CHI SQUARE TEST

Parametric and Nonparametric Tests

- This lecture introduces two **non-parametric thypothesis tests** using the chi-square statistic: the chi-square test for goodness of fit and the chi-square test for independence.

Parametric and Nonparametric Tests (cont.)

- The term "non-parametric" refers to the fact that the chi-square tests do not require assumptions about population parameters nor do they test hypotheses about population parameters.

- Previous examples of hypothesis tests, such as the t tests and analysis of variance, are **parametric tests** and they do include assumptions about parameters and hypotheses about parameters.

Parametric and Nonparametric Tests (cont.)

- The most obvious difference between the chi-square tests and the other hypothesis tests we have considered (t and Z) is the nature of the data.
- For chi-square, the data are frequencies rather than numerical scores.

Nonparametric Statistics

- A special class of hypothesis tests
- Used when assumptions for parametric tests are not met
  - Review: What are the assumptions for parametric tests?

Assumptions for Parametric Tests

- Dependent variable is a scale variable ➔ interval or ratio
  - If the dependent variable <u>is ordinal or nominal, it is a non-parametric test</u>
- Participants are randomly selected
  - <u>If there is no randomization, it is a non-parametric test</u>
- The underlying population distribution is normal
  - If the shape is not normal, it is a non-parametric test

Limitations of Nonparametric Tests

- Cannot easily use confidence intervals or effect sizes
- Have less statistical power than parametric tests
- Nominal and ordinal data provide less information
- More likely to commit type II error
  - Review: What is type I error? Type II error?

The Chi-Square Test for Goodness-of-Fit

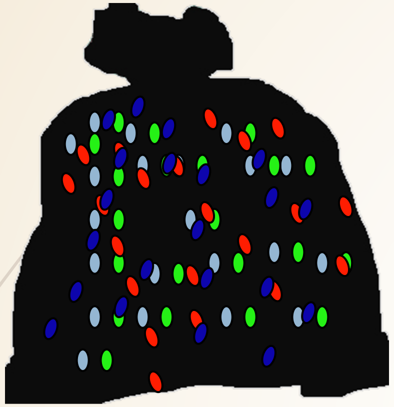
- The **chi-square test for goodness-of-fit** uses frequency data from a sample to test hypotheses about the shape or proportions of a population.
- Each individual in the sample is classified into one category on the scale of measurement.
- The data, called **observed frequencies**, simply count how many individuals from the sample are in each category.

The Chi-Square Test for Goodness-of-Fit (cont.)

- The null hypothesis specifies the proportion of the population that should be in each category.

- The proportions from the null hypothesis are used to compute **expected frequencies** that describe how the sample would appear if it were in perfect agreement with the null hypothesis.

Chi-Square test for goodness of fit- An Example



-Is the frequency of balls with different colors equal in our bag?

Observed Frequencies

50	30	30	10

Expected Frequencies

25%	25%	25%	25%

Chi-Square test for goodness of fit

Observed Frequencies

50	30	30	10	

Total: 120

Expected Frequencies

25%	25%	25%	25%

× 120 =

Expected Frequencies

30	30	30	30

H0

Chi-Square test for goodness of fit

Observed Frequencies

50	30	30	10

Expected Frequencies

30	30	30	30

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e}$$

Difference

Normalize

$$\chi^2 = \frac{(50-30)^2}{30} + \frac{(30-30)^2}{30} + \frac{(30-30)^2}{30} + \frac{(10-30)^2}{30} = 26.6$$

