REPORT 6

潘硕 PB24020526

2025 年 9 月 28 日

1 Question

已知 Gauss 分布 $exp(-ax^2)$, 类 Lorentz 型分布 $(1+bx^4)^{-1}$. 将这两个函数其一作为 $p(x)$, 另一为 $F(x)$, 其中常数 $a \neq b \neq 1$, 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.

2 Method

2.1 舍选抽样法

用舍选抽样法对样本进行抽样. 操作步骤如下:

1. 产生一对 $[0,1]$ 区间均匀分布的随机抽样值 (ξ_1, ξ_2)
2. 由已知的 $p(x), F(x)$ 得到取样点 (ξ_x, ξ_y), 满足:

$$\xi_1 = \int_a^{\xi_x} F(x)dx, \xi_y = \xi_2 F(\xi_x) \tag{1}$$

3. 若 $\xi_y < p(\xi_x)$, 则取得该点, 否则舍去.

2.2 函数, 参数选取

此问中,Gauss 函数和类 lorente 函数作为 $F(x)$, 在舍选法的步骤 2 很难写出 ξ_x 关于 ξ_1 的反函数关系, 故需使用另外的手段得到 $\xi_x(\xi_1)$。

对 Gauss 函数, 已经有较为成熟的方法进行采样, 如 Box-Muller 法, Marsaglia 法, 故此处取 $F(x)$ 为 Gauss 函数,$p(x)$ 为类 Lorentz 函数, 即:

$$p(x) = \frac{\sqrt{2}b^{1/4}}{\pi(1+bx^4)} \tag{2}$$

$$F(x) = \beta e^{-ax^2} \tag{3}$$

其中 a,β 为待定常数, 且有 $a,\beta,b>0$.

注意到当 x 时足够大时,$F(x)<p(x)$, 会对抽样结果带来误差, 可以通过选取合适的 a,β 减少该误差.

先考虑 $b=1$ 的情况, 此时有:

$$p(x)=\frac{\sqrt{2}}{\pi(1+x^4)} \tag{4}$$

计算采样点落在 $[-5,5]$ 区间的概率:

$$\frac{\int_{-5}^{5}p(x)dx}{\int_{-\infty}^{\infty}p(x)dx}=\frac{\int_{-5}^{5}\frac{1}{1+x^4}}{\int_{-\infty}^{\infty}\frac{1}{1+x^4}}\approx 0.9976 \tag{5}$$

故选定此时的 $a=a_0,\beta=\beta_0$, 使在 $[-5,5]$ 区间内,$F(x)\geqslant p(x)$ 恒成立.

抽样效率与 β_0 负相关, 故应在满足条件的情况下尽可能减少 β_0 的值.

此处可以取定 $\beta_0=1.6,a=0.258$, 绘制 $F(x),p(x)$ 图像如下所示, 在 $[-5,5]$ 区间满足条件

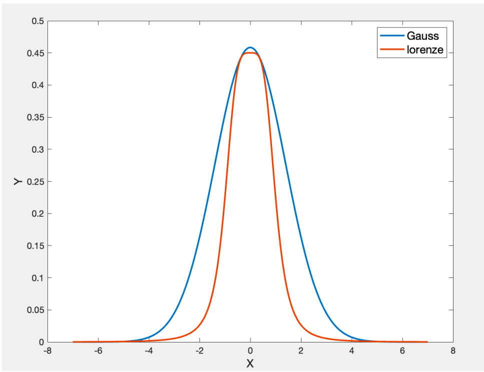

图 1: $\beta_0=1.6,a=0.258,b=1$ 时 $F(x),p(x)$ 曲线图

再考虑 $b\neq 1$ 的一般情况, 可做参量变换 $b^{1/4}x=t$ 进行转换. 对 $\forall b>0$ 时 a,β 的关系式:

$$\beta(b)=\beta_0,a(b)=a_0\sqrt{b} \tag{6}$$

3 Experiment

3.1 参数选取

此处使用 Marsaglia 法获取一维方向的 Gauss 分布, 概率密度函数 $\backsim e^{-ax^2}$. 将此作为采样点的横坐标 ξ_1, 具体代码实现封装于函数:

```
Random_Gauss_generator.m
```

