# Freeport AP Statistics

Chapter 11 : Inference for Distributions of Categorical Data

**11.1** Chi-Square Goodness-of-Fit Tests

OBJECTIVE(S):

- Students will learn how to compute expected counts, conditional distributions, and contributions to the chi-square statistic.
- Students will learn how to check the Random, Large Sample Size, and Independent conditions before performing a chi-square test.
- Students will learn how to Use a chi-square goodness-of-fit test to determine whether sample data are consistent with a specified distribution of a categorical variable.
- Students will learn how to examine individual components of the chi-square statistic as part of a follow-up analysis.

1. What is the difference between Observed Counts and Expected Counts?

## Chi-Square Statistic

2. Jenny made a six-sided die in her ceramics class and rolled it 60 times to test if each side was equally likely to show up on top.
    a. Assuming that her die is fair, calculate the expected counts for each side in the table for b.

    b. Here are the results of her ceramic die and the expected counts. Calculate the value of the chi-square statistic

| Outcome | Observed | Expected | Observed-Expected | (Observed-Expected) $^2$ | (Observed-Expected) $^2$ /Expected |
|---|---|---|---|---|---|
| 1 | 13 | | | | |
| 2 | 11 | | | | |
| 3 | 6 | | | | |
| 4 | 12 | | | | |
| 5 | 10 | | | | |
| 6 | 8 | | | | |
| Total | | | | | |

3. Casinos are required to verify that their games operate as advertised. American roulette wheels have 38 slots – 18 red, 18 black, and 2 green. In one casino, managers record data from a random sample of 200 spins of one of their American roulette wheels. The one-way table below displays the results

| **Color:** | Red | Black | Green |
|---|---|---|---|
| **Count:** | 85 | 99 | 16 |

    a. State appropriate hypotheses for testing whether these data give convincing evidence that the distribution of outcomes on this wheel is not what it should be.

b. Calculate the expected counts for each color. Show your work.

c. Calculate the chi-square test statistic.

4. Mars, Inc., reports that their M&M's Peanut Chocolate Candies are produced according to the following color distribution: 23% each of blue and orange, 15% each of green and yellow, and 12% each of red and brown. Joey bought a bag of Peanut Chocolate Candies and counted the colors of the candies in his sample: 12 blue, 7 orange, 13 green, 4 yellow, 8 red, and 2 brown.

    a. State appropriate hypotheses for testing the company's claim about the color distribution of peanut M&M's.

a. Calculate the expected count for each color, assuming that the company's claim is true. Show your work.

b. Calculate the chi-square statistic for Joey's sample. Show your work.

**Chi-Square Distributions -**

5. How do we find the mean and the mode of the chi-square distribution?

6. When Jenny rolled her ceramic die 60 times and calculated the chi-square statistic, she got $\chi^2$=3.4. Using the appropriate degrees of freedom, calculate the *P*-value. What conclusion can you make about Jenny's die?

7. Let's continue our analysis of Joey's sample of M&M's Peanut Chocolate Candies from question 3.
    a. Confirm that the expected counts are large enough to use a chi-square distribution. Which distribution (specify the degrees of freedom) should we use?

    b. Find the *P*-value and sketch a graph that shows the *P*-value.

c. What conclusion would you draw about the company's claimed color distribution for M&M's Peanut Chocolate Candies? Justify your answer.

8. What is the **Large Counts condition** that takes the place of the Normal condition for *z* and *t* procedures?

9. How is calculating the degrees of freedom slightly different for a chi-square distribution compared to a *t* distribution?

# <u>The Chi-Square Goodness-of-Fit Test</u>

*Conditions:*

- **Random**
- **Large Counts**
- **10%**

**NOTE:** Unlike the other significance tests we have learned so far, the *P*-value for a chi-square goodness-of-fit test is always found by calculating the area *to the right* of the observed chi-square statistic. Because we are squaring the differences between the observed and expected counts, the chi-square statistic can never be negative, and bigger values of the chi-square statistic indicate bigger differences from the hypothesized distribution. Fortunately, **Table C** is already set up to find areas to the right.

When checking conditions, simply stating the expected counts are at least 5 is not sufficient – students must list the expected counts to prove they have actually checked them.

