Spring 2021

BSTA 651
Linear and Generalized Linear models

Syllabus for First Half of the Course (Linear Models)

Time: Tu, Th: 1:30 PM- 2:50 PM   from Jan 21, 2021 – March 9, 2021
Location: Bluejeans

Instructors: Justine Shults, Ph.D (Linear Models) and Yong Chen, Ph.D (Generalized Linear Models)

Contact information:
Justine Shults, Ph.D,
Office: Blockley 626 (and Roberts 4471)
Phone: 215-518-8628
Email: shultsj@chop.edu
Office Hours: Fridays from 2-3 pm on Bluejeans

Teaching Assistant:
Office: Blockley 108/109
Email: Sarah.Weinstein@pennmedicine.upenn.edu
Office Hours: Wednesday 4:00 – 5:00 pm

Text-book: Linear Models in Statistics, 2nd Edition by Rencher and Schaalje   (Wiley ISBN: 978-0-471-75498-5)

Additional references: **1.** Linear Models by Searle (Wiley ISBN 0-471-18499-3)
                                  2. Generalized Linear Models, Second Edition by McCullagh and Nelder (Chapman & Hall/CRC Press ISBN-13: 978-0412317606)

Note: You do not need to purchase the references. A good text on matrix algebra (e.g. by Harville, ISBN 978-0-387-22677-4) will also be useful to have for future reference.

Prerequisites: BSTA 620 and BSTA 630. Both BSTA 621 and BSTA 631 may be taken concurrently, with permission of the instructor.

This course provides an introduction to the theory (primarily) and application of linear and nonlinear models. Topics covered in this course include: (1) multiple linear regression models; (2) analysis of variance models; and (3) generalized linear models (second half of the course).

Weekly (approximately) homework assignments will be given. Please work on these assignments independently, unless indicated otherwise.

In addition, there will be 1 examination for the first half of the course. Your final grades will be based on your work in both the homework problems and the two exams according to the following distribution:

(1) HW (40%); (2) Part I exam (25%); (3) Final (30%); (4) Class-Participation (5%).

Note regarding software: The emphasis of this course will be on the theory of linear models, but some applications will also be presented. Please be sure that your results for any analysis are clearly summarized, and any supporting code is included.

Spring Break:

No classes will be held on March 10 or March 11.

Classes and Topics for Part 1 (Linear Models) [This will be updated slightly as the course progresses]

Activity	Date	Reading	Topic	Homework: Assigned	Due
Lecture 1	Th: 1/21/2021	R: Ch 1-2	Introduction and Review of Matrix Algebra	HW1	
Lecture 2	Tu: 1/26/2021	R: Ch 6-7	Linear Regression: simple and multivariable		
Lecture 3	Th:1/28/2021	R: Ch 2,4 and 7.6	Multivariate Normal Distr. & Maximum likelihood Regression		
Lecture 4	Tu: 2/2/2021	R: Ch 5	Centered Model in Linear Models and Distributions of Quadratic Forms	HW2	HW1
Lecture 5	Th: 2/4/2021	R: Ch 8	Distribution of Quadratic Forms		
Lecture 6	Tu: 2/9/2021	R: Ch 8	Testing in Multivariable Linear Regression Models	HW3	HW2
Lecture 7	Th: 2/11/2021	R: Ch 12	Testing in Multivariable Full Rank Linear Regression Models		
Lecture 8	Tu: 2/16/2021	R: Ch 13	One way analysis of variance model- balanced case	HW4	HW3
Lecture 9	Th: 2/18/2021	R: Ch 14	Two way analysis of variance model- balanced case		
Lecture 10	Tu: 2/23/2021	R: Ch 15	Analysis of variance – cell means model for unbalanced data	HW5	HW4
Lecture 11	Th: 2/25/2021	R: Ch 16	Analysis of covariance		
Lecture 12	Tu: 3/2/2021	R: Ch 7	Generalized least squares		HW 5
Lecture 13	Th: 3/4/2021		Misc. and Review		
Midterm	Tu: 3/9/2021		Closed Book "in class" Midterm		
Spring Break: 3/10 – 3/11					

