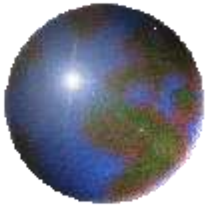
WELCOME



Well-known, so called “p” value..

Ahmet SALTIK, MD

Professor of Public Health

www.ahmetsaltik.net

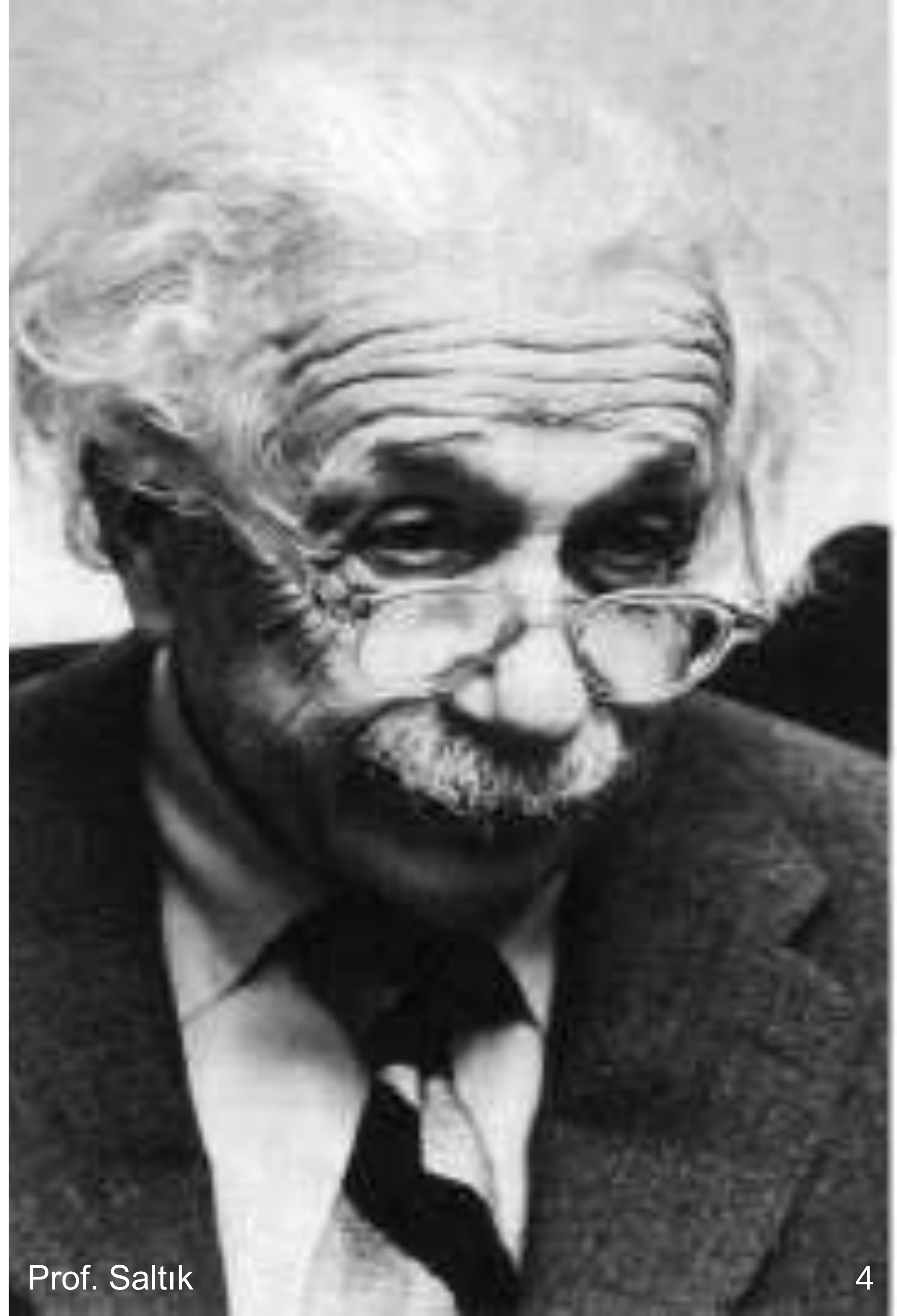
Ankara Univ. Medical School,
Dept. of Public Health

From data to model ??

- “OK, I’ve got some data..
- Now what?”
- **Data = Information + Error** (sampling errors and others such as measurement, recording, tagging etc.)
- <u>The exact action to be rendered :</u> Planning how to utilise each datum *a priori* even so, making ready the empty tables & graphs!

"If I had an hour to solve a problem and my life depended on the solution, I would spend the first 55 minutes determining the proper question to ask, for once I know the proper question, I could solve the problem in less than five minutes"

- Albert Einstein



A statistical test

.. is a procedure for deciding whether an assertion (e.g. a Hypothesis) about a quantitative feature of a population is true or false. We test an hypothesis of this sort by drawing a random sample from the population in question and calculating an appropriate statistic on its items. If, in doing so, we obtain a value of the statistic that would occur rarely when the hypothesis is true, we would have reason to reject the hypothesis.

A statistical test

With this procedure it is customary to reject the hypothesis tested when our statistic has a value that is among those that, theoretically, would be expected to occur no more than 5 out of every 100 times that a *random sample* (*of the same size*) is drawn from the population in question when the **hypothesis** is, in fact, true. Much of the text of this lecture is devoted to explanations of exactly how this kind of theoretical expectation is developed.

A statistical test

Finally, it is noteworthy that
the "*appropriate conduct of any statistical test*"
invariably requires many thoughtful decisions.
It is, for example, always necessary to decide
what statistic to be used,
what sample size to be employed
and what criteria to be established
for rejection of the **hypothesis** tested.

Hypothesis...

- At the end of the experiment, the most important hypothesis that would be tested is whether the difference between the groups treated originates by chance.

$$H_0 : \alpha_i = \alpha_{i'}$$

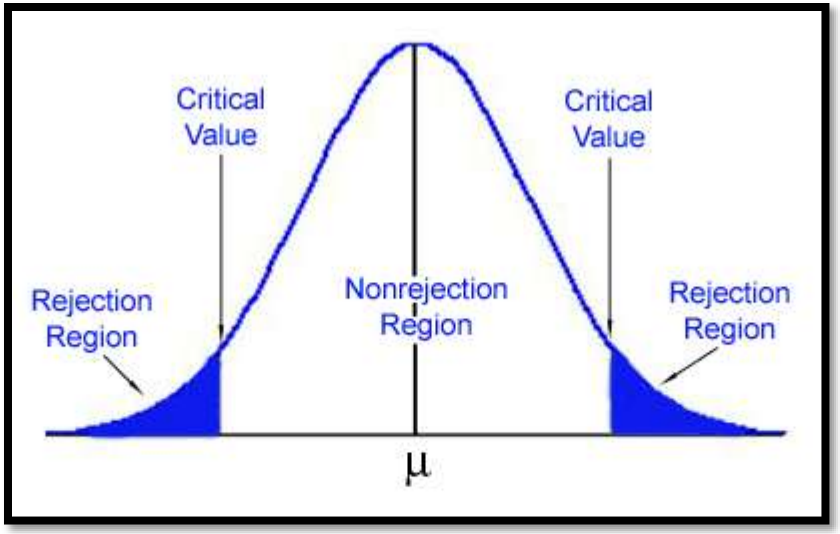
$$H_1 : \alpha_i \neq \alpha_{i'}$$

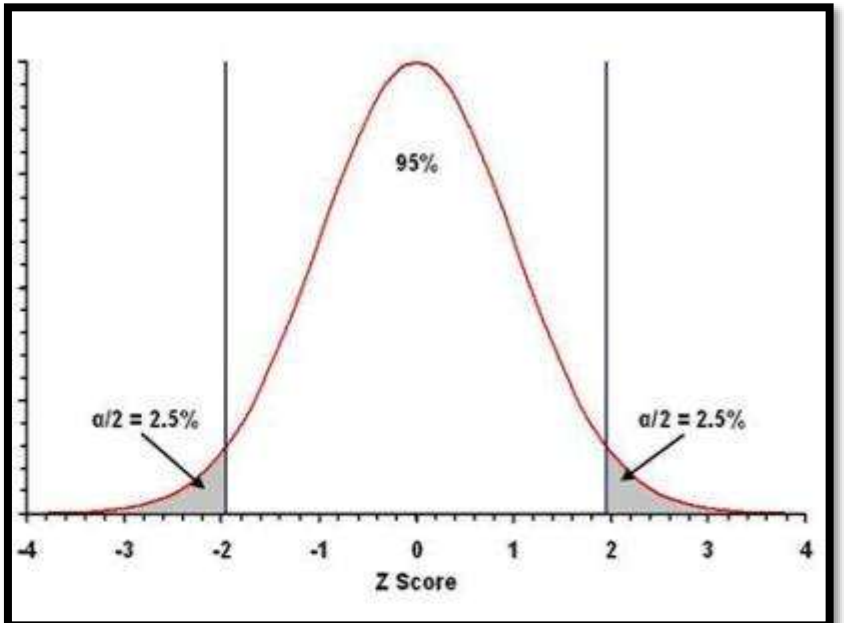
The null hypothesis

- The null hypothesis is a statement infering there is no difference between population parameters. That is, there is no relationship between independent and dependent variables in the population under study. Typically, this is not the anticipated outcome of an experiment. Usually the investigator conducts an experiment because he/she has reason to believe manipulation of the independent variable will influence the dependent variable. So, rejection of the null hypothesis is interpreted as a **significant finding**.

