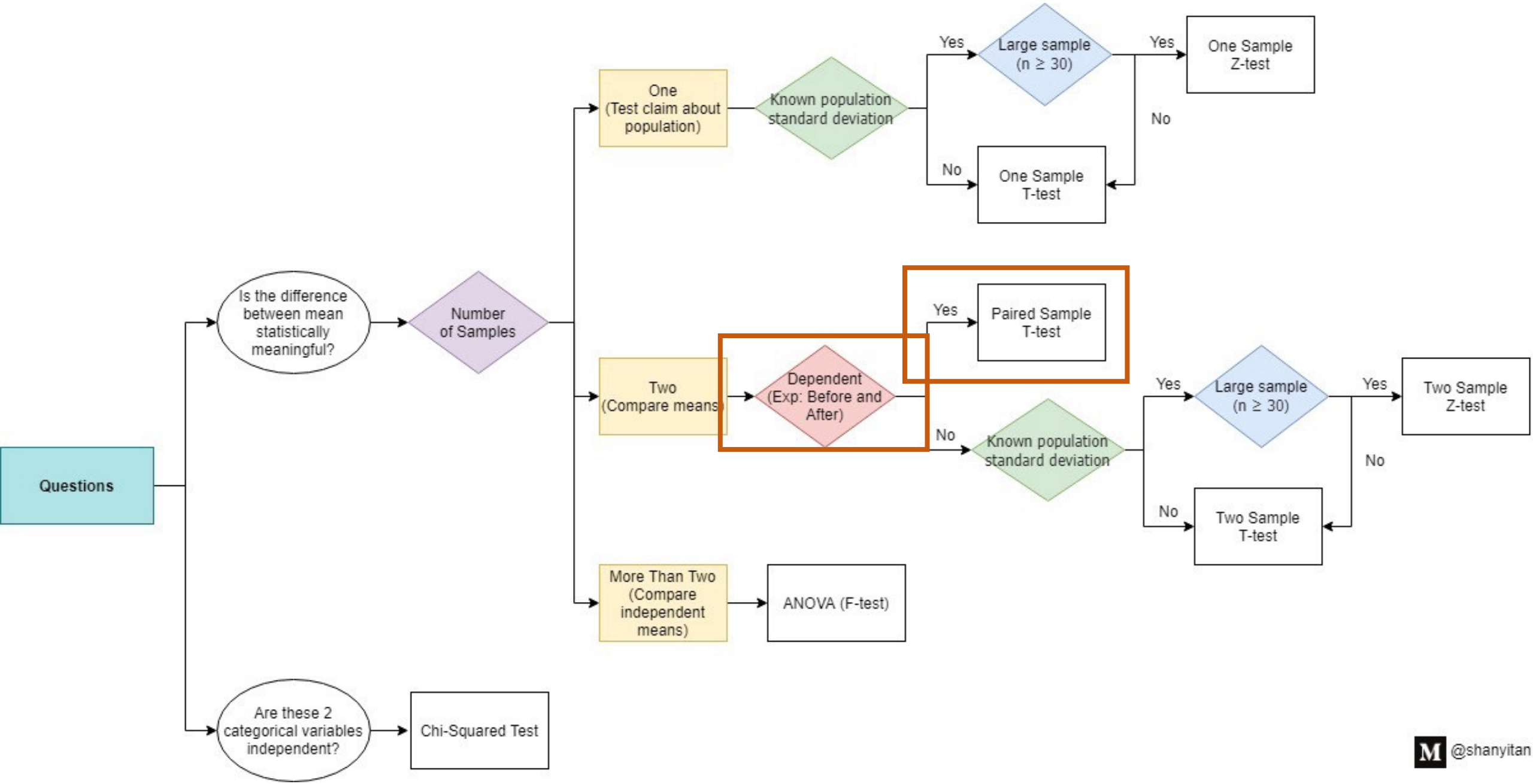
Biostatistics
Week IX

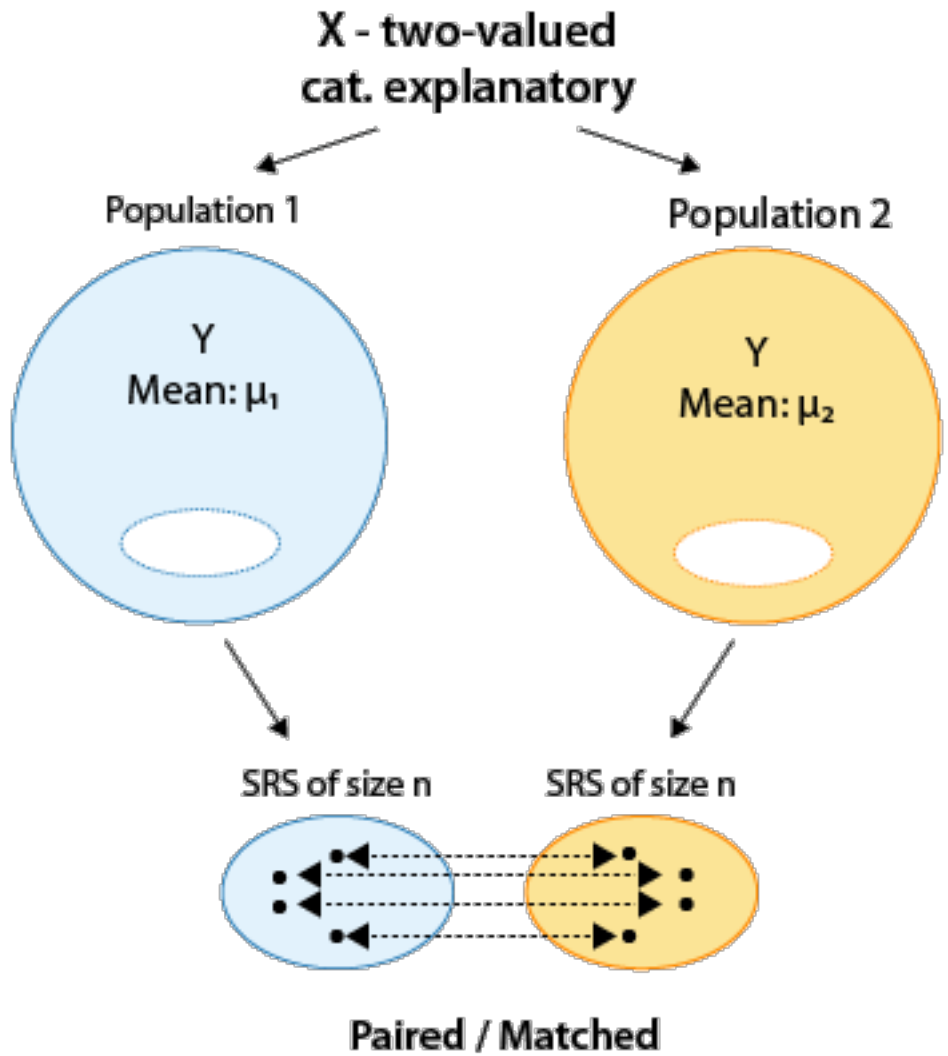
Ege Ülgen, MD, PhD

1 December 2022

ACIBADEM
MEHMET ALİ AYDINLAR ÜNİVERSİTESİ

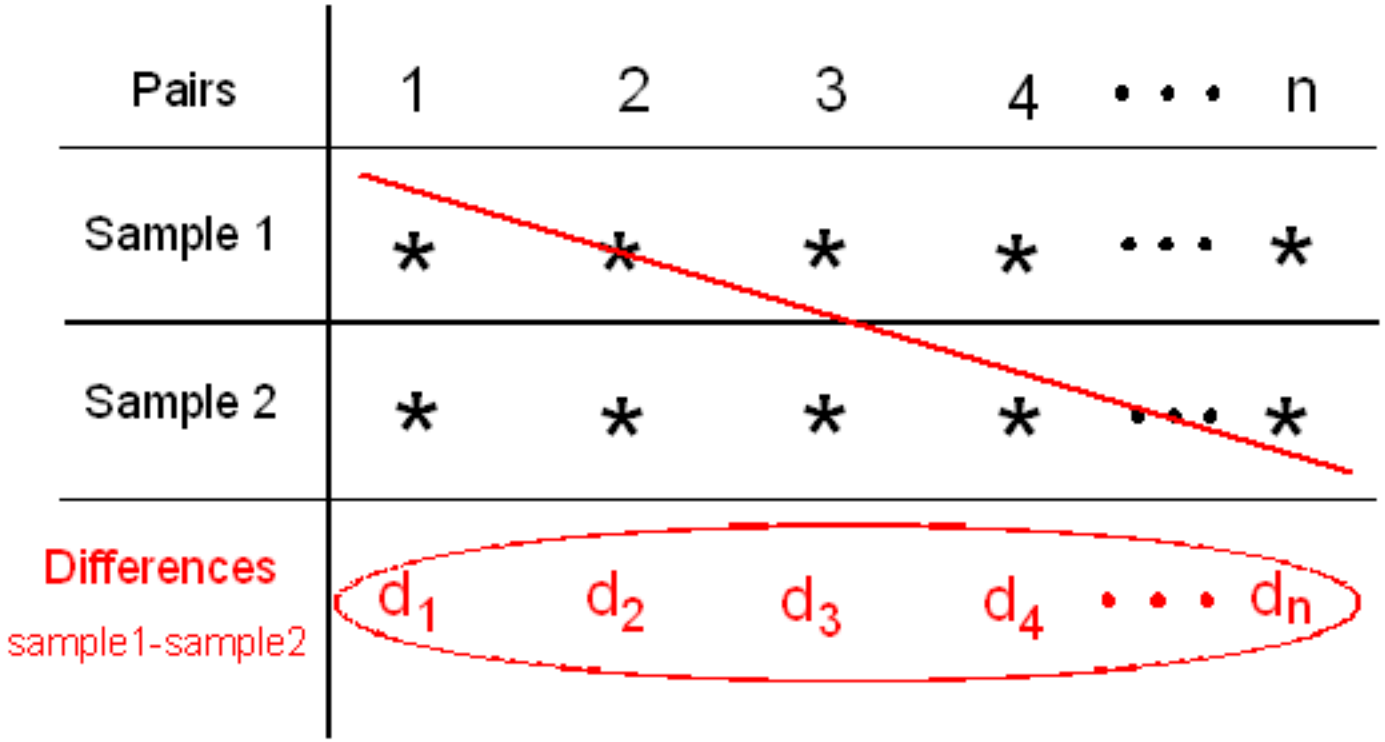

Questions
Is the difference between mean statistically meaningful?
Number of Samples
One (Test claim about population)
Known population standard deviation
Yes
Large sample (n ≥ 30)
Yes
One Sample Z-test
No
No
One Sample T-test
Two (Compare means)
Dependent (Exp: Before and After)
Yes
Paired Sample T-test
No
Known population standard deviation
Yes
Large sample (n ≥ 30)
Yes
Two Sample Z-test
No
No
Two Sample T-test
More Than Two (Compare independent means)
ANOVA (F-test)
Are these 2 categorical variables independent?
Chi-Squared Test
@shanyitan

