Name: _____ **Class:** _____ **Date:** _____ **ID: A**

practice questions correlation and chi-squared

Multiple Choice

Identify the choice that best completes the statement or answers the question.

1. A college professor reports that students who finish exams early tend to get better grades than students who hold on to exams until the last possible moment. The correlation between exam score and amount of time spent on the exam is an example of
   a. a positive correlation.
   b. a negative correlation.
   c. a correlation near zero.
   d. a correlation near one.

2. Which of the following sets of correlations correctly shows the highest to lowest degree of relationship?
   a. –0.91, +0.83, +0.10, –0.03
   b. –0.91, +0.83, –0.03, –0.10
   c. +0.83, +0.10, –0.91, –0.03
   d. +0.83, +0.10, –0.03, –0.91

3. For a sample of n = 20 individuals, how large a Pearson correlation is necessary to be statistically significant for a two-tailed test with $\alpha = .05$?
   a. 0.444
   b. 0.433
   c. 0.423
   d. 0.413

4. As the sample size gets larger, the size of the correlation necessary for significance
   a. also gets larger.
   b. gets smaller.
   c. stays constant.
   d. There is no consistent relationship between sample size and the critical value for a significant correlation.

5. A set of X and Y scores has $M_X = 4$, $SS_X = 10$, $M_Y = 5$, $SS_Y = 40$, and SP = 40. What is the regression equation for predicting Y from X?
   a. $\hat{Y} = 0.25X - 4$
   b. $\hat{Y} = 4X + 4$
   c. $\hat{Y} = X - 6$
   d. $\hat{Y} = 4X - 11$

Student ID _____

6. A set of X and Y scores has $SS_X = 10$, $SS_Y = 20$, and $SP = 8$. What is the slope for the regression equation?
   a. 8/10
   b. 8/20
   c. 10/8
   d. 20/8

7. Which of the following is a characteristic of a *nonparametric* test?
   a. A numerical score is required for each individual.
   b. The hypotheses concern population means and variances.
   c. The test requires assumptions about the population means or variances.
   d. None of the other options is a characteristic of a nonparametric test.

8. In a chi-square test for goodness of fit, the null hypothesis is used to create an ideal, hypothetical sample. What term is used for the frequencies in the ideal sample?
   a. expected frequencies
   b. theoretical frequencies
   c. observed frequencies
   d. null frequencies

9. What is referred to by the term *expected frequencies*?
   a. the frequencies found in the sample data
   b. the frequencies found in the population being examined
   c. the frequencies computed from the null hypothesis
   d. the frequencies that are hypothesized for the population being examined

10. Which of the following accurately describes the observed and expected frequencies in a chi-square test?
    a. Both the observed and expected frequencies will always be whole numbers.
    b. The observed frequencies are always whole numbers but the expected frequencies may be decimals or fractions.
    c. The expected frequencies are always whole numbers but the observed frequencies may be decimals or fractions.
    d. Both the observed and expected frequencies may be decimals or fractions.

11. A researcher obtains a value of –8.50 for a chi-square statistic. What can you conclude because the value is negative?
    a. The observed frequencies are consistently larger than the expected frequencies.
    b. The expected frequencies are consistently larger than the observed frequencies.
    c. There are large differences between the observed and expected frequencies.
    d. The researcher made a mistake. The value of chi-square cannot be negative.

Student ID _____

12. Ten years ago, only 15% of the U.S. population consisted of people more than 65 years old. A researcher plans to use a sample of n = 300 people to determine whether the population distribution has changed during the past ten years. If a chi-square test is used to evaluate the data, what is the expected frequency for the older-than-65 category?
    a. 15
    b. 30
    c. 45
    d. 150

13. A chi-square test for goodness of fit is used to examine the distribution of individuals across four categories, and a chi-square test for independence is used to examine the distribution of individuals across the six categories in a 2×3 matrix of categories. Which test has the larger value for df?
    a. the test for goodness of fit
    b. the test for independence
    c. Both tests have the same df value.
    d. The df value depends on the sizes of the samples that are used.

14. What is evaluated by the chi-square test for goodness of fit?
    a. the relationship between two variables
    b. the mean differences between two or more treatments
    c. the shape or proportions for a population distribution
    d. None of the other options are evaluated by the chi-square test.

15. A researcher would like to test the claim that 9 out of 10 doctors prefer Brand X. A sample of 60 doctors is obtained and each is asked to compare Brand X with another leading brand. The data show that 48 of the doctors picked Brand X. If these data are evaluated using a chi-square test for goodness of fit, what is the expected frequency for Brand X?
    a. 9
    b. 12
    c. 48
    d. 54

16. If the expected frequencies for a chi-square test for independence are displayed in a matrix, then ____.
    a. All rows of the matrix have exactly the same frequencies.
    b. All rows of the matrix have exactly the same totals.
    c. The proportions for the categories in one row are exactly the same as the proportions in every other row.
    d. None of the other options accurately describes the matrix.

17. A chi-square test for independence has df = 3. What is the total number of categories (cells in the matrix) that were used to classify individuals in the sample?
    a. 4
    b. 6
    c. 8
    d. 9

Student ID _____

18. A chi-square test for independence has df = 2. What is the total number of categories (cells in the matrix) that were used to classify individuals in the sample?
    a. 2
    b. 3
    c. 4
    d. 6

Other

19. Describe what is meant by a *negative* relationship and a *positive* relationship between two variables.

20. Compute the Pearson correlation for the following data.

X	Y
2	0
3	1
7	4
5	6
4	6
3	1

21. A sample of n = 20 pairs of scores (X and Y values) produces a correlation of r = –0.46. Are these sample data sufficient to conclude that there is a significant non-zero correlation between X and Y in the population? Use a two-tailed test at the .05 level of significance.

