VIII. STATISTICS

Part III

IN THIS CHAPTER:

- Specific measures of association and tests of significance for bivariate relationships
- Ways of testing the significance of differences between groups
- A summary chart for common statistics

TESTING: TAKING A DEEPER LOOK

WITHIN BIVARIATE ANALYSES, THERE ARE MANY PAIRINGS OF VARIABLE TYPES (VISUALIZED BELOW). THEREFORE, WE WILL SPLIT OUR DISCUSSION INTO THREE SECTIONS:

1) TESTS OF ASSOCIATION WITH NOMINAL VARIABLES

2) TESTS OF ASSOCIATION WITH ORDINAL OR INTERVAL VARIABLES

3) TESTING DIFFERENCES BETWEEN GROUPS WITH ORDINAL OR INTERVAL VARIABLES.

	Binary nominal	Non-binary nominal	Ordinal	Interval
Binary nominal				
Non-binary nominal				
Ordinal				
Interval				

SECTION 1: TESTS WITH NOMINAL VARIABLES

Data involving only nominal or ordinal variables with a few scale points are usually best represented in **contingency tables**. Contingency tables, also known as **cross-tabulations**, can be used to examine relationships between two nominal variables, or one nominal and one ordinal variable (provided that the scale for the ordinal variable doesn't contain that many points).

Let's look at an example. Let's start with a cross-tab with two binary nominal variables: gender (could be male or female) and primary subsistence role (could be primarily hunter *or* primarily gatherer— not both). Below is our hypothesis, stated along with the null hypothesis, as well as three examples of data:

- **Hypothesis**: Gender will be associated with subsistence role.
- **Null hypothesis**: There will be no association between gender and subsistence role.

EXAMPLE A

	Male	Female
Hunter	25	25
Gatherer	25	25

In this example, we see that the roles are evenly distributed between the sexes: 25 males are hunters, 25 males are gatherers, 25 females are hunters, and 25 females are gatherers. This suggests no relationship between gender and subsistence type. This result seems to be consistent with the null hypothesis.

EXAMPLE B

	Male	Female
Hunter	50	0
Gatherer	0	50

In this example, the roles are completely segregated by gender, suggesting that there is a one-to-one relationship between gender and subsistence role. This result seems to be consistent with the alternative hypothesis.

EXAMPLE C

	Male	Female
Hunter	37	9
Gatherer	13	41

In this example, the distribution seems to support the alternative hypothesis, but compared to example B it is a weaker relationship. Is it statistically significant? How strongly are the variables related? We discuss how to evaluate these questions shortly.

WHAT CAN BE SAID ABOUT THE CONTINGENCY TABLES IN THE LAST SLIDE?

First, we want to ask whether the results are **statistically significant**. That is, **are the results likely to be due to chance?** Second, if the results are significant, we want to ask **how strongly** the variables are associated.

For contingency tables with only nominal variables, we usually do one of two tests of significance. These will be covered in the following slides.

POSSIBLE TESTS OF SIGNIFICANCE WITH ONLY NOMINAL VARIABLES:

- **Fisher's Exact Test** (usually applies to cross-tabulations with two binary nominal variables). Fisher's Exact Test finds all the possible distributions of data in a 2x2 contingency table and then calculates the chances of getting your result or a stronger one. It is the most direct way to obtain a p-value, but it is generally only used for small sample sizes.
- **Chi-square** x^2 can be used with all pairs of nominal variables as long as the number of cases is not too small in each cell.

EXAMPLE A

	Male	Female
Hunter	25	25
Gatherer	25	25

$\chi^2 = 0$ d.f. = 1 $\varphi = 0.00$ p-value = 1 (Not significant)

EXAMPLE B

	Male	Female
Hunter	50	0
Gatherer	0	50

$\chi^2 = 100$ d.f. = 1 $\varphi = 1.00$ $p<.001$

EXAMPLE C

	Male	Female
Hunter	37	9
Gatherer	13	41

$\chi^2 = 31.562$ d.f. = 1 $\varphi = 0.56$ $p<.001$

In Examples A, B, and C, we show the x^2 score and the significance (p-value) of that score. Both Example B and Example C are statistically significant because p is less than .05.

