Basic Principles of X-ray Source Detection

March 4, 2004

1 Optical Astronomers have it easy!

Optical sources typically have high count rates, so gaussian statistics can be used to calculate signal to noise ratios very easily.

$$SNR = \frac{C}{noise} = \frac{C}{\sqrt{\sigma_C^2 + RN^2}} = \frac{T - B}{\sqrt{T + B + RN^2}} \qquad (1)$$

where C is the source counts and T the total counts in the aperture. B is the background counts estimated from the annulus, and RN is the read noise, which is typically ~ 5 counts. At low count rates the read noise dominates, but this requires a count level of ~ 25 which is very unusual, so normally this term can be neglected, at least in imaging observations.

2 X-ray sources - why so few counts?

X-ray astronomers have very few photons to use for their science. This is an intra-universe conspiracy, from particles through material scientists up to AGN and the universe itself.

- X-rays are not produced abundantly in the universe.
- We have to go to space to detect them.
- Focussing X-rays is complex and requires large mirrors.
- Mirrors are increasingly inefficient for higher energy X-rays
- The low energy efficiency is reduced by the optical/UV filters
- Off axis sources have large point-spread-functions, so merge into the background.
- 99.7% of detections are or could be particle background, and so are removed.

2.1 Removing the majority of the detections

We have one advantage over other astronomers in that because we have a low rate and high energy, we can not only determine the position of an event, but also the time of detection, energy and how it affects neighbouring pixels - the grade. This allows us to remove events which are likely to be background. Undoubtably many true X-rays will be removed, but the signal to noise is increased significantly. Events are removed if they fulfil any of the criteria below.

- Bad grade - such as all of the touching pixels being triggered, or a line of 3 pixels.
- Events recieved during times of particularly high background - these are mostly from solar flares.
- Events with energies less than 0.5keV or greater than 8keV.

These constitute over 99% of all recieved events.

2.2 Finding Sources

Sources are normally detected using wavelets. This will give some spurious detections which are just background fluctuations. Once the source is identified, the counts and background are estimated from an aperture and annulus as in the optical case. However, the size of the aperture (and annulus) is variable in the X-ray case.

3 What can we do about it? I - the Theory of Low Count Levels

The optical formulae (1) do not translate directly into the low-count regime. This is especially true of X-ray sources, where there is no read-noise, so getting accurate errors on the count rates is important. The formula assumes that

- The sky aperture is large enough to make errors in the sky count per pixel negligible.
- The 1σ error on the true value is a good estimator of the 68% confidence limit on the estimated value.
- $\sqrt{n}$ is the 1σ error on n counts.

All of these are untrue in the low count regime.

3.1 Background errors

There are two errors on the background counts in the annulus - the estimation of the background rate per pixel, and the application of this to the pixels in the annulus.

The background rate per pixel, B_p, estimated from an annulus of n_s pixels containing B_s counts is

$$B_p = \frac{B_s}{n_s} \pm \frac{\sigma_{B_s}}{n_s}$$

There is also an error on applying this to the aperture of n_a pixels to get the total, B;

$$B = B_p n_a \pm \sigma_{(B_p n_a)}$$

$$\text{therefore} \quad \mathrm{B} = \mathrm{B_s}\frac{\mathrm{n_a}}{\mathrm{n_s}} \pm \sqrt{(\sigma_{\mathrm{B_s}}\frac{\mathrm{n_a}}{\mathrm{n_s}})^2 + (\sigma_{(\mathrm{B_s}\frac{\mathrm{n_a}}{\mathrm{n_s}})})^2} \qquad (2)$$

3.2 Confidence limits

A confidence limit is defined as the region around the *measured* value in which there is an $x\%$ chance that the *true* value lies. It is therefore not always valid to use the distribution of the true value to estimate the errors.

