Università degli Studi di Padova
Department of Industrial Engineering

Design of Experiments
Lesson #3

Academic year 2025-2026

Prof. Pierantonio Facco
CAPE-Lab, Computer-Aided Process Engineering Laboratory

Email: pierantonio.facco@unipd.it
URL: https://research.dii.unipd.it/capelab/

Framing this lesson and the future ones

1. Does different treatments or different levels of a factor determine a statistically meaningful change on the response?
2. Is there a significant impact of factors/treatments on a response variable?
3. How much the variability of one factor determine the variability of the response?

→

1. Simple comparative experiments
2. ANOVA
3. DoE and response surface modelling

Simple comparative experiments

Objective

- Comparing experiments from two **different treatments/conditions**, often called **simple comparative experiments**
  - for example: do two different formulations lead to equivalent quality of the product?

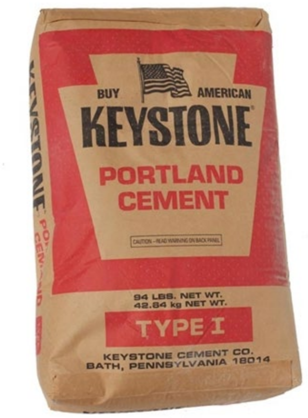
- Secondary aim:
  - review of **basic statistical concepts**:
    - random variables
    - probability distribution
    - random sampling
    - sampling distribution
    - hypothesis testing

Portland cement mortar example

- Evaluation of the effect of adding a polymer latex emulsion (used to decrease the cure time) on the tension bond strength of a Portland cement mortar
  - two different formulations are analyzed
    - different formulations are considered as two treatments/levels of the factor formulations

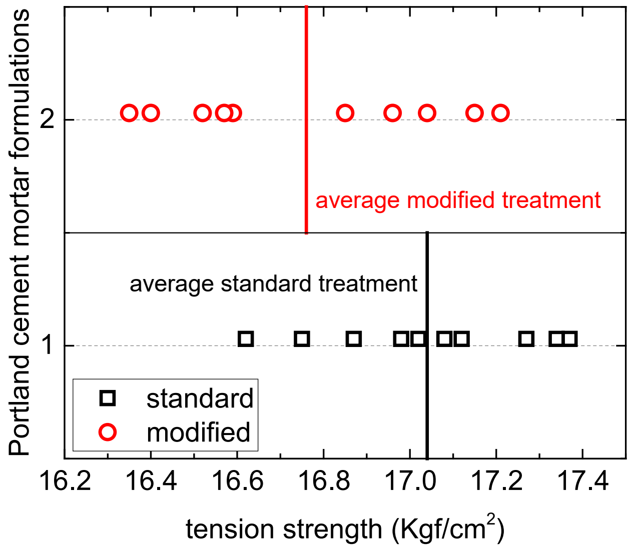
sample	modified mortar	standard mortar
1	16.85	16.62
2	16.40	16.75
3	17.21	17.37
4	16.35	17.12
5	16.52	16.98
6	17.04	16.87
7	16.96	17.34
8	17.15	17.02
9	16.59	17.08
10	16.57	17.27

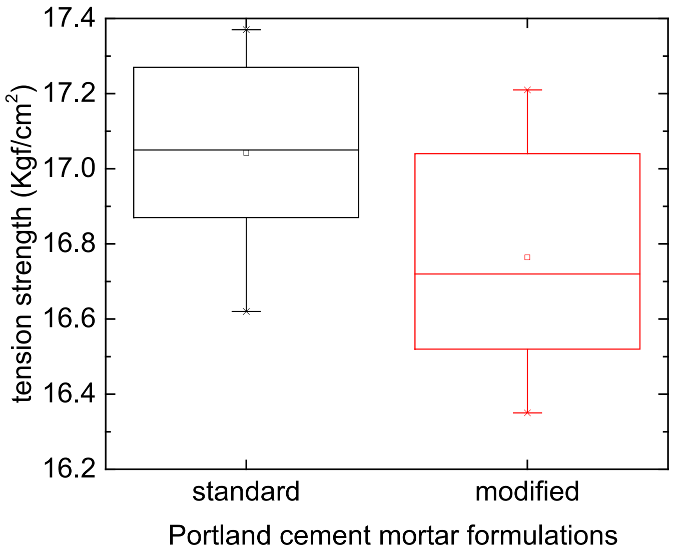
Ideas?



