Kelleher notes on Park & Feigenson paper on effect of powerpoint on juries - review of statistics principles to understand it

Monday, November 25, 2013
2:15 PM

http://en.wikipedia.org/wiki/F-test - F tests in statistics

http://ibgwww.colorado.edu/~carey/p7291dir/handouts/manova1.pdf: Instead of having an expectation around 1.0 (which F has under the null hypothesis), the
expectation under the alternative hypothesis will be sometihing greater than 1.0. Hence,
the larger the F statistic, the more likely that the null hypothesis is false.

F test - probability of hypothesis divided by probability of error - when F>1, then null hypothesis more likely to be false

he formula for the one-way **ANOVA** F-test statistic is

$$F = \frac{\text{explained variance}}{\text{unexplained variance}},$$

or

$$F = \frac{\text{between-group variability}}{\text{within-group variability}}.$$

Pasted from <http://en.wikipedia.org/wiki/F-test>

Note $F(x, y)$ denotes an F-distribution with x degrees of freedom in the numerator and y degrees of freedom in the denominator.

Pasted from <http://en.wikipedia.org/wiki/F-test>

In statistics, the number of degrees of freedom is the number of values in the final calculation of a statistic that are free to vary.[1]

The number of independent ways by which a dynamic system can move without violating any constraint imposed on it, is called degree of freedom. In other words, the degree of freedom can be defined as the minimum number of independent coordinates that can specify the position of the system completely.

Estimates of statistical parameters can be based upon different amounts of information or data. The number of independent pieces of information that go into the estimate of a parameter is called the degrees of freedom. In general, the degrees of freedom of an estimate of a parameter is equal to the number of independent scores that go into the estimate minus the number of parameters used as intermediate steps in the estimation of the parameter itself (which, in sample variance, is one, since the sample mean is the only intermediate step).[2]

Mathematically, degrees of freedom is the number of dimensions of the domain of a random vector, or essentially the number of 'free' components (how many components need to be known before the vector is fully determined).

The term is most often used in the context of linear models (linear regression, analysis of variance), where certain random vectors are constrained to lie in linear subspaces, and the number of degrees of freedom is the dimension of the subspace. The degrees of freedom are also commonly associated with the squared lengths (or "sum of squares" of the coordinates) of such vectors, and the parameters of chi-squared and other distributions that arise in associated statistical testing problems.

While introductory textbooks may introduce degrees of freedom as distribution parameters or through hypothesis testing, it is the underlying geometry that defines degrees of freedom, and is critical to a proper understanding of the concept. Walker (1940)[3] has stated this succinctly:

> For the person who is unfamiliar with *N*-dimensional geometry or who knows the contributions to modern sampling theory only from secondhand sources such as textbooks, this concept often seems almost mystical, with no practical meaning.

Pasted from <http://en.wikipedia.org/wiki/Degrees_of_freedom_(statistics)>

Eta-squared, η^2 - effect size measurement

Eta-squared, η^2[edit]

Eta-squared describes the ratio of variance explained in the dependent variable by a predictor while controlling for other predictors, making it analogous to the r^2. Eta-squared is a biased estimator of the variance explained by the model in the population (it estimates only the effect size in the sample). This estimate shares the weakness with r^2 that each additional variable will automatically increase the value of η^2. In addition, it measures the variance explained of the sample, not the population, meaning that it will always overestimate the effect size, although the bias grows smaller as the sample grows larger.

$$\eta^2 = \frac{SS_{\text{Treatment}}}{SS_{\text{Total}}}.$$

Pasted from <http://en.wikipedia.org/wiki/Effect_size>

Eta²

Eta² can be defined as the proportion of variance associated with or accounted for by each of the main effects, interactions, and error in an ANOVA study (see Tabachnick & Fidell, 2001, pp. 54-55, and Thompson, 2006, pp. 317-319). Formulaically, eta², or η^2, is defined as follows:

$$\eta^2 = \frac{SS_{effect}}{SS_{total}}$$

Where:

SS_{effect} = the sums of squares for whatever effect is of interest
SS_{total} = the total sums of squares for all effects, interactions, and errors in the ANOVA

