



ME 794

Statistical Design of Experiments

Chapter 2

Classical Design of Experiments

Importance of Experiments (Recap)



Why experimentation?

- To understand cause-and-effect relationships in a system/process.

Goal of experimentation:

- to determine which input variables are responsible for the observed changes in the response.
- to develop a model relating the response to the important input variables
- to use this model for process or system improvement or other decision-making.

Test: Each experimental run is a test

Experimentation plays an important role in technology commercialization and product realization activities.

Models: i) physics based: mechanistic model, ii) experiment based: empirical model

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

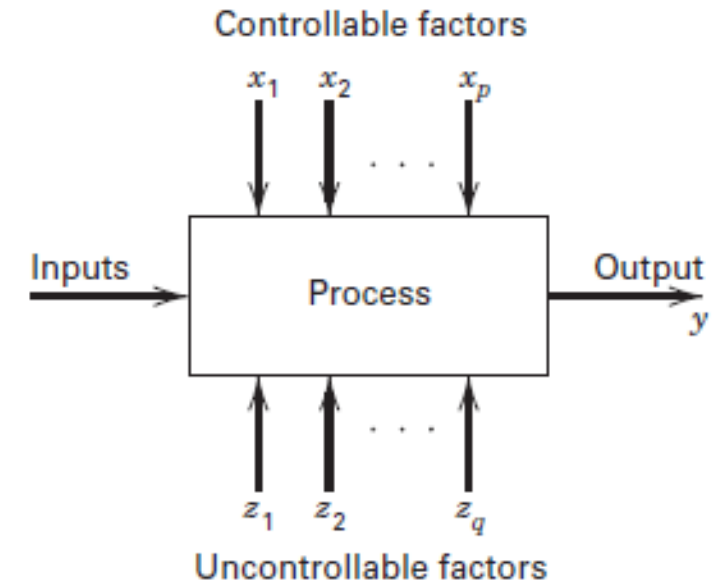
Objectives of Experiments



In general, experiments are used to study the performance of processes and systems.

The objectives of the experiment:

- Determining which variables are **most influential** on the response y
- Determining **where to set** the influential x 's so that y is almost always near the desired nominal value
- Determining where to set the influential x 's so that **variability in y is small**
- Determining where to set the influential x 's so that the **effects of the uncontrollable** variables z_1, z_2, \dots, z_q are minimized.



NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

Applications (Recap)



Experimental design is a critically important tool in the scientific and engineering world for improving the product realization process.

Process development

- Improved process yields
- Reduced variability and closer conformance to nominal or target requirements
- Reduced development time
- Reduced overall costs.

Design activities

- Evaluation and comparison of basic design configurations
- Evaluation of material alternatives
- Selection of design parameters so that the product will work well under a wide variety of field conditions, that is, so that the product is robust
- Determination of key product design parameters that impact product performance
- Formulation of new products.

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

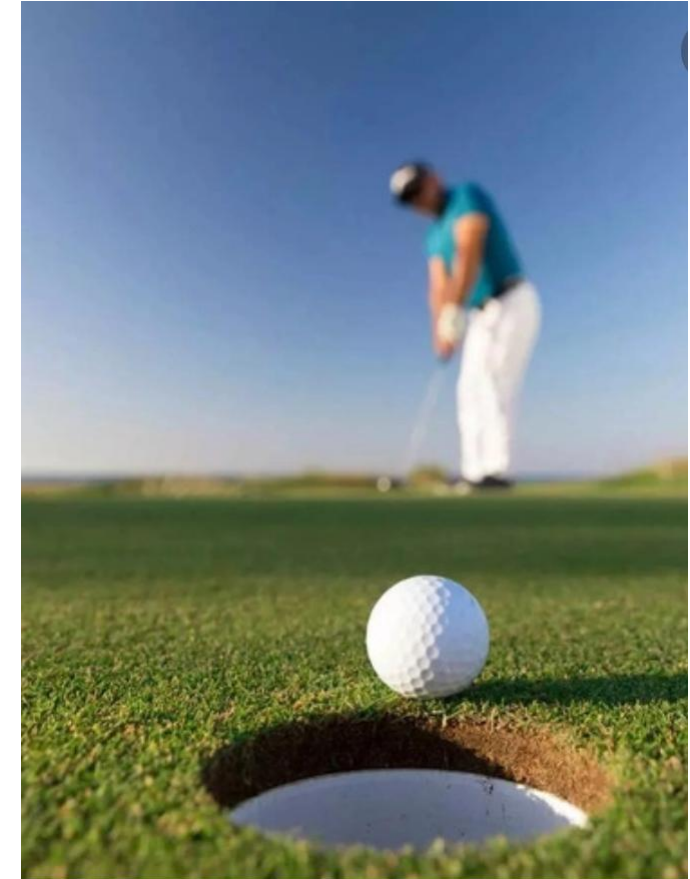
Example



Objective: Lowering golf score

Some of the factors that may be important, or that may influence golf score, are as follows:

