TEST OF SIGNIFICANCE

A STATISTICAL REPORT

SUBMITTED AS A PART OF UG STATISTICS SYLLABUS

BAHONA COLLEGE, BAHONA-JORHAT

NAME OF THE STUDENT

FAIZA ALMAS AHMED

EXAM ROLLNO.15120106

REGISTRATION NO.S1838037

YEAR:2018-2021 OF DIBRUGARH UNIVERSITY

PROJECT SUBMITTED TO

MR.LATIT KUMAR KAKOTI

HEAD OF THE DEPARTMENT

STATISTICS, BAHONA COLLEGE

2021

CERTIFICATE

The undersigned hereby certify that the project report entitled " TEST OF SIGNIFICANCE " is written and submitted by FAIZA ALMAS AHMED to the department of statistics (Bahona College) , represent our original work and interpretation which is based on material collected.

NAME OF STUDENT : MISS FAIZA ALMAS AHMED

ROLL NO. : 15120106

CLASS : B.SC 6th SEMESTER

SIGNATURE:

ACKNOWLEDGEMENT

I take these opportunities to express my profound gratitude and deep regards to my mentors and guiding sources, Mr. Lalit Kumar Kakoti, Dr. Aditi Baruah, Mrs. Shyamoli Dutta for their exemplary guidance, monitoring and constant encouragement throughout the course of this report. The preparation of this report would not have been possible without their valuable contribution in the form of blessings, help and guidance. I also express my thanks to department of statistics, Bahona College. I acknowledge my gratitude to various primary source which helped in collection of data about the topic.

Last but not the least I thank the almighty, my parents, friends and other fellow for their constant encouragement without which the report would not have seen the dawn.

INDEX

INTRODUCTION TO TEST OF SIGNIFICANCE:

In many situations, statisticians are called upon to make decisions about a statistical population on the basis of sample observations. In 1925, Ronald Fishers advanced the idea of statistical hypothesis testing, which he named as "Test of Significance".

In attempting to reach such decisions, it is necessary to make certain assumptions about the parameters or probability distributions. In other words we can simply say that it is a technique by means of which we may either accept or reject a null hypothesis. The acceptance does not establish the truth of the hypothesis rather it simply means that the test has no evidence against the null hypothesis.

OBJECTIVES:

- Define statistical inference.
- Steps in Testing for Statistical Significance
- Stating Hypothesis
- Errors in sampling
- Level of Significance and Critical Region
- One-tailed and Two-tailed test
- Critical Values or Significant Values
- Tests Of Significance for Large Sample
- Testing of Significance for Single Proportion.
- Testing of Significance for Difference of Proportions.
- Testing of Significance for Single Mean.
- Testing of Significance for Difference of Mean.
- Testing of Significance for the Difference of Standard deviation.
- T-Test
- F-Test
- Chi-Square Test

Statistical Inference

Confidence intervals are one of the two most common types of statistical inference. Researchers use a confidence interval when their goal is to estimate a population parameter. The second common type of inference, called a **test of significance**, has a different goal: to assess the evidence provided by data about some claim concerning a population.

*A **test of significance** is a formal procedure for comparing observed data with a claim (also called a hypothesis), the truth of which is being assessed.*

•The claim is a statement about a parameter, like the population proportion p or the population mean μ.

•The results of a significance test are expressed in terms of a probability that measures how well the data and the claim agree.

STEPS IN TEST OF SIGNIFICANCE:

I. Set up the null hypothesis H_0 and alternative hypothesis H_1.
II. Choose the appropriate level of significance (α).
III. Choose the test statistic and compute it under the null hypothesis.
IV. Find the critical region
V. Compare the computed value with the critical value. If modules of calculated values is smaller than the table value, null hypothesis is insignificant i.e., there is no evidence against H_0 and we should not reject H_0 at α% level of significance. Otherwise, we should reject H_0 at α% level.

Stating Hypothesis

The first step in conducting a test of statistical significance is to state the hypothesis.