Methods for visual inspection

- **Dot diagram** or **individual value plot**:
  - simple representation of points in their scale
  - identifies the range of data variability
- **Box-and-whiskers plot**:
  - rigorous representation of the sample distribution
  - describes indices of:
    - position
    - dispersion
    - skewness
  - the box represents:
    - inter-quantile range IQR
      - difference among first and third quartiles
    - mean and/or median
  - the whiskers represent:
    - either 2^{nd} and 98^{th} percentiles, or 1.5·IQR, or 3·IQR, or…
  - outliers may be also represented





Comparison among average (std. deviation)

- Average and standard deviation of the two formulations:

treatment	average	standard deviation
standard formulation	17.04(2) kgf/cm^2	0.32 kgf/cm^2
modified formulation	16.76(4) kgf/cm^2	0.25 kgf/cm^2

- Are these differences large enough to imply that the two formulations are different from the statistical point of view? Or are the formulations identical for what concerns the tension bond strength?

⬇

Hypothesis testing

Statistical hypothesis

- The **statistical hypothesis** on the Portland cement example reflects the following **conjecture** about the problem:
  - the mean tension bond strengths of the two formulations may be equal
  - the mean tension bond strengths of the two formulations may be different

- The hypothesis testing in this case is stated formally as:

$$H_0\colon \mu_1 = \mu_2$$
$$H_1\colon \mu_1 \neq \mu_2$$

  - H_0 is called the **null hypothesis**
  - H_1 is called the **alternative hypothesis**
    - this is a **two-sided alternative hypothesis** because it would be true if either $\mu_1 < \mu_2$ or if $\mu_1 > \mu_2$

Test statistics for hypothesis testing

- Assumption: the **variances of tension bond strengths are identical**
- The **appropriate test statistic to use for comparing two treatment means** is:

$$t_o = \frac{\bar{x}_1 - \bar{x}_2}{S_p\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}}$$

  - where $\bar{x}_1$ and $\bar{x}_2$ are the sample means, N_1 and N_2 are the sample sizes, and S_p is an estimate of the common variance computed from:

$$S_p^2 = \frac{(N_1 - 1)S_1^2 + (N_2 - 1)S_2^2}{N_1 + N_2 - 2}$$

  - S_1^2 and S_2^2 are the individual sample variances
  - $S_p\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}$ is called the **standard error**

Two sample t-test

- **Procedure** to carry out a **two-sample *t*-test**:
  - compare t_o to the t-distribution with $N_1 + N_2 - 2$ degrees of freedom to determine whether to reject the null hypothesis $H_0: \mu_1 = \mu_2$
  - test to reject H_0:

$$|t_o| > t_{N_1+N_2-2,\alpha/2}$$

    - where $t_{N_1+N_2-2,\alpha/2}$ is the upper $\alpha/2$ percentage point of the t-distribution with $N_1 + N_2 - 2$ degrees of freedom
  - if we reject H_0 ➡ the mean strengths of the two formulations of Portland cement mortar differ
  - if we fail to reject H_0 ➡ the mean strengths of the two formulations of Portland cement mortar are the same

Justification of the two-sample t-test procedure

- Samples are taken from independent normal distributions
- The distribution of $\bar{x}_1 - \bar{x}_2$ is:

$$\bar{x}_1 - \bar{x}_2 \sim N\left(\mu_1 - \mu_2, \sigma^2\left(\frac{1}{N_1} + \frac{1}{N_2}\right)\right)$$

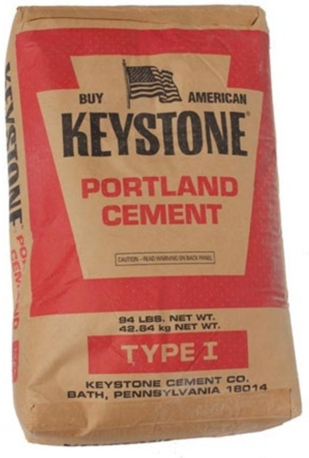
- In case σ^2 was known and $H_0: \mu_1 = \mu_2$ holds true:

$$z_o = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}} \sim N(0,1)$$

- Changing σ with S_p, the distribution becomes a $t_{N_1+N_2-2}$ distribution
  - if $H_0: \mu_1 = \mu_2$ is true
    - t_o is within the limits of $\pm t_{N_1+N_2-2,\alpha/2}$ and $100(1-\alpha)\%$ of the values of t_o are expected to fall within the limits
  - samples producing t_o outside these limits would be unusual and this is evidence that the null hypothesis should be rejected
- Note that α is the probability of type I error for the test
  - it is the **significance level** of the test

Example: Portland cement mortar

- In the Portland cement mortar example:
  - $S_p = 0.284$
  - $t_o = -2.2$
  - for $\alpha = 0.05$, $t_{18,0.025} = ?$



