Clear Sky

Background Light

Exercise adapted from Josh Wood
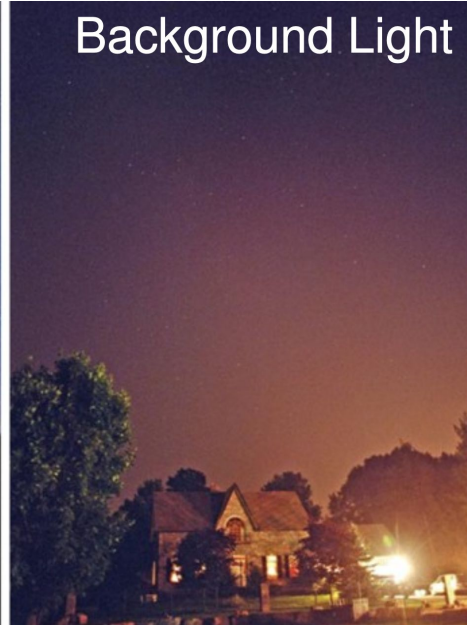
Point-Source Analyses and Likelihood

Alina Kochocki
June 4th, 2024
UW Madison – IceCube
Summer School

Evidence for Astrophysical Neutrinos

*https://doi.org/10.1126/science.1242856*
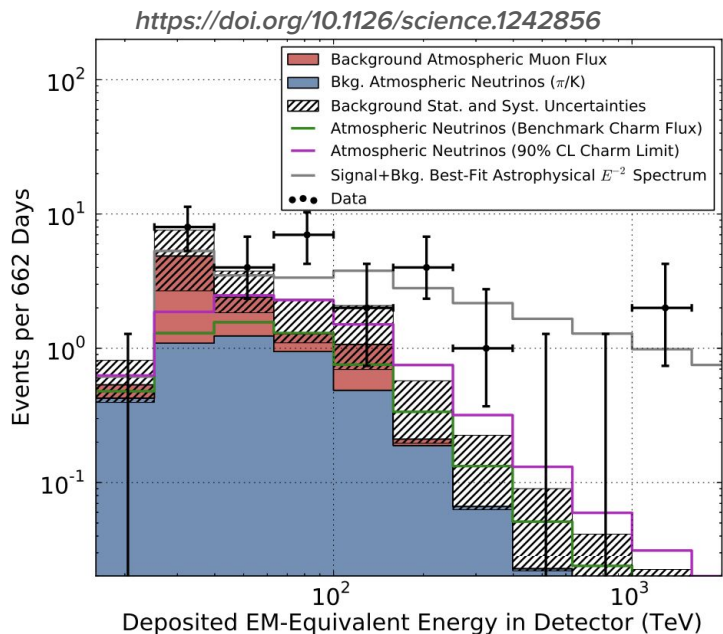
- Established background of atmospheric muons and neutrinos – knowledge of energy spectrum *predates IceCube*

- Around 2012/2013, individual, high-energy events observed, seemingly inconsistent with background

- Idea: model the 'astrophysical signal' with a new, isotropic distribution of events

- A model which included this astrophysical signal and background gave a superior fit to this excess of high-energy events compared to just background alone (*null hypothesis - no new signal*)

Alina Kochocki

2

Evolving 'Diffuse' into 'Nu-Sources'

- We have strong evidence for astrophysical neutrinos based on diffuse analyses (recent result pictured)

- Diffuse analyses *primarily* tell us about the energy and flavor of the composite set of astrophysical sources. Why go further?

- We have additional, existing knowledge about how sources of neutrinos should look

  - Example, more general: expect extragalactic neutrino sources to be point-like with continuous energy spectra

  - Example, less general: expect neutrinos from known, high-energy sources, like AGN, develop models for nu emission based on multiwavelength spectra and variability...

https://arxiv.org/pdf/1908.09551
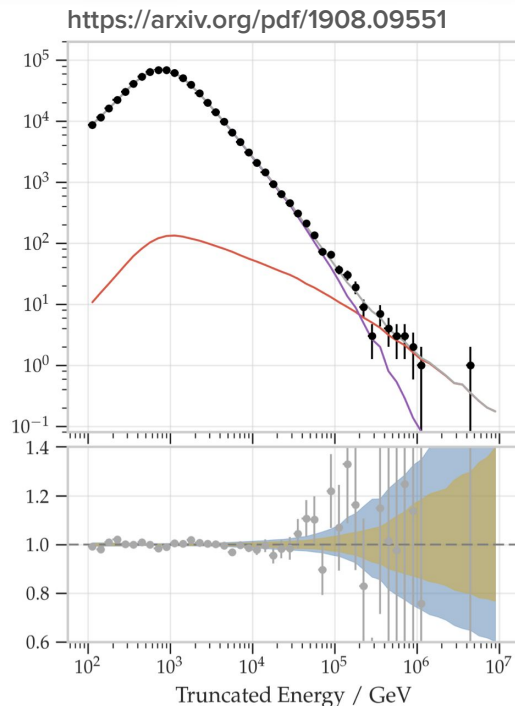
Evolving 'Diffuse' into 'Nu-Sources'

