References

Section 1.1: Introduction

Barbe, P., Genest, C., Ghoudi, K. Rémillard, B. (1996). On Kendall's process. *J. Multivariate Anal.* **58**, 197-229.

Cantelli, F. P. (1933). Sulla determinazione empirica delle leggi di probabilita. *Giorn. Ist. Ital. Attuari* **4**, 421 - 424.

Donsker, M. D. (1952). Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statist.* **23**, 277-281.

Dudley, R. M. (1978). Central limit theorems for empirical measures. *Ann. Probab.* **6**, 899-929.

Genest, C. and Rivest, L.-P. (1993). Statistical inference procedures for bivariate Archimedean copulas. *J. Amer. Statist. Assoc.* **88**, 1034-1043.

Glivenko, V. (1933). Sulla determinazione empirica della legge di probabilita. *Giorn. Ist. Ital. Attuari* **4**, 92-99.

Pollard, D. (1989). Asymptotics via empirical processes. *Statist. Sci.* **4**, 341 - 366.

Vapnik, V. N. and Chervonenkis, A. Ya. (1971). On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and Its Applications* **16**, 264 - 280.

Section 1.2: Weak Convergence: the fundamental theorems

Andersen, N. T. (1985). The central limit theorem for non-separable valued functions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **70**, 445 - 455.

Andersen, N. T. and Dobric, V. (1987) The central limit theorem for stochastic processes. *Ann. Probability* **15**. 164 - 177.

Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

Billingsley, P. (1999). *Convergence of Probability Measures.* Wiley, New York.

de la Peña, V. H., and Giné, E. *Decoupling. From dependence to independence.* Springer-Verlag, New York, 1999.

Hoffmann-Jørgensen, J. (1984). *Stochastic Processes on Polish Spaces.* Unpublished.

Hoffmann-Jørgensen, J. (1991). *Stochastic Processes on Polish Spaces.* Various Publication Series **39**. Aarhus Universeitet, Aarhus, Denmark.

Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes.* Springer-Verlag, New York.

Section 1.3: Maximal inequalities and chaining

de la Peña, V. H., and Giné, E. *Decoupling. From dependence to independence.* Springer-Verlag, New York, 1999.

Dudley, R. M. (1999). *Uniform Central Limit Theorems*. Cambridge Univ. Press, Cambridge.

Fernique, X. (1997). *Fonctions Aléatoires Gaussiennes, Vecteurs Aléatoires Gaussiens.* Publications du Centre de Recherches Mathématiques, Montréal.

Giné, E., Mason, D. M., and Zaitsev, A. Yu. (2003). The L_1-norm density estimator process. *Ann. Probability* **31**, 719 - 768.

Giné, E. and Zinn, J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* **12** (1984), 929-989.

Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces*. Springer-Verlag, Berlin.

Resnick, S. (1987). *Extreme Values, Regular Variation, and Point Processes.* Springer-Verlag, New York.

Section 1.4: Some Results for Gaussian Processes

Adler, R. J. (1990). An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes. *IMS Lecture Notes - Monograph Series* **12**. Institute for Mathematical Statistics, Hayard.

Borell, C. (1975). The Brunn-Minkowski inequality in Gauss space. *Inventiones Mathematicae* **30**, 205-216.

Dudley, R. M. (1999). *Uniform Central Limit Theorems*. Cambridge Univ. Press, Cambridge.

Ibragimov, I. A., Sudakov, V. N., and Tsirel'son, B.S. (1976). Norms of Gaussian sample functions. *Proceedings of the Third Japan-USSR Symposium on Probability Theory. Lecture Notes in Mathematics* **550**, 20 - 41. Springer-Verlag, New York.

Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces*. Springer-Verlag, Berlin.

Section 1.5: Inequalities for Sums of Independent Processes

de la Peña, V. H., and Giné, E. (1999). *Decoupling. From dependence to independence.* Springer-Verlag, New York.

