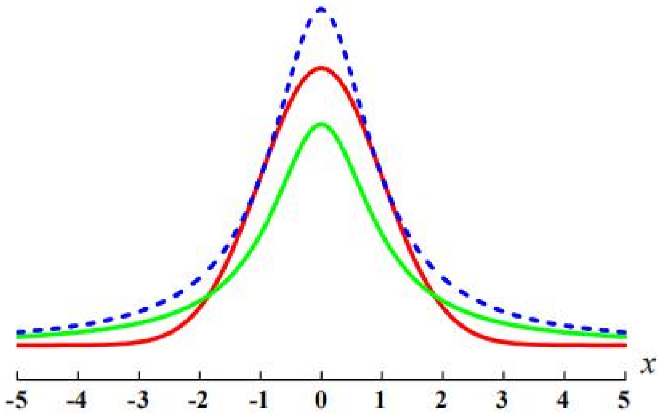
第 6 题作业报告

PB18000341 范玥瑶

A. 作业题目

对两个函数线型（Gauss 分布和类 Lorentz 型分布），设其一为 p(x)，另一为 F(x)，用舍选法对 p(x)抽样。将计算得到的归一化频数分布直方图与理论曲线 p(x)进行比较，讨论差异。讨论抽样效率。

$$Gaussian: \sim \exp\left(-\frac{x^2}{2}\right); Lorentzian\ like: \sim \frac{1}{1+x^4}$$



B. 算法及主要公式

B.1. 数学计算

$$\because \lim_{x\to\infty}\left|\frac{\exp\left(-\frac{x^2}{2}\right)}{\frac{1}{1+x^4}}\right| = 0$$

∴重要抽样取 p(x)为高斯分布，F(x)为类洛伦兹分布时，由 p(x)的归一化条件

$$p(x) = \frac{\exp\left(-\frac{x^2}{2}\right)}{\int_{-\infty}^{+\infty}\exp\left(-\frac{x^2}{2}\right)dx} = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$$

$$\because \lim_{x\to\infty}\left|\frac{p(x)}{\frac{1}{1+x^4}}\right| = 0$$

$$\therefore \left|\frac{p(x)}{\frac{1}{1+x^4}}\right| 在x \in R上有界,$$

$$\therefore 存在C = Const使得对F(x) = \frac{C}{1+x^4}, \left|\frac{p(x)}{F(x)}\right| < 1\ 在R上恒成立。$$

用 WolframAlpha 软件计算得$\max\limits_{x\in R}\frac{\exp\left(-\frac{x^2}{2}\right)}{\frac{1}{1+x^4}} \approx 2.30995$，取$C = \frac{3}{\sqrt{2\pi}}$。

假设（ξ_1，ξ_2）均为在（0，1）上均匀分布的随机数，由重要抽样方法，取随机变量（ξ_x，ξ_y）满足

$$\xi_1 = \frac{\int_{-\infty}^{\xi_x} F(x)dx}{\int_{-\infty}^{+\infty} F(x)dx} = \frac{-2\arctan[1-\sqrt{2}\xi_x] + 2\arctan[1+\sqrt{2}\xi_x] + \mathrm{Ln}[\frac{1+\sqrt{2}\xi_x+\xi_x^2}{1-\sqrt{2}\xi_x+\xi_x^2}]}{4\pi} + \frac{1}{2} (B.1)$$

$$\xi_y = \xi_2 F\left(\xi_x\right) = \xi_2 * \frac{3}{\sqrt{2\pi}} * \frac{1}{1+\xi_x^2} (B.2)$$

当$\xi_y \le p\left(\xi_x\right)$时取样本$x = \xi_x$.