Percentage points of the t-distribution

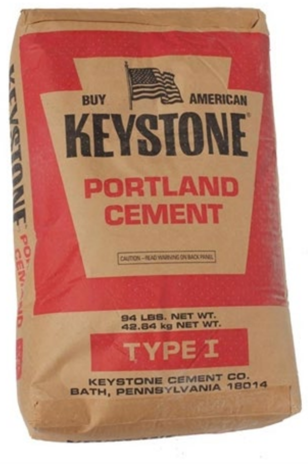
- Tables are available in the textbooks

II Percentage Points of the *t* Distribution[a]

ν \ α	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	14.089	23.326	31.598
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.265	0.727	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	4.019	4.785	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725

- Alternative:
  - calculation in Matlab®: `tinv(0.025,18)`

Example: Portland cement mortar



- In the Portland cement mortar example:
  - $S_p = 0.284$
  - $t_o = -2.2$
  - for $\alpha = 0.05$, $t_{18,0.025} = 2.101$
  - since $|t_o| > t_{18,0.025}$ the null hypothesis can be rejected
- Results of the two-sample t test:
  - the mean tension bond strengths of the *two formulations of Portland cement mortar are different* from the statistical point of view!
  - this is an important engineering finding:
    - the change in mortar formulation had the **desired effect of reducing the cure time**, but there is evidence that the change also affected the tension bond strength
    - the modified formulation **reduces the bond strength**
    - if the reduction in mean bond strength is of practical importance, then more development work and further experimentation will likely be required

Fixed significance level approach vs. p-value approach

1. **Fixed significance level testing** reports the results of a hypothesis test stating that the null hypothesis was (or was not) rejected at a specified value or **level of significance**
   - Portland cement mortar formulation example: the null hypothesis was rejected at the 0.05 significance level
   - this statement of conclusions is often **inadequate**:
     - **was the computed value of the test statistic just barely in the rejection region?**
     - or was it very far into this region?
2. **P-value approach** is adopted widely in practice to avoid these difficulties
   - the **P-value** is the **smallest level of significance that would lead to rejection of the null hypothesis** H_0
     - the P-value is the probability that the test statistic will take on a value that is at least as extreme as the observed value of the statistic when the null hypothesis is true
   - the P-value conveys information about the **weight of evidence against** H_0
     - a decision maker can draw a conclusion at any specified level of significance
     - we may think of the *P*-value as the smallest level at which the data are significant
       - it is customary to call the test statistic (and the data) significant when the null hypothesis H_0 is rejected

Calculation of the p-value

- It is not always easy to compute the exact *P*-value for a test
  - the computer programs for statistical analysis report P-values
  - Matlab® calculation: `2*(1-tcdf(2.2,18))` or `2*tcdf(-2.2,18)`
- Interpolate among the table values:
  - since $|t_o| > t_{18,0.025}$ the p-value is less than 0.05
  - for a t-distribution with 18 degrees of freedom and $t_o = -2.2$, the p-value should be between 2·0.01 and 0.05

II Percentage Points of the *t* Distribution[a]

ν \ α	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	14.089	23.326	31.598
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.265	0.727	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	4.019	4.785	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.258	0.691	1.341	1.753	2.1[illegible]1	2.602	2.947	3.286	3.733	4.073
16	0.258	0.690	1.337	1.746	2.[illegible]20	2.583	2.921	3.252	3.686	4.015
17	0.257	0.689	1.333	1.740	2.110	2.[illegible]67	2.898	3.222	3.646	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725

