*The Annals of Probability*
1974, Vol. 2, No. 2, 183-201

# TAILFREE AND NEUTRAL RANDOM PROBABILITIES AND THEIR POSTERIOR DISTRIBUTIONS[1]

BY KJELL DOKSUM

*University of California, Berkeley*

The random distribution function $F$ and its law is said to be neutral to the right if $F(t_1), [F(t_2) - F(t_1)]/[1 - F(t_1)], \cdots, [F(t_k) - F(t_{k-1})]/[1 - F(t_{k-1})]$ are independent whenever $t_1 < \cdots < t_k$. The posterior distribution of a random distribution function neutral to the right is shown to be neutral to the right. Characterizations of these random distribution functions and connections between neutrality to the right and general concepts of neutrality and tailfreeness (tailfreedom) are given.

**1. Introduction and summary.** Recently Ferguson (1973) introduced a class of random probabilities called Dirichlet processes and showed that they can be used to solve nonparametric decision theoretic problems. Let the probability distribution $\mathscr{P}$ of a Dirichlet process $P$ be called a Dirichlet distribution. The main properties of $P$ and $\mathscr{P}$ obtained by Ferguson are:

(a) $\mathscr{P}$ is nonparametric in the sense that it has a "large" or "nonparametric" class of probabilities as its suppport in the topology of weak convergence,

(b) if $P$ is viewed as a parameter with prior distribution $\mathscr{P}$, then the posterior distribution of $P$ given a sample also has a Dirichlet distribution, and

(c) $P$ is with probability one a discrete probability.

Ferguson's results apply to random probabilities on abstract measurable spaces. In Section 3, it is shown that similar results can be obtained for more general processes in the real case. Let $F(t) = P((-\infty, t])$ denote the random distribution function corresponding to $P$, then $F$, $P$ and $\mathscr{P}$ are said to be *neutral to the right* if the normalized increments

$$F(t_1), [F(t_2) - F(t_1)]/[1 - F(t_1)], \cdots, [F(t_k) - F(t_{k-1})]/[1 - F(t_{k-1})] \tag{1.1}$$

are independent for all $t_1 < \cdots < t_k$. Note that $F(t_1)$ is the normalized increment from $-\infty$ to $t$, and that (0/0) is defined here and throughout to be one. It is shown that:

(a′) random probabilities neutral to the right can be chosen to be "nonparametric" in the sense of Ferguson (Remark 3.3),

(b′) the posterior distribution of a random probability neutral to the right is neutral to the right (Theorem 4.2), and

Received May 11, 1972; revised July 26, 1973.

[1] Research supported in part by the National Science Foundation under Grants GP-29123 and GP-33697X.

*AMS* 1970 *subject classifications.* Primary 60K99; Secondary 62C10, 62G99.

*Key words and phrases.* Random probabilities, posterior distributions, processes, Dirichlet process, posterior mean of a process, Bayes estimates, tailfree, neutral.

(c′) if the "non-random part" of $F$ is removed, it is with probability one a discrete distribution function (Corollary 3.2). It is further shown (Theorem 3.1) that processes neutral to the right can be defined in terms of independent increment processes as follows: $F$ is neutral to the right if and only if it can be written in the form $F(t) = 1 - \exp[-Y(t)]$, where $Y(t)$ is an a.s. non-decreasing, independent increment process with $\lim_{t\to-\infty} Y(t) = 0$ a.s., $\lim_{t\to\infty} Y(t) = \infty$ a.s. and $Y(t^+) = Y(t)$ a.s. The Dirichlet process corresponds to one of these independent increment processes $Y(t)$. Section 4 contains formulas and examples which demonstrate how the posterior distribution of a process neutral to the right can be obtained.

The concept of neutrality can also be defined for random probabilities on abstract measurable spaces, which is done in Section 2. It turns out that the concept of a neutral random probability is an extension to a process of Connor and Mosimann's (1969) concept of neutrality for $k$-dimensional random vectors as Ferguson's Dirichlet process is an extension of the $k$-dimensional Dirichlet distribution to a process. Connor and Mosimann's concept of neutrality can be viewed as a special case of Freedman's (1963) concept of "tailfree." Fabius' (1964) tailfree or $F$-neutral processes are also defined in Section 2 and some recent ([7], [11]) results concerning Dirichlet, neutral and tailfree processes are stated.

Another class of random probabilities that contain the Dirichlet process has been defined by Antoniak (1969). Generalizations and characterizations of the Dirichlet process have been obtained by Blackwell (1973) and Blackwell and MacQueen (1973).

**2. Tailfree and neutral processes.** Let $\mathscr{X}$ be a set with a $\sigma$-field $\mathscr{A}$ of subsets. A *finitely additive random probability* on the measurable space $(\mathscr{X}, \mathscr{A})$ is a stochastic process $\{P(A) : A \in \mathscr{A}\}$ on some probability space $(\Gamma, \mathscr{S}, \lambda)$ such that

(i) $P(A)$ is a random variable with values in $[0, 1]$ for each $A \in \mathscr{A}$
(ii) $P(\mathscr{X}) = 1$ a.s.
(iii) $P$ is finitely additive in distribution, i.e.,

$$(P(\textstyle\bigcup_i A_{1,i}), \cdots, P(\bigcup_i A_{m,i})) =_{\mathscr{L}} (\sum_i P(A_{1,i}), \cdots, \sum_i P(A_{m,i}))$$

for every finite class $\{A_{j,i}\}$ of pairwise disjoint sets from $\mathscr{A}$, where $=_{\mathscr{L}}$ denotes equality in law.

As shown by Ferguson (1973), it follows from Kolmogorov's Extension Theorem that if a system of finite dimensional distributions of $P(A_{1,1}), \cdots, P(A_{m,k})$ are given for each finite class $A_{1,1}, \cdots, A_{m,k}$ of pairwise disjoint sets from $\mathscr{A}$ such that (i), (ii) and (iii) holds, then there exists a process $P$ with these finite dimensional distributions. The distribution of $P$ will be denoted by $\mathscr{P}$ and is a probability on $([0, 1]^{\mathscr{A}}, \sigma(\mathscr{B}^{\mathscr{A}}))$, where $\sigma(\mathscr{B}^{\mathscr{A}})$ denotes the $\sigma$-field generated by the field $\mathscr{B}^{\mathscr{A}}$ of Borel cylinder sets in $[0, 1]^{\mathscr{A}}$.

$P$ is said to be a *random probability* if it is also $\sigma$-additive in distribution, i.e., if $\mathscr{L}(P(A_k))$ tends to the law degenerate at zero as $k \to \infty$ for each decreasing sequence $\{A_k\}$ of sets from $\mathscr{A}$ with $\lim_{k\to\infty} A_k = \emptyset$. Since $P(A_1) \geqq P(A_2) \geqq \cdots$ a.s., it follows that this is equivalent to having $\lim_{k\to\infty} P(A_k) = 0$ a.s. Note that $\mathscr{P}$ is $\sigma$-additive whether or not $P$ is $\sigma$-additive in distribution. The above definitions can easily be shown to be equivalent to those of Ferguson (1973), however, the terminology is different in that here random probabilities are necessarily $\sigma$-additive in distribution unless they are called finitely additive.

Some connections between finitely additive random probabilities and their expectations are given in

LEMMA 2.1. *Let $\beta$ be the set function on $\mathscr{A}$ defined by $\beta(A) = E(P(A))$, then*

(a) *$\beta$ a finitely additive probability*

(b) *for $A \in \mathscr{A}$, $P(A) = 0$ a.s. if and only if $\beta(A) = 0$*

(c) *for $A_k \in \mathscr{A}$, $\mathscr{L}(P(A_k))$ converges to the law degenerate at $0$ as $k \to \infty$ if and only if $\beta(A_k) \to 0$ as $k \to \infty$, in particular, $P$ is $\sigma$-additive in distribution if and only if $\beta$ is $\sigma$-additive.*

PROOF. (a) and (b) are clear. (c) follows if one recalls that a sequence of uniformly integrable random variables converges to an integrable random variable in probability if and only if it converges in the mean to this variable.

We now turn to concepts and terminology needed in the definitions of $F$-neutral (tailfree) and neutral random probabilities or processes. The definition of $F$-neutral is essentially a definition of Fabius' term tailfree. Following a suggestion of Fabius (personal communication), the term tailfree is not used here since the definition does not in general depend on the tails of the real line. However, it does include Freedman's (1963) tailfree random probabilities in the discrete case (see Remark 2.2).

The processes or random probabilities are defined in terms of independence properties of $P(A_1), \cdots, P(A_m)$ for measurable partitions $A_1, \cdots, A_m$. These independence properties are desirable because they lead to processes whose probability distributions are nonparametric in the sense of Ferguson (1973) and one would expect them to have more or less tractable posterior distributions (Fabius (1964)). It should be noted that the type of independence properties considered are not necessarily the only ones that will lead to these desirable properties.

Let $\{\prod_m : m = 0, 1, \cdots\}$ denote a sequence of nested, measurable partitions with $\prod_0 = \{\mathscr{X}\}$. Let $\{A_{m,1}, \cdots, A_{m,k_m}\}$ denote $\prod_m$. Since the partitions are nested, then for $s < m$, there is one set in $\prod_s$ that contains $A_{m,i}$; this set will be denoted by $A_{s,l(i)}$. Note that $A_{0,l(i)} = \mathscr{X}$. For two measurable sets $A$ and $B$, $P(B \mid A)$ equals $P(A \cap B)/P(A)$ on the subset of $\Gamma$ where $P(A) > 0$ and, by convention, equals one elsewhere. Thus $P(B \mid A)$ is the conditional random probability of $B$ given $A$. It will sometimes be useful to regard the partition $\{A_{m,1}, \cdots, A_{m,k_m}\}$ not just as a collection of sets, but as a collection of sets with the order indicated, that is, $A_{m,1}$ precedes $A_{m,2}$, which precedes $A_{m,3}$ etc. We use the term *ordered partition* in this case. Finally, $A^c$ denotes the complement of the set $A$.

