Traditional Formulas For Decision Level Are Wrong For Small Numbers of Counts

Jay A. MacLellan[1] and Daniel J. Strom[2]
[1]Radiation and Health Technology and [2]Risk Analysis and Health Protection, Process Technology & Environmental Management Resources, Pacific Northwest National Laboratory, Richland, Washington 99352-0999 USA
jay.a.maclellan@pnl.gov or daniel.j.strom@pnl.gov

Abstract: Traditional formulas for decision level (*DL*) and minimum detectable amount (*MDA*) are given in numerous sources, including the recent HPS.ANSI N13.30-1995, *Performance Standard for Radiobioassay* and the *Multi-Agency Radiation Survey and /Site Investigation Manual (MARSSIM)*. These formulas do not adequately account for the discrete nature of the Poisson distribution for paired blank (equal count times for background and sample) measurements, especially at low numbers of counts. We calculate the false positive rates that occur using the traditional *DL* formula as a function of acceptable false positive rate α and Poisson mean $\mu = \rho t$, where ρ is the underlying Poisson rate and t is the counting time. False positive rates exceed α by significant amounts for $\alpha \leq 0.2$ and μ <100 counts, peaking at 25% at $\mu \cong 0.71$, nearly independent of α. Calculations were verified by Monte Carlo simulations. The original 1968 derivation of the *DL* was based on knowing a good estimate of the mean and standard deviation of background, ad case that does not hold for paired blanks and low background rates. We propose and evaluate several alternative decision levels. Many regulations, national standards, guidance documents, and texts will have to be corrected.

Introduction

When counting particles, such as in alpha spectroscopy for measurement of ^{239}Pu, one typically subtracts an estimate of background counts from the counts of an unknown. The resulting difference or net count value can then be compared to a statistic called decision level, *DL*. If the net count value is greater than the *DL*, then one makes the decision that there is activity present above background.

The traditional formula for the decision level *DL* for the paired blank scenario as a function of N_b observed background counts and the acceptable false positive rate α is

$$DL(N,\alpha) = k_\alpha \sqrt{2N_b}, \qquad \text{(Eq. 1)}$$

where k_α is found from the cumulative Normal distribution:

$$1-\alpha = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{k_\alpha} e^{-x^2/2} dx \qquad \text{(Eq. 2)}$$

The Poisson distribution for Poisson mean $\mu = \rho t$, where ρ is the underlying Poisson rate (e.g., counts per second) and t is the counting time (e.g., seconds), is

$$Poi(N \mid \mu) = \frac{e^{-\mu}\mu^N}{N!} \qquad \text{(Eq. 3)}$$

Note that while N is an integer, μ is a non-negative real number.

Setting the decision level for bioassay excreta analyses is an important function of the internal dosimetry program. If it is too high, potentially significant intakes will be missed. If it is too low, resources will be wasted on unnecessary resampling and reanalyses. The sampled population may also loose confidence in the program if they suspect the internal dosimetrist is "fishing" for the right answer. It is therefore important to verify that the decision level (*DL*; also known as the critical level, L_C) employed is providing the desired results. It is given by

$$DL\,(N_n) = k_\alpha \sqrt{\sigma_g^2 + \sigma_b^2} = k_\alpha \sqrt{N_g + N_b} \qquad \text{(Eq. 4)}$$

and when there is no activity in the sample

$$DL = k_\alpha \sqrt{2\sigma_b^2} = k_\alpha \sqrt{2N_b}\,, \qquad \text{(Eq. 5)}$$

as shown in Eq. (1).

Methods

Two methods were used to determine the actual false positive rates α' when Eq.(1) is used for the paired blank counting problem.

Monte Carlo Simulation. The first method was a Monte Carlo simulation. For each of 57 values of μ (0.01 through 100), a Poisson distribution was randomly sampled. This value was stored as the background observation. Then the same distribution, this time representing an unknown containing no analyte, was sampled again and stored as the unknown. A *DL* was computed using Eq.(1), and the net rate (i.e., unknown – background) was compared to it. If the net result was greater than or equal to the *DL*, for that α, then the decision was "activity was detected above background." All such decisions are false positives, since there is no net activity present. This procedure was repeated 10^6 times for each mean and for each of 18 values of α (0.5, 0.2, 0.1, 0.05, etc., down to 10^{-6}). The results were slightly noisy, but were in exact agreement with the analytical method described below.

