基本信息表

作品题目：**随机数的统计分布测试**

作品内容摘要：

本项目实现了一个基于Python的随机数生成器分析工具，旨在通过多维度测评方法评估不同随机数生成算法的性能与统计特性。系统采用模块化设计，集成了C风格rand()%N、均匀分布和正态分布三种随机数生成器，并构建了包含分布测试、正确性验证、统计检验、随机性检测、性能测试及熵值计算六大功能的测评框架。通过Tkinter构建的交互式GUI界面，用户可灵活设置参数（如数值范围N和样本量），并实时可视化呈现频率分布直方图、间隔分布曲线、Q-Q概率图及自相关系数等关键指标。系统创新性地结合了统计学检验（卡方检验、KS检验）与信息论指标（信息熵），可定量分析生成器的分布均匀性、随机性及生成效率。

关键词（五个）：

随机数生成器分析、伪随机数生成算法、统计检验框架、信息熵评估、交互式 GUI 界面

1. 作品功能与性能说明

功能说明

实现三种随机数生成算法：C 风格 rand()%N（基于 random 模块）、均匀分布（numpy.random.randint）及正态分布生成器（基于截断正态分布理论），支持动态切换以满足不同场景需求。

频率分布直方图：可视化统计各数值出现频次，验证生成结果的均匀性或正态性（如正态分布生成器的均值与方差是否符合理论值）。

间隔分布曲线：分析相同数值重复出现的时间间隔，评估序列的随机性。

卡方检验与 KS 检验：定量判断生成数据是否符合理论分布（此项目只实现对均匀分布的检测），输出 p 值作为显著性指标。

Q-Q 图对比：通过分位数-分位数图直观对比样本数据与理想分布的偏离程度。

自相关系数：计算滞后 1 阶的自相关性，低相关性表明序列无明显周期性。

游程检验：统计序列中增减趋势变化的次数，评估随机序列的波动复杂度。

生成速度测试：记录单位时间内生成的样本数量，对比不同算法的效率。

信息熵计算：量化随机数的不可预测性，实际值越接近理论最大熵表明随机性越强。

交互式 GUI 界面：基于 Tkinter 构建可视化操作面板，支持动态调整参数（如范围 N、生成样本量）并实时展示图表与文本结果，提升用户体验。

性能说明

算法选择影响性能 ：C 风格生成器（random.randint）因依赖全局状态，多线程场景下存在锁竞争问题;numpy 实现的均匀分布生成器基于 Mersenne Twister 算法，批量生成时吞吐量更高。

内存占用控制：采用生成器模式按需计算，避免一次性加载海量数据，实测 10,000 个样本内存占用约 39KB（均匀分布生成器）、78KB（正态分布生成器）和 273KB（C 风格生成器）。

边缘情况处理：正态分布生成器通过截断机制确保输出值在[0, N-1]范围内，避免极端值干扰统计结果。

模块化架构 ：随机数生成器与测评模块解耦，可灵活新增算法（如加密安全的 secrets 模块）或测试方法（如随机游走测试）。

跨平台兼容：依赖主流 Python 科学计算库（numpy、matplotlib、scipy），确保在 Windows/Linux/macOS 系统中无缝运行。

2.设计与实现方案

2.1 实现原理

软件流程：

参数输入 ：用户通过 GUI 设置数值范围（N 值）、样本量及随机数生成算法（C 风格、均匀分布、正态分布）。

触发测试 ：点击测试按钮后，系统调用 RandomnessTester 类生成指定数量的随机数样本。

数据生成与检验 ：

生成样本 ：根据选择的算法生成样本实现高效批量生成.

