Agri Articles
(e-Magazine for Agricultural Articles)
Volume: 04, Issue: 04 (JULY-AUG, 2024)
Available online at http://www.agriarticles.com
©Agri Articles, ISSN: 2582-9882

Understanding Correlations: Unveiling Connections in Data

(Harsha Sable, Vandana Gupta, Akanksha Singh, *Ruchi Nirmal, Mohan Singh Thakur and Asit Jain)
College of Veterinary Science and Animal Husbandry,
NDVSU, Jabalpur, Madhya Pradesh (India)
*Corresponding Author's email: ruchinirmal3@gmail.com

Abstract
Correlation is a methodology in statistics measuring the strength and direction of a direct relationship between two nonstop variables. This composition prospects correlation, exploring significance and types in multitudinous fields. When dealing with typically distributed data that have bivariate distribution, the Pearson correlation measure is employed. A Spearman rank correlation can be used as a measure of a monotonic association for nonstop data that isn't typically distributed, for ordinal data, or for data that contains meaningful outliers. We'll consider how to estimate correlation portions, produce smatter plots, and interpret the results. Knowing correlation facilitates the capability to make educated judgements and prognostications in variety of fields, including social lores, finance, wisdom and scientific exploration. This composition brings forth comprehensive overview of correlation, pressing its significance and applicability & operation in colorful fields.

Keywords: correlation coefficient, Karl Pearson's, variable, Rank method, Scatter plot.

Introduction
Correlation quantifies strength of the linear relationship between a pair of variables. Example, weight gain and feed given to an animal, using correlation and regression to arbitrate whether there is a relationship between feed intake (increase or decrease) and weight gain. In the sphere of statistics and data analysis, correlations play pivotal role in revealing relationships between variables. This article explores the concept of correlations, their types, significance, and implications.

What is Correlation?
Correlation deals with the strength and direction of the relationship among two variables. It directs how changes in one variable are related with changes in another.

Types of Correlation
Based on direction of change in variable

A. Positive Correlation- A positive correlation occurs when the variables move in the same direction. As one variable increases, the other variable also increases. For example, the relationship between eating junk food and gaining weight.

B. Negative Correlation- A negative correlation occurs when the variables move in opposite directions. As one variable increases, the other variable decreases. For example, the relationship between litter size and birth weight.

C. No Correlation-A zero correlation occurs when there is no relationship between the variables. For example, relationship between doing homework and gaining weight. There is no relationship between these variables.

Based on number of variables studied

Simple, Multiple and Partial Correlation:

A. Simple correlation When only two variables involved in correlation study. Of the two series, one modifying other is called independent or subject series and the other impacted is called dependent series.

B. Multiple correlation When the correlation involves three or more variables/data series itis **called multiple correlation**. one is dependent series, the others independent series.

C. Partial correlation When the correlation exists between two or more variables &the correlation is considered between any two of them keeping the other series constant, such a correlation is called **partial correlation.**

Based on proportion of change in variable:

Linear and Non-Linear: When relation between two set of variables is perfect degree of one. It states that two variables have an exact functional relationship correlation is linear. If two such variables are plotted on O-X and O-Y axis of graph paper, it will result in straight line graph. Opposing it, curve line will appear if change rates are not fixed on graph plotted appears called as **non-linear or curvilinear.**

How to collect correlational data

Primary data: Surveys, observations.

Secondary data: are the most often used data gathering techniques in the social and behavioral sciences**.**

How to analyze correlation data

Scatter Plot-. When two variables are connected to each other in any manner, a correlation between them is present. Starting with a scatterplot is the best option. A graph of the paired (x, y) sample data with a horizontal x-axis and a vertical y-axis is called a scatterplot, also known as a scatter diagram. Every single pair (x, y) is represented as a single point on the plot. The strength of the linear relationship between two variables is indicated by how close the spots on a scatter plot are to a straight line

Karl Pearson's correlation coefficient- Also called as product moment formula. It measures interdependence between 2 variables or it measures how the two variable covary Denoted by 'r'. where $\bar{x}$ is the mean of the x values, and $\bar{y}$ is the mean of the y values.

