SOC 3811/5811:
BASIC SOCIAL STATISTICS

Associations Between Categorical Variables

χ^2 Example

Is there an association between race and whether people have a gun in their home?

		Race		
		White	*Black*	*Total*
Own A Gun?	*Yes*	451	58	509
	No	655	195	850
	Total	*1,106*	*253*	*1,359*

χ^2 Example

State the null (H_0) and alternative (H_1) hypotheses

Again, we begin with the assumption that there is no association between X and Y

H_0: People's race and whether they have a gun in their home are statistically independent

H_1: People's race and whether they have a gun in their home are not statistically independent

χ^2 Example

Check that the sample data conform to basic assumptions; if they do not, then do not go any further

X and Y must be collected from a random sample of individuals from the population (OK)

Standard χ^2 testing procedures should be used with extreme caution — or not at all — if any cell frequency is less than 5 (OK)

χ^2 Example

Choose an α probability level … that is, a probability associated with incorrectly rejecting the null hypothesis

Let's use α=0.01 in our example

χ^2 Example

Determine the "critical value" … that is, how large the test statistic must be in order to reject the null hypothesis at the given α level

In our example, α=0.01 and df=(2—1)(2—1)=1

According to a χ^2 table, the critical value of χ^2 thus equals 6.63

Table entry for p is the critical value $(\chi^2)^*$ with probability p lying to its right.

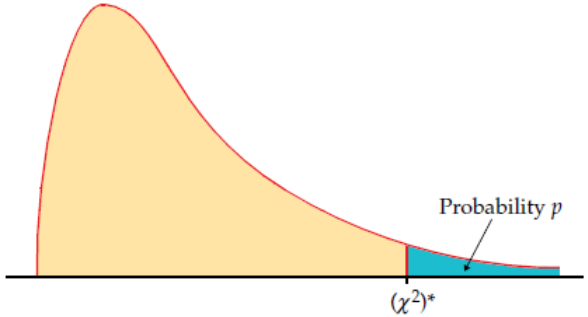


TABLE F

χ^2 distribution critical values

	Tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83	12.12
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82	15.20
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27	17.73
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47	20.00
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51	22.11
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46	24.10
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32	26.02
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12	27.87
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88	29.67
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59	31.42
11	13.70	14.63	15.77	17.28	19.68	21.92	22.62	24.72	26.76	28.73	31.26	33.14
12	14.85	15.81	16.99	18.55	21.03	23.34	24.05	26.22	28.30	30.32	32.91	34.82
13	15.98	16.98	18.20	19.81	22.36	24.74	25.47	27.69	29.82	31.88	34.53	36.48
14	17.12	18.15	19.41	21.06	23.68	26.12	26.87	29.14	31.32	33.43	36.12	38.11
15	18.25	19.31	20.60	22.31	25.00	27.49	28.26	30.58	32.80	34.95	37.70	39.72
16	19.37	20.47	21.79	23.54	26.30	28.85	29.63	32.00	34.27	36.46	39.25	41.31
17	20.49	21.61	22.98	24.77	27.59	30.19	31.00	33.41	35.72	37.95	40.79	42.88
18	21.60	22.76	24.16	25.99	28.87	31.53	32.35	34.81	37.16	39.42	42.31	44.43
19	22.72	23.90	25.33	27.20	30.14	32.85	33.69	36.19	38.58	40.88	43.82	45.97
20	23.83	25.04	26.50	28.41	31.41	34.17	35.02	37.57	40.00	42.34	45.31	47.50
21	24.93	26.17	27.66	29.62	32.67	35.48	36.34	38.93	41.40	43.78	46.80	49.01
22	26.04	27.30	28.82	30.81	33.92	36.78	37.66	40.29	42.80	45.20	48.27	50.51
23	27.14	28.43	29.98	32.01	35.17	38.08	38.97	41.64	44.18	46.62	49.73	52.00
24	28.24	29.55	31.13	33.20	36.42	39.36	40.27	42.98	45.56	48.03	51.18	53.48
25	29.34	30.68	32.28	34.38	37.65	40.65	41.57	44.31	46.93	49.44	52.62	54.95
26	30.43	31.79	33.43	35.56	38.89	41.92	42.86	45.64	48.29	50.83	54.05	56.41
27	31.53	32.91	34.57	36.74	40.11	43.19	44.14	46.96	49.64	52.22	55.48	57.86
28	32.62	34.03	35.71	37.92	41.34	44.46	45.42	48.28	50.99	53.59	56.89	59.30
29	33.71	35.14	36.85	39.09	42.56	45.72	46.69	49.59	52.34	54.97	58.30	60.73
30	34.80	36.25	37.99	40.26	43.77	46.98	47.96	50.89	53.67	56.33	59.70	62.16
40	45.62	47.27	49.24	51.81	55.76	59.34	60.44	63.69	66.77	69.70	73.40	76.09
50	56.33	58.16	60.35	63.17	67.50	71.42	72.61	76.15	79.49	82.66	86.66	89.56
60	66.98	68.97	71.34	74.40	79.08	83.30	84.58	88.38	91.95	95.34	99.61	102.7
80	88.13	90.41	93.11	96.58	101.9	106.6	108.1	112.3	116.3	120.1	124.8	128.3
100	109.1	111.7	114.7	118.5	124.3	129.6	131.1	135.8	140.2	144.3	149.4	153.2