Analytical Solution. The cumulative Poisson distribution up through M is the sum of the Poisson distribution values:

$$\text{CumulPoi}(M,\mu) = \sum_{N=0}^{M} \text{Poi}(N,\mu) \qquad \text{(Eq. 6)}$$

The function Trunc(x) returns the integer part of nonnegative real number x. The false positive rate for a Poisson mean μ and so-called Type I error probability α is given by summing over nonnegative integers N of the product of two probabilities: the probability of observing a background value of N counts given a Poisson mean of μ; and the probability of observing more than N plus the expected background counts in the sample count. The later probability is simply one minus the cumulative Poisson distribution up to ($N + DL(N, \alpha)$). In symbolic terms, we have the actual false positive rate α' as

$$\alpha'(\mu,\alpha) = \sum_{N=0}^{\infty} \text{Poi}(N,\mu)(1 - \text{CumulPoi}([\text{Trunc}(N + DL(N,\alpha))],\mu) \qquad \text{(Eq. 7)}$$

Results

Actual false positive rates from Eq. (5), when counting blanks, are plotted in Figure 1. The horizontal axis is the long-term number of mean background counts $\mu_b = \rho_b t_b$ that one is trying to estimate when one counts a reagent blank. When subsequent blanks are counted, of course, any and all decisions that activity has been detected are wrong decisions, i.e., "false positives." If the decision that activity has been detected is based on the use of the ANSI N13.30 decision level, the false positive rates are shown on the vertical axis for various levels of the acceptable Type I error rate α. If the ANSI N13.30 formula were correct, each curve would be horizontal line equal to the value of α, independent of background rate. Clearly, the ANSI N13.30 decision level formula is not correct.

Figure 1 shows that the false positive rate is essentially independent of α below 0.3 counts, and if $\alpha \leq 0.2$, this is true almost up to $\mu_b = 1$. Figure 1 also shows that for values of α of 0.1 or less, the claimed false positive rate, that is, α, is not even achieved with a background rate μ_b of 100! For very tiny values of α, the rate is not even close. For α equal 10^{-6}, the false positive rate at 100 background counts is 25.1×10^{-6}. For $\alpha \leq 0.1$, the maxima of the curves are about 0.25, and occur near 0.7 to 0.72 counts, depending only weakly on the value of α. In the interval $0.3 \leq \mu \leq 1.3$, the false positive rate is above 0.2 regardless of the value of α.

The false positive rate for very low background rates using the Currie *DL* is due almost entirely to the probability of observing zero background counts. Regardless of the α value applied the square root of zero is zero, and the *DL* will be zero. Therefore, any observed count will be interpreted as "detected." For very low background rates, e.g., 0.01, one observes zero in about 99% of cases, and one in the other 1% of cases. Similar rates pertain for a blank about which one is trying to make an inference. Thus for those 1 in 100 blanks for which one observes 1 count, the probability is 99% that the paired background measurement will have been 0, and that a false positive decision will be made. The false positive rate is then 0.99×0.01, or approximately 0.01. For very low background rates the probability of observing a false positive with Currie's rule is approximately equal to the probability of observing one or more counts in the counting period.

How Can So Many People Have Been Wrong For Over 30 Years?

Eq. (1) was popularized by, and is generally ascribed to, Currie (1968). It appeared earlier in a more general form for count rate with count times not necessarily equal, as Rule D_2 in Nicholson (1963):

$$DL_{\text{Currie}}(R_n) = k_\alpha \sqrt{\frac{N_b}{t_b}\left(\frac{1}{t_b} + \frac{1}{t_g}\right)} \qquad \text{Eq.(9)}$$

and has appeared in one form or another in countless publications since, including ANSI/HPS N13.30-1996 and MARSSIM (Brodsky 1986; Currie 1968; Currie 1984; Health Physics Society (HPS) 1996; Hickey et al. 1993; Lochamy 1976; Strom and Stansbury 1992; MARSSIM 1997).