1. The type of driver used (oversized or regular-sized)
2. The type of ball used (balata or three-piece)
3. Walking and carrying the golf clubs or riding in a golf cart
4. Drinking water or drinking “something else” while playing
5. Playing in the morning or playing in the afternoon
6. Playing when it is cool or playing when it is hot
7. The type of golf shoe spike worn (metal or soft)
8. Playing on a windy day or playing on a calm day.



NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

Statistical Design of Experiments



Statistical design of experiments refers to the process of planning the experiment so that appropriate data will be collected and analyzed by statistical methods, resulting in valid and objective conclusions.

There are two aspects to any experimental problem:

- The **design** of the experiment and
- The statistical **analysis** of the data.

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

Guidelines for Experiments



Guidelines for Designing an Experiment

- | | | |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---|------------------------------|
| <ol style="list-style-type: none">1. Recognition of and statement of the problem2. Selection of the response variable^a3. Choice of factors, levels, and ranges^a4. Choice of experimental design5. Performing the experiment6. Statistical analysis of the data7. Conclusions and recommendations | } | Pre-experimental
planning |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---|------------------------------|
-

^aIn practice, steps 2 and 3 are often done simultaneously or in reverse order.

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

- A clear statement of the problem often contributes substantially to better understanding of the phenomenon being studied and the final solution of the problem.
- Some (but by no means all) of the reasons for running experiments include:
 - a) Factor screening or characterization.
 - b) Optimization
 - c) Confirmation
 - d) Discovery
 - e) Robustness
- A **sequential approach** employing a series of smaller experiments, each with a specific objective, such as factor screening, is a better strategy.

ME 794 Statistical Design of Experiments

- In selecting the response variable, the experimenter should be certain that this variable really provides useful information about the process under study.
- Most often, the average or standard deviation (or both) of the measured characteristic will be the response variable.
- Multiple responses are NOT unusual.
- It is usually critically important to identify issues related to defining the responses of interest and how they are to be measured before conducting the experiment.

ME 794 Statistical Design of Experiments

3. Choice of factors, levels, and range.

- Factors can be classified as either potential design factors or nuisance factors.
- Potential design factors: design factors, held-constant factors, and allowed-to-vary factors.
- Nuisance factors, on the other hand, may have large effects that must be accounted for, yet we may not be interested in them in the context of the present experiment.
- Nuisance factors are often classified as controllable, uncontrollable, or noise factors.
- Next, the ranges over which these factors will be varied and the specific levels at which runs will be made.

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

3. Choice of factors, levels, and range.

- The cause-and-effect diagram (**fishbone diagram**) can be a useful technique for organizing some of the information generated in pre-experimental planning.

