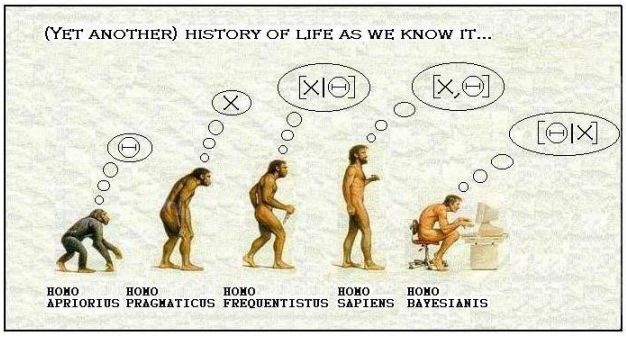
JOHN JAY COLLEGE OF CRIMINAL JUSTICE
The City University of New York
524 West 59th Street, New York, NY, 10019

Syllabus for: MAT/FOS 705, Applied Bayesian Data Analysis



credit: unknown, or maybe Søren Højsgaard 🤔...

Professor: Nicholas Petraco
Room: NA

Contact hours: Any time. Just email me at the address below.
Course Format: Asynchronous instruction with optional live Zoom tutorials and review sessions to be announced throughout the semester.
E-mail address: npetraco@gmail.com
Course website: https://npetraco.github.io/MATFOS705/

Course Description:
The classical "work horse" statistical methods learned in a first year undergraduate statistics course are known as "frequentist" and "Fisherian" methods. They are extremely useful and have been successfully applied for almost one hundred years. They have well known limits and flaws however. Complementary sets of statistical tools are known as "Bayesian" based methods. Related to Bayes' Theorem and relying on a definition of probability as "belief", they offer very intuitive interpretation and a naturally "coherent" methodological framework for inference. Within the last 20 years, as computing power has increased Bayesian methods have become practical for standard scientific applications.

The purpose of this course is to acquaint forensic science and forensic computing graduate students with Bayesian statistical tools that are applicable to data analysis problems they will encounter in their career. The course will also pay special attention to their relationship with classical methods, as well as issues, flaws and pitfalls to be aware of when applying Bayesian based methods.

This is an asynchronous instruction-based course. This means that the entire course is on-line and ***there is no officially scheduled or required live meeting time***. Course lecture materials will be posted on the following website:

https://npetraco.github.io/MATFOS705/

Announcements and important reminders will be emailed to you. As such **you must give me an email address that you check on a regular basis**. It will be our only official method of communication.

Most lectures posted to the course website will be accompanied by short videos which will be posted to YouTube. Links to the videos will also be sent to you in email when they are posted. Periodic live tutorial and review sessions over Zoom will be scheduled throughout the semester in email. They are not mandatory and will be recorded. After they have occurred, links to the recordings will be sent out to the class. Homework will be given though WebAssign. See below for details.

Course Learning Goals:

1. Recognize the importance of accuracy and objectivity in collecting/sampling data for applications to the law and policy.

2. Acquire an understanding of the types of data that can be recorded and analyzed (e.g. for quantitative trace, fire debris, toolmark, spectrochemical and digital evidence analysis).

3. Acquire an understanding of Bayesian statistical tools that can be used to analyze collected data.

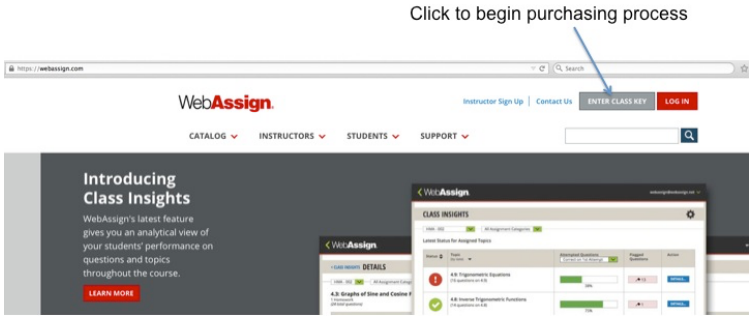
4. Understand the limitations of the Bayesian statistical methods used for data analytics and how not to misrepresent the capabilities of these methods to the courts or clients (ethics).

