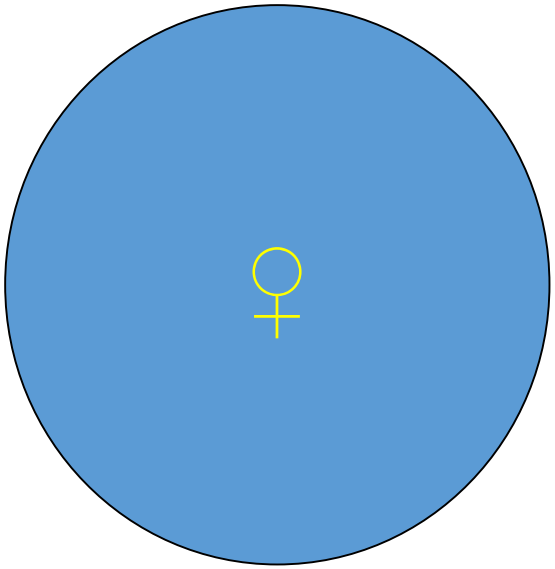
بسم الله الرحمن الرحيم

Inferential analysis

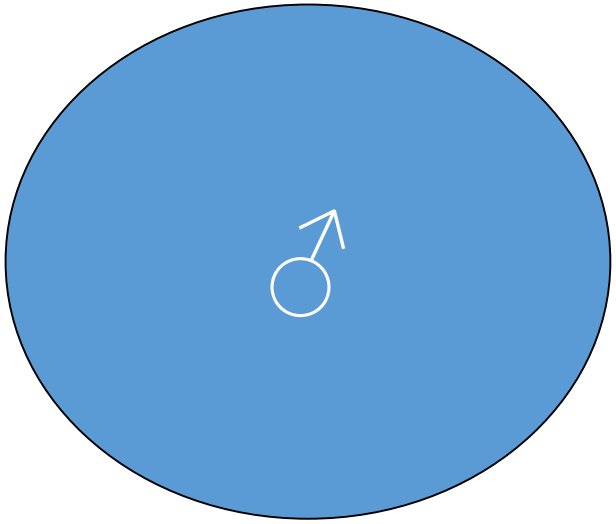
LX

25-7-2023

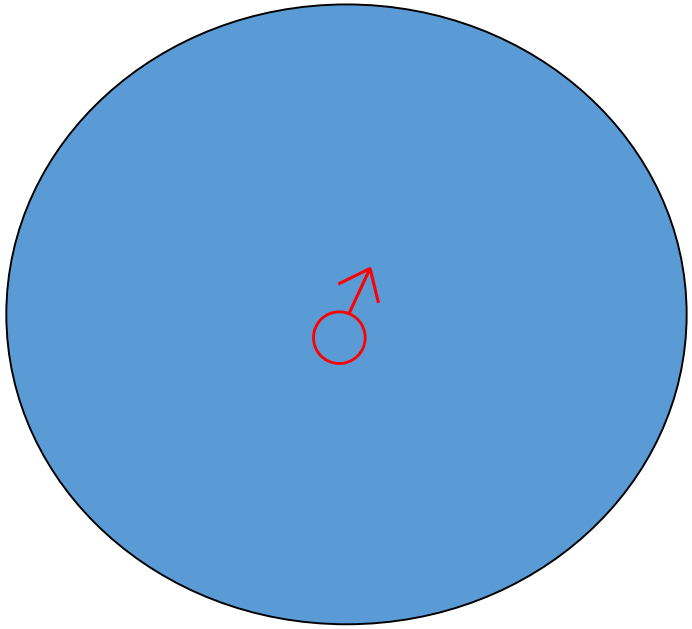
Prof DR. Waqar Al – Kubaisy



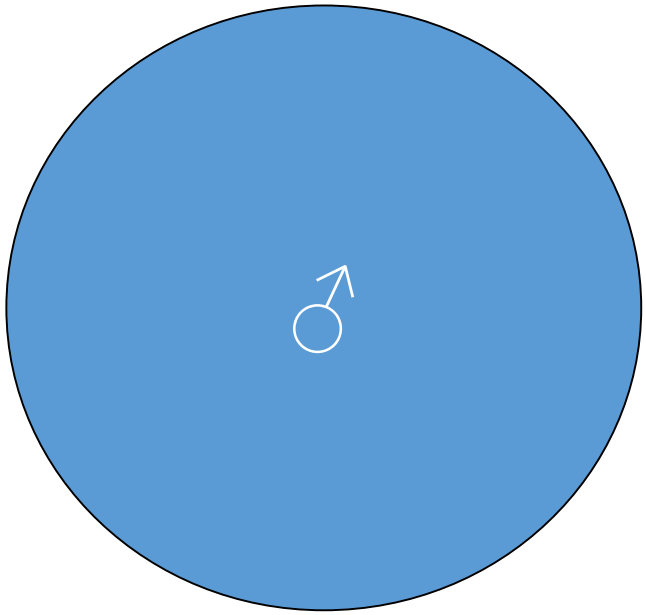
$\bar{X}$ 55kg



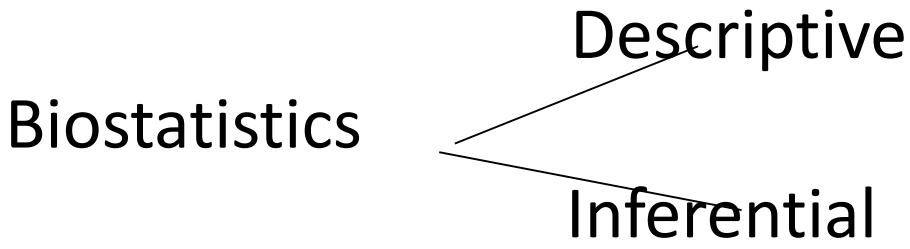
