

计算物理A作业8

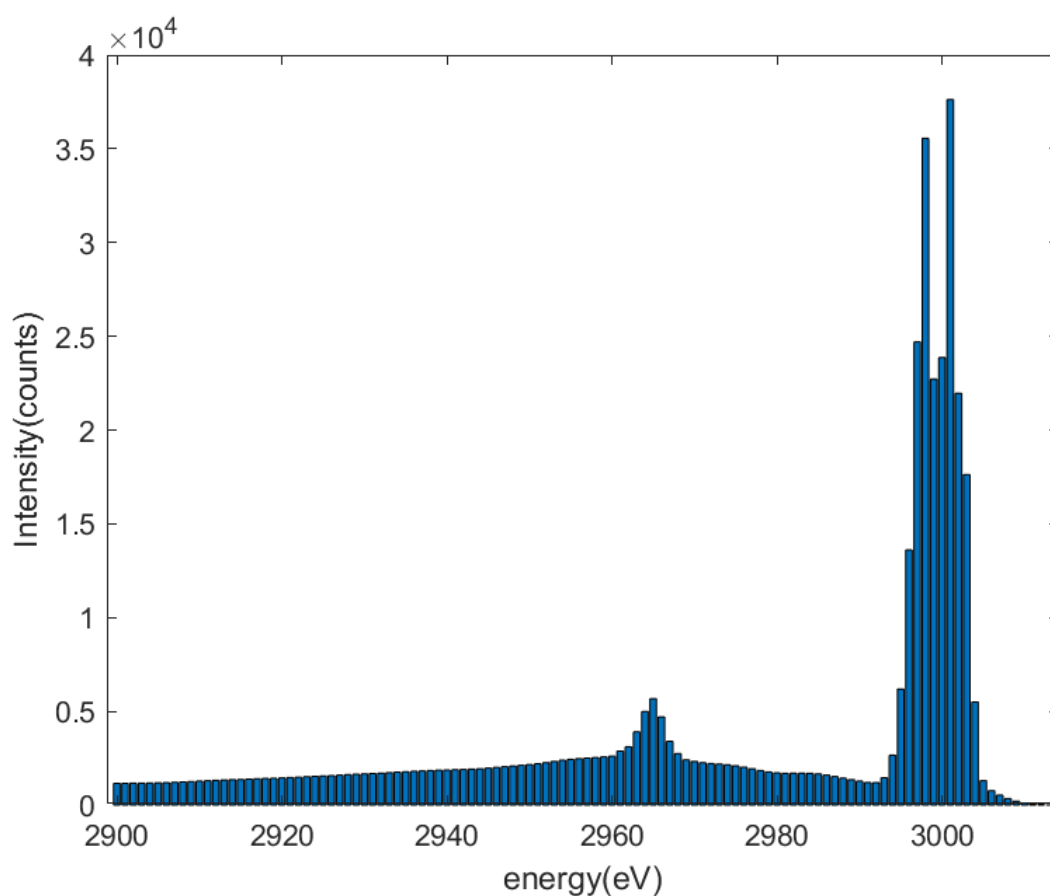
吕邀 PB19030789

1.作业题目

对一个实验谱曲线 $p(x)$ ，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.算法和主要公式

按题中所给数据画出直方图如下所示：



2.1 直接抽样

本题中所给数据横坐标是离散的能量数值，记为 X_i ，对应的纵坐标为 X_i 出现的频数 N_i ，由此我们可以得到能量数值在区间 $[x_i, x_{i+1}]$ 之间的概率 $p(i) = \frac{N_i}{\sum N_i}$ 。

