INTERNATIONAL CENTRE FOR ECONOMIC RESEARCH
ICER

WORKING PAPER SERIES

Antonio Lijoi, Igor Prünster, and Stephen G. Walker

ON CONSISTENCY OF NONPARAMETRIC NORMAL MIXTURES
FOR BAYESIAN DENSITY ESTIMATION

Working Paper no. 23/2004
October 2004

APPLIED MATHEMATICS
WORKING PAPER SERIES

On consistency of nonparametric normal mixtures for Bayesian density estimation

Antonio Lijoi[1], Igor Prünster[1,2] and Stephen G. Walker[3]

[1] Dipartimento Economia Politica e Metodi Quantitatavi, Università degli Studi di Pavia, via San Felice 5, 27100 Pavia, Italy. E-mail: lijoi@unipv.it, igor.pruenster@unipv.it

[2] ICER, Villa Gualino, viale Settimio Severo 63, 10133 Torino, Italy.

[3] Institute of Mathematics, Statistics and Actuarial Science, University of Kent, Kent CT2 7NZ, UK. E-mail: S.G.Walker@kent.ac.uk

Abstract. The past decade has seen a remarkable development in the area of Bayesian nonparametric inference both from a theoretical and applied perspective. As for the latter, the celebrated Dirichlet process has been successfully exploited within Bayesian mixture models leading to many interesting applications. As for the former, some new discrete nonparametric priors have been recently proposed in the literature: their natural use is as alternatives to the Dirichlet process in a Bayesian hierarchical model for density estimation. When using such models for concrete applications, an investigation of their statistical properties is mandatory. Among them a prominent role is to be assigned to consistency. Indeed, strong consistency of Bayesian nonparametric procedures for density estimation has been the focus of a considerable amount of research and, in particular, much attention has been devoted to the normal mixture of Dirichlet process. In this paper we improve on previous contributions by establishing strong consistency of the mixture of Dirichlet process under fairly general conditions: besides the usual Kullback–Leibler support condition, consistency is achieved by finiteness of the mean of the base measure of the Dirichlet process and an exponential decay of the prior on the standard deviation. We show that the same conditions are sufficient for mixtures based on priors more general than the Dirichlet process as well. This leads to the easy establishment of consistency for many recently proposed mixture models.

KEY WORDS: Bayesian nonparametrics, Density estimation, Mixture of Dirichlet process, Normal mixture model, Random discrete distribution, Strong consistency.

1. INTRODUCTION.

Consistency of Bayesian nonparametric procedures has been the focus of a considerable amount of research in recent years. Most contributions in the literature exploit the "frequentist" approach to Bayesian consistency, also termed the "what if" method according to Diaconis and Freedman (1986). It essentially consists in verifying what would happen to the posterior distribution if the data are generated from a "true" fixed density function f_0: does the posterior accumulate in suitably defined neighbourhoods of f_0?

Early papers on consistency are concerned with weak consistency. The reader is referred to, for example, Freedman (1963) and Diaconis and Freedman (1986) for some interesting examples of possible inconsistency. A sufficient condition for weak consistency, which is solely a support condition, is provided in Schwartz (1965).

When considering problems of density estimation, it is natural to ask for the strong consistency of posterior distributions. An early contribution in this area is represented by Barron (1988). Later developments combine techniques from the theory of empirical processes with results on uniformly consistent tests achieved in Barron (1988) and provide sufficient conditions for strong consistency relying upon the construction of suitable sieves. General results are derived in Barron, Schervish and Wasserman (1999) and Ghosal, Ghosh and Ramamoorthi (1999), whereas significant priors are studied in Petrone and Wasserman (2002), Choudhuri, Ghosal and Roy (2004), among others. The "sieve–approach" is treated in great detail in the monograph by Ghosh and Ramamoorthi (2003). See also Wasserman (1998) for a more concise account. Recently, a new approach to the study of strong consistency for Bayesian density estimation has been introduced in Walker (2004) where a simple sufficient condition for strong consistency, not relying on sieves, is obtained.

In the framework of Bayesian density estimation, one is naturally led to think of the mixture of Dirichlet process (MDP), a cornerstone in the area. This model was introduced by Lo (1984) and, later, popularized by Escobar (1988) and Escobar and West (1995) by the development of suitable simulation techniques. See also MacEachern (1994) and MacEachern and Müller (1998). The MDP is extensively reviewed in the book edited by Dey, Müller and Sinha (1998) and in Quintana and Müller (2004), where emphasis is put on applications and simulation algorithms. As far as consistency is concerned, the normal MDP model has been first analyzed in Ghosal et al. (1999) by exploiting the sieve–approach.

