The Analysis of Some Algorithms for Generating Random Variates with a Given Hazard Rate*

Luc Devroye
McGill University, School of Computer Science, McGill University, 805 Sherbrooke Street West, Montreal, Canada H3A 2K6

We analyze the expected time performance of two versions of the thinning algorithm of Lewis and Shedler for generating random variates with a given hazard rate on $[0,\infty)$. For thinning with fixed dominating hazard rate $g(x) = c$ for example, it is shown that the expected number of iterations is $cE(X)$ where X is the random variate that is produced. For DHR distributions, we can use dynamic thinning by adjusting the dominating hazard rate as we proceed. With the aid of some inequalities, we show that this improves the performance dramatically. For example, the expected number of iterations is bounded by a constant plus $E(\log_+(h(0)X))$ (the logarithmic moment of X).

1. INTRODUCTION

We consider the problem of the computer generation of random variables with given hazard rate h on $[0,\infty)$. If X is a random variable with density f and distribution function F, the hazard rate h and the cumulative hazard rate H are related by:

$$h(x) = \frac{f(x)}{1 - F(x)}, \quad H(x) = \int_0^x h(y)dy = -\log(1 - F(x)),$$

$$F(x) = 1 - e^{-H(x)}, \quad f(x) = h(x)e^{-H(x)}.$$

The main principles of random variate generation when f and/or F are given can be extrapolated to the case that only h and/or H are given. For example, inversion, composition, and rejection have straightforward generalizations, developed by several authors. In this article we would like to give some results regarding the expected time performance of these methods, and to discuss a fast method (called dynamic thinning) for generating random variates with decreasing hazard rate.

1.1. The Inversion Method

FACT: If E is exponentially distributed, then $H^{-1}(E)$ has hazard rate h. Conversely, if X has hazard rate h, $H(X)$ is exponentially distributed.

PROOF: If H is strictly increasing,

$$P(H^{-1}(E) \le x) = P(E \le H(x)) = 1 - e^{-H(x)} = F(x), \quad \text{all } x > 0.$$

*This research was supported by NSERC Grant No. A-3456 and by Quebec Ministry of Education Grant No. FCAC-EQ-1679. Part of this research was carried out while the author was on sabbatic leave at the Department of Statistics, Australian National University, GPO Box 4, Canberra, A.C.T. 2601, Australia.

Naval Research Logistics Quarterly, Vol. 33, pp. 281–292 (1986)
 CCC 0028-1441/86/020281-12$04.00

If H is not strictly increasing, the chain of equalities remains valid for any way of defining H^{-1}.

This method was mentioned in the context of nonhomogeneous Poisson point process simulation by Cinlar [3], Kaminsky, and Rumpf [6] and Lewis and Shedler [8], and more directly by Gaver [5].

EXAMPLES: Let $a > 0$ be a parameter. For the *Weibull* (a) density $ax^{a-1}e^{-x^a}$ on $[0,\infty)$, we have

$$h(x) = ax^{a-1}, H(x) = x^a.$$

Thus, Weibull (a) random variates can be obtained as $E^{1/a}$. The *Pareto* (a) density $a/(1 + x)^{a+1}$ on $[0,\infty)$ has

$$h(x) = a/(1 + x), H(x) = a \log(1 + x),$$

and random variates can be obtained as $\exp(E/a) - 1$. For the power function [or beta $(a,1)$] density on $[0,1]$, $f(x) = ax^{a-1}$, we obtain

$$h(x) = ax^{a-1}/(1 - x^a),$$

and H is easily invertible in the cases $a = 1$ (uniform density), and $a = 2$ (triangular density).

1.2. The Composition Method

FACT: If $h = h_1 + \ldots + h_n$ where h_i, $1 \le i \le n$, are individual hazard rate functions, then the random variable

$$\min(X_1, \ldots, X_n)$$

has hazard rate h when $X_1, \ldots, X_n$ are independent random variables with the given hazard rates.

PROOF: Since the cumulative hazard rates follow the same decomposition, $H = H_1 + \ldots + H_n$, we have for all x,

$$P(\min X_i \ge x) = \sum_{i=1}^{n} (1 - F_i(x)) = \sum_{i=1}^{n} e^{-H_i(x)} = e^{-H(x)},$$

which was to be shown.

