2013 年 11 月 重庆师范大学学报(自然科学版) Nov. 2013
第 30 卷 第 6 期 Journal of Chongqing Normal University (Natural Science) Vol. 30 No. 6

DOI:10.11721/cqnuj20130616

利用舍选抽样法生成随机数*

王丙参[1], 魏艳华[1], 孙永辉[2]

(1. 天水师范学院 数学与统计学院, 甘肃 天水 741001; 2. 河海大学 能源与电气学院, 江苏 南京 210098)

摘要:利用舍选法生成随机数的理论基础,借助几何概率描述了舍选法的直观意义,找出改进舍选法的途径,并给出优函数的选择标准,研究了优函数与接受概率的关系;特别讨论了压挤舍选抽样及自适应舍选抽样,并给出压挤函数及包络函数的选择标准,最后结合 Matlab 与 Sas 软件运用舍选法生成随机数,讨论了几种特殊密度函数随机数的生成算法,并举例给出了模拟程序。

关键词:随机数;舍选法;接受概率;压挤函数;包络函数

中图分类号:O212 **文献标志码**:A **文章编号**:1672-6693(2013)06-0086-06

用随机模拟方法解决实际问题时,首先要解决的是随机数的产生方法,然而,这项听起来简单的任务在计算机上并非很容易实现,即使能实现,因为需要调试计算机程序,所以真随机数也不可取。在调试程序过程中,经常必须对同一计算重复多次,这就要求重复产生一样的随机数序列。科学计算界广为接受的替代方法就是产生伪随机数。目前关于随机数生成的文献很多[1-5],而舍选法是非常有用的生成方法。Matlab 是计算功能最强大的软件,Sas 软件是最专业的统计软件,可以处理各种数理统计问题,进行数据分析。签于此,本文研究了舍选法生成随机数的理论基础,并给出优函数的选择标准,特别讨论了压挤舍选抽样及自适应舍选抽样,并给出压挤函数及包络函数的选择标准,最后结合 Matlab 与 Sas 软件运用舍选法生成随机数,并给出了程序。

1 利用舍选法生成非均匀随机数

对于数学性质不太好的分布可采用舍选法,它至少在理论上可从任意维数的给定概率分布抽样。舍选法不是对所产生的随机数都录用,而是建立一个检验条件,利用这一检验条件进行舍选得到所需的随机数。由于舍选法灵活、计算简单、使用方便而得到较为广泛的应用[6-7]。

定理 1 设 $f(x)$,$g(x)$为 pdf,$h(x)$为给定的函数,不一定是 pdf,如果按下法进行舍选抽样:

1)生成 $X\sim f(x)$,且 X,Y 相互独立;

2)若 $Y\leqslant h(X)$,令 $Z=X$,则 Z 的 pdf 为 $p(z)=\dfrac{f(z)G(h(z))}{\int_{-\infty}^{+\infty}f(y)G(h(y))\mathrm{d}y}$,其中 $G(y)=\int_{-\infty}^{y}g(x)\mathrm{d}x$ 。

证明 $P(Z\leqslant z)=P(X\leqslant z\mid Y\leqslant h(x))=\dfrac{P(X\leqslant z,Y\leqslant h(x))}{P(Y\leqslant h(x))}=\dfrac{\int_{-\infty}^{z}\int_{-\infty}^{h(x)}f(x)g(y)\mathrm{d}y\mathrm{d}x}{\int_{-\infty}^{+\infty}\int_{-\infty}^{h(x)}f(x)g(y)\mathrm{d}y\mathrm{d}x}=$

$\dfrac{\int_{-\infty}^{z}f(x)G(h(x))\mathrm{d}x}{\int_{-\infty}^{+\infty}f(x)G(h(x))\mathrm{d}x}$,求导可得结论成立。

若$(X,Y)\sim g(x,y)$,则此舍选法生成随机数的 pdf 形式为 $C\int_{-\infty}^{h(z)}g(z,y)\mathrm{d}y$,其中 C 为实数。