In this paper, we face the issue of consistency of the MDP by exploiting the approach set out in Walker (2004). This leads to a quite dramatic improvement on previous results. We essentially show that a MDP model is consistent if the base measure of the Dirichlet process has finite mean and the prior on the standard deviation has an exponentially decaying tail in a neighbourhood of 0. Our results carry over to normal mixture models, where the Dirichlet process is replaced by a general discrete nonparametric prior, thus establishing consistency of many models that have been recently proposed in the literature.

The paper is structured as follows. In Section 2, after a concise description of the normal mixture model, we state the main result. Afterwards we provide illustrations describing how the result applies to a variety of nonparametric priors. Finally, Section 3 contains a detailed proof.

2. THE CONSISTENCY RESULT.

2.1. The Bayesian normal mixture model.

Nowadays the most common use of Bayesian nonparametric procedures is represented by density estimation via a mixture model based on a random discrete distribution. In particular, attention has been focussed on normal mixtures, that is

$$\tilde{f}_{\sigma,\tilde{P}}(x) = \phi_\sigma * \tilde{P} = \int \phi_\sigma(x-\theta)\,\tilde{P}(\mathrm{d}\theta) \tag{1}$$

where, for each positive σ, ϕ_σ is the density function of the normal distribution with mean zero and variance σ^2. $\tilde{P}$ is a random probability distribution on $\mathbb{R}$ whose law, Λ, selects discrete distributions (almost surely). Moreover, σ has a prior distribution which we denote by μ. The model (1) can be equivalently expressed in hierarchical form as

$$\begin{aligned}
(X_i \mid \theta_i, \sigma_i) &\overset{\text{ind}}{\sim} N(X_i; \theta_i, \sigma_i^2), \quad i = 1, \dots, n \\
(\theta_i \mid \tilde{P}) &\overset{\text{iid}}{\sim} \tilde{P} \quad i = 1, \dots, n \\
\tilde{P} &\sim \Lambda \\
\sigma_i &\overset{\text{iid}}{\sim} \mu \quad i = 1, \dots, n.
\end{aligned}$$

where μ and Λ are independent and $N(\,\cdot\,;\theta,\sigma^2)$ stands for the normal distribution with mean θ and variance σ^2. Clearly, the MDP model is obtained when $\tilde{P}$ in (1) coincides with the Dirichlet process with parameter–measure α. Here α is a finite non-null measure, see, for example, Ghosh and Ramamoorthi (2003).

An important element in prior specification which will be considered later is the prior guess at the shape of $\tilde{P}$, that this

$$P_0(\,\cdot\,) = E\left[\tilde{P}(\,\cdot\,)\right].$$

2.2. A sufficient condition for strong consistency.

The relevance, both theoretical and applied, of normal mixture models motivates a study of their asymptotic properties. Among these properties, consistency plays a prominent role. Since the aim is density estimation, the appropriate notion to deal with is strong consistency. Consider a sequence of observations $(X_n)_{n\geq 1}$ each taking values in $\mathbb{R}$ and let $\mathscr{F}$ be the space of probability density functions with respect to the Lebesgue measure on $\mathbb{R}$. If Π stands for the (prior) distribution of the random density function $\tilde{f}_{\sigma,\tilde{P}}$ in (1), the observations X_n's are i.i.d. from f, given f selected from Π. The posterior distribution, given the observations $(X_1, \dots, X_n)$, coincides with

$$\Pi_n(B) = \frac{\int_B \prod_{i=1}^n f(X_i)\,\Pi(\mathrm{d}f)}{\int_{\mathscr{F}} \prod_{i=1}^n f(X_i)\,\Pi(\mathrm{d}f)}$$

for all measurable subsets B of $\mathscr{F}$. Hence, Π is said to be *strongly consistent* at f_0 in $\mathscr{F}$ if, for any $\varepsilon > 0$,

$$\Pi_n(A_\varepsilon) \to 1 \qquad \text{a.s. } [F_0^\infty]$$

as $n \to +\infty$, where A_ε is an L_1-neighbourhood of f_0 with radius ε, F_0 denotes the probability distribution whose density coincides with f_0 and F_0^∞ indicates the infinite product measure on $\mathbb{R}^\infty$.

