Zeitschrift für Operations Research, Volume 29, page 217 – 223

Coupling Methods in Connection with Poisson Process Approximation

By *D. Pfeifer*, Aachen[1]

Received November 1984
(Revised version April 1984)

Abstract: Motivated by a simple probabilistic model for the radioactive decay, we show that *Serfling's* [1978] approach to Poisson approximation using coupling techniques can in a natural way also be applied to Poisson process approximation. This provides at the same time uniform estimations for the deviation of a Markov-Bernoulli process from the approximating Poisson process with respect to the total variation distance. An application to quasirandom input queuing models is also given.

Keywords: Radioactive decay, quasirandom input queuing model, Poisson process approximation, coupling techniques

1 A Probabilistic Model of Radioactive Decay

In a famous experiment *Geiger* and *Rutherford* came to the conclusion that the number of particles emitted from a radioactive source within a short time period should approximately follow a Poisson distribution, the number of particles observed in non-overlapping time intervals being independent of each other. In the setting of stochastic processes, this means that the number of particles counted up to time $t > 0$ approximately follows a Poisson process with constant intensity $\lambda > 0$, say. Of course, this model cannot be appropriate for large values of t since after some finite time, the source will be exhausted which means that no further particles are observable.

Alternatively, due to the lack of memory property of radioactive decay, we might assume that actually a multiple of the empirical distribution function of an exponentially distributed population with mean $1/\mu > 0$ is observed, which is a sum of independent Markov-Bernoulli processes with states $\{0, 1\}$ and exponential jump time distribution. To be more precise, assume that the source consists of N atoms of the same type, acting independently of each other. Let T_i, $i = 1, 2, \ldots, N$ denote the

[1] *D. Pfeifer*, Institut für Statistik und Wirtschaftsmathematik, RWTH Aachen, Wüllnerstr. 3, D–5100 Aachen.

time of decay of atom i. Then $N_i(t) = I(T_i \leqslant t), t \geqslant 0$ represents the Markov-Bernoulli decay process for atom i, where $I(.)$ denotes the indicator random variable for the corresponding event. Let $N(t) = \sum_{i=1}^{N} N_i(t), t \geqslant 0$. Then $N(t), t \geqslant 0$ represents the decay process for the whole source, and

$$P(N_i(t) = 1) = P(T_i \leqslant t) = 1 - e^{-\mu t} \tag{1}$$

$$P(N(t) = n) = \binom{N}{n}(e^{\mu t} - 1)^n e^{-\mu t N}, \quad t \geqslant 0, n = 1, 2, \ldots, N. \tag{2}$$

Now assume that μt is small. Then $N_i(t)$ is close to a Poisson random variable with mean $-\log(e^{-\mu t}) = \mu t$ according to *Serfling* [1975, 1978; cf. also *Deheuvels/Pfeifer*] which means that $N(t)$ is approximately Poisson distributed with mean $\mu t N$. But in fact, much more can be said here; namely, it is possible to show that using Serfling's approach of coupling techniques, $N(t)$ can be approximated by a Poisson process $M(t)$ with parameter $\lambda = \mu N$ uniformly in a small neighbourhood of the origin, providing thus a logical explanation of Geiger's and Rutherford's results. This also becomes apparent from the fact that $\bar{N}(t) = 1/N\, N(t), t > 0$ is just the empirical distribution function of an exponentially distributed population with mean $E(\bar{N}(t)) = 1 - e^{-\mu t}$ and variance $\mathrm{Var}(\bar{N}(t)) \leqslant 1/N$ (which explains for the almost "deterministic" exponential law of radioactive decay since in practice N is of order $6 \cdot 10^{23}$ atoms per mole), while $E(\bar{M}(t)) = \mu t \sim 1 - e^{-\mu t} = E(\bar{N}(t))$ for small values of t. Here similarly, $\bar{M}(t) = 1/N\, M(t), t \geqslant 0$. From a practical point of view, it would thus be desirable to have some estimations for the distance between the Markov-Bernoulli decay process and the approximating Poisson process avalaible. A suitable distance measure here is the toal variation distance given in terms of non-negative integer-valued random variables X and Y by

$$d(X, Y) = \sup_{A \subseteq \mathbf{Z}^+} \{|P(X \in A) - P(Y \in A)|\} = \frac{1}{2} \sum_{k=0}^{\infty} |P(X = k) - P(X = k)|. \tag{3}$$

