Stats with jamovi

Dana Linnell

Invalid Date

Table of contents
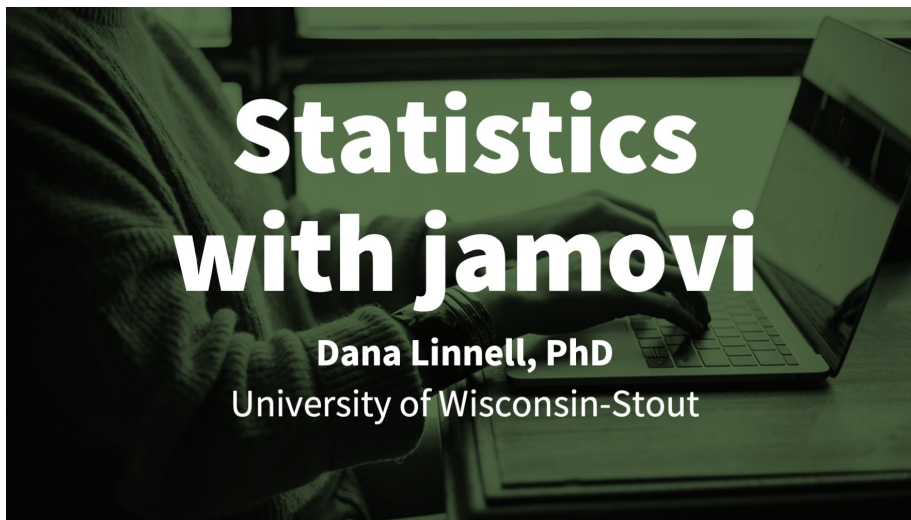
Chapter 1

Statistics with jamovi

Welcome

This is the website for PSYC 290 and PSYC 790 at the University of Wisconsin-Stout, taught by Dr. Dana Linnell. These resources are aimed at teaching you how to use jamovi and null hypothesis significance testing (NHST) to answer research questions.

This website is **free to use** and is licensed under a Creative Commons BY-SA (CC BY-SA) license version 4.0. This means you are *free to share* (i.e., copy and redistribute the material in any medium or format) and *free to adapt* (i.e., remix, transform, and build upon the material for any purpose, even commercially), provided that you *attribute* these resources by citing me, indicating if changes were made and you *share alike* (i.e., if you adapt, you must distribute your contributions under the same license as the original).

Many of the data examples come from "Learning statistics with jamovi: A tutorial for psychology students and other beginners" by Danielle J. Navarro and David R. Foxcroft, version 0.70.

Dedication: This book is dedicated to my graduate statistics professor Dr. Dale

Berger, who gave us a similar set of resources when he taught statistics at Claremont Graduate University. I still have my binder of handouts, homework assignments, and notes, which have been instrumental throughout my career. Thank you for showing me the joy of statistics!

Figure 1.1: Image of Dale Berger and Dana Linnell at her master's graduation ceremony

Part I

Overview

(PART) Overview

1. Introduction

This chapter will introduce you to the course, the instructor (Dr. Dana Linnell), and the textbook.

1.1 Getting help in this class

Come to student hours regularly! They are often available both in-person or online, depending on whether you are enrolled in an in-person or online section of the course. I am *always* available to help you. Depending on the semester, there may also be a graduate assistant (GA) who can support you as well. We will be scheduling regularly recurring student hours each week so you can come ask questions, get help on your homework, or just have a space to come together to work on your assignments in a dedicated online space.

If you have questions, you can message me on Teams or email me at linnellda@uwstout.edu. See your course on Canvas if you have any questions about Teams.

1.2 Dana, your instructor

My name is Dana Linnell (pronounced DAY-nuh lih-NELL) and I started teaching at UW-Stout in Fall 2019. I teach statistics, research methods, interpersonal effectiveness, and evaluation in the Department of Psychology. I *love* statistics! It is one way we can answer our research questions and test our hypotheses.
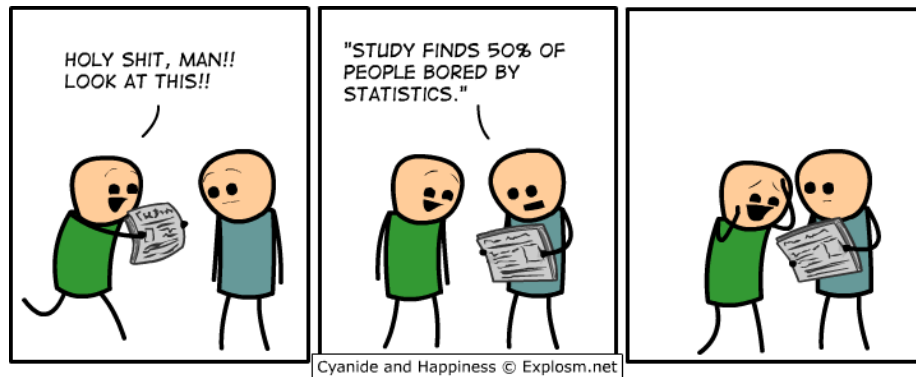
However, I know not everyone likes statistics. Some of you may not care much about it, and some of you may be apprehensive or anxious about taking this course. Please know that **I am here for you and I want to make this class an enjoyable learning experience**. If there is anything I can do to help make this class more enjoyable and to help you learn, please reach out to me.

