1. 根据上次作业中计算π的值的算法，设置投点步数 N=10000，分别写出置信水平为 68%和 95%的误差区间（对于误差函数，$\lambda=1$ 时值为 0.68，$\lambda=2$ 时值为 0.95）。

单次实验的标准差为$\sigma=\sqrt{p(1-p)}=\sqrt{\frac{\pi}{4}\left(1-\frac{\pi}{4}\right)}$，误差为$e=\lambda\frac{\sigma}{\sqrt{N}}$，置信水平为 68%即$\lambda=1$时有$e_1=\frac{\sqrt{\frac{\pi}{4}\left(1-\frac{\pi}{4}\right)}}{\sqrt{10000}}=0.00410546$，此时误差区间为$[\pi-4e_1,\pi+4e_1]$；置信水平为 95%即$\lambda=2$时有$e_2=2\frac{\sqrt{\frac{\pi}{4}\left(1-\frac{\pi}{4}\right)}}{\sqrt{10000}}=0.00821092$，计算$\pi$误差区间为$[\pi-4e_2,\pi+4e_2]$（或者写成计算概率的区间$\left[\frac{\pi}{4}-e_1,\frac{\pi}{4}+e_1\right]$，$\left[\frac{\pi}{4}-e_2,\frac{\pi}{4}+e_2\right]$）。

这里贴一段 MATLAB 代码供大家自己检验误差区间：

```
function [Pi] = Pi(N)
    Nh = 0;
    for i = 1 : N
        x = 2 * rand - 1;
        y = 2 * rand - 1;
        if x^2 + y^2 < 1
            Nh = Nh + 1;
        end
    end
    Pi = 4 * Nh / N;
end
clear;
% 每轮实验次数
n = 10000;
% 实验轮数
N = 10000;

% 误差计算
sigma = sqrt(pi/4*(1-pi/4));
e1 = 4 * sigma/sqrt(n);
e2 = 4 * 2 * sigma/sqrt(n);
Pi_trial = zeros(N, 1);
% 计数变量，用于计算落在误差范围内的实验轮数。
count1 = 0;
count2 = 0;
for i = 1 : N
```

```
    Pi_trial(i) = Pi(n);
    if abs(pi - Pi_trial(i)) < e1
        count1 = count1 + 1;
    end

    if abs(pi - Pi_trial(i)) < e2
        count2 = count2 + 1;
    end

end

% 输出误差区间内的实验轮数占比
disp(count1/10000)
disp(count2/10000)
```

```
>> count1/10000

ans =

    0.6858

>> count2/10000

ans =

    0.9544
```

评分标准（10 分）：标准差、误差都对得 10 分，标准差对、误差错得 7 分，其余答案基础分为 5 分，视过程是否详实加减分。如果有同学自己写程序估算了标准差那么也算正确，估算结果应与理论结果相差不大。

2．采用线性同余法[参见公式(2.2.3) $x_{n+1}=(ax_n+c)\ (\mathrm{mod}\ m)$]产生伪随机数。取$a=5, c=1$，$m=16$ 和$x_0=1$。记录下产生出的前 20 个数，它产生数列的周期是多少？

$x_{n+1}=(5x_n+1)\ (\mathrm{mod}\ 16)$，$x_0=1$

$x_1=6$，$x_2=15$，$x_3=12$，$x_4=13$，$x_5=2$，$x_6=11$，$x_7=8$，$x_8=9$，$x_9=14$，$x_{10}=7$，

$x_{11}=4$，$x_{12}=5$，$x_{13}=10$，$x_{14}=3$，$x_{15}=0$，$x_{16}=1$，$x_{17}=6$，$x_{18}=15$，$x_{19}=12$，$x_{20}=13$，

或者

$\xi_1=0.375$，$\xi_2=0.9375$，$\xi_3=0.75$，$\xi_4=0.8125$，$\xi_5=0.125$，$\xi_6=0.6875$，$\xi_7=0.5$，

$\xi_8=0.5625$，$\xi_9=0.875$，$\xi_{10}=0.4375$，$\xi_{11}=0.25$，$\xi_{12}=0.3125$，$\xi_{13}=0.625$，$\xi_{14}=0.1875$，

