POISSON PROCESSES WITH RANDOM ARRIVAL RATE[1]

By David A. Freedman

University of California, Berkeley

1. Introduction. Let F be a distribution function on $(0, \infty)$. A probability P_F on the integers defined by

$$P_F(n) = (n!)^{-1} \int_0^\infty e^{-\lambda}\lambda^n \, dF, \qquad n \geqq 0,$$

will be called a mixture of Poisson probabilities. Since [1] there is a 1-1 correspondence between P_F and F, any statistical question about F can, in principle, be answered by random sampling on P_F. However, F can be estimated more easily by random sampling on mixtures of laws of Poisson processes (to be defined below). Even then no unbiased estimate for F exists; but the Glivenko-Cantelli Lemma [2], p. 20 does hold for the natural estimate of a continuous F. These two results are proved in Section 3; Section 2 contains some preliminary material.

2. Independent realizations of mixtures. Let γ be a nonempty set, and $B(\gamma)$ a σ-algebra of subsets of γ. Let $\{P_{\lambda'} : \lambda' \varepsilon \Lambda\}$ be a family of probabilities defined on $B(\gamma)$. Take $B(\Lambda)$ to be the smallest σ-algebra of subsets of Λ over which all the functions $\{P_{\lambda'}(E) : E \varepsilon B(\gamma)\}$ are measurable. If μ is any probability on $B(\Lambda)$, define

$$P_\mu(E) = \int_\Lambda P_{\lambda'}(E) \, d\mu : E \varepsilon B(\gamma).$$

The set function P_μ is again a probability on $B(\gamma)$, and is called a mixture of the probabilities $P_{\lambda'}$. If X is any $B(\gamma)$-measurable function, and P any probability on $B(\gamma)$, define $E(X \mid P) = \int_\gamma X \, dP$. Then

Lemma 1. $E(X \mid P_\mu) = \int_\Lambda E(X \mid P_{\lambda'}) \, d\mu$, *in the sense that if either side exists, both do and they are equal.*

Proof. When X is the characteristic function of a measurable set, the lemma is a restatement of the definition. Hence the lemma holds for all simple functions by linearity, for nonnegative functions by a monotone passage to the limit, and finally for general functions by linearity.

The purpose of the next lemma is to describe mixtures on product spaces. Define ([2], pp. 90–91)

$$(\gamma^J, B(\gamma^J), P^J) = \prod_{j=1}^{J} (\gamma, B(\gamma), P)$$

$$(\Lambda^J, B(\Lambda^J), \mu^J) = \prod_{j=1}^{J} (\Lambda, B(\Lambda), \mu),$$

Received May 3, 1961; revised March 12, 1962.

[1] This paper was prepared with the partial support of the National Science Foundation Grant (G 14648).

where J is any natural number or ∞. If $s \varepsilon \gamma^J(\lambda \varepsilon \Lambda^J)$, its jth coordinate will be written $s_j(\lambda_j)$. The convention adopted here (for typographical reasons) is the following: λ is a point of a product space whose factors are Λ; λ' and λ_j are points of λ. Let $P_\lambda = \prod_{j=1}^J P_{\lambda_j}$ for $\lambda \varepsilon \Lambda^J$.

LEMMA 2. *For any $E \varepsilon B(\gamma^J)$, $P_\lambda(E)$ is $B(\gamma^J)$-measurable and*

$$P^J(E) = \int_{\Lambda^J} P_\lambda(E)\, d\mu^J.$$

PROOF. Both assertions hold for (finite dimensional) measurable rectangles by Fubini's theorem, and are preserved under complementation and countable unions of the sets E for which they are ture.

