Name: ______________________________________ Block: _____ Date: ______________________

<u>Lesson 13.1 - The χ^2 Statistic</u>

The χ^2 statistic is a goodness of fit test. The formula to find the χ^2 statistic is:

$$\chi^2 = \Sigma \frac{(f_O - f_E)^2}{f_E}$$

where f_O = observed frequency, f_E = expected frequency, and Σ means sum.

The χ^2 test is used to check if the frequencies of collected data values (observed) differ significantly from what you would 'expect' to find (expected). In other words, it can be used to assess how good the 'fit' is between the observed and expected frequencies.

EXAMPLE #1

Jon needs some random numbers for his project, but has left his calculator at school. He writes down 100 single figures chosen from the numbers 0 to 9. Are the figures he has written down really random? How good is the fit between the numbers he has chosen 'at random' and numbers that are truly randomly generated?

Number	0	1	2	3	4	5	6	7	8	9
Frequency	12	7	9	14	6	7	14	13	8	10

Step #1: Organize your data in a table and fill in the calculations for each column. Calculate the χ^2 statistic.

NUMBER	f_O	f_E	$f_O - f_E$	$(f_O - f_E)^2$	$\frac{(f_O - f_E)^2}{F_E}$
0	12	10	2	4	.4
1	7	10	-3	9	.9
2	9	10	-1	1	.1
3	14	10	4	16	1.6
4	6	10	-4	16	1.6
5	7	10	-3	9	.9
6	14	10	4	16	1.6
7	13	10	3	9	.9
8	8	10	-2	4	.4
9	10	10	0	0	0
			Total:		8.4

χ^2_{stat}

Step #2: Calculate the degrees of freedom. $df = (\#\ of\ rows - 1)(\#\ of\ columns - 1)$

$$df = (2-1)(10-1) = \boxed{9}$$

Step #3a: Look up the χ^2 critical value in the χ^2 table (we usually test data at the 5% significance level). Compare the χ^2 statistic to the critical value. (Critical Value Table - Page 404)

- If the χ^2 statistic > χ^2 critical value, then the result **is not a good fit** for the observed data.
- If the χ^2 statistic < χ^2 critical value, then the result **is a good fit** for the observed data.

$\chi^2_{crit} = 16.919$

$\chi^2_{stat} < \chi^2_{crit}$

Jon's data is a good fit.
His numbers were truly random.

EXAMPLE #2

Finn and Edda are playing a game with an eight-sided die. Edda wins if she rolls an odd number; Finn wins if he rolls an even number. Edda keeps losing the game, and is sure that the dice is biased towards even numbers. From the results below, is there enough evidence to support Edda's claim?

Number	1	2	3	4	5	6	7	8
Frequency	4	5	6	9	6	9	0	3

NUMBER	f_O	f_E	$f_O - f_E$	$(f_O - f_E)^2$
1	4	5.25	-1.25	1.5625
2	5	5.25	-.25	.0625
3	6	5.25	.75	.5625
4	9	5.25	3.75	14.0625
5	6	5.25	.75	.5625
6	9	5.25	3.75	14.0625
7	0	5.25	-5.25	27.5625
8	3	5.25	-2.25	5.0625
			Total:	63.5

$$\chi^2_{stat} = \frac{63.5}{5.25} = 12.1$$

$$df = (2-1)(8-1) = 7$$

$$\chi^2_{crit} = 14.067$$

$$\underline{\chi^2_{stat} < \chi^2_{crit}}$$

The data is a good fit.
Edda was just unlucky.