statinfer

Testing of Hypothesis

statinfer.com

Contents

- Testing of Hypothesis
- Null Hypothesis
- Test statistic
- P-Value

Intuition – Example1

- We have to buy 10 Kg Laddu (Sweet) - How do we do it?
- Assume it is good
- Take a sample to test
- taste the sample
- Buy or not buy based on the taste

Intuition – Example1

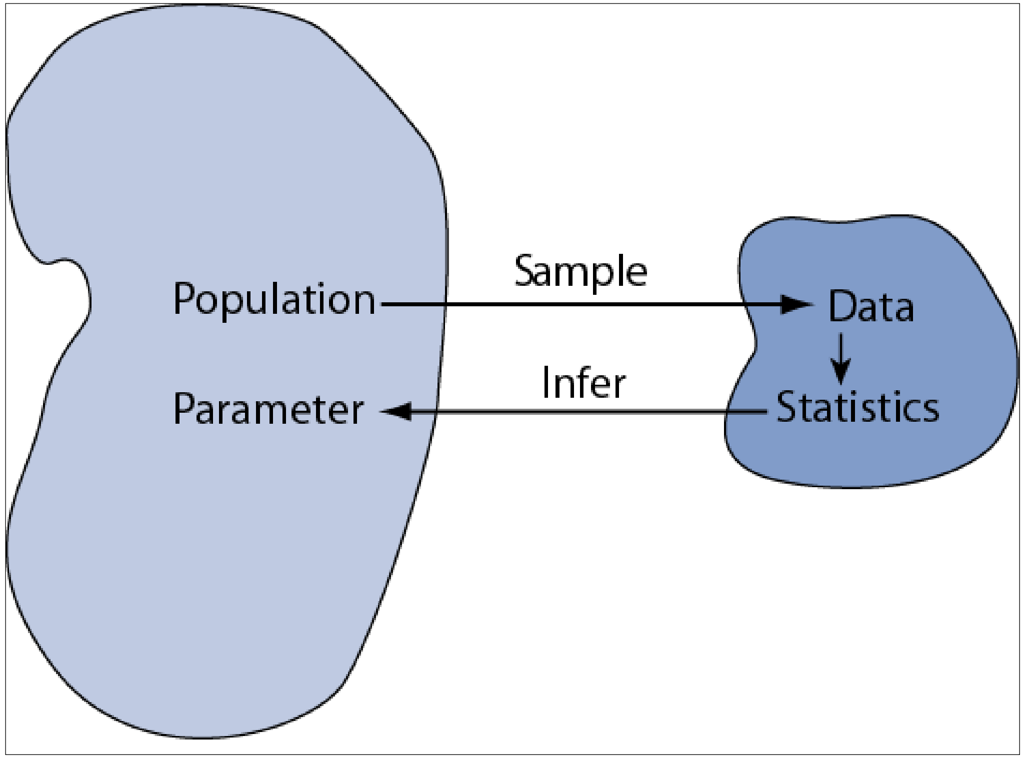
- 10 Kg Laddu - Population
- Assume it is good – Hypothesis (Null Hypothesis)
- Take a sample to test (Sample)
- taste the sample (Test Statistic)
- Buy or not buy based on the taste (Inference)

Intuition – Example2

- 1M Soaps - Population
- Assume Average weight is 250 grams– Hypothesis (Null Hypothesis)
- Take a sample of 100 soaps to test (Sample)
- Find the average weight(Test Statistic)
- Accept of reject null hypothesis based on sample average(Inference)
  - If sample average is 230 grams
  - If the sample average is 249.5 grams

Statistical Inference

- Inferences about a population are made on the basis of results obtained from a sample drawn from that population
- Want to talk about the larger population from which the subjects are drawn, not the particular subjects!



- A hypothesis test is a process that uses sample statistics to test a claim about the value of a population parameter.
- A verbal statement, or claim, about a population parameter is called a statistical hypothesis.

