

# STT 422 Exam 1: Summer 2004 SOLUTIONS

NAME: \_\_\_\_\_

PID: \_\_\_\_\_

1. Table 1 contains data on the family structure (authoritarian or democratic) and political views (SDS represents students who belonged to Students for a Democratic Society, a left-wing group; YAF represents students who belonged to Young Americans for Freedom, a right-wing group.) Answer the following questions based on the data in the table.

Family Structure	Political Affiliation	
	SDS	YAF
Authoritarian	29	33
Democratic	131	78

Table 1: Data for Problem 1.

- (a) What proportion of students from an Authoritarian family structure belonged to YAF?

**Solution:**  $33/62 \approx 0.532$ .

□

- (b) What proportion of students belonged to SDS?

**Solution:**  $160/271 \approx 0.590$ .

□

- (c) What proportion of students who belonged to SDS were from an Authoritarian family structure?

**Solution:**  $29/160 \approx 0.181$ .

□

- (d) What proportion of students were from a Democratic family structure and belonged to YAF?

**Solution:**  $78/271 \approx 0.288$ .

□

2. A random sample of 1005 registered voters was selected to determine the relationship between gender and voting. Each person in the sample was classified based on gender and by whether he or she voted in the most recent presidential election. The data are given in Table 2. We'll analyze the data using a chi-squared test of independence.

	Male	Female
Voted	225	255
Did not vote	280	245

Table 2: Gender and voting behavior for Problem 2.

- (a) What is the value of the  $X^2$  statistic for testing whether gender and voting status are independent?

**Solution:**

$$X^2 = \frac{(225 - 241.19)^2}{241.19} + \frac{(255 - 238.81)^2}{238.81} + \frac{(280 - 263.81)^2}{263.81} + \frac{(245 - 261.19)^2}{261.19} \approx 4.18.$$

□

- (b) Give bounds on the p-value for testing whether gender and voting status are independent.

**Solution:** The p-value is between 0.025 and 0.05.

□

- (c) Would you reject the null hypothesis of independence at the level  $\alpha = 0.05$ ?

**Solution:** Yes, since the p-value is less than 0.05.

□

3. A researcher is interested in whether music preferences are different for Republicans and Democrats. He obtained a random sample of 450 Republicans and asked them to choose their favorite musical genre from among classical, jazz, rock, and folk. He did the same for an independent random sample of 500 Democrats. The data were read into SAS, and **proc freq** was used to perform a chi-squared test of homogeneity. The SAS program is in Figure 1, and the output from **proc freq** is in Figure 2.

```
data elmusic;
  input party $ music $ count;
  cards;
  Repub Classical 150
  Repub Jazz 50
  Repub Rock 200
  Repub Folk 50
  Dem Classical 175
  Dem Jazz 100
  Dem Rock 125
  Dem Folk 100
  ;

proc freq data = elmusic;
  weight count;
  tables party * music / chisq nocum norow nocol nopercnt expected;

run;
```

Figure 1: SAS program for Problem 3.

- (a) What is the value of the  $X^2$  statistic?

**Solution:** 50.0712.

□

- (b) Would you reject the null hypothesis that the distribution of musical preferences is the same for Democrats and Republicans at the level  $\alpha = 0.01$ ?

**Solution:** Yes, since the p-value is less than 0.01.

□

- (c) What proportion of Republicans in the study have Folk music as their favorite?

**Solution:**  $50/450 \approx 0.11$ .

□

## The FREQ Procedure

Frequency Expected
-----------------------

Table of party by music					
party	music				Total
	Classica	Folk	Jazz	Rock	
Dem	175 171.05	100 78.947	100 78.947	125 171.05	500
Repub	150 153.95	50 71.053	50 71.053	200 153.95	450
Total	325	150	150	325	950

## Statistics for Table of party by music

Statistic	DF	Value	Prob
Chi-Square	3	50.0712	<.0001
Likelihood Ratio Chi-Square	3	50.7367	<.0001
Mantel-Haenszel Chi-Square	1	14.6593	0.0001
Phi Coefficient		0.2296	
Contingency Coefficient		0.2238	
Cramer's V		0.2296	

Sample Size = 950

Figure 2: SAS output for Problem 3.

4. For each of the following study designs, indicate whether a test of independence or a test of homogeneity is appropriate.

- (a) A random sample of 310 high-school students is asked whether they read the newspaper at least once a week. A separate random sample of 141 middle school students is asked whether they read the newspaper at least once a week. The goal is to determine whether the proportion of newspaper readers is the same or different between middle and high school students.

**Solution:** Homogeneity.

□

- (b) A random sample of 1021 registered voters is asked whether they are Republicans or Democrats, and whether they plan to vote in the next election. There are 422 Republicans and 599 Democrats. The goal is to determine whether voting behavior and political affiliation are related.

**Solution:** Independence.

□