Empirical Research Methods 1

Research methods: quantitative and qualitative, descriptive and inferential statistics

Miguel Rejón,

Room 1.14

m.rejon@edutech.uni-saarland.de

edutech

SAARLAND UNIVERSITY

Before we start 1/2: Info and tools

- Presentation will cover additional content
- Similar as last week, presentation + demo (volunteer or direct participation)
- MS Teams for comments, questions, etc
- SPSS !!!

Before we start 2/2: advice for your presentations

- Send the 2nd draft to the tutors **and me**
- When >1 presenter, assign roles. Maybe someone who shares the links in the chat.Ideally equal presenting times
- For the interactive stuff, ensure it works flawlessly. Remember some students have a class before this one
- Long answers can be said instead of written

Agenda

- Data distribution, normal distribution
- P-values
- (The additional topic) Calculating one-sample t-test and its assumptions with SPSS
- Q&A

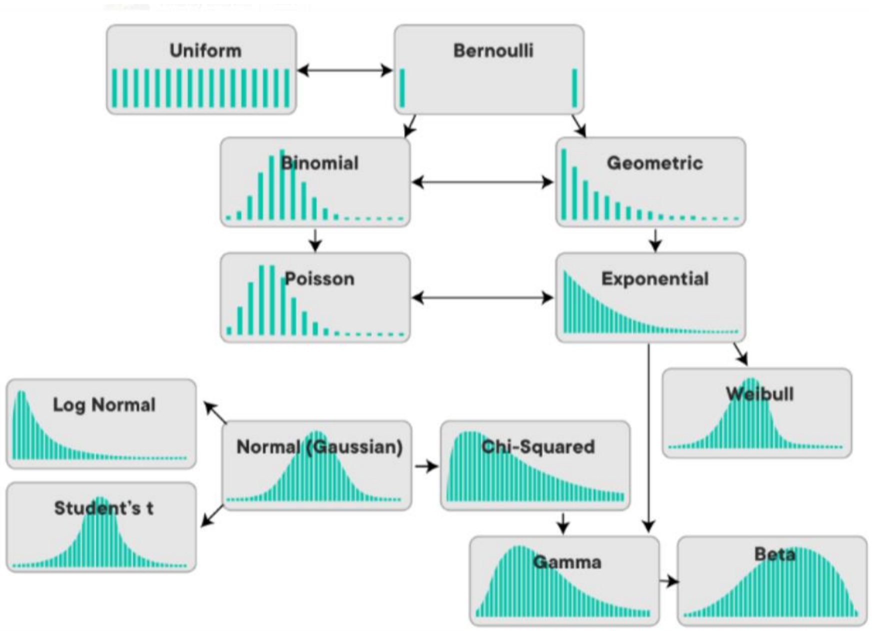
Data distribution

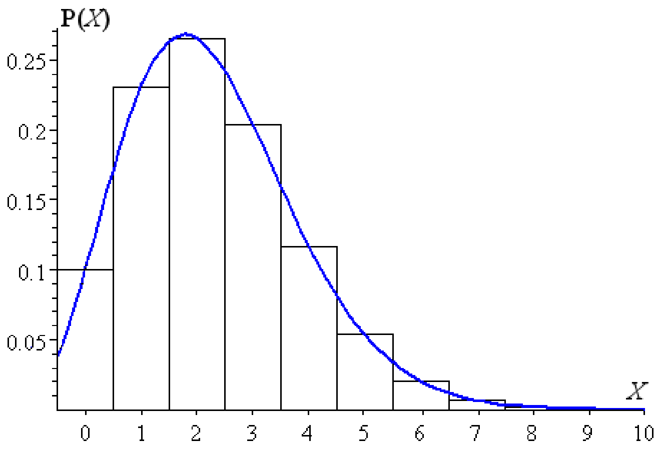
- “A **data distribution** is a function or a listing which shows all the possible values (or intervals) of the data.” (Stephanie from Statistics How To, 2017)
- “A **distribution** is a way to see all the <u>probable</u> values of our data.” (Me, 2020)

Reference: https://www.statisticshowto.com/data-distribution

Data distribution

There are different types of distributions which can be represented on different ways

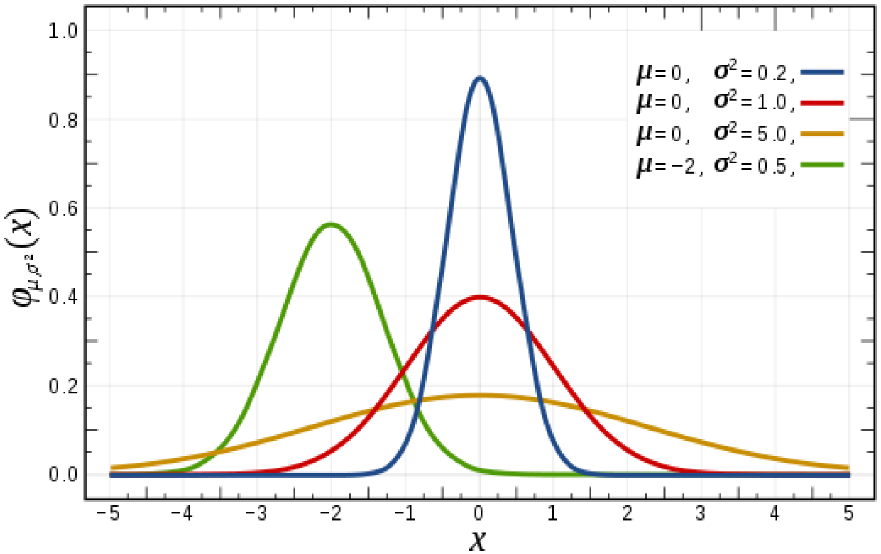


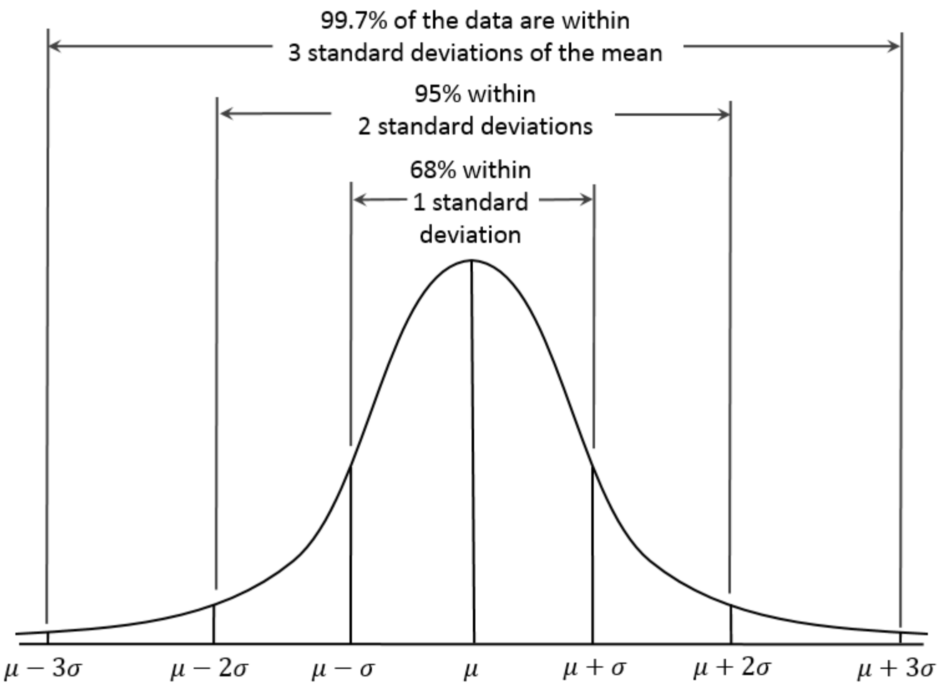


Images Reference: https://medium.com/mytake/understanding-different-types-of-distributions-you-will-encounter-as-a-data-scientist-27ea4c375eec

