Probability in the Engineering and Informational Sciences, **20**, 2006, 413–427. Printed in the U.S.A.

ON THE OUTCOME OF A CASCADING FAILURE MODEL

Claude Lefèvre
Institut de Statistique et de Recherche Opérationnelle
Université Libre de Bruxelles
B-1050 Bruxelles, Belgique
E-mail: clefevre@ulb.ac.be

This article is concerned with a loading-dependent model of cascading failure proposed recently by Dobson, Carreras, and Newman [6]. The central problem is to determine the distribution of the total number of initial components that will have finally failed. A new approach based on a closed connection with epidemic modeling is developed. This allows us to consider a more general failure model in which the additional loads caused by successive failures are arbitrarily fixed (instead of being constant as in [6]). The key mathematical tool is provided by the partial joint distributions of order statistics for a sample of independent uniform (0,1) random variables.

1. INTRODUCTION

In a recent article, Dobson, Carreras, and Newman [6] have proposed an interesting model to describe the occurrence of cascading failures in a closed system of n components. The basic points of the model can be summarized as follows. Initially, the n components have random loads $L_1, \ldots, L_n$ that correspond to independent uniform (0,1) random variables (r.v.'s). Following some disturbance, a load d is added to each of the n loads. Any component fails if its new load is larger than one. When a components fails, a new load p is added to each of the components that are still functioning. This can cause further failures in a cascade. The central problem is to determine the distribution of a r.v. N that represents the total number of components that will have finally failed. This loading-dependent model of cascading failure is motivated by the study of large blackouts of electric power transmission systems (see also [6] for a bibliography).

 0269-9648/06 $16.00

https://doi.org/10.1017/S0269964806060256 Published online by Cambridge University Press

Section 2 is mainly concerned with the same failure model. First, we reformulate this model by means of a cascade algorithm, different but equivalent to the original one, that allows us to deal with the occurrences of failures one by one. Moreover, it is convenient to introduce the r.v.'s $U_1 = 1 - L_1, \ldots, U_n = 1 - L_n$, which can be interpreted as the initial resistances of the n components. Obviously, these r.v.'s are still independent uniform (0,1); let $U_{1:n}, \ldots, U_{n:n}$ be the associated order statistics. We then show that the distribution of N can be expressed in terms of the joint distributions of the order statistics $\{U_{i:n}, 1 \leq i \leq k\}$ for $1 \leq k \leq n$, with respect to the linear boundary $\{d + (i-1)p, 1 \leq i \leq k\}$. As a corollary, we are able to derive, in an enlightening way, a key result obtained in [6], namely in case of nonsaturation N has a quasi-binomial distribution. Furthermore, we prove how this result can be extended to a model for which the additional loads after failures are independent and identically distributed (i.i.d.) r.v.'s.

In Section 3 we discuss a generalized failure model in which the additional loads are still constant but arbitrarily fixed (instead of being equal to each other). In particular, this extension covers those cases in which the initial resistances are i.i.d. r.v.'s with any given (not necessarily uniform) distribution. For this model, we can now write the distribution of N in terms of the joint distributions of the order statistics $\{U_{i:n}, 1 \leq i \leq k\}$ for $1 \leq k \leq n$, with respect to an arbitrary (instead of linear) nondecreasing boundary $\{s_i, 1 \leq i \leq k\}$. We then show that the state probabilities of N still have a remarkable structure and that they can be calculated using a simple recursion. As special situations, we examine in more detail a model built with several redundant components (so that the first failures will not generate any additional load) and a model for which the resistances are exponentially distributed and the additional loads are independent r.v.'s. Finally, some numerical illustrations and simple bounds are presented in Section 4.

The approach that we will follow to tackle this cascading failure is inspired by an argument used in epidemic modeling to determine the distribution of the final number of new infections for epidemics of the S (susceptible) $\rightarrow$ I (infected) $\rightarrow$ R (removed) schema (see, e.g., Ball and O'Neill [2], Lefèvre and Utev [8], and Picard and Lefèvre [9]). Very roughly, in an epidemic context, the n components represent the n initial susceptibles and the failure of a component corresponds to the infection of a susceptible (by one of the infectives present at that time). The additional load in case of failure means that once infected, an individual adds its infectiousness to the other infectives against the remaining susceptibles. An indirect contribution of our article is thus to point out the existence of a closed connection, within this framework, between reliability and epidemic modelings.

Applying the proposed method asks us to evaluate the partial joint distributions of order statistics for a sample of n independent uniform (0,1) r.v.'s. Much on this problem can be found in Shorack and Wellner [10, Chap. 9]; see also Denuit, Lefèvre, and Picard [5] and the references therein. In [5], it is established that the left tail distributions such as derived here rely on an underlying polynomial structure of the Abel–Gontcharoff form. This property, however, is not exploited in the present work, at least not in a direct or explicit way.

https://doi.org/10.1017/S0269964806060256 Published online by Cambridge University Press

2. THE BASIC MODEL

Consider a system made of n identical components. Initially, the n components are unfailed and have *random loads* $L_1, \dots, L_n$ that are independent and uniformly distributed on $(0,1)$. The cascade proceeds according to the following algorithm (Dobson et al. [6]):

Step 0

- Load increment. A disturbance d is added to the load of each of the n components.
- Failure test. For each component j, $1 \le j \le n$, if the load L_j is greater than one, j fails. Let M_0 be the number of failures.

