STATISTICS ISSUES FOR THE MINI-BOONE EXPERIMENT

Byron P. Roe and Michael Woodroofe; reported by B. Roe
University of Michigan, Ann Arbor, Michigan 48109, U.S.A.

Abstract
The Mini-BooNE experiment is about to run at Fermilab searching for $\nu_\mu \rightarrow \nu_e$ oscillations. Statistical methods for setting confidence intervals for parameters in this and related experiments will be discussed.

1 INTRODUCTION

In the mini-BooNE neutrino oscillation experiment at Fermilab, an 8 GeV proton beam hits a target producing pions and kaons. Some decay in a 50 (25) m decay length producing neutrinos. The detector is 500 m downstream. Most neutrinos crossing the detector are ν_μ. The experiment is a search for ν_μ oscillating into ν_e, in the same parameter range as the signal reported by the LSND experiment[1]. The oscillation probability is:

$$p_{\phi,\Delta m^2}(E) = \sin^2 2\phi \sin^2\left(\frac{1.27\Delta m^2 L}{E}\right),$$

where neutrino energy E is in MeV, Δm^2 is in eV^2, L is in meters, and ϕ, Δm^2 are parameters to be determined.

There are 3 sources of background. There is background from real ν_e from μ decay and K_{e3} decay, from $\nu_\mu N \rightarrow \mu^- N'$ events mistakenly identified as ν_e events, and also from $\nu_\mu N \rightarrow \nu_\mu \pi^0 N'$ events mistakenly identified as ν_e events.

The number and spectrum of these background events will be known (with some uncertainty). In two years of running at nominal intensity, 1500 intrinsic ν_e background, 500 π^0 mis-id background, and 500 ν_μ mis-id background are expected.

There may or may not be a signal. If LSND is confirmed, about 1000 oscillation signal events are expected.

2 SETTING LIMITS IF NO SIGNIFICANT SIGNAL IS SEEN

2.1 A Simplified Problem

Consider the simplified situation of a simple Poisson distribution. Several methods have been proposed over the last few years for setting limits.

2.11 The Unified Method

The Unified Method was introduced into particle physics by G. Feldman and R. Cousins[2]. Let f_μ be the PDF of a Poisson (μ) distribution. If θ signal and b background events are expected,

$$L(\theta|n) = f_{b+\theta}(n) = \frac{1}{n!}(b+\theta)^n e^{-(b+\theta)}.$$

Let

$$\hat{\theta} = \max(n-b, 0),$$

$$R_\theta(n) = \frac{L(\theta|n)}{L(\hat{\theta}|n)} = \left(\frac{b+\theta}{\max(n,b)}\right)^n e^{\max(\mathrm{n,b})-(\mathrm{b}+\theta)},$$

and

$$P_\theta\{n : R_\theta(n) \geq c_\theta\} \geq 1-\alpha.$$

Then the Unified intervals are $\{\theta : R_\theta(n) \geq c_\theta\}$. That is, the unified intervals consist of taking regions of high likelihood, relative to the maximum likelihood, in the space of the allowable parameters. Extensive tables are given by Feldman and Cousins. This method always finds a non-null interval, even if the experimental result is improbable. It goes automatically from a one sided limit to two-sided interval. The intervals are invariant under a change of variable. Regions of high relative likelihood are a general method for obtaining confidence sets that has long been recommended by statisticians, as noted in [2]. In the present context, the method is reliable if the observed n is much larger than the expected background b. However, it can give an anomalously low upper bound if n is smaller than b, as noted in [4]. Feldman and Cousins recognized this and suggested that the expected sensitivity be quoted as well as the limit.

The Karmen experiment[3] expected $1.25 \pm .25$ neutrino oscillation signal events given the LSND experiment results and expected 3 background events. Observing 0 events they quoted a Unified Method 90% UCB as 1.08. This value is too small. For if no signal events are observed, then there are no signal events and no background events; and if we simply observed no signal events, then the Neyman Pearson intervals and Baysian intervals with a uniform prior, described below, both produce an upper confidence bound of 2.3.

