Random Numbers

🞂 Randomness

🞂 Applications in scientific computing

🞂 Random variables and Pseudo Random Numbers 🞂 Properties of a good random number generator (RNG) 🞂 RNGs

🞂 Mid-square method

🞂 Linear Congruential Generator

🞂 Linear Shift Feedback Register

🞂 Quasi Random Numbers

Outline

🞂 Random number generation 🞂 Examples and limitations 🞂 Tests for RNGs

🞂 Description of practical exercise

Random Number Generation

Desirable Attributes: 🞂 Uniformity

🞂 Independence

🞂 Efficiency

🞂 Replicability

🞂 Long Cycle Length

3 Random Numbers

Random Number Generation (cont.)

Each random number Rt is an independent sample drawn from a continuous uniform distribution between 0 and 1

1 , 0 ≤ x ≤ 1

pdf: f(x) = 

0 , otherwise

4 Random Numbers

Random Number Generation (cont.)

PDF:

1

)

x

(f

0

x

5 Random Numbers

Techniques for Generating Random Number

MidSquare

Example:

X0 = 7182 (seed)

= 51581124

==> R1 = 0.5811

= (5811) 2 = 33767721

==> R2 = 0.7677

etc.

6 Random Numbers

Techniques for Generating Random Number (cont.)

Note: Cannot choose a seed that guarantees that the sequence will not degenerate and will have a long period. Also, zeros, once they appear, are carried in subsequent numbers.

Ex1: X0 = 5197 (seed) = 27008809 ==> R1 = 0.0088 = 00007744 ==> R2 = 0.0077

Ex2: X0 = 4500 (seed) = 20250000 ==> R1 = 0.2500 = 06250000 ==> R2 = 0.2500

7 Random Numbers

Techniques for Generating Random Number (cont.)

 Multiplicative Congruential Method:

Basic Relationship

Xi+1 = a Xi (mod m), where a ≥ 0 and m ≥ 0

Most natural choice for m is one that equals to the capacity of a computer word.

m = 2b (binary machine), where b is the number of bits in the computer word.

m = 10d (decimal machine), where d is the number of digits in the computer word.

8 Random Numbers

Techniques for Generating Random Number (cont.)

The max period(P) is:

 For m a power of 2, say m = 2b, and c ≠ 0, the longest possible period is P = m = 2b , which is achieved provided that c is relatively prime to m (that is, the greatest common factor of c and m is 1), and a = 1 + 4k, where k is an integer.

 For m a power of 2, say m = 2b, and c = 0, the longest possible period is P = m / 4 = 2b-2 , which is achieved provided that the seed X0 is odd and the multiplier, a, is given by a = 3 + 8k or a = 5 + 8k, for some k = 0, 1,...

9 Random Numbers

Techniques for Generating Random Number (cont.)

 For **m** a prime number and **c = 0**, the longest possible period is **P = m - 1**, which is achieved provided that the multiplier, **a**, has the property that the smallest integer k such that ak - 1 is divisible by m is k = m - 1,

10 Random Numbers

Techniques for Generating Random Number (cont.)

(Example)

Using the multiplicative congruential method, find the period of the generator for a = 13, m = 26, and X0 = 1, 2, 3, and 4. The solution is given in next slide. When the seed is 1 and 3, the sequence has period 16. However, a period of length eight is achieved when the seed is 2 and a period of length four occurs when the seed is 4.

11 Random Numbers

Techniques for Generating Random Number (cont.)

Period Determination Using Various seeds

i Xi Xi Xi Xi

0 1 2 3 4

1 13 26 39 52

2 41 18 59 36

3 21 42 63 20

4 17 34 51 4

5 29 58 23

6 57 50 43

7 37 10 47

8 33 2 35

9 45 7

10 9 27

11 53 31

12 49 19

13 61 55

14 25 11

15 5 15

16 1 3

12 Random Numbers

Techniques for Generating Random Number (cont.)

 Linear Congruential Method: Xi+1 = (aXi + c) mod m, i = 0, 1, 2....

(Example)

let X0 = 27, a = 17, c = 43, and m = 100, then X1 = (17\*27 + 43) mod 100 = 2

R1 = 2 / 100 = 0.02

X2 = (17\*2 + 43) mod 100 = 77

R2 = 77 / 100 = 0.77 .........

13 Random Numbers

Attributes of RNGs

1. Uniformity

2. Independence

3. Efficiency

4. Replicability

5. Long Cycle Length

Test for Random Numbers

1. ***Frequency test***. Uses the Kolmogorov-Smirnov or the chi-square test to compare the distribution of the set of numbers generated to a uniform distribution.

2. ***Runs test***. Tests the runs up and down or the runs above and below the mean by comparing the actual values to expected values. The statistic for comparison is the chi-square.

3. ***Autocorrelation test***. Tests the correlation between numbers and compares the sample correlation to the expected correlation of zero.

15 Random Numbers

Test for Random Numbers (cont.)

4. ***Gap test***. Counts the number of digits that appear between repetitions of a particular digit and then uses the Kolmogorov-Smirnov test to compare with the expected number of gaps.