推论 1 设 Z 的 pdf $p(z)\leqslant M(z)$,$\forall z\in\mathbf{R}$,令 $C=\int_{-\infty}^{+\infty}M(x)\mathrm{d}x$,$f(x)=\dfrac{M(x)}{C}$,$h(x)=\dfrac{p(x)}{M(x)}$,如果按下法

---

* **收稿日期**:2012-06-14 **网络出版时间**:2013-11-20 14:46

资助项目:国家自然科学基金(No. 61104045);甘肃省自然科学基金计划(No. 096RJZE106)

作者简介:王丙参,男,讲师,硕士,研究方向为随机过程和金融数学,E-mail:wangbingcan2004@163.com

网络出版地址:http://www.cnki.net/kcms/detail/50.1165.N.20131120.1446.201306.86_043.html

进行舍选抽样：1)生成 $X\sim f(x)$，$U\sim U[0,1]$，且 X,U 相互独立；2)若 $U\leqslant h(X)$，令 $Z=X$，则 $Z\sim p(z)$。

产生一对随机数 (X,U) 称作一次试验，一次试验不能保证产生一个随机数 $Z\sim p(z)$，一次试验产生随机数 Z 的概率叫做舍选法的接受概率[5]，记作 p_0，即 $p(z)$ 随机数 Z 在取舍原则中被选中的概率(舍选抽样法的效率)，经取舍原则首次接受时已取舍的次数记为 N，则

$$p_0=P(U\leqslant p(X)/M(X))=\int_a^b\frac{p(x)}{M(x)}f(x)\mathrm{d}x=\frac{1}{C}$$

$N\sim Ge(p_0)$，$EN=C$，可见 C 越小，取舍的效率越高。$M(z)$ 叫做函数 $p(z)$ 的优函数，不附加任何条件的优函数容易找到，好的优函数应该可快速产生 $X\sim M(x)/C$ 和高的接受概率，但两者往往相互制约，实用的优函数是两者的合理妥协，因此舍选法的关键是找出满足下述条件的优函数：1)$M(x)$ 应从 pdf $p(x)$ 的曲线上方尽量接近 $p(x)$；2)容易生成 $X\sim M(x)/C$。常数优函数产生 $X\sim M(x)/C$ 最快、最简单，但往往接受概率太低。

若 X 的取值在 $[a,b]$ 上，且 pdf $p(x)$ 满足 $\sup\limits_{x\in[a,b]}p(x)=f_0<\infty$，取 $M(x)=\begin{cases}f_0, & a\leqslant x\leqslant b\\ 0, & \text{其他}\end{cases}$，$X\sim U[a,b]$，$U\sim U[0,1]$ 且 X,U 相互独立，若 $U\leqslant p(X)/f_0$ 时，令 $Z=X$，则 $Z\sim p(z)$。可见，当 r.v Z 的 pdf $p(x)$ 在定义域 $[a,b]$ 上有界时，可采用此法生成随机数 $Z\sim p(x)$。显然 $p_0=\dfrac{1}{f_0(b-a)}$，$EN=f_0(b-a)$，可见 f_0 越小，取舍的效率越高。

综上所述，舍选抽样法的基本思想是：按照给定的 pdf $p(x)$，对易生成的随机数列 $\{r_i\}$ 进行舍选。舍选的原则是：在 $p(x)$ 大的地方，保留较多随机数 r_i，在 $p(x)$ 小的地方，保留较少随机数 r_i，使得到的子样本中 r_i 的分布满足 pdf $p(x)$ 的要求。

已知 $g(x)$ 在 $(s,s+h)$ 上取值，且近似线性，即 $a-\dfrac{b(x-s)}{h}\leqslant g(x)\leqslant b-\dfrac{b(x-s)}{h}$，则产生 $g(x)$ 随机数算法如下[7]：

