Improved approximations of Poissonian errors for high confidence levels

Harald Ebeling

Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA

Accepted 2002 December 20. Received 2002 December 19; in original form 2002 March 25

ABSTRACT

We present improved numerical approximations to the exact Poissonian confidence limits for small numbers $n$ of observed events following the approach of Gehrels. Analytic descriptions of all parameters used in the approximations are provided to allow their straightforward inclusion in computer algorithms for processing of large data sets. Our estimates of the upper (lower) Poisson confidence limits are accurate to better than 1 per cent for $n \lesssim 100$ and values of $S$, the derived significance in units of Gaussian standard deviations, of up to 7 (5). In view of the slow convergence of the commonly used Gaussian approximations towards the correct Poissonian values, in particular for higher values of $S$, we argue that, for $n \lesssim 40$, Poissonian statistics should be used in most applications, unless errors of the order of, or exceeding, 10 per cent are acceptable.

Key words: methods: numerical – methods: statistical.

1 INTRODUCTION

The need to assess the statistical significance of an observed small number of events is common in astrophysics and in virtually all other natural sciences. The number of neutrinos detected in an underground detector, the number of supernovae observed at $z > 1$, the number of photons in a faint X-ray point source – all of these are numbers that are probably in the Poisson regime, and knowing the correct errors of these measurements is obviously crucial to any scientific conclusions drawn from these numbers.

In a much noted paper Gehrels (1986, hereafter G86) provided analytic approximations to the correct Poisson confidence limits for small event numbers. A particularly useful new and improved approximation in G86 is for the Poisson lower confidence limit, $\lambda_l$ (see Section 2 below for an overview of the nomenclature used). However, two issues limit the applicability of these approximations to large data sets and/or applications that require high confidence levels. First, Gehrels’ analytic approximation to the correct value of $\lambda_l$ uses empirically determined parameters $\beta$ and $\gamma$, both of which are non-trivial functions of $S$, the desired significance in units of Gaussian standard deviations. G86 tabulates $\beta(S)$ and $\gamma(S)$ for 10 values of $S$ ranging from 1 to 3.3$\sigma$ but does not provide an analytic formula that would allow the reader to compute $\lambda_l$ for arbitrary $S$ values.\(^1\) Secondly, the validity of all approximations presented and discussed in G86 has only been verified in the same, relatively narrow range of confidence levels from 1 to 3.3$\sigma$. To our knowledge, no reasonably accurate approximations have been published for $S > 3.3\sigma$.

While these issues may be of limited importance for many, if not most, applications, they can become serious in cases where confidence limits need to be computed for large sets of event numbers, particularly if high confidence levels are required. Consider, for instance, X-ray astronomy, a traditionally photon-starved line of observational research. With large-area, high-resolution X-ray charge-coupled device (CCD) detectors now in use on-board the Chandra and XMM–Newton X-ray Observatories, X-ray images of dimensions $1024 \times 1024$ or even $4096 \times 4096$ pixels have become common. The numbers of photons registered in the vast majority of these pixels will be in the Poisson regime, and assessing accurately the significance of any features embedded in the very low and spatially non-uniform background measured with these large arrays requires the accurate computation of a considerable number of Poisson confidence limits. A real-life example of a scientific project relevant in this context is the compilation of a statistical sample of unresolved (single-pixel) point sources detected at the greater than 5$\sigma$ confidence level (i.e. $S = 5$) in a set of Chandra ACIS-I images.

It is for applications such as those outlined above that here we present modified and improved versions of G86’s approximations that are more accurate over an extended range of confidence levels ($S \lesssim 7\sigma$). To permit these approximations to be incorporated straightforwardly into computer algorithms for the processing of large data sets, we also provide analytic descriptions of all parameters used, thus allowing the computation of accurate Poisson confidence limits for a wide range of event numbers and confidence levels.

