Department of Statistics and Data Science at SUSTech

MAT7035: Computational Statistics

Tutorial 2: The SIR method and The mixture representation (MR) method

D. The SIR method

D.1 The background

- When the sampling from the pdf $f(x)$ with support $\mathcal{S}_X$ is very hard, we could find an importance sampling density $g(x)$, having the same support $\mathcal{S}_X$ and it is relatively easy to generate i.i.d. samples from $g(x)$.
- Then by adjusting the generated samples from $g(x)$, we can obtain *approximate* i.i.d. samples from $f(x)$.

D.2 The SIR algorithm

Step 1: Generate $X^{(1)}, \ldots, X^{(J)} \stackrel{\text{iid}}{\sim} g(\cdot)$;

Step 2: Select a subset $\{X^{(k_i)}\}_{i=1}^m$ from $\{X^{(j)}\}_{j=1}^J$ via resampling *without replacement* from the discrete distribution on $\{X^{(j)}\}$ with probabilities $\{\omega_j\}$, where

$$\omega_j = \frac{f(X^{(j)})/g(X^{(j)})}{\sum_{j'=1}^J f(X^{(j')})/g(X^{(j')})}, \quad j = 1, \ldots, J.$$

D.3 Remarks

(a) The SIR method generates samples approximately from $f(x)$; while the rejection method really produces samples exactly from $f(x)$.

(b) J/m should be large and $g(\cdot)$ should be close to $f(\cdot)$.

Example T2.1 (Standard arcsine distribution). To generate $X \sim f(x)$, where

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}}, \quad 0 < x < 1,$$

we consider a skewed beta density $\text{Beta}(x|2,4)$ as the importance sampling density $g(x)$. Thus, the importance ratio $\mathrm{w}(x) = f(x)/\text{Beta}(x|2,4)$. State the SIR algorithm and write an R code.

Solution: (i) The SIR algorithm is as follows:

Step 1: Generate $X^{(1)}, \ldots, X^{(J)} \overset{\text{iid}}{\sim} \text{Beta}(2,4)$;

Step 2: Select a subset $\{X^{(k_i)}\}_{i=1}^m$ from $\{X^{(j)}\}_{j=1}^J$ via resampling *without replacement* from the discrete distribution on $\{X^{(j)}\}$ with probabilities $\{\omega_j\}$, where

$$\omega_j = \frac{\mathrm{w}(X^{(j)})}{\sum_{j'=1}^J \mathrm{w}(X^{(j')})}, \quad j = 1, \ldots, J.$$

(ii) Let $J = 200,000$ and $m = 20,000$. The corresponding R codes are as follows:

```
> J <- 200000
> m <- 20000
> xJ <- rbeta(J, 2, 4)
> fxJ <- 1/sqrt(xJ*(1-xJ))
> w <- fxJ/dbeta(xJ, 2,4)
> om <- w/sum(w)
> sample(xJ, m, prob=om, replace=F)
```

We omit the plots here. ‖

E. The MR method

E.1 Basic idea

- Let $X \sim f_X(x)$. Statistically, we can write

$$f_X(x) = \int_{\mathbb{Y}} f_{(X,Y)}(x,y)\, \mathrm{d}y = \int_{\mathbb{Y}} f_Y(y) f_{(X|Y)}(x|y)\, \mathrm{d}y. \tag{T2.1}$$

- From $X|(Y=y) \sim f_{(X|Y)}(x|y)$, if we could find a function $q(y)$ such that the conditional distribution of $q(y)X|(Y=y)$ is free from y, i.e.,

$$q(y)X|(Y=y) \stackrel{\mathrm{d}}{=} W \sim f_W(w) \quad \text{and} \quad W \perp\!\!\!\perp Y, \tag{T2.2}$$

then we have $q(y)X|(Y=y) \stackrel{\mathrm{d}}{=} q(Y)X \stackrel{\mathrm{d}}{=} W$ or $X \stackrel{\mathrm{d}}{=} q^{-1}(Y) \cdot W$.

E.2 MR method for drawing $X \sim f_X(x)$ given by (T2.1) when Y is continuous

Step 1: Draw $Y = y \sim f_Y(y)$ and independently draw $W = w \sim f_W(w)$;

Step 2: Return $x = q^{-1}(y) \cdot w$.

