Glen Cowan
RHUL Physics
24 July, 2009

Comment on different likelihood ratios

A recent presentation at the ATLAS Statistics Forum [1] compared the test statistic used for exclusion in the CSC Higgs combination [2] with a modified statistic that provides an equivalently powerful test but allows one to compute accurate p-values more easily. The difference between the two statistics is that in the CSC version, the estimated number of signal events was effectively constrained to be positive, but in the modified version this constraint is not applied. In this note both test statistics are examined for a simple special case where the data consist of a single Gaussian distributed value.

The conclusions of this note agree with those of [1] on the question of using the modified statistic. Furthermore it is shown that both statistics are equivalent to using the likelihood ratio as defined in the LEP analyses, also in agreement with Ref. [1]. A second suggestion made in [1] was to use the CL_s method for exclusion limits; this question is not addressed here.

Note that the issues here have nothing to do with how one treats nuisance parameters. In fact, in the example here there are none. The concept of the profile likelihood relates to how one treats nuisance parameters, and therefore here the method used in the CSC Higgs combination is not referred to here as the “profile-likelihood” method. One could envisage applying the modification proposed here regardless of how one chooses to treat nuisance parameters.

1 Basic formalism and CSC-style likelihood ratio

Suppose the outcome of a measurement is a continuous variable x modeled as following a Gaussian distribution, with expectation value and variance given by

$$E[x] = \mu s + b \,, \tag{1}$$

$$V[x] = \sigma^2 \,. \tag{2}$$

Here s and b are specified constants denoting the contributions from signal and background, respectively. The standard deviation σ is also taken here as a known constant, and μ is a strength parameter. One wishes to make inferences about μ based on a single observed value of x. The likelihood function is

$$L(\mu) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2} \frac{(x-(\mu s + b))^2}{\sigma^2} \right] \,. \tag{3}$$

To test a value of μ, one constructs the likelihood ratio

$$\lambda(\mu) = \frac{L(\mu)}{L(\hat{\mu})} \,, \tag{4}$$

where $\hat{\mu}$ is the Maximum Likelihood (ML) estimator.

Suppose now that on physical grounds, μ should be positive. The maximum value of the likelihood from within the allowed parameter space for μ is therefore found from

$$\hat{\mu} = \begin{cases} \frac{x-b}{s} & x \geq b, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Rather than using the likelihood ratio directly, one usually uses the equivalent logarithmic variable $-2\ln\lambda(\mu)$. This is

$$-2\ln\lambda(\mu) = \begin{cases} \frac{(x-(\mu s+b))^2}{\sigma^2} & x \geq b\,, \\ \frac{(x-(\mu s+b))^2}{\sigma^2} - \frac{(x-b)^2}{\sigma^2} & \text{otherwise.} \end{cases} \tag{6}$$

For an upper limit on μ one defines the test statistic

$$q_\mu = \begin{cases} -2\ln\lambda(\mu) & \hat{\mu} \leq \mu\,, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

Putting together Eqs. (5), (6) and (7) gives

$$q_\mu = \begin{cases} \frac{(x-(\mu s+b))^2}{\sigma^2} - \frac{(x-b)^2}{\sigma^2} & x \leq b\,, \\ \frac{(x-(\mu s+b))^2}{\sigma^2} & b < x \leq \mu s + b\,, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

The p-value of a hypothesized value of μ is

$$p_\mu = \int_{q_{\mu,\mathrm{obs}}}^{\infty} f(q_\mu|\mu)\, dq_\mu\ , \tag{9}$$

where $f(q_\mu|\mu)$ is the pdf of q_μ under the assumption of μ. If x is Gaussian distributed with mean $\mu s + b$ and standard deviation σ, the quantity

$$\frac{(x-(\mu s+b))^2}{\sigma^2} \tag{10}$$

follows a chi-square distribution for one degree of freedom. From Eq. (8) it is clear that the pdf of q_μ does not, however, have this simple form but rather is more complicated.

2 Likelihood ratio without constraint on $\hat{\mu}$

According to the procedure proposed in Ref. [1], one defines an unphysical estimator $\hat{\mu}'$ which is allowed to be negative even if $x < b$, i.e.,

$$\hat{\mu}' = \frac{x-b}{s}\ . \tag{11}$$

This is used instead of $\hat{\mu}$ in the likelihood ratio (4). Defining a test statistic q'_μ in the same manner as above, one finds

$$q'_\mu = \begin{cases} \frac{(x-(\mu s+b))^2}{\sigma^2} & x \le \mu s + b\,, \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

Both test statistics, q_μ and q'_μ are shown as a function of x in Fig. 1.

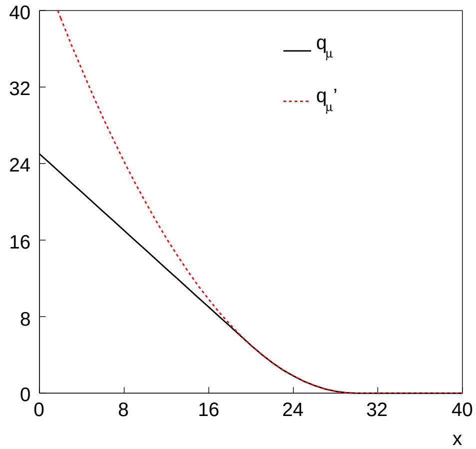

Figure 1: The statistics q_μ and q'_μ versus x (here $s = 10$, $b = 20$, $\sigma^2 = 20$).

