SDS 386D HW1

Saturday, February 3, 2018 2:43 PM

#1. Write the joint density:

$f(x,u) \propto M^{-1} 1(u < l(x)/M)\, h(x)\, 1(u \leq 1)$

where $l(x) = (-\log x)^2 \cdot x^2(1-x)/6$

$h(x) = 6x(1-x)$

Find a bound M for $l(x)$ on $(0,1)$

Although the analytical supremum of $l(x)$ is not easy to find, we can plot $l(x)$ numerically and pick $M = \frac{1}{60}$, which is not far from the sharp bound.

My algorithm:

1. Obtain X from p.d.f $h(x)$:

   the c.d.f of $h(x)$ is $3x^2 - 2x^3 =: H(x)$

   $H^{-1}(x)$ on $(0,1)$ is implemented by ~~using~~ a binary search between $(0,1)$ so that it can be easily adapted to higher order polynomials

   and obtain U from a uniform distribution on $(0,1)$

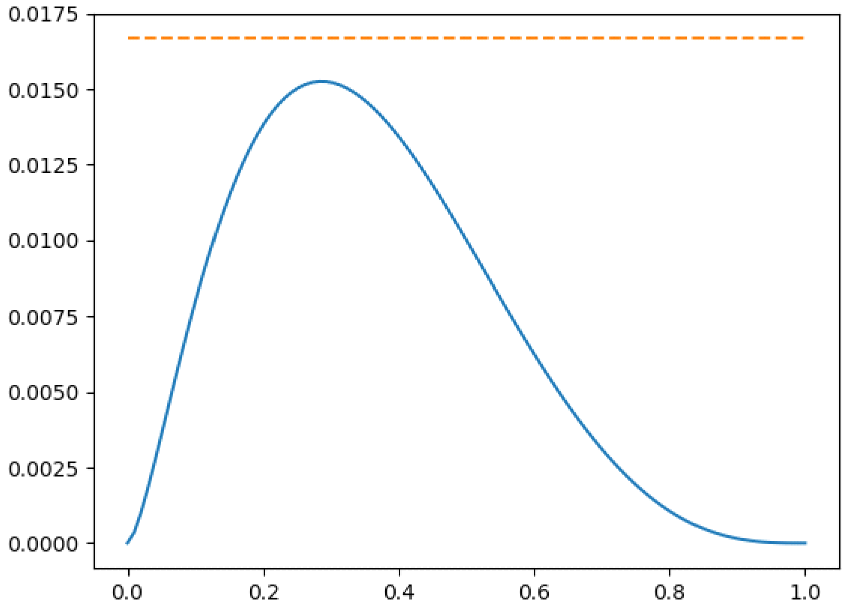
2. check whether $u < \frac{l(x)}{M}$

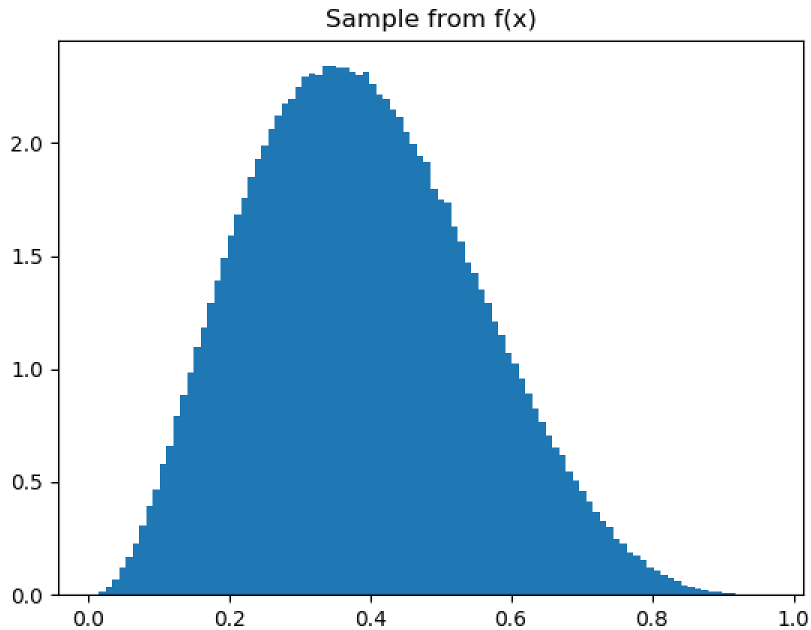
   if this holds, accept X as a sample from $f(x)$

   else: reject X and sample again.

