Choosing a Statistical Test Depending on Type of Variable and Study Design

In the following table, the "input variable" is the independent variable or factor; "output variable" is the dependent or response variable. Cells outlined in red are most common cases. χ^2=Chi-square test.

A statistics primer: https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one

Table & notes from BMJ: "13. Study design and choosing a statistical test"
https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/13-study-design-and-choosing-statisti
Accessed 7/14/18

Choice of statistical test for independent observations

		Outcome variable (= Dependent variable)					
		Nominal (2 Categories: Presence/ Absence)	Categorical (>2 Categories)	Ordinal (=Ranked)	Quantitative Discrete	Quantitative Non-Normal (Continuous variable)	Quantitative Normal (Continuous variable)
Input Variable (= Independent variable)	Nominal	χ^2 or Fisher's	χ^2	χ^2 trend or Mann-Whitney	Mann-Whitney	Mann-Whitney or log-rank (a)	Student's t test
	Categorical (>2 categories)	χ^2	χ^2	Kruskal-Wallis (b)	Kruskal-Wallis (b)	Kruskal-Wallis (b)	Analysis of variance (c)
	Ordinal (Ordered categories)	χ^2 -trend or Mann-Whitney	(e)	Spearman rank	Spearman rank	Spearman rank	Spearman rank or linear regression (d)
	Quantitative Discrete	Logistic regression	(e)	(e)	Spearman rank	Spearman rank	Spearman rank or linear regression (d)
	Quantitative non-Normal	Logistic regression	(e)	(e)	(e)	Plot data and Pearson or Spearman rank	Plot data and Pearson or Spearman rank and linear regression
	Quantitative Normal	Logistic regression	(e)	(e)	(e)	Linear regression (d)	Pearson and linear regression

Plot data to identify method to transform data to Normal, then use linear regression, otherwise use rank method

(a) If data are censored [not a usual case, see https://en.wikipedia.org/wiki/Censoring_(statistics)]
(b) The Kruskal-Wallis test is used for comparing ordinal or non-Normal variables for more than two groups, and is a generalisation of the Mann-Whitney U test. The technique is beyond the scope of this book, but is described in more advanced books and is available in common software (Epi-Info, Minitab, SPSS).
(c) Analysis of variance (ANOVA) is a general technique, and one version (one-way analysis of variance) is used to compare Normally-distributed variables for more than two groups, and is the parametric equivalent of the Kruskal-Wallis test.
(d) If the outcome variable is the dependent variable, then provided the residuals are plausibly Normal, then the distribution of the independent variable is not important.
(e) There are a number of more advanced techniques, such as Poisson regression, for dealing with these situations. However, they require certain assumptions and it is often easier to either dichotomise the outcome variable or treat it as continuous.