Appendix 5 Introductory and Advanced Electives Offered by the Department of Biostatistics

Introductory electives:

- **140.608 Health Administration Statistics**. 3rd Term. Methods appropriate for use in quantitative analysis of current problems facing public health administrators; classification and evaluation of mortality data, consideration of the limitations an users of illness data, and a discussion of the sources and uses of population data; techniques useful in the planning and evaluation of maternal and child health, medical care, population planning, and other service programs, collection and use of attitude data in public health studies; basic principles for the conduct of field investigations.

- **140.640 Statistical Methods for Sample Surveys**. 3rd Term. (*Requires 140.621-22 or 651-52*) Construction of sampling frames; area sampling; methods of estimation; stratified sampling; subsampling; sampling methods for surveys of human populations. Student evaluation is based on homework and a final exam.

- **140.641 Survival Analysis**. 3rd Term. (*Requires 140.621-22 or 651-52*) This course will discuss the basic concepts of survival analysis including hazard functions, survival functions, types of censoring, Kaplan-Meier estimates, log rank tests, and the generalized Wilcoxin tests. Parametric inference will include the exponential and Weibull distribution. The proportional hazard models and extensions to time dependent covariates will be discussed. Many clinical and epidemiological examples will be given to illustrate the various statistical procedures. Students will be evaluated on the basis of problem sets and an exam.

- **140.655 Analysis of Longitudinal Data**. 3rd Term. (*Requires 140.621-22 or 651-52*) This course covers statistical models for drawing scientific inferences from longitudinal data. Topics include longitudinal study design; exploring longitudinal data; linear and generalized linear regression models for correlated data, including marginal, random effects, and transition models; and handling missing data. This course is intended for doctoral students in quantitative sciences. Student evaluation is based on analysis of a longitudinal data set, presentation of results, and a written scientific report of the analysis methods and results.

- **140.662 Spatial Analysis and GIS I**. 3rd Term. (*Requires 611-12 or statistical equivalent*) This course will examine the use of ArcView Geographic Information System (GIS) software as a tool for integrating, manipulating, and displaying public health related spatial data. Topics covered will include mapping, geocoding, and manipulations related to data structures and topology. Selected

case studies will be used to demonstrate concepts. Focus will be on using GIS to generate and refine hypotheses about public health related spatial data in preparation for a formal statistical analysis. Although spatial statistical modeling is not a required part of the curriculum, related topics will be discussed throughout.

- **140.632 Introduction to the SAS Statistical Package**. 4th Term. (*Requires 140.621-22 or 651-52*) Through this course, the student will become an adept user of the SAS statistical package, mastering the skills needed for effective data management, data manipulation, and data analysis. The student will learn how to document his or her work, and make the work replicable. Graphical techniques for displaying data will be discussed. While this course will use the SAS statistical package exclusively, much of the technical knowledge and some of the computing techniques covered in this course will be applicable to any statistical package.

Advanced electives

- **140.642 Design of Clinical Experiments**. 1st term. (*Requires 621-22; offered alternate odd years*) Introduces the application of traditional experimental design theory to biomedical control experiments, including event time studies. Stresses methods of bias and variability, particularly randomization, blocking, factorial designs, stratification, and adequate sample size. Emphasizes clinical trials and other types of medical experiments likely to be encountered by biometric researchers. Discusses elements of analysis when they relate to the design principles. Students seeking a course with similar content are encouraged to take 390.612, Design and Analysis of Clinical Trials. Student evaluation based on problem sets and a short protocol for a designed experiment.

- **330.657** **Statistics for Psychosocial Research: Measurement**. 1st term. (*Requires 621-24 or 651-54; jointly offered with the Dept of Mental Hygiene*) Presents quantitative approaches to measurement in the psychological and social sciences. Topics include the principles of psychometrics, including reliability and validity; the statistical basis for latent variable analysis, including exploratory and confirmatory factor analysis and latent class analysis; and item response theory. Draws examples from the social sciences, including stress and distress, social class and socioeconomic status, personality; consumer satisfaction, functional impairment and disability, quality of life, and the measurement of overall health status. Intended for doctoral students. Student evaluation based on class participation, problem sets, and a final exam.

- **140.658** **Statistics for Psychosocial Research: Structural Models**. 2nd term. (*Requires 330.657*) This course presents quantitative approaches to theory construction in the context of multiple response variables, with models for both

continuous and categorical data. Students will become familiar with: the statistical basis for causal inference; principles of path analysis; linear structural equation analysis incorporating measurement models; latent class regression; and analysis of panel data with observed and latent variable models. Examples will be drawn from the social sciences, including the status attainment approach to intergenerational mobility, behavior genetics models of disease and environment, consumer satisfaction, functional impairment and disability, and quality of life. Student evaluation is based on: 1) problem sets requiring active manipulation of data sets provided by the instructors, using standard statistical packages as well as more specialized problems for latent variable analysis; 2) class participation; and 3) a final exam.

- **140.668 Special Topics in Genetics and Genomics**. 2nd Term. (*Requires 651-654*) Addresses statistical issues in genetics and genomics. Consists of two four-week modules, with revolving instructors and topics. Possible topics include the following: genetic mapping in experimental organisms; genetic map construction; linkage analysis in humans; linkage disequilibrium in humans; population genetics; phylogenetic inference; topics in protein structure; microarray analysis; and proteomics.

- **140.663 Spatial Analysis and GIS II**. 4th term. (*Requires 621-24 or 651-54)* This course will introduce the statistical techniques used to model, analyze, and interpret public health related spatial data. Generalized linear mixed models will be presented as a general framework for analyzing spatially dependent data. Some topics covered include the geostatistical techniques of kriging and variogram analysis and point process methods for spatial case control analyses. Although the focus will be on statistical modeling, some time will be spent covering topics related to clustering and cluster detection of disease events. Although helpful, knowledge of specific GIS software is not required. Instruction in the public domain statistical package R, to be used for analyses, will be provided.

- **140.665 Causal Inference**. 4th term. (*Requires 651-654*) Introduces and discusses role of potential outcomes and treatment assignment mechanisms for drawing causal inferences; studies completely randomized assignments; known unconfounded assignments, covariates, and role of Fisher's, Neyman's, and Bayesian methods; ignorable assignments, propensity scores and sensitivity analysis; nonignorable assignments arising from deviations to protocol, treatment-noncompliance, direct and indirect effects, methods of instrumental variables, loss to follow-up and methods of latent ignorability, and encouragement designs. Evaluation based on problem sets and a final project.

- **140.688 Statistics for Gene Expression**. 4th term. (*Requires 621-24*) Introduces statistical concepts and tools necessary to interpret and critically evaluate the literature on gene expression array data, and perform basic analysis of gene expression array data. Includes an overview of oligonucleotide analysis,

normalization, identification of differentially expressed gene, clustering, classification, and statistical pattern recognition. Interested students are encouraged to visit the course website. Student evaluation based on micorarray data analysis project.