A significance test starts with a careful statement of the claims being compared. The claim tested by a statistical test is called the null hypothesis (H0). The test is designed to assess the strength of the evidence against the null hypothesis. Often the null hypothesis is a statement of "no difference."
The claim about the population that evidence is being sought for is the alternative hypothesis (Ha).
The alternative is one-sided if it states that a parameter is larger or smaller than the null hypothesis value.
It is two-sided if it states that the parameter is different from the null value (it could be either smaller or larger).

- When using logical reasoning, it is much easier to demonstrate that a statement is false, than to demonstrate that it is true. This is because proving something false only requires one counterexample. Proving something true, however, requires proving the statement is true in every possible situation.
- For this reason, when conducting a test of significance, a **null hypothesis** is used. The term **null** is used because this hypothesis assumes that there is no difference between the two means or that the recorded difference is not significant. The notation that is typically used for the null hypothesis is H_0.
- The opposite of a null hypothesis is called the **alternative hypothesis**. The alternative hypothesis is the claim that researchers are actually trying to prove is true. However, they prove it is true by proving that the null hypothesis is false. If the null hypothesis is false, then its opposite, the alternative hypothesis, must be true. The notation that is typically used for the alternative hypothesis is H_1

XAMPLE

ull Hypothesis

o: IQ smple mean=100,Or IQ ample mean -100 = 0

n words:

o: The sample mean IQ is not ingnificantly different than 100.

Alternative Hypothesis

One-Sided H_1: IQ sample mean>100,Or H_1: IQ sample mean – 100 > 0 In word:

One-sided H_1: The sample mean IQ is significantly larger than 100.

[Two-sided H_1: IQ>≠100 or IQ-100≠0]

- In the example above, the null hypotheses states: "the sample mean is equal to 100" or "there is no difference between the sample mean and the population mean."

➢ The sample mean will not be exactly equal to the population mean. This null hypothesis is stating that the recorded difference is not a significant one.

➢ If researchers can demonstrate that this null hypothesis is false, then its opposite, the alternative hypothesis, must be true.

- In the example above, the alternative hypothesis states: "the sample mean *is* significantly different than 100" or "there *is* a significant difference between the sample mean and the population mean."
- If researchers are trying to prove that the mean IQ in the sample will specifically be *higher* or *lower* (just one direction) than the population mean, this is a **one-sided** alternative hypothesis because they are only looking at one direction in which the mean may vary. They are not interested in the other direction.
- If researchers suspect that the sample mean could be either lower *or* higher than 100, the alternative hypothesis would be **two-sided** because both directions in which mean IQ may vary are being tested.

❖ When conducting a significance test, the goal is to provide evidence to reject the null hypothesis. If the evidence is strong enough to reject the null hypothesis, then the alternative hypothesis can automatically be accepted. However, if the evidence is not strong enough, researchers fail to reject the null hypothesis.

Test-Statistic (z score)

Example:

Population mean: IQ=100

Population st dev=16

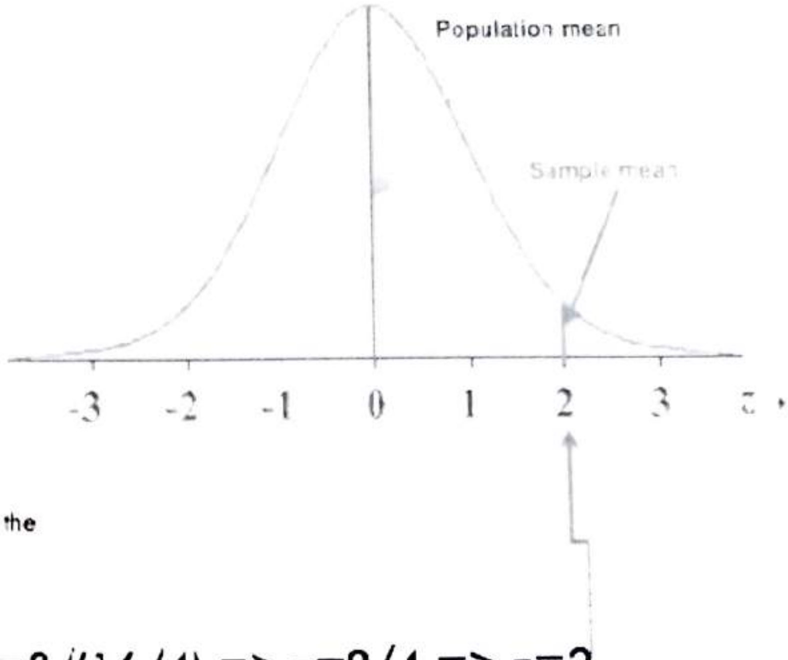
Sample mean: IQ=108

Sample size: N=16

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$

Test Statistic

z=(Sample Mean-Population Mean)/(st deviation in the population/√Sample size)



$$z=(108-100)/(16/\sqrt{16}) \Rightarrow z=8/(16/4) \Rightarrow z=8/4 \Rightarrow z=2$$

- In the above example, the mean IQ score for the sample is 108. This is slightly higher than the population mean, which is 100. The sample mean is obviously different from the population mean, but tests of significance must be done to determine if the difference is statistically significant. The difference could possibly be attributed to chance or to sampling error.
- The first step is to compute the **test statistic**. The test statistic is simply the z score for the sample mean. The only difference is that the population standard deviation is divided by the square root of *N*, just like when a confidence interval is computed.
- To compute the test statistic, the population standard deviation must be known for the variable. The population standard deviation for IQ is 16.
- To compute the test statistic, the sample size must also be known. In this case, it is 16. (In a real research scenario, the sample size would be larger. Small sample sizes are being use in this example to make calculations simpler).

Error in Sampling:

The main objective in sampling theory is to draw valid inferences about the population parameters on the basis of the sample results. In practice we decide to accept or reject the null hypothesis is made on the basis of the information supplied by the observed the sample observation. The conclusion drawn on the basis of a particular sample may not always be true in respect of the population. The four possible situation that arise in any test procedure are given in the following table :

True State		Decision from sample	
		Reject H_0	Accept H_0
	H_0 True	Wrong(Type(i)Error)	Correct
	H_0 False(H_1 True)	Correct	Wrong(Type ii Error)

Thus in testing of hypothesis we are likely to commit two error-
The error of rejecting H_0 where H_0 is true is known as ***Type [I] error.***
The error of accepting H_0 where H_0 is false is known as ***Type[II] error.***

P[Type I error] = P[rejecting H_0 when it is true]
$= \alpha$

P[Type II error] = P[accepting H_0 when it is false]
$= \beta$

Then α and β are also called as sizes of type[I] and type [II] error respectively.

Level of significance:

The probability of type [I] error is known as the develop significance. In other words, the maximum size of the type[I] error which we are prepared to risk is known as the level of significance. It is usually denoted by α and is given by -

$$P[\text{Type I error}] = P[\text{rejecting } H_0 \text{ when it is true}] = \alpha$$

The most commonly used level of significance in practice are- 5%(0.05) and 1%(0.01). If we adopted 5% level of significance it implies that 5 samples out of 100 we are likely to reject a true H_0. In other word these implies that we are 95% confident that our decision to reject H_0 is correct. Level of significance is always fixed in advance before collecting the sample information.

Critical Region : A region corresponding to a statistic "t" in the sample space "S" which amount to rejected to H_0 it terms as critical region. If ω is the critical region and if $t = t(x_1,x_2,x_3,\ldots\ldots\ldots,x_n)$ is the value of the statistic based on a random sample of size "n" then,

$$P(t \in \omega / H_0) = \alpha$$

$$P(t \in \overline{\omega} / H_1) = \beta$$

Where $\overline{\omega}$ is called a acceptation region.

One-tailed and Two-Tailed test :

In any test, the critical region is represented by a portion of the area under the probability curve of the sampling distribution of the test statistic.

A test of any statistical hypothesis where the alternatives hypothesis is one tailed (right tailed or left tailed) is called a one tailed test. For example , a test for testing the mean of a population $H_0 : \mu = \mu_0$ against the alternative hypothesis :

$H_1 : \mu = \mu_0$(Right-tailed) or $H_1 : \mu = \mu_0$(Left-tailed), is a single tailed test.