10. According to the 2000 census, of all U.S. residents aged 20 and older, 19.1% are in their 20s, 21.5% are in their 30s, 21.1% are in their 40s, 15.5% are in their 50s, and 22.8% are 60 and older. The table below shows the age distribution for a sample of U.S. residents aged 20 and older. Members of the sample were chosen by randomly dialing landline telephone numbers.

| **Category** | **Count** |
|---|---|
| 20-29 | 141 |
| 30-39 | 186 |
| 40-49 | 224 |
| 50-59 | 211 |
| 60+ | 286 |
| Total | 1048 |

Do these data provide convincing evidence that the age distribution of people who answer landline telephone surveys is not the same as the age distribution of all U.S. residents?

***P***arameters of interest

***H***ypothesis

$H_0$ :

$H_a$ :

***A***re the conditions met?

- **Random:**
- **Large Counts:**
- **10%:**

***N***ame of the test

***T***est of statistic

$$X^2 = \sum \frac{(O-E)^2}{E} =$$

*d.f.*=

***O***btain a *P*-value.

*P*-value =

***M***ake a decision about null

***S***tate your conclusion ( $H_a$ in context of the problem)

11. In his book *Outliers*, Malcolm Gladwell suggests that a hockey player's birth month has a big influence on his chance to make it to the highest levels of the game. Specifically, since January 1 is the cutoff date for youth leagues in Canada (where many National Hockey League [NHL] players come from), players born in January will be competing against players up to 12 months younger. The older players tend to be bigger, stronger, and more coordinated and hence get more playing time, more coaching, and have a better chance of being successful. To see if birth date is related to success (judged by whether a player makes it into the NHL), a random sample of 80 NHL players from the 2009-2010 season was selected and their birthdays were recorded. Overall, 32 were born in the first quarter of the year, 20 in the second quarter, 16 in the third quarter, and 12 in the fourth quarter. Do these data provide convincing evidence that the birthdays of NHL players are not uniformly distributed throughout the year?

***P***arameters of interest

***H***ypothesis

$H_0$ :

$H_a$ :

***A***re the conditions met?

- **Random:**

- **Large Counts:**

- **10%:**

***N***ame of the test

***T***est of statistic

$$X^2 = \sum \frac{(O-E)^2}{E} =$$

$d.f. =$

***O***btain a *P*-value.

*P*-value =

***M***ake a decision about null

***S***tate your conclusion ( $H_a$ in context of the problem)

12. Biologists wish to mate two fruit flies having genetic makeup RrCc, indicating that it has one dominant gene (R) and one recessive gene (r) for eye color, along with one dominant (C) and one recessive (c) gene for wing type. Each offspring will receive one gene for each of the two traits from both parents. The following table, often called a Punnett square, shows the possible combinations of genes received by the offspring.

**Parent 2 passes on**

| **Parent 1 passes on** | **RC** | **Rc** | **rC** | **rc** |
|---|---|---|---|---|
| **RC** | | | | |
| **Rc** | | | | |
| **rC** | | | | |
| **rc** | | | | |

Any offspring receiving an R gene will have red eyes; any offspring receiving a C gene will have straight wings. So based on this Punett square, the biologists predict a ratio of ___ red-eyed, straight-wing (x): ____ red-eyed, curly wing (y): ____ white-eyed, straight (z): ___ white-eyed, curly (w) offspring. In order to test their hypothesis about the distribution of offspring, the biologists mate a random sample of pairs of the fruit flies. Of 200 offspring, _____ had red eyes and straight wings, ____ had red eyes and curly wings, ____ had white eyes and straight wings, and ____ had white eyes and curly wings. Do these data differ significantly from what the biologists have predicted? Carry out a test at the $\alpha = 0.01$ significance level.

***P***arameters of interest

***H***ypothesis

$H_0$ :

$H_a$ :

***A***re the conditions met?

- **Random:**
- **Large Counts:**

| | |
|---|---|
| Red-eyed, straight wing: | 200(________) = |
| Red-eyed, curly-wing: | 200(________) = |
| White-eyed, straight wing: | 200(________) = |
| White-eyed, curly-wing: | 200(________) = |

Since all the expected cell counts are greater than ____, we can proceed with the test.

