Problem Set: Numerical Integration and Simulation

In this problem set, we will explore numerical integration and simulation. For simulation, we will illustrate importance sampling, and also the GHK recursive approach.

1. Consider computing the mean of a univariate standard normal distribution, truncated to the region [0,1]:

$$E(x) = \frac{\int_0^1 x\phi(x)dx}{\int_0^1 \phi(x)dx}$$

where $\phi(x)$ denotes the standard normal density.

(1a) Compute $E(x)$ explicitly, using quadrature routines (eg. `quad` in Matlab). Use this as the "true value" of $E(x)$.

(1b) Simulate using a "crude" accept-reject routine:

Take 50 random draws x^s from $N(0,1)$. Reject the draws which lie outside the truncation region. Take average of the non-rejected draws

$$\frac{\sum_s x^s * \mathbf{1}(x^s \in [0,1])}{\sum_s \mathbf{1}(x^s \in [0,1])}$$

as an estimate of $E(x)$.

Repeat this 100 times to get 100 estimates of $E(x)$. Characterize the distribution of these estimates (ie. mean, median, standard deviation, etc.).

(1c) Simulate using importance sampling:

Take 50 random draws x^s from $U[0,1]$. Multiply each draw by the importance sampling weight, which in this case is just $w^s = \frac{\phi(x^s)}{\int_0^1 \phi(s)ds}$. Take average of the weighted draws

$$\frac{1}{50} \sum_s x^s w^s$$

as an estimate of $E(x)$.

Repeat this 100 times to get 100 estimates of $E(x)$. Characterize the distribution of these estimates (ie. mean, median, standard deviation, etc.).

Compare the accuracy of the results in (1b) and (1c).

2. Now we consider a bivariate example. Let (x, y) be distributed according to the bivariate normal density $BN(0, 0, 1, 1, 0.5)$, the bivariate density with zero mean $[0, 0]'$ and variance matrix $\Sigma = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$, but truncated to the region $x \in [0,1]$; $y \in [-1,2]$.

Consider calculating

$$E(xy) = \frac{\int_0^1 \int_{-1}^2 xy\phi(x,y)dydx}{\int_0^1 \int_{-1}^2 \phi(x,y)dydx}$$

where $\phi(x,y)$ denotes the density of the bivariate normal distribution described above.

(2a) Compute $E(xy)$ explicitly, using quadrature routines. This requires double integration routines, such as `dblquad` in Matlab. Use this as the "true value" of $E(xy)$.

(2b) Simulate using a "crude" accept-reject routine:

Take 100 random draws x^s, y^s from $\phi(x,y)$. As we described in class, one way to do this is to draw u^s, v^s, independently from the $N(0,1)$ distribution, then transform them to the desired draws as

$$\begin{pmatrix} x^s \\ y^s \end{pmatrix} = \Sigma^{1/2} \cdot \begin{pmatrix} u^s \\ v^s \end{pmatrix}$$

where $\Sigma^{1/2}$ denotes the lower-triangular Cholesky factorization of Σ (make sure that you use the lower triangular version).

Then reject the draws which lie outside the truncation region. Take average of the non-rejected draws

$$\frac{\sum_s x^s \cdot y^s \cdot \mathbf{1}(x^s \in [0,1], y^s \in [-1,2])}{\sum_s \mathbf{1}(x^s \in [0,1], y^s \in [-1,2])}$$

as an estimate of $E(xy)$.

Repeat this 100 times to get 100 estimates of $E(xy)$. Characterize the distribution of these estimates (ie. mean, median, standard deviation, etc.).

(2c) Simulate using GHK:

Generate 100 random draws of (x^s, y^s) using the recursive GHK algorithm. For each draw, also calculate the associated truncation probability τ^s (as defined in lecture notes).

Estimate $E(xy)$ as

$$\frac{\sum_s x^s \cdot y^s \cdot \tau^s}{\sum_s \tau^s}$$

Repeat this 100 times to get 100 estimates of $E(xy)$. Characterize the distribution of these estimates (ie. mean, median, standard deviation, etc.).

Compare the accuracy of the results in (2b) and (2c).

5. Assume you have a simple discrete choice model where the choice of replacing the bus engine is given by the model:

$$y_t = 1(\beta * x_t > \epsilon_t)$$

where $y_t = 1$ means that the engine was replaced, and the coefficient $\beta = 0$. Also assume that ϵ_t is first-order Markov, and

$$\epsilon_t = 0.9\epsilon_{t-1} + \eta_t$$

and η_t is iid $N(0, 1)$.

Assume you observe the sequence

$$
\begin{array}{cc}
y_1 & 1 \\
y_2 & 0 \\
y_3 & 1 \\
y_4 & 1
\end{array}
$$

Assume $\epsilon_0 = 0$.

(a) Compute the exact likelihood for the observations, using quadrature methods.

(b) Simulate the likelihood for the observations, using a crude frequency simulator. Repeat 100 times, and report the mean and standard deviation.

(c) Simulate the likelihood for the observations, using particle filtering. Repeat 100 times, and report the mean and standard deviation.