A LARGE SAMPLE TEST FOR THE INDEPENDENCE OF TWO RENEWAL PROCESSES[1]

By Sidney C. Port and Charles J. Stone

University of California, Los Angeles

The purpose of this paper is to develop a large sample test for the independence of two renewal processes. Our original motivation was the suggestion of D. H. Perkel of The RAND Corporation that such tests are needed in connection with certain neurophysiological experiments. There are other areas, however, in which such a test is desired (e.g., in reliability and maintenance procedures for stochastically failing equipment).

Let V_n, $n \geqq 0$, be independent random variables such that V_n, $n \geqq 1$, are positive, identically distributed, and have finite mean. Let W_n, $n \geqq 0$, be another such sequence. Set $S_0 = T_0 = 0$ and for $n \geqq 1$ set $S_n = V_1 + \cdots + V_n$ and $T_n = W_1 + \cdots + W_n$. Also set

$$S_n' = V_0 + S_n,$$

$$T_n' = W_0 + T_n,$$

$$Z(t) = \min \{T_n - t \mid T_n > t\},$$

$$Z'(t) = \min \{T_n' - t \mid T_n' > t\},$$

and

$$X_n = Z'(S_n'), \qquad n \geqq 0.$$

If we think of S_n' and T_n' as the time of the $(n + 1)$st occurrence of two renewal processes, then X_n is the time elapsed between the $(n + 1)$st renewal of the first process and the next renewal of the second process.

Suppose now that V_n, $n \geqq 0$, and W_n, $n \geqq 0$, are independent. It then follows easily that X_n, $n \geqq 0$, is a Markov process on $(0, \infty)$ having stationary n-step transition functions given by

$$\mathrm{P}_n(x, A) = P(Z(S_n - x) \,\varepsilon\, A) + P(S_n \,\varepsilon\, x - A),$$

and furthermore that the process has an invariant probability measure Π defined by

$$\Pi(A) = \int_A (EW_1)^{-1} P(W_1 > \mathrm{t})\, dt.$$

If V_0 has the distribution with density $(EV_1)^{-1}P(V_1 > t)$ and W_0 has the distribution Π, then X_0 has the distribution Π and hence X_n, $n \geqq 0$, is a strictly stationary process. For a bounded function f on $(0, \infty)$, let $E_\pi f(X_0)$, $\mathrm{Var}_\pi f(X_0)$,

Received 2 September 1966.

[1] This research is sponsored by the United States Air Force under Project RAND-Contract No. AF 49(638)-1700.

and $\mathrm{Cov}_\Pi\,(f(X_0), f(X_k))$ denote, respectively, the mean of $f(X_0)$, the variance of $f(X_0)$, and the covariance of $f(X_0)$ and $f(X_k)$ when X_0 has as its distribution the invariant measure Π. In particular, if f is the characteristic function of a set A, then $E_\Pi f(X_0) = \Pi(A)$.

For any Borel set A let $N_n(A)$ denote the number of X_k, $0 \leqq k \leqq n$, that lie in A. If the two renewal processes are independent, then for n large, $N_n(A)$ should be close to $(n + 1)\Pi(A)$.

Consequently, an intuitively appealing procedure for testing the hypothesis that the two processes are independent is to choose n, A, and a positive constant C and reject the hypothesis if and only if

$$|N_n(A) - (n + 1)\Pi(A)| \geqq C.$$

More generally, we can let f denote a bounded measurable function and compare $\sum_0^n f(X_k)$ with $(n + 1)E_\Pi f(X_0)$. Our main result is that under the hypothesis of independence and under mild further conditions on the distributions of V_1 and W_1, a central limit theorem holds for $\sum_0^n f(X_k)$.

THEOREM. *Suppose the sequences V_n, $n \geqq 0$, and W_n, $n \geqq 0$, are independent. Suppose also that for some $m > 2$, $E\,|V_1|^m < \infty$ and $E\,|W_1|^m < \infty$ and that for some $n \geqq 1$, T_n has a non-singular distribution. Then for all bounded real-valued measurable functions f on $(0, \infty)$, as $n \to \infty$*

$$(n + 1)^{-\frac{1}{2}}[\textstyle\sum_{k=0}^n f(X_k) - (n + 1)E_\Pi f(X_0)] \to_{\text{dist}} N(0, \sigma_f^2),$$

where $N(0, \sigma_f^2)$ is the normal distribution with mean 0 and variance

$$\sigma_f^2 = \mathrm{Var}_\Pi f(X_0) + 2 \textstyle\sum_{k=1}^\infty \mathrm{Cov}_\Pi\,[f(X_0), f(X_k)].$$

The somewhat complicated proof to follow is necessary because the Markov process X_n does not in general satisfy Doeblin's condition. Indeed, Doeblin's condition fails to be satisfied whenever W_1 is unbounded.

