UTM
UNIVERSITI TEKNOLOGI MALAYSIA

Online Learning

Quantitative Research Design : Correlational Research

Adibah Abdul Latif | Aqeel Khan | Hee Jee Mei| Mahani Mokhtar | Narina Abu Samah | Zainudin Hassan

Correlational Research - Purpose

- The **purpose of correlational research** is to determine the nature of relations among variables or to use these relations to make predictions.
- Correlational studies often examine **numerous variables** believed to be related to complex variables (e.g., achievement).
  - Unrelated variables are discarded from future studies while those related may be examined further through causal-comparative or experimental studies.

Correlational Research

- High correlations among variables do not imply causation (e.g., self-concept and achievement).
- Correlational procedures are also used to examine reliability and validity.

Correlational Research Process

- Problem selection
  - Correlational studies are designed to explore whether and how variables are related.
  - Correlational studies are designed to test hypotheses regarding expected relations among variables.

Correlational Research Process

- Participant and instrument selection
  - Samples are derived from acceptable sampling methods.
  - Sample must include at least 30 participants.
  - When reliability and validity of instrumentation is lower, sample size must be larger.

Correlational Research Process

- Design and procedure
  - Correlational studies share a simple design. Scores for two or more variables of interest are obtained for each member of the sample and these scores are then correlated (e.g., self-concept and achievement).

Correlational Research Process

Correlation coefficients

- Correlation coefficients range from –1 to 1.
- A correlation of 0 indicates no relationship.
- Correlation coefficients between +.35 and -.35 represent a weak relationship or no relationship.

Correlational Research Process

Correlation coefficients

- Correlation coefficients between +.35 and +.65 or between -.35 and -.65 represent moderate relationships.
- Correlation coefficients between .65 and 1.0 or between -.65 and –1.0 represent strong relationships.

Relationship Studies

- The purpose of relationship studies is to gain insight into variables or factors that are related to a complex variable (e.g., retention, academic achievement).
- Relationship studies help researchers to determine which variables may be suitable for future research.

Relationship Studies

- Relationship studies provide insight into which variables should be controlled for in future causal-comparative or experimental studies.

Relationship Studies

- Data collection
  - Researchers first identify variables to be correlated.
  - Variables should be purposely identified.
  - A smaller number of carefully identified variables is preferable to a larger number (e.g., shotgun approach).

Relationship Studies

- Data collection
  - After identifying variables the researcher next identifies the appropriate population and sampling procedure to select participants for the study.
  - In some relationship studies data are collected all at one time or in several sessions conducted in close succession.

Relationship Studies

- Data analysis and interpretation
  - In relationship studies scores from one variable are correlated with scores for another variable; or scores for several variables are correlated with a particular variable of interest.
  - The result is a single correlation coefficient or a number of correlation coefficients.
  - The method of calculating a correlation coefficient depends upon the nature of the data.

Relationship Studies

- Data analysis and interpretation
  - Correlation coefficients
    - The Pearson *r* coefficient is the most common and most precise coefficient. Pearson r is used for continuous variables.
    - The Spearman rho coefficient is appropriate to use when one of the variables are represented by rank-order data.

Relationship Studies

- Data analysis and interpretation
  - Correlation coefficients
    - The phi coefficient is used when both variables are expressed as a categorical dichotomy.
    - Other correlation coefficients are appropriate given characteristics of the data collected, sample size, and underlying data distribution.
      - e.g., Kendall's tau, Biserial, Point biserial, Tetrachoric, Intraclass, eta.

Relationship Studies

- Data analysis and interpretation
  - Several factors may contribute to inaccurate estimates of relation among variables.
    - Underlying relationships that are curvilinear will effect coefficients.
    - Attenuation (mengurang) occurs when measures have low reliability and may provide inaccurate correlation coefficients.
    - Restricted range in scores generally leads to underestimates of relations.

Prediction Studies

- When two variables are highly related, scores on one variable can be used to predict scores on the other variable.
  - The variable used to predict is called the *predictor*.
  - The variable that is predicted is called the *criterion*.
- A prediction study is used to determine which variables are the most highly correlated with a criterion variable.
  - More than one variable can be used to make predictions.

Prediction Studies

- Data collection
  - In prediction studies all measures should be valid measures.
    - It is especially critical that the criterion variable be validly measured.
  - In prediction studies, sometimes the predictor variables are administered prior to the criterion variable (e.g., SAT and university GPA).
    - Attrition is a problem in some prediction studies.

Prediction Studies

- Data collection
  - Shrinkage, or the tendency to find less accuracy in predicting criterion variables in subsequent samples, is often noted in prediction studies.
  - Cross-validation is the process of conducting subsequent prediction studies with new samples to verify effects found in an initial prediction study.

Prediction Studies

- Data analysis and interpretation
  - In prediction studies, data analysis involves correlating predictor variables with the criterion variable.
    - The single variable prediction equation is:
      - $Y = a+bX$

    Y= Predicted criterion score for an individual

    X= An individual's score on the predictor variable

    a= A constant calculated from scores of all participants

    b= A coefficient that indicates the contribution of the predictor variable to the criterion variable

Single Variable Predictions

- For example, that we wanted to predict a student's college GPA using high school GPA. We know that the student's high school grade average (x) is 3.0, the coefficient (b) is 0.87, and the constant (a) is 0.15. The student's predicted score would be calculated as follows:

  $Y = a + bX$

  $Y = 0.15 + 0.87\ (3)$

  $Y = 2.76$ (predicted college GPA)

Prediction Studies

- Data analysis and interpretation
  - Multiple regression is used when a combination of variables is used to predict a criterion variable (e.g., Success in Algebra may be predicted by prior knowledge, prior achievement, aptitude, etc.)
  - Intervening variables may lower prediction accuracy (e.g., teacher).
  - The amount of common variance shared by predictors is the squared correlation of the predictors and the criterion and is referred to as the *coefficient of determination*.

Other Correlation-Based Analyses

- In discriminate function analysis, continuous variables are used to predict a categorical variable.
- Canonical analysis produces a correlation based upon a group of predictor variables and a group of criterion variables.
- Path analysis provides a diagram that illustrates how variables are related to one another.

Other Correlation-Based Analyses

- Structural equation modeling, or LISREL, extends path analysis and predicts relations among variables with added precision.
- Factor analysis is used to decrease the number of variables under consideration by grouping variables into clusters called factors.

Interpreting Correlation Coefficients

- The following considerations can assist researchers when interpreting correlation coefficients.
  - Was the proper correlation method used?
  - Do the variables have high reliabilities? Low reliabilities lower the chance of finding significant relations.
  - Is the validity of the variables strong? Invalid variables produce meaningless results.

Interpreting Correlation Coefficients

- Is the range of scores to be correlated restricted or extended? Narrow or restricted score ranges lower correlation coefficients, whereas broad or extended score ranges raise them.
- How large is the sample? The larger the sample the smaller the value needed to reach statistical significance. Large samples may yield correlations that are statistically significant but practically unimportant.

Online Learning

UTM
UNIVERSITI TEKNOLOGI MALAYSIA

Thank you