卓越大学出版联盟
学术出版精品工程

实用统计技术丛书

马尔科夫过程与实用随机模型

MARKOV PROCESSES AND APPLIED STOCHASTIC MODELS

李元章　何春雄　著

華南理工大學出版社
SOUTH CHINA UNIVERSITY OF TECHNOLOGY PRESS

实用统计技术丛书

马尔科夫过程与实用随机模型

李元章　何春雄　著

华南理工大学出版社
SOUTH CHINA UNIVERSITY OF TECHNOLOGY PRESS
·广州·

内 容 提 要

本书介绍针对随时间变化的随机现象的数学建模方法，主要介绍用随机过程的基本原理对实际问题建立数学模型，并借助随机模拟和SAS软件对模型求解的基本原理和方法.

本书适于欲了解概率论与数理统计及随机过程基础知识和具有使用计算机软件的基本经验的读者阅读. 可作为概率论与数理统计专业的研究生或数学专业高年级本科生的选修课教材，也可供用统计方法从事社会科学、自然科学研究的人员参考.

图书在版编目(CIP)数据

马尔科夫过程与实用随机模型/李元章，何春雄著. —广州：华南理工大学出版社，2018.4
(实用统计技术丛书)
ISBN 978-7-5623-5007-1

Ⅰ. ①马… Ⅱ. ①李… ②何… Ⅲ. ①马尔科夫过程 Ⅳ. ①O211.62

中国版本图书馆CIP数据核字(2017)第256589号

马尔科夫过程与实用随机模型
李元章 何春雄 著

出 版 人：卢家明
出版发行：华南理工大学出版社
(广州五山华南理工大学17号楼，邮编510640)
http://www.scutpress.com.cn E-mail:scutc13@scut.edu.cn
营销部电话：020—87113487 87111048(传真)
策划编辑：詹志青
责任编辑：詹志青
印 刷 者：虎彩印艺股份有限公司
开 本：787 mm×1092 mm 1/16 **印张**：10.5 **字数**：263千
版 次：2018年4月第1版 2018年4月第1次印刷
印 数：1～1000册
定 价：30.00元

前　言

随时间变化的随机现象普遍存在. 随机过程论研究这种现象的统计规律，包括有限维分布、数字特征和发展趋势等等. 对实际应用中遇到的随机现象建立数学模型，就是将其中随时间变化的变量视为随机过程，通过分析随机过程之间的关系和统计规律，建立数学模型并对模型求解，为相应的实际问题的解决或决策提供依据. 本书主要介绍用随机过程的基本原理对实际问题建立数学模型，并借助随机模拟和 SAS 软件对模型求解的基本原理和方法. 阅读本书需要读者具有概率论与数理统计及随机过程的基础知识，并有运用计算机软件的基本经验.

本书内容分为 9 章. 第 1 章简要叙述概率论与数理统计的基本概念和方法，并在简单介绍随机过程基本概念的基础上，总结泊松(Poisson)过程和马尔科夫(Markov)过程的基本性质. 第 2 章介绍生成随机数的基本原理和方法. 第 3 章介绍生成某种具体分布随机数的方法. 第 4 章介绍排队论模型，它是一种重要而常见的随机服务系统. 第 5 章介绍库存理论，主要介绍生产库存模型和流通库存模型. 第 6 章介绍排队网络模型，主要介绍开放 Jackson 网络系统和封闭 Jackson 网络系统，并简要介绍非 Jackson 网络系统. 第 7 章介绍更新与维修，主要介绍老化更换、维修保养时间的选择和成批更新优化. 第 8 章介绍非经典排队论模型，主要讨论批量到达的服务系统和相型分布模型. 第 9 章介绍马尔科夫决策过程，包括该类过程的定义以及在各种优化准则下优化策略的具体算法.

李元章教授曾在美国某研究院、兰州大学、华南理工大学等多地讲授过本书的内容，在实际问题的研究中用到本书中涉及的多种模型，并作为华南理工大学的客座教授，与何春雄教授合作，在华南理工大学举办实用统计技术系列讲座. 本书内容是在讲座“实用随机模型”基础上加工整理而成的. 在此我们特别感谢华南理工大学数学学院和研究生院的大力支持，并衷心感谢华南理工大学出版社的精诚合作. 由于作者水平有限，书中难免会有疏漏或不当之处，恳请同行和读者指正.

著　者

2018 年 4 月

目　录

1　预备知识 …… 1
1.1　概率论的基本概念 …… 1
1.2　数理统计的基本概念 …… 7
1.3　泊松过程及性质 …… 14
1.4　马尔科夫链及性质 …… 18
2　蒙特-卡罗数字模拟方法 …… 24
2.1　蒙特-卡罗方法的基本概念——人工模拟 …… 25
2.2　随机数与伪随机数 …… 28
2.3　乘法同余随机数生成器 …… 29
2.4　循环同余随机数生成器 …… 35
2.5　复合随机数生成器 …… 36
2.6　随机数的检验 …… 38
习题 2 …… 42
3　生成随机变元 …… 44
3.1　反函数方法 …… 44
3.2　生成离散分布随机数 …… 50
3.3　生成正态分布随机数 …… 53
习题 3 …… 54
4　排队论模型 …… 56
4.1　排队论模型的基本要素 …… 56
4.2　单一服务器系统 …… 58
4.3　排队系统稳定性 …… 61
4.4　有限容量排队系统 …… 66
4.5　多个服务器的服务系统 …… 69
4.6　系统的近似估计 …… 72
4.7　排队系统模拟 …… 74
习题 4 …… 80
5　库存理论 …… 83
5.1　小贩问题 …… 83
5.2　周期进货问题 …… 86
5.3　起始价问题 …… 88
5.4　多周期进货问题 …… 89
习题 5 …… 93

