Selecting Statistical Procedures

The choice of a univariate statistical test or methodology depends on several key factors.

The following points need to be considered:

1) The research question - whether we are interested in testing hypotheses about population differences or assessing the association between variables;
2) the level of measurement of the dependent variable - whether measurement is at the nominal, ordinal, or interval/ratio level;
3) the type of experimental or quasi-experimental design - whether it is a one, two, or more-than-two group design;
4) the sample type - whether the sample(s) arise from independent simple random sampling, paired sampling, or mixed sampling schemes.

We are most familiar with the parametric procedures, e.g., t-test, analysis of variance (ANOVA), analysis of covariance (ANCOVA), regression, and correlation. These methods require measurement at the interval/ratio level and make several assumptions about the nature of the population parameters of the underlying distributions. For example, in applying a two-sample t-test, one assumes that the population variances for the two groups are the same, i.e., equal within-group variances. If this assumption is not tenable, then one may transform the data to homogenize the variances or use a non-parametric statistical procedure that does not make this assumption, e.g., the rank sum test.

The choice of methodology is further complicated by the variety of procedures available. I know of no single reference that succinctly summarizes this process and facilitates test selection. Therefore, several years ago I undertook to develop a decision tree that would enable one to choose procedures. The decision tree that follows is the result of that effort. I have continued to revise and update it and have found it to be useful in my teaching and statistical consulting activities.

The top portion of the tree is concerned with detecting differences. For variables measured at the nominal level, interest is in testing for differences in proportions. For variables measured at the ordinal level, interest lies with detecting location (e.g. median) differences. For those variables measured at the interval or ratio level we are interested in detecting differences in population means.

The bottom portion of the tree is concerned with assessing associations or testing hypotheses about independence/dependence of variables.

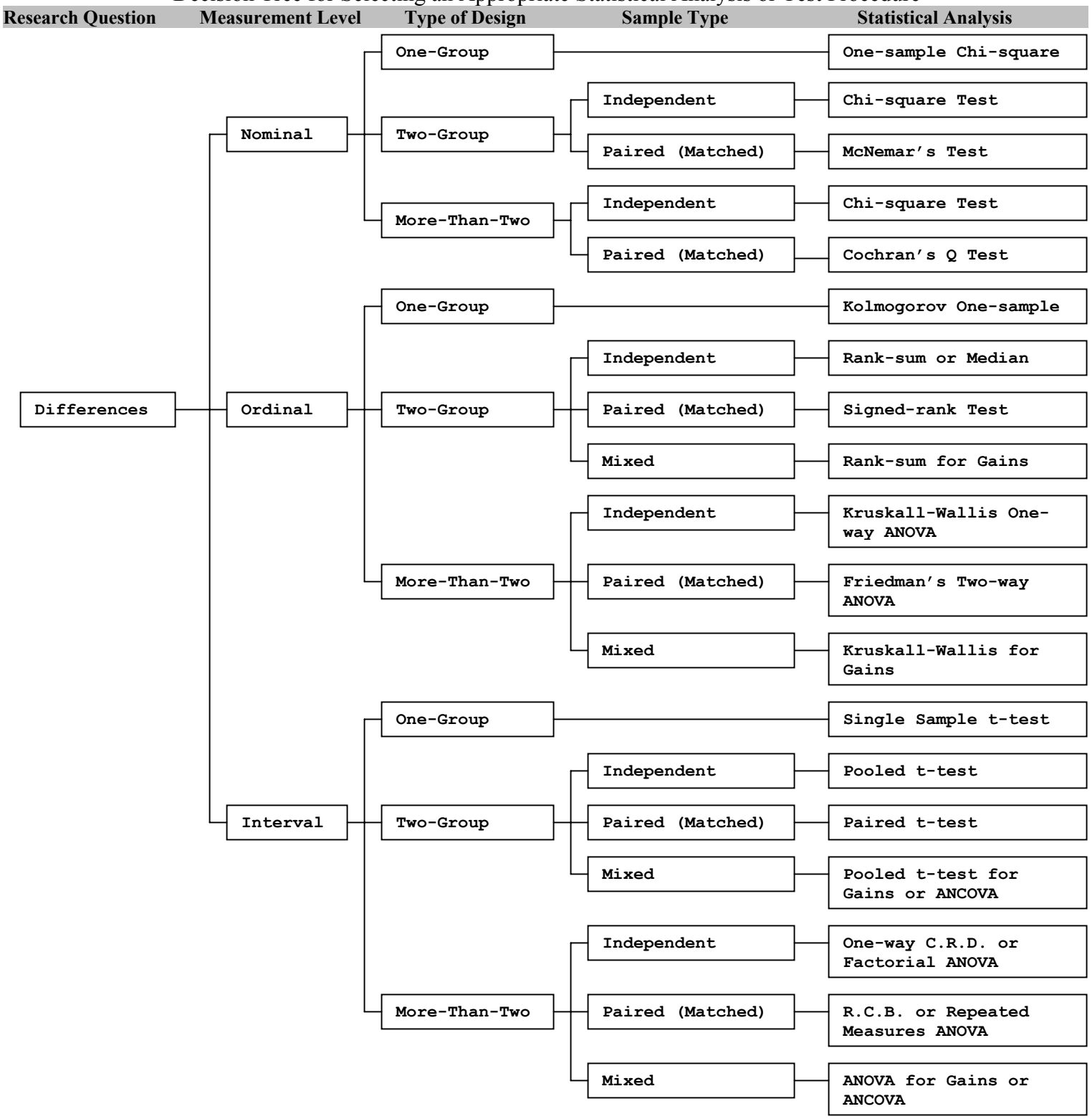
It should be pointed out that level of measurement refers to the level of measurement of the dependent or criterion variable under study. Also, mixed sample-type refers to samples that have independent groups where within each group there are repeated (correlated) measures on the dependent variable.

