

1. Quantitative and Qualitative research methods

2. Descriptive and Inferential statistics

3. Statistical significance; p-values and α levels

By Somphop Sukjaitham, Katarina Walther and Homai Faridi



edutech
SAARLAND UNIVERSITY



1. Qualitative and Quantitative research methods

- ❑ Overview
- ❑ Definition
- ❑ Matching Questions and Methods
- ❑ Advantages and Limitations

edutech
SAARLAND UNIVERSITY

OVERVIEW

Qualitative research

- Understand relations
- Describe phenomena
- Discover meaning

- Focus groups
- Observations
- Interviews with open end questions
- Words

- In depth insight in sample
- Develop hypothesis

Quantitative research

- Explain relations
- Discover facts
- Predict further outcomes

- Experiments
- Questionnaires/surveys with closed questions
- Numbers

- Projections to population
- Test hypothesis

edutech

DEFINITION

Qualitative research

Quantitative research

A type of research that is “explaining phenomena by collecting numerical data that are analyzed using mathematically based methods (in particular statistics)” (Creswell 1994).

“Multimethod in focus, involving an interpretive, naturalistic approach to its subject matter” (Denzin and Lincoln, 2005,).

edutech

WHICH METHOD(S) WOULD YOU CHOSE AND WHY?

Qualitative or Quantitative?

What strategies do tutors of EduTech use to deal with sleep deficit caused by work overload?

H1: The choice of a dark colour schema in a PowerPoint presentation does have a conductive effect on recipients' attention level.

"I wonder how former bullies feel in emotional challenging situations when working in a group."

Methods

- Focus groups
- Experiments
- Observations
- Questionnaires/surveys with closed questions
- Interviews with open end
- Questionnaires with open questions

ADVANTAGES AND DISADVANTAGES

Qualitative research

Quantitative research

"People can get quite passionate about which of these methods is best, which is a bit silly because they are **complementary**, **not competing**, approaches [...]" (Field, 2013).

The Advantages or Disadvantages of the methods depending on your research question!

edutech

Qualitative research

Quantitative research

Qualitative research		Quantitative research
Smaller	Number of participants	Higher
Lower → inconsistent data capture but flexible for unexpected beneficial variables	Control of extraneous variables	Higher → consistent data capture but might miss unexpected beneficial variables
Flexible	Design	Standardised
Hard	Replicability	Easy
Lower → need further quantitative research	Statistical representative	Easy

edutech



2 Descriptive and Inferential statistics

- Definition
- Goals

edutech
SAARLAND UNIVERSITY

DEFINITIONS

Descriptive Statistics: values that describe the characteristics of a sample or a population

Statistics: collection, organisation and interpretation of data

Inferential Statistics: values that infer results of a sample to the population from which the sample is drawn

GOALS

Descriptive Statistics

- Organizing, summarizing data
- Visualization of data in form of
  - Charts
  - Graphs

Inferential Statistics

- Determine the probability which we can make generalization from a sample to a population
- Test hypothesis