The problem arose from the basic assumption that one has a well-known estimate of the mean and standard deviation of the background. With low background rates one does not have a good estimate of either one. Currie's treatment of the "paired bland" case attempted to account for the increased uncertainty in the background when it was counted for only as long as the sample (as opposed to the well-known blank). But when the background rate is estimated form a measurement that is below the long-term mean, the use of Currie's *DL* (termed "critical level," L_c, in his 1968 paper) causes a large number of false positive decisions that are not offset by the fewer false positive decisions that result when the background rate is estimated form a measurement that is above the long-term mean. In particular, if the background is estimated from an observation of zero, one must decide that any result ≥ 1 results in a decision of "activity has been detected above background."

Currie stated on page 22 of his NUREG document (1984) that an assumption underlying the *DL* rule is that the estimated net signal is an independent random variable having a known distribution. Thus, knowing (or having a statistical estimate for) the standard deviation of the estimated net signal, one can calculate the *DL* given the distribution and alpha. He also stated on page 49 of the same document that if there are at least 5 counts in the background estimate, use of the Poisson variance as the estimate of the population variance is valid. Applying the rule to very low background count rates violates Currie's own assumptions for the *DL*. However, even at higher background levels, Currie's *DL* gives a false positive rate $\alpha' > \alpha$ because the estimate of the mean (and therefore the estimate of the variance) of the distribution is biased low. The background count is more like to give an estimate of the mean that is less than the true mean, than one that is larger than the true mean. The factor by which α' using Currie's *DL* overestimates α is particularly large for small α.

Other Decision Rules

Following our conclusion that Currie's formula does not give an unbiased estimate of false positive results, other methodologies were investigated.

Most Probable Value of Mean and Variance

Rainwater and Wu (1947) showed that the most probable values of the mean and variance are not the observed value of the mean, but a value larger than the observed value. Although not intuitively obvious, an example was given for clarification. If zero is observed, the mean is not necessarily also zero; therefore the average value of the mean that produces zero observations must be greater than zero, and the most probable value of the mean is larger than the observed value. One formal way of addressing this is to use an uninformative Bayesian prior probability distribution, which yields the result that the expectation value of background when *N* counts are observed is *N* + 1 (Friedlander and Kennedy 1955; Friedlander et al. 1963; Stevenson 1966; Little 1982). The variance is also *N* + 1. This leads to a decision level for the net count rate of

$$DL_{\text{Currie } N+1}(R_n) = k_\alpha \sqrt{\frac{(N_b+1)}{t_b}\left(\frac{1}{t_b} + \frac{1}{t_g}\right)} \qquad \text{Eq.(10)}$$

If the observed value *N* is much greater than one, the distinction between *N* and *N* + 1 is not important. One approach we took was to modify the Currie/ANSI N13.30 rule by using *N* + 1 as the estimate of the variance, instead of *N*.

Binomial Distribution

Nicholson (1963) and Sumerling and Darby (1981) describe equivalent processes for determining a decision level using the binomial distribution. As described in Sumerling and Darby (1981), if the background is not well known, the observation on the sample may be compared with the background observation, N_b counts in time t_b, to see if they are consistent with any single true count rate. They, and others, argue that using *both* the background and gross sample measurements to estimate the background increases the power of the test.

The probability of observing value N_g, when the mean of the quantity being measured is μ_g, is given by the Poisson distribution

$$P(N_g \mid \mu_g) = \frac{e^{-\mu_g}\mu_g^{N_g}}{N_g!} \qquad \text{(Eq. 11)}$$

The joint probability of making independent observations N_g and N_b when the respective means are μ_g and μ_b is given by

$$P(N_g, N_b \mid \mu_g, \mu_b) = \frac{e^{-\mu_g}\mu_g^{N_g}}{N_g!} \cdot \frac{e^{-\mu_b}\mu_b^{N_b}}{N_b!} \qquad \text{(Eq. 12)}$$

