Ann. Inst. Statist. Math.
38 (1986), Part A, 233-243

A REINFORCEMENT-DEPLETION URN MODEL: A CONTIGUITY CASE*

K. O. BOWMAN AND L. R. SHENTON

(Received July 14, 1984; revised Feb. 6, 1985)

Summary

An urn contains balls of s different colors. The problem of the reinforcement of a specified color and random depletion of balls has been considered by Bernard (1977, *Bull. Math. Biol.*, **39**, 463–470) and Shenton (1981, *Bull. Math. Biol.*, **43**, 327–340), (1983, *Bull. Math. Biol.*, **45**, 1–9). Here we consider a special relation between a reinforcement and depletion, leading to a hypergeometric distribution.

1. Introduction

An urn contains balls of two colors, red (σ) and white (ω); at the first stage (or cycle) r red balls and w white balls are added, and a random mixture of $r+w$ balls is removed. The process is repeated through m cycles. What is the distribution of red and white balls? Bernard [1] studied this model paying particular attention to means and variances when radioiodine is injected into animals. If the reinforcement of red balls in the urn at each cycle is large it seems likely that the variance of red balls at a subsequent stage will be large (small) if the number randomly removed at each stage is large (small).

The model has been generalized (Shenton, [4], [5]) so that balls of s colors can be considered. At each cycle each color receives a reinforcement (possibly zero) and any number of balls can be randomly removed (corresponding to a depletion). This includes the case when the depletion equals the reinforcement. Reinforcements for different cycles need not be equal. The sum of depletions can not exceed the number of balls available at any stage.

* Research sponsored in part by the Applied Mathematical Sciences Research Program, Office of Energy Research, U.S. Department of Energy under contract DE-AC05-840R21400 with the Martin Marietta Energy Systems, Inc.
Keywords: Contiguous factorials, factorial series distributions, hypergeometric distribution, moments.

The factorial moment generating function (fmgf) in the general case involves the multivariate finite difference calculus and depends on $ms+m+s$ parameters when there are m cycles. It is, except in trivial cases, impossible to develop the formulas mathematically, although low order factorial moments can be constructed; for example, the mean value of the number of balls of a specified color involves m terms, each of which has a product structure, altogether entailing $m(m+1)/2$ coefficients.

Our objective here then is to describe a model for balls of two colors for which the exact probability generating function (pgf) for one color can be found in closed form. The pgf is of hypergeometric form. This special model is derived from the general case by setting up a dependency between the reinforcements and depletions, one of which is assumed to be functionally defined. A certain property of contiguity is essential.

We are studying an application concerning the uptake of radioiodine by human subjects at different ages. Here the reinforcements and depletions over as many as 10950 cycles (30 years at 1 cycle a day) are of the order of 10^{12}, so that special care is needed in any computer implementation. As far as the distributions of the number of balls of a particular color are concerned there is special interest in the situation at the biological half-life of an element.

2. General theory for the three color urn

We turn to a brief description of the formulas for a three color urn scheme. It will be a slightly more detailed account than that given in Shenton [5] in the appendix. It is a simple matter to deduce the general formula for the mgf from that given for three colors.

For our purpose it is sufficient to consider a three color urn model with the following scheme, for $j=1, 2, \cdots, m$ $(m \geqq 1)$.

	Balls				
Colors	Red	White	Blue		
Initial numbers	σ	ω	β		$(\sigma+\omega+\beta=T_0)$
Cycle		Increments		Depletion	Totals
j	r_j	w_j	b_j	d_j	$T_j=r_j+w_j+b_j-d_j$
fmgf label	α_1	α_2			$(m \geqq 1)$

Notations: (a) Cumulative sums $(j=1, 2, \cdots)$

$$r_j^*=\sigma+r_1+\cdots+r_j\,, \quad w_j^*=\omega+w_1+\cdots+w_j\,, \quad b_j^*=\beta+b_1+\cdots b_j\,,$$

$d_j^* = d_1 + d_2 + \cdots + d_j$, $T_j^* = T_0 + T_1 + \cdots + T_j$; and

$(r_0^* = \sigma,\ w_0^* = \omega,\ b_0^* = \beta,\ d_0^* = 0,\ T_0^* = T_0 = \sigma + \omega + \beta)$.

