Name_____ Course Number:_____ Section Number:_____

Directions: Answer the questions and solve the problems in the spaces provided, or attach paper. Circle the correct choice for each response set. Where appropriate, use computer assistance.

Provide an appropriate response.

1) Describe parametric and nonparametric tests. Explain why nonparametric tests are important.

2) Define rank. Explain how to find the rank for data which repeats (for example, the data set: 4, 5, 5, 5, 7, 8, 12, 12, 15, 18).

3) Describe the runs test for randomness. What types of hypotheses is it used to test? Does the runs test measure frequency? What is the underlying concept?

4) Which of the following distribution–free tests has the lowest efficiency rating compared to its parametric counterpart?

A) rank correlation test
B) Wilcoxon signed–ranks test
C) Wilcoxon rank–sum test
D) Kruskal–Wallis test

5) The Kruskal–Wallis test statistic H has a distribution that can be approximated by which of the parametric distributions?

A) z distribution B) t distribution

C) F distribution D) chi–square distribution

Use the sign test to test the indicated claim.

6) An instructor gives a test before and after a lesson and results from randomly selected students are given below. At the 0.05 level of significance, test the claim that the lesson has no effect on the grade. Use the sign test.

Before	54	61	56	41	38	57	42	71	88	42	36	23	22	46	51
After	82	87	84	76	79	87	42	97	99	74	85	96	69	84	79

7) A researcher wishes to study whether music has any effect on the ability to memorize information. 88 randomly selected adults are given a memory test in a quiet room. They are then given a second memory test while listening to classical music. 68 people received a higher score on the second test, 19 a lower score, and 1 received the same score. At the 0.05 significance level, test the claim that the music has no effect on memorization skills.

Use the Wilcoxon signed–ranks test to test the claim that the matched pairs have differences that come from a population with a median equal to zero.

8) The systolic blood pressure readings of ten subjects before and after following a particular diet for a month are shown in the table. Use Wilcoxon's signed–ranks test and a 0.01 significance level to test the claim that the diet has no effect on systolic blood pressure.

Subject	A	B	C	D	E	F	G	H	I	J
Before	175	192	167	180	161	203	185	176	204	146
After	160	190	170	180	153	197	191	174	192	150

Use the Wilcoxon rank–sum test to test the claim that the two independent samples come from populations with equal medians.

9) Use the Wilcoxon rank–sum approach to test the claim that the sample student grade averages at two colleges come from populations with the same median. The sample data is listed below. Use a 0.05 level of significance, and assume that the samples were randomly selected.

College A	3.2	4.0	2.4	2.6	2.0	1.8	1.3	0.0	0.5	1.4	2.9
College B	2.4	1.9	0.3	0.8	2.8	3.0	3.1	3.1	3.1	3.5	3.5

Solve the problem.

10) The Wilcoxon signed–ranks test can be used to test the claim that a sample comes from a population with a specified median. The procedure used is the same as the one described in this section except that the differences are obtained by subtracting the value of the hypothesized median from each value.
The sample data below represent the weights (in pounds) of 12 women aged 20–30. Use a Wilcoxon signed–ranks test to test the claim that the median weight of women aged 20–30 is equal to 130 pounds. Use a significance level of 0.05. Be sure to state the hypotheses, the value of the test statistic, the critical values, and your conclusion.

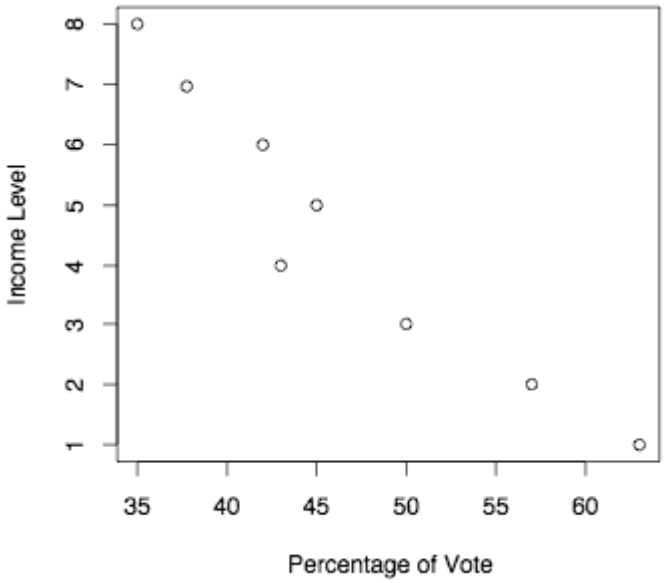
140	116	125	120	153	140
111	127	133	137	132	160

Use a Kruskal–Wallis test to test the claim that the samples come from populations with equal medians.

