Ann. Inst. Statist. Math.
Vol. 55, No. 1, 83–94 (2003)

EDGEWORTH EXPANSIONS FOR COMPOUND POISSON PROCESSES AND THE BOOTSTRAP

GUTTI JOGESH BABU[1]*, KESAR SINGH[2] AND YANING YANG[2]

[1] *Department of Statistics, The Pennsylvania State University, 326 Thomas Building, University Park, PA 16802-2111, U.S.A.*
[2] *Department of Statistics, Faculty of Arts and Sciences, Rutgers University, Hill Center Busch Campus, 110 Frelinghuysen Road, Piscataway, NJ 08903, U.S.A.*

(Received December 26, 2000; revised May 20, 2002)

Abstract. One-term Edgeworth Expansions for the studentized version of compound Poisson processes are developed. For a suitably defined bootstrap in this context, the so called one-term Edgeworth correction by bootstrap is also established. The results are applicable for constructing second-order correct confidence intervals (which make correction for skewness) for the parameter "mean reward per unit time".

Key words and phrases: Renewal reward processes, Poisson process, studentization, confidence interval, approximate cumulant, non-lattice distribution, one-term Edgeworth correction by bootstrap.

1. Introduction

The main objective of this investigation is to obtain one-term Edgeworth expansion and establish second-order correctness of a suitably defined bootstrap for the studentized compound Poisson process. We begin with the definition of a renewal reward process.

Let $\{N(t), t > 0\}$, be a renewal process with the inter-arrival times $T_1, T_2, \dots$. Here T_i, $i = 1, 2, \dots$ are positive i.i.d. r.v.'s. Thus $N(t)$ equals the number of arrivals up to time t. Suppose $X_1, X_2, \dots$ are i.i.d. random variables independent of $\{N(t), t > 0\}$. The renewal reward process is defined as $\sum_{i=0}^{N(t)} X_i$, where $X_0 = 0$. The classical example of a renewal reward process arises in a business setting where customers arrive according to a renewal process and X_i denotes the revenue directly due to the i-th customer. The total revenue up to time t gives rise to a renewal reward process.

Hipp (1985) obtained Edgeworth expansions for $(\sum_{i=0}^{N(t)} X_i - \lambda t\mu)/\sqrt{\lambda t}$, where $\mu = E(X_1)$ and $E(T_1) = \lambda^{-1} > 0$. In fact, he considered a more general case, where X_i are random vectors and X_i and T_i are allowed to be dependent. His results are derived under a Cramér's type condition for the distribution of (X_1, T_1). In the univariate case, one-term Edgeworth expansions for the standardized mean can be obtained under a weaker assumption that X_i has a non-lattice distribution.

Edgeworth expansions for pivotal quantities such as the studentized mean are difficult to derive for general renewal reward processes. We shall investigate the general case in a later paper. In this article we concentrate on the compound Poisson process

*Research work of Gutti Jogesh Babu was supported in part by NSF grants DMS-9626189 and DMS-0101360.

by making use of the relation between the variance and the mean of T_1. This relation facilitates studentization by a simple expression involving only $X_1, \ldots, X_{N(t)}$, when $\{N(t), t > 0\}$ is a Poisson process. It can be easily seen that $\sum_{i=0}^{N(t)} X_i^2$ is a natural consistent estimator of the variance of $(\sum_{i=0}^{N(t)} X_i - \lambda\mu t)$. One might think of deducing an Edgeworth expansion for the studentized

$$R_t = \left(\sum_{i=0}^{N(t)} X_i - \lambda\mu t\right)\left(\sum_{i=0}^{N(t)} X_i^2\right)^{-1/2}, \tag{1.1}$$

by expressing it as a smooth function of $\sum_{i=0}^{N(t)}(X_i, X_i^2)$, $\lambda\mu t$ and applying (the relevant) part (i) of Theorem 2.7 of Hipp (1985). To treat such functions, expansions for $P(\sum_{i=0}^{N(t)}((X_i, X_i^2) - E(X_i, X_i^2)) \in C)$ are needed uniformly for a class of non-convex sets C defined through inequalities involving a linear term plus a quadratic form. It is a non-trivial task to extend Hipp's result on convex sets to such a wider class of sets. We shall not pursue this line of investigation further here.

There are several related papers (Mykland (1992), (1993), (1995*a*, *b*), Yoshida (2001) and Kusuoka and Yoshida (2000)) on asymptotic expansions using nontrivial martingale approach and Malliavin calculus. However, our result is derived from the standard Edgeworth expansions in the i.i.d. case using simple and elementary estimates.

In Section 2, we shall develop an Edgeworth expansion for R_t. In Section 3, a result on second order correctness (similar to Singh (1981)) for a suitably defined bootstrap procedure for R_t is established. Some simulation studies are presented in Section 4 for standardized and studentized compound Poisson processes. The technical details, required in the proof of the main result, are presented in the Appendix.

We conclude this section by noting that the results obtained here can readily be applied to form confidence bounds for $\lambda\mu t$ or $\lambda\mu$. This is especially true for the second order bootstrap result on the studentized case. For a survey of the problems on estimation of the intensity function of a Poisson process see the monograph by Kutoyants (1998). Note that $\lambda\mu$ can be interpreted as 'mean reward' per unit time. Let $\hat{q}_\alpha(t)$ and $\hat{q}_{1-\alpha}(t)$ be the bootstrap estimates of the α-th and $(1-\alpha)$-th quantile for the statistic R_t. The interval

$$\left[\frac{1}{t}\sum_{i=0}^{N(t)} X_i - \hat{q}_{1-\alpha}(t)\frac{1}{t}\left(\sum_{i=0}^{N(t)} X_i^2\right)^{1/2},\ \frac{1}{t}\sum_{i=0}^{N(t)} X_i - \hat{q}_{\alpha}(t)\frac{1}{t}\left(\sum_{i=0}^{N(t)} X_i^2\right)^{1/2}\right]$$

is an asymptotically second order correct $100(1-2\alpha)$-level confidence interval for $\lambda\mu$. The proof of second order correctness can be carried out along the lines of Babu and Bose (1988), Bose and Babu (1991), or Hall (1988).