Hereafter we will assume that the density f_0 is in the Kullback–Leibler support of the prior Π. This means that Π assigns positive masses to any Kullback–Leibler neighbourhood of f_0. It is known that such an assumption is sufficient to ensure *weak consistency* of Π at f_0. See Schwartz (1965). Conditions for f_0 to be in the Kullback–Leibler support of the normal mixture model prior Π defined in (1) are given in Ghosal et al. (1999). However, since we aim at establishing the stronger property of L_1–consistency, the Kullback–Leibler support condition is not enough.

Note that Π is determined both by Λ and by the prior distribution for σ, which we have denoted by μ. As for the latter, from the point of view of consistency the most important values of σ are those included in a right–neighbourhood of zero. Thus, with no loss of generality, we can choose μ such that its support coincides with $(0, M]$ for some positive and finite M.

The main result on strong consistency of normal mixture models can be now stated. In the sequel $g(x) \sim h(x)$ as x tends to $+\infty$ means that $g(x)/h(x)$ tends to 1 as x tends to $+\infty$. The proof to the following can be found in Section 3.

Theorem 1 *Let f_0 be a density in the Kullback–Leibler support of* Π. *Suppose*

(i) $\int_{\mathbb{R}} |\theta|\, P_0(\mathrm{d}\theta) < +\infty$

(ii) $\mu\{\sigma < \sigma_k\} \le \exp\{-\gamma k\}$ *for some sufficiently large γ, where $(\sigma_k)_{k\ge 1}$ is any sequence such that $\sigma_k \sim k^{-1}$ as $k \to \infty$.*

Then Π *is consistent at f_0.*

By the above result, strong consistency follows from a simple condition on the prior guess P_0 combined with a condition on the probabilities assigned by μ on shrinking neighbourhoods of the origin. Note that the value for which γ can be considered sufficiently large is determined in the proof. See (7). Theorem 1 can be compared with the results obtained in Ghosal et al. (1999) for the MDP. Their theorem has three conditions (i)–(iii). Indeed, our condition (i) improves on their condition (i), which essentially requires α to have exponential tails. Moreover, our condition (ii) and their condition (ii) coincide. Finally, we have no need for their condition (iii).

Some comments on our (i) are in order. Notice, for instance, that it is satisfied even by heavy–tailed distributions, i.e. by those P_0's for which $P_0([-\theta, \theta]^c) \sim \theta^{-\gamma}$, for some

$\gamma > 1$, as $\theta \to +\infty$. Weakening the tail condition for P_0 from an exponential to a power law decay seems to be a quite remarkable achievement.

2.3. Illustrations.

In this section we show how condition (i) translates for a variety of normal mixture models, thus giving a simple criterion for establishing their (strong) consistency. It is worth stressing that strong consistency for the more general mixtures we are going to consider has not yet been considered in the literature. Note that Theorem 1 applies also to mixture models directed by random probability measures $\tilde{P}$ whose support contains continuous distributions. However, such cases seem not to be of particular interest since, commonly in applications, one wishes to exploit the clustering behaviour arising from discrete $\tilde{P}$.

First of all, recall that the celebrated MDP is recovered by setting $\tilde{P}$ to be the Dirichlet process with parameter–measure α. In this case, $P_0 = \alpha/\alpha(\mathbb{R})$ and condition (i) reduces to

$$\int_{\mathbb{R}} |\theta| \, \alpha(\mathrm{d}\theta) < \infty. \tag{2}$$

Recently a new class of random probability has been derived in Regazzini, Lijoi and Prünster (2003) via the normalization of random measures with independent increments. Such random probabilities include, as a special case, the Dirichlet process and will be denoted by the acronym NRMI. A wider class of mixtures can, then, be achieved by setting $\tilde{P}$ to be a NRMI as done in Nieto–Barajas, Prünster and Walker (2004). It can be shown that the prior guess reduces to

$$P_0(\mathrm{d}\theta) = \alpha(\mathrm{d}\theta) \int_0^{+\infty} \mathrm{e}^{-\psi(u)} \left\{ \int_0^{+\infty} \mathrm{e}^{-uv} \, v \, \rho(\mathrm{d}v|\theta) \, \mathrm{d}v \right\} \mathrm{d}u \tag{3}$$

