1 The Likelihood Formalism

This section describes the mathematical likelihood formalism used in Skylab. First it introduces the log-likelihood approach, second the likelihood ratio test and the used test statistic and then describes the used optimizations.

1.1 The Log-Likelihood Approach

Skylab implements the two-component likelihood approach with a likelihood function $\mathcal{L}(n_\mathrm{s}, \vec{p}_\mathrm{s} \mid D)$ of the form

$$\mathcal{L}(n_\mathrm{s}, \vec{p}_\mathrm{s} \mid D) = \prod_{i=1}^{N} \left[\frac{n_\mathrm{s}}{N} \mathcal{S}_i(\vec{x}_\mathrm{s}, \vec{p}_\mathrm{s}) + (1 - \frac{n_\mathrm{s}}{N}) \mathcal{B}_i \right], \tag{1}$$

where n_s is the number of signal events, hence, $(N - n_\mathrm{s})$ the number of background events in the dataset D of N total events. The set of signal model parameters is denoted as $\vec{p}_\mathrm{s}$. Signal model parameters are for instance the spectral index γ of the source. $\mathcal{S}_i(\vec{x}_\mathrm{s}, \vec{p}_\mathrm{s})$ is the value of the signal PDF assuming a signal source at location $\vec{x}_\mathrm{s}$ for the ith data event, whereas $\mathcal{B}_i$ is the value of background PDF of the ith data event.

The signal and background PDFs must incorperate the detector efficiency (yield) $\mathcal{Y}_i$, which usually is dependent on the data event's sky location, energy, and time.

For computational stability reasons the logarithm of the likelihood function of equation 1 is used in Skylab:

$$\log \mathcal{L}(n_\mathrm{s}, \vec{p}_\mathrm{s} \mid D) = \sum_{i=1}^{N} \log(...) \tag{2}$$

1.2 Likelihood Ratio Test and Test Statistic

For estimating the significance of an observation, the likelihood ratio Λ with respect to a null hypothesis of no observation, i.e. equation 1 at $n_\mathrm{s} = 0$ is tested:

$$\log \Lambda(n_\mathrm{s}, \vec{p}_\mathrm{s}) = \log \frac{L(n_\mathrm{s}, \vec{p}_\mathrm{s})}{L(n_\mathrm{s} = 0)} = \sum_{i=1}^{N} \log \left[1 + \frac{n_\mathrm{s}}{N} \left(\frac{\mathcal{S}_i(\vec{x}_\mathrm{s}, \vec{p}_\mathrm{s})}{\mathcal{B}_i} - 1 \right) \right] \tag{3}$$

By defining

$$\mathcal{X}_i(\vec{p}_\mathrm{s}) \equiv \frac{1}{N} \left(\frac{\mathcal{S}_i(\vec{x}_\mathrm{s}, \vec{p}_\mathrm{s})}{\mathcal{B}_i} - 1 \right), \tag{4}$$

this reads as:

$$\log \Lambda(n_\mathrm{s}, \vec{p}_\mathrm{s}) = \sum_{i=1}^{N} \log(1 + n_\mathrm{s} \mathcal{X}_i(\vec{p}_\mathrm{s})) \tag{5}$$

This leads to the test statistic TS

$$\mathrm{TS} = 2\mathrm{sgn}(n_\mathrm{s}) \log \Lambda(n_\mathrm{s}, \vec{p}_\mathrm{s}) \tag{6}$$

with separation of over- ($n_\mathrm{s} > 0$) and under-fluctuations ($n_\mathrm{s} < 0$).

1.3 Multiple Datasets

With Skylab a set of J different data samples (datasets) D_j can be analyzed at once. Each data sample has its own detector signal efficiency $\mathcal{Y}_{\mathrm{s},j}$.

The composite likelihood function is the product of the individual dataset likelihood functions:

$$\log \Lambda = \sum_{j=1}^{J} \log \Lambda_j \tag{7}$$

The total number of signal events n_s needs to get split-up into $n_{\mathrm{s},j}$ for the individual data samples. The distribution of n_s along the different data samples is based on the detector signal efficiency $\mathcal{Y}_{\mathrm{s},j}$ of each sample:

$$n_{s,j}(\vec{p}_\mathrm{s} \; | \vec{x}_\mathrm{s}) = n_\mathrm{s} \frac{\mathcal{Y}_{\mathrm{s},j}(\vec{x}_\mathrm{s}, \vec{p}_\mathrm{s})}{\sum_j \mathcal{Y}_{\mathrm{s},j}(\vec{x}_\mathrm{s}, \vec{p}_\mathrm{s})}, \tag{8}$$

where the parameter $\vec{x}_\mathrm{s}$ denotes the location(s) of the source(s). The parameter vector $\vec{p}_\mathrm{s}$ contains additional source hypothesis parameters, for instance the spectral index γ.

