计算物理第 6 次作业第 13 题

姓名：姚星宇　　学号：PB21000188　　班级：2021 级少年班学院四班

2022 年 11 月 10 日

1　题目重述

用 Metropolis-Hasting 抽样方法计算积分：

$$I=\int_0^{\infty}(x-\alpha\beta)^2 f(x)dx=\alpha\beta^2(\text{理论精确值}) \tag{1}$$

其中

$$f(x)=\frac{1}{\beta\Gamma(a)}(\frac{x}{\beta})^{\alpha-1}\exp(-\frac{x}{\beta}) \tag{2}$$

对于权重函数为

$$p_1(x)=f(x),p_2(x)=(x-\alpha\beta)^2 f(x) \tag{3}$$

分别进行计算。在给定 α,β 的情况下，选用不同的步长 i, 讨论精度与效率。

这里根据自己的理解稍微改动了题目的表述细节，最主要的改动是将原题中的 γ 改成了抽样的步长（用 i 表示）。如与原题意思不符合还请见谅。

2　题目分析

本题的所有程序均采用 Anaconda 3 编写。

本题的核心在于 Metropolis-Hasting 抽样。在参考文献 [1] 中，对于实现该类抽样的方法有清晰的表述：

设 p(x) 为所考虑的几率密度分布，问题在于是如何产生下一个抽样点 x_{n+1}。可以在上一个点附近构造一个试探解，$x_t=x_n+\delta$，δ 是试探步长，$\delta=(\xi-0.5)i$ 该点是否被选取决定于比值 $r=\frac{p(x_t)}{p(x_n)}$：若 r >1，则选取。否则，产生均匀分布的随机数 $\xi in[0,1]$：若 $\xi<r$ 则选取；否则，放弃。

由上述方法，去除热化部分后，可得到按被抽样函数 $p(x)$ 分布的数组 x_j。若待积分的式子形式为：

$$I=\int g(x)p(x)dx=\sum_x g(x)p(x)\Delta x \tag{4}$$

则在抽样点足够密集时，可由 $p(x)\Delta x$正比于$\frac{1}{N}\sum_{j\in\{j|x<x_j<x+\Delta x\}}$, 得到：

$$I=\frac{A}{N}\sum_{j=1}^{N}g(x_j) \tag{5}$$

其中 A_0 为一与 $g(x)$ 无关的常数，令 $g(x)=1$，则：

$$\int p(x)dx=I=A \tag{6}$$

所以积分方法为:

$$I=\frac{\int p(x)dx}{N}\sum_{j=1}^{N}g(x_j) \tag{7}$$

对于 $p_1 = f(x)$，由 Γ 函数相关定义，可得 $\int_0^\infty f(x)dx = 1$，即：

$$I_1 = \frac{1}{N}\sum_{j=1}^{N}(x_j - \alpha\beta)^2 \tag{8}$$

对于 $p_2 = f(x)(x-\alpha\beta)^2$，方法略复杂些。令 $g(x) = \frac{1}{(x-\alpha\beta)^2}$, 则：

$$1 = \int_0^\infty f(x)dx = I = \frac{\int_0^\infty f(x)(x-\alpha\beta)^2 dx}{N}\sum_{j=1}^{N}(x_j-\alpha\beta)^{-2} \tag{9}$$

即得：

$$\int_0^\infty f(x)(x-\alpha\beta)^2 dx = N(\sum_{j=1}^{N}(x_j-\alpha\beta)^{-2})^{-1} \tag{10}$$

即可计算出所需的积分。

3 结果

表 1：$\alpha = 2.5, \beta = 2.5$ 时积分值与抽样效率和步长的关系

i	0.1	0.2	0.4	0.8	1.6	3.2	6.4	12.8	25.6	51.2
(p_1)eff%	66.3	66.1	65.5	64.5	62.6	59.0	53.2	44.6	31.8	18.4
$(p_1)\lvert res-\alpha\beta^2\rvert$	4.82	5.62	2.87	1.41	0.35	0.45	0.37	0.67	4.48	7.35
(p_2)eff%	65.9	65.2	64.6	63.3	61.3	56.8	49.6	41.1	35.3	25.8
$(p_2)\lvert res-\alpha\beta^2\rvert$	3.20	5.92	2.81	0.61	1.14	1.71	0.55	0.47	0.84	2.54