Normal distribution

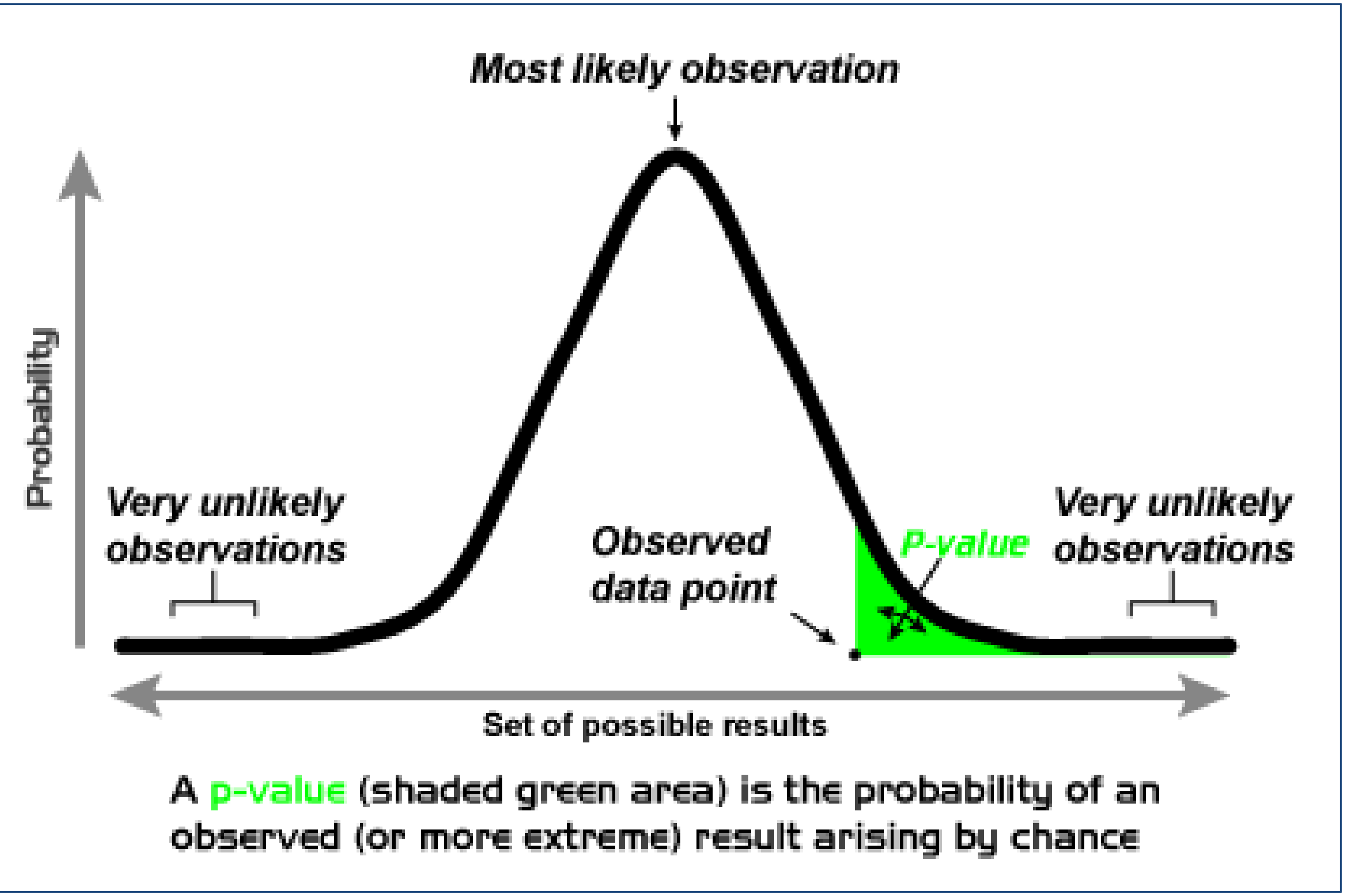
- A.k.a. Gaussian or Gauss or Laplace–Gauss distribution or **Bell curve**
- Data that produces a normal distribution can be called “normal data”
- Normality is a basic assumption/ requirement for parametric tests
- **We want normal data because we want parametric tests!**

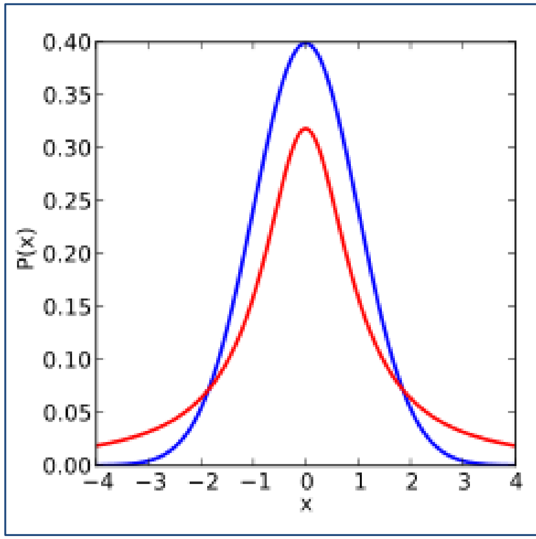


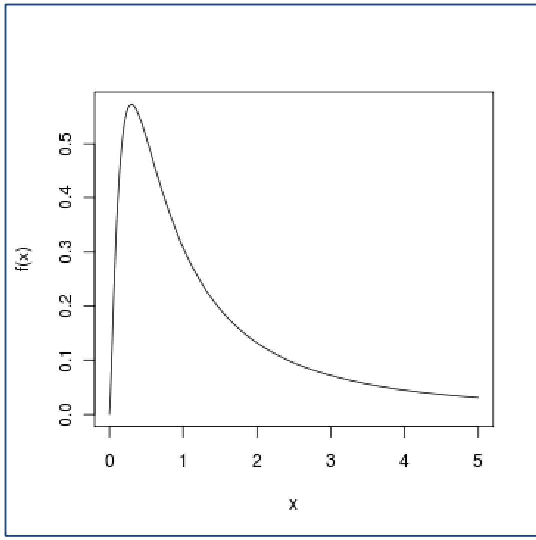


(**Normal distribution.** In *Wikipedia*. Retrieved December 09, 2019, from https://en.wikipedia.org/wiki/Normal_distribution)

p-value







https://www.thoughtco.com/definition-of-p-value-1148041
https://en.wikipedia.org/wiki/Student%27s_t-distribution
http://www.r-tutor.com/elementary-statistics/probability-distributions/f-distribution

p-value

- p = result from a statistical test. “Probability”
- The smaller it is, the less likely our data is under the assumption of the H0
- But: When are they “unlikely enough“ (→ accepting H1)?