$\bar{X}$ 70 kg

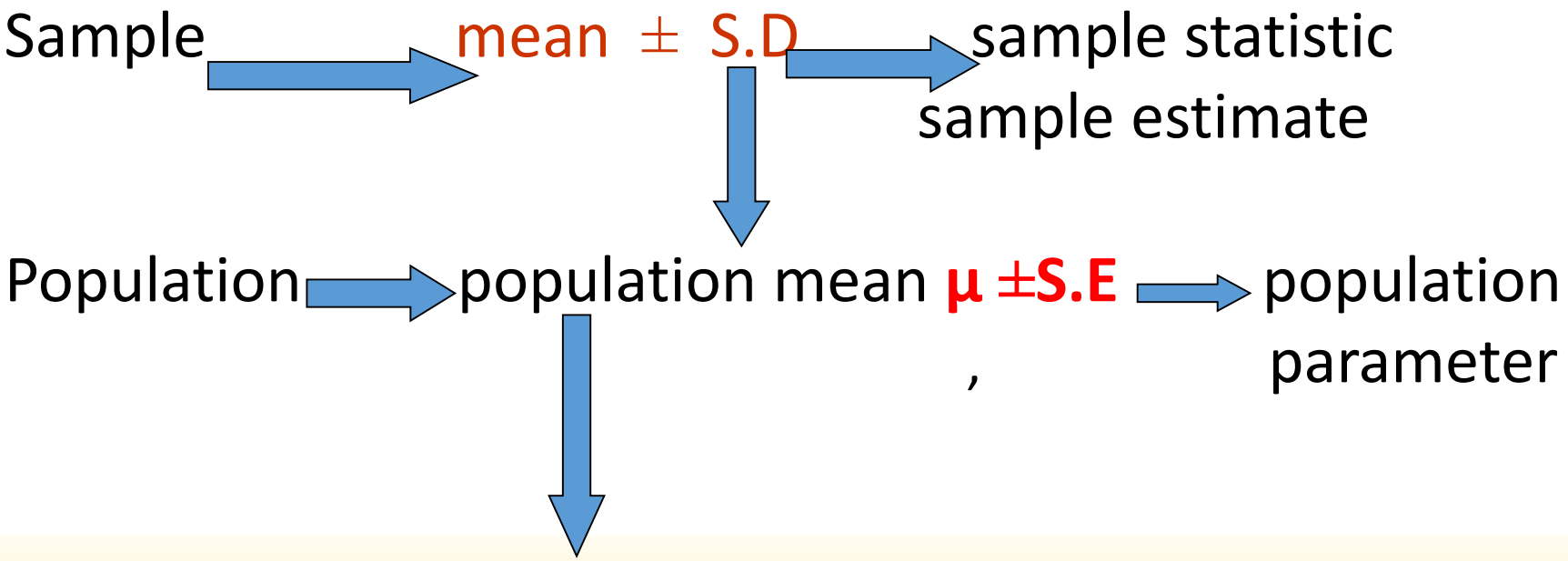


$\bar{X}$ 70 kg



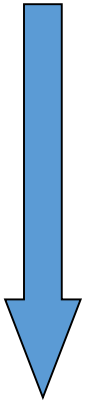
$\bar{X}$ 66kg





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

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



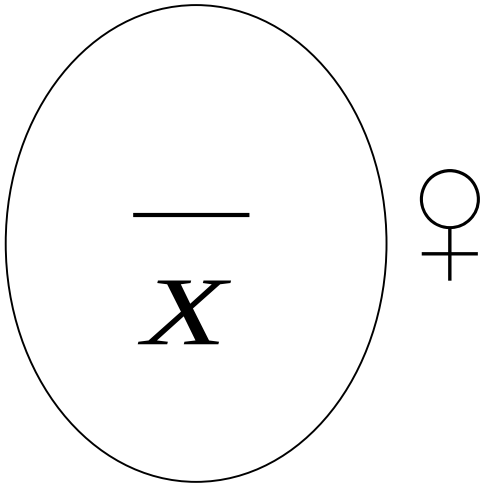
Inferential Biostatistics (Analysis).

