Instructor	Cecile Ane
Office	1208 MSC
email	ane@stat.wisc.edu

Required textbook: Data Analysis Using Regression and Multilevel/Hierarchical Models, by Gelman and Hill. Cambridge University Press (2006). I selected this book for its breadth and depth of coverage of modern statistical methods. However, examples in this book are geared towards social scientists. I will therefore provide biological examples from other sources in lecture and in homework assignments. Here are additional references:
[1] Linear Models with R, by J.J. Faraway. Chapman & Hall/CRC.
[2] Extending the Linear Model with R, by J.J. Faraway. Chapman & Hall/CRC.
[3] Ecological Models and Data in R, by B. Bolker. Princeton University Press (2008).

Course website: `www.stat.wisc.edu/courses/st572-ane/` for handouts, homework assignments and other course information. The `Learn@UW` website will be used for grades.

Computing. We will be using the statistical package R which is available for free download online at `cran.r-project.org`. R is free and available on all platforms (Windows, Mac and Linux). This tool is of growing importance in biology. There is a large and vibrant R community and over 25 books published since 2007 about statistical computing with R. A recent article about R was on the front page of the business section of the New York Times, 1/7/09[1]. There is also a multitude of on-line materials aimed at users of different skill levels and background. The book I selected demonstrates the use and output of R. I will assume that you have already had some basic prior experience with R such as from Stat/F&W Ecol./Hort. 571.

Some of you may already have experience with another statistical software package and may work in a discipline where there is another standard, such as SAS. I will not require the use of R for exams, but will provide instruction in the use of R to achieve course goals. If you wish to use an alternative statistical software, it is your responsibility to ensure both that the software is sufficient for the requirements of the course and that you are able to use the software effectively without assistance.

Exams. There will be one in-class midterm exam (75 min.), one take-home midterm exam (one week) and a final exam (2 hours). All exams will be open book, open notes. For the first and last exams, you will be allowed a calculator but not a laptop. The in-class midterm exam will be on **Thursday March 12**$^{\text{th}}$ (just before the break), the take-home midterm exam will be distributed on **Thursday April 23**$^{\text{d}}$ and due on **Thursday April 30**$^{\text{th}}$. The final exam will be on **Wednesday May 13**$^{\text{th}}$ from 2:45pm to 4:45pm, location TBA.

Grading. The final numerical score will be based on homework (20%), the in-class midterm (20%), the take-home exam (30%) and the final exam (30%). Letter grades will be based on this numerical score, using a curve. Letter grades will not be given for midterm exams, but information will be provided to let you know how you are doing in the class.

Discussion Sections. Attendance in discussion sections is strongly encouraged. You may attend any discussion section, so long as there are not too many students in any one discussion. We ask that you commit to attend the same discussion section most of the time. Sections will begin meeting the week of January 27-28.

[1] `http://www.nytimes.com/2009/01/07/technology/business-computing/07program.html?th&emc=th`

Homework will be assigned and posted on the course webpage on Thursdays, to be handed in by the following Friday by 4pm to your TA's mailbox. The mailboxes are in the hallway just inside the main University Ave. entrance to the Medical Sciences Center. Homework must either be typed or otherwise neat. You must show your work and organize it to receive credit. Late homework will be accepted only in extenuating circumstances. If possible, prior arrangements should be made with me in such cases. Unexcused late homework will not receive more than half credit. Late homework handed in after the solution is posted on the website will receive no credit.

Academic honesty. You are encouraged to work together with classmates, talk to your TA or to me about your homework. We are convinced it is very beneficial to share and discuss ideas, and share tips about R. However, you may not present other people's work as your own. If you work with other students solving problems, you still have to write up your own solution and produce computer output independently. You must work independently during exams. You may not share calculators or pass notes during exams.

Laptop policy. You may enjoy the wireless capability of the classroom so long as you stay on task. Advantages of using a laptop include: taking notes, viewing lecture notes rather than printing them, experimenting with R or other computing software, etc. There are also limitations; figures and sketches cannot be created on a notebook in the classroom for instance. However, activities such as emailing, web surfing and gaming are not allowed in class. They are a distraction to classmates: be respectful of others. Be sure the sound is off at the beginning of the class. In "no laptop times" laptop users will be asked to close their lids.

Tentative schedule –subject to change.

T 1/20	Intro and Probability	Chap. 2	Review: linear regression, R.
R 1/22	Multiple linear regression	Chap. 3-4	Model matrix, interpretation of coefficients, interactions
T 1/27	Multiple linear regression	Chap. 3-4	Transformations, polynomial regression, predictions
R 1/29	Multiple linear regression	Chap. 4	Anova, significance, correlated predictors
T 2/3	Multiple linear regression	Chap. 4	nested models, multiple R^2
R 2/5	Multiple linear regression		F-test, sequential tests
T 2/10	Multiple linear regression		model selection with adjusted R^2 and AIC
R 2/12	Logistic regression	Chap. 5	Binomial distribution, logit as link function, interpretation of coefficients, prediction
T 2/17	Logistic regression	Chap. 5	Deviance, model checking
R 2/19	Logistic regression	Chap. 5	chi-square and Wald tests, separation
T 2/24	Poisson regression	Chap. 6	Poisson process, link function, interpretation of coefficients
R 2/26	Poisson regression	Chap. 6	Model checking, overdispersion, significance
T 3/3	Generalized linear regression	Chap. 6	Case studies, multinomial data
R 3/5	Simulation for prediction	Chap. 7	
T 3/10	Simulation for inference	Chap. 7	
R 3/12	In-class midterm	Chap. 2-8	
T 3/17	break		
R 3/19	break		
T 3/24	Simulation for model checking	Chap. 8	
R 3/26	Causal inference	Chap. 9	Randomization, confounding, case study
T 3/31	Causal inference	Chap. 10	Imbalance, lack of overlap, matching, blocking
R 4/2	Multilevel structures and random effects	Chap. 11	Nesting, random effects, linear mixed effect models with `lmer`
T 4/7	Multilevel structures	Chap. 12	Predicting random effects: pooling
R 4/9	Multilevel structures	Chap. 12	Testing random effects, χ^2 tests and simulations
T 4/14	Multilevel structures	Ch. 12-13	Testing fixed effects, case study
R 4/16	CRD with random effects	Ch. 8 in [2]	Likelihood vs. Expected Mean Squares, sample size determination
T 4/21	CRD with subsampling	or Ch. 22	Expected Mean Squares approach, sub/sample size
R 4/23	Take-home exam distributed Case study		
T 4/28	RCBD and Split-plot designs		RCBD: Randomized Complete Block Design
R 4/30	Take-home exams returned Latin square designs		
T 5/5	Nested effects		
R 5/7	Review		
W 5/13	Final Exam	Cumulative	