The null hypothesis..(1)

Is a term that statisticians often use to indicate the statistical hypothesis tested. The purpose of most statistical tests, is to determine if the obtained results provide a reason to reject the hypothesis that they are merely a product of chance factors.





The null hypothesis..(2)

For example, in an experiment
in which two groups of randomly selected subjects
have received different treatments
and have yielded different means,
it is always necessary to ask if the difference
between the obtained means is among the
differences that would be expected
to occur by chance whenever
two groups are randomly selected.

The null hypothesis..(3)

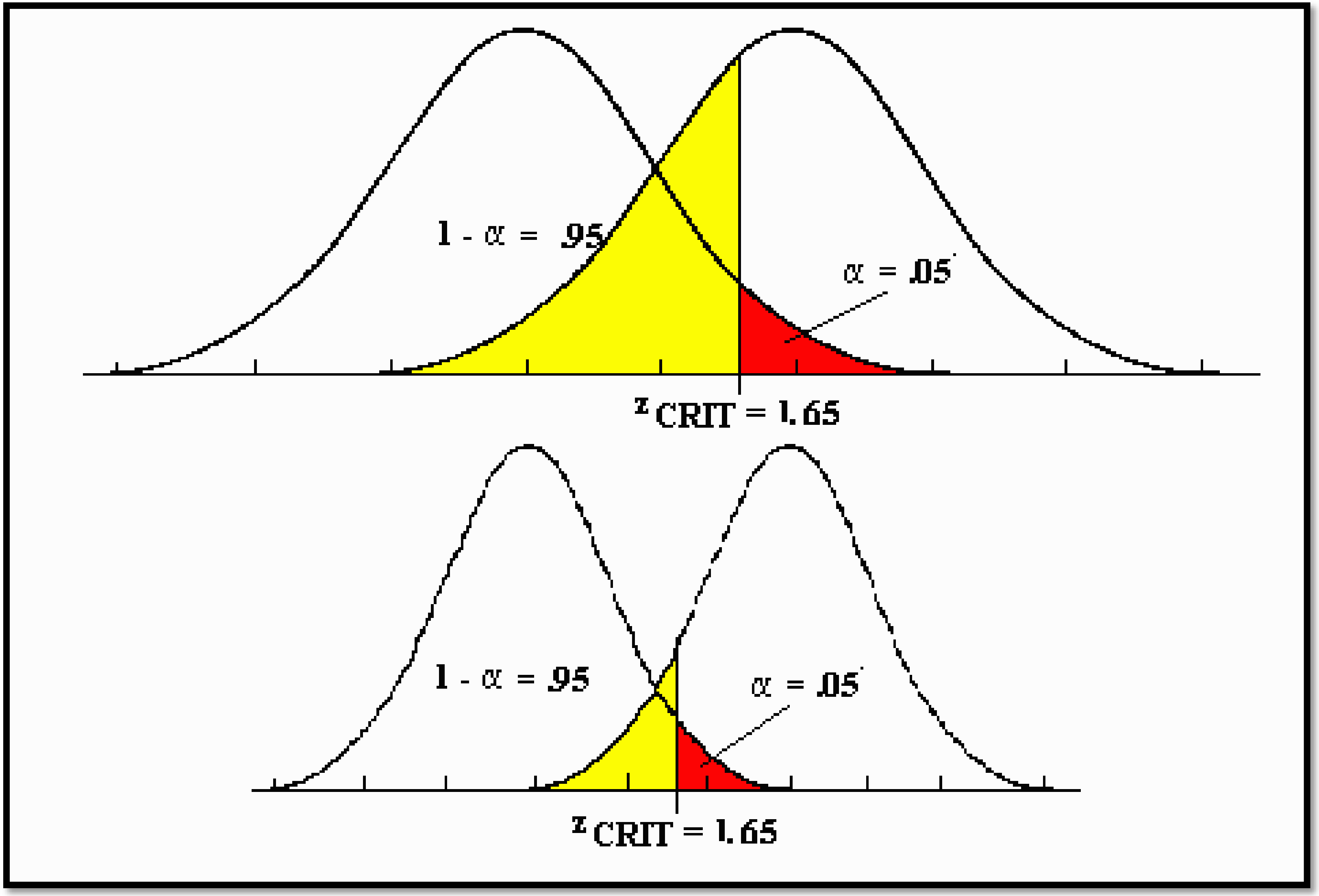
In this example, the hypothesis tested is that the two samples are from populations with the same mean.

Another way to say this is to assert that the investigator tests the null hypothesis that the difference between the means of the populations from which the samples were drawn, is zero.

The null hypothesis..(4)

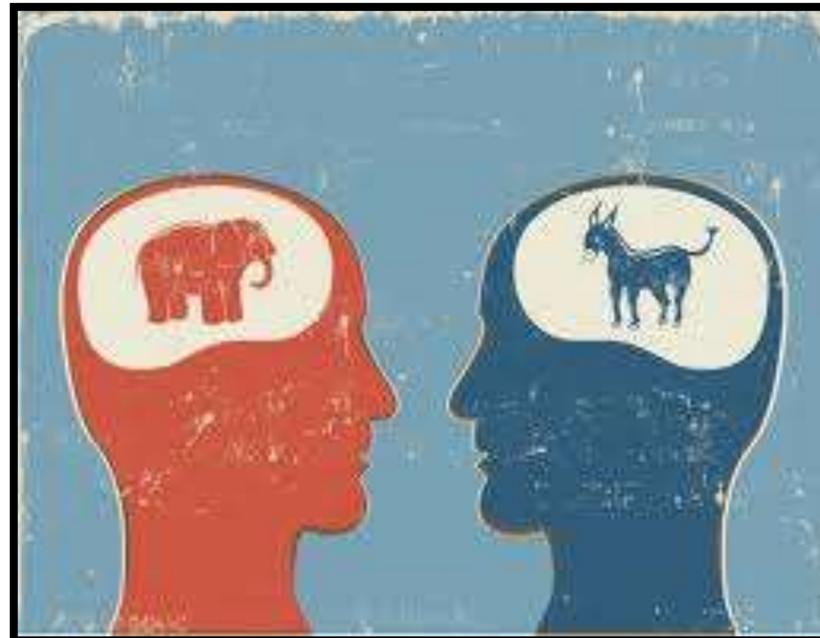
If the difference between the means of the samples is among those that would occur rarely by chance when the null hypothesis is true, the null hypothesis is rejected and the investigator describes the results as **statistically significant**.

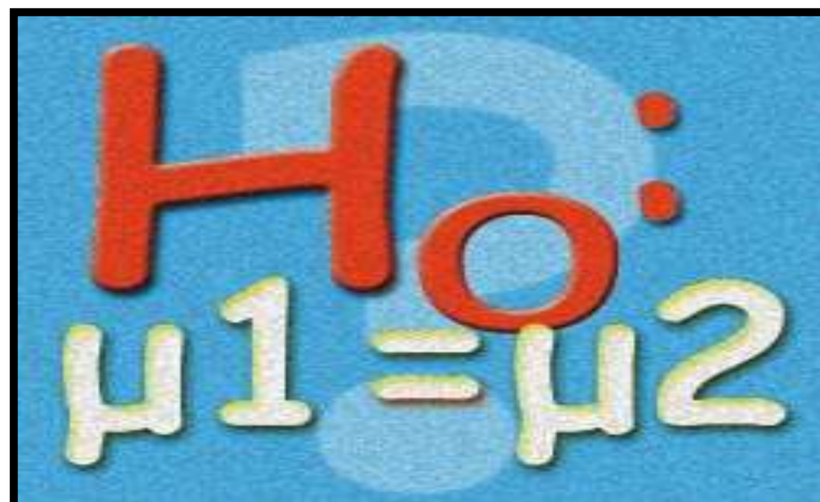
Rejection & Accepting Ho



The alternative hypothesis (1)

The test of a given statistical hypothesis entails an assessment of whether or not our sample (or samples) have yielded a statistic that is among those cases that would only occur α (alpha) proportion of the time if the hypothesis tested is true.



