Questions on Topic Fifteen: Tests of Significance—Chi-Square and Slope of Least Squares Line

Multiple-Choice Questions

Directions: The questions or incomplete statements that follow are each followed by five suggested answers or completions. Choose the response that best answers the question or completes the statement.

1. A random sample of mice is obtained, and each mouse is timed as it moves through a maze to a reward treat at the end. After several days of training, each mouse is timed again. The data should be analyzed using

   (A) a z-test of proportions.
   (B) a two-sample test of means.
   (C) a paired t-test.
   (D) a chi-square test.
   (E) a regression analysis.

2. To test the claim that dogs bite more or less depending upon the phase of the moon, a university hospital counts admissions for dog bites and classifies with moon phase.

	New moon	First quarter	Full moon	Last quarter
Dog bite admissions	32	27	47	38

   Which of the following is the proper conclusion?

   (A) The data prove that dog bites occur equally during all moon phases.
   (B) The data prove that dog bites do not occur equally during all moon phases.
   (C) The data give evidence that dog bites occur equally during all moon phases.
   (D) The data give evidence that dog bites do not occur equally during all moon phases.
   (E) The data do not give sufficient evidence to conclude that dog bites are related to moon phases.

3. For a project, a student randomly picks 100 fellow AP Statistics students to survey on whether each has either a PC or Apple at home (all students in the school have a home computer) and what score (1, 2, 3, 4, 5) each expects to receive on the AP exam. A chi-square test of independence results in a test statistic of 8. How many degrees of freedom are there?

   (A) 1
   (B) 4
   (C) 7
   (D) 9
   (E) 99

4. A random sample of 100 former student-athletes are picked from each of the 16 colleges which are members of the Big East conference. Students are surveyed about whether or not they feel they received a quality education while participating in varsity athletics. A 2×16 table of counts is generated. Which of the following is the most appropriate test to determine whether there is a difference among these schools as to the student-athlete perception of having received a quality education.

   (A) A chi-square goodness-of-fit test for a uniform distribution
   (B) A chi-square test of independence
   (C) A chi-square test of homogeneity of proportions
   (D) A multiple-sample z-test of proportions
   (E) A multiple-population z-test of proportions

5. According to theory, blood types in the general population occur in the following proportions: 46% O, 40% A, 10% B, and 4% AB. Anthropologists come upon a previously unknown civilization living on a remote island. A random sampling of blood types yields the following counts: 77 O, 85 A, 23 B, and 15 AB. Is there sufficient evidence to conclude that the distribution of blood types found among the island population differs from that which occurs in the general population?

   (A) The data prove that blood type distribution on the island is different from that of the general population.
   (B) The data prove that blood type distribution on the island is not different from that of the general population.
   (C) The data give evidence at the 1% significance level that blood type distribution on the island is different from that of the general population.
   (D) The data give evidence at the 5% significance level, but not at the 1% significance level, that blood type distribution on the island is different from that of the general population.
   (E) The data do not give evidence at the 5% significance level that blood type distribution on the island is different from that of the general population.

6. Is there a relationship between education level and sports interest? A study cross-classified 1500 randomly selected adults in three categories of education level (not a high school graduate, high school graduate, and college graduate) and five categories of major sports interest (baseball, basketball, football, hockey, and tennis). The χ^2-value is 13.95. Is there evidence of a relationship between education level and sports interest?

   (A) The data prove there is a relationship between education level and sports interest.
   (B) The evidence points to a cause-and-effect relationship between education and sports interest.
   (C) There is evidence at the 5% significance level of a relationship between education level and sports interest.
   (D) There is evidence at the 10% significance level, but not at the 5% significance level, of a relationship between education level and sports interest.
   (E) The *P*-value is greater than .10, so there is no evidence of a relationship between education level and sports interest.

7. A disc jockey wants to determine whether middle school students and high school students have similar music tastes. Independent random samples are taken from each group, and each person is asked whether he/she prefers hip-hop, pop, or alternative. A chi-square test of homogeneity of proportions is performed, and the resulting *P*-value is below .05. Which of the following is a proper conclusion?

   (A) There is evidence that for all three music choices the proportion of middle school students who prefer each choice is equal to the corresponding proportion of high school students.
   (B) There is evidence that the proportion of middle school students who prefer hip-hop is different from the proportion of high school students who prefer hip-hop.
   (C) There is evidence that for all three music choices the proportion of middle school students who prefer each choice is different from the corresponding proportion of high school students.
   (D) There is evidence that for at least one of the three music choices the proportion of middle school students who prefer that choice is equal to the corresponding proportion of high school students.
   (E) There is evidence that for at least one of the three music choices the proportion of middle school students who prefer that choice is different from the corresponding proportion of high school students.

8. A geneticist claims that four species of fruit flies should appear in the ratio 1:3:3:9. Suppose that a sample of 2000 flies contained 110, 345, 360, and 1185 flies of each species, respectively. Is there sufficient evidence to reject the geneticist's claim?