$\alpha = 2.5, \beta = 2.5$ 时，10^6 次抽样下，积分值与抽样效率和步长的关系如图 1，表 1 所示。由图表中的数据分析可得：对于两种不同的分布函数，均在步长接近 6.4 时，抽样效率接近 50%，且积分计算的精确度最高。

另一组 α, β 的图像如图二所示，表略去，分析结果类似：对于两种不同的分布函数，均在步长接近 6.4 时，抽样效率接近 50%，且积分计算的精确度最高。

由此可以得到初步结论：$\alpha, \beta \sim 10^0$ 时，抽样步长在 $3-10$ 的范围内效果最佳。

参考文献

[1] 丁泽军. 计算物理讲义 [M]

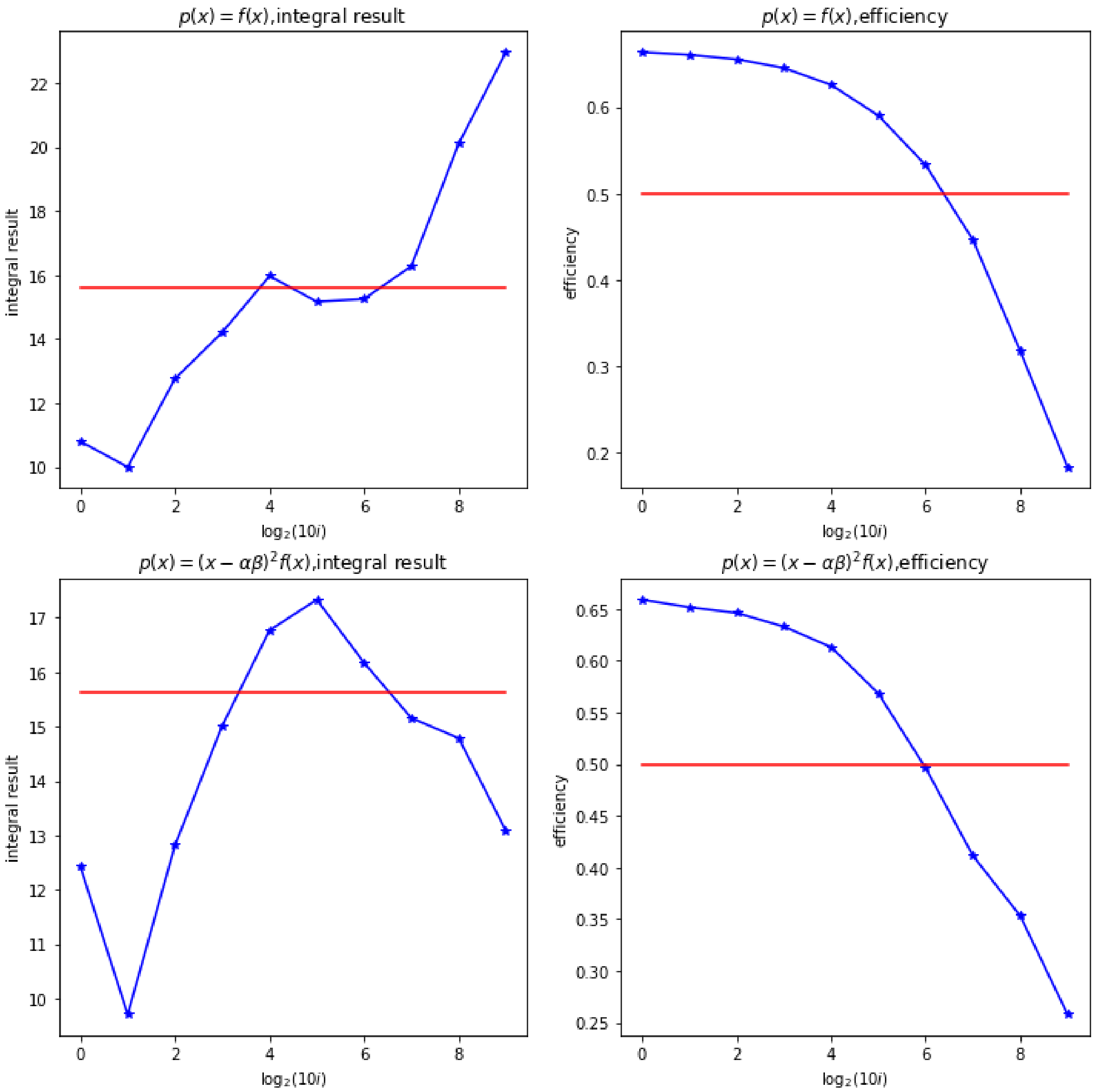

图 1: $\alpha = 2.5, \beta = 2.5$ 时积分值与抽样效率关于步长的曲线 (红线为积分理论值和抽样概率为 0.5 的参考值)

