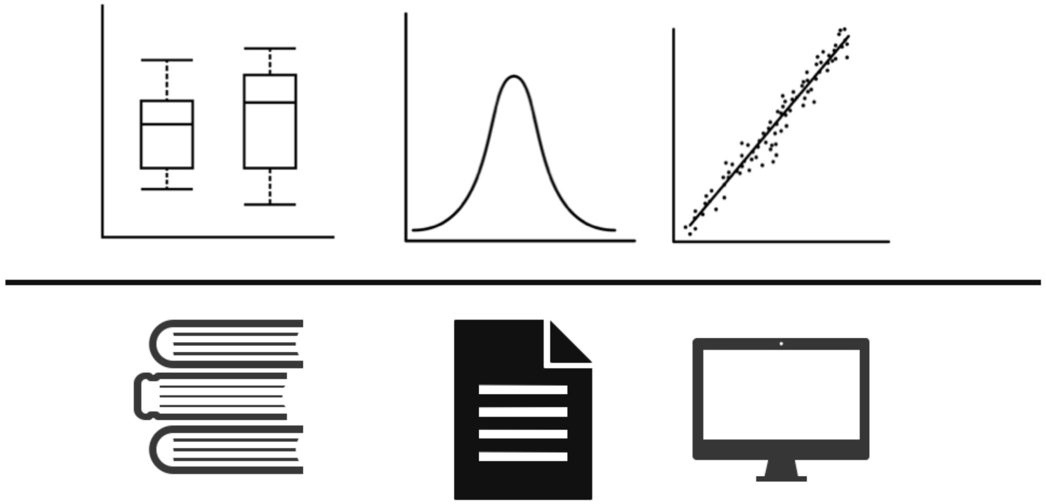
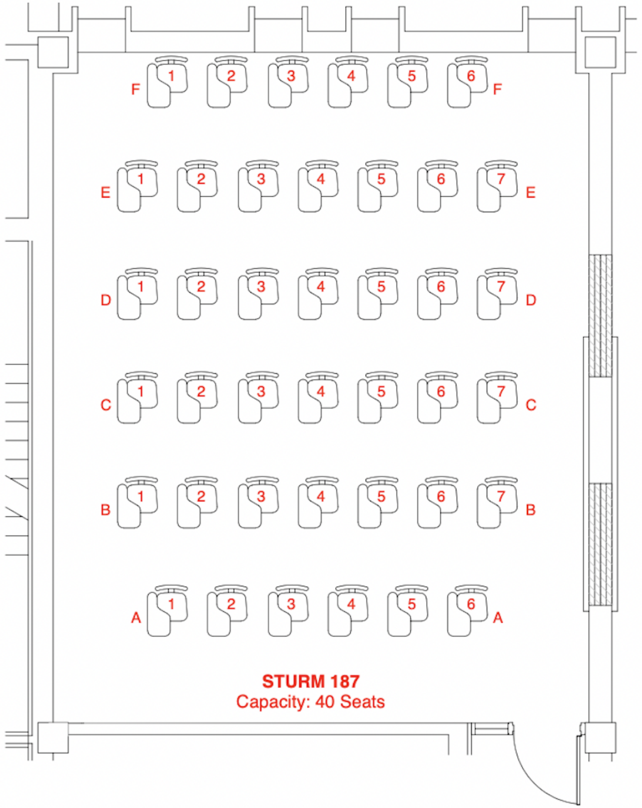
PSYC 2300

Introduction to Statistics



Lecture 05: Using Hypotheses to Test Questions

Announcements: Seating



Seat Assignments

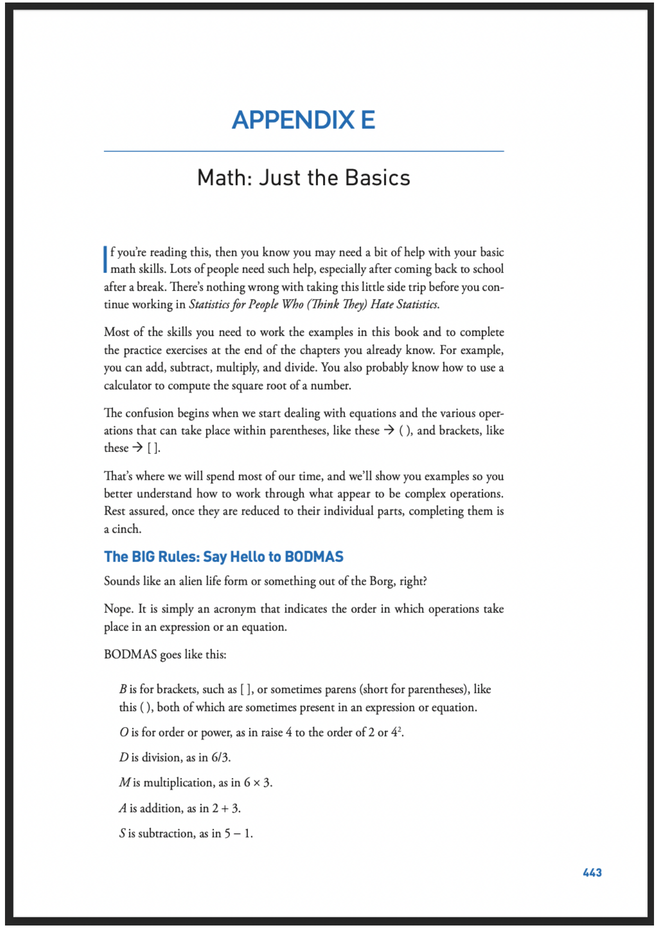
Announcements: Zoom

In-person classes will also be live-streamed

Announcements: Peer Notetaker

Please email me at chris.gunderson@du.edu

Announcements: Math



APPENDIX E

Math: Just the Basics

If you're reading this, then you know you may need a bit of help with your basic math skills. Lots of people need such help, especially after coming back to school after a break. There's nothing wrong with taking this little side trip before you continue working in *Statistics for People Who (Think They) Hate Statistics.*

Most of the skills you need to work the examples in this book and to complete the practice exercises at the end of the chapters you already know. For example, you can add, subtract, multiply, and divide. You also probably know how to use a calculator to compute the square root of a number.

The confusion begins when we start dealing with equations and the various operations that can take place within parentheses, like these → ( ), and brackets, like these → [ ].

That's where we will spend most of our time, and we'll show you examples so you better understand how to work through what appear to be complex operations. Rest assured, once they are reduced to their individual parts, completing them is a cinch.

The BIG Rules: Say Hello to BODMAS

Sounds like an alien life form or something out of the Borg, right?

Nope. It is simply an acronym that indicates the order in which operations take place in an expression or an equation.

BODMAS goes like this:

B is for brackets, such as [ ], or sometimes parens (short for parentheses), like this ( ), both of which are sometimes present in an expression or equation.

O is for order or power, as in raise 4 to the order of 2 or 4^2.

D is division, as in 6/3.

M is multiplication, as in 6×3.

A is addition, as in $2 + 3$.

S is subtraction, as in $5 - 1$.

443

p. 443 in your textbook

Outline for today

- **Open Questions**
- **Review portions of last class**
  - Correlations in JASP
  - Cronbach's Alpha in JASP
- **Hypothesis Testing**
  - The logic of hypothesis testing
  - z-test statistic

Open Questions

Do you have any theoretical or computational questions about the following?

- **Measures of Central Tendency**: Mean, Median, Mode, Skew
- **Scales of Measurement**: Nominal, Ordinal, Interval, Ratio
- **Measures of Variability**: Variance, Standard Deviation, Sum of Squares
- **Data Visualiztion**: Pie chart, bar graph, histogram, boxplot, line graph, scatterplot
- **Relationships in the world**: Pearson's r
- **Internal consistency**: Cronbach's α
- **Reliability**: Test-retest, Parallel forms, Interrater
- **Validity**: Content, criterion, construct

JASP

Correlations in JASP

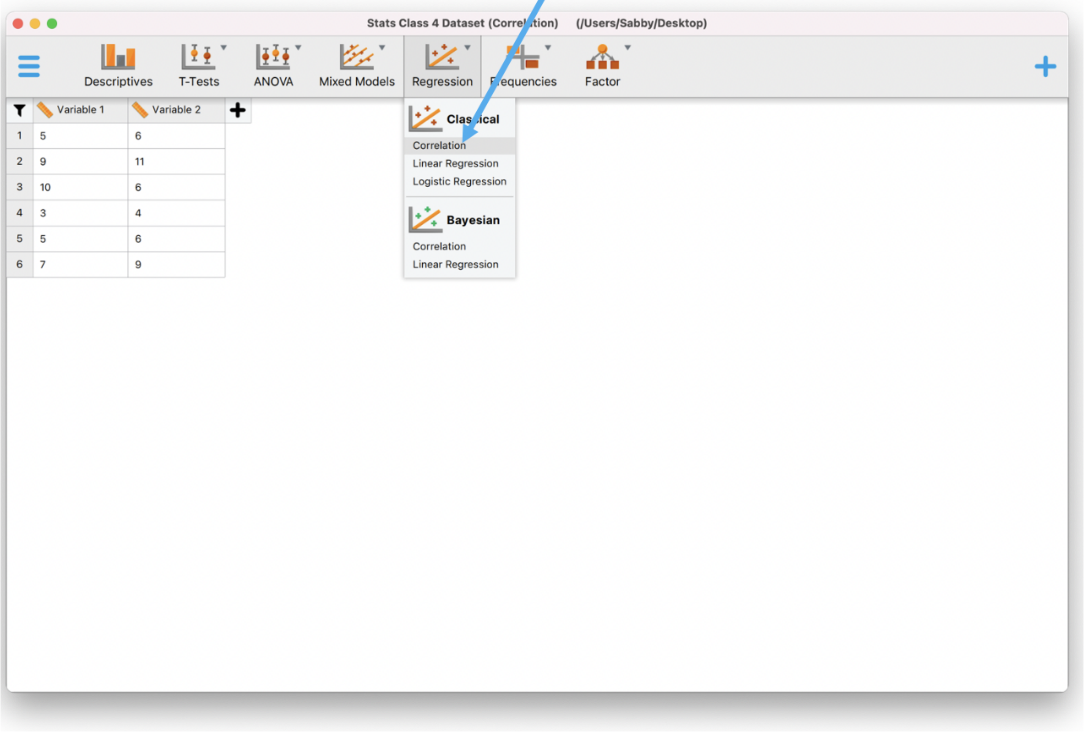
- Download `Stats Class 4 Dataset (Correlation).jasp` from Canvas

