Refining Poisson Confidence Intervals

by

George Casella and Christian Robert

Purdue University – Technical Report #88-7
Cornell University – Technical Report BU-903-M

Department of Statistics
Purdue University

March 1988
Revised November 1988

Refining Poisson Confidence Intervals

George Casella [1] and Christian Robert [2]

Cornell University

April 1986
Revised July 1987
Second Revision, March 1988
Third Revision, November 1988

Abstract

A computational method is presented which, when applied to an existing confidence procedure, produces a uniformly superior procedure. This method, called refinement, actually produces a family of procedures that constitute a complete class. Although details are only given for the Poisson distribution, the refinement process is applicable to any discrete distribution. Tables of refined Poisson confidence intervals are given for confidence levels of .95. An example of refining negative binomial confidence intervals is also included.

Key Words and Phrases Discrete distribution, Admissible, Complete class.

AMS Subject Classification (1985) 62F25, 62C07.

---

[1] George Casella is Associate Professor, Biometrics Unit, Cornell University, Ithaca, NY 14853. This research was supported by National Science Foundation Grant No. DMS 8501973, and was performed, in part, while Prof. Casella was Visiting Professor at Purdue University.

[2] Research supported by l'Institut National de la Statistique et des Etudes Economiques (Paris) and the U.S. Army Research Office through the Mathematical Sciences Institute of Cornell University. This research was performed while Prof. Robert was Visiting Assistant Professor at Purdue University.

1. Introduction

Construction of optimal confidence intervals for discrete distributions is a difficult problem to handle analytically, and standard techniques tend to break down. The procedure described here resorts to numerical rather than analytical optimization, and produces exact confidence intervals that are as small as possible yet still maintain the nominal confidence level.

The earliest derivation of Poisson confidence limits is given by Garwood (1936), who used a technique similar to that of Clopper and Pearson (1934), and derived intervals based on the chi–squared distribution. Later work was done by Crow and Gardner (1959), using an acceptance region construction. Good normal approximations are given by Peizer and Pratt (1968) and Pratt (1968). Randomized Poisson intervals are treated by Stevens (1957) and Blyth and Hutchinson (1961). These admit a simple mathematical treatment and exact analytical solutions. The practical consequences of such a practice, however, are quite disturbing: one must add random noise to the data to produce a confidence interval.

In this paper, non–randomized Poisson confidence limits are constructed by a process similar to that used by Casella (1986) for the binomial distribution. This process, called refinement, is a numerical procedure to be applied to an existing $1-\alpha$ confidence procedure. The result is a new confidence procedure whose endpoints have been shifted to produce the smallest possible intervals. In fact, the refined confidence procedure is a member of a complete class of confidence procedures, i.e., the class of all $1-\alpha$ procedures that are admissible when length is used as the loss function.

In Section 2, we describe a number of preliminary facts, including an inadmissibility result about procedures with interval endpoints that are not strictly increasing. Section 3 contains a description of the refinement process along with some properties of the refined intervals. Section 4 has some comments, in particular a discussion of the applicability of refinement to other discrete distributions.

2. Preliminaries

Let X be a Poisson random variable with mean θ, that is,

$$P(X = x|\theta) = e^{-\theta}\theta^x/x!, \ x = 0, 1, \ldots \tag{2.1}$$

A confidence procedure C is an infinite collection of intervals $C = \{[\ell_x, u_x), \ x = 0, 1, \ldots\}$. The coverage probability of C is the probability that the random interval $[\ell_X, u_X)$ covers θ, and is given by

$$P(\theta \varepsilon C|\theta) = \sum_{x=0}^{\infty} I_{[\ell_x, u_x)}(\theta) P(X = x|\theta). \tag{2.2}$$

Note. We consider right half–open (instead of closed) confidence intervals in order to avoid technical difficulties related to coincidental endpoints. (If the upper endpoint of a particular interval is identical to the lower endpoint of another interval, such a point is called *coincidental*. This is explained more fully in Section 3.2.) Evaluation of coverage probabilities, when θ is a coincidental endpoint, becomes computationally clumsy because one must examine $\theta \pm \varepsilon$, for small ε, to locate minima. Thus, for computational purposes, we suppose that the intervals are of the form $[\ell, u)$. For example, if $\ell_m < u_k < \ell_{m+1}$, then

$$P(\theta \varepsilon C|\ell_m < \theta < u_k) = \sum_{x=k}^{m} e^{-\theta}\theta^x/x!$$

$$P(\theta \varepsilon C|\theta = u_k) = \sum_{x=k+1}^{m} e^{-\theta}\theta^x/x!.$$

Note also that we cannot consider left half–open intervals, as $\theta = 0$ would not be covered at all. We stress that this modification has no effect on any theoretical properties of the intervals, with respect to length or coverage probabilities; it merely allows the minima to be attained at the endpoints.

It is customary to guarantee a minimum coverage probability, or confidence coefficient, for a procedure C. To say that C is a $1 - \alpha$ confidence procedure means

$$\inf_{\theta} P(\theta \varepsilon C|\theta) \geq 1 - \alpha. \tag{2.3}$$

One of the earliest Poisson confidence procedures is due to Garwood (1936), who actually derived a fiducial interval. Garwood's work is similar in spirit to that of Clopper

and Pearson (1934), and gives the $1-\alpha$ confidence procedure

$$\ell_x^G = \frac{1}{2}\chi^2_{2x,1-\alpha/2},\ u_x^G = \frac{1}{2}\chi^2_{2(x+1),\alpha/2}, \tag{2.4}$$

where $\chi^2_{\nu,\alpha}$ denotes the upper α cutoff point of a chi–squared distribution with ν degrees of freedom, and where we define $\chi^2_{0,\alpha} \equiv 0$.

The Garwood intervals are, in general, too wide, yielding coverage probabilities much greater than $1-\alpha$. Strictly speaking, however, they are exact $1-\alpha$ intervals in that $\inf_\theta P(\theta\varepsilon[\ell_X^G, u_X^G]|\theta) = 1-\alpha$ (Casella and McCulloch, 1984). Figure 1 shows a plot of the coverage probabilities of the Garwood intervals for $1-\alpha = .9$, from which it can be seen that the coverage probabilities are quite large.