5. ***Poker test***. Treats numbers grouped together as a poker hand. Then the hands obtained are compared to what is expected using the chi square test.

16 Random Numbers

Steps to Perform a Test of Hypothesis 🞂 State the null and alternative hypotheses 🞂 Select the distribution to use

🞂 Determine the rejection and non-rejection regions 🞂 Calculate the value of the test statistic

🞂 Make a decision

Test for Random Numbers (cont.)

In testing for uniformity, the hypotheses are as follows:

H0: Ri ~ U[0,1]

H1: Ri ≠ U[0,1]

The null hypothesis, H0, reads that the numbers are distributed uniformly on the interval [0,1].

18 Random Numbers

Test for Random Numbers (cont.)

In testing for independence, the hypotheses are as follows;

H0: Ri ~ independently

H1: Ri ≠ independently

This null hypothesis, H0, reads that the numbers are independent. Failure to reject the null hypothesis means that no evidence of dependence has been detected on the basis of this test. This does not imply that further testing of the generator for independence is unnecessary.

19 Random Numbers

Test for Random Numbers (cont.)

Level of significance α

α = P(reject H0 | H0 true)

Frequently, α is set to 0.01 or 0.05

(Hypothesis)

Actually True Actually False Accept 1 - α β

(Type II error)

Reject α 1 - β

(Type I error)

20 Random Numbers

Test for Random Numbers (cont.)

 The *Gap Test* measures the number of digits between successive occurrences of the same digit. (Example) length of gaps associated with the digit 3. 4, 1, 3, 5, 1, 7, 2, 8, 2, 0, 7, 9, 1, 3, 5, 2, 7, 9, 4, 1, 6, 3 3, 9, 6, 3, 4, 8, 2, 3, 1, 9, 4, 4, 6, 8, 4, 1, 3, 8, 9, 5, 5, 7 3, 9, 5, 9, 8, 5, 3, 2, 2, 3, 7, 4, 7, 0, 3, 6, 3, 5, 9, 9, 5, 5 5, 0, 4, 6, 8, 0, 4, 7, 0, 3, 3, 0, 9, 5, 7, 9, 5, 1, 6, 6, 3, 8 8, 8, 9, 2, 9, 1, 8, 5, 4, 4, 5, 0, 2, 3, 9, 7, 1, 2, 0, 3, 6, 3 Note: eighteen 3’s in list

==> 17 gaps, the first gap is of length 10

21 Random Numbers

Test for Random Numbers (cont.)

We are interested in the frequency of gaps. P(gap of 10) = P(not 3) ⋅⋅⋅P(not 3) P(3) , note: there are 10 terms of the type P(not 3)

= (0.9)10 (0.1)

The theoretical frequency distribution for randomly ordered digit is given by

F(x) = 0.1 (0.9)n = 1 - 0.9x+1

Note: observed frequencies for all digits are compared to the theoretical frequency using the Kolmogorov-Smirnov test.

22 Random Numbers

Test for Random Numbers (cont.)

(Example)

Based on the frequency with which gaps occur, analyze the 110 digits above to test whether they are independent. Use α= 0.05. The number of gaps is given by the number of digits minus 10, or 100. The number of gaps associated with the various digits are as follows:

Digit 0 1 2 3 4 5 6 7 8 9 # of Gaps 7 8 8 17 10 13 7 8 9 13

23 Random Numbers

Test for Random Numbers (cont.)

Gap Test Example

Relative Cum. Relative Gap Length Frequency Frequency Frequency F(x) |F(x) - SN(x)| 0-3 35 0.35 0.35 0.3439 0.0061 4-7 22 0.22 0.57 0.5695 0.0005 8-11 17 0.17 0.74 0.7176 0.0224 12-15 9 0.09 0.83 0.8147 0.0153 16-19 5 0.05 0.88 0.8784 0.0016 20-23 6 0.06 0.94 0.9202 0.0198 24-27 3 0.03 0.97 0.9497 0.0223 28-31 0 0.00 0.97 0.9657 0.0043 32-35 0 0.00 0.97 0.9775 0.0075 36-39 2 0.02 0.99 0.9852 0.0043 40-43 0 0.00 0.99 0.9903 0.0003 44-47 1 0.01 1.00 0.9936 0.0064

24 Random Numbers

Test for Random Numbers (cont.)

The critical value of D is given by D0.05 = 1.36 / √100 = 0.136

Since D = max |F(x) - SN(x)| = 0.0224 is less than D0.05, do not reject the hypothesis of independence on the basis of this test.

25 Random Numbers

Test for Random Numbers (cont.)

 *Run Tests* (Up and Down)

Consider the 40 numbers; both the Kolmogorov-Smirnov and Chi-square would indicate that the numbers are uniformly distributed. But, not so.

0.08 0.09 0.23 0.29 0.42 0.55 0.58 0.72 0.89 0.91 0.11 0.16 0.18 0.31 0.41 0.53 0.71 0.73 0.74 0.84 0.02 0.09 0.30 0.32 0.45 0.47 0.69 0.74 0.91 0.95 0.12 0.13 0.29 0.36 0.38 0.54 0.68 0.86 0.88 0.91

26 Random Numbers

Test for Random Numbers (cont.)