Cronbach's Alpha in JASP

- Download `Stats Class 5 Dataset (Cronbach's Alpha).jasp` from Canvas

JASP: Correlation

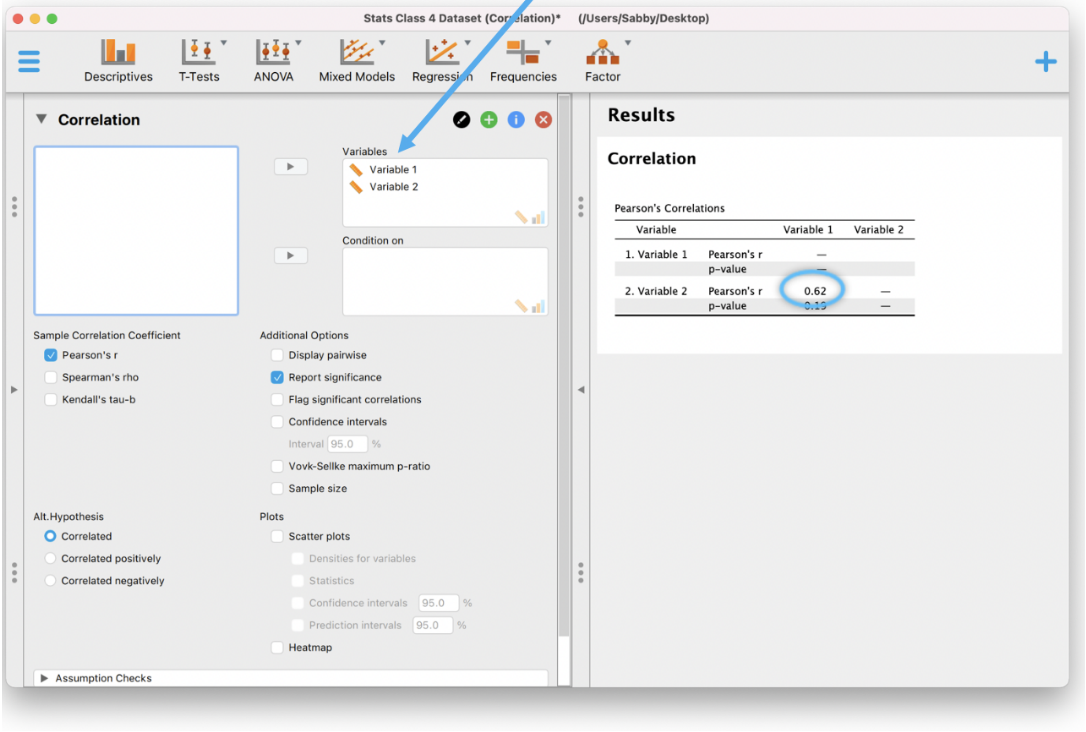
To begin to investigate correlations and to graph scatterplots, click on the “Regression” button and select “Correlation"



1

JASP: Correlation

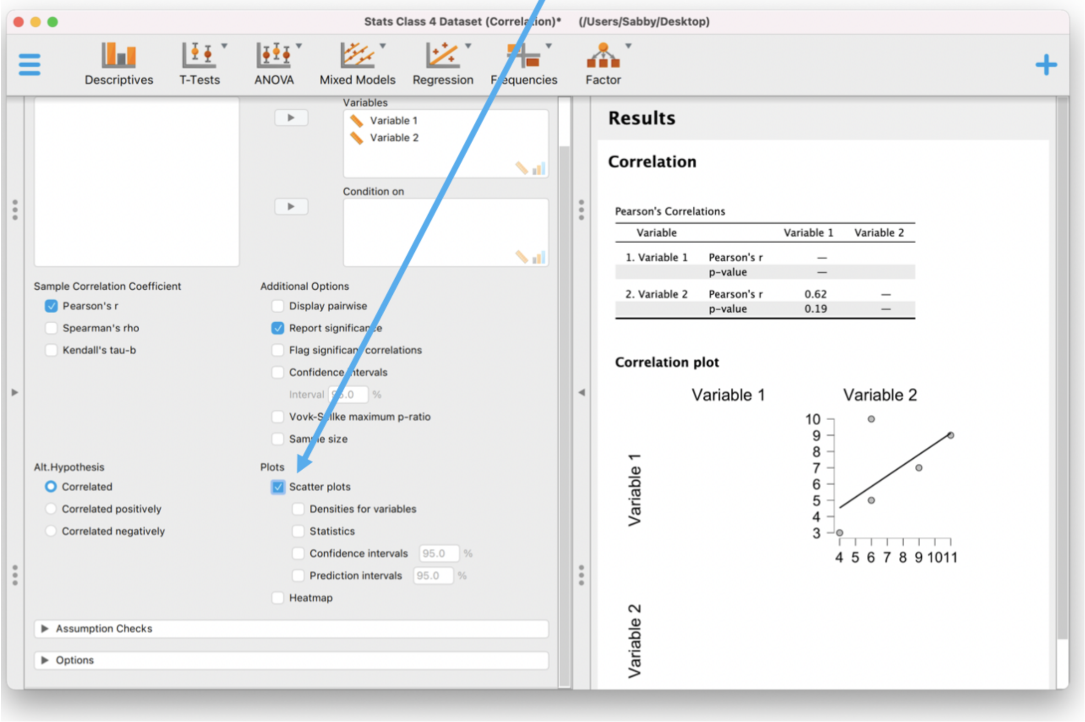
Next, drag your variables from the box on the left into the "Variables" box; JASP will generate a correlation table on the right



2

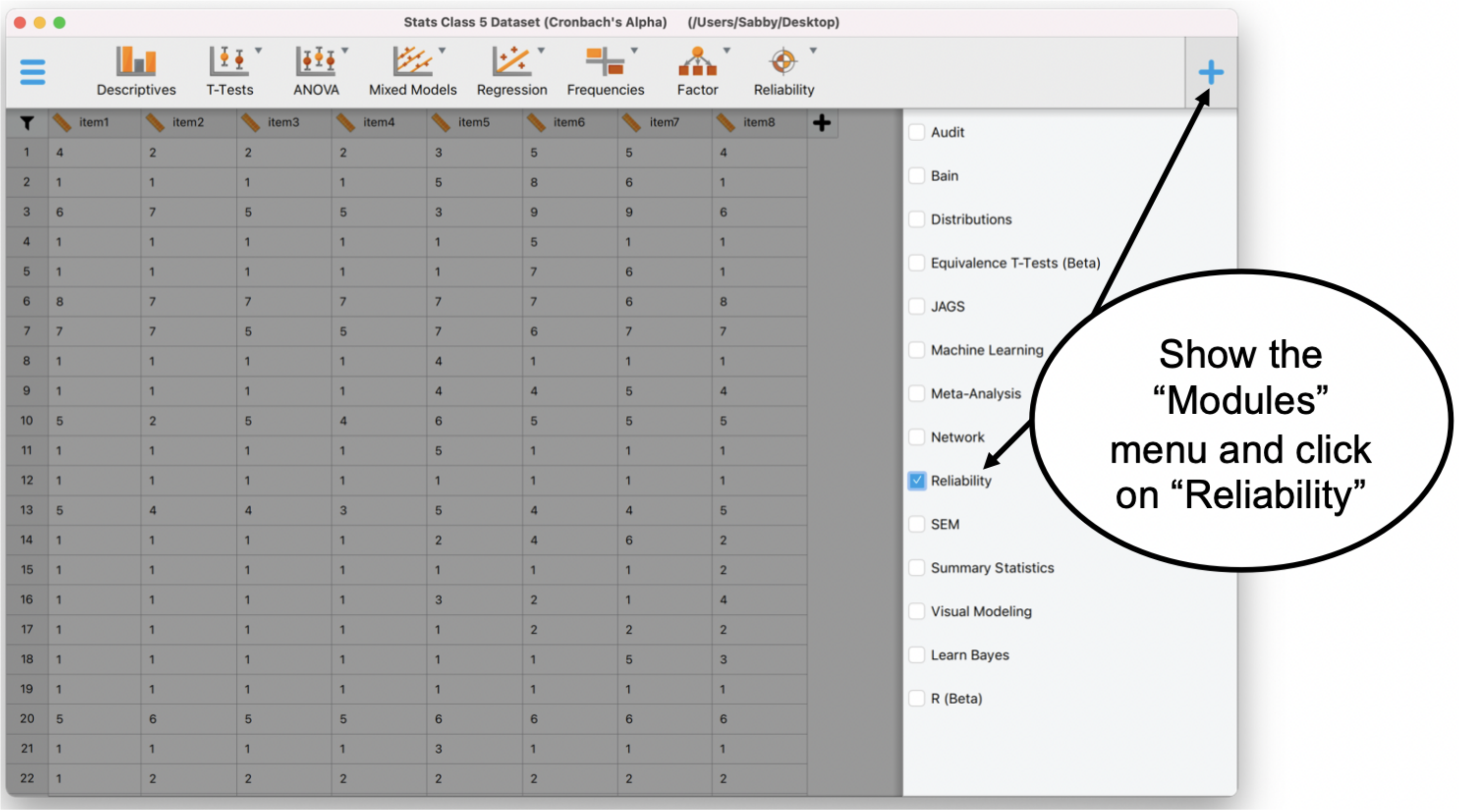
JASP: Correlation

Finally, click on "Scatter plots" under plots to make JASP generate a scatterplot (or multiple, if you have >2 variables)