1)独立产生 $U_1,U_2\sim U(0,1)$，令 $U=\min(U_1,U_2)$，$V=\max(U_1,U_2)$；

2)如果 $V\leqslant a/b$，转 4)；

3)如果 $V\leqslant U+\dfrac{1}{b}g(s+hU)$，转 1)；

4)令 $X=s+hU$，则 $X\sim g(x)$。

推论 1′　如果 r.v $U\sim U[0,1]$，且与 pdf 为 $g(x)$ 的 Y 独立，则 $P\left(Y\leqslant x\mid\dfrac{p(Y)}{Cg(Y)}\geqslant U\right)=\int_{-\infty}^{x}p(v)\mathrm{d}v$。显然推论 1 与推论 1′ 等价。令 $g(x)=\dfrac{M(x)}{C}$，则 $P\left(Y\leqslant x\mid\dfrac{p(Y)}{Cg(Y)}\geqslant U\right)=P\left(Y\leqslant x\mid\dfrac{p(Y)}{M(Y)}\geqslant U\right)=P(Y\leqslant x\mid U\leqslant h(Y))$，$p_0=1/C$，$EN=C$。取舍原则是：一个 Y 随机数 y 被接受当且仅当对于 U 随机数 u 有 $p(y)\geqslant Cg(y)u$。

利用逆函数法生成随机数往往需要很大的计算量，为此可利用一个简捷的取舍原则[6]：

1)独立生成随机数 $v_1,\cdots,v_n\sim g(x)$ 与随机数 $u_1,\cdots,u_n\sim U(0,1)$，其中 $g(x)$ 随机数容易生成，$g(x)$ 与 $p(x)$ 取值差不多且 $\exists C$ 使得 $p(x)\leqslant Cg(x)$；

2)对 $i=1,2,\cdots$，如果有 $\dfrac{p(v_i)}{Cg(v_i)}\geqslant u_i$，就保留 v_i，否则舍弃，则由推论 1′ 可知保留下来的 v_i 就成为一系列独立的 $p(x)$ 随机数。

设 $A\subset\Omega$，Ω 上的均匀分布 $U(\Omega)$ 容易生成，在 A 上均匀分布随机向量的舍选算法如下：

1)生成 $(U,V)\in U(\Omega)$；

2)若 $(U,V)\in A$，则令 $(X,Y)=(U,V)\sim U(A)$；$P((U,V)\in A)=\dfrac{\mu(A)}{\mu(\Omega)}$ 称为接受概率，该值越大，接受效率越高。

推论 2　设 $r.v\,Z$ 的 pdf $p(z)=Lh(z)f(z)$，其中 $L=\left(\int_{-\infty}^{+\infty}f(x)h(x)\mathrm{d}x\right)^{-1}>1$，$0\leqslant h(z)\leqslant 1$，$f(z)$ 是 $r.v\,X$ 的 pdf，如果按下法进行舍选抽样：

1)生成 $X\sim f(x)$，$U\sim U[0,1]$，且 X,U 相互独立；

2)若 $U \leqslant h(X)$,令 $Z=X$,则 $Z \sim p(z)$。

推论3　设 r.v Z 的 pdf $p(z)=L\int_{-\infty}^{h(z)} g(z,y)\mathrm{d}y$,其中 $g(z,y)$ 为随机向量 (X,Y) 的联合 pdf,$h(z)$ 在 Y 的定义域上取值,L 为规格化常量,如果按下法进行舍选抽样:1)生成 $(X,Y)\sim g(x,y)$;2)若 $Y\leqslant h(X)$,令 $Z=X$,则 $Z\sim p(z)$。

证明　由于

$$P(Z\leqslant z)=\frac{\int_{-\infty}^{z}\int_{-\infty}^{h(x)} g(x,y)\mathrm{d}y\mathrm{d}x}{\int_{-\infty}^{+\infty}\int_{-\infty}^{h(x)} g(x,y)\mathrm{d}y\mathrm{d}x}=\int_{-\infty}^{z}\left[L\int_{-\infty}^{h(x)} g(x,y)\mathrm{d}y\right]\mathrm{d}x$$

所以 $p(z)=L\int_{-\infty}^{h(z)} g(z,y)\mathrm{d}y$,$L=\left(\int_{-\infty}^{+\infty}\int_{-\infty}^{h(x)} g(x,y)\mathrm{d}y\mathrm{d}x\right)^{-1}$。

