The Annals of Statistics
1981, Vol. 9, No. 3, 604–614

ASYMPTOTIC INFERENCE IN LÉVY PROCESSES OF THE DISCONTINUOUS TYPE[1]

BY MICHAEL G. AKRITAS[2] AND RICHARD A. JOHNSON

University of Wisconsin, Madison

We establish contiguity of certain families of probability measures indexed by T, as $T \to \infty$, for classes of stochastic processes with stationary, independent increments whose sample paths are discontinuous. Many important consequences pertaining to properties of tests and estimates then apply. A new expression for the Radon-Nikodym derivative of these processes is obtained.

1. Introduction. We present a general treatment of parametric statistical procedures for Lévy processes. This wide class of stochastic models with stationary independent increments includes both compound Poisson processes and limits of jump processes. An important example of latter type is the process whose increments have gamma distributions.

Compound Poisson process have been treated in the statistical literature (see Lewis, 1972) but the general Lévy process has not been considered. Frost (1972) was first to study signal detection and estimation problems for independent increment processes (see also Segall and Kailath, 1975). Rubin and Tucker (1959) consider nonparametric estimation of quantities appearing in the characteristic function of Lévy processes. Recently, Basawa and Brockwell (1978) considered inference for gamma and stable processes based on jumps of size greater than ε.

In Section 2 we present the Radon-Nikodym derivative and Section 3 contains the major statistical implications. The derivation of the Radon-Nikodym derivative appears in Section 4 and the technical details for the other results in Section 5. Two applications of the theory to compound Poisson processes and a gamma process are given in Section 6.

2. The likelihood and the score functions. Let the process $\{X(t): t \in [0, \infty)\}$ have probability measure P_θ. We consider stationary, independent increment processes having characteristic function $f_t(u) = e^{t\psi_\theta(u)}$ where

$$\psi_\theta(u) = iu\beta(\theta) + \int_{\{0\}^c} \left(e^{iux} - 1 - \frac{iux}{1+x^2}\right) d\mu_\theta(x) \tag{2.1}$$

and for each $\theta \in \Theta$, $\mu_\theta((-\infty, -a] \cup [a, \infty)) < \infty$, for all $a > 0$, and $\int_{-1}^{1} x^2 d\mu_\theta(x) < \infty$. We suppose the process can be observed continuously from $0 \le t \le T$. When it cannot, the continuous case solution gives the limiting case. Let $P_{T,\theta}$ be the restriction of P_θ that pertains to $\{X(t): 0 \le t \le T\}$. The Radon-Nikodym derivative can be expressed in terms of the measure μ_θ in (2.1), and the process

$$X(B, t) = \textstyle\sum_{\tau \le t} \Delta X(\tau) I_B(\Delta X(\tau)) \tag{2.2}$$

Received June, 1978; revised July, 1979.

[1] This research was sponsored by the Naval Office of Research under Grant Number N00014-78C-0722.

[2] Now at M.I.T.

AMS 1970 *subject classifications.* Primary 62M07, 62M09; secondary 62G99.

Key words and phrases. Asymptotic inference, stochastic processes, independent increments, contiguity.

where $\Delta X(\tau) = X(\tau) - X(\tau-)$ and B is any Borel set bounded away from zero. The difficulties with small jumps in $X(t)$ is surmounted by considering a sequence of jump sizes determined from choices for B.

THEOREM 2.1. *Let θ_0, $\theta_1 \in \theta$, and assume that $P_{T,\theta_0} \approx P_{T,\theta_1}$. Then for any sequence $\{D_n\}$ of neighborhoods of the origin such that $D_n \downarrow \{0\}$,*

$$\frac{dP_{T,\theta_1}}{dP_{T,\theta_0}}(X(t)) = \prod_{m=1}^{\infty}\left[e^{-T(\mu_{\theta_1}-\mu_{\theta_0})(B_m)}\prod_{j=1}^{N_{mT}}\frac{d\mu_{\theta_1}}{d\mu_{\theta_0}}(Z_{mj})\right]$$

where $B_1 = D_1^c$, $B_m = D_{m-1} - D_m$, $m \geq 2$, and N_{mt} denotes the total number of jumps Z_{mj} of the process $X(B_m, t)$, $t \in [0, T]$.

After obtaining Theorem 2.1, we received a preprint of Brockett, Hudson and Tucker (1978) where they obtain the same expression for the Radon-Nikodym derivative. Their proof employing Kakutani's Theorem is much longer than ours; see Section 4 and Akritas and Johnson (1978).

Statistical inferences would be based on the likelihood determined by evaluating $dP_{T,\theta}/dP_{T,\theta_0}$ at the sample path $X(t)$, $0 \leq t \leq T$. We are concerned with optimal statistical procedures for the limiting case $T \to \infty$. The log-likelihood will be shown to be approximated by a score function whose terms are analogous to $\frac{\partial}{\partial\theta} \ln f_\theta(x) = 2\dot{\phi}$, say, in the case of independent observations having pdf $f_\theta(x)$. In our development, we employ quadratic mean derivatives, $\dot{\phi}$, and the measure μ_θ in (2.1) plays the primary role. For any θ_0, $\theta_1 \in \theta$ define

$$\phi(x; \theta_0, \theta_1) = \left(\frac{d\mu_{\theta_1}}{d\mu_{\theta_0}}(x)\right)^{1/2}. \tag{2.3}$$

Occasionally, when there is no danger of confusion, we will use the simpler notation $\phi(\theta_0, \theta_1)$ instead of $\phi(x; \theta_0, \theta_1)$. In the above notation we now formulate our assumptions.

(A1) The probability measures $\{P_{T,\theta}; \theta \in \Theta\}$ are mutually absolutely continuous for all $T \geq 0$.

(A2) For each $\theta \in \Theta$, there exists a k-dimensional random vector $\dot{\phi}(\theta)$ such that

(i) $\dot{\phi}(\theta)$ is $\mathscr{B} \times \mathscr{C}$-measurable, where $\mathscr{C}$ is the σ-field of Borel subsets of Θ.

(ii) $\int_{\{0\}^c} (h'\dot{\phi}(\theta))^2\, d\mu_\theta$ is finite for all $h \in \mathbb{R}^k$.

