The Annals of Statistics
1999, Vol. 27, No. 1, 361–373

EATON'S MARKOV CHAIN, ITS CONJUGATE PARTNER AND $\mathscr{P}$-ADMISSIBILITY

BY JAMES P. HOBERT[1] AND C. P. ROBERT

University of Florida and CREST, INSEE

Suppose that X is a random variable with density $f(x|\theta)$ and that $\pi(\theta|x)$ is a proper posterior corresponding to an *improper* prior $\nu(\theta)$. The prior is called $\mathscr{P}$-admissible if the generalized Bayes estimator of every bounded function of θ is almost-ν-admissible under squared error loss. Eaton showed that recurrence of the Markov chain with transition density $R(\eta|\theta) = \int \pi(\eta|x) f(x|\theta)\,dx$ is a sufficient condition for $\mathscr{P}$-admissibility of $\nu(\theta)$. We show that Eaton's Markov chain is recurrent if and only if its conjugate partner, with transition density $\tilde{R}(y|x) = \int f(y|\theta)\pi(\theta|x)\,d\theta$, is recurrent. This provides a new method of establishing $\mathscr{P}$-admissibility. Often, one of these two Markov chains corresponds to a standard stochastic process for which there are known results on recurrence and transience. For example, when X is Poisson(θ) and an improper gamma prior is placed on θ, the Markov chain defined by $\tilde{R}(y|x)$ is equivalent to a branching process with immigration. We use this type of argument to establish $\mathscr{P}$-admissibility of some priors when f is a negative binomial mass function and when f is a gamma density with known shape.

1. Introduction. We will be interested in the following univariate decision problem. Suppose that X is a single observation from the probability density (or mass) function $f(x|\theta)$ whose support $\mathscr{X} \subseteq \Re$ does not depend on the parameter θ, which we assume belongs to the set $\Theta \subseteq \Re$. Suppose that $\nu(\theta)$ is an improper prior density (or mass) function with support Θ that yields a proper posterior, $\pi(\theta|x)$; that is, we assume that for all $x \in \mathscr{X}$,

$$m(x) := \int_\Theta f(x|\theta)\nu(\theta)\,d\theta < \infty.$$

This setup is a special case of the decision problem considered by Eaton (1992), who first introduced what we call $\mathscr{P}$-admissibility.

DEFINITION 1. The prior, $\nu(\theta)$, is called $\mathscr{P}$-admissible if the generalized Bayes estimate of every bounded function of θ is almost-ν-admissible [Stein (1965)] under squared error loss.

Of course, the generalized Bayes estimate of $g(\theta)$ under squared error loss is the posterior expectation

$$\delta_g(x) := \int_\Theta g(\theta)\pi(\theta|x)\,d\theta.$$

Received August 1997; revised October 1998.

[1]Supported in part by the Center for Research in Economics and Statistics (CREST) at the French National Institute of Statistics and Economic Studies (INSEE), Paris, France.

AMS 1991 *subject classifications*. Primary 62C15; secondary 60J05.

Key words and phrases. Bilinear model, branching process with immigration, exponential family, improper prior, null recurrence, random walk, stochastic difference equation, transience.

Also, since ν has support Θ, almost-ν-admissibility implies admissibility as long as all finite valued risk functions are continuous [Berger (1985), Section 8.8].

Note that $\mathscr{P}$-admissibility is an endorsement of the improper prior itself and not just an individual estimator based on the prior. This is quite different from a typical result concerning admissibility where a *single* function of the parameter is of interest, and an improper prior is judged indirectly through its generalized Bayes estimator. See Eaton (1982) for more on this point of view.

Eaton (1992) established that $\mathscr{P}$-admissibility is implied by the local-ν-recurrence of a Markov chain constructed using f and ν. This sufficient condition is an essential tool for establishing $\mathscr{P}$-admissibility because a direct proof is often quite difficult to obtain. The Markov chain is now described. Suppose for the moment that f and ν are both density functions and note that, for fixed θ, the function $\pi(\eta|x)f(x|\theta)$ is a joint density in the variable (η, x) with support $\Theta \times \mathscr{X}$. Therefore, for any $\theta \in \Theta$,

$$R(\eta|\theta) := \int_{\mathscr{X}} \pi(\eta|x)f(x|\theta)\,dx \tag{1}$$

is a density function in the variable η with support Θ. Let $W = (W_0, W_1, W_2, \dots)$ be a time-homogeneous Markov chain on the infinite product space Θ^∞ with transition density given by $R(\cdot|\cdot)$; that is,

$$P(W_t \in A|w_{t-1}) = \int_A R(\eta|w_{t-1})\,d\eta.$$

Eaton (1992) showed that local-ν-recurrence of this Markov chain, which we call the R-chain, implies $\mathscr{P}$-admissibility of $\nu(\theta)$. More specifically, he developed a sufficient condition, which we call $\mathscr{C}$, for $\mathscr{P}$-admissibility of ν and then showed that $\mathscr{C}$ is equivalent to local-ν-recurrence of the R-chain. The condition $\mathscr{C}$ is described in Eaton (1992), page 1149, Theorem 1.1. [See Eaton (1997) for an introduction to these ideas.] Because we are not directly interested in $\mathscr{C}$, it will not be stated explicitly here.

