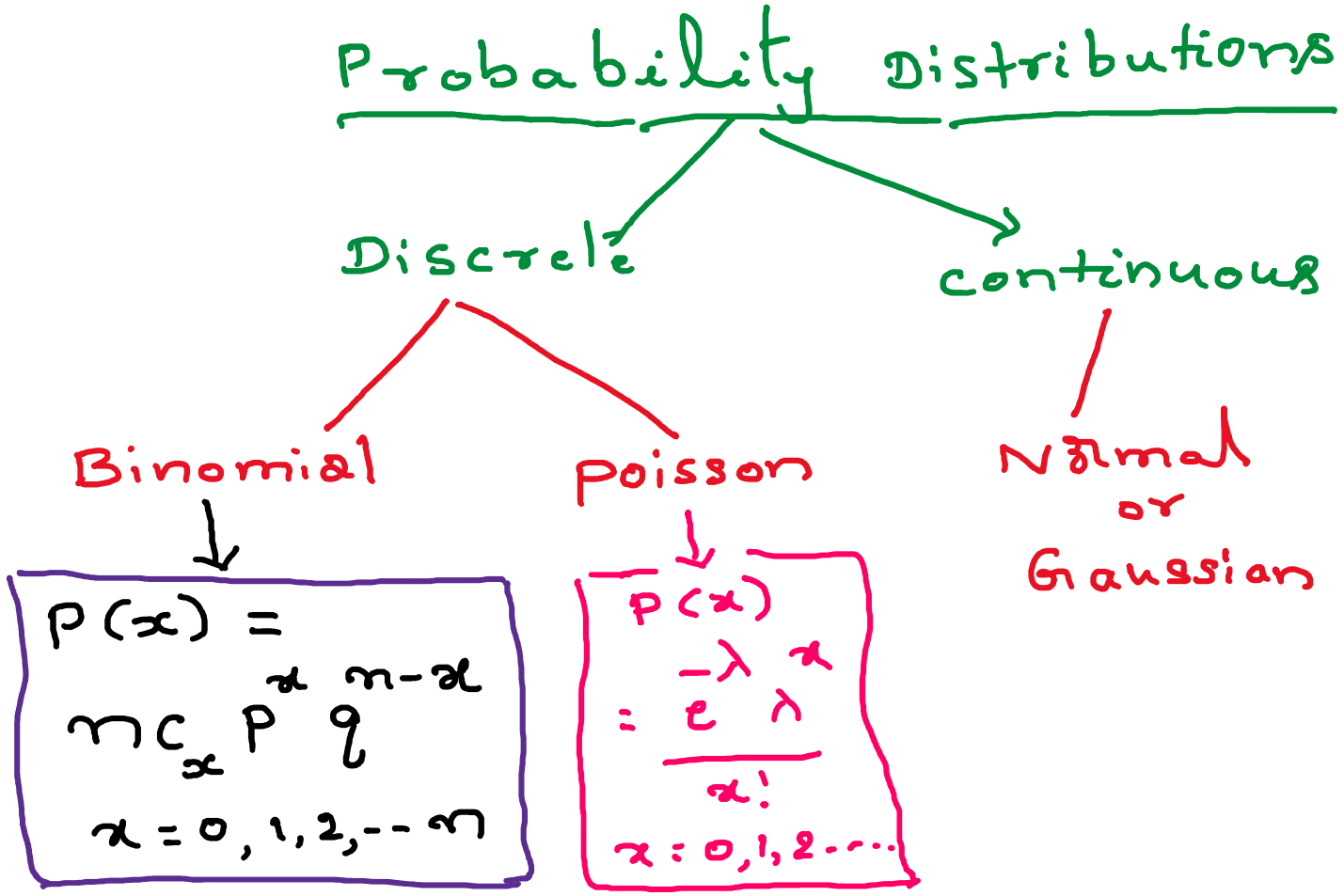
Inferential Statistics
Normal distribution
Sampling
Testing of Hypothesis


Probability Distributions
Discrete
continuous
Binomial
poisson
Normal
or
Gaussian
$P(x) = {}^{n}C_{x} p^{x} q^{n-x}$
$x = 0, 1, 2, \ldots n$
$P(x) = \frac{e^{-\lambda} \lambda^{x}}{x!}$
$x = 0, 1, 2 \ldots$

Normal Distributions

Normal distribution, also known as the **Gaussian distribution**, is a probability **distribution** that is symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean

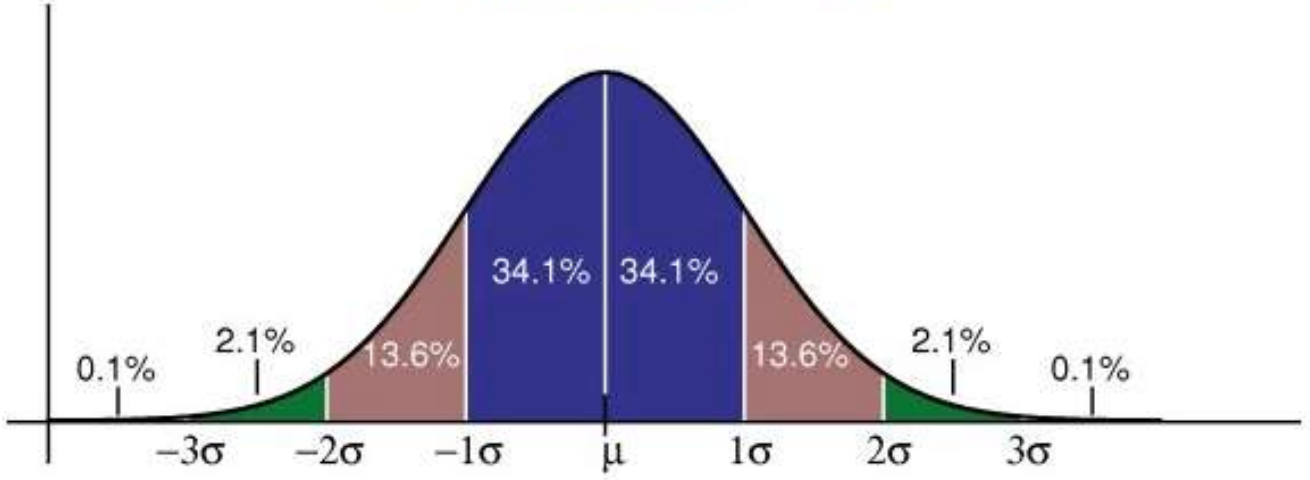
A normal distribution has some interesting properties: it has a bell shape, the mean and median are equal, and 68% of the data falls within 1 standard deviation.

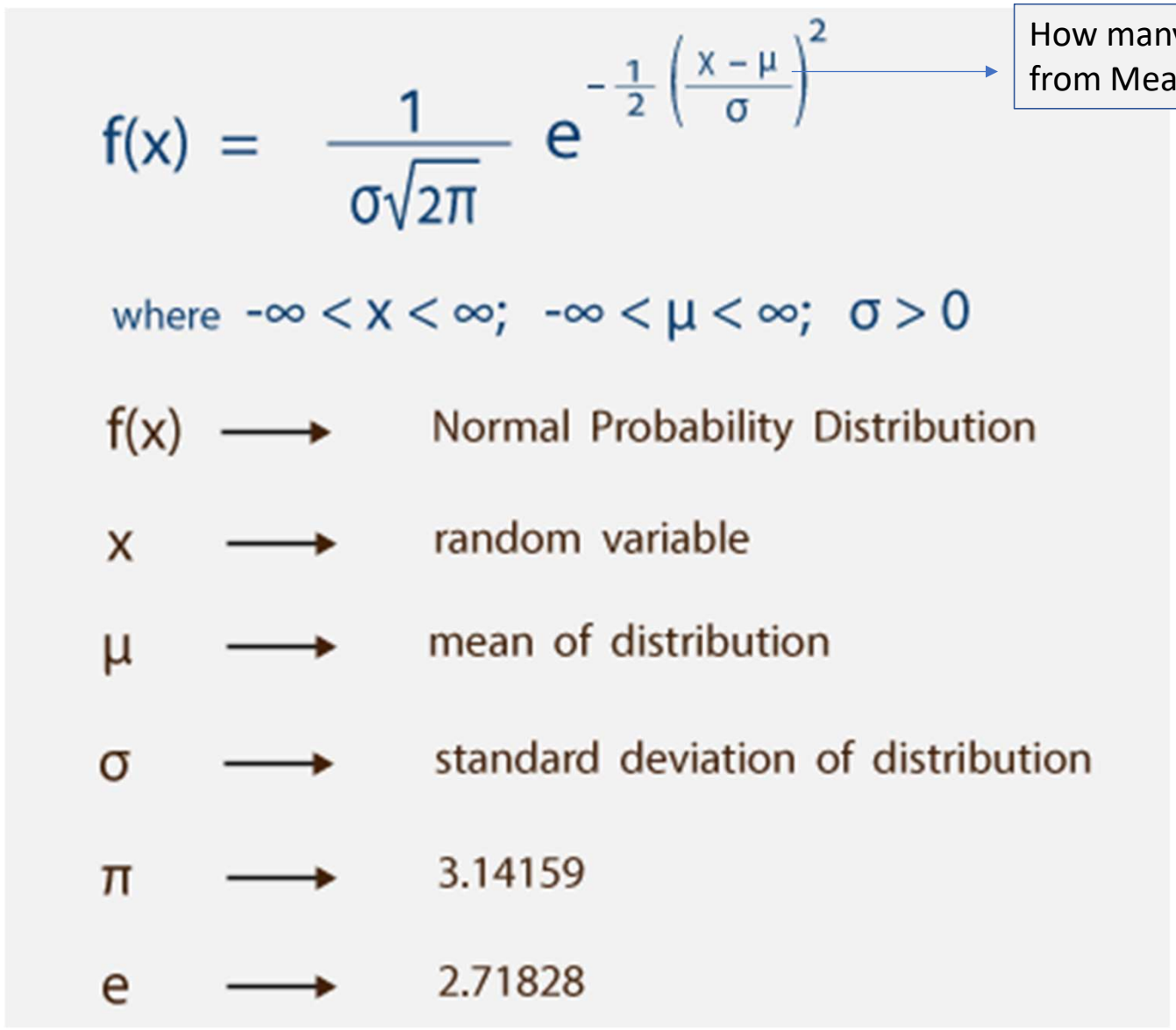
Many groups follow this type of pattern. That's why it's widely used in business, statistics and in government bodies like the FDA:

- Heights of people.
- Measurement errors.
- Blood pressure.
- Points on a test.
- IQ scores.
- Salaries

The empirical rule tells you what percentage of your data falls within a certain number of standard deviations from the mean:

- 68% of the data falls within one standard deviation of the mean.
- 95% of the data falls within two standard deviations of the mean.
- 99.7% of the data falls within three standard deviations of the mean.