由于符号计算软件无法给出$\xi_1(\xi_x)$的反函数的解析解，但是$\xi_1(\xi_x)$单调递增（函数图像如图 1），此后在程序中编写子函数采用二分法对反函数$\xi_x(\xi_1)$的数值进行求解。

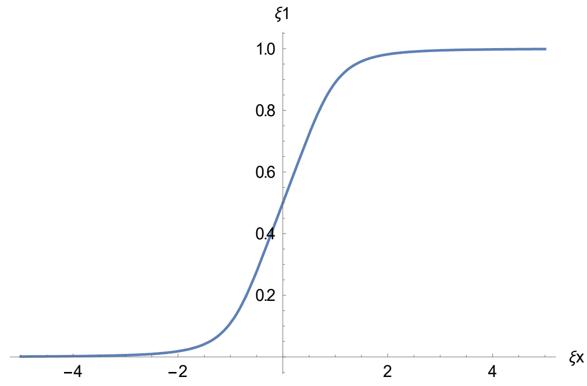


图 1ξ_1（ξ_x）函数图

ξ_x的取值范围是 R，不可能一一验证，由于生成的ξ_1精度为 1E-6，对$\xi_1 = 5*10^{-7}, 1-5*10^{-7}$用 WolframAlpha 解出$\xi_x \approx \pm 66.9511$,当$\left|\xi_x\right| > \pm 66.9511$时不同$\xi_x$对应的$\xi_1$在双精度浮点数表示中无法被区分，因此也不能从$\xi_1$变换得到$\left|\xi_x\right| > \pm 66.9511$的不同的$\xi_x$，因此不妨只考虑$\xi_x \in$（$-66.9511$，$66.9511$）的二分法计算。double 型的 5E-7 和 1-5E-7 分别输出为 0.000000，1.000000，因此$\xi_x \in (-66.9511$，$66.9511)$的抽样公式：

$$\xi_1 = \frac{\int_{-66.9511}^{\xi_x} F(x)dx}{\int_{-66.9511}^{+66.9511} F(x)dx}$$

$$\approx \frac{\int_{-\infty}^{\xi_x} F(x)dx}{\int_{-\infty}^{+\infty} F(x)dx}$$

$$= \frac{-2\arctan\left[1-\sqrt{2}\xi_x\right] + 2\arctan\left[1+\sqrt{2}\xi_x\right] + \mathrm{Ln}\left[\frac{1+\sqrt{2}\xi_x+\xi_x^2}{1-\sqrt{2}\xi_x+\xi_x^2}\right]}{4\pi} + \frac{1}{2} (B.3)$$

$$\xi_y = \xi_2 F\left(\xi_x\right) = \xi_2 * \frac{3}{\sqrt{2\pi}} * \frac{1}{1+\xi_x^2} (B.2)$$

B.2. 算法

B.2.1. 重要抽样方法

输入种子值 seed 和随机数的个数 N，调用自定义函数 RandomNumber 用 Schrage 方法得到随机数，以相邻两随机数为一个随机数对（ξ_1，ξ_2），在 RandomNumber 中调用子函数 ReverseFunction 生成ξ_x，代入抽样公式（B.1）、(B.2)得到随机数对（ξ_x，ξ_y），当$\xi_y \le p\left(\xi_x\right)$时输出抽样值$x = \xi_x$，同时样本数加 1.循环结束同时输出抽样效率$\eta = \frac{S}{N}$到屏幕和文件"Sampling1.txt"。

由于$\xi_1\left(\xi_x\right)$关于（$\frac{1}{2}$，0）中心对称，只编写$\frac{1}{2} < \xi_1 < 1$上的二分法。如果输入的ξ_1在（0，$\frac{1}{2}$）上则转换到（$\frac{1}{2}$，1）上的求解问题$\xi_1' = 1 - \xi_1, \xi_x' = ReverseFunction\left(\xi_1'\right), \xi_x = -\xi_x'$,化简得

$$\xi_x = -ReverseFunction\left(1 - \xi_1\right)$$

子函数 ReverseFunction 的算法为：二分法取样。若ξ_1对$1 > \xi_1 > \frac{1}{2}$,当$\xi_x > 2$即$\xi_1 > 0.981727$时，ξ_1随ξ_x变化比较缓慢，为提高取样效率，取样间隔会比较大，不妨在$\xi_1 > 0.981727$取ξ_x初始步长为 1，在$\frac{1}{2} < \xi_1 < 0.981727$取$\xi_x$初始步长为 0.01。找到$\xi_x$区间后缩小步长为上一步的十分之一直到求出$\xi_x$精度为 1E-6 的近似值。初始步长的值用宏表示，必要时（如希望提高计算速度时增大步长）。