What is Testing of Hypothesis

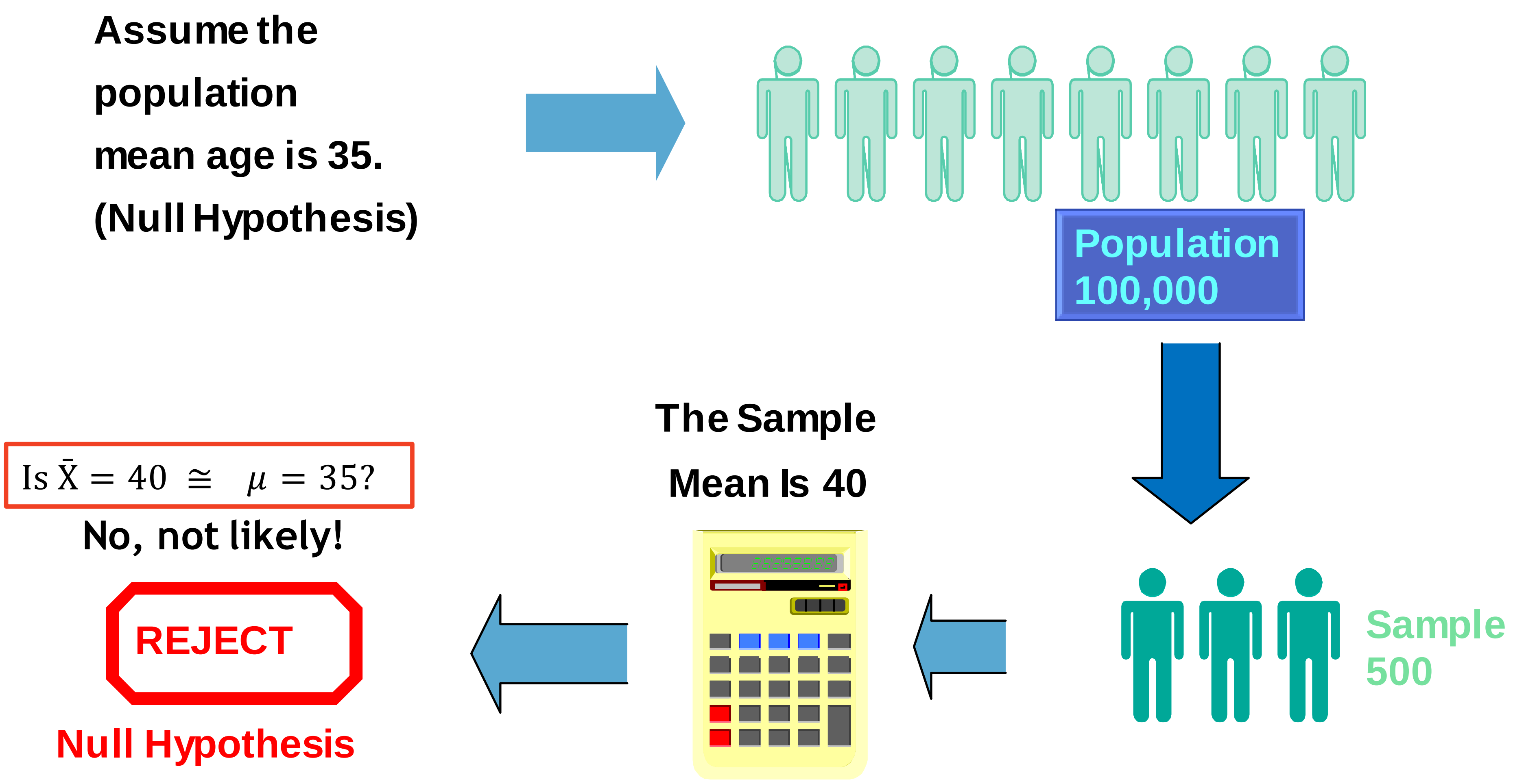
- What are we testing ? → A hypothesis on population parameter
- Hypothesis →A verbal statement, or claim, or an assumption about a population parameter is called a statistical hypothesis.
- How are we testing ? → using sample statistic.

Hypothesis Testing – An example

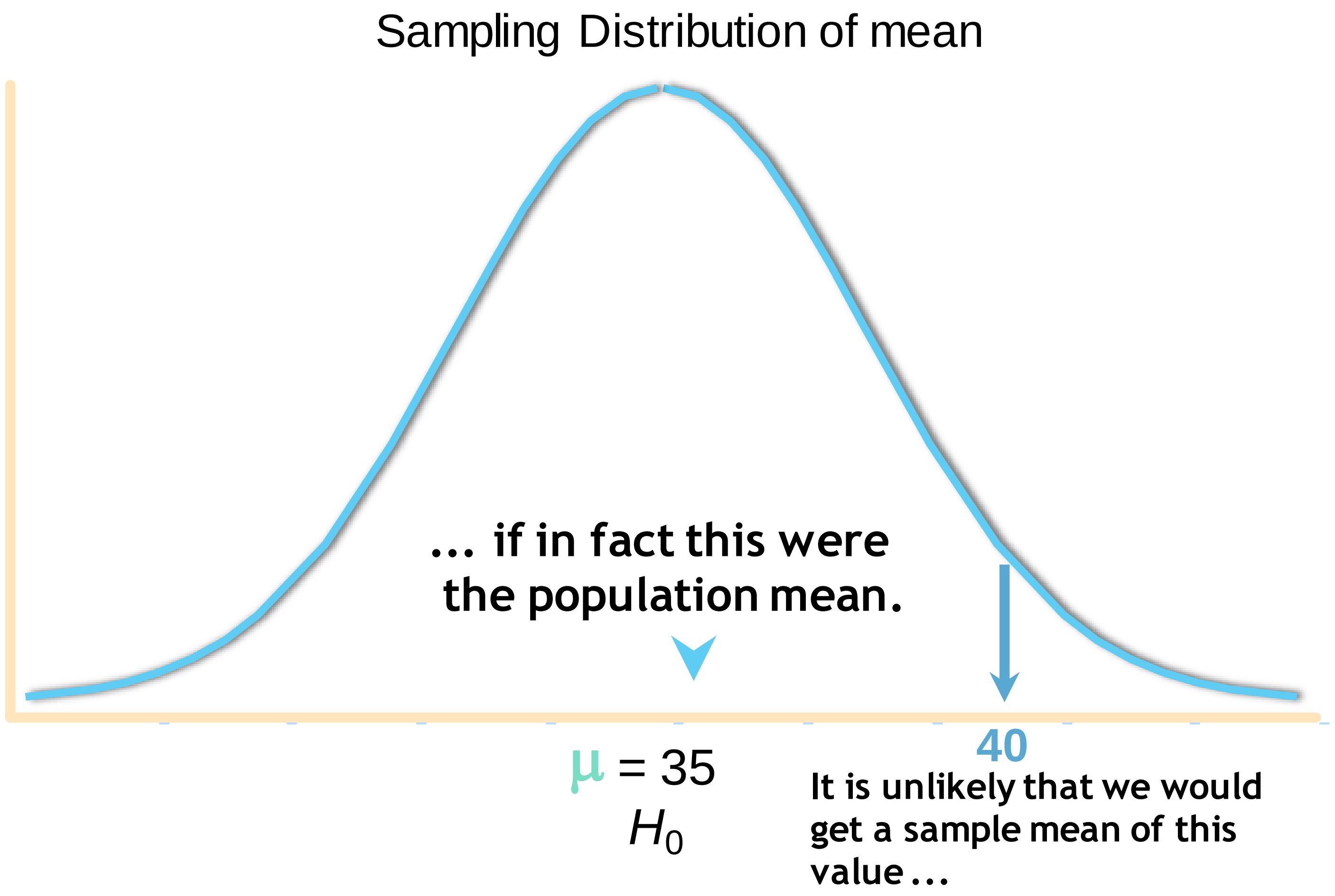
CEO of a national bank claims that employees mean age is 35. There are a total of 292,215 employees. How can we prove or disprove it?