A relatively easy–to–compute procedure, and a respectable performer, can be obtained from a normal approximation with continuity correction. Using the fact that $EX = \mathrm{Var}(X) = \theta$, we define the lower and upper bounds by

$$\begin{aligned} \ell_x^N &= \min\{\theta:\ (x-\theta-\frac{1}{2})^2/\theta < z^2_{\alpha/2}\} \\ u_x^N &= \max\{\theta:\ (x-\theta+\frac{1}{2})^2/\theta < z^2_{\alpha/2}\}, \end{aligned} \tag{2.5}$$

where z_α denotes the upper α cutoff point of the standard normal distribution. Numerical studies suggest that these intervals offer a small length improvement over the Garwood intervals, and do maintain $1-\alpha$ coverage (the normal approximation without the continuity correction does not maintain $1-\alpha$ coverage). Figure 2 shows the coverage probabilities for the intervals defined by (2.5) for $1-\alpha = .9$.

It is relatively easy to show that, for any *ordered* Poisson confidence procedure (i.e. one that satisfies $\ell_x \le \ell_y$ and $u_x \le u_y$ if $x < y$; we will see later why we can restrict ourselves to these intervals), the minimum coverage probabilities are attained at the interval endpoints. For example, consider what happens to the coverage probability for θ in the range $\ell_{x_2} \le \theta < u_{x_1}$, where $x_2 \ge x_1 \ge 1$ and there are no other endpoints between ℓ_{x_2} and u_{x_1}. From (2.1) and (2.2), the coverage probability over this interval is given by

$$P(\theta\varepsilon C|\theta) = \sum_{x=x_1}^{x_2} e^{-\theta}\theta^x/x!,\quad \ell_{x_2} \le \theta < u_{x_1}, \tag{2.6}$$

and after some algebra, it can be shown that

$$\frac{d}{d\theta}P(\theta \varepsilon C|\theta) = \left(\frac{e^{-\theta}\theta^{x_1-1}}{x_2!}\right)\left(\frac{x_2!}{(x_1-1)!} - \theta^{x_2-x_1+1}\right), \quad \ell_{x_2} \leq \theta < u_{x_1}. \tag{2.7}$$

Examination of (2.7) shows that as θ moves from ℓ_{x_2} to u_{x_1}, the derivative either has no sign change, or changes from positive to negative. In either case, the minimum of the coverage probability over the interval (ℓ_{x_2}, u_{x_1}) is attained at an endpoint. There are other possibilities to consider with respect to the ranges of θ, but they all work out in a similar manner.

Therefore, if one has a confidence procedure whose coverage probabilities do not equal $1-\alpha$ at the interval endpoints, it should be possible to improve on such an interval. Examination of Figures 1 and 2 shows that this is the case for the Garwood and normal intervals, so we expect to be able to improve upon them. In fact, from Figures 1 and 2 it is easy to see that both the Garwood and approximate normal intervals are inadmissible – one can obtain a better procedure by merely shortening any interval whose coverage probabilities do not equal $1-\alpha$ when θ equals an interval endpoint. This will produce a new procedure that maintains $1-\alpha$ confidence with uniformly shorter intervals.

Crow and Gardner (1959) indirectly exploited this fact. They didn't work with confidence intervals directly, but instead they constructed $1-\alpha$ confidence intervals by inverting families of α–level tests with shortest acceptance regions. Figure 3 shows coverage probabilities of the Crow–Gardner intervals for $1-\alpha = .9$, and it can be seen that their intervals attain $1-\alpha$ coverage probability at virtually all interval endpoints.

While the Crow–Gardner intervals give a length improvement over the other two intervals described here, they suffer from a somewhat arbitrary condition imposed to achieve uniqueness. For many parameter values, it is the case that the "shortest" acceptance region is not unique (here "shortest" refers to the fewest number of x values required in the acceptance region). To achieve uniqueness, Crow and Gardner choose the acceptance region to yield the smallest possible upper confidence bound, a strategy that they claim leads to a certain minimal length property (minimizing interval lengths for smaller values of x at the expense of larger values). However, their intervals have a counterintuitive property:

the interval endpoints are not strictly increasing in x. For example, for $1-\alpha = .9$ and $x = 7,8,9,10,11,12,13$, they give lower endpoints of 3.589, *4.532, 4.532, 5.976, 5.976, 7.512, 7.512*, respectively. In fact, this property is not only counterintuitive but also nonoptimal.

Firstly, we make the assumption that any confidence procedure under consideration satisfies

$$\forall x,\ x\varepsilon[\ell_x,\ u_x]. \tag{2.8}$$

This condition can be interpreted, in general, as requiring that the MLE of θ (which is x itself in the Poisson case) is in the interval constructed. While we do not prove length non–optimality if (2.8) is violated, we suspect this is so. More importantly, all procedures we have examined satisfy (2.8), and making this assumption greatly reduces technical difficulties.

We would, of course, like to compare two $1-\alpha$ confidence procedures by comparing their interval lengths directly; however, it is not clear if dominance results can be established. But we can compare procedures by comparing sums of lengths, a practice used by Crow and Gardner (1959). Although we cannot compare two confidence procedures according to the total sum of their lengths, because such a sum is infinite, we can compare them in the following sense. We say that a $1-\alpha$ confidence procedure $C' = \{[\ell'_x, u'_x),\ x = 0,1,\ldots\}$ *dominates* another $1-\alpha$ confidence procedure $C = \{[\ell_x, u_x),\ x = 0,1,\ldots\}$ if there exists an integer N_o such that, for all $N > N_o$, either

i. $$\sum_{x=0}^{N}(u'_x - \ell'_x) < \sum_{x=0}^{N}(u_x - \ell_x)$$

or

ii. $$\sum_{x=0}^{N}(u'_x - \ell'_x) = \sum_{x=0}^{N}(u_x - \ell_x) \text{ and}$$
$$P_\theta(\theta\varepsilon C') \geq P_\theta(\theta\varepsilon C), \text{ with strict inequality for some } \theta.$$

With this definition, we have the following results about the ordering of interval endpoints.