where $\nu_\alpha(\mathrm{d}v, \mathrm{d}\theta) = \rho(\mathrm{d}v|\theta) \, \alpha(\mathrm{d}\theta)$ is the Poisson intensity measure on $(0, +\infty) \times \mathbb{R}$ associated with the increasing additive process ξ which generates $\tilde{P}$. Moreover, ψ stands for the Laplace exponent of ξ which can be determined via the well-known Lévy–Khintchine representation theorem. For details see Regazzini et al. (2003) and James (2002). When $\rho(\mathrm{d}v|\theta) = \rho(\mathrm{d}v)$, for each $\theta \in \mathbb{R}$, $\tilde{P}$ is said to be homogeneous, the prior guess in (3) reduces to $\alpha/\alpha(\mathbb{R})$ and condition (i) coincides with (2). Apart from the Dirichlet process, the most notable prior within this class, which leads to explicit forms for quantities of statistical interest, is the so–called normal inverse Gaussian process studied in Lijoi, Mena and Prünster (2004).

Another interesting class of mixture models, first considered in Ishwaran and James (2001), arises when $\tilde{P}$ is chosen to be a stick–breaking prior. Such a prior depends upon the specification of a stick–breaking procedure and of a measure α which is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$. The prior guess P_0 coincides with $\alpha/\alpha(\mathbb{R})$ and, again, (i) becomes (2). Among these priors it is worth mentioning the two parameter Poisson–Dirichlet process. See Pitman (1996). We wish to remark that both the stick–breaking and homogeneous NRMI priors essentially belong to the class of species sampling models due to Pitman (1996), for which (i) is again equivalent to (2). However, the two classes we have been describing above are the only ones for which it is possible to assess effectively the weights in the species sampling representation. Hence, they are the most useful for concrete applications. Species sampling mixture models are dealt with in Ishwaran and James (2003).

These illustrations stress the usefulness of Theorem 1 in checking consistency of normal mixture models based on a number of alternatives to the Dirichlet process as a mixing distribution.

3. THE PROOF.

3.1. Preliminary result.

Recall that A_ε^c is the complement of the L_1–neighborhood of f_0 with radius ε. By separability of $\mathscr{F}$, such a set can be covered by a countable union of disjoint sets B_j, where $B_j \subseteq B_j^* := \{f : \|f - f_j\| < \eta\}$, f_j are densities in A_ε^c, $\|\cdot\|$ is the L_1–norm and η is any number in $(0, \varepsilon)$. An extension of a result in Walker (2004) can be stated as follows: if, for some $\beta \in (0,1)$ and for some covering $(B_j)_{j\geq 1}$ as above,

$$\sum_{j\geq 1} (\Pi(B_j))^\beta < +\infty, \tag{4}$$

then Π is consistent at f_0 with the proviso that f_0 is in the Kullback–Leibler support of Π. This result will be a key ingredient in the following proof.

3.2. The proof.

Let us first set some useful notation. For any $a > 0$ and $\sigma > 0$, let

$$\mathscr{F}^U_{\sigma,a,\delta} = \{\phi_\sigma * P : \, P([-a,a]) \geq 1-\delta\} \quad \text{and} \quad \mathscr{F}^L_{\sigma,a,\delta} = \{\phi_\sigma * P : \, P([-a,a]) < 1-\delta\}$$

and $\overline{\mathscr{F}}^M_{\sigma,a,\delta} = \bigcup_{\sigma<\sigma'<M} \mathscr{F}^U_{\sigma',a,\delta}$. For $\mathscr{G} \subset \mathscr{F}$ and $\eta > 0$, define $J(\eta, \mathscr{G})$ to be the L_1–metric entropy of the set $\mathscr{G}$. This means that $J(\eta,\mathscr{G})$ is the logarithm of the minimum number of L_1–balls of radius η which cover $\mathscr{G}$. From Ghosal et al. (1999) one has

$$J(\delta, \overline{\mathscr{F}}^M_{\sigma,a,\delta}) \leq \frac{Ca}{\sigma},$$

where C depends only on M and δ.