1.3 Navigating this website/book

> **i** Note
>
> This book was developed using a coding language called R in a graphical user interface called RStudio using a package called bookdown. This book is hosted on a platform called GitHub. You can see the code for this book here if you are interested.

There are some icons at the top of this book that you may find useful:

1. The first icon of the toolbar, which looks like three horizontal lines, toggles the visibility of the sidebar, which contains the table of contents. You can also hit the `S` key on your keyboard to toggle the sidebar.

2. The second icon of the toolbar, which looks like a magnifying glass, is the search feature, which you can use to search the entire book. You can also hit the `F` (Find) key on your keyboard.

3. The third icon, which is a capital A, is for font/theme settings, which you can use to change font size (smaller or bigger), font family (serif or sans serif), and theme (white, sepia, or night).

4. The fourth icon, which is a lower-case i, provides information on the keyboard shortcuts.

1.4 Technology Tips

To succeed in this course, you will need a computer capable of installing and running Microsoft Word, Microsoft Excel, and jamovi. The Stout Surplus store (Facebook; website) may have laptops and desktops they are selling for cheap. You may also find it useful to have a mouse in this course since we'll be pointing and clicking a lot in jamovi; you can get small portable ones to bring with you to campus if you're in my in-person section.

You may also find it beneficial to have two monitors in this course. When doing activities and assignments in this course, I recommend having jamovi open on one monitor and then the assignment and textbook open on the other monitor. The Stout Surplus store may have extra monitors they are selling for cheap.

If you are unable to get a second monitor in this course, learn how to snap windows side by side in Windows or Mac. Learn also how to use the Alt+Tab function on your computer as well to switch back and forth from open applications. For Macs, use Cmd+Tab.

When you are working on this course, I recommend that you only have open the application and windows you actually need for this course. Ample research has shown that multitasking does not work, and so when you switch from one activity to another (e.g., from working on another course to this one) you should wrap up your tasks, save your files, close them all, and then start fresh. This will clean up your taskbar and help you focus on this course. It will also free up your computer's resources so it won't start slowing down.

Another friendly reminder: restart your computer frequently, at least once a week. Your computer will begin slowing down the longer you wait in between restarts. Also, be sure to update your computer regularly.

Note that jamovi also has a Cloud version. The free version will not be sufficient for all this course, but you can pay for the cloud version if you wish. Please note that the desktop version is strongly recommended as it is free and should work fine with the minimum technical specifications of computers that Stout requires.

1.5 Errors, mistakes, and suggestions

I am human, therefore I err. If you find an error in the textbook or something you think might be a mistake, please let me know ASAP so I can update this for everyone else. Let me know which section you find the error or mistake in and what the error or mistake is. For example, if there was an error here you could say, "There was an error in chapter 1, last section, that the first sentence should really be 'To err is human (Alexander Pope, 1711).'"

In addition, if you have ideas to help make this textbook even better, please let me know. I would love to make this a useful resource to you both during our course and in your future research. Help me in making that a reality!

Chapter 2

———

———————————

Chapter 3

2. Statistics foundations

You have learned about both quantitative and qualitative methods. We will be focusing primarily on quantitative methods in this class and in this textbook. By quantitative methods, I mean methods that predominantly collect data that deals with numbers. We can then analyze that data using statistical procedures, which we will shorthand to "statistics." Understanding what we mean by statistics is the purpose of this chapter.

3.1   2.1 Describing our data

First, let's understand some basic statistics related to how we describe our data, including measures of central tendency (averages), measures of dispersion (spread), and measures of shape of the distribution (particularly a normal distribution). Here's a video walking you through what we learn in this chapter.

3.1.1   Measures of Central Tendency

There are multiple **measures of central tendency** (these are *all* averages so you must be careful when you say that word to explain which type you mean!):

- **Mean**: the sum of all points divided by the total number of points. The mean can be susceptible to outliers.

- **Median**: the middlemost value. The median less susceptible to outliers and best used when the data is skewed.

- **Mode**: the most frequently occurring score. Our data can also be **multimodal** when there are multiple modes or **bimodal** when there are two modes.