Example T2.2 (F distribution). Use the MR method to generate a positive random variable X with density

$$f_X(x) = \int_0^\infty \mathrm{e}^{-y} \cdot y\, \mathrm{e}^{-yx}\, \mathrm{d}y, \quad x > 0.$$

Solution: The joint density of X and Y is

$$f_{(X,Y)}(x,y) = \mathrm{e}^{-y} \cdot y\, \mathrm{e}^{-yx} = f_Y(y) \cdot f_{(X|Y)}(x|y), \quad x, y > 0$$

so that $Y \sim \text{Exponential}(1)$ and $X|(Y=y) \sim \text{Exponential}(y) = \text{Gamma}(1,y)$. We have

$$yX|(Y=y) \sim \text{Gamma}(1,1) = \text{Exponential}(1),$$

which is independent of $Y = y$, so $YX \stackrel{\mathrm{d}}{=} W \sim \text{Exponential}(1)$ and $W \perp\!\!\!\perp Y$. We obtain

$$\begin{aligned} X &\stackrel{\mathrm{d}}{=} \frac{W}{Y} = \frac{\text{Gamma}(1,1)}{\text{Gamma}(1,1)} = \frac{2 \cdot \text{Gamma}(1,1)}{2 \cdot \text{Gamma}(1,1)} \\ &= \frac{\text{Gamma}(2/2,1/2)}{\text{Gamma}(2/2,1/2)} = \frac{\chi^2(2)/2}{\chi^2(2)/2} \sim F(2,2). \end{aligned}$$

From Example 1.1 in Lecture Notes, we have $W \stackrel{\mathrm{d}}{=} -\log(U_1)$ with $U_1 \sim U(0,1)$, so that

$$X \stackrel{\mathrm{d}}{=} \frac{\log(U_1)}{\log(U_2)},$$

where $U_1, U_2 \stackrel{\mathrm{iid}}{\sim} U(0,1)$. The MR algorithm for generating $X \sim F(2,2)$ is as follows:

Step 1: Draw $U_1 = u_1, U_2 = u_2 \stackrel{\mathrm{iid}}{\sim} U(0,1)$;

Step 2: Return $x = \log(u_1)/\log(u_2)$. ‖

Example T2.3 (JTB distribution). Use the MR method to generate a random variable X following the Johnson–Tietjen–Beckman (JBT) distribution with density

$$f_X(x) = \int_{x^{1/r}}^{\infty} \frac{y^{\alpha-r-1}\,\mathrm{e}^{-y}}{\Gamma(\alpha)}\,\mathrm{d}y = \int_0^{\infty} \frac{y^{\alpha-r-1}\,\mathrm{e}^{-y}}{\Gamma(\alpha)} \cdot I(y > x^{1/r})\,\mathrm{d}y, \quad x > 0,$$

where $\alpha, r > 0$ are two shape parameters.

Solution: The joint density of X and Y is

$$\begin{aligned} f_{(X,Y)}(x,y) &= \frac{y^{\alpha-r-1}\,\mathrm{e}^{-y}}{\Gamma(\alpha)} \cdot I(0 < x < y^r) = \frac{y^{\alpha-1}\,\mathrm{e}^{-y}}{\Gamma(\alpha)} \cdot \frac{I(0 < x < y^r)}{y^r - 0} \\ &= f_Y(y) \cdot f_{(X|Y)}(x|y), \end{aligned}$$

so that $Y \sim \mathrm{Gamma}(\alpha, 1)$ and $X|(Y = y) \sim U(0, y^r)$. We have

$$\left.\frac{X - 0}{y^r - 0}\right|(Y = y) = y^{-r}X|(Y = y) \sim U(0,1),$$

which is independent of $Y = y$, so $Y^{-r}X \stackrel{\mathrm{d}}{=} W \sim U(0,1)$ and $W \perp\!\!\!\perp Y$. We obtain $X \stackrel{\mathrm{d}}{=} Y^r W$. The MR algorithm for generating $X \sim f_X(x)$ is as follows:

Step 1: Draw $Y = y \sim \mathrm{Gamma}(\alpha, 1)$ and independently draw $W = w \sim U(0,1)$;

Step 2: Return $x = y^r \cdot w$. ‖

F. From one finite mixture distribution to SR

F.1 The issue and goal

- In (T2.1), when Y is an integer random variable with pmf

$$f_Y(k) = \Pr(Y = k) = p_k, \quad k \in \mathbb{K},$$

we can write (T2.1) as

$$f_X(x) = \sum_{k \in \mathbb{K}} p_k f_k(x), \tag{T2.3}$$

where each $p_k > 0$ and $\sum_{k \in \mathbb{K}} p_k = 1$.

- The general finite mixture distribution can be expressed as

$$f_X(x) = \sum_{i=1}^{n} p_i f_{X_i}(x) \quad \text{or} \quad F_X(x) = \sum_{i=1}^{n} p_i F_{X_i}(x), \tag{T2.4}$$

where $X_i \sim f_{X_i}(\cdot)$ and $\{p_i\}_{i=1}^n$ are probability weights. The goal is to generate a sample from $X \sim f_X(x)$.

