Report13

> PB20020480 王润泽

1 Question

用 Metropolis-Hasting 抽样方法计算积分：

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

其中 $f(x)=\frac{1}{\beta\Gamma(\alpha)}(\frac{x}{\beta})^{\alpha-1}\exp(-x/\beta)$ 。设权重函数为：$p(x)=f(x)$和 $p(x)=(x-\alpha\beta)^2f(x)$。给定参数 α,β 用不同的 γ 值计算积分，讨论计算精度和效率。

2 Analysis

2.1 $p(x)=f(x)$

2.1.1 抽样算法

根据 Metropolis-Hasting 抽样规则，考虑到待抽样 $f(x)$的分布形式，取 $T(x\to x')=T(x')=\frac{1}{\gamma}\exp(-x'/\gamma)$为建议分布函数，与 $f(x)$形状比较接近

根据简单抽样方法，$T(x')$的累积分布函数为

$$F(x')=\int_0^{x'}T(t)dt=1-\exp(-x/\gamma)$$

所以，对于符合均匀分布[0，1]的随机数R来说，其可以对符合分布 $T(x')=\frac{1}{\gamma}\exp(-x'/\gamma)$ 的x'抽样

$$x'=-\gamma\ln(1-R)\Leftrightarrow-\gamma\ln R$$

于是Metropolis-Hasting抽样$f(x)$的算法为：

设起始点 $x_0=1$

1. 设X为Markov链集合 $\{x_k\}$，取最后一个点为 x_n
2. 在[0,1]均匀分布中抽样出随机点R，根据简单抽样法，抽样得到 $x'=-\gamma\ln R$
3. 根据Metropolis-Hasting方法，令

$$r_1=\frac{p_jT_{ji}}{p_iT_{ij}}=\frac{p(x')T(x_n)}{p(x_n)T(x')}=(\frac{x'}{x_n})^{\alpha-1}\exp[-\frac{x'-x_n}{\beta}]\exp[\frac{x'-x_n}{\gamma}]$$

4. 在[0,1]均匀分布中抽样出随机点ξ，若 $\xi<\min(1,r)$，取 $x_{n+1}=x'$；若$\xi>\min(1,r)$，取 $x_{n+1}=x_n$。
5. 将 x_{n+1}加入到 X集合，循环到 步骤1，直到集合总数大于 N

2.1.2 重要抽样法求积分

当得到有N个点的Markov链集合后，将序列中热化过程的前M个构型舍去，那么通过重要抽样方法可以得到积分数值表达式为

$$I_1 = \frac{1}{N-M}\sum_{i=M+1}^{N}(x_i-\alpha\beta)^2 \tag{2}$$

2.2 $p(x)=(x-\alpha\beta)^2 f(x)$

2.2.1 抽样算法

若令 p(x)为被积函数本身，则积分的结果为 p(x)的归一化因子 C，那么把 x满足的概率分布函数令为 $g(x)=\frac{p(x)}{C}$

而Metropolis-Hasting 抽样规则不依赖于归一的概率分布，依然可以对x进行抽样，得到的x符合 $g(x)$的概率分布。

具体抽样步骤只要将 **2.1**中 **步骤4**里的 $\mathbf{r}$ 改变，即

$$r_2=\frac{p_iT_{ji}}{p_iT_{ij}}=(\frac{x'-\alpha\beta}{x_n-\alpha\beta})^2(\frac{x'}{x_n})^{\alpha-1}\exp[-\frac{x'-x_n}{\beta}]\exp[\frac{x'-x_n}{\gamma}]$$

其他步骤不变，即可得到满足 g(x) 概率分布的抽样。

2.2.2 比值法求积分

为了得到数值积分结果，即得到归一化系数 C，对集合 X中的点进行统计，则可以得到 x 的近似概率分布函数$g^*(x)$，那么数值积分为

$$I_2=C=\frac{p(x)}{g^*(x)}$$

那么这样的点在实数域有无穷多，但考虑到对抽样结果统计时，总是离散化定义域的点 $\{x_i\}$取频率分布来计算，那么计算 I_2时也可以这样做，即对这些离散点 $\{x_i\}$处的值求平均

$$I_2=\frac{1}{n}\sum_{i=0}^{n}\frac{p(x_i)}{g^*(x_i)}$$

为了尽量使得结果准确可靠性，应当选择 $g(x)$的值较大的点 x_m，即概率出现较大的点，防止出现统计上的偏差，故设概率截断阈值为 $\epsilon=0.001$, 当 $g^*(x_i)<\epsilon$ 时舍去该点，剩下的 n'个点 $\{x_j\}$，故得到积分表达式为