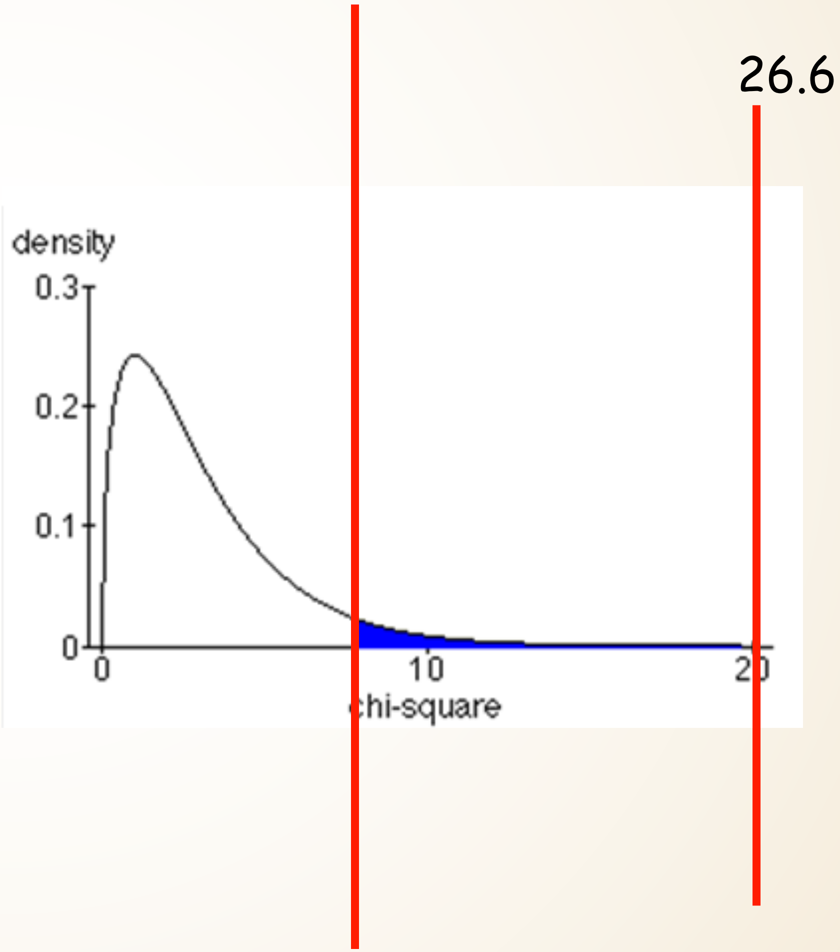
Chi-Square test for goodness of fit

$$\chi^2 = 26.6$$

$$df = C - 1 = 4 - 1 = 3$$

$\chi^2_{(3, n=120)} = 26.66$,
$p < 0.001$

Critical value = 7.81



Chi-Square test for Goodness of fit

- Chi-Square test for goodness of fit is like one sample t-test
- You can test your sample against any possible expected values

25%	25%	25%	25%	H0
10%	10%	10%	70%	H0

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14
15	5.229	7.261	8.547	11.037	14.339	18.25	22.31	25.00	30.58
16	5.812	7.962	9.312	11.912	15.338	19.37	23.54	26.30	32.00
17	6.408	8.672	10.085	12.792	16.338	20.49	24.77	27.59	33.41
18	7.015	9.390	10.865	13.675	17.338	21.60	25.99	28.87	34.80
19	7.633	10.117	11.651	14.562	18.338	22.72	27.20	30.14	36.19
20	8.260	10.851	12.443	15.452	19.337	23.83	28.41	31.41	37.57
22	9.542	12.338	14.041	17.240	21.337	26.04	30.81	33.92	40.29
24	10.856	13.848	15.659	19.037	23.337	28.24	33.20	36.42	42.98
26	12.198	15.379	17.292	20.843	25.336	30.43	35.56	38.89	45.64
28	13.565	16.928	18.939	22.657	27.336	32.62	37.92	41.34	48.28
30	14.953	18.493	20.599	24.478	29.336	34.80	40.26	43.77	50.89
40	22.164	26.509	29.051	33.660	39.335	45.62	51.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38

The Chi-Square Test for Independence

- The second chi-square test, the **chi-square test for independence**, can be used and interpreted in two different ways:

  1. Testing hypotheses about the relationship between two variables in a population, or

  2. Testing hypotheses about differences between proportions for two or more populations.

The Chi-Square Test for Independence (cont.)

- Although the two versions of the test for independence appear to be different, they are equivalent and they are interchangeable.

- The first version of the test emphasizes the relationship between chi-square and a correlation, because both procedures examine the relationship between two variables.

The Chi-Square Test for Independence (cont.)

- The second version of the test emphasizes the relationship between chi-square and an independent-measures t test (or ANOVA) because both tests use data from two (or more) samples to test hypotheses about the difference between two (or more) populations.

The Chi-Square Test for Independence (cont.)

- The first version of the chi-square test for independence views the data as one sample in which each individual is classified on two different variables.

- The data are usually presented in a matrix with the categories for one variable defining the rows and the categories of the second variable defining the columns.

The Chi-Square Test for Independence (cont.)

- The data, called **observed frequencies**, simply show how many individuals from the sample are in each cell of the matrix.

- The null hypothesis for this test states that there is no relationship between the two variables; that is, the two variables are independent.

The Chi-Square Test for Independence (cont.)

