

**Multiple Choice Questions:**

Use the following for questions 1 – 3:

Do male and female children respond differently to colors? A study of color association in children asked separate random samples of male and female fourth-graders what emotion they associated with the color red. Here are the results for each group:

	Emotion				Total
	Anger	Happiness	Love	Pain	
Female	27	19	39	17	102
Male	34	12	38	28	112
Total	61	31	77	45	214

1. Which of the following would be the appropriate null hypothesis for this test?

- (a) The distribution of emotional associations with the color red is the same for male and female fourth-graders.
- (b) Gender is dependent upon emotional association with the color red.
- (c) Emotional associations with the color red are independent of gender.
- (d) The number of observations in each cell is the same for each emotional association.
- (e) 25% of all fourth graders associate the color red with each of the four listed emotions.

2. Under the assumption that the null hypothesis is true, which of the following represents the expected count for female children who associate the color red with love?

- (a) 39
- (b)  $\frac{(77)(214)}{102}$
- (c)  $\frac{(77)(102)}{214}$
- (d)  $\frac{(39)(102)}{77}$
- (e)  $\frac{39^2}{214}$

3. The chi-square statistic for these data is  $X^2 = 4.629$ . Which of the following intervals contains the  $P$ -value for this test?

- (a)  $0.005 \leq P\text{-value} \leq 0.01$
- (b)  $0.01 \leq P\text{-value} \leq 0.025$
- (c)  $0.025 \leq P\text{-value} \leq 0.05$
- (d)  $0.05 \leq P\text{-value} \leq 0.1$
- (e)  $P\text{-value} \geq 0.1$

4. Is the accident rate for some car colors different than for other car colors? An insurance company selects a random sample of cars that it insures and records their color (using five categories: white, silver, black, red, or “all others”) and whether or not they have been involved in an accident in the last three years. They perform a chi-square test of association and obtain a test statistics of  $X^2 = 8.474$ , which yields a  $P$ -value of 0.0758. Using a significance level of  $\alpha = 0.05$ , which of the following is the appropriate conclusion for this test?

- (a) Reject  $H_0$ : there is convincing evidence of an association between car color and proportion of cars involved in accidents.
- (b) Accept  $H_0$ : there is convincing evidence that car color and proportion of cars involved in accidents are independent.
- (c) Reject  $H_0$ : there is insufficient evidence to establish an association between car color and proportion of cars involved in accidents.
- (d) Fail to reject  $H_0$ : there is insufficient evidence to establish an association between car color and proportion of cars involved in accidents.
- (e) Fail to reject  $H_0$ : there is convincing evidence that car color and proportion of cars involved in accidents are independent.

5. A market research firm wants to know if there is an obvious leader among the five most popular manufacturers of laptop computers, or whether they are all about equally popular. They ask a random

Most-trusted brand	Apple	Dell	HP/Compaq	Sony	Toshiba
Frequency	42	66	45	43	54

sample of 250 people which brand of computer they trust most. Which of the following are conditions that must be met in order to test this hypothesis using a chi-square test?

- I. Each respondent is randomly selected from the population.
- II. All expected cell counts are greater than 5.
- III. The population is normally distributed, or  $n$  is large enough for the central limit theorem to apply.

- (a) I and II only
- (b) II and III only
- (c) I and III only
- (d) II only
- (e) I, II, and III

6. Do male and female toddlers show a preference for certain colors? Researchers offer identical toys of four different colors to separate randomly selected samples of male and female children aged 15 to 20 months and record which color toy is picked up first. Which of the following is the appropriate null hypothesis for this test?

		Color of first toy chosen				
		Red	Blue	Yellow	Green	Total
Female		48	24	18	30	120
Male		55	34	9	18	116
Total		103	58	27	44	236

- (a) The distribution of first toy chosen is 25% for each color, regardless of gender.
- (b) The number of observations in each cell is the same for color of first toy chosen.
- (c) The distribution of first toy chosen is the same for male and female 15-to-20-month-old children.
- (d) Color of first toy chosen is associated with gender.
- (e) First toy chosen is independent of gender.

7. A consumer advocacy group surveys a random sample of car owners to determine if there are differences in the types of problems encountered with new cars made by different manufacturers. Each subject was asked what kind of car they drove and to identify the first problem they had that required repair. Here are the data for three different car brands and five different types of repairs. For the chi-square test for these data, what are the correct degrees of freedom?

		Repair type				
		Brakes	Electrical	Drive train	Body mechanical	Body cosmetic
Car Brand	Hudson	5	12	10	8	22
Packard	9	6	19	8	10	
Studebaker	36	8	10	24	12	

- (a) 3
- (b) 5
- (c) 8
- (d) 14
- (e) 15

8. Let  $C$  be a number such that  $P(X^2 > C) = 0.05$  for the  $X^2$  distribution with 5 degrees of freedom. Which of the following is true?

