# NOTE ON A STOCHASTIC RECURSION

DAVID SIEGMUND[1]

*Stanford University*

The method of Yakir and Pollak (1998) is applied heuristically to a stochastic recursion studied by Goldie (1991). An alternative derivation of the Goldie's tail approximation with a new representation for the constant and some related results are derived.

*AMS subject classifications:* 60H25, 62L10.
*Keywords and phrases:* ARCH process, change-point, tail approximation.

## 1 Introduction

The stochastic recursion

$$R_n = Q_n + M_n R_{n-1}, \tag{1.1}$$

has been studied by a number of authors. See, for example, Kesten (1973) and Goldie (1991), who obtained an expression for the tail behavior of the stationary distribution of (1.1), and de Haan, Resnick, Rootzén and de Vries (1989), who as an application of Kesten's result obtained *inter alia* the asymptotic distribution of $\max(R_1, \cdots, R_m)$. In these studies it is assumed that $(M_1, Q_1), (M_2, Q_2), \cdots$ are independent, identically distributed and satisfy

$$P\{M_n > 0\} = 1, \;\; E(\log M_n) < 0, \;\; P\{M_n > 1\} > 0 \tag{1.2}$$

along with other technical conditions. One motive for studying (1.1) is to obtain information about the ARCH(1) process, which has been proposed as a model for financial time series. It is defined by the recursion $X_n = \{\mu + \lambda X_{n-1}^2\}^{1/2}\epsilon_n$, where $\epsilon_1, \epsilon_2, \cdots$ are independent standard normal random variables. The process $X_n^2$ is a special case of (1.1) with $Q_n = \mu\epsilon_n^2$, $M_n = \lambda\epsilon_n^2$. See Embrechts, Klüppelberg and Mikosch (1997) for an excellent introduction to these and related ideas, and their applications. The special case of (1.1) having $Q_n = 1$ and $E(M_n) = 1$ has also been studied in numerous papers involving change-point detection, e.g., Shiryayev (1963), Pollak (1985, 1987).

[1] This research was supported by The National Science Foundation.

The process $R_n$ defined by (1.1) behaves similarly to the solution of the recursion

$$\log(R_n) = [\log(R_{n-1}) + \log(M_n)]^+, \tag{1.3}$$

which plays an important role in queueing theory and in change-point detection. The purpose of this note is to indicate the potential of a method motivated by change-point analysis (Yakir and Pollak, 1998; Siegmund and Yakir, 1999a, 1999b) and applied to processes similar to (1.3) to give insight into the results of Goldie and of de Haan, *et al.* The calculations are heuristic; rigorous justification appears to be a substantial undertaking. A by-product of this approach is a different and possibly more satisfactory expression for the constant $C$ in equation (2.2) below (cf. (3.6)).

## 2 The Kesten/Goldie approximation

Let $R$ denote the stationary solution of (1.1). For precise conditions under which this solution exists, see Vervaat (1979). Under the conditions (1.2), if $M_n$ has (positive) moments of all orders there will by convexity exist a unique $\theta > 0$ such that

$$E(M_n^\theta) = 1. \tag{2.1}$$

We assume that such a $\theta$ exists and that $I = \theta E[M_n^\theta \log(M_n)]$ is well defined and finite. The parameter I is the Kullback-Leibler information for testing the original distribution of $Y_n = \log(M_n)$ against the alternative having relative density $\exp(\theta Y_n) = M_n^\theta$ (cf. (2.1)). Kesten (1973) and Goldie (1991) showed that

$$P\{R > x\} \sim Cx^{-\theta}. \tag{2.2}$$

Although Kesten considered the more general case of a vector recursion, he did not characterize $C$. In the case of integer $\theta$ Goldie gave the constant $C$ explicitly in terms of mixed integer moments of $(M_n, Q_n)$. In general he characterized $C$ in terms of the distribution of $R$ itself. This characterization does not appear to be useful for evaluating $C$ in the case of non-integral $\theta$.