- The second version of the test for independence views the data as two (or more) separate samples representing the different populations being compared.
- The same variable is measured for each sample by classifying individual subjects into categories of the variable.
- The data are presented in a matrix with the different samples defining the rows and the categories of the variable defining the columns..

The Chi-Square Test for Independence (cont.)

- The data, again called **observed frequencies**, show how many individuals are in each cell of the matrix.

- The null hypothesis for this test states that the proportions (the distribution across categories) are the same for all of the populations

The Chi-Square Test for Independence (cont.)

- Both chi-square tests use the same statistic. The calculation of the chi-square statistic requires two steps:

1. The null hypothesis is used to construct an idealized sample distribution of **expected frequencies** that describes how the sample would look if the data were in perfect agreement with the null hypothesis.

The Chi-Square Test for Independence (cont.)

- For the goodness of fit test, the expected frequency for each category is obtained by

$$\text{expected frequency} = f_e = p \times n$$

(p is the proportion from the null hypothesis and n is the size of the sample)

- For the test for independence, the expected frequency for each cell in the matrix is obtained by

- $$\text{expected frequency} = f_e = \frac{(\text{row total})\text{x}(\text{column total})}{n}$$

The Chi-Square Test for Independence (cont.)

2. A chi-square statistic is computed to measure the amount of discrepancy between the ideal sample (expected frequencies from H_0) and the actual sample data (the observed frequencies = f_o).

A large discrepancy results in a large value for chi-square and indicates that the data do not fit the null hypothesis and the hypothesis should be rejected.

Chi-Square test for independence- An Example

- When we have two or more sets of categorical data (IV,DV both categorical)

F_O	None	Obama	McCain	
Male	10	50	35	95
Female	15	60	40	115
	25	110	75	210

Chi-Square test for independence- An Example

- Also called contingency table analysis
- H0: There is no relation between gender and voting preference (like correlation)

OR

- H0: There is no difference between the voting preference of males and females (like t-test)

- The logic is the same as the goodness of fit test: Comparing observed freq and Expected freq if the two variables were independent

Chi-Square test for independence- An Example

F_O	None	Obama	McCain	
Male	10	50	35	95
Female	15	60	40	115
	25	110	75	210

F_E	None	Obama	McCain	
Male				
Female				
	12%	52%	36%	100%

Chi-Square test for independence

In case of independence:

F_E	None	Obama	McCain		
Male	12%	52%	36%	×	95
Female	12%	52%	36%	×	115
	12%	52%	36%		100%

Finaly:

F_E	None	Obama	McCain
Male	11.4	49.4	34.2
Female	13.8	59.8	41.4

Chi-Square test for independence

- Another way:

$$f_e = \frac{f_c \times f_r}{n} = \frac{column \times row}{total}$$

F_E	None	Obama	McCain	
Male	$\frac{95 \times 25}{210}$			95
Female				
	25			210

Chi-Square test for independence

- Now we can calculate the chi square value :

