Introducción al Plan de Análisis de un Protocolo.

Introducción

- http://www.youtube.com/watch?v=sCxeUfroo-U

Objetivos de la Sesión

- Identificar cómo se redacta un plan de análisis en un protocolo de investigación

Secciones

- Manejo de los datos. Software utilizado.
- Análisis descriptivo
- Análisis bivariado
- Análisis de múltiples variables

- Ver STROBE statement
- https://www.strobe-statement.org/fileadmin/Strobe/uploads/checklists/STROBE_checklist_v4_combined.pdf
- Esto no fue diseñado para evaluar protocolos. Sin embargo, tiene varias secciones que pueden ser pertinentes.

Recordar

- Tipos de Variables:
  - Cuantitativas (p. ej. Edad en años, Peso en Kg, PAS en mmHg)
  - Categóricas (p. ej. Sexo M/F, Estado Civil S/C/V/D/Co, Grado de Enfermedad L/M/S)

Altman DG. Practical statistics for medical research. London: Chapman & Hall, 1991.

Análisis descriptivo

- Para variables categóricas, se utilizan %s o frecuencias relativas

- Para variables cuantitativas, se utilizan medidas de tendencia central (media, mediana) y de dispersión (D.E., IQR)

Análisis bivariado

- Var Categ vs. Var Categ: Chi cuadrado o Fisher
- Var Cuant vs Var Categ:
  - Var Categ (2 categorías): T de Student o U Mann Whitney
  - Var Categ (>= 3 categorías): ANOVA o Kruskal Wallis
- Var Cuant vs. Var Cuant: Correlación de Pearson o Spearman

Altman DG. Practical statistics for medical research. London: Chapman & Hall, 1991.

Flow Chart for Selecting Commonly Used Statistical Tests

Parametric Assumptions:
1. Independent, unbiased samples
2. Data normally distributed
3. Equal variances

Type of data?

Continuous → **Type of question**

Discrete, categorical → Chi-square tests one and two sample

Type of question

Relationships → Do you have a true independent variable?

Differences → Differences between what?

Do you have a true independent variable?

Yes → Regression Analyses

No → Correlation Analysis

Correlation Analysis

Parametric → Pearson's r

Nonparametric → Spearman's Rank Correlation

Differences between what?

Means →

Tests for Equal Variances → F_{max} test, Brown and Smythe's test, Bartlett's tests

How many treatment groups?

Two groups → Parametric assumptions satisfied?

More than two groups → Parametric assumptions satisfied?

Parametric assumptions satisfied? (Two groups)

Yes → Parametric → Student's unpaired t-test, Paired t-test

No → Nonparametric → Mann-Whitney U or Wilcoxon Rank sums test

Parametric assumptions satisfied? (More than two groups)

