Proceedings of the 1985 Winter Simulation Conference
D. Gantz, G. Blais, S. Solomon (eds.)

SCREENING TESTS OF PSEUDORANDOM NUMBER GENERATORS ON IBM PCs AND COMPATIBLES

John M. Gleason
College of Business Administration
Creighton University
Omaha, Nebraska 68178

ABSTRACT

Two of the more common tests for the statistical properties of pseudorandom number generators are the frequency test and the serial correlation test. This paper reports the results of frequency and serial correlation tests of random number generators available on two classes of microcomputers: the IBM PCs (the IBM PC, the IBM PC-XT, and the IBM portable) and the IBM PC compatible machines (the Compaq, the Olivetti/AT&T, the Sperry/Leading Edge, the ITT, and the IBM PC-AT). It is shown that the generators on the IBM compatibles perform better than the IBM generators.

INTRODUCTION

One important feature of microcomputers which has not been subject to much scrutiny is their ability to generate pseudorandom numbers. The use of a "random number generator" ranges from educational software with the smaller home computers to involved computer simulation projects with the larger personal computers. It is important, therefore, that a random number generator be capable of generating numbers that behave "statistically" as truly random numbers.

Two of the more common tests for the statistical properties of pseudorandom number generators are the frequency test and the serial correlation test. These tests are often used as "screening" devices. This is, the passing of these tests is only a minimum requirement, and the tests are "intended primarily to catch the main class of unacceptable generators" [1, p. 34]. Consequently, a generator which does not pass the tests is deemed unacceptable, and no further tests for the "goodness" of the generator are performed.

This paper reports the results of frequency and serial correlation tests of random number generators available in BASIC on two classes of microcomputers: the IBM PC (the IBM PC, the IBM PC-XT, and the IBM portable) and the IBM PC compatibles (the Compaq, the Olivetti/AT&T, the Sperry/Leading Edge, the ITT, and the IBM PC-AT). The results suggest that the generators on the IBM compatible machines perform better than the generators on the IBM machines.

THE STATISTICAL TESTS

The research of Hull and Dobell [1] is a fundamental source regarding the analysis of random number generators. The discussion herein is consistent with that research in terms of development and notation.

The test procedure is based on blocks of 500 numbers. For a given block, $f_i (i = 1,2,\ldots,10)$ is the number of numbers u which satisfy the condition $(i - 1)/10 \leq u < i/10$. A chi-square statistic is computed for the block

$$X_1^2 = \frac{1}{50} \sum_{i=1}^{10} (f_i - 50)^2,$$

and this statistic "for a truly random sequence is distributed approximately as X^2" with 9 degrees of freedom [1, p. 34]. This statistic is computed for each of 50 consecutive blocks, and we let $F_i (i = 1,2,\ldots,10)$ be the number of the resulting 50 values of X_1^2 which are between the (i - 1) and i deciles for the chi-square distribution with 9 degrees of freedom. The chi-square statistic for the frequency test is then

$$X_F^2 = \frac{1}{5} \sum_{i=1}^{10} (F_i - 5)^2.$$

For the serial correlation test, f_{ij} is the number of numbers u which satisfy $(i - 1)/10 \leq u < i/10$ and which are followed by a number v which satisfies $(j - 1)/10 \leq v < j/10$. A chi-square statistic is computed

$$X_2^2 = \frac{1}{5} \sum_{i=1}^{10} \sum_{j=i}^{10} (f_{ij} - 5)^2,$$

and for a truly random sequence it is known [2][3] "that $X_2^2 - X_1^2$ is distributed approximately as X^2 with 90 degrees of freedom" [1, p. 35]. The $X_2^2 - X_1^2$ statistic is computed for each of the 50 consecutive blocks, and we let $S_i (i = 1,2,\ldots,10)$ be the number of the resulting 50 values of $X_2^2 - X_1^2$ which are between the (i - 1) and i deciles for the chi-square distribution with 90 degrees of freedom. The chi-square statistic for the serial correlation test is then

$$X_S^2 = \frac{1}{5} \sum_{i=1}^{10} (S_i - 5)^2.$$

Generators are acceptable if the "values of X_F^2 and X_S^2 are not inconsistent with the hypothesis that they

are drawn at random from the X^2 distribution with 9 degrees of freedom" [1]. Thus, a generator is acceptable at the 99% level if the values of X_F^2 and X_S^2 do not exceed 21.7.

THE GENERATORS AND THE TESTS

Some observations are in order before the results of the tests are discussed. The IBM PC, the IBM PC-XT, and the IBM portable all have the same generator. The IBM PC-AT generator, however, differs from its corporate siblings. In fact, the generator on the IBM PC-AT is the same generator as that of the IBM PC compatibles. Thus, in this paper, the IBM PC-AT has been included with the group of IBM compatibles.

