CHI SQUARE

Hypothesis Testing

- To date, we have learned three methods for hypothesis testing:
  - One-Sample T-Tests
  - Two-Sample T-Tests
  - ANOVA

Hypothesis Testing

- Testing the relationship of two variables, whether the independent variable has an effect on the dependent variable (but not *cause & effect*!)
- Although they are different methods, they are accomplishing the same same task: testing means for difference

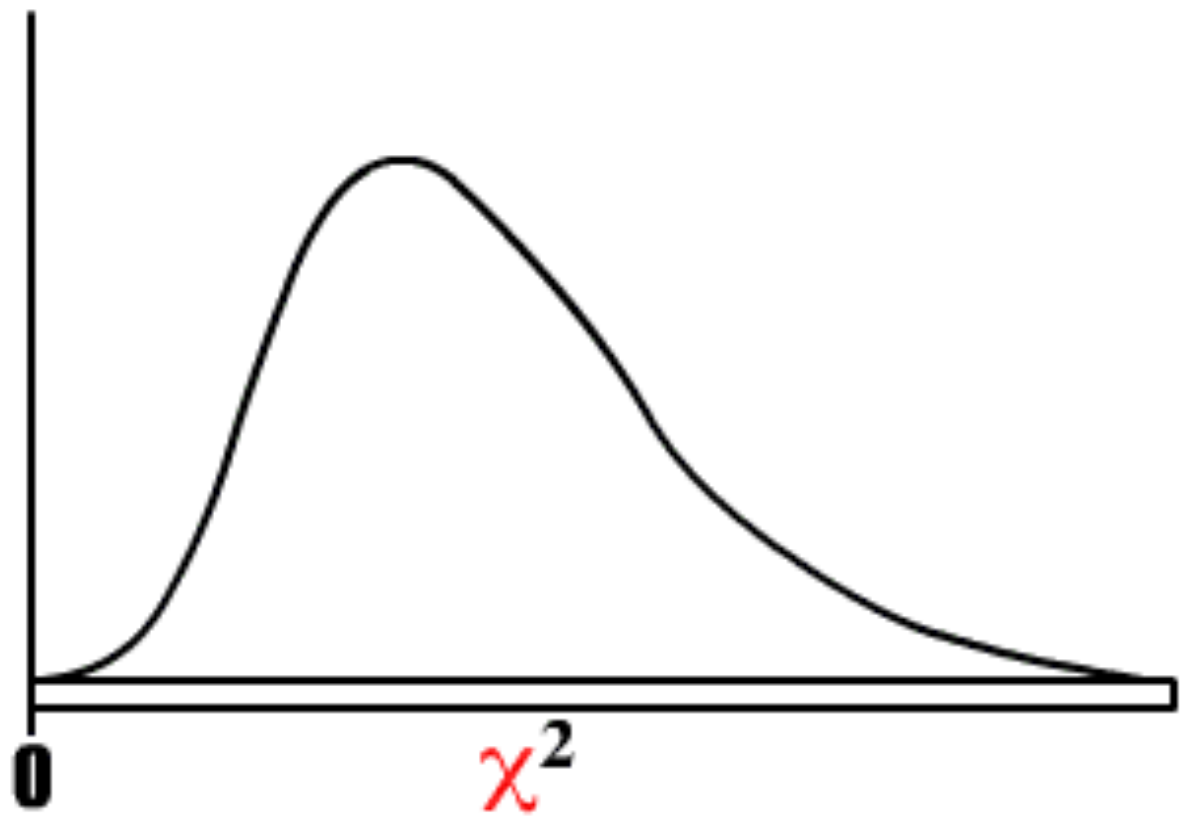
Chi-Square

- Also tests the research hypothesis that the dependent variable differs as the independent variable changes
- But Chi-Square is much more flexible

Chi-Square

- NOT testing means or proportions, rather levels in a cross tabulation (nonparametric)
- Requires that both the dependent and independent variables be the lowest level of measurement: nominal or ordinal (never Interval-Ratio)

Chi-Square Distribution



Quiz Yourself

- Would it be possible to conduct a Chi-Square test if you had interval-ratio level variables?

Variable Measurement Requirements

Statistic	**Dist**	**Dependent Variable**	**Independent Variable**
One-Sample Hypothesis	t/Z	Interval-ratio	Nominal or Ordinal (1 value, compared against population mean)
Two-Sample Hypothesis	t/Z	Interval-ratio	Nominal or Ordinal (2 values a.k.a. binary)
ANOVA	F	Interval-ratio	Nominal or Ordinal (3 or more values, but usually no more than 5)
Chi-Square	χ^2	Nominal or Ordinal	Nominal or Ordinal

Crosstabs Revisited

- A more complex frequency table, showing frequencies by *two* variables

Crosstabs Revisited

- What do you hypothesize is the relationship between training and learning?

Crosstabs Revisited

Learned Something?	Received Training?
Yes	Yes
Yes	No
No	No
No	No
Yes	Yes
Yes	Yes
No	Yes
No	No
Yes	Yes
No	No

Crosstabs Revisited

		Got Training	
		N	Y
Learned Something	N		
	Y		

Crosstabs Revisited

		Got Training	
		N	Y
Learned Something	N	4	1
	Y	1	4

Crosstabs Revisited

		Got Training		
		N	Y	Tot
Learned Something	N	4	1	5
	Y	1	4	5
	Tot	5	5	10

Crosstabs Revisited

		Got Training		
		N	Y	Tot
Learned Something	N	4 80%	1 20%	5 50%
	Y	1 20%	4 80%	5 50%
	Tot	5 100%	5 100%	10 100%

Do you see a relationship between training and learning?

