HYPOTHESIS TESTING FOR THE MEAN

Alison Kleffner

SCHEDULE

1. Basic Definitions
   - Population (Parameter) vs. Sample (Statistics)
2. Hypothesis Testing for a Single Mean
   - Setting up hypothesis
   - Conditions
   - Finding Standardized Statistics and Drawing Conclusion

POPULATION (PARAMETER) VS. SAMPLE (STATISTICS)

Most of what we do in Statistics is trying to understand a set of information. This set of information is from a . . .

- **<u>Population</u>:** The complete collection of ALL elements that are of interest for a given problem.

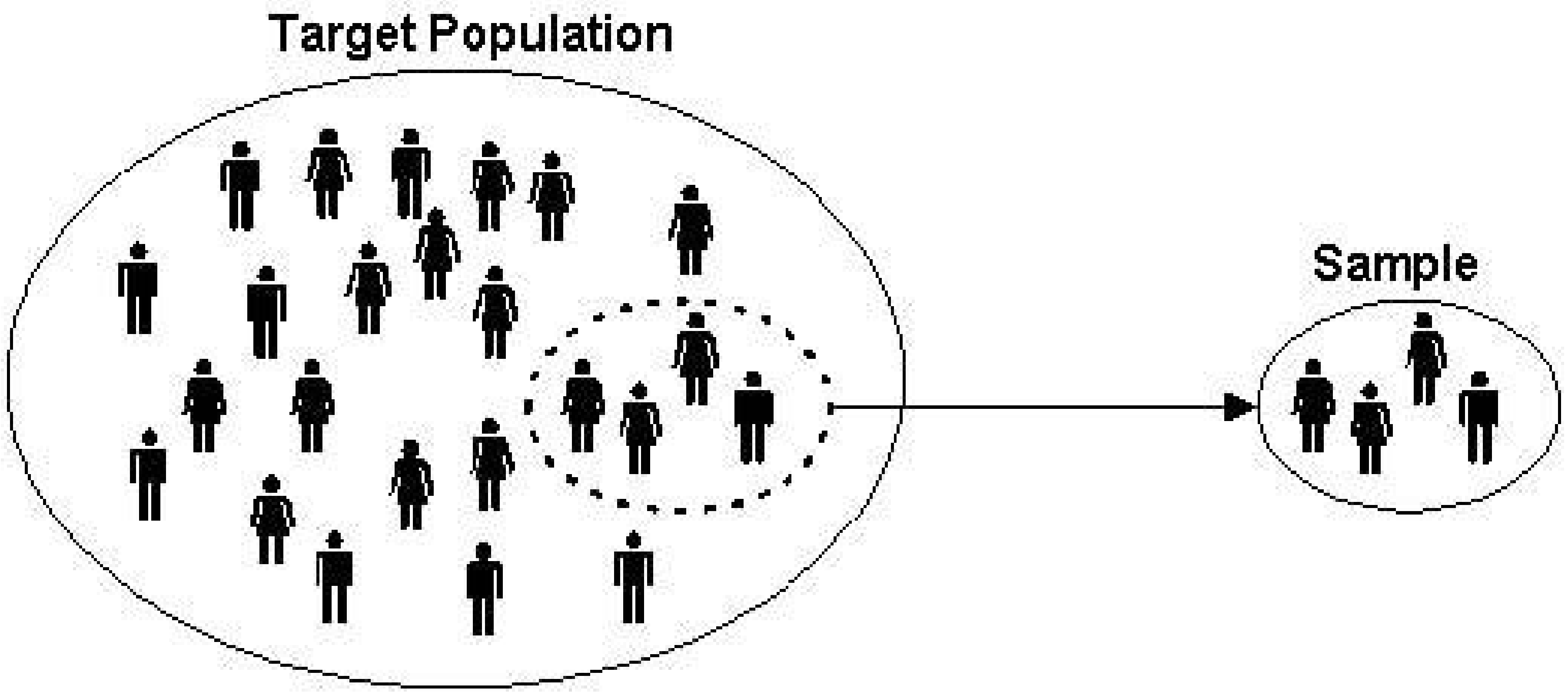
The population is often so big that obtaining all information about its elements is either difficult or impossible. So, we work with a more manageable set of data that we obtain from a . . .