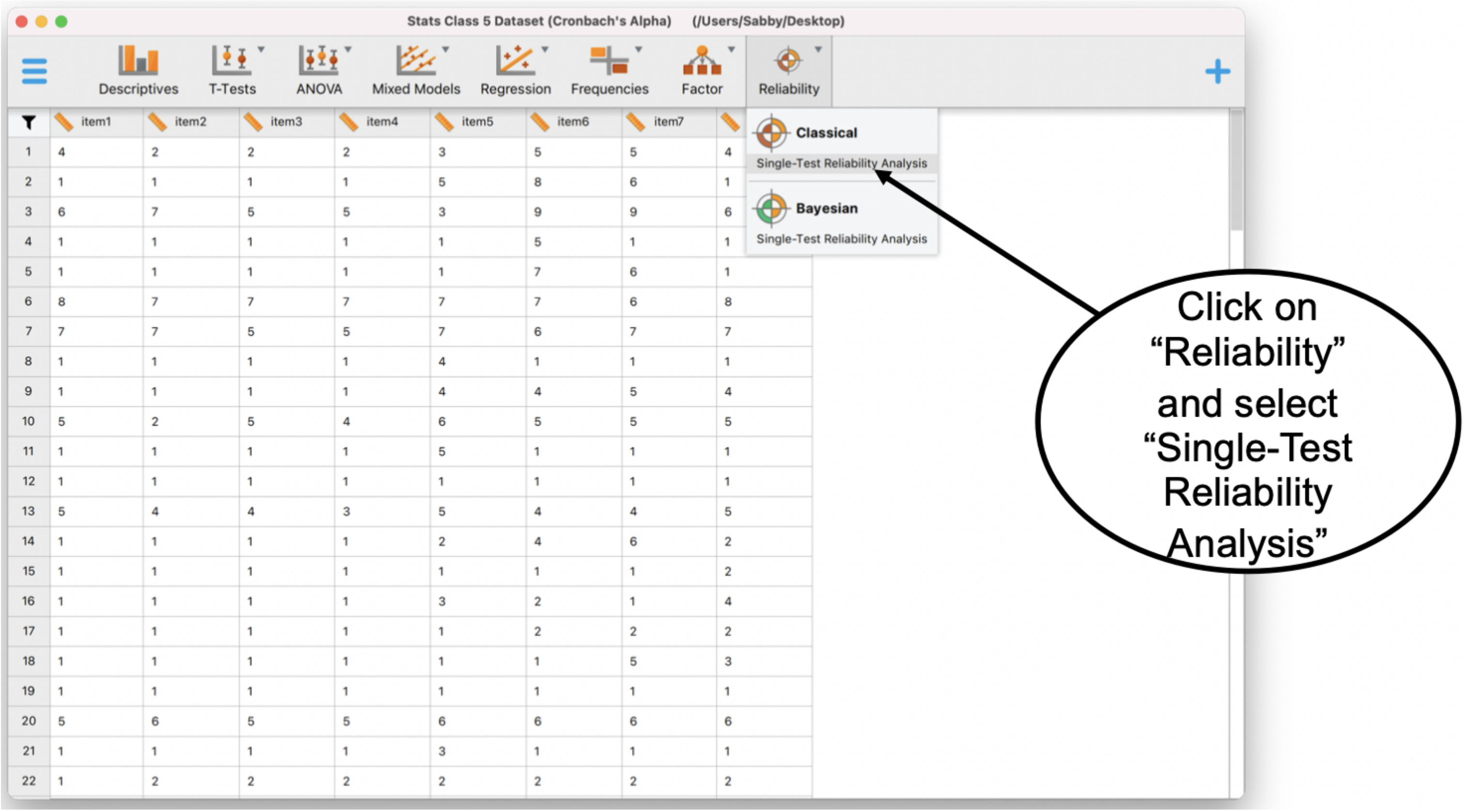
3

JASP: Cronbach's Alpha



1

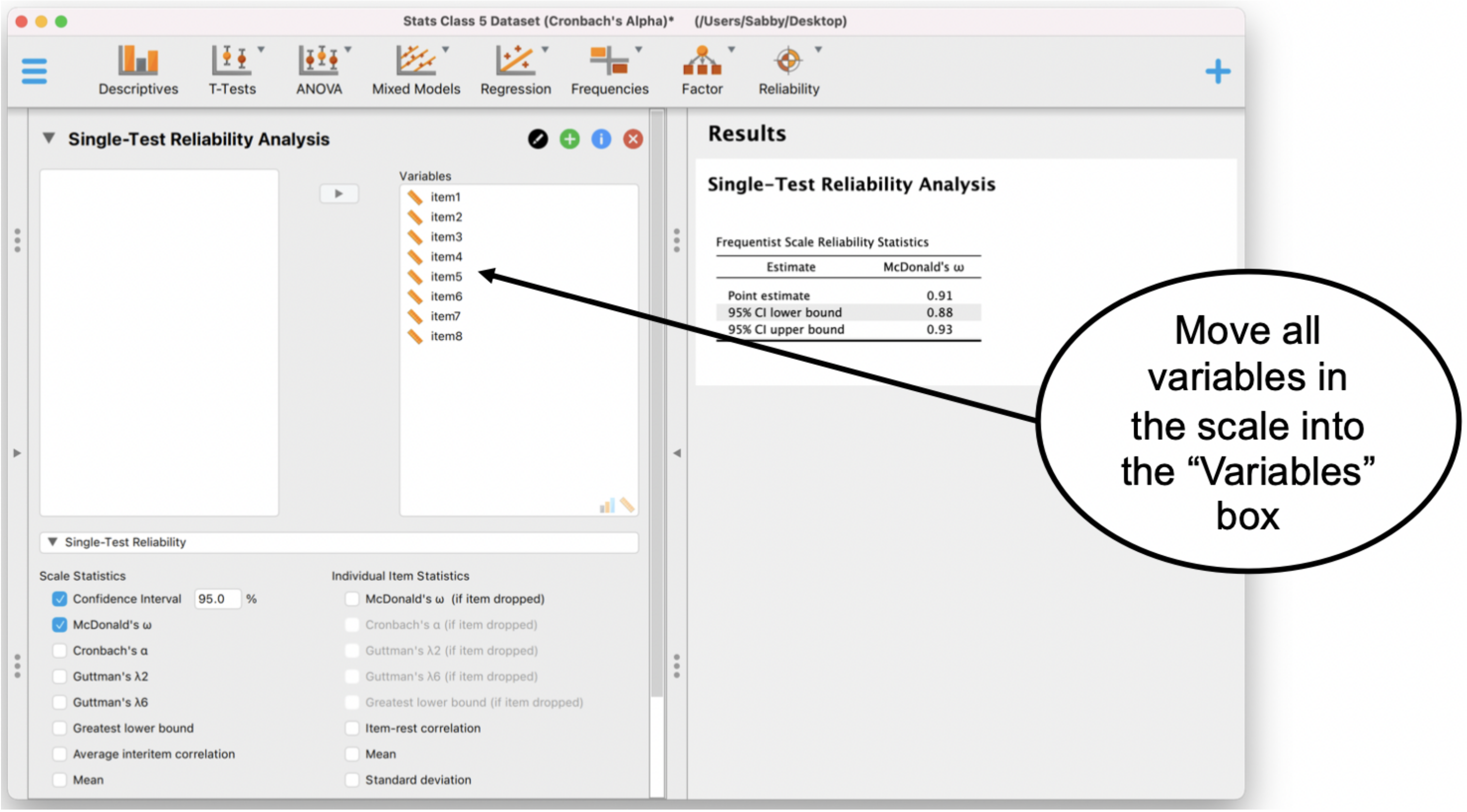
JASP: Cronbach's Alpha



	item1	item2	item3	item4	item5	item6	item7	
1	4	2	2	2	3	5	5	4
2	1	1	1	1	5	8	6	1
3	6	7	5	5	3	9	9	6
4	1	1	1	1	1	5	1	1
5	1	1	1	1	1	7	6	1
6	8	7	7	7	7	7	6	8
7	7	7	5	5	7	6	7	7
8	1	1	1	1	4	1	1	1
9	1	1	1	1	4	4	5	4
10	5	2	5	4	6	5	5	5
11	1	1	1	1	5	1	1	1
12	1	1	1	1	1	1	1	1
13	5	4	4	3	5	4	4	5
14	1	1	1	1	2	4	6	2
15	1	1	1	1	1	1	1	2
16	1	1	1	1	3	2	1	4
17	1	1	1	1	1	2	2	2
18	1	1	1	1	1	1	5	3
19	1	1	1	1	1	1	1	1
20	5	6	5	5	6	6	6	6
21	1	1	1	1	3	1	1	1
22	1	2	2	2	2	2	2	2