F_O

10	50	35
15	60	40

F_E

11.4	49.4	34.2
13.8	59.8	41.4

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e}$$

$$\chi^2 = \frac{(10-11.4)^2}{11.4} + \frac{(15-13.8)^2}{13.8} + \ldots = 0.35$$

$$df = (C-1)\times(R-1) = (3-1)\times(2-1) = 2$$

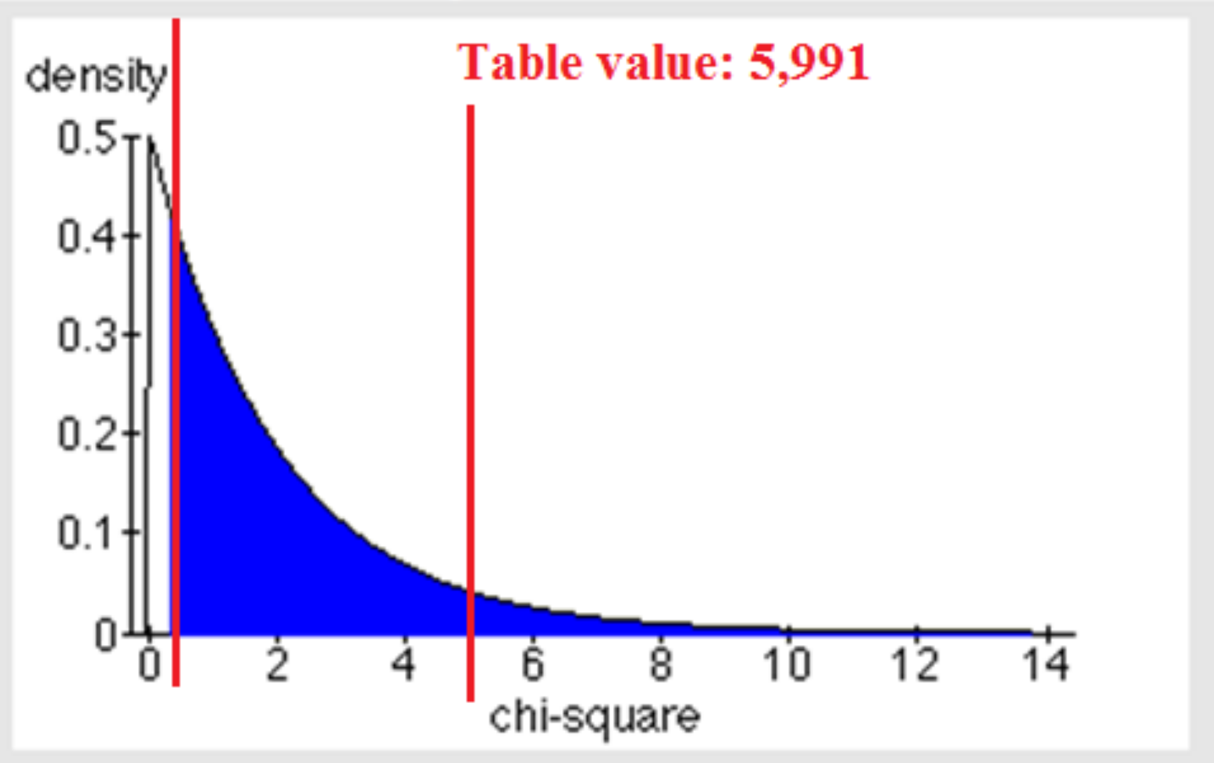
Chi-Square test for independence- An Example

$\chi^2(2, n=210) = 0.35, p= 0.83$

There is no significant effect of gender on vote preference

Or

We cannot reject the null hypothesis that gender and vote preference are independent



Contingency Table Example

Left-Handed vs. Gender

Dominant Hand: Left vs. Right

Gender: Male vs. Female

- 2 categories for each variable, so this is called a 2 x 2 table
- Suppose we examine a sample of 300 children

Contingency Table Example

(continued)

Sample results organized in a contingency table:

sample size = n = 300:

120 Females, 12 were left-handed

180 Males, 24 were left-handed

Gender	Hand Preference		
	Left	Right	
Female	12	108	120
Male	24	156	180
	36	264	300

χ^2 Test for the Difference Between Two Proportions

H_0: $\pi_1 = \pi_2$ (Proportion of females who are left handed is equal to the proportion of males who are left-handed)

H_1: $\pi_1 \neq \pi_2$ (The two proportions are not the same)

- If H_0 is true, then the proportion of left-handed females should be the same as the proportion of left-handed males
- The two proportions above should be the same as the proportion of left-handed people overall

The Chi-Square Test Statistic

The Chi-square test statistic is:

$$\chi^2_{STAT} = \sum_{all\ cells} \frac{(f_o - f_e)^2}{f_e}$$

- where:
  f_o = observed frequency in a particular cell
  f_e = expected frequency in a particular cell if H_0 is true

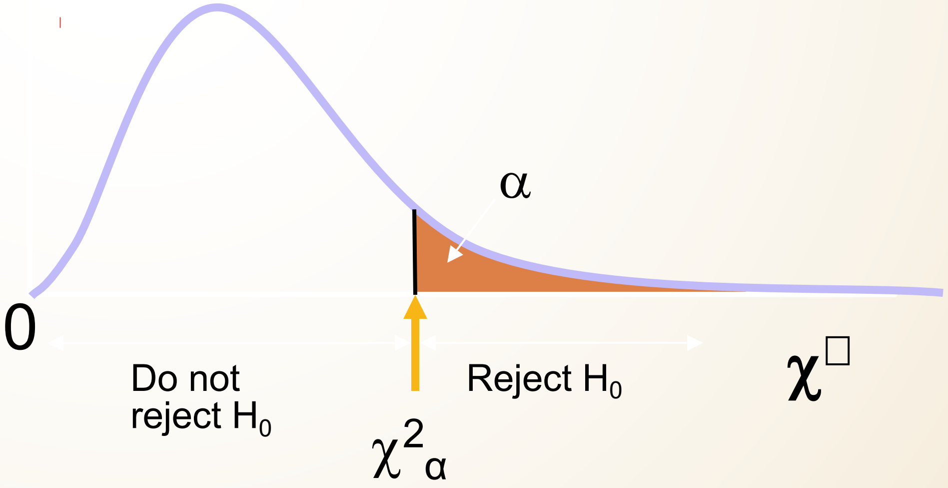
χ^2_{STAT} **for the 2 x 2 case has 1 degree of freedom**

