**Purpose**

The purpose of this list is to provide reminders and required information when a particular statistical test is included in a preregistration. It is a “living document” that will grow as more examples come in and we need a place to remind ourselves of what to look out for.

**Overview**

For most of the following tests, there are three pieces of information:

1. The situation in which a test is used.
2. The parameters that must be specified in order to properly “remove researcher degrees of freedom.”
3. The assumptions of the test.

A preregistration is only eligible for the Challenge if #2 is complete, 1 and 3 are partially for our help, but also can be used for advice under two circumstances

1. A preregistration is rejected for other reasons, go ahead and point out statistical issues or
2. A preregistration is allowed because everything required is included, but you can recommend submitting a second preregistration, given the “error” that you are noticing. Of course, this quickly leads down the rabbit hole of simply “preregister a million tests, publish the significant findings” but registrations are eventually surfaced, and we do other places encourage multiple preregistrations.

We should not reject preregistrations if issues around 1 and 3 exist (because we are likely to miss some issues, and because we are not making decisions on whether the study is any good, just if it is “complete”).

**Tests**

* Sequential analyses
  + Specify:
    - the number of samples at which each checkpoint will take place
    - the alpha level used to test for significance at each checkpoint
* T-test
  + Need to specify type of t-test
    - Student’s/Independent samples/Between groups T-test
      * Situation: One continuous DV and two independent groups (i.e. two separate samples)
      * Specify: the number of tails (one or two tailed)
        + if one tailed, need to specify the direction of the hypothesis test
      * Assumptions:
        + normal data distribution,
        + equal variance in groups (homoscedasticity)
        + continuous DV
        + random sampling
    - Welch’s/unequal variance t-test
      * Situation: One continuous DV and two independent groups (i.e. two separate samples) - used when have or think will have unequal variance
      * Specify: the number of tails (one or two tailed)
        + if one tailed, need to specify the direction of the hypothesis test
      * Assumptions:
        + normal data distribution
        + DV is continuous
        + random sampling
    - Paired/within subjects/repeated measures T-test
      * Situation: One continuous DV, measured twice on the same sample
      * Specify: the number of tails (one or two tailed)
        + if one tailed, need to specify the direction of the hypothesis test
      * Assumptions:
        + differences between pairs are normally distributed.
        + DV is continuous
        + random sampling
    - One-sample t-test
      * Situation: One group and one continuous DV whose average value is being tested against a given constant
      * Specify: the mean against which the group mean will be tested against
      * Assumptions:
        + DV is normally distributed
        + DV is continuous
        + Random sampling
* ANOVA
  + Situation: One continuous DV and one or more categorical IVs which can be either between or within/repeated
  + Specify:
    - The number of factors and the number of levels per factor
    - Which factors are between and which are within/repeated
    - For completely within, or mixed between/within designs where any factor has more than 2 levels:
      * Will they use repeated measures or multivariate ANOVA procedure? (see MANOVA below if they are using a MANOVA)
    - Should specify what tests will follow a significant ANOVA, unless they want all follow-up tests to be exploratory
      * specify what tests will be conducted
        + for complex contrasts, what are the contrast codes
      * what alpha correction/false discovery rate procedure will be used?
  + Assumptions of between subjects ANOVA:
    - DV is normally distributed
    - Homogeneity of variance between cells
    - The overall ANOVA is generally robust to violations of these two assumptions IF there number of samples are equal among groups
      * The follow-up tests are NOT generally robust to these violations though
  + Assumptions of mixed/repeated measures ANOVAs
    - DV is normally distributed
    - Sphericity (only for when have more than 2 levels for repeated factors)
* MANOVA (Multivariate ANOVA)
  + Situation: At least one categorical IV and more than one continuous DV (can also be used for repeated measured designs with only one DV)
  + Specify:
    - The number of factors and the number of levels per factor
    - Which factors are between and which are within/repeated
    - The DVs that will be included
    - Which multivariate test criterion will they use? (there are 4 possible ones, in the case where all factors/DVs have 2 levels, they will all come out the same)
    - How are they going to follow-up on a significant overall results?
      * need to specify this unless they want all later analyses to be exploratory
      * What specific pairwise comparisons/complex contrasts are they going to perform?
      * What (if any) method will they use to correct for false positives/false discovery rate?
  + Assumptions
    - Data are distributed multivariate-normal
    - Homogeneity of covariance structure
* ANCOVA (Analysis of Covariance)
  + Situation: 1 continuous DV, 1 or more categorical IVs, and at least one continuous covariate parameter
  + Specify:
    - The number of factors and the number of levels per factor
    - Which factors are between and which are within/repeated
    - What are the covariates, and what are you trying to covary? (main effect or interaction)
  + Assumptions:
* Pearson product-moment/Pearson product/Pearson correlation
  + Situation: Measures the linear relationship between two variables
  + Specify:
    - one tailed or two tailed test?
      * if one-tailed, what is the direction?
  + Assumptions:
* Regression
  + Situation: 1 DV and one or more predictor variables
  + Specify:
    - ALL terms that are going into the model (so all variables, and combinations of variables, such as interactions and quadratic terms)
    - Which are the variables of interest vs. control variables (if applicable)
    - for categorical predictors, what coding scheme will be used (e.g. dummy coding, summation coding, etc.)
      * What will the reference category be for each categorical variable?
    - If there are interaction terms present, what exact follow-up test will be performed to break down the interactions?
      * if they do not specify this, all those test will be exploratory
  + Need to specify the type of regression
    - Linear/Multiple linear regression
      * Specific Situation: One continuous DV
      * Specify:
        + All general specifications above
      * Assumptions:
        + predictors measured without error
        + Normal/multivariate normal distributed errors
        + error variance is constant (homoscedasticity)
        + Independence of errors (i.e. errors uncorrelated)
        + lack of multicollinearity of predictors
        + predictors linearly related to DV
    - Logistic regression
      * Situation: One categorical DV
      * Need to specify the type of logic regression:
        + logistic regression

