



---

# **Master's Degree Program in Biostatistics**

**Spring 2007 Syllabi**

**Department of Biostatistics,  
Bioinformatics & Biomathematics  
Georgetown University**

---

**BIST-512 Statistical Modeling I (4 credits)**

Lei Nie, PhD & Ying Zhang, MD, MS  
 8:00 – 9:40 AM, Mondays and Wednesdays  
 W402 Research Building

Date	Day	Topics	Chapters	Faculty
Jan. 10	Wednesday	Simple Regression Analysis (Inference)	1,2	YZ
		Sas Basic (1)		BS
Jan. 17	Wednesday	Simple Regression Analysis (Diagnosis)	3,4	YZ
Jan. 22	Monday	Multiple Regression (Inference)	6	YZ
		Sas Basic (2)		BS
Jan. 24	Wednesday	Multiple Regression (Inference)	7	YZ
Jan. 29	Monday	Model Diagnosis	9	YZ
Jan. 31	Wednesday	Analysis of Variance (Inference)	16	YZ
Feb. 5	Monday	Analysis of Variance: Diagnosis and Multiple Comparison	17-18	YZ
Feb. 7	Wednesday	Analysis of Variance (2 ways, part 1)	19	YZ?
Feb. 12	Monday	Analysis of Variance (2 ways, part 2)	20-21	YZ ?
Feb. 14	Wednesday	Analysis of Covariance	25	YZ?
Feb. 21	Wednesday	Model Selection	8	LN
Feb. 26	Monday	Analysis of Variance (random and mixed)	24	LN
Feb. 28	Wednesday	Nested Design and Split-plot Design	Handout	LN
Mar. 12	Monday	Review session (YZ)		
Mar. 14	Wednesday	<b>MIDTERM EXAM(YZ)</b>		
Mar. 19	Monday	Linear Mixed Models : Introduction	3* and 4*	LN
Mar. 21	Wednesday	Linear Mixed Models : Estimation	5*	LN
Mar. 26	Monday	Linear Mixed Models : Inference	6*,7*	LN
Mar. 28	Wednesday	Model Building and Local Inference	9*,10*,11*	LN
Apr. 2	Monday	Incomplete Data	14-16	LN
Apr. 4	Wednesday	Nonlinear Regression	13	LN
Apr. 11	Wednesday	Logistic Regression and Poisson Regression	14	LN
Apr. 16	Monday	Generalized Linear Model (I)	15	LN
Apr. 18	Wednesday	Generalized Linear Model (II)	15 and Handout	LN
Apr. 23	Monday	Generalized Estimating Equation (I)	Handout	LN
Apr. 25	Wednesday	Generalized Estimating Equation (II)	Handout	LN
Apr. 30	Monday	Review session (LN)		
		<b>FINAL EXAM (LN)</b>		

**Course Description:**

BIST-512 Statistical Modeling I is an applied course on statistical modeling of many data routinely met in health sciences. The first part of this course introduces the analysis of variance (ANOVA), simple and multiple regressions, and generalized-linear models, which are three of the most frequently used models. The second part focuses on longitudinal data analysis with linear mixed-effects models, generalized linear models, and Generalized Estimating equations.

Learning statistical software SAS will be one important task for this course. Students are expected to use SAS to complete their homework and projects.

**BIST-513 Statistical Modeling II (4 credits)**

Ionut Bebu, PhD & Antai Wang, PhD  
 8:00 – 9:40 AM, Tuesdays and Thursdays  
 W402 Research Building

Date	Day	Topics	Chapters	Faculty
		<b>MULTIVARIATE ANALYSIS</b>		
January 11	Thursday	Introduction to Multivariate Data	1-2	IB
January 16	Tuesday	Multivariate Normal Distribution	Handout	IB
January 18	Thursday	Multivariate Analysis of Variance	4	IB
January 23	Tuesday	Multivariate Analysis of Variance	4	IB
January 25	Thursday	Principal Components Analysis	6	IB
January 30	Tuesday	Applications + Computer Lab	Handout	IB
February 1	Thursday	Factor Analysis	7	IB
February 6	Tuesday	Discrimination and Classification	8	IB
February 8	Thursday	Canonical Correlation Analysis	10	IB
February 13	Tuesday	Applications + Computer Lab	Handout	IB
		<b>SURVIVAL ANALYSIS</b>		
February 15	Thursday	Introduction to Survival Analysis	C&O-1	AW
February 20	Tuesday	Censoring at Random, KM estimator	L-9.1	AW
February 22	Thursday	Failure Time Models	K&P-2	AW
February 27	Tuesday	Failure Time Models	K&P-2	AW
March 1	Thursday	Tests	L-9.3	AW
March 13	Tuesday	Cox Proportional Hazards Models	L-9.4	AW
March 15	Thursday	Cox Proportional Hazards Models	L-9.4	AW
March 20	Tuesday	EM Algorithm	Handout	AW
March 22	Thursday	Sample Size and Power Calculations	L-9.5	AW
March 27	Tuesday	Multivariate Survival Analysis	Handout	AW
March, 29	Thursday	Frailty Models	Handout	AW
April, 3	Tuesday	Applications + Computer Lab	Handout	AW
		<b>MARKOV MODELS</b>		
April, 5	Thursday	Introduction to Markov Models	HPS-1	IB
April, 10	Tuesday	Probabilistic Properties of Markov Chains	HPS-1	IB
April, 12	Thursday	Equilibrium Distributions	HPS-2	IB
April, 17	Tuesday	Estimation for Markov Chains	Handout	IB
April, 19	Thursday	Hidden Markov Models	DEKM-3	IB
April, 24	Tuesday	Estimation for Hidden Markov Models	DEKM-3	IB
April, 26	Thursday	Applications + Computer Lab	DEKM-4	IB