2.12 Conditional Confidence Intervals

Let N be the number of observed events and B the (random) background. If $N = n$, then clearly $B \leq n$. If it were known that $B \leq m$ apriori, then the likelihood would be

$$L(\theta|n) = \frac{1}{F_b(m)} f_{b+\theta}(n); \;\; n \leq m,$$

$$L(\theta|n) = \frac{1}{F_b(m)} \sum_{k=0}^{m} f_b(k) f_\theta(n-k); \;\; n > m,$$

where f_μ and F_μ are the PDF and DF of the Poisson(μ) distribution. The Conditional Confidence method[4] uses this likelihood with $m = n$ and applies the Unified Method to it.

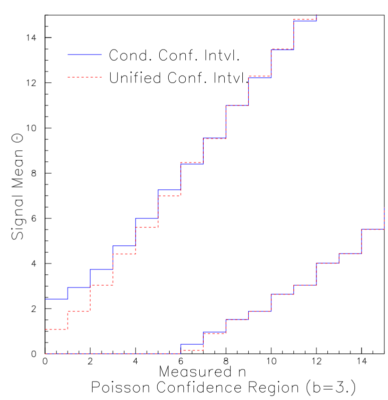

Fig. 1: Unified and Conditional Intervals

The confidence interval obtained is a conditional confidence interval. This method has the same advantages as does the Unified Method, and, in addition, gives sensible results if fewer events are seen than expected from background. However, a problem was found[5]. Suppose the method is applied to the continuous problem of measuring a parameter known to be greater than zero, with a normally distributed measurement error. The Conditional Confidence Method leads to a positive lower limit for all positive observations, in contrast to the sensible lower limit obtained with the Unified Method. The present method also has a lower limit problem for large numbers of events in the Poisson distribution problem.

2.13 Flat Prior Bayes Credible Interval

If Θ is given a uniform prior, let g and G denote the posterior density and DF of Θ. Thus

$$g(\theta|n) \propto f_{b+\theta}(n).$$

The Bayesian credible intervals are

$$[a_n, b_n] = \{\theta : g(\theta|n) \geq c_n\}$$

where c_n is so chosen that $G(b_n|n) - G(a_n|n) = 1 - \alpha$.

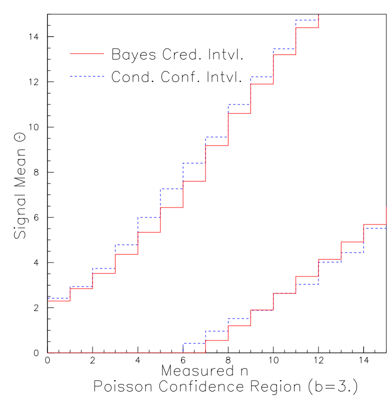

Fig. 2: Bayesian Intervals with Flat Prior

Examine the frequentist confidence level of this Bayesian credible interval. For the problem with a normally distributed measurement error, it can be shown that the frequentist coverage probability (FCP) of the Bayesian interval is always

$$FCP \geq \frac{1-\alpha}{1+\alpha}.$$

This limit is very conservative. For $\alpha = 0.1$, this limit gives 0.8182, but numerical evaluations show the limit to be ≥ 0.86. In addition, a small ad hoc modification can make the difference in values of the FCP and $1 - \alpha$ very small indeed. Similar results hold for the Poisson case[6].

Suppose the nuisance variable σ for the normally distributed measurement problem or b for Poisson problem has an uncertainty. This can be treated in a Bayesian manner: set the priors to be $1/\sigma^2$ for the normal and e^{-b} for the Poisson problem. For the normal case two independent random variables $X \sim$ Normal$[\theta, \sigma^2]$ and $W \sim \sigma^2\chi_r^2$ are observed. The second variable gives a measure of σ. The Bayesian credible limits are widened as expected. However, if examined from the frequentist view, the deviation of the frequentist confidence from Bayesian credible value is lessened[7]. With the nuisance variables smeared, the limits get wider for far negative results, as σ becomes less certain.

