J. Appl. Prob. **26**, 477–489 (1989)
Printed in Israel

ON REINFORCEMENT–DEPLETION COMPARTMENTAL URN MODELS

PETER DONNELLY,* *Queen Mary College, University of London*
WARD WHITT,** *AT&T Bell Laboratories*

Abstract

We verify and extend a conjecture of Purdue (1981) concerning the stochastic monotonicity of absorption times in a class of compartmental urn models. We also describe the effect of increased variability in the reinforcement sizes. Finally, we investigate variability in the content process for large populations. In many applications, compartmental models substantially under-represent the variability observed in the data, so that there has been considerable interest in modifying the model to increase the variability. We show that the squared coefficient of variation of the content is not asymptotically negligible when both the size and the variability of the reinforcements are of the same order as the initial population.

COMPARTMENTAL MODELS; FIRST-PASSAGE TIMES; STOCHASTIC ORDER; COUPLING; MAJORIZATION; LIMIT THEOREMS

1. Introduction

Compartmental models arise in a wide variety of physical and biological applications. A particular single compartment model with bulk arrivals and departures, the so-called reinforcement–depletion urn model, was introduced by Bernard (1977) in studying the dilution over time of a collection of radioactive atoms within some region. Suppose an urn initially contains b black balls and w white balls. At the nth stage ($n = 1, 2, \cdots$) the black balls in the urn are *reinforced* by the addition of R_n extra black balls. The contents of the urn are then mixed and *depletion* occurs, as R_n balls selected at random, independently of all other events, are removed from the urn. Sometimes reinforcement–depletions are assumed to occur in discrete time and sometimes at the instances of a time-homogeneous Poisson process. We follow Bernard (1977) and work with discrete time, but our results apply to the other situation too. We are interested in the number of white balls that remain in the urn after n stages, which we denote by W_n, and the time (stage) T when all the white balls have first been removed.

Received 3 April 1986; revision received 23 August 1988.
* Postal address: School of Mathematical Sciences, Queen Mary College, University of London, Mile End Road, London E1 4NS, UK.
Partially supported by the U.S. National Science Foundation under Grant No. DMS-86-08857.
** Postal address: AT&T Bell Laboratories, Room 2C-178, Murray Hill, NJ 07974, USA.

In Bernard's original model the size of each reinforcement–depletion is constant, say r, and it is possible to derive relatively simple expressions for the mean and variance of W_n, namely,

$$E(W_n) = w(1-\rho)^n \tag{1.1}$$

and

$$\mathrm{Var}(W_n) = w[(1-\rho)^n - (1-\alpha\rho)^n] + w^2[(1-\alpha\rho)^n - (1-\rho)^{2n}], \tag{1.2}$$

where

$$\rho = \frac{r}{w+b+r} \quad \text{and} \quad \alpha = 1 + \frac{w+b}{w+b+r-1}\,; \tag{1.3}$$

see Leitnaker and Purdue (1985). Shenton (1981) and Leitnaker and Purdue (1985) have also obtained the distribution of W_n,

$$P(W_n = k) = \binom{w}{k} \sum_{i=0}^{w-k} (-1)^i \binom{w-k}{i} \left[\frac{\binom{b+w+r-k-i}{r}}{\binom{b+w+r}{r}} \right]^n \tag{1.4}$$

Analogous more complicated expressions exist in the case of random independent and identically distributed reinforcements (Leitnaker and Purdue (1985)).

Another way to gain insight into this model is to seek qualitative results. Along these lines, Purdue (1981) conjectured that the absorption time T is stochastically decreasing in r for the case $P(R_n = r) = 1$ for all n. Our first purpose is to prove that conjecture and more general stochastic comparisons. We begin in Section 2 by considering one stage of a reinforcement–depletion process, allowing the initial number of balls in the urn and the size of the reinforcement to be random and dependent. In Section 3 we obtain stochastic comparisons for a sequence of reinforcement–depletions, including results for the first-passage time T, the process $\{W_n : n \geqq 0\}$ and other more descriptive processes that keep track of all substitutions. Throughout, we exploit ordered couplings, i.e., constructions of random elements on a common probability space that individually have prescribed distributions and are ordered w.p.1 (with probability 1); see Kamae *et al.* (1977).

In Section 4 we discuss the effect of more variable reinforcements. These variability comparisons show that, in a certain sense, the white balls are more likely to be removed when there is less variability in the reinforcement sizes. We consider both variability in the *distribution* of reinforcement sizes and variability over *time*. In particular, we show that the moments $E(W_n^m)$ conditioned on the reinforcement sizes are convex functions of each reinforcement size r_j for all n and m. Thus, among all reinforcement size vectors $(R_1, \cdots, R_n)$ with independent marginals and means $(r_1, \cdots, r_n)$, $E(W_n^m)$ is minimized by using the deterministic reinforcement sizes $(r_1, \cdots, r_n)$. We also show that the moments $E(W_n^m)$ conditional on the reinforcement sizes are Schur-convex; see Marshall and Olkin (1979). As a consequence, among all deterministic reinforcement sequences $(r_1, \cdots, r_n)$ with $r_1 + \cdots + r_n = nr$ for some integer r, $E(W_n^m)$ is minimized by using the constant deterministic sequence $(r,\cdots,r)$. Related stochastic comparisons can be obtained from stochastic majorization, see Chapter 11 of Marshall and Olkin (1979). For additional background on stochastic comparisons, see Ross (1983) and Stoyan (1983).