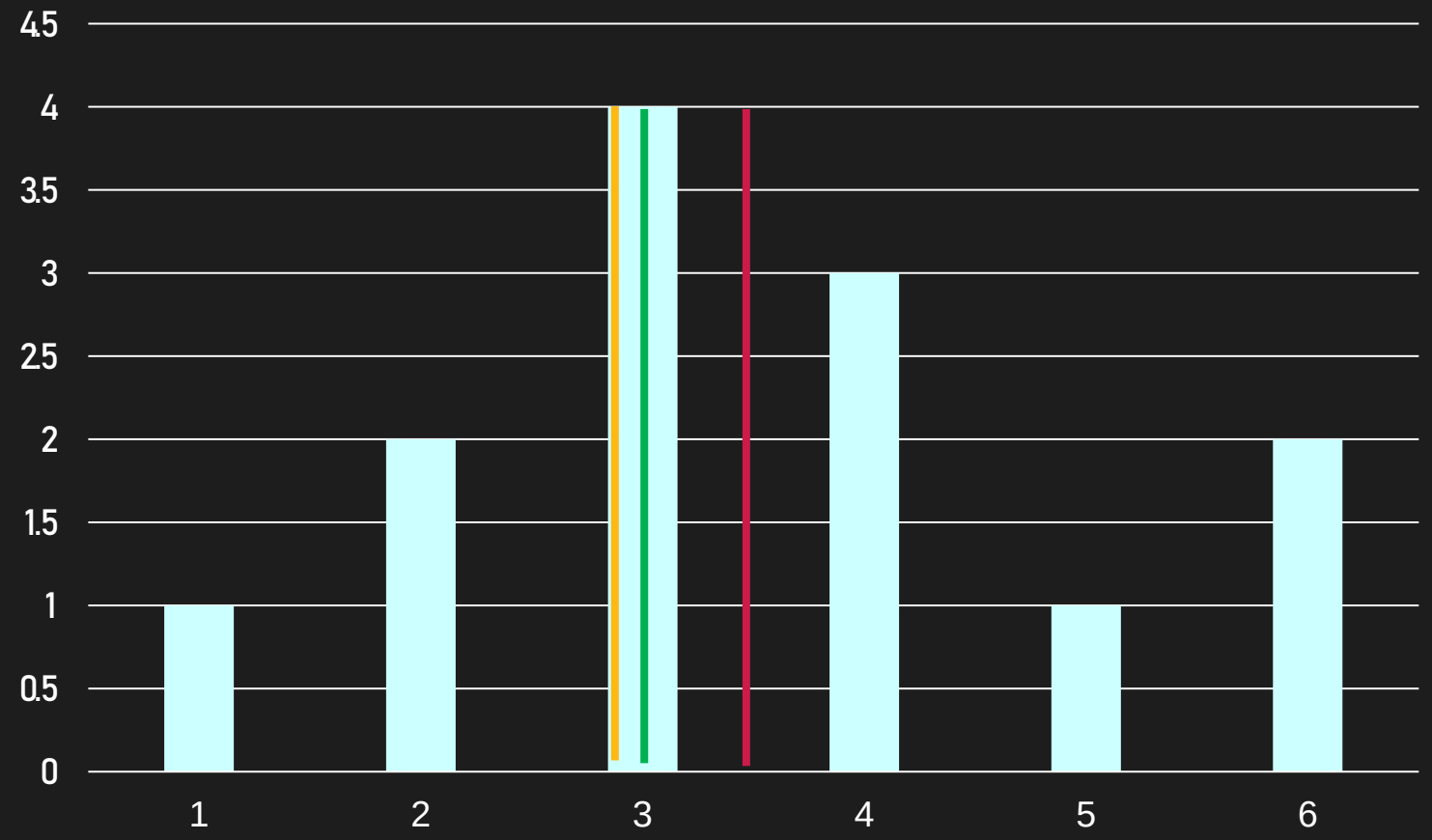
METHODOLOGIES

DESCRIPTIVE STATISTICS

- Measure of central tendency
  - Mean (average)
  - Median (midpoint)
  - Mode (most frequently occurring number)

- Measure of Variability
  - Range (difference between largest and smallest number)
  - Variance (average of squared numbers from the mean)
  - Standard deviation (how much variation exists relative to mean)

