

# Chapter 4 Bibliography

Jon A. Wellner

January 7, 2011

## References

### Section 4.1: Maximum likelihood and efficient likelihood estimation

- Buse, A. (1982). The likelihood ratio, Wald, and Lagrange multiplier tests: an expository note. *The American Statistician* **36**, 153 - 157.
- Efron, B. and Hinkley, D.V. (1978). Assessing the accuracy of the maximum likelihood estimator: Observed versus expected Fisher information (with discussion) *Biometrika* **65**, 457 - 487.
- Engle, R. F. (1984). Wald, Likelihood ratio, and Lagrange multiplier tests in econometrics. *Handbook of Econometrics* **II**, 775 - 826. Elsevier Science Publishers, Amsterdam
- Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philos. Trans. Roy. Soc. London A* **222**, 309-368.
- Fisher, R. A. (1925). Theory of statistical estimation. *Proc. Camb Phil. Soc.* **22**, 700 - 725.
- Rao, C. R. (1947). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proc. Camb Phil. Soc.* **44**, 50 - 57.
- Wald, A. (1943). Test of statistical hypotheses concerning several parameters when the number of observations is large *Trans. Amer. Math. Soc.* **54**, 426 - 482.
- Wilks, S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. Statist.* **9**, 60 - 62.

### Section 4.2: Likelihood ratio, Wald, and Rao (or score) tests

- Moran, P. A. P. (1970). On asymptotically optimal tests of composite hypotheses. *Biometrika* **57**, 47 - 55.
- Neyman, J. (1937). 'Smooth' test for goodness of fit. *Skand. Aktuarietidskrift* **20**, 149 - 199.

### Section 4.3: Examples

- Le Cam, L. (1990b). On the standard asymptotic confidence ellipsoids of Wald. *Int. Statist. Rev.* **58**, 129 - 152.
- Mäkeläinen, T., Schmidt, K., and Styan, G. P. H. (1981). On the existence and uniqueness of the maximum likelihood estimate of a vector-valued parameter in fixed-size samples. *Ann. Statist.* **9**, 758 - 767.
- Neyman, J. and Scott, E. (1948). Consistent estimates based on partially consistent observations. *Econometrica* **16**, 1 - 32.
- Reeds, J. (1985). Asymptotic number of roots of Cauchy location likelihood equations. *Ann. Statist.* **13**, 775 - 784.
- Rotnitzky, A., Cox, D. R. Bottai, M., Robins, J. (2000). Likelihood-based inference with singular information matrix. *Bernoulli* **6**, 243–284.
- SAMPSON, A. R. (1978). Simple BAN estimators of correlations for certain multivariate normal models with known variances. *J. Amer. Statist. Assoc.* **73** 859–862.
- SMALL, C. G., WANG, J. and YANG, Z. (2000). Eliminating multiple root problems in estimation. *Statist. Sci.* **15** 313–341. With comments by John J. Hanfelt, C. C. Heyde and Bing Li, and a rejoinder by the authors.

### Section 4.4: Consistency of Maximum Likelihood Estimators

- Bahadur, R. R. (1958). Examples of inconsistency of maximum likelihood estimates. *Sankhya* **20**, 207 - 210.
- Boyles, R. A., Marshall, A. W., and Proschan, F. (1985). Inconsistency of the maximum likelihood estimator of a distribution having increasing failure rate average. *Ann. Statist.* **13**, 413-417.
- Ferguson, T. S. (1982). An inconsistent maximum likelihood estimate. *J. Amer. Statist. Assoc.* **77**, 831 - 834.
- Van de Geer, S. (1993). Hellinger-consistency of certain nonparametric maximum likelihood estimators. *Ann. Statist.* **21**, 14 - 44.
- Kiefer, J. and Wolfowitz, J. (1956). Consistency of the maximum likelihood estimator in the presence of infinitely many nuisance parameters. *Ann. Math. Statist.* **27**, 887 - 906.
- Le Cam, L. (1990a). Maximum likelihood: an introduction *Int. Statist. Rev.* **58**, 153 - 171.
- Perlman, M. D. (1972). On the strong consistency of approximate maximum likelihood estimators. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **1**, 263 - 281.

