جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

Course Syllabus: Linear Models - STAT 230

Division	Computer, Electrical and Mathematical Sciences & Engineering
Course Number	STAT 230
Course Title	Linear Models
Academic Semester	Fall
Academic Year	2019/2020
Semester Start Date	08/25/2019
Semester End Date	12/10/2019
Class Schedule (Days & Time)	10:30 AM - 12:00 PM | Mon Thu

Instructor(s)

Name	Email	Phone	Office Location	Office Hours
Hernando Catequista Ombao	hernando.ombao@kaust.edu.sa	+966128080213	4126, 1, Al-Khawarizmi (bldg. 1)	Mondays, 1:00 pm - 2:00 pm Tuesdays, 11:00 am – 12:00 noon By Appointment

Teaching Assistant(s)

Name	Email
TBA	TBA

Course Information

Comprehensive Course Description	**Objective:** The course will cover theoretical and methodological aspects of linear models and its generalization: estimation and inference for multiple linear regression, analysis of variance, longitudinal and mixed models. By the end of the course the students are expected to know (a.) how to formulate models for dat analytic problem at hand; (b.) develop statistical inference (design tests of hypotheses and confidence intervals) on functionals of the model parameters; (c.) interpret results from the analysis of data. The tentative topics are the following: (1.) Linear models in matrix formulation. Estimation of the mean and variance. (2.) Specific linear model settings: simple and multiple linear regression, multiway analysis of variance, analysis of covariance. (3.) Basic time series models. Models for longitudinal data: estimating the mean function and the covariance structure. (4.) Penalized regression, principal components analysis. (5.) Nonparametric regression. (6.) Introduction to generalized linear models.
Course Description from Program Guide	This course is an introduction to the formulation and use of the general linear model, including parameter estimation, inference and the use of such models in a variety of settings. Emphasis will be split between understanding the theoretical formulation of the models and the ability to apply the models to answer scientific questions.Multivariate models; Inference about independence.

Goals and Objectives	The overall goal of this course is to master the theory of linear models (and their various generalizations covered during the course), their formulation, their estimation and inference, as well as their application to real datasets. In addition, the students will be able to perform simple data analyses in practice using the statistical software R, and be able to understand and correctly interpret the results and outputs (diagnostics, figures, tables, etc.) from R. All the material covered during the course (lectures and homework) will be mastered by the students.
Required Knowledge	Advanced and multivariate calculus, Linear algebra, Probability and Statistics.
Reference Texts	1. Christensen (2011) Plane Answers to Complex Questions: the Theory of Linear Models,Springer; ebook available 2. Wood (2015) Core Statistics, Cambridge University Press; e-book available 3. Seber and Lee (2003) Linear Regression Analysis, Wiley; e-book available 4. Hocking (1996) Methods and Applications of Linear Models: Regression and the Anal- ysis of Variance, Wiley 5. McCullagh and Nelder (1989) Generalized Linear Models, Chapman & Hall/CRC 6. Kariya and Kurata (2004) Generalized Least Squares, Wiley; e-book available 7. Hastie and Tibshirani (1990) Generalized Additive Models, Chapman & Hall/CRC 8. Davison (2003) Statistical Models, Cambridge University Press; e-book available 9. Faraway (2005) Linear Models with R, Chapman & Hall/CRC 10. Faraway (2006) Extending the Linear Model with R, Chapman & Hall/CRC
Method of evaluation	**40.00**% - Final exam **20.00**% - Quiz(zes) **20.00**% - Midterm exam **20.00**% - Homework /Assignments
Nature of the assignments	All graded work are to be done individually. There will be 4 homework problem sets; 4 in-class quizzes; 1 midterm exam to be scheduled outside of the regular class hours; 1 comprehensive final exam to be scheduled by the Registrar's office.
Course Policies	Assignments will be collected at the START of the class on the due date. Late assignments will not be accepted, unless prior arrangements have been made. Staple the pages together, submit the problems in order and write clearly (**illegible handwriting may not be graded**). Students are encouraged to work together on the homeworks, but collaboration with classmates is **strictly limited to discussing problems**, not writing them up, photocopying solutions, or sharing R code.
Additional Information	Class notices and course related information will be posted periodically on the STAT 230 dropbox folder. Please check regularly for important information. Also, there may be important email communications (like homework hints or a change in the exams date), so it is important to monitor your email account regularly.

Tentative Course Schedule *(Time, topic/emphasis & resources)*		
Week	**Lectures**	**Topic**
1	Mon 08/26/2019 Thu 08/29/2019	Formulating the linear model under specific contexts: simple linear regression and multi-group settings. Vector spaces and projections. Estimation of the mean and variance.
2	Mon 09/02/2019 Thu 09/05/2019	Distributions: Gaussian, chi-squared, student's-t. Least squares estimators: mean, variance, distributions.
3	Mon 09/09/2019 Thu 09/12/2019	Maximum likelihood estimation and inference.
4	Mon 09/16/2019 Thu 09/19/2019	Tests of significance. Nested models. F-distribution.
5	Mon 09/23/2019 Thu 09/26/2019	Sunday 23rd: Saudi National Day One-Way Analysis of Variance. Multiple comparisons.
6	Mon 09/30/2019 Thu 10/03/2019	Multi-way analysis of variance.
7	Mon 10/07/2019 Thu 10/10/2019	Analysis of covariance.
8	Mon 10/14/2019 Thu 10/17/2019	Random effects models. October 14. Midterm 1.
9	Mon 10/21/2019 Thu 10/24/2019	Mixed models. Basic time series models.
10	Mon 10/28/2019 Thu 10/31/2019	Sunday 28th: Mid-semester break Longitudinal data analysis (single curve). Weighted least squares.
11	Mon 11/04/2019 Thu 11/07/2019	Longitudinal data analysis (multiple curves). Model/variable selection.
12	Mon 11/11/2019 Thu 11/14/2019	Penalized regression. Principal components analysis.
13	Mon 11/18/2019 Thu 11/21/2019	Generalized linear models. Logistic regression analysis.
14	Mon 11/25/2019 Thu 11/28/2019	Nonparametric regression: kernel methods, basis approach.
15	Mon 12/02/2019 Thu 12/05/2019	Generalized additive models. Survival models.
16	Mon 12/09/2019	Exams

Note

The instructor reserves the right to make changes to this syllabus as necessary.