# 第3章 随机模拟

- **3.1 随机模拟介绍**
- **3.2 随机模拟积分**
- **3.3 重要抽样法**
- **3.4 分层抽样法**
- **3.5 方差缩减技术**
- **3.6 随机服务系统模拟**
- **3.7 Bootstrap方法**
- **3.8 MCMC**

# 3.1 随机模拟介绍

# 数学建模

- 忽略许多细节
- 增加一些可能很难验证的理论假设
- 使得模型比较简单
- 用数学理论进行分析研究

# 数学建模缺点

- 与实际情况有较大的差距
- 即使我们对模型进行了完美的理论分析，也不能保证分析结果是可信的
- 用随机模拟的方法解决

# 3.1 随机模拟介绍

- **模拟**

是指把某一现实的或抽象的系统的某种特征或部分状态， 用另一系统（称为模拟模型）来代替或模拟。

- **随机模拟法， 蒙特卡洛(Monte Carlo) 法**

为了解决某问题， 把它变成一个**概率模型**的求解问题， 然后产生符合模型的大量随机数， 对产生的随机数进行分析从而求解问题。

# 例 红绿灯信号最优控制

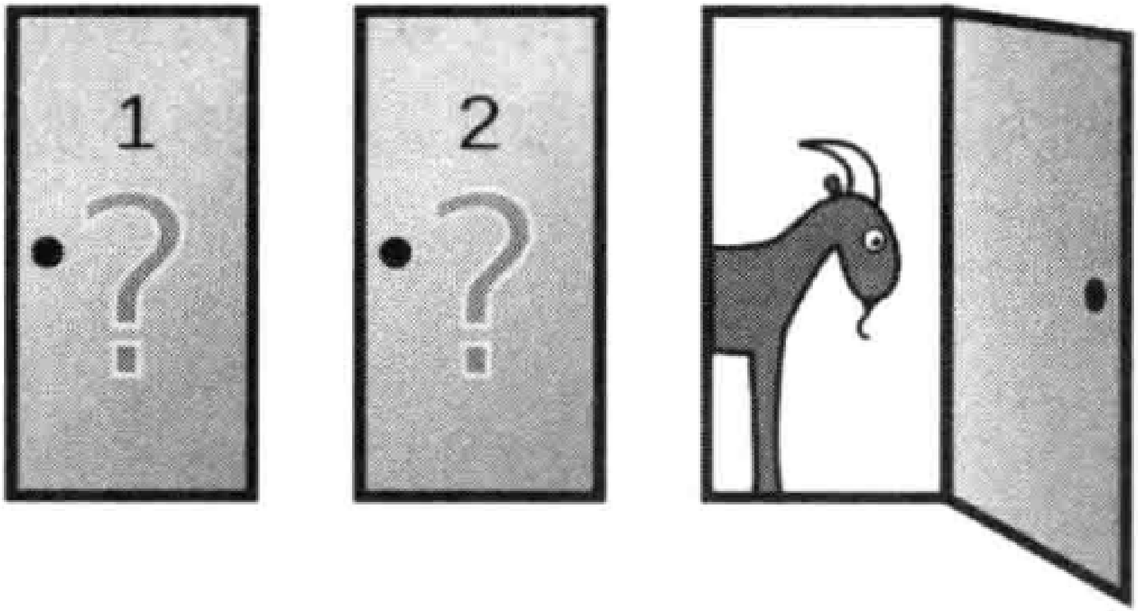
交通路口要找到一种**最优**的控制红绿灯信号的办法，使得：通过路口的汽车耽搁的平均时间最短，而行人等候过路的时间不超过某一给定的心理极限值

**解：**

- 用随机模拟方法
- 模拟汽车和行人到来的过程，
- 通过统计比较找出最优的控制方案

# 例 蒙提霍尔问题

- 蒙提霍尔问题(Monty Hall problem)
- 也称为三门问题



# 例 蒙提霍尔问题

参赛者面前有三扇关闭的门,其中一扇门的后面藏有一辆汽车,而另外两扇门的后面则各藏有一只山羊.参赛者从三扇门中随机选取一扇,若选中后面有车的那扇门就可以赢得该汽车.当参赛者选定了一扇门,但尚未开启它的时候,节目主持人会从剩下两扇门中打开一扇藏有山羊的门,然后问参赛者要不要更换自己的选择,选取另一扇仍然关上的门.

- 这个游戏涉及的问题是:参赛者更换自己的选择是否会增加赢得汽车的概率?

# 例 蒙提霍尔问题

## 随机模拟法

- 研究参赛者采用更换选择的策略是否有利?
- 在该策略下赢得汽车的机会有多大?
- 将两只山羊分别编号为"1"和"2",
- 将汽车编号为"3".

# 例 蒙提霍尔问题

## 随机模拟法

- 使用掷骰子的方法从数字1、2、3中随机选取一个数字.

- 骰子的点数为1和2的面对应于这里的数字"1",
- 骰子的点数为3和4的面对应于这里的数字"2",
- 骰子的点数为5和6的面对应于这里的数字"3".

# 例 蒙提霍尔问题

每次试验如此进行: 掷一次骰子,

- 若掷出的骰子是点数为1到4的面之一。

表明参赛者选中了1或2, 则采取更换选择的策略就必选中3, 即赢得汽车;

- 若掷出的骰子是点数为5或6的面之一。

表示他选中了3, 则采取更换选择的策略使得他失去了赢得汽车的可能.

# 例 蒙提霍尔问题

- 将这样的试验重复进行n次,
- 记录下在这些试验中掷出的骰子是点数为1到4的面的次数m

- m表示了采取更换选择策略而赢得汽车的次数

| 试验次数 | 10 | 100 | 1000 | 10000 | 100000 | 1000000 |
|---|---|---|---|---|---|---|
| 机会 $(m/n)$ | 0.7000 | 0.6600 | 0.6650 | 0.6600 | 0.6663 | 0.6666 |

# 随机模拟方法一般思路

应用随机模拟方法的关键步骤就是构造该问题的模拟系统

- 确定该模拟系统的状态是什么?

- 存在什么随机因素?

- 系统会出现哪些可能的结果?

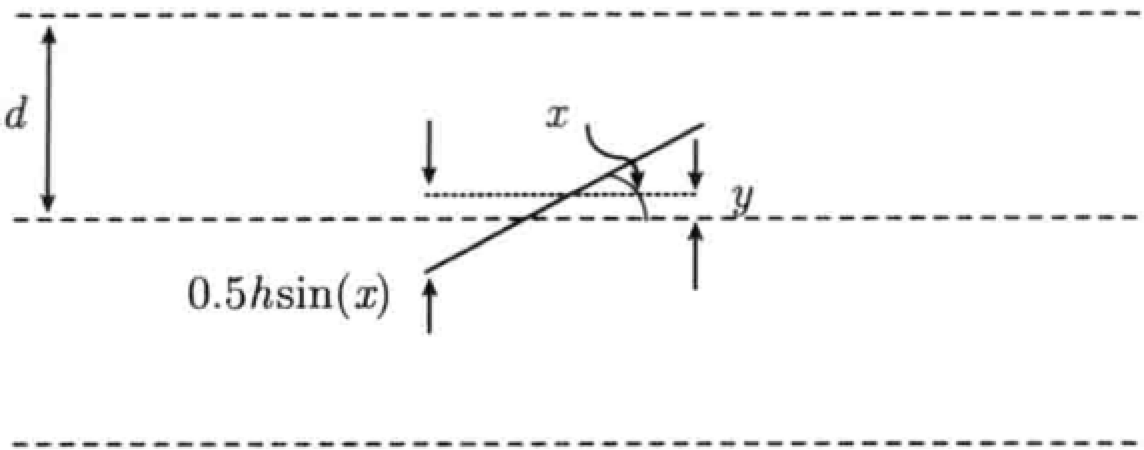
- 需要计算的系统的性能指标是什么?

# 例10.1 (1用蒲丰投针法求圆周率)

通过将针投到画有一组等间隔平行线的地面来统计针与线条相交的频率,并运用已知的数学关系式计算出圆周率π的值

| 实验者 | 年　份 | 投针次数 | π 的实验值 |
|---|---|---|---|
| 沃尔弗 (Wolf) | 1850 | 5000 | 3.1596 |
| 斯密思 (Smith) | 1855 | 3204 | 3.1553 |
| 福克斯 (Fox) | 1894 | 1120 | 3.1419 |
| 拉查里尼 (Lazzarini) | 1901 | 3408 | 3.1415929 |

# 例10.1 (1用蒲丰投针法求圆周率)




平面上画有间隔为d(d>0)的等距平行线,向平面内任意投掷一枚长为h(h<d)的针,统计针与任一平行线发生相交的**频率**

- 用y表示针的中点与最近的一条线的距离
- 用x表示针与此直线间的夹角

# 例10.1 (1用蒲丰投针法求圆周率)

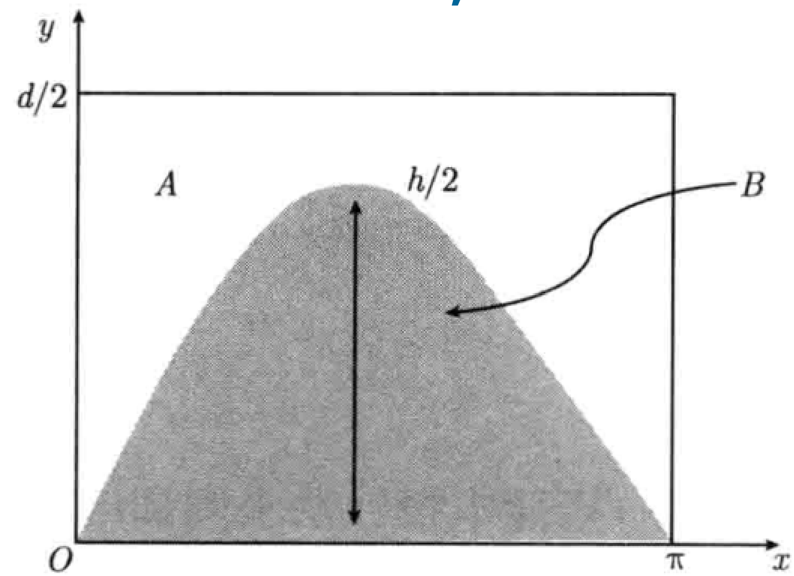
针与平行线发生相交的**概率**：

$$R=\frac{S_B}{S_A}=\frac{\int_0^{\pi}\frac{h}{2}\sin(x)dx}{d\,\pi/2}=\frac{2h}{d\pi}$$

