Error Analysis

Sadhana Dash

November 5, 2022

Outline

Statistical Errors
Systematic Errors
Error propagation
"It is better to be roughly right than precisely wrong ". (Alan Greenspan)

Statistical Uncertainties

measurement = (best estimate $\pm$ uncertainty) units
Error (of measurement) = result of a measurement - the value of the measurand (generally unknown)

Statistical Uncertainties can arise due to stochastic fluctuations. They are uncorrelated with the previous measurements and follow a well-developed statistical theory.

Its usually quantified by the standard deviation $= \sigma$

Bootstrap Method

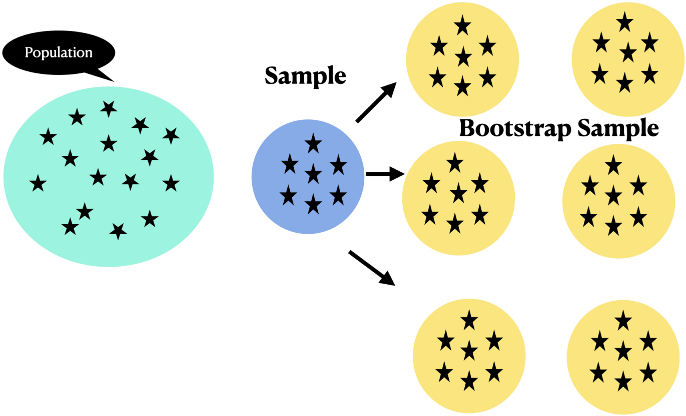

Figure: A schematic of bootstrap method

Sub-Sample Method

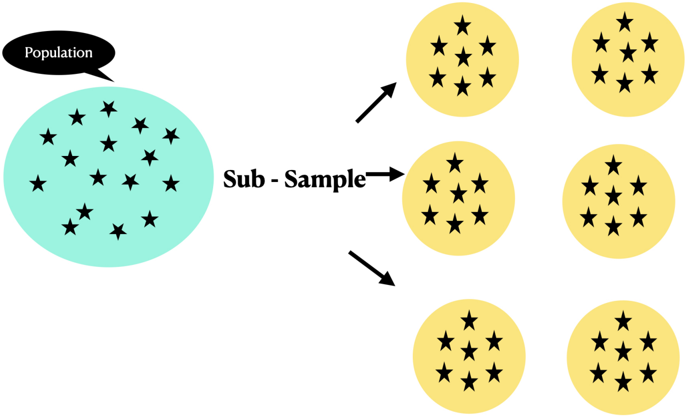

Figure: A schematic of subsample method

Limits and Confidence Intervals

Systematic Uncertainties

Systematic effects is a general category which includes effects such as background, scanning efficiency, energy resolution, variation of counter efficiency with beam position, and energy, dead time, etc. The uncertainty in the estimation of such a systematic effect is called a **systematic error** **It requires ingenuity, experience, cunning and a fair degree of paranoia. These are well provided in particle physics. It also requires honesty and bravery, which are scarcer.** —- R. Barlow

Try to differentiate between

1. **Manifest Systematic Errors** : factors from manifest sources with known errors
   Example : conversion factors and associated errors from calibration/test beam runs
   background contribution estimation on parameters used in MC models.
   Calculate error contribution
2. **Unsuspected systematic Errors** : Variation in detector performance, external conditions, noise, fit ranges etc
   Check for absence of systematic effects, but usually no error contribution.
   Do nothing if shifts of parameters are "small"

Try to repeat the analysis in different forms

1. varying the range of the data used to extract the result
2. varying cuts applied to ensure the purity and quality of the data,
3. including/excluding subsets of data taken under different experimental conditions,
4. using histograms with different bin sizes,
5. determining quantities by simple counting and by fits of a paramatrized curve,

Values obtained by different methods will disagree, even in the absence of systematic effects.

How to decide what is small?

Let us estimate a parameter θ from a data sample by two different methods :

first estimate $\hat{\theta}_1 \pm \sigma_1$

second estimate $\hat{\theta}_2 \pm \sigma_2$

Calculate

the difference between the estimates : $\Delta = \hat{\theta_1} - \hat{\theta_2}$

The error on this difference, $\sigma_\Delta = \sigma 2_1 + \sigma_2 2 - 2\rho\sigma_1\sigma_2$

where ρ is the correlation coefficient.

$\sigma_\Delta = \sqrt{\sigma_1 2 + \sigma_2 2}$. (uncorrelated)

correlated : $\sigma_\Delta = \sqrt{\sigma_1 2 - \sigma_2 2}$ (correlated)

Determine $|\frac{\Delta}{\sigma_\Delta}|$ and check whether $|\frac{\Delta}{\sigma_\Delta}| \leq 2.0$

If yes, do not consider for systematic effect else consider.