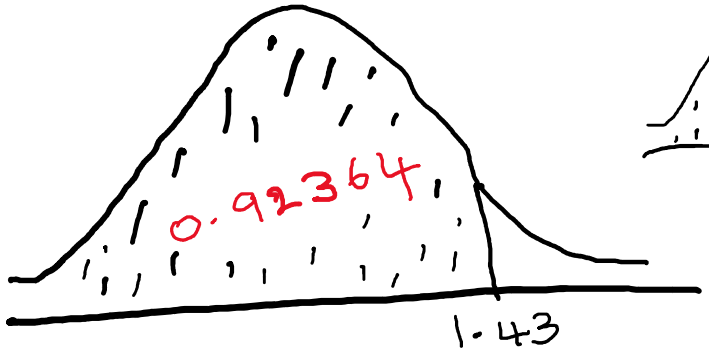

How many SD we are away from Mean? Aka **Z-score**
$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$
where $-\infty < x < \infty$; $-\infty < \mu < \infty$; $\sigma > 0$
f(x) → Normal Probability Distribution
x → random variable
μ → mean of distribution
σ → standard deviation of distribution
π → 3.14159
e → 2.71828

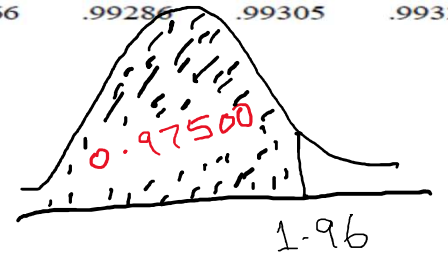
STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

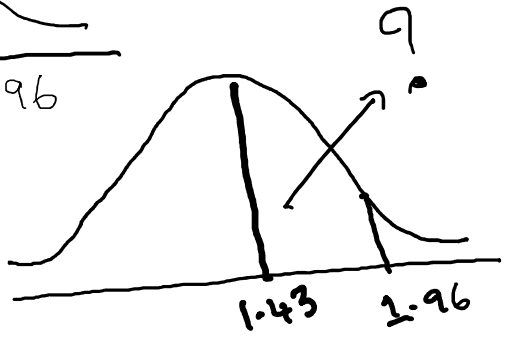
Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56749	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361

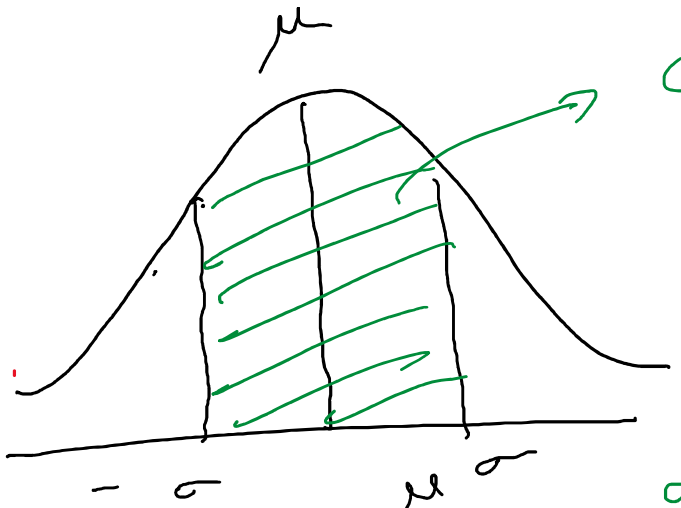
STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

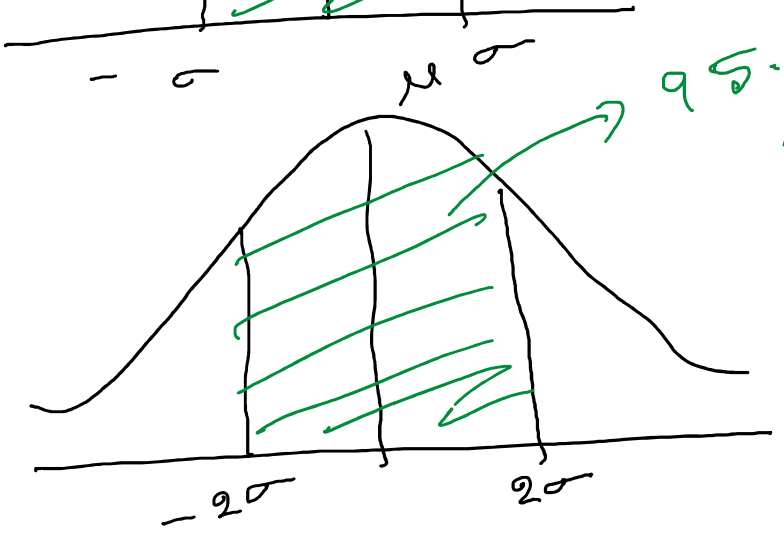
Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56749	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361

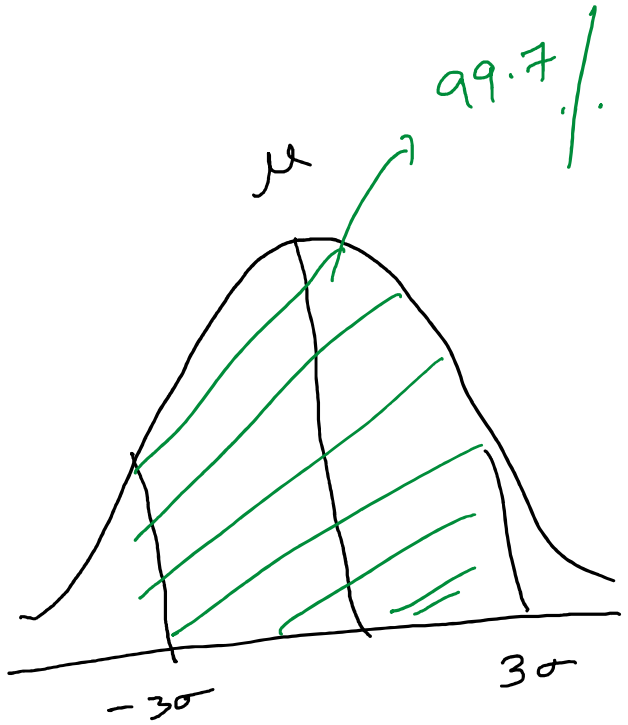







μ
68%
−σ
μ
σ


μ
95%
−2σ
2σ


μ
99.7%
−3σ
3σ

Example:-

If the salary of workers in a factory is assumed to follow a normal distribution with a mean of Rs 500 and a S.D of Rs 100.

Find number of workers whose salary vary between Rs 400 and Rs 650. (Total number of workers in the factory is 15,000)

Example:- (solution)

$$P(400 \le x \le 650)$$

$$z = \frac{x-\mu}{\sigma}$$

$$= \frac{400-500}{100} \qquad \frac{650-500}{100}$$

i.e. $P(-1 \le z \le 1.5)$

$$= F(1.5) - F(-1)$$

$$= 0.9332 - [1-F(1)]$$

$$= 0.9332 - [1-0.84134] = 0.9332 - 0.15866$$

$$= 0.77454$$

Number of students $= 0.77454 \times 15{,}000$

$$= 1161.81 \approx \boxed{1162}$$

Sampling

Definition Of Sampling

Application of certain queries to less than 100% of the population(group of all items that we are trying to observe and analyze) is known as Sampling. In simple terms, sampling is the process of selection of limited number of elements from large group of elements (population) so that, the characteristics of the samples taken is identical to that of the population. In above examples, suppose you choose 1000 students among 4 millions students. then:

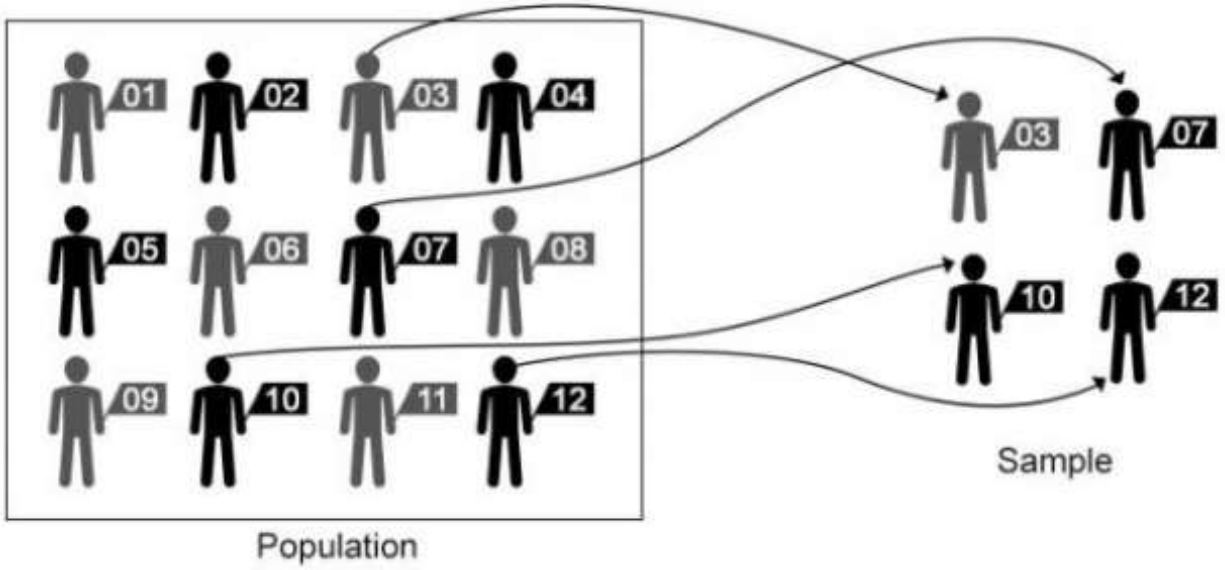
- 4 millions students is population
- 1000 is the size of sample

Sampling is a great tool if you have to deal with a huge volume of data and you have limited resources. When you have large population of the data, then it can also be the only option you have.

Although you do not subject all the data to your queries, the chance that you get the desired results is almost similar to that when you do thorough checking. Provided that your choice for the sampling techniques must be appropriate.

How Sampling Works?

First of all, we have to choose the basis of sampling, i.e. the rule that will determine whether a sample is chosen or not. After we are sure of the method which will be used for the process, you select the samples as specified in the previously set plan. The method used for choosing the samples as the very name suggests, is the most crucial part of the whole process, it defines whether the analysis accurately describes the entire population or not.



Advantages of Sampling

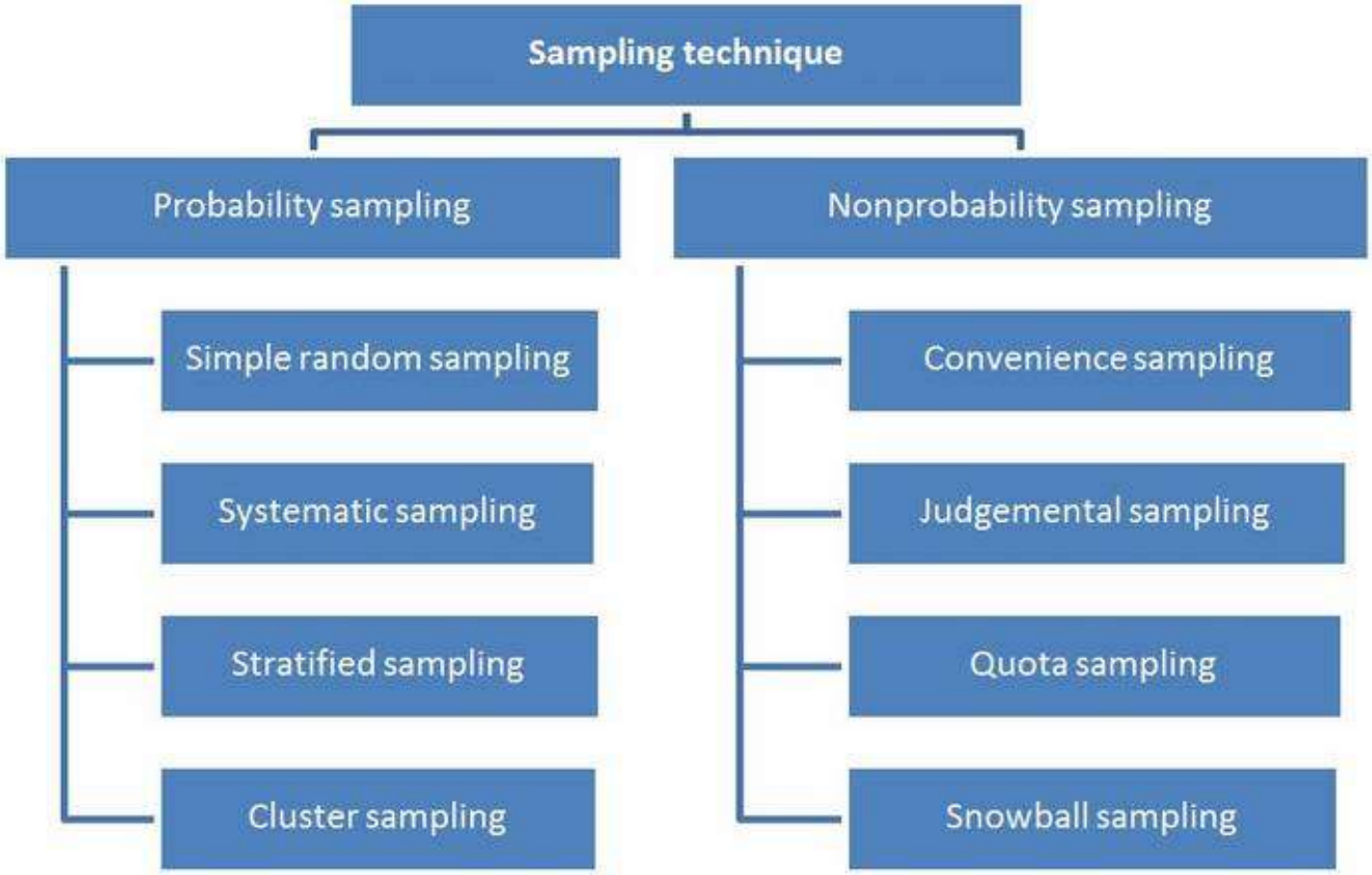
Sampling have various benefits to us. Some of the advantages are listed below:

- Sampling saves time to a great extent by reducing the volume of data. You do not go through each of the individual items.
- Sampling Avoids monotony in works. You do not have to repeat the query again and again to all the individual data.
- When you have limited time, survey without using sampling becomes impossible. It allows us to get near-accurate results in much lesser time
- When you use proper methods, you are likely to achieve higher level of accuracy by using sampling than without using sampling in some cases due to reduction in monotony, data handling issues etc.
- By using sampling, you can get detailed information on the data even by employing small amount of resources.