- **<u>Sample</u>:** A sub-collection of elements drawn from a population.

- **<u>Observation</u>:** The collection of measurements from a particular unit in a sample.
  - What subject you are taking data on

- **<u>Variable</u>:** Any measurable characteristic of an observation.

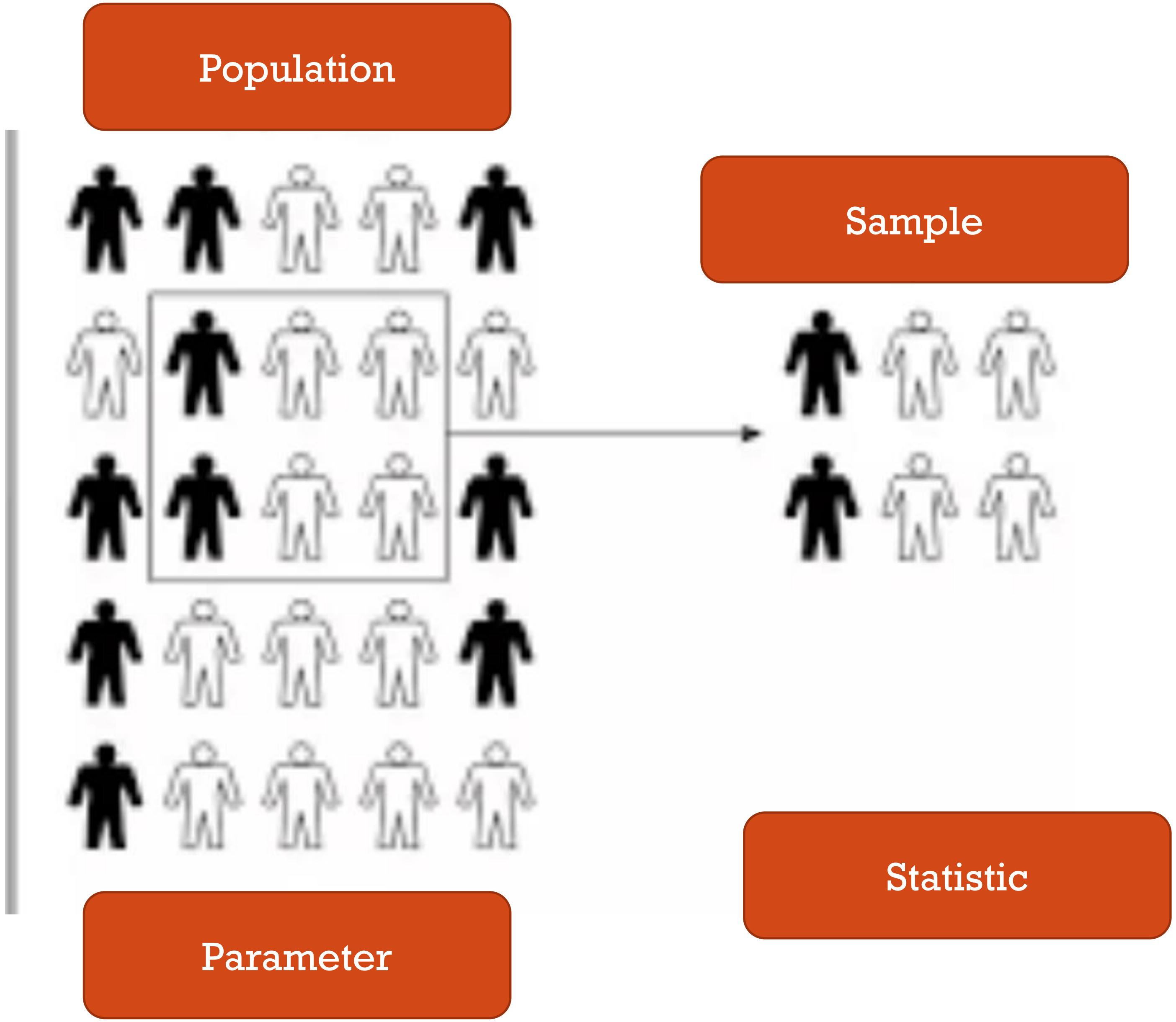
POPULATION VS SAMPLE



DEFINITIONS

- **Sample:** The set of observational units on which we collect data.
- **Sample size:** The number of observational units in the sample.
- **Statistic:** The number summarizing the result of the sample.
  - This is the number observed in the study.

- **Population:** The complete collection of ALL elements that are of interest for a given problem.
- **Parameter:** the long-run numerical property of the process.
  - "The long run average*…context*"
  - What is it that you want to know?



WANT TO USE STATISTIC TO ESTIMATE THE PARAMETER

Parameter → **P**opulation

Typically Unknown Value

Statistic → **S**ample

Known Value

(aka observed statistic)

STATISTICAL SIGNIFICANCE

- **Statistics** – Estimate broad populations
  - No way to collect all information
  - Want to know when is a sample Statistically Significant?

So What is Statistical Significance?

  - Asking ourselves: "Is our result unlikely to have occurred by *random chance?*"

How do we determine this? Hypothesis Testing

EXAMPLE