(b) Total number of balls in the urn at the completion of the j-th cycle (i.e. d_j^* balls have been removed)

$$T_j^* = \sum_{k=0}^{j} (r_k + w_k + b_k - d_k) .$$

(c) δ_{jk} is the Kronecker delta function ($\delta_{jj} = 1$, $\delta_{jk} = 0$, $j \neq k$).

(d) $x^{(j)} \equiv x(x-1) \cdots (x-j+1)$.

From Shenton ([5], Appendix) the bivariate factorial moment generating function (fmgf) for the number of red balls ($N_r^{[m]}$) and white balls ($N_w^{[m]}$) at the end of the m-th cycle, is

$$(1) \qquad F_m(\alpha_1, \alpha_2) = c_m L(\mathrm{E}_1, \mathrm{E}_2, \cdots, \mathrm{E}_m : \alpha_1, \alpha_2) \prod_{j=1}^{m} x_j^{(d_j)} ,$$

where $L(\cdot)$ is a function of the finite difference incremental operators $\mathrm{E}_1, \mathrm{E}_2, \cdots$, namely

$$(2) \qquad L(\cdot) = \prod_{j=1}^{m} (\alpha_1 + \mathrm{E}_m \mathrm{E}_{m-1} \cdots \mathrm{E}_j)^{r_j + \sigma\delta_{j1}} (\alpha_2 + \mathrm{E}_m \mathrm{E}_{m-1} \cdots \mathrm{E}_j)^{w_j + \omega\delta_{j1}} ,$$

and $x_j = b_j^* - d_{j-1}^*$ after the application of the operator. (The operator E_k only operates on the component in the product in (1) with subscript k. In the univariate case $Ef(x) = f(x+1) = (1+\Delta)f(x)$, and contrary to usage in the calculus it is conventional to use differences of numbers; for example, $\Delta^r 0^s$ (for positive integers r, s) refers to the differences of zero. In the multivariate case $\Delta_x^r \Delta_y^s f(x, y)$ is unambiguous, but when numbers only are involved we resort to a notation $\Delta_1^r \Delta_2^s 0_1^t 0_2^u$ in which each operator is completed before the subscripts are removed.) The constant c_m in (1) is given by

$$(3) \qquad 1/c_m = (T_1^* + d_1)^{(d_1)} (T_2^* + d_2)^{(d_2)} \cdots (T_m^* + d_m)^{(d_m)} .$$

Take the simple case (Shenton [5], p. 8) arranged as follows;

$\sigma = 3$	$\omega = 2$	$\beta = 1$	
$r_1 = 2$	$w_1 = 0$	$b_1 = 1$	$d_1 = 1$
$r_2 = 1$	$w_2 = 0$	$b_2 = 2$	$d_2 = 1$

With α_1, α_2 as parameters corresponding to red and white balls respectively,

$$F_2(\alpha_1, \alpha_2) = c_2(\alpha_1 + \mathrm{E}_2)(\alpha_1 + \mathrm{E}_2\mathrm{E}_1)^5 (\alpha_2 + \mathrm{E}_2\mathrm{E}_1)^2 2_1^{(1)} (4-1)_2^{(1)} ,$$

where $c_2 \mathrm{E}_2 (\mathrm{E}_2\mathrm{E}_1)^5 (\mathrm{E}_2\mathrm{E}_1)^2 2_1 3_2 = 1$, or $c_2 = 1/99$. For the probability generating function (pgf) define $t_1=1+\alpha_1$, $t_2=1+\alpha_2$. Then the pgf is

$$(2t_1^6+12t_1^6t_2+6t_1^6t_2^2+22t_1^5t_2+32t_1^5t_2^2+25t_1^4t_2^2)/99 ,$$

where the first two terms, for example, correspond to configurations (6, 0, 4), and (6, 1, 3) respectively.

