More Probability and “Error” Analysis

Road Map

- Probability (continued)
- Types of Uncertainty
- Propagating Uncertainty
Bayes Theorem with P.D.F.s

\[ g(y; x) = h(x; y) g_{\text{prior}}(y) \]

For Example: Take a Gaussian

\[ f(x, y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-y)^2}{2\sigma^2}} \]

Then we can define the conditional probability \( h(x; y) \) as

\[ h(x; y) = \frac{f(x, y)}{f_y(y)} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-y)^2}{2\sigma^2}} \]

\[ \int_{-\infty}^{\infty} dx' \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x'-y)^2}{2\sigma^2}} = f(x, y) \]

A little algebra gives

\[ g(y; x) = \frac{f(x, y) g_{\text{prior}}(y)}{\int_{-\infty}^{\infty} dy' f(x, y') g_{\text{prior}}(y')} \]
Example: The Bad Thermometer

We start with the assumption that the errors on the thermometer are Gaussian with a 5 K uncertainty:

\[
h(T_{obs} ; T_{true}) = f(T_{obs} , T_{true}) = \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{(T_{obs} - T_{true})^2}{2\sigma^2}}
\]

\[\sigma = 5 \text{ K}\]

We need to make a decision if the temperature is cold enough to liquefy helium, i.e. \(P(T_{true} < 4.2 \text{ K})\).

\[
g(T_{true} ; T_{obs}) = h(T_{obs} ; T_{true}) g^{prior}(T_{true})
\]

Since we know that the true temperature must be greater than zero, choose a prior

\[
g^{prior}(T_{true}) = \begin{cases} 1 & \text{for } T_{true} > 0 \\ 0 & \text{for } T_{true} \leq 0 \end{cases}
\]

So we get:

\[
g(T_{true} ; T_{obs}) = \frac{f(T_{obs} , T_{true})}{\int_0^\infty dT_{true} f(T_{obs} , T_{true}^{'})}
\]
The Inverted Probability

The temperature is **probably** below 4.2 K: Your decision?
Some Useful (obvious?) Identities

If A and B are independent, then

\[ P(A \text{ and } B) = P(A) P(B) \]

If A and B are not independent then

\[ P(A \text{ and } B) = P(A | B) P(B) = P(B | A) P(A) \]

\[ P(\text{not } A) = 1 - P(A) \]

\[ P(A \text{ or } \text{not } A) = 1 \]

\[ P(A \text{ and } \text{not } A) = 0 \]
Types of Uncertainty: Statistical

- **Statistical or Random Uncertainties**
  - Do to the intrinsic random nature of processes
  - You can't beat them
    - But, they are easy to handle with statistics
    - In fact, these are the only types of uncertainties that can be correctly handled with statistics
  - Examples: The variations in
    - The number of heads from 4 coin tosses
    - Number of radioactive decays per second
    - The number of people on a subway car at 8:05 am
Types of Uncertainty: Systematic

- Systematic or Instrumental Uncertainty
  - Unknown factors from random processes inside of your instrument
  - Unknown factors due to uncertain knowledge of the experimental conditions
  - Unknown factors due to approximations used during the measurement.
Systematic Uncertainty: Type 1

Unknown factors from random processes in your instruments.

Usually the result of a calibration measurement

- Calibration of a thermometer
- Calibration of an analog to digital converter

Any calibration measurement will (**must**) include an uncertainty

- Bizarrely, the uncertainty on a calibration constant can contain both statistical and systematic contributions!
Systematic Uncertainty: Type 2

- Uncertain knowledge about the experimental conditions
  - No matter what you do, you can't know exactly what is going on in the lab.
  - This will introduce fluctuations in your data that might not be random

- Example: Doing a precision Eotvos experiment
  - The absolute value of the gravitational constant is measured by very accurate torsion balances (Nelson et al, PRD42, 963 (1990))

They spent months chasing down a systematic variation that happened about 4am every other day: It turned out to be caused by the sprinklers going on outside of the lab.
Systematic Uncertainty: Type 3

- Uncertainty knowledge due to approximations and theoretical input
  - Example:
    - Nuclear corrections when measuring the neutrino cross section
    - We want to know the effect of a neutrino on a quark, but we measure the particles after they leave the nucleus
  - Must be treated with extra care since they often contain biases

- No particularly good way to include the effect in a measurement
  - Can't be handled in purely classical statistics → Absolutely subjective
  - Usually punt and treat them as a random uncertainty
    - Make a S.W.A.G at how good the approximation or theory is.
  - Be as “intellectually honest” as possible
    - Don't underestimate the uncertainty
    - It's just as bad to over estimate the uncertainty.
Real Life Systematics
Of Errors, Uncertainties and Mistakes

Mistakes (errors) are not an uncertainties.

- Do you level best to make sure you do not make a mistake
- Do not present a result until you are willing to defend it

However, mistakes do happen

- If a mistake is proven, publicly acknowledge it
- If at all possible correct the problem before anybody else!

Mistakes hurt your scientific credibility
Propagating Uncertainty
Functions of Random Variables

What happens when we need a function of a measurement
- Measure length and width to determine area

\[ A = lw \]

Consider a function \( x(u, v) \) where \( u \) and \( v \) are measurements
- The best estimate of the mean is \( \langle x \rangle = x(<u>, <v>) \)
- The best estimate of the uncertainty is:

\[
\sigma_x^2 = \sigma_u^2 \left( \frac{dx}{du} \right)^2 + \sigma_v^2 \left( \frac{dx}{dv} \right)^2 + \sigma_{uv}^2 \left( \frac{dx}{du} \right) \left( \frac{dx}{dv} \right)
\]

There is an implicit assumption that the uncertainties are Gaussian
Specific Formulas

- Simple Sums (variable and a constant)
  \[ x = u + a \]
  \[ \sigma_x = \sigma_u \]

- Weighted Sums (Add the values in quadrature)
  \[ x = au + bv \]
  \[ \sigma_x^2 = a^2 \sigma_u^2 + b^2 \sigma_v^2 + 2ab \sigma_{uv}^2 \]

- Simple Multiplication
  \[ x = au \]
  \[ \sigma_x = a \sigma_u \]

- Multiplication and Division
  \[ x = uv \]
  \[ \sigma_x^2 = (av \sigma_u)^2 + (bu \sigma_v)^2 \]
More Specific Formulas

- Fractions (Add the relative errors in quadrature)

\[ x = \frac{u}{v} \]

\[ \left( \frac{\sigma_x}{x} \right)^2 = \left( \frac{\sigma_u}{u} \right)^2 + \left( \frac{\sigma_v}{v} \right)^2 + 2 \frac{\sigma_{uv}^2}{uv} \]

- Powers

\[ x = a u^b \]

\[ \frac{\sigma_x}{x} = b \frac{\sigma_u}{u} \]

- Exponents

\[ x = a e^{bu} \]

\[ \frac{\sigma_x}{x} = b \sigma_u \]

- Logarithms

\[ x = a \ln(b u) \]

\[ \sigma_x = a b \frac{\sigma_u}{u} \]
Finally

- We've finished up the basic “theory” of probability

  - It's mostly conceptual. When we use it, there isn't a lot of math

- Always keep in mind the differences between statistical and systematic errors

  - The real work in physics is to determine your uncertainties.
  
  - Understanding systematics is often the hardest and problematic stage of a measurement

- Once you've gotten your uncertainties

  - We'll (almost) always assume they are Gaussian

The End