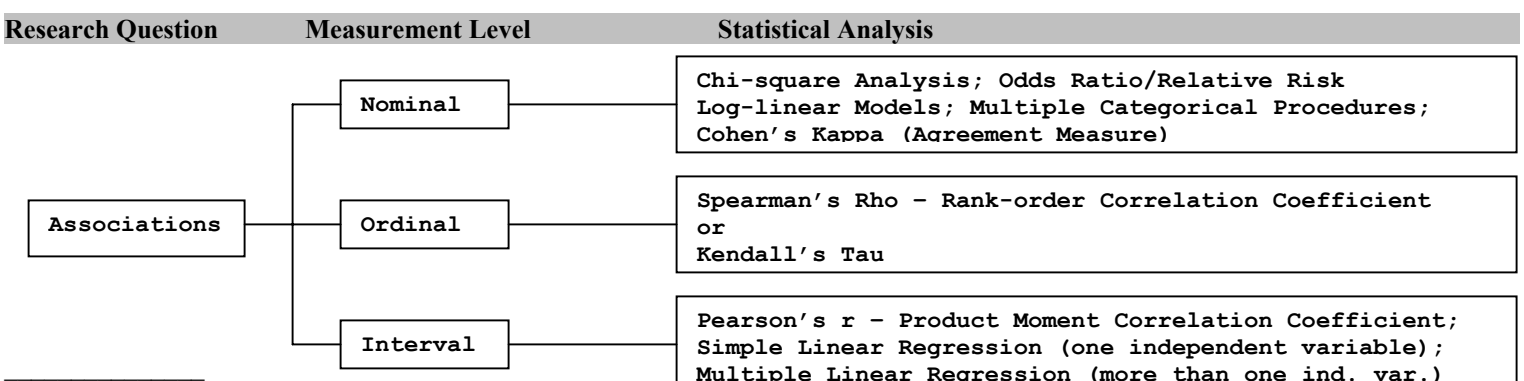
One way to analyze mixed samples is to form gain (or difference) scores within the groups and then apply independent sample procedures to the gain (or difference) scores. Alternately, if measurement is at least at the interval level, one can use ANCOVA procedures with one of the repeat measures as a covariate.

The statistical analysis column (the final branch of the tree) suggests one or two procedures appropriate to the situation. There are usually several methods to choose from that are both appropriate and well known. In suggesting a procedure, I have tried to balance my decision by suggesting simple, easy-to-use methods that are the most powerful procedures applicable to the situation. I have also considered the nature of assumptions implicit in a statistical procedure.

I hope that the decision tree will prove useful to you in your research activities.