Disadvantages of Sampling

Every coin has two sides. Sampling also have some demerits. Some of the disadvantages are:

- Since choice of sampling method is a judgmental task, there exist chances of biasness as per the mindset of the person who chooses it.
- Improper selection of sampling techniques may cause the whole process to defunct.
- Selection of proper size of samples is a difficult job.
- Sampling may exclude some data that might not be homogenous to the data that are taken. This affects the level of accuracy in the results.


Sampling technique
Probability sampling
Nonprobability sampling
Simple random sampling
Systematic sampling
Stratified sampling
Cluster sampling
Convenience sampling
Judgemental sampling
Quota sampling
Snowball sampling

Sampling Technique

1. **Probability sampling:** Probability sampling is a sampling technique where a researcher sets a selection of a few criteria and chooses members of a population randomly. All the members have an equal opportunity to be a part of the sample with this selection parameter.
2. **Non-probability sampling:** In non-probability sampling, the researcher chooses members for research at random. This sampling method is not a fixed or predefined selection process. This makes it difficult for all elements of a population to have equal opportunities to be included in a sample.

Probability Sampling

Simple random sampling: One of the best probability sampling techniques that helps in saving time and resources, is the Simple Random Sampling **method. It is a reliable method of obtaining information where every single member of a population is chosen randomly, merely by chance.** Each individual has the same probability of being chosen to be a part of a sample.
For example, in an organization of 500 employees, if the HR team decides on conducting team building activities, it is highly likely that they would prefer picking chits out of a bowl. In this case, each of the 500 employees has an equal opportunity of being selected.

Cluster sampling: Cluster sampling is a method where the researchers divide the entire population into sections or clusters **that represent a population. Clusters are identified and included in a sample based on demographic parameters like age, sex, location, etc.** This makes it very simple for a survey creator to derive effective inference from the feedback.
For example, if the United States government wishes to evaluate the number of immigrants living in the Mainland US, they can divide it into clusters based on states such as California, Texas, Florida, Massachusetts, Colorado, Hawaii, etc. This way of conducting a survey will be more effective as the results will be organized into states and provide insightful immigration data.

Probability Sampling

Systematic sampling: Researchers use the systematic sampling method to choose the sample members of a **population at regular intervals. It requires the selection of a starting point for the sample** and sample size that can be repeated at regular intervals. This type of sampling method has a predefined range, and hence this sampling technique is the least time-consuming.
For example, a researcher intends to collect a systematic sample of 500 people in a population of 5000. He/she numbers each element of the population from 1-5000 and will choose every 10th individual to be a part of the sample (Total population/ Sample Size = 5000/500 = 10).

Stratified random sampling: Stratified random sampling is a method in which the researcher divides the population into smaller groups that don't overlap but represent the entire population. While sampling, these groups can be organized and then draw a sample from each group separately.
For example, a researcher looking to analyze the characteristics of people belonging to different annual income divisions will create strata (groups) according to the annual family income. Eg – less than $20,000, $21,000 – $30,000, $31,000 to $40,000, $41,000 to $50,000, etc. By doing this, the researcher concludes the characteristics of people belonging to different income groups. Marketers can analyze which income groups to target and which ones to eliminate to create a roadmap that would bear fruitful results.

Non - Probability Sampling

Convenience sampling: This method is dependent on the ease of access to subjects such as surveying customers at a mall or passers-by on a busy street. It is usually termed as convenience sampling, because of the researcher's **ease of carrying it out and getting in touch with the subjects**. Researchers have nearly no authority to select the sample elements, and it's purely done based on proximity and not representativeness. This non-probability sampling method is used when there are time and cost limitations in collecting feedback. In situations where there are resource limitations such as the initial stages of research, convenience sampling is used.

For example, startups and NGOs usually conduct convenience sampling at a mall to distribute leaflets of upcoming events or promotion of a cause – they do that by standing at the mall entrance and giving out pamphlets randomly.

Judgmental or purposive sampling: Judgmental or purposive samples are formed by **the discretion of the researcher.** Researchers purely consider the purpose of the study, along with the understanding of the target audience. For instance, when researchers want to understand the thought process of people interested in studying for their master's degree. The selection criteria will be: "Are you interested in doing your masters in …?" and those who respond with a "No" are excluded from the sample.

Non - Probability Sampling

Snowball sampling: Snowball sampling is a sampling method that researchers apply when the subjects are difficult to trace. For example, it will be extremely challenging to survey shelterless people or illegal immigrants**. In such cases, using the snowball theory, researchers can track a few categories to interview and derive results.** Researchers also implement this sampling method in situations where the topic is highly **sensitive and not openly discussed**—for example, surveys to gather information about HIV Aids. Not many victims will readily respond to the questions. Still, researchers can contact people they might know or volunteers associated with the cause to get in touch with the victims and collect information.

Quota sampling: In Quota sampling, the selection of members in this sampling technique happens based on a pre-set standard. In this case, **as a sample is formed based on specific attributes, the created sample will have the same qualities found in the total population**. It is a rapid method of collecting samples.

Central Limit Theorem

The central limit theorem in statistics states that, given a sufficiently large sample size, the sampling distribution of the mean for a variable will approximate a normal distribution regardless of that variable's distribution in the population.

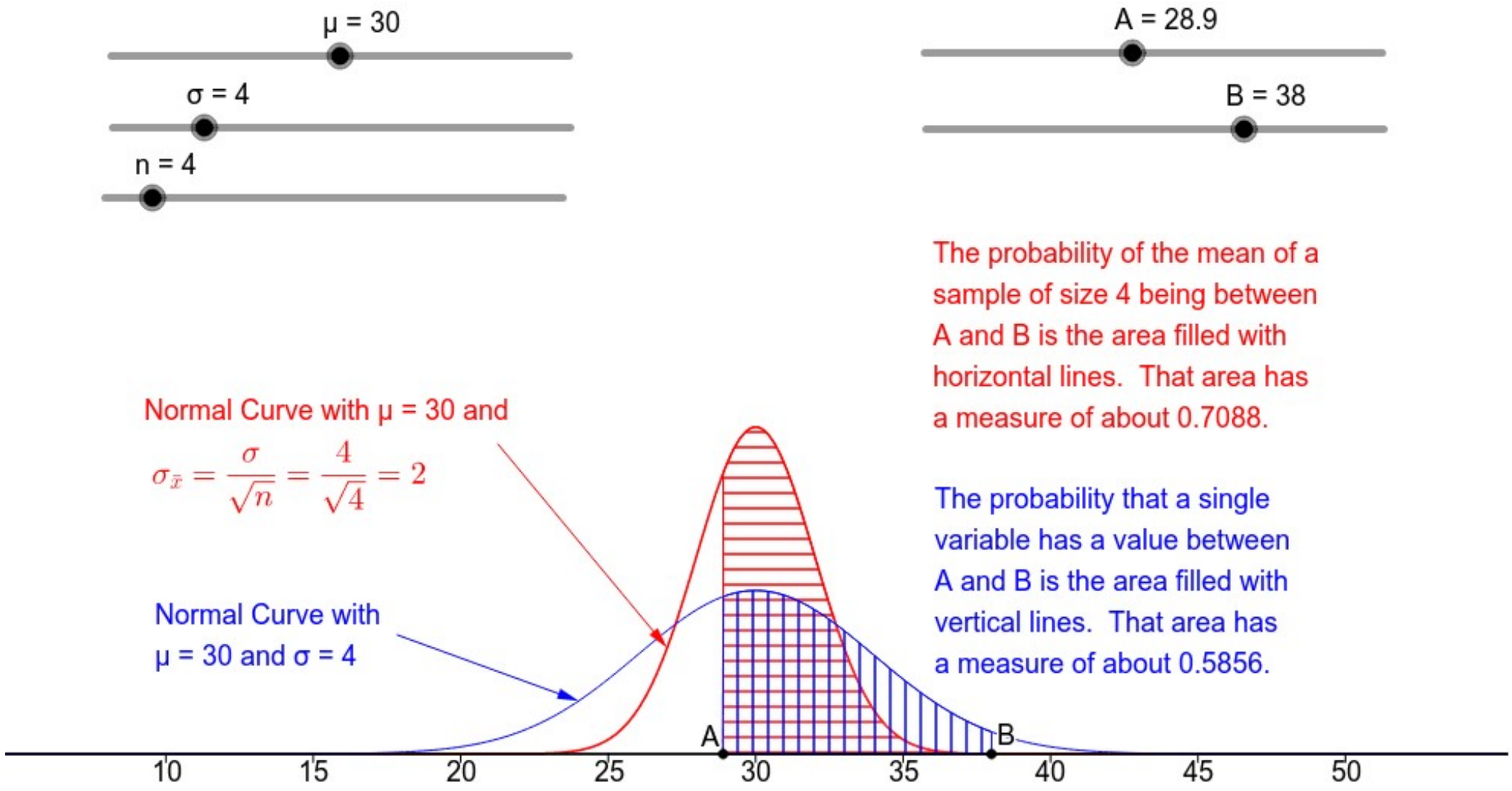
Central Limit theorem

Let $X_1, X_2 \cdots X_n$ be a random sample from a distribution with mean μ & variance σ^2. Then if n is sufficiently large, $\bar{X}$ has approximately a normal distribution with $\mu_{\bar{x}} = \mu$ and

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}.$$

The larger the value of n, better the approximation.

Can be any distribution..


μ = 30
σ = 4
n = 4
A = 28.9
B = 38
The probability of the mean of a sample of size 4 being between A and B is the area filled with horizontal lines. That area has a measure of about 0.7088.
Normal Curve with μ = 30 and
$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{4}{\sqrt{4}} = 2$
Normal Curve with
μ = 30 and σ = 4
The probability that a single variable has a value between A and B is the area filled with vertical lines. That area has a measure of about 0.5856.
A
B
10
15
20
25
30
35
40
45
50

Hypothesis Testing

Hypothesis

A hypothesis is an educated guess about something in the world around you. It should be testable, either by experiment or observation. For example:

- A new medicine you think might work.
- A way of teaching you think might be better.
- A possible location of new species.
- A fairer way to administer standardized tests.

It can really be anything at all as long as you can put it to the test.

If you are going to propose a hypothesis, it's customary to write a statement. Your statement will look like this:
"If I…(do this to an independent variable)….then (this will happen to the dependent variable)."
For example:

- If I (decrease the amount of water given to herbs) then (the herbs will increase in size).
- If I (give patients counseling in addition to medication) then (their overall depression scale will decrease).
- If I (give exams at noon instead of 7) then (student test scores will improve).
- If I (look in this certain location) then (I am more likely to find new species).