		Got Training		
		N	Y	Tot
Learned Something	N	4 80%	1 20%	5 50%
	Y	1 20%	4 80%	5 50%
	Tot	5 100%	5 100%	10 100%

What We' d Expect With *NO* Relationship

		Got Training		
		N	Y	Tot
Learned Something	N	2.5 50%	2.5 50%	5 50%
	Y	2.5 50%	2.5 50%	5 50%
	Tot	5 100%	5 100%	10 100%

Chi-Square is calculated on the Crosstab frequencies

- Step 1: Sample is EPSEM/Random and both variables are nominal or ordinal
- Step 2:
  - H_0: The two variables are independent.
  - H_1: The two variables are dependent.

Chi-Square

- Step 3:
  - Alpha = 0.05
  - df = (r-1)(c-1)
  - χ^2 critical = See appendix C

Chi-Square

- Step 4: Calculate the statistic
- Step 5: Interpret

- Compute by hand

Results in R

```
> table(learntrain$learned, learntrain$trained)

        FALSE TRUE
  FALSE     4    1
  TRUE      1    4
> chisq.test(table(learntrain$learned, learntrain$trained),
correct=FALSE)

     Pearson's Chi-squared test

data:  table(learntrain$learned, learntrain$trained)
X-squared = 3.6, df = 1, p-value = 0.05778

Warning message:
In chisq.test(table(learntrain$learned, learntrain$trained),
correct = FALSE) :
  Chi-squared approximation may be incorrect
```

Categorical Data

- Sometimes we have data consisting of the frequency of cases falling into unique categories
- Examples:
  - Number of people voting for different politicians
  - Numbers of students who pass or fail their degree in different subject areas.
  - Number of patients or waiting list controls who are ‘free from diagnosis’ (or not) following a treatment.

An Example: Dancing Cats and Dogs

- Analyzing two or more categorical variables
  - The mean of a categorical variable is meaningless
    - The numeric values you attach to different categories are arbitrary
    - The mean of those numeric values will depend on how many members each category has.
  - Therefore, we analyze frequencies.
- An example
  - Can animals be trained to line-dance with different rewards?
  - Participants: 200 cats
  - Training
    - The animal was trained using either food or affection, not both)
  - Dance
    - The animal either learnt to line-dance or it did not.
  - Outcome:
    - The number of animals (frequency) that could dance or not in each reward condition.
  - We can tabulate these frequencies in a **crosstab**.

Crosstab

TABLE 18.1 Contingency table showing how many cats will line dance after being trained with different rewards

		Training		
		Food as Reward	Affection as Reward	Total
Could They Dance?	Yes	28	48	76
	No	10	114	124
	Total	38	162	200

Pearson’s Chi-Square Test

- Use to see whether there’s a relationship between two categorical variables
  - Compares the frequencies you observe in certain categories to the frequencies you might expect to get in those categories by chance.
- The equation:

$$\chi^2 = \sum \frac{\left(Observed_{ij} - Model_{ij}\right)^2}{Model_{ij}}$$

  - i represents the rows in the contingency table and j represents the columns.
  - The observed data are the frequencies the contingency table
- The ‘Model’ is based on ‘expected frequencies’.
  - Calculated for each of the cells in the contingency table.
  - n is the total number of observations (in this case 200).

$$\text{Model}_{ij} = E_{ij} = \frac{\text{Row Total}_i \times \text{Column Total}_j}{n}$$

- Test Statistic
  - Checked against a distribution with (r − 1)(c − 1) degrees of freedom.
  - If significant then there is a significant association between the categorical variables in the population.
  - The test distribution is approximate so in small samples use **Fisher’s exact test**.

Important Points

- The chi-square test has two important assumptions:
  - Independence:
    - Each person, item or entity contributes to only one cell of the contingency table.
  - The expected frequencies should be greater than 5.
    - In larger contingency tables up to 20% of expected frequencies can be below 5, but there a loss of statistical power.
    - Even in larger contingency tables no expected frequencies should be below 1.
    - If you find yourself in this situation consider using Fisher's exact test.
- Proportionately small differences in cell frequencies can result in statistically significant associations between variables if the sample is large enough
  - Look at row and ***column* percentages** to interpret effects.

Output

```
> table(cat_train$dance, cat_train$training)

                 food affection
  didn't dance    10       114
  danced          28        48
> chisq.test(table(cat_train$dance, cat_train
$training), correct=FALSE)

     Pearson's Chi-squared test

data:  table(cat_train$dance, cat_train$training)
X-squared = 25.3557, df = 1, p-value = 4.767e-07
```

Interpretation

- There was a significant association between the type of training and whether or not cats would dance χ^2 (1) = 25.36, p< .001.

To Sum Up …

- We approach categorical data in much the same way as any other kind of data:
  - we fit a model, we calculate the deviation between our model and the observed data, and we use that to evaluate the model we've fitted.
- Two categorical variables
  - Pearson's chi-square test