2. Edgeworth expansions

To establish a one-term Edgeworth expansion for the studentized statistic R_t, recall that for the Poisson process $\{N(t), t > 0\}$ with rate $\lambda > 0$, the inter-arrival times T_i have the exponential distribution with

$$E(T_1) = \lambda^{-1} > 0, \quad \mathrm{var}(T_1) = \lambda^{-2}, \quad \text{and} \quad \nu_3 = E(T_1 - \lambda^{-1})^3 = 2\lambda^{-3}.$$

To state the main theorem, let

$$\mu = E(X_1), \quad \sigma^2 = \mathrm{var}(X_1), \quad \mu_3 = E(X_1 - \mu)^3 \quad \text{and} \quad \nu^2 = E(X_1^2) = \sigma^2 + \mu^2.$$

Unless otherwise stated, throughout this paper the limits are taken as $t \to \infty$.

THEOREM 2.1. *Suppose $E(X_1)^6 < \infty$ and the distribution of X_1 has a continuous component. Let $\{N(t), t > 0\}$ denote a Poisson process with rate $\lambda > 0$, and be independent of the sequence $\{X_n\}$. Then, uniformly in x, as $t \to \infty$,*

$$(2.1) \quad P\left(\frac{\sum_{i=0}^{N(t)} X_i - \lambda\mu t}{\sqrt{(\sum_{i=0}^{N(t)} X_i^2)}} \leq x\right) = \Phi(x) + \frac{1}{6\nu^3\sqrt{\lambda t}}(\mu_3(2x^2+1) - \mu^3(x^2-1) + 3\mu(\sigma^2 + x^2\sigma^2 + x^2\nu^2))\phi(x) + o(t^{-1/2}).$$

PROOF. Let $H_t = \{|N(t) - \lambda t| \leq (\lambda t/2)\}$,

$$(2.2) \quad W_t = (N(t) - \lambda t)/\sqrt{\lambda t}, \quad A_t = I_{H_t}, \quad C_t = H_t \cap \{|W_t| \leq \log t\}, \quad B_t = I_{C_t},$$

$$(2.3) \quad Z_t = Z_t(x) = \frac{1}{\sigma}(\nu x - \mu W_t), \quad \text{and} \quad V_t = V_t(x) = \frac{1}{\sigma}(\nu x - \mu W_t\sqrt{\lambda t/N(t)}).$$

If $\mu = 0$, then the result follows trivially from

$$(2.4) \quad \begin{aligned} P(H_t^c) &\leq \frac{4}{\lambda^2 t^2}E(N(t) - \lambda t)^2 = O(t^{-1}), \\ E(A_t(\sqrt{\lambda t/N(t)} - 1)) &= E(A_t((1 + (W_t/\sqrt{\lambda t}))^{-1/2} - 1)) \\ &= O(E(|W_t|)/\sqrt{t}) = O(t^{-1/2}), \end{aligned}$$

and Lemma 4 by taking $m_n = 0$.

From now on, without the loss of generality, we assume $\mu > 0$. We use Lemma 4 with $m_n = (n - \lambda t)/\sqrt{n}$, when $N(t) = n$. As $P(|W_t| > \log t) = o(t^{-1/2})$ by Lemma 3, it is easy to check using (2.4) and $P(|\overline{X_n^2} - \nu| > \nu/2) = O(n^{-1})$, that $P(|R_t| > 5\nu\sigma^{-1}\log t) = o(t^{-1/2})$. Hence it is enough to prove that (2.1) holds uniformly in $|x| \leq 5\nu\sigma^{-1}\log t$. Let

$$(2.5) \quad \psi(x, V_t) = 3x\theta\sigma\nu^{-1}V_t - \mu_3(V_t^2 - 1).$$

By Lemma 4, we have on C_t uniformly in $|x| \leq 5\nu\sigma^{-1}\log t$, that

$$(2.6) \quad P(R_t \leq x \mid N(t)) = \Phi(V_t) + \frac{1}{6\sigma^3\sqrt{\lambda t}}\psi(x, V_t)\phi(V_t)\sqrt{\lambda t/N(t)} + o(t^{-1/2}).$$

Since on C_t, $\psi(x, V_t)\phi(V_t) = O(|x|)$ and $\sqrt{\lambda t/N(t)} = 1 + O((\log t)t^{-1/2})$, we have for $|x| \leq 5\nu\sigma^{-1}\log t$ and on C_t,

$$(2.7) \quad \psi(x, V_t)\phi(V_t)(1 - \sqrt{\lambda t/N(t)}) = O((\log t)^2 t^{-1/2}).$$