11) The table below shows the lifetimes (in hours) of random samples of light bulbs of three different brands. Use a 0.01 significance level to test the claim that the different brands have the same median lifetime.

Brand A	Brand B	Brand C
190	182	203
220	170	210
230	203	199
215	175	200
224	178	196
231	181	197

The following scatterplot shows the percentage of the vote a candidate received in the 2004 senatorial elections according to the voter's income level based on an exit poll of voters conducted by CNN. The income levels 1–8 correspond to the following income classes: 1=Under $15,000; 2=$15–30,000; 3=$30–50,000; 4=$50–75,000; 5=$75–100,000; 6=$100–150,000; 7=$150–200,000; 8=$200,000 or more.



12) Use the election scatterplot to the find the value of the rank correlation coefficient r_S.

A) $r_S = -1$ B) $r_S = 1.9762$ C) $r_S = -0.9762$ D) $r_S = 0.9762$

Find the critical value. Assume that the test is two–tailed and that n denotes the number of pairs of data.

13) n =30, $\alpha = 0.01$

A) 0.467 B) ±0.467 C) −0.467 D) ±0.362

Use the rank correlation coefficient to test for a correlation between the two variables.

14) Given that the rank correlation coefficient, r_S, for 76 pairs of data is −0.526, test the claim of correlation between the two variables. Use a significance level of 0.05.

15) Ten luxury cars were ranked according to their comfort levels and their prices.

Make	Comfort	Price
A	5	1
B	8	7
C	9	3
D	10	5
E	4	4
F	3	2
G	2	10
H	1	9
I	7	6
J	6	8

Find the rank correlation coefficient and test the claim of correlation between comfort and price. Use a significance level of 0.05.

16) A placement test is required for students desiring to take a finite mathematics course at a university. The instructor of the course studies the relationship between students' placement test score and final course score. A random sample of eight students yields the following data.

Placement Score	Final Course Score
38	63
90	41
95	54
51	32
86	93
74	60
60	61
57	89

Compute the rank correlation coefficient, r_S, of the data and test the claim of correlation between placement score and final course score. Use a significance level of 0.05.

Use the runs test to determine whether the given sequence is random. Use a significance level of 0.05.

17) A true–false test had the following answer sequence.
T T T T F T F T F T F T T F T
T T F F F F F F F T F T F T F
Test the null hypothesis that the sequence was random.

18) The sequence of numbers below represents the maximum temperature (in degrees Fahrenheit) in July in one U.S. town for 30 consecutive years. Test the sequence for randomness above and below the median.

94 96 97 99 95 90 97 98 100 100
92 95 98 99 102 97 97 101 99 100
98 95 93 99 101 99 101 100 99 103

19) Test the sequence of digits below for randomness of odd and even digits.

0 4 7 3 6 0 9 7 4 8
7 2 8 5 7 3 9 6 4 6
4 7 9 1 6 1 9 5 8 3
7 8 5 7 3 5 2 9 3 8

Solve the problem.

20) When performing a rank correlation test, one alternative to using the *Critical Values of Spearman's Rank Correlation Coefficient* table to find critical values is to compute them using this approximation:

$$r_S = \pm\sqrt{\frac{t^2}{t^2 + n - 2}}$$

where t is the t–score from the *t Distribution* table corresponding to n – 2 degrees of freedom. Use this approximation to find critical values of r_S for the case where n = 17 and $\alpha = 0.05$.