$\xi_{15}=0$，$\xi_{16}=0.0625$，$\xi_{17}=0.375$，$\xi_{18}=0.9375$，$\xi_{19}=0.75$，$\xi_{20}=0.8125$。

周期为 16。

评分标准（10 分）：前 20 项 8 分，周期 2 分。

3. 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。

舍选法：

（1） 随机生成 1 到 6 的整数 k。

（2） 产生$\left(0,\frac{1}{6}\right)$之间的随机数$\xi$；

（3） 如果$\xi < p_{\max} = \frac{1}{6}$，则接受 k 作为抽样结果。

塔式抽样法：

（1） 产生(0,1)之间的随机数ξ。

（2） 设置分布概率$P = 0$；

（3） 设置循环变量i，i从 1 到 6;

（4） 计算$P = \frac{i}{6}$；

（5） 若$\xi < P$，则返回i，否则重复循环。

评分标准（10 分）：舍选法和塔式抽样法各 5 分。

4. 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)。

（1） 舍选法：

由于概率密度函数$f(x)$随$|x|$增大衰减很快，因此可自定$[-r,\ r]$作为定义域，并且$f(x)$在该区间上最大值为$f(0) = \frac{1}{\pi\Gamma}$；算法步骤为

(a) 构造$[-r,\ r]$上均匀分布的随机数$\delta = -r + 2r\xi_1$，其中ξ_1为$[0,1]$上均匀分布的随机数；

(b) 再选取独立的均匀分布于$[0,1]$区间上的随机数ξ_2，判断$\xi_2 \le \pi\Gamma f(\delta)$是否满足，如果满足则执行(c)，否则回到(a)

(c) 选取$\eta = \delta$作为一个抽样值。

反函数法：

(a) 求解分布函数 $F(x) = \int_{-\infty}^{x}\frac{\Gamma}{\pi}\frac{1}{x^2+\Gamma^2}dx = \frac{1}{\pi}\arctan\left(\frac{x}{\Gamma}\right) + \frac{1}{2}$，利用反函数法

$x = F^{-1}(y) = \Gamma \tan\left[\pi\left(y - \frac{1}{2}\right)\right]$；

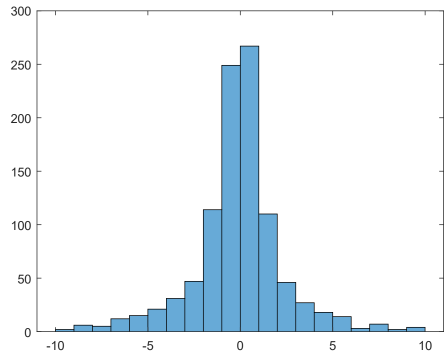
(b) 产生ξ为[0, 1]区间上的均匀分布的随机数，则所需抽样值为$\eta = \Gamma\tan\left(\pi\left(\xi - \frac{1}{2}\right)\right)$。

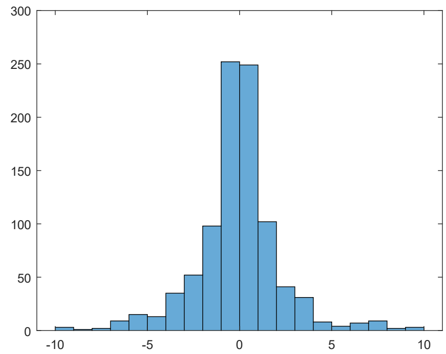
（2） 编写相应程序

（3） 直方图

MATLAB 代码：

```
clear;
% 抽样次数
N = 1000;

% 取 Gamma 为 1
Gamma = 1;

%% 舍选法
% 取 x 定义域为[-10,10]
r = 10;
eta = zeros(N, 1);
% 创建概率密度函数
syms x;
f = Gamma / pi * 1 / (x^2 + Gamma^2);

for i = 1 : N
    xi_1 = rand;
    xi_2 = rand;
    delta = - r + 2 * r * xi_1;
    while xi_2 > pi * Gamma * subs(f, x, delta)
        xi_1 = rand;
        xi_2 = rand;
        delta = - r + 2 * r * xi_1;
    end
    eta(i) = delta;
end
histogram(eta);

%% 反函数法
xi = rand(N, 1);
eta = Gamma * tan(pi * (xi - 1/2));
histogram(eta);
edges= - 10 : 1 : 10;
histogram(eta, edges);
```

