蒙特卡洛算法：抽样方法

计算物理b
高阳

核心问题

- 给定在 [0,1]上的均匀分布随机数

- 目标：满足某种所需分布的随机数列。

- 离散情形：$\xi \in \{x_i, i = 1,2, \dots, n\}, p_i = p(\xi = x_i)$ 且$\sum_i p_i = 1$

- 连续情形：$p(x)$。

离散情形：例子

- 周二早上九点半选择做什么事情，有若干种可能。1. 来上课，概率为$p_1 = 0.35$; 2. 去锻炼，概率为$p_2 = 0.1$; 3. 洗衣服，概率$p_3 = 0.05$；4. 睡觉，概率$p_4 = 0.25$; 5. 玩游戏，概率为$p_5 = 0.25$.

- 如何用概率算法决定应该做什么？

- 首先需产生满足此分布的随机数。

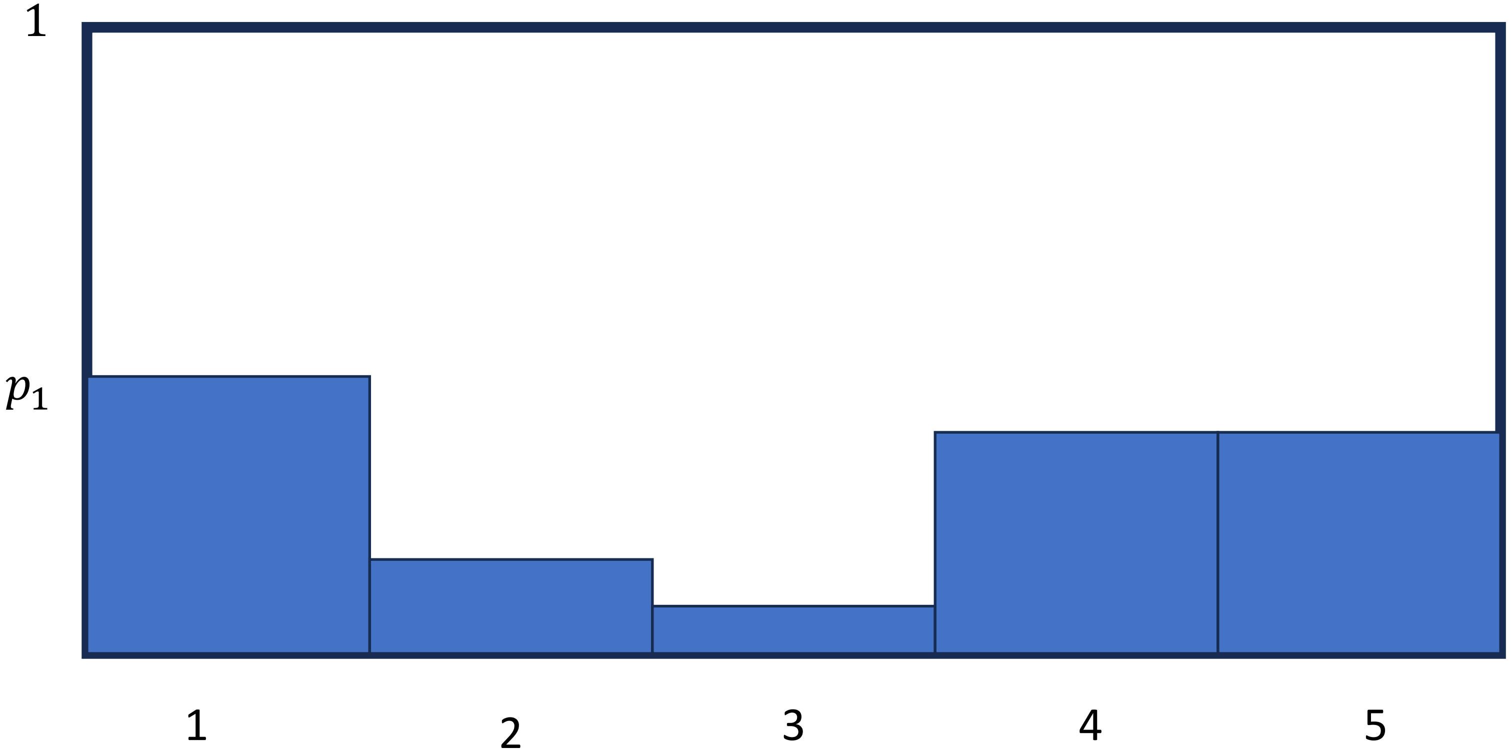
- 直接的方法：舍选法。

离散情形：舍选法

- 1. 随机生成从1到5的整数k（方法后面给出）。
- 2.产生(0,1)内的随机数ξ
- 3. 如果$\xi < p_k$,则输出k，也即做第k件事。
- 4. 否则回到第1步往下执行，直到生成某件事的标号为止。