We note that composition is usually an expensive operation since the cost incurred is the sum of the costs for the individual H_is.

1.3. The Thinning Method

FACT: Assume that $h \le g$, where g is another hazard rate on $[0,\infty)$. If $0 < Y_1 < Y_2 < \ldots$ is a nonhomogeneous Poisson point process with rate function g, and $U_1, U_2, \ldots$ is a sequence of independent uniform $[0,1]$ random variables, independent of the Y_is,

and if i is the smallest index for which $U_i g(Y_i) \leq f(Y_i)$, then the random variable $X = Y_i$ has hazard rate h.

PROOF: The statement follows from the following facts: (1) the subsequence of the Y_is for which $U_i g(Y_i) \leq f(Y_i)$ is a nonhomogeneous Poisson point process with rate function h (see Lewis and Shedler [8]); and (2) the first realization of such a process is a random variable with hazard rate h.

Lewis and Shedler [7], [9] discuss random variate generation for rate functions of the form $h(x) = \exp(a_0 + a_1 x + a_2 x^2)$, and illustrate the methods outlined above. We note here that the thinning method needs a nonhomogeneous Poisson point process generator for rate function g. If G is the cumulative rate function corresponding to g, this can be done as follows: let $E_1, E_2, \ldots$ be independent exponential random variables, and let Z_i be the partial sum $E_1 + \ldots + E_i$. Then the sequence $Y_i = G^{-1}(Z_i)$, $1 \leq i$, has the property needed for the thinning method.

2. EXPECTED TIME ANALYSIS OF THE THINNING METHOD

THEOREM 1: Let N be the number of Y_is needed in the thinning method. Then

$$E(N) = \int_0^\infty g(x)(1 - F(x))dx = \int_0^\infty f(x)G(x)dx.$$

PROOF: The random vectors $(Y_1, U_1 g(Y_1)), (Y_2, U_2 g(Y_2)), \ldots$ form a homogeneous Poisson process in the area bounded by the y-axis, the x-axis and the curve g. Thus, if X is the random variate generated by the thinning method, we have

$$\begin{aligned} E(N) &= 1 + E\left(\int_0^X (g(y) - h(y))dy\right) = 1 + E\left(\int_0^X g(y)dy\right) - E(H(X)) \\ &= E\left(\int_0^X g(y)dy\right) \quad \text{since } H(X) \text{ is exponentially distributed} \\ &= \int_0^\infty f(x)G(x)dx \\ &= \int_0^\infty g(x)(1 - F(x))dx \quad \text{by a change of integrals} \end{aligned}$$

Because $E(N)$ is a fair measure of the expected time taken to generate a random variate by the thinning method, the expressions of Theorem 1 can be used in the design of the "best" dominating hazard rate g within a class of dominating hazard rates.

The formula for $E(N)$ is very simple for most particular choices for g. For example, if $g(x) = \sum_{i=0}^n c_i x^i$, we have

$$E(N) = \sum_{i=0}^n (c_i/(i + 1))E(X^{i+1}).$$

Thus, for $g(x) = c$, the most important dominating hazard rate, we obtain

$$E(N) = cE(X),$$

where X is a random variable with density f. To get a feeling for the efficiency of the thinning method when g is close to h, we offer the following inequalities:

$$(1)\quad E(N) \leq \sup_{x>0} \frac{g(x)}{h(x)};$$

$$(2)\quad E(N) \leq \sup_{x>0} \frac{1 - F(x)}{1 - F^*(x)} \quad \text{where } F^* \text{ is the distribution function for } g$$

Inequality (1) follows from $E(N) = \int_0^\infty g(x)/h(x)f(x)dx$, while (2) follows from $E(N) = \int_0^\infty f^*(x)\,(1 - F(x))/(1 - F^*(x))dx$ where f^* is the density corresponding to g.

3. DHR DISTRIBUTIONS AND DYNAMIC THINNING

Distributions with decreasing hazard rate (DHR) are necessarily monotone on $[0,\infty)$ (because $f = he^{-H}$, $h \downarrow$ and $H \uparrow$). For example, the Pareto distribution, or the Weibull and gamma distributions with parameter $a \leq 1$ are DHR.