Measure of Tendency

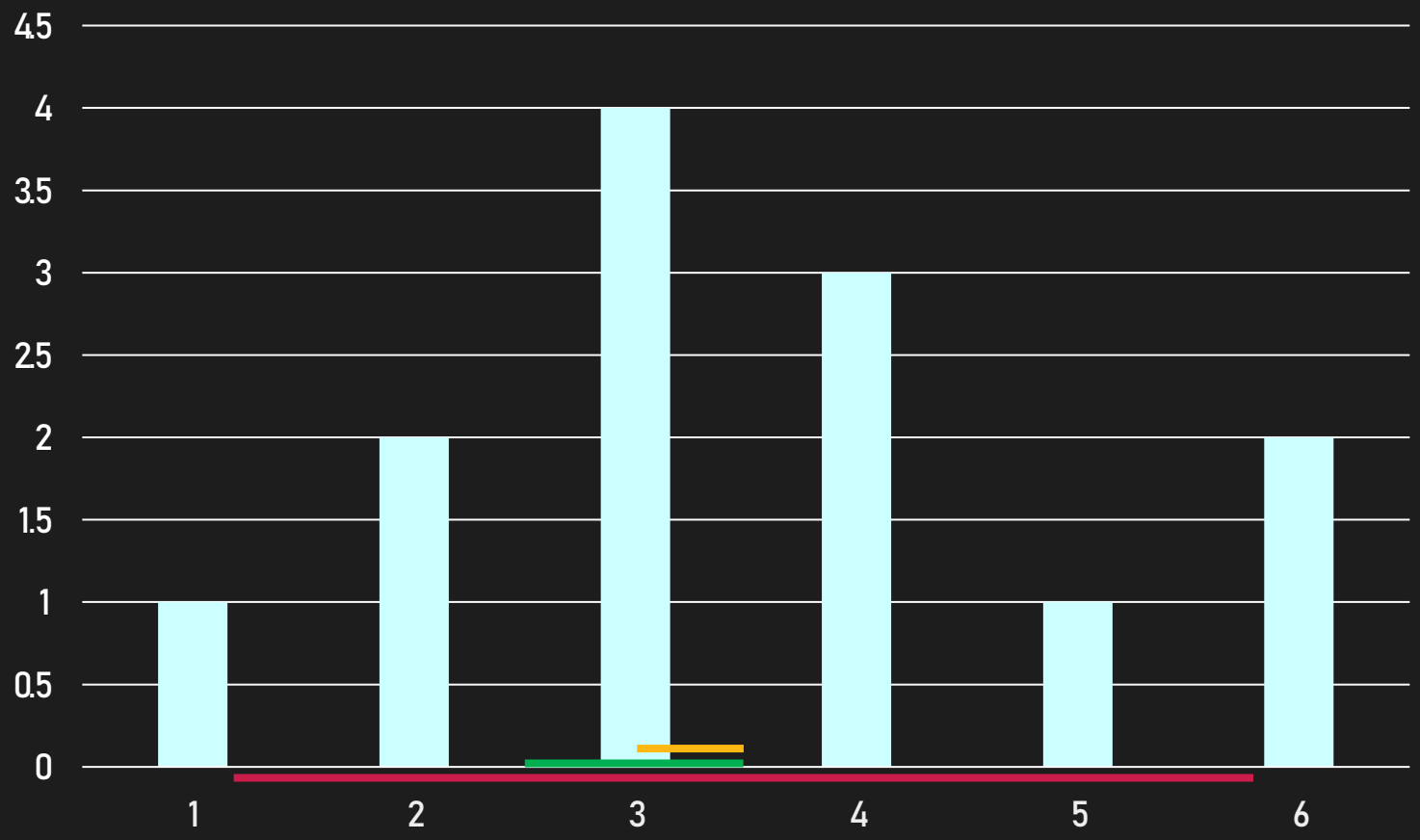


Mean = $\frac{46}{13}$ = 3.53

Median = 3

Mode = 3

Measure of Variability



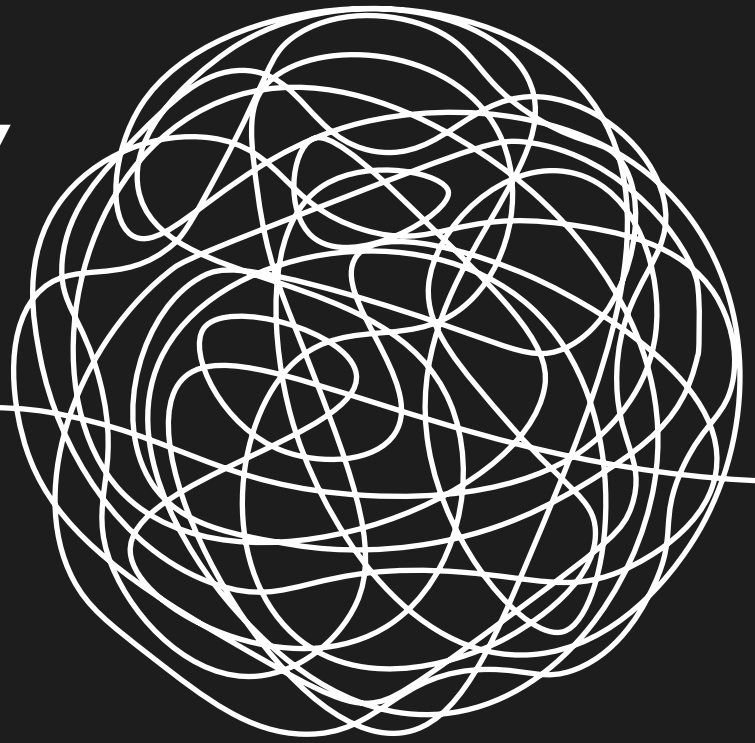
Range = 6 – 1 = 5

Variance = 1.425

Standard deviation = $\sqrt{1.425}$ = 1.194

One sample t test

(1 COMMON) METHODOLOGY INFERENTIAL STATISTICS



Will be explained later on – stay tuned

Limitations of the Statistics

Descriptive

Only allow you to make summations about the people or objects that you have actually measured

Inferential

Always a degree of **uncertainty** even in terms of the accuracy of the results

CONCLUSION

	Descriptive	Inferential
Goals	Organize, Simplify, Summarize and Describe	Generalize, Predict and Test Hypothesis
Data	Describe data which is already known	Make conclusion that is beyond the data available
Focus group's Size	Entire population	Sample from population
Certainty	Summary is certain	Summary is uncertain
Focus group's Properties	Properties of population are known as parameters	Properties of samples are known as statistics
Application	Cannot be extrapolated to other groups of data	Can be applied to a larger population of data and representative sample data
Accuracy	100% accurate	Not 100% accurate
User Assumption	Don't require user assumptions	(Some) Require user assumptions
Final Results	Shown as charts, table and graphs	Shown as the probability scores

What type of statistics ?

Descriptive or inferential ?

Descriptive	01	We would like to know how many freshmen students in EduTech experience high stress levels.
Inferential	02	We would like to know if the stress level among all freshmen in the university differs by gender
Inferential	03	We would like to know whether there is a relationship between freshmen students' study strategies and their academic results.
Descriptive	04	We would like to know what study strategies are used by first year EduTech students.

edutech



3. Statistical Significance

- ❑ t-Test
- ❑ alpha levels
- ❑ p-values

edutech
SAARLAND UNIVERSITY

What is a null hypothesis, directional and non-directional hypothesis ?