DEFINITION 2.1. $P$ (and $\mathscr{P}$) is *F-neutral* with respect to the sequence $\{\prod_m\}$ of nested, measurable partitions if there exist nonnegative random variables $Z_{m,i}$, $i = 1, \cdots, k_m$, $m = 1, 2, \cdots$ such that for each $m \geqq 1$, the families of random variables $\{Z_{1,i}: i = 1, \cdots, k_1\}, \cdots, \{Z_{m+1,i}: i = 1, \cdots, k_{m+1}\}$ are independent and

$$(2.1)\qquad (P(A_{m,1}), \cdots, P(A_{m,k_m})) =_{\mathscr{L}} (\prod_{s=1}^m Z_{s,l(1)}, \cdots, \prod_{s=1}^m Z_{s,l(k_m)}) .$$

DEFINITION 2.2. $P$ (and $\mathscr{P}$) is *neutral* with respect to the sequence $\{\prod_m\}$ of nested, measurable, ordered partitions if for each $m \geqq 1$ there exist nonnegative independent random variables $V_{m,1}, \cdots, V_{m,k_m}$ with $V_{m,k_m} = 1$ and

$$(2.2)\qquad (P(A_{m,1}), P(A_{m,2}), \cdots, P(A_{m,k_m})) \\ =_{\mathscr{L}} (V_{m,1}, V_{m,2}(1 - V_{m,1}), \cdots, V_{m,k_m} \prod_{j=1}^{k_m-1} (1 - V_{m,j})) .$$

Note that $F$-neutrality refers to independence properties between partitions and essentially means that the families

$$(2.3)\qquad \{P(A_{1,i} \mid A_{0,l(i)}): i = 1, \cdots, k_1\}, \cdots, \{P(A_{m+1,i} \mid A_{m,l(i)}): i = 1, \cdots, k_{m+1}\}$$

are independent; that is, when the sets in one partition, say $\prod_s$, are divided into new sets for the next partition $\prod_{s+1}$, the relative random probabilities assigned to these new sets are independent of the corresponding relative random probabilities assigned to the sets in other partitions. If $P$ is such that the families (2.3) are independent, then $P$ is $F$-neutral with respect to $\{\prod_m\}$ since we can define

$$(2.4)\qquad Z_{m+1,i} = P(A_{m+1,i} \mid A_{m,l(i)}) .$$

The term neutral refers to independence properties within partitions and essentially means that for each $m \geqq 1$, the random variables

$$(2.5)\qquad P(A_{m,1}), P(A_{m,2} \mid A_{m,1}^c), P(A_{m,3} \mid (A_{m,1} \cup A_{m,2})^c), \cdots, P(A_{m,k_m} \mid A_{m,k_m}) = 1$$

are independent. Thus as one moves to the right in an ordered partition, the relative random probability assigned to the next set is independent of the corresponding relative random probabilities assigned to the other sets in the partition. If $P$ is such that the random variables (2.5) are independent then $P$ is neutral since we can define

$$(2.6)\qquad V_{m,i} = P(A_{m,i} \mid A_{m,i}' \cup \cdots \cup A_{m,k_m}) .$$

The random probabilities that Kraft and van Eeden (1964) and Kraft (1964) call processes obtained by "$Z$-interpolation" or "independent interpolation" are $F$-neutral with respect to a tree of partitions made up of intervals. The processes considered by Dubins and Freedman (1966) are $F$-neutral with respect to a tree of partitions when their "base probability" assigns probability one to a vertical line in the unit square. Furthermore, the processes considered by Metivier (1971), which generalize the Dubins-Freedman processes, are $F$-neutral when the "$X_L$" of Metivier's paper are degenerate.

Freedman's (1963) "discrete tailfree" random probabilities are neutral when "$N$" in that paper is zero (see Remark 2.2). The word neutral was introduced by Connor and Mosimann (1969) who defined the concept for random vectors. It is possible to construct many interesting classes of neutral processes. One of these classes is the class of processes neutral to the right considered in Sections 3 and 4. Another class is given in Example 3.2.

Let $\alpha$ be a finite, nonnull, finitely additive measure on $(\mathscr{X}, \mathscr{A})$. The *Dirichlet process* with parameter $\alpha$ is a finitely additive random probability $P$ with the property that for every measurable partition $\{B_1, \cdots, B_k\}$ of $\mathscr{X}$, $(P(B_1), \cdots, P(B_k))$ has a Dirichlet distribution with parameter $(\alpha(B_1), \cdots, \alpha(B_k))$. When $\alpha(B_i) = 0$, the interpretation is that $P(B_i) = 0$ a.s. The Dirichlet process is $F$-neutral with respect to every sequence of nested, measurable partitions. To see this, note that for this process, (2.1) holds when, for each $r$, $Z_{1,l(r)}, \cdots, Z_{m,l(r)}$ are independent beta random variables with parameters $(\alpha(A_{1,l(r)}), \alpha(\mathscr{X})-\alpha(A_{1,l(r)})), \cdots, (\alpha(A_{m,l(r)}), \alpha(A_{m-1,l(r)}) - \alpha(A_{m,l(r)}))$. Similarly, the Dirichlet process is neutral with respect to every sequence of nested, measurable, ordered partitions, since if $P$ is a Dirichlet process, then (2.2) holds when, for each $m$, the $V_{m,i}$ are independent beta random variables with parameters $(\alpha(A_{m,i}), \sum_{j\geqq i+1} \alpha(A_{m,j}))$, $i = 1, \cdots, k_m$.

It turns out that the Dirichlet process is essentially the only random probability that is independent of the defining partitions in the sense of having the desired independence properties for all sequences of partitions. Let $C_1$ denote the class of all random probabilities $P$ such that either

(i) $P$ is degenerate at a given probability distribution $P_0$,
(ii) $P$ concentrates on a random point, or
(iii) $P$ concentrates on two nonrandom points.

THEOREM 2.1. *Suppose $P$ is not one of the random probabilities in $C_1$.*

(a) *If $P$ is $F$-neutral with respect to all sequences of nested, measurable partitions, then $P$ is a Dirichlet process.*

(b) *If $P$ is neutral with respect to all sequences of nested, measurable, ordered partitions, then $P$ is a Dirichlet process.*

This result follows immediately from Theorem 2 of Fabius (1973). It was obtained simultaneously and independently in [7] using the results and conditions of Darroch and Ratchiff (1971). These conditions are stronger than the ones above. If we combine the (a)-part of the result with Theorem 2.1 of Fabius (1964), we get an interesting corollary:

COROLLARY 2.1. *The Dirichlet process is the only process not in $C_1$ such that for each $A \in \mathscr{A}$, the posterior distribution of $P(A)$ given a sample $X_1, \cdots, X_n$ from $P$, depends only on the number $N_A$ of $X$'s that fall in $A$ (and not on where they fall within or outside of $A$).*

The property that the posterior distribution of $P(A)$ depends only on $N_A$ for the Dirichlet process is what makes the posterior distribution of this process so easy to handle. Otherwise, it is not necessarily a desirable property since the posterior distribution is rather insensitive to the values of the sample. For instance, if $\mathscr{X}$ is the real line and $A$ is an interval $[a, b]$, one may want the posterior distribution to give more weight to $[a, b]$ if $x$ falls in the middle of the set than when it falls on the boundary. Thus, in addition to the Dirichlet process, it is useful to consider processes that do not have the property of Corollary 2.1.

REMARK 2.1. A sequence $\{\prod_m\}$ of nested, measurable (ordered) partitions is said to be a *tree* of (ordered) partitions if the collection of all sets in the partitions generate $\mathscr{A}$. Trees of partitions exist if and only if $\mathscr{A}$ is separable (countably generated). Fabius' (1964) definition was in terms of trees of partitions. The definitions and results of this section may be restricted to such sequences of (ordered) partitions when $\mathscr{A}$ is separable. In this case, $F$-neutral processes can be constructed from $Z$'s satisfying the independence properties of Definition 2.1 by using (2.1) to define $P(B)$ for $B$ in $\bigcup_{m=1}^{\infty} \prod_m$ and setting $P(B) = \inf\{\sum_1^k P(A_i)\}$ for other $B \in \mathscr{A}$, where the infimum is over $A_1, \cdots, A_k$ in $\bigcup_{m=1}^{\infty} \prod_m$ satisfying $\bigcup_1^k A_i \supset B$.

REMARK 2.2. There are sequences of partitions for which the concept of $F$-neutrality coincides with neutrality. Let $B_1, B_2, \cdots$ be a measurable partition. Define $\{\prod_m : A_{m,1}, \cdots, A_{m,m+1}\}$ by $A_{m,i} = B_i$, $i = 1, \cdots, m$; $A_{m,m+1} = (\bigcup_{i=1}^m B_i)^c$. Then it is easy to check that $P$ is $F$-neutral with respect to the sequence of partitions $\{\prod_m\}$ if and only if $P$ is neutral with respect to the sequence of ordered partitions $\{\prod_m\}$. If $\mathscr{X}$ is the reals, or an interval, with the Borel $\sigma$-field, then the $\{\prod_m\}$ above can not be chosen to be a tree of partitions. However, suppose $\mathscr{X}$ is the positive integers and $\mathscr{A}$ the class of all subsets. Then if $B_i = \{i\}$, $\{\prod_m\}$ is a tree of partitions, moreover, if $P$ is $F$-neutral (neutral) with respect to $\{\prod_m\}$, then it is tailfree in the sense of Freedman (1963).

