Report13

胡琦浩　PB21000235

一、问题

用Metropolis-Hasting抽样方法计算积分：$I=\int_0^\infty(x-\alpha\beta)^2f(x)\,dx=\alpha\beta^2$　　其中 $f(x)=\frac{1}{\beta\Gamma(\alpha)}(\frac{x}{\beta})^{\alpha-1}exp(-x/\beta)$

设积分的权重函数为：$p(x)=f(x)$和$p(x)=(x-\alpha\beta)^2f(x)$

给定参数α,β，并用不同的γ值，分别计算积分，讨论计算的精度和效率

二、解决问题

在本题中，不妨选取$\alpha=2,\beta=3,$ 则$I=\alpha\beta^2=18$

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

设T与初态无关(即非对称的)：$T_{ij}=T(x\to x')=T(x')=\frac{1}{\gamma}exp(-x'/\gamma)$，与$f(x)$形状类似。

$F(x')=\int_0^{x'}T(t)dt=1-exp(-x'/\gamma)=1-R$，$R$为[0,1]的均匀随机数

得:$x'=-\gamma lnR$

设初始:$x_0=1$，$r=\frac{p_jT_{ji}}{p_iT_{ij}}=\frac{f(x')T(x_i)}{f(x_i)T(x')}=(\frac{x'}{x_i})^{\alpha-1}e^{-(x'-x_i)/\beta}e^{(x'-x_i)/\gamma}$

由Metropolis-Hasting方法：

$$x_{i+1}=\begin{cases}x',\ if\ R'<min(1,r)\\ x_i,\ if\ R'>min(1,r)\end{cases}$$

式中:R'为另一产生$[0,1]$的随机数

此时:$I=\frac{1}{N}\sum_{i=1}^N(x_i-\alpha\beta)^2$

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

由于此时的权函数未归一化，且其归一化常数为I，则$g(x)=\frac{p(x)}{I}$为其归一化后的函数。

同上面的抽样方法:初始$x_0=1$，$x'=-\gamma lnR$，
$r=\frac{f(x')T(x_i)}{f(x_i)T(x')}=\frac{(x'-\alpha\beta)^2}{(x_i-\alpha\beta)^2}(\frac{x'}{x_i})^{\alpha-1}e^{-(x'-x_i)/\beta}e^{(x'-x_i)/\gamma}$

然后利用Metropolis-Hasting抽样方法：

$$x_{i+1}=\begin{cases}x',\ if\ R'<min(1,r)\\ x_i,\ if\ R'>min(1,r)\end{cases}$$

得到满足$g(x)$分布的x抽样。

不过此时归一化常数正是我们要求的，因此不妨利用得到的样本X，得到近似的$g(x)$表达式：
$g(x)=\frac{N(x<X\leq x+\Delta x)}{N}$

为了提高精度，选取$\Delta x=0.1$且选取$g(x)$相对较大的区域，具体选择区域需根据样本X的概率直方图来确定。

这样，$I=\frac{1}{N^*}\sum\frac{p(x_i)}{g(x_i)}$，式中$N^*$为选择$g(x)$的数目

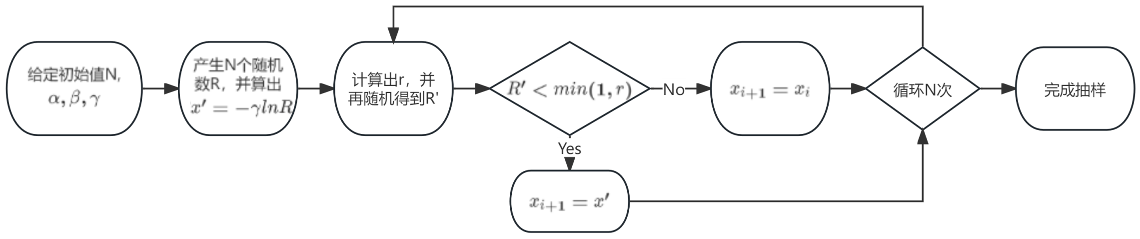

图1：抽样过程(在本题中$N=10^5$)

