PRINCIPAL SUBJECT AREA

STATISTICS

400 LEVEL COURSES

ST 401 Actuarial Statistics (2 credits)
(Prerequisites: ST 203)
Economics of uncertainty. Risk theory and utility. Jensen's inequality. Sums of random variables and convolutions. Applications to individual risk models. Failure rates and the force of mortality. Mixtures of random variables and mixtures of distributions. Loss distributions, Reinsurance. Risk models, Estimating distribution by simulation. Actuarial applications of statistical inference. Compound distribution. Collective risk models. Ruin theory. Lundberg's Inequality, Introduction to credibility theory. Compound stochastic processes. Applications of risk theory in insurance problems. No claims discounting. Run off triangles.

Recommended Texts:
1. *An Introduction to the Mathematics of Finance (Chapters 1-4),* J.J. McCutcheon and W.F. Scott.,(1998)
2. *Life Contingencies (Chapters 1-6),* A. Neill.(1999)
3. *Actuarial Mathematics (Chapters 3-8),* N.L. Bowers Jr, ... [et al.].(2001)
4. *Mathematical Models for the Growth of Human Populations,* .H. Pollard (1997)

ST 402 Statistical Data Mining (3 credits)
(Prerequisites: ST206, CS 409)
Classification –Neural Network Based Algorithms, Rule Based Algorithms, Combining Techniques.
Clustering – Similarity and Distance Measures, Hierarchical Algorithms, Partitional Algorithms, Clustering Large Data Bases, Clustering with Categorical Attributes, Comparison.
Association Rules – Large Item Sets, Basic Algorithms, Parallel and Distributed Algorithms, Comparing Approaches, Incremental Rules, Advanced Association Rule Techniques, Measuring the Quality of Rules.
Web Mining – Web Content Mining, Web Structure Mining, Web Usage Mining.
Spatial Mining – Spatial Data Overview, Spatial Data Mining Primitives, Generalization and Specialization, Spatial Rules, Spatial Classification Algorithm, Spatial Clustering Algorithms.
Temporal Mining – Modeling Temporal Events, Time Series, Pattern Detection, Sequences, Temporal Association Rules.
Some practical assignments will be given for this course
Recommended Texts:
1. *Data Mining Introductory and Advanced topics,* M.H. Dunham (2003)
2. *Predictive Data Mining,* Weiss SM & Indurkhya N, Morgan Kaufmann (1997)
3. *Principles of Data Mining, Hand DJ et al,* MIT Press (2001)

ST 403 Statistics for Bioinformatics (2 credits)
Review of the following in the context of bioinformatics: Basic probability, statistical inference, stochastic processes, computer intensive approaches to statistical inference, applications. Mathematical models and computational methods of statistical genetics including mendelian genetic traits, population genetics, pedigree relationships and gene identity, meiosis and recombination, linkage detection, multipoint linkage analysis. Course work involves some computation in a Unix environment.

Recommended Texts:
1. *Biological Sequence Analysis,* R. Durbin, S. Eddy, A. Krogh and G. Mitchison (1998)
2. *Statistical Methods in Bioinformatics, An Introduction,* W. J. Evans, G.R. Grant (2001)

ST 404 Stochastic Processes (2 credits)
(Prerequisites: ST201, ST 203)
Introduction to Stochastic processes: Markov Chains, Markov Processes with Discrete state space, Markov Processes with Continuous state space, Stationary processes, Branching Processes, Stochastic processes in Queueing and Reliability

Recommended Text:
1. *Stochastic Processes,* J. Medhi (1996)

ST 405 Multivariate Methods II (2 credits)
(Prerequisites: ST 305)
Discriminant analysis of two group and multiple groups, Principal component analysis (PCA). Interpretation using illustrative examples. Factor analysis. Comparison with PCA, factor loadings, rotations, Interpretation, Canonical correlation, Covariance structure models.
Some practical assignments will be given for this course

Recommended Texts:
1. *Multivariate Statistical inference & Applications,* A.C. Rencher(1990)
2. *Applied Multivariate Statistical Analysis,* R.A. Johnson and D.W. Wichern (1982)

ST 406 Bayesian Statistics II (2 credits)
(Prerequisites: ST 308)
Decision Rules; Making Decisions when data is not available: Specifying a prior distribution, Making decisions with only prior information; Making Decisions when data is available: Decision trees, Expected Value of Perfect Information (EVPI), Expected Value of Sample Information (EVSI), Non-informative and natural conjugate prior, Bayesian confidence intervals.