A brute force method based upon the numerical solution of the equation $H(X) = E$ for X (where E is exponentially distributed) has a particularly elegant implementation for DHR distributions: indeed, because H is concave (Barlow, Marshall, and Proschan, [2]), Newton–Raphson iterations started with 0 converge when $h(0) < \infty$:

$$Y \leftarrow 0.$$

$$\textit{Repeat } X \leftarrow Y,\ Y \leftarrow X + \frac{E - H(X)}{h(X)} \quad \text{Until } X = Y.$$

This algorithm takes, strictly speaking, infinite time, but will in some cases yield variates with reasonable accuracy after just a few iterations.

If a DHR distribution has a bounded density (i.e., $f(0) = h(0) < \infty$), the following algorithm (called *dynamic thinning*) can be employed:

(Initialize) $T \leftarrow 0$.
(Main body) Repeat Top $\leftarrow h(T)$;
Generate E exponential, and U uniform [0,1],
independent of E;
$T \leftarrow T + E/\text{Top}$
Until $U.\text{Top} \leq h(T)$.
(Exit) Exit with $X \leftarrow T$.

Here the top bound varies dynamically as we progress.

The method of dynamic thinning is only applicable when $h(0)$ is finite. We should point out that the general algorithms given in Devroye [4] remain applicable here since $f = h(1 - F)$ is $\downarrow$ when the distribution has the DHR property (see Barlow and Proschan [1] for more properties). Thus, when h and either F or H are given, we can handle the case $h(0) = \infty$ as well.

The dynamic thinning algorithm can of course be considered as a special case of Lewis' and Shedler's thinning algorithm with piecewise constant dominating curve. Unfortunately, because the dominating curve is not a priori fixed, the analysis of Theorem 1 is not applicable. In fact, the value of $E(N)$ is a complicated function of

d is only explicitly known in a few situations (such as for the exponential or o distributions; see Section 4). In Section 4, we will give useful inequalities for in terms of simple or easy-to-compute quantities.

4. INEQUALITIES FOR $E(N)$ IN DYNAMIC THINNING

n this section we offer upper bounds for $E(N)$ in terms of several quantities related ne distribution of X. Each of these bounds has its use, and except for one case, no nd is inadmissible (i.e., is totally dominated by another bound).

f we were to use standard thinning with constant dominating curve $h(0)$, we would ain

$$E(N) = \mu = E(h(0)X).$$

e value for $E(N)$ in dynamic thinning is thus always bounded from above by μ. It also helpful to compare the inequalities against values for $E(N)$ that are exactly own: perhaps the two most prominent cases are the *exponential distribution* ($h(x) =$ 0), all x), in which case dynamic thinning gives us $E(N) = 1$, and the *Pareto* (a) *stribution* ($h(x) = a/(x + 1)$), for which

$$E(N) = \left(\int_0^\infty e^{-z}\left(1 + \frac{z}{a}\right)^{-1} dz\right)^{-1}$$

see Section 4.1 for its derivation).

4.1. A Translation of $-h$ Inequality

THEOREM 2: Let X have a DHR distribution with $h(0) < \infty$. Then

$$E(N) \leq \frac{1}{1 - \beta},$$

where

$$\beta = \sup_{x \geq 0} \int_0^\infty e^{-yh(x)}(h(x) - h(x + y))dy.$$

(Note that in any case, $\beta \in [0,1]$, and that the right-hand side of the inequality should be read as "∞" when $\beta = 1$.)

PROOF: Let $E_1, E_2, \ldots$ be independent exponential random variables, and define the sequence $Y_0, Y_1, Y_2, \ldots$ by the recursive rule: $Y_0 = 0$, $Y_{i+1} = Y_i + E_{i+1}/h(Y_i)$. Note that the Y_i sequence corresponds to the sequence of values of "T" in the dynamic thinning algorithm *if the stopping rule were ignored*. If we take the stopping rule into account, we note that for $i \geq 1$,

$$P(N > i | Y_0, \ldots, Y_i) = \prod_{j=1}^{i}\left(1 - \frac{h(Y_j)}{h(Y_{j-1})}\right).$$

h, and is only explicitly known in a few situations (such as for the exponential or Pareto distributions; see Section 4). In Section 4, we will give useful inequalities for $E(N)$ in terms of simple or easy-to-compute quantities.