(iii) $\int_{\{0\}^c} \frac{1}{\|h\|^2}[\phi(\theta, \theta+h) - 1 - h'\dot{\phi}(\theta)]^2\, d\mu_\theta \to 0$, as $h \to 0$

where $\|\cdot\|$ denotes the usual Euclidean norm.

(iv) $\int_{\{0\}^c} \frac{1}{\|h\|^2}\,|(\phi(\theta, \theta+h) - 1)^2 - (h'\dot{\phi}(\theta))^2|\, d\mu_\theta \to 0$, as $h \to 0$.

(v) $\frac{1}{\|h\|^2}\int_{\{0\}^c} (\phi(\theta, \theta+h) - 1)^2\, d\mu_\theta \to \int_{\{0\}^c} (h'\dot{\phi}(\theta))^2\, d\mu_\theta$, as $h \to 0$.

Let $\Gamma(\theta)$ be the covariance function

$$\Gamma(\theta) = 4\int_{\{0\}^c} [\dot{\phi}(\theta)\dot{\phi}'(\theta)]\, d\mu_\theta. \tag{2.4}$$

(A3) $\dot{\Gamma}(\theta)$ is positive definite for all $\theta \in \Theta$.

REMARK 2.1. These assumptions, concerning the smoothness of $\phi(\theta, \theta + h)$, are similar in spirit and form to the usual assumptions required to obtain contiguity; see Roussas (1972) for the appropriate choice of ϕ in the discrete time Markov process case. Here, however, the measure μ_θ may be infinite and it is not clear that assumptions A2(iv), (v) follow from A2 (iii). The assumptions here are found to be satisfied in many examples including those of Section 6.

3. The main results. Let the process be observed from zero to a positive time T. We develop the asymptotic theory, $T \to \infty$, based on the log-likelihood ratios

$$
\begin{aligned}
\Lambda_T(\theta_0) &= \log \frac{dP_{T,\theta_T}}{dP_{T,\theta_0}} (X(t), t \in [0, T]) \\
&= \sum_{m=1}^{\infty} \left[T \int_{B_m} (1 - \phi^2(x; \theta_0, \theta_T))\, d\mu_{\theta_0}(x) + 2 \sum_{j=1}^{N_{mT}} \log \phi(Z_{mj}; \theta_0, \theta_T) \right]
\end{aligned}
\tag{3.1}
$$

where

$$\theta_T = \theta_0 + T^{-1/2} h_T, \; h_T \to h, \quad \text{as} \quad T \to \infty, h_T, h \in \mathbb{R}^k \tag{3.2}$$

and ϕ is defined in (2.3). In section 5, we show that the likelihood is approximated by the random vector

$$\Delta_T(\theta_0) = 2T^{-1/2} \sum_{m=1}^{\infty} \left[\sum_{j=1}^{N_{mT}} \dot{\phi}(Z_{mj}, \theta_0) - T \int_{B_m} \dot{\phi}(x; \theta_0)\, d\mu_{\theta_0}(x) \right] \tag{3.3}$$

where B_m, N_{mT} are defined in Theorem 2.1 and $\dot{\phi}(x; \theta_0)$ in assumption (A2). In particular

$$\Lambda_T(\theta_0) - h'\Delta_T(\theta_0) \to_{T\to\infty} - \tfrac{1}{2} h'\Gamma(\theta_0)h, \quad \text{in } P_{T,\theta_0} - \text{probability}. \tag{3.4}$$

Theorem 5.4 below gives the limiting distribution for the log likelihood ratios under both θ_0 and θ_T.

$$\mathscr{L}[h'\Delta_T(\theta_0) \mid P_{T,\theta_0}] \Rightarrow N(0, h'\Gamma(\theta_0)h) \tag{3.5}$$

and Theorem 5.4 gives the limiting distribution for the log likelihood ratios under θ_0 or under θ_T.

$$\mathscr{L}[\Lambda_T(\theta_0) \mid P_{T,\theta_T}] \Rightarrow N(\tfrac{1}{2} h'\Gamma(\theta_0)h, h'\Gamma(\theta_0)h). \tag{3.6}$$

We then conclude directly that the families $\{P_{T,\theta}\}$ and $\{P_{T,\theta_T}\}$ are contiguous. More importantly, several important conclusions pertaining to the properties of estimators and tests follow directly from general contiguity results.

Under conditions of the form (3.4) – (3.5), Hájek (1970) established that for any family V_T of estimators that belong in the class $\{\{V_T\};\ L[T^{1/2}(V_T - \theta_T) \mid P_{\theta_T}] \Rightarrow L_V(\theta_0)$, a probability measure, where $\theta_T = \theta_0 + hT^{-1/2}$, we must have $L_V(\theta_0) = L_1(\theta_0) * L_2(\theta_0)$ where $L_1(\theta_0) = N(0, \Gamma^{-1}(\theta_0))$ and $L_2(\theta_0)$ is a measure in R^k. Using the above representation, Hájek's results, stated in continuous time, become

(a) $\lim \sup_T P_{T,\theta_0}[T^{1/2}(V_T - \theta_0) \in C] \le \int_C d\Phi_\Gamma$ for all convex symmetric sets C in R^k, Φ_Γ is the cdf of $N(0, \Gamma^{-1}(\theta_0))$

(b) $\lim \inf \mathscr{E}_{\theta_0}[T^{1/2}h'(V_T - \theta_0)]^2 \ge h'\Gamma^{-1}(\theta_0)h$, all $h \in R^k$. If the limiting covariance D exists, $D - \Gamma^{-1}(\theta_0)$ is nonnegative definitive.

Some important testing hypotheses conclusions, for a one-dimensional parameter, are that the test which rejects H_0: $\theta = \theta_0$ in favor of H_1: $\theta > \theta_0$ if $\Delta_T(\theta_0) > c_T$, where c_T is an appropriately chosen constant, is asymptotically uniformly most powerful. The two sided test which rejects H_0: $\theta = \theta_0$ in favor of H_1: $\theta \neq \theta_0$ if $|\Delta_T(\theta_0)| > b_T$, for appropriately chosen b_T, is asymptotically uniformly most powerful unbiased. The conclusions regarding tests of multidimensional parameters, in turn, use the admissibility of the class of tests based on Δ_T^* which reject outside of convex sets.

