Statistical Theory (8 CFU, 48 hours)

Instructors: Prof. Marco Alfò, Dott.ssa Monia Ranalli

Lecture: Day… Hour…

Office hours: by appointment (marco.alfo@uniroma1.it, monia.ranalli@uniroma1.it)

Course objectives

The course is addressed to first year Ph.D. students in the Department of Statistics. It is assumed that students have taken statistics at an advanced undergraduate level and are familiar with mathematical ideas such as proofs.

We will cover a selection of topics in statistical theory, with a focus on extensions of classical inference to mis-specified models, quasi-likelihood and composite likelihood. One of the main goals is to provide students with some background to understand the ongoing statistical inference, which focus on dealing with non-standard cases.

Tentative list of topics

Module I (Marco Alfò)

- Basic statistical inference: likelihood functions, derived quantities (score, observed information matrix and expected information matrix)
- ML estimation: properties of MLE (consistency and asymptotic properties) under well-specified and mis-specified model
- Limiting Distributions and asymptotic tests
- Nuisance Parameters and profile likelihood.

Module II (Monia Ranalli)

- Quasi likelihood and genaralized estimating equations: mean response function, variance function, link function and dispersion parameter. Unbiased estimating equation and Bartlett identities. Asymptotic covariance matrix (Godambe Information matrix).
- Pseudo likelihood: definition and estimation.
- Composite likelihood: definition, derived quantities (composite score, sensitivity matrix, variability matrix, Godambe Information)
- Composite likelihood estimation: properties of CLE, asymptotic theory (limiting distribution) composite likelihood ratio test, nuisance parameters.
- Composite likelihood vs. weighted composite likelihood (asymptotic efficiency).
- Composite EM algorithm and composite model selection.

Main textbook

- Cox, D.R. and Hinkley, D.V. (1974) Theoretical Statistics

Additional References

Likelihood Basics

- Cox, D.R. (2006) Principles of Statistical Inference (Cox)

Misspecified Models

- Kent, J. T. (1982). Robust properties of likelihood ratio tests. Biometrika 69, 19-27
- White, H. (1982) Maximum likelihood estimation of misspecified models. Econometrica 50, 1--25.
- Liang, K.-Y. and Zeger, S.L. (1986) Longitudianl data analysis using generalized linear models. Biometrika 75, 13--22.
- Gouerieroux, C. et al. (1993) Indirect inference. J. Appl. Econometrics 8 S85-S118
- Smith, A.A. (2008). Indirect inference. in New Palgrave Dictionary of Economics

Quasi-likelihood

- Diggle, P.J., Heagerty, P., Liang K-Y., and Zeger, S.L. (2002). Analysis of Longitudinal Data (second edition). Oxford.
- Firth, D. (1993). Recent developments in quasi-likelihood methods. Bulletin International Statistical Institute 55, 341-358.
- Halekoh, U., Hojsgaard, S., and Yuan, J. (2006). The R package geepack for generalized estimating equations. Journal of Statistical Software 15 (2).
- Liang, K.-Y. and Zeger, S.L. (1986). Longitudinal data analysis using generalized linear models. Biometrika 73, 13-22.
- Wedderburn, R.W.M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. Biometrika 61, 439-447.

Pseudo likelihood

- Besag, J. (1975), "Statistical Analysis of Non-Lattice Data", The Statistician, 24 (3): 179–195

Composite likelihood

- Lindsay, B.G. (1988). Composite likelihood methods. Contemporary Mathematics 80, 221-239.
- Cox, D. and Reid, N. (2004). A note on pseudolikelihood constructed from marginal densities. Biometrika 91, 729-737.
- Lindsay, B. (1988). Composite likelihood methods.Contemporary Mathematics 80, 220-239.
- Molenberghs, G. and Verbeke, G. (2005). Models for Discrete Longitudinal Data. Springer
- Varin, C., Reid, N., and Firth, D. (2011). An overview of composite likelihood methods. Statistica Sinica 21, 5-42.

Final Exam