A NOTE ON A FURTHER GENERALIZATION OF THE EHRENFEST URN MODEL*

V. R. R. Uppuluri and Tommy Wright, Union Carbide Corporation

1. Introduction

The model discussed in this paper was motivated by the following problem. Assume a finite population which consists of N elements. It is known that w_0 of the elements have a particular characteristic and $N-w_0$ of the elements do not. The desire is to change the number of elements in the population which have the particular characteristic while keeping the overall population size, N, constant. This change may mean a decrease or an increase in w_0. Changes are accomplished by randomly selecting one element at a time and replacing it with a similar or dissimilar element.

The model presented is a generalization of the Ehrenfest Model (1907) and is appealing because at each trial it leaves the choice of replacement to chance while at the same time achieves a predetermined goal concerning the desired proportion of elements possessing the characteristic of interest.

In Section 2, a brief overview of relevant modifications of the Ehrenfest Model are presented. In Section 3, we present a new sampling scheme for the establishment of goals which is a further modification of the Ehrenfest Model. Numerical examples illustrating the new sampling scheme are given in Section IV.

2. The Ehrenfest Model and Some of Its Modifications

One model which can be used to address the general problem described in Section 1 is the Ehrenfest Model. In this section, we briefly review the Ehrenfest Model and some of its modifications as discussed by Johnson and Kotz (1977).

The basic Ehrenfest Model is presented under Model 1.

Model 1. Assume an urn with w_0 white balls and b_0 black balls ($w_0+b_0=N$). At each trial, a ball is drawn at random.

If it is white, replace it by a black ball with probability 1.
If it is black, replace it by a white ball with probability 1.

After the n^{th} trial, let W_n denote the number of white balls in the urn. It can be shown that as $n\to\infty$, the limiting distribution of W_n has a binomial distribution with parameters N and $^1/_2$. One can also show that

$$E(W_n) = \frac{N}{2} + (w_0 - \frac{N}{2})(1 - \frac{2}{N})^n \qquad (1)$$

and

$$Var(W_n) = \frac{N}{4}\{1 - (1 - \frac{2}{N})^{2n}\} + \frac{1}{4}\{N^2 - N\}\{(1 - \frac{4}{N})^n - (1 - \frac{2}{N})^{2n}\} \qquad (2)$$

A slight modification of Model 1 is given by Johnson and Kotz (1977) which permits some randomization before each replacement. This model is summarized under Model 2.

Model 2. As in Model 1, we have an urn with w_0 white balls and b_0 black balls. At each trial a ball is drawn at random.

If it is white, replace it by:
{a black ball with probability α.
{a white ball with probability $1-\alpha$.

If it is black, replace it by:
{a white ball with probability α.
{a black ball with probability $1-\alpha$.
$(0\leq\alpha\leq1)$

As in Model 1, W_n is the number of white balls in the urn after the n^{th} trial. It can be shown that as $n\to\infty$, the limiting distribution of W_n is again binomial with parameters N and $^1/_2$. Thus the limiting distribution of W_n does not depend on α. We also note that

$$E(W_n) = \frac{N}{2} + (w_0 - \frac{N}{2})(1 - \frac{2\alpha}{N})^n \qquad (3)$$

and

$$Var(W_n) = \frac{N}{4}\{1 - (1 - \frac{2\alpha}{N})^{2n}\} + \frac{1}{4}\{N^2 - N\}\{(1 - \frac{4\alpha}{N})^n - (1 - \frac{2\alpha}{N})^{2n}\} \qquad (4)$$