Building on earlier research of Cramér, Wald and others, Feller (1972) showed how a number of results in queueing and insurance risk theory could be elegantly derived by an application of the renewal theorem to an "associated" distribution. Kesten's and Goldie's methods of proof involve clever extensions of this idea along with substantial analysis. Goldie's associated distribution will appear in the calculation given below; but the motivation behind it is entirely different, and the renewal theorem has been replaced by a simple local limit theorem.

## 3 Alternative, heuristic derivation of (2.2)

Let $P$ denote the measure of Sections 1 and 2. Let $S_j = \Sigma_1^j Y_i$ and for $j \leq n$, put $\ell_{j,n} = \theta(S_n - S_j)$. Finally let the probability measure $P_{j,n}$ be defined by

$$dP_{j,n}/dP = \exp(\ell_{j,n}). \tag{3.1}$$

The change-point interpretation mentioned above is that under $P_{j,n}$ the random variables $Y_1, \cdots, Y_n$ are independent and have distribution in an exponential family with natural parameter $\xi = 0$ for $i = 1, \cdots, j$ and $\xi = \theta$ for $i = j+1, \cdots, n$.

To simplify the notation when there is no risk of confusion, I drop the subscript $n$ and write more concisely $\ell_j$ and $P_j$. It will also be convenient to let $P_j$ denote the extended measure under which the $Y_i$ are independent for all $-\infty < i < \infty$ and have distribution with parameter $\xi = 0$ for $i \leq j$ and $\xi = \theta$ for $i > j$.

Putting $R_{-1} = 0$, one obtains from the recursion (1.1) that

$$R_n = \Sigma_0^n Q_j \exp(S_n - S_j). \tag{3.2}$$

If $R_0 = Q_0$ has the stationary distribution of $R$, then $R_n$ also has this stationary distribution. Let $a = \theta \log(x)$. From (3.1) and (3.2) follows the identity

(3.3)
$$P\{R_n > x\} = \Sigma_j E[e^{\ell_j}/\Sigma_i e^{\ell_i}; \log(R_n^\theta) > a] = \Sigma_j E_j[1/\Sigma_i e^{\ell_i}; \log(R_n^\theta) > a].$$

Here the summation nominally extends over all $i$ and $j$ less than or equal to $n$. A rigorous proof would require showing that it suffices to sum over smaller subsets of these indices. For the moment it suffices to consider $i$ and $j$ such that $n-j$ and $n-i$ belong to the interval $[a/I - \epsilon a,\ a/I + \epsilon a]$ for a suitable small positive $\epsilon$; additional restrictions will be introduced below.

By straightforward algebra the term indexed by $j$ on the right hand side of (3.3) can be rewritten as

$$\begin{aligned} e^{-a} E_j\{ &\frac{[\Sigma_i Q_i \exp(S_j - S_i)]^\theta}{[\Sigma_i \exp[\theta(S_j - S_i)]]} \\ &\exp\{-[\ell_j - a + \theta \log(\Sigma_i Q_i \exp(S_j - S_i)]\}; [\cdots] > 0\}, \end{aligned} \tag{3.4}$$

where $[\cdots] > 0$ indicates that the expectation is taken over the event where the immediately preceding bracketed quantity is positive.

Under $P_j$ the random walks $S_j - S_i$ have negative drift both for $i > j$ and for $i < j$. (This is clear without calculation from the change-point interpretation, since this sum is proportional to the log likelihood ratio for testing that the change-point is at $j$ against the alternative that it is at $i$; and $j$ is the true change-point under $P_j$.) Hence with overwhelming probability the exponential of these sums is close to 0 unless $i$ is close to $j$, say $|i-j| < c\log(a)$. Also, $\ell_j = \theta(S_n - S_j)$ is the sum of approximately $a/I$ terms; and $a$ is assumed to be large. This means that the expressions involving $S_j - S_i$ and that involving $\ell_j$ are asymptotically independent, so (3.4)

$$(3.5) \qquad \sim e^{-a} E_j\{\frac{[\Sigma_i Q_i \exp(S_j - S_i)]^\theta}{[\Sigma_i \exp[\theta(S_j - S_i)]]}\} E_j\{\exp[-(\ell_j - a)]; \ell_j > a\}.$$