In the right tailed test($H_1 : \mu > \mu_0$) the critical region lies entirely in the right tail of the sampling distribution of $\bar{X}$, while for the left tailed test ($H_1 : \mu < \mu_0$),the critical region is entirely in the left tail of the distribution.

A test of statistical hypothesis where the alternative hypothesis is two-tailed such as: $H_0 : \mu \neq \mu_0$, against the alternative hypothesis $H_1 : \mu \neq \mu_0$($\mu > \mu_0$ and $\mu < \mu_0$)is known as two-tail test and in such a case the critical region is given by the area lying in both tails of the probability curve of the test statistics.

In a particular problem, whether one tailed or two tailed test is to be applied depends entirely on the nature of the alternative hypothesis. If the alternative hypothesis is two-tailed we apply two-tailed test and if alternative hypothesis is one-tailed, we apply one tailed test.

Critical Values or Significant Values:

The value of test statistic which separates the critical (or rejection) region and the acceptance region is called the critical value or significant value. It depends upon:

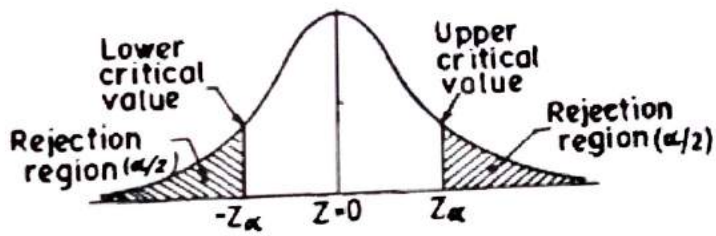
i. The level of significance used, and
ii. The alternative hypothesis, whether it is two-tailed or single-tailed

In the case of single tail alternatives, the critical value Z_α is determine so that the total area to right of it (for right tail test is α and for left-tailed test the total area to the left of $(-Z_\alpha)$ is α .

For Right-tail Test : $P(Z > Z_\alpha) = \alpha$

For Left-tail Test $P(Z < -Z_\alpha) = \alpha$

TWO-TAIL TEST
(Level of significance 'α')

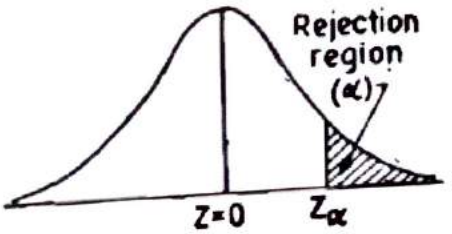


In case of single-tail alternative, the critical value z_α is determined so that total area to the right of it (for right-tailed test) is α and for left-tailed test the total area to the left of $-z_\alpha$ is α (See diagrams below), *i.e.*,

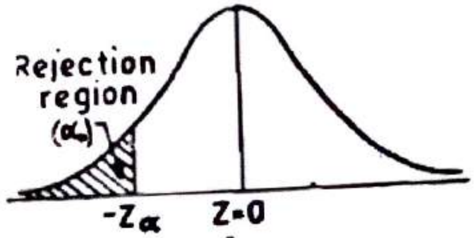
For Right-tail Test : $P(Z > z_\alpha) = \alpha$

For Left-tail Test : $P(Z < -z_\alpha) = \alpha$

RIGHT-TAILED TEST
(Level of Signifiance 'α')



LEFT-TAILED TEST
(Level of Significance 'α')



Thus *the significant* or *critical value of Z for a single-tailed test (left or right) at level of significance 'α' is same as the critical value of Z for a two-tailed test at level of significance '2α'.*

Critical Values (z_α)	CRITICAL VALUES (z_α) OF Z Level of significance (α) 1%	5%	10%
Two-tailed test	$\lvert Z_\alpha \rvert = 2.58$	$\lvert Z_\alpha \rvert = 1.96$	$\lvert Z_\alpha \rvert = 1.645$
Right-tailed test	$Z_\alpha = 2.33$	$Z_\alpha = 1.645$	$Z_\alpha = 1.28$
Left-tailed test	$Z_\alpha = -2.33$	$Z_\alpha = -1.645$	$Z_\alpha = -1.28$

Test of Significance for Large Samples:

In this section we will discuss the tests of significance when samples are large. We have seen that
for large values of *n,* the number of trials, almost all the distributions, *e.g,* binomial Poisson , negative binomial, etc ,are very closely approximated by normal distribution. Thus in this case we apply the *normal test,* which is based upon the following fundamental property *(area property)* of the normal
Probability curve.