Finally, in Section 5 we address a problem of concern to Bernard (1977) and the other authors, the variability of W_n relative to the mean. In many applications, compartmental models substantially under-represent the variability observed in the data, so that there has been considerable interest in modifying the model to increase variability, as measured by the squared coefficient of variation (variance divided by the square of the mean) of W_n, denoted by $c^2(W_n)$. Bernard (1977) observed that $c^2(W_n)$ is negligible when $P(R_n=r)=1$ for $r=1$ and w is very large, and proposed the reinforcement–depletion model with $P(R_n=r)=1$ for large r, e.g., of order w, to increase $c^2(W_n)$. However, Leitnaker and Purdue (1985), p. 195, showed that Bernard's scheme does not actually achieve its desired purpose. Indeed, as $w\to\infty$ with the ratios $b(w)/w\to\beta>0$ and $r(w)/w\to\eta>0$ in (1.3)

$$\rho(w)\to\bar\rho\equiv\eta/(1+\beta+\eta)\quad\text{and}\quad\alpha(w)\to\bar\alpha\equiv 1+[(1+\beta)/(1+\beta+\eta)]=2-\bar\rho,$$

so that

$$\mathrm{Var}(W_n)\sim w[(1-\bar\rho)^n-(1-\bar\rho)^{2n}-n\gamma\bar\rho(1-\bar\rho)^{2n-1}] \tag{1.5}$$

and

$$c^2(W_n)\equiv\frac{\mathrm{Var}(W_n)}{[E(W_n)]^2}\sim\frac{1-(1-\bar\rho)^n-n\gamma\bar\rho(1-\bar\rho)^{n-1}}{w(1-\bar\rho)^n}\to 0\quad\text{as } w\to\infty, \tag{1.6}$$

where

$$\gamma=\frac{1}{1+\beta+\eta}=\bar\rho/\eta, \tag{1.7}$$

for each fixed n, where $f(w)\sim g(w)$ means that $f(w)/g(w)\to 1$ as $w\to\infty$. The third term in (1.5) comes from the w^2 term in (1.2); i.e.,

$$(1-\alpha\rho)^n=\left[(1-\rho)^2-\frac{\rho(1-\rho)}{w+b+r-1}\right]^n, \tag{1.8}$$

so that

$$w^2[(1-\alpha\rho)^n-(1-\rho)^{2n}]=-wn\gamma\bar\rho(1-\bar\rho)^{2n-1}+O(1)\quad\text{as } w\to\infty. \tag{1.9}$$

Of course, (1.6) is a limit as $w\to\infty$ for fixed n. We might want to let n grow with w. However, from (1.8) we see that

$$(1-\alpha\rho)^n\leqq(1-\rho)^{2n}, \tag{1.10}$$

so that

$$\mathrm{Var}(W_n)\leqq E(W_n)\quad\text{for all } n \text{ and } w. \tag{1.11}$$

Hence, as long as $E(W_n)$ is large, $c^2(W_n)$ is relatively small.

To introduce additional variability in W_n, Leitnaker and Purdue (1985) let the reinforcement sizes R_n be random. (In addition, they let the successive reinforcement sizes be i.i.d.) They argue that $c^2(W_n)$ becomes non-negligible when $\mathrm{Var}(R_n) > 0$ and $w \to \infty$, with the distribution of R_n evidently not changing with w. However, in Section 5 we show that $c^2(W_n)$ is also negligible in this revised model. On the positive side, we show that non-negligible $c^2(W_n)$ can be obtained by letting $w \to \infty$ with $R_n = wX_n$, where X_n is a positive non-deterministic integer-valued random variable independent of w, i.e., when we have *both* the large size of R_n suggested by Bernard *and* the variability suggested by Leitnaker and Purdue. There are of course other ways to introduce more content variability into the model, and other issues to consider; e.g., see Ball and Donnelly (1989).

2. Ordered couplings of samples without replacement

In this section we treat the special case of *one stage of reinforcement and depletion*. The analysis is based on the following sampling problem: suppose that we select a sample of size k at random without replacement from n balls, and afterwards we decide that we actually want a sample of size $k + 1$ chosen at random from these n balls plus one other. Can we validly obtain the second sample by adding a yet unselected ball to the first sample?