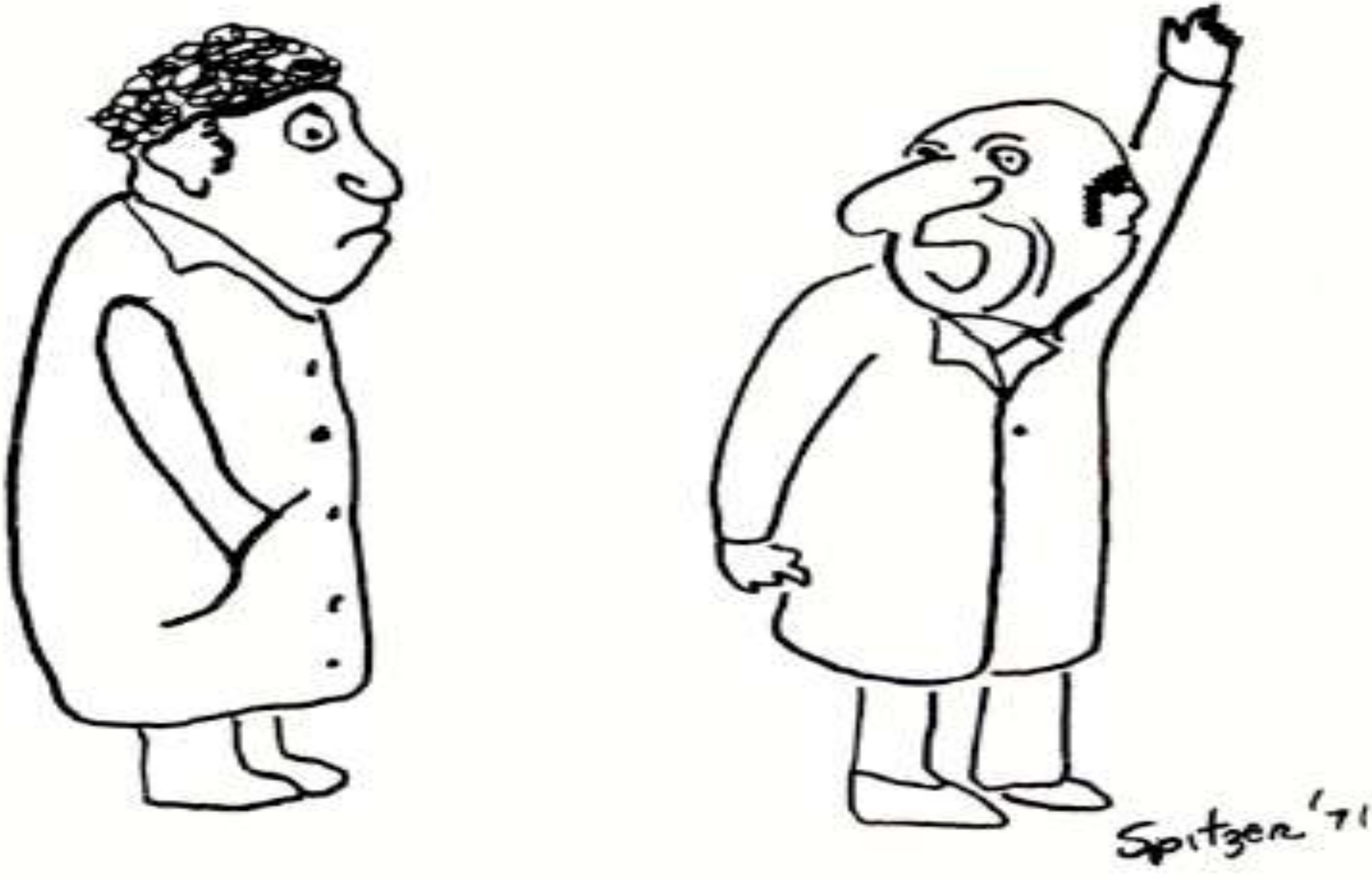


The alternative hypothesis (2)

In these circumstances we know the probability of rejecting the hypothesis tested when it is true *(that probability is equal to **alpha**)* but unless we have also specified an alternative hypothesis to the hypothesis tested, we have no idea of the probability of being in error, if our test has failed to yield a value that enables us to reject the hypothesis tested.

Insufficient methodological actions..

- Calculating sample size and its changing during process..
- Randomization and its details..
- Blind study..
- Carrying out of the Biostatistical evaluation by simple bivariates and making judgements on this basis; lack of in-depth research / investigation



"Four years of research, and now you tell me you forgot which is the control group!"

G. Spitzer, APA Monitor, August 1971.

A simple experiment by Galton

"If peas / chickpeas are shaked with a stable amplitude on a net filter above a box, how should they be heaped over?

- In case of replying with Aristo's minding :
- " *The fact of of the science can be expolerd by thinking and intuition...*"
- Let's try and see by that way : Figure it..
- If not, shall we put OBSERVATIONS and EXPERIMENT in front of our intuitions?
- The unique way for reaching the truth is reasoning and **observational-experimental science**!

Is intuition a resolution for all questions?

Everyone is "*imagining*" the heaping configuration by the peas in a different appearence..

Simply because, everyone's intiution is subjective..

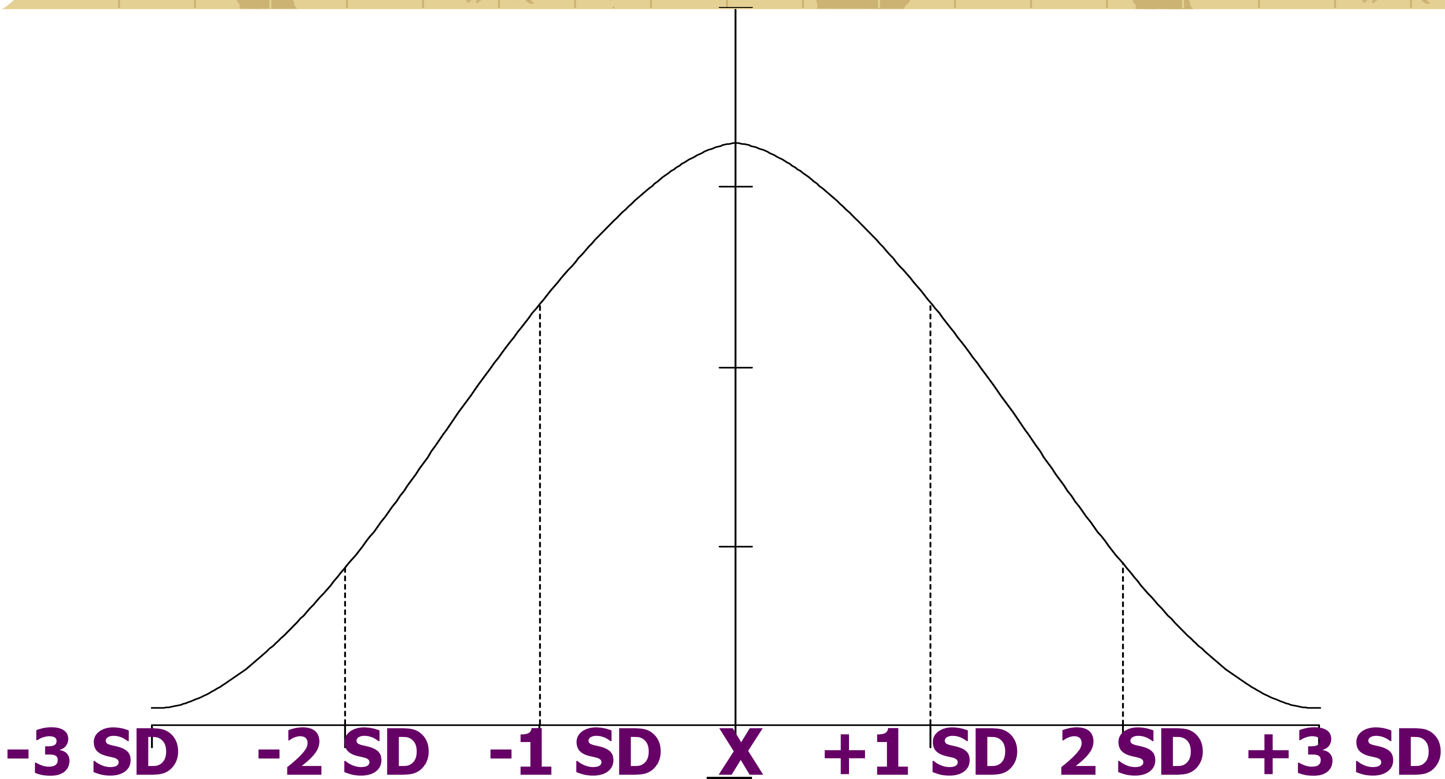
Aristo, with this minding, closed down the science to observation and experiment for 2000 years.

In fact, science *is general and objective*.

Is intuition a resolution for all questions?

Science is desperately to be based on,
repeatable <u>experiment-observation</u>,
confirming or denying itself.
The content; is the relative reality of the science
for which only valid at this moment,
and has to be changed in time.
The thing that will make this possible is,
Methodology
***i.e.* Epidemiology and Biostatistics.**

Galton's peas, formed a bell shape!



Standard normal distribution, is like a "***bell shape***".

It's been proved experimentelly and followed by theoretical calculations. At the two sides of the arithmetic mean (X) ± 1, 2 ve 3 Standard deviations are displayed.

Decisions Based on Hypothesis Tests and Errors

Contingency Options		Decision	
		H_0 rejection	H_0 cannot be rejected
Real situation	H_0 true	False decision Type 1 error α error	True decision 1- α
	H_0 false	True decision (Power of the Test) 1-β	False decision Type 2 error β error

Well known p value (1)

A finding *(for example the observed difference between the means of 2 random samples*) is described as statistically ***significant***, when it can be demonstrated that the probability of obtaining such a difference by chance only, is relatively low.

Well known p value (2)

In many domains, it is customary to describe one's finding as statistically significant, when the obtained result is among those that (theoretically) would occur no more than 5 out of 100 times when the only factors operating are the chance variations that occur whenever random samples are drawn.

Well known p value (3)

Statisticians use the Greek letter **alpha** to indicate the probability of rejecting the statistical hypothesis tested when in fact, that hypothesis is true.