另一方面，设m为针与水平线相交的次数,N为投针的总次数，则R=m/N

- 计算圆周率近似值为

$$\pi\approx\frac{2hN}{dm}$$



# 例10.1 (2用随机模拟估计圆周率)

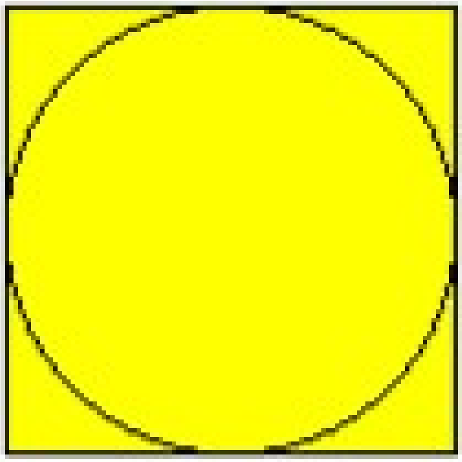
向正方形

$$D=\{(x,y): x\in[-1,1], y\in[-1,1]\}$$

内随机等可能投点， 落入单位圆

$$C=\{(x,y): x^2+y^2\le 1\}$$

的概率为面积之比 $p=\dfrac{\pi}{4}$



# 例10.1 (2用随机模拟估计圆周率)

如果独立重复地投了N个点
落入C中的点的个数 $\xi$

则由概率与频率近似相等，有

$$\frac{\xi}{N} \approx \frac{\pi}{4}$$

即

$$\pi \approx \hat{\pi} \approx \frac{4\xi}{N}$$

# 随机模拟误差

- 随机模拟方法结果的**随机性**导致的误差叫**随机模拟误差**

例10.1 $\xi \sim B(N, \frac{\pi}{4})$

方差 $Var(\hat{\pi}) = \frac{\pi(4-\pi)}{N}$

由中心极限定理 $\hat{\pi} \sim N(\pi, \frac{\pi(4-\pi)}{N})$

随机模拟误差的幅度大约在 $\pm 2\sqrt{\frac{\pi(4-\pi)}{N}}$

# 随机模拟方法特点

- 应用面广，适应性强
- 算法简单， 容易改变条件，但计算量大
- 结果具有随机性，精度较低
- 模拟结果的收敛服从概率论规律
- 对维数增加不敏感

# 随机模拟应用

- 探索性试验
- 说明新的模型或技术的有效性
- 蒙特卡洛检验
- bootstrap置信区间
- bootstrap偏差修正
- MCMC

# 3.2 随机模拟积分

3.2.1 随机投点法
3.2.2 平均值法
3.2.3 随机模拟估计误差的评估与样本量计算
3.2.4 高维定积分

# 3.2 随机模拟积分

- 用随机模拟解决非随机问题
- 也称为蒙特卡洛(Monte Carlo)积分
- 简称MC积分

# 3.2 随机模拟积分

- 双期望定理 $E[E(X|Y)] = EX$

证明：（连续型）

$$
\begin{aligned}
E[E[X \mid Y]] &= \int E[X \mid Y = u]P_Y(u))du \\
&= \int [\int tP_X(X = t|Y = u)dt]P_Y(u)du \\
&= \int \int tP_X(X = t \mid Y = u)P_Y(u)dtdu \\
&= \int t[\int P_X(t \mid Y = u)P_Y(u)du]dt \\
&= \int t[\int P_{XY}(t, u)du]dt \\
&= \int tP_X(t)dt \\
&= E[X]
\end{aligned}
$$

# 3.2 随机模拟积分

- 方差分解公式

$$Var(X) = \mathrm{E}[Var(X \mid Y)] + Var(\mathrm{E}[X \mid Y])$$

证明：

$$\begin{aligned}
&E[var(X|Y)] + var[E(X|Y)] \\
&= E[E(x^2|y) - E^2(x|y)] + E\left(E^2(x|y)\right) - \left[E\left(E(x|y)\right)\right]^2 \\
&= E[E(x^2|y)] - EE^2(x|y) + EE^2(x|y) - \left[E\left(E(x|y)\right)\right]^2 \\
&= Ex^2 - (Ex)^2 \\
&= var(X)
\end{aligned}$$

# 3.2 随机模拟积分

- 条件方差公式

$$
\begin{aligned}
&Var(X|Y)\\
&= E[(X - E(X|Y))^2|Y]\\
&= E(X^2|Y) - 2E(X * E(X|Y)|Y) + \left(E(X|Y)\right)^2\\
&= E(X^2|Y) - \left(E(X|Y)\right)^2
\end{aligned}
$$

# 3.2.1 随机投点法

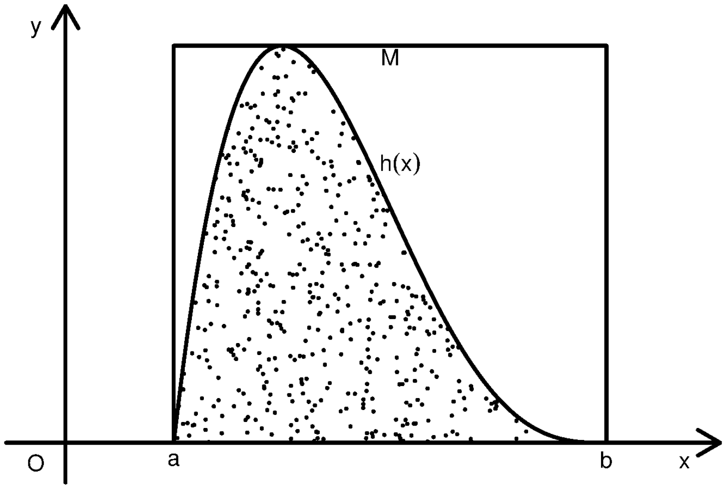
- 计算定积分 $I = \int_a^b h(x)dx$

其中函数*h(x)*在有限区间[a,b]上定义且有界

$$0 \leqslant h(x) \leqslant M$$

- 相当于计算曲线下的区域D的面积

$$D = \{(x,y): 0 \le y \le h(x), x = [a,b]\}$$



# 3.2.1 随机投点法

- 在 $G = [a, b] \times (0, M)$ 上均匀抽样N次
- 得随机点

$$Z_1, Z_2, \ldots, Z_N, Z_i = (X_i, Y_i), (i = 1,2, \ldots, N)$$

令

$$\xi_i = \begin{cases} 1, & Z_i \in D \\ 0, & \text{otherwise} \end{cases}, \quad (i = 1, \ldots, N)$$

是独立重复试验结果.

$$\{\xi_i\} iid \sim b(1, p)$$

而

$$p = P(Z_i \in D) = V(D) / V(G) = I / [M(b-a)] \qquad (11.1)$$

其中V(.)表示区域面积

# 3.2.1 随机投点法

- 从模拟产生的随机样本 $Z_1, Z_2, \dots, Z_N$
- 用这N个点中落入曲线下方区域D的百分比 $\hat{p}$ 来估计(11.1)中的概率$p$
- 由$I$=$p$M(b-a)得到定积分$I$的近似值

$$\hat{I} = \hat{p}M(b-a) \qquad (11.2)$$

# 随机投点法步骤

(1)独立地产生U(0,1)随机数:

$$u_i, v_i \ (i=1\ldots,N)$$

(2)计算$x_i=a+u_i(b-a)$,$y_i=Mv_i$,和$f(x)$

(3)统计$f(x_i)\geq y_i$的个数,计算 $\hat{p}$

(4)用(11.2)估计I.

$$\hat{I} = \hat{p}M(b-a) \qquad (11.2)$$

# 3.2.1 随机投点法

- 由强大数律可知

$$\hat{p} = \frac{\sum \xi_i}{N} \to p, \text{a.s.} \ (N \to \infty)$$

$$\hat{I} = \hat{p} M(b-a) \to pM(b-a) = I, \text{a.s.} \ (N \to \infty)$$

$N \to \infty$ 时，精度可以无限地提高

# 3.2.1 随机投点法

- 这样计算的定积分有随机性
- 误差中包含了**随机模拟误差**
- 由中心极限定理可知

$$\sqrt{N}(\hat{p}-p)/\sqrt{p(1-p)} \xrightarrow{d} \mathrm{N}(0,1), (N \to \infty)$$

$$\sqrt{N}(\hat{I}-I) = \sqrt{N}M(b-a)(\hat{p}-p) \xrightarrow{d} \mathrm{N}(0, [M(b-a)]^2 p(1-p)) \quad (11.3)$$