Hypothesis Testing

Hypothesis testing in statistics is a way for you to test the results of a survey or experiment to see if you have meaningful results. You're basically testing whether your results are valid by figuring out the odds that your results have happened by chance. If your results may have happened by chance, the experiment won't be repeatable and so has little use.

The best way to determine whether a statistical hypothesis is true would be to examine the entire population. **Since that is often impractical, researchers typically examine a random sample from the population. If sample data are not consistent with the statistical hypothesis, the hypothesis is rejected.**

There are two types of statistical hypotheses.

1. **Null hypothesis.** The null hypothesis, denoted by Ho, is usually the hypothesis that sample observations result purely from chance.
2. **Alternative hypothesis.** The alternative hypothesis, denoted by H1 or Ha, is the hypothesis that sample observations are influenced by some non-random cause.

Hypothesis Tests Standard Approach

Statisticians follow a formal process to determine whether to reject a null hypothesis, based on sample data. This process, called **hypothesis testing, consists of four steps**.

1. **State the hypotheses.** This involves stating the null and alternative hypotheses. The hypotheses are stated in such a way that they are mutually exclusive. That is, if one is true, the other must be false.

2. **Formulate an analysis plan.** The analysis plan describes how to use sample data to evaluate the null hypothesis. The evaluation often focuses around a single test statistic.

3. **Analyze sample data.** Find the value of the test statistic (mean score, proportion, t statistic, z-score, etc.) described in the analysis plan.

4. **Interpret results.** Apply the decision rule described in the analysis plan. If the value of the test statistic is unlikely, based on the null hypothesis, reject the null hypothesis.

Null & Alternate Hypothesis

For example, suppose we wanted to determine whether a coin was fair and balanced. A null hypothesis might be **that half the flips would result in Heads and half, in Tails**. *The alternative hypothesis might be that the number of Heads and Tails would be very different.* Symbolically, these hypotheses would be expressed as

Ho: $P = 0.5$
Ha: $P \neq 0.5$

Suppose we flipped the coin 50 times, resulting in 40 Heads and 10 Tails. Given this result, we would be inclined to reject the null hypothesis. We would conclude, based on the evidence, that the coin was probably not fair and balanced.

Decision Errors

Two types of errors can result from a hypothesis test.

- **Type I error.** A Type I error occurs when the researcher rejects a null hypothesis when it is true. The probability of committing a Type I error is called the significance level. **This probability is also called alpha, and is often denoted by α.**

- **Type II error.** A Type II error occurs when the researcher fails to reject a null hypothesis that is false. The probability of committing a Type II error is called Beta, and is often denoted by β. **The probability of not committing a Type II error is called the Power of the test.**

If we look at what can happen in a hypothesis test, we can construct the following contingency table:

	In Reality	
Decision	H_0 is TRUE	H_0 is FALSE
Accept H_0	OK	Type II Error β = probability of Type II Error
Reject H_0	Type I Error α = probability of Type I Error	OK

You should be familiar with type I and type II errors from your introductory course. It is important to note that we want to set α before the experiment (*a-priori*) because the Type I error is the more 'grevious' error to make. The typical value of α is 0.05, establishing a 95% confidence level. **For this course we will assume α =0.05.**

Decision Rules

The analysis plan includes decision rules for rejecting the null hypothesis. In practice, statisticians describe these decision rules in two ways - **with reference to a P-value or with reference to a region of acceptance**.

P-value. The strength of evidence in support of a null hypothesis is measured by the P-value. Suppose the test statistic is equal to S. The P-value is the probability of observing a test statistic as extreme as S, assuming the null hypothesis is true. If the P-value is less than the significance level, we reject the null hypothesis.

Region of acceptance. The region of acceptance is a range of values. If the test statistic falls within the region of acceptance, the null hypothesis is not rejected. The region of acceptance is defined so that the chance of making a Type I error is equal to the significance level.
The set of values outside the region of acceptance is called the region of rejection. If the test statistic falls within the region of rejection, the null hypothesis is rejected. In such cases, we say that the hypothesis has been **rejected at the α level of significance.**

These approaches are equivalent. Some statistics texts use the P-value approach; others use the region of acceptance approach.

One-Tailed and Two-Tailed Tests

A test of a **statistical hypothesis, where the region of rejection is on only one side of the sampling distribution, is called a one-tailed test.** For example, suppose the null hypothesis states that the mean is less than or equal to 10. The alternative hypothesis would be that the mean is greater than 10. The region of rejection would consist of a range of numbers located on the right side of sampling distribution; that is, a set of numbers greater than 10.

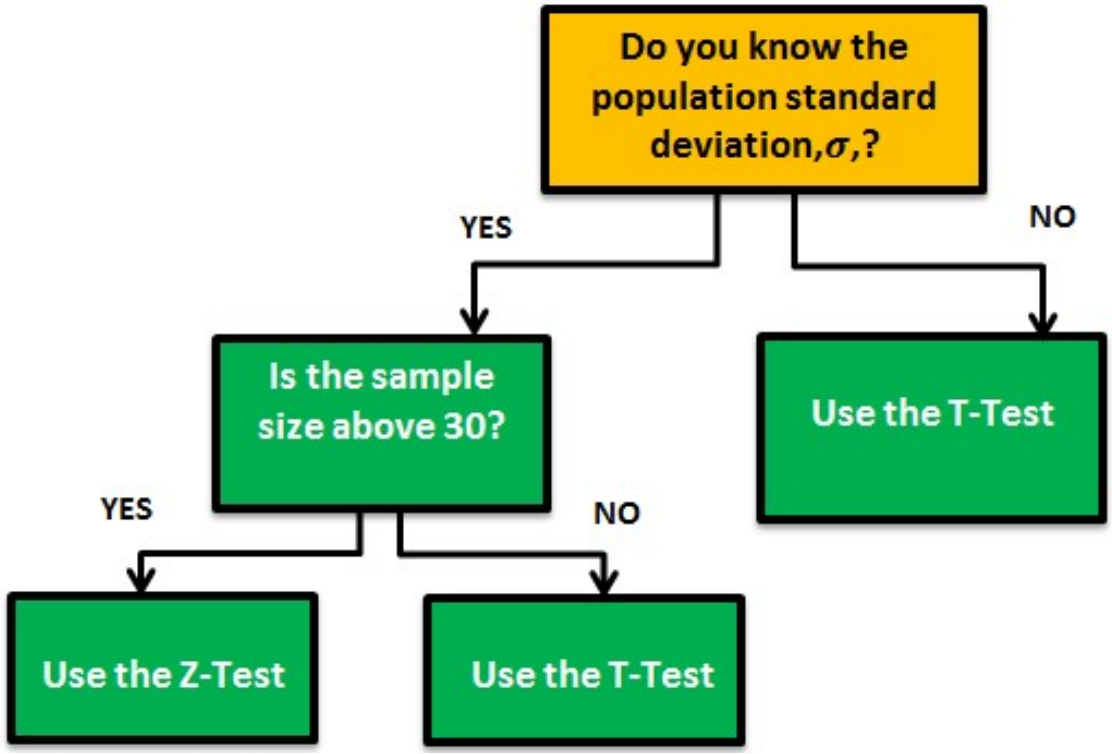
A test of a statistical hypothesis, where the **region of rejection is on both sides of the sampling distribution, is called a two-tailed test.** For example, suppose the null hypothesis states that the mean is equal to 10. The alternative hypothesis would be that the mean is less than 10 or greater than 10. The region of rejection would consist of a range of numbers located on both sides of sampling distribution; that is, the region of rejection would consist partly of numbers that were less than 10 and partly of numbers that were greater than 10.

Z Test Statistics Formula

$$Z\ Test = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

t-Test Formula

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$



One-Sample z-test

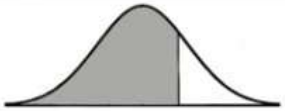
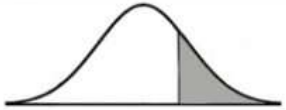
State Hypothesis

$H_0: \mu \geq M$ $H_\alpha: \mu < M$	$H_0: \mu \leq M$ $H_\alpha: \mu > M$	$H_0: \mu = M$ $H_\alpha: \mu \neq M$

Find Standardized Test Statistic

$$z = \frac{\bar{x} - M}{\frac{\sigma}{\sqrt{n}}}$$

Find the p-value

		
$H_\alpha: \mu < M$	$H_\alpha: \mu < M$	$H_\alpha: \mu \neq M$
On z-table $P(x < z)$ On calculator $NormCDF(-99, z, 0,1)$	On z-table $P(x > z)$ On calculator $NormCDF(z, 99,0,1)$	On z-table $2 \times P(x > \|z\|)$ On Calculator $2 \times [NormCDF(\|z\|, 99,0,1)]$

Conclusion

p-value$< \alpha$	Reject H_0
p-value$> \alpha$	Fail to Reject H_0

	one-tailed test		two-tailed test
hypothesis	$H_0 : \mu_1 \geq \mu_0$ $H_1 : \mu_1 < \mu_0$	$H_0 : \mu_1 \leq \mu_0$ $H_1 : \mu_1 > \mu_0$	$H_0 : \mu_1 = \mu_0$ $H_1 : \mu_1 \neq \mu_0$
test statistic (t distribution)	$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$		
deg. of freedom	$n - 1$		
rejection	reject H_0 if $t < -t_\alpha$	reject H_0 if $t > t_\alpha$	reject H_0 if $\lvert t \rvert > t_{\alpha/2}$

The **prerequisite** for the one-sample t-test is that the sample is normally distributed.

