

1 Statistical Inference: Test of Significance

2 I Introduction

- A. **Descriptive Statistics**
 1. Descriptive measures of variables in the sample are called *statistics*
 - a. \bar{X} , S , r , R ---univariate, bivariate, multivariate
 - b. These values will vary from sample to sample
 2. Descriptive measures of variables in the population are called *parameters*
 - a. μ , σ , ρ --- univariate, bivariate, multivariate
 - b. These have fixed, but unknown values
 3. Question: Are sample *statistics* an accurate reflection of the population *parameters*?

3

- B. **Statistical Inference**
 1. Statistical inference is the process of generalizing from the *sample statistics* to the *population parameters*
 2. You can infer your sample results **DO** or **DO NOT** apply to the population
 - a. **Whatever your conclusion, you may be wrong**
 - b. **You cannot eliminate the risk of being wrong**
 - c. **What we do is measure the risk we are taking when we make an inference (i.e., calculate probabilities)**

4 Introduction Cont'd

3. Test of Significance or "Statistical Tests" provide us with information (e.g., probabilities) and criteria (e.g., .05) to use in deciding what inference to make
4. Why in some "Tests of Significance" do we use inferential statistics instead of our descriptive statistics?
 - a. Use the sampling distribution of the sample statistic to calculate probabilities
 - b. Some sample statistics have unknown sampling distributions (e.g., r)
 - c. We convert these descriptive statistics to statistics with known sampling distributions
---Chi Square, T, F

5 Introduction Cont'd

- C. **What is to come?**
 1. Brief over view of statistical test
 2. Formal steps involved
 3. Issues

6 II General Overview of a Typical Statistical Test

- **Step 1** Formulate research hypothesis
(e.g., **Income and education are positively related**)
 1. Usually the research hypo is that there is some relationship between variables or some difference between groups
- **Step 2** We carry out the project and calculate an appropriate measure of association
(e.g., we find τ_b between education and income is .60)

7 General Overview Cont'd

- Step 3 Calculate the probability of obtaining a measure of association at least this large between our variables when there is no such relationship in the population

1. This is the significance value or probability value produced in most statistical programs

8 General Overview Cont'd

- **Step 4** Decide whether or not to infer the observed relationship in the sample, as measured by the measure of association, means there is a corresponding relationship in the population
 1. If the probability of getting the observed relationship in the sample when there is no relationship in the population is very small, (i.e., less than .05) then **we infer** that we got such a relationship in the sample because such a relationship exists in the population
 2. If the probability of getting the observed relationship in the sample when there is no relationship in the population is large (i.e., over .05), then **we infer** that the relationship in the sample simply occurred by chance.

9

III Formal Steps in a Modern Statistical Test

- **Step 1:** Formulate the research hypothesis (H_1)

(Note it is always stated in terms of specific parameters as they exist in the population)

Examples:

- a. The mean number of people in an American Catholic family is larger than the mean number of people in a Protestant family
- b. There is a nonzero tau between income and political participation
- c. The tau between income and political participation is greater than the tau between education and political participation

10 Modern Statistical Test Cont'd

– Symbols:

– Research hypothesis usually designated as H_1 : and the components are written in Greek letters

- $H_1: \mu_c > \mu_p$
- $H_1: \tau_{x_1y} \neq 0$
- $H_1: \tau_{x_1y} > \tau_{x_2y}$

11

- **Step 2:** Formulate the Null hypothesis (H_0 .)

– The Null hypothesis is that the research hypothesis is wrong --- one of the other mutually exclusive logical possibilities exists, not the one specified in the research hypothesis

Since H_1 is usually that there is a relationship or is some difference between groups, then the null hypothesis is usually that there is no relationship or no difference between groups.

– H_0 includes all logical outcomes other than H_1 .

- $H_1: \tau_{xy} \neq 0$ $H_0: \tau_{xy} = 0$
- $H_1: \mu_c > \mu_p$ $H_0: \mu_c \leq \mu_p$
- $H_1: \tau_{x_1y} > \tau_{x_2y}$ $H_0: \tau_{x_1y} \leq \tau_{x_2y}$

– We directly test the null hypothesis and by inference accept or reject the research hypothesis

12

- **Step 3:** Calculate the probability of getting the observed sample statistic assuming the null hypothesis is true
 - If the sampling distribution is known, this is used to generate the probability (e.g., lambda and Kendal's tau have normal sampling distributions)
 - If the sampling distribution is unknown, we use a test statistic--statistic such as t or F that has a known sampling distribution

13

- **Step 4: Decision**
 - If the probability is below **.05** we **REJECT** the null hypothesis
 - Implication, we accept the research hypothesis.
 - If the probability is above **.05** we **ACCEPT** the null hypothesis
 - Implication, we reject the research hypothesis

14

IV Formal Steps in the Classical Statistical Test

- **Step 1:** Formulate the research hypothesis (H_1)
 - $H_1: \mu_c > \mu_p$
 - $H_1: \tau_{x1y} \neq 0$
 - $H_1: \tau_{x1y} > \tau_{x2y}$
- **Step 2:** Formulate the null hypothesis (H_0)
 - $H_0: \mu_c \leq \mu_p$
 - $H_0: \tau_{x1y} = 0$
 - $H_0: \tau_{x1y} \leq \tau_{x2y}$