The probability of a measurement, a, drawn from a distribution with given (but here unknown) true mean, $\bar{a}$, is

$$P(a|\bar{a})$$

Suppose that we observe a value n. The chances of observing n or less are

$$P(a < n|\bar{a}) = \int_{-\infty}^{n} P(a|\bar{a})da$$

There is some value of $\bar{a}$, call it a_+, where this is equal to the confidence limit we are seeking, $1 - \beta$,

$$P(a < n|\bar{a} = a_+) = \int_{-\infty}^{n} P(a|a_+)da = \beta$$

Assuming that the shape of $P(a|\bar{a})$ doesn't change significantly between $\bar{a} = n$ and $\bar{a} = a_+$, we can say that

$$\beta = \int_{a_+}^{\infty} P(n|\bar{a})d\bar{a}$$

which is the value we are looking for, the probability that a_+ is greater than the true value of $\bar{a}$.

A similar derivation for the lower confidence limit, α gives

$$\int_{-\infty}^{a_-} P(n|\bar{a})d\bar{a} = \int_{n}^{\infty} P(a|a_-)da = \alpha$$

This is illustrated in figure 1, showing the required quantities on the left and the distributions of a_+ and a_- used to obtain these values. The shaded areas are β and α, the regions in which if $a_\pm$ was the true mean then a value below/above n would be expected less than $\beta\%$ or $\alpha\%$ of the time.

The most commonly used confidence interval is the 68.3% two sided level, corresponding to $\alpha = \beta = 100 - 84.13\%$. This is the 1σ gaussian level, but is used generally as a $''1\sigma''$ error in the general case.

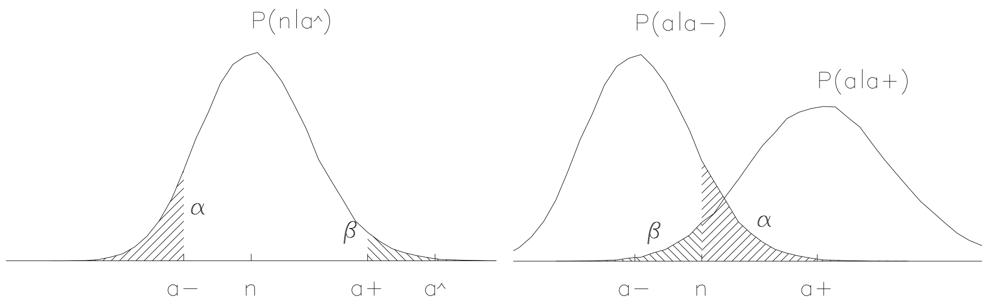

Figure 1: The probability of measuring n, given a true mean a; gives regions α and β on the left. The values of a_- and a_+ are calculated using $P(a|a_\pm)$ as shown on the right.

3.3 Poisson Errors

The Poisson distribution, which describes a small number of random events taken from a large parent distribution, is

$$P(n|\bar{a}) = \frac{\bar{a}^n \; e^{-\bar{a}}}{n!}$$

where n is the integer number of events recorded, and $\bar{a}$ is the true mean of the distribution. From the previous section, the confidence limits on this are

$$\beta = \sum_{x=0}^{n} \frac{{a_+}^x \; e^{-a_+}}{x!}, \qquad 1 - \alpha = \sum_{x=0}^{n-1} \frac{{a_-}^x \; e^{-a_-}}{x!}$$

so that the errors are $n^{+(a_+ - n)}_{-(n-a_-)}$. As this is a discrete distribution, this is the range in which there is *at least* a $1 - \alpha - \beta$ chance that the the true value lies. These equations cannot be solved exactly, but Gehrels (1985) calculates approximations which are valid in most cases.

The Poisson sum is related to the χ^2 distribution by

$$\sum_{x=0}^{n-1} \frac{a^x \; e^{-a}}{x!} = P(\chi^2|\nu)$$

where $\chi^2 = 2a$ and $\nu = 2n$.

