Test 13A   AP Statistics   Name:

Directions: *Work on these sheets. A chi-square table appears at the end of this test.*

Part 1: Multiple Choice. *Circle the letter corresponding to the best answer.*

1. A chi-square goodness of fit test is used to test whether a 0–9 spinner is "fair" (i.e., the outcomes are all equally likely). The spinner is spun 100 times, and the results are recorded. Which member of the chi-square family of curves is used?
   (a) $\chi^2(8)$ (b) $\chi^2(9)$ (c) $\chi^2(10)$ (d) $\chi^2(99)$ (e) None of the above

2. A study of accident records at a large engineering company in England reported the following number of injuries on each shift for 1 year:

Shift:	Morning	Afternoon	Night
Number of injuries:	1372	1578	1686

   Is there sufficient evidence to say that the numbers of accidents on the three shifts are not the same? Test at the 0.05, 0.01, and 0.001 levels.
   (a) There is sufficient evidence at all three levels to say that the numbers of accidents on each shift are not the same.
   (b) There is sufficient evidence at the 0.05 and 0.01 levels but not at the 0.001 level.
   (c) There is sufficient evidence at the 0.05 level but not at the 0.01 or 0.001 levels.
   (d) There is sufficient evidence at the 0.001 level but not at the 0.01 or 0.05 levels.
   (e) There is insufficient evidence at any of these levels.

The next set of questions refers to the following situation:
A study was conducted to determine if the fatality rate depends on the size of the automobile. The analysis of accidents is as follows (with some values hidden):

```
DEATH                 SIZE
FREQUENCY|    m    |    s    |    l    |  TOTAL
---------+---------+---------+---------+
no       |   63    |   128   |   46    |   237
---------+---------+---------+---------+
yes      |   26    |   95    |   16    |   137
---------+---------+---------+---------+
TOTAL        89       223       62       374

STATISTICS FOR TABLE OF DEATH BY SIZE
STATISTIC                       DF     VALUE      PROB
-----------------------------------------------------
CHI-SQUARE                       *     8.663      *****
LIKELIHOOD RATIO CHI-SQUARE      *     8.838      *****
```

3. Under a suitable null hypothesis, the expected frequency for the cell corresponding to fatal type of accident and small size automobile is:
   (a) 81.69
   (b) 67.00
   (c) 61.43
   (d) 63.41
   (e) 59.72

4. Which of the following is NOT CORRECT?
   (a) The accidents were cross-classified by size of automobile and fatality status. Each accident was counted in one and only one cell.
   (b) The null hypothesis is that the fatality status is independent of the size of the automobile.
   (c) The alternative hypothesis is that there is no association between fatality status and size of automobile.
   (d) If all expected cell counts are greater than five, then the distribution of the test statistic is an approximate chi-square distribution.
   (e) If we reject the null hypothesis then we have proven that the size of the automobile affects the chances of a fatality.

5. The null hypothesis will be rejected at $\alpha=0.05$ if the test statistic exceeds:
   (a) 12.59
   (b) 7.81
   (c) 5.99
   (d) 3.84
   (e) 9.49

6. The approximate P-value is:
   (a) less than 0.005
   (b) between 0.005 and 0.010
   (c) between 0.010 and 0.025
   (d) between 0.025 and 0.050
   (e) between 0.050 and 0.100

Part 2: Free Response
Communicate your thinking clearly and completely.

7. An experiment in chicken breeding results in offspring having very curly, slightly curly, or normal feathers. If this is the result of a single gene system, then the proportions of offspring in the three phenotypes should be 0.25, 0.50, and 0.25 respectively. In one such experiment, 93 chickens were born. 20 had normal feathers, 50 had slightly curly feathers, and 23 had very curly feathers. Carry out a test to determine whether the genetic model seems to hold in this setting.

8. The nonprofit group Public Agenda conducted telephone interviews with three randomly selected groups of parents of high-school children. There were 202 black parents, 202 Hispanic parents, and 201 white parents. The sample survey asked the parents to respond to the statement, "A college education has become as important as a high school diploma used to be." Here are the counts of responses:

	Black parents	Hispanic parents	White parents
Strongly agree	154	144	125
Somewhat agree	27	37	50
Somewhat disagree	11	13	18
Strongly disagree	10	8	8
TOTAL	202	202	201

Write a brief report on the similarities or differences among the three groups of parents in their attitudes towards the importance of a college education. Include a graph or graphs and a test of significance.

9. In a telephone survey of 800 registered voters, the data are cross-classified both by gender of respondent and by respondent's opinion on an environmental bond issue.

	Bond issue	
	For	Against
Men	450	150
Women	160	40

We want to know whether there is good evidence that one's gender influences whether a person is for or against the bond issue. Use the chi-square test to answer this question. State the hypotheses, discuss conditions, perform calculations, and report your conclusions.

(1) b (2) a (3) a (4) e (5) c (6) c

(7) **Step 1:** The hypotheses are $H_0 : p_{vc} = 0.25, p_{sc} = 0.50, p_n = 0.25$ vs. H_a : at least one of these proportions is incorrect. **Step 2:** We must consider these offspring as an SRS from all possible offspring that could be produced. Assuming the null hypothesis is true, our expected counts are: Very curly: $93(0.25) = 23.25$, Slightly curly: $93(0.50) = 46.50$, Normal: $93(0.25) = 23.25$. Since all expected counts are at least 5, we are safe using chi-square goodness of fit procedures. **Step 3:** The test statistic is $\chi^2 = \sum (O-E)^2/E = 0.72$. With df = 2, the P-value is 0.698. Since the P-value is "large," we have no reason to doubt the proposed genetic model.

(8) **Step 1:** Our hypotheses are H_0 : the proportions of parents who fall in each "agreement category" for this statement are the same in all three ethnic groups and H_a : the proportions of parents who fall in each "agreement category" for this statement differ according to ethnic group. **Step 2:** We are told that the data come from three independent random samples. They may not be SRSs, however. Assuming the null hypothesis is true, our expected counts are:

	Black	Hispanic	White
Strongly agree	141.23	141.23	140.53
Somewhat agree	38.06	38.06	37.87
Somewhat disagree	14.02	14.02	13.95
Strongly disagree	8.68	8.68	8.64

Since all expected counts are at least 5, we can safely use the chi-square homogeneity of populations procedures. **Step 3:** The test statistic is $\chi^2 = 12.253$ and the degrees of freedom are df = 6. The P-value is less than 0.057. **Step 4:** There is borderline evidence of a difference among the three ethnic groups. However, the P-value is not low enough to reject H_0 at the $\alpha = 0.05$ level. So we conclude that the attitudes toward the importance of a college education may not differ among the three ethnic groups of parents.

(9) **Step 1:** Our hypotheses are H_0 : there is no association between gender and one's position on this bond issue vs. H_a: There *is* an association between gender and one's position on this bond issue.
Step 2: The matrix of expected counts is:

Expected	For	Against
Men	457.5	142.5
Women	152.5	47.5

Since all expected counts are at least 5, the chi-square test for association/independence is applicable. We don't know how the sample was obtained, so our results may not generalize to the population of interest. **Step 3:** The chi-square test statistic is $\chi^2 = 2.07$ and the degrees of freedom are df = 1. The P-value is 0.15. **Step 4:** There is insufficient evidence to reject H_0. We can't conclude that gender influences whether a person is for or against the bond issue.