LEMMA 3. *For each $j \varepsilon Z$ (the set of nonnegative integers) let $\{Y_{j,n} : n \varepsilon Z\}$ be a stochastic process on $(\gamma^\infty, B(\gamma^\infty))$. Then $lim_{j,n\to\infty} Y_{j,n} = 0$ a.e. $[P_\mu^\infty]$, if and only if, for almost all sequences $\lambda \varepsilon \Lambda^\infty[\mu^\infty]$, $\lim_{j,n\to\infty} Y_{j,n} = 0$ a.e. $[P_\lambda]$.*

PROOF. For $m \varepsilon Z$ let $E_m = \{s : s \varepsilon \gamma^\infty$ and for any j_0 and n_0 there exist $j > j_0$ and $n > n_0$ with $|Y_{j,n}(s)| > 1/m\}$. Then each statement below is equivalent to the one following it.

(i) $Y_{j,n} \to 0$ a.e. $[P_\mu^\infty]$.

(ii) $P_\mu^\infty(E_m) = 0$, all m.

(iii) $\int_{\Lambda^\infty} P_\lambda(E_m)\, d\mu^\infty = 0$, all m.

(iv) $P_\lambda(E_m) = 0$, $[\mu^\infty]$, all m.

(v) $P_\lambda(E_m) = 0$, all m, $[\mu^\infty]$.

(vi) $Y_{j,n} \to 0$ a.e. $[P_\lambda][\mu^\infty]$.

3. Poisson processes. Let $B(Z)$ denote the family of all subsets of Z, and define $(\gamma, B(\gamma)) = \prod_1^\infty (Z, B(Z))$. If $s \varepsilon \gamma$, its nth coordinate will be written s_n. Define (using [2], p. 93, Theorem A) $P_{\lambda'}$ as the probability on $B(\gamma)$ making $X_n(s) = s_n$, $n \geqq 1$, $X_0(s) = 0$ a Poisson process with parameter λ' for $\lambda' \varepsilon \Lambda = (0, \infty)$; $B(\Lambda)$ is easily seen to consist of the Borel subsets of Λ. If μ is a probability on $B(\Lambda)$, its distribution function will be denoted by F.

When P_μ is constructed on $B(\gamma)$, the law of the process $\{X_n\}$ is called a mixture of laws of Poisson processes, with mixing distribution μ.

Recall that $(\gamma^\infty, B(\gamma^\infty)) = \prod_1^\infty (\gamma, B(\gamma)) = \prod_1^\infty \prod_1^\infty (Z, B(Z))$ so that γ^∞ is the space of infinite matrices of nonnegative integers.

For $s \varepsilon \gamma^\infty$, s_j is the jth coordinate of s and is a point of γ. Hence $s_{j,n}$ is the nth coordinate of the jth of s, and is an integer (namely, the entry in the jth row and nth column of s). Define $X_{j,n}(s) = s_{j,n}$, $n \geqq 1$; $X_{j,0}(s) = 0 : j \geqq 1$. In the balance of the paper, the probability on $(\gamma^\infty, B(\gamma^\infty))$ will be P_μ^∞.

Less formally, there is an unknown random mechanism which selects a parameter $\lambda_j \varepsilon \Lambda = (0, \infty)$ according to the prior probability μ. This is done repeatedly and independently, which corresponds to selecting a point $\lambda = \{\lambda_1, \lambda_2, \cdots\}$

from Λ^∞ according to probability μ^∞. Once a λ_j has been selected, the statistician observes the evolution of a discrete-time Poisson process $\{X_{j,n} : n = 1, 2, \cdots\}$ with arrival rate λ_j. The processes are independent in j. These processes are defined on the common probability space $(\gamma^\infty, B(\gamma^\infty), P_\mu^\infty)$, and are a sequence of independent realizations of a process whose law is a mixture of Poisson laws (with mixing distribution μ). Having observed $\{X_{j,n} : 1 \leqq n \leqq N\}$ for $1 \leqq j \leqq J$, the statistician wishes to make inferences about μ. The first theorem gives a limitation on the type of estimates available.

THEOREM 1. *For fixed $x > 0$, $F(x)$ has no unbiased estimate measurable on a finite number of the $X_{j,n}$.*

PROOF. Let J be the largest j-subscript available, and for $1 \leqq j \leqq J$ let n_j be the largest n-subscript for each j. Thus $1 \leqq J < \infty, 0 \leqq n_j < \infty$.

