△ Antithetic variables

Goal : reduce the variance (compared with simple MC)

SMC : ① generate u from Unif $(0,1)$

② Inverse transform

Steps : ① Generate $\frac{m}{2}$ replicates (Suppose we need m replicates)

$$\begin{cases} Y_j = g\left(F_X^{-1}(u_1^{(j)}), \cdots, F_X^{-1}(u_n^{(j)})\right) \\ Y_j' = g\left(F_X^{-1}(1-u_1^{(j)}), \cdots, F_X^{-1}(u_n^{(j)})\right) \end{cases}$$

② Calculate $\hat{\theta} = \frac{1}{m} \sum_{i=1}^{m/2} (Y_j + Y_j')$ sample average

eg:

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$= E\left\{ x e^{-\frac{x^2 u^2}{2}} \right\} \quad \text{where } u \sim \text{Unif}(0,1)$$

△ Control variates

$$\theta = E\{g(x)\}$$

If there is a function f with $\mu = E\{f(x)\}$ known

and $f(x)$ is correlated with $g(x)$ then define

$\hat{\theta}_c = g(x) + c\,(f(x) - \mu)$ is called an estimator

with control variates

Remarks : ① μ known

② correlated

③ $\hat{\theta}_c$ is unbiased

How to choose c to reduce the variance :

$$Var(\hat{\theta}_c) = Var(g(x)) + 2c \cdot Cov(g(x), f(x)) + c^2 Var(f(x))$$

is a quadratic function of c

$$c^* = -\frac{Cov(g(x), f(x))}{Var(f(x))}$$

$$Var(\hat{\theta}_{c^*}) = Var(g(x)) - \frac{\{Cov(g(x), f(x))\}^2}{Var(f(x))}$$

Percentage of variance reduction

$$\frac{Var(g(x)) - Var(\hat{\theta}_{c^*})}{Var(g(x))} \times 100\%$$

$$= \frac{\{Cov(g(x), f(x))\}^2}{Var(g(x)) \cdot Var(f(x))} \times 100\% = \{Corr(f(x), g(x))\}^2 \times 100\%$$

Remark: ① $f(x)$ is called the control variate

② $\hat{\theta}_c$ is good when $f(x)$ and $g(x)$ are strongly correlated

③ In practice, we need to estimate c,

estimate $Cov(f(x), g(x))$ and $Var(f(x))$

eg. 5.7. $\theta = E e^u = \int_0^1 e^x dx \qquad g(u) = e^u$

control variate $f(u) = u \sim Unif(0,1)$

$$c^* = -\frac{cov(e^u, u)}{Var(u)} = -1.69$$

$$\hat{\theta}_{c^*} = e^u - 1.69(u - \tfrac{1}{2})$$

eg. 5.8. $\theta = E g(x) = \int_0^1 \frac{e^{-x}}{1+x^2} dx$

$g(x) = \frac{e^{-x}}{1+x^2}$ $\quad f(x) \begin{cases} \text{close to } g(x) \text{ in shape} \\ \text{easy to obtain } E\{f(x)\} \end{cases}$

choose $f(x) = \frac{e^{-\frac{1}{2}}}{1+x^2}$ $\qquad \frac{g}{f} = \frac{e^{-x}}{e^{-\frac{1}{2}}}$ $\quad (\because X \sim Unif(0,1))$

$\therefore \frac{g}{f}$ 接近 const.

$$\mu = \int_0^1 \frac{e^{-\frac{1}{2}}}{1+x^2} dx = \frac{\pi}{4} e^{-\frac{1}{2}}$$

△ Antithetic variable as control variates

Control variate : linear combination of two unbiased estimators

Suppose two unbiased estimators $\hat{\theta}_1$ and $\hat{\theta}_2$

$$\hat{\theta}_c = c \cdot \hat{\theta}_1 + (1-c)\hat{\theta}_2 \quad (0 \le c \le 1) \text{ is unbiased for } \theta$$

eg. $\hat{\theta}_c = g(x) + c(f(x) - \mu)$

$$= (1-c)\underbrace{g(x)} + c\,\underbrace{(g(x) + f(x) - \mu)} \quad \text{both unbiased}$$

$$\theta = E\,g(x) = E(g(x) + f(x) - \mu)$$

$$Var(\hat{\theta}_c) = c^2 \, Var(\hat{\theta}_1) + 2c(1-c)\, Cov(\hat{\theta}_1, \hat{\theta}_2) + (1-c)^2 \, Var(\hat{\theta}_2)$$

Assume $\hat{\theta}_1$ and $\hat{\theta}_2$ are identically dist. with $corr(\hat{\theta}_1, \hat{\theta}_2) \triangleq r$

then $Var(\hat{\theta}_c) = [(2-2r)c^2 - (2-2r)c + 1]\, Var(\hat{\theta}_1)$

$$c^* = \frac{1}{2} \Rightarrow \hat{\theta}_c = \frac{1}{2}(\hat{\theta}_1 + \hat{\theta}_2)$$

