Part A Simulation and Statistical Programming Hilary 2014
Problem Sheet 2, due Wednesday 5pm Week 5 at SPR1

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) Implement your algorithm as an R function. Your function should take as input an integer $n > 0$ and return a vector of n iid realisations of $Y \sim p(y; 2)$. State briefly how you checked your code.

   (c) Give a rejection algorithm simulating $X \sim p(x) = p(x; s)$ for $s > 2$ using draws from $Y \sim q(y)$ with $q(y) = p(y; 2)$.

   (d) 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 $\zeta(s)$.

   (e) Implement your algorithm as an R function. Your function should take as input s and return as output $X \sim p(x; s)$ and the number of trials N it took to simulate X.

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 μ_ϕ given in terms of Y:

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

   (a) Show that $\widehat{\theta}_{1,n}$ and $\widehat{\theta}_{2,n}$ are unbiased and give their variances in terms of expectations of ϕ etc.

   (b) What range of values must σ 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 σ^2 and functions ϕ? (Hint: consider estimating $\mathbb{P}(X > 1)$ with $\sigma \ll 1$).

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

   (a) The arrival rate Λ for packets at an internet switch has a log-normal distribution LogNormal(μ, σ) with parameters μ and σ. The LogNormal(μ, σ) probability density is

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

   Show that if $V \sim N(\mu, \sigma^2)$ and we set $W = \exp(V)$ then $W \sim$ LogNormal(μ, σ).

   (b) Given an arrival rate $\Lambda = \lambda$, the number N of packets which actually arrive has a Poisson distribution, $N \sim$ Poisson(λ). Suppose we observe $N = n$. Show that the likelihood $L(\mu, \sigma; n)$ for μ and σ is

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

   (c) Give an algorithm simulating $\Lambda \sim$ LogNormal(μ, σ) using $Y \sim N(0, 1)$ as a base distribution, and explain how you could use simulated Λ-values to estimate $L(\mu, \sigma; n)$ by simulating values for Λ.

   (d) Suppose now we have m iid samples

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

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

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