It is a well-known fact that Millennials LOVE Avocado Toast. It's also a well-known fact that all Millennials live in their parent’s basements. Clearly, they aren't buying homes because they are buying too much Avocado Toast! Was the Avocadopocalypse of 2017 real? In the past, an avocado cost $1.36. To test the avocadopaocalypse, we collected a random sample of 94 avocados from 2017. The average price of an avocado was $1.44 with a sample standard deviation of 0.41

- Population?
  - Avocados from 2017
- Parameter?
  - The long run average price of an avocado in 2017.
- Sample?
  - 94 avocados from 2017
- Statistic?
  - 1.44
- Sample Size?
  - 94

YOUR TURN

It has been recommended US adults get 8 hours of sleep per night. A random sample of 32 college students was taken and was found that on average, they got 5.5 hours of sleep with a standard deviation of 0.71.

- Population?
  - College Students
- Parameter?
  - True average number of hours a college student sleeps a night
- Sample?
  - 32 college students
- Statistic?
  - 5.5
- Sample Size?
  - 32

SYMBOLS

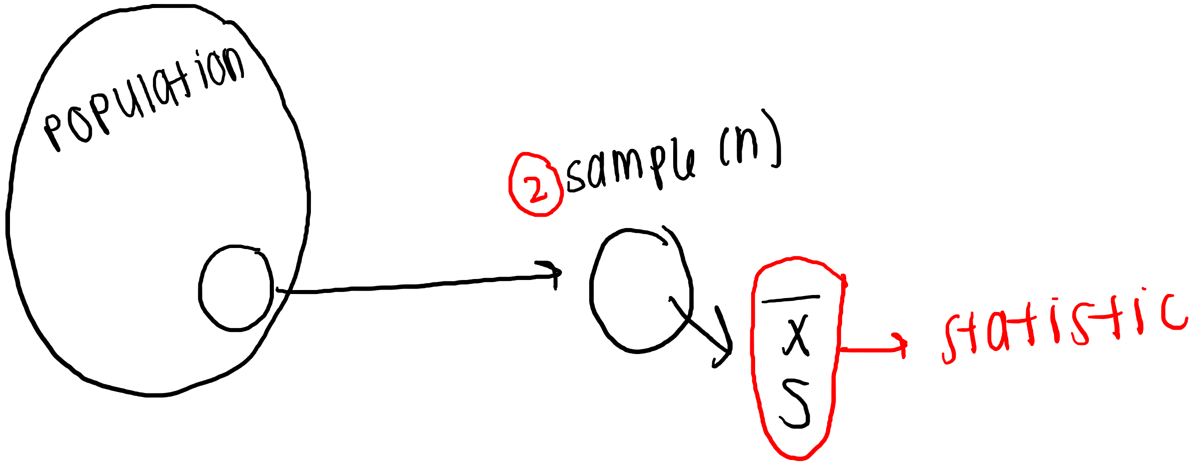
- μ is the **parameter**- the long run average
  - Since it is for the population we typically do not know this value
  - This is what we are interested in
- $\bar{x}$ is the **observed statistic**- the value we observe from our sample
- **n** is the sample size