It is not difficult to see that we can. Indeed, with probability $(k + 1)/(n + 1)$ we should augment our first sample by including the new $(n + 1)$th ball, and with probability $(n - k)/(n + 1)$ we should augment our first sample by choosing at random one of the $n - k$ balls among the first n not previously chosen. Clearly, every subset of size $k + 1$ from the $n + 1$ balls has equal probability with this scheme. This construction is of some interest, because *conditional* on having obtained the sample of size k from the first n balls, we are much more likely ($(k + 1)/(n + 1)$ versus $1/(n + 1)$) to choose the new ball than any one of the first $n - k$ balls not previously chosen.

We now apply this elementary construction to treat one stage of reinforcement–depletion with deterministic reinforcement sizes and initial conditions. Consider two sequences of balls each labelled by the positive integers. The original balls in the urn come from the first sequence and the reinforcements come from the second sequence. The balls in the first sequence may be white or black; that does not matter, because we keep track of each ball. Let $S(k, r)$ be the subset of r balls obtained by sampling at random without replacement from $k + r$ balls, where the set of $k + r$ balls contains the first k balls of the first sequence and the first r balls of the second sequence. Let $S^c(k, r)$ be the complement of $S(k, r)$, i.e., the k balls not selected from the designated set of $k + r$ balls. Let $I_{ij}(k, r) = 1$ if the jth ball in the ith sequence is *not* in the sample $S(k, r)$, and 0 otherwise, for $j = 1, 2, \cdots$ and $i = 1, 2$. Let $J_{ij}(k, r) = 1$ if the jth ball in the ith sequence is contained in $S^c(k, r)$, and 0 otherwise, for $j = 1, 2, \cdots$ and $i = 1, 2$. Note that $J_{1j}(k, r) = 0$ while $I_{1j}(k, r) = 1$ for $j \geqq k + 1$. Let $I(k, r)$ represent the matrix $\{I_{ij}(k, r) : i = 1, 2; j \geqq 1\}$ and similarly for $J(k, r)$. For any subset W of non-negative integers, let $N_{W1}(k, r)$ count the number of balls from the first k in the first sequence that are simultaneously in the set W and not in the sample $S(k, r)$, i.e.,

$$N_{W1}(k, r) = \sum_{\substack{j=1 \\ j \in W}}^{k} I_{1j}(k, r) = \sum_{\substack{j=1 \\ j \in W}}^{k} J_{1j}(k, r). \tag{2.1}$$

For example, W might be the set of indices of all white balls in the first sequence; then $N_{W1}(k, r)$ is W_1, the number of white balls remaining in the reference set of $k + r$ balls after the sample $S(k, r)$ has been removed (after one reinforcement–depletion cycle).

Repeated use of the sampling construction above immediately yields the following result. (Part (b) can be obtained from part (a) by relabeling.)

Theorem 2.1. It is possible to perform the sampling (yielding the proper distributions) so that either (a) or (b) below holds (but not both at once):

(a) For any $k \geqq 0$, $S(k, r) \subseteq S(k, r + 1)$, $I_{ij}(k, r) \geqq I_{ij}(k, r + 1)$ for $i = 1, 2$ and all $j \geqq 1$, and $N_{W1}(k, r) \geqq N_{W1}(k, r + 1)$ for all $r \geqq 0$ w.p.1.

(b) For any $r \geqq 0$, $S^c(k, r) \subseteq S^c(k + 1, r)$, $J_{ij}(k, r) \leqq J_{ij}(k + 1, r)$ for $i = 1, 2$ and all $j \geqq 1$, and $N_{W1}(k + 1, r) \geqq N_{W1}(k, r)$ for all $k \geqq 0$ w.p.1.

Theorem 2.1 immediately yields stochastic comparisons. As in Kamae *et al.* (1977), we say that stochastic order holds for random elements X_1 and X_2 of a complete separable metric space Σ endowed with a closed partial order $\leqq$, and write $X_1 \leqq_{\mathrm{st}} X_2$, if $Ef(X_1) \leqq Ef(X_2)$ for all non-decreasing measurable real-valued functions on $(\Sigma, \leqq)$ for which the expectations are well defined. This is a convenient definition because such stochastic order implies that there exist ordered couplings, i.e., that there exist new random elements $\tilde{X}_1$ and $\tilde{X}_2$ on a common probability space such that $P(\tilde{X}_1 \leqq \tilde{X}_2) = 1$, $\tilde{X}_1 \stackrel{\mathrm{d}}{=} X_1$ and $\tilde{X}_2 \stackrel{\mathrm{d}}{=} X_2$, where $\stackrel{\mathrm{d}}{=}$ means equality in distribution.

The following corollary gives stochastic comparisons for subsets of the sequences (let Σ be the set of all finite subsets from the two sequences with the partial order of set inclusion) and the matrices $I(k, r)$ and $J(k, r)$ (here $\Sigma = R^2 \times R^\infty$ with componentwise order, i.e, $\mathbf{z}^1 \equiv (z_{ij}^1) \leqq \mathbf{z}^2 \equiv (z_{ij}^2)$ if $z_{ij}^1 \leqq z_{ij}^2$ for all i and j). Let K and R, with superscript indices, be random variables yielding possible values for k and r.