三、实验结果

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

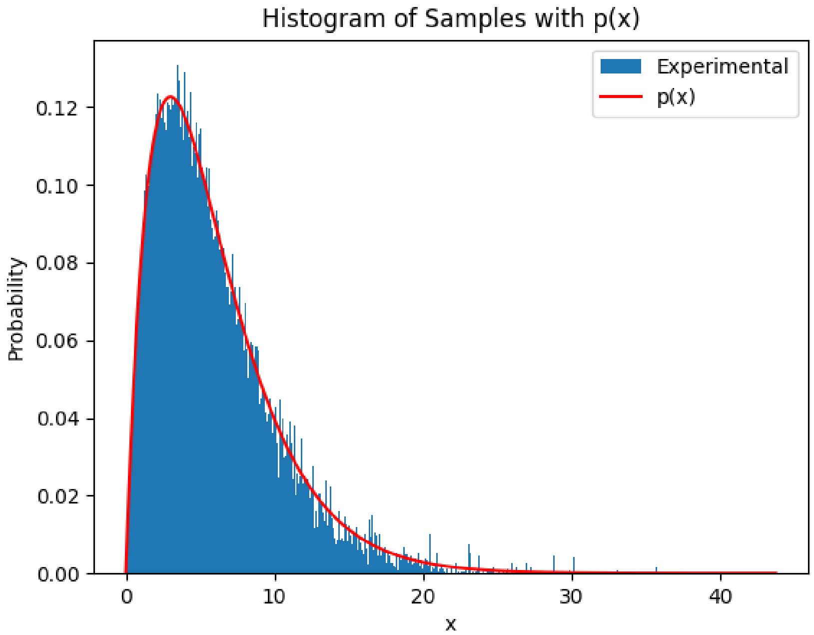

图2：$\gamma=3$时的抽样图案，$I=17.6893$

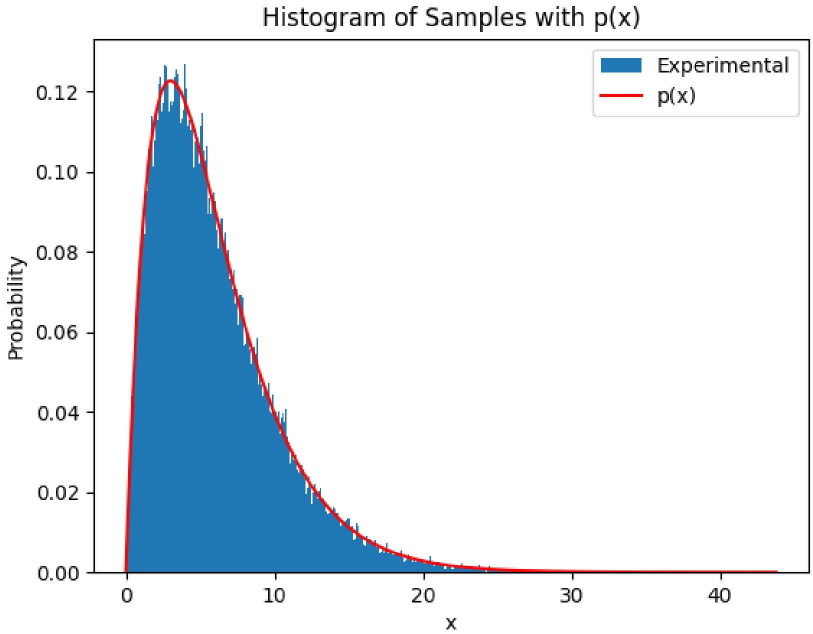

图3：$\gamma = 6$时的抽样图案，$I = 17.9303$

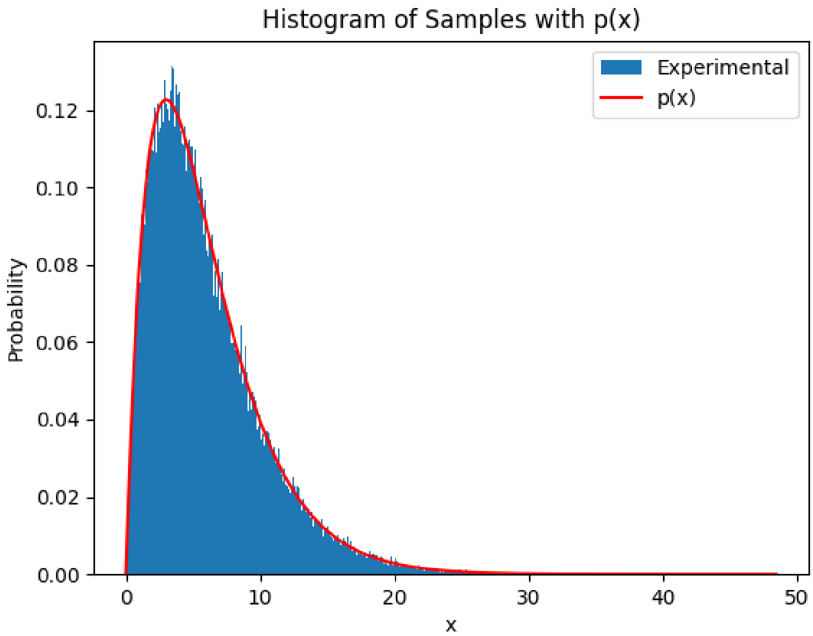

图4：$\gamma = 9$时的抽样图案，$I = 18.0285$

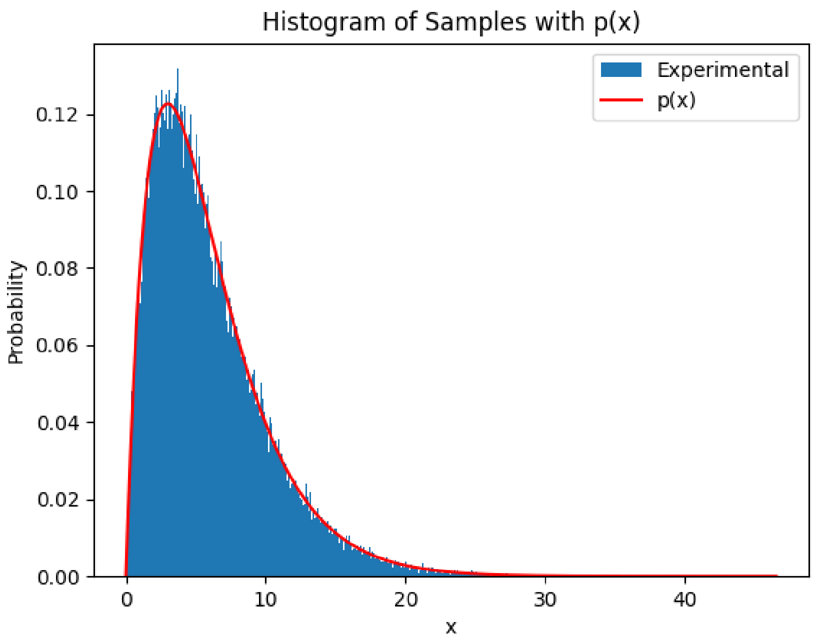

