Categorical Data Analysis: Chi-Squared Tests

13.6 Testing Categorical Probabilities: Two-Way (Contingency) Table

1. The U.S. Department of Health reported the following results from a 2007 study of drug use in the nation. Use the data and a 1% significance level to test the claim that gender and drug use are independent.

Sex	Have Used	Have Not Used	Totals
Male	324	316	640
Female	259	361	620
Totals	583	677	1260

2. The following table includes the numbers of insured and uninsured Americans by race according to the U.S. Census 2008 Annual Social and Economic Supplement. Use the data and a 10% significance level to test the claim that race and health insurance are independent (note: the test statistic is 27,662.857).

In '000s	Race			
	White	Hispanic	Black	**Totals**
Insured	193,333	19,001	26,961	239,295
Uninsured	20,548	14,770	7,372	42,690
Totals	213,881	33,771	34,333	281,985

3. A paper published in 2008 looked at ethnic, gender, and acculturation influences on sexual behaviors. A total of 1,348 undergraduate students (429 men, 919women) at a large, public Southwestern university participated in this study for course credit in an introductory psychology course. The sample was composed of 67% Euro-American, 17% Hispanic, and 16% Asian participants. Participants ranged from 18 to 42 years old with a mean age of 19.03 for men (range, 18–32) and 18.79 for women (range, 18–42). As part of the study, men and women were asked, ''with how many partners have you had sexual intercourse, or oral sex, in your lifetime?'' The results for women are included below:

	Women			
Number of lifetime sexual partners	Euro-American	Hispanic	Asian	**Totals**
0	92	23	55	170
1	111	31	36	178
2 – 5	253	66	44	363
6 – 10	111	19	9	139
More than 10	49	17	3	69
Totals	616	156	147	919

Use a 5% significance level to test the claim that a woman's ethnicity and number of sexual partners she has had are independent (note: the test stat is 61.133).

Answers:

1. Gender and drug use appear to be independent:

H_0 : The two categories are independent.
H_a : The two categories are dependent.
$TestStat: \chi^2 = 9.923$
$CriticalValue = \chi^2_{0.01,1} = 6.635$

Conclusion: The data allows us to reject the claim of independence.

Expected values:

296.13	343.87
286.87	333.13

2. Race and whether or not someone has health insurance appears to be dependent:

H_0 : The two categories are independent.
H_a : The two categories are dependent.
$TestStat: \chi^2 = 27,662.857$
$CriticalValue = \chi^2_{0.10,2} = 4.605$

Conclusion: The data does allow us to reject the claim of independence.

Expected values:

181,501.335	28,658.373	29,135.292
32,379.665	5,112.627	5,197.708

3. A woman's ethnicity and the number of sexual partners she has had appear to be dependent:

H_0 : The two categories are independent.

H_a : The two categories are dependent.

$TestStat : \chi^2 = 61.133$

$CriticalValue = \chi^2_{0.05,8} = 15.507$

Conclusion: The data does allow us to reject the claim of independence.

Expected values:

113.95	28.86	27.19
119.31	30.22	28.47
243.32	61.62	58.06
93.171	23.60	22.23
46.25	11.71	11.04