

## 11.2

### Inference for Relationships

**\*\*Turn in 11.1 #12,16 in the tray by  
tomorrow\*\***

#### **Chi-Square Test for Homogeneity (Used when sampling from multiple populations)**

Suppose the Random, Large Sample Size, and Independent conditions are met. You can use the chi-square test for homogeneity to test:

$H_0$  : There is *no difference* in the distribution of a categorical variable for several populations or treatments.

$H_A$  : There is *a difference* in the distribution of a categorical variable for several populations or treatments.

Start by finding the expected counts and then calculate the chi-square statistic.

## Chi-Square Test for Association/Independence (Used when sampling from one population)

Suppose the Random, Large Sample Size, and Independent conditions are met. You can use the chi-square test for association/independence to test:

$H_0$ : There is *no association* between two categorical variables in the population of interest or two categorical variables are independent in the population of interest.

$H_A$ : There is *an association* between two categorical variables in the population of interest or two categorical variables are not independent in the population of interest.

### 2013 AP<sup>®</sup> STATISTICS FREE-RESPONSE QUESTIONS

4. The Behavioral Risk Factor Surveillance System is an ongoing health survey system that tracks health conditions and risk behaviors in the United States. In one of their studies, a random sample of 8,866 adults answered the question “Do you consume five or more servings of fruits and vegetables per day?” The data are summarized by response and by age-group in the frequency table below.

Age-Group (years)	Yes	No	Total
18–34	231	741	972
35–54	669	2,242	2,911
55 or older	1,291	3,692	4,983
Total	2,191	6,675	8,866

Do the data provide convincing statistical evidence that there is an association between age-group and whether or not a person consumes five or more servings of fruits and vegetables per day for adults in the United States?

State:  $H_0$ : There is not an association between age group and whether a person consumes 5+ servings of fruits/veg. per day amongst US adults

$H_A$ : There is an association between age group and whether a person consumes 5+ servings of fruits/veg per day amongst US adults.

Plan: I will perform a  $\chi^2$  test of association if conditions are met:

Random ✓ stated in problem

Independent ✓ assume all <sup>US</sup> adults independent and 88660 or more adults

Large Sample Size ✓ All expected counts are  $\geq 5$

(Display them in ( ) in original)

Age-Group (years)	Yes	No	Total
18-34	231	741	972
35-54	669	2,242	2,911
55 or older	1,291	3,692	4,983
Total	2,191	6,675	8,866

D<sub>0</sub>:  $\chi^2 = 8.98$   
 $p\text{-val} = 0.011$   
 $df = 2$

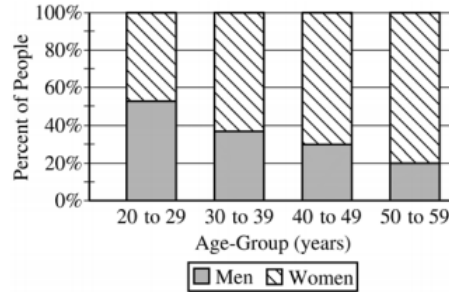
Conclude: Since  $p\text{-value} < \alpha$  (0.05), reject  $H_0$ .

There is sufficient evidence to conclude that there is an association b/tw age group and whether a person consumes 5+ servings of fruits/veg. per day amongst US adults

**2017 AP® STATISTICS FREE-RESPONSE QUESTIONS**

5. 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?

Stats:  $H_0$ : There is not an association b/tw age group and gender in the diagnosis of schizophrenia

$H_A$ : There is an association b/tw age group and gender in the diagnosis of schizophrenia

Plan: Perform a  $\chi^2$  test of association if conditions are met

Random ✓ stated in problem

Independent ✓ assume all patients are independent and at least 2070 patients

Large sample size ✓ all expected counts are  $\geq 5$

Do:  $\chi^2 = 10.88$

$p\text{-val} = 0.012$

$df = 3$

Conclude:

Since  $p\text{-value} < \alpha$   
(0.05), reject  $H_0$

There is sufficient evidence to conclude that there is an association b/tw age group and gender in the diagnosis of schizophrenia.