https://bolt.mph.ufl.edu/6050-6052/unit-4b/module-13/paired-t-test/

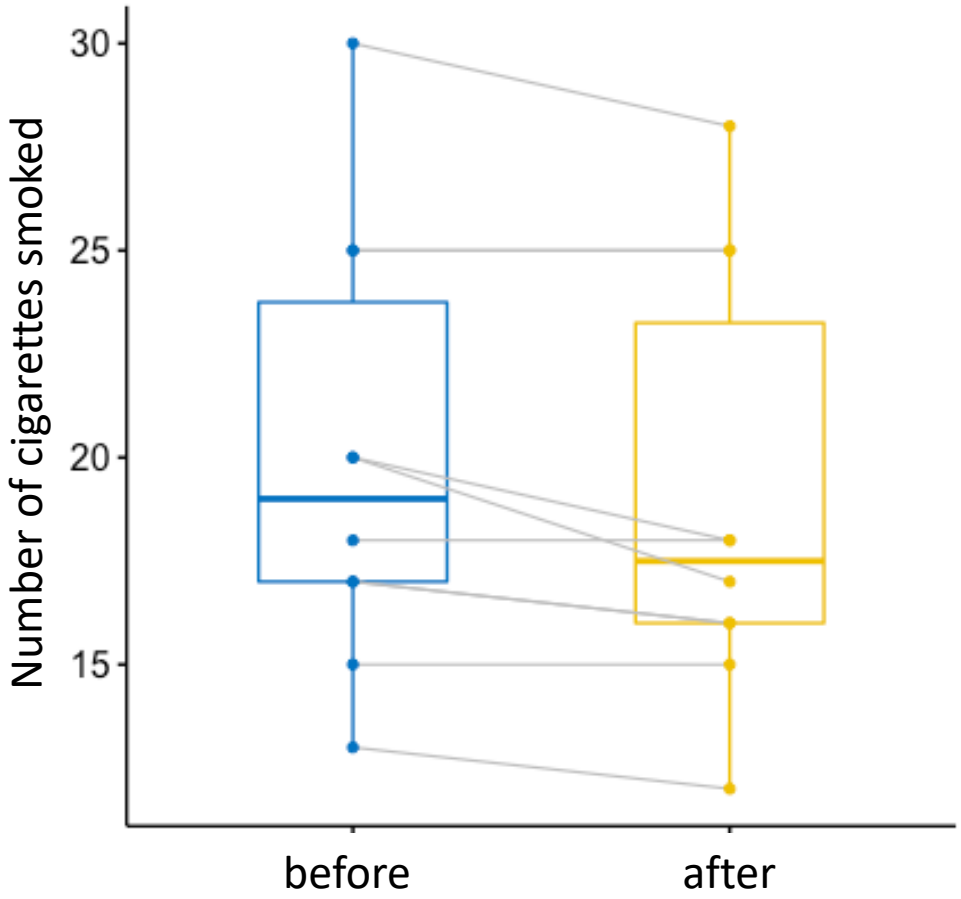


https://bolt.mph.ufl.edu/6050-6052/unit-4b/module-13/paired-t-test/

Paired/Matched t Test = essentially one-sample t Test



Paired/Matched t Test



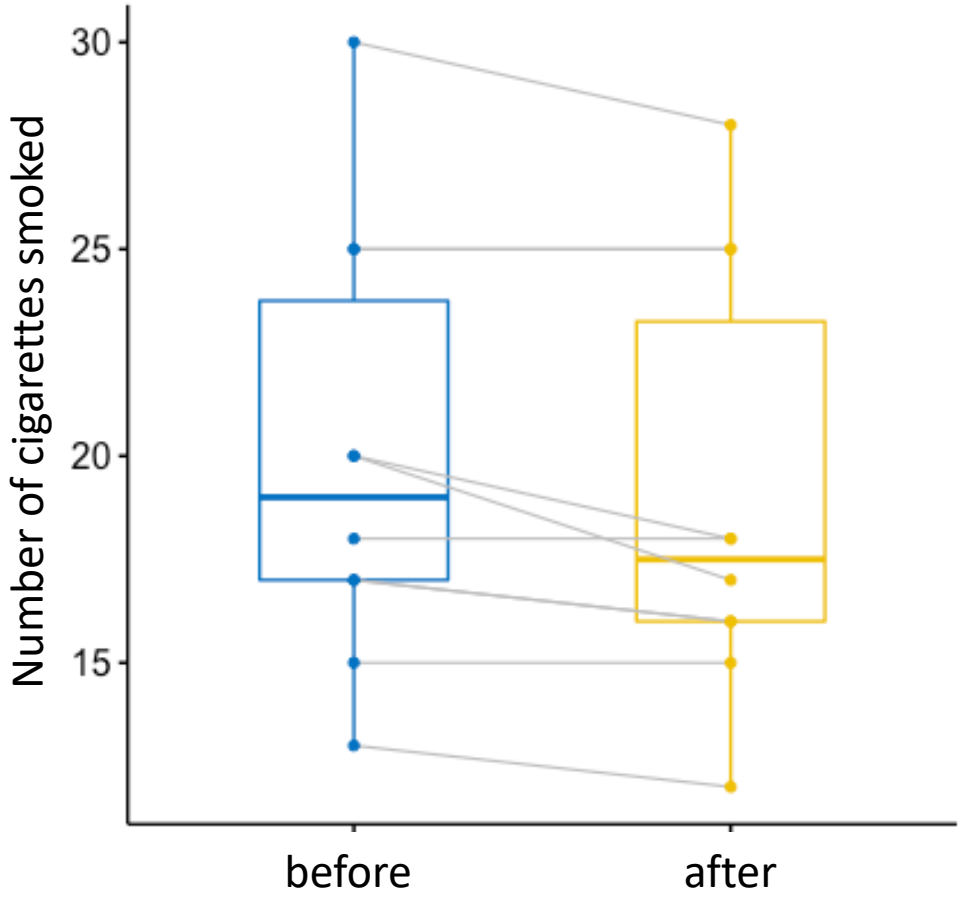
$$D = n_{before} - n_{after}$$

$$H_0: mean(D) = 0$$
$$H_a: mean(D) \neq 0$$

Paired/Matched t Test - Example

$D = X_1 - X_2$

X_{1i}	X_{2i}	D_j
30	28	2
25	25	0
25	25	0
20	18	2
20	17	3
18	18	0
17	16	1
17	16	1
15	15	0
13	12	1



