P137-练习5.13:

Find two importance functions f_1 and f_2 that are supported on $(1,\infty)$ and are "close" to

$$g(x) = \frac{x^2}{\sqrt{2\pi}}\, e^{-x^2/2}, \qquad x > 1.$$

Which of your two importance functions should produce the smaller variance in estimating

$$\int_1^\infty \frac{x^2}{\sqrt{2\pi}}\, e^{-x^2/2}\, dx$$

by importance sampling? Explain.

Obtain a Monte Carlo estimate of

$$\int_1^\infty \frac{x^2}{\sqrt{2\pi}}\, e^{-x^2/2}\, dx$$

by importance sampling.

$$g(x) = \begin{cases} \frac{x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, & x > 1 \\ 0, & x \le 1 \end{cases}$$

取:

$$f_1(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Big/ (1-\Phi(1)), & x > 1 \\ 0, & x \le 1 \end{cases}$$

$$F_1(x) = \begin{cases} \frac{\Phi(x)-\Phi(1)}{1-\Phi(1)}, & x > 1 \\ 0, & x \le 1 \end{cases}$$

$$F_1^{-1}(u) = \Phi^{-1}(u(1-\Phi(1))+\Phi(1)) \quad 0<u<1$$

$$f_2(x) = \begin{cases} x e^{-\frac{x^2}{2}} \Big/ e^{-\frac{1}{2}}, & x > 1 \\ 0, & x \le 1 \end{cases}$$

Rayleigh(1) 分布:

$f(x) = x e^{-\frac{x^2}{2}}, x \ge 0.$

$F(x) = \int_{-\infty}^{x} t e^{-\frac{t^2}{2}} dt$

$= -e^{-\frac{t^2}{2}}\Big|_{-\infty}^{x} = 1 - e^{-\frac{x^2}{2}}, \quad x>0$

$\int_1^{+\infty} f(x)\,dx = 1 - F(1) = e^{-\frac{1}{2}}$

$$F_2(x) = \begin{cases} \frac{e^{-\frac{1}{2}} - e^{-\frac{x^2}{2}}}{e^{-\frac{1}{2}}}, & x > 1 \\ 0, & x \le 1 \end{cases}$$

$F_2^{-1}(u) = \sqrt{1-2\ln(1-u)}, \quad 0<u<1$

$$\theta = \int_1^{+\infty} \frac{x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$= \begin{cases} \int_{-\infty}^{+\infty} \frac{g(x)}{f_1(x)} f_1(x)\,dx = E_X\left[\frac{g(X)}{f_1(X)}\right], & X \sim f_1 \\ \int_{-\infty}^{+\infty} \frac{g(x)}{f_2(x)} f_2(x)\,dx = E_X\left[\frac{g(X)}{f_2(X)}\right], & X \sim f_2. \end{cases}$$

① $u_1, \cdots, u_m$ i.i.d $\sim U(0, 1)$,

$x_i = \Phi^{-1}(u_i(1-\Phi(1)) + \Phi(1))$, $i = 1, \cdots, m$.

$$\hat{\theta}_{I,1} = \frac{1}{m} \sum_{i=1}^{m} \frac{g(x_i)}{f_1(x_i)}.$$

$$Var(\hat{\theta}_{I,1}) = \frac{1}{m} Var\left(\frac{g(X)}{f_1(X)}\right), \quad \widehat{Var}(\hat{\theta}_{I,1}) = \frac{1}{m}\left[\frac{1}{m-1}\sum_{i=1}^{m}\left(\frac{g(x_i)}{f_1(x_i)} - \hat{\theta}_{I,1}\right)^2\right]$$

② $u_1, \cdots, u_m$ i.i.d $\sim U(0, 1)$.

$x_i = \sqrt{1 - 2\ln(1-u_i)}$, $i = 1, \cdots, m$.

$$\hat{\theta}_{I,2} = \frac{1}{m} \sum_{i=1}^{m} \frac{g(x_i)}{f_2(x_i)}$$

$$\widehat{Var}(\hat{\theta}_{I,2}) = \frac{1}{m}\left[\frac{1}{m-1}\sum_{i=1}^{m}\left(\frac{g(x_i)}{f_2(x_i)} - \hat{\theta}_{I,2}\right)^2\right]$$

课本P137.

5.15 得到例5.13中的分层重要抽样估计并与例5.10中的结果进行比较.

