Stochastic Simulation Introduction

Bo Friis Nielsen

Applied Mathematics and Computer Science

Technical University of Denmark

2800 Kgs. Lyngby – Denmark

Email: bfn@imm.dtu.dk

Practicalities

- Reading material available online, with some suggestions for further reading
- Course evaluation is: passed/not passed - based on exercise reports and report over final projects. Some individual contributions are needed in the project report.
- Teachers:
  - Bo Friis Nielsen, e-mail bfni@dtu.dk
  - Nikolaj M Primault (npri@dtu.dk), Jonas Bruun Hubrechts (s194312@student.dtu.dk), Hannah Uhre Nielsen (s194335@student.dtu.dk), Anders Reenberg Andersen (arean@dtu.dk)/ Nikolaj Normann Holm (nnho@dtu.dk)

Significance

- One of the most (The most?)important Operations Research techniques
- Several modern statistical techniques rely on simulation

What is simulation?

- (From *Concise Oxford Dictionary*): To simulate:
  To pretend, to act like, to mimic, to imitate.
- Here: *Computer experiments with mathematical model*
- **Stochastic simulation**
  To (have a computer) simulate a system which is affected by randomness.

  Narrow sense: To generate (pseudo)random numbers from a prescribed distribution (e.g. Gaussian)
- Computer experiments with mathematical model
- General engineering technique
- Analytical/numerical solutions

Why simulate?

- Real system expensive
- Mathematical model too complex
- Get idea of dynamic behaviour

Related areas

- Statistics
  - Machine learning
- Computer science
- Operations research

Target group

- Methodology course of general interest
- Of special importance for students specialising in
  - Computer Science
  - Statistics/Machine Learning
  - Operations Research
  - Planning and management

Course goal

- Topics related to scientific computer experimentation
- Specialised techniques
  - Random number generation
  - Random variable generation
  - The event-by-event principle
  - Variance reduction methods
- Simulation based statistical techniques
  - Markov chain Monte Carlo
  - Bootstrap
- Validition and verification of models
- Model building

Recommended reading

- Sheldon M. Ross: Simulation, fifth edition, Academic Press 2013 *available online for DTU students*
- Søren Asmussen and Peter W. Glynn: Stochastic Simulation: Algorithms and Analysis, Springer 2007, *available online for DTU students*
- C.P. Robert and G. Casella: Introducing Monte Carlo Methods with R, Springer, 2010
- Reuven Y. Rubinstein and Benjamin Melamed: Modern Simulation and Modelling, John Wiley & Sons 1998, *First 50 pages available from the website. It is illegal to distribute these notes*
- Villy Bæk Iversen: Numerisk Simulation (In Danish), DTU, 2007

Supplementary reading

- Averill M. Law: Simulation Modeling and Analysis, McGraw-Hill 2015
- Jerry Banks, John S. Carson II, Barry L. Nelson, David M. Nicol: Discrete-Event System Simulation, Prentice and Hall 1999
- Brian Ripley: Stochastic Simulation, John Wiley & Sons 1987
- Jack P. C. Kleijnen: Statistical Tools for Simulation Practitioneers, Marcel Dekker 1987

Knowledge/science in simulation

- Modelling skill
- Statistical methods - it is necessary to understand statistical methodology
- OR - Stochastic Processes
- Technical skills
  - Random number generations
  - Sampling from distributions
  - Variance reduction techniques
  - Statistical techniques bootstrap/MCMC
- General purpose/and specialised simulation software

Discrete versus continuous

- Discrete event simulation
- as opposed to continuous simulation
- mixed models

Probability basics

- $0 \leq \mathsf{P}(A) \leq 1 \qquad \mathsf{P}(\Omega) = 1 \qquad \mathsf{P}(\emptyset) = 0$
- $A \cap B = \emptyset \Rightarrow \mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B)$
- Complement rule $\mathsf{P}(A^c) = 1 - \mathsf{P}(A)$
- Difference rule for $A \subset B$: $\mathsf{P}(B \cap A^c) = \mathsf{P}(B) - \mathsf{P}(A)$
- Inclusion, exclusion for 2 events

  $\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B)$
- Conditional probability: for A given B (partial information): $\mathsf{P}(A|B) = \frac{\mathsf{P}_{(A \cap B)}}{\mathsf{P}_{(B)}}$
- Multiplication rule: $\mathsf{P}(A \cap B) = \mathsf{P}(B)\mathsf{P}(A|B)$
- Law of total probability (B_i is a partitioning):

