Setting Confidence Belts

Byron P. Roe

University of Michigan

March 27, 2000

Introduction

This work has been done together with Michael B. Woodroofe.

We consider two simple problems for setting confidence belts

- Poisson distribution in the presence of background. For example counts from a radioactive source in the presence of background.
- Measurement of a parameter θ, known to be ≥ 0, with a measurement error.

Although simple problems, there has been considerable work and discussion on them in the last few years.

Most of the discussion has centered on the Poisson case, when the number of observed events is small.

This is an important region these days because of the many search experiments.

- Higgs particle searches
- Supersymmetry searches
- Neutrino oscillation searches (LSND, KARMEN, Mini-BooNE)

We will *briefly* describe some recent attempts, which have served to crystallize a number of desirable criteria for a methodology.

Is the conventional confidence definition appropriate?

Feldman-Cousins

The standard method for setting 90% CL bounds has been to select a region with 5% above and 5% below the bounds. About two years ago, Gary Feldman and Bob Cousins suggested a new method. For the Poisson case, for each possible value of the parameter θ, they looked at the ratio of the probablity of getting the observed n for that θ compared to the probability for the physically allowed θ giving the highest possible probability, and picked n's with the highest ratio to build a 90% region. This solved two problems with the old procedure.

In the old procedure, one decided after the experiment whether to set a confidence belt or an upper limit. This introduced a bias.

In the old procedure, the confidence region sometimes had no physical values in it. For example if no events were observed and there was a large known mean background, the entire region of θ in the confidence belt could be negative.

Desiderata

- Automatically go from upper limit to confidence belt.
- For any observation find a confidence belt in the physical region.

Problem with Feldman-Cousins

The F-C method had a problem if no events were observed. If 0 events were seen, there were 0 signal and 0 background events. The fact that 0 background events was seen is interesting, but irrelevant to the question of whether signal events were seen. The limit on θ should come from $p(0|\theta) \leq 10\%$, which sets a limit at $\theta = 2.3$. Such a case did occur in the Summer of 1998 from initial KARMEN results. They had 0 events with a background of 2.88, and using the F-C method obtained an upper limit of $\theta = 0.9$.

Roe-Woodroofe

Roe and Woodroofe introduced the concept of conditional coverage. One knows that the background was $\leq n_{observed}$.

- Use a sample space not of all events, but of events with background $\leq n_{observed}$.
- Consider the conditional probability of obtaining k events (k not necessarily $n_{observed}$) with this restricted background.
- Use this probability to set the ratio used by F-C and for the coverage probability.

This method had the advantage of giving an upper limit of about 2.42 for 0 observed events independently of the background mean b.

Desiderata

- Automatically go from upper limit to confidence belt.
- For any observation find a confidence belt in the physical region.
- For Poisson events with $n_{observed} = 0$, $u \approx 2.3$, for any background b.

Cousins Note

Just before the last Confidence Limit Workshop, Bob Cousins noted that the R-W procedure, had serious problems with its lower limit for the continuous case. If one measured a parameter $\theta \geq 0$ with a measurement error ϵ, $x = \theta + \epsilon$, then $\theta = 0$ should be a possibility as long as $p(x > x_{measured}|\theta = 0) > 10\%$. If ϵ has a normal distribution (0,1), then this corresponds to $x = 1.39$. The R-W procedure eliminated $\theta = 0$ for all $x > 0$! It appeared that the R-W upper limit and the F-C lower limit were needed.

In addition, the conditional coverage concept was unfamiliar (but we're trying to publicize it!).

Desiderata

- Automatically go from upper limit to confidence belt.
- For any observation find a confidence belt in the physical region.
- For Poisson events with $n_{observed} = 0$, $u \approx 2.3$, for any background b.
- For the continuous case, $\ell = 0$ for $x \leq 1.39$ for unit normal error.
- Conventional coverage is desirable as well as conditional coverage.

We have tried to find a method which was not ad-hoc, but followed from some clear principle, and which, in some sense, is optimal. Optimality is useful. Methods have been suggested which were so conservative as to be of limited use (limits for the $n = 0$ Poisson case of 3.5 or so).

