CHAPTER 14

More Nonparametric Tests

Summary _____

To use the nonparametric statistical tests in this chapter, the data must be ranks or be reduced to ranks.

Nonparametric statistics are similar to parametric statistics in some ways and different from them in others. They are similar in that they use the same hypothesis-testing logic (NHST) and that they require random assignment (or random sampling). They are different in that they do not assume that populations are normally distributed or that the variances are equal. They are also different because the null hypothesis is that the population *distributions* are identical rather than that the population *means* are identical.

To test the hypothesis that two population distributions are identical when data are from a two-group, *independent-samples design*, use the *Mann-Whitney U test*. Rank the scores for the entire data set, and then separately sum the ranks for each group. Using these sums in a formula, calculate two values of U. Use the smaller U. For designs with both N's less than 20, U is evaluated for significance with Table H. For larger samples, a z test based on the smaller U is used.

To test the hypothesis that two population distributions are identical when data are from a two-group, *paired-samples design*, use the *Wilcoxon matched-pairs, signed-ranks T test*. The difference between pairs of scores is found. The differences are ranked, with the smallest difference ranked 1. Each rank receives the algebraic sign of its difference. Positive ranks are summed to produce a T value. Negative ranks are summed to produce a second T value. Using the absolute value, the smaller of the two T's is evaluated for significance. If the number of pairs is less than 50, T is evaluated with a Table J. For larger samples, a z test based on the smaller T value is used.

Contrary to other statistical tests you have studied, calculated values for both U and T must be *less* than tabled critical value to be significant.

Chapter 14

To test the hypothesis that population distributions are identical when the data are from an independent-samples design with *more than two levels* of the independent variable, use the *Wilcoxon-Wilcox multiple comparisons test.* Rank the scores for the whole data set and then separately sum the ranks for each group. Find the difference for each pair of groups. Each difference is evaluated for significance with a table of critical differences for the Wilcoxon-Wilcox multiple comparisons test, Table K.

To find the *degree of relationship* between two variables, use the Spearman r_s. Rank the scores within one group and then rank the scores in the other group. Find the difference between the ranks for each pair. Square the differences, sum them, and use this sum of the squared differences in the formula for r_s.

To test the hypothesis that two populations have a correlation coefficient of .00, calculate r_s and evaluate its significance using Table L (for N 16) or Table A (for $N > 16$).

The parametric counterparts of the nonparametric tests in this chapter are:

Mann-Whitney U test – independent-samples t test
Wilcoxon matched-pairs, signed-ranks T test – paired-samples t test
Wilcoxon-Wilcox multiple comparisons test – Tukey HSD test
Spearman r_s – Pearson product-moment correlation coefficient, r

The textbook's final word is that statistical methods and techniques are useful in some situations but not in others. Unfortunately, no practice problems were given to help you sharpen your understanding of this issue. However, knowing what statistics can and cannot do is a good start. We wish you well as you continue to develop your skills.

Multiple-Choice Questions _____

1. The nonparametric tests in the text are based on sampling distributions of
   (1) means;
   (2) mean differences;
   (3) ranks;
   (4) variances.

Chapter 14

2. The nonparametric test that corresponds in design to the independent-samples *t* test is the
   (1) Mann-Whitney *U* test;
   (2) Wilcoxon matched-pairs singed-ranks *T* test;
   (3) Wilcoxon-Wilcox multiple-comparisons test;
   (4) r_s.

3. Suppose you find that three people tied for the top score in a Wilcoxon-Wilcox multiple-comparisons test. The correct procedure is to
   (1) assign a rank of 1 to all three;
   (2) assign a rank of 2 to all three;
   (3) assign a rank of 3 to all three;
   (4) randomly determine which scores get ranks 1, 2, and 3.

4. Nonparametric tests are used rather than a *t* test or an ANOVA when
   (1) the researcher does not know the specific value of the population parameters;
   (2) the data are in the form of ranks;
   (3) the assumption of random assignment is not justified;
   (4) both (2) and (3).

5. The null hypothesis for testing the significance of r_s is that the population correlation coefficient is
   (1) .00;
   (2) 1.00;
   (3) the statistic, r_s, calculated from the sample data;
   (4) none of the above.

6. Your text discussed two issues for which hard and fast rules don't exist on the choice of parametric or nonparametric tests. These are:
   (1) random assignment and scales of measurement;
   (2) sample size and power;
   (3) random assignment and sample size;
   (4) scales of measurement and power.