By defining the sample weight factor

$$f_j(\vec{p}_\mathrm{s} \; | \vec{x}_\mathrm{s}) \equiv \frac{\mathcal{Y}_{\mathrm{s},j}(\vec{x}_\mathrm{s}, \vec{p}_\mathrm{s})}{\sum_j \mathcal{Y}_{\mathrm{s},j}(\vec{x}_\mathrm{s}, \vec{p}_\mathrm{s})} \tag{9}$$

with the property

$$\sum_{j=1}^{J} f_j = 1 \tag{10}$$

equation 8 reads

$$n_{s,j}(\vec{p}_\mathrm{s} \; | \vec{x}_\mathrm{s}) = n_\mathrm{s} f_j(\vec{p}_\mathrm{s} \; | \vec{x}_\mathrm{s}) \tag{11}$$

The detector signal efficiency $\mathcal{Y}_{\mathrm{s},j}(\vec{p}_\mathrm{s}, \vec{x}_\mathrm{s})$ depends on the signal model parameters $\vec{p}_\mathrm{s}$ and the source location(s) $\vec{x}_\mathrm{s}$ and can be calculated via the detector effective area and the source flux.

For a single point source the sample weight factor can be calculated via the effective area $A_{\mathrm{eff},j}(E)|_{\vec{x}_\mathrm{s}}$ at the source location of each data sample, and the differential flux $\frac{\mathrm{d}\Phi_\mathrm{s}}{\mathrm{d}E}$ of the source.

$$f_j(\vec{p}_\mathrm{s} \; | \vec{x}_\mathrm{s}) = \frac{\int_0^\infty \mathrm{d}E A_{\mathrm{eff},j}(E)|_{\vec{x}_\mathrm{s}} \frac{\mathrm{d}\Phi_\mathrm{s}}{\mathrm{d}E}(E, \vec{p}_\mathrm{s})}{\sum_{i=1}^{J} \int_0^\infty \mathrm{d}E A_{\mathrm{eff},i}(E)|_{\vec{x}_\mathrm{s}} \frac{\mathrm{d}\Phi_\mathrm{s}}{\mathrm{d}E}(E, \vec{p}_\mathrm{s})} \tag{12}$$

Using the sample weight factor $f_j(\vec{p}_\mathrm{s}\,|\vec{x}_\mathrm{s})$ the likelihood ratio of equation (7) with equation (5) can now be written as

$$\log\Lambda(n_\mathrm{s},\vec{p}_\mathrm{s}) = \sum_{j=1}^{J}\sum_{i=1}^{N}\log(1+n_\mathrm{s}f_j(\vec{p}_\mathrm{s}\,|\vec{x}_\mathrm{s})\mathcal{X}_i(\vec{p}_\mathrm{s})) \tag{13}$$

For multiple point sources, i.e. a stacking of K point sources with positions $\vec{x}_{\mathrm{s},k}$, the sample weight factor of each single source needs to be taking into account. Thus, f_j can be written as the sum of the products of the sample weight factor $f_j(\vec{p}_\mathrm{s}\,|\vec{x}_{\mathrm{s},k})$ for source k and the relative strength $f_k(\vec{p}_\mathrm{s}\,|\vec{x}_{\mathrm{s},k})$ of the kth source in all data samples compared to all the other sources in all data samples.

$$f_j(\vec{p}_\mathrm{s}\,|\vec{x}_\mathrm{s}) = \sum_{k=1}^{K} f_j(\vec{p}_\mathrm{s}\,|\vec{x}_{\mathrm{s},k})f_k(\vec{p}_\mathrm{s}\,|\vec{x}_{\mathrm{s},k}) \tag{14}$$