执行检验 ：调用测评模块。

结果展示：通过多线程异步更新 GUI 界面，避免阻塞主线程，动态渲染图表与文本结果。

相关描述：

随机数生成器选择：

通过函数指针映射实现三种算法动态切换。

正态分布生成器基于截断正态分布理论，确保输出值在[0, N-1]范围内。

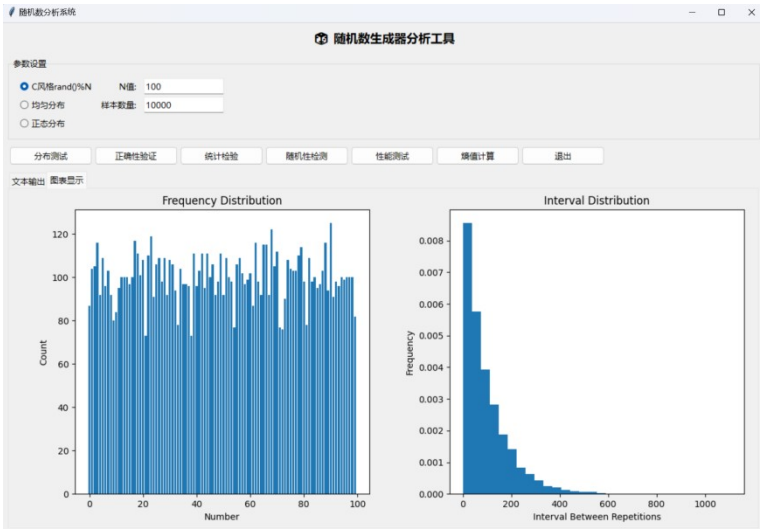
图表绘制：

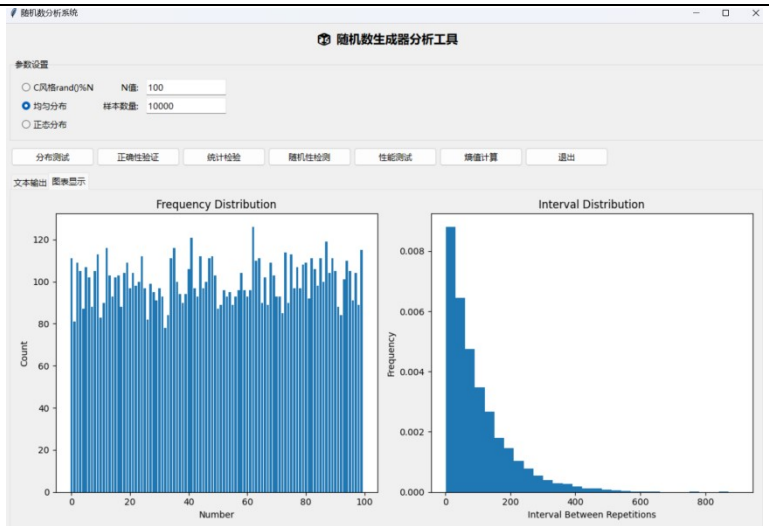
使用 matplotlib 嵌入 Tkinter 窗口，动态更新子图。

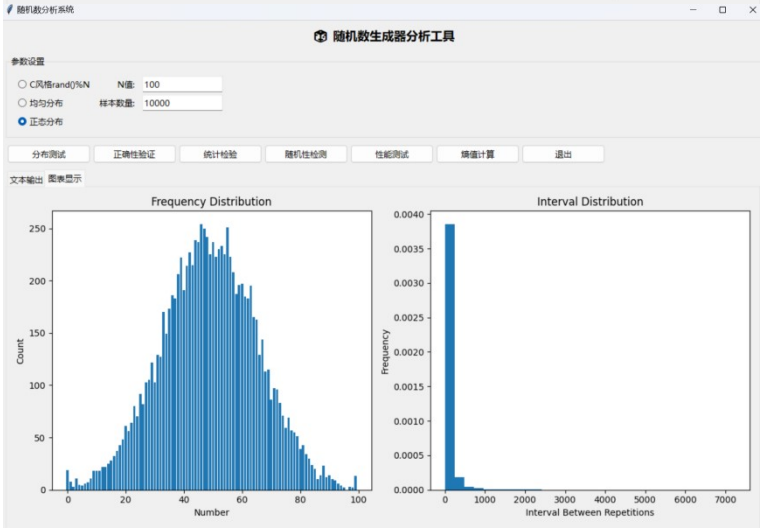
自相关图通过 acorr 函数计算并展示序列的滞后相关性。

