Math Methods and Experimental Physics

Probability and uncertainty

Ignacio Taboada

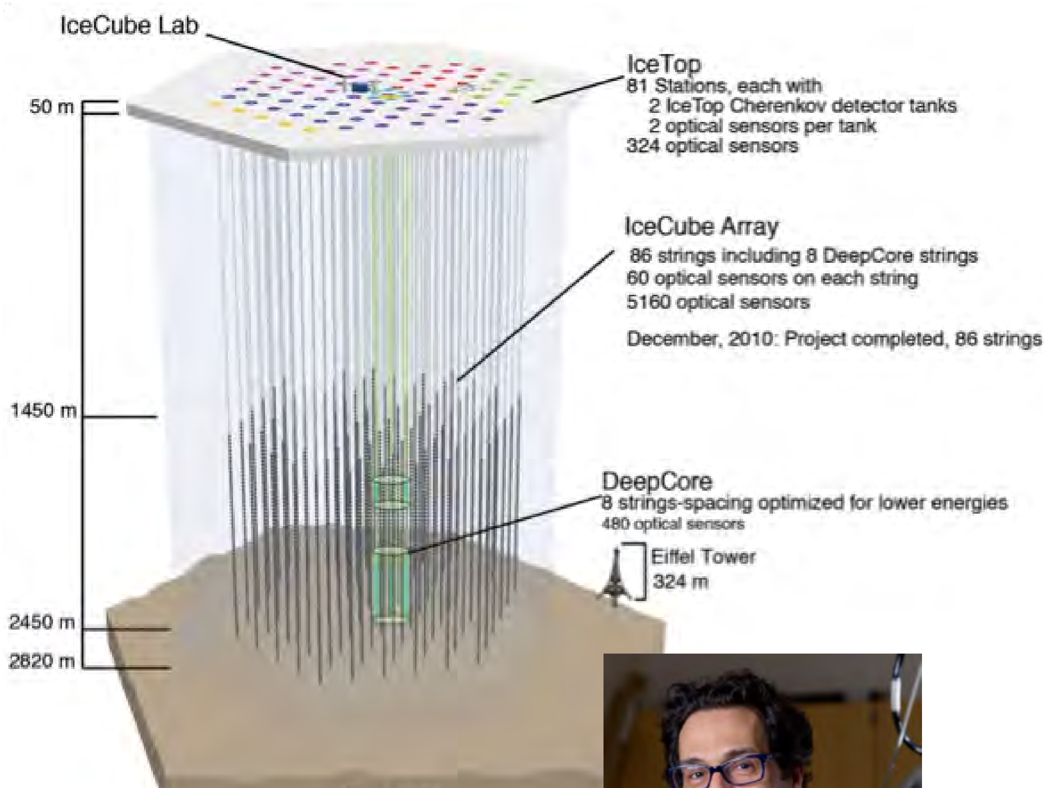
About me

Research: Particle Astrophysics - IceCube - HAWC.
Currently focusing on the search for astrophysical sources of neutrinos.

11 years at Georgia Tech.

Currently 3 PhD students:

Chun Fai (Chris) Tung
Pranav Dave
Chujie Chen



Ignacio Taboada

Shameless plug: **Data science for physicists. Fall 2020**

Course targeted to 2nd year grad students (you!) as well as juniors/senior majoring in physics. The course will be project driven. Probably 5 projects (2 weeks) plus a final project (4 weeks). Basic knowledge of python is assumed.

Topics:

- Python / numpy / matplotlib primer
- Uncertainty, probability, probability density functions, error matrix, etc.
- Random variables, pseudo-random numbers, distributions, moments, quantiles
- Parameter Fitting, least squares, maximum likelihood,
- Hypothesis testing, likelihood ratio, Wilk's theorem, K-S test, Cramer-Von Mises test
- Machine learning, neural networks, boosted decision trees (Final project topic)

Uncertainty (sometimes referred to as error)

Some instances of uncertainty:

- Theory is not deterministic (e.g. Quantum Mechanics)
- Random errors in measurement (e.g. Thermal noise) – aka statistical uncertainty
- Things which could be known but are not (limitations from cost, time) – aka systematic uncertainty

Uncertainties may be quantified using Probability

- The theory of probability is a branch of mathematics. From a specific set of axioms and definitions mathematicians build up the theory by deduction.
- For a physicist, contrast this with statistics, a branch of applied mathematics that is essentially inductive but connected with the theory of probability.

Statistical vs. Systematic errors

As we will see later, the statistical uncertainty of multiple measurements depends on $1/\sqrt{N}$, where N is the number of measurements.

