

1 Lecture 1:Jan 26

Today

- Introduction
- Introduce yourself
- Course logistics

What is this course about?

The term “linear models” describes a wide class of methods for the statistical analysis of multivariate data. The underlying theory is grounded in linear algebra and multivariate statistics, but applications range from biological research to public policy. The objective of this course is to provide a solid introduction to both the theory and practice of linear models, combining mathematical concepts with realistic examples.

Prerequisite

- **Must:** Introduction to Probability
- **Good to have:** Mathematical Statistics, Scientific Computation II

A hierarchy of linear models

- The linear mean model:

$$\underset{n \times 1}{\mathbf{y}} = \underset{n \times p}{\mathbf{X}} \underset{p \times 1}{\boldsymbol{\beta}} + \underset{n \times 1}{\boldsymbol{\epsilon}}$$

where $\mathbf{E}(\boldsymbol{\epsilon}) = \mathbf{0}$. Only assumption is that errors have mean 0.

- Gauss-Markov model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\mathbf{E}(\boldsymbol{\epsilon}) = \mathbf{0}$ and $\mathbf{Var}(\boldsymbol{\epsilon}) = \sigma^2 \mathbf{I}$. Uncorrelated errors with constant variance.

- Aitken model or general linear model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\mathbf{E}(\boldsymbol{\epsilon}) = \mathbf{0}$ and $\mathbf{Var}(\boldsymbol{\epsilon}) = \sigma^2 \mathbf{V}$. \mathbf{V} is fixed and known.

- Variance components models: $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma_1^2 \mathbf{V}_1 + \sigma_2^2 \mathbf{V}_2 + \dots + \sigma_r^2 \mathbf{V}_r)$ with $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_r$ known.
- General mixed linear Model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\mathbf{E}(\boldsymbol{\epsilon}) = \mathbf{0}$ and $\mathbf{Var}(\boldsymbol{\epsilon}) = \boldsymbol{\Sigma}(\theta)$.

- Generalized linear models (GLMs). Logistic regression, probit regression, log-linear model (Poisson regression), ... Note the difference from the general linear model. GLMs are generalization of the *concept* of linear models. They are covered in Math 7360 - Data Analysis class (<https://tulane-math7360.github.io/lectures/>).

Syllabus

Check course website frequently for updates and announcements.

<https://tulane-math-7260-2022.github.io/>

HW submission

Through Github with demo on Friday class.

Presentations

Let me know your pick by the end of Friday (01/28/2022).

Last year comments

1. Experience in this course

- Overall, I had a pretty good experience in this course. It moved quickly, but that is expected from this level of course. Sometimes it was hard to stay engaged with the lectures and to really absorb the course material. Because the lectures moved so fast, I really appreciated how Professor made the full notes available at the time of the lecture. I would have liked if there were a few more examples with the notes, as sometimes the homework felt disjoint from the notes.

Response: I will try to move slower this semester. I will start lab sessions earlier too.

- The professor is an extremely intelligent, kind, and understanding professor. He prioritizes in making sure that we understand the material and seeing how the material can be applied. His lecture notes were a godsend because the texts could be a bit ambiguous at times but he elucidated the material in such a comprehensible manner.

Response: I will try to fix the left-over typos.

- Mentioned in class from other students/internal evaluation, conveying the mathematical concepts through the presentation is not a good idea to follow the class in real-time. Prepared presentation can give rise to a distraction on what we have been going over.

Response: I am still delivering this class in hybrid-mode. I found the presentations fit online teaching better. I think the difficulty might be caused by (1) fast moving lecture (2) I only realized the need of reviewing basic concepts of probability almost a quarter into the semester...

- I found the setup of the course not very engaging. Additionally, many of the class notes came directly from the additional sources with no additional information or explanation, which I found to be not very helpful.

Response: I actually like them. I was the guinea pig to test them.

- Easily help us to understand the main course, and the notes and details are great.

Response: There will be notes.

- Moves very quickly and can be hard to keep up with. Sometimes instructions are unclear.

Response: I will try to slow down.

- Both the instructor and the TA were helpful. It was hard to follow along in class though.

Response: We don't have TA this time. Make use of the office hour. And I have to say, it needs effort to ace in this class.

2. Strong aspects of this course

- Having the lecture notes and labs available was very helpful. Professor was also always very nice and accommodating, and willing to meet with me when I needed help. He also always responded to student feedback, if we asked for an extra day or two on the homework or something like that.

Response: Here is an example of correctly using the office hours.

- His lecture notes and the lab sessions.

Response: They will be there again.

- Lab session is necessarily required to this class. A lot of computations in the class would be done by computer due to the complexity, and students are expected to handle with the computer programming properly at a desired level. The course can be an introduction to the statistical computation, which does not exist in the mathematics department.

Response: Hmm, there is a course Math 7360 Data Analysis that focuses more on the computational side.

- I appreciated the homework reviews in class and felt these helped clarify the material.

Response: Of course, the reviews will be there again. The purpose of the course is for you to learn.

- Grading was easy which made up for the rigor.

Response: Don't rely on this...

- Really appreciate that Professor Xiang made such a neat and tidy notes for us. It is really helpful for me to review. And notes have a great interaction with us, Professor Xiang also leaves some questions to help us think about the logic behind.

Response: Well, Xiang is my first name. Please call me Prof. X.

- Prof. Xiang was highly organized and wanted his students to understand the course content more than he made them worry about grades. I learned a lot about Linear Models and feel confident applying the course content professionally and academically. I wish most of the Math department had his teaching style and implemented his course documents and organization structure. Prof. Xiang made the course content in class digestible and if I needed to review the material I could easily find it through his course notes and textbook. I wish I could say the same about my other courses.

Response: Hmm, I like Prof. X. better.

- I really appreciated the emphasis on learning. It allowed for most students to take it at the pace that was good for them.

Response: Please don't let your score rely on this comment.

3. There will be an internal mid-term-ish evaluation for this course. Will remember to go over them.