If $X \sim N(\mu,\sigma^2)$ then $Z = \frac{X-\mu}{\sigma} = \frac{X-E(X)}{\sqrt{V(X)}} \sim N(0,1)$

Thus from the normal probability tables, we have-

$P(-3\leq Z\leq 3) = 0.9973$, i.e., $P(\lvert Z\rvert \leq 3) = 0.9973$

$\Rightarrow P(\lvert Z\rvert > 3) = 1 - P(\lvert Z\rvert \leq 3) = 0.0027$

i.e., in all probability we should expect a standard normal variate to lie between ± 3.

Also from the normal probability table, we get

$P(-1.96 \leq Z \leq 1.96) = 0.95$, i.e., $P(\lvert Z\rvert \leq 1.96) = 0.95$

$\Rightarrow P(\lvert Z\rvert > 1.96) = 1 - 0.95 = 0.05$

And $P(\lvert Z\rvert < 2.58) = 0.99$

$\Rightarrow P(\lvert Z\rvert > 2.58) = 0.01$

Thus the significant values of Z at 5% and 1% levels of significance for a two-tailed test are 1.96 and 2.58 respectively.

Thus the steps to be used in the normal test are as follows:

1. Compute the test statistic Z under H_0
2. If $|Z| > 3$, H_0 is always rejected
3. If $|Z| < 3$, we test its significance at certain level of significance, usually 5% and sometimes at 1% level of significance. Thus, for a two-tailed test it $|Z| > 1.96$, H_0 is rejected at 5% level of significance.

Similarly, if $|Z| > 2.58$, H_0 is contradicted at 1% level of significance and if $|Z| \leq 2.58$, H_0 may be accepted at 1% level of significance.
From the normal probability tables, we have :
$P(Z > 1.645) = 0.5 - P(0 \leq Z \leq 1.645) = 0.5-0.45=0.5-0.45=0.05$
$P(Z > 2.33) = 0.5 - P(0 \leq Z \leq 2.33) = 0.5-0.49=0.01$
Hence, for a single-tail test (Right tail or Left tail) we compare the computed values of $|Z|$ with 1.645 (at 5% level) and 2.33(at 1% level) and accept or reject H_0 accordingly.

Test of significance for single proportion:

If x is the number of of success i.e., the number of individuals or units possessing the given attribute in a random sample of size 'n' from a large population ,

E(X) =nP

V(X) =nPQ

P = Observed sample proportion of success

$=\frac{X}{n}$

$E(p) = E(\frac{X}{n})$

$= \frac{1}{n} E(\frac{X}{n})$

$= \frac{1}{n} \times nP$

Therefore, E(p) = P

$V(p) = V(\frac{X}{n})$

$= \frac{1}{n^2} V(X)$

$= \frac{1}{n^2} \times nPQ$

$= \frac{PQ}{n}$

(Standard Error) $SE(p) = \sqrt{V(p)}$

$= \sqrt{\frac{PQ}{n}}$

Hence for large sample the standard normal variate corresponding to the statistic p is,

$$Z = \frac{p-P}{\sqrt{\frac{PQ}{n}}} \sim N(0,1)$$

Test of significance for difference of proportions:

Suppose we want to compared two distinct population with respect to the prevalence of curtain attribute A among their members.Let,X_1 and X_2 be the number of person possessing the given attribute A in the random sample of size n_1 and $n_2$from the two population respectively. Then sample proportion are given by,

$$p_1 = \frac{X1}{n1} \text{ and } p_2 = \frac{X2}{n2}$$

If p_1 and p_2 are population proportion then,

$$E(p_1) = P_1 \quad ; \quad E(p_2) = P_2$$

And, $V(p_1) = \frac{P1Q1}{n1}$; $V(p_2) = \frac{P2Q2}{n2}$

Since for large sample p_1 and p_2 are independently normally distributed then p_1-p_2 also normally distributed,

$$Z = \frac{p1-p2-E(p1-p2)}{\sqrt{var}(p1-p2)} \sim n(0,1)$$

Under the null hypothesis H_0: $P_1 = P_2$ i.e., there is no significant difference between the sample proportion,

$$E(p_1-p_2) = E(p_1) = E(P_2)$$

$$= P_1 - P_2 = 0$$

$$V(p_1-p_2) = V(p_1) = V(p_1) + V(p_2)$$

$$= \frac{P1Q1}{n1} + \frac{P2Q2}{n2}$$

Since under H_0 : $P_1 = P_2 = P$

Therefore $Q_1 = Q_2 = Q$

And $V(p_1-p_2) = PQ(\frac{1}{n1} + \frac{1}{n2})$

So, $Z = \frac{p1-p2}{\sqrt{PQ}(\frac{1}{n1}+\frac{1}{n2})}$

Test of significance for single mean:

If $\{x_1, x_2, x_3 \ldots\ldots x_n\}$ is a random sample from a normal population with mean μ and variance σ^2 then the sample mean is distributed normally with mean μ and and $\frac{\sigma^2}{n}$ i.e., $X \sim N(\mu, \sigma^2)$.Thus for large sample the standard normal variate corresponding to x is

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

Under the null hypothesis H_0 then the sample has been drawn from a population with mean μ and variance σ^2 i.e, there is no significance difference between the sample mean x and population mean μ. The test statistic is,

$$Z = \frac{x-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

By comparing this calculated value of z with the table vale 1.96 and 2.58(for two tailed test), we can reject or accept the null hypothesis at 5% level of significance. The corresponding significance value for single (or one tail) test are 1.645 and 2.33 at 5% and 1% level of significance respectively.

Test of significance for difference of mean:

Let, x_1 be the mean of a sample of size n_1 from a population with μ_1 and variance σ_1^2 and the x_2 be the mean of n independent random sample of size n_2 from another population with mean μ_2 and variance σ_1^2 $\overline{x1} \sim N(\mu_1, \sigma_1^2/n_1)$ and $\overline{x2} \sim N(\mu_2, \sigma_2^2/n_2)$[sample are large]

So, their difference i.e $\overline{x1}$- $\overline{x2}$ is also a normal variant. The value of Z is,

$$Z = \frac{\overline{x1} - \overline{x2} - \mathrm{E}(\overline{x1} - \overline{x2})}{S.E(\overline{x1} - \overline{x2})} \sim N(0,1)$$

Under H_0: $\mu_1 = \mu_2$ i.e there is no significance difference below the sample mean we get,

$E(\overline{x1} - \overline{x2}) = E(\overline{x1}) - E(\overline{x2})$ [$E(\overline{x1}) = \mu_1$]

$= \mu_1 - \mu_2$

$var(\overline{x1} - \overline{x2}) = V(\overline{x1}) + V(\overline{x2})$

$= \frac{\sigma 1^2}{n1} + \frac{\sigma 1^2}{n2}$

$S.E((\overline{x1} - \overline{x2}) = \sqrt{\frac{\sigma 1^2}{n1} + \frac{\sigma 2^2}{n2}}$

Under, H_0: $\mu_1 = \mu_2$ i.e $\mu_1 - \mu_2 = 0$

The test statistic becomes,

$$Z = \frac{\overline{x1} - \overline{x2}}{\sqrt{\frac{\sigma 1^2}{n1} + \frac{\sigma 2^2}{n2}}} \sim N(0,1)$$

If $\sigma_1^2 = \sigma_1^2 = \sigma^2$ i.e if the sample have been drawn from the population with the common standard deviation σ than under H_0: $\mu_1 = \mu$

$$Z = \frac{\overline{x1} - \overline{x2}}{\sqrt{\sigma^2 (\frac{1}{n1} + \frac{1}{n2})}} = \frac{\overline{x1} - \overline{x2}}{\sigma \sqrt{(\frac{1}{n1} + \frac{1}{n2})}}$$

Test of significance for difference of standard deviations:

Let, S_1 and S_2 are the standard deviation of the two independent random sample of sizes n_1 and n_2 from two population with SD σ_1 and σ_2 respectively. The problem is to test $\sigma_1 = \sigma_2$ i.e if the sample SD's differ significantly or not.

