PSY 216 Exam 3 Review

Inferential statistics allow us to

- decide if sample results are representative of the population
- if a treatment probably had an effect

A point estimate is a sample statistic (e.g. $\overline{X}$, r, s) that is used to estimate the value of a population parameter (e.g. μ, ρ, σ)

A sampling distribution is a frequency distribution of sample statistics

- mean of sampling distribution equals the population mean
- standard deviation of sampling distribution equals the standard error of the mean
- will be normally distributed as long as the sample size is sufficiently large (thanks to the central limit theorem)

The standard error of the mean ($s_{\overline{X}}$)

- equals the sample standard deviation divided by the square root of the sample size ($s_{\overline{X}} = s_X / \sqrt{n}$)
- is the standard deviation of the sampling distribution
- tends to decrease as the sample size increases

A confidence interval (CI)

- is a range of numbers that likely contains the population parameter
- equals $\overline{X} \pm z \cdot s_{\overline{X}}$ where z is the z score that has one half the proportion of area not covered by the CI above it
- the CI either contains the population mean or it doesn't; 95% of 95% CIs will contain the population mean
- gets wider as the degree of confidence increases

Null (H_0) and alternative (H_1) hypotheses

- must be mutually exclusive: if H_0 is true, then H_1 cannot be true and if H_1 is true, then H_0 cannot be true
- must be exhaustive: for all possible outcomes, either H_0 is true, H_1 is true, or both are true
- can be non-directional or two-tailed; no particular ordering of the means is specified; hypothesis often has "different from" or "not equal to" in it; H_0: $\mu_1 = \mu_2$ H_1: $\mu_1 \neq \mu_2$
- can be directional or one-tailed; a particular ordering of the means is specified; hypothesis often has "greater", "better", "slower", etc. in it; H_0: $\mu_1 \leq \mu_2$ H_1: $\mu_1 > \mu_2$ or H_0: $\mu_1 \geq \mu_2$ H_1: $\mu_1 < \mu_2$
- are always written in terms of population parameters (e.g. μ, ρ, σ), never in terms of sample statistics (e.g. $\overline{X}$, r, s)
- null must contain an equal sign: =, ≤, or ≥
- alternative should correspond to the verbal hypothesis

A Type-I error

- occurs when you reject H_0 when H_0 really is true
- is also called an α error
- occurs with a probability specified by the α level (usually .05, but can be any value)

We reject H_0 when the probability of making a Type-I error is less than or equal to the pre-specified α level; if the probability of making a Type-I error is greater than the pre-specified α level, we fail to reject H_0; failing to reject H_0 does not imply that H_0 is true

A Type-II error

- occurs when you fail to reject H_0 when H_0 really is false
- is also called a β error
- occurs with a probability specified by the β level

For a given sample size, α and β are inversely related; as sample size increases, both can decrease

Single sample inferences are used in situations in which there is only one set of data; that data is compared to a pre-specified value

Steps in performing inferential statistics:

1. Write the null and alternative hypotheses; determine if they are one- or two-tailed
2. Specify the α level (usually .05)
3. Determine the appropriate statistic to use and calculate the observed (or calculated) value of that statistic
4. Determine the critical value from a table, or get the p value (significance value) from the output
5. Decide whether to reject H_0 or fail to reject H_0

z-score test

- is used with a one-sample data set in which the population mean and standard deviation are known and the sampling distribution is normally distributed (either because the sample size is sufficiently large or the population is normally distributed)
- $z = \dfrac{\overline{X} - \mu}{s_{\overline{X}}}$ $\quad s_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}}$
- critical value is gotten from a table of areas under the unit normal distribution; it corresponds to the z-score whose area above equals α / number of tails in the test
- reject H_0 if the observed or calculated z-score is larger than the critical z-score, or if the p value is $\leq \alpha$

Student's t-test (one-sample)

- is used with a one-sample data set in which the population standard deviation (and / or mean) are unknown
- estimates the population standard deviation (σ) based on the sample standard deviation (s) by using the formula $\hat{s} = \sqrt{\dfrac{n}{n-1}s^2}$
- t distributions are sampling distributions that account for the non-normality of small sample sizes; the particular t distribution used depends on the degrees of freedom
  - degrees of freedom (df) are the number of scores that are free to take on any value after certain restrictions have been placed on the data set
  - for a one-sample t-test, the df = n – 1 (where n is the sample size)
- $t = \dfrac{\overline{X} - \mu_0}{s_{\overline{X}}}$, $s_{\overline{X}} = \dfrac{\hat{s}}{\sqrt{n}}$, $\hat{s} = \sqrt{\dfrac{n}{n-1}s^2}$
- critical t-value is gotten from a table of critical t-values; find the row with the appropriate degrees of freedom and the column with the appropriate α levels and number of tails; the critical t is at the intersection of the row and column
- reject H_0 if the observed or calculated t-value is larger than the critical t-value, or if the p value is $\leq \alpha$