Transforming to new variables N_g and $N_{total} = N_g + N_b$,

$$P(N_g, N_{total} \mid \mu_{total}, Q) = \frac{e^{-\mu_{total}}\mu_{total}^{N_{total}}}{N_{total}!}\binom{N_{total}}{N_g} Q^{N_g}(1-Q)^{N_{total}-N_g} \qquad \text{(Eq. 13)}$$

where $\mu_{total} = \mu_g + \mu_b$ and $Q = \mu_g/(\mu_g + \mu_b)$. The probability of observing N_g conditional on a particular value of N_{total} is given by

$$P(N_g \mid N_{total}, \mu_{total}, Q) = \frac{P(N_g, N_{total} \mid \mu_{total}, Q)}{P(N_{total} \mid \mu_{total}, Q)} = \binom{N_{total}}{N_g} Q^{N_g}(1-Q)^{N_{total}-N_g} \qquad \text{(Eq. 14)}$$

Here, the binomial coefficient is denoted by $\binom{N_{total}}{N_g} = \frac{N_{total}!}{N_g!(N_{total}-N_g)!}$. This distribution is known as the binomial distribution with probability of success Q. If the sample is blank and N_g and N_b are both measurements of some unknown background with true count rate ρ_b, then $Q = Q_o$ where

$$Q_o = \frac{\rho_b t_g}{\rho_b t_g + \rho_b t_b} = \frac{t_g}{t_g + t_b} \qquad \text{(Eq. 15)}$$

For example, $Q_0 = ½$ when t_g equals t_b.

The inference about the presence of activity in the sample is based on the conditional distribution of N_g given N_{total}. Hence, the null hypothesis that the sample is blank is rejected if a blank sample would have produced a gross count as large or larger than the observed 100α% of the time or less, that is, if

$$\sum_{i=N_g}^{N_{total}} \binom{N_{total}}{N_g} Q_0^i (1-Q_0)^{N_{total}-i} \le \alpha \qquad \text{(Eq. 16)}$$

To use this rule in practice, one may simply compute the function on the left hand side of Eq. 16 to give the probability that the observed N_g was drawn from the same population as N_b for a given Q_0.

Confidence Interval of the Net Activity

The decision level may also be defined in terms of the confidence interval of the net activity. In his book *Atoms, Radiation, and Radiation Protection* Turner (1995) describes a decision level similar to one originally proposed by Altshuler and Pasternack, (1963). In this process, the decision is made on the basis of the difference in the gross (R_g) and background (R_b) count *rates*, the net count rate R_n:

$$\frac{N_g}{t_g} - \frac{N_b}{t_b} = R_g - R_b = R_n = k_\alpha\sqrt{\sigma_{gr}^2 + \sigma_{br}^2} = k_\alpha\sqrt{\frac{R_n + R_b}{t_g} + \frac{R_b}{t_b}} \qquad \text{(Eq. 17)}$$

Here, σ_{gr}^2 and σ_{br}^2 are the variances of the gross and background count rates, respectively. This is equivalent to Currie's detection level (minimum detectable count), when the decision level is set to zero. Solving the expression for R_n gives

$$DL_{\text{Turner}}(R_n) = \frac{k_\alpha^2}{2t_g} + \frac{k_\alpha}{2}\sqrt{\frac{k_\alpha^2}{t_g^2} + 4R_b\left(\frac{t_g + t_b}{t_g t_b}\right)}. \qquad \text{(Eq. 18)}$$

When $t_b = t_g$, the minimum significant count difference, Δ_1, is

$$\Delta_1 = k_\alpha\sqrt{2N_b}\left(\frac{k_\alpha}{\sqrt{8N_b}} + \sqrt{1 + \frac{k_\alpha^2}{8N_b}}\right) \qquad \text{(Eq. 19)}$$

Nicholson's D_1 Rule

Nicholson (1963) gives two other decision rules for the net count rate. Nicholson's D_1 Rule (1963) is k_α times the standard deviation of the net count rate:

$$DL_{\text{Nicholson } D_1}(R_n) = k_\alpha\sqrt{\frac{N_b}{t_b^2} + \frac{N_g}{t_g^2}} \qquad \text{Eq.(20)}$$

Nicholson states that D_1 uses the "obvious unbiased estimate of the variance with no restriction on" $\rho_n = \rho_g - \rho_b$.

Nicholson's D_3 Rule

Nicholson's D_3 Rule (1963) is k_α times the sum of the counts divided by the product of the count times:

$$DL_{\text{Nicholson } D_3}(R_n) = k_\alpha\sqrt{\frac{N_b + N_g}{t_b t_g}} \qquad \text{Eq.(21)}$$

Nicholson states that D_3 "optimally weights information about ρ_b in both t_b and t_b," but that its variance estimate is only

unbiased if the underlying net rate (due to activity in the sample), $\rho_n = 0$.

