Department of Statistics and Data Science at SUSTech

MAT7035: Computational Statistics

Tutorial 1: The inversion method and the rejection method

A. The inversion method for continuous distributions

A.1 The issue and basic idea

(a) Suppose that we want to generate a sample of a continuous random variable X having the *probability density function* (pdf) $f(x)$.

(b) Find the cdf $F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$. Let $U = u \sim U(0,1)$. From $u = F(x)$ to solve $x = F^{-1}(u)$.

A.2 The algorithm

Step 1: Draw $U = u \sim U(0,1)$;

Step 2: Set $x = F^{-1}(u)$.

A.3 Remark

Some distributions (e.g., normal distribution) don't have explicit cdf.

Example T1.1 (Uniform distribution). Use the inversion method to generate a random variable from the uniform distribution $U[a,b]$ with pdf:

$$f(x) = \frac{1}{b-a} \cdot I(a \leqslant x \leqslant b), \quad \text{where} \quad b > a.$$

Solution: The cdf of $X \sim U[a,b]$ is

$$F(x) = 0 \cdot I(x < a) + \frac{x-a}{b-a} \cdot I(a \leqslant x \leqslant b) + 1 \cdot I(x > b).$$

Let $u = F(x)$, we have $x = F^{-1}(u) = a + (b-a)u$. The algorithm is as follows:

Step 1: Draw $U = u \sim U[0,1]$;

Step 2: Set $x = a + (b-a)u$.

Comments: If $X \sim U[a,b]$ and $U \sim U[0,1]$, then

$$\frac{X-a}{b-a} \stackrel{\mathrm{d}}{=} U \sim U[0,1],$$

so that $X \stackrel{\mathrm{d}}{=} a + (b-a)U$ ‖

Example T1.2 (Piecewise constant distribution). Use the inversion method to generate a random variable from the piecewise constant distribution with pdf:

$$f(x) = \begin{cases} 0, & \text{if } x \leqslant x_0 \ \text{ or } \ x \geqslant x_2, \\ c_1, & \text{if } x_0 < x < x_1, \\ c_2, & \text{if } x_1 \leqslant x < x_2, \end{cases}$$

where $c_1 > 0$, $c_2 > 0$, $x_0 < x_1 < x_2$ and $c_1(x_1 - x_0) + c_2(x_2 - x_1) = 1$.

Solution: The cdf of $X \sim f(x)$ is

$$F(x) = \begin{cases} 0, & \text{if } x \leqslant x_0, \\ \displaystyle\int_{-\infty}^{x_0} f(t)\,\mathrm{d}t + \int_{x_0}^{x} f(t)\,\mathrm{d}t, & \text{if } x_0 < x < x_1, \\ \displaystyle\int_{-\infty}^{x_0} f(t)\,\mathrm{d}t + \int_{x_0}^{x_1} f(t)\,\mathrm{d}t + \int_{x_1}^{x} f(t)\,\mathrm{d}t, & \text{if } x_1 \leqslant x < x_2, \\ 1, & \text{if } x \geqslant x_2 \end{cases}$$

$$
= \begin{cases} 0, & \text{if } x \leqslant x_0, \\ c_1(x - x_0), & \text{if } x_0 < x < x_1, \\ c_1(x_1 - x_0) + c_2(x - x_1), & \text{if } x_1 \leqslant x < x_2, \\ 1, & \text{if } x \geqslant x_2. \end{cases}
$$

Let $u = F(x)$, then

$$
x = \begin{cases} x_0 + u/c_1, & \text{if } 0 < u < c_1(x_1 - x_0), \\ x_2 - (1 - u)/c_2, & \text{if } c_1(x_1 - x_0) \leqslant u < 1. \end{cases}
$$

The algorithm is as follows:

Step 1: Generate $U = u \sim U[0, 1)$;

Step 2: If $0 < u < c_1(x_1 - x_0)$, set $x = x_0 + u/c_1$; otherwise, set $x = x_2 - (1 - u)/c_2$. ‖

Example T1.3 (Laplace distribution). Use the inversion method to generate a random variable from the Laplace distribution with pdf:

$$
f(x) = \frac{1}{2\sigma} \exp\left(-\frac{|x - \mu|}{\sigma}\right), \quad -\infty < x < +\infty,
$$

where $-\infty < \mu < +\infty$ and $\sigma > 0$. [Hint: cf. Example 1.2 in Lecture Notes]

