Statistics

Jonathan A. Pedroza, PhD

Types of Statistics

- descriptive statistics are used to organize and summarize the data
  - can be very useful for giving an overview of your sample and to show any weird values about your sample
    - If you have a hook at the beginning of your research paper about the percentage of college students with depression (Ex: 21%) but then you look at the descriptive statistics of your sample and note that 46% of your sample meets depression criteria
- inferential statistics is the examination of relationships between two or more variables to see if relationships exist
  - these are used for generalizing findings from your sample to your population of interest
    - Ex: examining SONA Psychology participants to generalize to all Psychology students at CPP

Examples

- average age of CPP students
- getting a good night's rest is associated with better grades
- the percentage of students that are male
- the comparison between males and females in age

Terminology

- statistics are values that describe a sample
  - The average amount of students that failed the exam.
- parameters are values that describe a population
  - The percentage of students that failed PSY 3307 on their first attempt.

Descriptive Statistics

- frequency distribution are the counts for each participant for a chosen variable
  - Can be useful for nominal, ordinal, and numeric data
    - percentage of males and females
    - percentage of students that chose each value for your Likert-scale questions
    - number of students that are a specific age
- frequency distribution tables
  - SPSS gives you a table for each variable in the $Frequencies$ tab
  - can show you quick descriptive statistics for categorical data
- frequency distribution graphs
  - there are designated functions to provide either bar graphs or polygons
  - I'll show you how to read it on a histogram because SPSS likes to make it difficult to create multiple types of visuals at once

Descriptive Statistics

- histograms

- is a bar graph for continuous data
- it shows each score on a continuum and will group together to give you a general examination of your variable
- you will have one of these for your outcome to see whether or not your data is normally distributed
  - it should follow a bell-shaped curve

Descriptive Statistics

- polygons are useless don't use them
  - it shows a point with a line that connects the counts for each score
    - useless because it makes your eye follow the line like a trend
    - it is not a trend
- both histograms and polygons show frequencies of continuous and categorical data
  - for this purpose, I will only teach about histograms for your categorical data, although not 100% appropriate

Descriptive Statistics

- bar graphs are useful for categorical data because it is easier to see which categories have more participants (higher counts)
  - often stated that it is different from a histogram because there are spaces between the bars rather than all squished together
  - histogram = squished
  - bar graph = separated
- NO pie charts

Describing Interval and Ratio Data

- central tendency measures are measures that show the center of a distribution
  - there are three measures of central tendency
    - mean is the average value –> add up all scores for participants and then divide by the number of participants
    - JP: Can get a composite average score –> add up all items in a measure to then divide by the number of items to get an average score for a construct –> average depression value
    - median is the 50% percentile or the absolute middle of the distribution with half below this point and half above
    - useful for skewed data
    - mode is useless, it simply tells you what is the most frequent value in your distribution

Describing Interval and Ratio Data

- measures of dispersion show how much scores vary from the measure of central tendency use
  - will often be how much scores vary from the mean
  - standard deviation is the standardized way of stating how much scores vary from the mean
    - Ex: M = 10, SD = 1.5 means that at 1 standard deviation away, participants range from 8.5 to 11.5
    - standardized because scores are on the scale of whatever variable you are looking at
    - for age, a participant that is 1 standard deviation away will be 11.5 years old
  - variance is the same thing as standard deviation but in squared units
    - this value is the standard deviation squared
    - for age, a participant that is 11.5 years old, is 2.25 squared years older than the average participant

Describing Interval and Ratio Data

- degrees of freedom = N - 1
  - used to estimate a population from your sample by making a conservative estimation of your population
  - without this, you would overestimate how much variation is in your sample from the population
    - making it a biased sample estimate

Describing Nominal and Ordinal Data

- frequencies –> percentages
- get a frequency table and then report the percentages of each of those in your results

Using Visuals

- comparison between groups
  - we will cover this in SPSS
  - we're looking for a visual that provides an average score of your outcome for each of your groups
- relationships between IV and DV
  - with correlational data, you can show the relationship between two variables with a scatterplot
  - we won't cover this much for this class though
- line graphs
  - will be the most useful for showing your interaction
  - this will show the averages for treatment/control variable on the x-axis and then the individual colors will show your second variable
    - if those lines cross then you have yourself an interaction

Correlation

- a correlation is the examination of two scores/values to see what the strength of your relationship is
  - often used for preliminary data
- scatterplots show the relationship between two continuous variables
- Pearson correlation is a correlation that shows the linear relationship of two continuous variables
- Spearman correlation is the correlation of ordinal data
  - can also be used to examine the variables that are on a Likert-type scale
- lowest possible correlation $= 0$; strongest correlation $= 1$
  - can be in both directions