再使用 16807 产生器生成 $[0,1]$ 区间均匀分布的随机抽样值, 作为采样点的纵坐标 ξ_2. 具体代码实现封装于函数:

`Random_generator_16807.m`

对每个采样点 (ξ_1, ξ_2), 取 $\xi_x = \xi_1, \xi_y = \xi_2 F(\xi_1)$, 得到在 Gauss 函数下均匀分布的随机数点, 如下图所示:

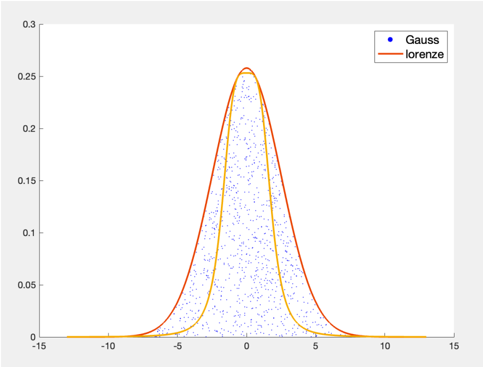

图 2: Gauss 分布采样示意图

对每个采样点, 若满足 $\xi_y \leqslant p(\xi_x)$, 取得该点; 若 $\xi_y > p(\xi_x)$, 则舍去. 取不同的 b 值进行采样, 计算抽样效率.

3.2 分析

取 $b = 0.1$ 及 $b = 10$, 做采样如下图所示:

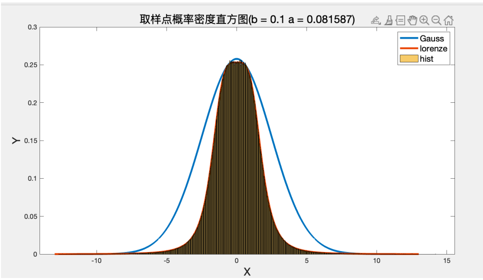

图 3: $b = 0.1$ 时采样点直方分布图

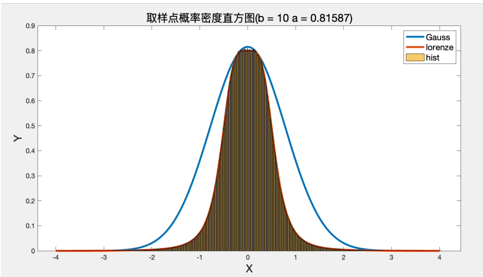

图 4: $b = 10$ 时采样点直方分布图

将抽样结果与原概率密度函数 $p(x)$ 做对比, 可见抽样结果较好, 在 $|x|$ 较大的地方抽样结果偏小, 但是该处点分布较少, 误差较小, 与理论预测相符,

接下来是对抽样效率的分析, 需要考虑两部分:

1. Marsaglia 法抽样 Gauss 分布:$\eta_1 = \pi/4 \approx 0.7854$

2. 舍取法抽样得到类 lorentz 分布:$\eta_2 \approx 1/\beta = 0.6250$

总抽样效率 $\eta = \eta_1\eta_2 \approx 0.4909$

对 $b = 0.1$ 的某次抽样:

```
Sample_first : 0.7854
Sample_second : 0.62389
Sample_total : 0.48994
```

图 5: 抽样效率计算图

与理论符合较好.

4 Summary

本实验使用舍取法, 以 Gauss 分布作为 $F(x)$, 得到满足类 lorentz 的抽样分布, 并对抽样效率做了计算. 该方法效率较高, 但在 x 较大处可能会出现偏差. 若要减少偏差, 可以调节 a, β 的值, 但也会牺牲掉一定的抽样效率.