Recommended Texts:
1. *Statistical Decision Theory and Bayesian Analysis,* J.O. Berger (1985)
2. *Bayes and Empirical Bayes methods for Data analysis,* B.P. Carlin and T.A. Louis (1996)

ST 407 Linear Models (3 Credits)
(Prerequisites: ST 105, ST 203)
Elementary Theorems on Linear and Matrix Algebra, Partitioned Matrices, Nonnegative Matrices; Generalized Inverses of Matrix; Solutions of Linear Equations; Idempotent Matrices, Trace of Matrices; Derivatives of Quadratic Forms, Expectation of Matrix, Multivariate Normal Distribution, Distribution of Quadratic Forms; General Linear Model, Optimal Estimation and Hypothesis Testing Procedures for the General Linear Model, Applications to Regression Models. Application of Optimal Inference Procedures for the General Linear Model to Multifactor Analysis of Variance, Experimental Design Models, Analysis of Covariance, Split-Plot Models, Repeated Measures Models, Mixed Models, Variance Component Models.

Recommended Texts:
1. *Matrices with Applications in Statistics,* F.A.Graybrill, Wadsworth
2. *Theory and Applications of the Linear Model,* F.A.Graybrill, Wadsworth

ST 408 Reliability Theory and survival analysis (3 Credits)
(Prerequisites: ST 203)
Reliability Theory: General introduction, Reliability concepts, Classical model, Weibull distribution: Censored observation, Parameter estimation, Asymptotic results.
Survival analysis: Introduction, Survival curves, Parametric modeling, Cox's proportional hazards model, Extensions.

Recommended Texts:
1. *Survival Analysis: A practical approach,* D Machin, Y.B. Cheung,M. Parmar, Wiley
2. *Statistical Methods for Reliability Data,* Meeker M. Q. and Escobar L. A, Wiley

CS 409 Neural networks and Fuzzy logics (3 Credits)
Fuzzy system models, Fuzziness and certainty, fuzzy sets, basic properties and characteristics, Domains, Alpha- level sets and support sets, Linear representation, Fuzzy set operators, Conventional (crisp) set operations, basic Zadeh type operations, intersection, union and complement of fuzzy sets, General algebraic operations, Fuzzy set hedges, Fuzzy reasoning, linguistic variables, Fuzzy models, Fuzzy systems and modeling, Design methodologies, modeling and utility software. Parallel and distributed processing, Neuron, Connectivity, Activation function, propagation rule, Learning rules, pattern Preparation, Perceptron, Multilayer perceptron, Associative memory, Hopfield neural networks, Self organizing map (SOM), Adaptive Resonance theory, topologies, Training methods, supervised and unsupervised learning

Recommended Texts:
1. *Artificial Intelligence: A Modern Approach* (Second Edition) by Stuart Russell and Peter Norvig.

ST 426 Research Project (6 Credits)
Students are expected to carry out an independent research project on a topic assigned to him/her under the supervision of a senior staff member. At the completion of the project students are expected to write a report and make a presentation.

Students, who do not offer Statistics as a major subject, can take the following course units after following the Applied Statistics (ST202) course. Note that ST 103 course can be followed concurrently with ST 202.

ST 103 Statistics Applications I
ST 204 Sampling Techniques
ST 303 Design and Analysis of experiments
ST 304 Non- parametrics and Categorical Data Analysis

ST 104 Statistics Applications II
ST 301 Regression Analysis
ST 302 Statistical Quality Control