UNIT-4

Statistical Inference

Dr. B. Mallikarjuna
Assistant Professor
Department of Mathematics
B.M.S.College of Engineering

Topics

- Introduction
- Procedure for testing of hypothesis
- Level of significance
- Test of significance for single mean (Large Sample)
- Test of significance for difference between two means (Large Sample)
- Test of significance for single proportions (Large Sample)
- Test of significance for difference between two proportions (Large Sample)
- Test of significance for single mean (Small Sample)-paired t-test
- Test of significance for difference between two means (Small Sample) – paired t-test
- Ratio of Variances – F-Distribution
- Chi-Square distribution-goodness of fit

Introduction

What do you think about this piece?

Introduction

- Population
- Sample
- Sampling
- Random Sampling
- Sampling Distribution
- Sampling Distributions of means
  - Sampling with replacement
  - Sampling without replacement

Introduction

- Population

  is a complete set of individuals or things or item or numerics or species or bacteria or etc.

- Sample

  Small part of the population is called sample

- Sampling

  Process of drawing samples is called sampling

- Random Sampling

  Process of drawing samples in which every member has same chance of occurrence is called Random Sampling

  Suppose a sample of size n from a finite population of size N, then possible number of samples = ?

- Sampling where a member of the population may be chosen more than once is called as ___________
- Number of samples=?
- If a member can't be chosen more than once is called ___________
- Number of samples=?

Introduction

- Population:

  Characteristics of population are called parameters:

  Example: Population mean, Population Variance, Population Standard Deviation, median and etc.

  Notations: Population mean = μ

  Population Standard Deviation = σ

- Samples:

  Characteristics of samples are called Statistics:

  Example: sample mean, sample Variance, sample Standard Deviation, median and etc.

  Notations: sample mean = $\overline{X}$

  sample standard deviation = $\sigma_{\bar{X}}$

Introduction

- Sampling Distributions
  - Random Variable: mean of samples
  - Choosing samples of size n out of N no. of populations
  - Find the statistics like mean, variance etc.
  - Find frequencies of the mean
- Generated frequency distribution is called as sampling distributions
- Example: Population={1,2,3}.

  Samples of size 2 with replacement and without replacement

  Write all possible samples=?

  Find mean and S.D of all samples and population

  Find frequency of these means

Sampling with replacement

- Example: Population={1,2,3}.

  Samples of size 2 with replacement and without replacement

  Write all possible samples=?

  Find mean and S.D of all samples and population

  Find frequency of these means

Sampling without replacement

- Example: Population={1,2,3}.

  Samples of size 2 with replacement and without replacement

  Write all possible samples=?

  Find mean and S.D of all samples and population

  Find frequency of these means

- Standard Error

The S.D of a sampling distribution is called the ***standard error***

- (N-n)/(N-1) is called the finite population correction factor
- The sampling distribution of large samples is assumed to be normal distributions

If N is very large, then (N-n)/(N-1) becomes 1

Note: If $n < 30$, then sample is called small sample otherwise it is called large sample

Statistical Inference

- ***Statistical inference*** is the process of reaching conclusions about characteristics of an entire population using data from a subset, or *sample*, of that population.

Statistical Inference

The process of making guesses about the truth about a population parameter from a sample statistic.

Sample statistics

$$\hat{\mu} = \overline{X}_n = \frac{\sum_{i=1}^{n} x}{n}$$

$$\hat{\sigma}^2 = s^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{X}_n)^2}{n-1}$$

Truth (not observable)

Population parameters

$$\mu = \frac{\sum_{i=1}^{N} x}{N}$$

$$\sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}$$

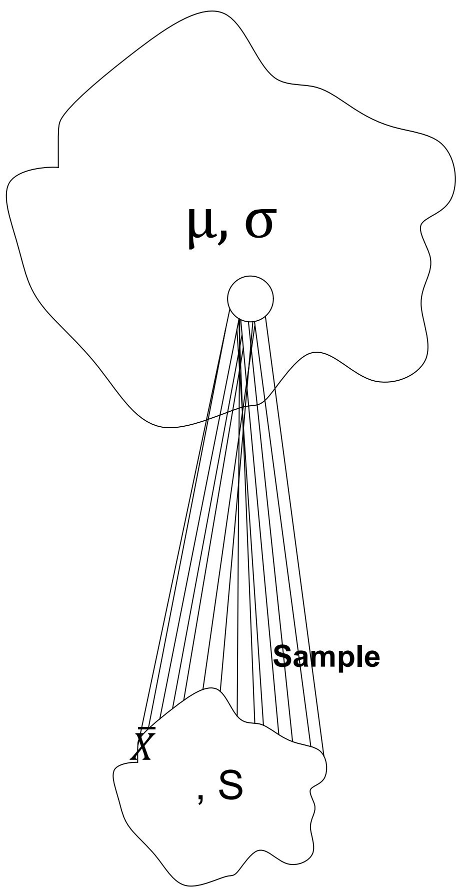
Sample (observation)

