

**Goals and Organization**

(Tue Sep 30, 2014 09:00 AM)

**PLS 206 Applied Multivariate Modeling in Agricultural and Environmental Sciences  
FALL 2014 4.0 Units****Goals**

To give graduate students in biology and related sciences an overview of regression analysis and multivariate statistical modeling.

The emphasis is on covering a wide variety of techniques and preparing students for further individual study according to their specific needs. Linear models are presented in detail. Students will learn R to complete assignments. My philosophy is to promote discovery and resourcefulness. For example, instead of showing ten different ways to do separation of means, we will: (1) try to understand why separation of means is an issue, (2) learn one reasonable way to do it, (3) make a plan for learning other methods if necessary. Students are expected to independently find and use information beyond what is presented during meetings.

**Organization**

Two 1.5-h lectures and a 1-h discussion per week. All meetings will be similar in format, regardless of whether it is a lecture or a

discussion meeting. Presentation of new materials will be interspersed with hands-on practice in all meetings.

Weekly

projects/homework are assigned. There will be a total of 8-10 weekly homework sets assigned through the SmartSite. Homework

is turned in electronically through the corresponding SmartSite assignment page.

**Grading**

Grading will be based on 8-10 homework sets (80%), and announced quizzes (18% and participation (2%).

Participation includes

attendance, asking and answering questions, contributing questions for quizzes, and any other activity that promotes an

environment of active learning and curiosity about the use of statistics.

**Homework**

Submitted work should include complete answers to the questions and the minimum code and output generated to fully support

the answers. For example, if a question asks if there is a significant difference, do not answer just "Yes." Include a statement of

how you reached the conclusion and the corresponding code and output. Late homework is NOT accepted.

Students must turn in

homework within one week of it being assigned or according to the due date specified in the assignment web page.

Name your files as: HW##emailID#.ext

where ## stands for the number of the homework, emailID is the part of your email address before @ucdavis.edu, # is the

sequential file number if you are submitting more than one file, and ext stands for the regular file extension for the application you

used to create the file. Try to submit all work in a single Word or compatible file. For example, if I were taking the course, my first homework file would be named HW01ealaca1.doc. There will be a 1% grade reduction for incorrect file names.

Students are encouraged to work in groups for the homework. Any resources available to students to complete homework is permitted, but sources must be cited, and the answers submitted must be prepared by each student individually. For example, a group of four people discuss a homework together and share calculations and ideas. Once the group reaches a final agreement on how to answer the questions and on what calculations to complete, each student completes these calculations and writes the answers independently. The idea is to learn to learn and be resourceful in a real-world environment, where you will be able to use as many resources as you have time to gather and understand. Do all the work yourself, even if it means that you have to re-run a code that the group already ran.

### Quizzes

Quizzes will be administered at the beginning of meetings to be announced. Students can use any books, notes and websites they want to complete the quizzes, but they cannot ask or communicate with anyone else except for the instructor and teaching assistant. Quizzes will be online, including multiple-choice, fill-in-the-blank, calculation and matching questions. Examples will be shown. Questions will be selected from those contributed by instructors and students.

Author Year Title Publisher Acronym

Faraway 2005 [Linear models with R](#) (Check [Faraway-PRA](#) in Readings) Chapman & Hall/CRC JF1

Faraway 2006 [Extending the linear model with R](#) Chapman & Hall/CRC JF2

Spector 2008 [Data Manipulation with R User R!](#) Springer DMwR

Zuur, Ieno & Smith 2007 [Analysing Ecological Data](#) Springer AEDZuur

Johnson & Wichern 2002 [Applied multivariate statistical analysis](#) Prentice Hall. JW

Kutner, Nachtsheim, &

Neter 2004 [Applied linear regression models](#) McGraw-Hill KNN

Kutner, Nachtsheim, Neter

& Li 2005 [Applied linear statistical models](#) McGraw-Hill KNNL

Rawlings, Pantula, & Dickey 1998 [Applied Regression Analysis](#) Springer RP

Schabenberger & Pierce 2001 [Contemporary Statistical Models for the Plant and Soil](#)

[Sciences](#) www.crcnetbase.com CSM

Sheather 2009 [A Modern Approach to Regression with R](#) Springer AMARR

Oksanen 2013 [Multivariate Analysis of Ecological Communities in R: vegan](#)

[tutorial](#) vegantutor.pdf vegant

Crawley 2007 [The R book](#) Wiley MC

### Topics

1. Review of probability and basic statistical concepts.
2. Simple and multiple linear regression.
3. Matrix approach to linear models.
4. Functions of estimated parameters and their statistical properties.
5. Types of sums of squares. General linear test.
6. Collinearity.
7. Model development and selection.

8. Principal components and principal component regression.
9. Bootstrapping and jackknifing.
10. Extending the linear model: mixed models.
11. Extending the linear model: generalized linear models.
12. Logistic regression. Zero inflated models.
13. Heterogeneity of variance and correlation of residuals.
14. Multidimensional scaling and Correspondence analysis.
15. ~~Path analysis and Structural Equation Modeling~~ - Dropped due to time constraints
16. Multivariate analysis of variance.
17. Discriminant analysis and Bayesian thinking.