Note that depending on the shape of the distribution, the mean, median, and mode may not be the same value. If we have a **normal distribution** then they

will be the exact same! However, if we have a positively skewed distribution, the mean and median will be pulled towards the positively skewed data, as shown in this figure by Peter Prevos.
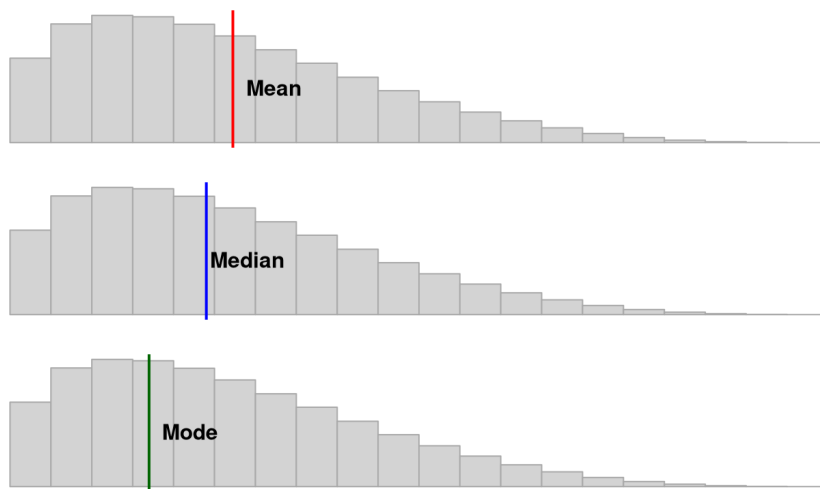
3.1.2   Measures of Dispersion

There are also multiple **measures of dispersion** that describe the spread of our data:

- **Range**: the difference between the maximum and minimum value (e.g., if the minimum score is 17 and the maximum is 49, then the range is 32)

- **Quartile**: when a dataset is divided into four equal parts, the first quartile (Q1) is at the 25th percentile, the second quartile (Q2) is at the 50th percentile, and the third quartile (Q3) is at the 75th percentile. Note that the median is also the 50th percentile!

  – **Interquartile range**: the middle 50% (Q1 to Q3)
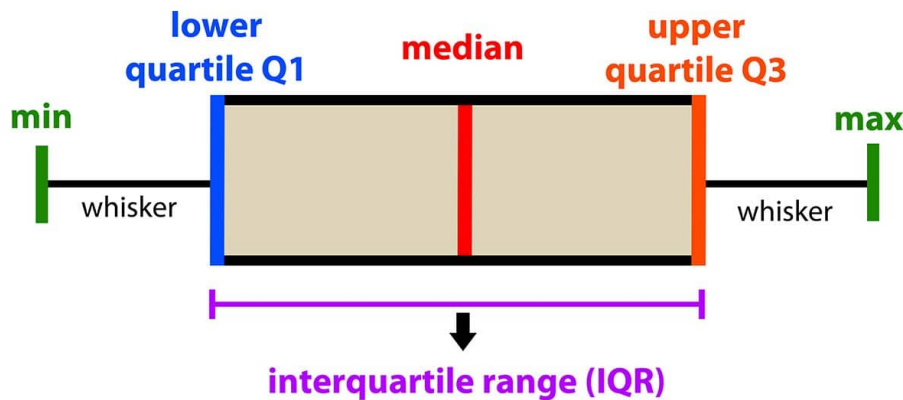
introduction to data analysis: Box Plot

Figure 3.1: Simply Psychology

- **Variance**: is the measurement of the spread between numbers in a dataset. More specifically, it is the sum of the squared deviations from the mean. This means first (a) calculating the mean, (b) subtracting each score from the mean (aka deviations from the mean), (c) squaring each of those deviations values, and (d) summing all those squared deviations. This is represented by the equation $\frac{\sum(X-\mu)^2}{N}$

- **Standard deviation**: is the square of the variance. This is represented by the equation $\sqrt{\frac{\sum(X-\mu)^2}{N}}$ if we are examining the whole population. If we only have a sample, we replace the denominator `N` with `N-1`.

> **ℹ Note**
>
> I sometimes present equations, like above, to help some folks better understand various things in this course. However, I very rarely expect you to calculate things, and when I do, I give you a lot of support in how to do so.

3.1.3 Measures of the Shape of the Distribution

We also have two main **measures of shape** that describe the shape of the distribution of our data:

- **Skew**: in a non-normal distribution, it is when one tail of the distribution is longer than another. Present in asymmetric distributions
    - **Negative skew**: when the tail points to the negative end of the spectrum; in other words, most of the values are on the right side of

the distribution

    – **Positive skew**: when the tail points to the positive end of the spectrum; in other words, most of the values are on the left side of the distribution
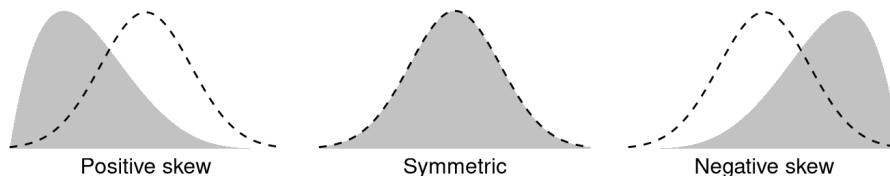

Figure 3.2: Figure by Peter Prevos.

- **Kurtosis**: the weight of the tails relative to a normal distribution. In other words, it's how flat or skinny the distribution is. There are some fancy terms related to kurtosis that you may hear about, but honestly I don't hear them used very frequently by researchers.

