

STATISTICS 592B
SEMIPARAMETRIC MODELS

Winter Quarter 1998

POTENTIAL TOPICS

1. Profile Likelihood.

- Severini, T. A. and Wong, W. H. (1992). Profile likelihood and conditionally parametric models. *Ann. Statist.* **20**, 1768-1862.
- Qin, J. and Wong, A. (1996). Empirical likelihood in a semiparametric model. *Scand. J. Statist.* **23**, 209-220.
- Murphy, S. and Van der Vaart, A. W. (1997). On profile likelihood. *Preprint*, Pennsylvania State University. [Available at: http://www.stat.psu.edu/murphy/research_th.html].
- Huang, J. and Wellner, J. A. (1996). Regression models with interval censoring (with J. Huang). (1996). In *Probability theory and mathematical statistics*, 269 - 296. I. A. Ibragimov and A. Yu. Zaitsev, editors. Gordon and Breach, Amsterdam. [Partial review of early work on profile likelihood.]

2. Testing in and between semiparametric models.

- Murphy, S. A. and Van der Vaart, A.W. (1995). Semiparametric likelihood ratio inference. *Ann. Statist.* **25**, 1471 - 1509.
- Choi, S., Hall, W. J., and Schick, A. (1996). Asymptotically uniformly most powerful tests in parametric and semiparametric models. *Ann. Statist.* **24**, 841 - 861.
- Cox, D., Kohn, E., Wahba, G., Yandell, B. S. (1988). Testing the (parametric) null model hypothesis in (semiparametric) partial and generalized spline models. *Ann. of Statistics* **16**, 113-119.
- Rotnitzky, A. and Jewell, N. P. (1990). Hypothesis testing of regression parameters in semiparametric generalized linear models for cluster correlated data. *Biometrika* **77**, 485-497.

3. Missing Data Models.

- BKRW: Chs. 4.6 & 6.6.
- Robins, J. M., Rotnitzky, A., and Zhao, L.P. (1994). Estimation of regression coefficients when some regressors are not always observed. *J. Amer. Statist. Assoc.* **89**, 846 - 866.
- Rotnitzky, A. and Robins, J. (1995). Semi-parametric estimation of models for means and covariances in the presence of missing data. *Scand. J. Statist.* **22**, 323 - 333.

- Robins, J. M., Hsieh, F., and Newey, W. (1995). Semiparametric efficient estimation of a conditional density with missing or mismeasured covariates. *J. R. Statist. Soc. B* **57**, 409 - 424.

4. Regression Models for Survival Data: alternatives to the Cox model.

- Lin, D. Y. & Ying, Z. (1994). Semiparametric analysis of the additive risk model. *Biometrika*, **81**, 61–71.
- Lin, D.Y. & Ying, Z. (1996). Semiparametric analysis of general additive-multiplicative intensity models for counting processes. *Ann. Statist.*, **23**, 1712–1734.
- Sasieni, P. (1992). Non-orthogonal projections and their application to calculating the information in a partly linear Cox model. *Scand. J. Statist.* **19**, 215 - 234.
- Huang, J. (1997). Efficient estimation in the partly linear additive Cox model. *Technical Report*, University of Iowa.

5. Attainment and Nonattainment of Information Bounds.

- Van der Vaart, A. W. (1995). Efficiency of infinite dimensional M -estimators. *Statistica Neerl.* **49**, 9 - 30.
- Ritov, Y. and Bickel, P. J. (1990). Achieving information bounds in non and semiparametric models. *Ann. Statist.* **18**, 925 - 938.
- Ibragimov, I. A. and Khas'minskii, R. Z. (1991). Asymptotically normal families of distributions and efficient estimation. *Ann. Statist.* **19**, 1681 - 1724.
- Birgé, L. and Massart, P. (1995). Estimation of integral functionals of a density. *Ann. Statist.* **23**, 11 - 29.
- Laurent, B. (1996). Efficient estimation of integral functionals of a density. *Ann. Statist.* **24**, 659 - 681.
- Efromovich, S. and Low, M. G. (1996). On Bickel and Ritov's conjecture about adaptive estimation of the intergral of the square of density derivative. *Ann. Statist.* **24**, 682 - 686.
- Laurent, B. (1997). Estimation of integral functionals of a density and its derivatives. *Bernoulli* **3**, 181-211.

