Name:_____ Date:_____

AP Statistics CH14 Inference for Categorical Variables: Chi Square Test

(When calculating the χ^2, technology is recommended. Make sure to write out the equation:)

1. A Large box of "M and M" was opened to see if the distribution of each of the five colours were equal. The distribution of the colours was as follow:
   BLUE:520 , RED: 480 , GREEN: 410 YELLOW: 615 ORANGE: 410

a) What is the expected count for each colour?

b) State the hypothesis: Ho and Ha in the context of this problem

c) Perform a Chi Square test for Goodness of Fit. Find the χ^2 statistics, DF, and P value. Interpret your results in the context of this problem

2. Is there a discrepancy between students scores from classes taught by TA vs Professors? The grade distribution of 100 students in a UBC taught by TA's in the last 2 semesters and 92 students in UBC taught by a professor of the same course were as follow and the grades :

Grade of Distribution from TA

Grade	*A*	*B*	*C*	*D / F*
Prob.	0.32	0.41	0.20	0.07

Grade of Distribution from Professor

Grade	*A*	*B*	*C*	*D / F*
Count	19	40	23	10

a) If we used the TA's distribution as a fixed probability distribution, what would the expected count be for each grade for the professor's distribution?

b) Perform a Chi-square test for homogeneity. Does it give good evidence that the distribution of grades between the professor and TA's are different? State the Null and Alternative Hypothesis in context

c) Calculate the Chi Square statistics χ^2, DF, and P value

d) Is your statistics significant at the 5% level? State and interpret your results

3. A large SRS of recent HS graduates were asked which field of study they are planning to enter in University: Arts, STEM, Business, Med. Using the data provided, is there convincing evidence that preferred field of study differ between the male and females? Perform a ChiSq test for Homogeniety.

	Arts	*Stem*	*Bus.*	*Med*
Male	50	120	70	60
Female	110	77	82	55

a) Find the totals and expected value for each cell

b) State your Hypothesis

c) Calculate the Chi Square statistics χ^2 and P value

d) Is your statistics significant at the 5% level? State and interpret your results

4. The table shows the daily average time spend on looking at a computer screen and whether if they were glasses. The screen time is split into high volume , Medium volume, and low volume screen time.

	YES	*NO*
High	87	63
Medium	18	13
Low	31	22

a) What are the variables and association of interest? Which chi-sq test would you perform? Why?

b) State the Null and Alternative hypothesis

c) Calculate the chi-sq statistics for this test

d) State and interpret your results

5. Does living next to a wiring pole cause cancer? Researchers visited the addresses of Children living in Denver who had died from some form of cancer (Leukemia, lymphoma, or other cancers) and classified their addresses as either High Current configuration (HCC) vs Low Current configuration (LCC). Data, Expected counts, and Chi Sq are below:

	Leukemia	*Lymphoma*	*Other*	*Total*
HCC	52(49.97)	10(11.39)	17(17.64)	
LCC	84(86.03)	21(19.61)	31(30.36)	
Total				

a) What is the DF and P-value?

b) Interpret and state your conclusion:

$$\chi^2 = 0.435$$

c) Is there an association between the level of Current configuration and type of cancer in child cancer patients?

Q6: AP (2016)

Product advertisers studied the effects of television ads on children's choices for two new snacks. The advertisers used two 30-second television ads in an experiment. One ad was for a new sugary snack called Choco-Zuties, and the other ad was for a new healthy snack called Apple-Zuties.

For the experiment, 75 children were randomly assigned to one of three groups, A, B, or C. Each child individually watched a 30-minute television program that was interrupted for 5 minutes of advertising. The advertising was the same for each group with the following exceptions.

- The advertising for group A included the Choco-Zuties ad but not the Apple-Zuties ad.
- The advertising for group B included the Apple-Zuties ad but not the Choco-Zuties ad.
- The advertising for group C included neither the Choco-Zuties ad nor the Apple-Zuties ad.

After the program, the children were offered a choice between the two snacks. The table below summarizes their choices.