Desiderata

- Automatically go from upper limit to confidence belt.
- For any observation find a confidence belt in the physical region.
- For Poisson events with $n_{observed} = 0$, $u \approx 2.3$, for any background b.
- For the continuous case, $\ell = 0$ for $x \leq 1.39$ for unit normal error.
- Conventional coverage is desirable as well as conditional coverage.
- The method should be optimum in some sense.
- The method should follow from some systematic approach, i.e., not be ad-hoc.

New Approach

- Meld Bayesian and Frequentist elements.
- Start with Bayesian approach with uniform improper prior, $p_{prior} = 1$. (Introduces small bias toward larger θ.)
- Examine coverage from frequentist and conditional frequentist points of view.

Variables

Continuous case:

- x = variable being measured.
- θ = parameter to estimate ≥ 0.
- ϵ = measurement error, normal (0,1)

Poisson case:

- n = number of events measured
- θ = mean value for signal, to be estimated.
- b = mean value for background, assumed known.
- x = measured parameter for each event with density $g(x)$ for signal events and $h(x)$ for background events

Continuous case

- Density function: $f_\theta(x) = \phi(x - \theta)$.
- Distribution function: $\Phi(x) = \int_{-\infty}^{x} \phi(y)dy$.
- Conditional density of θ given x:

$$g(\theta|x) = \frac{\phi(x-\theta)}{\Phi(x)}.$$

  Note this is just the conditional probability of x given that the measurement error is known to be $\leq x$.

For Bayes credible region $1 - \alpha$, we want ℓ and u such that $P\{\ell \leq \theta \leq u|x\} = 1 - \alpha$ and $u - \ell =$ minimum. This is solved by finding the set $[\ell, u] = \{\theta : g(\theta) \geq c\}$, i.e., selecting the interval to have the highest g's.

MICH

Details

We noted $g(\theta|x) = \frac{\phi(x-\theta)}{\Phi(x)}$. Then $g(\theta|x) \geq c$ iff $|\theta - x| \leq d$.

(For $\phi = (1/\sqrt{2\pi})e^{-\frac{1}{2}(x-\theta)^2}$,
$d = \sqrt{-2\ln c - \ln 2\pi - 2\ln\Phi(x)}$.)

In general there are two cases:

1: $d < x$. $d = \Phi^{-1}[\frac{1}{2} + \frac{1}{2}(1-\alpha)\Phi(x)]$,

2: $x \leq d$. $d = \Phi^{-1}[1 - \alpha\Phi(x)]$.

These curves meet at $x_0 = \Phi^{-1}[1/(1+\alpha)]$ and the interval $[\ell, u] = [\max(x-d, 0), x+d]$.

Coverage

By construction this method has exact Bayesian credible interval coverage and in the Bayesian sense has minimized the interval $[\ell, u]$.

Next consider frequentist properties.

- The error $\theta - x$ has exact conditional coverage. If x' is an independent result from the conditional sample, $Prob[\ell - x \leq \theta - x' \leq u - x] = 1 - \alpha$.
- The frequentist interval obtained this way is of minimum length among all intervals satisfying the conditional coverage condition.

Next look at the conventional (unconditional) coverage. Let c_1 be the u, i.e., the x for which $\theta = x + d(x)$, and c_2 be the ℓ, i.e., the x for which $\theta = x - \max(d(x), 0)$. We can then show that the conventional coverage

$$\Phi[c_2 - \theta] - \Phi[c_1 - \theta] \geq \frac{1-\alpha}{1+\alpha}.$$

For 90% C.L, $\alpha = .1$, and $(1-\alpha)/(1+\alpha) = .8182$. For a normally distributed measurement error, the minimum coverage is about 0.8607 and maximum about .955.