Proposition 2.1 *If, for some integer k, the confidence procedure C has $\ell_k > \ell_{k+1}$, or $u_k > u_{k+1}$, then C can be dominated.*

Proof. Assume that $\ell_k > \ell_{k+1}$ (the proof is similar if $u_k > u_{k+1}$). Define a new confidence procedure C' by exchanging ℓ_k and ℓ_{k+1}, that is, C' is exactly the same as C with the exception that $\ell'_k = \ell_{k+1}$ and $\ell'_{k+1} = \ell_k$. Note that $\ell'_{k+1} < u_{k+1}$ because $u_{k+1} \geq k+1 > k \geq \ell_k$.

It is straightforward to calculate

$$P_\theta(\theta \epsilon C') - P_\theta(\theta \epsilon C) = I_{[\ell_{k+1}, \ell_k)}(\theta) \frac{e^{-\theta}\theta^k}{k!}(1 - \frac{\theta}{k+1}).$$

For $\theta \leq \ell_k$, we have $\theta \leq k$ (since $k \epsilon [\ell_k, u_k]$ by assumption). Hence, $P_\theta\ (\theta \epsilon C') - P_\theta(\theta \epsilon C) \geq 0$ for all θ, with strict inequality if $\ell_{k+1} \leq \theta < l_k$. □

A confidence procedure can, thus, be immediately improved by fixing any endpoints that are out of order. However, the above proposition does not apply to the case of the Crow–Gardner intervals, the intervals that initiated this particular investigation. The following proposition covers that case.

Proposition 2.2 *If, for some integer k, the confidence procedure C has $\ell_k = \ell_{k+1}$, or $u_k = u_{k+1}$, then C can be dominated.*

Proof. We again will only consider the case of lower endpoints, upper endpoints being handled in a similar fashion. Let $r \geq 1$ be such that $\ell_k = \ldots = \ell_{k+r} < \ell_{k+r+1}$ and assume that u_m is the smallest upper endpoint such that $u_m \geq \ell_k$. Two cases exhaust the possibilities.

Case I: $\ell_k = \ell_{k+r} < u_m$

In this case, we can produce a dominating procedure with uniformly shorter interval lengths. The coverage probability at $\theta = \ell_k = \ell_{k+r}$ is

$$\begin{aligned} P(\theta \epsilon C | \theta = \ell_k) &= \sum_{x=m}^{k+r} P(X = x | \theta = \ell_k) \\ &\geq \sum_{x=m}^{k} P(X = x | \theta = \ell_k) + P(X = k+1 | \theta = \ell_k). \end{aligned}$$

If it can be shown that the first term on the *rhs* is strictly greater than $1 - \alpha$, it will immediately follow that there exists a $\delta > 0$ so that the confidence procedure having $\ell_k + \delta$ replacing $\ell_{k+1}, \ldots, \ell_{k+r}$ is a $1 - \alpha$ confidence interval with shorter interval length.

If $k = 0$, $\ell_0 = 0$ and $m = 0$; therefore, $P(\theta \varepsilon C | \theta = 0) = 1 > 1 - \alpha$. Otherwise, at $\theta = \ell_k - \varepsilon$, we must have $P(\theta \varepsilon C | \theta = \ell_k - \varepsilon) > 1 - \alpha$, since C is a $1 - \alpha$ confidence procedure and the mimimum is attained at the endpoints. It then follows that

$$\sum_{x=m}^{k-1} P(X = x | \theta = \ell_k) = \lim_{\varepsilon \to 0} \sum_{x=m}^{k-1} P(X = x | \theta = \ell_k - \varepsilon) \geq 1 - \alpha,$$

which implies that

$$\sum_{x=m}^{k} P(X = x | \theta = \ell_k) > 1 - \alpha.$$

Therefore, in both cases, there exists $\delta > 0$ such that moving $\ell_{k+1}, \ldots, \ell_{k+r}$ to $\ell_k + \delta$ produces a $1 - \alpha$ confidence procedure with shorter interval length.

Case II: $u_m = \ell_k = \ell_{k+r}$. Here we consider two subcases.

Case IIa: $\ell_k^{k-m} m!/k! \leq 1$.

First note that at $\theta = u_m = \ell_k = \ell_{k+r}$, the coverage probability is $\sum_{x=m+1}^{k+r} e^{-\theta} \ell_k^x / x!$, and, as $\theta \uparrow \ell_k$, we have

$$\lim_{\theta \uparrow \ell_k} P_\theta(\theta \varepsilon C) = \sum_{x=m}^{k-1} e^{-\ell_k} \ell_k^x / x! \geq 1 - \alpha,$$

since C is a $1 - \alpha$ confidence procedure. Our new confidence procedure C' has $u'_m = \ell'_{k+1} = \ldots = \ell'_{k+r} = \ell_k + \delta$.

If $\ell_k \leq \theta < \ell_k + \delta$, the coverage probability of C' is

$$\sum_{x=m}^{k-1} e^{-\theta} \theta^x / x! + e^{-\theta} \theta^k / k!,$$

and if δ is small enough, we can keep $P_\theta(\theta \varepsilon C') \geq 1 - \alpha$. Furthermore, if $r > 1$, the length is smaller and if $r = 1$ and $\theta \varepsilon [\ell_k, \ell_{k+1} + \delta)$, the coverage probability of C' will be strictly greater if

$$\sum_{x=m}^{k} e^{-\theta} \theta^x / x! > \sum_{x=m+1}^{k+1} e^{-\theta} \theta^x / x! \tag{2.9}$$

For δ small enough, (2.9) holds if $\ell_k^{k-m} m!/k! \leq 1$, as $\ell_k/(k+1) < 1$.

Case IIb: $\ell_k^{k-m} m!/k! > 1$. In this case, we have $\ell_k^k/k! > \ell_k^m/m!$. So

$$1-\alpha \le \sum_{x=m}^{k-1} e^{-\ell_k} \ell_k^x/x! < \sum_{x=m+1}^{k} e^{-\ell_k} \ell_k^x/x!,$$

and we now consider a new confidence procedure C' where $\ell'_{k+1} = \ldots = \ell'_{k+r} = \ell_k + \delta$. For δ small enough, C' is also a $1-\alpha$ confidence procedure, but clearly has uniformly smaller length.