6 排队网络系统 …… 96
6.1 开放 Jackson 网络系统 …… 96
6.2 封闭 Jackson 网络系统 …… 101
6.3 非 Jackson 网络系统 …… 104
习题 6 …… 115
7 更新与维修 …… 117
7.1 老化更换 …… 117
7.2 维修保养时间选择 …… 123
7.3 成批更新优化 …… 126
习题 7 …… 128
8 非经典排队论模型 …… 130
8.1 斐波纳契数列与差分方程 …… 130
8.2 批量到达的服务系统 …… 131
8.3 相型分布模型 …… 136
习题 8 …… 140
9 马尔科夫决策过程 …… 141
9.1 马尔科夫决策过程的定义 …… 141
9.2 稳定性策略 …… 145
9.3 折扣期望平均算法 …… 145
9.4 逐步优化策略折扣运算 …… 147
9.5 折扣准则的线性算法 …… 149
9.6 稳定平均指标算法 …… 151
9.7 策略选择平均指标算法 …… 152
9.8 平均指标的线性算法 …… 154
9.9 优化自动终止程序 …… 157
习题 9 …… 159
参考文献 …… 161

1 预备知识

随机模型是指模型中的有些因素是随机变化的. 随机建模就是基于这些随机因素随着时间推移的变化规律来估计模型解的统计特性或概率分布以及有关数字特征,从而为有关实际问题的解决或决策提供依据. 随机变化通常与时间有关,模型中的随机变元是一个时间序列或随机过程,所以随机模型的建模过程通常是在对历史数据分析的基础上进行模型设计或模拟计算. 随机模型的应用,最初起源于物理学(有时也称为蒙特卡罗方法),现在广泛应用于工业管理、统筹规划、生命科学和社会科学等诸多领域.

由于随机模型中基本的变量为随机变量或随机过程,为方便阅读后续章节,本章先简单回顾概率论与数理统计的有关概念和方法,并简单回顾泊松(Poisson)过程和马尔科夫(Markov)过程的基本性质. 本章内容取材于参考文献[1]和[2].

1.1 概率论的基本概念

本节简单回顾概率空间、随机变量的分布和数字特征等概率论的基本知识,并回顾数理统计的基本概念和方法,包括参数估计和假设检验.

1.1.1 概率空间与随机变量

概率论是研究随机现象的统计规律的科学,其核心问题是回答某种随机现象各个结果(或称随机事件)发生的可能性的大小,具体刻画这种可能性的大小的度量就是一个事件发生的概率,这也是用确定的数学科学研究随机现象的切入点.

为方便地运用数学学科的各种工具,人们想到人为地或自然地将随机试验的结果与数值对应起来,这个对应的结果就称为随机变量. 其含义是,在试验结果出来之前无法预知该变量取何值,也就是说该变量的取值是随机的,进而通过刻画随机变量的分布来描述一个随机事件发生的概率. 因此,概率论是研究随机变量分布的科学. 初等概率论研究的是独立随机变量,所以经典的数理统计大多基于独立随机变量,也就是数理统计中的简单随机样本.

为用确定性的数学研究随机现象这种非确定性现象,概率论在建立数学模型时,引进了三个最基本的概念,它们是随机试验、随机事件和概率.

所谓随机试验是指对随机现象的观测或实验,其基本意义是:① 试验可以在基本相同的条件下大量重复;② 试验会出现的那些结果是已知的;③ 每次试验将出现哪一个结果是无法预知的.

我们称随机试验的结果为随机事件(简称为事件),而事件的概率则是随机事件发生的可能性大小的一种度量. 如何规定或定义这个度量并非易事. 下文将进行简单总结.

例 1.1 掷一枚骰子,观察其落定后朝上的面出现的点数,这就是一个随机试验. 而

$A=\{$出现的点数为奇数$\}$，$B=\{$出现的点数为偶数$\}$，$C=\{$出现的点数小于 3$\}$等都是事件. 若骰子均匀，则 A、B、C 的概率应分别为 $P(A)=\frac{1}{2}$，$P(B)=\frac{1}{2}$，$P(C)=\frac{1}{3}$.

在初等概率论教材中见到的古典概型和几何概型，关于概率的定义都是基于某类特殊随机现象的，不能作为概率的数学定义或公理化定义，从而不能用来作一般的演绎推理. 经过数学家多年艰苦的探索，到 20 世纪 30 年代，以俄国数学家柯尔莫哥洛夫(A. H. Колмогоров)为代表的数学家引入概率的公理化定义，才使概率论建立在坚实的、严密的数学基础之上. 概率的公理化定义的基本思想是把随机事件看作集合，从而事件的和、积、对立及差等运算分别对应集合的并、交、求余和差等集合运算，而把概率定义为集合的测度(即集合大小的度量)，从而把概率论建立在测度论的基础之上，亦即概率论的所有推证或演绎都在概率空间的基础上进行. 在概率的公理化定义中，相应于前述随机试验、随机事件和概率三个直观概念的分别是基本事件空间 Ω(或称样本空间)、事件域 $\mathscr{F}$ 和概率 P.

1.1.1.1 基本事件空间(Ω)及事件的运算

定义 1.1(基本事件空间) 如果每次试验有且仅有 Ω 中的一个事件发生，则称 Ω 为基本事件空间(或样本空间)，并称 Ω 中的元素为基本事件(或样本点).

例 1.1(续 1) 记 $\omega_i=\{$出现的点数为 $i\}$，$i=1,2,\cdots,6$，则 $\Omega_1=\{\omega_1,\omega_2,\cdots,\omega_6\}$ 为一基本事件空间，而 $\Omega_2=\{A,B\}$ 也是基本事件空间.

设 A 和 B 为任意的两个事件，则 A 和 B 的积(记作 $A\cap B$ 或 AB)表示 A 和 B 都出现这样的结果；A 和 B 的和(记作 $A\cup B$)表示 A 和 B 至少出现一个这样的结果；A 和 B 的差(记作 $A\backslash B$ 或 $A-B$)表示 A 出现但 B 不出现这样的结果；A 的对立(记作 $\overline{A}$)表示 A 不出现这样的结果.

显然，$A\backslash B=A\overline{B}$. 另外，若 $AB=\varnothing$(不可能事件)，则称 A 与 B 互不相容；若 A 出现则 B 必然出现，则称 B 包含 A，记作 $A\subset B$.

例 1.1(续 2) 沿用例 1.1 的有关记号，我们知道，$A\cap C=\{$出现点数为 1$\}$，$A\cup C=\{$出现的点数为 1,2,3,5$\}$，$A\backslash C=\{$出现的点数为 3,5$\}$，$C\backslash A=\{$出现的点数为 2$\}$，$\overline{A}=B$，$\overline{B}=A$.