2.2 运行结果

分布测试：

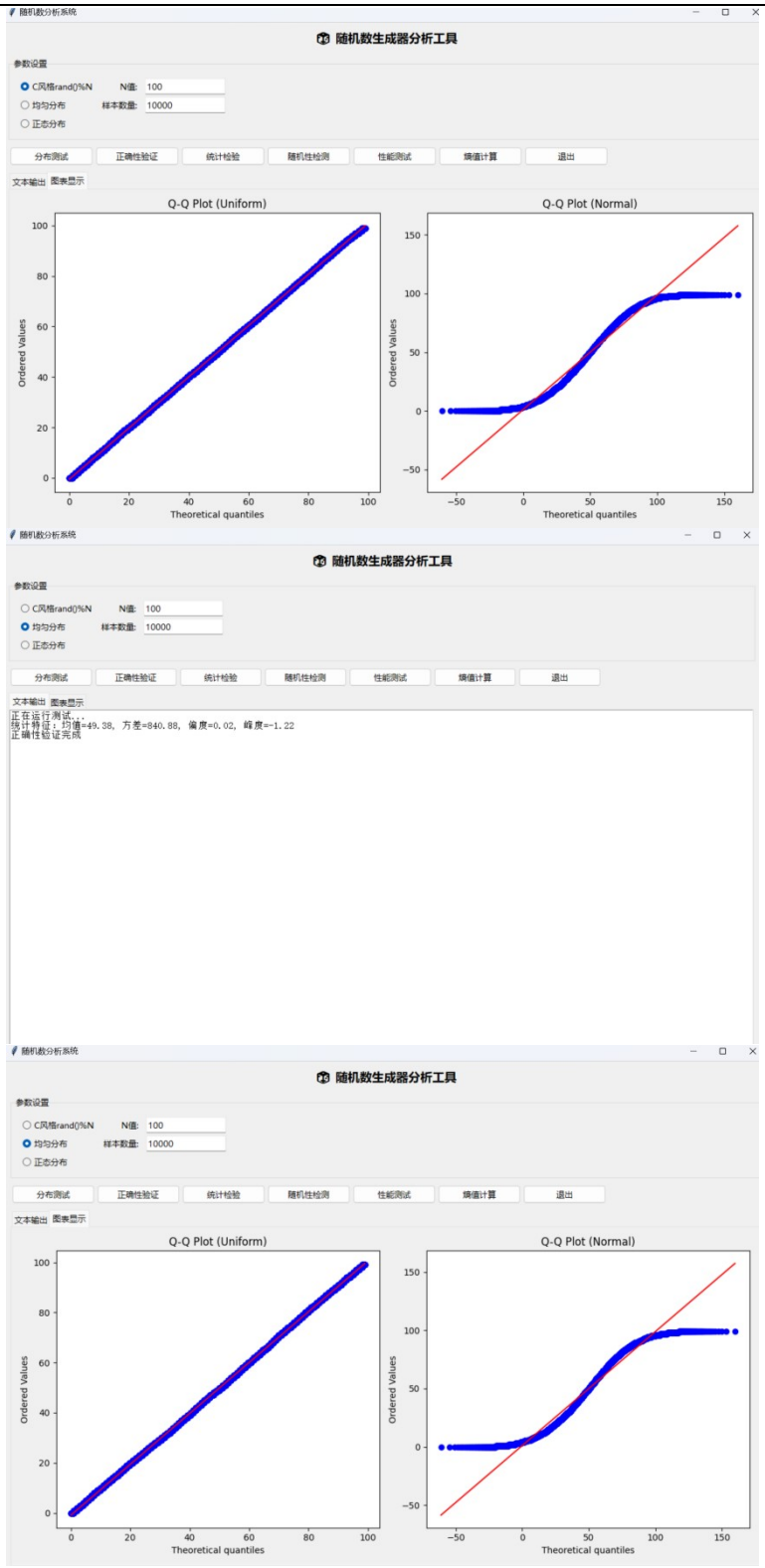



随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
均匀分布
正态分布
N值: 100
样本数量: 10000
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出
图表显示
Frequency Distribution
Count
Number
Interval Distribution
Frequency
Interval Between Repetitions