Hoffmann-Jørgensen, J. (1974). Sums of independent Banach space valued random variables. *Studia Mathematica* **52**, 159 - 186.

Johnson, W. B., Schechtman, G., and Zinn, J. (1985). Best constants in moment inequalities for linear combinations of independent and exchangeable random variables. *Ann. Probability* **13**, 234-253.

Rosenthal, H. P. (1970). On the subspaces of L_p $(p > 2)$ spanned by sequences of independent random variables. *Israel J. Math.* **8**, 273 - 303.

Section 1.6: Glivenko-Cantelli theorems

Blum, J. R. (1955). On the convergence of empiric distributions. *Ann. Math. Statist.* **26**, 527 - 529.

DeHardt, J. (1971). Generalizations of the Glivenko-Cantelli theorem. *Ann. Math. Statist.* **42**, 2050-2055.

Dudley, R. M. (1984). A Course on Empirical Processes (École d'Été de Probabilités de Saint-Flour XII-1982). *Lecture Notes in Mathematics* **1097**, 2 - 141 (P. L. Hennequin, ed.). Springer-Verlag, New York.

Dudley, R. M. (1998a). Consistency of M-estimators and one-sided bracketing. In *Proceedings of the First International Conference on High-Dimensional Probability*, 33 - 58. E. Eberlein, M. Hahn, and M. Talagrand, editors. Birkhäuser, Basel.

Dudley, R. M., Giné, E., and Zinn, J. (1991). Uniform and universal Glivenko-Cantelli classes. *J. Theoret. Probab.* **4**, 485-510.

Giné, E. and Zinn, J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* **12** (1984), 929-989.

Giné, E. and Zinn, J. (1986). Lectures on the central limit theorem for empirical processes. Probability and Banach Spaces, Zaragoza, Spain. *Lecture Notes in Math.* **1221**, 50-113.

Le Cam, L. (1953). On some asymptotic properties of maximum likelihood estimates and related estimates. *Univ. Calif. Publ. in Statist.* **1**, 277 - 330.

Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces*. Springer-Verlag, Berlin.

Pollard, D. (1981). Limit theorems for empirical processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **57**, 181 - 195.

Steele, M. (1978). Empirical discrepancies and subaddtive processes. *Ann. Probability* **6**, 118 - 127.

Van der Vaart, A. W. (1996). Efficient maximum likelihood estimation in semiparametric mixture models. *Ann. Statist.* **24**, 862-878.

Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer-Verlag, New York.

Van der Vaart, A. W. and Wellner, J. A. (2000). Preservation theorems for Glivenko-Cantelli and uniform Glivenko-Cantelli classes, pp. 115 - 134 In *High Dimensional Probability II*, Evarist Giné, David Mason, and Jon A. Wellner, editors, Birkhäuser, Boston.

Vapnik, V. N. and Chervonenkis, A. Ya. (1971). On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and Its Applications* **16**, 264 - 280.

Vapnik, V. N. and Chervonenkis, A. Ya. (1981). Necessary and sufficient conditions for the uniform convergence of means to their expectations. *Theory of Probability and Its Applications* **26**, 532 - 553.

Wald, A. (1949). Note on the consistency of the maximum likelihood estimate. *Ann. Math. Statist.* **20**, 595 - 601.

Section 1.7: Donsker theorems

Andersen, N. T., Giné, E., Ossiander, M., and Zinn, J. (1988). The central limit theorem and the law of iterated logarithm for empirical processes under local conditions. *Probability Theory and Related Fields* **77**, 271-0305.

Arcones, M. and Giné, E. (1993). Limit theorems for U-processes. *Ann. Probability* **21**, 1494 - 1452.

Durst, M. and Dudley, R. M. (1981). Empirical processes, Vapnik-Chervonenkis classes and poisson processes. In *Probab. Math. Statist.* (Wroclaw) **1**, 109-115.

Giné, E. and Guillou, A. (2001). On consistency of kernel density estimators for randomly censored data: rates holding uniformly over adaptive intervals. *Ann. I. H. Poincaré* **4** (2001), 503-522.