Solution: The cdf of $X \sim f(x)$ is

$$
\begin{aligned}
F(x) &= \begin{cases} \displaystyle\int_{-\infty}^{x} \frac{1}{2\sigma} \exp\left(\frac{t - \mu}{\sigma}\right) \mathrm{d}t, & \text{if } x \leqslant \mu, \\ \displaystyle\int_{-\infty}^{\mu} f(t)\,\mathrm{d}t + \int_{\mu}^{x} \frac{1}{2\sigma} \exp\left(-\frac{t - \mu}{\sigma}\right) \mathrm{d}t, & \text{if } x > \mu. \end{cases} \\
&= \begin{cases} \dfrac{1}{2} \exp\left(\dfrac{x - \mu}{\sigma}\right), & \text{if } x \leqslant \mu, \\ 1 - \dfrac{1}{2} \exp\left(-\dfrac{x - \mu}{\sigma}\right), & \text{if } x > \mu. \end{cases}
\end{aligned}
$$

Let $u = F(x)$, then

$$x = \begin{cases} \sigma \log(2u) + \mu, & \text{if } 0 < u \leqslant 0.5, \\ -\sigma \log[2(1-u)] + \mu, & \text{if } 0.5 < u < 1. \end{cases}$$

The algorithm is as follows:

Step 1: Generate $U = u \sim U(0, 1)$;

Step 2: If $0 < u \leqslant 0.5$, set $x = \sigma \log(2u) + \mu$; otherwise, set $x = -\sigma \log[2(1-u)] + \mu$. ‖

Example T1.4 (Triangular distribution). Use the inversion method to generate a random variable from the triangular distribution with pdf: for $a < b$,

$$f(x) = \begin{cases} 0, & \text{if } x < 2a \text{ or } x \geqslant 2b, \\ \dfrac{x-2a}{(b-a)^2}, & \text{if } 2a \leqslant x < a+b, \\ \dfrac{2b-x}{(b-a)^2}, & \text{if } a+b \leqslant x < 2b. \end{cases}$$

Solution: The cdf of X is

$$F(x) = \begin{cases} 0, & \text{if } x < 2a, \\ \displaystyle\int_{-\infty}^{2a} f(t)\,\mathrm{d}t + \int_{2a}^{x} \frac{t-2a}{(b-a)^2}\,\mathrm{d}t, & \text{if } 2a \leqslant x < a+b, \\ \displaystyle\int_{-\infty}^{2a} f(t)\,\mathrm{d}t + \int_{2a}^{a+b} \frac{t-2a}{(b-a)^2}\,\mathrm{d}t + \int_{a+b}^{x} \frac{2b-t}{(b-a)^2}\,\mathrm{d}t, & \text{if } a+b \leqslant x < 2b, \\ 1, & \text{if } x \geqslant 2b \end{cases}$$

$$= \begin{cases} 0, & \text{if } x < 2a, \\ 0 + \dfrac{(x-2a)^2}{2(b-a)^2} = \dfrac{(x-2a)^2}{2(b-a)^2}, & \text{if } 2a \leqslant x < a+b, \\ 0 + \dfrac{1}{2} - \dfrac{(2b-x)^2}{2(b-a)^2} + \dfrac{1}{2} = 1 - \dfrac{(2b-x)^2}{2(b-a)^2}, & \text{if } a+b \leqslant x < 2b, \\ 1, & \text{if } x \geqslant 2b. \end{cases}$$

Let $u = F(x)$, then

$$x = \begin{cases} 2a + (b-a)\sqrt{2u}, & \text{if } 0 \leqslant u < 0.5, \\ 2b - (b-a)\sqrt{2(1-u)}, & \text{if } 0.5 \leqslant u < 1. \end{cases}$$

The algorithm is as follows:

Step 1: Generate $U = u \sim U[0, 1)$;

Step 2: If $0 \leqslant u < 0.5$, set $x = 2a + (b-a)\sqrt{2u}$;

otherwise, set $x = 2b - (b-a)\sqrt{2(1-u)}$. ‖

Example T1.5 (Beta–shifted exponential piecewise distribution). Use the inversion method to generate a sample from the beta–shifted exponential piecewise distribution with pdf

$$\begin{aligned} g_\theta(x) &= c \cdot \theta x^{\theta-1} I(0 < x < 1) + (1-c) \cdot \mathrm{e}^{-(x-1)} I(x \geqslant 1) \qquad \text{(T1.1)} \\ &= \begin{cases} 0, & \text{if } x \leqslant 0, \\ c\,\theta x^{\theta-1}, & \text{if } 0 < x < 1, \\ (1-c)\,\mathrm{e}^{-(x-1)}, & \text{if } x \geqslant 1, \end{cases} \end{aligned}$$

where $c = 1/(1+\theta)$ and $\theta > 0$.

