吉林大学

统计计算

Statistical Computation

机器学习研究室
统计计算课程组

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

3.6 随机服务系统模拟

3.6.1 概述
3.6.2 离散事件模拟

3.6.1 概述

1 随机过程(Stochastic Process)

- 是依赖于参数的一组随机变量的全体，参数通常是时间。
- 随机变量是随机现象的数量表现，其取值随着偶然因素的影响而改变
- 例如，某商店在从时间t_0到时间t_K这段时间内接待顾客的人数，就是依赖于时间t的一组随机变量，即随机过程。

3.6.1 概述

1 随机过程(Stochastic Process)

- 随机过程的理论产生于20世纪初期
- 是应物理学、生物学、管理科学等方面的需要而逐步发展起来的
- 在自动控制、公用事业、管理科学等方面都有广泛的应用。

3.6.1 概述

1 随机过程(Stochastic Process)

- 设{X(t),t ∈T}表示一组随机变量的集合， 称为一个**随机过程**,指标集合T称为**时间参数集**
- X_t称为“**系统在时刻*t*的状态**”
- 设所有X_t均取值于**有限**集合S
- S称为**状态空间**

3.6.1 概述

1 随机过程(Stochastic Process)

例 (随机游动)
设质点在时刻t=0从原点出发沿x轴按如下规则移动：每个一个时间单位以概率p右移一格，以概率q=1-p左移一格。若用X(t)表示时刻t质点所处的位置，则{X(t),t=1,2,…}构成一随机变量序列。

3.6.1 概述

1 随机过程(Stochastic Process)

例 (排队问题)
顾客到银行办理个人业务，若窗口有人接受服务，后来者需排队等待。
若X(t)表示t时刻银行等待服务的人数，则{X(t),t≥0}是一个随机过程。

3.6.1 概述

1 随机过程(Stochastic Process)

例 (排队问题)
顾客到银行办理个人业务，若窗口有人接受服务，后来者需排队等待。
若X(t)表示t时刻银行等待服务的人数，则{X(t),t≥0}是一个随机过程。

3.6.1 概述

1 随机过程(Stochastic Process)

例 (股指波动)

- 记录证券交易所的股指，用X(n)表示第n天上证综合指数的开盘指数，

则{X(n),n=1,2,…}是一随机序列。

- 实时记录证券交易所的股指，用X(t)表示一天中t时刻上证综合指数，则{X(t),t≥0}表示一随机过程。

3.6.1 概述

1 随机过程(Stochastic Process)

- **随机过程的分类**

(按参数集与状态是连续还是离散分类)

（1）离散时间离散状态

（2）离散时间连续状态

（3）连续时间离散状态

（4）连续时间连续状态

3.6.1 概述

1 随机过程(Stochastic Process)

- 一维随机过程

给定随机过程 $\{X(t), t \in T\}$。

对每个固定的 $t \in T$，随机变量 $X(t)$的分布函数一般与 t有关，记为：

$$F_X(x,t) = P\{X(t) \leq x\},\ x \in R$$

称它为随机过程 $\{X(t), t \in T\}$的**一维分布函数**，

而 $\{F_X(x,t), t \in T\}$称为**一维分布函数族**。

3.6.1 概述

1 随机过程(Stochastic Process)

- 一维随机过程

对任意 n $(n = 2,3,\cdots)$个不同的时刻 $t_1, t_2, \cdots, t_n \in T$ ，

引入 n维随机变量 $\big(X(t_1), X(t_2), \cdots, X(t_n)\big)$，它的分布函数记为：

$$F_X(x_1, x_2, \cdots, x_n; t_1, t_2, \cdots, t_n) = P\{X(t_1) \leq x_1, X(t_2) \leq x_2, \cdots, X(t_n) \leq x_n\}$$

对于固定的 n ，称 $\{F_X(x_1, x_2, \cdots, x_n; t_1, t_2, \cdots, t_n), t_i \in T\}$为随机过程 $\{X(t), t \in T\}$的 ***n* 维分布函数族**。

3.6.1 概述

1 随机过程(Stochastic Process)

- 一维随机过程

给定随机过程 $X(t), t \in T$，固定 $t \in T, X(t)$ 是一维随机变量，它的均值一般与 t 有关，记为：

$$\mu_X(t) = E[X(t)]$$

称 $\mu_X(t)$的随机过程$\{X(t), t \in T\}$的**均值函数**。

3.6.1 概述

1 随机过程(Stochastic Process)

- 一维随机过程

分别记随机变量$X(t)$的二阶原点矩和二阶中心矩为：

$$\Psi_X^2(t) = E[X^2(t)]$$

$$\sigma_X^2(t) = D\_X(t) = Var[X(t)] = E[X(t) - \mu\_X(t)]^2$$

分别称为随机过程$\{X(t), t \in T\}$的**均方值函数**和**方差函数**

3.6.1 概述

1 随机过程(Stochastic Process)

- 如果对于每一个 $t \in T$ ，

随机过程$X(t), t \in T$ 的二阶矩 $E\left[X^2(t)\right]$ 都存在，

- 则称它为二阶矩过程。

3.6.1 概述

1 随机过程(Stochastic Process)

- 给定二阶矩过程 $\{X(t), t \geq 0\}$ ，

称随机变量$X(t) - X(s),\ 0 \leq s < t$为随机过程在区间$(s, t]$上的**增量**。

- 如果对任意选定的正整数 n 和任意选定的 $0 \leq t_0 \leq t_1 \leq \cdots \leq t_n$ ， n 个增量：

  $X(t_1) - X(t_0), X(t_2) - X(t_1), \cdots, X(t_n) - X(t_{n-1})$

相互独立，

则称$X(t), t \geq 0$ 为**独立增量过程**。

3.6.1 概述

1 随机过程(Stochastic Process)

- 若对任意的实数 h 和 $0 \le s+h < t+h, X(t+h) - X(s+h)$与 $X(t) - X(s)$ 具有相同的分布，

则称**增量具有平稳性**。

- 增量$X(t) - X(s)$的分布函数实际上只依赖于时间差$t - s\ (0 \le s < t)$，而不依赖于 t 和 s 本身。
- 当增量具有平稳性时，称相应的独立增量过程是齐次的或时齐的。

3.6.1 概述

1 随机过程(Stochastic Process)

- 泊松过程

定义

随机过程$\{N(t), t \geq 0\}$称为**计数过程**或**点过程**，

如果$N(t)$表示从时刻0到t某一特定事件A发生的次数， 它具备以下两个特点：

（1） $N(t) \geq 0$且取值为整数;

（2） $s < t$时， $N(s) \leq N(t)$且$N(t) - N(s)$表示$(s, t]$时间内事件A发生的次数。

3.6.1 概述

定义

计数过程$\{N(t), t \geq 0\}$称为参数为$\lambda(\lambda > 0)$的泊松过程， 如果

(1) $N(0) = 0$;

(2) 过程有独立增量;

(3) 对任意的$s, t \geq 0$,

$$P(N(t+s) - N(s) = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \ n = 0, 1, 2, \ldots$$

3.6.1 概述

2 马尔科夫链（Markov Chain, MC）

如果$\{X_t\}$满足

$$P(X_{t+1}=j \mid X_0=k_0,\ldots X_{t-1}=k_{t-1},X_t=i)$$
$$=P(X_{t+1}=j \mid X_t=i) \qquad (19.1)$$
$$=p_{ij},t=0,1,\ldots,k_0,\ldots,k_{t-1},i,j\in S$$

称$\{X_t\}$为**马尔科夫链**

- P_{ij}为转移概率
- 矩阵$P=(p_{ij})_{m\times m}$为**转移概率矩阵**

3.6.1 概述

2 马尔科夫链（Markov Chain, MC）

当转移概率 $P_{ij}(m, m+n)$ 只与 i,j 以及时间间隔 n 有关时，把它记为 $P_{ij}(n)$，即：

$$P_{ij}(m, m+n) = P_{ij}(n)$$

称此转移概率具有**平稳性**，

称此链是齐次的或时齐的。

3.6.1 概述

2马尔科夫链

- 对马氏链，

$$P(X_{t+k} = \mathrm{j} | X_t = i) = p_{ij}^{(k)}$$

不依赖于t， 称为**k步转移概率**

- 如果对任意 $i, j \in S, i \neq j$
  都存在$k>=1$使得 $p_{ij}^{(k)} > 0$

  则称$\{X_t\}$为**不可约马尔科夫链**

3.6.1 概述

2马尔科夫链

- 马氏链{X_t}的某个状态i, 如果存在$k>=0$使得

$$p_{ii}^{(k)} > 0, p_{ii}^{(k+1)} > 0$$

则称i是**非周期的**

- 如果一个马氏链所有状态都是非周期的，则该马氏链称为非周期的。

3.6.1 概述

2马尔科夫链

- 对状态*i*， 如从状态*i*出发总能再返回状态*i*，则称状态*i*是**常返的**(recurrent)

- 对常返状态*i*， 如果从*i*出发首次返回*i*的时间的期望有限， 称*i*是**正常返的**

3.6.1 概述

2马尔科夫链

- 对只有有限个状态的非周期不可约马氏链有

$$\lim_{n\to\infty} P(X_n = j \mid X_0 = i) = \pi_j, i, j = 1, 2, \ldots, m$$

其中 $\{\pi_j, j = 1, 2 \ldots m\}$ 为正常数，满足 $\sum_{j=1}^{m} \pi_j = 1$

称为$\{X_t\}$的**极限分布**

3.6.1 概述

2马尔科夫链

- 满足方程组

$$\begin{cases} \sum_{i=1}^{m} \pi_i p_{ij} = \pi_j,\ j = 1,2,\ldots,m \\ \sum_{j=1}^{m} \pi_j = 1 \end{cases} \qquad (19.2)$$

称满足(19.2)的分布$\{\pi_j\}$为**平稳分布**或不变分布

3.6.1 概述

3 随机服务系统(Random Service System Theory)

又称排队论（Queuing theory）

- 通过研究各种服务系统在排队等待现象中的概率规律性
- 解决服务系统最优设计与最优控制问题
- 一种理论

3.6.1 概述

- 起源于20世纪初的电话通话。1909-1920年丹麦数学家电气工程师**爱尔朗**(A.K.Erlang)用概率论方法研究电话通话问题

- 20世纪30年代，费勒(W.Feller)引进了生灭过程
- 20世纪50年代，D.G.Kendall用嵌入马尔可夫链方法

3.6.1 概述

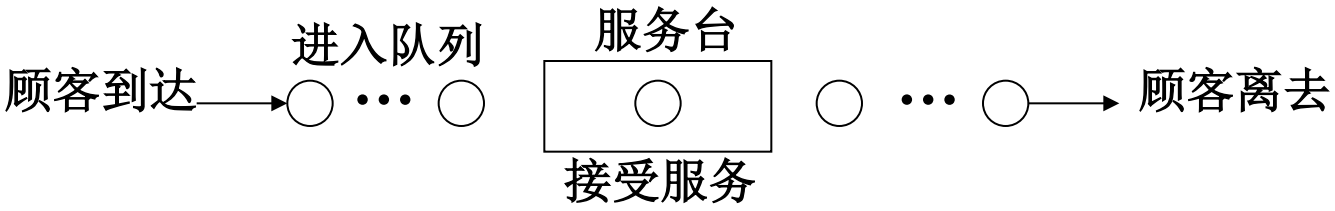
排队系统的例子

顾客	要求的服务	服务台
1．借书的学生	借书	图书管理员
2．打电话	通话	交换台
3．提货者	提货	仓库管理员
4．待降落的飞行器	降落	指挥塔台
5．储户	存款、取款	储蓄窗口、ATM取款机
6．河水进入水库	放水、调整水位	水库管理员
7．购票旅客	购票	售票窗口
8．十字路口的汽车	通过路口	红绿灯或交警