注:设 $X,Y\sim U(0,1)$ 相互独立,即 $g(x,y)=f(x)\varphi(y)$,则 $p(z)=Lh(z)f(z)$。

推论2　若 $Y\sim U(0,1)$,$X\sim f(x)=M(x)/L$,其中 $M(x)$ 是 $p(x)$ 的上界函数,则 $p(x)=M(x)h(x)$,此时正是推论1。可见推论1、2是推论3的特例。

2 压挤舍选抽样

一般的舍选抽样需要对每个备选抽样 Z 有一个 $p(x)$ 值,在 $p(x)$ 求值昂贵但舍选法却吸引人的时候,压挤舍选抽样可以提高模拟速度。若非负函数 $s(x)$ 在 $p(x)$ 的支撑上处处不超过 $p(x)$,则可选 $s(x)$ 作为压挤函数,像一般舍选法,也要用到包络 $M(x)\geqslant g(x)$,由推论1知算法如下:

1)生成 $X\sim f(x)$,$U\sim U[0,1]$,且 X,U 相互独立;

2)若 $U\leqslant\dfrac{s(X)}{M(X)}$,令 $Z=X$,则 $Z\sim p(z)$,然后转到5);

3)否则,确定是否有 $U\leqslant h(X)$,如果不等式成立,令 $Z=X$,则 $Z\sim p(z)$,然后转到4);

4)如果 X 仍未保留,拒绝其成为目标随机样本之一;

5)返回1),直到达到所需的样本量。

可见,压挤舍选抽样的总接受概率仍为 $1/C$,步骤2)基于 $s(x)$ 而非 $p(x)$ 决定时候保留 X,算法2),3)对应图1。当 $s(x)$ 紧紧靠在 $p(x)$ 的下方时,且 $s(x)$ 容易计算,则可大大减少计算量,避免计算 $p(x)$ 的比例为 $\int s(x)\mathrm{d}x/\int M(x)\mathrm{d}x$。

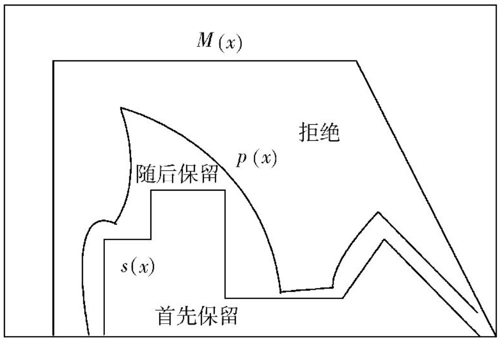

图1　压挤拒绝抽样图示

3 自适应舍选抽样

舍选抽样中的关键是构造合适的包络,Gilks 和 Wild 提出了一种针对支撑连通区域上连续、可导、对数凹密度的自动包络生成方法,也成为自适应舍选抽样。

令 $l(x)=\log p(x)$,假设在某实区间上 $p(x)>0$,可能取值无穷,$p(x)$ 是凹的,满足支撑区域内任意3点 $a<b<c$ 有 $l(a)-2l(b)+l(c)<0$。$l'(x)$ 存在且随 x 的增大单调递减,但可能有间断点。在点 $x_1<x_2<\cdots<x_k$ 处计算 $l(x)$,$l'(x)$,如果 $p(x)$ 的支撑延伸到 $-\infty$,选择 x_1 $s.t$ $l'(x_1)>0$,如果 $p(x)$ 的支撑延伸到 ∞,选择 x_k $s.t$ $l'(x_k)<0$。令 $T_k=\{x_1,\cdots,x_k\}$,T_k 上拒绝包络为 l 在 T_k 内各点处切线组成的分段线性上覆盖指数。$l(x)$ 在 x_i 的切线公式由点斜式可得 $l(x_i)+(x-x_i)l'(x_i)$,在 x_{i+1} 处得切线为 $l(x_{i+1})+(x-x_{i+1})l'(x_{i+1})$,两切线在点 $z_i=\dfrac{l(x_{i+1})-l(x_i)-x_{i+1}l'(x_{i+1})+x_il'(x_i)}{l'(x_{i+1})-l'(x_i)}$ 处相交,因此 l 的上覆盖为 $m_k^*=l(x_i)+(x-x_i)l'(x_i)$,$x\in[z_{i-1},z_i]$,且 $i=1,\cdots,k$,z_0,z_k 分别为 $p(x)$ 支撑区域的下界和上届,可能取无穷大。综上所述,拒绝包络 $M_k(x)=\exp\{m_k^*(x)\}$。