REMARK 2.3. It seems that the concept of neutrality is more useful than the concept of $F$-neutrality in the sense that even though it is not possible to find processes (other than the Dirichlet process or processes in $C_1$) neutral with respect to all sequences of nested, measurable, ordered partitions, it is possible to find processes neutral with respect to every member of large classes of sequences of nested, ordered partitions. See for instance Remark 3.1 (a) and Example 3.2. It is apparently not possible to do the same with the concept of $F$-neutrality for interesting classes of $\{\prod_m\}$ in the real case.

**3. Tailfree and neutral processes in the real case.** If one views the definitions of $F$-neutral and neutral as attempts to define general classes of finitely additive random probabilities that have certain independence properties with respect to all trees of partitions, then these attempts fail since in both cases, only the Dirichlet process have these properties. However, it turns out that in the real case, where $(\mathscr{X}, \mathscr{A})$ is the real line $R$ with the Borel sets $\mathscr{B}$, it is possible to define a wide class of "neutral" random probabilities which contains the Dirichlet process. Moreover, for the processes in the class, it is possible to compute the posterior distribution given a sample.

We will make use of the order properties of the real line, thus it is natural to introduce the random distribution function $F$ corresponding to the random probability $P$. It is defined by $F(t) = P((-\infty, t])$, and it follows that $F$ is a stochastic process with a separable version (a process with the same probability distribution as $F$) which satisfies

(I) $F$ is a.s. non-decreasing, i.e., $\mathscr{P}(F\colon F$ is non-decreasing$) = 1$,

(II) $\lim_{t\to-\infty} F(t) = 0$ a.s., $\lim_{t\to\infty} F(t) = 1$ a.s., and

(III) $\lim_{s\to t^+} F(s) = F(t)$ a.s. for each $t \in R$.

*F is continuous in probability (from the right)* if, for all $t$, whenever $s \to t$ ($s \to t^+$), $F(s)$ converges in probability to $F(t)$. If $F$ is separable, then it is said to be *continuous* a.s. (*from the right*) if

$$\{F\colon F(t) \neq \lim_{s\to t} F(s)\}(\{F\colon F(t) \neq \lim_{s\to t^+} F(s)\})$$

is a $\mathscr{P}$ null set for each $t$, and it is said to have a.s. *continuous sample paths* if $\mathscr{P}(F\colon F$ is continuous$) = 1$. Let $\beta(t) = \beta((-\infty, t]) = E(F(t))$, where $\beta(A)$, $A \in \mathscr{B}$, is as defined in Section 2, then $\beta(t)$ is a distribution function. From Lemma 2.1 and some standard arguments, we have:

PROPOSITION 3.1. *F is continuous in probability from the right. It is continuons in probability if and only if $\beta$ is continuous. There exists a separable version of F such that $\mathscr{P}(F\colon F$ is a distribution function$) = 1$; moreover, F is continuous* a.s. *if and only if $\beta$ is continuous.*

PROOF. In view of Lemma 2.1, the only part that requires some arguments is $\mathscr{P}(F\colon F$ is a distribution function$) = 1$. Since $F$ is a.s. non-decreasing, almost every sample path of $F$ has only jump discontinuities. Since $F$ is continuous in probability from the right, the result follows (e.g. Breiman (1968 page 299).

If $P$ is the Dirichlet process with parameter $\alpha$, then $\beta(t) = \alpha(t)/\alpha(R)$, and Ferguson (1973) has shown that there is a version of $F$ such that $\mathscr{P}(F\colon F$ is a discrete distribution function$) = 1$ whether or not $\beta(t)$ is continuous.

We have shown that for each random probability $P$, there is a separable version of the random distribution function $F$ corresponding to $P$ such that $\mathscr{P}(F\colon F$ is a distribution function$) = 1$. We now show that we can also go the other way.

PROPOSITION 3.2. *Suppose that F is a separable stochastic process satisfying* (I), (II) *and* (III); *then there is a version of F such that*:

(a) *The random set function $P_0$ defined on intervals of the form $(a, b]$, $a < b$, by $P_0((a, b]) = F(b) - F(a)$ can be extended to a random probability $P_0$ on $(R, \mathscr{B})$.*

(b) *$P_0$ can be chosen so that if $\mathscr{P}_0$ denotes the probability distribution of $P_0$, $\mathscr{P}_0\{P\colon P$ is a probability on $(R, \mathscr{B})\} = 1$.*

(c) *If $F$ is a version of the random distribution function corresponding to $P$, then $\mathscr{P}_0 = \mathscr{P}$.*

PROOF. Using Proposition 3.1, we choose a separable version of $F$ such that almost all the sample paths are distribution functions. For $A \in \mathscr{B}$, define $P_0(A) = \inf\{\sum_{i=1}^{k} [F(b_i) - F(a_i)]\}$, where the infimum is over all classes $\{(a_i, b_i], a_i < b_i, i = 1, \cdots, k\}$ of intervals with rational endpoints satisfying $\bigcup_{i=1}^{k} (a_i, b_i] \supset A$. $P_0(A)$ depends only on the values of $F$ at the rationals, thus $P_0(A)$ is measurable. For almost every $F$, $P_0$ is a probability, thus (a) and (b) follow. (c) is a consequence of $P((a, b]) = F(b) - F(a)$ a.s. for all $a \leqq b$.

Because of this result, we can call every separable stochastic process $F$ satisfying (I), (II) and (III) a random distribution function since it is the random distribution function corresponding to some random probability. Combining Theorem 2.1 (a) and the proof of Proposition 3.2, we have

COROLLARY 3.1. *If $P$ is $F$-neutral with respect to every tree of partitions made up of intervals of the form $(a, b]$, $a < b$, with endpoints in a dense subset of $R$, and if $P$ is not in $C_1$, then $P$ is a Dirichlet process.*

We now introduce the new processes.

DEFINITION 3.1. The random distribution function $F$ (and the corresponding $P$ and $\mathscr{P}$) is said to be *neutral to the right* if for each $k > 1$ and $t_1 < \cdots < t_k$, there exists nonnegative independent random variables $V_1, \cdots, V_k$ such that

$$(3.1) \qquad (F(t_1), F(t_2), \cdots, F(t_k)) =_{\mathscr{L}} (V_1, 1 - (1 - V_1)(1 - V_2), \cdots, 1 - \prod_{i=1}^{k} (1 - V_i)) .$$

The equations

$$(3.2) \qquad F(t_i) = 1 - \prod_{j=1}^{i} (1 - V_j) , \qquad i = 1, \cdots, k$$

yield

$$(3.3) \qquad F(t_i) - F(t_{i-1}) = V_i \prod_{j=1}^{i-1} (1 - V_j) , \qquad \text{and}$$

$$[F(t_i) - F(t_{i-1})]/[1 - F(t_{i-1})] = V_i , \qquad i = 1, \cdots, k;\ t_0 = -\infty .$$

Thus "$F$ is neutral to the right" essentially means that the normalized increments

$$(3.4) \qquad F(t_1), [F(t_2) - F(t_1)]/[1 - F(t_1)], \cdots, [F(t_k) - F(t_{k-1})]/[1 - F(t_{k-1})]$$

are independent for all $t_1 < \cdots < t_k$. Definition 3.1 is preferred because of the possibility of dividing by zeros in (3.4).

Remark 2.2 and (3.3) result in

REMARK 3.1. The following three conditions are equivalent to $F$ being neutral to the right.

(a) $P$ is neutral with respect to every sequence of nested ordered partitions of the form $\{\prod_m : A_{m,1}, \cdots, A_{m,k_m}\}$ with $A_{m,i} = (t_{m,i}, t_{m,i+1}]$, $i = 1, \cdots, k_m$; $-\infty = t_{m,1} < \cdots < t_{m,k_m+1} = \infty$; $m = 1, 2, \cdots$.

(b) For all $t_1 < \cdots < t_k$ there exists independent random variables $V_1, \cdots, V_k$ such that

$$(F(t_1), F(t_2) - F(t_1), \cdots, F(t_k) - F(_{k-1})) \\ =_{\mathscr{L}} (V_1, V_2(1 - V_1), \cdots, V_k \prod_{j=1}^{k-1} (1 - V_j))$$

(c) $P$ is $F$-neutral with respect to all nested sequences of partitions $\{\prod_m : A_{m,1}, \cdots, A_{m,m+1}\}$ with $A_{m,i} = (t_i, t_{i+1}]$, $i = 1, \cdots, m$; $A_{m,m+1} = (\bigcup_{i=1}^m A_{m,i})^c$, for some $-\infty = t_1 < t_2 < \cdots$.

The class of processes defined below contains those defined by Freedman [(1964), Definition 2] in the discrete case.

Definition 3.2. A random distribution function $F$ is *tailfree* with respect to the tail $(s, \infty)$ if for all $s = t_0 < \cdots < t_k$ there exists nonnegative independent random variables $V_1, \cdots, V_k$ independent of $\{F(t) : t \leqq s\}$ such that

$$(3.5) \qquad (F(t_1), \cdots, F(t_k)) \\ =_{\mathscr{L}} (F(s) + [1 - F(s)][1 - \prod_{j \leqq i} (1 - V_j)], i = 1, \cdots, k) .$$

Remark 3.2. It is easy to check that

(a) $F$ is neutral to the right if and only if $F$ is tailfree with respect to $(s, \infty)$ for all $s$ in $R$.

(b) If $F$ is tailfree with respect to $(s, \infty)$, then there is a version of $F$ such that for all $t_1 < t_2 < \cdots$ with $t_r = s$ for some $r$, $(F(t_1), \cdots, F(t_{r+k}))$ and

$$\left(\frac{F(t_{r+k+1}) - F(t_{r+k})}{1 - F(t_{r+k})}, \frac{F(t_{r+k+2}) - F(t_{r+k+1})}{1 - F(t_{r+k})}, \cdots\right)$$

are independent for each integer $k \geqq 0$.