- **Null hypothesis / H0 (stating no difference, no relationship)**

  → $M_C = M_W$ (The mean performance after drinking coffee is the same as the mean performance after drinking water)

- **Alternative hypothesis / H1 (stating a difference, a relationship):**

← $M_C > M_W$ (The mean performance after drinking coffee is higher than the mean performance after drinking water)

directional hypothesis

↓ $M_C < M_W$ (The mean performance after drinking coffee is lower than the mean performance after drinking water)

directional hypothesis

→ Or: $M_C \neq M_W$ (The mean performance after drinking coffee is different from the mean performance after drinking water

non-directional hypothesis

edutech

What is an alpha level? How it related to a null hypothesis?

Alpha and Beta errors recap and aliases

Alpha error = false positive = Type I error: Rejection of a true H0

Beta error = false negative = Type II error: Acceptance of a false H0

STATISTICAL SIGNIFICANCE

p-value & alpha-level

The significance level, also denoted as alpha or α, is the probability of rejecting the null hypothesis when it is true.

P-values are the probability of obtaining an effect at least as extreme as the one in your sample data, assuming the truth of the null hypothesis.

You feel like you ***deserve a treat*** today!

You go out and buy a ***bag of gummy bears***

Source: https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.candywarehouse.com%2Fharibo-gold-bears-gummy-bears-candy-5lb-bag&psig=AOvVaw2enok8YUCFDYAcqLvvBhsT&ust=1604911916990000&source=images&cd=vfe&ved=0CA0QjhxqFwoTCKiDxInJ8uwCFQAAAAAdAAAAABAD

And you ask yourself:

Are there really 50 red gummy bears in a bag on average?

For scientific research, please consider increasing your **sample size**

Source: https://www.artstar.com/products/kendyll-hillegas-gummy-bear-red (red), https://www.pinterest.de/pin/337347828311484714/ (yellow)

edutech

STATISTICAL SIGNIFICANCE

One Sample t-Test

To indicate whether your sample is statistical significance, you perform a one sample t-test

The test statistic for a one sample t test is denoted t, which is calculated using the following formula:

$$t = \frac{\bar{x} - \mu}{\frac{SD}{\sqrt{n}}}$$

μ = Proposed constant for the population mean
$\bar{x}$ = Sample mean
n = Sample size (i.e., number of observations)
SD = Sample standard deviation

edutech

SPSS is your best friend

P-value and α level

How to calculate the example of the presentation in SPSS

Problem Statement

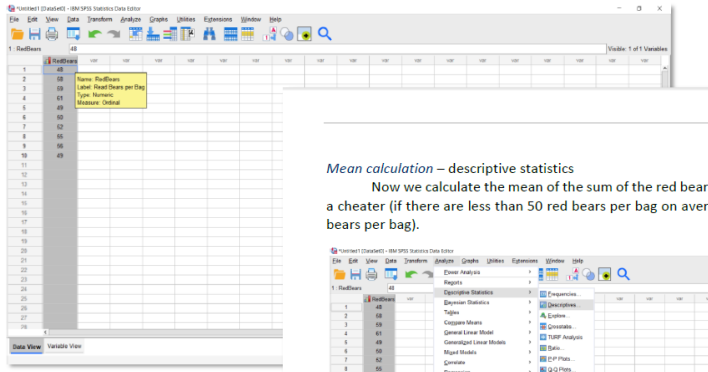
The brand says their 100 bear bags have an average of 50 red bears each. We, as critical scientists, formulate the following hypotheses:

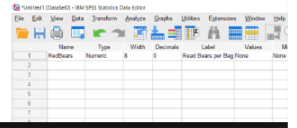
H0: The average of red bears in each bag is 50.
in other words: H0: $\mu = 50$

H1: The average of red bears in each bag is less than 50.
in other words: H1: $\mu < 50$

Data collection

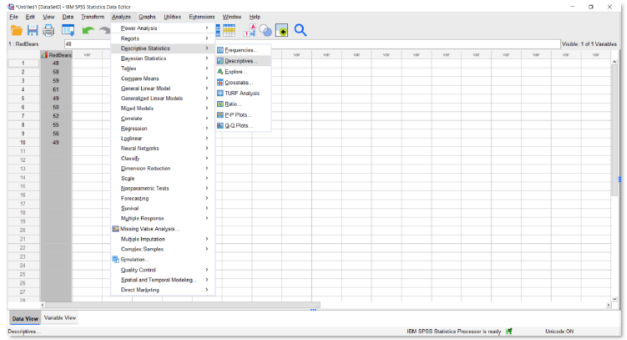
We buy 10 bags and count the red bears of each bag. Now we open SPSS and put the data (the number of red bears per bag) as a variable in and adjust its characteristics.