Other modifications exist in the literature and are discussed by Johnson and Kotz (1977). We make only brief mention of some of them. One model was presented by Karlin and McGregor (1965). They assume an urn with w_0 white balls and b_0 black balls. T_i, the time between the $(i-1)^{th}$ trial and the i^{th} trial, is a random variable which is independent of T_j (where $i\neq j$) with distribution given by $P(T_i>t) = e^{-\lambda t}$ $(\lambda,t>0)\forall i$. Under these assumptions $N_{(\tau)}$, the number of trials completed in a fixed time τ, is Poisson with parameter $\lambda\tau$. As each trial occurs, the experimenter proceeds as under Model 2. If W_τ is the number of white balls in the urn after time τ, one can show:

$$E(W_\tau) = \frac{N}{2} + (w_0 - \frac{N}{2})e^{-\frac{2\lambda\tau\alpha}{N}}, \qquad (5)$$

and

$$Var(W_\tau) = \frac{N}{4}(1 - \alpha e^{-\frac{4\lambda\tau}{N}}). \qquad (6)$$

The Karlin and McGregor (1965) model differs from models 1 and 2 because the number of trials is a random variable. In models 1 and 2, the experimenter has control over the number of trials, whereas in the Karlin and McGregor model, the experimenter has control over the time.

Another modification of the Ehrenfest Urn Model is given by Vincze (1964) where before each trial, the experimenter decides with probability π not to draw a ball. Thus after n trials, the effective number of selection is a binomial random variable with parameters n and $1-\pi$.

The Ehrenfest Model and some of its modifications have found application in areas including heat exchange problems and genetics.

The modification which we consider in the next section is a simple extension which is consistent with the spirit of models 1 and 2.

3. A New Sampling Scheme for the Establishment of Goals.

As in models 1 and 2 discussed in the previous section, we assume an urn with w_0 white balls and b_0 black balls. The total number of balls ($N=w_0+b_0$) in the urn is constant.

At the first trial, select a ball at random.

If it is white, replace it by:
- a black ball with probability α_1.
- a white ball with probability $1-\alpha_1$.

If it is black, replace it by:
- a white ball with probability α_2.
- a black ball with probability $1-\alpha_2$.

Below we define the random variables X_n, η_1, η_2, and W_n which will be used in finding $E(W_n)$, the expected number of white balls in the urn after n trials.

Let
- X_1 be the number of white balls selected on the first draw,
- η_1 be the number of white balls replaced if a white ball has been selected, and
- η_2 be the number of white balls replaced if a black ball has been selected.

It is easy to see that X_1, η_1, and η_2 are all Bernoulli random variables with parameters w_0/N, $1-\alpha_1$, and α_2 respectively. Further, let

- W_1 be the number of white balls in the urn after the first trial and
- B_1 be the number of black balls in the urn after the first trial. (Note $B_1=N-W_1$.)

We observe that

$$W_1 = w_0 - \{\text{number of white balls which were replaced by black balls}\} + \{\text{number of black balls which were replaced by white balls}\}$$
$$= w_0 - X_1(1-\eta_1) + (1-X_1)\eta_2 \qquad (7)$$

Since $N = w_0 + b_0 = W_1 + B_1$, we have

$$B_1 = b_0 + X_1(1-\eta_1) - (1-X_1)\eta_2 \qquad (8)$$

In matrix form, we have

$$\begin{bmatrix} W_1 \\ B_1 \end{bmatrix} = \begin{bmatrix} w_0 \\ b_0 \end{bmatrix} + \begin{bmatrix} -(1-\eta_1) & \eta_2 \\ (1-\eta_1) & -\eta_2 \end{bmatrix}\begin{bmatrix} X_1 \\ 1-X_1 \end{bmatrix} \qquad (9)$$

Assuming independence between the random variable X_1 and the random variables η_1 and η_2, we have

$$\mu_1 = E\begin{bmatrix} W_1 \\ B_1 \end{bmatrix} = \begin{bmatrix} w_0 \\ b_0 \end{bmatrix} + \begin{bmatrix} -(1-E(\eta_1)) & E(\eta_2) \\ 1-E(\eta_1) & -E(\eta_2) \end{bmatrix}\begin{bmatrix} E(X_1) \\ 1-E(X_1) \end{bmatrix}$$
$$= \begin{bmatrix} w_0 \\ b_0 \end{bmatrix} + \begin{bmatrix} -\alpha_1 & \alpha_2 \\ \alpha_1 & -\alpha_2 \end{bmatrix}\begin{bmatrix} \frac{w_0}{N} \\ \frac{b_0}{N} \end{bmatrix}$$
$$= \left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1}{N}\begin{bmatrix} -\alpha_1 & \alpha_2 \\ \alpha_1 & -\alpha_2 \end{bmatrix}\right\}\begin{bmatrix} w_0 \\ b_0 \end{bmatrix}$$
$$= \{I + \tfrac{1}{N}A\}\mu_0 \qquad (10)$$