    – **Leptokurtic**: light tails; values are more concentrated around the mean. Think of kangoroos which can be noted for "lepping".
    – **Platykurtic**: heavy tails; values are less concentrated around the mean. Think of a platypus!

Student (a pseudonym of William Gosset) wrote this in his 1927 paper).

* In case any of my readers may be unfamiliar with the term "kurtosis" we may define meso-kurtic as "having β_2 equal to 3," while platykurtic curves have $\beta_2 < 3$ and leptokurtic > 3. The important property which follows from this is that platykurtic curves have shorter "tails" than the

normal curve of error and leptokurtic longer "tails." I myself bear in mind the meaning of the words by the above *memoria technica*, where the first figure represents platypus, and the second kangaroos, noted for "lepping," though, perhaps, with equal reason they should be hares!

When skew and kurtosis are zero, then we have a special type of distribution called the **normal distribution** in which the data are symmetrical on both sides of the mean. We sometimes call this a "bell curve". Scribbr has a great article on the normal distribution I encourage you to read!

Standard normal distribution

Mode
Median
Mean
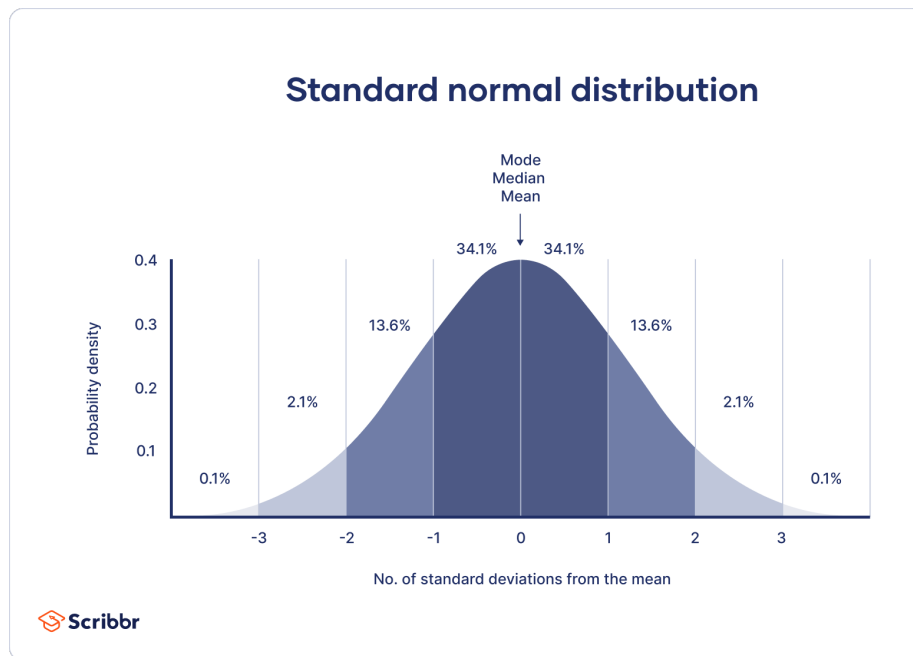
Although there are many ways we can visualize the distribution the two most common ways we will visualize the distribution using jamovi are with a **histogram** which could also be a density curve or a **boxplot**. We'll learn more about visualizing in Chapter 3 "Visualizing data".

3.2   2.2 Levels of measurement

This may or may not be refresher material for you, but it is extremely important you are familiar with the four levels of measurement.

Levels of Measurement

Categorical variables   Continuous variables

Binary, dummy-coded, or dichotomous variables are a special kind of nominal variable

Ratio

Named + Ordered + Proportionate intervals between levels + **Can accommodate Absolute Zero**

Interval

Named + Ordered + **Proportionate intervals between levels**

Ordinal

Named + **Ordered levels**

Nominal

Named levels

Analyses	Modes, frequencies, counts, percentages…	+ median, min/max, range, percentiles…	+ mean, variance, standard deviation.	
Examples	Occupation Political preference Favorite animal Place of residence	Likert scale items Rank order items (e.g., 1st, 2nd, 3rd)	Celsius/Fahrenheit Means of Likert scales Time (e.g., 1pm, 2pm)	Kelvin temperature Height, Weight Time duration (e.g., 20 mins, 1 hour)

Categorical: variables that have *categories* to the levels, but cannot be analyzed with a mean because the levels are not proportionate. There are two types of categorical variables:

- **Nominal**: a categorical variable in which each level of the variable is named but there is no order to them (e.g., breeds of dogs). Nominal variables can only be analyzed with frequencies, modes, counts, or percentages. Nominal variables with only two levels (generally coded as 0 or 1, although jamovi can handle named levels), we call these special nominal variables as either binary, dummy-coded, or dichotomous nominal variables. For example, binary variables could be coded as yes or no or as absence or presence.