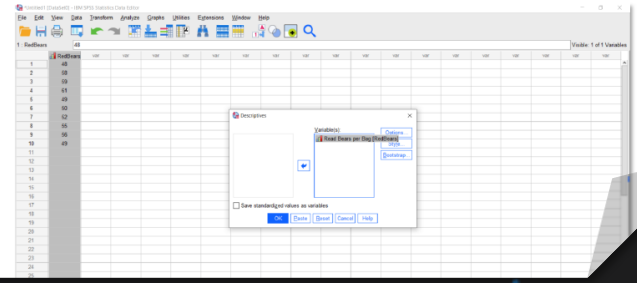




Mean calculation – descriptive statistics

Now we calculate the mean of the sum of the red bears per bag. This would allow us to call the seller a cheater (if there are less than 50 red bears per bag on average) or not (if there are more or exact 50 red bears per bag).





edutech

Step 1: Formulating your Conjecture

The average of red gummy bears in each bag is 50.

In other words: H_0: μ = 50
This is called a **Null Hypothesis** (H_0)

Rejection is a stronger proof,
because accepting the H_0 doesn't prove
it is true, only that you couldn't reject it!

Step 2: Determine Significance Level (α-Level)

Often set to α = 5%

Others are also common, like 1% or 0.1%
Depends on the application, and also the severity of the two errors one can make

Always try to set your α-Level **before experimenting** to not adjust your α-Level to fit your results

For α = 5%, this means in 1 of 20 cases you would **falsely reject** the Null hypothesis (known as **Type I error**) based on your observation

edu tech

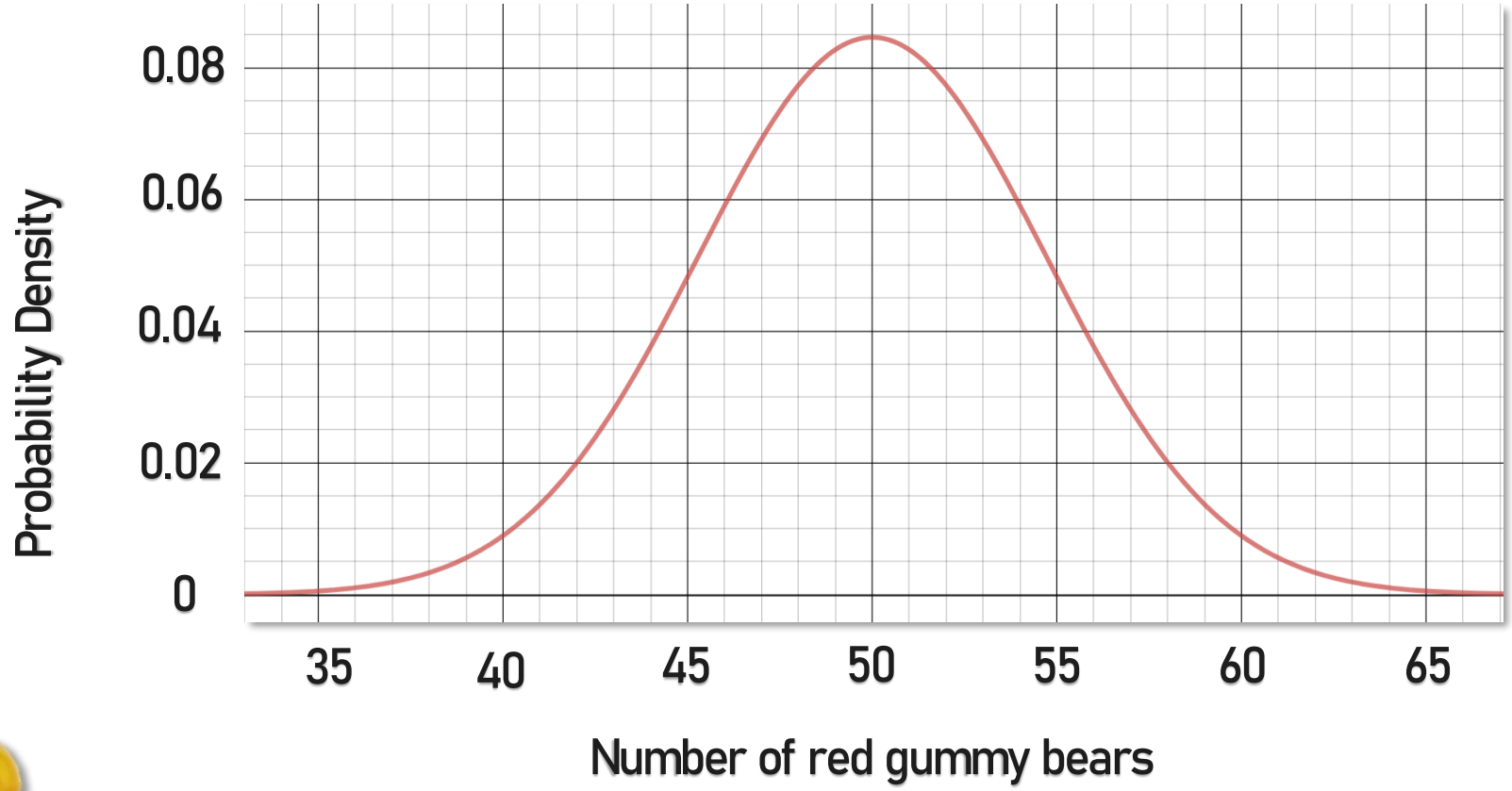
Step 3: Conduct the Experiment

You count the number of gummy bears in the 10 bags (n) and use SPSS to get a mean ($\bar{x}$) of 53.7 red gummy bears with a standard deviation (SD) of 4.715.
The manufacturer claims a population mean (μ) of 50*.