Corollary 2.1. (a) If $R^1 \leqq_{\mathrm{st}} R^2$, then $S(k, R^1) \leqq_{\mathrm{st}} S(k, R^2)$ and $I(k, R^2) \leqq_{\mathrm{st}} I(k, R^1)$ for each k.

(b) If $K^1 \leqq_{\mathrm{st}} K^2$, then $S^c(K^1, r) \leqq_{\mathrm{st}} S^c(K^2, r)$ and $J(K^1, r) \leqq_{\mathrm{st}} J(K^2, r)$ for each r.

Proof. Use the assumed stochastic order for $R^1 \leqq_{\mathrm{st}} R^2$ in (a) and $K^1 \leqq_{\mathrm{st}} K^2$ in (b) to construct new versions with the same distributions that are ordered w.p.l. Then apply Theorem 2.1 to get w.p.l order for the new versions of the displayed quantities. This w.p.l order in turn implies the order for the expectations in the definition of stochastic order. Since the coupling preserves the marginal distributions, the expectations are the same for all versions.

Our next result is proved in the same way as Corollary 2.1, but now we assume stochastic order for the random vectors $(K^j, -R^j)$. If K^j is independent of R^j for each j, then the condition is equivalent to $K^1 \leqq_{st} K^2$ and $R^2 \leqq_{st} R^1$ separately. However, the condition covers cases in which independence of K^j and R^j does *not* hold.

Corollary 2.2. If $(K^1, -R^1) \leqq_{st} (K^2, -R^2)$, then $N_{W1}(K^1, R^1) \leqq_{st} N_{W1}(K^2, R^2)$.

Proof. Use the stochastic order condition to construct an ordered coupling $(\tilde{K}^1, \tilde{R}^1, \tilde{K}^2, \tilde{R}^2)$ with $\tilde{K}^1 \leqq \tilde{K}^2$ and $\tilde{R}^1 \geqq \tilde{R}^2$ w.p.1, and $(\tilde{K}^j, \tilde{R}^j) \stackrel{d}{=} (K^j, R^j)$ for $j = 1, 2$. By Theorem 2.1(a), we can perform the reinforcement–depletion (tantamount to constructing a coupling) so that

$$N_{W1}(\tilde{K}^1, \tilde{R}^1) \leqq N_{W1}(\tilde{K}^1, \tilde{R}^2) \quad \text{w.p.l.} \tag{2.2}$$

By Theorem 2.1(b), we can perform the reinforcement–depletion (a different coupling) so that

$$N_{W1}(\tilde{K}^1, \tilde{R}^2) \leqq N_{W1}(\tilde{K}^2, \tilde{R}^2) \quad \text{w.p.l.} \tag{2.3}$$

We cannot combine (2.2) and (2.3) into a single coupling, but we can deduce the stochastic comparisons

$$N_{W1}(\tilde{K}^1, \tilde{R}^1) \leqq_{st} N_{W1}(\tilde{K}^1, \tilde{R}^2) \leqq_{st} N_{W1}(\tilde{K}^2, \tilde{R}^2).$$

Since $N_{W1}(\tilde{K}^i, \tilde{R}^j) \stackrel{d}{=} N_{W_1}(K^i, R^j)$ for each i, j, the desired stochastic comparison is established.

3. A general sequence of reinforcement–depletions

It is not difficult to extend Section 2 to a sequence of reinforcement–depletions where the successive sizes come from the general random sequence $\mathbf{R} \equiv \{R_n : n \geqq 1\}$. Our basic assumption is that, conditional on the sizes R_n, $n \geqq 1$, the successive reinforcement–depletions are mutually independent. We put this in the framework of Section 2 by successively repeating the sampling experiment there, sampling R_n balls from $K + R_n$ balls in the nth period, where the set of $K + R_n$ balls contains the first K (independent of n) balls from the first sequence and the first R_n balls from the second sequence.

To make a stochastic comparison, we assume that $\mathbf{R}^1 \leqq_{st} \mathbf{R}^2$, which is stochastic order for random elements of Z_+^∞ with componentwise ordering. If the reinforcement sizes R_n^j, $n \geqq 1$, are mutually independent for each j, then this condition is equivalent to $R_n^1 \leqq_{st} R_n^2$ for each n (without any common distribution condition). (A trivial but relevant case is the deterministic case $R_n^1 \equiv r_n^1 \leqq r_n^2 \equiv R_n^2$ for all n.) However, $\mathbf{R}^1 \leqq_{st} \mathbf{R}^2$ may hold without this independence, as is illustrated by results in Kamae et al. (1977), Sonderman (1980) and references cited there.

Given $\mathbf{R}^1 \leqq_{st} \mathbf{R}^2$, we can construct an ordered coupling $(\tilde{\mathbf{R}}^1, \tilde{\mathbf{R}}^2)$ with $\tilde{\mathbf{R}}^1 \leqq \tilde{\mathbf{R}}^2$ w.p.l and $\tilde{\mathbf{R}}^j \leqq_{st} \mathbf{R}^j$ for $j = 1, 2$. For this ordered coupling, then, Theorem 2.1(a) holds every period. For each k, the sequence of periods n in which ball j appears in the sample $S(k, \tilde{R}_n^1)$ is thus a *subsequence* of the sequence of periods n in which ball j appears in the sample $S(k, \tilde{R}_n^2)$. Stochastic conclusions then follow, just as in Section 2.

