Behavior Research Methods & Instrumentation
1983, 15 (5), 536-537

Tests of randomness for pseudorandom number generators

MICHAEL J STRUBE
Washington University, St. Louis, Missouri

This study examined the "randomness" of the numbers generated by the VIC-20 computer. Using eight standard tests, it was determined that the numbers provided by the VIC-20 are adequately random for practical purposes. The tests are applicable to other computer systems and provide a general means for evaluating random-number generators.

Small computers can serve a variety of useful functions, from stimulus presentation and response recording to small-scale statistical simulation. Many of these functions, however, require the adequate generation of random numbers. For example, stimulus presentation often requires the random ordering of multiple stimuli. Statistical simulation requires the generation of random-number distributions. Since computers must rely on a mathematical rule for generation of random numbers, the resultant values are more appropriately called "pseudo" random numbers. Moreover, a variety of generation rules exist, each producing "random numbers" with different characteristics. It is important that a random-number generator be tested for "randomness" prior to its serious use for problems requiring random numbers.

Described below are eight standard tests of randomness (Gruenberger & Jaffray, 1965) that were conducted on the Commodore VIC-20 computer. The VIC-20 is an inexpensive, flexible computer that makes an attractive instrument for both classroom and laboratory use. Its uses in my own laboratory have included stimulus presentation (requiring random ordering of stimuli and random screen placement), random determination of intertrial intervals for performance tasks, and statistical simulations. The tests, however, are general and can be used to test the random-number generators of other computer systems.

GAP TEST

The gap test examines the average interval between repetitions of numbers in a series of numbers. For each of the integers 0-9, 1,000 numbers were generated and the average interval was computed. This empirical average was then compared with the expected interval of 10, and a t test was computed. Furthermore, the variances of the intervals were calculated and compared with the expected value of 90 (see Gruenberger & Jaffray, 1965) via χ^2. The empirical interval means ranged from 8.55 to 11.16, and none differed reliably from the expected value (all ts $\leqslant$ 1.64). The empirical variances ranged from 58.04 to 107.5, and none differed reliably from the expected value (all ps $>$ .05).

CORRELATION TEST

The correlation test examines whether there are serial dependencies in a string of random numbers. For this test, 1,000 numbers (ranging between 0 and 1) were generated, and each number was correlated with its adjacent number, with the number two positions away in the series, and so forth, out to the 20th position. In other words, 20 correlations were computed by successively lagging the series 20 times. The obtained correlations ranged from −.044 to +.059, indicating the absence of serial dependencies over a range of 20 lags.

DISTANCE TEST

The distance test requires the generation of four numbers, which are paired to produce the coordinate points on a Cartesian plane. The distance between the points can then be determined using the Pythagorean theorem. The results from a large number of such distance computations can be compared with the theoretical cumulative distribution (assuming randomness: Gruenberger & Jaffray, 1965). To perform the distance test, 1,000 sets of four numbers (ranging between 0 and 1) were generated. The first two numbers in the set identified one point, and the second two numbers identified the second point. For each pair of points, the distance was calculated, and the resulting distribution of distance scores was compared with the expected distribution. A comparison of empirical and theoretical cumulative proportions (see Table 1) indicated that

This project was supported by BRSG SO7 RR07054-17, awarded by the Biomedical Research Support Grant Program, Division of Research Resources, National Institutes of Health. A complete copy of the program and results of this study is available on request. Correspondence should be addressed to Michael J Strube, Department of Psychology, Washington University, St. Louis, Missouri 63130.

 Copyright 1983 Psychonomic Society, Inc.

Table 1
Obtained and Expected Cumulative Proportions for the Distance Test

Distance Value	Proportion of Scores Less Than Selected Distance Values	
	Obtained	Expected
.1	.239	.235
.2	.404	.410
.3	.552	.549
.4	.683	.662
.5	.757	.753
.6	.816	.826
.7	.865	.882
.8	.911	.9252
.9	.945	.9556
1.0	.969	.9749
1.1	.981	.9857
1.2	.99	.99205
1.3	.992	.99579
1.4	.994	.99793
1.5	.997	.999080
1.6	.999	.999652
1.7	1.000	.999898
1.8	1.000	.999982
1.9	1.000	.999999
2.0	1.000	1.000

Note–Distance values represent the square of the distance between two points.

discrepancies never exceeded .021 ($p > .05$) (Kolmogorov-Smirnov one-sample test; Siegel, 1956).

POKER TEST

The poker test requires the generation of four-digit numbers and the tabulation of the frequency of instances of four matching digits, three matching digits, two sets of two matching digits, any two matching digits, and all digits different. For 1,000 four-digit sequences, the expected frequencies are 1, 36, 27, 432, and 504, respectively. The empirical frequencies were 3, 27, 27, 457, and 486, which do not differ reliably from the expected values [$\chi^2(4) = 8.34, p > .05$].

FREQUENCY TEST

The frequency test simply requires the tabulation of the frequency of the digits 0-9 in a series of numbers. In a sequence of 1,000 numbers, each should occur 100 times. The obtained frequencies for the digits 0-9 were, respectively, 108, 96, 100, 116, 112, 97, 96, 87, 98, and 90. These do not differ reliably from expectation [$\chi^2(9) = 7.78, p > .05$].

SERIAL TEST

Similar to the frequency test, the serial test examines the frequency of two-digit combinations from 00 to 99. In a list of 1,000 two-digit numbers, each should occur 10 times. The empirical frequencies ranged from 4 to 18 and did not differ reliably from expectation [$\chi^2(99) = 89.4, p > .05$].

MAXIMUM TEST

In the maximum test, a three-digit number is generated, and the frequency of instances in which the middle digit is the maximum is tabulated. Theoretically, this should occur 285 times in a set of 1,000 three-digit sequences. The empirical frequency was 263, not reliably different from the expected value [$\chi^2(1) = 1.70, p > .05$].

COUPON TEST

The coupon test requires the tabulation of the number of digits required to make up a complete set of the integers 0-9. The average empirical frequency over a series of 1,000 digits was 27.72, which does not differ reliably from the expected value of 29.3 ($t = 1.15$, $p > .05$).

In summary, the random numbers generated by the VIC-20 did not deviate from those expected on the basis of the laws of probability embodied in eight standard tests. It is safe to assume that, for practical purposes, the VIC-20 can supply numbers with adequately random characteristics. The application of the eight tests to random number generators from other computer systems will provide a basis for judging their quality and enhancing their utility as research tools.

REFERENCES

GRUENBERGER, F., & JAFFRAY, G. *Problems for computer solution.* New York: Wiley, 1965.

SIEGEL, S. *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill, 1956.

(Manuscript received August 16, 1983;
revision accepted for publication September 30, 1983.)