PROCESSES OF CLINICAL DATA ANALYSIS.

PRSENTED BY:

ONYIAPAT, JANE

@ GLOBAL RESEARCH NURSES' WORKSHOP.

OCTOBER 12, 2018

INTRODUCTION

- A necessary companion to a well designed clinical trial is its appropriate statistical analysis.
- Assuming that a clinical trial will produce data that could reveal differences in effects between two or more interventions,
- statistical analyses are used to determine whether such differences are real or are due to chance.

INTRODUCTION

- To analyze clinical data, researchers rely on tried and tested statistical methods,

- which have to be specified in a filling with the regulatory authorities before the trial even begins.

- The main statistical analysis should be clearly described in distinction from secondary statistical analyses.

INTRODUCTION

- In general, the main statistical analysis should be made for each of the categories assessed in clinical research

- which includes efficacy, safety, and usefulness evaluations.

Objectives of the presentation: to:

- explain the concept of data analysis
- discuss approaches to clinical research analysis
- discuss basic tests used in clinical research analysis

Data Analysis

- is a process of inspecting, cleaning, transforming, and modeling data with the goal of discovering useful information, suggesting conclusions, and supporting decision-making.

Data analysis:

Involves a rigorous scientific method for gaining insight into data.

Give an instant overall picture of data irrespective of the number of data points.

Besides data summarization, it provides information to make inference and predict relations of variables.

I hate statistics, why should I learn data analysis?

- Critique health related literature
- Apply study results to patient care
- Interpret vital statistics
- Understand epidemiologic problems
- Critique study protocols
- Participate in or direct research projects
- Become an entrepreneur

Onyiapat, 2018.

Approaches to clinical research analysis:

- Sequential analysis
- Hierarchical models, Bayesian analysis
- Decision analysis, statistical prediction
- Meta-analysis and other alternatives,

Sequen-tial analysis :

Analysis of data as they accumulate, with a view toward stopping the study as soon as the results become statistically compelling.

Data are analyzed as the result for each participant are obtained.

Sequential analysis:

Decision is made after each observation to: continue the study

by enrolling additional participants, stop the study & conclude

that there is a statistically significant difference between treatments or

there is not a statistically significant difference b/w the interventions.

Hierarchical models:

Provide a natural framework for combining information from a series of clinical trials.

Useful in longitudinal clinical trial, in which patients are randomly assigned to different treatments and are repeatedly evaluated over the course of the study.

bayesian analysis (Subjec-tivity):

conditional probability of a set of possible causes for a given observed event can be computed from knowledge of the probability of each cause & the conditional probability of the outcome of each cause

Bayesian analysis:

Before designing a study or collecting any data, the investigator acquires all

information about the activities of both the experimental & control treatments,

formulates a subjective opinion which is influenced by data collected.

combination of observed data & prior opinion provides an automatic update of the investigator's subjective opinion.

Decision analysis:

modelling technique that systematically considers all possible

management options for a problem. It uses probabilities and utilities to

explicitly define decisions. allow one to evaluate the importance of any

variable in decision-making process.

Statistical prediction:

The prediction problem involves construction of a limit or interval that

will contain one or more new measurements drawn from that same

distribution with a given level of confidence

Statistical prediction:

to determine whether a single new measurement obtained from onsite location is consistent with background measurements obtained from offsite.

if the new measurement lies within the interval, then one can conclude that the measurement from onsite is consistent .

Meta-analysis:

synthesis of results of independent studies. Used to summarize empirical research in the literature.

degree to which results of a synthesis can be generalized depends in part on the nature of the set of studies.

meta-analysis:

- In general, meta-analysis serves as a useful tool to answer questions for which single trials were underpowered or not designed to address.
- Sometimes refers to the entire process of synthesizing the results of independent studies, including the collection of studies, coding, abstracting, and so on, as well as the statistical analysis.

Basic clinical analysis:

- question of whether to use one- or two-sided test,
- when considering use of t-test whether to use separate or pooled version & what about testing for equality of variance first?

one-sided and two-sided tests:

- Tests are usually two-sided unless there are very good prior reasons, not observation or data based.
- The temptation to use a one-sided test is that it is more powerful for a given significance level

- Good editors of medical journal would almost certainly refuse to publish articles based on use of one-sided test.

Separate and pooled variance t-tests:

- 2 sample t-test looks for difference in sample means by an estimate of the standard deviation of that difference.

separate and pooled variance t-tests:

- The first is by estimating the variance of each sample separately and then combine the two
- The other is to pool all the data from the two samples and estimate a common variance.

- i) t = 

separate and pooled variance t-tests:

- ii)

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

Separate & pooled variance:

- Separate variance is primarily used because if the underlying populations variances are indeed the same,

- then the separate variance estimates is a good unbiased estimate of the common variance.

Separate & pooled variance:

- However, if you use the pooled variance estimate when the underlying population variances are unequal,

- then the resulting test statistic could be wrong neither generally conservative nor liberal, neither generally more nor less powerful, just incorrect.

Test of equality of variances

- It is natural to consider conducting a preliminary test of equality of variances and then on the basis of the outcome, decide whether to use a pooled or a separate variance estimate.

- The arguments against using such a preliminary test are; tests of equality of variance are very low powered without large quantities of data,

Test of equality of variances

- if the form of the t-test is chosen on the basis of a preliminary test using the same data then allowance needs to be made for the conditioning of the t-test distribution on the preliminary test.

Sampling distributions:

- When using a sample to estimate population characteristics, it is important to obtain a sample that is representative,
- and random sampling is the best means of securing such samples.
- Inferential statistics are based on the assumption of random sampling from populations.