- (a)  $P(X^2 > 2C) > P(X^2 > C)$
- (b)  $P(X^2 < -C) = P(X^2 > C)$
- (c)  $P(X^2 > C) < 0.05$  for the  $X^2$  distribution with 10 degrees of freedom
- (d)  $P(X^2 > (C+2)) < 0.05$
- (e)  $P(X^2 > C) = 0.05$  regardless of the degrees of freedom for the distribution

9. Children of three different ages—5 years, 7 years, and 9 years--were given a list of 5 different animals and asked which animal they would choose to be if they could be that animal for one day. A chi-square test for homogeneity on the data produced a test statistics of  $X^2 = 13.89$ , which yields a  $P$ -value of 0.0847. Using a significance level of  $X^2 = 0.05$ , which of the following is the appropriate conclusion for this test?

- (a) Reject  $H_0$ : there is convincing evidence that the distribution of animal choices is different for at least one of the age groups.
- (b) Accept  $H_0$ : there is convincing evidence the distribution of animal choices is the same for all three age groups.
- (c) Reject  $H_0$ : there is insufficient evidence to conclude that the distribution of animal choices is different for at least one of the age groups.

- (d) Fail to reject  $H_0$ : there is convincing evidence the distribution of animal choices is the same for all three age groups.  
 (e) Fail to reject  $H_0$ : there is insufficient evidence to conclude that the distribution of animal choices is different for at least one of the age groups.

**10.** A statistically-minded toll collector wonders if drivers are equally likely to choose each of the three lanes at his toll booth. He selects a random sample from all the cars that approach the booth when all three lanes are empty, so that the driver's choice isn't influenced by the cars already at the booth. Which of the following is the appropriate *alternative* hypothesis for addressing this question?

Lane	Left	Center	Right
Number of students	137	159	169

- (a) The observed number of cars choosing each lane is equal.  
 (b) The observed number of cars choosing each lane is different from the expected number of cars.  
 (c) The proportions of cars choosing each of the three lanes are equal.  
 (d) The proportions of cars choosing at least one of the lanes is different from the proportion choosing the other two lanes.  
 (e) The proportions of cars choosing each of the three lanes are all different.

**11.** A survey was conducted to investigate whether alcohol consumption and smoking are related. In an SRS of 300 smokers, 196 said they had consumed alcohol at least once in the past week. In an independent SRS of 300 non-smokers, 159 said they had consumed alcohol in the past week. If  $p_s$  is the proportion of smokers in the population who have had a drink in the past week and  $p_{ns}$  is the corresponding proportion of non-smokers, then a test of  $H_0: p_s - p_{ns} = 0$  against the two-sided alternative produces a test statistic of  $z = 3.07$  and a  $P$ -value of 0.002. If we had instead analyzed these results with a chi-square test of homogeneity, which of the following would be the test statistic and  $P$ -value?

- (a)  $X^2 = 9.42$ ;  $P$ -value = 0.002                      (b)  $X^2 = 9.42$ ;  $P$ -value = 0.004  
 (c)  $X^2 = 3.07$ ;  $P$ -value = 0.004                      (d)  $X^2 = 1.75$ ;  $P$ -value = 0.002  
 (e)  $X^2 = 1.75$ ;  $P$ -value = 0.004

**Free Response Questions:** (start on a new page or on the back of your current one):

**12.** OTL, a large national chain store, has one store in the city of Seireitei. One factor in deciding whether to build a second store in the city is whether the current store is serving all residents equally well, or whether unequal proportions of residents from different parts of town are using the store because it is located on one side of town. The Captain Commander, Genryusai Shigekuni Yamamoto, divides Seireitei into four geographical regions and determines the percentage of residents who live in each region. Here's what he finds:

Region	North	South	East	West
Percentage of population	40%	24%	22%	14%

Then the Captain Commander takes a simple random sample of 250 shinigami shoppers at Seireitei's OTL store and determines which part of town they come from by asking for their zip code when they are checking out:

Region	North	South	East	West
Number of shoppers	120	48	62	20

- a) Is Seireitei's OTL store used by a higher proportion of the residents in some parts of town than others? Support your conclusion with an appropriate statistical test.  
 b) Which region(s) of Seireitei has more shoppers to visit the store than others? Explain why.

**13.** A parent advisory board for a certain university was concerned about the effect of part-time jobs on the academic achievement of student attending the university. To obtain some information, the advisory board surveyed a simple random sample of 200 of the more than 20,000 students attending the university. Each student reported the average number of hours spent working part-time each week and his or her perception of the effect of part-time work on academic achievement. The data in the table below summarize the students' responses by average number of hours worked per week (less than 11, 11 to 20, more than 20) and perception of the effect of part-time work on academic achievement (positive, no effect, negative).

a) Perform an appropriate inference test for the association between the effect of part-time work on academic achievement and the average number of hours per week that student work. Be sure to identify the name of your method.

		Average Time Spent on Part-Time Jobs		
		Less Than 11 Hours per Week	11 to 20 Hours per Week	More Than 20 Hours per Week
Perception of the Effect of Part-Time Work on Academic Achievement	Positive Effect	21	9	5
	No Effect	58	32	15
	Negative Effect	18	23	19

b) Describe the type I and II errors of this setting.