- Nu-sources analyses are *specific tests* which *answer specific questions* about proposed neutrino sources

  - Results can tell us about source physics

- We are still in the 'discovery phase', building evidence to establish specific point-source populations

  - The more information you use to describe your astrophysical signal expectation, the better job you will do of differentiating from background

  - As you differentiate better from background, you will measure smaller neutrino fluxes with higher confidence

- At right, event rate as a function of energy for the *Northern Tracks* diffuse selection – track-like events from the northern hemisphere (lower background). Use this selection?

https://arxiv.org/pdf/1908.09551
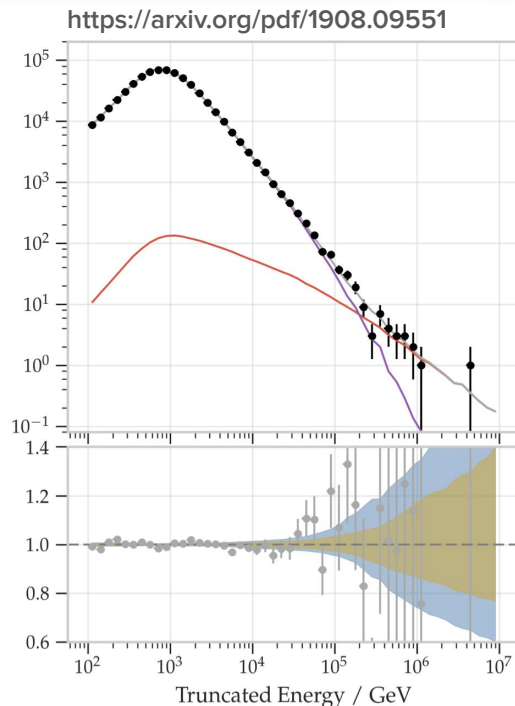
Poisson Distribution

- Why care about the 'confidence' of our measurements?

- Arrival of events described by Poisson distribution

  - Given an expected (mean) number of events, **M**, in time **T**, what is the probability of observing **N** events?

$$Pr(X = N) = \frac{M^N e^{-M}}{N!}$$

- Arrival of next event not controlled by arrival of previous!

- Probability distribution varies with **M** expected events

- *Standard deviation of distribution for N observed events is sqrt(N)*
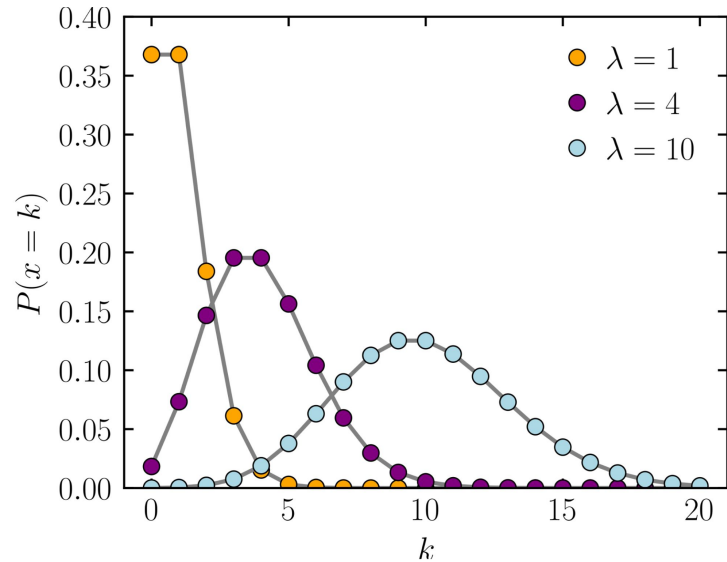
Poisson distributions for different, expected mean values. Can't observe negative events! Image credit: Wikipedia.