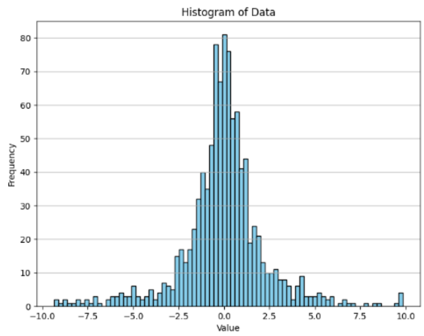




Python 代码：

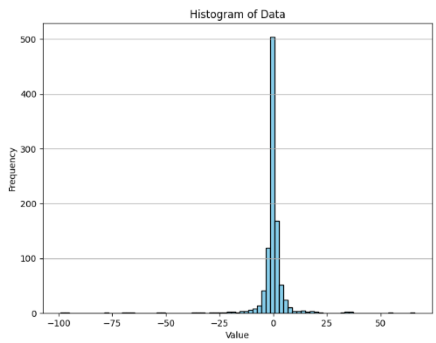
舍选法:

```
import math
import random
import matplotlib.pyplot as plt
import numpy as np

N = 10000
Num = []
num = 0
for i in range(N):
    x1 = 2 * random.random() - 1
    x2 = random.random()
    m = 1 / (math.pi * ((10 * x1)**2 + 1))
    if x2/math.pi < m:
        Num.append(10 * x1)
        num += 1
        if num == 1000:
            break

# 使用 matplotlib 绘制直方图
plt.figure(figsize=(8, 6))  # 设置图形大小
plt.hist(Num, bins=81, color='skyblue', edgecolor='black')  # bins 指定分区数
plt.title('Histogram of Data')  # 添加标题
plt.xlabel('Value')  # X 轴标签
plt.ylabel('Frequency')  # Y 轴标签
plt.grid(axis='y')  # 添加网格线
plt.show()
```



反函数法:

```
import math
import random
import matplotlib.pyplot as plt
import numpy as np

N = 1000
Num = []
for i in range(N):
    #y = random.uniform(0, 1)
    y = random.random()
    x = math.tan(math.pi * (y - 1 / 2))
    Num.append(x)

# 使用 matplotlib 绘制直方图
plt.figure(figsize=(8, 6))  # 设置图形大小
```



```
plt.hist(Num, bins=81, color='skyblue', edgecolor='black')  # bins 指定分区数
plt.title('Histogram of Data')  # 添加标题
plt.xlabel('Value')  # X 轴标签
plt.ylabel('Frequency')  # Y 轴标签
plt.grid(axis='y')  # 添加网格线
plt.show()
```

评分标准（10 分）：算法部分 5 分，程序 3 分，结果 2 分。

5.（附加）归一化黑体辐射的频谱为

$$f(x)\mathrm{d}x = \frac{15}{\pi^4}\left(\frac{x^3}{e^x-1}\right)\mathrm{d}x \qquad （其中 x = \frac{h\nu}{kT}）$$

证明如下抽样步骤得到的抽样分布满足上面的分布

抽样步骤：让 L 等于满足下面不等式的整数 l 的最小值

$$\sum_{j=1}^{l}\frac{1}{j^4} \geq \xi_1\pi^4/90$$

然后置 $x = -\frac{1}{L}\ln(\xi_2\xi_3\xi_4\xi_5)$，其中 ξ_i 为[0,1]区间均匀分布的伪随机数。

方法一：

利用 Gamma 函数分布：

$$g(x)\mathrm{d}x = \frac{a^n}{(n-1)!}x^{n-1}e^{-ax}\mathrm{d}x$$

其抽样方法为（需证明，可通过归纳法证明）

$$\eta = -\frac{1}{a}\ln(\xi_1\xi_2\cdots\xi_n)$$

归纳法步骤：

$n=1$ 时，对应于指数分布，成立。

设 $n=k$ 时成立，则 $\eta_k = -\frac{1}{a}\ln(\xi_1\xi_2\cdots\xi_k)$ 对应于

$$g_k(x) = \frac{a^k}{(k-1)!}x^{k-1}e^{-ax}$$

$n=k+1$ 时，$\eta_{k+1} = -\frac{1}{a}\ln(\xi_1\xi_2\cdots\xi_k\xi_{k+1}) = -\frac{1}{a}\ln(\xi_1\xi_2\cdots\xi_k) - \frac{1}{a}\ln\xi_{k+1}$，对应于 $g_k(x)$ 和一个指数分布，对他们的概率密度函数做卷积求联合概率密度