The random distribution function corresponding to the Dirichlet process with parameter $\alpha$ will be called the Dirichlet random distribution function with parameter $\alpha(t) = \alpha((-\infty, t])$. We now show that it is not the only random distribution function that is neutral to the right, in fact, random distribution functions neutral to the right can be constructed from certain independent increment processes $Y(t)$. We allow these processes $Y(t)$ to equal $\infty$ with positive probability. This will not lead to difficulties as long as we use the convention $\infty - \infty = \infty$.

Theorem 3.1. *$F(t)$ is a separable random distribution function neutral to the right if and only if it has the same probability distribution as*

$$(3.6) \qquad 1 - \exp[-Y(t)]$$

*for some separable*, a.s. *non-decreasing, right continuous* a.s., *independent increment process with* $\lim_{t\to\infty} Y(t) = 0$ a.s. *and* $\lim_{t\to\infty} Y(t) = \infty$ a.s.

Proof. Suppose that $Y(t)$ is as described in the result, then $F(t) = 1 - \exp[-Y(t)]$ satisfies Definition 3.1 with $V_i = 1 - \exp\{-[Y(t_i) - Y(t_{i-1})]\}$, $i = 1, \cdots, k$. Conversely, if $F$ is neutral to the right, then $Y(t) = -\log[1 - F(t)]$

has the stated properties since if $V_1, \cdots, V_k$ are as in Definition 3.1, then $(Y(t_1) - Y(t_0), \cdots, Y(t_k) - Y(t_{k-1})) =_{\mathscr{L}} (-\log(1 - V_1), \cdots, -\log(1 - V_k))$.

EXAMPLE 3.1. We show that a process $Y(t)$ satisfying the conditions of Theorem 3.1 can be constructed from each nonnegative, infinitely divisible random variable $Y$ and any distribution function $\beta_0$. Define $\gamma(t) = -\log[1 - \beta_0(t)]/\log 2$ (we allow $\gamma(t) = \infty$ from some point on). Let $\phi_Y(v) = E(\exp(ivY))$ denote the characteristic function of $Y$. Now let $Y(t)$ be the separable stochastic process defined by

(3.7) $Y(t)$ has independent increments and characteristic function $\phi_{Y(t)}(v) = [\phi_Y(v)]^{\gamma(t)}$.

It follows that for $t_1 < t_2$, $\phi_{Y(t_2)-Y(t_1)}(v) = [\phi_Y(v)]^{[\gamma(t_2)-\gamma(t_1)]}$. Note that for $F(t) = 1 - \exp[-Y(t)]$,

$$E(F(t)) = \beta(t) = 1 - [\phi_Y(i)]^{\gamma(t)} = 1 - [M_Y(-1)]^{\gamma(t)}, \tag{3.8}$$

where $M_Y(v)$ is the moment generating function of $Y$ ($M_Y(v)$ exists for all $v \leqq 0$). If we let $Y$ have the gamma distribution with $\phi_Y(v) = (1 - iv)^{-1}$, then $Y(t)$ is the gamma process with independent increments and we call $F(t)$ the *exponential gamma* process with parameter $\gamma$ (or $\beta_0$). For this process

$$E(F(t)) = \beta(t) = 1 - 2^{-\gamma(t)} = \beta_0(t). \tag{3.9}$$

There are processes $Y(t)$ satisfying the conditions of Theorem 3.1 that can not be defined by (3.7). If $F(t)$ is the Dirichlet process, then $Y(t) = -\log[1 - F(t)]$ is not of the form (3.7), but has a version that satisfies the conditions of Theorem 3.1.

It is known (e.g. [13]) that independent increment processes such as $Y(t)$ have the property that if the component in the Lévy representation of $\log E(\exp(-ivY(t))$ that corresponds to the random part of $Y(t)$ is zero, then it increases only in jumps with probability one. Thus from Theorem 3.1, we get

COROLLARY 3.2. *If the random distribution $F(t)$ is neutral to the right and if $-\log[1 - F(t)]$ has no nonrandom part, then*

$$\mathscr{P}(F: F \text{ is a discrete distribution function}) = 1. \tag{3.10}$$

*In particular,* (3.10) *holds for the exponential gamma process.*

REMARK 3.3. We say that a given distribution function $F_0$ is in the support of $\mathscr{P}$ if for each $\varepsilon > 0$ and any $t_1 < \cdots < t_k$, $\mathscr{P}\{|F(t_i) - F_0(t_i)| < \varepsilon, i = 1, \cdots, k\} > 0$. Consider the random distribution functions defined in Example 3.1. If the distribution function of $Y$ is strictly increasing on $(0, \infty)$ and if $\gamma$ is continuous and strictly increasing, then the support of $\mathscr{P}$ contains the class of all continuous distribution functions. Thus $\mathscr{P}$ is "nonparametric." More generally, if the distribution function of $Y$ is strictly increasing, then the support of $\mathscr{P}$ will contain the class of all distribution functions absolutely continuous with respect to $\beta(t) = E(F(t)) = 1 - [M_Y(-1)]^{\gamma(t)}$ (or $\gamma(t)$). Note that when $F_0$ is continuous, then the neighborhoods $\{F\colon |F(t_i) - F_0(t_i)| < \varepsilon, i = 1, \cdots, k\}$ are equivalent to sup norm neighborhoods.

$F$ is said to be *neutral to the left* if $P$ is neutral, in the sense of Section 2, with respect to those sequences of nested, ordered partitions where the ordered partitions are of the form $(t_k, \infty), (t_{k-1}, t_k], \cdots, (t_1, t_2], (t_1, -\infty)$ with $t_1 < \cdots < t_k$. Essentially, $F$ is neutral to the left if

$$F(t_k), [F(t_k) - F(t_{k-1})]/F(t_k), \cdots, [F(t_2) - F(t_1)]/F(t_2)$$

are independent. It is conjectured that the Dirichlet process is the only process which is both neutral to the right and neutral to the left (except for processes in $C_1$). However, it is possible to have processes neutral to the left for $t \leqq c$ and neutral to the right for $t > c$ as in the following example (where we have $c = 0$).

EXAMPLE 3.2. Let $Y_1(t)$ and $Y_2(t)$ be two independent, a.s. non-decreasing, right continuous a.s., independent increment processes defined on $(-\infty, 0]$ and $[0, \infty)$ respectively and satisfying $Y_1(0) = Y_2(0) = 0$ a.s., $\lim_{t\to-\infty} Y_1(t) = -\infty$ a.s., $\lim_{t\to\infty} Y_2(t) = \infty$ a.s. Let $V$ be a random variable with values in $[0, 1]$ independent of $Y_1(t)$ and $Y_2(t)$. Define the random distribution function $F(t)$ by

$$\begin{aligned} F(t) &= (1 - V)\exp[Y_1(t)] \qquad \text{for } t \leqq 0\,, \\ &= 1 - V\exp[-Y_2(t)] \qquad \text{for } t > 0\,. \end{aligned}$$

$F$ is neutral with respect to those sequences of nested, ordered partitions where the ordered partitions are of the form $(0, \infty), (s_k, 0], \cdots, (s_1, s_2], (s_1, -\infty)$ with $s_1 < \cdots < s_k < 0$ and with respect to those sequences where the ordered partitions are of the form $(-\infty, 0], (0, t_1], \cdots, (t_{k-1}, t_k], (t_k, \infty)$ with $0 < t_1 < \cdots < t_k$.

**4. The posterior distribution of tailfree and neutral processes.** $R$ is the real line and $\mathscr{B}$ is the $\sigma$-field of Borel sets. $P$ is a random probability (process) with distribution $\mathscr{P}$ and $F$ is the corresponding separable random distribution function which is a.s. a distribution function. $\mathscr{P}$ will be called the *marginal* or *prior* distribution of $F$. $(X_1, \cdots, X_n)$ is a random vector with values in $R^n$. We would like to say that the distribution of $(X_1, \cdots, X_n)$ given $P$ (or $F$) is that of a random sample from a population with probability distribution $P$. This is achieved, as usual, by defining the joint probability distribution $\mathscr{P}_2$ of $X_1, \cdots, X_n$ and $P$ as follows:

$$\mathscr{P}_2(P \in D, X_1 \in B_1, \cdots, X_n \in B_n) = E_{\mathscr{P}}(I_D(P) \textstyle\prod_{i\leqq n} P(B_i)) \tag{4.1}$$

where $D$ is in $\sigma(\mathscr{B}^{\mathscr{B}})$, $B_1, \cdots, B_n$ are in $\mathscr{B}$, and $I_D(P)$ is 1 if $P$ is in $D$ and is 0 otherwise. (4.1) determines a probability $\mathscr{P}_2$ on $(R^n \times [0, 1]^{\mathscr{B}}, \sigma(\mathscr{B}^n \times \mathscr{B}^{\mathscr{B}}))$ where $\mathscr{B}^n$ denotes the Borel cylinder sets of $R^n$.

