

Crack CSIR-NET 2021

(Study Notes on Data Analysis)



Data Analysis

Introduction:

It is the process by which useful information can be discovered through inspecting data. It is used in different business, social science domains and science. Nowadays, it plays a very important role in making decisions in a more scientific way. In other words, it is the process through which data is organized to draw conclusions. This analysis uses analytical as well as logical reasoning to gain information from data.

Mean and standard deviation:

Mean: It is an average or in other words, collection of numbers. It is also known by the name of expected value. It is the sum of all values divided by the number of numbers in collection. Its formula is:

$$\text{mean} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Standard deviation: It measures the amount of variation for a given set of values. When a standard deviation is low, it indicates the values are close to mean while high standard deviation indicates the values are spread over a wider range. Its formula is:

$$S = \sqrt{\frac{\sum(x_i - \bar{x})^2}{(n-1)}}$$

Absolute and relative error:

Absolute error: It is the difference between actual and measured value. If x is actual value and x_0 is the measured, then absolute error will be calculated as:

$$\Delta x = x_0 - x$$

Here, Δx is called an absolute error.

Relative error: It is the ratio of absolute error to actual value. If x , x_0 and Δx are actual, measured, and absolute quantities respectively, then relative error can be calculated as:

$$\text{Relative error} = (x_0 - x) / x = (\Delta x) / x$$

Covariance and correlation coefficient:

Covariance: It basically measures the relationship between two random variables. It can have both positive and negative value.

- Positive covariance indicates that two variables will move in the same direction.
- Negative covariance indicates that two variables will move in opposite directions.

Its formula is:

$$\text{Cov}(X, Y) = \frac{\sum((X - \mu)E(Y - v))}{n - 1}$$

Here,

- X is a random variable.
- $E(X) = \mu$ the mean of the random variable X.
- $E(Y) = v$ is the mean of the random variable Y.
- n = the number of items in the data set.

Correlation coefficient: It measures the strength of the linear relationship between two variables. The range of values is between -1.0 and 1.0. Its formula is:

$$r = \frac{\sum[(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{\sum(x_i - \bar{x})^2 * \sum(y_i - \bar{y})^2}}$$

Gradeup CSIR-NET Super Subscription

Features:

1. Memory Based Test Series of the actual exam paper.
2. All the CSIR NET Test Series based on the latest pattern and the trend that is followed.
3. Detailed performance analysis based on All India Rank after the completion of the test.
4. Mock Test are available in Hindi & English
5. Available on Mobile and Desktop

Gradeup Super Subscription, Enroll Now