Step 1

- Load increment. If there are failures, a disturbance $M_0 p$ is added to the load of each of the $n - M_0$ remaining components.
- Failure test. For each of these components j, if the new load $d + M_0 p + L_j$ is greater than one, j fails. Let M_1 be the number of failures.

Step 2

- Load increment. A disturbance $M_1 p$ is added to the load of each of the $n - M_0 - M_1$ remaining components.
- Failure test. For each of these components j, if the new load $d + (M_0 + M_1)p + L_j$ is greater than one, j fails. Let M_2 be the number of failures.

The next steps are similar. Failures can occur until time T, when there are no more new failures. Indeed, when $M_T = 0$, there is no new load added to the remaining components, so that the cascade process stops (the system is stabilized). The total number of failures is $N' = M_0 + M_1 + \cdots + M_{T-1}$, the number of unfailed components being $n - N'$.

An equivalent cascade algorithm. We now construct another algorithm that leads to the same total number of failures in the system. As indicated in Section 1, a similar approach is used in epidemic modeling to determine the final size distribution of an epidemic of the SIR schema. Let us introduce the order statistics $L_{1:n}, \dots, L_{n:n}$, which represent the initial loads of the n components arranged by increasing weights.

Step 0

- Load increment. A disturbance d is added to the load of each of the n components.
- Failure test. If the largest load $d + L_{n:n}$ is greater than one, the corresponding component fails.

https://doi.org/10.1017/S0269964806060256 Published online by Cambridge University Press

Step 1

- Load increment. If this occurs, a disturbance p is added to the load of each of the $n-1$ remaining components.
- Failure test. If the new load $d+p+L_{n-1:n}$ is greater than one, the corresponding component fails.

Step 2

- Load increment. A disturbance p is added to the load of each of the $n-2$ remaining components.
- Failure test. If the new load $d+2p+L_{n-2:n}$ is greater than one, the corresponding component fails.

Here too, failures stop when all of the loads of the remaining components are smaller than one. Let N be the associated total number of failures.

It can be easily understood that as long as the total number of failures is considered, both algorithms are quite equivalent; that is, N' and N have the same distribution. For instance, in the former formulation any given unfailed component at step 1 receives an additional load $M_0 p$, whereas in the latter formulation, that component receives the same load in several steps. This change of timescale, however, will not affect the final possible failure of the component.

The new version of the algorithm allows us to express the distribution of N in a simple way, using the key tool of order statistics. Indeed, we directly see that

$$P(N \geq k) = P[L_{n:n} > 1-d, L_{n-1:n} > 1-d-p, L_{n-2:n} > 1-d-2p, \ldots, L_{n-k+1:n} > 1-d-(k-1)p], \qquad k=1,\ldots,n. \tag{2.1}$$

In (2.1), it was implicitly assumed that $1-d-(n-1)p > 0$. If this is not true, define $j^* = \min\{j : 1-d-jp \leq 0\}$. Obviously, $P(N \geq k)$ is still given by (2.1) when $k=1,\ldots,j^*$, whereas $P(N \geq k) = P(N \geq j^*)$ when $k = j^*+1,\ldots,n$, meaning that the occurrence of j^* failures necessarily yields n failures. Such a consequence of cascading is called a saturation effect in [6]. To simplify the presentation, we will continue with (2.1) without saturation.

Now let us introduce the r.v.'s $U_1 = 1-L_1,\ldots,U_n = 1-L_n$. Obviously, these U_i's are still independent and uniformly distributed on $(0,1)$. By construction of the model, they can be interpreted as the initial *random resistances* of the n components. Denote by $U_{1:n},\ldots,U_{n:n}$ the associated order statistics (i.e., the n initial resistances arranged by increasing strengths). Then we can rewrite (2.1) as

$$P(N \geq k) = P[U_{1:n} < d, U_{2:n} < d+p, U_{3:n} < d+2p, \ldots, U_{k:n} < d+(k-1)p], \qquad k=1,\ldots,n. \tag{2.2}$$

From (2.2), we then get

$$P(N = k) = P[U_{1:n} < d, U_{2:n} < d + p, U_{3:n} < d + 2p, \ldots, U_{k:n} < d + (k-1)p, U_{k+1:n} > d + kp], \quad k = 0, \ldots, n, \quad \textbf{(2.3)}$$

where we put $U_{n+1:n} = 1$, say.

In order to evaluate the probability (2.3), we begin by deriving the preliminary result (2.4) below. That formula is well known (see, e.g., [5]); the simple proof given here will be adapted later to more general situations. For any *fixed size* k ($1 \leq k \leq n$), consider a sample $U_1, \ldots, U_k$ of independent and uniform (0,1) r.v.'s, and denote by $U_{1:k}, \ldots, U_{k:k}$ the associated order statistics. Let x be any real in $(0, d)$.