only 2 categories in DV

* + - * + Multinomial logistic regression

More than 2 categories in the DV

Specify:

What is the DV reference category?

* + - * + ordered logistic regression/proportional odds model/ordered logit

DV is ordinal

* + - Probit regression
      * Situation: One categorical DV
      * Need to specify type of probit regression:
        + Probit regression

only 2 categories in DV

* + - * + Multinomial probit regression

More than 2 categories in the DV

Specify:

What is the DV reference category?

* + - * + Ordered probit

DV is ordinal

* Non-parametric tests
  + Spearman/Spearman rank/Spearman rank-order/grade correlation
    - non-parametric test of the association between two variables
      * specifically, how well the relationship fits a monotonic function
  + Mann–Whitney *U* test/Mann–Whitney–Wilcoxon/Wilcoxon rank-sum test/Wilcoxon–Mann–Whitney test
    - Situation: 1 DV, 1 predictor with 2 levels
      * non-parametric equivalent of independent samples t-test
    - Specify:
      * one or two-tailed test (if one tailed, in which direction?)
    - Assumptions
      * DV is ordinal or continuous
      * all observations independent of each other
  + Wilcoxon signed-rank test
    - Situation: 1 DV, 1 predictor with 2 levels of paired/repeated measures data
      * non-parametric equivalent of dependent samples t-test
    - Specify:
      * one or two-tailed test (if one tailed, in which direction?)
    - Assumptions
      * DV is ordinal or continuous
  + Kruskal Wallis
    - Situation: 1 DV, 1 predictor with 2 or more levels
      * the non-parametric version of a one-way anova, used when DV is not normally distributed
    - Specify:
      * What follow-up tests will be conducted after a significant Kruskal Wallis (need to specify this unless they want all follow-up analyses to be exploratory)
        + what procedure (if any) will be used to keep false positive rate/false discovery rate in check
    - Assumptions:
      * Equal variance in all groups
      * DV is ordinal or continuous
* Pearson Chi-square/Goodness of Fit test
  + Situation: One or more variables where DV is categorical (non-ordered categorical), comparing observed frequency to expected frequence of categories
    - if have more than one variable, sometimes called a contingency table
  + Specify:
    - What is the expected frequencies being tested against?
    - If it’s larger than a 2x2 table, follow-up analyses likely
      * need to specify what they are unless all these will be exploratory
  + Assumptions:
    - observations independent of one another
    - inclusion of nonoccurences
* Fisher’s Exact Test
  + Situation: 2x2 contingency tables based on categorizing data along two different dimensions, used in place of the pearson’s chi-square when samples are small or when have very small predicted cell counts in some cells (i.e. less than 5 expected in any cell), but can be used with larger samples as well