It will be shown in the sequel that in our case,

$$d(N(t), M(t)) \leqslant 1 - (1 + \mu t)^N e^{-\mu t N} \leqslant \frac{N}{2}(\mu t)^2, \quad t \geqslant 0. \tag{4}$$

To give an example, consider a source consisting of one gram of the uranium isotope U^{238} which posseses a half-life of about $4.5 \cdot 10^9$ years [cf. *Wichmann*, chapter 7]. Then (4) indicates that with a maximum error of less than 10^{-3} a Poisson process approximation is justified within a time span of more than two days! (A further discussion on the probabilistic model of radioactive decay described above can be found in *Weise*).

Of course, similar considerations are possible when radioactive sources with different portions of different isotopes (implying varying half-lifes) are involved. For instance, if the source consists of N_1 atoms with exponentially distributed life lengths with mean $1/\mu_1$ and N_2 atoms with exponentially distributed life lengths with mean $1/\mu_2$, and $N_i^1(t)$ and $N_j^2(t)$, $t \geqslant 0$ represent the individual Markov-Bernoulli decay processes, $i = 1, 2, \ldots, N_1$, $j = 1, 2, \ldots, N_2$, then the decay process for the whole source, $N(t) = \sum_{i=1}^{N_1} N_i^1(t) + \sum_{j=1}^{N_2} N_j^2(t)$, $t \geqslant 0$, can similarly be approximated by a Poisson process $M(t)$, $t \geqslant 0$ which now has intensity $\lambda = \mu_1 N_1 + \mu_2 N_2$ [see e.g. *Çinlar*]. Then similarly to (4), we have an estimation of the form

$$d(N(t), M(t)) \leqslant 1 - (1 + \mu_1 t)^{N_1} e^{-\mu_1 t N_1} (1 + \mu_2 t)^{N_2} e^{-\mu_2 t N_2}$$
$$\leqslant \frac{N_1}{2} (\mu_1 t)^2 + \frac{N_2}{2} (\mu_2 t)^2, \; t \geqslant 0. \tag{5}$$

Extensions to more than two isotopes involved are obvious from this; see also the Theorem below.

2 The Coupling Techniques

We shall give a short account on couplings first. Let X and Y be arbitrary random variables with values in a real measurable space $(\mathbf{R}, \mathcal{B})$, and define analogously to (3) the (general) total variation distance d by

$$d(X, Y) = \sup_{a \in \mathcal{B}} \{ | P(X \in A) - P(Y \in A) | \}. \tag{6}$$

Due to a well-known inequality of Doeblin [see *Serfling*, 1975, 1978] we always have

$$d(X, Y) \leqslant P(X \neq Y); \tag{7}$$

if X and Y are additionally constructed in such a way that equality holds in(7), then (X, Y) is called a (maximal) coupling. Although *Strassen* has shown that it is always possible to construct a maximal coupling, the problem remains to find an easy way to do so in concrete cases. For Poisson process approximation we can proceed as follows.

Lemma 1

Let T denote the time of the first jump in a Poisson process $M(t)$, $t \geqslant 0$ with intensity $\lambda > 0$, and define

$$N(t) = I(T \leqslant t), \; t \geqslant 0. \tag{8}$$

Then $N(t)$ is a Markov-Bernoulli process, and $(N(t), M(t))$ is a maximal coupling for all $t \geqslant 0$.

Proof

Obviously, $N(t)$ forms a Markov process with states $\{0, 1\}$ and $P(N(t) = 1) = P(T \leqslant t) = 1 - e^{-\lambda t}$. Further, we have

$$P(N(t) \neq M(t)) = P(M(t) \geqslant 2) = 1 - (1 + \lambda t)\, e^{-\lambda t},\ t \geqslant 0 \tag{9}$$

which coincides with $d(N(t), M(t))$ by Lemma 4.1 in *Serfling* [1978]. This gives the desired result.