离散情形：抽样效率

- 抽样效率：产生有效的随机变量的概率



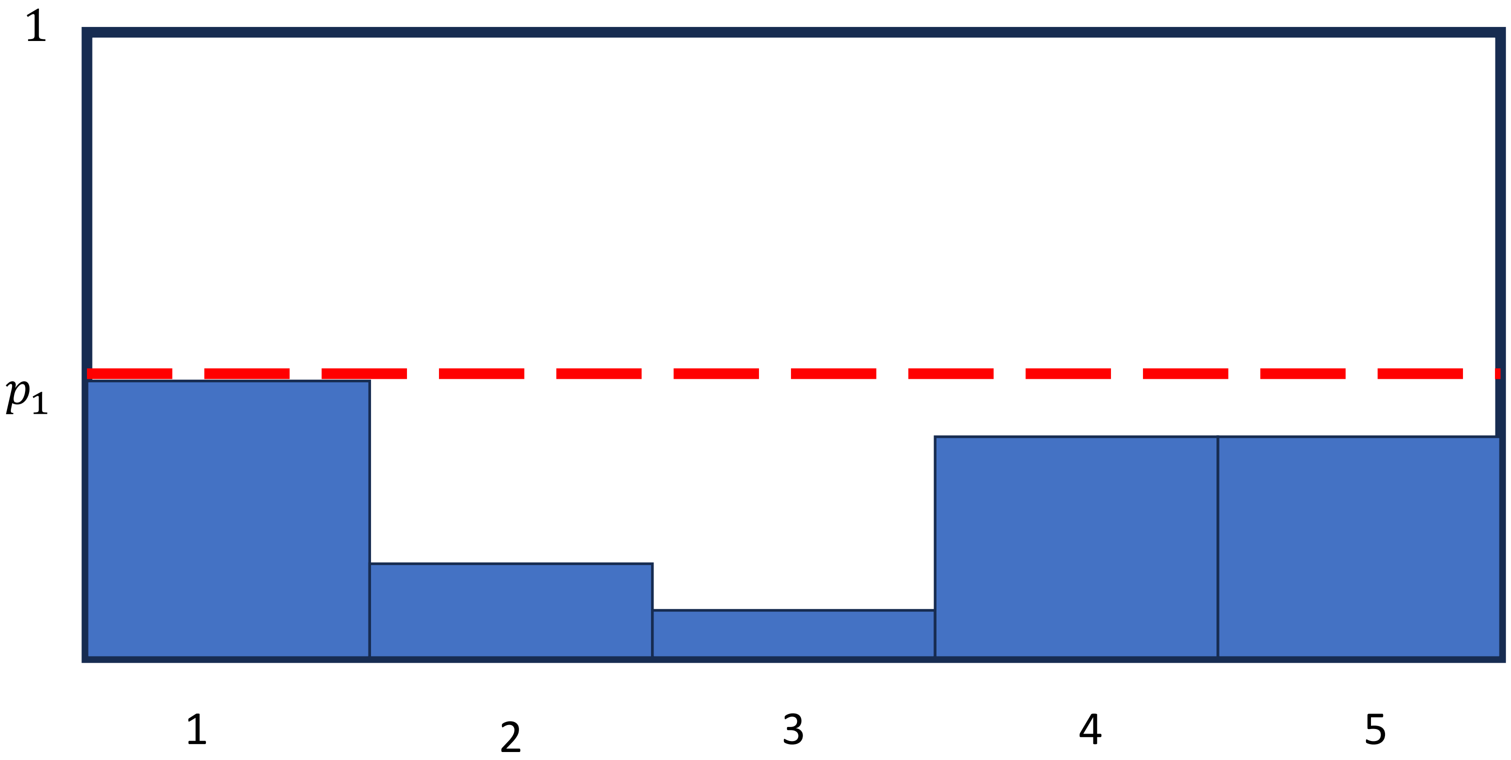
抽中的概率为：阴影面积除以总面积，
故抽样效率为 $\sum_i \frac{p_i}{5} = \langle p_i \rangle$

离散情形：优化算法

- 1. 找到概率最高的事，记其概率为p_{max}
- 2. 随机生成从1到5的整数k（方法后面给出）。
- 2.产生$(0, p_{max})$内的随机数ξ
- 3. 如果$\xi < p_k$,则输出k，也即做第k件事。
- 4. 否则回到第1步往下执行，直到生成某件事的标号为止。

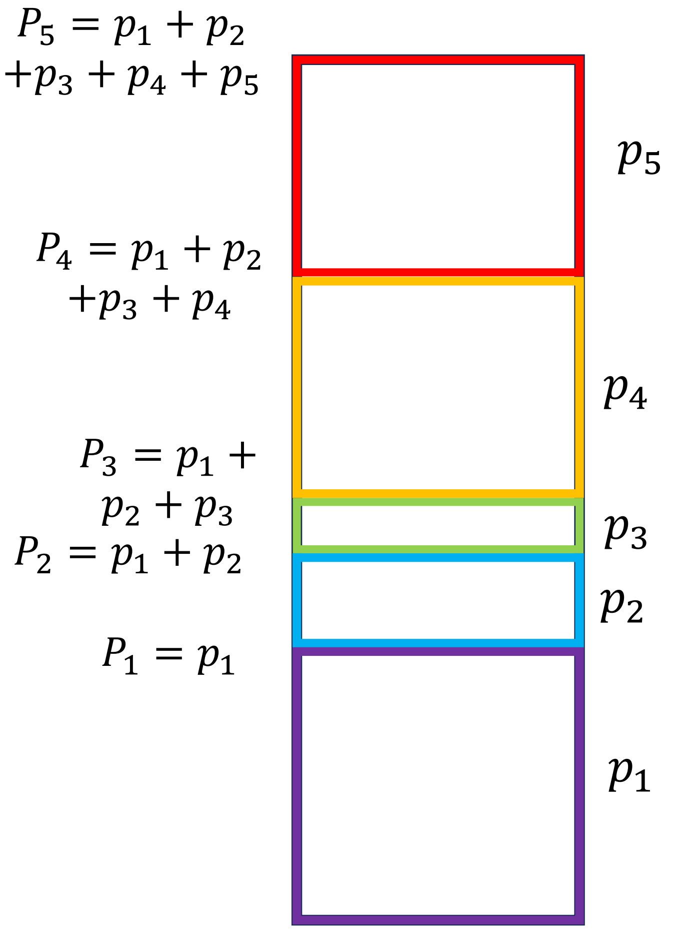
离散情形：抽样效率提高

- 抽样效率：产生有效的随机变量的概率



抽中的概率为：阴影面积除以总面积，
故抽样效率为
$\sum_i \frac{p_i}{5p_{max}} = \frac{\langle p_i \rangle}{p_{max}} > \langle p_i \rangle$

离散情形：塔式抽样



- 产生(0,1)之间的随机数ξ。
- 设置分布概率$P = 0$;
- 设置循环变量i, i从1到N，其中N为总事件数目;
- 计算$P = P + p_i$;
- 若$\xi < P$,则返回i, 否则重复循环。

注：实际上，可以事先求好节点分布概率P, 从而算法中只需搜索ξ在哪个区间即可。可用二分法加速搜索。事件顺序不重要，只需有编号（可数）即可。
这种塔式抽样在连续情形下即为反函数法。

连续情形：反函数法

- 变量及概率

$$(k, p_k) \to \big(x, p(x)\big)$$

- 分布概率

$$P_k = P_{k-1} + p_k \quad \to \quad P(x) = \int_{-\infty}^{x} dy\, p(y)$$

- 判断条件（可将连续函数无限分割以离散化）

$$P_{k-1} < \xi < P_k \quad \to \quad P(x) < \xi < P(x + dx)$$

或者，当dx足够小时，$P(x) = \xi$

- 也即$x = P^{-1}(\xi)$

例子1

- 产生满足如下概率密度分布的随机变量

  $$f(x) = \lambda\, e^{-\lambda\, x}, x > 0\ , \lambda > 0;$$
  $$f(x) = 0, otherwise.$$