LEMMA 4.1. *The marginal distribution of $P$ is $\mathscr{P}$, i.e., $\mathscr{P}_2(P \in D) = \mathscr{P}(P \in D)$. For $B_1, \cdots, B_n$ in $\mathscr{B}$, we have*

(a) *The conditional distribution of* $(X_1, \cdots, X_n)$ *satisfies*

(4.2) $$\mathscr{P}_2(X_1 \in B_1, \cdots, X_n \in B_n \mid P) = \prod_{j\le n} P(B_j) \quad \text{a.s.}$$

(b) *The marginal distribution of* $(X_1, \cdots, X_n)$ *is given by*

(4.3) $$\mathscr{P}_2(X_1 \in B_1, \cdots, X_n \in B_n) = E_{\mathscr{P}}(\prod_{j\le n} P(B_j)),$$
$$\textit{in particular, } \mathscr{P}_2(X_1 \le x) = \beta(x).$$

(c) *If* $B_1, \cdots, B_k$ *are pairwise disjoint Borel sets and if* $N_i$ *denotes the number of* $X$*'s in* $B_i$, $i = 1, \cdots, k$, *then for* $D \in \sigma(\mathscr{B}^{\mathscr{R}})$,

(4.4) $$\mathscr{P}_2(P \in D \mid N_1 = n_1, \cdots, N_k = n_k) = \frac{E_{\mathscr{P}}(I_D(P) \prod_{j\le k} P^{n_j}(B_j))}{E_{\mathscr{P}}(\prod_{j\le k} P^{n_j}(B_j))}$$

*provided the denominator is greater than zero and provided* $\sum n_i = n$.

Proof. (a) follows since if the right-hand side of (4.2) is integrated over the set $D$ with respect to the distribution of $P$, then we obtain (4.1).

(b) follows from (4.1) by letting $D = [0, 1]^{\mathscr{R}}$.

(c) is a consequence of

$$\mathscr{P}_2(N_1 = n_1, \cdots, N_k = n_k) = \frac{n!}{n_1! \cdots n_k!} E_{\mathscr{P}}(\prod_{j\le k} P^{n_j}(B_j)),$$

which follows from (b).

Note that here and throughout, $0^0$ is understood to be 1.

The following result shows that if $F$ is tailfree with respect to the tail $(s, \infty)$, then the posterior distribution of $\{F(t): t \le u\}$, $u > s$, depends on the $X$'s only through the $X$'s smaller than or equal to $u$. This motivates the use of the word "tailfree." The result extends results of Freedman [14] page 1401.

Theorem 4.1. *Let* $F$ *be tailfree with respect to the tail* $(s, \infty)$, *let* $t_1 < t_2 < \cdots$ *be a sequence of numbers with* $t_r = s$ *for some* $r \ge 1$ *and* $\lim_{j\to\infty} t_j = \infty$. *Let* $N_i$ *denote the number of* $X$*'s in* $(t_{i-1}, t_i]$, $i = 1, 2, \cdots$; $t_0 = -\infty$. *Then for* $k \ge 0$, $C \in \mathscr{B}^{r+k}$

$$\mathscr{P}_2((F(t_1), \cdots, F(t_{r+k})) \in C \mid N_1 = n_1, N_2 = n_2, \cdots)$$
(4.5) $$= a(r+k) E_{\mathscr{P}}(I_C(F(t_1), \cdots, F(t_{r+k}))$$
$$\times \{\prod_{i\le r+k} [F(t_i) - F(t_{i-1})]^{n_i}\}[1 - F(t_{r+k})]^{m_{r+k}})$$

*where* $m_i = n - \sum_{j\le i} n_j$, $I_C$ *is the indicator function,*

$$a^{-1}(r+k) = E_{\mathscr{P}}(\{\prod_{i\le r+k} [F(t_i) - F(t_{i-1})]^{n_i}\}[1 - F(t_{r+k})]^{m_{r+k}}),$$

*and it is assumed that* $b^{-1} > 0$, *where*

(4.6) $$b^{-1} = E_{\mathscr{P}}(\prod_i [F(t_i) - F(t_{i-1})]^{n_i}).$$

Proof. Let $M = \max\{i: n_i > 0\}$. When $M \le r + k$, the result is clear by Lemma 4.1(c); consider, therefore, the case where $r + k + 1 \le M$. If $G$ is any bounded measurable function of $F(t_1), \cdots, F(t_{r+k})$, then by Lemma 3.3(c),

$$(4.7)\qquad E_{\mathscr{P}}(G \textstyle\prod_{i=r+k+1}^{M} [F(t_i) - F(t_{i-1})]^{n_i}) = E_{\mathscr{P}}\left(\prod_{i=r+k+1}^{M} \left(\frac{F(t_i) - F(t_{i-1})}{1 - F(t_{r+k})}\right)^{n_i}\right) E_{\mathscr{P}}(G[1 - F(t_{r+k})]^{m_{r+k}}) .$$

By Lemma 4.1(c), $\mathscr{P}((F(t_1), \cdots, F(t_{r+k})) \in C \mid N_1 = n_1, N_2 = n_2, \cdots) = bE_{\mathscr{P}}(I_C(F(t_1), \cdots, F(t_{r+k})) \prod_{i=1}^{M} [F(t_i) - F(t_{i-1})]^{n_i})$. The result now follows by applying (4.7) with $G = I_C(F(t_1), \cdots, F(t_{r+k})) \prod_{i=1}^{r+k} [F(t_i) - F(t_{i-1})]^{n_i}$.

REMARK 4.1. The result holds if we have only a finite number of $t$'s, $t_1, \cdots, t_l$ with $t_l > s$. In this case, we define $N_i$ as before when $i \leqq l$, and define $N_{l+1}$ to be the number of $X$'s in $(t_l, \infty)$.

Let

$$V_i(F) = [F(t_i) - F(t_{i-1})/[1 - F(t_{i-1})] \qquad \text{where } (0/0) = 1 .$$

COROLLARY 4.1. *If $F$ is neutral to the right, if $t_1 < t_2 < \cdots$ either is a finite sequence or satisfies $\lim_{j\to\infty} t_j = \infty$, and if $b^{-1} > 0$, then for $k \geqq 1$, $C \in R^k$,*

$$(4.8)\qquad \begin{aligned} &\mathscr{P}_2((F(t_1), \cdots, F(t_k)) \in C \mid N_1 = n_1, N_2 = n_2, \cdots) \\ &\quad = a(k)E_{\mathscr{P}}(I_C(F(t_1), \cdots, F(t_k))\{\textstyle\prod_{i\leq k} [F(t_i) - F(_{i-1})^{n_i}\}[1 - F(t_k)]^{m_k}) \\ &\quad = \frac{E_{\mathscr{P}}(I_C(F(t_1), \cdots, F(t_k))(\prod_{i\leq k} V_i^{n_i}(F)[1 - V_i(F)]^{m_i}))}{E_{\mathscr{P}}(\prod_{i\leq k} V_i^{n_i}(F)[1 - V_i(F)]^{m_i})} . \end{aligned}$$

PROOF. The first equality follows from Lemma 3.3(a) and Theorem 4.1. The second equality is a consequence of the first equality, (3.3) and (3.4).

REMARK 4.2. Let $\Omega$ denote the domain of $X_1, \cdots, X_n$. Then from (4.1), it follows that the set $\Omega_0 = \{\omega : E_{\mathscr{P}}\{\prod_{i=1}^{M(\omega)} [F(t_i) - F(t_{i-1})]^{N_i(\omega)}) > 0\}$ has probability one. Thus we could avoid the condition $b^{-1} > 0$ in the preceding results by defining the conditional distribution in an arbitrary measurable manner on the complement of $\Omega_0$ or by writing a.s. (almost surely) after the equalities. Ferguson's posterior corresponds to using the same formula on the complement of $\Omega_0$ as in $\Omega_0$. This makes his Bayes procedures based on the Dirichlet prior consistent where they do not deserve to be. Compare Freedman (1963) who leaves the posterior undefined outside of $\Omega_0$.

The above results are useful for "guessing" the conditional distribution of a neutral process given $X_1, \cdots, X_n$. We need some additional notation. Recall that for $t_{j-1} < t_j$ and $F(t_{j-1}) < 1$ a.s.,

$$(4.9)\qquad V_j(F) = \frac{F(t_j) - F(t_{j-1})}{1 - F(t_{j-1})} .$$

For each Borel set $A$, $H_{A,m,t_j}$ is the extension to $(R^r, \sigma(\mathscr{B}^r))$ of the measure defined on $(R^r, \mathscr{B}^r)$, by

$$(4.10)\qquad H_{A,m,t_j}(\textstyle\prod_{i=1}^{r} B_i) = E_{\mathscr{P}}(I_A(V_j(F))[1 - F(t_j)]^m \prod_{i=1}^{r} P(B_i)) , \quad \prod_{i=1}^{r} B_i \in \mathscr{B}^r ,$$

where $1 \leqq r \leqq n$ and $0 \leqq m \leqq n$. $\beta_{m,t_j}$ is the measure defined on $(R^r, \sigma(\mathscr{B}^r))$ by $\beta_{m,t_j} = H_{R,m,t_j}$, and $\beta_n$ is the probability law of $X_1, \cdots, X_n$, i.e., $\beta_n(\prod_{i=1}^n B_i) = E_{\mathscr{P}}(\prod_{i=1}^n P(B_i))$. We note that $H_{A,m,t_j}$ is absolutely continuous with respect to $\beta_{m,t_j}$ and let $(dH_{A,m,t_j}/d\beta_{m,t_j})(x_1, \cdots, x_r)$ denote the $R-N$ (Radon–Nykodym) derivative of $H_{A,m,t_j}$ with respect to $\beta_{m,t_j}$ evaluated at $(x_1, \cdots, x_r)$. From (4.1) it is clear that

$$\text{(4.11 a)} \quad H_{A,m,t_j}(\textstyle\prod_{i=1}^r B_i) = \mathscr{P}_2(V_j(F) \in A, X_1 \in B_1, \cdots, X_r \in B_r, X_{r+1} > t_j, \cdots, X_{r+m} > t_j)\,.$$

and

$$\text{(4.11 b)} \quad \beta_{m,t_j}(\textstyle\prod_{i=1}^r B_i) = \mathscr{P}_2(X_1 \in B_1, \cdots, X_r \in B_r, X_{r+1} > t_j, \cdots, X_{r+m} > t_j)\,.$$

It follows that $(dH_{A,m,t_j}/d\beta_{m,t_j})(x_1, \cdots, x_r)$ is essentially the conditional probability distribution of $V_j(F)$ given that the first $r$ $X$'s equal $x_1, \cdots, x_r$ and that $m-r$ $X$'s exceed $t_j$. We let $M_j = n - \sum_{i \leq j} N_i$ be the number of $X$'s that exceed $t_j$. Recall that $B^c$ denotes the complement of the set $B$.