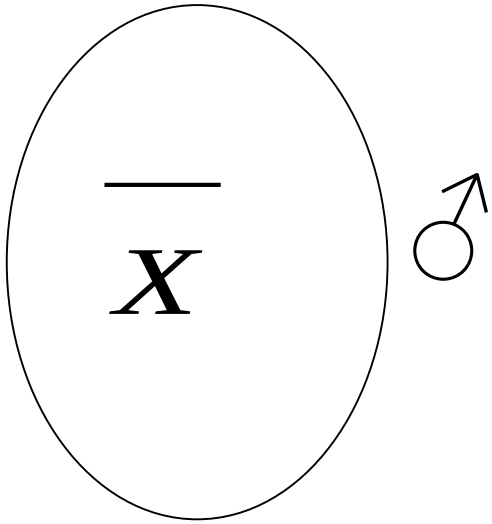
It is used to test specific hypothesis about population by using certain **test significance** .

We expect always that there is a difference between groups .

Mean body weight of ♂ = 70 kg .

Mean body weight of ♀ = 55 kg



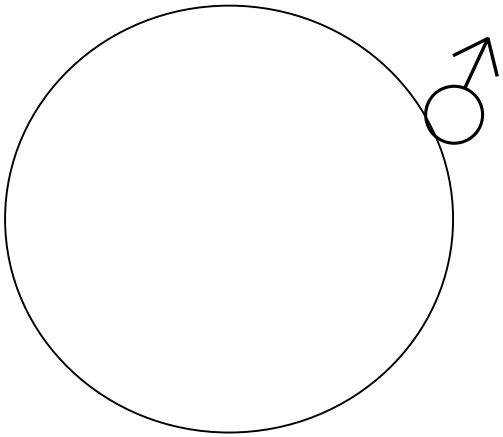


Mean body weight of ♀=55 kg

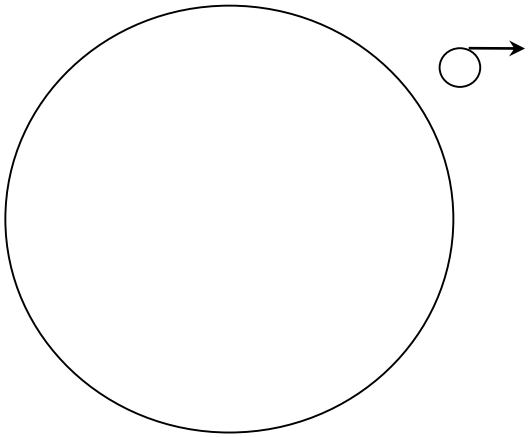
Mean body weight of ♂70 kg

Difference could be ????

Influencing factor



Group I



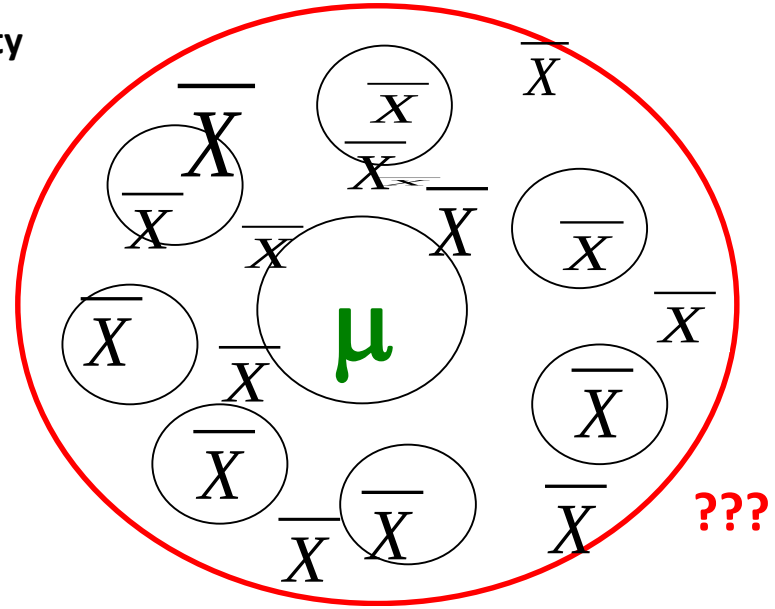
Group II

Mean body weight of grope I
= 65 kg .