5. Obtain skill with the general programming/computing/statistical software **R** (http://www.r-project.org/), parametric Bayesian software **Stan** (http://mc-stan.org/).

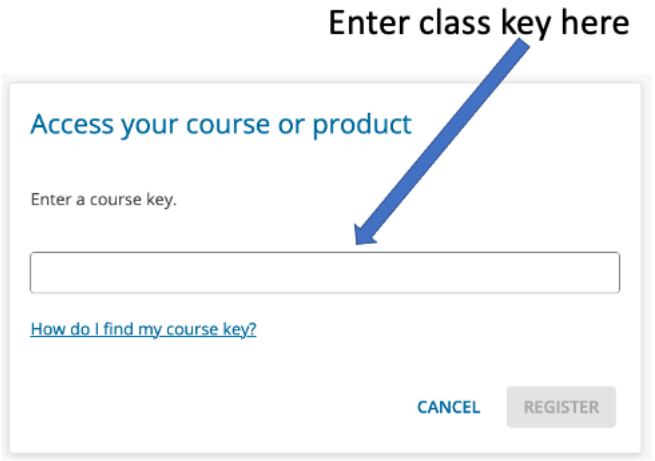
6. Develop oral and written communication skill as to how to present the results of sophisticated quantitative analysis to officers of the course and lay juries, in terms that are understandable to them.

Required Electronic Resources:

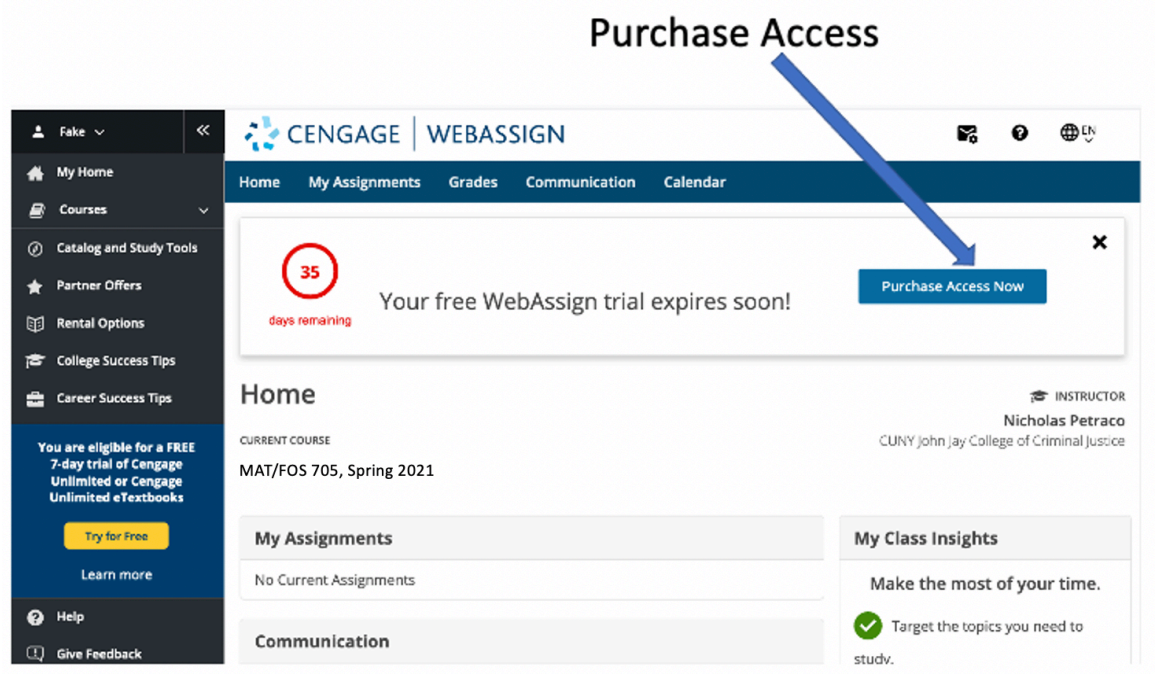
- **Webassign**: Can be purchased at:
    - https://webassign.com/
    - In order to purchase click on "**Enter Class Key**":