- 首先求分布函数 $F(x) = \int_{-\infty}^{x} dt\, \lambda e^{-\lambda\, t} = 1 - e^{-\lambda x}, x > 0$
- 反解：$F^{-1}(y) = -\frac{1}{\lambda}\ln(1-y)$
- 方法：若y为（0,1）的随机数，则所需的随机变量为

  $$x = -\frac{1}{\lambda}\ln(1-y)$$

- 注1：由于1-y也在（0,1）区间，所以上式可改为

  $$x = -\frac{1}{\lambda}\ln(y)$$

- 注2：对于y，应加语句排除左右端点。

反函数法：样本变换

- 反函数法的实质是积分与采样的统一性
- 上述函数变化可认为是两种不同的样本之间的变化：
  一种是所需样本，一种是（0,1）之间均匀分布的样本。
- 一般原则

$$\int_0^1 d\xi \xrightarrow{\text{积分变换}} const \int_a^b dx\, p(x)$$

$$(\xi\!:(0,1)\text{均匀分布} \xrightarrow{\text{样本变换}} (x\!:p(x))$$

示例

- 重看例子1

$$d\xi = const\, dx\, e^{-\lambda x}$$

- 从而有

$$\xi = const\, e^{-\lambda x} + const'$$

- 看边界 $0 = const + const'$

  $1 = const'$ 注：也可换边界$1 = const + const'$， $0 = const'$

- 从而有 $const = -const' = -1$

  故 $x = -\frac{1}{\lambda}\ln(1 - y)$

反函数法推广：变换抽样

- 为保证收敛性，直接反函数抽样不一定是最合适的；若分布密度比较复杂，直接抽样也会比较困难
- 变换抽样的思路：

$$p(x)dx = \phi(y)dy = \phi(y)\, h'(x)dx$$

- 其中：$y = h(x)$, $\qquad p(x) = \phi\big(h(x)\big)h'(x)$

- 也即满足$p(x)$的随机变量x，可用满足$\phi(y)$的随机变量y来抽样，在获得y后，$x = h^{-1}(x)$

示例

- 产生满足如下概率密度分布的随机变量
$$f(x) = \frac{1}{2}\left(3e^{-2x} + 1\right)e^{-x}, x > 0;$$
$$f(x) = 0, otherwise$$
- 取$y = h(x) = e^{-x}$; $\phi(y) = \frac{1}{2}(3y^2 + 1)$

- 我们有 $f(x)dx = \phi(y)dy$
- 从而我们只需产生满足$\phi(y) = \frac{1}{2}(3y^2 + 1)$分布的随机变量y即可，而$x = -\ln y$

与反函数法的关联

- 若$\phi(y)$为$(0,1)$内的均匀分布，则变换抽样退化为反函数法。
- 实际上，在变换抽样中，按照$\phi(y)$分布的随机变量原则上也可以由（0,1)内的均匀分布获得
- 一般的变换抽样: (0,1)内的均匀分布按照$\phi(y)$产生y，y按照h(x)产生x。
- 反函数法: (0,1)内的均匀分布直接按照h(x)产生x。