Mean body weight of **grope II**
=60 kg .

Difference could be ?? Chance factor
Sampling variability
Sampling error

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.

- We expect always that there is a difference between groups

Difference could be → Chance factor

Difference could be → influencing factor

So we expect always that, there is difference .

And by using these test of significance,

we assess whether that difference between groups is cause 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

- 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 (caring)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

❖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

Steps Of Testing Hypothesis

I- Data Nature of data (variable)

2-Assumption

3-Hypothesis formulation

4-Test statistics

5-Define Level of Significance

6-Apply The Proper Test of Significance

7-Statistical decision

8-P value

I- Data Nature of data (variable)
2-Assumption
3-Hypothesis formulation
4-Test statistics
5-Define Level of Significance
6-Apply The Proper Test of Significance
7-Statistical decision
8-P value

I-Data Nature of Data

Continuous

Discrete

2- Assumption

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

3-Hypothesis formulation

Formulate two statistical hypothesis **simultaneously**

A-Null hypothesis (H0)

B-Alternative hypothesis (HA)

Null hypothesis (HO)

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 H0

- ✓ **it states always that, there is no significance difference**
- ✓ -or **there is** no **influence** or **effect of influencing factor** .
- ❖ In testing hypothesis process , **the HO is either:**

-Reject or

-**Not reject** (accept) .

- **If Ho 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** = or accept.

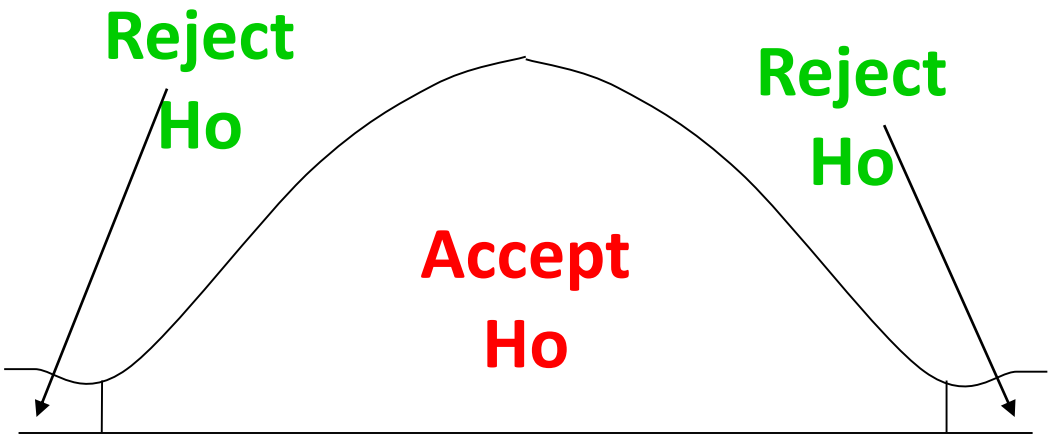
- If testing procedure leads to **rejection**,
  - we will conclude that, the data in our hand are
  - not compatible with Ho . but
  - **supporting** of **some other hypothesis** .
  - this hypothesis is known as
  - **Alternative hypothesis** (HA).

- The decision, to **reject or accept** the **Ho depends on the magnitude** (value) **of the test statistics** .

I- Data Nature of data (variable)
2-Assumption
3-Hypothesis formulation
4-Test statistics
5-Define Level of Significance
6-Apply The Proper Test of Significance
7-Statistical decision
8-P value

Test statistics

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

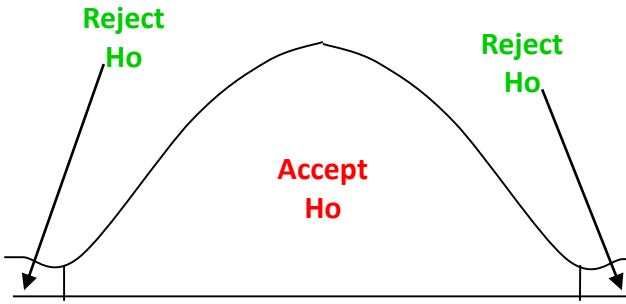


The distribution of test statistics, which is the key to the statistical inference

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 Ho is true.

