person → crime → court → evidence

guilty not guilty

Hypotheses Testing

guess → evidence

verify

1. set the purpose → hypothesis → guess
2. collect the data
3. decide population
4. create a sample
5. perform the experiment → statistical analysis
6. draw the conclusion
7. generalise the conclusion for the whole population → hypothesis testing

hypothesis testing
- accept
- reject

known data

sample → conclusion

generalize

population

unknown data

Introduction

- Statistical inference is that branch of statistics which is concerned with using probability concept to deal with uncertainty in decision-making
- It refers to the process of selecting and using a sample statistic to draw inference about a population parameter based on the set of attributes
- It treats two classes of problems
  - Hypotheses testing i.e. to test some hypothesis about parent population from which sample is drawn
  - Estimation i.e. to use the statistics obtained from the sample as estimate of the unknown parameter of the population from which population is drawn

guess / purpose

* sample statistic → characteristic of sample

* statistics → processes / methods for data operations

* population parameter → characteristic of population

Hypothesis Testing

- It begins with assumption called as hypothesis, that we make about the population parameter
- A hypothesis is supposition made as a basis for reasoning
- A hypothesis in statistics is simply a quantitative statement about a population
- There can be several types of hypotheses
- E.g.
  - A coin may be tossed 200 times and we may get heads 80 times and tails 120 times. We may now be interested in testing the hypothesis that the coin is unbiased.
  - We may study the average weight of 100 students of a college and get result as 110 lb. We may now be interested in testing the hypothesis that the sample has been drawn from a population with average weight of 115 lb.

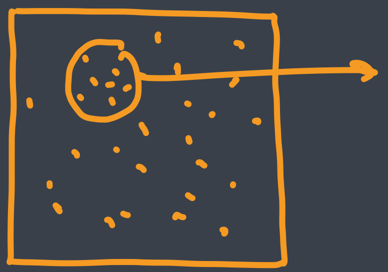
$p(H) = 0.5$, $p(H) = 0.4$

hypothesis = $p(H) \neq 0.5$

$[\mu = 5.5]$ ← quantitative statement

hypothesis

$[p(\text{rain}) = 0.5]$



$\bar{x} = 110$

$\bar{x} = \mu$

population $\mu = 115$

Procedure of Testing Hypothesis

Procedure of Testing Hypothesis

- Set up a hypothesis
  - Null hypothesis [H_0]
  - alternate hypothesis [H_1 or H_a]
- Set up a suitable significance level (α) 0.01, 0.05, 0.10
  - 99%, 95%, 90%
- Setting a test criterion
  - T-test, F-test ...
- Doing computations
- Making Decisions
  - accept Null hypothesis
  - reject Null hypothesis

Set up a hypothesis

- The first thing in hypothesis testing is to setup a hypothesis about the population parameter
- Then we collect sample data, produce sample statistics → measures [CT, variations, asymmetric..]
- Use this information to decide how likely it is that out hypothesized population parameter is correct
- E.g.
  - We assume certain value of a population mean
  - To test the validity of the assumption, we gather the sample data
  - Then we determine the difference between hypothesize mean (population mean) and actual value of sample mean
  - Then we judge whether the difference is significant
  - The smaller the difference, greater the likelihood that our hypothesized mean is correct or greater the difference, smaller the likelihood

Hypotheses Types

- The conventional approach to hypothesis testing is not to construct a single hypothesis about the population parameter, but rather to set up two hypotheses
- These hypotheses must be so constructed that if one hypothesis is accepted, the other is rejected by default and vice-a-versa
- The two hypotheses are normally referred as
  - Null hypothesis
  - Alternative hypothesis

null hypothesis = today it will rain = p(rain) = 80%.

altternate hypothesis = today it will Not rain = p(rain) ≠ 80%.

Null Hypothesis [H_0]

$\bar{x} = 110$, $\mu = 115$ $\quad$ $\mu = \bar{x}$

- It is a very useful tool in testing the significance of difference
- In its simplest form, the hypothesis asserts that there is no real difference in the sample and the population in under consideration
- Hence the word null, which means invalid, void or amounting to nothing that the difference found is accidental and unimportant arising out of fluctuations of sampling
- It is skin to the legal principle that a man is innocent until he is proved guilty
- It constitutes a challenge, and the function of the experiment is to give the facts to a chance to refute or failed to refute this challenge
- The rejection of null hypothesis indicates that the differences have statistical significance
- Acceptance of null hypothesis indicates that the differences are due to chance

Alternate Hypothesis

[H1 or Ha]