4. INEQUALITIES FOR $E(N)$ IN DYNAMIC THINNING

In this section we offer upper bounds for $E(N)$ in terms of several quantities related to the distribution of X. Each of these bounds has its use, and except for one case, no bound is inadmissible (i.e., is totally dominated by another bound).

If we were to use standard thinning with constant dominating curve $h(0)$, we would obtain

$$E(N) = \mu = E(h(0)X).$$

The value for $E(N)$ in dynamic thinning is thus always bounded from above by μ. It is also helpful to compare the inequalities against values for $E(N)$ that are exactly known: perhaps the two most prominent cases are the *exponential distribution* ($h(x) = h(0)$, all x), in which case dynamic thinning gives us $E(N) = 1$, and the *Pareto* (a) *distribution* ($h(x) = a/(x + 1)$), for which

$$E(N) = \left(\int_0^\infty e^{-z}\left(1 + \frac{z}{a}\right)^{-1} dz\right)^{-1}$$

(see Section 4.1 for its derivation).

4.1. A Translation of $-h$ Inequality

THEOREM 2: Let X have a DHR distribution with $h(0) < \infty$. Then

$$E(N) \leq \frac{1}{1 - \beta},$$

where

$$\beta = \sup_{x \geq 0} \int_0^\infty e^{-yh(x)}(h(x) - h(x + y))dy.$$

(Note that in any case, $\beta \in [0,1]$, and that the right-hand side of the inequality should be read as "∞" when $\beta = 1$.)

PROOF: Let $E_1, E_2, \ldots$ be independent exponential random variables, and define the sequence $Y_0, Y_1, Y_2, \ldots$ by the recursive rule: $Y_0 = 0$, $Y_{i+1} = Y_i + E_{i+1}/h(Y_i)$. Note that the Y_i sequence corresponds to the sequence of values of "T" in the dynamic thinning algorithm *if the stopping rule were ignored*. If we take the stopping rule into account, we note that for $i \geq 1$,

$$P(N > i | Y_0, \ldots, Y_i) = \prod_{j=1}^{i}\left(1 - \frac{h(Y_j)}{h(Y_{j-1})}\right).$$

Thus, for $i \geq 2$,

$$P(N > i|Y_0, \ldots, Y_{i-1}) = \prod_{j=1}^{i-1}\left(1 - \frac{h(Y_j)}{h(Y_{j-1})}\right)\int_0^\infty e^{-y}h(Y_{i-1})(h(Y_{i-1}) - h(Y_{i-1} + y))dy$$
$$\leq \beta \prod_{j=1}^{i-1}\left(1 - \frac{h(Y_j)}{h(Y_{j-1})}\right),$$

and we obtain, by a simple induction argument on i, that $P(N > i) \leq \beta^i$, $i \geq 0$. Thus,

$$E(N) = \sum_{i=0}^{\infty} P(N > i) \leq \frac{1}{1 - \beta}.$$

This concludes the proof of Theorem 2.

EXAMPLE: *The Pareto (a) distribution.* When $h(x) = a/(1 + x)$, $a > 0$, we note that the integral in the definition of β is independent of x:

$$\int e^{-ya/(1+x)}\left(\frac{a}{1+x} - \frac{a}{1+x+y}\right)dy = \int e^{-z}\left(1 - \left(1 + \frac{z}{a}\right)^{-1}\right)dz \left(\text{by a transformation } z = \frac{ay}{1+x}\right).$$

Thus, by carefully checking the induction argument in the proof of Theorem 2, we note that

$$P(N > i) = \beta^i, \quad i \geq 0,$$

and

$$E(N) = \left(\int_0^\infty e^{-z}\left(1 + \frac{z}{a}\right)^{-1}dz\right)^{-1} = \frac{1}{1 - \beta},$$

where

$$\beta = 1 - \int_0^\infty e^{-z}\left(1 + \frac{z}{a}\right)^{-1}dz.$$

Thus, the bound in Theorem 2 is sharp for all values of β. In the case of the Pareto (a) distribution, a comparison with standard thinning [with constant dominating curve at $h(0)$] is in order. We have the following chain of inequalities:

$$E(N) = \left(\int_0^\infty e^{-z}\left(1 + \frac{z}{a}\right)^{-1}dz\right)^{-1}$$
$$\leq \frac{a+1}{a} \left(\text{use Jensen's inequality, after noting that } \left(1 + \frac{z}{a}\right)^{-1} \text{ is convex in } z\right)$$
$$< \frac{a}{a-1} \text{ (for all } a > 1)$$
$$= \mu = E(h(0)X) \text{ (the expected number of iterations for thinning with constant dominating curve).}$$

