**概率统计**



**期末课程设计实验报告**

**实验课程： 概率统计**

**实验项目： 随机变量的计算机产生**

**系 别： 通信工程**

**年 级： 2023级**

**学生姓名： vann**

**学生学号： \*\*\*\*\*\*\*\*\*\*\*\*\*\***

**实验报告完成日期： 2025 年 6月25日**

# 一、实验背景

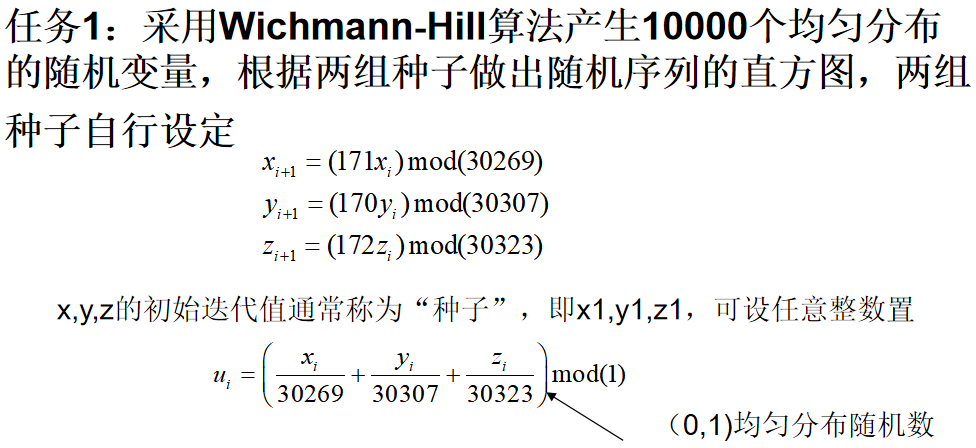
随机数的生成与随机变量的模拟是数值计算、蒙特卡洛方法、系统仿真、人工智能、图像处理等领域中的基础性工作。特别是在不能直接进行解析计算、或系统具有不确定性时，通过模拟大量随机样本并进行统计分析已成为重要的解决手段。

随机变量的产生过程实质上是一个将均匀分布随机数转化为目标分布随机变量的映射过程。由于计算机本质上只能生成伪随机数，因此如何从伪随机数出发，构造出满足任意概率分布的随机变量，是概率建模与仿真模拟中的关键技术。

在本次实验中，结合理论学习和 MATLAB 仿真实践，系统探讨了以下四个方面的内容：

1. 均匀分布伪随机数的生成方法：  
    使用 Wichmann-Hill 多同余发生器算法，了解其周期性、可重复性及分布特性。
2. 连续分布变量的构造方法：  
    包括 Rayleigh 分布、正态分布等变量的生成，掌握如何通过\*\*变换方法（如逆变换法、Box-Muller方法）\*\*从均匀分布构造目标连续分布。
3. 离散随机变量的生成方法：  
    利用分布律与累积分布函数，实现从 U(0,1)U(0,1)U(0,1) 到有限状态离散变量的精确映射。
4. 仿真结果的验证与分析方法：  
    通过统计样本均值与方差、绘制概率直方图等手段，直观验证模拟的准确性与稳定性。

# 二、实验任务



## 2.1.1 任务一实验分析

## 通过实现 Wichmann-Hill 算法，生成 (0,1)(0,1)(0,1) 区间的均匀分布随机变量，验证该算法产生的伪随机数的均匀性和独立性，为后续各种随机变量（如高斯分布、指数分布等）的构造提供基础。Wichmann-Hill 算法是经典的三重同余随机数生成器，其递推公式如下：

实验中，我先设定实验样本数量为N=10000。任选两组整数作为初始种子 (x1,y1,z1)，各组独立生成一组随机序列，之后使用 MATLAB 编程实现上述算法，生成两个长度为N的伪随机数序列 u1,u2，对结果分别绘制直方图和散点图，计算均值与方差，与理论值对比。