   (A) The data prove the geneticist's claim.
   (B) The data prove the geneticist's claim is false.
   (C) The data do not give sufficient evidence to reject the geneticist's claim.
   (D) The data give sufficient evidence to reject the geneticist's claim.
   (E) The evidence from this data is inconclusive.

9. A food biologist surveys people at an ice cream parlor, noting their taste preferences and cross-classifying against the presence or absence of a particular marker in a saliva swab test.

	Presence	Absence
Vanilla	32	12
Chocolate	15	7
Strawberry	24	19

Is there evidence of a relationship between taste preference and the marker presence?

(A) At the 10% significance level, the data prove that there is a relationship between taste preference and the presence of the marker.
(B) At the 10% significance level, the data prove that there is no relationship between taste preference and the presence of the marker.
(C) There is evidence at the 5% significance level of a relationship between taste preference and the presence of the marker.
(D) There is evidence at the 10% significance level, but not at the 5% significance level, of a relationship between taste preference and the presence of the marker.
(E) There is not evidence at the 10% significance level of a relationship between taste preference and the presence of the marker.

10. Random samples of 25 students are chosen from each high school class level, students are asked whether or not they are satisfied with the school cafeteria food, and the results are summarized in the following table:

	Freshmen	Sophomores	Juniors	Seniors
Satisfied	15	12	9	7
Dissatisfied	10	13	16	18

Is there evidence of a difference in cafeteria food satisfaction among the class levels?

(A) The data prove that there is a difference in cafeteria food satisfaction among the class levels.
(B) There is evidence of a linear relationship between food satisfaction and class level.
(C) There is evidence at the 1% significance level of a difference in cafeteria food satisfaction among the class levels.
(D) There is evidence at the 5% significance level, but not at the 1% significance level, of a difference in cafeteria food satisfaction among the class levels.
(E) With $P = .1117$ there is not evidence of a difference in cafeteria food satisfaction among the class levels.

11. Can dress size be predicted from a woman's height? In a random sample of 20 female high school students, dress size versus height (cm) gives the following regression results:

```
The regression equation is
Size = -48.8 + 0.374 Height

Predictor      Coef    SE Coef        T        P
Constant    -48.81      30.57    -1.60    0.128
Height      0.3736     0.1898     1.97    0.065

S = 4.46720      R-Sq = 17.7%     R-Sq(adj) = 13.1%
```

Is there statistical evidence of a linear relationship between dress size and height?

(A) No, because r^2, the coefficient of determination, is too small.
(B) No, because 0.128 is above any reasonable significance level.
(C) Yes, because by any reasonable observation, taller women tend to have larger dress sizes.
(D) Yes, because the computer printout does give a regression equation.
(E) There is evidence at the 10% significance level, but not at the 5% level.

12. To study the relationship between calories (kcal) and fat (g) in pizza, slices of 14 randomly selected major brand pizzas are chemically analyzed. Computer printout for regression:

```
Predictor       Coef    SE Coef       T        P
Constant       1.593      1.422    1.12    0.285
Calories    0.035881   0.004896    7.33    0.000

S = 1.27515      R-Sq = 81.7%     R-Sq(adj) = 80.2%
```

What is measured by `S = 1.27515`?

(A) Variability in calories among slices
(B) Variability in fat among slices
(C) Variability in the slope (g/kcal) of the regression line
(D) Variability in the y-intercept of the regression line
(E) Variability in the residuals

Free-Reponse Questions

Directions: You must show all work and indicate the methods you use. You will be graded on the correctness of your methods and on the accuracy of your final answers.

Seven Open-Ended Questions

1. A candy manufacturer advertises that their fruit-flavored sweets have hard sugar shells in five colors with the following distribution: 35% cherry red, 10% vibrant orange, 10% daffodil yellow, 25% emerald green, and 20% royal purple. A random sample of 300 sweets yielded the counts in the following table:

	Cherry red	Vibrant orange	Daffodil yellow	Emerald green	Royal purple
Observed counts	94	34	22	77	73

Is there evidence that the distribution is different from what is claimed by the manufacturer?

4. A poll, asking a random sample of adults whether or not they eat breakfast and to rate their morning energy level, results in the table:

	Morning energy level		
	Low	Medium	High
Breakfast	22%	24%	24%
No breakfast	12%	10%	8%

(a) If the sample size was $n = 500$, is there evidence of a relationship between eating breakfast and morning energy level?

(b) Does the answer to part (a) change if $n = 1000$ instead? Explain.

6. A study of 100 randomly selected teenagers, ages 13–17, looked at number of texts per waking hour versus age, yielding the computer regression output below:

```
Predictor      Coef    SE Coef        T        P
Constant     -1.055      2.815    -0.37    0.709
Age          0.4577     0.1866     2.45    0.016

S = 2.66501      R-Sq = 5.8%      R-Sq(adj) = 4.8%
```

   (a) Interpret the slope in context.

   (b) What three graphs should be checked with regard to conditions for a test of significance for the slope of the regression line?

   (c) Assuming all conditions for inference are met, perform this test of significance.

   (d) Give a conclusion in context, taking into account both your answer to the hypothesis test as well as the value of R-Sq.