$\mu \neq \bar{x}$

- It specifies these values that the researcher believes to hold true and of course, he hopes that the sample data lead to acceptance of this hypothesis true
- It may embrace the whole range of values rather than a single value
- E.g.
  - A certain psychologist who wishes to test whether or not a certain class of people have mean IQ higher than 100 might establish following hypotheses
  - Null hypothesis
    - $H_0: \mu = 100$
  - Alternate hypothesis
    - $H_a: \mu \neq 100$

α = level of significance

α = 1 - confidence level

α = risk taken when accepting null hypothesis

α = % of Error committed when accepting null hypothesis

α = Type I Error

↳ rejecting true hypothesis

Confidence level	90	95	99
Significance level (α)	0.10	0.05	0.01

Set up significance level (α)

$z(99) = 2.58$, $z(95) = 1.96$, $z(90) = 1.64$

- Having setup the hypotheses, the next step is to test the validity of null hypothesis against that of alternative one at a certain level of significance
- The confidence with which an experimenter rejects or remains a null hypothesis depends upon the significance level accepted
- It is expressed as percentage which is probability of rejecting the null hypothesis if it is true
- E.g. when a null hypothesis is accepted at the 5% level, the statistician is running the rick that in long run, he will be making the wrong decision about 5% of the time

confidence level = 90%, 95%, 99%

10% 5% 1%

0.10 0.05 0.01

$\boxed{\mu = \bar{x}}$

2

Standard Error (SE) $= \frac{\sigma}{\sqrt{n}}$

confidence levels

margin of error = z-score × SE

$= \text{z-score} \times \frac{\sigma}{\sqrt{n}}$

2

confidence interval $= \bar{x} \pm ME$

$= 115 \pm 2$

Setting up test criterion → hypothesis test

- The third step in hypothesis testing is to construct the test criterion
- It involves selecting an appropriate probability distribution for the particular test
- That is the probability distribution which can properly be applied for testing
- Some probability distributions which are commonly used
  - T distribution → Student T-test
  - F distribution → ANOVA
  - Chi Square distribution → Chi-Square test
- E.g. if only small sample information is available, the use of normal distribution would be inappropriate

Types

parametric — uses distribution

Non-parametric — No use of distribution

Doing computations

- This step involves performing the calculations from the random sample in which are necessary for the selected test
- These calculations include testing statistic and standard error of testing statistic

test statistic = value computed by using test criterion

[p-value]