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Rank method- Rank methods are used to measure the correlation between two variables when the data is not normally distributed. The most common rank method is **Spearman's rank correlation coefficient.** It is based on estimated value and actual value.

$$r = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

Correlation graph: The correlation graph approach, in which every dataset attribute is represented as a graph vertex. The correlation between the qualities is shown by the edges connecting the vertices**.**

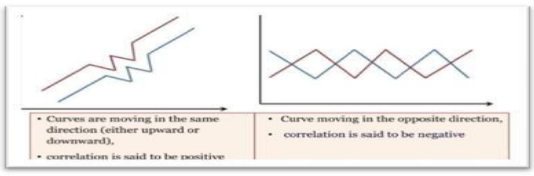

Image courtesy: www.slideshare.net

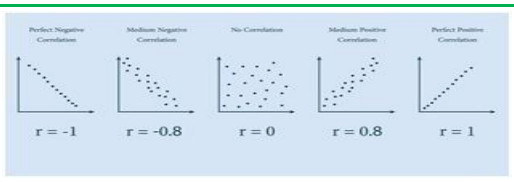

Image courtesy: www.shutterstock.com

Degree of Correlation

The coefficient of correlation, which has a value between + 1 and -1, expresses the degree of correlation. The prospect that the relationship between the two variables is equal to zero (null hypotheses; no relationship) is shown by the correlation coefficient's p-value. Since there is little chance that two variables are unrelated, strong correlations have low p-values.

Applications of correlation

In Agriculture: Role of quality of soil, fertilizers applied, environmental temperature and water in the crop yield. Farmers want to optimize their fertilizer application rates to maximize crop growth. They measure the levels of key soil nutrients (e.g. nitrogen, phosphorus, potassium) and correlate these with the amount of fertilizer applied per hectare. A strong positive correlation would suggest that increasing fertilizer leads to higher soil nutrient levels, guiding the farmers on appropriate application rates.

Milk Yield and Cow Health Indicators: Dairy farmers monitor various health indicators in their herd, such as body condition score, somatic cell count, and reproductive performance. They can calculate the correlation between these indicators and the individual milk yield of each cow. This helps identify which health factors have the strongest relationship with milk production, allowing the farmers to focus on the most important areas for herd management**.**

Stocks in the Same Sector: There is sometimes a positive connection between stocks in the same sector, such as technology or healthcare. For example, if a technology company climbs, it is probable that other stocks in the same industry would too. This is due to the fact that comparable market and economic developments have an impact on these equities. Market indices: There is a favorable link between individual equities and market indices. A stock that has a strong correlation with the S&P 500 index, for instance, will often move in the same direction as the index.

In veterinary science: Indiscriminate use of antimicrobials and development of resistance against pathogen in different animals. Use of vitamin C in wound healing hastens the wound healing by promoting tissue repair, minimizing the inflammation and protection against infection. Also, correlation between Oxidative Stress Markers and Periodontal Disease in Dogs, correlation between age and sexual behavioral can be evaluated.

Using negative correlation, investors employ hedging to control risk. For e, g, a stock in the gold mining sector often has a negative correlation with the broader market, meaning it tends to rise when the market falls.

Conclusion

To sum up, correlations are effective instruments for data analysis since they highlight relationships and trends that influence choices made by researchers in many fields. Even if they provide insightful information, it's important to evaluate them cautiously, taking potential limitations and contextual considerations into account. While data-driven methods advance, the ability to comprehend correlations is still essential for drawing insightful conclusions from large, complicated datasets. When it is neither feasible or desired to modify the independent variable of interest experimentally, correlational research is used to evaluate research hypotheses. It is also preferable since it makes it possible to examine behavior in contexts that occur normally.

References

1. Bhandari, P. (2023, June 22). Correlational Research When & How to Use. Scribbr. From https://www.scribbr.com/methodology/correlational-research/
2. Wubante, M. (2020). Review on correlation research. International Journal of English Literature and Culture, 8(4), 99-106.

3. https://journals.lww.com/picp/fulltext/2016/07040/common_pitfalls_in_statistical_analysis__the_use.9.aspx
4. https://milnepublishing.geneseo.edu/natural-resources-biometrics/chapter/chapter-7-correlation-and-simple-linear-regression/
5. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC374386/
6. https://www.westga.edu/academics/research/vrc/assets/docs/scatterplots_and_correlation_notes.pdf