THEOREM 4.2. *If $F$ is a random distribution function which is neutral to the right, then the posterior distribution of $F$ given $X_1, \cdots, X_n$ is neutral to the right. That is, if $-\infty = t_0 < \cdots < t_k < \infty$ are given with $\mathscr{P}(F(t_k) < 1) = 1$, then for Borel sets $A_1, \cdots, A_k$, $\mathscr{P}_2(V_1(F) \in A_1, \cdots, V_k(F) \in A_k \mid X_1, \cdots, X_n) = \Pr(W_1 \in A_1, \cdots, W_k \in A_k)$ a.s., where $W_1, \cdots, W_k$ are independent and $\Pr(W_j \in A_j) = \mathscr{P}_2(V_j(F) \in A_j \mid X_1, \cdots, X_n)$ a.s. Moreover, on the set $\Omega_1 = \{\omega : X_i(\omega) \in (t_{j-1}, t_j]^c, i = 1, \cdots, n\}$,*

$$\text{(4.12)} \quad \mathscr{P}_2(V_j(F) \in A_j \mid X_1, \cdots, X_n) = \frac{E_{\mathscr{P}}(I_{A_j}(V_j(F))[1-F(t_j)]^{M_j})}{E_{\mathscr{P}}([1-F(t_j)]^{M_j})} \quad \text{a.s.}$$

*On the set*

$$\Omega_2 = \{\omega : X_i(\omega) \in (t_{j-1}, t_j], i \in \{i_1, \cdots, i_r\} \neq \emptyset;\ X_i(\omega) \in (t_{j-1}, t_j]^c, i \in \{i_1, \cdots, i_r\}^c\}\,,$$

$$\text{(4.13)} \quad \mathscr{P}_2(V_j(F) \in A_j \mid X_1, \cdots, X_n) = \mathscr{P}_2(V_j(F) \in A_j \mid X_{i_1}, \cdots, X_{i_r}, M_j) = \frac{dH_{A_j,M_j,t_j}}{d\beta_{M_j,t_j}}(X_{i_1}, \cdots, X_{i_r}) \quad \text{a.s.}$$

PROOF. For convenience, we will write $V_j$ instead of $V_j(F)$. We first consider the case $n = 1$. It is enough to show that for all $B \in \mathscr{B}$,

$$\text{(4.14)} \quad \textstyle\int_B \prod_{j=1}^k \mathscr{P}_2(V_j \in A_j \mid x)\beta(dx) \doteq \mathscr{P}_2(V_1 \in A_1, \cdots, V_k \in A_k, X \in B)\,.$$

where the integrand is defined by (4.12) and (4.13), and $\beta = \beta_1$. This will follow if we can show that equality holds for all half open intervals $B_0$. We write $B_0$ as

$$\text{(4.15)} \quad B_0 = \textstyle\bigcup_{i=1}^k [(t_{i-1}, t_i] \cap B_0] \cup [(t_k, \infty) \cap B_0]\,.$$

If $B_{0,i} = [(t_{i-1}, t_i] \cap B_0]$ is nonempty, it will be denoted by $(s_{i-1}, s_i]$, $1 = 1, \cdots, k$, and we write $(s_k, s_{k+1}]$ for $B_{0,k+1} = [(t_k, \infty) \cap B_0]$ when this set is nonempty. We next show that (4.14) holds when we integrate over $(s_{i-1}, s_i]$, $i = 1, \cdots, k+1$. If $B_{0,i}$ is empty or $\mathscr{P}_2(X \in B_{0,i}) = 0$, then (4.14) trivially holds, so without loss of generality we assume that $B_{0,i}$ is nonempty and $\mathscr{P}_2(X \in (s_{i-1}, s_i]) > 0$, $i = 1, \cdots, k+1$. Let $Z_i = [F(s_i) - F(t_i)]/[1 - F(t_i)]$, $W_i = [F(s_i) - F(s_{i-1})]/[1 - F(s_{i-1})]$, then these quantities are a.s. finite and

$$\begin{aligned}\mathscr{P}_2(V_1 \in A_1, \cdots, V_k \in A_k, X \in (s_k, s_{k+1}]) &= E_{\mathscr{P}}(\textstyle\prod_{j=1}^k I_{A_j}(V_j)[\prod_{j=1}^k (1 - V_j)](1 - Z_k)W_{k+1}) \\ &= \textstyle\prod_{j=1}^k E_{\mathscr{P}}(I_{A_j}(V_j)(1 - V_j))E_{\mathscr{P}}((1 - Z_k)W_{k+1}) .\end{aligned}$$

On the other hand,

$$\begin{aligned}\int_{(s_k, s_{k+1}]} \textstyle\prod_{j=1}^k \mathscr{P}_2(V_j \in A_j \mid x)\beta(dx) &= \prod_{j=1}^k \frac{E_{\mathscr{P}}(I_{A_j}(V_j)(1 - V_j))}{E_{\mathscr{P}}(1 - V_j)} \mathscr{P}_2(X \in (s_k, s_{k+1}]) \\ &= \mathscr{P}_2(V_1 \in A_1, \cdots, V_k \in A_k, X \in (s_k, s_{k+1}])\end{aligned}$$

since $\mathscr{P}_2(X \in (s_k, s_{k+1}]) = \prod_{j=1}^k E_{\mathscr{P}}(1 - V_j)E_{\mathscr{P}}((1 - Z_k)W_{k+1})$. Similarly, for $i = 1, \cdots, k$.

$$\begin{aligned}&\mathscr{P}_2(V_1 \in A_1, \cdots, V_k \in A_k, X \in (s_{i-1}, s_i]) \\ &\quad = E_{\mathscr{P}}(\textstyle\prod_{j=1}^k I_{A_j}(V_j)[F(s_i) - F(s_{i-1})]) \\ &\quad = E_{\mathscr{P}}([\textstyle\prod_{j=1}^k I_{A_j}(V_j)][\prod_{j=1}^{i-1} (1 - V_j)](1 - Z_{i-1})W_i) \\ &\quad = \textstyle\prod_{j=1}^{i-1} E_{\mathscr{P}}(I_{A_j}(V_j)(1 - V_j))E_{\mathscr{P}}(I_{A_i}(V_i)(1 - Z_{i-1})W_i) \prod_{j=i+1}^k E_{\mathscr{P}}(I_{A_j}(V_j)) .\end{aligned}$$

Next note that as given, $\mathscr{P}_2(V_j \in A_j \mid x)$ is constant as a function of $x$ in $(t_{i-1}, t_i]$ except when $i = j$. We obtain

$$\begin{aligned}&\int_{(s_{i-1}, s_i]} \textstyle\prod_{j=1}^k \mathscr{P}_2(V_j \in A_j \mid x)\beta(dx) \\ &\quad = \prod_{j=1}^{i-1} \frac{E_{\mathscr{P}}(I_{A_j}(V_j)(1 - V_j))}{E_{\mathscr{P}}(1 - V_j)} \mathscr{P}_2(V_i \in A_i, X \in (s_{i-1}, s_i]) \prod_{j=i+1}^k E_{\mathscr{P}}(I_{A_j}(V_j)) \\ &\quad = \mathscr{P}_2(V_1 \in A_1, \cdots, V_k \in A_k, X \in (s_{i-1}, s_i])\end{aligned}$$

since

$$\begin{aligned}\mathscr{P}_2(V_i \in A_i, X \in (s_{i-1}, s_i]) &= E_{\mathscr{P}}(I_{A_i}(V_i)[\textstyle\prod_{j=1}^{i-1} (1 - V_j)](1 - Z_{i-1})W_i) \\ &= E_{\mathscr{P}}(I_{A_i}(V_i)(1 - Z_{i-1})W_i) \textstyle\prod_{j=1}^{i-1} E_{\mathscr{P}}(1 - V_j) .\end{aligned}$$

It follows that (4.14) holds for half open intervals $B_0$, and we turn to the case $n > 1$. By induction, we have that $V_1, \cdots, V_k$ are independent given $X_1, \cdots, X_n$. Thus we need only verify the formula for $\mathscr{P}_2(V_j \in A_j \mid X_1, \cdots, X_n)$, i.e., it is enough to show that

$$\begin{aligned}(4.16) \qquad \int_{B_n} \cdots \int_{B_1} \mathscr{P}_2(V_j \in A_j \mid x_1, \cdots, x_n)\beta_n(dx_1, \cdots, dx_n) \\ = \mathscr{P}_2(V_j \in A_j, X_1 \in B_1, \cdots, X_n \in B_n)\end{aligned}$$

for all half open intervals $B_i$, $i = 1, \cdots, n$. Since $B_i = (B_i \cap (-\infty, t_{j-1}]) \cup (B_i \cap (t_{j-1}, t_j]) \cup (B_i \cap (t_j, \infty))$, we can without loss of generality assume that for some $0 \leq l, r \leq n$, $B_i \subset (-\infty, t_{j-1}]$ for $i = 1, \cdots, l$, $B_i \subset (t_{j-1}, t_j]$ for $i = l + 1, \cdots, l + r$, and $B_i \subset (t_j, \infty)$ for $i = l + r + 1, \cdots, n$. We let $(r_i, s_i]$ denote $B_i$ and assume without loss of generality that $r_i < s_i$ and $\mathscr{P}_2(X_1 \in B_1, \cdots, X_n \in B_n) > 0$ since otherwise (4.16) holds trivially. Let $W(s, t) = [F(t) - F(s)]/[1 - F(s)]$,

