第三章 随机变量的生成

- **逆变法**

指数分布　　Exp(mean)---到达过程，服务过程
均匀分布　　U(min, max)---服务过程
三角分布　　Tria(min,mode,max)
威布尔分布　Weib(α,β)---故障时间

- **卷积法**

爱尔兰分布　Er(m, λ)---电话呼叫时间
泊松分布　Po(λ,k)---到达过程

- **函数变换法**　正态分布
- **近似法**
- **取舍法**

1. 逆变法(Inverse Transform Method)

如果$U \sim U(0, 1)$，而$F^{-1}(x)$是分布函数$F(x)$的反函数，则

$$X = F^{-1}(U) \sim F(x)。$$

故可由$U(0, 1)$随机数$\{u_i\}$直接生成规定分布$F(x)$的随机数$\{x_i\}$。该方法称为逆变法或反函数法，其算法为：

① 产生独立的$U(0, 1)$随机数$u_1, u_2, \ldots, u_n$；

② 令$x_i = F^{-1}(u_i)$ ($i = 1, 2, \ldots, n$)，则$x_1, x_2, \ldots x_n$就是给定分布$F(x)$的随机数序列。

(1) 均匀分布

$[a, b]$区间上均匀分布$U(a, b)$的概率密度函数为

$$f(x) = \begin{cases} 1/(b-a), & a \le x \le b \\ 0, & \text{其它} \end{cases}$$

分布函数为$F(x) = (x-a)/(b-a)$ $(a \le x \le b)$，由其反函数$F^{-1}(\cdot)$得抽样公式

$$X = F^{-1}(U) = (b-a)U + a。$$

于是可利用逆变法生成$U(a, b)$的随机数：

① 生成独立的$U(0, 1)$随机数$u_1, u_2, \ldots, u_n$；

② $x_i = (b-a)u_i + a$ ($i = 1, 2, \ldots, n$)即为$U(a, b)$随机数。

(2) (负)指数分布

指数分布E (λ)的概率密度函数为

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

其中$\lambda > 0$，分布函数为$F(x) = 1 - e^{-\lambda x}, x > 0$。由此得

$$X = F^{-1}(U) = -(1/\lambda)\ln(1-U)。$$

抽样公式常取为

$$X = -(1/\lambda)\ln(U)。$$

于是可利用逆变法生成指数分布的随机数：

① 生成独立的$U(0, 1)$随机数$u_1, u_2, \ldots, u_n$；

② $x_i = -(1/\lambda)\ln(u_i)$ ($i = 1, 2, \ldots, n$)即为指数分布的随机数。

例3.7 设随机变量X的概率密度函数为

$$f(x) = \begin{cases} 3x^2, & 0 < x \le 1 \\ 0, & \text{其它} \end{cases}$$

若有$U(0, 1)$随机数:

0. 38 0. 10 0. 60 0. 90 0. 88 0. 96 0. 01 0. 41 0. 86 0. 14

请由此生成随机变量X的随机数。

解：X分布函数为$F(x) = x^3$ $(0 \le x \le 1)$，由其反函数$F^{-1}(\cdot)$得抽样公式

$$X = F^{-1}(U) = U^{1/3}$$

于是可利用逆变法生成随机变量X随机数：

$$\sqrt[3]{0.38}\quad \sqrt[3]{0.10}\quad \sqrt[3]{0.60}\quad \sqrt[3]{0.90}\quad \sqrt[3]{0.88}$$
$$\sqrt[3]{0.96}\quad \sqrt[3]{0.01}\quad \sqrt[3]{0.41}\quad \sqrt[3]{0.86}\quad \sqrt[3]{0.14}$$