We now indicate how to apply this construction to the original reinforcement–depletion model. For model k, $k = 1, 2$, let W_{k0} and B_{k0} be the random numbers of white and black balls originally in the urn, let $K^k = W_{k0} + B_{k0}$, let W_{kn} be the number of white balls in the urn after n reinforcement–depletion cycles, let R_n^k be the size of the nth reinforcement of black balls, and let T_k be the first time that there are no white balls in the urn.

Theorem 3.1. If $\{W_{10}, K^1, -R_n^1 : n \geqq 1\} \leqq_{st} \{W_{20}, K^2, -R_n^2 : n \geqq 1\}$, then

$$\{T_1, W_{1n} : n \geqq 0\} \leqq_{st} \{T_2, W_{2n} : n \geqq 0\}.$$

Proof. To put this in the framework we have introduced, let the subset of indices corresponding to the white balls be W^k. Then W_{kn} can be expressed as

$$W_{kn} = \sum_{\substack{j \neq 1 \\ j \in W^k}}^{K^k} \prod_{m=1}^{n} I_{1j}(K^k, R_m^k), \qquad n \geqq 1. \tag{3.1}$$

The ordered coupling corresponding to the condition yields $\tilde{W}_{10} \leqq \tilde{W}_{20}$, $\tilde{K}^1 \leqq \tilde{K}^2$ and $\tilde{R}_n^1 \geqq \tilde{R}_n^2$ for all n. We thus can assume that $W^1 \subseteq W^2$ and do. Since $I_{1j}(k^1, r^1) \leqq I_{1j}(k^2, r^2)$ if $j \leqq k^1 \leqq k^2$ and $r^1 \geqq r^2$,

$$\begin{aligned}\tilde{W}_{1n} = \sum_{\substack{j=1 \\ j \in W^1}}^{\tilde{K}^1} \prod_{m=1}^{n} I_{1j}(\tilde{K}^1, \tilde{R}_m^1) &\leqq \sum_{\substack{j=1 \\ j \in W^1}}^{\tilde{K}^1} \prod_{m=1}^{n} I_{1j}(\tilde{K}^2, \tilde{R}_m^2) \\ &\leqq \sum_{\substack{j=1 \\ j \in W^2}}^{\tilde{K}^2} \prod_{m=1}^{n} I_{1j}(\tilde{K}^2, \tilde{R}_m^2) = \tilde{W}_{2n} \quad \text{for all } n \geqq 1,\end{aligned}$$

from which the stochastic comparison follows.

The conclusion of Theorem 3.1 implies the stochastic order of the first passage times $T_1 \leqq_{st} T_2$ conjectured by Purdue (1981); the conditions in Theorem 3.1 are also more general.

4. The effect of more variable reinforcement sizes

Recall that one random vector $(R_1^1, \cdots, R_n^1)$ is said to be *less variable* than another, $(R_1^2, \cdots, R_n^2)$, if

$$E[f(R_1^1, \cdots, R_n^1)] \leqq E[f(R_1^2, \cdots, R_n^2)]. \tag{4.1}$$

for all real-valued convex functions f; see p. 26 of Stoyan (1983). Note that for (4.1) to hold for all such f, and hence for two random vectors to be comparable in variability, we must have $E(R_i^1) = E(R_i^2)$, $i = 1, 2, \cdots, n$. (Consider the functions $f_i(x_1, \cdots, x_n) = x_i$ and $g_i = -f_i$, $i = 1, 2, \cdots, n$.) One example to bear in mind is that $(ER_1, \cdots, ER_n)$ is always less variable than $(R_1, \cdots, R_n)$.

Another comparison of random vectors is given by stochastic majorization. Say that $(R_1^1, \cdots, R_n^1)$ is *stochastically majorized* by $(R_1^2, \cdots, R_n^2)$ if

$$E[h(R_1^1, \cdots, R_n^1)] \leqq E[h(R_1^2, \cdots, R_n^2)] \tag{4.2}$$

for all real-valued, Schur-convex functions h. For a definition of Schur-convexity and further details about stochastic majorization, see Marshall and Olkin (1979). When the random vectors are deterministic, stochastic majorization reduces to ordinary majorization. For one vector to majorize another, it is necessary that they have identical sums of components. For example $(r, r, \cdots, r)$ is always majorized by $(r_1, \cdots, r_n)$ when $r_1 + \cdots + r_n = nr$.

In order to apply these ideas in this context, we first show that certain conditional moments of the content process W_n are convex, and Schur-convex in their arguments. As before, let W_0 and B_0 be the initial numbers of white and black balls in the urn.