图5：$\gamma=12$时的抽样图案，$I=17.8148$

由抽样图案可以看出，抽样出来的样本X与$p(x)$拟合效果非常好

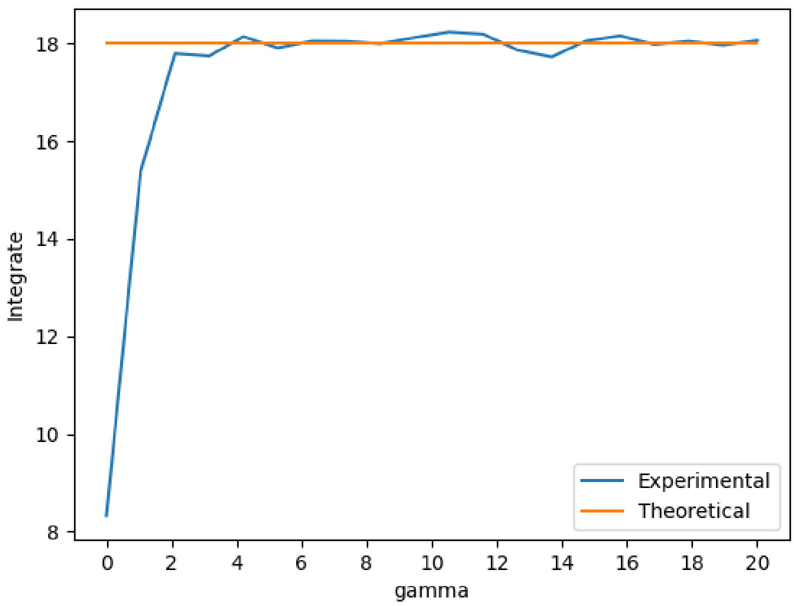

图6：取不同的γ值得到的积分结果

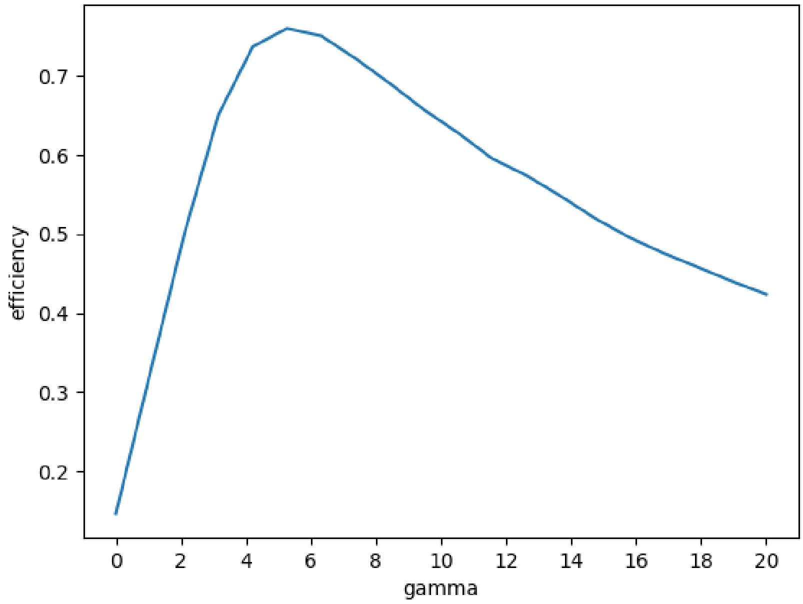

图7：取不同的γ值得到的抽样效率图

由图6与图7可以看出：当γ值较小时，抽样效率低，导致无法很好地按照$p(x)$进行抽样，导致积分结果偏差较大；当$\gamma \approx 6$时，抽样效率达到极值，此时积分结果也趋于稳定，在18附近微微跳动；当γ值继续增大时，x'逐渐变得更大，但由于$gamma$函数的性质，较大的抽样值选取概率降低，因此抽样效率下降。