(Assumed: each cell in the contingency table has expected frequency of at least 5)

Decision Rule

The χ^2_{STAT} test statistic approximately follows a chi-squared distribution with one degree of freedom

Decision Rule:
If $\chi^2_{STAT} > \chi^2_{\alpha}$, reject H_0, otherwise, do not reject H_0



Observed vs. Expected Frequencies

Gender	Hand Preference		
	Left	Right	
Female	Observed = 12 Expected = 14.4	Observed = 108 Expected = 105.6	120
Male	Observed = 24 Expected = 21.6	Observed = 156 Expected = 158.4	180
	36	264	300

The Chi-Square Test Statistic

Gender	Hand Preference		
	Left	Right	
Female	Observed = 12 Expected = 14.4	Observed = 108 Expected = 105.6	120
Male	Observed = 24 Expected = 21.6	Observed = 156 Expected = 158.4	180
	36	264	300

The test statistic is:

$$\chi^2_{STAT} = \sum_{\text{all cells}} \frac{(f_o - f_e)^2}{f_e}$$

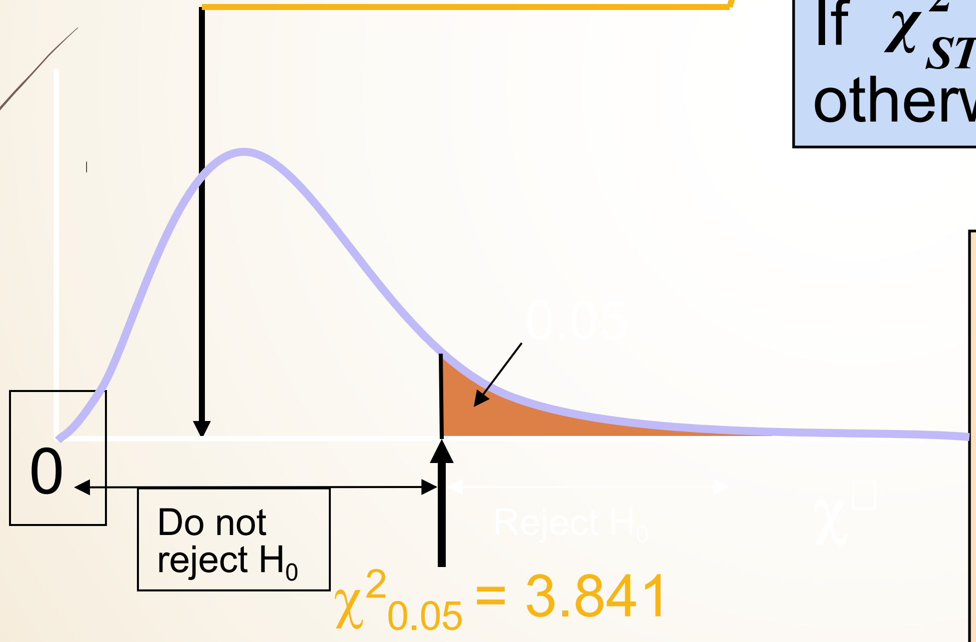
$$= \frac{(12-14.4)^2}{14.4} + \frac{(108-105.6)^2}{105.6} + \frac{(24-21.6)^2}{21.6} + \frac{(156-158.4)^2}{158.4} = 0.7576$$

Decision Rule

The test statistic is $\chi^2_{STAT} = 0.7576$; $\chi^2_{0.05}$ with 1 d.f. $= 3.841$

Decision Rule:

If $\chi^2_{STAT} > 3.841$, reject H_0, otherwise, do not reject H_0



Here,

$\chi^2_{STAT} = 0.7576 < \chi^2_{0.05} = 3.841$, so we do not reject H_0 and conclude that there is not sufficient evidence that the two proportions are different at $\alpha = 0.05$

χ^2 Test for Differences Among More Than Two Proportions

- Extend the χ^2 test to the case with more than two independent populations:

H_0: $\pi_1 = \pi_2 = \cdots = \pi_c$

H_1: Not all of the π_j are equal ($j = 1, 2, \cdots, c$)