Decision Rule Comparisons

Each decision rule described above was evaluated with the Monte Carlo simulation method described for the Currie decision rule. However, only six values of α (0.05, 0.02, 0.01, 0.005, 0.002, and 0.001) were used. The results are shown in Figures 2 through 7.

For the ANSI N13.30 rule using $N + 1$ as the estimator of the background mean and variance, the nominal alpha values consistently underestimate the observed false positive rates for all $\mu_b > 2$. At background means less than one, the rule overestimates the false positives. This test is considered inadequate.

For the Nicholson/Sumerling and Darby binomial decision rule, the nominal alpha values overestimate the observed false positive rates for all background means, and grossly overestimate the false positives for $\mu_b < 10$. That is, using this decision rule results in a far smaller proportion of false positives than α. This test is considered inadequate for low background counting.

The Turner decision rule produces false positive rates that are relatively unbiased estimates of α down to $\mu_b = 4$ for the lowest α evaluated (0.001), and down to $\mu_b = 2$ with $\alpha = 0.05$. For $\mu_b < 2$, using the Turner decision rule results in a far smaller proportion of false positives than α. For equal count times, Nicholson's D_3 rule produces identical results, but not for unequal count times.

We have not completed our evaluation of Nicholson's D_1 rule, but it gives lower decision levels than D_3 or the binomial rule although greater decision levels than Currie's rule.

Recommendations

None of the rules evaluated provides an unbiased estimate of the false positive rate at all background means. Turner's decision rule comes the closest, but α is much greater than the actual proportion of false positives at the lowest background means. We are therefore presented with the conundrum in that our lowest background detectors may not be the most sensitive. Future work will investigate:

- Other existing decision rules
- The effect of varying the background and sample count time ratios
- Bayesian applications
- The effect on minimum detectable activity

References

Altshuler, B.; Pasternack, B. Statistical Measures of the Lower Limit of Detection of a Radioactivity Counter. Health Phys. 9:293-298; 1963.

Brodsky, A. Accuracy And Detection Limits For Bioassay Measurements In Radiation Protection - Statistical Considerations. NUREG-1156. Washington, DC: U.S. Nuclear Regulatory Commission; 1986.

Currie, L.A. Limits for Qualitative Detection and Quantitative Determination. Application to Radiochemistry. Analytical Chemistry 40(3):586-593; 1968.

Currie, L.A. Lower Limit Of Detection: Definition And Elaboration Of A Proposed Position For Radiological Effluent And Environmental Measurements. NUREG/CR-4007. Washington, DC: U.S. Nuclear Regulatory Commission; 1984.

Friedlander, G.; Kennedy, J.W. Nuclear and Radiochemistry. New York: John Wiley & Sons, Inc.; 1955.

Friedlander, G.; Kennedy, J.W.; Miller, J.M. Nuclear and Radiochemistry. 2nd edition. New York: John Wiley & Sons, Inc.; 1963.

Health Physics Society (HPS). Performance Criteria for Radiobioassay. An American National Standard. HPS N13.30-1996. New York: American National Standards Institute; 1996.

Hickey, E.E.; Stoetzel, G.A.; Strom, D.J.; Cicotte, G.R.; Wiblin, C.M.; McGuire, S.A. Air Sampling in the Workplace. Final Report. NUREG-1400. Washington, DC: U.S. Nuclear Regulatory Commission; 1993.

Little, R.J.A. The Statistical Analysis of Low-Level Radioactivity in the Presence of Background Counts. Health Phys. 43(5):693-703; 1982.

Lochamy, J.C. The Minimum-Detectable-Activity Concept. pp. 169-172 in: Fivozinsky, S.P., ed. Measurements for Safe Use of Radiation. NBS Special Publication 456; Washington,D.C.: National Bureau of Standards; 1976.

Nicholson, W.L. Fixed Time Estimation of Counting Rates with Background Corrections. AEC R and D Report HW-76279; 1963. See also Nicholson, W.L. Statistics of Net-Counting-Rate Estimation with Dominant Background Corrections. Nucleonics 24(8):118-121; 1966.