**Course Description:**

*Multivariate Analysis:* multivariate normal distribution, multivariate analysis of variance, principal components, canonical correlation, factor analysis, discrimination and classification.

*Survival Models:* concepts and methods for analyzing survival time data, life tables, survival distributions, mathematical and graphical methods for evaluating goodness of fit, comparison of treatment groups, regression models, proportional hazards models.

*Markov Models:* probabilistic properties of Markov chains, equilibrium distributions, Hidden Markov models, parameter estimation, applications in biology.

**BIST-531 Pattern Recognition (4 credits)**

Habtom Resson, PhD

4:00 – 5:40 PM, Mondays and Wednesdays

W402 Research Building

Date	Day	Topics	Chapters
Jan. 10	Wednesday	Introduction to Pattern Recognition	1
Jan. 15		<i>Martin Luther King, Jr. Day</i>	
Jan. 17	Wednesday	Review of Statistical Methods for Pattern Recognition	2,3
Jan. 22	Monday	Probabilistic Neural Networks, K-Nearest Neighbor Rule	4
Jan. 24	Wednesday	Fuzzy Classification	4
Jan. 29	Monday	Linear Discriminant Functions	5
Jan. 31	Wednesday	Support Vector Machines	5
Feb. 5	Monday	Multilayer Perceptrons	6
Feb. 7	Wednesday	Radial Basis Function Networks, Recurrent Networks	6
Feb. 12	Monday	MATLAB Session	handout
Feb. 14	Wednesday	Applications of Classification Methods in Bioinformatics	handout
Feb. 19		<i>President's Day</i>	
Feb. 21	Wednesday	<b>EXAM ONE</b>	
Feb. 26	Monday	Evolutionary Methods	7
Feb. 28	Wednesday	Swarm Intelligence	handout
Mar. 5-9		<i>Spring Break</i>	
Mar. 12	Monday	Decision Trees	8
Mar. 14	Wednesday	CART, ID3, C4.5	8
Mar. 19	Monday	Rule-Based Methods	8
Mar. 21	Wednesday	Bagging, Boosting, Cross-Validation, Jackknife, Bootstrap	9
Mar. 26	Monday	Feature Extraction	handout
Mar. 28	Wednesday	Feature Selection	handout
Apr. 2	Monday	<b>EXAM TWO</b>	
Apr. 4	Wednesday	Unsupervised Learning and Cluster Analysis	10
Apr. 9		<i>Easter Break</i>	
Apr. 11	Wednesday	K-Means	10
Apr. 16	Monday	Fuzzy K-Means	10
Apr. 18	Wednesday	Hierarchical Clustering	10
Apr. 23	Monday	Principal Component Analysis, Multidimensional Scaling	10
Apr. 25	Wednesday	Self Organizing Maps	10
Apr. 30	Monday	Applications of Pattern Recognition in Bioinformatics	handout
May 9		<b>FINAL PROJECTS DUE</b>	

**Course Description:**

BIST-531 Pattern Recognition course will introduce you to the fundamentals of pattern recognition and its application in extracting biological knowledge from high dimensional data. The course will discuss several supervised and unsupervised algorithms and how they can be applied for various purposes including feature extraction, feature selection, dimensionality reduction, clustering, and classification. Particular emphasis will be given to computational methods such as linear discriminant functions, nearest neighbor rule, weighed voting, artificial neural networks, fuzzy logic, support vector machines, genetic algorithms, and swarm intelligence. The course will look into applications of pattern recognition in genomics and proteomics (e.g., DNA base calling, analysis of microarray and mass spectral data, etc.).