3.6.1 概述

根据服务台的数量及排队方式，排队系统可以分为四类

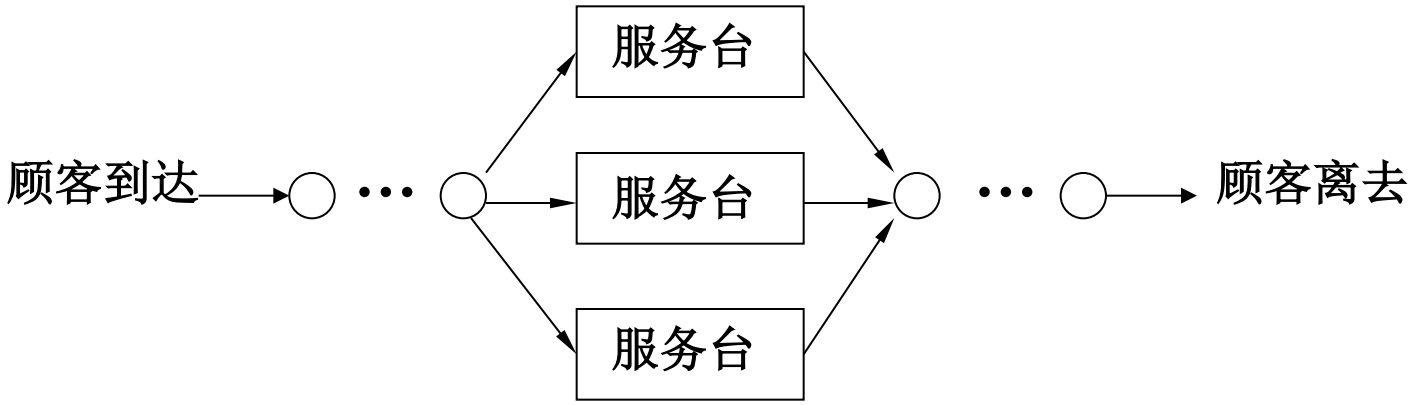
(1)单服务台单队



3.6.1 概述

根据服务台的数量及排队方式，排队系统可以分为四类

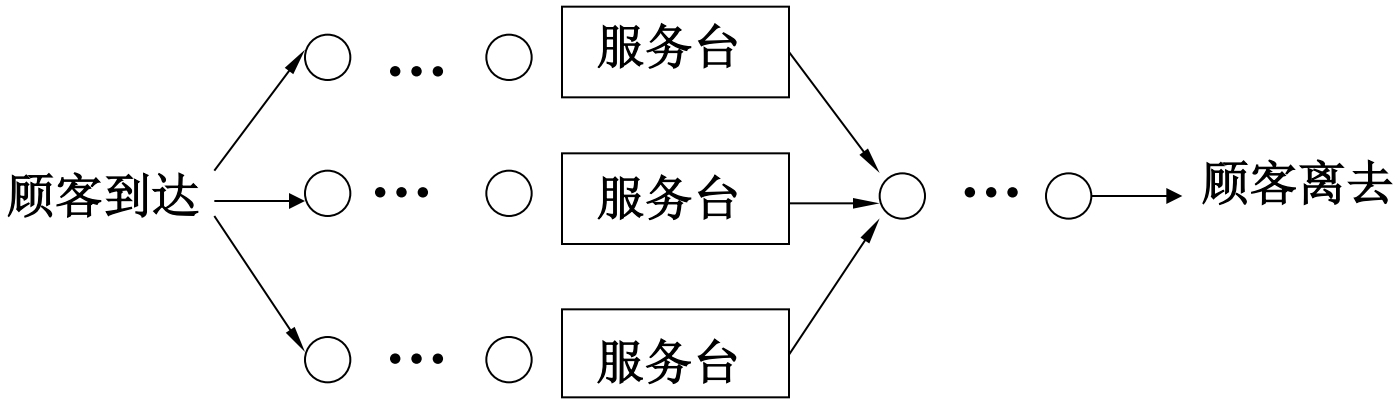
(2)多服务台单队



3.6.1 概述

根据服务台的数量及排队方式，排队系统可以分为四类

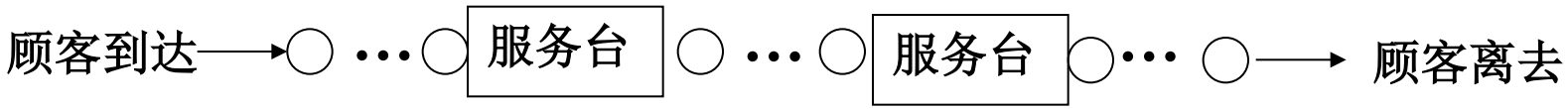
(3)多队多服务台



3.6.1 概述

根据服务台的数量及排队方式，排队系统可以分为四类

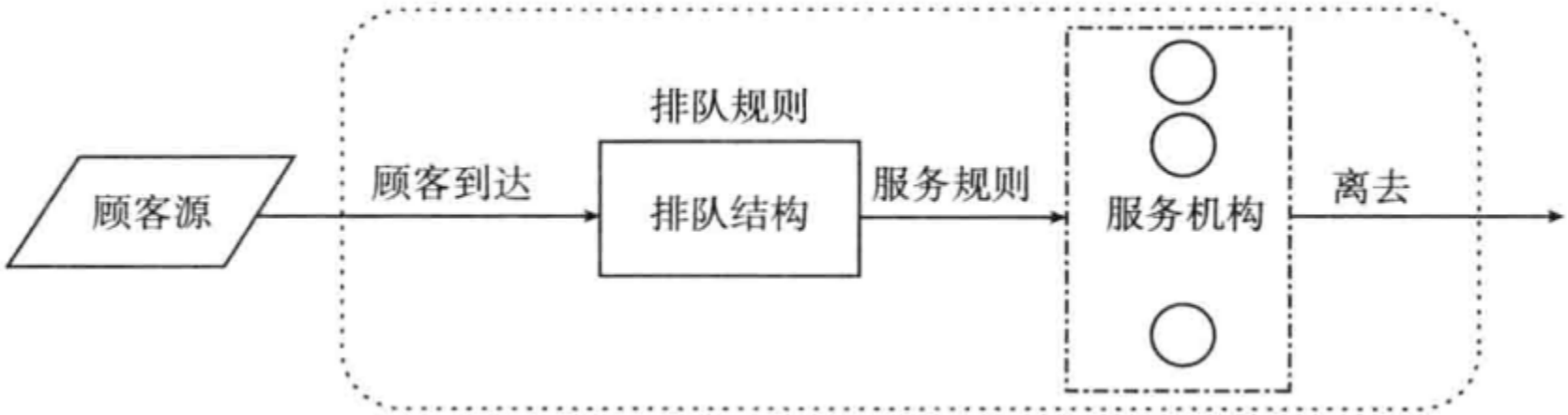
(4)多服务台串联服务



3.6.1 概述

随机服务系统模型应该包括三要素

- **输入过程**, 顾客按照怎样的规律到达
- **排队规则**, 顾客按照怎样的规则排队等待服务
- **服务机构**, 服务机构的设置,服务台的数量,服务的方式,服务时间分布等



3.6.1 概述

- 排队模型的分类

X/Y/Z/A/B/C

X — 顾客相继到达的间隔时间的分布；

Y — 服务时间的分布（M —负指数分布、D — 确定时间(deterministic)、Ek — k 阶埃尔朗分布、G — 一般分布等）；

Z — 服务台个数；

A — 系统容量限制（默认为 ∞）；

B — 顾客源数目（默认为 ∞）；

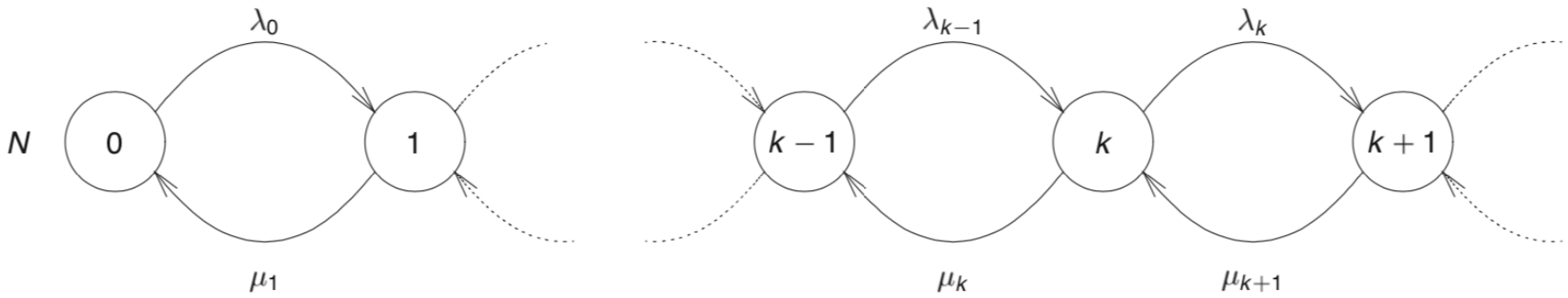
C — 服务规则 （默认为先到先服务 FCFS）。

3.6.1 概述

排队系统的衡量指标

- 服务队长 Ls — 正在接受服务的顾客数；
- 排队长 Lq — 在队列中等待的顾客数；
- 总队长 L = Ls + Lq — 系统中的顾客总数；
- 服务时间 Ws — 顾客在服务中消耗的时间；
- 等待时间 Wq — 顾客在队列中等待的时间
- 总时间 W = Ws + Wq—顾客在系统中的总逗留时间；
- 忙期 — 服务机构两次空闲的时间间隔；
- 服务强度 ρ；
- 稳态 — 系统运行充分长时间后，初始状态的影响基本消失，系统状态不再随时间变化

M/M/1排队系统



State diagram for an M/M/1 queue

M/M/1排队系统

对**到达速率**λ

P(在时间间隔 [t,t+Δt] 内，有且仅有一位顾客到达)=λΔt

P(在时间间隔 [t,t+Δt] 内，没有顾客到达)=1−λΔt

P(在时间间隔 [t,t+Δt] 内，多于一位顾客到达)=0

对**服务速率**μ

P(在时间间隔 [t,t+Δt] 内，有且仅有一位顾客服务完成)=μΔt

P(在时间间隔 [t,t+Δt] 内，没有顾客服务完成)=1−μΔt

P(在时间间隔 [t,t+Δt] 内，多于一位顾客服务完成)=0

M/M/1排队系统

假设：

$P_k(t)$表示在时间间隔t内，系统内有k位顾客的概率;

$p_{i,j}(\Delta t)$表示在时间间隔Δt内，系统内人数由i变为j的概率

时间间隔$t+\Delta t$内系统内人数为k的概率可由时间间隔t内系统内人数为k和$k-1$的概率表示：

$$P_k(t+\Delta t)=P_k(t)p_{k,k}(\Delta t)+P_{k-1}(t)p_{k-1,k}(\Delta t)+P_{k+1}(t)p_{k+1,k}(\Delta t)$$

$$P_0(t+\Delta t)=P_0(t)p_{0,0}(\Delta t)+P_1(t)p_{1,0}(\Delta t)$$ k=0(起始状态)