Similarly, for n=2, we have

$$W_2 = W_1 - X_2(1-\eta_1) + (1-X_2)\eta_2$$

and

$$B_2 = B_1 + X_2(1-\eta_1) - (1-X_2)\eta_2$$

In matrix form, we have

$$\begin{bmatrix} W_2 \\ B_2 \end{bmatrix} = \begin{bmatrix} W_1 \\ B_1 \end{bmatrix} + \begin{bmatrix} -(1-\eta_1) & \eta_2 \\ 1-\eta_1 & -\eta_2 \end{bmatrix}\begin{bmatrix} X_2 \\ 1-X_2 \end{bmatrix}$$

Hence

$$\mu_2 = E\begin{bmatrix} W_2 \\ B_2 \end{bmatrix} = \mu_1 + \frac{1}{N}\begin{bmatrix} -\alpha_1 & \alpha_2 \\ \alpha_1 & -\alpha_2 \end{bmatrix}\begin{bmatrix} E(W_1) \\ E(B_1) \end{bmatrix}$$
$$= \{I + \tfrac{1}{N}A\}\mu_1$$
$$= \{I + \tfrac{1}{N}A\}^2\mu_0$$

In general after the n^{th} trial, we have

$$\begin{bmatrix} W_n \\ B_n \end{bmatrix} = \begin{bmatrix} W_{n-1} \\ B_{n-1} \end{bmatrix} + \begin{bmatrix} -(1-\eta_1) & \eta_2 \\ (1-\eta_1) & -\eta_2 \end{bmatrix}\begin{bmatrix} X_n \\ 1-X_n \end{bmatrix}$$

where X_n, the number of white balls selected on the n^{th} draw, has a Bernoulli conditional distribution with parameter W_{n-1}/N and W_n is the number of white balls in the urn after the n^{th} draw and randomized replacement. It follows that

$$\mu_n = E\begin{bmatrix} W_n \\ B_n \end{bmatrix} = \{I + \tfrac{1}{N}A\}^n \mu_0 \qquad (11)$$

The formula (11) gives the expected number of white balls (and black balls) that one can expect to have in the urn after n trials.

Thus the sampling scheme is a tool for the attainment of certain goals which one may want to accomplish over time. The tool is appealing because it leaves the choice of replacement at each trial to chance while at the same time it achieves a predetermined goal concerning the desired proportion of white balls.

Note that the new sampling scheme is a generalization of models 1 and 2. Recall that as $n \to \infty$, $E(W_n) \to N/2$ for every value of α. This suggests that the experimenter has severely limited control over the achievement of predetermined goals under models 1 and 2. It also suggests that if a goal is achieved at trial n-1, one will need to constantly vary α in order to maintain the desired proportion of white balls in the urn.

The sampling scheme which we have presented gives the experimenter greater control over the rate at which a goal is achieved by having him choose α_1 and α_2 instead of, say simply α.

Given that we have W_{n-1} white balls at trial n-1, one can find the conditional distribution of $W_n|W_{n-1}$ from the following table:

W_n	$W_{n-1}-1$	W_{n-1}	$W_{n-1}+1$
$P(W_n \mid W_{n-1})$	$\frac{\alpha_1 W_{n-1}}{N}$	$1 - \frac{\alpha_1 W_{n-1}+\alpha_2 B_{n-1}}{N}$	$\frac{\alpha_2 B_{n-1}}{N}$

