作业第六题报告

PB18020539 黄韫飞

1.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: $\sim\exp(-\frac{x^2}{2})$，Lorentz like:$\sim\frac{1}{1+x^4}$

2.公式与算法

取高斯分布为$p(x)$，类Lorentz分布为$F(x)$。讨论的范围取[-5,5]，将$p(x)$归一化：

$$\int_{-5}^{5}\exp(-x^2/2)\mathrm{d}x=2.506627\quad\Rightarrow\quad p(x)=0.398943\exp(-x^2/2)$$

要使$F(x)$在[-5,5]范围内始终大于$p(x)$，设$F(x)=\frac{B}{1+x^4}$，为了使抽样效率尽可能高，B要尽可能小，经检验$B=0.95$满足条件，可能存在比0.95更小的数满足条件，但是0.95已经足够进行较高效率的抽样。

将舍选抽样与变换法结合：

先求出$F(x)$的分布函数$G(x)$：

$$G(x)=\int_{-5}^{x}\frac{B}{1+x^4}\mathrm{d}x=B(\Phi(x)-\Phi(-5))$$
$$\Phi(x)=\frac{1}{4\sqrt{2}}\ln\frac{x^2+\sqrt{2}x+1}{x^2-\sqrt{2}x+1}+\frac{1}{2\sqrt{2}}(\arctan(1+\sqrt{2}x)-\arctan(1-\sqrt{2}x))$$

从上式可知，$G(x)$单调递增，但是无法将其反函数解析地写出，在程序中编写了一个GInvFunc()函数来求解随机数对应的x值，算法是二分法。

抽样：设ξ_1,ξ_2是[0,1]上的均匀随机数

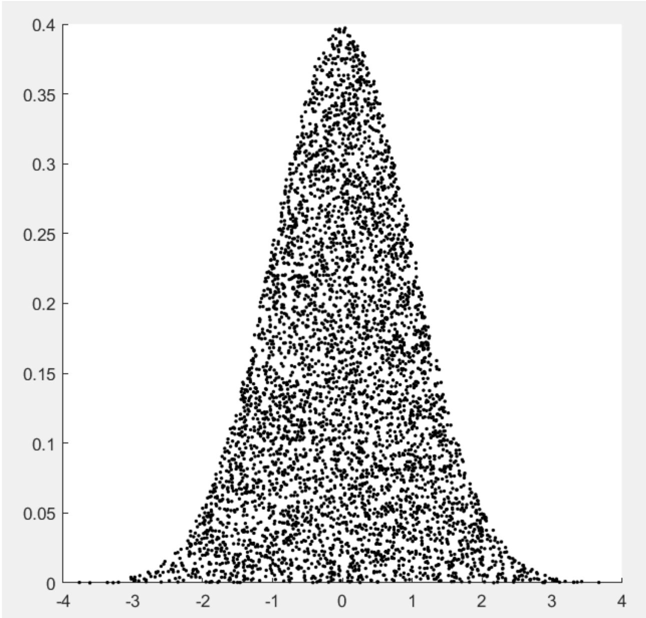
$$\xi_1=\int_{-5}^{\xi_x}F(x)\mathrm{d}x/\int_{-5}^{5}F(x)\mathrm{d}x,\xi_y=\xi_2F(\xi_x)$$