M/M/1排队系统

在泊松过程假设下顾客到达与离开的四种情况

情　况	时期 $[t, t+1)$ 内顾客数的变化	时期 $[t, t+\Delta t)$ 内		转移概率
		到　达	离　去	
A	$k \to k$	×	×	$(1-\lambda\Delta t)(1-\mu\Delta t)$
B	$k+1 \to k$	×	✓	$(1-\lambda\Delta t)(\mu\Delta t)$
C	$k-1 \to k$	✓	×	$(\lambda\Delta t)(1-\mu\Delta t)$
D	$k \to k$	✓	✓	$(\lambda\Delta t)(\mu\Delta t)$

$$P_k(t+\Delta t) = P_k(t)p_{k,k}(\Delta t) + P_{k-1}(t)p_{k-1,k}(\Delta t) + P_{k+1}(t)p_{k+1,k}(\Delta t)$$

M/M/1排队系统

将泊松过程的模型假设公式和M/M/1 排队系统的假设公式，代入上述公式：

$$P_k(t+\Delta t) = P_k(t)(1-\lambda\Delta t)(1-\mu\Delta t) + P_{k-1}(t)(\lambda\Delta t)(1-\mu\Delta t) + P_{k+1}(t)(\mu\Delta t)(1-\lambda\Delta t)$$

$$P_0(t+\Delta t) = P_0(t)(1-\lambda\Delta t) + P_1(t)(\mu\Delta t)(1-\lambda\Delta t), k=0$$

当Δt→0时有：

$$\frac{dP_k(t)}{dt} = -(\lambda+\mu)P_k(t) + \lambda P_{k-1}(t) + \mu P_{k+1}(t), k \geq 1$$

$$\frac{dP_0(t)}{dt} == -\lambda P_0(t) + \mu P_1(t)$$

考虑系统处于稳定状态(steady state)的情况，有：

$$\frac{dP_k(t)}{dt} = 0, k \geq 0$$

M/M/1排队系统

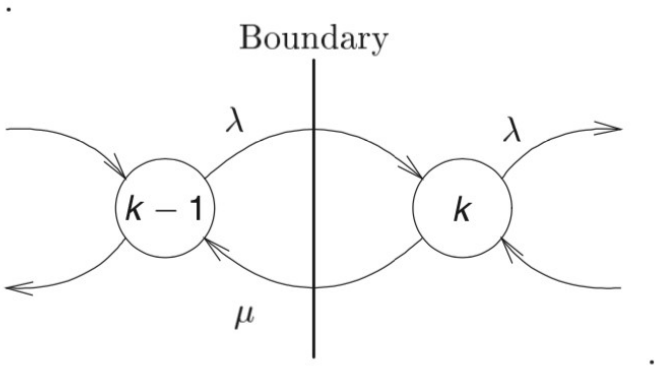
代入上式可得：

$$(\lambda+\mu)P_k-\lambda P_{k-1}-\mu P_{k+1}=0, k\geq 1$$

$$\lambda P_0-\mu P_1=0, k=0$$

解得：

$$\lambda P_{k-1}=\mu P_k$$



Flow rates at a boundary for an M/M/1 queue.

M/M/1排队系统

$$P_k = \frac{\lambda}{\mu} P_{k-1}$$

递推可得：

$$P_k = (\frac{\lambda}{\mu})^k P_0 = \rho^k P_0, \text{其中 } \rho = \frac{\lambda}{\mu}$$

下面计算P_0的值，由于：

$$\sum_{k=0}^{\infty} P_k = 1$$

将$P_k = \rho^k P_0$代入上式有：

$$P_0 \sum_{k=0}^{\infty} \rho^k = 1$$

M/M/1排队系统

故：

$$P_0 = (1 - \rho)$$

综上，处于稳定状态的M/M/1 排队系统的概率分布为：

$$P_k = \rho^k(1 - \rho), \rho < 1$$

3.6.1 概述

排队系统的应用----互联网容量保障

1. 设定合理的 CPU 利用率预警阈值

2. 利用排队论进行容量规划
设一个服务的平均流量为λ，它平均每秒能处理μ个请求，我们最多能够接受 20% 的请求处于排队状态，请问需要多少台服务器能够达到这个目标。

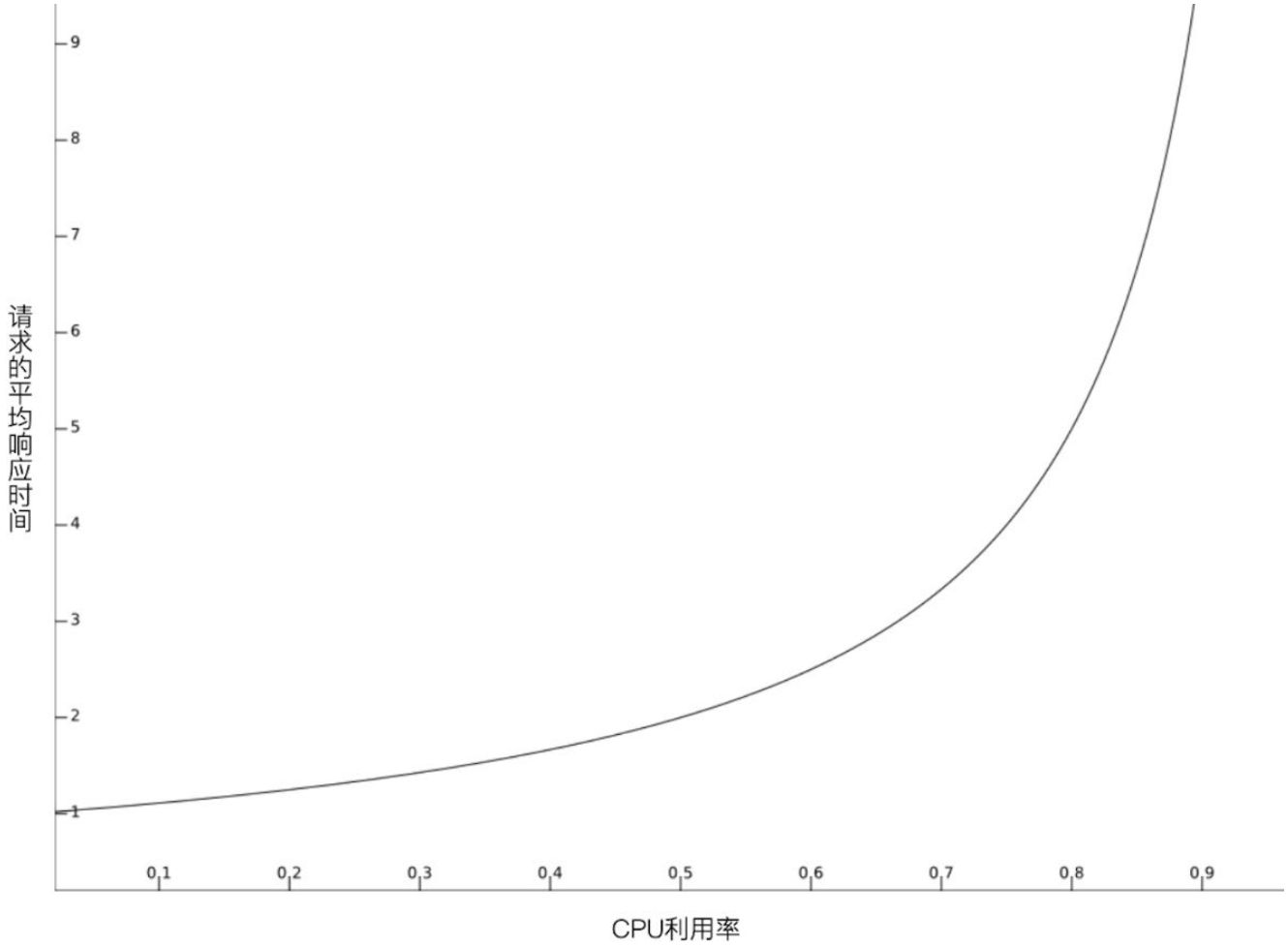
M/M/s 模型的三大基本公式

- **系统利用率公式** $\rho = \lambda / s\mu$：其中λ指的是到达率，每秒到达的请求数，也可以认为是流量。μ为服务率，也就是每秒能处理的请求数，sμ就是系统总共能处理的请求数。

- **资源需求公式** $R = \lambda / \mu$：

- **平均等待时间公式** $E[T] = 1 / (s\mu - \lambda)$

设定合理的 CPU 利用率预警阈值



- 平方根配置规则

给定一个 M/M/s 模型，
令 k 表示确保 “请求排队概率小于α”所需的最少服务器数，
则有 $k \approx R + c\sqrt{R}$，
其中 c 是方程的解 $c\Phi(c)/\varphi(c)=(1-\alpha)/\alpha$，
$\Phi(c)$ 表示标准正态的累积分布函数，
$\varphi(c)$ 表示它的概率密度函数。

α=0.8	α=0.5	α=0.2	α=0.1
c=0.173	c=0.506	c=1.06	c=1.42

莫尔·哈肖尔 - 巴尔特. 计算机系统的性能建模与设计：排队论实战[M]. 方娟，蔡旻，张佳玥译. 机械工业出版社，2020：187-189.

3.6.2 离散事件模拟

离散事件模拟
(Discrete Event Simulation, DES)

在模拟随机服务系统时，按时间顺序记录发生的事件，

如顾客到来、顾客接受服务、顾客结束服务等

3.6.2 离散事件模拟

离散事件模拟算法可以分为三类：

- 活动模拟
- 事件模拟
- 过程模拟

3.6.2 离散事件模拟

事件模拟算法必须考虑的变量包括：

- 当前时刻t;

- 随时间而变化的计数变量， 如t时刻时到来顾客人数、已离开人数；

- 系统状态，比如是否有顾客正在接受服务、排队人数、队列中顾客序号。

3.6.2 离散事件模拟

用随机服务系统进行建模和模拟研究一般步骤如下

- 提出问题
- 建立模型
- 数据收集与处理
- 建立模拟程序
- 模拟模型的正确性确认
- 模拟试验
- 模拟结果分析

例15.1

- 设银行仅有一个柜员， 并假设银行不休息
- 顾客到来间隔的时间服从独立指数分布Exp(λ) (1/λ为间隔时间的期望值)
- 如果柜员正在为先前的顾客服务，新到顾客就排队等待
- 柜员为顾客服务时间服从均值为1/μ的指数分布
- 设u= λ /μ<1
- 设X_t表示t时刻在银行内的顾客数（包括正在服务的和正在排队的）
- 则X_t是一个连续时马氏链
- 这是一个生灭过程马氏链

例15.1

- 当系统处于稳定状态时

$$P(X_t = i) = u^i(1-u), i = 0,1,2,\ldots$$

- 设随机变量*N*服从X_t的平稳分布
- 则银行中**平均顾客数**为

$$EN = \frac{u}{1-u}$$

例15.1

- 平均队列长度EQ等于EN减去平均正在服务人数
- 正在服务人数Y_t为

$$Y_t = \begin{cases} 1, & \text{as } X_t > 0 \\ 0, & \text{as } X_t = 0 \end{cases}$$

- 则**平均队列长度**为

$$EQ = EN - EY_t = \frac{u}{1-u} - u = \frac{u^2}{1-u}$$

例15.1

- 设顾客平均滞留时间为ER，由关系式

$$EN = \lambda \cdot ER$$

- 则**平均滞留时间**为

$$ER = \frac{u}{\lambda(1-u)} = \frac{1}{\mu - \lambda}$$

- 顾客滞留时间*R*服从均值为1/(μ-λ)的指数分布

例15.2

- 用事件模拟的方法来模拟例15.1。
- 估计平均滞留时间ER

{初始化当前时钟$t \leftarrow 0$, 柜员忙标志$B \leftarrow 0$, 当前排队人数$L \leftarrow 0$,

最新到来顾客序号$i \leftarrow 0$, 正在服务顾客序号$j \leftarrow 0$, 已服务顾客数$n \leftarrow 0$ }

从Exp(λ)抽取X, 设置下一顾客来到时间$A \leftarrow X$

repeat {

  if($B = 0$ **or** ($B = 1$ **and** $A < E$)) { # E是正在服务的顾客结束时刻

    $t \leftarrow A$

  } **else** {

    $t \leftarrow E$

  }

  if ($t > T_1$) **break** # T_1是预先确定的模拟时长

例15.2

- 用事件模拟的方法来模拟例15.1。
- 估计平均滞留时间ER

```
if(t == A) { # 待处理到达事件
    L ← L + 1
    i ← i + 1, 记录第i位顾客到来时间a_i ← t
    从Exp(λ)抽取X, A ← t + X
    if(B == 0 ) { # 不用排队，直接服务
        B ← 1, L ← L − 1
        j ← j + 1, 置第j位顾客开始服务时间s_j ← t
        从Exp(μ)抽取Y, 置E ← t + Y
    }
} else { # 待处理结束服务事件
```