6. Lower bounds for “inverse” problems and nonstandard asymptotics.

- Groeneboom and Wellner (1992), Part I, Chapter 2.
- Groeneboom, P. and Jongbloed, G. (1995). Isotonic estimation and rates of convergence in Wicksell's problem. *Ann. Statist.* **23**, 1518 - 1542.
- Groeneboom, P. (1996). Lectures on inverse problems. In *Lectures on Probability Theory and Mathematical Statistics, Ecole d'Ete de Probabilites de St. Flour XXIV-1994, Lecture Notes in Mathematics* **1648**, 67 - 164.

- Jongbloed, G. (1995). *Three Statistical Inverse Problems*. Ph. D. dissertation, Delft University.
- Donoho, D. L. and Low, M. G. (1990). Renormalization exponents and optimal pointwise rates of convergence. *Ann. Statist.* **20**, 944 - 970.
- Huber, C. (1997). Lower bounds for function estimation. In *Festschrift for Lucien Le Cam*, pp. 245 - 258. D. Pollard, E. Torgersen, and G. L. Yang, editors. Springer-Verlag, New York.
- Yu, B. (1997). Assouad, Fano, and Le Cam. In *Festschrift for Lucien Le Cam*, pp. 245 - 258. D. Pollard, E. Torgersen, and G. L. Yang, editors. Springer-Verlag, New York.

7. Estimation for binary choice and interval censoring models.

- Cosslett, S.R. (1983). Distribution-free maximum likelihood estimator of the binary choice model. *Econometrica* **55**, 559 - 585.
- Huang, J. and Wellner, J. (1997). Interval censored survival data: a review of recent progress. In *Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis, Lecture Notes in Statistics* **123**, 123 - 169.
- Groeneboom, P. (1996). Lectures on inverse problems. In *Lectures on Probability Theory and Mathematical Statistics, Ecole d'Ete de Probabilites de St. Flour XXIV-1994, Lecture Notes in Mathematics* **1648**, 67 - 164.
- Lin, D. Y., Oakes, D. & Ying, Z. (1997). Additive hazards regression with current status data. *Biometrika*, in press.

8. Information bounds for multiphase and case-cohort designs.

- Breslow, N. E. and Holubkov, R. (1997) Maximum likelihood estimation of logistic regression parameters under two-phase, outcome-dependent sampling. *J. R. Statist. Soc. B* **59**, 447 - 461.
- Lawless, J.F., Wild, C. J., and Kalbfleisch, J. D. (1997). Semiparametric methods for response-selective and missing data problems in regression. *Preprint*; submitted to JRSS?
- Scott, A. J. and Wild, C. J. (1991). Fitting logistic models in stratified case-control studies. *Biometrics* **47**, 497-510.
- Scott, A. J. and Wild, C. J. (1997). Fitting regression models to case-control data by maximum likelihood. *Biometrika* **84**, 57 - 84.
- Prentice, R. L. (1986). A case-cohort design for epidemiologic cohort studies and disease prevention trials. *Biometrika*, **73**, 1-11.
- Self, S. G. & Prentice, R. L. (1988). Asymptotic distribution theory and efficiency results for case-cohort studies. *Ann. Statist.*, **16**, 64-81.
- Goldstein, L. & Langholz, B. (1992). Asymptotic theory for nested case-control sampling in the Cox regression model. *Ann. Statist.*, **20**, 1903-1928.