二维推广

- 若需要产生满足联合概率密度f(x,y)的随机变量x与y
- 若满足g(u,v)的随机变量(u,v)容易产生
- 且我们可找到如下变换

$$u = h_1(x,y),\ v = h_2(x,y)$$

使得 $f(x,y) = g(u,v)\,J,$

其中J为上述变换的雅可比行列式

$$|J| = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix}$$

- 则我们可生成满足g(u,v)的u与v，利用h_1与h_1的逆生成x与y即可
- 本质仍为样本变换。

示例：正态分布的产生

- 可用极限近似法，利用中心极限定理
- 生成N个(0,1)区间均匀分布的随机数，
- 单个随机数的均值为1/2, 方差为 1/12
- 考察N个随机数的均值

$$\xi = \frac{1}{N}\sum_i \xi_i$$

- 则如下变量近似服从正态分布

$$\frac{\xi - \frac{1}{2}}{1/\sqrt{12N}}$$

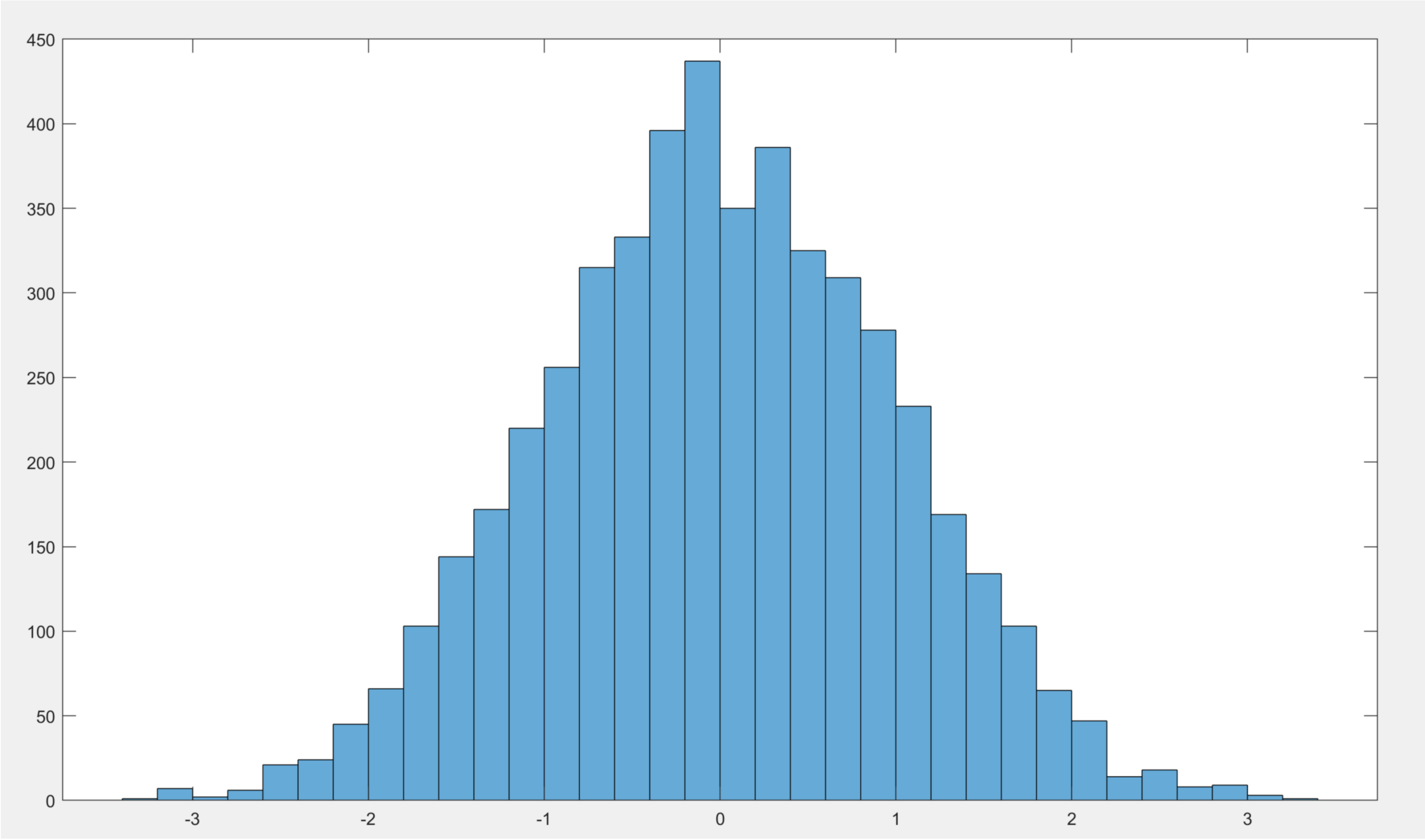
- 注：此法产生的随机数有范围限制。

代码及效果

```
function [x] = mynorm1(Nt,N1)
%MYNORM 此处显示有关此函数的摘要
% 此处显示详细说明


temp=rand(N1,Nt);
x=(sum(temp)/N1-0.5)/(12^(-0.5) * N1^(-0.5));


end
```



示例：变换抽样法

- 考虑如下变换

$$f(x)f(y)dxdy = \frac{1}{2\pi}\ e^{-\frac{x^2+y^2}{2}}dxdy = \frac{1}{2\pi}\ e^{-\frac{r^2}{2}}rdrd\phi$$

$$= \frac{1}{4\pi}\ e^{-\frac{r^2}{2}}dr^2d\phi = \frac{1}{4\pi}\ e^{-\frac{u}{2}}dud\phi$$

- 我们只需产生在$u \in (0, +\infty)$ 与$v \in (0,2\pi)$内满足$\frac{1}{4\pi}\ e^{-\frac{u}{2}}$分布的随机数即可
- 做法如下：产生(0,1)内的两个随机数ξ_1，ξ_2
- 取 $\phi = 2\pi\xi_1, u = -2\ln\xi_2$
- 则$x = \sqrt{u}\cos(\phi)$，$y = \sqrt{u}\sin(\phi)$即为满足正态分布的两个随机变量。

示例：去除三角函数

- 做法如下：产生(0,1)内的两个随机数ξ_1，ξ_2
- 计算$w = (2\xi_1 - 1)^2 + (2\xi_2 - 1)^2$
- 记$a = 2\xi_1 - 1, b = 2\xi_2 - 1,$则a,b在(-1,1)上均匀分布
- 因为$dadb = rdr\, d\phi = \frac{1}{2} dwd\phi,$考虑单位圆内的w，其在(0,1)上均匀分布
- 若w<1, 取 $x = \sqrt{-\frac{2\ln w}{w}}a,\ y = \sqrt{-\frac{2\ln w}{w}}b$

  即为满足正态分布的两个随机变量。

连续情形：舍选法

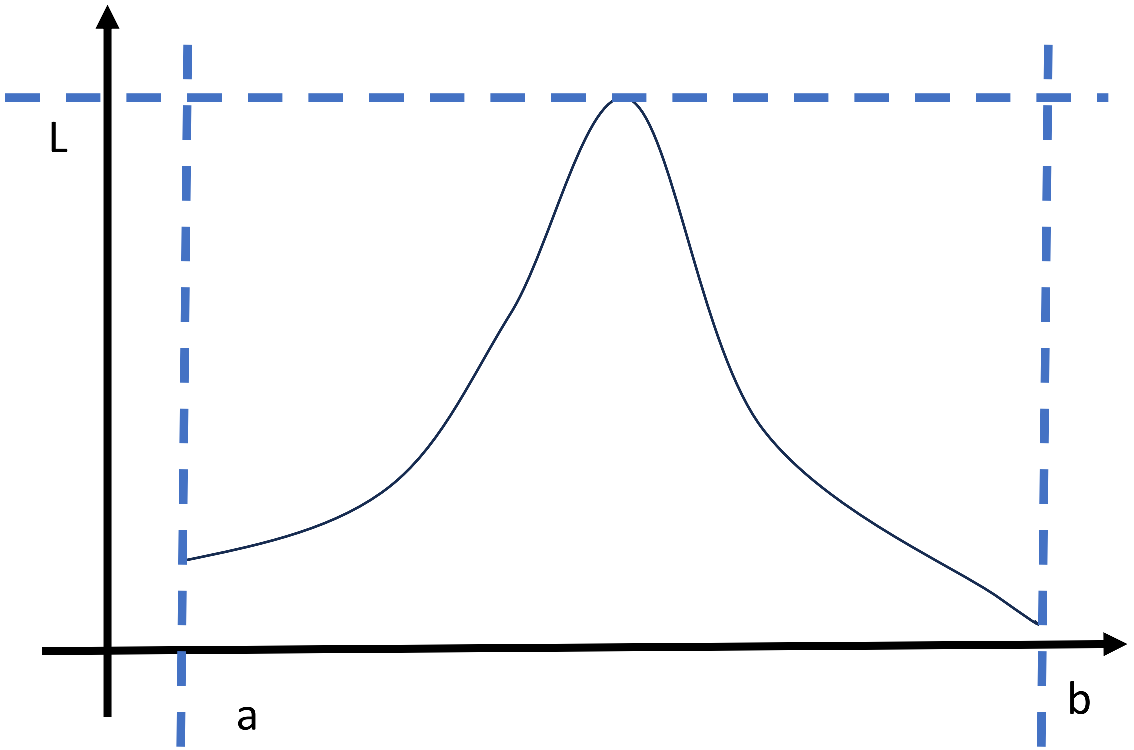
- 变量及概率$(k, p_k) \to \left(x, p(x)\right)$
- 找到p_k的最大值p_{max} 变为 找到$p(x)$最大值$p_0 = L$
- k的取值范围为N， x的取值范围为(a,b)
- 在取值范围内产生随机整数k变为在取值范围内产生随机数x
- 产生$(0, p_{max})$内的随机数ξ变为产生$(0, L)$内的随机数ξ
- 判断条件（可将连续函数无限分割以离散化）

