HYPOTHESIS TESTING

- In inferential statistics, you learnt how to draw inferences about the population mean from a small sample. As these inferences are based on multiple assumptions, hypothesis testing helps you validate its likelihood through different statistical tests (such as the critical value method and the p-value method).
- The statistical analysis learnt in Inferential Statistics will enable you to make inferences about the population mean from the sample data. However, sometimes, you have some initial assumptions about the population mean, and you want to confirm those using sample data. Hypothesis testing allows us to test claims about the population and find out how likely they are to be true.
- You gained an understanding of some theoretical topics of hypothesis testing, such as the types of hypotheses, types of tests and types of errors.
- You also learnt how to formulate a hypothesis and make a decision using the critical value method or the p-value method.
- Then, you learnt about the industry-related application of hypothesis testing.

Common Interview Questions

1. How do you assess the statistical significance of an insight?
2. When should you use a t-test vs a z-test?
3. How do you describe p-value?
4. What is the empirical rule?
5. Give some examples of some random sampling techniques.
6. What is the difference between type I error and type II error?
7. What are the types of sampling in statistics?
8. What are null and alternative hypotheses?
9. What is the relationship between the confidence level and the significance level in statistics?
10. Name three widely used symmetric distributions.
11. What is one sample t-test?
12. What is two sample t-test?

Hypothesis Testing

- Understanding hypothesis testing
- Null and alternative hypothesis
- Decision making
- Critical value method

- p-value method

Type of errors

- Type I error
- Type II error

Industry demonstration

- T distribution
- A/B testing

HYPOTHESIS TESTING

What is hypothesis testing?

To confirm your conclusion (or hypothesis) about the population parameter (which you know from EDA or your intuition).

Hypothesis testing starts with the formulation of two hypothesis:

- **Null Hypothesis(H_0):** The status quo
- **lternate Hypothesis(H_1):** The challenge to the status quo
- The **null hypothesis** always has the followings signs: = OR ≤ OR ≥
- The **alternate hypothesis** always has the followings signs: ≠ OR < OR >

Making a Decision

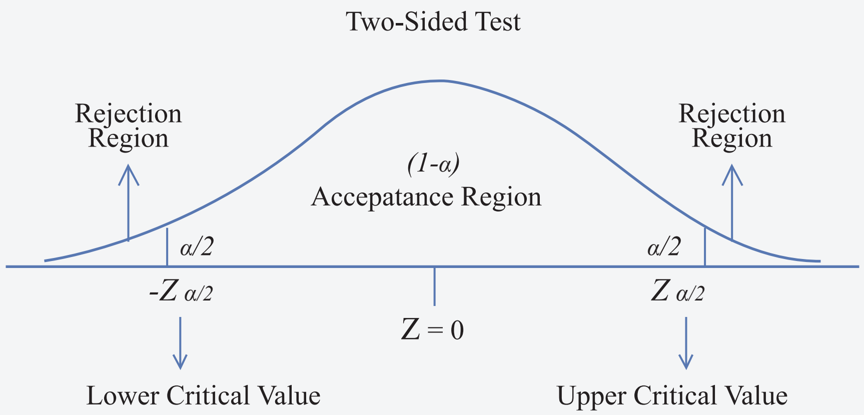
- Reject the Null Hypothesis
- Fail to reject the Null Hypothesis
- Critical Region: A set of values of the test statistic for which the null hypothesis is rejected

Types of errors

- **Type-I Error (α):** occurs when you reject a true null hypothesis
- **Type-II Error (β):** occurs when you fail to reject a false null hypothesis

	H_0 is true	H_0 is false
Reject H_0	Type I error	Correct decision
Do not Reject H_0	Correct decision	Type II error

Types of tests



- Two-Tailed Test: (≠ in H_1) Rejection region on both sides of the distribution
- Lower-Tailed Test: (< in H_1) Rejection region on left side of the distribution
- Upper-Tailed Test: (> in H_1) Rejection region on right side of the distribution

Critical value method

- Significance Level (α): Probability of rejecting the null hypothesis when it is true
- Critical values (UCV and LCV): A cut-off value that is compared with a test statistic in hypothesis testing to check whether the null hypothesis should be rejected or not.
- It defines the upper and lower bounds of the confidence interval

1. Calculate the value of Z-critical value (Z_c) from the given value of α (significance level)
2. Calculate the critical values (UCV and LCV) from the value of Z_c
3. Make the decision based on the value of the sample mean x with respect to the critical values (UCV AND LCV)

DATA WRANGLING

How to calculate p-value

- Calculate the z-score for the sample mean point on the distribution
- Calculate the p-value from the cumulative probability for the given z- score using the z-table
- Make a decision based on the p-value (multiply it by 2 in a two-tailed test) with respect to the given value of α (significance value)

p-value method

p-value: Probability that the null hypothesis will be not be rejected

A higher p-value implies a higher probability of failing to reject the null hypothesis. A lower p-value implies a higher probability of the null hypothesis getting rejected

P-Value vs Critical Value

Tail area, defines sensitivity: Eg 5% = will reject true null hypothesis 5% of time.