$$t = \frac{\bar{x} - \mu}{\frac{SD}{\sqrt{n}}} = \frac{53.7 - 50}{\frac{4.715}{\sqrt{10}}} = 2.481$$

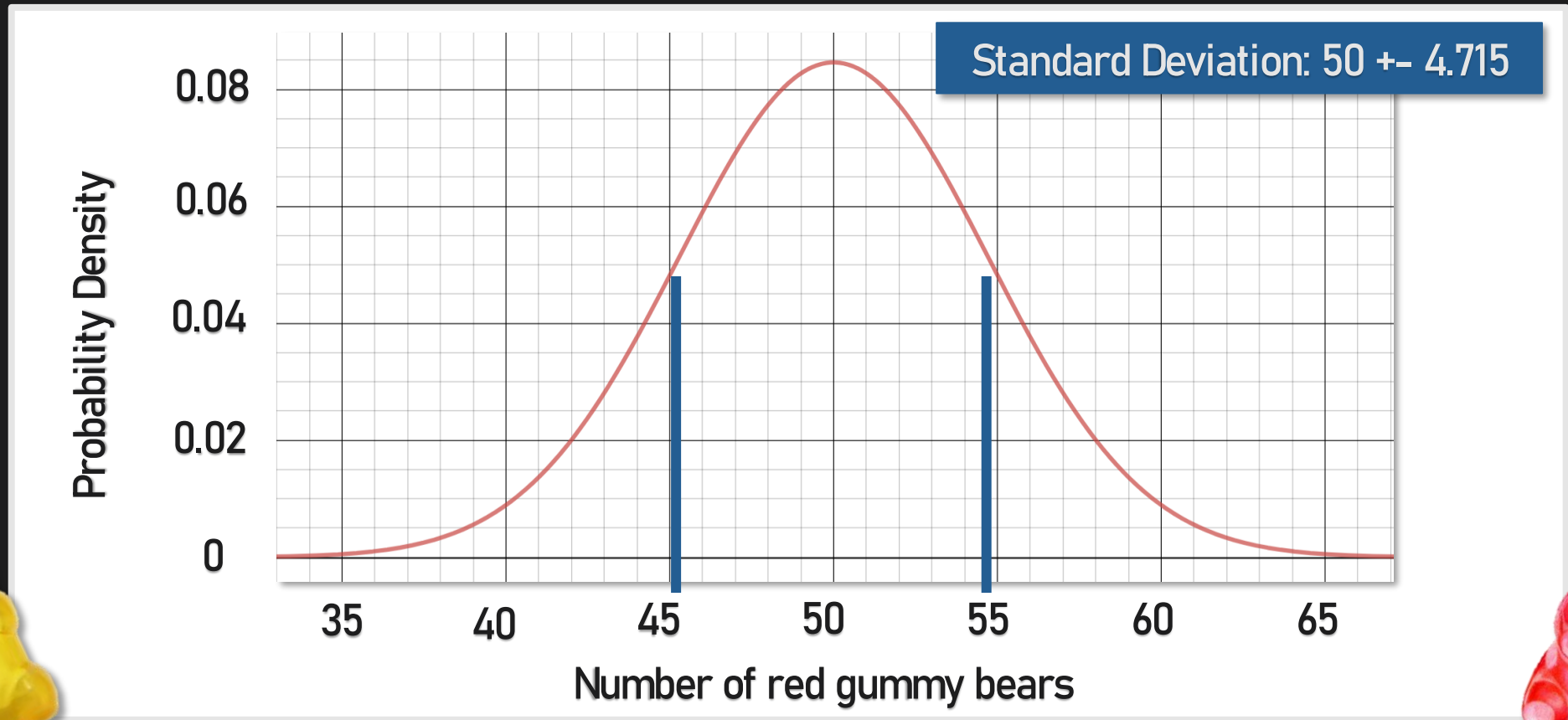
*for details check the cheat sheet provided

Assumed **normal distribution** of gummy bears of our sample with 10 bags, assuming the Null Hypothesis to be true : **μ = 50**, sample standard deviation SD = 4.715

