

HOW COMMON RANDOM NUMBERS AFFECT MULTINOMIAL SELECTION

J. O. Miller
Kenneth W. Bauer, Jr.

Department of Operational Sciences
Air Force Institute of Technology
Wright-Patterson AFB, OH 45433, U.S.A.

ABSTRACT

We consider the effect of common random numbers (CRNs) among simulated systems in two different multinomial selection procedures: the classical procedure BEM (Bechhofer, Elmaghraby, and Morse) and procedure AVC (all vector comparisons). We examine a simple two population scenario and demonstrate analytically how CRNs affect the probability of correct selection (PCS). In addition, various levels of positive correlation are induced to approximate the effect of using CRNs with complex simulated systems. We present analytical and simulation results to show the effect correlation has on the probability of correct selection (PCS) for each solution procedure. This paper responds to Jim Wilson's question raised at WSC'96 about how the use of CRNs affects multinomial selection procedures.

1 INTRODUCTION AND MOTIVATION

Motivating Example: As tactical war planning analysts, we are directed to provide the Joint Task Force Commander with the best plan to cripple the enemy's command and control. "Best" means achieving the highest level of cumulative damage expectancy (CDE) against a selected set of targets given current intelligence estimates of enemy defense capabilities and available friendly forces. Using a theater-level combat simulation model, our team prepares four independent attack plans and we simulate r replications across all four plans. For each replication we compare the CDE between each of the four plans. Since the chosen plan can only be executed a single time, we select as the best plan the one that has the largest CDE in most of the replications.

In such a scenario, the argument can be raised that CRNs should be used to test the different plans under the same experimental conditions—where our experimental conditions are the random variates used to

determine various system reliabilities, probabilities of target damage, etc. When estimating the difference in system performance, CRNs reduce the variance of the estimator by inducing a positive correlation among the individual system responses. In a multinomial selection problem (MSP), we only care about which system is best in each trial and not the magnitude of the difference between the responses. Mata (1993) shows the use of CRNs increases the PCS in an empirical study, or in other words, sharpens the comparison between some simple systems in an MSP. We look at one of Mata's examples and demonstrate analytically how CRNs can increase the PCS with BEM.

The use of CRNs introduces a dependency among the systems, with the level of this dependency varying with the complexity of the systems and how they are modeled. We explore the effects of CRNs in MSPs involving complex systems through inducing selected levels of correlation among simple simulated systems.

The paper is organized as follows: We first provide a brief review of the MSP and two solution procedures, BEM and AVC. We then present analytical results on the effects of CRNs working with a model from Mata (1993) and with some simple modifications to this model. Empirical results are also given for these examples. Finally, additional empirical results show the effects on PCS for both BEM and AVC over a range of positive correlations.

2 BACKGROUND

We consider the general problem of selecting the best of $k \geq 2$ independent populations, $\pi_1, \pi_2, \dots, \pi_k$, where in our context "populations" is taken to mean simulated systems. This is known as the MSP. Let $\mathbf{X}_i = (X_{1i}, X_{2i}, \dots, X_{ki})$ represent a vector of independent observations of some common performance measure across all populations on the i^{th} replication. For each i , the best population is the population with

the largest X_{ji} . The goal is to find the population that is most likely to be the best performer among the populations, as opposed to identifying the best average performer over the long run.

Let $Y_{ji} = 1$ if $X_{ji} > X_{\ell i}$, for $\ell = 1, 2, \dots, k$, but $\ell \neq j$; and let $Y_{ji} = 0$ otherwise. In other words, $Y_{ji} = 1$ if X_{ji} is the largest observation in \mathbf{X}_i . In case of a tie for the largest value, we randomly select one of the tied populations as the best. Suppose that there are v independent replications across all populations, and let $Y_j = \sum_{i=1}^v Y_{ji}$ represent the number of times population j wins out of these v replications. So $\sum_{j=1}^k Y_j = v$ and the k -variate discrete random variable $\mathbf{Y} = (Y_1, Y_2, \dots, Y_k)$ follows a multinomial distribution with $p_j = \Pr\{\pi_j \text{ wins}\}$, $j = 1, 2, \dots, k$, where $0 < p_j < 1$, $j = 1, 2, \dots, k$, and $\sum_{j=1}^k p_j = 1$.