Signal-to-Noise Mini Demo

- Example: observe a nearly steady rate (some seasonal variations) of atmospheric background events in **Northern Tracks** selection, say 73000 events/year
    - Astrophysical rate is ~**150 per year**
    - As a function of time, compare astrophysical rate to atmospheric *variance*

- Now, consider that 5% of the astrophysical flux comes from > 100 TeV, but only 1/10,000 of the atmospheric does

- How does the astrophysical rate compare with the atmospheric Poisson uncertainty, sqrt(N), in either case?
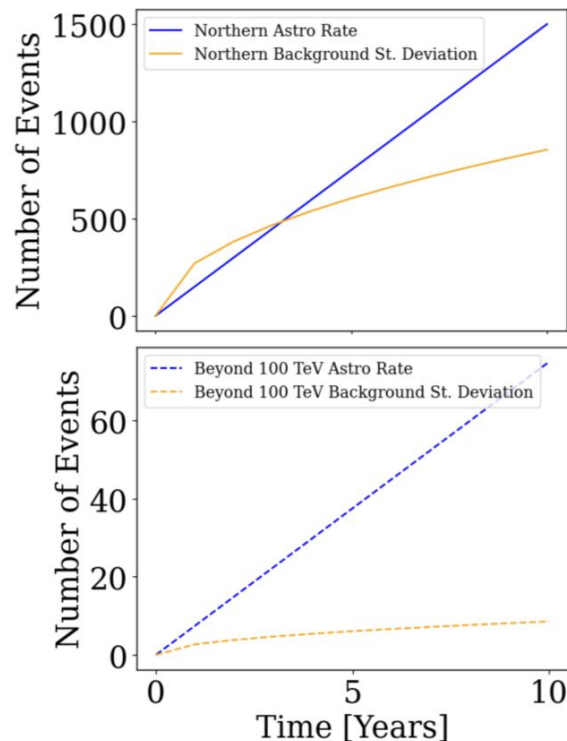
We exploit the fact that signal and background have different energy distributions!

We become more confident of our signal sooner

Exploit Event Observables

- **Reconstructed IceCube events characterized by:**

  - Direction
  - Energy
  - Interaction Point (interaction containment?)
  - Topology (linked to flavor)
  - Time

- **These are all *observables* which have some uncertainty from reconstruction**

- **Either through selections or through signal model descriptions, we exploit the above observables to further differentiate our astrophysical signal from our background**

- **Test specific signal hypotheses ➝ benefit of becoming sensitive to smaller signals, generally**

DOI: 10.1126/science.abg3395
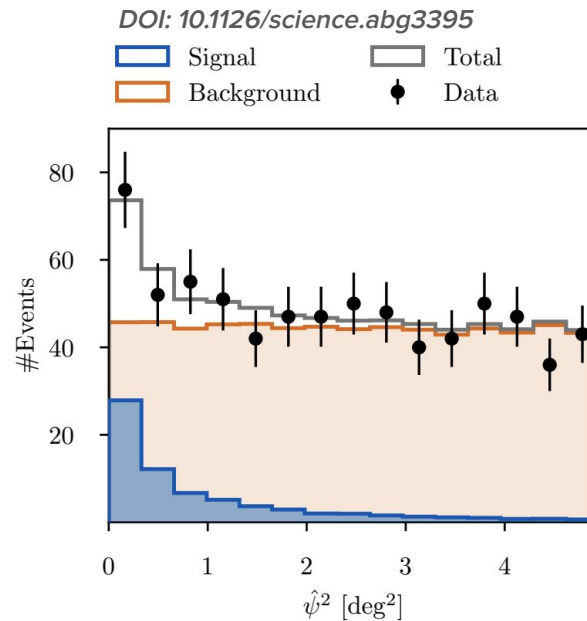
Spatial excess of events around NGC 1068, utilizes good pointing of Northern Tracks

Point-Source Exercise

Distant neutrino sources, especially outside of our own galaxy, will appear as 'point sources'. Any extension associated with the size of the source is negligible compared to our directional reconstruction uncertainty.

We will now build a 'point-source' analysis, following an exercise developed by Josh Wood. This is an introduction to the statistical methods used in the neutrino-sources group...

What to do? Think about it!