It suffices to consider functions of the sufficient statistic $\{X_{j,n_j} : 1 \leqq j \leqq J\}$. Let T be a function on Z^J, and suppose that, contrary to the theorem,

$$E[T(X_{j,n_j} : 1 \leqq j \leqq J) \mid P_\mu^J] = F(x),$$

for all μ (or even all μ with carriers of J points).

By Lemmas 1 and 2, the finiteness of $E(T \mid P_\mu^J)$ implies the finiteness of $E(T \mid P_\lambda)$ for almost all vectors $\lambda \varepsilon \Lambda^J[\mu^J]$. Since this holds for all μ, $|E(T \mid P_\lambda)| < \infty$ for all $\lambda \varepsilon \Lambda^J$. Then $E(T \mid P_\lambda)$ is $\exp(-n_1\lambda_1 - \cdots - n_J\lambda_J)$ times a multiple power series, absolutely and uniformly convergent on any bounded set in Λ^J; and is a continuous function of λ.

Let μ_y assign mass 1 to $y > 0 : F_y$ being the corresponding distribution function. Then $\int_{\Lambda^J} E(T \mid P_\lambda)\, d\mu_y^J = F_y(x)$, and the left-hand side is continuous, as a function of y, while the right-hand side, as a function of y, is discontinuous at $y = x$; a contradiction which completes the proof.

On the other hand, for large n, $\{n^{-1}X_{j,n} : 1 \leqq j \leqq J\}$ is approximately a random sample from F, and the sample cumulative distribution function provides a natural estimate for F. Let $f(y) = f(y, x) = 1(y \leqq x)$ and $O(y > x)$ and put $F_{J,n}(x) = J^{-1} \sum_{j=1}^J f(n^{-1}X_{j,n})$.

THEOREM 2. *If F is continuous at x, $\lim_{J,n\to\infty} F_{J,n} = F(x)$* a.e. $[P_\mu^\infty]$.

PROOF. By Lemma 3, this is equivalent to showing that $F_{J,n}(x) \to F(x)$ a.e. $[P_\lambda]$ for almost all $[\mu^\infty]$ sequences $\lambda \varepsilon \Lambda^\infty$.

Choose $\epsilon > 0$ and $\delta > 0$ so that $F(x + \delta) - F(x - \delta) < \epsilon$. The idea of the proof is to discard $\lambda_j \varepsilon (x - \delta, x + \delta)$, committing only a small error. Outside this region the Markov inequality gives sharp enough estimates to secure the theorem.

The construction is in terms of the following functions (whose dependence on ϵ, δ, and x is understood):

$$f_1(\lambda) = 1, \lambda \leqq x - \delta; = 0, \qquad \lambda > x - \delta$$

$$f_2(\lambda) = 0, \lambda < x + \delta; = 1, \qquad \lambda \geqq x + \delta$$

$$f_3(\lambda) = 1, \lambda \varepsilon (x - \delta, x + \delta); = 0, \quad \lambda \not\varepsilon (x - \delta, x + \delta)$$