- Take a random sample (500) and find their age
- If it is near 35 then we say there is no evidence to reject that hypothesis
- What if sample average age is lower or higher than 35 ?
- How far is really far? We want to quantify the severity of deviation……
- Lets find the probability of this occurrence
- It if it is really low, then we say we reject null hypothesis

Hypothesis testing process



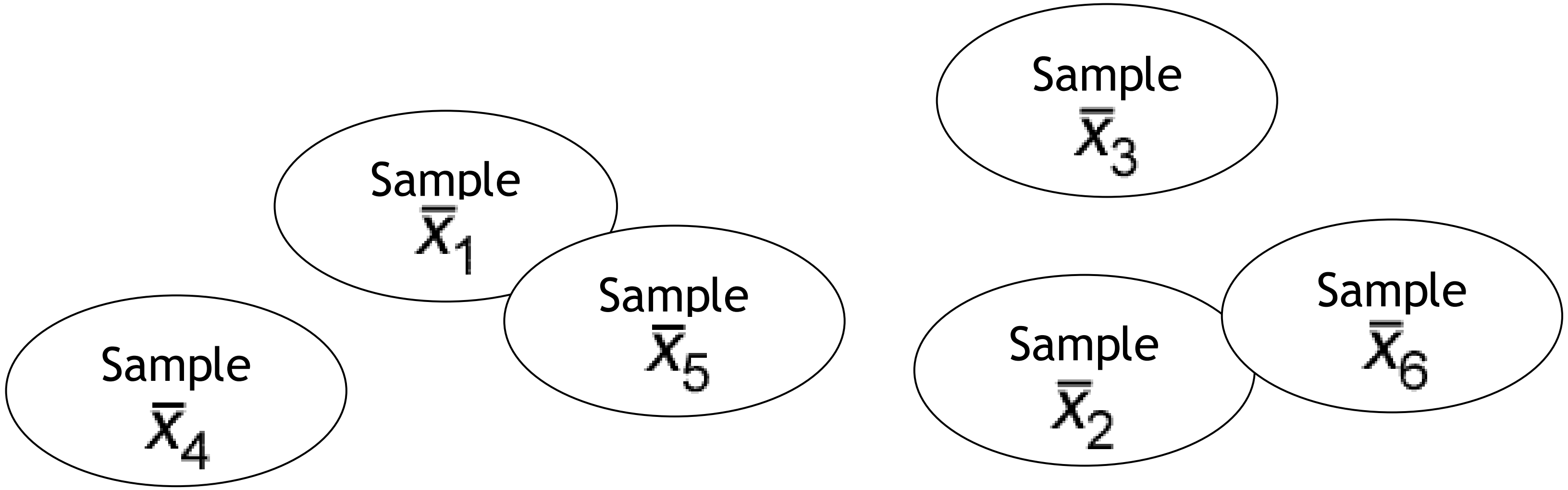
Reason for Rejecting H0



... Therefore, we reject the null hypothesis that μ = 35.

What is sampling distribution

- A sampling distribution is the probability distribution of a sample statistic that is formed when samples of size n are repeatedly taken from a population.
- If the sample statistic is the sample mean, then the distribution is the sampling distribution of sample means.