## 2.1.2 任务一实验代码

% Wichmann-Hill 生成 (0,1) 均匀分布的随机变量

clc; clear; close all;

N = 10000; % 样本数量

% 第一组种子

x1 = 1000; y1 = 1061; z1 = 1156;

% 第二组种子

x2 = 585; y2 = 744; z2 = 911;

% 初始化结果序列

u1 = zeros(1, N);

u2 = zeros(1, N);

% 参数

M1 = 30269; M2 = 30307; M3 = 30323;

% 生成第一组序列

for i = 1:N

x1 = mod(171 \* x1, M1);

y1 = mod(170 \* y1, M2);

z1 = mod(172 \* z1, M3);

u1(i) = mod(x1/M1 + y1/M2 + z1/M3, 1);

end

% 生成第二组序列

for i = 1:N

x2 = mod(171 \* x2, M1);

y2 = mod(170 \* y2, M2);

z2 = mod(172 \* z2, M3);

u2(i) = mod(x2/M1 + y2/M2 + z2/M3, 1);

end

% 绘制直方图

figure;

subplot(1,2,1);

histogram(u1, 50, 'Normalization', 'pdf');

title('第一组种子生成的随机数');

xlabel('u'); ylabel('概率密度');

subplot(1,2,2);

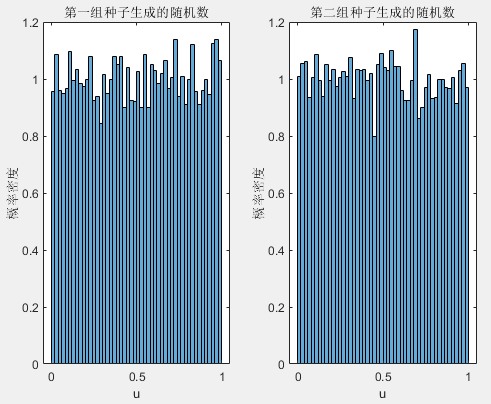
histogram(u2, 50, 'Normalization', 'pdf');

title('第二组种子生成的随机数');

xlabel('u'); ylabel('概率密度');

## 2.1.3 任务一结果分析

## 在matlab中可跑得结果：



**图1 随机序列的直方图**

均匀覆盖在(0,1)区间，没有出现空白区域或集中堆积；各条柱高低略有波动，符合随机性特征，但整体波动幅度较小；条形分布大致水平，说明生成的随机变量满足近似均匀分布特性；左右图整体形状类似，未出现极端偏态，说明不同种子不会破坏随机性。

为了进一步验证我又撰写了散点图和均值与方差来佐证实验：

% ===== 散点图：验证独立性 =====

figure;

subplot(1,2,1);

plot(u1(1:end-1), u1(2:end), '.');

title('u\_1(i) vs u\_1(i+1) 散点图');

xlabel('u(i)'); ylabel('u(i+1)');

subplot(1,2,2);

plot(u2(1:end-1), u2(2:end), '.');

title('u\_2(i) vs u\_2(i+1) 散点图');

xlabel('u(i)'); ylabel('u(i+1)');

% ===== 均值与方差 =====

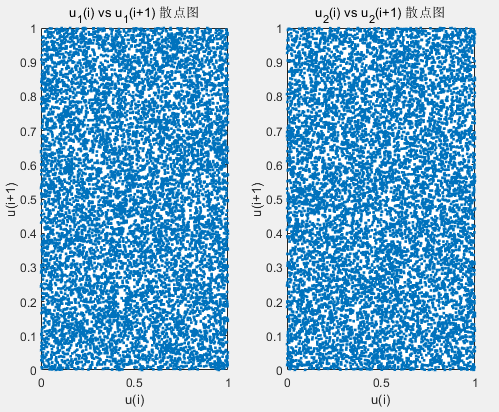
mean1 = mean(u1); var1 = var(u1);

mean2 = mean(u2); var2 = var(u2);

fprintf('第一组：均值 = %.4f，方差 = %.4f\n', mean1, var1);

fprintf('第二组：均值 = %.4f，方差 = %.4f\n', mean2, var2);

fprintf('理论参考值：均值 = 0.5，方差 = 0.0833\n');