The conclusions regarding asymptotically optimal tests of hypotheses follow from the development in Johnson and Roussas (1969, 1970, 1971), since the relevant proofs there use neither the Markovian character of the observations, nor the discreteness of time. To summarize; it may be shown from standard arguments (see Section 4 of Johnson and Roussas, 1970) that a truncated version Δ_T^* may be constructed in such a way the $P_{\theta_0}[\Delta_T^* \neq \Delta_T] \to 0$, as $T \to \infty$. Based on Δ_T^*, an exponential family $dR_{T,h} = e^{-B_T(h)} e^{h\Delta_T^*} dP_{T,\theta_0}$ is constructed, which approximates P_{T,θ_T}. As in Theorem 5.1 of Johnson and Roussas (1970), it may be shown that $\sup\{\| R_{T,h} - P_{T,\theta_T} \|, h \in \text{bounded set}\} \to 0$, as $T \to \infty$, where $\| \cdot \|$ is the total variation. That is Δ_T^* is *asymptotically sufficient.* Also, it may be shown that $\mathscr{E}_{\theta_0}[Z_T \mid \Delta_T]$ has the same local power, asymptotically, as any test function Z_T.

Two examples are given in Section 6 following the technical details in the next two sections.

4. Derivation of the Radon-Nikodym derivative. The primary aim of the present section is to present a proof of Theorem 2.1. For $\theta_1, \theta_0 \in \Theta$ and $0 < T < \infty$, we derive the Radon-Nikodym derivative of P_{T,θ_1} with respect to P_{T,θ_0}, under the assumption that they are mutually absolutely continuous ($\approx$). Here $P_{T,\theta}$ denotes the restriction of P_θ on $\sigma\{X(t), t \in [0, T]\}$. In particular, let Ω be the space $D([0, \infty))$ of all real valued, right continuous functions $X(t)$, $t \in [0, \infty)$, that have finite left hand limits, and let $\mathscr{A}$ be the σ-field of cylinder sets in Ω. For each $\theta \in \Theta \subset R^k$, open, let P_θ be a probability measure on $(\Omega, \mathscr{A})$ and assume that, under P_θ, the coordinate process $\{X(t), t \in [0, \infty)\}$ has stationary independent increments and characteristic function $\exp(t\psi_\theta(u))$ with $\psi_\theta(u)$ given by (2.1). It follows that the process $X(t)$ is continuous in probability (see Breiman, 1968, page 304) and $X(0) = 0$ a.s. $[P_\theta]$. The function $\psi_\theta(u)$ is called the *exponent function* and μ_θ the *Lévy measure* of the process. If μ_θ is finite, $X(t)$ is called a *jump process*, or a *compound Poisson process* and its sample paths are step functions with probability one. If μ_θ is infinite, the process is called a *limit of jump processes* and its sample paths have, with probability one, infinitely many jumps over any finite interval of time. Necessary and sufficient conditions for $P_{T,\theta_1} \approx P_{T,\theta_0}$ are given in Theorem 4.1. Our derivation of the Radon-Nikodym derivative is based on a method inspired by Striebel (1959). By a rather simple observation we obtain Proposition 4.1 that shows that the log-likelihood has an infinitely divisible distribution and specifies its Lévy measure. We first state

THEOREM 4.1. *For each $\theta \in \Theta$, let $P_{T,\theta}$ be a probability measure on $D([0, T])$ such that the coordinate process has exponent function given by* (2.1). *Then $P_{T,\theta_1} \approx P_{T,\theta_0}$ if and only if*

(L1) $\mu_{\theta_1} \approx \mu_{\theta_0}$

(L2) $\int_{[0]^c} (1 - (d\mu_{\theta_1}/d\mu_{\theta_0})^{1/2})^2 \, d\mu_{\theta_0} < \infty$

(L3) $\beta(\theta_1) - \beta(\theta_0) - \int_{[0]^c} \frac{x}{1+x^2} \, d(\mu_{\theta_1} - \mu_{\theta_0})(x) = 0$

REMARK 2.1 Newman (1973) showed that condition (L2) implies

$$\int_{[0]^c} \frac{|x|}{1+x^2} \, d|\mu_{\theta_1} - \mu_{\theta_0}|(x) < \infty,$$

so that, the integral appearing in (L3) is well defined. The proof of Theorem 2.1 composed of three steps which we state as lemmas.

LEMMA 4.1. *Let $P_{n,T,\theta}$ be the measure on $\mathscr{B}_{n,T} = \sigma\{X(D_n^c, t), 0 \le t \le T\}$ induced by P_θ. Then the sequence $\left\{\left(\frac{dP_{n,T,\theta_1}}{dP_{n,T,\theta_0}}, \mathscr{B}_{n,T}\right), n \ge 1\right\}$ is a martingale in $(\Omega, \mathscr{A}, P_{\theta_0})$.*

PROOF. Let $f_n = dP_{n,T,\theta_1}/dP_{n,T,\theta_0}$ and assume for the time being that $m > n$ implies $\mathscr{B}_{n,T} \subseteq \mathscr{B}_{m,T}$. Then, for $\mathscr{A}_n \in \mathscr{B}_{n,T}$, $\int_{A_n} f_n \, dP_{n,T,\theta_0} = \int_{\mathscr{A}_n} f_m \, dP_{m,T,\theta_0}$ and since f_n is $\mathscr{B}_{n,T}$-measurable, it follows that $\mathscr{E}_\theta(f_m \mid \mathscr{B}_{n,T}) = f_n$, which proves the lemma. It remains to show that $\mathscr{B}_{n,T} \subseteq \mathscr{B}_{m,T}$. By the right continuity of $X(D_n^c, t)$, we may define the rv's $Z_k^{(n)}, \rho_k^{(n)}$, $k \geq 1$, on $(\Omega, \mathscr{A})$ such that, if $0 \leq t < \rho_1^{(n)}$, $X(D_n^c, t) = Z_1^{(n)}$, and if $\rho_1^{(n)} + \cdots + \rho_{k-1}^{(n)} \leq t < \rho_1^{(n)} + \cdots + \rho_k^{(n)}$, $X(D_n^c, t) = Z_k^{(n)}$, $k \geq 2$. It follows that the sequence $\{(Z_k^{(n)}, \rho_k^{(n)}),\ k \geq 1\}$ determines, for each $\omega \in \Omega$, the sample path induced by $\{X(D_n^c, t),\ t \in [0, T]\}$, and conversely. It is easy to see that for $m > n$, $\sigma\{(Z_k^{(n)}, \rho_k^{(n)}),\ k \geq 1\} \subseteq \sigma\{(Z_k^{(m)}, \rho_k^{(m)}),\ k \geq 1\}$ so that, by the equivalence of $\{(Z_k^{(n)}, \rho_k^{(n)}),\ k \geq 1\}$ and $\{X(D_n^c, t), t \in [0, T]\}$, $\mathscr{B}_{n,T} \subseteq \mathscr{B}_{m,T}$. □