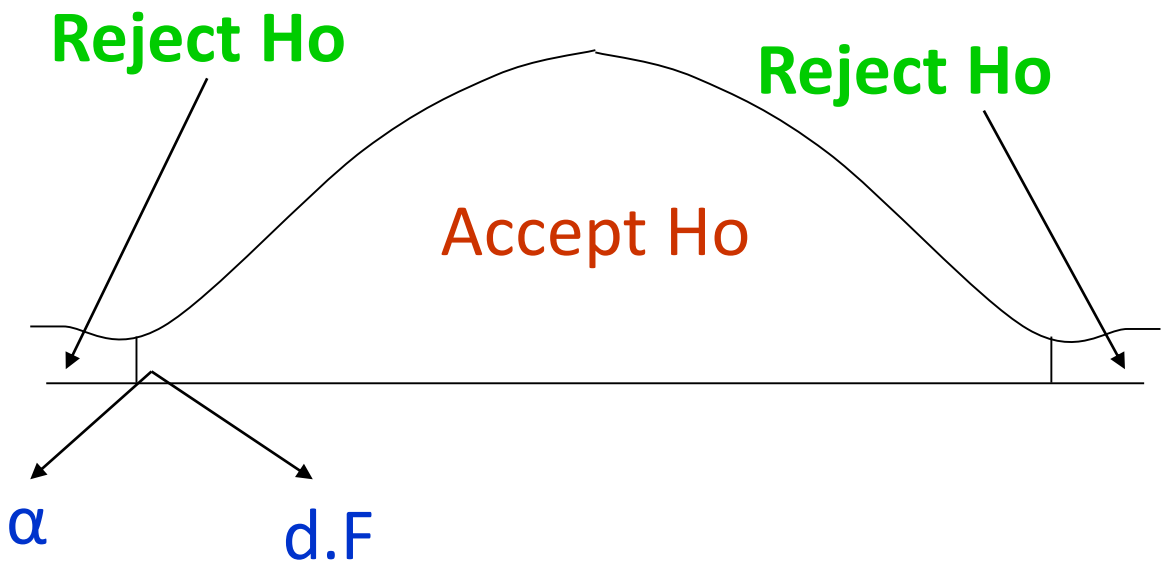
- While the values making up the accept region are more likely to occur if Ho is true.

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 .



1- Data Nature of data (variable)
2-Assumption
3-Hypothesis formulation
4-Test statistics
5-Define Level of Significance
6-Apply The Proper Test of Significan
7-Statistical decision
8-P value

