

# References

---

- [1] Andersen, P.K., Borgan, O., Gill, R.D., and Keiding, N. (1992). *Statistical Models Based on Counting Processes*. Springer, Berlin.
- [2] Arcones, M.A., and Giné, E. (1993). Limit theorems for  $U$ -processes. *Annals of Probability* **21**, 1494–1542.
- [3] Bahadur, R.R. (1967). An optimal property of the likelihood ratio statistic. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability (1965/66)* **I**, 13–26. University of California Press, Berkeley.
- [4] Bahadur, R.R. (1971). Some limit theorems in statistics. *Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics* **4**. Society for Industrial and Applied Mathematics, Philadelphia.
- [5] Barndorff-Nielsen, O.E., and Hall, P. (1988). On the level-error after Bartlett adjustment of the likelihood ratio statistic. *Biometrika* **75**, 378–378.
- [6] Bauer, H. (1981). *Probability Theory and Elements of Measure Theory*. Holt, Rinehart, and Winston, New York.
- [7] Bentkus, V., Götze, F., van Zwet, W.R. (1997). An Edgeworth expansion for symmetric statistics. *Annals of Statistics* **25**, 851–896.
- [8] 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.
- [9] Bickel, P.J., and Ghosh, J.K. (1990). A decomposition for the likelihood ratio statistic and the Bartlett correction—a Bayesian argument. *Annals of Statistics* **18**, 1070–1090.
- [10] Bickel, P.J., and Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. *Annals of Statistics* **1**, 1071–1095.
- [11] Billingsley, P. (1968). *Convergence of Probability Measures*. John Wiley, New York.
- [12] Birgé, L. (1983). Approximation dans les espaces métriques et théorie de l'estimation. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **65**, 181–238.
- [13] Birgé, L. (1997). Estimation of unimodal densities without smoothness assumptions. *Annals of Statistics* **25**, 970–981.
- [14] Birgé, L., and Massart, P. (1993). Rates of convergence for minimum contrast estimators. *Probability Theory and Related Fields* **97**, 113–150.
- [15] Birgé, L., and Massart, P. (1997). From model selection to adaptive estimation. *Festschrift for Lucien Le Cam*. Springer, New York, 55–87.
- [16] Birman, M.S., and Solomjak, M.Z. (1967). Piecewise-polynomial approximation of functions of the classes  $W_p$ . *Mathematics of the USSR Sbornik* **73**, 295–317.
- [17] Brown, L. (1987). *Fundamentals of Statistical Exponential Families with Applications in Statistical Decision Theory*. Institute of Mathematical Statistics, California.
- [18] Brown, L.D., and Fox, M. (1974). Admissibility of procedures in two-dimensional location parameter problems. *Annals of Statistics* **2**, 248–266.
- [19] Cantelli, F.P. (1933). Sulla determinazione empirica delle leggi di probabilità. *Giornale dell'Istituto Italiano degli Attuari* **4**, 421–424.

- [20] Chernoff, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Mathematical Statistics* **23**, 493–507.
- [21] Chernoff, H. (1954). On the distribution of the likelihood ratio statistic. *Annals of Mathematical Statistics* **25**, 573–578.
- [22] Chernoff, H., and Lehmann, E.L. (1954). The use of maximum likelihood estimates in  $\chi^2$  tests for goodness of fit. *Annals of Mathematical Statistics* **25**, 579–586.
- [23] Chow, Y.S., and Teicher, H. (1978). *Probability Theory*. Springer-Verlag, New York.
- [24] Cohn, D.L. (1980). *Measure Theory*. Birkhäuser, Boston.
- [25] Copas, J. (1975). On the unimodality of the likelihood for the Cauchy distribution. *Biometrika* **62**, 701–704.
- [26] Cramér, H. (1938). Sur un nouveau théorème-limite de la théorie des probabilités. *Actualités Scientifiques et Industrielles* **736**, 5–23.
- [27] Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press, Princeton.
- [28] Csörgő, M. (1983). *Quantile Processes with Statistical Applications*. CBMS-NSF Regional Conference Series in Applied Mathematics **42**. Society for Industrial and Applied Mathematics (SIAM), Philadelphia.
- [29] Dacunha-Castelle, D., and Duflo, M. (1993). *Probabilités et Statistiques, tome II*. Masson, Paris.
- [30] Davies, R.B. (1973). Asymptotic inference in stationary Gaussian time-series. *Advances in Applied Probability* **4**, 469–497.