As a consequence of Lemma 4.1 we have that $\lim_{n\to\infty}(dP_{n,T,\theta_1}/dP_{n,T,\theta_0})$ exists a.s. $[P_{\theta_0}]$. However, it must still be verified that this limit equals $dP_{T,\theta_1}/dP_{T,\theta_0}$. From Skorohod (1965, page 100) it follows that $\lim_{n\to\infty} (dP_{n,T,\theta_1}/dP_{n,T,\theta_0})$ will equal $dP_{T,\theta_1}/dP_{T,\theta_0}$ provided that $X(D_n^c, t) \to X(t)$ in P_{T,θ_0}-probability, for all $t \in [0, T]$. But from Breiman (1968, page 313) it follows that $[X(D_n^c, t) + c_n t] \to X(t)$ in P_{T,θ_0}-probability, for all $t \in [0, T]$, where $\{c_n\}$ is a sequence of constants. Thus we need

LEMMA 4.2. *Let c be any constant and $\tilde{P}_{n,T,\theta}$ be the probability measure on $\tilde{\mathscr{B}}_{n,T} = \sigma\{X(D_n^c, t) + ct,\ 0 \leq t \leq T\}$ induced by P_θ. Then*

$$\frac{d\tilde{P}_{n,t,\theta_1}}{d\tilde{P}_{n,T,\theta_0}}(X(D_n^c, t) + ct) = \frac{dP_{n,T,\theta_1}}{dP_{n,T,\theta_0}}(X(D_n^c, t)).$$

PROOF. For $A \in \tilde{\mathscr{B}}_{n,T}$,

$$\begin{aligned}
&\int_{A - \{ct, t\in[0,T]\}} \frac{d\tilde{P}_{n,T,\theta_1}}{d\tilde{P}_{n,T,\theta_0}}(X(D_n^c, t) + ct)\, dP_{T,\theta_0}(X(D_n^c, t)) \\
&= \tilde{P}_{n,T,\theta_1}(\{X(D_n^c, t) + ct,\ t \in [0, T]\} \in A) \\
&= P_{n,T,\theta_1}(\{X(D_n^c, t),\ t \in [0, T]\} \in A - \{ct,\ t \in [0, T]\}) \\
&= \int_{A - \{ct, t\in[0,T]\}} \frac{dP_{n,T,\theta_1}}{dP_{n,T,\theta_0}}(X(D_n^c, t),\ t \in [0, T])\, dP_{T,\theta_0}(X(D_n^c, t)).
\end{aligned}$$

Since A was arbitrary, the proof of the lemma is complete. □

LEMMA 4.3. *Let $P_{n,T,\theta}$ be as defined in Lemma* 4.1. *Then,*

$$\frac{dP_{n,T,\theta_1}}{dP_{n,T,\theta_0}}(X(D_n^c, t)) = \exp[-T(\mu_{\theta_1} - \mu_{\theta_0})(D_n^c)] \prod_{j=1}^{\tilde{N}_{nT}} \frac{d\mu_{\theta_1}}{d\mu_{\theta_0}}(Z_j)$$

where Z_j, $j = 1, \ldots, \tilde{N}_{nt}$ are the jumps of the process $X(D_n^c, t)$, $t \in [0, T]$.

PROOF. This follows from the fact that $\{X(D_n^c, t),\ 0 \leq t \leq T\}$ is a jump process (see Breiman, page 312) and the well-known expression for the Radon-Nikodym derivative of measures corresponding to jump processes (Skorokhod, 1957). □

The proof of Theorem 2.1 now follows by a simple synthesis of the results of the above lemmas.

Before closing this section we present Proposition 4.1 which enables us to show that the logarithm of the Radon Nikodym derivative has an infinitely divisible distribution. This proposition, which generalizes Proposition 4.25 of Breiman (1968), is used repeatedly in our derivations in Section 5.

PROPOSITION 4.1. *Let $\{X(t),\ t \geq 0\}$ be a process with stationary, independent increments defined on $(\Omega, \mathscr{A}, P)$ and let μ denote its Lévy measure. Let g be a Borel measurable function such that $\int_{\{0\}^c} |g(x)|\, d\mu(x) < \infty$. Define*

$$Y(t) = \textstyle\sum_{\tau \leq t} g(\Delta X(\tau)) I_{\{0\}^c}(\Delta X(\tau)) = \sum_j^{(t)} g(Z_j)$$

where, the index $\sum_j^{(t)}$ denotes the summation over all jumps of the process $X(\tau)$, $\tau \in [0, t]$. Then, $Y(t)$ exists and is finite. Moreover $\{Y(t),\ t \geq 0\}$ is a process with stationary, independent increments and Lévy measure $\mu \circ g^{-1}$.

