A Definition of Statistics….

- Statistics have been defined as “a collection of methods for planning experiments, obtaining data, and then analyzing, interpreting and drawing conclusions based on the data” (Triola, 1992, p.4).
  - This definition illustrates the strong link between research methodology and statistics
  - The research design is the foundation of a good study
  - No statistic can fix an inferior research design
  - “Garbage In …………..Garbage Out”

Review of Levels of Measurement

- Levels of measurement can be broken down into a hierarchy with four categories:
  - Nominal
  - Ordinal
  - Interval
  - Ratio
  - Lowest Level
  - Highest Level

- Statistics are appropriate or not appropriate depending on the levels of measurement of your data

Descriptive & Inferential Statistics

- All types of statistics can be categorized into two subgroups:
  - Descriptive statistics
    - Describe large amounts of data in an abbreviated way
    - Describe important characteristics of your data
  - Inferential statistics
Review of Descriptive Statistics

- Descriptive statistics are commonly grouped into these categories:
  - Tables & graphs of data
  - Measures of central tendency
    - Mean, Mode, Median
  - Measures of dispersion or variance
    - Range, Standard Deviation, Variance

Descriptive & Inferential Statistics

- All types of statistics can be categorized into two subgroups:
  - Descriptive Statistics
  - Inferential Statistics
    - Goes beyond mere description
    - Use sample data to draw conclusions and make inferences about a population

Samples and Populations

- Inferential statistics are used to determine whether researchers can make statements that the results reflect what would happen if the experiment were conducted again and again with multiple samples
  - Can we infer results taken from a sample to the population?

Probability Theory

- Probability is the likelihood of occurrence of some event or outcome.
- In probability theory, we know the population and are trying to see how well the sample represents the population
- Probability in inferential statistics looks at how likely an event will occur if there is no difference in the population because we are using a sample to make inferences about the population.
Sampling Theory
- The sample is at the crux of inferential statistics.
- Sampling theory is concerned with the probability that the sample is representative of the population.
- Of course, the selection of the sample is crucial to the representativeness of the sample.
- In research, when drawing samples, random selection is suggested:
  - Randomization assumes the error is distributed randomly among the groups.
- However, more often researchers tend to use non-random sampling procedures that influence the generalizability of results.

Sampling Distributions
- Inferential statistics uses probability and sampling theory to create sampling distributions.
- These distributions represent the population of all sample statistics and are used to represent "how likely it is that the sample is representative of the larger population."
- The mathematical theorem that supports this concept is called the central limit theorem.

Central Limit Theorem
- If the sample size is sufficiently large (n>30), the sampling distribution of sample means tends to be a normal distribution no matter what shape of the original "parent" population.
- Also, the central limit theorem states that the sampling distribution of smaller sample sizes will be a normal distribution if the parent population is a normal distribution.
- So what does all this mean?

Sampling Distribution & Inferential Statistics
- As the size of the sample increases, there is an increase in the likelihood that the sample data obtained are an accurate estimate of the true population value.
- Researchers use the sampling distribution to test their research hypotheses, by setting an alpha level or error rate.

Let’s test this on a bag of M&M’s.
Probability and Sampling Distributions

- What is the probability of obtaining the observed results if ONLY random error is operating?
  - Probability is the likelihood of the occurrence of some event or outcome
  - Statistical significance = the probability the difference in sample means is due to error

Probability and Sampling Distributions (con't)

- Probability required for significance is called the alpha level
  - While three common alpha levels exist, most researchers use the .05 level
  - This means the researcher has a 5% chance of saying there is a difference when there is not
  - If probability is low (.05 or less), we say there is a difference
  - If probability is high (over .05), we say there is no difference

Probability and Sampling Distributions (con't)

- Sampling distributions
  - Probability distribution = binomial distribution
  - All statistical significance decisions are based on probability distributions such as this one
    - Such distributions are called sampling distributions
Inferential Statistics

- Researchers rarely study entire populations.
- Used to determine what would happen if we conducted the experiment on subsequent occasions with multiple samples.
- Central question with using samples:
  - How do we know that the differences we observe are due to “real differences” or are due to error?
    - The difference between samples will NEVER be zero, even when all principles of experimental design are used.
    - The difference in sample means reflect any true difference in the population (effects of the IV) plus any random error.
- Inferential statistics allow for inferences to be made about the true difference in the population on the basis of sample data.

Statistical Inference

- Begins with a statement about the null hypothesis and the research hypothesis.
  - Null hypothesis – population means are equal, differences observed due to random error, the IV had no effect.
  - Research hypothesis – population means are not equal, differences observed are due to the effect of the IV.
  - The alpha level represents the mistake of rejecting the null hypothesis when it is true.
    - If our alpha level is .05, then we have a 5% chance that we are rejecting the null hypothesis when it is true.

Null and Research Hypotheses

- Null hypothesis
  - Ho: Population means are equal.
- Research hypothesis
  - H1: Population means are not equal.