- 当N很大时，$\hat{I}$ 近似服从分布

$$\mathrm{N}(I, [M(b-a)]^2 p(1-p)/N)$$

称此近似分布的方差 $[M(b-a)]^2 p(1-p)/N$ 为 $\hat{I}$ 的**渐近方差**

# 3.2.2 平均值法

- 计算定积分 $$I = \int_a^b h(x)dx$$
- 取 $U \sim \mathrm{U}(a,b)$

期望 $$E[h(U)] = \int_a^b h(u)\frac{1}{b-a}du = \frac{I}{b-a}$$

则

$$I = (b-a) \cdot Eh(U)$$

## 3.2.2 平均值法

- 若取

$$\{U_i, i = 1, \ldots, N\} iid \sim U(a, b)$$

则

$$Y_i = h(U_i), (i = 1,2, \ldots \mathrm{N})$$

是iid随机变量列。

# 3.2.2 平均值法

- 由强大数律

$$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} h(U_i) \to Eh(U) = \frac{I}{b-a}, \quad \text{a.s.}\,(N \to \infty)$$

于是

$$\hat{I} = \frac{b-a}{N}\sum_{i=1}^{N} h(U_i) \qquad (11.4)$$

是$I$的强相合估计。

# 平均值法步骤

(1)独立地产生随机数N个U(0,1)随机数:

$u_i$, (i=1…,N)

(2)计算$x_i=a+u_i(b-a)$,和$h(x_i)$

(3) 用(11.4)估计I.

$$\hat{I}=\frac{b-a}{N}\sum_{i=1}^{N}h\left(X_i\right) \qquad (11.4)$$

# 3.2.2 平均值法

- 由中心极限定理有

$$\sqrt{N}(\hat{I}-I) \xrightarrow{d} \mathrm{N}(0,(b-a)^2\mathrm{Var}(h(U)))$$

其中

$$Var[h(U)] = \int_a^b \left[h(u) - Eh(U)\right]^2 \frac{1}{b-a} du$$

仅与h有关，有

$$\hat{I} - I = O_p(\frac{1}{\sqrt{N}})$$

# 3.2.2 平均值法

- **求随机变量函数期望**

求一般随机变量$X$的函数$h(X)$的期望$I=\mathrm{E}h(X)$

对$X$的随机数$X_i$ (i=1,2,…,N), 令$Y_i = h(X_i)$

也可以用平均值法来估计$\mathrm{E}h(\mathrm{X})$

$$\hat{I} = \frac{1}{N}\sum_{i=1}^{N} h(X_i)$$

是$I$的无偏估计和强相合估计

# 例11.1

设X~N(0,1)，求 $I = E|X|^{\frac{3}{2}}$

解：

用平均值法估计I, 取抽样样本量N=10000,

产生标准正态随机数$X_i$ (i=1,2,...,N),令 $Y_i = |X_i|^{\frac{3}{2}}$

则 $\hat{I} = \overline{Y} = \frac{1}{N}\sum_{i=1}^{N}|X_i|^{\frac{3}{2}}$ $S_N^2 = \frac{1}{N}\sum_{i=1}^{N}(|X_i|^{\frac{3}{2}} - \hat{I})^2$

一次模拟的结果得到 $\hat{I}$=0.8754,

$S_N$=0.09409

# 3.2.3 随机模拟估计误差的评估与样本量计算

对 $\theta = EY$

设 $\mathrm{Var}(Y) = \sigma^2$

Yi(i=1,...,N)独立同分布,是来自Y的样本

则θ的估计 $\hat{\theta} = \bar{Y}_N = \frac{1}{N}\sum_{i=1}^{N} Y_i$ (11.7)

令 $S_N^2 = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \bar{Y}_N)^2$

则 $N \to \infty, \hat{\theta} \to \theta \; a.s.$

$S_N^2 \to \sigma^2 \; a.s.$

# 3.2.3 随机模拟估计误差的评估与样本量计算

并且 $$\hat{\theta} \sim \mathrm{N}\left(\theta, \frac{\sigma^2}{N}\right) a.s. \qquad (11.8)$$

若$\sigma^2$可以用$S_N^2$近似，则

$$\frac{\hat{\theta}-\theta}{S_N / \sqrt{N}} \xrightarrow{\mathrm{d}} \mathrm{N}(0,1)$$

所以N充分大时

$$\hat{\theta} \sim \mathrm{N}\left(\theta, \frac{S_N^2}{N}\right) a.s.$$

# 估计误差的一些度量: 估计的根均方误差

定义估计的根均方误差为

$$\text{RMSE} = \sqrt{E(\hat{\theta} - \theta)^2}. \tag{11.9}$$

在以上用独立同分布样本均值估计期望值$\theta = EY$的做法中，因为$E\hat{\theta} = \theta$，所以

$$\text{RMSE} = \sqrt{\text{Var}(\hat{\theta})} = \sqrt{\text{Var}(\bar{Y}_n)} = \frac{\sigma}{\sqrt{N}}, \tag{11.10}$$

- 根均方误差可以用$\frac{S_N}{\sqrt{N}}$估计

# 估计误差的一些度量:估计的平均绝对误差

定义估计的平均绝对误差为

$$\text{MAE} = E\left|\hat{\theta} - \theta\right|. \tag{11.11}$$

当N充分大时由$\hat{\theta}$的近似正态分布式(11.8)可知

$$\text{MAE} = E\left|\hat{\theta} - \theta\right| = \frac{\sigma}{\sqrt{N}} E\left|\frac{\hat{\theta} - \theta}{\sigma/\sqrt{N}}\right| \approx \sqrt{\frac{2}{\pi}} \frac{\sigma}{\sqrt{N}} \approx 0.7979 \frac{\sigma}{\sqrt{N}} \tag{11.12}$$

- 平均绝对误差可以用$0.80 \frac{S_N}{\sqrt{N}} = 0.80\text{RMSE}$估计

# 估计误差的一些度量:估计的平均相对误差

定义估计的平均相对误差为

$$\mathrm{MRE} = E\left|\frac{\hat{\theta}-\theta}{\theta}\right| = \frac{\mathrm{MAE}}{\theta} \tag{11.11}$$

- 平均相对误差可以用$0.80\,\frac{s_N}{\sqrt{N}\hat{\theta}} = 0.80\,\frac{\mathrm{RMSE}}{\hat{\theta}}$估计

# 例11.2

对例11.1用平均值法做的估计进行误差分析。

解：

取N=10000个样本点的一次模拟的结果得到

$$\hat{I} = 0.8754, S_N = 0.9409$$

根均方误差的估计为 $\mathrm{RMSE} = \frac{S_N}{\sqrt{N}} = 0.0094$

平均绝对误差的估计为 $\mathrm{MAE} = 0.80\mathrm{RMSE} = 0.0075$

平均相对误差的估计为 $\mathrm{MRE} = \mathrm{MAE} / \hat{I} = 0.0087$

# 3.2.4 高维定积分

$d$元函数 $h(x_1, x_2, \ldots, x_d)$

定义于超矩形

$$C = \{(x_1, x_2, \ldots, x_d): a_i \le x_i \le b_i, i = 1,2, \ldots, d\}$$

且 $0 \le h(x_1, \ldots, x_d) \le M, \forall x \in C$

令

$$D = \{(x_1, x_2, \ldots, x_d, y): (x_1, x_2, \ldots, x_d) \in C, 0 \le y \le h(x_1, x_2, \ldots, x_d)\}$$

$$G = \{(x_1, x_2, \ldots, x_d, y): (x_1, x_2, \ldots, x_d) \in C, 0 \le y \le M\}$$

# 3.2.4 高维定积分

计算*d*维定积分

$$I=\int_{a_d}^{b_d}\cdots\int_{a_2}^{b_2}\int_{a_1}^{b_1}h(x_1,x_2,\ldots,x_d)dx_1dx_2\cdots dx_d \quad (11.19)$$

- **1 随机投点法**，产生 $\mathbf{Z}_1,\mathbf{Z}_2,\ldots,\mathbf{Z}_N$

是服从*d*+1维空间中的超矩形G内的均匀分布的独立抽样

令

$$\xi_i=\begin{cases}1, \mathbf{Z}_i\in D\\ 0, \mathbf{Z}_i\in G-D\end{cases}, \quad i=1,2,\ldots,N$$

则 $\xi_i$ iid *b*(1,p)

# 3.2.4 高维定积分

于是N个随机点中落入D的百分比

$$\hat{p} = \frac{\sum \xi_i}{N} \to p, \text{a.s.}(N \to \infty)$$

另一方面

$$p = P(\mathbf{Z}_i \in D) = \frac{V(D)}{V(G)} = \frac{I}{MV(C)}$$

$$= \frac{I}{M\prod_{j=1}^{d}(b_j - a_j)}$$

其中V表示区域体积。

# 3.2.4 高维定积分

- 积分I的估计

$$\hat{I}_1 = \hat{p}V(G) = \hat{p} \cdot MV(C) \tag{11.20}$$

$$= \hat{p} \cdot M \prod_{j=1}^{d} (b_j - a_j)$$

是I的无偏估计和强相合估计

# 3.2.4 高维定积分

- 由中心极限定理知

$$\sqrt{N}(\hat{p}-p)/\sqrt{p(1-p)} \xrightarrow{d} \mathrm{N}(0,1).$$

$$\sqrt{N}(\hat{I}_1 - I) \xrightarrow{d} \mathrm{N}\left(0, \left(M\prod_{j=1}^{d}(b_j - a_j)\right)^2 p(1-p)\right)$$