T_k 上压挤函数为 l 在 T_k 内各相邻点的弦组成的分段线性下覆盖指数。$l(x)$ 的下覆盖由 $s_k^*(x)=\dfrac{(x_{i+1}-x_i)l(x_i)+(x-x_i)l(x_{i+1})}{x_{i+1}-x_i}$,$x\in[x_i,x_{i+1}]$,$i=1,\cdots,k$ 给出。当 $x<x_1$ 或 $x>x_k$ 时,令 $s_k^*(x)=-\infty$。这

样压挤函数为 $s_k(x)=\exp\{s_k^*(x)\}$。

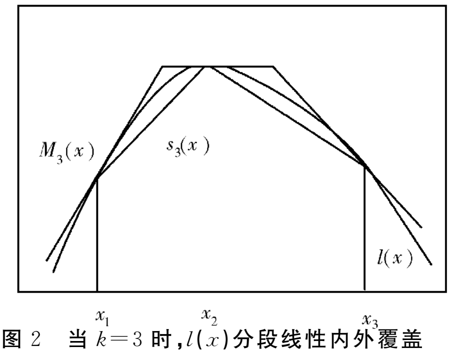

图 2　当 $k=3$ 时，$l(x)$ 分段线性内外覆盖

由图 2 可知，拒绝包络与压挤函数都是分段指数函数，包络具有在 $p(x)$ 尾部之上的指数尾部，压挤函数具有有界支撑。

自适应舍选抽样通过选择一个适合的 k 和相应的网格 T_k 来初始化。第一次迭代像压挤舍选抽样一样进行，分别用 $M_k(x)$，$s_k(x)$ 作为包络及压挤函数。当一个备选抽样被接受时，如果满足压挤条件，就不用计算即可直接接受。不过，它也可能在第二阶段被接受，这是需要计算备选抽样出的 $l(x)$，$l'(x)$，同时接受点加到 T_k 中，得到 T_{k+1}，并计算函数 $M_k(x)$，$s_k(x)$。迭代继续。如果一个新的接受点与 T_k 中的点重合，则不必更新 T_k 与 $M_k(x)$，$s_k(x)$。

若对点集 T_k，定义 $L_i(x)$ 为连接 $(x_i,l(x_i))$ 和 $(x_{i+1},l(x_{i+1}))$ 的直线函数，其中 $i=1,\cdots,k-1$，则包络函数 $m_k^*(x)=\begin{cases}\min\{L_{i-1}(x),L_{i+1}(x)\},x\in[x_i,x_{i+1}]\\ L_1(x),x<x_1\\ L_{k-1}(x),x>x_k\end{cases}$，并约定 $L_0(x)=L_k(x)=\infty$，则 $m_k^*(x)$ 是 $l(x)$ 的上覆盖，$M_k(x)=\exp\{m_k^*(x)\}$ 为 $p(x)$ 的包络函数。这样生成包络函数可以避免计算 $l'(x)$。笔者希望在 $p(x)$ 取最大值的附近网格点最密集，幸运的是，这将自动发生。因为这样的点在迭代中最可能保留，从而被更新到 T_k 中。

无导数自适应舍选法的迭代和前面方法一样类似进行，每当新点保留时，更新 T_k、包络和压挤函数。图 3 演示了无导数自适应舍选算法，显然包络不如使用 $l'(x)$ 有效。