Take $(a_n)_{n\geq1}$ to be any increasing sequence of positive numbers such that $\lim_n a_n = +\infty$ and let $(\sigma_n)_{n\geq1}$ be a decreasing sequence of positive numbers such that $\lim_n \sigma_n = 0$. For our purposes it is useful to consider sets of the type

$$\mathscr{G}_{\sigma,a_j,\delta} := \{\phi_\sigma * P : \, P([-a_j,a_j]) \geq 1-\delta, \;\; P([-a_{j-1},a_{j-1}]) < 1-\delta\}.$$

These sets are pairwise disjoint and $\lim_j \mathscr{G}_{\sigma,a_j,\delta} = \varnothing$ for any positive σ and δ. This definition entails the following inclusions: $\mathscr{G}_{\sigma,a_j,\delta} \subset \mathscr{F}^U_{\sigma,a_j,\delta}$ and $\mathscr{G}_{\sigma,a_j,\delta} \subset \mathscr{F}^L_{\sigma,a_{j-1},\delta}$. Moreover, $\mathscr{F}^L_{\sigma,a_j,\delta} \downarrow \varnothing$ as j tends to $+\infty$. Thus, for any $\eta > 0$ there exists an integer N such that for any $j \geq N$

$$J(\eta, \mathscr{F}^L_{\sigma,a_j,\delta}) \leq J(\eta, \mathscr{F}^U_{\sigma,a_N,\delta}).$$

Set

$$\mathscr{G}^M_{\sigma_k,a_j,\delta} = \bigcup_{\sigma_k<\sigma<M} \mathscr{G}_{\sigma,a_j,\delta},$$

and note that

$$\bigcup_{j,k\geq 1} \mathscr{G}^M_{\sigma_k,a_j,\delta} = \mathscr{F}.$$

Since $\mathscr{G}^M_{\sigma_k,a_j,\delta}$ is included in $\bigcup_{\sigma_k<\sigma<M} \mathscr{F}^L_{\sigma,a_j,\delta}$, one has

$$J(\eta, \mathscr{G}^M_{\sigma_k,a_j,\delta}) \leq \frac{Ca_N}{\sigma_k} \tag{5}$$

for any $j \geq N$. On the other hand, the inclusion $\mathscr{G}^M_{\sigma_k,a_j,\delta} \subset \overline{\mathscr{F}}^M_{\sigma_k,a_j,\delta}$ entails that (5) holds true also for any $j < N$. These findings can be summarized by saying that $\mathscr{G}^{\sigma_{k-1}}_{\sigma_k,a_j,\delta}$ has a finite η–covering $\{C_{j,k,l} : \, l = 1, 2, \ldots, N_{j,k}\}$ where $N_{j,k} \leq [\exp(Ca_N/\sigma_k)] + 1$. Here we use $[x]$ to denote the integer part of a real number x. Define, now, the following sets

$$B_{j,\delta} = \{P : \, P([-a_j, a_j]) \geq 1 - \delta, \;\; P([-a_{j-1}, a_{j-1}]) < 1 - \delta\} \qquad j \geq 1.$$

The condition for convergence (4) would be implied by

$$\begin{aligned} \sum_{j,k\geq 1} \sum_{l=1}^{N_{j,k}} (\Pi(C_{j,k,l}))^\beta &\leq \sum_{j,k\geq 1} N_{j,k} \left\{ \Pi(\mathscr{G}^{\sigma_{k-1}}_{\sigma_k,a_j,\delta}) \right\}^\beta \\ &\leq \sum_{k\geq 1} \mathrm{e}^{\frac{Ca_N}{\sigma_k}} \, \{\mu(\sigma_k < \sigma \leq \sigma_{k-1})\}^\beta \sum_{j\geq 1} \{\Lambda(B_{j,\delta})\}^\beta < +\infty, \end{aligned} \tag{6}$$

where $\sigma_0 = M$. Consider now the part concerning the mixing measure Λ above. Let $A_j = (-\infty, -a_{j-1}) \cup (a_{j-1}, +\infty)$ and note that

$$B_{j,\delta} \subset \{P : \, P(A_j) > \delta'\}$$

with $\delta' > \delta$. Hence, by Markov's inequality

$$\Lambda(B_{j,\delta}) \leq \Lambda(\{P : \, P(A_j) > \delta'\}) \leq \frac{1}{\delta'} \, P_0(A_j),$$

and (6) becomes

$$\sum_{k\geq 1} \mathrm{e}^{\frac{C'}{\sigma_k}} \, \{\mu(\sigma_k < \sigma \leq \sigma_{k-1})\}^\beta \sum_{j\geq 1} \{P_0(A_j)\}^\beta < +\infty.$$

At this stage, we can fix $a_j \sim j$ as $j \to +\infty$. Condition (i) is then equivalent to $P_0(A_j) = O(j^{-(1+r)})$ which, in turn, ensures the convergence of $\sum_{j\geq 1}\{P_0(A_j)\}^\beta$ for any β such that $(1+r)^{-1} < \beta < 1$. Moreover, take

$$\gamma > C'/\beta \tag{7}$$

so that condition (ii) implies the prior μ to be such that

$$\sum_{k\geq 1} \mathrm{e}^{\frac{C'}{\sigma_k}} \{\mu(\sigma_k < \sigma \leq \sigma_{k-1})\}^\beta$$

converges. The proof of Theorem 1 is complete.

