Adv. Appl. Prob. **19,** 512–514 (1987)
Printed in N. Ireland

THE BERRY–ESSEEN BOUND FOR THE POISSON SHOT-NOISE

JOHN A. LANE,* *University College of Wales, Aberystwyth*

Abstract

This note provides a useful extension of the Berry–Esseen bound on the error in the normal approximation for shot-noise. The special cases treated are of particular interest in the statistical analysis of Poisson processes and cluster point processes.

POISSON PROCESS; CLUSTER POINT PROCESSES

The Berry–Esseen theorem provides a uniform bound on $|F_Z - \Phi(x)|$ where Φ is the standard normal cumulative distribution function (c.d.f.) and F_Z is the c.d.f. of Z, a random variable suitably normalised and centred. When Z is Poisson shot-noise such bounds have been considered by Papoulis (1971), Lane (1979) and Heinrich and Schmidt (1985). With minimal extra effort the latter result can usefully be extended by removing the requirements that the primary Poisson process be stationary and the secondary random functions be translation-invariant.

Suppose A is a metric space with its Borel σ-algebra, $\mathscr{B}$. Let $N(\cdot)$ denote the counting measure of a Poisson point process on $(A, \mathscr{B})$ with σ-finite intensity measure $\Lambda(\cdot) = \mathbb{E}N(\cdot)$. The points of the Poisson process may be enumerated as τ_k, $k = 1, 2, \cdots$. Suppose also that $\{H(s): s \in A\}$ is a family of independent real random variables which is independent of $N(\cdot)$. The Poisson shot-noise discussed here is the sum

$$(1) \qquad Y = \sum_{i=1}^{N(A)} H(\tau_i) = \int_A H(s)N(ds).$$

The characteristic function of Y takes the well-known form

$$\mathbb{C}_Y(\theta) = \exp\left\{\int [\mathbb{C}_H(\theta; s) - 1]\Lambda(ds)\right\}$$

where $\mathbb{C}_H(\theta; s) = \mathbb{E} \exp [i\theta H(s)]$. From this it follows that the mean and variance of Y are respectively $\mu_Y = \int_A \mathbb{E}H(s)\Lambda(ds)$ and $\sigma_Y^2 = \int_A \mathbb{E}H^2(s)\Lambda(ds)$. We assume throughout that $\rho_Y = \int_A \mathbb{E} |H(s)|^3 \Lambda(ds) < \infty$. Let $Z = (Y - \mu_Y)/\sigma_Y$ have c.d.f. F_Z and characteristic function $\mathbb{C}_Z(\theta)$. From the expansion of $\exp \{i\theta H(s)/\sigma_Y\}$, we see that

$$|\log \mathbb{C}_Z(\theta) + \tfrac{1}{2}\theta^2| \leqq |\theta|^3 \rho_Y/(6\sigma_Y^3).$$

The argument leading to Theorem 7 of Heinrich and Schmidt (1985) now yields that

$$|F_Z(x) - \Phi(x)| < 2{\cdot}21\rho_Y/\sigma_Y^3.$$

Frequently in practice Λ (or H) will be taken to be 0 outside some subspace $Q \subset A$.

Received 24 October 1986.

* Postal address: Department of Mathematics, The University College of Wales, Aberystwyth, SY23 3BZ, UK.

Suppose that as the expected number of points $\Lambda(Q)\rightarrow\infty$, we have $\rho_Y=O(\Lambda(Q))$ and $\Lambda(Q)/\sigma_Y^2=O(1)$. Then $|F_Z(x)-\Phi(x)|=O(\Lambda(Q)^{-\frac{1}{2}})$. Note however that, as in Lane (1984), a non-normal limit can arise even when Var Y and higher moments are finite.

Example 1. In a cluster point process the primary points τ_k are cluster centres, each triggering off a secondary point process or cluster. Let $N_s(Q)$ be the number of cluster points in Q, $(Q\in\mathscr{B})$, triggered off by a primary point at s. Then $H(s)=N_s(Q)$ and Y denotes the total number of cluster points in Q.

The following construction defines a cluster point process of Neyman–Scott type (Neyman and Scott (1958), Vere-Jones (1970)). Suppose $N_s(A)$, the total cluster size triggered from s, has probability generating function $P(\cdot)$ for all s and let m_j, $j=1,2,3$ denote its first three factorial moments. Given $N_s(A)$, the cluster points are laid down independently with conditional probability $G_Q(s)$ that a cluster point lies in Q. It is easily verified that

$$\mathbb{C}_H(\theta;s)=P[G_Q(s)(\exp(i\theta)-1)+1]$$

and that

$$\mu_Y=I_1(Q),\qquad \sigma_Y^2=I_1(Q)+I_2(Q)$$
$$\rho_Y=I_1(Q)+3I_2(Q)+I_3(Q)$$

where $I_j(Q)=m_j\int_A G_Q(s)^j\Lambda(ds)$. Now $\rho_Y/\sigma_Y^2\leqq(m_1+3m_2+m_3)/(m_1+m_2)$. In special cases, the rate of normal convergence $\sigma_Y^{-\frac{1}{2}}$, will reduce to $\Lambda(Q)^{-\frac{1}{2}}$. Thus, if both $\int_{A\setminus Q}G_Q(s)\Lambda(ds)$ and $\int_Q G_{A\setminus Q}(s)\Lambda(ds)$ are $o(\Lambda(Q))$ then $\sigma_Y^2\sim\Lambda(Q)$. These conditions reflect a tendency for clusters triggered off from within Q to be concentrated in Q and to form the dominant component of Y.