(3) 离散分布

设随机变量X的概率分布为

$$P\{X = x_i\} = p_i \quad (i = 1, 2, \ldots),$$

不妨设$x_1 < x_2 < x_3 < \ldots$，X的分布函数为

$$F(x) = P\{X \le x\} = \sum_{x_i \le x} p_i$$

算法：

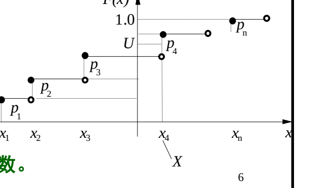
① 生成U(0, 1)随机数u；

② 若$u \le p_1$，令$x = x_1$；否

则当

$$\sum_{k=1}^{i-1} p_k < u \le \sum_{k=1}^{i} p_k \quad (i = 2, 3, \cdots)$$

时，令$x = x_i$，则x为$F(x)$随机数。



例3.8 设随机变量X的概率分布为

X_i	1	2	3	4	5	6
P_i	0.1	0.2	0.2	0.3	0.15	0.05

请利用例3.7中的$U(0, 1)$随机数，生成X的随机数。

解：注意到 $F(x)=P\{X\le x\}=\sum_{x_i\le x} p_i$

若 $u\le p_1$，令 $x=x_1$；否则当 $\sum_{k=1}^{i-1} p_k < u \le \sum_{k=1}^{i} p_k (i=2,3,\cdots)$

时，令$x=x_i$。故有

u_i	0.38	0.10	0.60	0.90	0.88	0.96	0.01	0.41	0.86	0.14
x_i	3	1	4	5	5	6	1	3	5	2

2. 卷积法(Convolution Method)

若随机变量Y可以表示为独立随机变量$\{X_i\}$ $(i=1, 2, \ldots, n)$的和，则Y的概率密度是$\{X_i\}$的概率密度的卷积。因此可以利用$\{X_i\}$生成随机变量Y，抽样公式为

$$Y=X_1+X_2+\ldots+X_n$$

爱尔朗分布

n 阶爱尔朗(Erlang)分布$E_n(\lambda)$的概率密度为

$$f(x)=\frac{\lambda(\lambda x)^{n-1}}{(n-1)!}e^{-\lambda x}(x>0;\lambda>0)$$

若X_1，X_2，…，X_n独立且同服从指数分布$E(\lambda)$，令$Y=X_1+X_2+\ldots+X_n$，则Y 服从 n 阶爱尔朗分布。

利用该性质生成 n 阶爱尔朗分布随机数的抽样方法为

① 生成独立的U(0, 1)随机数 $u_1, u_2, \ldots, u_n$；

② 计算 $u=u_1\cdot u_2\cdot \cdots \cdot u_n$；

③ 令$y=-(1/\lambda)\ln(u)$，则 y 即为 n 阶爱尔朗分布的随机数。

3. 函数变换法与近似法

函数变换法是关于随机变量的函数(仍为随机变量)的抽样法。

设常用随机变量X的分布函数为$F(x)$，X的函数$Y=g(X)$也是随机变量，其分布函数为

$$\begin{aligned}G(y)&=P\{Y\le y\}\\&=P\{g(X)\le y\}\\&=P\{X\le g^{-1}(y)\}\\&=F[g^{-1}(y)]\end{aligned}$$

再利用逆变法可得函数变换法的抽样公式为

$$Y=g(X)$$

于是由$F(x)$的随机数生成$G(x)$的随机数的步骤为：

① 生成独立的$F(x)$随机数 $x_1, x_2, \ldots, x_n$；

② 令 $y_i=g(x_i)(i=1, 2,\ldots,n)$，则 $y_1, y_2, \ldots, y_n$ 就是$G(y)$的随机数序列。

逆变法是一种特殊的函数变换法。

(1) 用函数变换法生成正态分布随机数
(Box－Muller方法)

正态分布$N(\mu, \sigma^2)$ 的概率密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\{-\frac{1}{2\sigma^2}(x-\mu)^2\}(-\infty<x<+\infty;-\infty<\mu<+\infty,\sigma>0)$$

