CHI SQUARE ANALYSIS

INTRODUCTION TO NON-PARAMETRIC ANALYSES

HYPOTHESIS TESTS SO FAR…

- We've discussed
  - One-sample t-test
  - Dependent Sample t-tests
  - Independent Samples t-tests
  - One-Way Between Groups ANOVA
  - Factorial Between Groups ANOVA
  - One-Way Repeated Measures ANOVA
  - Correlation
  - Linear Regression
- What do all of these tests have in common?

PARAMETRIC VS. NON-PARAMETRIC

- **Parametric Tests** – Statistical tests that involve assumptions about or estimations of population parameters.
  - (what we've been learning)
  - E.g., normal distribution, interval/ratio level measurement, homogeneity of variance
- **Nonparametric Tests**
  - Also known as distribution-free tests
  - Statistical tests that do not rely on assumptions of distributions or parameter estimates
  - E.g., does not assume interval/ratio, no normality assumption
  - (what we're going to be introducing today)

SOME NON-PARAMETRIC TESTS

- Frequency Data
  - Chi-Square (χ^2) Analysis
    - χ^2 Goodness-of-Fit test (one variable)
    - χ^2 Test of Independence (2 or more variables)
- Non-normal Data (e.g., ordinal)
  - Mann-Whitney U (NP analogue of Independent Samples t-test)
  - Wilcoxon Signed Ranks Tests (NP analogue of Dependent Samples t-test)
  - Kruskal-Wallis One-Way Analysis of Variance (Between)
  - Friedman's Rank Test for *K* correlated samples (Within)

CHI-SQUARE

- The χ^2 Goodness-of-Fit test
  - Used when we have distributions of frequencies across two or more categories on *one* variable.
  - Test determines how well a hypothesized distribution fits an obtained distribution.
- The χ^2 test of independence.
  - Used when we compare the distribution of frequencies across categories in two or more independent samples.
  - Used in a single sample when we want to know whether two categorical variables are related.

CHI-SQUARE GOODNESS OF FIT TEST

- Quarter Tossing
  - Probability of Head?
  - Probability of Tails?
  - How can you tell if a Quarter is unfair when tossed?
  - Imagine a flipped a quarter 50 times, what would we expect?

Heads	Tails
25	25

CHI-SQUARE GOODNESS OF FIT TEST

- Which of these scenarios seems probable with a "fair" coin?

Heads	Tails
20	30

Heads	Tails
15	35

Heads	Tails
10	40

Heads	Tails
5	45

CHI-SQUARE GOODNESS OF FIT TEST

- We can compare it to our expectation about "fair" coins

	Heads	Tails
Observed	17	33
Expected	25	25
O-E	-8	8

CHI-SQUARE GOODNESS OF FIT TEST

- We can test to see if our observed frequencies "Fit" our expectations
- This is the χ^2 Goodness-of-Fit test

$$\chi^2 = \sum \frac{(O-E)^2}{E}; df = \#categories - 1$$

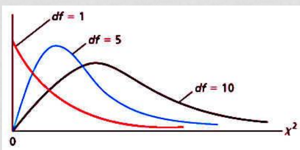
- This converts the difference between the frequencies we observe and the frequencies we expect to a distribution with known probabilities

CHI-SQUARE GOODNESS OF FIT TEST

- Hypothesis Test
  1. H_0: P(heads) = .5
  2. H_1: P(heads) ≠ .5
  3. a = .05
  4. Type of test = χ^2 goodness-of-fit

CHI SQUARE DISTRIBUTION

5. DF = 2 – 1 = 1;



See Chi-square table

$\chi^2(1)$ = 3.841; If χ^2 observed is larger than 3.841, reject the null hypothesis

CHI-SQUARE GOODNESS OF FIT TEST

6. Do the test: For our coin example

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(17-25)^2}{25} + \frac{(33-25)^2}{25} =$$

$$\chi^2 = \frac{(-8)^2}{25} + \frac{(8)^2}{25} = \frac{64}{25} + \frac{64}{25} = 2.56 + 2.56 = 6.55$$

7. Since 6.55 > 3.84, reject the null.
   There is evidence that the coin is not fair.

	Heads	Tails
Observed	17	33
Expected	25	25

CHI-SQUARE TEST OF INDEPENDENCE

- Used when we want to know if frequency responses of one categorical depend on another categorical variable (sounds like an interaction, right?)

	Pro Choice	Pro Life
Democrats	↑	↓
Republicans	↓	↑

CHI-SQUARE TEST OF INDEPENDENCE

- We compare observed vs. expected frequencies as in the goodness-of-fit test but the expectant frequencies aren't as easy to figure out because of the row and column totals.

	Column 1	Column 2	Column 3	
Row 1				Total R_1
Row 2				Total R_2
Row 3				Total R_3
	Total C_1	Total C_2	Total C_3	Total

CHI-SQUARE TEST OF INDEPENDENCE

- Expectant frequencies for each cell is found by multiplying row and column totals then dividing by the grand total.

$$E = \frac{RxC}{T}$$

CHI-SQUARE TEST OF INDEPENDENCE

- Example: Researchers stood on a corner and watched drivers come to a stop sign. They noted their gender and the type of stop they made.

	Male	Female	
Full Stop	8	15	23
Rolling Stop	17	5	22
No Stop	5	1	6
	30	21	51

CHI-SQUARE TEST OF INDEPENDENCE

	Male	Female	
Full Stop	8	15	23
Rolling Stop	17	5	22
No Stop	5	1	6
	30	21	51

	Male	Female	
Full Stop			23
Rolling Stop			22
No Stop			6
	30	21	51