Pasted from <http://jalt.org/test/bro_28.htm>

Eta² values are easy to calculate. Simply add up all the sums of squares (*SS*), the total of which is 64.24 in the example; then, divide the *SS* for each of the main effects, the interaction, and the error term by that total. The results will be as follows:

$$\eta^2_{Anxiety} = \frac{SS_{Anxiety}}{SS_{Total}} = \frac{0.08}{64.24} = 0.00124533 \approx 0.0012$$

$$\eta^2_{Tension} = \frac{SS_{Tension}}{SS_{Total}} = \frac{2.08}{64.24} = 0.03237858 \approx 0.0324$$

$$\eta^2_{AxT} = \frac{SS_{AxT}}{SS_{Total}} = \frac{18.75}{64.24} = 0.291874221 \approx 0.2919$$

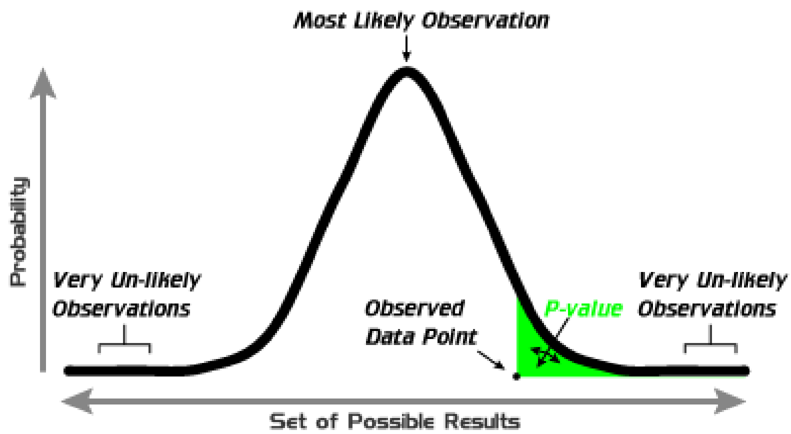
$$\eta^2_{Error} = \frac{SS_{Error}}{SS_{Total}} = \frac{43.33}{64.24} = 0.674501867 \approx 0.6745$$

Interpretation of these values is easiest if the decimal point is moved two places to the right in each case, the result of which can be interpreted as percentages of variance associated with each of the main effects, the interaction, and error. Starting with Anxiety, the value of 0.0012 indicates that a mere 0.12% of the variance is accounted for by Anxiety, whereas Tension accounts for 3.24%, the Anxiety x Tension (A x T) interaction accounts for a much larger 29.19%, and a whopping 67.45% is accounted for by Error. Now let's consider the A x T interaction and Error separately in more detail.

Pasted from <http://jalt.org/test/bro_28.htm>

Multivariate analysis of variance (MANOVA) is a statistical test procedure for comparing multivariate (population) means of several groups. Unlike ANOVA, it uses the variance-covariance between variables in testing the statistical significance of the mean differences. It is a generalized form of univariate analysis of variance (ANOVA). It is used when there are two or more dependent variables. It helps to answer : 1. do changes in the independent variable(s) have significant effects on the dependent variables; 2. what are the interactions among the dependent variables and 3. among the independent variables.[1] Statistical reports however will provide individual p-values for each dependent variable, indicating whether differences and interactions are statistically significant.

Pasted from <http://en.wikipedia.org/wiki/MANOVA>



A p-value (shaded green area) is the probability of an observed (or more extreme) result arising by chance

p-value

From Wikipedia, the free encyclopedia

This article **may be too technical for most readers to understand**. Please help improve this article to make it understandable to non-experts, without removing the technical details. The talk page may contain suggestions. *(February 2013)*

In statistical significance testing, the ***p*-value** is the probability of obtaining a test statistic at least as extreme as the one that was actually observed, assuming that the null hypothesis is true.[1] A researcher will often "reject the null hypothesis" when the *p*-value turns out to be less than a certain significance level, often 0.05[2][3] or 0.01. Such a result indicates that the observed result would be highly unlikely under the null hypothesis (that is, the observation is highly unlikely to be the result of random chance alone). Many common statistical tests, such as chi-squared tests or Student's t-test, produce test statistics which can be interpreted using *p*-values.

The *p*-value is a key concept in the approach of Ronald Fisher, where he uses it to measure the weight of the data against a specified hypothesis, and as a guideline to ignore data that does not reach a specified significance level. Fisher's approach does not involve any alternative hypothesis, which is instead a feature of the Neyman–Pearson approach.