2 NOMENCLATURE

In the following we adopt Gehrels’ nomenclature and definitions. Specifically, the upper limits $\lambda_u$ and the lower limits $\lambda_l$ for Poisson statistics are defined by

\[^{\star}\text{E-mail: ebeling@ifa.hawaii.edu}\]

\[^{\dag}\text{Especially relevant in view of the fact that }\gamma(S)\text{ actually becomes singular close to }S = 1,\text{ as we show in Section 5.}\]
\[ \sum_{x=0}^{n} \frac{\lambda^x e^{-\lambda}}{x!} = 1 - CL, \quad n \geq 0 \] (1)

and

\[ \sum_{x=0}^{n-1} \frac{\lambda^x e^{-\lambda}}{x!} = CL, \quad n \geq 1, \] (2)

where \( n \) is the number of events observed, and \( CL \) is the desired confidence level (equations 1 and 2 of G86). (For \( n = 0 \), \( \lambda = 0 \) for all values of \( CL \).) For Gaussian statistics (i.e. a probability distribution that is normal) \( CL \) is related to \( S \), the equivalent Gaussian number of \( \sigma \), by

\[ CL(S) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{S} e^{-t^2/2} dt. \] (3)

For ease of presentation we shall use \( S \) to parametrize a large range of \( CL \) values.

For \( n = 0 \) to 100 and selected values of \( S \) ranging from 1 to 3.3 G86 tabulates \( \lambda_u \) and \( \lambda_l \) as obtained from equations (1) and (2); Gehrels also presents analytic and numerical approximations accurate to better than 2 per cent for \( S < 3.3 \). In this paper, we test the Gaussian approximation and those discussed in G86 over a larger range of confidence levels (\( S \leq 7 \)). We then modify Gehrels’ approximations to improve their performance specifically for large values of \( S \) and, finally, present numerical (polynomial) descriptions of all parameters used, to allow the computation of approximate values of \( \lambda_u(n, S) \) and \( \lambda_l(n, S) \) for all values of \( n \) up to at least 100 and \( S \leq 7 \).

3 THE GAUSSIAN APPROXIMATION

For a normal probability distribution (i.e. in the limit \( \rightarrow \infty \)) \( \lambda_u \) and \( \lambda_l \) are given by the well-known expressions \( n \pm S\sqrt{n} \), which are also commonly used to approximate the not straightforwardly computable Poisson limits for ‘reasonably’ large values of \( n \). What ‘reasonable’ means in this context is subject to debate; in practice, values of \( n \) in excess of 20 (and sometimes even \( n > 10 \)) are often deemed sufficiently large to justify the use of the Gaussian approximation.

In Fig. 1, we show the percentage error of the Gaussian approximation \( n \pm S\sqrt{n} \) when applied to event numbers in the Poisson regime (loosely defined as \( n < 100 \)). For low to moderate confidence levels (\( S \leq 2 \)) the Gaussian approximation is accurate to better than 10 per cent for \( n = 20 \) (but not for \( n = 10 \)), which may or may not be sufficient for a given application. However, the error of the Gaussian approximation, in particular for the lower limit, increases rapidly for higher confidence levels. Already at \( S = 3, n > 45 \) is required to limit the error in \( \lambda_l \) to 10 per cent; for \( S = 5 \) even \( n = 100 \) is insufficient if 10 per cent accuracy is sought. 10 per cent accuracy may or may not be nearly good enough when in applications where errors of many independent measurements are propagated. In the case of X-ray spectral fitting, for instance, theoretical models are fitted simultaneously to events registered in hundreds of independent energy channels. A systematic error of the order of 10 per cent in the errors on the counts in all spectral bins introduced by the use of Gaussian approximations to the true Poisson errors can lead to best-fitting parameter values that may be erroneous by significantly more than the formal, statistical errors.

4 APPROXIMATION OF THE POISSONIAN UPPER LIMIT

G86 derives two analytic approximations to the true Poissonian upper limit, namely

\[ \lambda_u \approx (n + 1) \left[ 1 - \frac{1}{9(n + 1)} + \frac{S}{3\sqrt{n + 1}} \right]^3 \] (G86 eq. 9)

and

\[ \lambda_u \approx n + S\sqrt{n + 1} + \frac{S^2 + 2}{3}. \] (G86 eq. 10)

where the latter expression is simply the previous one expanded and limited to the dominant terms in \( n + 1 \).