### Books and readings

(Mon Sep 29, 2014 02:00 AM)

### PLS 206. Recommended books for consultation

Most books and readings are available to students free of charge. I intend to provide links for all reading materials. Let the instructor know if you cannot find a link for a reading, or if a link is not working. Many Springerlink books are available for free in electronic form to computers with campus IP addresses. In other words, if you want to get the pdf file for a book, use a computer on campus.

Note that the KNN book is listed as "Required" at the bookstore, but it is NOT required. We are trying to maximize the use of books that are freely available to students.

### References

Author	Year	Title	Publisher	Acronym
Faraway	2005	<a href="#">Linear models with R</a> (Check <a href="#">Faraway-PRA</a> in Readings)	Chapman & Hall/CRC	JF1
Faraway	2006	<a href="#">Extending the linear model with R</a>	Chapman & Hall/CRC	JF2
Spector	2008	<a href="#">Data Manipulation with R</a>	User R! Springer	DMwR
Zuur, Ieno & Smith	2007	<a href="#">Analysing Ecological Data</a>	Springer	AEDZuur
Johnson & Wichern	2002	<a href="#">Applied multivariate statistical analysis</a>	Prentice Hall.	JW
Kutner, Nachtsheim, & Neter	2004	<a href="#">Applied linear regression models</a>	McGraw-Hill	KNN
Kutner, Nachtsheim, Neter & Li	2005	<a href="#">Applied linear statistical models</a>	McGraw-Hill	KNNL
Rawlings, Pantula, & Dickey	1998	<a href="#">Applied Regression Analysis</a>	Springer	RP
Schabenberger & Pierce	2001	<a href="#">Contemporary Statistical Models for the Plant and Soil Sciences</a>	www.crcnetbase.com	CSM
Sheather	2009	<a href="#">A Modern Approach to Regression with R</a>	Springer	AMARR
Oksanen	2013	<a href="#">Multivariate Analysis of Ecological Communities in R: vegan tutorial</a>	vegantutor.pdf	vegant
Crawley	2007	<a href="#">The R book</a>	Wiley	MC
Stevens	2009	<a href="#">A primer of Ecology with R</a>	User R! Springer	APER
Zuur, Ieno & Meesters	2009	<a href="#">A Beginner's Guide to R</a>	User R! Springer	ABGR
Husson & Pagès	2010	<a href="#">Exploratory Multivariate Analysis by Example using R</a>	www.crcnetbase.com	EMAER
Borcard, Gillet & Legendre	2011	<a href="#">Numerical Ecology with R</a>	User R! Springer	NEwR
Everitt & Hothorn	2011	<a href="#">An Introduction to Applied Multivariate Analysis with R</a>	User R! Springer	AIAMAR
Shahbaba	2012	<a href="#">Biostatistics with R</a>	User R! Springer	BwR
Tabachnik & Fidell	2001	<a href="#">Using multivariate statistics</a>	Allen & Bacon	TF

Scheiner & Gurevitch	1993	<a href="#">Design and analysis of ecological experiments</a>	Oxford University Pres	SG
Zuur et al.	2009	<a href="#">Mixed Effects Models and Extensions in Ecology with R</a>	Springer	MEMEER
Albert & Rizzo	2012	<a href="#">R by Example</a>	Springer	AR
Kuhn & Johnson	2013	<a href="#">Applied Predictive Modeling</a>	Springer	KJ

### **Recommended Readings by Topic:**

- Review of probability and basic statistical concepts. KNN App. A, AIAMAR Ch1 and Ch2 - **CLASSNOTES 1 & 2**
  - Simple and multiple linear regression. CSM Ch01, AEDZurrCh05 & AR Ch07 - **CLASSNOTES 1 & 2**
  - Matrix approach to linear models. Faraway-PRA Ch02, KNN 5, 8 - **CLASSNOTES 3**
  - Functions of estimated parameters and their statistical properties. - **CLASSNOTES 2**
  - Types of sums of squares. General linear test. - Custom Tests.pdf, **CLASSNOTES 3**
  - Collinearity.
  - Model development and selection. - KNN Ch3, 10 & 11. - **CLASSNOTES 4, 8**
  - Principal components and principal component regression. - RP\$##PCA.pdf, PCAMcCune.pdf, vegantutor.pdf, paperPCR.pdf
  - Bootstrapping and jackknifing. - KNN 11.5 & SG Boots.pdf & BootstrappingFox2002.pdf
  - Extending the linear model: mixed models. MEMEER Ch13
  - Extending the linear model: generalized linear models. MEMEER Ch13
  - Logistic regression. Zero inflated models. - KNN 14, MEMEER Ch. 9 & 11
  - Heterogeneity of variance and correlation of residuals. MEMEER Ch. 4, 6, 7
- 
- Multidimensional scaling and Correspondence analysis. [AEDZuurCh13.pdf](#), [AEDZuurCh32.pdf](#), [vegantutor.pdf](#)
  - Path analysis and Structural Equation Modeling. KNN 7.5 & SG Ch10 PA.pdf
  - Multivariate analysis of variance. SG MANOVA.pdf & JW 6, MANOCAdetailsJMP.pdf
  - Discriminant analysis and Bayesian thinking. JW 11, AEDZuurCh14