There are problems related to the use of the generators with both classes of microcomputers. Specifically, the RANDOMIZE statement and the RND function do not work as described in the BASIC manuals which accompany the microcomputers, and this can lead to a variety of problems in attempting to use the generators. For example, Modianos et. al. [4] have noted problems with the random number generator on the IBM PC. Furthermore, it is interesting to note that an example program, in the BASIC manual for the IBM PCs, which illustrates the use of the RANDOMIZE statement does not even yield the same results as shown in the manual when the program is run on an IBM PC.

The problems with the use of the generators will not be discussed in this paper. However, it should be noted that the problems are of such significance that they can invalidate the results of analyses in which the analyst is not aware of the drawbacks of RANDOMIZE and RND. Specifically, the user cannot be assured that the RANDOMIZE statement reseeds the generator in the manner that the analyst expects. Thus, it is possible that many studies are flawed by the way in which the RANDOMIZE statements are used.

The research reported herein recognizes that many analysts may be unaware of the problems with RANDOMIZE and RND. Consequently, although results are presented for fifty different generator seeding processes, the exact seed values are reported for only nine of the seeding processes. These nine processes resulted from seeding by using a single RANDOMIZE statement at the beginning of the program. The other 41 processes resulted from seeding by using the RANDOMIZE statement in a manner that is consistent with the literature, but which results in bizarre seeding which is not the type of seeding that the analyst would expect to occur. For this reason, the actual seed values are not reported for these 41 runs. That is, the actual seed values are not of importance in these runs; instead, it is the behavior of the generators on the two classes of machines, under the exact same sort of seeding processes, that is of interest.

For those who are aware of the problems with the use of the generators, and who are concerned with using seeds which yield numbers which pass the fundamental screening tests for randomness, results for nine specific seeds are also presented.

RESULTS

The first 41 seeding processes are similar to those which were used successfully to test random number generators for specific seeds on other microcomputers [5][6]. Although the processes are consistent with the literature regarding the use of generators with IBM PCs and compatibles, the processes result in bizarre seeding because of the vagaries of RANDOMIZE and RND. The results of frequency and serial correlation tests for these processes for the two classes of microcomputers are shown in Table 1. For the frequency test on the IBM PC, the chi-square values ranged from 4.4 to 23.6. Processes 25 and 37 resulted in chi-square values of 23.6 and 22.8, respectively. These were the only two processes which resulted in chi-square values greater than the critical value of 21.7 at the 99% level. Serial correlation chi-square values for the IBM ranged from 3.2 to 28. The value of 28 occurred with seeding process 31, which was the only process which resulted in a chi-square value greater than the critical value of 21.7.

TABLE 1: Results of the Seeding Processes

	IBM PCs		IBM PC COMPATIBLES	
Seeding Process	Frequency	Serial Correlation	Frequency	Serial Correlation
1	10.8	11.6	7.2	8
2	8	6	18.8	12.4
3	8.4	8.4	9.2	18.4
4	12.4	16	10.4	15.6
5	10.8	10.8	8	5.6
6	8	10.4	9.6	16.4
7	5.2	9.6	11.6	3.6
8	7.6	13.2	8	14
9	12	9.6	6	11.6
10	7.2	9.2	9.2	5.2
11	4.4	9.6	4.4	6.4
12	6.8	8.8	6	14.8
13	10.8	13.6	4.4	6.8
14	10	9.6	8	14.8
15	9.2	7.6	4	10.8
16	8	6	6	8
17	7.2	3.2	6.4	9.6
18	7.2	9.2	4.8	5.2
19	10.8	6.8	5.2	9.6
20	11.6	16.4	5.6	8
21	4.4	10.4	15.2	9.6
22	8.4	3.6	5.6	1.6
23	5.6	12	14.8	10
24	8.8	3.6	10.4	9.2
25	23.6	13.2	8.8	10.8
26	12	6.4	10.8	14.8
27	4.4	13.2	12	7.6
28	14.8	13.6	11.2	13.2
29	9.6	7.2	8	16
30	8	16.4	4.8	14
31	5.6	28	12	2.8
32	8.8	11.6	17.2	6
33	7.6	6	9.6	5.6
34	10.4	7.6	6.4	2
35	7.2	12	7.6	9.2
36	14.8	5.6	2.8	8
37	22.8	11.2	6	3.2
38	5.2	5.6	9.2	2.4
39	12.8	12.8	11.2	3.2
40	6.4	14.8	16.4	7.2
41	14.8	5.6	6	2

For the frequency test on the IBM PC compatibles, the chi-square values range from 2.8 to 18.8. All seeding processes passed the test, based on a critical value of 21.7. Similarly, all seeding processes passed the serial correlation test, with chi-square values ranging from 1.6 to 18.4.

