Homework 4
Chi-Square
Answers

1. Assume that we are interested in the outcome of the election described above. 96 people were polled about who they were voting for. 42 said they were voting for Sleazy and 55 said they were voting for Slimy. Assuming that the group sizes are expected to be equal, use the goodness of fit test to determine whether the voting pattern significantly differs from chance.
A. Report your answer in APA format, report the highest level of significance achieved, and report whether the Null and Research Hypotheses should be rejected or not.
B. In your own words, explain what these result indicate.

	Sleazy	Slimy	Sum
O	41	55	96
E	48	48	96
O-E	-7	7	0
$(O-E)^2$	49	49	98
$(O-E)^2/E$	1.020833	1.020833	2.041667

With Rounding at 4th Digit= 2.0416
Without Rounding = 2.0216
$x^2(1, N = 96) = 2.04, p > .05, ns$
$x^2_{\text{critical } (p = .05, \text{ df} = 1)} = 3.84$; x^2 obt < x^2 crit = *ns.*
Reject the Alternative Hypothesis (H_A) & Fail to reject the Null Hypothesis (H_O)

B. The number people endorsing Sleazy or Slimy does not differ from what we would expect by chance. The rate at which each candidate is endorsed is essentially random.

2. If we double the sample size of problem 1 and find the same pattern of responding (84 for Sleazy and 110 for Slimy), what happens to the Chi-Square statistic?
A. Report your answer in APA format, report the highest level of significance achieved, and report whether the Null and Research Hypotheses should be rejected or not.
B. What does this suggest about the relationship between Sample Size and Significance?

	Sleazy	Slimy	Sum
O	82	110	192
E	96	96	192
O-E	-14	14	0
$(O-E)^2$	196	196	392
$(O-E)^2/E$	2.041667	2.041667	4.083333

With rounding at 4th digit= 4.0834
Without rounding = 4.0832
$x^2(1, N = 192) = 4.08, p < .05, ns.$
$x^2_{\text{critical } (p = .05, \text{ df} = 1)} = 3.84$; x^2 obt > x^2 crit.
Reject the Null Hypothesis (H_O) & Fail to reject the Alternative Hypothesis (H_A)

B. The chi square doubled in size and we found a significant difference between our observed and the expected frequencies. The larger the sample, the easier it is to find significance [given that the pattern (ratio) of frequencies remains the same].

3. In a different sample we found that 35 people supported Sleazy and 61 people supported Slimy. Again Assuming that the group sizes are expected to be equal, use the goodness of fit test to determine whether the voting pattern significantly differs from chance.
A. Report your answer in APA format, report the highest level of significance achieved, and report whether the Null and Research Hypotheses should be rejected or not.
B. In your own words, explain what these results indicate.

	Sleazy	Slimy	Sum
O	35	61	96
E	48	48	96
O-E	-13	13	0
$(O-E)^2$	169	169	338
$(O-E)^2/E$	3.5208	3.5208	7.0416

With rounding = 7.0417
Without rounding = 7.0416
$x^2(1, N = 96) = 7.04, p < .01$
$x^2_{\text{critical } (p = .05, \text{ df} = 1)} = 3.84$; x^2 obt > x^2 crit = significant.
$x^2_{\text{critical } (p = .01, \text{ df} = 1)} = 6.63$; x^2 obt > x^2 crit = significant.
Reject the Null Hypothesis (H_O) & Fail to reject the Research Hypothesis (H_A)

B. The number of people supporting Sleazy or Slimy is significantly different from what we would expect from chance alone. Specifically, there were fewer people supporting sleazy and more people supporting slimy than would be expected by chance.

4. Based on extensive sampling of the population, we are quite sure that 40% of the population supports Sleazy and 60% of the population supports Slimy. Use the goodness of fit chi-square to test the hypothesis that the sample collected in problem 3 (35 Sleazy; 61 Slimy) significantly differs from the general population.
A. Report your answer in APA format, report the highest level of significance achieved, and report whether the Null and Research Hypotheses should be rejected or not.
B. In your own words, explain what these results indicate.

	Sleazy	Slimy	Sum
O	35	61	96
E	38.4	57.6	96
O-E	-3.4	3.4	0
$(O-E)^2$	11.56	11.56	23.12
$(O-E)^2/E$	0.3010	0.2006	0.5017

With rounding = .5017
Without rounding = .5016
$x^2(1, N = 96) = .50, p > .05$
$x^2_{\text{critical } (p = .05, \text{ df} = 1)} = 3.84$; x^2 obt > x^2 crit = significant.
Reject the Alternative Hypothesis (H_A) & Fail to reject the Null Hypothesis (H_O)

B. Given that the population ratio is 40/60 (sleazy/slimy), the number of people supporting sleazy or slimy in our sample does not differ from what we expect in the population. Thus, our sample does not significantly differ from the population.