从例 1.1(续 2)可以看出，如果将事件看成集合，$A=\{\omega_1,\omega_3,\omega_5\}$，$B=\{\omega_2,\omega_4,\omega_6\}$ 和 $C=\{\omega_1,\omega_2\}$，那么事件的和、积、差和对立等运算，分别对应于集合的并、交、差和求余等运算.

正如实数经加、减、乘、除等运算之后仍为实数一样，即实数关于这几个运算封闭. 事件经运算后也是事件，即如果把一个试验的结果看成一个集合类，那么该集合类应当对事件运算封闭，这就引出下面关于事件域的定义.

1.1.1.2 事件域($\mathscr{F}$)

定义 1.2(事件域) 设 Ω 是基本事件空间，$\mathscr{F}$ 是由 Ω 的一些子集为元素所组成的集合，如果满足下列条件：

(1) $\Omega\in\mathscr{F}$；

(2) 若 $A\in\mathscr{F}$，则 $\overline{A}\in\mathscr{F}$；

(3) 若 $A_n\in\mathscr{F}$，$n=1,2,\cdots$，则 $\bigcup\limits_{n=1}^{\infty}A_n\in\mathscr{F}$.

则称 $\mathscr{F}$ 为事件域，$\mathscr{F}$ 中的元素称为事件，称 Ω 为必然事件、$\varnothing$ 为不可能事件.

事件域之所以这样定义，一方面它抽象出了一个由事件构成的集合关于事件运算封闭所需的条件，另一方面也为灵活运用概率论研究各种随机现象提供一个基本框架.

例 1.1(续 3) 沿用例 1.1(续 1)的记号，则 $\mathscr{F}_1=\{A\mid A\subset\Omega_1\}$(通常记作 2^{Ω_1})和 $\mathscr{F}_2=\{\Omega_2,\varnothing,A,B\}$以及 $\mathscr{F}_3=\{\Omega_1,\varnothing,A,B\}$都为事件域.

1.1.1.3 概率(P)及其性质

定义 1.3(概率) 设 Ω 是随机试验的基本事件空间，P 为定义在事件域 $\mathscr{F}$ 到实数集 $\mathbf{R}$ 的映射，满足：

(1)(非负性)对任一事件 A，有 $0\leqslant P(A)\leqslant 1$；

(2)(规范性)$P(\Omega)=1$；

(3)(可列可加性)若 $A_1,A_2,\cdots$ 互不相容，则

$$P\Big(\bigcup_{i=1}^{\infty}A_i\Big)=\sum_{i=1}^{\infty}P(A_i).$$

则称 P 为事件域 $\mathscr{F}$ 上的概率，而称 $P(A)$ 为事件 A 的概率.

概率的这种定义是长度、面积和体积等几何度量的抽象，也是我们日常生活中称重和量长度的一种抽象. 称$(\Omega,\mathscr{F},P)$为概率空间. 为方便应用，下面简单讨论概率测度的性质，其证明可在一般概率书中找到.

概率的性质：

(1)$P(\varnothing)=0$；

(2)(有限可加性) 若 $A_1,A_2,\cdots,A_n$ 互不相容，则

$$P\Big(\bigcup_{i=1}^{n}A_i\Big)=\sum_{i=1}^{n}P(A_i);$$

(3)$P(\bar{A})=1-P(A)$；

(4)(单调性和可减性) 若 $A\subset B$，则有

$$P(A)\leqslant P(B),\ \text{且}\ P(B\backslash A)=P(B)-P(A);$$

(5)加法公式：

$$P(A\cup B)=P(A)+P(B)-P(AB).$$

1.1.2 随机变量及其分布与数字特征

我们知道，概率测度 P 是定义在事件域 $\mathscr{F}$ 到实数集 $\mathbf{R}$ 的映射，它不是经典的函数. 为有效地应用分析数学工具来分析和研究随机现象，人们想到把基本事件变换成数，从而把事件的概率用函数值来表达，这就是随机变量.

1.1.2.1 随机变量及其分布函数

定义 1.4(随机变量与分布函数) 设$(\Omega,\mathscr{F},P)$为概率空间，记 $\mathbf{R}$ 为实数集. 如果对任意 $x\in\mathbf{R}$，有

$$\{\omega\mid\xi(\omega)\leqslant x\}\in\mathscr{F},$$

则称映射 $\xi:\Omega\to\mathbf{R}$ 为随机变量.

随机变量 ξ 的分布函数定义为

$$F_\xi(x)=P(\{\omega\in\Omega\mid\xi(\omega)\leqslant x\}),\forall\, x\in\mathbf{R}.$$

由定义 1.4，显然有

$$P(a<\xi\leqslant b)=F_\xi(b)-F_\xi(a),\forall a<b\in\mathbf{R}.$$

引入随机变量的分布函数后，我们所关心的有关事件的概率都可以用其分布函数来表达.

描述实际随机现象时，有两类重要的随机变量：一类是至多取可数多个不同的值，称之为离散型随机变量；另一类是连续取值的，它的分布函数可以表示为 $F_\xi(x)=\int_{-\infty}^{x} f_\xi(t)\mathrm{d}t$，称之为连续型随机变量，而称 f_ξ 为 ξ 的分布密度函数.

现实中在刻画一个随机现象时，随机试验的结果往往需要维数高于一维的数组来记录或表达，这就需要引入随机向量. 数理统计中的样本就是随机向量.

定义 1.5(随机向量与联合分布函数) 设$(\Omega,\mathscr{F},P)$为概率空间，记 $\mathbf{R}^n$ 为 n 维实数空间. 若 ξ_i 为随机变量，$i=1,2,\cdots,n$，则称向量$(\xi_1,\xi_2,\cdots,\xi_n)$为随机向量. 其联合分布函数定义为

$$F_{\xi_1,\xi_2,\cdots,\xi_n}(x_1,x_2,\cdots,x_n)=P\Big(\bigcap_{i=1}^{n}\{\omega\mid\xi_i(\omega)\leqslant x_i\}\Big),\quad\forall(x_1,x_2,\cdots,x_n)\in\mathbf{R}^n.$$

与前述随机变量平行地，若分布函数可以表示为

$$F_{\xi_1,\xi_2,\cdots,\xi_n}(x_1,x_2,\cdots,x_n)=\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n} f_{\xi_1,\xi_2,\cdots,\xi_n}(t_1,t_2,\cdots,t_n)\mathrm{d}t_1\mathrm{d}t_2\cdots\mathrm{d}t_n,$$

则称$(\xi_1,\xi_2,\cdots,\xi_n)$为连续型随机向量，而称 $f_{\xi_1,\xi_2,\cdots,\xi_n}$ 为其分布密度函数. 若$(\xi_1,\xi_2,\cdots,\xi_n)$的每个分量都至多取可数多个不同的值，则称之为离散型随机向量.

