DRAFT 0.0

Glen Cowan
6 June, 2011

Some Properties of Limits

This note summarizes some properties of several limit procedures that have come under consideration for use in LHC searches. It is not in its current form intended to represent a consensus view, but rather only that of the author. Others are invited to suggest modifications. For now only frequentist limits are described. Bayesian limits (and their frequentist properties) will considered at a future point.

1 Introduction and prototype analysis

As a useful prototype analysis, we consider a measured quantity x is modeled as following a Gaussian distribution with mean μ and standard deviation σ. Taking parameter $\mu = 0$ corresponds to the null or "no-signal" or "background-only" model. Often we consider signal models where μ cannot be negative.

Rejecting the background-only model is the key step towards claiming discovery of new physics. In addtion to testing the $\mu = 0$ hypothesis, one may also want to test nonzero μ and see which values can be rejected. The largest (smallest) values not reject may be reported as upper (lower) limits on the parameter.

To carry out the procedure outlined above, one defines for each value of μ a statistical test. This can be done by taking a fixed significance level α and finding the corresponding critical region. If the data fall in the critical region, the parameter value is rejected. Those parameter values not rejected form a confindence interval with confidence level $1 - \alpha$.

Equivalently one may for every hypothesized parameter value specify a rule by which the data outcomes are ordered in terms of their level of compatibility with the hypothesis. The p-value of the hypothesized parameter is the probability, under assumption of that hypothesis, to find an outcome in the region of equal or lesser compatibility relative to what was observed in the real data. If the p-value of a hypothesized parameter value μ is found below α, then this corresoponds to finding the data in the critical region and μ is rejected. Thus the edges of the confidence interval are those values of the parameter for which the p-value is just at the threshold α.

2 Limits for different kinds of analyses

The test chosen as the basis for the confidence interval will depend in general on the nature of the analysis. We could consider, e.g., the following scenarios:

1. The signal is not yet established and physically restricted to $\mu > 0$. (Here "established" could be taken in a loose sense meaning "broadly recognized as a real process" or could correspond to a specific numerical criterion such as a 5σ discovery threshold.)

2. The signal is not yet established and there is no restriction on the sign of μ.

3. The signal is well established. That is, values of μ both above and below some range have been excluded, and in particular $\mu = 0$ has been rejected in convincing enough fashion that there is a general conensus that the signal exists.

4. The signal is not well established and the signal model is characterized by more than one parameter of interest, e.g., a strength parameter μ and a particle mass m_H, or the case of neutrino oscillations with a mass difference squared Δm^2 and a mixing parameter $\sin^2 2\theta$.

3 Criteria to consider when constructing limits

When constructing a limit, one may wish to consider the follow criteria:

1. Coverage. The probability for the confidence interval to include the true parameter value should be greater than or equal to the confidence level $1 - \alpha$.

2. Power. One may wish to construct the test of each parameter value so as to maximize the probability to reject it under assumption of some stated alternative.

3. Choice of alternative for test. For example, when testing a value μ_0 one may be regard the alternative to be $\mu \neq \mu_0$, or one might have a one-sided alternative, e.g., $\mu < \mu_0$.

4. Conditional coverage. One may wish to construct the interval so that it does not contain identifiable data outcomes such that the maximum coverage probability over all values of μ and conditional on the data outcome is less than the stated confidence level (which would indicate the presence of negatively biased relevant subsets).

5. Adapted conditional coverage. One may wish to require that the maximum conditional coverage not be below the confidence level for a specified subset of parameter values.

6. Bayesian correspondence. One may wish to require that the resulting confidence interval matches, at least approximately, the interval that would result from a Bayesian analysis for some prior.

7. Flip-flopping. One would like the procedure to be such that it does not allow one to alter the way that the result is reported depending on the outcome such that the coverage properties are formally violated.

PCL issues: Mmin, bands, etc.

References

[1] Robert D. Cousins, Draft note on Negatively Biased Relevant Subsets and references therein.

[2] Glen Cowan, Kyle Cranmer, Eilam Gross and Ofer Vitells, *Power-Constrained Limits*, arXiv:1105.3166.

[3] T. Junk, Nucl. Instrum. Methods Phys. Res., Sec. A **434**, 435 (1999); A.L. Read, J. Phys. G **28**, 2693 (2002).