The Annals of Mathematical Statistics
1970, Vol. 41, No. 3, 1064–1067

A COMPARISON BETWEEN THE MARTIN BOUNDARY THEORY AND THE THEORY OF LIKELIHOOD RATIOS[1]

BY ALLAN F. ABRAHAMSE

University of Southern California

1. Summary. Given a sequence $\{X_n; n \geqq 0\}$ of random variables, let $\{P^\theta; \theta \in \pi\}$ be a parameterized class of probability measures, with respect to each of which the sequence is a Markov chain. Under conditions which make it appropriate to define likelihood ratios, the parameter set π can be identified with a subset of the Martin boundary for the space-time chain $\{(n, X_n); n \geqq 0\}$, so that each parameter θ can also be considered as a point of this Martin boundary. Then for each θ, the space-time sample paths converge in the Martin boundary topology to θ, almost surely with respect to the probability measure P^θ. Moreover, the likelihood ratio corresponding to the parameter θ is the same as the minimal regular function corresponding to the parameter θ, and the probability measure P^θ is the relativised probability measure corresponding to the point θ.

2. Results. Let $(\Omega, \mathscr{F})$ be a measurable space, let E denote the integers, let I denote the nonnegative integers, and let $\{X_n; n \in I\}$ be a sequence of measurable functions from Ω into E. For each $n \in I$, let $\mathscr{F}_n$ denote the Borel σ-field of subsets of Ω induced by the collection $\{X_m; m \leqq n\}$, and let $\mathscr{B}_n$ be the σ-field induced by the collection $\{X_m; m \geqq n\}$. We assume that $\mathscr{F}$ is the smallest σ-field containing $\mathscr{F}_n$ for each $n \in I$, and we set $\mathscr{F}_\infty = \bigcap_{n \in I} \mathscr{B}_n$, which we call the *tail field*. Let π be some collection of parameters, and let $\{P^\theta; \theta \in \pi\}$ be a collection of probability measures on $\mathscr{F}$, with respect to each of which, $\{X_n; n \in I\}$ is a temporally homogeneous Markov chain.

Let θ_0 be fixed in π, and let P^{θ_0} be denoted by P. We make the following assumptions concerning the parameterized class of Markov chains just introduced:

ASSUMPTION 1. For each $\theta \in \pi$, for each $n \in I$, P^θ is absolutely continuous with respect to P over $\mathscr{F}_n$.

ASSUMPTION 2. For each $n \in I$, X_n is sufficient on $\mathscr{F}_n$ for the parameter θ. (The meaning of sufficiency is given below.)

ASSUMPTION 3. For each $\theta \in \pi$, the tail field $\mathscr{F}_\infty$ is P^θ-trivial, i.e., for each $\Lambda \in \mathscr{F}_\infty$, $P^\theta(\Lambda)$ is either zero or one.

Assumption 1 allows us to define the likelihood ratio $L_n(\theta)$, the Radon–Nikodym derivative of the measure P^θ restricted to $\mathscr{F}_n$, taken with respect to the measure P restricted to $\mathscr{F}_n$. That is, if $\Lambda \in \mathscr{F}_n$, then

$$P^\theta(\Lambda) = \int_\Lambda L_n(\theta)\, dP. \tag{1}$$

Received June 20, 1969.

[1] Prepared with the partial support of the National Science Foundation GP–14798.

Assumption 2 means that for each $\theta \in \pi$, there is a function $k[\cdot, \theta]$ defined on $I \times E$ such that $k[(n, X_n), \theta] = L_n(\theta)$ almost surely with respect to the measure P. Finally, Assumption 3 implies in effect that the "true" value of the parameter uniquely determines, and is uniquely determined by, the asymptotic behavior of the chain.