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

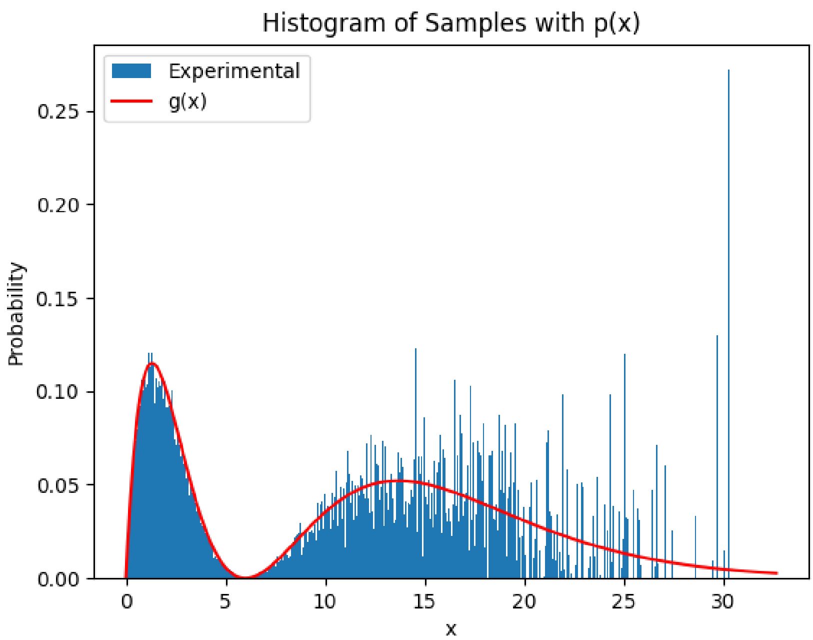

图8：$\gamma = 3$时的抽样图案

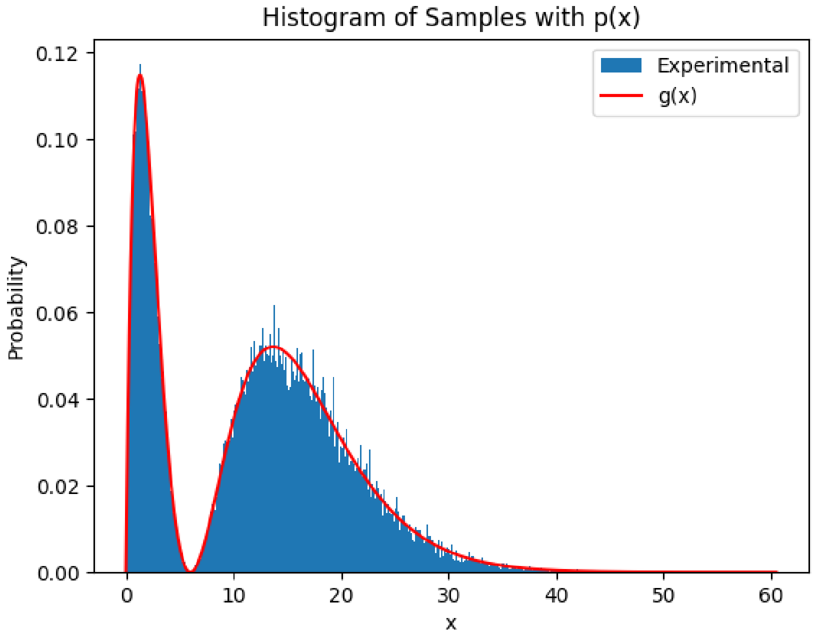

图9：$\gamma = 8$时的抽样图案

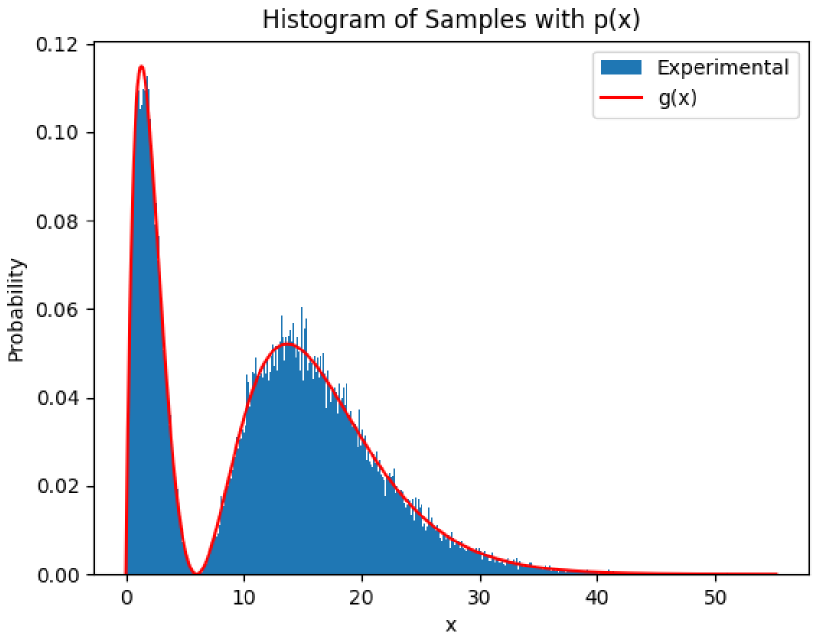