Imagine you're an astronomer looking for a point source in the presence of uniform background.

source

background light/neutrinos

light/neutrinos
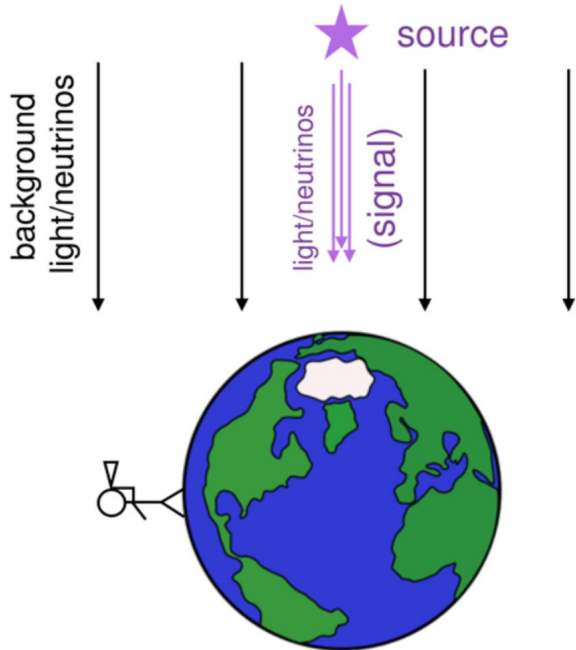
(signal)

def: **signal** is a particle that came from the source you're looking for

def: **background** is a particle that did not come from the source but looks identical to a particle emitted by the source
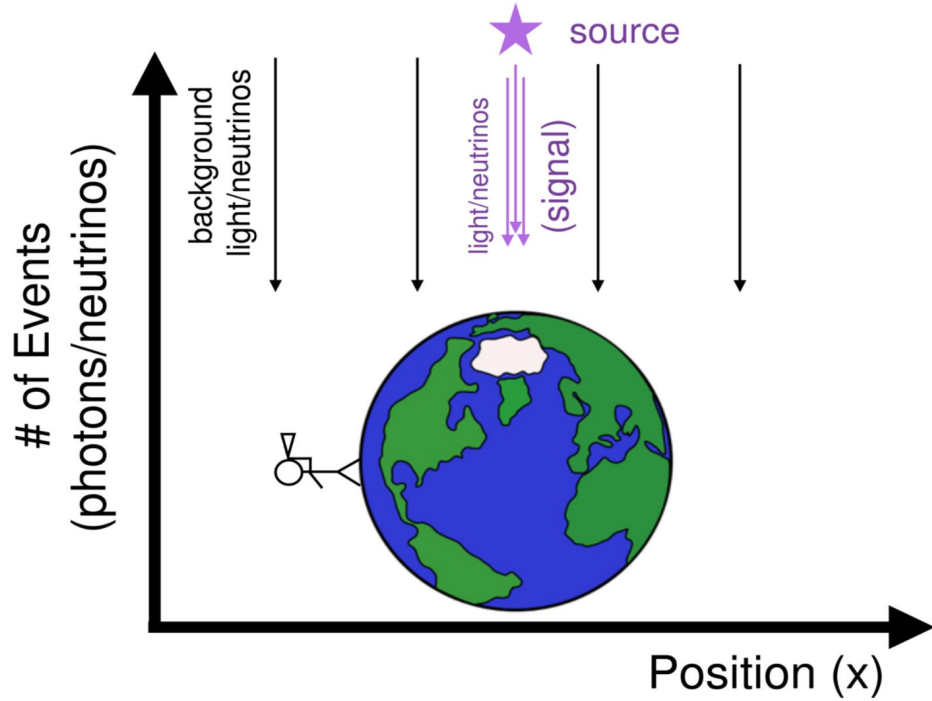
Ex. photon/neutrino with same energy as one from the source

def. **event** a detected particle. Can be photon, neutrino etc.
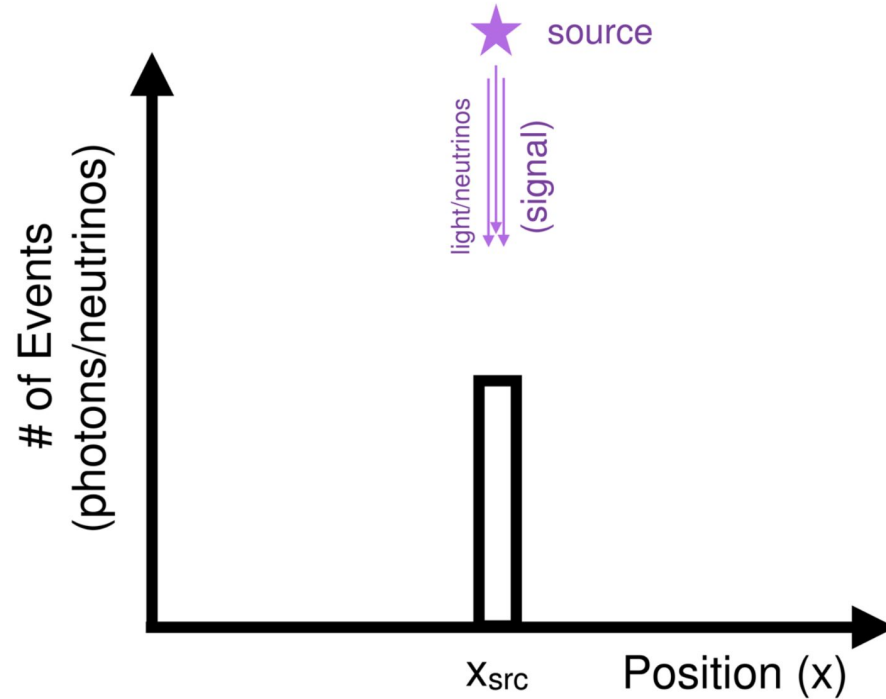
Formalism Part I: Spatial Distributions

