

STAT3003: 统计算法基础

蒙特卡洛方法 (1)

曾靖

中国科学技术大学管理学院

大纲

大纲

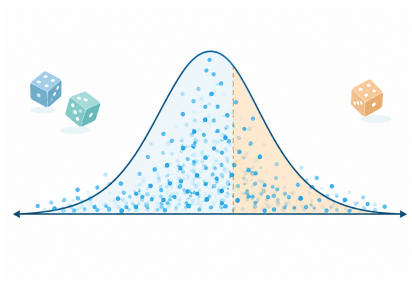
1. 蒙特卡洛方法
2. 蒙特卡洛积分

本节内容来自于《统计计算使用 R 》第 5 章,《统计方法》3.2 小节

1 蒙特卡洛方法

1.1 蒙特卡洛方法

- Monte Carlo 方法又称随机模拟方法



- Monte Carlo 方法（或称 Monte Carlo 实验）是一类广泛的计算算法，依靠重复随机抽样获得数值结果。
- 其核心思想是利用随机性解决原则上可以是确定性的问题。

1.2 历史

- 20 世纪 40 年代末, *Stanislaw Ulam* 在洛斯阿拉莫斯国家实验室研究核武器项目中的中子扩散问题。传统确定性方法难以处理这一问题, 他提出用随机实验。
- 1947 年 3 月, *John von Neumann* 提议用统计方法和计算机研究裂变材料中的中子扩散; *Nicholas Metropolis* 以 Monte Carlo 赌场为灵感提出这一名称。
- 20 世纪 50 年代, Monte Carlo 方法成为曼哈顿计划和氢弹研制中的重要模拟工具, 随后推广到物理、物理化学和运筹学。

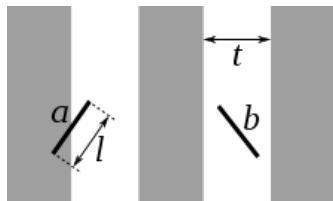
蒙特卡洛方法：一场随机的突破



1.3 蒲丰投针问题

Monte Carlo 最早起可用于蒲丰投针问题来近似估计 π :

假设地面由宽度相同、相互平行的木条组成，每条木条宽度为 t 。现将一根长度为 ℓ 的针随机投到地面上。求这根针跨过两条木条之间分界线的概率。



1.3 蒲丰投针问题：解答

设针长 $\ell \leq t$ 。令 h 为针中心到最近分界线的距离， Θ 为针与木条方向的夹角。则

$$h \sim \text{Unif}(0, t/2), \quad \Theta \sim \text{Unif}(0, \pi/2),$$

且二者相互独立。

针跨过分界线的事件记为 A 。给定 $\Theta = \theta$ 时，

$$\Pr(A \mid \Theta = \theta) = \Pr\left(h \leq \frac{\ell}{2} \sin \theta\right) = \frac{\ell}{t} \sin \theta.$$

因此

$$\Pr(A) = \frac{2}{\pi} \int_0^{\pi/2} \frac{\ell}{t} \sin \theta \, d\theta = \frac{2\ell}{\pi t}.$$

若独立投针 N 次，记跨线次数为 K ，则

$$\hat{p} = \frac{K}{N}, \quad \hat{\pi} = \frac{2\ell}{t\hat{p}} = \frac{2\ell N}{tK} \quad (K > 0).$$

1.4 蒙特卡洛方法流程

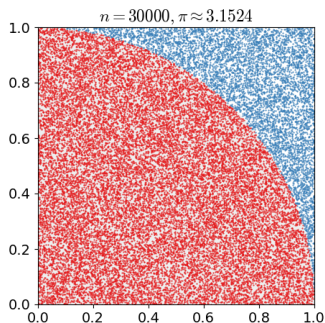
◦ Monte Carlo 方法的主要流程如下：

1. 把某问题转化为求解概率模型问题；
2. 产生大量符合概率模型的随机样本；
3. 对产生的随机样本进行分析从而解决问题。

核心思想是利用随机方法解决非随机的问题，而这些问题往往很难用其他方法求解。

例 1.1: 圆周率 π 的估计

$$\begin{aligned} & \Pr(\text{A randomly dropped point is located in the sector}) \\ &= \frac{\text{Area of the sector}}{\text{Area of the square}} = \frac{\pi^2}{4}. \end{aligned}$$



1.5 应用

- 数值积分：将积分转化为期望，用随机样本估计积分值。
- 参数估计：通过重复抽样评估估计量的均方误差。
- 置信区间：通过重复试验估计区间的覆盖率。
- 假设检验：估计一类错误率和检验功效。
- 贝叶斯推断：利用 MCMC 从后验分布中抽样。

2 Monte Carlo 积分

2.1 Monte Carlo 积分

Question: 对于函数 $g(x)$, 如果积分 $\theta = \int_A g(x)dx$ 存在, 如何通过数值方法计算该积分?