  $\mathsf{P}(A) = \sum_i \mathsf{P}(B_i)\mathsf{P}(A|B_i)$
- Bayes theorem: (B_i is a partitioning): $\mathsf{P}(B_i|A) = \frac{\mathsf{P}_{(A|B_i)}\mathsf{P}_{(B_i)}}{\sum_j \mathsf{P}_{(A|B_j)}\mathsf{P}_{(B_j)}}$
- independence: $\mathsf{P}(A|B) = \mathsf{P}(A|B^c) \qquad (\mathsf{P}(A \cap B) = \mathsf{P}(A)\mathsf{P}(B))$

Random variables

- Mapping from sample space to the real line
- Probabilities defined in terms of the preimage
- Most probabilitistic calculations are performed with only a slight reference to the underlying sample space

Random variables

- Random variables: *maps* *outcomes* to real values
  - Distribution $\mathsf{P}(X=x)$ $\qquad \sum_x \mathsf{P}(X=x)=1$
  - Joint distribution
    $\mathsf{P}(X=x,Y=y) \qquad \sum_{x,y}\mathsf{P}(X=x,Y=y)=1$
  - Marginal distribution $\mathsf{P}_X(X=x)=\sum_y \mathsf{P}(X=x,Y=y)$
  - Conditional distribution $\mathsf{P}(Y=y|X=x)=\frac{\mathsf{P}(X=x,Y=y)}{\mathsf{P}_X(X=x)}$
  - independence
    $\mathsf{P}(Y=y,X=x)=\mathsf{P}_X(X=x)\mathsf{P}_Y(Y=y), \quad \forall(x,y)$
- Mean value $\mathsf{E}(X)=\sum x\cdot\mathsf{P}(X=x)$
- General expectation $\mathsf{E}(g(X))=\sum_x g(x)\cdot\mathsf{P}(X=x)$
- Linearity $\mathsf{E}(aX+bY+c)=a\mathsf{E}(X)+b\mathsf{E}(Y)+c$

Continuous random variables

- Uniform distribution of two variables: $\mathsf{P}\left((x,y)\in C\right)=\frac{A(C)}{A(D)}$
- Continuous random variables
  - ◇ Density: $f(x)\geq 0,\quad \int f(x)\mathrm{d}x=1,\quad \mathsf{P}(X\in \mathrm{d}x)\stackrel{\sim}{=} f(x)\mathrm{d}x$
  - ◇ Mean, variance (moments): $\mathsf{E}(X)=\int xf(x)\mathrm{d}x$

    $\mathsf{E}(g(X))=\int g(x)f(x)\mathrm{d}x,\quad \mathsf{E}\left(X^k\right)=\int x^k f(x)\mathrm{d}x$
- Normal distribuion: $f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\quad Z=\frac{X-\mu}{\sigma}$
- Joint densities

$f(x,y)\mathrm{d}x\mathrm{d}y\stackrel{\sim}{=}\mathsf{P}(x\leq X\leq x+\mathrm{d}x, y\leq Y\leq y+\mathrm{d}y), f(x,y)\geq 0$

- Joint distribution

$$F(x,y)=\mathsf{P}(X\leq x, Y\leq y)=\int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)\mathrm{d}u\mathrm{d}v$$

Continuous random variables continued

- Conditional continous distributions $f_Y(y|X=x) = \frac{f(x,y)}{f_X(x)}$
- Integral version of law of total probability
  $\mathsf{P}(A) = \int \mathsf{P}(A|X=x) f_X(x) \mathsf{d}x$
- Conditional expectation $\mathsf{E}(Y) = \mathsf{E}(\mathsf{E}(Y|X))$
- Covariance/corellation

$$\mathsf{Cov}(X,Y) = \mathsf{E}[(X-\mathsf{E}(X))(Y-\mathsf{E}(Y))] = \mathsf{E}(XY) - \mathsf{E}(X)\mathsf{E}(Y)$$

$$\mathsf{Corr}(X,Y) = \frac{\mathsf{Cov}(X,Y)}{\mathsf{SD}(X)\mathsf{SD}(Y)}$$

- (X,Y) independent $\Rightarrow \mathsf{Corr}(X,Y) = 0$
- Variance of sum of variables
  $\mathsf{Var}\left(\sum_{k=1}^{N} X_k\right) = \sum_{k=1}^{n} \mathsf{Var}(X_k) + 2\sum_{1\leq j<k\leq n} \mathsf{Cov}(X_j, X_k)$
- Bilinearity of covariance
  $\mathsf{Cov}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{m} b_j Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \mathsf{Cov}(X_i, Y_j)$