Similarly on C_t,

$$(2.8) \quad V_t - Z_t = (\mu/\sigma)W_t(1 - \sqrt{\lambda t/N(t)})$$

$$
\begin{aligned}
&= (\mu/\sigma)W_t\left(\frac{W_t}{2\sqrt{\lambda t}} + O\left(\frac{1}{t}W_t^2\right)\right) \\
&= \frac{\mu}{2\sigma\sqrt{\lambda t}}W_t^2 + O((\log t)^3 t^{-1}) \\
&= \frac{\sigma}{2\mu\sqrt{\lambda t}}(Z_t - (\nu x/\sigma))^2 + O((\log t)^3 t^{-1}),
\end{aligned}
$$

and

$$
(V_t - Z_t)^2 = O((\log t)^4 t^{-1}). \tag{2.9}
$$

By (2.8) and (2.9), there exists a ξ_t between V_t and Z_t such that on C_t,

$$
\begin{aligned}
\Phi(V_t) &= \Phi(Z_t) + (V_t - Z_t)\phi(Z_t) + O((V_t - Z_t)^2 \xi_t \phi(\xi_t)) \\
&= \Phi(Z_t) + \frac{\sigma}{2\mu\sqrt{\lambda t}}(Z_t - (\nu x/\sigma))^2\phi(Z_t) + O((\log t)^4 t^{-1}),
\end{aligned} \tag{2.10}
$$

uniformly in $|x| \le 5\nu\sigma^{-1}\log t$.

For any polynomial Q, we have on C_t,

$$
\begin{aligned}
|Q(x, V_t)\phi(V_t) - Q(x, Z_t)\phi(Z_t)| &= |V_t - Z_t| \cdot |q_x(\xi_t) - \xi_t Q(x, \xi_t)|\phi(\xi_t) \\
&= O((\log t)^r t^{-1/2})
\end{aligned} \tag{2.11}
$$

uniformly in x, for some positive integer r depending on the degree of the polynomial Q. Here $q_x(y) = \frac{\partial Q}{\partial y}(x, y)$ and ξ_t is a number between V_t and Z_t. Thus from (2.6)–(2.11), we have, uniformly in $|x| \le 5\nu\sigma^{-1}\log t$,

$$
\begin{aligned}
&E\Bigg(B_t\Bigg|P(R_t \le x \mid Z_t) - \Phi(Z_t) \\
&\quad - \frac{1}{6\sqrt{\lambda t}}\left(\frac{3\sigma}{\mu}\left(Z_t - \frac{\nu x}{\sigma}\right)^2 + \frac{1}{\sigma^3}\psi(x, Z_t)\right)\phi(Z_t)\Bigg|\Bigg) = o(t^{-1/2}).
\end{aligned} \tag{2.12}
$$

We have by Lemmas 1–3, uniformly in $|x| \le 5\nu\sigma^{-1}\log t$,

$$
\begin{aligned}
&E\left(\left(\frac{3\sigma}{\mu}\left(Z_t - \frac{\nu x}{\sigma}\right)^2 + \frac{1}{\sigma^3}\psi(x, Z_t)\right)\phi(Z_t)B_t\right) \\
&= E\Bigg(\Bigg(\left(\frac{3\sigma}{\mu} - \frac{\mu_3}{\sigma^3}\right)(Z_t^2 - 1) \\
&\qquad + \left(\frac{3\theta}{\nu\sigma^2} - \frac{6\nu}{\mu}\right)xZ_t + \frac{3\sigma}{\mu}\left(1 + \frac{\nu x}{\sigma}\right)^2\Bigg)\phi(Z_t)B_t\Bigg) \\
&= \Bigg(\left(\frac{3\sigma}{\mu} - \frac{\mu_3}{\sigma^3}\right)\frac{\sigma^3}{\nu^3}(x^2 - 1) \\
&\qquad + \left(\frac{3\theta}{\nu\sigma^2} - \frac{6\nu}{\mu}\right)x^2\frac{\sigma^2}{\nu^2} + \frac{3\sigma^2}{\mu\nu}\left(1 + \frac{\nu^2 x^2}{\sigma^2}\right)\Bigg)\phi(x) + o(1) \\
&= \left(\frac{\mu_3}{\nu^3}(2x^2 + 1) + \frac{3\mu\sigma^2}{\nu^3} + \frac{3\mu x^2}{\nu^3}(\nu^2 + \sigma^2)\right)\phi(x) + o(1).
\end{aligned} \tag{2.13}
$$

Finally to estimate $E(\Phi(Z_t))$, let $Z \sim N(0, 1)$, U and the process $\{N(t), t > 0\}$ be independent, where U is uniformly distributed on $(-1/2, 1/2)$. Then $W = (x\nu - \sigma Z)\mu^{-1}$

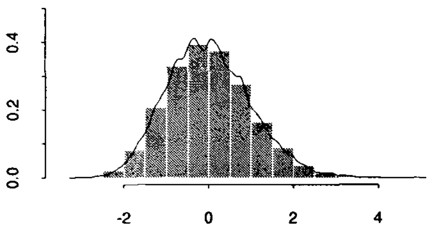

Fig. 1. Exponential, normalized, true distribution.

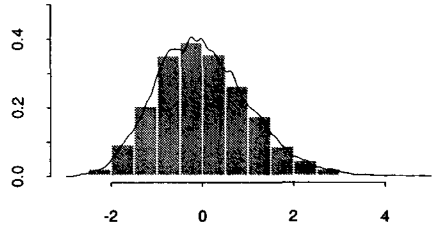

Fig. 2. Exponential, normalized, bootstrap distribution.

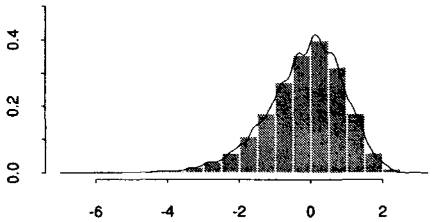

Fig. 3. Exponential, studentized, true distribution.