- [31] Dembo, A., and Zeitouni, O. (1993). *Large Deviation Techniques and Applications*. Jones and Bartlett Publishers, Boston.
- [32] Deuschel, J.D., and Stroock, D.W. (1989). *Large Deviations*. Academic Press, New York.
- [33] Devroye, L., and Györfi, L. (1985). *Nonparametric Density Estimation: The  $L_1$ -View*. John Wiley & Sons, New York.
- [34] Diaconis, P., and Freedman, D. (1986). On the consistency of Bayes estimates. *Annals of Statistics* **14**, 1–26.
- [35] Donald, S.G., and Newey, W.K. (1994). Series estimation of semilinear models. *Journal of Multivariate Analysis* **50**, 30–40.
- [36] Donoho, D.L., and Johnstone, I.M. (1994). Idea spatial adaptation by wavelet shrinkage. *Biometrika* **81**, 425–455.
- [37] Donoho, D.L., and Liu, R.C. (1991). Geometrizing rates of convergence II, III. *Annals of Statistics* **19**, 633–701.
- [38] Donsker, M.D. (1952). Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems. *Annals of Mathematical Statistics* **23**, 277–281.
- [39] Doob, J. (1948). Application of the theory of martingales. *Le Calcul des Probabilités et ses Applications. Colloques Internationales du CNRS Paris*, 22–28.
- [40] Drost, F.C. (1988). *Asymptotics for Generalized Chi-Square Goodness-of-Fit Tests*. CWI tract **48**. Centrum voor Wiskunde en Informatica, Amsterdam.
- [41] Dudley, R.M. (1976). *Probability and Metrics: Convergence of Laws on Metric Spaces*. Mathematics Institute Lecture Notes Series **45**. Aarhus University, Denmark.
- [42] Dudley, R.M. (1989). *Real Analysis and Probability*. Wadsworth, Belmont, California.
- [43] Dupač, V., and Hájek, J. (1969). Asymptotic normality of simple linear rank statistics under alternatives. *Annals of Mathematical Statistics* **40**, 1992–2017.
- [44] Efron, B., and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*. Chapman and Hall, London.
- [45] Fahrmeir, L., and Kaufmann, H. (1985). Consistency and asymptotic normality of the maximum likelihood estimator in generalized linear models. *Annals of Statistics* **13**, 342–368. (Correction: *Annals of Statistics* **14**, 1643.)
- [46] Farrell, R.H. (1972). On the best obtainable asymptotic rates of convergence in estimation of a density function at a point. *Annals of Mathematical Statistics* **43**, 170–180.
- [47] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications, vol. II*. John Wiley & Sons, New York.

- [48] Fisher, R.A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London, Series A* **222**, 309–368.
- [49] Fisher, R.A. (1924). The conditions under which  $\chi^2$  measures the discrepancy between observations and hypothesis. *Journal Royal Statist. Soc.* **87**, 442–450.
- [50] Fisher, R.A. (1925). Theory of statistical estimation. *Proceedings of the Cambridge Philosophical Society* **22**, 700–725.
- [51] van de Geer, S.A. (1988). *Regression Analysis and Empirical Processes*. CWI Tract **45**. Centrum voor Wiskunde en Informatica, Amsterdam.
- [52] Ghosh, J.K. (1994). *Higher Order Asymptotics*. Institute of Mathematical Statistics, Hayward.
- [53] Gill, R.D. (1989). Non- and semi-parametric maximum likelihood estimators and the von-Mises method (part I). *Scandinavian Journal of Statistics* **16**, 97–128.
- [54] Gill, R.D. (1994). Lectures on survival analysis. *Lecture Notes in Mathematics* **1581**, 115–241.
- [55] Gill, R.D., and Johansen, S. (1990). A survey of product-integration with a view towards application in survival analysis. *Annals of Statistics* **18**, 1501–1555.
- [56] Gill, R.D., and van der Vaart, A.W. (1993). Non- and semi-parametric maximum likelihood estimators and the von Mises method (part II). *Scandinavian Journal of Statistics* **20**, 271–288.
- [57] Giné, E., and Zinn, J. (1986). Lectures on the central limit theorem for empirical processes. *Lecture Notes in Mathematics* **1221**, 50–113.
- [58] Giné, E., and Zinn, J. (1990). Bootstrapping general empirical measures. *Annals of Probability* **18**, 851–869.