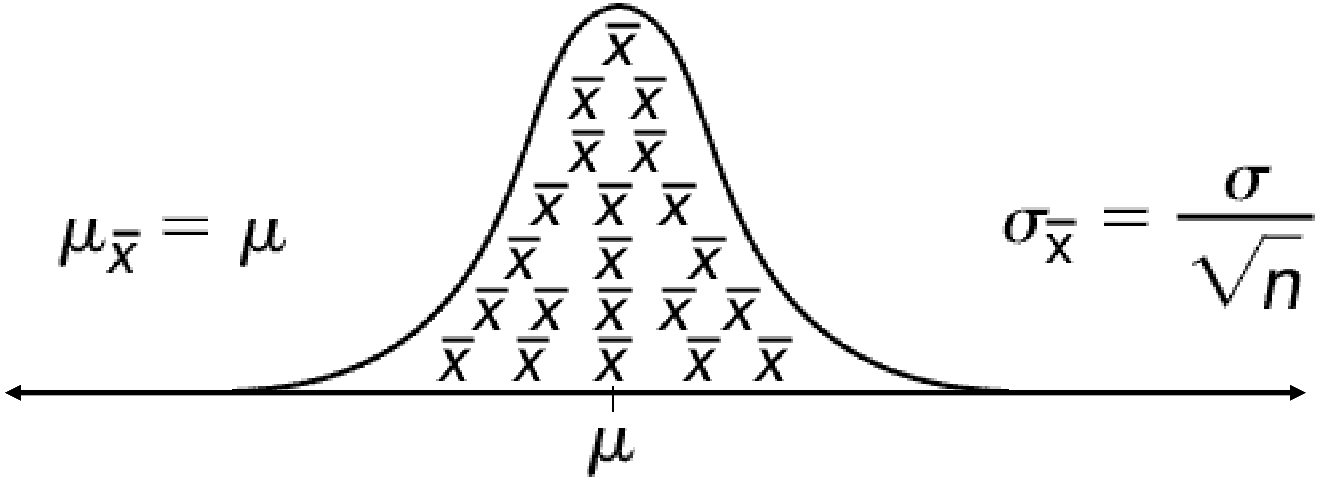
The sampling distribution consists of the values of the sample means, $\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4, \bar{x}_5, \bar{x}_6, \ldots$

Central Limit theorem

If a sample *n* (30) is taken from a population with
any type distribution that has a mean = μ and standard deviation = σ

the ***sample means*** will have a **normal distribution** $\mu_{\bar{x}} = \mu$ **and s.d** $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

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



$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

In short

- You start with some assumption (Null Hypothesis)
- You select a random **sample** or assume that the sample you have is random
- You do a test on the same(calculate the sample statistic)
- If the sample is taken from population then you expect the sample statistic to be within certain limits. (Selection criterion)
- If it is beyond those limits then we reject our assumption(Inference)

Five Step in Testing of Hypothesis

1. Make Assumptions and meet test requirements.
2. State the null hypothesis.
3. Select the sampling distribution and establish the critical region.
4. Compute the test statistic.
5. Make a decision and interpret results.

Step 1: Make assumptions and meet test requirements

- Random sampling
  - Hypothesis testing assumes samples were selected using random sampling.
  - In this case, the sample of 500 cases was randomly selected from all major branches.
- Level of Measurement is Interval-Ratio.
  - Yes age is not a categorical variable
- Sampling Distribution is normal in shape.
  - What is the sampling distribution of age?
- This is a “large” sample ($n \geq 100$).

Step 2 State the Null Hypothesis

- H0: $\mu = 35$
- In other words, Ho: No difference between the sample mean and the population parameter
- In other words, The sample mean of 40 is really the same as the population mean of 35 – the difference is not real but is due to chance.
- In other words, The sample of 500 comes from a population that has average age of 35
- In other words, The difference between 35 and 40 is trivial and caused by random chance.

Step 2 (cont.) State the Alternate Hypothesis

- H1: $\mu \neq 35$
- Or H1: There is a difference between the sample mean and the population parameter
- Or The sample of 500 comes a population that does not have average age 35 In reality, it comes from a different population.
- Or The difference between 40 and 35 reflects an actual difference
- Or the average age of the population is more than 35

Step 3 Select Sampling Distribution and Establish the Critical Region

- What is the sampling distribution of population mean?
- What is alpha?
  - Probability of rejecting H0 when it is true
- α is the indicator of “rare” events.
- Any difference with a probability less than α is rare and will cause us to reject the H0.
- We started with H0 as true, we still want to reject the null if the statistic is beyond a certain value,
- We already know about some unlikely values of test statistic when null hypothesis is true
- for example if the average age of the sample is 60, we definitely want to reject null

Step 4: Use Formula to Compute the Test Statistic - Z for large samples (≥ 100)

- We got the sample average as 40, the age according to null hypothesis is 35, there is a difference of 5, is it due to chance?
- How bad is this difference of 5?

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{N}}$$

When the Population σ is not known, use the following formula:

$$Z = \frac{\bar{X} - \mu}{s/\sqrt{n-1}}$$

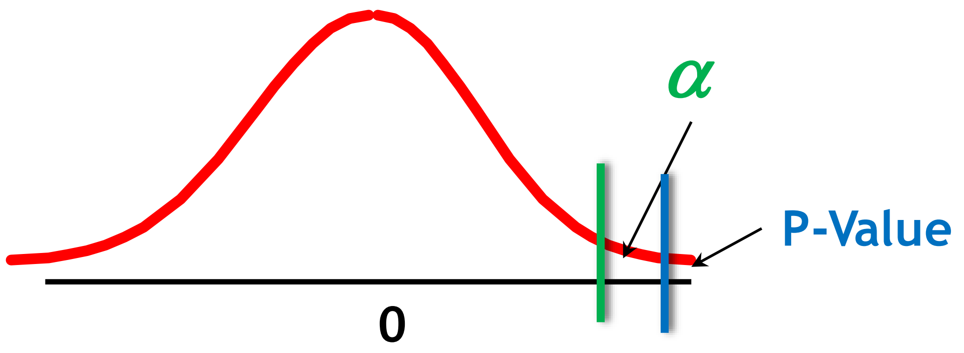
$$Z = \frac{40-35}{7.86/\sqrt{500-1}}$$

Step 5 Make a Decision and Interpret Results

- The obtained Z score fell in the Critical Region, so we reject the H0.
  - If the H0 were true, a sample outcome of 14 would be unlikely.
  - Therefore, the H0 is false and must be rejected.