Paired/Matched t Test - Example

1. Check assumptions, determine H_0 ve H_a , select α
   - $H_0: \mu_D \leq 0$ $H_a: \mu_D > 0$ $D = X_1 - X_2$
   - $\alpha = 0.05$

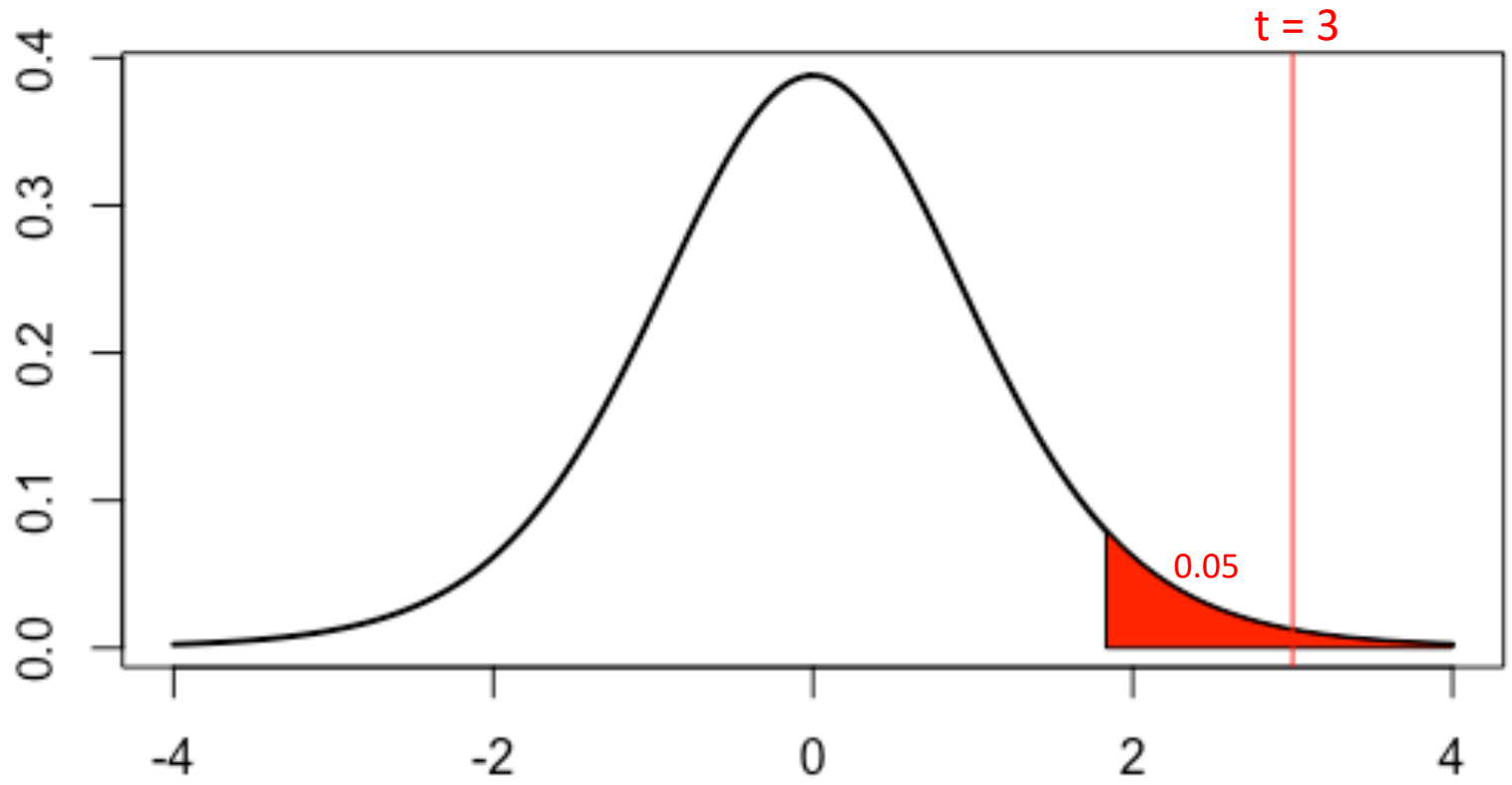
2. Calculate test statistic

$$t = 3 \quad (\sim t_{10-1=9})$$

$D = X_1 - X_2$

Paired/Matched t Test - Example