- *Example:* Avocadopocalypse
  - Parameter (μ) – The long run average price of an avocado in 2017.
  - Observed Statistic –$\bar{x} = 1.44$
  - Sample Size – $\text{n} = 94$

HYPOTHESIS TESTING

HYPOTHESIS TESTING STEPS

1. Make an initial assumption.
2. Collecting data (sample) – find statistic
3. Based on the sample, reject or fail to reject the initial assumption.



SETTING UP HYPOTHESIS

- Summarize research question and parameter
- $H_o \rightarrow$ Null: "random chance alone"
  - Chance model, historical data
  - Equal to (=)
- $H_A \rightarrow$ Alternative: "there is an effect"
  - What researchers hope to support
  - Pick <, >, or ≠ based on what the researchers would like to show

- *Example:* Avocadopocalypse
  - Question: Is the price of an avocado in 2017 higher than the price of an avocado in the past?
  - In Words:
    - Null: The long run average price of an avocado in 2017 is equal to \$1.36
    - Alternative: The long run average price of an avocado in 2017 is greater than \$1.36
  - In Symbols:
    - $H_o: \mu = 1.36$
    - $H_A: \mu > 1.36$

- Write out your Hypotheses:
  - In Words
  - H0: The true average number of hours a college student sleeps a night is equal to 8.
  - HA: The true average number of hours a college student sleeps a night is less than 8.

  - In Symbols
  - H0: $\mu = 8$
  - HA: $\mu < 8$

YOUR TURN

It has been recommended US adults get 8 hours of sleep per night. A random sample of 32 college students was taken and was found that on average, they got 5.5 hours of sleep with a standard deviation of 0.71. Is the average less than the recommended 8 hours?

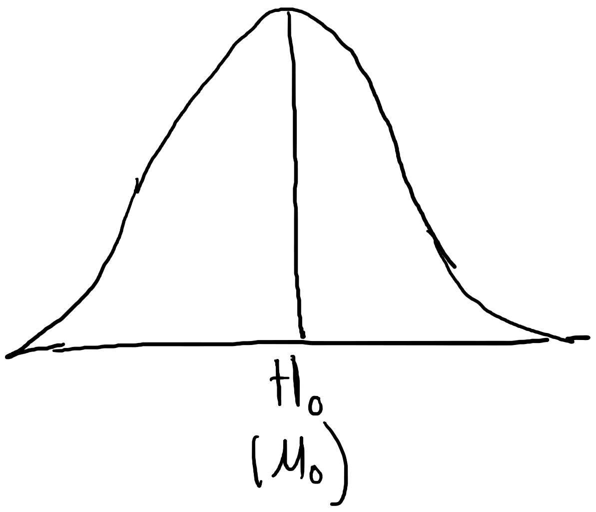
COMPARING OUR STATISTIC

Need something (null distribution) to compare to our statistic

- With mathematical theory we should be able to predict what the pattern of the null distribution will look like
  - Null Distribution: is the probability distribution of the test statistic when the null hypothesis is true.
- From the theoretical distribution we will then be able to find p-values and standardized statistics

NULL DISTRIBUTION

- **Null Distribution Commonalities**
  - Sample size (helps to determine variability)
  - Symmetric ‘bell’ shaped distribution
  - Centered at null hypothesis value
- **Can predict things about null…**
  - Where it will be centered
  - If its bell shaped (normally distributed)
  - Standard deviation of the null distribution



