Bayesian Evaluation of (Astrophysical) Source Strength

Nick van Eijndhoven

nick@icecube.wisc.edu   http://www.iihe.ac.be

Vrije Universiteit Brussel - IIHE(ULB-VUB)
Pleinlaan 2, B-1050 Brussel, Belgium

Contents

Source Strength and Fluxes

Some definitions

Source Luminosity (L) : Total emitted amount of radiation per unit of time

Flux (F) : Amount of radiation passing per unit area per unit of time

* Note : Radiation can represent energy as well as number of particles
* Redshift effects are neglected and will be discussed later in this lecture series

Flux measurements

- Consider a spherical source with radius R_s and luminosity L that radiates photons isotropically from its surface (e.g. a star)
- Consider a sphere with radius $r > R_s$ with the source at its center

  $\rightarrow$ Photon flux at the surface of this sphere: $F = \dfrac{L}{4\pi r^2}$
- Imagine a detector at some distance $d > R_s$ from the center of the source

  $\rightarrow$ Observed flux F will decrease as d^{-2}

If distance unknown $\rightarrow$ L can not be determined from the observed flux

Source Strength and Fluxes

What can we infer about the source strength if the distance is not known ?

- Surface area on a sphere with radius r : $\mathrm{d}A = r^2 \sin(\theta)\mathrm{d}\theta\mathrm{d}\phi$

  Solid angle : $\mathrm{d}\Omega \equiv \sin(\theta)\mathrm{d}\theta\mathrm{d}\phi$ which yields $\mathrm{d}A = r^2\mathrm{d}\Omega$

- At a distance $d \gg R_s$ the source covers a solid angle $\Omega_s = \int \mathrm{d}\Omega \approx \dfrac{\pi R_S^2}{d^2}$ sr

  $\rightarrow$ Flux received from a unit of solid angle : $\Phi = \dfrac{F}{\Omega_s} = \dfrac{L}{4\pi(\pi R_S^2)}$

  This does not depend on the distance from the source !

  $\rightarrow$ Good observable to characterize the source strength

- No distance dependence $\rightarrow$ Same value at the radiating source

  As seen from the source : $\Phi =$ Flux emitted within a unit of solid angle

  Source Intensity (I) : Flux received c.q. emitted per unit of solid angle

∗ For an extended source that can be resolved, intensity measurements allow to characterize the brightness of different source areas

Source Strength and Fluxes

How to measure the source intensity ?

- Consider a detector with area A cm^2 aimed at a source

  In a time interval ΔT seconds the detector records N photons
  and a deposited energy of E GeV

  Seen from the detector the source spans a circle with radius $\alpha \ll 1$ radians

  $$\rightarrow \Omega_s = \int_0^{\alpha} \sin(\theta)\, \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \approx 2\pi \int_0^{\alpha} \theta\, \mathrm{d}\theta = \pi\alpha^2$$

  Photons : $F = \dfrac{N}{A\,\Delta T}$ ph cm^{-2} s^{-1} $\qquad I = \dfrac{N}{A\,\Delta T\,\pi\alpha^2}$ ph cm^{-2} s^{-1} sr^{-1}

  Energy : $F = \dfrac{E}{A\,\Delta T}$ GeV cm^{-2} s^{-1} $\qquad I = \dfrac{E}{A\,\Delta T\,\pi\alpha^2}$ GeV cm^{-2} s^{-1} sr^{-1}

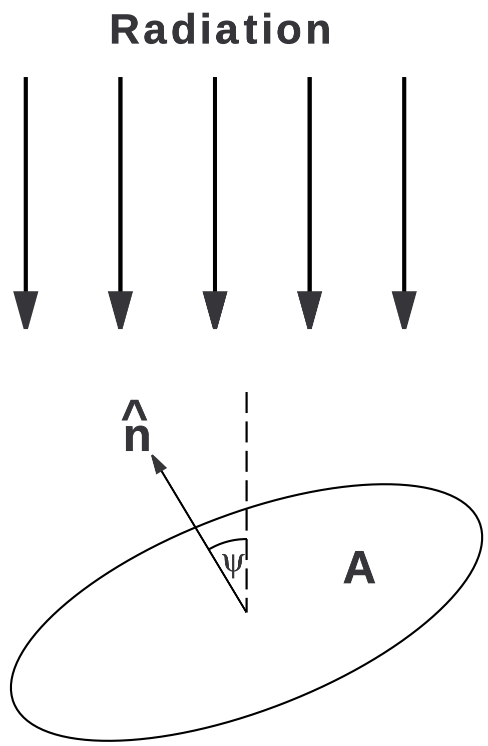
$*$ For (short) transient events one often denotes the **Fluence**