由此我们可以得到直接抽样的方法：在区间 $[0, 1]$ 上均匀抽取随机数 ξ ，如果 $P(k-1) < \xi < P(k)$ ，则能量取值就为 X_k 。

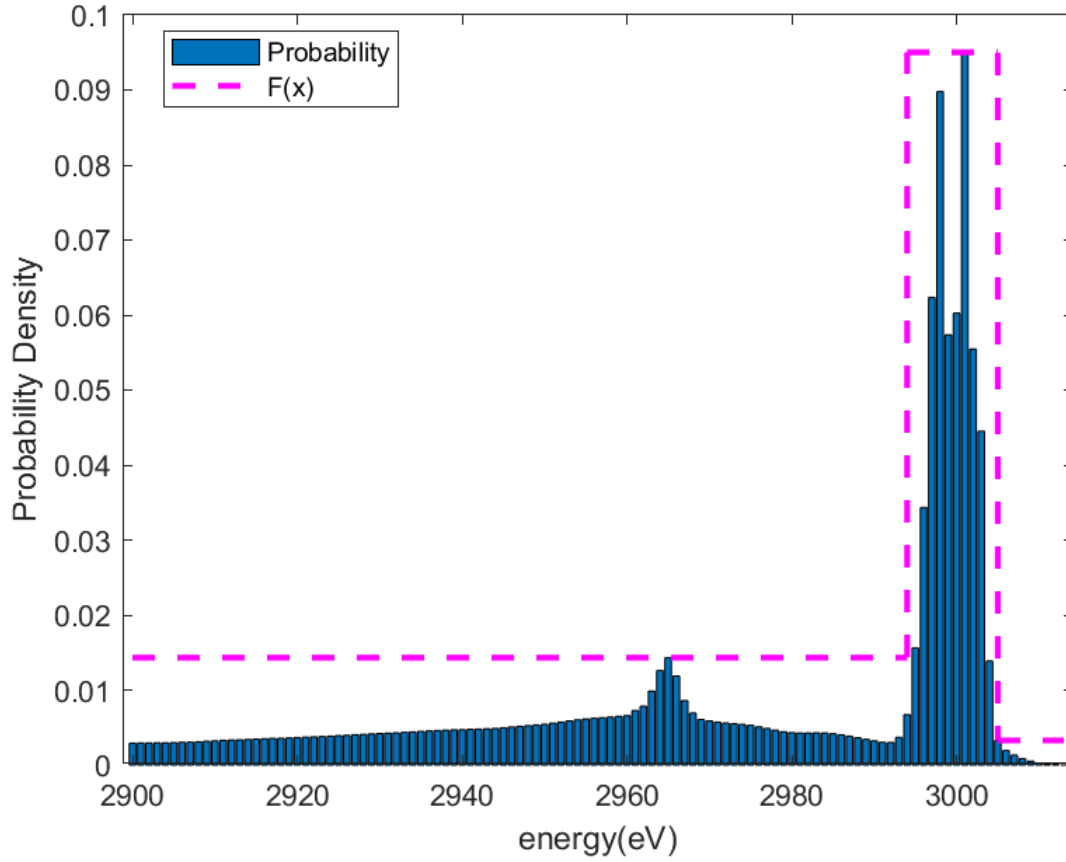
其中 $P(k) = \sum_{i=2900}^k p(i)$ 。

直接抽样的取值我们在程序direct_sample.c中进行。

2.2 舍选抽样

舍选抽样需要选择一个合适的比较函数，我们选取一个比较简单的常值阶梯函数 $F(x)$ ：

$$F(x) = \begin{cases} 0.0143232, & 2900 \leq x \leq 2994 \\ 0.0950248, & 2995 < x \leq 3004 \\ 0.0032778, & 3005 < x \leq 3013 \end{cases}$$



曲线 $F(x)$ 所包围的面积为 $\int_{2900}^{3013} F(x) dx = 2.41559$

对 $F(x)$ 积分可求得累计函数（已归一化）为

$$G(x) = \begin{cases} 0.00592948x - 17.195492, & 2900 \leq x \leq 2994 \\ 0.0393381x - 117.811, & 2995 < x \leq 3004 \\ 0.00135694x - 3.16618, & 3005 < x \leq 3013 \end{cases}$$