Let's start by adding some axes to our example

Formalism Part I: Spatial Distributions

Let's start by adding some axes to our example

Formalism Part I: Spatial Distributions
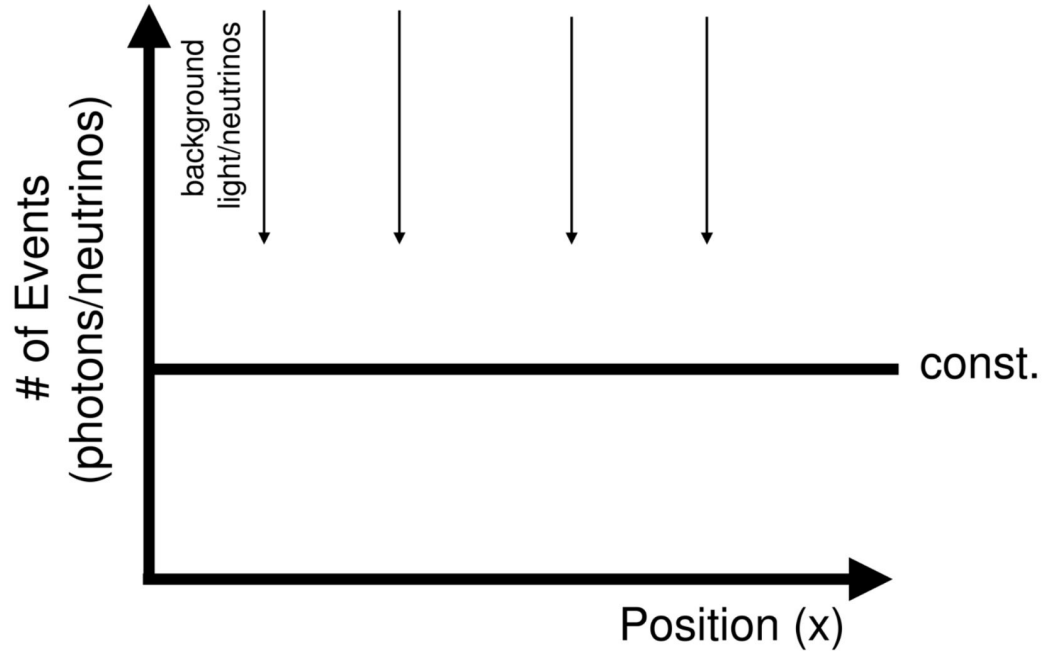
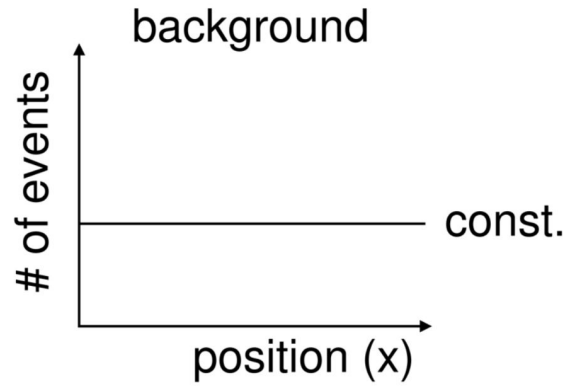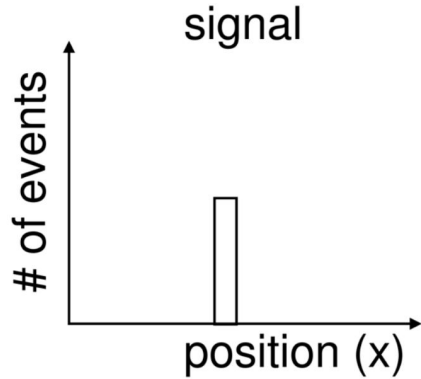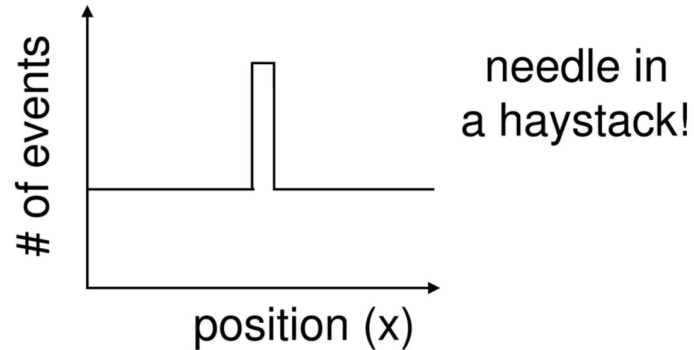Let's start by adding some axes to our example