Fluence (S) : Time integrated flux

$$\rightarrow \qquad S = \frac{N}{A} \text{ ph cm}^{-2} \qquad\qquad S = \frac{E}{A} \text{ GeV cm}^{-2}$$

Source Strength and Fluxes

- In case the detector area A is mis-aligned w.r.t. the incident radiation
  $\rightarrow$ A **has to be replaced by** $A\cos(\psi)$



- The same holds for the emission when the source area is mis-aligned

Source Strength and Fluxes

- Very often not every incident particle will be recorded by the detector

  This is accounted for by replacing A with the so called Effective Area (A_{eff})

$$\text{Particles : } A_{eff} = \frac{\text{Observed count rate}}{\text{Incoming flux}} = \frac{\text{\# Observed counts}}{\text{Incoming fluence}}$$

$$\text{Energy : } A_{eff} = \frac{\text{Recorded power}}{\text{Incoming flux}} = \frac{\text{Recorded energy}}{\text{Incoming fluence}}$$

$*$ A_{eff} can only be determined by simulations or a calibration source

Flux densities $F(\nu)$ and $F(E)$

$F(\nu)\,\mathrm{d}\nu \equiv$ Flux in the frequency interval $[\nu, \nu + \mathrm{d}\nu]$

$F(E)\,\mathrm{d}E \equiv$ Flux in the energy interval $[E, E + \mathrm{d}E]$

$\rightarrow$ This leads also to $I(\nu)$, $S(\nu)$, $I(E), S(E)$ etc.

$*$ One often uses the notation F_ν, I_ν, S_ν etc.

Special flux density unit : Jansky (Jy)

$$1\ \text{Jy} = 10^{-26}\ \text{J s}^{-1}\ \text{m}^{-2}\ \text{Hz}^{-1} = 10^{-23}\ \text{erg s}^{-1}\ \text{cm}^{-2}\ \text{Hz}^{-1}$$

$$= 6.25 \cdot 10^{-21}\ \text{GeV s}^{-1}\ \text{cm}^{-2}\ \text{Hz}^{-1}$$

Formalism of a Bayesian rate analysis

The problem to be addressed

- Cosmic rays impinge on the atmosphere of the Earth, producing a constant rate of atmospheric high-energy neutrinos, homogeneously distributed over the celestial sphere. We call these neutrinos (atmospheric) background.
- Astrophysical sources may yield an additional, constant high-energy neutrino rate at specific locations. We call these neutrinos (cosmic) signal.

∗ The essential questions :

  - Can we identify a possible signal by measurements ?
  - Can we determine (a limit on) the source strength ?
  - What is our degree of belief in the presence of a source ?

- To investigate we study a certain patch on the sky over a time interval Δt.

  This will result in observing n neutrinos.

∗ All of the above is called our prior information I.

- Note : The following reasoning works also for stacked observations.

Formalism of a Bayesian rate analysis

- Based on our prior information we know that the pdf for the number of observed neutrinos n is the Poisson distribution with constant rate r

$$p(n|rI) = \frac{(r\Delta t)^n e^{-r\Delta t}}{n!}$$

- But : We actually want to determine the rate $r \rightarrow p(r|nI)$

Bayesian Logical Inference

- Consider two **propositions** H **and** D and some **prior information** I

  Product rule : $p(HD|I) = p(H|I)p(D|HI) = p(D|I)p(H|DI)$

- From the product rule we have : $p(H|I)p(D|HI) = p(D|I)p(H|DI)$

  which can be written as :

$$\boxed{p(H|DI) = p(H|I)\,\frac{p(D|HI)}{p(D|I)}}$$

(Theorem of Bayes)

Formalism of a Bayesian rate analysis

- The Bayes theorem directly yields $p(r|nI) = p(r|I)\,\dfrac{p(n|rI)}{p(n|I)}$

  $p(r|I)$ is some prior pdf for the rate

  $p(n|rI)$ is the Poisson pdf from before

  $p(n|I)$ is some normalisation which can be determined as follows :