- Predictions used to calculate the standardized statistic

CENTRAL LIMIT THEOREM

- If the sample size (n) is large enough, the distribution of the sample means will be bell-shaped (or approximately normal), centered at the long run average (μ_o), with a standard deviation of

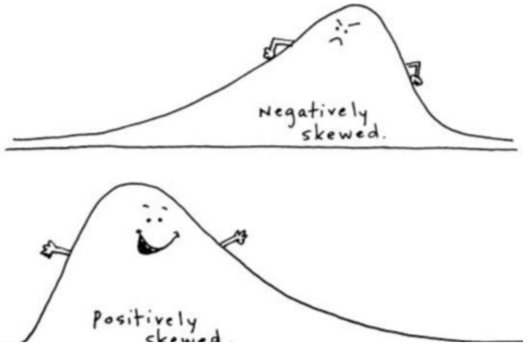
$$\text{standard deviation of the null} = \frac{\sigma}{\sqrt{n}}$$

- Going to be used as the ***standard deviation of the null*** in the denominator of the standardized statistic

  - $\mathbf{z} = \dfrac{statistic - hypothesized\ value}{\boldsymbol{standard\ deviation\ of\ the\ null}} = \dfrac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

CONDITIONS

- **If you want to use Central Limit Theorem need to meet certain conditions**
  - Need a <u>sample size large enough</u>, in order to say the null will follow the Normal Distribution
    - $n \geq 30$
    - Sampling distribution is symmetric
  - Assume each observation is independent of each other
    - $n \leq 10\%$ of population

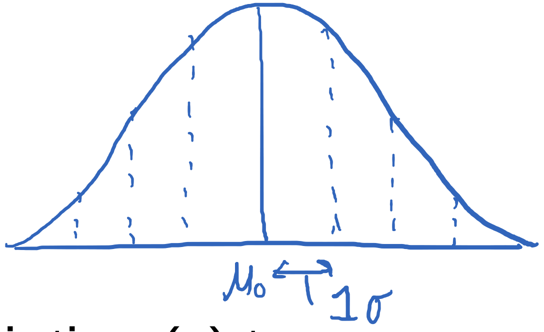


- **Doesn’t always work**
  - If the sample size is small …
    - null distribution will **not** be approximately bell-shaped and the resulting standardized statistics are **not accurate**

STANDARDIZED STATISTIC

- Standardized Statistic- calculates how many standard deviation from the null value that the statistics falls

- For quantitative variables:
  - $z = \ Standardized\ statistic\ = \frac{sample\ mean\ - null\ value}{standard\ deviation\ of\ sample\ mean} = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$