The referee has pointed out that our theorem follows from Theorem 5.1 of Orey [4]. This theorem is not correct as stated—the conclusion holds only under the additional assumption that, in Orey's notation, $\sigma^2 = E_{Q_A}\{\bar{f}^*(Y_1))^2\} < \infty$. It is easy to construct examples such that $\sigma^2 = \infty$ and all the assumptions of Theorem 5.1 hold. We know of no direct way of showing that under our assumptions $\sigma^2 < \infty$. Any such method would be most interesting and would of course yield an alternative proof of the above theorem.

PROOF OF THEOREM. If X_0 has the invariant distribution Π, then $f(X_n)$, $n \geqq 0$, is a bounded strictly stationary process. A central limit theorem for such processes, assumed to satisfy some further conditions, has been obtained by Ibragimov [1], p. 365. The main part of the proof below is to show that these further conditions are satisfied. In doing so we will use some estimates of the renewal function of Stone [3] and some estimates of probabilities of large deviations obtained by Nagaev [2]. The case of arbitrary initial distributions will easily be reduced to the stationary case.

We assume below that all the conditions of the theorem hold.

LEMMA 1. *As* $t \to \infty$,

$$\int_0^\infty |P(Z(t) \varepsilon dy) - \Pi(dy)| = O(\int_{t/2}^\infty uP(W_1 \varepsilon du) + t^{-2}).$$

PROOF OF LEMMA 1. Let ν denote the renewal measure for the $\{T_n\}$ process defined by

$$\nu(A) = \sum_0^\infty P(T_n \varepsilon A).$$

Since $E \quad |W_1|^2 < \infty$ and, for some $n \geqq 1$, T_n has a non-singular distribution, it follows from a theorem of Stone [3] that $\nu(ds) = \nu_1(ds) + p(s)\,ds$, where ν_1 is a finite measure having finite second moment, p is continuous and non-negative, and as $s \to \infty$

$$|p(s) - (EW_1)^{-1}| = O(\int_s^\infty uP(W_1 \varepsilon du) + s^{-2}).$$

From the formula $P(Z(t) \varepsilon A) = \int_0^t P(W \varepsilon A + t - s)\nu(ds)$ we obtain a decomposition of the distribution of $Z(t)$ as follows:

$$P(Z(t) \varepsilon A) = \psi_t(A) + \psi_t'(A),$$

where

$$\psi_t(A) = \int_0^t P(W \varepsilon A + t - s)\nu_1(ds);$$
$$\psi_t'(A) = \int_0^t P(W \varepsilon A + t - s)p(s)\,ds.$$

Then

$$\int_0^\infty |P(Z(t) \varepsilon dy) - \Pi(dy)| \leqq \psi_t([0, \infty)) + 2 \sup_A |\psi_t'(A) - \Pi(A)|.$$

But

$$\begin{aligned}\psi_t([0, \infty)) &= \int_0^{t/2} \nu_1(ds)P(W_1 \geqq t - s) + \int_{t/2}^t \nu_1(ds)P(W_1 \geqq t - s)\\ &\leqq \nu_1([0, \infty))P(W_1 \geqq t/2) + \nu_1([t/2, \infty))\\ &= O(t^{-2}).\end{aligned}$$

On the other hand, $\psi_t'(dy) = q_t(y)\,dy$, where

$$q_t(y) = \int_y^{t+y} p(t + y - u)P(W_1 \varepsilon du).$$

To see this, note that

$$\begin{aligned}\int_x^\infty dy \int_y^{t+y} p(t + y - u)P(W_1 \varepsilon du) &= \int_0^t P(W_1 \geqq x + t - s)p(s)\,ds\\ &= \int_x^{x+t} P(W_1 \geqq u)p(x + t - u)\,du\\ &= \int_x^{x+t} \int_u^\infty P(W_1 \varepsilon ds)p(x + t - u)\,du.\end{aligned}$$

By Fubini's theorem the left hand side can be rewritten as

$$\int_x^{x+t} P(W_1 \varepsilon du) \int_x^u p(t + y - u)\,dy + \int_{x+t}^\infty P(W_1 \varepsilon du) \int_{u-t}^u p(t + y - u)\,dy.$$

By Fubini's theorem the right hand side can be written in the same form.