Formalism Part I: Spatial Distributions

signal

# of events

position (x)

background

# of events

const.

position (x)

data = signal + background

# of events

needle in a haystack!

position (x)

Formalism Part II: Probabilities

Now that we know what signal & background distributions look like, we can formulate them in terms of **probabilities**

<u>def.</u> **probability** is the chance of getting a given result out of the total number of outcomes.

—> ranges 0 to 1 (never to always)
—> sum of all outcomes must be 1

This way we can ask the question: what is the probability that our data are consistent with background + signal versus the case of background only?

—> quantify if a point source is present in data

Formalism Part II: Probabilities

Ok, let's turn our distributions of events into probability densities —> scale such that integral of distribution is 1

signal

# of events

position (x)
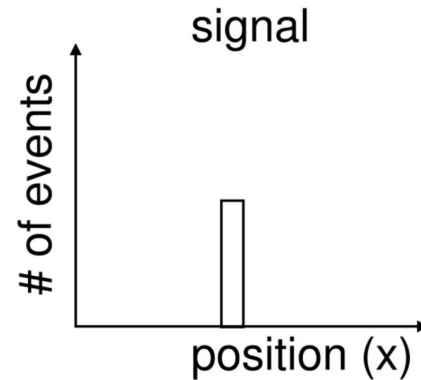
Formalism Part II: Probabilities

Ok, let's turn our distributions of events into probability densities —> scale such that integral of distribution is 1

signal

$S(x)$ = probability density of finding signal at x

$S(x)\ dx$ = probability of finding signal within dx of x

Formalism Part II: Probabilities

Ok, let's turn our distributions of events into probability densities —> scale such that integral of distribution is 1
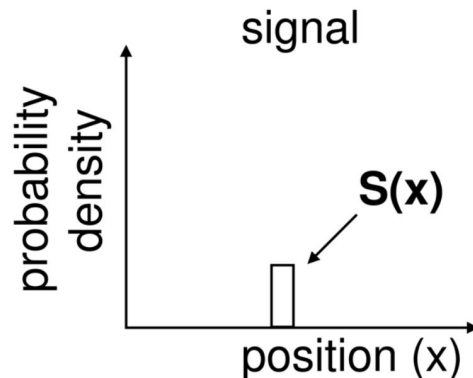
background

const.

position (x)

# of events

Formalism Part II: Probabilities

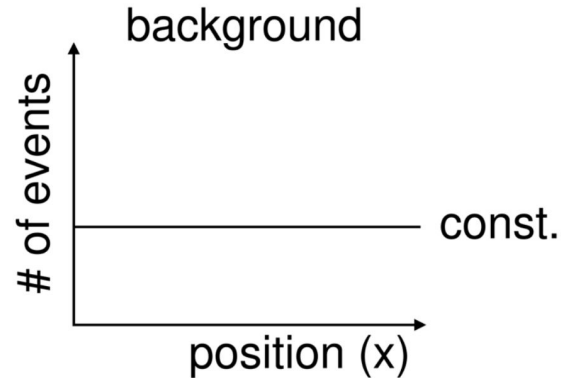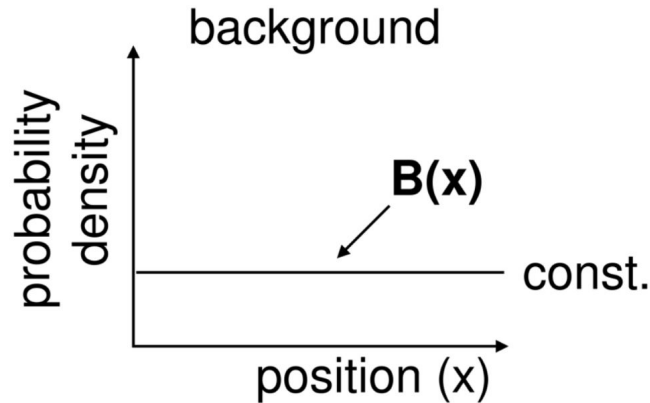Ok, let's turn our distributions of events into probability densities —> scale such that integral of distribution is 1