LEMMA 2.1:

$$P[x < U_{1:k} < d, U_{2:k} < d + p, U_{3:k} < d + 2p, \ldots, U_{k:k} < d + (k-1)p] = (d - x)(d + kp - x)^{k-1}, \qquad k = 1, \ldots, n. \quad \textbf{(2.4)}$$

PROOF: We will proceed by induction. Obviously, for $k = 1$, $P[x < U_{1:1} < d] = d - x$. Suppose that (2.4) holds for $k = 1, \ldots, l$ $(\leq n - 1)$. For $k = l + 1$, we start by representing the probability on the left-hand side of (2.4) as follows. Consider the $l + 1$ possible choices U_i, $1 \leq i \leq l + 1$, for the first ordered value $U_{1:l+1}$, and denote by $U_{1:l}^{(i)}, \ldots, U_{l:l}^{(i)}$ the order statistics of the original sample deprived of U_i (which reduces to a sample of size l). The probability in (2.4) can be expressed as

$$\sum_{i=1}^{l+1} P[x < U_i < d, U_i < U_{1:l}^{(i)} < d + p, U_{2:l}^{(i)} < d + 2p, \ldots, U_{l:l}^{(i)} < d + lp]\, dz. \quad \textbf{(2.5)}$$

By conditioning on U_i, we can then rewrite (2.5) as

$$(l + 1) \int_x^d P[z < U_{1:l} < d + p, U_{2:l} < d + 2p, \ldots, U_{l:l} < d + lp]\, dz. \quad \textbf{(2.6)}$$

However, by induction, (2.6) is equivalent to

$$(l + 1) \int_x^d (d + p - z)(d + p + lp - z)^{l-1}\, dz, \quad \textbf{(2.7)}$$

which, after integration, yields the right-hand side of (2.4) for $k = l + 1$. ∎

It now becomes easy to show from (2.3) that N has a quasi-binomial distribution (in the sense of Consul [3]). This result was derived in [6]; as indicated in [6], it was also obtained in other contexts and by different methods. For related problems, see, for example, Stadje [11] and Zacks [12].

PROPOSITION 2.2:

$$P(N=k)=\binom{n}{k}d(d+kp)^{k-1}(1-d-kp)^{n-k}, \qquad k=0,\ldots,n. \tag{2.8}$$

PROOF: Since the n initial components are i.i.d., a number of $N=k$ failures can arise in an analogous fashion among all of the sets of k components, hence the factor $\binom{n}{k}$ in (2.8). Fix any given set of k components as the set having those k failures. Looking at (2.3), we first see that the other $n-k$ components will not fail with the probability $(1-d-kp)^{n-k}$, as indicated in (2.8). Furthermore, when $k \geq 1$, the failure of all of the k components of the set now means that $[U_{1:k}<d, U_{2:k}<d+p,\ldots,U_{k:k}<d+(k-1)p]$ with a sample of size k (instead of n), and this will occur with a probability given by (2.4) evaluated at $x=0$, hence the remaining factor in (2.8). ■

The noncascade situation arises when $p=0$. As expected, (2.8) then reduces to the binomial distribution for a sample size n and a success (i.e., failure) probability d.

A randomized extension. The functioning of the components can be influenced by various factors, internal or external. In such situations, the additional loads imposed to the system are not known with certainty but can be viewed as random elements of the model. So, let us suppose that the initial disturbance corresponds to a r.v. D and that, independently of D, the successive loads after failures correspond to i.i.d. r.v.'s $W_1, W_2, \ldots, W_{n-1}$. Denote the cumulated additional loads by $P_1 = W_1, P_2 = W_1 + W_2, \ldots, P_{n-1} = W_1 + \cdots + W_{n-1}$. To avoid saturation, we assume that $D + P_{n-1} < 1$ a.s. (this hypothesis is satisfied if, for instance, D and the W_i's are bounded by $1/n$). Now (2.8) for the law of N can be extended as follows (see Lefèvre and Picard [7] for a similar result derived in an epidemic context).

PROPOSITION 2.3:

$$P(N=k)=\binom{n}{k}E[D(D+P_k)^{k-1}(1-D-P_k)^{n-k}], \qquad k=0,\ldots,n. \tag{2.9}$$

PROOF: For any given k, let us condition the event $(N=k)$ with respect to the two r.v.'s D and P_k. By adapting the argument followed for (2.8), we then get

$$P(N=k)=\binom{n}{k}E[\theta_k(0)(1-D-P_k)^{n-k}], \qquad k=0,\ldots,n, \tag{2.10}$$

where $\theta_0(0) \equiv 1$ and the other $\theta_k(0)$'s represent the following conditional probabilities evaluated at $x=0$:

https://doi.org/10.1017/S0269964806060256 Published online by Cambridge University Press

$$\theta_k(x) = P[x < U_{1:k} < D, U_{2:k} < D + P_1, \ldots, U_{k:k} < D + P_{k-1} \mid D, P_k], \quad \textbf{(2.11)}$$

x being any real in $(0, D)$. Now we will prove that

$$\theta_k(x) = (D - x)(D + P_k - x)^{k-1}, \quad k = 1, \ldots, n, \quad \textbf{(2.12)}$$

which leads to the announced formula (2.9).