The relative strength of source k can be written as

$$f_k(\vec{p}_\mathrm{s}\,|\vec{x}_{\mathrm{s},k}) = \frac{\sum_{i=1}^{J}\int_0^\infty \mathrm{d}E A_{\mathrm{eff},i}(E)|_{\vec{x}_{\mathrm{s},k}}\frac{\mathrm{d}\Phi_\mathrm{s}}{\mathrm{d}E}(E,\vec{p}_\mathrm{s})}{\sum_{\kappa=1}^{K}\sum_{i=1}^{J}\int_0^\infty \mathrm{d}E A_{\mathrm{eff},i}(E)|_{\vec{x}_{\mathrm{s},\kappa}}\frac{\mathrm{d}\Phi_\mathrm{s}}{\mathrm{d}E}(E,\vec{p}_\mathrm{s})} \tag{15}$$

Combining equation 12 with $\vec{x}_\mathrm{s}\equiv\vec{x}_{\mathrm{s},k}$ and 15 leads to the final expression for f_j for multiple sources:

$$f_j(\vec{p}_\mathrm{s}\,|\vec{x}_\mathrm{s}) = \frac{\sum_{k=1}^{K}\int_0^\infty \mathrm{d}E A_{\mathrm{eff},j}(E)|_{\vec{x}_{\mathrm{s},k}}\frac{\mathrm{d}\Phi_\mathrm{s}}{\mathrm{d}E}(E,\vec{p}_\mathrm{s})}{\sum_{i=1}^{J}\sum_{k=1}^{K}\int_0^\infty \mathrm{d}E A_{\mathrm{eff},i}(E)|_{\vec{x}_{\mathrm{s},k}}\frac{\mathrm{d}\Phi_\mathrm{s}}{\mathrm{d}E}(E,\vec{p}_\mathrm{s})} \tag{16}$$

One should note that the numerator of equation (16) is one of the summands of the sum in the denumerator, i.e. for $i=j$.

1.4 Optimizations

For point-source like signal hypothesis most of the events in the data sample will be far away from the hypothesised source, hence, the value of the signal PDF $\mathcal{S}_i$ will be zero or very close to zero. By selecting only the signal-contributing N' events from the sample, the likelihood ratio $\log\Lambda$ reads

$$\log\Lambda = \log\Lambda_{N'} + (N-N')\log(1-\frac{n_\mathrm{s}}{N}) \tag{17}$$

The used minimizer (L-BFG-S) operates most stable and fast when provided with gradients of the likelihood function.

TODO: Derive the gradients of the LH function.

2 Detector Signal Efficiency

The detector signal efficiency $\mathcal{Y}_{\mathrm{s},j}(\vec{x}_{\mathrm{s},k},\vec{p}_\mathrm{s})$ of a data sample j for a source k is defined as the integral over the energy of the product of the detector effective area and the differential flux $\frac{\mathrm{d}\Phi}{\mathrm{d}E}$ of the source

$$\mathcal{Y}_{\mathrm{s},j}(\vec{x}_{\mathrm{s}}, \vec{p}_{\mathrm{s}}) \equiv \int_0^\infty \mathrm{d}E A_{\mathrm{eff},j}(E)|_{\vec{x}_{\mathrm{s},k}} \frac{\mathrm{d}\Phi}{\mathrm{d}E}(E, \vec{p}_{\mathrm{s}}) T_{\mathrm{live},j} \quad (18)$$

2.1 Effective Area

In Skylab the effective area $A_{\mathrm{eff},j}$ of a data sample is not calculated separately in order to avoid binning effect. The effective area is calculated using the monte-carlo weights `mcweight`[1] of the simulation events. The monte-carlo weights have the unit GeV cm^2 sr. Using the monte-carlo weight, w_m, of the mth event the effective area is given by

3 Implemented Likelihood Models

This section describes the implemented likelihood models.

3.1 Time Dependent Point-Source Flare

The `TimeDepPSFlareLHModel` class provides the likelihood model for searching for a point source with unknown time-dependence. The search is based on the formulism described in [1].

The model utilizes a two-component likelihood function with signal and background events.

References

[1] Jim Braun, Mike Baker, Jon Dumm, Chad Finley, Albrecht Karle, Teresa Montaruli. Time-Dependent Point Source Search Methods in High Energy Neutrino Astronomy. *Astropart.Phys.*, 33:175–181, 2010.

[1] In IceCube known as "OneWeight".