- **Ordinal**: a categorical variable in which each level of the variable is named and there is an order to them (e.g., ranks). We can analyze ordinal variables with frequencies and percentages, like nominal variables, but we can *also* analyze them using medians, minimum and maximum values, ranges, and percentiles.

Continuous: variables with proportionate intervals between the levels, meaning they can be analyzed with a mean, SD, variance. There are two types of continuous variables (although for the purpose of this course we will simply call them continuous variables):

- **Interval**: a continuous variable that has intervals that are directly proportionate (e.g., the distance between 2-3 is the same as the distance between 5-6). We can analyze interval variables using means, variances, and standard deviations.

- **Ratio**: a continuous variable like an internal variable but can accommodate an absolute zero, meaning a zero is actually possible (e.g., weight,

temperature in Kelvin, reaction time). We can also analyze ratio variables with means, variances, and standard deviations, but we can also conduct arithmetic operations like fractions and ratios.

Note that in this class, I won't ask you to differentiate between interval and ratio levels of measurement because the differences between the two are usually not meaningful in our work. You'll even find that jamovi doesn't differentiate between the two, either. Instead, you just need to know whether the variables are continuous, ordinal, or nominal.

3.2.1 Examples of levels of measurement

Confused still on the levels of measurement? Maybe this will help!

One semester, a student asked, "Isn't time a continuous variable?" To which I responded, "It depends on how it is measured!" If time was measured in a simple pre/post design, such as the start of the semester and the end of the semester, then it's a nominal variable, and a specific type of nominal variable that we can call binary or dichotomous. If time was measured by month, January through December, then it would be ordinal because January is before (earlier than) February, for example. There is an order to the months, and a calendar that went March, October, April, January, etc. would make no sense. If time was measured as a continuous variable, it could be something like the exact day, exact time, response time, or time remaining.

Here's another example. Notice that studying can be measured at different levels. Depending on the nature of the question and response options, it might be nominal, ordinal, or continuous! Here's an example of data at the continuous, ordinal, and nominal level.

Name	Study_Continuous	Study_Ordinal	Study_Nominal
Name (Character)	*Hours studied per day*	*Likert scale of amount of studying*	*Whether or not they study every day*
Jesus	5.0	A great deal	Yes
Nicky	4.5	A great deal	Yes
Bradford	3.2	A moderate amount	Yes
Sylvia	1.7	A small amount	Yes
Martha	0.2	Rarely	Yes
Lillian	0.0	Never	No
Trayvon	0.0	Never	No

We can make any continuous variable into an ordinal and nominal variable and any ordinal variable into a nominal variable. But if we have a nominal variable we cannot make it ordinal, nor can we make an ordinal variable continuous. In other words, continuous variables *contain more information* than categorical

variables. Often, we want to avoid losing information and so we should aim to keep the variable at the highest level of measurement.

Because continuous variables have more information, we want to avoid doing things like mean or median splits or "collapsing" categories when you can. A mean or median split, which involves finding the mean or median value and splitting the data so it's above the mean/median or below the mean/median, is making a continuous variable as nominal, which is removing information. Collapsing categories may also further reduce the information.

Another thing to keep in mind is that just because we put numbers to something does not necessarily make it continuous! Be careful and think critically. If I said cat = 1, dog = 2, and frog = 3 it doesn't make it an ordinal or continuous variable.


Figure 3.3: XKCD: Assinging Numbers

3.3   2.3 Descriptive vs inferential statistics

There are basically two different types of statistics:

1. **Descriptive statistics** are used to summarize, organize, and overall *describe* our sample data. Typically, we do so using measures of central tendency (e.g., mean, median, mode), measures of dispersion (e.g., range, standard deviation, variance), and shape (e.g., skew, kurtosis). We may also visualize the data using tables or graphs.

2. **Inferential statistics** are what we use when we collect data about a sample and see how well that sample *infers* things about the population from which the sample comes from. Typically, we do so with statistical tests like the t-test, ANOVA, correlation, chi-square, regression, and more.

We can visualize the relationship between the population, sample, descriptive statistics, and inferential statistics (see figure below). We are typically interested in a **population** of interest but may not be able to collect data from the entire population because of budget, time, access, or other constraints. Therefore, we typically **sample** from the population; ideally, we do so randomly, but there are other types of sampling methods available that are covered in research methods. We then use **descriptive statistics** to describe our sample data, and we and **inferential statistics** to make generalizations about the population from which they were selected.
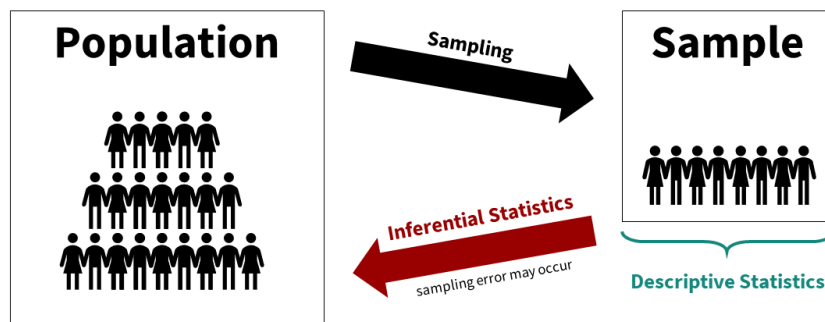
Note that we typically use both descriptive and inferential statistics. However, some studies may be purely descriptive with no inference back to the population. We typically always describe the sample when we are performing inferential statistics though.

