Chi Square Examples

Question 1 (11.13) A computer science major claimed to have written a program that randomly generates integers from 1 to 10. The program generates the following data. Use a 5% level of significance to test the claim.

Integers	Frequency
1-10	6
11-20	6
21-30	13
31-40	9
41-50	13
51-60	11
61-70	8
71-80	12
81-90	10
91-100	12

Chi-square Statistics - From table / Program

Enter alpha 0.05
Enter Degree of Freedom 9
1 - alpha = 0.95
Chi-square critical value 16.9189776 **(1- alpha)**

Enter chi-sq test statistics 6.4
p-value is 0.699312575

Chi Square Goodness-of-fit Test

chi-square= 6.4

	Observed	Expected p	Expected Freq.	Difference		
Criterion	(O)	(p)	(E)	O-E	(O-E)^2	(O-E)^2/E
10-Jan	6	0.1	10	-4	16	1.6
20-Nov	6	0.1	10	-4	16	1.6
21-30	13	0.1	10	3	9	0.9
31-40	9	0.1	10	-1	1	0.1
41-50	13	0.1	10	3	9	0.9
51-60	11	0.1	10	1	1	0.1
61-70	8	0.1	10	-2	4	0.4
71-80	12	0.1	10	2	4	0.4
81-90	10	0.1	10	0	0	0
91-100	12	0.1	10	2	4	0.4
Totals	100	1	100	0	64	6.4

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$ = 6.4

k= 10
(k-1)=d.f. 9

Chi Square Examples

Question 2 (11.19) A group of employees in a large firm claimed that gender and yearly salary are dependent. For the following data, test the claim using a 10% level of significance.

	Salary (thousands of dollars)			
	Less than 30	At least 30 but less than 35	At least 35 but less than 40	At least 40
Females	*12*	*30*	*20*	*13*
Males	*7*	*26*	*31*	*27*

Chi Square Statistics

Enter alpha 0.1

Enter Degree of Freedom 3

1 - alpha = 0.9

Chi-square critical value 6.251388746 (1- alpha)

Enter chi-sq test statistics **7.4**

p-value is 0.060184327

Expected Frequency Contingency Table (row X column)/Grand Total

		Criterion A			Row Freq. Total		
Criterion B	1	2	3	4			
1	8.584337349	25.30120482	23.04216867	18.07228916	75		
2	10.41566265	30.69879518	27.95783133	21.92771084	91		
Column Totals	19	56	51	40	166	Grand Total	

Independent Characteristics Table: p(mk)=p(m) x p(k)

		Criterion A			Row Totals	Probability, Row	
Criterion B	1	2	3	4			
1	0.051712876	0.152416897	0.138808245	0.108869212	75	*0.4518*	
2	0.062744956	0.184932501	0.168420671	0.132094644	91	*0.5482*	
Column Totals	19	56	51	40	166	1	Sum(pm)
Prob.Col	*0.114457831*	*0.337349398*	*0.307228916*	*0.240963855*		1	Sum(pk)

Chi-Square Statistics Ta (O-E)^2/E)

		Criterion A			Row Totals		
Criterion B	1	2	3	4			
1	1.359074192	0.872633391	0.401645799	1.42362249	4.056975871		
2	1.120116092	0.719203344	0.331026757	1.173315239	3.343661432		
Column Totals	2.479190283	1.591836735	0.732672556	2.596937729	7.4006373	chi-square stat	

Chi Square Examples

Question 3 (11.26 - 11.22) Former patients from 3 treatment centers were sampled and asked whether they were satisfied with the care they received in these centers. The results are as follows: Significance level of 5%.

	Are you satisfied with treatment?	
	Yes	**No**
Center A	***54***	***21***
Center B	***58***	***32***
Center C	***42***	***38***

Chi Square Statistics

Enter alpha 0.05

Enter Degree of Freedom 2

1 - alpha = 0.95

Chi-square critical value 5.991464547 (1- alpha)

Enter chi-sq test statistics **6.4581**

p-value is 0.039595096

3 x 2 - Chi-Square Test for Independent Charateristics					
chi-square =	**6.458139083**				
Observed Frequencies Contingency Table					
		Criterion A		Row Totals	
Criterion B	1	2	3		
1	*54*	*58*	*42*	154	
2	*21*	*32*	*38*	91	
Column Totals	75	90	80	245	Grand Total
Expected Frequency Contingency Table (row X column)/Grand Total					
		Criterion A		Row Totals	
Criterion B	1	2	3		
1	*47.14285714*	*56.57142857*	*50.28571429*	154	
2	*27.85714286*	*33.42857143*	*29.71428571*	91	
Column Totals	75	90	80	245	Grand Total
Chi-Square Statistics Table: S (O-E)^2/E)					
		Criterion A		Row Totals	
Criterion B	1	2	3		
1	*0.997402597*	*0.036075036*	*1.36525974*	2.398737374	
2	*1.687912088*	*0.061050061*	*2.31043956*	4.059401709	
Column Totals	2.685314685	0.097125097	3.675699301	6.4581391	chi-sq stat

Chi Square Examples

Question 4. Compute and interpret the **effect size** for Question 3 above.

Question 3 Summary: ***n = 245***, Number of Columns, ***C = 3***, Number of Rows, ***R = 2***, and $\chi^2 = 6.4581$.

Effect size measure for Chi-square analysis is a r^2 measure. Basically, r^2 measures report the percentage of variation in one variable that can be accounted for by the variation in another variable.

Cramer's V is the magnitude of the effect for chi-square or a measure of the magnitude of effect in a contingency table (larger than 2 x 2).

$$ES \Rightarrow V = \sqrt{\frac{\chi^2}{n(k-1)}} = \sqrt{\frac{6.4581}{245(2-1)}} = 0.1624$$

Where n is the sample size and k is the *smaller* of number of rows or number of columns.

If we square this measure, as we would a correlation coefficient, we get $r^2 = 0.02636$ or 2.64 percent of the variation in patient's satisfaction (Yes of No) is accounted for by the treatment center that care was received. While this effect is significant it is not very strong.