ACKNOWLEDGMENTS

A. Lijoi and I. Prünster were partially supported by the Italian Ministry of University and Research (MIUR), research project "Bayesian nonparametric methods and their applications". The work of S.G. Walker is financed by an EPSRC Advanced Research Fellowship.

REFERENCES

Barron, A. (1988). The exponential convergence of posterior probabilities with implications for Bayes estimators of density functions. Technical Report N. 7, Dept. Statistics, Univ. Illinois, Champaign.

Barron, A., Schervish, M.J. and Wasserman, L. (1999). The consistency of distributions in nonparametric problems. *Ann. Statist.* **27**, 536–561.

Choudhuri, N., Ghosal, S. and Roy, A. (2004). Bayesian estimation of the spectral density of a time series. *J. Amer. Stat. Assoc.*, to appear.

Dey, D., Müller, P. and Sinha, D. (1998). *Practical Nonparametric and Semiparametric Bayesian Statistics.* Springer, New York.

Diaconis, P. and Freedman, D. (1986). On the consistency of Bayes estimates. *Ann. Statist.* **14**, 1–26.

Escobar, M.D. (1988). *Estimating the means of several normal populations by nonparametric estimation of the distribution of the means.* Unpublished Ph.D. dissertation, Department of Statistics, Yale University.

Escobar, M.D. and West, M.(1995). Bayesian density estimation and inference using mixtures. *J. Amer. Statist. Assoc.* **90**, 577–588.

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

Ghosal, S., Ghosh, J.K. and Ramamoorthi, R.V. (1999). Posterior consistency of Dirichlet mixtures in density estimation. *Ann. Statist.* **27**, 143–158.

Ghosh, J.K. and Ramamoorthi, R.V. (2003). *Bayesian Nonparametrics.* Springer, New York.

Ishwaran, H. and James, L. F. (2001). Gibbs sampling methods for stick-breaking priors. *J. Amer. Stat. Assoc.*, **96**, 161–173.

Ishwaran, H. and James, L. F. (2003). Generalized weighted Chinese restaurant processes for species sampling mixture models. *Statist. Sinica*, **13**, 1211–1235.

James, L.F. (2002). Poisson Process Partition Calculus with applications to exchangeable models and Bayesian Nonparametrics. *Mathematics ArXiv*, math.PR/0205093.

Lijoi, A., Mena, R.H. and Prünster, I. (2004) Hierarchical mixture modelling with normalized inverse Gaussian priors. Tech. Report, ICER, Turin.

Lo, A.Y. (1984). On a class of Bayesian nonparametric estimates: I Density estimates. *Ann. Statist.* **12**, 351–357.

MacEachern, S.N. (1994). Estimating normal means with a conjugate style Dirichlet process prior. *Commun. Statist. Simulation Comp.*, **23**, 727–741.

MacEachern, S.N. and Müller, P. (1998). Estimating Mixture of Dirichlet Process Models. *Journal of Computational and Graphical Statistics*, **7**, 223–239.

Müller, P. and Quintana, F. A. (2004). Nonparametric Bayesian data analysis. *Statist. Sci.* **19**, 95–110.

Nieto-Barajas, L.E., Prünster, I. and Walker, S.G. (2004). Normalized Random Measures driven by Increasing Additive Processes. *Ann. Statist.*, to appear.

Petrone, S. and Wasserman, L. (2002). Consistency of Bernstein polynomial posteriors. *J. Roy. Stat. Soc. B* **64**, 79–100.

Pitman, J. (1996). Some developments of the Blackwell-MacQueen urn scheme. In *Statistics, Probability and Game Theory. Papers in honor of David Blackwell* (Eds. Ferguson, T.S., *et al.*). Lecture Notes, Monograph Series, **30**, 245–267. Institute of Mathematical Statistics, Hayward.