Solution: It is easy to verify that $\int_0^\infty g_\theta(x)\,\mathrm{d}x = 1$. When $\theta \neq 1$, from (T1.1), we can see that $g_\theta(x)$ is continuous at both $x = 0$ and $x = 1$. The cdf of $X \sim g_\theta(x)$ is

$$\begin{aligned} G_\theta(x) &= \begin{cases} 0, & \text{if } x \leqslant 0, \\ \displaystyle\int_{-\infty}^0 g_\theta(t)\,\mathrm{d}t + \int_0^x c\,\theta t^{\theta-1}\,\mathrm{d}t, & \text{if } 0 < x < 1, \\ \displaystyle\int_{-\infty}^0 g_\theta(t)\,\mathrm{d}t + \int_0^1 c\,\theta t^{\theta-1}\,\mathrm{d}t + \int_1^x (1-c)\,\mathrm{e}^{-(t-1)}\,\mathrm{d}t, & \text{if } x \geqslant 1 \end{cases} \\ &= \begin{cases} 0, & \text{if } x \leqslant 0, \\ 0 + c\,x^\theta = c\,x^\theta, & \text{if } 0 < x < 1, \\ 0 + c + (1-c)\big[-\mathrm{e}^{-(x-1)} + 1\big] = 1 - (1-c)\,\mathrm{e}^{-(x-1)}, & \text{if } x \geqslant 1 \end{cases} \end{aligned}$$

Let $u = G_\theta(x)$, then

$$x = \begin{cases} (c^{-1}u)^{1/\theta}, & \text{if } 0 < u < c, \\ 1 + \log[(1-c)/(1-u)], & \text{if } c \leqslant u < 1. \end{cases}$$

The algorithm is as follows:

Step 1: Generate $U = u \sim U(0,1)$;

Step 2: If $0 < u < c$, set $x = (c^{-1}u)^{1/\theta}$; otherwise, set $x = 1 + \log[(1-c)/(1-u)]$.

Discontinuous version: In fact, we can construct another $g_\theta(x)$ such that it is discontinuous at $x = 1$. To this end, we rewrite (T1.1) as

$$g^*_\theta(x) = c\,x^{\theta-1} I(0 < x < 1) + c\,\mathrm{e}^{-x} I(x \geqslant 1), \tag{T1.2}$$

where $c = (\theta^{-1} + \mathrm{e}^{-1})^{-1}$. It is easy to verify that $g^*_\theta(x)$ is discontinuous at $x = 1$ and $\int_0^\infty g^*_\theta(x)\,\mathrm{d}x = 1$. The cdf of $X \sim g^*_\theta(x)$ is

$$\begin{aligned} G^*_\theta(x) &= \begin{cases} 0, & \text{if } x \leqslant 0, \\ \int_{-\infty}^0 g^*_\theta(t)\,\mathrm{d}t + \int_0^x c\,t^{\theta-1}\,\mathrm{d}t, & \text{if } 0 < x < 1, \\ 0 + \int_0^1 c\,t^{\theta-1}\,\mathrm{d}t + \int_1^x c\,\mathrm{e}^{-t}\,\mathrm{d}t, & \text{if } x \geqslant 1 \end{cases} \\ &= \begin{cases} 0, & \text{if } x \leqslant 0, \\ 0 + c\,x^\theta/\theta = c\,x^\theta/\theta, & \text{if } 0 < x < 1, \\ 0 + c/\theta + c(-\,\mathrm{e}^{-x} + \mathrm{e}^{-1}) = 1 - c\,\mathrm{e}^{-x}, & \text{if } x \geqslant 1. \end{cases} \end{aligned}$$

We can see that $G^*_\theta(x)$ is continuous at both $x = 0$ and $x = 1$ with $G^*_\theta(1) = c/\theta$.