Bechhofer, Elmaghraby and Morse (1959) describe a single-stage procedure for selecting the multinomial event (population) which has the largest success probability. BEM requires the specification of P^* (where $1/k < P^* < 1$), a minimum probability of correctly identifying the population with the largest success probability (i.e., the best population), and θ^* (where $1 < \theta^* < \infty$), a minimum ratio of the largest success probability to the second largest success probability. The procedure consists of the following steps:

Procedure BEM

1. For given k and θ^* , find the minimum value of v , denoted as v^* , that guarantees that the PCS is at least P^* .
2. Generate v^* independent replications for each population.
3. Compute $Y_j = \sum_{i=1}^{v^*} Y_{ji}$, for $j = 1, 2, \dots, k$.
4. Let $Y_{(1)} \leq Y_{(2)} \leq \dots \leq Y_{(k)}$ be the ranked sample counts from step 3. Select the population associated with the largest count, $Y_{(k)}$, as the best population. In case of a tie for the largest count, randomly select one of the tied populations as the best.

To determine the appropriate v^* in step 1, let $p_{[1]} \leq p_{[2]} \leq \dots \leq p_{[k]}$ denote the ranked success probabilities for the k populations. Since only values of the ratio $\theta = p_{[k]}/p_{[k-1]}$ greater than or equal to θ^* are of interest, we are indifferent between the best and the next-best population for values of $\theta < \theta^*$. A procedure of this type is referred to as an *indifference-zone approach*. Select v^* as the minimum number of independent vector observations required to achieve a probability of correct selection (PCS) greater than

or equal to P^* whenever $\theta \geq \theta^*$. We refer to the PCS with BEM as PCS^{bem} .

Miller, Nelson, and Reilly (1997) introduced a method shown to provide a PCS greater than or equal to PCS^{bem} (for specific small sample cases and asymptotically in general) using the same replications \mathbf{X}_i , $i = 1, 2, \dots, v$. With BEM parameters k , P^* , and θ^* , the first step of BEM is used to find a value of v^* . Then a total of $(v^*)^k$ pseudo-replications are formed by associating each X_{ji} ($j = 1, 2, \dots, k$; $i = 1, 2, \dots, v^*$), with all possible combinations of the remaining $X_{\ell h}$ ($\ell = 1, 2, \dots, k$; $\ell \neq j$; $h = 1, 2, \dots, v^*$). Each such pseudo-replication contains one observation from each population. Note that the $(v^*)^k$ pseudo-replications include the v^* independent replications from which the pseudo-replications are formed.

Procedure AVC

1. Given values for k , P^* , and θ^* , use step 1 of procedure BEM to determine a value for v^* .
2. Generate v^* independent replications for each population and construct the additional $(v^*)^k - v^*$ pseudo-replications possible with one value from each of the populations.
3. Compute Z_j as follows:

$$Z_j = \sum_{a_1=1}^{v^*} \sum_{a_2=1}^{v^*} \dots \sum_{a_k=1}^{v^*} \prod_{\ell=1, \ell \neq j}^k \phi(X_{ja_\ell}, X_{\ell a_\ell})$$

for $j = 1, 2, \dots, k$, where $\phi(X_{ja_\ell}, X_{\ell a_\ell})$ is an indicator function that takes the value 1 only if X_{ja_ℓ} is larger than $X_{\ell a_\ell}$.

4. Let $Z_{(1)} \leq Z_{(2)} \leq \dots \leq Z_{(k)}$ be the ranked sample counts from step 3. Select the population associated with the largest count, $Z_{(k)}$, as the best population. In case of a tie for the largest count, randomly select one of the tied populations as the best.

We refer to the PCS with AVC as PCS^{avc} .

3 ANALYTICAL RESULTS

We restrict our analysis to $k = 2$ populations and consider the model used for Example 1 in Mata (1993). Let X represent an observation from $\pi_{[2]}$ (the best population) and let O represent an observation from $\pi_{[1]}$. Let $X \sim \exp(\lambda)$ and $O \sim U(0, a)$ where $\lambda > 0$

and $a > 0$. To calculate $\Pr\{X > O\}$ we have the following expression conditioning on O :

$$\int_0^a e^{-\lambda o} \frac{1}{a} do = \frac{1 - e^{-\lambda a}}{\lambda a}. \quad (1)$$

In a similar fashion to find $\Pr\{O > X\}$ we have

$$\int_0^a (1 - e^{-\lambda o}) \frac{1}{a} do = \frac{e^{-\lambda a} + \lambda a - 1}{\lambda a}. \quad (2)$$

We then combine expressions from (1) and (2) to obtain

$$\theta = \frac{\Pr\{X > O\}}{\Pr\{O > X\}} = \frac{1 - e^{-\lambda a}}{e^{-\lambda a} + \lambda a - 1}. \quad (3)$$

Using $\theta = 1.3$ to match Mata (1993), we set $\lambda = 1.0$ and solve (3) for a to arrive at $a = 1.2747$. We use these parameter values in our simulation runs of this model.