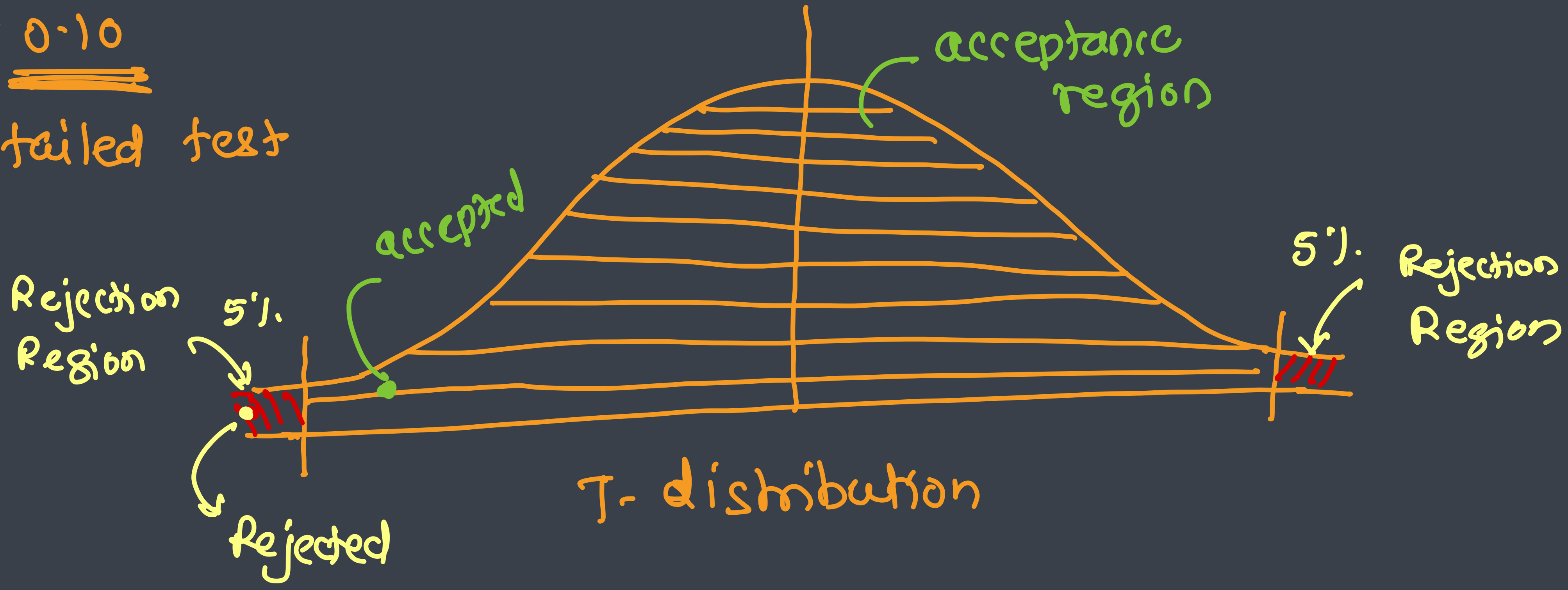
T-test formula

Making Decisions

- Finally with respect to the test calculations, we may draw the statistical conclusions and take decisions
- A statistical conclusion or decision is a decision either to reject or to accept the null hypothesis
- The decision will be dependent on whether the computed value of the test criterion falls in the accepted region or rejection region

