PhD Course in Statistical Sciences
cycle XXXV, 2020

Theory and Methods of Inference

Instructors: **Alessandra Salvan, Nicola Sartori, Luigi Pace**

Syllabus

Some prerequisites. Empirical distribution function. Convergence of sums of r.v.'s. Order statistics. Density functions. Scale and location families. Exponential families. Multivariate normal distributions. Parametric inference: basics.

Statistical models and uncertainty in inference. Statistical models. Paradigms of inference: the Bayesian and frequentist paradigms. Prior specification. Model specification (data variability). Levels of model specification. Problems of distribution (variability of statistics). Simulation. Asymptotic approximations and delta method.

Generating functions, moment approximations, transformations. Moments, cumulants and their generating functions. Generating functions of sums of independent random variables. Edgeworth and Cornish-Fisher expansions. Notations $O_p(\cdot)$ and $o_p(\cdot)$. Approximations of moments and transformations. Laplace approximation.

Likelihood: observed and expected quantities, exact properties. Dominated statistical models. Sufficiency. Likelihood: observed quantities. Examples: a two-parameter model, grouped data, censored data, sequential sampling, Markov chains, Poisson processes. Likelihood and sufficiency. Invariance properties. Expected likelihood quantities and exact sampling properties. Reparameterizations.

Likelihood inference: first-order asymptotics. Likelihood inference procedures. Consistency of the maximum likelihood estimator. Asymptotic distribution of the maximum likelihood estimator. Asymptotic distribution of the log-likelihood ratio: simple null hypothesis, likelihood confidence regions, comparisons among asymptotically equivalent forms, non-null asymptotic distributions, composite null hypothesis (nuisance parameters), profile likelihood, asymptotically equivalent forms and one-sided versions, testing constraints on the components of the parameter. Non-regular models.

Bayesian Inference. Noninformative priors. Inference based on the posterior distribution. Point estimation and credibility regions. Hypothesis testing and the Bayes factor. Linear models.

Likelihood and Bayesian inference in R. A scalar parameter example: log likelihood, plot of the log likelihood, MLE and observed/expected information, Wald confidence intervals, deviance confidence regions, simulation, numerical optimization methods, significance function. A vector parameter example: plot of the log likelihood, parameter estimates, simulation. Parameter of interest and profile likelihood. Examples in the Weibull model. Deviance intervals: simulation. Stratified models. EM algorithm with applications to censored data and mixture models. Bayesian inference: posterior summaries, simulation from the posterior (rejection sampling).

Estimating equations and pseudolikelihoods. Misspecification. Estimating equations. Quasi likelihood. Pairwise likelihood. Empirical likelihood.

Data and model reduction by marginalization and conditioning. Distribution constant statistics. Completeness. Ancillary statistics. Data and model reduction with nuisance parameters: lack of information with nuisance parameters, pseudo-likelihoods. Marginal likelihood. Conditional likelihood.

Decision paradigms.. Statistical decision problems and Bayes decision rules. Efficient estimators: Cramér-Rao lower bound, asymptotic efficiency, Godambe efficiency, Rao-Blackwell-Lehmann-Scheffé theorem, other constraints for point estimation. Optimal tests: Neyman-Pearson lemma, composite hypotheses: families with monotone likelihood ratio, locally most powerful tests, two-sided alternatives, other constraint criteria. Optimal confidence regions.

Exponential families, Exponential dispersion families, Generalized linear models. Exponential families of order 1. Mean value mapping and variance function. Multiparameter exponential families. Marginal and conditional distributions. Sufficiency and completeness. Likelihood and exponential families: likelihood quantities, conditional likelihood, profile likelihood and mixed parameterization. Procedures with finite sample optimality properties. First-order asymptotic theory. Exponential dispersion families. Generalized linear models.

Group families. Groups of transformations. Orbits and maximal invariants. Simple group families and conditional inference. Composite group families and marginal inference.

References

Barndorff-Nielsen, O.E. and Cox, D.R. (1994). *Inference and Asymptotics.* Chapman and Hall, London.

Bickel, P. and Doksum, K. (2015). *Mathematical Statistics: Basic Ideas and Selected Topics*, Volumes I (Second Edition) and II. Chapman & Hall/CRC Texts in Statistical Science, London.