- **10%:**

***N***ame of the test

***T***est of statistic

$$X^2 = \sum \frac{(O-E)^2}{E} =$$

$d.f. =$

***O***btain a *P*-value.

*P*-value =

***M***ake a decision about null

***S***tate your conclusion ( $H_a$ in context of the problem)

13. How do you determine which component contributes the most to the chi-square statistic?

**DAY 2**

## **Freeport AP Statistics**

Chapter 11 : Inference for Distributions of Categorical Data

**11.2** Inference for Relationships

OBJECTIVE(S):

- Students will learn how to check the Random, Large Sample Size, and Independent conditions before performing a chi-square test.
- Students will learn how to use a chi-square test for homogeneity to determine whether the distribution of a categorical variable differs for several populations or treatments.
- Students will learn how to interpret computer output for a chi-square test based on a two-way table.
- Students will learn how to examine individual components of the chi-square statistic as part of a follow-up analysis.
- Students will learn how to show that the two-sample z test for comparing two proportions and the chi-square test for a 2-by-2 two-way table give equivalent results.
- Students will learn how to use a chi-square test of association/independence to determine whether there is convincing evidence of an association between two categorical variables.
- Students will learn how to distinguish between the three types of chi-square tests.

14. An article in the *Journal of the American Medical Association* (vol. 287, no. 14, April 10, 2002) reports the results of a study designed to see if the herb Saint-John's-wort is effective in treating moderately severe cases of depression. The study involved 338 subjects who were being treated for major depression. The subjects were randomly assigned to receive one of three treatments – Saint-John's-wort, Zoloft (a prescription drug), or a placebo-for an eight-week period. The table below summarizes the results of the experiment.