Theorem 4.1. For each $m \geqq 1$ and $n \geqq 1$, the conditional mth moment $E(W_n^m \mid W_0 = w, B_0 = b, R_1 = r_1, \cdots, R_n = r_n)$ is a convex function of r_j for each j and a Schur-convex function of $(r_1, \cdots, r_n)$.

Proof. For $1 \leqq i \leqq w$, let $X_i = 1$ if the ith white ball remains after n stages, and 0 otherwise. Let A be the event $\{W_0 = w, B_0 = b, R_1 = r_1, \cdots, R_n = r_n\}$. Then

$$E(W_n^m \mid A) = \Sigma E(X_{i_1} X_{i_2} \cdots X_{i_m} \mid A)$$

where the sum extends over all possible m-tuples of integers with $1 \leqq i_j \leqq w$. The number of factors in the product reduces when the same index appears more than once. Each reduced term containing l different X_i variables is of the form

$$E(X_{i_1} X_{i_2} \cdots X_{i_l} \mid A) = \prod_{j=1}^{n} \left(\frac{w+b}{w+b+r_j}\right)\left(\frac{w+b-1}{w+b+r_j-1}\right) \cdots \left(\frac{w+b-l+1}{w+b+r_j-l+1}\right)$$

which is a convex function of r_j for each j. For each j, use the fact that the product of two, and thus any number, of positive decreasing convex differentiable real-valued functions is convex. For Schur-convexity, apply B.1.d. on p. 62 of Marshall and Olkin (1979).

Theorem 4.1 allows us to make further comparisons between processes. Consider two sequences of reinforcements $R_1^1, R_2^1, \cdots$, and $R_1^2, R_2^2, \cdots$ and common initial numbers W_0, B_0 of white and black balls, and denote the resulting processes by $\{W_{1n} : n = 1, 2, \cdots\}$ and $\{W_{2n} : n = 1, 2, \cdots\}$.

Corollary 4.1. Suppose that W_0, B_0, R_n^k, $n \geqq 1$, are mutually independent for $k = 1, 2$. If R_i^1 is less variable than R_i^2 for $i = 1, 2, \cdots, n$, then $E(W_{1n}^m) \leqq E(W_{2n}^m)$ for all m.

Corollary 4.2. If $(R_1^1, \cdots, R_n^1)$ is stochastically majorized by $(R_1^2, \cdots, R_n^2)$, then $E(W_{1n}^m) \leqq E(W_{2n}^m)$ for all m.

Here are two important special cases:

(i) Among all reinforcement size vectors $(R_1, \cdots, R_n)$ with independent marginals and means $(r_1, \cdots, r_n)$, $E(W_n^m)$ is minimized by using the deterministic reinforcement sizes $(r_1, \cdots, r_n)$.

(ii) Among all deterministic reinforcement sequences $(r_1, \cdots, r_n)$ with $r_1 + \cdots + r_n = nr$ for some integer r, $E(W_n^m)$ is minimized by using the constant sequence $(r, \cdots, r)$.

Remarks. (1) It is not difficult to see that $E(W_n^m \mid W_0 = w_1,\ B_0 = b, R_1 = r_1, \cdots, R_n = r_n)$ is *not* convex as a function of $(r_1, \cdots, r_n)$, so we need the independence in Corollary 4.1.

(2) The probabilities $P(W_n = 0) = P(T \leqq n)$ do not necessarily decrease in response to increased variability. For example, let $w = 2$, $P(R_1^1 = 1) = 1$ and $P(R_1^2 = 0) = P(R_1^2 = 2) = 1/2$. Then R_1^1 is less variable than R_1^2, but $P(W_{11} = 0) = 0 < P(W_{21} = 0)$.

5. Variability of the compartment contents

As indicated in the introduction, Bernard (1977), Purdue (1981) and Leitnaker and Purdue (1985) were interested in the variability of $W_n \equiv W_n(w)$ relative to the mean for large values of w, as measured for example by the squared coefficient of variation of W_n. Assume that there are initially b black balls and w white balls in the urn, and let the sequence of successive reinforcement sizes $\{R_n : n \geqq 1\}$ be i.i.d. The expressions for the mean and variance of W_n on p. 200 of Leitnaker and Purdue (1985) can be used to describe the asymptotic behavior of $c^2(W_n)$ as $w \to \infty$; they are

$$E(W_n) = w\left[\sum_{i=0}^{\infty} (-1)^i \frac{E(R_1^i)}{(b+w)^i}\right]^n \tag{5.1}$$

and

$$\operatorname{Var}(W_n) = V_1 + V_2 - V_3, \tag{5.2}$$

where

$$\begin{aligned} V_1 &= w\left[\sum_{i=0}^{\infty} (-1)^i \frac{E(R_1^i)}{(b+w)^i}\right]^n \\ V_2 &= w(w-1)\left[\sum_{i=0}^{\infty}\sum_{k=0}^{\infty} (-1)^{i+k} \frac{E(R_1^{i+k})}{(b+w)^i(b+w-1)^k}\right]^n \\ V_3 &= w^2\left[\sum_{i=0}^{\infty} (-1)^i \frac{E(R_1^i)}{(b+w)^i}\right]^{2n}. \end{aligned} \tag{5.3}$$

We assume that there is a constant r such that

$$E(R_1^i) \leqq r^i \quad \text{for all } i \geqq 1, \tag{5.4}$$

so that all the sums in (5.1)–(5.3) are absolutely convergent for sufficiently large w, and we can easily do asymptotic analysis as $w \to \infty$. An obvious sufficient condition for (5.4) is $P(R_1 \leqq r) = 1$.