Using standard approximate relations for χ^2, we find that for the "1σ" case

$$a_+ = n + 1 + \sqrt{n + \frac{3}{4}}, \qquad a_- = n - \sqrt{n - \frac{1}{4}}$$

The expression for a_+ is commonly refered to as the 'Gehrels error' , although other approximations are derived by Gehrels which are more complex, but more accurate for various ranges of n.

A comparison of the true, gehrels and $\sqrt{n}$ errors is given below

n	+True	-True	Gehrels+	Gehrels-	$\sqrt{n}$
0	+1.84	-0.00	+1.87	-0.00	±0.00
1	+2.30	-0.83	+2.32	-0.67	±1.00
2	+2.63	-1.92	+2.66	-1.33	±1.41
3	+2.92	-1.63	+2.94	-1.66	±1.73
4	+3.16	-1.91	+3.18	-1.94	±2.00
5	+3.38	-2.16	+3.40	-2.18	±2.24
10	+4.27	-3.11	+4.28	-3.12	±3.16
50	+8.12	-7.05	+8.12	-7.05	±7.07
100	+11.00	-9.98	+11.00	-9.99	±10.00
10000	+101.00	-100.00	+101.00	-100.00	±100.00

For most purposes it is sufficient to use the approximation $\sigma = 1 + \sqrt{n + 0.75}$ for the upper limit, which is accurate to $< 1\%$ for $n > 0$.

4 What can we do about it? II - Application to X-ray sources

Replacing all σ_n with the Gehrels error is conservative as it doesn't take into account the smaller value of the lower error. However, it is far more accurate than using $\sqrt{n}$, and is sufficient for most purposes.

There are two statistics of importance when evaluating the significance of a sources.

- The probability that the source is true, i.e. not a background fluctuation. This is the source significance.
- The signal to noise ratio of the source. This is a measure of how accurately the flux of the source can be measured, rather than whether it exists.

We will look at each of these in turn.

4.1 Source Significance

The chance of a source being a background fluctuation is given by the probability that n is drawn from the distribution of background points. This is the significance, sig for a $sig\sigma$ detection.

$$B_a + sig\,\sigma^+_{B_a} = T$$

$$\text{therefore} \quad \mathrm{sig} = \frac{\mathrm{T} - \mathrm{B_s}\frac{\mathrm{n_a}}{\mathrm{n_s}}}{\sigma_\mathrm{B}}$$

$$\text{where} \quad \sigma_\mathrm{B}^2 = \left(\frac{\mathrm{n_a}}{\mathrm{n_s}}\right)^2 \left(1 + \sqrt{\mathrm{B_s} + 0.75}\right)^2 + \left(1 + \sqrt{\mathrm{B_s}\frac{\mathrm{n_a}}{\mathrm{n_s}} + 0.75}\right)^2$$

This uses the background errors for low count regimes (equation 2). As before;
B_s = background counts in annulus
n_s = number of pixels in annulus
T = total counts in aperture
n_a = number of pixels in aperture

4.2 Signal to Noise Ratio

Following the optical example in equation 1, but ignoring read noise and using the gehrels error, we obtain

$$SNR = \frac{T - B_s\frac{n_a}{n_s}}{\sqrt{(\sigma_T^2 + \sigma_B^2)}}$$

$$\text{where} \quad \sigma_\mathrm{T} = 1 + \sqrt{(\mathrm{T} + 0.75)}$$

5 Calculating flux and luminosity

The method for calculating flux requires knowledge of many properties of the telescope. In particular the mirror effective area, detector quantum efficiency and filter transmission all affect the calculation from counts to flux. The various functions are combined into an exposure map which is both energy and position dependent. The method is

- Select a significant source
- Find source counts (and error)
- Assume a source spectrum and convert counts to ergs.
- Create an exposure map at the point of the source
- Convert *ergs* to $erg/cm^2/sec$ by dividing by the exposure map

If required, this can be converted to a luminosity using a distance from measurements in other wavebands.