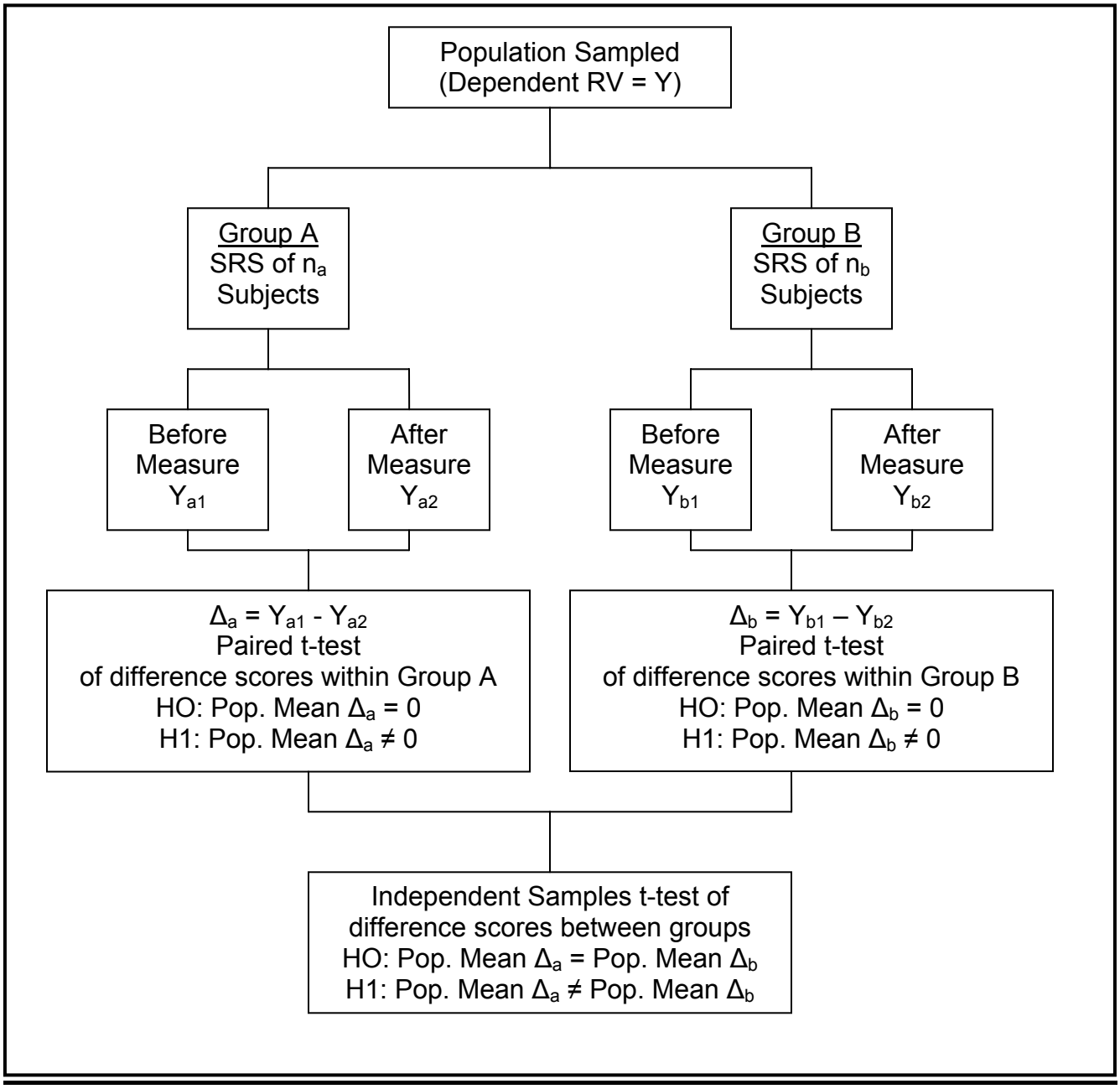
Decision Tree for Selecting an Appropriate Statistical Analysis or Test Procedure





Decision Tree.pdf © James K. Dias, Ph.D., 2001

Two-Group Mixed-Sample Longitudinal RCT



Note: Each subject has a before and after measure. Therefore, we have paired/correlated observations within groups. The between-group difference scores are independent. The difference score can be calculated as either a “gain” or “loss” depending on ones definition of “improvement”.