Over the last 50 years, computational techniques have become increasingly important in statistics. Simulation of stochastic systems may be used, for example, to *evaluate* new statistical methodology. Further, it has become recognised that sophisticated simulation methods can be used as the *basis* of new techniques, in statistics, as well as in problems that arise in optimisation and operations research.

This course provides an account of such simulation methods, and covers areas from basic techniques of random variate generation, to modern computational techniques in statistics, such as Markov chain Monte Carlo (MCMC) methods, and bootstrap techniques. Material from M2S1 forms a firm foundation.

The course will consist of two (roughly equal) elements. The first will consist of simulation techniques, and the second will focus on the use of these methods in statistical inference.

The starting point of our discussion, and the basis for everything we cover, is the generation of **pseudo-random numbers** (a deterministic sequence of numbers which has the characteristics of a stream of random variables, uniformly distributed on $[0, 1]$). Then the key objective becomes that of converting these, in an efficient manner, into **random variables from general distributions**: various slick ideas are used. In the first bit of the course we will discuss also ideas of **Monte Carlo integration**, and of the various trick ideas ('computer swindles') that can be utilised with the objective of **variance reduction**, getting the same computational accuracy from a smaller simulation.

The key uses of simulation in statistical inference that we will discuss are: **Monte Carlo tests**, **MCMC techniques in Bayesian inference** and **bootstrap inference**.

Monte Carlo tests use simulation to build up the distribution of a test statistic under some null hypothesis being tested on data (so replacing the need to 'look up tables'). MCMC techniques are based on the idea that we can sample from some probability distribution of interest (such as the *posterior* distribution in a Bayesian context) by the (strange looking) device of constructing, then simulating, a Markov chain which has the desired distribution as its *equilibrium distribution*. [We will describe all the necessary elements of Markov chain theory we need]. In bootstrap inference, data from some unknown probability distribution are used to construct an *empirical sampling model*: statistical inference is then performed on the basis of samples simulated from the empirical model.

There will be one assessed project, accounting for $\sim 25\%$ of the available course credit. The dead line for the course work is Friday, **13 Dec 2019**. The remaining $\sim 75\%$ will be allocated to a final examination in the summer term. For M3 students, this will be a 1.5-hour exam consisting of three questions; M4 students will sit a 2-hour exam consisting of four questions.

I am planning on setting several pieces of work for formative feedback.

Additional course information:

- Lecture slides and notes will be updated on Blackboard shortly after each lecture.
- Problem sheets will be published on Blackboard before the tutorials; solutions will be posted around a week later.
- The timetable is as follows:
  - Mondays (lecture) @ 3pm in Huxley 139;
  - Mondays (tutorial) @ 4pm in Huxley 414;
  - Fridays (lecture) @ 1pm in Huxley 340 (week 1: Huxley 144).
  - Office hour on Fridays @ 2pm
- My email: e.mccoy@imperial.ac.uk
- My office: Huxley 552.