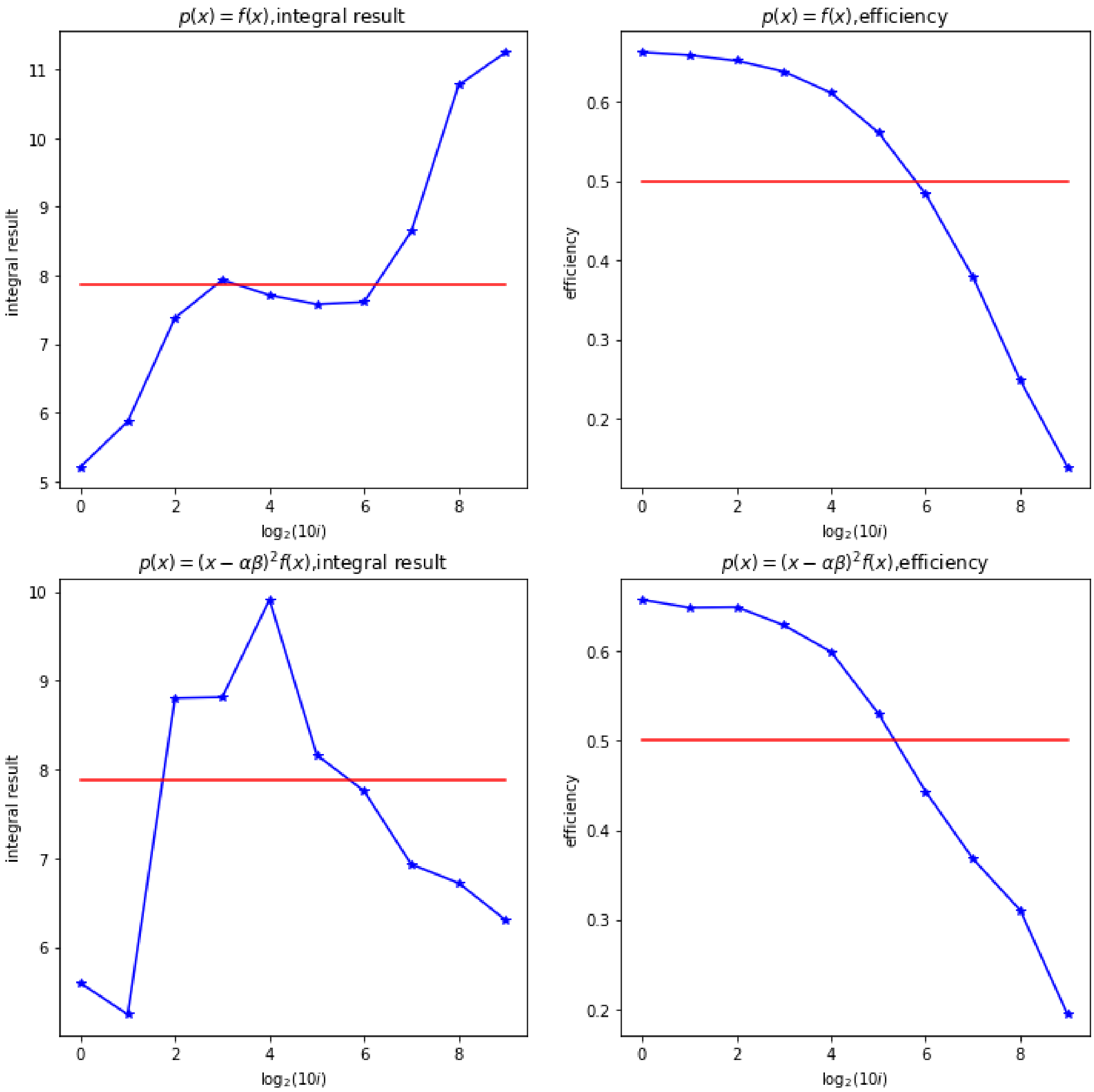

图 2: $\alpha = 1.5, \beta = 3.5$ 时积分值与抽样效率关于步长的曲线 (红线为积分理论值和抽样概率为 0.5 的参考值)