5. One critic, of the election research we have conducted thus far, notes that we only ask participants to chose between one of two candidates. To address this, with a new sample we have included the opportunity to support Screwy, the third party candidate, or to report that they are undecided. We find that 45 people support Sleazy, 48 support Slimy, 44 support Screwy, and 15 were undecided. Assuming that the group sizes are expected to be equal, use the goodness of fit test to determine whether the voting pattern significantly differs from chance.
A. Report your answer in APA format, report the highest level of significance achieved, and report whether the Null and Research Hypotheses should be rejected or not.
B. In your own words explain what these result indicate.

	Sleazy	Slimy	Screwy	Undecided	Sum
O	45	48	44	15	152
E	38	38	38	38	76
O-E	7	10	6	-23	17
$(O-E)^2$	49	100	36	529	149
$(O-E)^2/E$	1.2894	2.6315	0.9473	13.9210	18.7894

With rounding = 18.7896
Without rounding = 18.7892
$x^2(3, N = 152) = 18.79$, $p < .01$
$x^2_{\text{critical } (p = .05,\ df = 3)} = 7.81$; x^2 obt > x^2 crit = significant.
$x^2_{\text{critical } (p = .01,\ df = 3)} = 11.3$; x^2 obt > x^2 crit = significant.
Reject the Null Hypothesis (H_o) & Fail to reject the Alternative Hypothesis (H_A)

B. The number of people endorsing sleazy, slimy, or screwy, or report that they are undecided is significantly different from what we would expect by chance alone. Specifically, there are far fewer people who report being undecided than were expected.

6. With respect to the McDonald's cartoon, let's say we are now interested in the gender of the cow for the Angus breed in terms of visiting McDonald's. Our findings are reported in Table 1.
A. Using the Chi-square test of independence (Pearson's Chi-Square), does gender have a significant effect on visits to McDonald's? Report your answer in APA format, report the highest level of significance achieved, and report whether the Null and Research Hypotheses should be rejected or not.
B. Find Cramer's V (or phi). What % of the variance in trips to McDonald's is predicted/explained by gender of the cow? Is this a small, medium or large effect (according to Cohen's standards)?
C. In your own words explain what these result indicate.

Table 1 Visited McDonald's by Gender of Angus Cow

		Gender	
		Male	*Female*
Visited McDonald's	Yes	28	25
	No	32	31

			Male	Femal	row total
McYes		O	28	25	53
	E		27.4137	25.5862	
	O-E		0.5863	-0.5862	
	$O\text{-}E^2$		0.3437	0.3436	
	$O\text{-}E^2/E$		0.0125	0.0134	
McNo		O	32	31	63
	E		32.5862	30.4137	
	O-E		-0.5862	0.5863	
	$O\text{-}E^2$		0.3436	0.3437	
	$O\text{-}E^2/E$		0.0105	0.0113	
Colum Total			60	56	116

Chi-Square = 0.0477
V = 0.02 V^2 0.0004
k = 2

Without rounding = .0477

A. $x^2(1, N = 116) = .05$, $p > .05$, *ns*.
$x^2_{\text{critical } (p = .05, \text{ df} = 1)} = 3.84$; x^2 obt < x^2 crit = *ns*.
Reject the Alternative Hypothesis (H_A) & Fail to reject the Null Hypothesis (H_O)

B. $$V_{cramer} = \sqrt{\frac{\chi^2}{n(k-1)}} = \sqrt{\frac{.0476}{116(2-1)}} = \sqrt{\frac{.0476}{116(1)}} = \sqrt{\frac{.0476}{116}} = \sqrt{.0004} = .02$$

V^2_{cramer} = % variance accounted for (effect size) = .0004 or .04%
This is an exceptionally small effect.

C. The Rate at which male and female cow get sent to McDonald's does not significantly differ from chance. That is, there is no significant association between being one gender or another, and being sent to McDonald's.

7. With respect to the McDonald's cartoon, let's say we are now interested in the gender of the cow for a different breed (Holstein) in terms of visiting McDonald's. Our findings are reported in Table 2.
A. Using the Chi-square test of independence (Pearson's Chi-Square), does gender have a significant effect on visits to McDonald's? Report your answer in APA format, report the highest level of significance achieved, and report whether the Null and Research Hypotheses should be rejected or not.
B. Find Cramer's V (or phi). What % of the variance in trips to McDonald's is predicted/explained by gender of the cow? Is this a small, medium or large effect (according to Cohen's standards)?
C. In your own words explain what these result indicate.