As for Lemma 2.1, we proceed by recurrence and consider (2.11) for $k = l + 1$. First, (2.6) is replaced by

$$\theta_{l+1}(x) = (l+1) \int_x^D P[z < U_{1:l} < D + P_1, U_{2:l} < D + P_2, \ldots, U_{l:l} < D + P_l \mid D, P_{l+1}]\, dz. \quad \textbf{(2.13)}$$

Now, for the term $P[\cdots]$ in (2.13), let us condition on the r.v. P_1, and put $\tilde{P}_1 = W_2, \tilde{P}_2 = W_2 + W_3, \ldots, \tilde{P}_{n-2} = W_2 + \cdots + W_{n-1}$. We obtain that

$$\theta_{l+1}(x) = (l+1) \int_x^D E\{P[z < U_{1:l} < D + P_1, U_{2:l} < D + P_1 + \tilde{P}_1, \ldots, U_{l:l} < D + P_1 + \tilde{P}_{l-1} \mid D, P_1, \tilde{P}_l] \mid D, P_{l+1}\}\, dz. \quad \textbf{(2.14)}$$

However, by the hypothesis of recurrence (for $k = l$) and using $P_1 + \tilde{P}_l = P_{l+1}$, (2.14) then becomes

$$\begin{aligned} \theta_{l+1}(x) &= (l+1) \int_x^D E[(D + P_1 - z)(D + P_1 + \tilde{P}_l - z)^{l-1} \mid D, P_{l+1}]\, dz \\ &= (l+1) \int_x^D [D + E(P_1 \mid P_{l+1}) - z](D + P_{l+1} - z)^{l-1}\, dz. \quad \textbf{(2.15)} \end{aligned}$$

To close, we observe that $E(P_1 \mid P_{l+1}) = P_{l+1}/(l+1)$, so that (2.15) is analogous to (2.7) with $P_{l+1}/(l+1)$ substituted for p, hence (2.12) for $k = l + 1$. ■

3. A GENERALIZED MODEL

In this section we discuss another generalization of the model in [6] by assuming this time that the additional loads in case of failures are still constant but arbitrarily fixed. Such an extension allows us to widen the field of applications of the model.

Thus, a large system is often built, for security reasons, with several redundant components. Then the first failures of components will not influence the functioning conditions applied to the other components. In the model of Section 2, this means that a load p is added to each unfailed component only when there are at least $c + 1$ failures (instead of one), with $0 \leq c \leq n - 1$. The probability (2.2) is then modified as follows: If $k = 1, \ldots, c$,

$$P(N \geq k) = P[U_{k:n} < d], \quad \textbf{(3.1)}$$

and if $k = c + 1, \ldots, n$,

$$P(N \geq k) = P[U_{c+1:n} < d, U_{c+2:n} < d + p, \ldots, U_{k:n} < d + (k - c - 1)p]. \quad \textbf{(3.2)}$$

More generally, let us assume that the additional loads are successively of given weights $w_1, \ldots, w_{n-1}$. Denote the cumulated weights by $p_1 = w_1, \ldots, p_{n-1} = w_1 + \cdots + w_{n-1}$, and to avoid saturation, suppose that $d + p_{n-1} < 1$. Then (2.2) is changed in

$$P(N \geq k) = P[U_{1:n} < d, U_{2:n} < d + p_1, \ldots, U_{k:n} < d + p_{k-1}], \qquad k = 1, \ldots, n. \quad \textbf{(3.3)}$$

Such an extension is also relevant when in (2.2) the initial resistances of the n components are independent but no longer uniformly distributed. More precisely, let us suppose that these resistances are represented by i.i.d. random variables $X_1, \ldots, X_n$, with an arbitrary continuous distribution function F. Let $X_{1:n}, \ldots, X_{n:n}$ be the associated order statistics. The initial disturbance is still denoted by d, and the successive loads are still denoted by $p_1, \ldots, p_{n-1}$; assume that d and $d + p_{n-1}$ are in the support of F. Then, instead of (2.2), we have

$$P(N \geq k) = P[X_{1:n} < d, X_{2:n} < d + p_1, \ldots, X_{k:n} < d + p_{k-1}], \qquad k = 1, \ldots, n. \quad \textbf{(3.4)}$$

However, since the vector $[X_{1:n}, \ldots, X_{n:n}]$ has the same distribution as $[F^{-1}(U_{1:n}), \ldots, F^{-1}(U_{n:n})]$, (3.4) is equivalent to

$$P(N \geq k) = P[U_{1:n} < s_1, \ldots, U_{k:n} < s_k], \qquad k = 1, \ldots, n, \quad \textbf{(3.5)}$$

where we put $s_1 = F(d), s_2 = F(d + p_1), \ldots, s_n = F(d + p_{n-1})$. Of course, (3.5) covers all of the previous cases.

Now let us determine from (3.5) the exact probability of N. To begin, we will extend Lemma 2.1. Specifically, we consider again any fixed sample of size k $(1 \leq k \leq n)$ and define

$$\xi_k(x) = P[x < U_{1:k} < s_1, U_{2:k} < s_2, \ldots, U_{k:k} < s_k], \qquad k = 1, \ldots, n, \quad \textbf{(3.6)}$$

x being any real in $(0, s_1)$, with $\xi_0 \equiv 1$. These ξ_k's are no longer known explicitly but they can be computed recursively from (3.7).

