Assignment Topic №2: Univariate Simulation

Dmytro Marushkevych

September 6, 2022

The second assignment topic is univariate simulation. If you draw the topic "Univariate Simulation" at the oral exam, you will have to present a solution of one of the two assignments below.

Remember the five points:

- How can you test that your implementation is correct?
- Can you implement alternative solutions?
- Can the code be restructured e.g. by modularization, abstraction or object oriented programming to improve readability?
- How does the implementation perform (benchmarking)?
- Where are the bottlenecks (profiling), and what can you do about them?

When comparing simulation algorithms it may be necessary to use the same pseudo random numbers (set the seed) or to compute results of comparable accuracy as measured by e.g. the standard error.

Both of the two simulation assignments are made very concrete. You can solve the assignments by writing code that is very specific, and this is a good way to get started. But you should (re)write the final solution to be more generic.

Assignment 1: Rejection sampling

This assignment uses the Poisson data available from Absalon. It is a table with 100 rows and two variables, z and x, where z takes integer values and x takes positive real values.

The purpose of this assignment is to sample from the probability distribution on $[0, \infty)$ with density

$$f(y) \propto \prod_{i=1}^{100} \exp\left(y z_i x_i - e^{y x_i}\right), \quad y \geq 0.$$

Find a Gaussian envelope of f and implement rejection sampling from the distribution with density f using this envelope.

Implement the adaptive rejection sampling algorithm that uses a piecewise log-affine envelope and compare it with the one based on the Gaussian envelope.

You are welcome to invent other envelopes and try to optimize the choice of envelope. However, for the exam the focus should be on the implementations and their comparisons and not on theoretical derivations of envelopes.

For those interested, the distribution given by f above is a simple example of a Bayesian posterior distribution in a Poisson regression model.

Assignment 2: Importance sampling

Let $X_1, X_2, \ldots$ be i.i.d. uniformly distributed on $(-1.9, 2)$ and define

$$S_n = 30 + \sum_{k=1}^{n} X_k.$$

Think of X_k as the capital that a company or a gambler earns or loses at time n. Starting with a total capital of 30, S_n is the total capital at time n. We are interested in computing the *ruin probabilities* before time n:

$$p(n) = P(\exists k \leq n : S_k \leq 0).$$

Implement Monte Carlo integration to compute $p(100)$.

Define

$$g_{\theta,n}(x) = \frac{1}{\varphi(\theta)^n} \exp\left(\theta \sum_{k=1}^{n} x_k\right)$$

for $x \in (-1.9, 2)^n$ where

$$\varphi(\theta) = \int_{-1.9}^{2} e^{\theta z} \mathrm{d}z.$$

Implement importance sampling to compute $p(100)$ by sampling from $g_{\theta,100}$ for suitable θ. Is there an optimal choice of θ?

Note that you should consider number of samples as well as run time to discuss what "optimal" means.