For $j$ in the critical interval satisfying $n - j \in [a/I - \epsilon a, a/I + \epsilon a]$, under the assumption that the $P_j$-distribution of $Y_{j+1}$ is sufficiently smooth for a local central limit theorem to apply, the second expectation in (3.5) is asymptotic to

$$\varphi[(a-(n-j)I)/\sigma(n-j)^{1/2}]/\sigma(n-j)^{1/2},$$

where $\varphi$ denotes the standard normal density and $\sigma^2 = \theta^2 \mathrm{Var}_j(Y_{j+1})$. The first expectation in (3.5) is easily seen to be asymptotically independent of $j$, so the sum over $j$ in (3.3) involves only the second expectation and can be evaluated approximately to obtain

$$(3.6) \qquad P\{R > x\} \sim I^{-1} x^{-\theta} E_j\{\frac{[\Sigma_i Q_i \exp(S_j - S_i)]^\theta}{[\Sigma_i \exp[\theta(S_j - S_i)]]}\}.$$

In the asymptotic (large $a$) limit the index $i$, which was previously constrained to satisfy $|i - j| < c\log a$, now extends over all integers, and $P_j$ now denotes the probability under which the independent random variables $Y_i$ have distribution with parameter $\xi = 0$ for $-\infty < i \le j$ and parameter $\xi = \theta$ for $i > j$.

This is Goldie's approximation (2.2) with a new representation for the constant $C$. Unlike Goldie's form of this constant, which involves the distribution of $R$, the form given in (3.6) involves only the distribution of $(M_1, Q_1)$, albeit in a complicated way. Evaluation of this constant is considered in the following section.

## 4 Evaluation of the expectation in (3.6)

In general the expectation in (3.6) seems difficult to evaluate explicitly. However, the form given is quite suitable for simulation, unlike the expression given by Goldie (1991), which involves the distribution of $R$ itself. See Yakir and Pollak (1998) for a related numerical example. For the special case of integral $\theta$ Goldie gives an explicit evaluation. The expectation in (3.6) can be rewritten and then shown to equal Goldie's expression, although the details are messy when $\theta$ is large. The first part of this calculation comes from Siegmund and Yakir (1999a) and does not require that $\theta$ be an integer.

Let $m$ be an arbitrary positive integer. The expectation in (3.6) can be expressed as the limit as $m \to \infty$ of the same expression with the range of $i$ restricted to $1, \cdots, m$ and $j$ any integer which satisfies $j \to \infty$ and $m - j \to \infty$. Hence this expectation also equals the limit of

$$E_j\{\frac{[\Sigma_i Q_i \exp(S_m - S_i)]^\theta}{[\Sigma_i \exp[\theta(S_m - S_i)]]}\}. \tag{4.1}$$

Since the limit of (4.1) is the same uniformly in $j$ (provided $j$ is far from 1 and from $m$), it also equals in the limit the average over $j$ of these expectations, *viz.*

$$m^{-1}\Sigma_{j=1}^m E_j\{\frac{[\Sigma_i Q_i \exp(S_m - S_i)]^\theta}{[\Sigma_i \exp[\theta(S_m - S_i)]]}\}. \tag{4.2}$$

Recalling that $P_{j,m}$ is defined by the likelihood ratio $dP_{j,m}/dP = \exp[\theta(S_m - S_j)]$, one sees that (4.2) equals

$$m^{-1}E\{[\Sigma_i Q_i \exp(S_m - S_i)]^\theta\}. \tag{4.3}$$

For the special case $\theta = 1$, (4.3) equals $E(Q_1)$, so the constant multiplying $x^{-\theta}$ in (3.6) is of the form given by Goldie (1991). For the special case $\theta = 2$, (4.3) equals

$$\begin{aligned} m^{-1}\Sigma_{i=1}^m & E[Q_i^2 \exp\{2(S_m - S_i)\} \\ &+ m^{-1}2\Sigma_{i=1}^m \Sigma_{k=1}^{i-1} E[Q_i Q_k \exp\{2(S_m - S_i) + S_i - S_k\}] \\ &= EQ_1^2 + m^{-1}2\Sigma_{i=1}^m \Sigma_{k=1}^{i-1} EQ_k E(Q_i M_i)[E(M_1)]^{i-k-1} \\ &\to EQ_1^2 + 2E(Q_1)E(Q_1 M_1)/[1 - EM_1], \end{aligned}$$

which again is of the form given by Goldie (1991). Moreover, one sees that with some effort similar expansions can be obtained for arbitrary integral values of $\theta$.