From (1) the first and second order factorial moments of say the number $N_t^{[m]}$ of red balls at the end of the m-th cycle are

$$E(N_r^{[m]})=\sum_{j=1}^{m}(r_j+\sigma\delta_{j1})\prod_{\lambda=j}^{m}\left(\frac{T_\lambda^*}{T_\lambda^*+d_\lambda}\right) \tag{4}$$

and

$$\begin{aligned} E(N_r^{[m]}(N_r^{[m]}-1))=&\sum_{j=1}^{m}(r_j+\sigma\delta_{j1})(r_j+\sigma\delta_{j1}-1)\prod_{\lambda=j}^{m}\left(1+\frac{d_\lambda}{T_\lambda^*}\right)^{-1} \\ &\times\left(1+\frac{d_\lambda}{T_\lambda^*-1}\right)^{-1}+2\sum_{1\leq k<j\leq m}(r_j+\sigma\delta_{j1})(r_k+\sigma\delta_{k1}) \\ &\times\prod_{\lambda=k}^{j-1}\left(1+\frac{d_\lambda}{T_\lambda^*}\right)^{-1}\prod_{\mu=j}^{m}\left(1+\frac{d_\mu}{T_\mu^*}\right)^{-1}\left(1+\frac{d_\mu}{T_\mu^*-1}\right)^{-1} \end{aligned} \tag{5}$$

respectively.

Also, for the expectation of a cross product, we have

$$\begin{aligned} E(N_r^{[m]}N_w^{[m]})=&\sum_{t\leq k<j\leq m}(r_j+\sigma\delta_{j1})(w_j+\omega\delta_{j1})\prod_{\lambda=j}^{m}\left(1+\frac{d_\lambda}{T_\lambda^*}\right)^{-1} \\ &\times\left(1+\frac{d_\lambda}{T_\lambda^*-1}\right)^{-1}\prod_{\mu=k}^{j-1}\left(1+\frac{d_\mu}{T_\mu^*}\right)^{-1} \\ &+\sum_{1\leq j<k\leq m}(r_j+\sigma\delta_{j1})(w_j+\omega\delta_{j1})\prod_{\lambda=k}^{m}\left(1+\frac{d_\lambda}{T_\lambda^*}\right)^{-1} \\ &\times\left(1+\frac{d_\lambda}{T_\lambda^*-1}\right)^{-1}\prod_{\mu=j}^{k-1}\left(1+\frac{d_\mu}{T_\mu^*}\right)^{-1} \\ &+\sum_{1\leq j=k\leq m}(r_j+\sigma\delta_{j1})(w_j+\omega\delta_{j1})\prod_{\lambda=j}^{m}\left(1+\frac{d_\lambda}{T_\lambda^*}\right)^{-1} \\ &\times\left(1+\frac{d_\lambda}{T_\lambda^*-1}\right)^{-1}. \end{aligned} \tag{6}$$

Variances and covariances follow from (4), (5) and (6) through the known transformation formulae.

In the general case the structure of the pgf for a single color makes mathematical simplification almost impossible. However the factorial moments can be converted to central moments and a four-moment approximating distribution can be fitted; the two main possibilities here are the Pearson system of distributions, and Johnson's translational system (Johnson and Kotz [2]). In both cases percentage points are readily approximated or evaluated exactly.