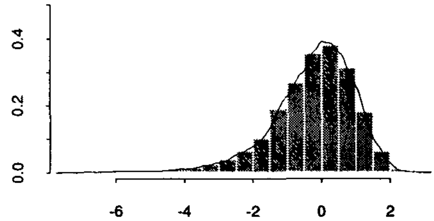

Fig. 4. Exponential, studentized, bootstrap distribution.

has the normal distribution with mean $x\nu/\mu$ and variance $\sigma^2\mu^{-2}$. Since $E(U) = 0$, we have by Lemma 3 (see (A.12)),

$$\begin{aligned} E(\Phi(Z_t)) &= E(\Phi(Z_t - U(\mu/\sigma\sqrt{\lambda t}))) + o(t^{-1/2}) \\ &= P(Z \le \nu\sigma^{-1}x - \mu\sigma^{-1}W_t - (U\mu/\sigma\sqrt{\lambda t})) + o(t^{-1/2}) \\ &= E(P(W_t + (U/\sqrt{\lambda t}) \le W \mid W)) + o(t^{-1/2}) \\ &= E(\Phi(W)) - \frac{1}{6\sqrt{\lambda t}}E((W^2-1)\phi(W)) + o(t^{-1/2}). \end{aligned}$$

Lemma 3 is essentially used, for the expansion, only here. By Lemma 1, $E(\Phi(W)) = \Phi(x)$ and $E((W^2-1)\phi(W)) = (\mu/\nu)^3(x^2-1)\phi(x)$. Therefore

$$E(\Phi(Z_t)) = \Phi(x) - \frac{\mu^3}{6\nu^3\sqrt{\lambda t}}(x^2-1)\phi(x) + o(t^{-1/2}) \tag{2.14}$$

uniformly in x. Theorem 2.1 now follows from (2.12)–(2.14).

3. Bootstrapping

To describe the bootstrap procedure for R_t, let $T_1, T_2, \ldots, T_{N(t)}$ and $X_1, X_2, \ldots, X_{N(t)}$ be the observed data. Let

$$\hat\mu = \frac{1}{N(t)}\sum_{i=0}^{N(t)} X_i \quad \text{and} \quad \frac{1}{\hat\lambda} = \frac{1}{N(t)}\sum_{i=0}^{N(t)} T_i.$$

Let $T_1^*, \ldots, T_{m^*}^*$ be i.i.d. exponential random variables with mean $\hat\lambda^{-1}$ satisfying

$$\sum_{i\le m^*-1} T_i^* \le t < \sum_{i\le m^*} T_i^*.$$

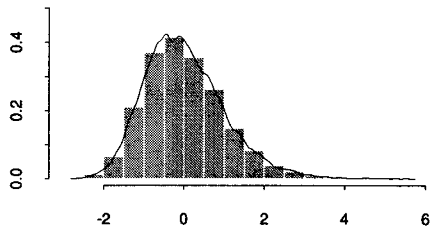

Fig. 5. Lognormal, normalized, true distribution.

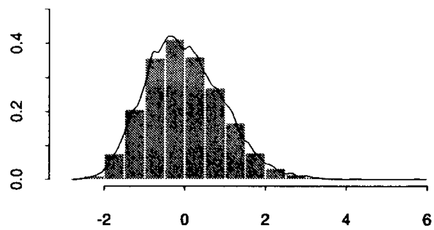

Fig. 6. Lognormal, normalized, bootstrap distribution.

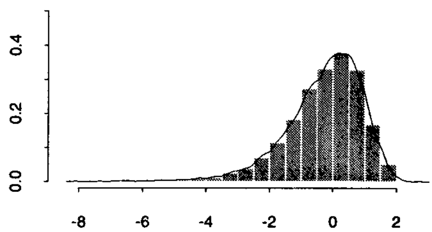

Fig. 7. Lognormal, studentized, true distribution.

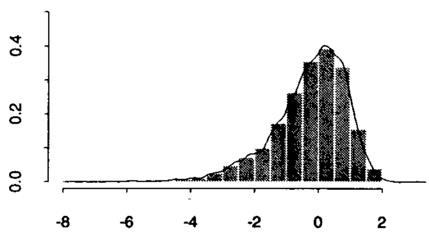

Fig. 8. Lognormal, studentized, bootstrap distribution.

Here $N^*(t) = m^* - 1$. Thus $N^*(t)$ is a Poisson random variable with mean $\hat{\lambda}t$. Also, note that $\hat{\nu}_3 = E^*(T_1^* - \hat{\lambda}^{-1})^3 = 2\hat{\lambda}^{-3}$. Furthermore, let $X_1^*, X_2^*, \dots, X_{N^*(t)}^*$ be random draws with replacement from $X_1, X_2, \dots, X_{N(t)}$.

THEOREM 3.1. *Under the conditions of Theorem* 2.1 *we have, for almost all sample sequences* $\{(X_i, T_i)\}$ *and* $N(t)$, *that*

$$\sqrt{t}\sup_x \left| P\left(\frac{\sum_{i\le N(t)} X_i - \lambda\mu t}{\sqrt{\sum_{i\le N(t)} X_i^2}} \le x\right) - P^*\left(\frac{\sum_{i\le N^*(t)} X_i^* - \hat{\lambda}\hat{\mu}t}{\sqrt{\sum_{i\le N^*(t)} X_i^{*2}}} \le x\right)\right| \to 0,$$

as $t \to \infty$.

