<u>Instructor</u>: Prof. Badredine Arfi
<u>Office</u>: 221 Anderson Hall   <u>Phone</u>: (352) 273 2357
<u>Office Hours</u>: Wed 1:00-4:00pm   <u>email</u>: barfi@ufl.edu

COURSE DESCRIPTION AND OBJECTIVES

Imagine that you have a data that you would like to analyze starting with a set of possible explanations. Perhaps, before getting the data you might have believed that some theories are plausible explanations. After you acquire the data you decide to examine whether any of those explanations fits the data. Your belief in the credibility of your explanations, given the data, is based on using Bayes' rule from probability theory: you update predicted probabilities of an event by using whatever new information comes your way. *This is the essence of doing Bayesian analysis*. And this is what we in fact do every day in reallocating credibility across various possibilities that face us when encountering situations that call for a decision to be made in favor of some of these possibilities – *Bayesian decision is in effect a manifestation of our daily intuitive behavior*. As a matter of fact, you have been doing Bayesian analysis without knowing it. For example, your intuitive interpretation of the usual p-values and confidence intervals are Bayesian, through and through. More generally, the parameters of any parametric statistical model can be estimated using Bayesian methods in a very intuitive way. In sum: two shifts of focus are made in explicitly doing Bayesian data analysis:

(1)   We go from a frequentist notion of probability (that is, as a property of the outside world) to a belief-based one (that is, as an observer's belief about observed uncertainty);

(2)   We go from point-value hypothesis testing to estimating parameter values and uncertainty by:

   a. Specifying a probability model with some prior knowledge about parameters

   b. Updating our knowledge about the parameters through a conditioning of this probability model on observed data

   c. Assessing the fit of the model to the data and checking the sensitivity of the conclusions to the starting assumptions

Wait: Bayesian analysis is however not a panacea! It does not automatically produce THE correct interpretations of the data. Yet Bayesian analysis is built on the observation that all

statistical models are subjective – *we always make decisions* about variable specifications, significance thresholds, functional forms, error distributions in a nonobjective way, etc.!

The purpose of this course is to introduce and train students to thinking in such a 'Bayesian intuitive' way when doing scientific data analysis. The course starts with the basic concepts of Bayesian analysis and incrementally goes into somewhat more advanced computational methods. Students are expected to acquire enough skills and understanding of Bayesian statistics such that by the end of the semester they will be able to:

1. Acquire a good understanding of Bayesian methods including Bayesian model specification, Bayesian posterior inference, and model assessment.

2. Use the acquired knowledge of Bayesian statistics to develop and estimate linear and non-linear Bayesian models as well as have enough exposure to MCMC (Markov Chain Monte Carlo) computation.

3. Deploy this knowledge in analyzing data in their respective research fields of interest.

The course will be using the R software (and stata; should anyone want to do that too). It is assumed that students have prior knowledge of basic statistics as taught in a first-year graduate course.

RECOMMENDED TEXTS

1. Jeff Gill. 2015. Bayesian Methods: A Social and Behavioral Sciences Approach. Third Edition. CRC Press. (On reserve at Library West).
2. Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, Donald B. Rubin. 2013. Bayesian Data Analysis. Third Edition. CRC Press. (On reserve at Library West in progress).
3. John K. Kruschke. 2015. Doing Bayesian Data Analysis. A Tutorial with R, JAGS, and Stan. Edition 2. Elsevier. (On reserve at Library West in progress).
4. James E. Morgan III. 2015. Political Analysis Using R. Springer International Publishing. (On reserve at Library West in progress).
5. Jeff Gill. 2006. Essential Mathematics for Political and Social Research. Cambridge University Press. (On reserve at Library West).
6. R-Software materials freely available online (suggestions will be posted on canvas).
7. John Thompson. 2014. Bayesian Analysis with Stata. Stata Press. (On reserve at Library West).
8. Additional materials will be made available o students.

REQUIREMENTS AND ASSESSMENT

The requirement for this course is simple (as always!): work diligently and persistently. This includes attending classes, doing the readings carefully before the seminar meets, and working regularly on the computer applications, occasional extra readings, and the research

paper. Each student should expect to be spending many hours learning how to excel in using the R software and use it to estimate models discussed in class.

There will be a number of homework assignments that the students must complete and upload to canvas. The homework assignments are due on the specified dates; no late submission is acceptable (except with a valid excuse).

A major component of the course evaluation will be a term research paper. Each student will produce a manuscript of high quality using an appropriate modelling strategy.

DISTRIBUTION OF GRADES

1. **35%: Weekly homework exercises**. All assignments are to be uploaded to canvas before the beginning of class on their respective due dates. No late submission will be accepted for any reason (except when justified with university sanctioned documentation). The problem sets will be assigned at the end of the lectures depending on what we cover in the lecture sessions.

2. **15%: Each student will be assigned "presentations**" for the practice session of the course which will consist in presenting the weekly assigned homework (this will be fully explained on the first day of class). A schedule of these presentations will be created on the first day of class.

3. **40%: Research Paper**
   a. Each student is required in consultation with the instructor (see down below) to replicate a paper (published in the last 5 years or so) using Bayesian analysis, or
   b. The goal is to produce a high-quality, potentially publishable manuscript, using a Bayesian model (or models) discussed in the course.

4. **10%: Research Paper Presentation**: Each student will present his/her paper on the last day of classe**s** of the semester. The presentation will consist of a ppt presentation for about ten minutes followed by five minutes of Q & A.