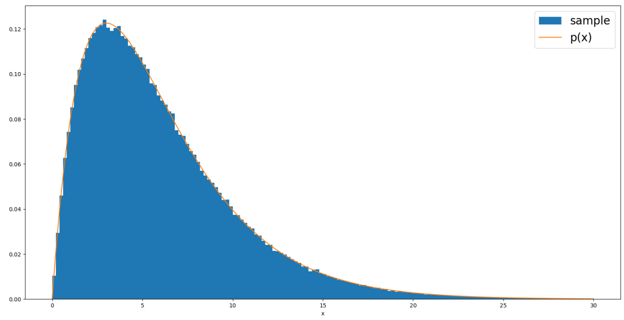
$$I_2=\frac{1}{n'}\sum_{j=0}^{n'}\frac{p(x_j)}{g^*(x_j)} \tag{3}$$

3 Experiment

3.1 $p(x) = f(x)$

3.1.1 实验结果

利用Metropolis方法对p(x)抽样结果如下所示



Metropolis方法抽样结果不错

根据公式（1）（2），**定义相对误差为**：$error = \frac{|I_0 - I_1|}{I_1}$

根据抽样时第4步，$\xi < \min(1, r)$是否成立，决定Markov链下一步 x_{i+1}是否更新。**定义效率为**：$rate = \frac{n}{N}$，其中n为 $x_i \neq x_{i+1}$的个数，N为X集合总数

选择 $\alpha = 2, \beta = 3, \gamma = 6$, `hw13.py`输出结果如下

```
I=18 ,I1 = 17.976104210004017
relative error = 0.001327543888665714
sample rate = 0.7629607843137255
```

结论1：由程序输出可见，相对误差较低，至少有3位有效数字，效率为0.7，较为不错

3.1.2 改变 γ，观察曲线变化

实验中，尝试了 $\alpha = 2, \beta = 3$的取值情况，改变 γ取值，修改建议分布参数，得到以下曲线

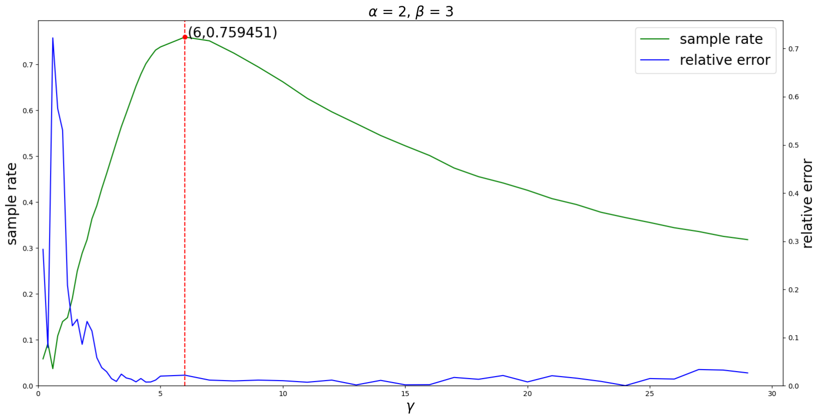


图1：γ 与相对误差和接受效率关系

从图像上可以很明显的看出，其接受效率（舍选效率）先上升后下降，在 $\gamma = 6$时取最大。

而其相对误差在 γ较小时由较大的波动，且相对误差较大；在 γ很大时，其相对误差较小，但也会有增长的趋势；只有在 $\gamma = 6$附近处，相对误差保持在较小的水平。

那么有理由认为 $\gamma = 6$，建议分布$T(x') = \frac{1}{\gamma}\exp(-x'/\gamma)$抽样的结果较好。

如图2所示

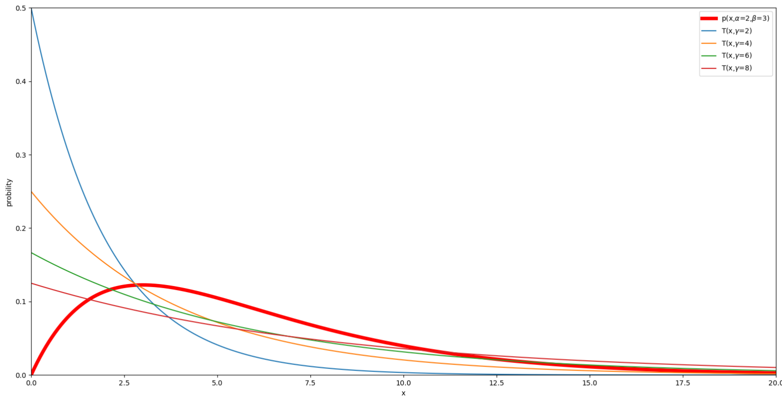


图2：α,β与不同γ取值下待抽样分布与建议分布图像

由图2可以解释图1中效率先上升后下降，而误差随着 γ增大而减小的原因。

对于指数分布来说，γ 越小，概率密度衰减越快。即γ较小时指数分布主要集中于 x 较小的地方；γ 增大时 x 取大值的可能性也会增高。而我们待求的分布 f(x)是一个有极值点的函数，它在 x 轴上某一段区间上分布较为集中。