随机投点法渐近方差

$$\frac{\left(M\prod_{j=1}^{d}(b_j - a_j)\right)^2 p(1-p)}{N} \tag{11.21}$$

# 3.2.4 高维定积分

- 多重积分的随机投点法步骤

①赋初值:试验次数n=0,成功次数m=0;规定随机投点试验的总次数N;

②向k+1维立方体{0≤xi≤1(i=1,...,k),0≤y≤1}内随机投点,即产生k+1个相互独立的均匀随机数ξ=(ξ1,...,ξk,η), 置n=n+1;

③判断n≤N是否成立,若成立则转④;否则停止模拟试验,然后转⑤;

④检验k+1维空间的点ξ是否落入V中,即检验条件η≤f(ξ1,...,ξk)是否成立,若成立即试验成功,置m=m+1,然后转②;否则转②;

⑤计算θ1=m/N, 其中m是N次试验中成功的总次数,则I=θ1

# 3.2.4 高维定积分

计算d维定积分

$$I = \int_{a_d}^{b_d} \cdots \int_{a_2}^{b_2} \int_{a_1}^{b_1} h(x_1, x_2, \ldots, x_d) dx_1 dx_2 \cdots dx_d \qquad (11.19)$$

- **2平均值法**

设 $\mathbf{U}_i \sim iid\ U(C), (i = 1, 2, \ldots, N)$

则

$$h(\mathbf{U}_i), i = 1,2, \ldots, N$$

是iid随机变量列

# 3.2.4 高维定积分

期望

$$Eh(\mathbf{U}_i) = \int_C h(\mathbf{u}) \frac{1}{\prod_{j=1}^{d}(b_j - a_j)} d\mathbf{u} = \frac{I}{\prod_{j=1}^{d}(b_j - a_j)}$$

I估计为

$$\hat{I}_2 = \prod_{j=1}^{d}(b_j - a_j) \cdot \frac{1}{N}\sum_{i=1}^{N} h(U_i) \qquad (11.22)$$

# 3.2.4 高维定积分

由强大数律

$$\hat{I}_2 \to \prod_{j=1}^{d}(b_j - a_j)Eh(U) = I, \text{ a.s.}(N \to \infty)$$

渐近方差为

$$\frac{(V(C))^2 \mathrm{Var}(h(\mathbf{U}))}{N} = \frac{\left(\prod_{j=1}^{d}(b_j - a_j)\right)^2 \mathrm{Var}(h(\mathbf{U}))}{N} \quad (11.23)$$

$N \to \infty$ 时的误差阶也不受维数d的影响

# 3.2.4 高维定积分

- 多重积分的平均值法步骤

首先赋初值：ξ落入D的次数m=0，试验次数n=0，并规定试验总次数
N.

①产生k个相互独立服从[a,b]区间上的均匀随机数ξ=(ξ1,...,ξk)，置
n=n+1;

②判n≤N是否成立;若成立转③;否则停止抽样,转④;

③检验k维空间的点ξ是否落入积分区域D,若ξ∈D,置m=m+1,并令ηm=ξ,计算f(ηm),转①;否则舍去,转①,重新产生k维均匀随机数:

④计算$V_D \approx \frac{m}{N}(b-a)^k$，$E[f(\xi)|\xi \in D] \approx \frac{1}{m}\sum_{i=1}^{m} f(\eta_i)$

则$I_k \approx \frac{1}{N}(b-a)^k \sum_{i=1}^{m} f(\eta_i)$

# 3.2.4 高维定积分

- 比较随机投点法(11.20)与平均值法(11.22)

设随机投点法的估计为$\hat{I}_1$

平均值法的估计为$\hat{I}_2$

$$0 \le h(\mathbf{x}) \le M$$

不妨设上界M=1：

$$0 \le h(\mathbf{x}) \le 1$$

# 3.2.4 高维定积分

令$\boldsymbol{X}_i \sim$ iid U($C$), $\eta_i = h(\boldsymbol{X}_i)$, $Y_i \sim$ iid U(0,1)与$\{\boldsymbol{X}_i\}$独立,

$$\xi_i = \begin{cases} 1, & \text{as} Y_i \leq h(\boldsymbol{X}_i) \\ 0, & \text{as} Y_i > h(\boldsymbol{X}_i) \end{cases} \quad i = 1, 2, \ldots, N,$$

这时有

$$\hat{I}_1 = V(C)\frac{1}{N}\sum_{i=1}^{N}\xi_i, \qquad \hat{I}_2 = V(C)\frac{1}{N}\sum_{i=1}^{N}\eta_i$$

$$\text{Var}(\hat{I}_1) = \frac{1}{N}V^2(C)\cdot\frac{I}{V(C)}\left(1-\frac{I}{V(C)}\right)$$

$$\text{Var}(\hat{I}_2) = \frac{1}{N}V^2(C)\cdot\left(\frac{1}{V(C)}\int_C h^2(\boldsymbol{x})d\boldsymbol{x} - \left(\frac{I}{V(C)}\right)^2\right)$$

$$\text{Var}(\hat{I}_1) - \text{Var}(\hat{I}_2) = \frac{V(C)}{N}\int_C \left\{h(\boldsymbol{x}) - h^2(\boldsymbol{x})\right\} d\boldsymbol{x} \geq 0$$

- 可见平均值法精度更高

# 例11.3

考虑

$$f(x_1, x_2, \ldots, x_d) = x_1^2 x_2^2 \cdots x_d^2, 0 \le x_1 \le 1, 0 \le x_2 \le 1, \ldots, 0 \le x_d \le 1$$

的积分

$$I = \int_0^1 \cdots \int_0^1 \int_0^1 x_1^2\, x_2^2 \cdots x_d^2 dx_1 dx_2 \cdots dx_d$$

以d=8为例

比较**网格点法**、**随机投点法**、**平均值法**的精度

# 例11.3

- 网格点法

$$I_0 = \frac{1}{N} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_d=1}^{n} f\left(\frac{2i_1 - 1}{2n}, \frac{2i_2 - 1}{2n}, \dots, \frac{2i_d - 1}{2n}\right)$$

- 随机投点法

先在$(0,1)^d \times (0,1)$均匀抽样$(\xi_i^{(1)}, \xi_i^{(2)}, \dots, \xi_i^{(d)}, y_i), i = 1, \dots, N$,

令$\hat{I}_1$为$y_i \leq f(\xi_i^{(1)}, \xi_i^{(2)}, \dots, \xi_i^{(d)})$成立的百分比。

- 平均值法

$$\hat{I}_2 = \frac{1}{N} \sum_{i=1}^{N} f\left(\xi_i^{(1)}, \xi_i^{(2)}, \dots, \xi_i^{(d)}\right)$$

# 3.3 重要抽样法

3.3.1 标准化重要抽样法
3.3.2 带有舍选控制的重要抽样法

# 3.3 重要抽样法

在定积分 $$I = \int_C h(\mathbf{x}) d\mathbf{x}$$

- 积分区域C可能是任意形状的，也可能无界
- $h(\boldsymbol{x})$在C内各处的取值大小差异可能很大
- 用非均匀抽样

设$g(\boldsymbol{x}), \boldsymbol{x} \in C$是一个密度，其形状与$|h(\boldsymbol{x})|$相近，

且当$g(\boldsymbol{x}) = 0$时$h(\boldsymbol{x}) = 0$，

当$\|\boldsymbol{x}\| \to \infty$时$h(\boldsymbol{x}) = o(g(\boldsymbol{x}))$。

称$g(\boldsymbol{x})$为**试投密度或重要抽样密度**。

# 3.3 重要抽样法

设 $\mathbf{X}_i iid \sim g(\mathbf{x}), i = 1,2,\ldots,N$

令

$$\eta_i = \frac{h(\mathbf{X}_i)}{g(\mathbf{X}_i)}, i = 1,2,\ldots,N$$

则

$$E\eta = \int_C \frac{h(\mathbf{x})}{g(\mathbf{x})} g(\mathbf{x}) d\mathbf{x} = \int_C h(\mathbf{x}) d\mathbf{x} = I$$

# 3.3 重要抽样法

$\{\eta_i, i=1,2,\ldots,N\}$ 的**样本平均值**来估计$I$，即

$$\hat{I}_3 = \bar{\eta} = \frac{1}{N}\sum_{i=1}^{N}\frac{h(\mathbf{X}_i)}{g(\mathbf{X}_i)} \tag{12.1}$$

是$I$的无偏估计和强相合估计

渐近方差

$$\mathrm{Var}(\hat{I}_3) = \mathrm{Var}\left(\frac{h(\mathbf{X})}{g(\mathbf{X})}\right)\frac{1}{N} = O(\frac{1}{N}) \tag{12.2}$$

# 3.3 重要抽样法

- **求随机变量函数的期望问题**

设 $\mathbf{Y} \sim f(\mathbf{y})$

求期望值 $Eh(\mathbf{Y}) = \int h(\mathbf{y}) f(\mathbf{y}) d\mathbf{y}$

- 用重要抽样法估计E$h(\boldsymbol{Y})$
- 找试投密度$g(x)$
- 从$g(x)$产生随机数 $\mathbf{X}_i, i = 1,2,\ldots,N$

- 则有

$$\hat{I}_{3.1} = \frac{1}{N}\sum_{i=1}^{N} h(\mathbf{X}_i)\frac{f(\mathbf{X}_i)}{g(\mathbf{X}_i)} \qquad (12.4)$$

# 3.3 重要抽样法

- **求随机变量函数的期望问题**

易见

$$E\hat{I}_{3.1} = Eh(\mathbf{Y})$$

设

$$W_i = \frac{f(\mathbf{X}_i)}{g(\mathbf{X}_i)}$$

为**重要性权重**.