  $\int p(r|nI)\,\mathrm{d}r = 1 \rightarrow p(n|I) = \int p(r|I)\,p(n|rI)\,\mathrm{d}r$

* But : The rate r consists of independent signal and background $\rightarrow r = r_s + r_b$

  So we get : $p(r_s r_b|nI) = p(r_s r_b|I)\,\dfrac{p(n|r_s r_b\,I)}{p(n|I)}$

  where : $p(n|r_s r_b\,I) = \dfrac{([r_s + r_b]\Delta t)^n e^{-[r_s + r_b]\Delta t}}{n!}$ (Poisson for rate $r = [r_s + r_b]$)

  Product rule : $p(r_s r_b|I) = p(r_b|I)p(r_s|r_b\,I) = p(r_b|I)p(r_s|I)$

  $p(r_b|I)$ is some prior pdf for the bkg rate (e.g. $p(r_b|n_b\,I)$ from off-source)
  $p(r_s|I)$ is some prior pdf for the signal rate (e.g. based on previous limits)

- Similar as above : $p(n|I) = \int\int p(r_b|I)\,p(r_s|I)\,p(n|r_s r_b\,I)\,\mathrm{d}r_b\,\mathrm{d}r_s$

Formalism of a Bayesian rate analysis

- Conclusion :

  Given some priors $p(r_b|I)$ and $p(r_s|I)$ we can determine $p(r_s r_b|nI)$

  But : We want to determine the pure signal rate $p(r_s|nI)$

* Thanks to the Bayesian logic : Marginalisation

  Without loss of statistical information : $p(r_s|nI) = \int p(r_s r_b|nI)\,\mathrm{d}r_b$

  We can determine the full posterior signal rate pdf $p(r_s|nI)$ from data alone

- Going to credible regions, upper limits and all that

* x% credibility region $[r_{min}, r_{max}]$ for r_s : $\displaystyle\int_{r_{min}}^{r_{max}} p(r_s|nI)\,\mathrm{d}r_s =$ x%

  with $r_{min} < \hat{r}_s < r_{max}$ and $p(r_{min}|nI) = p(r_{max}|nI)$

  $\rightarrow$ r_{min} and r_{max} form the x% credibility region of the signal rate r_s

* x% upper limit r_{max} for r_s : $\displaystyle\int_{0}^{r_{max}} p(r_s|nI)\,\mathrm{d}r_s =$ x%

  $\rightarrow$ r_{max} is the x% credible upper limit for the signal rate r_s

Formalism of a Bayesian rate analysis

- **Decision between credibility region or upper limit**

  Degree of belief in signal presence via hypothesis testing c.q. ψ value

  (NvE, Astroparticle Physics 28 (2008) 540, arXiv:astro-ph/0702029)

  See also my lectures on "Logical Data Analaysis"

- **Going from an observed signal rate r_s to an incoming flux value F_s**

  Just plug in the Effective Area (A_{eff})

  $A_{eff} \equiv$ observed event rate / incoming flux (from simulations)

  which yields : $F_s = \dfrac{r_s}{A_{eff}} \rightarrow p(F_s|nI) = \dfrac{p(r_s|nI)}{A_{eff}}$

  We get the full posterior source flux pdf $p(F_s|nI)$ directly from $p(r_s|nI)$

- **Derivation of the (distance independent) source intensity I_s**

  In case the solid angle Ω_s of the source can be determined : $I_s = \dfrac{F_s}{\Omega_s}$

The total signal and background rate

- To summarise the previous we have

$$p(r_s r_b|nI) = p(r_b|I)\, p(r_s|I)\, \frac{p(n|r_s r_b\, I)}{p(n|I)}$$

consisting of the following contributions :

$p(r_b|I)$ is some (prior) pdf for the background rate

$p(r_s|I)$ is some prior pdf for the signal rate

$p(n|r_s r_b\, I) = \dfrac{([r_s + r_b]\Delta t)^n e^{-[r_s+r_b]\Delta t}}{n!}$ (Poisson for rate $r = [r_s + r_b]$)

$p(n|I) = \int \int p(r_b|I)\, p(r_s|I)\, p(n|r_s r_b\, I)\, \mathrm{d}r_b\, \mathrm{d}r_s$

$*$ And the pure (i.e. background independent) signal rate is given by

$$p(r_s|nI) = \int p(r_s r_b|nI)\, \mathrm{d}r_b$$

Working out the mathematics

Background rate determination

- The background rate pdf can be obtained via an "off-source" measurement observing n_o (background) events over a certain time period t_o
- Note : This measurement is different from the actual "on-source" one