随机变量(向量)除了可用其分布函数完整刻画外，还可用其数字特征来刻画其某些方面的特性. 下面简单回顾一下.

1.1.2.2 随机变量的数字特征

数学期望是刻画随机变量取值的平均程度的一个数字特征.

我们知道，一个随机变量 ξ 取何值是无法预知的，它的取值因各次试验出现的结果而有所不同，但它平均取何值呢?

如果 ξ 为离散型随机变量，其分布列为

ξ	a_1	a_2	$\cdots$	a_n	$\cdots$
P	p_1	p_2	$\cdots$	p_n	$\cdots$

则考虑到 ξ 取值的可能性的大小，该均值应为加权平均，所以此时定义随机变量 ξ 的数学期望 $E[\xi]$为

$$E[\xi]=\sum_{i=1}^{\infty}a_ip_i.$$

由于 ξ 的统计特性由其分布函数 F_ξ 确定，并考虑到 ξ 的平均值应当与其取值 $a_i(i=1,2,\cdots)$的排列顺序无关，而有如下较为抽象的定义.

定义 1.6(数学期望) 设$(\Omega,\mathscr{F},P)$为概率空间，ξ 为随机变量，其分布函数为 F_ξ，如果 $\int_{-\infty}^{\infty}|x|\mathrm{d}F_\xi(x)<\infty$，则称

$$E[\xi]=\int_{-\infty}^{\infty}x\mathrm{d}F_\xi(x)$$

为 ξ 的数学期望(或称均值).

可以证明，若 g 为定义在实直线上的实值函数，ξ 为随机变量，$\eta=g(\xi)$，并且 $\int_{-\infty}^{\infty}|y|\mathrm{d}F_\eta(y)<\infty$，则

$$E[\eta]=\int_{-\infty}^{\infty}g(x)\mathrm{d}F_\xi(x).$$

此时，若 ξ 为离散型随机变量，取值为 $a_i, i=1,2,\cdots$，$p_i=P(\xi=a_i)$，$i=1,2,\cdots$，则

$$E[\xi]=\sum_{i=1}^{\infty}g(a_i)p_i;$$

若 ξ 为连续型随机变量，分布密度函数为 f_ξ，则

$$E[\xi]=\int_{-\infty}^{\infty}g(x)f_\xi(x)\mathrm{d}x.$$

方差是刻画随机变量取值分散程度的又一个重要数字特征，它定义为随机变量与其均值差距平方的平均值，即

$$\mathrm{Var}[\xi]=E[(\xi-E[\xi])^2].$$

另外，设 $m\geqslant 1$ 为整数，若 $E[|\xi|^m]<\infty$，则称 $E[\xi^m]$ 为 ξ 的 m 阶原点矩，而称 $E[(\xi-E[\xi])^m]$ 为 ξ 的 m 阶中心矩.

对于随机向量 $(\xi_1,\xi_2,\cdots,\xi_n)$，我们定义其数学期望(向量)和协方差(矩阵)为

$$E[(\xi_1,\xi_2,\cdots,\xi_n)]=(E[\xi_1],E[\xi_2],\cdots,E[\xi_n]);$$

$$\mathrm{Var}[(\xi_1,\xi_2,\cdots,\xi_n)]=(\mathrm{Cov}(\xi_i,\xi_j))_{1\leqslant i,j\leqslant n}.$$

其中，

$$\mathrm{Cov}(\xi_i,\xi_j)=E[(\xi_i-E[\xi_i])(\xi_j-E[\xi_j])],$$

称为 ξ_i 与 ξ_j 的协方差.

若 $\xi_1,\xi_2,\cdots,\xi_n$ 都有有限的数学期望，则对任意实数 $k_1,k_2,\cdots,k_n$，有

$$E[k_1\xi_1+k_2\xi_2+\cdots+k_n\xi_n]=k_1E[\xi_1]+k_2E[\xi_2]+\cdots+k_nE[\xi_n].$$

1.1.2.3 常见随机变量的分布及其数字特征

为方便读者阅读，将数理统计中常见的随机变量列于表 1-1.

表 1-1 常见随机变量的分布及其数字特征

分布名称	分 布	记号	参 数	均值	方 差
两点分布	$P(\xi=k)=p^k(1-p)^{1-k}$，$k=0,1$	$B(1,p)$	$0<p<1$	p	$p(1-p)$
二项分布	$P(\xi=k)=\binom{n}{k}p^k(1-p)^{n-k}$，$k=0,1,2,\cdots,n$	$B(n,p)$	$n>0,0<p<1$	np	$np(1-p)$
Poisson 分布	$P(\xi=k)=\frac{\mathrm{e}^{-\lambda}\lambda^k}{k!}$，$k=0,1,2,\cdots$	$\mathrm{Poi}(\lambda)$	$\lambda>0$	λ	λ
几何分布	$P(\xi=k)=(1-p)^{k-1}p$，$k=1,2,\cdots$	$\mathrm{Geo}(p)$	$0<p<1$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
负二项分布	$P(\xi=k)=\binom{k-1}{l-1}p^l(1-p)^{k-l}$，$k=l,l+1,\cdots$	$\mathrm{Pas}(l,p)$	$l>0,0<p<1$	$\frac{l}{p}$	$\frac{l(1-p)}{p^2}$