一开始γ 较小时，指数分布使得抽样点集中在 x 较小的区间，和 p(x)所集中的区间不重合，使得采样率较低，并且导致试探点x' 不能在整个区间内分布，会使得抽样的结果并不接近 $p(x)$,导致误差偏大

随着 γ 的增大指数分布所集中的区间和 p(x)开始重合，这时候采样率有所上升，误差有所下降。

而当 γ 太大时指数分布几乎成为很大一段区间上的均匀分布，和 f(x)所集中的区间相差太远，使得采样率有所下降，但误差保持在较低的水平

结论2：随着 γ增大，接受效率（舍选效率）先上升后下降；相对误差先剧烈波动，后逐渐下降并趋于稳定

3.1.3 最优 γ^*取值

为了寻找 γ取最优时的表达式，故改变 α, β 的取值，定义最优 γ^* 为接受效率最高处，得到下表

a\b	**3**	**4**	**5**	**6**	**7**
3	9	11.5	15	19.5	20.5
4	12.5	15.5	21	26.5	29.5
5	14.5	17.5	24	32.5	33
6	19.5	22.5	30	35	43
7	22	28	**39**	**48**	49

表1：γ* 与α和β关系（步长为0.5）

由表中数据可以看到最优 γ^*大致都在 $\alpha * \beta$附近，大部分误差不超过 ± 4，故猜测 $\gamma^* = \alpha\beta$

考虑到表1数据，采样时步长为0.5，选取可能较大，且搜索范围较宽，选择出 γ^*精度较低

故为了进一步验证 γ^*表达式，重新选择更小的步长0.1，在 $\alpha\beta$范围附近（± 4），重新采样，得到下表

a\b	3	4	5	6	7
3	9	11	14.9	19.4	21.7
4	11.3	16.8	22.3	24.1	26.9
5	14.1	20.9	25.1	27.9	36.6
6	16.8	24.8	30.9	34.2	38.1
7	18.7	28.4	35.6	44.7	49

表2：γ* 与α和β关系（步长为0.1）

可以更加具体的看到 γ^*取值在 $\alpha\beta$ 范围附近，可以得到以下结论

结论3：建议分布T(x)其参数 γ选择在 $\alpha\beta$附近有较好的结果

3.2 $p(x) = (x - \alpha\beta)^2 f(x)$

3.2.1 实验结果

依然取 $\alpha = 2, \beta = 3$，则 $p(x)$图像如下

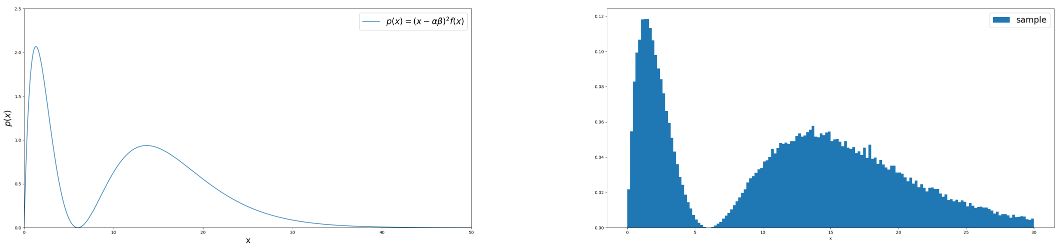


图3：p(x)与抽样图像

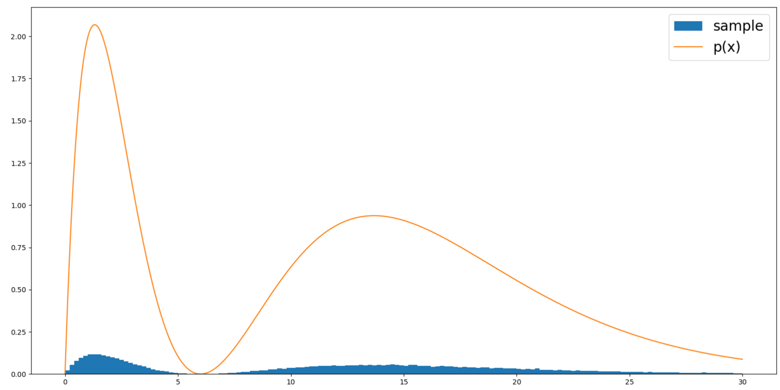


图4：p(x)与抽样图像对比 由图3，图4可见抽样的形状是一致的，但放在一起，可以看出 p（x）未归一化

根据公式（3），程序得到以下输出

```
I=18 ,I2 = 17.95174627809675
relative error = 0.00268076232795838
sample rate = 0.5950916334661355
```