Rainwater, L.J.; Wu, C.S. Applications of Probability Theory to Nuclear Particle Detection. Nucleonics 1(October):60-69; 1947.

Stevenson, P.C. Processing of Counting Data. NAS-NS-3109. Livermore, California: National Academy of Sciences -- National Research Council; 1966.

Strom, D.J.; Stansbury, P.S. Minimum Detectable Activity when Background Is Counted Longer than the Sample. Health Phys. 63(3):360-361; 1992.

Sumerling, T.J.; Darby, S.C. Statistical Aspects of the Interpretation of Counting Experiments Designed to Detect Low Levels of Radioactivity. NRPB-R113. Chilton UK: National Radiological Protection Board; 1981.

Turner, J.E. Atoms, Radiation, and Radiation Protection. 2nd edition. New York: John Wiley & Sons; 1995.

U.S. NRC; U.S. EPA; U.S. DOE; U.S. DoD. Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM; http://www.epa.gov/radiation/marssim/). Washington, DC: U.S. Government Printing Agency; 1997

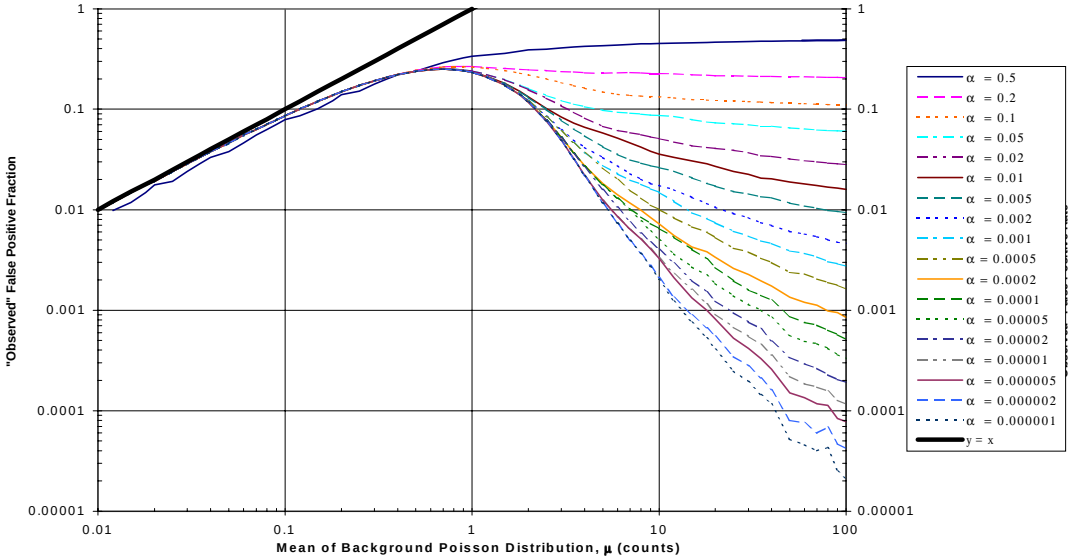

Figure 1 "Observed" False Positives with Given ANSI N13.30 Alpha Values

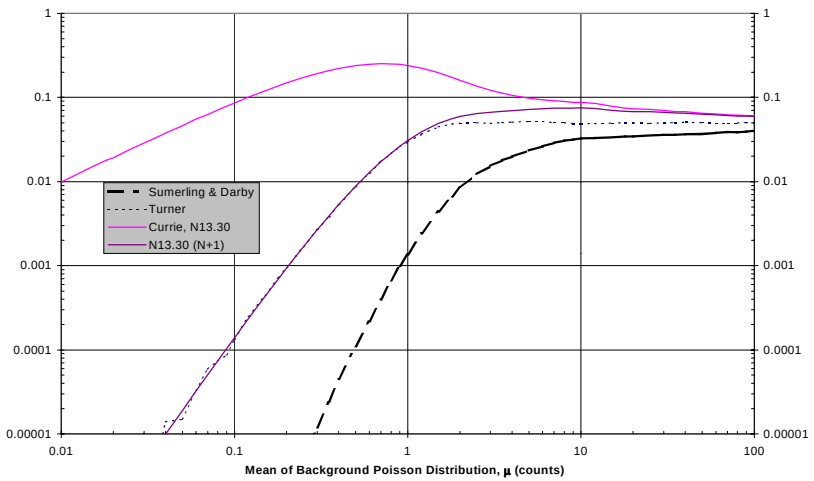

Figure 2 "Observed" False Positives with Various Decision Rules (α = 0.05)