Testing of Hypothesis

H_0 : H_1

$\alpha = ?$

Decide one tailed / two tailed

$H_1: \mu \neq$ → Two

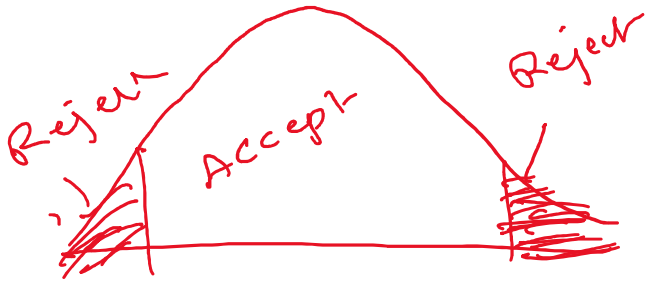
$H_1: \mu \geq$ } one
$\mu \leq$ }

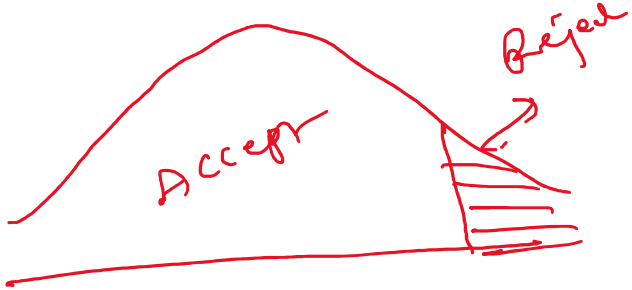
→ Choose the test

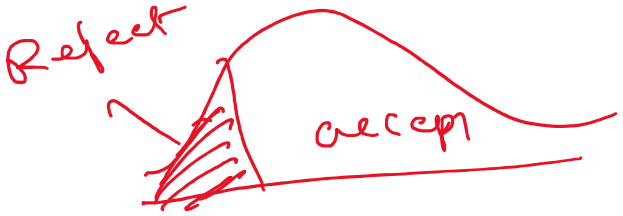
→ Calculate Z / t

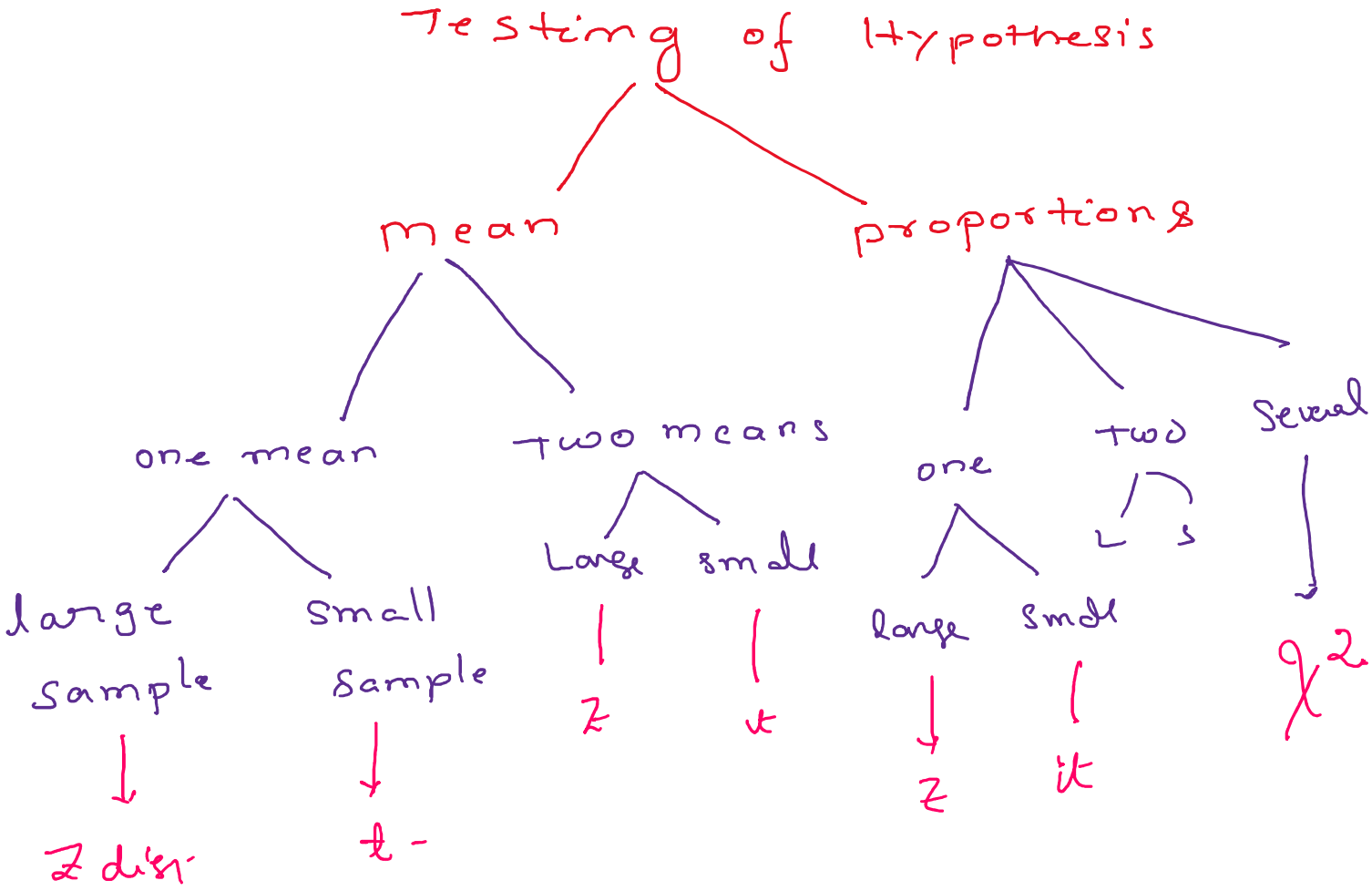
→ taking decision

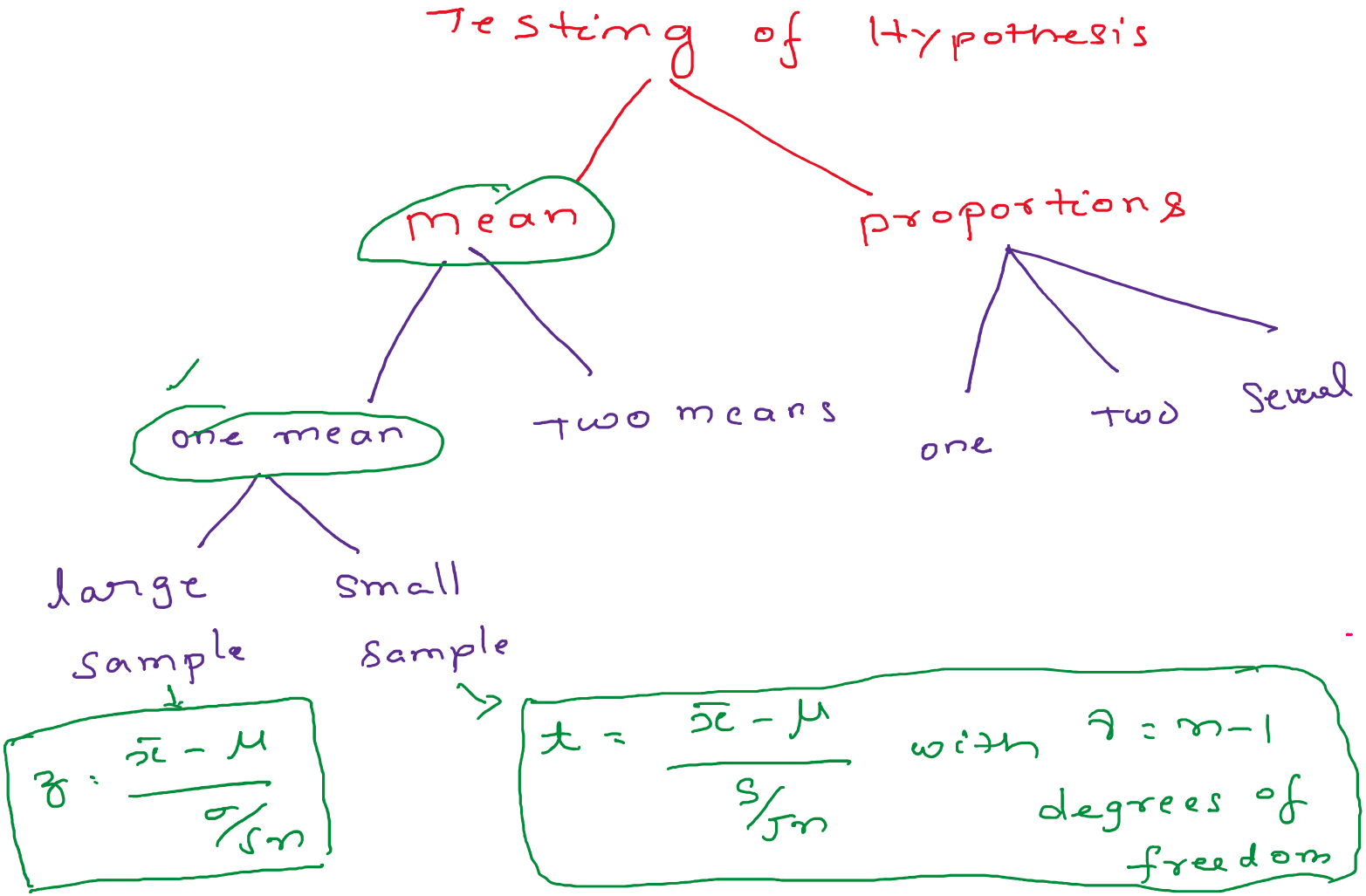
↓ Accept or reject H_0

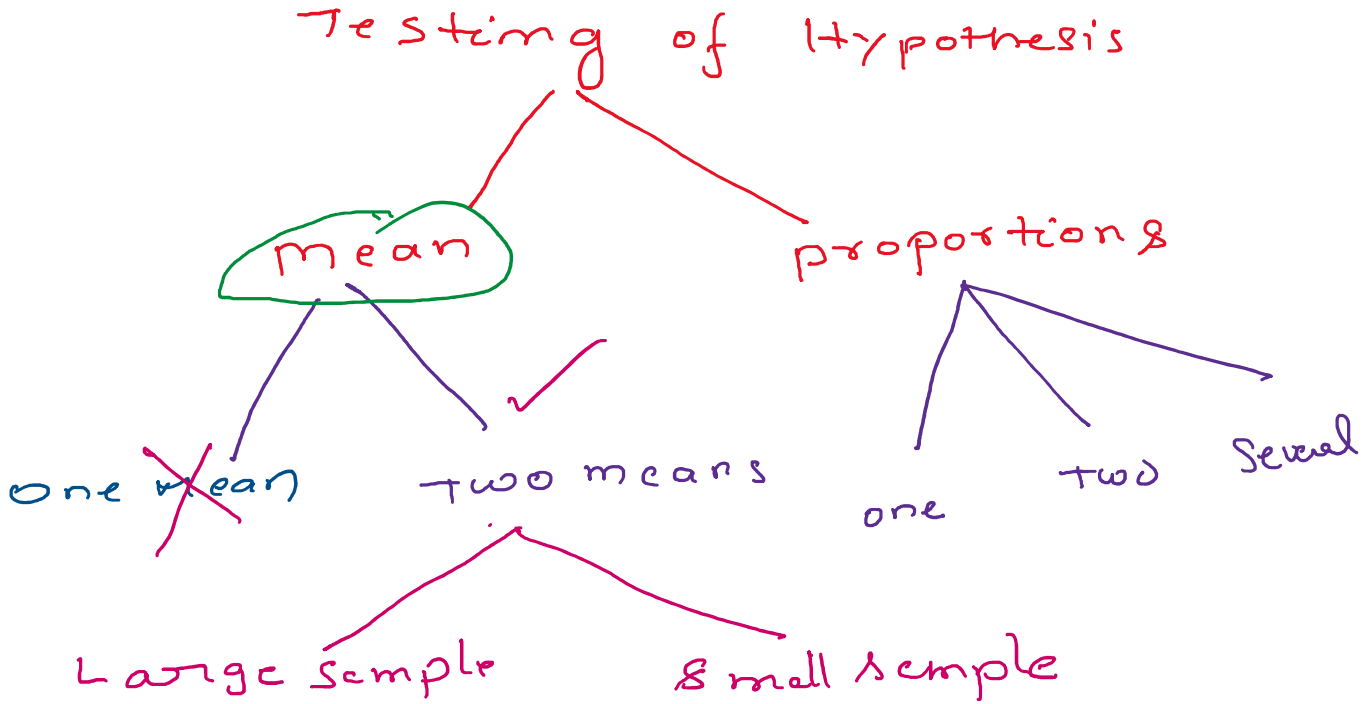




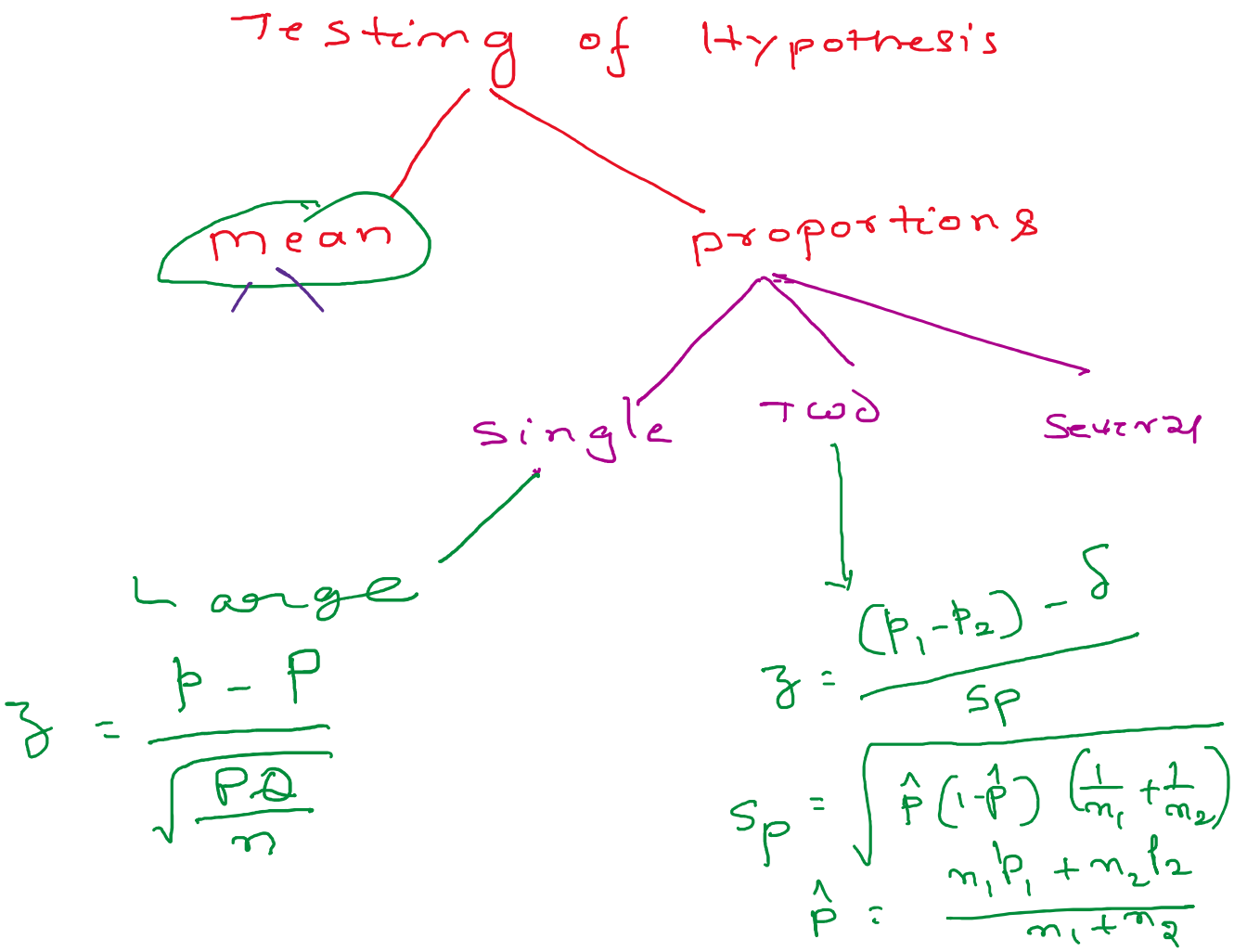


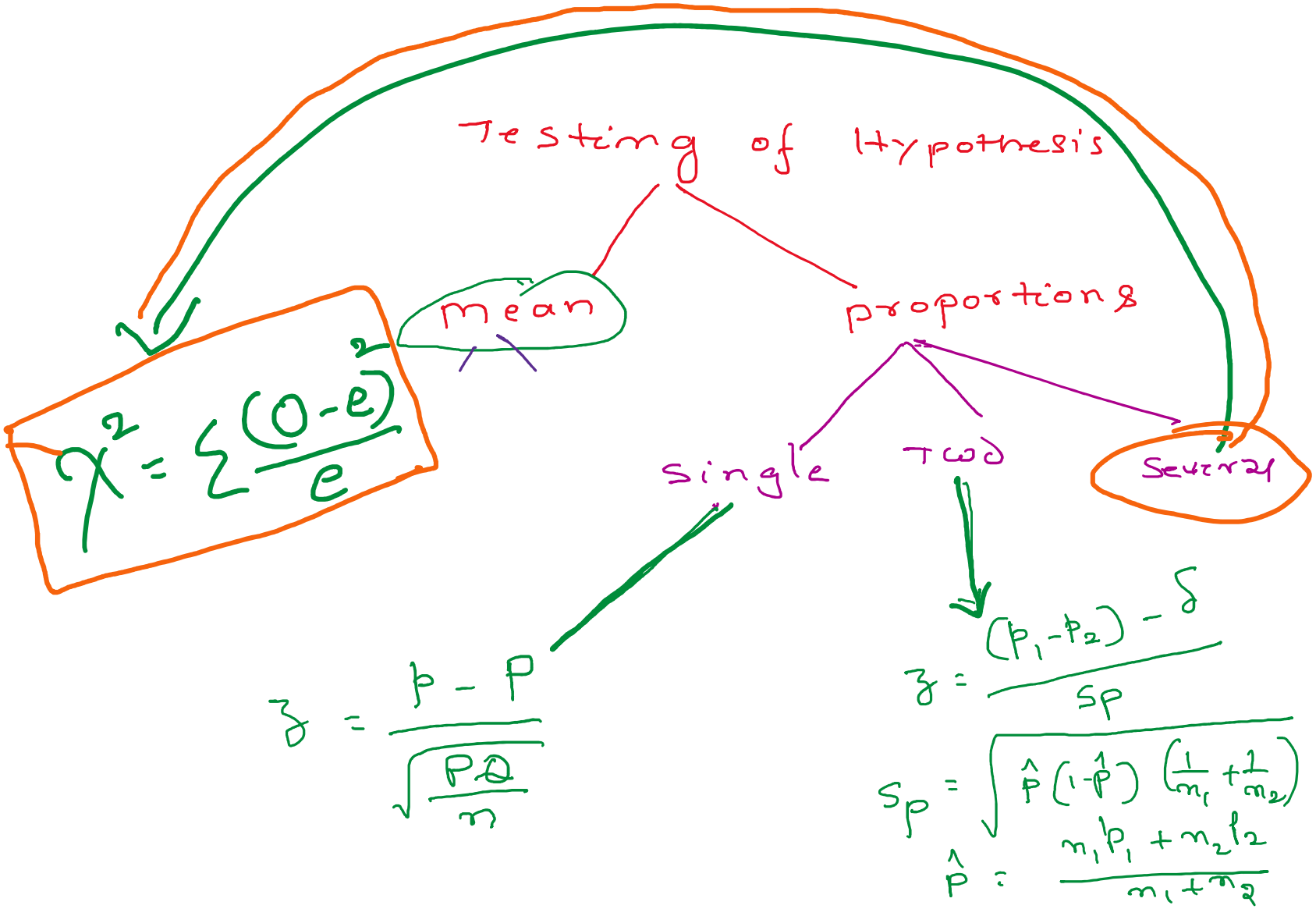

Testing of Hypothesis
mean
proportions
one mean
two means
one
two
several
large
small
Large
small
large
small
sample
sample
Z
t
Z
t
Z
t
χ^2
Z dist.
t -


Testing of Hypothesis
mean
proportions
one mean
two means
one
two
Several
large sample
small sample
$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$
$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$ with $\nu = n-1$ degrees of freedom



$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

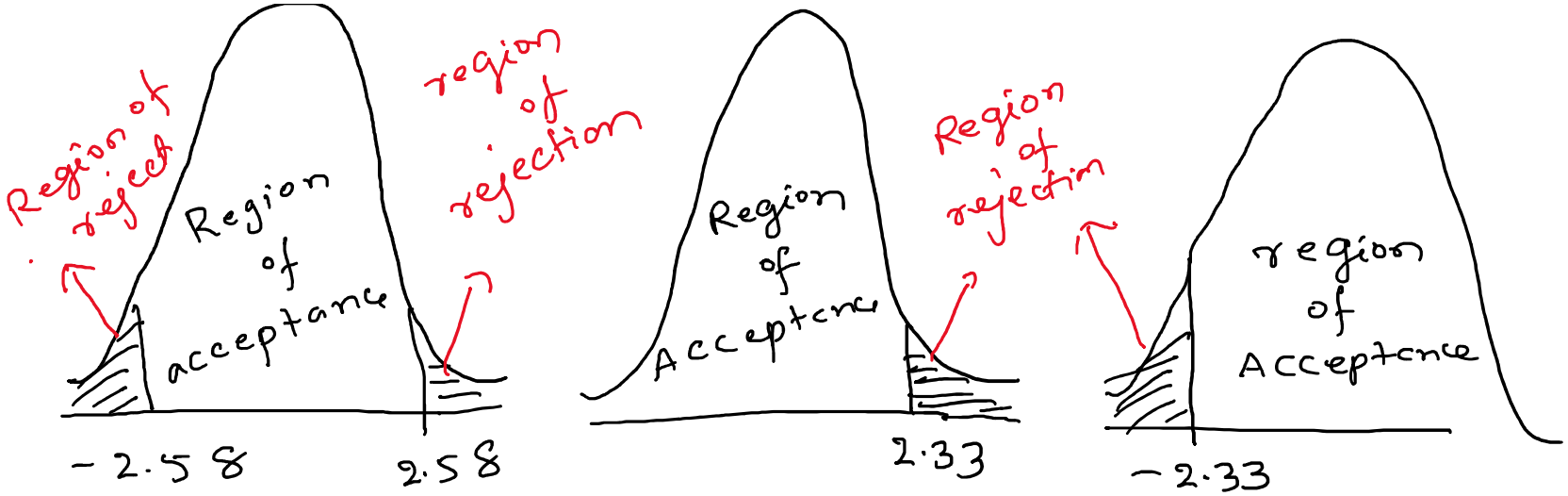

Testing of Hypothesis
mean
proportions
single
Two
Several
Large
$z = \frac{\bar{p} - P}{\sqrt{\frac{PQ}{n}}}$
$z = \frac{(p_1 - p_2) - \delta}{s_p}$
$s_p = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$
$\hat{p} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$

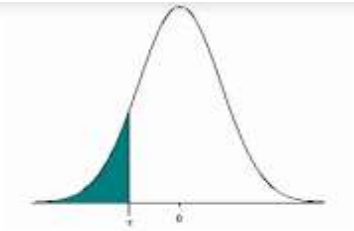

Testing of Hypothesis
mean