Group	Type of Ad	Number Who Chose Choco-Zuties	Number Who Chose Apple-Zuties
A	Choco-Zuties only	21	4
B	Apple-Zuties only	13	12
C	Neither	22	3

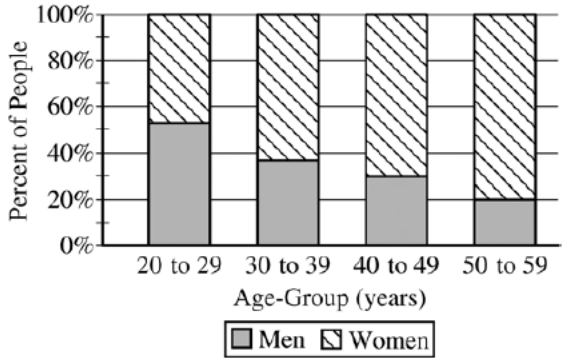
(a) Do the data provide convincing statistical evidence that there is an association between type of ad and children's choice of snack among all children similar to those who participated in the experiment?

(b) Write a few sentences describing the effect of each ad on children's choice of snack.

Q7: AP (2017)

The table and the bar chart below summarize the age at diagnosis, in years, for a random sample of 207 men and women currently being treated for schizophrenia.

	Age-Group (years)				
	20 to 29	30 to 39	40 to 49	50 to 59	Total
Women	46	40	21	12	119
Men	53	23	9	3	88
Total	99	63	30	15	207



Do the data provide convincing statistical evidence of an association between age-group and gender in the diagnosis of schizophrenia?

SOLUTION:

Solution to Q6: (2016)

Solution

Part (a):

Step 1: States a correct pair of hypotheses.

H_0 : The proportion of children who would choose each snack is the same regardless of which type of ad is viewed.

H_a : The proportion of children who would choose each snack differs based on which type of ad is viewed.

Step 2: Identifies a correct test procedure (by name or formula) and checks appropriate conditions.

The appropriate procedure is a chi-square test of homogeneity.

The conditions for this test are satisfied because (1) the question states that the children were randomly assigned to groups, and (2) expected counts for the six cells of the table are all at least 5, as seen in the following table that lists expected counts beside observed counts.

Group	Choco-Zuties	Apple-Zuties	Total
A	21 (18.67)	4 (6.33)	25
B	13 (18.67)	12 (6.33)	25
C	22 (18.67)	3 (6.33)	25
Total	56	19	75

Step 3: Calculates the appropriate test statistic and p-value.

The test statistic is calculated as $\chi^2 = \sum \frac{(O-E)^2}{E}$, which is

$$\chi^2 \approx 0.292 + 0.860 + 1.720 + 5.070 + 0.595 + 1.754 \approx 10.291.$$

The p-value is $P(\chi^2_{df2} \geq 10.291) \approx 0.006$.

Step 4: States a correct conclusion in the context of the study, using the result of the statistical test.

Because the p-value is very small (for instance, much smaller than $\alpha = 0.05$), we reject the null hypothesis at the 0.05 level (and at the 0.01 level). The data provide convincing statistical evidence that the proportions who would choose each snack differ based on which ad is viewed.

Part (b):

When neither ad was viewed, $\frac{22}{25}$ or 88 percent of the children chose Choco-Zuties whereas only 12 percent chose Apple-Zuties.

When the Choco-Zuties ad was viewed, 84 percent of the children chose Choco-Zuties, which was very similar to the percentage that chose them without viewing any ad. So watching the Choco-Zuties ad did not affect the snack choice very much.

When the Apple-Zuties ad was viewed, only $\frac{13}{25}$ or 52 percent of the children chose Choco-Zuties, and 48 percent chose Apple-Zuties. Watching the Apple-Zuties ad seemed to increase the proportion of children choosing Apple-Zuties.

Scoring

This question is scored in four sections. Section 1 consists of steps 1 and 2 in part (a); section 2 consists of step 3 in part (a); section 3 consists of step 4 in part (a); and section 4 consists of part (b). Sections 1, 2, 3 and 4 are scored as essentially correct (E), partially correct (P), or incorrect (I).

Section 1 is scored as follows:

Essentially correct (E) if the response includes the following three components:

1. Both hypotheses are stated correctly with at least one in context.
2. Identifies the correct test procedure (by name or by formula).
3. The technical conditions are checked (all expected counts are greater than or equal to 5).

Partially correct (P) if the response includes only two of the three components.

Incorrect (I) if the response correctly includes at most one of the three components.

Note for component 2: It is acceptable if the test is identified as a test of independence instead of a test of homogeneity.

