

Course Announcement MATH 485: Applied Statistical Methods: Fall 2010 MWF  
11:00-11:50 WHAM 0317

Following the course catalog, M485 will cover Sampling Methods and Categorical Data Analysis, and the prerequisite is a calculus based introduction to statistics course such as Math 483. The presentation of the categorical data analysis material is closely related to my presentation of multiple linear regression material in M484, but the response variable is categorical instead of quantitative.

The text for this course is *An Introduction to Categorical Data Analysis* by Alan Agresti  $\approx$  \$100. An optional text is *Introduction to Survey Sampling* by Graham Kalton  $\approx$  \$15, or check out a sampling text from the library (a list of texts will be given).

I plan to use the software packages SAS, Minitab, ARC, and R. SAS is the industry standard, and the four packages are in the Necker's computer lab, where class will be held several times to demonstrate the software. ARC and R are free programs and can be downloaded on a PC.

Sampling techniques include the simple random sample, stratified and cluster samples, and ratio estimation. We will also show why survey results in media are often garbage. For example, according to a Time.com poll, Ronnie O'Brien, an Irish soccer player, was the most important person of the 20th century. Time narrowed its field of candidates to 100 persons, and Elvis won! As another example, Ann Landers asked her readers "if you had it to do over again, would you have children?" 70% of 10000 parents said that they would not have children if they could make the choice again. In a statistical poll, 91% of the parents said that they would have children again.

Most of the course will cover categorical data techniques including the chi-square test, logistic regression, generalized linear models and loglinear models. The basic idea is how do you get information about the categorical variable  $y$  of interest given several explanatory variables  $x_1, \dots, x_p$ . For example, can the survival of a cancer patient be predicted from the age and gender of the patient and the size of the cancer tumor?

*Why should you take Math 485?*

- 1) Math 485 is a lot easier than Math 483.
- 2) If you are interested in a career in statistics, then sampling and categorical data are important courses.
- 3) Being able to list 4 software packages on your resumé will make you more marketable.
- 4) If you are an engineer, get a minor in Math. If you had calculus and M305, get a double major by taking Math 221 and 5 math courses above the 300 level of which at least 3 are above the 400 level (see Dr. Mary Wright in Neckers 357 for more information).
- 5) Learn how to understand bad statistics jokes: i) According to recent surveys, 51 percent of the people are in the majority. ii) Did you hear the one about the politician who promised that, if he was elected, he'd make certain that everybody would have an above average income? iii) A statistician is someone who can have his head in an oven and his feet in ice, and say that on the average he feels great. iv) It is proven that the celebration of birthdays is healthy. Statistics show that those people who celebrate the most birthdays will be the oldest. v) 42.35 percent of statistics are meaningless.

For more information contact David Olive, 261 Neckers.  
Phone: 453-6566, Email: dolive@math.siu.edu

Math 485, Section 1, Fall 2010. MWF 11:00-11:50 WHAM 0317

*Instructor:* David Olive

*Text:* Agresti, A. (2007), *An Introduction to Categorical Data Analysis*, 2nd ed., Wiley, Hoboken, NJ. ISBN: 978-0471226185

*Optional text:* Graham Kalton (1982), *Introduction to Survey Sampling*, Sage Publications, Beverley Hills CA. ISBN: 0803921268

*Office:* (J.W.) Neckers 261 (wing A) *Phone:* (618)-453-6566

*email:* dolive@math.siu.edu *Office hours:* MWThF 10:00-10:30, 2:10-3:10

I am also available by appointment and on a walkin basis, eg before class.

This course covers simple random samples, stratified and cluster samples, and ratio estimation. Categorical data techniques include the chi-square test, logistic regression, generalized linear models and loglinear models.

Cumulative Final: Monday, Dec. 13, 12:50-2:50 PM

The grading and syllabus below are tentative. Last day to drop: Sunday, Oct. 17.

Students receive a WF if they stop attending class and fail. An INC is given if for reasons beyond their control, students engaged in *passing* work are unable to complete all class assignments. One HW may be turned in one class period late with 25% penalty except for the last week of classes.

Week of	MON	WED	FRI
Aug 23	SRS	SRS	SRS
Aug 30	Stratified random	sample	cluster HW1, Q1
Sept 6	no class	sample,	ratio HW2, Q2
Sept 13	estimation	1.1, 1.2, Q3	1.3, 1.4, HW3
Sept 20	2.1, 2.2,	Ex1	2.4
Sept 27	lab?	2.4,4.1,Q4	4.1,4.2,HW4
Oct 4	4.2,4.4	4.4,Q5	5.1, 5.2,HW5
Oct 11	5.2	E2	5.2
Oct 18	lab?	4.3	4.2, HW6
Oct 25	4.4	4.4,Q6	4.3,HW7
Nov 1	3.1,3.2	3.3,Q7	GAM?2.3,HW8
Nov 8	lab?	2.3,Q8	2.3,HW9
Nov 15	2.7	Ex3	2.7
Nov 22	no class	no class	no class
Nov 29	2.7	7.1,Q9	7.1, HW10
Dec 6	7.2	7.2, Q10, HW11	review

*Grading:*

HW	300	drop 1 HW	Quizzes	100	
exam1	100	exam 2	100	exam 3	100
final	300			total	1000
min. grade	points	min. grade	points	min. grade	points
A	900-1000	B	800-899	C	700-799
D	550-699				