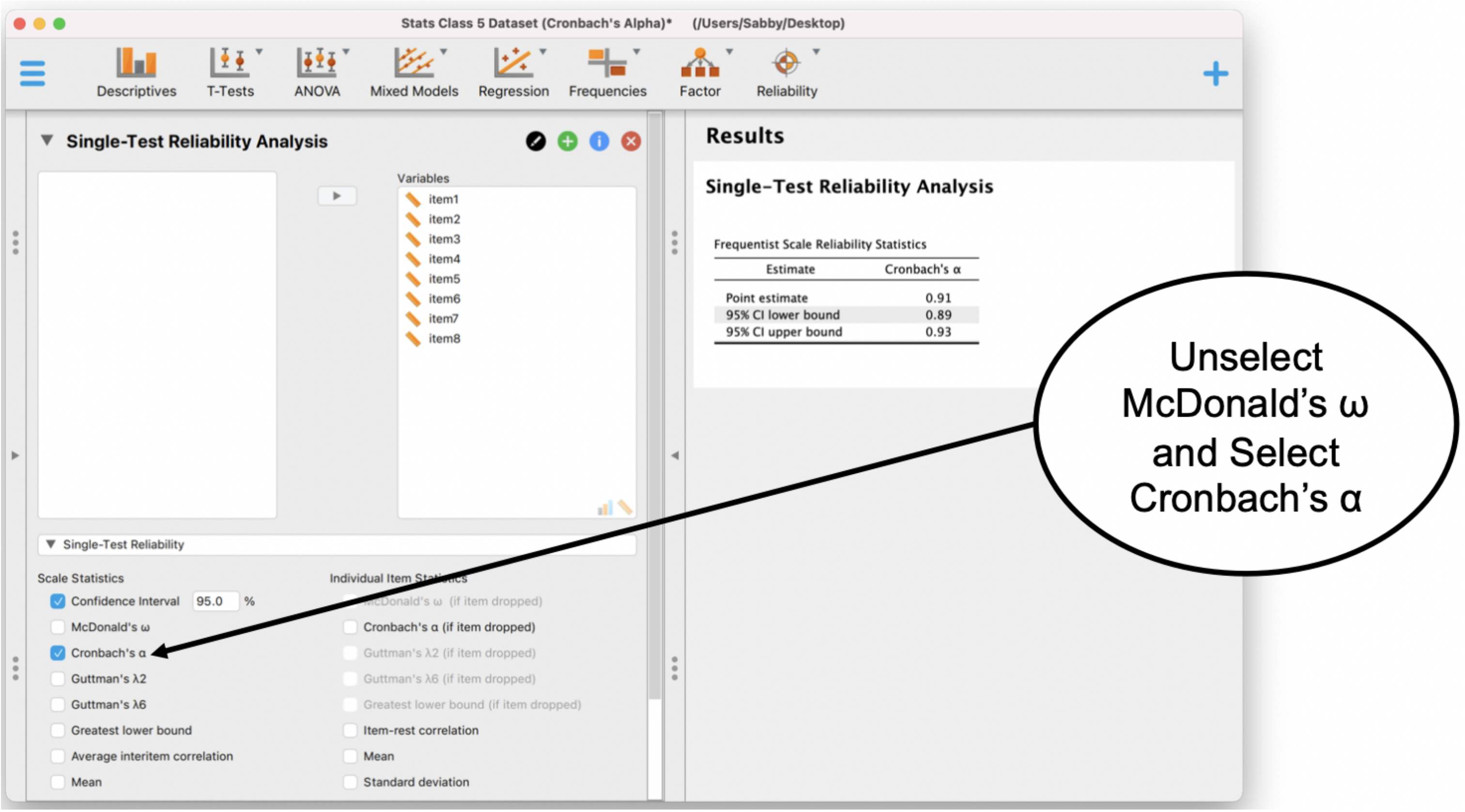
2

JASP: Cronbach's Alpha



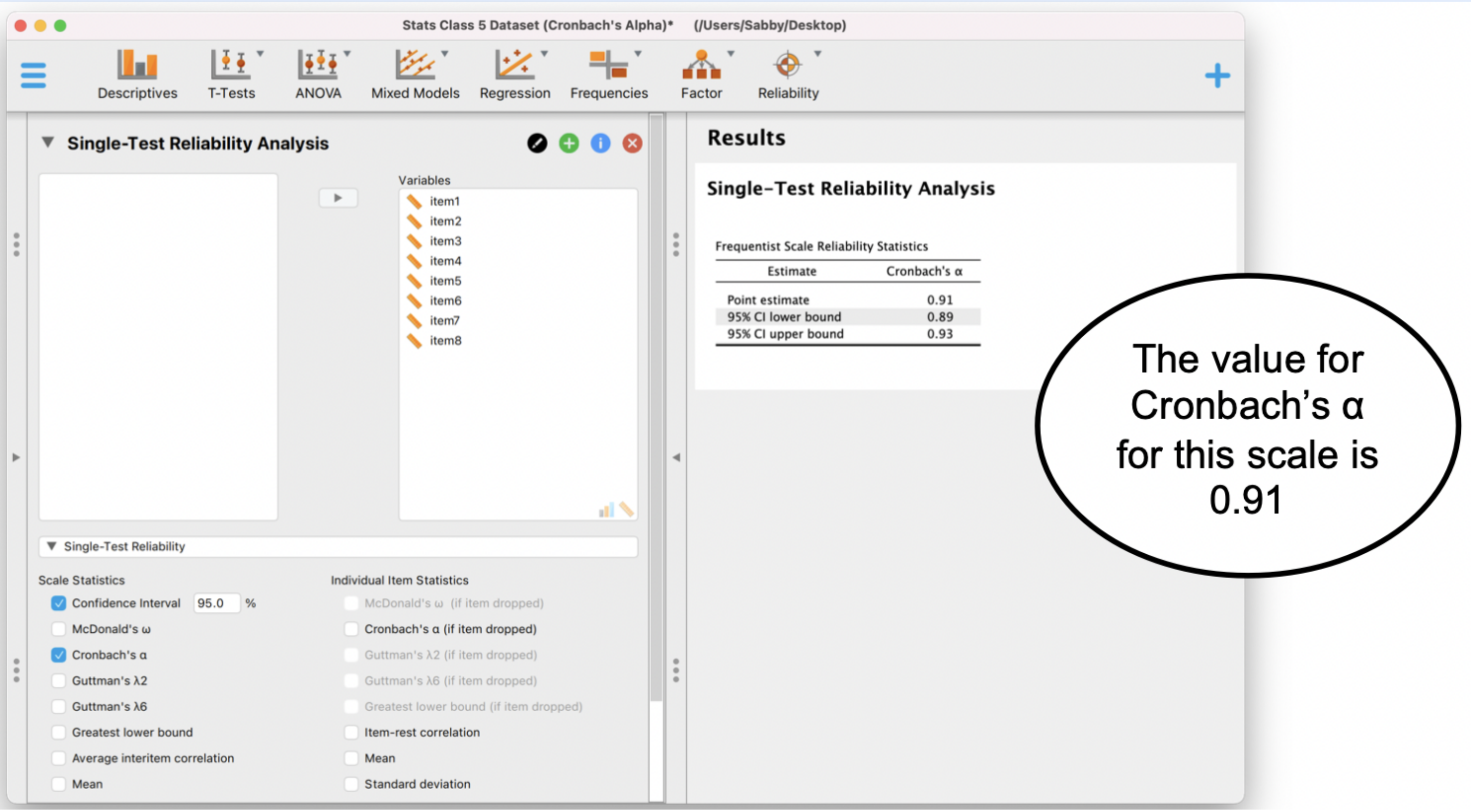
3

JASP: Cronbach's Alpha



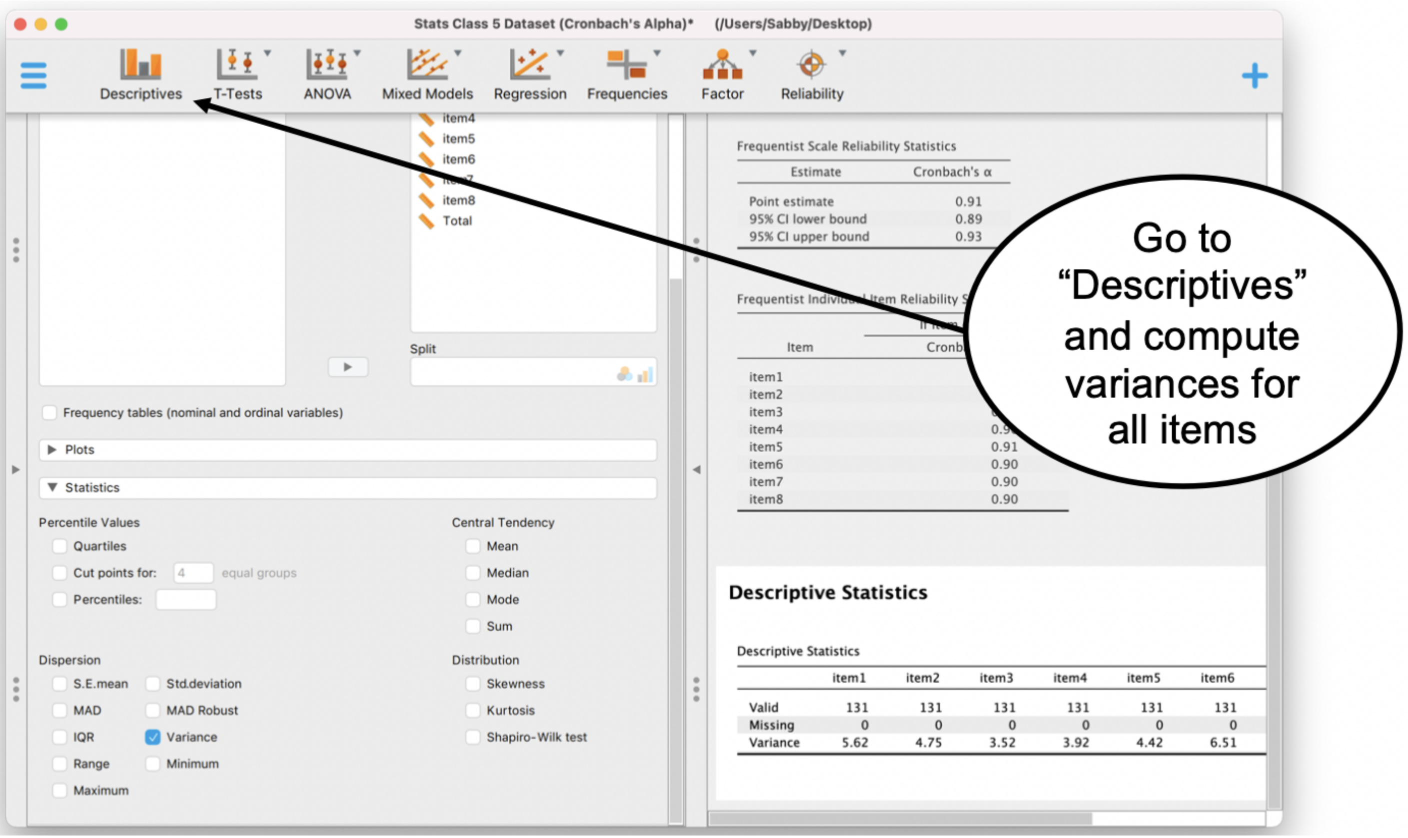
4

JASP: Cronbach's Alpha



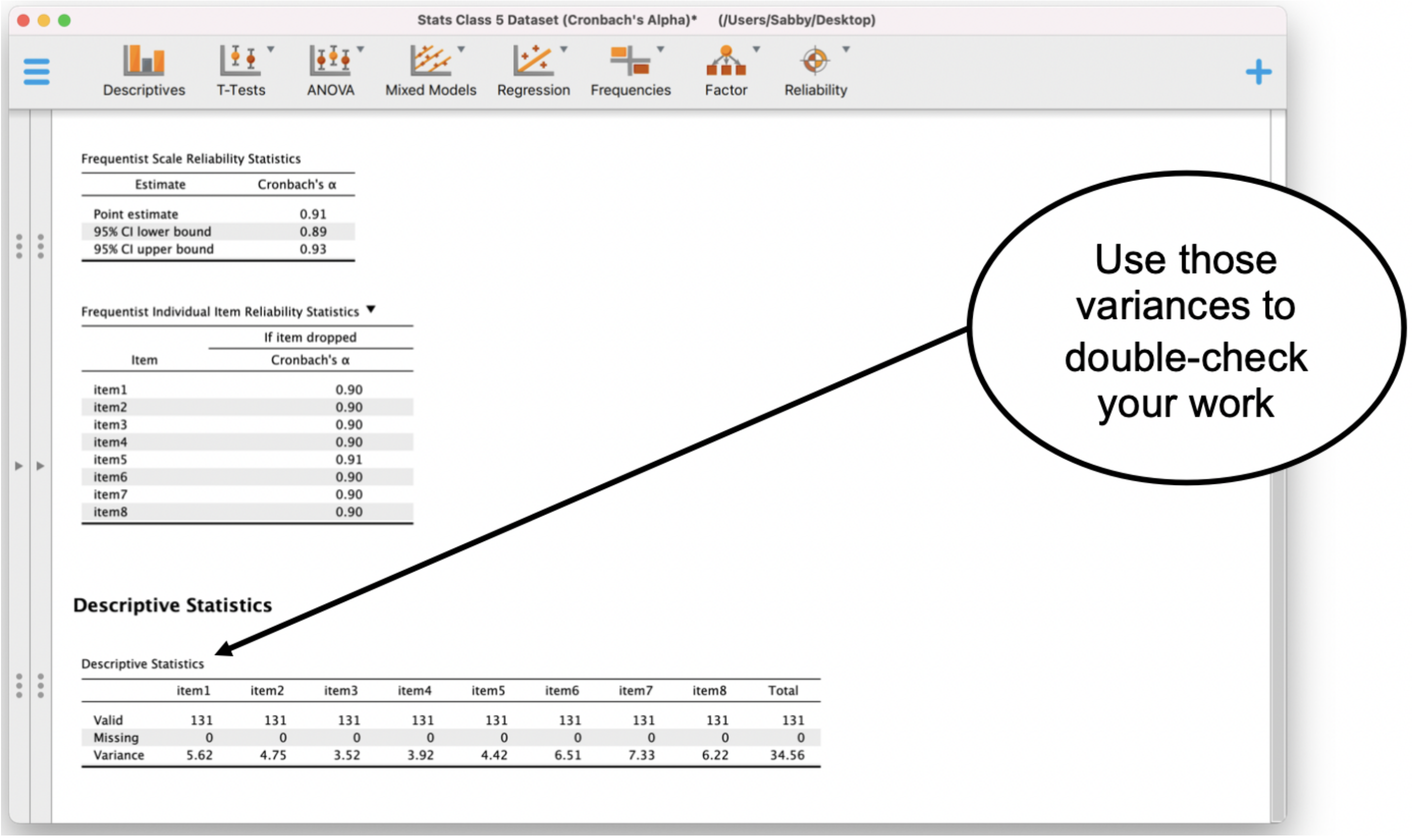
5

JASP: Cronbach's Alpha



6

JASP: Cronbach's Alpha



Frequentist Scale Reliability Statistics

Estimate	Cronbach's α
Point estimate	0.91
95% CI lower bound	0.89
95% CI upper bound	0.93

Frequentist Individual Item Reliability Statistics

	If item dropped
Item	Cronbach's α
item1	0.90
item2	0.90
item3	0.90
item4	0.90
item5	0.91
item6	0.90
item7	0.90
item8	0.90

Descriptive Statistics

Descriptive Statistics

	item1	item2	item3	item4	item5	item6	item7	item8	Total
Valid	131	131	131	131	131	131	131	131	131
Missing	0	0	0	0	0	0	0	0	0
Variance	5.62	4.75	3.52	3.92	4.42	6.51	7.33	6.22	34.56

7

The Logic of Hypothesis Testing

Hypothesis

Hypothesis: a specific, clear, and testable proposition or predictive statement about the possible outcome of a scientific research study

Simply put, it is an educated guess

The Basics

Two types of statistics

Descriptive Statistics: Used to organize or describe the characteristics of a dataset

Inferential Statistics: Used to make inferences (reach a conclusion) about the data

The Basics

Collecting samples of data

Sample: a subset of a population

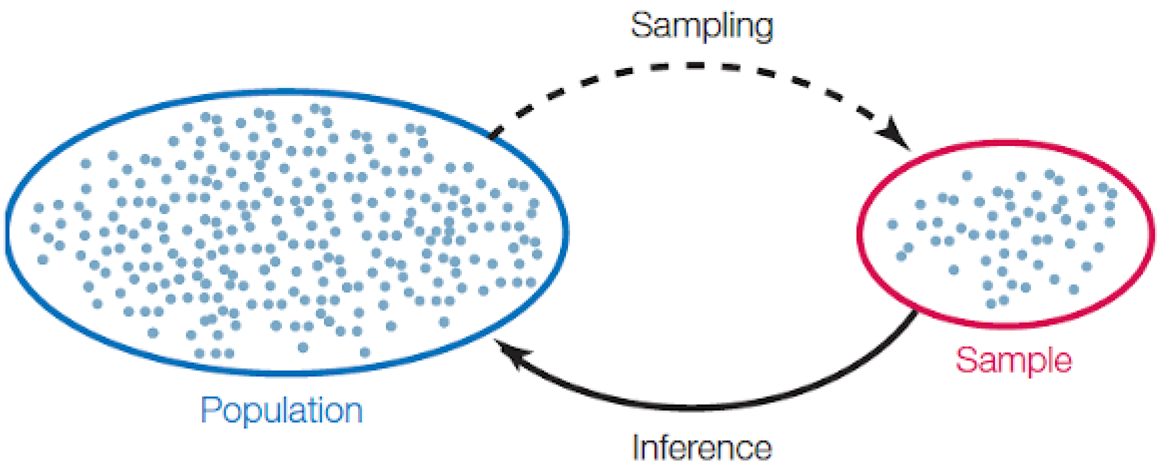