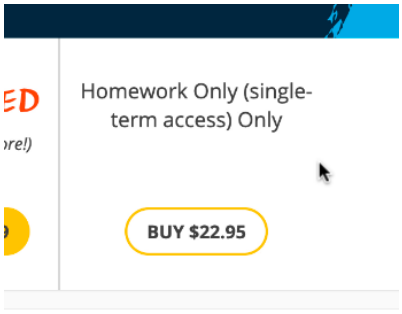
- The WebAssign website will probably prompt you to log into your account. If you don't have an account, create one.
- When you log in, you should see a place to enter the class key:
    - Class Key: **jjay.cuny 7307 3187**



- After logging in/creating-account, select your class and the website will prompt you to purchase the required materials:



- Purchase "MAT/FOS 705", which should be ~$22.95:



(Strongly Suggested) Textbooks:
Bayesian Data Analysis, 3rd Edition.
http://www.stat.columbia.edu/~gelman/book/
Electronic: http://www.stat.columbia.edu/~gelman/book/BDA3.pdf
ISBN-10: 1439840954
The hardcover new is about $60 if you want it.

A Student's Guide to Bayesian Statistics, 1st Edition.
ISBN-10: 1473916364
The hardcover new is about $46 if you want it.
Author's website: https://ben-lambert.com/
Author's YouTube Channel: https://www.youtube.com/user/SpartacanUsuals/featured
Author's Book YouTube Playlist:
https://www.youtube.com/playlist?list=PLwJRxp3blEvZ8AKMXOy0fc0cqT61GsKCG

Web resources:
Stan: https://mc-stan.org/
http://mc-stan.org/documentation/

Grading:

The grades for this course are based on homework (100%). We'll have about HW set one week.

Topics and Rough Schedule:

Week	Lecture Topics	Sub Topics	HW Due Dates
Feb 1-5	Introduction	Introduction to Bayesian Perspectives Review: Common Defininitions in Statistics Review: Basic Graphing	
Feb 8-12	Introduction and Tutorial for R		Feb 12: HW Set 1 Due
Feb 15-19	Review: Summary Statistics 1-3	Review: Measurs of Central Tendency Review: Standard Measures of Uncertainty Review: Range and Quantiles	Feb 22: HW Set 2 Due
Feb 22-26	Probability 1-3	Definitions Axioms and Some Useful Theorems Conditional Probability and Bayes' Theorem	Mar 1: HW Set 3 Due
March 1-5	Discrete Moments Important Distributions 1-2	Discrete Moments Important Discrete Distributions Important Continuous Distributions	Mar 8: HW Set 4 Due
March 8-12	Introduction and Tutorial for rstan		
March 15-19	Single Parameter Models 1-2	Intro to single parameter inference Inference by MCMC Sampling	Mar 22: HW Set 5 Due
March 22-26	Multiparameter Parameter Models 1-2	Multiparameter parameter inference Graphing Results, Diagnostics and Troubleshooting	Apr 5: HW Set 6 Due
Mar 29-Apr 2	Spring Break/Catch-up		
April 5-9	Jacobians 1-2	Definitions and Theory Exmples	Apr 12: HW Set 7 Due
April 12-16	Introduction to Regression Modeling 1-3	Basic Bayesian Regression Exmples	Apr 26: HW Set 8 Due
April 19-23	Model Checking	Basic Regression Model Checking	
April 26-30	Hierarchical Regression 1-3	Definitions and Theory Linear model example: 8 schools Advanced example: Radon	May 3: HW Set 9 Due
May 3-7	Hierarchical Regression 4-5	Generalized Linear Models: Logistic Model Example Generalized Linear Models: Poisson Model Example	May 10: HW Set 10 Due
May 10-14	Model Practical Assessment 1-2	Estimates of Predctive Error Examples computing WAIC and LOO estimates	May 17: HW Set 11 Due
May 17-21	Formal Model Comparison 1-3	Marginal Likelihood and Bayes Factors Bridge Sampling Examples	May 21: HW Set 12 Due

Keeping Up and Studying

This course is packed with information and our schedule will be tight. If you do R practice and a little home work every night you should be fine.

Academic Honesty:

Cheating and plagiarism will not be tolerated. If a student is suspected of either of these offenses he or she will be turned into the Provost. I encourage you to collaborate and help each other on projects, however the work you develop and present must be yours. Upon being found guilty a grade of F will be recorded for the course.