  $\rightarrow$ Different prior information I_o

- From the previous we have : $p(r_b|n_o\,I_o) = p(r_b|I_o)\,\dfrac{p(n_o|r_b\,I_o)}{p(n_o|I_o)}$

$$p(n_o|r_b\,I_o) = \frac{(r_b\,t_o)^{n_o}e^{-r_b\,t_o}}{n_o!} \qquad \text{(Off-source bkg Poisson pdf)}$$

$p(r_b|I_o)$ is some prior pdf for the background rate

$p(n_o|I_o) = \int p(r_b|I_o)\,p(n_o|r_b\,I_o)\,\mathrm{d}r_b$

$*$ So we have $p(n_o|I_o) = \displaystyle\int p(r_b|I_o)\,\frac{(r_b\,t_o)^{n_o}e^{-r_b\,t_o}}{n_o!}\,\mathrm{d}r_b$

$\rightarrow$ Specification of the prior $p(r_b|I_o)$ yields the bkg rate pdf $p(r_b|n_o\,I_o)$

Working out the mathematics

The background rate prior $p(r_b|I_o)$

Since $r_b \geq 0$ it qualifies as a scale parameter $\rightarrow$ Jeffrey's prior

$$p(r_b|I_o) = \frac{1}{r_b \ln\left[r_{bmax}/r_{bmin}\right]}$$

But ... r_b may vanish $\rightarrow r_{bmin} = 0$

$*$ Use modified Jeffrey's prior : $p(r_b|I_o) = \dfrac{1}{(\beta + r_b)\ \ln\left[(\beta + r_{bmax})/\beta\right]}$

where β is a constant

This behaves as a uniform prior for $r_b < \beta$ and a Jeffrey's prior for $r_b \geq \beta$

- So we get :

$$p(n_o|I_o) = \int_0^{r_{bmax}} \frac{1}{(\beta + r_b)\ \ln\left[(\beta + r_{bmax})/\beta\right]} \frac{(r_b\, t_o)^{n_o} e^{-r_b\, t_o}}{n_o!}\, \mathrm{d}r_b$$

Exact but complicated integral : Analytical expression ?

Working out the mathematics

- **Try with the simpler uniform prior :** $p(r_b|I_o) = \dfrac{1}{r_{bmax} - r_{bmin}}$

As mentioned before $r_{bmin} = 0 \to \boldsymbol{p(r_b|I_o) = \dfrac{1}{r_{bmax}}}$

Which yields : $\boldsymbol{p(n_o|I_o) = \int_0^{r_{bmax}} \dfrac{1}{r_{bmax}} \dfrac{(r_b\,t_o)^{n_o} e^{-r_b\,t_o}}{n_o!}\,\mathrm{d}r_b}$

$*$ **Note :** $\int_0^x y^n e^{-y}\,\mathrm{d}y = \gamma(n+1,x)$ **(incomplete gamma function)**

So we get : $\boldsymbol{p(n_o|I_o) = \dfrac{\gamma(n_o+1, r_{bmax}\,t_o)}{r_{bmax}\,t_o\,n_o!}}$

which finally yields for the posterior background rate pdf

$$\boxed{p(r_b|n_o\,I_o) = \frac{t_o\,(r_b\,t_o)^{n_o} e^{-r_b\,t_o}}{\gamma(n_o+1, r_{bmax}\,t_o)}}$$

$*$ **Note : If** $r_{bmax}\,t_o \gg n_o$ **then** $\gamma(n_o+1, r_{bmax}\,t_o) \approx \Gamma(n_o+1) = n_o!$

Working out the mathematics

The signal and background rate of the actual "on-source" observation

- In the actual "on-source" measurement we observe a total of n events over a certain time period t

Knowledge of the "off-source" measurement is part of the prior info I !