3. A two-color urn model

The urn initially has σ red balls and ω white balls. At the j-th cycle, only the white balls receive an increment w_j (the number of red balls is not increased) and d_j balls are randomly removed. Then from (4) the factorial moments of the number of red balls (see note on this choice of color in Appendix A) at the m-th cycle are given by:

$$\mu_{[r]}^{[m]}=\prod_{\lambda=0}^{r-1}(\sigma-\lambda)\prod_{j=1}^{m}\left(1+d_j\Big/\Big(\sigma+\omega+\sum_{k=1}^{j}(w_k-d_k)\Big)\right)^{-1}. \tag{7}$$

If the reinforcement at each cycle equals the depletion, and is independent of the cycle, $(w_j=w,\ j=1,2,\cdots,m)$, the factorial moments of the number of red balls at the m-th cycle are given by

$$\mu_{[r]}^{[m]}=\prod_{\lambda=0}^{r-1}(\sigma-\lambda)\left(1+\frac{w}{\sigma-\lambda+\omega}\right)^{-m}. \tag{8}$$

From (8) we are able to find (for $r=1$) (a) how many cycles are needed to reach the half-life of the initial number of red balls and (b) how many cycles are needed to almost completely deplete the mean number of red balls in the urn. In particular, we have for (a), setting $\alpha=w/(\sigma+\omega)$, $\sigma/2=\sigma(1+\alpha)^{-m}$, so that the half-life cycle number is approximately $\ln 2/\ln(1+\alpha)$, and depends only on α.

For (b), approximate complete depletion will occur when

$$\mu_{[1]}^{[m]}\leqq 1 \quad \text{or} \quad m>\ln\sigma/\ln(1+\alpha). \tag{9}$$

The above remarks can find applications in problems concerning the uptake of radioiodine by humans—here the red ball numbers can be large. So in the present simple model, suppose that the urn contain 10^{15} red balls initially. Then Table 1 gives illustrations of the cycles needed to reach half-life and complete depletion of the red balls for varying α and w. $(\omega=0)$. The results are approximations only, being based on the mean number of red balls.

It will be seen that if the white ball reinforcement exceed the number of red balls (α large), then half-life scarcely exists; on the other hand if the number of red balls exceeds the white ball reinforcement (α small) then the half-life is increased and occurs after many cycles.

Table 1. Half-life and depletion of red balls

α	cycles for half-life	cycles for $\mu_{[1]}^{[m]} \leqq 1$	w
100000	<1	3	10^{20}
10000	<1	4	10^{19}
1000	<1	5	10^{18}
100	<1	8	10^{17}
10	<1	15	10^{16}
1	1	50	10^{15}
0.1	8	363	10^{14}
0.01	70	3472	10^{13}

An idea of the distribution of the red balls (mean, variance, skewness, and kurtosis) is given in Table 2.