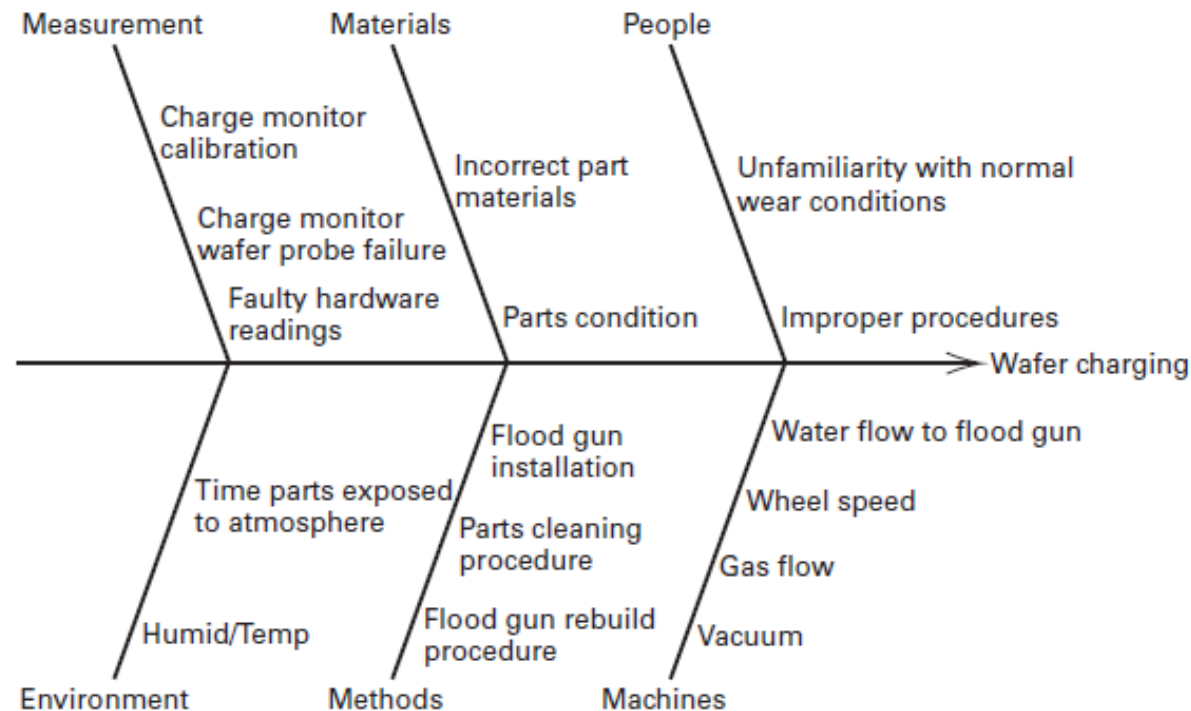


Fig. A cause-and-effect diagram for the etching process experiment

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

4. Choice of experimental design

- Design selection also involves thinking about and selecting a tentative empirical model to describe the results.
- Some examples of models:

- First order model: two variables with main effects

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

- First order model with interaction

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$$

- Second order model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon$$

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

5. Performing the experiment

- When running the experiment, it is vital to monitor the process carefully to ensure that everything is being done according to plan.
- Someone should be assigned to check factor settings before each run.
- Coleman and Montgomery (1993) suggest that prior to conducting the experiment a few trial runs or pilot runs are often helpful.
- In initiating an experimental study, we should probably plan in the first experiment to do **no more than 25% of the total experiments** we have resources for
- One big experiment is not only inefficient but leaves us with no resources left if our initial conjecture was in error.
- We should begin by looking at many factors in a somewhat superficial fashion and more toward the examination of only the few, most relevant, factors in a more comprehensive fashion.

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

5. Performing the experiment

- The three basic principles of experimental design are:
 - Blocking
 - Randomization
 - Replication

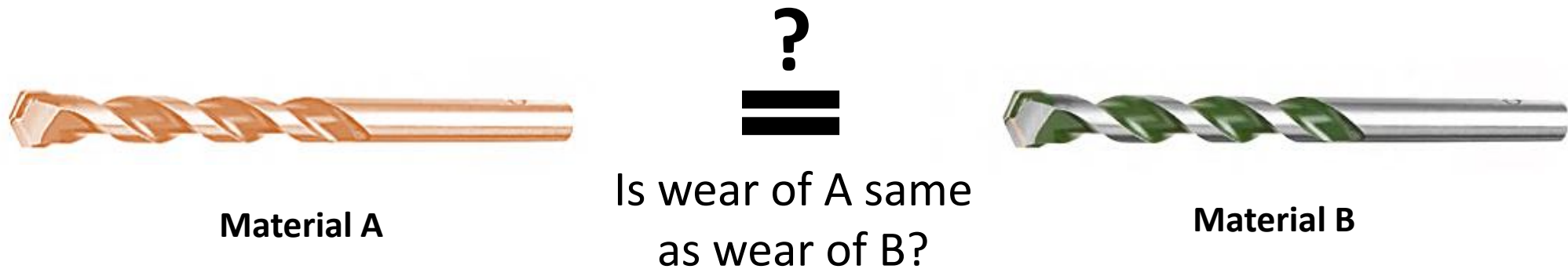
NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

Guidelines: Blocking



- When **known sources of extraneous/unwanted variation** can be identified, we can design the experiment in such a way as to eliminate their influence and provide more sensitive test of significance test

Example: Imagine two tool materials being tested to determine whether or not there is a real difference in their wear characteristics

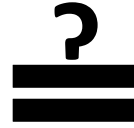


NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

Guidelines: Blocking



Material A



Is wear of A same
as wear of B?