→ α-level (or, level of significance):

  - usually 5% = .05

- → if $p \leq .05$: statistically significant result and rejection of H0 (which implies acceptance of H1 but <u>not</u>: proof)

p-value: summary

- A p-value is a measure of the probability that an observed difference could have occurred just by random chance. That's why we want p-values to be as small as possible.
- A p-value indicates evidence against the null hypothesis, indicating the probability the null is correct (and the results are random).
- A p-value of ≤ 0.05 is statistically significant. It indicates strong evidence against the null hypothesis, as there is less than a 5% probability the null is correct (and the results are random). Therefore, we reject the null hypothesis, and accept the alternative hypothesis.
- However, this does not mean that there is a 95% probability that the research hypothesis is true. The p-value is conditional upon the null hypothesis being true is unrelated to the truth or falsity of the research hypothesis.

https://www.investopedia.com/terms/p/p-value.asp
https://www.simplypsychology.org/p-value.html

Example

Process analysis was run to help us interpret the results with reference how learning processes were influenced. It showed a significant main effect for argumentation script on argument quality $F(1,77) = 4.7$, $p = .033$, ${\eta_p}^2 = .06$, and argument structure $F(1,77) = 10.46$, $p = .002$, ${\eta_p}^2 = .12$, meaning that conditions with the argumentation script created arguments of better quality and structure. There were no significant variance effects on argument

Tsovaltzi, D., Judele, R., Puhl, T., & Weinberger, A. (2015). Scripts, individual preparation and group awareness support in the service of learning in Facebook: How does CSCL compare to social networking sites? *Computers in Human Behavior*, *53*, 577–592. http://doi.org/10.1016/j.chb.2015.04.067

Exercise: Significant or not?

- p = .50 
- p = .03 
- p = .27 
- p = .006 
- p = .051 

http://www.academiaobscura.com/still-not-significant/

Significant…?

- appeared to be marginally significant (p<0.10)
- approached acceptable levels of statistical significance (p=0.054)
- approached but did not quite achieve significance (p>0.05)
- approached but fell short of significance (p=0.07)
- approached conventional levels of significance (p<0.10)
- approached near significance (p=0.06)
- approached our criterion of significance (p>0.08)
- approached significant (p=0.11)
- approached the borderline of significance (p=0.07)
- approached the level of significance (p=0.09)
- approached trend levels of significance (p0.05)
- approached, but did reach, significance (p=0.065)
- approaches but fails to achieve a customary level of statistical significance (p=0.154)
- approaches statistical significance (p>0.06)
- approaching a level of significance (p=0.089)
- approaching an acceptable significance level (p=0.056)
- approaching borderline significance (p=0.08)
- approaching borderline statistical significance (p=0.07)
- approaching but not reaching significance (p=0.53)
- approaching clinical significance (p=0.07)
- approaching close to significance (p<0.1)
- approaching conventional significance levels (p=0.06)
- approaching conventional statistical significance (p=0.06)
- approaching formal significance (p=0.1052)
- approaching independent prognostic significance (p=0.08)
- approaching marginal levels of significance p<0.107)
- approaching marginal significance (p=0.064)
- approaching more closely significance (p=0.06)
- approaching our preset significance level (p=0.076)
- approaching prognostic significance (p=0.052)
- approaching significance (p=0.09)
- approaching the traditional significance level (p=0.06)

- modestly significant (p=0.09)
- narrowly avoided significance (p=0.052)
- narrowly eluded statistical significance (p=0.0789)
- narrowly escaped significance (p=0.08)
- narrowly evaded statistical significance (p>0.05)
- narrowly failed significance (p=0.054)
- narrowly missed achieving significance (p=0.055)
- narrowly missed overall significance (p=0.06)
- narrowly missed significance (p=0.051)
- narrowly missed standard significance levels (p<0.07)
- narrowly missed the significance level (p=0.07)
- narrowly missing conventional significance (p=0.054)
- near limit significance (p=0.073)
- near miss of statistical significance (p>0.1)
- near nominal significance (p=0.064)
- near significance (p=0.07)
- near to statistical significance (p=0.056)
- near/possible significance(p=0.0661)
- near-borderline significance (p=0.10)
- near-certain significance (p=0.07)
- nearing significance (p<0.051)
- nearly acceptable level of significance (p=0.06)
- nearly approaches statistical significance (p=0.079)
- nearly borderline significance (p=0.052)
- nearly negatively significant (p<0.1)
- nearly positively significant (p=0.063)
- nearly reached a significant level (p=0.07)
- nearly reaching the level of significance (p<0.06)
- nearly significant (p=0.06)
- nearly significant tendency (p=0.06)
- nearly, but not quite significant (p>0.06)

Exercise: One-sample t-test and the pursue of normality

- Have your (VPN and) SPSS ready!
- I'll explain the exercise and then a live demo. First the what and why and then the how
- Same data and problem as the one presented by Homai, Totor and Katarina
- I will upload the dataset and output in the moodle

Exercise: One-sample t-test and the pursue of normality

- T-test family of tests: When you want to find out if 2, and only 2, **means** are statistically significantly different
  - Three members of the t-test family: independent-samples, repeated-samples, and one-sample
  - This is a family of parametric tests!
- **One-sample t-test**: when you only got 1 mean (from your data collection) that you want to compare with a population or hypothetical mean

Exercise: One-sample t-test and the pursue of normality. Hypothesis time!

- Problem statement: The brand says their 100 bear bags have an average of 50 red bears each.
- Research question: Does the 100 gummy bear bag contain, on average, 50 red gummy bears?
- Hypotheses:
  - H0: The avg of red bears per bag is 50 → H0: $\mu = 50$
  - H1: The avg of red bears per bag is different than 50 → H1: $\mu \neq 50$

Quick knowledge check:
Q: Is our H1 a directional or a non-directional hypothesis?

Quick knowledge check:
A: It's a non-directional hypothesis

Source: Group 5 cheatsheet

Exercise: One-sample t-test and the pursue of normality. Looking for normality

- Before the t-test analysis, we need to ensure our data is normal (because t-tests are parametric)
- Parametric requirements (or assumptions) for the t-test:

1. Interval or ratio dependent variable
2. Independence of observations = random sampling
3. Normality
4. No outliers

Exercise: One-sample t-test and the pursue of normality. Looking for normality

- Assumptions for the t-test:

1. Interval or ratio variable
2. Independence of observations = random sampling
3. Normality (a.k.a. Normal distribution)
4. No outliers

- Our experiment:

1. Amount of red bears (we're good)
2. 10 random bags we're taken (we're good)
3. Testing needed. To be determined
4. Testing needed. To be determined

Exercise: One-sample t-test and the pursue of normality. Looking for normality: Setting up the data