如果点(ξ_x,ξ_y)满足$\xi_y<p(\xi_x)$，则取该点，否则舍去该点。取的ξ_x即为$p(x)$的随机抽样

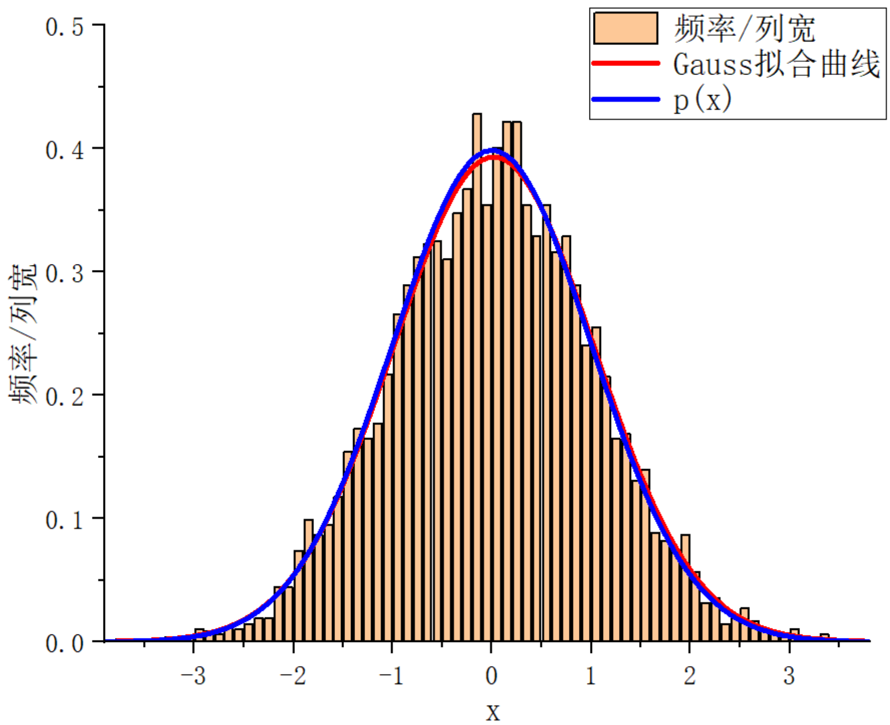
3.计算结果与讨论

取$N=10^4$

计算得到的(ξ_x,ξ_y)做出散点图，可以看出分布与正态分布比较符合。



将得到的ξ_x导入ORIGIN2018，以0.1为步长，统计ξ_x在各区间内的频率，画出频率分布直方图。并用Gauss拟合$y=y_0+\frac{A}{\omega\sqrt{\pi/2}}\exp(-2(\frac{x-x_c}{\omega})^2)$，得到$y=0.000278+0.393080\exp(-\frac{(x-0.01933)^2}{2.05191})$，把理论曲线与拟合曲线画在同一张图。拟合曲线与$p(x)=0.398943\exp(-x^2/2)$符合的较好，但是在部分区域仍然可以看到差别，这可能是由于列宽太大或者点数不足导致的。



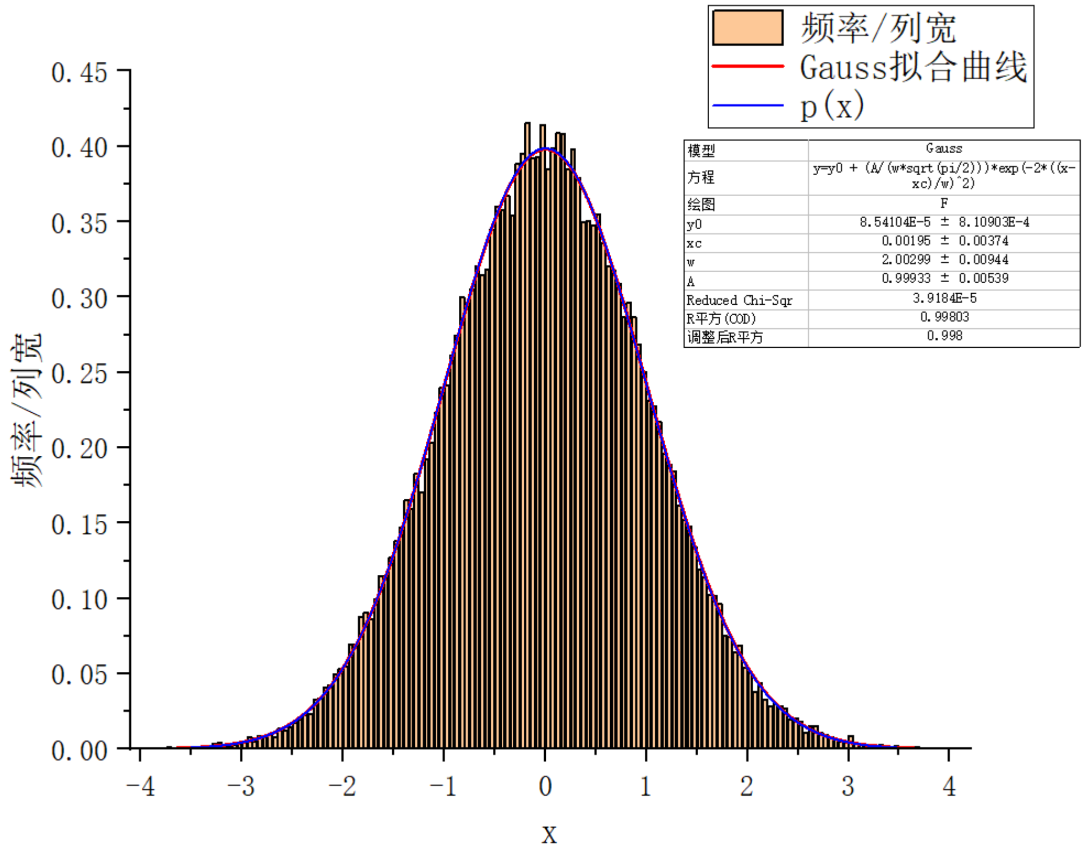
10000个点中有效点为4741个，抽样效率为0.4741

产生10000组，有效4741组，效率:0.474100

取$N=10^5$并减小直方图的列宽，做出直方图。拟合曲线为$y=0.000085+0.398080\exp(-\frac{(x-0.00195)^2}{2.005984})$，与理论曲线在图上几乎完全重合。抽样效率为0.474560

Microsoft Visual Studio 调试控制台

产生100000组，有效47456组，效率:0.474560



4.结论

以类Lorentz型分布为$F(x)$对正态分布函数进行舍选抽样的效果比较好，当点数比较多(10^4以上)时，抽取的ξ_x的分布与理论曲线的分布差别很小。但是因为Lorentz型函数与正态分布函数在$x=0$附近的差别仍然很大，所以抽样效率只能达到47.4%左右，如果使用性质更好的$F(x)$，可能会有更高的抽样效率。而且$1/(1+x^4)$积分得到的$G(x)$的反函数不好求，采用二分法使得程序运行时间增加。（猜测$\frac{1}{1+x^2}$的性质可能会更好，因为$1/(1+x^2)$的形状与正态分布更接近，并且积分得到的arctan函数的反函数比较好求，所以不管是在运算时间还是在抽样效率上都应该会更高）