Material B

Experimental Design:

- Twenty tools are made, ten with material A and ten with material B.
- Twenty machine operators are chosen at random, given the tools, and told to use them in machining as they normally do
- At the end of the experiments the mean amount of wear is determined based on the ten measurements for each material type.
- The mean difference is calculated and examined by a statistical test of significance.
- No real difference is found

Do you agree with the results? OR Do you see any problems with this experiment?

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

Guidelines: Blocking



Problems with First Experiments:

- The twenty operators are very different in terms of their skills and experiences
- Hence, a sizeable nuisance variation is introduced within the measurements comprising the mean wear for each material.
- This nuisance variation markedly increases the level of “chance” variation and hence may be “hiding” the presence of a real difference in materials

Revised Experiment:

- Give each operator a pair of tools with one tool of material A and one tool of material B.
- At the end of experiment measure the wear on each tool for each operator and calculate the difference in wear within each operator.
- Average these differences across all operators and perform a test of significance on this average difference.
- **Operator-to-operator nuisance variation is in this way blocked** from consideration since only relative differences within each operator block are examined

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

Guidelines: Randomization



- Unknown sources of nuisance variation are always present but hard to identify – trends from day to day.
- Systematic conduct of tests may violate basic statistical assumptions for significance tests, i.e., Normality/Independence.
- Random sampling imply that the data are independently distributed about their respective means.
- **Block for what you can identify, Randomize for what you cannot.**

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

- In the analysis of the results of comparative experiments, it is necessary to assess the magnitude of the test statistic of interest, say a difference in the mean results of two methods, in light of the natural variability inherent in the experiment.
- The question we are really asking is: **could this experimental result--difference--have arisen solely due to chance causes or is there a real difference in the methods?**
- To answer this question **we must measure the level of chance variation**, usually through **genuine test replication**, and then determine the probability that we could have observed the result before us if only these chance causes were at work.

Summary:

- Experiments should be comparative.
- There should be genuine **replication**: replicates can provide an accurate measure of errors.
- Whenever appropriate, **blocking** (pairing) should be used to reduce error.
- **Randomization** is needed for homogeneity or independence.
- Experiments should be designed in such a way as to be able to determine the interaction effects of factors.

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

6. Statistical analysis of data

- Statistical methods should be used to analyze the data so that results and conclusions are objective rather than judgmental in nature.
- Remember that statistical methods cannot prove that a factor (or factors) has a particular effect. They only provide guidelines as to the reliability and validity of results.
- The primary advantage of statistical methods is that they add objectivity to the decision-making process.

7. Conclusions and recommendations

- Once the data have been analyzed, the experimenter must draw practical conclusions about the results and recommend a course of action.
- Follow-up runs and confirmation testing should also be performed to validate the conclusions from the experiment.

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

Hypothesis Testing



india.com

Badam Dhoodh Aside, 5 Reasons to Try Walnut Milk in Winters

Yesterday • Jigyasa Sahay

The Indi
Health be
2 days ago

TOI IndiaTin
10 nutrien

TOI IndiaTin
Here's

7 hours ago

Hindustan Times

Magnesium supports brain health: Here's how

Yesterday



DT&NEXT

Gender-specific brain changes due to ageing make older women feel more pain,...