则(12.4)改成

$$\hat{I}_{3.1} = \frac{1}{N}\sum_{i=1}^{N} W_i\, h(\mathbf{X}_i) \qquad (12.5)$$

# 3.3.1 标准化重要抽样法

$\mathbf{Y} \sim f(\cdot)$ 很难直接抽样，或者

$f(\cdot)$ 本身未知

已知 $\tilde{f}(\mathbf{x}) = cf(\mathbf{x})$

其中：常数C未知，且计算困难

定义重要性权重为

$$W_i = \frac{\tilde{f}(\mathbf{X}_i)}{g(\mathbf{X}_i)}$$

公式(12.5)可以改成

$$\hat{I}_4 = \frac{\sum_{i=1}^{N} W_i h(\mathbf{X}_i)}{\sum_{i=1}^{N} W_i} \quad (12.7)$$

是$Eh(\boldsymbol{Y})$的强相合估计，但不是无偏的。

# 3.3.1 标准化重要抽样法

**定义12.1 (适当加权抽样)**

随机变量序列$\{(\mathbf{X}_i, W_i), i = 1,2,\ldots,N\}$称为关于密度$f(\cdot)$的适当加权抽样,

如果对于任何平方可积函数$h(\cdot)$都有

$$E[h(X_i)W_i] = cE_f[h(X)] = c\int h(\mathbf{x})f(\mathbf{x})\,d\mathbf{x},\ i = 1,2,\ldots,N,$$

其中$c$是归一化常数。

# 3.3.1 标准化重要抽样法

设随机变量$(\mathbf{X}, W)$联合密度为$g(\mathbf{x}, w)$，

则$(\mathbf{X}, W)$的样本为密度$f(\cdot)$的适当加权抽样的充分必要条件是

$$E_g(W|\mathbf{x}) = E_g(W)\frac{f(\mathbf{x})}{g(\mathbf{x})},\ \forall \mathbf{x},$$

其中：$E_g(W)$是关于$W$的边缘密度的期望，

$E_g(W|\mathbf{x})$是在$(\mathbf{X}, W)$的联合密度下条件期望$E(W|\mathbf{X})$在$\mathbf{X} = \mathbf{x}$处的值。

# 3.3.2 带有舍选控制的重要抽样法

- 舍弃权重太小的样本点， 重新抽样替换，称为**带有舍选控制的重要抽样法**
- 需要预先选定权重的一个阈值C，并计算

$$
\begin{aligned}
p_c &= \int \min\left\{1, \frac{f(\mathbf{x})}{c g(\mathbf{x})}\right\} g(\mathbf{x}) d\mathbf{x} \\
&= \int \min\left\{g(\mathbf{x}), \frac{f(\mathbf{x})}{c}\right\} d\mathbf{x}
\end{aligned}
$$

# 3.3.2 带有舍选控制的重要抽样法

- 在产生每个抽样点$X_i$时，计算权重$W_i$，
- 当权重$W_i \geq C$时接受此抽样点，但权重改为

$$W_i^* = p_c W_i$$

- 当$W_i < C$时仅以$W_i/C$的概率接受此抽样点，并调整权重为$P_c C$，
- 如果没有被接受则重新从g(.)中抽取

# 3.3.2 带有舍选控制的重要抽样法

**带有舍选控制的重要抽样法**

```
for(i in 1 : N) {
    repeat {
        从g(·)抽取ξ_i
        计算W_i ← f(ξ_i)/g(ξ_i)
        if(W_i ≥ c) {
            令X_i ← ξ_i, W_i^* ← p_c W_i
            break
        } else {
            从U(0,1)抽取U
            if(U < W_i/c) {
                令X_i ← ξ_i, W_i^* ← p_c c
                break
            }
        }
    } # repeat
} # for
```

输出$(X_i^*, W_i^*)$, $i = 1, 2, \ldots, N$

# 例12.1

- 用MC积分法计算

$$I = \int_0^1 e^x \, dx = e - 1 \approx 1.718$$

# 例12.1

**解：** 对被积函数$h(x)=e^x$做泰勒展开得

$$e^x=1+x+\frac{x^2}{2!}+\cdots$$

取 $g(x)=c(1+x)=\frac{2}{3}(1+x)$

要产生$g(x)$的随机数可以用逆变换法，

密度$g(x)$的分布函数$G(x)$的反函数为

$$G^{-1}(y)=\sqrt{1+3y}-1,\ 0<y<1$$

# 例12.1

因此，取$U_i iid U(0,1)$,

令$X_i = \sqrt{1+3U_i} - 1, i = 1,2, \ldots, N$,

则重要抽样法的积分公式为

$$\hat{I}_3 = \frac{1}{N} \sum_{i=1}^{N} \frac{e^{X_i}}{\frac{2}{3}(1+X_i)}$$

渐近方差为

$$\mathrm{Var}(\hat{I}_3) = \frac{1}{N}\left(\frac{3}{2}\int_0^1 \frac{e^{2x}}{1+x}dx - I^2\right) \approx 0.02691/N.$$

# 例12.1

如果用**平均值法**，估计公式为

$$\hat{I}_2 = \frac{1}{N}\sum_{i=1}^{N} e^{U_i},$$

渐近方差为

$$\mathrm{Var}(\hat{I}_2) = \frac{1}{N}\left(\int_0^1 e^{2x}\, dx - I^2\right) \approx 0.2420/N \approx 9.0 \times \mathrm{Var}(\hat{I}_3)$$

是重要抽样法方差的9倍。

# 例12.1

如果用随机投点法，

$h(x) = e^x \leq e(0 < x < 1)$,

取上界$M = e$,

向$[0,1] \times [0,M]$随机投点，

落到$f(x)$下方的概率为

$$p = I/\big(M(b-a)\big) = (e-1)/e,$$

# 例12.1

设投$N$点落到$h(x)$下方的频率为$\hat{p}$,

用随机投点法估计I的公式为

$$\hat{I}_1 = \hat{p} \cdot M(b-a) = e\hat{p},$$

渐近方差为

$$\mathrm{Var}(\hat{I}_1) = e^2 p(1-p)/N = (e-1)/N \approx 1.718/N$$
$$\approx 7.1 \times \mathrm{Var}(\hat{I}_2) \approx 64.8 \times \mathrm{Var}(\hat{I}_3)$$

- 重要抽样法只需要随机投点法的1/65的样本量

# 例12.2

- 用MC积分法计算二重定积分

$$I=\int_{-1}^{1}\int_{-1}^{1}f(x,y)dxdy$$

其中

$$f(x,y)=\exp\{-45(x+0.4)^2-60(y-0.5)^2\}$$
$$+0.5\exp\{-90(x-0.5)^2-45(y+0.1)^4\}$$

# 例12.2

**解：** 取试投密度为

$$g(x,y) \propto \tilde{g}(x,y) = \exp\{-45(x+0.4)^2 - 60(y-0.5)^2\} + 0.5\exp\{-90(x-0.5)^2 - 10(y+0.1)^2\},$$

$$-\infty < x < \infty, -\infty < y < \infty$$

# 例12.2

记$N(\mu,\sigma^2)$的分布密度为 $f(x;\mu,\sigma^2)$，

对$\tilde{g}(x,y)$积分得 $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\tilde{g}(x,y) =$

$$\sqrt{2\pi/90}\int_{-\infty}^{\infty}f(x;-0.4,90^{-1})dx\cdot\sqrt{2\pi/120}\int_{-\infty}^{\infty}f(y;0.5,120^{-1})dy$$

$$+0.5\sqrt{2\pi/180}\int_{-\infty}^{\infty}f(x;0.5,180^{-1})dx\cdot\sqrt{2\pi/20}\int_{-\infty}^{\infty}f(y;-0.1,20^{-1})dy$$

$$=\sqrt{2\pi/90}\sqrt{2\pi/120}+0.5\sqrt{2\pi/180}\sqrt{2\pi/20}$$

$$\approx 0.1128199$$

# 例12.2

于是令

$$g(x,y)=\tilde{g}(x,y)/0.1128199=$$

$$0.5358984f(x;-0.4,90^{-1})f(y;0.5,120^{-1})$$

$$+0.4641016f(x;0.5,180^{-1})f(y;-0.1,20^{-1})$$

$$-\infty<x<\infty,-\infty<y<\infty$$

- 用复合抽样法对g(x,y)抽样
- 然后用重要抽样法得到估计值

# 3.4 分层抽样法

# 3.4 分层抽样法

$$\int_C h(\mathbf{x})d\mathbf{x}$$

- 把C上的**积分分解为**若干个子集上的积分，
- 使得h(x)在每个子集上变化不大，
- 分别计算各个子集上的积分再求和，
- 可以提高估计精度。
- 这种方法叫做**分层抽样法**。

# 例13.1

- 求定积分 $$I = \int_0^1 h(x)dx$$

其中

$$h(x) = \begin{cases} 1 + \dfrac{x}{10}, & 0 \le x \le 0.5 \\ -1 + \dfrac{x}{10}, & 0.5 < x \le 1 \end{cases}$$

# 例13.1

- 在(0,1)区间随机抽取N点用**平均值法**得 $\hat{I}_2$
- 渐近方差

$$\mathrm{Var}(\hat{I}_2)=\frac{\mathrm{Var}(h(U))}{N}=\frac{143}{150N}\approx\frac{0.9533}{N}$$