Using the formula for $q_t(y)$, we obtain

$$\begin{aligned}|\psi_t'(A) - \Pi(A)| &\leqq \int_0^\infty |q_t(y) - (EW_1)^{-1}P(W_1 \geqq y)|\,dy \\ &\leqq \int_0^\infty dy \int_y^{t+y} |p(t + y - u) - (EW_1)^{-1}|\,P(W_1\,\varepsilon\,du) \\ &\quad + (EW_1)^{-1}\int_0^\infty P(W_1 \geqq t + y)\,dy \\ &= O(\textstyle\int_{t/2}^\infty uP(W_1\,\varepsilon\,du) + t^{-2} + \int_0^\infty P(W_1 \geqq y + t/2)\,dy) \\ &= O(\textstyle\int_{t/2}^\infty uP(W_1\,\varepsilon\,du) + t^{-2}).\end{aligned}$$

In this last step we have used properties of p and the fact that

$$\begin{aligned}\int_0^\infty P(W_1 \geqq y + t/2)\,dy &= \int_0^\infty dy \int_{y+t/2}^\infty P(W_1\,\varepsilon\,du) \\ &\leqq \int_{t/2}^\infty uP(W_1\,\varepsilon\,du).\end{aligned}$$

LEMMA 2. *Choose c such that* $0 < c < EW_1$ *and set* $\delta = \min(1, m - 2) > 0$. *Then there is a constant* $K > 0$ *such that for* $n \geqq 1$,

$$\int_0^\infty |P_n(x, dy) - \Pi(dy)| \leqq Kn^{-(1+\delta)}, \qquad x \leqq cn.$$

PROOF OF LEMMA 2. Choose a such that $0 < a < EW_1 - c$. For $x \leqq cn$,

$$P(S_n\,\varepsilon\,x - A) \leqq P(S_n \leqq x) \leqq P(S_n \leqq cn) = O(n^{1-m}),$$

the last step following from a result of Nagaev [2], p. 215. Now $P(Z(S_n - x)\,\varepsilon\,A) = \int_x^\infty P(S_n\,\varepsilon\,dt)P(Z(t - x)\,\varepsilon\,A)$ and thus

$$\begin{aligned}&\int_0^\infty |P_n(x, dy) - \Pi(dy)| \\ &\quad\leqq 2\sup_A |P_n(x, A) - \Pi(A)| \\ &\quad\leqq 2\sup_A \textstyle\int_x^\infty P(S_n\,\varepsilon\,dt)|P(Z(t - x)\,\varepsilon\,A) - \Pi(A)| + 4P(S_n \leqq cn) \\ &\quad\leqq 2\textstyle\int_x^\infty P(S_n\,\varepsilon\,dt)\int_0^\infty |P(Z(t - x)\,\varepsilon\,dy) - \Pi(dy)| + 4P(S_n \leqq cn) \\ &\quad= O(P(S_n \leqq (c + a)n) + \textstyle\int_{an/2}^\infty uP(W_1\,\varepsilon\,du) + n^{-2}) \\ &\quad= O(n^{1-m} + n^{-2}) = O(n^{-1-\delta}),\end{aligned}$$

the last step following from Lemma 1 and Nagaev's result.

For an event A in the sample space of X_n, $n \geqq 0$, let $P_\Pi(A)$ denote the probability of A if X_0 has distribution Π, and let $P_x(A)$ denote the probability of A if $X_0 = x$.

