

Name Key  
 (Please ALSO write your name on the back of the last page.)

**Ag 400 - Quiz 3**  
**Fall 2006**

- A. A random sample of 30 Mudville residents was interviewed to determine how Mudville citizens view a variety of social issues. To ascertain whether the sample differs significantly from the total Mudville population in regard to the distribution of age, the following data were compiled.

<u>Age</u>	<u>Total Mudville Population</u>	<u>Sample (N=30)</u>	<u><math>f_c</math></u>
< 30 years	1,000	6	3
30-49 years	2,000	6	6
50-64 years	5,000	15	15
65 years and older	<u>2,000</u>	<u>3</u>	<u>6</u>
	10,000	30	

- (24) 1. Test the statistical significance of the difference between the sample and the total Mudville population in regard to age. Use the .05 level to determine statistical significance. Indicate the following:

- a. State the Null Hypothesis and the Alternative Hypothesis in words, not statistical symbols.

$H_0$ : the distrib of age in the total Mudville population is the same as the distrib of age in the population we sampled. ~~is the sample.~~

$H_A$ : not true.

- b. Show the necessary calculations.

$f_o$	$f_c$	$f_o - f_c$	
12	9	+3	$(9)/9 = 1.00$
15	15	0	0
3	6	-3	$(9)/6 = 1.50$
		0	<u>2.50</u>

- c. Indicate degrees of freedom.

2

- d. Report the critical value of chi-square at the .05 level.

5.99

- e. Reject or do not reject the Null Hypothesis.

Do Not Reject  $H_0$

- f. What is the probability that you have made a Type I error in (e) above?  
(Be as precise as you can be.)

$\alpha = 0$

- g. What is the probability that you have made a Type II error in (e) above?  
(Be as precise as you can be.)

$\beta > 0$

- h. State your conclusion precisely in terms of the problem. What can you conclude about the representativeness of the sample in regard to age?

Based on these data we cannot conc. that the sample is biased in regard to age.

(10) 2. Indicate whether each of following statements is true (T) or false (F) in terms of the above data. If any part of a statement is untrue, it should be marked false (F). Add comments if you wish to clarify your answers.

T a. The statistical unit in the above analysis is a Mudville resident.

F b. The independent variable in the above analysis is age; the dependent variable is number of persons.

T c. Only 10% of the sample members but 20% of the Mudville population are 65 years of age or older.

T d. The above test of statistical significance is what we have called a One Sample Chi Square or Goodness of Fit test.

F e. If the exact age in years of every person in the Mudville population and the sample were known, you could use a t-test for the difference between two independent sample means to test the statistical significance of the difference between the sample member's mean age and the mean age of all Mudville residents.

- B. Using data from the sample described in (A) above you compile the following table to ascertain whether residents of different ages express differing views about the use of lethal force by police. Each survey participant was asked whether he/she believed that police shootings of individuals involved in a felony crime is "generally justified," "seldom justified," or "never justified."

Use of Lethal Force	Age of Respondent					
	<50 Years	50-64 yrs.	65 years and over			
Never justified	10	2	0	1	15	
Seldom justified	8	4	5	0		
Generally justified	2	2	11	13	2	15
	12		18			

- (14) 1. Test the statistical significance of the relationship between age of respondent and opinions about the use of lethal force by police. Use the .05 level to determine statistical significance. Report the following:

- a. Show your calculations.

$f_o$	$f_e$	$f_o - f_e$	$ f_o - f_e  - .5$	$( )^2$	
10	6	4	3.5	12.25 / 6	2.04
2	6	4	3.5	12.25 / 6	2.04
5	9	4	3.5	12.25 / 9	1.36
13	9	4	3.5	12.25 / 9	1.36
$\frac{30}{30}$	$\frac{30}{30}$				<u>6.80</u>

- b. Report degrees of freedom and the critical value of chi-square at the .05 level.

1

- c. State your conclusion precisely in terms of the problem. What can you conclude about this relationship?

Reject a statement of no rel. between age and opinion.

Persons over 50 yrs of age were more likely than those < 50 yrs to report that the use of lethal force was generally justified (72% vs 17%)

	<50	50+
Nov S	83%	28%
Gen.	17%	72%

(20) 2. Indicate whether each of the following statements is true (T) or false (F) in terms of the above data. If any part of a statement is untrue, it should be marked false (F). Add comments, if you wish, to clarify your answers.

F a. The Null Hypothesis for the above test can be stated as follows:

$H_0$ : In the population, there is no difference between the ages of respondents and their stated opinions to the question of the use of lethal force as measured here.

F b. The relationship between age and opinions about police use of lethal force in the sample could be described as follows:

Persons 50 years of age and older were more likely to indicate that police use of lethal force during a felony crime was "generally justified" than they were to report that it was "seldom or never justified."

T c. In the sample, respondents who were less than 50 years of age were less likely than those 50 years and older to report that the use of lethal force was "generally justified."