Table 2. Distribution of number of red balls in urn at each cycle

α	cycle	μ_1	μ_2	$\sqrt{\beta_1}$	β_2
10,000	2	$9.998 \cdot 10^6$	$9.998 \cdot 10^6$	0	3.00
	3	$9.997 \cdot 10^2$	$9.997 \cdot 10^2$	0.03	3.00
1,000	2	$9.98 \cdot 10^8$	$9.98 \cdot 10^8$	0	3.00
	3	$9.97 \cdot 10^5$	$9.97 \cdot 10^5$	0.001	3.00
	4	$9.96 \cdot 10^2$	$9.96 \cdot 10^2$	0.032	3.00
100	2	$9.8 \cdot 10^{10}$	$9.8 \cdot 10^{10}$	0	3.00
	3	$9.7 \cdot 10^8$	$9.7 \cdot 10^8$	0	3.00
	4	$9.6 \cdot 10^6$	$9.6 \cdot 10^6$	0	3.00
	5	$9.5 \cdot 10^4$	$9.5 \cdot 10^4$	0	3.00
	6	$9.4 \cdot 10^2$	$9.4 \cdot 10^2$	0.03	3.00
	7	9.3	9.3	0.33	3.11
10	2	$8.3 \cdot 10^{12}$	$8.1 \cdot 10^{12}$	0	3.00
	4	$6.8 \cdot 10^{10}$	$6.8 \cdot 10^{10}$	0	3.00
	6	$5.6 \cdot 10^8$	$5.6 \cdot 10^8$	0	3.00
	8	$4.7 \cdot 10^6$	$4.7 \cdot 10^6$	0	3.00
	10	$3.9 \cdot 10^4$	$3.9 \cdot 10^4$	0.005	3.00
	12	$3.2 \cdot 10^2$	$3.2 \cdot 10^2$	0.056	3.00
	14	2.6	2.6	0.616	3.38
1	5	$3.13 \cdot 10^{13}$	$2.78 \cdot 10^{13}$	0	3.00
	10	$9.77 \cdot 10^{11}$	$9.71 \cdot 10^{11}$	0	3.00
	15	$3.05 \cdot 10^{10}$	$3.05 \cdot 10^{10}$	0	3.00
	20	$9.54 \cdot 10^8$	$9.54 \cdot 10^8$	0	3.00
	25	$2.98 \cdot 10^7$	$2.98 \cdot 10^7$	0	3.00
	30	$9.31 \cdot 10^5$	$9.31 \cdot 10^5$	0.001	3.00
	35	$2.91 \cdot 10^4$	$2.91 \cdot 10^4$	0.006	3.00
	40	$9.09 \cdot 10^2$	$9.09 \cdot 10^2$	0.033	3.00
	45	$2.84 \cdot 10$	$2.84 \cdot 10$	0.188	3.04
	49	1.78	1.78	0.750	3.56
0.1	7	$5.13 \cdot 10^{14}$	$8.23 \cdot 10^{13}$	0	7.04
	8	$4.67 \cdot 10^{14}$	$9.06 \cdot 10^{13}$	0	3.05
	50	$8.52 \cdot 10^{12}$	$8.12 \cdot 10^{12}$	0	3.04
	100	$7.26 \cdot 10^{10}$	$7.25 \cdot 10^{10}$	0	3.00
	150	$6.18 \cdot 10^8$	$6.18 \cdot 10^8$	0	3.00
	200	$5.27 \cdot 10^6$	$5.27 \cdot 10^6$	0	3.00
	250	$4.49 \cdot 10^4$	$4.49 \cdot 10^4$	0.005	3.00
	300	$3.82 \cdot 10^2$	$3.82 \cdot 10^2$	0.051	3.00
	350	3.26	3.26	0.554	3.31
	362	1.04	1.04	0.982	3.96

Thus if the white ball reinforcement is large there is a rapid reduction in the red ball mean, whereas the decrease in red balls is greatly retained when black and white reinforcements are small compared to the initial number of red balls. Notice that the actual distribution of red ball is in general nearly normal with mean practically equal to the variance.

The case where the reinforcements exceed the depletions and increase with the cycle, is an open problem.

4. A contiguity case

4.1 *General considerations*

Consider the urn scheme of Section 3. There is contiguity in the operand $(x_1^{(d_1)} x_2^{(d_2)} \cdots x_m^{(d_m)})$ (see Appendix B) if

$$w_j = d_{j-1} + d_j \qquad (j=2, 3, \cdots, m),$$

in which case the fmgf for red balls after m cycles reduces to

$$F_m(\alpha) = c_m(\alpha + \mathrm{E}_m \mathrm{E}_{m-1} \cdots \mathrm{E}_1)^\sigma (\omega + w)^{(d_1)} \prod_{k=2}^{m} \left(\omega + w_1 + \sum_{j=2}^{k} d_j\right)^{(d_k)} \tag{10}$$

where E_1 operates on $\omega + w_1$, E_j on $\omega + w_1 + d_2 + d_3 + \cdots + d_j$, $j=2, 3, \cdots, m$. For the probability generating function (pgf) we have

$$P_m(t) = c_m(t+\Delta)^\sigma \left(\omega + w_1 + \sum_{j=2}^{m} d_j\right)^{(d_1+d_2+\cdots+d_m)} \tag{11}$$

where

$$c_m^{-1} = \left(\sigma + \omega + w_1 + \sum_{j=2}^{m} d_j\right)^{(d_1+d_2+\cdots+d_m)}.$$