proportions
$\chi^2 = \sum \frac{(O-e)^2}{e}$
single
Two
Several
$z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$
$z = \frac{(p_1 - p_2) - \delta}{SP}$
$S_p = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$
$\hat{p} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$

Critical regions

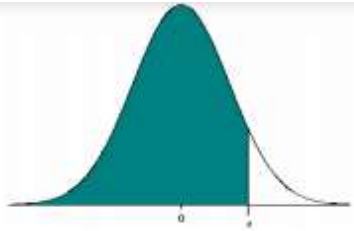
(Large Sample)

	1%	5%	10%
Two tailed test	(−2.58, 2.58)	(−1.96, 1.96)	(−1.645, 1.645)
Right tailed test	2.33	1.645	1.28
Left tailed test	−2.33	−1.645	−1.28





z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Note that the probabilities given in this table represent the area to the LEFT of the z-score.
The area to the RIGHT of a z-score = 1 – the area to the LEFT of the z-score

EXAMPLE 1

EXAMPLE

The mean lifetime of a sample of 100 items of a product by a company is 1,560 hours with a S.D of 90 hours.

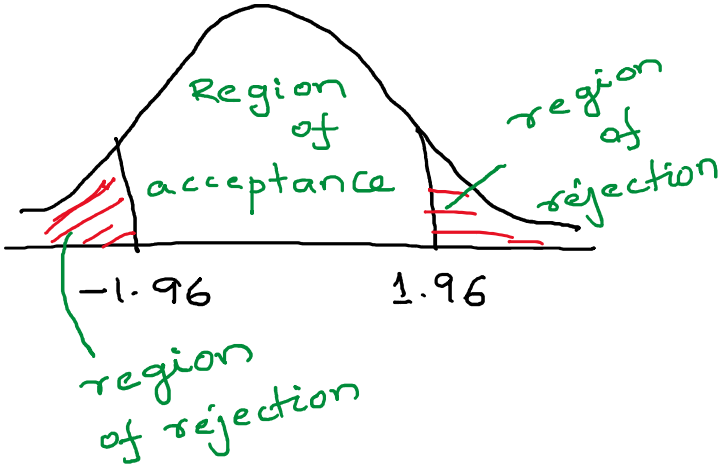
Test the hypothesis that the mean lifetime of the product is 1,580 hours.

EXAMPLE (Discussion)

Null hypothesis: $H_0 : \mu = 1580$

Alternate Hypo: $H_1 : \mu \neq 1580$

L.O.S : $\alpha = 5\%$



Calculations :-

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

$$= \frac{(1560 - 1580)}{90/\sqrt{100}}$$

$$= -2.22$$

Reject H_0

EXAMPLE 2

EXAMPLE:

The mean life time of a sample of 400 fluoresent light bulbs produced by a company is found to be 1600 hours with a S.D of 150 hours.