□

Proposition 2.2 shows that the Crow–Gardner rule of choosing the smallest upper bound from the non–unique acceptance regions results in an inadmissible procedure if the sum of lengths criterion is used. This proposition can also be interpreted as saying that a necessary condition for admissibility is that a confidence procedure have strictly increasing interval endpoints. We also note that Propositions 2.1 and 2.2 can be shown to apply to any discrete distribution with the property that the local minima of the coverage probability occur at the interval endpoints.

3. Refining Poisson Intervals

3.1 Difficulties.

Refinement is a method for systematically shortening a given $1-\alpha$ confidence procedure until it can be shortened no more. Refinement actually results in a family of intervals that constitute a complete class of non–randomized procedures.

In Casella (1986) a procedure for refining binomial confidence intervals was developed. The procedure developed here extends the previous one, and must deal with two additional difficulties. First, there is no natural invariance structure to take advantage of, as in the binomial case, but this turns out not to be too great of a problem. We must merely apply the refinement procedure to all endpoints, where in the binomial problem only lower endpoints need to be refined.

The second problem is that the sample space for the Poisson distribution is infinite, necessitating, in theory, the refinement of an infinite number of endpoints. In practice, we sidestep this issue by appealing to the normal approximation with continuity correction for

large X, as described in (2.5). This approximation is remarkably close to the refined intervals for large values of X, which was also noted by Crow and Gardner (1959). Therefore, we will only carry out the refinement procedure for a fixed number of endpoints.

3.2 Coincidental Endpoints.

The essence of the refinement procedure is to decrease upper endpoints and increase lower endpoints until they can be moved no further. In certain cases, an upper endpoint will be decreased until it hits a lower endpoint. When this happens, it may be the case that the upper endpoint can be decreased beyond the lower endpoint and $1-\alpha$ confidence can be maintained. If not, then the upper and lower endpoints become *coincidental*, i.e., they must have the same value. In Table 1, for $1-\alpha = .9$, the endpoints u_0 and ℓ_6 are coincidental and the tabled value is given as $u_0 = \ell_6 = 2.79$. However, this value is not unique, and u_0 can take any value in the interval $2.79 \pm .36$, as long as ℓ_6 takes the same value. This phenomenon of coincidental endpoints corresponds to the fact that in the two–sided testing problem there is sometimes more than one "shortest" acceptance region ("shortest" in the sense of containing the fewest x values). Coincidental endpoints appear in the Crow–Gardner tables, but are given unique values based on their rule of always choosing the acceptance region that yields the smallest upper endpoint. However, coincidental endpoints can occupy a range of values and $1-\alpha$ confidence can still be maintained. Therefore, refining a $1-\alpha$ confidence procedure actually results in a family of procedures, and an experimenter can choose the member of the family that best suits his needs.

3.3 The Refinement Procedure.

Starting with a $1-\alpha$ confidence procedure $C = \{[\ell_x, u_x),\ x = 0, 1, \ldots\}$, we create a refined procedure $C^* = \{[\ell_x^*, u_x^*),\ x = 0, 1, \ldots\}$. There are many ways to accomplish this, and if the original confidence procedure is very conservative, things tend to get somewhat complicated because there will be many instances when upper endpoints get decreased past lower endpoints. The following method is quite easy to implement, and seems to hold confusion to a minimum. The strategy is simple: starting with the smallest upper

endpoint, u_0, decrease it as far as possible while maintaining $1-\alpha$ confidence (this will probably entail decreasing it past some lower endpoints). Once u_0 is decreased as far as possible, increase, as far as possible, all of the lower endpoints below u_0. Now move to u_1 and repeat this process, then to u_2, etc. A more formal statement is:

The Refinement Procedure: To create the refined $1-\alpha$ confidence procedure $C^* = \{[\ell_x^*, u_x^*), x = 0, 1, \ldots\}$ from the $1-\alpha$ confidence procedure $C = \{[\ell_x, u_x), x = 0, 1, \ldots\}$.

0. Set $k = 0$.

1. Set $u = u_k$ and find m such that $\ell_m \leq u < \ell_{m+1}$.

2. Check if $u = \ell_m$ maintains $1-\alpha$ confidence.

   i. Calculate $T(\ell_m) = \sum_{x=k+1}^{m} e^{-\ell_m} \ell_m^x / x!$, the coverage probability, at $\theta = \ell_m$, of the confidence procedure with $u = \ell_m$.

   ii. If $T(\ell_m) \geq 1-\alpha$, set $u = \ell_m$ and go to 3.

   iii. If $T(\ell_m) < 1-\alpha$, solve for u^*, $\ell_m < u^* < u_k$, $T(u^*) = 1-\alpha$. Set $u_k^* = u^*$ and go to 4.

3. Check if $u < \ell_m$ maintains $1-\alpha$ confidence.

   i. Calculate $U(\ell_m) = \sum_{x=k+1}^{m-1} e^{-\ell_m} \ell_m^x / x!$.

   ii. If $U(\ell_m) > 1-\alpha$ then u can be decreased past ℓ_m. Find $0 < \delta < \ell_m - \ell_{m-1}$ such that $U(\ell_m - \delta) \geq 1-\alpha$. Set $u_k = \ell_m - \delta$ and go to 1.

   iii. If $U(\ell_m) \leq 1-\alpha$ then the endpoints u_k and ℓ_m are coincidental. Set $u_k^* = \ell_m^* = \ell_m$ and go to 4.

4. Next, refine all lower endpoints between u_k^* and u_{k-1}^* ($u_{-1} = 0$).

   i. Find n such that $\ell_{m-n-1} \leq u_{k-1}^* < \ell_{m-n}$.

   ii. For $j = 0, 1, \ldots, n$, solve for ℓ such that $\sum_{x=k}^{m-j-1} e^{-\ell} \ell^x / x! = 1-\alpha$. Set $\ell_{m-j}^* = \ell$.

   iii. Increment k and go to 1 if $k \leq K$.