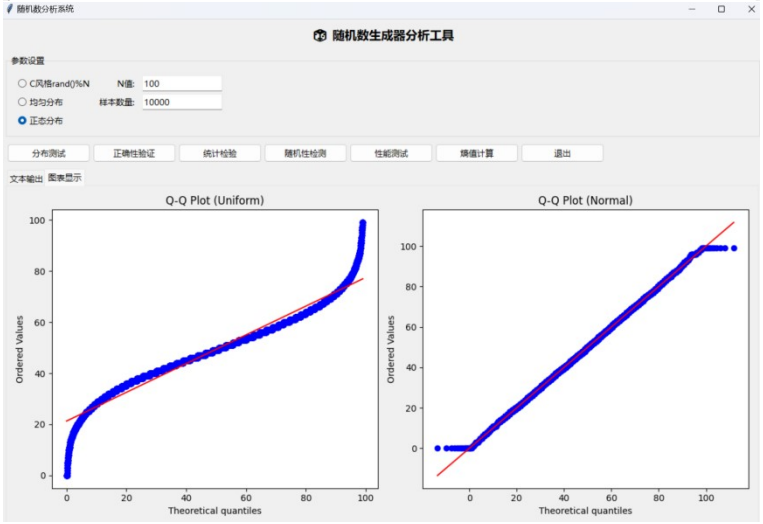

随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
均匀分布
正态分布
N值: 100
样本数量: 10000
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出
图表显示
Frequency Distribution
Count
Number
Interval Distribution
Frequency
Interval Between Repetitions

正确性验证 ：


随机数生成器分析工具
参数设置
C风格rand()%N
均匀分布
正态分布
N值: 100
样本数量: 10000
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出
图表显示
正在运行测试...
统计特征：均值=49.72，方差=840.53，偏度=0.00，峰度=-1.20
正确性验证完成


随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
均匀分布
正态分布
N值: 100
样本数量: 10000
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出
图表显示
Q-Q Plot (Uniform)
Q-Q Plot (Normal)
Ordered Values
Theoretical quantiles
正在运行测试...
统计特征：均值=49.38, 方差=840.88, 偏度=0.02, 峰度=-1.22
正确性验证完成


随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
均匀分布
正态分布
N值: 100
样本数量: 10000
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出
图表显示
正在运行测试...
统计特征：均值=49.12, 方差=269.55, 偏度=0.03, 峰度=-0.08
正确性验证完成


随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
均匀分布
正态分布
N值: 100
样本数量: 10000
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出
图表显示
Q-Q Plot (Uniform)
Q-Q Plot (Normal)
Ordered Values
Theoretical quantiles