（续表 1-1）

分布名称	分　布	记号	参数	均值	方差
均匀分布	$f_\xi(x)=\begin{cases}\frac{1}{b-a} & \text{当 } x\in(a,b)\\ 0 & \text{其他}\end{cases}$	$U(a,b)$	$a<b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
正态分布	$f_\xi(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$N(\mu,\sigma^2)$	$-\infty<\mu<\infty$, $\sigma>0$	μ	σ^2
指数分布	$f_\xi(x)=\begin{cases}\lambda e^{-\lambda x} & \text{当 } x>0\\ 0 & \text{其他}\end{cases}$	$\mathrm{Exp}(\lambda)$	$\lambda>0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Γ 分布	$f_\xi(x)=\begin{cases}\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x} & \text{当 } x>0\\ 0 & \text{其他}\end{cases}$	$\Gamma(\alpha,\lambda)$	$\alpha>0,\lambda>0$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$
Weibull 分布	$f(x)=\begin{cases}\frac{\beta}{\alpha^\beta}x^{\beta-1}e^{-(\frac{x}{\alpha})^\beta} & \text{当 } x\geqslant 0\\ 0 & \text{当 } x<0\end{cases}$	$\mathrm{Weib}(\alpha,\beta)$	$\alpha>0,\ \beta>0$	$\alpha\Gamma(\frac{1}{\beta}+1)$	$\alpha^2\left[\Gamma(\frac{2}{\beta}+1)-\left(\Gamma(\frac{1}{\beta}+1)\right)^2\right]$

1.1.3　条件概率与独立性

条件概率与独立性是概率论中的两个十分重要的概念，我们在此简单复习.

1.1.3.1　条件概率

定义 1.7（条件概率）　设$(\Omega,\mathscr{F},P)$为概率空间，$A,B\in\mathscr{F}$，且 $P(B)>0$，则定义

$$P(A\mid B)=\frac{P(AB)}{P(B)}$$

为 B 发生的条件下 A 发生的条件概率.

显然，若取定 B 且 $P(B)>0$，则 $P_B(\cdot)=P(\cdot\mid B)$为$(\Omega,\mathscr{F})$上的概率，亦即$(\Omega,\mathscr{F},P_B)$也是概率空间.

命题 1.1（乘法公式）　设$(\Omega,\mathscr{F},P)$为概率空间，$A_i\in\mathscr{F}$，$i=1,2,\cdots,n$，且

$$P(A_1A_2\cdots A_{n-1})>0,$$

则

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2\mid A_1)P(A_3\mid A_1A_2)\cdots P(A_n\mid A_1A_2\cdots A_{n-1}).$$

命题 1.2（全概率公式）　设$(\Omega,\mathscr{F},P)$为概率空间，$B_i\in\mathscr{F}$，$P(B_i)>0$，$i=1,2,\cdots$，$B_i\cap B_j=\varnothing$，$i\neq j$，且 $A\subset\bigcup_{i=1}^{\infty}B_i$，则

$$P(A)=\sum_{i=1}^{\infty}P(B_i)P(A\mid B_i).$$

命题 1.3（Bayes 公式）　在命题 1.2 的条件下，若 $P(A)>0$，则

$$P(B_j\mid A)=\frac{P(B_j)P(A\mid B_j)}{\sum_{i=1}^{\infty}P(B_i)P(A\mid B_i)}.$$

1.1.3.2　独立性

关于随机事件和随机变量的独立性，下面用一个定义给出.

定义 1.8　设$(\Omega,\mathscr{F},P)$为概率空间，记 $I=\{1,2,\cdots,n\}$.

(1) 如果对$\{1,2,\cdots,n\}$的任意非空子集$\{i_1,i_2,\cdots,i_m\}$,有

$$P\Big(\bigcap_{k=1}^{m}A_{i_k}\Big)=\prod_{k=1}^{m}P(A_{i_k}),$$

则称事件$A_i\in\mathscr{F}(i=1,2,\cdots,n)$相互独立.

(2) 如果

$$F_{\xi_1,\xi_2,\cdots,\xi_n}(x_1,x_2,\cdots,x_n)=F_{\xi_1}(x_1)F_{\xi_2}(x_2)\cdots F_{\xi_n}(x_n),$$

则称随机变量$\xi_1,\xi_2,\cdots,\xi_n$相互独立.

对于独立随机变量,有:

命题 1.4 设$(\Omega,\mathscr{F},P)$为概率空间,$\xi_1,\xi_2,\cdots,\xi_n$为随机变量.

(1) 若$\xi_1,\xi_2,\cdots,\xi_n$独立,且都有有限的数学期望,则

$$E\Big[\prod_{i=1}^{n}\xi_i\Big]=\prod_{i=1}^{n}E[\xi_i].$$

(2) 若$\xi_1,\xi_2,\cdots,\xi_n$独立,且都有有限的方差,则

$$\mathrm{Var}\Big[\sum_{i=1}^{n}\xi_i\Big]=\sum_{i=1}^{n}\mathrm{Var}[\xi_i].$$

1.1.4 大数定律与中心极限定理

本节简单叙述大数定律和中心极限定理的有关主要结果,数理统计的许多处理方法的思想都基于这些基本事实.

命题 1.5(弱大数定律) 设$\xi_1,\xi_2,\cdots$为独立同分布随机变量序列,具有有限的数学期望μ和方差σ^2,则$n^{-1}\sum_{k=1}^{n}\xi_k$依概率收敛于$\mu$,即对任意$\varepsilon>0$,有

$$\lim_{n\to\infty}P\Big(\Big|n^{-1}\sum_{k=1}^{n}\xi_k-\mu\Big|\geqslant\varepsilon\Big)=0.$$

命题 1.6(强大数定律) 设$\xi_1,\xi_2,\cdots$为独立同分布随机变量序列,具有有限的数学期望μ,则$n^{-1}\sum_{k=1}^{n}\xi_k$几乎必然收敛于$\mu$,即

$$P\Big(\lim_{n\to\infty}n^{-1}\sum_{k=1}^{n}\xi_k=\mu\Big)=1.$$

命题 1.7(中心极限定理) 设$\xi_1,\xi_2,\cdots$为独立同分布随机变量序列,具有有限的数学期望μ和方差σ^2,则有

$$\lim_{n\to\infty}P\left([\sigma\sqrt{n}]^{-1}\Big[\sum_{k=1}^{n}\xi_k-n\mu\Big]\leqslant x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-\frac{u^2}{2}}\,\mathrm{d}u.$$

1.2 数理统计的基本概念

本节简单回顾数理统计的基本概念和方法,包括参数估计和假设检验.

