计算物理 作业报告9

PB14203209 张静宁 2017.11.06

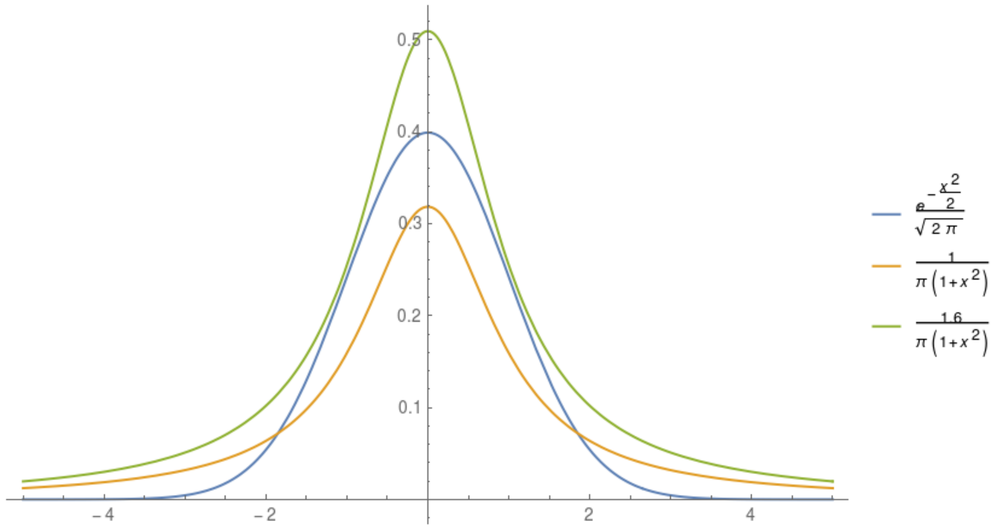
第九题

对两个函数线型（Gauss 分布和 Lorentz 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.

$Gauss: \frac{1}{\sqrt{2\pi}}exp(-\frac{x^2}{2})$; $Lorentz: \frac{1}{\pi(1+x^2)}$

公式推导和算法

设所要实现的抽样是 $p(x)=\frac{1}{\sqrt{2\pi}}exp(-\frac{x^2}{2})$，比较函数为$F(x)=\frac{1.6}{\pi(1+x^2)}$，这里把 $Lorentz$ 分布乘以1.6，使得 $F(x)$ 曲线可以把 $p(x)$ 盖住. 如下图



变换抽样

引入变换抽样，为了提高抽样效率.

比较函数 $F(x)$ 曲线与 x 轴所围成的封闭区域($x\in[-\infty,+\infty]$)，其面积为$S=\int_{-\infty}^{+\infty}F(x)dx=1.6$.

在该区域产生随机点 (ξ_x,ξ_y) 的方法如下：

$\xi_1=\int_{-\infty}^{\xi_x}F(x)dx/\int_{-\infty}^{+\infty}F(x)dx=\frac{1}{2}+\frac{Arctan(\xi_x)}{\pi}$，即 $\xi_x=tan(\pi(\xi_1-\frac{1}{2}))$,

$\xi_y=\xi_2F(\xi_x)$,

ξ_1,ξ_2为$[0,1]$区间均匀随机取样.

通过以上做法，使得ξ_x满足$F(\xi_x)$分布，而ξ_y满足$[0,F(\xi_x)]$的均匀分布，产生的随机点(ξ_x,ξ_y)处在 xdx 条带区域的概率为 $F(x)dx$，处在 ydy 条带区域的概率为 $\frac{1}{F(x)}dy$，则点处在 (x,y)点上面积元$dxdy$区域的概率为 $F(x)dx*\frac{1}{F(x)}dy=dxdy$.

即单位面积上(ξ_x,ξ_y)点分布仍然是均匀的，且全部在$F(x)$曲线的面积内.

舍选法

在以上变换抽样后，采用舍选法，具体步骤是：

（1）在比较函数内产生随机点 (ξ_x, ξ_y)

（2）如果该点在$p(x)$的面积区域内，即$\xi_y < p(\xi_x)$，则取 ξ_x 为$p(x)$的随机抽样

（3）否则，舍弃该点，返回（1）.

抽样效率

比较函数的具体形式不影响抽样的准确性，但抽样效率即有效的点数等于$p(x)$与$F(x)$的面积比.

