专业课辅导+真题资料+微信公众号：432应统书院

高等学校教材

概率论及数理统计

第4版　下册

中山大学 邓集贤 杨维权 司徒荣 邓永录 编著

高等教育出版社

专业课辅导+真题资料+微信公众号：432应统书院

ISBN 978-7-04-026630-6

9 787040 266306 >

定价 17.60 元

专业课辅导+真题资料+微信公众号：432应统书院

高等学校教材

概率论及数理统计

下册

第 4 版

中山大学

邓集贤　杨维权　司徒荣　邓永录　编著

高等教育出版社

专业课辅导+真题资料+微信公众号：432应统书院

图书在版编目(CIP)数据

概率论及数理统计．下册/邓集贤等编著．—4版．
—北京:高等教育出版社,2009.7
ISBN 978-7-04-026630-6

Ⅰ．概…　Ⅱ．邓…　Ⅲ．①概率论-高等学校-教材
②数理统计-高等学校-教材　Ⅳ．O21

中国版本图书馆CIP数据核字（2009）第087920号

策划编辑　李　蕊　　责任编辑　崔梅萍　　封面设计　刘晓翔
责任绘图　吴文信　　版式设计　王　莹　　责任校对　殷　然
责任印制　朱学忠

出版发行	高等教育出版社	购书热线	010-58581118
社　　址	北京市西城区德外大街4号	咨询电话	400-810-0598
邮政编码	100120	网　　址	http://www.hep.edu.cn
总　　机	010-58581000		http://www.hep.com.cn
		网上订购	http://www.landraco.com
经　　销	蓝色畅想图书发行有限公司		http://www.landraco.com.cn
印　　刷	河北新华印刷一厂	畅想教育	http://www.widedu.com
		版　　次	1980年3月第1版
开　　本	850×1168　1/32		2009年7月第4版
印　　张	12.375	印　　次	2009年7月第1次印刷
字　　数	310 000	定　　价	17.60元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换。

版权所有　侵权必究

物料号　26630-00

专业课辅导+真题资料+微信公众号：432应统书院

内 容 简 介

本书是在中山大学统计科学系梁之舜等五人编著的《概率论及数理统计》(第三版)的基础上修订而成的,具有适应面广、便于自学的特点。本次修订删除了第五章的内容,其他各章保留原有的特点、结构和基本内容,进行了适当的修改和补充,习题也作了更新修订,使本书更适应当前的教学。全书共十二章,仍分上、下两册出版。

本书可作为综合性大学、师范院校及其他院校的数学类专业教材,也可作为其他有关专业的教材或教学参考书。

专业课辅导+真题资料+微信公众号：432应统书院

第四版说明

本书《概率论及数理统计》(上、下册)作为全国高等院校中概率统计课程的通用教材,高等教育出版社于1980年出版(1988年第二版,2005年第三版)。根据当前高等院校教学的具体情况和需要,我们对本书重新修订,作为第四版继续出版发行。本书编者讨论了第四版修订的基本原则,确定本书原有的内容、结构与特点基本保持不变,对某些章节(内容与习题)作了适当的调整、修改与补充。

概率论及数理统计是研究随机现象统计规律性及其应用的一门数学学科,它像所有其他数学学科一样有基本概念和逻辑推理的严谨性,更有应用的广泛性。我们在编写中既重视各种概念、定义、定理、公式的严谨叙述和推理,也注重对理论的直观背景的解析和实际应用的示例。

学习本书只要求微积分与高等代数的基础知识,对于偏难的有关内容,采用星号"*"或指出参考书目。本书(上、下册)共设十二章,第一章至第九章可作为概率统计课程的基本内容。不同专业可按其教学计划的要求,从本书(上、下册)中选讲有关章节内容。

本书第一、四、五、十一各章由邓集贤编写;第二、三各章由司徒荣编写;第六、七、八、九、十各章由杨维权编写;第十二章由邓永录编写。编者非常感谢多年来采用本书的广大读者的厚爱和对本书作过指导与帮助的专家学者,衷心感谢高等教育出版社的大力支持和李蕊同志的具体帮助。本书(上、下册)再版后,可能还会有许多不妥之处,恳请读者及专家批评指正。

编　者

2009年2月于广州中山大学

专业课辅导+真题资料+微信公众号：432应统书院

目　录

第六章　数理统计的基本概念 …… 1
§6.1　基本概念 …… 1
一、总体、个体、简单随机样本 …… 1
二、统计量 …… 3
三、小样问题与大样问题 …… 5
§6.2　样本的数字特征及其分布 …… 6
*一、经验分布与格列汶科定理 …… 6
二、样本的数字特征 …… 8
三、样本数字特征的分布 …… 9
§6.3　抽样分布定理 …… 15
习题 …… 23
第七章　参数估计 …… 26
§7.1　矩法与极大似然法 …… 27
一、矩法 …… 27
二、极大似然法 …… 31
§7.2　无偏性与优效性 …… 39
一、无偏性 …… 40
二、优效性 …… 48
三、相合性 …… 53
*§7.3　充分性与完备性 …… 57
一、充分性 …… 58
二、完备性 …… 66
§7.4　区间估计 …… 72
习题 …… 77
第八章　假设检验 …… 81
§8.1　基本概念 …… 81

专业课辅导+真题资料+微信公众号：432应统书院

§8.2 参数假设检验 …… 86
一、正态总体数学期望 a 的检验问题 …… 87
二、正态总体方差 σ^2 的检验问题 …… 94
*三、非正态总体的参数假设检验 …… 99
*四、讨论两个问题 …… 100
§8.3 非参数的检验 …… 103
一、分布函数的拟合检验 …… 103
二、独立性的检验 …… 114
*§8.4 最佳检验 …… 118
一、两类错判 …… 118
二、功效函数 …… 122
三、最佳检验 …… 125
*§8.5 样本容量 n 的确定 …… 135
一、参数估计与检验中 n 的确定 …… 135
二、最佳检验中 n 的确定 …… 139
三、验收抽样方案中 n 的确定 …… 141
习题 …… 143
第九章 回归分析与方差分析 …… 148
§9.1 线性回归分析 …… 148
§9.2 最小二乘法估计 …… 153
一、参数的最小二乘法估计 …… 153
二、最小二乘法估计量的性质 …… 155
§9.3 例题 …… 162
一、讨论三个例题 …… 162
二、将曲线方程线性化 …… 174
*§9.4 假设检验 …… 182
一、线性模型的假设检验 …… 182
二、回归系数的假设检验 …… 183
*§9.5 单因子方差分析 …… 184
习题 …… 189
第十章 统计决策及贝叶斯统计 …… 197

专业课辅导+真题资料+微信公众号：432应统书院

§10.1 极大极小估计 …… 199
一、决策论的基本概念 …… 199
二、极大极小估计 …… 200
§10.2 贝叶斯统计 …… 202
一、贝叶斯估计 …… 203
二、区间估计 …… 210
三、假设检验 …… 212
*四、共轭先验分布 …… 214
§10.3 应用事例 …… 218
习题 …… 227
第十一章 随机过程引论 …… 233
§11.1 随机过程的概念 …… 233
一、随机过程的直观背景和定义 …… 233
二、随机过程的有穷维分布函数族 …… 236
§11.2 几类重要的随机过程简介 …… 243
一、独立增量过程(可加过程) …… 243
二、正态随机过程(高斯过程) …… 244
三、维纳过程 …… 244
四、泊松过程 …… 245
五、随机点过程与计数过程 …… 246
§11.3 马氏链 …… 256
一、定义及例 …… 256
二、齐次马氏链 …… 259
三、遍历性、最终分布与平稳分布 …… 262
四、分支过程 …… 267
五、销售市场决策中应用的例子 …… 268
§11.4 连续时间马尔可夫链 …… 272
一、定义 …… 273
二、生灭过程 …… 275
§11.5 均方微积分与随机微分方程 …… 276
一、随机序列的均方收敛 …… 276

专业课辅导+真题资料+微信公众号：432应统书院

二、随机过程的均方连续 …… 278
三、随机过程的均方积分 …… 279
四、随机过程的均方导数 …… 282
五、随机微分方程 …… 287
§11.6 平稳随机过程 …… 291
一、定义及例 …… 291
二、相关函数 …… 293
三、弱平稳随机过程的功率谱密度 …… 296
四、遍历性定理 …… 300
§11.7 时间序列与离散鞅 …… 303
一、时间序列分析 …… 303
二、中国消费与积累的非线性模型 …… 306
三、离散鞅 …… 309
第十二章 概率统计在计算方法中的一些应用 …… 314
§12.1 蒙特卡罗方法与均匀分布随机数 …… 314
§12.2 连续型随机变量的一般模拟方法 …… 317
一、反函数法 …… 317
二、舍选法 …… 318
§12.3 连续型随机变量的特殊模拟方法 …… 322
一、正态分布随机数 …… 322
二、瑞利分布随机数 …… 324
三、指数分布随机数 …… 324
四、Γ 分布和 χ^2 分布随机数 …… 326
§12.4 离散型随机变量的模拟 …… 330
一、一般方法 …… 330
二、基于伯努利试验概型的方法 …… 332
三、其他方法 …… 333
§12.5 随机向量和随机过程的模拟 …… 334
一、随机向量的模拟 …… 334
二、齐次泊松过程的模拟 …… 336
三、马尔可夫链的模拟 …… 337
§12.6 定积分的概率计算方法 …… 339

专业课辅导+真题资料+微信公众号：432应统书院

一、常用的两种算法 …… 339
二、重积分的计算 …… 345
§12.7 某些方程的概率解法 …… 347
一、线性方程组的求解 …… 347
二、一些偏微分方程的求解 …… 350
下册习题答案 …… 355
附表 …… 361
表1 χ^2-分布的上侧临界值表 …… 361
表2 t-分布的双侧临界值表 …… 363
表3 F检验的临界值(F_α)表 …… 365
表4 检验相关系数$\rho=0$的临界值(r_α)表 …… 375
表5 随机数表 …… 376
译名对照表 …… 380
参考文献 …… 381

专业课辅导+真题资料+微信公众号：432应统书院

第六章　数理统计的基本概念

前述各章，是概率论的引论，讲述了概率论的基本内容，为数理统计学建立了重要的数学基础. 从本章起的接连五章，是数理统计学的初步，主要讲述基本概念、估计与检验等原理、回归分析与方差分析、贝叶斯统计等统计方法.

数理统计学是运用概率论的基本知识，对要研究的随机现象进行多次观察或试验，研究如何合理地获得数据资料，建立有效的数学方法，根据所获得的数据资料，对所关心的问题作出估计与检验. 数理统计学的重要分支有统计推断、多元统计分析、试验设计等，其具体方法甚多，应用相当广泛，已成为各学科从事科学研究及生产、经济等部门进行有效工作的必不可少的数学工具.

在这一章中，我们从数理统计学的基本概念讲起，讨论抽样分布及其重要的定理. 这些抽样分布及其几个重要的定理，在前述各章中尚未提到，而在后述各章中却经常要用到它们.

§6.1　基本概念

一、总体、个体、简单随机样本

总体、个体、样本是数理统计学中三个最基本的术语. 我们把对某一个问题的研究对象的全体称为总体（或母体），组成总体的每个基本成员称为个体，从总体中随机抽取的 n 个个体称为容量为 n 的样本（或子样）. 例如，把某月的整批产品视为一总体，则每个产品为个体；又如把某批灯泡视为一总体，则每个灯泡为个体；某地在某季度内每天的日平均气温的全体视为一总体，则其中某天的日平均气温为个体. 在数理统计学中，我们是对总体成员的一个或者若干个数量表征进行研究. 如日平均气温的摄氏度数用 ξ

专业课辅导+真题资料+微信公众号：432应统书院

表示，灯泡的使用寿命（小时）用 η 表示，ξ 及 η 都为随机变量，这样对总体的研究就归结为讨论随机变量 ξ 或 η 的分布函数及其主要数字特征（如数学期望 $E(\xi)$、方差 $D(\xi)$ 等）的研究. 又如某月的整批产品为一总体，若只要检查产品的质量，可用数量表征 ζ 来反映. 产品为一等品，令 ζ 取值为 1；二等品，ζ 取值为 2；不合格品，ζ 取值为 3. 我们研究这批产品的质量规律，就归结为讨论随机变量 ζ 的分布函数及其主要数字特征来反映.

今后我们常用"总体 ξ 服从什么分布"这样的术语，它是指总体的某个具体数量表征 ξ 服从什么分布规律. 我们说对总体进行 n 次独立的重复试验或 n 次独立的观察，就是从总体中随机地抽取容量为 n 的样本，对应的数量表征以 $\xi_1,\xi_2,\cdots,\xi_n$ 表示，总体 ξ 及 $\xi_1,\xi_2,\cdots,\xi_n$ 都是随机变量，$(\xi_1,\xi_2,\cdots,\xi_n)$ 是 n 维随机向量. 例如我们说整批灯泡的寿命分布如何？是指灯泡的寿命 ξ 服从什么分布. 从整批灯泡中随机抽取容量为 n 的样本，是指随机抽取 n 个灯泡，观察这 n 个灯泡的寿命，用 $\xi_1,\xi_2,\cdots,\xi_n$ 表示，其中 ξ_i 表示随机抽取到的第 i 个灯泡的寿命，显然 ξ_i 也是一随机变量（$i=1,2,\cdots,n$）. 为区别总体与样本的概念，我们今后常用诸如"总体 ξ 服从什么分布"、"从总体 ξ 中抽取样本 $\xi_1,\xi_2,\cdots,\xi_n$"等这类术语.

我们在对总体进行研究中，当总体相当大时（指个体相当多时）或当对个体进行试验具有破坏性（如炮弹能否引发爆炸）或费时耗资大时，只可能从总体中抽取部分个体来作研究，即抽取容量为 n 的样本来作研究. 我们要从样本的观察或试验结果的特性来对总体的特性作出估计与推断，一方面自然要研究应该怎样从总体中抽取样本，使得样本在尽可能大的程度上反映总体的特性；同时必须建立一整套的方法，使能根据所选取的样本的性质，来对总体的特性进行估计与推断. 因此，我们在抽取样本来对总体作出估计与推断时，从总体中抽取样本必须是随机的，即每一个个体都有同等概率被抽取（当总体中的个体数目较少时，要用有返回抽取方式）. 其具体要求为两个方面：独立性，是指 $\xi_1,\xi_2,\cdots,\xi_n$ 相互独

专业课辅导+真题资料+微信公众号：432应统书院

立；代表性，是指 $\xi_1,\xi_2,\cdots,\xi_n$ 中每一个都与总体 ξ 有相同分布.

定义 6.1.1（简单随机样本） 设 $\xi_1,\xi_2,\cdots,\xi_n$ 为来自总体 ξ 的容量为 n 的样本，如果 $\xi_1,\xi_2,\cdots,\xi_n$ 相互独立且每一个都是与总体 ξ 有相同分布的随机变量，则称 $\xi_1,\xi_2,\cdots,\xi_n$ 为总体 ξ 的容量为 n 的简单随机样本，简称为简单样本或样本.

在这一章中，我们所讲的样本 $\xi_1,\xi_2,\cdots,\xi_n$，无特别声明的话，都是指简单样本，即 $\xi_1,\xi_2,\cdots,\xi_n$ 相互独立且每一个都与总体 ξ 有相同的分布. 引入简单随机样本，是基于要从样本所获得的概率统计特性来对总体的概率统计特性作出估计与推断. 对于简单随机样本，我们可以应用概率论中对独立随机变量的情形所建立的许多重要的定理，这些重要的结论为数理统计学提供了必要的基础.

二、统计量

样本 $\xi_1,\xi_2,\cdots,\xi_n$ 也可用 n 维随机向量 $(\xi_1,\xi_2,\cdots,\xi_n)$ 表示. 记 x_i 为 ξ_i 的一次观察值，并称 $(x_1,x_2,\cdots,x_n)$ 为样本 $\xi_1,\xi_2,\cdots,\xi_n$ 的一次观察值.

定义 6.1.2（统计量） 设 $\xi_1,\xi_2,\cdots,\xi_n$ 为总体 ξ 的样本，T 为 n 维实值函数，作样本 $\xi_1,\xi_2,\cdots,\xi_n$ 的函数 $T=T(\xi_1,\xi_2,\cdots,\xi_n)$（不带未知参数的随机变量），$T$ 的取值记为 $t=T(x_1,x_2,\cdots,x_n)$，称 T 或 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 为统计量.

我们在用样本 $\xi_1,\xi_2,\cdots,\xi_n$ 获得的信息来对总体 ξ 作出估计与推断时，是按不同的统计问题的要求而规定样本的各种函数. 在这本教材中，所涉及的样本的各种函数 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 都是随机变量.

例 6.1.1 设 $\xi_1,\xi_2,\cdots,\xi_n$ 为总体 ξ 的样本，其容量为 n. 记

$$\bar{\xi}=\frac{1}{n}\sum_{i=1}^{n}\xi_i,\qquad S^2=\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2,$$

则 $\bar{\xi}$ 及 S^2 都是统计量，称 $\bar{\xi}$ 及 S^2 分别为样本 $\xi_1,\xi_2,\cdots,\xi_n$ 的平均值

专业课辅导+真题资料+微信公众号：432应统书院

及方差. 样本的观察值为 $x_1, x_2, \cdots, x_n$, $\bar{\xi}$ 及 S^2 的观察值分别记作

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n} x_i, \qquad s^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2.$$

今后,大写的 S^2 表示统计量,小写的 s^2 表示统计量 S^2 的观察值.

定义 6.1.3(顺序统计量) 设 $\xi_1, \xi_2, \cdots, \xi_n$ 为总体 ξ 的样本,今由样本建立 n 个函数:

$$\xi_k^*=\xi_k^*(\xi_1, \xi_2, \cdots, \xi_n), \qquad k=1,2,\cdots,n,$$

其中 ξ_k^* 为这样的统计量,它的观察值为 x_k^*, x_k^* 为样本 $\xi_1, \xi_2, \cdots, \xi_n$ 的观察值 $x_1, x_2, \cdots, x_n$ 中由小至大排列(即 $x_1^* \leqslant \cdots \leqslant x_k^* \leqslant \cdots \leqslant x_n^*$)后的第 k 位数值,则称 $\xi_1^*, \xi_2^*, \cdots, \xi_n^*$ 为顺序统计量.

易见,$\xi_1^*=\min\{\xi_1, \xi_2, \cdots, \xi_n\}$, $\xi_n^*=\max\{\xi_1, \xi_2, \cdots, \xi_n\}$. 称 ξ_1^* 为最小项统计量,ξ_n^* 为最大项统计量. 若 n 为奇数,则称 $\xi_{\frac{n+1}{2}}^*$ 为样本的中值;若 n 为偶数,则称 $\xi_{\frac{n}{2}+1}^*$ 为样本的中值.

定义 6.1.4(极差) 设 $\xi_1, \xi_2, \cdots, \xi_n$ 为总体 ξ 的样本,则称统计量 $D_n^*=\xi_n^*-\xi_1^*$ 为样本的极差.

极差是样本中最大值与最小值之差,反映了样本观察值的波动幅度. 它同方差一样是反映观察值离散程度的数量指标,而且计算方便.

例 6.1.2 设 $\xi_1, \xi_2, \cdots, \xi_5$ 为总体 ξ 的容量为 5 的样本,今对这个样本作了三次观察,其值如表 6.1.1 所示,试求 $\bar{\xi}$、S^2 及 D_n^* 的观察值.

表 6.1.1

x \ ξ	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5
1	3	1	10	5	6
2	2	6	7	2	8
3	8	3	9	10	5

专业课辅导+真题资料+微信公众号：432应统书院

解 计算结果列于表 6.1.2.

表 6.1.2

ξ_1^*	ξ_2^*	ξ_3^*	ξ_4^*	ξ_5^*	$\overline{x}$	s^2	D_n^*
1	3	5	6	10	5	9.2	9
2	2	6	7	8	5	6.4	6
3	5	8	9	10	7	6.8	7

三、小样问题与大样问题

统计量是我们对总体 ξ 的分布函数或数字特征进行估计与推断最重要的基本概念，求出统计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 的分布函数是数理统计学的基本问题之一. 统计量的分布，称为抽样分布.

设总体 ξ 的分布函数表达式已知，对于任一正整数 n，如能求出给定统计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 的分布函数，这分布称为统计量 T 的精确分布. 求出统计量 T 的精确分布，这对于数理统计学中的所谓小样问题（即在样本容量 n 比较小的情况下所讨论的各种统计问题）的研究是很重要的.

但一般说来，要确定一个统计量的精确分布其难度比较大. 只对一些重要的特殊情形，如总体 ξ 服从正态分布时，已求出 t 统计量、χ^2 统计量、F 统计量等的精确分布. 它们在参数的估计及检验中起很重要的作用.

若统计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 的精确分布求不出来，或其表达式非常复杂而难以应用，但如能求出它在 $n\to\infty$ 时的极限分布，那么这个统计量的极限分布对于数理统计学中的所谓大样问题（即在样本容量 n 比较大的情况下讨论的各种统计问题）的研究很有用. 但要注意，在应用极限分布时，要求样本的容量 n 比较大. 如第八章的 §8.3 所讨论的非参数性假设检验问题，是用检验统计量的极限分布，因而样本容量 n 应取得比较大才行.

专业课辅导+真题资料+微信公众号：432应统书院

§6.2 样本的数字特征及其分布

* 一、经验分布与格列汶科定理

在实际工作中遇到各种各样的随机变量，怎样确定它的分布函数 $F(x)$？在概率论中，我们介绍了常用的几种分布函数以及它们的一些性质，在那里我们假定它们都是事先给定了的. 现在，我们介绍用独立重复试验的方法，利用样本建立一定的概率模型，用由此所获得的概率统计特性来对总体 ξ 的分布函数 $F(x)$ 等作出统计推断.

定义 6.2.1（经验分布函数） 从总体 ξ 中抽取容量为 n 的样本 $\xi_1,\xi_2,\cdots,\xi_n$，当顺序统计量 $\xi_1^*,\xi_2^*,\cdots,\xi_n^*$ 的值给定时，对任何实数 x，我们定义函数 $F_n^*(x)$：

$$F_n^*(x)=\begin{cases}0, & x\leqslant x_1^*;\\ \dfrac{k}{n}, & x_k^*<x\leqslant x_{k+1}^*,k=1,2,\cdots,n-1;\\ 1, & x>x_n^*.\end{cases}\tag{6.2.1}$$

称 $F_n^*(x)$ 为总体 ξ 的经验分布函数①

易见，对于每一组观察值 $\xi_i^*=x_i^*,i=1,2,\cdots,n$，$F_n^*(x)$ 单调、非降、左连续且在 $x=x_k^*$ 点有间断点，在每个间断点上跳跃量都是 $\dfrac{1}{n}$（如图 6.2.1）. 显然，$0\leqslant F_n^*(x)\leqslant 1$，并具有分布函数的其他性质.

由定义 6.2.1 知，对于 x 的每一数值而言，经验分布函数

① 对于每一 ω，即样本的一次观察值，由（6.2.1）定义的 $F_n^*(x)$ 是实数 x 的函数，$x\in\mathbf{R}_1$. 而对于每一 x 值，$F_n^*(x)$ 依赖于样本的观察值，即为 ω 的函数，因而 $F_n^*(x)$ 为一统计量. 这里，如果 $x_k<x_{k+1}=x_{k+2}=\cdots$，而 $x_k<x=x_{k+1}=x_{k+2}=\cdots$，则应在（6.2.1）中把 $\dfrac{k}{n}$ 改为 $\dfrac{N_k(x)}{n}$，其中 $N_k(x)$ 是样本中小于 x_{k+1} 的观察值的个数.

专业课辅导+真题资料+微信公众号：432应统书院

图 6.2.1

$F_n^*(x)$为样本 $\xi_1,\xi_2,\cdots,\xi_n$ 的函数，它是一统计量，即为一随机变量，其可能取值为 $0,\frac{1}{n},\cdots,\frac{n-1}{n},1$. 事件"$F_n^*(x)=\frac{k}{n}$"发生的概率，由于 $\xi_1,\xi_2,\cdots,\xi_n$ 相互独立且有相同的分布函数 $F(x)$，因而它等价于 n 次独立重复试验的伯努利概型中事件"$\xi<x$"发生 k 次而其余 $n-k$ 次不发生的概率，即有：

$$P\left\{F_n^k(x)=\frac{k}{n}\right\}=C_n^k\{F(x)\}^k\{1-F(x)\}^{n-k}, \qquad (6.2.2)$$

其中 $F(x)=P\{\xi<x\}$，它是总体 ξ 的分布函数.

我们从第五章所讲的大数定律知道，在一定的条件下，事件发生的频率依概率收敛于这个事件发生的概率. 人们自然要问，总体 ξ 的经验分布函数 $F_n^*(x)$，即"$\xi<x$"事件发生的频率，当 n 足够大时，是否也渐近于事件"$\xi<x$"发生的概率，即总体 ξ 的分布函数 $F(x)$呢？格列汶科于 1933 年作了肯定的回答.

格列汶科定理 设总体 ξ 的分布函数为 $F(x)$，经验分布函数为 $F_n^*(x)$，对于任何实数 x，记

$$D_n=\sup_{-\infty<x<+\infty}|F_n^*(x)-F(x)|, \qquad (6.2.3)$$

则有

$$P\{\lim_{n\to\infty}D_n=0\}=1. \qquad (6.2.4)$$

我们知道 $F_n^*(x)$为一统计量，因而 D_n 也为一统计量. D_n 这个

专业课辅导+真题资料+微信公众号：432应统书院

统计量用来衡量 $F_n^*(x)$ 同 $F(x)$ 之间在所有的 x 值上最大的差异程度. 格列汶科定理证明了统计量 D_n 以概率为 1 地收敛于零. 通俗地说,就是当 n 足够大时,对于所有的 x 值,$F_n^*(x)$ 同 $F(x)$ 之差的绝对值都很小这个事件发生的概率等于 1. 但是格列汶科定理还未阐明统计量 D_n 服从什么分布或以什么分布为其极限分布. 柯尔莫哥洛夫定理和斯米尔诺夫定理. 找到了统计量 D_n 的极限分布表达式,这些定理在分布函数的假设检验中有重要的作用(在此从略).

二、样本的数字特征

1. 样本矩

设 $F_n^*(x)$ 为总体 ξ 的经验分布函数,记

$$A_r=\int_{-\infty}^{\infty}x^r\mathrm{d}F_n^*(x)=\frac{1}{n}\sum_{i=1}^{n}\xi_i^r, \tag{6.2.5}$$

$$B_r=\int_{-\infty}^{\infty}(x-\bar{\xi})^r\mathrm{d}F_n^*(x)=\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^r. \tag{6.2.6}$$

称 A_r 及 B_r 分别为样本的 r 阶原点矩及 r 阶中心矩,其中 r 为正整数. 显然 $A_1=\bar{\xi}, B_2=S^2$. 由于 A_r 及 B_r 都是样本的不带未知参数的函数,因而都是统计量.

2. 协方差与相关系数

设 $(\xi_1,\eta_1),(\xi_2,\eta_2),\cdots,(\xi_n,\eta_n)$ 为二维总体 (ξ,η) 的样本,记

$$\bar{\xi}=\frac{1}{n}\sum_{i=1}^{n}\xi_i,\qquad S_1^2=\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2;$$

$$\bar{\eta}=\frac{1}{n}\sum_{i=1}^{n}\eta_i,\qquad S_2^2=\frac{1}{n}\sum_{i=1}^{n}(\eta_i-\bar{\eta})^2. \tag{6.2.7}$$

$$S_{12}=\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\bar{\xi})(\eta_i-\bar{\eta}),$$

称统计量 S_{12} 为样本的协方差,称统计量

专业课辅导+真题资料+微信公众号：432应统书院

$$R=\frac{S_{12}}{S_1 \cdot S_2} \tag{6.2.8}$$

为样本的相关系数.

三、样本数字特征的分布

我们在这一段叙述样本的平均值、极值与极差的分布，样本方差 S^2 的分布在下一节中叙述.

1. 样本平均值的分布

设 $\varphi(t)$ 为总体 ξ 的特征函数，$\xi_1,\xi_2,\cdots,\xi_n$ 为总体 ξ 的样本，则样本平均值 $\bar{\xi}$ 的特征函数为：

$$\varphi_1(t)=[\varphi_0(t)]^n=\left[\varphi\left(\frac{t}{n}\right)\right]^n, \tag{6.2.9}$$

其中 $\varphi_0(t)=\varphi\left(\frac{t}{n}\right)$，它是 $\frac{1}{n}\xi$ 的特征函数.

例 6.2.1 设总体 ξ 服从正态 $N(a,\sigma)$，求样本平均值 $\bar{\xi}$ 的分布.

解 因 ξ 的特征函数为

$$\varphi(t)=\exp\left[-\frac{\sigma^2}{2}t^2+\mathrm{j}at\right],$$

则 $\bar{\xi}$ 的特征函数为

$$\begin{aligned}\varphi_1(t)&=\left\{\exp\left[-\frac{\sigma^2}{2}\left(\frac{t}{n}\right)^2+ja\frac{t}{n}\right]\right\}^n\\&=\exp\left[-\frac{t^2}{2}\left(\frac{\sigma}{\sqrt{n}}\right)^2+\mathrm{j}at\right],\end{aligned}$$

即 $\bar{\xi}$ 服从正态 $N\left(a,\frac{\sigma}{\sqrt{n}}\right)$.

例 6.2.2 设总体 ξ 服从具有参数 λ 的泊松分布，求样本平均值 $\bar{\xi}$ 的分布.

解 因 ξ 的特征函数为

$$\varphi(t)=\exp[\lambda(\mathrm{e}^{\mathrm{j}t}-1)],$$

专业课辅导+真题资料+微信公众号：432应统书院

则$\bar{\xi}$的特征函数为

$$\varphi_1(t)=\exp[n\lambda(e^{j\frac{t}{n}}-1)],$$

即$\bar{\xi}$的概率函数为

$$P\left\{\bar{\xi}=\frac{k}{n}\right\}=e^{-n\lambda}\frac{(n\lambda)^k}{k!},k=0,1,2,\cdots.$$

右端是参数为 $n\lambda$ 的泊松分布.

例 6.2.3 设总体 ξ 服从具有参数为 λ 的指数分布,求样本平均值$\bar{\xi}$的分布.

ξ 的密度函数为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\lambda>0;\\ 0, & x\leqslant 0.\end{cases}$$

解 ξ 的特征函数为

$$\varphi(t)=\frac{\lambda}{\lambda-jt}=\left(1-j\frac{t}{\lambda}\right)^{-1},\lambda>0,$$

则$\bar{\xi}$的特征函数为

$$\varphi_1(t)=\left(1-j\frac{t}{n\lambda}\right)^{-n}.$$

可见样本平均$\bar{\xi}$服从具有参数 $\alpha=n-1,\beta=\frac{1}{n\lambda}$的 Γ-分布,记作 $\Gamma\left(n-1,\frac{1}{n\lambda}\right)$.

例 6.2.4 设总体 ξ 服从参数为 p 的二项分布,$0<p<1$,N 为正整数,概率函数为:

$$P\{\xi=k\}=C_N^k p^k(1-p)^{N-k},\qquad k=0,1,\cdots,N,$$

求样本平均值$\bar{\xi}$的极限分布.

解 因 $E(\xi)=Np$, $\qquad D(\xi)=Np(1-p)\leqslant\frac{N}{4}$,

则统计量

专业课辅导+真题资料+微信公众号：432应统书院

$$\eta_n=\frac{\sum_{i=1}^{n}\xi_i-E\left(\sum_{i=1}^{n}\xi_i\right)}{\sqrt{D\left(\sum_{i=1}^{n}\xi_i\right)}}=\frac{n\bar{\xi}-nNp}{\sqrt{nNp(1-p)}}$$

$$=\frac{\bar{\xi}-Np}{\sqrt{Np(1-p)/n}}$$

有极限分布 $N(0,1)$，称$\bar{\xi}$有渐近正态分布

$$N\left(Np,\sqrt{\frac{Np(1-p)}{n}}\right).$$

*2. 极值的分布

极值分布在水文、气象、地震等预报问题中，有其重要的作用. 在大型工程的设计中，我们需要对一些带有严重破坏性的自然灾害进行必要的估计与预测. 如在建造桥梁时，为了防止洪水冲塌桥梁这类事故发生，设计时就必须预先考虑到在使用期间该河流可能爆发的最高水位；在建水库、河堤时也有这样的问题. 又如，在建造高大建筑物时，也要考虑到今后若干年内的最高水位、最大风压、最大的地震震级等. 考察了解这些随机变量的概率分布，就是极值的分布问题.

设总体 ξ 的分布函数为 $F(x)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，样本中的最大项 $\xi_n^*=\max\{\xi_1,\xi_2,\cdots,\xi_n\}$ 的分布函数记为 $F_n(u)$，则有：

$$F_n(u)=P\{\xi_n^*<u\}$$

$$=P\{\xi_1<u,\xi_2<u,\cdots,\xi_n<u\}=[F(u)]^n. \tag{6.2.10}$$

样本中的最小项 $\xi_1^*=\min\{\xi_1,\xi_2,\cdots,\xi_n\}$ 的分布函数记为 $F_1(v)$，由于

$$P\{\xi_1^*\geqslant v\}=P\{\xi_1\geqslant v,\xi_2\geqslant v,\cdots,\xi_n\geqslant v\}$$

$$=[P\{\xi\geqslant v\}]^n=[1-F(v)]^n,$$

所以

$$F_1(v)=1-[1-F(v)]^n. \tag{6.2.11}$$

专业课辅导+真题资料+微信公众号：432应统书院

如果总体 ξ 有分布密度函数 $f(x)$，则 ξ_n^* 及 ξ_1^* 的分布密度函数分别为

$$f_n(u)=nf(u)[F(u)]^{n-1}, \tag{6.2.12}$$

$$f_1(v)=nf(v)[1-F(v)]^{n-1}. \tag{6.2.13}$$

一般地，若 $\xi_1^*<\xi_2^*<\cdots<\xi_n^*$ 为样本 $\xi_1,\xi_2,\cdots,\xi_n$ 的顺序统计量，由小到大的第 k 位统计量为 ξ_k^*，其分布函数记为 $F_k(y)$，$k=1,2,\cdots,n$. 我们用 $P\{y\leqslant\xi_k^*<y+\Delta y\}$ 表示事件"在样本 $\xi_1,\xi_2,\cdots,\xi_n$ 中有一个落在 $[y,y+\Delta y)$，$k-1$ 个落在 $(-\infty,y)$，其余 $n-k$ 个落在 $[y+\Delta y,\infty)$"的概率，于是

$$\begin{aligned}\Delta F_k(y)&=P\{y\leqslant\xi_k^*<y+\Delta y\}\\&=C_n^1\Delta F(y)\cdot C_{n-1}^{k-1}[F(y)]^{k-1}\cdot[1-F(y+\Delta y)]^{n-k},\end{aligned}$$

因此

$$F_k(y)=nC_{n-1}^{k-1}\int_{-\infty}^{y}[F(x)]^{k-1}\cdot[1-F(x)]^{n-k}\mathrm{d}F(x). \tag{6.2.14}$$

特别地有

$$F_1(y)=1-[1-F(y)]^n,$$

$$F_n(y)=[F(y)]^n.$$

*3. 极差的分布

设总体 ξ 的分布函数为 $F(x)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，样本极差 $D_n^*=\xi_n^*-\xi_1^*$ 的分布函数记为 $F_{D_n^*}(y)$，其中 $y\geqslant0$. 欲求 $F_{D_n^*}(y)$，一般地，可先导出顺序统计量 $(\xi_1^*,\xi_2^*,\cdots,\xi_n^*)$ 的联合分布，或者 (ξ_1^*,ξ_n^*) 的联合分布，从而导出样本极差 D_n^* 的分布函数. 现在，我们来直接导出 D_n^* 的分布函数 $F_{D_n^*}(x)$.

现对连续型随机变量 ξ 作讨论. 设总体 ξ 有密度函数 $f(x)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，用 A 表示"样本 $\xi_1,\xi_2,\cdots,\xi_n$ 中有一个落在区间 $(v,v+\mathrm{d}v)$，而另 $n-1$ 个同时落在区间 $(v,v+y)$"这个事件，那么事件 A 发生的概率为

$$C_n^1f(v)\mathrm{d}v\cdot[F(v+y)-F(v)]^{n-1},$$

专业课辅导+真题资料+微信公众号：432应统书院

其中$(v,v+y)$为样本观察值的范围，y为极差D_n^*的观察值的上界. 于是

$$F_{D_n^*}(y)=P\{D_n^*<y\}$$

$$=\int_{-\infty}^{\infty} n[F(v+y)-F(v)]^{n-1}f(v)\,\mathrm{d}v. \tag{6.2.15}$$

其密度函数为

$$f_{D_n^*}(y)=\frac{\mathrm{d}F_{D_n^*}(y)}{\mathrm{d}y}$$

$$=n(n-1)\int_{-\infty}^{\infty}\left[\int_{v}^{v+y}f(x)\,\mathrm{d}x\right]^{n-2}\cdot f(v+y)f(v)\,\mathrm{d}v. \tag{6.2.16}$$

样本极差D_n^*的数学期望及方差分别为

$$c_n=E(D_n^*)=\int_0^{\infty} yf_{D_n^*}(y)\,\mathrm{d}y, \tag{6.2.17}$$

$$v_n^2=D(D_n^*)=\int_0^{\infty}(y-c_n)^2f_{D_n^*}(y)\,\mathrm{d}y. \tag{6.2.18}$$

例 6.2.5 设总体ξ服从参数为λ的指数分布，其密度函数及分布函数分别为

$$f(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & 0<x<\infty,\lambda>0,\\ 0, & \text{其他},\end{cases}$$

$$F(x)=\begin{cases}1-\mathrm{e}^{-\lambda x}, & 0<x<\infty,\lambda>0,\\ 0, & \text{其他}.\end{cases}$$

求样本$\xi_1,\xi_2,\cdots,\xi_n$的极值与极差的分布.

解 由(6.2.12)、(6.2.13)及(6.2.16)分别求出

$$f_n(u)=\begin{cases}n\lambda\mathrm{e}^{-\lambda u}[1-\mathrm{e}^{-\lambda u}]^{n-1}, & 0<u<\infty,\\ 0, & \text{其他};\end{cases} \tag{6.2.19}$$

$$f_1(v)=\begin{cases}n\lambda\mathrm{e}^{-n\lambda v}, & 0<v<\infty,\\ 0, & \text{其他};\end{cases} \tag{6.2.20}$$

$$f_{D_n^*}(y)=\begin{cases}(n-1)[1-\mathrm{e}^{-\lambda y}]^{n-2}\lambda\mathrm{e}^{-\lambda y}, & 0<y<\infty,\\ 0, & \text{其他}.\end{cases} \tag{6.2.21}$$

专业课辅导+真题资料+微信公众号：432应统书院

可见，样本的最小值 ξ_1^* 服从参数为 $n\lambda$ 的指数分布.

在一般情况下，样本的极值与极差这些统计量的分布函数并不是那么容易求出来的. 例如当总体 ξ 服从正态 $N(a,\sigma)$ 或总体 ξ 服从分布 $\Gamma(\alpha,\beta)$，它们的样本极值的分布函数表达式都相当复杂. 格涅坚科在他的 1943 年的文章中具体地讨论了样本最大值 ξ_n^* 的极限分布的三种类型，并找到了向这些极限分布收敛的充要条件（在此从略）.