Lemma 3.1: *For any real y,*

$$\xi_k(x) = (y - x)^k - \sum_{j=0}^{k-1} \binom{k}{j} \xi_j(x)(y - s_{1+j})^{k-j}, \qquad k = 1, \ldots, n. \quad \textbf{(3.7)}$$

https://doi.org/10.1017/S0269964806060256 Published online by Cambridge University Press

PROOF: First, we observe that as for (2.6), one can write, for $0 \le l \le n-1$,

$$\xi_{l+1}(x) = (l+1)\int_x^{s_1} P[z < U_{1:l} < s_2, U_{2:l} < s_3, \ldots, U_{l:l} < s_{l+1}]\,dz. \tag{3.8}$$

Now let us proceed by induction and suppose that (3.7) is true for $k = 1, \ldots, l$ ($\le n-1$). For $k = l+1$, applying induction to the probability on the right-hand side of (3.8) yields

$$\xi_{l+1}(x) = (l+1)\int_x^{s_1} (y-z)^l\,dz - (l+1)\sum_{j=0}^{l-1}\binom{l}{j}(y-s_{2+j})^{l-j} \times \int_x^{s_1} P[z < U_{1:j} < s_2, U_{2:j} < s_3, \ldots, U_{j:j} < s_{j+1}]\,dz. \tag{3.9}$$

However, computing the first integral in (3.9) and again using (3.8) for the second integral, we find that (3.9) reduces to

$$\xi_{l+1}(x) = (y-x)^{l+1} - (y-s_1)^{l+1} - \sum_{j=0}^{l-1}\binom{l+1}{j+1}(y-s_{2+j})^{l-j}\xi_{j+1}(x),$$

hence (3.7) for $k = l+1$. ■

The recursion (3.7) can sometimes be slightly simplified by choosing an appropriate value for y. For instance, $y = 1$ will guarantee that the coefficients in the recursion are all positive. Several applications are given in the following.

The distribution (2.8) for N can now be generalized by (3.10) in terms of the probabilities $\xi_k(0)$. The reader is referred to [5] for an alternative proof based on Abel–Gontcharoff polynomials.

PROPOSITION 3.2:

$$P(N = k) = \binom{n}{k}\xi_k(0)(1 - s_{1+k})^{n-k}, \qquad k = 0, \ldots, n. \tag{3.10}$$

PROOF: As for (2.3), we have

$$P(N = k) = P[U_{1:n} < s_1, \ldots, U_{k:n} < s_k, U_{k+1:n} > s_{k+1}], \qquad k = 0, \ldots, n.$$

Thus, it suffices to follow the argument given in the proof of Proposition 2.2 using (3.6) with $x = 0$. ■

Formula (3.10) with the recursion (3.7) can be combined to give the single recursion (3.11).

https://doi.org/10.1017/S0269964806060256 Published online by Cambridge University Press

Corollary 3.3:

$$\binom{n}{k} = \sum_{j=0}^{k} \binom{n-j}{k-j} P(N=j) \frac{1}{(1-s_{1+j})^{n-k}}, \qquad k=0,\ldots,n. \qquad \textbf{(3.11)}$$

Proof: Equation (3.7) with $x = 0$ and $y = 1$ yields

$$1 = \sum_{j=0}^{k} \binom{k}{j} \xi_j(0)(1-s_{1+j})^{k-j}, \qquad k=0,\ldots,n. \qquad \textbf{(3.12)}$$

After multiplication by $\binom{n}{k}$, (3.12) can be rewritten as

$$\binom{n}{k} = \sum_{j=0}^{k} \binom{n-j}{k-j}\binom{n}{j} \xi_j(0)(1-s_{1+j})^{n-j} \frac{1}{(1-s_{1+j})^{n-k}}, \qquad k=0,\ldots,n. \quad \textbf{(3.13)}$$

However, applying (3.10) inside the sum of (3.13) allows us to exhibit $P(N=j)$, which then leads to (3.11). ■

Notice that the relations (3.11) constitute a triangular system of $n+1$ linear equations in the $n+1$ unknown probabilities $P(N=j)$. For instance, (3.11) for $k=0$ gives $P(N=0) = (1-s_1)^n$, and for $k=n$, $P(N=0) + \cdots + P(N=n) = 1$.