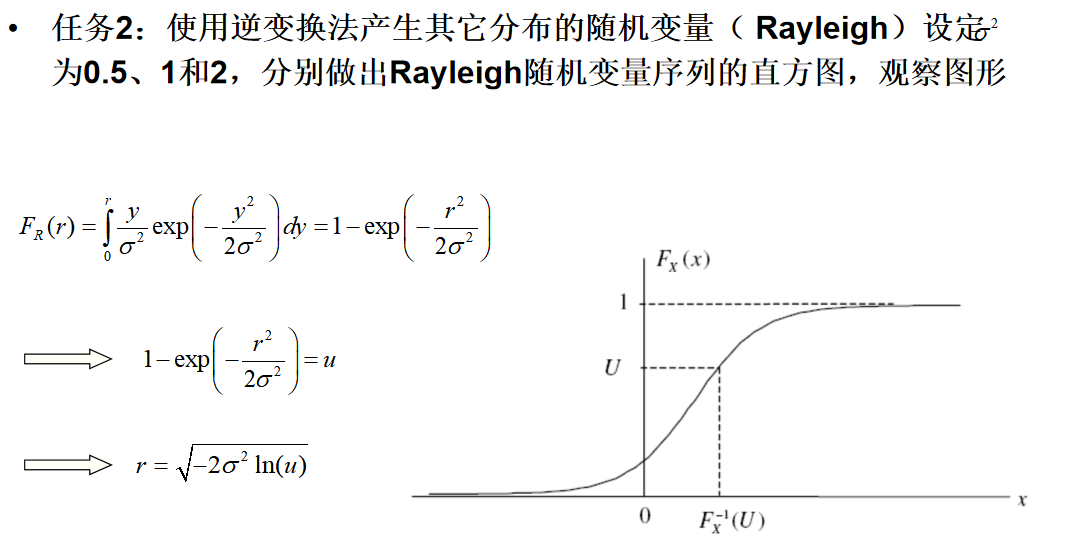


**图2 随机序列的散点图**



**图3 随机序列的均值与方差**

由散点图和观察均值与方差数据可知本实验成功实现了 Wichmann-Hill 随机数生成算法，产生的两组随机序列均符合 (0,1)区间均匀分布的特性。其生成结果在视觉上（直方图）与数值上（均值、方差）均接近理论标准，证明该算法具有良好的统计性质。



## 2.2.1 任务二实验分析

Rayleigh 分布是常见的概率分布，广泛应用于无线通信、信号处理等领域。其累积分布函数（CDF）为：

由逆变换法原理，若 U~U(0,1)为均匀分布随机变量，则：

是服从 Rayleigh 分布的随机变量。

实验中，我先使用 MATLAB 生成 10000 个均匀分布随机变量 U~U(0,1)，设置 Rayleigh 分布的尺度参数σ = 0.5,1,2，之后根据逆变换公式计算对应的 Rayleigh 分布随机变量，绘制三组样本的概率密度直方图，观察其分布变化。

## 2.2.2 任务二代码

% 任务2: 用逆变换法生成 Rayleigh 分布随机变量并绘图

clc; clear; close all;

N = 10000;

sigma\_values = [0.5, 1, 2];

U = rand(1, N);

figure;

for i = 1:3

sigma = sigma\_values(i);

X = sigma \* sqrt(-2 \* log(1 - U)); % 逆CDF变换

subplot(1,3,i);