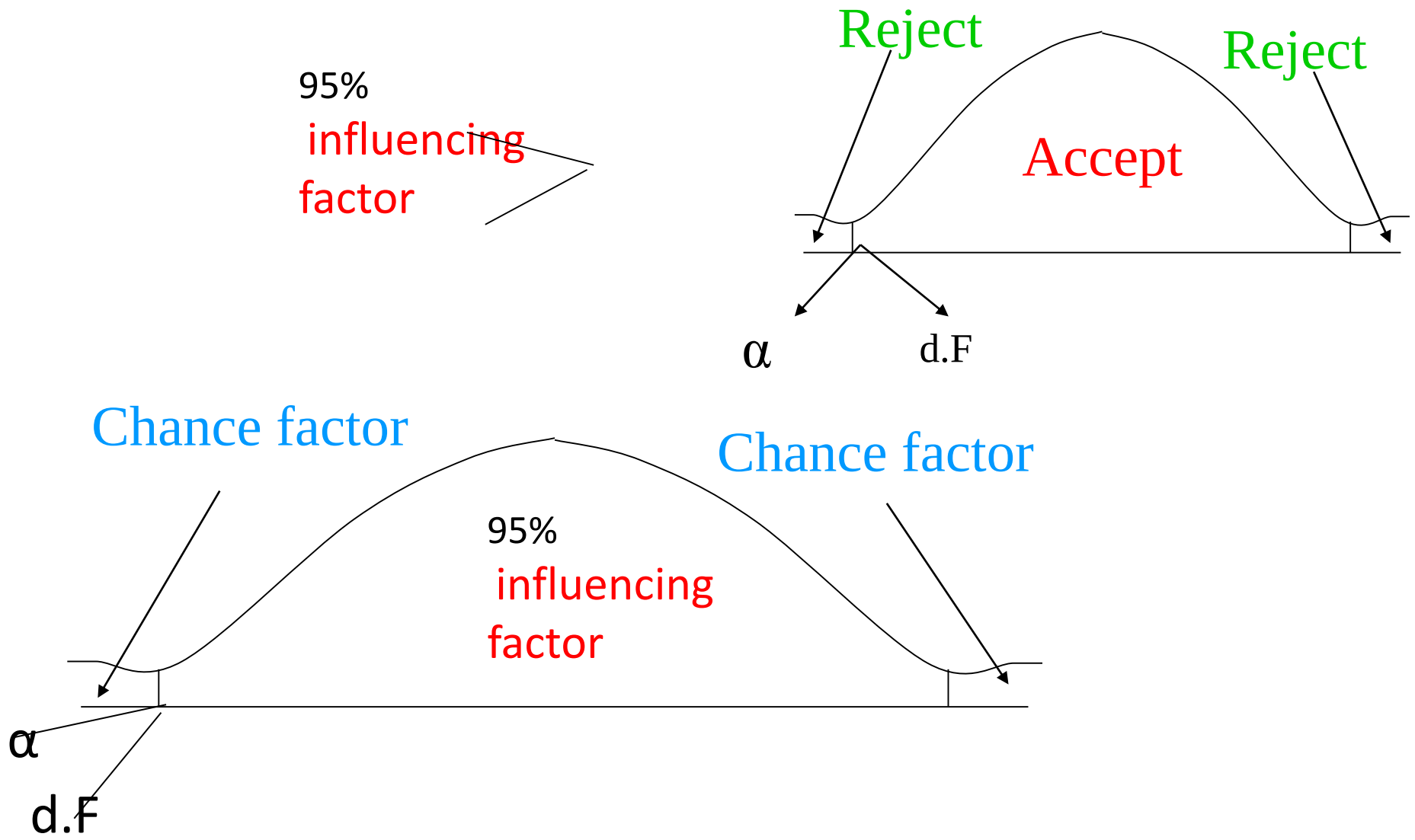
Define Level of Significance

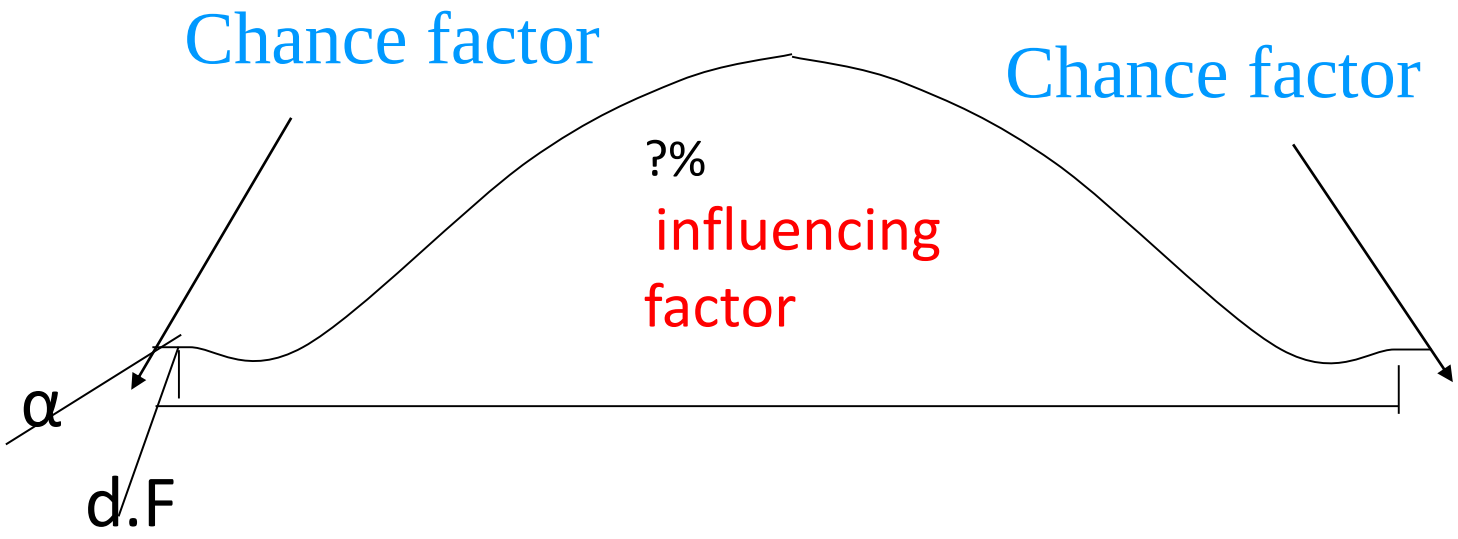
Level of significance it is the probability level, **According to NDC at which we either accept or reject Ho**