7. When a sample size is large, the U value from a Mann-Whitney U test is evaluated using the
   (1) t distribution;
   (2) normal distribution;
   (3) F distribution;
   (4) χ^2 distribution.

8. Which of the following statistical tests is most similar to the Mann-Whitney U test?
   (1) t test;
   (2) ANOVA;
   (3) χ^2 test;
   (4) r_s.

9. To test for a significant difference between paired samples, use a
   (1) Mann-Whitney U test;
   (2) Wilcoxon matched-pairs signed-ranks T test;
   (3) Wilcoxon-Wilcox multiple-comparisons test;
   (4) r_s.

10. Which of the following could *not* be analyzed with a Wilcoxon-Wilcox multiple-comparisons test?
   (1) three independent samples;
   (2) $N_1 = 10$, $N_2 = 20$, $N_3 = 30$;
   (3) the data for three groups that consisted of ranks in college;
   (4) all of the above.

11. What is the rank of 4 in the following distribution?
   1, 2, 2, 3, 3, 4, 4, 5, 5.
   (1) 4;
   (2) 5.5;
   (3) 6;
   (4) 7.
   (5) none of the above, answer is _____.

Chapter 14

12. Power is the likelihood of
    (1) rejecting H_o when it is true;
    (2) rejecting H_o when it is false;
    (3) retaining H_o when it is true;
    (4) retaining H_o when it is false.

13. Which of the following factors has an influence on power?
    (1) sample size;
    (2) effect size;
    (3) statistical test;
    (4) all can influence power.

14. Suppose you found, for the 26 people in your wing of the dorm, a Spearman r_s of .38 between the number of breakfasts eaten during the term and grade point average. You may conclude that there is
    (1) no significant relationship;
    (2) a significant relationship at the .05 level;
    (3) a significant relationship at the .01 level;
    (4) a significant relationship at the .001 level.

15. If you have one difference score of 0 it should be kept in the analysis for
    (1) Wilcoxon matched-pairs signed-rank *T* test;
    (2) Spearman r_s;
    (3) both (1) and (2);
    (4) neither (1) nor (2).

16. One common use of statistics is to compile a composite score for each city on a list of cities. Suppose you thought that large cities would rank higher than small cities. To make sure that the age of the city didn't influence the comparison, you matched each large city with a small one that was founded at about the same time. Which statistic below should you use to test your idea about the size of cities and their rank?
    (1) Mann-Whitney *U* test;
    (2) Wilcoxon matched-pairs signed-ranks *T*;
    (3) Wilcoxon-Wilcox multiple comparisons test;
    (4) Spearman's r_s.

17. Which of the following captures the theme of the E. F. Schumacher anecdote at the end of the book?
    (1) Accurate counting is necessary for statistics;
    (2) Inferential statistics is central to all decision making;
    (3) Nonparametric statistics show that there will be new developments in statistics;
    (4) Some problems are not solved by being translated into numbers.

18. The _____ the value of *U* and the _____ the value of *T*, the more likely you are to reject the null hypothesis.
    (1) larger, larger;
    (2) larger, smaller;
    (3) smaller, smaller;
    (4) smaller, larger.

Interpretation _____

1. Which nonparametric test should be used to analyze data from the following studies?
    A. Some experimenters, impressed by the intimidation power of a steady, direct gaze by dominant primates, decided to investigate the effect in *Homo sapiens*. They rode a motor scooter up next to a car stopped at an intersection with a stoplight and either stared directly at the driver or stared ahead. The dependent variable was the time it took the driver to cross the intersection. (Results: staring increases speed for *Homo sapiens* too.)
    B. Countries are ranked on infant mortality rates. Suppose you had the ranks for 15 countries and the rank for each of the countries for population density. What test would determine whether there was a relationship between the two variables?
    C. To determine if there was sex discrimination in salaries at Old State U., a statistician began with 15 female professors. On the basis of degree, discipline, and years of experience, each was matched with a male professor.

2. Which nonparametric test should be used to analyze data from the following studies?
   A. To produce high-frustration, subjects were forced to wait 15 minutes for a late participant before beginning an experiment. Low-frustration subjects started on time. During the experiment aggression was measured as the dependent variable.
   B. To find out if there was a relationship between a basketball team's height and its rank in the conference, the average height and conference rank were determined for the 10 schools in the Central Collegiate Conference.
   C. In a before-and-after study, participants rated their views on abortion. Between conditions, they watched a propaganda movie.