Then we have

$$E(W_n|W_{n-1}) = \left(1 - \frac{\alpha_1+\alpha_2}{N}\right)W_{n-1} + \alpha_2 \qquad (12)$$

and

$$\mathrm{Var}(W_n|W_{n-1}) = \left(\frac{3\alpha_1+\alpha_2}{N}\right)W_{n-1} + (\alpha_2 - 2W_{n-1})(1-\alpha_2) - \left(\frac{\alpha_1+\alpha_2}{N}\right)^2 W_{n-1}^2 \qquad (13)$$

When the desired goal is achieved, it can be maintained on the average by selecting new α_1 and α_2 as indicated below. Assume that the desired goal is achieved on the $(n-1)^{th}$ trial and that it is W^*_{n-1}. If we want α_1 and α_2 which will give $E(W_n) = W^*_{n-1}$, we take (12) and obtain

$$W^*_{n-1} = \left(1 - \frac{\alpha_1+\alpha_2}{N}\right)W^*_{n-1} + \alpha_2 .$$

Solving for α_1 gives

$$\alpha_1 = \left(\frac{N}{W^*_{n-1}} - 1\right)\alpha_2 \qquad (14)$$

Thus to maintain the goal achieved on the $(n-1)^{th}$ trial, take $\alpha_2' = \alpha_2$ and

$$\alpha_1' = \left(\frac{N}{W^*_{n-1}} - 1\right)\alpha_2' . \qquad (15)$$

In practice in order to establish the stated goal of $E(W_n)$, for given N, w_0, and n, the experimenter chooses α_1 and α_2 which will satisfy (11). Once that goal is achieved, it is maintained by choosing α_1 and α_2 according to (15). Some numerical examples are given in the next section.

Further efforts are being made to find α_1 which will give the minimum n for $\alpha_2 = \alpha_1 + d$, where d is some positive real number such that $\alpha_2 - \alpha_1 \leqslant d$ and $\alpha_2 > \alpha_1$.

4. A Numerical Example

Blatz (1968) obtained a general expression for any integer power of a 2×2 matrix

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

He showed that for any positive integer n,

$$M^n = \frac{\lambda_1^n - \lambda_2^n}{\lambda_1 - \lambda_2} M - \frac{\lambda_1^{n-1} - \lambda_2^{n-1}}{\lambda_1 - \lambda_2} \lambda_1 \lambda_2 I \qquad (16)$$

where λ_1 and λ_2 are eigenvalues of M and I is the 2×2 identity matrix. Applying (16), one can show that the matrix

$$\left(I + \frac{1}{N}A\right)^n$$

is equal to

$$\frac{1}{\lambda_1-\lambda_2}\begin{bmatrix} (\lambda_1^n-\lambda_2^n)a - (\lambda_1^{n-1}-\lambda_2^{n-1})\lambda_1\lambda_2 & (\lambda_1^n-\lambda_2^n)b \\ (\lambda_1^n-\lambda_2^n)c & (\lambda_1^n-\lambda_2^n)d - (\lambda_1^{n-1}-\lambda_2^{n-1})\lambda_1\lambda_2 \end{bmatrix}$$

where

$$a = 1 - \frac{\alpha_1}{N},\quad b = \frac{\alpha_2}{N},\quad c = \frac{\alpha_1}{N},\quad d = 1 - \frac{\alpha_2}{N},$$

$$\lambda_1 = \frac{(a+d)+\sqrt{(a-d)^2+4bc}}{2} \text{ and } \lambda_2 = \frac{(a+d)-\sqrt{(a-d)^2+4bc}}{2}$$

Thus from (11), we have

$$E(W_n) = \frac{1}{\lambda_1-\lambda_2}\left[\{(\lambda_1^n-\lambda_2^n)a - (\lambda_1^{n-1}-\lambda_2^{n-1})\lambda_1\lambda_2\}w_0 + \{(\lambda_1^n-\lambda_2^n)b\}b_0\right]. \qquad (17)$$

Consider an urn with $w_0 = 90$ white balls and $b_0 = 10$ black balls (N=100). We first assume that the experimenter wants to achieve a goal of $10(=E(W_n))$ white balls after n=500 trials. To determine α_1 and α_2 which will give the result on the average, we choose values for α_1 between 0 and 1 and solve (17) for corresponding values of α_2. The results are indicated in Figure 1 for n=500. Similar results are also given for n=1,000 and n=1,500.

