Totally Random – at least almost

Tim Chartier

CSC 231 @

DAVIDSON

April 17, 2009

Random Numbers

- In ancient times, rolling dice, flipping a coin, or shuffling playing cards were among the methods that served to generate random results.

- Today, we concern ourselves with generating a sequence of numbers with the intent that the sequence does not have any (or less rigorously, no easily discernable) patterns.

DAVIDSON

Living in sin

- Distinguishing whether a *random number generator* truly produces a random sequence is often difficult to decide.
- However, any deterministic computation is generally considered not to be a “true” random number generator.
- Encapsulating this reality, John von Neumann once said

> *“Anyone who uses arithmetic methods to produce random numbers is in a state of sin.”*

DAVIDSON

Pseudorandomness

- All is not lost since pseudorandom number generators exist that satisfy all the statistical properties of true random numbers. However, the output is deterministic and in that sense predictable.
- Note, today will be only an introduction to this area of computer science.

DAVIDSON

LCGs

- We introduce to create a sequence of pseudorandom numbers with linear congruential generators (LCGs).
- We will see their sensitivity to input parameters but also succeed in producing a uniform distribution.
- Many applications require more sophisticated techniques.

DAVIDSON

Defining an LCG

An LCG is defined by the recursive relation:

$$x_{i+1} = (ax_i + c) \bmod m \quad (i = 0, 1, 2, \ldots)$$

In order to produce the sequence of numbers, a, c, m and x_0 (also known as the *seed*) must be specified.

Then the random number between 0 and 1 is defined as:

$$R_i = \frac{x_i}{m} \quad (i = 0, 1, 2, \ldots)$$

While it is possible to generate a 0, a random number from this method cannot equal 1.

DAVIDSON

Example

Again, $x_{i+1} = (ax_i + c) \bmod m \quad (i = 0, 1, 2, \ldots)$ and $R_i = \frac{x_i}{m} \quad (i = 0, 1, 2, \ldots)$.

Let $x_0 = 30, a = 13, c = 55$, and $m = 100$.

$R_0 = 0.30 \Rightarrow x_1 = (13 * 30 + 55) \bmod 100 = 45$. Therefore $R_1 = 45/100 = 0.45$.

DAVIDSON

Your Turn

Find R_2, R_3 and R_4.

Your Turn

Find R_2, R_3 and R_4.

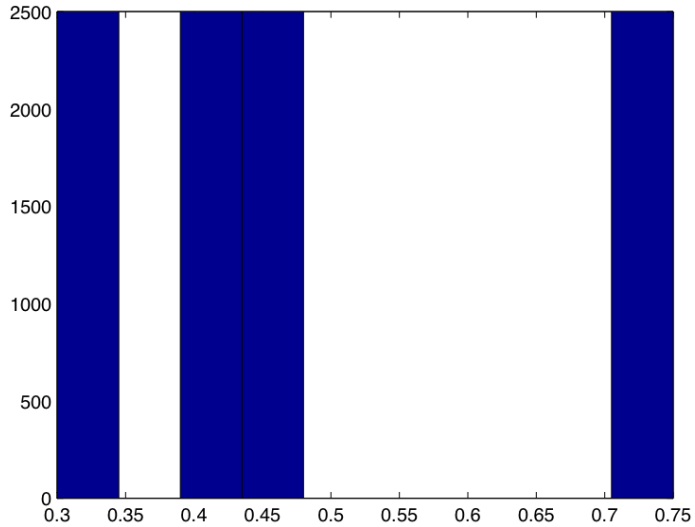
- So, $x_2 = 13 * x_1 + 55 \bmod 100 = 5840 \bmod 100 = 40$. Therefore, $R_2 = 0.40$.
- Next, $x_3 = 13 * x_2 + 55 \bmod 100 = 575 \bmod 100 = 75$. Therefore, $R_3 = 0.75$.
- Finally, $x_4 = 13 * x_3 + 55 \bmod 100 = 1030 \bmod 100 = 30$. Therefore, $R_4 = 0.30$.

