LIVEWIRE
FOR LIVE CAREERS

BUSINESS ANALYTICS

CONFIDENCE INTERVAL

- Describe the amount of uncertainty associated with a sample estimate of a population parameter
- Confidence Interval Data Requirements
  - Confidence level (90%, 95%, or 99%)
  - Statistic (Sample mean)
  - Margin of error

    (Margin of error = Critical value * Standard deviation of statistic)
    (Margin of error = Critical value * Standard error of statistic)
- Range of the confidence interval is defined by the sample statistic ± margin of error

Critical value:

- Factor used to compute the margin of error
- To find the critical value:
    - Compute alpha (α): α = 1 - (confidence level / 100)
    - Critical probability (p*): p* = 1 - α/2
    - Z score

Margin of Error:

- Expresses the maximum expected difference between the true population parameter and a sample estimate of that parameter

Statistic:

- A statistic is a characteristic of a sample
- Generally, a statistic is used to estimate the value of a population parameter

Four Steps to find CONFIDENCE INTERVEL

- Identify a sample statistic (Mean Value)
- Select a confidence level (99%, 95%, 90%)
- Calculate margin of error (1%, 5%, 10%)

$$Z_{\alpha/2} * \frac{\sigma}{\sqrt{n}}$$

σ = SD

Z=Z value

α=CI Value

- Range of CI – Output can vary within the rage
  (Confidence interval = sample statistic ± Margin of error)

Note: Most widely used Confidence Interval is 95%

P-Value

- Function of the observed sample result that is used for testing a statistical hypothesis
- Before analytics, a threshold value (significance level of the test) will be chosen
- Traditionally 5% or 1% and denoted as α
- Threshold value – Proportion of false alarms that we are willing to tolerate
- P<0.05, it is significant, then reject the Null Hypothesis
- P>0.05, it is not significant, then accept the Null Hypothesis

Note: Various tests or methods are there to calculate p value

TEST OF SIGNIFICANCE

- Methods of inference used to support or reject claims based on sample data
- A good scientific practice, Significance level is usually 0.05
- Setting up and testing hypothesis is an essential part of statistical inference
- Hypothesis has 2 options – it will happen or not
- Two type of Hypothesis
  - Null Hypothesis [H0]
  - Alternate Hypothesis [H1]
- Based on P-value, can analysis whether it is significant or not

Hypothesis

- In statistical hypothesis testing, two hypotheses are compared
- Null hypothesis states that there is no relation between the Variables
- Alternative hypothesis states that there is some kind of relation
- Hypotheses are not treated on an equal basis, special consideration is given to the null hypothesis.

Example:

An experiment is done, to reject or accept the null hypothesis:

- H0: there is no difference in taste between coke and diet coke against H1: there is a difference.
- If the p-value is <0.05, then it is significant, states we have got strong evidence against the null hypothesis, then reject the null hypothesis: no difference in taste between coke flavours.
- If the p-value is >0.05, then it is not significant, accept the null hypothesis

Z score

90%	95%	99%
1.645	1.96	2.576

CORRELATION

- Mutual relationship or connection between two variables
- Strength of relationship, how variable is depended on another
- If correlation is same, it will be 100% which is equal to 1
- If co-efficient of correlation (r):
  - r>0 -- positive relationship
  - r<0 -- negative relationship
  - r=0 -- No relationship
- Disadvantages:
  - Only 2 variable can compare
  - Cannot conclude or form a model based on correlation