Let $u = G^*_\theta(x)$, then

$$x = \begin{cases} (\theta u/c)^{1/\theta}, & \text{if } 0 < u < c/\theta, \\ \log(c) - \log(1-u), & \text{if } c/\theta \leqslant u < 1. \end{cases}$$

The algorithm is as follows:

Step 1: Generate $U = u \sim U(0,1)$;

Step 2: If $0 < u < c/\theta$, set $x = (\theta u/c)^{1/\theta}$; otherwise, set $x = \log(c) - \log(1-u)$. ‖

B. The inversion method for discrete distributions

B.1 The issue and basic idea

(a) Suppose that we want to generate a sample from a discrete r.v. X with pmf

$$f(x_i) = \Pr(X = x_i) = p_i, \quad p_i > 0, \quad i = 1, \ldots, d, \quad \sum_{i=1}^{d} p_i = 1,$$

where d is finite or $+\infty$. When d is finite, we write $X \sim \text{FDiscrete}_d(\{x_i\}, \{p_i\})$.

(b) The cdf of $X \sim \text{FDiscrete}_d(\{x_i\}, \{p_i\})$ is

$$F(x) = \Pr(X \leqslant x) = \sum_{x_i \leqslant x} f(x_i) = \begin{cases} 0, & \text{if } x < x_1, \\ p_1, & \text{if } x_1 \leqslant x < x_2, \\ p_1 + p_2, & \text{if } x_2 \leqslant x < x_3, \\ \vdots & \vdots \\ p_1 + \cdots + p_{d-1}, & \text{if } x_{d-1} \leqslant x < x_d, \\ 1, & \text{if } x \geqslant x_d. \end{cases}$$

(c) To this end, we first generate a random number $U = u \sim U(0,1)$, and then set

$$X = x = \begin{cases} x_1, & \text{if } u \leqslant p_1, \\ x_2, & \text{if } p_1 < u \leqslant p_1 + p_2, \\ \vdots & \vdots \\ x_{d-1}, & \text{if } \sum_{i=1}^{d-2} p_i < u \leqslant \sum_{i=1}^{d-1} p_i, \\ x_d, & \text{if } \sum_{i=1}^{d-1} p_i < u \leqslant 1. \end{cases}$$

B.2 The algorithm

— Step 1: Draw $U = u \sim U(0,1)$;

— Step 2: If $u \leqslant p_1$, set $X = x_1$ and stop;

If $u \leqslant p_1 + p_2$, set $X = x_2$ and stop;

$\vdots$

If $u \leqslant \sum_{j=1}^{d-1} p_i$, set $X = x_{d-1}$ and stop;

If $u \leqslant 1$, set $X = x_d$ and stop.

B.3 Remark

Let $X \sim \text{FDiscrete}_d(\{x_i\}, \{p_i\})$. The built-in R function sample(x, N, prob= p, replace = F) produces a vector of length N randomly chosen from $\boldsymbol{x} = (x_1, \dots, x_d)^\top$ with corresponding probabilities $\boldsymbol{p} = (p_1, \dots, p_d)^\top$ without replacement.

C. The rejection method

C.1 The background

- When the sampling from the pdf $f(x)$ with support $\mathcal{S}_X$ is very hard, or inefficient, we could find an envelope density $g(x)$, having 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 i.i.d. samples from $f(x)$.

C.2 The algorithm

Step 1: Draw U from $U(0,1)$ and independently draw Y from $g(\cdot)$;

Step 2: If $U \leqslant \frac{f(Y)}{cg(Y)}$, return $X = Y$; otherwise go to Step 1.

C.3 Remarks

(a) Find the constant $c(>1) = \max_{x \in \mathcal{S}_X} \frac{f(x)}{g(x)}$.

(b) The acceptance probability is $1/c$.

(c) If the potential envelope $g(\cdot)$ comes from a family $g_\theta(\cdot)$ with $\theta \in \Theta$. Then

$$c_{\text{opt}} = \min_{\theta \in \Theta} \left\{ \max_{x \in \mathcal{S}_X} \frac{f(x)}{g_\theta(x)} \right\}.$$

Example T1.6 (Beta distribution). Use the uniform pdf $g(x) = 1$ for $x \in (0,1)$ as the envelope function to generate a random variable having the beta distribution $\text{Beta}(3,1)$ with density

$$f(x) = 3x^2, \quad 0 < x < 1,$$

by the rejection method. What is the acceptance probability for this algorithm?

Solution: (i) The ratio

$$\frac{f(x)}{g(x)} = 3x^2$$

is an increasing function for $x \in (0,1)$ so the maximal value of this ratio is reached at $x = 1$. Hence

$$c = \max_{0<x<1} \frac{f(x)}{g(x)} = 3 \quad \text{and} \quad \frac{f(x)}{cg(x)} = \frac{3x^2}{3} = x^2.$$

(ii) The rejection method is as follows:

Step 1. Draw $U \sim U(0,1)$ and independently draw $Y \sim U(0,1)$;

Step 2. If $U \leqslant Y^2$, set $X = Y$; otherwise, go to Step 1.