SDS 386D HW1

Saturday, February 3, 2018 2:43 PM





Run the rejection sampling algorithm until a million samples are generated. The samples drawn are shown in the above histogram. We can see that it fits well with curve of f(x).

SDS 386D HW1

Saturday, February 3, 2018 2:43 PM

Acceptance probability:

The numerical test shows that the acceptance probability is 51.03%.

To get the actual probability, we need to know the normalizing constant in this problem

$$c = \int_0^1 (-\log x)^2 x^3 (1-x)^2$$

This can be obtained by { numerical integration, Monte Carlo integration, Analytically

$c = \frac{919}{108000}$ from the third method and a MC integration using samples from $f(x)$ confirms this result.

$\Rightarrow \Pr\left(u < \frac{l(x)}{M}\right) = \frac{c}{M} \approx 0.5105$ which is consistent with the test result.

An implementation of adaptive rejection sampling is also tested: this implementation follows the method described in the lecture. The proposed piecewise probability enveloping the original probability is restricted to have have 50 pieces at maximum.

SDS 386D HW1

Saturday, February 3, 2018 2:43 PM

Generate a million accepted samples as in rejection sampling. Only 15089 samples are rejected throughout the whole process and the histogram is similar.

SDS 386D HW1

Saturday, February 3, 2018 2:43 PM

#2. Use the sampling method from #1, a MC integration can be carried out on this problem:

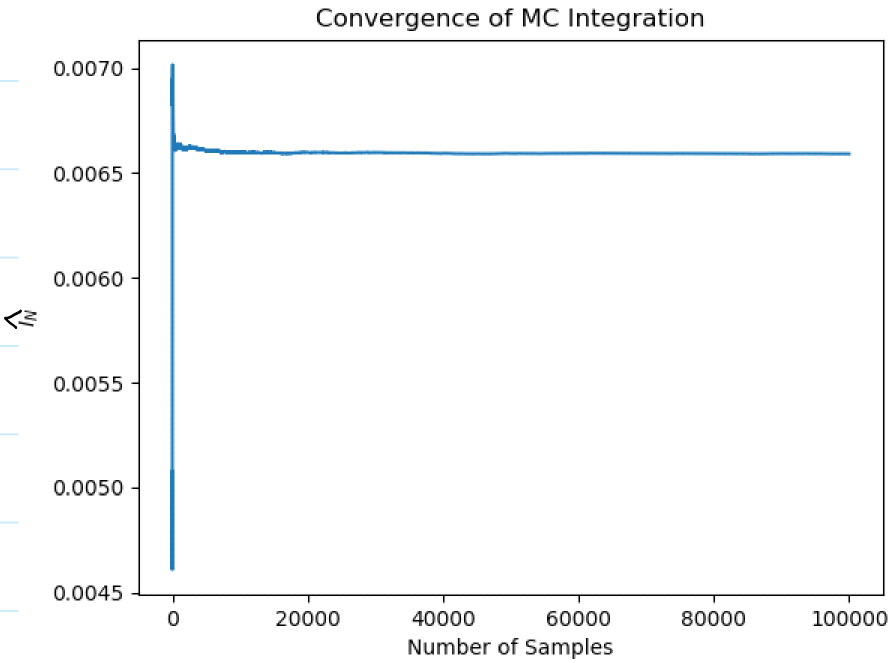
$$I = \int_0^1 c\, f(x)\, (1-x)^{1/2}\, dx.$$

where $c = 919/108000$

Take 10^5 samples from $f(x)$ and use

$$\hat{I}_N = \frac{1}{N}\sum_{i=1}^{N} c\,(1-X_i)^{1/2}$$

The numerical ~~cost~~ gives 0.00659 which is close to the exact solution.