1.2.1 总体、样本和统计量

数理统计的基本特征是用局部推断整体. 这个整体在数理统计中称为总体,也就是为了

某一目的需要研究的对象的全体，而将其中的每个对象称为个体或样本. 但是，实际中我们往往关心的是研究对象某方面的数量特征，比如灯泡的寿命、一台机器正常工作的持续时间、某种药物的疗效等. 由于在对一个个体进行试验或观测结束之前，我们无法预知该数量的取值，这一点类似于前面所述的随机变量. 所以，可以认为总体就是一个随机变量，每次试验后，获得的一个个体的具体取值就是该随机变量的一次观测值. 另外，若抽取 n 个个体进行试验或观测，在试验或观测结束之前，也无法预知各个体的取值，因此，从概率分析的角度来说，它们的取值也应该为随机变量.

总之，总体为一个随机变量(比如 ξ)，需要通过观测 n 个样本来推断 ξ 的分布或数字特征. 通常记 ξ 的分布为 $F_\xi(x,\theta)$，其中 $x\in(-\infty,\infty)$ 为实变元，$\theta(\in\Theta)$ 为参数，Θ 称为参数空间，它可能包含多个参数. 样本为 n 维随机向量 $(\xi_1,\xi_2,\cdots,\xi_n)$，$n$ 称为样本容量. 初等数理统计中通常都假定 ξ_i $(i=1,2,\cdots,n)$ 与 ξ 同分布，且相互独立，即 $(\xi_1,\xi_2,\cdots,\xi_n)$ 为简单随机样本. 此时有

$$F_{\xi_1,\xi_2,\cdots,\xi_n}(x_1,x_2,\cdots,x_n;\theta)=F_\xi(x_1,\theta)F_\xi(x_2,\theta)\cdots F_{\xi_n}(x,\theta).$$

其中等号两端可以是分布函数，也可以是分布密度函数(对于连续型随机变量)或分布列(对于离散型随机变量).

综上所述，总体 ξ 是我们研究的目标，而出发点是样本 $(\xi_1,\xi_2,\cdots,\xi_n)$，那么研究的途径或手段就是所谓的统计量. 它的直观意思是要通过观测值的处理来推断总体的分布、数字特征、参数等. 因此，统计量就是样本 $(\xi_1,\xi_2,\cdots,\xi_n)$ 的一个函数，但其中不能含未知参数.

常用的统计量有以下几种：

(1) 样本均值：$\bar{\xi}=\dfrac{1}{n}\sum\limits_{i=1}^{n}\xi_i$.

(2) 样本方差：$S_n^2=\dfrac{1}{n}\sum\limits_{i=1}^{n}(\xi_i-\bar{\xi})^2$.

修正样本方差：$S_n^{*2}=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(\xi_i-\bar{\xi})^2$.

(3) 样本 k 阶原点矩：$\bar{\xi^k}=\dfrac{1}{n}\sum\limits_{i=1}^{n}\xi_i^k$.

(4) 样本 k 阶中心矩：$\dfrac{1}{n}\sum\limits_{i=1}^{n}(\xi_i-\bar{\xi})^k$.

(5) 顺序统计量：$\xi_{(1)}\leqslant\xi_{(2)}\leqslant\cdots\leqslant\xi_{(n)}$，其中 $\xi_{(1)}=\min\{\xi_1,\xi_2,\cdots,\xi_n\}$，$\xi_{(n)}=\max\{\xi_1,\xi_2,\cdots,\xi_n\}$，而 $\xi_{(k)}$ 是将 $\xi_1,\xi_2,\cdots,\xi_n$ 的取值从小到大排列的第 k 位的值.

(6) 样本中位数：

$$\tilde{\xi}=\begin{cases}\xi_{\left(\frac{n+1}{2}\right)} & \text{当 } n \text{ 为奇数}\\ \dfrac{1}{2}\left(\xi_{\left(\frac{n}{2}\right)}+\xi_{\left(\frac{n}{2}+1\right)}\right) & \text{当 } n \text{ 为偶数}\end{cases}$$

(7) 样本极差：$R_n^\xi=\xi_{(n)}-\xi_{(1)}$.

1.2.2 参数的点估计和区间估计

简单地说，所谓参数的点估计，就是找一个合适的统计量，将样本观测值代入该统计量得到的值作为该参数的估计，而参数的区间估计则是找两个统计量，以其中一个为左端点、另一个为右端点，构成一个可能包含该参数的随机区间.

1. 矩法估计

所谓矩法估计，就是用样本 k 阶矩估计总体 k 阶矩. 它的想法来自大数定律，即当样本容量较大时，样本 k 阶矩 $n^{-1}\sum_{i=1}^{n}\xi_i^k$ 与总体 k 阶矩 $E[\xi^k]$ 差别很小(参见命题 1.5). 通常总体各阶矩都与总体参数有关，从而可以通过用样本 k 阶矩估计总体 k 阶矩来实现对参数的点估计.

2. 最大似然估计

参数的最大似然估计的想法基于大家普遍接受的一个事实，即“小概率事件在一次试验中几乎不可能发生”. 换言之，在一次试验中发生的事件，其发生的概率应该比较大. 所以，设总体 $\xi\sim F_\xi(\cdot,\theta)$，当有了一组样本观测值 $(x_1,x_2,\cdots,x_n)$ 时，θ 的取值应该使得样本观测值 $(x_1,x_2,\cdots,x_n)$ 出现的可能性较大. 为确定出 θ 的具体估计量，我们要求 θ 的取值应该使得样本观测值 $(x_1,x_2,\cdots,x_n)$ 出现的可能性达到最大.

下面以总体 ξ 为连续型随机变量的情形来解释如何得到参数的最大似然估计.

设总体 $\xi\sim f_\xi(\cdot,\theta)$(连续型时为分布密度函数，离散型时为分布列)，要求 θ 的取值使得 $(\xi_1,\xi_2,\cdots,\xi_n)$ 的联合密度函数在样本观测值 $(x_1,x_2,\cdots,x_n)$ 处取到最大. 记 $(\xi_1,\xi_2,\cdots,\xi_n)$ 的联合密度函数为 $L(x_1,x_2,\cdots,x_n;\theta)$，则

$$L(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n}f_\xi(x_i,\theta).$$

称 L 为 ξ 的似然函数.