Giné, E., Koltchinskii, V., and Wellner, J. A. (2003). Ratio limit theorems for empirical processes. In *Stochastic Inequalitites and Applications*, Proceedings of the Euro-Conference on Stochastic Inequalities helds at Barcelona, June 18-22, 2002, 249 - 278. E. Giné, Christian Houdré, and David Nulart, eds.

Giné, E. and Zinn, J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* **12** (1984), 929-989.

Giné, E. and Zinn, J. (1986). Lectures on the central limit theorem for empirical processes. Probability and Banach Spaces, Zaragoza, Spain. *Lecture Notes in Math.* **1221**, 50-113.

Le Cam, L. (1983). A remark on empirical measures. In *Festschrift for E. L. Lehmann*, pages 305-327. Eds. P. Bickel, K. doksum, and J. Hodges. Wadsworth, Belmont.

Loève, M. (1978). *Probability, II.* Fourth Edition. Springer, New York.

Ossiander, M. (1987). A central limit theorem under metric entropy with L_2 bracketing. *Ann. Probability* **15**, 897 - 919.

Pisier, G. (1984). Remarques sur les classes de Vapnik-Chervonenkis. *Ann. Inst. H. Poincaré* **20**, 287-298.

Pollard, D. (1982). A central limit theorem for empirical processes. . *J. austral. Math. Soc. Ser. A* **33**, 235-248.

Koltchinskii, V. I. (1981). On the central limit theorem for empirical measures. *Theor. Probab. Math. Statist.* **24**, 63 - 75.

Section 1.8: VC - theory: bounding uniform covering numbers

Ball, K. and Pajor, A. (1990). The entropy of convex bodies with “few” extreme points. In *Geometry of Banach Spaces, Proceedings of the conference held in Strobl, Austria, 1989*, *London Mathematical Society Lecture Note Series* **158**, 25 - 32, P.F.X. Müller and W. Schachermayer, eds.

Carl, B. (1997). Metric entropy of convex hulls in Hilbert spaces. *Bull. London Math. Soc.* **29** 452 - 458.

Carl, B., Kyrezl, I. and Pajor, A. (1999). Metric entropy of convex hulls in Banach spaces. *J. London Math. Soc.* **60**, 871 - 896.

Dudley, R. M. (1978). Central limit theorems for empirical measures. *Ann. Probab.* **6**, 899-929.

Dudley, R. M. (1987). Universal Donsker classes and metric entropy. *Ann. Probability* **15**, 1306 - 1326.

Dunford, N. and Schwartz, J. T. (1958). *Linear Operators. Part I.* Wiley-Interscience, New York.

Haussler, D.(1995). Sphere packing numbers for subsets of the Boolean $n-$cube with bounded Vapnik-Chervonenkis dimension. *J. Comb. Theory* **A 69**, 217 - 232.

Laskowski, M. C. (1992). Vapnik-Chervonenkis classes of definable sets. *J. London Math. Soc.* **45**, 377 - 384.

Mendelson, S. (2001). On the size of convex hulls of small sets. *J. Machine Learning Res.* **2**, 1 - 18.

Nolan, D. and Pollard, D. (1987). U-processes: rates of convergence. *Ann. Statist.* **15**, 780 - 799.

Nolan, D. and Pollard, D. (1988). Functional limit theorems for U-processes. *Ann. Probab.* **16**, 1291 - 1298.

Olshen, R. A., Biden, E. N., Wyatt, M., and Sutherland, D. H. (1989). Gait analysis and the bootstrap. *Ann. Statist.* **17**, 1419 - 1440.

Rudin, W. (1973). *Functional Analysis.* McGraw-Hill, New York.

Song, S. and Wellner, J. A. (2002). An upper bound for uniform entropy numbers. Technical Report No. **406**, Department of Statistics, University of Washington. Available at *http://www.stat.washington.edu/www/research/reports/ #2002/tr409.ps*

Stengle, G., and Yukich, J. E. (1989). Some new Vapnik-Chervonenkis classes. *Ann. Statist.* **17**, 1441-1446.

Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes.* Springer-Verlag, New York.

Section 1.9: Bracketing numbers

Birman, M. S. and Solomjak, M. Z. (1967). Piece-wise polynomial approximations of functions in the classes W_p^α. *Mathematics of the USSR Sbornik* **73**, 295-317.

Birgé, L. and Massart, P. (2000). An adaptive compression algorithm in Besov spaces. *Const. Approx.* **16**, 1 - 36.

Bolthausen, E. (1978). Weak convergence of an empirical process indexed by the closed convex subsets of I^2. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **43**, 173 - 181.

Bronstein, E. M. (1976). Epsilon-entropy of convex sets and functions. *Siberian Math. J.* **17** 393-398.

Clements, G. F. (1963). Entropies of sets of functions of bounded variation. *Canadian J. Math.* **15**, 422 - 432.

DeVore, R. and Lorentz, G. G. (1993). *Constructive Approximation.* Springer-Verlag, Berlin.

Hanson, D. L., Pledger, G., and Wright, F. T. (1973). On consistency in monotonic regression. *Ann. Statist.* **1**, 401 - 421.

Kolmogorov, A. N. and Tikhomirov, V. M. (1959). $\epsilon-$entropy and $\epsilon-$capacity of sets in function spaces. *Uspekhi Mat. Nauk* **14**, 3 - 86. (Engl. Translation in *American Mathematical Society Translations* **(2) 17** (1961), 277 - 364.

Lorentz, G. G. (1966). *Approximation of Functions*. Holt, Rhinehart, Winston, New York.

Strassen, V. and Dudley, R. M (1969). The central limit theorem and epsilon-entropy. *Lecture Notes in Mathematics* **89**, 224-231. Springer-Verlag, Berlin.

Van de Geer, S. (1991). The entropy bound for monotone functions. *Technical Report* **91-10**, University of Leiden.

Van de Geer, S. (2000). *Empirical Processes in M-Estimation*. Cambridge University Press, Cambridge.

Van der Vaart, A. W. (1994). Bracketing smooth functions. *Stoch. Proc. Applic.* **52**, 93-105.

Van der Vaart, A. W. (1996). New Donsker classes. *Ann. Probability*, **24**, 2128–2140.

Wright, F. T. (1981). The empirical discrepancy over lower layers and a related law of large numbers. *Ann. Probab.* **9**, 323 - 329.

Section 1.10: Multiplier inequalities and the multiplier CLT

Giné, E. and Zinn, J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* **12** (1984), 929-989.

Giné, E. and Zinn, J., *Lectures on the central limit theorem for empirical processes.* Probability and Banach Spaces, Zaragoza, Spain. Lecture Notes in Math., **1221** (1986), 50-113.

Ledoux, M. and Talagrand, M. (1986). Conditions d'intégrabilité pour les multiplicateurs dans le TLC Banachque. *Ann. Probability* **14**, 916-921.

Ledoux, M. and Talagrand, M. (1988). Un critère sur les petite boules dans le théorème limite central. *Probability Theory and Related Fields* **77**, 29 - 47.

Section 1.11: Further Developments: Material Not Covered

Alexander, K. S. (1987a). Central limit theorems for stochastic processes under random entropy conditions. *Probab. Theory Related Fields* **75**, 351-378.

Alexander, K. S. (1987b). Rates of growth and sample moduli for weighted empirical processes indexed by sets. *Probab. Theory Related Fields* **75**, 379-423.

Alexander, K. S. (1987c). The central limit theorem for weighted empirical processes indexed by sets. *J. Multivariate Anal.* **22**, 313-339.

Andersen, N. T., Giné, E., Ossiander, M., and Zinn, J. (1988). The central limit theorem and the law of iterated logarithm for empirical processes under local conditions. *Probability Theory and Related Fields* **77**, 271-0305.