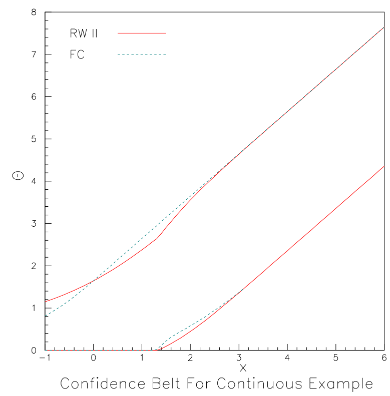

Fig. 3: Confidence and Credible Belts for the Continuous Example

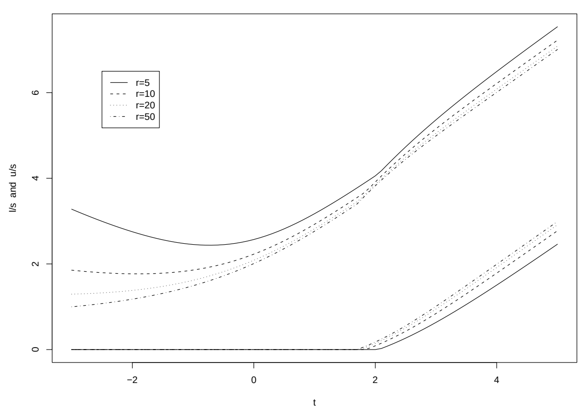

Fig. 4: 90% Credible Intervals for Various r

This method has most of the good advantages of both the Unified Method and the Conditional Confidence Method. However, the frequentist confidence limits are not completely constant (varying from 0.86 to $>$ .9) for the 90% credible limit Poisson case, and the intervals are not invariant under a change of variable since the intervals have been chosen to minimize length in the θ-scale. (Other statistical tools also lack invariance, .e.g., the Cramer-Rao inequality.)

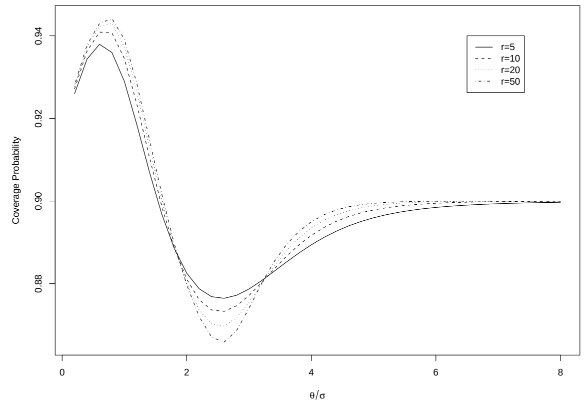

Fig. 5: Frequentist Coverage Probability

2.2 Poisson Problem with Extra Information for Each Event

Consider next the less trivial case where there is partial information for each event as to what kind of event it is. This can be modelled as a "marked Poisson" process[8]. The energy of the events are the "marks" with intensity

$$\lambda_{\phi,\Delta m^2}(e,x) = N[\theta g(e,x) + bh(e,x)],$$

where g and h are normalized to integrate to 1; θ depends on ϕ and Δm^2 and g depends on Δm^2. (ϕ is usually called θ, but I have changed the name here to avoid confusion.) N is proportional to the overall length of the run. The normalization of θ and b are then independent of run length. g and h are also dependent on x, an event shape parameter, such as the output of a neural net distinguishing ν_e from ν_μ or neutral current events with a π^0.

The "marked Poisson" process corresponds to binning with bin width approaching zero. If $0 < e_1 < e_2 < e_3 \cdots < e_n < \infty$, then the probability of observing n events with $e_i \le E \le e_i + de_i$, and $x_i \le X \le x_i + dx_i, \;\; i = 1, \cdots, n$ is

$$L = \left(\prod_{i=1}^{n} \lambda_{\phi,\Delta m^2}(e_i,x_i) \times de_i dx_i\right) \times e^{-\Lambda(\phi,\Delta m^2)},$$

where,

$$\Lambda(\phi,\Delta m^2) = \int_{E_{min}}^{E_{max}} \int \lambda_{\phi,\Delta m^2}(e,x) dx de = \theta + b.$$

If a Bayes method with uniform prior is used, the marginal PDF is

$$f(n,e,x) = \prod_{k=1}^{n} h(e_k,x_k) \int_0^\infty e^{-(b+\theta)} \prod_{k=1}^{n} [b + \theta r(e_k,x_k)] e^{-(b+\theta)} d\theta,$$

with $r(e_k,x_k) = g(e_k,x_k)/h(e_k,x_k)$.

