20 Probability

20.2 Importance Sampling and Fast Simulation (5 units)

This project assumes some basic knowledge of discrete-time Markov Chains, as covered in the Part IB course.

1 Importance sampling

Importance sampling is a technique for simulating random variables. Suppose that X is a random variable with density π, and suppose we wish to use simulation to estimate the probability that X takes a value in some set A. If X is difficult to simulate, or if the event that X is in A is very rare, then this might be hard to do.

Suppose that we have some other random variable Y with density π', such that $\pi'(x) > 0$ whenever $\pi(x) > 0$. Let L be the likelihood ratio,

$$L(y) = \frac{\pi(y)}{\pi'(y)}.$$

Let $\gamma = \mathbb{P}(X \in A)$, and consider the estimator

$$\hat{\gamma} = L(Y)1[Y \in A]$$

where $1[\cdot]$ is the indicator function. We call π' the *twisted density.*

> **Question 1** Show that $\hat{\gamma}$ is an unbiased estimator of γ. How could you run a simulation to estimate γ using this fact?

This method for estimating γ is called *importance sampling.* Now let X be an exponential random variable with mean 3, and consider the event $B = \{X > 30\}$. Suppose we wish to estimate $\mathbb{P}(B)$ by importance sampling, using as our twisted distribution an exponential with mean λ^{-1}.

> **Question 2** What is $\mathbb{P}(B)$? Write a program to estimate this probability using importance sampling. Simulate for different values of λ and include in your report some typical outputs for each λ. *(Programming hint. If U is a uniform random variable on $[0, 1]$, then $-\log U$ is an exponential random variable with mean 1.)*

You should find that some values of λ lead to better estimators than others.

> **Question 3** Modify your program to estimate how long a simulation you need in order to obtain a good estimate, for a range of values of λ. Explain your method. What value of λ seems best? Give an intuitive reason why it is so.

> **Question 4** Prove that for $\lambda \geqslant \frac{2}{3}$ the simulation is useless. Calculate the optimal value of λ.

This shows that importance sampling isn't automatically good—it is important to choose the distribution carefully.

Of course, the probability that an exponential random variable exceeds some amount doesn't need simulating, but it is still useful to be able to do importance sampling of exponential random variables for the following reason:

The exponential distribution is often used to model the time until an event occurs. Now imagine trying to simulate, say, breakdowns in a large system. It is useful to be able to increase the rate of breakdowns while leaving the rate of other events unchanged. Importance sampling can be used to do this.

2 Fast simulation

Importance sampling can be applied to Markov chains, when it is often called *fast simulation*. Suppose we have a discrete-time Markov chain with jump probabilities P_{ij}. The path which the Markov chain takes can be regarded as a random variable in the space of sample paths, and we might be interested in the probability that the path is in some particular set of paths.

Consider another Markov chain with *twisted jump probabilities* P'_{ij} such that $P'_{ij} > 0$ whenever $P_{ij} > 0$.

> **Question 5** If we observe a path $x_0 \to x_1 \to \cdots \to x_n$, what is the likelihood ratio L of that path?

Consider a simple random walk $(X_n)_{n\geqslant 0}$ on the non-negative integers with up probability $p < 1/2$ and down probability $1-p$, except up probability 1 when the walk is at 0. We might be interested in the event that, on a single excursion away from 0, the random walk hits some high level C before it returns to 0.

> **Question 6** Calculate the probability of this event, for $p = 1/4$ and $C = 30$. That is, find $\mathbb{P}(T_C < T_0 | X_0 = 0)$, where $T_x = \inf\{n \geqslant 1 : X_n = x\}$, for these values of the parameters.

This probability is too small to simulate directly. Now consider a fast simulation using another random walk with up probability $p' > 1/2$ and down probability $1-p'$.

> **Question 7** What is the probability that the twisted random walk reaches C on a single excursion away from 0? How does this change as C becomes large, compared to the random walk with up probability p? What consequences does this have for simulation?

> **Question 8** Carry out a simulation to estimate the probability referred to in Question 6. Try a range of values of $p' > 1/2$. Comment on your results, keeping in mind that the estimate should be unbiased.

> **Question 9** In the case $p' = 1-p$, calculate the exact distribution of your estimator. How close would you expect your estimator to be to the true value, after $10,000$ trials?

This model might represent the behaviour of a buffer in a computer network, where we would be interested in the probability that the buffer became full and started losing messages. Again, a single buffer does not need simulating, but in a more complicated network, it may not be possible to calculate the overflow probability analytically. The sort of overflow probability which one sees in real computer networks is often of the order of 10^{-8}, so there is a genuine need for fast simulation in these situations.