设总体 ξ 服从正态 $N(0,1)$，即有密度函数

$$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$$

对于样本容量的不同 n 值，利用(6.2.17)及(6.2.18)两式，可算出 c_n 及 v_n^2 的数值，如表 6.2.1 所示.

表 6.2.1

n	c_n	$1/c_n$	v_n
2	1.12838	0.8862	0.853
3	1.69257	0.5908	0.888
4	2.05875	0.4857	0.880
5	2.32593	0.4299	0.864
6	2.53441	0.3946	0.848
7	2.70436	0.3698	0.833
8	2.84720	0.3512	0.820
9	2.97003	0.3367	0.808
10	3.07751	0.3249	0.797

若总体 ξ 服从正态 $N(a,\sigma)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，极差为

$$D_n^*=\max_{1\leqslant i\leqslant n}\xi_i-\min_{1\leqslant i\leqslant n}\xi_i. \tag{6.2.22}$$

专业课辅导+真题资料+微信公众号：432应统书院

作标准化变换

$$\xi'=\frac{\xi-a}{\sigma},$$

$$\xi_i'=\frac{\xi_i-a}{\sigma},\quad i=1,2,\cdots,n,$$

则 ξ' 服从正态 $N(0,1)$，$\xi_1',\xi_2',\cdots,\xi_n'$ 为其样本，并有

$$\begin{aligned}\frac{D_n^*}{\sigma}&=\max_{1\leqslant i\leqslant n}\left(\frac{\xi_i-a}{\sigma}\right)-\min_{1\leqslant i\leqslant n}\left(\frac{\xi_i-a}{\sigma}\right)\\&=\max_{1\leqslant i\leqslant n}\xi_i'-\min_{1\leqslant i\leqslant n}\xi_i'=D_n^{*\prime},\end{aligned}$$

于是

$$E\left(\frac{D_n^*}{\sigma}\right)=E(D_n^{*\prime})=c_n,$$

其中 c_n 是在总体 ξ' 服从正态 $N(0,1)$ 的假定下，由(6.2.17)式计算得到的(它在表(6.2.1)中可查得到). 于是

$$E\left(\frac{D_n^*}{c_n}\right)=\sigma,D(D_n^*)=v_n^2\sigma^2. \tag{6.2.23}$$

在第七章的§7.2中将看到，统计量 D_n^*/c_n 是 σ 的无偏估计量.

§6.3 抽样分布定理

现在，我们来讨论几个重要的有关抽样分布的定理，这些定理在估计理论、假设检验及方差分析等数理统计学的基本内容中都有重要的作用. 这些定理，都是在总体 ξ 服从正态 $N(a,\sigma)$ 的基本假定下而得到的.

定理 6.3.1 设总体 ξ 服从正态 $N(a,\sigma)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，样本的平均值与方差分别记为 $\bar{\xi}$ 与 S^2，则 $\bar{\xi}$ 服从正态 $N\left(a,\frac{\sigma}{\sqrt{n}}\right)$，$\frac{nS^2}{\sigma^2}$ 服从自由度为 $n-1$ 的 χ^2-分布，简记作 $\frac{nS^2}{\sigma^2}\sim\chi^2_{(n-1)}$，而且 $\bar{\xi}$ 与 S^2 相互独立.

证 先就 $a=0$ 的情形证明所需要的结论. 今对样本作正交线

专业课辅导+真题资料+微信公众号：432应统书院

性变换如下：

$$\begin{cases} \eta_1=\dfrac{1}{\sqrt{1\cdot 2}}(\xi_1-\xi_2); \\ \eta_2=\dfrac{1}{\sqrt{2\cdot 3}}(\xi_1+\xi_2-2\xi_3); \\ \quad\vdots \\ \eta_{n-1}=\dfrac{1}{\sqrt{(n-1)\cdot n}}[\xi_1+\xi_2+\cdots+\xi_{n-1}-(n-1)\xi_n]; \\ \eta_n=\dfrac{1}{\sqrt{n}}(\xi_1+\xi_2+\cdots+\xi_n). \end{cases} \tag{6.3.1}$$

我们来证 $\eta_1,\eta_2,\cdots,\eta_n$ 相互独立且有相同正态 $N(0,\sigma)$.

(1) 由于正交变换保持长度不变,所以

$$\sum_{i=1}^{n}\eta_i^2=\sum_{i=1}^{n}\xi_i^2. \tag{6.3.2}$$

(2) 注意到$(\xi_1,\xi_2,\cdots,\xi_n)$的联合分布是 n 维正态的,其联合密度函数为

$$\begin{aligned} f(x_1,x_2,\cdots,x_n)&=\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n\prod_{i=1}^{n}\mathrm{e}^{-\frac{x_i^2}{2\sigma^2}} \\ &=\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n\mathrm{e}^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{n}x_i^2}. \end{aligned} \tag{6.3.3}$$

由于(6.3.1)变换的雅可比行列式的绝对值为 1,所以$(\eta_1,\eta_2,\cdots,\eta_n)$的联合密度函数为

$$g(y_1,y_2,\cdots,y_n)=\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n\mathrm{e}^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{n}y_i^2}. \tag{6.3.4}$$

于是$(\eta_1,\eta_2,\cdots,\eta_n)$的联合分布是 n 维正态分布. 我们知道,联合正态分布的边缘分布也是正态的,且在正态情形,不相关性与独立性是等价的,因而 $\eta_1,\eta_2,\cdots,\eta_n$ 是相互独立且有相同正态$N(0,\sigma)$的随机变量.

(3) 在(6.3.1)式中,$\eta_n=\dfrac{n}{\sqrt{n}}\bar{\xi}=\sqrt{n}\bar{\xi}$,即

专业课辅导+真题资料+微信公众号：432应统书院

$$\eta_n^2 = n\bar{\xi}^2,$$

另外
$$nS^2 = \sum_{i=1}^{n}(\xi_i - \bar{\xi})^2 = \sum_{i=1}^{n}\xi_i^2 - n\bar{\xi}^2$$
$$= \sum_{i=1}^{n}\eta_i^2 - \eta_n^2 = \sum_{i=1}^{n-1}\eta_i^2,$$

于是
$$\frac{nS^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n-1}\eta_i^2 \sim \chi^2_{(n-1)}, \tag{6.3.5}$$

而且 nS^2 与 η_n 相互独立，也就是 nS^2 与 $\bar{\xi}$ 相互独立，这就证明了我们所需的结论.

除去 $a=0$ 这个假定，定理的结论仍然成立. 事实上，我们可作变换：
$$\zeta = \xi - a, \zeta_i = \xi_i - a, i = 1, 2, \cdots, n. \tag{6.3.6}$$

由 ξ 服从正态 $N(a,\sigma)$ 知，ζ 服从正态 $N(0,\sigma)$. 另外
$$\bar{\zeta} = \bar{\xi} - a,$$
$$\frac{1}{n}\sum_{i=1}^{n}(\zeta_i - \bar{\zeta})^2 = \frac{1}{n}\sum_{i=1}^{n}(\xi_i - \bar{\xi})^2 = S^2,$$

对 ζ 应用上述证明的结果可知，$\sum_{i=1}^{n}(\zeta_i - \bar{\zeta})^2$ 与 $\bar{\zeta}$ 相互独立，从而 nS^2 与 $\bar{\xi}$ 相互独立，而且 $\frac{nS^2}{\sigma^2}$ 服从自由度为 $n-1$ 的 χ^2-分布，即 $\frac{nS^2}{\sigma^2} \sim \chi^2_{(n-1)}$.

推论 设总体 ξ 服从正态 $N(a,\sigma)$，$\xi_1, \xi_2, \cdots, \xi_n$ 为样本，则
$$T = \sqrt{n-1}\frac{\bar{\xi} - a}{S} \tag{6.3.7}$$

服从自由度为 $n-1$ 的 t-分布，记作 $T \sim t_{(n-1)}$.

证 由定理 6.3.1 知，下述两个随机变量
$$\zeta = \frac{\bar{\xi} - a}{\sigma/\sqrt{n}}, \tag{6.3.8}$$

专业课辅导+真题资料+微信公众号：432应统书院

$$\eta=\frac{nS^2}{\sigma^2}$$

相互独立，而且 ζ 服从正态 $N(0,1)$，η 服从自由度为 $n-1$ 的 χ^2-分布，由第二章的 §2.4 知，

$$\frac{\zeta}{\sqrt{\eta/(n-1)}}=\sqrt{n-1}\frac{\bar{\xi}-a}{S}$$

服从自由度为 $n-1$ 的 t-分布，记作 $\sqrt{n-1}\frac{\bar{\xi}-a}{S}\sim t_{(n-1)}$.

定理 6.3.2 设总体 ξ 服从正态 $N(a_1,\sigma)$，$\xi_1,\xi_2,\cdots,\xi_{n_1}$ 为其样本，总体 η 服从正态 $N(a_2,\sigma)$，$\eta_1,\eta_2,\cdots,\eta_{n_2}$ 为其样本，而且这两个样本是相互独立的①. 记

$$\bar{\xi}=\frac{1}{n_1}\sum_{i=1}^{n_1}\xi_i,\qquad S_1^2=\frac{1}{n_1}\sum_{i=1}^{n_1}(\xi_i-\bar{\xi})^2;$$

$$\bar{\eta}=\frac{1}{n_2}\sum_{j=1}^{n_2}\eta_j,\qquad S_2^2=\frac{1}{n_2}\sum_{j=1}^{n_2}(\eta_j-\bar{\eta})^2,$$

则有

$$(1)\ \frac{(n_2-1)n_1S_1^2}{(n_1-1)n_2S_2^2}\sim F_{(n_1-1,n_2-1)},\tag{6.3.9}$$

即左端的统计量服从第一自由度为 n_1-1、第二自由度为 n_2-1 的 F-分布.

$$(2)\ \sqrt{\frac{n_1n_2(n_1+n_2-2)}{n_1+n_2}}\cdot\frac{(\bar{\xi}-\bar{\eta})-(a_1-a_2)}{\sqrt{n_1S_1^2+n_2S_2^2}}\sim t_{(n_1+n_2-2)},\tag{6.3.10}$$

即左端的表达式服从自由度为 n_1+n_2-2 的 t-分布.

证 先证(6.3.9)式是正确的. 由定理 6.3.1 知

① 随机向量 $(\xi_1,\xi_2,\cdots,\xi_{n_1})$ 与随机向量 $(\eta_1,\eta_2,\cdots,\eta_{n_2})$ 之间是相互独立的.

专业课辅导+真题资料+微信公众号：432应统书院

$$\frac{n_1 S_1^2}{\sigma^2} \sim \chi^2_{(n_1-1)},$$

$$\frac{n_2 S_2^2}{\sigma^2} \sim \chi^2_{(n_2-1)}.$$

由于两个样本是相互独立的，因而$\frac{n_1 S_1^2}{\sigma^2}$与$\frac{n_2 S_2^2}{\sigma^2}$也相互独立. 由第二章§2.4知

$$\frac{\left(\frac{n_1 S_1^2}{\sigma^2}\right)\Big/(n_1-1)}{\left(\frac{n_2 S_2^2}{\sigma^2}\right)\Big/(n_2-1)} = \frac{(n_2-1)n_1 S_1^2}{(n_1-1)n_2 S_2^2} \sim F_{(n_1-1,n_2-1)},$$

即左端的统计量服从第一自由度为 n_1-1，第二自由度为 n_2-1 的 F-分布.

现证(6.3.10)式是正确的. 由于两个样本是相互独立的，因而统计量$\bar{\xi}$与 $\bar{\eta}$ 相互独立，而且

$$E(\bar{\xi}-\bar{\eta}) = a_1 - a_2,$$

$$D(\bar{\xi}-\bar{\eta}) = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}.$$

容易证明

$$(\bar{\xi}-a_1)-(\bar{\eta}-a_2) = (\bar{\xi}-\bar{\eta})-(a_1-a_2)$$

服从正态 $N\left(0, \sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$. 由于两个样本是相互独立的，且$\frac{n_1 S_1^2}{\sigma^2} \sim \chi^2_{(n_1-1)}$，$\frac{n_2 S_2^2}{\sigma^2} \sim \chi^2_{(n_2-1)}$. 由第二章§2.4知，

$$\frac{n_1 S_1^2}{\sigma^2} + \frac{n_2 S_2^2}{\sigma^2} = \frac{n_1 S_1^2 + n_2 S_2^2}{\sigma^2} \sim \chi^2_{(n_1+n_2-2)}.$$

如果能证明$(\bar{\xi}-\bar{\eta})-(a_1-a_2)$与$(n_1 S_1^2 + n_2 S_2^2)/\sigma^2$ 是相互独立的，则由第二章§2.4知，

专业课辅导+真题资料+微信公众号：432应统书院

$$\frac{[(\bar{\xi}-\bar{\eta})-(a_1-a_2)]\Big/\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}{\sqrt{\frac{n_1S_1^2+n_2S_2^2}{\sigma^2}\Big/(n_1+n_2-2)}}$$

$$=\sqrt{\frac{n_1n_2(n_1+n_2-2)}{n_1+n_2}}\cdot\frac{(\bar{\xi}-\bar{\eta})-(a_1-a_2)}{\sqrt{n_1S_1^2+n_2S_2^2}}\sim t_{(n_1+n_2-2)}.$$

下面证明$(\bar{\xi}-a_1)-(\bar{\eta}-a_2)$与$(n_1S_1^2+n_2S_2^2)/\sigma^2$是相互独立的. 由定理6.3.1知道,$(\bar{\xi}-a_1)$与$n_1S_1^2/\sigma^2$相互独立,$(\bar{\eta}-a_2)$与$n_2S_2^2/\sigma^2$相互独立. 由于两个样本是相互独立的,因而$(\bar{\xi}-a_1)$与$n_2S_2^2/\sigma^2$相互独立. $(\bar{\eta}-a_2)$与$n_1S_1^2/\sigma^2$相互独立,于是$(\bar{\xi}-a_1)-(\bar{\eta}-a_2)$与$n_1S_1^2/\sigma^2$相互独立,并与$n_2S_2^2/\sigma^2$相互独立,因而与$(n_1S_1^2+n_2S_2^2)/\sigma^2$相互独立.

利用定理6.3.3,对这个结论可作严格的证明.

定理6.3.3(柯赫伦分解定理) 设$\xi_1,\xi_2,\cdots,\xi_n$为n个相互独立的随机变量,每一个$\xi_i(i=1,2,\cdots,n)$都服从正态$N(0,1)$. 如果

$$Q=\sum_{i=1}^{n}\xi_i^2=\sum_{l=1}^{k}Q_l,$$

其中Q_l是秩为n_l的$\xi_1,\xi_2,\cdots,\xi_n$的二次型,则下述结论

(1) $Q_1,Q_2,\cdots,Q_k$相互独立,

(2) $Q_l\sim\chi^2_{n_l},\qquad l=1,2,\cdots,k$

成立的充要条件为$\sum_{l=1}^{k}n_l=n$.

* **证** 必要性 设$Q_1,Q_2,\cdots,Q_k$相互独立且分别服从自由度为$n_1,n_2,\cdots,n_k$的χ^2-分布,即

$$Q_l\sim\chi^2_{n_l}\qquad(l=1,2,\cdots,k),$$

则由定理2.4.2知

专业课辅导+真题资料+微信公众号：432应统书院

$$\sum_{l=1}^{k} Q_l \sim \chi^2_{\sum_{l=1}^{k} n_l}.$$

由定理的假设知

$$Q=\sum_{i=1}^{n} \xi_i^2 \sim \chi_n^2,$$

且

$$Q=\sum_{i=1}^{n} \xi_i^2=\sum_{l=1}^{k} Q_l, \tag{6.3.11}$$

所以

$$n=\sum_{l=1}^{k} n_l.$$

充分性　若 $n=\sum_{l=1}^{k} n_l$，我们来证(1)及(2)成立. 由于 Q_l 是秩为 n_l 的 $\xi_1,\xi_2,\cdots,\xi_n$ 的二次型，可以证明(留作练习) Q_l 能够表成下式：

$$Q_l=\sum_{j=1}^{n_l} b_{lj}\eta_{lj}^2, \qquad b_{lj}=1 \text{ 或 } -1, \tag{6.3.12}$$

其中 η_{lj} 是 $\xi_1,\xi_2,\cdots,\xi_n$ 的线性函数($j=1,2,\cdots,n_l$). 于是

$$Q=\sum_{i=1}^{n} \xi_i^2=\sum_{l=1}^{k}\sum_{j=1}^{n_l} b_{lj}\eta_{lj}^2. \tag{6.3.13}$$

记

$$\boldsymbol{x}'=(\xi_1,\xi_2,\cdots,\xi_n),$$

$$\boldsymbol{y}'=(\eta_{11},\cdots,\eta_{1n_1},\cdots,\eta_{k1},\cdots,\eta_{kn_k}),$$

则(6.3.13)式可以写为矩阵形式：

$$\boldsymbol{x}'\boldsymbol{x}=\boldsymbol{y}'\boldsymbol{B}\boldsymbol{y}.$$

其中 $\boldsymbol{B}$ 是对角线元素为±1的 $n\times n$ 的对角矩阵，$n=\sum_{l=1}^{k} n_l$. 作变换

$$\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x},$$

它是由正定二次型 $\boldsymbol{x}'\boldsymbol{x}$ 变到 $\boldsymbol{y}'\boldsymbol{B}\boldsymbol{y}$ 的变换式. 由

$$\boldsymbol{x}'\boldsymbol{x}=\boldsymbol{y}'\boldsymbol{B}\boldsymbol{y}=\boldsymbol{x}'\boldsymbol{A}'\boldsymbol{B}\boldsymbol{A}\boldsymbol{x}$$

得到 $\boldsymbol{A}'\boldsymbol{B}\boldsymbol{A}=\boldsymbol{I}$，由此推知 $\boldsymbol{A}$ 是非奇异的，于是 $\boldsymbol{A}\boldsymbol{A}'$ 是正定的. 因此，

专业课辅导+真题资料+微信公众号：432应统书院

$\boldsymbol{B}=\boldsymbol{I}$ 及 $\boldsymbol{A}'\boldsymbol{A}=\boldsymbol{I}$. 从而 $\boldsymbol{A}$ 为正交矩阵. 由定理 6.3.1 知，η_{lj} 之间相互独立而且都服从正态 $N(0,1)$ $(l=1,2,\cdots,k;j=1,2,\cdots,n_l)$. 所以 $Q_l\sim\chi^2_{n_l}$，而且 $Q_1,Q_2,\cdots,Q_k$ 是相互独立的.

推论 设 $\xi_1,\xi_2,\cdots,\xi_k$ 为 k 个相互独立的且每个都服从正态 $N(0,1)$ 的随机变量，今从这 k 个总体中分别抽取容量为 n_i 的样本 $\xi_{i1},\xi_{i2},\cdots,\xi_{in_i}(i=1,2,\cdots,k)$，记

$$\bar{\xi}_i=\frac{1}{n_i}\sum_{j=1}^{n_i}\xi_{ij},$$

$$S_i^2=\frac{1}{n_i}\sum_{j=1}^{n_i}(\xi_{ij}-\bar{\xi}_i)^2,$$

$$n=\sum_{i=1}^{k}n_i,$$

$$\bar{\xi}=\frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\xi_{ij}=\frac{1}{n}\sum_{i=1}^{k}n_i\bar{\xi}_i,$$

则

(1) $$S_{总}^2=\sum_{i=1}^{k}\sum_{j=1}^{n_i}(\xi_{ij}-\bar{\xi})^2=\sum_{i=1}^{k}n_iS_i^2+\sum_{i=1}^{k}n_i(\bar{\xi}_i-\bar{\xi})^2;\tag{6.3.14}$$

(2) 记 $Q_e=\sum_{i=1}^{k}n_iS_i^2$,

$$U=\sum_{i=1}^{k}n_i(\bar{\xi}_i-\bar{\xi})^2,$$

那么

$$\frac{U/(k-1)}{Q_e/(n-k)}\sim F_{(k-1,n-k)}.\tag{6.3.15}$$

*证 先验证恒等式(6.3.14)，称(6.3.14)式为离差平方和的分解式.

$$S_{总}^2=\sum_{i=1}^{k}\sum_{j=1}^{n_i}(\xi_{ij}-\bar{\xi})^2$$

专业课辅导+真题资料+微信公众号：432应统书院

$$
\begin{aligned}
&= \sum_{i=1}^{k} \sum_{j=1}^{n_i} [\xi_{ij}-\bar{\xi}_i+\bar{\xi}_i-\bar{\xi}]^2 \\
&= \sum_{i=1}^{k} \sum_{j=1}^{n_i} [(\bar{\xi}_i-\bar{\xi})^2+2(\bar{\xi}_i-\bar{\xi})(\xi_{ij}-\bar{\xi}_i)+(\xi_{ij}-\bar{\xi}_i)^2] \\
&= \sum_{i=1}^{k} n_i(\bar{\xi}_i-\bar{\xi})^2+\sum_{i=1}^{k} n_i S_i^2 \\
&= U+Q_e,
\end{aligned}
$$

其中

$$
\begin{aligned}
&\sum_{i=1}^{k} \sum_{j=1}^{n_i} (\bar{\xi}_i-\bar{\xi})(\xi_{ij}-\bar{\xi}_i) \\
&= \sum_{i=1}^{k} (\bar{\xi}_i-\bar{\xi})\left[\sum_{j=1}^{n_i} (\xi_{ij}-\bar{\xi}_i)\right]=0.
\end{aligned}
$$

现证(2). 由定理 6.3.1 可知，$S_{总}^2 \sim \chi_{(n-1)}^2$. 我们知道，在 Q_e 中有 k 个独立的线性关系式，即

$$\sum_{j=1}^{n_i} (\xi_{ij}-\bar{\xi}_i)=0 \quad (i=1,2,\cdots,k),$$

Q_e 的自由度为 $\sum_{i=1}^{k} n_i-k=n-k$. 在 U 中有一个线性关系式，即

$$\sum_{i=1}^{k} n_i(\bar{\xi}_i-\bar{\xi})=0.$$

因而 U 的自由度为 $k-1$. 于是由定理 6.3.3 知，Q_e 与 U 相互独立且有 $Q_e \sim \chi_{(n-k)}^2$ 及 $U \sim \chi_{(k-1)}^2$，由定理 2.4.5 知

$$\frac{U/(k-1)}{Q_e/(n-k)} \sim F_{(k-1,n-k)}.$$

习　题

1. 设总体 ξ 服从正态 $N(5,2)$，$\xi_1,\cdots,\xi_9$ 为其样本，试求样本的平均值 $\bar{\xi}$ 大于 8 的概率。

专业课辅导+真题资料+微信公众号：432应统书院

2. 设总体 ξ 在 $\left(\theta-\frac{1}{2},\theta+\frac{1}{2}\right)$ 上服从均匀分布，$\xi_1,\xi_2,\cdots,\xi_9$ 为其样本，试指出 $\bar{\xi}-\frac{1}{2}$，$\bar{\xi}+\theta$ 中哪个是统计量.

3. 设总体 ξ 服从正态 $N(0,\sigma)$，ξ_1,ξ_2,ξ_3,ξ_4 为其样本，试问 $\eta=(\xi_1-\xi_2)^2/(\xi_3+\xi_4)^2$ 服从什么分布.

4. 设总体 ξ 服从正态 $N(1,2)$，ξ_1,ξ_2,ξ_3,ξ_4 为其样本，记 $\eta=k\left[\sum_{i=1}^{4}\xi_i-4\right]^2$，试问 k 取何值，使得 η 服从 χ_m^2 分布，自由度 m 是何值.

5. 设总体 ξ 服从正态 $N(3,2)$，$\xi_1,\xi_2,\cdots,\xi_{16}$ 为其样本，$\bar{\xi}$ 及 S^2 分别为样本的均值与方差，试建立 t-分布的统计量.

6. 设总体 ξ 服从正态 $N(5,6)$，n 及 $\bar{\xi}$ 分别为样本容量及样本均值，试问 n 至少应取多大，使得 $\bar{\xi}$ 位于区间 $(3,7)$ 的概率不小于 0.90.

7. 设总体 ξ 服从具有参数 $\lambda>0$ 的指数分布，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，$\bar{\xi}$ 为样本均值：

(1) 试求 $\eta=2n\lambda\bar{\xi}$ 的分布；

(2) 若 $n=1$，试问 $P\{\eta>6\}$ 是何值(查表).

8. 设总体 ξ 服从正态 $N(12,2)$，今抽取容量为 5 的样本 $\xi_1,\xi_2,\cdots,\xi_5$，试问：

(1) 样本的平均值 $\bar{\xi}$ 大于 13 的概率是多少？

(2) 样本的极小值小于 10 的概率是多少？

(3) 样本的极大值大于 15 的概率是多少？

9. 设电子元件的寿命(时数) ξ 服从以 $\lambda=0.001\,5$ 为参数的指数分布，即有密度函数 $f(x)=0.001\,5e^{-0.001\,5x}$，$x>0$. 今测试 6 个元件，并记录下它们各自失效的时间(单位：h). 试问：

(1) 至 800 小时时没有一个元件失效的概率是多少？

(2) 至 3 000 小时时所有元件都失效的概率是多少？

10. 设总体 ξ 服从正态 $N(20,\sqrt{3})$，今从中抽取容量为 10 及 15 的两个独立样本，试问这两个样本的平均值之差的绝对值大于 0.3 的概率是多少？

*11. 设总体 ξ 服从正态 $N(a,\sigma)$，ξ_1,ξ_2 为其样本，试求样本极差的分布，

专业课辅导+真题资料+微信公众号：432应统书院

极大值与极小值的分布.

* 12. 设总体 ξ 服从参数为 λ 的指数分布，ξ_1,ξ_2 为其样本，试求样本的极大值、极小值与极差的分布.

* 13. 设 $\xi_1,\xi_2,\cdots,\xi_n$ 是 n 个相互独立的且都是服从正态 $N(0,1)$ 的随机变量，$\xi_1,\xi_2,\cdots,\xi_n$ 到 $\eta_1,\eta_2,\cdots,\eta_n$ 的变换为正交变换. 试证 $\eta_1,\eta_2,\cdots,\eta_n$ 是 n 个相互独立的且都是服从正态 $N(0,1)$ 的随机变量.

14. 设总体 ξ 服从正态 $N(a,\sigma)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，$\bar{\xi}$ 及 S_n^2 分别为样本的平均值及方差. 又设 ξ_{n+1} 服从正态 $N(a,\sigma)$，且与 $\xi_1,\xi_2,\cdots,\xi_n$ 相互独立. 试求统计量

$$\eta=\frac{\xi_{n+1}-\bar{\xi}}{S_n}\sqrt{\frac{n-1}{n+1}}$$

的抽样分布.

15. 设 $\xi_1,\xi_2,\cdots,\xi_n$ 相互独立且分别服从正态 $N(a_i,\sigma_i)$，试证 $\eta=\sum_{i=1}^{n}c_i\xi_i$ 服从正态 $N\left(\sum_{i=1}^{n}c_ia_i,\sqrt{\sum_{i=1}^{n}c_i^2\sigma_i^2}\right)$.

* 16. 设总体 ξ 在 $\left(\theta-\frac{1}{2},\theta+\frac{1}{2}\right)$ 上服从均匀分布，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，$\xi_1^*,\xi_2^*,\cdots,\xi_n^*$ 为顺序统计量. 试求 ξ_1^*,ξ_n^* 及 (ξ_1^*,ξ_n^*) 的分布.

专业课辅导+真题资料+微信公众号：432应统书院

第七章　参数估计

估计理论是数理统计学的重要内容之一. 我们知道,服从泊松分布的随机变量,其概率分布 $P(x;\lambda)$ 由一个参数 $\lambda>0$ 确定;服从正态分布的随机变量,其概率分布 $N(a,\sigma)$ 由一对参数 a 及 $\sigma>0$ 确定. 就是说,对所要研究的随机变量 ξ,当它的概率分布的类型为已知时,还需确定分布函数中的参数是什么值,这样随机变量 ξ 的分布函数才能完全确定. 例如,纺织厂细纱机上的断头次数 ξ 服从泊松分布,如果要求知道每只纱锭在某一时间间隔内断头次数为 k 次的概率是多少,那么就要首先确定泊松分布中的参数 λ 取什么值,然后才能得出所要求的概率. 这就提出了参数的估计问题. 设总体 ξ 的分布函数为 $F(x;\theta)$,其中参数 θ 为未知,试问怎样由样本 $\xi_1,\xi_2,\cdots,\xi_n$ 提供的信息建立样本的函数,即统计量来对未知参数 θ 作出估计,并且估计量的"最佳"准则又怎样确定? 这类统计问题,称为参数的估计问题. 另外在有些实际问题中,事先并不知道随机变量 ξ 服从什么分布,而要对其数字特征,如数学期望 $E(\xi)$ 及方差 $D(\xi)$ 等作出估计. 例如,某砖瓦厂生产了一批青砖,为了评定这批青砖的质量情况,需要了解这批青砖的平均抗压强度. 一般地说抗压强度为一随机变量,记作 ξ,现在的问题是希望了解 $E(\xi)$ 是多大? 我们知道,随机变量 ξ 的数字特征同它的概率分布中的参数有一定关系,因而对数字特征的估计问题,也称之为参数的估计问题.

参数估计有点估计和区间估计两方面的问题. 参数点估计有许多方法,在这一章中讨论矩法、极大似然法,在第九章中将讨论最小二乘法估计,第十章介绍贝叶斯估计.

专业课辅导+真题资料+微信公众号：432应统书院

§7.1 矩法与极大似然法

设总体ξ的分布函数为$F(x;\theta)$，其中θ为未知参数，θ的取值范围记为Ω，Ω称为参数空间①. 例如，若随机变量ξ服从二项分布$b(x;p)$，p为参数，则$p\in\Omega=(0,1)$；若ξ服从泊松分布$P(x;\lambda)$，λ为参数，则$\lambda\in\Omega=(0,\infty)$；若$\xi$服从正态$N(a,\sigma)$，其中$a$及$\sigma$都是参数，则$\Omega=\{(a,\sigma):a\in\mathbf{R}_1=(-\infty,\infty),\sigma\in(0,\infty)\}$.

定义7.1.1（点估计） 设总体ξ的分布函数$F(x;\theta)$中参数θ未知，$\theta\in\Omega$，Ω为参数空间，今由样本$\xi_1,\xi_2,\cdots,\xi_n$建立统计量$T(\xi_1,\xi_2,\cdots,\xi_n)$，对于样本观察值$(x_1,x_2,\cdots,x_n)$，若将$T(x_1,x_2,\cdots,x_n)=t$作为$\theta$的估计值，则称$T(\xi_1,\xi_2,\cdots,\xi_n)$为$\theta$的估计量，通常记作$\hat{\theta}=T(\xi_1,\xi_2,\cdots,\xi_n)$. 建立一个这样的统计量$T(\xi_1,\xi_2,\cdots,\xi_n)$作为$\theta$的估计量，称之为参数$\theta$的点估计.

如果随机变量ξ的分布函数$F(x;\theta_1,\cdots,\theta_l)$中有$l$个不同的未知参数，则要由样本$\xi_1,\xi_2,\cdots,\xi_n$建立$l$个统计量作为这$l$个未知参数的估计量，分别记之为$\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_l$.

现在叙述求参数点估计的矩法与极大似然法.

一、矩法

皮尔逊所引入的矩法是较早提出的求参数点估计的方法. 我们从辛钦大数定律知道，若总体ξ的数学期望$E(\xi)$有限，则样本的平均值$\bar{\xi}$依概率收敛于$E(\xi)$. 这就启发我们想到，在利用样本所提供的信息来对总体ξ的分布函数中未知参数θ作估计时，可以用样本矩作为总体矩的估计.

定义7.1.2（矩法估计量） 设总体ξ的分布函数$F(x;\theta_1,\theta_2,\cdots,\theta_l)$中有$l$个未知参数$\theta_1,\theta_2,\cdots,\theta_l$，假定总体$\xi$的$l$阶原点绝对矩

① 在第一章中我们以Ω表示样本空间；在数理统计各章中，我们用Ω表示参数空间.

专业课辅导+真题资料+微信公众号：432应统书院

有限,并记 $\nu_k=E(\xi^k)\quad(k=1,2,\cdots,l)$,$\nu_k$ 同参数有关,记作

$$\nu_k=g_k(\theta_1,\theta_2,\cdots,\theta_l)\quad(k=1,2,\cdots,l).$$

假定求得

$$\theta_k=H_k(\nu_1,\nu_2,\cdots,\nu_l)\quad(k=1,2,\cdots,l).$$

现用样本 k 阶原点矩作为总体 k 阶原点矩 ν_k 的估计量 $\hat{\nu}_k$,即

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i^k=\hat{\nu}_k\quad(k=1,2,\cdots,l),\tag{7.1.1}$$

将 $\hat{\nu}_1,\hat{\nu}_2,\cdots,\hat{\nu}_l$ 代入 $H_1,H_2,\cdots,H_l$ 中,得到

$$\hat{\theta}_k=H_k(\hat{\nu}_1,\hat{\nu}_2,\cdots,\hat{\nu}_l)=h_k(\xi_1,\xi_2,\cdots,\xi_n)\quad(k=1,2,\cdots,l),$$

则称 $\hat{\theta}_k$ 为未知参数 θ_k 的矩法估计量.

若 $\hat{\theta}$ 为 θ 的矩法估计量,$g(\theta)$ 为 θ 的连续函数,则也称 $g(\hat{\theta})$ 为 $g(\theta)$ 的矩法估计量. 矩法估计比较直观、简便,我们也把用样本的数字特征作为总体数字特征的估计看作矩法估计. 例如,不论总体 ξ 服从什么分布,若 $E(\xi)=a$ 及 $D(\xi)=\sigma^2$ 是有限的,但其值未知,则由

$$\begin{cases}\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\xi_i=\hat{\nu}_1;\\[2ex]\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\xi_i^2=\hat{\nu}_2\end{cases}$$

解得:

$$\hat{a}=\hat{\nu}_1=\bar{\xi},$$

$$\hat{\sigma}^2=\hat{\nu}_2-(\hat{\nu}_1)^2=\frac{1}{n}\sum_{i=1}^{n}\xi_i^2-(\bar{\xi})^2$$

$$=\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2=S^2.\tag{7.1.2}$$

就是说,样本的平均值$\bar{\xi}$及方差 S^2 分别为总体的数学期望 a 及方差 σ^2 的矩法估计量.

例 7.1.1 求事件概率 p 的矩法估计量.

专业课辅导+真题资料+微信公众号：432应统书院

记事件 A 的概率 $P(A)=p$，今要求 p 的估计量. 我们用 ξ 表示事件 A 在一次试验中是否出现的随机变量，即：

$$\xi=\begin{cases}1, A\text{ 出现};\\0, A\text{ 不出现},\end{cases}$$

则 $P\{\xi=1\}=P(A)=p, P\{\xi=0\}=1-p$. 因为 $E(\xi)=p$，所以 p 的矩法估计量为：

$$\hat{p}=\frac{1}{n}\sum_{i=1}^{n}\xi_i=\frac{\mu}{n}, \tag{7.1.3}$$

其中 μ 为事件 A 在 n 次独立试验中出现的次数. 也就是说，在 n 次独立试验中，用事件 A 出现的频率 $\frac{\mu}{n}$ 作为事件 A 出现的概率 p 的矩法估计量.

例 7.1.2 设总体 ξ 在 $[\theta_1,\theta_2]$ 上服从均匀分布，其密度函数为

$$f(x;\theta_1,\theta_2)=\begin{cases}\dfrac{1}{\theta_2-\theta_1}, & x\in[\theta_1,\theta_2];\\ 0, & x\notin[\theta_1,\theta_2],\end{cases}$$

其中 $\theta_2>\theta_1$，试求 θ_1 及 θ_2 的矩法估计量.

解 我们知道，$E(\xi)=\frac{1}{2}(\theta_1+\theta_2)$，$D(\xi)=\frac{1}{12}(\theta_2-\theta_1)^2$. 由方程组

$$\begin{cases}\bar{\xi}=\dfrac{1}{2}(\hat{\theta}_1+\hat{\theta}_2);\\ S^2=\dfrac{1}{12}(\hat{\theta}_2-\hat{\theta}_1)^2\end{cases}$$

解得

$$\begin{cases}\hat{\theta}_1=\bar{\xi}-\sqrt{3}S;\\ \hat{\theta}_2=\bar{\xi}+\sqrt{3}S,\end{cases} \tag{7.1.4}$$

则 $\hat{\theta}_1,\hat{\theta}_2$ 分别为 θ_1 及 θ_2 的矩法估计量.

专业课辅导+真题资料+微信公众号：432应统书院

例 7.1.3 设总体 ξ 服从 Γ-分布 $\Gamma(\alpha,\beta)$，其密度函数为：

$$f(x)=\begin{cases}\dfrac{\beta^{-(\alpha+1)}}{\Gamma(\alpha+1)}x^{\alpha}e^{-\frac{x}{\beta}}, & x>0;\\ 0, & x\leqslant 0.\end{cases}\tag{7.1.5}$$

试求 α 及 β 的矩法估计量.

解 我们知道

$$E(\xi)=\beta(\alpha+1),D(\xi)=\beta^2(\alpha+1),$$

由方程组

$$\begin{cases}\bar{\xi}=\hat{\beta}(\hat{\alpha}+1);\\ S^2=\hat{\beta}^2(\hat{\alpha}+1).\end{cases}$$

解得

$$\hat{\alpha}=\frac{(\bar{\xi})^2}{S^2}-1,\hat{\beta}=\frac{S^2}{\bar{\xi}}.\tag{7.1.6}$$

例 7.1.4 求相关系数 ρ 的矩法估计量.

解 记

$$\bar{\xi}=\frac{1}{n}\sum_{i=1}^{n}\xi_i,S_1^2=\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2,$$

$$\bar{\eta}=\frac{1}{n}\sum_{i=1}^{n}\eta_i,S_2^2=\frac{1}{n}\sum_{i=1}^{n}(\eta_i-\bar{\eta})^2.$$

用(6.2.13)所定义的样本相关系数 R 作为 ρ 的矩法估计量，即

$$\hat{\rho}=R=\frac{S_{12}}{S_1\cdot S_2}.$$

这可理解为：用 $\bar{\xi}$ 及 $\bar{\eta}$ 分别作为 $E(\xi)$ 及 $E(\eta)$ 的矩法估计量，用 S_1^2 及 S_2^2 分别作为 $D(\xi)$ 及 $D(\eta)$ 的矩法估计量，用样本协方差 S_{12} 作为总体的协方差 $\mathrm{cov}(\xi,\eta)=E[\xi-E(\xi)][\eta-E(\eta)]$ 的矩法估计量，因而可用 R 作为 ρ 的矩法估计量，其中 ρ 为总体的相关系数

$$\rho=\frac{\mathrm{cov}(\xi,\eta)}{\sqrt{D(\xi)}\sqrt{D(\eta)}}.$$

专业课辅导+真题资料+微信公众号：432应统书院

矩法估计是最古老的点估计方法，它直观而又简便，特别是在对总体 ξ 的数学期望及方差数字特征作估计时，并不一定要知道 ξ 的分布函数 $F(x;\theta)$. 但是，矩法估计要求随机变量 ξ 的原点矩存在，如柯西分布的原点矩不存在，那就不能用矩法了. 再者，样本矩的表达式同总体 ξ 的分布函数 $F(x;\theta)$ 的表达式无关，因而矩法还没有充分利用 $F(x;\theta)$ 对参数 θ 所提供的信息.