1. Create the variable and set it up

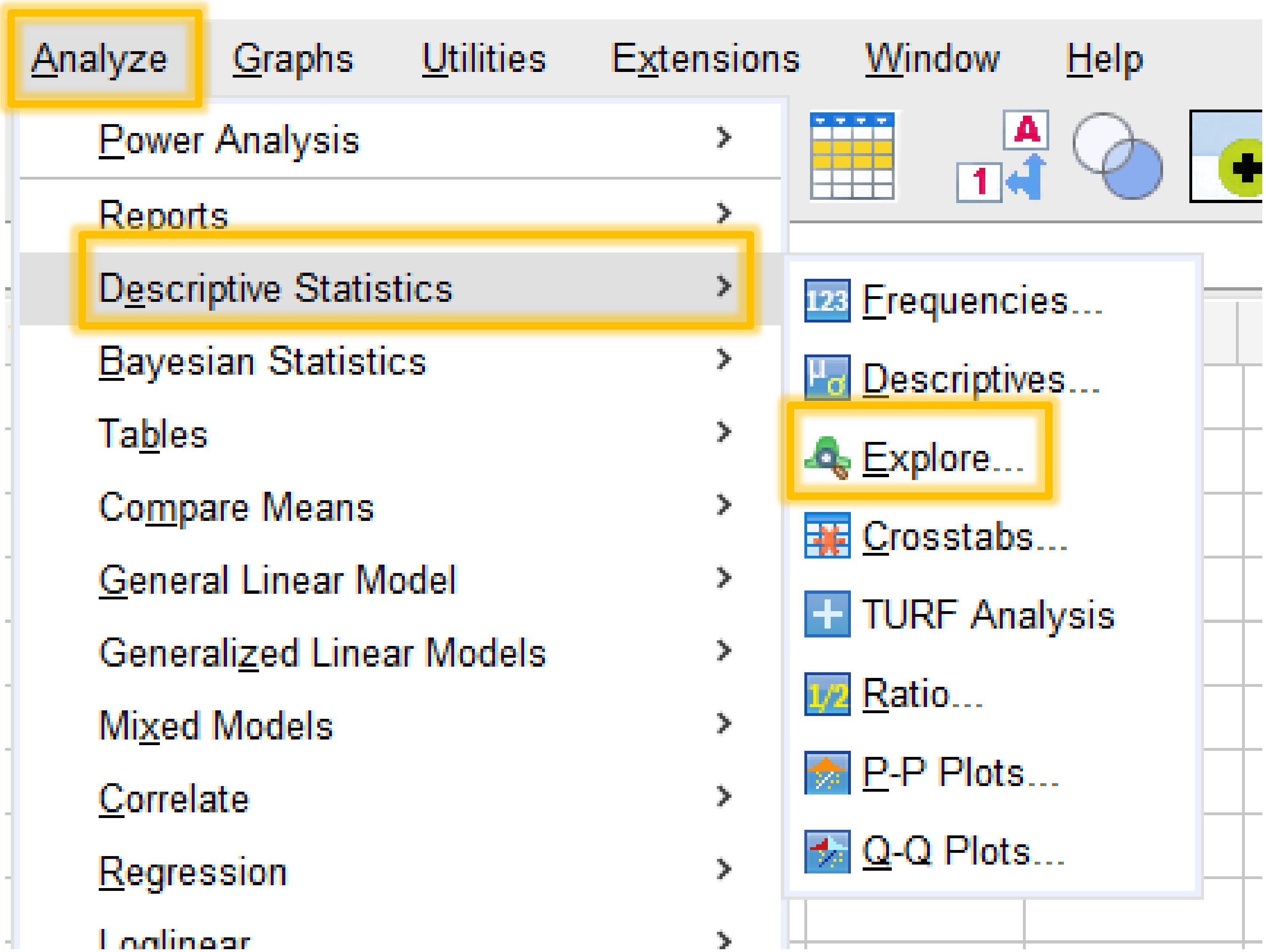
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	RedBears	Numeric	8	0		None	None	8	Right	Scale	Input

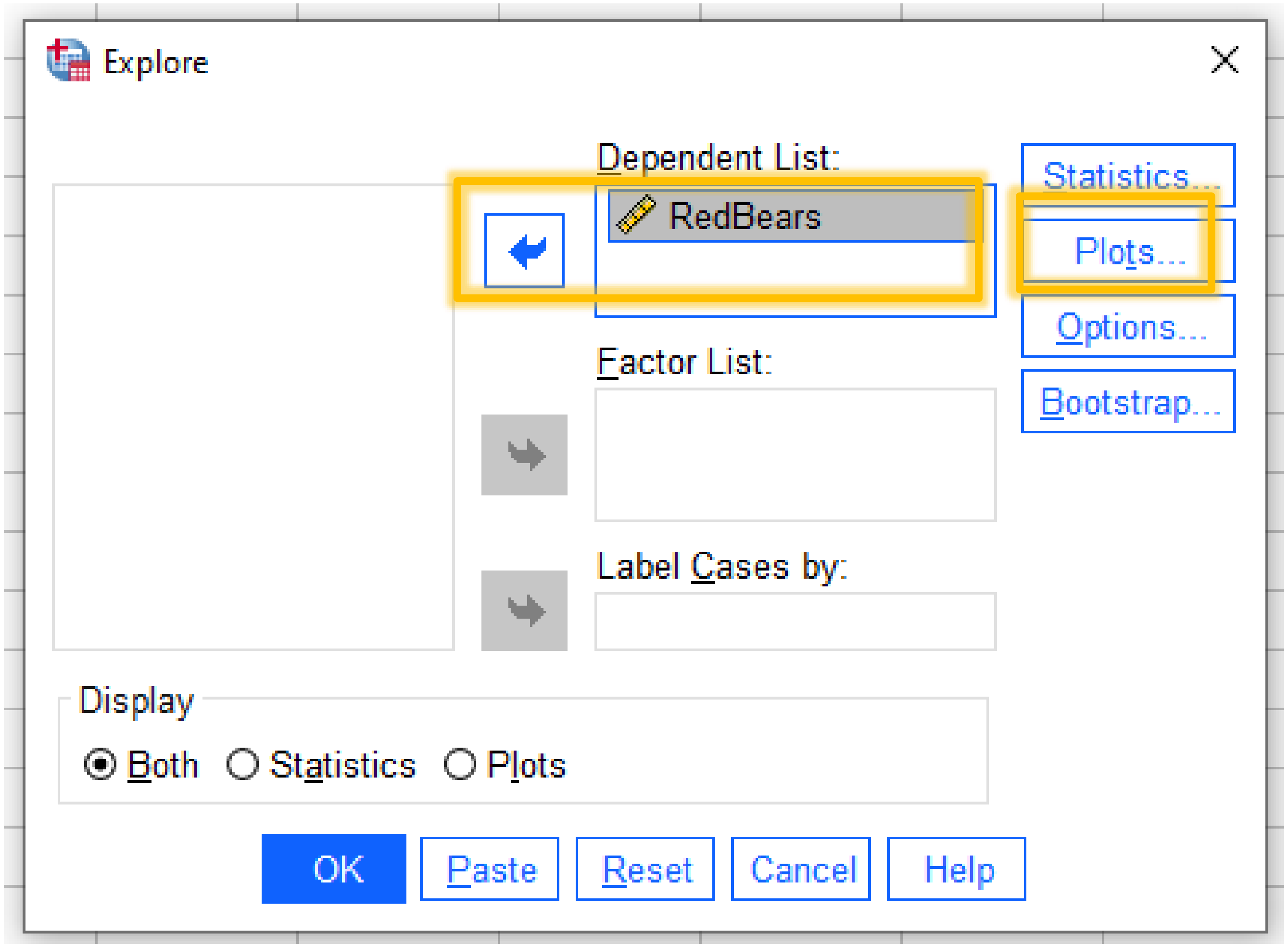
2. Add the values (observations)