由程序输出可见，相对误差较低，至少有3位有效数字，效率为0.7，较为不错

3.2.2 改变 γ，观察曲线变化

实验中，尝试了 $\alpha=2,\beta=3$的取值情况，改变 γ取值，修改建议分布参数，得到以下曲线

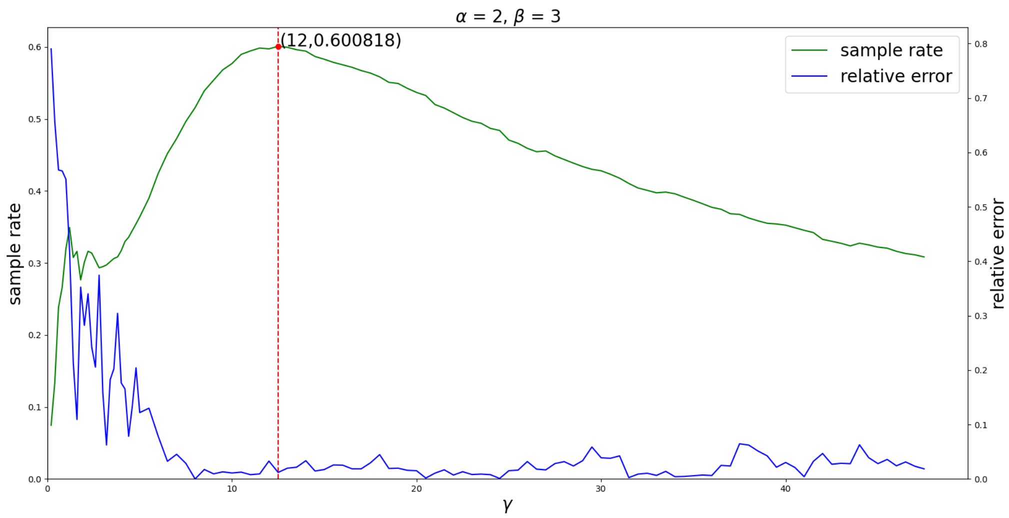


图5：γ 与相对误差和接受效率关系

与 **3.1.2** 类似，**随着 γ增大，接受效率（舍选效率）先上升后下降；相对误差先剧烈波动，后逐渐下降并趋于稳定**

3.2.3 最优 γ^*取值

为了寻找 γ取最优时的表达式，故改变 α,β 的取值，定义最优 γ^* 为接受效率最高处，得到下表

a\b	3	4	5	6	7
3	15.5	21.5	29	32.5	34.5
4	19.5	23	31	36	43.5
5	21.5	33.5	41.5	39.5	**41.5**
6	25	34.5	46	47	**47**
7	25	34.5	44	**47.5**	**46**

表3：γ* 与α和β关系（步长为0.5）

由表中数据可以看到最优 γ^*大致都在 $(\alpha+2)\beta$或$\alpha(\beta+2)$附近，大部分误差不超过 ±5，在 α，β值较大时，与预期有较大偏差（**标粗值**），故可以猜测**当 α,β不大时，最优γ^*=$(\alpha+2)\beta$或** $\alpha(\beta+2)$

4 Summary

本次实验利用 Metropolis-Hasting 方法，分别以 p(x)=f(x)和 p(x)= $(x-\alpha\beta)$ 2$f(x)$计算积分。

当 $p(x)=f(x)$时，存在一个最优的$\gamma=\alpha\beta$使得整体抽样率最高，误差较小；熟悉了重要抽样方法求积分的过程。

当 $p(x)=(x-\alpha\beta)^2 f(x)$时，当 α,β不大时，存在一个最优的$\gamma=(\alpha+2)\beta$或$\alpha(\beta+2)$使得整体抽样率最高，误差较小；同时掌握了求解归一化系数的方法（比值法）

在两个实验中，固定 α,β 改变 γ的值，都会出现类似的曲线，即随着 γ增大，接受效率（舍选效率）先上升后下降；相对误差先剧烈波动，后逐渐下降并趋于稳定。故在数值积分时，选择恰当的建议分布 $T(x)$ 对积分精度与效率会有明显的改善。