We now recall some definitions from the potential theory for Markov chains. The reader is referred to [1] for a more complete exposition. The definitions given here are specialized for the space-time chain $\{(n, X_n); n \in I\}$ over the probability space $(\Omega, \mathscr{F}, P)$. The state-space for this chain is the subset E' of $I \times E$ defined by

$$E' = \{(n, x) \in I \times E \mid P(X_n = x) > 0\}.$$

A nonnegative function f on E' is said to be *regular* (for the space-time chain) if for each $(n, x) \in E'$,

$$f(n, x) = E(f(n+1, X_{n+1}) \mid X_n = x).$$

We wish to consider only those regular functions f satisfying the additional condition that $E(f(0, X_0)) = 1$. The word "regular" will henceforth imply that this condition holds.

If f is a regular function, we can define a probability measure P^f over $\mathscr{F}$ called a *relativised* measure. It is sufficient to define P^f over $\mathscr{F}_n$ for each $n \in I$. When $\Lambda \in \mathscr{F}_n$, define $P^f(\Lambda)$ by

$$P^f(\Lambda) = \textstyle\int_\Lambda f(n, X_n)\, dP. \tag{2}$$

The regularity of f guarantees that this is a consistent definition.

A regular function f is *minimal* if the relation $f = \frac{1}{2}(g+h)$, for g, h regular, implies that $f = g = h$. It is not difficult to show that the regular function f is minimal if and only if the tail field $\mathscr{F}_\infty$ is trivial with respect to the relativised measure P^f.

The Martin potential kernel for the space-time chain is the function K over $E' \times E'$, defined by

$$\begin{aligned} K[(m, x), (n, y)] &= P(X_m = x \mid X_n = y)/P(X_m = x) && m \leqq n \\ &= 0 && m < n. \end{aligned}$$

We now state one of the central results of the Martin boundary theory for Markov chains. There is a set S, called the *minimal boundary*. For each $s \in S$, there is a uniquely determined minimal function $K[\cdot, s]$ on E'. For each $s \in S$, there is a set $\Omega_s \in \mathscr{F}$, such that $P^{K[\cdot, s]}(\Omega_s) = 1$. We call the elements of Ω_s *s-paths*, and for each s-path $\omega_0 \in \Omega_s$, we have

$$\lim_{n \to \infty} K[(m, x), (n, X_n(\omega_0))] = K[(m, x), s]$$

for each $(m, x) \in E'$. This amounts to saying that the sequence $\{(n, X_n(\omega_0)); n \in I\}$ converges to s in a special topology placed on $E' \cup S$ called the *Martin boundary topology*.

We can now state the main point of this paper: The minimal boundary S and the parameter set π have essentially the same properties. Specifically:

(i) For each $\theta \in \pi$, the likelihood function $k[\cdot, \theta]$ is a minimal regular function and furthermore, the measure P^{θ} is just the relativised measure $P^{k[\cdot,\theta]}$.

(ii) For each $\theta \in \pi$, there is a subset $\Omega_{\theta} \in \mathscr{F}$, such that $P^{\theta}(\Omega_{\theta}) = 1$, and for $\omega_0 \in \Omega_{\theta}$,

$$\lim_{n \to \infty} K[(m, x), (n, X_n(\omega_0)] = k[(m, x), \theta]$$

for each $(m, x) \in E'$.

These assertions imply that Martin boundary for the chain $\{(n, X_n); n \in I\}$ can be constructed in such a way that the minimal boundary contains the parameter set π. Then for each $\theta \in \pi$, the likelihood functions $k[\cdot, \theta]$ and the Martin potential kernel $K[\cdot, \theta]$ are the same function. We note that the minimal boundary may be strictly larger than π.

Brief proofs of these assertions now follow. It is well known that the system $\{(L_n(\theta), \mathscr{F}_n); n \in I\}$ is a martingale with respect to the measure P. From the definition of the functions $k[\cdot, \theta]$, it follows that $k[\cdot, \theta]$ is regular. That $P^{K[\cdot,\theta]} = P^{\theta}$ then follows from equations (i) and (ii). Finally, Assumption 3 implies $k[\cdot, \theta]$ is minimal. This proves Assertion (i).