	RedBears
1	48
2	58
3	59
4	61
5	49
6	50
7	52
8	55
9	56
10	49

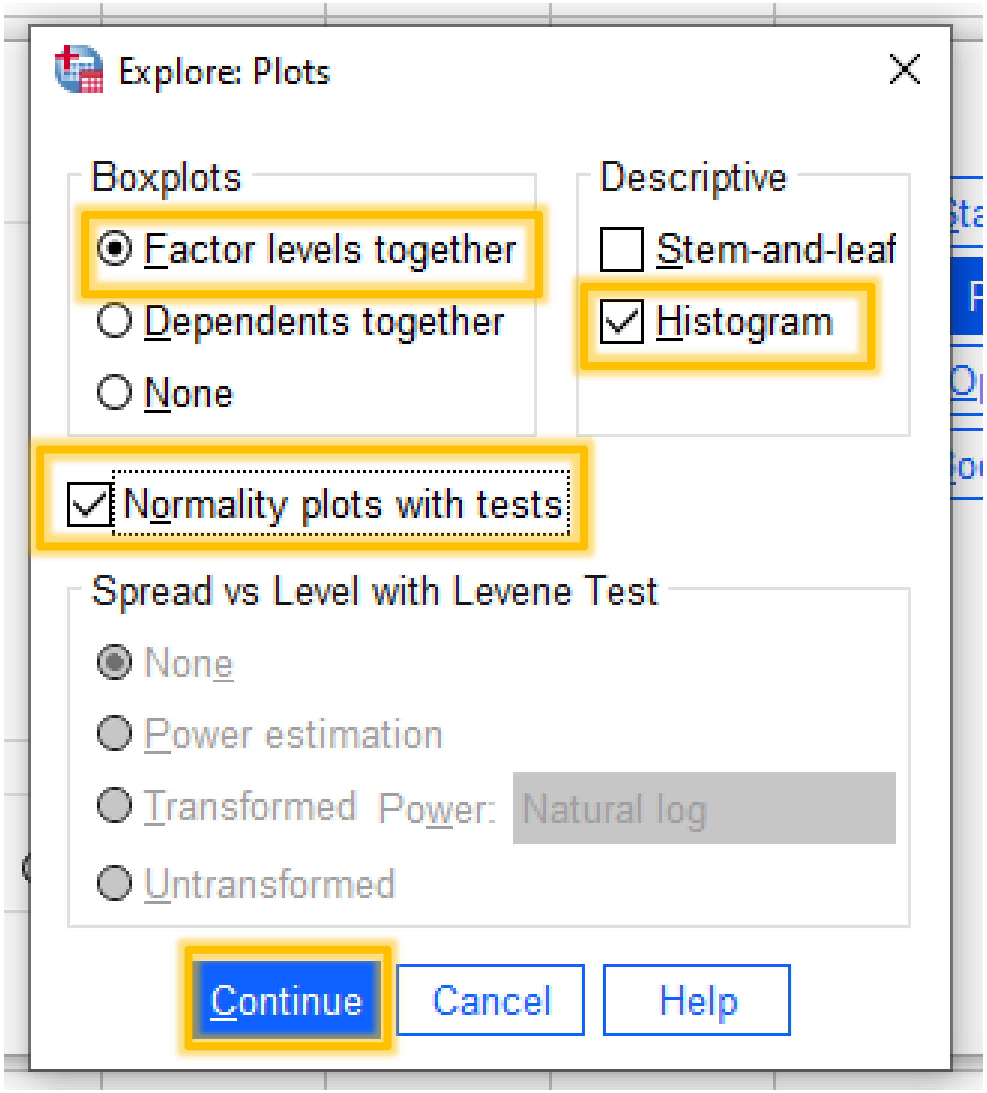
Exercise: One-sample t-test and the pursue of normality. Looking for normality: Normality testing



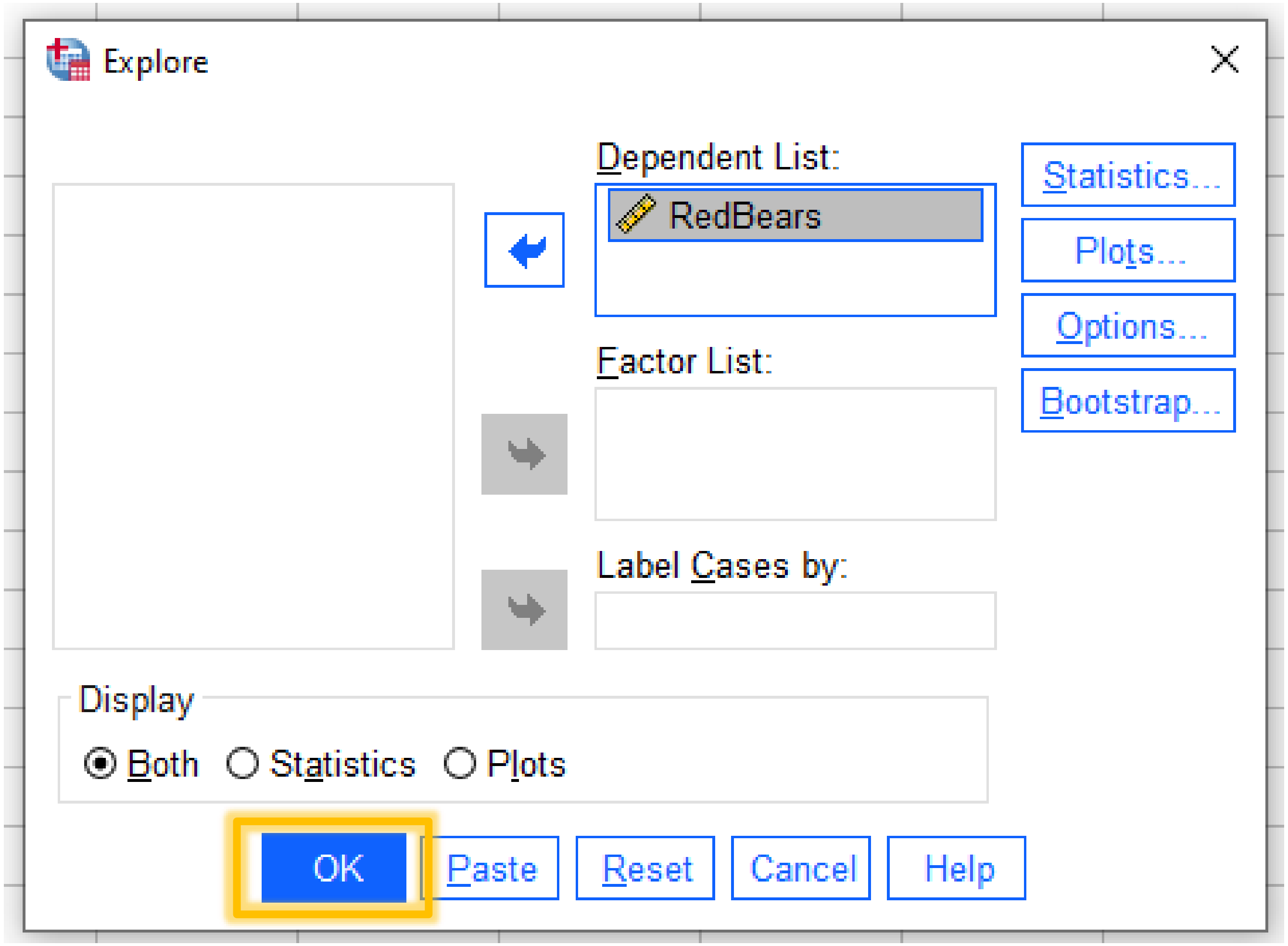
Exercise: One-sample t-test and the pursue of normality. Looking for normality: Normality testing



Exercise: One-sample t-test and the pursue of normality. Looking for normality: Normality testing



Exercise: One-sample t-test and the pursue of normality. Looking for normality: Normality testing



Exercise: One-sample t-test and the pursue of normality. Looking for normality: Normality testing

Descriptives

			Statistic	Std. Error
redBears	Mean		53.70	1.491
	95% Confidence Interval for Mean	Lower Bound	50.33	
		Upper Bound	57.07	
	5% Trimmed Mean		53.61	
	Median		53.50	
	Variance		22.233	
	Std. Deviation		4.715	
	Minimum		48	
	Maximum		61	
	Range		13	
	Interquartile Range		9	
	Skewness		.243	.687
	Kurtosis		-1.589	1.334

We need to divide these 2 values and the result (i.e. the z-value) must be between -1.96 and +1.96:

0.243/0.687 = 0.35

We're good!