- [59] Glivenko, V. (1933). Sulla determinazione empirica della leggi di probabilità. *Giornale dell'Istituto Italiano degli Attuari* **4**, 92–99.
- [60] Greenwood, P.E., and Nikulin, M.S. (1996). *A Guide to Chi-Squared Testing*. John Wiley & Sons, New York.
- [61] Groeneboom, P. (1980). *Large Deviations and Bahadur Efficiencies*. MC tract **118**, Centrum voor Wiskunde en Informatica, Amsterdam.
- [62] Groeneboom, P. (1985). Estimating a monotone density. *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer* **2**, 539–555. Wadsworth, Monterey, California.
- [63] Groeneboom, P. (1988). Brownian Motion with a parabolic drift and Airy functions. *Probability Theory and Related Fields* **81**, 79–109.
- [64] Groeneboom, P., Lopuhaä, H.P. (1993). Isotonic estimators of monotone densities and distribution functions: basic facts. *Statistica Neerlandica* **47**, 175–183.
- [65] Groeneboom, P., Oosterhoff, J., and Ruymgaart, F. (1979). Large deviation theorems for empirical probability measures. *Annals of Probability* **7**, 553–586.
- [66] de Haan, L. (1976). Sample extremes: An elementary introduction. *Statistica Neerlandica* **30**, 161–172.
- [67] Hájek, J. (1961). Some extensions of the Wald-Wolfowitz-Noether theorem. *Annals of Mathematical Statistics* **32**, 506–523.
- [68] Hájek, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. *Annals of Mathematical Statistics* **39**, 325–346.
- [69] Hájek, J. (1970). A characterization of limiting distributions of regular estimates. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **14**, 323–330.
- [70] Hájek, J. (1972). Local asymptotic minimax and admissibility in estimation. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* **1**, 175–194.
- [71] Hájek, J., and Šidák, Z. (1967). *Theory of Rank Tests*. Academic Press, New York.
- [72] Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. Springer Series in Statistics. Springer-Verlag, New York.
- [73] Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., and Stahel, W.A. (1986). *Robust Statistics: the Approach Based on Influence Functions*. Wiley, New York.
- [74] Halmers, R. (1982). *Edgeworth Expansions for Linear Combinations of Order Statistics*. Mathematical Centre Tracts **105**. Mathematisch Centrum, Amsterdam.
- [75] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Annals of Mathematical Statistics* **19**, 293–325.

- [76] Hoffmann-Jørgensen, J. (1991). *Stochastic Processes on Polish Spaces. Various Publication Series 39*. Aarhus Universitet, Aarhus, Denmark.
- [77] Huang, J. (1996). Efficient estimation for the Cox model with interval censoring. *Annals of Statistics* **24**, 540–568.
- [78] Huber, P. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* **1**, 221–233. University of California Press, Berkeley.
- [79] Huber, P. (1974). *Robust Statistics*. Wiley, New York.
- [80] Ibragimov, I.A., and Has'minskii, R.Z. (1981). *Statistical Estimation: Asymptotic Theory*. Springer-Verlag, New York.
- [81] Jagers, P. (1975). *Branching Processes with Biological Applications*. John Wiley & Sons, London-New York-Sydney.
- [82] Janssen, A., and Mason, D.M. (1990). *Nonstandard Rank Tests. Lecture Notes in Statistics 65*. Springer-Verlag, New York.
- [83] Jensen, J.L. (1993). A historical sketch and some new results on the improved log likelihood ratio statistic. *Scandinavian Journal of Statistics* **20**, 1–15.
- [84] Kallenberg, W.C.M. (1983). Intermediate efficiency, theory and examples. *Annals of Statistics* **11**, 498–504.
- [85] Kallenberg, W.C.M., and Ledwina, T. (1987). On local and nonlocal measures of efficiency. *Annals of Statistics* **15**, 1401–1420.
- [86] Kallenberg, W.C.M., Oosterhoff, J., and Schriever, B.F. (1980). The number of classes in chi-squared goodness-of-fit tests. *Journal of the American Statistical Association* **80**, 959–968.
- [87] Kim, J., and Pollard, D. (1990). Cube root asymptotics. *Annals of Statistics* **18**, 191–219.
- [88] Kolčinskii, V.I. (1981). On the central limit theorem for empirical measures. *Theory of Probability and Mathematical Statistics* **24**, 71–82.
