بسم الله الرحمن الرحيم

السلام عليكم ورحمة الله وبركاته

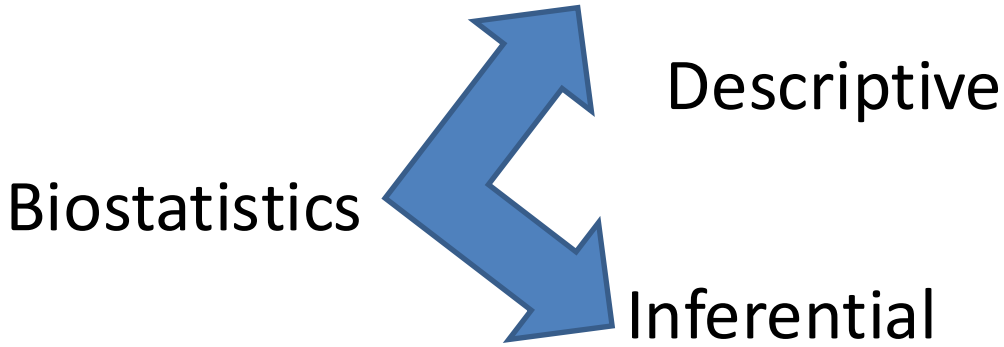
Biostatistics

Inferential analysis

L IX

PROF. DR. WAQAR AL-KUBAISY

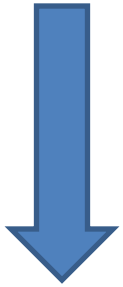
18-8-2024



Sample: **mean ± S.D** **sample statistic**
sample estimate

Population: **population mean μ ±S.E** **population parameter**

Sound generalized information about the population from which the sample has been drawn depending on the evidence of the sample .

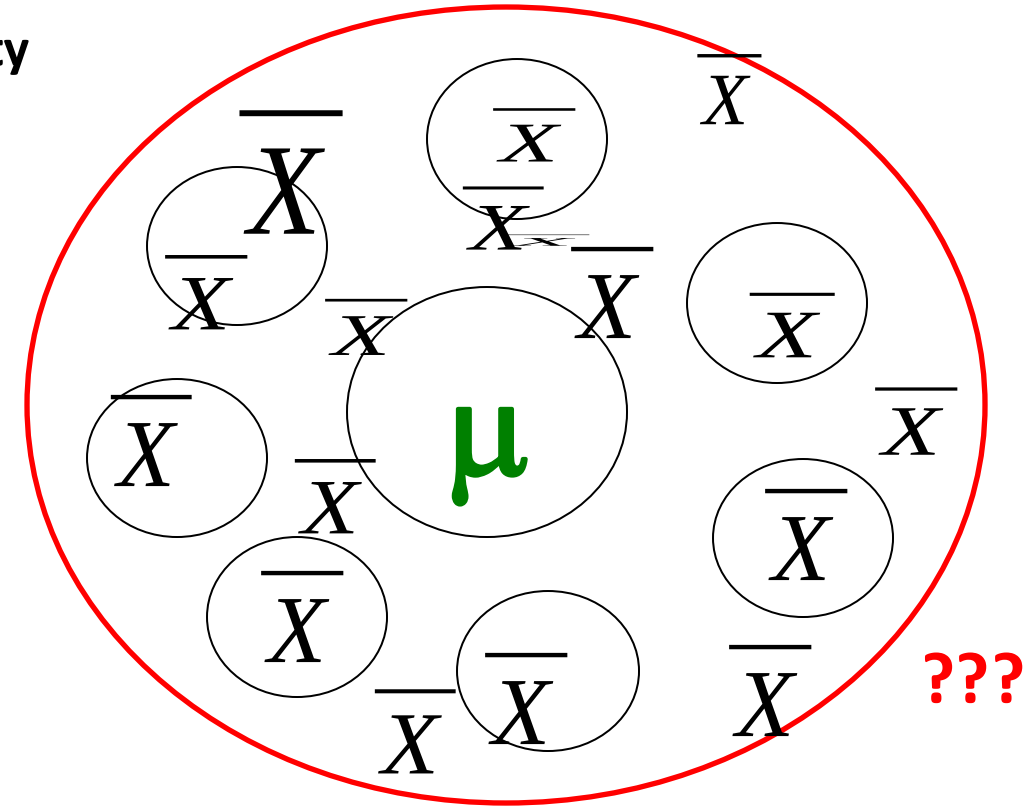


Inferential Biostatistics (Analysis).