$$\xi < p_k \quad \to \quad \xi < p(x)$$

- 若满足判断条件，则x即为抽样值

抽样效率

- 抽样效率：产生有效的随机变量的概率



$$E = \frac{\int_a^b dx\, p(x)}{L(b-a)} = \frac{1}{L(b-a)}$$

从这里可看出，定出p(x)的最大值这一步不是必须的。实际上任何p(x)的上界都可以。定出最大值的目的只是为了提高效率。

例子

- 产生满足如下分布的随机变量

$$f(x) = 2x, x \in (0,1);$$
$$f(x) = 0, otherwise$$

- 反函数法：$x = \xi^{1/2}$
- 舍选法：产生(0,1)内均匀分布的两个随机变量ξ_1, ξ_2
- 则$2\xi_1$在(0,2)上均匀分布
- 判断：$2\xi_1 < f(\xi_2) = 2\xi_2$,也即 $\xi_1 < \xi_2$
- 若满足，ξ_2即为抽样值，若不满足，则重新产生随机数进行判断。
- 优化：实际上，ξ_1, ξ_2并无区别，将前者作为待定抽样值也没问题，故在$\xi_1 < \xi_2$时只需将ξ_1返回作为抽样值即可，这可以保证每次试探均有可选值。

推广

- 产生满足如下分布的随机变量

$$f(x) = nx^{n-1}, x \in (0,1)\ (n \geq 1);$$
$$f(x) = 0, otherwise$$

- 产生n个满足(0,1)内均匀分布的随机数

$$\xi_1, \xi_2, \cdots, \xi_n$$

- 取$\xi = \max(\xi_1, \xi_2, \cdots, \xi_n)$
- ξ即为所需抽样值
- 计算$\xi = x$的概率：则在n个随机数中，1个为x,其余比x小

$$p(\xi = x) = C_n^1\, x^{n-1} = nx^{n-1}$$

第二类舍选法

- 需要原概率密度函数可分解
  $$f(x) = g(x)h(x)$$
- 其中，满足$h(x)$分布的随机变量容易抽样获得。
- 假设$g(x)$最大值为L，则可用如下舍选法：
- 1. 产生[0,1]中的两个随机变量ξ_1, ξ_2。
- 2. 利用ξ_1按照反函数法获得满足$h(x)$分布的随机变量η
- 3. 判断$L\xi_2 < g(\eta)$
- 4. 若满足，则η即为所需抽样。
- 注，也可如课本中将L提出，本质是一样的。

原理：条件概率

- η满足的概率如下：
  选中η需：按照$h(\eta)$的概率先选出η，然后需通过判定条件
- 根据条件概率

$$p(\eta) \propto h(\eta)g(\eta)$$

- 由归一化可定出前置常数L

例子：正态分布

- 产生满足标准正态分布的随机变量

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}};$$

- 利用偶函数的性质，将x限制为正，此时$f(x) \rightarrow 2f(x)$
- 选$h(x) = e^{-x}$
- 则 $g(x) = \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}+x} = \sqrt{\frac{2e}{\pi}} e^{-\frac{(x-1)^2}{2}}$
- 产生(0,1)上的两个随机变量ξ_1, ξ_2
- 按照$h(x)$产生中间随机变量：$\eta = -\ln \xi_1$
- 判断：$\xi_2 < e^{-\frac{(\eta-1)^2}{2}}$
- 若不满足，则重新产生随机数进行判断。
- 若满足，再产生一个(0,1)上的随机变量ξ_3
- 若$\xi_3 > 0.5$, 则η为抽样值，否则$-\eta$为抽样值

第三类舍选法

- 需要原概率密度函数可分解

$$f(x) = L\int_{-\infty}^{h(x)} dy\, g(x,y)$$

$$\frac{1}{L} = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{h(x)} dy\, g(x,y)$$

- 其中，$g(x,y)$需要是一个合理的联合概率密度函数。
- 方法：1. 根据$g(x,y)$产生随机向量(ξ_1,ξ_2)

  2. 判断$\xi_2 < h(\xi_1)$

  3. 若满足，则ξ_1即为所需抽样值。

抽样效率为1/L

若$g(x,y) = h(x)\, t(y)$，且$t(y) = 1$,则退回第二类舍选法。

例子

- 产生各向同性方位角余弦的抽样:
- 方法1：可产生$(0,2\pi)$的均匀分布的角度，然后求余弦。
- 若不想求余弦，则可用舍选法。此余弦的概率密度为

$$f(x) = \frac{1}{\pi}\frac{1}{\sqrt{1-x^2}},\ |x| < 1$$

- 产生(0,1)内的两个随机数ξ_1, ξ_2,并计算

$$x = \frac{\xi_1^2-\xi_2^2}{\xi_1^2+\xi_2^2}, \qquad y = \xi_1^2 + \xi_2^2$$

- 反解得：$\xi_1 = \sqrt{\frac{1}{2}y(1+x)}, \xi_2 = \sqrt{\frac{1}{2}y(1-x)}$
- 获得：$g(x,y) = f(\xi_1, \xi_2)|J| = \frac{1}{4}\frac{1}{\sqrt{1-x^2}}, |\mathrm{x}| < 1, 0 < \mathrm{y} < 2$
- 从而：$f(x) = \frac{4}{\pi}\int_{-\infty}^{1} g(x,y)dy,\ L = \frac{4}{\pi}, h(x) = 1.$
- x为余弦，正弦为$\frac{2\xi_1\xi_2}{\xi_1^2+\xi_2^2}$.
- 这里正弦一直是正的，若想获得负的需做修改。实际上，前面的课程已给出正方内接圆的方法。