求似然函数的最大值点，往往通过求它的驻点，即关于 θ 的导数为 0 的点. 似然函数为 n 个密度函数的乘积，求导比较繁琐. 通常可利用对数函数 $\ln(\cdot)$ 的单增性，将问题转化为求 $\ln(L(x_1,x_2,\cdots,x_n;\theta))$ 的最大值点.

3. 参数的区间估计

与参数的点估计不同，区间估计是以两个统计量 T_1 和 T_2 为左右端点构成一个随机区间，并且要求该随机区间包含待估计参数的概率为 $1-\alpha$，即

$$P\big(T_1(\xi_1,\xi_2,\cdots,\xi_n)\leqslant g(\theta)\leqslant T_2(\xi_1,\xi_2,\cdots,\xi_n)\big)=1-\alpha.$$

此时称 $[T_1(\xi_1,\xi_2,\cdots,\xi_n),T_2(\xi_1,\xi_2,\cdots,\xi_n)]$ 为 θ 的置信度为 $1-\alpha$ 的区间估计 $(0<\alpha<1)$.

由于对于一般总体，统计量的分布难以得到闭形式表达式，往往用正态总体来近似. 正态总体参数的区间估计见表 1-2.

表 1-2 正态总体参数的区间估计

总体数目	待估计参数	用到的样本函数及其分布	区间估计
一个总体，$\xi\sim N(\mu,\sigma^2)$，$\bar{\xi}=\frac{1}{n}\sum_{i=1}^{n}\xi_i$，$S_n^2=\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2$	$\mu(\sigma^2=\sigma_0^2$ 已知$)$	$\dfrac{\bar{\xi}-\mu}{\sqrt{\frac{\sigma_0^2}{n}}}\sim N(0,1)$	$\left[\bar{\xi}\pm u_{1-\frac{\alpha}{2}}\sqrt{\dfrac{\sigma_0^2}{n}}\right]$
	$\mu(\sigma^2$ 未知$)$	$\dfrac{\bar{\xi}-\mu}{S_n}\sqrt{n-1}\sim t(n-1)$	$\left[\bar{\xi}\pm t_{1-\frac{\alpha}{2}}(n-1)\dfrac{S_n}{\sqrt{n-1}}\right]$
	$\sigma^2(\mu=\mu_0$ 已知$)$	$\sum_{i=1}^{n}\dfrac{(\xi_i-\mu_0)^2}{\sigma^2}\sim\chi^2(n)$	$\left[\dfrac{\sum_{i=1}^{n}(\xi_i-\mu_0)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)},\dfrac{\sum_{i=1}^{n}(\xi_i-\mu_0)^2}{\chi^2_{\frac{\alpha}{2}}(n)}\right]$
	$\sigma^2(\mu$ 未知$)$	$\dfrac{nS_n^2}{\sigma^2}\sim\chi^2(n-1)$	$\left[\dfrac{nS_n^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)},\dfrac{nS_n^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right]$

（续表 1-2）

总体数目	待估计参数	用到的样本函数及其分布	区间估计
两个总体，$\xi\sim N(\mu_1,\sigma_1^2)$，$\bar{\xi}=\frac{1}{m}\sum_{i=1}^{m}\xi_i$，$S_{1m}^2=\frac{1}{m}\sum_{i=1}^{m}(\xi_i-\bar{\xi})^2$. $\eta\sim N(\mu_2,\sigma_2^2)$，$\bar{\eta}=\frac{1}{n}\sum_{j=1}^{n}\eta_j$，$S_{2n}^2=\frac{1}{n}\sum_{j=1}^{n}(\eta_i-\bar{\eta})^2$.	$\mu_1-\mu_2$ （σ_1^2,σ_2^2 已知）	$\frac{(\bar{\xi}-\bar{\eta})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}\sim N(0,1)$	$\left[(\bar{\xi}-\bar{\eta})\pm u_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}\right]$
	$\mu_1-\mu_2$ （$\sigma_1^2=\sigma_2^2$ 未知）	$\frac{(\bar{\xi}-\bar{\eta})-(\mu_1-\mu_2)}{S_{\xi\eta}}\sim t(m+n-2)$ $S_{\xi\eta}=\sqrt{\frac{m+n}{(m+n-2)mn}}\cdot\sqrt{mS_{1m}^2+nS_{2n}^2}$	$\left[\bar{\xi}-\bar{\eta}\pm t_{1-\alpha/2}(m+n-2)S_{\xi\eta}\right]$
	$\frac{\sigma_1^2}{\sigma_2^2}$（$\mu_1,\mu_2$ 已知）	$\frac{n\sum_{i=1}^{m}(\xi_i-\mu_1)^2}{m\sum_{j=1}^{n}(\eta_j-\mu_2)^2}\cdot\frac{\sigma_2^2}{\sigma_1^2}\sim F(m,n)$	$\left[\frac{1}{F_{1-\alpha/2}(m,n)}\frac{n\sum_{i=1}^{m}(\xi_i-\mu_1)^2}{m\sum_{j=1}^{n}(\eta_j-\mu_2)^2},\ \frac{1}{F_{\alpha/2}(m,n)}\frac{n\sum_{i=1}^{m}(\xi_i-\mu_1)^2}{m\sum_{j=1}^{n}(\eta_j-\mu_2)^2}\right]$
	$\frac{\sigma_1^2}{\sigma_2^2}$（$\mu_1,\mu_2$ 未知）	$\frac{(n-1)\sum_{i=1}^{m}(\xi_i-\bar{\xi})^2}{(m-1)\sum_{j=1}^{n}(\eta_j-\bar{\eta})^2}\cdot\frac{\sigma_2^2}{\sigma_1^2}\sim F(m-1,n-1)$	$\left[\frac{1}{F_{1-\alpha/2}(m-1,n-1)}\cdot\frac{(n-1)\sum_{i=1}^{m}(\xi_i-\bar{\xi})^2}{(m-1)\sum_{j=1}^{n}(\eta_j-\bar{\eta})^2},\ \frac{1}{F_{\alpha/2}(m-1,n-1)}\cdot\frac{(n-1)\sum_{i=1}^{m}(\xi_i-\bar{\xi})^2}{(m-1)\sum_{j=1}^{n}(\eta_j-\bar{\eta})^2}\right]$