Casella, G. and Berger, R.L. (1990, 2002). *Statistical Inference.* Wadsworth publishing Co., Belmont, CA.

Cox, D.R. (2006). *Principles of Statistical Inference.* Cambridge University Press, Cambridge.

Cox, D.R. and Hinkley, D.V. (1974). *Theoretical Statistics.* Chapman and Hall, London.

Davison, A.C. (2003). *Statistical Models.* Cambridge University Press, Cambridge.

Efron, B. and Hastie, T. (2016). *Computer Age Statistical Inference: Algorithms, Evidence and Data Science.* Cambridge University Press, Cambridge.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. and Rubin, D.B. (2014). *Bayesian Data Analysis* (Third edition). Chapman & Hall/CRC Texts in Statistical Science, London.

Hoff, P. (2009). *A First Course in Bayesian Statistical Methods*, Springer, New York.

Lehmann, E.L. (1983). *Theory of Point Estimation.* Wiley, New York.

Lehmann, E.L. (1986). *Testing Statistical Hypotheses.* 2-nd ed.. Wiley, New York.

Liseo, B. (2007). *Introduzione alla Statistica Bayesiana.*

McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models.* 2-nd ed.. Chapman and Hall, London.

O'Hagan, A. and Forster, J. (2004). *Bayesian Inference.* 2-nd ed.. Edward Arnold, London.

Pace, L. and Salvan, A. (1996). *Teoria della Statistica: Metodi, Modelli, Approssimazioni Asintotiche.* Cedam, Padova.

Pace, L. and Salvan, A. (1997). *Principles of Statistical Inference from a neo-Fisherian Perspective.* World Scientific, Singapore.

Robert C.P. and Casella G. (2004). *Monte Carlo Statistical Methods* (Second edition). Springer, New York

Severini, T.A. (2000). *Likelihood Methods in Statistics.* Oxford University Press, Oxford.

Severini, T.A. (2005). *Elements of Distribution Theory.* Cambridge University Press, Cambridge.

van der Vaart, A.W. (1998). *Asymptotic Statistics.* Cambridge University Press, Cambridge.

Young, G.A. and Smith, R.L. (2005). *Essentials of Statistical Inference.* Cambridge University Press, Cambridge.

Lectures and topics (tentative schedule)

date		topic	instructor
25/02/20		basic statistics prerequisites	AS
03/03/20		course presentation, some prerequisites	AS
05/03/20		exercises	AS
10/03/20		statistical models and uncertainty in inference: Bayesian models and frequentist inference, model specification and problems of distribution	AS
12/03/20		generating functions, approximation of moments, transformations	AS
17/03/20		likelihood: observed and expected quantities, exact sampling properties	AS
19/03/20		likelihood: observed and expected quantities, exact sampling properties	AS
24/03/20		likelihood inference: first-order asymptotics	AS
26/03/20		likelihood inference: first-order asymptotics	AS
31/03/20		Bayesian inference: noninformative priors, point and interval estimation	AS
02/04/20		Bayesian inference: hypothesis testing and the Bayes factor; normal linear model	AS
07/04/20		students' seminars and exercises	NS
09/04/20		estimating equations and pseudolikelihoods	NS
16/04/20		likelihood and Bayesian inference in `R`	NS
21/04/20		likelihood and Bayesian inference in `R`	NS
23/04/20		data and model reduction by marginalization and conditioning	AS
28/04/20		data and model reduction by marginalization and conditioning	NS
30/04/20		decision paradigms	AS
05/05/20		decision paradigms	AS
07/05/20		exponential families	AS
08/05/20		exponential families	AS
12/05/20		exponential dispersion models, generalized linear models	AS
14/05/20		group families	AS
15/05/20		group families	AS
21/05/20		writing papers and reports + paper assignment	AS
??/06/20	3 hours	written exam	
??/06/20	3 hours	practical exam	
??/07/20		oral presentations	

Instructors: AS is Alessandra Salvan, NS Nicola Sartori.

AS, NS and Luigi Pace (LP) collaborate in lecture material, in the final exam and in paper assignment and evaluation.