- Example: A researcher wants to know if people higher on emotional intelligence are better lie detectors.

Click here to find out

Hypothesis Testing

Hypothesis testing involves using samples to make inferences about the population

Equivalently, we gather observed scores to make inferences about true scores while simultaneously reducing error scores



Hypothesis Testing

When we use *samples* to approximate *populations*, however, we always have **sampling error** (difference between the sample statistic and the population parameter)

Sample Statistic $\overline{x}$

Population Parameter μ

Sampling Error ϵ

Hypothesis Testing

Hypothesis testing is an application of decision theory, designed to help us make reasonable inferences about populations *under uncertainty*

Hypothesis Testing

Hypothesis Test: A statistical method that uses sample data to evaluate a hypothesis about a population

Hypothesis Testing

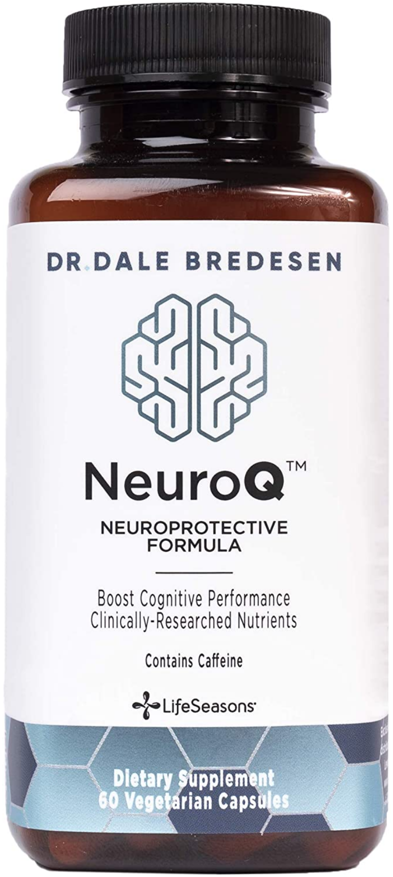
1. Form statistical hypotheses
2. Make predictions and establish a standard of evidence
3. Test hypotheses by collecting data and calculating statistics
4. Evaluate our hypotheses in light of evidence

Hypothesis Testing

1. **Form statistical hypotheses**
2. Make predictions and establish a standard of evidence
3. Test hypotheses by collecting data and calculating statistics
4. Evaluate our hypotheses in light of evidence

Designing an Experiment

Research Question: Does NeuroIQ actually change peoples' IQ?