Two different notions of recurrence are considered in this paper: Eaton's (1992) local-ν-recurrence and the Meyn and Tweedie (1993), page 496 version, which is referred to simply as "recurrence." As we explain in Section 2, recurrence implies local-ν-recurrence for the models considered in this paper. Thus, for these models, recurrence of the R-chain is a sufficient condition for $\mathscr{P}$-admissibility of ν.

EXAMPLE 1.1. Let $X|\theta \sim N(\theta, 1)$ and $\nu(\theta) \propto \exp\{-b\theta^2/2 + ab\theta\}$. Then the posterior is normal with mean $(x+ab)/(b+1)$ and variance $1/(b+1)$. Thus, we have an improper prior yielding a proper posterior whenever $b \in (-1, 0]$. Using (1) it follows that

$$\eta|\theta \sim N\left(\frac{\theta+ab}{b+1}, \frac{b+2}{(b+1)^2}\right).$$

Therefore, the R-chain can be written as an autoregressive process of order one; that is, for $t = 0, 1, 2, \ldots,$

$$W_{t+1} = \frac{1}{b+1} W_t + U_{t+1},$$

where $U_1, U_2, U_3, \ldots$ is an independent and identically distributed (iid) sequence of normal random variables with mean $ab/(b+1)$ and variance $(b+2)/(b+1)^2$. When $b = 0$, this is a standard random walk with a mean-zero increment, and is therefore recurrent. When $b \in (-1, 0)$, the coefficient of W_t is larger than 1, and W is transient, that is, not recurrent [Meyn and Tweedie (1993), page 221]. Thus, $\mathscr{P}$-admissibility holds when $b = 0$, that is, when the prior is Lebesgue measure. Indeed, Eaton (1992), page 1157, shows that, under very minimal conditions, $\mathscr{P}$-admissibility holds when Lebesgue measure is used as a prior for a translation parameter. We return to this example below.

Other works regarding relationships between admissibility and recurrence properties include Brown (1971), who related admissibility of estimators of the multivariate normal mean to recurrence properties of associated diffusions, and Johnstone (1984), who reported similar results for Poisson means and associated birth and death processes [see also Johnstone (1986)]. Both of these authors considered continuous time Markov processes on the *sample space* while, in contrast, Eaton (1992) looked at discrete time Markov chains on the *parameter space*. We show that an obvious counterpart to the R-chain, which lives on the sample space, is recurrent if and only if the R-chain is recurrent. Therefore, at least in our context, Eaton's results can be viewed as involving a Markov chain on the sample space. The counterpart of Eaton's R-chain is now described.

Using arguments similar to those above, it follows that for any $x \in \mathscr{X}$,

$$\tilde{R}(y|x) := \int_{\Theta} f(y|\theta)\pi(\theta|x)\, d\theta \tag{2}$$

is a density function in the variable y with support $\mathscr{X}$. Let $V = (V_0, V_1, V_2, \ldots)$ be a time-homogeneous Markov chain on the infinite product space $\mathscr{X}^\infty$ with transition density $\tilde{R}(\cdot|\cdot)$. Eaton (1992) did not study this Markov chain, which we call the $\tilde{R}$-chain, but did mention it (page 1171).

EXAMPLE 1.1 (*Continued*). Using (2), we find that the transitions for the $\tilde{R}$-chain are given by

$$y|x \sim N\left(\frac{x+ab}{b+1}, \frac{b+2}{b+1}\right).$$

Thus, the $\tilde{R}$-chain is equal in distribution to the following autoregressive process of order one:

$$V_{t+1} = \frac{1}{b+1} V_t + U'_t,$$

where $U'_1, U'_2, U'_3, \dots$ is an iid sequence of normal random variables with mean $ab/(b+1)$ and variance $(b+2)/(b+1)$. Except for a slight difference in the variance of the error sequence $[(b+2)/(b+1)^2$ versus $(b+2)/(b+1)]$, the R and $\tilde{R}$-chains are the same in this case, and it follows that the $\tilde{R}$-chain is recurrent when $b = 0$ and transient when $b \in (-1, 0)$.

While the R and $\tilde{R}$-chains are always either both recurrent or both transient (see Section 2), the similarity in form exhibited in the previous example is not typical. In many situations one of the chains is much easier to analyze than the other, and this is the case in the following example.

EXAMPLE 1.2. Suppose that $X|\theta \sim \text{Poisson}(\theta)$ and $\nu(\theta) \propto \theta^{a-1}e^{-b\theta}$. Then $\theta|x \sim \text{Gamma}(x+a, b+1)$. Thus, we have an improper prior yielding a proper posterior when $a > 0$ and $b \in (-1, 0]$. A simple calculation shows that

$$R(\eta|\theta) = \sum_{x=0}^{\infty} \frac{e^{-\theta}\theta^x}{x!} \frac{(b+1)^{x+a}}{\Gamma(x+a)} \eta^{x+a-1} \exp(-\eta(b+1)) \tag{3}$$

for $\theta, \eta \in (0, \infty)$. Feller [(1971), page 58] calls (3) a *randomized gamma* density. When $b = -1/2$ and a is a multiple of $1/2$, it is the noncentral chi-squared density. Eaton [(1992), page 1165] considered the case in which $b = 0$ and proved directly that the condition $\mathscr{C}$ holds when $a \in (0, 1]$. It follows that, in this specific case, $\mathscr{P}$-admissibility holds and the R-chain is locally-ν-recurrent.