PROOF. Using Theorem 1 of Babu and Singh (1984), a result similar to Lemma 4 can be established for the bootstrapped version. We also note that for any $c, C > 0$, estimate (A.3) of Lemma 3 holds uniformly for $c < \lambda < C$ and $|\nu_3| < C$. By strong law of large numbers, the empirical versions of $\nu^2, \mu_3, \lambda, \sigma^2, \mu$ converge to the corresponding parameters. This leads to

$$P^*\left(\frac{\sum_{i\le N^*(t)} X_i^* - \hat{\lambda}\hat{\mu}t}{\sqrt{\sum_{i\le N^*(t)} X_i^{*2}}} \le x\right) = \Phi(x) + \frac{1}{6\nu^3\sqrt{\lambda t}}(\mu_3(2x^2+1) - \mu^3(x^2-1) + 3\mu(\sigma^2 + \sigma^2x^2 + \nu^2x^2))\phi(x) + o(t^{-1/2})$$

uniformly in x for almost all sample sequences. Now the result follows from (2.1).

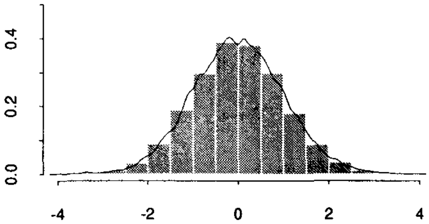

Fig. 9. Normal, normalized, true distribution.

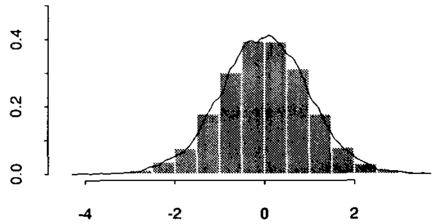

Fig. 10. Normal, normalized, bootstrap distribution.

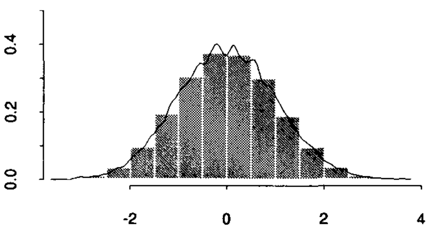

Fig. 11. Normal, studentized, true distribution.

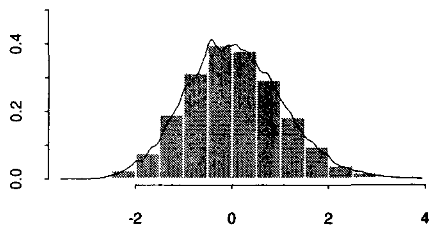

Fig. 12. Normal, studentized, bootstrap distribution.

4. Simulations

Figures 1–12 give simulation results on the true distribution, bootstrap distribution for normalized and studentized compound Poisson processes. Taking $\lambda = 1$, the Poisson process is simulated up to time $t = 40$. Given a realization of Poisson process $N(t)$, $t \leq 40$, the reward random variables $X_1, X_2, \ldots, X_{N(40)}$ are generated from exponential distribution with mean $\mu = 1$ (Figs. 1–4), log-normal distribution $LN(0,1)$ (Figs. 5–8) and normal distribution $N(0,1)$ (Figs. 9–12), respectively. Each figure is based on 10,000 runs. Figures 1, 5, 9 are the simulated true distributions of the normalized processes

$$\frac{\sum_{i=1}^{N(t)} X_i - \lambda\mu t}{\sqrt{\lambda t(\sigma^2 + \mu^2)}}.$$

Figures 2, 6, 10 are the bootstrap distributions for the normalized processes

$$\frac{\sum_{i=1}^{N^*(t)} X_i^* - \hat{\lambda}\hat{\mu} t}{\sqrt{\hat{\lambda} t(\hat{\sigma}^2 + \hat{\mu}^2)}}.$$

Figures 3, 7, 11 are the simulated distributions of the studentized process R_t and Figs. 4, 8, 12 of their bootstrapped versions. From these histograms, it can be seen that the normalized Poisson processes are approximately normally distributed for $t = 40$ whatever the rewards are. But, the studentized ones are seriously skewed to the left for exponential or log-normal rewards. The skewness in each case is almost perfectly captured by the corresponding bootstrap distribution. For normal reward, studentization of the process does not incur any evident skewness.

Acknowledgements

The authors would like to thank the associate editor and the referees for many constructive comments, which helped in improving the paper.

Appendix

We now present some technical lemmas needed in the proof of Theorem 2.1.

LEMMA 1. *If* $X \sim N(x\sqrt{1+\beta^2}, \beta^2)$ *for* β *and* x *real, then*

$$E(\Phi(X)) = \Phi(x) \tag{A.1}$$
$$E(\phi(X)) = (1+\beta^2)^{-1/2}\phi(x)$$
$$E(X\phi(X)) = (1+\beta^2)^{-1}x\phi(x)$$
$$E((X^2-1)\phi(X)) = (1+\beta^2)^{-3/2}(x^2-1)\phi(x).$$

PROOF. Let Z_1, Z_2 be i.i.d. standard normal random variables. If $X = x\sqrt{1+\beta^2} + \beta Z_1$, then $(Z_2 - \beta Z_1)(1+\beta^2)^{-1/2}$ is a standard normal random variable, and

$$E(\Phi(X)) = P(Z_2 \le \beta Z_1 + x(1+\beta^2)^{1/2}) = P((Z_2 - \beta Z_1)(1+\beta^2)^{-1/2} \le x) = \Phi(x).$$

The rest of the three equations follow by taking derivatives of (A.1) with respect to x.