复合抽样法

- 所需分布密度函数与另一变量有关：

$$f(x) = \int_{-\infty}^{+\infty} g(x|y)\, h(y) dy$$

- 首先按照$h(y)$产生随机变量y0（舍选法，反函数法）
- 当y0已产生，按照$g(x|y0)$ 产生随机变量x
- 原理为条件概率。

加分布抽样

- 所需分布密度函数可写为:

$$f(x) = \sum_i p_i g_i(x)$$

- 这对应之前的一般性方法中$h(y)$取离散值的情况。
- 物理中的对应：哈密顿量的本征值为ϵ_n,本征态为$\psi_n(x)$

 当我们已知处于某个本征态的概率p_i（比如费米分布），则系统处于某种混合态。此时，在某个位置x处的概率密度为

$$p(x) = \sum_i p_i |\psi_i(x)|^2$$

方法

- 下标取n的概率为p_n
- 故可用塔式抽样法首先确定随机变量n的值
- 当n的值定好之后，按照概率密度$g_n(x)$产生随机变量x即可

- 在实际应用中需注意，每个$g_n(x)$最好是归一化的。虽然前面很多方法里不需要概率密度整体归一，但对于几个分布的和，如果有某个不归一，可能会影响应对大小，从而出现系统误差。

例子

球壳均匀分布的抽样。对于在内径为R_0外径为R_1的球壳内均匀分布的点，其半径为r的概率密度为

$$f(r)=\frac{3r^2}{R_1^3-R_0^3}, R_0<r<R_1;$$

- 如何产生r的抽样值？
- 可直接用反函数法：产生(0,1)内均匀分布的随机数ξ，

  则$x=\left(\left(R_1^3-R_0^3\right)\xi+R_0^3\right)^{1/3}$
- 也可用舍选法：产生两个(0,1)内的随机数ξ_1,ξ_2
- 计算$\eta=(R_1-R_0)\xi_1+R_0$
- 判断$\xi_2<\frac{\eta^2}{R_1^2}$，若满足，则$\eta$为所需抽样值。

加分布方法

- 做变换 $r=(R_1-R_0)x+R_0$,则x在（0,1）内取值。
- 定义$\lambda=R_1^2+R_1R_0+R_0^2$
- 则 $f(r)=\frac{(R_1-R_0)^2}{R_1^3-R_0^3}\ 3x^2+\frac{3R_0(R_1-R_0)}{R_1^3-R_0^3}\ 2x+\frac{3R_0^2}{R_1^3-R_0^3}$
- 而由此可得，x的分布为

$$g(x)=f(r)|J|=\frac{(R_1-R_0)^2}{\lambda}\ 3x^2+\frac{3R_0(R_1-R_0)}{\lambda}\ 2x+\frac{3R_0^2}{\lambda}$$

- 只需对三种简单分布抽样即可，对应方法前面已讲过。

算法步骤

- 1. 产生四个在(0,1)上均匀分布的随机数ξ_1,ξ_2,ξ_3,ξ_4。
- 2. （塔式分布确认n值）若$\xi_1 < \frac{3R_0^2}{\lambda}$,则$\eta = \xi_2$
  (和此概率对应的是均匀分布)
- 3.否则，若$\xi_1 < \frac{3R_0R_1}{\lambda}$,则$\eta = \max(\xi_2,\xi_3)$
  （和此概率对应的概率分布是$2x$）
- 4.否则，$\eta = \max(\xi_2,\xi_3,\xi_4)$
  （和此概率对应的概率分布是$3x^2$）
- 所需的抽样值为 $r = (R_1 - R_0)\eta + R_0$

减分布抽样

- 所需分布密度函数可写为：

$$f(x) = A_1 g_1(x) - A_2 g_2(x)$$

- 严格来讲这个不可用前面的一般理论解释，因为概率需要为正。
- 物理中的对应：有些玻色子满足的运动方程为

$$J i \partial_t \psi = H\psi$$

其中J和H都为厄米矩阵，但二者不需对易。一个有代表性的例子为$J = \sigma_z$

此时的本征值问题为 $\omega\psi = J^{-1} H\psi$,需要考虑奇异值分解。

假设哈密顿量的一个本征态为$\psi(x) = \left(\psi_1(x), \psi_2(x)\right)$

则系统处于此态时，在某个位置x处的概率密度为

$$p(x) = |\psi_1(x)|^2 - |\psi_2(x)|^2$$

方法

- 核心是消掉“负概率”
- 首先获得$g_2(x)/g_1(x)$的下届，记为m。注意，这里不需要是最高下届。
- 则$0 < f(x) = g_1(x)\left(A_1 - \frac{A_2 g_2}{g_1}\right) \le g_1(x)(A_1 - A_2 m)$
- $A_1 - A_2 m > 0$必须成立，并且$\frac{f(x)}{(A_1 - A_2 m) g_1} \le 1$
- 故我们可定义$f(x) = (A_1 - A_2 m) g_1(x) h_1(x)$

  并获得 $h_1(x) = \frac{A_1}{A_1 - A_2 m}\left(1 - \frac{A_2 g_2}{A_1 g_1}\right)$，按照第二类舍选法抽样即可，

抽样效率为$\frac{1}{A_1 - A_2 m}$

方法补充

- 也可提取$g_2(x)$

$$0 < f(x) = g_2(x)\left(\frac{A_1 g_1}{g_2} - A_2\right) \le g_2(x)\left(\frac{A_1}{m} - A_2\right)$$

- 故$\frac{mf(x)}{(A_1 - A_2 m)g_2} \le 1$
- 我们可定义$f(x) = \frac{(A_1 - A_2 m)}{m} g_2(x) h_2(x)$

并获得 $h_2(x) = \frac{A_2 m}{A_1 - A_2 m}\left(\frac{A_1 g_1}{A_2 g_2} - 1\right)$，按照第二类舍选法抽样即可，抽样效率为$\frac{m}{A_1 - A_2 m}$