Andrews, D. W. K. and Pollard, D. (1994). An introduction to functional central limit theorems for dependent stochastic processes. *Int. Statist. Rev.* **62**, 119 - 132.

Arcones, M. A. and Yu, Bin (1994). Limit theorems for empirical processes under dependence. In *Chaos expansions, multiple Wiener-Ito integrals and their applications*, C. Houdre and V. Perez-Abreu, eds. CRC Press, Boca Raton.

Bercu, B., Gassiat, E., Rio, E. (2002). Concentration inequalities, large and moderate deviations for self-normalized empirical processes. *Ann. Probab.* **30**, 1576-1604.

Dedecker, J. and Rio, E. (2000). On the functional central limit theorem for stationary processes. *Ann. Inst. H. Poincaré Probab. Statist.***36**, 1-34.

Dehling, H., Mikosch, T., and Sorensen, M. (2002). *Empirical Process Techniques for Dependent Data*. Birkhäuser, Boston.

de la Peña, V. H., and Giné, E. *Decoupling. From dependence to independence.* Springer-Verlag, New York, 1999.

Doukhan, P., Massart, P., and Rio, E. (1994). The functional central limit theorem for strongly mixing processes. *Ann. Inst. Henri Poincare* **30**, 63 - 82.

Dudley, R. M. and Koltchinskii, V. I. (1995). Envelope moment conditions and Donsker classes. *Theory Probab. Math. Statist.* **51**, 39-49.

Dudley, R. M., and Philipp, W. (1983). Invariance principles for sums of Banach space valued random elements and empirical processes. *Zeitchsrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **62** 509-552.

Giné, E. (1997). Decoupling and limit theorems for U-statistics and U-processes. In *Lectures on Probabiity Theory and Mathematical Statistics*, *Lectures Notes in Mathematics* **1665**, 1 - 36. Springer, Berlin.

Koltchinskii, V. I. (1994). Komlos-Major-Tusnady approximation for the general empirical process and Haar expansions of classes of functions. *J. Theoret. Probab.* **7**, 73-118.

Ledoux, M. and Talagrand, M. (1989). Comparison theorems, random geometry and some limit theorems for empirical processes. *Ann. Probab.* **17**, 596 - 631.

Pollard, D. (1995). Uniform ratio limit theorems for empirical processes. *Scand. J. Statist.* **22**, 271 - 278.

Rio, E. (1994). Local invariance principles and their application to density estimation. *Probab. Theory Related Fields* **98**, 21-45.

Rio, E. (1998). Processus empiriques absolument réguliers et entropie universelle. [Absolutely regular empirical processes and universal entropy] *Probab. Theory Related Fields* **111**, 585 - 608.

Section 2.1: Consistency of Maximum Likelihood Estimators

Balabdaoui, F. (2003). Estimation in some mixture models: direct and inverse problems. Ph.D. dissertation in progress, University of Washington.

Bauer, H. (1972). *Probability Theory and Elements of Measure Theory*. Holt, Rinehart, and Winston, New York.

Birgé, L. and Massart, P. (1993). Rates of convergence for minimum contrast estimators. *Prob. Theory and Related Fields* **97**, 113-150.

Groeneboom, P., Jongbloed, G., and Wellner, J. A. (2001). Estimation of a convex function: characterizations and asymptotic theory. *Ann. Statist.* **29**, 1653-1698.

Jewell, N. P. (1982) Mixtures of exponential distributions. *Ann. Statist.* **10**, 479 - 484.

Jewell, N. P. and Kalbfleisch, J. D. (2001) Maximum likelihood estimation of ordered multinomial parameters. Working Paper 2001-09, Department of Statistics and Acturial Science, University of Waterloo.

Jewell, N. P., Van der Laan, and Henneman, T. (2001). Nonparametric estimation from current status data with competing risks. Technical Report 97, Department of Biostatistics, University of California, Berkeley.

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.