COMPUTER SOFTWARE

All models covered in this class will be estimated using the R software package, the reason being that most published Bayesian works are done using in R. The software R is freely available online at https://cran.r-project.org . Students are thus required to install the R software on their laptops as well as an editor for R; many scholars use RStudio (free open source edition) available at https://rstudio.com/products/rstudio/download/#download (this will be explained on the first day of classes). Any student who is interested can, in addition, learn how to use stata to do Bayesian analysis (this will not be covered in class; a recommended book is suggested for that purpose). Stata software is available through UFapps for all UF students.

<div align="center">SPECIFICS ON THE RESEARCH PAPER</div>

In order for the instructor to provide guidance in the preparation of the paper, you will be required to turn in various brief intermediate papers throughout the semester.

Each student must:
1. Find a published paper that interests you and that applies a Bayesian statistical method comparable to the material covered in this course. **Date: January 31**

2. Obtain the data from a database or the author. **Date: February 14**

3. Replicate the published results as nearly as possible. **Date: March 13**

4. You **must extend** the analysis in some way. You could, for example:
   - Suggest a more appropriate functional form for the estimation and re-estimate.
   - Argue that one or a set of important variables were omitted and conduct the analysis anew.
   - Argue that the results are likely to be sensitive to sample selection or variable measurement etc. and then conduct appropriate analyses to address that possibility.
   - Extend the data or use a different data set to test the theory.
   - Any other good idea that you might have.

   The final paper should be 15-20 pages long, including the bibliography.
   Date: April 17.

Final Submission of the Paper: Students are required to submit to canvas (in addition to the paper) an annotated log life displaying their complete work, an annotated file listing all the commands that one would use to replicate the results of the paper, and the final data file used for the paper (including both replication and extension works)

<div align="center">IMPORTANT DATES</div>

Classes Begin (for this course)	**Fri, January 10**
Holidays *No classes*	**January 20: Martin Luther King, Jr.** **February 29 – March 7: Spring Break**
Last Class (for this course)	**Wed, April 17**
ISA Annual Convention – Honolulu (no class)	**March 27**

Session	TOPIC
1	• Background and Introduction
2	• Specifying Bayesian Models
3	• The Normal and Student's-t Models
4	• The Bayesian Prior
5	• The Bayesian Linear Model
6	• Assessing Model Quality
7	• Bayesian Hypothesis Testing and the Bayes Factor
8	• Basics of Markov Chain Monte Carlo
9	• Implementing Bayesian Models with Markov Chain Monte Carlo
10	• Bayesian Hierarchical Models
11	• Utilitarian Markov Chain Monte Carlo (MCMC)
12	• Students Presentations

Most lectures mostly from Jeff Gill's Bayesian Methods and supplemented with the other recommended books.

IMPORTANT NOTES

- The instructor reserves the right to change or amend any part or aspect of this document should a need for doing so emerge at any point in time during the semester.

- **Academic Honesty**
  UF students are bound by The Honor Pledge which states, "We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment." The Honor Code (sccr.dso.ufl.edu/process/student-conduct-code/) specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor or TAs in this class.

- **Class Attendance, Make-up Exams, Assignments**
  Requirements for class attendance and make-up exams, assignments, and other work in this course are consistent with university policies that can be found at: catalog.ufl.edu/UGRD/academic-regulations/attendance-policies/

- **UF Grading Policies**
  Information on current UF grading policies for assigning grade points. This may be achieved by including a link to the web page: catalog.ufl.edu/UGRD/academic-regulations/grades-grading- policies/.

- **Online Course Evaluation Process**
  Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at gatorevals.aa.ufl.edu/students/. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via ufl.bluera.com/ufl/. Summaries of course evaluation results are available to students at gatorevals.aa.ufl.edu/public-results/

- **Services for Students with Disabilities**
  Students with disabilities who experience learning barriers and would like to request academic accommodations should connect with the disability Resource Center by visiting https://disability.ufl.edu/students/get-started/. It is important for students to share their accommodation letter with their instructor and discuss their access needs, as early as possible in the semester.

- **Campus Helping Resources**

Health and Wellness

- U Matter, We Care: If you or someone you know is in distress, please contact umatter@ufl.edu, 352-392-1575, or visit umatter.ufl.edu/ to refer or report a concern and a team member will reach out to the student in distress.
- Counseling and Wellness Center: Visit counseling.ufl.edu/ or call 352-392-1575 for information on crisis services as well as non-crisis services.
- Student Health Care Center: Call 352-392-1161 for 24/7 information to help you find the care you need, or visit shcc.ufl.edu/.
- University Police Department: Visit police.ufl.edu/ or call 352-392-1111 (or 9-1-1 for emergencies).
- UF Health Shands Emergency Room / Trauma Center: For immediate medical care call 352-733-0111 or go to the emergency room at 1515 SW Archer Road, Gainesville, FL 32608; ufhealth.org/emergency-room-trauma-center.

Academic Resources

- E-learning technical support: Contact the UF Computing Help Desk at 352-392-4357 or via e-mail at helpdesk@ufl.edu.
- Career Connections Center: Reitz Union Suite 1300, 352-392-1601. Career assistance and counseling services career.ufl.edu/.
- Library Support: cms.uflib.ufl.edu/ask various ways to receive assistance with respect to using the libraries or finding resources.
- Teaching Center: Broward Hall, 352-392-2010 or to make an appointment 352-392-6420.
- General study skills and tutoring. teachingcenter.ufl.edu/
- Writing Studio: 2215 Turlington Hall, 352-846-1138. Help brainstorming, formatting, and writing papers. writing.ufl.edu/writing-studio/
- Student Complaints On-Campus: sccr.dso.ufl.edu/policies/student-honor- code-student-conduct-code/
- On-Line Students Complaints: distance.ufl.edu/student-complaint-process/