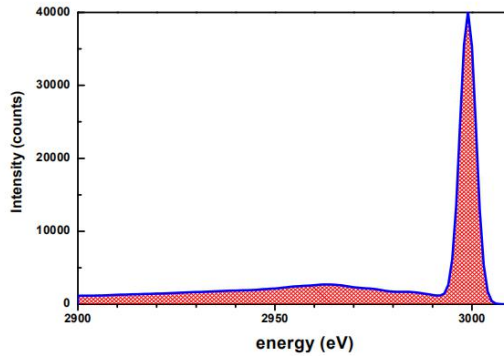


第 7 题作业报告

PB18000341 范玥瑶

A. 作业题目

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。



谱值数据见 data.txt

B. 算法及主要公式

B.1. 直接抽样

因为给出的 $p(x)$ 的能量 E 是步长 1 (eV) 的离散值，所以不妨按照离散值进行抽样。通过

$$p_i = N_i / \sum_i N_i$$

其中 N_i 是该能量对应粒子数目，得到某个能量的几率值 p_i 。由随机数产生器生成 (0, 1) 上均匀分布的随机数 ξ ，均匀抽样表达式为：

$$\sum_{i=0}^{n-1} p_i < \xi \leq \sum_{i=0}^n p_i \quad (n = 1, 2, \dots)$$

当 ξ 满足上述表达式时抽样结果为 $E = E_n = 2900 + n(\text{eV})$ 。 $0 < \xi \leq p_0$ 时 $E_0 = 2900(\text{eV})$ 。

程序的算法为：输入种子值和数据数目 Q ，在 main 函数里用 Schrage 方法生成随机数 x ，调用子函数 Sample 返回能量，在主函数中将能量输出到 sample1.txt。子函数 Sample 的算法是从文件导入 N_i ，存储在数组里，求出 $N = \sum_i N_i$ ，（不导入 E 是因为 E 和数组的编号 n 满足 $E = (2900 + n)\text{eV}$ ），在循环语句中通过检验 $N * \xi$ 和 $\sum_{i=0}^n N_i$ 的大小关系检验 ξ 和 $\sum_{i=0}^n p_i$ 的大小关系，当 $N * \xi \leq \sum_{i=0}^n N_i$ 时退出循环，此时 $E = 2900 + n$ ，返回能量 E 。

B.2. 舍选法

用 Schrag 方法生成 (0, 1) 上均匀分布的随机数列，取相邻两随机数组成数对 x, y ，舍选法抽样值

$$X = 2900 + [(3013 - 2900) * x]$$

$$Y = N_{max} * y$$

当 $Y \leq E(X)$ 时能量值抽样为 X ，有效抽样的次数加 1。

B.3. 重要抽样方法

计算机程序的算法流程为：导入 data，将能量和对应的数目存储于 data[114][2]的前两列，同时求第二列的和 N，调用子函数 Romberg 计算 $\int_{2899}^{data[n][0]} F(\xi)d\xi$ 存入一维数组 z[114]。由 Schrage 方法生成 Q 个随机数，取相邻两随机数为随机数对 X,Y，由 (X,Y) 生成随机数对 (x,y)，(x,y) 满足

$$X = \frac{\int_{2899}^x F(\xi)d\xi}{\int_{2899}^{3013} F(\xi)d\xi} = \frac{\int_{2899}^x F(\xi)d\xi}{1.24747}, \quad y = Y * F(x)$$

调用子函数 sample1 生成 x, 返回主函数进而求出 y。对 (x,y) 判断是否满足 $y \leq p(x)$ 即 $N*y \leq N(x)$ ， $p(x)$ 是实验谱中能量 x 出现的频率，N 是实验谱中的例子总数。若成立则能量的抽样为 x，成功抽样计数 J 增加 1。抽样结束后计算抽样效率 J/Q 输出。

根据对不同函数作图的简单试验，选取

$$F(x) = 0.1 * \exp\left(-\frac{(x - 3000)^2}{50}\right)$$

用 Schrage 方法生成两个在 (0, 1) 上均匀分布随机数 X,Y 进行重要抽样。

子函数 sample1 的算法为：

由 (X,Y) 生成 (x,y) 时，因为实验数据的能量精度是 1eV，所以由 X 求 x 时，x 取值范围是 {2900, 2901, …, 3013}。由于 $\int_{2899}^x F(\xi)d\xi$ 在 (2900, 3013) 上单调递增，使用循环语句中比较 $X' = X * \int_{2899}^{3013} F(\xi)d\xi$ 与 $\int_{2899}^x F(\xi)d\xi$ 直至 $X' \leq \int_{2899}^x F(\xi)d\xi$ ，取 x。由于随机数点的个数较多，为提高运算速度应先计算好积分值存储在数组 z [114] 中，求反函数时直接调用即可。

自定义函数 Romberg 的算法为 Romberg 方法。求反函数的条件 $\int_{2899}^{x-1} F(\xi)d\xi \leq X' \leq \int_{2899}^x F(\xi)d\xi$ ，精度控制值 $e = \Delta X' = \Delta X * 1.24747 \approx 5E - 7 * 1.24747$ ，向上取为 $7E - 7$ 。下求取点个数 N'。最大区间长为 3013-2900=113，要求 $\left(\frac{113}{N'}\right)^{2N'+2} < e$ ，不妨取 N'=121。