Pfanzagl, J. (1988). Consistency of maximum likelihood estimators for certain nonparametric families, in particular: mixtures. *J. Statist. Planning and Inference* **19**, 137 - 158.

Prakasa Rao, B. L. S. (1969). Estimation of a unimodal density. *Sankhya Ser. A* **31**, 23 - 36.

Schick, A. and Yu, Q. (2000), Consistency of the GMLE with Mixed Case Interval-Censored Data, *Scand. J. Statist.* **27**, 45–55.

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. *J. Nonparametric Statist.* **6**, 293 - 310.

Section 2.2: M-Estimators: the Argmax Continuous Mapping Theorem

Davies, L. (1992). The asymptotics of Rousseeuw's minimum volume ellipsoid estimator. *Ann. Statist.* **20**, 1828 - 1843.

Grübel, R. (1988). The length of the shorth. *Ann. Statist.* **16**, 619-628.

Kim, J. and Pollard, D. (1990). Cube root asymptotics. *Ann. Statist.* **18**, 191 - 219.

Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.

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

Section 2.3: Rates of Convergence

Birgé, L. and Massart, P. (1993). Rates of convergence for minimum contrast estimators. *Prob. Theory and Related Fields* **97**, 113-150.

Van de Geer, S. (1996). Rates of convergence for the maximum likelihood estimator in mixture models. *J. Nonparametric Statist.* **6**, 293 - 310.

Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer-Verlag, New York.

Wong, W. H. and Shen, X. (1995). Probability inequalities for likelihood ratios and convergence rates for sieve MLEs. *Ann. Statist.* **23**, 339 - 362.

Section 2.4: M-Estimators and Z-Estimators

de la Peña, V. H., and Giné, E. *Decoupling. From dependence to independence.* Springer-Verlag, New York, 1999.

Giné, E. (1996). Empirical processes and applications. *Bernoulli* **2**, 1 - 28.

Giné, E. (1997). Lectures on some aspects of the bootstrap. In *Lectures on Probabiity Theory and Mathematical Statistics*, *Lectures Notes in Mathematics* **1665**, 37 - 152. Springer, Berlin.

Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1**, 221 - 233. Univ. California Press.

Pollard, D. (1985). New ways to prove central limit theorems. *Econometric Theory* **1**, 295 - 314.

Van der Vaart, A. W. (1995). Efficiency of infinite-dimensional M-estimators. *Statistica Neerl.* **49**, 9 - 30.

Section 2.5: Bootstrap Empirical Processes

Giné, E. and Zinn, J. (1990). Bootstrapping general empirical measures. *Ann. Probability* **18**, 851 - 869.

Klaassen, C. A. J. and Wellner, J. A. (1992). Kac empirical processes and the bootstrap. *Proceedings of the Eighth International Conference on Probability in Banach Spaces*, 411 - 429 (eds. M. Hahn and J. Kuelbs), Birkhäuser, New York.

Praestgaard, J. and Wellner, J. A. (1993). Exchangeably weighted bootstraps of the general empirical process. *Ann. Probablity* **21**, 2053 - 2086.

Section 2.6: Bootstrapping M- and Z- estimators

Arcones, M. A. and Giné, E. (1992). Bootstrap of M estimators and related statistical functionals. *Exploring the Limits of Bootstrap*, 14-17, R. LePage and L. Billard, eds. Wiley, New York.

Wellner, J. A. and Zhan, Y. (1996). Bootstrapping Z-estimators. Technical Report 308, Department of Statistics, University of Washington.

Section 2.7: Semiparametric Mixture Models

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.

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.

Lindsay, B. G. (1983). Efficiency of the conditional score in a mixture setting. *Ann. Statist.* **11**, 486 - 497.

Pfanzagl, J. (1988). Consistency of maximum likelihood estimators for certain nonparametric familes, in particular: mixtures. *J. Statist. Plann. Inference* **19**, 137 - 158.

Pfanzagl, J. (1990). *Estimation in Semiparametric Models*. Springer, New York.