图10：$\gamma = 13$时的抽样图案

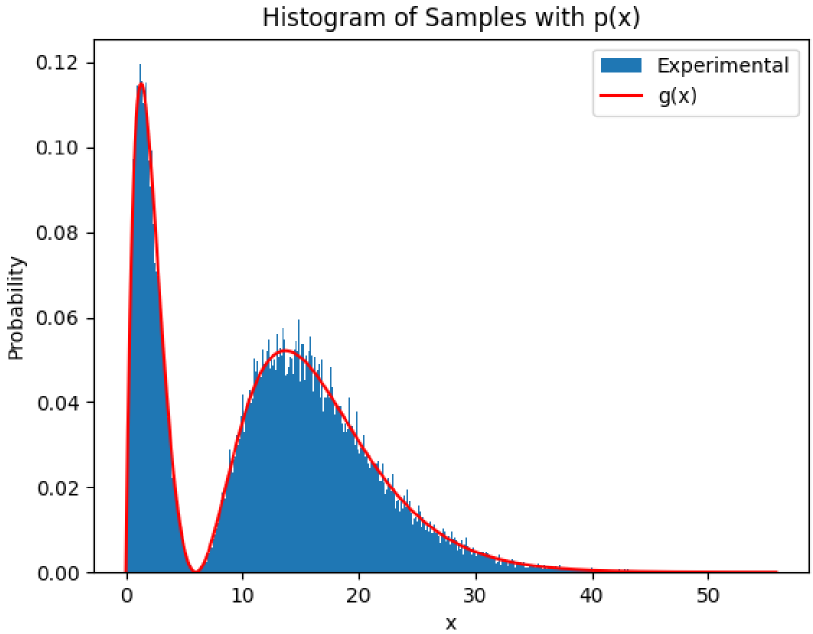

图11：$\gamma = 18$时的抽样图案

由抽样图案可以看出，样本X与$g(x)$符合较好，且由图可以看出$[1,5]$和$[10,20]$区间内，g(x)的值相对较大，故取这两个区间内的样本X的统计值来估计$g(x)$，结果如下：