3.3.1   An example

This has been pretty abstract so far. Let's go through a fairly simple research study to walk through all of this.

Imagine we're conducting an experimental study examining whether watching Schitt's Creek–a very good show–versus watching video lessons on studying techniques–useful, but boring–improved test performance in UW-Stout students.

Our population of interest is therefore all UW-Stout students, roughly 9,500 students total. We cannot include them all in our study; it wouldn't be feasible for us to collect all that data and probably not possible to get the university to get on board with the study of the entire student body. Therefore, we smartly decide to only collect data from a sample of the student body.

Who might our sample be? Ideally, we'd gather a random sample of the 9,500 students. However, to do that we'd likely need to still get university approval and get a list of a portion of student emails for recruitment purposes (over-sampling because our response rate is unlikely to be 100%). I just want to do this study to show what descriptive and inferential statistics are, so I just use students in my two sections of introduction to psychology classes (around 80 students total) as my population. This is definitely not a random sample, but a fine study for our illustrative purposes.

We conduct our study–let's assume we're fabulous researchers and it worked out perfectly. We randomly assign half our students to watch Schitt's Creek as part of their studying, and the other half watch video lessons on studying techniques. They have an exam a week later and we measure their accuracy on that exam. We then want to know: which group performed better on the exam?

First, let's describe the sample. We would likely visualize our results, perhaps as a histogram of all test scores, maybe separated by which group they were in. This would help us look at whether our data is normally distributed (more on this in a subsequent chapter on assumption checking). We would get the descriptive statistics: probably the mean, maybe the median if our data is skewed, the standard deviation and variance, and the range. If we wrote up our results and didn't share a visualization, this information would give a good sense of our data to our readers.

But what we really want to know is: which group performed better on the test? For that, we need our mean, standard deviations, and sample sizes for both groups. We then plug the numbers into the equation for this particular inferential statistic (in this case, an independent t-test, but we'll learn about that later) or–even better–we perform the statistic in our statistical software (jamovi). It spits out our statistical value and our p-value and we can then infer what the results mean for our population and answers our research question.[1]

---

[1]You might be wondering: well, what were the results? Which group performed better? As much as I love Schitt's Creek, most students don't know how to study well, and so the students who watched the video lessons on studying techniques far outperformed the students who watched Schitt's Creek.

Interested in better techniques for studying? Check out The Learning Scientists. This article does a good job of summarizing the research on effective study practices.

Chapter 4

2.4 Design & Methods Key Terms

The following covers some terminology that will be helpful to keep in mind throughout the semester.

4.0.1 Variables

We tend to talk about two different types of variables in our studies:

1. **Independent variable** (IV; also known as the predictor variable or exposure variable): this is the variable that is thought to be the cause of some effect. In experimental research, it is the variable that is manipulated.
2. **Dependent variable** (DV; also known as the outcome variable): this is the variable that is thought to be affected by changes in the IV. In research, it is the variable that is measured.

There are other types of variables we may be interested in. Note that these variables are not considered independent variables because we would not manipulate them or try to examine how they cause the dependent variable.

- **Confounding variable**: a variable that affects or is related to both the independent and dependent variable
- **Covariate**: a variable that only affects or is only related to the dependent variable
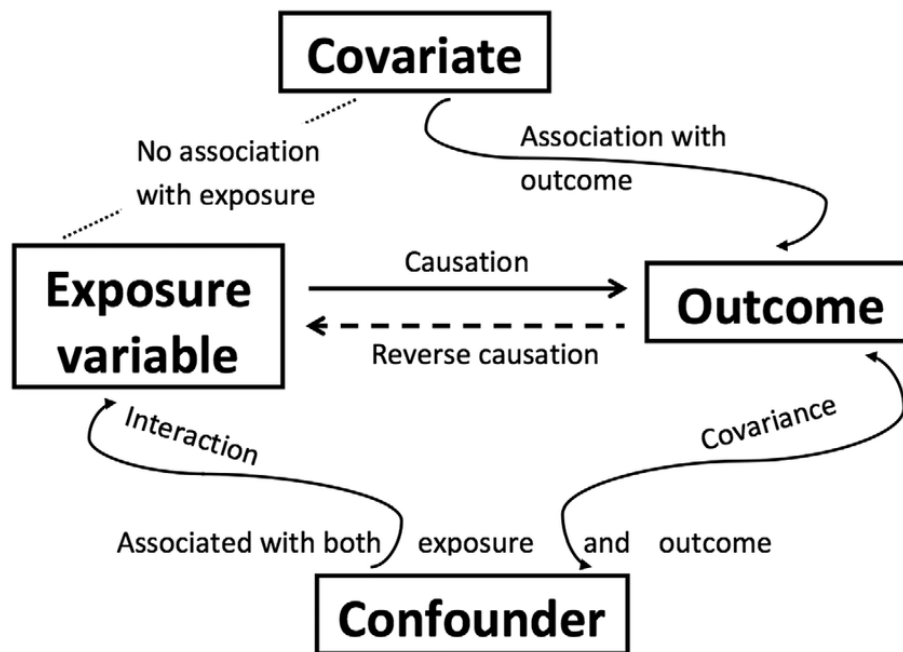
Covariate