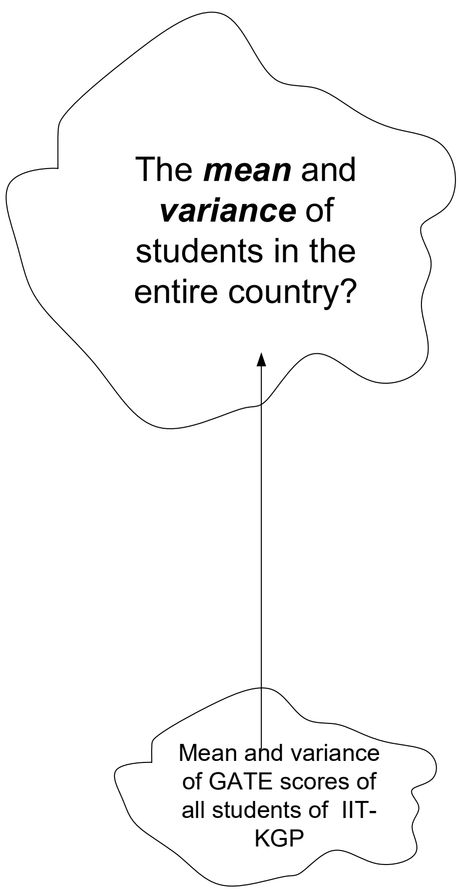
*hat notation ^ is often used to indicate "estimate"

Make guesses about the whole population

Objective

The primary objective of statistical analysis is to use data from a sample to make inferences about the population from which the sample was drawn.





This lecture aims to learn the basic procedures for making such inferences.

Problem

Determine the mean and s.d of the sampling distribution of means of 300 random samples each of size n=36 are drawn from a population of N=1500 which is normally distributed with mean μ=22.4 and s.d σ of 0.048, if sampling is done (a) with replacement (b) without replacement.

Determine the expected number of random samples having their means (a) between 22.39 and 22.41 (b) greater than 22.42 c). Less than 22.37 d) less than 22.38 o rmore than 22.41.

Two different types of procedures

- Estimating population parameters
  - Point estimation
    - Using a sample statistic to estimate a population parameter
  - Interval estimation
    - Estimation of the amount of variability in a sample statistic when many samples are repeatedly taken from a population.
- Hypothesis testing
  - The comparison of sample results with a known or hypothesized population parameter

Basic Approaches

Approach 1: Hypothesis testing

- We conduct **test on hypothesis**.
  - We hypothesize that one (or more) parameter(s) has (have) some specific value(s) or relationship.
- Make our decision about the parameter(s) based on one (or more) sample statistic(s)
- Accuracy of the decision is expressed as the probability that the decision is incorrect.

Approach 2: Confidence interval measurement

- We estimate one (or more) parameter(s) using sample statistics.
  - This estimation usually done in the form of an interval.
- Accuracy of the decision is expressed as the **level of confidence** we have in the interval.

Hypothesis Testing

Statistical inference

Sample

Null hypothesis

Alternative hypothesis

Hypothesis Testing

What is Hypothesis?

- "A hypothesis is an educated prediction that can be tested" (study.com).
- "A hypothesis is a proposed explanation for a phenomenon" (Wikipedia).
- "A hypothesis is used to define the relationship between two variables" (Oxford dictionary).
- "A supposition or proposed explanation made on the basis of limited evidence as a starting point for further investigation" (Walpole).

- **Example: Avogadro's Hypothesis**

"The volume of a gas is directly proportional to the number of molecules of the gas."

$$V = aN$$

Statistical Hypothesis

- If the hypothesis is stated in terms of population parameters (such as mean and variance), the hypothesis is called statistical hypothesis.
- Data from a sample (which may be an experiment) are used to test the validity of the hypothesis.
- A procedure that enables us to agree (or disagree) with the statistical hypothesis is called a **TEST OF THE HYPOTHESIS.**

Example:

1. To determine whether the wages of men and women are equal.
2. A product in the market is of standard quality.
3. Whether a particular medicine is effective to cure a disease.