In Fig. 2 we show the percentage error of these two approximations for a range of \( n \) and \( S \) values. Both expressions represent a considerable improvement over the Gaussian approximations (cf. Fig. 1). At low values of \( n \) the remaining error of several per cent may, however, be too high for certain applications. Following the approach taken by Gehrels to improve his approximation to \( \lambda_l \) (see the following section) we therefore apply a heuristic correction to G86 equation (9) in the form of an additional term \( b(n + 1)^c \):

\[ \lambda_u \approx (n + 1) \left[ 1 - \frac{1}{9(n + 1)} + \frac{S}{3\sqrt{n + 1}} + b(n + 1)^c \right]^3. \] (4)

We determine \( b = b(S) \) from the requirement that the above approximation should be an identity for \( n = 0 \), i.e.

\[ b(S) = \lambda_u(0, S)^{1/3} - 1 + 9/9 - S/3, \] (5)

thus forcing better performance for low values of \( n \). For each value of \( S \) from 0 to 7, \( c = c(S) \) is then chosen such that the error of the above approximation is minimized for \( 0 \leq n \leq 100 \). The runs of \( b \) and \( c \) as functions of \( S \) are shown in Fig. 3. We find \( c \) to be negative for all values of \( S \) with one-sided singularities at the roots of equation (5), which lie at \( S_{0.1} = 0.50688 \) (corresponding to a confidence level of 69.388 per cent) and \( S_{0.2} = 2.27532 \) (confidence level of 98.856 per cent).

To allow the evaluation of equation (4) for any number of observed events, \( n \), and any confidence level, \( S \), specifically in the proximity of the mentioned singularities, we fit piecewise polynomial functions to \( b(S) \) and \( c(1/S) \), or \( c(\log_{10} S) \), with the degree of the polynomial being determined by the requirement that the absolute of the residuals be less than 1 per cent over the full \( S \) range of the fit, except at the locations of the mentioned singularities. We find acceptable polynomial descriptions of \( b(S) \) and \( c(S) \) as follows:

\[ b(S) = \sum_{i=0}^{8} b_i S^i \] (6)

and

\[ c(S) = \begin{cases} \sum_{i=0}^{4} c_{1,i} \left( \frac{1}{S - S_{0.1}} \right)^i & S < S_{0.1} \\ \sum_{i=0}^{3} c_{2,i} \log_{10}(S - S_{0.1})^i & S_{0.1} < S < 1.2 \\ \sum_{i=0}^{3} c_{3,i} \left( \frac{1}{S - S_{0.2}} \right)^i & 1.2 < S < S_{0.2} \\ \sum_{i=0}^{4} c_{4,i} \log_{10}(S - S_{0.2})^i & S > S_{0.2} \end{cases} \] (7)
5 APPROXIMATION OF THE POISSONIAN LOWER LIMIT

As demonstrated by Fig. 1 the Gaussian approximation \( \lambda_l(n, S) \approx n - S \sqrt{n} \) is a poor one for all but the lowest values of \( S \). G86 explores the behaviour of several more sophisticated analytic approximations before resorting to modifying the most promising of them by adding

---

Figure 1. Percentage error of the Gaussian approximations \( \lambda_u \approx n + S \sqrt{n} \) (top) and \( \lambda_l \approx n - S \sqrt{n} \) (bottom) as a function of \( n \). In each panel the \( S \) values of the shown curves vary from lower left to upper right in steps of 0.5 as indicated; the thick lines correspond to \( S = 1, 2 \) and 3. Note how, for \( n = 10 \) (marked by the dotted line), the errors still reach and exceed 10 per cent for all but the lowest values of \( S \).
a heuristic power-law term $\beta n^\gamma$ (we used the same approach in the preceding section to improve the approximation to $\lambda_u$):

$$
\lambda_1 \approx n \left(1 - \frac{1}{9n} - \frac{S}{3\sqrt{n}} + \beta n^\gamma\right)^3.
$$

(G86 eq. 14)

To find $\beta(S)$ and $\gamma(S)$ we proceed similarly to before for our approximation to $\lambda_u$ and define $\beta(S)$ as

$$
\beta(S) = \lambda_1(1,S)^{1/3} - 1 + 1/9 + S/3,
$$

(8)

and then determine $\gamma(S)$ such that the error of the above approximation is minimized for $0 \leq n \leq 100$. The result, $\beta$ and $\gamma$ as functions of $S$, is shown in Fig. 5 which, in the overlap region, agrees with fig. 1 of G86. $\gamma$ is negative for all values of $S$ with a one-sided singularity at the only root of equation (8) at $S_0 = 0.93876$, corresponding to a confidence level of 82.607 per cent.