F.2 From density representation to SR

- The density representation (T2.4) is equivalent to the following random variables representation:

$$X = \begin{cases} X_1, & \text{with probability } p_1, \\ X_2, & \text{with probability } p_2, \\ \vdots & \quad\vdots \\ X_n, & \text{with probability } p_n. \end{cases} \tag{T2.5}$$

- When $n = 2$, (T2.5) becomes

$$X = \begin{cases} X_1, & \text{wp } \ p, \\ X_2, & \text{wp } \ 1-p \end{cases} \quad \text{or} \quad X - X_1 = \begin{cases} 0, & \text{wp } \ p, \\ X_2 - X_1, & \text{wp } \ 1-p \end{cases} \tag{T2.6}$$

with the following SR

$$X - X_1 \stackrel{\mathrm{d}}{=} Z(X_2 - X_1) \quad \text{or} \quad X \stackrel{\mathrm{d}}{=} (1-Z)X_1 + ZX_2, \tag{T2.7}$$

where $X_i \sim f_{X_i}(\cdot)$ for $i = 1, 2$; $Z \sim \text{Bernoulli}\,(1-p)$ and $Z \perp\!\!\!\perp \{X_1, X_2\}$.

- **Proof of (T2.7):** To verify (T2.7), we need to show that the cdf of $(1-Z)X_1 + ZX_2$ is identical to $F_X(x)$ defined in (T2.4) with $n = 2$. In fact, the cdf of $(1-Z)X_1 + ZX_2$ is

$$\begin{aligned}
& \Pr\{(1-Z)X_1 + ZX_2 \leqslant x\} \\
= & \sum_{z=0}^{1} \Pr(Z = z) \cdot \Pr\{(1-Z)X_1 + ZX_2 \leqslant x | Z = z\} \\
= & \Pr(Z = 0) \cdot \Pr(X_1 \leqslant x | Z = 0) \\
& + \Pr(Z = 1) \cdot \Pr(X_2 \leqslant x | Z = 1) \\
= & \; p\Pr(X_1 \leqslant x) + (1-p)\Pr(X_2 \leqslant x) = pF_{X_1}(x) + (1-p)F_{X_2}(x),
\end{aligned}$$

which is identical to $F_X(x)$ defined in (T2.4) with $n = 2$. $\square$

- The SR of X defined by (T2.5) is

$$X \stackrel{\mathrm{d}}{=} Z_1 X_1 + \cdots + Z_n X_n, \tag{T2.8}$$

where $X_i \sim f_{X_i}(\cdot)$, $1 \leqslant i \leqslant n$; $\mathbf{z} = (Z_1, \ldots, Z_n)^\top \sim \text{Multinomial}(1; p_1, \ldots, p_n)$, and $\mathbf{z} \perp\!\!\!\perp \{X_1, \ldots, X_n\}$.

F.3 SR method for drawing $X \sim f_X(x)$ given by (T2.4) with $n = 2$

Step 1: Draw $X_i = x_i \sim f_{X_i}(\cdot)$ for $i = 1, 2$ and independently draw $Z = z \sim \text{Bernoulli}\,(1-p)$;

Step 2: Return $x = (1-z)x_1 + zx_2$.

F.4 SR method for drawing $X \sim f_X(x)$ given by (T2.4)

Step 1: Draw $X_i = x_i \sim f_{X_i}(\cdot)$ for $i = 1, \ldots, n$ and independently draw $\mathbf{z} = \boldsymbol{z} = (z_1, \ldots, z_n)^\top \sim \text{Multinomial}(1; p_1, \ldots, p_n)$;

Step 2: Return $x = z_1 x_1 + \cdots + z_n x_n$.

Example T2.4 (Polynomial distribution). Use the SR method to generate a random variable X following the polynomial distribution with density

$$f_X(x) = \sum_{i=1}^{n} c_i x^{i-1}, \quad 0 < x < 1,$$

where $\{c_i\}$ are positive constants such that $\sum_{i=1}^{n} \frac{c_i}{i} = 1$.

Solution: (i) We can write

$$f_X(x) = \sum_{i=1}^{n} \frac{c_i}{i} \cdot i x^{i-1} = \sum_{i=1}^{n} p_i f_{X_i}(x),$$

where $X_i \sim \text{Beta}(i, 1)$ or $X_i \stackrel{\text{d}}{=} U_i^{1/i}$ with $U_i \stackrel{\text{iid}}{\sim} U(0, 1)$ for $i = 1, \ldots, n$. We can see that $f_X(x)$ is a mixture of n beta distributions.

(ii) The SR method for generating $X \sim f_X(x)$ is as follows:

Step 1: Draw $U_i = u_i \stackrel{\text{iid}}{\sim} U(0, 1)$, set $x_i = u_i^{1/i}$ for $i = 1, \ldots, n$ and independently draw $\mathbf{z} = \boldsymbol{z} = (z_1, \ldots, z_n)^\top \sim \text{Multinomial}(1; p_1, \ldots, p_n)$;

Step 2: Return $x = z_1 x_1 + \cdots + z_n x_n$.