The Hypotheses

- The main purpose of statistical hypothesis testing is to choose between two competing hypotheses.

Example:

One hypothesis might claim that wages of men and women are equal, while the **alternative** might claim that men make more than women.

- Hypothesis testing start by making a set of two statements about the parameter(s) in question.
- The hypothesis actually to be tested is usually given the symbol H_0 and is commonly referred as the **null hypothesis.**
- The other hypothesis, which is assumed to be true when null hypothesis is false, is referred as the **alternate hypothesis** and is often symbolized by H_1
- The two hypotheses are exclusive and exhaustive.

The Hypotheses

Example:

Ministry of Human Resource Development (MHRD), Government of India takes an initiative to improve the country's human resources and hence set up **23 IIT's** in the country.

To measure the engineering aptitudes of graduates, MHRD conducts GATE examination for a mark of 1000 in every year. A sample of 300 students who gave GATE examination in 2018 were collected and the mean is observed as 220.

In this context, statistical hypothesis testing is to determine the mean mark of the all GATE-2018 examinee.

The two hypotheses in this context are:

$$H_0: \mu = 220$$

$$H_1: \mu < 220$$

The Hypotheses

Note:

1. As null hypothesis, we could choose $H_0: \mu \leq 220$ or $H_0: \mu \geq 220$
2. It is customary to always have the null hypothesis with an equal sign.
3. As an alternative hypothesis there are many options available with us.

Examples:

I. $H_1: \mu > 220$

II. $H_1: \mu < 220$

III. $H_1: \mu \neq 220$

4. The two hypothesis should be chosen in such a way that they are **exclusive** and **exhaustive**.
   - One or other must be true, but they cannot both be true.

The Hypotheses

One-tailed test

- A statistical test in which the alternative hypothesis specifies that the population parameter lies entirely above or below the value specified in H_0 is called a one-sided (or one-tailed) test.

Example.

$$H_0: \mu = 100 \qquad H_1: \mu > 100$$

Two-tailed test

- An alternative hypothesis that specifies that the parameter can lie on their sides of the value specified by H_0 is called a two-sided (or two-tailed) test.

Example.

$$H_0: \mu = 100 \qquad H_1: \mu <> 100$$

The Hypotheses

Note:

In fact, a 1-tailed test such as:

$$H_0: \mu = 100$$

$$H_1: \mu > 100$$

is same as

$$H_0: \mu \leq 100$$

$$H_1: \mu > 100$$

In essence, $\mu > 100$, it does not imply that $\mu > 80, \mu > 90$, etc.

Hypothesis Testing Procedures

The following five steps are followed when testing hypothesis

1. Specify H_0 and H_1, the null and alternate hypothesis, and an **acceptable level of α.**
2. Determine an appropriate sample-based test statistics and the **rejection region** for the specified H_0.
3. Collect the sample data and calculate the test statistics.
4. Make a decision to either reject or fail to reject H_0.
5. Interpret the result in common language suitable for practitioners.

Hypothesis Testing Procedure

- In summary, we have to choose between H_0 and H_1
- The standard procedure is to assume H_0 is true.
  (Just we presume innocent until proven guilty)
- Using statistical test, we try to determine whether there is sufficient evidence to declare H_0 false.
- We reject H_0 only when the **chance is small** that H_0 is true.
- The procedure is based on probability theory, that is, there is a chance that we can make errors.

Errors in Hypothesis Testing

In hypothesis testing, there are two types of errors.

Type I error: A type I error occurs when we incorrectly reject H_0 (i.e., we reject the null hypothesis, when H_0 is true).

Type II error: A type II error occurs when we incorrectly fail to reject H_0 (i.e., we accept H_0 when it is not true).

	Observation	
Decision	H_0 is true	H_0 is false
H_0 is accepted	Decision is correct	Type II error
H_0 is rejected	Type I error	Decision is correct

Probabilities of Making Errors

Type I error calculation

$\boldsymbol{\alpha}$**:** denotes the probability of making a Type I error

$$\boldsymbol{\alpha} = \mathbf{P}(\text{Rejecting } H_0 | H_0 \text{ is true})$$

Type II error calculation

$\boldsymbol{\beta}$**:** denotes the probability of making a Type II error

$$\boldsymbol{\beta} = \mathbf{P}(\text{Accepting } H_0 | H_0 \text{ is false})$$

Note:

- $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are not independent of each other as one increases, the other decreases
- When the sample size increases, both to decrease since sampling error is reduced.
- In general, we focus on Type I error, but Type II error is also important, particularly when sample size is small.

