The Logic of Statistical Inference -- Testing Hypotheses

- Confirming your research hypothesis (relationship between 2 variables) is dependent on ruling out:
  - Rival hypotheses
  - Research design problems (e.g. measurement error, non-representative sample), and/or
  - Chance—sampling *error--the* natural tendency of any sample to differ from the population from which it was drawn

Null Hypothesis

- A statement that there is no relationship between two variables of interest.

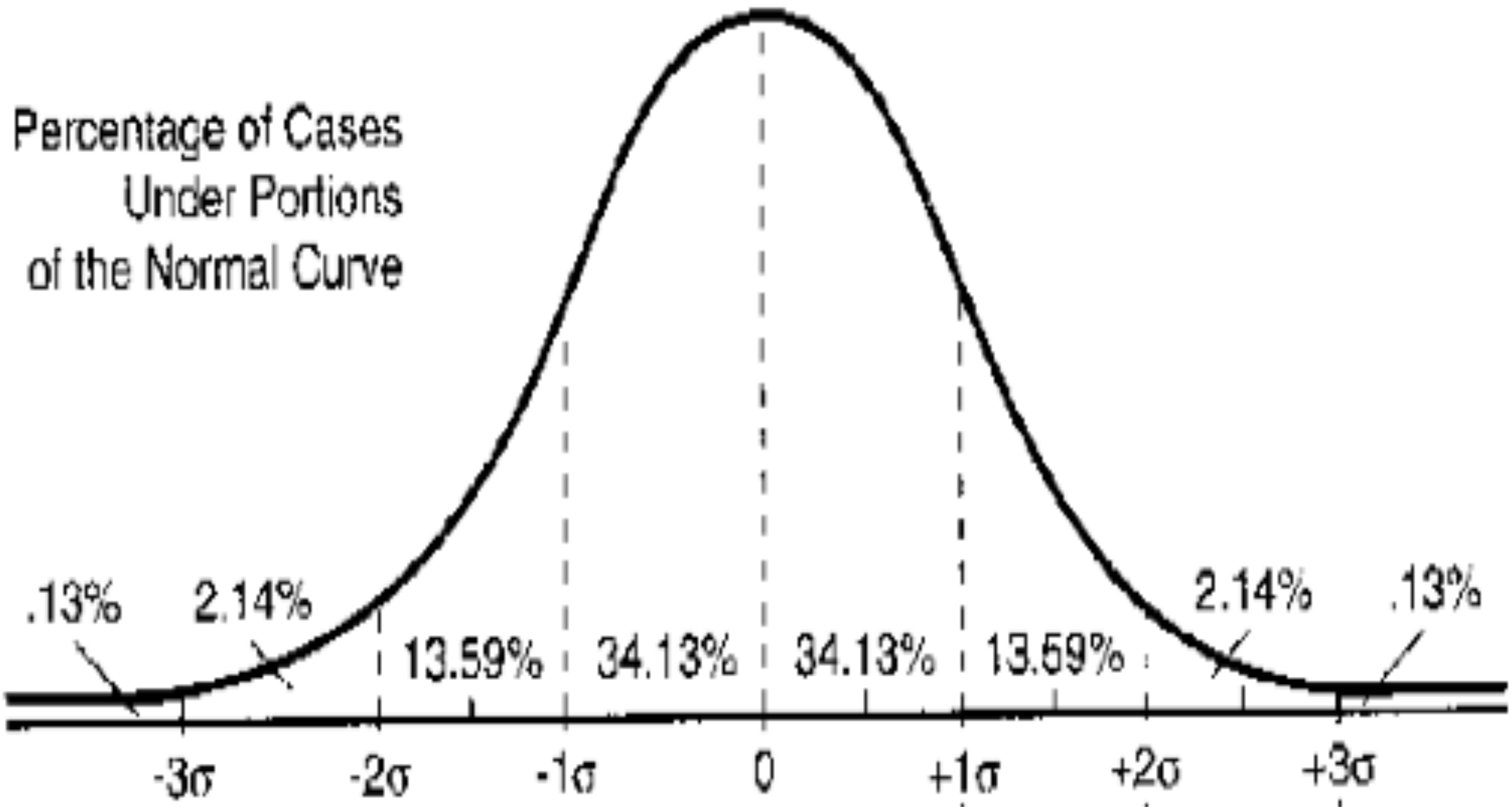
  Another way of saying it:

- Any relationship between these variables is only due to *chance,* not a real relationship that exists *in the population* (i.e. sampling error)

Reject the Null Hypothesis If:

- The “research hypothesis”, a.k.a “alternative hypothesis” proves correct
  - “There is a difference between these two variables (e.g. “There is a difference in outcomes, comparing the experimental and non-experimental groups’”), OR
  - “The experimental treatment will result in an improved outcome”

Area Under the Normal Curve and Standard Deviation Units



Standard Deviation Units

Statistical Inference

- Inferring whether or not a relationship between variables exists in the population, from your sample, requires disproving or rejecting the Null Hypothesis
  - By calculating (or computing) a test statistic
  - Then locating where the statistic falls in the theoretical sampling distribution, and from that
  - Determining the likelihood (probability) that the statistical result you found is due to chance alone (sampling error)

What is a *p* value?

- Probability: the likelihood that an event will occur (# actual events ÷ # possible events)
- How do we use probability in inference testing?
  - To quantify our confidence that our statistical result is not just due to sampling error (chance)
  - To confirm or disconfirm our hypotheses

Interpreting the *p* value

- Each statistic result is accompanied by a *p* value
- SPSS gives you the actual *p* value by using the statistic's computation formula and the distribution tables for the statistical test you've chosen
- If your actual *p* value (from SPSS) equals or is smaller than your alpha, then we can say the null hypothesis can be rejected

Summary—the 8 steps to hypothesis testing

1. Identify your independent variable(s)
2. Identify your dependent variable
3. State the Null Hypothesis
4. State the Alternative Hypothesis
5. Identify appropriate statistical test and alpha level
6. Review results (SPSS output)
7. Describe results & decision to reject or not reject Null
8. Discuss results and implications

Which Statistics?

- Using the area under the Normal curve to determine this “critical region” has an important requirement—the data must be “normally distributed” in the population, e.g. when plotted on a frequency polygon the line should follow the normal curve.
- At the very least, the data must be ratio or *interval*
- Relevant statistics for these data include t-tests, ANOVA, and linear regression

Application of Appropriate Statistical Tests

- **Chi-Square**: used with variables (both independent and dependent) that are categorical (nominal or ordinal) and with other samples that are clearly not distributed normally.

- **Dependent t-test**: used when you are working with two dependent groups that have an independent variable that is categorical and a dependent variable that is interval or ratio.

- **Independent t-test**: used when you have two dichotomous independent groups with an independent variable that is categorical and a dependent variable that is ratio or interval.

- **ANOVA** - Analysis of Variance – used when the independent variable has more than two attributes and is categorical and where the dependent variable is ratio or interval.

How ANOVA Works

- ANOVA is an "omnibus" test—it only tests the Null hypothesis of "no difference between the means"
- The ANOVA statistic and associated *p* value does not prove or disprove your research hypothesis by singling out one of the means as "significantly different than the others"

How does ANOVA answer the research question?

- With the use of multiple comparisons (called *post-hoc* tests)—each group's mean is contrasted with each other group's mean
- This is only done if the ANOVA test results in a p value less than our alpha. (If not, *game over!)*
- The multiple comparisons constitute the second of two types of statistical tests

First, the *F* Ratio

- The ANOVA statistic is called the "F ratio." It has the same function as the *t* statistic and the Chi Square value, and it has its own distribution table (built into SPSS) so it can also be associated with a *p* value -- Except the *F* ratio is easier to interpret:
  - The *F* ratio reflects the variation of means *between* the groups divided by the variation of means *within* the groups. So it tells you the % of variation that's related to the difference between the groups

Next, the Multiple Comparisons

- The *post-hoc* tests the Null Hypothesis that there are no differences between the two means in each comparison
- The *post-hoc* procedure adjusts for the inflated risk of making a Type I error, so that the combined probability of falsely rejecting the Null (Type I error) among all the comparisons is no more than your intended alpha (such as .05)

Reporting the ANOVA

- You would report the overall results of the ANOVA as: F = _____, $p < .05$
- And to address your research question (alternate hypothesis) you report on the multiple comparisons results with the associated *p* values.
- See lab assignments and readings for more examples