当$X\sim N(0, 1)$时，$Y=\sigma X+\mu\sim N(\mu,\sigma^2)$，故只需考虑生成$N(0, 1)$随机数的方法。

设 u_1, u_2 是相互独立的$U(0, 1)$分布随机数，令

$$\begin{cases}z_1=\sqrt{-2\ln u_1}\cos 2\pi u_2\\ z_2=\sqrt{-2\ln u_1}\sin 2\pi u_2\end{cases}$$

则 z_1, z_2 为相互独立$N(0, 1)$的随机数。

Box-Muller

设(X,Y)是一对相互独立的服从正态分布N(0,1)的随机变量，则有概率密度函数$f_{(X,Y)}=\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}$,令$x=R\cos\theta, y=R\sin\theta$，其中$\theta\in[0,2\pi]$，则

$$P(R<r)=\int_0^{2\pi}\int_0^r\frac{1}{2\pi}e^{-\frac{u^2}{2}}udud\theta=1-e^{-\frac{r^2}{2}}$$

$$P(\theta<\phi)=\int_0^{\phi}\int_0^{\infty}\frac{1}{2\pi}e^{-\frac{u^2}{2}}udud\theta=\frac{\phi}{2\pi}$$

令$F_R(R<r)=1-e^{-\frac{r^2}{2}}$，则分布函数的反函数为

$R=F_R^{-1}(Z)=\sqrt{-2ln(1-Z)}$

选取两个服从[0, 1]上均匀分布的随机变量U_1, U_2，则R可由$\sqrt{-2lnU_1}$模拟生成，令θ为$2\pi U_2$，代入

$$x=R\cos\theta, y=R\sin\theta$$

可得

$$x=\sqrt{-2lnU_1}\cos 2\pi U_2$$
$$y=\sqrt{-2lnU_1}\sin 2\pi U_2$$

例3.9 请利用Box—Muller方法及例3.7中前2个$U(0, 1)$随机数，生成两个$N(0, 1)$的随机数。

解：注意到

$$\begin{cases} z_1=\sqrt{-2\ln u_1}\cos 2\pi u_2 \\ z_2=\sqrt{-2\ln u_1}\sin 2\pi u_2 \end{cases}$$

有

u_i	0.38	0.10
x_i	$\sqrt{-2\ln 0.38}\cos(2\pi 0.1)$	$\sqrt{-2\ln 0.38}\sin(2\pi 0.1)$

(2) 用近似法生成正态分布随机数

由中心极限定理知，若U_1，U_2，…，U_n 独立同服从$U(0, 1)$分布，则

$$\bar{U}=\frac{1}{n}\sum_{i=1}^{n}U_i\sim N(\frac{1}{2},\frac{1}{12n})$$

从而可得$N(0, 1)$随机数的近似抽样公式为

$$X=\frac{\bar{U}-0.5}{\sqrt{1/12n}}=\sqrt{12n}(\bar{U}-0.5)$$

过去常取 $n=6$或 $n=12$。

例3.10 请利用近似法及例3.7中前6个$U(0, 1)$随机数，生成1个$N(0, 1)$的随机数。

解：注意到近似抽样公式为

$$X=\frac{\bar{U}-0.5}{\sqrt{1/12n}}=\sqrt{12n}(\bar{U}-0.5)$$

计算可得：

$$\bar{U}=\frac{1}{6}\sum_{i=1}^{n}U_i=0.64$$

由近似抽样公式生成1个N(0, 1)随机数：

$$X=\frac{\bar{U}-0.5}{\sqrt{1/12n}}=\sqrt{72}(0.64-0.5)=0.84\sqrt{2}$$

4. 取舍法(Acceptance /Rejection Technique)

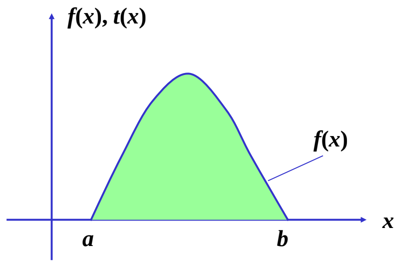
设$f(x)$为所求随机数的概率密度函数，取舍法要求选定一个覆盖函数$t(x)$，满足