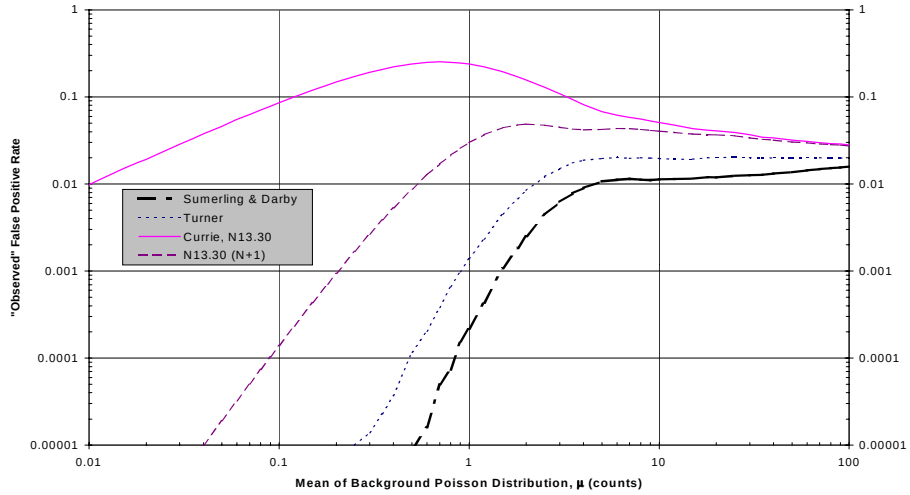

Figure 3 "Observed" False Positives With Various Decision Rules (α= 0.02)

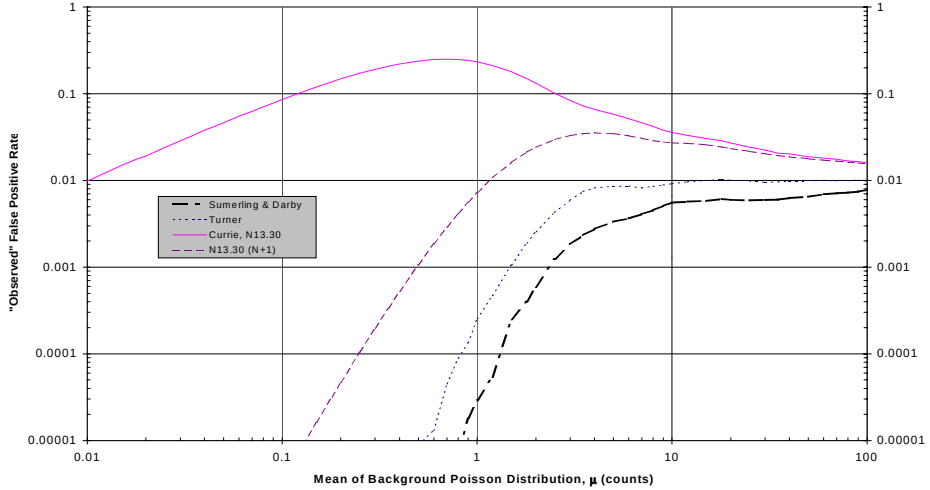

Figure 4 "Observed" False Positives With Various Decision Rules (α= 0.01)