二、极大似然法

费希尔引进的极大似然法，从理论观点来看，至今仍然是参数点估计中最重要的方法. 以后将会知道，这种估计方法，是利用总体 ξ 的分布函数 $F(x;\theta)$ 的表达式及样本所提供的信息，建立未知参数 θ 的估计量 $\hat{\theta}(\xi_1,\xi_2,\cdots,\xi_n)$.

极大似然法的想法同矩法一样也是直观的. 今举一个通俗的例子：有两位同学一起进行实弹射击，两人共同射击一个目标，事先并不知道谁的技术较好，让每人各打一发，有一人击中目标，那么认为击中的同学的技术比击不中的技术较好，显然是合理的。又举一例：有一事件，我们知道它发生的概率 p 只可能是 0.01 或 0.09，在一次观察中这事件发生了，试问这事件发生的概率是什么？当然人们会认为它发生的概率是 0.09 而不是 0.01. 我们再举一个实例，由此例便于引进极大似然法.

例 7.1.5 设在一个布袋中盛放着许多个白球和黑球，但不知道是黑球多还是白球多，只知道两种球的数目之比为 1∶3，就是说抽取到黑球的概率是 $\frac{1}{4}$ 或 $\frac{3}{4}$，我们希望通过实验来判断黑球占的比例是 $\frac{1}{4}$ 或 $\frac{3}{4}$.

解 今用有放回抽取的方法从布袋中抽取 n 个球，其中黑球的个数记为 ξ，则 ξ 服从二项分布

$$P\{\xi=x\}=C_n^x p^x(1-p)^{n-x}. \tag{7.1.7}$$

今就 $n=3$ 的情形讨论如下. 怎样通过样本的观察值也即 x 的

专业课辅导+真题资料+微信公众号：432应统书院

取值来估计参数 p 呢？换句话说，在什么情况下取 $p=\frac{1}{4}$ 而在另外的情况下取 $p=\frac{3}{4}$ 更为合理呢？为此，我们就 $p=\frac{1}{4}$ 或 $\frac{3}{4}$ 为参数值的二项分布计算得到的概率列如下表 7.1.1.

表 7.1.1

x	0	1	2	3
$P\left(x;\frac{3}{4}\right)$	$\frac{1}{64}$	$\frac{9}{64}$	$\frac{27}{64}$	$\frac{27}{64}$
$P\left(x;\frac{1}{4}\right)$	$\frac{27}{64}$	$\frac{27}{64}$	$\frac{9}{64}$	$\frac{1}{64}$

我们从对样本所下的定义中知道，样本来自总体并很好地反映了总体的概率分布特征. 因此，我们在对总体的分布函数的参数作估计时，应该从样本的观察值出发来考虑. 在这个例子中，如果我们观察到的黑球个数 $x=0$，由 $P\left(0,\frac{1}{4}\right)=\frac{27}{64}$ 及 $P\left(0,\frac{3}{4}\right)=\frac{1}{64}$ 知，显然 $P\left(0,\frac{1}{4}\right)>P\left(0,\frac{3}{4}\right)$，这表明使 $x=0$ 的样本从 $p=\frac{1}{4}$ 为参数的总体中抽取比从参数 $p=\frac{3}{4}$ 的总体中抽取更有可能发生，因而取 $\frac{1}{4}$ 作为 p 的估计比取 $\frac{3}{4}$ 作为 p 的估计更为合理. 类似地，当 $x=1$ 时，取 $\frac{1}{4}$ 作为 p 的估计比取 $\frac{3}{4}$ 更为合理. 而当 $x=2$ 或 3 时，取 $\frac{3}{4}$ 作为 p 的估计比取 $\frac{1}{4}$ 更合理. 综上所述，确定参数 p 的估计量为：

$$\hat{p}(x)=\begin{cases}\frac{1}{4}, \text{当 } x=0,1;\\ \frac{3}{4}, \text{当 } x=2,3.\end{cases} \tag{7.1.8}$$

专业课辅导+真题资料+微信公众号：432应统书院

也就是说，对于每个 x 值，选取 $\hat{p}(x)$ 使得

$$P(x,\hat{p}(x)) \geqslant P(x,p(x)),$$

其中 $p(x)$ 是不同于 $\hat{p}(x)$ 的任一估计量. 这就是按极大似然法原理估计参数的基本思想.

更一般地，总体 ξ 为连续型随机变量，具有密度函数 $f(x;\theta)$，$\theta \in \Omega$，θ 为未知参数. 又设 $(x_1,x_2,\cdots,x_n)$ 是样本 $(\xi_1,\xi_2,\cdots,\xi_n)$ 的一个观察值，那么样本 $(\xi_1,\xi_2,\cdots,\xi_n)$ 落在点 $(x_1,x_2,\cdots,x_n)$ 的邻域里的概率为 $\prod_{i=1}^{n} f(x_i;\theta)\mathrm{d}x_i$. 可见，$\theta$ 的取值不同，直接影响到 $\prod_{i=1}^{n} f(x_i;\theta)\mathrm{d}x_i$，因而它是 θ 的函数. 极大似然法原理就是选取使得样本落在观察值 $(x_1,x_2,\cdots,x_n)$ 的邻域里的概率 $\prod_{i=1}^{n} f(x_i;\theta)\mathrm{d}x_i$ 达到最大的数值 $\hat{\theta}(x_1,x_2,\cdots,x_n)$ 作为参数 θ 的估计值. 对于样本的每个观察值 $(x_1,x_2,\cdots,x_n)$，都是选取使 $\prod_{i=1}^{n} f(x_i;\theta)\mathrm{d}x_i$ 达到最大的数值 $\hat{\theta}(x_1,x_2,\cdots,x_n)$ 作为参数 θ 的估计值. 而 $\hat{\theta}(x_1,x_2,\cdots,x_n)$ 随样本观察值 $(x_1,x_2,\cdots,x_n)$ 的不同而可能取不同的数值，因此 $\hat{\theta}$ 为样本的函数，可记作 $\hat{\theta}(\xi_1,\xi_2,\cdots,\xi_n)$，称它为参数 θ 的估计量. 当 ξ 是离散型随机变量时，用 ξ 的概率函数 $P(x;\theta)$ 代替 $f(x;\theta)$，可按上述原理同样进行讨论.

定义 7.1.3（极大似然法估计） 设总体 ξ 的密度函数为 $f(x;\theta_1,\theta_2,\cdots,\theta_l)$，其中 $\theta_1,\theta_2,\cdots,\theta_l$ 为未知参数，参数空间 Ω 是 l 维的. $\xi_1,\xi_2,\cdots,\xi_n$ 为样本，它的联合密度函数为 $f(x_1,x_2,\cdots,x_n;\theta_1,\theta_2,\cdots,\theta_l)$. 称

$$L(\theta_1,\theta_2,\cdots,\theta_l)=\prod_{i=1}^{n} f(x_i;\theta_1,\theta_2,\cdots,\theta_l) \qquad (7.1.9)$$

为 $\theta_1,\theta_2,\cdots,\theta_l$ 的似然函数. 若有 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_l$ 使得下式成立：

专业课辅导+真题资料+微信公众号：432应统书院

$$L(\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_l)=\max_{(\theta_1,\theta_2,\cdots,\theta_l)\in\Omega}\{L(\theta_1,\theta_2,\cdots,\theta_l)\},\tag{7.1.10}$$

则称 $\hat{\theta}_j=\hat{\theta}_j(x_1,x_2,\cdots,x_n)$ 为 θ_j 的极大似然法估计值，相应地称 $\hat{\theta}_j(\xi_1,\xi_2,\cdots,\xi_n)$ 为参数 θ_j 的极大似然法估计量.$(j=1,2,\cdots,l)$.

我们知道，

$$\ln L(\theta_1,\theta_2,\cdots,\theta_l)=\sum_{i=1}^{n}\ln f(x_i;\theta_1,\theta_2,\cdots,\theta_l),$$

由于 $\ln x$ 是 x 的单调上升函数，因而 $\ln L$ 与 L 有相同的极大值点. 称下式

$$\left.\frac{\partial\ln L(\theta_1,\theta_2,\cdots,\theta_l)}{\partial\theta_j}\right|_{\theta_j=\hat{\theta}_j}=0,j=1,2,\cdots,l\tag{7.1.11}$$

为似然方程，由它解得 $\hat{\theta}_j=\hat{\theta}_j(\xi_1,\xi_2,\cdots,\xi_n)$，当满足(7.1.10)式时，则 $\hat{\theta}_j$ 为 θ_j 的极大似然法估计量$(j=1,2,\cdots,l)$.①

若 ξ 为离散型随机变量，概率函数为 $P(x;\theta_1,\theta_2,\cdots,\theta_l)$，则似然函数为

$$L(\theta_1,\theta_2,\cdots,\theta_l)=\prod_{i=1}^{n}P(x_i;\theta_1,\theta_2,\cdots,\theta_l).\tag{7.1.12}$$

由似然方程组

$$\left.\frac{\partial\ln L(\theta_1,\theta_2,\cdots,\theta_l)}{\partial\theta_j}\right|_{\theta_j=\hat{\theta}_j}=0,j=1,2,\cdots,l,$$

解得 $\hat{\theta}_j=\hat{\theta}_j(\xi_1,\xi_2,\cdots,\xi_n)$，若满足(7.1.10)式，则 $\hat{\theta}_j$ 为 θ_j 的极大似然法估计量$(j=1,2,\cdots,l)$.

注 在例 7.1.6—例 7.1.11 中写似然函数时，为叙述简化，直接用样本 $\xi_1,\xi_2,\cdots,\xi_n$。

例 7.1.6 设总体 ξ 服从正态 $N(a,\sigma)$，试求 a 及 σ 的极大似

① 由数学分析求极值的原理知，(7.1.11)的解是否为极大值的点，还要验证二阶导数在该点是否小于零.

专业课辅导+真题资料+微信公众号：432应统书院

然法估计量.

解 我们知道，a 及 σ 的似然函数为

$$L(a,\sigma)=\prod_{i=1}^{n}\left[\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(\xi_i-a)^2}{2\sigma^2}}\right]$$

$$=\frac{1}{(\sqrt{2\pi}\sigma)^n}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\xi_i-a)^2}.$$

似然方程为：

$$\begin{cases}\dfrac{\partial\ln L(a,\sigma)}{\partial a}=\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(\xi_i-a)=0;\\[2ex]\dfrac{\partial\ln L(a,\sigma)}{\partial\sigma}=-\dfrac{n}{\sigma}+\dfrac{1}{\sigma^3}\sum\limits_{i=1}^{n}(\xi_i-a)^2=0.\end{cases}$$

解得：

$$\hat{a}=\frac{1}{n}\sum_{i=1}^{n}\xi_i=\bar{\xi},$$

$$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2=S^2.$$

则$\bar{\xi}$及 S^2 分别为 a 及 σ^2 的极大似然法估计量.

例 7.1.7 设总体 ξ 服从具有参数 $\lambda>0$ 的泊松分布，概率函数为

$$P(x;\lambda)=\frac{\lambda^x}{x!}e^{-\lambda},$$

$\lambda>0,x=0,1,2,\cdots$. 试求参数 λ 的极大似然估计量.

解 我们知道，参数 λ 的似然函数为

$$L(\lambda)=\prod_{i=1}^{n}\left[\frac{\lambda^{\xi_i}}{\xi_i!}e^{-\lambda}\right]=e^{-n\lambda}\frac{\lambda^{\sum_{i=1}^{n}\xi_i}}{\prod_{i=1}^{n}(\xi_i!)},$$

似然方程为

$$\frac{\partial\ln L(\lambda)}{\partial\lambda}=-n+\frac{1}{\lambda}\sum_{i=1}^{n}\xi_i=0.$$

专业课辅导+真题资料+微信公众号：432应统书院

解得

$$\hat{\lambda}=\frac{1}{n}\sum_{i=1}^{n}\xi_i=\bar{\xi},$$

即样本平均值$\bar{\xi}$为参数 λ 的极大似然法估计量.

例 7.1.8 设总体 ξ 服从参数为 λ 的指数分布,密度函数为

$$f(x;\lambda)=\lambda e^{-\lambda x},$$

$x>0,\lambda>0$. 试求参数 λ 的极大似然法估计量.

解 我们知道,λ 的似然函数为

$$L(\lambda)=\prod_{i=1}^{n}(\lambda e^{-\lambda\xi_i})=\lambda^n e^{-\lambda\sum_{i=1}^{n}\xi_i},$$

则

$$\ln L(\lambda)=n\ln\lambda-\lambda\sum_{i=1}^{n}\xi_i.$$

似然方程为

$$\frac{\partial\ln L(\lambda)}{\partial\lambda}=\frac{n}{\lambda}-\sum_{i=1}^{n}\xi_i=0,$$

解得:

$$\hat{\lambda}=\frac{1}{\bar{\xi}}.$$

例 7.1.9 设总体 ξ 在$[\theta_1,\theta_2]$上服从均匀分布,求 θ_1 及 θ_2 的极大似然估计量.

解 由 ξ 的密度函数

$$f(x;\theta_1,\theta_2)=\begin{cases}\dfrac{1}{\theta_2-\theta_1}, & x\in[\theta_1,\theta_2];\\ 0, & x\notin[\theta_1,\theta_2],\end{cases}$$

知 θ_1 及 θ_2 的似然函数为:

$$L(\xi_1,\xi_2,\cdots,\xi_n;\theta_1,\theta_2)=\begin{cases}\left(\dfrac{1}{\theta_2-\theta_1}\right)^n, & \theta_1\leqslant\xi_i\leqslant\theta_2,(i=1,2,\cdots,n);\\ 0, & \text{其他}.\end{cases}$$

专业课辅导+真题资料+微信公众号：432应统书院

似然方程为

$$\begin{cases}\dfrac{\partial \ln L}{\partial \theta_1}=\dfrac{n}{\theta_2-\theta_1}=0;\\ \dfrac{\partial \ln L}{\partial \theta_2}=\dfrac{-n}{\theta_2-\theta_1}=0.\end{cases} \tag{7.1.13}$$

我们看到，从(7.1.13)不可能解得 θ_1 及 θ_2 的极大似然估计量.

现在，我们从似然函数的定义出发来确定 θ_1 及 θ_2 的极大似然估计量. 我们知道，要使似然函数 $L(\xi_1,\xi_2,\cdots,\xi_n;\theta_1,\theta_2)$ 非零，必须要求 $\theta_1\leqslant\xi_1^*=\min\{\xi_1,\xi_2,\cdots,\xi_n\}$，$\xi_n^*=\max\{\xi_1,\xi_2,\cdots,\xi_n\}\leqslant\theta_2$，由于

$$L(\xi_1,\xi_2,\cdots,\xi_n;\theta_1,\theta_2)=\frac{1}{(\theta_2-\theta_1)^n}\leqslant\frac{1}{(\xi_n^*-\xi_1^*)^n},$$

今取

$$\hat{\theta}_1=\xi_1^*,\hat{\theta}_2=\xi_n^*, \tag{7.1.14}$$

则有

$$L(\xi_1,\xi_2,\cdots,\xi_n;\theta_1,\theta_2)\leqslant L(\xi_1,\xi_2,\cdots,\xi_n;\hat{\theta}_1,\hat{\theta}_2).$$

因而 $\hat{\theta}_1$ 及 $\hat{\theta}_2$ 为 θ_1 及 θ_2 的极大似然估计量.

下面再举一个借助极大似然法的思想估计参数的例子.

例 7.1.10 钓鱼问题. 为了估计湖中的鱼数 N，同时自湖中钓出 r 条鱼做上记号后放回湖中，然后再自湖中同时钓出 s 条鱼，结果发现这 s 条中有 x 条标有记号. 这里 N 是未知数，r 及 s 是已知常数，试问应如何估计 N 的值?

解 第二次钓出的标有记号的鱼数 ξ 是随机变量，且 ξ 服从超几何分布：

$$P\{\xi=x\}=\frac{C_r^x C_{N-r}^{s-x}}{C_N^s}, \tag{7.1.15}$$

其中 x 为整数，且 $\max(0,s-(N-r))\leqslant x\leqslant\min(r,s)$. 今用 $L(x;N)$ 表示(7.1.15)式右端，则取使 $L(x;N)$ 达到极大值的 $\hat{N}$ 作为 N 的估

专业课辅导+真题资料+微信公众号：432应统书院

计量. 但是,直接对 N 求导数的方法相当困难,现用下述方法. 考虑比值

$$A(x;N)=\frac{L(x;N)}{L(x;N-1)}=\frac{N-r}{N}\cdot\frac{N-s}{(N-r)-(s-x)}$$
$$=\frac{N^2-(r+s)N+rs}{N^2-(r+s)N+Nx}. \tag{7.1.16}$$

从(7.1.16)式看到,当且仅当 $N<\frac{rs}{x}$, $L(x;N)>L(x;N-1)$,当且仅当 $N>\frac{rs}{x}$, $L(x;N)<L(x;N-1)$,因此 $L(x;N)$ 在 $\frac{rs}{x}$ 附近取极大值,于是 N 的估计量为:

$$\hat{N}=\left[\frac{rs}{x}\right]. \tag{7.1.17}$$

其右端的方括号表示取整数部分.

极大似然估计量有许多重要的性质. 因此,当碰到似然方程难解时,为了获得参数的极大似然法估计量,设法求似然方程的近似解.

例 7.1.11 设总体 ξ 服从 Γ-分布 $\Gamma(\alpha,\beta)$,求 α 及 β 的极大似然估计量.

解 由(7.1.9)式知,参数 α 及 β 的似然函数为

$$L(\xi_1,\xi_2,\cdots,\xi_n;\alpha,\beta)=\frac{\beta^{-n(\alpha+1)}}{[\Gamma(\alpha+1)]^n}\left[\prod_{i=1}^{n}\xi_i\right]^{\alpha}\mathrm{e}^{-\frac{1}{\beta}\sum_{i=1}^{n}\xi_i},$$

则

$$\ln L=-n(\alpha+1)\ln\beta-n\ln\Gamma(\alpha+1)+\alpha\sum_{i=1}^{n}\ln\xi_i-\frac{1}{\beta}\sum_{i=1}^{n}\xi_i,$$

$$\frac{\partial\ln L}{\partial\alpha}=-n\ln\beta-\frac{n\Gamma'(\alpha+1)}{\Gamma(\alpha+1)}+\sum_{i=1}^{n}\ln\xi_i,$$

$$\frac{\partial\ln L}{\partial\beta}=-\frac{n(\alpha+1)}{\beta}+\frac{n\bar{\xi}}{\beta^2}.$$

解得 $\hat{\beta}=\frac{\bar{\xi}}{\hat{\alpha}+1}$,其中 $\hat{\alpha}$ 由下式确定:

专业课辅导+真题资料+微信公众号：432应统书院

$$\ln(\alpha+1)-\frac{\Gamma'(\alpha+1)}{\Gamma(\alpha+1)}=\ln\bar{\xi}-\frac{1}{n}\sum_{i=1}^{n}\ln\xi_i. \qquad (7.1.18)$$

切普曼(D. G. Chapman)提出了解超越方程(7.1.18)的方法(1956年),并给出了函数 $\Gamma'(\alpha+1)/\Gamma(\alpha+1)$ 数值表供查用(在此从略).

记

$$Q(\alpha+1)=\ln(\alpha+1)-\frac{\Gamma'(\alpha+1)}{\Gamma(\alpha+1)},$$

$$G(\xi_1,\xi_2,\cdots,\xi_n)=\ln\bar{\xi}-\frac{1}{n}\sum_{i=1}^{n}\ln\xi_i,$$

用"试差法"求 α 的估计 $\hat{\alpha}$. 先给定 α 的初始值 α_0,若 $Q(\alpha_0+1)=G$,则 α_0 为 α 的估计值. 如果 $Q(\alpha_0+1)\neq G$,则选取 α_1,直至选出使等式成立或两端很接近的 $\hat{\alpha}$ 值,作为 α 的估计值.

§7.2 无偏性与优效性

我们从上面已经看到,对于总体 ξ 的数学期望 $E(\xi)$,根据矩法可用样本的算术平均值 $\bar{\xi}$ 作为它的估计量. 但是,我们也可用另一种方法,例如用样本的加权平均值 $\sum_{i=1}^{n}c_i\xi_i$ 作为它的估计量,其中 $c_i\geqslant 0$, $\sum_{i=1}^{n}c_i=1$. 甚至更简便地用容量为1的样本 ξ_1 作为 $E(\xi)$ 的估计. 问题在于哪个估计量为它的"最佳"估计,而"最佳"的准则又怎样确定? 我们知道,未知参数 θ 的估计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 为一不带有未知参数的统计量,是一随机变量. 而它的分布函数与总体 ξ 的分布函数有关,即与未知参数 θ 有关. 如果 $E_\theta(T)$ 等于或渐近于被估计的参数 θ,$D_\theta(T)$ 很小或渐近于零,那么估计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 由样本观察值 $(x_1,x_2,\cdots,x_n)$ 而计算得到的数值,虽然不能都等于 θ,但其平均值等于或渐近于 θ,而且其离散度很小. 这时,用 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 作为未知参数 θ 的估计,人们自然感到满意. 下面,我们从估计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 的数学期望及方差这两个

专业课辅导+真题资料+微信公众号：432应统书院

重要的数字特征出发，引进无偏性、优效性、相合性等概念，从不同角度来衡量估计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 作为 θ 的"最佳"准则，即讨论估计量的优良性.

注意，并非未知参数 θ 的每个估计量都具有上述性质. 对于柯西分布而言，由于它的数学期望不存在，因而谈参数估计量的无偏性及优效性将失去意义.

一、无偏性

定义 7.2.1（无偏估计量） 若参数 θ 的估计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 对一切 n 及 $\theta\in\Omega$，有

$$E_\theta[T(\xi_1,\xi_2,\cdots,\xi_n)]=\theta, \tag{7.2.1}$$

则称 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 为参数 θ 的无偏估计量. 记

$$E_\theta(T)-\theta=b_n, \tag{7.2.2}$$

称 b_n 为估计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 的偏差. 若 $b_n\neq 0$，则称 T 为 θ 的有偏估计量. 如果

$$\lim_{n\to\infty} b_n=0, \tag{7.2.3}$$

则称 T 为 θ 的渐近无偏估计量.

对于参数 θ 的任一实值函数 $g(\theta)$，如果 $g(\theta)$ 的无偏估计量存在，也就是说有估计量 T，使得

$$E_\theta(T)=g(\theta), \tag{7.2.4}$$

则称 $g(\theta)$ 为可估计函数.

例 7.2.1 设总体 ξ 的 k 阶原点矩存在，即 $\nu_k=E(\xi^k)$ 有限，则样本的 k 阶原点矩为总体的 k 阶原点矩的无偏估计量.

证 设样本的 k 阶原点矩为

$$A_k=\frac{1}{n}\sum_{i=1}^{n}\xi_i^k,$$

则有

$$E(A_k)=\frac{1}{n}\sum_{i=1}^{n}E(\xi_i^k)=\nu_k. \tag{7.2.5}$$

特别地，若 $E(\bar{\xi})=a$，则 $A_1=\bar{\xi}$ 为 a 的无偏估计量.

专业课辅导+真题资料+微信公众号：432应统书院

例 7.2.2 设总体 ξ 的方差 $D(\xi)=\sigma^2$ 是有限的，试求 σ^2 的无偏估计量.

解 样本的方差 S^2 是 σ^2 的矩法估计，由于

$$S^2=\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2=\frac{1}{n}\sum_{i=1}^{n}\xi_i^2-(\bar{\xi})^2.$$

$$\begin{aligned}E(S^2)&=E\left[\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2\right]=E\left[\frac{1}{n}\sum_{i=1}^{n}\xi_i^2-(\bar{\xi})^2\right]\\&=\frac{1}{n}\sum_{i=1}^{n}E(\xi_i^2)-E[(\bar{\xi})^2]=E(\xi^2)-[E(\bar{\xi})^2]\\&=\{D(\xi)+[E(\xi)]^2\}-\{D(\bar{\xi})+[E(\bar{\xi})]^2\}\\&=D(\xi)-D(\bar{\xi})=\sigma^2-D(\bar{\xi}).\end{aligned}$$

其中

$$\begin{aligned}D(\bar{\xi})&=D\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(\xi_i)=\frac{1}{n^2}\sum_{i=1}^{n}D(\xi)\\&=\frac{1}{n}D(\xi)=\frac{\sigma^2}{n},\end{aligned}$$

于是

$$E(S^2)=\sigma^2-\frac{\sigma^2}{n}=\frac{n-1}{n}\sigma^2. \tag{7.2.6}$$

因此，样本方差 S^2 不是总体的方差 σ^2 的无偏估计量，当然，S^2 是 σ^2 的渐近无偏估计量. 现将有偏估计量 S^2 无偏化，记

$$S^{*2}=\frac{nS^2}{n-1}=\frac{1}{n-1}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2, \tag{7.2.7}$$

则

$$E(S^{*2})=\sigma^2. \tag{7.2.8}$$

即 S^{*2} 为 σ^2 的无偏估计量.

由上述两例可见，不论总体 ξ 服从什么分布，只要 $E(\xi)=a$ 及 $D(\xi)=\sigma^2$ 是有限的话，那么

$$\bar{\xi}=\frac{1}{n}\sum_{i=1}^{n}\xi_i,\quad S^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2$$

专业课辅导+真题资料+微信公众号：432应统书院

分别为 a 及 σ^2 的无偏估计量.

注意,当 $E(\hat{\theta})=\theta$ 时,不一定有 $E[g(\hat{\theta})]=g(\theta)$,其中 $g(\theta)$ 为 θ 的实值函数. 也就是说,当 $\hat{\theta}$ 为 θ 的无偏估计量时,$g(\hat{\theta})$ 不一定是 $g(\theta)$ 的无偏估计量.

例 7.2.3 设 $D(\xi)\neq 0$,若 $E(\bar{\xi})=a$,则 $E(\bar{\xi}^2)=a^2$ 不成立.

这是因为:

$$E(\bar{\xi}^2)=D(\bar{\xi})+[E(\bar{\xi})]^2=\frac{1}{n}D(\xi)+a^2,$$

所以

$$E(\bar{\xi}^2)\neq a^2.$$

定义 7.2.2(最小方差无偏估计量) 设 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 为可估计函数 $g(\theta)$ 的无偏估计量,若对于 $g(\theta)$ 的任一无偏估计量 $T'(\xi_1,\xi_2,\cdots,\xi_n)$,有

$$D_\theta(T)\leqslant D_\theta(T'),\quad 对一切\ \theta\in\Omega,$$

则称 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 为 $g(\theta)$ 的最小方差无偏估计量. 最小方差无偏估计量简称为最优无偏估计量.

为便于下述两个定理的叙述,记

$$U=\{T:E_\theta(T)=\theta,D_\theta(T)<\infty,对一切\ \theta\in\Omega\},\tag{7.2.9}$$

即 U 为参数 θ 的方差有限的无偏估计量的集合.

又记

$$U_0=\{T_0:E_\theta(T_0)=0,D_\theta(T_0)<\infty,对一切\ \theta\in\Omega\},$$

U_0 是参数 θ 的数学期望为零、方差有限的估计量的集合.

定理 7.2.1 设 U 是非空的集合,有一 $T\in U$,则 T 为 θ 的最优无偏估计量的充要条件为对每个 $T_0\in U_0$,有

$$E_\theta(T\cdot T_0)=0,\tag{7.2.10}$$

其中 $\theta\in\Omega$.

证 必要性 用反证法,设 $T\in U$ 为 θ 的最优无偏估计量,但条件(7.2.10)不满足,即存在 $T_0\in U_0$ 及 $\theta_0\in\Omega$,使得

专业课辅导+真题资料+微信公众号：432应统书院

$$E_{\theta_0}(T\cdot T_0)\neq 0.$$

因为 $E_\theta(T_0)=0$，所以对一切 c 值，$T-cT_0\in U$. 于是

$$E_{\theta_0}(T-cT_0)^2=E_{\theta_0}(T^2)+c^2E_{\theta_0}(T_0^2)-2cE_{\theta_0}(T\cdot T_0).$$

由于 $E_{\theta_0}(T\cdot T_0)\neq 0$，一定能找到 c_0 值，例如

$$c_0=\frac{E_{\theta_0}(T\cdot T_0)}{E_{\theta_0}(T_0^2)},$$

使得

$$E_{\theta_0}(T-c_0T_0)^2<E_{\theta_0}(T^2),$$

所以

$$D_{\theta_0}(T-c_0T_0)<D_{\theta_0}(T).$$

这与 T 为最优无偏估计量的假定相矛盾，因此说明条件(7.2.10)是必要的.

充分性 设有某个 $T\in U$，使得(7.2.10)成立，我们来证 T 为 θ 的最优无偏估计量. 若(7.2.10)成立，则对任一估计量 $T'\in U$，显然 $T'-T\in U_0$，因而对一切 $\theta\in\Omega$，有

$$E_\theta[T(T'-T)]=0.$$

由施瓦兹不等式知，对一切 $\theta\in\Omega$，

$$E_\theta(T^2)=E_\theta(T\cdot T')\leqslant\sqrt{E_\theta(T^2)}\sqrt{E_\theta(T'^2)},$$

所以

$$E_\theta(T^2)\leqslant E_\theta(T'^2).$$

由于 $E_\theta(T)=E_\theta(T')=\theta$，因而对一切 $\theta\in\Omega$，

$$D_\theta(T)\leqslant D_\theta(T').$$

可见 T 为 θ 的最优无偏估计量.

推论 设 T_1 和 T_2 分别是参数 θ 的可估计函数 $g_1(\theta)$ 和 $g_2(\theta)$ 的最优无偏估计量，则 $b_1T_1+b_2T_2$ 是 $b_1g_1(\theta)+b_2g_2(\theta)$ 的最优无偏估计量，其中 b_1 和 b_2 是常数.

证明留作练习.

定理 7.2.2 设 U 是(7.2.9)式所定义的非空集合，则对参数

专业课辅导+真题资料+微信公众号：432应统书院

θ 而言至多存在一个最优无偏估计.

证 假定 T 及 T' 是 θ 的两个最优无偏估计量,即对一切 $\theta\in\Omega$,有

$$E_\theta(T)=E_\theta(T')=\theta,$$
$$D_\theta(T)=D_\theta(T').$$

因此对一切 $\theta\in\Omega$,有

$$E_\theta(T-T')=0.$$

即 $T-T'\in U_0$. 由定理 7.2.1 知,对一切 $\theta\in\Omega$,有:

$$E_\theta[T(T-T')]=0,$$
$$E_\theta[T'(T-T')]=0.$$

于是对一切 $\theta\in\Omega$,有:

$$E_\theta(T-T')^2=E_\theta[T(T-T')]-E_\theta[T'(T-T')]=0,$$

亦即对一切 $\theta\in\Omega$,有:

$$D_\theta(T-T')=0.$$

由切比雪夫不等式知,对一切 $\theta\in\Omega$,有:

$$P_\theta\{T=T'\}=1.$$

这就说明参数 θ 的最优无偏估计量在概率为 1 的意义下是唯一的.

例 7.2.4 设总体 ξ 服从正态 $N(a,\sigma)$,其中 a 及 σ^2 都是未知参数,试求 a 及 σ^2 的最优无偏估计量.

解 我们知道,若 $\xi_1,\xi_2,\cdots,\xi_n$ 为总体 ξ 的样本,则 $\bar{\xi}$,S^{*2} 分别为 a,σ^2 的无偏估计量,试问它们是否为最优无偏估计量? 我们应用定理 7.2.1 看一看 $\bar{\xi}$、S^{*2} 是否满足(7.2.10)式的条件,即

$$E(\bar{\xi}\cdot T_0)=0,$$
$$E(S^{*2}\cdot T_0)=0$$

对每个 $T_0\in U_0$ 是否成立.

设 $T_0\in U_0$,则有 $E(T_0)=0$,即

专业课辅导+真题资料+微信公众号：432应统书院

$$\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} T_0(x_1, x_2, \cdots, x_n)$$

$$\times \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-a)^2\right] \mathrm{d}x_1 \cdots \mathrm{d}x_n = 0. \qquad (7.2.11)$$

将(7.2.11)对未知参数 a 求导数,则有

$$\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} T_0(x_1, x_2, \cdots, x_n)\left[\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-a)\right]$$

$$\times \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-a)^2\right] \mathrm{d}x_1 \cdots \mathrm{d}x_n = 0.$$

将非零系数消去,上式等价于

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} T_0(x_1, x_2, \cdots, x_n)\left(\sum_{i=1}^{n} x_i\right)$$

$$\times \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-a)^2\right] \mathrm{d}x_1 \cdots \mathrm{d}x_n = 0.$$

于是有

$$E\left(T_0 \cdot \sum_{i=1}^{n} \xi_i\right) = 0,$$

即有

$$E(T_0 \cdot \bar{\xi}) = 0.$$

由于$\bar{\xi}$是 a 的无偏估计量. 应用定理 7.2.1 知,$\bar{\xi}$为参数 a 的最优无偏估计量.

将(7.2.11)式对参数 a 求导两次得到

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} T_0(x_1, x_2, \cdots, x_n)\left(\sum_{i=1}^{n} x_i\right)^2$$

$$\times \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-a)^2\right] \mathrm{d}x_1 \cdots \mathrm{d}x_n = 0,$$

将(7.2.11)对参数 σ^2 求导便得

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} T_0(x_1, x_2, \cdots, x_n)\left[\sum_{i=1}^{n}(x_i-a)^2\right]$$

专业课辅导+真题资料+微信公众号：432应统书院

$$\times\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-a)^2\right]\mathrm{d}x_1\cdots\mathrm{d}x_n=0$$

（其中积分号前的非零系数省略了）. 由于

$$\sum_{i=1}^{n}(x_i-a)^2=\sum_{i=1}^{n}(x_i-\bar{x})^2+n(\bar{x}-a)^2$$

$$=\sum_{i=1}^{n}(x_i-\bar{x})^2+\frac{1}{n}\left(\sum_{i=1}^{n}x_i-na\right)^2,$$

因而可得

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}T_0(x_1,x_2,\cdots,x_n)\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]$$

$$\times\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-a)^2\right]\mathrm{d}x_1\cdots\mathrm{d}x_n=0.$$

由于 $\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2=nS^2$ 是 $(n-1)\sigma^2$ 的无偏估计量，应用定理7.2.1知 nS^2 是参数 $(n-1)\sigma^2$ 的最优无偏估计量，再应用定理的推论知，$\frac{nS^2}{n-1}=\frac{1}{n-1}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2=S^{*2}$ 是参数 σ^2 的最优无偏估计量.

我们说参数 θ 的估计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 是样本 $\xi_1,\xi_2,\cdots,\xi_n$ 的线性函数，是指

$$T(\xi_1,\xi_2,\cdots,\xi_n)=\sum_{i=1}^{n}c_i\xi_i,$$

其中 $c_1,c_2,\cdots,c_n$ 为给定的常数.

定义 7.2.3（最小方差线性无偏估计） 可估计函数 $g(\theta)$ 的估计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 称为它的最小方差线性无偏估计量，是指：如果

（1）$T(\xi_1,\xi_2,\cdots,\xi_n)$ 为样本的线性函数；

（2）$E_\theta[T(\xi_1,\xi_2,\cdots,\xi_n)]=g(\theta)$；

（3）对于满足（1）及（2）的任一估计量 $T'(\xi_1,\xi_2,\cdots,\xi_n)$，对一切 $\theta\in\Omega$ 有

$$D_\theta(T)\leqslant D_\theta(T').$$

专业课辅导+真题资料+微信公众号：432应统书院

定理 7.2.3 设 $T_1, T_2, \cdots, T_m$ 为可估计函数 $g(\theta)$ 的 m 个相互独立的线性无偏估计量，而且它们有相同的方差 $D_\theta(T_j)=\sigma^2<\infty$，$j=1,2,\cdots,m$，则统计量

$$\overline{T}=\frac{1}{m}\sum_{j=1}^{m}T_j$$

是 $T_1, T_2, \cdots, T_m$ 的线性组合类中可估计函数 $g(\theta)$ 的最小方差线性无偏估计量，且有 $D_\theta(T)=\dfrac{\sigma^2}{m}$.

证 令

$$T=\sum_{j=1}^{m}b_jT_j+b_0,$$

其中 b_0 及 $b_1, b_2, \cdots, b_m$ 为常数. 我们来考虑在什么情况下，估计量 T 为 $g(\theta)$ 的最小方差线性无偏估计量. 首先要 T 为 $g(\theta)$ 的无偏估计量，则由

$$E_\theta(T)=\sum_{j=1}^{m}b_jg(\theta)+b_0=g(\theta)\sum_{j=1}^{m}b_j+b_0,$$

知 $b_0=0$，$\sum\limits_{j=1}^{m}b_j=1$. 现进一步考虑估计量 T 的方差.

$$\begin{aligned}D_\theta(T)&=\sum_{j=1}^{m}b_j^2E_\theta[T_j-E_\theta(T_j)]^2\\&=\sum_{j=1}^{m}b_j^2\sigma^2=\sigma^2\sum_{j=1}^{m}b_j^2.\end{aligned}$$

欲要求 $D(T)$ 在 $\sum\limits_{j=1}^{m}b_j=1$ 的约束条件下达到极小值，我们通过下述方法来讨论. 记：

$$\varphi(c)=\sigma^2\left(\sum_{j=1}^{m}b_j^2\right)-c\left(\sum_{j=1}^{m}b_j-1\right),$$

c 为拉格朗日因子，由

$$\frac{\partial\varphi(c)}{\partial b_j}=0\quad(j=1,2,\cdots,m),$$

专业课辅导+真题资料+微信公众号：432应统书院

知

$$b_j=\frac{c}{2\sigma^2}\quad(j=1,2,\cdots,m).$$

又由 $\sum_{j=1}^{m} b_j=1$，因而 $c=\frac{2\sigma^2}{m}$，于是

$$b_j=\frac{1}{m}\quad(j=1,2,\cdots,m).$$

也就是说，当 $b_j=\frac{1}{m}$ 时，$D(T)$ 达到最小值，因而 $\overline{T}=\frac{1}{m}\sum_{j=1}^{m} T_j$ 是 T_1，$T_2,\cdots,T_m$ 的线性组合类中 $g(\theta)$ 的最小方差线性无偏估计，且有 $D(\overline{T})=\sigma^2\sum_{j=1}^{m}\frac{1}{m^2}=\frac{\sigma^2}{m}$.

推论 设总体 ξ 的数学期望 a 及方差 σ^2 都存在，$\xi_1,\xi_2,\cdots,\xi_n$ 为样本，则样本平均值 $\overline{\xi}$ 为 a 的最小方差线性无偏估计量.

证明留作练习.

最小方差线性无偏估计，简称作最优线性无偏估计. 对于线性无偏估计类，有类似于定理 7.2.1 及定理 7.2.2 的结论.