It will be seen that the distribution of red balls in this 2-color contiguous urn model depends only on the three parameters

$$d_m^* = d_1 + d_2 + \cdots + d_m, \qquad M = \omega + w_1 - d_1, \qquad H_m = M + d_m^* + \sigma,$$

so that the pgf becomes

$$P_m(t) = (1/H_m^{(d_m^*)})(t+\Delta)^\sigma (M + d_m^*)^{(d_m^*)}, \tag{12}$$

a three parameter (d_m^*, M, H_m) hypergeometric distribution. If $R^{[m]}$ is the random variable, then

$$\Pr(R^{[m]} = x) = \binom{\sigma}{x} d_m^{*(\sigma - x)} (M + d_m^*)^{(d_m^* + x - \sigma)} / H_m^{(d_m^*)} \tag{13}$$

$$(x = \sigma, \sigma - 1, \cdots, \max(0, \sigma - d_m^*)).$$

Notice that the total reinforcement in the m cycles is $w_j + d_1 + 2(d_2 + d_3 + \cdots + d_{m-1}) + d_m$ which is approximately twice the total depletion d_m^*

provided w_1 is not large. Notice also that the probability of zero red balls at the m-th cycle is zero unless the total depletion (a determinate value) is at least equal to the initial number of red balls.

The pgf in (12) for a hypergeometric distribution seems to have been overlooked. Moments, up to the fourth central, are given in Kendall and Stuart [3].

4.2 *Numerical illustration*

Let there be 100 red balls initially, and no white balls. For the depletion assume a linear monotonic increasing form $d_j = j$. Then for contiguity, the white ball reinforcements must be

$$w_j = 2j - 1\,, \qquad (j = 2, 3, \cdots)$$

with w_j arbitrary; take it to be 2. Then the pgf after m cycles is

$$P_m(t) = c_m(t+\Delta)^{100}(K_m+1)^{(K_m)}\,, \qquad (m=1, 2, \cdots)\,,$$

where

$$K_m = m(m+1)/2\,, \qquad c_m^{-1} = (K_m+101)^{(K_m)}\,.$$

Moment parameters (Table 3), using expressions in the Appendix B, demonstrate the decrease in the mean number of red balls as the cycles increase and the near-normal form of the distribution, especially at $m = 14$ which is close to the half-life of the initial number of red balls.

Tabe 3. Moment parameters for the distribution of red balls (contiguous case)

Cycle (m)	Mean	Standard deviation	Skewness $\sqrt{\beta_1}$	Kurtosis β_2
5	87.1	1.25	0.436	3.014
13	52.6	3.47	0.006	2.989
14	49.0	3.59	0.000	2.990
25	23.7	3.72	0.075	2.986
35	13.8	3.21	0.164	3.000
40	11.0	2.95	0.207	3.014
50	7.34	2.51	0.291	3.052
100	1.96	1.37	0.673	3.415

(When $m=5$, the red balls lie between 100 and 85; similarly when $m=13$, the red balls lie between 100 and 9. For $m \geqq 14$ the range is 0 to 100.)

Suppose the depletions for this case are doubled, so that $d_j = 2j$ and the reinforcements are $w_j = 4j - 2$, the initial number being the same. Then the total reinforcement after m cycle is $2m^2$, and total depletion $m^2 + m$. Thus the flushing-out process is stronger and after

14 cycles, the four moment parameters are 32.3, 3.85, 0.033, and 2.987; the general form of the distribution has changed little but the mean is significantly reduced. Similarly if the depletions are lighter ($d_j=1$, $w_j=2$) then after 14 cycles, the moment parameters are 87.8, 1.19, 0.480, and 3.037. This suggests that in general, as is intuitively clear, the stronger the flushing-out the more emphatic the decrease in the mean number of red balls.