Exercise: One-sample t-test and the pursue of normality. Looking for normality: Normality testing

Descriptives

			Statistic	Std. Error
redBears	Mean		53.70	1.491
	95% Confidence Interval for Mean	Lower Bound	50.33	
		Upper Bound	57.07	
	5% Trimmed Mean		53.61	
	Median		53.50	
	Variance		22.233	
	Std. Deviation		4.715	
	Minimum		48	
	Maximum		61	
	Range		13	
	Interquartile Range		9	
	Skewness		.243	.687
	Kurtosis		-1.589	1.334

We need to divide these 2 values and the result (i.e. the z-value) must be between -1.96 and +1.96:

-1.589/1.334 = -1.19

We're good! (x2)

Exercise: One-sample t-test and the pursue of normality. Looking for normality: Normality testing

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RedBears	.184	10	.200*	.915	10	.320

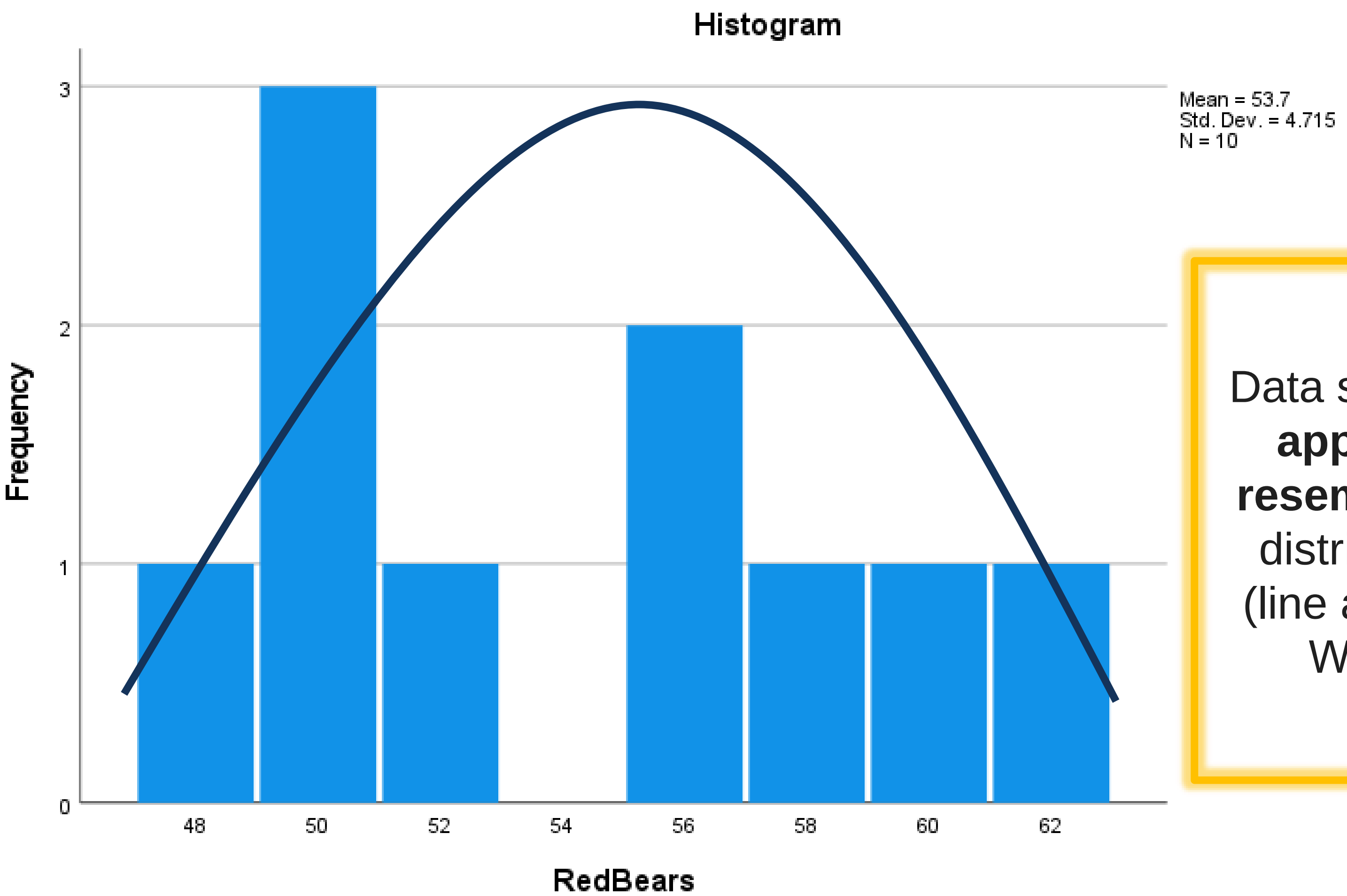
*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

For these tests, the H0 is that the data is normally distributed, therefore here we want to accept the H0:
Is 0.200 less than 0.05? No!
Is 0.320 less than 0.05? No!

