KTH Electrical Engineering

Electric Power Systems Lab
EG2080 MONTE CARLO METHODS IN ENGINEERING
14 January 2012

Formulae

Random variables

$$P(X \leq x) = F_X(x) = \int_{-\infty}^{x} f_X(x)dx$$

$$E[X] = \int_{-\infty}^{\infty} f_X(x)x dx$$

$$Var[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2 = \int_{-\infty}^{\infty} f_X(x)(x - E[X])^2 dx$$

Random numbers

U is $U(0, 1)$-distributed, $Y = F_Y^{-1}(U) \Rightarrow Y$ has the distribution function $F_Y(x)$

Simple sampling

$m_X = \frac{1}{n}\sum_{i=1}^{n} x_i$ is an estimate of $E[X]$

$s_X^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m_X)^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - m_X^2$ is an estimate of $Var[X]$

$$Var[M_X] = \frac{Var[X]}{n}$$

Coefficient of variation: $a_X = \frac{s_X}{m_X\sqrt{n}}$

M_X normally distributed $\Rightarrow m_X \pm \delta$, where $\delta = \frac{t_\alpha \cdot s_X}{\sqrt{n}}$, is a 100α % confidence interval

($t_{0.95} = 1.9600$, $t_{0.99} = 2.5758$, $t_{0.999} = 3.2905$)

$T_1 \cdot Var[M_{X1}] < T_2 \cdot Var[M_{X2}] \Rightarrow$ Simulation method 1 is more efficient than method 2

Complementary random numbers

$u_i \in U(0, 1), y_i = F_Y^{-1}(u_i),\ y_i^* = F_Y^{-1}(1 - u_i),\ x_i = g(y_i), x_i^* = g(y_i^*), i = 1, \ldots, n$

$$\Rightarrow \quad m_X = \frac{1}{2n}\sum_{i=1}^{n}(x_i + x_i^*) \text{ is an estimate of } E[X]$$

Dagger sampling

$$f_Y(x) = \begin{cases} p & \text{if } x = a, \\ 1 - p & \text{if } x = b, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{where } p < 0.5$$

$$F^{\ddagger}_{Yj}(x) = \begin{cases} a & \text{if } (j-1)p \le x < jp, \\ b & \text{otherwise,} \end{cases} \quad \text{for } j = 1, \ldots, S$$

(where S is the largest integer such that $S \le 1/p$)

U is $U(0, 1)$-distributed, $Y_j = F^{\ddagger}_{Yj}(U) \Rightarrow Y_j$ has the frequency function $f_Y(x)$

Control variates

$\mu_Z = E[\tilde{g}(Y)],\ x_i = g(y_i), z_i = \tilde{g}(y_i),\ i = 1, \ldots, n$

$$\Rightarrow \quad m_X = \frac{1}{n}\sum_{i=1}^{n}(x_i - z_i) + \mu_Z \text{ is an estimate of } E[X]$$

Correlated sampling

$x_{1,i} = g_1(y_i), x_{2,i} = g_2(y_i), i = 1, \ldots, n$

$$\Rightarrow \quad m_{(X1-X2)} = \frac{1}{n}\sum_{i=1}^{n}(x_{1i} - x_{2i}) \text{ is an estimate of } E[X_1 - X_2]$$

Importance sampling

$$\int_{-\infty}^{\infty} f_Z(\psi)d\psi = 1,\ f_Z(\psi) > 0\ \forall\ \psi \in \Upsilon$$

$u_i \in U(0, 1), y_i = F_Z^{-1}(u_i),\ x_i = g(y_i), i = 1, \ldots, n$

$$\Rightarrow \quad m_X = \frac{1}{n}\sum_{i=1}^{n}\frac{f_Y(y_i)}{f_Z(y_i)}x_i \text{ is an estimate of } E[X]$$

$$s_X^2 = \frac{1}{n}\sum_{i=1}^{n}\frac{f_Y(y_i)}{f_Z(y_i)}x_i^2 - m_X^2 \text{ is an estimate of } Var[X]$$

$$Var[M_X] = \frac{1}{n}\left(\int_{\psi \in \boldsymbol{R}} g^2(\psi)\frac{f_Y^2(\psi)}{f_Z(\psi)}d\psi - \mu_X^2\right)$$

$$f_Z(\psi) = \frac{g(\psi)f_Y(\psi)}{\mu_X} \Rightarrow Var[M_X] = 0$$

Stratified sampling

$$\mathcal{X}_h \subset \mathcal{X},\ \bigcup_{h=1}^{L} \mathcal{X}_h = \mathcal{X},\ \mathcal{X}_h \cap \mathcal{X}_k = \varnothing \text{ if } h \neq k$$

$\omega_h = P(X \in \mathcal{X}_h)$, m_{Xh} is an estimate of $E[X_h]$, s_{Xh}^2 is an estimate of $Var[X_h]$

$$\Rightarrow \quad m_X = \sum_{h=1}^{L} \omega_h m_{Xh} \text{ is an estimate of } E[X]$$

$$Var[M_X] = \sum_{h=1}^{L} \omega_h^2 Var[M_{Xh}]$$

Neyman allocation: $n_h = n\dfrac{\omega_h \sigma_{Xh}}{\sum_{k=1}^{L} \omega_k \sigma_{Xk}}$