Notes for component 3:

- The random assignment condition was stated so need not be explicitly mentioned.
- Stating the expected count condition is met is not sufficient for component 3. The condition must be checked by reporting the expected counts and either:
    - noting that all are greater than or equal to 5;

      OR

- noting that all are greater than or equal to 1 and at most 20 percent are less than 5.
- Noting that the smallest expected count is 6.33 and that it is greater than or equal to 5 is sufficient to satisfy this component.
- Component 3 is not satisfied if the expected counts are reported as integers.
- This component is not satisfied if the response includes any of the following inappropriate conditions:
  - The response implies that a random sample was taken, e.g., "SRS – check."
  - The response refers to independence of groups or independence of ads as a required condition.
  - The response indicates that a sample size greater than 30 ensures normality or the response implies normality as a condition.
- A response stating that children are independent can be ignored in the scoring of this component.

Section 2 is scored as follows:

Essentially correct (E) if the response includes the following three components:
1. Correct chi-square test-statistic
2. Correct degrees of freedom
3. Correct *p*-value

Partially correct (P) if the response includes only two of the three components.

Incorrect (I) if the response correctly includes at most one of the three components.

Notes:
- If the response makes an error in one calculation, future calculations are considered correct if they follow correctly from the initial miscalculation.
- A chi-square critical value approach is acceptable: The critical value for $\alpha = 0.05$ is $\chi^2 = 5.99$.

Section 3 is scored as follows:

Essentially correct (E) if the response provides a correct conclusion about the alternative hypothesis in context, *AND* provides justification based on linkage between the *p*-value and conclusion.

Partially correct (P) if the response provides a correct conclusion in context, but without justification based on linkage to the *p*-value;
OR
if the response provides a correct conclusion, with linkage to the *p*-value, but not in context;
OR
if the response provides a correct decision stated in terms of the null hypothesis in context, with linkage to the *p*-value, but no conclusion is made about the alternative hypothesis.

Incorrect (I) if the response does not meet the criteria for E or P.

Notes:

- If the conclusion is consistent with the *p*-value from section 2, and also in context with justification based on linkage to the *p*-value, section 3 is scored as E.
- If no alpha level is given, the solution must be explicit about the linkage by giving a correct interpretation of the *p*-value or by explaining how the conclusion follows from the *p*-value such as saying: "Because the *p*-value is small, we reject the null hypothesis." or "Because the *p*-value is large, we do not reject the null hypothesis."
- If a conclusion contains language that suggests that the response refers to the sample data, the conclusion component is not correct, unless the same error occurred in the statement of hypotheses in section 1.

Section 4 is scored as follows:

Essentially correct (E) if the response, in context, concludes that ad type A had little effect and ad type B had an effect, both supported by the observed proportions (or counts) from the table.

Partially correct (P) if the response correctly concludes that ad type A had little effect, and ad type B had an effect but does not provide correct numerical justification;
OR
if the response compares all of the proportions (or counts) as required, but without correctly concluding ad effectiveness;
OR
if the response correctly describes the effect of one of the ad types A or B (in context, with correct numerical justification) but not the other;
OR
if the response compares all of the proportions (or counts) as required, but not in context.

Incorrect (I) if the response does not meet the criteria for E or P.

Each essentially correct (E) section counts as 1 point, and a partially correct (P) section counts as ½ point.

4 **Complete Response**

3 **Substantial Response**

2 **Developing Response**

1 **Minimal Response**

If a response is between two scores (for example, 2½ points), use a holistic approach to decide whether to score up or down, depending on the strength of the response and communication.

SOLUTION Q7 (2017)

Intent of Question

The primary goal of this question was to assess a student's ability to identify, set up, perform, and interpret the results of an appropriate hypothesis test to address a particular question. More specific goals were to assess a student's ability to (1) state appropriate hypotheses; (2) identify the appropriate statistical test procedure and check appropriate conditions for inference; (3) calculate the appropriate test statistic and p-value; and (4) draw an appropriate conclusion, with justification, in the context of the study.

Solution

Step 1: State a correct pair of hypotheses.

The null hypothesis is that age group at diagnosis and gender are independent (that is, they are not associated) for the population of people currently being treated for schizophrenia.

The alternative hypothesis is that age group at diagnosis and gender are not independent for the population of people currently being treated for schizophrenia.

Step 2: Identify a correct test procedure (by name or formula) and check appropriate conditions.

The appropriate test is a chi-square test of independence.

The conditions for this test are satisfied because:

1. The question states that the sample was randomly selected.
2. The expected counts for the eight cells of the table are at least 5, as seen in the following table, with expected counts shown below observed counts.

	Age at Diagnosis				
	20 to 29	30 to 39	40 to 49	50 to 59	Total
Women	46	40	21	12	119
	56.91	36.22	17.25	8.62	
Men	53	23	9	3	88
	42.09	26.78	12.75	6.38	

Step 3: Find the value of the test statistic and the p-value.