需比较此法与上法那个抽样效率高，采取效率高的方法。

例子

随机变量x满足的概率密度为

$$f(x) = \frac{3}{4}x^2 - \frac{3}{4}, 1 < r < 2;$$

- 如何产生x的抽样值?
- 因为$g_1(x) = \frac{3}{7}x^2, g_2(x) = 1,$故$\frac{7}{12} \leq \frac{g_2}{g_1} \leq \frac{7}{3}$我们有$m = \frac{7}{12} < 1,$故应采取第一种方法$, A_1 = \frac{7}{4}, A_2 = \frac{3}{4},$
  $A_1 - A_2 m = \frac{21}{16}, h_1(x) = \frac{A_1}{A_1 - A_2 m}\left(1 - \frac{A_2 g_2}{A_1 g_1}\right) = \frac{4}{3}\left(1 - \frac{1}{x^2}\right)$
- 采用第二类舍选法：产生两个(0,1)内的随机数ξ_1, ξ_2
- 按照$g_1(x)$产生x的抽样:$x = (7\xi_1 + 1)^{1/3}$
- 判断$\xi_2 < h_1(x)$，若满足，则x为所需抽样值。

乘加分布抽样

- 加分布问题： $f(x)=\sum_i p_i g_i(x)$
- 乘加分布问题：$f(x)=H_1(x)g_1(x)+H_2(x)g_2(x)$ 每项的函数不归一。

- 可将其化为加分布：找到p_1,p_2满足

$$p_1=\int_a^b H_1(x)g_1(x)dx\,,\quad p_2=\int_a^b H_2(x)g_2(x)dx$$

- 则

$$f(x)=H_1(x)g_1(x)+H_2(x)g_2(x)=p_1h_1(x)+p_2h_2(x)$$

其中 $\int_a^b h_1(x)dx=1\,,\int_a^b h_2(x)dx=1$

乘减分布抽样

- 乘减分布问题：$f(x) = H_1(x)g_1(x) - H_2(x)g_2(x)$ 每项的函数不一定归一。

- 原理仍是减分布：

$$m = \min\frac{H_2 g_2}{H_1 g_1} \qquad M = \max H_1$$

$$0 < f(x) = H_1(x)g_1(x)(1-m) \le M g_1(x)(A_1 - A_2 m)$$

- 则

$$f(x) = M\,(1-m)h_1(x)g_1(x), \quad h_1(x) = \frac{1}{M(1-m)}\left(H_1 - \frac{H_2 g_2}{g_1}\right)$$

注：1. 按照第二类舍选法进行抽样，需要两个函数里有一个可以用直接法（反函数法）获得

2. 也可以利用H_2的上届来做。

基于直方图（频数图）的抽样

- 问题：我们有某一随机变量的实验值，希望从理论上对此变量的一些性质做预测。则我们需要用代码来模拟满足此分布的随机变量。

- 如果随机变量本身取值是离散的，这就是塔式抽样。

- 否则，若此变量本身是连续的，但受实验次数限制无法连续化，而是以直方图的方式表达，则可利用塔式抽样进行简化。

- 方法1：计算出塔式分布的节点对应的分布函数的值，设第i个节点的值为F_i，则我们可沿用塔式抽样的方法，判断$F_i < x < F_{i+1}$，则和i对应的取值即为抽样值。
- 方法2：因为随机变量本身应该连续，当直方图对应的取值区间足够小时，我们可以做线性插值：$x_i' = x_i + \frac{\xi - F(x_i)}{F(x_{i+1}) - F(x_i)}(x_{i+1} - x_i)$，以$x_i'$而不是$x_i$作为抽样值。这是近似求解反函数的方法。

多维随机向量抽样

- 问题：希望获得随机向量$(\eta_1, \eta_2, \cdots, \eta_n)$, 使其分布满足联合概率密度函数 $f(\eta_1, \eta_2, \cdots, \eta_n)$
- 多维积分中的常见问题.多粒子统计问题也可遇到。

- 舍选法：假设向量在某个有限区间上取值，也即
  $\eta_1 \in [a_1, b_1], \eta_2 \in [a_2, b_2], \cdots, \eta_n \in [a_n, b_n]$
  且$L = \max f(\eta_1, \eta_2, \cdots, \eta_n) < +\infty$
  则我们可生成[0,1]上的n+1个随机数$\xi_1, \xi_2, \cdots, \xi_n, \xi_{n+1}$
  若 $L\xi_{n+1} < f((b_1 - a_1)\xi_1 + a_1, (b_2 - a_2)\xi_2 + a_2, \cdots, (b_n - a_n)\xi_n + a_1)$
  则$\eta_i = (b_i - a_i)\xi_i + a_i$是一组可取抽样值。

这种方法抽样效率过低（$\frac{1}{L(b_1-a_1)\cdots(b_n-a_n)}$）。

条件密度法

- 以二维问题为例。
- 需考察条件概率

$$f(\eta_2|\eta_1) = f(\eta_1,\eta_2)/f(\eta_1)$$

$$f(\eta_1) = \int_{-\infty}^{+\infty} d\eta_2 \ f(\eta_1,\eta_2)$$

- 则

$$f(\eta_1,\eta_2) = f(\eta_1)f(\eta_2|\eta_1)$$

- 方法：按照$f(\eta_1)$抽取η_1，当有η_1之后，按照$f(\eta_2|\eta_1)$抽取η_2

推广

- n维情况

$$f(\eta_1,\eta_2,\cdots,\eta_n)=f(\eta_1)f(\eta_2|\eta_1)f(\eta_3|\eta_1,\eta_2)\cdots f(\eta_n|\eta_1,\eta_2,\cdots,\eta_{n-1})$$

$$f(\eta_1)=\int_a^b d\eta_2\cdots d\eta_n\ \ f(\eta_1,\eta_2,\cdots,\eta_n)$$

$$f(\eta_2|\eta_1)=\frac{\int_a^b d\eta_3\cdots d\eta_n\ \ f(\eta_1,\eta_2,\cdots,\eta_n)}{f(\eta_1)}$$

$$f(\eta_3|\eta_1,\eta_2)=\frac{\int_a^b d\eta_4\cdots d\eta_n\ \ f(\eta_1,\eta_2,\cdots,\eta_n)}{f(\eta_1)f(\eta_2|\eta_1)}$$

- 方法：按照$f(\eta_1)$抽取η_1，当有η_1之后，按照$f(\eta_2|\eta_1)$抽取η_2, 以此类推。

例子（1）

角度θ,ϕ满足联合概率密度

$$f(\theta,\phi)=\frac{1}{\alpha}\left(1+\sqrt{3}sin\phi sin\theta\right)sin\phi\sin^2\theta,\frac{\pi}{2}>\phi\geq 0,\frac{\pi}{2}>\theta\geq 0;\alpha=\frac{3+2\sqrt{3}}{12}\pi$$