Calculating α

Assuming that we have the results of random sample. Hence, we use the characteristics of sampling distribution to calculate the probabilities of making either Type I or Type II error.

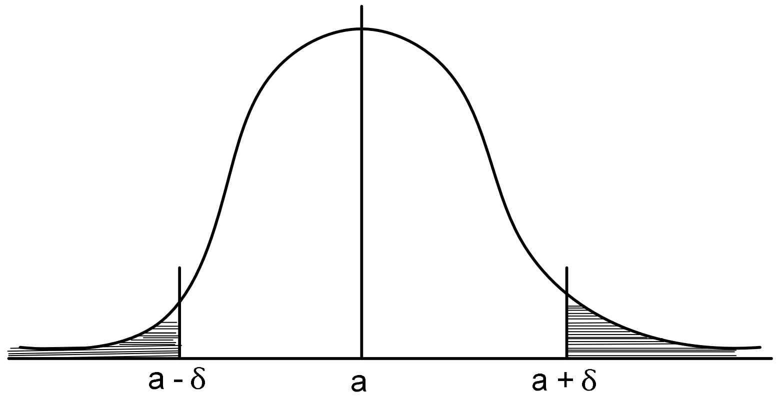
Example:

Suppose, two hypotheses in a statistical testing are:

$$H_0: \mu = a$$
$$H_1: \mu \neq a$$

Also, assume that for a given sample, population obeys normal distribution. A threshold limit say $a \pm \delta$ is used to say that they are significantly different from a.

Calculating α



Here, shaded region implies the probability that, $\overline{\mathrm{X}} < a - \delta$ *or* $\overline{\mathrm{X}} > a + \delta$

Thus the null hypothesis is to be rejected if the mean value is less than $a - \delta$ or greater than $a + \delta$.

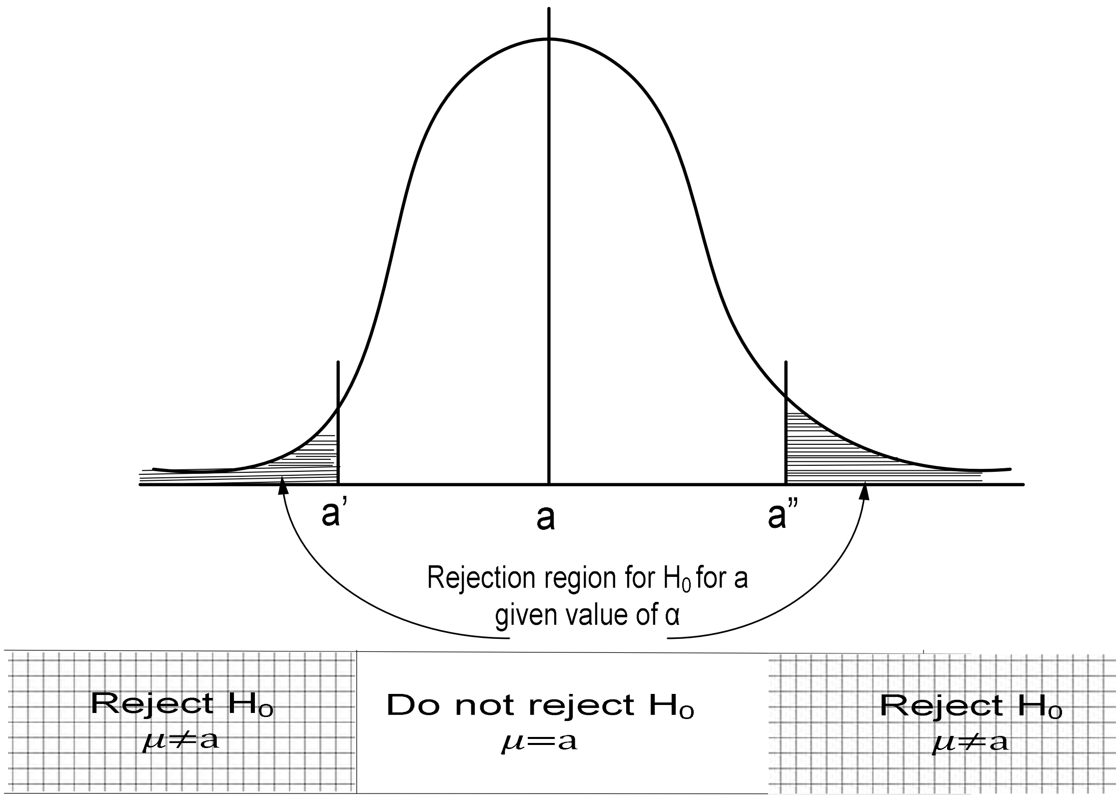
If $\overline{\mathrm{X}}$ denotes the sample mean, then the Type I error is

$$\alpha = P(\overline{\mathrm{X}} < a - \delta \ or \ \overline{\mathrm{X}} > a + \delta, \qquad when \ \mu = a \ , \qquad \text{i.e.}, \ H_0 \text{ is true})$$