The test statistic is calculated as $\chi^2 = \sum \frac{(O-E)^2}{E}$, or

$$\begin{aligned}\chi^2 &= 2.093 + 0.395 + 0.817 + 1.322 \\ &\quad + 2.830 + 0.534 + 1.105 + 1.788 \\ &= 10.884.\end{aligned}$$

The p-value is $P(\chi^2 \geq 10.884) = 0.012$, based on $(4-1) \times (2-1) = 3$ degrees of freedom.

Step 4: State the conclusion in context, with linkage to the p-value.

Because the p-value is very small (for instance much smaller than $\alpha = 0.05$), we reject the null hypothesis and conclude that the sample data provide strong evidence that there is an association between age group at diagnosis and gender for the population currently being treated for schizophrenia.

Scoring

This question is scored in three sections. Section 1 consists of steps 1 and 2 (stating the correct hypotheses, identifying the appropriate test procedure, and checking the technical conditions); section 2 consists of step 3 (performing the correct mechanics); and section 3 consists of step 4 (stating a correct conclusion with justification). Sections 1, 2, and 3 are scored as essentially correct (E), partially correct (P), or incorrect (I).

Section 1 is scored as follows:

Essentially correct (E) if the response correctly includes the following three components:

1. States BOTH hypotheses correctly with context included in at least one of them
2. Identifies a chi-square test of independence by name or formula
3. Verifies appropriate conditions that minimally include the condition for the expected counts and do not include any incorrect conditions (such as normality)

Partially correct (P) if the response includes only two of the three components.

Incorrect (I) if the response includes at most one of the three components.

Notes:

- Stating the expected count condition is not sufficient; the condition must be checked by reporting the expected counts, or minimally by showing that the smallest expected count is at least 5.
- The random sample condition was stated in the stem so need not be explicitly checked.
- If the null and alternative hypothesis are correctly stated in terms of population proportions, component 1 is satisfied. For example:

  $H_0 : p_1 = p_2 = p_3 = p_4$, where p_i is the population proportion of women at each indicated age group, 1, 2, 3, or 4, who are currently being treated for schizophrenia.

  H_a : At least one of the population proportions, p_1, p_2, p_3, p_4, differs from the other three.

  OR

  H_a : The population proportions for the four age groups are not all the same.

Section 2 is scored as follows.

Essentially correct (E) if the response correctly calculates the following two values:

1. The value of the chi-square test statistic
2. The p-value, critical value, or p-value range from chi-square table

Partially correct (P) if the response correctly calculates only one of the two values.

Incorrect (I) if the response does not satisfy the criteria for E or P.

Notes:

- If the response makes an error in one calculation, subsequent calculations are considered correct if they follow correctly from the initial miscalculation.
- With 3 degrees of freedom, the correct critical value is 7.81 for a significance level of 0.05 and 11.34 for a significance level of 0.01.
- Work does not have to be shown for calculations of test statistic or p-value. However, if incorrect work (other than minor arithmetic/transcription errors) is shown it is considered to be an incorrect calculation of the respective component, even if the correct value is given.
- If a response provides a test statistic that is not a chi-square test statistic, section 2 is scored I.

Section 3 is scored as follows.

Essentially correct (E) if the response includes the following three components:

1. A correct conclusion about the alternative hypothesis.
2. Justification of the conclusion based on linkage between the p-value and a reasonable alpha (or linkage between test statistic and critical value).
3. The conclusion is stated in context.

Partially correct (P) if the response includes only two of the three components.

Incorrect (I) if the response includes at most one of the three components.

Notes:

- If the response provides a correct decision, in context, with linkage to the p-value, but the decision is stated in terms of the null hypothesis with no conclusion about the alternative hypothesis, component 1 is not satisfied.
- Incorrect statistical statements are considered incorrect conclusions for the hypothesis test and do not satisfy component 1.
- If the conclusion is consistent with the p-value from section 2, and also in context with justification based on linkage to the p-value, section 3 is scored E.
- If no alpha level is given, the solution must be explicit about the linkage by giving a correct interpretation of the p-value or explaining how the conclusion follows from the p-value. For example, stating that because the p-value is small, we reject the null hypothesis or stating that because the p-value is large, we do not reject the null hypothesis.
- A decision about the null hypothesis (reject H_0 or fail to reject H_0) is not required, but if such a statement is given the scoring of the decision is considered in component 2.