理论值 $\eta = \frac{N_{accepted}}{N_{all}} = \int_{-\infty}^{+\infty} p(x)dx / \int_{-\infty}^{+\infty} F(x)dx = 1/1.6 = 0.625$，其中$\int_{-\infty}^{+\infty} p(x)dx = 1$

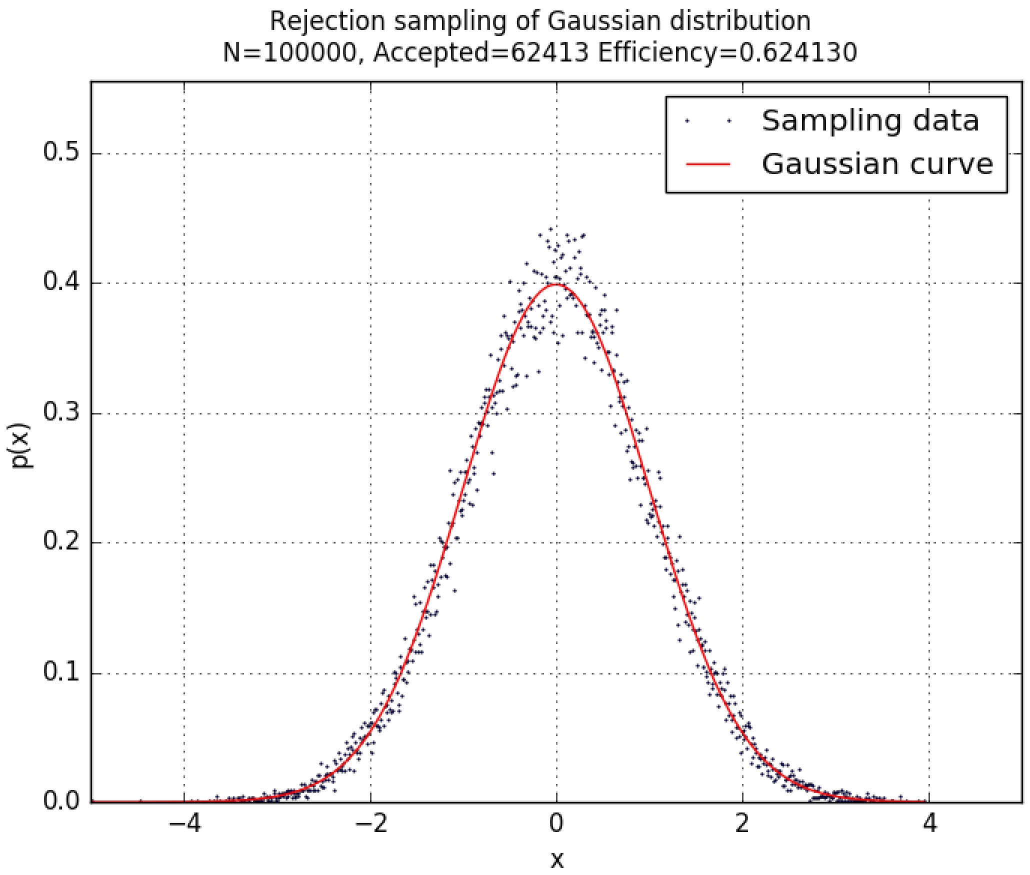
程序说明

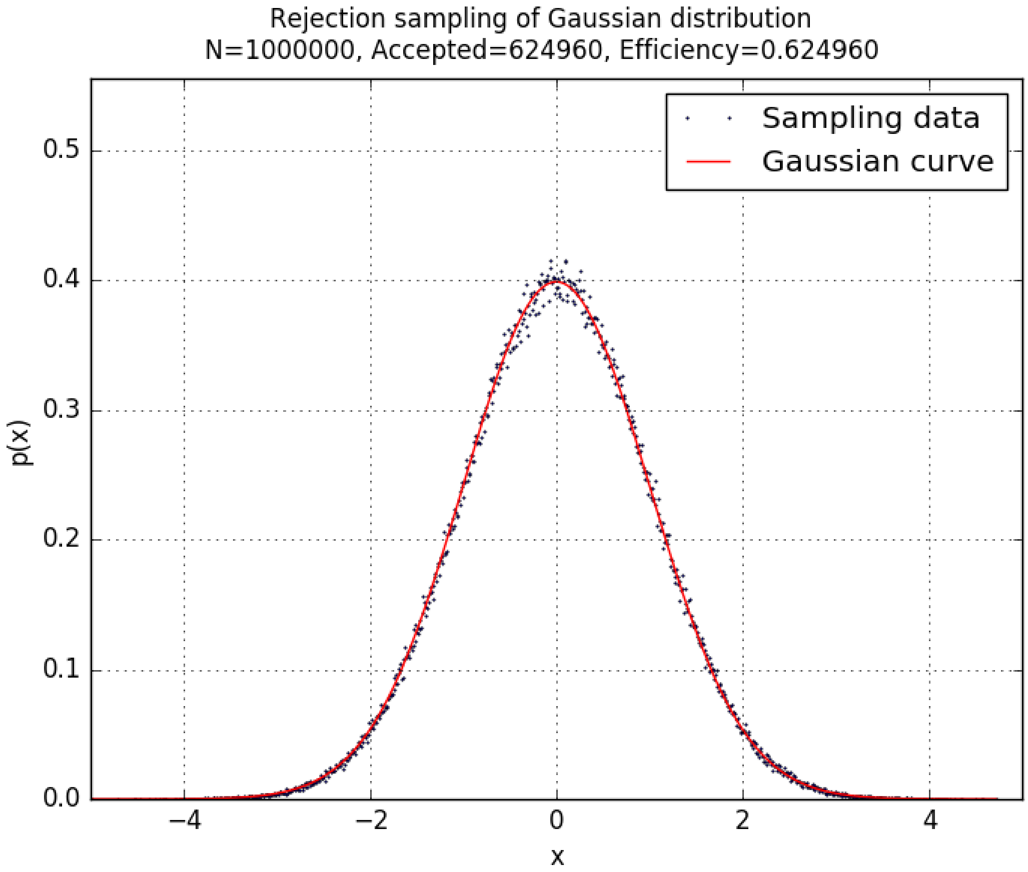
- rejection.c
  本程序采用16807随机数生成器，用舍选法实现高斯分布抽样，输出为有效的抽样值、总的抽样个数和有效抽样个数
- rejection.out
  为Linux下的可执行文件