二、优效性

我们知道，最优无偏估计量的方差总不会大于最优线性无偏估计量的方差. 对于未知参数 θ，在它的许多不同的无偏估计量中，它们的方差可以小到怎样的程度（当然方差总是非负的）？另外，在参数 θ 的无偏估计类中，最优无偏估计量是否存在，怎样寻找最优无偏估计量？在线性无偏估计类中，又怎样寻找最优线性无偏估计量？这些问题在定理 7.2.1 ~ 定理 7.2.3 中已作了初步讨论. 在这一段中，我们引进拉奥-克拉默不等式，指出无偏估计类中方差的下界，达到这个下界的无偏估计量称为优效估计量.

定理 7.2.4（拉奥-克拉默不等式） 设总体 ξ 为连续型随机变量，密度函数为 $f(x;\theta)$，θ 为未知参数，$\theta\in\Omega$，$\xi_1,\xi_2,\cdots,\xi_n$ 为样

专业课辅导+真题资料+微信公众号：432应统书院

本，$T(\xi_1,\xi_2,\cdots,\xi_n)$为可估计函数$g(\theta)$的无偏估计量，如果

(1) $\dfrac{\partial f(x;\theta)}{\partial\theta}$存在，且有 $E_\theta\left[\dfrac{\partial}{\partial\theta}\ln f(\xi;\theta)\right]^2=I(\theta)>0$；

(2) $$\frac{\partial}{\partial\theta}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left[\prod_{i=1}^{n}f(x_i;\theta)\right]\mathrm{d}x_1\cdots\mathrm{d}x_n$$

$$=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\frac{\partial}{\partial\theta}\left[\prod_{i=1}^{n}f(x_i;\theta)\right]\mathrm{d}x_1\cdots\mathrm{d}x_n;$$

(3) $\dfrac{\partial g(\theta)}{\partial\theta}$存在，且有

$$\frac{\partial g(\theta)}{\partial\theta}=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}T(x_1,x_2,\cdots,x_n)\frac{\partial}{\partial\theta}\left[\prod_{i=1}^{n}f(x_i;\theta)\right]\mathrm{d}x_1\cdots\mathrm{d}x_n,$$

则有

$$D_\theta(T)\geqslant\frac{[g'(\theta)]^2}{nI(\theta)},\tag{7.2.12}$$

其中 $g'(\theta)=\dfrac{\partial g(\theta)}{\partial\theta}$. 特别当 $g(\theta)=\theta$ 时，上式可简化为

$$D_\theta(T)\geqslant\frac{1}{nI(\theta)}.\tag{7.2.13}$$

证明 若 $I(\theta)=\infty$ 或 $D_\theta(T)=\infty$，则(7.2.12)式显然成立，因而可假定 $I(\theta)<\infty$，$D_\theta(T)<\infty$.

首先，我们知道

$$\int_{-\infty}^{\infty}f(x_i;\theta)\mathrm{d}x_i=1\quad(i=1,2,\cdots,n);$$

$$g(\theta)=E(T)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}T(x_1,x_2,\cdots,x_n)\prod_{i=1}^{n}[f(x_i;\theta)\mathrm{d}x_i].$$

将上述两式分别对 θ 求导数，则有

$$0=\int_{-\infty}^{\infty}\frac{\partial}{\partial\theta}f(x_i;\theta)\mathrm{d}x_i=\int_{-\infty}^{\infty}\frac{\partial}{\partial\theta}[\ln f(x_i;\theta)]f(x_i;\theta)\mathrm{d}x_i$$

$$=E\left[\frac{\partial}{\partial\theta}\ln f(\xi_i;\theta)\right]\quad(i=1,2,\cdots,n);$$

专业课辅导+真题资料+微信公众号：432应统书院

$$
\begin{aligned}
g'(\theta) &= \frac{\partial}{\partial\theta}\left\{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} T(x_1,x_2,\cdots,x_n)\prod_{i=1}^{n}[f(x_i;\theta)\mathrm{d}x_i]\right\} \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} T(x_1,x_2,\cdots,x_n)\left[\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\ln f(x_i;\theta)\right] \\
&\quad \times\prod_{i=1}^{n}[f(x_i;\theta)\mathrm{d}x_i] \\
&= E[T\cdot Z] = E(T)E(Z)+\mathrm{cov}(T,Z) \\
&= E(T)E(Z)+\rho\sqrt{D(T)}\sqrt{D(Z)}.
\end{aligned}
$$

其中

$$
Z=\sum_{i=1}^{n}\frac{\partial}{\partial\theta}[\ln f(\xi_i;\theta)],\ T=T(\xi_1,\xi_2,\cdots,\xi_n),
$$

$$
\rho=\frac{\mathrm{cov}(T,Z)}{\sqrt{D(T)}\sqrt{D(Z)}}.
$$

由于

$$
E(Z)=\sum_{i=1}^{n}E\left[\frac{\partial}{\partial\theta}\ln f(\xi_i;\theta)\right]=0,
$$

$$
\begin{aligned}
D(Z)=E(Z^2) &= \sum_{i=1}^{n}E\left[\frac{\partial}{\partial\theta}\ln f(\xi_i;\theta)\right]^2 \\
&= nE\left[\frac{\partial}{\partial\theta}\ln f(\xi_i;\theta)\right]^2,
\end{aligned}
$$

于是

$$
\rho^2=\frac{[g'(\theta)]^2}{D(T)\cdot D(Z)}.
$$

因为 $\rho^2\leqslant 1$，所以

$$
\begin{aligned}
D(T) &\geqslant \frac{[g'(\theta)]^2}{D(Z)} \\
&= \frac{[g'(\theta)]^2}{nE\left[\frac{\partial}{\partial\theta}\ln f(\xi;\theta)\right]^2}.
\end{aligned}
$$

若总体 ξ 为离散型，其概率函数 $P(x;\theta)$ 满足定理 7.2.4 中相

专业课辅导+真题资料+微信公众号：432应统书院

应的(1)、(2)及(3)三个条件,用求和式代积分式,则有

$$D(T)\geqslant\frac{[g'(\theta)]^2}{nE\left[\frac{\partial}{\partial\theta}\ln P(\xi;\theta)\right]^2}.$$

定义 7.2.4(正规分布)　若总体 ξ 的密度函数(或概率函数)满足定理 7.2.4 中的(1)及(2)两个条件,则称 ξ 的分布函数为正规分布,满足条件(3)的估计量称为正规估计量.(对于 $f(x;\theta)$ 用积分式,对于 $P(x;\theta)$ 用求和式).

定义 7.2.5(优效估计)　如果参数 θ 的函数 $g(\theta)$ 的某个正规无偏估计量 $U(\xi_1,\xi_2,\cdots,\xi_n)$ 的方差达到拉奥-克拉默不等式的下界,则称这个无偏估计量 $U(\xi_1,\xi_2,\cdots,\xi_n)$ 为 $g(\theta)$ 的优效估计量.

就是说,参数 $g(\theta)$ 的正规无偏估计量 $U(\xi_1,\xi_2,\cdots,\xi_n)$ 是优效估计量,当且仅当 $U(\xi_1,\xi_2,\cdots,\xi_n)$ 的方差为:

$$D(U)=\frac{[g'(\theta)]^2}{nE\left[\frac{\partial}{\partial\theta}\ln f(\xi;\theta)\right]^2}. \tag{7.2.14}$$

这是连续型随机变数 ξ 的情形;当 ξ 为离散型时,则为

$$D(U)=\frac{[g'(\theta)]^2}{nE\left[\frac{\partial}{\partial\theta}\ln P(\xi;\theta)\right]^2}. \tag{7.2.15}$$

定义 7.2.6(有效率)　设 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 为参数 $g(\theta)$ 的任一正规无偏估计量,$U(\xi_1,\xi_2,\cdots,\xi_n)$ 为 $g(\theta)$ 的优效估计量,则称

$$e_n=\frac{D(U)}{D(T)} \tag{7.2.16}$$

为正规无偏估计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 的有效率.

易见,有效率 e_n 满足关系式:$0\leqslant e_n\leqslant 1$. 如果正规无偏估计量 T 的有效率 e_n 满足:

$$\lim_{n\to\infty}e_n=1, \tag{7.2.17}$$

则称 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 为参数 θ 的渐近优效估计量.

专业课辅导+真题资料+微信公众号：432应统书院

可以证明:若

$$\frac{\partial}{\partial\theta}\int\frac{\partial f(x;\theta)}{\partial\theta}\mathrm{d}x=\int\frac{\partial^2 f(x;\theta)}{\partial\theta^2}\mathrm{d}x,$$

则(7.2.12)式中的 $I(\theta)$ 等价于

$$I(\theta)=-E\left[\frac{\partial^2}{\partial\theta^2}\ln f(\xi;\theta)\right]. \tag{7.2.18}$$

例 7.2.5 设总体 ξ 服从 $\Gamma(\alpha,\beta)$,其中 α 为已知常数,试求参数 β 的优效估计.

解 由例 7.1.3 及例 7.1.11 知,$\hat{\beta}=\frac{\bar{\xi}}{\alpha+1}$,我们来证 $\hat{\beta}=\frac{\bar{\xi}}{\alpha+1}$ 为 β 的优效估计量.由于

$$E(\xi)=\beta(\alpha+1),D(\xi)=\beta^2(\alpha+1),$$

所以

$$E(\bar{\xi})=\beta(\alpha+1),D(\bar{\xi})=\frac{\beta^2(\alpha+1)}{n},$$

则有

$$E(\hat{\beta})=\beta,D(\hat{\beta})=\frac{\beta^2}{n(\alpha+1)}.$$

可以验证,估计量 $\hat{\beta}$ 是正规的.现计算

$$\begin{aligned}&n\int_0^\infty\left[\frac{\partial}{\partial\beta}\ln f(x,\beta)\right]^2 f(x,\beta)\mathrm{d}x\\&=n\int_0^\infty\left(\frac{x}{\beta^2}-\frac{\alpha+1}{\beta}\right)^2 f(x,\beta)\mathrm{d}x\\&=n\beta^{-4}E[\xi-E(\xi)]^2=n\beta^{-4}D(\xi)=n\beta^{-2}(\alpha+1).\end{aligned}$$

今取 $g(\beta)=\beta$.由(7.2.14)式知,β 的优效估计的方差应该等于

$$\frac{1}{nE\left[\frac{\partial}{\partial\beta}\ln f(x,\beta)\right]^2}=\frac{\beta^2}{n(\alpha+1)}.$$

由此可见,$\hat{\beta}=\frac{\bar{\xi}}{\alpha+1}$ 为 β 的优效估计.

专业课辅导+真题资料+微信公众号：432应统书院

例 7.2.6 设总体 ξ 服从正态 $N(a,\sigma)$，a 及 σ 都是未知参数，$\xi_1,\xi_2,\cdots,\xi_n$ 为样本，则 S^{*2} 为 σ^2 的渐近优效估计量，其中

$$S^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2.$$

证 由定理 6.3.1 知，$\frac{(n-1)S^{*2}}{\sigma^2}\sim\chi^2_{(n-1)}$，于是

$$E\left[\frac{(n-1)S^{*2}}{\sigma^2}\right]=n-1,\ D\left[\frac{(n-1)S^{*2}}{\sigma^2}\right]=2(n-1),$$

因而

$$E(S^{*2})=\sigma^2,D(S^{*2})=\frac{2\sigma^4}{n-1}.$$

由不等式(7.2.12)的右端，计算优效估计量的方差. 在本例中，$\theta=\sigma^2$，考虑 $g(\theta)=\theta$ 的情形，因而(7.2.12)式右端的分子为 1，现在来求分母.

$$\ln f(x;a,\sigma^2)=-\frac{(x-a)^2}{2\sigma^2}-\ln\sqrt{2\pi\sigma^2},$$

$$\frac{\partial}{\partial\sigma^2}\ln f(x;a,\sigma^2)=\frac{(x-a)^2}{2\sigma^4}-\frac{1}{2\sigma^2},$$

$$\frac{\partial}{\partial\sigma^2}\left[\frac{\partial}{\partial\sigma^2}\ln f(x;a,\sigma^2)\right]=-\frac{(x-a)^2}{\sigma^6}+\frac{1}{2\sigma^4},$$

所以

$$-E\left\{\frac{\partial}{\partial\sigma^2}\left[\frac{\partial}{\partial\sigma^2}\ln f(\xi;a,\sigma^2)\right]\right\}=\frac{1}{2\sigma^4}.$$

由此可见，参数 σ^2 的优效估计量的方差为 $\frac{2\sigma^4}{n}$，那么 S^{*2} 的有效率 $e_n=\frac{n-1}{n}$，可见 S^{*2} 为 σ^2 的渐近优效估计量.

三、相合性

在无偏估计类中，我们以估计量的方差大小作为衡量估计量为“最优”的准则，作了较为充分的讨论. 但是，无偏估计类中方差

专业课辅导+真题资料+微信公众号：432应统书院

为最小或较小的估计量，不一定比某个有偏差的估计量的方差来得小. 无偏与有偏是反映估计量 $\hat{\theta}(\xi_1,\xi_2,\cdots,\xi_n)$ 的数学期望是否等于被估计的真参数值 θ；方差 $D(\hat{\theta})$ 的大小反映 $\hat{\theta}$ 的观察值以其 $E(\hat{\theta})$ 为中心的离散程度. 一个估计量，它依赖于样本观察值而求得的数值，即使其平均值等于 θ，但离散程度很大，那么这个估计量用来估计 θ 时，仍然不大理想. 因此，人们希望在偏差性（有偏或无偏）与离散性（方差的大小）两者兼顾的原则下来建立估计量为"最优"的准则，先讲两个例题，后引进相合性的概念.

例 7.2.7 设总体 ξ 在 $(0,\theta)$ 上服从均匀分布，密度函数为

$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta}, & 0<x<\theta;\\ 0, & \text{其他}.\end{cases}$$

我们可以求出参数 θ 的矩法估计量 $\hat{\theta}_1=2\bar{\xi}$，且有

$$E(\hat{\theta}_1)=E(2\bar{\xi})=\theta,$$

$$D(\hat{\theta}_1)=D(2\bar{\xi})=4D(\bar{\xi})$$

$$=\frac{4}{n}D(\xi)=\frac{\theta^2}{3n}.$$

我们又可求出参数 θ 的极大似然法估计量 $\hat{\theta}_2=\max\limits_{1\leqslant i\leqslant n}(\xi_1,\xi_2,\cdots,\xi_n)$，且有

$$E(\hat{\theta}_2)=\frac{n\theta}{n+1}, \tag{7.2.19}$$

$$D(\hat{\theta}_2)=\frac{n\theta^2}{(n+1)^2(n+2)}. \tag{7.2.20}$$

可见 $\hat{\theta}_2(\xi_1,\xi_2,\cdots,\xi_n)$ 是 θ 的渐近无偏估计量. (7.2.19) 及 (7.2.20) 两式的证明留作练习.

从上述等式看到，$\hat{\theta}_1$ 是无偏估计量，$\hat{\theta}_2$ 是有偏估计量，但 $\hat{\theta}_2$ 的方差比 $\hat{\theta}_1$ 来得小，这是因为在 $n\geqslant 1$ 时，有

专业课辅导+真题资料+微信公众号：432应统书院

$$\frac{n}{(n+1)^2(n+2)}<\frac{1}{3n},$$

特别当样本容量 n 很大时更为显著，因为这时

$$\lim_{n\to\infty}\frac{D(\hat{\theta}_2)}{D(\hat{\theta}_1)}=\lim_{n\to\infty}\frac{3n^2}{(n+1)^2(n+2)}=0.$$

例 7.2.8 设总体 ξ 服从正态 $N(a,\sigma)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为样本，a 及 σ 都为未知参数，我们用统计量

$$S_k^2=\frac{1}{k}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2 \tag{7.2.21}$$

作为 σ^2 的估计量，其中 $\bar{\xi}$ 为样本平均值. 由定理 6.3.1 知

$$\frac{kS_k^2}{\sigma^2}\sim\chi^2_{(n-1)},$$

所以

$$E\left(\frac{kS_k^2}{\sigma^2}\right)=n-1,$$

$$D\left(\frac{kS_k^2}{\sigma^2}\right)=2(n-1).$$

因而

$$E(S_k^2)=\frac{n-1}{k}\sigma^2, \tag{7.2.22}$$

$$D(S_k^2)=\frac{2(n-1)}{k^2}\sigma^4. \tag{7.2.23}$$

特别地，当 $k=n-1$ 时，则有

$$E(S_{n-1}^2)=\sigma^2,$$

$$D(S_{n-1}^2)=\frac{2\sigma^4}{n-1}.$$

从(7.2.22)及(7.2.23)两式看到，估计量 S_k^2 的偏差 $b_k=\left(\frac{n-1}{k}-1\right)\sigma^2$，当 k 取值自 $n-1$ 增大时，偏差 b_k 的绝对值增大，而

专业课辅导+真题资料+微信公众号：432应统书院

方差却减小. 例如

$$E(S_{n+1}^2)=\frac{n-1}{n+1}\sigma^2,$$

$$D(S_{n+1}^2)=\frac{2(n-1)}{(n+1)^2}\sigma^4,$$

$D(S_{n+1}^2)<D(S_{n-1}^2)$，但 S_{n+1}^2 是 σ^2 的有偏估计量. 试问：怎样衡量 S_{n-1}^2 与 S_{n+1}^2 哪个为 σ^2 的较优估计量？

对于(7.2.21)式所定义的 σ^2 的估计量 S_k^2，我们可计算估计量 S_k^2 与被估计参数 σ^2 之间的均方误差为：

$$\begin{aligned}E[S_k^2-\sigma^2]^2&=E[S_k^2-E(S_k^2)+E(S_k^2)-\sigma^2]^2\\&=D(S_k^2)+b_k^2,\end{aligned}\tag{7.2.24}$$

其中

$$\begin{aligned}&E\{[S_k^2-E(S_k^2)][E(S_k^2)-\sigma^2]\}\\&=[E(S_k^2)-\sigma^2]E[S_k^2-E(S_k^2)]=0.\end{aligned}$$

于是

$$E[S_k^2-\sigma^2]^2=\frac{2(n-1)}{k^2}\sigma^4+\left(\frac{n-1}{k}-1\right)^2\sigma^4.\tag{7.2.25}$$

当 $k=n-1$ 时，有

$$E[S_{n-1}^2-\sigma^2]^2=\frac{2}{n-1}\sigma^4,$$

当 $k=n+1$ 时，有

$$E[S_{n+1}^2-\sigma^2]^2=\frac{2}{n+1}\sigma^4,$$

于是

$$E[S_{n+1}^2-\sigma^2]^2<E[S_{n-1}^2-\sigma^2]^2.$$

(7.2.24)式的均方误差，它由估计量的方差及偏差平方两项组成，当左端达到最小时，其方差及偏差的绝对值都有可能很小，特别是在左端趋于零时，右端的两个非负项都分别趋于零. 在均方误差的数学期望为最小的原则下，参数 σ^2 的估计量取 S_{n+1}^2 较 S_{n-1}^2

专业课辅导+真题资料+微信公众号：432应统书院

为优，而这个原则是在兼顾了方差及偏差绝对值都相对地小的情况下确定的.

定义 7.2.7 相合性估计. 设统计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 为待估函数 $g(\theta)$ 的估计量，若对任一 $\varepsilon>0$，有

$$\lim_{n\to\infty}P\{|T(\xi_1,\xi_2,\cdots,\xi_n)-g(\theta)|\leqslant\varepsilon\}=1,$$

则称 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 为 $g(\theta)$ 的弱相合估计. 如果下式成立：

$$P\{\lim_{n\to\infty}T(\xi_1,\xi_2,\cdots,\xi_n)=g(\theta)\}=1,$$

则称 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 为 $g(\theta)$ 的强相合估计.

我们知道，对于 $g(\theta)$ 的任何估计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$，下式都成立：

$$E[T(\xi_1,\xi_2,\cdots,\xi_n)-g(\theta)]^2=D(T)+b_n^2, \tag{7.2.26}$$

由于对任给 $\varepsilon>0$，下式成立

$$P\{|\hat{\theta}-\theta|>\varepsilon\}\leqslant\frac{1}{\varepsilon^2}E(\hat{\theta}-\theta)^2,$$

所以当 $n\to\infty$ 时，若方差 $D(T)$ 及偏差 b_n 都趋于零，则有

$$\lim_{n\to\infty}P\{|T(\xi_1,\xi_2,\cdots,\xi_n)-g(\theta)|>\varepsilon\}=0.$$

因此，$T(\xi_1,\xi_2,\cdots,\xi_n)$ 为 $g(\theta)$ 的弱相合估计.

如果估计量 $\hat{\theta}=T(\xi_1,\xi_2,\cdots,\xi_n)$ 为参数 θ 的相合性估计量，则 $\hat{\theta}$ 依概率收敛于 θ，即对任给 $\varepsilon>0$，有：

$$\lim_{n\to\infty}P\{|\hat{\theta}-\theta|>\varepsilon\}=0, \tag{7.2.27}$$

或等价地有

$$\lim_{n\to\infty}P\{|\hat{\theta}-\theta|\leqslant\varepsilon\}=1.$$

在有些教材中，常称 $\hat{\theta}$ 为参数 θ 的一致估计量.

相合性或一致性的概念，是在极限的意义上引进的，适用于大样本情形.

* §7.3 充分性与完备性

在前两节中引进了求参数点估计的矩法与极大似然法，讨论

专业课辅导+真题资料+微信公众号：432应统书院

了点估计量的无偏性与优效性，这是点估计的最基本知识. 从这两节的内容讨论中，人们会发现还有一些问题，有必要给予进一步的探讨. 如对于所获得的容量为 n 的样本 $\xi_1,\xi_2,\cdots,\xi_n$，常用样本的一个函数作为参数的点估计，试问这样建立点估计有没有充分利用样本所获得的信息？也就是说，用矩法与极大似然法确定参数的点估计，在哪些情况下充分利用了样本获得的信息？又如，怎样寻找参数的最小方差无偏估计及优效估计？在这节里，将引进充分性与完备性的概念，从而完整地解决寻找最小方差无偏估计及优效估计的问题.

充分性与完备性是数理统计学中很重要的概念，在估计理论与假设检验中都很有用.

一、充分性

定义 7.3.1（充分统计量） 设总体 ξ 的密度函数为 $f(x,\theta)$，θ 为未知参数，$\theta\in\Omega$，$\xi_1,\xi_2,\cdots,\xi_n$ 为样本，$T(\xi_1,\xi_2,\cdots,\xi_n)$ 为不带有未知参数的统计量，若在给定统计量 $T(x_1,x_2,\cdots,x_n)=t$ 时，$(\xi_1,\xi_2,\cdots,\xi_n)$ 的条件密度函数 $f(x_1,x_2,\cdots,x_n|t)$ 与 θ 无关，则称 $T=T(\xi_1,\xi_2,\cdots,\xi_n)$ 为 θ 的充分统计量（本定义也适合离散型）.

例 7.3.1 设总体 ξ 服从两点分布，即有

$$P\{\xi=x\}=\begin{cases}p, & 当\ x=1;\\ 1-p, & 当\ x=0,\end{cases}$$

其中 $0<p<1$. $\xi_1,\xi_2,\cdots,\xi_n$ 为样本，则样本平均值 $\bar{\xi}=\frac{1}{n}\sum_{i=1}^{n}\xi_i$ 为参数 p 的充分统计量.

证 由于

$$\bar{\xi}=\frac{1}{n}\sum_{i=1}^{n}\xi_i=\frac{k}{n},$$

其中 k 为样本 $\xi_1,\xi_2,\cdots,\xi_n$ 中取值为 1 的个数，则样本关于 $\bar{\xi}$ 的条件概率函数为

$$P\left\{\xi_1=\varepsilon_1,\cdots,\xi_n=\varepsilon_n\,|\,\bar{\xi}=\frac{k}{n};p\right\}$$

专业课辅导+真题资料+微信公众号：432应统书院

$$=\frac{P\left\{\xi_1=\varepsilon_1,\cdots,\xi_n=\varepsilon_n;\bar{\xi}=\frac{k}{n};p\right\}}{P\left\{\bar{\xi}=\frac{k}{n};p\right\}}$$

$$=\frac{P\left\{\xi_1=\varepsilon_1,\cdots,\xi_n=\varepsilon_n;\sum_{i=1}^{n}\xi_i=k;p\right\}}{P\left\{\sum_{i=1}^{n}\xi_i=k;p\right\}}$$

$$=\frac{p^k(1-p)^{n-k}}{C_n^k p^k(1-p)^{n-k}}=\frac{1}{C_n^k},$$

其中 $\varepsilon_i=1$ 或 $0,i=1,2,\cdots,n$. 于是在给定统计量$\bar{\xi}$的取值为$\frac{k}{n}$时，样本的条件分布与参数 p 无关，因而$\bar{\xi}$为 p 的充分统计量.

充分统计量的直观背景是很有意思的. 我们知道，容量为 n 的样本 $\xi_1,\xi_2,\cdots,\xi_n$ 的观察值 $x_1,x_2,\cdots,x_n$ 中每个 x_i 都反映了总体 ξ 的特征，而在用样本平均$\bar{\xi}$估计 $E(\xi)$ 时，只用$\bar{\xi}$的观察值 $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ 作为它的估计值. 就是说，对 n 个数据 $x_1,x_2,\cdots,x_n$ 进行了“压缩”和“精简”，试问这样做是否合理？亦即$\bar{\xi}$是否已经包含了样本关于 $E(\xi)$ 的全部信息？或者说$\bar{\xi}$作为 $E(\xi)$ 的估计是否充分利用了样本关于 $E(\xi)$ 的全部信息？从几何学上来看，样本的观察值$(x_1,x_2,\cdots,x_n)$是 n 维的样本值域 $\mathscr{X}$ 中的一点，而当已知$\bar{\xi}=\bar{x}$时，样本$(\xi_1,\xi_2,\cdots,\xi_n)$是 $n-1$ 维空间的点了，就是说在已知$\bar{\xi}=\bar{x}$时，进一步知道$(\xi_1,\xi_2,\cdots,\xi_n)$的各自取何值，对于被估计的 $E(\xi)$ 来说，能否获得更多的信息？为说明简便，我们就 $n=2$ 的特殊情形，举例说明.

例 7.3.2 设总体 ξ 服从正态 $N(a,1)$，要求估计 a. ξ_1,ξ_2 为总体 ξ 的样本，用$\bar{\xi}=\frac{1}{2}(\xi_1+\xi_2)$估计 a. 此时，样本值域 $\mathscr{X}$ 可用平

专业课辅导+真题资料+微信公众号：432应统书院

面来表示，那么在给定$\bar{\xi}$的取值为y时，$\frac{1}{2}(x_1+x_2)=y$，表示该平面(x_1,x_2)上的一条直线（如图 7.3.1）. 在这平面上，(ξ_1,ξ_2)服从二维正态分布，其密度函数为

$$f(x_1,x_2;a)=\frac{1}{2\pi}\exp\left[-\frac{1}{2}\sum_{i=1}^{2}(x_i-a)^2\right],$$

图 7.3.1

样本平均值$\bar{\xi}$服从正态$N\left(a,\frac{1}{\sqrt{2}}\right)$，其密度函数为

$$f(\bar{x};a)=\frac{1}{\sqrt{\pi}}\exp[-(\bar{x}-a)^2].$$

在直线$y=\frac{1}{2}(x_1+x_2)$上，(ξ_1,ξ_2)的分布可以用条件分布密度函数

$$f(x_1,x_2\mid\bar{x}=y)=\begin{cases}\dfrac{f(x_1,x_2)}{f(y)}, & 当\ \bar{x}=y;\\ 0, & 当\ \bar{x}\neq y\end{cases}$$

来描述，其中$\bar{x}$为$\bar{\xi}$的观察值，而y为$\bar{\xi}$的给定值. 所以

$$f(x_1,x_2\mid\bar{x}=y)$$

专业课辅导+真题资料+微信公众号：432应统书院

$$=\begin{cases}\dfrac{1}{2\sqrt{\pi}}\exp\left[-\dfrac{1}{2}\sum\limits_{i=1}^{2}(x_i-y)^2\right], & 当\ \bar{x}=y;\\ 0, & 当\ \bar{x}\neq y,\end{cases}$$

其中 $\bar{x}=\frac{1}{2}(x_1+x_2)$.

从上面条件密度函数的表示式中看到，它与参数 a 无关，这表明在已知 $\bar{x}=y$ 时，进一步知道样本观察值 (x_1,x_2) 在这条直线上的具体位置（即 x_1 及 x_2 各取何值使得 $\bar{x}=y$），丝毫没有增加关于参数 a 的任何新的信息. 换句话说，$\bar{x}$ 已充分地从样本 ξ_1,ξ_2 中提取了所包含的关于参数 a 的全部信息，因而称 $\bar{\xi}=\frac{1}{2}\sum\limits_{i=1}^{2}\xi_i$ 为参数 a 的充分统计量.

若用参数 θ 的充分统计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 作为 θ 的点估计，则称 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 为参数 θ 的充分估计量.

定理 7.3.1（费希尔-奈曼因子分解定理） 若总体 ξ 为连续型随机变量，密度函数为 $f(x,\theta)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为样本，则统计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 为参数 θ 的充分估计量的充要条件是：样本的联合密度函数可表示为

$$\prod_{i=1}^{n}f(x_i,\theta)=f_T(t,\theta)h(x_1,x_2,\cdots,x_n),\qquad(7.3.1)$$

其中 $f_T(t,\theta)$ 是仅通过 $T(x_1,x_2,\cdots,x_n)=t$ 而依赖于样本且与 θ 有关的非负函数，$h(x_1,x_2,\cdots,x_n)$ 是样本 $\xi_1,\xi_2,\cdots,\xi_n$ 的非负函数但与参数 θ 无关. 特别地，$f_T(t,\theta)$ 可为统计量 T 的密度函数.

若总体 ξ 为离散型随机变量，则密度函数代之以概率函数 $P(x,\theta)$，有类似于(7.3.1)的分解式.

注：本节各定理的证明，参看书目[5].

如果 θ 是参数向量，如正态 $N(a,\sigma)$ 中的 a 及 σ^2 都是未知参数，则记 $\theta=(a,\sigma^2)$；统计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 是随机向量，而且定理 7.3.1 的条件相应地成立，则称 T 为关于 θ 的联合充分统计量.

专业课辅导+真题资料+微信公众号：432应统书院

例 7.3.3 设总体 ξ 服从如下均匀分布

$$f(x,\theta)=\begin{cases}\dfrac{1}{\theta}, & 0<x<\theta;\\ 0, & \text{其他}.\end{cases}$$

$\xi_1,\xi_2,\cdots,\xi_n$ 为样本，求参数 θ 的充分估计量.

解 样本的联合密度函数为

$$\prod_{i=1}^{n} f(x_i,\theta)=\begin{cases}\dfrac{1}{\theta^n}, & \text{当 } 0\leqslant x_i^*\leqslant x_n^*\leqslant\theta;\\ 0, & \text{其他}\end{cases}$$

$$=\frac{1}{\theta^n}\chi[0\leqslant x_i^*\leqslant x_n^*\leqslant\theta],$$

其中 χ 表示集合 $[0\leqslant x_i^*\leqslant x_n^*\leqslant\theta]$ 的示性函数. 由定理 7.3.1 知，统计量 $T(\xi_1,\xi_2,\cdots,\xi_n)=\max\limits_{1\leqslant i\leqslant n}\xi_i=\xi_n^*$ 为 θ 的充分估计量.

例 7.3.4 设总体 ξ 服从泊松分布 $P(x;\lambda)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为样本，求参数 λ 的充分估计量.

解 样本的联合概率函数为

$$P(x_1,x_2,\cdots,x_n;\lambda)=\frac{e^{-n\lambda}}{x_1!\cdots x_n!}\lambda^{\sum\limits_{i=1}^{n}x_i}$$

取

$$T(\xi_1,\xi_2,\cdots,\xi_n)=\frac{1}{n}\sum_{i=1}^{n}\xi_i=\bar{\xi};$$

$$h(x_1,x_2,\cdots,x_n)=\frac{1}{x_1!\cdots x_n!};$$

$$f_T(t,\lambda)=\lambda^{nt}e^{-n\lambda},$$

则有

$$P(x_1,x_2,\cdots,x_n;\lambda)=f_T(t,\lambda)h(x_1,x_2,\cdots,x_n).$$

因而统计量 $T=\bar{\xi}$ 为参数 λ 的充分估计量.

例 7.3.5 设总体 ξ 服从正态 $N(a,\sigma)$，其中 a 及 σ 都是未知参数，求 $\theta=(a,\sigma^2)$ 的充分估计量.

专业课辅导+真题资料+微信公众号：432应统书院

解 样本的联合密度函数为

$$\prod_{i=1}^{n} f(x_i,\theta)=\frac{1}{(\sqrt{2\pi}\sigma)^n}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-a)^2\right]$$

$$=\frac{1}{(\sqrt{2\pi}\sigma)^n}\exp\left\{-\frac{n}{2\sigma^2}[s^2+n(\bar{x}-a)^2]\right\}.$$

记二维统计量 $T(\xi_1,\xi_2,\cdots,\xi_n)=(\bar{\xi},s^2)$，由定理 7.3.1 知，$(\bar{\xi},s^2)$ 是参数向量 $\theta=(a,\sigma^2)$ 的充分估计量.

定理 7.3.2 参数 θ 的正规无偏估计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 是优效估计量的充要条件为

(1) $T(\xi_1,\xi_2,\cdots,\xi_n)$ 为 θ 的充分估计量；

(2) 统计量 T 的密度函数(或概率函数，视总体 ξ 为连续型或离散型随机变量而定)概率为 1 地满足下述方程：

$$\frac{\partial \ln f_T(t,\theta)}{\partial\theta}=C(\theta)[T(x_1,x_2,\cdots,x_n)-\theta], \tag{7.3.2}$$

其中 $C(\theta)$ 与样本无关，$T(x_1,x_2,\cdots,x_n)$ 与 θ 无关.

证 必要性 假定 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 是 θ 的优效估计量，由(7.2.13)式有：

$$D_\theta(T)=\frac{1}{nI(\theta)}.$$

我们来证条件(1)和(2)是必要的. 在不等式(7.2.13)的证明过程中看到，等号成立的条件是 T 与 Z 的相关系数 $|\rho|=1$，即 T 与 Z 之间的线性关系式

$$Z=C(\theta)(T-\theta)$$

几乎处处成立，其中 $C(\theta)$ 只是与参数 θ 有关的数，不是统计量，也就是

$$Z=\frac{\partial \ln L(\xi_1,\xi_2,\cdots,\xi_n;\theta)}{\partial\theta}=C(\theta)(T-\theta).$$

将上式在 (θ_0,θ) 上关于 θ 积分，θ_0 为参数空间 Ω 中的任一固定值，得到

专业课辅导+真题资料+微信公众号：432应统书院

$$L(\xi_1,\xi_2,\cdots,\xi_n;\theta)=\exp\{T\cdot\varphi_1(\theta)-\varphi_2(\theta)\},$$

其中

$$\varphi_1(\theta)=\int_{\theta_0}^{\theta}C(\theta)\mathrm{d}\theta,\varphi_2(\theta)=\int_{\theta_0}^{\theta}\theta C(\theta)\mathrm{d}\theta.$$

由定理 7.3.1 知，$T(\xi_1,\xi_2,\cdots,\xi_n)$ 为 θ 的充分估计量. 因而

$$\frac{\partial\ln f_T(t,\theta)}{\partial\theta}=C(\theta)[T(x_1,x_2,\cdots,x_n)-\theta].$$

充分性　如果条件(1)及(2)满足，则由定理 7.3.1 知

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left\{\frac{\partial}{\partial\theta}\left[\ln\prod_{i=1}^{n}f(x_i,\theta)\right]\right\}^2\prod_{i=1}^{n}[f(x_i,\theta)\mathrm{d}x_i]$$

$$=C^2(\theta)\int_{-\infty}^{\infty}(t-\theta)^2f_T(t,\theta)\mathrm{d}t=C^2(\theta)D_\theta(T),$$

$$1=\int_{-\infty}^{\infty}t\frac{\partial}{\partial\theta}[f_T(t,\theta)]\mathrm{d}t=\int_{-\infty}^{\infty}(t-\theta)\frac{\partial}{\partial\theta}[f_T(t,\theta)]\mathrm{d}t$$

$$=\int_{-\infty}^{\infty}(t-\theta)\frac{\partial}{\partial\theta}[\ln f_T(t,\theta)]f_T(t,\theta)\mathrm{d}t$$

$$=C(\theta)D_\theta(T),$$

因此，

$$nI(\theta)=\frac{1}{D_\theta(T)}.$$

定理 7.3.3　设总体 ξ 的密度函数 $f(x;\theta)$ 中只含一个未知参数 θ，如果 θ 的充分估计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 存在，且似然方程

$$\frac{\partial\ln L(\xi_1,\xi_2,\cdots,\xi_n;\theta)}{\partial\theta}=0 \tag{7.3.3}$$

有解，则其解一定是 $t=T(x_1,x_2,\cdots,x_n)$ 的函数.

证　如果 T 是 θ 的充分估计量，则由(7.3.1)式知

$$L(\xi_1,\xi_2,\cdots,\xi_n;\theta)=\prod_{i=1}^{n}f(\xi_i;\theta)=f_T(t,\theta)h(\xi_1,\xi_2,\cdots,\xi_n),$$

且

$$\frac{\partial\ln L(\xi_1,\xi_2,\cdots,\xi_n;\theta)}{\partial\theta}=\frac{\partial\ln f_T(t,\theta)}{\partial\theta}.$$

专业课辅导+真题资料+微信公众号：432应统书院

因而(7.3.3)的解一定是 t 的函数,即依赖于统计量 T 的观察值 t.

这个定理表明,如果参数 θ 的充分估计存在,且似然方程有解并为 θ 的极大似然估计,那么这个极大似然估计具有充分估计量的优良性.

定理 7.3.4 如果 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 是参数 θ 的优效估计量,那么 $t=T(\xi_1,\xi_2,\cdots,\xi_n)$ 在满足 $f_T(t,\theta)>0$ 的区域内,概率为 1 地都是似然方程的解.

证 由定理 7.3.2 知,$f_T(t,\theta)$ 满足(7.3.2)式,并有

$$\frac{\partial \ln L(\theta)}{\partial \theta}=C(\theta)(T-\theta), \tag{7.3.4}$$

所以统计量 T 是似然方程的解.

这个定理表明,如果参数 θ 的优效估计存在,且似然方程有解并为 θ 的极大似然估计,那么这个极大似然估计具有优效估计的优良性.

在这一段中,首先看到常见的分布如泊松分布、均匀分布及正态分布、二项分布,其未知参数的极大似然法估计量都是充分估计量,接着证明了几个定理.从几个定理的结论看到,优效估计量一定是充分估计量,极大似然法估计量一定是充分估计量的函数.可以证明,充分估计量的单值函数也是一个充分估计量.下面再介绍一个定理,这个定理阐明寻找优效估计量与最优无偏估计量的途径.

定理 7.3.5 设总体 ξ 的分布函数为 $f(x;\theta),\theta\in\Omega,\xi_1,\xi_2,\cdots,\xi_n$ 为其样本,如果

(1) $\hat{\theta}$ 是 θ 的无偏估计量且方差有限,即对一切 $\theta\in\Omega$,

$$E_\theta(\hat{\theta})=\theta,\quad D_\theta(\hat{\theta})<\infty;$$

(2) T 是 θ 的充分统计量.记

$$\hat{\theta}^*=E_\theta(\hat{\theta}|T), \tag{7.3.5}$$

它是 $\hat{\theta}$ 关于 T 的条件数学期望.