H_0: $\mu = 35$

H_1: $\mu > 35$



What does z of 14 mean? The probability of z being more than 14 is less than 0.000000001

- It is like getting more than 25 heads in a row when you toss a coin
- It is like drawing the same card more than 6 times from a shuffled deck
- If the average age of 40 is just by chance then compare that chance with above examples

In simple terms

- Ok, population average age is 35
- I took a random sample(500) and found the average age to be 40
- But it is extremely unlikely that sample age is 40 when population age is 35, in fact the probability of this event is less than 1 in billion
- How did I find the probability
  - Coz I know that sample distribution of mean follows normal distribution
  - Using that I quantified the probability of mean >=40, which is nothing but p value

What is P-Value

- If the observed statistic happens to be just a chance, p-values tells us what is the probability of that chance
- **The P-value answer the question:** What is the probability of the observed test statistic or one more extreme when H0 is true?
- Given H0, probability of the current value or extreme than this
- Given H_0 is true, probability of obtaining a result as extreme or more extreme than the actual sample
- The **observed significance level,** or **p-value** of a test of hypothesis is the probability of obtaining the observed value of the sample statistic, or one which is even more supportive of the alternative hypothesis, under the assumption that the null hypothesis is true.

Individual Impact of the variables - Regression

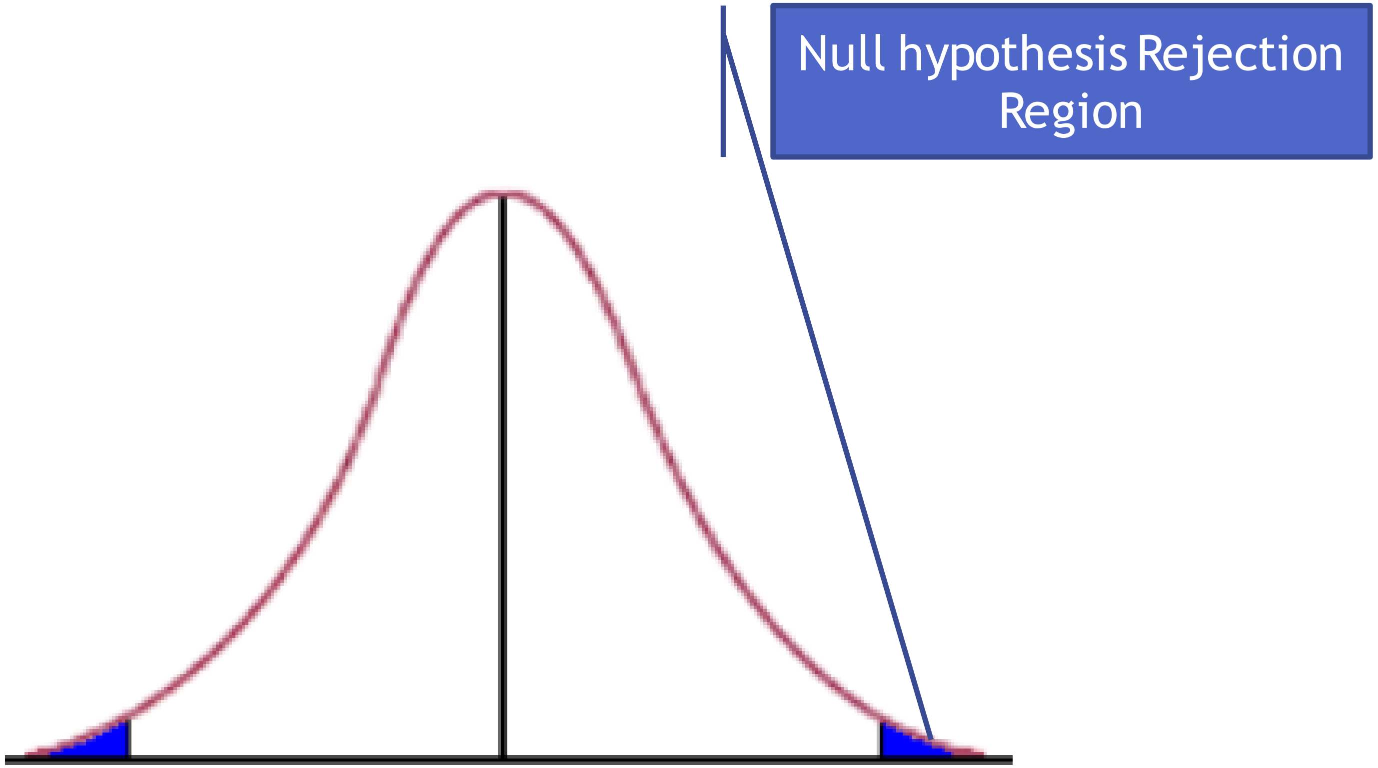
To test

$H_0: \beta_i = 0$ No Impact

$H_a: \beta_i \neq 0$ Impactful

Test statistic: $t = \dfrac{b_i}{s(b_i)}$

Reject H_0 if P-value is >0.05



Testing of Hypothesis in Regression

- Look at the P-value
- Probability of the hypothesis being right.
- Individual variable coefficient is tested for significance
- Beta coefficients follow t distribution.
- Individual P values tell us about the significance of each variable
- A variable is significant if P value is less than 5%. Lesser the P-value, better the variable
- Note it is possible all the variables in a regression to produce great individual fits, and yet very few of the variables be individually significant.

