

FREQUENCY ANALYSIS

- Categorical Variables
- Event and Exposure
- Contingency Table
- 2x2 Table
- rxc Table
- Chi-square Test
- Fisher's Exact Test
- Odds Ratio
- Risk Ratio / Relative Risk

OBJECTIVE

To identify basic characteristics & analysis for categorical outcomes with a categorical predictor

PURPOSE

To review unadjusted categorical outcome analysis (no covariates) & the interpretation of commonly reported measures for contingency tables

VARIABLE TYPES

- **Outcomes (Y, Dependent Variable, Response): A**
Dichotomous Yes/No, e.g. Event, Failure, Cure, Relapse, Remission
OR Multiple Levels, e.g. HIV-, HIV+, AIDS
- **Predictors (X, Independent):**
Treatment Groups, Intervention Arms, Exposure levels, Race, Sex, Age Group, Disease Classes

CONTINGENCY TABLES & TESTS

CONTINGENCY TABLE: a table used to display counts and percentages, n(%), of two categorical variables in r (rows) x c (columns)

2x2 table		Outcome		
		Yes	No	Row Totals
Predictor	Treatment A	a	b	a+b
	Treatment B	c	d	c+d
Column Totals		a+c	b+d	N = a+b+c+d

- **Row & Column Percents:**
 - Row % are conditioned on the row, i.e. Out of Treatment A, x% had the outcome and y% did not have the outcome:
 - $x = a / (a+b)$ $y = b / (a+b)$
 - Column % are conditioned on the column, i.e. For those that had the Outcome, j% were in Treatment A and k% Treatment B:
 - $j = a / (a+c)$ $k = c / (a+c)$
- **Percentages vs Proportions:**
 - Percentages range from 0-100%
 - Proportions range from 0-1 (proportion x100 → percentage)
- **Tests of Association:**
 - Chi-Square (χ^2) test or Fisher's Exact test (for small N)
 - H0: No association between Treatment Group and Outcome
 - HA: There is a difference in the % with Outcome by Treatment

ODDS & RISK RATIOS

ODDS RATIO: The odds of having a Yes Outcome for Treatment B relative to Treatment A: $OR = \frac{bc}{ad}$

- Range 0-1: OR = 1 no difference in likelihood of event by treatment
- OR < 1: Treatment B has a lower odds of (less likely to) having the Outcome compared to Treatment A, also called a protective effect
- OR > 1: Treatment B has a higher odds compared to Treatment A

RISK RATIO: The risk of having a Yes Outcome for Treatment B relative to Treatment A: $RR = (\frac{c}{c+d}) / (\frac{a}{a+b})$

- Range 0 to infinity: RR = 1 means no difference in risk of having the event by treatment group (also called Relative Risk)
- RR < 1: Treatment B has a lower risk than Treatment A
- RR > 1: Treatment B has a higher risk than Treatment A

ODDS vs RISK: When the event rate is small, OR and RR are similar values but still have different interpretations

- OR is used for cross-sectional or retrospective studies, **NOT prospective studies**. Answers the question: how likely is the event for one treatment compared to another?
- RR is used for prospective studies. Answers the question: Looking forward for a given time period, what is the risk of the event for one treatment compared to another?