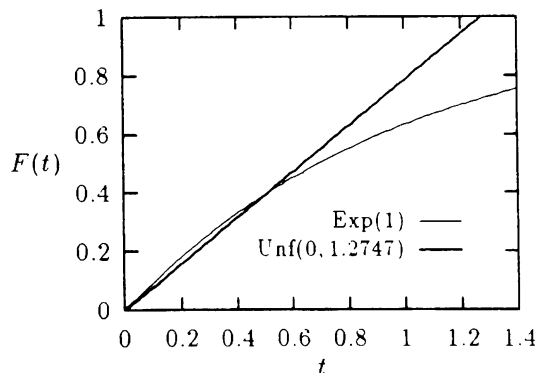


Figure 1: $F(t)$ for Best Exponential with Uniform at $\theta = 1.3$

Figure 1 shows the cumulative density functions (CDFs), denoted as $F(t)$, for the exponential and uniform populations of Example 1 from Mata (1993). By construction, these CDFs provide us with the desired values for $\Pr\{X > O\}$ (0.5652) and $\Pr\{O > X\}$ (0.4348) to yield $\theta = 1.3$. Generating random variates from each distribution using an inverse transform with independent random number streams, we observe our simulated PCS^{bem} results match standard table values for $\theta = 1.3$ (Bechhofer, 1959).

A closer look at Figure 1 illustrates what happens when using CRNs. Note that the two CDFs cross at approximately $F(t) = 0.4$, with the exponential CDF larger for values of $F(t)$ less than this point, and the uniform CDF larger for values above this point. Generating random variates from each distribution

using an inverse transform and CRNs, approximately 60% of the time (for any random number > 0.4) the resulting exponential variate will be larger than the uniform variate. So by using CRNs we have increased $\Pr\{X > O\}$ and decreased $\Pr\{O > X\}$. The “blocking effect” offered by CRNs has provided greater clarification of the selection problem. In other words, CRNs has increased our θ from 1.3 to $.6/.4 = 1.5$. We use this new θ value to compare standard PCS^{bem} at $\theta = 1.5$ with simulation results at $\theta = 1.3$ using CRNs in the following section.

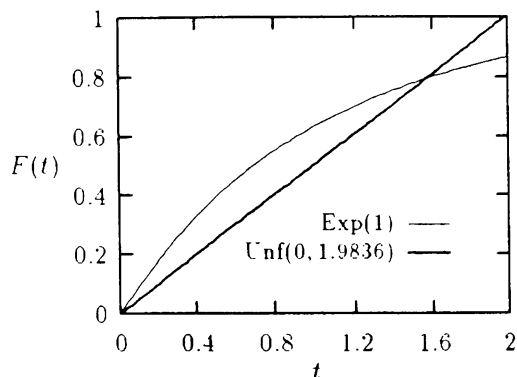


Figure 2: $F(t)$ for Best Uniform with Exponential at $\theta = 1.3$

For another example, we use an exponential and a uniform population, but this time model the uniform population as best. We again set $\lambda = 1.0$ and $\theta = 1.3$ and solve for a , yielding $a = 1.9836$. Simulation results for PCS^{bem} using independent random number streams match standard tables at $\theta = 1.3$. In estimating the effect of CRNs in this scenario from Figure 2, we observe that approximately 79% of the time the resulting uniform variate will be larger than the exponential variate. Here our θ increases from 1.3 to about 3.76. We compare simulation results in the following section with standard BEM values to verify this predicted increase in θ .

As a final illustration of the effect of CRNs, consider the populations used in the first example above but reduce θ to 1.2. Again setting $\lambda = 1.0$ and solving for a , we obtain $a = 1.3653$. In plotting the CDFs we see the same relationship as in Figure 1, except the point where the CDFs cross is now approximately $F(t) = 0.4825$, which gives us $\theta = 0.5175/0.4825 \approx 1.07$. So for this model the use of CRNs has reduced θ , which in turn will reduce PCS^{bem}. As with our previous examples, we compare simulation results in the following section to verify this decrease in θ .