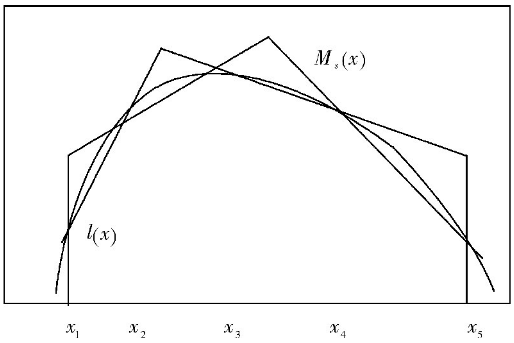

图 3　当 $k=5$ 且无导数时，$l(x)$ 分段线性外覆盖

总之，不加约束的包络函数 $M(x)$ 很容易选择，比如常数包络函数，但接受效率太低，因此它选择标准是：1）容易生成 $M(x)/C$ 随机数，且易计算 $M(x)$，从而提高运算速度；2）$M(x)$ 应从上方尽可能接近 $p(x)$，从而提高接受效率。有时候 pdf $p(x)$ 计算复杂，为较少它的计算次数，可引进压挤函数 $s(x)$ 来提高计算速度，它的选择原则是 $s(x)$ 应从下方尽可能接近 $p(x)$，且容易计算。

4 举例说明及软件实现

例 4.1　生成 $Beta(a,b)$ 的随机数。

解：由于 pdf $p(x)=\frac{1}{B(a,b)}x^{a-1}(1-x)^{b-1}$，$0<x<1$，求导可得在 $x=\frac{a-1}{a+b-2}$ 处取得最大值 $D=\frac{1}{B(a,b)}\cdot\left(\frac{a-1}{a+b-2}\right)^{a-1}\left(1-\frac{a-1}{a+b-2}\right)^{b-1}$。故舍选算法如下：

1）生成 $X\sim U[0,1]$，$U\sim U[0,1]$，且 X,U 相互独立；

2）若 $U\leqslant p(X)/D$，令 $Z=X$，则 $Z\sim p(z)$。若 $(a-1)/(a+b-2)<0$，最大点在 0 或 1 达到，相应改变 D，上面方法同样可行；当无最大点时，如 $a=1,b<1$ 或者 $a<1,b=1$，则按幂分布产生随机数。

Matlab 擅长处理数值运算，而 Sas 擅长统计分析，因此可采用二者所长，更好处理实际问题。若 $a=4,b=5$，编程如下：

```
format long; a=4,b=5;
d=1/beta(a,b)*((a-1)/(a+b-2))^(a-1)*((b-1)/(a+b-2))^(b-1)
data ex1; seed=678;a=4;b=5;d=2.350041224319791;
do I=1 to 100;
do until (u<=1/d*x**(a-1)*(1-x)**(b-1));
u=ranuni(seed);x=ranuni(seed); end;
z=x; output;  end;  run;
```

例 4.2　生成 $\Gamma(\alpha,1)$ 随机数。

解:由于 $pdf\ p(x)=\frac{1}{\Gamma(\alpha)}x^{\alpha-1}e^{-x}$,$\alpha>0$,$x\geqslant 0$,若 $0<\alpha<1$,令 $f(x)=\frac{1}{\alpha}e^{-x/\alpha}$,$x>0\sim\exp(1/\alpha)$,$L=\frac{e^{\alpha-1}\alpha^{\alpha}}{\Gamma(\alpha)}>1$,$0\leqslant h(x)=\left(\frac{x/\alpha}{\exp(x/\alpha+1)}\right)^{\alpha-1}\leqslant 1$,则舍选抽样步骤如下:

1)产生独立的 $U,V\sim U(0,1)$,令 $Y=-\frac{\ln V}{\alpha}$;

2)若 $U\leqslant\left(\frac{Y/\alpha}{e^{Y/\alpha+1}}\right)^{\alpha-1}$,令 $X=Y$,则 $X\sim\Gamma(\alpha,1)$;等价:1)产生独立的 $U,V\sim U(0,1)$,令 $Y=-\ln V$;