$$u_i(\lambda) = \lambda(\lambda - x)^{-4} f_i(\lambda), \qquad i = 1, 2$$

$$v_i(\lambda) = \lambda^2(\lambda - x)^{-4} f_i(\lambda), \qquad i = 1, 2$$

$$Q(n, \lambda) = P_\lambda(X_{j,n} \leqq nx).$$

By the strong law of large numbers [2], p. 239, $J^{-1} \sum_{j=1}^{J} \xi(\lambda_j) \to \int_\Lambda \xi \, d\mu$ a.e. $[\mu^\infty]$ simultaneously for $\xi = f_i$, $i = 1, 2, 3$, $= u_i$, $= v_i$, $i = 1, 2$, since these functions are bounded and therefore summable. Let $N_{\epsilon,\delta}$ be the exceptional μ^∞-null set. Select a sequence $\epsilon_k \to 0$ with corresponding $\delta_k \to 0$. Put $N = \bigcup_{k=1}^{\infty} N_{\epsilon_k,\delta_k}$, so that $\mu^\infty(N) = 0$. In what follows, $\lambda \, \varepsilon \, \Lambda^\infty - N$, while (ϵ, δ) takes values in the sequence (ϵ_k, δ_k). Two estimates of $P(n, \lambda)$ are required. By the Markov inequality [2], p. 158, with $r = 4$ and $(n^{-1}X_{j,n} - \lambda)$ for X, if $\lambda \leqq x - \delta$

$$(1) \qquad 1 - Q(n, \lambda) \leqq n^{-3} u_1(\lambda) + 3n^{-2} v_1(\lambda),$$

while if $\lambda \geqq x + \delta$,

$$(2) \qquad Q(n, \lambda) \leqq n^{-3} u_2(\lambda) + 3n^{-2} v_2(\lambda).$$

These estimates can be used to prove

$$(3) \qquad \lim_{J,n\to\infty} J^{-1} \sum_{j=1}^{J} Q(n, \lambda_j) = F(x) \text{ a.e. } [\mu^\infty].$$

Indeed, let $A_i = J^{-1} \sum_{j=1}^{J} Q(n, \lambda_j) f_i(\lambda_j)$, $i = 1, 2, 3$. By inequality (1), with $0 \leqq \theta \leqq 1$,

$$A_1 = J^{-1} \sum_{j=1}^{J} f_1(\lambda_j) - \theta n^{-3} J^{-1} \sum_{j=1}^{J} u_1(\lambda_j) - 3\theta n^{-2} J^{-1} \sum_{j=1}^{J} v_1(\lambda_j),$$

the first term converging to $F(x - \delta)$, the second and third to 0. Similarly, (2) implies that $A_2 \to 0$, and clearly $\limsup_{J,n\to\infty} A_3 < \epsilon$. In summary

$$\liminf_{J,n\to\infty} J^{-1} \sum_{j=1}^{J} Q(n, \lambda_j) \geqq F(x - \delta) > F(x) - \epsilon,$$

$$\limsup_{J,n\to\infty} J^{-1} \sum_{j=1}^{J} Q(n, \lambda_j) < F(x - \delta) + \epsilon < F(x) + \epsilon.$$

Allowing $k \to \infty$, so that $\epsilon \to 0$, completes the proof of (3).

The next step is to prove

$$(4) \qquad \lim_{J,n\to\infty} J^{-1} \sum_{j=1}^{J} [f(n^{-1}X_{j,n}) - Q(n, \lambda_j)] = 0 \text{ a.e. } [P_\lambda].$$

As before, put $B_i = J^{-1} \sum_{j=1}^{J} [f(n^{-1}X_{j,n}) - Q(n, \lambda_j)] f_i(\lambda_j)$, $i = 1, 2, 3$. Then

$$E\{[f(n^{-1}X_{j,n}) - Q(n, \lambda_j)] f_2(\lambda_j) \mid P_{\lambda_j}\} = 0$$

and

$$E\{[f(n^{-1}X_{j,n}) - Q(n, \lambda_j)]^2 f_2(\lambda_j) \mid P_{\lambda_j}\} = Q(n, \lambda_j)(1 - Q(n, \lambda_j))f_2(\lambda_j)$$

$$\leqq Q(n, \lambda_j)f_2(\lambda_j) \tag{5}$$

$$\leqq n^{-3}u_2(\lambda_j) + 3n^{-2}v_2(\lambda_j)$$

by (2).