Lemma 1 essentially says that in a maximal coupling, the Markov-Bernoulli process is a truncated Poisson process, both processes coinciding for $t < T_2$ where T_2 is the time of the second jump of the Poisson process. Choosing $t = 1$ we see that $N(1)$ is a binomial random variable with mean $p = 1 - e^{-\lambda}$ while $M(1)$ is a Poisson random variable with mean $\lambda = -\log(1 - p)$. This explains why a Poisson distribution with mean $\lambda = -\log(1-p)$ is closer to a binomial with mean p than a Poisson with mean p; [see *Serfling*, 1978; or *Deheuvels/Pfeifer*].

Theorem

Let $\{N_i(t); t \geqslant 0\}$, $i = 1, 2, \ldots, N$ be independent Markov-Bernoulli processes with exponentially distributed jump times with mean $1/\mu_i$, and $\{M(t); t \geqslant 0\}$ a Poisson process with intensity $\lambda = \sum_{i=1}^{N} \mu_i$. Then

$$d\left(\sum_{i=1}^{N} N_i(t), M(t)\right) \leqslant 1 - \prod_{i=1}^{N} \{(1 + \mu_i t)\, e^{-\mu_i t}\} \leqslant \frac{1}{2} \sum_{i=1}^{N} (\mu_i t)^2,\ t \geqslant 0. \tag{10}$$

Proof

Since the total variation distance only depends on the distribution of the random variables involved, we may assume that $\{M_i(t); t \geqslant 0\}$, $i = 1, 2, \ldots, N$ are Poisson processes which are maximally coupled with $N_i(t)$, $t \geqslant 0$ each according to Lemma 1. Then $M(t) = \sum_{i=1}^{N} M_i(t)$, $t \geqslant 0$ is a superposition of Poisson processes with intensities μ_i each, giving again a Poisson process with intensity $\lambda = \sum_{i=1}^{N} \mu_i$. The estimation (10) now follows from *Deheuvels/Pfeifer*.

It would be interesting to see whether estimation (10) is sharp in some sense. In fact, by Theorem 1.1 in *Deheuvels/Pfeifer*, and the arguments used above, it follows that

$$d\left(\sum_{i=1}^{N} N_i(t), M(t)\right) \geqslant \frac{1}{2} \sum_{i=1}^{N} (\mu_i t)^2 \exp\left(-\sum_{i=1}^{N} \mu_i t\right) \tag{11}$$

as long as $t \leqslant \{\sum_{i=1}^{N} \mu_i\}^{-1}$ which means that for small values of t, estimation (10) cannot essentially be improved.

Finally, it should be pointed out that by Poisson process approximation, a maximal coupling can also easily be constructed starting with the Markov-Bernoulli process rather than the Poisson process itself. For this purpose, let $\{E_n; n \in \mathbf{N}\}$ be a sequence of i.i.d. exponentially distributed random variables with mean $1/\lambda > 0$, independent from a two-state Markov Bernoulli process with exponentially distributed jump time T with mean $1/\lambda$. Define

$$T_1 = T, T_{n+1} = T_n + E_n, n \geqslant 1. \tag{12}$$

Then $\{T_n; n \in \mathbf{N}\}$ is the arrival time sequence of a Poisson process $M(t)$, $t \geqslant 0$, say, with intensity λ, which is maximally coupled with the given Markov-Bernoulli process $N(t) = I(T \leqslant t)$, $t \geqslant 0$. This follows from the fact that by our construction, $N(t)$ is just the truncation of $M(t)$, $t \geqslant 0$ according to Lemma 1.

The last remarks also allow for a similar approach to simple Poisson approximation, i.e. a construction of a maximal coupling for a single $\{0, 1\}$-valued Bernoulli random variable X with mean $p \in (0, 1)$, say. We only have to use an appropriate imbedding technique as follows. Let $\lambda = -\log(1-p)$ and again $\{E_n; n \in \mathbf{N}\}$ be an i.i.d. sequence of exponentially distributed random variables with mean $1/\lambda$. Define (frac = fractional part)

$$T_1 = (1 + E_1) I(X = 0) + \text{frac}(E_1) I(X = 1),$$

$$T_{n+1} = T_n + E_{n+1}, n \geqslant 1. \tag{13}$$

If $N(t)$, $t \geqslant 0$ denotes the Markov-Bernoulli process with jump time T_1 and $M(t)$, $t \geqslant 0$ the Poisson process with arrival-time sequence $\{T_n; n \in \mathbf{N}\}$, then according to what has been said above, $N(t)$ and $M(t)$ are maximally coupled for each t; but also, $X = N(1)$, hence $(X, M(1))$ is the desired maximal coupling. We only have to prove that T_1 in fact follows an exponential distribution, which is obvious since for $x \geqslant 0$,

