

**STAT 6390-501: BIOSTATISTICS**  
**Course information and syllabus**

**Goal:** Learn the fundamentals of how to analyze *categorical data* and *survival data*. The focus will be on learning the basic methodology, data analysis and interpretation. We will use the free statistical software R (available from <http://www.r-project.org/>) for data analysis.

**Instructor:** Pankaj Choudhary, ECSN 3.908, 972-883-4436, [pankaj@utdallas.edu](mailto:pankaj@utdallas.edu)

**Class:** MW, 7 – 8.15 p.m., [CB1.122](#)

**Office hours:** W 3.00 – 4.00 p.m. or by appointment

**Texts:**

1. Categorical Data Analysis, 2<sup>nd</sup> edition, A. Agresti, 2002, Wiley  
Book website: <http://www.stat.ufl.edu/~aa/cda/cda.html>
2. Survival Analysis – Techniques for Censored and Truncated Data, 2<sup>nd</sup> edition, J. P. Klein and M. L. Moeschberger, 2003, Springer  
Book website: <http://www.biostat.mcw.edu/homepgs/klein/book.html>

Both the book websites are very useful. They provide datasets used in the books, software, errata, and other useful information.

**Pre-requisite:** Familiarity with the following topics: maximum likelihood estimation, confidence intervals, tests of hypotheses (at the level of STAT 5351 & 5352 or equivalent) and multiple linear regression (at the level of STAT 5353 or equivalent).

**Course website:** <http://www.utdallas.edu/~pankaj/6390/SP06/> Homework, datasets, announcements, etc., will be posted here.

**Topics to be covered:**

Module I: Analysis of categorical data – first 14 lectures.

Module II: Analysis of survival data – next 14 lectures.

**Grades:** Six 15-minute Quizzes (10% each) = 60%  
Two data analysis projects (20% each) = 40%

- Incomplete grade is possible only in the case of a documented serious medical emergency near the end of the semester, with at least 70% of work completed at an on-going passing grade.
- No make-up or late quizzes. However, it may be possible to take it early.
- Homework will be assigned but will not be collected or graded. Solutions will not be provided. However, the first 10-15 minutes of every class will be devoted to discussion of HW problems. In addition, the solutions of many odd numbered exercises in Agresti's book appear on the book's website. Also, the solutions of odd numbered exercises in K&M's book are given in the book itself.
- You can view your grades on WebCT: <http://webct.utdallas.edu/>
- One page of notes allowed for each quiz.

**Tentative schedule:** (It may change slightly during the semester. The quiz & project dates are firm.)

		# classes	Chapters*
Categorical data: Jan 9 – Feb 27	Jan 9 – Jan 11	2	1
Quiz 1: Jan 30	Jan 18 – Jan 25	3	2-3
Quiz 2: Feb 8	Jan 30 – Feb 1	2	4
Quiz 3: Feb 22	Feb 6 – Feb 20	5	5-6
Project 1 given on: Feb 27 due on: Mar 13	Feb 22 – Feb 27	2	7
Survival data: Mar 1 – Apr 24	Mar 1 – Mar 15	3	2-3
Quiz 4: Mar 22	Mar 20 – Mar 22	2	4
Quiz 5: Apr 5	Mar 27 – Mar 29	2	7
Quiz 6: Apr 19	Apr 3 – Apr 10	3	8
Project 2 given on: Apr 10; due on: Apr 24	Apr 12 – Apr 17	2	11
	Apr 19 – Apr 24	2	12

\* Agresti's book for categorical data and K&M's book for survival data.