

Statisztika és információelmélet

3/1/0/f/5

Tárgyfelelős: Bolla Marianna
További oktatók: Györfi László

Becslések és hipotézisvizsgálat többdimenziós paramétertérben: Fisher-információs-mátrix, likelihood-hányados-próba. Hipotézisvizsgálat többdimenziós Gauss-modellben: Mahalanobis-távolság, Wishart-, Hotelling-, Wilks-eloszlások. Lineáris becslések, Gauss–Markov-tétel. Regresszióanalízis, egy- és többszemponos varianciaanalízis, mint lineáris modell. ANOVA-táblázatok, Fisher–Cochran-tétel. Főkomponens- és faktoranalízis. Faktorok becslése és forgatása, hipotézisvizsgálatok a faktorok számára.

Hipotézisvizsgálat és I-divergencia (diszkrét eset).

I-vetületek, exponenciális eloszláscsalád esetén a maximum likelihood becslés, mint I-vetület. A megfelelő I-divergencia-statisztika határeloszlása. Kontingenciátáblázatok analízise információelméleti módszerrel, loglineáris modellek. Információelméleti alapú statisztikai algoritmusok: iteratív arányos illesztés, EM-algoritmus. Maximális entrópia módszere.

Irodalom:

M. Bolla, A. Krámlí: Statisztikai következtetések elmélete, Typotex, Budapest, 2005
I. Csiszár, P. C. Shields: Információelmélet és statisztika. Oktatási segédanyag (angolul). Alapok és trendek a kommunikáció- és információelméletben c. kiadványnak 420-525. oldala, Now Publ. Inc., Hollandia, 2004. (Szintén elérhető a Rényi Intézet www.renyi.hu honlapján, Csiszár Imre oktatási segédanyagainál.)

Statistics and information theory

3/1/0/f/5

Course coordinator: Marianna Bolla
Other instructors: László Györfi

Multivariate statistical inference in multidimensional parameter spaces: Fisher's information matrix, likelihood ratio test. Testing hypotheses in multivariate Gauss model: Mahalanobis' distance, Wishart's, Hotelling's, Wilks' distributions. Linear statistical inference, Gauss–Markov theorem. Regression analysis, one- and two-way analysis of variance as a special case of the linear model. ANOVA tables, Fisher–Cochran theorem. Principal component and factor analysis. Estimation and rotation of factors, testing hypotheses for the effective number of factors.

Hypothesis testing and I-divergence (the discrete case). I-projections, maximum likelihood estimate as I-projection in exponential families. The limit distribution of the I-divergence statistic. Analysis of contingency tables by information theoretical methods, loglinear models. Statistical algorithms based on information geometry: iterative scaling, EM algorithm. Method of maximum entropy.

References:

Bolla, M., Krámlí, A.: Theory of statistical inference (in Hungarian), Typotex, Budapest, 2005
Csiszár, I., Shields, P. C.: Information Theory and Statistics. A tutorial. In: Foundations and Trends in Communications and Information Theory, 420-525. Now Publ. Inc., The Netherlands, 2004