To facilitate the evaluation of equation (14) of G86 for a wide range of values of $n$ and $S$, we attempt to find analytical expressions for $\beta(S)$ and $\gamma(S)$ (G86 quotes the values of either function only at 10 locations between $S = 1$ and 3.291). In analogy to the approach taken in the preceding section, we fit piecewise polynomial functions to $\beta(S)$, and to $\gamma(S)$, $\gamma(1/S)$ or $\gamma(\log_{10} S)$, with the degree of the polynomial being determined by the requirement that the absolute
Improved approximations of Poissonian errors

Figure 3. Run of parameters \( b \) (top) and \( c \) (bottom) of equation (4) as a function of \( S \), the equivalent number of Gaussian \( \sigma \). \( c \) exhibits singularities at \( S = 0.507 \) and 2.275 where \( b = 0 \). The bullets mark the locations at which \( b(S) \) and \( c(S) \) were computed, the solid lines mark polynomial fits to the data (see the text for details).

of the residuals be less than 2 per cent over the full \( S \) range of the fit (less than 0.1 per cent at the high-\( S \) end where high accuracy is critical). We find acceptable polynomial descriptions of \( \beta(S) \) and \( \gamma(S) \) as follows:

\[
\beta(S) = \begin{cases} 
\sum_{i=0}^{5} \beta_{1i} S^i & S \leq 3 \\
\sum_{i=0}^{4} \beta_{2i} S^i & S > 3 
\end{cases}
\]

and

\[
\gamma(S) = \begin{cases} 
\sum_{i=0}^{5} \gamma_{1i} \log_{10}(S_0 - S) & S \leq S_0 \\
\sum_{i=0}^{5} \gamma_{2i} \left( \frac{1}{S - S_0} \right)^j & S_0 < S \leq 2.7 \\
\sum_{i=0}^{2} \gamma_{3i} S^i & S > 2.7 
\end{cases}
\]
Table 1. Coefficients of polynomial fits to $b(S)$ and $c(S)$ of equation (4), as defined in equations (6) and (7).

| $i$ | $b_i$     | $c_{1,i}$ | $c_{2,i}$ | $c_{3,i}$ | $c_{4,i}$ |
|-----|-----------|-----------|-----------|-----------|-----------|
| 0   | -3.8954E−03 | -2.0799E+2 | -1.4354E+1 | -8.4098E−01 | -1.0120E+1 |
| 1   | +6.2328E−6  | -7.1925E−01 | -6.3188E−01 | +6.8766E−01 | -2.8853E−01 |
| 2   | +5.2345E−5  | -4.0064E−01 | -1.6177E−01 | +2.0358E−02 | +4.2013E−04 |
| 3   | -5.3096E−03 | -7.3386E−02 | -5.6966E−01 | +3.9965E−01 | -5.3310E−02 |
| 4   | +1.3093E−1  | -5.4791E−03 | -2.2835E−01 | -1.6319E−02 |          |
| 5   | -2.0344E−04 |          |          | +4.8667E−04 |          |
| 6   | +2.0393E−2  |          |          | -5.5299E−02 |          |
| 7   | -1.974E−06  |          |          | -3.3361E−02 |          |
| 8   | +3.1161E−3  |          |          |          |          |

Figure 4. Percentage error of the approximations of equation (4) with $b(S)$ and $c(S)$ as computed (top), and using the polynomial fits of equations (6) and (7) (bottom), as a function of $n$. In each panel the $S$ values of the shown curves vary from 0.5 to 7 as annotated. For all values of $n$ and $S$ explored here the error remains below the 0.5 per cent level.
Improved approximations of Poissonian errors