The Rejection Region

The rejection region comprises of value of the test statistics for which

1. The probability when the null hypothesis is true is less than or equal to the specified α.
2. Probability when H_1 is true are greater than they are under H_0.



Two-Tailed Test

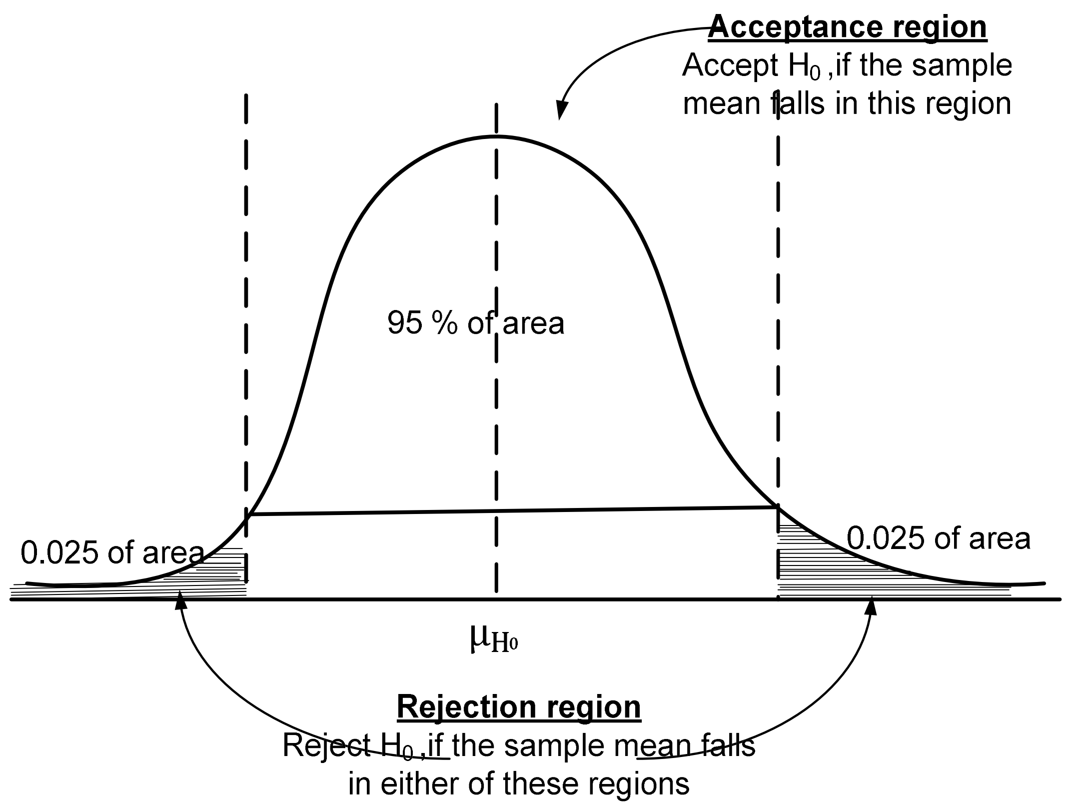
For two-tailed hypothesis test, hypotheses take the form

$$H_0: \mu = \mu_{H_0}$$
$$H_1: \mu \neq \mu_{H_0}$$

In other words, to reject a null hypothesis, sample mean $\mu > \mu_{H_0}$ or $\mu < \mu_{H_0}$ under a given α.

Thus, in a two-tailed test, there are two rejection regions (also known as critical region), one on each tail of the sampling distribution curve.

Two-Tailed Test



Acceptance and rejection regions in case of a two-tailed test with 5% significance level.