F d. About 87% of those respondents 50 years of age and older felt that the use of lethal force was "generally justified" during a felony crime, compared with only 13% of those less than 50 years of age who believed such force was "generally justified."

T e. If Cramer's V were calculated for the above table, V would be greater than zero and less than +1.00.

T f. The statistic calculated in B1a above is an inferential statistic since it is used in making conclusions about the Null Hypothesis.

F g. The Alternative Hypothesis for the test in B1 above could be stated as follows:

$H_A$ : there is a significant relationship between age and opinion about the use of the lethal force.

T h. The "expected frequencies" in the Chi-square analysis in B1 above are the frequencies that would be "expected" in the cells of the table if there was no relationship between age and stated opinion about lethal force.

- F i. If one of the 9 cells in this 3x3 contingency table had no cases in it that cell could be deleted from the chi-square analysis without loss of information.
- F j. Although a chi-square for contingency is usually calculated using frequencies, the relationship is described using percentages, and chi-square could be calculated using the percentage of the total number of cases in each cell instead of frequencies.

- (32) C. In exit interviews with a random sample of voters leaving the polling peop;e were asked whether they voted for or against a local issue in Mudville. The attached output reports analysis of the relationship of the voter's political affiliation and whether they voted "no" or "yes" on this issue.

The following variables were used:

Political Affiliation

- 1 Republican
- 2 Democrat

Vote on Issue

- 1 Yes
- 2 No

Answer the following questions in terms of the attached output by indicating whether each statement is true (T) or false (F). If any part of a statement is untrue, it should be marked false (F).

T 1. The statistical unit here is a voter.

T 2. The independent variable in the problem, political affiliation, is measured by a two-category nominal scale.

F 3. The Null Hypothesis tested by the Chi-Square test reported on this output can be stated as follows:

$H_0$ : there is no relationship between Republicans and Democrats in how they voted on this issue.

F 4. The relationship between political affiliation and vote in this sample can be described as follows:

Republican voters were more likely to vote "no" than "yes" on this issue; Democrats were less likely to vote "yes" than "no."

F 5. Sixty-eight percent of the Democrats voted "yes" on this issue while 32% of the Republicans voted "yes."

F 6. The figures reported on this output are statistics, except for the Chi-Square values which are parameters because they provide information to make conclusions about the population.

- T 7. If you reject the Null Hypothesis of no relationship between political affiliation and vote on this issue, the probability that you have made a Type I error is .028.
- F 8. If you do not reject the Null Hypothesis of no relationship between political affiliation and vote on this issue, you are wrong and  $\beta = 1.00$ .
- F 9. The Case Processing Summary reports that 5 cases are "Missing." This means that 5 people refused to answer how they voted on this issue.
- F 10. The Chi-square value of 5.845 should be used here to test the statistical significance of the relationship between political affiliation and vote since none of the cells have expected frequencies of less than 5.0.
- T 11. The Chi-square value using Yates Correction for Continuity is smaller than the (uncorrected) Pearson Chi-Square for these data. The corrected Chi-square is never greater than uncorrected chi-square.
- F 12. The output reports the results of Fisher's Exact Test. This test should not be used here since none of the expected frequencies in the table is less than 5.0.
- T 13. A phi coefficient could correctly be used to measure the strength of the relationship between political affiliation and vote on this issue in the sample data.
- T 14. For this analysis, Phi = Cramer's V. This is true for any analysis where the number of rows and/or columns in the contingency table both equal 2.
- T 15. For this analysis, the phi coefficient is larger than the (uncorrected) contingency coefficient. The phi-coefficient is never smaller than the contingency coefficient for a 2x2 table.
- F 16. For this analysis, Cramer's V = .179.

```

CROSSTABS
  /TABLES=polit BY vote
  /FORMAT= AVALUE TABLES
  /STATISTIC=CHISQ CC PHI
  /CELLS= COUNT EXPECTED ROW COLUMN .

```

## Crosstabs

### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
POLIT * VOTE	150	96.8%	5	3.2%	155	100.0%

### POLIT \* VOTE Crosstabulation

			VOTE		Total
			1 No	2 Yes	
POLIT	1 republican	Count	73	8	81
		Expected Count	67.5	13.5	81.0
		% within POLIT	90.1%	9.9%	100.0%
		% within VOTE	58.4%	32.0%	54.0%
2 democrat	Count	52	17	69	
	Expected Count	57.5	11.5	69.0	
	% within POLIT	75.4%	24.6%	100.0%	
	% within VOTE	41.6%	68.0%	46.0%	
Total	Count	125	25	150	
	Expected Count	125.0	25.0	150.0	
	% within POLIT	83.3%	16.7%	100.0%	
	% within VOTE	100.0%	100.0%	100.0%	

### Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	5.845 <sup>b</sup>	1	.016		
Continuity Correction <sup>a</sup>	4.831	1	.028		
Likelihood Ratio	5.898	1	.015		
Fisher's Exact Test				.027	.014
Linear-by-Linear Association	5.806	1	.016		
N of Valid Cases	150				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 11.50.