B.2.2.舍选法抽样

考察

$$p(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

定义域(-66.9511，66.9511)，值域（0，0.398942）。

（x,y）为（0，1）上均匀分布的两个随机数，随机抽样值

$$X = -66.9511 + 2 * 66.9511 * x$$
$$Y = 0.398942 * y$$

当 p(X)<Y 时舍去将抽样值 x=X，否则将其输出到 sampling2.txt 里，样本数加 1.循环结束同时输出抽样效率 η=S/N 到屏幕和文件。

C. 计算结果及具体分析、讨论

C.1. 重要抽样方法

取 seed=1,N=10000 运行得到数据 Sampling1.txt，导入 Origin 作归一化频数分布直方图如图 2。抽样效率$\eta = 37.15\%$不是很高。

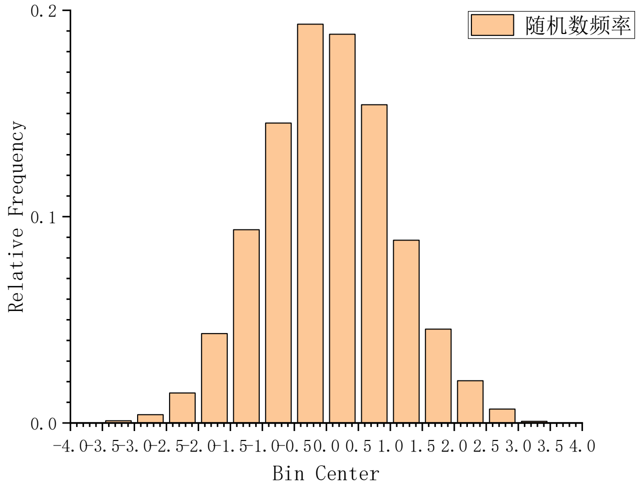


图2 N=1E4归一化频数分布直方图

取 seed=1,N=1E7 频率分布直方图与理论曲线$p(x)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$比较如图 3。

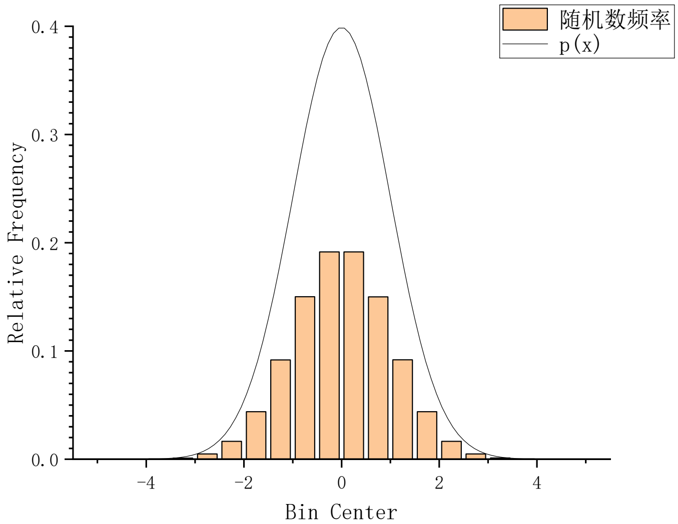


图3 重要抽样方法随机数归一化频数分布直方图与p(x)比较

C.2.舍选法抽样

取 seed=1,N=10000 运行得到数据 sampling2.txt，导入 Origin 作归一化频数分布直方图如图 4。抽样效率$\eta' = 2.03\% \ll \eta$，重要抽样方法的抽样效率比较高。