例15.2

- 用事件模拟的方法来模拟例15.1。
- 估计平均滞留时间ER

$B \leftarrow 0$

$n \leftarrow n+1$, 记录第n个顾客结束服务时间$e_n \leftarrow t$

if($L > 0$) { # 排队顾客开始服务

$L \leftarrow L-1$

$B \leftarrow 1$

$j \leftarrow j+1$, $s_j \leftarrow t$

从Exp(μ)抽取Y, 置$E \leftarrow t+Y$

}

}

}

令$I=\{i:\ T_0 \le s_i \le T_1\}$, 求$\{e_i - a_i,\ i \in I\}$的平均值作为$ER$估计

例 医院预检处

在一个中等规模的医院都设有预检处,预检员的工作主要是帮助患者或访客对病类的甄别,使他们能正确地挂号,到医院的相关诊疗科室就诊.
随着计算机的普及,医院会在接待大厅设置计算机自助挂号的触摸屏来取代人工接待,患者或访客要使用计算机提供的导航信息为自己挂号.但有些患者或访客可能不会操作或因不熟练的操作而引起耽误,所以往往会因此出现排队现象.这种现象就需要医院的信息管理工程师来评估计算机挂号的效率及可能引起的意外耽误程度.

例 医院预检处

```
Total_time=10;%总迭代时间
lambda=65;mu=60;%到达率与服务率
arr_mean=1/lambda;ser_mean=1/mu;%平均到达时间与平均服务时间
%可能到达的最大顾客数(round:四舍五入求整数)
arr_num=round(Total_time*lambda*2);
guests=zeros(5,arr_num);%定义顾客信息的数组
%按指数分布产生各顾客到达的时间间隔
guests(1,:)=exprnd(arr_mean,1,arr_num);
%各顾客的到达时刻等于时间间隔的累积和
guests(1,:)=cumsum(guests(1,:));
%按指数分布产生各顾客服务时间
guests(2,:)=exprnd(ser_mean,1,arr_num);
%计算模拟的顾客个数,即到达时刻在模拟时间内的顾客数
1en_sim=sum(guests(1,:)<=Total_time);
%初始化第1个顾客的信息
guests(3,1)=0; %第1个顾客进入系统后直接接受服务,无需等待
%其离开时刻等于其到达时刻与服务时间之和
guests(4,1)=guests(1,1)+guests(2,1);
guests(5,1)=0;%此时系统内没有其他顾客,故附加信息为0
member=[1];%其进入系统后,系统内已有成员序号为1
```

例 医院预检处

```
%计算第i个顾客的信息
for i=2:arr_num
%如果第i个顾客的到达时间超过了迭代时间,则跳出循环
    if guests(1,i)>Total_time
        break;
   else
%如果第i个顾客的到达时间在迭代时间内,则计算在其到达时刻系统中已有的顾客数
       number=sum(guests(4,member)>guests(1,i));
       if number==0
%如果系统为空,则第i个顾客直接接受服务,其等待时间为0
           guests(3,i)=0;
%其离开时刻等于到达时刻与服务时间之和
           guests(4,i)=guests(1,i)+guests(2,i);
           guests(5,i)=0; %其附加信息是0
           member=[member,i];
       else
%如果系统有顾客正在接受服务,且系统等待队列未满,则第i个顾客进入系统
           1en_mem=length(member);
%其等待时间等于队列中前一个顾客的离开时刻减去其到达时刻
           guests(3,i)=guests(4,member(len_mem))-guests(1,i);
%其离开时刻等于队列中前一个顾客的离开时刻加上其服务时间
           guests(4,i)=guests(4,member(len_mem))+guests(2,i);
%附加信息表示其进入系统时在他之前系统已有的顾客数
           guests(5,i)=number;
           member=[member,i];
       end
   end
end
len_mem=length(member); %模拟结束时,进入系统的总顾客数
```

例 医院预检处

```
%输出结果%
%绘制在模拟时间内,进入系统的所有顾客的到达时刻和离开时刻的曲线图
figure(1)
stairs(1:len_mem,guests(1,member));
hold on
stairs(1:len_mem,guests(4,member),'.r-')
legend('到达时间','离开时间')
hold off
grid on
%绘制在模拟时间内,进入系统的所有顾客的停留时间和等待时间的曲线图
figure(2)
plot(1:len_mem,guests(3,member),'r-
',1:len_mem,guests(2,member)+guests(3,member),'k-')
legend('等待时间','停留时间')
grid on
```

3.7 Bootstrap方法

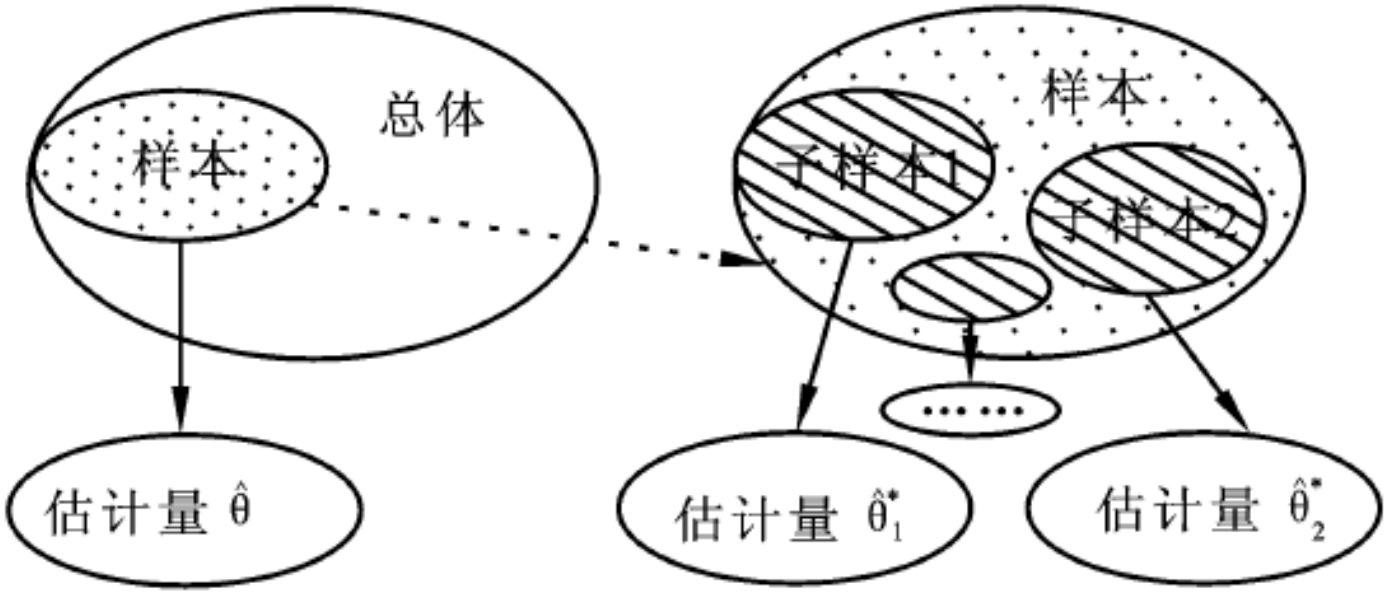
3.7.1 BOOTSTRAP方法引入
3.7.2 非参数BOOTSTRAP方法
3.7.3 参数BOOTSTRAP方法

3.7.1 Bootstrap方法引入

- 统计量的分布是关键
- 统计量的分布很难求
- 计算参数估计的标准误差不一定总有简单公式
- 线性模型估计
- 最大似然估计问题

3.7.1 Bootstrap方法引入

重抽样方法思想示意图



3.7.1 Bootstrap方法引入

- 在对观测值重复抽样，每次用重抽样数据生成一个经验分布函数
- 对每个重抽样数据集或者等价地经验分布函数，可以计算统计量的一个新的取值,收集这些值可以给出受关注的统计量的抽样分布的一个估计
- Bootstrap用在独立同分布的数据上。也可用在不独立的数据上,比如回归残差或者时间序列数据（需要先处理数据）.

3.7.1 Bootstrap方法引入

设总体X服从某个未知分布F(x)，$\boldsymbol{X}=(x_1,x_2,\ldots,x_n)$是X的一个样本，

ϕ是F的一个参数，(可以是标准误差、置信区间或者p值

可以把ϕ看成F的一个泛函$\phi(F)$，用统计量$\hat{\phi}=g(\boldsymbol{X})$估计$\phi$，

设$\psi=\psi(g,F,n)$是统计量$\hat{\phi}$的某种分布特征($\hat{\phi}$的抽样分布的数字特征)。

例如$\psi=\sqrt{\mathrm{Var}(\bar{X})}$为统计量$\bar{X}$的标准误差，

又如取$\psi=E\hat{\phi}-\phi$为统计量$\hat{\phi}$的偏差。

可以用**随机模拟**的方法估计ψ。

3.7.1 Bootstrap方法引入

用随机模拟方法估计ψ的**步骤**如下。

1. 从样本$\boldsymbol{X}$估计总体分布F为$\hat{F}$;
2. 从$\hat{F}$抽取B个独立样本$\boldsymbol{Y}^{(b)}, b = 1, \ldots, B$，每一个$\boldsymbol{Y}^{(b)}$样本量为n, 称$\boldsymbol{Y}^{(b)}$为**bootstrap样本**。
3. 从每个bootstrap样本$\boldsymbol{Y}^{(b)}$可以**估计**得到$\hat{\phi}^{(b)} = g(\boldsymbol{Y}^{(b)}), (b = 1, \ldots, B)$。
4. $\hat{\phi}^{(b)}, (b = 1, \ldots, B)$是$g(\boldsymbol{Y})$在$\hat{F}$下的独立同分布样本, 可以用标准的估计方法估计关于$g(\boldsymbol{Y})$在$\hat{F}$下的分布特征$\hat{\psi} = \psi(g, \hat{F}, n)$, 估计结果记作$\tilde{\psi}$，并以$\tilde{\psi}$作为统计量$\hat{\phi}$的抽样分布的数字特征$\psi(g, F, n)$的估计值。

