Unconference – Incorporating Stats/DS in Physics/STEM curricula

- Theoretical knowledge, practical skills, awareness
- Probability Distributions (uni/multi-variate, mixtures, marginal, conditional, (copula-based) joint distributions)
  Dichotomies between samples vs. populations, distribution properties vs. inter-distribution relations, critical-values vs. probability-values, sampling strategies vs. sampling distributions
- Optimization (objective function & algorithmic)
  Optimization is critical in most AI/ML/DS applications, but often overlooked
- Scientific Inference
  (empirical) Data-driven, (theoretical) model-based, and (computational) MCMC inference. Most DS, AI & ML techniques involve quantifying odds, risk assessment, likelihood estimation
- Student-cohort & Class-specific Modular Embedding
  Inductive vs. deductive concept presentation, order of theory & practice presentation
- Role of coding/IDE/frameworks, EDA vs. CDA

PHYSICS 166/266 @ Stanford

2 Learning Goals

- Understand common probability distributions (e.g. binomial, Poisson, Gaussian, Chi-square etc.), their key properties, and examples of where such distributions occur in physics (and why)
- Be able to state and derive (analytically) key results in probability and statistics, such as the Central Limit Theorem and Cramér–Rao inequality, and verify and understand them conceptually by writing computer simulations
- Be able to define statistical and systematic errors, identify them in real physics research context, and explain how errors are propagated throughout data analysis while properly taking into account correlations
- Understand the theoretical limits of the precision of a given physics measurement and how these can be approached with a given data-set and statistical analysis
- Write codes to perform simple Monte Carlo simulations, parameter estimation, confidence-interval calculation, and hypothesis testing for real physics data analysis
- Be able to interpret statistical data analysis results from physics experiments (e.g. histograms, contour plots, confidence intervals, exclusion limits, etc.)

4 List of Topics

1. Basics of probability theory: Random variables. Conditional probability. Independence and correlation. Bayes' theorem. Mean value and variance. Probability and density distribution. Multivariate probability densities and covariance.
2. Important probability distributions: Definition and applications of the Binomial, Poisson, Uniform, Normal, Exponential, Gamma, Beta, and Cauchy distributions. Binormal distribution and covariance ellipse. Sampling distributions: Chi-squared, Student's t. Characteristic function. Central limit theorem.
3. Measurement errors: Functions of random variables. Types of measurement uncertainty: statistical and systematic errors. Error propagation. Averaging of uncorrelated and correlated measurements.
4. Theory of Estimators: General properties: consistency, convergence, bias, efficiency, sufficiency. Accuracy and Precision. Fisher Information. Lower bound for the variance: Cramer-Rao inequality. Maximum likelihood estimators. The least squares method. Linear regression. The least square method with errors in both variables and with non-linear functions.
5. Confidence intervals: Frequentist confidence intervals. Covariance ellipse. Confidence intervals and belts. Upper and lower limits. Bayesian intervals. Frequentist and Bayesian interpretation of probability. Priors and Posteriors.
6. Hypothesis testing: General properties. Critical region and significance level. Type I, II errors. Power of a test. p-values. The Neyman-Pearson test. The likelihood ratio test.