State Null and Alternate Hypotheses

↓

Choose Alpha (Significance) Level

↓

Calculate Test Statistic Based on Sample

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

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

↓

Choose Approach

P-Value Approach
Compare areas

↓

Compute p-value

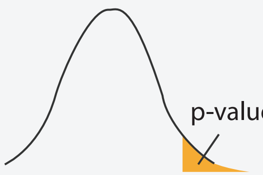
Probability value = tail area

↓

Compare p-value to alpha level

↓

Reject null if p-value ≤ alpha level



"Classical" Critical Value Approach
Compare areas

↓

Find critical value
$\alpha,\ \alpha/2,\ t\alpha\ ,\ \alpha/2$

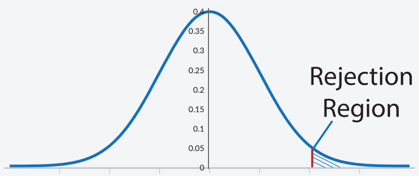
Cut off value, splits graph in two with a "rejection region".

↓

Compare test statistic to critical value

↓

Reject null if test value falls in ejection region



HYPOTHESIS TESTING

How to find z-score from z-table

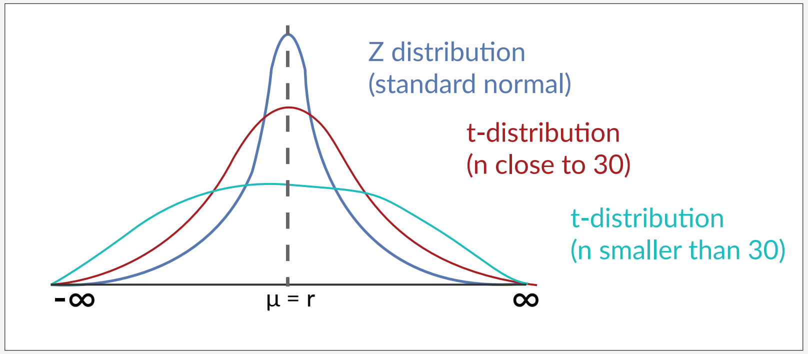
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5398	0.5478	0.5517	0.5557	0.5596	0.5636	0.5636	0.6714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6100	0.6141
0.3	0.6170	0.6217	0.6255	0.6203	0.6331	0.6368	0.6406	0.0443	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	0.8051	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

Z Score table

HYPOTHESIS TESTING

T-Distribution

- t-distribution is used when the standard deviation of a population is unknown
- The degrees of freedom of a t-distribution is equal to the sample size -1 = (n - 1)
- For a sample size of ≥30, t-distribution becomes the same as normal distribution
- The output values and results of both t-test and z-test are the same for a sample size of ≥30



Two-Sample Proportion Test

- It is used when your sample observations are categorical, with two categories
- A/B testing is a direct industry application of the two-sample proportion test
- It provides a way to test two different versions of the same element and see which one performs better

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1 - \hat{q}_1}{n_1} + \frac{\hat{p}_2 - \hat{q}_2}{n_2}}}$$

where $\hat{p}_1$ = proportion of success in sample one

$\hat{p}_2$ = proportion of success in sample two

$\hat{q}_1 = 1 - \hat{p}_1$

$\hat{q}_2 = 1 - \hat{p}_2$

n_1 = size of sample one

n_2 = size of sample two

and

$\sqrt{\frac{\hat{p}_1 - \hat{q}_1}{n_1} + \frac{\hat{p}_2 - \hat{q}_2}{n_2}}$ = the standard error of difference between two sample proportions."

Two-Sample Mean Test

- Paired: It is used when your sample observations are from the same individual or object. During this test, you are testing the same subject twice
- Unpaired: It is used when your sample observations are independent. During this test, you are not testing the same subject twice

Chi-Squared Test

- To conduct hypothesis tests with categorical variables to see if there is an effect of one on another
- To determine whether or not there is a significant relationship between two nominal (categorical) variables
- O_i - Observed value, E_i - Expected Value $X^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

ANOVA

- Analysis of variance (ANOVA) can determine whether the means of three or more groups are different
- ANOVA uses F-tests to statistically test the equality of means