

# Computing flux upper-limits for non-detections

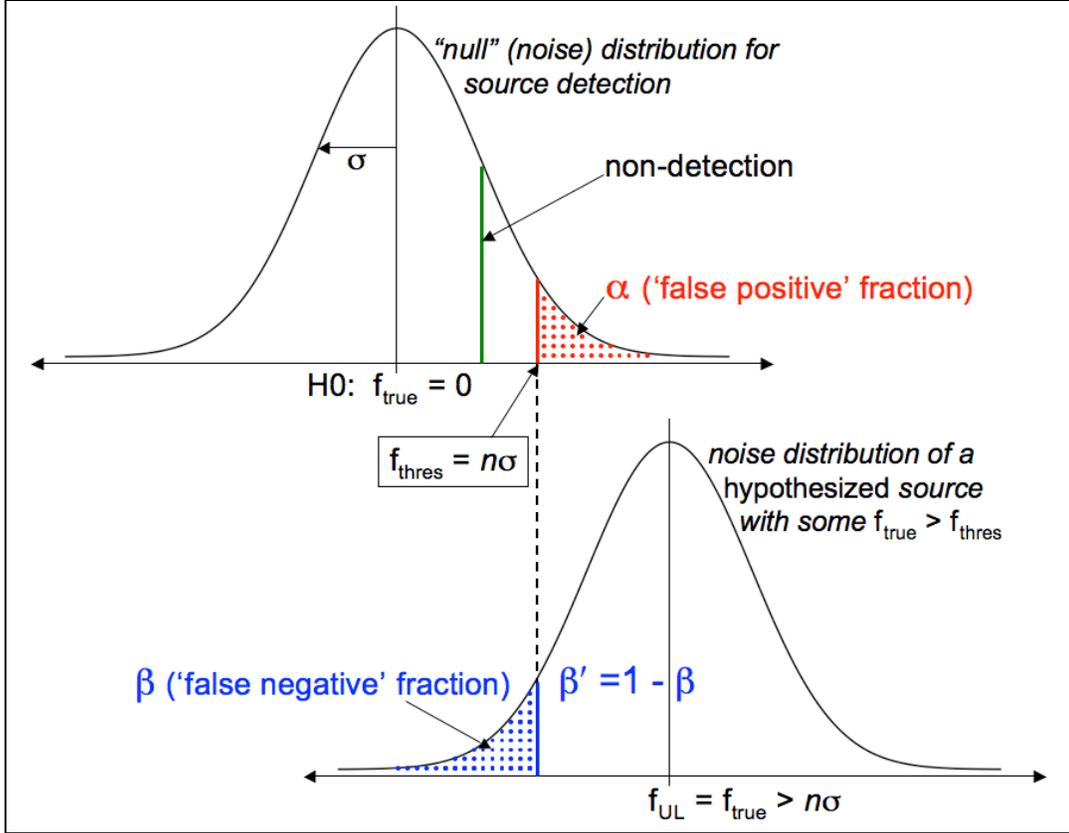
F. Masci, 10/25/2011

If a source is not detected above some flux threshold (when expected to be present at a target position), how can one assign a plausible upper limit to its flux such that (i) it has precise probabilistic meaning; (ii) has a degree of confidence that is not subjective or arbitrarily conservative; and (iii) has no dependence on any prior knowledge about the flux? Unfortunately, the astronomical literature is not consistent in how upper limits are derived, and there are many confusing interpretations. Although “safe” and irrefutable when compared to “real” detections (e.g., in a spectrum), the quoted upper limits often lack probabilistic power.

One of the difficulties stems from the fact that formal confidence intervals can only be assigned if a source is *detected* in the first place, i.e., has flux measurement  $>$  some threshold set according to some maximum tolerable fraction of spurious detections (false positives) in the noise distribution. This noise distribution defines the null hypothesis ( $H_0$ ) of no-detection (source flux consistent with zero). If a source is detected (and declared to have significant non-zero flux), upper and lower flux-limits (or a confidence interval) that *probably* contains the true flux can then be assigned using the uncertainties. If a source is not detected however, we cannot easily proceed because our flux measurement is consistent with zero according to the detection process. This process only decides if a noise excursion is probably not noise but a real source with significance defined *with respect to  $H_0$  only*. It says nothing about possible bounds on its estimated flux relative to the true value.