统计检验结果：

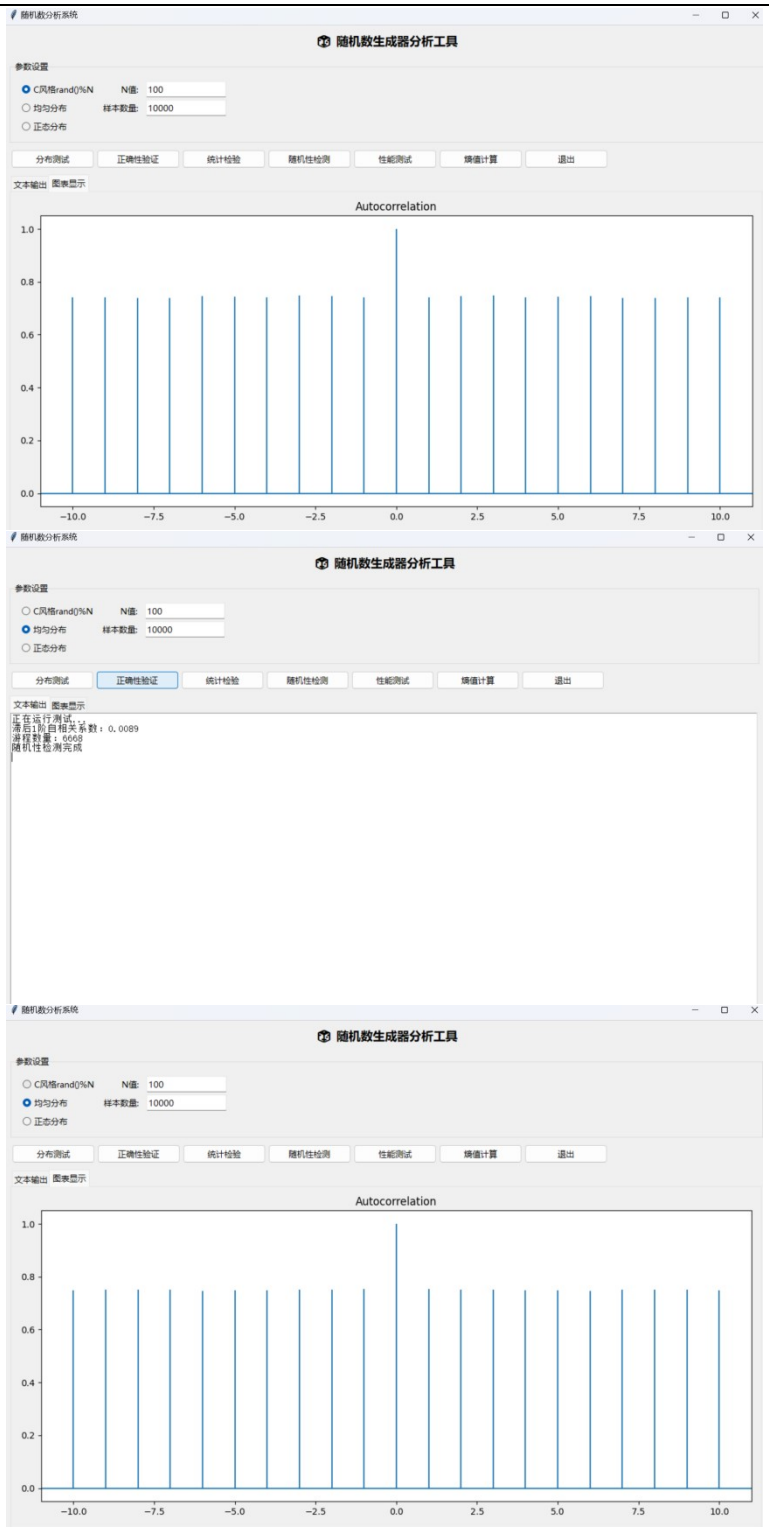

随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
均匀分布
正态分布
N值: 100
样本数量: 10000
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出
图表显示
正在运行测试...
卡方检验：χ²=86.28, p值=0.8154
KS检验：D=0.0177, p值=0.0872
统计分布检验完成


随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
均匀分布
正态分布
N值: 100
样本数量: 10000
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出 图表显示
正在运行测试...
卡方检验：X²=88.02，p值=0.7774
KS检验：D=0.0173，p值=0.1003
统计分布检验完成

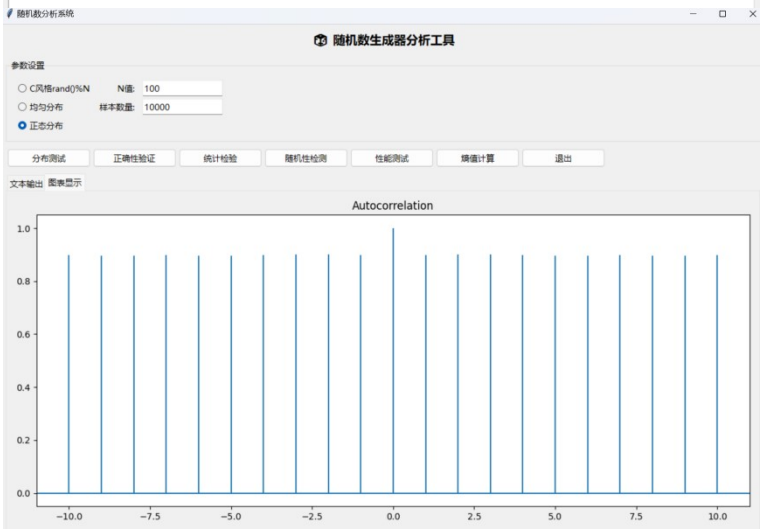

随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
均匀分布
正态分布
N值: 100
样本数量: 10000
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出 图表显示
正在运行测试...
卡方检验：X²=7257.42，p值=0.0000
KS检验：D=0.1940，p值=0.0000
统计分布检验完成

随机性检测 ：


随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
均匀分布
正态分布
N值: 100
样本数量: 10000
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出 图表显示
正在运行测试...
滞后1阶自相关系数：-0.0051
滞后数量：6666
随机性检测完成