Inferential Statistics

- Importance of ensuring that the groups are equivalent.
  - Achieved by experimental control and/or randomization.
  - Key aspects of experimental method:
    - If groups are equivalent, any differences in the dependent variable will be due to the manipulation of the independent variable.
- True score + random error
  - True score – what we expect the population score to be.
  - Random error will be responsible for some difference in the means.
  - The observed score = true score + random error.
Null and Research Hypotheses (con’t)

- The null hypotheses ……
  - Null hypothesis is a very precise statement
  - If we can determine that the null hypothesis is incorrect, then we can reject it and accept the research hypothesis
  - Null hypothesis is rejected when there is a low probability that the results could be due to random error = **statistical significance**

Type I and Type II Errors

- **Type 1 error** – reject the null hypothesis when the null is true
  - Saying there is a difference when there is really no difference
- **Type 2 error** – accept the null hypothesis when null is false
  - Saying there no difference when there really is a difference

Example: The $t$ and $F$ Tests

- Statistical tests allow us to use probability to decide whether or not to reject the null hypothesis
  - Examples: $t$-test and $F$-test (ANOVA)

- $t = \text{group difference (difference between the sample group means)}$ within-group variability (amount of random error in the sample)

  We would expect the group difference to be zero if the Null Hypothesis is true!

Example: The $t$ and $F$ Tests (con’t)

- **One-tailed versus two-tailed tests**
  - **One-tailed** = research hypothesis specified a direction of difference between the groups
  - **Two-tailed** = research hypothesis did not specify a predicted direction of difference
t-test

- t-test allow for the comparison among two sample means
- t-tests are based on two aspects
  - The level of measurement of the dependent variable, and
  - The type of independent variable(s)
- The combination of these two elements create the different t-test procedures
  - Parametric (interval or ratio level of measurement) versus non-parametric (ordinal level or measurement)
  - This course will only look at parametric t-tests
  - Independent groups versus repeated measures

**t-test (con’t)**

- Type of independent variable(s)
  - Independent groups:
    - Subjects are tested under only one level of the independent variable(s)
  - Repeated measures:
    - Subjects are tested under all levels (repeatedly) of the independent variable(s)
    - For example, any data collected repeatedly for a person (e.g. pre- and post-test).

**t-test: Basic Assumptions**

- Dependent variable must be at least interval level of measurement.
- Independent variable must be nominal or ordinal in nature and must classify subjects into separate groups
- The variances of the dependent variable scores must be similar across both levels of the independent variable (aka homogeneity of variance)
- Dependent variable scores should be essentially normally distributed
- The sample size in each of the samples should be similar
- Subjects should be randomly selected

**t-test: Interpreting Significance**

- Statistical significance
  - Goal of the test is to allow you to make a decision about whether your obtained results are reliable
  - Statistical significance level called alpha
    - Alpha is a predetermined value set by the researcher
    - The alpha you choose indicates how confident you are with the results
    - Three popular levels = .01, .05 and .10
      - A .05 significance level is most common and says that you are 95% sure of the reliability of the findings; however, there is a 5% chance you are wrong
Type I and Type II Errors Revisited

- **Type 1 error** – reject the null when it is true
  - Determined by choice of alpha level; lowering alpha will decrease the probability of Type 1 error
- **Type 2 error** – accept the null when it is false
  - Generally, research should be designed to minimize Type 2 errors; depends on type of research

**Statistical reasons for Type II errors:**
- Stringent alpha
- Small sample sizes that cannot detect differences
- Effect size is small

Interpreting Nonsignificant Results

- Research is designed to show that a relationship between variables does exist
- Results of a single study may be nonsignificant when a relationship does exist
  - Type II error
    1. Procedures used (smaller alpha level)
    2. Sample size
    3. Small effect size

The Importance of Replication

- Instead of applying a statistical test to determine whether the results would hold up again and again, researchers look at the results of studies that replicate previous investigations.
**F Tests (ANOVA)**

- **F test**
  - Analysis of variance (F test) is an extension of the t-test
  - t-test and F test identical when there is one IV with two levels
  - F statistic is a ratio of two types of variance
    - Systematic variance (between groups)
    - Error variance (within groups)

---

**Significance of Pearson r Correlation Coefficient**

- Is the relationship statistically significant?
  - Ho: $r = 0$ and H1: $r \neq 0$

---

**Computer Analysis of Data**

- SPSS
- SAS
- Minitab
- Siestas
- BMDP
- Excel

---

**Selecting the Appropriate Statistical Test**

- One IV – two groups only
  - Nominal scale data = chi square
  - Ordinal scale data
    - Independent groups = Mann-Whitney U test
    - Repeated measures = Wilcoxon's T or the sign test
  - Interval or ratio scale data
    - Independent groups = t-test for independent samples
    - Repeated measures = Correlated t-test
Selecting the Appropriate Statistical Test (con't)

- One IV – three groups or more
  - Nominal scale data = chi square
  - Ordinal scale data
    - Independent groups = Kruskal Wallace H test
    - Repeated measures = Friedman T test
  - Interval or ratio scale data
    - One-way analysis of variance for independent
    - groups or repeated measures

Selecting the Appropriate Statistical Test (con't)

- Two or more IV’s
  - Nominal scale data = chi square
  - Ordinal scale data = no appropriate test is available
  - Interval or ratio data = two-way analysis of variance

The End