Figure 5. Run of parameters $\beta$ (top) and $\gamma$ (bottom) of G86 equation (14) as a function of $S$, the equivalent number of Gaussian $\sigma$. $\gamma$ exhibits a singularity at $S = 0.939$ where $\beta = 0$. The bullets mark the locations at which $\beta(S)$ and $\gamma(S)$ were computed, the solid lines mark polynomial fits to the data (see the text for details).

with $-50 \leq \gamma(S) \leq 0$ overriding the above where necessary, and coefficients $\beta_{1,i}$, $\beta_{2,i}$, $\gamma_{1,i}$, $\gamma_{2,i}$ and $\gamma_{3,i}$ as listed in Table 2. The results of these fits are shown as the solid lines in Fig. 5.

Equation (14) of G86 indeed yields greatly reduced errors when compared with the exact values of $\lambda_i$. For $n < 100$ and $1 < S < 3.291$ G86 quotes an accuracy of better than 2 per cent for the above approximation. Fig. 6 confirms this, but also demonstrates that the errors become unacceptably large ($>10$ per cent) for higher confidence levels and small to moderate values of $n$.

We now attempt to improve on equation (14) of G86 by adding a second, higher-order correction term. As illustrated by Fig. 6 such an additional term would have to improve the performance of the approximation particularly in the high-$S$ regime for which equation (14) of G86 was not optimized. This goal can be achieved by
Table 2. Coefficients of polynomial fits to $\beta(S)$ and $\gamma(S)$ of G86, equation (14), as defined in equations (9) and (10).

| $i$ | $\beta_{1,i}$ | $\beta_{2,i}$ | $\gamma_{1,i}$ | $\gamma_{2,i}$ | $\gamma_{3,i}$ |
|-----|----------------|----------------|----------------|----------------|----------------|
| 0   | $-3.8605899 \times 10^{-3}$ | $+3.4867327 \times 10^{-01}$ | $-1.7480435$ | $-0.6347351$ | $-2.7517416 \times 10^{00}$ |
| 1   | $-6.6002964 \times 10^{-03}$ | $-4.0996949 \times 10^{-01}$ | $-1.8895824$ | $-4.6707845$ | $+3.1692400 \times 10^{-01}$ |
| 2   | $+6.5798149 \times 10^{-03}$ | $+1.6514495 \times 10^{-01}$ | $-3.0808786$ | $+6.1602866$ | $-8.7788310 \times 10^{-03}$ |
| 3   | $+2.8172041 \times 10^{-03}$ | $-1.5783156 \times 10^{-02}$ | $-5.5164953$ | $-4.3543401$ | $+1.4470675$ |
| 4   | $+2.9892915 \times 10^{-03}$ | $+5.2768918 \times 10^{-01}$ | $-3.9940504$ | $+1.4470675$ | $+1.4470675$ |
| 5   | $-5.4387574 \times 10^{-04}$ | | $-1.0248451$ | $-0.1870896$ | |

Figure 6. Percentage error of the approximation to $\lambda_1$ given by G86 equation (14) with $\beta(S)$ and $\gamma(S)$ as computed (top), and using the polynomial fits of equations (9) and (10) (bottom), as a function of $n$. In each panel the $S$ values of the shown curves vary from 1 to 5 as annotated. While the approximation is good for low to moderate confidence levels ($S < 3.5$) it fails for $S > 4$ where errors approaching and exceeding 10 per cent are observed for small values of $n$. © 2003 RAS, MNRAS 340, 1269–1278
...by design, vanishes for \( n \approx 0 \) for small values of \( S \). With \( \beta(S) \) unchanged and \( \gamma(S) \) virtually indistinguishable from the data shown in Fig. 5 we can focus on \( \delta(S) \), which shows a complex behaviour (Fig. 7). We do not attempt to model the run of \( \delta(S) \) for small values of \( S \) where the function remains close to zero. Instead, we fit a high-order polynomial to the high-\( S \) end and set \( \delta(S) = 0 \) for \( S < 1.2 \):

\[
\delta(S) = \sum_{i=0}^{8} \delta_i S^i \quad S \geq 1.2
\]

introducing a (totally ad hoc) sinusoidal term that adds only one additional parameter \( \delta \):

\[
\lambda_1 \approx n \left[ 1 - \frac{1}{9 n} - \frac{S}{3 \sqrt{n}} + \beta n^2 + \delta \sin \left( \frac{5}{n + 1/4} \frac{\pi}{2} \right) \right]^3 \quad (11)
\]