One-Tailed Test

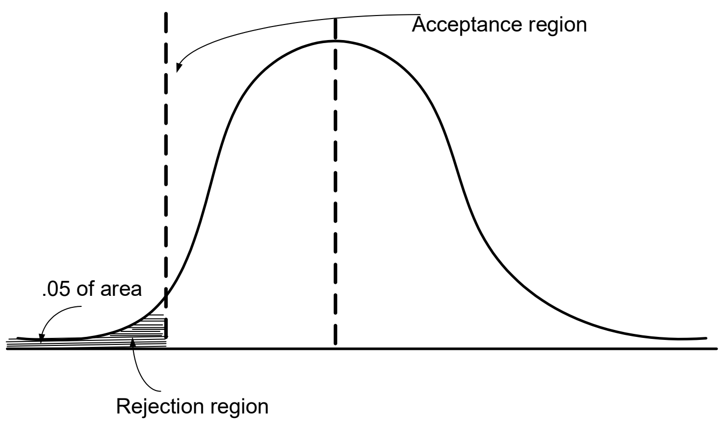
A one-tailed test would be used when we are to test, say, whether the population mean is either lower or higher than the hypothesis test value.

Symbolically,

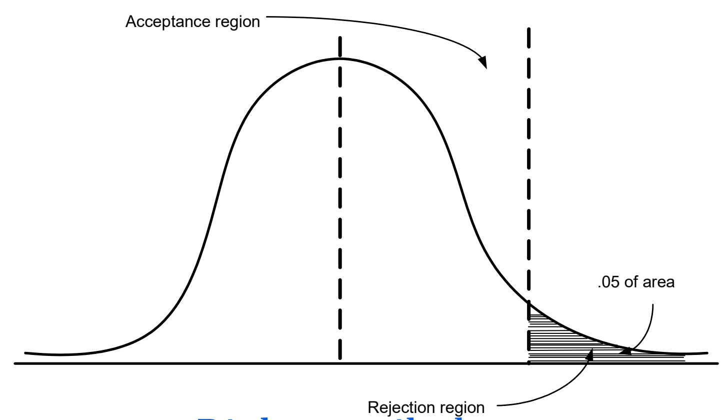
$$H_0: \mu = \mu_{H_0}$$

$$H_1: \mu < \mu_{H_0} \qquad [or\ \mu > \mu_{H_0}]$$

Wherein there is one rejection region only on the left-tail (or right-tail).



Left – tailed test



Right – tailed test

Example: Calculating α

Consider the two hypotheses are

The null hypothesis is

$$H_0: \mu = 8$$

The alternative hypothesis is

$$H_1: \mu \neq 8$$

Assume that given a sample of size 16 and standard deviation is 0.2 and sample follows normal distribution.

Example: Calculating α

We can decide the rejection region as follows.

Suppose, the null hypothesis is to be rejected if the mean value is less than 7.9 or greater than 8.1. If $\overline{\mathrm{X}}$ is the sample mean, then the probability of Type I error is

$$\alpha = P(\overline{\mathrm{X}} < 7.9 \; or \; \overline{\mathrm{X}} > 8.1, \text{when } \mu = 8)$$

Given σ, the standard deviation of the sample is 0.2 and that the distribution follows normal distribution.

Thus,

$$P(\overline{\mathrm{X}} < 7.9) = P\left[Z = \frac{7.9 - 8}{0.2/\sqrt{16}}\right] = P[Z < -2.0] = 0.0228$$

and

$$P(\overline{\mathrm{X}} > 8.1) = P\left[Z = \frac{8.1 - 8}{0.2/\sqrt{16}}\right] = P[Z > 2.0] = 0.0228$$

Hence, $\alpha = 0.0228 + 0.0228 = 0.0456$

Level of Significance (α)

- L.O.S. of a test denoted by α is the probability of committing type I error.
- Thus L.O.S. measures the amount of risk or error associated in taking decisions.
- It is customary to fix α before sample information is collected
- We choose generally α as 0.05 or 0.01.
- L.O.S. α =0.01 is used for high precision and
- L.O.S. α =0.05 is used for moderate precision
- L.O.S. can also be expressed as percentage
- Thus L.O.S. α=5% means there are 5 chances in 100 that N.H. is rejected when it is true or one is 95% confident that a right decision is made.
- L.O.S. is also known as size of test.

Critical Region (C.R.)

- C.R. is the region of rejection of N.H.
- The area of the critical region equals to the level of significance α.

Note: C.R is always depending on the nature of A.H.,

C.R. may lie on ons side or both sides of the tails

Hypothesis Testing : 5 Steps

The following five steps are followed when testing hypothesis

1. Specify H_0 and H_1, the null and alternate hypothesis, and an **acceptable level of α.**
2. Determine an appropriate sample-based test statistics and the **rejection region** for the specified H_0.
3. Collect the sample data and calculate the test statistics.
4. Make a decision to either reject or fail to reject H_0.
5. Interpret the result in common language suitable for practitioner.