This is already fairly good, but can be improved, with little cost, if the upper limit u is replaced by $u'(x) = \max[u(x), x + \Phi^{-1}(1-(1/2)\alpha)]$, corresponding to an $x = c_1'$. One can then show that the coverage satisfies

$$\Phi[c_2 - \theta] - \Phi[c_1' - \theta] =$$

$$\frac{1}{2} + \frac{1}{2}(1-\alpha)\Phi(c_2) - \min[\frac{1}{2}, \Phi(c_1)]\alpha.$$

The conventional coverage is then ≥ 0.900 to three significant figures.

The Poisson Case

The Poisson case is treated similarly to the continuous case. Suppose the unknown signal mean is θ and the known background mean is b. For the uniform improper prior,

$$p(\theta + b|n) = \frac{p(n|\theta + b)}{P(n|b)}, \text{ where}$$

$$p(n|\theta + b) = \frac{e^{-(\theta+b)}(\theta + b)^n}{n!} \quad \text{and}$$

$$P(n|b) \equiv \sum_{k=0}^{n} p(k|b).$$

We wish to find the credible region defined as $P\{\ell \leq \theta \leq u\} = 1 - \alpha$, where ℓ is the lower limit and u is the upper limit for θ. To minimize the interval, we want $[\ell, u] = \{\theta : p(\theta + b|n) \geq c\}$.

Details for The Poisson Case

The Bayesian integral probability for $\theta + b \le y$ is $G(y|n)/P(n|b)$, where

$$G(y|n) \equiv \int_0^y \frac{x^n}{n!} e^{-x} dx.$$

The Bayes upper (u_n) and lower (ℓ_n) limits for a given n satisfy

$$G(u_n + b|n) - G(\ell_n + b|n) = (1-\alpha)P(n|b).$$

This is equivalent to:

$$P(n|\ell_n + b) - P(n|u_n + b) = (1-\alpha)P(n|b),$$

where

$$P(n|\lambda) \equiv \sum_{k=0}^{n} p(k|\lambda).$$

Either

- $p(b + u_n|n) = p(b + \ell_n|n)$ or
- $p(b + u_n|n) \ge p(b|n)$ and $\ell_n = 0$.

More Details for The Poisson Case

The maximum of $p(\theta + b|n)$ occurs at θ_{max} =max$(n - b, 0)$. The maximum value of c is then

$$
\begin{aligned}
c_{max} &= \frac{e^{-n} n^n}{n! \sum_{k=0}^{n} (b^k e^{-b}/k!)} \quad \text{if } n - b \geq 0 \\
&= \frac{e^{-b} b^n}{n! \sum_{k=0}^{n} (b^k e^{-b}/k!)} \quad \text{otherwise.}
\end{aligned}
$$

Let c_{min} be the c corresponding to $\theta = 0$.

$$c_{min} = \frac{e^{-b} b^n}{n! \sum_{k=0}^{n} (b^k e^{-b}/k!)}$$

MICH

If $n \leq b$, the lower limit is $\ell_n = 0$.

If $n > b$, there may be a 2-sided or a 1-sided limit. These limits can be found by a double iteration, first get the interval from a given starting θ_{min}. Given the interval then find the probability of the credible region, and iterate the starting θ.

Coverage

- Conditional coverage $\approx 1 - \alpha$ except for discreteness.
- Conventional coverage can be as low as 0.86 for a 90% C.L. for $b = 3$.
- Tentative conservative version. Consider an alternate upper limit of θ given n as the one-sided limit with a higher C.L. = C.L.$'$. Take the maximum of the Bayesian and alternate upper limits. For a background of 3, and a C.L. of 90%, a C.L.$'$ of 92%, gives minimum coverage better than .88. For a fixed C.L.$'$ the limit for $n_{observed} = 0$ is fixed (2.53 for the case above) independent of the background. However, it is probably necessary to tune C.L.$'$ for the background expected.

Suppose we observe more...

Suppose that, in addition to the number of events, we observe a statistic x_i with each event.

- $p(x) = g(x)$ for signal events
- $p(x) = h(x)$ for background events.
- Observe $N = S + B$ events.
- Assume a Poisson distribution with mean $\theta + b$. θ and b are the expected number of signal and background events respectively.
- b is assumed known.