9. Information Bounds for Models with Dependence.

- Greenwood, P. E. and Wefelmeyer, W. (1995). Efficiency of empirical estimators for Markov chains. *Ann. Statist.* **23**, 132-143.
- Ibragimov, I. A. and Khas'minskii (1975). Local asymptotic normality for non-identically distributed observations. *Theor. Prob. Applic.* **20**, 246 - 260.
- Jeganathan, P. (1981). On a decomposition of the limit distribution of a sequence estimators. *Sankhya* **43**, 26 - 36.
- Jeganathan, P. (1982). On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal *Sankhya*, **44**, 173 - 212.
- Jeganathan, P. (1995). Some aspects of asymptotic theory with applications to time series models. *Econometric Theory* **11**, 818 - 887.
- Schick, A. On estimation in LAMN families when there are nuisance parameters present *Sankhya* , **50**, 249-268.

10. Semiparametric generalized linear models.

- Chen, Hung. (1995). Asymptotically efficient estimation in semiparametric generalized linear models. *Ann. Statist.* **23**, 1102 - 1129.
- Mammen, E. and Van de Geer, S. (1997). Penalized quasi-likelihood estimation in partial linear models. *Ann. Statist.* **25**, 1014-1035.
- Emond, M. J. and Self, S. (1993). An efficient estimator for the generalized semilinear model. *J. Amer. Statist. Assoc.* **92**, 1033-1040.

11. General M - and Z - theorems.

- Huber, P. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1**, 221-233, University of California Press, Berkeley.
- Pollard, D. (1985). New ways to prove central limit theorems. *Econometric Theory* **1**, 295 - 314.
- Pakes, A. and Pollard, D. (1989). Simulation and the asymptotics of optimization estimators. *Econometrica* **57**, 1027 - 1057.
- Gill, R. D. (1989). Non- and semi-parametric maximum likelihood estimators and the von-Mises method - I. *Scan. J. Statist.* **16**, 97 - 128.
- Gill, R. D. and Van der Vaart, A. W. (1993). Non- and semi-parametric maximum likelihood estimators and the von-Mises method - II. *Scan. J. Statist.* **20**, 271 - 288.
- Wong, W. H. and Severini, T. A. (1991). On maximum likelihood estimation in infinite-dimensional parameter spaces. *Ann. Statist.* **19**, 603 - 632.
- Van der Vaart, A. W. (1995). Efficiency of infinite dimensional M-estimators. *Statistica Neerl.* **49**, 9 - 30.

- Huang, J. (1996). Efficient estimation for the proportional hazards model with interval censoring. *Ann. Statist.*, **24**, 540–568.
- Newey, W. (1994). The asymptotic variance of semiparametric estimators. *Econometrica* **62**, 1349 - 1382.
- Hu, Huilin. (1997). Large sample theory for pseudo-maximum likelihood estimation. *Unpublished Ph.D. dissertation*, University of Washington.

12. Empirical Process Methods in Estimation of General Parameters

- Birgé, L. and Massart, P. (1993). Rates of convergence for minimum contrast estimators. *Probability Theory and Related Fields* **97**, 113 - 150.
- Birgé, L. and Massart, P. (1997). From model selection to adaptive estimation. In *Festschrift for Lucien Le Cam*, pp. 55 - 87. D. Pollard, E. Torgersen, and G. L. Yang, editors. Springer-Verlag, New York.
- Shen, X., and Wong, W. H. (1994). Convergence rate of sieve estimates. *Ann. Statist.* **22**, 580–615.
- Wong, W. H., and Shen, X. (1995). Probability inequalities for likelihood ratios and convergence rates of sieve MLE's. *Ann. Statist.* **23**, 339 - 363.
- Van de Geer, S. (1993). Hellinger-consistency of certain nonparametric maximum likelihood estimators. *Ann. Statist.* **21**, 14 - 44.
- Van de Geer, S. (1996). Rates of convergence for the maximum likelihood estimator in mixture models. *Nonparametric Statistics* **6**, 293 - 310.

13. Mixture Models.

- BKRW, sections 4.5 and 6.5.
- Van der Vaart, A. W. (1995). Efficient maximum likelihood estimation in semi-parametric mixture models. *Ann. Statist.* **24**, 862-878.
- Roeder, K., Carroll, R. J., and Lindsay, B. G. (1996). A semiparametric mixture approach to case-control studies with errors in covariables. *J. Amer. Statist. Assoc.* **91**, 722-732.