则有

专业课辅导+真题资料+微信公众号：432应统书院

(1) $\hat{\theta}^*$ 是不含有 θ 的并为 θ 的无偏估计量，即对一切 $\theta \in \Omega$，有 $E_\theta(\hat{\theta}^*)=\theta$；

(2) $D_\theta(\hat{\theta}^*) \leqslant D_\theta(\hat{\theta})$，

当且仅当 $P_\theta\{\hat{\theta}=\hat{\theta}^*\}=1$ 时，等号成立.

证 利用上述几个定理的结论，可以扼要地叙述这个定理的证明. 由于估计量 $\hat{\theta}$ 不带有未知参数 θ，样本在充分统计量 $T=t$ 的条件下的条件分布与 θ 无关，因而 $\hat{\theta}^*=E_\theta(\hat{\theta}|T)$ 不带有未知参数 θ，是一统计量，可作为 θ 估计量. 从第三章讲到的条件数学期望性质知，对一切 $\theta \in \Omega$

$$E_\theta(\hat{\theta}^*)=E_\theta[E_\theta(\hat{\theta}|T)]=E_\theta(\hat{\theta})=\theta.$$

另一方面，

$$\begin{aligned}D_\theta(\hat{\theta})&=E_\theta[(\hat{\theta}-\hat{\theta}^*)+(\hat{\theta}^*-\theta)]^2\\&=E_\theta(\hat{\theta}-\hat{\theta}^*)^2+2E_\theta(\hat{\theta}-\hat{\theta}^*)(\hat{\theta}^*-\theta)+D_\theta(\hat{\theta}^*)\\&=E_\theta(\hat{\theta}-\hat{\theta}^*)^2+D_\theta(\hat{\theta}^*)\geqslant D_\theta(\hat{\theta}^*),\end{aligned}$$

其中等号成立的充要条件为

$$E_\theta(\hat{\theta}-\hat{\theta}^*)^2=0,$$

即有

$$P_\theta\{\hat{\theta}=\hat{\theta}^*\}=1.$$

(在推导上述不等式过程中，应用了关系式：

$$E_\theta(\hat{\theta}-\hat{\theta}^*)(\hat{\theta}^*-\theta)=0,$$

这个等式分别在连续型与离散型的情况下直接可得到证明.)

二、完备性

由上所知，最优无偏估计量可局限在充分无偏估计类中寻找. 如果充分无偏估计量是唯一的，那么它就是最优无偏估计了. 试问，在什么情况下，它是唯一的呢？为此引进另一个重要的概念，即完备性.

专业课辅导+真题资料+微信公众号：432应统书院

定义 7.3.2(完备性统计量) 设总体 ξ 的分布函数为 $F(x;\theta)$，$\theta\in\Omega$，$g(\xi)$ 为一随机变量，如果对一切 $\theta\in\Omega$，

$$E_\theta[g(\xi)]=0$$

成立时，对于一切 $\theta\in\Omega$ 必有

$$P_\theta\{g(\xi)=0\}=1,$$

则称 $F(x,\theta)$ 是完备的. 若 $\xi_1,\xi_2,\cdots,\xi_n$ 为样本，统计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$ 的分布函数是完备的，则称 T 为完备统计量.

定理 7.3.6 设总体 ξ 的分布函数为 $F(x,\theta)$，$\theta\in\Omega$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，$T(\xi_1,\xi_2,\cdots,\xi_n)$ 是 θ 的充分完备统计量. 如果 θ 的无偏估计量存在，记为 $\hat{\theta}$，则 $\hat{\theta}^*=E(\hat{\theta}\mid T)$ 是唯一的最优无偏估计量.

证 设 $\hat{\theta}_1$ 及 $\hat{\theta}_2$ 是参数 θ 的任意两个无偏估计量，由定理 7.3.5 知，$E(\hat{\theta}_1\mid T)$ 及 $E(\hat{\theta}_2\mid T)$ 也是参数 θ 的无偏估计量，即对一切 $\theta\in\Omega$，有：

$$E_\theta[E(\hat{\theta}_1\mid T)]=E_\theta[E(\hat{\theta}_2\mid T)]=\theta,$$

而且对一切 $\theta\in\Omega$，有

$$D_\theta[E(\hat{\theta}_1\mid T)]\leqslant D_\theta(\hat{\theta}_1),$$

$$D_\theta[E(\hat{\theta}_2\mid T)]\leqslant D_\theta(\hat{\theta}_2).$$

由于对一切 $\theta\in\Omega$，有：

$$E_\theta[E(\hat{\theta}_1\mid T)-E(\hat{\theta}_2\mid T)]=0,$$

且 T 是完备统计量，所以对一切 $\theta\in\Omega$，

$$P_\theta\{E(\hat{\theta}_1\mid T)=E(\hat{\theta}_2\mid T)\}=1.$$

因此参数 θ 的充分无偏估计量(在几乎处处相等的意义下)是唯一的. 由定理 7.3.5 知，若 $\hat{\theta}$ 为参数 θ 的无偏估计量，则 $\hat{\theta}^*=E(\hat{\theta}\mid T)$ 是参数 θ 的最优无偏估计量.

对于一个估计量 $\hat{\theta}$，直接计算它的数学期望，可以判断它是否为参数 θ 的无偏估计量，但是要判断它是否为 θ 的充分完备估计

专业课辅导+真题资料+微信公众号：432应统书院

量，一般是较麻烦的. 现在引进一类很重要的分布族，它包含了常用的一些分布，如泊松分布、正态分布、指数分布、二项分布、Γ 分布等，对这类分布族，寻找参数的充分完备估计量是方便的.

定义 7.3.3（指数族分布） 设总体 ξ 的密度函数为 $f(x;\theta)$，$\theta\in\Omega$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，样本的联合密度函数具有形式

$$f(X;\theta)=C(\theta)\exp\{b(\theta)T(X)\}h(X), \tag{7.3.6}$$

其中 $X=(x_1,x_2,\cdots,x_n)$，$C(\theta)$、$b(\theta)$ 只与参数 θ 有关而与样本无关，$h(X)$、$T(X)$ 只与样本有关而不带未知参数 θ，则称 $f(x;\theta)$ 为指数族分布. 对于离散型随机变量，如果它的概率函数 $P\{x;\theta\}$ 具有(7.3.6)式所定义的形式，也同样称它为指数族分布.

如果分布函数为 $F(x;\theta_1,\theta_2,\cdots,\theta_k)$，即有 k 维参数向量 $\widetilde{\boldsymbol{\theta}}=(\theta_1,\theta_2,\cdots,\theta_k)$，则(7.3.6)式中的指数部分为 $\exp\left\{\sum_{j=1}^{k}b_j(\widetilde{\boldsymbol{\theta}})\cdot T_j(X)\right\}$ 形式，其中 $T_j(X)$ 为统计量，$j=1,2,\cdots,k$.

定理 7.3.7 设总体 ξ 的密度函数 $f(x;\widetilde{\boldsymbol{\theta}})$ 为指数族分布，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，联合密度函数具有如下形式：

$$f(X;\widetilde{\boldsymbol{\theta}})=C(\widetilde{\boldsymbol{\theta}})\exp\left\{\sum_{j=1}^{k}b_j(\widetilde{\boldsymbol{\theta}})T_j(X)\right\}h(X),$$

其中 $\widetilde{\boldsymbol{\theta}}=(\theta_1,\theta_2,\cdots,\theta_k)$，$X=(x_1,x_2,\cdots,x_n)$，$\widetilde{\boldsymbol{\theta}}\in\Omega$. 如果 Ω 中包含有一个 k 维矩形，而且 $\boldsymbol{B}=(b_1,b_2,\cdots,b_k)$ 的值域包含有一个 k 维开集，则

$$\boldsymbol{T}(X)=(T_1(X),T_2(X),\cdots,T_k(X))$$

是 k 维参数向量 $\widetilde{\boldsymbol{\theta}}$ 的充分完备估计量.

例 7.3.6 设总体 ξ 服从参数 $\lambda>0$ 的泊松分布 $P(x;\lambda)$，即有

$$P(x;\lambda)=\frac{\lambda^x}{x!}e^{-\lambda}\quad(x=0,1,2,\cdots).$$

$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，它的联合概率函数为

专业课辅导+真题资料+微信公众号：432应统书院

$$P\{\xi_1=x_1,\cdots,\xi_n=x_n;\lambda\}=\frac{\lambda^{\sum_{i=1}^{n}x_i}e^{-n\lambda}}{\prod_{i=1}^{n}(x_i!)}$$

$$=e^{-n\lambda}\exp\left\{\sum_{i=1}^{n}x_i\cdot\ln\lambda\right\}\frac{1}{\prod_{i=1}^{n}(x_i!)}.$$

与(7.3.6)式对比,有

$$C(\lambda)=e^{-n\lambda};$$

$$h(X)=\frac{1}{\prod_{i=1}^{n}(x_i!)};$$

$$T(X)=\bar{x};$$

$$b(\lambda)=n\ln\lambda.$$

因此,样本平均值$\bar{\xi}=\frac{1}{n}\sum_{i=1}^{n}\xi_i$是参数$\lambda$的充分完备估计量. 由于$\bar{\xi}$是$\lambda$的无偏估计量,因而$\bar{\xi}$是$\lambda$的最优无偏估计量.

例 7.3.7 设总体ξ服从两点分布,即有:

$$P\{\xi=x\}=\begin{cases}p, & x=1;\\ 1-p, & x=0,0<p<1.\end{cases}$$

$\xi_1,\xi_2,\cdots,\xi_n$为其样本,它的联合概率函数为

$$P\{\xi_1=x_1,\cdots,\xi_n=x_n;p\}$$

$$=p^{\sum_{i=1}^{n}x_i}(1-p)^{n-\sum_{i=1}^{n}x_i}$$

$$=(1-p)^n\left(\frac{p}{1-p}\right)^{\sum_{i=1}^{n}x_i}$$

$$=(1-p)^n\exp\left\{n\bar{x}\cdot\ln\left(\frac{p}{1-p}\right)\right\}.$$

与(7.3.6)对比,有

$$C(p)=(1-p)^n;$$

专业课辅导+真题资料+微信公众号：432应统书院

$$h(X)=1;$$

$$b(p)=n\ln\left(\frac{p}{1-p}\right);$$

$$\boldsymbol{T}(X)=\bar{x}.$$

因此，样本平均值 $\bar{\xi}=\frac{1}{n}\sum_{i=1}^{n}\xi_i$ 是参数 p 的充分完备估计量. 由于 $\bar{\xi}$ 是 p 的无偏估计量，因而 $\bar{\xi}$ 是 p 的最优无偏估计量.

例 7.3.8 设总体 ξ 服从均匀分布，其密度函数为

$$f(x;\theta)=\begin{cases}\frac{1}{\theta}, & 0<x<\theta;\\ 0, & \text{其他},\end{cases}$$

$\theta\in\Omega=(0,1)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，它的联合密度函数为

$$f(x_1,x_2,\cdots,x_n;\theta)=\begin{cases}\frac{1}{\theta^n}, & 0<x_1,x_2,\cdots,x_n<\theta;\\ 0, & \text{其他}.\end{cases}$$

这时不便于与(7.3.6)式对比寻找参数 θ 的充分完备估计量. 我们在例 7.3.3 中已指出 $T(\xi_1,\xi_2,\cdots,\xi_n)=\max\limits_{1\leqslant i\leqslant n}\xi_i=\xi_n^*$ 是参数 θ 的充分估计量，下面证明它也是 θ 的完备估计量.

样本极大值 $\xi_n^*=\max\limits_{1\leqslant i\leqslant n}\xi_i$ 的分布函数为

$$F_n(t)=P\{\xi_n^*<t\}=[P\{\xi<t\}]^n$$

$$=\begin{cases}0, & t\leqslant 0;\\ \left(\frac{t}{\theta}\right)^n, & 0<t\leqslant\theta;\\ 1, & t>\theta.\end{cases}$$

密度函数为

$$f_T(t)=\begin{cases}\frac{nt^{n-1}}{\theta^n}, & 0<t<\theta;\\ 0, & \text{其他}.\end{cases}$$

如果有函数 $g(t)$ 对于一切 $0<\theta<1$，使得

专业课辅导+真题资料+微信公众号：432应统书院

$$\int_0^\theta g(t)f_T(t;\theta)\,\mathrm{d}t=\frac{n}{\theta^n}\int_0^\theta g(t)t^{n-1}\mathrm{d}t\equiv 0,$$

我们来证对一切 $0<\theta<1$，有

$$g(t)\equiv 0.$$

记

$$h(\theta)=\int_0^\theta g(t)t^{n-1}\mathrm{d}t=0,$$

则

$$h'(\theta)=g(\theta)\theta^{n-1}=0$$

几乎处处成立. 因为 $\theta\neq 0$，从而推知对一切 $0<\theta<1$，

$$g(\theta)=0.$$

于是 $g(t)$ 在 $(0,1)$ 上几乎处处为零，因此 ξ_n^* 是参数 θ 的完备估计量.

但是，$\xi_n^*=\max\limits_{1\leqslant i\leqslant n}\xi_i$ 不是 θ 的无偏估计量，而由 $E(\xi_n^*)=\dfrac{n\theta}{n+1}$ 知，$\dfrac{n+1}{n}\xi_n^*$ 为 θ 的无偏估计量. 因而

$$\frac{n+1}{n}\xi_n^*=\frac{n+1}{n}\max_{1\leqslant i\leqslant n}\xi_i$$

为 θ 的最优无偏估计量.

例 7.3.9 设总体 ξ 服从正态 $N(a,\sigma)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，它的联合密度函数为

$$\begin{aligned}f(x_1,x_2,\cdots,x_n)&=\frac{1}{(\sqrt{2\pi}\sigma)^n}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n(x_i-a)^2\right\}\\&=\frac{1}{(2\pi\sigma^2)^{n/2}}\mathrm{e}^{-\frac{na^2}{2\sigma^2}}\\&\quad\times\exp\left\{-\frac{n}{2\sigma^2}\left(\frac{1}{n}\sum_{i=1}^n x_i^2\right)+\frac{na}{\sigma^2}\bar{x}\right\}.\end{aligned}$$

与(7.3.6)式对比，$\boldsymbol{\theta}=(a,\sigma^2)$，有

专业课辅导+真题资料+微信公众号：432应统书院

$$C(\boldsymbol{\theta})=\frac{1}{(2\pi\sigma^2)^{n/2}}e^{-\frac{na^2}{2\sigma^2}};$$

$$\boldsymbol{T}(X)=\left(\bar{x},\frac{1}{n}\sum_{i=1}^{n}x_i^2\right),$$

$$h(X)=1.$$

因此，样本平均值$\bar{\xi}=\frac{1}{n}\sum_{i=1}^{n}\xi_i$及方差$S^2=\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2$是$a$及$\sigma^2$的充分完备估计量，$\bar{\xi}$及$S^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2$为$a$及$\sigma^2$的最优无偏估计量.

§7.4 区间估计

设总体ξ的分布函数$F(x;\theta)$中的θ为未知参数，$\xi_1,\xi_2,\cdots,\xi_n$为其样本，现建立两个统计量$T_1(\xi_1,\xi_2,\cdots,\xi_n)$及$T_2(\xi_1,\xi_2,\cdots,\xi_n)$，并满足不等式：

$$T_1(\xi_1,\xi_2,\cdots,\xi_n)<T_2(\xi_1,\xi_2,\cdots,\xi_n),$$

则称$[T_1,T_2]$为随机区间. 随机区间的端点及区间长度都是样本的函数，因而都是统计量.

定义7.4.1(区间估计)　设α为一给定的常数，满足$0<\alpha<1$，如果关系式：

$$P\{T_1\leqslant\theta\leqslant T_2\}=1-\alpha \tag{7.4.1}$$

成立，并用这个随机区间作为参数θ的估计，则称$[T_1,T_2]$是参数θ的置信水平为$1-\alpha$的区间估计，称α为显著性水平，称$T_2(\xi_1,\xi_2,\cdots,\xi_n)$、$T_1(\xi_1,\xi_2,\cdots,\xi_n)$分别为上、下置信限.

参数θ的区间估计的意义可解释为：建立统计量$T_1(\xi_1,\xi_2,\cdots,\xi_n)$及$T_2(\xi_1,\xi_2,\cdots,\xi_n)$后，$T_1$及$T_2$的观察值依赖于样本$\xi_1,\xi_2,\cdots,\xi_n$的观察值，现对样本作了$N$次观察，每次样本的观察值记为$(x_{1k},x_{2k},\cdots,x_{nk})$，$k=1,2,\cdots,N$，它对应于$t_{1k}$及$t_{2k}$观察值所组成的一个区间$[t_{1k},t_{2k}]$，这$N$个区间不一定都包含着参数$\theta$的

专业课辅导+真题资料+微信公众号：432应统书院

真值，有些区间包含着 θ 的真值而另一些区间不包含它，当(7.4.1)式成立时，随机区间 $[T_1,T_2]$ 包含着参数 θ 真值的频率近似为 $1-\alpha$. 此时，若认为"区间 $[T_1,T_2]$ 包含着参数 θ 的真值"，那么这种"认为"犯错误的概率为 α. 也就是说，随机区间 $[T_1,T_2]$ 以 $1-\alpha$ 的概率包含着参数 θ，这是参数 θ 的区间估计的本质所在. 注意，我们并非说参数 θ 以 $1-\alpha$ 的概率落入随机区间 $[T_1,T_2]$，因为参数 θ 是非随机的.

例 7.4.1 设总体 ξ 服从 $N(a,\sigma)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，试求 a 及 σ^2 的区间估计.

解 （一）a 的区间估计

(1) 在方差 $\sigma^2=\sigma_0^2$ 为已知时，求 a 的区间估计.

我们知道，$\bar{\xi}$ 服从正态 $N\left(a,\dfrac{\sigma_0}{\sqrt{n}}\right)$，因而

$$U=\frac{\bar{\xi}-a}{\sigma_0/\sqrt{n}}$$

服从正态 $N(0,1)$. 给定置信水平 $1-\alpha$，在标准正态分布表中查得临界值 u_α，使得

$$P\left\{\left|\frac{\bar{\xi}-a}{\sigma_0/\sqrt{n}}\right|\leqslant u_\alpha\right\}=1-\alpha,$$

亦即

$$P\left\{\bar{\xi}-u_\alpha\frac{\sigma_0}{\sqrt{n}}\leqslant a\leqslant\bar{\xi}+u_\alpha\frac{\sigma_0}{\sqrt{n}}\right\}=1-\alpha. \tag{7.4.2}$$

就是说，随机区间 $\left[\bar{\xi}-u_\alpha\dfrac{\sigma_0}{\sqrt{n}},\bar{\xi}+u_\alpha\dfrac{\sigma_0}{\sqrt{n}}\right]$ 是数学期望 a 的置信水平为 $1-\alpha$ 的区间估计.

对于不同的置信水平 $1-\alpha$，a 的区间估计随之不同. 如表 7.4.1 所示.

专业课辅导+真题资料+微信公众号：432应统书院

表 7.4.1

α	u_α	$\left[\bar{\xi}-u_\alpha\frac{\sigma_0}{\sqrt{n}},\bar{\xi}+u_\alpha\frac{\sigma_0}{\sqrt{n}}\right]$
0.01	2.58	$\left[\bar{\xi}-2.58\frac{\sigma_0}{\sqrt{n}},\bar{\xi}+2.58\frac{\sigma_0}{\sqrt{n}}\right]$
0.05	1.96	$\left[\bar{\xi}-1.96\frac{\sigma_0}{\sqrt{n}},\bar{\xi}+1.96\frac{\sigma_0}{\sqrt{n}}\right]$
0.10	1.65	$\left[\bar{\xi}-1.65\frac{\sigma_0}{\sqrt{n}},\bar{\xi}+1.65\frac{\sigma_0}{\sqrt{n}}\right]$

(2) 在方差 σ^2 未知时，求 a 的区间估计.

当方差 σ^2 未知时，可用样本方差 S^2 估计 σ^2. 由(6.3.7)式知道，

$$\sqrt{n-1}\frac{\bar{\xi}-a}{S}\sim t_{(n-1)}.$$

给定置信水平 $1-\alpha$，$0<\alpha<1$，在 t 分布表中查自由度 $n-1$ 对应的临界值 $t_{(n-1)}(\alpha)$，使得

$$P\left\{\left|\sqrt{n-1}\frac{\bar{\xi}-a}{S}\right|\geqslant t_{(n-1)}(\alpha)\right\}=\alpha,$$

亦即

$$P\left\{\bar{\xi}-t_{(n-1)}(\alpha)\frac{S}{\sqrt{n-1}}\leqslant a\leqslant\bar{\xi}+t_{(n-1)}(\alpha)\frac{S}{\sqrt{n-1}}\right\}=1-\alpha. \quad (7.4.3)$$

就是说，$\left[\bar{\xi}-t_{(n-1)}(\alpha)\frac{S}{\sqrt{n-1}},\bar{\xi}+t_{(n-1)}(\alpha)\frac{S}{\sqrt{n-1}}\right]$ 是 a 的置信水平为 $1-\alpha$ 的区间估计.

(二) 方差 σ^2 的区间估计

由第六章的定理 6.3.1 知，若总体 ξ 服从正态 $N(a,\sigma)$，则有：

专业课辅导+真题资料+微信公众号：432应统书院

$$\frac{nS^2}{\sigma^2} \sim \chi^2_{(n-1)}.$$

给定置信水平 $1-\alpha, 0<\alpha<1$，在 χ^2-分布临界值表中查自由度 $n-1$ 对应的两个数值 $\chi^2_{(n-1)}\left(\frac{\alpha}{2}\right)$ 及 $\chi^2_{(n-1)}\left(1-\frac{\alpha}{2}\right)$ 使得

$$P\left\{\frac{nS^2}{\sigma^2} \geqslant \chi^2_{(n-1)}\left(\frac{\alpha}{2}\right)\right\} = \frac{\alpha}{2};$$

$$P\left\{\frac{nS^2}{\sigma^2} \geqslant \chi^2_{(n-1)}\left(1-\frac{\alpha}{2}\right)\right\} = 1-\frac{\alpha}{2}.$$

即有

$$P\left\{\chi^2_{(n-1)}\left(1-\frac{\alpha}{2}\right) \leqslant \frac{nS^2}{\sigma^2} \leqslant \chi^2_{(n-1)}\left(\frac{\alpha}{2}\right)\right\} = 1-\alpha;$$

$$P\left\{\frac{nS^2}{\chi^2_{(n-1)}\left(\frac{\alpha}{2}\right)} \leqslant \sigma^2 \leqslant \frac{nS^2}{\chi^2_{(n-1)}\left(1-\frac{\alpha}{2}\right)}\right\} = 1-\alpha. \quad (7.4.4)$$

就是说，$\left[\frac{nS^2}{\chi^2_{(n-1)}\left(\frac{\alpha}{2}\right)}, \frac{nS^2}{\chi^2_{(n-1)}\left(1-\frac{\alpha}{2}\right)}\right]$ 为方差 σ^2 的置信水平 $1-\alpha$ 的区间估计.

上述的区间估计是在总体 ξ 服从正态 $N(a,\sigma)$ 的情况下讨论的. 当总体 ξ 不服从正态分布时，若其数学期望 $E(\xi)$ 及方差 $D(\xi)$ 都有限，则统计量 $\eta_n = \frac{\bar{\xi}-a}{\sigma/\sqrt{n}}$ 的极限分布为正态 $N(0,1)$，即当 n 足够大时，样本平均 $\bar{\xi}$ 渐近服从正态 $N\left(a, \frac{\sigma}{\sqrt{n}}\right)$. 此时，可应用(7.4.2)近似地给出 a 的区间估计. 如果方差 σ^2 未知，要用样本方差 S^2 估计 σ^2，则可应用(7.4.3)式近似地给出 a 的区间估计.

下面的例子，则不必用(7.4.2)及(7.4.3)两式作近似计算，它可直接导出样本平均 $\bar{\xi}$ 的精确分布，从而求出参数的区间估计.

例 7.4.2 设总体 ξ 服从指数分布，密度函数为：

专业课辅导+真题资料+微信公众号：432应统书院

$$f(x;\lambda)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\lambda>0;\\ 0, & 其他,\end{cases}$$

求参数 λ 的置信水平 $1-\alpha$ 的上置信限.

解 我们知道,ξ 的特征函数为

$$\varphi(t)=\left(1-j\frac{t}{\lambda}\right)^{-1},$$

于是,样本平均$\bar{\xi}$的特征函数为

$$\varphi_{\bar{\xi}}(t)=\left(1-j\frac{t}{n\lambda}\right)^{-n}.$$

记 $\eta=2n\lambda\bar{\xi}$,则 η 的特征函数为

$$\varphi_{\eta}(t)=(1-2jt)^{-n}.$$

可见,样本函数 η 服从自由度为 $2n$ 的 χ^2 分布. 给定置信水平 $1-\alpha$,查得临界值 $\chi^2_{2n}(\alpha)$,使得

$$P\{2n\lambda\bar{\xi}\leqslant\chi^2_{2n}(\alpha)\}=1-\alpha,$$

亦即

$$P\left\{0<\lambda\leqslant\frac{1}{2n\bar{\xi}}\chi^2_{2n}(\alpha)\right\}=1-\alpha, \tag{7.4.5}$$

因此 $\left(0,\dfrac{1}{2n\bar{\xi}}\chi^2_{2n}(\alpha)\right)$ 是参数 λ 的置信水平为 $1-\alpha$ 的区间估计.

也可以这样选取 λ 的区间估计:

$$P\left\{2n\lambda\bar{\xi}\geqslant\chi^2_{2n}\left(\frac{\alpha}{2}\right)\right\}=\frac{\alpha}{2};$$

$$P\left\{2n\lambda\bar{\xi}\geqslant\chi^2_{2n}\left(1-\frac{\alpha}{2}\right)\right\}=1-\frac{\alpha}{2};$$

$$P\left\{\chi^2_{2n}\left(1-\frac{\alpha}{2}\right)\leqslant2n\lambda\bar{\xi}\leqslant\chi^2_{2n}\left(\frac{\alpha}{2}\right)\right\}=1-\alpha;$$

专业课辅导+真题资料+微信公众号：432应统书院

$$P\left\{\frac{1}{2n\bar{\xi}}\chi^2_{2n}\left(1-\frac{\alpha}{2}\right)\leqslant\lambda\leqslant\frac{1}{2n\bar{\xi}}\chi^2_{2n}\left(\frac{\alpha}{2}\right)\right\}=1-\alpha.$$

习　题

1. 设某地区的风速 ξ 服从 $\Gamma(\alpha,\beta)$，密度函数为

$$f(x,\alpha,\beta)=\begin{cases}\dfrac{\beta^{-(\alpha+1)}}{\Gamma(\alpha+1)}x^{\alpha}e^{-\frac{x}{\beta}}, & x>0;\\ 0, & x\leqslant 0.\end{cases}$$

$(\alpha+1>0,\beta>0)$，试按频率估计概率的原理，在 $\beta=\frac{1}{9}$，$\alpha=2$，$n=5$ 时求百年一遇的最大风速 M 值（即 ξ 大于 M 的概率为 1%）.

2. 对任一地区，地震的震级数 y 与其发生的次数 n 之间有经验公式：$n(y)=e^{\alpha-\beta y}$，$\alpha>0,\beta>0,y\geqslant 0$，试按频率估计概率的原理，求震级 ξ 的分布函数.

3. 设总体 ξ 服从正态 $N(a,1)$，今观察了二十次，只记录是否为负值，若事件"$\xi<0$"出现了十四次，试按频率估计概率的原理，试求 a 的估计值.

4. 设总体 ξ 的密度函数为 $f(x;\theta)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，求参数 θ 的极大似然估计量.

(1) $f(x;\theta)=\frac{1}{2}e^{-|x-\theta|}$，$|x|>\theta,-\infty<\theta<0,-\infty<x<\infty$；

(2) $f(x;\theta)=\begin{cases}\theta x^{\theta-1}, & 0<x<1,\quad 0<\theta<\infty,\\ 0, & \text{其他};\end{cases}$

(3) $f(x;\theta)=\begin{cases}\dfrac{1}{\theta}, & 0<x\leqslant\theta,\quad 0<\theta<\infty,\\ 0, & \text{其他};\end{cases}$

(4) $f(x;\theta)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}}, & 0<x<\infty,0<\theta<\infty,\\ 0, & \text{其他}.\end{cases}$

5. 设总体 ξ 服从二项分布 $b(N,p)$，$0<p<1$，N 为正整数，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，求 N 及 p 的矩法估计量.

专业课辅导+真题资料+微信公众号：432应统书院

6. 设总体 ξ 服从对数正态分布，密度函数为

$$f(x;a,\sigma)=\frac{\lg e}{\sqrt{2\pi}\sigma x}\exp\left\{-\frac{1}{2\sigma^2}(\lg x-a)^2\right\},$$

$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，求 a 及 σ 的矩法估计量.

7. 设 T 为电子元件的失效时间(单位:h)，其密度函数为

$$f(t)=\beta e^{-\beta(t-t_0)},\quad t>t_0>0$$

(即随机变量 T 具有在左边 t_0 截头的，参数为 β 的指数分布). 假定 n 个元件独立地试验并记录其失效时间分别为 $T_1,T_2,\cdots,T_n$.

(1) 当 t_0 为已知时，求 β 的极大似然法估计量.

(2) 当 β 为已知时，求 t_0 的极大似然法估计量.

8. 设总体 ξ 服从正态 $N(a,\sigma)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本.

(1) 求 k，使 $\hat{\sigma}=\frac{1}{k}\sum_{i=1}^{n}|\xi_i-a|$ 为 σ 的无偏估计量.

(2) 求 k，使 $\hat{\sigma}^2=\frac{1}{k}\sum_{i=1}^{n-1}(\xi_{i+1}-\xi_i)^2$ 为 σ^2 的无偏估计量.

9. 设总体 ξ 的数学期望为 a，方差为 σ^2，$\xi_1,\xi_2,\cdots,\xi_n$ 是它的样本，$T(\xi_1,\xi_2,\cdots,\xi_n)$ 为 a 的任一线性无偏估计量. 证明样本平均 $\bar{\xi}$ 与 T 的相关系数为 $\sqrt{D(\bar{\xi})/D(T)}$.

10. 设总体服从正态 $N(a,1)$，$-\infty<a<\infty$，ξ_1,ξ_2,ξ_3 为其样本，试证下述三个估计量

(1) $\hat{a}_1=\frac{1}{5}\xi_1+\frac{3}{10}\xi_2+\frac{1}{2}\xi_3$;

(2) $\hat{a}_2=\frac{1}{3}\xi_1+\frac{1}{4}\xi_2+\frac{5}{12}\xi_3$;

(3) $\hat{a}_3=\frac{1}{3}\xi_1+\frac{1}{6}\xi_2+\frac{1}{2}\xi_3$

都是 a 的无偏估计量，并求出每一估计量的方差，问哪一个最小?

*11. 设总体 ξ 的数学期望为 a，$\hat{a}_1$ 及 $\hat{a}_2$ 分别为参数 a 的两个无偏估计量，它们的方差分别为 σ_1^2 及 σ_2^2，相关系数为 ρ，试确定常数 $c_1>0$，$c_2>0$，$c_1+c_2=1$，使得 $c_1\hat{a}_1+c_2\hat{a}_2$ 有最小方差.

* 12. 设总体 ξ 服从正态 $N(a_1,1)$，总体 η 服从正态 $N(a_2,2)$，$\xi_1,\xi_2,\cdots,$

专业课辅导+真题资料+微信公众号：432应统书院

ξ_{n1} 为总体 ξ 的样本，$\eta_1,\eta_2,\cdots,\eta_{n2}$ 为总体 η 的样本，且这两个样本相互独立.

(1) 试求 $a=a_1-a_2$ 的无偏估计量 $\hat{a}$.

(2) 如果 $n_1+n_2=n$ 固定，问 n_1 及 n_2 如何配置，可使 $\hat{a}$ 的方差达到最小.

* 13. 设总体 ξ 及 η 的数学期望及方差分别为 a_1,a_2 及 σ_1^2,σ_2^2，$\xi_1,\xi_2,\cdots,\xi_n$ 及 $\eta_1,\eta_2,\cdots,\eta_n$ 分别为它们的样本，这两个样本相互独立，$\bar{\xi}=\frac{1}{n}\sum_{i=1}^{n}\xi_i$，$\bar{\eta}=\frac{1}{n}\sum_{i=1}^{n}\eta_i$，试证 $\bar{\xi}-\bar{\eta}$ 为 a_1-a_2 的最优线性无偏估计量.

* 14. 设总体 ξ 的密度函数为

$$f(x;\theta)=\begin{cases}\frac{1}{\theta}, & 0<x\leqslant\theta,0<\theta<\infty;\\ 0, & 其他,\end{cases}$$

ξ_1,ξ_2,ξ_3 为其样本. 试证 $\frac{4}{3}\max\limits_{1\leqslant i\leqslant 3}\xi_i$ 及 $4\min\limits_{1\leqslant i\leqslant 3}\xi_i$ 都是参数 θ 的无偏估计量，问哪个较有效？

*15. 设 $\hat{\theta}_1$ 及 $\hat{\theta}_2$ 是参数 θ 的两个独立的无偏估计量，且 $\hat{\theta}_1$ 的方差为 $\hat{\theta}_2$ 的方差的两倍，试确定常数 c_1 及 c_2，使得 $c_1\hat{\theta}_1+c_2\hat{\theta}_2$ 为参数 θ 的无偏估计量，并且在所有这样的线性估计中方差最小.

*16. 设 T_1 及 T_2 分别是参数 θ 的可估计函数 $g_1(\theta)$ 及 $g_2(\theta)$ 的最优无偏估计量，试证 $b_1T_1+b_2T_2$ 是 $b_1g_1(\theta)+b_2g_2(\theta)$ 的最优无偏估计量，其中 b_1 和 b_2 是常数.

17. 设总体 ξ 服从正态 $N(a_1,\sigma_1)$，总体 η 服从正态 $N(a_2,\sigma_2)$，$\xi_1,\xi_2,\cdots,\xi_{n_1}$ 及 $\eta_1,\eta_2,\cdots,\eta_{n_2}$ 分别为其样本，并且这两个样本相互独立.

(1) 试建立 σ_1^2/σ_2^2 的置信水平为 $1-\alpha$ 的区间估计；

(2) 假定 $\sigma_1=\sigma_2$，试建立 a_1-a_2 的置信水平为 $1-\alpha$ 的区间估计.

18. 设总体 ξ 服从正态 $N(a_1,\sigma)$，总体 η 服从正态 $N(a_2,\sigma)$，其中 σ 未知，$\xi_1,\xi_2,\cdots,\xi_n$ 及 $\eta_1,\eta_2,\cdots,\eta_n$ 分别为其样本，并且这两个样本相互独立，求 a_1-a_2 的置信水平为 0.95 的区间估计.

19. 设总体 ξ 服从正态 $N(a,\sigma)$，已知 $\sum_{i=1}^{15}x_i=8.7$，$\sum_{i=1}^{15}x_i^2=25.05$，试分别求置信水平为 0.95 的 a 及 σ^2 的区间估计.

专业课辅导+真题资料+微信公众号：432应统书院

20. 若从自动车床加工的一批零件中随机抽取 10 个，测得其尺寸与规定尺寸的偏差（单位：μm）分别为：2，1，-2，3，2，4，-2，5，3，4. 零件尺寸的偏差记作 ξ，假设总体 ξ 服从正态 $N(a,\sigma)$，试求 a 及 σ^2 的无偏估计值，并求置信水平为 0.9 的区间估计.

专业课辅导+真题资料+微信公众号：432应统书院

第八章　假设检验

§8.1　基本概念

假设检验同参数估计一样，都是数理统计学的重要内容. 设总体 ξ 的分布函数 $F(x;\theta)$ 中 θ 为未知参数，现对未知参数 θ 提出假设"θ_0 为其真值"，试问怎样利用样本 $\xi_1,\xi_2,\cdots,\xi_n$ 提供的信息来检验这个假设应否定还是不应否定，"最佳"检验的准则又怎样确定？这类统计问题，称之为参数的假设检验问题. 参数检验与参数估计，是利用样本 $\xi_1,\xi_2,\cdots,\xi_n$ 对总体的概率统计特性提供的信息，建立样本的函数——估计量或检验统计量，从不同的角度来处理未知参数 θ 的两种统计方法.

除参数的假设检验问题之外，非参数的假设检验问题也相当重要. 例如，若总体 ξ 的分布函数表达式 $F(x)$ 不知道时，假设总体 ξ 的分布函数为某个指定的分布函数 $F_0(x)$，问怎样利用样本 $\xi_1,\xi_2,\cdots,\xi_n$ 提供的信息来对假设作出判断，是否定或不否定，就是一个非参数性假设检验. 又如随机变量 ξ 与 η 之间的相关性或独立性等问题的假设检验，也归属于非参数性的假设检验.

关于检验法的准则，将引进两类错判、最佳检验、一致最佳检验等概念.

现在，我们从几个例子谈起.

例 8.1.1　某工厂金工车间生产一种铆钉，铆钉直径的标准定为 $a_0=2$ cm. 现为了提高产量，采用了一种新工艺，现从新工艺生产的铆钉中随机抽取 100 个，分别测其直径，计算得平均值 $\bar{x}=1.978$ cm. 我们要问，$\bar{x}$ 与 a_0 之间的差异，纯粹是试验或测试的误差造成的，还是反映了工艺条件的改革使得铆钉直径发生了显著

专业课辅导+真题资料+微信公众号：432应统书院

性的变化.

上述例子是关于数学期望 $E(\xi)=2$ 的假设检验问题. 凡对随机变量 ξ 的数字特征或分布函数中的参数提出假设作检验的问题,都称之为参数性的假设检验问题. 又如下面的问题,是非参数性的假设检验问题,它是对随机变量 ξ 的分布函数表达式提出假设作检验的问题.

例 8.1.2 认为某工厂生产的灯泡其光通量 ξ 服从正态分布,是否正确,如何检验;又认为某电话交换台在某段时间接到的呼唤次数 η 服从泊松分布,是否正确,如何检验?

关于假设检验的思想方法,没有参数估计那样直观. 现在,我们着重讨论例 8.1.1,阐述假设检验这种统计方法的基本思想及主要步骤.

