Topics in statistics III/IV (MATH3361/4071) Michaelmas term, 2018-2019

Description of the course

Lecturer: Georgios Karagiannis georgios.karagiannis@durham.ac.uk

The following details concern Michaelmas term only. The description below is informal, but more details, and aims to help the student organize his/her study. The formal description of the course can be found in the Faculty handbook here for the Level III, and here for the Level IV.

1 Description of the course

The course discusses two main topics in the frequentist Statistics framework.

Statistical inference: We will study methods and tools for likelihood inference. In particular we will study Maximum Likelihood (ML) based methods/tools used to perform statistical inference. Also, we will study the method of moments for estimation as well as its connections with the ML inference. Our focus will be on the theory, methods, and implementation in basic models and contingency tables.

Contingency tables analysis: We will study the categorical data analysis with special focus on contingency tables and log-linear models.

Intended learning outcomes

The students will be able to

[ILO1] Explain, extend, and apply the method of moments for statistical inference, its asymptotic behavior (with respect to the data-set size), and association to the ML inferential tools.

[ILO2] Explain and extend the asymptotic behavior of the ML based inference with respect to the data-set size, as well as the conditions required, in a mathematically rigorous matter.

[ILO3] Identify and explain problems with multiple parameters, and nuisance parameters

[ILO4] Explain, apply, and extend frequentist inference tools, such as ML (and related) tools, in problems with multiple unknown parameters, and nuisance parameters

[ILO5] Explain and design contingency tables, as well as identify problems that can be analyzed with these concepts

[ILO6] Explain and apply ML based inference to analyze problems with consistency tables and log-linear models. In particular, assumptions, development of tools, implementation, and output.

[ILO7] Identify, explain, and use R packages for the implementation of ML inference on log-linear models to analyze contingency tables problems

in frequentist statistics framework

Teaching and learning activities

[TLA1] Lectures

Students will be introduced to the theory, and be exposed to a small number of examples. For further details will be directed to the learning material which is available from the library.

- Major focus [ILOs 1-6]

[TLA2] Problems Class

Students will learn how to implement the theory and methods introduced in Lectures. Examples will be given, and exercises will be solved in the blackboard.

- Major focus [ILOs 1-6]

[TLA3] Practical classes

Students will be instructed on how to use R software and packages to analyze applications with contingency tables based on the theory introduced in Lectures, and the implementation exercised in Problems Classes.

- Major focus [ILO 6, 7]

[TLA4] Office hours at 12.00-13.00 & 14.00-15.00 on Fridays, at Office CM126b.

Students will ask further questions. Students should have itemize their questions to be discussed in a bullet list on a piece of paper.

- Major focus [ILO 1-7]

Assessment tasks / activities

Formative assessment

[FA1] Four homework assignments will be assigned.

Homework problems and solutions will be available from DUO (Tab: Formative assessment).

Major focus [ILOs 1-6].

Summative assessment

[SA1] See here for the Level III, and here for the Level IV.

Major focus [ILOs 1-7] regarding the Michaelmas term.

2 Syllabus

Revision & catch up:

Calculus:

- Mean value theorem; Taylor series; Lagrange multipliers

Probability:

- Modes of convergence, and their properties ; characteristic functions ; central limit theorems ; Edworth expansions ; Mann-Wald notation ; laws of large numbers ; uniform strong law of large numbers ; Slutsky theorems

Method of moments:

- asymptotic distribution ; asymptotic efficiency ; strong consistency

Improving estimators:

- One step estimation: Fisher scoring; Newton-Raphson alg. ; asymptotic efficiency

Likelihood methods:

Estimation:

- strong consistency of MLE ; asymptotic normality of MLE ; asymptotic distribution of MLE functions ; Cramer-Rao lower Bound ; asymptotic efficiency ; Delta method ; variance stabilising transformations

Inferential tools:

- asymptotic distribution of likelihood ratio (ML) ratio ; Wald statistic ; Score statistic
- asymptotic hypothesis tests and confidence intervals based on the above

Information criteria:

- Akaike information criterion ; Bayesian information criterion

Nuisance parameters

- profile likelihood methods ; profile likelihood ratio test

Contingency tables:

- $I \times J$, and $I \times J \times K$ contigency tables:
  - design and description of contingency tables; sampling schemes (multinomial, product multinomial, Poisson) ; derivation of MLE ; Goodness of fit tests ; analysis of residuals ; Odds ratio and their asymptotic distributions ; Fisher's exact test ; Mantel-Haenzsel test ; types of independency.
- log-linear models:
  - log-linear equation ; modeling contingency $I \times J \times K$ tables, and higher order ones ; hierarchical model parametrization ; identification constraints ; model selection (deviance or likelihood ratio, Pearson χ^2, AIC, BIC, stepwise algorithm, forward, backward elimination) ; MLE (Poisson and Multinomial sampling scheme); Newton-Raphson algorithm; Iterative proportion fitting method.