

REVIEW

AP Statistics

χ^2 - Procedures

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.

- 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
- 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.

- There is sufficient evidence at all three levels to say that the numbers of accidents on each shift are not the same.
- There is sufficient evidence at the 0.05 and 0.01 levels but not at the 0.001 level.
- There is sufficient evidence at the 0.05 level but not at the 0.01 or 0.001 levels.
- There is sufficient evidence at the 0.001 level but not at the 0.01 or 0.05 levels.
- 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 FREQUENCY	SIZE			TOTAL
	m	s	l	
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	*****

- 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.

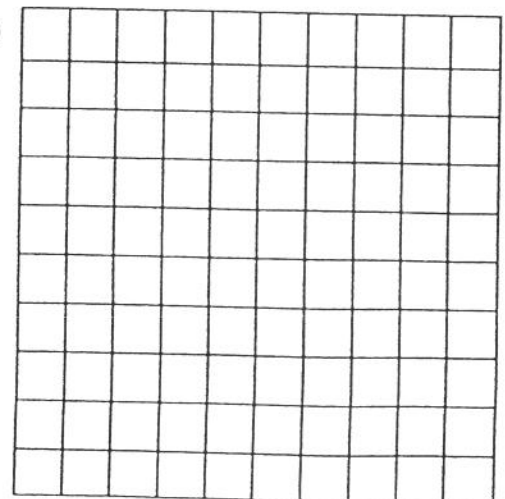
The Advanced Placement (AP) Statistics examination was first administered in May 1997. Students' papers are graded on a scale of 1–5, with 5 being the highest score. Over 7,600 students took the exam in the first year, and the distribution of scores was as follows (not including exams that were scored late).

Score	5	4	3	2	1
Percent	15.3	22.0	24.8	19.8	18.1

A distance learning class that took AP Statistics via satellite television had the following distribution of grades:

Score	5	4	3	2	1
Frequency	7	13	7	6	2

- 9.. Calculate marginal percents and make a bar graph of the population scores and the sample scores on the same grid, so that the two distributions can be compared visually.



10. Carry out an appropriate test to determine if the distribution of scores for students enrolled in the distance learning program is significantly different from the distribution of scores for all students who took the inaugural exam. (Don't forget to include all of the necessary steps.)

11. 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.

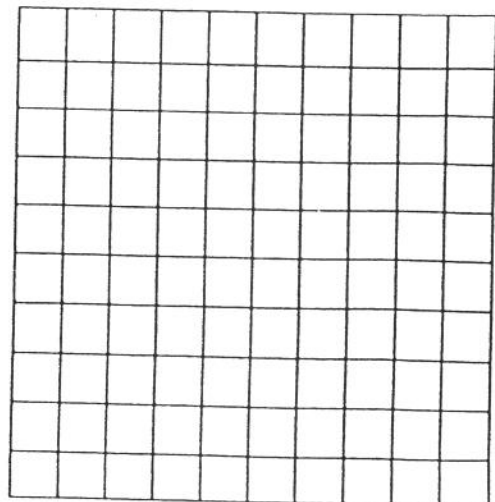
	Bond issue	
	For	Against
Men	450	150
Women	160	40

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.

The paper "Linkage Studies of the Tomato" (*Trans. Royal Canad. Inst. (1931)*) reported the accompanying data on phenotypes resulting from crossing tall cut-leaf tomatoes with dwarf potato-leaf tomatoes. We wish to investigate if the frequencies below are consistent with genetic laws which state the phenotypes should occur in the ratio 9:3:3:1.

	Phenotype			
	Tall Cut leaf	Tall Pot leaf	Dwarf Cut leaf	Dwarf Pot leaf
Frequency	926	288	293	104

12. State appropriate null and alternative hypotheses for investigating the genetic model.
13. Verify conditions for performing inference in this setting.
14. Construct an appropriate graphical display to compare the results of the experiment with the hypothesized values.
15. Calculate the test statistic and the P -value.



16. What conclusion would you draw?