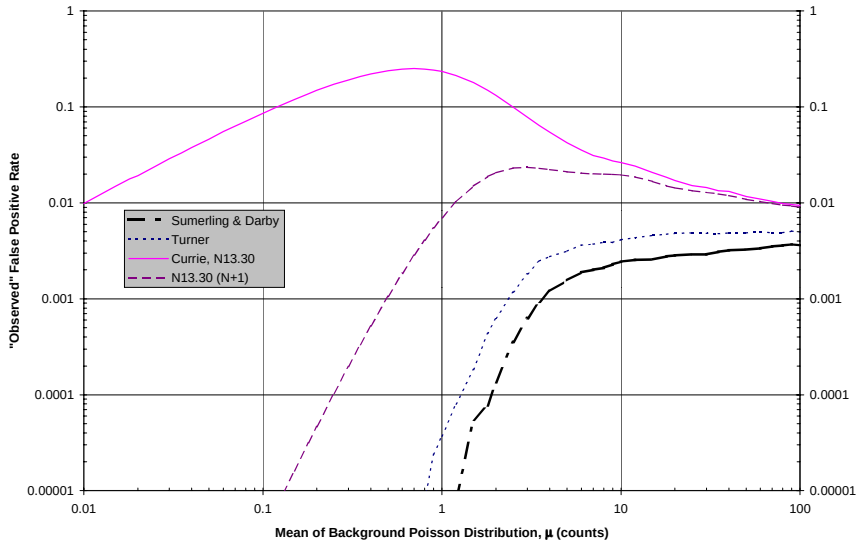

Figure 5 "Observed" False Positives With Various Decision Rules (α= 0.005)

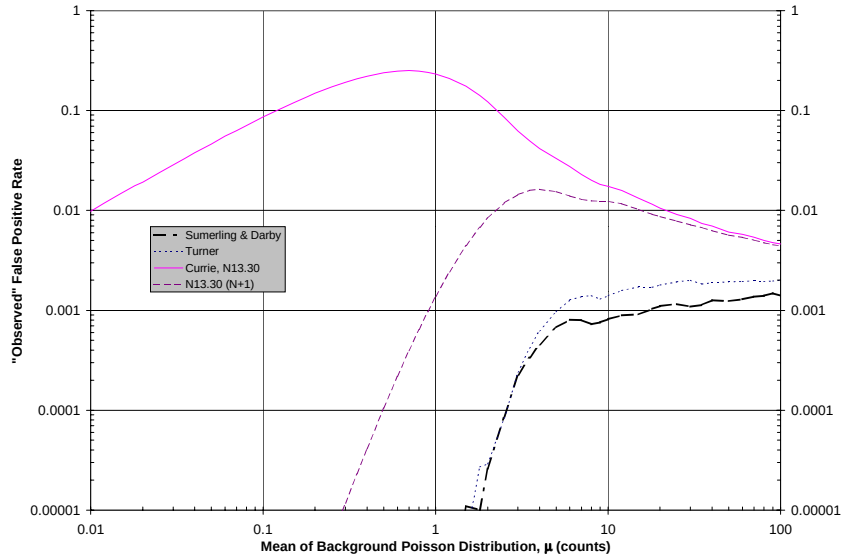

Figure 6 "Observed" Fase Positives With Various Decision Rules (α= 0.002)

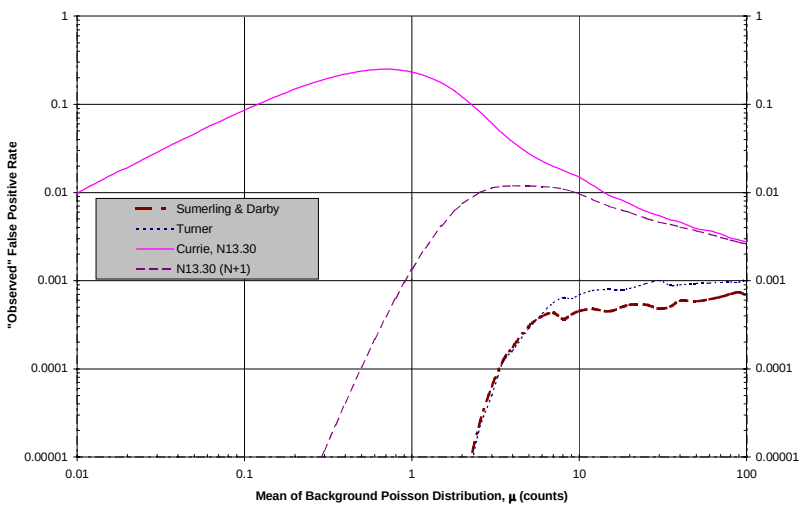

Figure 7 "Observed" False Positives With Various Decision Rules (α= 0.001)