

HW3 第 7 题 Data 文件抽样报告

王启骅 PB20020580

2022 年 10 月 6 日

1 题目

对一个实验谱数值曲线 $p(x)$ ，自设 $F(x)$ ，分别用直接抽样和舍选法对 $p(x)$ 抽样。比较原曲线和抽样得到的曲线以验证。讨论抽样效率。

2 算法原理

2.1 舍选法

观察 data 数据，决定取分段阶梯函数作为抽样 $F(x)$ 曲线。首先将 data 数据归一化为概率分布数据，即每一能量点数据数除以总点数。取

$$F(x) = \begin{cases} 0.015 & , 2900 \leq x \leq 2993 \text{ or } 3005 \leq x \leq 3013 \\ 0.1 & , 2993 < x < 3005 \end{cases} \quad (1)$$

取 ξ_1, ξ_2 为 $[0,1]$ 均匀分布的随机数列

$$\xi_1 = \frac{\int_{2900}^{\xi_x} F(x) dx}{\int_{2900}^{3013} F(x) dx} \quad (2)$$

反解可得

$$\xi_x = \begin{cases} \frac{2.715}{0.015} \xi_1 + 2900 & , [0, 93/181] \\ \frac{2.715\xi_1 - 2.595}{0.1} + 2993 & , (93/181, 173/181) \\ \frac{2.715\xi_1 - 2.595}{0.015} + 3005 & , [173/181, 1] \end{cases} \quad (3)$$

$$\xi_y = \xi_2 F(x) \quad (4)$$

判断若有 $\xi_y < p(\xi_x)$ 则取 $x = \xi_x$ 。这里由于 p 为离散的函数，则把每一个 ξ_x 转化为其最接近的整数点。

2.2 直接抽样法

取 $[0,1]$ 均匀分布的随机数列 x_{i1} ，并对 $p(x)$ 从 2900 开始挨个求和，并比较与 ξ_1 的大小，当求和刚好大于 ξ_1 时，返回上一个能量点即为所对应的 ξ_x 。这里是由于离散点，必须通过求和的方式代替积分。

3 结果

如图 1 是 data 文件的归一化后数据图，舍选法结果与直接抽样法结果。

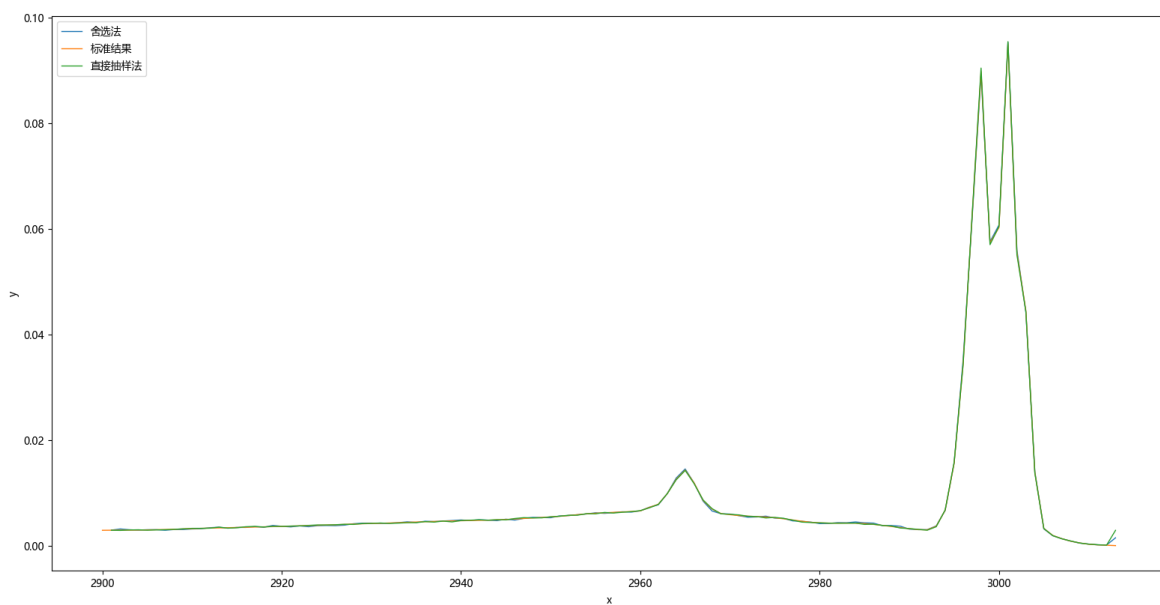


图 1: data 文件抽样对比图

rate: 0.368016988

图 2: 舍选法抽样效率

如图 2 是舍选法抽样效率，达到 0.368 以上。

4 结论

可得舍选法与直接抽样法与原 data 数据图基本吻合。由于本身数据是离散点，所以将数据点近似到整数，用求和代替积分的方法会产生一定的误差。

也可以通过更加细化接近原函数曲线外形的阶梯函数，分成更多区间段，来提高抽样效率。