Designing an Experiment

In the population

μ = 100

σ = 15

Give NeuroIQ to the population

This is what we would like to do, but is unrealistic

Designing an Experiment

In the population

μ = 100

σ = 15

Our sample

$$n = 15$$

$$\overline{X}_{IQ} = 105.9$$

$$s_{IQ} = 15.10$$

$$\begin{array}{c} 101, 122, 132 \\ 94, 129, 89 \\ 109, 92, 100 \\ 125, 103, 91 \\ 94, 116, 92 \end{array}$$

Designing an Experiment

Two possibilities

1. There is **no effect** of NeuroIQ on IQ scores. The 5.9-point increase is due only to sampling error, ϵ

2. There is a **real effect** of NeuroIQ on IQ scores, $\tau + \epsilon$

Epsilon
"error"

ϵ

Tau
"treatment"

τ

Statistical Hypotheses

Null Hypothesis: there is no change, no difference, or no relationship between variables

The difference observed is due only to *sampling error*, and not to any real effects in the population

Null Hypothesis

$$H_0 = \epsilon$$

Statistical Hypotheses

Research Hypothesis: there is change, a difference, or a relationship between variables

The difference observed is due to both sampling error and a real effect

Research Hypothesis

$$H_1 = \tau + \epsilon$$

The research hypothesis is sometimes called the alternative hypothesis, H_a

Statistical Hypotheses

Null Hypothesis

H_0: No real effect exists

Mutually exclusive

- Only one of these hypotheses can be true in the world

Research Hypothesis

H_1: A real effect does exist

Exhaustive

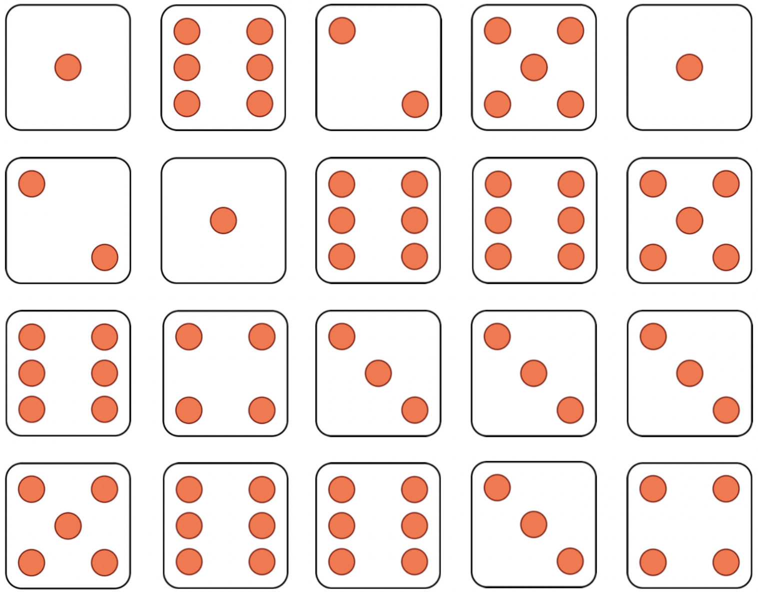
- One of these hypotheses is true in the world and must occur

As a result, discrediting one explanation gives us reason to believe in the other one

Dice Example

Null Hypothesis H_0

The die is normal



Alternative Hypothesis H_a

The die is loaded

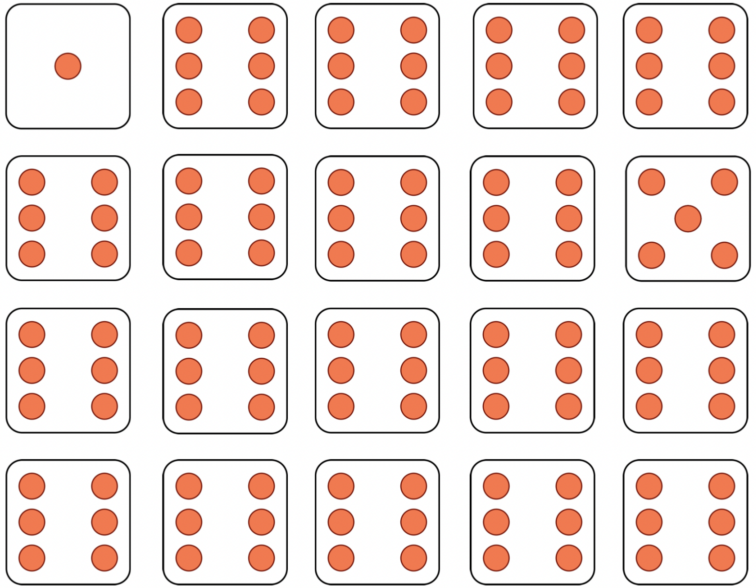
Number on die	Number of rolls
1	3
2	2
3	4
4	2
5	3
6	6

The evidence here **would probably *not* be strong enough** to reject the null hypothesis

Dice Example

Null Hypothesis H_0

The die is normal



Alternative Hypothesis H_a

The die is loaded

Number on die	Number of rolls
1	1
2	0
3	0
4	0
5	1
6	18

The evidence here **would probably be strong enough** to reject the null hypothesis

Statistical Hypotheses

Null Hypothesis H_0

$$\mu_{treatment} = \mu_{without\ treatment}$$

Alternative Hypothesis H_a

$$\mu_{treatment} \neq \mu_{without\ treatment}$$

- Note, however, that we aren't able to treat an entire population; rather, we treat **samples**

- Hypotheses are about **population parameters**, but we evaluate them with data from **samples**

Hypothesis Testing

1. Form statistical hypotheses

2. **Make predictions and establish a standard of evidence**

3. Test hypotheses by collecting data and calculating statistics

4. Evaluate our hypotheses in light of evidence

Back to our experiment

In the population

$\mu_{IQ} = 100$

$\sigma_{IQ} = 15$

Our sample

$n = 15$

$\overline{X}_{IQ} = 105.9$

$s_{IQ} = 15.10$

$$\begin{array}{c} 101, 122, 132 \\ 94, 129, 89 \\ 109, 92, 100 \\ 125, 103, 91 \\ 94, 116, 92 \end{array}$$

Back to our experiment

Null Hypothesis H_0

$$\mu_{NeuroIQ} = 100$$

Alternative Hypothesis H_a

$$\mu_{NeuroIQ} \neq 100$$

The Rationale

- In science, we seek to **falsify** (or disprove) our hypotheses
- H_1 is broad and would be hard to falsify by itself, so we choose to test it against H_0

Standard of Evidence

If we're testing against H_0, then, we need to determine *how extreme* (how far from 100) the sample mean would need to be for us to reject H_0

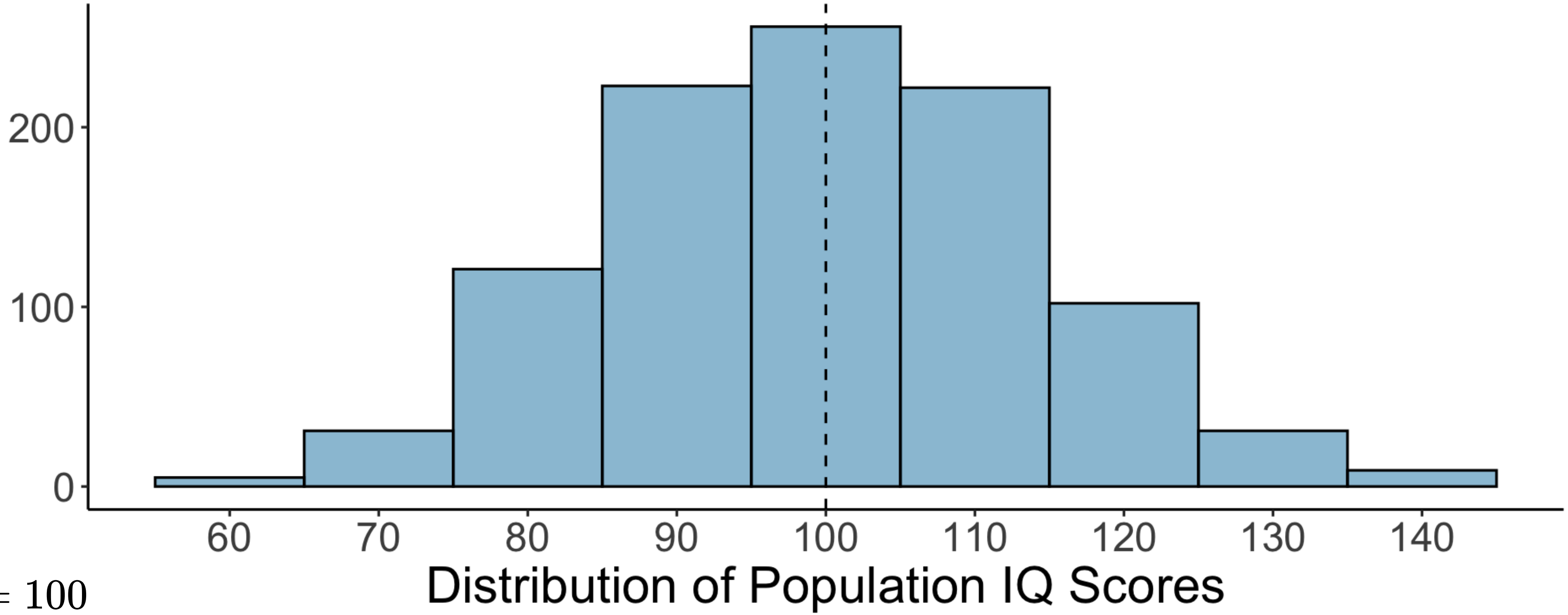
$\bar{x}_{NeuroIQ} = 150$?