Before conducting any statistical test, it is important to establish a value for **alpha**. For many other scientists, it is customary to set alpha at **0.05**.

Well known p value (4)

This is the equivalent of asserting that you will reject the hypothesis tested if the obtained statistic is among those that would occur only 5 out of 100 times that random samples are drawn from a population in which the hypothesis is true. If your obtained statistic leads you to reject the hypothesis tested, it's not because you believe that the obtained statistic could not have occurred by chance.

Well known p value (5)

It's that you are asserting that the odds of obtaining that statistic by chance only are sufficiently low (one out of twenty) that it reasonable to conclude that your results are not due to chance.

Could you be in error?

Of course you could, but at least you know the probability of such an error.
It is exactly equal to the value you have previously established for alpha.

- ✓ Interpret the results by comparing with the conclusions of other studies.

How long the expected life time was extended?

Hypothesis 1

Ho: d = 0

Ha: d >0

t statistics = 8,1

p value = 0,00

Ho: There is no difference between life span in 1980 and 2005.

Ha: In 2005 expected life span is longer than what it was in 1980.

…othesis 2

… = 1

… > 1

…stics = 2,95

…e = 0,00

Ha: In 2005, expected life span is 1 year longer than what it was in 1980.

Hypothesis 3

Ho: d = 2

Ha: d > 2

t statistcs =-2,31

p value = 0,988

Ha: In 2005, expected life span is 2 years longer than what it was in 1980.

How long the expected life time was extended FOR WOMEN?

Hypothesis 1

Ho: d = 1

Ha: d > 1

t statistcs = 2,48

p value = 0,008

Ha: In 2005, expected life span is 1 year longer than what it was in 1980.

othesis 2

2

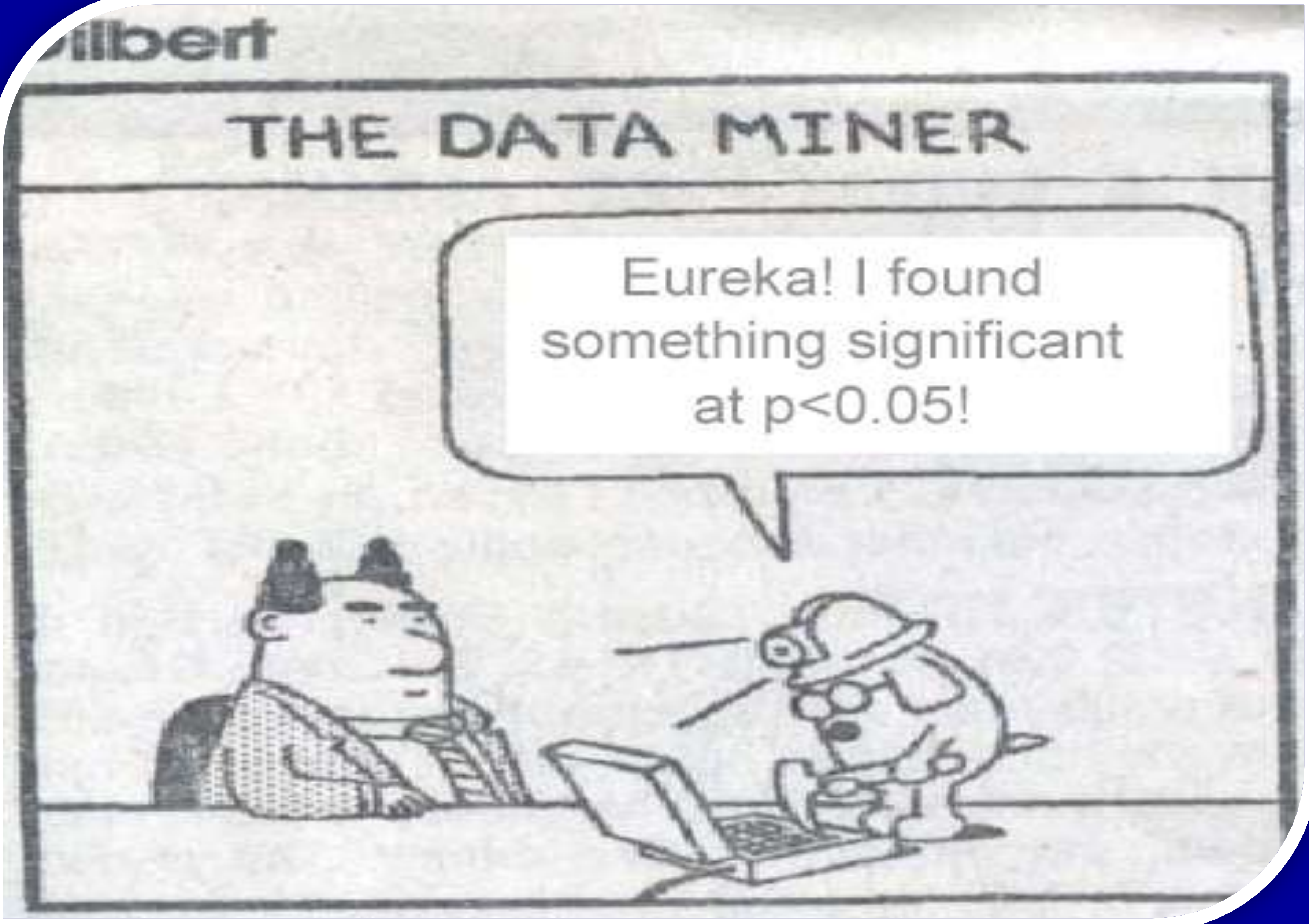
2

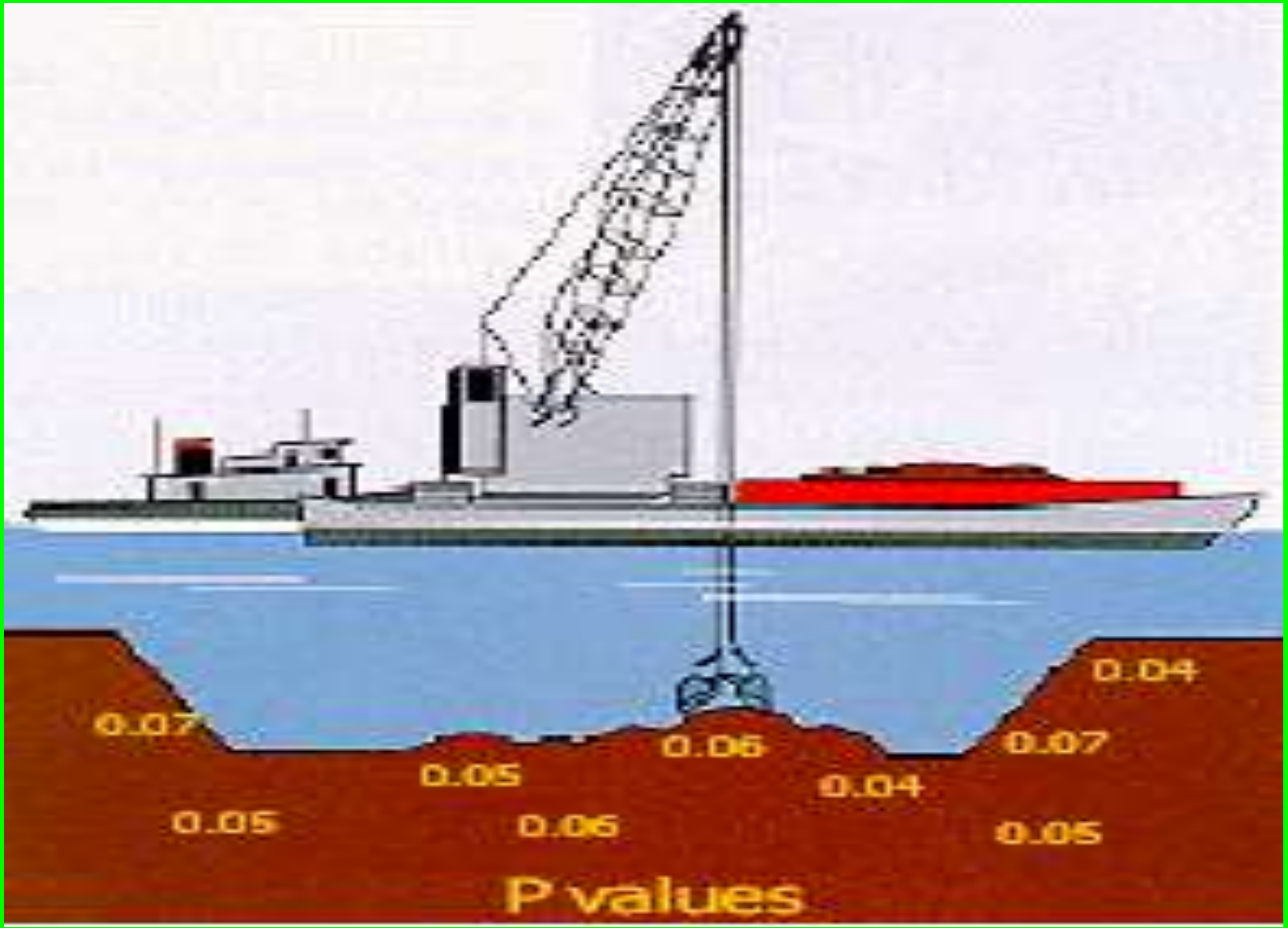
t statistics =-2,31

p value = 0,994

Ha: In 2005, expected life span is 2 years longer than what it was in 1980 for women.