### 2003 AP<sup>®</sup> STATISTICS FREE-RESPONSE QUESTIONS

5. A random sample of 200 students was selected from a large college in the United States. Each selected student was asked to give his or her opinion about the following statement.

“The most important quality of a person who aspires to be the President of the United States is a knowledge of foreign affairs.”

Each response was recorded in one of five categories. The gender of each selected student was noted. The data are summarized in the table below.

	Response Category				
	Strongly Disagree	Somewhat Disagree	Neither Agree nor Disagree	Somewhat Agree	Strongly Agree
Male	10	15	15	25	25
Female	20	25	25	25	15

Is there sufficient evidence to indicate that the response is dependent on gender? Provide statistical evidence to support your conclusion.

State :  $H_0$ : Response and gender are independent ) For  
 $H_A$ : Response and gender are dependant ) students  
@  
this  
college

Plan: Perform  $\chi^2$  test of Independence if conditions  
are met:

Random ✓ stated

Large Sample Size ✓ all expected counts are  $\geq 5$

Independant ✓ assumed



Do :  $\chi^2 = 8.92$

p-val = 0.06

df = 4

conclude : Since p-val > 0.05, fail to reject  $H_0$   
there is not sufficient evidence to conclude that  
response and gender are dependent for students @  
this college.

## 2016 AP® STATISTICS FREE-RESPONSE QUESTIONS

2. 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.

State:  $H_0$ : There is not an association b/tw type of ad and children's choice of snack among all children similar to those who participated in the experiment

$H_A$ : There is an association b/tw type of ad and choice of snack among all children similar to those who participated in the exper

Plan :  $\neq$  will perform a  $\chi^2$  test of association if conditions are met

Random  $\checkmark$  Random assignment

Independent  $\checkmark$  Assumed

Large Sample Size  $\checkmark$  All expected counts are  $\geq 5$   
(show in table!)

Do :  $\chi^2 = 10.29$   
 $p\text{-val} = 0.006$   
 $df = 2$

Conclude: since  $p\text{-val} < \alpha$  (0.05), reject  $H_0$

There is sufficient evidence to determine that an association exists b/tw ad type and snack choice for all children similar to those who participated in the experiment.

### 2004 AP<sup>®</sup> STATISTICS FREE-RESPONSE QUESTIONS

5. A rural county hospital offers several health services. The hospital administrators conducted a poll to determine whether the residents' satisfaction with the available services depends on their gender. A random sample of 1,000 adult county residents was selected. The gender of each respondent was recorded and each was asked whether he or she was satisfied with the services offered by the hospital. The resulting data are shown in the table below.

	Male	Female	Total
Satisfied	384	416	800
Not Satisfied	80	120	200
Total	464	536	1,000

- (a) Using a significance level of 0.05, conduct an appropriate test to determine if, for adult residents of this county, there is an association between gender and whether or not they were satisfied with services offered by the hospital.
- (b) Is  $\frac{800}{1,000}$  a reasonable estimate for the proportion of all adult county residents who are satisfied with the services offered by this hospital? Explain why or why not.

State:  $H_0$ : There is not an association btw gender and satisfaction w/ hospital services for adult residents of this country

$H_A$ : There is an association btw gender and satisfaction w/ hospital services for adults of this country

Plan: Use a  $\chi^2$  test of association if conditions are met:

Random ✓ stated

Independent ✓ assume

Large Sample Size ✓ all expected counts are  $\geq 5$   
(Displayed in table)

Do:

$$\chi^2 = 4.12$$

$$p\text{-val} = 0.04$$

$$df = 1$$

Conclusion: Since  $p\text{-value} < \alpha$ , reject  $H_0$   
There is sufficient evidence to conclude an association  
exists b/w gender and satisfaction w/ hospital services  
for adults in this country.

## Attachments

---

11.2.gif