$$\begin{aligned} P(T_1 > x) &= (1-p) P(1 + E_1 > x) + pP(\text{frac}(E_1) > x) \\ &= e^{-\lambda} \{e^{-\lambda(x-1)} I(x > 1) + I(x \leqslant 1)\} + (1 - e^{-\lambda}) \frac{e^{-\lambda x} - e^{-\lambda}}{1 - e^{-\lambda}} I(x \leqslant 1) \\ &= e^{-\lambda x}. \end{aligned}$$

3 Applications to Quasirandom Input Queuing Models

In many birth-death queuing models it is assumed that requests for service occur according to a Poisson process with intensity λ, say; this case is then also referred to as completely random input [cf. Ch. 3, Sec. 3–7, to 3–12. in *Cooper*]. On the other hand, it is some times more realistic to think of a system in which the requests for service are generated by a finite number of sources. In such a system the probability of an arrival in a small time interval $(t, t + h)$ will not be independent of the system state at time t, but will depend on the number of sources idle at time t. For instance, if we consider a finite-source system with an equal number of sources and servers, then no new requests can occur as long as all the servers are occupied, since there are no idle sources to generate new requests. The probability of blocking hence is zero, whereas the portion of time all servers are busy can take any value between zero and one. Such a kind of finite-source input is often called quasirandom input; to be more precise, it is assumed that the probability for any particular source to g generate a request for service in a small time interval $(t, t + h)$ is $\mu h + o(h)$ for $h \to 0$ if the source is idle at time t, and zero if the source is not idle at time t, independently of the states of any other sources. From this it follows that, if a particular source is idle at time t, the distribution of time from t until the source next generates a request for service is exponential with mean $1/\mu$. Equivalently, the arrival process is just a Markov-Bernoulli process of the form (1) and (2) as is the radioactive decay process which thus represents a special case of a quasirandom input queuing model. By the Theorem above we now see that for a short time period, such a quasirandom input model can be approximated by a completely random input model, allowing at the same time for an estimation of the "distance" between both models (in a probabilistic sense). Consider, as an example, a job shop consisting of N machines and a single repairman, and suppose that the amount of time a machine runs before breaking down is exponentially distributed with rate μ and the amount of time it takes the repairman to fix any broken machine is exponential with rate τ [This is Example 5.5 (b) in *Ross*]. Then we have a quasirandom input birthdeath model with birth and death intensities μ_n and τ_n, respectively, given by

$$\begin{aligned} \mu_n &= \begin{cases} (N-n)\,\mu & n \leqslant N \\ 0 & n > N \end{cases} \\ \tau_n &= \tau \qquad\qquad\qquad n \geqslant 1. \end{aligned} \tag{15}$$

For all small time period, the repair system hence behaves approximately like a M/M/1 queue, since by the Theorem, the arrival process is close to a Poisson process with rate $N\mu$.

References

Geiger, H., and *E. Rutherford*: The probability variations in the distribution of α particles. (With a note by H. Bateman). Phil Mag. 6, 20, 1910, 698–707.

Çinlar, E.: Introduction to Stochastic Processes. Prentice-Hall Inc. Englewood Cliffs, N.J. 1975.

Cooper, R.B.: Introduction to Queuing Theory. Macmillan, N.Y. 1972.

Deheuvels, P., and *D. Pfeifer*: A semi-group approach to Poisson approximation. To appear in: Ann. Prob., 1985.

Ross, S.M.: Stochastic Processes. Wiley, N.Y. 1983.

Serfling, R.J.: A general Poisson approximation theorem. Ann. Prob. **3**, 1975, 726–731.
–: Some elementary results on Poisson approximation in a sequence of Bernoulli trials. SIAM Review **20**, 1978, 567–579.
Strassen, V.: The existence of probability measures with given margnials. Ann. Math. Statist. **36**, 1965, 423–439.
Weise, L.: Statistische Auswertung von Kernstrahlungsmessungen. R. Oldenbourg, München 1971.
Wichmann, E.H.: Quantum Physics. Berkeley Physics Course, Vol. 4. McGraw-Hill, N.Y. 1971.