It is used to test specific hypothesis about population by

using certain test significance

Cont. …Sampling Variability

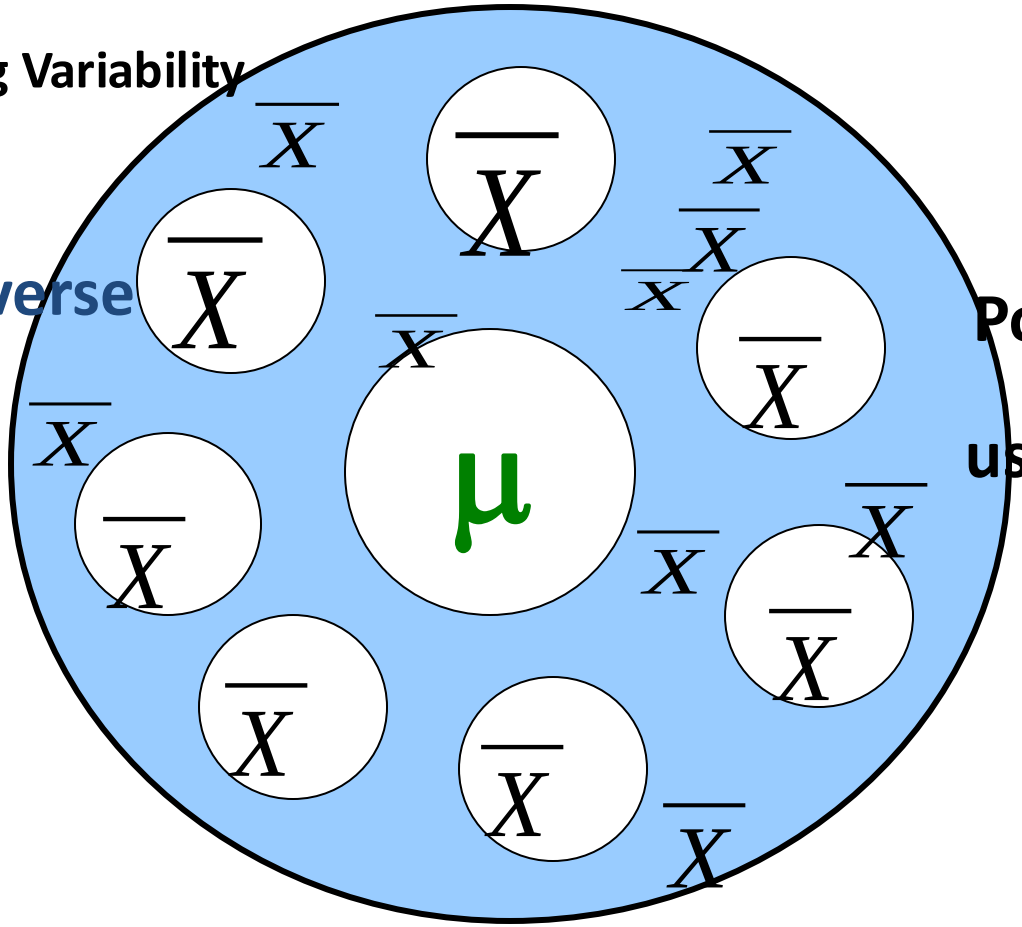


Different samples → different $\overline{X}_S$ even if the samples size are equal

There is a variation in the $\overline{X}_S$ of different samples
This variation is due to sampling variation.

Cont. …Sampling Variability

mean of universe
true mean



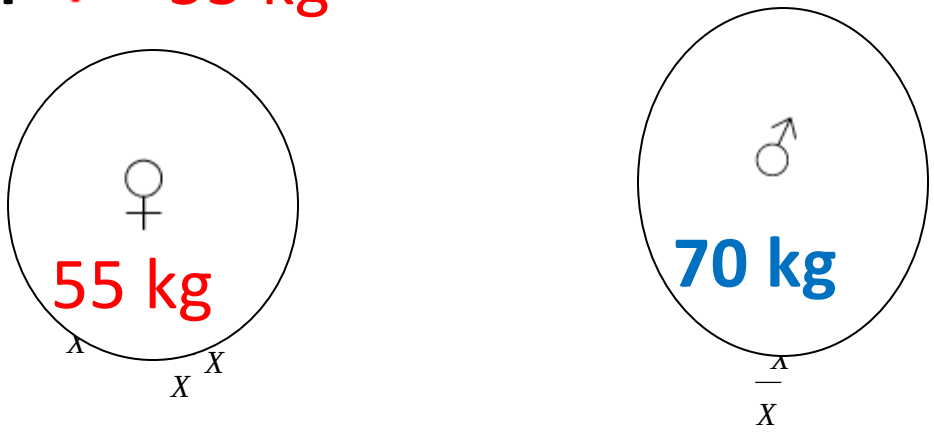
Population mean
is
usually unknown

the sample measurement (mean± S.D) is
not exactly reflect its population .
There is a difference between sample mean $\overline{X}$ and
population mean μ

We expect always that there is a difference between groups .