Let $f_\theta(n,e,x)$ be the density function for given θ. Then the posterior PDF is

$$q(\theta|n,x) = \frac{f_\theta(n,e,x)}{f(n,e,x)}.$$

The appropriate Bayes credible interval/limit can then be found by iteration techniques. The present situation is a little more complicated than that described in Roe and Woodroofe[6], in that g is now a function of the parameters of interest.

3 Determining Values of Parameters if a Signal is Seen

If the LSND results are correct, of the order of 1000 signal events and 2500 background events are expected. Since $g = g_{\Delta m^2}$, define different variables

$$\theta(\phi, \Delta m^2) g_{\Delta m^2}(e, x) = a(e, x) \sin^2 2\phi \sin^2 \left(\frac{1.27 \Delta m^2 L}{e} \right),$$

where $a(e, x)$ is *not* a function of $\phi, \Delta m^2$. If LSND is correct the results will not be near the background boundary. The Unified Method should work well here. Let $\omega = \ln L$,

$$L = \prod_{k=1}^{n} N[\theta g(e_k, x_k) + bh(e_k, x_k)] e^{-\Lambda},$$

where

$$\Lambda = \int_{E_{min}}^{E_{max}} \int N dx de \left[a(e, x) \sin^2 2\phi \sin^2 \left(\frac{1.27 \Delta m^2 L}{e} \right) + bh(e, x) \right].$$

$$\omega = n \ln N - \Lambda + \sum_{k=1}^{n} \ln \left[a(e_k, x_k) \sin^2 2\phi \sin^2 \left(\frac{1.27 \Delta m^2}{e_k} \right) + bh(e_k, x_k) \right].$$

The maximum likelihood corresponds to the minimum of $-\omega$. A grid of nearby points can then be examined using many Monte Carlo simulations of the experiment at each point. Find for each Monte Carlo experiment the value of $R_{test} = L_{test}/L_{max-test}$ such that $100\alpha\%$ of the time the R found would be lower than R_{test}. Plot out the region(s) of $\phi, \Delta m^2$ where R_{test} corresponds to the R obtained in the experiment for that $\phi, \Delta m^2$ point.

Energy smearing can be included in the MC experiments, as can the effect of uncertainties in nuisance variables g, h, and b. The latter would be included as systematic errors, somewhat alleviating the conceptual problem of mixing Bayes and frequentist concepts.

ACKNOWLEDGEMENTS

I would like to thank the conference organizers and especially the local organizers James Stirling and Mike Whalley, for their gracious hospitality and for the fine organization of the conference. The considerable efforts of the conference secretary, Linda Wilkinson, were very much appreciated.

4 REFERENCES

References

[1] A. Aguilar et al., *Phys. Rev.* **D64**, 112007 (2001).

[2] G.J. Feldman and R.D. Cousins, *Phys. Rev.* **D57**, (1998), 3773-3789.

[3] B. Zeitnitz et. al. Neutrino Oscillation Results from KARMEN. *Progress in Particle and Nuclear Physics*, **40**, (1998), 169-181.

[4] Byron P. Roe and Michael B. Woodroofe, *Phys. Rev.* **D60**, (1999),053009.

[5] Robert D. Cousins, *Phys. Rev.* **D60**, 053009 (1999).

[6] Byron P. Roe and Michael B. Woodroofe, *Phys. Rev.* **D63**, (2001),013009.

[7] Tonglin Zhang and Michael B. Woodroofe (2001), to appear in J. Statdist. Plan. Inf. (2002).

[8] Edward Kao, *An Introduction to Statistical Processes*, Duxbury, 1997. (Page 80)