In Remark 5.1, Eaton (1992) reverifies local-ν-recurrence of the R-chain (when $b = 0$ and $a \in (0, 1]$), using a theorem of Lamperti (1960). There is a problem with this application, however, as Lamperti's regularity condition (3.11) is not satisfied. In a Ph.D. dissertation, Lai (1996) proved a modified version of Lamperti's theorem and used it to reverify local-ν-recurrence of the R-chain when $b = 0$ and $a \in (0, 1]$. (Lai also shows how this theorem can be used to establish admissibility results in the multivariate Poisson problem.) See Kersting (1986) for another extension of Lamperti's (1960) results.

In contrast to the rather complicated form of the R-chain in this example, the $\tilde{R}$-chain has a simple interpretation, and results concerning its recurrence and transience have been around for over 25 years. The transition density for the $\tilde{R}$-chain is given by

$$\tilde{R}(y|x) = \frac{\Gamma(y+x+a)}{y!\Gamma(x+a)} p^{x+a}(1-p)^y,$$

where $p = (b+1)/(b+2)$ and $x, y \in \{0, 1, 2, \dots\}$. This is a *generalized* negative binomial mass function [Feller (1968), page 269]. Indeed, when a is a positive integer, it is the usual negative binomial mass function. If Z is a random variable supported on the nonnegative integers and

$$P(Z = z) = \frac{\Gamma(z+c)}{z!\Gamma(c)} d^c (1-d)^z$$

for $d \in (0, 1)$ and $c > 0$, we write $Z \sim \text{NB}(c, d)$. Note that $E[Z] = c(1-d)/d$ and $\text{Var}(Z) = c(1-d)/d^2$. It follows from the form of the probability generating function that if $Z_1, Z_2, \ldots, Z_n$ are independent random variables with $Z_i \sim \mathrm{NB}(c_i, d)$, then $\sum Z_i \sim \mathrm{NB}(\sum c_i, d)$. Using these facts, it is clear that the $\tilde{R}$-chain can be written as a *branching process with immigration*; that is,

$$V_{t+1} = \sum_{i=1}^{V_t} N_{i,t} + M_{t+1},$$

where $N_{1,t}, N_{2,t}, \ldots, N_{V_t,t}$ are iid $\mathrm{NB}(1, p)$ and M_{t+1} (independent of the $N_{i,t}$'s) is $\mathrm{NB}(a, p)$. Think of the chain as follows. At generation t, there are V_t animals in the population. Each animal, independently of all others, has a random number of offspring whose distribution is $\mathrm{NB}(1, p)$. Also at generation t, a random number ($\mathrm{NB}(a, p)$) of animals migrate into the society. The population at the $(t+1)$th generation consists of all of those offspring and the immigrants. The mean and variance of the offspring distribution are $1/(b+1)$ and $(b+2)/(b+1)^2$, respectively.

Results in Pakes (1971) show that this branching process is recurrent if $b = 0$ and $a \in (0, 1]$ and is transient otherwise. This is intuitively reasonable because when $b \in (-1, 0)$, the mean number of offspring per animal is larger than one (supercritical case) and the population explodes. On the other hand, when $b = 0$ each animal averages a single offspring (critical case), and the stability of the population depends upon the rate of immigration. We have extended Eaton's (1992) analysis of this example by showing that $\mathscr{P}$-admissibility cannot be established through recurrence of the R-chain when $b = 0$ and $a > 1$ nor when $b \in (-1, 0)$.

The rest of the paper is laid out as follows. The theorem showing that the R and $\tilde{R}$-chains are always either both recurrent or both transient is presented in Section 2. In Section 3 we consider the case in which X has a gamma density with known shape parameter and an improper gamma prior is placed on the unknown scale. Here the $\tilde{R}$-chain turns out to be a *bilinear model* [Meyn and Tweedie (1993), page 30], and recent results of Babillot, Bougerol and Elie (1997) can be used to establish recurrence, and hence, $\mathscr{P}$-admissibility. Section 4 studies the case of the $\mathrm{NB}(k, p)$ distribution with k known and an improper beta prior on the unknown success probability, p. A transformation of the R-chain gives another bilinear model and the results of Babillot, Bougerol and Elie (1997) can again be applied. Because the success probability is bounded, our results yield a class of admissible estimators of p. Finally, conclusions and avenues for future research are discussed in Section 5.

2. Recurrence duality. Let $\mathscr{B}(\Theta)$ and $\mathscr{B}(\mathscr{X})$ be the appropriate σ-algebras on Θ and $\mathscr{X}$; that is, the Borel σ-algebra if the set is uncountable and the set of all subsets in the countable case. Let μ^{Θ} be Lebesgue or counting measure on $\mathscr{B}(\Theta)$ as appropriate, and define $\mu^{\mathscr{X}}$ similarly. Put $S = \mathscr{X} \times \Theta$ and let μ^S be the product measure on the product σ-algebra $\mathscr{B}(S)$.