Mean body weight of ♂ = 70 kg

Mean body weight of ♀ = 55 kg



Mean body weight of grope I (Jordanian) = 65 kg .

Mean body weight of grope II (Iraqi) =60 kg .

Difference could be

influencing factor

Chance factor

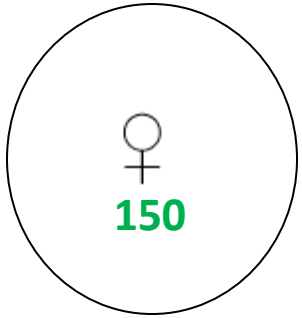
- **So we expect always that, there is difference .**
- **And by using these test of significance,**
- we assess whether **that difference between groups is caused** by
  - **specific factor, that we are interest about** or it
  - caused by **chance factor?**

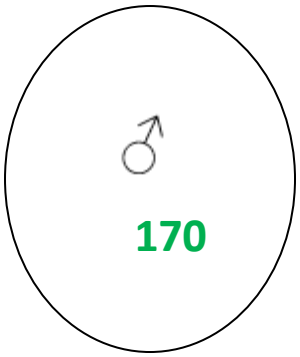
Is the difference caused by **variation of sex** ??

Or it is due **to chance factor** .

So we are testing the significance effect of the sex on the mean body weight .

Or the influence of sex on the body weight of human .





150/170=88.23%

170/200=85 %

- Inferential statistics is used to test specific hypothesis by certain test of significance .

The purpose of testing hypothesis is to aid the clinician, researcher, administer in reaching a decision concerning population, basis on examination of sample from that pop. .

Hypothesis

A statement about one or more population .

Hypothesis is usually **concerned with the parameter of pop.** about which the statistics is made .

Drug A is better than drug B.

COVID 19 infection more in Health Care Workers (HCWs)

- **So by mean** of **hypothesis testing**
- we are going to decide or determine **whether or not such statement is compatible with available data in**
- **sample , through using appropriate test of sign .**

I- **Data Nature of data** 

II- Assumption * **Random sample (RS).**

* **Independent or dependent R .S .**
* **Equal variance (various equality) .**
* **Normality of pop. Distribution .**

III. Hypothesis formulation

Formulate two statistical hypothesis simultaneously ;

1. **Null hypothesis (H_O)**
2. **Alternative hypothesis (H_A).**

Null hypothesis (H_O)

- **Hypothesis of no difference** .

Since it is a statement of agreement with true condition in the population of interest.

- **Consequently the opposite of the conclusion that the researcher is seeking to reach, become the statement of the null hypothesis** .

In H_0 it states always that, there is **no significance difference or there is no influence or effect of influencing factor .**

In testing hypothesis process , **the H_O is either**

Reject or

Not reject (accept) .

Difference could be ⟷ **influencing factor** / **Chance factor**

If H_o not rejected, we will say, that, the data in our hand

~~influencing factor~~

Difference could be

Chance factor

If H_o not rejected,

we will say that, the data in our hand (or which the test is based on) **not provide sufficient evidence to cause rejection**

If testing procedure leads **to rejection**, we will conclude that, the data in our hand **are not compatible with H_o** . **But supporting of some other hypothesis** .

