Electronic Journal of Statistics > Vol. 3 (2009)

open journal systems

A Bernstein-Von Mises Theorem for discrete probability distributions

Stéphane Boucheron, *Université Paris 7*
Elisabeth Gassiat, *Université Paris-Sud 11*

Abstract

We investigate the asymptotic normality of the posterior distribution in the discrete setting, when model dimension increases with sample size. We consider a probability mass function $theta_0$ on $mathbbm{N} setminus {0}$ and a sequence of truncation levels $(k_n)_n$ satisfying $k_n^3 leq n inf_{ileq k_n} theta_0(i).$ Let $hat{theta}$ denote the maximum likelihood estimate of $(theta_0(i))_{ileq k_n}$ and let $Delta_n(theta_0)$ denote the k_n-dimensional vector which i-th coordinate is defined by begin{math} sqrt{n} (hat{theta}_n(i)-theta_0(i)) end{math} for $1leq ileq k_n.$ We check that under mild conditions on $theta_0$ and on the sequence of prior probabilities on the k_n-dimensional simplices, after centering and rescaling, the variation distance between the posterior distribution recentered around $hat{theta}_n$ and rescaled by $sqrt{n}$ and the k_n-dimensional Gaussian distribution $mathcal{N}(Delta_n(theta_0), I^{-1}(theta_0))$ converges in probability to $0.$ This theorem can be used to prove the asymptotic normality of Bayesian estimators of Shannon and R'enyi entropies.
The proofs are based on concentration inequalities for centered and non-centered Chi-square (Pearson) statistics. The latter allow to establish posterior concentration rates with respect to Fisher distance rather than with respect to the Hellinger distance as it is commonplace in non-parametric Bayesian statistics.

AMS 2000 subject classifications: Primary 60K35, 60K35; secondary 60K35.

Keywords: Bernstein-Von Mises Theorem, Entropy estimation, non-parametric Bayesian statistics, Discrete models, Concentration inequalities.

Full Text: PDF

Boucheron, Stéphane, Gassiat, Elisabeth, A Bernstein-Von Mises Theorem for discrete probability distributions, Electronic Journal of Statistics, 3, (2009), 114-148 (electronic). DOI: 10.1214/08-EJS262.

References

A. Antos and I. Kontoyiannis. Convergence properties of functional estimates for discrete distributions. Random Struct. & Algorithms, 19(3-4): 163–193, 2001. MR1871554

S. Boucheron, A. Garivier, and E. Gassiat. Coding over infinite alphabets. IEEE Trans. Inform. Theory, 55:to appear, 2009. MR2236169

B. Clarke and A. Barron. Information-theoretic asymptotics of bayes methods. IEEE Trans. Inform. Theory, 36:453–471, 1990. MR1053841

B. Clarke and A. Barron. Jeffrey’s prior is asymptotically least favorable under entropy risk. J. Stat. Planning and Inference, 41:37–60, 1994. MR1292146

T. Cover and J. Thomas. Elements of information theory. John Wiley & sons, 1991. MR1122806

I. Csiszár and J. Körner. Information Theory: Coding Theorems for Discrete Memoryless Channels. Academic Press, 1981. MR0666545

J. Doob. Application of the theory of martingales. In Le Calcul des Probabilités et ses Applications., Colloques Internationaux du Centre National de la Recherche Scientifique, no. 13, pages 23–27. Centre National de la Recherche Scientifique, Paris, 1949. MR0033460

D. Dubhashi and D. Ranjan. Balls and bins: A study in negative dependence. Random Struct. & Algorithms, 13(2):99–124, 1998. MR1642566

R. M. Dudley. Real analysis and probability, volume 74 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2002. MR1932358

J. Fan. Local linear regression smoothers and their minimax efficiency. Annals of Statistics, 21:196–216, 1993. MR1212173

J. Fan and Y. K. Truong. Nonparametric regression with errors in variables. Annals of Statistics, 21(4):1900–1925, 1993. MR1245773

J. Fan, C. Zhang, and J. Zhang. Generalized likelihood ratio statistics and wilks phenomenon. Annals of Statistics, 29(1):153–193, 2001. MR1833962

D. A. Freedman. On the asymptotic behavior of Bayes' estimates in the discrete case. Ann. Math. Statist., 34:1386–1403, 1963. MR0158483

D. A. Freedman. On the asymptotic behavior of Bayes estimates in the discrete case. II. Ann. Math. Statist., 36:454–456, 1965. MR0174146

R. G. Gallager. Information theory and reliable communication. John Wiley & sons, 1968.

S. Ghosal. Asymptotic normality of posterior distributions for exponential families when the number of parameters tends to infinity. J. Multivariate Anal., 74(1):49–68, 2000. MR1790613

S. Ghosal and A. van der Vaart. Convergence rates of posterior distributions for non-i.i.d. observations. Annals of Statistics, 35(1):192–223, 2007a. MR2332274

S. Ghosal and A. van der Vaart. Posterior convergence rates of Dirichlet mixtures at smooth densities. Annals of Statistics, 35(2):697–723, 2007b. MR2336864

S. Ghosal and A. W. van der Vaart. Entropies and rates of convergence for maximum likelihood and Bayes estimation for mixtures of normal densities. Annals of Statistics, 29 (5):1233–1263, 2001.

S. Ghosal, J. Ghosh, and A. van der Vaart. Convergence rates of posterior distributions. Annals of Statistics, 28(2):500–531, 2000. MR1790007

L. Gyorfi, I. Pali, and E. van der Meulen. On universal noiseless source coding for infinite source alphabets. Eur. Trans. Telecommun. & Relat. Technol., 4(2):125–132, 1993.

J. C. Kieffer. A unified approach to weak universal source coding. IEEE Trans. Inform. Theory, 24(6):674–682, 1978. MR0514346

L. Le Cam and G. Yang. Asymptotics in Statistics: Some Basic Concepts. Springer, 2000. MR1784901

A. Lo. A large sample study of the Bayesian bootstrap. Ann. Statist., 15 (1):360–375, 1987. MR0885742

A. Lo. A Bayesian bootstrap for a finite population. Ann. Statist., 16(4): 1684–1695, 1988. MR0964946

P. Massart. Ecole d'Eté de Probabilité de Saint-Flour XXXIII, chapter Concentration inequalities and model selection. LNM. Springer-Verlag, 2003. MR2319879

L. Paninski. Estimating entropy on m bins given fewer than m samples. IEEE Trans. Inform. Theory, 50(9):2200–2203, 2004. MR2097210

S. Portnoy. Asymptotic behavior of likelihood methods for exponential families when the number of parameters tends to infinity. Annals of Statistics, 16:356–366, 1988. MR0924876

D. Rubin. The Bayesian bootstrap. Annals of Statistics, 9(1):130, 1981. MR0600538

L. Schwartz. On Bayes procedures. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 4:10–26, 1965. MR0184378

A. Tsybakov. Introduction à l'estimation non-paramétrique, volume 41 of Mathématiques & Applications. Springer-Verlag, Berlin, 2004. MR2013911

A. van der Vaart. The statistical work of Lucien Le Cam. Annals of Statistics, 30(3):631–682, 2002. MR1922537

A. van der Vaart. Asymptotic statistics. Cambridge University Press, 1998. MR1652247

A. van der Vaart and J. Wellner. Weak convergence and empirical processes. Springer Series in Statistics. Springer-Verlag, New York, 1996. MR1385671

C.-S. Weng. On a second-order asymptotic property of the Bayesian bootstrap mean. Ann. Statist., 17(2):705–710, 1989. MR0994261

Electronic Journal of Statistics. ISSN: 1935-7524