Strength of association is usually assessed if the variables are ordinal, interval or ratio. However, with nominal variables we can imagine that a stronger relationship is one in which one variable perfectly predicts the other. Recalling the 2x2 example slides as seen above, Example B is clearly perfectly predictive and is a stronger relationship than Example C. A number of coefficients of association have been devised to assess strength of relationship for nominal variables. Some measures can be found on the following slides. We have computed a phi coefficient for each table. Example C demonstrated a moderate phi correlation of 0.56.

POSSIBLE MEASURES OF ASSOCIATION WITH NOMINAL VARIABLES*

- **Phi** (used only for 2x2 tables)
- **Yule's Q** (used only for 2 x 2 tables)
- **Cramer's V**
- **Lambda** (based on how well you can predict Y from X and X from Y)

*These can all be tested for statistical significance

SECTION 2: TESTS OF ASSOCIATION WITH ORDINAL OR INTERVAL VARIABLES

SITUATION A: TWO ORDINAL VARIABLES

If you want to test for a relationship between two ordinal variables, there are a number of measures of association you can use. The useful thing about some of these measures of association is that they can be tested for significance. See the list below:

Possible measures of association for two ordinal variables (all can be tested for significance:

- **Spearman's rho**
- **Kendall's tau**
- **Gamma**

Additionally, if you have one ordinal and one interval/ratio variable, you can use any of the above tests.

SITUATION B: TWO INTERVAL/RATIO VARIABLES

If you have two interval/ratio variables, one of the primary options available is to run what is called a **linear regression**.

This involves plotting your cases on a graph with an X and Y axis and then examining the scatterplot of points. The arrangement of points may resemble a line, a curve, or really no pattern at all. If you suspect a curve, there are measures (such as eta) you can use to investigate the relationship between the variables.

In this course, we will focus on the scatterplots that seem to be arranged around a line, as the analysis is more straightforward in these cases.

The primary method to use in that scenario is linear regression (ordinary least squares regression is the typical type). When you run a regression with a statistics program, the computer will calculate a line of best fit for your scatterplot and provide you with an equation for that line:

$$\text{Expected Y} = b * X + c$$

Where X and Y are the variables, *b* is the slope of the line, and *c* is the Y-intercept. With this equation, you can calculate the expected value of the Y variable given any X value. The actual point you see on the graph shows the *observed* Y value at a given X value, and it may not match the expected Y value. The more clear the pattern in the data, however, the more the expected and observed values will correspond.

In addition to the regression equation, a statistics program will usually also produce Pearson's r, a measure of association that varies from -1 to 1.

A negative *r* represents a negative association; a positive *r* represents a positive association. Pearson's *r* can be tested for significance—a critical step in ensuring a non-zero relationship between your two interval/scale variables.

SECTION 3: TESTING DIFFERENCES BETWEEN GROUPS

The following questions are not about degree of association, but rather whether the differences between groups on some ordinal, interval, or ratio measure are statistically significant. For example:

- Do agriculturalists tend to live in signifcantly wetter regions than pastoralists?
- Do agriculturalists and pastoralists differ significantly in population density on average?
- Do agriculturalists, pastoralists, and hunter-gatherers tend to differ in terms of average rainfall in their habitats? If so, are these difference significant on average?*

*See the next slide for hypothetical data.

The first two questions are comparing two groups: agriculturalists and pastoralists. The last is comparing three groups. Determination of the appropriate statistical tests to use depends on the number of groups compared (two vs. two or more) and whether the dependent variable is interval/ratio or ordinal.
The hypothesis will usually state the expected difference (e.g. "Agriculturalists and pastoralists will differ in population density, with agriculturalists generally having greater populaton density than pastoralists").

Table: Average Annual Precipitation (mm) for 30 Societies

Hunter-Gatherers	Pastoralists	Agriculturalists
16.0	24.0	306.0
211.0	123.0	27.0
43.0	111.0	64.0
98.0	88.0	344.0
112.0	52.0	23.0
13.0	132.0	155.0
330.0	12.0	5.0
51.0	23.0	182.0
55.0	14.0	55.0
144.0	198.0	96.0
Mean: 107.3 mm	**Mean: 77.7 mm**	**Mean: 125.7 mm**

For this type of data, we do not run measures of association; we only run tests of significance to determine whether the difference in means is due to chance. Some potential tests are explored in the next slide.