5. After steps 1–4 have been run for $k = 0, 1, \ldots, K$, there will be a number of coincidental endpoints. As mentioned before, there is an allowable range of values for

each pair of coincidental endpoints. These ranges are now calculated. For each pair of coincidental endpoints, say $u_n^* = \ell_{n'}^* = t$, the upper and lower ends of the range must be found. This is most easily done by calculating two probabilities, one which considers the interval open at $\ell_{n'}$, and closed at u_n, and the other which considers the opposite situation. Calculate

$$P_1(r) = \sum_{x=n}^{n'-1} e^{-r} r^x / x!$$

$$P_2(r) = \sum_{x=n+1}^{n'} e^{-r} r^x / x!$$

$$m(r) = \min\{P_1(r),\ P_2(r)\}.$$

The allowable range for the coincidental endpoints is given by $[r_*, r^*]$ where

$$r^* = \max\{r :\ m(r) \geq 1 - \alpha\}$$

$$r_* = \min\{r :\ m(r) \geq 1 - \alpha\}.$$

Formally, one has to choose a value of K, the value at which to stop the algorithm. For large values of X, numerical studies indicate that the quadratic normal approximate confidence interval is quite close to the refined procedure with centered coincidental endpoints. Thus, one really only needs to refine intervals for small values of X. On the other hand, since computing is relatively cheap and quick, one can always find the refined interval for any observed X using the procedure in an interactive way.

Table 1 gives refined Poisson confidence intervals for $\alpha = .05$, and $x = 0, 1, \ldots, 49$. In each case of coincidental endpoints, the value tabled is the midpoint of the range. In addition, Figure 4 is a plot of the coverage probabilities of a refined procedure for $1 - \alpha = .9$. The procedure plotted is the one for which the coincidental endpoints are set to the midpoint of their range, and results in a procedure with fairly regular coverage probabilities.

3.4 Properties of Refined Intervals.

Because of the occurrence of coincidental endpoints, the refinement procedure actually produces a family of refined intervals. If we denote this family by R_α, we see that

the confidence procedures in R_α all agree at non–coincidental endpoints, but may have coincidental endpoints anywhere in the allowable range.

It is easy to see that the family R_α is a complete class, i.e., it contains all the admissible procedures. If $C \notin R_\alpha$, then it must be the case that C has at least one non–coincidental endpoint different from those procedures in R_α. Therefore, if the confidence procedure C is refined, there will be a non–trivial decrease in the length of at least one interval and hence the procedure C is inadmissible. The question of whether R_α is a *minimal* complete class (contains only admissible procedures) is still open. In general, the procedures in R_α are not comparable, but there is a possibility that some procedures in R_α are not admissible. Even if this is the case, however, the method of construction of the refined intervals make it seem unlikely that any major improvement can easily be made.

The observation that procedures in R_α are not comparable can be exploited by an experimenter to choose an optimal procedure from R_α. For example, if the experimenter had some prior knowledge that θ was near a certain value, he could set the coincidental endpoints to minimize interval size near the prior value. In the absence of any such information, choosing the coincidental endpoints at the middle of their range will produce a reasonable interval, one whose left and right–hand tail probabilities are approximately equal.

As mentioned before, Crow and Gardner (1959) regard the sum of the interval lengths as a criterion to consider, and assert that their intervals minimize the lengths for small x values at the expense of larger x values. The procedures in R_α are essentially equivalent on this criterion, since coincidental endpoints will cancel out of the sum of the interval lengths. Moreover, any confidence procedure that attains a minimum on this sum–of–length criteria must be in R_α.

4. Comments

The refinement procedure outlined in Section 3.3, although somewhat computer-intensive, describes a method of constructing a class of optimal confidence procedures. The method, with minor modification, can be adapted to any discrete distribution. For

example, consider the negative binomial distribution

$$P(X = x|\theta) = \binom{r+x-1}{r-1}\theta^r(1-\theta)^x, \qquad x = 0, 1, \ldots \tag{4.1}$$

where r is known and θ is unknown. The mean of this distribution is $\mu = r(1-\theta)/\theta$, and one can construct conservative $1-\alpha$ confidence intervals given by

$$xF_{2x,2(r+1),1-\alpha/2} \leq \mu \leq (x+1)F_{2(x+1),2(r+1),\alpha/2},$$

where $F_{m,n,\alpha}$ denotes the upper α cutoff from a F distribution with m and n degrees of freedom, and we define $F_{0,n,\alpha} \equiv 0$. The method of construction is similar to that of Clopper and Pearson or Garwood, and exploits the relationship between the binomial and beta distributions. Recall the relationship (which can be established by integrating the *rhs* by parts),

$$P(W \geq w|\theta) = P(T \leq \theta),$$

where W is binomial (n, θ) and T is beta $(w, n-w+1)$. One then uses the negative binomial-binomial relationship and the F-beta relationship to construct the intervals.

An easier method of construction, which seems to work reasonably well in practice, is to use a normal approximation with continuity correction. Recall that if the distribution of X is given by (4.1), then $EX = \mu$ and $\mathrm{Var}X = \mu + \frac{1}{r}\mu^2$. Lower and upper bounds are defined by

$$\begin{aligned} \ell_x^N &= \min\{\mu\colon\ (x - \mu - \frac{1}{2})^2/(\mu + \frac{1}{r}\mu^2) \leq z_{\alpha/2}^2\} \\ u_x^N &= \max\{\mu\colon\ (x - \mu + \frac{1}{2})^2/(\mu + \frac{1}{2}\mu^2) \leq z_{\alpha/2}^2\}. \end{aligned} \tag{4.2}$$

A plot of the coverage probabilities, for $r = 5$ and $1-\alpha = .9$, is given in Figure 5. The conservative nature of the intervals is clearly evident, although the intervals do, in fact, attain the $1-\alpha$ infimum. The refinement procedure was applied here, and coverage probabilities for the resulting intervals are given in Figure 6. It is easy to see that a large improvement is possible, especially for small values of x.