LEMMA 2. *Let* F *and* G *be two probability distribution functions. Suppose* f *is a differentiable function with derivative* f' *and satisfying*

$$\int |f'(y)|dy < \infty, \qquad \int |f|dF < \infty, \qquad \int |f|dG < \infty.$$

Then for any real α *and* β, *we have uniformly in* x,

$$\left|\int f(\alpha x + \beta y)dF(y) - \int f(\alpha x + \beta y)dG(y)\right| \le \sup_y |F(y) - G(y)| \int |f'(u)|du. \tag{A.2}$$

PROOF. If $\beta = 0$, then the left side of (A.2) is zero. For $\beta \neq 0$, the use of integration by parts yields the lemma.

To prove Theorem 2.1, we require a result on one-term Edgeworth expansion for the Poisson Process. However, we derive a result on Edgeworth expansions for a general renewal processes, as it is of interest on its own. Though, Lemma 3 appears to be a standard result, a short and simple proof of it is presented here for the sake of completeness. For a renewal process $\{N(t), t > 0\}$, the inter-arrival times $T_1, T_2, \ldots$ are i.i.d. positive random variables. Let $E(T_1) = \lambda^{-1} > 0$, $\tau^2 = \text{var}(T_1)$, and $\nu_3 = E(T_1 - \lambda^{-1})^3$. Recall that for the Poisson process with rate λ, T_i are exponentially distributed with $\tau^2 = \lambda^{-2}$.

LEMMA 3. *For a renewal process $\{N(t), t > 0\}$, if $\tau > 0$, $E|T_1|^3 < \infty$ and T_1 is a non-lattice random variable, then as $t \to \infty$,*

$$P\left(\frac{N(t)-\lambda t}{\lambda\tau\sqrt{\lambda t}} \le x\right) - \Phi(x) - \frac{\nu_3 - 3\lambda\tau^4}{6\tau^3\sqrt{\lambda t}}(x^2-1)\phi(x) + \frac{h_t(x)}{\lambda\tau\sqrt{\lambda t}}\phi(x) = o(t^{-1/2}) \tag{A.3}$$

uniformly in x, where $h_t(x) = \frac{1}{2}(\lambda\tau)^2 - 1 + \{\lambda t + \lambda\tau x\sqrt{\lambda t}\}$, $\{a\} = a - [a]$, and $[a]$ is the integer part of a. Further, if U is a uniformly distributed random variable on $(-1/2, 1/2)$ and is independent of the process $\{N(t), t > 0\}$, then uniformly in x,

$$P\left(\frac{N(t)-\lambda t + U}{\lambda\tau\sqrt{\lambda t}} \le x\right) - \Phi(x) - \frac{\nu_3 - 3\lambda\tau^4}{6\tau^3\sqrt{\lambda t}}(x^2-1)\phi(x) + \frac{\lambda^2\tau^2 - 1}{2\lambda\tau\sqrt{\lambda t}}\phi(x) = o(t^{-1/2}). \tag{A.4}$$

PROOF. For any real x, define $a_t = a_t(x) = 1 + \lambda t + \lambda\tau x\sqrt{\lambda t}$, $n_t = [a_t]$. We will estimate the left side of (A.3) on $x < -\sqrt{\lambda t}/(2\lambda\tau)$ and $x \ge -\sqrt{\lambda t}/(2\lambda\tau)$ respectively. Uniformly for $x < -\lambda t/(2\lambda\tau)$, we obviously have $\Phi(x) + (x^2+1)\phi(x) = O(t^{-1})$ and by the renewal theorem, that

$$P(N(t) - \lambda t \le x\lambda\tau\sqrt{\lambda t}) \le P(2|N(t) - \lambda t| \ge \lambda t) \le (2/\lambda t)^2 E(N(t) - \lambda t)^2 = O(t^{-1}).$$

So we only consider the case $x \ge -\sqrt{\lambda t}/(2\lambda\tau)$. Note that in this case $n_t > \lambda t/2$. In addition, let $|x| \le \log t$,

$$S_n = \sum_{i=1}^{n} T_i, \quad \text{and} \quad x_t = \frac{n_t - \lambda t}{\lambda\tau\sqrt{n_t}}.$$

By applying the usual Edgeworth expansions for sums of i.i.d. random variables, we have

$$\begin{aligned} P(N(t) - \lambda t \le x\lambda\tau\sqrt{\lambda t}) &= P(N(t) \le a_t - 1) = P(N(t) \le n_t - 1) \\ &= P(S_{n_t} > t) = P(-(S_{n_t} - \lambda^{-1}n_t) < x_t\tau\sqrt{n_t}) \\ &= \Phi(x_t) + \frac{\nu_3}{6\tau^3\sqrt{n_t}}(x_t^2 - 1)\phi(x_t) + o(t^{-1/2}) \end{aligned} \tag{A.5}$$

uniformly for $x \ge -\sqrt{\lambda t}/(2\lambda\tau)$. As $n_t - \lambda t = O(1 + |x|\sqrt{t})$ and $|x| \le \log t$, we have by Taylor series expansion

$$1 - (n_t/\lambda t)^{-1/2} = ((n_t - \lambda t)/2\lambda t) + O(((n_t - \lambda t)/\lambda t)^2)$$