(iii) The acceptance probability is $1/c = 1/3 \approx 0.3333$. ‖

Example T1.7 (Semi-circle distribution). Use the uniform pdf $g(x) = 1/(2r)$ for $x \in (-r, r)$ as the envelope function to generate a random variable having the semi-circle distribution with density

$$f(x) = \frac{2}{\pi r^2}\sqrt{r^2 - x^2}, \quad x \in (-r, r),$$

by the rejection method. Calculate the expected number of iterations until one acceptance and the value of the acceptance probability.

Solution: (i) By differentiating the ratio

$$\frac{f(x)}{g(x)} = \frac{4}{\pi r}\sqrt{r^2 - x^2}$$

with respect to x and setting the resultant derivative equal to zero, we obtain the maximal value of this ratio at $x = 0$. Hence

$$c = \max_{-r<x<r} \frac{f(x)}{g(x)} = \frac{4}{\pi} \quad \text{and} \quad \frac{f(x)}{cg(x)} = \frac{\sqrt{r^2 - x^2}}{r}.$$

(ii) The rejection method is as follows:

Step 1. Draw $U \sim U(0, 1)$ and independently draw $Y \sim U(-r, r)$;

Step 2. If $U \leqslant \sqrt{r^2 - Y^2}/r$, set $X = Y$; otherwise, go to Step 1.

(iii) The acceptance probability is $1/c = \pi/4 \approx 0.7854$. ‖

Example T1.8 (Half-normal distribution). Use a density, selected from the following family of exponential densities

$$g_\theta(x) = \theta\,\mathrm{e}^{-\theta x}, \quad x \geqslant 0, \quad \theta > 0,$$

as the optimal envelope function (i.e., with the largest acceptance probability) to generate a random variable having the half-normal distribution with pdf

$$f(x) = \sqrt{\frac{2}{\pi}}\,\mathrm{e}^{-x^2/2}. \quad x \geqslant 0,$$

by the rejection method. Calculate the value of the acceptance probability.

Solution: (i) The ratio is

$$\frac{f(x)}{g_\theta(x)} = \frac{\sqrt{\frac{2}{\pi}}\, \mathrm{e}^{-x^2/2}}{\theta\, \mathrm{e}^{-\theta x}} = \sqrt{\frac{2}{\pi}}\theta^{-1}\, \mathrm{e}^{-x^2/2+\theta x}.$$

Let

$$0 = \frac{\mathrm{d}}{\mathrm{d}x}\log\left[\frac{f(x)}{g_\theta(x)}\right] = \frac{\mathrm{d}}{\mathrm{d}x}\left[\text{constant} - \frac{x^2}{2} + \theta x\right] = -x+\theta,$$

we obtain that the maximal value of this ratio is arrived at $x=\theta$. Thus

$$c_\theta = \frac{f(\theta)}{g_\theta(\theta)} = \sqrt{\frac{2}{\pi}}\theta^{-1}\, \mathrm{e}^{\theta^2/2}, \quad \text{and} \quad c_{\text{opt}} = \min_{\theta>0} c_\theta = \min_{\theta>0}\left[\sqrt{\frac{2}{\pi}}\theta^{-1}\, \mathrm{e}^{\theta^2/2}\right].$$

Let

$$H(\theta) = \log\left[\theta^{-1}\, \mathrm{e}^{\theta^2/2}\right] = -\log\theta + \frac{\theta^2}{2},$$

and set

$$0 = H'(\theta) = -\frac{1}{\theta} + \theta,$$

we have $\theta = 1$. Hence

$$c_{\text{opt}} = \sqrt{\frac{2\,\mathrm{e}}{\pi}} \quad \text{and} \quad \frac{f(x)}{c_{\text{opt}}\, g_\theta(x)} = \frac{f(x)}{c_{\text{opt}}\, g_1(x)} = \mathrm{e}^{-x^2/2+x-0.5}.$$

(ii) The rejection method:

Step 1. Draw $U_1, U_2 \stackrel{\text{iid}}{\sim} U(0,1)$ and set $Y = -\log(U_2)$.

Step 2. If $U_1 \leqslant \exp(-\frac{Y^2}{2} + Y - 0.5)$, return $X = Y$; otherwise, go to Step 1.

(iii) The acceptance probability is $1/c_{\text{opt}} = 0.7602$. ‖