ORNL –DWG 81–18268

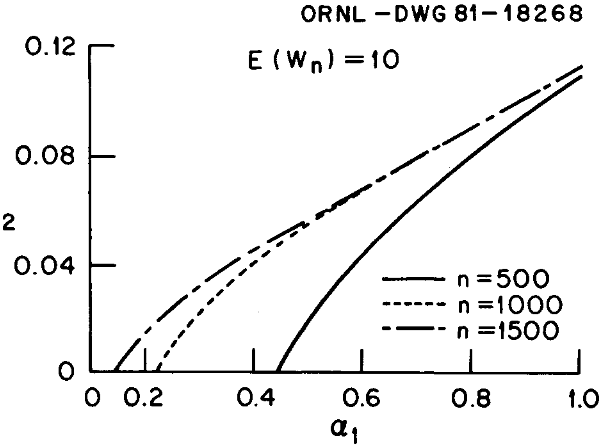


FIGURE 1

From Figure 1, we observe for n=500, that one pair that will accomplish the stated goal on the average is (α_1,α_2) = (.6,.043). For fixed α_1, as n increases, we observe that α_2 increases. Note for n=500, that if $\alpha_1 < .4$, then there is no value of α_2 which will accomplish the goal on the average. Similar results are true for n=1,000 and $\alpha_1<.2$ and for n=1,500 and $\alpha_1<.1$. This is not surprising when one considers the definition of α_1 and α_2 and the stated goal.

In Figure 2, we consider the same initial urn conditions, but the goal is 25 white balls after n trials. As in Figure 1, we take n=500; 1,000; and 1,500. Similar results are given in Figures 3, 4, 5, and 6 for the goals of 50, 75, 90 and 95 respectively.

Generalizations to urn models with more than two categories appear to be straightforward.