△ Extension to multiple control variates

If $E\,\hat{\theta}_i = \theta \quad i = 1, \cdots, k$

constants $C = (c_1, \cdots, c_k) \quad \sum c_i = 1, \ c_i \ge 0$

then $\sum_{i=1}^{k} c_i \hat{\theta}_i$ is unbiased

Correspondingly, $\hat{\theta}_c = g(x) + \sum_{i=1}^{k} c_i^* (f_i(x) - \mu_i)$, where $\mu_i = E\{f_i(x)\}$

△ Duality between control variate to LR (linear regression)

Simple LR : Observe $(X_1, Y_1), \cdots, (X_n, Y_n)$

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Assumptions : Linear

$E(\varepsilon | X) = 0$

$Cov(\vec{\varepsilon} | X) = \sigma^2 I$ with σ unknown const.

LS: $\min_{\beta_0, \beta_1} \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_i)^2$

$$\begin{cases} \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \\ \hat{\beta}_1 = \dfrac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})^2} = \dfrac{\widehat{Cov}(X, Y)}{\widehat{Var}(X)} \end{cases}$$

Observe, $(f(X_1), g(X_1)), \cdots, (f(X_n), g(X_n))$

$$g(X) = \beta_0 + \beta_1 f(X) + \varepsilon$$

$$\begin{cases} \hat{\beta}_0 = \overline{g(X)} - \hat{\beta}_1 \overline{f(X)} \\ \hat{\beta}_1 = \dfrac{\widehat{Cov}(f(X), g(X))}{\widehat{Var}(f(X))} = -\hat{c}^* \end{cases}$$

Remark: ① Predicted value at $\mu = E\{f(X)\}$

$$\hat{\beta}_0 + \hat{\beta}_1 \cdot \mu = \overline{g(X)} - \hat{\beta}_1 \overline{f(X)} + \hat{\beta}_1 \mu$$
$$= \overline{g(X)} - \hat{c}^* (\overline{f(X)} - \mu) = \hat{\theta}_{\hat{c}^*}$$

② Variance of control variate estimator

$$Var(\overline{g(X)} + \hat{c}^* (\overline{f(X)} - \mu))$$
$$= \frac{1}{n} Var(g(X) + \hat{c}^* (f(X) - \mu))$$
$$= \frac{1}{n} Var(g(X) - \hat{\beta}_1 f(X) - \hat{\beta}_0) = \frac{1}{n} Var(\text{residual}) = \frac{1}{n} \hat{\sigma}_\varepsilon^2$$

③ Proportion of the variance reduction

$$Corr(g(X), f(X))^2 = R^2$$

means the proportion of the variance of $g(X)$ explained by $f(X)$

Extension to multiple LR and multiple control variate

$$Y = \beta_0 + \sum_{j=1}^{p} \beta_j X_j + \varepsilon$$

eg. $g(x) = \dfrac{e^{-x}}{1+x^2}$, $f(x) = \dfrac{e^{-\frac{1}{2}}}{1+x^2}$ see in the code

△ Importance Sampling

$$\theta = \int_a^b g(x)\,dx = E\,g(x)$$

$\frac{1}{b-a}\int_a^b g(x)\,dx$: Average value of $g(x)$ over $[a,b]$

$E\,g(x)$. $X \sim \text{Unif}(a,b)$

SMC: ① Generate from $\text{Unif}(a,b)$

② Calculate $\frac{b-a}{m}\sum_{i=1}^{m} g(x_i)$

Drawback: what if $g(x)$ not be uniform over $[a,b]$

Solution: choose a non-uniform weighted function

Def: Let $X \sim f$. such that $f(x) > 0$ on the support of g

Set $Y = \frac{g(X)}{f(X)}$ then $\int g(x)\,dx = \int \frac{g(x)}{f(x)} f(x)\,dx = E_f(Y)$

IS steps: ① Generate $X_1, \cdots, X_m$ from f

② Calculate $\frac{1}{m}\sum_{i=1}^{m} Y_i = \frac{1}{m}\sum_{i=1}^{m}\frac{g(X_i)}{f(X_i)} \xrightarrow{SLLN} \int g(x)\,dx$

We call f the importance function or envelope

$\text{Var}\left(\frac{1}{m}\sum Y_i\right) = \frac{1}{m}\text{Var}(Y)$ where $Y = \frac{g(X)}{f(X)}$

$\text{Var}(Y)$ small $\Leftrightarrow$ ① $\frac{g(X)}{f(X)}$ close to const.

② f should be easy to stimulate.