# 1. Concepts

When do we run one-sample t-tests versus z-tests?

* We run z-tests when we know population parameters: Standard Deviation. We run the one sample t-test when our SD is unknown, using a different significant value in its place.

When are inferential analyses statistically significant?

What do we mean by “power” in statistical analyses?

* The probability that a true significant effect will be represented in our dataset. Increases as the sample size increases. Power is limited when sample size is too small.

What is a Type I error and how is that error level set?

* Type one error is a false positive (when the sample appears to show a significant effect though there is none). We fail to affirm our null-hypothesis.
* We set the error level through the alpha-variable(a)

What is a Type II error?

Type two error is a false negative, when our alternative hyppthesis is true, but we fail to reject the null-hypothesis

* What is the relationship between Type II error and power?
  + If our power is too low, a type II error may occur.
* What effects the size of your power/type II error rates?
  + The sample size, and the treatment effect.

What is the relationship between Type I and Type II error?

Type I error is a false negative, whereas type II error is a false positive. Each error depends on our failing to reject a hypothesis.

What is a within-subjects (repeated measures) versus between subjects design?

Between-subject design divides our sample into test groups, giving each subject one treatment or the other. Within-subject design obtains data for both treatments by giving each subject both treatments.

Experimental vs. Non-Experimental design

* Experimental design
  + Intended to answer Qs of cause and effect between variables
  + Participants randomly assigned to diff groups or condition
    - Independent var manipulated by experimenter
  + Measurement
    - Measurement of a dependent variable in relation to the independent var.
  + Control
    - Other variables in the study that must be controlled since they may influence the DV
* Non-Experimental design
  + Demonstrates relationships between variables, but not causality
  + no random assignment, participants are observed and measured without manipulation of variables

When do you run the different tests (z-test versus one-sample t-test, versus paired samples t-test, versus independent samples t-test)?

What is an effect size?

What does the p-value mean and how do you interpret it?

When do you run a one-tailed versus two-tailed test? What is the difference between them (which is more conservative)?

Can you identify whether a researcher has a directional or non-directional hypothesis and do you know how that relates to whether you run a one-tailed or two-tailed test?

Can you identify a null hypothesis and alternative hypothesis from a research scenario?

* Null: there is no effect
* Alternate: there is an effect

Can you identify the independent vs. dependent variable?

* Independent: the variable we manipulate
* Dependent: the variable we measure to see the effect of the independent variable.

*Levels* of independent variable:

What are the different conditions of the independent variable? What treatments are being compared?

Can you identify a control group/condition versus an experimental group/condition?

Do you know what we mean by “manipulation” in an experiment? Can you identify which variable was manipulated?

When do we retain the null hypothesis versus reject the null hypothesis?

Can you draw information from, and interpret, APA-style written results?

Can you draw information from, and interpret, JASP statistical output?

Can you report information from JASP statistical output into APA-style written results?

# 2. Identifying directionality of hypotheses and appropriate test to run

**For each example, answer**

**a) is it a directional or non-directional hypothesis?**

**b) should you run a z-test, one-sample t-test, paired samples t-test, or independent samples t-test?**

**c) would you run a one-tailed or two-tailed test?**

Panera wants to know whether putting photos of the food on the menu leads people to spend more money on food than having menus without the photos. They assign some locations to have the photo menu and some to have the traditional menu without photos and then compare the amount in sales made.

A teacher examines whether students report differing levels of confidence if wrong answers get marked in red or blue ink. She marks some students’ assignments using a red pen, other students’ assignments using a blue pen, and then asks everyone to rate their confidence on a scale from 1 (not at all confident) to 7 (extremely confident).

A street musician wants to know whether people will throw more money into their donation box when they play upbeat songs versus slower (more melancholy) songs. They alternate playing the two types of songs to the same audience and keep track of how much money got thrown in for each song.

A researcher is going to study how socially anxious individuals perform on a task. First, the researcher needs to determine if his sample is socially anxious. The researcher concludes that a score of 7 (out of 10) or greater on the social anxiety questionnaire would indicate social anxiety, so he tests whether his sample has a score greater than 7.

# 3. Practicing interpreting output

Imagine the following output from a paired-samples t-test comparing students’ scores on a quiz taken on a rainy day versus a snowy day.

| **Paired Samples T-Test** | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Measure 1** | |  | | **Measure 2** | | **t** | | **df** | | **p** | | **Cohen's d** | |
| Rain |  | - |  | Snow |  | -0.108 |  | 16 |  | 0.915 |  | -0.026 |  |
|  | | | | | | | | | | | | | |
| Note.  Student's t-test. | | | | | | | | | | | | | |

### Descriptives

| **Descriptives** | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **N** | | **Mean** | | **SD** | | **SE** | |
| Rain |  | 17 |  | 5.882 |  | 1.364 |  | 0.331 |  |
| Snow |  | 17 |  | 5.941 |  | 1.638 |  | 0.397 |  |

How many students were tested? (How do you know?)

Was there a statistically significant difference between their scores on a rainy vs. snowy day? (How do you know?)

| **Independent Samples T-Test** | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **t** | | **df** | | **p** | | **Cohen's d** | |
| Distance |  | 2.057 |  | 16 |  | 0.028 |  | 0.970 |  |
|  | | | | | | | | | |
| Note.  For all tests, the alternative hypothesis specifies that group Control is greater than group Treatment . | | | | | | | | | |
| Note.  Student's t-test. | | | | | | | | | |

### Now imagine the following output is from an analysis comparing how many feet away people were willing to stand from a large spider. The researcher compared the distance away that people with arachnophobia stood compared to those without arachnophobia

| **Group Descriptives** | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | **Group** | | **N** | | **Mean** | | **SD** | | **SE** | |
|  |  | Phobics | |  | 9 |  | 6.222 |  | 1.856 |  | 0.619 |  |
|  |  | Non-phobics | |  | 9 |  | 4.444 |  | 1.810 |  | 0.603 |  |
|  | | | | | | | | | | | | |

How many students were tested? (How do you know?)

Was there a statistically significant difference between the distance away from a spider that those with arachnophobia stood versus those without arachnophobia? (How do you know?)

If there is a statistically significant difference, who stood farther away from the spider? (How do you know?)