5. Concluding remarks

The hypergeometric model described here serves as an approximation to the 2-color urn model especially when reinforcements of white balls are about double the random depletions. It has the advantage that complicated expressions involving sums of products which appear in the general case for factorial moments, leading to serious loss of accuracy when converted to central moments, are avoided.

In the practical application, there is interest in the distribution of red balls at half-life, that is when the mean is reduced to $\sigma/2$ approximately. With the contiguous model this occurs when the total depletion at a cycle equals $M+\sigma=\omega+w_1-d_1+\sigma$. From the appendix this corresponds to $p=1/2$ and zero skewness; we assume that the depletions are positive integers, and that there is a positive integer solution to the equation $d_m^*=M+\sigma$. We find for the variance at the half-life

$$\mu_2=(\sigma/4)[(2M+\sigma)/(2M+2\sigma-1)] ,$$

or approximately $\sigma/8$ as $\sigma\to\infty$. Similarly the asymptotic kurtosis is

$$\beta_2\approx 3[1+(7/6-M)/\sigma+(25/6-16M/3+2M^2)/\sigma^2 +(25/2-25M+46M^2/3-4M^3)/\sigma^3] \qquad (\sigma\to\infty) .$$

This expression suggests that the half-life distribution of red balls is nearly normal $N(\sigma/2, \sigma/8)$, provided M is fixed and the initial number is large.

Acknowledgements

The authors wish to thank the referees for their constructive comments.

OAK RIDGE NATIONAL LABORATORY
UNIVERSITY OF GEORGIA

REFERENCES

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

[2] Johnson, N. L. and Kotz, S. (1970). *Continuous Univariate Distributions*—1, Houghton Mifflin Co., Boston.

[3] Kendall, M. G. and Stuart, A. (1969). *The Advanced Theory of Statistics*, Vol. 1, (3rd ed.), Charles Griffin & Co., London and High Wycombe; Hafner Press, New York.

[4] Shenton, L. R. (1981). A reinforcement-depletion urn problem—I. Basic theory, *Bull. Math. Biol.*, **43**, 327-340.

[5] Shenton, L. R. (1983). A reinforcement-depletion urn problem—II. Application and generalization, *Bull. Math. Biol.*, **45**, 1-9.

Appendix A

The 2-color case of Section 3

Because of the linear relation $N_r^{[m]}+N_w^{[m]}=T_m^*$ between the numbers of red and white balls at the m-th cycle one would consider the fmgf for the color whose reinforcements are of smallest density to avoid complicated expressions. Compare for example the form of the fmgf when white balls are considered, namely

$$F_m(0, \beta)=c_m\left(\prod_{j=1}^{m} (\beta+\mathrm{E}_m\,\mathrm{E}_{m-1}\cdots\mathrm{E}_j)^{w_j+\omega\delta_{j1}}\right)x_1^{(d_1)}x_2^{(d_2)}\cdots x_m^{(d_m)}\,,$$

$$(x_j=\sigma-d_{j-1}^*,\ \ j=1, 2,\cdots, m)$$

to that for red balls, namely

$$F_m(\alpha, 0)=c_m(\alpha+\mathrm{E}_m\,\mathrm{E}_{m-1}\cdots\mathrm{E}_1)^{\sigma}y_1^{(d_1)}y_2^{(d_2)}\cdots y_m^{(d_m)}\qquad (y_j=w_j^*-d_{j-1}^*)\,.$$