Since the sinusoidal term, by design, vanishes for \( n = 1 \), equation (8) still holds, and continues to define \( \beta(S) \), \( \gamma(S) \) (slightly different from that determined from equation 14 of G86) and \( \delta(S) \) are again obtained by iteratively minimizing the absolute error of the approximation for \( 0 \leq n \leq 100 \). With \( \beta(S) \) unchanged and \( \gamma(S) \) virtually indistinguishable from the data shown in Fig. 5 we can focus on \( \delta(S) \), which shows a complex behaviour (Fig. 7). We do not attempt to model the run of \( \delta(S) \) for small values of \( S \) where the function remains close to zero. Instead, we fit a high-order polynomial to the high-\( S \) end and set \( \delta(S) = 0 \) for \( S < 1.2 \):

\[
\delta(S) = \sum_{i=0}^{8} \delta_i S^i \quad S \geq 1.2
\]

with coefficients \( \delta_i \) as listed in Table 3. The results of these fits are shown as the solid line in Fig. 7.

Fig. 8 demonstrates that equation (11) provides an approximation to \( \lambda_1 \) that is accurate to better than 1 per cent when the polynomial fits to \( \beta(S) \), \( \gamma(S) \) and \( \delta(S) \) (equations 9, 10 and 12) are used, except for \( n = 1 \) where an error of just over 1 per cent is observed.

6 SUMMARY

The Gaussian approximation \( \lambda_0 \approx n + S \sqrt{n} \) to the true Poissonian upper confidence limit is acceptable for low confidence levels (\( S < 3 \sigma \)) and \( n > 40 \), but becomes increasingly inaccurate for higher values of \( S \). The situation is worse for the Gaussian approximation \( \lambda_1 \approx n - S \sqrt{n} \) to the true Poissonian lower confidence limit, which is off by more than 10 per cent at \( S = 5 \) even at \( n = 100 \). The approximations proposed by G86 greatly improve upon the Gaussian estimates but are still inaccurate at the 10 per cent level for low values of \( n \) and high confidence levels.

Building on Gehrels’ work we present improved algebraic approximations that reduce the error with respect to the true Poissonian confidence limits to under 1 per cent for \( S \leq 7 \) (Poisson upper...

![Figure 7. Run of \( \delta \) of equation (11) as a function of \( S \), the equivalent number of Gaussian \( \sigma \). The bullets mark the locations at which \( \delta(S) \) was computed, the solid line marks a polynomial fit to the data (see the text for details).](image-url)
Figure 8. Percentage error of the approximations of equation (11) with $\beta(S)$, $\gamma(S)$ and $\delta(S)$ as computed (top), and using the polynomial fits of equations (9), (10) and (12) (bottom) as a function of $n$. In each panel the $S$ values of the shown curves vary from 0.5 to 5 as annotated. For all values of $n$ (except $n = 1$) and $S$ explored here the error remains below the 1 per cent level.

limit) and $S \leq 5$ (Poisson lower limit). Although we have tested these equations only for $n \leq 100$, their analytic behaviour suggests that they hold for all values of $n$ (cf. Figs 4 and 8).

To allow the numerical computation of approximate Poissonian confidence limits for arbitrary combinations of $n$ and $S$ within the quoted ranges, we provide the coefficients of piecewise polynomial fits to all parameters used in the definition of either approximation.

All figures of this paper were produced using the Interactive Data Language (IDL); the IDL source code of the approximations POISSON_UPLIM (equation 4) and POISSON_JOLIM (equation 11) is available from the author.

ACKNOWLEDGMENTS

HE gratefully acknowledges financial support from NASA LTSA grant NAG 5-8253 and NASA ADP grant NAG 5-9238.

REFERENCE

Gehrels N., 1986, ApJ, 303, 336 (G86)

This paper has been typeset from a TeX/LaTeX file prepared by the author.