

Többváltozós statisztika 3/1/0/v/5

Tárgyfelelős: Bolla Marianna

Többdimenziós Centrális Határeloszlás Tétel és alkalmazásai. A statisztikában használt véletlen mátrixok (Wishart-, Wigner-mátrixok) sűrűsége, spektruma és aszimptotikus eloszlása. Sajátértékekre és szinguláris értékekre vonatkozó szeparációs tételek alkalmazása a főkomponens-, faktor-, kanonikus korreláció- és korrespondenciaanalízisben. Faktoranalízis, mint alacsony rangú reprezentáció, reprezentáció és metrikus klaszterező eljárások kapcsolata. Klasszifikációs módszerek: diszkriminanciaanalízis, hierarchikus, k-közép és gráfelméleti módszerek a klaszteranalízisben. Gráfok spektruma és becsülhető paraméterfüggvényei. Algoritmikus modellek, tanulóalgoritmusok. EM-, ACE-algoritmus, Kaplan–Meier-becslések. Újramintavételezési eljárások: bootstrap és jackknife. Adatbányászati alkalmazások, randomizált módszerek nagyméretű problémákra. A többváltozós statisztikai módszerek használatának és angol nevezéktanának elsajátítása egy programcsomag segítségével (SPSS vagy S+), output eredmények alkalmazásorientált elemzése.

Irodalom:

Bolla, M., Krámli, A.: Theory of statistical inference (in Hungarian), Typotex, Budapest, 2005

Mardia, K. V.; Kent, J. T., Bibby, J. M.: Multivariate Analysis, Academic Press, Elsevier Science, 1979, 2003

Multivariate statistics 3/1/0/v/5

Responsible: Marianna Bolla

Multivariate central limit theorem and its applications. Density, spectra and asymptotic distribution of random matrices in multivariate statistics (Wishart-, Wigner-matrices). How to use separation theorems for eigenvalues and singular values in the principal component, factor, and correspondence analysis. Factor analysis as low rank representation, relations between representations and metric clustering algorithms. Methods of classification: discriminatory analysis, hierarchical, k-means, and graph theoretical methods of cluster analysis. Spectra and testable parameters of graphs.

Algorithmic models, statistical learning. EM algorithm, ACE algorithm, Kaplan–Meier estimates. Resampling methods: bootstrap and jackknife. Applications in data mining, randomized methods for large matrices. Mastering the multivariate statistical methods and their nomenclature by means of a program package (SPSS or S+), application oriented interpretation of the output data.

References:

Bolla, M., Krámli, A.: Theory of statistical inference (in Hungarian), Typotex, Budapest, 2005

Mardia, K. V.; Kent, J. T., Bibby, J. M.: Multivariate Analysis, Academic Press, Elsevier Science, 1979, 2003