Recall : $p(r_s r_b|nI) = p(r_b|I)\, p(r_s|I)\, \dfrac{p(n|r_s r_b\, I)}{p(n|I)}$

$p(r_b|I)$ is some (prior) pdf for the background rate

* Use the "off-source" posterior pdf : $p(r_b|I) = \dfrac{t_o\, (r_b\, t_o)^{n_o} e^{-r_b\, t_o}}{\gamma(n_o+1, r_{bmax}\, t_o)}$

$p(r_s|I)$ is some prior pdf for the signal rate

$p(n|r_s r_b\, I) = \dfrac{([r_s + r_b]t)^n e^{-[r_s+r_b]t}}{n!}$ (Poisson for rate $r = [r_s + r_b]$)

$\rightarrow p(n|I) = \displaystyle\int\int p(r_s|I)\, \frac{t_o\, (r_b\, t_o)^{n_o} e^{-r_b\, t_o}}{\gamma(n_o+1, r_{bmax}\, t_o)}\, \frac{([r_s + r_b]t)^n e^{-[r_s+r_b]t}}{n!}\, \mathrm{d}r_b\, \mathrm{d}r_s$

Working out the mathematics

The signal rate prior $p(r_s|I)$

Similar approach as for the background "off-source" case

- Modified Jeffrey's prior : $p(r_s|I) = \dfrac{1}{(\sigma + r_s)\ \ln\left[(\sigma + r_{smax})/\sigma\right]}$

where σ is a constant

Again a complicated integral for $p(n|I)$: Analytical expression ?

- Uniform prior : $p(r_s|I) = \dfrac{1}{r_{smax}}$ which yields

$$p(n|I) = \int_0^{r_{smax}} \int_0^{r_{bmax}} \frac{1}{r_{smax}} \frac{t_o\,(r_b\,t_o)^{n_o} e^{-r_b\,t_o}}{\gamma(n_o+1, r_{bmax}\,t_o)} \frac{([r_s+r_b]t)^n e^{-[r_s+r_b]t}}{n!}\,\mathrm{d}r_b\,\mathrm{d}r_s$$

Evaluation of this integral allows to determine $p(r_s r_b|nI)$

$*$ BUT ... our interest is the pure (bkg independent) signal rate

$$p(r_s|nI) = \int_0^{r_{bmax}} p(r_s r_b|nI)\,\mathrm{d}b = \frac{\int_0^{r_{bmax}} p(r_b|I)\,p(n|r_s r_b\,I)\,\mathrm{d}r_b}{r_{smax}\,p(n|I)}$$

Ratio of similar integrals for $p(r_s|nI)$ can be solved

Working out the mathematics

- Writing $p(r_s|nI) \equiv A/B$ we have as expressions :

$$A = \int_0^{r_{bmax}} (r_b)^{n_o}\, e^{-r_b t_o}\, (r_s + r_b)^n\, e^{-(r_s+r_b)\,t}\, \mathrm{d}r_b$$

$$B = \int_0^{r_{smax}} \int_0^{r_{bmax}} (r_b)^{n_o}\, e^{-r_b t_o}\, (r_s + r_b)^n\, e^{-(r_s+r_b)\,t}\, \mathrm{d}r_b\, \mathrm{d}r_s = \int_0^{r_{smax}} A\, \mathrm{d}r_s$$

- Binomial expansion : $(r_s + r_b)^n = \sum_{i=0}^{i=n} \frac{n!}{i!\,(n-i)!} (r_s)^i\, (r_b)^{(n-i)}$

which yields : $A = \sum_{i=0}^{i=n} \frac{n!}{i!\,(n-i)!} (r_s)^i\, e^{-r_s t} \int_0^{r_{bmax}} (r_b)^{n+n_o-i}\, e^{-r_b(t+t_o)}\, \mathrm{d}r_b$

$$\rightarrow A = \sum_{i=0}^{i=n} \frac{n!}{i!\,(n-i)!} (r_s)^i\, e^{-r_s t}\, \frac{\gamma([n+n_o-i+1], r_{bmax}[t+t_o])}{(t+t_o)^{(n+n_o-i+1)}}$$

$$* \; A = \frac{n!\, e^{-r_s t}}{(t+t_o)^{(n+n_o+1)}} \sum_{i=0}^{i=n} \frac{(r_s)^i\, (t+t_o)^i\, \gamma([n+n_o-i+1], r_{bmax}[t+t_o])}{i!\,(n-i)!}$$