1.2.3 假设检验

假设检验是统计推断的一个重要方面. 它所针对的问题，从实际应用意义上来说，往往从经验可以大体判断出某种随机变化的量的分布类型，或者参数在某个范围内，但无法从数学上说明其正确性，此时通过观测样本就可以用假设检验来推断该经验结论是否正确，或者在哪个置信度下，该经验的结论是可信的. 用数理统计的术语来说，分为非参数假设检验和参数假设检验. 前者是对总体的分布提出假设（比如，H_0（原假设或 0 假设）：$\xi\sim F_0(\cdot,\theta)$，$H_1$（备选假设）：$\xi\nsim F_0(\cdot,\theta)$）；后者是总体的分布类型是已知的，其统计假设是关于参数的（比如，$H_0:\theta\leqslant\theta_0$；$H_1:\theta>\theta_0$），需要我们通过样本$(\xi_1,\xi_2,\cdots,\xi_n)$来对假设做出推断（比如，是拒绝 H_0 还是接受 H_0？ 在何种显著性水平下拒绝或接受？）.

由于一次样本观测值为 $\mathbf{R}^n$ 中的一个点，因此假设检验问题的回答，最终是将 $\mathbf{R}^n$ 分成互不相交的两部分，一部分是使得小概率事件 A 发生的点的集合，称其为该假设检验问题

的拒绝域;另一部分是拒绝域的余集,称其为该假设检验问题的接受域.

假设检验的基本思路也是基于大家的一个共识,那就是"小概率事件在一次试验中几乎不会发生".具体来说,如果某假设 H_0 成立,则某事件 A 发生的可能性很小.但样本观测的结果是 A 发生了,这说明 A 不是小概率事件,这个矛盾说明假设 H_0 应该是不成立的.

对一个假设检验问题作统计推断,就是在原假设 H_0 成立的前提下,从样本($\xi_1,\xi_2,\cdots,\xi_n$)构造一个概率不超过 α 的小概率事件 A,然后基于一次观测的样本观测值看该小概率事件是否发生,来决定是否拒绝原假设 H_0.显然,这种统计推断可能会犯两类错误,即拒真(第一类错误)和受伪(第二类错误).

实际应用中样本容量不可能无限制大,从而同时使两类错误都很小是不可能的.一般都是在控制第一类错误的概率不超过某值 α_0 前提下,使犯第二类错误的概率尽可能小.另外,有一种检验方法是只控制第一类错误而不控制第二类错误,这种检验方法称为显著性检验.也就是说,当原假设 H_0"显著地"不真或不正确时就拒绝 H_0,否则不拒绝或勉强接受 H_0.

对于显著性检验,需要强调指出的是,这种检验方法实际上是"保护"(不轻易拒绝)原假设 H_0 的,这是因为当小概率事件发生时才拒绝它,但小概率事件通常几乎不发生.另外,如果显著性检验的结果是拒绝原假设 H_0,那么该推断的可信程度较高,而如果检验的结果是不拒绝原假设 H_0,那么此时接受原假设 H_0 的推断,对原假设 H_0 的成立是没有说服力的,这是因为一个事件发生概率较大,它在一次试验中发生是应该的.因此,当想用显著性检验对某一猜测结论作强有力的支持时,应将该猜测结论的反面作为原假设.

1.2.3.1 正态总体参数的假设检验

正态总体参数的假设检验问题的有关结果见表 1-3～表 1-6,其中有关统计量的记号参见表 1-2.

表 1-3 单个正态总体均值 μ 的假设检验的拒绝域(显著性水平为 α)

序号	H_0	H_1	σ^2 已知	σ^2 未知
Ⅰ	$\mu=\mu_0$	$\mu\neq\mu_0$	$\dfrac{\lvert\bar{\xi}-\mu_0\rvert}{\sigma/\sqrt{n}}\geqslant u_{1-\alpha/2}$	$\dfrac{\lvert\bar{\xi}-\mu_0\rvert}{S_n/\sqrt{n-1}}\geqslant t_{1-\alpha/2}(n-1)$
Ⅱ	$\mu=\mu_0$	$\mu>u_0$	$\dfrac{\bar{\xi}-\mu_0}{\sigma/\sqrt{n}}\geqslant u_{1-\alpha}$	$\dfrac{\bar{\xi}-\mu_0}{S_n/\sqrt{n-1}}\geqslant t_{1-\alpha}(n-1)$
Ⅲ	$\mu\leqslant\mu_0$	$\mu>u_0$		
Ⅳ	$\mu=\mu_0$	$\mu<u_0$	$\dfrac{\bar{\xi}-\mu_0}{\sigma/\sqrt{n}}\leqslant u_{1-\alpha}$	$\dfrac{\bar{\xi}-\mu_0}{S_n/\sqrt{n-1}}\geqslant t_{1-\alpha}(n-1)$
Ⅴ	$\mu\geqslant\mu_0$	$\mu<u_0$		

注:统计软件在对假设检验问题作显著性检验时,往往不事先给定显著性水平,而是在打印出有关统计量的值的同时打印出一个 p 值,它是此时拒绝原假设时所犯错误(即第一类错误)的概率,也就是显著性水平.

表 1-4 单个正态总体方差 σ^2 的假设检验的拒绝域(显著性水平为 α)

序号	H_0	H_1	μ 已知	μ 未知
Ⅰ	$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\dfrac{\sum_{i=1}^{n}(\xi_i-\mu)^2}{\sigma_0^2}\leqslant\chi^2_{\alpha/2}(n)$ 或 $\dfrac{\sum_{i=1}^{n}(\xi_i-\mu)^2}{\sigma_0^2}\geqslant\chi^2_{1-\alpha/2}(n)$	$\dfrac{\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2}{\sigma_0^2}\leqslant\chi^2_{\alpha/2}(n-1)$ 或 $\dfrac{\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2}{\sigma_0^2}\geqslant\chi^2_{1-\alpha/2}(n-1)$