3)若 $U\leqslant\left(\frac{Y}{e^{Y+1}}\right)^{\alpha-1}$,令 $X=\alpha Y$,则 $X\sim\Gamma(\alpha,1)$;

Ahrens-Dieter(1974)给出的舍选法为:

1)独立产生 $U,V\sim U(0,1)$,$c=1+\alpha/e$,令 $Y=cV$;

2)当 $Y\leqslant 1$ 时,令 $X=Y^{\alpha-1}$,若 $U>e^{-X}$,转(1);

3)令 $X=\ln[\alpha/(c-Y)]$,若 $U>X^{\alpha-1}$,转 1),则 $X\sim\Gamma(\alpha,1)$;

Wallace(1974)提出:记 $m=[a]$,$p=a-m\in(0,1)$,令 $h(x)=p\frac{x^{m-1}e^{-x}}{(m-1)!}+(1-p)\frac{x^m e^{-x}}{m!}$,$x>0$,$L=\frac{(m-1)!\ m^p}{\Gamma(a)}$,$f(x)=\left(\frac{x}{m}\right)^p\left[1+(1-p)\left(\frac{x}{m}-1\right)\right]^{-1}$,由此舍选法为:

1)独立产生 $U,R\sim U(0,1)$;

2)若 $U\leqslant p$,独立抽 m 个变量 $U_1,\cdots,U_m\sim U(0,1)$,计算 $V=-\ln(\prod_{i=1}^{m}U_i)$,转 4);

3)若 $U>p$,独立抽 $m+1$ 个变量 $U_1,\cdots,U_{m+1}\sim U(0,1)$,计算 $V=-\ln(\prod_{i=1}^{m+1}U_i)$;

4)若 $R\leqslant\left(\frac{V}{m}\right)^p\left[1+\left(\frac{V}{m}-1\right)(1-p)\right]^{-1}$,则 $X=V$,停止;否则转 1);

例 4.3 生成 $Z=|\zeta|$,其中 $\zeta\sim N(0,1)$。

解:因为 $p(z)=\sqrt{\frac{2}{\pi}}e^{-\frac{z^2}{2}}=\sqrt{\frac{2}{\pi}}e^{-\frac{(z-1)^2}{2}}e^{-z}$,$z\in[0,\infty)$,令 $f(z)=e^{-z}$,$z>0\sim\exp(1)$,$h(z)=e^{-\frac{(z-1)^2}{2}}\in[0,1]$,$L=\sqrt{\frac{2e}{\pi}}\approx 1.32>1$,检验条件 $U\leqslant h(X)\Leftrightarrow -\ln U\geqslant\frac{(X-1)^2}{2}\Leftrightarrow\frac{(X-1)^2}{2}\leqslant Y\triangleq -\ln U$,则舍选抽样步骤如下:

1)产生 $U_1\sim U(0,1)$,$U\sim U(0,1)$且相互独立;

2)令 $X=-\ln U_1$,$Y=-\ln U$,若 $\frac{(X-1)^2}{2}\leqslant Y$,令 $Z=X$,则 $Z\sim|\zeta|$。显然舍选效率 $p_0=\frac{1}{L}=0.760$。

法 2:令 $M(y)=e^{-y}$,$y\in[0,\infty)\sim\exp(1)$,由 $\frac{p(y)}{M(y)}=\sqrt{\frac{2e}{\pi}}\exp\left\{-\frac{(y-1)^2}{2}\right\}\leqslant\sqrt{\frac{2e}{\pi}}$,$y\in[0,\infty)$,故取 $C=\sqrt{\frac{2e}{\pi}}\approx 1.32$,则产生随机数 Z 的步骤如下:

1)独立产生随机数 $\eta\sim Exp(1)$的和 $\xi\sim U(0,1)$;

2)若 $\xi\leqslant\exp\{-(\eta-1)^2/2\}$,或 $-\ln\xi\leqslant\frac{(\eta-1)^2}{2}$,则令 $\theta=\eta$,并转到下一步(3),否则返回 1);