Inferential Statistics

- sampling error is the naturally occurring difference between your sample statistic and the population parameter
  - reductions in sampling error can be done by making sure your sample is representative of your population
- hypothesis tests are statistical procedures to use sample data to attempt to state if there is a relationship in your sample that represents the population
  - or if your data shows that your sample is no different from the population
- null hypothesis is when you state that there
  - are no differences between groups
  - no relationship between variables
  - no difference between your sample and your population

Inferential Statistics

- standard error is the "standard deviation" of your population
  - the distance between your sample and the population on a sampling distribution
- test statistic the summary value that measures what your analyses/model can account for and what is unknown/noise/error
- alpha level or level of statistical significance is the probability that your result was obtained by chance or if there is a true difference between your sample and your population
  - $p < .05$
  - less than 5% chance that your finding was due to chance

Errors in Hypothesis Testing

- Type I Errors are when you run too many tests and you end up create an opportunity where your finding is simply due to chance
  - false positives
  - if you run 20 tests and find 1 statistically significant finding, that is due to chance

$$\frac{1}{20} = .05 \; or \; 5\%$$

- Type II Errors are when you don't detect a real significant finding even if there really is one
  - usually the result of not having enough participants
  - false negatives

Factors That Influence Outcome Of Hypothesis Test

- number of scores in a sample
  - more participants will give you a better estimation in your sample
  - JP: If you are worried about your sample size, I can create some fake data for you
- size of the variance
  - if there is not enough variation in your scores, you will likely not find a statistically significant finding

Supplementing Hypothesis Tests with Effect Sizes

- Cohen's d
  - .2 = small effect size
  - .5 = medium/moderate effect size
  - .8 = large effect size
- percentage of variance
  - R^2 and η^2
    - .01 = small effect size
    - .09 = medium/moderate effect size
    - .25 = large effect size
- confidence intervals

Data Structures

- one variable for your sample of participants
  - descriptive statistics
- two variables that you are looking at a relationship between for your participants
  - correlation/regression
- two or more groups/conditions that you comparing
  - types of ANOVA

Scales of Measurement

- ratio
- interval
- ordinal
- nominal

Scales of Measurement

- ratio
  - true zero, deals with numeric values
    - Ex: Weight
- interval
  - no true zero, deals with numeric values
    - Ex: money in a bank account
    - Ex: temperature
- ordinal
  - ordered/ranked categories
    - Ex: places in a race
- nominal
  - categories
    - Ex: race/ethncity, sex/gender, etc.

Descriptive Statistics

- want to get general knowledge about your variables
- can get descriptive statistics of all of your variables
  - all the variables that make your outcome
  - if a scale score then get descriptives for each variable that makes it up that construct
  - if using a quiz/test then you'd want to see how many participants get every question correct
- shows how much of your data is missing for each question

Correlational Analyses

- if you have a second construct to compare to your outcome, you could use a correlation
- we will not be using correlations/regressions for your experiments

Group Comparisons

- can be used to see differences between your groups (control and experimental) or conditions (pre- and posttest)
- between-subjects designs
  - comparing 2+ groups on your outcome
  - factorial design
    - comparing 2+ groups on your outcome by a second variable comparing 2+ groups
- within-subjects design
  - comparing 2+ time points/conditions on your outcome
  - factorial design - mixed design
    - comparing 2+ time points/conditions on your outcome with the inclusion of between-subjects variable comparing 2+ groups

Group Comparisons

- one-way ANOVA
  - comparing groups/levels of one IV/factor and seeing if there are differences in your outcome

two-way ANOVA

- comparing groups/levels of one IV/factor and seeing if there are differences in your outcome (main effect 1)
- comparing groups/levels of one IV/factor and seeing if there are differences in your outcome (main effect 2)
- comparing groups/levels of one IV/factor based specific group/level of second IV/factor to see if there are differences in your outcome (interaction)

Special Statistics

- Cronbach's alpha
  - is used to assess the internal reliability or also called the internal consistency
  - calculates the items' variation and combines them for the sample to see whether items show a consistent pattern
  - easiest way to assess whether your measure is reliable
- Test-retest reliability
  - those that are using within-subjects designs can test whether the scores differed from pretest to posttest
  - use the composite score (average) to see if your measure was reliable when comparing the first time to the second time
    - more than 2 time points would require intraclass correlations, but we're not going to need that for you all