AST3100: Lecture Series in Specialized Topics: Statistics and Inference in Astrophysics

Spring 2018

(last updated: January 31, 2018)

Course description

This graduate-level course provides an introduction to statistics and probabilistic inference with an eye towards applications in astrophysics. This course will provide you with a good understanding of the basic ingredients of statistical inference: probability distributions, the likelihood and its applications, Bayes' theorem and Bayesian inference, Markov Chain Monte Carlo Methods, statistical significance, and goodness-of-fit. This course will also give you some hands-on experience with modern software packages for doing statistical modeling.

Logistics

- **Meeting time / place:** Fridays, 2pm–4pm, Room: AB 113 (AB 88 in the first week).
- **Instructor:** Jo Bovy, AB 229.
- **Email:** jo.bovy@utoronto.ca
- **Office hours:** Stop by my office or by appointment.
- **Course website:** `http://astro.utoronto.ca/~bovy/teaching.html`.

Learning objectives

After this course you should understand

- The three steps of probabilistic data analysis: setting up the model, fitting the model, evaluating and checking the model.
- Basics of discrete and continuous probabilities and the basic rules for working with conditional probabilities (Bayes' theorem).
- Properties of important basic probability distributions: Uniform, Gaussian, Poisson, χ^2, ...
- Basics of sampling random variables.
- Maximum likelihood inference (and its relation to "chi-square").
- The main ingredients of Bayesian inference: likelihood, prior, posterior, evidence and how to set them up or derive them.
- Monte Carlo methods for sampling distributions; the theory and practice of Markov Chain Monte Carlo (MCMC)
- How *emcee* works.

- Hamiltonian Monte Carlo methods for efficient MCMC: theory and practice.
- Non-parametric methods for determining uncertainties: bootstrap and jackknife.
- Goodness-of-fit and model selection methods ("chi-square per degree of freedom", AIC, BIC, cross-validation).

Reading

The following book contains most of what we will discuss:

- Ivezić, Connolly, VanderPlas, & Gray, *Statistics, Data Mining, and Machine Learning in Astronomy*, 2014, Princeton University Press. Errata can be found here.

Other useful books are:

- Gelman, Carlin, Stern, Dunson, Vehtari, & Donald Rubin, *Bayesian Data Analysis*, 3rd Edition, 2013, Chapman & Hall/CRC Texts in Statistical Science.
- MacKay, *Information Theory, Inference and Learning Algorithms*, 2003, Cambridge University Press.
- Sivia & Skilling, *Data Analysis: A Bayesian Tutorial*, 2011, Oxford University Press.

and a further list of readings is

- Hogg, Bovy, & Lang, *Data analysis recipes: Fitting a model to data*, 2010, `https://arxiv.org/abs/1008.4686`
- Hogg, *Data analysis recipes: Probability calculus for inference*, 2012, `https://arxiv.org/abs/1205.4446`
- Hogg & Foreman-Mackey, *Data analysis recipes: Using Markov Chain Monte Carlo* , 2017, `https://arxiv.org/abs/1710.06068`

Grading scheme

This course will be graded based on a single assigment and a presentation on a statistical topic in astrophysics during the last lecture.

- **Assignment:** 50 %
- **Presentation:** 50 %

Schedule

- **Week 1:** Introduction; rules of probability calculus and of conditional probability; important probability distributions; sampling random numbers.
- **Week 2:** The likelihood; maximum-likelihood estimation; Bayesian inference.
- **Week 3:** Monte Carlo sampling and Markov Chain Monte Carlo methods; non-parametric methods (bootstrap, jackknife).
- **Week 4:** Goodness-of-fit and model selection, cross-validation, statistical significance; intro to advanced topics.
- **Week 5:** Student presentations.

Academic integrity

From Appendix D of the Academic Integrity Handbook:

> Academic integrity is one of the cornerstones of the University of Toronto. It is critically important both to maintain our community which honours the values of honesty, trust, respect, fairness, and responsibility and to protect you, the students within this community, and the value of the degree towards which you are all working so diligently.
>
> According to Section B of the University of Toronto's Code of Behaviour on Academic Matter (`http://www.governingcouncil.utoronto.ca/policies/behaveac.htm`) which all students are expected to read and by which they are expected to abide, it is an offence for students to:
>
> - Use someone else's ideas or words in their own work without acknowledging explicitly that those ideas/words are not their own with a citation and quotation marks, i.e. to commit plagiarism.
> - Include false, misleading, or concocted citations in their work.
> - Obtain unauthorized assistance on any assignment.
> - Provide unauthorized assistance to another students. This includes showing another student your own work.
> - Submit their own work for credit in more than one course without the permission of the instructors.
>
> There are other offenses covered under the Code, but these are the most common. You are instructed to respect these rules and the values that they protect.