B(x) = probability density of finding background at x

In astronomy, we typically work on surface of a sphere $\longrightarrow$ uniform **B(x)** = $1/4\pi$

Formalism Part II: Probabilities

For a dataset with:
- → **N** total events
- → **n$_s$** signal events
- → **x$_i$** is the position where we detect the ith event (i $\in$[1, N])

total ith signal prob.   total ith background prob.

$$\frac{n_s}{N} * S(x_i) \quad + \quad (1 - \frac{n_s}{N}) * B(x_i)$$

probability density   probability density
of signal at **x$_i$**   of background at **x$_i$**

probability ith event   probability ith event
is a signal event   is a background event

Formalism Part II: Probabilities

For a dataset with:
→ **N** total events
→ $\mathbf{n_s}$ signal events
→ $\mathbf{x_i}$ is the position where we detect the ith event (i $\in$[1, N])

total i^{th} signal prob.   total i^{th} background prob.

$$\frac{\mathbf{n_s}}{\mathbf{N}} * \mathbf{S(x_i)} \qquad + \qquad (1 - \frac{\mathbf{n_s}}{\mathbf{N}}) * \mathbf{B(x_i)}$$

↑   ↑

probability density   probability density
of signal at $\mathbf{x_i}$ of background at $\mathbf{x_i}$

probability i^{th} event   probability i^{th} event
is a signal event   is a background event

Formalism Part II: Probabilities

For a dataset with:
→ **N** total events
→ **n$_s$** signal events
→ **x$_i$** is the position where we detect the ith event (i $\in$[1, N])

$$L(n_s) = \prod_{i=1}^{N} \frac{n_s}{N} * S(x_i) + (1 - \frac{n_s}{N}) * B(x_i)$$

How to interpret the likelihood optimization?

L(n$_s$) is the total probability of the dataset containing **n$_s$** signal events. It is called a **likelihood function**.

The best estimate for the true value of **n$_s$** is the value which **maximizes the likelihood function**.

Formalism Part II: Probabilities

For a dataset with:
- $\longrightarrow$ **N** total events
- $\longrightarrow$ **n_s** signal events
- $\longrightarrow$ **x_i** is the position where we detect the ith event ($i \in [1, N]$)

$$L(n_s) = \prod_{i=1}^{N} \frac{n_s}{N} * S(x_i) \; + \; (1 - \frac{n_s}{N}) * B(x_i)$$

Each event contributes to the joint probability of our hypothesis data description (relative scale between background and signal PDFs tuned by ns). This is a fit to the data set.

L(n_s) is the total probability of the dataset containing **n_s** signal events. It is called a **likelihood function**.

The best estimate for the true value of **n_s** is the value which **maximizes the likelihood function**.

$$L(n_s) = \prod_{i=1}^{N} \frac{n_s}{N} * S(x_i) + \left(1 - \frac{n_s}{N}\right) * B(x_i)$$

- **Consider our uniform background PDF and expected signal in a range dx**
  - **Say, 10 of 100 events fall into the specific range dx. All events correspond to a location of nonzero background probability, but only our ten events can have nonzero signal probability**

- **In one extreme, set ns to zero (top right, red). Background expectation is too high, will inflate probability contribution for all events outside of dx, *but* events within dx now contribute substantially less, this drives the probability product down**

- **In the other extreme, set ns to N. Now, all events are seen as signal, and background expectation is zero. The signal term will be large for each of the ten events, but zero for all others. This will drive the likelihood down (evaluating to zero for our extreme signal PDF)**

# of events

position (x)

# of events

position (x)
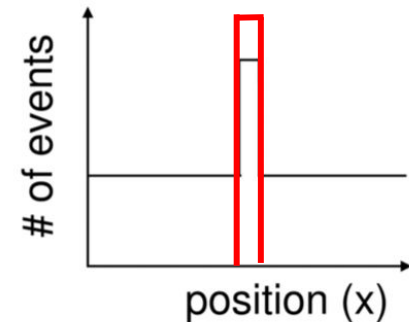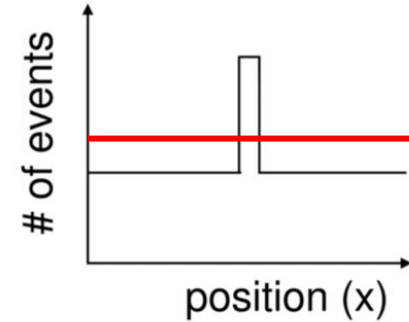
Formalism Part III: Voodoo

Working with **ratios of likelihoods** has some nice statistical properties.

