Statistics GU4221 & GR5221
Time Series Analysis
Spring 2023

Instructor

- Gabriel Young (PhD)
  Email: gjy2107@columbia.edu
  Office: Room 614, 6th floor of Watson Hall, 612 West 115th Street

Teaching Assistants

- Reed Palmer
  Email: wrp2110@columbia.edu

- Diane Lu
  Email: dl3213@columbia.edu

Office Hours

- Gabriel Young: **TBA**
- Reed Palmer: **TBA**
- Diane Lu: **TBA**

Course Description

In numerous fields of study related to statistics, data is often recored over time and hence, exhibits a temporal structure more complex than "*independent and identically distributed*." A few examples follow:

- The closing price of a financial instrument (P_t) heavily depends on the previous observation (P_{t-1}), where t could be any time index such as: month, day, minute, ...etc.
- Windspeed (W_t) recorded on a minute basis. Clearly the windspeed (W_t) at minute t heavily depends on the previous case (W_{t-1}).
- In classical design of experiments, respondents are often recored over time.
- In the field of machine learning, speech recognition falls within the scope of time series analysis.

Time series analysis (GU4221/GR5221), extends the traditional linear regression setting to include important aspects of temporal dependence. By the end of the semester, students should understand fundamental theoretical aspects of time series models. Further, students should also have the ability to analyze a non-trivial time series dataset. Topics include: Least squares smoothing and prediction, trend estimation, seasonality, autocorrelation/autocovariance function, stationarity, fundamental time series models (AR, MA, ARMA, ARIMA, GARCH), forecasting, Fourier series, spectral density and spectral analysis. Some statistical machine learning topics are also included: Long short-term memory (LTSM) neural network and the Kahlam filter (time permitting). Emphasis on theory, applications and computing are all integral components of the course.

Lecture

- **Days:** Tuesday and Thursday 6:10pm-7:25pm
- **Location:** Location: Mathematics Building, Room 207
- **Class Structure:** The class follows a traditional lecture environment, i.e., the professor will handwrite notes for each lecture. In-class examples and other relevant course materials will be posted on Canvas.
- **Attendance:** Students should ideally attend each class meeting. Even though attendance is not required, I will frequently give examples or hints during lecture that may show up on assessments. **Attendance is required on exam days.**

Prerequisites

- A previous course in statistics, elementary probability (GU4203/GR5203), multivariate calculus, linear algebra and linear regression models (GU4205/GR5205).

Reading Material

- Peter J. Brockwell, Richard A. Davis. *Introduction to Time Series and Forecasting, 3rd Edition.* **Note that Richard Davis is a current professor in our department!**
- Maybe more references...

Software

`R` will be used in lecture but students can use any computing language. `ITSM` is used in the recommended text but is not used in lecture or HW.

Homework

GU4221/GR5221 homework will be a combination of theory and application. Students should expect homework 4-5 assignments given during the semester. HW submission and its required format will be discussed during lecture.

Midterm

The midterm exam is 30% of the final course grade. You will have **TBA** minutes to complete the midterm and allowable materials will be discussed one week before the exam. The midterm exam is scheduled for

- 03/09/2023

Final

The final exam is 35% of the final course grade. You will have **TBA** minutes to complete the final and allowable materials will be discussed one week before the exam. You must take the final at the scheduled time. The final exam is tentatively scheduled for

- **TBA** (scheduled by the office of the registrar)

Project

Details of the GU4221/GR5221 will be discussed closer to the midterm.

Exam Absences

Make-up exams will not be given routinely. If you have a legitimate conflict with an exam date, it is incumbent upon you to make arrangements with the instructor. An exam missed due to a documented illness or other unforeseeable (and documented) extraordinary circumstances must be made up before the test papers are returned to the class.

Grading and Academic Integrity

I take the honor code very seriously; students caught cheating or otherwise in violation will face disciplinary action. Please note the Barnard honor code text:

"We... resolve to uphold the honor of the College by refraining from every form of dishonesty in our academic life. We consider it dishonest to ask for, give, or receive help in examinations or quizzes, to use any papers or books not authorized by the instructor in examinations, or to present oral work or written work which is not entirely our own, unless otherwise approved by the instructor.... We pledge to do all that is in our power to create a spirit of honesty and honor for its own sake."

For more information please see:

- `http://barnard.edu/node/2875`
- `https://www.college.columbia.edu/academics/academicintegrity`

Homework	30%
Project	10%
Midterm	25%
Final	35%

Grading

Grading Scale

93 or more	A
90 to 92	A-
87 to 89	B+
83 to 86	B
80 to 82	B-
77 to 79	C+
70 to 76	C
60 to 69	D
Below 59	F

- Final grades are subject to a curve. The above scale will guarantee the displayed letter grades.
- The letter grade $A+$ will be given to a handful of **top** students in the class (assuming they earned a final grade of at least 93%).

Approximate Lecture Outline

The GU4221/GR5221 course outline is displayed on the next page. This outline is projected and we will likely deviate from the displayed timeframe. Hence I need to heavily emphasize that this schedule is **tentative**. The topics with the symbol **(*)** next to them are time permitting. The midterm and spring recess will not change from the displayed schedule.

Date	Content	Chapter
01/17/2023	Introduction to time series analysis and the basic AR(1) model	CH1
01/19/2023	Stochastic process, autocorrelation/autocovariance function (ACF/ACFV), general approach to time series analysis	CH1 CH1
01/24/2023	Stationarity, AR(1), random walk , MA(1)	CH1
01/26/2023	Seasonality and trend, differencing, backward shift operator	CH1
01/31/2023	Stationary models: linear process, ARMA model	CH2
02/02/2023	Stationary models: central limit theorem, long-run variance	CH2
02/07/2023	Stationary models: prediction operator	CH2
02/09/2023	Stationary models: Durbin-Levinson and innovations algorithms	CH2
02/14/2023	ARMA model: formal definition	CH3
02/16/2023	ARMA model: autocovariance function (ACFV), ARMA model: partial auto covariance function (PACF)	CH3 CH3
02/21/2023	ARMA model: forecasting	CH3
02/23/2023	ARMA modeling and forecasting	CH5
02/28/2023	ARMA modeling and forecasting: Yule-Walker estimation	CH5
03/02/2023	ARMA modeling and forecasting: innovations algorithm	CH5
03/07/2023	ARMA modeling and forecasting: MLE	CH5
03/09/2023	**Midterm**	
03/14/2023 03/16/2023	**Spring recess** **Spring recess**	
03/21/2023	ARMA modeling and forecasting: MLE continued	CH5
03/23/2023	ARMA modeling and forecasting: model order selection	CH5
03/28/2023	Non-stationary and seasonal time series models: ARIMA model	CH6
03/30/2023	Non-stationary and seasonal time series models: unit root tests	CH6
04/04/2023	ARCH and GARCH time series models	CH7
04/06/2023	Multivariate time series **(*)**	CH8
04/11/2023	Review of neural networks **(*)**	Notes
04/13/2023	Long short-term memory neural networks (LTSM) **(*)**	Notes
04/18/2023	Spectral analysis: spectral density **(*)**	CH4
04/20/2023	Spectral analysis: linear combination of sinusoids, Fourier series **(*)**	CH4
04/25/2023	Spectral analysis: discrete Fourier transform **(*)**	CH4
04/27/2023	Catch-up and final review **(*)**	
05/02/2023	**Study week**	
05/04/2023	**Study week**	
05/09/2023 05/11/2023	**Finals week** **Finals week**	