To test

$$H_0 : \beta_i = 0$$
$$H_a : \beta_i \neq 0$$

Test statistic:

$$t = \frac{b_i}{s(b_i)}$$

Reject H_0 if

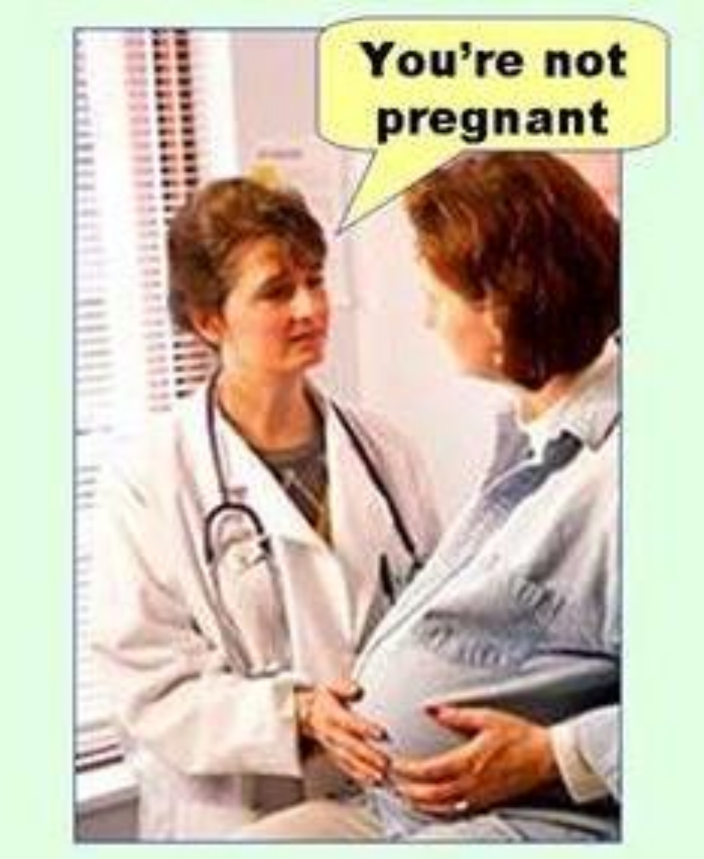
$$t > t(\frac{\alpha}{2}; n-k-1) \quad or$$
$$t < -t(\frac{\alpha}{2}; n-k-1)$$

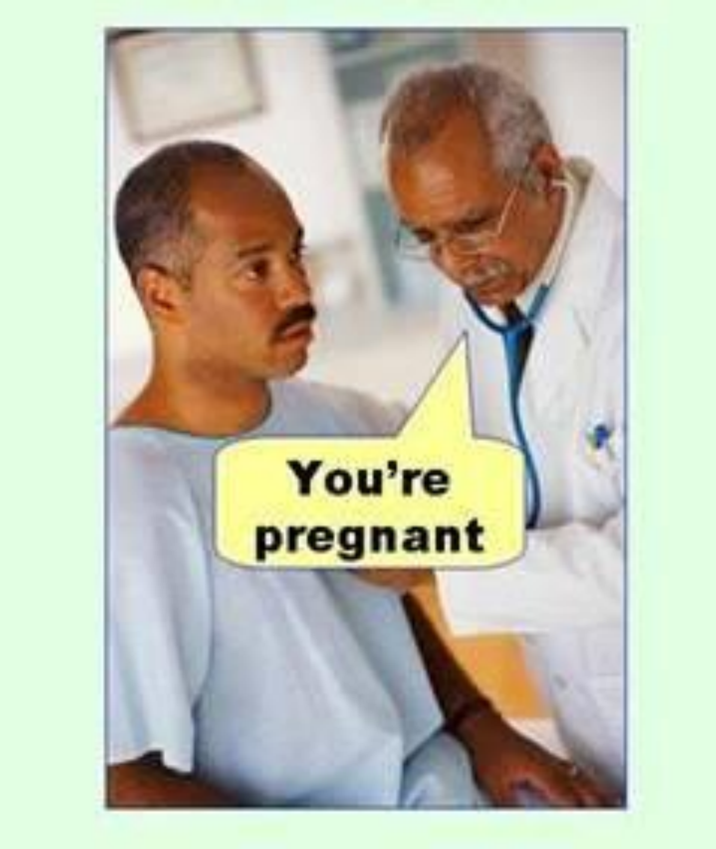
There will be errors

- Testing is done based on samples
- There will always be some error
- There are two types of errors in testing
  - Rejection error
  - Acceptance error