Case with redundant components. Let us go back to the previous particular model specified by (3.1), (3.2), with a threshold of $c+1$ failures. Here thus,

$$s_1 = \cdots = s_c = d \quad \text{and} \quad s_k = d + (k-c-1)p, \qquad k = c+1,\ldots,n. \qquad \textbf{(3.14)}$$

Corollary 3.4: *Under (3.14), the law of N is given by (3.10), where, if $k = 1,\ldots,c$,*

$$\xi_k(0) = d^k, \qquad \textbf{(3.15)}$$

and if $k = c+1,\ldots,n$,

$$\xi_k(0) = \sum_{l=0}^{k-c-1} \binom{k}{l} d^{k-l} p^l (k-c)^{l-1}(k-c-l). \qquad \textbf{(3.16)}$$

Proof: Equation (3.15) is straightforward and we will focus on (3.16). For $k = c+1,\ldots,n$, choosing $x=0$ and $y=d$ in (3.7) then yields

$$\xi_k(0) = d^k - \sum_{j=c+1}^{k-1} \binom{k}{j} \xi_j(0)[-p(j-c)]^{k-j}, \qquad k=1,\ldots,n. \qquad \textbf{(3.17)}$$

Proceeding by recurrence, let us suppose that (3.16) is true for $\xi_j(0)$ with $j = c+1,\ldots,k-1$. Then, by substitution in (3.17), and after rearrangement and permutation of the two sums, we obtain

$$\xi_k(0) = d^k - \sum_{j=c+1}^{k-1} \binom{k}{j} (-p)^{k-j} \sum_{l=c+1}^{j} \binom{j}{l} d^l p^{j-l} (j-c)^{k-l-1}(l-c)$$
$$= d^k - \sum_{l=c+1}^{k-1} \binom{k}{l} d^l p^{k-l}(l-c) \sum_{j=0}^{k-l-1} \binom{k-l}{j} (l-c+j)^{k-l-1}(-1)^{k-l-j}. \tag{3.18}$$

However, for any real x,

$$\sum_{j=0}^{m} \binom{m}{j} (x+j)^{m-1}(-1)^{m-j} = 0, \qquad m = 1,2,\ldots, \tag{3.19}$$

since the left-hand side of (3.19) corresponds to $\Delta^m(x^{m-1})$, where Δ denotes the usual forward difference operator. Thus, using (3.19) we see that the second sum in (3.18) reduces to $-(k-c)^{k-l-1}$. Finally, (3.18) becomes

$$\xi_k(0) = \sum_{l=c+1}^{k} \binom{k}{l} d^l p^{k-l}(l-c)(k-c)^{k-l-1},$$

hence (3.16), as indicated. ■

Case with exponential resistances. Let us assume that the resistances have an exponential law, with parameter μ (> 0) say. Thus, $s_1 = 1 - \exp(-\mu d)$, and for $k = 1, \ldots n-1$, $s_{1+k} = 1 - \exp[-\mu(d + p_k)]$, with $p_k = w_1 + \cdots + w_k$. In other words,

$$s_k = 1 - q_0 q_1 \cdots q_{k-1}, \qquad k = 1, \ldots, n, \tag{3.20}$$

where $q_0 \equiv \exp(-\mu d)$ and $q_k \equiv \exp(-\mu w_k)$, $k = 1, \ldots, n-1$. This particular situation is of the type met in epidemic theory with the so-called Reed–Frost model. The homogeneous case where $q_1 = \cdots = q_{n-1} \equiv q$ corresponds to that model in its standard version (see, e.g., Andersson and Britton [1] and Daley and Gani [4]); the nonhomogeneous case with different q_k's gives a nonstationary version of the model (see [7]). Under (3.20), (3.11) becomes

$$\binom{n}{k} = \sum_{j=0}^{k} \binom{n-j}{k-j} P(N=j) \frac{1}{(q_0 \cdots q_j)^{n-k}}, \qquad k = 0, \ldots, n. \tag{3.21}$$

An advantage of this case is that the additional loads imposed on the system can be considered, without difficulty, as random elements. Specifically, let us suppose that the initial disturbance and the successive loads in case of failure correspond to independent r.v.'s, D and $W_1, \ldots, W_{n-1}$, all possibly with different laws. In an epidemic context, $W_1, \ldots, W_n$ are i.i.d. r.v.'s that represent lengths of infectious periods; the model then corresponds to the randomized version of the Reed–Frost model (e.g., [1,4]). Now, for this randomized case, the $n+1$ relations (3.21) can

https://doi.org/10.1017/S0269964806060256 Published online by Cambridge University Press

still be extended into a linear system of similar structure. More precisely, for $j = 0,\dots,n$, define the parameters

$$q_0^{(j)} = E[\exp(-\mu jD)] \quad \text{and} \quad q_k^{(j)} = E[\exp(-\mu jW_k)], \qquad k = 1,\dots,n-1. \quad \textbf{(3.22)}$$

COROLLARY 3.5: *Under randomized (3.20),*

$$\binom{n}{k} = \sum_{j=0}^{k} \binom{n-j}{k-j} P(N=j)\, \frac{1}{q_0^{(n-k)} \cdots q_j^{(n-k)}}, \qquad k = 0,\dots,n. \quad \textbf{(3.23)}$$