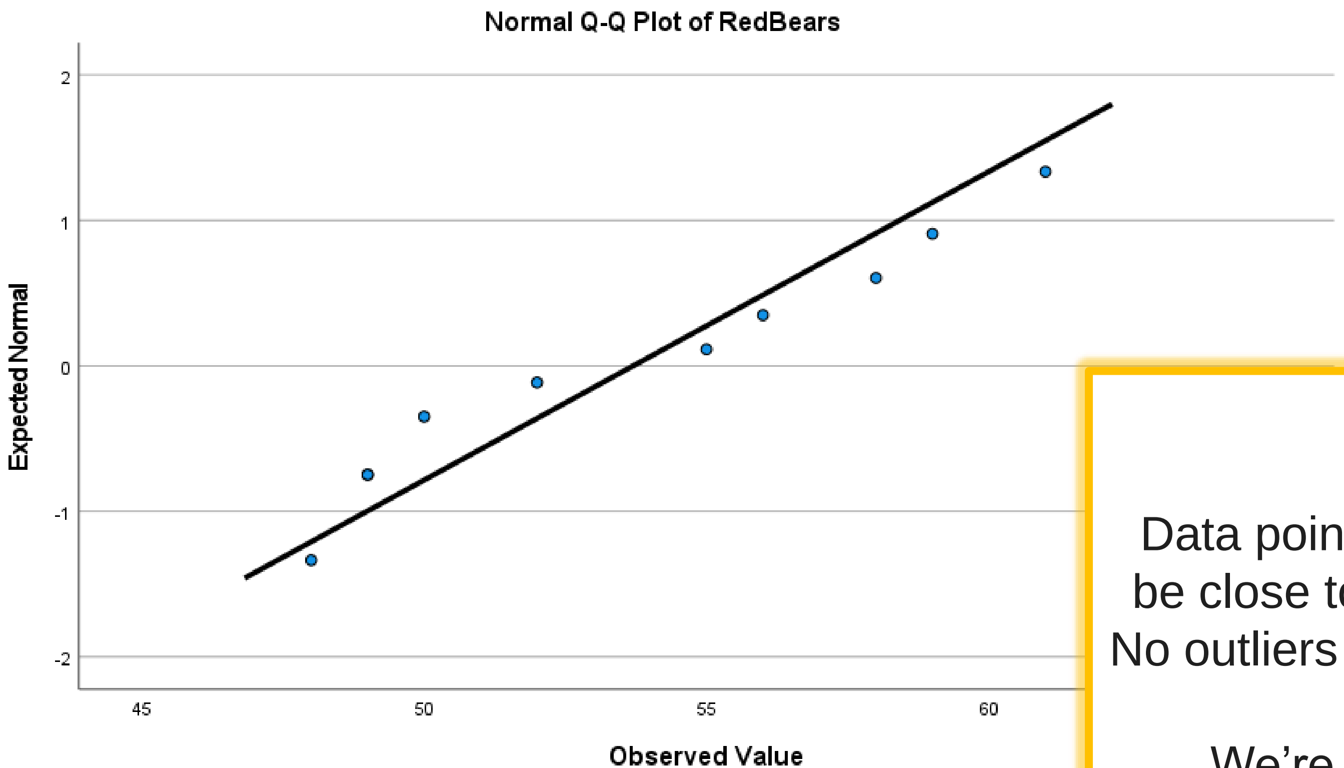
We can keep our H0 and assume the data is normally distributed!
We’re good!

Exercise: One-sample t-test and the pursue of normality. Looking for normality: Normality testing



Data should visually **approximately resemble** a normal distribution curve (line added by me) We're good!

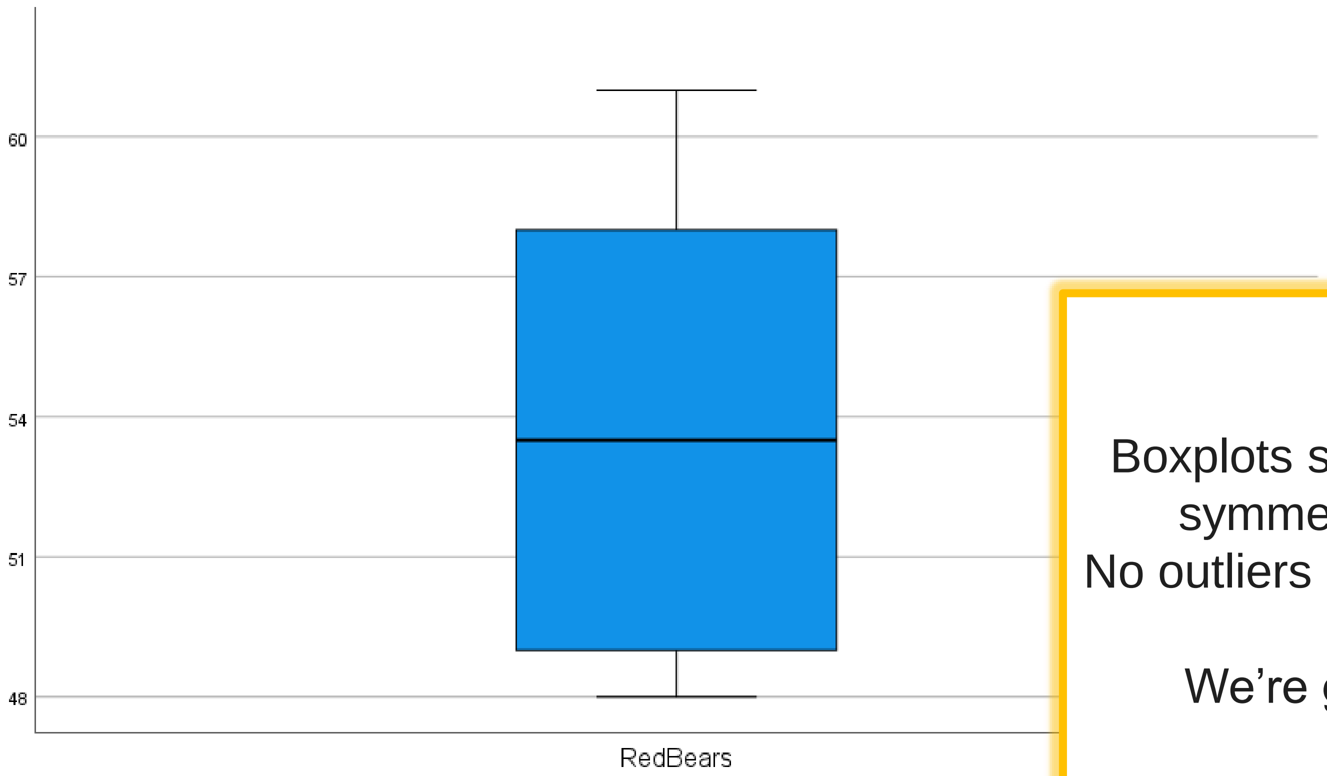
Exercise: One-sample t-test and the pursue of normality. Looking for normality: Normality testing



Data points should be close to the line. No outliers are visible

We're good!

Exercise: One-sample t-test and the pursue of normality. Looking for normality: Normality testing



