

Basic Statistical Tests

Data Characteristics that help determine statistical test used:

Type(s) of data used

Two groups versus more than 2 groups

Whether groups are matched ("paired") or unmatched

Whether groups are small or large

Whether the data are normally distributed (continuous data)

Independent variable (IV) versus Dependent variable (DV)

IV - varied by and under control of the experimenter

DV - variable that responds to "experimental" manipulation

* frequently researcher must rely on natural variation of both variables and look for a relationship between the two. Example: smoking and lung cancer

Parametric versus Non-parametric statistics

Parametric - the dependent variable is a continuous variable, so it makes sense to calculate a mean and SD

Non-parametric - the DV is a count (nominal data) or a ranking (ordinal data) so it doesn't make sense to measure a means i.e "The average gender of Americans is 1.25."

Parametric:

Comparison of Two Means

Student t test - Example: Days to recovery with 2 different treatments.

Paired t test - Used when you have repeated measures on the same individual
Example: Effectiveness of a diet, weigh patients pre and post diet and compare see if there is a significant difference. This takes away the effect of individual differences between subjects, resulting in an increase in the precision of the measurement of the difference.

ANalysis Of VAriance (ANOVA) - comparison among many means

Factorial ANOVA - allows inclusion of multiple factors in a single study and evaluation of each factor without committing the cardinal sin of distorting the overall probability of a chance difference.

ANOVA with Repeated measures - an extension of the paired t-test

Correlation - measures the degree of association. Does not make judgments as to whether one affects or predicts the other. Example: maternal smoking is associated with the weight of babies

Regression Analysis - Relationship b/w 2 continuous variables

Simple - predicts one to predict the dependent variable, given the independent variable. $Y = \text{slope of the line (X)} + b$ (intercept of the line with the Y axis).
Example: babies weight = m (degree of maternal smoking) + b .

Multiple - multiple IVs and their relationship to DV. Example: Peak bilirubin levels, days on ventilator, and birth weight (IVs) on IQ (DV).

Non-Parametric:

Chi Square (X^2) - uses nominal data. Checks to see if there are significant differences between/among groups compared to what would be expected.
Example - compare a new therapy for BPD. Our DV would be "+ BPD" versus "no BPD" and our IV would be "new therapy" versus "conventional therapy."

Logistic Regression - Use if **DV is dichotomous** and **categorical** (ordinal or nominal), regardless of whether the IV is continuous, ordinal, or nominal.
Example: Does maternal age (IV - continuous) and/or prior use of birth control pills (IV - nominal) have an effect on ability to get pregnant (DV - nominal).