Working out the mathematics

- The expression for $B = \int_0^{r_{smax}} A \, \mathrm{d}r_s$ will then become

$$B = \frac{n!}{(t+t_o)^{(n+n_o+1)}}$$

$$\times \sum_{j=0}^{j=n} \frac{(t+t_o)^j \, \gamma([n+n_o-j+1], r_{bmax}[t+t_o])}{j! \, (n-j)!} \int_0^{r_{smax}} (r_s)^j \, e^{-r_s \, t} \, \mathrm{d}r_s$$

$$\rightarrow B = \frac{n!}{(t+t_o)^{(n+n_o+1)}}$$

$$\times \sum_{j=0}^{j=n} \frac{(t+t_o)^j \, \gamma([n+n_o-j+1], r_{bmax}[t+t_o])}{j! \, (n-j)! \, t^{j+1}} \gamma(j+1, r_{smax} \, t)$$

Working out the mathematics

- So this finally yields for the **posterior signal rate pdf**

$$p(r_s|nI) = \frac{t\,e^{-r_s t}\,\sum_{i=0}^{i=n} \frac{(r_s)^i\,(t+t_o)^i\,\gamma([n+n_o-i+1],r_{bmax}[t+t_o])}{i!\,(n-i)!}}{\sum_{j=0}^{j=n} \frac{(1+\frac{t_o}{t})^j\,\gamma([n+n_o-j+1],r_{bmax}[t+t_o])\,\gamma(j+1,r_{smax}\,t)}{j!\,(n-j)!}}$$

Note that :

∗ To obtain good prior coverage r_{bmax} and r_{smax} should be chosen large enough

∗ r_{smax} may be obtained from an existing upper limit on the signal rate (or flux) determined in a previous measurement

∗ In the case that $r_{bmax}[t+t_o] \gg n+n_o$
we have $\gamma([n+n_o-i+1, r_{bmax}[t+t_o]) \approx (n+n_o-i)!$
and $\gamma([n+n_o-j+1, r_{bmax}[t+t_o]) \approx (n+n_o-j)!$

Summary

Background rate determination

- The background rate r_b can be obtained via an "off-source" measurement observing n_o (background) events over a certain time period t_o

Posterior background rate pdf

$$\boxed{p(r_b|n_o,t_o) = \frac{t_o\,(r_b\,t_o)^{n_o} e^{-r_b\,t_o}}{\gamma(n_o+1, r_{bmax}\,t_o)}}$$

* For good prior coverage r_{bmax} should be chosen large enough
* If $r_{bmax}\,t_o \gg n_o$ then $\gamma(n_o+1, r_{bmax}\,t_o) \approx \Gamma(n_o+1) = n_o!$

- See function GetBackgroundRatePDF() of the NcAstrolab facility in the NCFS-Pack analysis framework (https://nick-ve.github.io/ncfs/docs)

Summary

Signal rate determination

- In the actual "on-source" measurement we observe a total of n events over a certain time period t

  Together with the "off-source" measurement (n_o, t_0) this yields

Posterior signal rate pdf

$$p(r_s|n,t,n_o,t_o) = \frac{t\, e^{-r_s t} \sum_{i=0}^{i=n} \frac{(r_s)^i\, (t+t_o)^i\, \gamma([n+n_o-i+1], r_{bmax}[t+t_o])}{i!\,(n-i)!}}{\sum_{j=0}^{j=n} \frac{(1+\frac{t_o}{t})^j\, \gamma([n+n_o-j+1], r_{bmax}[t+t_o])\, \gamma(j+1, r_{smax}\, t)}{j!\,(n-j)!}}$$

* r_{smax} may be obtained from an existing upper limit on the signal rate (or flux)
* In the case that $r_{bmax}[t+t_o] \gg n+n_o$
  we have $\gamma([n+n_o-i+1, r_{bmax}[t+t_o]) \approx (n+n_o-i)!$
  and $\gamma([n+n_o-j+1, r_{bmax}[t+t_o]) \approx (n+n_o-j)!$

- See function GetSignalRatePDF() of the NcAstrolab facility

  in the NCFS-Pack analysis framework (https://nick-ve.github.io/ncfs/docs)

Summary

Going from the observed signal rate pdf to an incoming flux pdf

- Just plug in the Effective Area (A_{eff})

Posterior flux pdf

$$\boxed{p(F_s|n,t,n_o,t_o) = \frac{p(r_s|n,t,n_o,t_o)}{A_{eff}}}$$

Credibility Region and Upper Limit determination

- Can be directly determined from the signal pdf $p(r_s|n,t,n_o,t_o)$
- See GetCredibleInterval() and GetUpperLimit() of the NcAstrolab facility in the NCFS-Pack analysis framework (https://nick-ve.github.io/ncfs/docs)