Regazzini, E., Lijoi, A. and Prünster, I. (2003). Distributional results for means of random measures with independent increments. *Ann. Statist.* **31**, 560-585.

Schwartz, L. (1965). On Bayes procedures. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **4**, 10–26.

Walker, S.G. (2004). New approaches to Bayesian consistency. *Ann. Statist.* **32**, 2028–2043.

Wasserman, L. (1998). Asymptotic properties of nonparametric Bayesian procedures. In *Practical Nonparametric and Semiparametric Bayesian Statistics* (D. Dey, P. Muller and D. Sinha, eds.), 293–304. Lecture Notes in Statistics N.133, Springer, New York.

INTERNATIONAL CENTRE FOR ECONOMIC RESEARCH
APPLIED MATHEMATICS WORKING PAPER SERIES

1. Luigi Montrucchio and Fabio Privileggi, "On Fragility of Bubbles in Equilibrium Asset Pricing Models of Lucas-Type," *Journal of Economic Theory* 101, 158-188, 2001 (ICER WP 2001/5).
2. Massimo Marinacci, "Probabilistic Sophistication and Multiple Priors," *Econometrica* 70, 755-764, 2002 (ICER WP 2001/8).
3. Massimo Marinacci and Luigi Montrucchio, "Subcalculus for Set Functions and Cores of TU Games," *Journal of Mathematical Economics* 39, 1-25, 2003 (ICER WP 2001/9).
4. Juan Dubra, Fabio Maccheroni, and Efe Ok, "Expected Utility Theory without the Completeness Axiom," *Journal of Economic Theory*, forthcoming (ICER WP 2001/11).
5. Adriana Castaldo and Massimo Marinacci, "Random Correspondences as Bundles of Random Variables," April 2001 (ICER WP 2001/12).
6. Paolo Ghirardato, Fabio Maccheroni, Massimo Marinacci, and Marciano Siniscalchi, "A Subjective Spin on Roulette Wheels," *Econometrica*, forthcoming (ICER WP 2001/17).
7. Domenico Menicucci, "Optimal Two-Object Auctions with Synergies," July 2001 (ICER WP 2001/18).
8. Paolo Ghirardato and Massimo Marinacci, "Risk, Ambiguity, and the Separation of Tastes and Beliefs," *Mathematics of Operations Research* 26, 864-890, 2001 (ICER WP 2001/21).
9. Andrea Roncoroni, "Change of Numeraire for Affine Arbitrage Pricing Models Driven By Multifactor Market Point Processes," September 2001 (ICER WP 2001/22).
10. Maitreesh Ghatak, Massimo Morelli, and Tomas Sjoström, "Credit Rationing, Wealth Inequality, and Allocation of Talent", September 2001 (ICER WP 2001/23).
11. Fabio Maccheroni and William H. Ruckle, "BV as a Dual Space," *Rendiconti del Seminario Matematico dell'Università di Padova*, 107, 101-109, 2002 (ICER WP 2001/29).
12. Fabio Maccheroni, "Yaari Dual Theory without the Completeness Axiom," *Economic Theory*, forthcoming (ICER WP 2001/30).
13. Umberto Cherubini and Elisa Luciano, "Multivariate Option Pricing with Copulas," January 2002 (ICER WP 2002/5).
14. Umberto Cherubini and Elisa Luciano, "Pricing Vulnerable Options with Copulas," January 2002 (ICER WP 2002/6).
15. Steven Haberman and Elena Vigna, "Optimal Investment Strategies and Risk Measures in Defined Contribution Pension Schemes," *Insurance: Mathematics and Economics* 31, 35-69, 2002 (ICER WP 2002/10).
16. Enrico Diecidue and Fabio Maccheroni, "Coherence without Additivity," *Journal of Mathematical Psychology*, forthcoming (ICER WP 2002/11).
17. Paolo Ghirardato, Fabio Maccheroni, and Massimo Marinacci, "Ambiguity from the Differential Viewpoint," April 2002 (ICER WP 2002/17).

