X}^2 \geq 0.01) = 9.21$.

Thus, with $H = 8.92$, we conclude that $0.01 \leq p\text{−value} \leq 0.025$.

We can also use Excel to calculate the exact *p*-value as $=\mathrm{CHISQ.DIST.RT}(8.92{,}2){=}0.0116$.

Because $p\text{−value} \leq \alpha = 0.05$, we reject H_0 and conclude that the three populations of performance ratings are not identical and differ significantly depending upon the college.

Because the sum of the ranks is relatively low for the managers who graduated from college B, the company should thoroughly evaluate college B graduates before making a hiring decision.

18.5 Rank Correlation

The Pearson correlation coefficient, r, is a measure of the linear association between two variables for which interval or ratio data are available.

The **Spearman rank-correlation coefficient**, r_s, is a measure of association between two variables when only ordinal or rank-ordered data are available.

The Spearman rank-correlation coefficient is defined as

$$\boldsymbol{r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}}$$

Where

n = number of observations being ranked

x_i = rank of observation i with respect to the first variable

y_i = rank of observation i with respect to the second variable

$d_i = x_i - y_i$

18.5 Testing for Rank Correlation Significance

We can use the sample rank correlation, r_s, to make an inference about the population rank correlation coefficient, ρ_s, by testing the following hypotheses:

$$\boldsymbol{H_0: \rho_s = 0}$$

$$\boldsymbol{H_a: \rho_s \neq 0}$$

Under the assumption that H_0 is true and $\rho_s = 0$, the sampling distribution of r_s follows a normal distribution with mean and standard deviation defined as

$$\boldsymbol{\mu_{r_s} = 0}$$

$$\boldsymbol{\sigma_{r_s} = \sqrt{\frac{1}{n-1}}}$$

Provided that the sample size, $n \geq 10$.

18.5 Spearman Rank-Correlation Coefficient

A company's personnel director reviewed the hiring files of $n = 10$ members of the sales force and ranked them in terms of their potential for success at the time of employment by assigning the salesperson who had the most potential the rank of 1.

A second ranking of the 10 salespersons, based on the actual sales records, was also obtained.

The table provides the ranks based on potential (x_i), the ranks based on the actual performance (y_i), as well as the calculations for $\sum_{i=1}^{n} d_i^2$.

We can calculate r_s as follows

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2+1)} = 1 - \frac{6(44)}{10(10^2+1)} = 0.739$$

Salesperson	x_i = Ranking of Potential	y_i = Ranking Sales Perf.	$d_i = x_i - y_i$	d_i^2
A	2	1	1	1
B	4	3	1	1
C	7	5	2	4
D	1	6	-5	25
E	6	7	-1	1
F	3	4	-1	1
G	10	10	0	0
H	9	8	1	1
I	8	9	-1	1
J	5	2	3	9
				$\Sigma d_i^2 = 44$

CENGAGE

18.5 *p*-Value Approach to a Rank Correlation Test

With $n = 10$, we can approximate the sampling distribution of r_s with a normal distribution, having mean and standard deviation equal to

$$\mu_{r_s} = 0 \quad \text{and} \quad \sigma_{r_s} = \sqrt{\frac{1}{n-1}} = \sqrt{\frac{1}{10-1}} = \frac{1}{3} = 0.333$$

We conduct a test on the Spearman rank-correlation coefficient using a 0.05 significance level.

We calculate the *p*-value multiplying by two the upper tail probability of having $r_s \geq 0.739$.

Using Table 2 of Appendix B, we have:

$$p\text{−value} = 2P(r_s \geq 0.739) = 2P\left(z \geq \frac{0.739 - 0}{0.333}\right) = 2P(z \geq 2.20) = 2[1 - P(z \leq 2.20)] = 0.0278$$

With $p\text{−value} = 0.0278 \leq \alpha = 0.05$, we reject H_0 and conclude that there is a significant rank correlation between potential at the time of employment and actual sales performance.

Summary

- In this chapter, we have presented statistical procedures that are classified as nonparametric methods and can be applied to categorical data as well as quantitative data.
- The sign test is a nonparametric procedure for testing a hypothesis about a population median or for testing a hypothesis with matched samples.
- The Wilcoxon signed-rank test analyzes matched samples from two populations when quantitative data are available.
  - It can also be used to make inferences about the median of a symmetric population.
- The Mann-Whitney-Wilcoxon (MWW) test is a nonparametric procedure for the difference between two populations; the test uses the combined ranks from two independent samples.
  - The MWW test can also provide an inference about the difference between the medians of the two populations having the same shape.
- The Kruskal-Wallis test extends the MWW test to three or more populations.
- Finally, we introduced the Spearman rank-correlation coefficient as a measure of association between two variables based on rank-ordered data.