For $0 < a \leq 1$, the comparison is even more dramatic, since $\mu = \infty$! We can put the expected time performance of the dynamic thinning algorithm differently: for the Pareto distribution with mean μ (note: $\mu = a/a - 1$), we have:

$$\sup_{\text{mean}\,\mu<\infty} E(N) \leq 2.$$

EXAMPLE: *The exponential distribution.* Our bound is sharp too for the exponential distribution $h(x) = h(0)$, since $\beta = 0$, and thus $E(N) = 1$.

4.2. A Ratio-of-h Inequality

THEOREM 3: Let X have a DHR distribution with $h(0) < \infty$. Then

$$E(N) \leq \frac{e}{e-1}\gamma,$$

and, when h is convex,

$$E(N) \leq \gamma,$$

where

$$\gamma = \sup_{x \geq 0} \frac{h(x)}{h\left(x + \frac{1}{h(x)}\right)}.$$

PROOF: The inequalities can be obtained by bounding β, defined in Theorem 2, from above. Let $c \in R$ be a constant, and let x be fixed. Then

$$\begin{aligned}
\int_0^\infty e^{-yh(x)}(h(x) - h(x+y))dy &\leq \int_{y>c/h(x)} + \int_0^{c/h(x)} \\
&\leq \int_c^\infty e^{-z}dz + \int_0^{c/h(x)} e^{-yh(x)}(h(x) - h(x + c/h(x)))dy \\
&= e^{-c} + (1 - e^{-c})\left(1 - h\left(x + \frac{c}{h(x)}\right)\Big/ h(x)\right) \\
&= 1 - (1 - e^{-c})h\left(x + \frac{c}{h(x)}\right)\Big/ h(x).
\end{aligned}$$

The first inequality follows upon taking $c = 1$. The second inequality follows by a straightforward application of Jensen's inequality:

$$\begin{aligned}
&\int_0^\infty e^{-yh(x)}h(x)\left(1 - \frac{h(x+y)}{h(x)}\right)dy \\
&\quad \leq 1 - \frac{1}{h(x)}h\left(x + \int_0^\infty e^{-yh(x)}h(x)y\,dy\right) \\
&\quad = 1 - \frac{1}{h(x)}h\left(x + \frac{1}{h(x)}\right).
\end{aligned}$$

This concludes the proof of Theorem 3.

EXAMPLES: For the convex hazard rate $h(x) = a/(1 + x^b)$, $0 < b \leq 1$, $a > 0$, we obtain with a little work, the inequality

$$E(N) \leq \gamma = 1 + a^{-b}.$$

For the Pareto (a) distribution, this yields a bound already obtained in Section 4.1 (set $b = 1$). We also note that for the exponential distribution, the inequality is sharp, since $E(N) = \gamma = 1$.

The inequality of Theorem 3 is always weaker than or equal to the inequality of Theorem 2, but the parameter γ has the advantage that it is in most cases much easier to evaluate, numerically and analytically, than β.

4.3. A Moment Inequality

Theorems 2 and 3 do not directly show what the relation is between the tail of X (as measured for example by the moments of X) and $E(N)$. Since standard thinning has the property that $E(N) = \mu = E(h(0)X)$ when a constant dominating curve is used for a DHR distribution we expect that there should be inequalities linking $E(N)$ for dynamic thinning and μ, that improve over $E(N) \leq \mu$. Without attempting to obtain a "sharp" inequality, we are able to show that $E(N)$ cannot increase faster than $\sqrt{\mu}$:

THEOREM 4: Let X have a DHR distribution with $H(0) < \infty$. Then

$$E(N) \leq \text{Min}(\mu, (8\mu)^{1/2} + 4(8\mu)^{1/4}),$$

where

$$\mu = E(h(0)X)$$

(note that $\mu \geq 1$ in all cases).