Let (V_n, W_n), $n = 0, 1, 2, \ldots$, be the bivariate, discrete time, time homogeneous Markov chain on the product space S^{∞} defined by the Markov transition density

$$M(y, \eta | x, \theta) = \pi(\eta | y) f(y | \theta), \tag{4}$$

where we use (x, θ) and (y, η), instead of (v_n, w_n) and (v_{n+1}, w_{n+1}), respectively, to avoid excessive subscripting. This Markov chain, which we call the *extended* chain, is similar to that used in data augmentation [Tanner and Wong (1987); Liu, Wong and Kong (1994)]. For any fixed $(x, \theta) \in S$, $M(y, \eta | x, \theta)$ is strictly positive on S. Thus, the extended chain is μ^S-irreducible and aperiodic. Let $j(x, \theta) = f(x|\theta)\nu(\theta)$ and note that

$$j(y, \eta) = \int_S M(y, \eta | x, \theta) j(x, \theta) \mu^S(d(x, \theta)), \tag{5}$$

which shows that j is an invariant density for the extended chain.

It is clear from (4) that given W_n, (V_{n+1}, W_{n+1}) is conditionally independent of V_n; that is, x does not appear on the right side of (4). Similarly, given V_n, (W_n, V_{n+1}) is conditionally independent of W_{n-1}. As a result, $\{W_n\colon n = 0, 1, 2, \ldots\}$ and $\{V_n\colon n = 0, 1, 2, \ldots\}$ are both univariate Markov chains and, following Liu, Wong and Kong (1994), we call them *conjugate* Markov chains. Indeed, $\{W_n\colon n = 0, 1, 2, \ldots\}$ is the R-chain and its Markov transition density is

$$R(\eta|\theta) = \int_{\mathscr{X}} \pi(\eta|x) f(x|\theta) \mu^{\mathscr{X}}(dx).$$

It follows directly from the μ^S-irreducibility of the extended chain that the R-chain is μ^Θ-irreducible and aperiodic. The R-chain satisfies the *detailed balance condition*

$$R(\eta|\theta)\nu(\theta) = R(\theta|\eta)\nu(\eta) \tag{6}$$

from which it follows that $\nu(\eta) = \int_\Theta R(\eta|\theta)\nu(\theta)\mu^\Theta(d\theta)$; that is, the prior, ν, is an invariant density for the R-chain. [Another way to establish that ν is invariant is to integrate both sides of (5) with respect to y.] We note in passing that Eaton refers to a chain satisfying (6) as *ν-symmetric*, while other authors [e.g., Lyons (1983)] use the term *reversible* despite the fact that $\nu(\cdot)$ is improper [Kelly (1979), page 5].

Analogously, $\{V_n\colon n = 0, 1, 2, \ldots\}$ is the $\tilde{R}$-chain whose transition density is

$$\tilde{R}(y|x) = \int_\Theta f(y|\theta)\pi(\theta|x)\mu^\Theta(d\theta).$$

The $\tilde{R}$-chain is $\mu^{\mathscr{X}}$-irreducible and aperiodic, and has $m(\cdot)$ as an invariant density. Note that the invariant densities for the R and $\tilde{R}$-chains are the marginals of the invariant density for the extended chain.

We now formally define recurrence. Let $\mathscr{B}^+(\Theta)$ be the class of sets in $\mathscr{B}(\Theta)$ with positive measure [and define $\mathscr{B}^+(\mathscr{X})$ similarly]. The R-chain is *recurrent* if, for any $A \in \mathscr{B}^+(\Theta)$ and any starting value,

$$\sum_{n=1}^{\infty} E[I_A(W_n)] = \infty.$$

(Of course, recurrence of the $\tilde{R}$-chain is defined analogously.) This is a standard definition of recurrence for irreducible Markov chains [Meyn and Tweedie (1993), page 496]. In order to "... circumvent a discussion of irreducibility issues ...," Eaton (1992) used a slightly different version called local-ν-recurrence, which is defined in terms of hitting times. Under our assumptions, recurrence implies local-ν-recurrence. This follows from a comparison of Eaton's (1992) Definition A.2, with Meyn and Tweedie's [(1993), page 497] hitting time characterization of recurrence and the fact that our prior measure is absolutely continuous with respect to μ^{Θ}.

A Markov chain that is irreducible and recurrent possesses a unique (up to constant multiples) invariant measure. When the invariant measure is finite, the chain is called *positive* [Meyn and Tweedie (1993), Chapter 10]. A consequence of uniqueness is that a Markov chain possessing an invariant measure with infinite mass cannot be positive recurrent. Such a chain is either null recurrent or transient. Thus, the R-chain is positive recurrent only when ν is proper and, similarly, the $\tilde{R}$-chain is positive recurrent only when m is proper. [This makes sense from a decision-theoretic standpoint because unique proper Bayes estimators are admissible [Lehmann and Casella (1998), page 323] and Eaton's (1992) results deal with the more challenging case where ν is improper.] Since m is proper if and only if ν is proper, it follows that the R-chain is positive recurrent if and only if the $\tilde{R}$-chain is positive recurrent. This correspondence can be viewed as a special case of the *duality principle* [Diebolt and Robert (1994)]. Our main result shows that this connection between the chains can be extended.

