

Hw4

Rainzor

1 Question

设 pdf 函数满足如下关系式

$$p'(x) = a\delta(x) + b\exp(-cx), \quad x \in [-1, 1], a \neq 0$$

讨论该函数的性质并给出抽样方法

2 Algorithm

2.1 表达式

对一维随机变量, 设 ξ 是 $[0, 1]$ 上的均匀分布, 对 x 的分布函数 $p(x)$ 满足

$$p'(x) = a\delta(x) + b\exp(-cx), x \in [-1, 1], a \neq 0$$
$$p(x) = \int_{-\infty}^x p'(t) dt$$

由于 $\delta(x)$ 积分后得到的阶梯函数 $H(x)$, 在 $x=0$ 处不连续, 所以对 $p(x)$ 分段处理

$$p(x) = \begin{cases} \frac{b}{c}(\exp c - \exp(-cx)) & -1 \leq x < 0 \\ a + \frac{b}{c}(\exp c - \exp(-cx)) & 0 \leq x \leq 1 \end{cases}$$

由概率密度要求: 非负且积分归一。且因为 $p(x)$ 分段单调, 那么分别有

1. $p(-1) = 0 \geq 0$
2. $p(0-) = b/c * (\exp c - 1) \geq 0$
3. $p(0+) = a + b/c * (\exp c - 1) \geq 0$
4. $p(1) = a + b/c * (\exp c - \exp(-c)) \geq 0$
5. $P(x) = \int p(x) dx = (\int_{-1}^0 + \int_0^1) p(x) dx = a + \frac{2b(c \exp c - \sinh c)}{c^2} = 1$

由上面的条件可得, 一个关于 a, b, c 的充分条件

$$a = 1 - \frac{2b}{c}e^c + \frac{2b}{c^2}\sinh c$$
$$b > 0$$
$$c > 0$$

$$1 - \frac{b}{c}(1 + \exp c) + \frac{2b}{c^2}\sinh c \geq 0$$

$$1 - \frac{2b}{c}(\cosh c) + \frac{2b}{c^2}\sinh c \geq 0$$

可以取

$$b = 0.5, c = 0.5, a = 1 - \frac{2b \exp(-c)}{c} \approx -0.213061$$

2.2 舍选法

考虑到 $p(x)$ 的不连续性，所以无法写出累积函数的解析表达式，采取舍选法：

1. 设 $p(x)$ 在 $[-1, 0)$ 上界 $p(0-) < M_1 = 0.65$ ； $[0, 1]$ 上界 $p(1) < M_2 = 0.85$
2. 随机选择两个在 $[0, 1]$ 之间均匀分布的随机抽样 (ξ_1, ξ_2) ,
3. 注意由于 M_1, M_2 不同，在 $[-1, 0]$ 与 $(0, 1]$ 区间内随机数比例不同，满足 $M_1 : M_2$
4. 判断条件：

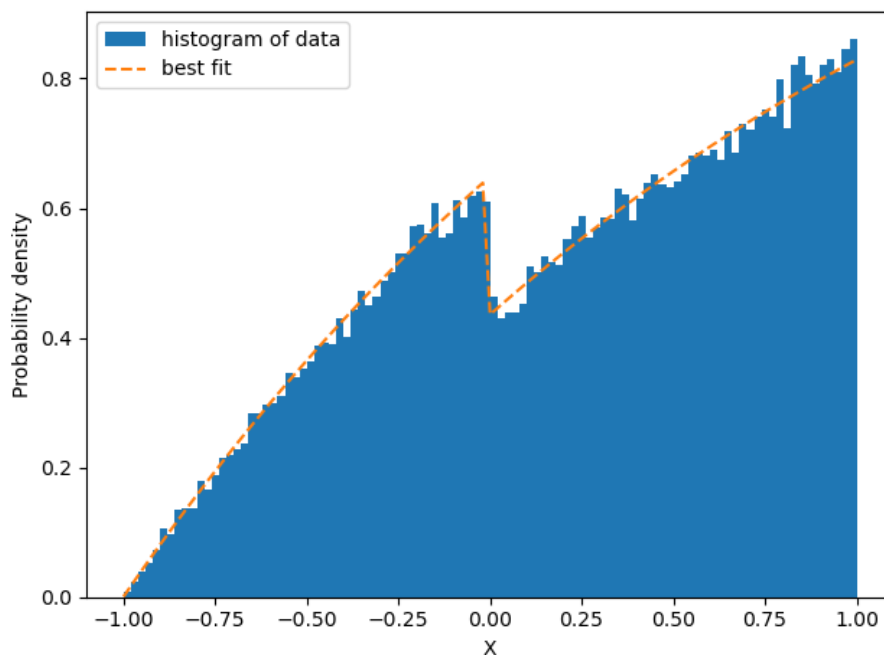
$$\begin{aligned} M_1 \xi_2 &\leq p(-1 + 2\xi_1) & -1 + 2\xi_1 &\in [-1, 0) \\ M_2 \xi_2 &\leq p(-1 + 2\xi_1) & -1 + 2\xi_1 &\in [0, 1] \end{aligned}$$

是否成立

5. 否，则舍去；是，则取 $x = -1 + 2\xi_1$

3 Experiment

在实验中定义 pdf 作为定义的函数，抽样 $1e5$ 个数据点，且在不同区间按照一定比例分布，于是画出直方图与概率分布函数：



由图可见，这样的参数选取，通过舍选抽样法得到的抽样频数分布和密度分布函数性质相近，可见所得抽样确实服从 pdf 函数所代表的分布。

Summary

本实验主要对概率密度函数性质进行熟悉，同时掌握了对于复杂分布的抽样方法，即舍选法。并且利用分段的表示，提高了抽样效率。