The proof is based upon the following Lemma:

LEMMA 1: Let $x \in R$, $p > 2$, and $m \in \{0,1, \ldots, n\}$. Then

$$P(N > n) \leq P(X > x) + h(0)x/p^{n-m} + \left(1 - \frac{1}{p}\right)^m, \quad \text{all } n > 0.$$

PROOF OF LEMMA 1: Let $U_1, U_2, \ldots$ be independent uniform [0,1] random variables, and let $E_1, E_2, \ldots$ be independent exponential random variables. Set $Y_1 \leftarrow E_1/h(0), \ldots, Y_{n+1} \leftarrow Y_n + E_{n+1}/h(Y_n)$. Let $X = Y_N$ be the first Y_i for which $h(Y_{i-1})U_i \leq h(Y_i)$. Clearly, X is a random variable with density f and hazard rate h obtained by dynamic thinning. Let Y_0 be 0. Then, if

$$N_2 = \sum_{i=1}^{n} I_{h(Y_i) > h(Y_{i-1})/p} \quad (I \text{ is the indicator function}),$$

and

$$N_1 = \sum_{i=1}^{n} I_{h(Y_i) \leq h(Y_{i-1})/p},$$

then we can write the following inclusion of events:

$$[N > n] \subseteq [X > x] \cup [X \leq x, N_1 \geq n - m, N > n] \cup [N_2 \geq m, N > n].$$

Now,

$$P(X \le x, N_1 \ge n - m, N > n) \le P(E_1/(h(0)/p^{n-m}) \le x) \le h(0)x/p^{n-m},$$

and

$$P(N_2 \ge m, N > n) \le P(N > n | N_2 \ge m) \le \left(1 - \frac{1}{p}\right)^m.$$

PROOF OF THEOREM 4: In Lemma 1, we will take x_n random, independent of X, and uniformly distributed on $[n/(h(0)C),(n+1)/(h(0)C)]$, where $C > 0$ is a constant to be chosen later. Let p be a constant independent of n, and let $m = \text{ceil}(n/2)$, where ceil(.) is the ceiling function (the smallest integer at least equal to . . .). Using the formula $E(N) = \sum_{n=0}^{\infty} P(N > n)$, and using the bound of Lemma 1, averaged over x_n, we obtain

$$\sum_{n=0}^{\infty} P(X > x_n) = \sum_{n=0}^{\infty} \int_n^{n+1} P(Ch(0)X > t)dt = \int_0^{\infty} P(Ch(0)X > t)dt = CE(h(0)X) = C\mu.$$

Also,

$$\sum_{n=0}^{\infty} \left(1 - \frac{1}{p}\right)^{m_n} = 1 + 2\sum_{j=1}^{\infty} \left(1 - \frac{1}{p}\right)^j = 1 + \frac{2\left(1 - \frac{1}{p}\right)}{\frac{1}{p}} = 2p - 1.$$

Finally,

$$\frac{1}{C}\sum_{n=0}^{\infty} \left(\int_n^{n+1} \text{tdt}\right) p^{-(n-m_n)} = \frac{1}{C}\sum_{n=0}^{\infty} \frac{2n+1}{2} p^{-(n-m_n)}$$

$$= \frac{1}{C}\left(\frac{1}{2} + \frac{3}{2} + \frac{5}{2}p^{-1} + \frac{7}{2}p^{-1} + \ldots\right) = \frac{2}{C}\sum_{n=0}^{\infty} (2n+1)p^{-n}$$

$$= \frac{2}{C}\left(\frac{1}{1 - \frac{1}{p}} + 2\sum_{n=0}^{\infty} np^{-n}\right) = \frac{2}{C}\left(\left(1 - \frac{1}{p}\right)^{-1} + \frac{2}{p}\left(1 - \frac{1}{p}\right)^{-2}\right)$$

$$= \frac{2}{C}\left(1 + \frac{1}{p}\right)\left(1 - \frac{1}{p}\right)^{-2}.$$

A simple addition of the three terms would give us an inequality to work with, were it not for the fact that with little extra work, we can obtain a slightly better upper bound: just use $E(N) \le 1 + \sum_{n=1}^{\infty} (P(X > x_n) + h(0)x_n p^{-(n-m_n)} + (1 - 1/p)^{m_n})$, and average over x_n. This gives, with the estimates computed above,

$$E(N) \le 1 + C\mu - P(X > x_0) + 2(p-1) + \frac{2}{C}\left(\frac{p(p+1)}{(p-1)^2} - \frac{1}{4}\right).$$