THEOREM 1. *The R-chain is recurrent if and only if the $\tilde{R}$-chain is recurrent.*

PROOF. Suppose that the $\tilde{R}$-chain is recurrent. For $A \in \mathscr{B}^{+}(\Theta)$, there exists an $\varepsilon > 0$ such that $\mu^{\mathscr{X}}(B_{\varepsilon}) > 0$ where

$$B_{\varepsilon} := \left\{x \in \mathscr{X}\colon \int_A \pi(\eta|x)\mu^{\Theta}(d\eta) > \varepsilon\right\}.$$

This is true since if no such ε exists, then $\int_A \pi(\eta|v)\mu^{\Theta}(d\eta) = 0$ almost everywhere ($\mu^{\mathscr{X}}$), which contradicts the fact that $A \in \mathscr{B}^{+}(\Theta)$. Therefore,

$$\begin{aligned}\sum_{n=1}^{\infty} E[I_A(W_n)] &= \sum_{n=1}^{\infty} E[E[I_A(W_n)|V_n]]\\ &\geq \sum_{n=1}^{\infty} E[I_{B_{\varepsilon}}(V_n)E[I_A(W_n)|V_n]]\\ &\geq \varepsilon \sum_{n=1}^{\infty} E[I_{B_{\varepsilon}}(V_n)] = \infty.\end{aligned}$$

This shows that recurrence of the $\tilde{R}$-chain implies recurrence of the R-chain. The reverse implication can be shown using the same type of argument. $\square$

Theorem 1 shows that, at least for the models that we consider, Eaton's (1992) result may be regarded as concerning a Markov chain on the sample space, which is the domain of the Markov processes constructed by Brown (1971) and Johnstone (1984). More importantly, our result is useful in situations where recurrence of the $\tilde{R}$-chain is easier to establish than recurrence of the R-chain. In Example 1.2, the $\tilde{R}$-chain turned out to be a standard Markov chain for which the transience–recurrence behavior is already known. The next section concerns a similar example.

3. Gamma-gamma model. Suppose that $X|\theta \sim \text{Gamma}(\alpha, \theta)$ ($\alpha > 0$ known) and $\nu(\theta) \propto \theta^{a-1}e^{-b\theta}$. Then $\theta|x \sim \text{Gamma}(\alpha+a, b+x)$. In this case, we have an improper prior yielding a proper posterior when $b = 0$ and $a > -\alpha$ and when $b > 0$ and $a \in (-\alpha, 0]$. Most of the admissibility and domination results for gamma models concern estimators of θ or the scale parameter $1/\theta$ [e.g., Berger (1980, 1985), pages 255, 305; Das Gupta (1984)], whereas our results concern $\mathscr{P}$-admissibility of improper conjugate priors.

The transition density for the R-chain is given by

$$R(\eta|\theta) = \frac{\theta^{\alpha}\eta^{\alpha+a-1}\exp(-\eta b)}{\Gamma(\alpha+a)\Gamma(\alpha)} \int_0^{\infty} (b+x)^{\alpha+a} x^{\alpha-1} \exp(-x(\eta+\theta))\, dx,$$

for $\eta, \theta \in (0, \infty)$. This density involves an integral that cannot be written in closed form unless $b = 0$. On the other hand, the transition density for the $\tilde{R}$-chain has a closed form expression for all values of a and b,

$$\tilde{R}(y|x) = \frac{\Gamma(2\alpha+a)(b+x)^{\alpha+a}}{\Gamma(\alpha+a)\Gamma(\alpha)} \frac{y^{\alpha-1}}{(x+y+b)^{2\alpha+a}}$$

for $x, y \in (0, \infty)$. We now state the result.

THEOREM 2. *The Markov chains are recurrent if and only if* $a = 0$. *Thus,* $\mathscr{P}$-*admissibility holds when* $a = 0$.

PROOF. Consider the $\tilde{R}$-chain. By noting that the random variable $Z = Y/(x+b)$ has a density that is free of x, we may write the $\tilde{R}$-chain as a *bilinear model* [Meyn and Tweedie (1993), page 30],

$$V_{t+1} = (V_t + b)Z_{t+1}, \tag{7}$$

where $Z_1, Z_2, Z_3, \dots$ is an iid sequence of random variables with density

$$f_Z(z) = \begin{cases} \dfrac{\Gamma(2\alpha+a)}{\Gamma(\alpha+a)\Gamma(\alpha)} \dfrac{z^{\alpha-1}}{(z+1)^{2\alpha+a}}, & \text{if } z > 0, \\ 0, & \text{otherwise.} \end{cases}$$

The random variable $(\alpha+a)Z/\alpha$ has an $F(2\alpha, 2(\alpha+a))$ distribution, which implies that Z has an infinite mean whenever $\alpha + a \le 1$.

First, consider the case $b = 0$. A log transformation of (7) leads to the following random walk on $\Re$:

$$L_{t+1} = L_t + \log Z_{t+1}, \tag{8}$$

where $L_t = \log V_t$, $t = 0, 1, 2, \dots$. Suppose $B \in \mathscr{B}(\mathscr{X})$, then $V_t \in B$ if and only if $L_t \in B^*$ where $B^* = \{y \in \Re\colon e^y \in B\}$. Then since B has positive Lebesgue measure if and only if B^* has positive Lebesgue measure, it follows that (7) and (8) are either both recurrent or both transient.