随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
均匀分布
正态分布
N值: 100
样本数量: 10000
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出
图表显示
Autocorrelation
1.0
0.8
0.6
0.4
0.2
0.0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
均匀分布
正态分布
N值: 100
样本数量: 10000
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出
图表显示
正在运行测试...
滞后1阶自相关系数：0.0089
游程数量：6668
随机性检测完成
随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
均匀分布
正态分布
N值: 100
样本数量: 10000
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出
图表显示
Autocorrelation
1.0
0.8
0.6
0.4
0.2
0.0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0


随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
N值: 100
均匀分布
样本数量: 10000
正态分布
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出
图表显示
正在运行测试...
滞后1阶自相关系数：0.0051
游程数量：6710
随机性检测完成


随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
N值: 100
均匀分布
样本数量: 10000
正态分布
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出
图表显示
Autocorrelation
1.0
0.8
0.6
0.4
0.2
0.0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0

性能测试：


随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
N值: 100
均匀分布
样本数量: 10000
正态分布
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出
图表显示
正在运行测试...
生成速度：3059631 samples/sec
内存占用：约273.44 KB for 10000 samples
性能测试完成


随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
均匀分布
正态分布
N值: 100
样本数量: 10000
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出 图表显示
正在运行测试...
生成速度：13315207 samples/sec
内存占用：约39.06 KB for 10000 samples
性能测试完成
随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
均匀分布
正态分布
N值: 100
样本数量: 10000
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出 图表显示
正在运行测试...
生成速度：50051360 samples/sec
内存占用：约78.12 KB for 10000 samples
性能测试完成

熵值计算 ：


随机数分析系统
随机数生成器分析工具
参数设置
C风格rand()%N
均匀分布
正态分布
N值: 100
样本数量: 10000
分布测试
正确性验证
统计检验
随机性检测
性能测试
熵值计算
退出
文本输出 图表显示
正在运行测试...
信息熵：6.6367（理论最大熵：6.6439）
熵值计算完毕

2.3 技术指标

分布测试 ： 图像化表示生成样本的分布和重复出现的间隔的分布情况

正确性验证 ：计算样本均值、方差、偏度和峰度，同时绘制 Q-Q 图比对随机数样本对均匀分布和正态分布的满足情况（偏度用于衡量数据分布的不对称性 ，帮助判断生成的随机数是否符合理论分布的对称性要求；峰度用于衡量数据分布的尾部厚度与峰值尖锐程度 ，反映数据与理想分布的形态差异）

统计检验结果 ： 对样本进行卡方检验和 KS 检验（卡方检验通过比较实际观测频数与理论期望频数的差异，判断生成的随机数是否符合目标分布；KS 检验通过比较样本的累计分布函数与理论分布的差异，判断生成的随机数是否符合特定连续分布）

随机性检测 ：计算滞后 1 阶自相关系数和游程数据（滞后 1 阶自相关系数衡量当前随机数与其前一个数值之间的线性相关性，取值范围为[-1, 1]；游程是指序列中连续递增或递减的子序列数量。游程检验通过统计游程数量判断序列是否随机）