CANNOT BE REJECTED


ilbert
THE DATA MINER
Eureka! I found something significant at p<0.05!


0.04
0.07
0.06
0.07
0.05
0.04
0.05
0.06
0.05
P values

Confounding factor -1

In a well designed psychology experiment, an investigator will randomly assign subjects to two or more groups and except for differences in the experimental procedure applied to each group, the groups will be treated exactly alike.

Confounding factor -2

Under these circumstances any differences between the groups that are statistically significant are attributed to differences in the treatment conditions.

This, of course assumes that except for the various treatment conditions the groups were, in fact, treated exactly alike.

Confounding factor -3

Unfortunately, however, it is always possible that despite an experimenter's best intentions there was some unsuspected systematic differences in the way the groups were treated in addition to the intended treatment conditions.

Statisticians describe systematic differences of this sort as confounding factors or ***confounding variables***.

Confounding factor -4

If, for example, subjects in one group are simultaneously tested in a room with the heat set at 70 degrees whereas subjects in another group are simultaneously tested in a nearby identically appointed room with the heat set at 60 degrees, the obtained differences in performance could be attributed to any of three factors.

Confounding factor -5

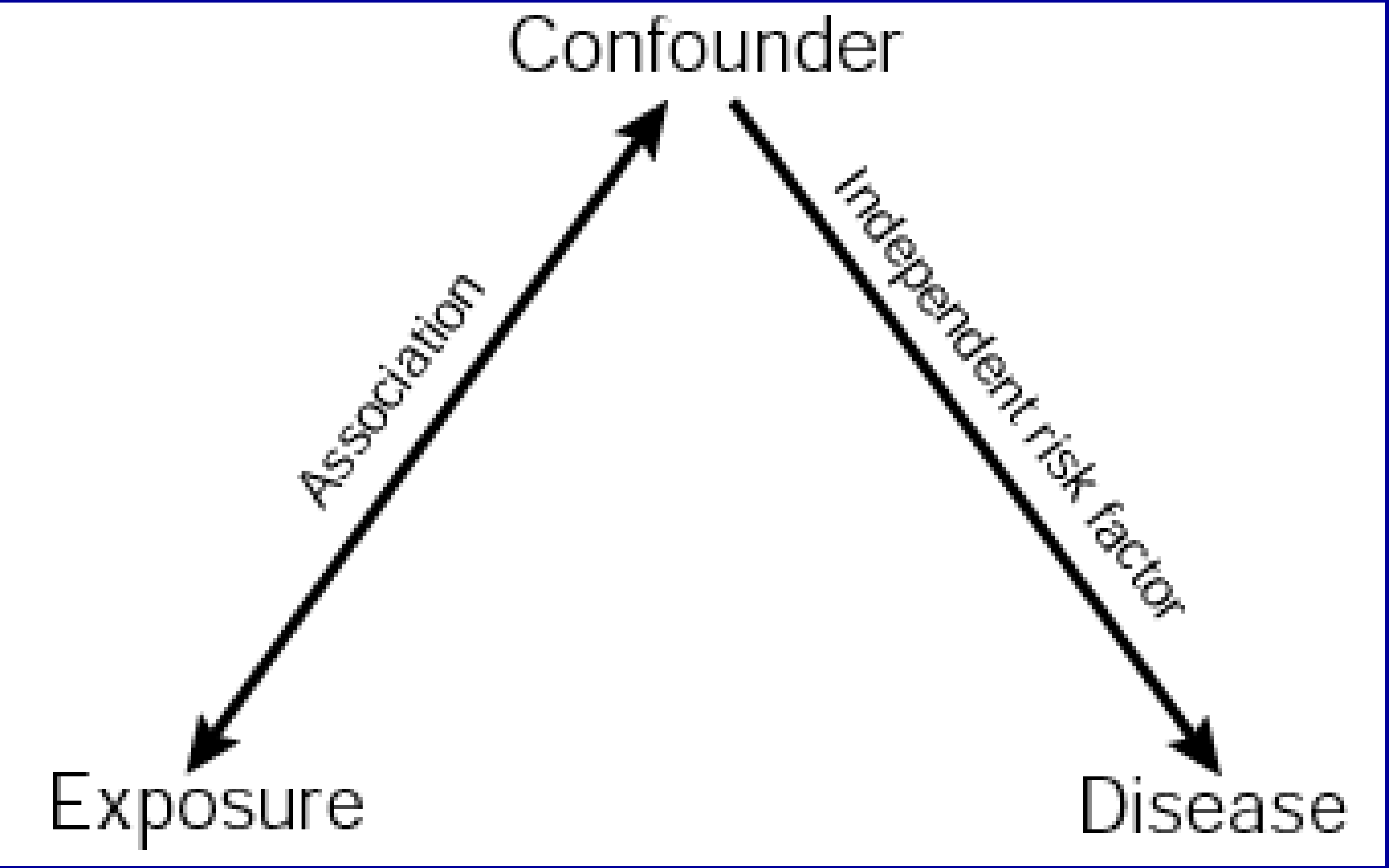
It could be due to the random assignment of subjects (i.e. to chance). It could be due to the different temperatures in the two rooms.

It could, however, be due to some confounding factor such as differences in ambient illumination that result from unnoticed differences in the orientation of each room with respect to the sun.

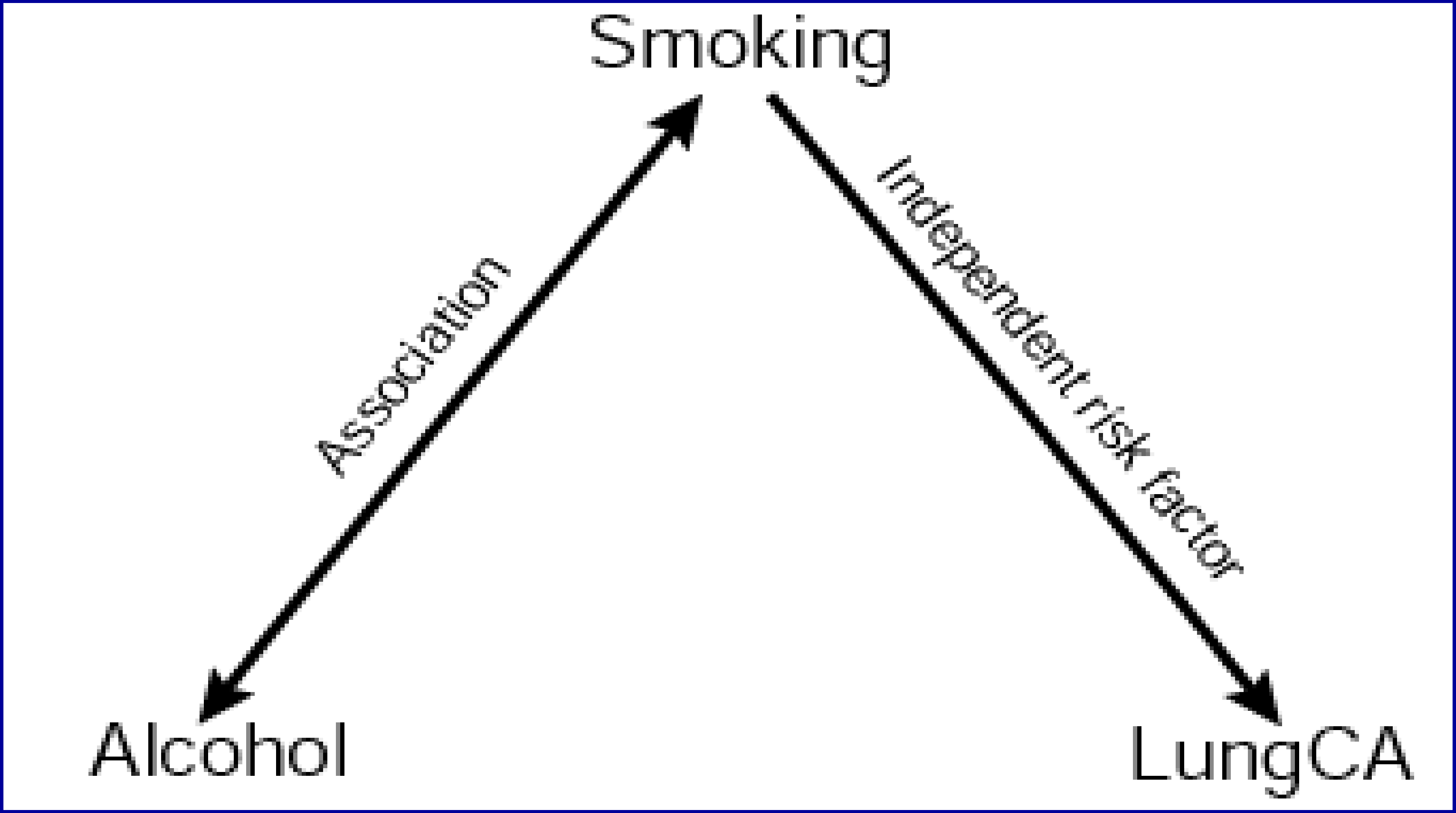
Confounding factor -6

In any experiment an appropriate statistical test can help in the decision as to whether or not to attribute the results to chance, but only the most careful analysis of the actual conditions of the experiment can suggest whether or not the results might be due to a confounding factor.