The results of the nine tests with specific seed values are shown in Table 2. The seeds were 1, 2, 5, 100, 199, 1000, 9000, 9999, and 10000. The frequency test results for the IBM PCs ranged from 4.4 to 30.8. The seed of 2 resulted in the chi-square value of 30.8, and this was the only seed which resulted in a chi-square value greater than the critical value of 21.7. Serial correlation values for the IBM PCs ranged from 7.2 to 14.4; thus, all seeds passed the test.

TABLE 2: Results for Specific Seeds

	IBM PCs		IBM PC COMPATIBLES	
Seeding Process	Frequency	Serial Correlation	Frequency	Serial Correlation
1	10.4	9.2	4.4	4
2	30.8	12	11.6	3.6
5	11.6	14.4	6.4	16.8
100	16	14	3.6	9.6
199	9.2	13.2	11.2	10
1,000	6.4	13.6	11.6	4
9,000	17.6	10.8	6.4	11.6
9,999	5.2	7.2	10.8	6.4
10,000	4.4	11.6	11.2	8

The frequency test results for the nine specific seeds with the IBM PC compatibles ranged from 3.6 to 11.6, and all of the seeds passed the test. Similarly, all of the seeds passed the serial correlation test with serial correlation values ranging from 3.6 to 16.8.

CONCLUSIONS

The results of the 50 seeding processes suggest that the generators on the IBM PC compatibles perform better than the generators on the IBM PCs. Not only did the IBM PC compatibles fail no test (whereas the IBM PCs failed both frequency and serial correlation tests), but the range of chi-square values, and the lower and upper chi-square values for those ranges for the frequency and serial correlation tests are smaller than the respective values for the tests on the IBM PCs. Furthermore, it is interesting that the IBM PC-AT has the same (seemingly better) generator as the IBM PC compatibles, whereas the IBM PC, the IBM PC-XT, and the IBM portable share a different generator.

Finally, it should be noted that a program that is written on an IBM PC compatible machine can be expected to yield different results when run on an IBM PC if the program uses the random number generator. More importantly, seeding processes which yield numbers which pass statistical tests for randomness on the IBM PC compatible machines may result in numbers which fail those tests when the program is run on an IBM PC. For example, if the seed of 2 were used in a program, the numbers generated on the IBM PC compatibles would pass statistical tests for frequency (numbers uniformly distributed over the interval) and first order serial correlation. However, if that same program were run on one of the machines in the IBM PC class, the resulting numbers are not random because they do not pass the frequency test--they are not uniformly distributed over the interval.

ACKNOWLEDGEMENTS

This research was supported by a grant from Creighton University. The author expresses his appreciation to the following individuals for making hardware available for this research: Steve Comer, Department of Mathematics, The Citadel; Roger Hayen, Department of Decision Sciences, University of Nebraska at Omaha; Mike Glynn, Overland Computer Services, Omaha; and Terry Begley, Creighton University.

REFERENCES

[1] Hull, T. E. and Dobell, A. R., "Mixed Congruential Random Number Generators for Binary Machines," Journal of the Association for Computing Machinery, Vol. II, No. 1, January, 1964, pp. 31-40.

[2] Good, I. J., "The Serial Test for Sampling Numbers and Other Tests for Randomness," Proc. Camb. Phil. Soc., Vol. 49, 1953, pp. 276-284.

[3] Good, I. J., "On the Serial Test for Random Sequences," Ann. Math. Stat., Vol. 28, 1957, pp. 262-264.

[4] Modianos, Doan T., Scott, Robert C. and Cornwell, Larry W., "Random Number Generation on Microcomputers," Interfaces, Vol. 14, No. 4, 1984, pp. 81-87.

[5] Gleason, John M., "Screening Tests of Microcomputer Pseudorandom Number Generators: The Inexpensive Machines," American Institute for Decision Sciences Midwest Conference Proceedings, Indianapolis, May 1984, pp. 123-124.

[6] Gleason, John M., "The Apple IIe Random Number Generator: Are the Numbers Really Random?," Joint National ORSA/TIMS Meeting, Dallas, November 1984.

JOHN M. GLEASON is Professor of Decision Sciences, College of Business Administration, Creighton University. He received the B.S. (mathematics) and M.B.A. degrees from the University of Missouri at Kansas City, and the D.B.A. degree from Indiana University. He has served on the faculties of Texas Tech University, the University of Nebraska at Omaha, and The Citadel, where he held the W. Frank Hipp endowed chair as Distinguished Professor of Business Administration and Public Policy.

College of Business Administration
Creighton University
Omaha, Nebraska 68178
(402) 280-2624