Feller [(1971), page 50] shows that Z has the same distribution as $B^{-1} - 1$ where B is Beta$(\alpha + a, \alpha)$, from which it follows that the moment generating function of $\log Z$ exists. Therefore, the random walk (8) is recurrent if and only if $\log Z$ has mean zero [Meyn and Tweedie (1993), page 247]. Now,

$$\begin{aligned} E[\log Z] &= \frac{\Gamma(2\alpha + a)}{\Gamma(\alpha + a)\Gamma(\alpha)}\left[\int_0^1 \frac{z^{\alpha-1}\log z}{(z+1)^{2\alpha+a}}\,dz + \int_1^\infty \frac{z^{\alpha-1}\log z}{(z+1)^{2\alpha+a}}\,dz\right] \\ &= \frac{\Gamma(2\alpha + a)}{\Gamma(\alpha + a)\Gamma(\alpha)}\int_0^1 \frac{z^{\alpha-1}\log z}{(z+1)^{2\alpha+a}}(1 - z^a)\,dz, \end{aligned}$$

which implies that the mean of $\log Z$ is negative when $a > 0$, zero when $a = 0$ and positive when $a < 0$. Thus, (7) is recurrent when $a = b = 0$ and transient when $b = 0$ and $a \neq 0$.

We now present a result of Babillot, Bougerol and Elie (1997) that will enable us to deal with the remaining cases; that is, when $b > 0$. Let $\{(A_t, B_t)\}_{t \geq 1}$ be an iid sequence of random variables taking values in $\Re_+ \times \Re$. Consider the Markov chain $X_0, X_1, X_2, \dots,$ defined by the stochastic difference equation $X_{t+1} = A_{t+1}X_t + B_{t+1}$, $t = 0, 1, 2, \dots$. The following proposition is part of Babillot, Bougerol and Elie's (1997) Corollary 4.2.

PROPOSITION 1. *Suppose that*:

(i) *For all* $x \in \Re$, $P(A_1 x + B_1 = x) < 1$;
(ii) *For some* $\delta > 0$,

$$E[(|\log A_1| + (\log|B_1|)^+)^{2+\delta}] < \infty,$$

where $x^+ = \max(0, x)$; *and*
(iii) $E[\log A_1] = 0$ *and* $P(A_1 = 1) < 1$.

Then the Markov chain $X_0, X_1, X_2, \dots$ *is recurrent*.

When $E[\log A_1] < 0$, the chain is positive recurrent [Brandt (1986)]. On the other hand, if $E[\log A_1] > 0$, then the random walk $\sum_{i=1}^t \log A_i$, and hence $\prod_{i=1}^t A_i$, diverge to infinity w.p. 1. Note that

$$X_t = X_0 \prod_{i=1}^t A_i + B_t + \sum_{i=1}^{t-1} B_i \prod_{j=i+1}^t A_j.$$

Therefore, when $E[\log A_1] > 0$ and $P(B_1 \geq 0) = 1$, the Markov chain $X_0, X_1, X_2, \dots$ is transient.

Now consider again the $\tilde{R}$-chain. When $b > 0$, (7) fits into the Babillot, Bougerol and Elie (1997) framework with $A_t = Z_t$ and $B_t = bZ_t$. Assumptions (i) and (ii) of Proposition 1 are clearly satisfied, and we conclude that, in the case of strictly positive b, the $\tilde{R}$-chain is recurrent when $a = 0$ and transient when $a \in (-\alpha, 0)$. □

Note that when $\alpha = 1$, $f(x|\theta)$ is an exponential density, which is a scale family, and the Markov chains are recurrent under the prior $\nu(\theta) = 1/\theta$. This is actually a special case of a general result in Eaton [(1992), page 1159] concerning scale families and the right invariant Haar prior [Berger (1985), page 409].

4. Negative binomial-beta model. Suppose that $X|\theta \sim \mathrm{NB}(k, \theta)$ and that $\nu(\theta) \propto \theta^{a-1}(1-\theta)^{b-1}$. Then $\theta|x \sim \mathrm{Beta}(a+k, b+x)$, and we have an improper prior yielding a proper posterior as long as $b > 0$ and $a \in (-k, 0]$.

THEOREM 3. *For $b \geq k$, the Markov chains are recurrent if and only if $a = 0$. Thus, $\mathscr{P}$-admissibility holds when $b \geq k$ and $a = 0$.*

PROOF. The transition density for the R-chain is given by

$$R(\eta|\theta) = \frac{\eta^{k+a-1}(1-\eta)^{b-1}\theta^k}{\Gamma(k+a)\Gamma(k)} \sum_{x=0}^{\infty} \frac{\Gamma(k+a+b+x)\Gamma(x+k)}{x!\Gamma(b+x)} [(1-\eta)(1-\theta)]^x$$

for $\eta, \theta \in (0,1)$. When $k = b$, the summand is the kernel of a $\mathrm{NB}(k+a+b, 1-(1-\eta)(1-\theta))$ mass function, and we have

$$R(\eta|\theta) = \frac{\Gamma(k+a+b)}{\Gamma(k+a)\Gamma(k)} \frac{\eta^{k+a-1}(1-\eta)^{b-1}\theta^k}{(1-(1-\eta)(1-\theta))^{k+a+b}}.$$