3. **Critical region(s)/** p value
4. Decide to reject/failt to reject H_0



p = 0.0074782

Hypothesis testing steps

1. **Check assumptions**, determine H_0 ve H_a , select α

   …

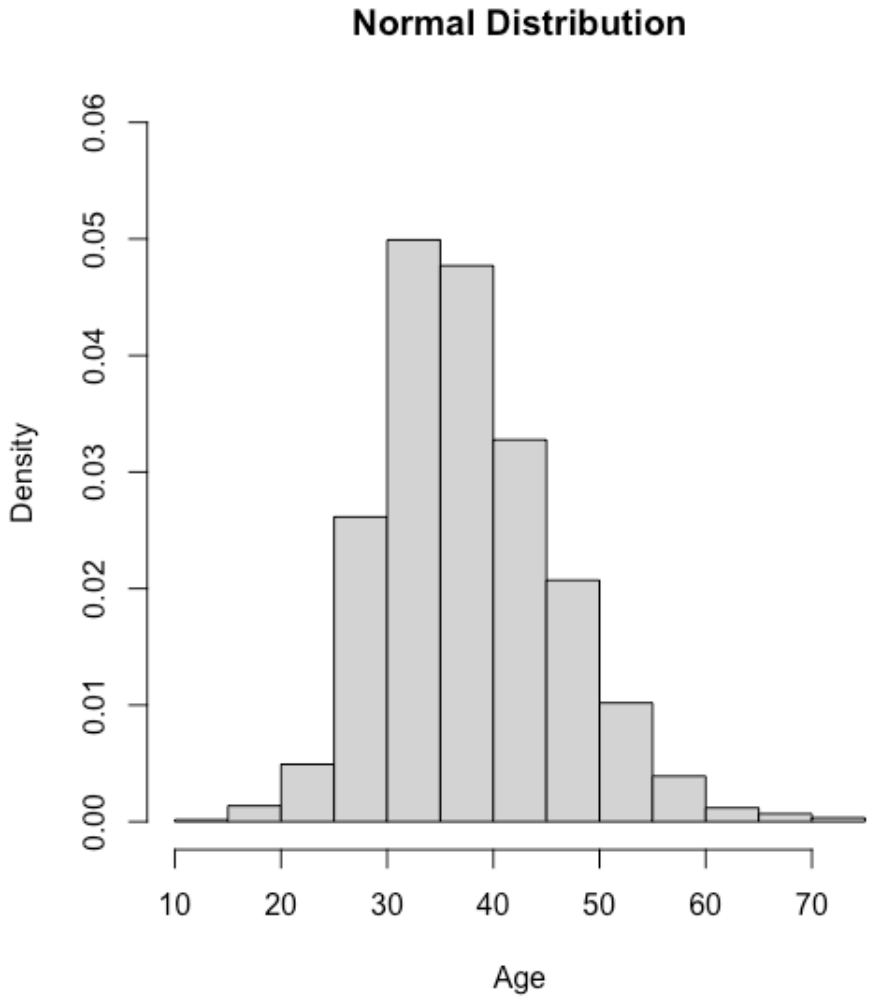
General Assumptions of Parametric Tests

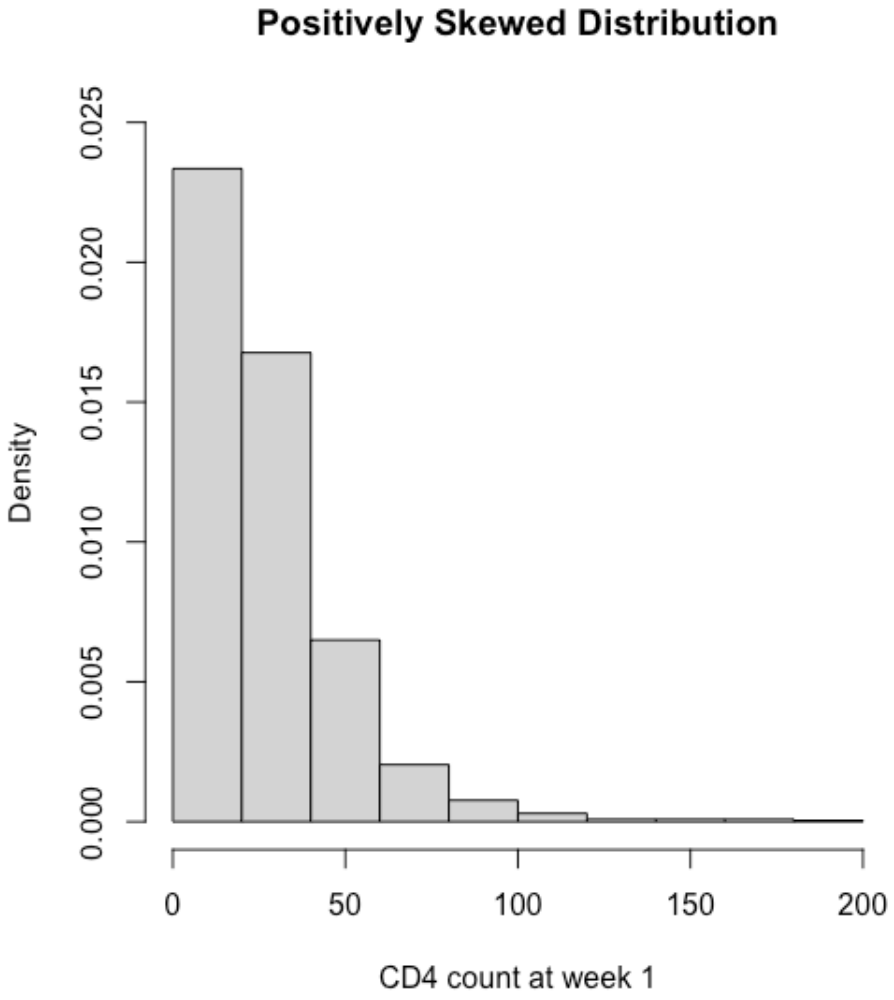
- The population(s) are **normally distributed**