The Errors

H0: I am pregnant





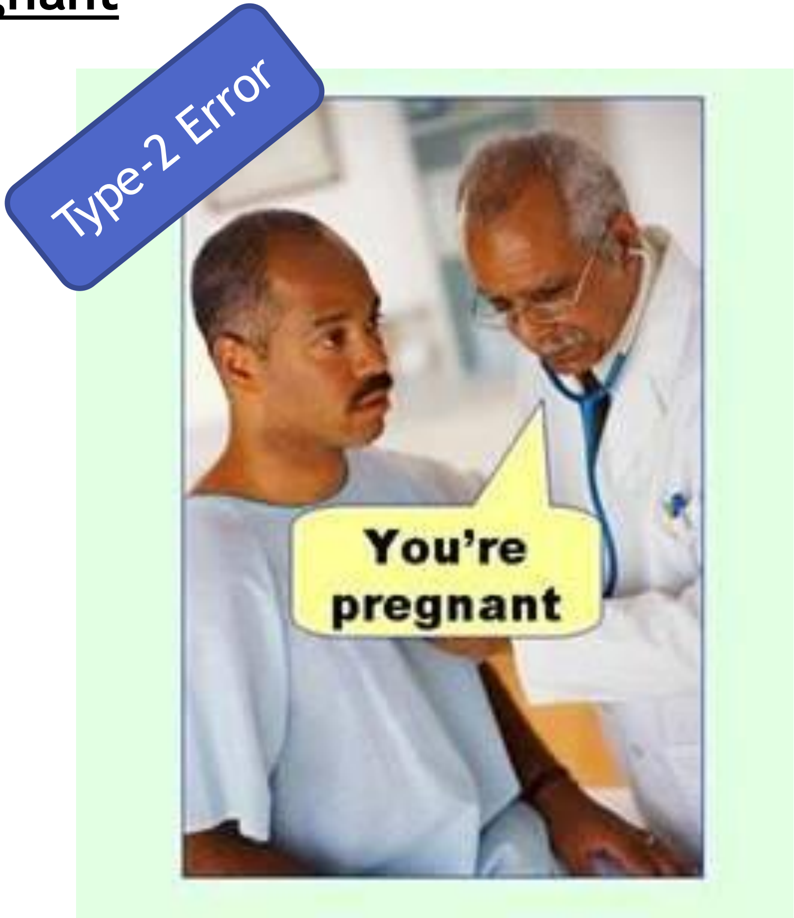
The Errors

<u>H0: I am pregnant</u>

Wrongly rejecting H0

Rejection Error

Type-1 Error



Wrongly accepting H0

Acceptance Error

Type-2 Error

The errors

- What of, the population was perfect and the problem is only in my sample. What if I rejected H0, wrongly
  - Wrongly rejecting H0
  - Rejecting H0 when it is true
- What of, the population was not right and the sample passes my test and I accept H0 wrongly
  - Wrongly accepting H0
  - Accepting H0 when it is true
- A **type I error** occurs if the null hypothesis is rejected when it is true.
- A **type II error** occurs if the null hypothesis is not rejected when it is false.

Error Types

	Test Result	
Reality	**Don't Reject H_0**	**Reject H_0**
H_0 True	Correct	Type I Error
H_0 False	Type II Error	Correct

- As usual we don't want any error, but we start with setting up the type-1 error
- I.e the error incurred in rejecting H0 when true
- This is also known as level of significance
- Eg:By fixing the level of significance at 5%, we are ready to reject 5% of the extreme values, since we feel they are most unlikely

Which error is bad?

- Both are bad. Depends on the situation
- False negative
  - Miss what could be important
    - Testing a metal whether it is gold or not
    - Are these samples going to be looked at again?
- False positive
  - Waste resources following dead ends
    - Test whether a drug is deadly or not
- Reducing one error increases the other error. Adjust level of significance based on situation

Level of Significance α

- Defines unlikely values of sample statistic if Null Hypothesis is true
  - Called Rejection Region of Sampling Distribution
- Designated α (alpha)
- Typical values are 0.01, 0.05, 0.10
- Selected by the Researcher at the Start Provides the Critical Value(s) of the Test
- P(Type I error)