## 5 Distribution of $\max(R_1 \cdots R_m)$

From the tail approximation (2.2) one can use any of several methods to obtain the approximate distribution of $\max(R_1 \cdots R_m)$ (e.g., de Haan *et al.* (1989) or Woodroofe (1976)). It may be of interest to see briefly how the present method would deal with this problem–without requiring prior knowledge of (2.2). Assume that $mE[M_1^\theta \log(M_1)]/\log(x) \to \infty$ and $mx^{-\theta} \to 0$. This puts $P\{\max(R_1 \cdots R_m) > x\}$ into the domain of large deviations. A Poisson approximation can be derived by an auxiliary argument. In terms of the probabilities $P_{j,n}$ defined in (3.1) the argument leading to the equations (3.3)-(3.4) now yields

$$P\{\max(R_1 \cdots R_m) > x\}$$

$$= e^{-a}\Sigma_{n=1}^m \Sigma_j E_{j,n}\{\frac{\exp[\max_{n'} \theta(S_{n'} - S_n)][\Sigma_{j'} Q_{j'} \exp(S_j - S_{j'})]^\theta}{[\Sigma_{n'}\Sigma_{j'} \exp[\theta(S_j - S_{j'} + S_{n'} - S_n)]]}$$

$$(5.1)\quad \times \exp\{-[\ell_{n,j} - a + \theta \log(\max_{n'} \Sigma_{j'} Q_{j'} \exp(S_j - S_{j'} + S_{n'} - S_n))]\}; [\cdots] > 0\}.$$

The observation that the important values of $n'$ are close to $n$ (say to within $c \log a$) and the important values of $j'$ are close to $j$ combined with similar asymptotic analysis to that given above yields the asymptotic approximation

$$(5.2)\quad \begin{aligned} &\sim mI^{-1}e^{-a}E_{-\infty,n}\{\frac{\exp[\max_{n'} \theta(S_{n'} - S_n)]}{\Sigma_{n'} \exp[\theta(S_{n'} - S_n)]}\} \\ &\qquad \times E_{j,+\infty}\{\frac{[\Sigma_{j'} Q_{j'} \exp(S_j - S_{j'})]^\theta}{[\Sigma_{j'} \exp[\theta(S_j - S_{j'})]]}\}. \end{aligned}$$

Under the probability $P_{-\infty,n}$ the independent $Y_i$ have parameter $\xi = \theta$ for $i \le n$ and $\xi = 0$ for $i > n$; under $P_{j,+\infty}$ they have parameter $\xi = 0$ for $i \le j$ and $\xi = \theta$ for $i > j$. The first expectation on the right hand side of (5.2) is in the form obtained by Yakir and Pollak (1998) and Siegmund and Yakir (1999a,b). The second is the same as that obtained above. Since the first expectation does not involve the $Q_i$, one can use the argument of Siegmund and Yakir (1999a) to infer its equivalence to the corresponding expression obtained by de Haan *et al.* (1989) or to convert it into one of the equivalent expressions given by Siegmund (1985), which are more suitable for numerical computation.

## 6 Discussion

The preceding calculations indicate how one might study the stochastic recursion (1.1) via the changes of measure indicated in (3.3) and (5.1). Note that this change of measure does not make use of the linear ordering of the indexing set, and hence is particularly useful for problems involving multidimensional time (e.g., Siegmund and Yakir (1999a,b)).