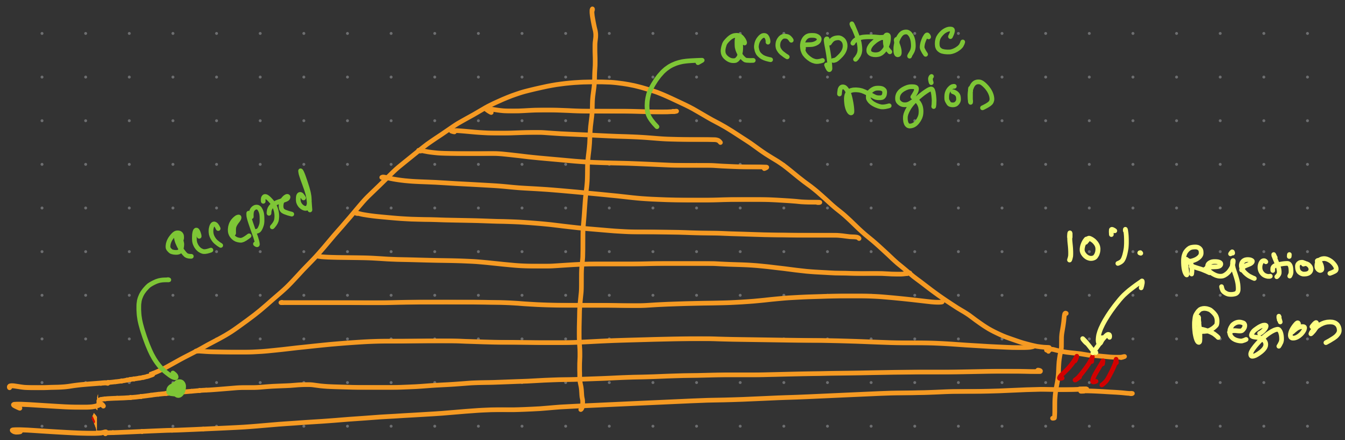
$\alpha = 0.10$

two tailed test

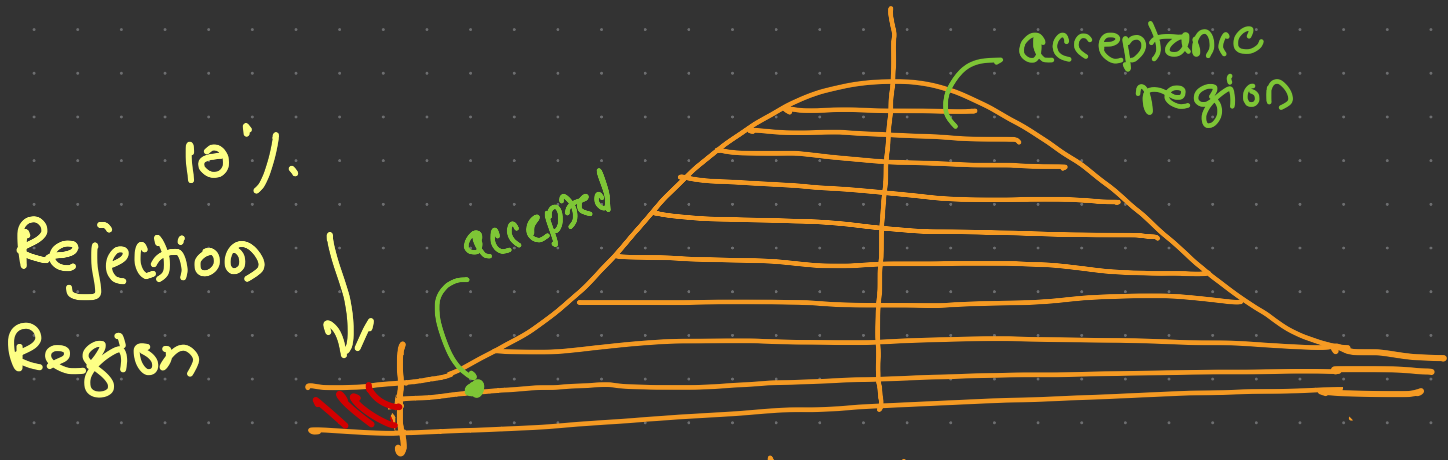


One tailed test

$\alpha = \underline{0.10}$



T- distribution



T- distribution

Testing Error

- When a statistical hypothesis is tested, there are four possibilities
  - The hypothesis is true, but our test rejects it (Type I error)
  - The hypothesis is false, but our test accepts it (Type II error)
  - The hypothesis is true, and our test accepts it (correct decision)
  - The hypothesis is false, and our test rejects it (correct decision)
- While testing the hypothesis, the aim is to reduce both the types of error

confusion matrix

	Accept	Reject
H_0 is true	Correct	Error
H_0 is false	Error	Correct

Rejecting true H_0 (Type I Error)

Accepting false H_0 (Type II Error)

Type I Error

- In hypothesis testing, the Type I error is committed by rejecting the null hypothesis when it is true
- The probability of committing a Type I error is denoted by α where
  - α = probability (Type I error)
  - α = probability (rejecting H_0 | H_0 is true)

Type II Error

- The Type II error is committed by not rejecting (i.e. accepting) the null hypothesis when it is false
- The probability of committing a Type II error is denoted by β where
  - β= probability (Type II error)
  - β= probability (not rejecting H_0 | H_0 is false)

probability (accepting Ho | Ho is false)

30% α

10 Errors

100 samples

$$\alpha + \beta = 1$$

$$\alpha = 1 - \beta$$

Trade-off between errors

- The probability of making one type of error can only be reduced if we are willing to increase the probability of making the other type of error
- In other words, to get low β, we will have to put up with a high α
- It is more dangerous to accept a false hypothesis (Type II error) than to reject a correct one *
- Hence, we keep the probability of Type I error at a certain level called as level of significance (α)
- The level of significance also known as size of the rejection region or size of critical region or simply the size of the test is traditionally denoted by α
- In most cases, the level of significance is generally fixed as 5% which means that the probability of accepting true hypothesis is 95%

Hypothesis Tests

Two-tailed tests of hypothesis

p-value

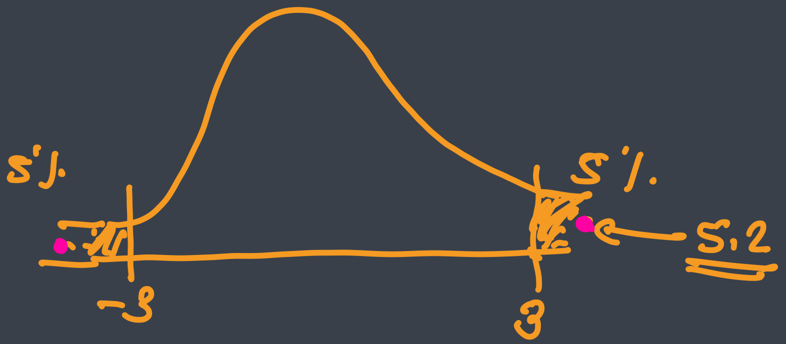
- A two tailed hypothesis test will reject the null hypothesis, if the sample statistic is significantly higher or lower than the hypothesized population parameter
- Thus in the two tailed test, the rejection region is located on both the sides
- If we want to reduce the risk of committing type I error, we have to reduce the size of rejection region
- For this the hypothesis may be tested at lower level of significance
- As we decrease the size of rejection region, we increase the probability of accepting hypothesis
- Two-tailed test is appropriate when the alternative hypothesis has equal-to or not-equal-to sign
  - $H_a: \mu \neq 100$

$t = -5.2$

$\alpha = 0.1$

$t = 5.2$

z-score = 3



$5.2 > 3$

One-tailed test

- It is so called because the rejection region will be located only on one side which may be either right or left depending on the test and hypothesis formed
- One-tailed test is used when alternate hypothesis has >, >=, < or <= sign
  - $H_a: \mu > 100$
  - $H_a: \mu < 100$

$\alpha = 0.05$