PROOF. Let $\{D_n\}$ be a sequence of neighborhoods of the origin such that $D_n \downarrow \{0\}$ and let

$$Y_n(t) = \textstyle\sum_{\tau \leq t} [g(\Delta X(\tau))\, I_{D_n^c}(\Delta X(\tau)) = \sum_{\tau \leq t} g_n(\Delta X(\tau)).$$

It is easy to see that $Y_n(t)$ has an infinitely divisible distribution. Moreover, $\{Y_n(\tau),\ \tau \in [0, t]\}$, is a jump process whose number of jumps obey a Poisson process with intensity $\mu(D_n^c)$. Then, in order to find its Lévy measure, it suffices to find the distribution of the size of its jumps. The measure that corresponds to this distribution is easily seen to be $G_n(A) = \mu \circ g_n^{-1}(A)/\mu(D_n^c)$, for any Borel set $A \subseteq \{0\}^c$. Therefore, the Lévy measure of $Y_n(t)$ is $\mu(D_n^c) \cdot G_n = \mu \circ g_n^{-1}$. Next, $Y(t)$ is infinitely divisible since it is the almost sure limit of the $Y_n(t)$ which have this property. Further the Lévy measure of $Y(t)$ is given by $\lim_n \mu \circ g_n^{-1}(A) = \mu \circ g^{-1}(A)$, for any Borel set A bounded away from zero. Here, the interchange of limit and integration is guaranteed by the assumption $\int_{\{0\}^c} |g(x)\ d\mu(x) < \infty$. Finally, $Y(t)$ is finite a.s. $[P]$ by the assumption just mentioned and Breiman (1968, page 313). □

COROLLARY 4.1 *Let $Y(T)$ be the logarithm of the Radon-Nikodym derivative of Theorem 2.1. Then $Y(T)$ has an infinitely divisible distribution with exponent function*

$$iuK(\theta_1, \theta_0) + \int_{\{0\}^c} \left(e^{iux} - 1 - \frac{iux}{1 + x^2}\right) d\mu_{\theta_0} \circ g^{-1}(x)$$

where $g(x) = \log \dfrac{d\mu_{\theta_1}}{d\mu_{\theta_0}}(x)$ *and* $K(\theta_1, \theta_0) = \sum_{m=1}^{\infty} [\int_{B_m} \frac{x}{1 + x^2}\, d\mu_{\mu_0} \circ g^{-1}(x) - (\mu_{\theta_1} - \mu_{\theta_0})(B_m)]$ *with B_m defined in Theorem 2.1.*

See Akritas and Johnson (1978) for a proof.

5. The contiguity results. We now develop the proofs regarding contiguity through a series of lemmas. Let the log-likelihood Λ_T be defined as in (3.1). In the remainder of the paper, and when there is no danger of confusion, we use the notation $\Lambda_T = \Lambda_T(\theta_0)$, $\phi_{Tmj} = \phi_{Tmj}(\theta_0) = \phi(Z_{mj};\ \theta_0, \theta_T)$, $\phi_{Tj} = \phi_{Tj}(\theta_0) = \phi(Z_j;\ \theta_0, \theta_T)$, $\dot{\phi}_{mj} = \dot{\phi}_{mj}(\theta_0) = \dot{\phi}(Z_{mj};\ \theta_0)$, $\dot{\phi}_j = \dot{\phi}_j(\theta_0) = \dot{\phi}(Z_j;\ \theta_0)$, where Z_{mj}, $m \geq 1, j = 1, \ldots, N_{mT}$, denotes jumps belonging to B_m and the alternate notation Z_j, $j \geq 1$, denotes the complete sequence of jumps of $\{X(t),\ t \in [0, T]\}$ without regard to their size groupings.

LEMMA 5.1. $\max_{m,j} |\phi_{Tmj}(\theta_0) - 1| \to_{T \to \infty} 0$, *in P_{T,θ_0} – probability.*

PROOF. Set $\phi_{Tmj} - 1 = T^{-1/2} h' \dot{\phi}_{mj} + T^{-1/2} R_{Tmj}$, so that, by assumption (A2) (iii), $\int_{\{0\}^c} R_T^2\, d\mu_{\theta_0} \to 0$, as $T \to \infty$. Next,

$$P_{\theta_0}(\max_{m,j} |\phi_{Tmj} - 1| > \varepsilon)$$

$$\leq P_{\theta_0}\left(\max_{m,j} |h' \dot{\phi}_{mj}| > \frac{\varepsilon T^{1/2}}{2}\right) + P_{\theta_0}\left(\max_{m,j} |R_{Tmj}| > \frac{\varepsilon T^{1/2}}{2}\right). \tag{5.1}$$

Define a set A_{1T} by $A_{1T} = \left[\, | h'\dot{\phi}(\theta_0) | > \frac{\varepsilon T^{1/2}}{2} \right]$. Then,

$$(5.2) \qquad P_{\theta_0}\left(\max_{m,j} | h'\dot{\phi}_{mj} | > \frac{\varepsilon T^{1/2}}{2}\right) = P_{\theta_0}(\omega \in D([0, T]) = 1 - \exp[-T\mu_{\theta_0}(A_{1T})]:$$

(there is at least one jump whose size belongs to A_{1T})

But since, $A_{1T} \downarrow$ to a μ_{θ_0} − null set, as $T \to \infty$, assumption (A2) (ii) implies

$$(5.3) \qquad T\mu_{\theta_0}(A_{1T}) = \frac{4}{\varepsilon^2} \int_{A_{1T}} \frac{\varepsilon^2}{4} T \, d\mu_{\theta_0} \le \frac{4}{\varepsilon^2} \int_{A_{1T}} | h'\dot{\phi} |^2 \, d\mu_{\theta_0} \to_{T\to\infty} 0.$$

Similarly, let $A_{2T} = \left[\, | R_T | > \frac{\varepsilon T^{1/2}}{2} \right]$, so that

$$(5.4) \qquad P_{T,\theta_0}\left(\max_{m,j} | R_{Tj} | > \frac{\varepsilon T^{1/2}}{2}\right) = 1 - \exp[-T\mu_{\theta_0}(A_{2T})]$$

and

$$(5.5) \qquad T\mu_{\theta_0}(A_{2T}) \le \frac{4}{\varepsilon^2} \int_{A_{2T}} \frac{\varepsilon^2}{4} T \, d\mu_{\theta_o} \le \frac{4}{\varepsilon^2} \int_{A_{2T}} | R_T |^2 \, d\mu_{\theta_0} \to_{T\to\infty} 0.$$