But since $h(0)X$ is obviously stochastically greater than an exponential random variate, we have

$$P(X > x_0) = \int_0^1 P(Ch(0)X > t)dt \geq \int_0^1 e^{-t/C}dt$$
$$= C\int_0^{1/C} e^{-z}dz = C(1 - e^{-1/C}) \geq 1 - 1/(2C).$$

Thus, we obtain

$$E(N) \leq C\mu + 2(p - 1) + \frac{2}{C}\frac{p(p + 1)}{(p - 1)^2}.$$

The optimal choice for C is $(2p(p + 1)/\mu(p - 1)^2)^{1/2}$, which after substitution gives

$$E(N) \leq 2(p - 1) + \sqrt{8\mu}\left(\frac{p(p + 1)}{(p - 1)^2}\right)^{1/2}$$
$$< 2(p - 1) + \sqrt{8\mu}\frac{p + 1}{p - 1} = 2(p - 1) + \frac{2\sqrt{8\mu}}{p - 1} + \sqrt{8\mu}.$$

The right-hand side is minimal for $p - 1 = (8\mu)^{1/4}$, and this choice gives the bound of Theorem 4.

REMARK: By taking $p = 3$ in the second-to-last inequality of the proof of Theorem 4, we obtain $E(N) \leq 4 + \sqrt{24\mu}$. This bound is better than μ for all μ at least equal to 35. Bounds with different cross-over points can be obtained by using different values of p in the proof.

4.4. A Logarithmic Moment Inequality

Theorem 4 cannot be used when μ is infinite. In fact, we can ask under what conditions $E(N)$ is finite in general. One sufficient condition seems to be the finiteness of the *logarithmic moment of* X,

$$\lambda = E(\log_+(h(0)X)),$$

where $u_+ = \max(u,0)$. We have

THEOREM 5: Let X have a DHR distribution with $h(0) < \infty$. Then

$$E(N) \leq \inf_{p>2}\left(4p + 2 + \frac{2\lambda}{\log(p - 1)}\right)$$
$$\leq 10 + \frac{2\lambda}{\log^2(1 + \lambda)} + \frac{2\lambda}{\log\left(1 + \frac{\lambda/2}{\log^2(1 + \lambda)}\right)}$$

REMARK: The upper bound of Theorem 5 increases as $2\lambda/\log(\lambda)$ as $\lambda \to \infty$. But because

$$\lambda = E(\log_+(h(0)X)) \leq E(\log(h(0)X + 1)) \leq \log(1 + \mu),$$

we see that the bound of Theorem 5 increases at worst as $2 \log(\mu)/\log\log(\mu)$ as $\mu \to \infty$. This, of course, is a formidable improvement over the asymptotic behavior of the bound of Theorem 4. We should also note that λ is finite for most useful distributions.

PROOF OF THEOREM 5: In Lemma 1, replace n by $2j$, and sum over j. Set $m_{2j} = j$, $p_{2j} = p > 2$, and $h(0)x_{2j} = (p-1)^j$. Since for any random variable Z, $\sum_{j=0}^{\infty} P(Z > j) \le 1 + \int_0^{\infty} P(Z > t)dt = 1 + E(Z_+)$, we see that

$$\begin{aligned} E(N) &\le 2\sum_{j=0}^{\infty} P(N > 2j) \le 2\sum_{j=0}^{\infty}\left(P(h(0)X > (p-1)^j) + 2\left(1 - \frac{1}{p}\right)^j\right) \\ &= 2\sum_{j=0}^{\infty} P(Y/\log(p-1) > j) + 4p \\ &\le 2E(Y)/\log(p-1) + 4p + 2, \end{aligned}$$

where $Y = \log_+(h(0)X)$.

The first inequality is sharpest for the following choice of p : p is the solution of $(p-1)\log^2(p-1) = \lambda/2$. Because we want $p > 2$, and because we want a good choice of p for large values of λ, we can use instead the formula

$$p = 2 + \frac{\lambda}{2\log^2(1+\lambda)},$$

obtained by functional iteration. Resubstitution yields the desired result.

