"Likelihood" Method for Gamma-Ray Burst and other Source Searches

T.J. Haines
LANL, December 20, 1999

There has been some effort of late to fold more information that we know about GRBs into the analysis technique. In particular, at the New Hampshire collaboration meeting, G. Sinnis showed what was termed a likelihood method that included the positional information from BATSE into the analysis; this technique has been applied and extended by J. McCullough. In this brief note, I will first propose what I think is a correct method for performing the likelihood (or at least likelihood-like) analysis in this specific example. I will then review the method now being used by J. McCullough/G. Sinnis, and illustrate the weakness of that method. I will finish by giving some examples that compare the two methods. In another note to follow, I will be summarizing ways to extend the approach to include even more information about the GRB itself, like the statistical error in background estimation, and information about the point-spread-function of Milagrito/Milagro.

Introduction & Jargon

Until the last few months, the search for a signal from GRBs with Milagrito/Milagro data has proceeded in a fairly straightforward manner, essentially identical to the method used by the CYGNUS collaboration [D.E. Alexandreas et al., Astroph. J. 426(1994)], and perhaps others. To summarize, for each GRB detected by BATSE, the entire "error circle" is searched for a non-statistical excess of events (called N_o in this note) above the expected background level (called μ_b), with the search being performed with fixed overlapping bins. The Poisson probability (called $P(N_o|\mu_b)$) is used to judge the probability that a particular position within the error circle is a candidate GRB position. Sinnis/McCullough try to extend this to include the fact that, on average, one expects the GRB's true location to more often to be located near the center of the BATSE "error circle."

Let me here define a few more symbols.

μ_s	True level of signal in the GRB region
μ_b	True background flux in the GRB region
N_b	Number of background events in region
N_s	Number of signal events in region
N_o	Total number of events observed in the region

With these definitions, it is clear that $N_o = N_s + N_b$ is our constraint; all of the above numbers are by assumption (or by prior choice in a Baysian's view) positive. Let

$P_{BATSE}(r)$ represent the *properly normalized* probability density function of the probability of finding the true location of a particular GRB an angular distance r from the center of the BATSE bin. Here I will only concern myself with the analysis for a particular GRB, so I will ignore the (obvious) fact that e.g. all of the above background and signal levels, $P_{BATSE}(r)$ etc. vary on a GRB-by-GRB basis; that will be left as an exercise to the student/programmer.

Suggested Likelihood Methodology

With the above definitions, we can proceed to develop an appropriate likelihood test. I will take the view that we are performing this analysis by searching in overlapping bins throughout the region given by BATSE, to replicate the spirit of the McCullough/Sinnis work. In this case, $P_{BATSE}(r)$ is independent of the signal and background levels, and thus really does not effect the formalism. Said differently, I want to specify the likelihood (really likelihood ratio) that a particular bin in the entire BATSE region is the location of a Milagrito/Milagro detected GRB. $P_{BATSE}(r)$ will simply multiply the likelihood I will develop below.

We want to maximize the ratio of the probability that our observation is due to a signal of true "strength" μ_s and true background μ_b compared to the probability that our observation is only due to background ($\mu_s = 0$). Note here that for now I am assuming that the true background level is determined with perfect precision by the background estimation method; I will generalize to remove this assumption in a later note. The only unknown parameter that we are concerned about (so far) is the signal level; we will maximize the likelihood with respect to this variable, so we must write down the probability of having a particular μ_s given our observation of N_o. The probability we want is then for a particular unknown N_s

$$\mathrm{Prob}(N_s,\mu_s) = \mathrm{Prob}(N_s|\mu_s) * \mathrm{Prob}(N_b|\mu_b) = \mathrm{Prob}(N_s|\mu_s) * \mathrm{Prob}(N_o - N_s\,|\mu_b)$$

where we have invoked our constraint that $N_b + N_s = N_b$. Now, we do not know the actual number of signal or background events that we observed, we only know the total number that we observed. So, we must include all possibilities for the actual number of signal (and thus background) events that composed our observation. Said differently, our observation may have resulted from observing $N_s = 0$, $N_b = N_o$ events, or $N_s = 1$, $N_b = N_o - 1$ events, or $N_s = 2$, $N_b = N_o - 2$ events, or … etc. Thus the total probability is

$$P(\mu_s) = \sum_{N_b=0}^{N_o} P(N_b \mid \mu_b) * P(N_o - N_b \mid \mu_s)$$

where again $P(N|\mu)$ is the properly normalized Poisson probability of observing N given an expected average value μ. It should not make a difference if the sum is performed over the value of the background counts or the signal counts. It is straightforward to show that this can be re-written as

$$P(\mu_s) = \sum_{N_s=0}^{N_o} P(N_s \mid \mu_s) * P(N_o - N_s \mid \mu_b)$$

consistent with one's expectation.