编写程序得出结果，调试发现由于浮点数精度影响，Romberg 积分输出恒为 0。改用符号计算软件得到数值积分文件 z.txt 进行抽样，较小的 x 数值积分均为 0，因此使用 F (x) z 作为舍选函数不合理。用 Romberg 积分的程序附在压缩包中。

C. 计算结果及具体分析、讨论

C.1. 直接抽样

编写程序取 seed=1 计算了 1E5 个点的能量抽样作频数分布直方图如图 1，以 data.txt 作原曲线如图 2。由算法，抽样效率为 100%。

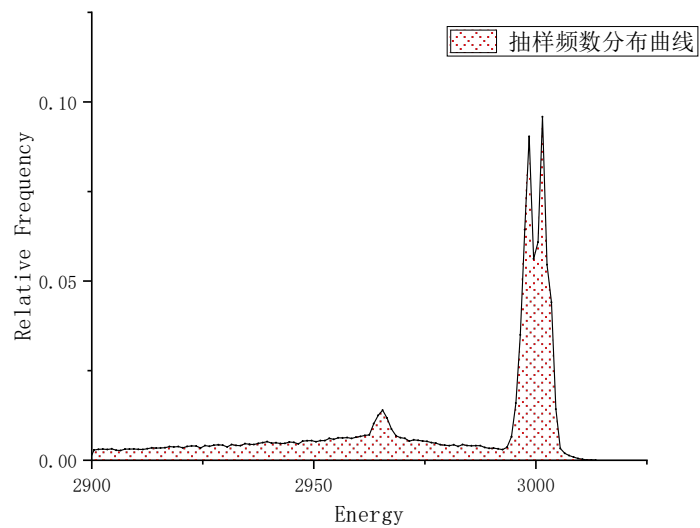


图1 简单抽样能谱

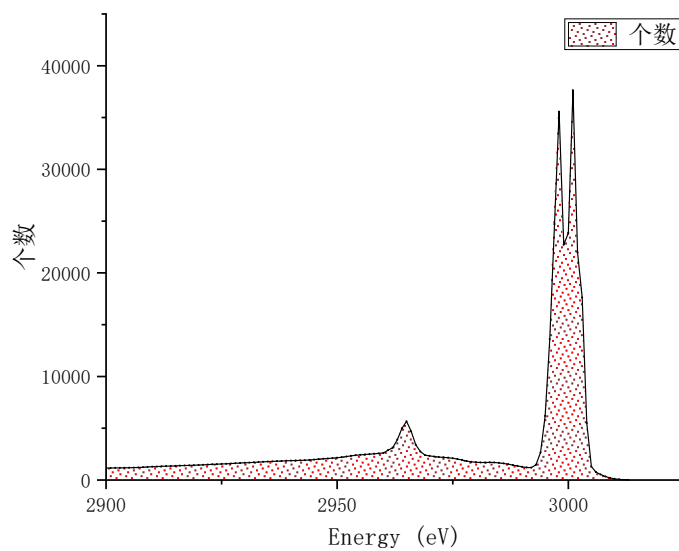
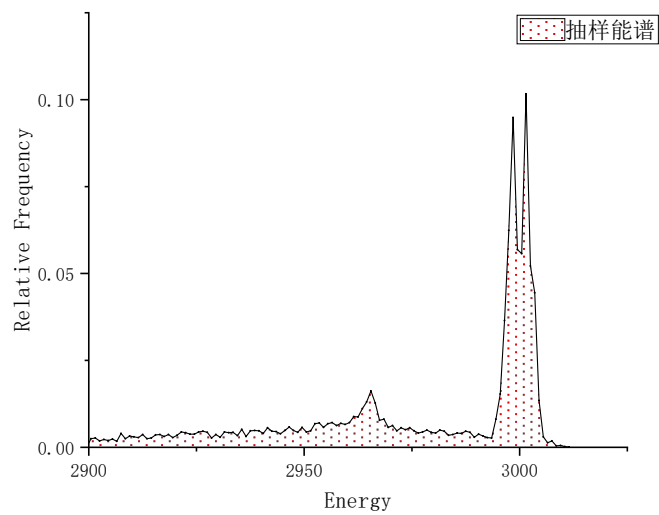


图2 能谱实验数据

形状基本一致，抽样效果好。

C.2.简单舍选法

编写程序，取 seed=1 计算了 1E5 个点的能量抽样作频数分布直方图如图 3，和图 2 对比比较接近，但是能量较高两峰之间的部分不如图 1 相似度高，抽样效果略差于直接抽样。抽样效率 $9.2650\% < 100\%$ 。



D. 总结

本次作业中选用了离散情形直接抽样和舍选法抽样两种抽样方式，以实验测得的能谱 data.txt 为分布函数进行了能谱的蒙特卡洛计算，分别统计了三种方式的抽样效率。简单抽样的抽样效率最高，但是用时长；舍选法抽样效率低；理论上重要抽样方法效率在两者之间，但是可以兼顾抽样效率和计算速度。