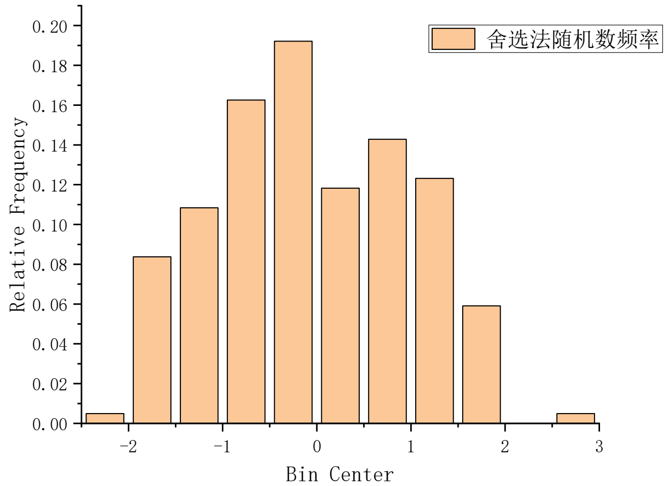


图4 直接舍选法N=1E4归一化频数分布直方图

取 seed=1,N=1E7 频率分布直方图与理论曲线$p(x)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$比较如图 5。

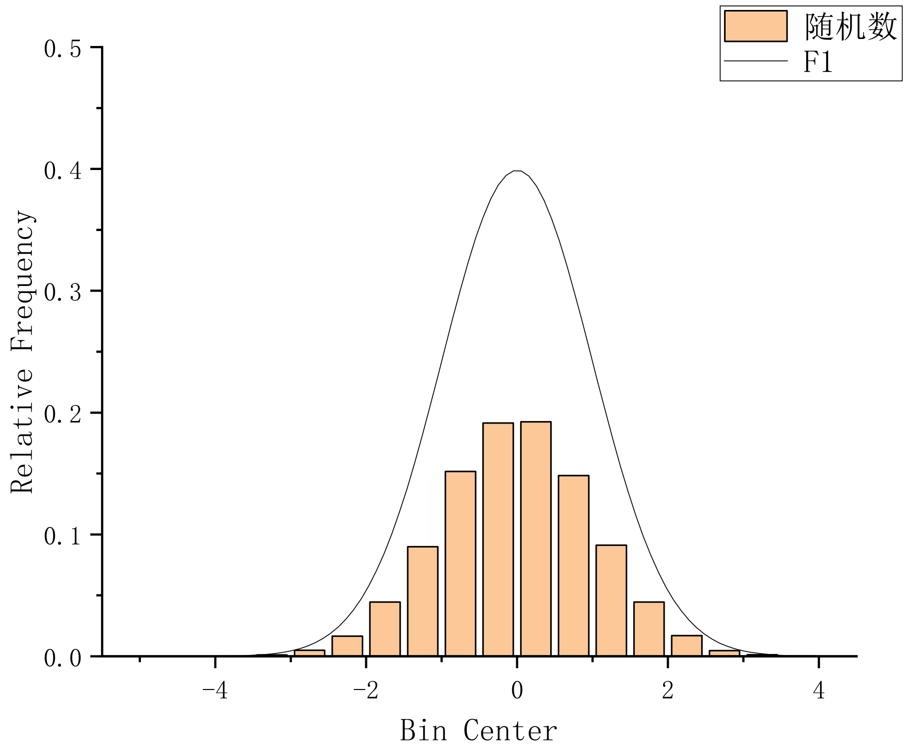


图5 舍选法归一化频数分布直方图与理论曲线p(x)比较

D. 总结

本次作业中通过数学的推到与编写程序对$p(x)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$进行了重要抽样和舍选法抽样，将计算得到的归一化频数分布直方图与理论曲线 p(x)进行比较，基本一致而且 N 比较大时更接近 p(x)。直观上重要抽样方法的抽样效率比直接使用舍选法更高，运行数据也与

结论相符。