The proof of the lemma then follows from relations (5.1)–(5.5). □

Next, consider the expansion $\log x = (x - 1) - \frac{1}{2}(x - 1)^2 + c(x - 1)^3$, $| c | \le 3$, which is valid for $| x - 1 | \le \frac{1}{2}$. To replace x by ϕ_{Tmj}, we use Lemma 5.1 to conclude that the set $A_{3T} = [\max_{m,j} | \phi_{Tmj} - 1 | > \varepsilon]$ satisfies $P_{T,\theta_0}(A_{3T}^c) > 1 - \varepsilon$, for all sufficiently large T. Thus, on A_{3T}^c, we may rewrite the log-likelihood of relation (3.1) as

$$(5.6) \qquad \begin{aligned} \Lambda_T(\theta_0) = \textstyle\sum_{m=1}^{\infty} & \left[T \int_{B_m} (1 - \phi_T^2(\theta_0)) \, d\mu_{\theta_0} + 2\textstyle\sum_{j=1}^{N_{mT}} (\phi_{Tmj} - 1) \right] \\ & - \textstyle\sum_{m=1}^{\infty} \sum_{j=1}^{N_{mT}} (\phi_{Tmj} - 1)^2 + 2 \sum_{m=1}^{\infty} \sum_{j=1}^{N_{mT}} c_{Tmj}(\phi_{Tmj} - 1)^3. \end{aligned}$$

The fact that the above sums are well defined and finite follows from Proposition 4.1, Lemma 5.1 and Theorem 2.1.

Let $\sum_j^{(T)} g(Z_j)$ denote the summation $\sum_{m=1}^{\infty} \sum_{j=1}^{N_{mT}} g(Z_{mj})$. Then,

LEMMA 5.2. $\sum_j^{(T)} c_{Tj}(\phi_{Tj}(\theta_0) - 1)^3 \to_{T\to\infty} 0$, *in* P_{T,θ_0} − *probability.*

PROOF. This follows from $\sum_j^{(T)} | c_{Tj}(\phi_{Tj} - 1)^3 | \le 3 \max_{m,j} | \phi_{Tmj} - 1 | \sum_j^{(T)} (\phi_{Tj} - 1)^2$ and Lemma 5.3. □

LEMMA 5.3. $\sum_j^{(T)} (\phi_{Tj}(\theta_0) - 1)^2 \to_{T\to\infty} \int_{\{0\}^c} (h'\dot{\phi}(\theta_0))^2 \, d\mu_{\theta_0}$, *in* P_{T,θ_0} − *probability.*

PROOF. Set $Y_1(T) = \sum_j^{(T)} (h'\dot{\phi}_j)^2$, so that from Proposition 4.1, assumption A2(ii) and well known properties of Lévy processes,

$$(5.7) \qquad \frac{1}{T} Y_1(T) \to_{T\to\infty} \int_{\{0\}^c} (h' \dot{\phi})^2 \, d\mu_{\theta_0} \text{ in } P_{T,\theta_0} - \text{probability.}$$

Thus it suffices to show that $| \sum_j^{(T)} (\phi_{Tj} - 1)^2 - \frac{1}{T} Y_1(T) | \to_{T\to\infty} 0$, in P_{T,θ_0} − probability. Indeed it is easy to see that

$$P_{T,\theta_0}\left[\,|\textstyle\sum_j^{(T)} (\phi_{Tj} - 1)^2 - \frac{1}{T} Y_1(T)\,| > \varepsilon\right) \le \frac{T \int_{\{0\}^c} |\,T(\phi_T - 1)^2 - (h'\dot{\phi})^2\,|\, d\mu_{\theta_0}}{T\varepsilon}$$

and the expression on the right-hand side tends to zero, as $T \to \infty$, by assumption (A2) (iv). □

LEMMA 5.4. *As* $T \to \infty$,

$$\textstyle\sum_{m=1}^{\infty} \left[T \int_{B_m} (1 - \phi_T^2(\theta_0))\, d\mu_{\theta_0} + 2T \int_{B_m} (\phi_T(\phi_0) - 1)\, d\mu_{\theta_0}\right] \to -\int_{\{0\}^c} (h'\dot{\phi}(\theta_0))^2\, d\mu_{\theta_0}.$$

PROOF. The conclusions follows, from the identity $(1 - \phi_T^2) + 2(\phi_T - 1) = -(\phi_T - 1)^2$ and assumption A2(v).

LEMMA 5.5.

$$2 \textstyle\sum_{m=1}^{\infty} \left[\sum_{j=1}^{N_{mT}} (\phi_{Tmj}(\theta_0) - 1) - T \int_{B_m} (\phi_T(\theta_0) - 1)\, d\mu_{\theta_0}\right]$$
$$- 2T^{-1/2} \textstyle\sum_{m=1}^{\infty} \left[\sum_{j=1}^{N_{mT}} h'\dot{\phi}_{mj}(\theta_0) - T \int_{B_m} h'\dot{\phi}(\theta_0)\, d\mu_{\theta_0}\right] \to 0$$

as $T \to \infty$, *in* P_{T,θ_0} *– probability.*

PROOF. We first note that $\sum_{m=1}^{\infty} [\sum_{j=1}^{N_{mT}} h'\dot{\phi}_{mj}(\theta_0) - T \int_{B_m} h'\dot{\phi}(\theta_0)\, d\mu_{\theta_0}]$ is a well defined random variable having an infinitely divisible distribution with Lévy measure $\mu_{\theta_0} \circ [h'\dot{\phi}]^{-1}$ (see Akritas and Johnson, 1978 for details). Next, by Chebyshev's inequality,

$$P_{T,\theta_0}\left\{\left|2 \textstyle\sum_{m=1}^{\infty} \left[\sum_{j=1}^{N_{mT}} (\phi_{Tmj} - 1 - T^{-1/2} h'\dot{\phi}_{mj})\right.\right.\right.$$
$$\left.\left.\left. - T \int_{B_m} (\phi_T - 1)\, d\mu_{\theta_0} + T^{1/2} \int_{B_m} h'\dot{\phi}\, d\mu_{\theta_0}\right]\right| > \varepsilon\right\}$$
$$\le \frac{4}{\varepsilon^2} \int_{\{0\}^c} [T^{1/2}(\phi_{Tj} - 1) - h'\dot{\phi}]^2\, d\mu_{\theta_0} \to 0,$$

by assumption A2(iii). □

We are now ready to establish the main results.