Van der Vaart, A. W. (1988). Estimating a real parameter in a class of semiparametric models. *Ann. Statist.* **16**, 1450 - 1474.

Van der Vaart, A. W. (1996). Efficient maximum likelihood estimation in semiparametric mixture models. *Ann. Statist.* **24**, 862-878.

Section 2.8: Profile Likelihood

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. (1999). Observed information in semiparametric models. *Bernoulli* **5**, 381 - 412.

Murphy, S. and Van der Vaart, A. W. (2000). On profile likelihood. *J. Amer. Statist. Assoc.* **95**, 449 - 485.

Van der Vaart, A. W., Murphy, S. A., and Wellner, J. A. (1999). Current status regression. *Mathematical Methods of Statistics* **8**, 407 - 425.

Qin, J. and Wong, A. (1996). Empirical likelihood in a semiparametric model. *Scand. J. Statist.* **23**, 209-220.

Severini, T. A. and Wong, W. H. (1992). Profile likelihood and conditionally parametric models. *Ann. Statist.* **20**, 1768-1862.

Section 2.9: Further Developments: Topics Not Covered

Barron, A. Birgé, L., Massart, P. (1999). Risk bounds for model selection via penalization. *Probab. Theory Related Fields* **113**, 301-413.

Birgé, L., Massart, P. (1997). From model selection to adaptive estimation. *Festschrift for Lucien Le Cam*, 55–87, Springer, New York.

Birgé, L., and Massart, P. (1998). Minimum contrast estimators on sieves: exponential bounds and rates of convergence. *Bernoulli* **4**, 329-375.

Birgé, L. and Massart, P. (2001). Gaussian model selection. *J. Eur. Math. Soc. (JEMS)* **3**, 203–268.

Ghosal, S., and van der Vaart, A. W. (2001) Entropies and rates of convergence for maximum likelihood and Bayes estimation for mixtures of normal densities. *Ann. Statist.* **29**, 1233 - 1263.

Massart, P. (2000). Some applications of concentration inequalities to statistics. *Probability theory. Ann. Fac. Sci. Toulouse Math.* **(6) 9**, 245–303.

Van der Vaart, A. W. (1994a). Maximum likelihood estimation with partially censored data. *Ann. Statist.* **22**, 1896 - 1916.

Section 3.1:

Azuma, K. (1967). Weighted sums of certain dependent random variables. *Tokuku Math J.* **19**, 357 - 367.

Birnbaum, Z. W. and McCarty, R. C. (1958). A distribution-free upper confidence bound for $P(Y < X)$ based on independent samples of X and Y. *Ann. Math. Statist.* **29**, 558 - 562.

Dvoretzky, A., Kiefer, J., and Wolfowitz, J. (1956). Asymptotic minimax character of the samle distribution functions and of the classical multinomial estimator. *Ann. Math. Statist.* **27**, 642-669.

Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**, 13-30.

Kiefer, J. (1961). On large deviations of the empiric d.f. of vector chance variables and a law of the iterated logarithm. *Pacific J. Math.* **11**, 649 - 660.

Koltchinskii, V. (2001). Rademacher penalties and structural risk minization. *IEEE Trans. Inform. Theory***47**, 1902–1914.

Massart, P. (1990). The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. *Ann. Probability* **18**, 1269 - 1283.

McDiarmid, C. (1989). On the method of bounded differences. In *Surveys in Combinatorics, 1989*, pages 148 - 188. Cambridge University Press, Cambridge.

Talagrand, M. (1994). Sharper bounds for Gaussian and empirical processes. *Ann. Probability* **22**, 28 -76.

Yurinskii, Y.V. (1974). Exponential bounds for large deviations. *Theor. Probab. Appl.* **19**, 154 - 155.

Yurinskii, Y.V. (1976). Exponential inequalities for sums of random vectors. *J. Mult. Anal.* **6**, 473 - 499.