χ^2 Example

Is there an association between race and whether people have a gun in their home?

Calculate the test statistic ... χ^2

		Race		
		White	*Black*	*Total*
Own A Gun?	*Yes*	451	58	509
		414	95	
	No	655	195	850
	Total	*1,106*	*253*	*1,359*

$$\hat{f}_{i,j} = \frac{(f_{i\bullet})(f_{\bullet j})}{N} \qquad \frac{(509)(1106)}{1359} \qquad \frac{(509)(253)}{1359}$$

χ^2 Example

Is there an association between race and whether people have a gun in their home?

Calculate the test statistic ... χ^2

		Race		
		White	*Black*	*Total*
Own A Gun?	*Yes*	451 414	58 95	509
	No	655 692	195 158	850
	Total	*1,106*	*253*	*1,359*

χ^2 Example

Calculate the test statistic ... χ^2

$$\chi^2 = \text{Sum of } \frac{(\text{Expected} - \text{Observed})^2}{\text{Expected}} = \sum_{i=1}^{R} \sum_{j=1}^{C} \frac{\left(\hat{f}_{ij} - f_{ij}\right)^2}{\hat{f}_{ij}}$$

χ^2 Example

Is there an association between race and whether people have a gun in their home?

Calculate the test statistic ... χ^2

		Race		
		White	*Black*	*Total*
Own A Gun?	*Yes*	451 414	58 95	509
	No	655 692	195 158	850
	Total	*1,106*	*253*	*1,359*

$$\chi^2 = \text{Sum of } \frac{(\text{Expected} - \text{Observed})^2}{\text{Expected}} \approx \frac{(414-451)^2}{414} + \frac{(95-58)^2}{95} + - + -$$

$$= 28.01$$

χ^2 Example

Is there an association between race and whether people have a gun in their home?

Calculate the test statistic ... χ^2

		Race		
		White	*Black*	*Total*
Own A Gun?	*Yes*	451 414	58 95	509
	No	655 692	195 158	850
	Total	*1,106*	*253*	*1,359*

$$\chi^2 = \text{Sum of } \frac{(\text{Expected} - \text{Observed})^2}{\text{Expected}}$$

$$= \frac{(414-451)^2}{414} + \frac{(95-58)^2}{95} + \frac{(692-655)^2}{692} + \frac{(158-195)^2}{158} = 28.01$$

χ^2 Example

Compare the test statistic to the critical value

1. If the test statistic is as large or larger than the critical value, then reject H_0 (with probability of α of doing so even though H_0 should not actually be rejected)
2. If the test statistic is less than the critical value, then do no reject H_0 (with probability of β of doing so even though H_0 should be rejected)

We can restate the hypotheses as

H_0: X & Y are independent ➔ Fail to Reject if $\chi^2 \leq 6.63$

H_1: X & Y not independent ➔ Reject if $\chi^2 > 6.63$

Since $\chi^2 = 28.01$, we reject H_0

WORKSHEET

Is there an association between people's favorite pet and whether they are right or left handed? Use α=0.05.