$s \leqq t$, and $W(s_i) = W(r_i, s_i)$. Repeatedly using the independence properties defining neutral processes, we obtain

$$\begin{aligned}
\beta_n(\prod_{i=1}^n B_i) &= E_{\mathscr{P}}(\prod_{i=1}^n [F(s_i) - F(r_i)]) \\
&= E_{\mathscr{P}}(\{\prod_{i=1}^l P(B_i)\}\{[\prod_{i=l+1}^{l+r} W(s_i)(1 - W(t_{j-1}, r_i))][1 - F(t_{j-1})]\} \\
&\qquad \times \{[\prod_{i=l+r+1}^n W(s_i)[1 - W(t_j, r_i)](1 - V_j)[1 - F(t_{j-1})]\}) \\
&= E_{\mathscr{P}}([\prod_{i=1}^l P(B_i)][1 - F(t_{j-1})]^{m_{j-1}}) \\
&\qquad \times E_{\mathscr{P}}(\{\prod_{i=l+1}^{l+r} W(s_i)[1 - W(t_{j-1}, r_i)]\}[1 - V_j]^{m_j}) \\
&\qquad \times E_{\mathscr{P}}(\prod_{i=l+r+1}^n W(s_i)[1 - W(t_j, r_i)]) \\
&= \beta^{(1)}(\prod_{i=1}^l B_i)\beta^{(2)}(\prod_{i=l+1}^{l+r} B_i)\beta^{(3)}(\prod_{i=l+r+1}^n B_i)
\end{aligned}$$

where $m_{j-1} = n - l$, $m_j = n - (l + r)$, and $\beta^{(1)}$, $\beta^{(2)}$ and $\beta^{(3)}$ are the three measures defined on the Borel sets of $(-\infty, t_j]^l$, $(t_{j-1}, t_j]^r$ and $(t_j, \infty)^{n-(l+r)}$ respectively, by

$$\begin{aligned}
\beta^{(1)}(\prod_{i=1}^l B_i) &= E_{\mathscr{P}}([\prod_{i=1}^l P(B_i)][1 - F(t_{j-1})]^{m_{j-1}}) \\
\beta^{(2)}(\prod_{i=l+1}^{l+r} B_i) &= \frac{E_{\mathscr{P}}([\prod_{i=l+1}^{l+r} P(B_i)][1 - F(t_j)]^{m_j})}{E_{\mathscr{P}}([1 - F(t_{j-1})]^{m_{j-1}}} \\
\beta^{(3)}(\prod_{i=l+r+1}^n B_i) &= \frac{E_{\mathscr{P}}(\prod_{i=l+r+1}^n P(B_i))}{E_{\mathscr{P}}([1 - F(t_j)]^{m_j})} .
\end{aligned}$$

Since we can write $\mathscr{P}_2(V_j \in A_j, X_1 \in B_1, \cdots, X_n \in B_n) = E_{\mathscr{P}}(I_{A_j}(V_j) \prod_{i=1}^n [F(s_i) - F(s_{i-1})])$, the above computation shows that

$$\begin{aligned}
&\mathscr{P}_2(V_j \in A_j, X_1 \in B_1, \cdots, X_n \in B_n) \\
&\qquad = \beta^{(1)}(\prod_{i=1}^l B_i) \frac{H_{A_j, m_j, t_j}(\prod_{i=l+1}^{l+r} B_i)}{E_{\mathscr{P}}([1 - F(t_{j-1})]^{m_{j-1}}} \beta^{(3)}(\prod_{i=r+l+1}^n B_i) .
\end{aligned}$$

Note that as given, $\mathscr{P}_2(V_j \in A_j \mid x_1, \cdots, x_n)$ is constant as a function of $x_i$ in $B_i$ except when $i \in \{l + 1, \cdots, l + r\}$. Using the representation $\beta_n = \beta^{(1)}\beta^{(2)}\beta^{(3)}$ and Fubini's Theorem, we obtain for $x_i = X_i(\omega)$, $\omega \in \Omega_1$ ($r = 0$ and $m_{j-1} = m_j = M_j(\omega)$),

$$\begin{aligned}
&\int_{B_n} \cdots \int_{B_1} \mathscr{P}_2(V_j \in A_1 \mid x_1, \cdots, x_n)\beta_n(dx_1, \cdots, dx_n) \\
&\qquad = \frac{E_{\mathscr{P}}(I_{A_j}(V_j)[1 - F(t_j)]^{m_j}}{E_{\mathscr{P}}([1 - F(t_{j-1})]^{m_j})} \beta^{(1)}(\prod_{i=1}^l B_i)\beta^{(3)}(\prod_{i=l+1}^n B_i) \\
&\qquad = \mathscr{P}_2(V_j \in A_j, X_1 \in B_1, \cdots, X_n \in B_n) .
\end{aligned}$$

For $\omega \in \Omega_2$ ($r > 0$), the computation becomes

$$\begin{aligned}
&\int_{B_n} \cdots \int_{B_1} \mathscr{P}_2(V_j \in A_j \mid x_1, \cdots, x_n)\beta_n(dx_1, \cdots, dx_n) \\
&\quad = \mathscr{P}_2(V_j \in A_j, X_{l+1} \in B_{l+1}, \cdots, X_{l+r} \in B_{l+r}, X_{l+r+1} > t_j, \cdots, X_n > t_j) \\
&\qquad \times \beta^{(1)}(\prod_{i=1}^l B_i)\beta^{(3)}(\prod_{i=l+r+1}^n B_i)/E_{\mathscr{P}}([1 - F(t_{j-1})]^{m_{j-1}}) \\
&\quad = H_{A_j, m_j, t_j}(\prod_{i=l+1}^{l+r} B_i)\beta^{(1)}(\prod_{i=1}^l B_i)\beta^{(3)}(\prod_{i=l+r+1}^n B_i)/E_{\mathscr{P}}([1 - F(t_{j-1})]^{m_{j-1}}) \\
&\quad = \mathscr{P}_2(V_j \in A_j, X_1 \in B_1, \cdots, X_n \in B_n) .
\end{aligned}$$

COROLLARY 4.1. *If $F$ is neutral to the right, if $\omega \in \Omega_1$, and if $\mathscr{P}(F(t_j) < 1) = 1$ then the posterior law $\lambda_j$ of $V_j(F)$ is absolutely continuous with respect to the prior law $\mu_j$ of $V_j(F)$ and*

$$\frac{d\lambda_j}{d\mu_j}(v) = \frac{(1-v)^{M_j(\omega)}}{E_{\mathscr{P}}([1 - V_j(F)]^{M_j(\omega)})} . \tag{4.17}$$

Note that the posterior distribution of $F(t)$ can be obtained from Theorem 4.2 since if $t = t_1$ then $V_1(F) = F(t)$ a.s. The joint posterior distribution of $F(t_1), \cdots, F(t_k)$ is obtained by observing that $F(t_j) = 1 - \prod_{i \leq j}(1 - V_i(F))$, $j = 1, \cdots, k$.

REMARK 4.3. It is useful to state Theorem 4.2 in terms of limits instead of Radon–Nikodym derivatives. For any collection of numbers $x_1, \cdots, x_n$, let $n_0$ denote the number of distinct $x$'s, let $x_{(1)} < \cdots < x_{(n_0)}$ be the ordered distinct $x$'s, $n_{(j)}$ be the number of $x$'s equal to $x_{(j)}$, let $m_{(j)} = n - \sum_{i \leq j} n_{(i)}$, $j = 1, \cdots, n_0$, let $n_0(t)$ denote the number of distinct $X$'s less than or equal to $t$, let $m_0(t)$ be the number of $x$'s greater than $t$, and $t_0 = x_{(n_0(t))}$. We define $x_{(0)} = -\infty$ and

$$W_j = W(x_{(j-1)}, x_{(j)}] = \frac{F(x_{(j)}) - F(x_{(j-1)})}{1 - F(x_{(j-1)})} , \qquad j = 1, \cdots, n_0 . \tag{4.18}$$

Now we can write

$$F(t) = 1 - [\textstyle\prod_{j=1}^{n_0(t)} (1 - W_j)](1 - W_{0,t}) , \tag{4.19}$$

$$\text{where } W_{0,t} = W(t_0, t] .$$

The $W$'s are independent both under the prior and the posterior distribution of $F$. Moreover, if $\beta(x_i) - \beta(x_i - \delta) = E[F(x_i) - F(x_i - \delta)] > 0$, which is the case a.s., then we can typically obtain

$$\begin{aligned} &\mathscr{P}_2(W_j \leq w \mid x_1, \cdots, x_n) \\ &= \lim_{\delta\to 0} \frac{E_{\mathscr{P}}([1 - F(x_{(j)})]^{m(j)} I_{[0,w]}(W_j)[F(x_{(j)}) - F(x_{(j)} - \delta)]^{n(j)}}{E_{\mathscr{P}}([1 - F(x_{(j)})]^{m(j)}[F(x_{(j)}) - F(x_{(j)} - \delta)]^{n(j)}} \quad \text{a.s.} \end{aligned} \tag{4.20}$$

when this limit exists. If $t \in \{x_1, \cdots, x_n\}$, then $W_{0,t} = 0$. If $t \notin \{x_1, \cdots, x_n\}$, then

$$\mathscr{P}_2(W_{0,t} \leq w \mid x_1, \cdots, x_n) = \frac{E_{\mathscr{P}}([1 - F(t)]^{m_0(t)} I_{[0,w]}(W_{0,t}))}{E_{\mathscr{P}}([1 - F(t)]^{m_0(t)})} \quad \text{a.s.} \tag{4.21}$$