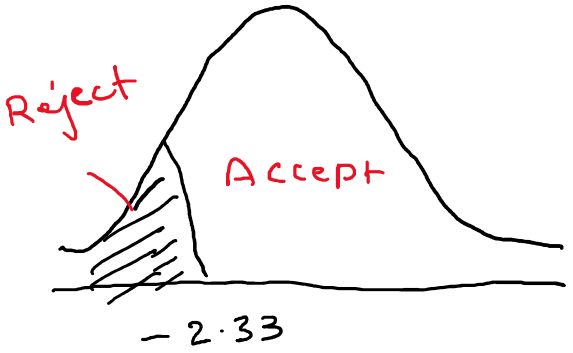
Test the hypothesis that the mean life time of the bulbs produced is higher than 1570 hours at $\alpha = 0.01$.

EXAMPLE (solution)

$H_0 : \mu \geq 1570$

$H_1 : \mu < 1570$

$\alpha : 1\%$



calculation :-

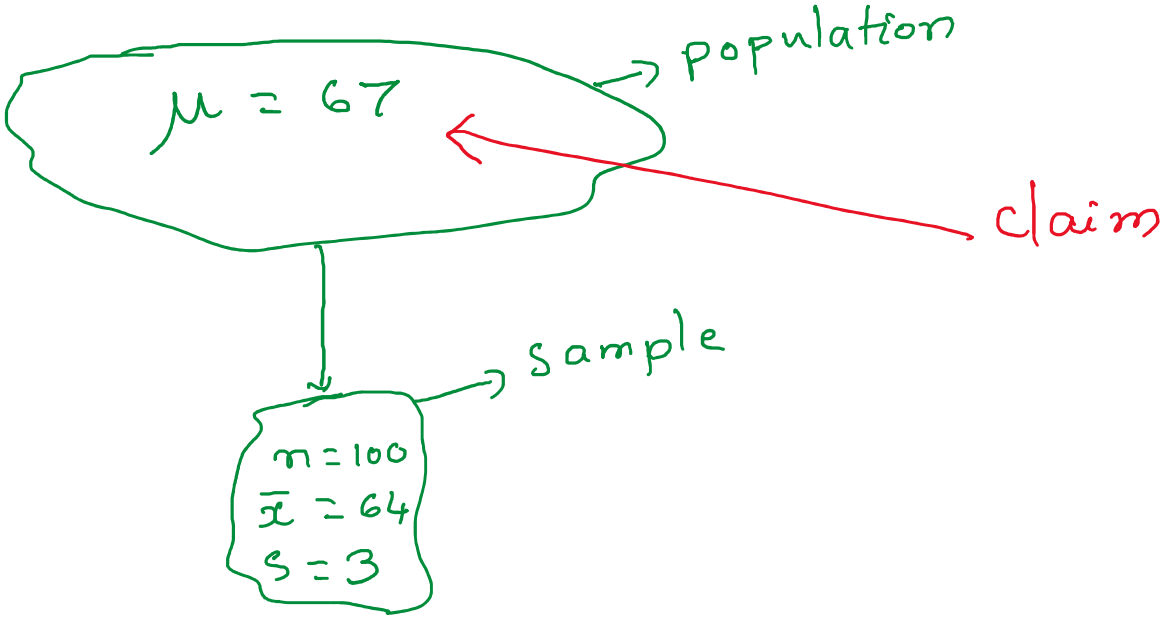
$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{1600 - 1570}{150/\sqrt{400}}$$

$$= 4$$

EXAMPLE 3

Example:.

- It is claimed that the mean of the population is 67 gms at 5% level of significance. Mean obtained from a random sample of size 100 is 64 with SD 3. Validate the claim.



Example:— Discussion

- It is claimed that the mean of the population is 67 gms at 5% level of significance. Mean obtained from a random sample of size 100 is 64 with SD 3. Validate the claim.

$H_0 : \mu = 67$

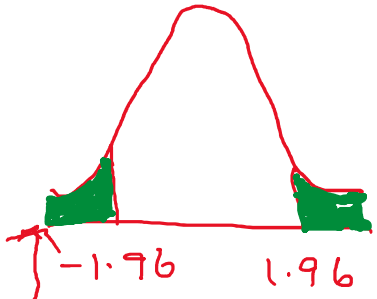
$H_1 : \mu \neq 67$

$\alpha = 5\%$

population: $\mu = 64$

sample: $n = 100$, $\bar{x} = 64$, $s = 3$

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{64 - 67}{3/\sqrt{100}} = -10$$



So H_0 is rejected

EXAMPLE 4

Example

- There is an assumption that there is no significant difference between boys and girls wrt intelligence. Tests are conducted on two groups and the following are the observations
- | | Mean. | SD. | Size |
- | Girls. | 75. | 8. | 60 |
- | Boys. | 73. | 10. | 100 |
- Validate this at 5% LoS

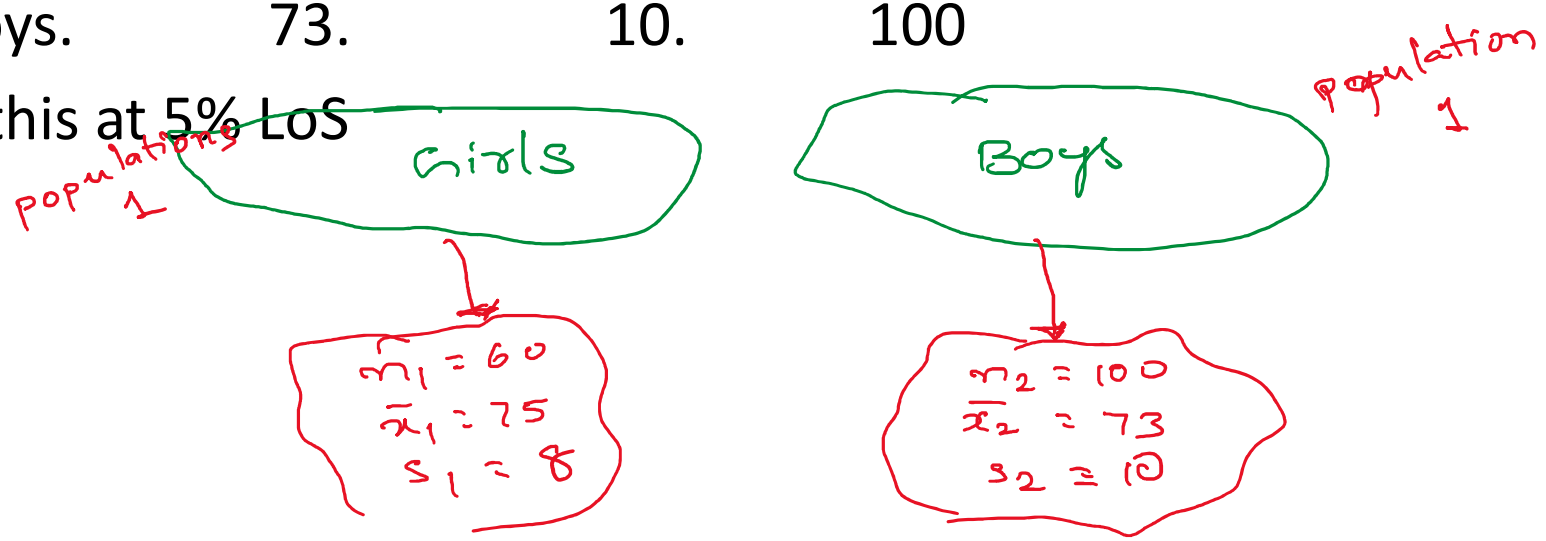
Example

Discussion

- There is an assumption that there is no significant difference between boys and girls wrt intelligence. Tests are conducted on two groups and the following are the observations

	Mean.	SD.	Size
Girls.	75.	8.	60
Boys.	73.	10.	100

- Validate this at 5% LoS



Solution :-

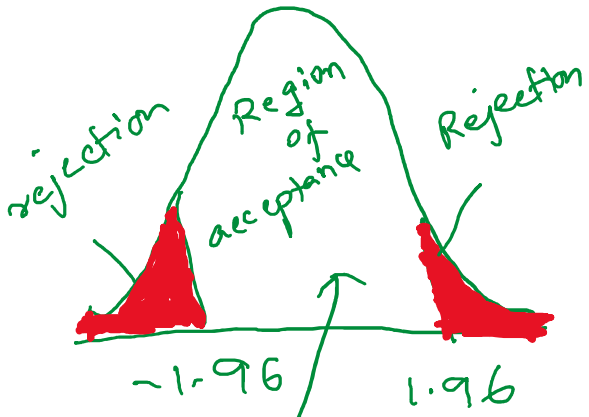
$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2 \rightarrow$ two

$\alpha = 5\%$

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$= \frac{75 - 73}{\sqrt{\frac{8^2}{60} + \frac{10^2}{100}}} = 1.3912$$

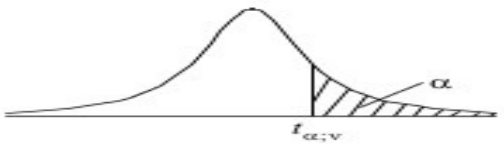


	one-tailed test		two-tailed test
hypothesis	$H_0 : \mu_1 \geq \mu_0$ $H_1 : \mu_1 < \mu_0$	$H_0 : \mu_1 \leq \mu_0$ $H_1 : \mu_1 > \mu_0$	$H_0 : \mu_1 = \mu_0$ $H_1 : \mu_1 \neq \mu_0$
test statistic (t distribution)	$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$		
deg. of freedom	$n - 1$		
rejection	reject H_0 if $t < -t_\alpha$	reject H_0 if $t > t_\alpha$	reject H_0 if $\lvert t \rvert > t_{\alpha/2}$

The **prerequisite** for the one-sample t-test is that the sample is normally distributed.

Table of the Student's *t*-distribution

The table gives the values of $t_{\alpha;\nu}$ where
$\Pr(T_\nu > t_{\alpha;\nu}) = \alpha$, with ν degrees of freedom

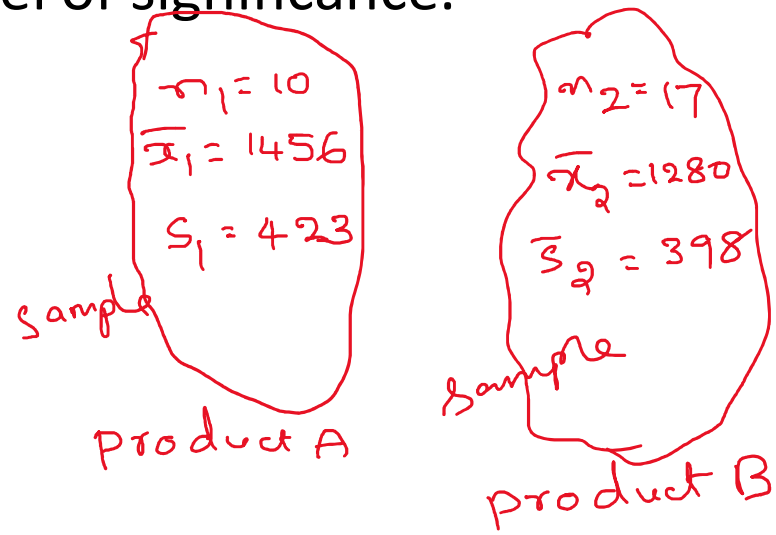


ν \ α	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.078	6.314	12.076	31.821	63.657	318.310	636.620
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

EXAMPLE 5

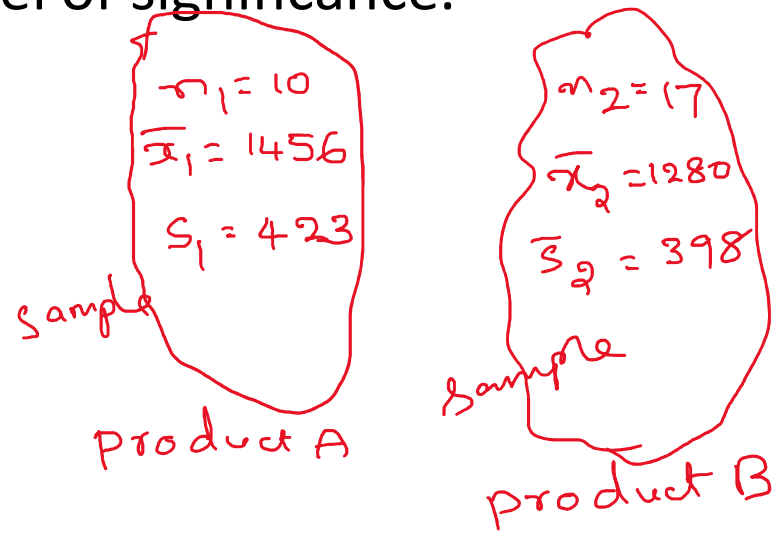
Example

- As a manager of finance you are assigned the task of choosing the best product in terms of life of the product.
- Product A: Mean 1456 hours with SD 423, size 10
- Product B: Mean 1280 hours with SD 398, size 17
- Use 5% level of significance.