this hypothesis is known as

- **Alternative hypothesis (H_A).**

Difference could be

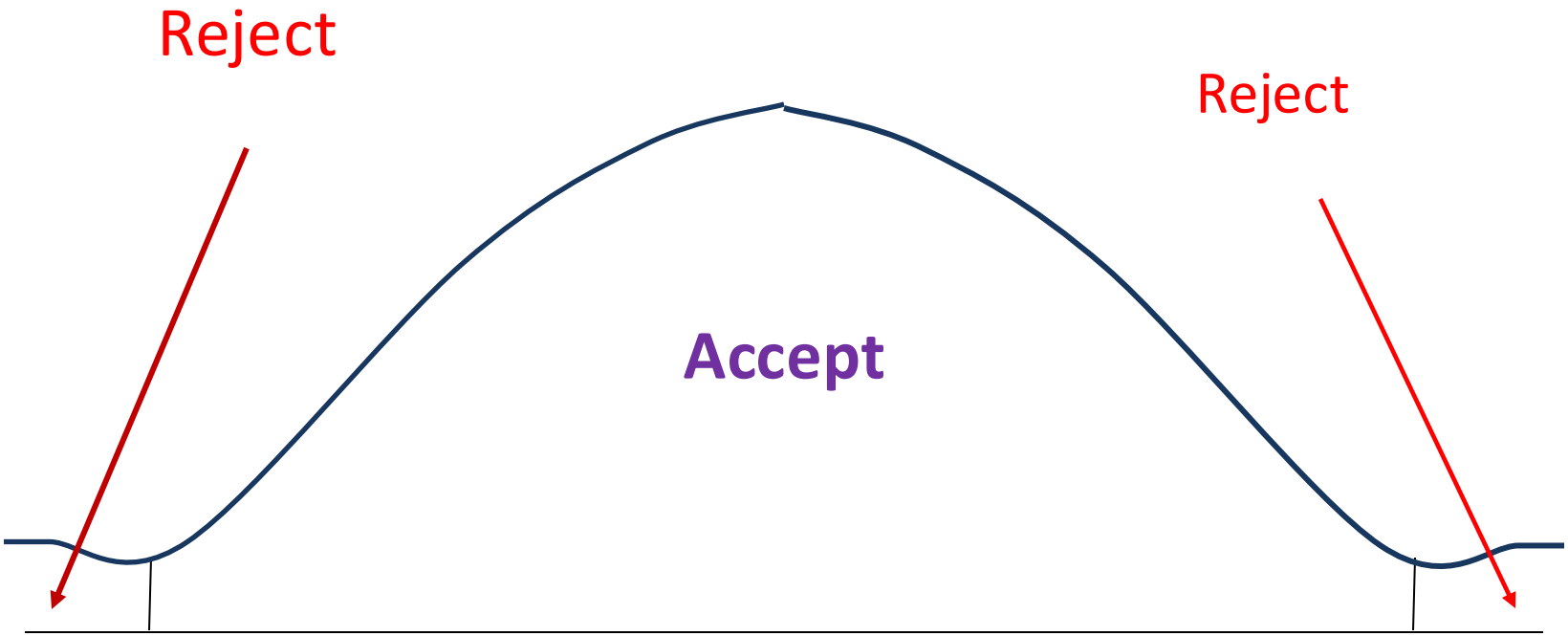
influencing factor

~~Chance factor~~

The decision, to reject or accept the H_o depends on the magnitude (value) of the test statistics

Test statistics

Serve as a decision maker for rejecting or not rejecting the Null Hypothesis.



The distribution of test statistics, which is the key to the statistical inference **area under the carve divided into two groups or areas:**

*** Rejection area (region)**

*** Acceptance area (region)**

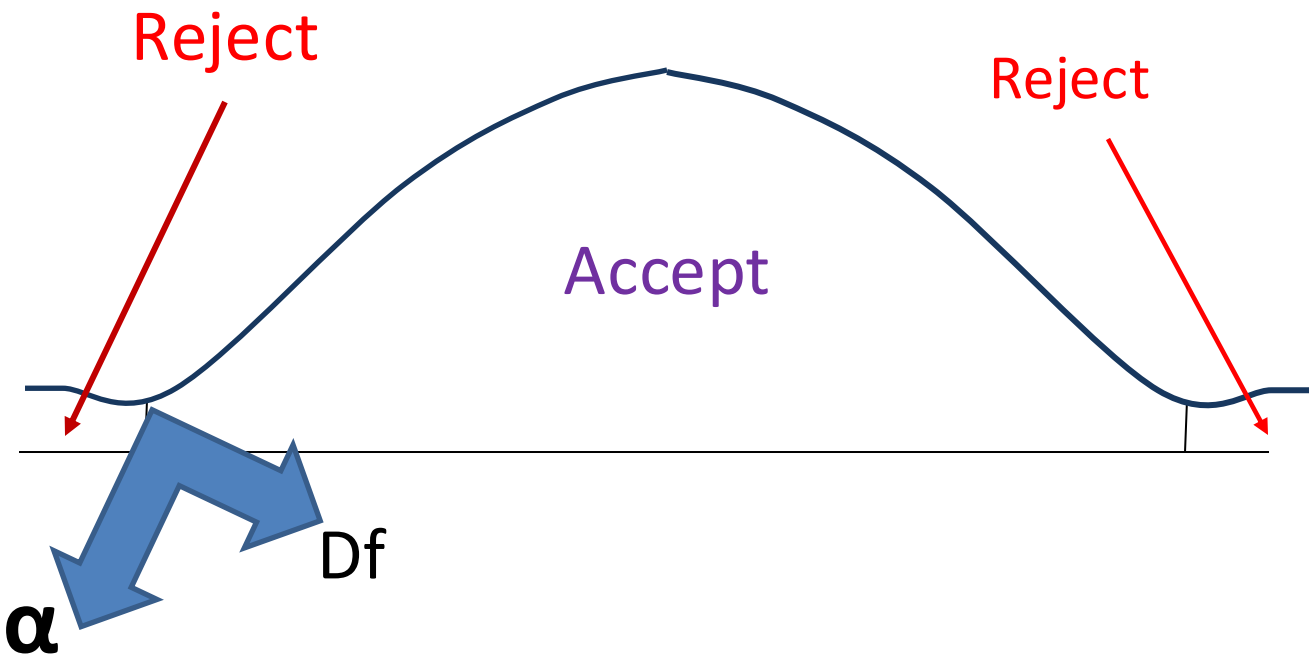
The decision as to which value go into the rejection and which one go to the accept region is made on the basis of the desired level of significance designated by (α) .