22 hours ago

The Week

Brain changes behind pain sensitivity may affect older women more, study shows

23 hours ago

ThePrint

Researchers find brain changes behind pain sensitivity affect older women more

2 days ago

Medical Xpress

Research suggests chronic pain is different for males and females

24 Jan • Gillian Rutherford

Full coverage

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

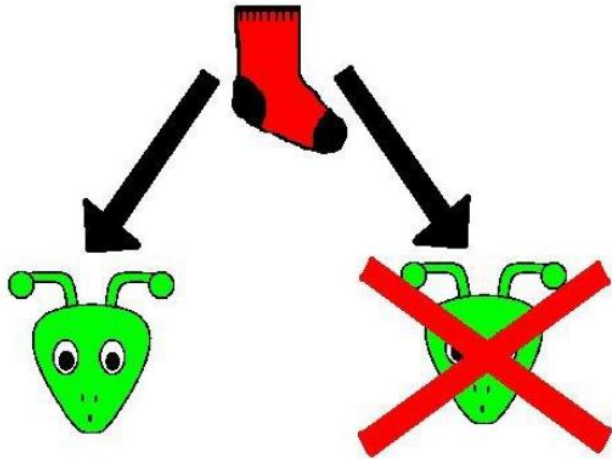
Hypothesis Testing



- Often we are called on to *make decisions* or *draw conclusions* on the new design or improvement in performance of a given process *based on sampled data*.
- In making a decision, we typically *form a hypothesis* concerning what *we believe* is true and then *collect data to prove or disapprove* the hypothesis.
 - “Eating chocolate everyday can lower the risk of heart disease”
 - “Indian population has better immunity against COVID-19”
 - “Music can enhance plant growth” ...
- In statistical hypothesis testing, we generally formulate two hypotheses.
 - **The null hypothesis (H_0):** will be rejected or nullified if the sample data do not support it.
 - **Alternative hypothesis (H_1):** any hypothesis that is different from H_0 is called an alternative hypothesis, denoted by H_1 . ***Any time, H_0 is rejected, H_1 will be considered accepted.***

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

Q. Where have all my socks gone?



Alternate Hypothesis

Null Hypothesis

=

Extra-terrestrial beings have transported themselves into my house in order to steal my socks.

=

Aliens are not to blame. There is some other explanation for the disappearing socks.

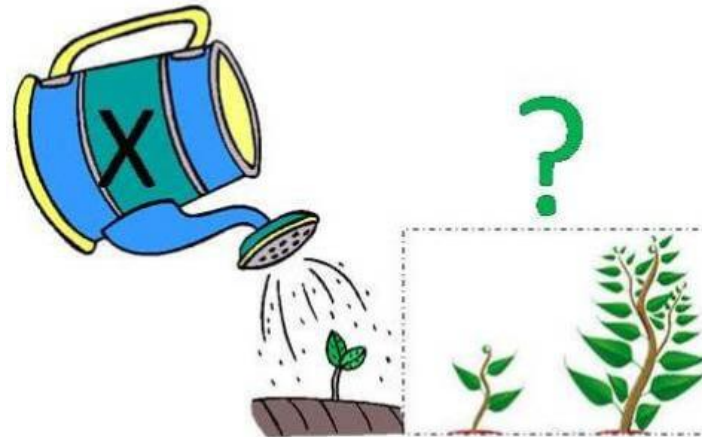
https://www.jcu.edu.au/__data/assets/pdf_file/0018/115344/Basic-Statistics-8_Hypothesis-Testing.pdf

× Effect of Bio-fertilizer 'x' on Plant growth

www.majordifferences.com

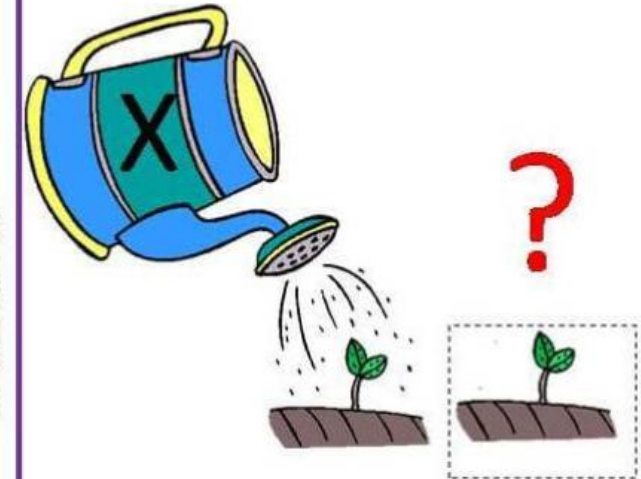
Alternative Hypothesis

H_1 : Application of bio-fertilizer 'x' increase plant growth.



Null Hypothesis

H_0 : Application of bio-fertilizer 'x' do not increase plant growth.



NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

Hypothesis Testing of Population Means



- There are three ways to set up the alternative hypothesis (H_1)

$$\text{Method 1 : } H_0: \mu_x \leq \mu_0 \quad H_1: \mu > \mu_0$$

$$\text{Method 2 ; } H_0: \mu_x \geq \mu_0 \quad H_1: \mu < \mu_0$$

$$\text{Method 3 : } H_0: \mu_x = \mu_0 \quad H_1: \mu \neq \mu_0$$

A statistical hypothesis test consists of the following six steps:

1. State the null and alternative hypotheses. Define the test statistic used to analyze the situation.
2. Determine the significance level, ' α ', at which the test will be made.
3. Collect the data and calculate the test statistic result.
4. Define the reference distribution for the test statistic.
5. Compare the test statistic and its reference distribution under H_0 . Carry out the necessary analysis of data.
6. Assess the risk.

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

Errors and Significance Level



Two kinds of errors may be committed when testing hypotheses. If the null hypothesis is rejected when it is true, a type I error has occurred. If the null hypothesis is *not* rejected when it is false, a type II error has been made. The probabilities of these two errors are given special symbols

$$\alpha = P(\text{type I error}) = P(\text{reject } H_0 | H_0 \text{ is true})$$
$$\beta = P(\text{type II error}) = P(\text{fail to reject } H_0 | H_0 \text{ is false})$$

Sometimes it is more convenient to work with the **power** of the test, where

$$\text{Power} = 1 - \beta = P(\text{reject } H_0 | H_0 \text{ is false})$$

The general procedure in hypothesis testing is to specify a value of the probability of type I error α , often called the **significance level** of the test, and then design the test procedure so that the probability of type II error β has a suitably small value.

Type I and Type II Error

Null hypothesis is ...	True	False
Rejected	Type I error False positive Probability = α	Correct decision True positive Probability = $1 - \beta$
Not rejected	Correct decision True negative Probability = $1 - \alpha$	Type II error False negative Probability = β



The power of a binary hypothesis test is the probability that the test correctly rejects the null hypothesis when a specific alternative hypothesis is true. It represents the chances of a true positive detection conditional on the actual existence of an effect to detect.

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.

Significance Level

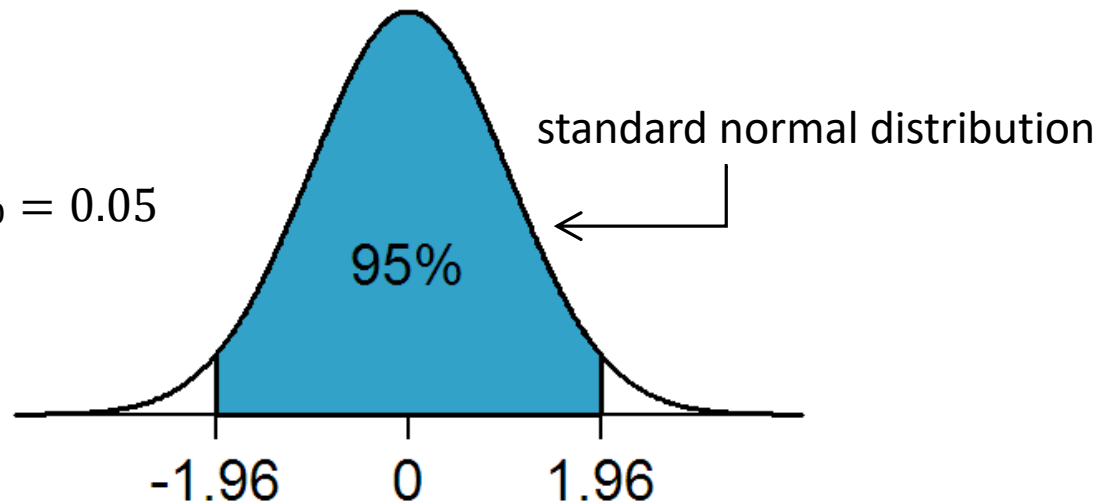


In statistical hypothesis testing, a result has *statistical significance* when it is very unlikely to have occurred given the null hypothesis.

- More precisely, a study's defined *significance level, α* , is the *probability of the study rejecting the null hypothesis*, given that the null hypothesis was assumed to be true
- **p-value** of a result, is the probability of obtaining a result at least as extreme, *given that the null hypothesis is true*

The result is statistically significant when $p \leq \alpha$

Example, when $\alpha = 5\% = 0.05$



TRY THIS:

<https://www.mathsisfun.com/data/standard-normal-distribution-table.html>

NOTE: You do NOT have permission to share this file or any of its contents with anyone else, and/or upload it on internet or any of the platforms where it can be accessed by others.