5 Monte Carlo (AJ Chapter 8)

The (*) questions from sheets 5 to 9 form homework 2.

1. [The Monte Carlo method for calculating $\int h(x)f(x)dx$ where f is a p.d.f is to sample $X_1, \ldots X_n \sim f$ then calculate $\frac{1}{n}\sum_{i=1}^{n} h(X_i)$.]

    (a) Estimate the absolute value of a normal $\mathrm{N}(0,1)$ random variable. Estimate how accurate your answer is.

    (b) How does $\int_{[0,1]^n}\int_{[0,1]^n} \|x-y\|_p \,\mathrm{d}x\mathrm{d}y$ behave when n is large (and $p \in \{1,2\}$, say). [This is related to *the curse of dimensionality.*]

    (c) Suppose some statistic X_n is measured annually ($n = 2014, 2015, ...$; the number of accidents per year or something). Suppose the X_n and independent samples from the Poisson(λ) distribution. How large does λ have to be for the percentage change (this year compared to last year) to be typically less than 10%.

2. (*) [Inversion sampling: The c.d.f. for a random variable X is

    $$F(x) = \mathbb{P}(X \leq x).$$

    The generalized inverse c.d.f. is

    $$F^{-1}(u) = \inf\{u : F(x) \geq u\}.$$

    If $U \sim \text{Uniform}(0,1)$ then $F^{-1}(U)$ has c.d.f F.]
    Use the inversion method, and the R function `runif`, to sample from the geometric distribution on $1, 2, \ldots$ with mean $1/p$,

    $$\mathbb{P}(X = k) = (1-p)^{k-1}p, \qquad k = 1, 2, \ldots$$

3. (*) Let U denote the set

    $$\{(x_1, x_2, x_3) \in (0,1)^3 : \sqrt{x_1} + \sqrt[3]{x_2} + \sqrt[4]{x_3} < 1\}.$$

    We can estimate the volume of U using the rectangle method, applied to the indicator function of U

```
n=100; i=(2*(1:n)-1)/(2*n)
a=0
for (j1 in i) {
  for (j2 in i) {
    for (j3 in i) {
      if (j1**(1/2)+j2**(1/3)+j3**(1/4)<1) a=a+1
    }
  }
}
print(a/n**3)
[1] 0.000724
```

(a) Sampling from the uniform distribution on $[0,1]^2$ or $[0,1]^3$, use Monte Carlo integration to estimate the volume of U. Estimate the accuracy of your answers.

(b) [Rejection sampling: Given a two distributions f and g such that $f(x) \leq cg(x)$ for all x, to sample from the f distribution:

- sample from the g distribution
- accept the sample with probability $f(x)/(cg(x))$. If you reject, try again.]

Use rejection sampling to sample (X_1, X_2, X_3) from the uniform distribution on U. Estimate the expected values of X_1, X_2 and X_3, and the accuracy of your estimates.

(c) [Importance sampling: Estimate $\int h(x)f(x)dx$ by sampling $Y_1, \ldots Y_n \sim g$ and then calculate either

$$\frac{1}{n}\sum_{i=1}^{n} h(Y_i)f(Y_i)/g(Y_i)$$

or (if you only know f up to a multiplicative constant)

$$\sum_{i=1}^{n} [h(Y_i)f(Y_i)/g(Y_i)] \,/ \sum_{i=1}^{n} [f(Y_i)/g(Y_i)]$$

]

Sample Y_1, Y_2, Y_3 independently from the exponential distribution, with mean values as computed in part (b). Use Y_1, Y_2, Y_3 and importance sampling to estimate the expected values of X_1, X_2, X_3 when (X_1, X_2, X_3) is uniformly distributed on U. Estimate the accuracy of your estimates.

4. [Sequential importance sampling with rejection control: SIS is an incremental version of importance sampling. It can be combined with rejection sampling in a variety of clever ways to save you from exploring parts of the sample space with low weights.

(a) Let $\pi = \pi(x_1, \ldots, x_d) = \pi(x_{1:d})$ denote a probability density function for $X = (X_1, \ldots, X_d) = X_{1:d} \in \mathbb{R}^d$.