With no detection (hence knowledge of the probable flux with respect to “truth”), the process of defining an upper limit must be inextricably tied to the detection process. The process of deciding if a flux measurement exceeds some signal-to-noise (S/N) threshold where “N” is the noise-distribution  $\sigma$  is well defined. The noise-distribution for  $H_0$  can be defined either empirically or from a model. Figure 1 shows a schematic of the detection process and how a flux upper-limit can be assigned with a non-subjective probability. The  $\alpha$  value for some threshold  $S/N = n$  specifies the maximum tolerable fraction of “false positives” (or *unreliable* fraction from noise alone) such that a fluctuation with  $S/N > n$  is deemed too rare to be a noise spike and hence declared a source. E.g., if  $n = 3$ , then the (one-tailed) probability that it's a spurious noise spike under a Gaussian  $H_0$  is  $< 0.135\%$ . One can then reject  $H_0$  at the  $\alpha = 0.135\%$  (or  $3\sigma$ ) significance level.

The  $\beta$  value in Figure 1 refers to the fraction of “false negatives”, that is, the fraction of real sources that go undetected. This value can be used to define a statistically meaningful measure for the flux upper-limit  $f_{UL}$  on a non-detection since it indicates the maximum tolerable *incompleteness* of a population of sources with true flux  $f_{true} = f_{UL}$  subject to noise alone. Therefore, we define  $f_{UL}$  as the *minimum* expected flux ( $f_{true}$ ) such that  $\beta$  achieves some  $\beta_{max}$ , where  $\beta_{max}$  is the maximum fraction of sources with  $f_{true} = f_{UL}$  we can tolerate as going undetected. In other words,  $100(1 - \beta_{max})\%$  of the sources with flux  $f_{UL}$  will have flux measurements  $> n\sigma$ . This also represents the minimum desired *completeness* level inflicted by noise alone. Therefore, given that we can predict in advance what *is possible* with our observational design and detection process (according to  $n\sigma$  and  $\beta_{max}$ ),  $f_{UL}$  as defined here provides a plausible upper limit for a non-detection.



**Figure 1:** schematic illustrating the definition of an upper limit for a non-detection, i.e., with flux  $\leq n\sigma$  and hence consistent with the null hypothesis (H0) of zero flux. A flux upper-limit ( $f_{UL}$ ) is assigned according to one's choice of the probability or confidence ( $\beta'$ ) that an actual source with flux =  $f_{UL}$  will be detected. Equivalently,  $f_{UL}$  is the *minimum* expected flux of a source such that the probability of *not* detecting it does not exceed some  $\beta$ . See text for details.

For the common case of Gaussian-distributed noise (which is also a good approximation to Poisson-distributed noise when the detector counts exceed  $\sim 30$ ), we can provide some estimates for  $f_{UL}$  in terms of the detection threshold  $n\sigma$  for different choices of the detection probability  $\beta' = \Pr(f > n\sigma)$ . This probability can be written:

$$\begin{aligned} \beta' &= \frac{1}{\sqrt{2\pi}} \int_z^{\infty} e^{-t^2/2} dt \\ &= \frac{1}{2} \left[ 1 - \operatorname{erf}\left(\frac{z}{\sqrt{2}}\right) \right], \end{aligned} \quad (1)$$

where

$$z = \frac{n\sigma - f_{UL}}{\sigma} \quad (2)$$

and  $\operatorname{erf}(x)$  is the error function.

We have defined the population mean flux  $\mu$  (or true flux) of a fiducial source with  $f_{\text{true}} > n\sigma$  to be the flux upper limit sought for,  $f_{\text{UL}}$ . Note: we have assumed that the  $\sigma$  of this fiducial source is the same as that used to define the initial detection threshold ( $n\sigma$ ) under a  $H_0$  of  $f_{\text{true}} = 0$ . I.e., we assume this  $\sigma$  does not depend on the source flux, as is the case for Poisson noise in the source-photon dominated (high flux) regime. However, given detection thresholds generally correspond to low fluxes (where background-noise and/or detector read-noise likely dominates), and since our fiducial source flux used to define  $f_{\text{UL}}$  is not expected to be too far above  $n\sigma$ , the same  $\sigma$  should suffice throughout.

Combining Equations (1) and (2), we have:

$$\beta' = \frac{1}{2} \left[ 1 - \operatorname{erf} \left( \frac{n\sigma - f_{\text{UL}}}{\sigma\sqrt{2}} \right) \right]. \quad (3)$$

We can invert (3) and express  $f_{\text{UL}}$  in terms of all other parameters and the inverse error function,  $\operatorname{erf}^{-1}(x)$ :

$$f_{\text{UL}} = \left[ n + \sqrt{2} \operatorname{erf}^{-1}(2\beta' - 1) \right] \sigma. \quad (4)$$

The flux upper limit for a non-detection therefore depends only on the detection parameters ( $n$ ,  $\sigma$ ) and our desired value for the probability  $\beta'$  of detecting a fiducial source with flux =  $f_{\text{UL}}$ . Table 1 presents some possible  $\beta'$  values of interest and the corresponding flux upper limits. These are expressed in terms of the detection S/N threshold  $n$  and noise  $\sigma$ . As a reminder, these values only apply to a detection process where measurements are affected by Gaussian noise (the basis of Equation 4).