No association with exposure

Association with outcome

Exposure variable

Causation

Reverse causation

Outcome

Interaction

Covariance

Associated with both exposure and outcome

Confounder

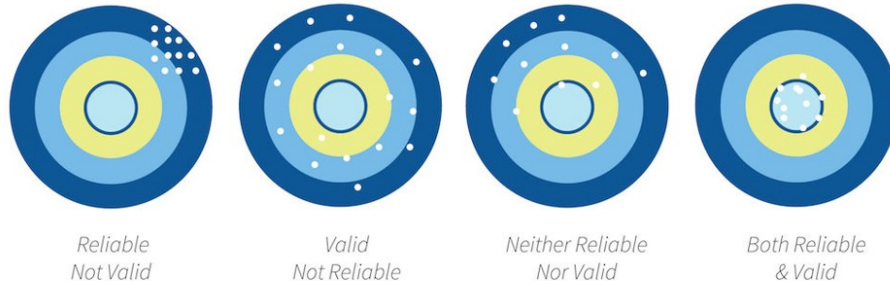Figure 4.1: Wilson & Wilson (2016)

4.0.2 Study design terms

Some terms you should be familiar with:

- **Between-group/subject design**: in a study, participants are only exposed to a single condition. In other words, there are different people in each condition.

- **Correlational research**: a study in which causality cannot be claimed; correlation does not infer causation! It is, however, one of three necessary conditions to infer causality. It is a *necessary* but *insufficient alone* condition.

- **Cross-sectional research**: also called non-experimental research; the IV is not manipulated and there is no random assignment. Furthermore, data is only collected at one time point (as opposed to longitudinal research)

- **Experimental research**: the IV is manipulated and there is random assignment

- **Falsification**: A key way we separate science from pseudo-science is that we attempt to *falsify* our hypotheses as opposed to *verify* our hypotheses. Null hypothesis significance testing (NHST) is about falsifying the null hypothesis; we can never truly verify our alternative hypothesis.

- **Hypothesis**: What we think the answer to our research question is (often our alternative hypothesis). The alternative and null hypotheses must be mutually exclusive (a result can't satisfy both) and exhaustive (all possible results are specified)

  – **Alternative hypothesis**: Often that the IV had *an* effect on the DV; can be specified as a two-tailed (an effect) or one-tailed (greater/less than) hypothesis

  – **Null hypothesis**: The null hypothesis is the opposite of the alternative hypothesis and encompasses all the rest of the possibilities. A special type of null hypothesis is the *nill* hypothesis which is that there is *no* effect on the DV.

- **Qualitative methods**: Broadly, methods that focus on words and meaning (e.g., interviews)

- **Quantitative methods**: Broadly, methods that focus on numbers and statistics (e.g., Likert scales)

- **Quasi-experimental research**: the IV is manipulated but there is no random assignment

- **Randomization**: participants are randomly assigned to conditions

- **Repeated-measures design**: participants are repeatedly measured on the dependent variable, either across conditions or across time

- **Theory**: A description of a behavior that makes predictions about future behaviors

- **Variation**:

  – **Systematic**: researcher introduces error systematically into the study, especially into one condition over another. For example, by randomly assigning participants into one of two conditions, we are introducing systematic variability between participants. However, it could be unintentional systematic variation; for example, perhaps we have two researchers collecting data and one is mean and the other is nice, and so participants respond differently depending on which researcher collects data from them.
  – **Unsystematic**: random variation

- **Within-group/subject design**: in a study, participants are exposed to every condition. In other words, there are the same people in each condition.

4.0.3   Reliability and validity

Reliable   Valid   Neither Reliable   Both Reliable
Not Valid   Not Reliable   Nor Valid   & Valid

- **Reliability**: the consistency of a measure by time (test-retest reliability), across items (internal consistency) or across different researchers (inter-rater reliability). Check out the chapter "Reliability" for more information on measuring reliability.