$$= (\lambda\tau x/2\sqrt{\lambda t}) + O((1+x^2)t^{-1}) \tag{A.6}$$

$$= O((1+x^2)t^{-1/2}). \tag{A.7}$$

Estimates (A.6) and (A.7) yield,

$$\begin{aligned} x_t - x - \frac{2(1 - \{a_t\}) - \lambda^2\tau^2x^2}{2\lambda\tau\sqrt{\lambda t}} &= \frac{1 - \{a_t\}}{\lambda\tau}\left(\frac{1}{\sqrt{n_t}} - \frac{1}{\sqrt{\lambda t}}\right) - x\left(1 - \sqrt{\frac{\lambda t}{n_t}} - \frac{\lambda\tau x}{2\sqrt{\lambda t}}\right) \\ &= O(t^{-1}(1 + |x|^3)), \end{aligned} \tag{A.8}$$

and hence $|x_t - x| = O(t^{-1/2}(1+|x|^3))$. By another Taylor series expansion

$$\Phi(x_t) = \Phi(x) + (x_t - x)\phi(x) - \frac{1}{2}(x_t - x)^2\phi(\xi_{t,1})\xi_{t,1}, \tag{A.9}$$

$$(x_t^2 - 1)\phi(x_t) = (x^2-1)\phi(x) + O((x_t - x)(3\xi_{t,2} - \xi_{t,2}^3)\phi(\xi_{t,2})), \tag{A.10}$$

for some $\xi_{t,1}, \xi_{t,2}$ between x_t and x. Since $y^k\phi(y)$ is uniformly bounded for any $k \geq 0$, we have by (A.8),

$$\begin{aligned} &(x_t - x)^2\phi(\xi_{t,1})\xi_{t,1} = O(t^{-1}) \quad \text{and} \\ &(x_t - x)(3\xi_{t,2} - \xi_{t,2}^3)\phi(\xi_{t,2}) = O(t^{-1/2}), \end{aligned} \tag{A.11}$$

uniformly for $x \geq -\sqrt{\lambda t}/(2\lambda\tau)$ and $|x| \leq \log t$. In particular, it implies that for $x \leq -\log t$,

$$P(N(t) - \lambda t \leq x\sqrt{\lambda t}) \leq P(N(t) - \lambda t \leq -\sqrt{\lambda t}\log t) = o(t^{-1/2}),$$

and for $x \geq \log t$,

$$P(N(t) - \lambda t \leq x\sqrt{\lambda t}) \geq P(N(t) - \lambda t \leq \sqrt{\lambda t}\log t) = 1 - o(t^{-1/2}).$$

Consequently, (A.3) holds uniformly for $x \geq -\sqrt{\lambda t}/(2\lambda\tau)$, by (A.5), (A.8)–(A.11). The expansion (A.4) follows as $E\{a + U\} = 1/2$ for any real number a. This completes the proof of Lemma 3.

Remark. For the Poisson process $N(t)$, the results reduce to

$$P(N(t) - \lambda t \leq x\sqrt{\lambda t}) = \Phi(x) - \frac{1}{6\sqrt{\lambda t}}(x^2 - 4 + 6\{\lambda t + x\sqrt{\lambda t}\})\phi(x) + o(t^{-1/2}),$$

and

$$P(N(t) - \lambda t + U \leq x\sqrt{\lambda t}) = \Phi(x) - \frac{1}{6\sqrt{\lambda t}}(x^2 - 1)\phi(x) + o(t^{-1/2}), \tag{A.12}$$

uniformly in x.

Although the Edgeworth expansions for studentized random variables are well known, the explicit forms of expansions for 'perturbed' studentized random variables are not easily available. Before stating the next lemma on such expansions, we establish some notation. Let $\{m_n\}$ be a sequence of real numbers satisfying $m_n = O(\log n)$. Define

$$\begin{aligned} &\theta = E(X_1^3) - \mu\nu = \mu_3 + 2\mu\sigma^2, \quad \sigma_n^2 = (\sigma/\nu)^2 - (\theta\mu/\sqrt{n})m_n\nu^{-4}, \\ &\theta_n = (\mu/\nu)m_n - (\theta/2\sqrt{n})\nu^{-3}, \quad \gamma_3 = \mu_3\nu^{-3} - 3\sigma^2\nu^{-5}\theta, \quad Z_{n,x} = (\nu x - \mu m_n)\sigma^{-1}, \\ &\overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i, \quad W_n = \sqrt{n}(\overline{X_n} - \mu), \quad \text{and} \quad \overline{X_n^2} = \frac{1}{n}\sum_{i=1}^n X_i^2. \end{aligned}$$

LEMMA 4. *Suppose $E(X_1^6) < \infty$ and the distribution of X_1 has a continuous component. Then, as $n \to \infty$,*

$$\begin{aligned} \sup_x |P(\sqrt{n}(\overline{X_n} - \mu) + m_n\mu \leq x\sqrt{\overline{X_n^2}}) - \Phi(Z_{n,x}) \\ + (6\sigma^3\sqrt{n})^{-1}(\mu_3(Z_{n,x}^2 - 1) - 3\theta x\sigma\nu^{-1}Z_{n,x})\phi(Z_{n,x})| = o(n^{-1/2}). \end{aligned}$$

PROOF. We start with the computation of the first three "approximate cumulants" of

$$(W_n + \mu m_n)(\overline{X_n^2})^{-1/2} = Z_n + O(U_n^2(|W_n| + \log n)/n), \tag{A.13}$$

where

$$U_n = \sqrt{n}(\overline{X_n^2} - \nu^2), \quad \text{and} \quad Z_n = \frac{1}{\nu}W_n - \frac{1}{2\nu^3\sqrt{n}}W_nU_n + \frac{\mu}{\nu}m_n - \frac{\mu}{2\nu^3\sqrt{n}}m_nU_n.$$