The denominator of the likelihood ratio should be the numerator evaluated at $\mu_s = 0$. $P(N_o - N_b|0) = 0$ for all values of $N_o - N_b$ in the above sum *except* $N_o - N_b = 0$, where it has a value of 1. Thus $P(\mu_s = 0) = P(N_o|\mu_b)$, simply the Poisson probability for the null hypothesis of no-source. This is as one would have expected. The likelihood ratio can then be written explicitly as

$$L(\mu_s) = \frac{P(\mu_s)}{P(\mu = 0\ \ )} = e^{-\mu_s} \left(\frac{\mu_s}{\mu_b} \right)^{N_o} \sum_{N_b=0}^{N_o} \left(\frac{\mu_s}{\mu_b} \right)^{N_b} \frac{1}{N_b!(N_o - N_b)!} \qquad eq.1$$

where terms independent of N_b have been pulled out of the summation, and a few terms from the denominator and numerator have been cancelled. Unfortunately, I have not found a closed-form analytic solution for the maximum of this likelihood function. The full likelihood, as mentioned above, would include an additional multiplicative term $P_{BATSE}(r)$, which depends upon the position of the bin of interest with respect to the BATSE-determined position.

The McCullough/Sinnis Method

Using the above notation, the likelihood ratio as shown at the New Hampshire meeting and in a few subsequent meetings at LANL is

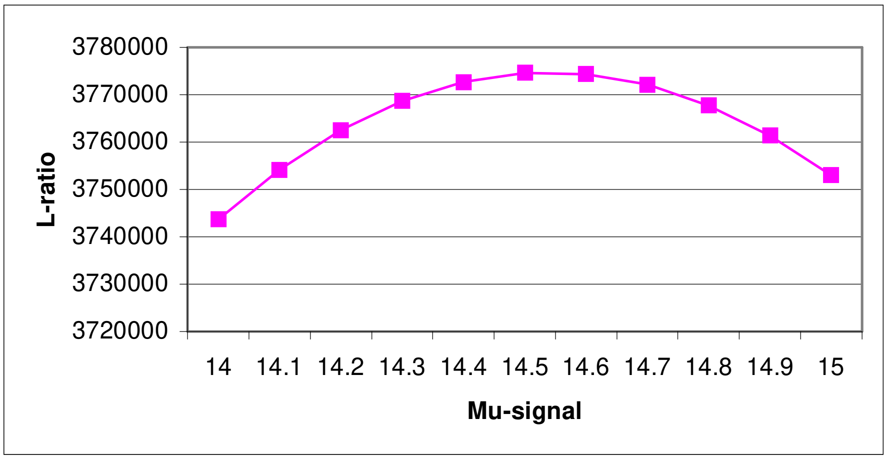
$$L_{Joe}(N_s, \mu_s) = \frac{\mathrm{P}^{\mathrm{Joe}}_{\mathrm{BATSE}}(\mathrm{r}) * \mathrm{P}(\mathrm{N_b} \mid \mu_\mathrm{b}) * \mathrm{P}(\mathrm{N_s} \mid \mu_\mathrm{s})}{P(N_o \mid \mu_b)}$$

where I have explicitly put the term P^{Joe}_{BATSE} in this expression to point out that this term is not properly normalized; instead in this implementation so far presented, it is normalized to $P^{Joe}_{BATSE}(r = 0) = 1$, so that it does not have unit total area. It should be normalized properly, as pointed out above. Inspection of the above shows that this likelihood corresponds to only one term in eq. 1.

Ignoring BATSE-position term for the moment, the Joe-likelihood is maximized with respect to both N_s and μ_s. Joe has noticed that this leads to a solution $N_s \cong \mu_s$, and points out that one can immediately see by noticing that the last term in the numerator maximizes as this value. In fact, I have analytically (for large N such that N is approximately a continuous variable) proven that the maximum for this situation occurs at $N_o - N_s = \mu_b$ (or $N_b = \mu_b$) and $\mu_s = N_o - N_b = N_o - \mu_b$. Thus, practically speaking, it will maximize at the integer nearest $N_o - \mu_b$. In general, and specifically for smaller N, the true likelihood will not maximize near this integer; the property of the Poisson distribution is that it becomes quite narrow and centered on N only for large N.

A Specific Example

I want to illustrate the difference between the two methods with an example, perhaps the most interesting is taken from GRB 970417, where we observed (using the numbers from the GRB paper) 18 events ($N_o = 18$) and expected 3.46 as background (μ_b). Performing a grid-search in both N_s and μ_s, the maximum of L_{Joe} is at $N_s = 15$ and $\mu_s = 15.0$. Using eq. 1, the likelihood ratio as a function of μ_s is shown



where the maximum is found at $\mu_s = 14.54$. In fact, for a variety of examples, I am finding numerically the solution that $\mu_s = N_o - \mu_b$, though I am still trying to find the analytic solution to the problem. I have also found numerically the solution $\mu_s \cong 0$ in the case that $N_o < \mu_b$, that is to say when one observes fewer events than the expected background (though I did not explore the 'non-physical' part of the parameter space $\mu_s < 0$).

Summary

In this note I have proposed a method for performing a "binned" likelihood analysis on Poisson distributed data. The solution to the max-L problem numerically behaves as one would expect. Subsequent notes will expand on this method and explore its ramifications.