在区间 $[0, 1]$ 上产生两个随机序列 ξ_1 和 ξ_2 。在 x 方向上，按照 $F(x)$ 的分布抽样，即 $\xi_1 = G(\xi_x)$ ，可以反解出 ξ_x ：

$$\xi_x = \begin{cases} 168.649\xi_1 + 2900, & 0 \leq \xi_1 \leq 0.557371 \\ 25.4206\xi_1 + 2984.832, & 0.557371 < \xi_1 \leq 0.989978 \\ 733.952\xi_1 + 2282.32, & 0.989978 < \xi_1 \leq 1 \end{cases}$$

在 y 方向上，有 $\xi_y = F(\xi_x)\xi_2$ ，

最后需要比较大小

$$\begin{cases} \xi_y < p(\xi_x) \\ \xi_y \geq p(\xi_x) \end{cases}$$

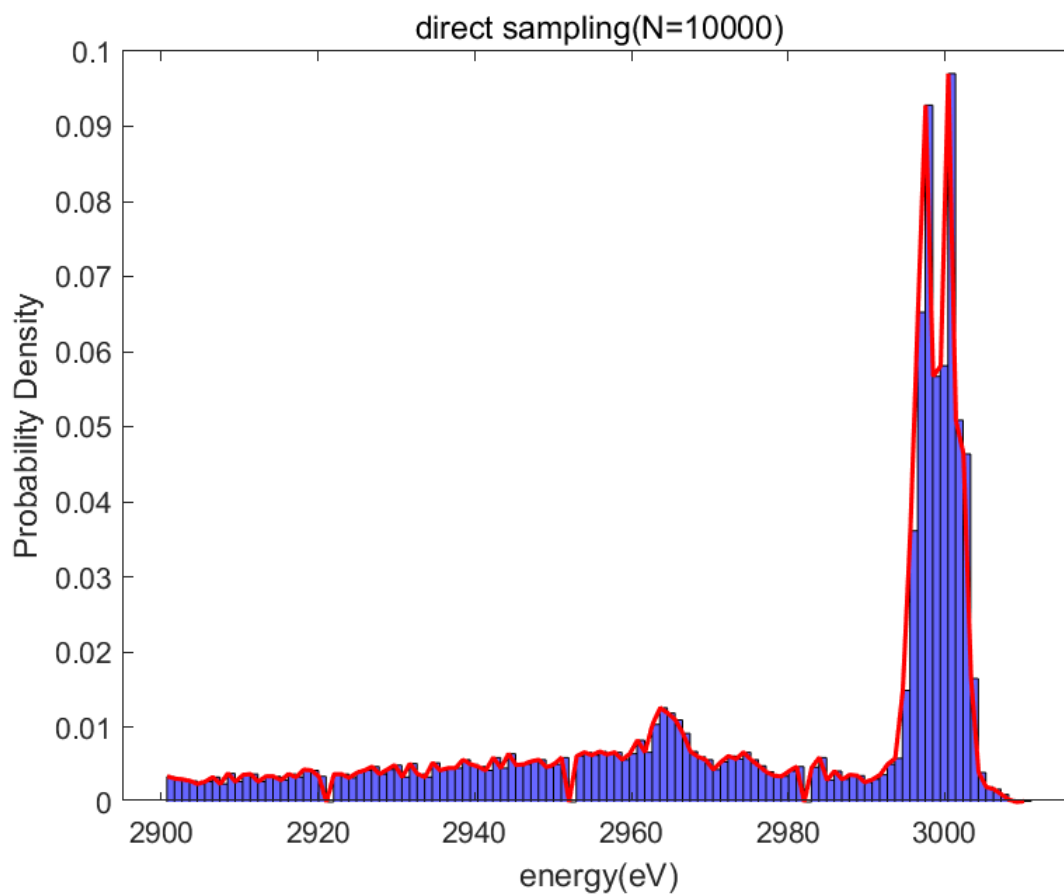
上式中小于则取，大于等于则舍。

由于舍选抽样的效率不高，这次我们适当增加随机点的总数，取 $N = 10000000$ 。

