

1. About Multivariate Methods

In most studies there are one or more outcome (or response) variables and several explanatory variables together with a variety of variables, which add a characteristic particularity to the situation. Thus, for example, in a study of infant feeding and growth in children, growth is the outcome variable and the method of feeding the explanatory variable. All additional information like mother's age, her social class, education, income number of siblings and so on, provide the particularity of each mother infant pair.

The t test and χ^2 test are the commonly employed statistical tests in biomedical research. They analyze the effects of single variable at a time, and are part of Univariate and Bivariate methods of data analysis. Medical research is to a large extent aimed at determining associations between two variables. For example, we may wish to study the association between bottle feeding and diarrhoea. In real life, however, several factors act simultaneously. Some may influence the outcome whereas others may just be confounders. In order to obtain a correct assessment of what is happening one has to resort to methods, which would take into account several variables acting together. These are the methods of Multivariate Analysis. With multivariate analysis the researcher is able to study the effects of several variables acting simultaneously instead of singly.

In most studies the researcher is attempting to describe the relationship between an outcome variable and its several determinants. The purpose may be to measure the sensitivity of a response to its determinants, or to predict the response under varying conditions. For all such purposes data on a number of subject attributes are collected. If one then wishes to examine the interactions and relations between a large number of determinants and the response variable simultaneously, simple univariate methods would not do. They would provide only a one-sided or biased view.

Multivariate methods are based on multiple linear regression, which is a multiple variable approach using the principles of simple linear regression. Multivariate methods have the advantage of bringing in more information to bear on a specific outcome. They allow one to take into account the continuing relationship among several variables. This is especially valuable in observational studies where total control is never possible. In the case of intervention studies the researcher can exercise some control over intervention and yet total control as in the case of laboratory experiments is never possible.

The specific advantages of multivariate studies are as follows:

1. They resemble closely how the researcher thinks about the data.
2. They allow easier visualization and interpretation of the data.
3. More data can be analyzed simultaneously, thereby providing greater statistical power.
4. Regression models can give more insight into relationships between variables.
5. The focus is on relationships among variables rather than on isolated individual factors.

The main methods of Multivariate analysis are set out in the table below.

Method	Type of Dependent Variable	Type of Independent variable	Purpose
Multiple Regression Analysis	Continuous (numerical)	Mostly continuous but in practice categorical can be used	To describe the extent, direction, and strength of the relationship between several independent variables and a continuous dependent variable
Logistic Regression Analysis	Categorical dichotomous	Continuous and categorical	To describe how many times more likely is the event in one group compared to the other
Analysis of variance	Continuous	All nominal	To describe the relationship between a continuous dependent variable and one or more nominal independent variables
Discriminant Analysis	Nominal (polychotomous)	Commonly all continuous, but in practice a mixture of various types can be used as long as some are continuous	To determine how one or more independent variables can be used to discriminate among different categories of a nominal dependent variable
Factor analysis	Commonly continuous, but in practice may be of any type. The variables are not initially identified as dependent or independent, but the resulting factors may be used as dependent or independent variables in a later analysis.	Commonly continuous, but in practice may be of any type. The variables are not initially identified as dependent or independent, but the resulting factors may be used as dependent or independent variables in a later analysis.	To define one or more new composite variables called factors.
Analysis of covariance	Continuous	Mixture of nominal and continuous variables. The continuous variables are used as control variables	To describe the relationship between a continuous dependent and one or more nominal independent variables, controlling for the effect of one or more variables