- The selected sample is **representative of general population**
- The data is **continuous**

Assessing Normality

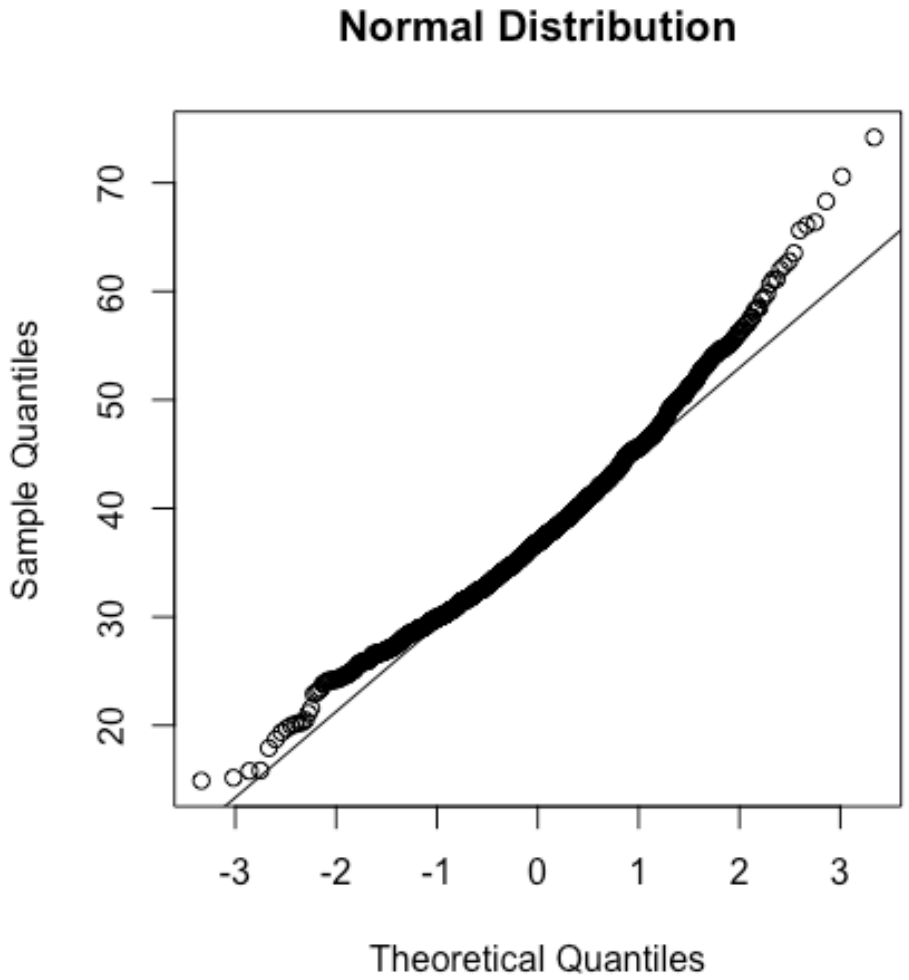
- Inspecting the **histogram** of the variable
- **Quantile-quantile plots**
- **Shapiro-Wilk test**
  - $p < 0.05$ indicates normal distribution
- …