We know that in sampling from large normal population,sampling distribution of S is normal with

var(s)=$\sigma^2/2n$

S.E(s)=$\sqrt{\sigma^2/2n} = \sigma/\sqrt{2n}$

Provided n is large

Therefore var(s_1-s_2)=var(s_1)+var(s_2)

$=\sigma^2/2n_1+\sigma^2/2n_2$

The covariance term vanished because samples are independent.

Hence, under null hypothesis , H_0:the sample SD does not differ significantly, the test statistic is given by

$$Z=\frac{S1-S2}{\sqrt{\sigma^2/2n1+\sigma^2/2n2}} \sim N(0,1)$$

σ_1^2 and σ_2^2 are usually unknown in that case we use their estimates. Hence test statistic is reduce to ,

$$Z= \frac{S_1-S_2}{S_1^2/2n_1+S_2^2/2n_2} \quad [\sigma_1 \sim s_1, \sigma_2 \sim s_2] \text{ for large sample.}$$

Some more types of test of significance

T-Test

A t-test is a type of inferential statistic used to determine if there is a significant difference between the means of two groups, which may be related in certain features. It is mostly used when the data sets, like the data set recorded as the outcome from flipping a coin 100 times, would follow a normal distribution and may have unknown variances. A t-test is used as a hypothesis testing tool, which allows testing of an assumption applicable to a population. A t-test looks at the t-statistic, the t-distribution values, and the degrees of freedom to determine the statistical significance. To conduct a test with three or more means, one must use an analysis of variance.

A t-test allows us to compare the average values of the two data sets and determine if they came from the same population.

Mathematically, the t-test takes a sample from each of the two sets and establishes the problem statement by assuming a null hypothesis that the two means are equal. Based on the applicable formulas, certain values are calculated and compared against the standard values, and the assumed null hypothesis is accepted or rejected accordingly. If the null hypothesis qualifies to be rejected, it indicates that data readings are strong and are probably not due to chance. The t-test is just one of many tests used for this purpose. Statisticians must additionally use tests other than the t-test to examine more variables and tests with larger sample sizes. For a large sample size, statisticians use **a z-test.** Other testing options include the chi-square test and the f-test.

F-test

An F-test is any statistical test in which the test statistic has an F-distribution under the null hypothesis. It is most often used when comparing statistical models that have been fitted to a data set, in order to identify the model that best fits the population from which the data were sampled. Exact "F-tests" mainly arise when the models have been fitted to the data using least squares. The name was coined by George W. Snedecor, in honour of Sir Ronald A. Fisher. Fisher initially developed the statistic as the variance ratio in the 1920s.

The F- test determine whether or not the two independent estimates of the population variances are homogeneous in nature. The researcher uses the F-test to test the Significance of an observed multiple correlation coefficient. It is also Used to test the significance of an observed sample correlation ratio .The F-test have some association with T-test. According to this association, if a statistic t follows a student's t distribution with 'n' degrees of freedom, then the square of this statistic will follow Snedecor's F distribution with 1
And n degrees of freedom. The F- test also has some other associations, like the association between chi square
distribution.

Due to such relationships, the F-test has many properties, like chi square. The F-values are all non negative. The F distribution in the F-Test is always non-symmetrically distributed. The mean in F-distribution in the F-test is Approximately one. There are two independent degrees of freedom in F distribution, one in the numerator and the other
In the denominator. There are many different F distributions in the F-test, one for every pair of degree of freedom.

TheF-test is a parametric test that helps the researcher draw out an inference about the data that is drawn from a particular population. The F-test is called a parametric test because of the presence of parameters in the F-test. These parameters in the F-test are the mean and variance. The mode of the F-test is the value that is most frequently in a data set and it is always less than unity. According to Karl Pearson's coefficient of skewness, the F-test is highly positively skewed. The probability distribution of F increases steadily before reaching the peak, and then it starts decreasing in order to become tangential at infinity. Thus, we can say that the axis of F is asymptote to the right.