Confounding



Confounding



t statistics-1

This statistic is a measure on a random sample *(or pair of samples)* in which a mean *(or pair of means)* appears in the numerator and an estimate of the numerator's standard deviation appears in the denominator.

The later estimate is based on the calculated s square or s squares of the samples.

t statistics-2

- If these calculations yield a value of (t) that is sufficiently different from zero, the test is considered to be statistically significant.
- The **<u>t test</u>** employs the statistic (**t**) to test a given statistical hypothesis about the mean of a population *(or about the means of two populations*).

Type 1 error

You have committed a Type One error if you have rejected the hypothesis tested when it was true. In a given statistical test, the probability of a type 1 error is equal to the value you have set for **alpha**.

Type 2 error

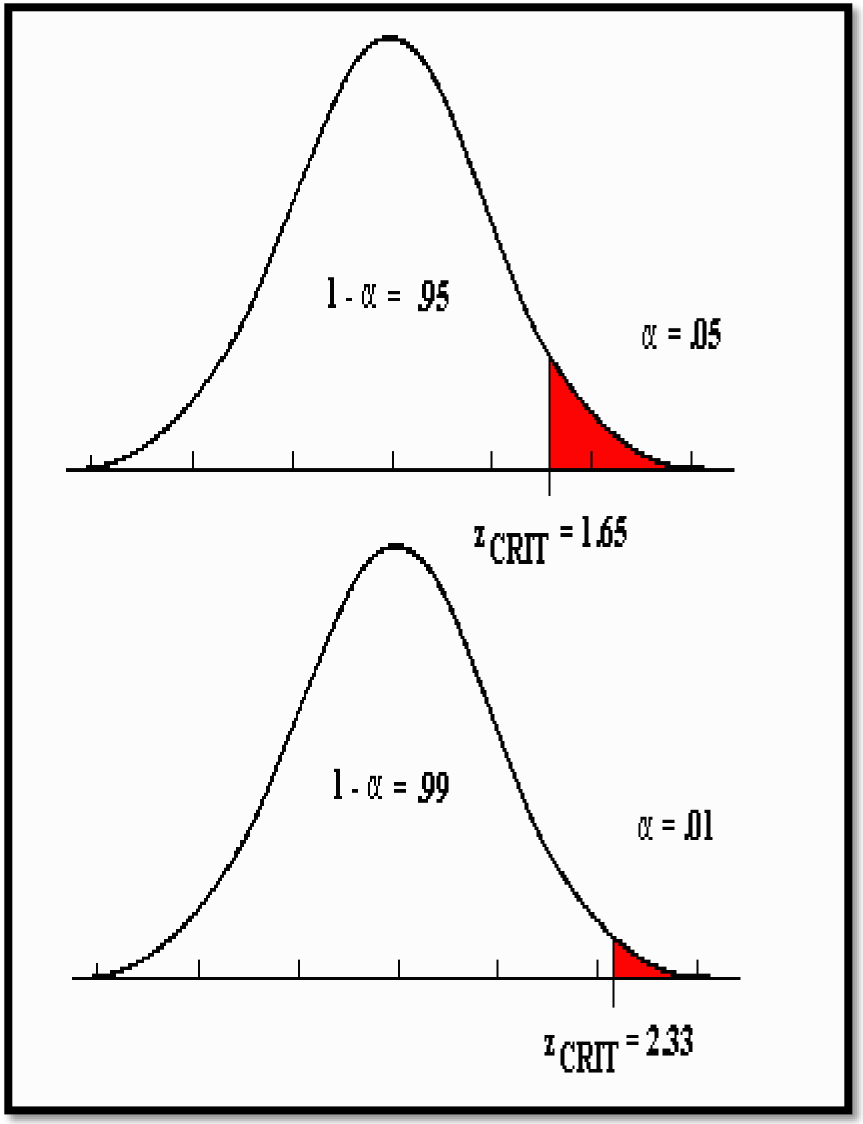
You have committed a **Type II error** if you failed to reject the hypothesis tested when a given alternative hypothesis was true. In a given statistical test, the probability of a type II error is equal to the value calculated for **Beta**.

Type 2 error

- Type II- alpha- there is a difference but it is not detected
- Type I- beta- measure a difference (effect) that is not so
- Reliability -accuracy, remember, random error
- Validity -systemmatic error- measuring right thing

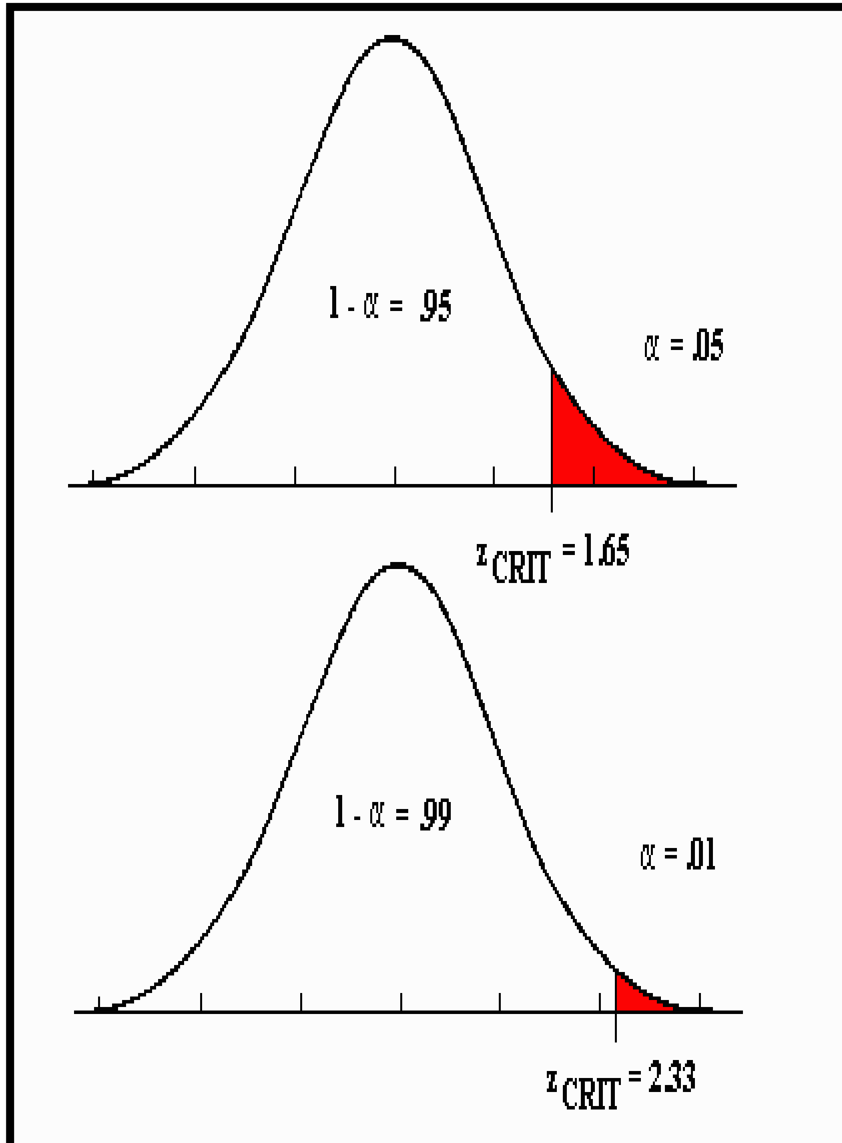
ß value

Statisticians use the Greek letter **beta** to indicate the probability of failing to reject the hypothesis tested when that hypothesis is false and a specific alternative hypothesis is true.