Yes → Parametric → **ANOVA**

No → Nonparametric → Kruskal-Wallis Test

ANOVA → If significant, do a *post hoc* test, e.g. Tukey's or Bonferroni's

Kruskal-Wallis Test → If significant, do a Dunn's Test

http://abacus.bates.edu/~ganderso/biology/resources/statistics.html

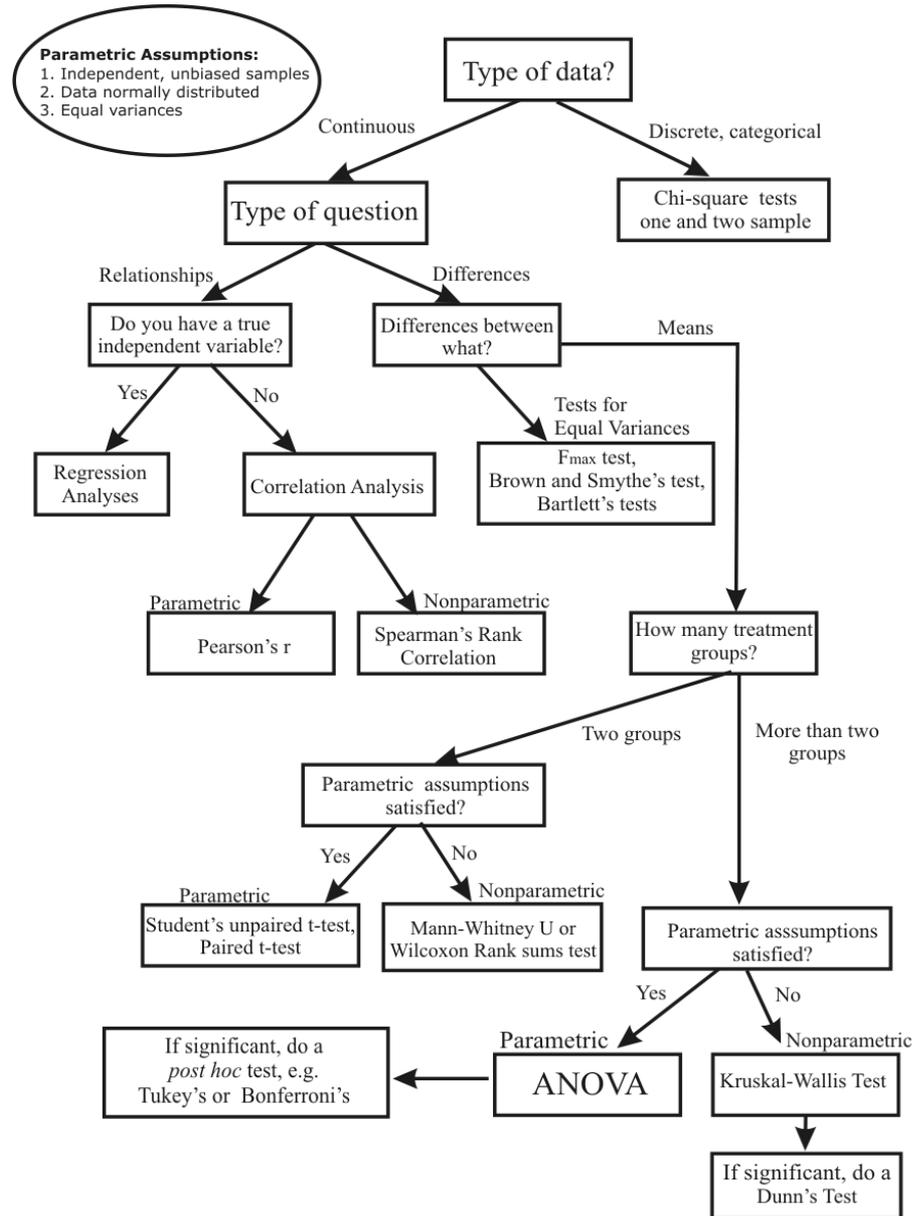**Table 1** Some commonly used statistical tests

Parametric test	Example of equivalent non-parametric test	Purpose of test	Example
Two sample (unpaired) t test	Mann-Whitney U test	Compares two independent samples drawn from the same population	To compare girls' heights with boys' heights
One sample (paired) t test	Wilcoxon matched pairs test	Compares two sets of observations on a single sample	To compare weight of infants before and after a feed
One way analysis of variance (F test) using total sum of squares	Kruskall-Wallis analysis of variance by ranks	Effectively, a generalisation of the paired t or Wilcoxon matched pairs test where three or more sets of observations are made on a single sample	To determine whether plasma glucose level is higher one hour, two hours, or three hours after a meal
Two way analysis of variance	Two way analysis of variance by ranks	As above, but tests the influence (and interaction) of two different covariates	In the above example, to determine if the results differ in male and female subjects
χ^2 test	Fisher's exact test	Tests the null hypothesis that the distribution of a discontinuous variable is the same in two (or more) independent samples	To assess whether acceptance into medical school is more likely if the applicant was born in Britain
Product moment correlation coefficient (Pearson's r)	Spearman's rank correlation coefficient (r_s)	Assesses the strength of the straight line association between two continuous variables.	To assess whether and to what extent plasma HbA$_1$ concentration is related to plasma triglyceride concentration in diabetic patients
Regression by least squares method	Non-parametric regression (various tests)	Describes the numerical relation between two quantitative variables, allowing one value to be predicted from the other	To see how peak expiratory flow rate varies with height
Multiple regression by least squares method	Non-parametric regression (various tests)	Describes the numerical relation between a dependent variable and several predictor variables (covariates)	To determine whether and to what extent a person's age, body fat, and sodium intake determine their blood pressure

Greenhald, 1998

A. Múltiples variables

- Desenlace cuantitativo: Regresión Múltiple
- Desenlace categórico binario: Regresión Logística; Regresión Log Binomial, Regresión de Poisson
- Desenlace tiempo hasta evento: Regresión de Cox

Para consolidar lo aprendido

- Greenhaigh T. How to read a paper. Statistics for the non-statistician I: Different types of data need different statistical tests. *BMJ* 1997; 215(8):364-8.

- ¿Preguntas?