We first describe the situation considered by Leitnaker and Purdue. We omit the proof using (5.1)–(5.4).

Theorem 5.1. Suppose that the distribution of R_1 is independent of w. If $w \to \infty$ with $w/(b+w) \to p > 0$, then

$$E(W_n) - w \sim -pnE(R_1) \tag{5.5}$$

and

$$\text{Var}(W_n) \sim n\sigma^2 \equiv n[p(1-p)E(R_1) + p^2 \text{Var}(R_1)], \tag{5.6}$$

so that

$$c^2(W_n) \equiv \frac{\text{Var}(W_n)}{[E(W_n)]^2} \sim \frac{n\sigma^2}{w^2} \tag{5.7}$$

and

$$c^2(w - W_n) \equiv \frac{\text{Var}(w - W_n)}{(E(w - W_n))^2} \sim \frac{\sigma^2}{np^2[E(R_1)]^2} . \tag{5.8}$$

Hence, for any fixed n, the squared coefficient of variation of the number of white balls *remaining* is asymptotically negligible as $w \to \infty$, while the squared coefficient of variation of the number of white balls *removed* is not.

These asymptotics should be intuitively clear, because as w gets large with n and the distribution of R_i fixed, the number of white balls removed, $w - W_n(w)$, can obviously be bounded above by a quantity that is independent of w, i.e.,

$$w - W_n(w) \leqq R_1 + \cdots + R_n \quad \text{for all } w. \tag{5.9}$$

Hence, we trivially have

$$w^{-1}W_n(w) \to 1 \quad \text{w.p.1} \quad \text{as } w \to \infty, \tag{5.10}$$

$$\begin{aligned} \text{Var}(W_n) = \text{Var}(w - W_n) &\leqq E[(w - W_n)^2] \leqq E[(R_1 + \cdots + R_n)^2] \\ &\leqq n \text{Var}(R_1) + n^2[E(R_1)]^2, \end{aligned} \tag{5.11}$$

$$E(W_n) \geqq w - nE(R_1), \tag{5.12}$$

and

$$c^2(W_n) \leqq \frac{n \text{Var}(R_1) + n^2[E(R_1)]^2}{[w - nE(R_1)]^2} = O(w^{-2}). \tag{5.13}$$

Moreover, the number of white balls removed, $w - W_n(w)$, obviously approaches the number of successes in a random number of Bernoulli trials; i.e., we have the following elementary asymptotic result, which we also state without proof. Let $\Rightarrow$ denote convergence in distribution.

Theorem 5.2. As $w \to \infty$,

$$(w - W_n(w)) \Rightarrow U_1 + \cdots + U_{(R_1 + \cdots + R_n)}, \tag{5.14}$$

where $\{U_n : n \geqq 1\}$ is a sequence of i.i.d. random variables, independent of $\{R_n : n \geqq 1\}$, with $P(U_n = 1) = 1 - P(U_n = 0) = p$.

Note that the mean and variance of the limit in (5.14) agree with (5.5) and (5.6), as they should, because (5.14) and (5.4) imply convergence of moments.

We now show that we cannot obtain interesting behavior for W_n as $w \to \infty$ if we simply let n grow with w, so that the stage is in the same scale as the content size. From (5.7), we expect that $c^2(W_n) \sim w^{-1}$ as $w \to \infty$ with $n/w \to \xi > 0$ and $w/(b+w) \to p > 0$. In fact, this can be shown to be the case, again using (5.1)–(5.4). We use the fact that $(1 + n^{-1}x_n)^n \to e^x$ as $n \to \infty$ if $x_n \to x$ as $n \to \infty$; see p. 169 of Chung (1974).

Theorem 5.3. If $w \to \infty$ with $n/w \to \xi > 0$ and $w/(b+w) \to p > 0$, then

$$E(W_n) \sim w \exp(-p\xi E(R_1)), \tag{5.15}$$

and

$$\begin{aligned} \operatorname{Var}(W_n) \sim\; & w \exp(-p\xi E(R_1)) \\ & \times (1 - \exp(-p\xi E(R_1))(1 + p^2\xi[E(R_1) - \operatorname{Var}(R_1)])), \end{aligned} \tag{5.16}$$

so that

$$c^2(W_n) \equiv \frac{\operatorname{Var}(W_n)}{[E(W_n)]^2} = O(w^{-1}) \quad \text{as } w \to \infty. \tag{5.17}$$

So far, we have kept the distribution of R_1 fixed as $w \to \infty$. Significantly greater variability is seen if the distribution of R_1 grows with w. To treat this case, assume that

$$R_j = wX_j, \tag{5.18}$$

where $X_j, j \geqq 1$, are i.i.d. integer-valued random variables independent of w.