Spring 2021

BSTA 651
Linear and Generalized Linear models

Syllabus for <u>Second Half</u> of the Course (<u>Generalized Linear Models</u>)

Time: Tu, Th: 1:30 PM- 2:50 PM   from March 16, 2021 – April 27, 2021
Location: Zoom (<u>link will be distributed to the class</u>)

Instructors: Justine Shults, Ph.D (Linear Models) and Yong Chen, Ph.D (Generalized Linear Models)

Contact information:
Yong Chen, Ph.D,
Office: Blockley 602
Phone: 215-746-8155
Email: ychen123@upenn.edu
Office Hours: by appointment

Teaching Assistant:
Email: Sarah.Weinstein@pennmedicine.upenn.edu
Office Hours: Mondays 3:30-4:30pm

Text-book: <u>Generalized Linear Models, Second Edition</u> by McCullagh and Nelder (Chapman & Hall/CRC Press ISBN-13: 978-0412317606)

Prerequisites: BSTA 620 and BSTA 630. Both BSTA 621 and BSTA 631 may be taken concurrently, with permission of the instructor.

The <u>second half of this course</u> provides an introduction to the theory of generalized linear models from the perspective of estimating functions. Topics covered in this course include: (1) generalized linear models: components and formulation; (2) quasi-likelihoods and estimating functions; and (3) marginal models and conditional models.

Weekly (approximately) homework assignments will be given. Please work on these assignments independently, unless indicated otherwise.

In addition, there will be 1 take-home examination at the end of the course. Your final grades will be based on your work in both the homework problems and the two exams according to the following distribution:

(1) HW (40%); (2) Part I exam (25%); (3) Final (30%); (4) Class-Participation (5%).

Note regarding software: The emphasis of this course will be on the theory of generalized linear models. The students are expected to learn and get familiar with software (e.g., R packages/functions 'glm', 'gee', 'geepack') by themselves.

Classes and Topics for Part 2 (Generalized Linear Models) [This will be updated slightly as the course progresses]

Activity	Date	Reading: McCullagh & Nelder	Topic	Homework: Assigned	Due
Lecture 1	Tu: 3/16/2021	Appendix A & C; Ch 2	Review of MLE theory; GLM: a unified framework	HW1	
Lecture 2	Th: 3/18/2021	Ch 2	GLM: three key components		
Lecture 3	Tu: 3/23/2021	Ch 2	Canonical links		
Lecture 4	Th: 3/25/2021	Ch 9	Quasi likelihoods	HW2	HW1
Lecture 5	Tu: 3/30/2021	Ch 8	Review of properties of GLM		
Lecture 6	Th: 4/1/2021	Ch 8	Algorithms: basics	HW3	HW2
Lecture 7	Tu: 4/6/2021	Ch 4	GLM: binary outcomes		
Lecture 8	Th: 4/8/2021	Ch 13	Conditional logistic regression; logistic regression under retrospective sampling	HW4	HW3
Lecture 9	Tu: 4/13/2021	Ch 5 and 6	Stratified data; polytomous logistic regression; Poisson regression		
Lecture 10	Th: 4/15/2021	References will be distributed	Marginal model: Generalized estimating equations		HW4
Lecture 11	Tu: 4/20/2021	References will be distributed	Conditional models: Generalized linear mixed effect models (GLMM) and transition models		
Lecture 12	Th: 4/22/2021	References will be distributed	Advanced topic 1: TBD	Take-home final will be distributed (due on 11:59 pm 04/30/2021)	
Lecture 13	Tu: 4/27/2021	References will be distributed	Advanced topic 2: TBD		
Final			Take-home final	4/22/2021	11:59 pm 04/30/ 2021