Statistics, Source Detection, and Noise at High Energies

- How do we know that a source is a real? With what degree of certainty?

- How bright is the source? What is the uncertainty in the brightness?

- Factors include the source intensity, size, shape, separation from other sources; and how the source is observed and measured.

Foundations

- In quantitative scientific research, we make measurements and observations to infer something about the physical world.
- These are always subject to uncertainties originating as
  (a) random measurement errors, and
  (b) systematic errors in measurement and modeling.
- We will mostly stick to discussion of random errors, and focus on source detection and statistics, but the same formalism applies to any sort of signal detection (e.g., a spectral line).

Counting Statistics

- The fundamental measurement in astronomy involves the counting of events (1,2… N) associated with incident photons in a number of bins.
- These bins may be some spatial region in an image, or wavelength region in a spectrum.

Basic Counting Statistics

- N_{source} : number of counts from some source
- N_{back} : number of counts in the background
- N_{total} : total number of counts, $N_{total}=N_{source}+N_{back}$
- The standard deviation for any N, is $\sigma=N^{1/2}$.
- If N, is moderately large N ($N>10$), Gaussian statistics apply, and σ is a measure of the uncertainty in N.
- In this case, the signal-to-noise ratio, $SNR=N/\sigma$

Basic Counting Statistics, contd.

- In the case where there the background is negligible,

$$SNR=N_{source}/\sigma=N_{source}/N_{source}^{1/2}=N_{source}^{1/2}$$

- In the presence of a background with N_{back} counts,

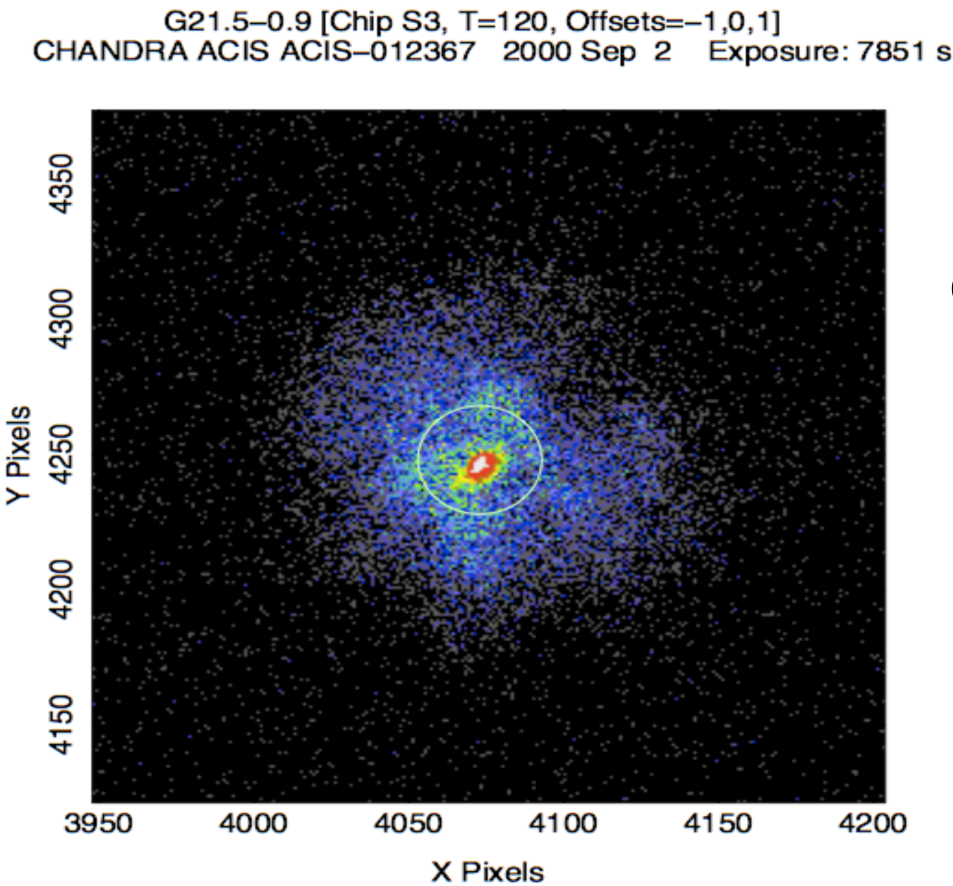
$\sigma=(N_{total}+N_{back})^{1/2}$, and

