

Exam Questions

1. Basic concepts: sample, sample realization, statistics. Examples: mean, standard deviation, ordered sample, empirical distribution function, etc. Glivenko-Cantelli Theorem. Theory of Estimation: unbiased estimation, consistent estimation, efficient estimation, sufficient estimation. Examples. Fisher-Neyman factorization theorem. Rao-Blackwell-Kolmogorov Theorem. Cramer-Rao Inequality, the Fisher information. Maximum likelihood- and the momentum methods.
2. Distributions derived from normal distribution: Chi-square, Student, F (Fisher). Independency of the mean and the standard deviation in normal case. Multivariate Normal Distribution. Central region of the multivariate normal distribution.
3. Confidence intervals. Examples in normal case. Basic concepts of hypothesis theory: nul hypothesis, types of error, significance level. Parametric Tests: u- and t-tests. Welch-test. F-test. Bartlett test.
4. Nonparametric tests 1. Chi-square tests. Connection between multinomial- and the chi-square distribution. Chi-square Goodnes of fit tests. Chi square homogeneity test, Chi-square test for independence. Kolmogrov-Smirnov tests: Goodness of fit and homogeneity tests.
5. Nonparametric tests 2. Mann-Whitney U test. Kruskal-Wallis test. Wilcoxon test. Friedman test. Levene test. Run test. Median test. Sign test. Sequential tests. Wald's Theorem. Exact tests.
6. ANOVA. Cochran theorem. One-Way ANOVA, Post Hoc tests. Welch ANOVA test. Two-way ANOVA without interactions and with interactions. The latin square design.
7. Regression Analysis. The conditional expectation. Properties of the conditional expectation. Regression in normal case. Linear regression between two variables. Method of the least squares. Gauss-Markov Theorem. Polynomial regression. Two parametrical non-linear regressions between two variables. R-square, the coefficient of determination. Nadaraya-Watson regression.
8. Multivariate linear regression. Beta coefficients. Partial F-test. Automated Model Building Algorithms: Backward, Forward, Remove (with ENTER), Stepwise. Adjusted Coefficient of Determination. Checking the validity of the model. Heteroscedasticity. Multicollinearity. Exploring of the points which effect significantly to the regression. Binary logistic regression.

9. Factor and Principal Components Analysis (PCA). The k-factors model. Kaiser-Meyer-Olkin statistics, Measure of Sampling Adequacy. Cummunality and uniqueness. Rotations: Varimax, Equamax, Quartimax. Optimal dimensionality reduction with PCA.
10. Pattern Recognition. Basic concepts: feature vector, clas.sification, decision function, risk, cost , Bayes decision, Bayes risk. Discriminant Analysis. Linear Discriminant Analysis. Wilk's lambda. Nearest Neighbor Rule. Cover-Hart Theorem. Quickly Searching the nearest neighbor.
11. Cluster Analysis. The wording of the problem. Types of Clustering: partitioning (k-means) and hierarchical clustering. The McQUEEN algorithm, McQueen Theorem. Dendrogram, hierarchical algorithms.
12. Sampling tecnics. The representative sample. Random sample, systematic random sample, stratified random sampling, cluster sampling, snowball sampling, census, more stage samplings. Determining the Sample Size. Chernoff-, Bernstein-, Hoeffding Inequalities.