Chi-Square

Are two ways of Categorizing people or things related?

Both Variables Qualitative/Categorical/Membership

Step 1: Arrange data into a frequency/contingency table

Step 2: Compute Expected Frequencies
Based Upon Null Hypothesis

Step 3: Compare Obtained Frequencies to Expected Frequencies
Do they Agree?

1: Contingency Table

Are Abortion Attitudes Related to Gender?					
	Abortion Attitude				
	Acceptable	Uncceptable		**Row Total**	
Women	59	29		88	
Men	15	37		52	
Column Total	74	66		**140**	**Grand Total**

2: Expected Frequencies

For each Cell: (Row Total x Column Total) / Grand Total

Are Abortion Attitudes Related to Gender?					
	Abortion Attitude				
	Acceptable	Uncceptable		**Row Total**	
Women	59	29		88	
Men	15	37		52	
Column Total	74	66		**140**	**Grand Total**

74	88	140	46.51
66	88	140	41.49
74	52	140	27.49
66	52	140	24.51

46.51	41.49
27.49	24.51

3: Compare

Error

Estimated Expected

$$\chi^2 = \Sigma\left[\frac{(O - E)^2}{E}\right]$$

↑
Evaluation
↓

For Each Cell

Do O and E Agree?

O	E	$O - E$	$(O - E)^2$	$\frac{(O - E)^2}{E}$
59	46.51	12.49	156.00	3.35
15	27.49	−12.49	156.00	5.67
29	41.49	−12.49	156.00	3.76
37	24.51	12.49	156.00	6.36
				$\chi^2 = 19.14$

The more the Observed frequencies differ from the Expected Frequencies,

- The Larger X^2
- The Lower the probability of the outcome, given H_0

Is It Significant?

df =
(Rows-1)*(Columns-1)

Significance:

Equal to or Greater
Than Critical Value

TABLE E Chi square distribution*

	α levels				
df	.10	.05	.02	.01	.001
1	2.71	3.84	5.41	6.64	10.81
2	4.60	5.99	7.82	9.21	13.82
3	6.25	7.82	9.84	11.34	16.27
4	7.78	9.49	11.67	13.28	18.46
5	9.24	11.07	13.39	15.09	20.52
6	10.64	12.59	15.03	16.81	22.46
7	12.02	14.07	16.62	18.48	24.32
8	13.36	15.51	18.17	20.09	26.12
9	14.68	16.92	19.68	21.67	27.88
10	15.99	18.31	21.16	23.21	29.59
11	17.28	19.68	22.62	24.72	31.26
12	18.55	21.03	24.05	26.22	32.91
13	19.81	22.36	25.47	27.69	34.53
14	21.06	23.68	26.87	29.14	36.12
15	22.31	25.00	28.26	30.58	37.70
16	23.54	26.30	29.63	32.00	39.25
17	24.77	27.59	31.00	33.41	40.79
18	25.99	28.87	32.35	34.80	42.31
19	27.20	30.14	33.69	36.19	43.82
20	28.41	31.41	35.02	37.57	45.32
21	29.62	32.67	36.34	38.93	46.80
22	30.81	33.92	37.66	40.29	48.27
23	32.01	35.17	38.97	41.64	49.73
24	33.20	36.42	40.27	42.98	51.18
25	34.38	37.65	41.57	44.31	52.62
26	35.56	38.88	42.86	45.64	54.05
27	36.74	40.11	44.14	46.96	55.48
28	37.92	41.34	45.42	48.28	56.89
29	39.09	42.56	46.69	49.59	58.30
30	40.26	43.77	47.96	50.89	59.70

Who Cares?

Is the Relationship non-Trivial?

$$\phi = \sqrt{\frac{\chi^2}{N}}$$

For a 2x2 Chi-Square

0 = No Relationship
1 = Perfect Relationship (What would that be?)

$\phi = 0.10$	Small Effect
$\phi = 0.30$	Medium Effect
$\phi = 0.50$	Large Effect

$$\phi = \sqrt{\frac{\chi^2}{N}} = \sqrt{\frac{19.14}{140}} = \sqrt{.1367} = .37$$

Hypothesis Testing:
Goodness of Fit
A One-Group Chi-Square

1. Specify Some Expected Probabilities/Proportions in Advance

2. Collect some data

3. Convert your Expected Proportions into Expected Frequencies
   Based upon the Total number of subjects assessed

4. Compare your Expected Frequencies to your Obtained Frequencies
   df = # of Categories - 1

Is Handedness Distributed Randomly in Monkeys

	R	L	
Expected Probability	50%	50%	
Obtained Frequency	15	5	N=20
Expected Frequencies	10	10	N=20

$$\chi^2 = \Sigma\left[\frac{(O-E)^2}{E}\right]$$

$(15\text{-}10)^2/10 \;+\; (5\text{-}10)^2/10$

$25/10 \;+\; 25/10 \;= 5$

df = K-1 = 2-1 = 1

Critical $X^2_{\text{2-tail,1df}} = 3.841$

5 > 3.841 ⇒ Reject Null Hypothesis

What If Outcome Was 14,6?

	R	**L**	
Expected Probability	50%	50%	
Obtained Frequency	14	6	N=20
Expected Frequencies	10	10	N=20

$$\chi^2 = \Sigma\left[\frac{(O - E)^2}{E}\right]$$

$(14\text{-}10)^2/10 \ + \ (6\text{-}10)^2/10$

$16/10 \quad + \quad 16/10 \quad = 3.2$

df = K-1 = 2-1 = 1

Critical $X^2_{\text{2-tail,1df}} = 3.841$

$3.2 < 3.841 \Rightarrow$ Retain Null Hypothesis

What If Hypothesis was Monkeys Lateralized to Right?

	R	L	
Expected Probability	50%	50%	
Obtained Frequency	14	6	N=20
Expected Frequencies	10	10	N=20

$$\chi^2 = \Sigma\left[\frac{(O-E)^2}{E}\right]$$

$(14-10)^2/10 + (6-10)^2/10$

$16/10 + 16/10 = 3.2$

df = K-1 = 2-1 = 1

What if More Monkeys were Lefties?

Critical $X^2_{1\text{-tail},1df} = 2.706$

$3.2 < 2.706 \Rightarrow$ Reject Null Hypothesis

What If Hypothesis was Monkeys Are Not Like Us?

	R	**L**	
Expected Probability	90%	10%	
Obtained Frequency	15	5	N=20
Expected Frequencies	18	2	N=20

$$\chi^2 = \Sigma\left[\frac{(O-E)^2}{E}\right]$$

$(15\text{-}18)^2/18 \;+\; (5\text{-}2)^2/2$

$9/18 \quad + \quad 9/2 \quad = 0.5 + 4.5 = 5$

df = K-1 = 2-1 = 1

Critical $X^2_{2\text{-tail},1df} = 3.841$

5 > 3.841 $\Rightarrow$ Reject Null Hypothesis