3. A P. E. instructor designed a CVR (cardiovascular-respiratory) fitness program that gradually increased the duration and intensity of exercise over a 6-month period. The 9 participants in the program were measured for general fitness at the beginning and at the end of the program on such measures as percent body fat, vital lung capacity, and blood pressure and heart rate. Each participant was assigned a "fitness" score on a scale of 1 (a physical disaster) to 10 (superb specimen). A Wilcoxon matched-pairs signed-ranks *T* test produced a *T* value of 5.5. Was the program effective in improving CVR fitness?

4. A social psychologist examined the effect of sex education in the public schools. She randomly selected 25 college sophomore women and asked them if they had a sex education course in elementary school. She then asked them to complete a questionnaire concerning their sexual activity. Two refused to complete the questionnaire. She used the questionnaire results to rank-order the women from most (rank = 1) to least (rank = 23) sexually active. The sum of ranks for those who had not had a sex education course (N = 14) was 135. For those who had taken a course (N = 9), the sum was 141. The smaller *U* value was 30. Complete the analysis and write an interpretation regarding elementary school sex education courses.

5. Students face the practical question of how to use study time efficiently (that is, how to have more time to play). Gates conducted an early study (1917), which was used as a model for the data that follow. Twenty students were divided into five groups. The students all studied an article on dinosaurs, but each group spent a different proportion of the time in "self-recitation" (looking away from the article and mentally reciting what had been read). Afterward, each student took a 100-point test on the material in the article. The sums of the ranks are shown below. (The best score was ranked 1.) Finish the analysis and write an explanation that the data support.

Proportion of Study Time Spent in Self-Recitation				
0	20	40	60	80
73	54	44.5	27.5	11

6. Work Problem 1 that follows and then come back to this interpretation problem.

In addition to tests that simulated pilots' eye-hand coordination tasks, printed tests were used. A Spearman r_s between a general information test (with items such as, *name the five Great Lakes*) and the pilot competency test was .49. Compare this correlation coefficient to that in Problem 1 and draw a conclusion.

Problems

1. During WW II, enlistees were selected for pilot training on the basis of test scores. Some of the tests were printed, and some were eye-hand coordination tests that simulated tasks that pilots actually performed. One important eye-hand coordination test was the pursuit rotor, which requires tracking a target. The data that follow show time on target for the pursuit rotor during the fifth trial of practice and a "pilot competency score," which is based on actual flying skill. Find r_s (which will be the same as that found by Air Force researchers).

Seconds on Target Max. = 30	Pilot Competency Score
18	37
15	57
28	63
25	41
9	31
17	51
23	42
11	45

2. Many nurses are reluctant to approach patients diagnosed as having psychological problems. To help alleviate the problem, a Director of Nursing developed a training program to change nurses' attitudes toward such patients. To assess the effectiveness of the program, two matched groups were formed. Only one group completed the program, but both groups completed an attitude test. High scores indicate positive attitudes toward patients with psychological problems. Analyze the data with a Wilcoxon matched-pairs signed-ranks *T* test, and comment on the success or failure of the program.

	Attitude Scores	
Matched Pairs	Untrained	Trained
1	21	23
2	12	18
3	17	22
4	23	23
5	16	17
6	21	24
7	19	27
8	14	13

3. Karl Lashley carried out a 30-year search for the engram (the physical basis of memory). One of the many variables he studied was *place*. In one study, Lashley removed 40 percent of a rat's brain from either the frontal lobe or the parietal lobe. The following errors were produced by rats when their memory was tested. Analyze them with an appropriate nonparametric test. The sentence explanation you write will be a paraphrase of Lashley's conclusion.

Frontal	Parietal
54	47
31	30
18	12
26	17
39	21
42	25
45	35

4. A consumer advocate compared the cleanliness of four chains of supermarkets by devising a 50-point rating scale and inspecting six stores of each chain. The higher the score, the cleaner the store. Analyze the data, and make your shopping recommendations to the public.

Score	Chain
35	A
39	C
29	B
43	C
50	D
26	A
41	C
33	D
49	D
27	B
31	B
46	A
34	B
42	C
32	D

38	A
47	C
44	A
48	D
28	B
37	A
36	C
45	D
40	B

5. If you were interested in determining if there was a difference between two groups of students on ranked IQ tests, what test would you use? Assume you have the following data. Do the analysis and write an interpretation.

Group 1	Group 2
1	2
3	6
4	7
5	8
9	10