Also, addition is easier and faster than multiplication so we define a **test statistic (TS):**

$$TS = 2 \log(L(n_s) / L(n_s = 0))$$

$$TS = 2 \sum_{i=1}^{N} \log\left[\frac{n_s}{N} * \frac{S(x_i)}{B(x_i)} + (1 - \frac{n_s}{N}) \right]$$

Finding the value of n_s which maximizes TS is equivalent to maximizing the likelihood. TS = 0 means consistent with background only. TS ~ 25 typically proof of a point source.

Existing Resources for Nu-Sources Analyses

- Analysis software with likelihood implementations:

    - *csky* – https://github.com/icecube/csky – likelihood implementations support common use cases within the collaboration, optimization in C++

    - *skyllh* – https://github.com/icecube/skyllh – easy access to vary signal and background PDFs

    - *skylab* – https://github.com/icecube/skylab – historical software optimized for certain use cases, slowly being retired

- When your analysis is developed and approaching unblinding, related scripts for the study will be reviewed and archived at: https://github.com/icecube/wg-nu-sources

    - This repository is also a great resource for studying example source analyses!

Existing Resources for Nu-Sources Analyses

- Common analysis-level event selections, further info at the nu-sources selection page, https://wiki.icecube.wisc.edu/index.php/Nu_Sources_Datasets:
  - *Point-Source Tracks* – all-sky sample of track events
  - *Northern Tracks* – tracks events from the northern hemisphere, cut substantial muon background in the south
  - *ESTES (Enhanced Starting Track Event Selection)* – all-sky sample of starting tracks, likely neutrino interaction in detector
  - *DNN Cascades* – all-sky cascades with improved reconstruction
  - *GFU (Gamma-ray Follow-up)* – all-sky tracks intended for fast response analyses
  - *GRECO* – all flavor, low-energy selection (5 GeV - 10 TeV)

- If you target source is not yet established at high-significance, consider theory predictions for your source class, and choose the selection that optimizes sensitivity to your source hypothesis

- You may also benefit from combining datasets for improved statistics and signal coverage

Point-Source Exercise

1. Log in to your cobalt computing account
   a. ssh username@pub.icecube.wisc.edu
   b. ssh cobalt
2. Clone this exercise repository to your account
   a. git clone https://github.com/icecube/SummerSchool2024.git
3. Now, go to https://jupyterhub.icecube.wisc.edu/ and log in to access computing environment through browser :)
   a. Select CPU environment
   b. Navigate to SummerSchool2024/point_sources directory
   c. Open likelihood-example.ipynb (do not look at likelihood-example-answers.ipynb!)
   d. Select the python environment in upper-right corner. Choose py3-v4.3.0: icetray/v1.9.2

Generalizing the PDFs for Multiple Observables

$$L(n_s) = \prod_{i=1}^{N} \frac{n_s}{N} * S(x_i) + (1 - \frac{n_s}{N}) * B(x_i)$$

- We have just worked through a spatial likelihood implementation. The signal and background PDFs were probabilities of a spatial observable

- We can generalize this formula to utilize information from multiple observables. The signal and background PDFs will become joint PDFs of multiple observable distributions

$$S(\alpha_i, \delta_i, E_i, t_i) = S(\alpha_i)S(\delta_i)S(E_i)S(t_i)$$

$$B(\alpha_i, \delta_i, E_i, t_i) = B(\alpha_i)B(\delta_i)B(E_i)B(t_i)$$

Above, position, x, is replaced by astronomical units, right ascension and declination, as well as energy and time

$$\int S(\alpha)d\alpha = \int B(\alpha)d\alpha = 1$$

$$\int S(\delta)d\delta = \int B(\delta)d\delta = 1$$

$$\int S(E)dE = \int B(E)dE = 1$$

$$\int S(T)dT = \int B(T)dT = 1$$

Alina Kochocki