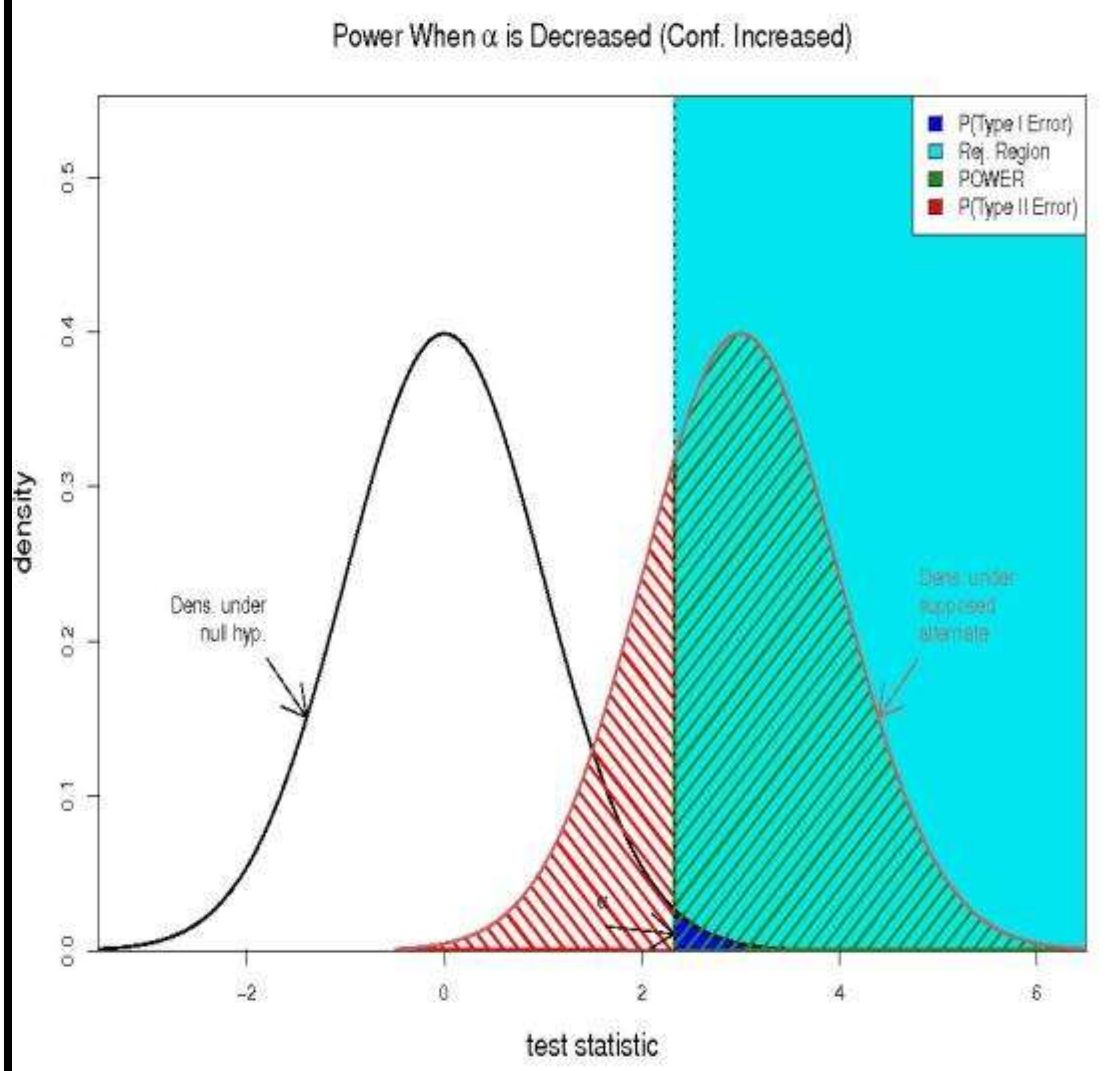
ß value

For a given test, the value of beta is determined by the previously elected value of ***alpha***, certain features of the statistic that is being calculated (particularly with the ***sample size***) and the specific alternative hypothesis that is being entertained.



1-ß value = Power of the test

While it is possible to carry out a statistical test without entertaining a specific alternative hypothesis, neither beta nor power can be calculated if there is no specific alternative hypothesis.



1-ß value = Power of the test

It is relevant to note here that power (the probability that the test will reject the hypothesis tested when a specific alternative hypothesis is true) is always equal to one minus beta. (i.e. **Power = 1 - beta**)

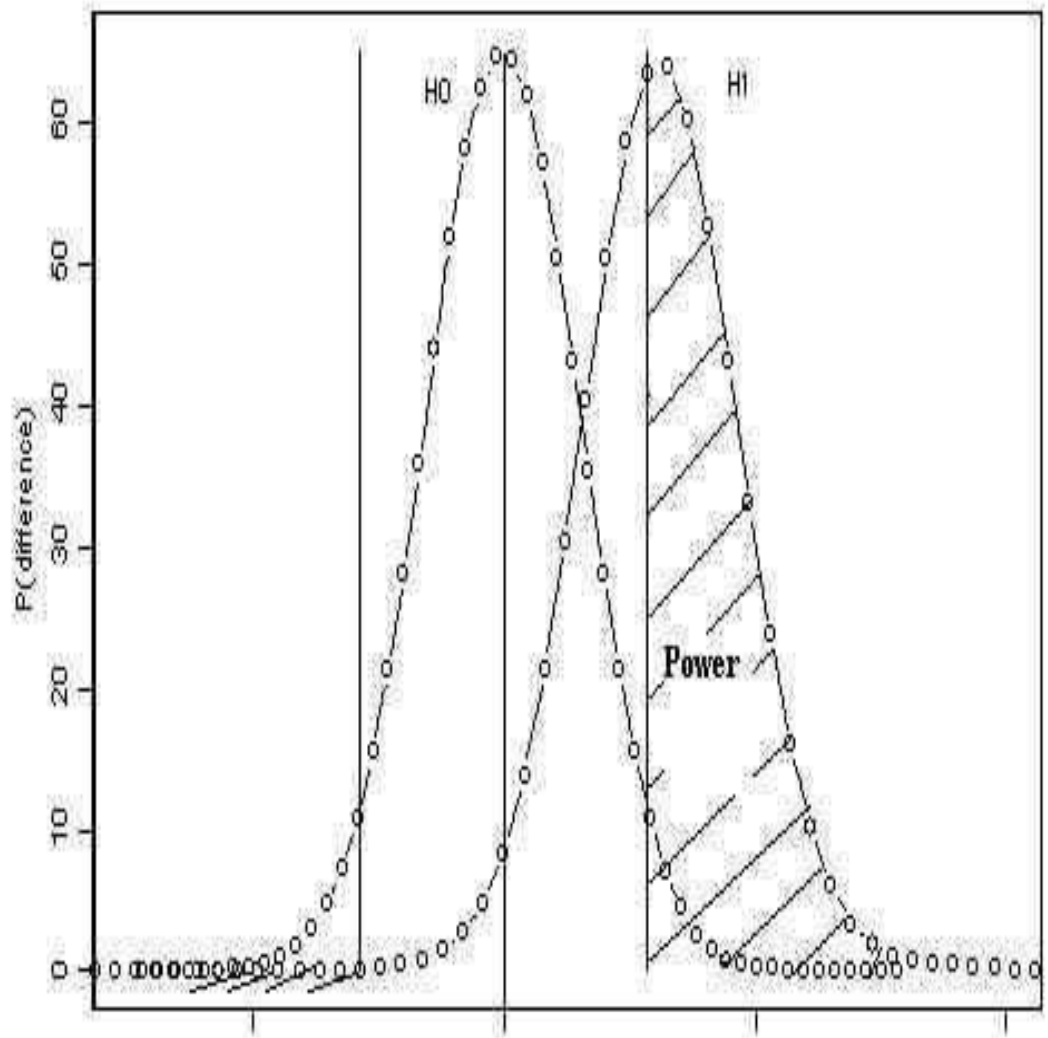


Figure 1.

Power of the test

For a statistician, **the power of a test** is the probability that the test will reject the hypothesis tested when a specific alternative hypothesis is true. To calculate the **power** of a given test it is necessary to specify alpha (*the probability that the test will lead to the rejection of the hypothesis tested when that hypothesis is true*) and to specify a specific **alternative hypothesis**.

Power of the test

A statistical test is a procedure for deciding whether an assertion (e.g. a Hypothesis) about a quantitative feature of a population is true or false. We test an hypothesis of this sort by drawing a random sample from the population in question and calculating an appropriate statistic on its items. If, in doing so, we obtain a value of the statistic that would occur rarely when the hypothesis is true we would have reason to reject the hypothesis.

Power of the test

With this procedure it is customary to reject the hypothesis tested when our statistic has a value that is among those that, theoretically, would be expected to occur no more than 5 out of every 100 times that a random sample (of the same size) is drawn from the population in question when the hypothesis is, in fact, true. Much of the text of this tutorial is devoted to explanations of exactly how this kind of theoretical expectation is developed.

Thoughtful decisions

Finally, it is noteworthy that
the appropriate conduct of
any statistical test invariably requires
many thoughtful decisions.
It is, for example, always necessary
to decide what statistic to use,
what sample size to employ and
what criteria to establish for rejection
of the hypothesis tested.

SE & The Confidence Interval

Sometimes, a confidence interval may be computed from theory alone. For instance, means of large, random samples tend to be unbiased and normally distributed. Therefore, the 95-percent confidence interval (95 % CI) for any such mean is just m ± 1.96 SE, where m is the observed mean and SE is the *standard error* of the mean, as given by the equation below

$$SE = \frac{\sigma}{\sqrt{N}}$$

where σ is the population standard deviation for the data and n is the sample size.

- **THE SECREET of the NUMBERS…**
- **1 x 8 + 1 = 9**
  12 x 8 + 2 = 98
  123 x 8 + 3 = 987
  1234 x 8 + 4 = 9876
  12345 x 8 + 5 = 98765
  123456 x 8 + 6 = 987654
  1234567 x 8 + 7 = 9876543
  12345678 x 8 + 8 = 98765432
  123456789 x 8 + 9 = 987654321

What is being a scientist?

" Being a scientist is not a professional job; but rather, is a life style. "

Prof. Dr. Cahit ARF
Founder of "Arf's Theorem"

One anectode from Hodja..

One day, Nasrettin Hodja takes his «saz» *(a wired popular musical instrument)* to play. Keeping hand on «one note's interval» insistingly, plays the wires monotonly.

His wife gets tired and angry due to monotony and warns Hodja :

- Hey Hodja, according to my knowledge, your hand pressing on the notes must move up and down through the handle of Saz»??

But your hand is almost sticky!?

Hodja response very rapidly :

- Shut up hey woman! They all are looking for the exact point which I've been pressing..