3)产生 $\xi_1\sim U(0,1)$,令 $Z=\begin{cases}\theta, & 若\ \xi_1\leqslant\frac{1}{2}\\ -\theta, & 若\ \xi_1>\frac{1}{2}\end{cases}$,则 $Z\sim|\zeta|$。

评价生成算法的原则有准、快、少,舍选法从理论上保证生成的随机数严格地具有所要求的密度,生成一个非均匀随机数所需均匀随机数平均个数取决于接受概率,这与优函数的选择相关,因此优函数尽量满足:快速产生 $X\sim M(x)/C$ 和高的接受概率。

参考文献:

[1] 魏艳华,王丙参,宋立新.均匀分布的优良特性及其应用[J].四川理工学院学报:自然科学版,2010,23(4):385-387.

Wei Y H,Wang B C,Song L X. Properties of uniform distribution and its applications[J]. Journal of Sichuan University of Science& Engineering: Natural Science Edition, 2010,23(4):385-387.

[2]王丙参,魏艳华,张云.利用反函数及变换抽样法生成随机数[J].重庆文理学院学报:自然科学版,2011,30(5):9-12.
Wang B C, Wei Y H, Zhang Y. Generate random number by using inverse function and transform sampling method [J]. Journal of Chongqing University of Arts and Sciences : Natural Science Edition ,2011,30(5):9-12.

[3]杨振海.张国志.随机数生成[J].数理统计与管理,2006,25(2):244-252.
Yang Z H, Zhang G Z. Generating random variables[J]. Application of Statistics and Management, 2006, 25 (2): 244 -252.

[4]杨振海,程维虎.非均匀随机数产生[J].数理统计与管理,2006,25(6):750-756.
Yang Z H, Cheng W H. The common method of generating random number for the nonuniform distribution[J]. Application of Statistics and Management, 2006, 25(6):750-756.

[5]程维虎,杨振海.舍选法几何解释及曲边梯形概率密度随机数生产算法[J].数理统计与管理,2006,25(4):494-504.
Cheng W H, Yang Z H. The geometric sense of acceptance rejection method and generating random number with uniform distribution on the trapezoid with curve side[J]. Application of Statistics and Management, 2006, 25(4): 494-504.

[6]龚光鲁.随机微分方程及其应用概要[M].北京:清华大学出版社,2008,2:6-10.
Gong G L. Summary of stochastic differential equations and its applications[M]. Beijing: Tsinghua University Press, 2008, 2: 6-10.

[7]高惠旋.统计计算[M].北京:北京大学出版社,1995,7:80-160.
Gao H X. Statistical calculations[M]. Beijing: Beijing University Press, 1995, 7: 80-160.

[8]Givens G H, Hoeting J A.计算统计[M].王兆军,刘民千,邹长亮,等译.北京:人民邮电出版社,2009,9:118-150.
Givens G H, Hoeting J A. Computational statistics[M]. Wang Z J, Liu M Q, Zou C L, et al. Beijing: Posts & Telecom Press, 2009, 9: 118-150.

Generate Random Number by Using Acceptance Rejection Method

WANG Bing-can[1], WEI Yan-hua[1], SUN Yong-hui[2]

(1. School of Mathematics and Statistics, Tianshui Normal University, Tianshui Gansu, 741001;
2. College of Energy and Electrical Engineering, Hehai University, Nanjing 210098, China)

Abstract: It discussed the theory of acceptance rejection method, described the geometric scene of acceptance and reject method by using geometric probability, and then, found out the improvement about acceptance rejection method, given evaluation criteria of superior function, discussed the relationship between superior function and the probability of acceptance. It studied squeeze and adaptive sampling method, given evaluation criteria of squeeze function and envelope function, used acceptance rejection method to generate non-uniform random numbers by Matlab and Sas sofeware , discussed algorithm of generation random number of several special density function , given the program by an example.

Key words: random number; acceptance rejection method; acceptance probability; squeeze function; envelope function

(责任编辑　欧红叶)