$$SNR=N_{source}/(N_{total}+N_{back})^{1/2}$$

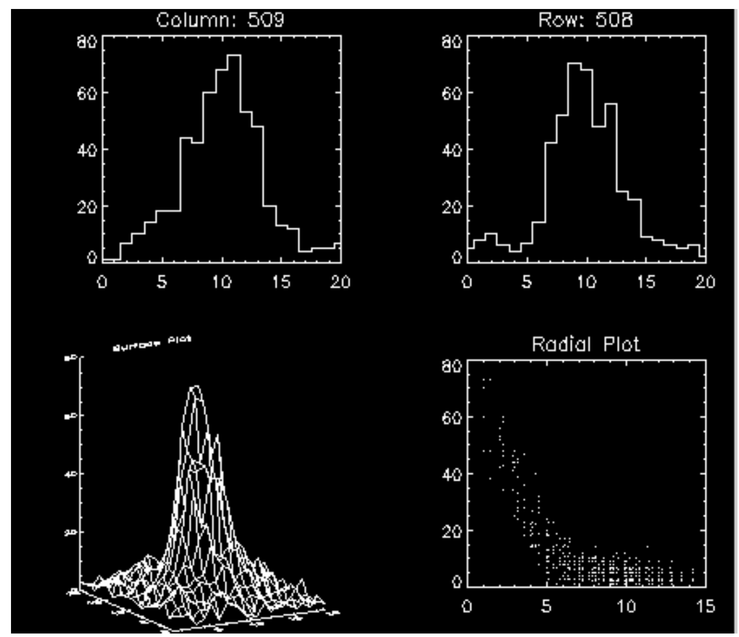
- If there are additional sources of noise (e.g., instrumental), these must be added to the denominator in ``quadrature'':

$$SNR=N_{source}/(N_{total}+N_{back}+\sigma_{other}^{2})^{1/2}$$

An Example: *Chandra* Image of the Pulsar Wind Nebula G21.5-0.9



Extract the counts from a regions chosen based on the PSF



Here, we collect 8000 total counts in a circle with radius of 20 pixels.

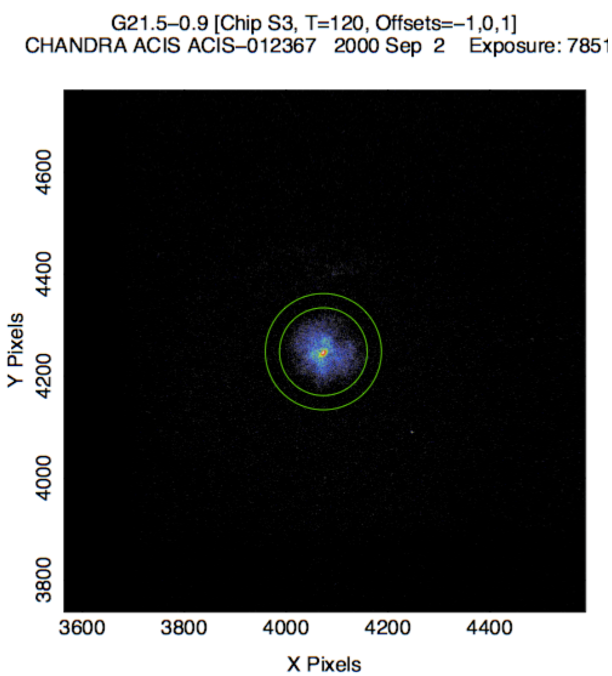
G21.5-0.9 Example contd.

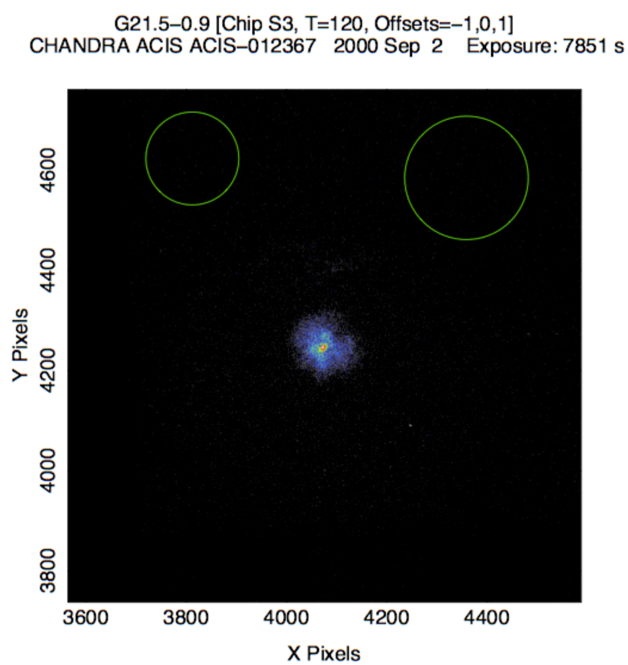
Background (off-source) region options include:

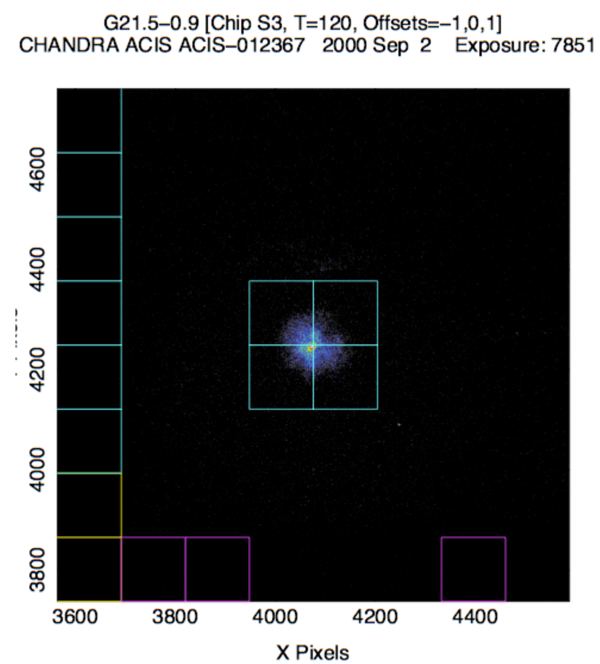
an annular region centered on the source (varying bkgd)

regions of low background (if unresolved sources)

some clever algorithm that cleverly averages







For the annulus, we collect 1000 total counts in a ring with inner and outer radii of 86 and 114 pixels.

G21.5-0.9 Example contd.

Let's crunch the numbers…

$N_{on\text{-}source}$=8000±89 ($8000^{1/2}$=89), $N_{off\text{-}source}$=1000±32

Now, N_{total}=$N_{on\text{-}source}$ -- the total number of counts in the source region.

But, since we want to know the background in the source region, we must correct for the different aperture sizes.

$$N_{back}=(N_{off\text{-}source}/n_{off\text{-}source}) \times n_{on\text{-}source},$$

where $n_{on\text{-}source}$ and $n_{off\text{-}source}$ are the number of on- and off-source pixels, respectively.

G21.5-0.9 Example contd.

$n_{on\text{-}source}$ $=\pi\times 20^2=1256.6$

$n_{off\text{-}source}$ $\pi\times(114^2\text{-}86^2)^2=17593$

Therefore,

$N_{back}= N_{off\text{-}source}\times (n_{on\text{-}source} / n_{off\text{-}source}) = (1000\pm 32) /14$
$=71.4\pm 2.26$

$N_{source}=N_{total}-N_{back}=7929$,

$\sigma_{source}=(8000+1000/14^2)^{1/2}=89.5$, and

$SNR=N_{source}/\sigma_{source}=88.6$

Source Detection

- A workhorse detection algorithm for X-ray images is the "sliding cell" method. Counts are summed in boxes that scan over the image, and compared to those in "background" frames. At each point, SNR is is computed. If this ratio is above some detection threshold, a candidate source is identified.

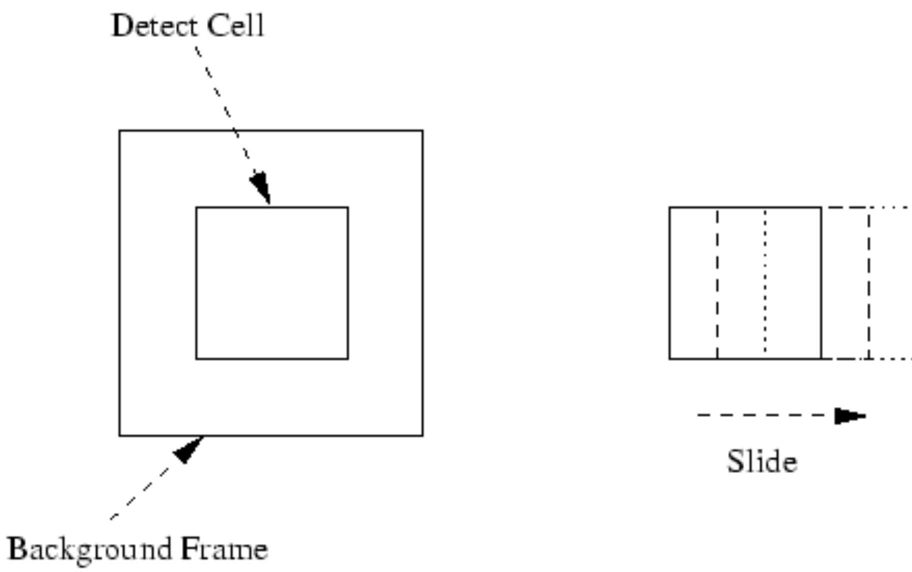


Figure 1: The sliding Detect Cell.