The Chi-Square Test Statistic

The Chi-square test statistic is:

$$\chi^2_{STAT} = \sum_{\textit{all}\ \text{cells}} \frac{(f_o - f_e)^2}{f_e}$$

- Where:

  f_o = observed frequency in a particular cell of the 2 x c table

  f_e = expected frequency in a particular cell if H_0 is true

$$\chi^2_{STAT} \ \textbf{for the 2 x c case has (2 - 1)(c - 1)} = \textbf{c - 1 degrees of freedom}$$

(Assumed: each cell in the contingency table has expected frequency of at least 1)

χ^2 Test of Independence

- Similar to the χ^2 test for equality of more than two proportions, but extends the concept to contingency tables with r rows and c columns

H_0: The two categorical variables are independent
(i.e., there is no relationship between them)
H_1: The two categorical variables are dependent
(i.e., there is a relationship between them)

χ^2 Test of Independence

(continued)

The Chi-square test statistic is:

$$\chi^2_{STAT} = \sum_{all\ \text{cells}} \frac{(f_o - f_e)^2}{f_e}$$

- where:

  f_o = observed frequency in a particular cell of the r x c table

  f_e = expected frequency in a particular cell if H_0 is true

χ^2_{STAT} for the r x c case has (r - 1)(c - 1) degrees of freedom

(Assumed: each cell in the contingency table has expected frequency of at least 1)

Expected Cell Frequencies

- Expected cell frequencies:

$$f_e = \frac{\text{row total} \times \text{column total}}{n}$$

Where:

row total = sum of all frequencies in the row

column total = sum of all frequencies in the column

n = overall sample size

Decision Rule

- The decision rule is

> If $\chi^2_{STAT} \square \chi^2_{\alpha}$, reject H_0,
>
> otherwise, do not reject H_0

Where χ^2_{α} is from the chi-square distribution with (r – 1) (c – 1) degrees of freedom

Example

- The meal plan selected by 200 students is shown below:

Class Standing	Number of meals per week			Total
	20/week	10/week	none	
Fresh.	24	32	14	70
Soph.	22	26	12	60
Junior	10	14	6	30
Senior	14	16	10	40
Total	70	88	42	200

Example

(continued)

- The hypothesis to be tested is:

H_0: Meal plan and class standing are independent
(i.e., there is no relationship between them)
H_1: Meal plan and class standing are dependent
(i.e., there is a relationship between them)

(continued)

Example: Expected Cell Frequencies

Observed:

Class Standing	Number of meals per week			Total
	20/wk	10/wk	none	
Fresh.	24	32	14	70
Soph.	22	26	12	60
Junior	10	14	6	30
Senior	14	16	10	40
Total	70	88	42	200

Expected cell frequencies if H_0 is true:

Class Standing	Number of meals per week			Total
	20/wk	10/wk	none	
Fresh.	24.5	30.8	14.7	70
Soph.	21.0	26.4	12.6	60
Junior	10.5	13.2	6.3	30
Senior	14.0	17.6	8.4	40
Total	70	88	42	200

Example for one cell:

$$f_e = \frac{\text{row total} \times \text{column total}}{n}$$

$$= \frac{30 \times 70}{200} = 10.5$$

Example: The Test Statistic

(continued)

- The test statistic value is:

$$\chi^2_{STAT} = \sum_{all\ cells} \frac{(f_o - f_e)^2}{f_e}$$

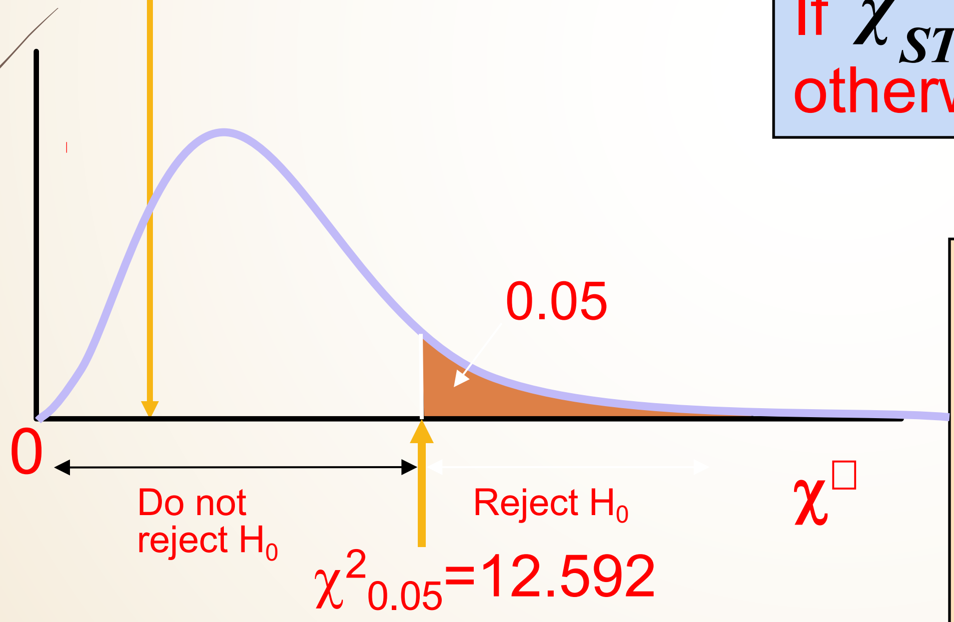
$$= \frac{(24-24.5)^2}{24.5} + \frac{(32-30.8)^2}{30.8} + \cdots + \frac{(10-8.4)^2}{8.4} = 0.709$$