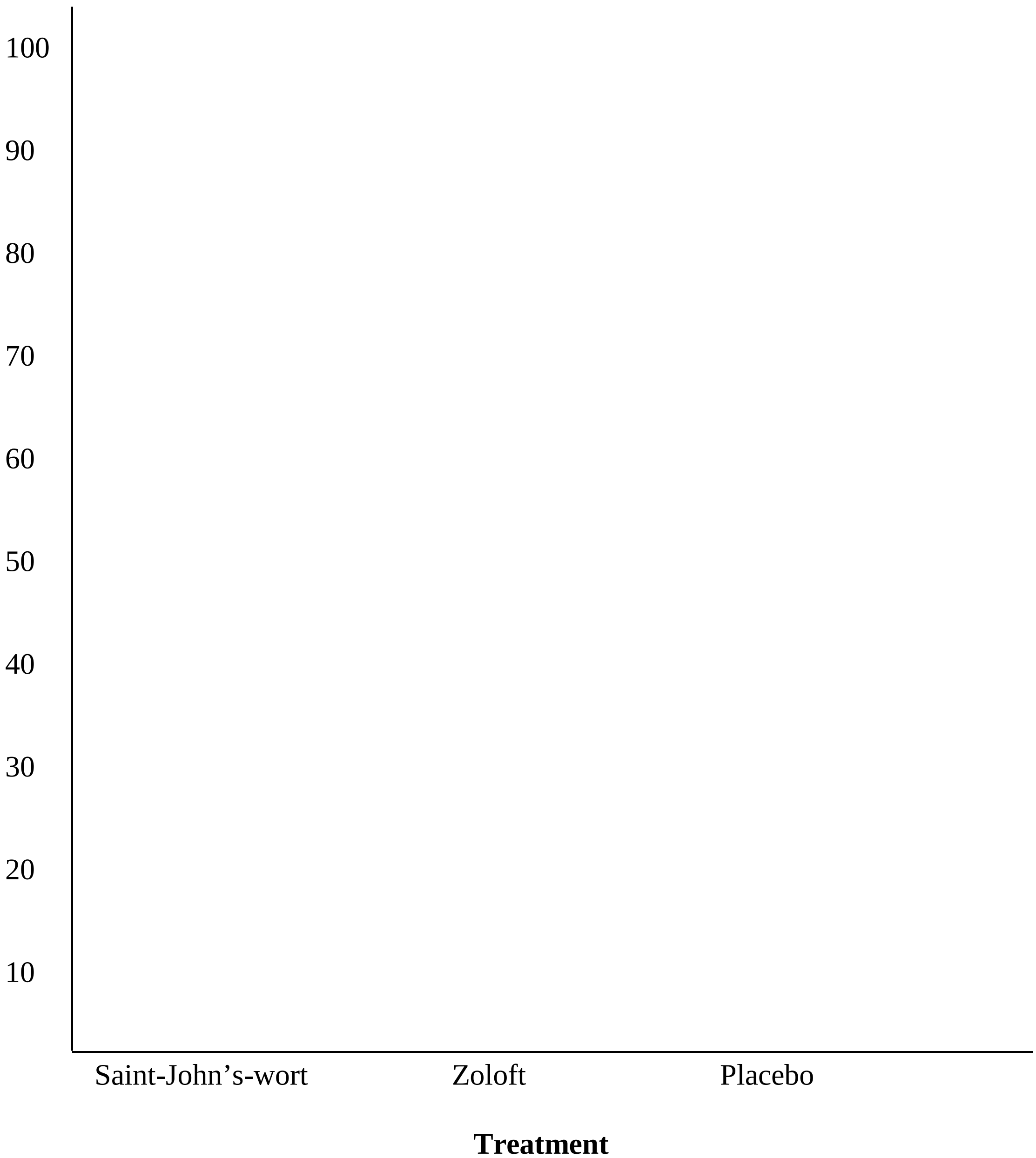
**Observed Counts**

| | **Saint-John's-wort** | **Zoloft** | **Placebo** | **Total** |
|---|---|---|---|---|
| Full response | 27 | 27 | 37 | 91 |
| Partial response | 16 | 26 | 13 | 55 |
| No response | 70 | 56 | 66 | 192 |
| Total | 113 | 109 | 116 | 338 |

a. Calculate the conditional distribution (in proportions) of the type of response for each treatment.

| | Saint-John's-wort treatment | Zoloft treatment | placebo treatment |
|---|---|---|---|
| Full: | | | |
| Partial: | | | |
| No: | | | |

b. Make an appropriate graph for comparing the conditional distributions in a.



c. Compare the distributions of response for each treatment.

Statistical methods for dealing with multiple comparisons usually have two parts:

-
-

d. Calculate the expected counts for the three treatments, assuming that all three treatments are equally effective.

**Expected Counts**

| | **Saint-John's-wort** | **Zoloft** | **Placebo** |
|---|---|---|---|
| Full response | | | |
| Partial response | | | |
| No response | | | |

e. Calculate the chi-square statistic. Show your work.

## Finding Expected Counts

## Chi-Square Test for Homogeneity

$H_o$ :

$H_a$ :

*Conditions:*

- **Random**
- **Large Counts**
- **Independent (10%)**

f. State the hypotheses.

$H_o$:

$H_a$:

g. Verify that the conditions for this test are satisfied.

- 
- 
- 

h. Calculate the *P*-value for this test.

i. Interpret the *P*-value in context?