As mentioned before, practical improvements are obtained from the refinement procedure only for small values of x. As $x \to \infty$, the central limit theorem takes over and even fairly crude normal approximations are close to optimal. Intervals for small x values, however, are not well approximated, and that is the region where the refined intervals yield their greatest improvement. Producing tables of refined intervals for many discrete distributions would be quite unwieldy, since in many cases there are many parameters to vary (e.g., negative binomial, hypergeometric). However, it is quite easy and fast to produce the needed intervals from a computer program.

References

Blyth, C. R. and Hutchinson, D. W. (1961). Tables of Neyman-shortest unbiased confidence intervals for the Poisson parameter. *Biometrika* **48**, 191–194.

Casella, G. (1986). Refining binomial confidence intervals. *Canadian Journal of Statistics* **14**, 113–129.

Casella, G. and McCulloch, C. E. (1984). Confidence intervals for discrete distributions. Biometrics Unit Technical Report. No. BU-839-M. Cornell University, Ithaca, N.Y.

Clopper, C. J. and Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of binomial. *Biometrika* **26**, 404–413.

Crow, E. L. and Gardner, R. S. (1959). Confidence intervals for the expectation of a Poisson variable. *Biometrika* **46**, 441–453.

Garwood, F. (1936). Fiducial limits for the Poisson distribution. *Biometrika* **28**, 437–442.

Peizer, D. B. and Pratt, J. W. (1968). A normal approximation for binomial, F, beta, and other common, related tail probabilities, I. *J. Amer. Statist. Assoc.* **63**, 1416–1456.

Pratt, J. W. (1968). A normal approximation for binomial, F, beta, and other common, related tail probabilities, II. *J. Amer. Statist. Assoc.* **63**, 1457–1483.

Stevens, W. L. (1957). Shorter intervals for the parameter of the binomial and Poisson distributions. *Biometrika* **44**, 436–440.

Table 1: Refined Poisson confidence limits, $1-\alpha=.95$. For each value of x the first entry is the lower limit and the second entry is the upper limit. If an endpoint is coincidental, the x value of the coincidental point is noted and the range (±) is given.

x	Endpts	Coinc.Pts.	Range	x	Endpts	Coinc.Pts.	Range
0	0.00	–	–	25	16.77	–	–
	3.54	8	.25		36.59	49	.55
1	.05	–	–	26	16.98	9	.21
	5.49	11	.16		37.97	51	.29
2	.36	–	–	27	18.10	10	.47
	7.04	13	.35		39.03	52	.86
3	.82	–	–	28	19.05	–	–
	8.56	15	.45		40.25	53	1.23
4	1.37	–	–	29	19.51	11	.46
	10.04	17	.45		41.48	55	1.15
5	1.97	–	–	30	20.77	12	1.26
	11.51	19	.34		42.54	56	.79
6	2.61	–	–	31	21.35	12	1.17
	12.98	21	.16		43.95	58	.02
7	3.28	–	–	32	22.02	13	.80
	14.26	22	1.28		44.99	59	1.13
8	3.54	0	.25	33	23.43	–	–
	15.54	24	.66		46.05	60	.77
9	4.46	–	–	34	23.45	14	.02
	16.98	26	.21		47.46	62	.05
10	5.32	–	–	35	24.52	15	.81
	18.10	27	.47		48.49	63	1.15
11	5.49	1	.16	36	25.94	–	–
	19.51	29	.46		49.54	64	.80
12	6.69	–	–	37	25.95	16	.01
	20.77	30	1.26		50.95	66	.01
13	7.04	2	.35	38	27.01	17	.88
	22.02	32	.80		51.97	67	1.22
14	8.11	–	–	39	28.14	18	.39
	23.45	34	.02		53.03	68	.88
15	8.56	3	.45	40	28.97	–	–
	24.52	35	.81		54.12	69	.42
16	9.59	–	–	41	29.49	19	.52
	25.95	37	.01		55.46	71	.46
17	10.04	4	.45	42	30.57	20	.56
	27.01	38	.88		56.49	72	1.00
18	11.18	–	–	43	31.68	–	–
	28.14	39	.39		57.59	73	.58
19	11.51	5	.34	44	31.97	21	.30
	29.49	41	.52		58.96	75	.23
20	12.82	–	–	45	33.04	22	.76
	30.57	42	.56		59.99	76	1.16
21	12.98	6	.16	46	34.41	–	–
	31.97	44	.30		61.04	77	.79
22	14.26	7	1.28	47	34.45	23	.05
	33.04	45	.76		62.25	78	1.21
23	14.92	7	1.07	48	35.51	24	.98
	34.45	47	.05		63.47	80	1.32
24	15.54	8	.66	49	36.59	25	.55
	35.51	48	.98		64.51	81	1.02