Theorem 5.4. If (5.18) holds and $w \to \infty$ with $w/(b+w) \to p > 0$, then

$$E(W_n) \sim w[E(Y)]^n \tag{5.19}$$

and

$$\operatorname{Var}(W_n) \sim w^2\{[E(Y^2)]^n - [E(Y)]^{2n}\}, \tag{5.20}$$

so that

$$c^2(W_n) \sim E(Y^2)^n[E(Y)]^{-2n} - 1, \tag{5.21}$$

where

$$Y = \frac{1}{1 + pX_1}. \tag{5.22}$$

Proof. To establish (5.19), instead of (5.1), we use

$$w^{-1}E(W_n) = \left[E\left(\frac{b+w}{b+w+R_1}\right)\right]^n; \tag{5.23}$$

see (22) of Leitnaker and Purdue. By (5.18), (5.23) and the dominated convergence theorem

$$w^{-1}E(W_n) = \left[E\left(\frac{1}{1 + [w/(b+w)]X_1}\right)\right]^n \to [E(Y)]^n \quad \text{as } n \to \infty.$$

Similarly, using (23) of Leitnaker and Purdue,

$$w^{-2}\operatorname{Var}(W_n) = w^{-1}\left[E\left(\frac{b+w}{b+w+R_1}\right)\right]^n + \frac{(w-1)}{w}\left[E\left(\frac{b+w}{b+w+R_1}\right)\left(\frac{b+w-1}{b+w-1+R_1}\right)\right]^n - \left[E\left(\frac{b+w}{b+w+R_1}\right)\right]^{2n} \to [E(Y^2)]^n - [E(Y)]^{2n}.$$

Paralleling Theorems 5.1 and 5.2, we can also generalize Theorem 5.4 by obtaining asymptotic properties of the distribution of W_n.

Theorem 5.5. Under the conditions of Theorem 5.4,

$$w^{-1}W_n \xrightarrow{\text{p}} Y_1Y_2\cdots Y_n \quad \text{as } w \to \infty,$$

where Y_n, $n \geqq 1$, are i.i.d. and distributed as Y in (5.22).

Proof. By the law of large numbers for sampling without replacement, as $w \to \infty$ the proportion of available white balls removed on the ith reinforcement–depletion cycle is asymptotically Y_i, as in (5.22), with successive cycles being independent. (Effectively, wX_i balls are sampled without replacement from $w(p^{-1} + X_i)$, where $w \to \infty$.) A direct proof follows from (1.1), (1.5) and Chebychev's inequality, after first conditioning on the X_i.

From Theorem 5.4, we see that $c^2(W_n) > 0$ if and only if $\operatorname{Var}(Y) > 0$ or equivalently, if and only if $\operatorname{Var}(X_1) > 0$, in which case $c^2(W_n)$ is increasing with $c^2(W_n) \to \infty$ as $n \to \infty$.

A convenient special case occurs when $P(X_1 = c) = q = 1 - P(X_1 = 0)$. Then

$$E(Y^k) = (1-q) + \frac{1}{(1+pc)^k}\,. \tag{5.24}$$

Acknowledgement

We thank Peter Purdue for communicating his conjecture and helping to bring about this collaboration.

References

Ball, F. and Donnelly, P. (1989) A unified approach to variability in compartmental models. *Biometrics*.

Bernard, S. R. (1977) An urn model study of variability within a compartment. *Bull. Math. Biol.* **39**, 463–470.

Chung, K. L. (1974) *A Course in Probability*, 2nd edn. Academic Press, New York.

KAMAE, T., KRENGEL, U. AND O'BRIEN, G. L. (1977) Stochastic inequalities on partially ordered spaces. *Ann. Prob.* **5**, 899–912.

LEITNAKER, M. G. AND PURDUE, P. (1985) Non-linear compartmental systems: extensions of S. R. Bernard's urn model. *Bull. Math. Biol.* **67**, 193–204.

MARSHALL, A. W. AND OLKIN, I. (1979) *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York.

PURDUE, P. (1981) Variability in a single compartment system: a note on S. R. Bernard's model. *Bull. Math. Biol.* **43**, 111–116.

ROSS, S. M. (1983) *Stochastic Processes.* Wiley, New York.

SHENTON, L. R. (1981) A reinforcement–depletion urn problem–I. Basic theory. *Bull. Math. Biol.* **43**, 327–340.

SHENTON, L. R. (1983) A reinforcement–depletion urn problem–II. Application and generalization. *Bull. Math. Biol.* **45**, 1–9.

SONDERMAN, D. (1980) Comparing semi-Markov processes. *Math. Operat. Res.* **5**, 110–119.

STOYAN, D. (1983) *Comparison Methods for Queues and Other Stochastic Models*, ed. D. J. Daley Wiley, New York.