IF THE NOMINAL VARIABLE HAS ONLY TWO GROUPS:

• & THE DEPENDENT VARIABLE IS AN INTERVAL OR RATIO SCALE

The **t-test** is ideally used when four assumptions about the data are met:

- The data are measured on **interval or ratio scales.**
- The populations have **"normal" distributions** on the measured variables: that is, the distributions follow a bell-shaped curve.
- The variances for the two populations of the binary nominal variable are approximately the same.
- The cases were **randomly sampled** from their respective populations.

The second and third assumptions for the t-test make it a **parametric test,** meaning that the test works on certain assumptions about the underlying distributions of populations. **Non-parametric tests** do not require such assumptions.

• & THE DEPENDENT VARIABLE IS AN ORDINAL SCALE OR PARAMETRIC ASSUMPTIONS ARE NOT MET

The **Mann-Whitney U** or **Kolmogorov-Smirnov Z** are ideal.

IF THE NOMINAL VARIABLE HAS THREE OR MORE GROUPS:

- ## & THE DEPENDENT VARIABLE IS AN INTERVAL OR RATIO

ANOVA (Analysis of Variance) is ideal. However, if the nominal variable has **three or more categories**, as in the above example, a test for significant difference in means only gets you so far. If a test of the data above yielded a significant difference in rainfall among the three subsistence types, we would still need to run more tests to determine whether each group was different from the others, or whether only one group was significantly different from the other two. This would be possible on most computer statistics programs.

- ## & THE DEPENDENT VARIABLE IS ORDINAL

Kruskal-Wallis is ideal.

REVIEW OF STATISTICAL TESTS

On the following slide is a reference table with information about **Tests of Significance** (in yellow) and **Measures of Association** (in purple). Again, this list is not comprehensive; it merely includes a handful of commonly used tests and measures you are likely to encounter in reading or use in your own research.

Test	Notation	Description
Fisher's exact test		Nonparametric test of significance for a contingency table with two binary nominal variables. Generally used with a small sample. Yields an exact p-value.
Chi-square	χ^2	Nonparametric test of significance for a contingency table with nominal variables and more than 20 to 30 cases. Used to calculate a p-value.
t-test	t	Parametric test used with one binary nominal variable and one interval variable. Can be used to calculate p-value. Reported with degrees of freedom.
Mann-Whitney U test	u	Non-parametric test analogous to the t-test. Used with one binary nominal variable and one ordinal variable. Can be used to calculate p-value.
Kolmogorov-Smirnov test		Non-parametric test analogous to the t-test. Used with one binary nominal variable and one ordinal variable. Can be used to calculate p-value.
ANOVA		Parametric test used with one non-binary nominal variable and one interval/ratio variable Can be used to calculate p-value.
Kruskal-Wallis One-Way Analysis of Variance		Nonparametric test analogous to ANOVA, used with one non-binary nominal variable and one ordinal variable. Can be used to calculate p-value.
Phi	φ	Nonparametric measure of association for binary nominal variables in 2x2 tables.
Goodman and Kruskal's Gamma	G	Nonparametric measure of association between two ordinal variables. Yields a value between -1 and 1. A Gamma can achieve a value of 1 (or -1) if the relationship is asymmetrical.
Yule's Q	Q	Nonparametric measure of association for binary nominal variables in 2x2 tables; special case of Gamma for ordinal variables; will be a +1 or -1 if there is one zero cell in a table.
Spearman's rho	ρ	Nonparametric measure of association for two ordinal variables, or one ordinal and one interval/ratio. Yields a value between -1 and 1.
Kendall's tau	τ	Nonparametric measure of association for two ordinal variables, or one ordinal and one interval/ratio.Yields a value between -1 and 1.
Pearson's r	r	Parametric correlation between two interval/ratio variables. Yields a value between 1 and -1.
Cramer's V	V	Nonparametric measure of association for nominal variables based on chi-square; for more than 2 x 2 tables

SUMMARY

- **Specific bivariate tests of significance** and **measures of association** depend on:
    - The **combination of variables** being considered in the hypothesis
    - The **assumptions about the distributions** of variables (parametric vs. nonparametric)
    - Whether you are looking for **relationships between variables or differences between groups**

CITATION AND TERMS OF USE

Using this course: The material in this course is intended for educational purposes only--either by individuals for personal use or for classroom use by instructors with appropriate attribution. For any commercial use or other uses please contact HRAF. (Email: <hraf@yale.edu>).

This chapter is from: Carol R. Ember. 2016. Introducing Cross-Cultural Research. Human Relations Area Files. <http://hraf@yale.edu/ccc/>