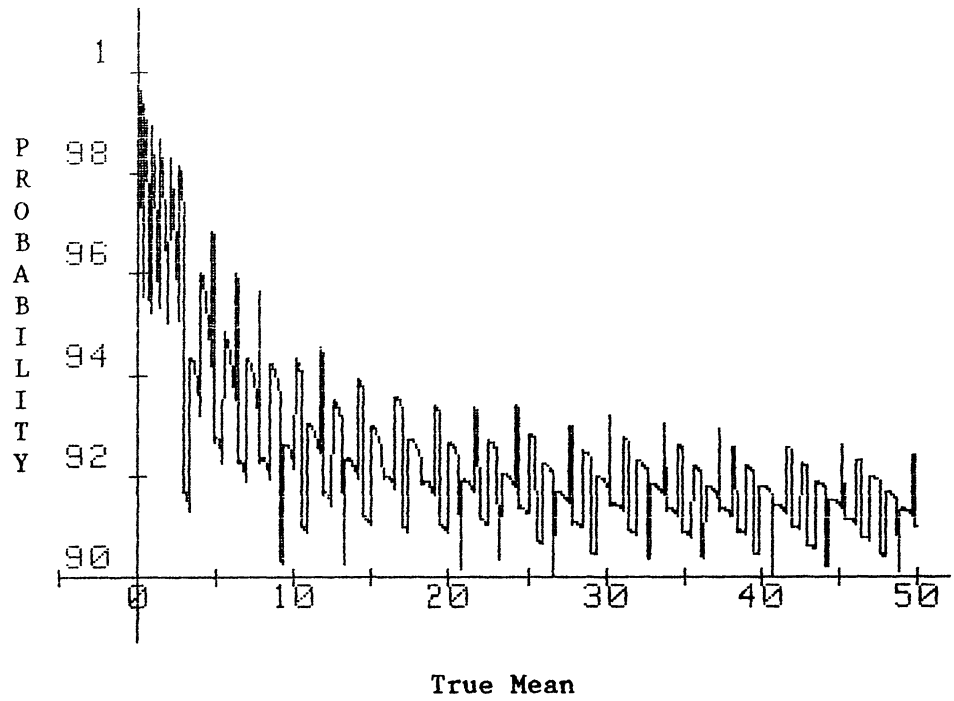

Figure 1: Coverage Probabilities for the Garwood Intervals, $1-\alpha=.9$

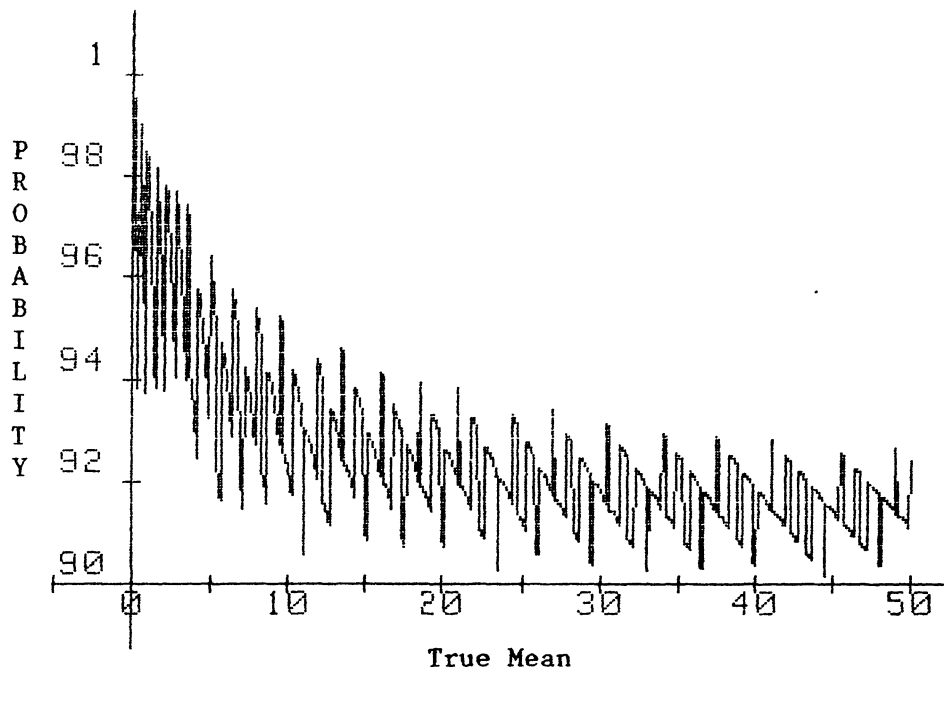

Figure 2: Coverage Probabilities for the approximate normal intervals given in (2.5), $1-\alpha=.9$.