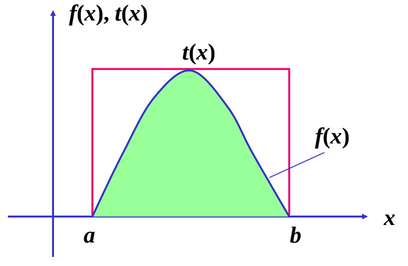
$$f(x)\le t(x),\quad C=\int_{-\infty}^{+\infty}t(x)dx<+\infty$$

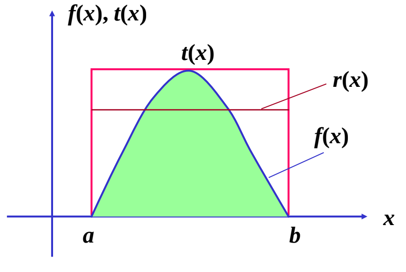
令$r(x)=t(x)/C$，则

$$\int_{-\infty}^{+\infty}r(x)dx=\int_{-\infty}^{+\infty}t(x)/C\,dx=1$$

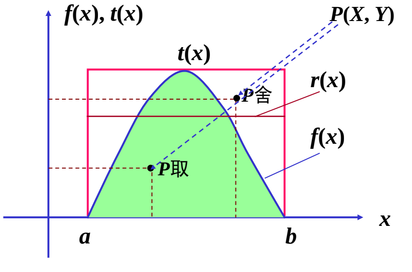
故是$r(x)$ 一个概率密度函数。



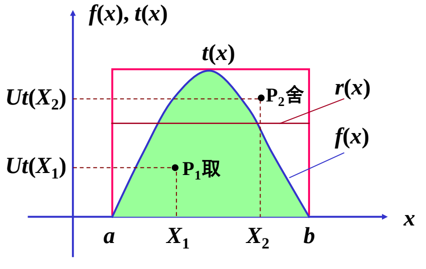




取舍法的思路是从以$t(x)$为顶的曲边梯形中随机地抽取一点P，若P点落在以$f(x)$为顶的曲边梯形中，则选取该点(该点的横坐标即为所求)；否则舍弃该点。



取舍法的思路是从以$t(x)$为顶的曲边梯形中随机地抽取一点$P(X, U\,t(X))$，若P点落在以$f(x)$为顶的曲边梯形中，则选取该点(该点的横坐标即为所求)；否则舍弃该点。



取舍法生成$F(x)$的随机数的算法为：

① 生成$r(x)$的随机数 x；

② 生成$U(0, 1)$随机数 u，且 x 与 u 独立；

③ 若 $u \le f(x) / t(x)$，令$y=x$；否则转到①重新抽样，则 y 就是$F(x)$的随机数。

取舍法不被舍弃的点的概率为

$$p = \frac{\int_a^b f(x)dx}{\int_a^b t(x)dx} = \frac{1}{C}$$

贝塔分布

贝塔(Beta)分布$\beta(a, b)$的概率密度函数为

$$f(x)=\frac{1}{B(a,b)}x^{a-1}(1-x)^{b-1}(0<x<1;a>0,b<0)$$

$f(x)$的最大值为

$$M=\frac{1}{B(a,b)}\left(\frac{a-1}{a+b-2}\right)^{a-1}\left(\frac{b-1}{a+b-2}\right)^{b-1}$$

取 t(x)＝M，显然r(x)为U(0, 1)分布的密度函数。取舍法抽样过程如下：

① 生成独立的U(0, 1)随机数 u_1, u_2；

② 如果 $u_2 \le f(u_1)/M$，则 $x=u_1$ 是分布为β(a, b)的随机数；否则转①重新抽样。