What says p value us? (1)

In «p value» abbreviation,
it abbreviates term «probability».
For instance, in a situation where a given test statistic and p value were in hand ; such as p = .002, t = 7.89.. «p value» delineates us that :

Look : You have a test statistics t in size of 7.89 and my «hight» is 0.002. If you say. «2 variables are in connection», «there must be a significant relationship between the 2 variables» so on.. (similar expressions) The probability of accuracy for that way judgement equals to .998 (1-p). You my rely on me, in that way causality construction, really, the probability of mistake is equals to me, i.e to %0.002!

What says p value us? (2)

On the contrary, in case of interpreting «no significant tie» between those 2 variables depicts an inaccurate decision with a probability of 99.8 % *(remember the basic rule of probability; p+q=1)*.

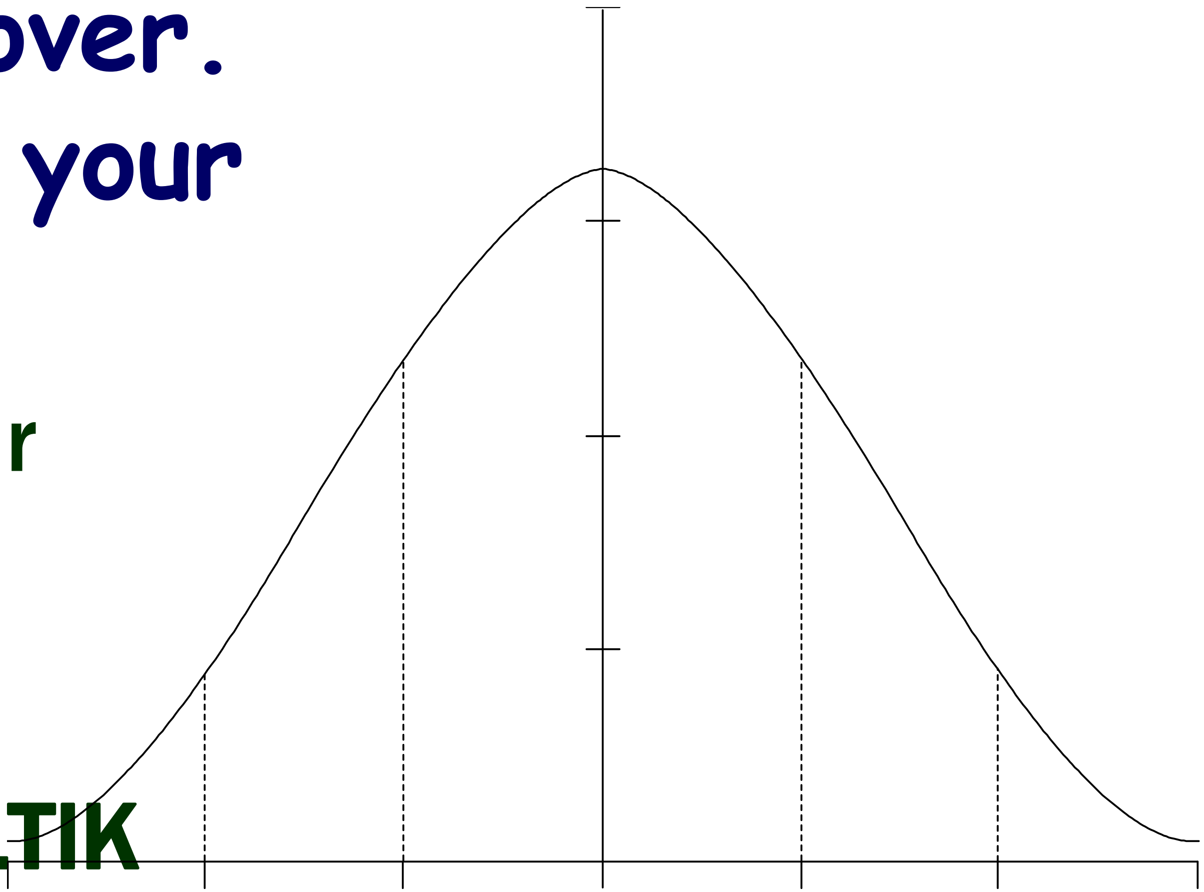
As a conclusion, 1-p = q is the probability of type 1 (alpha) and 2 error (beta) if one falls; but just opposite the reliability level for statistical inference without type 1 and 2 error.

In correct causality reasoning, p says the possibility of «real true» situation in the universe.

Well done, lovely and so wise, skillfull «p value»!

Lecture is over. Thanks for your interest.

We expect your questions and comments..

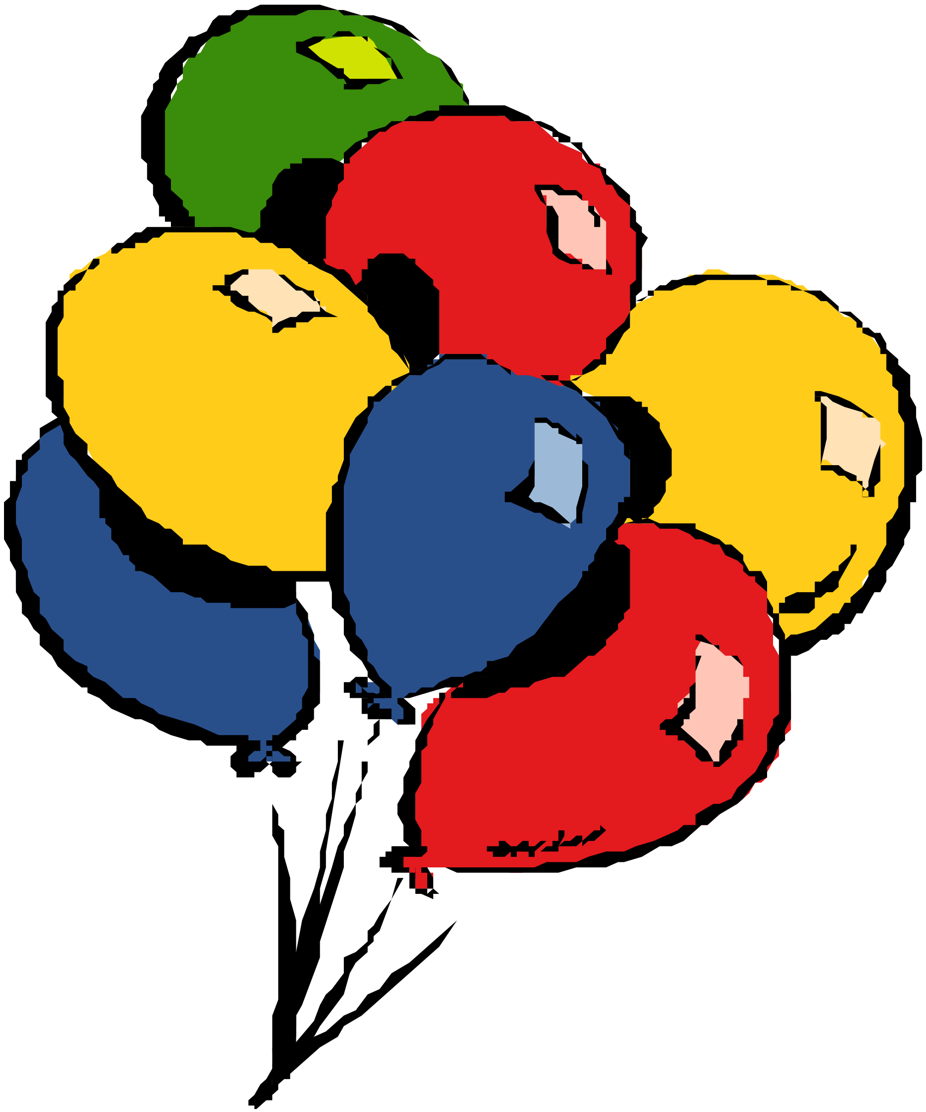


Dr. Ahmet SALTIK

www.ahmetsaltik.net

Your moderator;

- Thanks you sincerely for your contribution and patience.
- Greets you with love and respect..



Dr. Ahmet SALTIK
www.ahmetsaltik.net

PEACE at HOME PEACE in the WORLD!

Gazi Mustafa Kemal ATATÜRK

Key words : Nasrettin Hodja, one anectode, interval estimation, point estimation, Being a scientist, Cahit Arf, Arf's Theorem, SECREET of the NUMBERS, Confidence Interval, Thoughtful decisions, Power of the test, alternative hypothesis, 1-ß value, ß beta value, Type 2 error, beta error, Type 1 error, alpha error, pair of means, pair of samples, Confounding, Confounder, Confounding factor, Confounding variable, well designed experiment, p value, well known p value, data mining, data miner, Eureka, **statistically significant**, Interpret the results by comparing, Decisions Based on Hypothesis Tests, Galton's experiment, Galton's peas formed a bell shape, Standard normal distribution, Epidemiology and Biostatistics, research methodology, science is general and objective, control group, Randomization, blind study, alternative hypothesis, null hypothesis, statistical assessment, Rejection & Accepting Ho, independent and dependent variables, statistical test, sample size, rejection of a hypothesis, From data to model, Albert Einstein, what p value says, causality, statistical inference, reality of the universe, p+q=1, basic rule of probaility, **www.ahmetsaltik.net, Ankara Univ. Medical School,**