



A small circular icon representing a user profile, followed by the text 'rishabh-mishra' and a downward-pointing arrow indicating a dropdown menu.

Course > Course 4: AI Statistics - Python > Module 2: Descriptive Statistics > Reading: Define and Demonstrate Correlation

Reading: Define and Demonstrate Correlation

Bookmark this page

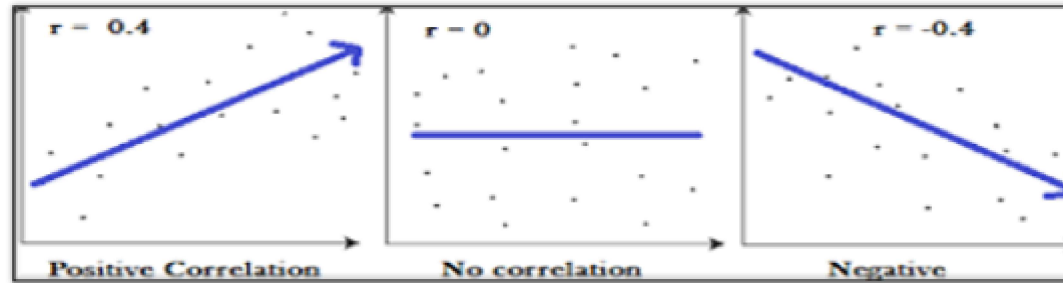
After completing this topic, you will have knowledge about:

1. Statistical Correlation
2. Statistical Dependence
3. Correlation and Causality

1. Statistical Correlation:

Correlation is a statistical measure that expresses the extent to which two variables are linearly related (meaning they change together at a constant rate). How is the correlation measured? The sample correlation coefficient, r , quantifies the strength of the relationship. Correlations are also tested for statistical significance.

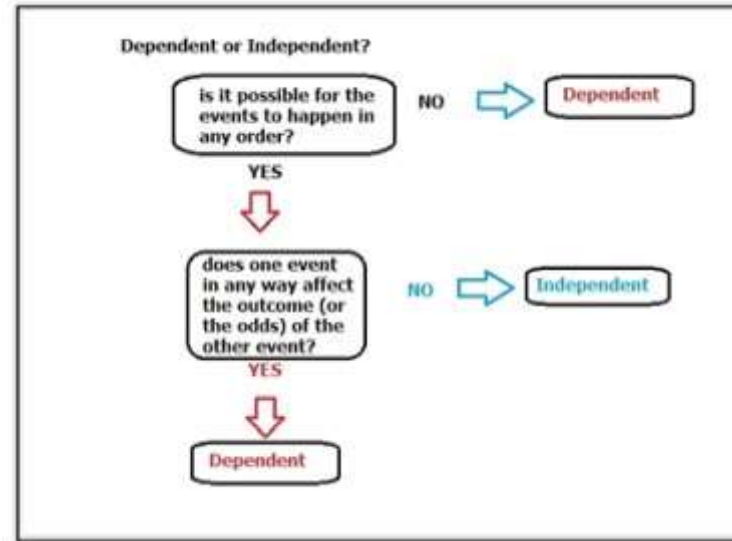
Correlation:



- Pearson correlation is the one most commonly used in statistics. This measures the strength and direction of a linear relationship between two variables.
- Values always range between -1 (strong negative relationship) and +1 (strong positive relationship). Values at or close to zero imply weak or no linear relationship.
- Correlation coefficient values less than +0.8 or greater than -0.8 are not considered significant.

2. Statistical Dependence:

A condition in which two random variables are not independent. X and Y are positively dependent if the conditional probability, $P(X | Y)$, of X given Y is greater than the probability, $P(X)$, of X, or equivalently if $P(X \& Y) > P(X) \cdot P(Y)$. They are negatively dependent if the inequalities are reversed.



3. Correlation and Causality

Two or more variables considered to be related, in a statistical context, if their values change so that as the value of one variable increases or decreases so does the value of the other variable (although it may be in the opposite direction).

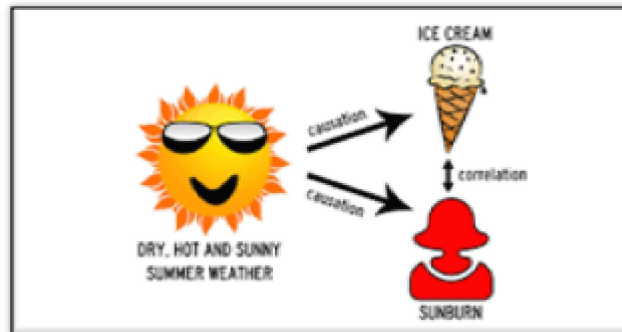
Correlation is a statistical measure (expressed as a number) that describes the size and direction of a relationship between two or more variables. A correlation between variables, however, does not automatically mean that the change in one variable is the cause of the change in the values of the other variable.

Causation indicates that one event is the result of the occurrence of the other event; i.e. there is a causal relationship between the two events. This is also referred to as cause and effect.

How can causation be established?

Causality is the area of statistics that is commonly misunderstood and misused by people in the mistaken belief that because the data shows a correlation that there is necessarily an underlying causal relationship.

The use of a controlled study is the most effective way of establishing causality between variables. In a controlled study, the sample or population is split in two, with both groups being comparable in almost every way. The two groups then receive different treatments, and the outcomes of each group are assessed.



In today's modern age of disruption, SkillUp Online is your ideal learning platform that enables you to upskill to the most in-demand technology skills like Data Science, Big Data, Artificial Intelligence, Cloud, Front-End Development, DevOps & many more. In your journey of evolution as a technologist, SkillUp Online helps you work smarter, get to your career goals faster and create an exciting technology led future.

Corporate

- ▶ Home
- ▶ About Us

- ▶ Enterprise
- ▶ Blog
- ▶ Press

Support

- ▶ Contact us
- ▶ Terms of Service
- ▶ Privacy Policy

Copyright ©2020 [Skillup](#). All Rights Reserved