$\beta'$	$1 - \beta'$	$f_{\text{UL}}$
0.50	0.50	$n\sigma$
0.75	0.25	$[n + 0.674]\sigma$
0.80	0.20	$[n + 0.842]\sigma$
0.85	0.15	$[n + 1.036]\sigma$
0.90	0.10	$[n + 1.282]\sigma$
0.95	0.05	$[n + 1.645]\sigma$
0.98	0.03	$[n + 2.054]\sigma$
0.99	0.01	$[n + 2.326]\sigma$
0.999	0.001	$[n + 3.090]\sigma$

**Table 1: suggested detection probability values ( $\beta'$ ) and corresponding flux upper limits ( $f_{\text{UL}}$ ) to assume for a non-detection with flux  $\leq n\sigma$  under a simple Gaussian-noise model.**

The first row in Table 1 shows that assuming the actual detection threshold  $n\sigma$  as the flux upper limit (a choice often encountered in the literature) implies a detection probability of  $\beta' = 50\%$ . This limit for a non-detection is speculative. We recommend a more conservative value for  $f_{\text{UL}}$ , e.g., that corresponding to  $\beta' = 98\%$  is a good choice. A S/N threshold of  $n = 3$  then implies  $f_{\text{UL}} \approx 5\sigma$  from Table 1 (highlighted red). With this assumption, a survey designed to select sources with  $S/N > 3$  will detect 98% of sources with true flux  $\approx 5\sigma$ . This flux therefore represents a plausible upper limit for a non-detection with  $S/N \leq 3$ .

If we were to go deeper and attempt to detect this  $\leq 3\sigma$  source above some new threshold  $3\sigma_{\text{new}}$  (e.g., by beating down the noise with multiple independent exposures to give some  $\sigma_{\text{new}} < \sigma$ ), our initially assigned  $f_{\text{UL}} = 5\sigma$  will still apply. Once detected (with flux presumably satisfying  $3\sigma_{\text{new}} < f \leq 3\sigma$ ), our  $5\sigma$  upper limit will be more conservative, i.e., larger relative to the measured flux since  $\sigma_{\text{new}} < \sigma$ . For example, if we need  $\sigma_{\text{new}} = 0.1\sigma$  to detect the source, our initial  $f_{\text{UL}}$  will imply an upper limit of  $(5/0.1)\sigma_{\text{new}} = 50\sigma_{\text{new}}$ ! The least-conservative (worst case) scenario is if the source resides just below the initial  $3\sigma$  threshold, e.g., at  $2.999999\sigma$ . We can then recover it with  $\sigma_{\text{new}} = 0.9999995\sigma$ , giving a detection of  $3.000001\sigma_{\text{new}}$ . The upper limit now becomes  $(5/0.9999995)\sigma_{\text{new}} = 5.000003\sigma_{\text{new}}$ . This limit is  $[5.000003 - 3.000001]\sigma_{\text{new}} \approx 2\sigma_{\text{new}}$  above the measured flux and corresponds to a confidence of  $\approx 97.7\%$  for the *true* flux for Gaussian noise. Therefore, the choice  $f_{\text{UL}} \approx [n + 2.054]\sigma$  will be reasonably conservative for all possible fluxes our non-detection may have. I doubt anyone will argue with this given it has a clear probabilistic interpretation when the noise-distribution is known.

The simple construction above ignores the fact that true source populations follow a luminosity function (or observed number-count law) whereby there are relatively more faint sources than bright ones. Noise fluctuations will systematically cause more sources just below a flux threshold to be bumped above it and fewer to be thrown below it (and hence become non-detected). In qualitative terms, this will bias the counts upwards just above the flux threshold with the consequence of increasing the detection probability ( $\beta'$ ). This is the well known ‘Eddington Bias’. This bias works in the preferred direction when it comes to assigning flux upper limits to non-detections. Its ignorance (i.e., assuming source numbers symmetrically fluctuate about a threshold) will make  $f_{\text{UL}}$  slightly larger (more conservative) than it needs to be for a given  $\beta'$ . An exercise for the reader is to derive the necessary correction to  $f_{\text{UL}}$  assuming a power-law source-flux distribution that is hit by Gaussian-noise. You might as well also derive it for the case of Poisson-noise in the low photon-count regime.