- Pfanzagl, J. (1988). Consistency of maximum likelihood estimators for certain nonparametric families, in particular: mixtures. *J. Statist. Planning and Inference* **19**, 137 - 158.
- van der Vaart, A. W. (1996). Efficient maximum likelihood estimation in semiparametric mixture models. *Ann. Statist.* **24**, 862-878.
- Wald, A. (1949). Note on the consistency of maximum likelihood estimates *Ann. Math. Statist.* **20**, 595-601.

#### **Section 4.5: The EM-algorithm and related methods**

- Boyles, R. A. (1983). On the convergence of the EM algorithm. *JRSSB* **45**, 47-50.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *JRSS B* **39**, 1 - 38.
- Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. *J. Roy. Statist. Soc. B* **44**, 226 - 233.
- McLachlan, G. J. and Krishnan, T. (1996). *The EM Algorithm and Extensions*. Wiley, New York.
- Meilijson, I. (1989). A fast improvement of the EM - algorithm on its own terms. *JRSS B* **51**, 127 - 138.
- VARDI, Y., SHEPP, L. A. and KAUFMAN, L. (1985). A statistical model for positron emission tomography. *J. Amer. Statist. Assoc.* **80** 8-37. With discussion.
- Wu, C. F. (1983). On the convergence properties of the EM algorithm. *Ann. Statist.* **11**, 95 - 103.

#### **Section 4.6: Nonparametric maximum likelihood estimation**

- Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Ann. Statist.* **10**, 1100-1120.
- Andersen, P. K., Borgan, O., Gill, R. D., and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag, New York.
- Barlow, R. E., Bartholomew, D. J., Bremner, and Brunk, H. D. (1972). *Statistical Inference under Order Restrictions*. Wiley, New York.
- Bailey, K. R. (1984). Asymptotic equivalence between the Cox estimator and the general ML estimators of regression and survival parameters in the Cox model. *Ann. Statist.* **12**, 730 - 736.
- Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *Ann. Statist.* **2**, 437 - 453.

- Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B* 34, 187-220.
- Fleming, T. R. and Harrington, D. P. (1991). *Counting Processes and Survival Analysis*. Wiley, New York.
- Gill, R. D. (1983). Large sample behavior of the product-limit estimator on the whole line. *Ann. Statist.* **11**, 49 - 58.
- GILL, R. D. (1984). Understanding Cox's regression model: a martingale approach. *J. Amer. Statist. Assoc.* **79** 441-447.
- Gill, R. D., Vardi, Y., and Wellner, J. A. (1988). Large sample theory of empirical distributions in biased sampling models. *Ann. Statist.* **16**, 1069 - 1112.
- Grenander, U. (1956). On the theory of mortality measurement, Part II. *Skand. Akt.* **39**, 125 - 153.
- Groeneboom, P. (1985). Estimating a monotone density. *Proceedings of the Conference in Honor of Jerzy Neyman and Jack Kiefer II*, 539 - 555. Wadsworth, Belmont L. Le Cam and R. A Olshen, editors.
- Groeneboom, P. and Wellner, J. A. (1992). *Information Bounds and Nonparametric Maximum Likelihood Estimation*. DMV Seminar Band 19, Birkhäuser, Basel.
- Jewell, N. P. (1982). Mixtures of exponential distributions. *Ann. Statist.* **10**, 479 - 484.
- Johansen, S. (1978). The product limit estimator as maximum likelihood. *Scand. J. Statist.* **5**, 195 - 199.
- Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53**, 457-581.
- Kiefer, J. and Wolfowitz, J. (1976). Asymptotically minimax estimation of concave and convex distribution functions. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **34**, 73 - 85.
- Kim, J. and Pollard, D. (1990). Cube root asymptotics. *Ann. Statist.* **18**, 191-219.
- Millar, P. W. (1979). Asymptotic minimax theorems for the sample distribution function. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **48**, 233 - 252.
- Robertson, T., Wright, F. T., Dykstra, R. L.(1988). *Order Restricted Statistical Inference*. Wiley, New York.
- Scholz, F. W. (1980). Towards a unified definition of maximum likelihood *Canad. J. Statist.* **8**, 193 - 203
- Vardi, Y. (1985). Empirical distributions in selection bias models *Ann. Statist.* **13**, 178 - 203.