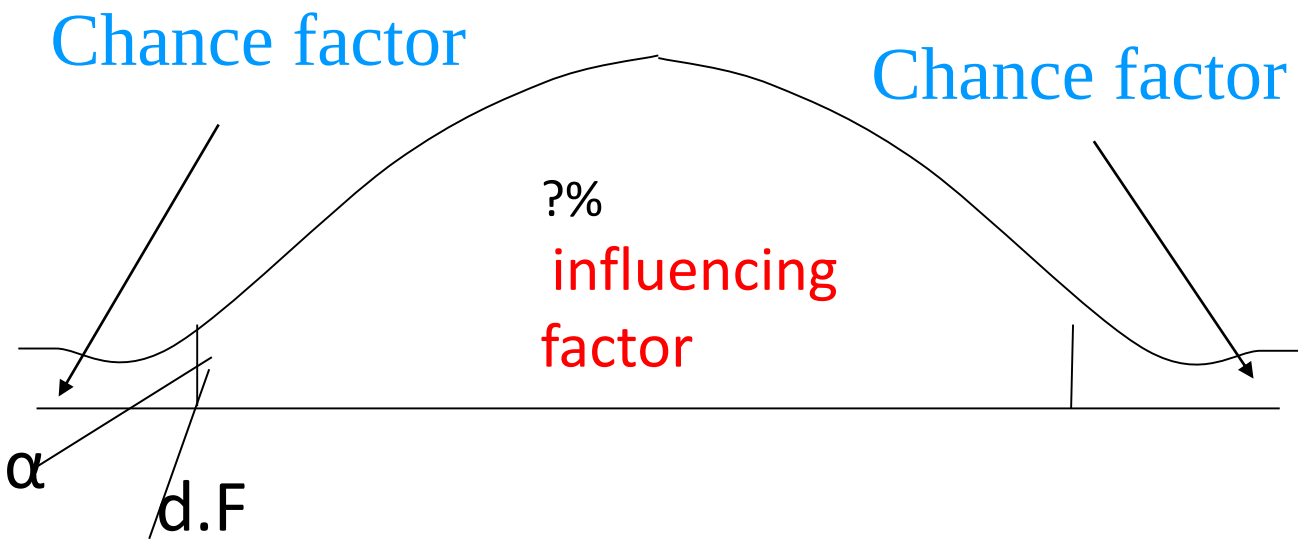
- According to N.D.C we can assume that, **95%** of the difference between groups are caused by the
- influencing factor .
- the remaining **5%** (**2.5% on each side**) are caused by **chance factor** **so**

❖ **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) .**


95%
influencing
factor
Reject
Reject
Accept
α
d.F
Chance factor
Chance factor
95%
influencing
factor
α
d.F


Chance factor
Chance factor
?%
influencing
factor
α
d.F


Chance factor
Chance factor
?%
influencing
factor
α
d.F

I- Data Nature of data (variable)
2-Assumption
3-Hypothesis formulation
4-Test statistics
5-Define Level of Significance
6-Apply The Proper Test of Significance
7-Statistical decision
8-P value

Apply The Proper Test of Significance

Compute test statistics for each set of observation (data) or (study),

we might use different test of significance.

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

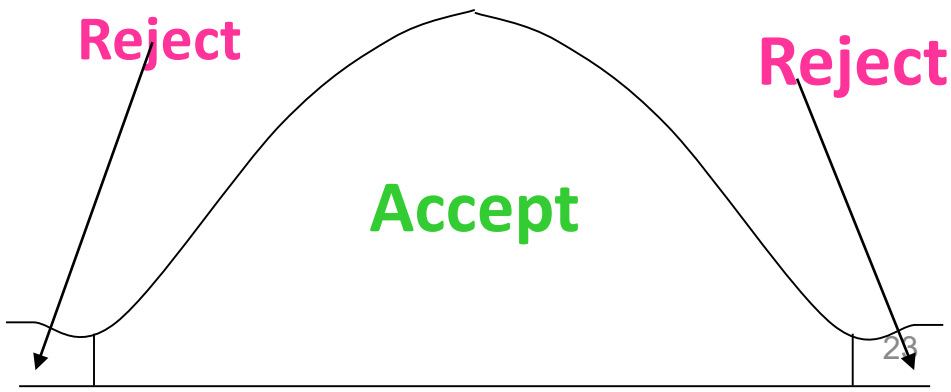
Continuous Discrete .

2- we will compute the value of test statistics.

3- Compare with accept or reject region.

4- Then by using test of significance.

5-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 Ho , or

- ✓ **the difference between these groups is not significance.**

And

There is a **chance factor causing** the difference beside the influencing factor.

P > 0.05 → **accept Ho** → **no significance difference** .

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

I- Data Nature of data (variable)
2-Assumption
3-Hypothesis formulation
4-Test statistics
5-Define Level of Significance
6-Apply The Proper Test of Significance
7-Statistical decision
8-P value

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 < 0.05.**

This means that the

- **influencing factor** has **significant** effect

- **P < 0.05 → reject Ho → significant difference.**

I- Data Nature of data (variable)
2-Assumption
3-Hypothesis formulation
4-Test statistics
5-Define Level of Significance
6-Apply The Proper Test of Significance
7-Statistical decision
8-P value

Statistical decision

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

If computed value of test statistical fall in the reject region we reject Ho and taking HA

or **not rejected if the computed value of test statistical fall in the accept region , we accept Ho .**

If Ho is rejected clinical decision is compatible to the HA .

