PhD School in Statistics
XXV cycle, 2010

Theory and Methods of Statistical Inference

PART I — Frequentist likelihood methods

(A. Salvan, N. Sartori, L. Pace)

Syllabus

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

Statistical models: data variability and uncertainty in inference: Statistical models. Paradigms of inference. The Fisherian paradigm. Model specification (data variability). Levels of specification. Problems of distribution (statistical variability). 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.

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 condence regions, comparisons among asymptotically equivalent forms, non-null asymptotic distributions, composite null hypothesis (nuisance parameters), prole likelihood, asymptotically equivalent forms and one-sided versions, testing constraints on the components of the parameter. Non-regular models.

Likelihood: numerical and graphical aspects in R – Scalar and vector parameters: 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.

Likelihood: numerical and graphical issues in R – Parameter of interest and profile likelihood: Parameter of interest and profile likelihood. Examples in the Weibull model. Deviance intervals: simulation. Stratified models.

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.

The frequency-decision paradigm: Statistical decision problems. Efficient estimators: Cramer-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.

Higher order asymptotics: Laplace expansions. Approximation of marginal likelihood. Bayesian interpretation of approximate conditional likelihood. Saddlepoint expansion for density functions. Improvements over classical asymptotic theory. The p^* formula. Tail probabilities and modified directed likelihood. Bartlett adjustment. Modified profile likelihood.

References

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

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

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

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

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.

Wasserman, L. (2004). *All of Statistics.* Springer, New York.

Welsh, A.H. (1996). *Aspects of Statistical Inference.* Wiley, New York.

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

Lectures and topics

date		**topic**	instructor
26/02/10	9.30–11.30	course presentation, some prerequisites	AS
02/03/10	9.30–11.30	statistical models: data variability and sampling variability	AS
05/03/10	9.30–11.00	*Sweave* =`R x LaTeX`^{2}	NS
05/03/10	11.30–13.00	statistical models: data variability and sampling variability	AS
09/03/10	9.30–11.30	generating functions, approximation of moments, transformations	AS
12/03/10	9.30–11.30	likelihood: observed and expected quantities, exact sampling properties	AS
16/03/10	9.30–11.30	likelihood: observed and expected quantities, exact sampling properties	AS
17/03/10	9.30–11.30	likelihood inference (first-order asymptotics)	NS
19/03/10	9.30–11.30	likelihood inference (first-order asymptotics)	NS
30/03/10	9.30–11.30	review-exercises	AS
12/04/10	9.30–11.30	review-exercises	NS
13/04/10	9.30–11.30	lab. on "likelihood: graphical and numerical techniques with `R`, I "	NS
16/04/10	11.00–13.00	estimating equations and pseudolikelihoods	NS
21/04/10	9.30–11.30	data and model reduction by marginalization and conditioning	AS
27/04/10	9.30–11.30	the frequency decision paradigm	AS
30/04/10	9.30–11.30	review-exercises	AS, NS
04/05/10	9.30–11.30	lab. on "likelihood: graphical and numerical techniques with `R`, II "	NS
07/05/10	9.30–11.30	exponential families	AS
11/05/10	9.30–11.30	exponential dispersion models, generalized linear models	AS
13/05/10	11.30–13.30	review-exercises	AS
14/05/10	9.30–11.30	exponential dispersion models, generalized linear models	AS
18/05/10	9.30–11.30	group families	AS
19/05/10	9.30–11.30	group families	AS
21/05/10	9.30–11.30	higher-order likelihood theory	NS
25/05/10	9.30–11.30	higher-order likelihood theory	NS
01/06/10	10.00–13.00	written exam	AS, NS
10/06/10	14.00–15.00	writing papers or reports + paper assignment	LP,AS, NS
07/09/10	14.00–18.00	final seminars	LP,AS, NS

Instructors: AS is Alessandra Salvan, NS is Nicola Sartori.
Luigi Pace (LP) collaborates in lecture material for the first part, in the final exam and in paper assignment and evaluation.

PART II — Bayesian Inference

(L. Ventura, B. Liseo, A. Tancredi, E. Moreno)

Syllabus

Statistical models and prior information. Inference based on the posterior distribution. Choice of the prior distribution. Point and interval estimation.

Hypothesis testing and the Bayes factor.

Linear models. Multivariate normal model. Generalized linear models. Hierarchical models. Model selection.

Posterior simulation and computational methods for Bayesian inference.

Markov Chain Monte Carlo methods for Bayesian inference.

References

Albert, J. (2009). *Bayesian Computation with R.* (Second edition). Springer, New York.

Ghosh, J.K., Delampady, M., Tapas, S. (20)6). *An Introduction to Bayesian Analysis.* Springer, New York.

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

Lee, P. (2004). *Bayesian Statistics: an Introduction.* Oxford University Press, New York.

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

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

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

Lectures and topics

date		topic	instructor	location
23/03/10	9.30–11.30	Introduction. Prior and posterior distributions.	LV	PD
23/03/10	14.30–16.30	Point and interval estimation. Hypothesis testing.	LV	PD
25/03/10	15.00–18.00	Choice of prior distributions. Noninformative priors.	BL	BO
26/03/10	9.00–13.00	Linear models. Posterior simulation.	BL	BO
08/04/10	9.30–17.00	Markov Chain Monte Carlo methods. GLM case studies.	AT	BO
09/04/10	10.30–13.00	Markov Chain Monte Carlo methods. GLM case studies.	AT	BO
20/04/10	11.00–15.30	Multivariate Normal model and Hierarchical models - Model selection	BL	BO
23/04/10	10.30–16.30	Objective Bayesian model selection statistical procedures.	EM	PD

Lectures in Bologna will be part of the PhD course *Introduction to Bayesian Statistics* organized by the Bologna PhD in Statistics, PhD School in Economic and Statistical Sciences (course organizer Prof. Daniela Cocchi).

Instructors: LV is Laura Ventura, BL is Brunero Liseo, AT is Andrea Tancredi, EM is Elias Moreno.