From the form of (12) one can see that its pdf under the assumption of the strength parameter μ is a half-chi-square distribution. That is, if $x > 0$, which occurs half the time, one has $q'_\mu = 0$. The other half of the time when $x \le \mu s + b$, q'_μ follows a chi-square pdf for one degree of freedom. Therefore one recovers the simple result for the significance (see [2]),

$$Z = \sqrt{q'_\mu} = \begin{cases} \frac{\mu s+b-x}{\sigma} & x \le \mu s + b\,, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

A crucial point is now that q_μ and q'_μ are connected by a monotonic relation, as can be seen in Fig. 2. This means that they are equivalent test statistics. If one were to determine, e.g., using Monte Carlo, the exact sampling pdf $f(q_\mu|\mu)$ and determine from it the significance Z for a given observed x, then it would by construction agree with the value found from the simple formula (13).

Therefore there is a clear advantage from using the modified statistic q'_μ. It allows for a more accurate determination of the p-value, and hence significance, since one can exploit the chi-square properties of its pdf. The study in Ref. [1] applied this to the more realistic example of a Higgs search with Poisson distributed data and found essentially the same result.

3 LEP-style likelihood ratio

As a further step one can also show that the likelihood ratio (4) is in this example equivalent to the test statistic used at LEP and the Tevatron. There one effectively tests a value of μ using

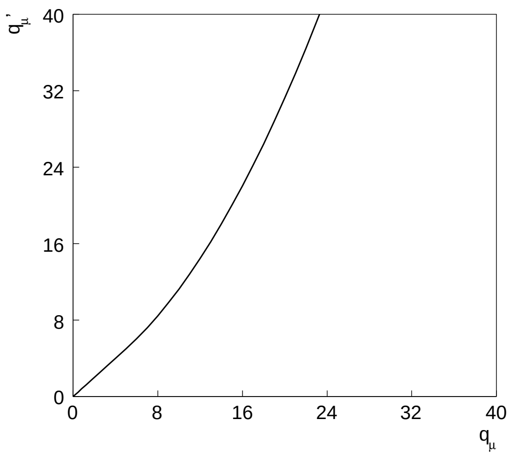

Figure 2: The statistic q'_μ versus q_μ. Different points on the curve correspond to different values of x.

$$
\begin{aligned}
q_{\mu 0} &= -2\ln\frac{L(\mu)}{L(0)} \\
&= \frac{(x-(\mu s+b))^2}{\sigma^2} - \frac{(x-b)^2}{\sigma^2} \\
&= \frac{-2\mu s x + (\mu s+b)^2 - b^2}{\sigma^2} \ . \qquad (14)
\end{aligned}
$$

From the last line in (14) one sees that this statistic is a linear function of x and therefore itself must follow a Gaussian distribution. Its expectation value and variance assuming a strength parameter μ are

$$
\begin{aligned}
E[q_{\mu 0}|\mu] &= \frac{-2\mu s(\mu s+b) + (\mu s+b)^2 - b^2}{\sigma^2} = -\frac{(\mu s)^2}{\sigma^2} \\
V[q_{\mu 0}|\mu] &= \frac{4\mu^2 s^2}{\sigma^4} V[x] = \frac{4\mu^2 s^2}{\sigma^2} \ , \qquad (15)
\end{aligned}
$$

and therefore the standardized variable

$$
y = \frac{q_{\mu 0} - E[q_{\mu 0}|\mu]}{\sqrt{V[q_{\mu 0}|\mu]}} = \frac{\mu s + b - x}{\sigma} \qquad (16)
$$

will follow a Gaussian with mean of zero and unit variance. The p-value of a hypothesized value of μ is therefore

$$
p_\mu = \int_{q_{\mu 0,\mathrm{obs}}}^{\infty} f(q_{\mu 0}|\mu)\, dq_\mu = 1 - \Phi(y) \ , \qquad (17)
$$

and the corresponding significance is

$$
Z = \Phi^{-1}(1-p) = \Phi^{-1}(\Phi(y)) = \frac{\mu s + b - x}{\sigma} \ . \qquad (18)
$$

This is therefore equivalent to the significance found using q'_μ above.

According to the Neyman-Pearson lemma, the statistic $q_{\mu 0}$ will give the most powerful test, and in other more complicated examples one could expect that the statistic $L(\mu)/L(\hat{\mu})$ is not quite as powerful.

4 Median significance

In addition to finding the significance from a given observation of x, one can ask for the median significance with which one can reject a certain strength μ assuming data distributed according to a different strength μ'. For exclusion limits this means finding the median, assuming $\mu = 0$ of the significance with which one can reject μ.

In the example presented above, the significance Z from Eq. (13) is a linear function of x, which is itself Gaussian distributed with mean $\mu s + b$. Therefore the median signifance is given by

$$\mathrm{med}[Z|\mu'] = \frac{\mu s + b - (\mu' s + b)}{\sigma} = \frac{(\mu - \mu')s}{\sigma} \; . \tag{19}$$

In this case the Asimov data value $x_{\mathrm{A}} = \mu' s + b$ clearly gives the same result, since this is equal to the median value of x.

References

[1] N. Andari et al,. *Proposal to change limits for exclusion*, presentation at the ATLAS Statistics Forum, 8 July, 2008.

[2] G. Aad et al. (ATLAS Collaboration), *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, e-print arXiv:0901.0512, CERN-OPEN-2008-020 (2008).