THEOREM 5.1. *In the above notation*

$$\Lambda_T(\theta_0) - h'\Delta_T(\theta_0) \to_{T\to\infty} -\tfrac{1}{2}\, h'\Gamma(\theta_0)h, \text{ in } P_{T,\theta_0} - \text{probability.}$$

PROOF. From relation (5.6) and Lemmas 5.2, 5.3, 5.4 it easily follows that

$$\Lambda_T(\theta_0)$$
$$- 2 \textstyle\sum_{m=1}^{\infty} \left[\sum_{j=1}^{N_{mT}} (\phi_{Tmj}(\theta_0) - 1) - T \int_{B_m} (\phi_T(\theta_0) - 1)\, d\mu_{\theta_0}\right]$$
$$\to_{T\to\infty} - 2 \int_{\{0\}^c} (h'\dot{\phi}(\theta_0))^2\, d\mu_{\theta_0}.$$

The proof of the theorem then follows by adding the result of Lemma 5.5 to the above relation and by recalling the definition of $\Gamma(\theta_0)$. □

THEOREM 5.2. *For all $h \in R^k$,*

$$\mathscr{L}[h'\Delta_T(\theta_0) \mid P_{T,\theta_0}] \Rightarrow N(0, h'\Gamma(\theta_0)h).$$

PROOF. Since $\mathscr{E}_{\theta_0} \{\sum_{m=1}^n [\sum_{j=1}^{N_{mT}} h'\dot{\phi}_{mj} - T \int_{B_m} h'\dot{\phi}\, d\mu_{\theta_0}]\}^2, \leq \int_{\{0\}^c} (h'\dot{\phi})^2\, d\mu_{\theta_0} < \infty$ by assumption A2(ii), it follows that the random variables $\{\sum_{m=1}^n [\sum_{j=1}^{N_{mT}} h'\dot{\phi}_{mj} - T \int_{B_m} h'\dot{\phi}\, d\mu_{\theta_0}]\}^r$, $n \geq 1$ are uniformly integrable for all $r < 2$ (see Loève, 1963, page 184) so that

$$\mathscr{E}_{\theta_0}(h'\Delta_T) = \lim_n \mathscr{E}_{\theta_0}\left\{2T^{-1/2} \sum_{m=1}^n \left[\sum_{j=1}^{N_{mT}} h'\dot{\phi}_{mj} - T \int_{B_m} h'\dot{\phi} d\mu_{\theta_0}\right]\right\} = 0.$$

To complete the proof of the theorem, note that for T large and some $t_0 > 0$, we may set $h'\Delta_T = \sum_{j=1}^n Y_j + R_n$, where $n = [T/t_0]$ and, for $T_j = jt_0$, $Y_j = h'\Delta_{T_j} - h'\Delta_{T_{j-1}}$, and apply Slutsky's Theorem together with the classical Central Limit Theorem. □

As a consequence of Theorems 5.1 and 5.2 one has

THEOREM 5.3. $\mathscr{L}[\Lambda_T(\theta_0) \mid P_{T,\theta_0}] \Rightarrow N(-\tfrac{1}{2} h'\Gamma(\theta_0)h, h'\Gamma(\theta_0)h)$.

Theorem 5.3 and the fact that $\int \exp(x) d\mathscr{L}(x) = 1$, where $\mathscr{L} = N(-\tfrac{1}{2}\sigma^2, \sigma^2)$, imply that the families $\{P_{T,\theta}\}$ and $\{P_{T,\theta_T}\}$ are contiguous (see Proposition 3.1, page 11, in Roussas, 1972).

THEOREM 5.4. *Let θ_T, $\Lambda_T(\theta_0)$ and $\Delta_T(\theta_0)$ be defined by* (3.2), (3.1) *and* (3.3) *respectively. Then*

(*i*) $\Lambda_T(\theta_0) - h'\Delta_T(\theta_0) \to -\tfrac{1}{2} h'\Gamma(\theta_0)h$, *in* P_{T,θ_T} *– probability,*

(ii) $\mathscr{L}[\Delta_T(\theta_0) \mid P_{T,\theta_T}] \Rightarrow N(\Gamma(\theta_0)h, \Gamma(\theta_0))$,

(iii) $\mathscr{L}[\Lambda_T(\theta_0) \mid P_{T,\theta_T}] \Rightarrow N(\tfrac{1}{2} h'\Gamma(\theta_0)h, h'\Gamma(\theta_0)h)$.

PROOF. Part (i) follows from Theorem 5.1 and the definition of contiguity (see Roussas, 1972, page 7). Part (ii) follows from Theorems 5.1, 5.2 and Theorem 7.2, page 38 in Roussas (1972). Part (iii) follows from Theorem 5.3 and Corollaries 7.1, 7.2 pages 34, 35 in Roussas (1972). □

6. Some examples. In this section, we illustrate the applicability of our assumptions by two examples.

EXAMPLE 6.1. *Compound Poisson processes.* For compound Poisson processes, the Lévy measure is of the form $\lambda \cdot F$ where, λ is the intensity and F denotes both the cdf of the jumps and the corresponding measure. The cdf F may be taken to belong to almost any of the standard parametric families. Condition (L1) of Theorem 4.1 and assumptions (A2), (A3) have been checked for several distributions in Roussas (1972). Here, of course, the intensity λ may be an additional parameter. Further, conditions (L2) and (L3) are always true for jump processes with exponent function given by (2.1) so that assumption (A1) is also satisfied. Consequently our conclusions apply, quite generally, to compound Poisson processes.

EXAMPLE 6.2. *The Gamma process.* The process $X(t)$ is said to be a gamma process if its characteristic function is $f_t(u) = (1 - iu/\theta)^{-t}$, $\theta > 0$. The exponent function is given by (2.1) with (see Feller (1971), page 567) $\beta(\theta) = \int_{(0,\infty)} [e^{-\theta x}/(1 + x^2)]\, dx$ and $\mu_\theta(A) = \int_{A \cap (0,\infty)}$