- 分层抽样法

把I拆分为[0,0.5]和[0.5,1]上的积分，即

$$I=a+b=\int_0^{0.5}h(x)dx+\int_{0.5}^{1}h(x)dx$$

# 例13.1

## 对a和b分别用平均值法，得

$$\hat{a} = \frac{0.5}{N/2}\sum_{i=1}^{N/2} h\left(0.5U_i\right) = \frac{0.5}{N/2}\sum_{i=1}^{N/2}\left(1 + 0.05U_i\right),$$

$$\hat{b} = \frac{0.5}{N/2}\sum_{i=(N/2)+1}^{N} h\left(0.5 + 0.5U_i\right) = \frac{0.5}{N/2}\sum_{i=(N/2)+1}^{N}\left(-1 + 0.05 + 0.05U_i\right),$$

$$\hat{I}_5 = \hat{a} + \hat{b}$$

## 渐近方差

$$\mathrm{Var}(\hat{I}_5) = \mathrm{Var}(\hat{a}+\hat{b}) = \mathrm{Var}(\hat{a}) + \mathrm{Var}(\hat{b})$$

$$= 0.25\frac{\mathrm{Var}(1+0.05U)}{N/2} + 0.25\frac{\mathrm{Var}(-0.95+0.05U)}{N/2}$$

$$= \frac{1/4800}{N}$$

- 分层后的估计方差远小于不分层的结果，可以节省样本量约4500倍

# 3.4 分层抽样法

- 设积分 $$I=\int_C h(\mathbf{x})d\mathbf{x}$$

可以分解为m个不交的子集$C_j$上的积分，即

$$I=\int_C h(\mathbf{x})d\mathbf{x}=\int_{C_1} h(\mathbf{x})d\mathbf{x}+\int_{C_2} h(\mathbf{x})d\mathbf{x}+\cdots+\int_{C_m} h(\mathbf{x})d\mathbf{x}$$

在$C_j$投$n_j$个随机点 $X_{ji}\sim \mathrm{U}(C_j), i=1,\ldots,n_j$

则$I$的$m$个部分可以分别用平均值法估计

# 3.4 分层抽样法

由此得I的分层估计为

$$\hat{I}_5 = \sum_{j=1}^{m} \frac{V(C_j)}{n_j} \sum_{i=1}^{n_j} h(X_{ji})$$

记 $\sigma_j^2 = \mathrm{Var}(h(X_{j1}))$

划分子集时应使每一子集内h(.)变化不大，

即 $\sigma_j^2$ 较小，有

$$\mathrm{Var}(\hat{I}_5) = \sum_{j=1}^{m} \frac{V^2(C_j)\sigma_j^2}{n_j}$$

# 3.4 分层抽样法

若 $\sigma_j^2$ 可估计，应取$n_j$使

$$n_j \propto V(C_j)\sigma_j \qquad (13.1)$$

即

$$n_j = N\frac{V(C_j)\sigma_j}{\sum_{k=1}^{m} V(C_k)\sigma_k}, j=1,2,\ldots,m$$

这样取的样本量$(n_1, n_2, \ldots, n_m)$

在所有满足 $n_1 + n_2 + \cdots + n_m = N$

的取法中使得渐近方差最小。

# 3.4 分层抽样法

- 求随机变量函数期望

设X为随机变量，求X的函数h(X)的数学期望

$$\theta = Eh(X)$$

假设存在离散型随机变量Y

$$p_j = P\left(Y = y_j\right), j = 1,2,\ldots,m$$

在$Y=y_j$条件下可以从X的条件分布抽样，则

$$E[h(X)] = E\{E[h(X)|Y]\} = \sum_{j=1}^{m} E\left[h(X)\middle|Y = y_j\right]p_j$$

# 3.4 分层抽样法

- 求随机变量函数期望 $\theta = Eh(X)$

如果在Y=$y_j$条件下生成X的$N_j$=$Np_j$个抽样值，设为 $X_i^{(j)}, (i = 1,2, \dots, N_j)$

则可以用

$$\frac{1}{N_j}\sum_{i=1}^{N_j} h\left(X_i^{(j)}\right)$$

估计

$$E\left[h\left(X\right) \mid Y = y_j\right]$$

# 3.4 分层抽样法

- 求随机变量函数期望 $\theta = Eh(X)$

估计

$$\hat{\theta} = \sum_{j=1}^{m} \frac{1}{N_j} \sum_{i=1}^{N_j} h\left(X_i^{(j)}\right) p_j = \frac{1}{N} \sum_{j=1}^{m} \sum_{i=1}^{Np_j} h\left(X_i^{(j)}\right) \qquad (13.2)$$

这是$\theta$的**无偏和强相合估计**， 且估计方差

$$\mathrm{Var}(\hat{\theta}) = \frac{1}{N^2} \sum_{j=1}^{m} Np_j \mathrm{Var}[h(X) \mid Y = y_j] = \frac{1}{N} \sum_{j=1}^{m} \mathrm{Var}[h(X) \mid Y = y_j] p_j$$

$$= \frac{1}{N} E\{\mathrm{Var}[h(X) \mid Y]\} \le \frac{1}{N} \mathrm{Var}[h(X)] \qquad (13.3)$$

- 比直接用平均值法估计E h(X)的方差小

# 例13.2

设 $U \sim \mathrm{U}(0,1)$

用分层抽样法估计 $\theta = Eh(U) = \int_0^1 h(x)dx$

解：令$Y =$ ð³⁄₄ĩřŤ ě$(mU)$, 即当且仅当$\frac{j-1}{m} < U \le \frac{j}{m}$时$Y = j$， $j = 1,2,\ldots,m$,

可以按照$Y$分层抽样估计$\theta$:

$$\theta = E[h(U)] = \sum_{j=1}^{m} E\,[h(U)|Y = j]\,P(Y = j) = \frac{1}{m}\sum_{j=1}^{m} E\,[h(U)|Y = j]$$

易见$Y = j$条件下$U$服从$\left(\frac{j-1}{m}, \frac{j}{m}\right)$上的均匀分布，

设$U_1, U_2, \ldots, U_{\mathrm{m}}$是$U(0,1)$的独立抽样，则用分层抽样法取每层$N_j = 1$估计$\theta = Eh(U)$为

$$\hat{\theta} = \frac{1}{m}\sum_{j=1}^{m} h\left(\frac{j-1+U_j}{m}\right).$$

# 3.5 方差缩减技术

3.5.1 控制变量法
3.5.2 对立变量法
3.5.3 条件期望法
3.5.4 随机数复用

# 3.5 方差缩减技术

- 随机模拟方法**优点**：适用性广、方法简单
- 随机模拟方法**缺点**：精度低、计算量大
- 减小随机模拟误差方差，　可以有效节省随机模拟时间

估计随机变量X的**期望**$\theta=EX$

目标是降低$\theta$的估计量的**渐近方差**

# 3.5.1 控制变量法

从X中抽取N个独立样本值

$$X_1, X_2, \ldots, X_N$$

用样本平均值

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$$

估计$\theta$=EX

# 3.5.1 控制变量法

设随机变量Y满足 $EY=0,\quad \mathrm{Cov}(X,Y)<0$

令$Z = X+Y$, 则

$$E(Z)=\theta$$

$$\mathrm{Var}(Z)=\mathrm{Var}(X)+\mathrm{Var}(Y)+2\mathrm{Cov}(X,Y)$$

只要

$$\mathrm{Var}(Y)+2\mathrm{Cov}(X,Y)<0$$

则

$$\mathrm{Var}(Z)<\mathrm{Var}(X)$$

# 3.5.1 控制变量法

为利用Y与X相关性，令 $Z(b) = X + bY$

则
$$EZ(b) = EX = \theta$$

$$\mathrm{Var}(Z(b)) = \mathrm{Var}(X) + 2b\mathrm{Cov}(X,Y) + b^2\mathrm{Var}(Y)$$

求$\mathrm{Var}(Z(b))$关于$b$的最小值点， 得

$$b = -\mathrm{Cov}(X,Y)/\mathrm{Var}(Y) = -\rho_{X,Y}\sqrt{\mathrm{Var}(X)/\mathrm{Var}(Y)}$$

这时

$$\mathrm{Var}(Z(b)) = (1-\rho_{X,Y}^2)\mathrm{Var}(X) \le \mathrm{Var}(X)$$

# 例14.1

估计

$$I=\int_0^1 e^t\,dt$$

设 $U\sim \mathrm{U}(0,1)$ $X=e^U$ 则$I=Ee^U=EX$

**法1：**用平均值法估计I为

$$\hat{I}_1=\frac{1}{N}\sum_{i=1}^{N}e^{U_i}$$

其方差为

$$\mathrm{Var}(\hat{I}_1)=\frac{1}{N}\mathrm{Var}(e^U)=\frac{1}{N}(-\frac{1}{2}e^2+2e-\frac{3}{2})\approx\frac{0.2420}{N}$$

# 例14.1

**法2**：令

$$Y = U - \frac{1}{2}$$

则EY=0，X与Y正相关

$$\mathrm{Cov}(X,Y) \approx 0.14086$$

$$\mathrm{Var}(Y) = 1/12$$

$$b = -\mathrm{Cov}(X,Y)/\mathrm{Var}(Y) = -1.690$$

$$Z(b) = e^U - 1.690\left(U - \frac{1}{2}\right)$$

# 例14.1

**法2：** 用控制变量法估计I为

$$\hat{I}_2 = \frac{1}{N}\sum_{i=1}^{N}\left[e^{U_i} - 1.690\left(U_i - \frac{1}{2}\right)\right]$$