Hypothesis testing:

- Statistical hypothesis testing provides objective criteria for deciding whether research hypotheses should be accepted as true or rejected as false.

- Suppose we hypothesized that maternity patients exposed to a teaching film on breastfeeding would breastfeed longer than mothers who did not see the film.

Hypothesis testing:

- We find that the mean number of days of breastfeeding is 131.5 for 25 experimental subjects and 125.1 for 25 control subjects.

- Should we conclude that the hypothesis has been supported? True, group differences are in the predicted direction, but perhaps in another sample the group means would be nearly identical.

Hypothesis testing:

- The observed outcome may have occurred due to the following explanations:

- The film is truly effective in encouraging breastfeeding.

- The difference was due to chance factors.
- Hypothesis testing will make objective decision about the study result.

Type I and type II errors:

- An investigator makes a type I error by rejecting the null hypothesis when it is, in fact, true.

- If we conclude that the film was effective in promoting breast feeding, when in fact, group differences were due to sampling error – we would have made a type 1 error (false positive conclusion).

Type I and type II errors:

- Acceptance of a false null hypothesis is called a type II error- a false negative conclusion

Level of significance:

- Researchers control the degree of risk in making a type I error by selecting a level of significance.

- Level of significance is the term used to signify the probability of making a type I error.

Level of significance:

- The two most frequently used levels of significance (referred to as alpha) are .05 and .01.

- With .05 sig level, we accept the risk that out of 100 samples, a true null hypothesis would be wrongly rejected 5 times.

Type 11 error (beta):

- Researchers can reduce the risk of a type 11 error by simply increasing their sample size.
- This can be estimated through power analysis.

Parametric and nonparametric tests:

- Parametric tests have assumption that the variables are normally distributed in the population.

- Nonparametric tests, by contrast, do not estimate parameters and involve less restrictive assumptions about the shape of the distribution of the critical variables.

Hypothesis testing procedures:

- Selecting an appropriate test statistic: Researchers select a test based on such factors as whether a parametric test is justified, which levels of measurement were used, and how many groups are being compared.

Hypothesis testing procedures:

- Selecting the level of significance: An alpha level of 0.5 is usually chosen, but sometimes the level is set more stringently at 0.1.

- Computing a test statistic: Researchers then calculate a test statistic based on the collected data.

hypothesis testing procedures:

- Determining degrees of freedom: The term degrees of freedom (df) is used throughout hypothesis testing to refer to the number of observations free to vary about a parameter.

- The concept is too complex for full elaboration here, but computing degrees of freedom is easy.

hypothesis testing procedures:

- Comparing the test statistic to a tabled value:

- Theoretical distributions have been developed for all the test statistics, and values for these distributions are available in tables for specified degrees of freedom and level of significance.

hypothesis testing procedures:

- The tabled value enables researchers to determine whether the computed value of the statistics is beyond what is probable if the null hypothesis is true.

hypothesis testing procedures:

- If the absolute value of the computed statistic is larger than the table value, the results are statistically significant; ($p < .05$)

- if the computed value is smaller, the results are nonsignificant; ($p > .05$)

Bivariate statistical tests:

- T-tests (mean of 2 groups)
- Analysis of variance (3 or more groups)
- Chi-squared test (contingency table)*
- Correlation coefficients (relationship)
- * non parametric test

Multivariate statistical analysis (3 or more):

- Multiple regression (> 1 independent variable to predict dependent variable)
- Analysis of covariance (ANOVA + MR)
- Discriminant function analysis (predicts group membership eg smoker vs non smoker)

Multivariate statistical analysis (3 or more):

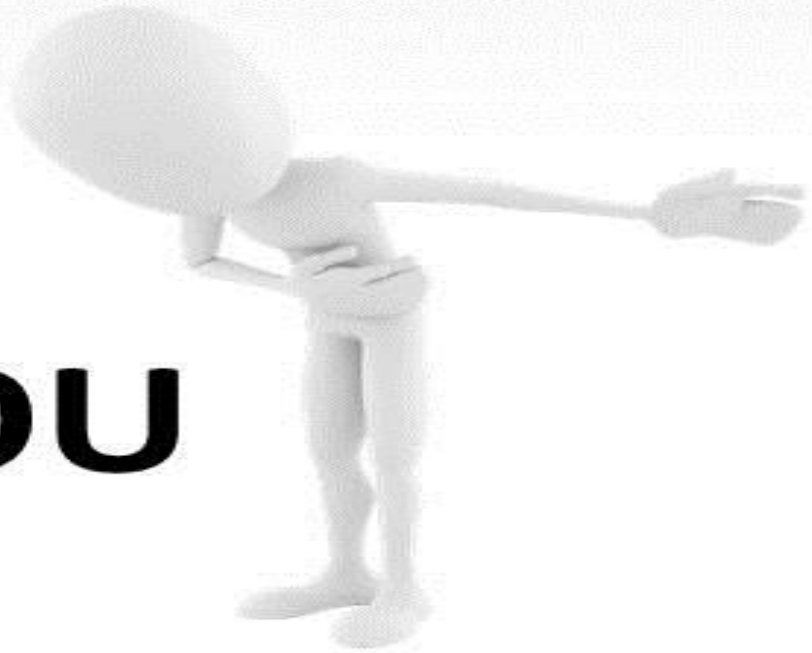
- Logistic regression (multiple independent variable & nominal level dependent variable)
- Factor analysis (develop, refine or validate complex instrument)
- Multivariate analysis of variance (MANOVA) (extension of ANOVA to > 1 dependent variable)

Wisdom.....

- Failures or limitations of the past have no control over the greatness in you.
- To him that believes, all things are possible.
- Don't wait for everybody to believe in your ability and dreams.

Power Resources

THANK YOU



YOUR LOGO