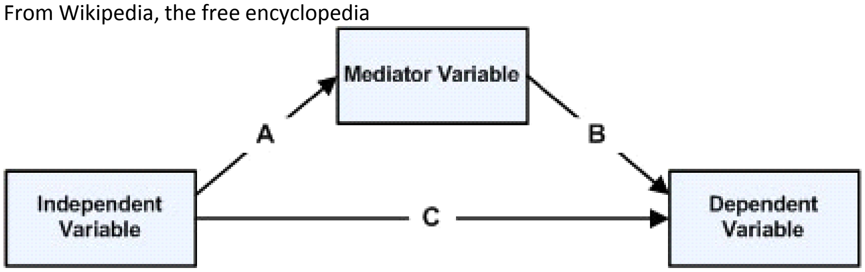
The *p*-value should not be confused with the Type I error rate [false positive rate] α in the Neyman–Pearson approach. Although α is also called a "significance level" and is often 0.05, these two "significance levels" have different meanings. Their parent approaches are incompatible, and the numbers p and α cannot meaningfully be compared. Fundamentally, the *p*-value does not in itself support reasoning about the probabilities of hypotheses, nor choosing between different hypotheses–it is simply a measure of how likely the data were to have occurred by chance, assuming the null hypothesis is true.

Statistical hypothesis tests making use of *p*-values are commonly used in many fields of science and social sciences, such as economics, psychology,[4] biology, criminal justice and criminology, and sociology.[5]

Depending on which style guide is applied, the "p" is styled either italic or not, capitalized or not, and hyphenated or not (*p*-value, *p* value, *P*-value, *P* value, p-value, p value, P-value, P value).

Pasted from <http://en.wikipedia.org/wiki/P-value>

Mediation (statistics)
From Wikipedia, the free encyclopedia



A simple statistical mediation model.
Instatistics, a **mediation** model is one that seeks to identify and explicate the mechanism or process that underlies an observed relationship between anindependent variable and a dependent variable via the inclusion of a third explanatory variable, known as a mediator variable. Rather than hypothesizing a direct causal relationship between the independent variable and the dependent variable, a mediational model hypothesizes that the independent variable influences the mediator variable, which in turn influences the dependent variable. Thus, the mediator variable serves to clarify the nature of the relationship between the independent and dependent variables.[1] In other words, mediating relationships occur when a third variable plays an important role in governing the relationship between the other two variables.

Pasted from <http://en.wikipedia.org/wiki/Mediation_(statistics)>

In statistics, **maximum-likelihood estimation** (**MLE**) is a method of estimating theparameters of a statistical model. When applied to a data set and given a statistical model, maximum-likelihood estimation provides estimates for the model's parameters.
The method of maximum likelihood corresponds to many well-known estimation methods in statistics. For example, one may be interested in the heights of adult female penguins, but be unable to measure the height of every single penguin in a population due to cost or time constraints. Assuming that the heights are normally (Gaussian) distributed with some unknown mean and variance, the mean and variance can be estimated with MLE while only knowing the heights of some sample of the overall population. MLE would accomplish this by taking the mean and variance as parameters and finding particular parametric values that make the observed results the most probable (given the model).
In general, for a fixed set of data and underlying statistical model, the method of maximum likelihood selects the set of values of the model parameters that maximizes the likelihood function. Intuitively, this maximizes the "agreement" of the selected model with the observed data, and for discrete random variables it indeed maximizes the probability of the observed data under the resulting distribution. Maximum-likelihood estimation gives a unified approach to estimation, which is well-defined in the case of the normal distribution and many other problems. However, in some complicated problems, difficulties do occur: in such problems, maximum-likelihood estimators are unsuitable or do not exist.[

Pasted from <http://en.wikipedia.org/wiki/Maximum_likelihood>

Main Effect
In the design of experiments and analysis of variance, a **main effect** is the effect of an independent variable on a dependent variable averaging across the levels of any other independent variables. The term is frequently used in the context of factorial designs and regression models to distinguish main effects from interaction effects.
Relative to a factorial design, under an analysis of variance, a main effect test will test the hypotheses

expected such as H0, the null hypothesis. Running a hypothesis for a main effect will test whether there is evidence of an effect of different treatments. However a main effect test is nonspecific and will not allow for a localization of specific mean pairwise comparisons (simple effects). A main effect test will merely look at whether overall there is something about a particular factor that is making a difference. In other words a test examining differences amongst the levels of a single factor (averaging over the other factor and/or factors). Main effects are essentially the overall effect of a factor.

Pasted from <http://en.wikipedia.org/wiki/Main_effect>