- **Validity**: the extent to which a test measures what it claims to measure

  - **Construct validity**: validity of inferences about the higher order constructs that represent sampling particulars. There are multiple types of construct validity; here are a few:

    * **Content validity**: experts using their judgment that something measures what it is supposed to measure
    * **Convergent validity**: correlations among two theoretically related constructs (or measurements) are strong and positive
    * **Divergent validity**: correlations among two theoretically not-related constructs (or measurements) are zero/null
    * **Criterion validity**: content on one test (predictor) correlates with performance on relevant criterion measures (outcome)

  - **Statistical validity**: validity of inferences about the correlation between treatment and outcome

  - **Internal validity**: validity about whether the observed relationship between A and B reflects a causal relationship between A and B

  - **External validity**: validity of inferences about whether the cause-effect relationship holds over variation in persons, places, treatment variables, and measurement variables

4.0.4   Statistical terms

- **Degrees of freedom (df)**: Degrees of freedom are the number of values that are free to vary and are not constrained.

  - The easiest way to understand them is with an example of the mean. If you know the sum of 4 numbers, then 3 of those numbers are free

to vary but the fifth value is dependent on (and constrained) to the values of the other. If your sum is 10, you could have three random values of 2, 3, and 1 (which are free to vary) but your fourth number *must* be 4. Therefore, the degrees of freedom in this example would be 3, which is n - 1.

– The statistical tests all have different ways of calculating degrees of freedom. For example, the one-sample t-test is n - 1, the independent t-test is $n_1 + n_2$ - 2, and the dependent t-test is n - 1 where n is the number of pairs.

- **Confidence intervals**: A confidence interval is such that in repeated sampling (e.g., replicating a study over and over and over again) then 95% of those confidence intervals should contain the true population parameter. You can see this visualized, and learn more about it, with this great visualization and read this great blog post. You can also read a chapter on this topic by Daniel Lakens.

4.0.5   Other terms

If other terms come up in the course of the semester that you believe should belong in this key term website, include it in your weekly reflection so I can update this page!

Part II

Intro to jamovi

(PART) Intro to jamovi

3. Overview of jamovi

jamovi is a free and open statistical software that helps us run our descriptive and inferential statistics. Why are we using jamovi and not another program?

1. Did I mention it's free? You won't ever have to pay a dime to use the software in the future, unless you use the paid version of jamovi Cloud.

2. It's open source, meaning that the statistical community helps support and improve the program. As jamovi says, "jamovi is made by the scientific community, for the scientific community."

3. It's built on top of the R statistical language, meaning you can begin learning how to code (if you want). I do all of my statistical analyses using R in a different program called RStudio (actually this book was developed in RStudio and hosted on GitHub!). It's a very powerful tool which is also free and open source.

4. It's incredibly easy to learn and use. I have taught statistics using both SPSS and jamovi, and students greatly prefer jamovi.

5. It promotes reproducibility. jamovi will save your data, analyses, options, and results all in one file so you can easily pick up where you left off. This will make your homework and future data analyses a breeze.

And if you don't want to hear it from me, here's another person's description of what jamovi is and why it's a great alternative to SPSS that provides a useful path toward learning R.

3.1 Installing jamovi

To install jamovi, go to https://www.jamovi.org/download.html. I recommend downloading the "current" version with the latest features, but you may want to get the "solid" version that is more stable.

jamovi has more documentation on how to install jamovi per operating system. Installation on macOS and Windows should be pretty straightforward. Here's some guidance if you are getting an error on Windows that "The app you're trying to install isn't a Microsoft verified app."

Chromebooks may have an issue installing jamovi if your version of Chromebook does not use Linux. Watch this video for support on installing jamovi on a Chromebook.

3.2 Navigating the jamovi interface

I recommend you watch this video by Alexander Swan on navigating the jamovi interface in jamovi to understand everything. Note that he is using a Mac in his demonstrations and that he has downloaded a ton of add-on modules which is why his Analyses tab looks different from yours.

Opening data in jamovi

I recommend you watch this video by Alexander Swan on opening files in jamovi.

Opening SPSS (.sav) files in jamovi

Here's another video walking you through how to open .sav files in jamovi.

https://www.youtube.com/watch?v=e-09OfLVs2U

Entering data in jamovi

Occasionally you need to manually enter data into jamovi instead of opening a pre-existing data file. I recommend you watch this video by Alexander Swan on entering data manually in jamovi.

Annotating output in jamovi

Sometimes I may ask you to submit a file of your output with annotations. I recommend you watch this video by Alexander Swan on annotating output in jamovi.

3.3 Additional jamovi videos and resources

Here are some additional resources on jamovi I recommend you bookmark and use when you are struggling or need additional support:

- jamovi 2022 tutorials by Alexander Swan on YouTube

- jamovi docs provide more support for using jamovi, including tips for transitioning from SPSS to jamovi and from jamovi to R

- caution: these use an older version of jamovi! Introduction to jamovi LinkedIn Learning course by Barton Poulson, founder of datalab.cc (you can find the video on YouTube and other places now, too, if you do not have LinkedIn Learning).

- Want more practice with jamovi? Want to do so using real data with a focus on DEI? Check out Neumann, Hood, & Neumann's (2013) datasets and problem sets!

Part III

NHST

Part IV

Inferential Statistics

Part V

Psychometrics

Part VI

Appendices