A) ±0.311 B) ±0.480 C) ±0.482 D) ±0.411

1) Parametric tests require assumptions about the nature or shape of the populations involved. Most of the tests we have worked with have required that the populations be normal. Nonparametric tests do not have requirements regarding the parent population. Nonparametric tests are important because they can be applied to ordinal data and nominal data and answer questions about medians of populations.
2) A rank is a number assigned to an individual sample item according to its order in the ranked list. (Ranked lists are arranged in order by some criterion such as smallest to largest.) For repeating data points, you find the mean of the ranks involved. For the data point 5 in the list above, the rank for each would be 3 (which is the average of ranks 2, 3, and 4). For the data point 12 in the list above, the rank for each would be 7.5 (which is the average of 7 and 8). So, corresponding to the given data, the ranks are 1, 3, 3, 3, 5, 6, 7.5, 7.5, 9, 10.
3) The runs test for randomness is a procedure for testing the randomness of data (with only two characteristics) using the concepts of runs. A run is a sequence of data that exhibit the same characteristic. For example, the data set M M M M M F F F M M F F F F has four runs. The null hypothesis is that the sequence is random and the alternate hypothesis is that the sequence is not random. The runs test is based only on the order in which the data occur; it does not test the frequency of the data. The underlying concept is that if the number of runs is very low or very high, randomness is lacking.
4) A
5) D
6) H_0: There is no difference between before and after grades.
   H_1: There is a difference between before and after grades.
   Test statistic: x = 0. Critical value: x = 2.
   Reject the null hypothesis of no difference. There is sufficient evidence to warrant rejection of the claim that the lesson has no effect on grade.
7) H_0: the music has no effect on memorization skills.
   H_1: the music has an effect on memorization skills.
   Convert x = 19 to the test statistic z = −5.15. Critical values: z = ±1.96.
   Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that music has no effect on memorization skills.
8) H_0: The diet has no effect on systolic blood pressure. H_1: The diet has an effect on systolic blood pressure. Test statistic T = 12.5. Critical value: T = 2.
   Fail to reject the null hypothesis that the population of differences has a median of zero.
9) H_0: Sample student grade averages at the two colleges come from populations with the same median. H_1: Sample student grade averages at the two colleges come from populations with different medians. $\mu_R = 126.5$, $\sigma_R = 15.2288$, $R_1 = 108.5$, $R_2 = 144.5$, z = −1.18.
   Test statistic: z = −1.18. Critical values z = ±1.96.
   Fail to reject the null hypothesis that the populations have the same median.
10) H_0: The sample comes from a population with a median of 130 pounds.
    H_1: The sample comes from a population with a median different from 130 pounds.
    Test statistic: T = 32.5
    Critical value: 14
    Do not reject the null hypothesis. There is not sufficient evidence to reject the hypothesis that the sample comes from a population with a median of 130 pounds.

11) H_0: The three brands come from populations which have the same median lifetime.
H_1: The three brands come from populations which don't have the same median lifetime.
Test statistic: H = 10.371. Critical value: $\chi^2 = 9.210$.
Reject the null hypothesis of equal medians. There is sufficient evidence to warrant rejection of the claim that the that the three brands come from populations which have the same median lifetime.
12) C
13) B
14) $r_s = -0.526$. Critical values: $r_s = \pm 0.226$.
Reject the null hypothesis $\rho_s = 0$. There appears to be a correlation between the two variables.
15) $r_s = -0.285$. Critical values: $r_s = \pm 0.648$.
Fail to reject the null hypothesis $\rho_s = 0$. There does not appear to be a correlation between comfort and price.
16) $r_s = -0.167$. Critical values: $r_s = \pm 0.738$.
Fail to reject the null hypothesis $\rho_s = 0$. There does not appear to be a correlation between placement score and final course score.
17) $n_1 = 15$, $n_2 = 15$, G = 18, 5% cutoff values: 10, 22.
Fail to reject the null hypothesis of randomness.
18) $n_1 = 15$, $n_2 = 15$, G = 10, 5% cutoff values: 10, 22.
Reject the null hypothesis of randomness.
19) $n_1 = 17$, $n_2 = 23$, G = 19, $\mu_G = 20.55$, $\sigma_G = 3.0494$.
Test statistic: z = −0.51. Critical values: z = ±1.96.
Fail to reject the null hypothesis of randomness.
20) C