By noting that the density of $\eta\theta^{-1}(1-\eta)^{-1}$ is free of θ, we may write the R-chain as a *nonlinear state space model* [Meyn and Tweedie (1993), page 29]

$$W_{t+1} = \frac{W_t}{Z_{t+1} + W_t}, \tag{9}$$

where $Z_1, Z_2, Z_3, \ldots$ is an iid sequence of random variables such that $(k+a)Z_1/k$ has an $F(2k, 2(k+a))$ distribution. Letting $I_t = 1/W_t$, (9) becomes

$$I_{t+1} = I_t Z_{t+1} + 1$$

for $t = 0, 1, 2, \ldots$ and by an argument similar to that used in Section 3, this Markov chain is recurrent if and only if the R-chain is recurrent. We know from Section 3 that $E[\log Z] = 0$ when $a = 0$ and is strictly positive when $a < 0$. Appealing to Proposition 1, the chain I_t, and hence the R-chain, are recurrent when $a = 0$ and transient when $a < 0$.

Now consider the case in which $b > k$. Let $\eta_1 \sim \mathrm{Beta}(a+k, b-k)$ and $\eta_2 \sim \mathrm{Beta}(a+b, k+x)$. It is straightforward to show that if η_1 and η_2 are independent, then $\eta_1\eta_2$ has a Beta$(a+k, b+x)$ distribution. We can therefore write the R-chain as

$$W_{t+1} = B_{t+1}\frac{W_t}{Z_{t+1}+W_t}, \tag{10}$$

where $Z_1, Z_2, Z_3, \dots$ is an iid sequence of random variables such that $(a+b)Z_1/k$ has an $F(2k, 2(a+b))$ distribution, and $B_1, B_2, B_3, \dots$ is an iid sequence of Beta$(a+k, b-k)$ random variables, which are independent of the Z's. Again, letting $I_t = 1/W_t$, (10) becomes

$$I_{t+1} = I_t\frac{Z_{t+1}}{B_{t+1}} + \frac{1}{B_{t+1}}$$

for $t = 0, 1, 2, \dots$. The behavior of this chain depends upon the expectation of the logarithm of Z_{t+1}/B_{t+1}. Using moment generating functions, one can show that

$$E\left[\log\left(\frac{Z_{t+1}}{B_{t+1}}\right)\right] = \psi(k) - \psi(a+k),$$

where ψ is the derivative of the log gamma function; that is, $\psi(x) = \Gamma'(x)/\Gamma(x)$. Now, since $\psi(\cdot)$ is increasing, $E[\log(Z_{t+1}/B_{t+1})] = 0$ when $a = 0$ and is strictly positive when $a \in (-k, 0)$. Another application of Proposition 1 yields the result. □

Most of the admissibility and domination results for negative binomial models concern estimators of θ [e.g., Hwang (1982a, b)]. Since θ is a bounded parameter in this example, Theorem 3 shows that if $b \geq k$,

$$\delta(X) = k/(X+k+b)$$

is an admissible estimator of θ under squared error loss. Although admissibility of generalized Bayes estimates of the negative binomial success probability based on conjugate priors seems like an obvious question, we were unable to find this result in the literature. (A direct computation of the generalized Bayes risk shows that it is finite, but $\mathscr{P}$-admissibility implies that admissibility also holds for every bounded transform of θ.)

Finally, the transition density for the $\tilde{R}$-chain is given by

$$\tilde{R}(y|x) = \frac{\Gamma(y+k)\Gamma(k+a+b+x)\Gamma(2k+a)\Gamma(b+x+y)}{y!\Gamma(k)\Gamma(k+a)\Gamma(b+x)\Gamma(2k+a+b+x+y)}$$

for $y, x \in \{0, 1, 2, \dots\}$. It is called the *generalized Waring* distribution [Johnson, Kotz and Kemp (1992), page 242]. When b and k are positive integers and $a \in \{-k+1, -k+2, \dots, -1, 0\}$, the $\tilde{R}$-chain has a Pólya urn representation [Panaretos and Xekalaki (1986)] which is now described. Suppose that V_t represents the current state of the Markov chain. Consider an urn containing $k+a$ white balls and k black balls. A ball is drawn at random, its color is noted and the ball is replaced along with one additional ball of the same color before the next ball is drawn. This procedure is continued until $V_t + b$ white balls have been drawn, and V_{t+1} is defined to be the number of black balls drawn before the $(V_t + b)$th white ball is drawn. Note that a negative value of a would mean more black balls than white at the start. Thus, it makes sense that the chain is recurrent when $a = 0$ and transient when a is negative.

5. Discussion. We have shown that Eaton's (1992) Markov chain is recurrent if and only if its conjugate partner is recurrent. This result may be viewed as an extension of Diebolt and Robert's (1994) duality principle. It is interesting from a theoretical standpoint in that Eaton's (1992) results can now be viewed as concerning a Markov chain on the sample space, which is the domain of the Markov processes constructed by Brown (1971) and Johnstone (1984). From a practical point of view, this result is useful because it allows one to prove $\mathscr{P}$-admissibility by establishing the recurrence of the $\tilde{R}$-chain, which can be much easier than doing the same for the R-chain. Finally, we have used Eaton's (1992) theory and our extensions to establish that certain priors for the gamma scale parameter and the negative binomial success probability are $\mathscr{P}$-admissible.

