**Unit 1**

* Understand how study design affects inferences that can be made

-Observational vs. Experimental design

-An observational study does not allow researcher control over the independent variable – the dependent variable is measured in a natural state

-An experimental study allows the researcher to control the independent variable

* Apply the difference between experiment and observational study to different types of inferences

-Causal inference: an inference that the independent variable affects the dependent variable

-Causal inference can only be drawn in an experimental study, not in an observational study

* Learn terms such as sample, population, statistic, parameter

-Population: the group of interest for a study

-Sample: a smaller subset of the population created through a selection mechanism for the purpose of the study

-Parameter: a numeric observation of a population

-Statistic: a numeric observation of a sample created during a study

* Understand the difference between random assignment and random sampling

-Random sampling: the mechanism by which individuals are selected from the population to be part of the study sample

-Random assignment: the mechanism by which members of the sample are partitioned into research groups within the study

* Understand the principles behind permutation tests

-Permutation tests randomize outcomes across the treatment group

**Unit 2**

* Use appropriate inferential tools for the independent sample and paired sample design settings

-Histograms and basic stats work well for this

* Understand the conceptual basis for confidence intervals and *p*-values in a one- and two-sample setting

-A way to measure how large or small a population parameter value could be, estimated using the sample mean

-Multiple “answers” will result from multiple sample means – CI’s indicate where the answers should lie for the answers to be statistically significant

-Alpha dictates the confidence interval – 95% CI means that 95% of the means should fall within 1.96 SD intervals of the population means (provided H0 is false). Increasing confidence widens the interval.

-Formula uses population standard deviation instead of sample, this reduces error

-"With 95% confidence, the true mean μ is between Lower CI and Upper CI, based on a sample n from this population."

-Four steps to setting test: correct significance level (alpha), directionality (one or two tailed), quality assumption (quantitative, reliable, normal distribution, outliers accounted for), randomly sampled

* Calculate *p*-values and confidence intervals using the *t*-distribution

-CI=X(bar)±tcrit(s/√n)

-p-value: the chance of obtaining a certain sample result if H0 is true (smaller disproves H0)

-t=(x(bar)−μ0)/(s/√n)

-p-value is the area past the t-value

Error types:

|  |  |  |
| --- | --- | --- |
|  | H0 is True | H0 is False |
| Reject H0 | Type 1 error | Correct |
| Accept H0 | Correct | Type 2 error |

**Unit 3**

* Use appropriate graphics and statistics to examine mathematical assumptions
* Understand the effect of violations of assumptions on the validity of *t*-distribution based tests

-Sample must be randomly drawn

-Observations must be independent

-Sample must be representative of the populations

-Observations must be normally distributed

-Two sample t-tests must have equal variances

* Learn and apply methods of correcting for assumption violations

-Log transform can correct skewed data

-Power increases as alpha increases

**Unit 4**

* Determine when analysis methods based on the t-tools are inappropriate for a data set

-Nonparametric methods- does not assume that data comes from normal, etc, population, only necessary assumption is that the samples are independent

-Wilcoxon Rank Sum test – places all m and n observations, sorts regardless of membership, assigns numbers 1:N(where N=m+n) where ties receive average rank, sums should be relatively equal

-Fully parametric: find all possible permutations, build vector of ranks from each permutation, determine CI’s and p-value, and calculate number of sum ranks above observed value and divide by permutations

-True p value is the relative frequency

-Permutation test – randomly assign observations to 2 treatments and compute differences between means; continue reshuffling until a k number of times, compute difference D between means for each treatment and p value

-Permutation tests can be used to test subsets of means, medians, mean ratios, variance ratios, or sums of items, and the null distribution is built from test statistic

* Understand the principles behind the rank-sum and signed-rank tests

-Kruskal-Wallis Test – assume that experimental units are selected from k populations, observations from one treatment will be generally larger than the other

-KW HA holds that at least one pair of samples will differ in means

-Signed rank test runs against paired data (ie the same treatment measured twice on the same population), rank sum runs against non-paired data

* Apply such tests when appropriate and interpret their results

-Wilcoxon Rank Sum: apply in non-normal distribution when samples are independent

**Unit 5**

* Examine model assumptions

ANOVA – analysis of variance; measures quality of means across >2 populations

Assumptions – all populations are normally distributed, all populations have equal variances, observations are independent (both within and among samples)

* Make statistical inferences for multiple samples and interpret the results in context

-Full model compares all means in context, reduced model compares a subset of means by altering the null hypothesis

-Random effect – actual values from observations are from a random sample themselves (ex – people with medical afflictions present certain parameters as a population but a researcher might be interested in a broader group)

-Inflates probability of type 1 and type 2 error if the model does not examine error variance for the full model

* Understand the extra sum of squares principle

-Extra sum of squares (numerator of linear F-stat) used to reduce sum squared error

* Use the extra sum of squares principle to obtain an analysis of variance table
* Apply the concept of full and reduced model to test different combinations of means

**Unit 6**

* Understand weakness of the alternative hypothesis for an ANOVA *F*-test

-ANOVA F-test can indicate that at least one of the population means will differ from the rest but does not indicate which one. Further tests are needed to indicate which.

* Formulate appropriate linear combinations of means

-Tests for all pairwise differences in means:

-Bonferroni: overall type 1 error rate divided by number of tests, adjust significance level and compare to subsequent p-values, simple but conservative – generates higher type 1 error

-Tukey’s HSD: uses studentized range stats(q = largest x(bar)-smallest x(bar)/sqrt(m\*s(w)\*(1/n))) to obtain simultaneous confidence intervals for each pair of population means, unequal sample sizes work well with this test if 2\*n(i)\*n(j)/n(i)+n(j) replaces n

-REGWQ: controls maximum experimentwise error rate under complete or partial H0, and has good power, arranges means in descending order, rejects homogeneity (H0) if difference of smaller and larger sample means is greater than or equal to the adjusted p value (gamma p) where p = difference of largest and smallest samples sizes plus 1

-Test for pairwise difference from control:

-Dunnett: compares treatment to control, computes D value which, if larger than the difference between treatment and control means, rejects H0, can be one or two sided

-Test for all possible differences: Scheffe

* Understand the multiple comparison problem and how to correct for it

-More tests can lead to higher chance of type 1 error

-Bonferroni correction is best method to keep familywise error rate lower than total alpha

* Use the formulation to make appropriate inferences
* -False causality – small p-value does not imply causality (only experimental design with small p-value is causal)

-Null hypothesis cannot be “accepted”, only fail to reject

**Key terms:**

-Beta: The probability of making a Type II error when the null hypothesis is false

-Power: The probability of rejecting the null hypothesis for two alts when the null hypothesis is false

-Alpha: The probability of rejecting the null hypothesis for two alts when the null hypothesis is true

-Z-score: The number of standard deviations away from the mean

-Central Limit Theorum: the mean of the sampling distribution of the sample average will be approximately the same as the mean of the original distribution from which the sample was pulled, the sampling distribution of the sample average will be approximately normally distributed (no matter what the distribution of the overall population is) – SD will differ