Student ID _____

22. For the following data:

X	Y
4	1
7	16
3	4
5	7
6	7

a. Compute the Pearson correlation.
b. Find the regression equation for predicting Y from X.
c.

Student ID _____

23. Explain how the chi-square tests differ from parametric tests (such as t tests or ANOVA) with respect to the hypotheses, the data, and the assumptions underlying the test.

24. A researcher is interested in the relationship between IQ and longevity. A sample of n = 200 obituaries is selected from the newspaper and old school records are consulted to determine the individuals' IQ scores when they were between the ages of 10 and 20. The individuals were then classified into two IQ groups and three age groups at the time of death. The data are as follows. Based on these results, is there a significant relationship between IQ and longevity? Test with $\alpha = .05$.

	Age at the time of death		
	< 60	60 to 75	> 75
IQ < 110	41	63	36
IQ ≥ 110	9	27	24

Student ID _____

25. A researcher would like to know whether there is a consistent, predictable relationship between verbal skills and math skills for high school students. A sample of 200 students is obtained and each student is given a standardized English test and a standardized math test. Based on the test results, students are classified as high or low for verbal skills and for math skills. The results are summarized in the following frequency distribution:

	Verbal Skills	
	High	Low
High Math	50	30
Low Math	50	70

a. Based on these results, can the researcher conclude that there is a significant relationship between verbal skills and math skills? Test at the .05 level of significance.

practice questions correlation and chi-squared
Answer Section

MULTIPLE CHOICE

1. ANS: B PTS: 1 REF: p. 469 TOP: Section 15.1
2. ANS: A PTS: 1 REF: p. 469 TOP: Section 15.1
3. ANS: A PTS: 1 REF: p. 483 TOP: Section 15.4
4. ANS: B PTS: 1 REF: p. 483 TOP: Section 15.4
5. ANS: D PTS: 1 REF: p. 502 TOP: Section 15.6
6. ANS: A PTS: 1 REF: p. 502 TOP: Section 15.6
   MSC: www
7. ANS: D PTS: 1 REF: p. 522 TOP: Section 16.1
8. ANS: A PTS: 1 REF: p. 526 TOP: Section 16.2
9. ANS: C PTS: 1 REF: p. 526 TOP: Section 16.2
10. ANS: B PTS: 1 REF: p. 527 TOP: Section 16.2
11. ANS: D PTS: 1 REF: p. 528 TOP: Section 16.2
12. ANS: C PTS: 1 REF: p. 526 TOP: Section 16.2
13. ANS: A PTS: 1 REF: p. 539 TOP: Section 16.3
14. ANS: D PTS: 1 REF: p. 523 TOP: Section 16.2
    MSC: www
15. ANS: D PTS: 1 REF: p. 526 TOP: Section 16.2
16. ANS: C PTS: 1 REF: p. 535 TOP: Section 16.3
17. ANS: C PTS: 1 REF: p. 539 TOP: Section 16.3
18. ANS: D PTS: 1 REF: p. 539 TOP: Section 16.3
    MSC: www

OTHER

19. ANS:
    In a negative relationship there is a tendency for one variable to decrease as the other variable increases. The two variables tend to change in opposite directions. In a positive relationship, as one variable increases, the other also tends to increase. They tend to change in the same direction.

    PTS: 1
20. ANS:
    $SS_X = 16$, $SS_Y = 36$, $SP = 16$, $r = +0.67$.

    PTS: 1 MSC: www

21. ANS:
H_0: $\rho = 0$. For these data, df = 18 and the critical value for r from Table B.5 is 0.444. H_0 is rejected, $p < .05$.

PTS: 1

22. ANS:

a. For these data, $SS_X = 10$, $SS_Y = 126$, and SP = 30. The correlation is r = 0.845.

b. The regression equation is $\hat{Y} = 3X - 8$.

c. _____

PTS: 1

23. ANS:
As nonparametric tests, the chi-square tests state general hypotheses about the entire population without any reference to a specific population parameter. The data for a chi-square test consist of frequencies, but the data for a t test or ANOVA consist of scores that can be added, multiplied, squared, etc. Finally, the chi-square tests do not require any assumptions about population parameters. The t tests and ANOVA require normal populations and homogeneity of variance for tests with independent-measures designs.

PTS: 1

24. ANS:
The null hypothesis states that there is no relationship between IQ and longevity. With $\alpha = .05$ and df = 2, the critical value is 5.99. The expected frequencies are as follows:

	Age at the time of death		
	< 60	60 to 75	> 75
IQ < 110	35	63	42
IQ ≥ 110	15	27	18

For these data, chi-squared = 6.29. Reject H_0 and conclude that there is a significant relationship between IQ and longevity.

PTS: 1

25. ANS:

a. The null hypothesis states that there is no relationship between the two variables. With $\alpha = .05$ and df = 1, the critical value is 3.84. The expected frequencies are as follows:

	Verbal Skills	
	High	Low
High Math	40	40
Low Math	60	60

For these data, chi-squared = 8.33. Reject H_0. The data indicate a significant relationship between math skills and verbal skills.

PTS: 1 MSC: www