Appendix B

A hypergeometric-type model and contiguous factorials

The parameters $x_1, x_2,\cdots, x_m$ (see for example, expressions in Appendix A) that are operated upon by the incremental operator $\mathrm{E}_1, \mathrm{E}_2, \cdots, \mathrm{E}_m$ are usually non-negative increasing integers. A single factorial such as

$$x^{(s)}=x(x-1)\cdots(x-s+1)\qquad (x, s=1, 2,\cdots)$$

has contiguity because all integers between x and $x-s+1$ (inclusive) appear once and once only. Similarly a product of such factors, say

$$\varphi(\boldsymbol{x}, \boldsymbol{d})=x_1^{(d_1)}x_2^{(d_2)}\cdots x_m^{(d_m)}$$

where $0<x_1<x_2<\cdots<x_m$, will be called contiguous if it is reducible;

i.e.

$$\varphi(\boldsymbol{x}, \boldsymbol{d}) = x_m^{(d_m^*)} \qquad \left(d_m^* = \sum_1^m d_j\right).$$

For example

$$x_1^{(d_1)}(x_1+d_2)^{(d_2)}(x_1+d_2+d_3)^{(d_3)} = (x_1+d_2+d_3)(x_1+d_2+d_3-1)\cdots(x_1-d_1+1)$$

all elements being positive integers.

Now consider the homogeneous incremental operator

$$E(x_1, x_2, \cdots, x_m) = \mathrm{E}_{x_1}\mathrm{E}_{x_2}\cdots\mathrm{E}_{x_m}.$$

Clearly

$$\begin{aligned}[E(x_1, x_2, \cdots, x_m)]^h x_1^{(d_1)} x_2^{(d_2)} \cdots x_m^{(d_m)} \\ = (x_1+h)^{(d_1)}(x_2+h)^{(d_2)}\cdots(x_m+h)^{(d_m)} = (x_m+h)^{(d_m^*)}.\end{aligned}$$

Thus under the homogeneous operator $E(x_1, x_2, \cdots, x_m)$ a contiguous factorial remains contiguous. The corresponding difference operator δ in the multivariate case has similar properties to the advancing operator Δ in the univariate case provided we operate on contiguous factorials. Thus

$$\begin{aligned}\delta^h(x_1, x_2, \cdots, x_m) &= [E(x_1, x_2, \cdots, x_m) - 1]^h x_1^{(d_1)} x_2^{(d_2)} \cdots x_m^{(d_m)} \\ &= \sum_{j=0}^{h} (-1)^j \binom{h}{j} E^{h-j}[x_1^{(d_1)} x_2^{(d_2)} \cdots x_m^{(d_m)}] \\ &= \Delta^h x_m^{(d_m^*)} = (d_m^*)^{(h)} x_m^{(d_m^*-h)}.\end{aligned}$$

As for moments of the hypergeometric distribution the first three are derived easily enough using the algebraic language "REDUCE". The fourth had to be rearranged from a development based on the moment generating function. For example if $X = R^{[m]}$ then

$$\begin{aligned}E[\exp(X-\mu_1^*)] &= (1/H_m^{(d_m^*)})(e^{\alpha q} + \Delta e^{-\alpha p})^{\sigma}(M+d_m^*)^{(d_m^*)} \\ &= (1/H_m^{(d_m^*)})(\mathrm{E}^{-1}e^{\alpha q} + \mathrm{E}^{-1}\Delta e^{-\alpha p})^{\sigma} H_m^{(d_m^*)}\end{aligned}$$

where

$$p = (M+\sigma)/(M+\sigma+d_m^*), \qquad q = 1-p.$$

The operational component expands in the form

$$[1 + \alpha\,\mathrm{E}^{-1}(q-p\Delta) + \alpha^2\,\mathrm{E}^{-1}(q^2+p^2\Delta)/2! + \alpha^3\,\mathrm{E}^{-1}(q^3-p^3\Delta)/3! + \cdots]^{\sigma}.$$

This approach requires the development of an algebra based on elements such as

$$\{\mathrm{E}^{-r-s-t\cdots}[(q-p\Delta)^r(q^2+p^2\Delta)^s(q^3-p^3\Delta)^t\cdots]\}H_m^{(d_m^*)}.$$

$r, s, t, \cdots$ being non-negative integers.