舍选抽样的过程在rejection_sample.c中进行，在抽样的同时计数并计算了舍选抽样的效率。

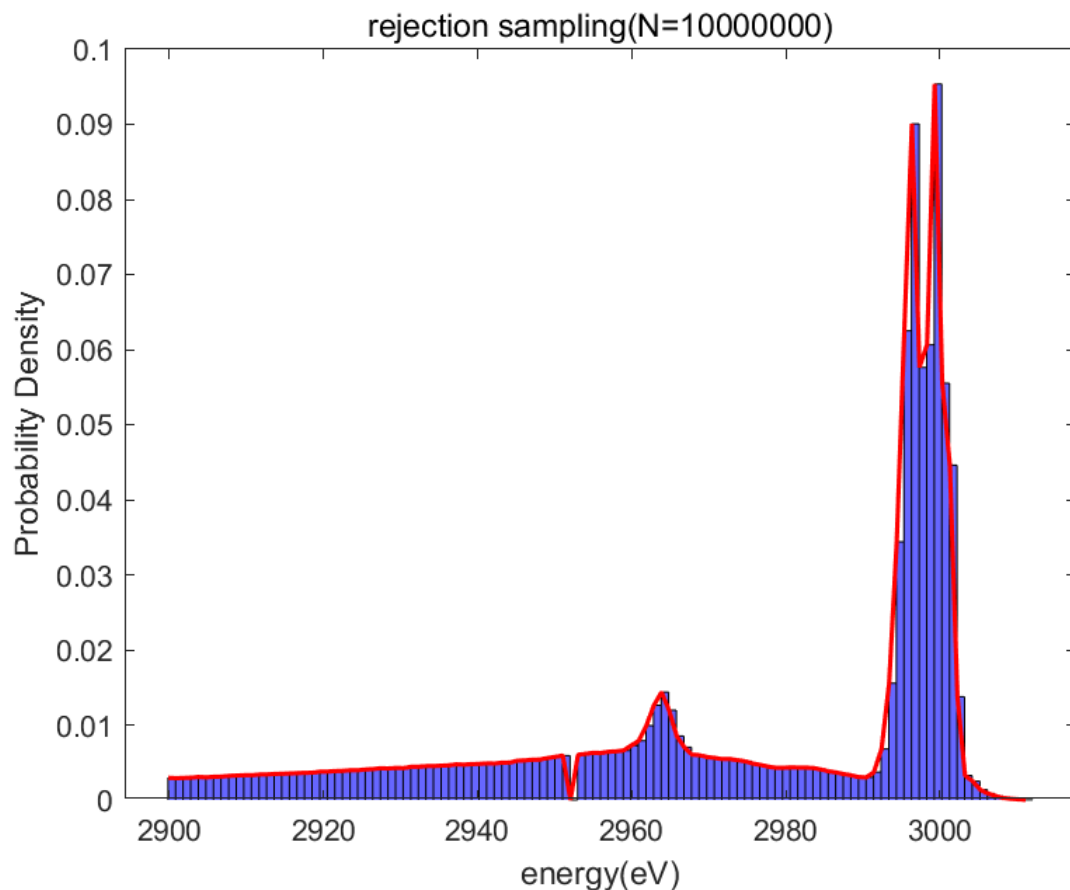
3.计算结果及分析

3.1 直接抽样



根据直接抽样的原理，我们可以认为直接抽样的效率为1。

3.2 舍选抽样



我们在程序中计数计算出舍选抽样的效率为0.412642，理论上的效率应该是 $\frac{1}{\int_{2990}^{3013} F(x) dx} = 0.413978$ ，可见二者很接近，舍选抽样的方法用在这里是合理的。

但我们同时可以看到舍选抽样的效率其实是比较低的，所以想更好地拟合出原来的谱曲线必须增加总的取点数 N 。

4.总结

- (1) 直接抽样和舍选抽样都能得到与原曲线比较接近的拟合曲线。
- (2) 直接抽样的效率为1，而舍选抽样的效率则比较低，只有大约40%。所以对于舍选抽样来说，想要得到比较准确的结果，需要增加 取点的数目。
- (3) 舍选抽样的效率与比较函数的选取有关。很显然这本题中如果我们将比较函数分的更细，是可以得到更高的抽样效率。但这样做会使比较函数本身变得非常复杂，同样不利于计算。这可能是舍选抽样一个不足的地方。