The p-value which corresponds to $t_o = -2.2$ is $p = 0.0411$

- this means that **the null hypothesis is rejected at any significance level**

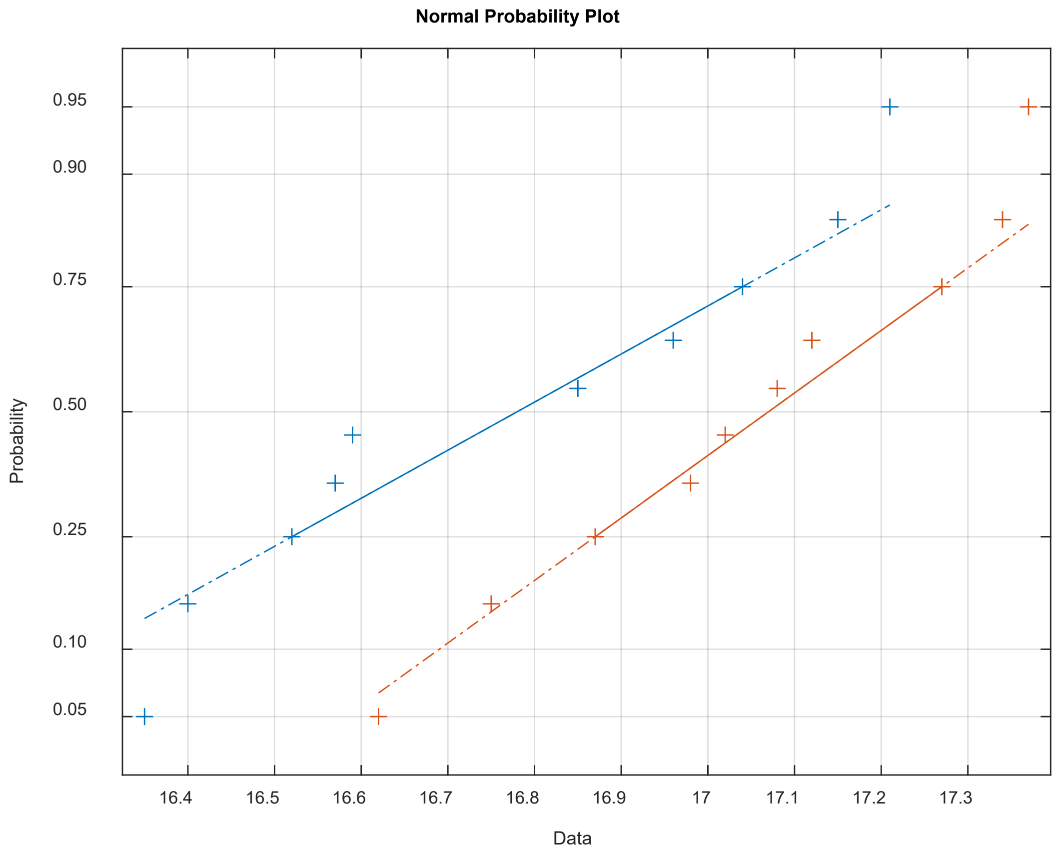
$$\boldsymbol{\alpha > 0.0411}$$

t-test assumptions verification

- The assumptions at the basis of the paired t-test are:
  - samples are **random samples** are drawn from **independent populations** described by a **normal distribution**
  - the **standard deviation** or variances of the populations are **equal**
  - if these assumptions are violated the t-test is affected
    - moderate departures from normality do not seriously affect the results
- The assumption of independence is usually satisfied if the run order is randomized
  - if appropriate experimental units and materials are selected at random
- This type of procedure is called a **randomization test** (if the samples observed are equally likely)
  - the *t*-tests can be used without extensive concern about the normality assumption
- How could we test if the variances are equal, and the samples are normally distributed?

Normal probability plot for Portland cement mortar

- The **normal probability plot** shows that:
  - the distribution of the **two samples are normal**
    - the points are approximately on the straight line
  - the standard deviation of the two distribution is approximately the same
    - the fitting **lines have approximately the same slope**



Alternative approach on the confidence limits

- Hypothesis testing is useful, but it sometimes does not tell the entire story
- **Confidence interval approach**: it is often preferable to provide an interval within which the value of the variable of interest would be expected to lie
  - the experimenter already knows that the means μ_1 and μ_2 differ, so he/she is more interested in knowing how much the means differ
  - to define a **confidence interval**, suppose that θ is an unknown parameter
    - we need to find two statistics L and U such that the probability statement
      $$\mathrm{P}(L \leq \theta \leq U) \leq 1 - \alpha$$
      holds true
    - the interval is called a $100(1-\alpha)\%$ **confidence interval** if, in repeated random samplings, $100(1-\alpha)\%$ of them will contain the true value of θ
    - L and U are the **lower and upper confidence limits**, respectively
    - $(1-\alpha)$ is called the **confidence coefficient**
- Confidence intervals have a frequency interpretation
  - if the statement is true for a specific sample is not known
  - the method yields correct statements $100(1-\alpha)\%$ of the time

Confidence interval for the Portland Cement mortar

- The formal solution of the problem is:

$$P\left(-t_{N_1+N_2-2,\alpha/2} \leq \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{N_1}+\frac{1}{N_2}}} \leq t_{N_1+N_2-2,\alpha/2}\right) = 1-\alpha$$

$$(\bar{x}_1-\bar{x}_2) - t_{N_1+N_2-2,\alpha/2} S_p\sqrt{\frac{1}{N_1}+\frac{1}{N_2}} \leq \mu_1-\mu_2 \leq (\bar{x}_1-\bar{x}_2) + t_{N_1+N_2-2,\alpha/2} S_p\sqrt{\frac{1}{N_1}+\frac{1}{N_2}}$$