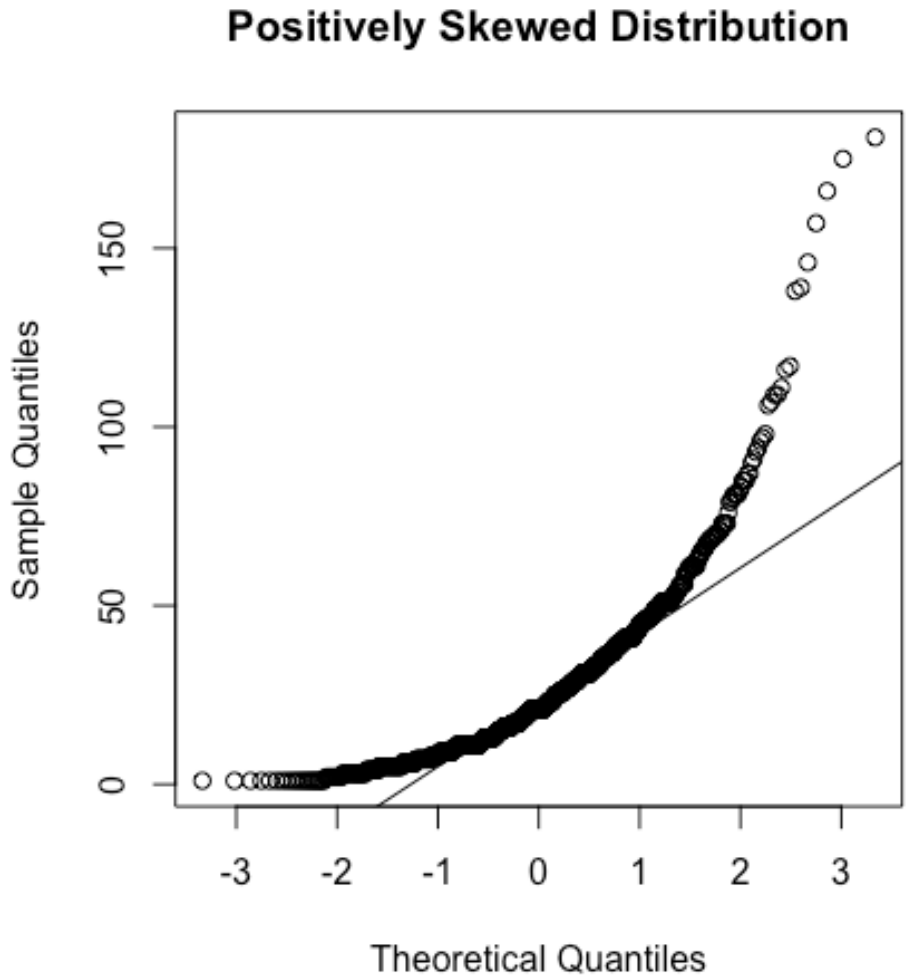
Inspecting Histogram





Quantile-Quantile Plots





Non-parametric Tests

- Used when assumptions of parametric tests are not met
- **Not dependent on the distribution**
- **Less assumptions**
  - e.g., they do not depend on the assumption of normality

- **Less statistical power** compared to parametric tests
  - Higher risk of type II errors (e.g., high probability of accepting there is no difference between the groups where there is a difference)

Non-parametric Tests

- **χ^2 test**
- **Wilcoxon rank-sum test (Mann–Whitney U test)** ~ t-test
- **Kruskal-Wallis test** ~ANOVA
- **Spearman's rank correlation test** ~ Pearson correlation test
- …

Brief Summary

- Normality of a variable can be assessed using
  - Histogram
  - Q-Q plot
  - Shapiro-Wilk test
- Non-parametric tests have **fewer assumptions** but also have **less statistical power** compared to parametric tests