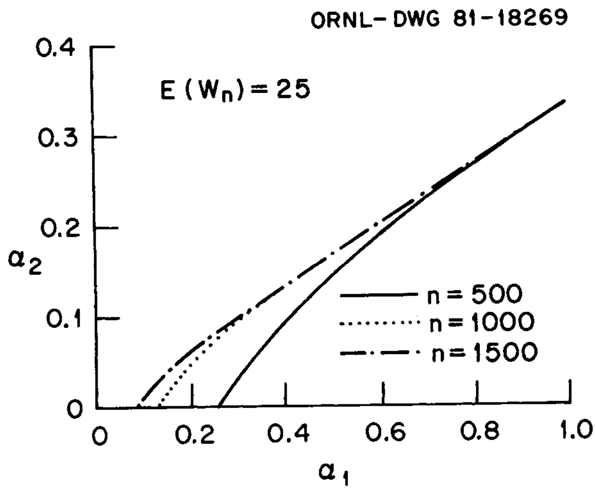


FIGURE 2

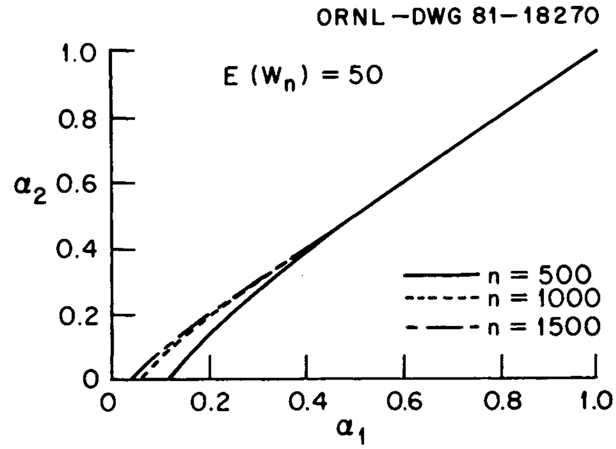


FIGURE 3

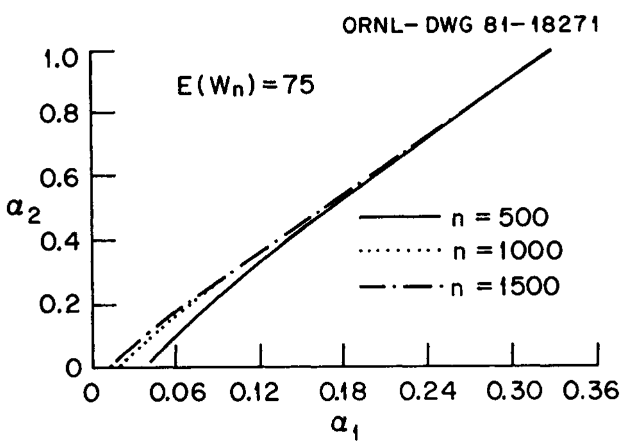


FIGURE 4

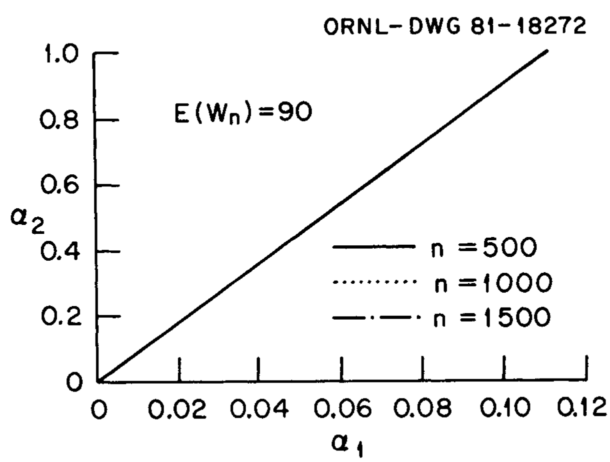
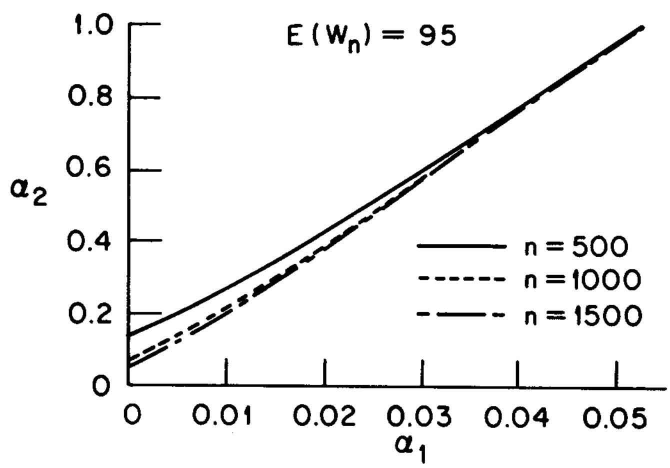


FIGURE 5

ORNL-DWG 81-18273



REFERENCES

1. Blatz, P.J. (1968) On the Arbitrary Power of an Arbitrary 2×2 Matrix, American Mathematical Monthly, 75, 57-58.

2. Ehrenfest, P. and Ehrenfest, T. (1907) Uber Zwei beKannte Einwände gegen das Boltzmannsche H-Theorem, Physik. Z, 8, 311-314.

3. Johnson, N.L. and Kotz, S. (1977) Urn Models and Their Applications, New York: John Wiley and Sons.

4. Karlin, S. and McGregor, J. (1965) Ehrenfest Urn Models, J. Appl. Prob., 2, 352-376.

*Research sponsored jointly by the Applied Mathematical Sciences Research Program, Office of Energy Research, U.S. Department of Energy under contract W-7405-eng-26 with the Union Carbide Corporation and by the Office of Nuclear Regulatory Research of the Nuclear Regulatory Commission.