The most important observation to make from the above examples, is that the use of CRNs does not invalidate any of the assumptions necessary for the MSP. We still have independence between replications and the probability that population j wins, p_j , $j = 1, 2$, is the same on every replication. However, the values of the p_j s are changed through the use of CRNs. For the examples used here, the resulting changes can either increase or decrease θ , resulting in an increase or decrease respectively in PCS^{bem} for a given v .

At this time we have not come up with an approach to examine the effect on PCS^{avc} analytically. Intuitively, we expect a similar change in PCS^{avc} for the above examples since the same v independent replications are included in the v^k pseudo-replications. However, as v gets even moderately large, the remaining $v^k - v$ pseudo-replications without CRNs will significantly outnumber the v replications with CRNs, diluting any apparent effect derived from CRNs.

It is also important to note that the output of our simulation is exactly defined by the CDFs used for our random variate generation. Therefore, the full effect of CRNs is translated into our model results. Clearly, with even the most simple simulated systems, the effect of CRNs will be filtered by other interactions in the model.

In the following section, we explore what effect CRNs might have on more complex systems, by inducing a range of correlations between our simple system outputs to approximate various levels of filtering the effect of CRNs.

4 EMPIRICAL RESULTS

Mirroring our analytical discussion we focus on $k = 2$ populations, and initially present results with $X \sim \exp(1)$ and $O \sim U(0, 1.2747)$. As a first step, we calculate PCS^{bem} and PCS^{avc} with and without CRNs. All simulation results are for 100,000 macro-replications yielding PCS values with standard errors of approximately 0.0015. We use a maximum v of 27 to coincide with Mata's study (Mata, 1993). Our results are presented in Table 1 and clearly show an increase in both PCS^{bem} and PCS^{avc} with CRNs.

It is important to note here that the PCS^{bem} results without CRNs correspond to Mata as well as with standard BEM tables (Bechhofer, 1995). However, the PCS^{bem} values using CRNs are significantly different from Mata's results. Specifically, Mata reports a simulated $\text{PCS}^{\text{bem}} = 1$ at $v = 27$, while our simulated results show $\text{PCS}^{\text{bem}} = 0.86$.

Our simulation results are supported by our analytical results. In fact our simulated results using CRNs

at $\theta = 1.3$ are nearly identical (average absolute difference of 0.005) to the exact values at $\theta = 1.5$ for $v = 1, 2, \dots, 27$ (Bechhofer, 1959). A similar comparison of simulated results using CRNs at $\theta = 1.2$ with exact values at $\theta = 1.07$ (generated using a FORTRAN code provided by Goldsman 1995), yield an average absolute difference of 0.004. Finally from our model with the uniform population as best, our simulated PCS^{bem} values at $\theta = 1.3$ using CRNs have an average absolute difference of 0.0008 with the exact PCS^{bem} values at $\theta = 1.5$ for $v = 1, 2, \dots, 27$. These simulation results all support our predicted changes in θ with CRNs. In addition, the very small difference between the PCS values over a range of v also supports our conclusion that we still have an MSP.

Table 1: PCS^{bem} and PCS^{avc} With and Without CRNs

v	PCS^{bem}		PCS^{avc}	
	No CRN	CRN	No CRN	CRN
1	0.5657	0.6004	0.5658	0.6004
3	0.5995	0.6509	0.6085	0.6505
5	0.6229	0.6859	0.6326	0.7160
7	0.6424	0.7150	0.6530	0.7761
9	0.6579	0.7400	0.6714	0.8232
11	0.6718	0.7604	0.6875	0.8613
13	0.6855	0.7766	0.7009	0.8911
15	0.6982	0.7937	0.7145	0.9144
17	0.7110	0.8068	0.7256	0.9332
19	0.7218	0.8198	0.7371	0.9472
21	0.7315	0.8321	0.7475	0.9578
23	0.7401	0.8435	0.7574	0.9666
25	0.7476	0.8533	0.7665	0.9736
27	0.7567	0.8617	0.7756	0.9787

To consider the effect of CRNs on more complex systems, we use our two populations from above and induce correlation levels of 0.2, 0.4, 0.6, and 0.8. In order to generate an exponential and a uniform variate with the desired correlation, we start with a multivariate normal random variate with the desired correlation using a FORTRAN subroutine from Dagpunar (1988). We then use a numerical approximation to obtain the value of the normal CDF for each variate, and use these values as correlated uniform random numbers to generate the exponential and uniform random variates. Table 2 shows the input correlations to the multivariate normal generator and the resulting sample correlations for the random numbers and for the exponential and uniform random variates.