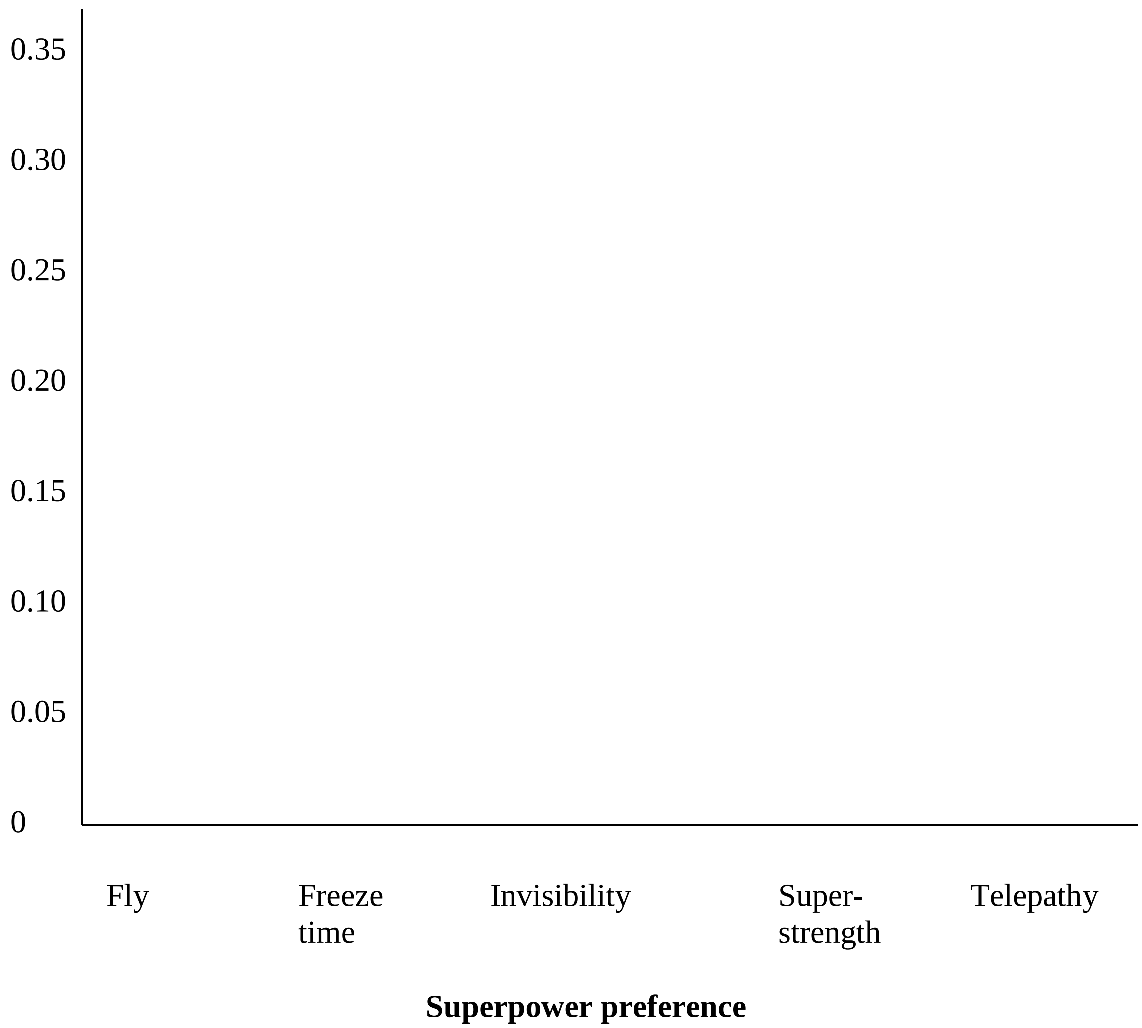
j. What is your conclusion?

15. Do American children have the same superpower preferences as children from the U.K. To find out, a random sample of 200 children from the U.K. and 215 children from the U.S. were selected (ages 9-17).

**Observed Counts**

| | **U.K.** | **U.S.** | **Total** |
|---|---|---|---|
| **Fly** | 54 | 45 | 99 |
| **Freeze time** | 52 | 44 | 96 |
| **Invisibility** | 30 | 37 | 67 |
| **Super strength** | 20 | 23 | 43 |
| **Telepathy** | 44 | 66 | 110 |
| **Total** | 200 | 215 | 415 |

a. Construct an appropriate graph to compare the distribution of superpower preference for U.K. and U.S. children



b. Do these data provide convincing evidence that the distribution of superpower preference differs among U.S. and U.K. children?

***P***arameters of interest

***H***ypothesis

$H_0$ :

$H_a$ :

***A***re the conditions met?

- 

- 

**Expected Counts**

| | **U.K.** | **U.S.** |
|---|---|---|
| **Fly** | | |
| **Freeze time** | | |
| **Invisibility** | | |
| **Super strength** | | |
| **Telepathy** | | |

- 

***N***ame of the test

***T***est of statistic

$$X^2 = \sum \frac{(O-E)^2}{E} =$$

$d.f. =$

***O***btain a *P*-value.

*P*-value =

***M***ake a decision about null

***S***tate your conclusion ( $H_a$ in context of the problem)

16. In a study reported by the Annals of Emergency Medicine (March 2009), researchers conducted a randomized, double-blind clinical trial to compare the effects of ibuprofen and acetaminophen plus codeine as a pain reliever for children recovering from arm fractures. There were many response variables recorded, including the presence of any adverse effect, such as nausea, dizziness, and drowsiness. Here are the results:

| | **Ibuprofen** | **Acetaminophen plus codeine** | **Total** |
|---|---|---|---|
| Adverse effects | 36 | 57 | 93 |
| No adverse effects | 86 | 55 | 141 |
| Total | 122 | 112 | 234 |

a. Explain why it was important to investigate this question with a randomized, double-blind clinical trial.

b. Is the difference between the two groups statistically significant? Conduct an appropriate chi-square test to find out.