- [89] Koul, H.L., and Pflug, G.C. (1990). Weakly adaptive estimators in explosive autoregression. *Annals of Statistics* **18**, 939–960.
- [90] Leadbetter, M.R., Lindgren, G., and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer-Verlag, New York.
- [91] Le Cam, L. (1953). On some asymptotic properties of maximum likelihood estimates and related Bayes estimates. *University of California Publications in Statistics* **1**, 277–330.
- [92] Le Cam, L. (1960). Locally asymptotically normal families of distributions. *University of California Publications in Statistics* **3**, 37–98.
- [93] Le Cam, L. (1969). *Théorie Asymptotique de la Décision Statistique*. Les Presses de l'Université de Montréal, Montreal.
- [94] Le Cam, L. (1970). On the assumptions used to prove asymptotic normality of maximum likelihood estimates. *Annals of Mathematical Statistics* **41**, 802–828.
- [95] Le Cam, L. (1972). Limits of experiments. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* **1**, 245–261. University of California Press, Berkeley.
- [96] Le Cam, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. Springer-Verlag, New York.
- [97] Le Cam, L.M., and Yang, G. (1990). *Asymptotics in Statistics: Some Basic Concepts*. Springer-Verlag, New York.
- [98] Ledoux, M., and Talagrand, M. (1991). *Probability in Banach Spaces: Isoperimetry and Processes*. Springer-Verlag, Berlin.
- [99] Lehmann, E.L. (1983). *Theory of Point Estimation*. Wiley, New York.
- [100] Lehmann, E.L. (1991). *Testing Statistical Hypotheses, 2nd edition*. Wiley, New York.
- [101] Levit, B.Y. (1978). Infinite-dimensional informational lower bounds. *Theory of Probability and its Applications* **23**, 388–394.
- [102] Mammen, E., and van de Geer, S.A. (1997). Penalized quasi-likelihood estimation in partial linear models. *Annals of Statistics* **25**, 1014–1035.
- [103] Massart, P. (1990). The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. *Annals of Probability* **18**, 1269–1283.

- [104] von Mises, R. (1947). On the asymptotic distribution of differentiable statistical functions. *Annals of Mathematical Statistics* **18**, 309–348.
- [105] Murphy, S.A., Rossini, T.J., and van der Vaart, A.W. (1997). MLE in the proportional odds model. *Journal of the American Statistical Association* **92**, 968–976.
- [106] Murphy, S.A., and van der Vaart, A.W. (1997). Semiparametric likelihood ratio inference. *Annals of Statistics* **25**, 1471–1509.
- [107] Murphy, S.A., and van der Vaart, A.W. (1996). Semiparametric mixtures in case-control studies.
- [108] Murphy, S.A., and van der Vaart, A.W. (1996). Likelihood ratio inference in the errors-in-variables model. *Journal of Multivariate Analysis* **59**, 81–108.
- [109] Noether, G.E. (1955). On a theorem of Pitman. *Annals of Mathematical Statistics* **25**, 64–68.
- [110] Nussbaum, M. (1996). Asymptotic equivalence of density estimation and Gaussian white noise. *Annals of Statistics* **24**, 2399–2430.
- [111] Ossiander, M. (1987). A central limit theorem under metric entropy with  $L_2$  bracketing. *Annals of Probability* **15**, 897–919.
- [112] Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine, Series 5* **50**, 157–175. (Reprinted in: *Karl Pearson's Early Statistical Papers*, Cambridge University Press, 1956.)
- [113] Pfanzagl, J., and Wefelmeyer, W. (1982). *Contributions to a General Asymptotic Statistical Theory. Lecture Notes in Statistics* **13**. Springer-Verlag, New York.
- [114] Pfanzagl, J., and Wefelmeyer, W. (1985). *Asymptotic Expansions for General Statistical Models. Lecture Notes in Statistics* **31**. Springer-Verlag, New York.
- [115] Pflug, G.C. (1983). The limiting loglikelihood process for discontinuous density families. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **64**, 15–35.
- [116] Pollard, D. (1982). A central limit theorem for empirical processes. *Journal of the Australian Mathematical Society A* **33**, 235–248.
- [117] Pollard, D. (1984). *Convergence of Stochastic Processes*. Springer-Verlag, New York.
- [118] Pollard, D. (1985). New ways to prove central limit theorems. *Econometric Theory* **1**, 295–314.