Section 3.2:

Bousquet, O. (2002). Concentration inequalities and empirical process theory applied to the analysis of learning algorithms. Ph.D. dissertation, Ecole Polytechnique, Paris.

Bousquet, O. (2002). A Bennett concentration inequality and its application to suprema of empirical processes. *C. R. Acad. Sci. Paris Ser.* **I 334**, 495-500.

Bousquet, O. (2003). Concentration inequalities for sub-additive functions using the entropy method. *Proceedings of Stochastic Inequalities 2002.* E. Gine and C. Houdre, eds. Birkhäuser, Boston. (in press). Available at *http://www.kyb.tuebingen.mpg.de/publication.html?user=bousquet.*

Borell, C. (1975). The Brunn-Minkowski inequality in Gauss-space. *Invent. Math.* **30**, 207-216.

Dembo, A. (1997). Information inequalities and concentration of measure. *Ann. Probability* **25** 927 - 939.

Diaconis, P. and Freedman, D. (1987). A dozen de Finetti-style results in search of a theory. *Ann. Inst. H. Poincaré* **1**, 397 - 423.

Erhard, A. (1983). Symmetrisation dans l'espace de Gauss. *Math. Scand.* **53**, 281-301.

Erhard, A. (1984). Inégalités isopérimétriques et intégrales de Dirichlet gaussiennes. *Ann. Scient. Ec. Norm. Sup.* **17**, 317 - 332.

Erhard, A. (1986). Elements extrémaux pour les inégalités de Brunn-Minkowski gaussiennes. *Ann. Inst. H. Poincaré* **22**, 149 - 168.

Giné, E., Koltchinskii, V., and Wellner, J. A. (2003). Ratio limit theorems for empirical processes. *Proceedings of Stochastic Inequalities 2002.* E. Giné and C. Houdre, eds. Birkhäuser, Boston. (in press). Available at *http://www.stat.washington.edu/jaw/RESEARCH/PAPERS/ available.html/ratios-BCN3-final.pdf.*

Klein, T. (2002). Une inégalité de concentration à gauche pour les processus empiriques. *C. R. Math. Acad. Sci. Paris* **334**, 501-504.

Ledoux, M. (1996a). Isoperimetry and Gaussian Analysis. *Ecole d'Eté de Probabilités de St-Flour*, *Lecture Notes in Math.* **1648**, 165 - 294.

Ledoux, M. (1996b). On Talagrand's deviation inequalities for product measures. *ESAIM: Probability and Statistics* **1**, 63 - 87. http://www.emath.fr/ps/.

Ledoux, M. (2001). *The Concentration of Measure Phenomenon. Mathematical Surveys and Monographs* **89**. American Mathematical Society.

Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces.* Springer-Verlag, Berlin. Chapter 1, pages 14 -34.

Massart, P. (2000a). About the constants in Talagrand's concentration inequalities for empirical processes. *Ann. Probability* **28**, 863-884.

Massart, P. (1998). Optimal constants for Hoeffding type inequalities. Preprint, Univerité de Paris-Sud, Mathematiques.

Massart, P. (2000b). Some applications of concentration inequalities to statistics. Probability theory. *Ann. Fac. Sci. Toulouse Math.* **6**, 245 - 303.

Rio, E. (2001). Inégalités de concentration pour les processus empiriques de classes de parties. *Probab. Theory Related Fields* **119**, 163 - 175.

Talagrand, M. (1989). Isoperimetry and integrability of the sum of independent Banach space valued random variables. *Ann. Probability* **17**, 1546-1570.

Talagrand, M. (1992). A new ispoerimetric inequality for product measure and the tails of sums of independent random variables. *Geometric and Functional Analysis* **1**, 211 - 223.

Talagrand, M. (1995). Concentration of measure and isoperimetric inequalities in product spaces. *Publications Mathématiques de l'I.H.E.S.* **81**, 73 - 205.

Talagrand, M. (1996). New concentration inequalities in product spaces. *Invent. Mathematicae* **126**, 505-563.

Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes.* Springer-Verlag, New York.