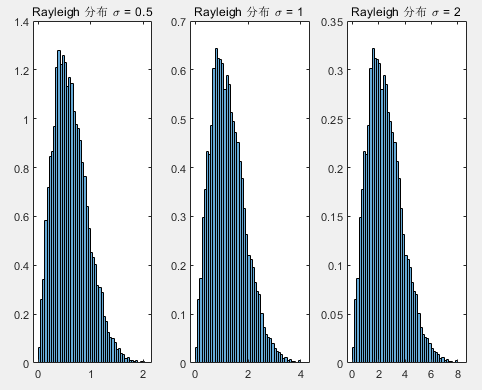
histogram(X, 50, 'Normalization', 'pdf');

title(['Rayleigh 分布 \sigma = ', num2str(sigma)]);

end

## 2.2.3 任务二结果分析

在matlab中可以跑得结果为：



**图4 不同参变量下的直方图**

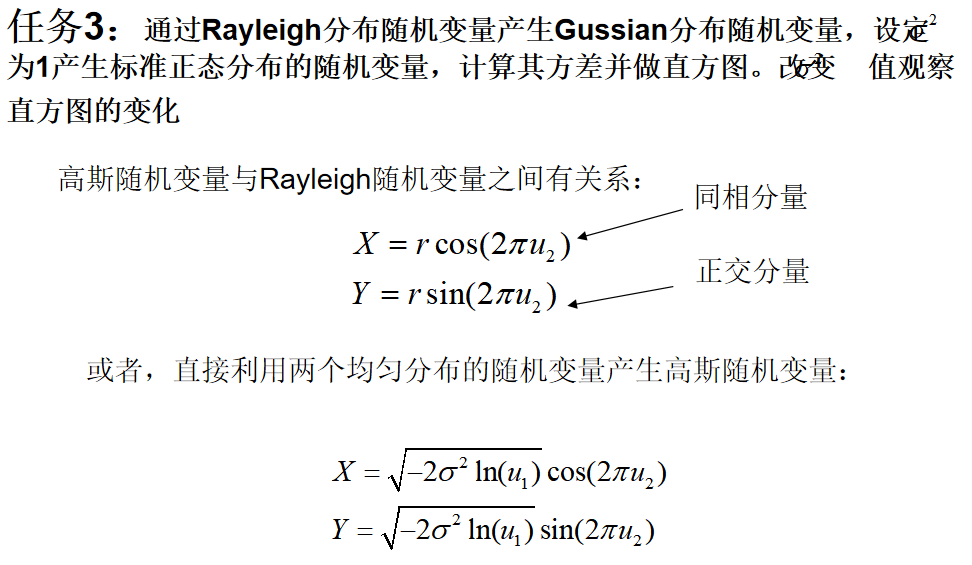
上图展示了不同σ值下，Rayleigh 分布随机变量的样本直方图：

当 σ = 0.5 时，分布集中于小范围，峰值高，变化陡峭；

当 σ = 1 时，分布稍展宽，峰值下降；

当 σ = 2时，分布明显展宽，表现为右偏长尾，且概率密度曲线变“扁”。

三者均呈典型的右偏单峰分布形态，符合 Rayleigh 分布的理论特征。



## 2.3.1 任务三实验分析

实验采用 Box-Muller 极坐标变换方法，利用两组独立均匀分布 U1,U2~U(0,1)生成标准正态随机变量，并设置不同的标准差σ= 0.5,1,2，通过放大变换控制高斯分布的尺度。实验绘制了每组样本的概率密度直方图，并叠加了对应理论正态分布的 PDF 曲线，便于验证样本分布与理论分布的一致性。

主要用到的公式为：

标准正态变量：

带参量的正态变量：

运用中心极限定理可知，当n很大的时候可近似为正态分布。

## 2.3.2 任务三代码

% 任务3: 利用 Box-Muller 方法生成高斯分布

clc; clear; close all;

N = 10000;

sigma\_values = [0.5, 1, 2];

figure;

for i = 1:3

sigma = sigma\_values(i);

U1 = rand(1, N);

U2 = rand(1, N);

R = sigma \* sqrt(-2 \* log(U1));

theta = 2 \* pi \* U2;