15 Classical Statistical Test Cont'd

- **Step 3.** Calculate the sampling distribution of the sample statistic (If it has a known sampling distribution) or the sampling distribution of a test statistic such as Chi Square, t, or F
 1. We actually never do this-----like in sampling we make use of our knowledge of what this would be if we constructed it

16

2. If the null hypothesis says no relationship, we construct a sampling distribution of all possible outcomes assuming the null hypothesis true
 - For example,
 - a. We know that only a few samples will have taus of exactly zero
 - b. Some will have taus over 0
 - c. Others will have taus less than 0
 - d. These are simply the result of sampling error

17

- **Step 4.** Select a significance level or define the region(s) of rejection
 - $H_0: \tau_{xy} = 0$ [2 tail test -- 2 regions of rejection]

- $H_0: \tau_{x,y} \leq 0$ [1 tail test --1 region of rejection]
1. Alternatively this is called
 - a. Designating the critical region(s)
 - b. Designating the region(s) of rejection
 - c. Dividing the outcomes into those that mean the null hypothesis is accepted and those that mean the null hypothesis is rejected

18

19 **Classical Statistical Test** Cont'd

3. If the sampling distribution is normal---as for tau and Lambda--then the critical regions
 - a. Are greater than 1.96 SD for a two tailed test
 - b. Is greater than 1.65 SD for a one tailed test

20

- Step 5. Compute the sample statistic or test statistic and decide on the null hypothesis and by inference
 1. If sample statistic falls in the critical region(s), we reject the null hypothesis
 2. If sample statistic is outside of the critical region(s), we accept the null hypothesis

21

V Additional Topics in Inferential Statistics

• A. Type I and Type II Errors

1. **Type I** is rejecting a true null hypothesis----thus we falsely conclude the research hypothesis of a relationship exists
 - a. **Probability of Type I Error is Level of Significance!!!**
 - With .05 level of sig, 5 times out of 100 we get values in the tails simply by chance alone

22

V Additional Topics in Inferential Statistics

• A. Type I and Type II Errors

1. **Type I** is rejecting a true null hypothesis----thus we falsely conclude the research hypothesis of a relationship exists
 - a. **Probability of Type I Error is Level of Significance!!!**
 - With .01 level of Sig, 1 time out of 100 we get values in the tails simply by chance alone

23 **Type I and Type II Errors** Cont'd

2. **Type II** Error is accepting a false null hypothesis----thus incorrectly concluded there is no relationship or difference in the population when there is
 - a. **Probability of Type II Error is inversely related to the probability of Type I Error**
 - If we lower the probability of Type I we increase the probability of a Type II error

- 24 **Type I and Type II Errors** Cont'd
2. **Type II Error** is accepting a false null hypothesis----thus incorrectly concluded there is no relationship or difference in the population when there is
 - a. **Probability of Type II Error is inversely related to the probability of Type I Error**
 - If we lower the probability of Type I we increase the probability of a Type II error
- 25 **Type I and Type II Errors** Cont'd
2. **Type II Error** is accepting a false null hypothesis----thus incorrectly concluded there is no relationship or difference in the population when there is
 - a. **Probability of Type II Error is inversely related to the probability of Type I Error**
 - If we lower the probability of Type I we increase the probability of a Type II error
- 26 **Additional Topics** Cont'd
- **B. Test Statistics**
 1. Some sample statistics (e.g., r) do not have known sampling distributions, thus we use a statistic (called inferential statistic) that has a known distribution (e.g., t , F)
 2. Logic of the test is always the same
 - a. Is the probability of getting the observed sample statistic or difference less than .05 when the null hypothesis is true?
- 27 **Test Statistics** Cont'd
3. When the sampling distribution is not known
 - a. **We then convert our sample statistic to an inferential statistic that has a known distribution**
 - (E.G. Chi Square, F test, T test)
 - b. Like descriptive statistics, these test statistics have certain limitations and should be used in certain situations
 - (E.G. Chi Square is good for nominal data, F is good for statistics using variation from the mean -- r , b , β , R , etc)
 - c. Mechanics of calculating the various test statistics differ, but the logic of the test is the same
- 28 **Additional Topics** Cont'd
- **C. Levels of Significance**
 1. Now we simply calculate the probability of achieving the sample statistics if the null hypothesis is true and then make a decision
 2. Previously .05, .01, and .001 were popular benchmarks
 3. These are just guides----you make the decision on what inference is to be made
 - a. **What will you do if r_{x_1y} has a sig level of .049, whereas r_{x_2y} has a sig level of .051?**
 - b. **Will you say one exists and the other does not exist?**
- 29 **Additional Topics** Cont'd
- **D. Use on Population**
 1. Some argue we can use tests of sig on entire population
 2. Why?
 - a. **Critic says your measure of association for the population occurred by chance**
 - i. You can say here is the exact probability it occurred by chance
 - b. **You argue that your population is really a sample when we take time into account**
- 30 **Additional Topics** Cont'd
- **E. Statistical Significance vs Substantive Significance**
 1. Relatively simple----statistical significance does not mean substantive significance
 2. We accept the research hypothesis of some difference or some relationship, **BUT HOW IMPORTANT IS THIS DIFFERENCE OR RELATIONSHIP??**