3.7.1 Bootstrap方法引入

- 非参数bootstrap方法
- 参数bootstrap方法
- 可用于当人们对总体知之甚少的情况
- 是近代统计中的一种用于数据处理的重要的实用方法

3.7.2 非参数bootstrap方法

1，估计量的标准误差的bootstrap估计

2，估计量的均方误差及偏差的 bootstrap估计

3，bootstrap置信区间-分位数法

4，bootstrap置信区间-bootstrap-t法

1，估计量的标准误差的bootstrap估计

- **标准误差**

设总体$X \sim F(x, \phi), \phi \in \Theta$，

ϕ 是总体的一个参数

$\hat{\phi}$是ϕ的估计量，

称$\mathrm{SE} = \sqrt{\mathrm{Var}(\hat{\phi})}$为$\hat{\phi}$的**标准误差**。

1，估计量的标准误差的bootstrap估计

例17.1

设总体$X \sim F(x)$

$X_i, (i = 1, \ldots, n)$ 是来自总体的样本

- 样本平均值$\hat{\phi} = \bar{X} = \frac{1}{n}\sum_i X_i$为$\phi = EX$的点估计
- $\mathrm{SE}(\bar{X}) = \sqrt{\mathrm{Var}(X)/n}$，可以用$S/\sqrt{n}$估计SE($S^2$为样本方差)
- 根据中心极限定理和强大数律，当样本量n较大时可以取EX 的近似95%置信区间为$\bar{X} \pm 2\,\mathrm{SE}(\bar{X})$

1，估计量的标准误差的bootstrap估计

例17.2 线性模型参数估计的SE

设模型为$\mathbf{Y} = \mathbf{X\beta} + \boldsymbol{\varepsilon}$

其中: $\boldsymbol{\varepsilon} \sim \mathrm{N}(0, \sigma^2 I_n)$, σ^2未知， $\boldsymbol{\beta}$是未知系数向量, X是已知的$n \times p$数值矩阵, n > p。

在X列满秩时， $\boldsymbol{\beta}$的最小二乘估计为

$$\hat{\boldsymbol{\beta}} = (X^T X)^{-1} X^T \mathbf{Y}$$

1，估计量的标准误差的bootstrap估计

例17.2 线性模型参数估计的SE

$\hat{\boldsymbol{\beta}}$的协方差阵为$\mathrm{Var}(\hat{\boldsymbol{\beta}}) = \sigma^2 (X^T X)^{-1}$

第j个系数β_j的**标准误差**可估计为

$$\mathrm{SE}(\hat{\beta}_j) = \hat{\sigma}\sqrt{a^{(jj)}}$$

其中$\hat{\sigma}$是σ的估计，$a^{(ij)}$为$(X^T X)^{-1}$的(i,j)元素

1，估计量的标准误差的bootstrap估计

设总体$X \sim F(x,\phi), \phi \in \Theta$，

ϕ是我们感兴趣的一个总体的参数

$\hat{\phi}$=g(X)是总体参数ϕ的估计量

设分布**F未知**,

设$x_1, x_2, \ldots, x_n$是来自F的样本

求$\hat{\phi}$的**标准误差**SE的bootstrap估计

1，估计量的标准误差的bootstrap估计

- 设F_n是相应的经验分布函数.
- 当n很大时, F_n接近F
- 用F_n代替 F, 在F_n中抽样.
- 在F_n中抽样,就是在原始样本$x_1, x_2, \ldots, x_n$中每次随机地取一个个体作**有放回抽样**

1，估计量的标准误差的bootstrap估计

如此抽取B个独立样本$\mathbf{Y}^{(b)}, b = 1, \ldots, B$，

每一个$\mathbf{Y}^{(b)}$样本量为n

称$\mathbf{Y}^{(b)}$为**bootstrap样本**

- 从每个bootstrap样本$\mathbf{Y}^{(b)}$可以估计得到 $\hat{\phi}^{(b)} = g\left(\mathbf{Y}^{(b)}\right), b = 1, \ldots, B$

$\hat{\phi}^{(b)}, b = 1, \ldots, B$是$g(\mathbf{Y})$在$\hat{F}$下的独立同分布**样本**

求SE的bootstrap估计的步骤:

1° 自原始数据样本$x = (x_1, x_2, \dots, x_n)$按放回抽样的方法, 抽得容量为n的样本$Y = (Y_1, Y_2, \dots, Y_n)$

2° 相继地、独立地求出B个(B≥1000)容量为n的bootstrap样本,$Y^{(i)} = \left(Y_1^{(i)}, Y_2^{(i)}, \dots, Y_n^{(i)}\right)$,i=1,2...,B.

对于第i个bootstrap样本,计算$\hat{\phi}^{(i)}$

3° 计算$SE = \sqrt{\frac{1}{B-1}\sum_{i=1}^{B}\left(\hat{\phi}^{(i)} - \bar{\phi}\right)^2}$

其中$\bar{\phi} = \frac{1}{B}\sum_{i=1}^{B}\hat{\phi}^{(i)}$

例17.5

设(H,W)为某地小学五年级学生的身高和体重的总体，$(H,W) \sim F(\cdot,\cdot)$,

求估计H和W的**相关系数ϕ**估计量的标准误差估计。

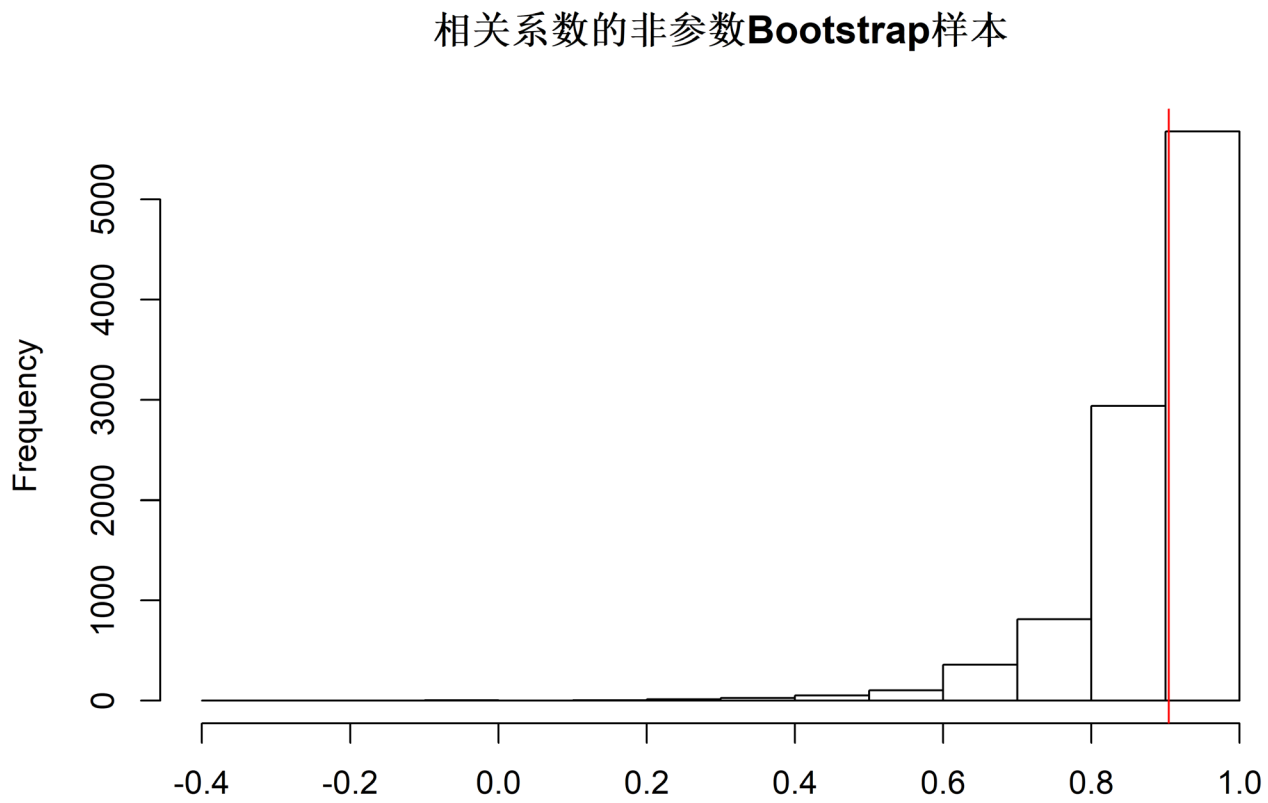
设调查了n=10个学生的身高和体重的数据$(h_i, w_i), (i = 1,2,\ldots,n)$:

h_i	144	166	163	143	152	169	130	159	160	175
w_i	38	44	41	35	38	51	23	51	46	51

例17.5 计算步骤如下：

1. 从F_n中作n=10次独立抽样，即从$\{(h_1, w_1), \ldots, (h_n, w_n)\}$中**有放回**独立抽取n次，得到$\hat{F} = F_n$的一组样本$\mathbf{Y}^{(1)} = \left(\left(h_1^{(1)}, w_1^{(1)}\right), \ldots, \left(h_n^{(1)}, w_n^{(1)}\right)\right)$;
2. 重复第(1)步, 直到获取了**B组**bootstrap样本$\mathbf{Y}^{(b)}, b = 1, \ldots, B$;
3. 对每一样本$\mathbf{Y}^{(b)}$计算样本相关系数$\hat{\phi}^{(b)} = g\left(\mathbf{Y}^{(b)}\right)$
4. 把$\hat{\phi}^{(b)}, b = 1, \ldots, B$作为$\hat{F}$下n=10 的样本相关系数的简单随机样本，估计其样本标准差S，以S作为$\psi\left(g, \hat{F}, n\right)$的估计, 进而用S估计$\hat{\phi}$在真实的总体分布F下的标准误差$\mathrm{SE}\left(\hat{\phi}\right)$。

例17.5


相关系数的非参数Bootstrap样本
Frequency
0
1000
2000
3000
4000
5000
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8
1.0

1，估计量的标准误差的bootstrap估计

例1：某种基金的年回报率是具有分布函数F的连续型随机变量, F未知, F的中位数θ是未知参数.

现有以下的数据(%率):18.2, 9.5, 12.0, 21.1, 10.2

以样本中位数作为总体中位数θ的估计.

试求**中位数估计的标准误差的bootstrap估计**.

1，估计量的标准误差的bootstrap估计

解 将原始样本自小到大排序,中间一个数为12.0,得样本中位数为12.0.

相继地、独立地在上述5个数据中,按放回抽样的方法取样, 取B=10得到下述10个bootstrap样本:

样本1 9.5,18.2,12.0,10.2,18.2
样本2 21.1, 18.2, 12.0, 9.5, 10.2
样本3 21.1, 10.2, 10.2, 12.0, 10.2
样本4 18.2, 12.0, 9.5, 18.2, 10.2
样本5 21.1, 12.0, 18.2, 12.0, 18.2
样本6 10.2, 10.2, 9.5, 21.1, 10.2
样本7 9.5, 21.1, 12.0, 10.2, 12.0
样本8 10.2, 18.2, 10.2, 21.1, 21.1
样本9 10.2, 10.2, 18.2, 18.2, 18.2
样本10, 18.2, 10.2, 18.2, 10.2, 10.2

1，估计量的标准误差的bootstrap估计

解

对以上每个bootstrap样本, 求得样本中位数分别为
12.0,12.0,10.2,12.0,18.2,10.2,12.0,18.2,18.2,10.2
是以原始样本确定的样本中位数$\hat{\theta} = 12.0$作为总体中位数θ的估计,

其标准误差的bootstrap估计为

$$\hat{\psi} = \sqrt{\frac{1}{9}\sum_{i=1}^{10}(\hat{\theta}^{(i)} - \bar{\theta})^2} = 3.4579$$

2 估计量的均方误差及偏差的 bootstrap估计

例2 均方误差

设金属元素铂的升华热是具有分布函数F的连续型随机变量,F的中位数是未知参数,现测得以下的数据(以kcal/mol计)

136.3	136.6	135.8	135.4	134.7	135.0	134.1	143.3	147.8
148.8	134.8	135.2	134.9	149.5	141.2	135.4	134.8	135.8
135.0	133.7	134.4	134.9	134.8	134.5	134.3	135.2	

求中位数θ的**均方误差**$E\left(\hat{\theta}-\theta\right)^2$的bootstrap估计.