Now, rearrange and there is less reason to doubt independence.

0.41 0.68 0.89 0.84 0.74 0.91 0.55 0.71 0.36 0.30 0.09 0.72 0.86 0.08 0.54 0.02 0.11 0.29 0.16 0.18 0.88 0.91 0.95 0.69 0.09 0.38 0.23 0.32 0.91 0.53 0.31 0.42 0.73 0.12 0.74 0.45 0.13 0.47 0.58 0.29

27 Random Numbers

Test for Random Numbers (cont.)

Concerns:

 Number of runs

 Length of runs

Note: If N is the number of numbers in a sequence, the maximum number of runs is N-1, and the minimum number of runs is one.

If “a” is the total number of runs in a sequence, the mean and variance of “a” is given by

28 Random Numbers

Test for Random Numbers (cont.)

∝a = (2n - 1) / 3

= (16N - 29) / 90

For N > 20, the distribution of “a” approximated by a normal distribution, N(∝a , ). This approximation can be used to test the independence of numbers from a generator. Z0 = (a - ∝a) / σa

29 Random Numbers

Test for Random Numbers (cont.)

Substituting for ∝a and σa ==>

Za = {a - [(2N-1)/3]} / {√(16N-29)/90}, where Z ~ N(0,1)

Acceptance region for hypothesis of independence -Zα/2 ≤ Z0 ≤ Zα/2

30 Random Numbers

Test for Random Numbers (cont.)

(Example)

Based on runs up and runs down, determine whether the following sequence of 40 numbers is such that the hypothesis of independence can be rejected where α = 0.05.

0.41 0.68 0.89 0.94 0.74 0.91 0.55 0.62 0.36 0.27 0.19 0.72 0.75 0.08 0.54 0.02 0.01 0.36 0.16 0.28 0.18 0.01 0.95 0.69 0.18 0.47 0.23 0.32 0.82 0.53 0.31 0.42 0.73 0.04 0.83 0.45 0.13 0.57 0.63 0.29

31 Random Numbers

Test for Random Numbers (cont.)

The sequence of runs up and down is as follows: + + + - + - + - - - + + - + - - + - + - - + - - + - + + - - + + - + - - + + - There are 26 runs in this sequence. With N=40 and a=26, ∝a = {2(40) - 1} / 3 = 26.33 and

= {16(40) - 29} / 90 = 6.79

Then,

Z0 = (26 - 26.33) / √(6.79) = -0.13

Now, the critical value is Z0.025 = 1.96, so the independence of the numbers cannot be rejected on the basis of this test.

32 Random Numbers

Test for Random Numbers (cont.)

 *Poker Test* - based on the frequency with which certain digits are repeated.

Example:

0.255 0.577 0.331 0.414 0.828 0.909

Note: a pair of like digits appear in each number generated.

33 Random Numbers

Test for Random Numbers (cont.)

In 3-digit numbers, there are only 3 possibilities. P(3 different digits) =

(2nd diff. from 1st) \* P(3rd diff. from 1st & 2nd) = (0.9) (0.8) = 0.72

P(3 like digits) =

(2nd digit same as 1st) \* P(3rd digit same as 1st) = (0.1) (0.1) = 0.01

P(exactly one pair) = 1 - 0.72 - 0.01 = 0.27

34 Random Numbers

Test for Random Numbers (cont.)

(Example)

A sequence of 1000 three-digit numbers has been generated and an analysis indicates that 680 have three different digits, 289 contain exactly one pair

of like digits, and 31 contain three like digits. Based on the poker test, are these numbers

independent?

Let α = 0.05.

The test is summarized in next table.

35 Random Numbers

Test for Random Numbers (cont.)

Observed Expected (Oi - Ei)2

Combination, Frequency, Frequency, ----------- i Oi Ei Ei Three different digits 680 720 2.24 Three like digits 31 10 44.10  Exactly one pair 289 270 1.33 ------ ------ ------- 1000 1000 47.65

The appropriate degrees of freedom are one less than the number of class intervals. Since χ20.05, 2 = 5.99 < 47.65, the independence of the numbers is rejected on the basis of this test.

36 Random Numbers

Perspective

Diehard Tests (George Marsaglia)

1. Birthday spacing

2. Overlapping permutations

3. Rank of matrices

4. Monkey tests

5. Count the 1s

6. Parking lot test

7. Minimum distance test

8. Random spheres test

9. The squeeze test

10. Overlapping sums test

11. Runs test

12. The craps test

DieHarder

Random number generator in stochastic simulations in physics/biology/chemistry: Mersenne Twister

Practical Exercise

🞂 Generate random numbers using a linear congruential generator

🞂 Determine the frequencies of the (N-1) runs up and runs down in the sequence

References

🞂 Simulation and Modeling Analysis, Law & Kelton (1991) 🞂 CSE 808 Modeling and Discrete Simulation- H. Hughes, Carnegie Mellon Univ.