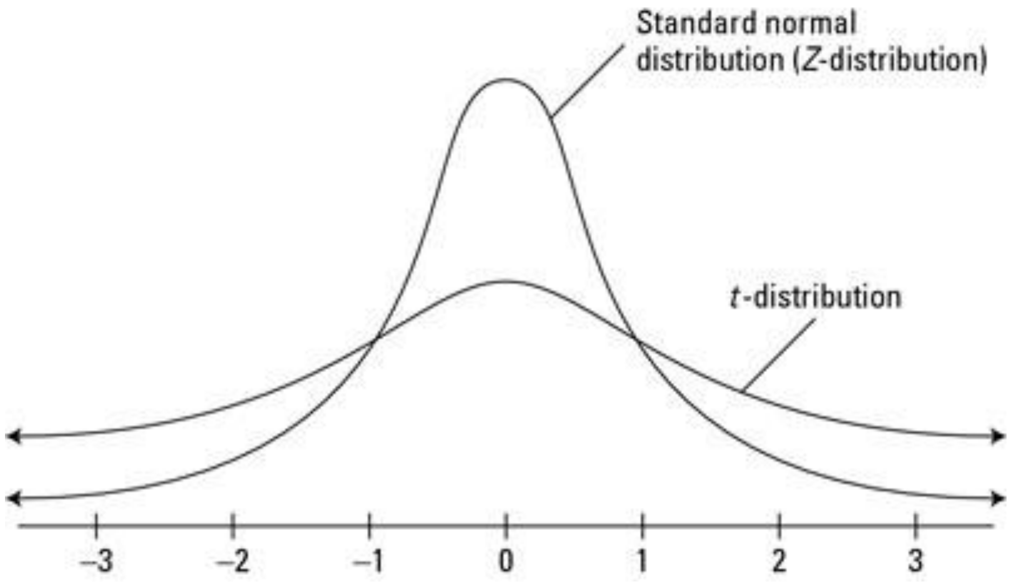


- Since σ is not usually known, we use the sample standard deviation (s) to approximate the standard deviation of sample mean
  - This approximation is called the **standard error** (SE) of sample mean

  - t = standardized statistic = $\frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$

T STATISTIC VS. Z STATISTIC

- Like with the normal approximation, need a reference distribution
  - With categorical, using z (standard normal) distribution
  - Using standard error instead of standard deviation results in no longer using normal distribution
  - Different reference distribution: t distribution



SUMMARY OF CALCULATIONS

- Validity conditions:
  - **Sample size ≥ 30 and sampling distribution is not strongly skewed**
- Standard Error = $\frac{s}{\sqrt{n}}$ where s is our sample standard deviation
- Standardized Statistic = $\frac{\bar{x}-\mu}{SE} = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$

AVOCADOPAOCALYPSE

μ - The long run average price of an avocado in 2017

$$H_o: \mu = 1.36$$
$$H_A: \mu > 1.36$$

To test the avocadopaocalypse, we collected a random sample of 94 avocados from 2017. The average price of an avocado was $1.44 with a sample standard deviation of 0.41

- Sample Size:
  - n = 94
- Can the theory-based approach be used? Justify your answer.
  - Yes. More than 30 in the sample
- Statistic:
  - $\bar{x} = 1.44$
- Sample Standard Deviation:
  - s = 0.41
- Standard error of the null

$$SE = \frac{s}{\sqrt{n}} = \frac{0.41}{\sqrt{94}} = 0.042$$

Standardized statistic

$$t = \frac{\bar{x} - \mu}{SE} = \frac{1.44 - 1.36}{0.042} = 1.9$$

YOUR TURN

It has been recommended US adults get 8 hours of sleep per night. A random sample of 32 college students was taken and was found that on average, they got 5.5 hours of sleep with a standard deviation of 0.71. Is the average less than the recommended 8 hours?

- Sample Size:
  - n= 32
- Can the theory-based approach be used? Justify your answer.
  - **Yes. More than 30 observations (32) (*and the distribution is not strongly skewed*)**
- Statistic:
  - $\bar{x} = 5.5$
- Sample Standard Deviation:
  - s = 0.71
- Standard error of the null

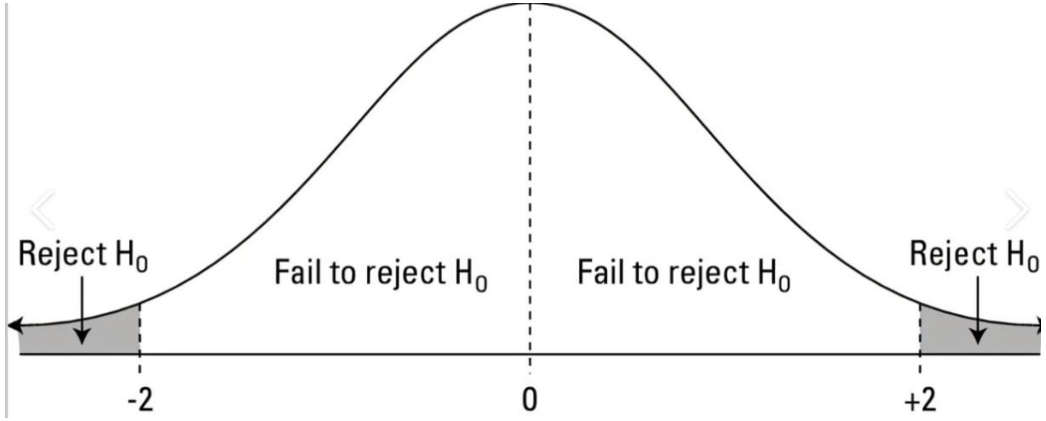
$$SE = \frac{s}{\sqrt{n}} = \frac{0.71}{\sqrt{32}} = 0.126$$