Let $B(J, n) = JB_2$. To show $\lim_{J,n\to\infty} B_2 = 0$, a.e. $[P_\lambda]$ select $\eta > 0$ and define the sets $C_{r,n} = \{s : s \in \gamma^\infty$ and there is a J with (i) $2^{r-1} < J \leqq 2^r$, (ii) $B(J, n) \geqq J\eta\}$, $D_{r,n} = \{s : s \in \gamma^\infty$ and there is a J with (i) $1 \leqq J \leqq 2^r$, (ii) $B(J, n) \geqq \eta/2\cdot 2^r\}$. Then $C_{r,n} \subset D_{r,n}$ so that $\sum_{r,n} P_\lambda(C_{r,n}) \leqq \sum_{r,n} P_\lambda(D_{r,n})$. Now by Kolmogorov's inequality [2], p. 235, and the estimate (5)

$$P_\lambda(D_{r,n}) \leqq 4\eta^{-2}2^{-2r} \sum_{j=1}^{2^r} [n^{-3}u_2(\lambda_j) + 3n^{-2}v_2(\lambda_j)],$$

so that $\sum_{r,n} P_\lambda(C_{r,n}) < \infty$. Hence by the Borel-Cantelli Lemma [2], p. 228, $B_2 \geqq \eta$ only finitely often $[P_\lambda]$; allowing $\eta \to 0$ through some sequence of values shows that $\lim_{J,n\to\infty} B_2 = 0$ a.e. $[P_\lambda]$. Similarly, $\lim_{J,n\to\infty} B_1 = 0$ a.e. $[P_\lambda]$ and clearly $\limsup_{J,n\to\infty} |B_3| < \epsilon$ a.e. $[P_\lambda]$. These facts show that

$$\limsup_{J,n\to\infty} J^{-1} \sum_{j=1}^{J} [f(n^{-1}X_{j,n}) - Q(n, \lambda_j)] < \epsilon \text{ a.e. } [P_\lambda],$$

and allowing $k \to \infty$ establishes (4).

Finally, combining (3) and (4) gives

$$\lim_{J,n\to\infty} F_{J,n}(x) =$$

$$\lim_{J,n\to\infty} J^{-1} \sum_{j=1}^{J} Q(n, \lambda_j) + \lim_{J,n\to\infty} J^{-1} \sum_{j=1}^{J} [f(n^{-1}X_{j,n}) - Q(n, \lambda_j)]$$

$$= F(x), \text{ a.e. } [P_\lambda],$$

which is the requisite conclusion.

COROLLARY. *If F is continuous,*

$$\lim_{J,n\to\infty} (\sup_{-\infty<x<\infty} |F_{J,n}(x) - F(x)|) = 0 \text{ a.e. } [P_\mu^\infty].$$

The condition that F be continuous is indispensable. Indeed, by the central limit theorem, $\lim_{n\to\infty} Q(n, x) = \frac{1}{2}$; moreover, for a rapidly increasing sequence n_ν, the events $n_\nu^{-1}X_{J,n_\nu} \leqq x$ are almost independent. Put $f_4(\lambda) = 1$ or 0 according as $\lambda = x$ or not. Then by a slight modification of the Borel-Cantelli Lemma, for any J

$$\sum_{j=1}^{J} f(n^{-1}X_{j,n_\nu})f_4(\lambda_j) / \sum_{j=1}^{J} f_4(\lambda_j)$$

is equal to 0 infinitely often and equal to 1 infinitely often as $\nu \to \infty$, $[P_\lambda]$. Hence

$$\liminf_{J,n\to\infty} F_{J,n} = F(x-),$$

$$\limsup_{J,n\to\infty} F_{J,n} = F(x) \text{ a.e. } [P_\mu^\infty].$$

5. Acknowledgments. This paper is based on the author's doctoral dissertation, written under the very helpful supervision of Profssor W. Feller. The author also wishes to acknowledge the generous financial support provided by McGill University and the RAND Corporation.

REFERENCES

[1] FELLER, W. (1943). On a class of contagious distributions. *Ann. Math. Statist.* **14** 389–399.

[2] LOÈVE, M. (1960). *Probability Theory*, 2nd Ed. Van Nostrand, Princeton.