What statistical analysis should I use?

The following table shows general guidelines for choosing a statistical analysis. We emphasize that these are general guidelines and should not be construed as hard and fast rules. Usually your data could be analyzed in multiple ways, each of which could yield legitimate answers. The table below covers a number of common analyses and helps you choose among them based on the number of dependent variables (sometimes referred to as outcome variables), the nature of your independent variables (sometimes referred to as predictors). You also want to consider the nature of your dependent variable, namely whether it is an interval variable, ordinal or categorical variable, and whether it is (approximately) normally distributed (see What is the difference between categorical, ordinal and interval variables? for more information on this).[1] The table then shows one or more statistical tests commonly used given these types of variables (but not necessarily the only type of test that could be used) and links showing how to do such tests using SAS, Stata and SPSS.

Number of Dependent Variables	Nature of Independent Variables	Nature of Dependent Variable(s)	Test(s)	How to SAS	How to Stata	How to SPSS	How to R
1	0 IVs (1 population)	interval & normal	one-sample t-test	SAS	Stata	SPSS	R
		ordinal or interval	one-sample median	SAS	Stata	SPSS	R
		categorical (2 categories)	binomial test	SAS	Stata	SPSS	R
		categorical	Chi-square goodness-of-fit	SAS	Stata	SPSS	R
	1 IV with 2 levels (independent groups)	interval & normal	2 independent sample t-test	SAS	Stata	SPSS	R
		ordinal or interval	Wilcoxon-Mann Whitney test	SAS	Stata	SPSS	R
		categorical	Chi-square test	SAS	Stata	SPSS	R
			Fisher's exact test	SAS	Stata	SPSS	R
	1 IV with 2 or more levels (independent groups)	interval & normal	one-way ANOVA	SAS	Stata	SPSS	R
		ordinal or interval	Kruskal Wallis	SAS	Stata	SPSS	R
		categorical	Chi-square test	SAS	Stata	SPSS	R
	1 IV with 2 levels (dependent/matched groups)	interval & normal	paired t-test	SAS	Stata	SPSS	R
		ordinal or interval	Wilcoxon signed ranks test	SAS	Stata	SPSS	R
		categorical	McNemar	SAS	Stata	SPSS	R
	1 IV with 2 or more levels (dependent/matched groups)	interval & normal	one-way repeated measures ANOVA	SAS	Stata	SPSS	R
		ordinal or interval	Friedman test	SAS	Stata	SPSS	R
		categorical	repeated measures logistic regression	SAS	Stata	SPSS	R

	2 or more IVs (independent groups)	interval & normal	factorial ANOVA	SAS	Stata	SPSS	R
		ordinal or interval	ordered logistic regression	SAS	Stata	SPSS	R
		categorical	factorial logistic regression	SAS	Stata	SPSS	R
	1 interval IV	interval & normal	correlation	SAS	Stata	SPSS	R
		interval & normal	simple linear regression	SAS	Stata	SPSS	R
		ordinal or interval	non-parametric correlation	SAS	Stata	SPSS	R
		categorical	simple logistic regression	SAS	Stata	SPSS	R
	1 or more interval IVs and/or 1 or more categorical IVs	**interval & normal**	multiple regression	SAS	Stata	SPSS	R
			analysis of covariance	SAS	Stata	SPSS	R
		categorical	multiple logistic regression	SAS	Stata	SPSS	R
			discriminant analysis	SAS	Stata	SPSS	R
2+	**1 IV with 2 or more levels (independent groups)**	interval & normal	one-way MANOVA	SAS	Stata	SPSS	R
	2+	interval & normal	multivariate multiple linear regression	SAS	Stata	SPSS	R
	0	interval & normal	factor analysis	SAS	Stata	SPSS	R
2 sets of 2+	**0**	interval & normal	canonical correlation	SAS	Stata	SPSS	R
Number of Dependent Variables	**Nature of Independent Variables**	**Nature of Dependent Variable(s)**	**Test(s)**	**How to SAS**	**How to Stata**	**How to SPSS**	**How to R**

This page was adapted from Choosing the Correct Statistic developed by James D. Leeper, Ph.D. We thank Professor Leeper for permission to adapt and distribute this page from our site.

[1]Technically, it is the residuals from these analyses that should be normally distributed; it does not matter if the dependent variable is normally distributed or not.

How to cite this page Report an error on this page or leave a comment