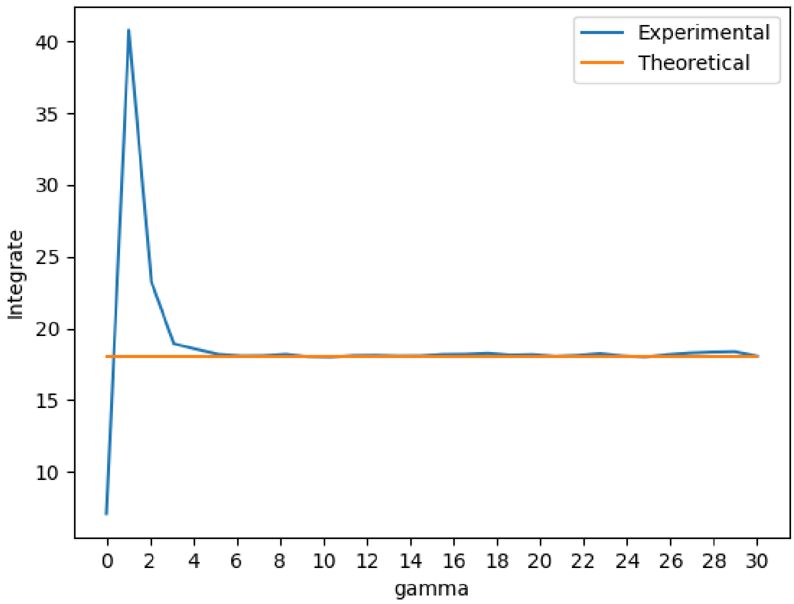

图12：取不同的γ值得到的积分结果

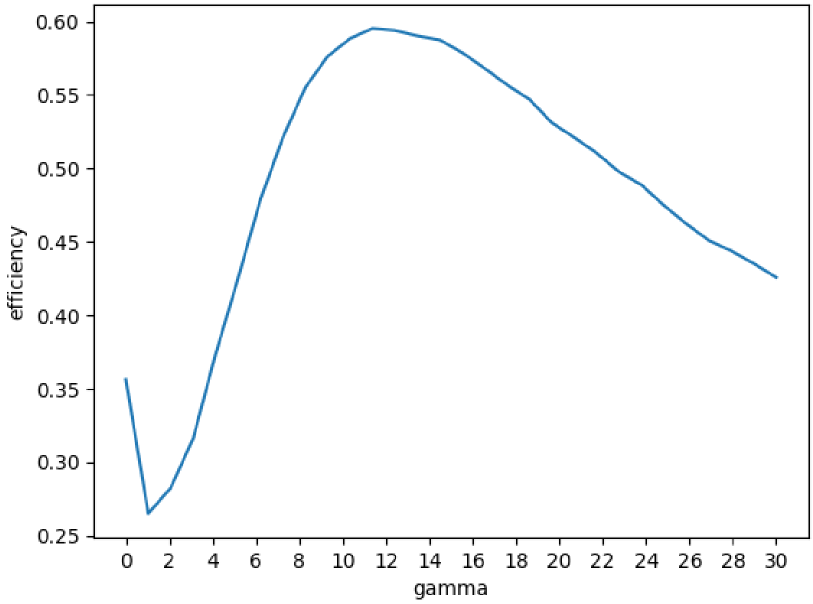

图13：取不同的γ值得到的抽样效率

分析同第一种情况类似，γ较小时，由于抽样效率低导致积分结果偏差较大；在$\gamma = 12$附近时，抽样效率达到极值，积分结果也趋于稳定。

四、总结

本题细致了解了$Metropolis - Hasting$方法在求解积分值时的应用。

在求解数值积分时，选择合适的$T(x)$尤为重要，在本题中可以看出不同的γ值对结果的影响，因此我们需要对函数性质有充分的了解后，选择恰当的数值，能显著提升结果的精度和效率