		Handed?		
		Right	*Left*	*Total*
Best	*Not Dog*	100	20	120
Pet?	*Dog*	200	50	250
	Total	*300*	*70*	*370*

H_0:
H_a

$\frac{(120)(300)}{370}$ $\frac{(120)(70)}{370}$

97.297 22.703

202.703 47.297

$\chi^2_{critical} = 3.84$

$\Sigma = \frac{(97.297-100)^2}{97.297} + \ldots + + \ldots = .5873$

WORKSHEET

Is there an association between people's favorite pet and whether they are right or left handed? Use α=0.05.

$df = (\text{Rows}-1)(\text{Cols}-1) = (2-1)(2-1) = 1$

		Handed?		
		Right	*Left*	*Total*
Best Pet?	*Not Dog*	100 **97.297**	20 **22.703**	120
	Dog	200 **202.703**	50 **47.297**	250
	Total	*300*	*70*	*370*

$\frac{(\text{Row})(\text{Col})}{\text{Total}}$

$$\chi^2 = \text{Sum of } \frac{(\text{Expected}-\text{Observed})^2}{\text{Expected}}$$

$$= \frac{(97-100)^2}{97} + \frac{(23-20)^2}{23} + \frac{(203-200)^2}{203} + \frac{(47-50)^2}{47} = 0.5873$$

Associations Between Discrete Variables

Is there an association in the population?

χ^2 analysis

How strong is the association? In what direction is the association?

Gamma, relative risk, odds ratios

Association Between Discrete Variables

Measures of Association

Statistics that show the direction and/or magnitude of a relationship between pairs of variables

When X and Y are Both Ordinal:

Gamma

Others (Not discussed today)

When X and Y are Both Dichotomous:

Gamma

Relative Risk (RR)

Odds Ratio (OR)

Others (Not discussed today)

Gamma

The possible values of gamma range from -1 to +1

Gamma = -1 = a perfect negative association

Gamma = +1 = a perfect positive association

Gamma = 0 = no association

Gamma = +0.986

	x=1	*x=2*	*x=3*
y=3	1	1	100
y=2	1	100	1
y=1	100	1	1

Gamma = -0.986

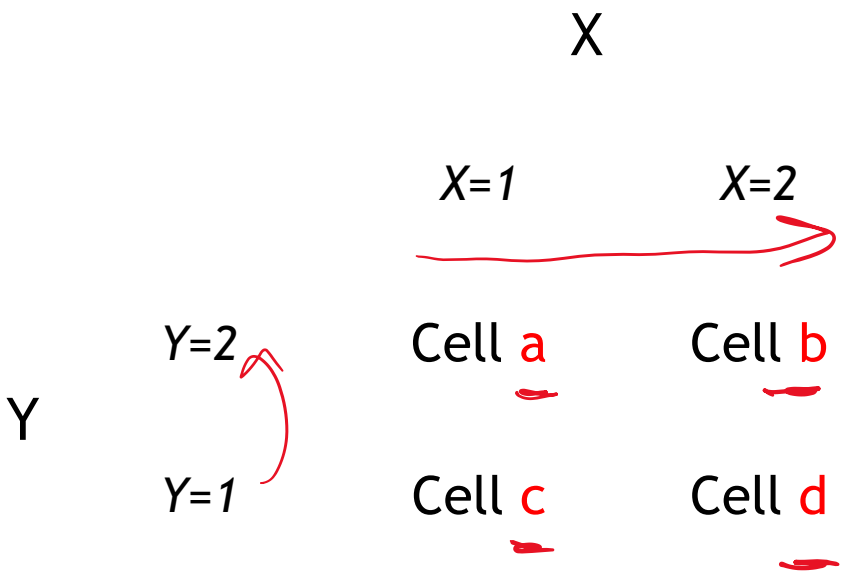
	x=1	*x=2*	*x=3*
y=3	100	1	1
y=2	1	100	1
y=1	1	1	100

Gamma = 0.000

	x=1	*x=2*	*x=3*
y=3	1	1	1
y=2	1	1	1
y=1	1	1	1

Association in 2x2 Tables

Each measure relies on a particular naming convention for the four cells in the cross-tabulation



χ^2 Example

Is there an association between race and whether people have a gun in their home?