$\chi^2_{0.05}$ = 12.592 from the chi-square distribution with (4 – 1)(3 – 1) = 6 degrees of freedom

Example: Decision and Interpretation *(continued)*

The test statistic is $\chi^2_{STAT} = 0.709$; $\chi^2_{0.05}$ with 6 d.f. $= 12.592$

Decision Rule:
If $\chi^2_{STAT} > 12.592$, reject H_0, otherwise, do not reject H_0



Here,
$\chi^2_{STAT} = 0.709 < \chi^2_{0.05} = 12.592$,
so do not reject H_0
Conclusion: there is not sufficient evidence that meal plan and class standing are related at $\alpha = 0.05$

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14
15	5.229	7.261	8.547	11.037	14.339	18.25	22.31	25.00	30.58
16	5.812	7.962	9.312	11.912	15.338	19.37	23.54	26.30	32.00
17	6.408	8.672	10.085	12.792	16.338	20.49	24.77	27.59	33.41
18	7.015	9.390	10.865	13.675	17.338	21.60	25.99	28.87	34.80
19	7.633	10.117	11.651	14.562	18.338	22.72	27.20	30.14	36.19
20	8.260	10.851	12.443	15.452	19.337	23.83	28.41	31.41	37.57
22	9.542	12.338	14.041	17.240	21.337	26.04	30.81	33.92	40.29
24	10.856	13.848	15.659	19.037	23.337	28.24	33.20	36.42	42.98
26	12.198	15.379	17.292	20.843	25.336	30.43	35.56	38.89	45.64
28	13.565	16.928	18.939	22.657	27.336	32.62	37.92	41.34	48.28
30	14.953	18.493	20.599	24.478	29.336	34.80	40.26	43.77	50.89
40	22.164	26.509	29.051	33.660	39.335	45.62	51.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38

Award Preference & SAT

The data in **StudentSurvey** includes two categorical variables:

Award = Academy, Nobel, or Olympic

HigherSAT = Math or Verbal

Do you think there is a relationship between the award preference and which SAT is higher? If so, in what way?

Award Preference & SAT

HigherSAT	**Academy**	**Nobel**	**Olympic**	Total
Math	21	68	116	205
Verbal	10	79	61	150
Total	31	147	177	355

Data are summarized with a 2×3 table for a sample of size n=355.

H_0 : Award preference is not associated with which SAT is higher

H_a : Award preference is associated with which SAT is higher

If H_0 is true $\Longrightarrow$ The award distribution is expected to be the same in each row.

Expected Counts

$$\text{Expected Count} = \frac{\text{row total} \times \text{column total}}{n}$$

HigherSAT	**Academy**	**Nobel**	**Olympic**	Total
Math				205
Verbal				150
Total	31	147	177	355

Note: The expected counts maintain row and column totals, but redistribute the counts as if there were *no* association.