18. Massimo Marinacci and Luigi Montrucchio, “A Characterization of the Core of Convex Games through Gateaux Derivatives,” *Journal of Economic Theory*, forthcoming (ICER WP 2002/18).
19. Fabio Maccheroni and Massimo Marinacci, “How to Cut a Pizza Fairly: Fair Division with Decreasing Marginal Evaluations,” *Social Choice and Welfare*, 20, 457-465, 2003 (ICER WP 2002/23).
20. Erio Castagnoli, Fabio Maccheroni and Massimo Marinacci, “Insurance Premia Consistent with the Market,” *Insurance: Mathematics and Economics* 31, 267-284, 2002 (ICER WP 2002/24).
21. Fabio Privileggi and Guido Cozzi, “Wealth Polarization and Pulverization in Fractal Societies,” September 2002 (ICER WP 2002/39).
22. Paolo Ghirardato, Fabio Maccheroni, and Massimo Marinacci, “Certainty Independence and the Separation of Utility and Beliefs,” December 2002 (ICER WP 2002/40).
23. Salvatore Modica and Marco Scarsini, “The Convexity-Cone Approach to Comparative Risk and Downside Risk”, January 2003 (ICER WP 2003/1).
24. Claudio Mattalia, “Existence of Solutions and Asset Pricing Bubbles in General Equilibrium Models”, January 2003 (ICER WP 2003/2).
25. Massimo Marinacci and Luigi Montrucchio, “Cores and Stable Sets of Finite Dimensional Games”, March 2003 (ICER WP 2003/7).
26. Jerome Renault, Sergio Scarlatti, and Marco Scarsini, “A Folk Theorem for Minority Games”, April 2003 (ICER WP 2003/10).
27. Peter Klibanoff, Massimo Marinacci, and Sujoy Mukerji, “A Smooth Model of Decision Making under Ambiguity”, April 2003 (ICER WP 2003/11).
28. Massimo Marinacci and Luigi Montrucchio, “Ultramodular Functions”, June 2003 (ICER WP 2003/13).
29. Erio Castagnoli, Fabio Maccheroni, and Massimo Marinacci, “Choquet Insurance Pricing: a Caveat”, June 2003 (ICER WP 2003/14).
30. Thibault Gajdos and Eric Maurin, “Unequal Uncertainties and Uncertain Inequalities: an Axiomatic Approach, June 2003 (ICER WP 2003/15).
31. Thibault Gajdos and John A. Weymark, “Multidimensional Generalized Gini Indices”, June 2003 (ICER WP 2003/16).
32. Thibault Gajdos, Jean-Marc Tallon, and Jean-Christophe Vergnaud, “Decision Making with Imprecise Probabilistic Information”, June 2003 (ICER WP 2003/18).
33. Alfred Müller and Marco Scarsini, “Archimedean Copulae and Positive Dependence”, July 2003 (ICER WP 2003/25).
34. Bruno Bassan, Olivier Gossner, Marco Scarsini, and Shmuel Zamir, “Positive Value of Information in Games”, *International Journal of Game Theory*, forthcoming (ICER WP 2003/26).
35. Marco Dall’Aglio and Marco Scarsini, “Zonoids, Linear Dependence, and Size-Biased Distributions on the Simplex”, *Advances in Applied Probability* forthcoming (ICER WP 2003/27).
36. Taizhong Hu, Alfred Muller, and Marco Scarsini, “Some Counterexamples in Positive Dependence”, *Journal of Statistical Planning and Inference*, forthcoming (ICER WP 2003/28).
37. Massimo Marinacci, Fabio Maccheroni, Alain Chateauneuf, and Jean-Marc Tallon, “Monotone Continuous Multiple Priors” August 2003 (ICER WP 2003/30).

38. Massimo Marinacci and Luigi Montrucchio, “Introduction to the Mathematics of Ambiguity”, October 2003 (ICER WP 2003/34).
39. Fabio Maccheroni, Massimo Marinacci and Aldo Rustichini, “Variational representation of preferences under ambiguity”, June 2004 (ICER WP 2004/5).
40. Antonio Lijoi, Ramsés H. Mena, and Igor Prünster, “Hierarchical mixture modelling with normalized Gaussian priors”, November 2004 (ICER WP 2004/12).
41. Stephen G. Walker, Antonio Lijoi, and Igor Prünster, “Contributions to the understanding of Bayesian consistency”, November 2004 (ICER WP 2004/13).
42. Antonio Lijoi, Igor Prünster, and Stephen G. Walker, “On consistency of nonparametric normal mixtures”, October 2004 (ICER WP 2004/23).
43. Stephen G. Walker, Antonio Lijoi, and Igor Prünster, “On rates of convergence for posterior distributions”, October 2004 (ICER WP 2004/24).