A simple algebra leads to

$$\text{(A.14)} \quad E(U_nW_n) = \theta, \quad Z_n - \theta_n = \frac{1}{\nu}W_n - \frac{1}{2\nu^3\sqrt{n}}(\mu m_n U_n + (W_nU_n - \theta)),$$
$$E(Z_n) = \theta_n, \quad \mathrm{Var}(Z_n) = \sigma_n^2 + O((1 + m_n^2)/n),$$

and

$$\text{(A.15)} \quad \nu^3 E(Z_n - \theta)^3 = E(W_n^3) - \frac{3}{2\nu^3\sqrt{n}}E(W_n(W_nU_n - \theta))$$
$$+ O(n^{-1}(1 + m_n^2) + n^{-3/2}E|U_nW_n|^3).$$

So the third "approximate cumulant" of (A.13) is given by

$$\frac{1}{\sqrt{n}}(\mu_3\nu^{-3} - 3\sigma^2\theta\nu^{-5}) = \frac{\gamma_3}{\sqrt{n}}.$$

To establish the validity of the formal one term Edgeworth expansion, note that the distribution of (X_1, X_1^2) is strongly non-lattice. So by Theorem 20.8, 24.2 and Lemma 24.1 of Bhattacharya and Ranga Rao (1986), it follows that the distribution of (W_n, U_n) has a valid one term Edgeworth expansion. Hence by (A.14) and (A.15), as in Lemma 3 of Babu and Singh (1984) or Theorem 2 of Babu and Singh (1985), it follows that

$$\text{(A.16)} \quad P(W_n + \mu m_n \le x\sqrt{\overline{X_n^2}}) = \Phi((x - \theta_n)/\sigma_n) + o(n^{-1/2})$$
$$- \frac{\gamma_3}{6\sigma_n^3\sqrt{n}}(((x - \theta_n)/\sigma_n)^2 - 1)\phi((x - \theta_n)/\sigma_n),$$

uniformly in x. See also Bhattacharya and Ghosh (1978). Since

$$\frac{1}{\sigma_n} = \frac{\nu}{\sigma}\left(1 - \frac{\theta\mu m_n}{\sigma^2\nu^2\sqrt{n}}\right)^{-1/2} = \frac{\nu}{\sigma}\left(1 + \frac{\theta\mu m_n}{2\sigma^2\nu^2\sqrt{n}}\right) + O\left(\frac{m_n^2}{n}\right),$$

and

$$\frac{x - \theta_n}{\sigma_n} = Z_{n,x} + Z_{n,x}\frac{\mu m_n \theta}{2\sigma^2\nu^2\sqrt{n}} + \frac{\theta}{2\nu^2\sigma\sqrt{n}} + O\left(\frac{1}{n}(1 + |x|m_n^3)\right),$$

the lemma now follows from (A.16) and the estimate

$$\Phi\left(\frac{x - \theta_n}{\theta_n}\right) = \Phi(Z_{n,x}) + \frac{1}{2\sqrt{n}}\left(\frac{Z_{n,x}\theta\mu m_n}{\sigma^2\nu^2} + \frac{\theta}{\nu^2\sigma}\right)\phi(Z_{n,x}) + o(n^{-1/2}),$$

which holds uniformly in x. This completes the proof.

References

Babu, G. J. and Bose, A. (1988). Bootstrap confidence intervals, *Statist. Probab. Lett.*, **7**, 151–160.

Babu, G. J. and Singh, K. (1984). On one term Edgeworth correction by Efron's bootstrap, *Sankhyā Ser. A*, **46**, 219–232.

Babu, G. J. and Singh, K. (1985). Edgeworth expansions for sampling without replacement from finite populations, *J. Multivariate Anal.*, **17**, 261–278.

Bhattacharya, R. N. and Ghosh, J. K. (1978). On the validity of formal Edgeworth expansion, *Ann. Statist.*, **6**, 434–451.

Bhattacharya, R. N. and Ranga Rao, R. (1986). *Normal Approximation and Asymptotic Expansions*, 2nd ed., Wiley, New York.

Bose, A. and Babu, G. J. (1991). Accuracy of the bootstrap approximation, *Probab. Theory Related Fields*, **90**, 301–316.

Hall, P. (1988). Theoretical comparison of bootstrap confidence intervals (with discussions), *Ann. Statist.*, **16**, 927–985.

Hipp, C. (1985). Asymptotic expansions in the central limit theorem for compound and Markov processes, *Z. Wahrsch. verw. Gebiete.*, **69**, 361–385.

Kusuoka, S. and Yoshida, N. (2000). Malliavin calculus, geometric mixing, and expansions of diffusion functionals, *Probab. Theory Related Fields*, **116**, 457–484.

Kutoyants, Yu. A. (1998). *Statistical Inference for Spatial Poisson Processes*, Springer, New York.

Mykland, P. A. (1992). Asymptotic expansions and bootstrapping distributions for dependent variables: A martingale approach, *Ann. Statist.*, **20**, 623–654.

Mykland, P. A. (1993). Asymptotic expansions for martingales, *Ann. Probab.*, **21**, 800–818.

Mykland, P. A. (1995*a*). Martingale expansions and second order inference, *Ann. Statist.*, **23**, 707–731.

Mykland, P. A. (1995*b*). Embedding and asymptotic expansions for martingales, *Probab. Theory Related Fields*, **103**, 475–492.

Singh, K. (1981). On asymptotic accuracy of Efron's bootstrap, *Ann. Statist.*, **9**, 1187–1195.

Yoshida, N. (2001). Malliavin calculus and martingale expansion, *Bull. Sci. Math.*, **125**, 431–456.