$[e^{-\theta x}/x]\,dx$, $A \in \mathscr{B}$. Conditions (L1), (L2) and (L3) are easily seen to be satisfied so that assumption (A1) is true. Next it is easily seen that $\phi(z;\theta_0\,\theta_1) = \exp[-(\theta_1-\theta_0)z/2]$, so that, $\dot\phi(z;\theta_0) = -z/2$. Assumptions (A2) (i), (ii) are clearly satisfied. Further, $\Gamma(\theta) = \frac{1}{4}\int_{(0,\infty)} ze^{-\theta z}dz = \frac{1}{4}\frac{1}{\theta^2}$ so that (A3) holds. To show (A2) (iii) note that

$$\int_{\{0\}^c}\frac{1}{h^2}[\phi(\theta,\theta+h)-1-h\dot\phi(\theta)]^2\,d\mu_\theta$$

$$=\int_{(0,1)}+\int_1^\infty \frac{1}{h^2}[\phi(\theta,\theta+h)-1-h\dot\phi(\theta)]^2\,d\mu_\theta \le h\int_{(0,1)} ze^{-\theta z}\,dz + I_1(h).$$

Then, $h\int_{(0,1)} ze^{-\theta z}dz \to 0$ as $h\to 0$, and $I_1(h)$ can be shown to converge to zero by Vitali's Theorem. To show (A2) (iv) we again split the integral $\int_{\{0\}^c}\frac{1}{h^2}\,|(\phi(\theta,\theta+h)-1)^2-(h'\dot\phi(\theta))^2|\,d\mu_\theta = \int_{(0,1)}+\int_1^\infty = I_2(h)+I_3(h)$. But $I_2(h)\to 0$, as $h\to 0$ by the dominated convergence theorem since $|(e^{-hz/2}-1)/h| \le 2z$ and $I_3(h)\to 0$ by Vitali's Theorem and the fact that $I_1(h)\to 0$. To verify (A2) (v), we first note that by $I_1(h)\to 0$ and Vitali's Theorem, $\frac{1}{h^2}\int_1^\infty(\phi(\theta,\theta+h)-1)^2\,d\mu_\theta \to \int_1^\infty(\dot\phi(\theta))^2\,d\mu_\theta$, as $h\to 0$. Next, $\frac{1}{h^2}\int_{(0,1)}(\phi(\theta,\theta+h)-1)^2\,d\mu_\theta \to \int_{(0,1)}(\dot\phi(\theta))^2\,d\mu_\theta$, as $h\to 0$, by the dominated convergence theorem. Finally, by $\int_0^\infty|\dot\phi(z;\theta_0)|\,d\mu_{\theta_0}(z) = \frac{1}{2\theta_0} < \infty$, and Proposition 4.1, it follows that the random vector $\Delta_T(\theta_0)$, defined in (3.3), takes the form $\Delta_T(\theta_0) = -T^{-1/2}\left(X(T) - T\frac{1}{\theta_0}\right)$. Hence as it is pointed out in Section 3, the test which rejects H_0: $\theta=\theta_0$ in favor of H_1: $\theta>\theta_0$ if $\Delta_T(\theta_0) > c_T$, is asymptotically uniformly most powerful and the test which rejects H_0: $\theta=\theta_0$ in favor of H_1: $\theta\ne\theta_0$ if $\Delta_T(\theta_0) < a_T$ or $\Delta_T(\theta_0) > a_T$, is asymptotically uniformly most powerful unbiased, where c_T and a_T are appropriately chosen constants. These results provide a rigorous justification for the rather curious limits in Basawa and Brockwell (1978).

REFERENCES

AKRITAS, M. G. and JOHNSON, R. A. (1978). Asymptotic Inference in Lévy Processes of the Discontinuous Type. Tech. Rep. No 532, Univ. of Wisconsin, Madison.

BASAWA, I. V. and BROCKWELL, P. J. (1978). Inference for gamma and stable processes. *Biometrika* **65** 129–133.

BREIMAN, L. (1968). *Probability*. Addison-Wesley, Reading, Massachusetts.

BROCKETT, P. L., HUDSON, W. N. and TUCKER, H. G. (1978). The distribution of the likelihood ratio for additive processes. *J. Multivariate Anal.* **8** 233–243.

CHUNG, K. L. (1974). *A Course in Probability Theory*. Academic, New York.

FELLER, W. (1971). *An Introduction to Probability Theory and its Applications II*. Wiley, New York.

FROST, P. (1972). Optimal estimation and detection for conditionally independent increment processes. *IEEE Internat. Symp. Information Theory*.

JOHNSON, R. A. and ROUSSAS, G. G. (1969). Asymptotically most powerful tests in Markov processes. *Ann. Math. Statist.* **40** 1207–1215.

JOHNSON, R. A. and ROUSSAS, G. G. (1970). Asymptotically optimal tests in Markov processes. *Ann. Math. Statist.* **41** 918–938.

JOHNSON, R. A. and ROUSSAS, G. G. (1971). Applications of contiguity to multiparameter hypothesis testing. *Proc. 6th Berkeley Symp. of Math. Statist. and Prob.* **1** 195–226.

LEWIS, P. A. W. (1972). *Stochastic Point Processes: Statistical Analysis, Theory and Applications*. Wiley-Interscience, New York.

LOÉVE, M. (1963). *Probability Theory*. Van Nostrand, Princeton.

NEWMAN, C. M. (1973). On the orthogonality of independent increment processes. In *Topics in Probability Theory* (D. W. Stroock and S. R. S. Varadhan, eds.). Courant Institute of Math. Science, N.Y.

RÉNYI, A. (1970). *Foundations of Probability*. Holden-Day, San Francisco.

ROUSSAS, G. G. (1972). *Contiguity of Probability Measures: Some Applications in Statistics.* Cambridge Univ. Press.

RUBIN, H. and TUCKER, H. G. (1959). Estimating the Parameters of a Differential Process. *Ann. Math. Statist.* **30** 641–658.

SEGALL, A. and KAILATH, T. (1975). Radon-Nikodym derivatives with respect to measures induced by discontinuous independent increment processes. *Ann. Probability* **3** 449–464.

SKOROKHOD, A. V. (1957). On differentiability of measures which correspond to stochastic processes I. Processes with independent increments. *Theor. Probability Appl.* **2** 407–432.

SKOROKHOD, A. V. (1965). *Studies in the Theory of Random Processes.* Addison-Wesley, Reading, Massachusetts.

STRIEBEL, C. T. (1959). Densities of stochastic processes. *Ann. Math. Statist.* **30** 559–567.

DEPARTMENT OF MATHEMATICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS 02139

DEPARTMENT OF STATISTICS
UNIVERSITY OF WISCONSIN-MADISON
1210 WEST DAYTON STREET
MADISON, WISCONSIN 53706