So the value of test statistics fall in the rejection region are those that are less likely to occur if H_o is true. While the values making up the accept region are more likely to occur if H_o is true

when Test statistics that fall in the rejection region is said to be significant.

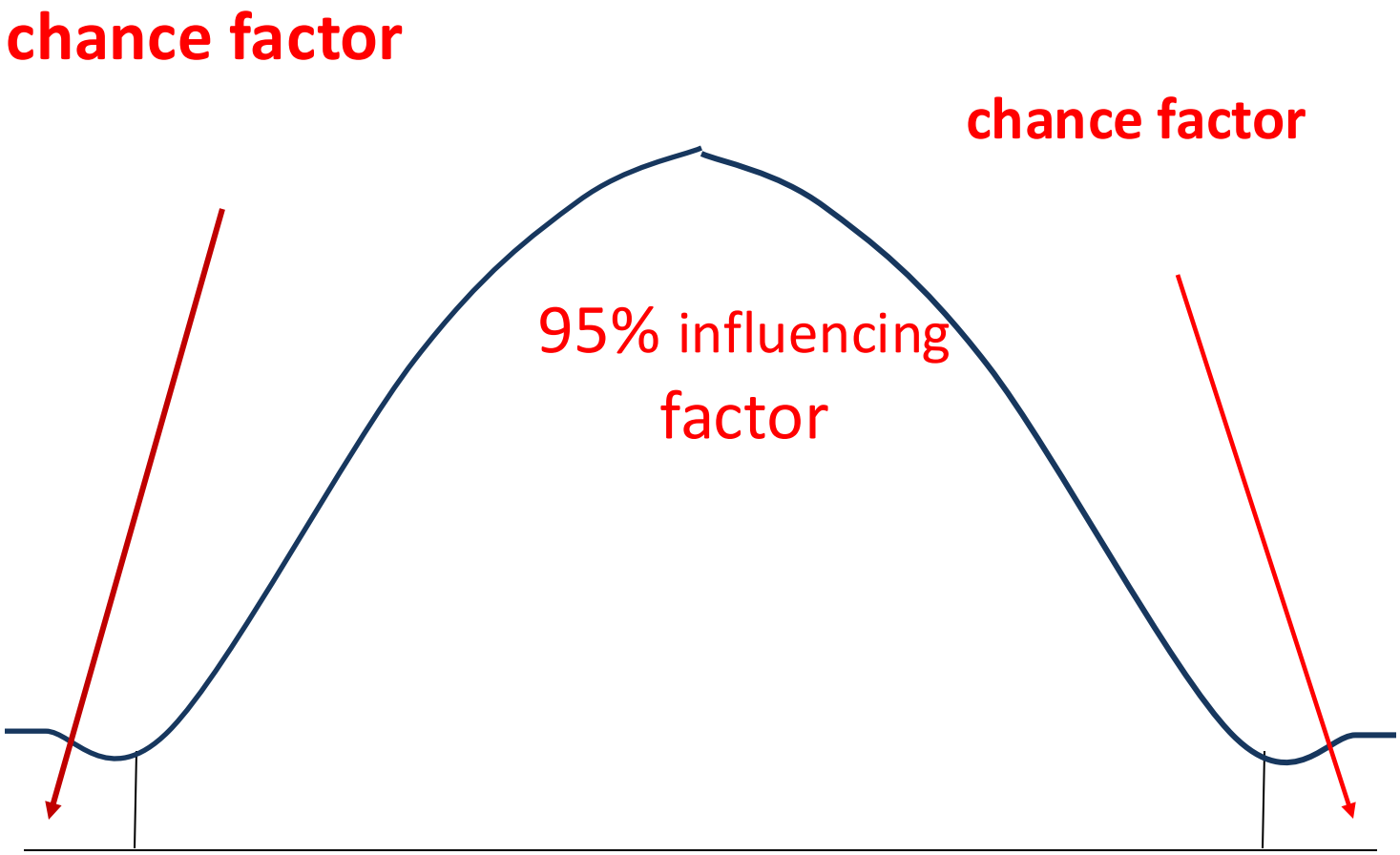
So the level of signify (α) is specify the area under the curve of the distribution of the test statistics.