Example 2. Statistics of the form (1) occur extremely often in statistical problems concerning Poisson processes observed over a subspace $Q\subset A=\mathbb{R}^d$; see for example Krickeberg (1982). Typically $H(\cdot)$ is a deterministic function which is 0 outside Q.

In the 'natural' linear model, Λ depends on real parameters γ through

$$\Lambda(\cdot;\gamma)=\int\lambda(s;\gamma)\,ds,\qquad \log\lambda(s;\gamma)=\gamma^T\mathbf{g}(s)$$

where $\mathbf{g}(s)$ is a real vector function on A and we take ds to denote Lebesgue measure (or some appropriate dominating measure) on $(A,\mathscr{B})$. The log likelihood is

$$\int_Q\log\lambda(s;\gamma)N(ds)-\Lambda(Q;\gamma).$$

Its first derivative, or score, vector $\mathbf{U}$ is of interest for, among other things, the development of locally most powerful tests. Now

$$\mathbf{U}=\int_Q\mathbf{g}(s)N(ds)-\int_Q\mathbf{g}(s)\lambda(s;\gamma)\,ds$$

which of course has mean 0. Denote the covariance matrix of U by

$$\mathbf{\Sigma}=\mathbf{\Sigma}(Q)=\int_Q\mathbf{g}(s)\mathbf{g}(s)^T\lambda(s;\gamma)\,ds,$$

which we assume is positive definite with smallest eigenvalue $m(Q)>0$ and let $h(Q)=\sup_{s\in Q}\|\mathbf{g}(s)\|$.

Consider the distribution of $W = \mathbf{l}^T \mathbf{\Sigma}^{-\frac{1}{2}} \mathbf{U}$ where, without loss of generality, $\mathbf{l}^T \mathbf{l} = 1$. Now

$$\sigma_w^2 = \int_Q (\mathbf{l}^T \mathbf{\Sigma}(Q)^{-\frac{1}{2}} \mathbf{g}(s))^2 \lambda(s; \gamma)\, ds = 1$$

so that

$$|F_w(x) - \Phi(x)| \leqq 2{\cdot}21 \rho_w$$

where

$$\begin{aligned} \rho_w &= \int_Q |\mathbf{l}^T \mathbf{\Sigma}(Q)^{-\frac{1}{2}} \mathbf{g}(s)|^3 \lambda(s; \gamma)\, ds \\ &\leqq \|\Sigma(Q)^{-\frac{1}{2}}\|\, h(Q) \sigma_w^2 \\ &= m(Q)^{-\frac{1}{2}} h(Q). \end{aligned}$$

Let $\{Q_t\}$ be a sequence of sets in $\mathscr{B}$ for which $\Lambda(Q_t) \to \infty$ as $t \to \infty$. Under the condition

(2) $$m(Q_t) \sim k h(Q_t)^2 \Lambda(Q_t)$$

as $t \to \infty$, for some $k > 0$, we have

$$|F_w(x) - \Phi(x)| = O(\Lambda(Q_t)^{-\frac{1}{2}}).$$

Since $\mathbf{l}$ is arbitrary, $\mathbf{U}$ is asymptotically multivariate normal $\mathrm{N}(\mathbf{0}, \mathbf{\Sigma}(Q))$. The condition (2) is more transparent in the case where $|g|$ is an increasing scalar function on $\mathbb{R}$ and $Q_t = [0, t]$. For an extensive range of plausible models

$$m([0, t]) = \int_0^t g(s)^2 \lambda(s)\, ds \sim k g(t)^2 \Lambda(t).$$

On the other hand, for $\Lambda(\mathrm{s}) = \log[1 + \log(1 + s)]$, (2) fails as

$$\int_0^t |\log \lambda(s)|^r \lambda(s)\, ds \sim \exp(r\Lambda(t))/r \qquad (r = 1, 2, \ldots).$$

Here $\rho_w / \sigma_w^3 \sim 2^{3/2}/3$ but nonetheless asymptotic normality can be proved via Theorem 3 of Lane (1984).

References

Heinrich, L. and Schmidt, V. (1985) Normal convergence of multidimensional shot-noise and rates of this convergence. *Adv. Appl. Prob.* **17,** 709–730.

Krickeberg, K. (1982) Processus Ponctuels en Statistique. *Lecture Notes in Mathematics* **929,** 206–313. Springer-Verlag, Berlin.

Lane, J. A. (1979) Some Limit Theorems for Shot-noise Processes. Unpublished Ph.D. Thesis, University of London.

Lane, J. A. (1984) The central limit theorem for the Poisson shot-noise process. *J. Appl. Prob.* **21,** 287–301.

Neyman, J. and Scott, E. L. (1958) Statistical approach to the problems of cosmology. *J.R. Statist. Soc.* B **20,** 1–43.

Papoulis, A. (1971) High density shot-noise and gaussianity. *J. Appl. Prob.* **8,** 118–127.

Vere-Jones, D. (1970) Stochastic models for earthquake occurrence. *J.R. Statist. Soc.* B **32,** 1–62.