UNIVERSITY OF DEBRECEN
Doctoral School of Informatics

Complex exam major subject

Stochastic models

Syllabus

Basic topics of probability theory. Probability space, Caratheodory's extension theorem. Random variables, probability distribution functions, Lebesgue-Stieltjes measure, independence, Fubini's theorem. Kolmogorov's extension theorem.

Limit theorems of probability theory. Weak and strong laws of large numbers. Characteristic functions, central limit theorems. Convergence of probability distributions in function spaces, the C and D spaces.

The role of normal distribution in statistics. The definitions and properties of univariate and multivariate normal distributions. Distributions derived from the normal law. Conditional distribution in the joint normal case. Gaussian stochastic process (Gaussian system). Best linear prediction for Gaussian system.

Staistical foundations of data mining. Preprocessing data, data exploration, descriptive statistics, visualization. Applications of classification methods for data mining.

Statistical estimators. Properties of point estimators: unbiased, consistent and efficient estimators. Estimation methods, maximum likelihood estimator. Rao-Cramer inequality. Sufficient statistics, Fisher-Neyman factorization theorem, Rao-Blackwell theorem. The exponential family of distributions. Confidence intervals. Nonparametric estimators.

Statistical tests. Null-hypothesis, alternative hypothesis, critical region, power function. Randomized tests, Neyman-Pearson lemma. Parametric tests. Nonparametric tests, Kolmogorov-Smirnov tests.Testing normality.

Variance and regression analysis. Linear regression. The Fisher-Cochran theorem, models in analysis of variance. Linear model. Least squares and maximum likelihood estimators in the linear model, Gauss-Markov theorem. Testing hypotheses in the linear model. Connection of the linear model with analysis of variance and regression analysis.

Classification methods. Bayes's rule classifier, logistic regression, discriminant analysis, nearest-neighbour classifier, cluster analysis, neural networks. Examples.

Dimensionality reduction. Principal component analysis, canonical correlation analysis, exploratory factor analysis. Examples.

Neural networks. Multilayer preceptron, error back-propagation, optimization methods. Convolutional netwoks. Radial basis function network, penalty functions. Support vector machines, Kuhn-Tucker theorem. Recurrent networks. Fitting functions and classification by neural networks.

Discrete time martingales. Definition of a martingale, submartingale and supermartingale, examples, basic properties, Doob's decomposition of submartingales, stopping times, optional sampling theorem, Wald's identity, Doob's maximal inequality, Doob's upcrossing lemma, convergence theorems for (sub)martingales.

Definitions of discrete time Markov chains, classification of states, periodicity, discrete renewal theorem, Pólya's theorem on symmetric random walk, stationary distribution and measure, Perron's theorem, ergodic theorems.

Stochastic calculus. Continuous time martingales (Doob's maximal inequality, (sub)martingale convergence theorem, quadratic variation). Wiener process (its definition, existence, properties, martingale characterization), Markov processes. Stochastic integrals (definition, properties, Itô's formula). Stochastic differential equations (existence of weak and strong solutions, uniqueness of the solution). Linear stochastic differential equations (Wiener bridge, Ornstein-Uhlenbeck process).

Time series analysis. ARMA, ARIMA models, lag operator, stationarity, autocovariance, autocorrelation, prediction, Box-Jenkins method, deterministic trend. VAR (vector-autoregression) models and cointegration.

Network models, random graphs. Erdős-Rényi, Barabási-Albert and Watts-Strogatz models. Degree distribution (scalefree graphs), small world, clustering.

Bibliography

1. Ash, R. B.: Real analysis and probability. Academic Press, New York, 1972.
2. Ash, R. B., Gardner, M. F.: Topics in stochastic processes. Academic Press, New York, 1975.
3. Athreya, K. B., Lahiri, S. N.: Measure Theory and Probability Theory, Springer, 2006.
4. Barabási, Albert-László: Network science. Cambridge University Press, 2018.
5. Bauer, H.: Probability Theory, Walter de Gruyter, Berlin, 1996.
6. Billingsley, P.: Convergence of probability measures. Wiley, New York, 1999.
7. Christensen, R., Linear models for multivariate, time series, and spatial data. Springer, New York, 1991.
8. Chung, K. L.: Markov chains. Springer, New York, 1967.
9. Chow, Y. S., Teicher H.: Probability Theory, 3rd edition, Springer 1978.
10. Durrett, R.: Random Graph Dynamics. Cambridge University Press, 2007.
11. Everitt, B., Hothorn, T.: An Introduction to Applied Multivariate Analysis with R. Springer, 2011.
12. Gibbons, J. D.: Nonparametric statistical inference. McGraw-Hill, New York,1971.
13. Gikhman, I. I., Skorokhod, A. V.: The Theory of Stochastic Processes. Springer, 2004.
14. Giri, N. C., Introduction to probability and statistics. Dekker, New York,1985.
15. Goodfellow, I., Bengio, Y., Courville, A. Deep Learning. MIT Press, 2016.
16. Gut, A.: Probability: A Graduate Course. Springer, 2005.
17. Guttorp, P.: Stochastic Modeling of Scientific Data. Chapman and Hall, London, 1995.
18. Hamilton, J. D.: Time Series Analysis, Princeton University Press, Princeton, New Jersey, 1994.

19. Haykin, S.: Neural Networks and Learning Machines. Pearson, 2009.
20. Izenman, A. J.: Modern Multivariate Statistical Techniques. Regression, Classification, and Manifold Learning. Springer, 2008.
21. Karatzas, I., Shreve, S. E.: Brownian Motion and Stochastic Calculus, 2nd edition, Springer-Verlag, 1991.
22. Lehmann, E. L., Romano, J. P.: Testing statistical hypotheses. Springer, 2005.
23. Maddala, G. S., Lahiri, K.: Introduction to Econometrics, Wiley, 2010.
24. Mardia, K. V., Kent, J. T.: Bibby, J. M., Multivariate analysis. Academic Press, London, 1979.
25. Medvegyev, P.: Stochastic Integration Theory, Oxford University Press, 2007.
26. Protter, P. E.: Stochastic Integration and Differential Equations, 2nd edition, Springer, 2005.
27. Rényi, A.: Probability Theory, Dover Publications, Inc, 2007.
28. Rohatgi, V. K.: Statistical Inference. Wiley, New York, 1984.
29. Scheffé, H.: The analysis of variance. Wiley, New York, 1999.
30. Schervish, M. J.: Theory of statistics, Springer, New York, 1995.
31. Shiryayev, N.: Probability, 2nd edition, Springer-Verlag, 1995.
32. Tan, Pang-Ning; Steinbach, M., Kumar, V.: Introduction to Data Mining. Pearson Education Inc. 2006.
33. Timm, N. H.: Applied Multivariate Analysis. Springer, 2002.

Compulsory subjects for this major subject	Machine Learning; Selected topics in probability
Recommended subjects for this major subject	Topics in the theory of stochastic processes; Random graphs and networks; Convergence of Probability Measures; Financial mathematics