#### Section 4.7: Limit theory for the statistical agnostic: $P \notin \mathcal{P}$

- Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. *Proc. Fifth Berk. Symp. Math. Statist. and Prob.* **1**, 221 - 233. University of California Press, Berkeley.
- PATILEA, V. (2001). Convex models, MLE and misspecification. *Ann. Statist.* **29** 94–123.
- Pollard, D. (1985). New ways to prove central limit theorems. *Econometric Theory* **1**, 295 - 314.
- Pollard, D.(1989). Asymptotics via empirical processes (with discussion). *Statistical Science* **4**, 341-366.
- Pollard, D. (1990). *Empirical Processes: Theory and Applications*. Conference Board of the Mathematical Sciences, Regional Conference Series in Applied Mathematics, Washington, D.C. .
- Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer, New York.
- White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica* **50**, 1 - 25.

#### **Section 4.8: Profile and Composite Likelihood**

- Barndorff-Nielsen, O. E. (1991) Likelihood theory, In *Statistical Theory and Modeling, In honour of Sir David Cox, FRS*, chapter 10, pages 232 - 264. Chapman and Hall, London. D. V. Hinkley, N. Reid, and E. J. Snell, editors.
- Bickel, P. J., Klaassen, C. A. J., Ritov, Y., and Wellner, J. A. (1993). *Efficient and Adaptive Estimation for Semiparametric Models*. Johns Hopkins University Press, Baltimore.
- Kalbfleisch, J. D. and Sprott, D. A. (1970). Application of likelihood methods to models involving large numbers of parameters. *J. Roy. Statist. Soc. B* **32**, 175 - 208.
- Kale, B. K. (1963). Some remarks on a method of maximum-likelihood estimation proposed by Richards. *J. Roy. Statist. Soc. B* **25**, 209 - 212.
- LINDSAY, B. G. (1988). Composite likelihood methods. In *Statistical inference from stochastic processes (Ithaca, NY, 1987)*, vol. 80 of *Contemp. Math.* Amer. Math. Soc., Providence, RI, 221–239.
- Murphy, S. A. and van der Vaart, A. W. (1997). Semiparametric likelihood ratio inference. *Ann. Statist.* **25**, 1471-1509.
- Murphy, S. A., and van der Vaart, A. W. (2000). On profile likelihood. With comments and a rejoinder by the authors. *J. Amer. Statist. Assoc.* **95**, 449-485.
- PARNER, E. T. (2001). A composite likelihood approach to multivariate survival data. *Scand. J. Statist.* **28** 295–302.

- Patefield, W. M. (1977). On the maximized likelihood function. *Sankhyā Ser. B* **39**, 92 - 96.
- Richards, F. S. G. (1961). A method of maximum-likelihood estimation. *J. Roy. Statist. Soc.* **23**, 469 - 475.
- Severini, T. A. and Wong, W. H. (1992). Profile likelihood and conditionally parametric models. *Ann. Statist.* **20**, 1768 - 1802.
- VARIN, C., REID, N. and FIRTH, D. (2011). An overview of composite likelihood methods. *Statistica Sinica* **21** 5–42.
- VARIN, C. and VIDONI, P. (2005). A note on composite likelihood inference and model selection. *Biometrika* **92** 519–528.
- WIUF, C. (2006). Consistency of estimators of population scaled parameters using composite likelihood. *J. Math. Biol.* **53** 821–841.
- ZHAO, Y. and JOE, H. (2005). Composite likelihood estimation in multivariate data analysis. *Canad. J. Statist.* **33** 335–356.