Table 2 Visited McDonald's by Gender of Holstein Cow

		Gender	
		Male	*Female*
Visited McDonald's	Yes	29	10
	No	11	30

		Male	Femal	row total
McYes	O	29	10	39
	E	19.5	19.5	
	O-E	9.5	-9.5	
	$O-E^2$	90.25	90.25	
	$O-E^2/E$	4.6282	4.6282	
McNo	O	11	30	41
	E	20.5	20.5	
	O-E	-9.5	9.5	
	$O-E^2$	90.25	90.25	
	$O-E^2/E$	4.4024	4.4024	
Colum Total		40	40	80
Chi-Square		18.0612		
V		0.4750		V^2 0.2257
k 2				

With rounding = 18.0613
Without rounding = 18.0612

A. $x^2(1, N = 152) = 18.06, p < .01$
$x^2_{\text{critical } (p = .05, \text{ df} = 1)} = 3.84$; x^2 obt < x^2 crit = significant.
$x^2_{\text{critical } (p = .01, \text{ df} = 1)} = 6.63$; x^2 obt < x^2 crit = significant.
Reject the Null Hypothesis (H_o) & Fail to reject the Alternative Hypothesis (H_A)

B. $$V_{cramer} = \sqrt{\frac{\chi^2}{n(k-1)}} = \sqrt{\frac{18.0612}{80(2-1)}} = \sqrt{\frac{18.0612}{80(1)}} = \sqrt{\frac{18.0612}{80}} = \sqrt{.2257} = .4750$$

V^2_{cramer} = % variance accounted for (effect size) = .4750 or 47.5%
This is an exceptionally large effect.

C. Among Holstein cows, the rate at which male and female cows get sent to McDonald's does differ from chance alone. Specifically, male cows are over-represented among the cows that get sent to McDonald's while female cows are over-represented among the cows that do not "visit" McDonald's. Further, the association between cow gender and "visiting" McDonald's is quite strong.

8. With respect to the McDonald's cartoon, let's say we are now interested in the breed of cow in terms of visiting McDonald's. Our findings are reported in Table 3.
A. Using the Chi-square test of independence (Pearson's Chi-Square), does breed have a significant effect on visits to McDonald's? Report your answer in APA format, report the highest level of significance achieved, and report whether the Null and Research Hypotheses should be rejected or not.
B. Find Cramer's V (or phi). What % of the variance in trips to McDonald's is predicted/explained by gender of the cow? Is this a small, medium or large effect (according to Cohen's standards)?
C. In your own words explain what these result indicate.

Table 3	Visited McDonald's by Breed of Cow			
		Gender		
		Angus	*Holstein*	*Guernsey*
Visited McDonald's	Yes	5	8	5
	No	7	4	6

		Angus	Holstein	Guernsey	row total
McYes	O	5	8	5	18
	E	6.171429	6.171429	5.657143	
	O-E	-1.17143	1.828571	-0.65714	
	$O\text{-}E^2$	1.372245	3.343673	0.431837	
	$O\text{-}E^2/E$	0.222354	0.541799	0.076335	
McNo	O	7	4	6	17
	E	5.828571	5.828571	5.342857	
	O-E	1.171429	-1.82857	0.657143	
	$O\text{-}E^2$	1.372245	3.343673	0.431837	
	$O\text{-}E^2/E$	0.235434	0.573669	0.080825	
Colum Total		12	12	11	35

With rounding = 1.7303
Without rounding = 1.7301

A. $x^2(2, n = 24) = 1.73$, $p > .05$, *ns*.
$x^2_{\text{critical } (p = .05,\ df = 2)} = 5.99$; x^2 obt < x^2 crit = *non-ssinificant.*
Reject the Alternative Hypothesis (H_o) & Fail to reject the Null Hypothesis (H_A)

B. $$V_{cramer} = \sqrt{\frac{\chi^2}{n(k-1)}} = \sqrt{\frac{1.7301}{35(2-1)}} = \sqrt{\frac{1.7301}{35(1)}} = \sqrt{\frac{1.7301}{35}} = \sqrt{.0494} = .2222$$

V^2_{cramer} = % variance accounted for (effect size) = .0494 or 4.94%
This is a small to medium effect.

C. The frequency of visiting vs not visiting McDonald's, with respect to the breed of cow, does not significantly differ from the frequencies expected by chance alone. That is, the breed of cow does not significantly predict whether one visits McDonald's or not. However, given the moderate effect size this relationship is likely to be significant with a larger sample. There is a non-significant trend for Holsteins to be more likely to visit McDonald's, compared to the other breeds.