# Part A Simulation Trinity 2013, Problem Sheet 2

1. The random variable $X$ has probability mass function

$$p(x;s) = \frac{1}{\zeta(s)}\frac{1}{x^s}, \qquad \text{for } x = 1, 2, 3, ....$$

   (a) The normalising constant $\zeta(s)$ is hard to calculate, However, when $s = 2$ we do have $\zeta(2) = \pi^2/6$. Give an algorithm to simulate $Y \sim p(y; 2)$ by inversion.

   (b) Give a rejection algorithm simulating $X \sim p(x) = p(x; s)$ for $s > 2$ using the rejection algorithm and draws from $Y \sim q(y)$ where the proposal is $q(y) = p(y; 2)$. You will need to derive the upper bound $M' \geq \tilde{p}(x)/\tilde{q}(x)$ for all $x$.

   (c) Compute the expected number of simulations of $Y \sim q$ for each simulated $X \sim p$ in the previous part question, giving your answer in terms of $\zeta(s)$.

2. Suppose $X \sim N(0, \sigma^2)$ and we want to estimate $\mu_\phi = \mathbb{E}(\phi(X))$ for some function $\phi(x)$ known to have finite mean and variance. Suppose we have samples $Y = (Y_1, ..., Y_n)$ with $Y_i \sim N(0, 1), i = 1, 2, ..., n$ iid. Here are two estimators for $\mu_\phi$ given in terms of $Y$:

$$\widehat{\theta}_{1,n} = \frac{1}{n}\sum_{i=1}^{n}\phi(\sigma Y_i)$$

and

$$\widehat{\theta}_{2,n} = \frac{1}{n\sigma}\sum_{i=1}^{n} e^{-Y_i^2(1/2\sigma^2 - 1/2)}\phi(Y_i).$$

   (a) Show that $\widehat{\theta}_{1,n}$ and $\widehat{\theta}_{2,n}$ are unbiased and give the expression of their variances.

   (b) What range of values must $\sigma$ be in for $\widehat{\theta}_{2,n}$ to have finite variance? Can you give a weaker condition if it is known that $\int_{-\infty}^{\infty}\phi^2(x)\,dx < \infty$?

   (c) Why might we prefer $\widehat{\theta}_{2,n}$ to $\widehat{\theta}_{1,n}$, for some values of $\sigma^2$ and functions $\phi$? (Hint: consider estimating $\mathbb{P}(X > 1)$ with $\sigma \ll 1$).

3. We are interested in performing inference about the parameters of internet traffic model.

   (a) The arrival rate $\Lambda_t$ for packets at an internet switch varies with time $t$ according to a log-normal distribution with parameters $\mu$ and $\sigma$. Let $\Lambda_t$ be the arrival rate in the $t$'th millisecond interval $(t, t+1]$. The probability density $p_\Lambda(\lambda)$ of log-normal $\Lambda_t$ is

$$p_\Lambda(\lambda; \mu, \sigma) = \frac{1}{\lambda\sqrt{2\pi\sigma^2}}\exp\left(-(\log(\lambda) - \mu)^2/2\sigma^2\right),$$

   independent of time $t$ and of events in all other intervals. Show that if $V \sim N(\mu, \sigma^2)$ and we set $W = \exp(V)$ then $W \sim \text{LogNormal}(\mu, \sigma)$.

   (b) Given an arrival rate $\Lambda_t = \lambda_t$, the number $N_t$ of packets which actually arrive in the $t$'th time interval has a Poisson distribution, $N_t \sim \text{Poisson}(\lambda_t)$, and again, this is conditionally independent from one interval to the next. Observations $N_1 = n_1, ..., N_k = n_k$ are made of the numbers of arrivals. Let $N = (N_1, ..., N_k)$ and $n = (n_1, ..., n_k)$. We wish to estimate $\mu$ and $\sigma$, the parameters of the underlying rate variation. Show that the likelihood for $\mu$ and $\sigma$ is

$$\mathbb{P}(N = n|\mu, \sigma) \propto \prod_{t=1}^{k} L(\mu, \sigma; n_t)$$

   with

$$L(\mu, \sigma; n_t) \propto \mathbb{E}(\Lambda^{n_t}\exp(-\Lambda)|\mu, \sigma).$$

(c) Explain how you would estimate $L(\mu, \sigma; n_t)$ by simulating values for $\Lambda$ from its prior.

(d) Suppose now we have $m$ iid samples

$$\Lambda^{(j)} \sim \text{LogNormal}(\mu, \sigma), j = 1, 2, ..., m$$

for one pair of ($\mu$,$\sigma$)-values. Give an importance sampling estimator for $L(\mu', \sigma'; n_t)$ at new parameter values $(\mu', \sigma') \neq (\mu, \sigma)$, in terms of the $\Lambda^{(j)}$'s.

(e) For what range of $\mu', \sigma'$ values can the $\Lambda^{(j)}$-realisation be safely 'recycled' in this way?