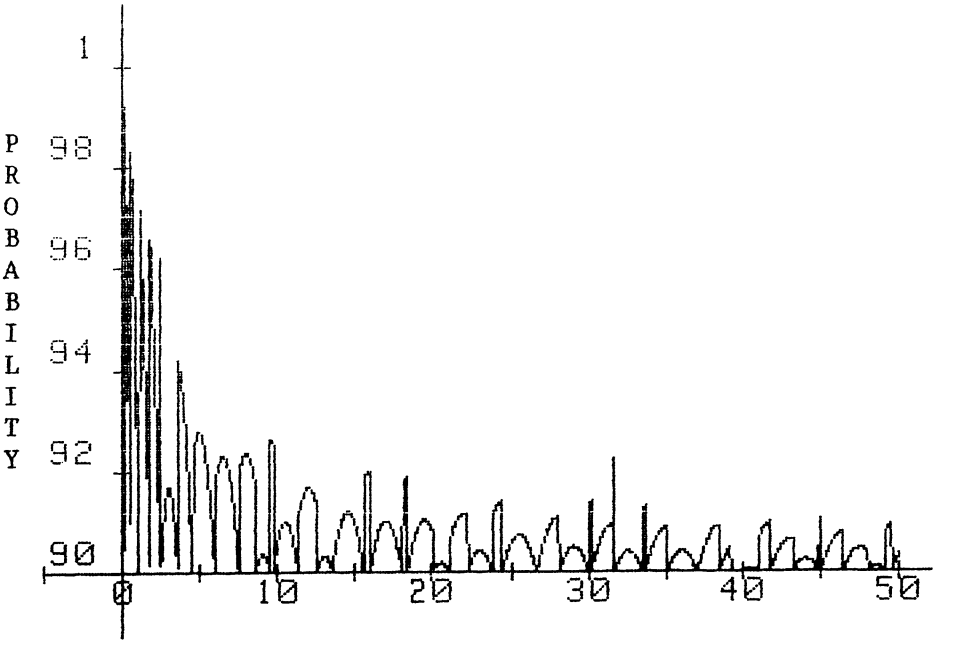

Figure 3: Coverage probabilities for the Crow-Gardner intervals, $1-\alpha=.9$

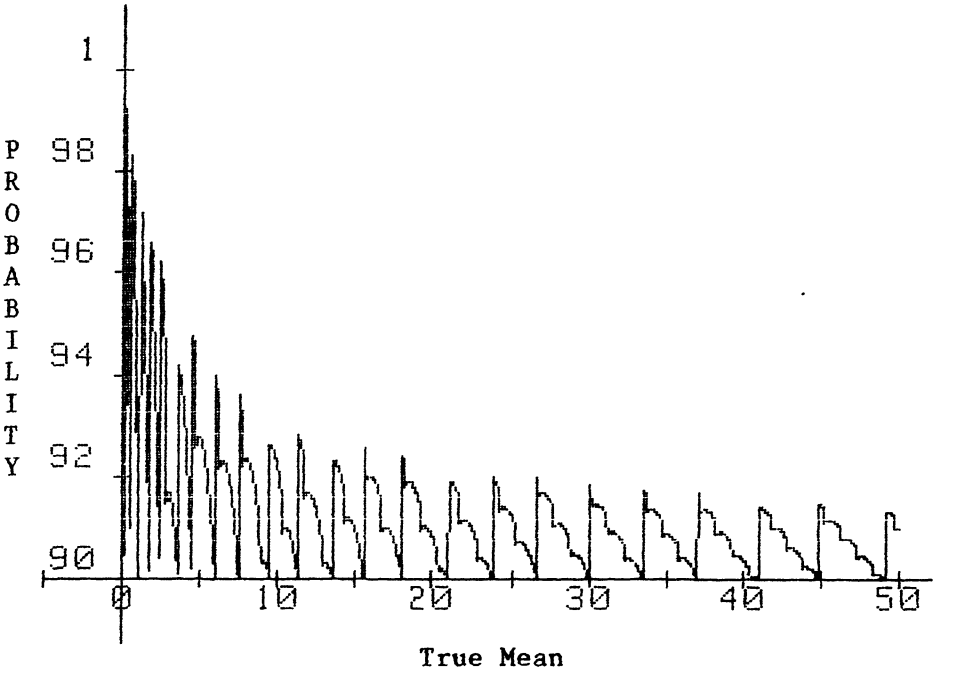

Figure 4: Coverage probabilities of the refined intervals, $1-\alpha=.9$.

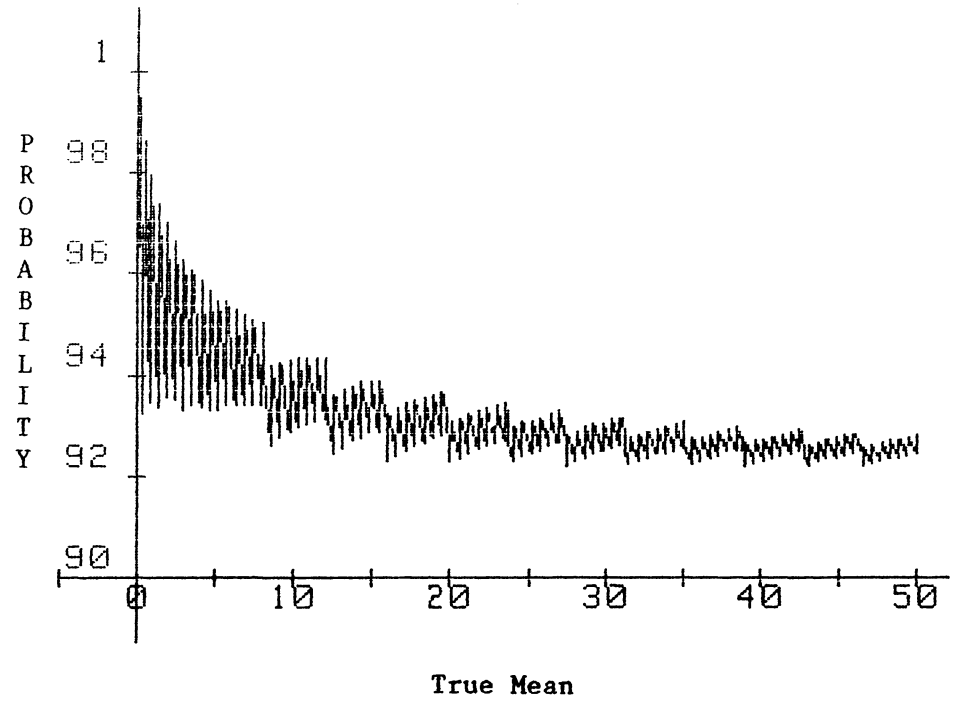

Figure 5: Coverage probabilities of the normal approximation (4.2) for the negative binomial distribution, $1-\alpha=.9$, $r=5$.

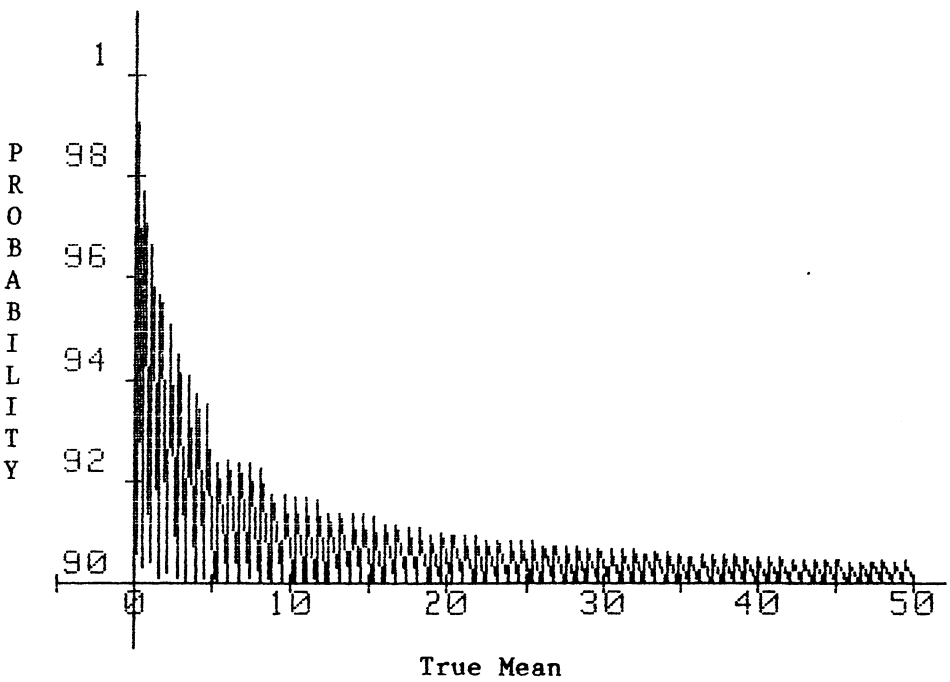

Figure 6: Coverage probabilities of the refined negative binomial intervals, $1-\alpha=.9$, $r=5$.