

Choosing the correct test

The following is a guide to help you think about your study and the type of analysis you should be using.

Independent and dependent variables

An *independent* variable, sometimes called the experimental or predictor variable, is a variable that is being manipulated in an experiment in order to observe the effect on a *dependent* variable, sometimes called an outcome variable. In observational studies, the independent variable (often called an *explanatory* variable) is not manipulated, but observed. The independent variable is hypothesized to contribute to the development of the outcome of interest.

What is the nature of your independent variable?

Statistical analyses assume that the variables have a specific levels of measurement. Consider gender, you cannot analyze average gender! It does not make sense. What we are saying is that calculating an average of a categorical variable provides you with a nonsensical result. Therefore it's important to know what type of variable you have so you can select the appropriate analysis for your data.

Categorical variables

A categorical variable (sometimes called a nominal variable) is one that has two or more categories, with no intrinsic ordering to the categories. Consider gender for example, we can split gender into 2 categories, males and females. There is, however, no intrinsic ordering in the categories (i.e. males or females can come first). This is also true for things like eye colour, race, political party etc. The key is that there is no way to order these from highest to lowest, they are purely categories *without* order. If, however, we can 'order' the categories then we call the variable and ordinal variable.

Ordinal variables

An ordinal variable is similar to a categorical variable, with the key difference being that there is a clear ordering of the variables. Consider socioeconomic status grouped into 3 categories; low, medium and high. Individuals can be classified into 1 of the 3 categories (as the case for categorical variables) but we can now *order* the categories. It would may no sense to order these 3 categories as *medium low high*. It is far more reasonable to order them as *low medium high*. Similarly, if we consider a variable like education; primary school, high school, and university education. These also can be ordered as primary school is the lowest level of education, followed by high school, and then university (bachelor). Note that even though we are ordering these variables from lowest to highest, they may not be evenly spaced (primary school is 7 years, high school is 5 year and university may range between 3-4 years); this is Ok because they are grouped variables. If the categories were equally spaced, then the variable would be considered an interval variable.

Interval variables

In this example, the intervals or spaces between variables are equally spaced. Lets consider income brackets for \$30,000-\$35,000, \$45,000-\$50,000 and \$60,000-\$65,000. There is a \$10,000 interval *between* each income bracket.

Continuous variables

A continuous variable has a changing value that can vary from one sample to the next having an infinite number of possible values. Variables like weight, age, concentration of drug, measurement of gene/protein expression are all examples of a continuous variable.

Choosing the correct test

Number of dependent variables	Nature of independent variables (IV)	Nature of dependent variable	Test(s)
1	0 IVs (1 population)	continuous	one-sample t-test
		ordinal or interval	one-sample median
		categorical (2 categories)	binomial test
		categorical	Chi-square goodness-of-fit
	1 IV with 2 levels (independent groups)	continuous	2 independent sample t-test
		ordinal or interval	Wilcoxon-Mann Whitney test
		categorical	Chi-square test
			Fisher's exact test
	1 IV with 2 or more levels (independent groups)	continuous	one-way ANOVA
		ordinal or interval	Kruskal Wallis
		categorical	Chi-square test
	1 IV with 2 levels (dependent/matched groups)	continuous	paired t-test
		ordinal or interval	Wilcoxon signed ranks test
		categorical	McNemar
	1 IV with 2 or more levels (dependent/matched groups)	continuous	one-way repeated measures ANOVA
		ordinal or interval	Friedman test
		categorical	repeated measures logistic regression
	2 or more IVs (independent groups)	continuous	factorial ANOVA
		ordinal or interval	ordered logistic regression
		categorical	factorial logistic regression
	1 interval IV	continuous	correlation
			simple linear regression
		ordinal or interval	non-parametric correlation
	categorical	simple logistic regression	
1 or more interval Ivs and/or 1 or more categorical Ivs	continuous	multiple regression	
		analysis of covariance	
	categorical	multiple logistic regression	
		discriminant analysis	
2 or more	1 IV with 2 or more levels (independent groups)	continuous	one-way MANOVA
	2 or more	continuous	multivariate multiple linear regression
2 sets of 2 or more	0	continuous	canonical correlation
2 or more		continuous	factor analysis