$\bar{x}_{NeuroIQ} = 109$?

$\bar{x}_{NeuroIQ} = 60$?

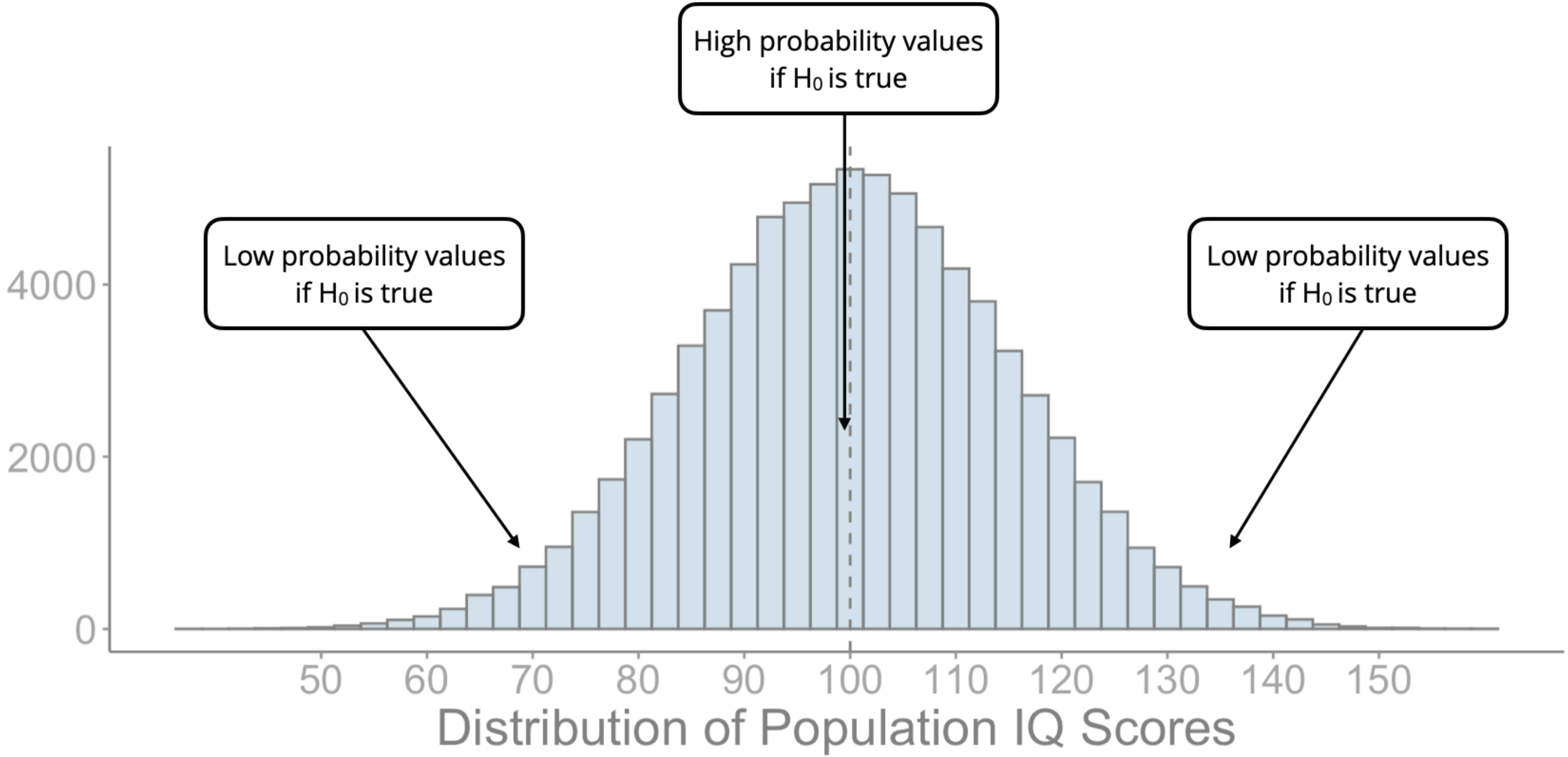
$\bar{x}_{NeuroIQ} = 92$?

Standard of Evidence



$\mu = 100$
$\sigma = 15$

Standard of Evidence



Standard of Evidence

Alpha: the probability value that is used to define which sample outcomes are considered very unlikely if H_0 is true

- To be precise about which values are extreme enough, we establish an alpha level
- This is also referred to as Significance level and is our "line in the sand"

Alpha Level

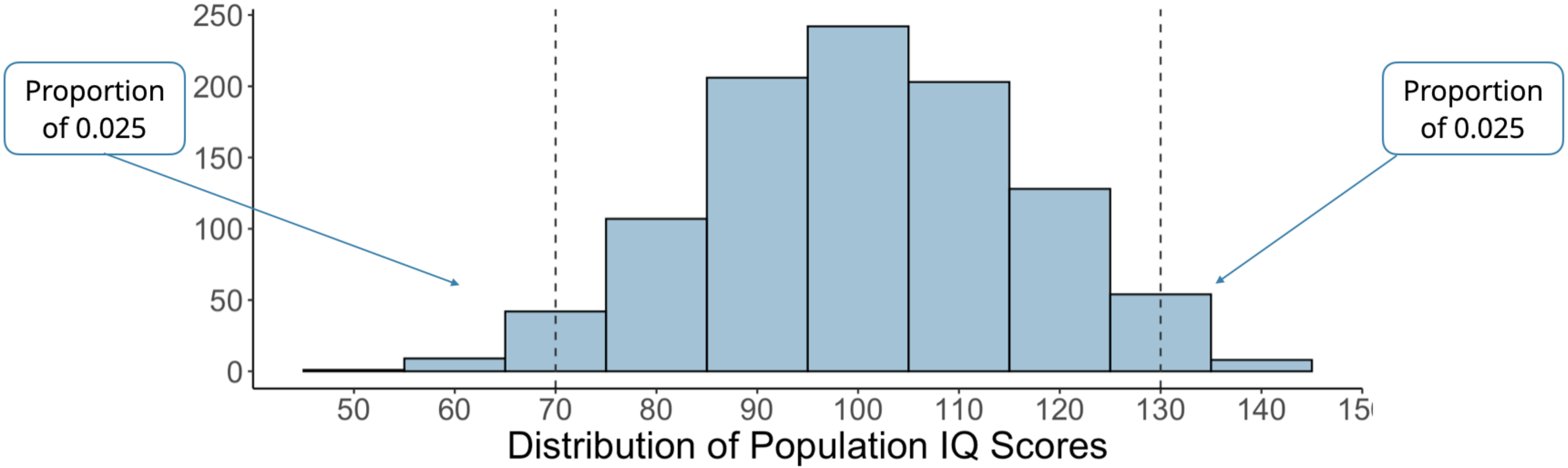
α

Standard of Evidence

The most common $\alpha = .05$

Interpretation: The least likely 5% of samples (in the extreme tails) would be evidence enough to reject H_0