Case Study: **Machine Testing**

A medicine production company packages medicine in a tube of 8 ml. In maintaining the control of the amount of medicine in tubes, they use a machine. To monitor this control a sample of 16 tubes is taken from the production line at random time interval and their contents are measured precisely. The mean amount of medicine in these 16 tubes will be used to test the hypothesis that the machine is indeed working properly.

Case Study 2: Step 1

Step 1: Specification of hypothesis and acceptable level of α

The hypotheses are given in terms of the population mean of medicine per tube.

The null hypothesis is

$$H_0: \mu = 8$$

The alternative hypothesis is

$$H_1: \mu \neq 8$$

We assume α, the significance level in our hypothesis testing $\approx$ 0.05.
(This signifies the probability that the machine needs to be adjusted less than 5%).

Case Study 2: **Step 2**

Step 2: Sample-based test statistics and the rejection region for specified H_0

Rejection region: Given $\alpha = 0.05$**,** which gives $|Z| > 1.96$ (obtained from standard normal calculation for $n(Z: 0{,}1) = 0.025$ for a rejection region with two-tailed test).

Case Study 2: Step 3

Step 3: Collect the sample data and calculate the test statistics

Sample results: $n = 16$, $\bar{x} = 7.89$, $\sigma = 0.2$

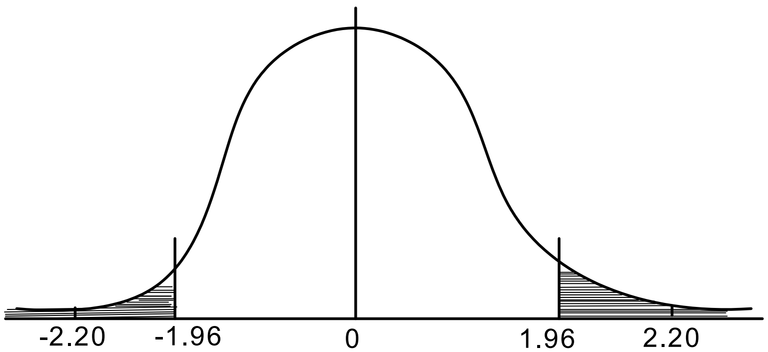
With the sample, the test statistics is

$$Z = \frac{7.89 - 8}{0.2/\sqrt{16}} = -2.20$$

Hence, $|Z| = 2.20$

Case Study 2: Step 4

Step 4: Make a decision to either reject or fail to reject H_0



Since $Z > 1.96$, we reject H_0

Case Study 2: **Step 5**

Step 5: Final comment and interpret the result

We conclude $\mu \neq 8$ and recommend that the machine be adjusted.

Case Study 2: **Alternative Test**

Suppose that in our initial setup of hypothesis test, if we choose $\alpha = 0.01$ instead of 0.05, then the test can be summarized as:

1. $H_0: \mu = 8$, $H_1: \mu \neq 8$ $\alpha = 0.01$
2. Reject H_0 if $Z > 2.576$
3. Sample result n =16, σ = 0.2, $\overline{X}$=7.89, Z $= \frac{7.89-8}{0.2/\sqrt{16}} = -2.20$, $|Z| = 2.20$
4. $|Z| < 2.20$, we fail to reject H_0= 8
5. We do not recommend that the machine be readjusted.

Problem

The length of life X of certain compupters is approximately normally distributed with mean 800 hours and S.D 40 hours. If a random sample of 30 computers has an average life of 788 hours, test the null hypothesis that $\mu=800$ hours at (a) 5% (b) 1% (c) 10% (d) 15% level of significance.

Problem

The length of life X of certain computers is approximately normally distributed with mean 800 hours and S.D 40 hours. If a random sample of 30 computers has an average life of 788 hours, test the null hypothesis that $\mu=800$ hours against alternative hypothesis that $\mu \neq 800$ at (a) 5% (b) 1% (c) 10% (d) 15% level of significance.

Solution:

Null Hypothesis: $\mu=800$

Alternate Hypothesis: $\mu \neq 800$

Level of Significance = $\alpha=5\% =0.05$

Critical region:

since alternate hypothesis is $\neq$ type, the test is two tailed and critical region is $-2.81 < Z < 2.81$ (Indicate in standard normal curve)

Calculation of statistic:

Mean of sample =788

Sample size = 30

S.D=40. So Z=?