		Race		
		White	*Black*	*Total*
Own A Gun?	*Yes*	451 **a**	58 **b**	509
	No	655 **c**	195 **d**	850
	Total	*1,106*	*253*	*1,359*

Relative Risk

When X=1, the "Risk" that Y=2 equals a/(a+c) … which is the same as P(Y=2|X=1)

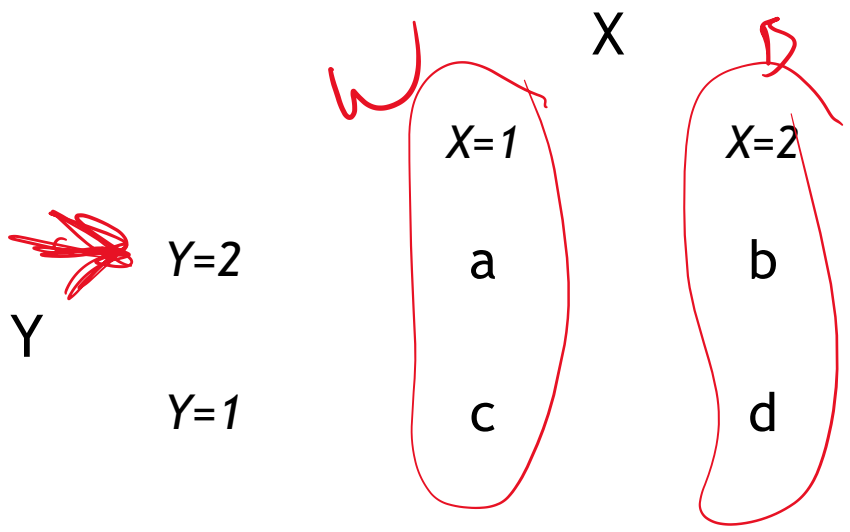
When X=2, the "Risk" that Y=2 equals b/(b+d) … which is the same as P(Y=2|X=2)

The **relative risk** (RR)

is the ratio of the

two risks:

$$RR = \frac{b/(b+d)}{a/(a+c)}$$



Relative Risk

Relative Risk can range from zero to infinity

RR 0.00 to 1.00 means that the risk that Y=2 is reduced when you move from X=1 to X=2

RR = 1.00 means that the risk that Y=2 is unchanged when you move from X=1 to X=2

RR > 1.00 means that the risk that Y=2 is increased when you move from X=1 to X=2

Relative Risk

Is there an association between race and whether people have a gun in their home?

		Race		
		White	*Black*	*Total*
Own A Gun?	*Yes*	451 **a**	58 **b**	509
	No	655 **c**	195 **d**	850
	Total	*1,106*	*253*	*1,359*

$$RR = \frac{b/(b+d)}{a/(a+c)} = \frac{58/(58+195)}{451/(451+655)} = 0.56$$

Odds Ratio

When X=1, the “Odds” that Y=2 equals a/c

When X=2, the “Odds” that Y=2 equals b/d

The “Odds Ratio” (OR) is the ratio of the two odds:

$$OR = \frac{b/d}{a/c} = \frac{bc}{ad}$$

		X=1	X=2
Y	Y=2	a	b
	Y=1	c	d

Odds Ratio

Odds Ratios can range from zero to infinity

OR 0.00 to 1.00 means that the odds that Y=2 is reduced when you move from X=1 to X=2

OR = 1.00 means that the odds that Y=2 is unchanged when you move from X=1 to X=2

OR > 1.00 means that the odds that Y=2 is increased when you move from X=1 to X=2

Odds Ratio

Is there an association between race and whether people have a gun in their home?

		Race		
		White	*Black*	*Total*
Own A Gun?	*Yes*	451 **a**	58 **b**	509
	No	655 **c**	195 **d**	850
	Total	*1,106*	*253*	*1,359*

$$\text{OR} = \frac{b/d}{a/c} = \frac{bc}{ad} = \frac{(58)(655)}{(451)(195)} = 0.43$$

1. **Gamma = 0.6552**
2. **Compute relative risk and odds ratio**
3. **Interpret all 3 measures**

		Ever Tried Meditation? *No*	*Yes*	*Total*
Ever Done Yoga?	*Yes*	18	54	72
	No	16	10	26
	Total	*34*	*64*	*98*

$$RR = \frac{b/(b+d)}{a/(a+c)} = \frac{54/(54+10)}{18/(18+16)} = \frac{.84}{.52} = 1.59$$

$$OR = \frac{b/d}{a/c} = \frac{bc}{ad} = \frac{54/10}{18/16} = \frac{5.4}{1.125} = 4.8$$