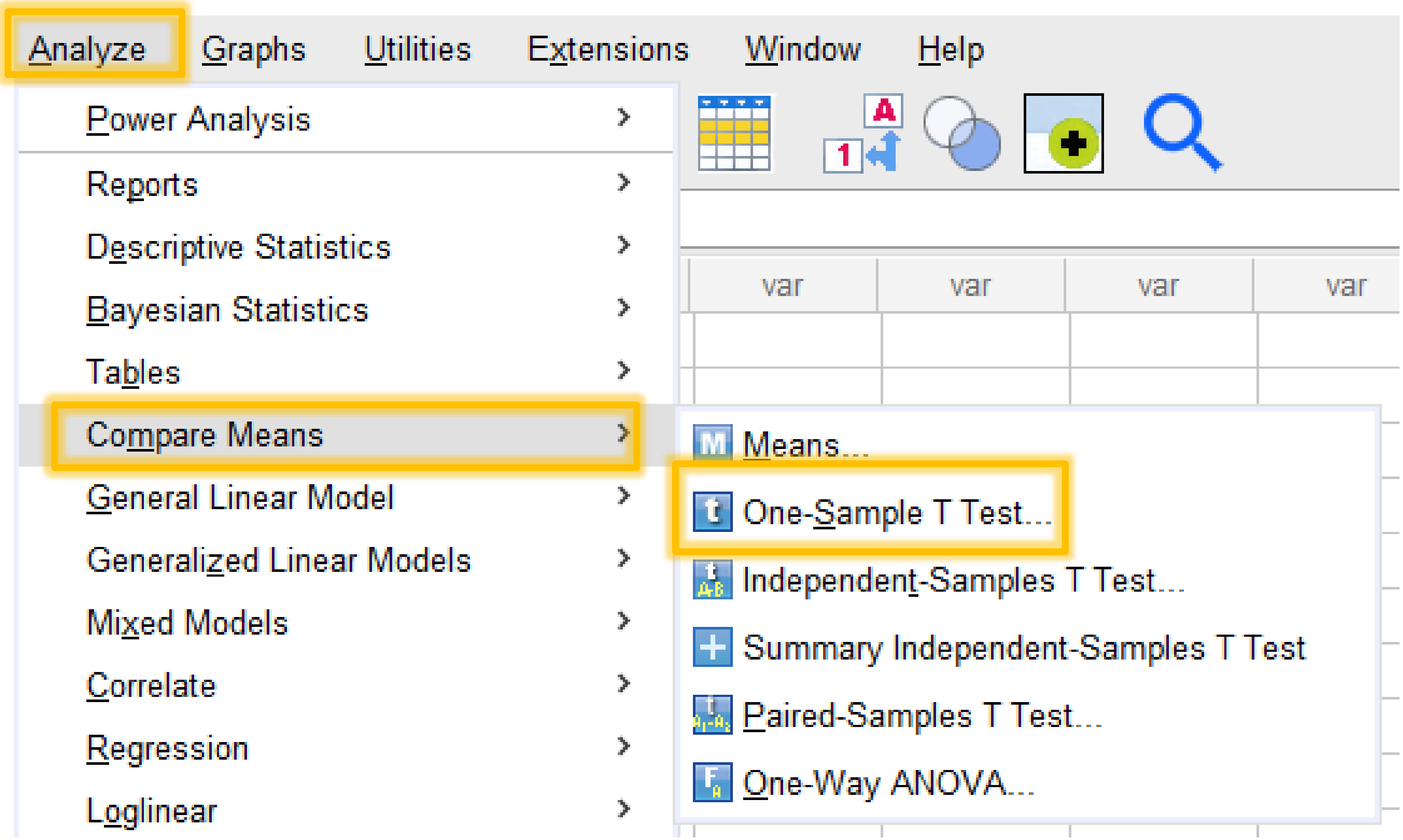
Boxplots should be symmetrical.
No outliers are visible

We’re good!

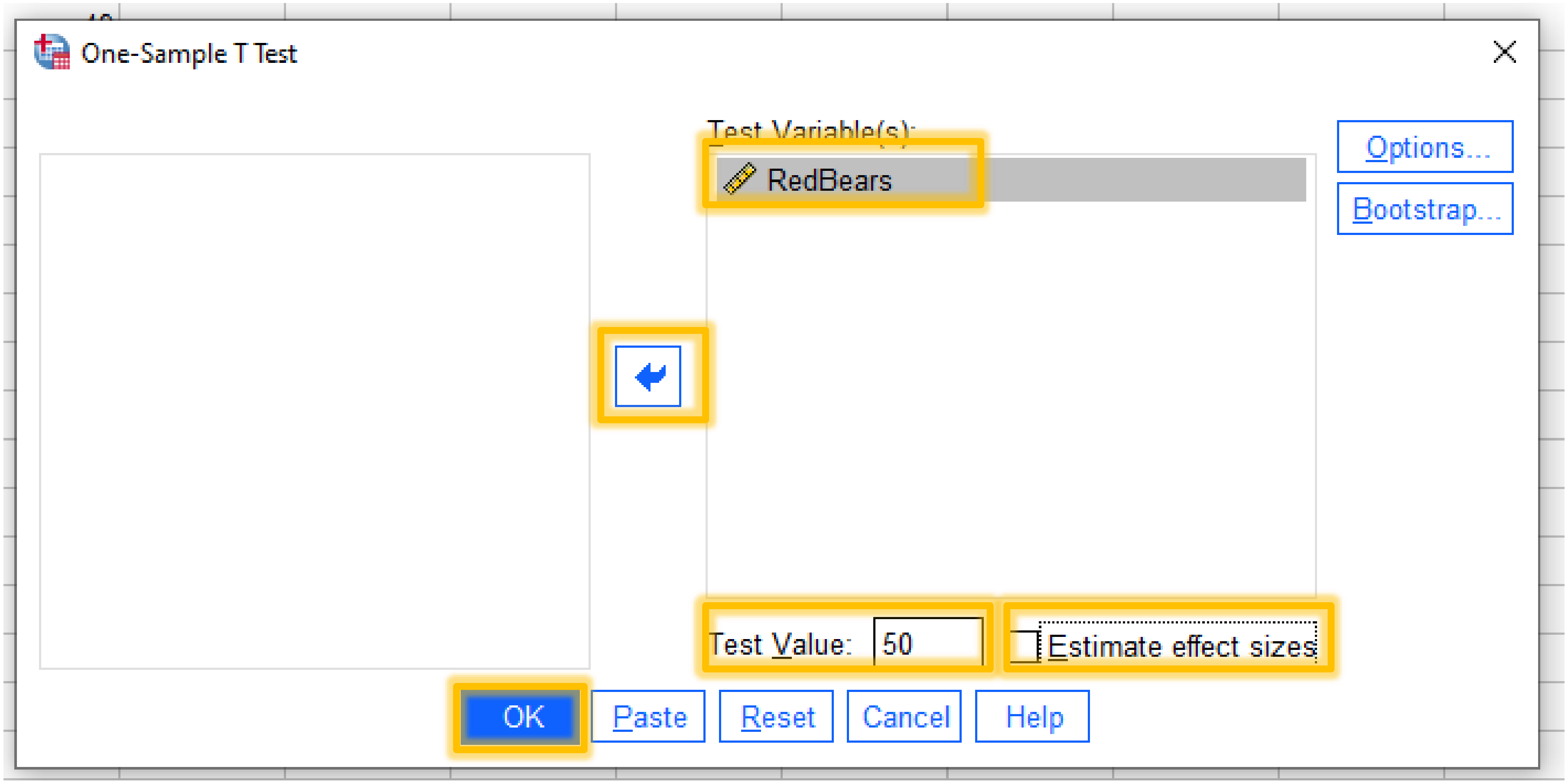
Exercise: One-sample t-test and the pursue of normality. Looking for normality: Normality testing

- Conclusion: our data is normal and we have proof for that!
- Then we can use it for a parametric test such as the one-sample t-test
- So let’s move on with that!

Exercise: One-sample t-test and the pursue of normality.
Looking for normality: One-sample t-test



Exercise: One-sample t-test and the pursue of normality.
Looking for normality: One-sample t-test



Exercise: One-sample t-test and the pursue of normality.
Looking for normality: One-sample t-test

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
RedBears	10	53.70	4.715	1.491

One-Sample Test

	Test Value = 50					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
RedBears	2.481	9	.035	3.700	.33	7.07

Exercise: One-sample t-test and the pursue of normality. Looking for normality: One-sample t-test

- If our p-value is 0.035, can we accept or reject the H0?
- We can reject the H0! (Which implies accepting the H1)
- Conclusion: after ensuring the normality of our data, we were able to prove that the average amount of red gummy bears is (statistically) significantly different from the claimed number of 50. Actually, we get more red gummy bears on each bag!

Exercise: One-sample t-test and the pursue of normality.
Looking for normality: One-sample t-test

- Possible caveats:
  1. True random sampling?
  2. Low sample size

Exercise: One-sample t-test and the pursue of normality.
Looking for normality: One-sample t-test

Now let’s do it!

Q&A:
TODAY'S TOPIC; ERM1 IN GENERAL

WEBCAMS ON FOR THE GOODBYE