Note there is a separate column for the sample correlation for the random variates depending on whether the exponential or uniform population was modeled as best (also indicates which was the first of the two random variates generated). The sample correlations were computed using the first 100 macro-replications of the 27 random numbers/variates generated (2700 total observations for each population).

Table 2: Input Correlations and Resulting Sample Correlations

Input ρ	Random Numbers ρ	Random Variates Exp Best $\hat{\rho}$	Random Variates Unif Best ρ
0.0	0.013	0.024	-0.011
0.2	0.206	0.202	0.172
0.4	0.396	0.373	0.358
0.6	0.588	0.539	0.538
0.8	0.788	0.703	0.707
CRN	1.000	0.863	0.863

As expected, the use of independent random number streams provides a lower bound on both PCS values for each case (exponential as best or uniform as best) over the range of correlations induced. Figures 3–6 illustrate increasing PCS^{avc} and PCS^{bem} with increasing correlation. When the uniform population is modeled as best, the PCS values using CRNs provide an upper bound on the other PCS values. When the exponential population is modeled as best, the PCS^{bem} values using CRNs do not provide a similar upper bound, but instead fall between the PCS^{bem} values using input correlations of 0.4 and 0.6. This may be explained in part by looking back at Figure 1. Notice how close the two CDFs are together for random numbers less than 0.4. It is in this region where the uniform population will always win under CRNs. But in this same region using highly correlated random numbers (roughly greater than 0.6), the uniform population is no longer guaranteed to win. In some cases where both random numbers are less than 0.4 and the larger of the pair is used to generate the exponential variate, the exponential population may pick up an additional win. These additional wins increase the PCS^{bem} values at correlations such as 0.6 and 0.8 beyond the PCS^{bem} values using CRNs. The PCS^{avc} values under CRNs also fail to provide an upper bound for values of $v \leq 5$. For values of $v = 6$ and beyond, the PCS^{avc} under CRNs does bound PCS^{avc}

at all induced correlations.

PCS^{avc} values behave as predicted when the uniform population is modeled as best at all induced correlations and with CRNs. This means we see less of an increase in PCS^{avc} with increasing correlation than we see in the corresponding PCS^{bem} values. However, when the exponential population is modeled as best, PCS^{avc} values appear to benefit more from the use of CRNs and from increased correlation. Explaining this behavior is an area of future research.

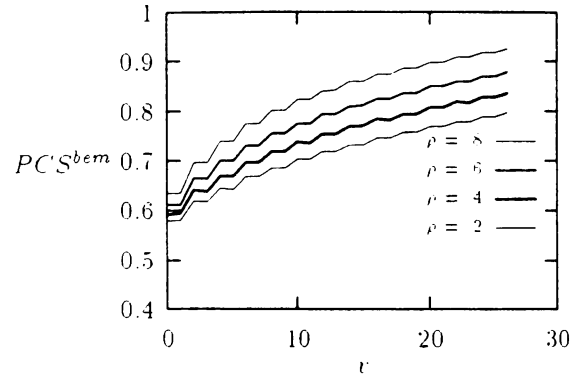


Figure 3: Correlation Effect on PCS^{bem} with Best Population Exponential at $\theta = 1.3$

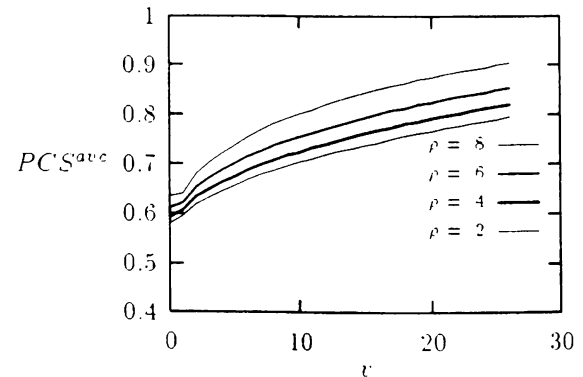


Figure 4: Correlation Effect on PCS^{avc} with Best Population Exponential at $\theta = 1.3$

5 CONCLUDING REMARKS

In this paper we have examined both analytically and empirically the effect of CRNs in some simple two population MSPs. The models used for our populations allowed us to represent the simulation out-