- 如何抽样产生$x=cos\phi$与 $y=cos\theta$?
- 首先写出x与y的分布密度函数，

$$g(x,y)=\frac{f(\theta,\phi)}{sin\theta sin\phi}=\frac{1}{\alpha}\left(1+\sqrt{3}\sqrt{1-x^2}\sqrt{1-y^2}\right)\sqrt{1-y^2}$$

- 从而有 $g(x)=\int dy\, g(x,y)=\int d\theta\,\frac{f(\theta,\phi)}{sin\phi}=\frac{1}{\alpha}\left(\frac{\pi}{4}+\frac{2\sqrt{3}}{3}\sqrt{1-x^2}\right)$

$$=p_1+p_2*\frac{4}{\pi}\sqrt{1-x^2}$$

$$p_1=\frac{3}{3+2\sqrt{3}}\quad,\qquad p_2=\frac{2\sqrt{3}}{3+2\sqrt{3}},\ p_1+p_2=1$$

例子（2）

- $g(y|x) = g(x,y)/g(x)$
  $$= p_3 * \frac{4}{\pi}\sqrt{1-y^2} + p_4 * \frac{3}{2}\left(1-x^2\right)$$
  $$p_3 = \frac{\pi/4}{\frac{\pi}{4}+\frac{2\sqrt{3}}{3}\sqrt{1-x^2}}, \qquad p_4 = \frac{2\sqrt{3(1-x^2)}/3}{\frac{\pi}{4}+\frac{2\sqrt{3}}{3}\sqrt{1-x^2}} \quad ,p_3 + p_4 = 1.$$

- 从而$g(x)$ 与$g(y|x)$都可用复合分布的抽样方法来抽样

算法（1）：对x抽样

- （1）产生(0,1)上的随机数ξ_1;
- （2）（用塔式抽样确定第几项）若$\xi_1 < p_1$, 则产生(0,1)上的随机数ξ_2, $x = \xi_2$;
- （3）否则，产生（0,1）上的随机数ξ_2 与ξ_3；
- （4）若$\xi_2^2 + \xi_3^2 < 1$, 则$x = \xi_2$，否则返回第（3）步执行。

- 注：这里用到了第二项中函数的最大值为$4/\pi$,并选用了舍选法对其进行抽样。

算法（2）：对y抽样

- （1）产生(0,1)上的随机数ξ_4;
- （2）（用塔式抽样确定第几项）若$\xi_4 < p_3$, 则产生(0,1)上的随机数ξ_5与ξ_6,
- （3）若$\xi_5^2 + \xi_6^2 < 1$, 则$y = \xi_5$，否则返回第（2）步执行。
- （4）若第（2）步中的判断不成立，则产生(0,1)上的随机数ξ_5与ξ_6
- （5）若$\xi_5^2 + \xi_6 < 1$则$y = \xi_5$，否则返回第（4）步执行。

- 注：这里对第二项的抽样仍然用的舍选法，但也可以采用反函数法，这时需要解三次方程。

多维正态分布随机向量

- 若随机向量$(\eta_1,\eta_2,\cdots,\eta_n)$满足标准正态分布，也即

$$f(\eta_1,\eta_2,\cdots,\eta_n)=\frac{1}{\sqrt{2\pi}}e^{-\frac{\eta_1^2}{2}}*\frac{1}{\sqrt{2\pi}}e^{-\frac{\eta_2^2}{2}}*\cdots*\frac{1}{\sqrt{2\pi}}e^{-\frac{\eta_n^2}{2}}$$

- 我们可对每个分量按照产生一维正态分布的方法独立生成（由于此算法是直接产生两个满足正态分布的变量，我们也可以一对一对的生成）
- 但若随机向量满足的正态分布不是标准形式：

$$f=(2\pi)^{-n/2}*|M|^{-1/2}*e^{-\frac{1}{2}(\boldsymbol{\eta}-\boldsymbol{\mu})^T M^{-1}(\boldsymbol{\eta}-\boldsymbol{\mu})}$$

其中，$\boldsymbol{\mu}=\langle\boldsymbol{\eta}\rangle$是期望值，M是正定且对称的协方差矩阵

$$M_{ij}=\langle(\eta_i-\mu_i)(\eta_j-\mu_j)\rangle=M_{ji}$$

- 如何对$\boldsymbol{\eta}$进行抽样？

矩阵分解

- 对于正定且对称的矩阵，可做如下分解
  $M = AA^T$ 其中A为下三角矩阵。
  则$\det(M) = \det\left(AA^T\right) = (\det A)^2$
- 若我们做如下变换$\boldsymbol{\eta} = \boldsymbol{\mu} + A\boldsymbol{\xi}$ 则$\boldsymbol{\xi}$满足标准正态分布。
- 证明：雅可比部分显然消掉了$|M|^{-1/2}$，而$\boldsymbol{\mu}$消掉了中心值，对于协方差阵，我们有
  $A^T * M^{-1} * A = A^T * \left(A^{-1}\right)^T * A^{-1} * A = I$
- 也可如下分解:用正交矩阵将M对角化，$M = ADA^T$，其中D为对角矩阵。则我们定义$S = A\sqrt{D}$,从而$M = SS^T$。注意此时S不一定是上三角或者下三角的。
- 我们仍有$\det(M) = \det\left(SS^T\right) = (\det S)^2$
- 若我们做如下变换$\boldsymbol{\eta} = \boldsymbol{\mu} + \boldsymbol{S}\boldsymbol{\xi}$ 则$\boldsymbol{\xi}$满足标准正态分布。
- 证明完全同上。

作业（9月24号）

1. 正方体色子，每面有一个不同的数字，分别为1到6。用舍选法和塔式抽样给出掷色子的结果这一随机变量的抽样。请写出算法步骤。

作业（9月29号）

1. 随机变量x满足如下概率密度分布

$$f(x) = \frac{\Gamma}{\pi}\frac{1}{x^2+\Gamma^2}$$

请用舍选法和反函数法对x进行抽样。要求：（1）写出算法步骤；（2）编写相应程序；（3）按照程序运行，并获得1000个抽样值，将抽样值的频数以直方图形式画出(matlab中可用histogram命令；只需给出直方图即可，不用附数据）。

2.（附加）黑体辐射是量子物理的起源问题之一。请完成教材第二章第10题的证明（只需做证明部分），从中了解黑体辐射中的频率抽样原理。