(1) 假设检验的基本思想

在例 8.1.1 中,从新工艺生产的铆钉来看,$\bar{x}$ 比 a_0 小了些,其差为 $a_0-\bar{x}=0.022$ cm. 前面提出这个差异是新工艺造成的还是纯粹由于随机因素引起的问题,可以用"假设检验"这种统计方法对此作出判断. 假设检验的基本思想是这样:用 a_0 表示原工艺生产的铆钉直径总体 ξ 的数学期望 $E(\xi)$,假设"新工艺对铆钉直径没有显著影响",那么从新工艺生产的铆钉中抽取的样本,则可认为是从原工艺总体 ξ 中抽取的. 假定总体 ξ 服从正态 $N(a_0,\sigma_0)$,其中 σ_0 为已知(如:由以往生产情况确定),那么统计量

$$U=\frac{\bar{\xi}-a_0}{\sigma_0/\sqrt{n}} \tag{8.1.1}$$

服从正态 $N(0,1)$. 由于 $P\{|U|\leqslant 1.96\}=0.95$,所以

$$P\left\{a_0-1.96\frac{\sigma_0}{\sqrt{n}}\leqslant\bar{\xi}\leqslant a_0+1.96\frac{\sigma_0}{\sqrt{n}}\right\}=0.95. \tag{8.1.2}$$

因此,若"假设"是正确的,那么从新工艺生产的铆钉总体中随机抽取容量为 n 的样本 $\xi_1,\xi_2,\cdots,\xi_n$,其样本平均 $\bar{\xi}$ 的观察值 $\bar{x}$ 落在这

专业课辅导+真题资料+微信公众号：432应统书院

个区间之外的概率为 0.05. 对此可作如下解释："在抽取样本 20 次中，大约只有 1 次$\bar{x}$值落在这个区间之外". 如果在一次实际抽取的容量为 n 的样本的平均值$\bar{x}$落在这个区间外面，那么我们就说从新工艺总体中抽取的样本不能"认为"来自原工艺生产的铆钉总体 ξ，即认为"假设"是不正确的，则"假设"应该被否定或者说被拒绝.

(2) 判断"假设"的根据

上述判断假设是否定或不否定，是根据所谓小概率事件的原理. 小概率事件原理是："小概率事件(或概率很小的事件)在一次试验(或观察)中是几乎不可能发生的". 我们再详细解释一下，设有某假设 H_0 需要检验，先假设 H_0 是正确的，在此"假设"之下，构造某一个事件 A，它在 H_0 为正确的条件下的概率很小，例如 $P\{A|H_0\}=0.05$，现在进行一次试验，如果事件 A 发生了，就说一个小概率事件居然发生了，这就不能不使人怀疑 H_0 的正确性，因而很可能要否定 H_0.

下面，再举一例子予以说明.

例 8.1.3 某箱子中有白球及黑球，总数为 100 个，但不知白球及黑球各占多少. 现在提出假设 H_0："其中 99 个是白的". 下面用刚才的想法来判断这个假设 H_0 是否正确. 暂设 H_0 正确，那么从箱子中任取一球，"得黑球"的概率为 0.01，我们认为这是一小概率事件. 如果抽一球居然抽得是黑球，那么自然使人们要否定 H_0，就是说白球的个数不是 99.

(3) 第一类错误

按上述原则否定 H_0，不免要犯错误. 我们知道，当 H_0 正确时，小概率事件 A 也有可能发生，并非 A 绝对不发生. 因此，当 H_0 本来是正确的，而在事件 A 发生后，却错误地否定了 H_0，这种"弃真"的错误，称之为犯第一类错误. 由上所述，犯第一类错误的概率为 α，即 $P\{A|H_0\}=\alpha$. 例如在检验新工艺的例 8.1.1 中，由(8.1.2)式知道，犯第一类错误的概率为 $\alpha=0.05$.

专业课辅导+真题资料+微信公众号：432应统书院

现将上述讨论，归纳出假设检验问题的三个主要步骤：

第一步　提出原假设

$$H_0: a=a_0 \quad 或 \quad H_0: F(x)=F_0(x).$$

如例 8.1.1 中，$H_0: a=2$. 例 8.1.2 中，$H_0: F(x)=F_0(x)$，对于光通量 ξ 来说，$F_0(x)$ 是正态分布，对于呼唤次数 η 来说，$F_0(x)$ 是泊松分布. 可见，原假设（又称作零假设）H_0 是我们所要进行检验的对象.

第二步　建立检验统计量

建立检验统计量是假设检验中重要的环节. 比如例 8.1.1 中，在总体 ξ 服从正态 $N(a,\sigma_0)$ 的假定下，当原假设 $H_0: a=a_0$ 成立时，建立检验统计量 $U=\dfrac{\bar{\xi}-a_0}{\sigma_0/\sqrt{n}}$，$U$ 服从标准正态 $N(0,1)$. 注意，检验统计量是样本的函数，要求不带有任何未知参数.

对于总体 ξ 的分布函数为正态 $N(a,\sigma)$ 的基本假定下，常用的检验统计量有 t-分布 χ^2-分布、F-分布，这些适合于小样问题. 如果总体 ξ 不服从正态分布，或总体 ξ 的分布函数未知，这时检验统计量的精确分布难于求出或相当复杂，如有可能求出其渐近分布，则只适用于大样问题. 非参数性的检验问题，一般都是大样问题，如 §8.3 中所讨论的检验问题.

第三步　确定 H_0 的否定域

如例 8.1.1 中，当原假设 H_0 成立时，检验统计量 U 服从正态 $N(0,1)$，那么给定满足 $0<\alpha<1$ 的 α 值，在标准正态分布表中查得临界值 u_α，使得

$$P\{|U|\geqslant u_\alpha\}=\alpha,$$

或者

$$P\{-u_\alpha\leqslant U\leqslant u_\alpha\}=1-\alpha. \tag{8.1.3}$$

若由样本 $\xi_1,\xi_2,\cdots,\xi_n$ 的观察值 $x_1,x_2,\cdots,x_n$ 算得统计量 U 的观察值 u 落在 $(-\infty,-u_\alpha)$ 或 (u_α,∞) 时则否定 H_0，H_0 的否定域为 A

专业课辅导+真题资料+微信公众号：432应统书院

$=\{(x_1,x_2,\cdots,x_n):|U|\geqslant u_\alpha\}$，称 u_α 为临界值.

对原假设 H_0 作出否定或不否定的判断，通常称之为对 H_0 作显著性检验. 称 α 为显著性水平（或检验水平），$1-\alpha$ 为置信水平. 以后常用"在显著性水平 α 下对 H_0 作显著性检验"这类术语. 值得强调的是，我们对 H_0 作出判断，是冒着犯第一类错误的风险的.

对于 H_0，给定不同的显著性水平 α，对应有不同的临界值 u_α，相应地有不同的否定域，因而有不同的判断结论，这点必须注意，如图 8.1.1 所示.

图 8.1.1

$$\alpha_1=0.01,\ u_{\alpha_1}=2.58;$$

$$\alpha_2=0.05,\ u_{\alpha_2}=1.96;$$

$$\alpha_3=0.10,\ u_{\alpha_3}=1.65.$$

如果检验统计量 U 的观察值 u 落在阴影部分，在给定显著性水平 $\alpha_3=0.10$时，作出否定 H_0 的结论；而在给定显著性水平 $\alpha_2=0.05$ 时，作出不否定 H_0 的结论.

还必须指出，不否定 H_0，绝不意味着不必再作进一步的考察就可以断然接受 H_0 这个假设，即并不认为 H_0 一定是正确的. 这个观点必须注意，我们将在 §8.4 中作较深入的讨论. 在此，用图 8.1.2作简略的说明（$H_0:a=a_0$）.

如果检验统计量 U 的观察值 $u\in(-u_\alpha,u_\alpha)$，则不否定 H_0，但

专业课辅导+真题资料+微信公众号：432应统书院

图 8.1.2

还不能认为 H_0 是正确的. 因为当 u 落在阴影部分时,样本平均值 $\bar{\xi}$ 来自服从正态 $N(a_0,\sigma)$ 的总体 ξ 的可能性很大,但来自服从正态 $N(a_1,\sigma)$ 的可能性也很大,也就是说,总体的数学期望 $E(\xi)=a$ 可能是取 a_0 值,但不能排斥取 a_1 值的可能性.

§8.2 参数假设检验

参数假设检验的一般提法是:设总体 ξ 的分布函数 $F(x;\theta)$ 中,θ 为未知参数,$\theta\in\Omega$,Ω 为参数空间. 我们将参数空间 Ω 分解为互不相交的两个部分 Ω_0 及 $\Omega-\Omega_0$,即 $\Omega_0\cap(\Omega-\Omega_0)=\varnothing$,$\Omega_0\cup(\Omega-\Omega_0)=\Omega$. 考虑检验问题:

$$H_0:\theta\in\Omega_0,\qquad H_1:\theta\in\Omega-\Omega_0,$$

Ω_0 为非空子集,Ω_0 是假设检验的对象,称 H_0 为原假设(或零假设),称 H_1 为备选假设.

如果 Ω 只含有两个点,即若 $\Omega=\{\theta_0,\theta_1\}$,$\Omega_0=\{\theta_0\}$,$\Omega-\Omega_0=\{\theta_1\}$,则有

$$H_0:\theta=\theta_0,\qquad H_1:\theta=\theta_1.$$

这时称 H_0 及 H_1 分别为简单原假设及简单备选假设.

如果 Ω 多于两个点,$\Omega_0=\{\theta_0\}$,而 $\Omega-\Omega_0=\{\theta\neq\theta_0\}$ 为非单点集,即有:

$$H_0:\theta=\theta_0,\qquad H_1:\theta\neq\theta_0,$$

专业课辅导+真题资料+微信公众号：432应统书院

则称 H_0 为简单原假设，H_1 为复合备选假设.

若 Ω_0 及 $\Omega-\Omega_0$ 都是非单点集，则称 H_0 及 H_1 都是复合的.

一、正态总体数学期望 a 的检验问题

为了简便叙述，对每一种检验问题，着重给出检验统计量，并举例说明其应用.

（一）一个总体时 a 的检验：

（甲） $H_0: a=a_0$， $H_1: a\neq a_0$.

（1）设方差 $D(\xi)=\sigma^2$ 为已知，记 $\sigma^2=\sigma_0^2$

假定总体 ξ 服从正态 $N(a,\sigma_0)$，那么由于样本均值 $\bar{\xi}$ 服从正态 $N\left(a,\dfrac{\sigma_0}{\sqrt{n}}\right)$，所以

$$U^*=\frac{\bar{\xi}-a}{\sigma_0/\sqrt{n}}$$

服从正态 $N(0,1)$. 因此，当原假设 H_0 成立时，统计量

$$U=\frac{\bar{\xi}-a_0}{\sigma_0/\sqrt{n}} \tag{8.2.1}$$

服从正态 $N(0,1)$，U 不带有未知参数，它可作为判断 H_0 的检验统计量. 这种检验法称之为 u 检验法.

（2）设方差 $D(\xi)=\sigma^2$ 为未知

假定总体 ξ 服从正态 $N(a,\sigma)$，由定理 6.3.1 的推论知

$$T^*=\sqrt{n-1}\frac{\bar{\xi}-a}{S}\sim t_{(n-1)}.$$

因此，当原假设 H_0 成立时，则

$$T=\sqrt{n-1}\frac{\bar{\xi}-a_0}{S}\sim t_{(n-1)}, \tag{8.2.2}$$

即统计量 T 服从自由度为 $n-1$ 的 t-分布，且不带有未知参数，它可作为判断 H_0 的检验统计量，这种检验法，称之为 t 检验法.

例 8.2.1 某工厂生产的电灯泡的使用时数用 ξ 表示，假定 ξ

专业课辅导+真题资料+微信公众号：432应统书院

服从正态 $N(a,\sigma)$，其中 a 及 σ 都是未知参数，现在观察 $n=20$ 个灯泡，测得 20 个灯泡使用时数 $x_1,x_2,\cdots,x_{20}$，并由此算得 $\bar{x}=1\ 832$，$S=497$，试问该厂电灯泡的平均使用时数为"$a=2\ 000$(h)"这个结论是否成立？

解 考虑检验问题如下：

$$H_0: a=2\ 000,\qquad H_1: a\neq 2\ 000.$$

由(8.2.2)式计算得 $t_{19}=\sqrt{19}\dfrac{1\ 832-2\ 000}{497}=-1.473$，给定显著性水平 $\alpha=0.05$，对于自由度 19，在 t 分布临界值表中，查得临界值 $t_{19}(0.05)=2.093$. 由于 $|t_{19}|=1.473<2.093$，因而不否定 H_0.

（乙）$H_0: a\leqslant a_0$　　$H_1: a>a_0$

假定总体 ξ 服从正态 $N(a,\sigma)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为样本，$\bar{\xi}$ 为样本均值，S^2 为样本方差.

（1）设方差 σ^2 为已知，记 $\sigma=\sigma_0^2$

记 $U^*=\dfrac{\bar{\xi}-a}{\sigma_0/\sqrt{n}}$，　$U=\dfrac{\bar{\xi}-a_0}{\sigma_0/\sqrt{n}}$.

由定理 6.3.1 知，随机变量 U^* 服从 $N(0,1)$，给定检验水平 α，由下式确定临界值 u_α：

$$P\{U^*\geqslant u_\alpha\}=\alpha. \tag{8.2.3}$$

当 H_0 为真时，$U\leqslant U^*$，则有

$$P\{U\geqslant u_\alpha\}\leqslant P\{U^*\geqslant u_\alpha\}=\alpha,$$

即 H_0 的检验水平为 α 的否定域为 $A_1=\{(x_1,x_2,\cdots,x_n): U\geqslant u_\alpha\}$，由样本观察值计算样本均值 $\bar{x}$，求得检验统计量 U 的观察值 $u=\sqrt{n}(\bar{x}-a_0)/\sigma_0$，若 $u\geqslant u_\alpha$，则否定 H_0，若 $u<u_\alpha$，则不否定 H_0.

（2）设方差 σ^2 为未知，用 S^2 估计 σ^2

记　$T^*=\dfrac{\bar{\xi}-a}{S/\sqrt{n-1}}$，　$T=\dfrac{\bar{\xi}-a_0}{S/\sqrt{n-1}}$，

由定理 6.3.1 知，随机变量 T^* 服从 $t_{(n-1)}$ 分布，给定检验水平 α，由

专业课辅导+真题资料+微信公众号：432应统书院

下式确定临界值 $t_{(n-1)}(\alpha)$：

$$P\{T^* \geqslant t_{(n-1)}(\alpha)\} = \alpha. \tag{8.2.4}$$

当 H_0 为真时，$T \leqslant T^*$，则有

$$P\{T \geqslant t_{(n-1)}(\alpha)\} \leqslant P\{T^* \geqslant t_{(n-1)}(\alpha)\} = \alpha,$$

即 H_0 的检验水平为 α 的否定域为 $A_2 = \{(x_1, x_2, \cdots, x_n): T \geqslant t_{(n-1)}(\alpha)\}$，由样本观察值计算样本均值 $\bar{x}$、样本方差 S^2，求得检验统计量 T 的观察值 $t = \sqrt{n-1}(\bar{x}-a_0)/S$，若 $t \geqslant t_{(n-1)}(\alpha)$，则否定 H_0，若 $t < t_{(n-1)}(\alpha)$，则不否定 H_0.

例 8.2.2 设总体 ξ 服从正态 $N(a, \sigma)$，a 及 σ 未知，随机抽取容量为 $n=17$ 的样本，由样本观察值计算得到 $\bar{x}=23$，$S^2=14.91$，给定检验水平 $\alpha=0.05$，检验假设：

$$H_0: a \leqslant 21, \quad H_1: a > 21.$$

解 由(8.2.4)式知，$t_{16}(0.05)=1.746$，计算检验统计量 T 的观察值为

$$t = \frac{\bar{x}-21}{\sqrt{14.91}/\sqrt{16}} = 2.07.$$

由于 $t=2.07 > t_{16}(0.05)=1.746$，于是否定 H_0.

（二）两个总体时 a 的比较：

$$H_0: a_1 = a_2, \qquad H_1: a_1 \neq a_2.$$

现在讨论两个总体的数学期望 a_1 及 a_2 是否相等的假设检验，假定总体 ξ 服从正态 $N(a_1, \sigma_1)$，总体 η 服从正态 $N(a_2, \sigma_2)$，其中 a_1, σ_1 及 a_2, σ_2 都是未知的. 现在分别从总体 ξ 及 η 中抽取样本 $\xi_1, \xi_2, \cdots, \xi_{n_1}$ 及 $\eta_1, \eta_2, \cdots, \eta_{n_2}$，记

$$\bar{\xi} = \frac{1}{n_1}\sum_{i=1}^{n_1} \xi_i, \quad S_1^2 = \frac{1}{n_1}\sum_{i=1}^{n_1} (\xi_i - \bar{\xi})^2;$$

$$\bar{\eta} = \frac{1}{n_2}\sum_{i=1}^{n_2} \eta_i, \quad S_2^2 = \frac{1}{n_2}\sum_{i=1}^{n_2} (\eta_i - \bar{\eta})^2.$$

设这两个样本相互独立，讨论 H_0 的显著性检验.

专业课辅导+真题资料+微信公众号：432应统书院

(1) 设 $\sigma_1^2=\sigma_2^2=\sigma^2$，$\sigma^2$ 未知.

由(6.3.10)式知，

$$\sqrt{\frac{n_1n_2(n_1+n_2-2)}{n_1+n_2}}\cdot\frac{(\bar{\xi}-\bar{\eta})-(a_1-a_2)}{\sqrt{n_1S_1^2+n_2S_2^2}}\sim t_{(n_1+n_2-2)}.$$

因此，当 H_0 成立时，统计量

$$\sqrt{\frac{n_1n_2(n_1+n_2-2)}{n_1+n_2}}\cdot\frac{(\bar{\xi}-\bar{\eta})}{\sqrt{n_1S_1^2+n_2S_2^2}}\sim t_{(n_1+n_2-2)}. \tag{8.2.5}$$

它不带有未知参数，因而可作为 H_0 的检验统计量，这是 t 检验法.

当 $n_1=n_2=n$ 时，则有 $\sqrt{n-1}\dfrac{\bar{\xi}-\bar{\eta}}{\sqrt{S_1^2+S_2^2}}\sim t_{2n-2}$.

(2) 配对数据的 t 检验

设总体 (ξ,η) 服从二维正态分布，$(\xi_1,\eta_1),(\xi_2,\eta_2),\cdots,(\xi_n,\eta_n)$ 为样本，$H_0:a_1=a_2$（这是数据成对地出现的情形），ξ 服从 $N(a_1,\sigma_1)$，η 服从 $N(a_2,\sigma_2)$，σ_1 及 σ_2 都为未知值，记

$$Z=\xi-\eta,Z_i=\xi_i-\eta_i\quad(i=1,2,\cdots,n). \tag{8.2.6}$$

将 Z 看成一维的总体，$z_1,z_2,\cdots,z_n$ 为其样本，Z 服从 $N(a_1-a_2,\sigma)$，记 $a_1-a_2=a$. 此时，原来的数学期望是否相等的检验，变成为对总体 Z 的数学期望 a 是否等于零的检验，即有

$$H_0:a=0,\qquad H_1:a\neq0.$$

记

$$\bar{Z}=\frac{1}{n}\sum_{i=1}^{n}Z_i,\qquad S^2=\frac{1}{n}\sum_{i=1}^{n}(Z_i-\bar{Z})^2,$$

方差 $D(Z)=\sigma^2$ 为未知，用 S^2 估计 σ^2，当 H_0 为真时，

$$T=\frac{\bar{Z}}{S/\sqrt{n-1}}\sim t_{(n-1)}. \tag{8.2.7}$$

给定显著性水平 α，由下式确定临界值 $t_{(n-1)}\left(\dfrac{\alpha}{2}\right)$，

$$P\left\{|T|\geqslant t_{(n-1)}\left(\frac{\alpha}{2}\right)\right\}=\alpha.$$

专业课辅导+真题资料+微信公众号：432应统书院

H_0 的否定域为 $A=\left\{(z_1,\cdots,z_n):|T|\geqslant t_{(n-1)}\left(\frac{\alpha}{2}\right)\right\}$. 由样本观察值 $z_1,z_2,\cdots,z_n$ 计算得到 $\overline{Z}$ 及 S^2 值，求得检验统计量 T 的观察值 t，当 $|t|\geqslant t_{(n-1)}\left(\frac{\alpha}{2}\right)$ 时，否定 H_0，当 $|t|<t_{(n-1)}\left(\frac{\alpha}{2}\right)$ 时，不否定 H_0.

例 8.2.3 设有甲、乙两种安眠药，考虑比较它们的治疗效果，以 ξ 表示失眠病者服甲种药后睡眠时间延长的时数；以 η 表示服乙种药后睡眠时间延长的时数. 现在独立观察 20 个病者，其中 10 人服甲种药，另 10 人服乙种药，延长的时数记在表 8.2.1 中.

表 8.2.1

时数 \ 人数 / 药物	1	2	3	4	5	6	7	8	9	10
x	1.9	0.8	1.1	0.1	0.1	4.4	5.5	1.6	4.6	3.4
y	0.7	−1.6	−0.2	−1.2	−0.1	3.4	3.7	0.8	0.0	2.0

假定 ξ 及 η 分别服从正态 $N(a_1,\sigma)$ 及 $N(a_2,\sigma)$，其中方差 σ^2 相同，试问这两种药物的疗效有无显著差异.

解 将上述问题归结为考虑如下检验问题：

$$H_0:a_1=a_2,\qquad H_1:a_1\neq a_2.$$

这可用(8.2.5)式建立的统计量 $t_{(n_1+n_2-2)}$ 对 H_0 作显著性检验.

实际计算结果如下：

$$\overline{x}=2.35,\quad S_1^2=3.52,\quad n_1=10;$$

$$\overline{y}=0.75,\quad S_2^2=2.88,\quad n_2=10.$$

将 $\overline{x}$、S_1^2 及 $\overline{y}$、S_2^2 各值代入(8.2.5)式计算得 $t_{18}=1.90$. 给定显著性水平 $\alpha=0.05$，对于自由度 18，在 t-分布临界值表中查得临界值 $t_{18}(0.05)=2.101$. 由于 $1.90<2.101$，因而不否定 H_0. 在实际工作中，就认为这两种药物的疗效没有显著差异.

例 8.2.4 某工厂生产某种电器材料，要检验原标准材料的

专业课辅导+真题资料+微信公众号：432应统书院

平均疲劳寿命同新试验材料的平均疲劳寿命之间有无显著性的差异. 试验所得的数据列在表 8.2.2 中.

表 8.2.2

标准材料的疲劳寿命/h	40	110	150	65	90	210	270
新试验材料的疲劳寿命/h	60 450	150 110	220 175	310	380	350	250

用 ξ 及 η 分别表示标准材料及新试验材料的疲劳寿命，一般情况下，可假定 ξ 及 η 分别服从对数正态分布，即有 $X=\lg\xi$，$Y=\lg\eta$，其中 X 及 Y 分别服从正态 $N(a_1,\sigma)$ 及正态 $N(a_2,\sigma)$（在此又假定了 X 及 Y 的方差相同）.

解 将上述问题归结为考虑下述检验问题：

$$H_0: E(X)=E(Y), \qquad H_1: E(X)\neq E(Y).$$

可用(8.2.5)式建立的统计量 $t_{(n_1+n_2-2)}$ 对 H_0 作显著性检验. 现将 X 及 Y 的试验值（试验结果的数据）列在表 8.2.3 中.

表 8.2.3

$x=\lg\xi$	1.602	2.041	2.176	1.813	1.954	2.322	2.431
$y=\lg\eta$	1.778 2.653	2.176 2.041	2.342 2.243	2.491	2.580	2.544	2.398

实际计算结果如下：

$$\bar{x}=2.0484,\quad S_1^2=\frac{0.501}{7}\approx 0.072,\ n_1=7;$$

$$\bar{y}=2.3246,\quad S_2^2=\frac{0.663}{10}=0.0663,\ n_2=10;$$

$$t_{(7+10-2)}\approx -2.01.$$

专业课辅导+真题资料+微信公众号：432应统书院

给定显著性水平 $\alpha=0.05$，对于自由度 15，在 t-分布临界值表中查得临界值 $t_{15}(0.05)=2.131$. 由于 $|t_{15}|=2.01<2.131$，因而不否定 H_0. 在实际工作中，就说还不能认为这两种材料的疲劳寿命有显著差异.

例 8.2.5 现要比较甲、乙两种橡胶制成的轮胎的耐磨性. 今从甲、乙两种轮胎中各随机抽取八个，各取一个组成一对，再随机选取八架飞机，将八对轮胎随机地搭配给八架飞机，作耐磨性试验，经过了一定时间的起落后，测得轮胎磨损量（单位：mg）数据如表 8.2.4.

表 8.2.4

甲	4 900	5 220	5 500	6 020	6 340	7 660	8 650	4 870
乙	4 930	4 900	5 140	5 700	6 110	6 880	7 930	5 010

试问这两种轮胎的耐磨性有无显著差异？

解 用 ξ 及 η 分别表示甲、乙两种轮胎的磨损量，假定 ξ 及 η 分别服从正态 $N(a_1,\sigma)$ 及正态 $N(a_2,\sigma)$，其中方差假定为相同，提出假设检验问题：

$$H_0: a_1=a_2, \qquad H_1: a_1\neq a_2.$$

记 $Z=\xi-\eta$，用（8.2.7）式建立的统计量 $t_{(n-1)}$ 作 H_0 的检验统计量，将表 8.2.4 中的数据对应地相减，将 Z 的数值列在表 8.2.5中.

表 8.2.5

Z	-30	320	360	320	230	780	720	-140

假设检验问题化为

$$H_0: d=0, \qquad H_1: d\neq 0.$$

实际计算结果如下：

专业课辅导+真题资料+微信公众号：432应统书院

$$\overline{Z}=320,\ S^2=89\ 425,$$

$$t_7=\sqrt{7}\frac{\overline{Z}}{S}\approx 2.83.$$

给定显著性水平 $\alpha=0.05$，对于自由度 $8-1=7$，在 t 分布临界值表中查得临界值 $t_7(0.05)=2.365$. 由于 $2.83>2.365$，因而否定 H_0，即认为这两种轮胎的耐磨性有显著差异. 事实上从表 8.2.4 看到乙轮胎明显地较耐磨.

两个总体的数学期望是否相等的显著性检验，其实际意义还在于它是一种选优的统计方法. 若 H_0 被否定，数学期望之间的差异是显著的，这时就有比较、选优的意义. 至于多个总体的数学期望之间差异性的显著性检验，是建立 F 统计量作检验统计量，用所谓方差分析法进行检验，将在第九章的* §9.5 中作介绍.

二、正态总体方差 σ^2 的检验问题

（一）一个总体时 σ^2 的检验：

（甲）　　$H_0:\sigma^2=\sigma_0^2,\qquad H_1:\sigma^2\neq\sigma_0^2.$

假定总体 ξ 服从正态 $N(a,\sigma)$，其中 a 及 σ 都是未知的，由(6.3.5)式知

$$\frac{nS^2}{\sigma^2}\sim\chi^2_{(n-1)}.$$

因此，当 H_0 成立时，统计量

$$\frac{nS^2}{\sigma_0^2}\sim\chi^2_{(n-1)},\tag{8.2.8}$$

即 $\frac{nS^2}{\sigma_0^2}$ 服从自由度为 $n-1$ 的 χ^2-分布，且不带有未知参数，因此可作为 H_0 的检验统计量，用它来进行显著性检验的方法称为 χ^2 检验法.

给定显著性水平 α，在 χ^2-分布临界值表中查自由度 $n-1$ 对应的两个临界值 $\chi^2_{(n-1)}\left(\frac{\alpha}{2}\right)$ 及 $\chi^2_{(n-1)}\left(1-\frac{\alpha}{2}\right)$，使得

专业课辅导+真题资料+微信公众号：432应统书院

$$P\left\{\frac{nS^2}{\sigma_0^2}\geqslant\chi^2_{(n-1)}\left(\frac{\alpha}{2}\right)\right\}=\frac{\alpha}{2},$$

$$P\left\{\frac{nS^2}{\sigma_0^2}\geqslant\chi^2_{(n-1)}\left(1-\frac{\alpha}{2}\right)\right\}=1-\frac{\alpha}{2},$$

即

$$P\left\{\chi^2_{(n-1)}\left(1-\frac{\alpha}{2}\right)\leqslant\frac{nS^2}{\sigma_0^2}\leqslant\chi^2_{(n-1)}\left(\frac{\alpha}{2}\right)\right\}=1-\alpha. \tag{8.2.9}$$

当样本的观察值由(8.2.8)式算得$\frac{nS^2}{\sigma_0^2}$之值大于$\chi^2_{(n-1)}\left(\frac{\alpha}{2}\right)$或小于$\chi^2_{(n-1)}\left(1-\frac{\alpha}{2}\right)$时，则否定$H_0$.

例 8.2.6 某纺织厂生产的维尼纶纤度用ξ表示，在稳定生产情况下，可假定ξ服从正态$N(a,\sigma)$，其标准差σ按往常资料暂定为0.048. 现在随机抽取5根纤维，测得其纤度为：1.32，1.55，1.36，1.40，1.44. 试问总体ξ的方差σ^2有没有显著变化.

解 将上述问题归结为下述检验问题：

$$H_0:\sigma=0.048,\qquad H_1:\sigma\neq0.048.$$

利用(8.2.8)式建立的统计量$\frac{nS^2}{\sigma_0^2}\sim\chi^2_{(n-1)}$对$H_0$作显著性检验.

实际计算结果如下：

$$\bar{x}=1.414,$$

$$\chi_4^2=\frac{1}{0.048^2}(0.094^2+0.136^2+0.054^2+0.014^2+0.026^2)$$
$$=13.5.$$

给定显著性水平$\alpha=0.10$，在χ^2-分布临界值表中查得数值$\chi_4^2(0.05)=9.488$及$\chi_4^2(0.95)=0.711$，使得

$$P\{0.711\leqslant\chi_4^2\leqslant9.488\}=0.90.$$

由于13.6不在区间[0.711，9.488]内，因而否定H_0，即认为方差显著地改变了.

专业课辅导+真题资料+微信公众号：432应统书院

（乙）$H_0:\sigma^2\leqslant\sigma_0^2$，　　$H_1:\sigma^2>\sigma_0^2$

假定总体 ξ 服从正态 $N(a,\sigma)$，参数 a 及 σ 都是未知的，ξ_1，$\xi_2,\cdots,\xi_n$ 为样本，$\bar{\xi}$为样本均值，S^2 为样本方差，记

$$S^2=\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2,K^*=\frac{nS^2}{\sigma^2},\quad K=\frac{nS^2}{\sigma_0^2}.\qquad(8.2.10)$$

由定理 6.3.1 知，随机变量 K^* 服从 $\chi^2_{(n-1)}$ 分布. 给定检验水平 α，由下式确定临界值 $\chi^2_{(n-1)}(\alpha)$

$$P\{K^*\geqslant\chi^2_{(n-1)}(\alpha)\}=\alpha.\qquad(8.2.11)$$

当 H_0 为真时，$K\leqslant K^*$，则有

$$P\{K\geqslant\chi^2_{(n-1)}(\alpha)\}\leqslant P\{K^*\geqslant\chi^2_{(n-1)}(\alpha)\}=\alpha,$$

即 H_0 的检验水平为 α 的否定域为 $A=\{(x_1,x_2,\cdots,x_n):K\geqslant\chi^2_{(n-1)}(\alpha)\}$. 由样本观察值计算样本平均$\bar{x}$，样本方差 S^2，求得检验统计量 K 的观察值 $k=nS^2/\sigma_0^2$，当 $k\geqslant\chi^2_{(n-1)}(\alpha)$时，否定 H_0，若 $k<\chi^2_{(n-1)}(\alpha)$，则不否定 H_0.

（二）两个总体的方差比较问题

（甲）　　$H_0:\sigma_1^2=\sigma_2^2$，　　$H_1:\sigma_1^2\neq\sigma_2^2$.

假定总体 ξ 及 η 分别服从正态 $N(a_1,\sigma_1)$ 及正态 $N(a_2,\sigma_2)$，现在分别抽取样本 $\xi_1,\xi_2,\cdots,\xi_{n_1}$ 及 $\eta_1,\eta_2,\cdots,\eta_{n_2}$，记

$$\bar{\xi}=\frac{1}{n_1}\sum_{i=1}^{n_1}\xi_i,\qquad S_1^2=\frac{1}{n_1}\sum_{i=1}^{n_1}(\xi_i-\bar{\xi})^2;$$

$$\bar{\eta}=\frac{1}{n_2}\sum_{j=1}^{n_2}\eta_j,\qquad S_2^2=\frac{1}{n_2}\sum_{j=1}^{n_2}(\eta_j-\bar{\eta})^2,$$

设这两个样本相互独立，则由(6.3.5)式知

$$\frac{n_1S_1^2}{\sigma_1^2}\sim\chi^2_{(n_1-1)};$$

$$\frac{n_2S_2^2}{\sigma_2^2}\sim\chi^2_{(n_2-1)}.$$

由定理 6.3.2 知

专业课辅导+真题资料+微信公众号：432应统书院

$$\left[\frac{\frac{n_1S_1^2}{\sigma_1^2}}{(n_1-1)}\right]\Bigg/\left[\frac{\frac{n_2S_2^2}{\sigma_2^2}}{(n_2-1)}\right]\sim F_{(n_1-1,n_2-1)},$$

因此，当 H_0 成立时，则有

$$\frac{(n_2-1)n_1S_1^2}{(n_1-1)n_2S_2^2}\sim F_{(n_1-1,n_2-1)}, \tag{8.2.12}$$

即$\frac{(n_2-1)n_1S_1^2}{(n_1-1)n_2S_2^2}$服从第一自由度为 n_1-1、第二自由度为 n_2-1 的F-分布，且不带有未知参数，因而它可作为判断 H_0 的检验统计量. 用它作为检验统计量进行显著性检验的方法称为F 检验法.

给定显著性水平 α，怎样选择临界值呢？我们知道

$$E\left(\frac{n_1S_1^2}{n_1-1}\right)=\sigma_1^2,\quad E\left(\frac{n_2S_2^2}{n_2-1}\right)=\sigma_2^2.$$

若 H_0 正确，则 $F_{(n_1-1,n_2-1)}$ 的比值大约为 1，因此这个 F 值太大或太小，都表明 H_0 可能不正确. 一般书本上附的 F 分布表的临界值都是下述这种表达式所确定的：

$$P\{F\geqslant F(\alpha)\}=\alpha.$$

于是选取临界值使得

$$P\left\{F\geqslant F_{(n_1-1,n_2-1)}\left(\frac{\alpha}{2}\right)\right\}=\frac{\alpha}{2},$$

$$P\left\{F\geqslant F_{(n_1-1,n_2-1)}\left(1-\frac{\alpha}{2}\right)\right\}=1-\frac{\alpha}{2}.$$

当(8.2.12)式的 F 统计量的观察值大于 $F_{(n_1-1,n_2-1)}\left(\frac{\alpha}{2}\right)$或者小于 $F_{(n_1-1,n_2-1)}\left(1-\frac{\alpha}{2}\right)$时，否定 H_0.

例 8.2.7 设用两种不同方法冶炼某种金属材料，分别抽样测定其杂质的含量(单位：%)，数据如表 8.2.6 及表 8.2.7. 试问这两种冶炼法的杂质含量的差异性是否显著不同.

专业课辅导+真题资料+微信公众号：432应统书院

表 8.2.6

原冶炼方法:杂质含量						
26.9	22.3	27.2	25.1	22.8	24.2	30.2
25.7	26.1	24.5	23.0	26.4	29.5	

表 8.2.7

新冶炼方法:杂质含量								
22.6	24.3	23.4	22.5	21.9	20.6	20.6	23.5	23.2

解 原方法与新方法的杂质含量分别记作 ξ 及 η,假定 ξ 服从正态 $N(a_1,\sigma_1)$,η 服从正态 $N(a_2,\sigma_2)$,于是提出如下假设检验问题:

$$H_0:\sigma_1=\sigma_2,\qquad H_1:\sigma_1\neq\sigma_2.$$

实际计算结果如下:

$$n_1=13,\bar{x}=25.68,S_1^2=\frac{1}{13}\sum_{i=1}^{13}(x_i-\bar{x})^2=5.411;$$

$$n_2=9,\bar{y}=22.51,S_2^2=\frac{1}{9}\sum_{i=1}^{9}(y_i-\bar{y})^2=1.459.$$

由(8.2.12)式知

$$F_{(12,8)}=\frac{n_1S_1^2/(n_1-1)}{n_2S_2^2/(n_2-1)}=\frac{5.862}{1.641}\approx3.57.$$

给定显著性水平 $\alpha=0.10$,在 F 分布表中查第一自由度 12、第二自由度 8 的临界值为 $F_{(12,8)}(0.05)=3.28$,由于 $3.57>3.28$,因而否定 H_0,即认为这两种冶炼法的杂质含量的差异性是显著不同的.

(乙) $H_0:\sigma_1^2\leqslant\sigma_2^2,\qquad H_1:\sigma_1^2>\sigma_2^2$

假定总体 ξ 服从正态 $N(a_1,\sigma_1)$,$\xi_1,\xi_2,\cdots,\xi_{n_1}$ 为样本;总体 η 服从正态 $N(a_2,\sigma_2)$,$\eta_1,\eta_2,\cdots,\eta_{n_2}$ 为样本.设这两个样本相互独立,$\bar{\xi}$ 及 $\bar{\eta}$ 分别为样本均值,S_1^2 及 S_2^2 分别为样本方差,记

专业课辅导+真题资料+微信公众号：432应统书院

$$F^*=\frac{n_1S_1^2/[(n_1-1)\sigma_1^2]}{n_2S_2^2/[(n_2-1)\sigma_2^2]},\qquad F=\frac{n_1S_1^2/(n_1-1)}{n_2S_2^2/(n_2-1)},$$

由定理 6.3.2 知，随机变量 F^* 服从 $F_{(n_1-1,n_2-1)}$ 分布，给定检验水平 α，由下式确定临界值 $F_{(n_1-1,n_2-1)}(\alpha)$：

$$P\{F^*\geqslant F_{(n_1-1,n_2-1)}(\alpha)\}=\alpha.$$

当 H_0 为真时，$F\leqslant F^*$，则有

$$P\{F\geqslant F_{(n_1-1,n_2-1)}(\infty)\}\leqslant P\{F^*\geqslant F_{(n_1-1,n_2-1)}(\alpha)\}=\alpha, \tag{8.2.13}$$

即 H_0 的检验水平为 α 的否定域为 $A=\{F\geqslant F_{(n_1-1,n_2-1)}(\alpha)\}$，由样本观察值 $x_1,x_2,\cdots,x_{n_1}$ 及 $y_1,y_2,\cdots,y_{n_2}$，计算样本均值 $\bar{x}$ 及 $\bar{y}$，样本方差 S_1^2 及 S_2^2，求得检验统计量 F 的观察值 F，当 $F\geqslant F_{(n_1-1,n_2-1)}(\alpha)$ 时，则否定 H_0，若 $F<F_{(n_1-1,n_2-1)}(\alpha)$，则不否定 H_0.

*三、非正态总体的参数假设检验

例 8.2.8 某工厂生产一批产品，质量检查规定：其次品率 $p_0\leqslant 0.05$，则这批产品可出厂，否则不能出厂. 现从这批产品中抽查 50 件产品，发现有 4 件产品是次品，试问这批产品能否出厂？我们将此问题归结为如下的参数检验问题：

$$H_0:p_0\leqslant 0.05,\qquad H_1:p_0>0.05. \tag{8.2.14}$$

这个问题将放在 *§8.4 中讨论.