LEMMA 3. *If* $n \geqq 1$ *and A is an event depending only on* $\{X_k, k \geqq n\}$, *then*

$$|P_x(A) - P_\Pi(A)| \leqq Kn^{-(1+\delta)}, \qquad x \leqq cn.$$

PROOF OF LEMMA 3. We need only observe that

$$\begin{aligned}|P_x(A) - P_\Pi(A)| &= |\textstyle\int_0^\infty P(A \mid X_n = y)(P_n(x, dy) - \Pi(dy))| \\ &\leqq \textstyle\int_0^\infty |P_n(x, dy) - \Pi(dy)| \leqq Kn^{-(1+\delta)}.\end{aligned}$$

LEMMA 4. *Let* $n \geqq 1$ *and* $0 \leqq k < \infty$. *Let A be an event depending only on* $\{X_j \mid j \leqq k\}$ *and let B be an event depending only on* $\{X_j \mid j \geqq n + k\}$. *Then*

$$|P_{\Pi}(AB) - P_{\Pi}(A)P_{\Pi}(B)| \leqq K'n^{-(1+\delta)},$$

where K' is independent of A, B, n, and k.

PROOF OF LEMMA 4. Let 1_A denote the characteristic function of the set A. Then

$$|P_{\Pi}(AB) - P_{\Pi}(A)P_{\Pi}(B)| \leqq E_{\Pi}[1_A\,|P(B|X_k) - P_{\Pi}(B)|]$$
$$\leqq \Pi((cn, \infty)) + Kn^{-(1+\delta)}.$$

but

$$\Pi((cn, \infty)) = (EW_1)^{-1} \textstyle\int_{cn}^{\infty} P(W_1 \geqq y)\, dy$$
$$\leqq (EW_1)^{-1} \textstyle\int_{cn}^{\infty} uP(W_1 \,\varepsilon\, du),$$

from which Lemma 4 follows immediately.

We can now complete the proof of the theorem. If X_0 has distribution Π, then the theorem follows immediately from Lemma 4 and Theorem 1.6 of Ibragimov [1], p. 365, at least in the case $\sigma_f > 0$. If $\sigma_f = 0$, then a simple computation shows that $\mathrm{Var}_{\Pi}\,(\sum_{k=0}^{n} f(X_k)) = o(n^{-1})$ and hence the conclusion still holds.

In general, let f be a bounded measurable function on $(0, \infty)$ and $-\infty < \theta < \infty$. Choose $\epsilon > 0$. Then

$$|E \exp [i\theta(n+1)^{-\frac{1}{2}} \textstyle\sum_{j=k}^{n} f(X_j)] - E_{\Pi} \exp [i\theta(n+1)^{-\frac{1}{2}} \sum_{j=k}^{n} f(X_j)]|$$
$$\leqq Kk^{-(1+\delta)} + P(X_0 \geqq ck) \leqq \epsilon$$

for k sufficiently large and all $n \geqq k$. This inequality follows from Lemma 3 or, alternatively, by a proof similar to that of Lemma 3. For $n \geqq k$ and n sufficiently large,

$$|E \exp [i\theta(n+1)^{-\frac{1}{2}} \textstyle\sum_{j=0}^{n} f(X_j)] - \mathrm{E} \exp [i\theta(n+1)^{-\frac{1}{2}} \sum_{j=k}^{n} f(X_j)]|$$
$$\leqq E|\exp [i\theta(n+1)^{-\frac{1}{2}} \textstyle\sum_{j=0}^{k-1} f(X_j)] - 1| \leqq \epsilon$$

and, similarly,

$$|E_{\Pi} \exp [i\theta(n+1)^{-\frac{1}{2}} \textstyle\sum_{j=0}^{n} f(X_j)] - E_{\Pi} \exp [i\theta(n+1)^{-\frac{1}{2}} \sum_{j=k}^{n} f(X_j)]| \leqq \epsilon.$$

Combining the above inequalities, we see that

$$\lim_{n\to\infty} E \exp [i\theta(n+1)^{-\frac{1}{2}} \textstyle\sum_{j=0}^{n} f(X_j)] = \lim_{n\to\infty} E_{\Pi} \exp [i\theta(n+1)^{-\frac{1}{2}} \sum_{j=0}^{n} f(X_j)],$$

and the conclusion of the theorem holds, as desired.

REFERENCES

[1] IBRAGIMOV, I. A. (1962). Some limit theorems for stationary processes. *Theor. Prob. Appl.* **7** 349–382.

[2] NAGEV, S. V. (1965). Some limit theorems for large deviations. *Theor. Prob. Appl.* **10** 214–235.

[3] STONE, C. (1966). On absolutely continuous components and renewal theory. *Ann. Math. Statist.* **37** 271–275.

[4] OREY, S. (1959). Recurrent Markov chains. *Pacific J. Math.* **9** 805–827.