(b) For $i = 1, \ldots, d$, let $\pi_i = \pi_i(x_{1:i})$ denote positive functions on $\mathbb{R}^i$ such that $\pi \propto \pi_d$.

(c) For $i = 1, \ldots, d$, let g_i denote proposal probability density functions on $\mathbb{R}^i$.

(d) Let $c_1, \ldots, c_d$ denote threshold values.

Taking the threshold values equal to zero gives ordinary importance sampling with respect to the distribution $g_1(x_1)g_2(x_2 \mid x_1) \ldots g_d(x_d \mid x_{1:d-1})$. In that case, $\bar{w}_i \equiv w_i$ for all i.

Step 1:
Sample $x_1 \sim g_1(x_1)$.
Let $w_1 = \pi_1(x_1)/g_1(x_1)$.
Accept x_1 with probability $\min\{1, w_1/c_1\}$. Let $\bar{w}_1 = \max\{w_1, c_1\}$.

Step $t+1$: $(t = 1, \ldots, d-1)$
Sample $x_{t+1} \sim g_t(x_{t+1} \mid x_{1:t})$.
Let $w_{t+1} = \bar{w}_t \, \pi_{t+1}(x_{1:t+1}) / \big[g_{t+1}(x_{t+1} \mid x_{1:t})\pi_t(x_{1:t})\big]$.
Accept x_{t+1} with probability $\min\{1, w_{t+1}/c_{t+1}\}$.
Let $\bar{w}_{t+1} = \max\{w_{t+1}, c_{t+1}\}$.

If Y is a sample from the SIS procedure then

$$\mathbb{E}_{X\sim\pi}[h(X)] = \frac{\mathbb{E}[h(Y)\bar{w}_d]}{\mathbb{E}[\bar{w}_d]}.$$

]
Suppose that the logarithm of the size of a population of fish at times $t = 1, \ldots, d$ is given by the vector $X = (X_1, \ldots, X_d)$. Assume that X is an auto-regressive process ($|\lambda| < 1$),

$$X_1 = Z_1, \qquad Z_1 \sim \text{Normal}(\mu, \sigma^2/(1-\lambda^2))$$

$$X_t = Z_t + \lambda X_{t-1}, \quad Z_t \sim \text{Normal}((1-\lambda)\mu, \sigma^2), \qquad t = 2, \ldots, d$$

The population X cannot be observed directly. Instead we observe the number of fish Y caught in nets

$$Y_t \sim \text{Poisson}(\exp(X_t)/100), \qquad t = 1, \ldots, d.$$

Let f denote the joint distribution of X and Y

$$f(x, y) = f(x)f(y \mid x).$$

We can use SIS to to sample from the posterior distribution

$$\pi(x) := f(x \mid y) \propto f(x)f(y \mid x).$$

For $t = 1, \ldots, d$ let

$$\pi_t(x_{1:t}) = f(x_{1:t})f(y_{1:t} \mid x_{1:t}), \qquad g_t(x_{1:t}) = f(x_{1:t}).$$

The importance weights are given by

$$w_1 = \frac{\pi_1(x_1)}{g_1(x_1)} = f(y_1 \mid x_1).$$

$$w_{t+1} = \bar{w}_t \frac{\pi_{t+1}(x_{1:t+1})}{\pi_t(x_{1:t})g_{t+1}(x_{t+1} \mid x_{1:t})} = \bar{w}_t f(y_{t+1} \mid x_{t+1})$$

```
ag_process<-function(mu,sigma,lambda,n) {
  x=c()
  x[1]=rnorm(1,mu,sigma/sqrt(1-lambda**2))
  for (i in 2:n)
    x[i]=x[i-1]*lambda + rnorm(1,mu*(1-lambda),sigma)
  x
}
#Generate observations Y
x=ag_process(10,0.3,0.75,10) #Pretend you cannot see this!
y=rpois(10,exp(x)/100)
```

(a) Use SIS (without rejection control) to sample x conditional on y. Produce an empirical plot of the distribution of the log weights.

(b) Implement SIS with rejection control.

(c) How can part (b) be made more efficient? [Rejection Control and Sequential Importance Sampling. Jun S. Liu; Rong Chen; Wing Hung Wong.http://www.people.fas.harvard.edu/~junliu/TechRept/98folder/liu98_s.pdf]