Randomness Characterization in Computing and Stochastic Simulations

J. A. Rod Blais, Un. of Calgary, blais@ucalgary.ca

Introduction

• Physics: classical (deterministic) → modern (quantum physics)
• Mathematics: Lebesgue measures → probabilistic measures
• Computations: deterministic → pseudo random and stochastic
• Simulations: computations → Monte Carlo and stochastic
• Predictions: simulations → stochastic and probabilistic
• Practical considerations: exact rigorous computations are not always possible
  stochastic simulations often offer practical solutions
• Applications to direct and inverse problems

True Random Number Sequences

Basic: physical phenomena 'known' to be random
Examples:
  - HotBits service based on radioactive decay
  - Quants generated by quantum mechanical process
  - Random generated by atmospheric noise (radio static)

Expected Variances in Monte Carlo Simulations

Numerical Recipes state:

\[ \int f(x) dx \approx \frac{1}{N} \sum_{i=1}^{N} f(x_i) \]

implying a variance \( O(1/N) \)

In general, then with \( N \) data values,

• True Random Numbers \( \Rightarrow O(1/N) \) error variance
• Pseudo Random Numbers \( \Rightarrow O(1/N) \) error variance
• Chaotic Random Numbers \( \Rightarrow O(1/N) \) error variance
• Quasi Random Numbers \( \Rightarrow O((\ln N)^2/N) \) error variance

but

• Exact rigorous computations are not always possible

Randomness

• Mathematics
  - Randomness applies only to processes
  - 'Lawlessness' = algorithmic incompressibility
  - Axiomatization in terms of non-deterministic processes

• Physics
  - Unpredictable chaotic random processes
  - Unpredictable quantum random processes

• Computational Science
  - Unreproducible computations
  - Algorithmic probabilistic entropy
  - Computer code of shortest description

Pseudo Random Number Sequences

Basic: computational rounding-off or related errors
Common Methodology:
  - using some linear congruential model applied recursively
  - or lagged Fibonacci congruential sequence, such as
    \[ x_n \parallel x_{n-p} \parallel x_{n-q} \parallel \text{mod} \ p \]
  - where \( p, q \) are chosen so that \( x_n \) has period \( \prod_{i=1}^{p} i \) modulo \( p \)

Importance Sampling

Variable of integration may be transformed for better results

Axiomatization in terms of non-deterministic processes

Distributional Aspects

Randomness in data sequences does not obviously imply low discrepancy nor a uniform distribution as can easily be seen in sample spatial and spectral plots.

Spatial plots of random numbers show clumping effects in places and open gaps in other places. Such spatial discrepancies are usually unrelated to the distributional properties of the sequences. Some quasi random sequences are especially designed to have low discrepancy characteristics (without necessarily being random).

Well known procedures can be used to transform a random variate \( y \) with a distribution \( p(y) \) into another variate \( x \) with distribution \( p(x) \). The transformation follows the usual approach with Jacobians in integrals:

\[ p(y) dy = p(x) \left| \frac{dx}{dy} \right| \]

which can obviously be simplified with uniform distributions.

Gibbs Sampler

A Gibbs sampler is a technique for generating random variables indirectly from some (marginal) distribution without calculating the density.

In conventional Monte Carlo applications, random variables are required with some assumed distribution often derived somehow from other random variables having known distributional characteristics. Most random number generators are designed to produce a uniform distribution of random numbers over the unit interval \((0, 1)\).

In practice, it really depends on the application context to decide on the most appropriate Gibbs sampler. For example, in digital image restoration, the Gibbs sampler is often based on immediate pixel neighborhoods for the Markov random field (see e.g. Geman & Geman, 1984).

Quasi Random Number Sequences

Basic: computational sequences of low discrepancy

Markov Chain Monte Carlo Modeling

Markovian properties imply that only the immediate past transition probabilities need to be considered in current simulations. This greatly simplifies the modeling and the analysis.

The transition probabilities are often modeled in terms of decreasing ‘temperatures’ to simulate annealing processes converging to some appropriate uniform distribution. This is described as ‘stoichiometric relaxation’ in digital image and similar restoration.

Applications and Conclusions

In Monte Carlo volume estimation and stochastic simulations, the randomness requirements can be quite different:

In the former, randomness is often secondary to the distributional aspects of the data sequences. In fact, quasi random numbers of the deterministic type can give the best results, essentially \( O(1/N^2) \) with \( N \) data values.

In the latter, however, randomness can be critical for the probabilistic aspects of the simulations. For instance, in digital image restoration, the equivalence of the Gibbs distribution and the Markov random field is explicitly used in the stochastic modeling and restoration.

Applications abound in geomatics, gecosience and elsewhere (see e.g. Blais & Zhang, 2011; Blais, 2010; Blais, 2009; Blais et al, 2008).