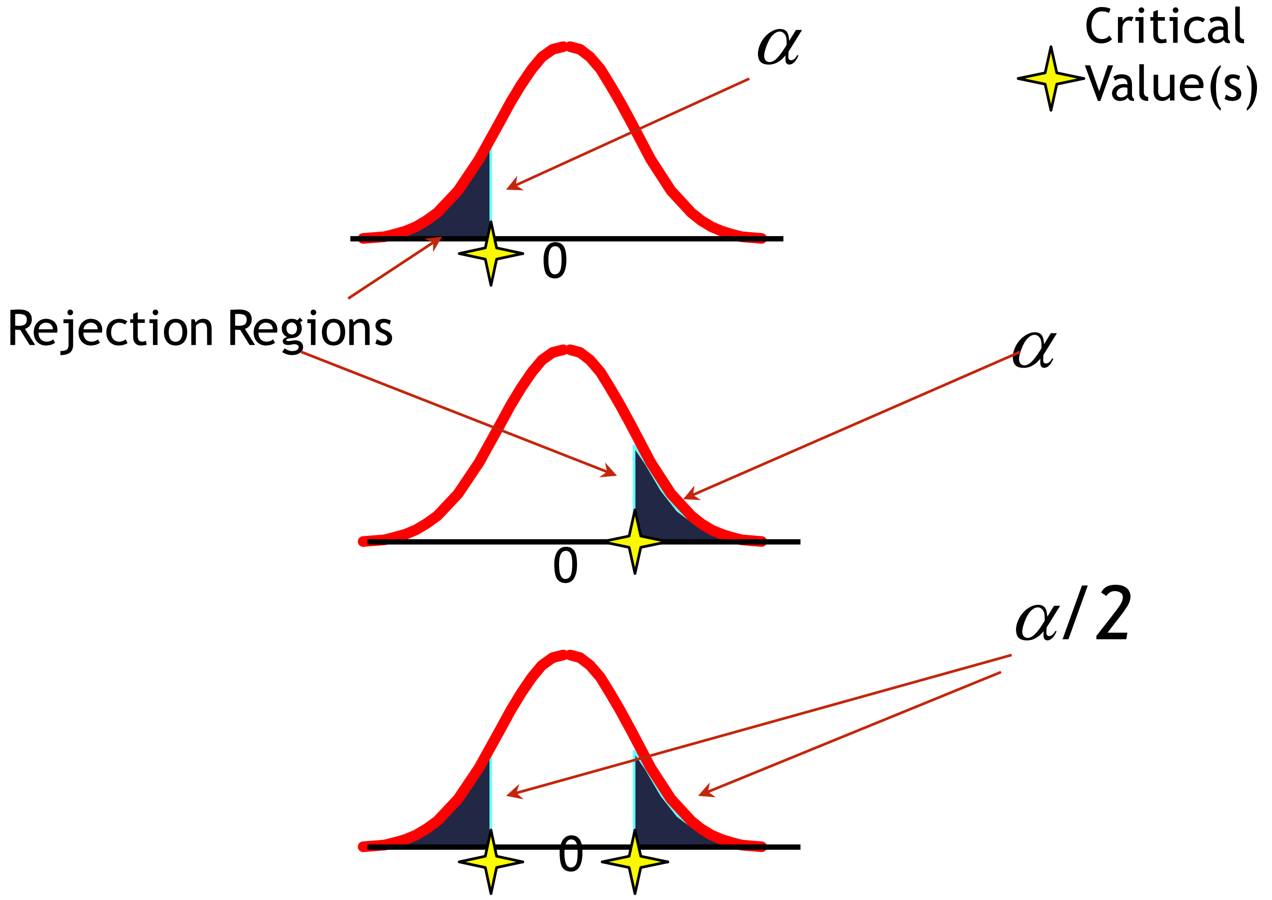
Level of Significance, α and the Rejection Region

H_0: $\mu = 35$
H_1: $\mu < 35$

H_0: $\mu = 35$

H_1: $\mu > 35$

H_0: $\mu = 35$
H_1: $\mu \neq 35$



LAB – Z-test for single mean

Take a sample of 100 from the Rossman Store Sales data. Test the below null hypothesis

Test-1:

- Null Hypothesis : Average number of customers = 750/700/650
- Alt Hypothesis : Average number of customers ≠750 /700 /650

Test-2:

- Null Hypothesis : Average sales = 5600
- Alt Hypothesis : Average sales ≠ 5600

Other Tests

- Z-test for two means
  - Tests whether the means of two independent large samples are significantly different.
- Z-test for proportions
  - Tests whether the means of two independent large samples are significantly different.
- F- test
  - Testing the equality of means across several groups. Also Known as ANOVA test
- T-tests
  - Same as Z-test for small samples
- Chi-Square test of independence
  - Testing the independence between two categorical variables

Testing Multiple Means

Z- Test for two means

- Test statistic

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

LAB – Testing two means

- Test-1
  - In Rossman sales data, is there a significant difference in average sales when Promo=0 vs Promo=1
- Test-2
  - In Rossman sales data, is there a significant difference in average sales when School_Holiday=0 vs School_Holiday =1

Use z-test for two means

Testing Multiple Means

F-test

Source of Variance	Sum of Squares	Degrees of freedom (df)	Mean Squares (SS/df)	F-Ratio (F-Statistic)
(Between) Group or Treatment	$SSTr = \sum n_i \bar{x}_i^2 - N\bar{\bar{x}}^2$	k-1	$MSTr = SSTr / (k-1)$	$F = MSTr / MSE$
(Within Group) or Error	$SSE = SSTo - SSTr$	N-k	$MSE = SSE / (N-k)$	
Total	$SSTo = \sum x^2 - N\bar{\bar{x}}^2$	N-1		

LAB : F-test ANOVA

- Test-1
  - In Rossman sales data, is there a significant difference in average sales between store type a, b, c and d ?
- Test-2
  - In Rossman sales data, is there a significant difference in average sales between the days 2, 3 and 4?

Chi square test for Independence

- Chi square test of independence
- Gender and Exam result – Are they independent?

Exam Result

Gender

	Pass	Fail	Total
M	100	400	500
F	500	0	500
Total	600	400	1000

What are the expected values?

What are the expected values if the variables are independent

Exam Result

Gender

	Pass	Fail	Total
M	?	?	500
F	?	?	500
Total	600	400	1000

Chi square test for Independence

Exam Result

Gender

	Pass	Fail	Total
M	100	400	500
F	500	0	500
Total	600	400	1000

These are observed values

When Exam Result is Independent of Gender

Expected Values

Exam Result

Gender

	Pass	Fail	Total
M	300	200	500
F	300	200	500
Total	600	400	1000

Chi-square statistic

- Summarize closeness of $\{f_o\}$ and $\{f_e\}$ by

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

where sum is taken over all cells in the table.

- When H_0 is true, sampling distribution of this statistic is approximately (for large *n*) the *chi-squared probability distribution*.

LAB: Chi-Square Test

- Perform the Chi-square test on all the cross tables in bank market data

Chi-Square Test – Issues

- Very Susceptible to sample size.
- For small samples very likely to show independence.
- For larger samples very likely to show dependence.

statinfer

Thank you