

R-Lab

吴海旭 2016013223 软件61

R-Lab

- 1. 实验描述
- 2. 实验方案
- 3. 实验结果
- 4. 结果分析
 - 4.1 生成结果正确性
 - 期望
 - 分布
 - 4.2 抽样效率

1. 实验描述

指定分段函数为概率密度函数，进行随机变量的生成。

分段函数如下：

$$\tilde{f}(x) = \begin{cases} \frac{\sqrt{x(2-x)}}{\frac{\pi}{2}+2}, & x \in [0, 2], \\ \frac{x-2}{\frac{\pi}{2}+2}, & x \in (2, 4], \end{cases}$$

2. 实验方案

实验使用舍取法进行数据生成。具体实现如下：

- 首先，生成[0,4]的随机变量 `x` 与 [0,1]的随机变量 `y`
- 判断下式是否成立

$$y > \tilde{f}(x)$$

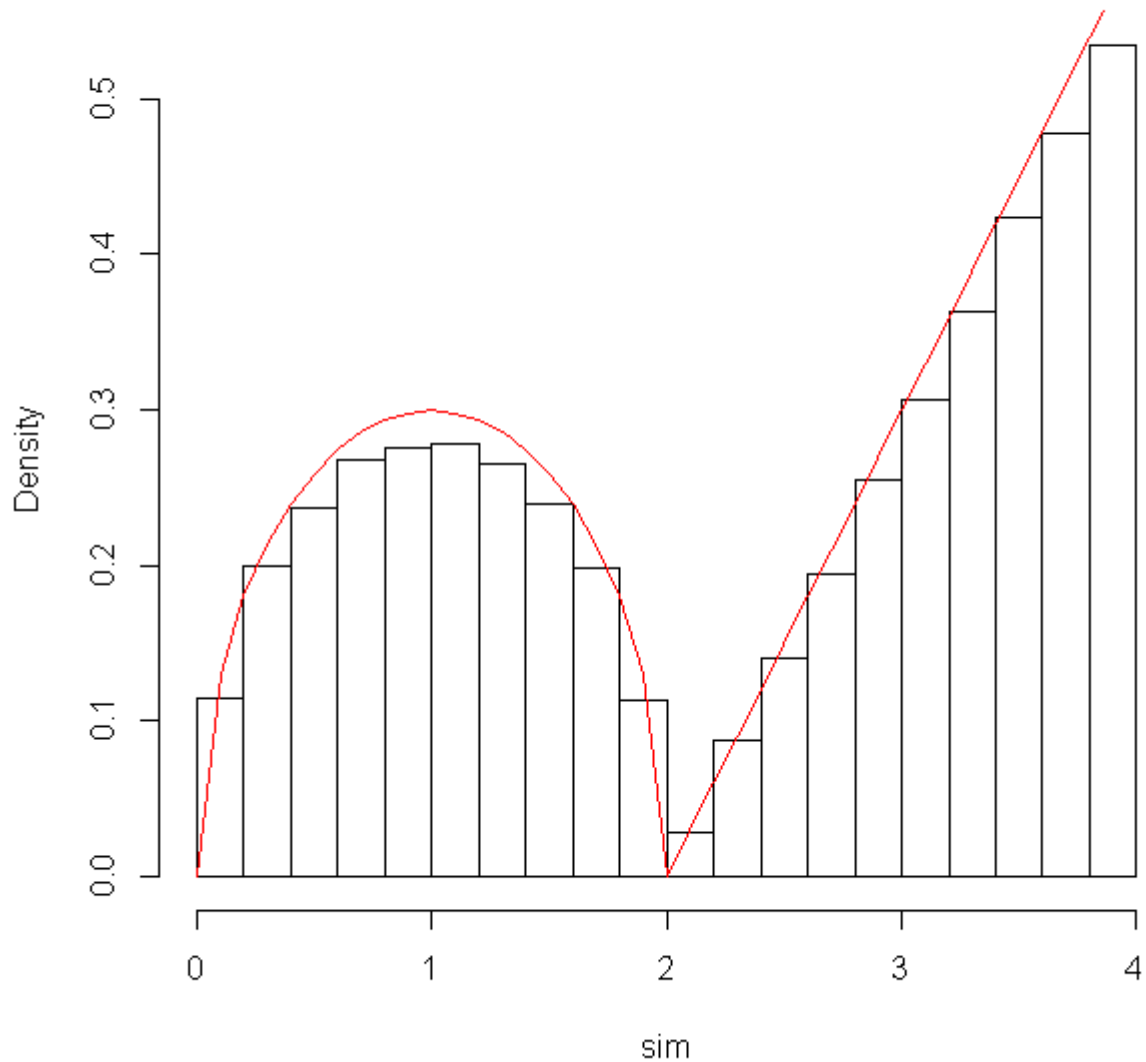
- 若成立，则重复生成 `x` 和 `y`
 - 若不成立，则保存 `x`
- 重复上述过程，获取到一定数量的随机变量 `x`
- 绘制图像

3. 实验结果

我们通过上述实验方案获得以下实验结果。

生成随机变量 `x` 100000个。

Histogram of sim



- 黑色部分为生成变量 x 的频率直方图。
- 红色部分为概率密度函数的图像。

4. 结果分析

4.1 生成结果正确性

期望

随机生成的100000个变量均值为2.30601

而我们实际计算变量的期望为

$$\begin{aligned}\int_0^2 x \sqrt{x(2-x)} dx + \int_2^4 x(x-2) dx &= \int_{\frac{\pi}{2}}^0 2 \cos^2 \theta \sqrt{2 \cos^2 \theta (2 - 2 \cos^2 \theta)} d(2 \cos^2 \theta) + \frac{20}{3} \\ &= \frac{\pi}{2} + \frac{20}{3}\end{aligned}$$

最中结果为

$$\frac{\frac{\pi}{2} + \frac{20}{3}}{\frac{\pi}{2} + 2} \approx 2.30689$$

非常接近，所以可以认为期望拟合正确。

分布

根据上图，我们发现在随机变量数目为100000时，概率密度函数图像与频率分布图基本吻合。

所以，随机变量生成的正确性得到验证。

4.2 抽样效率

num 为数据量， count of operation 为产生随机数的操作数目。

num	count of operation
100000	400790

生成数据的效率大概为25%

这是因为

$$\int_0^2 \sqrt{x(2-x)}dx + \int_2^4 (x-2)dx = \frac{\pi}{2} + 2$$
$$\frac{4(\frac{\pi}{2} + 2)}{\frac{\pi}{2} + 2} = 4$$

即，因为横坐标为均匀分布，故效率为可以选中的概率均值，即为概率密度函数的面积与总体面积的比值，为25%。