- Z的方差为

$$\mathrm{Var}(Z(b)) = [1-\rho_{X,Y}^2]\mathrm{Var}(X) = (1-0.9919^2)\mathrm{Var}(X) = 0.016\mathrm{Var}(X) = 0.0039$$

- 用控制变量法估计I的方差为

$$\mathrm{Var}(\hat{I}_2) = \frac{1}{N}\mathrm{Var}(Z(b)) = \frac{0.0039}{N}$$

# 例14.2 (系统可靠性估计)

考虑由n个部件组成的一个系统，

用$S_i$表示第$i$个部件是否正常工作， 1表示正常工作，0表示失效。

设$S_i \sim \mathrm{B}(1, p_i)$且各$S_i$相互独立。

用$Y$表示系统是否工作正常，1表示工作正常， 0表示系统失效。

设$Y$为$S_1, S_2, \dots, S_n$的函数$\phi(S_1, S_2, \dots, S_n)$且$\phi$关于每个$S_i$单调不减，

称$\phi$为系统的结构函数。

令$R = P(Y = 1) = E\phi(S_1, S_2, \dots, S_n)$， 称R为**系统可靠度**。

# 例14.2 (系统可靠性估计)

- 用平均值法估计R

记$\mathbf{S}=(S_1,S_2,\ldots,S_n)$, $X=\phi(\mathbf{S})$, 对$\mathbf{S}$独立抽取N个点

$$\mathbf{S}^{(j)}=\left(S_1^{(j)}S_2^{(j)},\ldots,S_n^{(j)}\right), j=1,2,\ldots,N$$

R可以用平均值法估计为

$$\hat{R}_1=\frac{1}{N}\sum_{j=1}^{N}\phi\left(\mathbf{S}^{(j)}\right)$$

# 例14.2 (系统可靠性估计)

- 用控制变量法估计R

令$Y = \sum_{i=1}^{n}(S_i - p_i)$, 则$EY = 0$, Y与X正相关。

用一个小的抽样先近似估计$\mathrm{Cov}(X, Y), \mathrm{Var}(Y)$得到$b$的近似值，

则可得到方差缩减的估计量

$$\hat{R}_2 = \frac{1}{N}\sum_{j=1}^{N}\left[\phi\left(\mathbf{S}^{(j)}\right) + b\sum_{i=1}^{n}\left(S_i^{(j)} - p_i\right)\right]$$

# 3.5.2 对立变量法

**定理14.1**

设g为单调函数， $U \sim U(0,1)$

则 $\mathrm{Cov}(g(U), g(1-U)) \le 0$

# 定理14.1证明

$\forall u_1, u_2 \in [0,1]$，由g单调可知

$$(g(u_1)-g(u_2))\,(g(1-u_1)-g(1-u_2)) \leq 0$$

设$U_2$服从 U(0,1)且与U独立，

令$X_1 = g(U), Y_1 = g(1-U), X_2 = g(U_2), Y_2 = g(1-U_2)$,

则$X_1, Y_1, X_2, Y_2$的分布相同，且

$$\begin{aligned}
&E(X_1 - X_2)(Y_1 - Y_2)\\
=&\operatorname{Cov}(X_1 - X_2, Y_1 - Y_2)\\
=&\operatorname{Cov}(X_1, Y_1) + \operatorname{Cov}(X_2, Y_2) - \operatorname{Cov}(X_1, Y_2) - \operatorname{Cov}(X_2, Y_1)\\
=&2\operatorname{Cov}(X_1, Y_1)
\end{aligned}$$

注意$(X_1 - X_2)(Y_1 - Y_2) \leq 0$

所以$\operatorname{Cov}(X_1, Y_1) \leq 0$，即

$\operatorname{Cov}(g(U), g(1-U)) \leq 0$。

# 3.5.2 对立变量法

## 定理14.2

设$h(x_1, x_2, \ldots, x_n)$是关于每个自变量单调的函数，$U_1, U_2, \ldots, U_n$相互独立，

则$\mathrm{Cov}(h(U_1, U_2, \ldots, U_n),\ h(1-U_1, 1-U_2, \ldots, 1-U_n)) \le 0$

# 3.5.2 对立变量法

**定理14.3 (对立变量法)**

设F(x)为连续分布函数,

$$U \sim \mathrm{U}(0,1),\ X = F^{-1}(U),\ Y = F^{-1}(1-U),\ Z = \frac{X+Y}{2},$$

则 X与Y同分布F(x)且

$$\mathrm{Cov}(X,Y) \le 0$$

$$\mathrm{Var}(Z) \le \frac{1}{2}\mathrm{Var}(X)$$

# 定理14.3 (对立变量法)

因为U和1-U同分布

所以$X=F^{-1}(U)$和$Y=F^{-1}(1-U)$同分布。

由定理14.1, 令$g(\cdot)=F^{-1}(\cdot)$

可知$\mathrm{Cov}(X,Y)=\mathrm{Cov}(g(U),g(1-U))\leq 0$，从而

$$
\begin{aligned}
\mathrm{Var}(Z) &= \frac{\mathrm{Var}(X)+\mathrm{Var}(Y)+2\mathrm{Cov}(X,Y)}{4} \\
&= \frac{\mathrm{Var}(X)+\mathrm{Cov}(X,Y)}{2} \leq \frac{1}{2}\mathrm{Var}(X)
\end{aligned}
$$

# 3.5.2 对立变量法

- 对立变量法

为了估计$I = EX$，产生$U_1, U_2, \ldots, U_N$后用

$$Z_i = \frac{1}{2}\left(F^{-1}(U_i) + F^{-1}(1 - U_i)\right),\ i = 1,2,\ldots,N$$

的样本平均值估计$I$可以提高精度。

# 例14.3

再次考虑例14.1的问题。
用对立变量法改善估计方差。

$$I = \int_0^1 e^t \, dt$$

设$U \sim \mathrm{U}(0,1)$，$X = e^U$, 令$Y = e^{1-U}$,

用对立变量法估计$I$为

$$\hat{I}_3 = \frac{1}{N}\sum_{i=1}^{N}\frac{e^{U_i} + e^{1-U_i}}{2}$$

方差为

$$\begin{aligned}\mathrm{Var}(\hat{I}_3) &= \frac{1}{N}\frac{\mathrm{Var}(e^U) + \mathrm{Cov}(e^U, e^{1-U})}{2} \\ &= \frac{1}{N}\left(-\frac{3}{4}e^2 + \frac{5}{2}e - \frac{5}{4}\right) \approx \frac{0.003913}{N}\end{aligned}$$

# 例14.4

再次考虑例14.2的可靠度估计问题。

- 用对立变量法改善估计方差。

设$\{U_k\}$为标准均匀分布随机数列，取

$$S_i^{(j)} = \begin{cases} 1 & \text{当} U_{n(j-1)+i} \le p_i \\ 0 & \text{其它} \end{cases}$$

则$S_i^{(j)}$是$U_{n(j-1)+i}$的单调非增函数。

# 例14.4

- R用**平均值法**估计为

$$\hat{R}_1 = \frac{1}{N}\sum_{j=1}^{N}\phi\left(S_1^{(j)}, S_2^{(j)}, \dots, S_n^{(j)}\right)$$

# 例14.4

- R用**对立变量法**估计

令$h(U_1,U_2,\ldots,U_n)=\phi(S_1,S_2,\ldots,S_n)$,

则h关于每个自变量是单调非增函数，于是

$$\mathrm{Cov}(h(U_1,U_2,\ldots,U_n),\ h(1-U_1,1-U_2,\ldots,1-U_n))\leq 0$$

估计系统可靠度R为

$$\hat{R}_3=\frac{1}{N}\sum_{j=1}^{N}\frac{h(U_1^{(j)},U_2^{(j)},\ldots,U_n^{(j)})+h(1-U_1^{(j)},1-U_2^{(j)},\ldots,1-U_n^{(j)})}{2}$$

# 3.5.3 条件期望法

设变量X与Y不独立，
根据Rao-Blackwell不等式

$$\mathrm{Var}\left\{\mathrm{E}(Y \mid X)\right\} \leq \mathrm{Var}(Y)$$

又

$$\mathrm{E}\left\{\mathrm{E}(Y \mid X)\right\} = \mathrm{E}Y = I$$

对$Z$=E ($Y$|$X$)抽样，用$Z$的样本平均值来估计$I$=E$Y$比直接用Y的样本平均值的精度更高

- 条件期望法或Rao-Blackwell方法

# 例14.5

设$X \sim p(x)$，$\varepsilon \sim \mathrm{N}(0, \sigma^2)$且与X独立，

$$Y = \psi(X) + \varepsilon$$

估计$I = \mathrm{E}Y$

# 例14.5 法1：平均值法

可以用**条件分布抽样法**对二元随机向量$\boldsymbol{Z}=(X,Y)$抽样产生Y的样本。

从$p(\cdot)$抽样得$X_1,X_2,\ldots,X_N$, 独立地从$N(0,\sigma^2)$抽样得$\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_N$， 令

$$Y_i=\psi(X_i)+\varepsilon_i,\ i=1,2,\ldots,N$$

然后用$Y_1,Y_2,\ldots,Y_N$的样本平均值估计EY:

$$\hat{I}_1=\frac{1}{N}\sum_{i=1}^{N}Y_i,$$

估计方差为

$$\text{Var}(\hat{I}_1)=\frac{\text{Var}(Y_1)}{N}=\frac{\text{Var}(\psi(X_1))}{N}+\frac{\sigma^2}{N}.$$