Acknowledgments. The authors are grateful to George Casella, William E. Strawderman and Martin T. Wells for suggesting an investigation of the $\tilde{R}$-chain during conversations at Cornell University in the summer of 1995. The authors are also grateful to Andrew Rosalsky for some helpful discussions and to James O. Berger and an anonymous referee for constructive comments and suggestions.

REFERENCES

BABILLOT, M., BOUGEROL, P. and ELIE, L. (1997). The random difference equation $X_n = A_n X_{n-1} + B_n$ in the critical case. *Ann. Probab.* **25** 478–493.

BERGER, J. O. (1980). Improving on inadmissible estimators in continuous exponential families with applications to simultaneous estimation of gamma scale parameters. *Ann. Statist.* **8** 545–571.

BERGER, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Springer, New York.

BRANDT, A. (1986). The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Adv. Appl. Probab.* **18** 211–220.

BROWN, L. D. (1971). Admissible estimators, recurrent diffusions, and insoluble boundary value problems. *Ann. Math. Statist.* **42** 855–904.

DAS GUPTA, A. (1984). Admissibility in the gamma distribution: two examples. *Sankhyā Ser. A* **46** 395–407

DIEBOLT, J. and ROBERT, C. P. (1994). Estimation of finite mixture distributions by Bayesian sampling. *J. Roy. Statist. Soc. Ser. B* **56** 363–375.

EATON, M. L. (1982). A Method for evaluating improper prior distributions. In *Statistical Decision Theory and Related Topics III* **1** (S. S. Gupta and J. O. Berger, eds.). Academic, New York, 329–352.

EATON, M. L. (1992). A statistical diptych: admissible inferences-recurrence of symmetric Markov chains. *Ann. Statist.* **20** 1147–1179.

EATON, M. L. (1997). Admissibility in quadratically regular problems and recurrence of symmetric Markov chains: Why the connection? *J. Statist. Plann. Inference* **64** 231–247.

FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications* **1**, 3rd ed. Wiley, New York.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications* **2**, 2nd ed. Wiley, New York.

HWANG, J. T. (1982). Improving upon standard estimators in discrete exponential families with applications to Poisson and negative binomial cases. *Ann. Statist.* **10** 857–867.

HWANG, J. T. (1982). Semitail upper bounds on the class of admissible estimators in discrete exponential families with applications to Poisson and negative binomial distributions. *Ann. Statist.* **10** 1137–1147.

JOHNSON, N. L., KOTZ, S. and KEMP, A. W. (1992). *Univariate Discrete Distributions*, 2nd ed. Wiley, New York.

JOHNSTONE, I. (1984). Admissibility, difference equations, and recurrence in estimating a Poisson mean. *Ann. Statist.* **12** 1173–1198.

JOHNSTONE, I. (1986). Admissible estimation, dirichlet principles, and recurrence of birth–death chains on Z_t^p. *Probab. Theory Related Fields* **71** 231–269.

KELLY, F. P. (1979). *Reversibility and Stochastic Networks*. Wiley, New York.

KERSTING, G. (1986). On recurrence and transience of growth models. *J. Appl. Probab.* **23** 614–625.

LAI, W.-L. (1996). Admissibility and the recurrence of Markov chains with applications. Technical Report 612, School of Statistics, Univ. Minnesota.

LAMPERTI, J. (1960). Criteria for the recurrence or transience of stochastic processes I. *J. Math. Anal. Appl.* **1** 314–330.

LEHMANN, E. L. and CASELLA, G. (1998). *Theory of Point Estimation*, 2nd ed. Springer, New York.

LIU, J. S., WONG, W. H. and KONG, A. (1994). Covariance structure of the Gibbs sampler with applications to the comparisons of estimators and augmentation schemes. *Biometrika* **81** 27–40.

LYONS, T. (1983). A simple criterion for transience of a reversible Markov chain. *Ann. Probab.* **11** 393–402.

MEYN, S. P. and TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, London.

PAKES, A. G. (1971). On the critical Galton–Watson process with immigration. *J. Austral. Math. Soc.* **12** 476–482.

PANARETOS, J. and XEKALAKI, E. (1986). On some distributions arising from certain generalized sampling schemes. *Comm. Statist. Theory Methods* **15** 873–891.

STEIN, C. (1965). Approximation of improper prior measures by prior probability measures. In *Bernoulli–Bayes–Laplace Festschrift* (J. Neyman and L. Le Cam, eds.) 217–240. Springer, New York.

TANNER, M. A. and WONG, W. H. (1987). The calculation of posterior distributions by data augmentation (with discussion). *J. Amer. Statist. Assoc.* **52** 528–550.

DEPARTMENT OF STATISTICS
UNIVERSITY OF FLORIDA
GAINESVILLE, FLORIDA 32611
E-MAIL: jhobert@stat.ufl.edu

LABORATOIRE DE STATISTIQUE
CREST, INSEE
75675 PARIS CEDEX 14
FRANCE
E-MAIL: robert@ensae.fr