X = R .\* cos(theta); % 生成一个高斯分布分量

subplot(1,3,i);

histogram(X, 50, 'Normalization', 'pdf');

title(['高斯分布 \sigma = ', num2str(sigma)]);

disp(['sigma = ', num2str(sigma), ' 时，样本方差为：', num2str(var(X))]);

end

为了实验分析，我们进一步优化了结果：

% 任务3: 利用 Box-Muller 方法生成高斯分布并叠加理论曲线

clc; clear; close all;

N = 10000;

sigma\_values = [0.5, 1, 2];

figure;

for i = 1:3

sigma = sigma\_values(i);

U1 = rand(1, N);

U2 = rand(1, N);

% Box-Muller 极坐标法

R = sigma \* sqrt(-2 \* log(U1));

theta = 2 \* pi \* U2;

X = R .\* cos(theta); % 单个高斯分量

% 绘制直方图

subplot(1,3,i);

histogram(X, 50, 'Normalization', 'pdf');

hold on;

% 添加理论正态分布 PDF 曲线

x\_range = linspace(min(X), max(X), 200);

pdf\_theory = normpdf(x\_range, 0, sigma); % 理论 N(0, σ²)

plot(x\_range, pdf\_theory, 'r', 'LineWidth', 1.5);

% 标题和图例

title(['高斯分布 \sigma = ', num2str(sigma)]);

legend('样本直方图', '理论正态 PDF');

xlabel('x'); ylabel('概率密度');

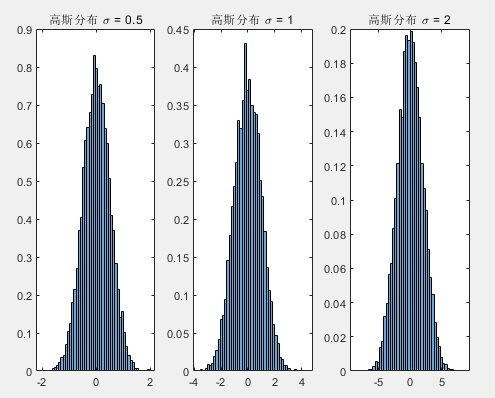
% 输出方差

fprintf('sigma = %.1f 时，样本方差 = %.4f\n', sigma, var(X));