性能测试 ：计算函数的样本生成速度和内存占用

熵值计算 ：计算样本信息熵，同时比对理论最大熵（信息熵是用于量化随机数生成器输出序列不可预测性的核心指标）

3.系统测试与结果

3.1 测试方案

选取 3 种生成函数，使用 6 种工具分别进行测试，每组测试重复 5 次，取平均值以减少随机误差。N 值固定为 100，样本量固定为 10000。

分布测试 ：

C 风格随机数生成函数：观察是否存在频次异常的数值。

均匀分布随机数生成器：频率分布应均匀（各数值频次≈100），间隔分布曲线应呈指数衰减。

正态分布随机数生成器：频率分布应符合正态分布（均值≈49.5，标准差≈16.5）。

正确性验证 ：

C 风格随机数生成函数：需对比 Q-Q 图偏离程度及统计特征异常值。

均匀分布随机数生成器：对应 Q-Q 图为直线，均值≈49.5，方差≈833.3，偏度≈0，峰度≈-1.2。

正态分布随机数生成器：Q-Q 图（正态）应接近直线，均值≈49.5，标准差≈16.5，偏度≈0，峰度≈0。

统计检验结果 ：

均匀分布随机数生成器：卡方检验 p 值>0.05（接受分布一致假设），KS 检验 D<0.02。

卡方检验的 p 值越高表明分布越均匀,KS 检验的 D 值越低表明分布越匹配。

随机性检测 ：

所有生成器：自相关系数绝对值<0.05（无显著短期依赖），游程数量接近理论值（(2n-1)/3≈6666）。

性能测试 ：

理论上使用 numpy 的均匀分布随机数生成器快于正态分布随机数生成器快于 C 风格随机数生成器，并且依次占用内存逐渐增大。

熵值计算 ：

熵值与理论最大值的差距越小越好。

3.2 测试数据与结果

	均值	方差	偏度	峰度	卡方	P（卡方）	D	P（D）
C	49.436	830.652	-0.002	-1.202	93.316	0.61882	0.01486	0.24586
均匀	49.576	838.16	0	-1.202	99.668	0.47002	0.01838	0.14208
正态	48.914	269.926	0.012	-0.086	无	无	无	无

	滞后1阶自相关系数	游程数量	生成速度（samples/s）	内存占用（kb/10000samples）	信息熵（比特）
C	0.00502	6628	4894120.4	273.44	6.63674
均匀	-0.005	6669.8	124708112	39.06	6.63708
正态	0.00122	6689	57843159.8	78.12	6.07906

结果总结:

C 风格随机数生成函数和均匀分布随机数生成器的间隔分布曲线呈指数衰减；正态分布随机数生成器的间隔分布曲线与前两者相比分布更为集中。

正态分布随机数生成器的峰度绝对值小于另两者的峰度绝对值。C 风格随机数生成函数和均匀分布随机数生成器的 Q-Q 图（均匀分布）呈直线；正态分布随机数生成器的 Q-Q 图（正态分布）中间呈直线，最边缘两边平行于 x 坐标轴（受到截断影响）。C 风格随机数生成函数生成的样本基本符合均匀分布。

三者的滞后 1 阶自相关系数绝对值都<0.05;在游程数量上，均匀分布随机数生成器好于正态分布随机数生成器好于 C 风格随机数生成函数。

生成速度：均匀分布随机数生成器>正态分布随机数生成器>C 风格随机数生成函数。

内存占用：均匀分布随机数生成器<正态分布随机数生成器<C 风格随机数生成函数。

信息熵：均匀分布随机数生成器和 C 风格随机数生成函数的信息熵显著大于正态分布随机数生成器（理论最大熵 6.6439）。均匀分布随机数生成器的信息熵距理论最大熵最近。

4.应用前景

本工具可作为概率论与数理统计课程的教学演示平台，通过直观的图表（如频率分布直方图、Q-Q 图）帮助学生理解随机数分布特性。同时，其集成的统计检验框架（卡方检验、KS 检验）和熵值计算功能，可为科研人员提供基础数据验证工具，辅助分析生成模型的随机性质量。

尽管当前功能聚焦于离散分布测试，但其模块化设计为未来扩展提供了空间。例如，通过集成连续分布检验（如正态分布的 KS 检验）或引入机器学习模型预测生成器缺陷，可进一步适配多种使用场景。

目前工具尚未覆盖加密安全生成器的深度测试，且依赖用户手动解读统计指标。未来可通过增加自动化评估报告、支持多语言接口或结合云端测试平台实现远程协作，提升实用性。

5. 结论

本项目设计并实现了一款随机数统计分布测试工具，集成了 C 风格、均匀分布及正态分布随机数生成算法，并通过可视化图表（频率分布直方图、Q-Q 图）、统计检验（卡方检验、KS 检验）及熵值计算等方法，系统评估生成序列的随机性质量。测试结果表明：