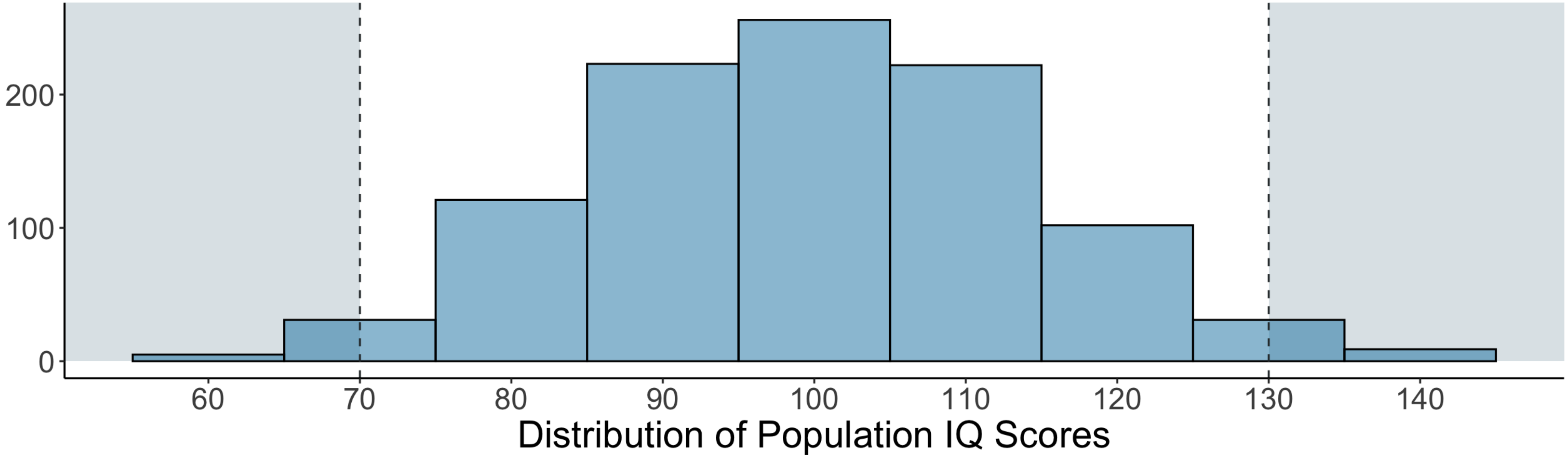
Standard of Evidence



$$\alpha = .05$$

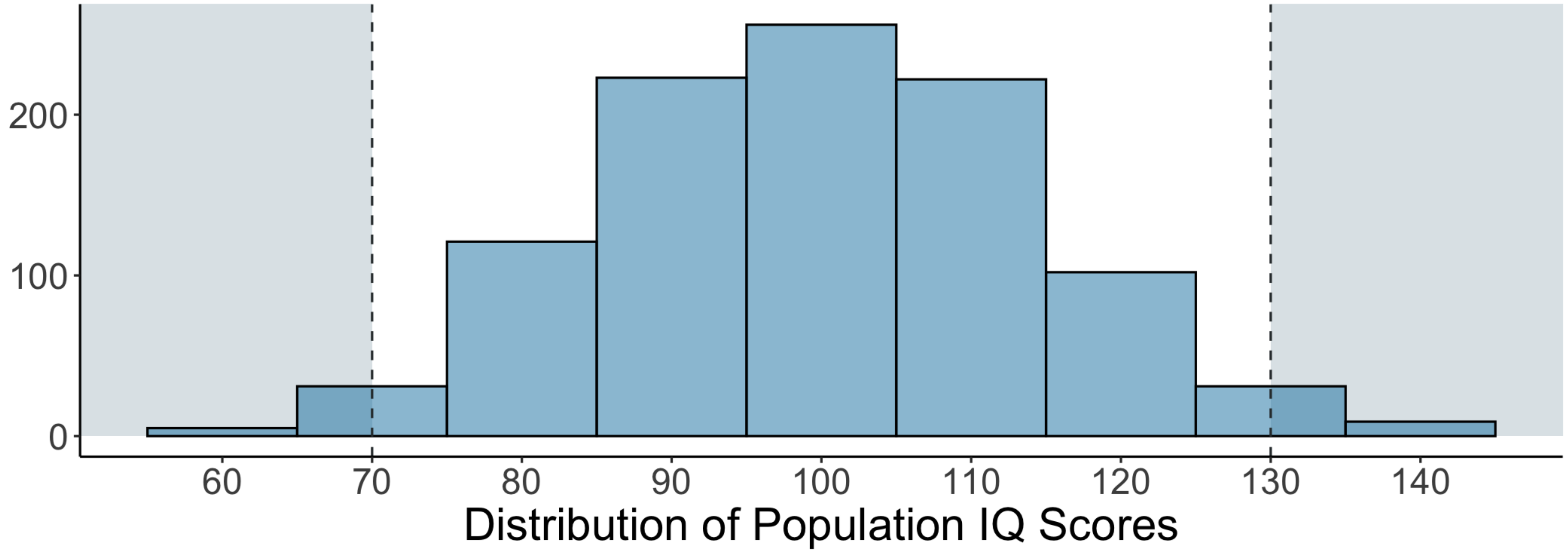
Standard of Evidence

Critical Region: The region (of the sampling distribution) that contains the sample outcomes that are considered *very unlikely* if H_0 is true



Standard of Evidence

If we get a test statistic extreme enough to be in the critical region, we can reasonably **reject** H_0



Hypothesis Testing

1. Form statistical hypotheses
2. Make predictions and establish a standard of evidence
3. **Test hypotheses by collecting data and calculating statistics**
4. Evaluate our hypotheses in light of evidence

p-value

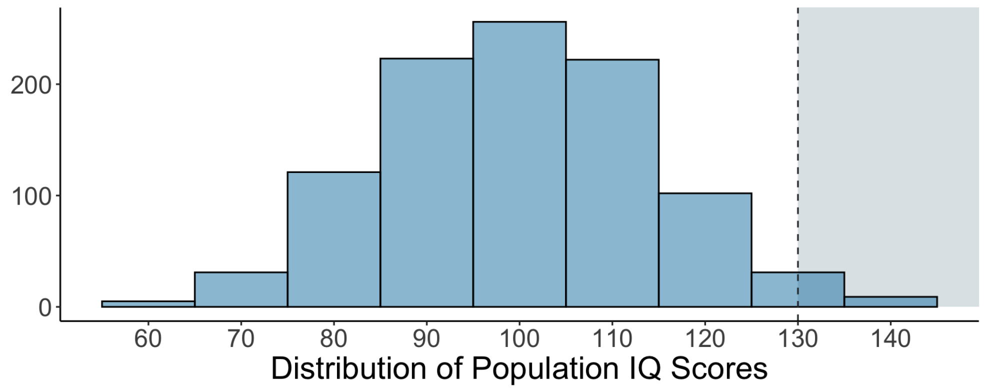
***p*-value**: the probability of getting the observed or more extreme data, *assuming* the null hypothesis is true

- Our first line of defense against being fooled by randomness
- *p*-values tell you how surprising the data is, *assuming* there is no effect

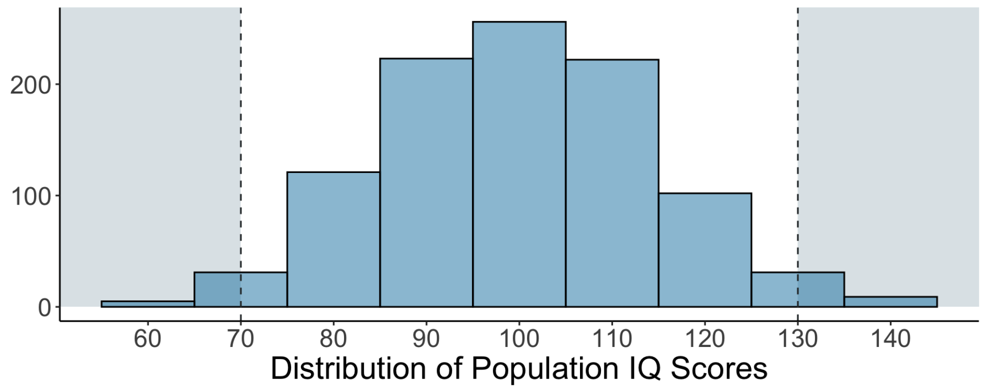
***p*-value**
p

p-value

One-tailed *p*-value: Proportion of distribution more extreme in one tail



Two-tailed *p*-value: Proportion of distribution more extreme in both tails



Using a one- or two-tailed *p*-value will depend on the research hypothesis

Test Statistic

Test statistic: A numerical summary of the degree to which a sample is unlike the samples predicted by the null hypothesis, H_0

***z*-test statistic**

$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}}$$

z-test Statistic

z-test statistic

$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}}$$

z-test Statistic

The sample mean we collected

Population mean assuming H_0 is true

z-test statistic

$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}}$$

Standard error

z-test Statistic

z-test statistic

$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}}$$

Standard Error

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

$\overline{x}$ = sample mean
$\mu_{\overline{x}}$ = population mean when H_0 is true
$\sigma_{\overline{x}}$ = standard error

σ = population standard deviation
n = sample size

Back to our experiment

In the population

$\mu = 100$

$\sigma = 15$

Our sample

$$n = 15$$

$$\overline{X}_{IQ} = 105.9$$

$$s_{IQ} = 15.10$$

$$101, 122, 132$$
$$94, 129, 89$$
$$109, 92, 100$$
$$125, 103, 91$$
$$94, 116, 92$$

Back to our experiment

In the population

μ = 100

σ = 15

Our sample

$$n = 15$$

$$\overline{X}_{IQ} = 105.9$$

$$s_{IQ} = 15.10$$

$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

$$z_{\overline{x}} = \frac{105.9 - 100}{\sigma_{\bar{x}}}$$

Back to our experiment

In the population

$\mu = 100$

$\sigma = 15$

Our sample

$$n = 15$$

$$\overline{X}_{IQ} = 105.9$$

$$s_{IQ} = 15.10$$

$$z_{\overline{x}} = \frac{105.9 - 100}{\sigma_{\bar{x}}}$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

$$\sigma_{\overline{x}} = \frac{15}{\sqrt{15}}$$

$$\sigma_{\overline{x}} = 3.87$$

Back to our experiment

In the population

μ = 100

σ = 15

Our sample

$$n = 15$$

$$\overline{X}_{IQ} = 105.9$$

$$s_{IQ} = 15.10$$

$$z_{\overline{x}} = \frac{105.9 - 100}{3.87}$$

$$z_{\overline{x}} = \frac{5.90}{3.87}$$

$$z_{\overline{x}} = 1.53$$

Hypothesis Testing

1. Form statistical hypotheses
2. Make predictions and establish a standard of evidence
3. Test hypotheses by collecting data and calculating statistics
4. **Evaluate our hypotheses in light of evidence**

Evaluating the Evidence

Results

$z_{\overline{x}} = 1.53$

$p = .126017$

$\alpha = .05$

Decision

If $p < \alpha$, reject the null hypothesis

If $p > \alpha$, fail to reject the null hypothesis

Since $p = .13 > \alpha = .05$, we fail to reject the null hypothesis

> *"We tested the hypothesis that NeuroIQ would change peoples' IQ. Our results suggest that there is no evidence that NeuroIQ changes peoples' IQ, z = 1.53, p = .13."*

Interpreting *p*-values

- If $p < \alpha$, the data you observe is **not** likely due to just sampling error
- If $p > \alpha$, the data is likely due to sampling error

Why?

- A *p*-value of .0001 means that only a tiny proportion of the distribution is more extreme than what we found
  - Small *p*-values would be **surprising** if the null hypothesis is true
- A *p*-value of .70 means that 70% of the distribution is more extreme than what we found
  - Large *p*-values are not surprising if the null hypothesis is true

Hypothesis Testing

Form statistical hypotheses

NeuroIQ changes peoples’ IQ

Make predictions and establish a standard of evidence

H_0: $\mu_{NeuroIQ} = 100$
H_a: $\mu_{NeuroIQ} \neq 100$
$\alpha = .05$

Test hypotheses by collecting data and calculating statistics

$z_{\bar{x}} = 1.53$

Evaluate our hypotheses in light of evidence

$p = .13 > \alpha = .05$. Fail to reject the null hypothesis.

Next time

Lecture

- Probability and the Normal Curve

Reading

- Chapter Seven

Quiz 1

- Due tonight at 11:59pm MT
  - Lecture 1-4, Ch.1-5