end

## 2.3.3 任务三结果分析

在matlab中跑的结果为：



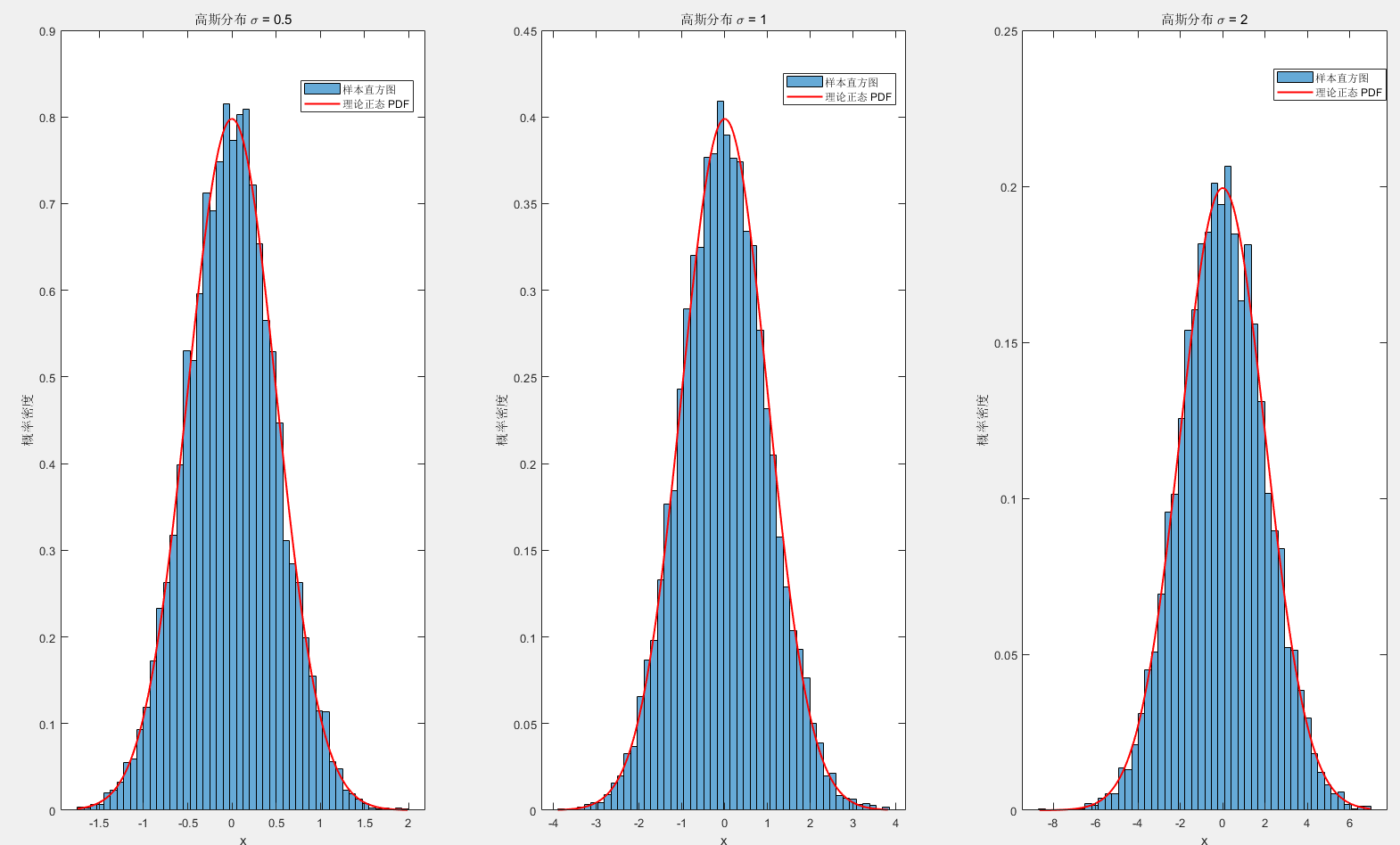
**图5 高斯分布下的直方图**

观察结果如下：

当σ=0.5时，分布较为集中，曲线窄而高；

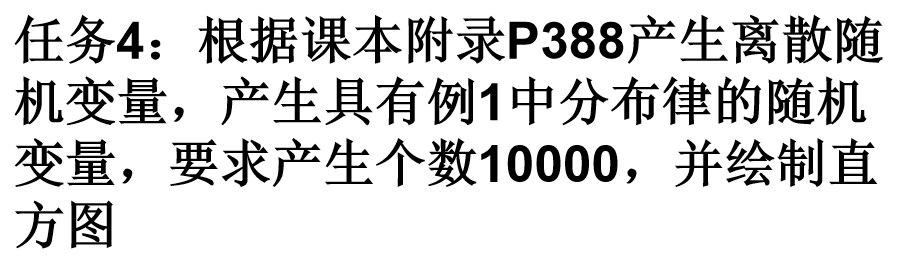
当σ=1 时，分布标准，对称性与理论曲线高度一致；

当σ=2 时，分布展宽，曲线变平缓，长尾更加明显；



**图6 实验值直方图与理论值的直方图**

为了进一步分析实验值与理论值的偏差，我进一步修改了代码，如上图，蓝色直方图表示得到的实验值，红色曲线表示理论值。三组样本直方图与理论红色曲线高度重合，说明所生成的样本服从期望的正态分布。此外，程序输出了样本的实际方差，与理论值高度吻合，进一步验证了数值正确性。