That is above the value on the horizontal axis constituting the rejection so (α) is probability of rejecting the true H_o .



Define Level of Significance

Level of significance it is the probability level,
According to N.D at which we either accept or reject H_o .
According to N.D.C
we can assume that, 95% of the difference between groups are caused by the influencing factor .
the remaining 5% (**2.5% on each side**) are caused by
chance factor
so in biological research including medical research,
level of significance is 95% (it is probability of influencing the factor understudy)
the remaining 5% is the probability of effect of chance factor it is also called (P value)


chance factor
chance factor
95% influencing factor

Apply The Proper Test of Significance

Compute test statistics for each set of observation (data) or (study), we might use different test of sign.

1-Depending on the **variable** that we deal with Whether data is

Continuous **Discrete .**

2- we will **compute** the value of **test statistics**.

3- **compare** with accept or reject region.

4- Then by using **test of significance**

We will able to quantify (measure) the **amount of (α)** error or **(P)** value

If by using test of sing.

we found that calculated (P) value is larger than 5% (0.05),

this means **that chance factor affect more then 5%,**

in another word,

the influencing factor is affecting the difference less than 95%

in this we accept the H_o ,

or the difference between these groups is not significance.

And

There are chance factor causing the difference beside the influencing factor.

P > 0.05 → accept H_o → no significance difference

This mean that the effect of **influencing factor** is **not significance**.

If the calculated P value is smaller than 5%

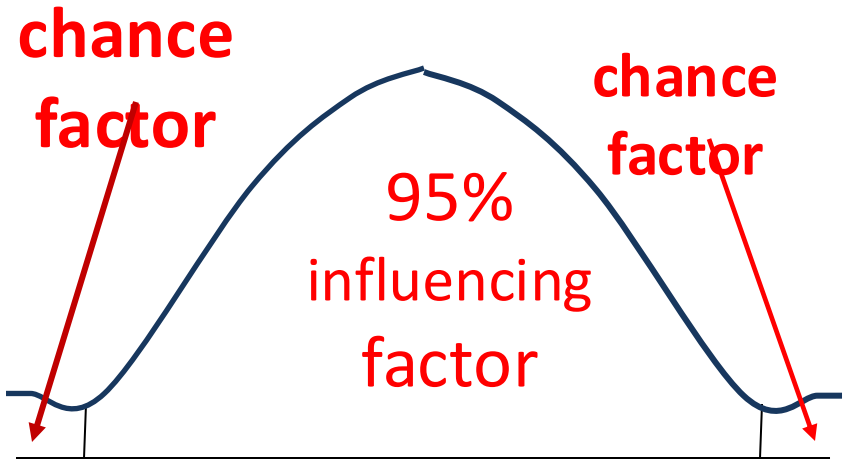
If the calculated P value is smaller than 5% (P<0.05) it means that

the **effect of the factor** under study **is larger than 95%** (0.95)

or the chance **factor is minimal effect**.

This **means that the influencing factor has significant effect**

P < 0.05 → reject H_o → significant difference.



Statistical decision

Statistical decision, consist **of rejecting or not rejecting (accepting) H_o.**

If computed value of test statistical fall in the reject region

or not rejected if the computed value of test stat.
fall in the accept region.

If H_o is rejected clinical decision is compatible to the H_A

.

If H_o is not reject, the clinical decision may take other from such a decision to collect more data .

P value


Thank you!