例5.10：计算积分$\theta=\int_0^1 \frac{e^{-x}}{1+x^2}dx$

$$= \int_{-\infty}^{+\infty} \frac{g(x)}{f(x)} f(x)\,dx, \quad g(x)=\begin{cases} \frac{e^{-x}}{1+x^2}, & 0<x<1 \\ 0, & o.w. \end{cases}, \quad f(x)=\begin{cases} \frac{e^{-x}}{1-e^{-1}}, & 0<x<1 \\ 0, & o.w. \end{cases}$$

$$= E_X\left[\frac{g(X)}{f(X)}\right], \quad X\sim f.$$

重要抽样法 $\hat\theta = 0.52506988$, $se(\hat\theta) = 0.09658794$ 为 σ 的估计.

$x_1, \cdots, x_M$ i.i.d $\sim f$, $\quad \hat\theta = \frac{1}{M}\sum_{i=1}^{M} g(x_i)/f(x_i)$, $\quad Var(\hat\theta) = \frac{1}{M}\underbrace{Var\left(\frac{g(X)}{f(X)}\right)}_{\sigma^2}$, $X\sim f$

$$\begin{cases} u_1, \cdots, u_M \text{ i.i.d} \sim U(0,1) \\ x_i = F^{-1}(u_i) = -\log[1-(1-e^{-1})u_i], \; i=1,\cdots,M \end{cases}$$

Example 6.14 (Example 6.11, cont.). In Example 6.11 our best result was obtained with importance function $f_3(x) = e^{-x}/(1-e^{-1})$, $0 < x < 1$. From 10000 replicates we obtained the estimate $\hat\theta = 0.5257801$ and an estimated standard error 0.0970314. Now divide the interval (0,1) into five subintervals, $(j/5, (j+1)/5)$, $j = 0, 1, \ldots, 4$.

Then on the j^{th} subinterval variables are generated from the density

$$\boxed{\frac{5e^{-x}}{1-e^{-1}}}, \qquad \frac{j-1}{5} < x < \frac{j}{5}.$$

$$f_j(x)=\begin{cases} \frac{e^{-x}}{e^{-\frac{(j-1)}{5}}-e^{-\frac{j}{5}}}, & \frac{j-1}{5}<x<\frac{j}{5} \\ 0, & o.w. \end{cases}$$

The implementation is left as an exercise. ◇

分层重要抽样：$\theta = \int_0^1 \frac{e^{-x}}{1+x^2}dx = \sum_{j=1}^{5}\int_{\frac{j-1}{5}}^{\frac{j}{5}} \frac{e^{-x}}{1+x^2}dx \triangleq \sum_{j=1}^{5}\theta_j$

对 $j=1,\cdots,5$，

$$\theta_j = \int_{-\infty}^{+\infty}\frac{g_j(x)}{f_j(x)}f_j(x)\,dx, \quad g_j(x)=\begin{cases} \frac{e^{-x}}{1+x^2}, & \frac{j-1}{5}<x<\frac{j}{5} \\ 0, & o.w. \end{cases}$$

$$= E_X\left[\frac{g_j(X)}{f_j(X)}\right], \quad f_j(x)=\begin{cases} \frac{e^{-x}}{e^{-\frac{(j-1)}{5}}-e^{-j/5}}, & \frac{j-1}{5}<x<\frac{j}{5} \\ 0, & o.w. \end{cases}$$

$X\sim f_j$

故 $\hat\theta_j = \frac{1}{m}\sum_{k=1}^{m}\frac{g_j(x_k)}{f_j(x_k)}$, $\quad x_1, \cdots, x_m$ i.i.d $\sim f_j$, $\quad m=\frac{M}{k}$

$$\begin{cases} u_1, \cdots, u_m \text{ i.i.d} \sim U(0,1) \\ x_i = -\log\left[e^{-\frac{(j-1)}{5}} - u\left(e^{-\frac{j-1}{5}} - e^{-\frac{j}{5}}\right)\right], \; i=1,\cdots,m. \end{cases}$$

$$\hat\theta = \sum_{j=1}^{5}\hat\theta_j$$

$$\widehat{Var}(\hat\theta) = \sum_{j=1}^{5}\hat\sigma_j^2/m = \frac{k}{M}\sum_{j=1}^{5}\hat\sigma_j^2, \quad \sigma_j^2 = Var_X\left(\frac{g_j(X)}{f_j(X)}\right), \; X\sim f_j$$