From the definition of the Martin potential kernel, for $m <$ and$n$ $(m, x) \in E'$,

$$K[(m, x), (n, X_n)] = P(X_m = x \mid X_n)/P(X_m = x).$$

Hence, the system $\{(K[(m, x), (n, X_n)], \mathscr{B}_n); n \geqq m\}$ is a martingale with respect to P, hence it converges a.s., (with respect to P). The limit is a.s. given by

$$\lim_{n \to \infty} K[(m, x), (n, X_n)] = P(X_m = x \mid \mathscr{F}_{\infty})/P(X_m = x).$$

Since $\mathscr{F}_{\infty}$ is trivial with respect to $P = P^{\theta_0}$, this limit is a.s. equal to 1, hence

$$\begin{aligned} \lim_{n \to \infty} K[(m, x), (n, X_n)] &= 1 \\ &= k[(m, x), \theta_0]. \end{aligned}$$

P^{θ_0}—a.s., for all $(m, x) \in E'$. Hence, a set $\Omega_{\theta_0} \in \mathscr{F}$ with the required properties can be obtained.

For an arbitrary $\theta \in \pi$, define the Martin potential kernel K^{θ} for the space-time chain over $(\Omega, \mathscr{F}, P^{\theta})$. It is easy to see that

$$K^{\theta}[(m, x), (n, y)]k[(m, x), \theta] = K[(m, x), (n, y)].$$

Following the reasoning of the previous paragraph, $\lim K^{\theta}[(m, x), (n, X_n)] = 1$, P^{θ}—a.s., hence, $\lim_{n \to \infty} K[(m, x), (n, X_n)] = k[(m, x), \theta]$, P^{θ}—a.s., and so the set Ω_{θ} can be constructed, thus proving assertion (ii).

3. Examples. We exhibit a large class of familiar statistical models in which Assumptions 1, 2 and 3 hold. Suppose for each $\theta \in \pi$, there is a probability distribution function $f(x, \theta)$ on E, such that

$$P^{\theta}(X_0 = 0) = 1$$

$$P^{\theta}(X_{n+1} - X_n = x) = f(x, \theta) n \in I.$$

Then for each $\theta \in \pi$, $\{X_n;\, n \in I\}$ is a random walk with a fixed starting point over $(\Omega, \mathscr{F}, P^\theta)$, hence Assumption 3 holds.

LEMMA. *A necessary and sufficient condition for Assumptions* 1 *and* 2 *is that the families of probability distribution functions* $\{f(x, \theta);\, \theta \in \pi\}$ *be of exponential form, i.e., there exist a function* f *on* E, *and functions* g *and* h *on* π, *such that*

$$f(x,\theta) = f(x)g(\theta)^x h(\theta).$$

PROOF. Proving sufficiency is a simple matter, and is left to the reader.

When Assumptions 1 and 2 hold, $k[\cdot\,, \theta]$ is a minimal regular function for the space-time chain. It is shown in [2] that such a function is of the form $k[(n, x), \theta] = g(\theta)^x h(\theta)^n$. Hence,

$$\begin{aligned} f(x,\theta) &= P^\theta(X_1 = x) \\ &= P^{\theta_0}(X_1 = x)k[(1,x),\theta] \\ &= f(x,\theta_0)g(\theta)^x h(\theta) \end{aligned}$$

and this proves the lemma.

The binomial, exponential and negative binomial distributions are of exponential form.

REFERENCES

[1] DOOB, J. L. (1959). Discrete potential theory and boundaries. *J. Math. Mech.* **8** 433–458.

[2] DOOB, J. L., SNELL, J. L. and WILLIAMSON, R. E. (1960). Application of boundary theory to sums of independent random variables in *Contributions to Probability and Statistics*. Stanford Univ. Press, 182–197.