Chi-Square Statistic

HigherSAT	**Academy**	**Nobel**	**Olympic**	Total
Math	21 (17.9)	68 (84.9)	116 (102.2)	205
Verbal	10 (13.1)	79 (62.1)	61 (74.8)	150
Total	31	147	177	355

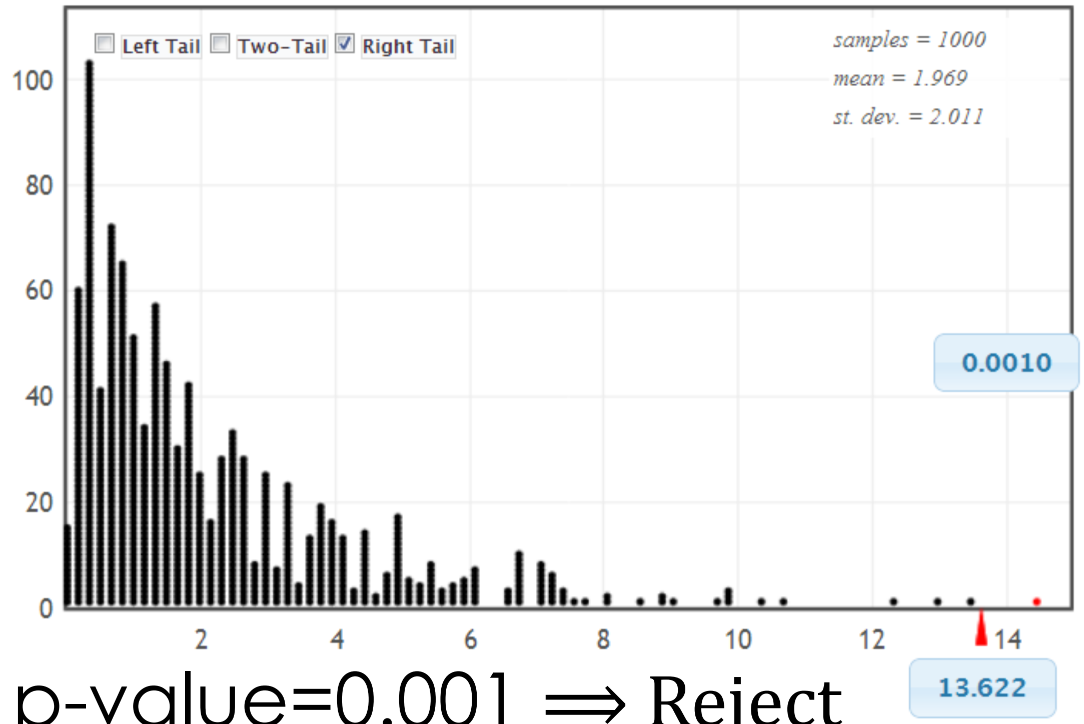
HigherSAT	**Academy**	**Nobel**	**Olympic**
Math			
Verbal			

$$\chi^2 = \sum \frac{(\text{observed - expected})^2}{\text{expected}}$$

Randomization Test

www.lock5stat.com/statk

Randomization Dotplot of χ^2 , Null hypothesis: No Association



Original Sample Show Details

$n = 355, \chi^2 = 13.622$

	Academy	Nobel	Olympic	Total
Math	21	68	116	205
Verbal	10	79	61	150
Total	31	147	177	355

Randomization Sample Show Details

$n = 355, \chi^2 = 9.577$

	Academy	Nobel	Olympic	Total
Math	15	99	91	205
Verbal	16	48	86	150
Total	31	147	177	355

p-value=0.001 $\Rightarrow$ Reject H_0

We have evidence that award preference is associated with which SAT score is higher.

Chi-Square (χ^2) Distribution

- If each of the expected counts are at least 5, AND if the null hypothesis is true, then the χ^2 statistic follows a χ^2 –distribution, with degrees of freedom equal to

df = (number of rows – 1)(number of columns – 1)

- Award by HigherSAT:

$$df = (2 - 1)(3 - 1) = 2$$

Chi-Square Test for Association

Note: The χ^2-test for two categorical variables only indicates **if** the variables are associated. Look at the contribution in each cell for the possible nature of the relationship.

Detailed Sample Table

	Academy	Nobel	Olympic	Total
Math	21 17.9 0.536	68 84.9 3.36	116 102.2 1.86	205
Verbal	10 13.1 0.733	79 62.1 4.591	61 74.8 2.542	150
Total	31	147	177	355

Observed, Expected, Contribution to χ^2

Chi-Square Test for Association

1. H_0 : The two variables are not associated
   H_a : The two variables are associated

2. Calculate the expected counts for each cell:

$$\text{Expected Count} = \frac{\text{row total} \times \text{column total}}{n}$$

3. Calculate the χ^2 statistic:

$$\chi^2 = \sum \frac{(\text{observed - expected})^2}{\text{expected}}$$

4. Compute the p-value as the area in the tail above the χ^2 statistic using either a randomization distribution, or a χ^2 distribution with df = (r – 1) (c – 1) if all expected counts > 5

5. Interpret the p-value in context.