Although the ARCH(1) process does not itself satisfy (1.1), the marginal tail probability of its stationary distribution is easily inferred from (2.2): one simply replaces $x$ by $x^2$ and $C$ by $C/2$. However, (5.2) requires an auxiliary argument to produce an approximation for the maximum of an ARCH(1) process. This argument is straightforward, but it seems intrinsically one dimensional; and the methods described above do not seem helpful. Let $T = \min\{n : X_n > x\}$. Let $T_0 = \min\{n : R_n > x^2\}$, and for $k = 1, 2, \cdots$ let $T_k = \min\{n : n > T_{k-1}, R_n > x^2\}$. Also let $\nu = \min\{k : \epsilon_{T_k} > 0\}$. From the representation $T = T_\nu$ and the approximation (5.2) one can derive, for example by the method of Woodroofe (1976), a tail probability approximation for $\max(X_1, \cdots, X_m)$. Except for some details of the calculation, this is closely related to the argument of de Haan *et al.* (1989). It leads to still a third constant, which is similar to the first expectation appearing in (5.2) in the sense that it is a functional of a random walk with increments $Y_i$. More precisely, under the conditions of the preceding section one obtains

$$P\{T < m\} \sim mI^{-1}x^{-2\theta}C_1C_2,$$

where $C_1$ is the product of the two expectations on the right hand of (5.2) and

$$C_2 = 2\int_0^\infty e^{-\theta x}[1 - E2^{-N_x}]dx,$$

with $N_x = \Sigma_0^\infty 1\{S_n > -x\}$.

The constant $C_2$ can be calculated by simulation or possibly by repeated numerical integration as follows. Let $u(x) = E(2^{-N_x})$ and $h(x) = 1/2^{1\{x>0\}}$. Also let $Q$ denote the operator defined by $Qf(x) = Ef(x + Y_1)$. Then $u$ satisfies $u = h\ Qu$ and can be obtained recursively as $\lim_{n\to\infty} u_n$, where $u_0 = h \in (u, 1]$ and $u_n = h\ Qu_{n-1}$.

**Acknowledgements** I would like to thank Benny Yakir for several helpful discussions and The University of Cambridge for their hospitality.

## REFERENCES

Embrechts, Klüppelberg, C. and Mikosch, T. (1997). *Modeling Extremal Events for Insurance and Finance*, Springer-Verlag, Berlin.

Feller, W. (1972). *An Introduction to Probability Theory and Its Applications, Vol. II*, John Wiley and Sons, New York.

Goldie, C.M. (1991). Implicit renewal theory and tails of solutions of random equations, *Ann. Appl. Probab.* **1**, 126-166.

Haan, L. de, Resnick. S.I., Rootzén, H. and Vries, C.G. de (1989). Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes, *Stoch. Proc. Appl.* **32**, 213-224.

Kesten, H. (1973) Random difference equations and renewal theory for products of random matrices, *Acta Math.* **131**, 207-248.

Pollak, M. (1985). Optimal detection of a change in distribution, *Ann. Statist.* **13** 206-227.

Pollak, M. (1987). Average run lengths of an optimal method of detecting a change in distribution, *Ann. Statist.* **15** 749-779.

Shiryayev, A. (1963). On optimal methods in quickest detection problems, *Theory Probab. Appl.* **8** 22-46.

Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals*, Springer-Verlag, New York.

Siegmund, D. and Yakir, B. (1999a). Tail probabilities for the null distribution of scanning statistics, *Bernoulli*, to appear.

Siegmund, D. and Yakir, B. (1999b). Approximate p-values for sequence alignments. Unpublished manuscript.

Vervaat, W. (1979). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables, *Adv. Appl. Probab.* **11** 750-783.

Yakir, B. and Pollak, M. (1998). A new representation for a renewal-theoretic constant appearing in asymptotic approximations of large deviations, *Ann. Appl. Probab.* **8** 749-774.

Woodroofe, M. (1976). Frequentist properties of Bayesian sequential tests, *Biometrika* **63**, 101–110.

DEPARTMENT OF STATISTICS
SEQUOIA HALL
STANFORD UNIVERSITY
790 SERRA ST.
STANFORD, CA 94305
USA
*dos@stat.stanford.edu*