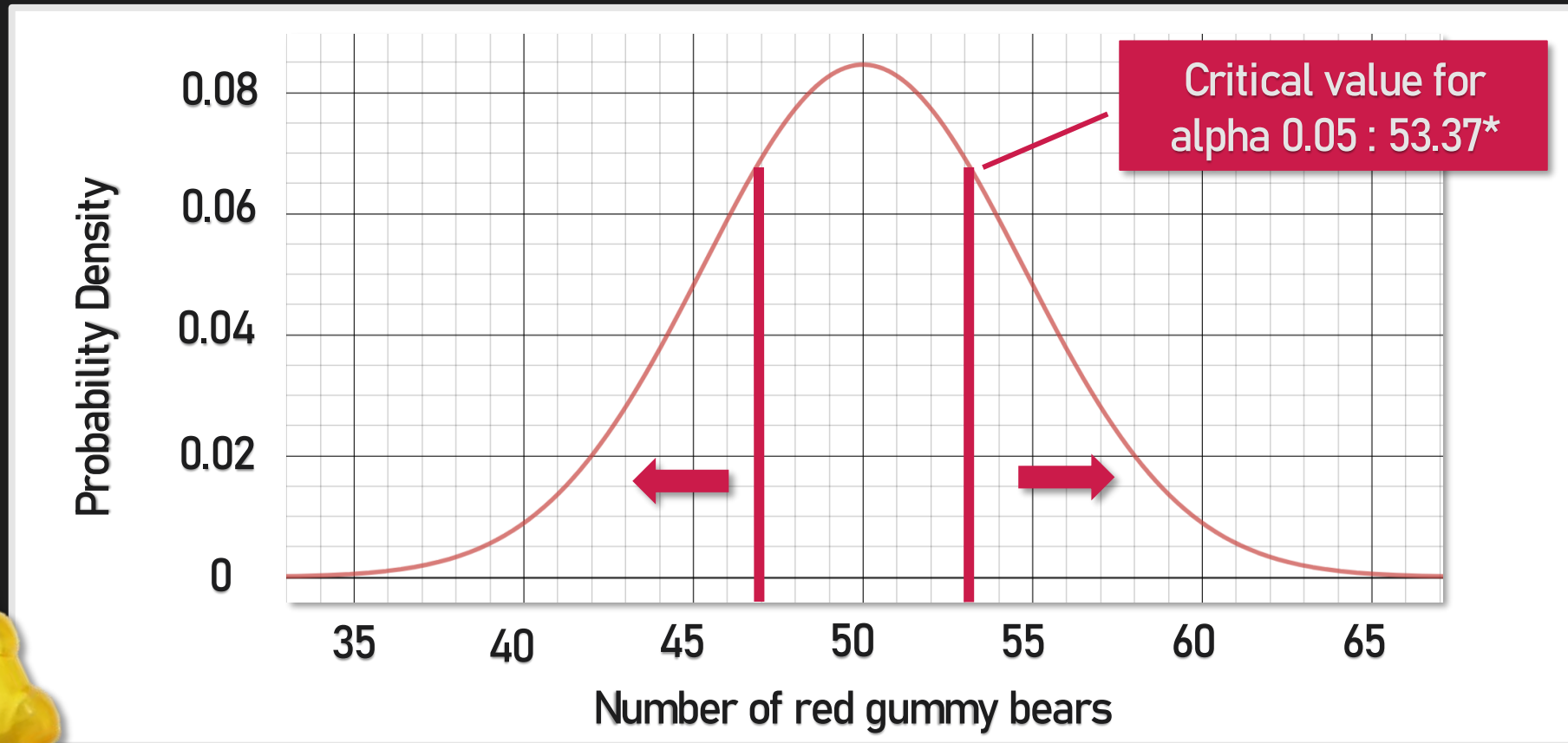


Graph Source: https://www.desmos.com/calculator/0x3rpqtgrx

Excursion: Drawing in the standard deviation



Marking critical value for alpha 0.05



*For calculating the critical values, you use the "t-table" for 5% (two-tailed) and calculate $\bar{x}$ using the equation for the t-value

edutech

Step 3 : Conduct the Experiment

With t = 2.481 and DF = 9*, you look up the p-value for two-tailed-testing in a table or use SPSS