# 例14.5 法2：条件期望法

另一方面，注意$E(Y|X)=\psi(X)$，也可以只对X抽样然后用条件期望法估计EY:

$$\hat{I}_2 = \frac{1}{N}\sum_i \psi(X_i),$$

估计方差为

$$\mathrm{Var}(\hat{I}_2) = \frac{\mathrm{Var}(\psi(X_1))}{N} < \mathrm{Var}(\hat{I}_1).$$

# 例14.6 法1：随机投点法

## 考虑例10.1中用随机模拟方法估计π的改进问题

设(X, Y)服从正方形$D=[-1,1]^2$上的均匀分布，

令$\eta=1$表示(X,Y)落入单位圆$C=\{(x,y):x^2+y^2\le 1\}$,

$\eta=0$表示未落入单位圆,

则$\eta\sim \mathrm{B}(1,\pi/4)$。

设$(X_i,Y_i), i=1,2,\ldots,N$是(X,Y)的独立重复抽样，

$\eta_i$表示$X_i^2+Y_i^2\le 1$是否发生，则$\pi$的估计为

$$\hat{\pi}_1=\frac{4}{N}\sum_{i=1}^{N}\eta_i,$$

方差为$\mathrm{Var}(\hat{\pi}_1)=\pi(4-\pi)/N\approx\frac{2.6968}{N}$。

# 例14.6 法2： 条件期望法

## 考虑例10.1中用随机模拟方法估计π的改进问题

用$\zeta = E(\eta|X)$的样本代替$\eta$来估计$E\eta = \pi/4$，则由

$$
\begin{aligned}
E(\eta|X=x) =& P(X^2+Y^2 \le 1 \,|\, X=x) = P(Y^2 \le 1-x^2) \\
=& P(-\sqrt{1-x^2} \le Y \le \sqrt{1-x^2}) \\
=& \sqrt{1-x^2} \quad (\text{notice that } Y \sim \mathrm{U}(-1,1))
\end{aligned}
$$

可知$\zeta = \sqrt{1-X^2}$。其方差为

$$
\mathrm{Var}(\zeta) = E(1-X^2) - (E\zeta)^2 = \frac{2}{3} - \frac{\pi^2}{16},
$$

# 例14.6

## 考虑例10.1中用随机模拟方法估计π的改进问题

估计$\pi$为 $\hat{\pi}_2 = \frac{4}{N}\sum_{i=1}^{N}\sqrt{1-X_i^2}$,

方差为 $\mathrm{Var}(\hat{\pi}_2) \approx 0.7971/N$,

$\mathrm{Var}(\hat{\pi}_1)/\mathrm{Var}(\hat{\pi}_2) \approx 3.4$。

另外，$\zeta$仅依赖于$X^2$,

容易发现若$U \sim \mathrm{U}(0,1)$则$X^2$和$U^2$同分布,

所以可取$\xi = \sqrt{1-U^2}$, 其中$U \sim \mathrm{U}(0,1)$。

# 例14.6 法3：对立变量法

## 考虑例10.1中用随机模拟方法估计π的改进问题

函数$h(u)=\sqrt{1-u^2}$是$u\in(0,1)$的单调函数，

可以利用**对立变量法**，构造$\pi$的估计量为

$$\hat{\pi}_3 = \frac{4}{N}\sum_{i=1}^{N}\frac{\sqrt{1-U_i^2}+\sqrt{1-(1-U_i)^2}}{2},$$

其中$U_1,U_2,\ldots,U_N$为U(0,1)随机数，

则$\mathrm{Var}(\hat{\pi}_3)\approx 0.11/N$,

$\mathrm{Var}(\hat{\pi}_1)/\mathrm{Var}(\hat{\pi}_3)\approx 25$。

# 例14.6 法4：　控制变量法

## 考虑例10.1中用随机模拟方法估计π的改进问题

令$W = U^2 - \frac{1}{3}$, 则EW = 0且W与$\xi$负相关。

可以先进行一个小规模的模拟估计$\mathrm{Cov}(\xi, W)$和$\mathrm{Var}(W)$得到

$b = -\mathrm{Cov}(\xi, W)/\mathrm{Var}(W)$的近似值，

用$\xi(b) = \xi + bW$代替$\xi$进行抽样，　可以减小$\hat{\pi}_2$的方差。

# 3.5.4 随机数复用

- 公共随机数法
- （Common Random Number)
- 亦称“相关采样”、“匹配流”
- 方差缩减技术之一
- 在模拟过程中针对不同的系统或配置**采用完全同样的随机数序列**
- 由于随机数序列完全相同，不同系统的运行结果之间存在正相关性，从而缩小了对两个系统差异的**估计方差**

# 3.5.4 随机数复用

例：假设有n件工作可由两台相同机器中的任一个进行处理。

- 有两种不同工作安排顺序的策略：
  1，长工期工作优先策略
  2，短工期工作优先策略
- 比较两种策略处理完所有工作所需的时间

**解：**

通过随机模拟比较在这两种不同策略下,完工时间期望的差异

# 3.5.4 随机数复用

**解：**

- $T_j$：处理第j个工作所用的时间($j=1\ldots,n$)
  则$T_1,\ldots,T_n$是分布为F的随机变量
- $g(t_1,\ldots,t_n)$是采用**长工期**工作优先策略下,处理需要时间分别为$t_1,\ldots,t_n$的n个工作所用的时间
- $h(t_1,\ldots,t_n)$是采用**短工期**工作优先策略下,处理需要时间分别为$t_1,\ldots,t_n$的n个工作所用的时间
- $\theta_1=E[g(T)]$, $\theta_2=E[h(T)]$, $T=(t_1,\ldots,t_n)$

则要用**随机模拟估计**：$\theta=\theta_1-\theta_2$

# 3.5.4 随机数复用

**解:**

- 通过随机产生向量T来计算*g*(T)
- 计算*h*(T)是用同样的值，还是重新产生？
- 设用$T^*=(T_1^*,\dots,T_n^*)$估计$\theta_2$
- 则$\theta$的估计*g*(T)-*h*(T*)的方差是:

$$Var(g(T)-h(T^*))$$
$$=Var(g(T))+Var(h(T^*))-2\mathrm{Cov}((g(T),h(T^*))$$
$$=Var(g(T))+Var(h(T))-2Cov((g(T),h(T^*))$$

- 重复使用进行计算效果较好

# 3.5.4 随机数复用

例14.8 对正态分布总体 $X \sim N(\mu, \sigma^2)$

设样本 $X_1, X_2, \ldots, X_n$

则方差有两种不同的公式估计:

$$\hat{\sigma}_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

$$\hat{\sigma}_2^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

- 比较两个估计量的偏差和均方误差

# 3.5.4 随机数复用

## 例14.8

- 偏差 $b_1 = E\hat{\sigma}_1^2 - \sigma^2, \quad b_2 = E\hat{\sigma}_2^2 - \sigma^2,$
- 均方误差 $s_1 = E(\hat{\sigma}_1^2 - \sigma^2)^2, \quad s_2 = E(\hat{\sigma}_2^2 - \sigma^2)^2$
- 随机模拟法作比较

# 3.5.4 随机数复用

## 例14.8

重复地生成N组样本$(X_1^{(j)}, X_2^{(j)}, \ldots, X_n^{(j)})$, $j = 1, 2, \ldots, N$。

对每组样本分别计算$\hat{\sigma}_{1j}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i^{(j)} - \bar{X}^{(j)})^2$ 和$\hat{\sigma}_{2j}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i^{(j)} - \bar{X}^{(j)})^2$

得到偏差和均方误差的估计

$$\hat{b}_1 = \frac{1}{N}\sum_{j=1}^{N}\hat{\sigma}_{1j}^2 - \sigma^2, \quad \hat{b}_2 = \frac{1}{N}\sum_{j=1}^{N}\hat{\sigma}_{2j}^2 - \sigma^2,$$

$$\hat{s}_1 = \frac{1}{N}\sum_{j=1}^{N}(\hat{\sigma}_{1j}^2 - \sigma^2)^2, \quad \hat{s}_2 = \frac{1}{N}\sum_{j=1}^{N}(\hat{\sigma}_{2j}^2 - \sigma^2)^2.$$

- 两种方法使用相同的模拟样本得到的偏差、均方误差的估计精度与每种方法单独生成模拟样本得到的估计精度相同

# 3.5.4 随机数复用

## 例14.8

- 下面比较：$\Delta s = s_1 - s_2$

$$\mathrm{Var}(\hat{s}_1 - \hat{s}_2) = \mathrm{Var}(\hat{s}_1) + \mathrm{Var}(\hat{s}_2) - 2\mathrm{Cov}(\hat{s}_1, \hat{s}_2) \quad (14.1)$$

- 用不同的样本

$$\mathrm{Var}(\hat{s}_1 - \hat{s}_2) = \frac{1}{N}\left[\mathrm{Var}((\hat{\sigma}_1^2 - \sigma^2)^2) + \mathrm{Var}((\hat{\sigma}_2^2 - \sigma^2)^2)\right] \quad (14.2)$$

- 用相同的样本

$$\mathrm{Var}(\hat{s}_1 - \hat{s}_2) = \frac{1}{N}\left[\mathrm{Var}((\hat{\sigma}_1^2 - \sigma^2)^2) + \mathrm{Var}((\hat{\sigma}_2^2 - \sigma^2)^2) - 2\mathrm{Cov}((\hat{\sigma}_1^2 - \sigma^2)^2, (\hat{\sigma}_2^2 - \sigma^2)^2)\right] \quad (14.3)$$

- 利用**相同样本**的估计精度更好