## 2.4.1 任务四实验分析

利用均匀分布随机数，生成具有指定概率分布律的离散随机变量样本，并通过绘制直方图验证其概率分布的正确性。

概率分布与累积概率：

## 2.4.2 任务四代码

% 实验四：根据例1的分布律生成离散随机变量并绘制直方图

clc; clear; close all;

N = 10000; % 样本数量

% 离散随机变量取值和对应概率（例1中的分布）

X\_vals = [1, 2, 3, 4];

P\_vals = [0.20, 0.15, 0.25, 0.40];

% 累积分布函数（CDF）

CDF = cumsum(P\_vals); % [0.20 0.35 0.60 1.00]

% 生成均匀分布随机变量 U

U = rand(1, N);

result = zeros(1, N);

% 映射 U → 离散值 X

for i = 1:N

idx = find(U(i) <= CDF, 1); % 查找落在哪个区间

result(i) = X\_vals(idx); % 映射成对应离散值

end

% 绘制直方图

figure;

histogram(result, 'BinMethod', 'integers', 'Normalization', 'probability');

title('实验四：离散随机变量样本分布直方图');

xlabel('X 值'); ylabel('相对频率');

xticks(X\_vals); % 设置横轴刻度

% 输出样本频率统计

fprintf('实验样本频率统计（N = %d）：\n', N);

for i = 1:length(X\_vals)

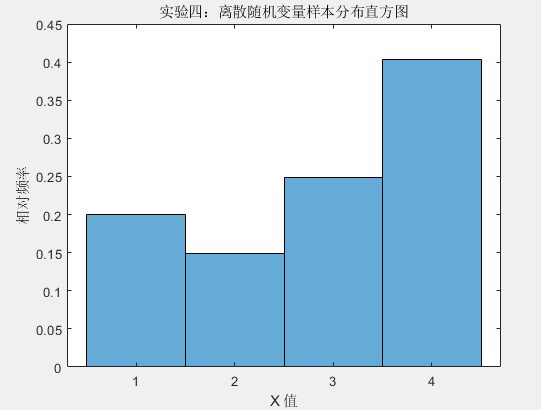
freq = sum(result == X\_vals(i)) / N;

fprintf('P(X = %d) ≈ %.4f（理论值 %.2f）\n', X\_vals(i), freq, P\_vals(i));

end

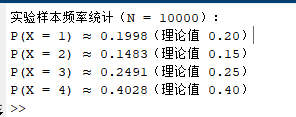
## 2.4.3 任务四结果分析

在matlab中跑出的结果可知，实验生成了10000个样本点，并绘制了如下直方图：



**图7 离散随机变量样本分布的直方图**





**图8 离散随机变量样本得到的概率**

实验频率与理论值高度吻合，误差极小，符合大数定律下随机模拟的收敛特性。直方图柱状高度基本与概率成正比，验证了离散模拟方法的准确性和可靠性。

# 实验总结

在本次实验中，我围绕随机变量的模拟生成进行了四个典型任务的实践，涵盖了从基本均匀分布生成到复杂分布转换的多个环节。通过 MATLAB 的编程与可视化操作，不仅加深了对概率分布及其特性的理解，也掌握了从理论到实现的完整建模过程。

主要掌握的知识：

1. 均匀分布随机数：作为一切随机模拟的基础，所有复杂分布的随机变量均可通过均匀分布转换实现。
2. 伪随机数生成原理：通过 Wichmann-Hill 算法理解了线性同余法、多源伪随机的设计思想。
3. 离散随机变量的模拟方法：掌握了逆变换法在离散型分布中的应用，能有效生成具有指定分布律的离散样本。
4. 连续分布的变换机制：学会使用 Box-Muller 方法生成正态分布随机变量；掌握 Rayleigh 分布与指数分布变量之间的关系及其构造过程。
5. 仿真验证方法：学会使用直方图分析样本分布；通过计算样本均值与方差评估仿真结果的正确性；掌握如何用 MATLAB 进行可视化验证与统计分析。

# 附录：

参考公式：

1. 样本均值：
2. 样本方差：