（以样本中位数作为总体中位数θ的估计）

2 估计量的均方误差及偏差的 bootstrap估计

解将原始样本中位数为(135.0+135.2)=135.1.以135.1作为总体中位数θ的估计,

相继地、独立地抽取5个bootstrap样本如下:

Sample 1: 134.3, 135.4, 134.1, 136.3, 136.6, 134.8, 134.8, 149.5, 135.8, 135, 136.3, 134.4, 134.4, 148.8, 134.9, 149.5, 148.8, 134.1, 147.8, 141.2, 134.4, 134.7, 147.8, 134.3, 149.5, 134.7 | Median: 135.2

Sample 2: 134.7, 136.6, 148.8, 134.3, 135.2, 141.2, 136.3, 133.7, 135.8, 135, 148.8, 148.8, 135, 141.2, 133.7, 133.7, 135.2, 135.2, 134.3, 134.8, 135, 134.3, 143.3, 135.2, 134.3, 134.8 | Median: 135.1

Sample 3: 134.5, 135, 135.2, 134.1, 135, 148.8, 135, 134.5, 149.5, 136.6, 134.9, 135.8, 134.9, 149.5, 135, 136.3, 135, 135, 143.3, 135, 134.8, 135.2, 135, 143.3, 134.1, 134.8 | Median: 135

Sample 4: 134.8, 135.4, 134.9, 143.3, 134.3, 135.4, 134.3, 134.8, 149.5, 133.7, 134.1, 134.9, 134.9, 135.2, 135.4, 136.3, 134.9, 134.4, 135, 134.8, 147.8, 134.1, 134.8, 135, 134.9, 135.8 | Median: 134.9

Sample 5: 135.2, 135.4, 134.8, 148.8, 134.1, 135, 136.6, 148.8, 136.3, 134.8, 135.2, 141.2, 135.2, 134.5, 148.8, 135.4, 134.4, 135.2, 134.1, 134.5, 134.4, 134.9, 143.3, 136.3, 135, 135.2 | Median: 135.2

2 估计量的均方误差及偏差的 bootstrap估计

将上述5个数取平均值得到MSE的bootstrap估计为

$$\frac{1}{5}\sum_{i=1}^{5}\left(\hat{\theta}^{(i)}-135.1\right)^2 = \square.\square\square\square$$

即得中位数的均方误差MSE的bootstrap估计为0.014

2 估计量的均方误差及偏差的 bootstrap估计

例 3 偏差

试在例2中, 求中位数θ的偏差的bootstrap估计.

解，设$X = (X_1, X_2, \ldots, X_n)$是来自总体F的样本,

$\hat{\theta} = \hat{\theta}(X_1, X_2, \ldots, X_n)$是中位数θ的估计量.

以样本中位数作为总体中位数θ的估计

由例2知**原始样本**的中位数为135.1.

以135.1作为总体中位数θ的估计,

取$\psi = \hat{\theta} - \theta$, 需估计$E(\hat{\theta} - \theta)$

2 估计量的均方误差及偏差的 bootstrap估计

bootstrap: 134.3 135.4 134.1 136.3 136.6 134.8 134.8 149.5 135.8 135 136.3 134.4 134.4 148.8 134.9 149.5 148.8 134.1 147.8 141.2 134.4 134.7 147.8 134.3 149.5 134.7 median: 135.2
bootstrap: 134.7 136.6 148.8 134.3 135.2 141.2 136.3 133.7 135.8 135 148.8 148.8 135 141.2 133.7 133.7 135.2 135.2 134.3 134.8 135 134.3 143.3 135.2 134.3 134.8 median: 135.1
bootstrap: 134.5 135 135.2 134.1 135 148.8 135 134.5 149.5 136.6 134.9 135.8 134.9 149.5 135 136.3 135 135 143.3 135 134.8 135.2 135 143.3 134.1 134.8 median: 135
bootstrap: 134.8 135.4 134.9 143.3 134.3 135.4 134.3 134.8 149.5 133.7 134.1 134.9 134.9 135.2 135.4 136.3 134.9 134.4 135 134.8 147.8 134.1 134.8 135 134.9 135.8 median: 134.9
bootstrap: 135.2 135.4 134.8 148.8 134.1 135 136.6 148.8 136.3 134.8 135.2 141.2 135.2 134.5 148.8 135.4 134.4 135.2 134.1 134.5 134.4 134.9 143.3 136.3 135 135.2 median: 135.2

2 估计量的均方误差及偏差的 bootstrap估计

将上述1000个数取平均值得到偏差b的bootstrap估计为

$$b^{\star} = \frac{1}{10000}\sum_{i=1}^{10000}\left(\hat{\theta}^{(i)} - 135.1\right)$$

$$= \frac{1}{10000}\sum_{i=1}^{10000}\hat{\theta}^{(i)} - 135.1 = 135.14 - 135.1$$

$$= 0.04$$

3，bootstrap置信区间-分位数法

- 从样本$x = (x_1, x_2, \ldots, x_n)$中抽出B个容量为n的**bootstrap样本**
- 对于每个bootstrap样本求出θ的**bootstrap估计**:

$$\hat{\theta}_1^\star, \hat{\theta}_2^\star, \ldots, \hat{\theta}_B^\star$$

- 将它们自小到大排序,得

$$\hat{\theta}_{(1)}^\star, \hat{\theta}_{(2)}^\star, \ldots, \hat{\theta}_{(B)}^\star$$

3，bootstrap置信区间-分位数法

用对应的$\hat{\theta}^*$的分布作为$\hat{\theta}$ 的分布的近似，

求出$\hat{\theta}^*$的分布的近似分位数$\hat{\theta}^*_{\frac{\alpha}{2}}$和$\hat{\theta}^*_{1-\frac{\alpha}{2}}$

使 $$P\{\hat{\theta}^*_{\frac{\alpha}{2}} < \hat{\theta}^* < \hat{\theta}^*_{1-\frac{\alpha}{2}}\} = 1-\alpha$$

于是近似地有

$$P\{\hat{\theta}^*_{\frac{\alpha}{2}} < \theta < \hat{\theta}^*_{1-\frac{\alpha}{2}}\} = 1-\alpha$$

若记$k_1 = \left[B \times \frac{\alpha}{2}\right], k_2 = \left[B \times \left(1-\frac{\alpha}{2}\right)\right]$,

3，bootstrap置信区间-分位数法

以$\hat{\theta}^{\star}_{(k_1))}$和$\hat{\theta}^{\star}_{(k_2)}$分别作为分位数$\hat{\theta}^{\star}_{\frac{\alpha}{2}}$和$\hat{\theta}^{\star}_{1-\frac{\alpha}{2}}$的**估计**，

得到近似等式

$$P\{\hat{\theta}^{\star}_{(k_1)} < \theta < \hat{\theta}^{\star}_{(k_2)}\} = 1 - \alpha$$

于是由上式就得到θ的置信水平为1-α的近似置信区间$\left(\hat{\theta}^{\star}_{(k_1)}, \hat{\theta}^{\star}_{(k_2)}\right)$

这一区间称为θ的置信水平为1-α的**bootstrap置信区间.**

这种求置信区间的方法称为**分位数法**

3，bootstrap置信区间-分位数法

例4 在例2中

以样本中位数作为总体中位数θ的估计求θ的置信水平为0.95的**bootstrap 置信区间**

3，bootstrap置信区间-分位数法

解 n=26,B=10000,原始样本以及10000个模拟bootstrap样本见例2.

对于每一个bootstrap样本算出中位数$M_1^\star, M_2^\star, \ldots, M_{10000}^\star$.

将它们自小到大排序得到

$$M_{(1)}^\star \le M_{(2)}^\star \le \cdots \le M_{(250)}^\star \le M_{(251)}^\star \le \cdots \le M_{(9750)}^\star \le M_{(9751)}^\star \le \cdots \le M_{(10000)}^\star$$

由B=10000,1-a=0.95,a=0.05,

k1=|10000×0.05/2|=250,k2=|10 000×(1-0.05/2)1=9750.

bootstrap置信区间为$\left(M_{(250)}^\star, M_{(9750)}^\star\right) = (134.8,135.8)$.

4，bootstrap置信区间-bootstrap-t法

以**求期望**ϕ的bootstrap置信区间为例。

枢轴量法是构造置信区间的最基本的方法。

设ϕ是**总体**$F(\cdot)$的一个参数（比如期望）。

$\mathbf{X} = (x_1, x_2, \ldots, x_n)$为来自总体的样本，容量为n

均值和方差均为未知参数，我们要利用样本值来估计期望ϕ

考虑函数$g(\mathbf{X})$为与ϕ有关系的一个统计量， 经常是ϕ的估计量。

$$g(\mathbf{X}) = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

假设总体F具有正态分布, 此时$g(\mathbf{X})$的分布与参数ϕ无关, 它是一个**枢轴量**而且有$g(\mathbf{X})$~t(n-1),利用枢轴量$g(\mathbf{X})$,就能求得ϕ的置信区间.

4，bootstrap置信区间-bootstrap-t法

用bootstrap方法来求ϕ的近似置信区间.

以原始样本$\mathbf{X}=(x_1,x_2,\ldots,x_n)$的样本均值$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$作为$\phi$的估计，考虑与$g(\mathbf{X})$相应的枢轴量

$$W^*=\frac{\bar{X}^*-\bar{x}}{S^*/\sqrt{n}}$$

其中：$\bar{X}^*,S^*$分别为与$\bar{X}$，S相应的bootstrap样本均值与样本标准差.

用$\boldsymbol{W}^*$的分布近似$g(\mathbf{X})$的分布，求出W^*的近似分位数为$w^*_{\frac{\alpha}{2}}$和$w^*_{1-\frac{\alpha}{2}}$

$$P\left\{w^*_{\frac{\alpha}{2}}<\frac{\bar{X}^*-\bar{x}}{S^*/\sqrt{n}}<w^*_{\frac{1-\alpha}{2}}\right\}=1-\alpha$$

4，bootstrap置信区间-bootstrap-t法

于是近似地有

$$P\left\{w^*_{\frac{\alpha}{2}} < \frac{\bar{X}-\phi}{S/\sqrt{n}} < w^*_{1-\frac{\alpha}{2}}\right\} = 1-\alpha$$

$$P\left\{\bar{X} - w^*_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n}} < \phi < \bar{X} - w^*_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right\} = 1-\alpha$$

将W^*的B个bootstrap值自小到大排序

$$w^*_{(1)} \le w^*_{(2)} \le \cdots \le w^*_{(B)}$$

4，bootstrap置信区间-bootstrap-t法

记$k_1 = \left[B \times \frac{\alpha}{2}\right], k_2 = \left[B \times \frac{1-\alpha}{2}\right]$, 作为分位数的估计

$$P\left\{\bar{X} - w^*_{(k_2)}\frac{S}{\sqrt{n}} < \phi < \bar{X} - w^*_{(k_1))}\frac{S}{\sqrt{n}}\right\} = 1 - \alpha$$

得到ϕ的置信水平为$1-\alpha$的bootstrap置信区间。

$$\left(\bar{X} - w_{(k_2)}\frac{S}{\sqrt{n}}, \bar{X} - w_{(k_1)}\frac{S}{\sqrt{n}}\right)$$

4，bootstrap置信区间-bootstrap-t法

例5 有30窝仔猪出生时各窝猪的存活只数为

9	8	10	12	11	12	7	9	11	8	9	7	7	8	9	7
9	9	10	9	9	9	12	10	10	9	13	11	13	9		

用bootstrap-t法求均值μ的置信水平为0.90的置信区间

4，bootstrap置信区间-bootstrap-t法

解：在原始样本中

n=30, $\bar{x} = 9.53$,s=1.72,$s^2 = 2.95$.