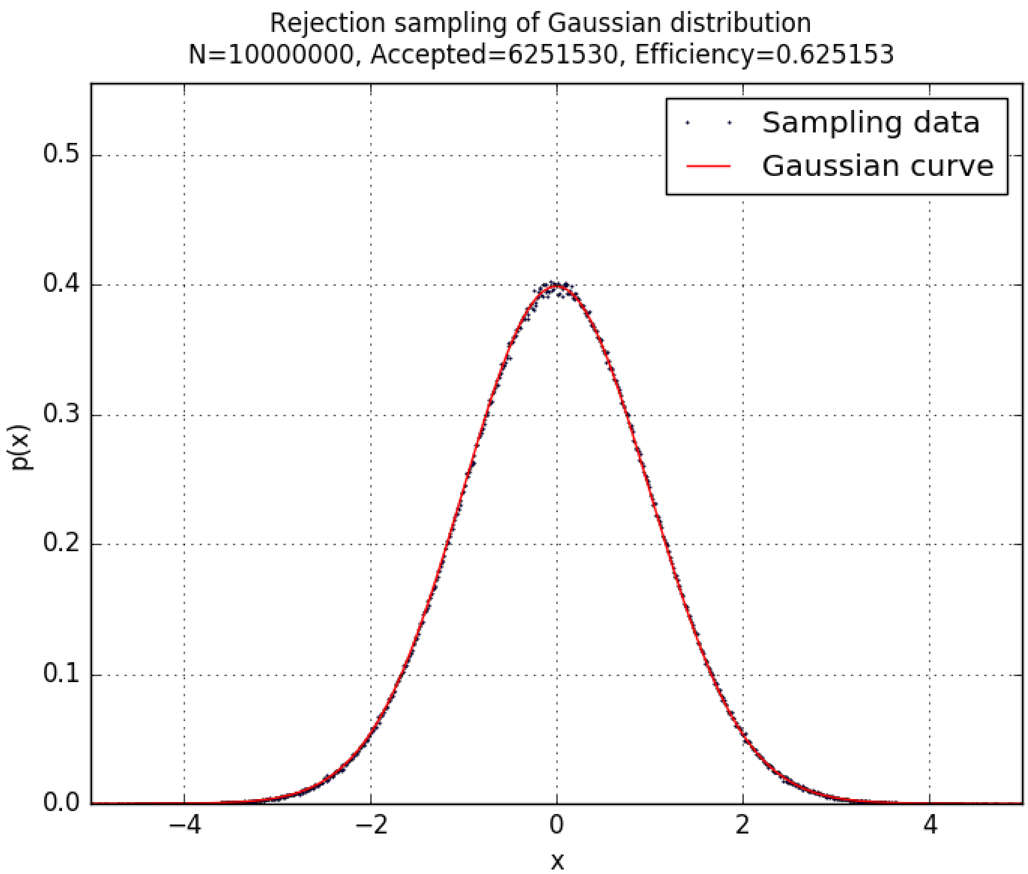
计算结果

计算了粒子总数分别为$N = 10^5, 10^6, 10^7$的情况下舍选法抽样的结果，得到归一化散点图如下：

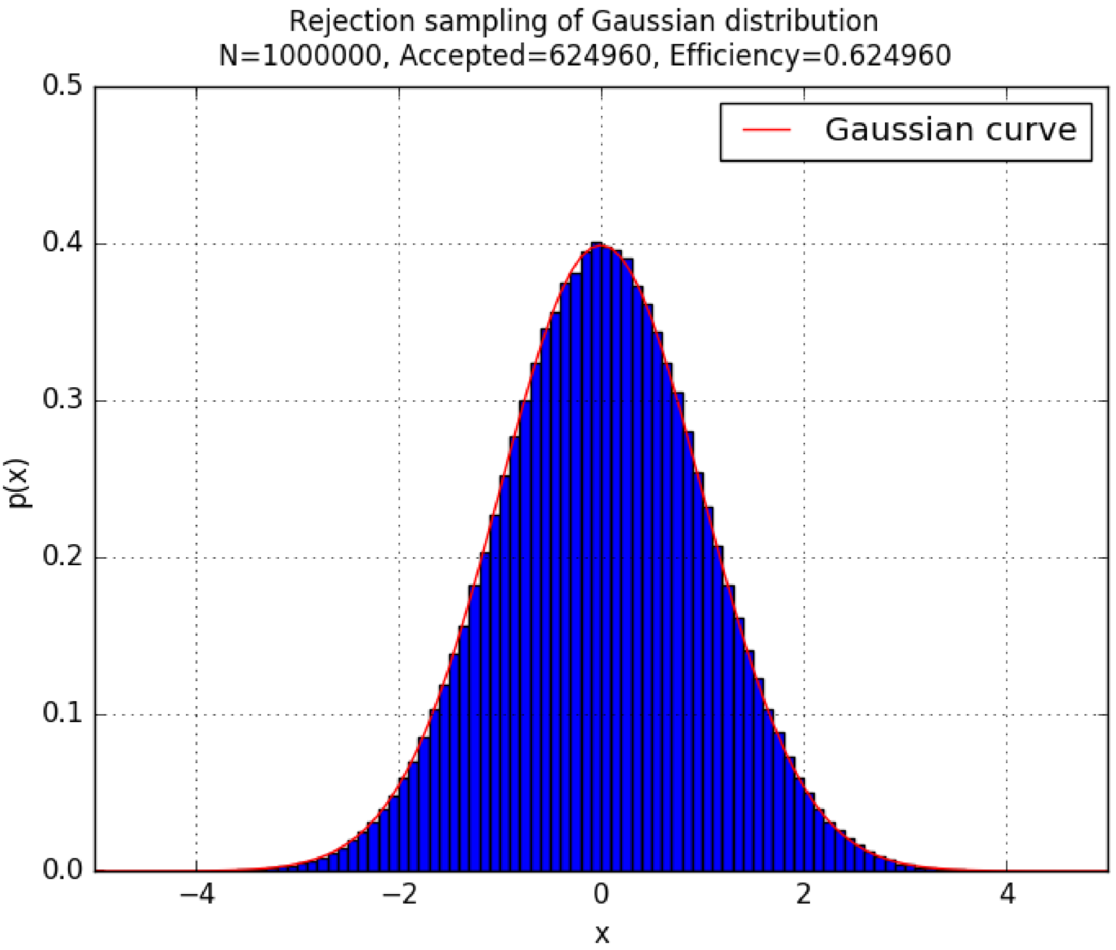
(图中的每一个数据点横坐标为抽样 x，间隔为 0.01，纵坐标为 x 出现的频数乘以100.)







得到$N = 10^6$时归一化的频数分布直方图如下



分析、总结

1. 计算得到的归一化的频数分布直方图与理论曲线$p(x)$较吻合
2. 随着抽样点数N数量级的增加，抽样值x与对应的归一化频数越趋向于理论曲线$p(x)$
3. 实际抽样效率分别为0.624130，0.624960，0.625153与理论值0.625都非常接近

实际上抽样效率的限制来自于比较函数的选取.

设 $F(x)=\frac{C}{\pi(1+x^2)}$，有$\int_{-\infty}^{+\infty}F(x)dx=C$，$\int_{-\infty}^{+\infty}p(x)dx=1$

为了使$F(x)$覆盖住$p(x)$，要求常数$C\geq1.5$，则抽样效率理论值

$\eta=\int_{-\infty}^{+\infty}p(x)dx/\int_{-\infty}^{+\infty}F(x)dx=1/C\leq0.6536$

可见，选取一个合适的比较函数形式，对抽样效率的提高十分重要。对于高斯分布的舍选法抽样实现，0.6536的抽样效率已经算是比较高了。