Example

- As a manager of finance you are assigned the task of deciding if both the product are same in terms of life.
- Product A: Mean 1456 hours with SD 423, size 10
- Product B: Mean 1280 hours with SD 398, size 17
- Use 5% level of significance.



two population
— two means
— ?
— which test

Solution :-

$H_0 : \mu_1 = \mu_2 \quad : \quad H_1 : \mu_1 \neq \mu_2$

$$S^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{9(423)^2 + 16(398)^2}{10+17-2}$$

$$= 1,65,793$$

i.e. $S = 407.18$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{(1456-1280) - 0}{407.18 \cdot \sqrt{\frac{1}{10} + \frac{1}{17}}}$$

$$= 1.085$$

d.o.f $= n_1 + n_2 - 2 = 10 + 17 - 2 = 25$ at 5% LOS

2.06

$\therefore \quad t = 1.085 < 2.06$

H_0 accepted.

Chi squared test

The Chi Square statistic is commonly used for testing relationships between categorical variables. The null hypothesis of the Chi-Square test is that no relationship exists on the categorical variables in the population; they are independent.

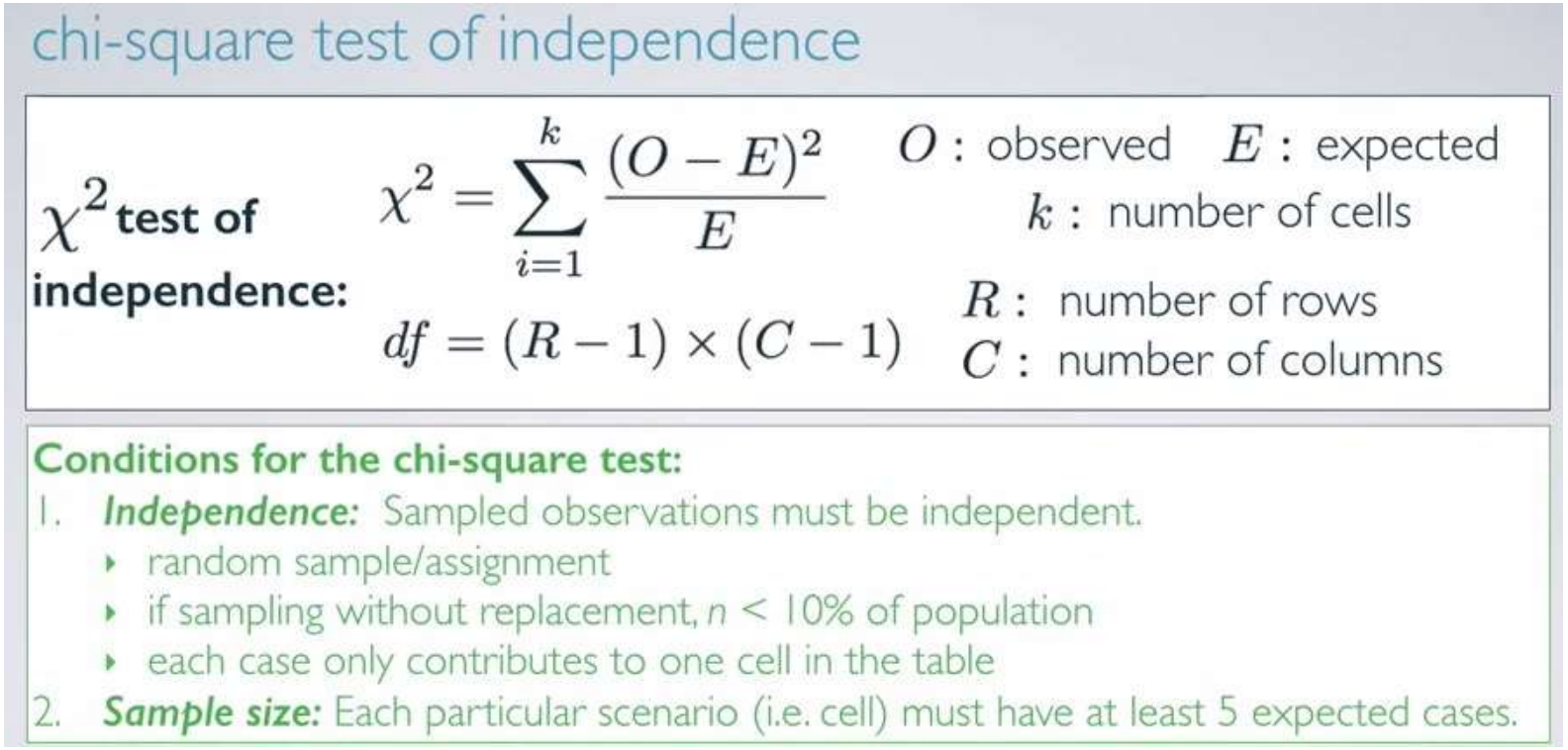


Table II
Chi-square Distribution

The following table presents selected quantiles of chi-square distribution; i.e, the values x such that

$$P(X \leq x) = \int_0^x \frac{1}{\Gamma(r/2)2^{r/2}} w^{r/2-1} e^{-w/2}\, dw,$$

for selected degrees of freedom r.

	$P(X \leq x)$							
r	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
1	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345
4	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277
5	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086
6	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812
7	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475
8	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090
9	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666
10	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209
11	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725
12	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217
13	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688
14	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141
15	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578
16	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000
17	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409
18	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805
19	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191
20	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566
21	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932
22	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289
23	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638
24	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980
25	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314
26	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642
27	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963
28	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278
29	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588
30	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892

EXAMPLE 6

Example:-

Following is the record of number of accidents took place during the various days of the week.

Monday	Tuesday	wednesday	Thursday	Friday	Saturday	Sunday
184	148	145	153	150	154	116

Can we conclude that accident is independent of the day in a week?

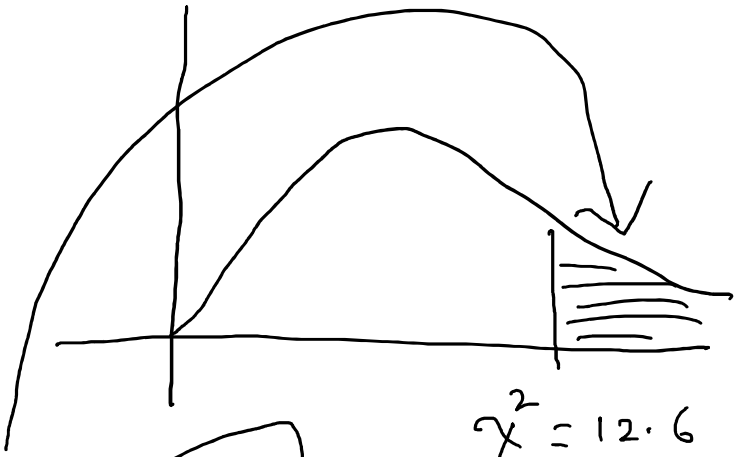
$$\chi^2 = \sum \frac{(O-e)^2}{e}$$

$1050 \times \frac{1}{7} = 150$

O	e	$(O-e)^2/e$
184	150	7.71
148	150	0.03
145	150	0.17
153	150	0.06
150	150	0
154	150	0.11
116	150	7.71
Total 1050		$\chi^2 = 15.77$

H_0: accident is indept of the day wise

H_1: Not



$\chi^2 = 12.6$ at 5% LOS & 6 d.o.f

Reject H_0

EXAMPLE 7

Example:

	with cancer	without cancer	total
Smokers	400	300	700
Non-Smokers	300	500	800
total	700	800	1500

Can we conclude that smoking causes cancer?

Discussion: H_0: Smoking causes cancer

H_1: No ——.

	with Cancer	without cancer	total
Smokers	400 / 326	300 / 373	700
Non-Smokers	300 / 373	500 / 426	800
Total	700	800	1500

$\frac{700 \times 700}{1500}$, $\frac{700 \times 800}{1500}$

$\frac{700 \times 800}{1500}$, $\frac{800 \times 800}{1500}$

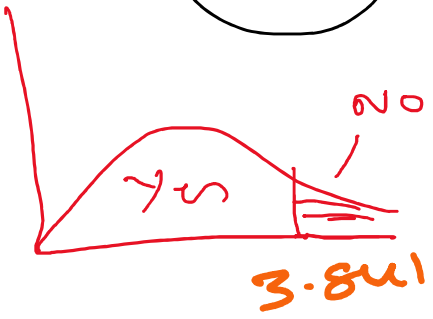
$$\chi^2 = \frac{(O-e)^2}{e} = \frac{(400-326)^2}{326} + \frac{(300-373)^2}{373} + \frac{(300-373)^2}{373} + \frac{(500-426)^2}{426} =$$

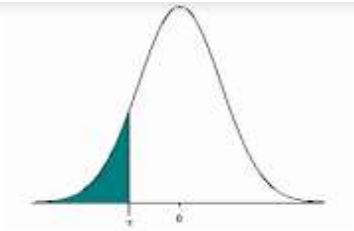
	with Cancer	without cancer	total
Smokers	400 / 326	300 / 373	700
Non-Smokers	300 / 373	500 / 426	800
total	700	800	1500

From tables,

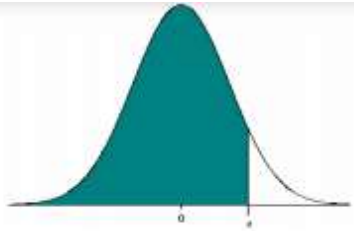
$\alpha = 5\%$

d.o.f : $(2-1)(2-1) = 1$





z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

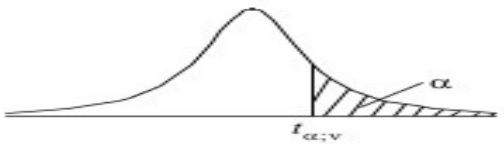


z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Note that the probabilities given in this table represent the area to the LEFT of the z-score.
The area to the RIGHT of a z-score = 1 – the area to the LEFT of the z-score

Table of the Student's *t*-distribution

The table gives the values of $t_{\alpha;\nu}$ where
$\Pr(T_\nu > t_{\alpha;\nu}) = \alpha$, with ν degrees of freedom



ν \ α	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.078	6.314	12.076	31.821	63.657	318.310	636.620
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Table II
Chi-square Distribution

The following table presents selected quantiles of chi-square distribution; i.e, the values x such that

$$P(X \leq x) = \int_0^x \frac{1}{\Gamma(r/2)2^{r/2}} w^{r/2-1} e^{-w/2}\, dw,$$

for selected degrees of freedom r.

	$P(X \leq x)$							
r	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
1	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345
4	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277
5	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086
6	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812
7	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475
8	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090
9	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666
10	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209
11	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725
12	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217
13	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688
14	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141
15	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578
16	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000
17	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409
18	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805
19	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191
20	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566
21	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932
22	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289
23	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638
24	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980
25	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314
26	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642
27	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963
28	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278
29	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588
30	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892

Thanks