对于第i个(i=1,2,…,10000=B)bootstrap样本,

求出它的均值$\bar{x}_i^\star$和样本标准差$S_i^\star$,

从而得到$w^\star$的第i个值:

$$w_i^\star = \frac{\bar{x}_i^\star - \bar{x}}{s_i^\star/\sqrt{n}}, i = 1,2,\ldots,10000$$

其中$\bar{x}$是由原始样本确定的样本均值.

4，bootstrap置信区间-bootstrap-t法

将$w_i^\star$自小到大排序得到

$$w_{(1)}^\star \le w_{(2)}^\star \le \cdots \le w_{(10000)}^\star$$

取置信水平1-α=0.90,此时α=0.10,α/2=0.05,1-α/2=0.95,

取$k_1 = \left[B \times \frac{\alpha}{2}\right] = 500, k_2 = \left[B \times \left(1 - \frac{\alpha}{2}\right)\right] = 9500,$

得

$$w_{(500)}^\star = -1.7813, w_{(9500)}^\star = 1.62999$$

于是得到u的置信水平为0.90的bootstrap-t置信区间为

$$\left(9.53 - 1.6299 \times \frac{1.72}{\sqrt{30}}, 9.53 + 1.7813 \times \frac{1.72}{\sqrt{30}}\right) = (9.0182, 10.0894)$$

3.7.3 参数bootstrap方法

假设所研究的总体的分布函数F(x;β)的形式已知,但其中包含未知参数β(β可以是向量).

现在已知有一个来自F(x;β)的样本$X_1, X_2, \dots, X_n$

利用这一样本求出β的最大似然估计$\hat{\beta}$.

在F(x;β)中以代替$\hat{\beta}$得到F(x;$\hat{\beta}$),

接着在F(x;$\hat{\beta}$)中产生容量为n的样本

$$X_1^\star, X_2^\star, \dots, X_n^\star \sim F\left(x; \hat{\beta}\right)$$

这种样本可以产生很多个,例如产生B个(B≥1000),就可以利用这些样本对总体进行统计推断,其做法与非参数bootstrap方法一样.

这种方法称为**参数bootstrap法**.

3.7.3 参数bootstrap方法

例 据Hardy-Weinberg定律,若基因频率处于平衡状态,则在一总体中个体具有血型M、MN、N的概率分别是$(1-\theta)^2$,2θ(1-θ),θ^2,其中0<θ<1.

据1937年对某地区的调查有以下的数据:

血型	M	MN	N	
人数	342	500	187	共 1 029

(1)求θ的最大似然估计;

(2)求θ的置信水平为0.90的bootstrap置信区间

3.7.3 参数bootstrap方法

解

分别记x_1,x_2,x_3为具有血型为M,MN,N的人数,记$x_1+x_2+x_3=n$.似然函数为

$$L=[(1-\theta)^2]^{x_1}[2\theta(1-\theta)]^{x_2}[\theta^2]^{x_3}$$
$$=2^{x_2}\theta^{x_2+2x_3}(1-\theta)^{2x_1+x_2}$$

$$lnL=x_2ln2+(x_2+2x_3)ln\theta+(2x_1+x_2)ln(1-\theta)$$

$$\frac{d}{d\theta}lnL=\frac{x_2+2x_3}{\theta}+\frac{-(2x_1+x_2)}{1-\theta}=0$$

$$\hat{\theta}=\frac{x_2+2x_3}{2x_1+2x_2+2x_3}=\frac{x_2+2x_3}{2n}$$

3.7.3 参数bootstrap方法

以数据x1=342,x2=500,x3=187,n=1029,代入得到$\hat{\theta}$=0.4247.

以$\hat{\theta}$代替θ,得到$(1-\theta)^2$=0.331, $2\theta(1-\theta)$=0.489, θ^2=0.180.

于是血型的近似分布律为

血型	M	MN	N
概率	0.331	0.489	0.180

以此分布律产生10000个bootstrap样本, 从而得到θ的10000个bootstrap 估计

$\hat{\theta}_1^\star, \hat{\theta}_2^\star, \ldots, \hat{\theta}_{10000}^\star$.

将这10000个数按自小到大的次序排序得到

$$\hat{\theta}_{(1)}^\star \le \hat{\theta}_{(2)}^\star \le \cdots \le \hat{\theta}_{(500)}^\star \le \cdots \le \hat{\theta}_{(9500)}^\star \le \cdots \le \hat{\theta}_{(10000)}^\star.$$

取$(\hat{\theta}_{(500)}^\star, \hat{\theta}_{(9500)}^\star) = (1.83, 1.92)$为$\theta$的置信水平为0.90的bootstrap 置信区间.

3.8 MCMC

3.8.1马氏链和MCMC介绍
3.8.2 METROPOLIS-HASTING抽样
3.8.3 GIBBS抽样

3.8.1马氏链和MCMC介绍

例子：设状态转移矩阵

$$P = \begin{bmatrix} 0.65 & 0.28 & 0.07 \\ 0.15 & 0.67 & 0.18 \\ 0.12 & 0.36 & 0.52 \end{bmatrix}$$

假设初始概率分布为$\pi_0 = [0.21, 0.68, 0.11]$

0 0.210 0.680 0.110

1 0.252 0.554 0.194

2 0.270 0.512 0.218

3 0.278 0.497 0.225

4 0.282 0.490 0.226

5 0.285 0.489 0.225

6 0.286 0.489 0.225

7 0.286 0.489 0.225

8 0.289 0.488 0.225

9 0.286 0.489 0.225

10 0.286 0.489 0.225

3.8.1马氏链和MCMC介绍

初始概率分布 $\pi_0 = [0.75,0.15,0.1]$

0 0.75 0.15 0.1

1 0.522 0.347 0.132

2 0.407 0.426 0.167

3 0.349 0.459 0.192

4 0.318 0.475 0.207

5 0.303 0.482 0.215

6 0.295 0.485 0.220

7 0.291 0.487 0.222

8 0.289 0.488 0.225

9 0.286 0.489 0.225

10 0.286 0.489 0.225

3.8.1马氏链和MCMC介绍

计算 P^n

$$P^{20} = P^{21} = \cdots = P^{100}$$

$$= \cdots$$

$$= \begin{bmatrix} 0.286 & 0.489 & 0.225 \\ 0.286 & 0.489 & 0.225 \\ 0.286 & 0.489 & 0.225 \end{bmatrix}$$

3.8.1马氏链和MCMC介绍

- 马尔科夫链的收敛性质

如果一个非周期的马尔科夫链有状态转移矩阵P, 并且它的任何两个状态是连通的，那么 $\lim_{n\to\infty} P_{ij}^{n}$ 与i无关，

1） $\lim_{n\to\infty} P_{ij}^{n} = \pi(j)$

2)

$$\lim_{n\to\infty} P^{n} = \begin{pmatrix} \pi(1) & \pi(2) & \dots & \pi(j) & \dots \\ \pi(1) & \pi(2) & \dots & \pi(j) & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \pi(1) & \pi(2) & \dots & \pi(j) & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

3） $\pi(j) = \sum_{i=0}^{\infty} \pi(i) P_{ij}$

4）π是方程$\pi P = \pi$的唯一非负解，其中：$\pi = [\pi(1), \pi(2), \dots, \pi(j), \dots]$ $\sum_{i=0}^{\infty} \pi(i) = 1$

3.8.1马氏链和MCMC介绍

- 设正常返的不可约马氏链的平稳分布为π,
- 设h(·)是状态空间S上的有界函数， 则

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n}h\left(X_k\right)=\sum_{x\in S}\pi_x\,h(x)\right)=1 \qquad (19.3)$$

3.8.1马氏链和MCMC介绍

- 非周期正常返的不可约马氏链存在极限分布，极限分布就是平稳分布：

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j,\ \forall i,j \in S$$

- 如果存在$\{\pi_j, j \in S\}, \pi_j \geq 0, \sum_{j\in S}\pi_j = 1$，使得

$$\pi_i p_{ij} = \pi_j p_{ji},\ \forall i \neq j,$$

称这样的马氏链为**细致平稳的**(detailed balance)

- $\{\pi_j\}$是$\{X_t\}$的平稳分布

3.8.1马氏链和MCMC介绍

- 目标平稳分布π(x)和某一个马尔科夫链状态转移矩阵Q不满足细致平稳条件

$$\pi(i)Q(i,j) \neq \pi(j)Q(j,i)$$

- 引入一个α(i,j),使得

$$\pi(i)Q(i,j)\alpha(i,j) = \pi(j)Q(j,i)\alpha(j,i)$$

只要α(i,j)满足

$$\alpha(i,j) = \pi(j)Q(j,i)$$
$$\alpha(j,i) = \pi(i)Q(i,j)$$

3.8.1马氏链和MCMC介绍

- 令P为平稳分布π(x)对应的状态转移矩阵，取

$$P(i,j) = Q(i,j)\alpha(i,j)$$

- α(i,j) 称为**接受率**
- 取值在[0,1]之间，
- 可以理解为一个概率值

3.8.1马氏链和MCMC介绍

- MCMC抽样步骤：

1）输入我们任意选定的状态转移矩阵Q，平稳分布π(x)，设定状态转移次数阈值n_1，需要的样本个数n_2

2）从任意简单概率分布抽样得到初始状态值x_0

3）for t = 0 to n_1 +n_2-1:

a) 从条件概率分布$Q(x|x_t)$中抽样得到样本x_*

b) 从均匀分布抽样$u \sim U[0,1]$

c) 如果$u < \alpha(x_t, x_*) = \pi(x_*)Q(x_*, x_t)$, 则接受转移$x_t \to x_*$，即$x_{t+1} = x_*$

d) 否则不接受转移，即$x_{t+1} = x_t$

样本集$(x_{n_1}, x_{n_1+1}, \ldots, x_{n_1+n_2-1})$即为我们需要的平稳分布对应的样本集

3.8.2 Metropolis-Hasting抽样

- 这种MCMC抽样，有时接受率过低
- 接受率做如下改进

$$\alpha(i,j) = min\left\{\frac{\pi(j)Q(j,i)}{\pi(i)Q(i,j)}, 1\right\}$$

即得 Metropolis-Hasting抽样

3.8.2 Metropolis-Hasting抽样

- **Metropolis-Hasting抽样步骤：**

1）输入任意选定的状态转移矩阵Q，平稳分布π(x)，设定状态转移次数阈值n_1，需要的样本个数n_2

2）从任意简单概率分布采样得到初始状态值x_0

3）for t = 0 to $n_1 + n_2 - 1$:

a) 从条件概率分布$Q(x|x_t)$中采样得到样本x_*

b) 从均匀分布采样$u \sim uniform[0,1]$

c) 如果$u < \alpha(x_t, x_*) = min\{\frac{\pi(j)Q(j,i)}{\pi(i)Q(i,j)}, 1\}$, 则接受转移$x_t \rightarrow x_*$，即$x_{t+1} = x_*$

d) 否则不接受转移，即$x_{t+1} = x_t$

样本集$(x_{n_1}, x_{n_1+1}, \ldots, x_{n_1+n_2-1})$即为平稳分布对应样本集

3.8.2 Metropolis-Hasting抽样

- 如果转移矩阵Q是对称的，即满足Q(i,j) = Q(j,i),
- 这时我们的接受率可以进一步简化为：

$$\alpha(i,j) = min\{\frac{\pi(j)}{\pi(i)}, 1\}$$

- 相应的算法称为**Metropolis抽样法**

- https://github.com/chi-feng/mcmc-demo

3.8.2 Metropolis-Hasting抽样

连续型分布的随机游动MH算法

连续型的目标分布:

设$\pi(x)$为目标分布的密度,

$T(y|x)$为给定x条件下的试抽样密度

这个试转移概率函数较难找到，用随机游动Metropolis抽样：

设X的目标分布$\pi(\boldsymbol{x})$取值于欧式空间$\mathcal{X}=\mathbb{R}^d$。

从$\boldsymbol{x}^{(t)}$出发试转移，令

$$\boldsymbol{y}=\boldsymbol{x}^{(t)}+\boldsymbol{\varepsilon}_t$$

其中$\boldsymbol{\varepsilon}_t \sim g(\boldsymbol{x};\sigma)$对不同$t$是独立同分布的,

试转移概率函数$T(\boldsymbol{y}|\boldsymbol{x})=g(\boldsymbol{y}-\boldsymbol{x})$。

设g是关于$\boldsymbol{x}=\mathbf{0}$对称的分布， 则$T(\boldsymbol{y}|\boldsymbol{x})=T(\boldsymbol{x}|\boldsymbol{y})$。

常取g为$N(\mathbf{0},\sigma^2 I)$或半径为σ的中心为$\mathbf{0}$的球内的均匀分布。

3.8.2 Metropolis-Hasting抽样

随机游动MH算法

转移法则为:

从$\boldsymbol{x}^{(t)}$出发试转移到$\boldsymbol{y}$后,

若$\pi(\boldsymbol{y}) > \pi(\boldsymbol{x}^{(t)})$

则令$\boldsymbol{x}^{(t+1)} = \boldsymbol{y}$;

否则,

独立地抽取$U \sim \mathrm{U}(0,1)$, 取

$$\boldsymbol{x}^{(t+1)} = \begin{cases} \boldsymbol{y}, & \text{当 } U \leq \pi(\boldsymbol{y})/\pi(\boldsymbol{x}^{(t)}), \\ \boldsymbol{x}^{(t)}, & \text{其它.} \end{cases}$$

3.8.2 Metropolis-Hasting抽样

例19.3 考虑如下的简单气体模型

在平面区域$G = [0, A] \times [0, B]$内有K个直径为 d的刚性圆盘

随机向量$\mathbf{X} = (x_1, y_1, \ldots, x_K, y_K)$为这些圆盘的位置坐标。

分布$\pi(\mathbf{x})$是G内所有允许位置的均匀分布。

希望**对π抽样**。

3.8.3 Gibbs抽样

- 对二维的数据分布

假设$\pi(x_1, x_2)$是一个二维联合数据分布，

观察第一个特征维度相同的两个点

$A\left(x_1^{(1)}, x_2^{(1)}\right)$和$B\left(x_1^{(1)}, x_2^{(2)}\right)$，

容易发现下面两式成立：

$$\pi\left(x_1^{(1)}, x_2^{(1)}\right)\pi\left(x_2^{(2)}|x_1^{(1)}\right) = \pi\left(x_1^{(1)}\right)\pi\left(x_2^{(1)}|x_1^{(1)}\right)\pi\left(x_2^{(2)}|x_1^{(1)}\right)$$

$$\pi\left(x_1^{(1)}, x_2^{(2)}\right)\pi\left(x_2^{(1)}|x_1^{(1)}\right) = \pi\left(x_1^{(1)}\right)\pi\left(x_2^{(2)}|x_1^{(1)}\right)\pi\left(x_2^{(1)}|x_1^{(1)}\right)$$

3.8.3 Gibbs抽样

由于两式的右边相等，因此：

$$\pi\left(x_1^{(1)}, x_2^{(1)}\right)\pi\left(x_2^{(2)}|x_1^{(1)}\right) = \pi\left(x_1^{(1)}, x_2^{(2)}\right)\pi\left(x_2^{(1)}|x_1^{(1)}\right)$$

也就是：

$$\pi(A)\pi\left(x_2^{(2)}|x_1^{(1)}\right) = \pi(B)\pi\left(x_2^{(1)}|x_1^{(1)}\right)$$

- 同理，对点$A\left(x_1^{(1)}, x_2^{(1)}\right)$和点$\mathrm{C}\left(x_1^{(2)}, x_2^{(1)}\right)$，有

$$\pi(A)\pi\left(x_1^{(2)}|x_2^{(1)}\right) = \pi(\mathrm{C})\pi\left(x_1^{(1)}|x_2^{(1)}\right)$$

3.8.3 Gibbs抽样

在$x_2 = x_2^{(1)}$这条直线上，如果用条件概率分布$\pi\left(x_1|x_2^{(1)}\right)$作为状态转移概率，则任意两个点之间的转移也满足**细致平稳条件**

在$x_1 = x_1^{(1)}$这条直线上，如果用条件概率分布$\pi\left(x_2|x_1^{(1)}\right)$作为状态转移概率，则任意两个点之间的转移满足**细致平稳条件**

3.8.3 Gibbs抽样

这样构造分布$\pi(x_1, x_2)$的状态转移矩阵P：

$$P(A \to B) = \pi\left(x_2^{(B)} | x_1^{(1)}\right) \quad if \ x_1^{(A)} = x_1^{(B)} = x_1^{(1)}$$

$$P(A \to C) = \pi\left(x_1^{(C)} | x_2^{(1)}\right) \quad if \ x_2^{(A)} = x_2^{(C)} = x_2^{(1)}$$

$$P(A \to D) = 0 \quad else$$

3.8.3 Gibbs抽样

- 二维Gibbs抽样步骤：

1）输入平稳分布$\pi(x_1, x_2)$，设定状态转移次数阈值n_1，需要的样本个数n_2

2）随机初始化初始状态值$x_1^{(0)}$和$x_2^{(0)}$

3）for t = 0 to $n_1 + n_2$-1:

a) 从条件概率分布$P\left(x_1 | x_2^{(t)}\right)$中抽样得到样本$x_1^{t+1}$

b) 从条件概率分布$P\left(x_2 | x_1^{(t+1)}\right)$中抽样得到样本$x_2^{t+1}$

样本集$\{\left(x_1^{(n_1)}, x_2^{(n_1)}\right), \left(x_1^{(n_1+1)}, x_2^{(n_1+1)}\right), \ldots, \left(x_1^{(n_1+n_2-1)}, x_2^{(n_1+n_2-1)}\right)\}$即为平稳分布对应的样本集。

3.8.3 Gibbs抽样

- 例

一个二维正态分布$N(\mu, \Sigma)$,其中：

$$\mu = (\mu_1, \mu_2) = (5, -1)$$

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$$

3.8.3 Gibbs抽样

状态转移条件分布为：

$$P(x_1|x_2) = N(\mu_1 + \rho\sigma_1\sigma_2(x_2 - \mu_2), (1-\rho^2)\sigma_1^2)$$

$$P(x_2|x_1) = N(\mu_2 + \rho\sigma_2\sigma_1(x_1 - \mu_1), (1-\rho^2)\sigma_2^2)$$

3.8.3 Gibbs抽样

```
    from mpl_toolkits.mplot3d import Axes3D
    from scipy.stats import multivariate_normal
    samplesource = multivariate_normal(mean=[5,-1], cov=[[1,1],[1,4]])
    def p_ygivenx(x, m1, m2, s1, s2):
        return (random.normalvariate(m2 + rho * s2 / s1 * (x - m1), math.sqrt((1 - rho ** 2) *
(s2**2))))
    def p_xgiveny(y, m1, m2, s1, s2):
        return (random.normalvariate(m1 + rho * s1 / s2 * (y - m2), math.sqrt((1 - rho ** 2) *
(s1**2))))
    N = 5000
    K = 20
    x_res = []
    y_res = []
    z_res = []
    m1 = 5
    m2 = -1
    s1 = 1
    s2 = 2
    rho = 0.5
    y = m2
```

3.8.3 Gibbs抽样

```
for i in xrange(N):
    for j in xrange(K):
        x = p_xgiveny(y, m1, m2, s1, s2)
        y = p_ygivenx(x, m1, m2, s1, s2)
        z = samplesource.pdf([x,y])
        x_res.append(x)
        y_res.append(y)
        z_res.append(z)
num_bins = 50
plt.hist(x_res, num_bins, normed=1, facecolor='green', alpha=0.5)
plt.hist(y_res, num_bins, normed=1, facecolor='red', alpha=0.5)
plt.title('Histogram')
plt.show()
```

例19.5

设目标分布为

$$\pi(x,y) \propto \binom{n}{x} y^{x+\alpha-1}(1-y)^{n-x+\beta-1},$$
$$x = 0,1,\dots,n,\ 0 \le y \le 1,$$

则$X|Y \sim B(n,y), Y|X \sim Beta(x+\alpha, n-x+\beta)$。

易见Y的边缘分布为$Beta(\alpha,\beta)$。

用**Gibbs抽样**方法模拟生成(X,Y)的样本链。

3.8.3 Gibbs抽样

- 对多维的数据分布

1）输入平稳分布$\pi(x_1, x_2, \ldots, x_n)$或者对应的所有特征的条件概率分布，设定状态转移次数阈值n_1，需要的样本个数n_2

2）随机初始化初始状态值$(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})$

3）for t = 0 to $n_1 + n_2$-1:

a) 从条件概率分布$P\left(x_1 | x_2^{(t)}, x_3^{(t)}, \ldots, x_n^{(t)}\right)$中抽样得到样本$x_1^{t+1}$

b) 从条件概率分布$P\left(x_2 | x_1^{(t+1)}, x_3^{(t)}, x_4^{(t)}, \ldots, x_n^{(t)}\right)$中抽样得到样本$x_2^{t+1}$

c)...

d) 从条件概率分布$P\left(x_j | x_1^{(t+1)}, x_2^{(t+1)}, \ldots, x_{j-1}^{(t+1)}, x_{j+1}^{(t)} \ldots, x_n^{(t)}\right)$中抽样得到样本$x_j^{t+1}$

e)...

f) 从条件概率分布$P\left(x_n | x_1^{(t+1)}, x_2^{(t+1)}, \ldots, x_{n-1}^{(t+1)}\right)$中抽样得到样本$x_n^{t+1}$

样本集$\{\left(x_1^{(n_1)}, x_2^{(n_1)}, \ldots, x_n^{(n_1)}\right), \ldots, \left(x_1^{(n_1+n_2-1)}, x_2^{(n_1+n_2-1)}, \ldots, x_n^{(n_1+n_2-1)}\right)\}$即为平稳分布对应的样本集。

3.8.3 Gibbs抽样

- **系统扫描Gibbs抽样算法:**

从$\pi(\mathbf{x})$的取值区域任意取一个初值$\mathbf{X}^{(0)}$

$\text{for}\left(t \text{ in } 0:(N-1)\right)\{$

$\text{for}(i \text{ in } 1:n)\{$

从条件分布$p\left(x_i|X_1^*,\ldots,X_{i-1}^*,X_{i+1}^{(t)},\ldots,X_n^{(t)}\right)$抽取$X_i^*$

$\}$

$\mathbf{X}^{(t+1)} \leftarrow (X_1^*,\ldots,X_n^*)$

$\}$

3.8.3 Gibbs抽样

- **随机扫描Gibbs抽样算法:**。

从$\pi(\mathbf{x})$的取值区域任意取一个初值$\mathbf{X}^{(0)}$

$\text{for}\left(t \text{ in } 0:(N-1)\right)\{$

按概率$\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$随机抽取下标$i$

从条件分布$p\left(x_i|\mathbf{X}_{(-i)}^{(t)}\right)$抽取$X_i^*$

$\mathbf{X}^{(t+1)} \leftarrow \left(X_1^{(t)}, \ldots, X_{i-1}^{(t)}, X_i^*, X_{i+1}^{(t)}, \ldots, X_n^{(t)}\right)$

$\}$

其中下标的抽样概率$\boldsymbol{\alpha}$为事先给定