$$\begin{aligned} g_{k+1}(x) &= \int_0^x \frac{a^k}{(k-1)!}t^{k-1}e^{-at}\cdot ae^{-a(x-t)}\mathrm{d}t \\ &= \int_0^x \frac{a^{k+1}}{(k-1)!}t^{k-1}e^{-ax}\mathrm{d}t \\ &= \frac{a^{k+1}}{k!}x^k e^{-ax} \end{aligned} \quad ，证毕。$$

本题中$n=4$，$a=L$，因此

$$g(x)=\frac{L^4}{6}x^3e^{-Lx}$$

结合

$$P(l=L)=\frac{90}{\pi^4}\frac{1}{L^4}$$

可以得到

$$\begin{aligned}
f(x)&=\sum_{L=1}^{\infty}P(l=L)g(x)\\
&=\sum_{L=1}^{\infty}\frac{90}{\pi^4}\frac{1}{L^4}\cdot\frac{L^4}{6}x^3e^{-Lx}\\
&=\sum_{L=1}^{\infty}\frac{15}{\pi^4}x^3e^{-Lx}\text{(几何级数)}\\
&=\frac{15}{\pi^4}\left(\frac{x^3}{e^x-1}\right)
\end{aligned}$$

方法二：

直接求解：

$$G(x)=P\left(-\frac{1}{L}\ln(\xi_2\xi_3\xi_4\xi_5)\leq x\right)=\int_{e^{-Lx}}^{1}\mathrm{d}\xi_2\int_{\frac{e^{-Lx}}{\xi_2}}^{1}\mathrm{d}\xi_3\int_{\frac{e^{-Lx}}{\xi_2\xi_3}}^{1}\mathrm{d}\xi_4\int_{\frac{e^{-Lx}}{\xi_2\xi_3\xi_4}}^{1}\mathrm{d}\xi_5$$

计算积分，

$$\begin{aligned}
I&=\int_{e^{-Lx}}^{1}\mathrm{d}\xi_2\int_{\frac{e^{-Lx}}{\xi_2}}^{1}\mathrm{d}\xi_3\int_{\frac{e^{-Lx}}{\xi_2\xi_3}}^{1}\mathrm{d}\xi_4\left(1-\frac{e^{-Lx}}{\xi_2\xi_3\xi_4}\right)\\
&=\int_{e^{-Lx}}^{1}\mathrm{d}\xi_2\int_{\frac{e^{-Lx}}{\xi_2}}^{1}\mathrm{d}\xi_3\left[1+\frac{e^{-Lx}\left(\ln\frac{e^{-Lx}}{\xi_2\xi_3}-1\right)}{\xi_2\xi_3}\right]\\
&=\int_{e^{-Lx}}^{1}\mathrm{d}\xi_2\left[1-\frac{e^{-Lx}}{\xi_2}+\frac{e^{-Lx}(Lx+\ln\xi_2+1)}{\xi_2}\ln\frac{e^{-Lx}}{\xi_2}+\frac{e^{-Lx}\left(\ln\frac{e^{-Lx}}{\xi_2}\right)^2}{2\xi_2}\right]\\
&=1-e^{-Lx}-Lxe^{-Lx}-\frac{1}{2}e^{-Lx}L^2x^2-\frac{1}{6}e^{-Lx}L^3x^3\\
&=1-\frac{1}{6}e^{-Lx}(6+6Lx+3L^2x^2+L^3x^3)
\end{aligned}$$

则概率密度函数为

$$g(x)=\frac{\mathrm{d}G(x)}{\mathrm{d}x}=\frac{L^4}{6}x^3e^{-Lx}$$

其余步骤与方法一相同。

评分标准（10 分）：使用 Gamma 分布需要给出证明过程，无证明则扣 2 分。