例 8.2.9 已知某工厂生产的元件的寿命服从指数分布，现在抽取一些元件进行测试，依据所得的寿命数据去判定产品是否符合“元件的平均寿命不小于 5 000 小时”的质量标准？

用 ξ 代表元件的寿命，即总体 ξ 服从指数分布，有密度函数

$$f(x;\lambda)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\lambda>0;\\ 0, & \text{其他},\end{cases}$$

$E\xi=1/\lambda$. 设有从该总体中抽出的样本 $\xi_1,\cdots,\xi_n$（即取到的元件的寿命）. 现在我们已经可以把考虑的问题化成以下的检验问题：

$$H_0:1/\lambda\geqslant 5\,000,\qquad H_1:1/\lambda<5\,000.$$

专业课辅导+真题资料+微信公众号：432应统书院

解 参照例 7.4.2，记 $\eta=2n\lambda\bar{\xi}$，则 η 服从自由度为 $2n$ 的 χ^2 分布. 给定检验水平 α，$0<\alpha<1$，由下式确定临界值 $\chi^2_{2n}(1-\alpha)$：

$$P\{2n\lambda\bar{\xi}\leqslant\chi^2_{2n}(1-\alpha)\}=\alpha.$$

若 H_0 为真，则

$$P\left\{\bar{\xi}\leqslant\frac{\chi^2_{2n}(1-\alpha)}{2n(1/5\ 000)}\right\}\leqslant P\left\{\bar{\xi}\leqslant\frac{\chi^2_{2n}(1-\alpha)}{2n\lambda}\right\}$$

$$=P\{2n\lambda\bar{\xi}\leqslant\chi^2_{2n}(1-\alpha)\}=\alpha.$$

H_0 的检验水平 α 的否定域为

$$A=\left\{(x_1,x_2,\cdots,x_n):\bar{\xi}\leqslant\frac{2\ 500}{n}\chi^2_{2n}(1-\alpha)\right\}.$$

如 $\alpha=0.05$，$n=15$，则临界值 $\chi^2_{30}(0.95)=18.493$，H_0 的否定域为 $A_0=\{(x_1,x_2,\cdots,x_{15}):\bar{\xi}\leqslant 3\ 082.2\}$. 由样本观察值计算样本均值 $\bar{x}$，若 $\bar{x}\leqslant 3\ 082.2$，则否定 H_0，若 $\bar{x}>3\ 082.2$，则不否定 H_0.

*四、讨论两个问题

（一）单边检验的原假设的选择

在实际问题中，对单边检验的原假设的不同的选择，有时会有不同的结果.

例 8.2.10 某工厂生产的一种绳索，其断裂强度 ξ 服从正态分布 $N(a,\sigma_0)$，σ_0 已知. 某客户要从这工厂进一批货，要求断裂强度的质量指标是 $a\geqslant a_0$. 从这批货中抽样检查，得到样本 $\xi_1,\xi_2,\cdots,\xi_n$，怎样从这一样本判断这批货的质量指标是否符合要求.

解法 1 检验假设 $H_0^{(1)}:a\leqslant a_0$，$H_1^{(1)}:a>a_0$.

由(8.2.3)式知，$H_0^{(1)}$ 的检验水平 α 的否定域为 $A_1=\{(x_1,x_2,\cdots,x_n):U\geqslant u_\alpha\}$，其中检验统计量 U 为

$$U=\frac{\bar{\xi}-a_0}{\sigma_0/\sqrt{n}},\quad \bar{\xi}=\frac{1}{n}\sum_{i=1}^{n}\xi_i,$$

等价地有 $A_1=\{(x_1,x_2,\cdots,x_n):\bar{\xi}\geqslant a_0+u_\alpha\sigma_0/\sqrt{n}\}$，由样本观察值计

专业课辅导+真题资料+微信公众号：432应统书院

算样本均值$\bar{x}$，若$\bar{x} \geqslant a_0 + u_\alpha \sigma_0/\sqrt{n}$，则否定$H_0^{(1)}$；若$\bar{x} < a_0 + u_\alpha \sigma_0/\sqrt{n}$，则不否定$H_0$.

解法 2 检验假设 $H_0^{(2)}: a \geqslant a_0$，$H_1^{(2)}: a < a_0$.

$$U^* = \frac{\bar{\xi} - a}{\sigma_0/\sqrt{n}}, \quad U = \frac{\bar{\xi} - a_0}{\sigma_0/\sqrt{n}},$$

随机变量U^*服从$N(0,1)$，给定检验水平α，则有

$$P\{U^* \leqslant -u_\alpha\} = \alpha.$$

当$H_0^{(2)}$为真时，$U^* \leqslant U$，则有

$$P\{U \leqslant -u_\alpha\} \leqslant P\{U^* \leqslant -u_\alpha\} = \alpha. \tag{8.2.15}$$

$H_0^{(2)}$的检验水平α的否定域为

$$\begin{aligned} A_2 &= \{(x_1, x_2, \cdots, x_n): U \leqslant -u_\alpha\} \\ &= \{(x_1, x_2, \cdots, x_n): \bar{\xi} \leqslant a_0 - \sigma_0 u_\alpha/\sqrt{n}\}. \end{aligned}$$

由样本观察值计算样本均值$\bar{x}$，若$\bar{x} \leqslant a_0 - \sigma_0 u_\alpha/\sqrt{n}$，则否定$H_0^{(2)}$；若$\bar{x} > a_0 - \sigma_0 u_\alpha/\sqrt{n}$，则不否定$H_0^{(2)}$.

从A_1及A_2看到，若$a_0 - \sigma_0 u_\alpha/\sqrt{n} < \bar{x} < a_0 + \sigma_0 u_\alpha/\sqrt{n}$，按解法 1，不否定$H_0^{(1)}$，即认为质量指标不符合要求；按照解法 2，不否定$H_0^{(2)}$，即认为质量指标符合要求. 可见检验的结果与原假设的选定有关.

原假设的选定应根据具体的问题而定. 比如，如果对该工厂的产品质量不大了解，或了解到该厂的产品质量不很好，或这质量不符合指标会造成较大的损失，则可以采取较保守的态度，采用解法 1，把原假设定为$H_0^{(1)}$. 而如果已经知道该厂产品的质量比较可靠，而质量指标离合格标准相差不太远又不会造成大的损失，则可以采取较大胆的态度，采用解法 2 把原假设定为$H_0^{(2)}$. 此外一个可供选择的办法是当同时出现接受$H_0^{(1)}$和$H_0^{(2)}$的时候，不马上作结论，而是继续抽样增加样本容量后再作进一步的检验.

（二）检验统计量的p值

专业课辅导+真题资料+微信公众号：432应统书院

上述各参数假设检验的基本程序为：对于提出的零假设 H_0，建立检验统计量 $T(\xi_1,\xi_2,\cdots,\xi_n)$，给定检验水平 $\alpha(0<\alpha<1)$，确定临界值 λ_α. 由样本观察值 $x_1,x_2,\cdots,x_n$，计算检验统计量的观察值 $T(x_1,x_2,\cdots,x_n)=t$，通过 t 值与 λ 值作比较，对 H_0 作出否定或不否定的判断.

在 SAS 或 SPSS 等统计分析计算软件中，假设检验的基本程序为：如例 8.2.1，有了检验统计量的观察值 $T(x_1,x_2,\cdots,x_n)=t$ 后，用 T 的密度函数计算得 p 值为 $p=P\{T\geqslant t\}$，通过 p 值与检验水平 α 值作比较，若 $p\leqslant\alpha$，则否定 H_0；若 $p>\alpha$，则不否定 H_0.

事实上这两个基本程序是等价的.

例如 F 检验：$f(x)$ 为 F 分布的概率密度函数，记

$$P\{F\geqslant\lambda_\alpha\}=\int_{\lambda_\alpha}^{\infty}f(x)\,\mathrm{d}x=\alpha,$$

图 8.2.1

给定显著性水平 $\alpha=0.05$，若自由度分别为 3 及 15，则 $\lambda_\alpha=3.29$. 当计算得到 F_0 值为 3.5 时，即 $F_0=3.5>\lambda_\alpha=3.29$，则否定 H_0，即认为是显著的（如图 8.2.1）. 而用计算软件 SAS 或 SPSS，则计算得到概率 p 值，即有：

$$P\{F\geqslant F_0\}=P\{F\geqslant 3.5\}=p.$$

可见 $p<0.05$，则否定 H_0，即认为是显著的，因为

$$p=P\{F\geqslant 3.5\}<P\{F\geqslant 3.29\}=0.05.$$

专业课辅导+真题资料+微信公众号：432应统书院

§8.3 非参数的检验

现在，我们来讨论非参数的假设检验，这有两方面的问题，一是总体 ξ 的分布函数 $F(x)$ 的拟合检验，二是随机变量之间的独立性的检验.

一、分布函数的拟合检验

考虑如下假设检验问题：

$$H_0: F(x)=F_0(x), \qquad H_1: F(x)\neq F_0(x).$$

对 H_0 作显著性检验，通常称之为分布函数的拟合检验. 在此，$F_0(x)$ 为需要检验的某个已知的分布函数，$F_0(x;\theta)$ 中也可以含有未知参数 θ. 如例 8.1.2，对于呼唤次数 η，假定 η 服从泊松分布 $P(x;\lambda)$ 的话，那么 $F_0(x)$ 就是泊松分布函数，λ 为参数. 关于总体的分布函数，怎样提得较为准确的假设，即 $F_0(x)$ 的函数表达式怎样提得较为准确，这可以由样本 $\xi_1,\xi_2,\cdots,\xi_n$ 作经验分布函数 $F_n^*(x)$，从中得到启发；也可以从第二章介绍的几种常用的概率分布的物理模型中得到启发.

正态概率纸的应用，只限于判断检验对象 $F_0(x)$ 为正态分布这个假设是否正确. 对于较为一般的检验对象 $F_0(x)$，如 $F_0(x)$ 为泊松分布、均匀分布等情形，需要建立较为通用的检验统计量. 为此，下面介绍皮尔逊定理，这个定理所建立的统计量的极限分布为 χ^2-分布，不论总体 ξ 的分布函数 $F(x)$ 是什么类型，定理的结论都适用于作检验判断，这种检验法称之为皮尔逊 χ^2 检验.

例 8.3.1 某工厂生产一种 220 伏 25 瓦的白炽灯泡，其光通量（单位：lm）用 ξ 表示，ξ 为一随机变量，假设 ξ 是服从正态 $N(a,\sigma)$，试问这个假设是否正确.

考察如下假设检验问题：

$$H_0: F(x)=F_0(x), \qquad H_1: F(x)\neq F_0(x).$$

其中 $F_0(x)$ 为正态 $N(a,\sigma)$ 的分布函数.

现以例 8.3.1 为例叙述皮尔逊 χ^2 检验法的五个步骤：

专业课辅导+真题资料+微信公众号：432应统书院

（一）抽取样本

现在从总体 ξ 中抽取容量为 $n=120$ 的样本（对于有限总体，即个体是有限的情形，一定要用有返回抽取方式，随机地独立抽取样本），进行观察得光通量 ξ 的 120 个观察值，亦即随机地抽取 120 个灯泡测得其光通量的数据，列表 8.3.1.

表 8.3.1　白炽灯泡测试数据　　单位：lm

216	203	197	208	206	209	206	208	202	203
206	213	218	207	208	202	194	203	213	211
193	213	208	208	204	206	204	206	208	209
213	203	206	207	196	201	208	207	213	208
210	208	211	211	214	220	211	203	216	224
211	209	218	214	219	211	208	221	211	218
218	190	219	211	208	199	214	207	207	214
206	217	214	201	212	213	211	212	216	206
210	216	204	221	208	209	214	214	199	204
211	201	216	211	209	208	209	202	211	207
202	205	206	216	206	213	206	207	200	198
200	202	203	208	216	206	222	213	209	219

（二）作频数分布表

对于表 8.3.1 中的 120 个数据，可取 189.5 为下界，228.5 为上界，将（189.5，228.5）按等间距离为 3 划分为 13 个小区间，发现前三个小区间及后三个小区间的 v_i 值（即频数值）都太小，应适当并小区间，使得每个小区间的 v_i 值都不小于 5. 经适当并小区间后，成为 9 个小区间，第一个小区间为（189.5，198.5），第 9 个小区

专业课辅导+真题资料+微信公众号：432应统书院

间为[219.5,228.5). 第 2 至 8 这七个小区间的间距为 3. 这里假定测量精确度为个位的 1,小区间端点都带 0.5 是为了计算频数方便. 详见表 8.3.2.

表 8.3.2　频数分布表

编　号	小　区　间	频　数	累积频数	组中值 x_i
1	$(-\infty,198.5)$	6	6	194
2	[198.5,201.5)	7	13	200
3	[201.5,204.5)	14	27	203
4	[204.5,207.5)	20	47	206
5	[207.5,210.5)	23	70	209
6	[210.5,213.5)	22	92	212
7	[213.5,216.5)	14	106	215
8	[216.5,219.5)	8	114	218
9	$[219.5,\infty)$	6	120	224

以样本观察值作横坐标,以相应的频数为纵坐标,作出直方图,如图 8.3.1.

图 8.3.1　直方图

（三）计算理论频数

专业课辅导+真题资料+微信公众号：432应统书院

如果 H_0 成立，由给定的分布函数 $F_0(x)$，计算得到：

$$p_i=F_0(y_i)-F_0(y_{i-1})\quad(i=1,2,\cdots,m).$$

我们知道，对于正态分布，a 及 σ^2 的极大似然法估计量为：

$$\hat{a}=\bar{\xi},\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2.$$

现在用表 8.3.1 中的数据 x_i 求出 $\bar{\xi}$ 及 S_n^2 的观察值 $\bar{x}$ 及 s^2，作为 a 及 σ^2 的估计值，计算得到：

$$\hat{a}=\bar{x}\approx 209,$$

$$\hat{\sigma}^2=s^2\approx 42.77,\quad \hat{\sigma}=s\approx 6.5.$$

于是 $F_0(x)$ 为正态 $N(209,6.5)$ 的分布函数，可算得

$$\begin{aligned}
p_1&=F_0(198.5)-F_0(-\infty)\\
&=P\{-\infty<\xi<198.5\}\\
&=P\left\{-\infty<\frac{\xi-209}{6.5}<-1.62\right\}\\
&=\Phi(-1.62)-\Phi(-\infty)=1-\Phi(1.62)\\
&=1-0.947\,38=0.052\,62;\\
p_2&=F_0(210.5)-F_0(198.5)\\
&=P\{198.5\leqslant\xi<201.5\}\\
&=P\left\{-1.62\leqslant\frac{\xi-209}{6.5}<-1.15\right\}\\
&=\Phi(-1.15)-\Phi(-1.62)\\
&=\Phi(1.62)-\Phi(1.15)\\
&=0.947\,38-0.874\,9=0.072\,48,
\end{aligned}$$

其中 $\Phi(x)$ 为标准正态 $N(0,1)$ 的分布函数，类似于 p_1 及 p_2 的算法，可逐一计算 $p_3,p_4,\cdots,p_9$ 从而得到 $np_i(i=1,2,\cdots,9)$ 各值如表 8.3.3.

专业课辅导+真题资料+微信公众号：432应统书院

表 8.3.3 频 数 表

编 号	$[y_{i-1},y_i)$	v_i	np_i
1	$(-\infty,198.5)$	6	6.3
2	$[198.5,201.5)$	7	8.7
3	$[201.5,204.5)$	14	14.5
4	$[204.5,207.5)$	20	19.7
5	$[207.5,210.5)$	23	21.8
6	$[210.5,213.5)$	22	19.7
7	$[213.5,216.5)$	14	14.5
8	$[216.5,219.5)$	8	8.7
9	$[219.5,\infty)$	6	6.3

其中 $0<p_i<1$，$\sum_{i=1}^{m} p_i=1$，$i=1,2,\cdots,m$，称 np_i 为样本 $\xi_1,\xi_2,\cdots,\xi_n$ 落入第 i 个小区间的理论频数. 实测频数 v_i 依赖于样本 $\xi_1,\xi_2,\cdots,\xi_n$ 的观察值. v_i 与 np_i 的差异越小，H_0 假设的接受程度越高.

社会经济统计学中常用加权平均法计算样本平均 $\bar{x}$ 与样本方差 S^2. 利用表 8.3.2 中的实测频数 v_i 与组中值 x_i，计算得到：

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{9} v_i x_i=209.125,$$

$$S^2=\frac{1}{n}\sum_{i=1}^{9} v_i(x_i-\bar{x})^2=44.175,\quad S=6.646,$$

其中 $n=\sum_{i=1}^{9} v_i=120$.

（四）建立检验统计量

$$\eta=\sum_{i=1}^{m}\frac{(v_i-np_i)^2}{np_i}=\sum_{i=1}^{m}\frac{v_i^2}{np_i}-n, \tag{8.3.1}$$

专业课辅导+真题资料+微信公众号：432应统书院

其中 η 依赖于 v_i 及 p_i，因而它与 $F_0(x)$ 建立了一定的关系，它可作为判断 H_0 的检验统计量. 利用表 8.3.3 的数据，计算得到表 8.3.4.

表 8.3.4　η 值计算表

编　号	v_i	v_i^2	v_i^2/np_i
1	6	36	5.701
2	7	49	5.632
3	14	196	13.517
4	20	400	20.305
5	23	529	24.266
6	22	484	24.568
7	14	196	13.517
8	8	64	7.356
9	6	36	5.701

我们来分析(8.3.1)式所建立的统计量 η 的实际意义. 当 H_0 成立时，由第五章的博雷尔强大数定律知道，以概率为 1 地使得 $\frac{v_i}{n}\to p_i$，即当 n 足够大时 $v_i\approx np_i$，则 η 以概率为 1 地取值于零的附近，因而由(8.3.1)式所建立的统计量 η 可以用来判断 $F(x)$ 与 $F_0(x)$ 之间的差异性是否显著. 正因为这样，这个统计量 η(它不带有未知参数)可作为判断 H_0 的检验统计量. 试问，统计量 η 服从什么分布？由于统计量 η 依赖于 m 及 n，如果固定 m，当 $n\to\infty$ 时，η 的极限分布是什么？皮尔逊于 1900 年证明了如下定理：

皮尔逊定理　不论 $F_0(x)$ 是什么分布，当 H_0 正确时，则由(8.3.1)式建立的统计量 η 以自由度 $m-1$ 的 χ^2-分布为极限分布，其中 $F_0(x)$ 不带有未知参数.

证　我们就 $m=2$ 的情形先作证明.

$$\eta=\frac{(v_1-np_1)^2}{np_1}+\frac{(v_2-np_2)^2}{np_2}$$

专业课辅导+真题资料+微信公众号：432应统书院

$$=\frac{(v_1-np_1)^2}{np_1}+\frac{[(n-v_1)-n(1-p_1)]^2}{n(1-p_1)}$$

$$=\frac{(v_1-np_1)^2}{np_1}+\frac{(v_1-np_1)^2}{n(1-p_1)}$$

$$=\frac{(v_1-np_1)^2}{np_1(1-p_1)}. \tag{8.3.2}$$

引进记号

$$v_1=\sum_{k=1}^{n}\zeta_k,$$

其中 $\zeta_1,\zeta_2,\cdots,\zeta_n$ 是 n 个相互独立的如下定义的随机变量：

$$\zeta_k=\begin{cases}1, & \xi_k\in[y_0,y_1];\\ 0, & \xi_k\bar{\in}[y_0,y_1],\end{cases}\quad k=1,2,\cdots,n,$$

则

$$P\{\zeta_k=1\}=F_0(y_1)-F_0(y_0)=p_1;$$

$$E(v_1)=np_1,\qquad D(v_1)=np_1(1-p_1).$$

由中心极限定理知，统计量

$$\frac{v_1-np_1}{\sqrt{np_1(1-p_1)}} \tag{8.3.3}$$

的极限分布为 $N(0,1)$，因而统计量 η 的极限分布为 χ_1^2.

如果 H_0 中的 $F_0(x;\theta_1,\theta_2,\cdots,\theta_l)$ 含有 l 个未知参数，则首先用这 l 个未知参数的极大似然法估计量 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_l$ 来代替 $\theta_1,\theta_2,\cdots,\theta_l$，使 $F_0(x;\theta_1,\theta_2,\cdots,\theta_l)$ 不含未知参数，然后应用(8.3.1)式建立的统计量 $\eta(\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_l)$，再用上面介绍的方法进行显著性检验. 但注意此时以 $\chi^2_{(m-1-l)}$-分布为其极限分布，要求 $m>1+l$.

（五）H_0 的显著性检验

给定了显著性水平 α，对于(8.3.1)式建立的检验统计量 η，怎样选择临界值确定 H_0 的否定域？由于 η 的精确分布不知道，

专业课辅导+真题资料+微信公众号：432应统书院

我们是用它的极限分布 $\chi^2_{(m-1)}$ 近似地选择临界值. 同时还应注意到,(8.3.1)式的 χ^2 检验同(8.2.10)式及(8.2.11)式选择临界值及否定域是有所不同的. 粗略地看到,若(8.3.1)式计算统计量 η 的观察值很大,H_0 很可能被否定. 因此,应该这样选择临界值 $\chi^2_{(m-1)}(\alpha)$,使得

$$P\{\eta>\chi^2_{(m-1)}(\alpha)\}=\alpha. \tag{8.3.4}$$

即否定域由区间$(\chi^2_{(m-1)}(\alpha),\infty)$确定. 就是说,由(8.3.1)式计算得统计量 η 的观察值,若大于临界值 $\chi^2_{(m-1)}(\alpha)$,则在显著性水平 α 下否定 H_0. 由(8.3.1)式计算得到:

$$\eta=\sum_{i=1}^{9}(v_i^2/np_i)-120\approx 0.563.$$

给定显著性水平 $\alpha=0.05$,由于自由度等于 $9-1-2=6$,查得临界值为 $\chi^2_6(0.05)=12.592$,由于 $0.563<12.592$,所以不否定 H_0,在实际工作中可认为光通量 ξ 服从正态 $N(209,6.5)$.

例 8.3.2 某箱子中盛有 10 种球,现在从中有返回地随机抽取 200 个,其中第 i 种球共取得 v_i 个($i=1,2,\cdots,10$),数据记录在表 8.3.5 中.

表 8.3.5

种　别	v_i	v_i^2	np_i	v_i^2/np_i
1	35	1 225	20	61.25
2	16	256	20	12.8
3	15	225	20	11.25
4	17	289	20	14.45
5	17	289	20	14.45
6	19	361	20	18.05
7	11	121	20	6.05

专业课辅导+真题资料+微信公众号：432应统书院

续表

种　别	v_i	v_i^2	np_i	v_i^2/np_i
8	16	256	20	12.8
9	30	900	20	45
10	24	576	20	28.8
$\sum_{i=1}^{10}$	200		200	224.9

H_0:箱子中各种球的个数相同,H_1:至少有两种球的个数不同. 检验这个假设是否正确.

解　此时 $m=10, y_0=-\infty, y_i=i(i=1,2,\cdots,10), y_{11}=\infty$. 若 H_0 正确,采用有返回抽取方式抽得每种球的概率都相同,皆为$\frac{1}{10}$,亦即用有返回抽取方式抽取一球为第 i 种球的概率 $p_i=\frac{1}{10}(i=1,2,\cdots,10)$. 又 $n=200$,则 $np_i=20(i=1,2,\cdots,10)$,并记录在表 8.3.5 中.

给定显著性水平 $\alpha=0.05$,由自由度为 $10-1=9$ 的 χ^2-分布表查得临界值 $\chi_9^2(0.05)=16.9$. 由于 η 的观察值为 $224.9-200=24.9>16.9$,所以否定 H_0. 事实上,从表 8.3.5 中 v_i 的观察数据看到,第 1 及第 9 这两种的个数明显地较多.

例 8.3.3　卢瑟福(Rutherford. 1920 年)与盖革(Geiger)作了一个著名的实验①,他们观察了长为 7.5 s 的时间间隔里到达某个计数器的由某块放射物质放出的 α 质点数,共观察了 $n=2\ 608$ 次. 表 8.3.6 中的第 1 列给出的是质点数 i,第 2 列表示有 i 个质点到达计数器的时间间隔数 v_i(每个时间间隔都是长为 7.5 s). 试

① Rutherford Chadwick and Ellis, Radiations from Radivative Substances Cambridge 1920. p. 172.

专业课辅导+真题资料+微信公众号：432应统书院

问这种分布规律是否服从泊松分布？即考察如下检验问题：

$$H_0: F(x)=F_0(x), \quad H_1: F(x)\neq F_0(x),$$

其中 $F_0(x)$ 是泊松分布

$$P(x;\lambda)=\frac{\lambda^x}{x!}e^{-\lambda}, \quad x=0,1,2,\cdots$$

的分布函数，$\lambda>0$ 为其参数.

解 从第二章的 §2.1 中讲到的满足泊松分布的四个条件来看，可以认为原假设 H_0 是正确的. 现利用 χ^2 检验法对 H_0 作出判断，主要是看一看理论频数 np_i 与观察频数 v_i 之间的差异性是大还是小. 参数 λ 的极大似然法估计值为：

$$\hat{\lambda}=\frac{1}{n}\sum_i iv_i=\frac{10\ 086}{2\ 608}=3.87.$$

表 8.3.6 中第 3 列的数值 np_i，利用泊松分布表计算 p_i 值而得到，其中

$$p_i=\frac{(3.87)^i}{i!}e^{-3.87} \qquad i=0,1,2,\cdots.$$

如 $p_0=0.020\ 858$，$np_0=54.398$.

表 8.3.6

i	v_i	np_i
0	57	54.398
1	203	210.523
2	383	407.361
3	525	525.496
4	532	508.418
5	408	393.515
6	273	253.817
7	139	140.325

专业课辅导+真题资料+微信公众号：432应统书院

续表

i	v_i	np_i
8	45	67.882
9	27	29.189
≥10	16	17.075
	2 608	2 608.000

读者可以像表 8.3.4 那样计算 v_i^2 及 v_i^2/np_i 各值. 应当注意到,在用 χ^2 检验法时,自由度为 $m-1-1$,而 $m=11$. 在给定显著性水平 α 的时候,查自由度 9 所对应的临界值 $\chi_9^2(\alpha)$. H_0 的否定域由 $(\chi_9^2(\alpha),\infty)$ 确定.

例 8.3.4 在表 8.3.7 中给出了在 $n=267$ 次观察中电话接错次数 i 的观察频数 v_i 及理论频数 np_i 各值①. 提出同例 8.3.3 那样的检验问题.

表 8.3.7

i	v_i	np_i
0~2	1	2.05
3	5	4.76
4	11	10.39
5	14	18.16
6	22	26.45
7	43	33.03
8	31	36.09

① F. Thorndike, Applications of Poisson's Probability Summation, *The Bell System Technical Journal*, Vol. 5(1926).

专业课辅导+真题资料+微信公众号：432应统书院

续表

i	v_i	np_i
9	40	35.04
10	35	30.63
11	20	24.34
12	18	17.72
13	12	11.92
14	7	7.44
15	6	4.33
≥16	2	4.65
	267	267.00

解 参数 λ 的极大似然法估计值为

$$\hat{\lambda}=\frac{1}{n}\sum_i iv_i=8.74.$$

在用 χ^2 检验法时，自由度为 $15-1-1=13$. 给定显著性水平 $\alpha=0.05$ 时，查得 $\chi^2_{13}(0.05)=22.362$.

二、独立性的检验

我们知道，对于任何两个随机变量 ξ 及 η，ξ 与 η 相互独立是指：若 (ξ,η) 的联合分布函数为 $F(x,y)$，ξ 及 η 的边缘分布函数为 $F_1(x)$ 及 $F_2(y)$，则对一切 x 及 y 下式成立：

$$F(x,y)=F_1(x)\cdot F_2(y).$$

因此，两个随机变量之间独立性的检验问题应写为：

$$H_0: F(x,y)=F_1(x)F_2(y),\ H_1: F(x,y)\neq F_1(x)F_2(y).$$

在实际工作中，如果总体的特征是由两个指标 ξ 及 η 联合来反映，这就要考虑二维随机变量 (ξ,η) 了. 从这个总体中，随机抽取容量为 n 的二维样本 $(\xi_1,\eta_1),(\xi_2,\eta_2),\cdots,(\xi_n,\eta_n)$，将 ξ 及 η 的可能取值的范围分别分成 r 个及 s 个互不相交的小区间，用 n_{ik}

专业课辅导+真题资料+微信公众号：432应统书院

表示这个样本中“ξ 属于第 i 个小区间、η 属于第 k 个小区间”的个数($i=1,2,\cdots,r;k=1,2,\cdots,s$). 记：

$$n_{i\cdot}=\sum_{k=1}^{s}n_{ik},\quad n_{\cdot k}=\sum_{i=1}^{r}n_{ik},\quad n=\sum_{i=1}^{r}\sum_{k=1}^{s}n_{ik}.$$

下面的表 8.3.8 称为列联表.

表 8.3.8

i \ k (n_{ik})	1	2	$\cdots$	s	$n_{i\cdot}$
1	n_{11}	n_{12}	$\cdots$	n_{1s}	$n_{1\cdot}$
2	n_{21}	n_{22}	$\cdots$	n_{2s}	$n_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
r	n_{r1}	n_{r2}	$\cdots$	n_{rs}	$n_{r\cdot}$
$n_{\cdot k}$	$n_{\cdot 1}$	$n_{\cdot 2}$	$\cdots$	$n_{\cdot s}$	n

从总体(ξ,η)中任意抽取一个元素(即一个个体),它的“第一坐标 x 属于第 i 个区间,第二坐标 y 属于第 k 个小区间”这一事件的概率记为 $p(i,k)$,以 $p(i,\cdot)$ 及 $p(\cdot,k)$ 分别记相应的边缘概率. 则有：

$$p(i,\cdot)=\sum_{k=1}^{s}p(i,k),p(\cdot,k)=\sum_{i=1}^{r}p(i,k),$$

$$\sum_{i=1}^{r}\sum_{k=1}^{s}p(i,k)=\sum_{i=1}^{r}p(i,\cdot)=\sum_{k=1}^{s}p(\cdot,k)=1.$$

此时,独立性的检验就简化为检验下述原假设：

$$H_0:p(i,k)=p(i,\cdot)p(\cdot,k),$$

$i=1,2,\cdots,r;\ k=1,2,\cdots,s.$

现在,要建立 H_0 的检验统计量. 首先,我们按极大似然法确定未知参数 $p(i,\cdot)$ 及 $p(\cdot,k)$ 的估计量. 由于

$$p(r,\cdot)=1-\sum_{i=1}^{r-1}p(i,\cdot),$$

专业课辅导+真题资料+微信公众号：432应统书院

$$p(\cdot,s)=1-\sum_{k=1}^{s-1}p(\cdot,k),$$

所以实际上要求 $r+s-2$ 个未知参数的估计量. 当 H_0 成立时,未知参数的似然函数为:

$$L=\prod_{i=1}^{r}\prod_{k=1}^{s}[p(i,k)]^{n_{ik}}=\prod_{i=1}^{r}\prod_{k=1}^{s}[p(i,\cdot)]^{n_{ik}}[p(\cdot,k)]^{n_{ik}}$$

$$=\left[\prod_{i=1}^{r}p(i,\cdot)^{n_{i\cdot}}\right]\left[\prod_{k=1}^{s}p(\cdot,k)^{n_{\cdot k}}\right]$$

$$=\left[1-\sum_{i=1}^{r-1}p(i,\cdot)\right]^{n_{i\cdot}}\left[1-\sum_{k=1}^{s-1}p(\cdot,k)\right]^{n_{\cdot k}}$$

$$\times\left[\prod_{i=1}^{r-1}p(i,\cdot)\right]^{n_{i\cdot}}\left[\prod_{k=1}^{s-1}p(\cdot,k)\right]^{n_{\cdot k}}.$$

由似然方程

$$\begin{cases}\dfrac{\partial\ln L}{\partial p(i,\cdot)}=0, & i=1,2,\cdots,r-1;\\[2ex] \dfrac{\partial\ln L}{\partial p(\cdot,k)}=0, & k=1,2,\cdots,s-1\end{cases}$$

解得

$$\hat{p}(i,\cdot)=\frac{n_{i\cdot}}{n},\qquad \hat{p}(\cdot,k)=\frac{n_{\cdot k}}{n},$$

$i=1,2,\cdots,r-1;k=1,2,\cdots,s-1$,则 $\hat{p}(i,\cdot)$ 及 $\hat{p}(\cdot,k)$ 分别为 $p(i,\cdot)$ 及 $p(\cdot,k)$ 的极大似然法估计量.

利用(8.3.1)式,导出 H_0 的检验统计量

$$\eta=\sum_{i=1}^{r}\sum_{k=1}^{s}\left(n_{ik}-\frac{n_{i\cdot}n_{\cdot k}}{n}\right)^2\Big/\frac{n_{i\cdot}n_{\cdot k}}{n}$$

$$=n\sum_{i=1}^{r}\sum_{k=1}^{s}\left(n_{ik}-\frac{n_{i\cdot}n_{\cdot k}}{n}\right)^2\Big/n_{i\cdot}n_{\cdot k},\qquad(8.3.5)$$

其中 $n\hat{p}(i,k)=n\hat{p}(i,\cdot)\hat{p}(\cdot,k)=n_{i\cdot}n_{\cdot k}/n$,自由度为 $rs-1-(r+s-2)=(r-1)(s-1)$. 可见,当 H_0 为真时,(8.3.5)所建立的统计量 η 的极限分布为 $\chi^2_{(r-1)(s-1)}$. 给定显著性水平 α,在 χ^2-分布临

专业课辅导+真题资料+微信公众号：432应统书院

界值表中查自由度$(r-1)(s-1)$所对应的临界值$\chi^2_{(r-1)(s-1)}(\alpha)$，当由(8.3.5)式计算得$\eta$的数值大于这个临界值时，则在显著性水平$\alpha$下否定$H_0$. 这种检验法又称列联表检验法.

特别地，当$r=s=2$时，统计量为

$$\eta=\frac{n(n_{11}n_{22}-n_{12}n_{21})^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}.$$

η的极限分布是自由度为1的χ^2-分布.

例 8.3.5 现有三个工厂生产同一种产品，产品分为1,2,3三个等级. 为了考察各工厂质量是否一致，从这三个工厂中分别随机地抽出产品109件，100件和91件，每件鉴定其质量等级，结果如表8.3.9：

表 8.3.9

等级＼工厂	1	2	3	合计
1	58	38	32	128
2	28	44	45	117
3	23	18	14	55
合计	109	100	91	300

给定检验水平$\alpha=0.01$，试作独立性的假设检验.

解 “各工厂质量一致”这个假设，可看作是“工厂”和“质量等级”这两个属性独立的假设，用H_0表示. 用(8.3.5)式的检验统计量η对H_0作显著性检验，$r=s=3$，η的极限分布是自由度为4的χ^2-分布，查得临界值为$\chi^2_4(0.01)=12.3$，用(8.3.5)式计算检验统计量η的观察值$v=13.59$，其中$n=300$，

$$n_{1\cdot}=128,\quad n_{2\cdot}=117,\quad n_{3\cdot}=55,$$

$$n_{\cdot 1}=109,\quad n_{\cdot 2}=100,\quad n_{\cdot 3}=91.$$

因为$v=13.59>\chi^2_4(0.01)=12.3$，于是在检验水平$\alpha=0.01$下否定

专业课辅导+真题资料+微信公众号：432应统书院

H_0，即认为各工厂该产品的质量水平不一致.

§8.4 最佳检验

假设检验的一般理论，在数理统计学上占有重要的地位，以上我们讨论了参数及非参数两方面的显著性检验问题. 但是，对假设检验问题还未作深入的研究. 例如对于§8.2中所讨论到的参数检验问题，可以发现还存在一系列的问题有待解决. 例如，当原假设不真时而作出判断不否定 H_0 的可能性有多大；为什么要选择那样的检验统计量；给定了显著性水平 α，确定 H_0 的否定域时有多少种方法；怎样给定检验法的优劣准则，而最佳的检验法在什么条件下存在，又怎样寻找？要解决这些问题，必须引入一些新的概念. 在这一节，我们仅以参数的假设检验问题为例，讨论最优势检验（简称作最佳检验）.

一、两类错判

例 8.4.1 设总体 ξ 服从正态 $N(a,\sigma)$，其中 a 及 σ 都是未知参数，现在提出如下检验问题：

$$H_0: a=0, \qquad H_1: a\neq 0.$$

当原假设成立时，由(8.2.2)式知，建立了统计量

$$T=\sqrt{n-1}\,\frac{\bar{\xi}}{S}\sim t_{(n-1)},$$

它可作为对 H_0 的显著性检验的检验统计量. 给定显著性水平 α，可选临界值 $t_{(n-1)}(\alpha)$，使得

$$P\{|T|>t_{(n-1)}(\alpha)\}=\alpha. \tag{8.4.1}$$

我们也可这样选择临界值 δ，使得

$$P\{|T|<\delta\}=\alpha. \tag{8.4.2}$$

如图 8.4.1 所示，$g(t)$ 为检验统计量 T 的密度函数.

由此可见，在参数的显著性检验中，对于相同的检验统计量及相同的显著性水平 α，可以有不同的临界值，相应地有不同的否定域.

专业课辅导+真题资料+微信公众号：432应统书院

图 8.4.1

例 8.4.2 设总体 ξ 服从正态 $N(a,\sigma_0)$，其中 σ_0 为已知，考虑如下检验问题：

$$H_0: a=a_0,\quad H_1: a\neq a_0.$$

当原假设 H_0 成立时，由(8.2.1)式知，建立了统计量

$$U=\frac{\bar{\xi}-a_0}{\sigma_0/\sqrt{n}},\tag{8.4.3}$$

U 服从正态 $N(0,1)$，它可作为对 H_0 的显著性检验的检验统计量. 给定显著性水平 α，可选临界值 u_α，使得

$$P\{|U|>u_\alpha\}=\alpha,\tag{8.4.4}$$

即

$$P\left\{na_0-\sqrt{n}\sigma_0u_\alpha\leqslant\sum_{i=1}^{n}\xi_i\leqslant na_0+\sqrt{n}\sigma_0u_\alpha\right\}=1-\alpha.\tag{8.4.5}$$

记

$$B_0=(-\infty,na_0-\sqrt{n}\sigma_0u_\alpha)\cup(na_0+\sqrt{n}\sigma_0u_\alpha,\infty),\tag{8.4.6}$$

从(8.4.6)式看到，当样本 $\xi_1,\xi_2,\cdots,\xi_n$ 的观察值 $(x_1,x_2,\cdots,x_n)$ 属于集合

$$\mathscr{X}_0=\left\{(x_1,x_2,\cdots,x_n):\sum_{i=1}^{n}x_i\in B_0\right\}\tag{8.4.7}$$

专业课辅导+真题资料+微信公众号：432应统书院

时，则在显著性水平 α 下否定 H_0. 因此，(8.4.7)式建立的区域 $\mathscr{X}_0$，也是 H_0 的显著性水平为 α 的否定域. (8.4.7)式建立的 $\mathscr{X}_0$，实际上是将样本值域 $\mathscr{X}$ 剖分为互不相交的两部分，即

$$\mathscr{X}=\mathscr{X}_0\cup(\mathscr{X}-\mathscr{X}_0),$$

$$\mathscr{X}_0\cap(\mathscr{X}-\mathscr{X}_0)=\varnothing,$$

其中 $\mathscr{X}_0$ 为 H_0 的显著性水平为 α 的否定域.