PROOF: First, we condition on fixed values d for D and $w_1,\dots,w_{n-1}$ for $W_1,\dots,W_{n-1}$. The conditional state probabilities of N thus satisfy the relations (3.21). We observe that $P(N=j)$, $0 \le j \le n$, depends only on the parameters $q_0,\dots,q_j$. Let us multiply both sides of (3.21) by $(q_0 \cdots q_k)^{n-k}$, and on the right-hand side, replace $(q_0 \cdots q_k)^{n-k}/(q_0 \cdots q_j)^{n-k}$ by $(q_{1+j} \cdots q_k)^{n-k}$. Then, to eliminate the condition, we take the expectation, which yields, using the parameters (3.22),

$$\binom{n}{k} q_0^{(n-k)} \cdots q_k^{(n-k)} = \sum_{j=0}^{k} \binom{n-j}{k-j} P(N=j)\, q_{1+j}^{(n-k)} \cdots q_k^{(n-k)}, \qquad k = 0,\dots,n,$$

hence the system (3.23). ■

4. EXAMPLES AND BOUNDS

For illustration, we consider a special case of nonregular type for the model examined in Section 3. Specifically, (3.5) is assumed to hold now with

$$s_k = [d + p(k-1)]^\theta, \qquad k = 1,\dots,n, \quad \textbf{(4.1)}$$

where $\theta > 0$. This arises when the resistances are i.i.d. r.v.'s with a power-function distribution [i.e., $F(x) = x^\theta$ for $x \in (0,1)$ with $\theta > 0$] and when the initial disturbance is equal to d and each additional load is equal to p, with $d + p(n-1) < 1$. If $\theta = 1$, the model corresponds to the case in Section 2.

We have evaluated the probability law of N, using the recursion (3.10), (3.12), when the system has $n = 50$ components and for different sets of parameters (with nonsaturation). Figure 1 gives this distribution on a logarithmic scale if $d = p = \frac{1}{51}$ and $\theta = 0.5, 0.7, 1$, or 1.1. As expected, the power parameter θ plays a crucial role; in particular, the mode is at $k = 50$ for $\theta = 0.5$ or 0.7 and at $k = 0$ if $\theta = 1$ or 1.1. In Figure 2, the parameters are $\theta = 0.5$ and $d = 0.1p$, $0.75p$, or p, with $p = \frac{1}{51}$. We observe that varying the initial disturbance d moderately affects the distribution; of course, the influence exerted by d depends also on the value of θ.

It is worth pointing out that for n large, the recursion is not always efficient and can even generate negative values for certain probabilities of N.

To close, we indicate that for arbitrary $s_1,\dots,s_n$, the distribution of N can be approximated by finding upper and lower bounds for the probabilities $\xi_k(0)$ in (3.10). Such simple bounds are provided in Property 4.1.

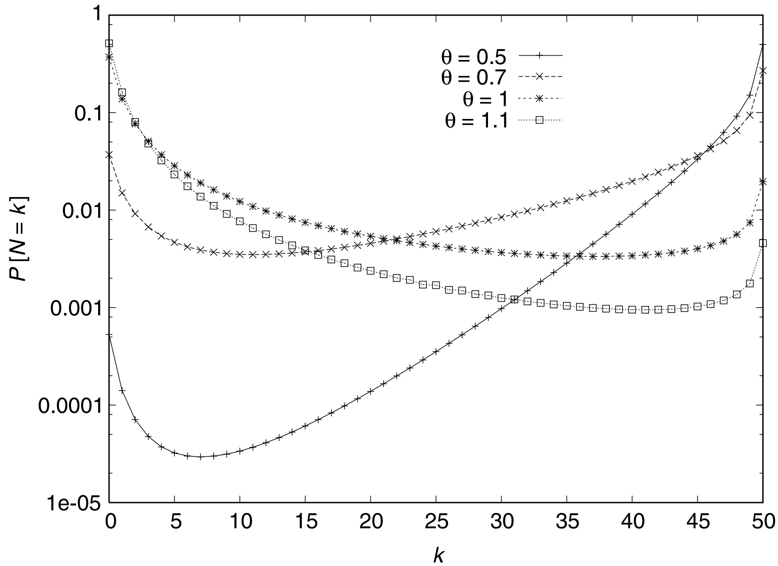

FIGURE 1. Log plot of the distribution of the number N of failed components, under (4.1) with $n = 50$ and $d = p = \frac{1}{51}$ and for four different θ's.

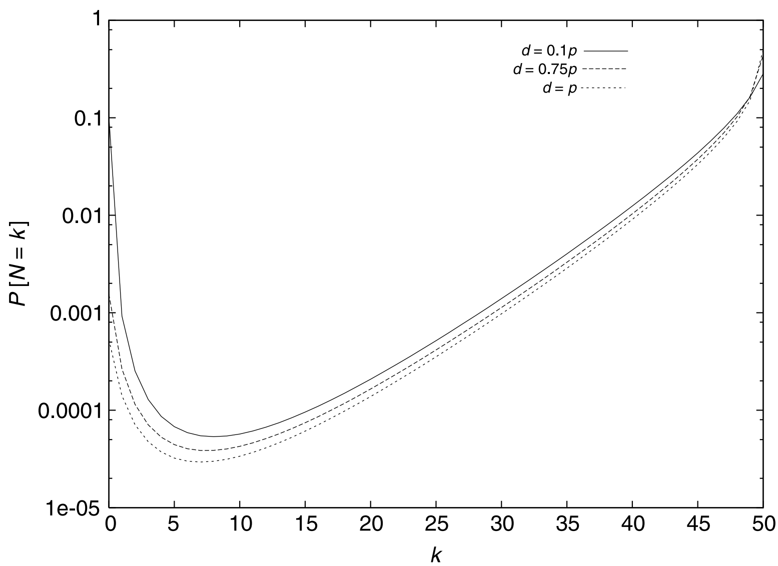

FIGURE 2. Log plot of the distribution of the number N of failed components, under (4.1) with $n = 50$, $\theta = 0.5$, and $p = \frac{1}{51}$ and for three different d's.