REMARK: We can push our technique to the limit, and show that the existence of $E(Y/\log(1+Y))$ implies the finiteness of $E(N)$: in fact, there exists a universal constant A such that

$$\begin{aligned} E(N) &\le A + E\left(\frac{Y}{\log(1+Y)}\left(1 + 6\frac{\log(1+\log(1+Y))}{\log(1+Y)}\right)\right) \\ &\le A + 7E\left(\frac{Y}{\log(1+Y)}\right) \end{aligned}$$

where $Y = \log_+(h(0)X)$ (this will be proved below). The following example shows that the latter bound can be of some use: choose $f(x)$ with a decreasing hazard rate such that $f(x)$ is asymptotic to $(x\log^2 x \log\log x)^{-1}$ as $x \to \infty$. Then, assuming that $h(0) < \infty$, we see that $Y = \log_+(h(0)X)$ has a density that decreases as constant divided by $y^2 \log y$ as $y \to \infty$. Clearly, $E(Y) = \infty$ (so that Theorem 5 is useless), but $E(Y/\log(1+Y)) < \infty$.

For the proof of this inequality, we use Lemma 1 once more, with $E(N) \le \sum_{n=0}^{\infty} P(N > n)$, and constants $p_n = n/\log^3(n+1)$, $m_n = \text{int}(n/\log(n+1))$ and $x_n = h(0)^{-1}\exp(n\log n - 4n\log\log(n+1))$. Let n_0 be so large that all three constants fall in the desired ranges, and note that $E(N) \le n_0 + \sum_{n=n_0}^{\infty}$ (Upper bound). The terms $(1 - 1/p_n)^{m_n} \le \exp(-m_n/p_n) \le \exp(-\log^2(n+1) + 1/n\log^3(n+1))$ sum to a constant. The terms $h(0)x_n/p_n^{n-m_n} = \exp(n\log n - 4n\log\log(n+1) - (n-m_n)(\log n - 3\log\log(n+1))) \le \exp(-n\log\log(n+1) + m_n\log n) \le \exp(n -$

$n \log \log(n + 1))$ sum to a constant too. Thus, we are done if for all n at least equal to n_0,

$$\frac{y_n}{\log(1 + y_n)} + 6\frac{y_n \log(1 + \log(1 + y_n))}{\log^2(1 + y_n)} \quad (y_n = \log_+(h(0)x_n)) > n.$$

(This follows from the fact that this expression is an increasing function of y_n for $y_n > 0$.) It is a straightforward but tedious exercise to verify that as $n \to \infty$, the given function of y_n minus n varies as $n \log \log n / \log n$ as $n \to \infty$. This concludes the proof, since we can take n_0 large enough, and independent of the distribution of X.

REFERENCES

[1] Barlow, R. E. and Proschan, F., *Mathematical Theory of Reliability,* Wiley, New York, 1965.

[2] Barlow, R. E., Marshall, A. W., and Proschan, F., "Properties of Probability Distributions with Monotone Hazard Rate," *Annals of Mathematical Statistics,* **34,** 375–389 (1963).

[3] Cinlar, E., *Introduction to Stochastic Processes,* Prentice-Hall, Englewood Cliffs, NJ, 1975.

[4] Devroye, L., "Random Variate Generation for Unimodal and Monotone Densities," *Computing,* **32,** 43–68 (1984).

[5] Gaver, D. P., "Analytical Hazard Representations for use in Reliability, Mortality and Simulation Studies," *Communications in Statistics,* **B8,** 91–111 (1979).

[6] Kaminsky, F. C., and Rumpf, D. L., "Simulating Nonstationary Poisson Processes: A Comparison of Alternatives Including the Correct Approach," *Simulation,* **28,** 17–20 (1977).

[7] Lewis, P. A. W., and Shedler, G. S., "Simulation of Non-Homogeneous Poisson Processes with Log-Linear Rate Function," *Biometrika,* **63,** 501–505 (1976).

[8] Lewis, P. A. W., and Shedler, G. S., "Simulation of Nonhomogeneous Poisson Processes by Thinning," *Naval Research Logistics Quarterly,* **26,** 403–413 (1979a).

[9] Lewis, P. A. W., and Shedler, G. S., "Simulation of Non-Homogeneous Poisson Processes with Degree-Two Exponential Polynomial Rate Function," *Operations Research,* **27,** 1026–1041 (1979b).