1. 假设我们可以找到一个随机变量 X , 其密度函数为 $f(x)$, 支撑集为 A , 使得

$$\theta \propto \int_A g(x)f(x)dx = E(g(X)).$$

2. 利用蒙特卡罗方法 (Monte-Carlo) 估计 $E(g(X))$, 进而得到积分 θ 的估计。

算法流程:

- 生成 X 的 i.i.d. 样本 $x_1, \dots, x_m$;
- 计算样本均值

$$\hat{\theta} = \frac{1}{m} \sum_{i=1}^m g(x_i).$$

根据大数定律可知, $\hat{\theta}$ 是积分 θ 的无偏估计。

2.2 简单 Monte Carlo 积分

简单 Monte Carlo 积分是基于均匀分布的 Monte Carlo 积分方法。

- 考虑积分 $\theta = \int_a^b g(x)dx$.
- 令 $X \sim \text{Unif}(a, b)$, 则 $g(X)$ 的期望为

$$E(g(X)) = \frac{1}{b-a} \int_a^b g(x)dx,$$

即

$$\theta = (b-a)E(g(X)).$$

- 设 $x_1, \dots, x_m$ 服从均匀分布 $\text{Unif}(a, b)$, 那么 Monte Carlo 估计量为

$$\hat{\theta} = (b-a) \frac{1}{m} \sum_{i=1}^m g(x_i).$$

例 2.1

计算

$$\theta = \int_2^4 e^{-x} dx$$

```
m <- 10000  
x <- runif(m, min=2, max=4)  
theta.hat <- mean(exp(-x)) * 2  
print(theta.hat)
```

```
## [1] 0.117247
```

```
print(exp(-2) - exp(-4))
```

```
## [1] 0.1170196
```

例 2.1 (续)

利用指数分布直接生成样本

```
m <- 10000
x <- rexp(m)
theta.hat <- mean((x >= 2) & (x <= 4))
print(theta.hat)
```

```
## [1] 0.1165
```

```
print(exp(-2) - exp(-4))
```

```
## [1] 0.1170196
```

两种方法的积分估计

假设 $g(x)$ 是容易抽样的密度函数，考虑

$$\theta = \int_a^b g(x)dx.$$

- 方法 1: 简单 *Monte Carlo* 积分 (均匀分布)。生成 $X_1, \dots, X_m \sim \text{Unif}(a, b)$, 使用

$$\hat{\theta}_U = (b - a) \frac{1}{m} \sum_{i=1}^m g(X_i).$$

- 方法 2: 直接从 g 抽样。生成 $X_1, \dots, X_m \sim g$, 使用

$$\hat{\theta}_g = \frac{1}{m} \sum_{i=1}^m I(a \leq X_i \leq b).$$

例 2.2

Exercise: 基于如下分布，使用 Monte Carlo 方法计算积分

$$\theta = \int_0^1 \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt.$$

- 均匀分布
- 正态分布

Question: 对于不同的样本生成方法和积分估计方法，如何评估哪种方法更好？

2.3 方差和效率

方差是比较两种无偏估计量优劣的评判标准之一。

- 对于两个 θ 的无偏估计量 $\hat{\theta}_1, \hat{\theta}_2$ ，如果

$$\text{Var}(\hat{\theta}_1) \leq \text{Var}(\hat{\theta}_2),$$

我们称 $\hat{\theta}_1$ 更有效率。

2.4 方差估计

Question: 如何估计一个 *Monte Carlo* 估计量 $\hat{\theta}$ 的方差?

1. 对于平均值估计量, 可以利用样本均值的方差公式。若 $\hat{\theta} = m^{-1} \sum_{i=1}^m h(X_i)$, 则

$$\text{Var}(\hat{\theta}) = \frac{\text{Var}(h(X))}{m}.$$

用 $h(X_i)$ 的样本方差估计 $\text{Var}(h(X))$ 。

2. 对于一般估计量, 则重复 *Monte Carlo* 实验。独立重复整个估计过程 m' 次, 得到 $\hat{\theta}^{(1)}, \dots, \hat{\theta}^{(m')}$, 用这些估计量的样本方差估计 $\text{Var}(\hat{\theta})$:

$$\bar{\hat{\theta}} = \frac{1}{m'} \sum_{j=1}^{m'} \hat{\theta}^{(j)}, \quad \widehat{\text{Var}}(\hat{\theta}) = \frac{1}{m' - 1} \sum_{j=1}^{m'} \left(\hat{\theta}^{(j)} - \bar{\hat{\theta}} \right)^2.$$

2.4 方差估计

对前面的两个平均值估计量 $\hat{\theta}_U$ 和 $\hat{\theta}_g$,

$$\widehat{\text{Var}}(\hat{\theta}_U) = \frac{(b-a)^2}{m} \widehat{\text{Var}}(g(X)), \quad \widehat{\text{Var}}(\hat{\theta}_g) = \frac{\hat{\theta}_g(1-\hat{\theta}_g)}{m}.$$

其中, 令 $y_i = g(X_i)$, 则

$$\widehat{\text{Var}}(g(X)) = \frac{1}{m} \sum_{i=1}^m (y_i - \bar{y})^2, \quad \bar{y} = \frac{1}{m} \sum_{i=1}^m y_i.$$