- This means that:
  - if $\mu_1-\mu_2 \leq -0.5451$ or $\mu_1-\mu_2 \geq -0.0109$, we can say that $\mu_1-\mu_2 \neq 0$
  - $\mu_1-\mu_2 = 0$, otherwise, if the interval in which $\mu_1-\mu_2$ is defined comprises zero:

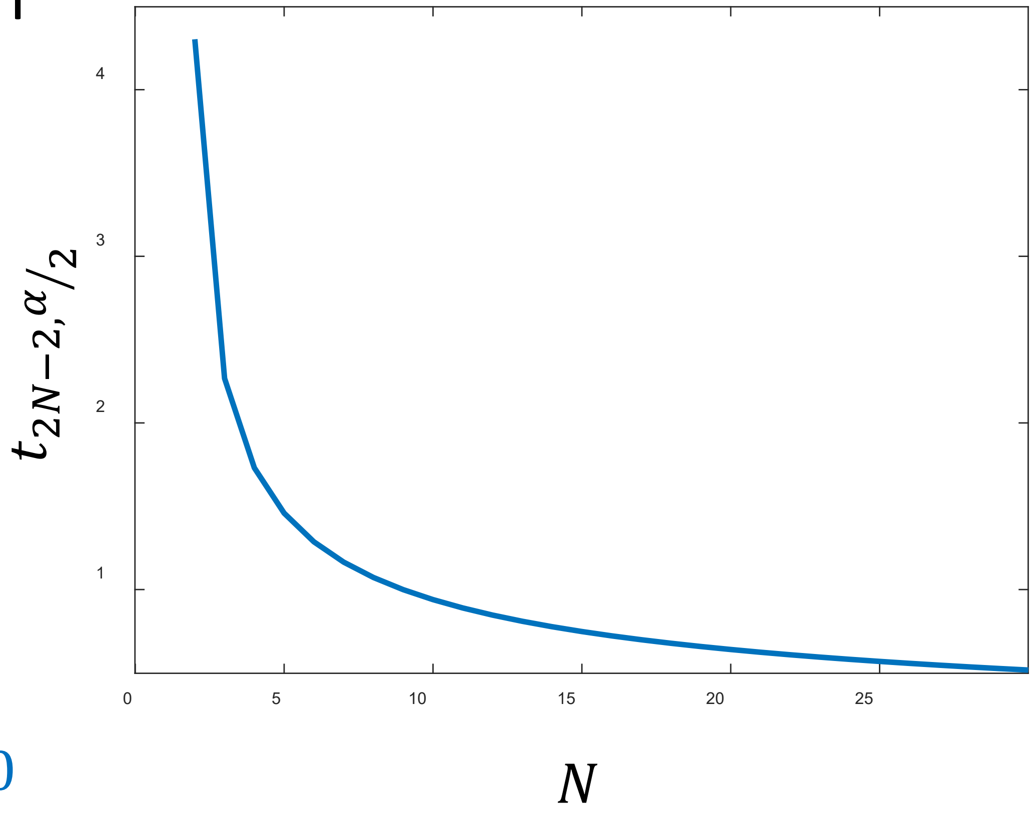
$$-0.5451 \leq \mu_1-\mu_2 \leq -0.0109$$

Sample size from confidence intervals

- The size of the considered sample is one of the most important choices in the experimentation
- Assume that $N_1 = N_2 = N$, the confidence interval is a function of S_p and N (where S_p cannot be controlled):

$$t_{2N-2,\alpha/2} S_p \sqrt{\frac{2}{N}} = f\left(S_p, N\right)$$

  - this function can be represented as a variable of N
  - this function does not vary too much if $N > 8 - 10$



Next lesson: flipped lecture

- The next lecture will be a **flipped classroom**
  - the learning procedure is inverted
- Please, complete the following procedure before Lesson #21:
  1. **attend the video lecture #20 available in Moodle**
  2. **self-assess your learning**
  3. **read the following book chapter**
     - Montgomery, D. (2013). Design and Analysis of Experiments. *J. Wiley & Sons.*
       - Chapter 3
  4. **prepare questions** and anything you need to discuss with the teacher and your mates in the next lecture
- Next lecture will be held in the following manner:
  - 1/2 Q&A:
    - questions (of the students) and answers (of the teacher)
  - 1/2 DoE example

… per sempre a fianco a me!