Z-Test:

A z-test is a statistical test used to determine whether two population means are different when the variances are known and the sample size is large.

The test statistic is assumed to have a normal distribution, and nuisance parameters such as standard deviation should be known in order for an accurate z-test to be performed.

The z-test is also a hypothesis test in which the z-statistic follows a normal distribution. The z-test is best used for greater-than-30 samples because, under the central limit theorem, as the number of samples gets larger, the samples are considered to be approximately normally distributed.

When conducting a z-test, the null and alternative hypotheses, alpha and z-score should be stated. Next, the test statistic should be calculated, and the results and conclusion stated. A z-statistic, or z-score, is a number representing how many standard deviations above or below the mean population a score derived from a z-test is.

Z-tests are closely related to t-tests, but t-tests are best performed when an experiment has a small sample size. Also, t-tests assume the standard deviation is unknown, while z-tests assume it is known. If the standard deviation of the population is unknown, the assumption of the sample variance equaling the population variance is made.

Chi-Square Test:

chi-square ($\chi 2$) statistic is a test that measures how a model compares to actual
observed data. The data used in calculating a chi-square statistic must be random, raw, mutually exclusive, drawn
from independent variables, and drawn from a large enough sample. For example, the results of tossing a fair coin
meet these criteria.

Chi-square tests are often used in hypothesis testing. The chi-square statistic compares the size any discrepancies
Between the expected results and the actual results, given the size of the sample and the number of variables in the
relationship. For these tests, degrees of freedom are utilized to determine if a certain null hypothesis can be rejected
based on the total number of variables and samples within the experiment. As with any statistic, the larger the sample size, the more reliable the results.

There are two main kinds of chi-square tests :the test of independence, which asks question of relationship, such as,
"Is there a relationship between student sex and course choice?";
and the goodness-of-fit test, which asks something
Like "How well does the coin in my hand match a theoretically fair coin?"

Independence

when considering students expand course choice, a $\chi 2$ test for independence could be used. To do this test, the
Researcher would collect data on the two chosen variables (sex and courses picked) and then compare the frequencies at
Which male and female students select among the offered classes using the formula given above and a $\chi 2$ statistical table. If there is no relationship between sex and courses selection (that is, if they are independent),the n the actual frequencies at which male and female

students select each offered course should be expected to be approximately equal ,or conversely, the proportion of male and female students in any selected course should be approximately equal to the proportion of male and female students in the sample. A χ2 test for independence can tell us how likely it is that random chance can explain any observed difference between the actual frequencies in the data and these theoretical expectations .

Goodness of fit

χ2 provides away to test how well a sample of data matches the (known or assumed) characteristics of the larger population that the sample is intended to represent. If the sample data donot fit the expected properties of the population that we are interested in, then we would not want to use this sample to draw conclusions about the larger population. For example consider an imaginary coin with exactly 50/50 chance of landing heads or tails and a real coin that you toss 100 times .If this real coin has an is fair ,then it will also have an equal probability of landing on either side, and The expected result of tossing the coin 100 times is that heads will come up 50 times and tails will come up 50 times. In this case ,χ2 can tell us how well the actual results of 100 coin flips compare to the theoretical model that a fair Coin will give 50/50 results. The actual toss could come up 50/50, or 60/40,or even 90/10. The farther away the actual results of the 100 tosses is from 50/50, the less good the fit of this set of tosses is to the theoretical expectation of 50/50 and the more likely we might conclude that this coin is not actually a fair coin.

Conclusion

Once sample data has been gathered through an observational study or experiment statistical significance allow analyst to assess evidence in favor or some claim about the population from which the sample has been drawn the method of test of significance use to support or reject claims based on sample data.

References:

- A First Course in STATISTICS-II, By Manoj K. Bhowal & Pranab Baruah.
- Fundamentals of Mathematical Statistics, By S.C Gupta & V.K Kapoor.
- 502 College Notes Copy.
- Wikipedia.