https://doi.org/10.1017/S0269964806060256 Published online by Cambridge University Press

PROPERTY 4.1: *In general,*

$$s_1 \cdots s_k \le \xi_k(0) \le s_k^k - (s_k - s_1)^k, \qquad k = 1, \dots, n. \tag{4.2}$$

If the sequence $s_1, \dots, s_k$ *is concave (resp. convex), then*

$$\xi_k(0) \ge (\le) s_1 [(k s_k - s_1)/(k-1)]^{k-1}, \qquad k = 1, \dots, n, \tag{4.3}$$

and for any given $k' \in \{1, \dots, k-1\}$,

$$\xi_k(0) \le (\ge) [s_{k'+1} - k'(s_{k'+1} - s_{k'})] \times [s_{k'+1} + (k - k')(s_{k'+1} - s_{k'})]^{k-1}, \qquad k = 1, \dots, n, \tag{4.4}$$

provided, in the convex case, that $s_{k'+1} - k'(s_{k'+1} - s_{k'}) > 0$.

PROOF: Let us first derive (4.2). By definition (3.6),

$$\xi_k(0) \ge P(U_1 \le s_1, \dots, U_k \le s_k) = s_1 \cdots s_k,$$

whereas choosing $x = 0$ and $y = s_k$ in (3.7) yields

$$\xi_k(0) = s_k^k - \sum_{j=0}^{k-1} \binom{k}{j} \xi_j(0)(s_k - s_{1+j})^{k-j} \le s_k^k - (s_k - s_1)^k.$$

Now if the sequence $s_1, \dots, s_k$ is concave, then a lower linear approximation to this sequence is given by

$$s_i \ge s_i^{(l)} \equiv s_1 + (i-1)(s_k - s_1)/(k-1), \qquad i = 1, \dots, k,$$

and an upper linear approximation is given by

$$s_i \le s_i^{(u)} \equiv s_{k'} + (i - k')(s_{k'+1} - s_{k'}), \qquad i = 1, \dots, k,$$

for any $k' \in \{1, \dots, k-1\}$. Moreover, it is obvious that

$$P[U_{1:k} < s_1^{(l)}, \dots, U_{k:k} < s_k^{(l)}] \le \xi_k(0) \le P[U_{1:k} < s_1^{(u)}, \dots, U_{k:k} < s_k^{(u)}]. \tag{4.5}$$

It then remains to evaluate the two bounds in (4.5) by applying (2.4), which leads to (4.3) and (4.4). The convex case can be treated similarly, under the constraint for (4.4) that $s_1^{(u)}$ defined earlier is positive. ■

Acknowledgments

This research was supported in part by the Banque Nationale de Belgique. I am grateful to I. Coulibaly (ISRO) for his help with the numerical part. I also wish to thank the Editor for helpful remarks.

References

1. Andersson, H. & Britton, T. (2000). *Stochastic epidemic models and their statistical analysis*. Lecture Notes in Statistics 151. New York: Springer-Velag.
2. Ball, F.G. & O'Neill, P.D. (1999). The distribution of general final state random variables for stochastic epidemic models. *Journal of Applied Probability* 36: 473–491.
3. Consul, P.C. (1974). A simple urn model dependent on pre-determined strategy. *Sankhya* 36: 391–399.
4. Daley, D.J. & Gani, J. (1999) *Epidemic modelling: An introduction*. Cambridge: Cambridge University Press.
5. Denuit, M., Lefèvre, C., & Picard, P. (2003). Polynomial structures in order statistics distributions. *Journal of Statistical Planning and Inference* 113: 151–178.
6. Dobson, I., Carreras, B.A., & Newman, D.E. (2005). A loading-dependent model of probabilistic cascading failure. *Probability in the Engineering and Informational Sciences* 19: 15–32.
7. Lefèvre, C. & Picard, P. (2005). Non-stationarity and randomization in the Reed–Frost epidemic model. *Journal of Applied Probability* 42: 950–963.
8. Lefèvre, C. & Utev, S. (1996). Comparing sums of exchangeable Bernoulli random variables. *Journal of Applied Probability* 33: 285–310.
9. Picard, P. & Lefèvre, C. (1990). A unified analysis of the final state and severity distribution in collective Reed–Frost epidemic processes. *Advances in Applied Probability* 22: 269–294.
10. Shorack, G.R. & Wellner, J.A. (1986). *Empirical processes with applications to statistics*. New York: Wiley.
11. Stadje, W. (1993). Distribution of first-exit times for empirical counting and Poisson processes with moving boundaries. *Stochastic Models* 9: 91–103.
12. Zacks, S. (1991). Distributions of stopping times for Poisson processes with linear boundaries. *Stochastic Models* 7: 233–242.