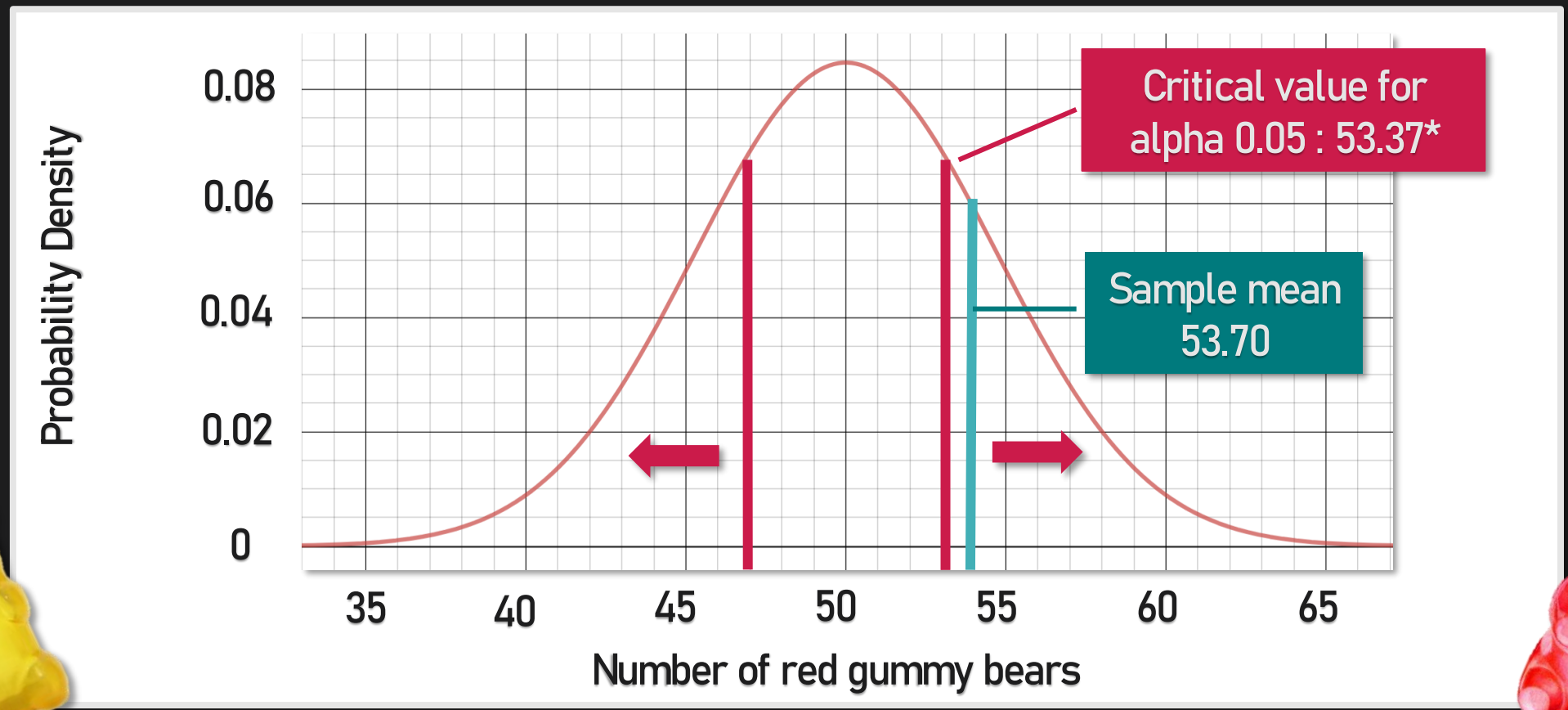
*DF = n – 1 = 10 – 1 = 9

The p-value denotes the probability for a test result to be as extreme or more extreme (on both sides / tails for two-tailed-testing) than the result observed

How high is the chance of obtaining our test result with 53.7 or more red gummy bears assuming the Null Hypothesis to be true?
For two-tailed testing, the **p-value** is:

p = 0.035

Marking the sample mean



Step 4: Assess the Null Hypothesis

Comparing the p-value (0.035) to the α-Level (0.05), you find:

$$p < \alpha$$

This means, the observed result is statistically significant and you reject the Null Hypothesis

To Summarize

$p \leq \alpha$

The observed result is *statistically significant* and we *reject* the Null Hypothesis (H_0)

$p > \alpha$

The observed result is *not statistically significant* and we *fail to reject* the Null Hypothesis (H_0)

In performing a hypothesis test the null hypothesis is defined as $\mu = 6.9$.
It can be assumed that the population is normally distributed and $\alpha = 0.05$.
After running our significance test, we get a p-value of 0.156.

01

Reject H0

02

Fail to reject H0


ANY
QUESTION???
meme-arsenal.ru


THANK YOU
FOR YOUR ATTENTION
meme-arsenal.ru

REFERENCES

Creswell, J.W., Research Design: Qualitative & Quantitative Approaches, London: SAGE Publications, 1

Cumming, G. (2014), "The New Statistics: Why and How," Psychological Science, 25, 7–29. DOI: 10.1177/0956797613504966.

Denzin, N. K., & Lincoln, Y. S. (2005). *Introduction: The Discipline and Practice of Qualitative Research.* In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (p. 1–32).

Edgeworth, F. Y. (1885), "Methods of Statistics," *Journal of the Statistical Society of London*, Jubilee Volume, 181–217. [Google Scholar]

Field, A. (2013). Discovering Statistic using IBM SPSS Statistics. London. Sage

Jim Frost. (2018). *Introduction to Statistics: An Intuitive Guide. https://statisticsbyjim.com/basics/descriptive-inferential-statistics/*

Punch, K. (1998). Introduction to Social Research: Quantitative and Qualitative Approaches. London: Sage

The sources given in the ERM1 course

edutech