To test H_0: $\rho = 0$, use the usual procedure but calculate t with $t = \dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}}$, df = n – 2

Two-sample inferences are used when there are two sets of data

- one set corresponds to the control condition which did not receive the treatment
- one set corresponds to the experimental or treatment condition which did receive the treatment

Two-sample t-test

- is conceptually identical to one-sample t-test except
  - there are two sample means instead of a sample mean and a population mean
  - there are two estimates of the population standard deviation instead of one
- $t = \dfrac{\overline{X}_{treatment} - \overline{X}_{control}}{s_{\overline{X}_{exp\,erimental} - \overline{X}_{control}}}$, $s_{\overline{X}_{exp\,erimental} - \overline{X}_{control}} = \sqrt{s^2_{\overline{X}_{exp\,erimental}} + s^2_{\overline{X}_{control}}}$, df = $n_{experimental} + n_{control} - 2$ (note $n_{experimental}$ must equal $n_{control}$ for these equations)
- $s_{\overline{X}_{exp\,erimental} - \overline{X}_{control}}$ is called the standard error of the difference of the means; it is the standard deviation of the sampling distribution
- H0: $\mu_{experimental} = \mu_{control}$ $\quad H_1$: $\mu_{experimental} \neq \mu_{control}$ (two-tailed)
  H0: $\mu_{experimental} \leq \mu_{control}$ $\quad H_1$: $\mu_{experimental} > \mu_{control}$ (one-tailed)
  H0: $\mu_{experimental} \geq \mu_{control}$ $\quad H_1$: $\mu_{experimental} < \mu_{control}$ (one-tailed)
- critical t-value and decision are made in the same manner as a one-sample t-test

Between-subjects designs

- occur when different participants participate in each condition of the study
- should have sample standard deviations that are independent of each other

Within-subjects (repeated measures) designs

- occur when the same participants participate in every condition of the study
- should have correlated sample standard deviations
- uses $s_{\overline{X}_{experimental} - \overline{X}_{Control}} = \sqrt{s^2_{\overline{X}_{experimental}} + s^2_{\overline{X}_{control}} - 2 \cdot r \cdot s_{\overline{X}_{experimental}} \cdot s_{\overline{X}_{control}}}$ for the standard error of the difference of the means; this will usually result in a smaller standard error of the difference of the means than would occur in a between-subjects design; this increases the statistical power of the statistical test
    - statistical power is the ability to reject H_0 when H_0 is false; we want statistical power to be large
- uses df = $n_{\text{pairs of scores}} - 1$; this results in a smaller degrees of freedom than would occur in a between-subjects design; this decreases the statistical power of the statistical test
- usually within-subjects designs have greater statistical power than their between-subjects counterparts

Always use the appropriate statistical test with a given experimental design

Student's t (one-sample and two-sample varieties) makes several assumptions:

- assumptions should be checked; violating the assumption can make the t-test give invalid results
- homogeneity of variance assumption states that the variability in each group is approximately equal
    - Levene's test for homogeneity of variance tests H_0: $\sigma^2_{\text{experimental}} = \sigma^2_{\text{control}}$; typically α is set to .1 or .25 for Levene's test as we are attempting to accept H_0; if Levene's p value is $\leq \alpha$, then reject H_0 and the assumption of homogeneity of variance has probably been violated; otherwise the assumption has probably been satisfied
    - if the assumption is violated, there is a special type of t-test that can be used; this special t-test has lower statistical power
- normal distribution assumption – data in each condition should be normally distributed
    - often true for social science research
    - t-test is robust to non-extreme violations of this assumption
        - robust means the statistic will give valid results even if the assumption is violated
- independent samples assumption – you should not be able to predict the data in one sample from the data in the other sample (that is, the correlation coefficient of the two samples should be close to 0; this can be tested using the procedure discussed above for H_0: $\rho = 0$)
    - this assumption is automatically violated in within-subjects designs
    - t-test is robust to violations of this assumption

The exam covers

- material presented on 10/12 through and including 11/5
- chapters 11, 12, and 13 of the text
- SPSS for all types of t-tests (one-sample, independent samples t-test with single value groups, independent samples t-test with cut point groups, and paired samples t-tests)

Any formulae needed for the exam will be given to you; however, you should know how to interpret and use the formulae. You may be asked to perform a z-score test or a one-sample t-test by hand. You will not be asked to perform a two-sample t-test by hand.