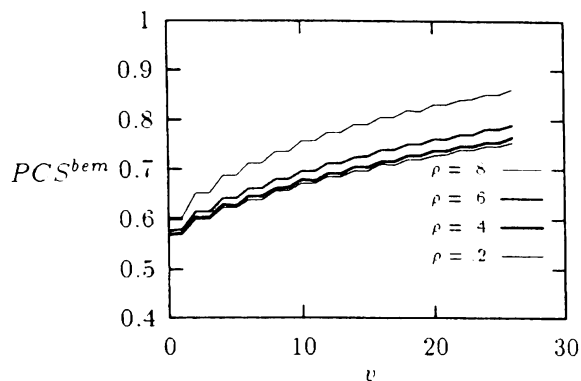


Figure 5: Correlation Effect on PCS^{bem} with Best Population Uniform at $\theta = 1.3$

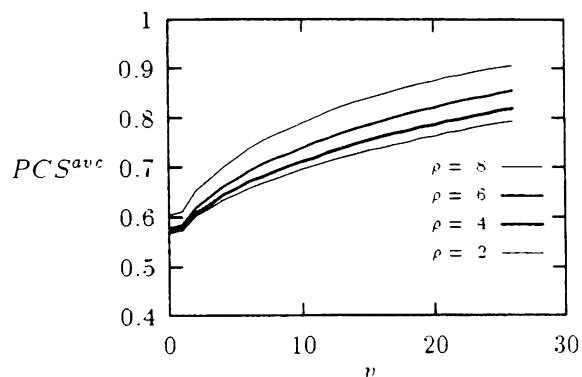


Figure 6: Correlation Effect on PCS^{avc} with Best Population Uniform at $\theta = 1.3$

put from each of our populations as CDFs. We were then able to directly quantify the effect of CRNs on PCS^{bem} values. For our examples we showed that the use of CRNs does not invalidate any required conditions for an MSP, and we were able to quantify the resulting change in PCS^{bem} through the changes in θ . We saw some similar changes in PCS^{avc} , but the results are far from predictable at this point in time. The AVC results are included to provide a comparison with the BEM results. Additional research is needed to understand how CRNs affect PCS^{avc} . The significant variation in the magnitude as well as the direction of the change in θ among our examples, illustrates that the effect of CRNs on an MSP is sensitive to the underlying distributions of the individual systems. This observation is significant since without the use of CRNs, BEM results for a given v depend only on the individual p_j s. However, AVC results do

show a weak distributional dependence on the underlying performance measures (Miller et al. 1997).

In attempting to examine the effect of CRNs on MSPs involving more complex systems, we induced various levels of correlation among the populations used in our first two examples. Our results in general showed an increase in PCS^{bem} and PCS^{avc} values with increasing correlation. However, our results also demonstrate that for our simple systems, the effect of CRNs in an MSP cannot be accurately captured through induced correlation in the system inputs. Even when matching our induced level of correlation in the system inputs with the resulting sample correlation in the system outputs with CRNs, we obtained significantly different results.

REFERENCES

- Bechhofer, R. E., S. Elmaghraby, and N. Morse. 1959. A single-sample multiple-decision procedure for selecting the multinomial event which has the highest probability. *Annals of Mathematical Statistics* 30:102-119.
- Bechhofer, R. E., T. Santner, and D. Goldsman. 1995. *Design and analysis of experiments for statistical selection, screening and multiple comparisons*. New York: John Wiley & Sons, Inc.
- Dagpunar, J. 1988. *Principles of random variate generation*. New York: Oxford University Press.
- Goldsman, D. 1995. Personal communication.
- Mata, F. 1993. Common random numbers and multinomial selection. *Computers and Industrial Engineering* 25:33-36.
- Miller, J. O., B. Nelson, and C. Reilly. 1997. Efficient multinomial selection in simulation. Submitted to *Naval Research Logistics*.

AUTHOR BIOGRAPHIES

J.O. MILLER is a Lieutenant Colonel in the United States Air Force and an Assistant Professor in the Department of Operational Sciences at the Air Force Institute of Technology. His research interests include simulation, nonparametric statistics, and ranking and selection.

KENNETH W. BAUER, JR. is a retired United States Air Force Lieutenant Colonel and a Professor in the Department of Operational Sciences at the Air Force Institute of Technology. His research interests include the statistical aspects of neural networks and simulation modeling.