Source Detection, contd.

- The cell size should be comparable to the size of a point source (e.g., the FWHM).
- The background frames may follow the options discussed above.
- The detection threshold SNR is typically set at ~3-5.
- The method breaks down for extended sources, crowded fields.

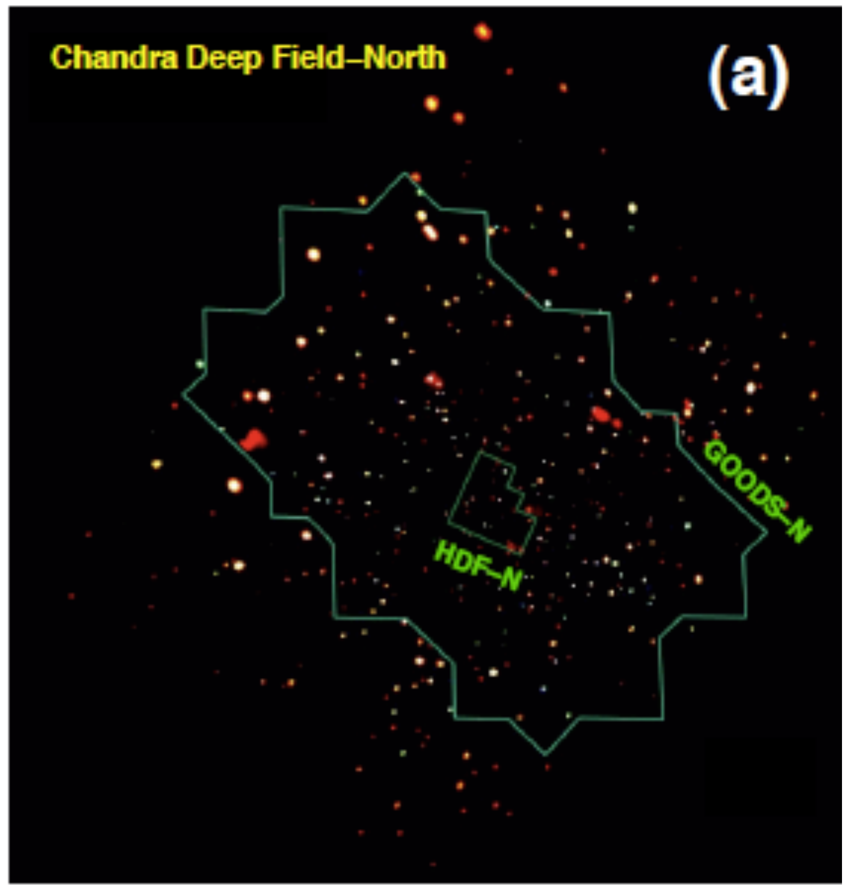
One Astronomer's Background…

- The background can be instrumental, associated with the earth's atmosphere (not a problem for X-ray astronomy), or astrophysical.

- At X-ray wavelengths, the astrophysical background comes from (local and Milky Way) hot interstellar gas and unresolved (mostly extragalactic) sources.

Chandra Deep Surveys

- One of the great early successes of the *Chandra* X-ray Observatory was the resolution of the X-ray background into individual sources – mostly AGN, and mostly at intermediate redshift ($z \approx 1$).



We will use portions of this image in the lab and homework assignment.

W.N. Brandt and G. Hasinger, 2005,
Annual Reviews in Astronomy and Astrophysics, 43, 827

IDL Programming Basics

- IDL programs consists of sequences of IDL commands that are compiled and executed.
- One class of programs are **functions** that receive an argument and return a result via an assignment statement. You've already utilized some examples of built-in functions such as *mean* and *histogram*.
- For now, we will look at **procedures**…

The Most Basic IDL Procedure Format

```
Pro <procedure name>, <in_arg1>, <in_arg2>...
<statement 1>
<statement 2>
...
End
```

input arguments

- Often there are input arguments and output statements.
- If an argument is a *string* it must be delineated by single quotes.
- Procedures are compiled within IDL using **.run**.

IDL Programming Example

```
pro basic, fitsfile
image = readfits(fitsfile,hdr)
image_mean=mean(image)
print, image_mean
end
```

A file procedure called "basic.pro"

```
IDL> .run basic
% Compiled module: BASIC.
IDL> basic, 'xray_image.fits'
% Compiled module: READFITS.
% Compiled module: SXPAR.
% Compiled module: GETTOK.
% Compiled module: VALID_NUM.
% READFITS: Now reading 32 by 32
   array
% Compiled module: MEAN.
% Compiled module: MOMENT.
   6.47363
```

Compile and run the "basic.pro" procedure inside IDL