The probability can be shown to be

$$dp\{N = n, x \leq X \leq x + dx|\theta\}$$

$$= \quad \frac{(\theta + b)^n}{n!} e^{-(\theta+b)} \times \Pi_{i=1}^n \frac{\theta g(x_i) + bh(x_i)}{\theta + b} dx$$

$$\frac{d^n p}{dx_1 \cdots dx_n} \quad = \quad \frac{1}{n!} e^{-(\theta+b)} \times \Pi_{i=1}^n [\theta g(x_i) + bh(x_i)].$$

Bayes Treatment

Let the prior probability be the improper prior $p(\theta) = 1$. Then we can show

$$g(\theta|n,x) = \frac{e^{-\theta}\Pi_{i=1}^{n}[\theta g(x_i) + bh(x_i)]}{n!\sum_{m=0}^{n}\frac{b^{n-m}}{(n-m)!}C_{n,m}},$$

$$C_{n,m} =$$

$$\left[1/\binom{n}{m}\right]\sum_{j_1+j_2+\cdots+j_n=m}\Pi_{k=1}^{n}g^{j_k}(x_k)h^{1-j_k}(x_k).$$

We want upper and lower limits u, ℓ for θ such that $Prob\{\ell \leq \theta \leq u|n,x\} = 1-\alpha$ and the interval is minimized. Hence, $[\ell, u] = \{\theta : g(\theta|n,x) \geq c\}$.

Find the θ_{max} which maximizes $g(\theta|n,x)$. It is the solution of

$$\sum_{j=1}^{n}\left[\frac{g(x_j)}{\theta_{max}g(x_j) + bh(x_j)}\right] = 1,$$

MICH

which defines θ_{max} if θ_{max} is > 0. Otherwise, $\theta_{max} = 0$.

Let

$$G(a) = \int_0^a g(\theta|n, x) d\theta$$

After algebra

$$G(a) = \frac{1}{D} \sum_{m=0}^{n} b^{n-m} C_{n,m} \binom{n}{m} m! \int_0^a \frac{\theta^m}{m!} e^{-\theta} d\theta,$$

where

$$D = n! \sum_{m=0}^{n} \frac{b^{n-m}}{(n-m)!} C_{n,m}$$

The integral is the incomplete gamma function, available in CERNLIB. We can also use

$$1 - G(a) = \frac{e^{-a} \sum_{i=0}^{n} \frac{b^{n-i} C_{n,i}}{(n-i)!} \sum_{l=0}^{i} \frac{a^l}{l!}}{\sum_{m=0}^{n} \frac{b^{n-m}}{(n-m)!} C_{n,m}}.$$

- $\theta_{max} = 0$. By iteration find a such that $G(a) = 1 - \alpha$. This is the upper limit. The lower limit is 0.

- $\theta_{max} > 0$. Find $g(\theta = 0|n, x)$. By iteration find $g([\{\theta = a\} > 0]|n, x) = g(0|n, x)$. Calculate $G(a)$.

- If $G(a) \leq 1 - \alpha$, then the lower limit is 0 and proceed as for $\theta_{max} = 0$..

- If $G(a) > 1 - \alpha$, then there is a two-sided limit. Iterate trying larger lower limits, finding an upper limit with the same probability to zero in on Probability $=$ $G(u) - G(\ell) = 1 - \alpha$.

Summary

We have outlined a method for handling Poisson observations with a background and for handling the measurement of a parameter with an error.

This method satisfies a number of the desiderata that have been recognized over the past few years.

The method has been extended for the Poisson case to include results in which parameters are measured which help distinguish whether a particular event is signal or background.

Desiderata

- Automatically go from upper limit to confidence belt.
- For any observation find a confidence belt in the physical region.
- For Poisson events with $n_{observed} = 0$, $u \approx 2.3$, for any background b.
- For the continuous case, $\ell = 0$ for $x \leq 1.39$ for unit normal error.
- Conventional coverage is desirable as well as conditional coverage.
- The method should be optimum in some sense.
- The method should follow from some systematic approach, i.e., not be ad-hoc.