Decision: ?

Hypothesis Testing Strategies

- The hypothesis testing determines the validity of an assumption (technically described as null hypothesis), with a view to choose between two conflicting hypothesis about the value of a population parameter.

- There are two types of tests of hypotheses
  - ✓ Non-parametric tests (also called distribution-free test of hypotheses)
  - ✓ Parametric tests (also called standard test of hypotheses).

Parametric Tests : Applications

- Usually assume certain properties of the population from which we draw samples.
  - Observation come from a normal population
  - Sample size is small
  - Population parameters like mean, variance, etc. are hold good.
  - Requires measurement equivalent to interval scaled data.

Parametric Tests

Important Parametric Tests

The widely used sampling distribution for parametric tests are

- $Z - test$
- $t - test$
- $\chi^2 - test$
- $F - test$

Note:

All these tests are based on the assumption of normality (i.e., the source of data is considered to be normally distributed).

Parametric Tests : Z-test

$\boldsymbol{Z-test}$**:** This is most frequently test in statistical analysis.

- It is based on the normal probability distribution.
- Used for judging the significance of several statistical measures particularly the mean.
- It is used even when $binomial\, distribution$ or $t-distribution$ is applicable with a condition that such a distribution tends to normal distribution when n becomes large.
- Typically it is used for comparing the mean of a sample to some hypothesized mean for the population in case of large sample, or when **population variance** is known.

Parametric Tests : t-test

$\boldsymbol{t-test}$: It is based on the t-distribution.

- It is considered an appropriate test for judging the significance of a sample mean or for judging the significance of difference between the means of two samples in case of
    - small sample(s)
    - **population variance is not known** (in this case, we use the variance of the sample as an estimate of the population variance)

Parametric Tests : χ^2 -test

$\chi^2 - test$**:** It is based on Chi-squared distribution.

- It is used for comparing a sample variance to a theoretical population variance.

Parametric Tests : $\boldsymbol{F}$ -test

$\boldsymbol{F - test}$: It is based on F-distribution.

- It is used to compare the variance of two independent samples.
- This test is also used in the context of analysis of variance (ANOVA) for judging the significance of more than two sample means.

Hypothesis Testing : Assumptions

Case 1: Normal population, population infinite, sample size may be large or small, variance of the population is known.

$$z = \frac{\bar{X} - \mu_{H_0}}{\sigma/\sqrt{n}}$$

Case 2: Population normal, population finite, sample size may large or small………variance is known.

$$z = \frac{\bar{X} - \mu_{H_0}}{\sigma/\sqrt{n}[\sqrt{(N-n)/(N-1)}]}$$

Case 3: Population normal, population infinite, sample size is small and variance of the population is unknown.

$$t = \frac{\bar{X} - \mu_{H_0}}{S/\sqrt{n}} \qquad with\ degree\ of\ freedom = (n-1)$$

and

$$s = \sqrt{\frac{\sum(X_i - \bar{X})^2}{(n-1)}}$$

Hypothesis Testing

Case 4: Population finite

$$t = \frac{\bar{X} - \mu_{H_0}}{\sigma/\sqrt{n}[\sqrt{(N-n)/(N-1)}]} \; with\ degree\ of\ freedom = (n-1)$$

Note: If variance of population (σ) is known, replace S by σ. Population normal, population infinite, sample size is small and variance of the population is unknown.

Hypothesis Testing : Non-Parametric Test

- ***Non-Parametric tests***
  - ✓Does not under any assumption
  - ✓Assumes only nominal or ordinal data

Note: Non-parametric tests need entire population (or very large sample size)

Reference

- The detail material related to this lecture can be found in

Probability and Statistics for Engineers and Scientists (8^{th} Ed.) by Ronald E. Walpole, Sharon L. Myers, Keying Ye (Pearson), 2013.

Any question?

You may post your question(s) at the "Discussion Forum" maintained in the course Web page!

Questions of the day…

1. In a hypothesis testing, suppose H_o is rejected. Does it mean that H_1 is accepted? Justify your answer.

2. Give the expressions for z, t and χ^2 in terms of population and sample parameters, whichever is applicable to each. Signifies these values in terms of the respective distributions.

3. How can you obtain the value say $P(z = a)$? What this values signifies?

4. On what occasion, you should consider z-distribution but not t-distribution and vice-versa?

5. Give a situation when you should consider χ^2 distribution but neither z- nor t-distribution.