Standardized statistic

$$t = \frac{\bar{x} - \mu}{SE} = \frac{5.5 - 8}{0.126} = \text{-}19.92$$

ONE QUANTITATIVE VARIABLE CONCLUSIONS

- SS "outside the 2's" → Reject the Null
  - We have evidence to conclude that the true average ___________ is (not equal to/ less than/ greater than) _________.
- SS "between the 2's" → Fail to Reject the Null
  - We do not have evidence to conclude that the true average ___________ is (not equal to/ less than/ greater than) _________.



Standardized statistic

$$t = 1.9$$

Decision and why?

- Fail To Reject -> Between the 2's

Conclusion

- We do not have evidence the true average price of an avocado in 2017 is less than $1.36

AVOCADOPAOCALYPSE

μ - The long run average price of an avocado in 2017

$$H_o: \mu = 1.36$$
$$H_A: \mu > 1.36$$

To test the avocadopaocalypse, we collected a random sample of 94 avocados from 2017. The average price of an avocado was $1.44 with a sample standard deviation of 0.41

Standardized statistic

$$t = -19.92$$

Decision and why?

- Reject- > Outside the 2's

Conclusion

- We have evidence the true average number of hours a college student sleeps a night is less than 8

YOUR TURN

It has been recommended US adults get 8 hours of sleep per night. A random sample of 32 college students was taken and was found that on average, they got 5.5 hours of sleep with a standard deviation of 0.71. Is the average less than the recommended 8 hours?

PRACTICE (TIME PERMITTING)

Suppose a company wanted to test if the machine that is filling the milk bottles it aims to sell, is filling the bottles the correct amount. The mean amount of milk in each bottle is supposed to be 30 Oz. To check if the machine is operating properly, 44 filled bottles were randomly sampled. The average amount of milk in each bottle is 32.2 with a standard deviation of 3.4 and the distribution is not strongly skewed. Suppose they want to know if the average amount of milk in each bottle is more than 30 oz.

- Identify the following:
  - Population
  - Parameter
  - Sample
  - Statistic
  - Sample Size
- Set up the correct Null and Alternative Hypothesis
- Determine if the assumptions for our test is met. If so, calculate the standardized statistic.
- Based on the standardized statistic, what is your conclusion?

Population: milk bottles

Parameter (μ): The true average amount of milk in a bottle.

Sample: 44 milk bottles, $n = 44$

Statistic ($\bar{x}$): 32.2

$H_0: \mu = 30$ vs. $H_A: \mu > 30$

Yes ($44 > 30$ & distribution Not skewed)

$$\frac{32.2 - 30}{\frac{3.4}{\sqrt{44}}} = 4.29 \text{ (R)}$$

We have evidence the true average amount of milk in a bottle is more than 30 oz.

SOLUTIONS

Suppose a company wanted to test if the machine that is filling the milk bottles it aims to sell, is filling the bottles the correct amount. The mean amount of milk in each bottle is supposed to be 30 Oz. To check if the machine is operating properly, 44 filled bottles were randomly sampled. The average amount of milk in each bottle is 32.2 with a standard deviation of 3.4 and the distribution is not strongly skewed. Suppose they want to know if the average amount of milk in each bottle is more than 30 oz.

REVIEW

- Parameter vs Statistics
- Single Quantitative Test
  - Hypotheses (μ)
    - Null: "random chance alone"
    - Alternative: "there is an effect"
  - Take a sample -> obtain statistic
  - Compare to Null Distribution
  - $SS = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$
    - Is it outside or inside the 2's?