As shown in the convergence plot, the MC approx converges to the exact integration as $N \uparrow$

SDS 386D HW1

Saturday, February 3, 2018 2:43 PM

Variance:

From lecture, the variance of $\hat{I}_N$ is

$$Var(\hat{I}_N) = \frac{1}{N} Var(g(X_1))$$

$$= \frac{1}{N} \cdot \left[E(g(X)^2) - E(g(X))^2 \right]$$

$$= \frac{1}{N} \left[\int g^2(X) f(X) \, dX - \left(\int g(X) f(x) dx \right)^2 \right]$$

$$\approx \frac{1}{N} \left[\sum_{i=1}^{N} g^2(X_i) / N - \hat{I}_N^2 \right]$$ where X_i is from p.d.f $f(x)$.

Numerical test shows the $Var(\hat{I}_N)$ at $N=1000$ is around 8×10^{-7} (Using the same sample $\{X_i\}$ and $\{g(X_i)\}$)

SDS 386D HW1

Saturday, February 3, 2018 2:43 PM

#3. $l(x) = f(x)/h(x) = \pi(1+x^2) \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

$$0 = l'(x) = \pi(1+x^2)\frac{1}{\sqrt{2\pi}} \cdot e^{-x^2/2}(-x) + \pi(2x)\frac{1}{\sqrt{2\pi}} \cdot e^{-x^2/2}$$

$$\Rightarrow \quad (1+x^2)x = 2x$$

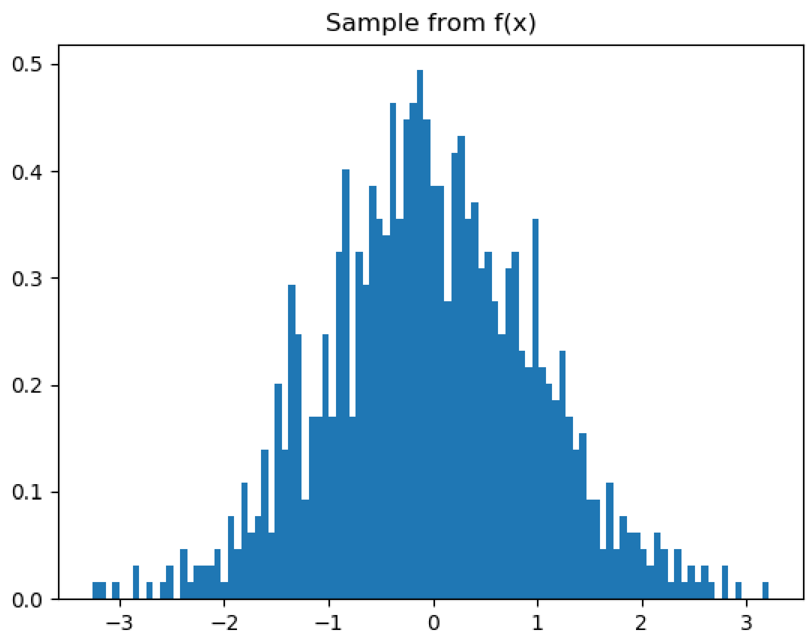
$x=0$ is a saddle point

$x \neq 0$: $\quad 1+x^2 = 2 \quad x = \pm 1$

$$l(\pm 1) = \sqrt{\frac{2\pi}{e}} = M$$

$$P_r = \frac{1}{M} = \sqrt{\frac{e}{2\pi}} \approx 65.8\%$$

Get a 1000 sample from simulation : the acceptance probability is about 64% ~ 68%, and the histogram is shown below. As the number of samples is increased, the acceptance probability will converge to the exact value.



SDS 386D HW1

Saturday, February 3, 2018 2:43 PM