***P***arameters of interest

***H***ypothesis

$H_0$ :

$H_a$ :

***A***re the conditions met?

- 

- 

**Expected Counts**

| | **Ibuprofen** | **Acetaminophen plus codeine** |
|---|---|---|
| Adverse effects | | |
| No adverse effects | | |

- 

***N***ame of the test

***T***est of statistic

$$X^2 = \sum \frac{(O-E)^2}{E} =$$

$d.f. =$

***O***btain a *P*-value.

*P*-value =

***M***ake a decision about null

***S***tate your conclusion ( $H_a$ in context of the problem)

**NOTE:** Anytime you conduct a chi-square test for a 2-by-2 table, you can also use a two-sample *z* test for a difference in proportions. Here are some things to keep in mind:

- The chi-square test is always two-sided. That is, it only tests for a difference in the two proportions. If you want to test whether one proportion is larger than the other, use the two-sample *z* test.
- If you want to estimate the difference between two proportions, use a two-sample *z* interval. There are no confidence intervals that correspond to chi-square tests.
- If you are comparing more than two treatments or the response variable has more than two categories, you must use a chi-square test.
- You can also use a chi-square goodness-of-fit test in place of a one-sample *z* test for a proportion if the alternative hypothesis is two-sided. The chi-square test will use two categories (success and failure) and have *df* = 2-1 = 1.

# The Chi-Square Test for Association/Independence

*Conditions:*

- **Random**
- **Large Counts**
- **Independent (10%)**

| | **Goodness-of-Fit Test** | **Chi-Square Test for Homogeneity** | **Chi-Square Test for Association/Independence** |
|---|---|---|---|
| **Number of Populations/Groups** | 1 | 2 or more | 1 |
| **Number of Categorical Variables** | 1 | 1 | 2 |

17. Is there an association between gender and allergies in the population of U.S. high school students who filled out the CensusAtSchool survey? Here is a two-way table that summarizes the sample data:

| | **Female** | **Male** | **Total** |
|---|---|---|---|
| Allergies | 10 | 8 | 18 |
| No allergies | 13 | 9 | 22 |
| Total | 23 | 17 | 40 |

a. Do these data provide convincing evidence of an association between gender and having allergies for U.S. high school students who filled out the CensusAtSchool survey?

***P***arameters of interest

***H***ypothesis

$H_0$ :

$H_a$ :

***A***re the conditions met?

- 

- 

**Expected Counts**

| | **Female** | **Male** |
|---|---|---|
| Allergies | | |
| No allergies | | |

- 

***N***ame of the test

***T***est of statistic

$$X^2 = \sum \frac{(O-E)^2}{E} =$$

$d.f. =$

***O***btain a $P$-value.

$P$-value =

***M***ake a decision about null

***S***tate your conclusion ( $H_a$ in context of the problem)

b. If the results were statistically significant, could we conclude that there is a gender determines whether an individual will have allergies?

18. An article in the *Arizona Daily Star* (April 9, 2009) included the following table:

| **Age (years):** | 18-24 | 25-34 | 35-44 | 45-54 | 55-64 | 65+ | **Total** |
|---|---|---|---|---|---|---|---|
| **Use online social networks:** | 137 | 126 | 61 | 38 | 15 | 9 | 386 |
| **Do not use online social networks:** | 46 | 95 | 143 | 160 | 130 | 124 | 698 |
| **Total:** | 183 | 221 | 204 | 198 | 145 | 133 | 1084 |

Suppose that you decide to analyze these data using a chi-square test. However, without any additional information about how the data were collected, it isn't possible to know which chi-square test is appropriate.

a. Explain how you know that a goodness-of-fit test is not appropriate for analyzing these data.

b. Describe how these data could have been collected so that a test for homogeneity is appropriate.

c. Describe how these data could have been collected so that a test for association/independence is appropriate.

# $\chi^2$ TEST

| | GOF | HOMOGENEITY | ASSOCIATION/INDEPENDENCE |
|---|---|---|---|
| P | | | |
| H | | | |
| A | | | |
| N | | | |
| T | | | |
| O | | | |
| M | | | |

| | | | |
|---|---|---|---|
| **S** | | | |
| | | | |
| **Which one to use?** | | | |