- [119] Pollard, D. (1989). A maximal inequality for sums of independent processes under a bracketing condition.
- [120] Pollard, D. (1990). *Empirical Processes: Theory and Applications. NSF-CBMS Regional Conference Series in Probability and Statistics* **2**. Institute of Mathematical Statistics and American Statistical Association. Hayward, California.
- [121] Prakasa Rao, B.L.S. (1983). *Nonparametric Functional Estimation*. Academic Press, Orlando.
- [122] Qin, J., and Lawless, J. (1994). Empirical likelihood and general estimating equations. *Annals of Statistics* **22**, 300–325.
- [123] Rao, C.R. (1973). *Linear Statistical Inference and Its Applications*. Wiley, New York.
- [124] Reed, M., and Simon, B. (1980). *Functional Analysis*. Academic Press, Orlando.
- [125] Reeds, J.A. (1976). *On the Definition of von Mises Functionals*. Ph.D. dissertation, Department of Statistics, Harvard University, Cambridge, MA.
- [126] Reeds, J.A. (1985). Asymptotic number of roots of Cauchy location likelihood equations. *Annals of Statistics* **13**, 775–784.
- [127] Révész, P. (1968). *The Laws of Large Numbers*. Academic Press, New York.
- [128] Robins, J.M., and Ritov, Y. (1997). Towards a curse of dimensionality appropriate (CODA) asymptotic theory for semi-parametric models. *Statistics in Medicine* **16**, 285–319.
- [129] Robins, J.M., and Rotnitzky, A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS Epidemiology–Methodological Issues*, 297–331, eds: N. Jewell, K. Dietz, and V. Farewell. Birkhäuser, Boston.
- [130] Roussas, G.G. (1972). *Contiguity of Probability Measures*. Cambridge University Press, Cambridge.
- [131] Rubin, H., and Vitale, R.A. (1980). Asymptotic distribution of symmetric statistics. *Annals of Statistics* **8**, 165–170.
- [132] Rudin, W. (1973). *Functional Analysis*. McGraw-Hill, New York.



- [133] Shiriyayev, A.N. (1984). *Probability*. Springer-Verlag, New York-Berlin.
- [134] Shorack, G.R., and Wellner, J.A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.
- [135] Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.
- [136] Stigler, S.M. (1974). Linear functions of order statistics with smooth weight functions. *Annals of Statistics* **2**, 676–693. (Correction: *Annals of Statistics* **7**, 466.)
- [137] Stone, C.J. (1990). Large-sample inference for log-spline models. *Annals of Statistics*, **18**, 717–741.
- [138] Strasser, H. (1985). *Mathematical Theory of Statistics*. Walter de Gruyter, Berlin.
- [139] van der Vaart, A.W. (1988). *Statistical Estimation in Large Parameter Spaces*. CWI Tracts **44**. Centrum voor Wiskunde en Informatica, Amsterdam.
- [140] van der Vaart, A.W. (1991). On differentiable functionals. *Annals of Statistics* **19**, 178–204.
- [141] van der Vaart, A.W. (1991). An asymptotic representation theorem. *International Statistical Review* **59**, 97–121.
- [142] van der Vaart, A.W. (1994). *Limits of Experiments*. Lecture Notes, Yale University.
- [143] van der Vaart, A.W. (1995). Efficiency of infinite dimensional  $M$ -estimators. *Statistica Neerlandica* **49**, 9–30.
- [144] van der Vaart, A.W. (1994). Maximum likelihood estimation with partially censored observations. *Annals of Statistics* **22**, 1896–1916.
- [145] van der Vaart, A.W. (1996). Efficient estimation in semiparametric models. *Annals of Statistics* **24**, 862–878.
- [146] van der Vaart, A.W., and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes*. Springer, New York.
- [147] Vapnik, V.N., and Červonenkis, A.Y. (1971). On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and Its Applications* **16**, 264–280.
- [148] Wahba, G. (1990). *Spline models for observational data*. CBMS-NSF Regional Conference Series in Applied Mathematics **59**. Society for Industrial and Applied Mathematics, Philadelphia.
- [149] Wald, A. (1943). Test of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society* **54**, 426–482.
- [150] Wilks, S.S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Annals of Mathematical Statistics* **19**, 60–62.
- [151] van Zwet, W.R. (1984). A Berry-Esseen bound for symmetric statistics. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **66**, 425–440.