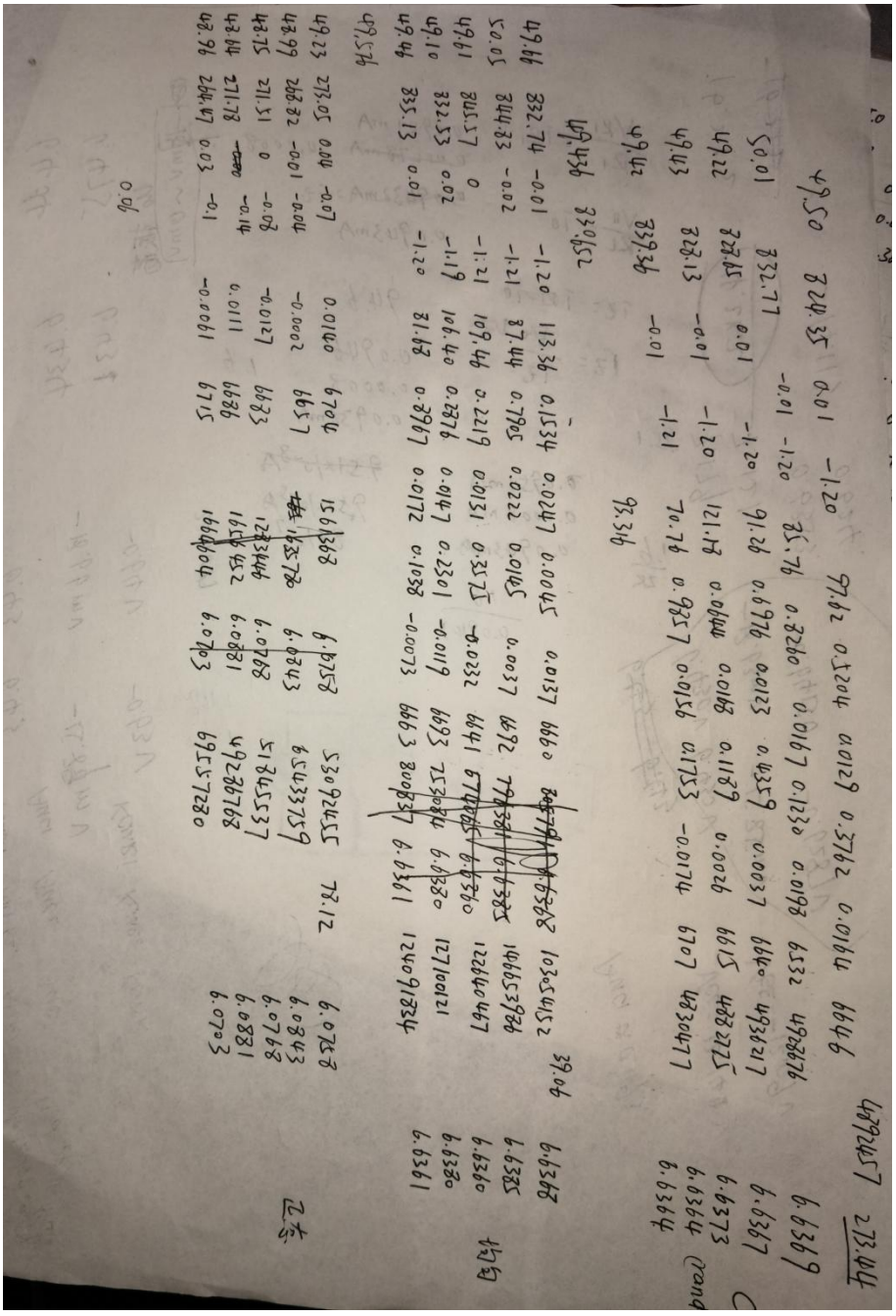
性能与适用性 ： numpy 均匀分布生成器在生成速度和信息熵上表现最优，适用于高吞吐量和高随机性场景；正态分布生成器通过截断机制有效控制数值范围，适合需特定分布的场景；C 风格生成器因全局状态依赖，在多线程场景下性能受限。

随机性验证 ： 三类生成器的自相关系数绝对值均低于 0.05，游程数量接近理论值，表明序列具备良好短期无相关性和波动复杂度。卡方检验中，均匀分

布生成器和 C 风格生成器 p 值>0.05，符合分布一致性假设。

扩展性与应用 ：模块化设计支持快速集成新算法及测试方法，可作为概率统计教学演示工具或科研基础验证平台。未来可通过自动化评估报告生成、云端协作测试等功能提升实用性，进一步适配加密安全领域需求。

附件（数据记录）



另附

项目网址：https://github.com/DEZY5545/crypto/tree/master#

未找到中科国昱密评工具箱资源，询问后可不写