So: take a ruler measure an object one. Say you get 12.1 ± 0.1

If you repeat this measurement, you get: 12.1 ± 0.1

Why is it incorrect to combine both mesurements to get 12.10 ± 0.07

Probability

Mathematical Probability is described by Kolmogorov's Axioms (1933). Let P be a set operation that returns a real number.

If A and B are subsets of S. The axioms are:

$$P(S) = 1$$

For all $A \subset S$; $P(A) \geq 0$

if $A \cap B = \emptyset,\ P(A \cup B) = P(A) + P(B)$

Where $\emptyset$ is the null set.

Any set operation P that satisfies these axioms is a probability.

Frequentist probability

Let’s have a possible outcome A for an experiment. We repeat the experiment n times (*trials*). Let r be the number of times that A is the measured outcome.

Then:

$$P(A) = \lim_{n \to \infty} \frac{r}{n}$$

… see? How frequently does A happen? -> Frequentist probability

Usually, experiments are not performed on infinite sets. For a finite number of trials, n trials of the experiment result in a

point estimate:

$$\tilde{p} = \frac{r}{n}$$

Corollaries

$$P(\overline{A}) = 1 - P(A)$$

$$P(A \cup \overline{A}) = 1$$

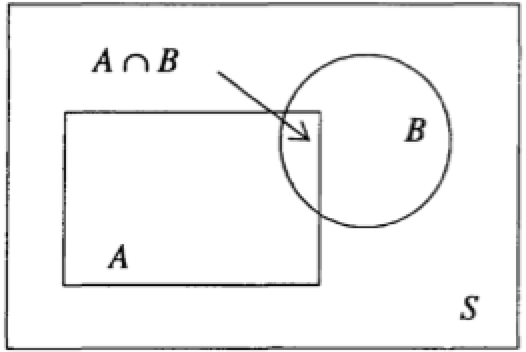
$$0 \leq P(A) \leq 1$$

$$P(\emptyset) = 0$$

$A \subset B$ If $P(A) \leq P(B)$ then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Conditional Probability



Consider A and B as 2 subsets of S.

How can one express the probability of the subset B relative to the subset A? This is what is called the conditional probability of B relative to A

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

← a.k.a.: Prob. of B, given A

Example – for a dice roll:

$$P(n < 3|n_{even}) = \frac{P((n < 3) \cap n_{even})}{P(even)} = \frac{1/6}{3/6} = \frac{1}{3}$$

Conditional Probability

A and B are *defined* to be independent if

$$P(A \cap B) = P(A)P(B)$$

Do not confuse independent with $A \cap B = \emptyset$

For A and B independent:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)P(B)}{P(A)} = P(B)$$

Example. Two dice roll: Probability of dice 1 being even given that dice 2 is odd

$$P(D_1\ even|D_2\ odd) = \frac{P(D_1\ even)P(D_2\ odd)}{D_2\ odd} = P(D_1\ even) = \frac{1}{3}$$

Bayes Theorem

We know that:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{and} \quad P(B|A) = \frac{P(B \cap A)}{P(A)}$$

Because $P(A \cap B) = P(B \cap A)$

This is Bayes theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Looks simple – but it's very important.

Bayes theorem - an example

Let’s describe an unusual illness and a test to diagnose it.

Let’s have this prior probability: $P(\text{disease}) = 0.001$

From which trivially: $P(\text{no disease}) = 0.999$

The prior probability could be found, e.g. via a frequentist measurement.

Bayes theorem - an example

Let's have a test that can be positive or negative with, also prior, conditional probabilities:

$$P(+|\text{disease}) = 0.98$$

$$P(+|\text{no disease}) = 0.03$$

It also follows trivially that:

$$P(-|\text{disease}) = 0.02$$

$$P(-|\text{no disease}) = 0.97$$

These probabilities can be found in clinical trials. Given a set of people that are known to have the disease and a control group with no disease, we use frequentist prob. to find the numbers above. If we go to the Dr. what we really want to know is: $P(\text{disease}|+)$

Bayes theorem - an example

Recall: $$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

So: $$P(\text{disease}|+) = \frac{P(+|\text{disease})P(\text{disease})}{P(+|\text{disease})P(\text{disease}) + P(+|\text{no disease})P(\text{no disease})}$$

Which results in: $$\boxed{P(\text{disease}|+) = 0.032}$$

Thought train: "This has to be wrong, this is a good test". "Oh, I get it. Yes, this makes sense for rare diseases". "OMG!!!! I hope my Dr. has heard of Bayes theorem?".

Well, at least: $$P(\text{no disease}|-) = 0.99998$$

Summary on probability

Probability is central to physics. Some processes are intrinsically random (quantum mechanics). But even for deterministic processes, measurement uncertainty is described in probabilistic language.