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

P value

$P < 0.05$

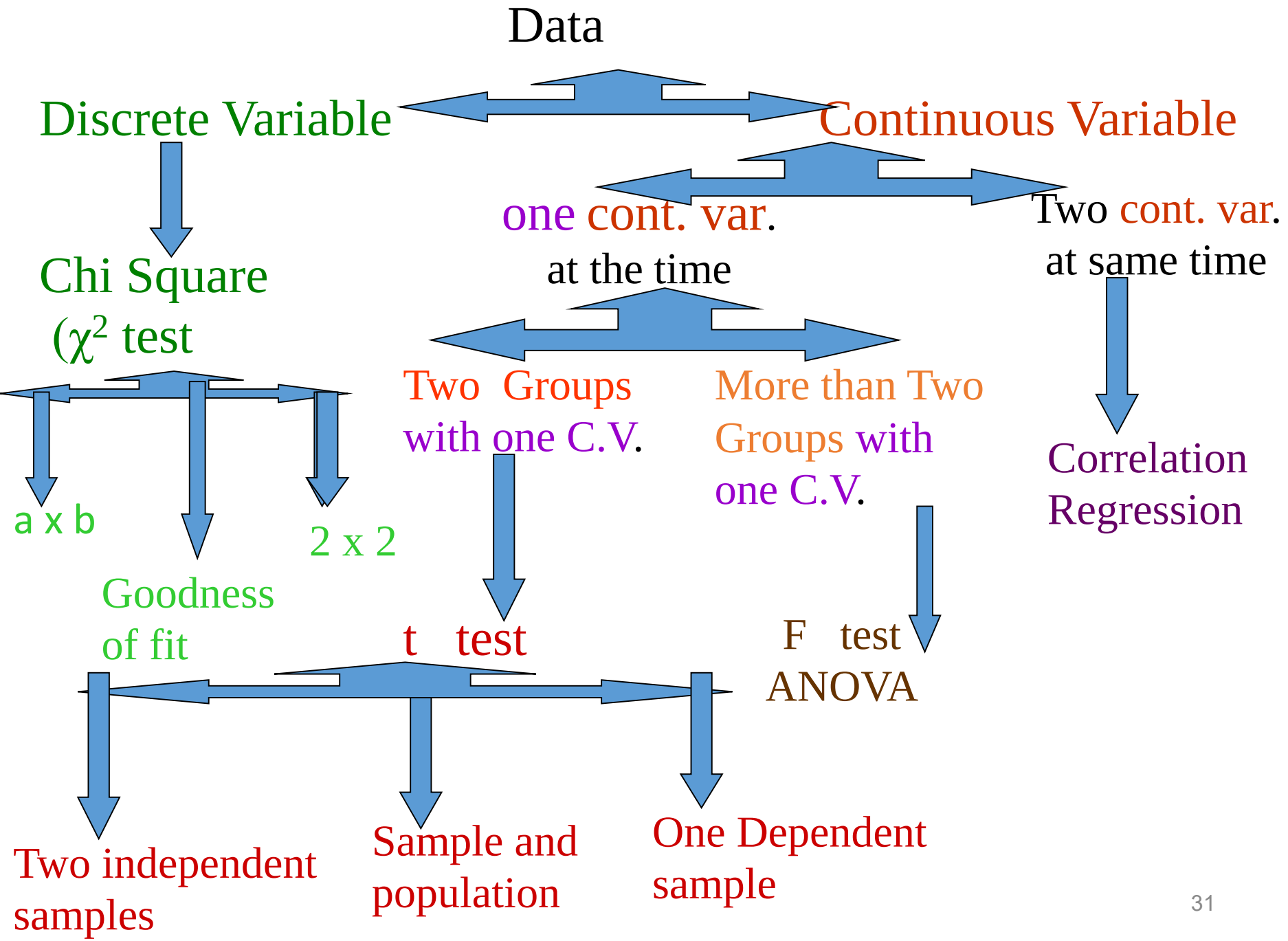
$P > 0.05$

t distribution critical values

df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965

18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.663.	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.15	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646

40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.295	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
inf.	.674	.841	1.036	1.282	1.64	1.960	2.054	2.326	2.576	2.807	3.091	3.291


Data
Discrete Variable
Continuous Variable
Chi Square
(χ^2 test
one cont. var.
at the time
Two cont. var.
at same time
Correlation
Regression
a x b
2 x 2
Goodness
of fit
Two Groups
with one C.V.
More than Two
Groups with
one C.V.
t test
F test
ANOVA
Two independent
samples
Sample and
population
One Dependent
sample


THANK YOU ALL