Also note that $E_{\mathscr{P}_2}(W_j \mid x_1, \cdots, x_n)$ can be obtained by replacing $I_{[0,w]}(W_j)$ in (4.20) by $W_j$, while $E_{\mathscr{P}_2}(W_{0,t} \mid x_1, \cdots, x_n)$ results if we replace $I_{[0,w]}(W_{0,t})$ by $W_{0,t}$ in (4.21). Sometimes it is useful to write $(1 - W_j) = (1 - W_{j\delta})(1 - Z_{j\delta})$, where $(1 - W_{j\delta}) = [1 - F(x_{(j)} - \delta)]/[1 - F(x_{(j-1)})]$, $(1 - Z_{j\delta}) = [1 - F(x_{(j)})]/[1 - F(x_{(j)} - \delta)]$, and $\delta \in (0, x_{(j)} - x_{(j-1)})$. $W_{j\delta}$ and $Z_{j\delta}$ are independent given $x_1, \cdots, x_n$ and we obtain

$$\begin{aligned} &E_{\mathscr{P}_2}(1 - W_j \mid x_1, \cdots, x_n) \\ &= \lim_{\delta\to 0+} \frac{E_{\mathscr{P}}([1 - W_{j\delta}]^{m(j-1)+1})}{E_{\mathscr{P}}([1 - W_{j\delta}]^{m(j-1)})} \lim_{\delta\to 0+} \frac{E_{\mathscr{P}}([1 - Z_{j\delta}]^{(m(j)+1)} Z_{j\delta}^{n(j)})}{E_{\mathscr{P}}([1 - Z_{j\delta}]^{m(j)} Z_{j\delta}^{n(j)})} . \end{aligned} \tag{4.22}$$

EXAMPLE 4.1. Let $F(t) = 1 - \exp[-Y(t)]$ be one of the random distribution functions defined in Example 3.1 by (3.8) and (3.9). We will compute $E_{\mathscr{P}_2}(F(t) \mid X_1, \cdots, X_n)$, which is the Bayes estimate of $F$ for the loss function $L(F, \hat{F}) = \int (F(t) - \hat{F}(t))^2 \lambda(dt)$, where $\lambda$ is a finite measure on $(R, \mathscr{B})$.

Letting $E$ denote $E_{\mathscr{P}}$, we obtain

$$E([1 - W_{j\delta}]^m) = E(\exp\{-m[Y(x_{(j)} - \delta) - Y(x_{(j-1)})]\}) \\ = [M_Y(-m)]^{[\gamma(x_{(j)}-\delta)-\gamma(x_{(j-1)})]} ,$$

and

$$\lim_{\delta\to 0+} \frac{E([1 - W_{j\delta}]^{m_{(j-1)}+1})}{E([1 - W_{j\delta}]^{m_{(j-1)}})} = \left\{\frac{M_Y[-(m_{(j-1)} + 1)]}{M_Y[-m_{(j-1)}]}\right\}^{[\gamma(x_{(j)})-\gamma(x_{(j-1)})]}$$

where $\gamma$ is assumed to be continuous at $x_{(j)}$.

Using the binomial expansion for $Z_{j\delta}^{n_{(j)}} = [1 - (1 - Z_{j\delta})]^{n_{(j)}}$, we have

$$E([1 - Z_{j\delta}]^{m_{(j)}} Z_{j\delta}^{n_{(j)}}) \\ = E(\textstyle\sum_{i=0}^{n_{(j)}} \binom{n_{(j)}}{i}(-1)^i(1 - Z_{j\delta})^{m_{(j)}+i}) \\ = E(\textstyle\sum_{i=0}^{n_{(j)}} \binom{n_{(j)}}{i}(-1)^i \exp\{-(m_{(j)} + i)[Y(x_{(j)}) - Y(x_{(j)} - \delta)]\}) \\ = \textstyle\sum_{i=0}^{n_{(j)}} \binom{n_{(j)}}{i}(-1)^i\{M_Y[-(m_{(j)} + i)]\}^{[\gamma(x_{(j)})-\gamma(x_{(j)}-\delta)]} .$$

It follows that if $\gamma$ has a derivative $\gamma'$ with $\gamma'(x_{(j)}) > 0$, then

$$\lim_{\delta\to 0+} \frac{E([1 - Z_{j\delta}]^{m_{(j)}+1} Z_{j\delta}^{n_{(j)}})}{E([1 - Z_{j\delta}]^{m_{(j)}} Z_{j\delta}^{n_{(j)}})} \\ = \frac{\sum_{i=0}^{n_{(j)}} \binom{n_{(j)}}{i}(-1)^i \log M_Y[-(m_{(j)} + i + 1)]}{\sum_{i=0}^{n_{(j)}} \binom{n_{(j)}}{i}(-1)^i \log M_Y[-(m_{(j)} + i)]} .$$

We note that the right-hand side is independent of $\gamma$, and denote it by $C(n_{(j)}, m_{(j)}, M_Y)$. Thus, for the case that all the $x_i$ are different

$$C(1, n - j, M_Y) = \frac{\log M_Y[-(n - j + 1)] - \log M_Y[-(n - j + 2)]}{\log M_Y[-(n - j)] - \log M_Y[-(n - j + 1)]} .$$

We also need (see (4.19) and (4.21))

$$E_{\mathscr{P}_2}((1 - W_{0,t}) \mid x_1, \cdots, x_n) = \left\{\frac{M_Y[-m_0(t) + 1)]}{M_Y[-m_0(t)]}\right\}^{[\gamma(t)-\gamma(t_0)]} .$$

Putting these results together, we have

$$E_{\mathscr{P}_2}([1 - F(t)] \mid x_1, \cdots, x_n) \\ = \textstyle\prod_{j=1}^{n_0(t)} C(n_{(j)}, m_{(j)}, M_Y) \left\{\frac{M_Y[-(m_{(j-1)} + 1)]}{M_Y[-m_{(j-1)}]}\right\}^{[\gamma(x_{(j)})-\gamma(x_{(j-1)})]} \\ \times \left\{\frac{M_Y[-(m_0(t) + 1)]}{M_Y[-m_0(t)]}\right\}^{[\gamma(t)-\gamma(t_0)]} . \tag{4.23}$$

If $Y(t)$ is a gamma process with independent increments, i.e., $M_Y(v) = (1 - v)^{-1}$, and if $n_{(j)} = 1, j = 1, \cdots, n$, then (4.23) simplifies to

$$E_{\mathscr{P}_2}([1 - F(t)] \mid x_1, \cdots, x_n)$$

$$= \left\{ \prod_{j=1}^{nF_n(t)} \frac{\log\left(\frac{n-j+2}{n-j+3}\right)}{\log\left(\frac{n-j+1}{n-j+2}\right)} \left\{\frac{n-j+3}{n-j+2}\right\}^{-[\gamma(x_{(j)})-\gamma(x_{(j-1)})]} \right\}$$

$$\times \left\{\frac{n - nF_n(t) + 2}{n - nF_n(t) + 1}\right\}^{-[\gamma(t)-\gamma(t_0)]}$$

where $F_n(t) = n^{-1}[n_0(t)]$ is the sample distribution function. When $n = 1$, this yields

$$E_{\mathscr{P}_2}(F(t) \mid x) = 1 - (\tfrac{3}{2})^{-\gamma(t)} \qquad \text{for} \quad x > t$$

$$= 1 - (.58)(\tfrac{3}{4})^{-\gamma(x)}2^{-\gamma(t)} \qquad \text{for} \quad x \leqq t .$$

## REFERENCES


[1] Antoniak, C. (1969). Mixtures of Dirichlet processes with applications to some nonparametric problems. Ph. D. Dissertation, Univ. of California, Los Angeles.

[2] Blackwell, D. (1973). Discreteness of Ferguson selections. *Ann. Statist.* **1** 356-358.

[3] Blackwell, D. and MacQueen, J. B. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Statist.* **1** 353-355.

[4] Breiman, L. (1968). *Probability.* Addison-Wesley, Reading.

[5] Connor, R. J. and Mosimann, J. E. (1969). Concepts of independence of proportions with a generalization of the Dirichlet distribution. *J. Amer. Statist. Assoc.* **64** 194-206.

[6] Darroch, J. N. and Ratcliff, D. (1971). A characterization of the Dirichlet distribution. *J. Amer. Statist. Assoc.* **66** 641-643.

[7] Doksum, K. A. (1971). Tailfree and neutral processes and their posterior distributions. ORC Report 71-72, Univ. of Calif., Berkeley.

[8] Doksum, K. A. (1972). Decision theory for some nonparametric models. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **1**. Univ. of California, Berkeley, 331-343.

[9] Dubins, L. E. and Freedman, D. A. (1966). Random distribution functions. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **2**. Univ. of California, Berkeley, 183-214.

[10] Fabius, J. (1964). Asymptotic behavior of Bayes estimates. *Ann. Math. Statist.* **35** 846-856.

[11] Fabius, J. (1973). Neutrality and Dirichlet distributions. *Transactions of the Sixth Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes.* 175-181.

[12] Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209-230.

[13] Ferguson, T. S. and Klass, M. J. (1972). A representation of independent increment processes without Gaussian components. *Ann. Math. Statist.* **43** 1634-1643.

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

[15] Kraft, C. H. (1964). A class of distribution function processes which have derivatives. *J. Appl. Probability* **1** 385-388.

[16] Kraft, C. H. and van Eeden, C. (1964). Bayesian bio-assay. *Ann. Math. Statist.* **35** 886-890.

[17] Metivier, M. (1971). Sur la construction de mesures aléatoires presque sûrement absolument continues par rapport à une measure donnée. *Z. Wahrsheinlichkeitstheorie und Verw. Gebiete* **20** 332-344.



Department of Statistics
University of California
Berkeley, California 94720