DAVIDSON

uniform or !(uniform)

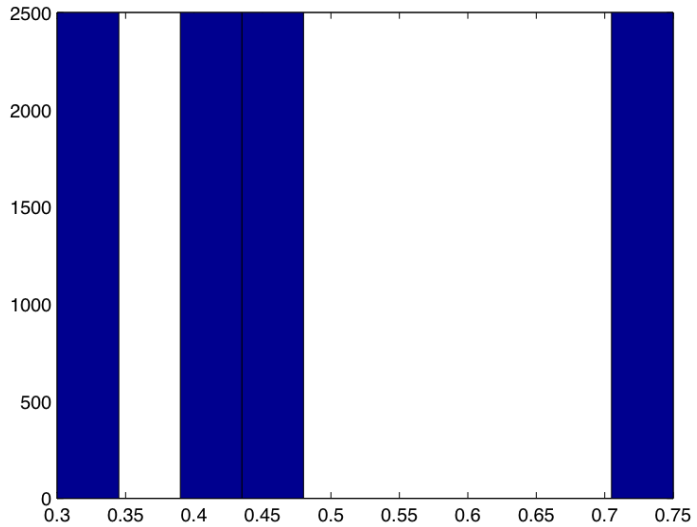
Goal: A uniform distribution between 0 and 1. That is, all numbers between 0 and 1 are equally likely to be produced by our random number generator.

We will determine this, at least empirically, by producing 10,000 random numbers from this recurrence relation. The histogram of the sequence of 1000 numbers produced by this recurrence is given below.



DAVIDSON

Far from uniform



- Clearly, this is not uniform. Without looking at the next slide, why? Look at the previous slide and see if you can see it.
- Our goal of a uniform distribution failed. Note that $R_4 = 0.30 = R_0$, which leads to $R_i = R_{i+4}$ for $i \geq 0$.
- This LCG is said to have period 4.

DAVIDSON

Java code

- Java code that will compute random numbers with an LCG is contained on Blackboard.
- You will find the code in `LcgGenerator.java`.
- This code will not produce a histogram. You will learn to adapt this code for visualization using *Easy Java Simulation* software with Dr. Christian.

DAVIDSON

More insight

An LCG will have a full period (note the period is always less than or equal to m) if:

- c and m are relatively prime (do not contain common prime factors),
- $a - 1$ is divisible by all prime factors of m,
- $a - 1$ is a multiple of 4 if m is a multiple of 4,
- $m > \max(a, c, x_0)$, and
- $a > 0, b > 0$.

LCGs are sensitive to the choice of c, m and a.

DAVIDSON

Your Turn

Take a few moments to see if you can determine which of these properties was violated in the previous example.

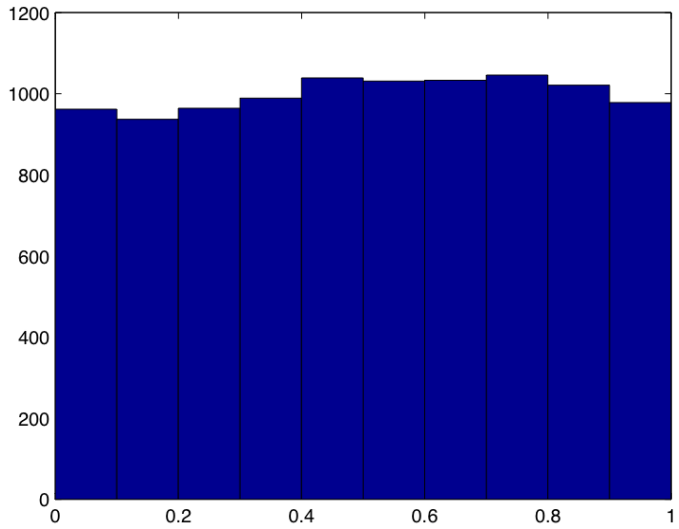
Your Turn

Take a few moments to see if you can determine which of these properties was violated in the previous example.

In the previous example, $a - 1 = 12$ is not divisible by 5 which is a prime factor of $m = 100$.

Example

- Consider $a = 1664525$, $b = 1013904223$, $m = 2^{32}$ and again $x_0 = 30$.
- As before, we produce 10,000 random numbers for these parameters to the LCG and plot the resulting histogram:



DAVIDSON

Place for LCGs

- LCGs are not reliable for applications requiring high-quality randomness. For example, Monte Carlo simulations, which we will look at briefly to motivate another topic, depend on randomness and LCGs do not ensure quality in the results. Cryptographic applications also require more sophisticated generators.
- Still, LCGs have a place. For example, a video game console may find an LCG to be a suitable and efficient generator.
- Note that rigorous numerical analysis is often needed to have confidence in pseudorandom generators.

DAVIDSON