我们就 $n=2$ 的简单情形，对否定域 $\mathscr{X}_0$ 作几何解释. 在平面上作两条平行的直线 L_1 及 L_2，

$$L_1:\ x_1+x_2=2a_0+\sqrt{2}\sigma_0 u(\alpha):$$

$$L_2:\ x_1+x_2=2a_0-\sqrt{2}\sigma_0 u(\alpha).$$

将样本值域 $\mathscr{X}$ 剖分为互不相交的两部分 $\mathscr{X}=\mathscr{X}_0\cup(\mathscr{X}-\mathscr{X}_0)$，若样本观察值 $(x_1,x_2)\in\mathscr{X}_0$，则否定 H_0，当样本观察值 $(x_1,x_2)\bar{\in}\mathscr{X}_0$ 时，则不否定 H_0(图 8.4.2 所示).

图 8.4.2

综上所述，在对 H_0 作显著性检验中，给定显著性水平 α 后，可以有不同的临界值，相应地有不同的否定域，这等价于将样本值域剖分为互不相交的两部分的不同剖分法，一个否定域对应于一种剖分法. 因此，在 §8.2 中讨论到的 H_0 的显著性检验法还很不完善.

专业课辅导+真题资料+微信公众号：432应统书院

定义 8.4.1（两类错误） 对于原假设 H_0，设 $\mathscr{X}_0$ 为其否定域. 当 H_0 成立时，若样本观察值 $(x_1, x_2, \cdots, x_n) \in \mathscr{X}_0$，则作出否定 H_0 的判断，这时则称作犯第一类错误（弃真）. 犯第一类错误的概率记为

$$P\{(\xi_1, \xi_2, \cdots, \xi_n) \in \mathscr{X}_0 \mid \Omega_0\} = \alpha. \tag{8.4.8}$$

当 H_0 不成立时，若样本观察值 $(x_1, x_2, \cdots, x_n) \bar{\in} \mathscr{X}_0$，而作出不否定 H_0 的判断，这时则称作犯第二类错误（取伪）. 犯第二类错误的概率记为

$$P\{(\xi_1, \xi_2, \cdots, \xi_n) \bar{\in} \mathscr{X}_0 \mid \Omega_1\} = \beta, \tag{8.4.9}$$

其中 $\Omega_1 = \Omega - \Omega_0$，$\Omega$ 为参数空间. 如 a 为参数，$\Omega_0 = \{a = 0\}$，$\Omega_1 = \{a \neq 0\}$. 对 H_0 的判断情况用表 8.4.1 作说明.

表 8.4.1

决策 \ 假设	H_0 为真	H_0 不真
否定 H_0	第一类错误	正确
不否定 H_0	正确	第二类错误

为了较直观地理解定义 8.4.1，我们仅就 $a_0 < a_1$ 的一种情形，作图说明如下：

图 8.4.3

图 8.4.3 中以斜线部分表示 α 的大小，以网格的部分表示 β 的大小，其中

专业课辅导+真题资料+微信公众号：432应统书院

$$B_0=\left(-\infty,a_0-\frac{\sigma_0}{\sqrt{n}}u_\alpha\right)\cup\left(a_0+\frac{\sigma_0}{\sqrt{n}}u_\alpha,\infty\right),$$

$$\mathscr{X}_0=\{(x_1,x_2,\cdots,x_n):\bar{x}\in B_0\},$$

$$P\{(\xi_1,\xi_2,\cdots,\xi_n)\in\mathscr{X}_0\mid a=a_0\}=\alpha;$$

$$P\{(\xi_1,\xi_2,\cdots,\xi_n)\in\mathscr{X}_0\mid a=a_1\}=\beta.$$

注意，$\beta\neq1-\alpha$，当 n 固定时，α 减小等价于 u_α 增大，从而等价于 β 增大. 当 α 固定时，则增大 n 等价于减小 β.

在实际工作中这两类错误造成的影响常常是不一样的，在不同的场合允许 α 及 β 的大小也常常不一样. 例如在检查病人是否患有某种疾病时，无病当作有病与有病当作无病，其造成的后果就不一样. 又如产品检查验收中，危及健康的药品验收，希望不合格而当作合格验收的可能性尽量小些，即犯第二类错误的概率 β 要求很小. 而对于价格高昂的商品，生产者希望验收者犯第一类错误的可能性尽量小些，即犯第一类错误的概率 α 要求很小.

当然，我们总希望所作出的检验能使犯两类错误的概率同时尽可能地小，最好是全为零，就是说，H_0 为真时，能作出不否定 H_0 的判断；而当 H_0 不真时，能作出否定 H_0 的判断. 这种理想化的检验法，由于样本 $\xi_1,\xi_2,\cdots,\xi_n$ 的随机性，在实际上是不存在的. 当样本的容量 n 给定后，犯这两类错误的概率难于同时被控制. 在假设检验的一般理论中，是在这样原则下进行讨论的，即固定一个 α（或 β）而使得另一个 β（或 α）尽可能地小. 通常是限制犯第一类错误的概率 α，即选定一个数 α，$0<\alpha<1$，要求犯第二类错误的概率 β 尽可能地小. 当然尽量增大 n，可使 α 及 β 都尽可能地小.

二、功效函数

设总体 ξ 的分布函数 $F(x;\theta)$ 中含有未知参数 θ，参数空间记作 Ω，即 $\theta\in\Omega$. 考虑如下假设检验问题：

$$H_0:\theta\in\Omega_0,\quad H_1:\theta\in\Omega-\Omega_0.$$

将样本值域 $\mathscr{X}$ 剖分为互不相交的两部分 $\mathscr{X}_0$ 及 $\mathscr{X}-\mathscr{X}_0$，引进记号 $M(\mathscr{X}_0,\theta)$：对于每一 $\theta\in\Omega$，令：

专业课辅导+真题资料+微信公众号：432应统书院

$$M(\mathscr{X}_0,\theta)=P_\theta\{\mathscr{X}_0\}=P\{(\xi_1,\xi_2,\cdots,\xi_n)\in\mathscr{X}_0\mid\theta\}. \tag{8.4.10}$$

由(8.4.10)式知,

$$M(\mathscr{X}_0,\theta)=\begin{cases}\alpha(\theta), & \theta\in\Omega_0;\\ 1-\beta(\theta), & \theta\in\Omega-\Omega_0.\end{cases} \tag{8.4.11}$$

即当 H_0 为真时,否定 H_0 的概率为 $\alpha(\theta)$,当 H_0 不真时,否定 H_0 的概率为 $1-\beta(\theta)$. $M(\mathscr{X}_0,\theta)$是否定 H_0 的概率.

定义 8.4.2 势函数(或功效函数)　设 $\mathscr{X}_0$ 为原假设 $H_0(\theta\in\Omega_0)$的否定域,称由(8.4.10)定义的函数 $M(\mathscr{X}_0,\theta)$为具有否定域 $\mathscr{X}_0$ 的势函数(或功效函数). 若给定常数 α,$0<\alpha<1$,否定域 $\mathscr{X}_0$ 满足下述不等式:

$$M(\mathscr{X}_0,\theta)=P_\theta\{\mathscr{X}_0\}\leqslant\alpha\quad(\theta\in\Omega_0),$$

则称 α 为否定域 $\mathscr{X}_0$ 的检验水平.

势函数(或功效函数)$M(\mathscr{X}_0,\theta)$依赖于 $\mathscr{X}_0$ 及参数 θ,当选定一个否定域 $\mathscr{X}_0$ 时,它是参数 θ 的一元函数. 势函数的意义在于当 H_0 不真时,反映否定 H_0 的功效大小. 因此,寻找 H_0 的最佳检验法,可归结为选择这样的否定域 $\mathscr{X}_0$,当 H_0 不真时,使得否定 H_0 的功效最大.

如果参数空间 $\Omega=\{\theta_0,\theta_1\}$,$\Omega_0=\{\theta_0\}$,$\Omega_1=\{\theta_1\}$,那么势函数为

$$M(\mathscr{X}_0,\theta)=\begin{cases}\alpha, & \theta=\theta_0,\\ 1-\beta, & \theta=\theta_1.\end{cases} \tag{8.4.12}$$

例 8.4.3　设总体 ξ 服从正态 $N(a,\sigma)$,σ 为已知,记作 σ_0. 考虑简单原假设及简单备选假设

$$H_0:\ a=0,\quad H_1:\ a=a_1(\neq 0).$$

试求 u 检验法的势函数及犯第二类错误的概率.

解　对于显著性水平 $\alpha=0.05$,当原假设 H_0 成立时,由(8.4.3)式知道,u 检验法的否定域为

专业课辅导+真题资料+微信公众号：432应统书院

$$\mathscr{X}_0=\left\{(x_1,x_2,\cdots,x_n):\frac{|\bar{x}|}{\sigma_0/\sqrt{n}}\geqslant 1.96\right\}$$

$$=\left\{(x_1,x_2,\cdots,x_n):|\bar{x}|\geqslant\frac{1.96\sigma_0}{\sqrt{n}}\right\}.$$

对应的势函数为

$$M(\mathscr{X}_0,a)=P_a\{\mathscr{X}_0\}=P_a\left\{\frac{|\bar{\xi}|}{\sigma_0}\sqrt{n}\geqslant 1.96\right\}$$

$$=P_a\left\{\frac{\bar{\xi}}{\sigma_0}\sqrt{n}\leqslant -1.96\right\}+P_a\left\{\frac{\bar{\xi}}{\sigma_0}\sqrt{n}\geqslant 1.96\right\}$$

$$=P_a\left\{\frac{\bar{\xi}-a}{\sigma_0/\sqrt{n}}\leqslant -1.96-\frac{a}{\sigma_0/\sqrt{n}}\right\}$$

$$+P_a\left\{\frac{\bar{\xi}-a}{\sigma_0/\sqrt{n}}\geqslant 1.96-\frac{a}{\sigma_0/\sqrt{n}}\right\}$$

$$=1-\Phi\left(1.96-\frac{a}{\sigma_0/\sqrt{n}}\right)+\Phi\left(-1.96-\frac{a}{\sigma_0/\sqrt{n}}\right),$$

其中 $\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$. 当 H_0 为真时，即 $a=0$ 时，则

$$M(\mathscr{X}_0,0)=1-\Phi(1.96)+\Phi(-1.96)=0.05,$$

当 H_1 为真时，即 $a=a_1(\neq 0)$ 时，则有

$$M(\mathscr{X}_0,a_1)=1-\Phi\left(1.96-\frac{a_1}{\sigma_0/\sqrt{n}}\right)$$

$$+\Phi\left(-1.96-\frac{a_1}{\sigma_0/\sqrt{n}}\right),$$

即得

$$\beta=1-M(\mathscr{X}_0,a_1)$$

$$=\Phi\left(1.96-\frac{a_1}{\sigma_0}\sqrt{n}\right)-\Phi\left(-1.96-\frac{a_1}{\sigma_0}\sqrt{n}\right).$$

对于一般显著性水平 $0<\alpha<1$，u 检验的势函数为

专业课辅导+真题资料+微信公众号：432应统书院

$$M(\mathscr{X}_0,a)=1-\Phi\left(u_\alpha-\frac{a}{\sigma_0}\sqrt{n}\right)+\Phi\left(-u_\alpha-\frac{a}{\sigma_0}\sqrt{n}\right),$$

它与 α、a、n 都有关.

三、最佳检验

定义 8.4.3(最优势检验)　设参数空间为 Ω,考虑

$$H_0:\ \theta_0\in\Omega_0,\quad H_1:\ \theta_1\in\Omega_1=\Omega-\Omega_0,$$

称具有检验水平 α 的否定域 $\mathscr{X}_0$ 所确定的检验为最优势检验(简称作最佳检验),是指对于具有检验水平 α 的任一否定域 $\mathscr{X}'_0$,对每一 $\theta_1\in\Omega_1$,有

$$\begin{aligned}P_{\theta_1}\{\mathscr{X}_0\}&=M(\mathscr{X}_0,\theta_1)\\&\geqslant M(\mathscr{X}'_0,\theta_1)=P_{\theta_1}\{\mathscr{X}'_0\}\end{aligned}$$

其中

$$M(\mathscr{X}_0,\theta_0)\leqslant\alpha,\quad \theta_0\in\Omega_0,$$
$$M(\mathscr{X}'_0,\theta_0)\leqslant\alpha,\quad \theta_0\in\Omega_0.$$

奈曼-皮尔逊定理　设总体 ξ 的分布函数 $F(x;\theta)$ 中含未知参数 $\theta,\theta\in\Omega=\{\theta_0,\theta_1\}$,则对任一数 $\alpha,0<\alpha<1$,对于假设检验问题:

$$H_0:\theta=\theta_0,\quad H_1:\theta=\theta_1.$$

必存在水平为 α 的最佳检验.

如果备选假设是复合的,最佳检验不一定存在. 如果存在,称它为一致最佳检验. 应用似然比法确定的最佳否定域,如果它与备选假设中的 θ_1 无关,那么它就是一致最佳否定域,也即为一致最佳检验.

奈曼-皮尔逊定理对于连续型分布及离散型分布其结论都同样成立. 下面叙述连续型分布的情形. 在叙述过程中,如用概率函数 $P(x,\theta)$ 代替密度函数 $f(x,\theta)$ 的位置,用求和符号代替积分符号,就成为对离散型分布的叙述.

若总体 ξ 有密度函数 $f(x,\theta)$,$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本,θ 为未知参数,$\theta\in\Omega$. 在第七章的 §7.1 中引进了似然函数的概念,即

$$L(\theta)=\prod_{i=1}^{n}f(\xi_i;\theta)$$

为未知参数 θ 的似然函数. 现称

专业课辅导+真题资料+微信公众号：432应统书院

$$\frac{L(\theta_1)}{L(\theta_0)}=\frac{\prod_{i=1}^{n}f(\xi_i,\theta_1)}{\prod_{i=1}^{n}f(\xi_i,\theta_0)} \tag{8.4.13}$$

为似然比. 用似然比将样本值域 $\mathscr{X}$ 剖分为互不相交的两部分，其中

$$\mathscr{X}_0=\{(x_1,x_2,\cdots,x_n):L(\theta_1)\geqslant cL(\theta_0)\}, \tag{8.4.14}$$

$$\mathscr{X}-\mathscr{X}_0=\{(x_1,x_2,\cdots,x_n):L(\theta_1)<cL(\theta_0)\},$$

其中 c 为非负的实数. 定义实数 c 的函数 $\varphi(c)$ 为

$$\varphi(c)=P_{\theta_0}\{\mathscr{X}_0\}=P_{\theta_0}\left\{\frac{L(\theta_1)}{L(\theta_0)}\geqslant c\right\}. \tag{8.4.15}$$

*引理　由(8.4.15)所定义的函数 $\varphi(c)(c\geqslant 0)$，为一左连续，单调不增的函数，而且

$$\varphi(0)=1,\ 0\leqslant\varphi(c)\leqslant\frac{1}{c},$$

$$\lim_{c\to\infty}\varphi(c)=0.$$

由此结论，可见 $1-\varphi(c)$ 为一分布函数. 并且对任一 α，$0<\alpha<1$，存在常数 c_α，使得

$$\varphi(c_\alpha+0)\leqslant\alpha\leqslant\varphi(c_\alpha),$$

其中

$$\varphi(c_\alpha+0)=P_{\theta_0}\left\{\frac{L(\theta_1)}{L(\theta_0)}>c_\alpha\right\}.$$

证　由似然函数的非负性及(8.4.15)式的定义知道，$\varphi(0)=1$，$0\leqslant\varphi(c)\leqslant 1$. 而对于 $c>0$，有

$$\begin{aligned}\varphi(c)&=\int\cdots\int_{\mathscr{X}_0}\prod_{i=1}^{n}f(x_i,\theta_0)\,\mathrm{d}x_1\cdots\mathrm{d}x_n\\&\leqslant\int\cdots\int_{\mathscr{X}_0}\frac{1}{c}\prod_{i=1}^{n}f(x_i,\theta_1)\,\mathrm{d}x_1\cdots\mathrm{d}x_n\\&\leqslant\frac{1}{c},\end{aligned}$$

专业课辅导+真题资料+微信公众号：432应统书院

因而

$$\lim_{c\to\infty}\varphi(c)=0.$$

又由(8.4.15)定义知，若 $0\leqslant c_1<c_2$，则

$$\varphi(c_1)\geqslant\varphi(c_2).$$

还可证明，当 $c_n\uparrow c$ 时，则有 $\lim\limits_{n\to\infty}\varphi(c_n)=\varphi(c)$.

函数 $\varphi(c)$ 的图形如图 8.4.4 所示.

图 8.4.4

***奈曼-皮尔逊定理的证明**

(1) 若存在常数 c_α，使得 $\varphi(c_\alpha)=\alpha$，则检验水平为 α 的最佳否定域为

$$\mathscr{X}_0=\{(x_1,x_2,\cdots,x_n):L(\theta_1)\geqslant c_\alpha L(\theta_0)\}. \quad (8.4.16)$$

证 设 $\mathscr{X}_0'$ 是检验水平为 α 的任一否定域，即

$$P_{\theta_0}\{\mathscr{X}_0\}=P_{\theta_0}\{\mathscr{X}_0'\}=\alpha.$$

记

$$\mathscr{X}_0=[\mathscr{X}_0\cap\mathscr{X}_0']\cup[\mathscr{X}_0\cap(\mathscr{X}-\mathscr{X}_0')],$$
$$\mathscr{X}_0'=[\mathscr{X}_0'\cap\mathscr{X}_0]\cup[\mathscr{X}_0'\cap(\mathscr{X}-\mathscr{X}_0)].$$

易见

$$P_{\theta_1}\{\mathscr{X}_0\}-P_{\theta_1}\{\mathscr{X}_0'\}=\int_{\mathscr{X}_0}\cdots\int\prod_{i=1}^{n}[f(x_i,\theta_1)\mathrm{d}x_i]$$

专业课辅导+真题资料+微信公众号：432应统书院

$$-\int_{\mathscr{X}_0}\cdots\int\prod_{i=1}^{n}[f(x_i,\theta_1)\mathrm{d}x_i]$$

$$=\int_{\mathscr{X}_0\cap(\mathscr{X}-\mathscr{X}_0)}\cdots\int\prod_{i=1}^{n}[f(x_i,\theta_1)\mathrm{d}x_i]$$

$$-\int_{\mathscr{X}_0\cap(\mathscr{X}-\mathscr{X}_0)}\cdots\int\prod_{i=1}^{n}[f(x_i,\theta_1)\mathrm{d}x_i].$$

由(8.4.16)式知，若对于样本的每一次观察值$(x_1,x_2,\cdots,x_n)\in\mathscr{X}_0$，则有

$$\prod_{i=1}^{n}f(x_i,\theta_1)\geqslant c_\alpha\prod_{i=1}^{n}f(x_i,\theta_0).$$

因此

$$\int_{\mathscr{X}_0\cap(\mathscr{X}-\mathscr{X}_0)}\cdots\int\prod_{i=1}^{n}[f(x_i,\theta_1)\mathrm{d}x_i]$$

$$\geqslant c_\alpha\int_{\mathscr{X}_0\cap(\mathscr{X}-\mathscr{X}_0)}\cdots\int\prod_{i=1}^{n}[f(x_i,\theta_0)\mathrm{d}x_i],$$

及

$$\int_{\mathscr{X}_0\cap(\mathscr{X}-\mathscr{X}_0)}\cdots\int\prod_{i=1}^{n}[f(x_i,\theta_1)\mathrm{d}x_i]$$

$$<c_\alpha\int_{\mathscr{X}'_0\cap(\mathscr{X}-\mathscr{X}_0)}\cdots\int\prod_{i=1}^{n}[f(x_i,\theta_0)\mathrm{d}x_i],$$

于是

$$P_{\theta_1}\{\mathscr{X}_0\}-P_{\theta_1}\{\mathscr{X}'_0\}$$

$$\geqslant c_\alpha\left\{\int_{\mathscr{X}_0\cap(\mathscr{X}-\mathscr{X}'_0)}\cdots\int\prod_{i=1}^{n}[f(x_i,\theta_0)\mathrm{d}x_i]\right.$$

$$\left.-\int_{\mathscr{X}'_0\cap(\mathscr{X}-\mathscr{X}_0)}\cdots\int\prod_{i=1}^{n}[f(x_i,\theta_0)\mathrm{d}x_i]\right\}$$

$$=c_\alpha\left\{\int_{\mathscr{X}_0}\cdots\int\prod_{i=1}^{n}[f(x_i,\theta_0)\mathrm{d}x_i]\right.$$

专业课辅导+真题资料+微信公众号：432应统书院

$$-\int\cdots\int_{\mathscr{X}_0'}\prod_{i=1}^{n}[f(x_i,\theta_0)\mathrm{d}x_i]\Big\}$$

$$=c_\alpha(\alpha-\alpha)=0.$$

因此，

$$P_{\theta_1}\{\mathscr{X}_0'\}\leqslant P_{\theta_1}\{\mathscr{X}_0\}.$$

定理得证.

(2) 若存在常数 c_α，满足

$$\varphi(c_\alpha+0)=\alpha\leqslant\varphi(c_\alpha),$$

则检验水平 α 的最佳否定域为

$$\mathscr{X}_0^*=\{(x_1,x_2,\cdots,x_n):L(\theta_1)>c_\alpha L(\theta_0)\}. \quad (8.4.17)$$

证明方法与(1)相同.

(3) 若为 $\varphi(c_\alpha+0)<\alpha<\varphi(c_\alpha)$，则检验水平 α 的最佳否定域为

$$\mathscr{X}_0^{**}=\mathscr{X}_0^*\cup B_0^*, \quad (8.4.18)$$

其中 $\mathscr{X}_0^*$ 由(8.4.17)式给出，$B_0^*\subset B_0^r\subset B_0$，都为 n 维空间中的子集，满足条件

$$P_{\theta_0}\{B_0^*\}=\sup_r[P_{\theta_0}\{B_0^r\}<\alpha-\varphi(c_\alpha+0)]$$

$$\leqslant\alpha-\varphi(c_\alpha+0),$$

其中

$$B_0=\{(x_1,x_2,\cdots,x_n):L(\theta_1)=c_\alpha L(\theta_0)\},$$

$$B_0^r=\{(-\infty,r_1)\times\cdots\times(-\infty,r_n)\}\cap B_0.$$

证 令

$$g(r)=P_{\theta_0}\{B_0^r\}, \quad (8.4.19)$$

则 $g(r)$ 具有如下性质：对任一 $r^*=(r_1^*,r_2^*,\cdots,r_n^*)$，

$$\lim_{\substack{r_i\to r_i^*\\1\leqslant i\leqslant n}}g(r)=g(r^*).$$

又令

$$\alpha^*=\sup_r\{g(r):g(r)<\alpha-\varphi(c_\alpha+0)\},$$

则依上确界定义知，存在 $r^m\nearrow$（是指 n 个分量都递增），使得

专业课辅导+真题资料+微信公众号：432应统书院

$$\alpha^* - \frac{1}{m} < g(r^m) \leqslant \alpha^*.$$

于是

$$\lim_{m\to\infty} g(r^m) = \alpha^*.$$

另一方面，由单调性知必存在某个 r^*，使得 $r^m \nearrow r^*$，从而

$$\alpha^* = \lim_{m\to\infty} g(r^m) = g(r^*) \leqslant \alpha - \varphi(c_\alpha + 0).$$

最后，令

$$B_0^* = B_0^{r^*} = B_0 \cap \{(-\infty, r_1^*) \times \cdots \times (-\infty, r_n^*)\},$$

则 B_0^* 为所求，也就是说，$\mathscr{X}_0^*$ 是检验水平 α 的否定域，即有

$$P_{\theta_0}\{\mathscr{X}_0^{**}\} = P_{\theta_0}\{\mathscr{X}_0^*\} + P_{\theta_0}\{B_0^*\} \leqslant \alpha.$$

例 8.4.4 设总体 ξ 服从正态 $N(a, \sigma_0)$，其中 σ_0 为已知，考虑下列各种情形的假设检验问题：

(A) $H_0: a = a_0, H_1: a = a_1 (> a_0)$；

(B) $H_0: a = a_0, H_1: a = a_1 (< a_0)$；

(C) $H_0: a = a_0, \quad H_1: a > a_0$；

(D) $H_0: a \leqslant a_0, \quad H_1: a > a_0$；

(E) $H_0: a \geqslant a_0, \quad H_1: a < a_0$.

给定显著性水平 α，$0<\alpha<1$，试求最佳否定域.

解 由(8.4.13)式知，似然比为

$$\frac{L[a_1]}{L[a_0]} = \exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}\left[(\xi_i - a_1)^2 - (\xi_i - a_0)^2\right]\right\}$$

$$= \exp\left\{\frac{n(a_1 - a_0)\bar{\xi}}{\sigma_0^2} - \frac{n(a_1^2 - a_0^2)}{2\sigma_0^2}\right\}$$

$$= \exp\left\{\frac{\sqrt{n}(a_1 - a_0)}{\sigma_0}\left[\frac{\sqrt{n}(\bar{\xi} - a_0)}{\sigma_0} - \frac{\sqrt{n}(a_1 - a_0)}{2\sigma_0}\right]\right\}$$

专业课辅导+真题资料+微信公众号：432应统书院

$$=e^{kU-\frac{1}{2}k^2}, \tag{8.4.20}$$

其中

$$U=\sqrt{n}\frac{\bar{\xi}-a_0}{\sigma_0}, \qquad k=\sqrt{n}\frac{a_1-a_0}{\sigma_0}.$$

$$\varphi(c)=P_{a_0}\{(\xi_1,\xi_2,\cdots,\xi_n):e^{kU-\frac{1}{2}k^2}\geqslant c\},$$

k 为常数，U 服从正态分布.

现对于给定的显著性水平 α，$0<\alpha<1$，讨论(A)及(B)的情形，确定最佳否定域.

(A)：由

$$\int_b^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\alpha$$

可确定 b 值，记 $b=u_\alpha$，选取 $c_\alpha=e^{ku_\alpha-\frac{1}{2}k^2}$ 时，则

$$\mathscr{X}_0=\{(x_1,x_2,\cdots,x_n):L(a_1)\geqslant c_\alpha L(a_0)\}$$

为 H_0 的最佳否定域. 事实上，由于 $k>0$，则

$$\mathscr{X}_0=\{e^{kU-\frac{1}{2}k^2}\geqslant e^{ku_\alpha-\frac{1}{2}k^2}\}=\{U\geqslant u_\alpha\}. \tag{8.4.21}$$

当 H_0 为真时，则 U 服从 $N(0,1)$，于是

$$\varphi(c_\alpha)=P_{a_0}\{\mathscr{X}_0\}=P_{a_0}\{U\geqslant u_\alpha\}=\alpha.$$

(B)：由

$$\int_{-\infty}^{b}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\alpha$$

可确定 b 值，记 $b=-u_\alpha$. 选取 $c_\alpha=e^{-ku_\alpha-\frac{1}{2}k^2}$ 时，则

$$\mathscr{X}_0=\{(x_1,x_2,\cdots,x_n):L(a_1)\geqslant c_\alpha L(a_0)\}$$

为 H_0 的最佳否定域. 事实上，由于 $k<0$，则

$$\mathscr{X}_0=\{e^{kU-\frac{1}{2}k^2}\geqslant e^{-ku_\alpha-\frac{1}{2}k^2}\}=\{U\leqslant -u_\alpha\}. \tag{8.4.22}$$

当 H_0 为真时，则 U 服从正态 $N(0,1)$，于是

$$\varphi(c_\alpha)=P_{a_0}\{\mathscr{X}_0\}=P_{a_0}\{U\leqslant -u_\alpha\}=\alpha.$$

下面，利用(A)及(B)的结论，试求(C)、(D)、(E)的一致最

专业课辅导+真题资料+微信公众号：432应统书院

佳否定域.

(C):对于 H_1,任取 a_1,满足 $a_1>a_0$,由(A)知,$H_0:a=a_0$,$H_1':a=a_1(>a_0)$,最佳否定域由(8.4.21)式确定.由于这个最佳否定域与 a_1 具体值无关,只要求 $a_1>a_0$,因而它为(C)的一致最佳否定域.

(D):对于 $H_0':a=a_0$,$H_1:a>a_0$ 而言,由(C)知,检验水平为 α 的一致最佳否定域为

$$\mathscr{X}_0=\{(x_1,x_2,\cdots,x_n):U\geqslant u_\alpha\}.$$

可以证明 $\mathscr{X}_0$ 也是(D)的一致最佳否定域.事实上,对任何 $a\leqslant a_0$,有

$$\begin{aligned}P_a\{\mathscr{X}_0\}&=P\left\{\frac{\bar{\xi}-a_0}{\sigma_0}\sqrt{n}\geqslant u_\alpha\mid a\right\}\\&=P\left\{\frac{\bar{\xi}-a}{\sigma_0}\sqrt{n}\geqslant u_\alpha+\frac{a_0-a}{\sigma_0}\sqrt{n}\mid a\right\}\\&\leqslant P\left\{\frac{\bar{\xi}-a}{\sigma_0}\sqrt{n}\geqslant u_\alpha\mid a\right\}=\alpha,\end{aligned}$$

$$\sup_{a\leqslant a_0}[P_a\{\mathscr{X}_0\}]=P_{a_0}\{\mathscr{X}_0\}=\alpha.$$

设 $\mathscr{X}_0'$ 是水平为 α 的任一否定域,即对每一 $a\leqslant a_0$,

$$P_a\{\mathscr{X}_0'\}\leqslant\alpha,$$

特别地有

$$P_{a_0}\{\mathscr{X}_0'\}\leqslant\alpha,$$

即 $\mathscr{X}_0'$ 为 H_0'、H_1 的水平为 α 的另一否定域.由(C)知,对任何 $a>a_0$,有

$$P_a\{\mathscr{X}_0'\}\leqslant P_a\{\mathscr{X}_0\},$$

从而 $\mathscr{X}_0$ 为(D)的水平为 α 的一致最佳否定域.

(E):按(C)及(D)的证明方法可知,(8.4.22)式的 $\mathscr{X}_0$ 为(E)的水平为 α 的一致最佳否定域.

例 8.4.5 设总体 ξ 服从两点分布,其概率分布为

专业课辅导+真题资料+微信公众号：432应统书院

$$P\{x;p\}=\begin{cases}p, & x=1;\\ 1-p, & x=0,\end{cases}$$

其中 $0<p<1$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本. 考虑假设检验问题：

$$H_0:p=p_0,\quad H_1:p=p_1\quad(p_1>p_0).$$

试求最佳检验即最佳否定域.

解 由(8.4.13)式知，似然比为

$$\frac{L(p_1)}{L(p_0)}=\frac{\prod_{i=1}^{n}P(\xi_i,p_1)}{\prod_{i=1}^{n}P(\xi_i,p_0)}=\frac{\prod_{i=1}^{n}[p_1^{\xi_i}(1-p_1)^{1-\xi_i}]}{\prod_{i=1}^{n}[p_0^{\xi_i}(1-p_0)^{1-\xi_i}]}$$

$$=\left(\frac{p_1}{p_0}\right)^{\sum_{i=1}^{n}\xi_i}\left(\frac{1-p_1}{1-p_0}\right)^{n-\sum_{i=1}^{n}\xi_i}.$$

记

$$\zeta=\sum_{i=1}^{n}\xi_i,$$

由 $p_1>p_0$ 知，$p_1(1-p_0)>p_0(1-p_1)$，因此

$$\frac{L(p_1)}{L(p_0)}=\left(\frac{p_1}{p_0}\right)^{\zeta}\left(\frac{1-p_1}{1-p_0}\right)^{n-\zeta}$$

$$=\left(\frac{1-p_1}{1-p_0}\right)^{n}\left[\frac{p_1(1-p_0)}{p_0(1-p_1)}\right]^{\zeta}\geqslant c,$$

它等价于

$$\zeta\geqslant\frac{n\log\left[\dfrac{c^{\frac{1}{n}}(1-p_0)}{(1-p_1)}\right]}{\log\left[\dfrac{p_1(1-p_0)}{p_0(1-p_1)}\right]}.$$

如果存在常数 c_α，使得

$$\varphi(c_\alpha)=P_{p_0}\{\mathscr{X}_0\}=\alpha,$$

其中

$$\mathscr{X}_0=\{(x_1,x_2,\cdots,x_n):L(p_1)\geqslant c_\alpha L(p_0)\},$$

专业课辅导+真题资料+微信公众号：432应统书院

这等价于存在一个数 b，满足

$$P_{p_0}\{\zeta \geqslant b\} = \alpha,$$

其中 $b = n\log\left[c_\alpha^{\frac{1}{n}}\left(\frac{1-p_0}{1-p_1}\right)\right] \Big/ \log\left[\frac{p_1(1-p_0)}{p_0(1-p_1)}\right]$，而统计量 ζ 服从二项分布，即有

$$P\{\zeta = k\} = C_n^k p^k (1-p)^{n-k} \quad (k = 0, 1, \cdots, n).$$

现举例说明：$n=5$，$\alpha=0.05$，$p_0=0.5$，$p_1=0.6$，则有

$$P_{p_0}\{\zeta = 5\} = \frac{1}{32}; \qquad P_{p_0}\{\zeta = 4\} = \frac{5}{32}.$$

于是

$$P_{p_0}\{\zeta \geqslant 5\} = \frac{1}{32} < 0.05,$$

$$P_{p_0}\{\zeta \geqslant 4\} = \frac{6}{32} > 0.05.$$

即有

$$P_{p_0}\{\zeta \geqslant 5\} = \frac{1}{32} < 0.05 < P_{p_0}\{\zeta \geqslant 4\} = \frac{6}{32}.$$

选取 $c_\alpha = 4$，则水平为 $\alpha = 0.05$ 的最佳否定域为

$$\mathscr{X}_0^{**} = \mathscr{X}_0^* = \{\zeta_n \geqslant 5\},$$

因为形如 $\{\zeta_n = 4\}$ 的集合是由形如

$$\{\xi_1 = 1,\ \xi_2 = 1,\ \xi_3 = 1,\ \xi_4 = 1,\ \xi_5 = 0\}$$

的集合构成（即五维实空间中的 5 个点），则它的任一子集，其概率至少为 $\frac{1}{32}$，而

$$\alpha - \varphi(c_\alpha + 0) = 0.05 - \frac{1}{32} = 0.018\ 75 < \frac{1}{32},$$

于是 B_0^* 为空集.

最后指出，$\mathscr{X}_0 = \{\zeta_n > 4\}$ 还是检验问题

$$H_0: p \leqslant 0.5, \quad H_1: p > 0.5$$

的一致最佳检验，从而提供了对于例 8.2.8 求一致最佳检验的

专业课辅导+真题资料+微信公众号：432应统书院

方法.

按照例 8.4.4 的证明方法，可以完善上述检验问题的讨论. 对于 $H_0: p=p_0, H_1: p>p_0$，先求出 $H_0: p=p_0, H_1': p=p_1(>p_0)$ 的最佳否定域为 $\{\zeta_n>4\}$，它与 p_1 无关，因而它为 H_0、H_1 的一致最佳否定域. 对于 $H_0: p\leqslant p_0, H_1: p>p_0$，先讨论 $H_0': p=p_0, H_1: p>p_0$ 的最佳检验，继而得出 $\{\zeta_n>4\}$ 也为 $H_0: p\leqslant p_0, H_1: p>p_0$ 的一致最佳否定域. 在此 $p_0=0.5$. 对于一般给定的 p_0 值，上述讨论也成立.

*§8.5 样本容量 n 的确定

样本 $\xi_1, \xi_2, \cdots, \xi_n$ 的容量 n 的确定，在许多科学实验工作中是经常遇到的问题. 我们知道，容量 n 太小，估计问题不那么精确，检验问题不那么可靠；容量 n 太大，又会造成人力物力的浪费. 在此，我们所指的样本都是简单样本. 下面，我们就几种场合遇到的样本容量 n 的问题，作扼要的讨论.

一、参数估计与检验中 n 的确定

假定总体 ξ 服从正态 $N(a,\sigma)$，其中 a 及 σ 都是未知参数，给定显著性水平 α，要判断原假设 $H_0: a=a_0$. 当 H_0 成立时，由(8.2.2)知，

$$\sqrt{n-1}\frac{\bar{\xi}-a_0}{S}\sim t_{(n-1)}.$$

选择临界值 $t_{(n-1)}(\alpha)$，使得

$$P\left\{\left|\sqrt{n-1}\frac{\bar{\xi}-a_0}{S}\right|\geqslant t_{(n-1)}(\alpha)\right\}=\alpha, \tag{8.5.1}$$

亦即

$$P\left\{\left|\sqrt{n-1}\frac{\bar{\xi}-a_0}{S}\right|\leqslant t_{(n-1)}(\alpha)\right\}=1-\alpha,$$

其中

专业课辅导+真题资料+微信公众号：432应统书院

$$\bar{\xi}=\frac{1}{n}\sum_{i=1}^{n}\xi_i,\quad S^2=\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2.$$

于是在 σ 为未知时，给出了 a 的置信水平为 $1-\alpha$ 的区间估计

$$\left(\bar{\xi}-\frac{S}{\sqrt{n-1}}t_{(n-1)}(\alpha),\ \bar{\xi}+\frac{S}{\sqrt{n-1}}t_{(n-1)}(\alpha)\right).$$

这个置信区间的长度记作 2Δ，其中

$$\Delta=\frac{S}{\sqrt{n-1}}t_{(n-1)}(\alpha). \tag{8.5.2}$$

在实际工作中，常称 Δ 为估计精度或误差精度、试验精度.

在(8.5.2)式中，如果事先给定 Δ 值，则有

$$n=1+\frac{S^2}{\Delta^2}t^2_{(n-1)}(\alpha). \tag{8.5.3}$$

又当总体的方差 σ^2 由以往经验能确定为 σ_0^2，用 σ_0^2 代替 S^2；给定显著性水平 α，能查得临界值 $t_{(n-1)}(\alpha)$. 这时，万事俱备，n 值才能由(8.5.3)式所确定. 但是，确定临界值 $t_{(n-1)}(\alpha)$ 本身，事先就要知道 n 值，即自由度 $n-1$ 的值. 因此，(8.5.3)式还没有真正解决计算 n 值的问题. 然而，我们从 t 分布临界值表中看到，对于显著性水平 $\alpha\leqslant 0.05$ 的情形，当 $n>30$ 时，其临界值 $t_{(n-1)}(\alpha)\approx 2$，这个临界值对于大于 30 的各个 n 值的影响不太大. 因此，我们先采用近似公式

$$n=1+\frac{4S^2}{\Delta^2} \tag{8.5.4}$$

计算 n. 当计算出的 n 大大超过 30 时，这就与 $t_{(n-2)}(\alpha)\approx 2$ 不矛盾了.

在实际工作中是这样确定 n 的：给定 S 及 Δ 的值，按(8.5.4)式计算 n，若 n 值大于 30，就以这个 n 值作为样本的容量；若 n 值不大于 30，则采用一种叫“试差法”来确定样本的容量 n. 试差法是这样：先由(8.5.4)式计算出的 n，作为第二次查临界值 $t_{(n-1)}(\alpha)$ 时用的 n，将查得的临界值 $t_{(n-1)}(\alpha)$ 代入(8.5.3)式计算

专业课辅导+真题资料+微信公众号：432应统书院

n 值;再由这个 n 值作为第三次查临界值 $t_{(n-1)}(\alpha)$ 时用的 n,又将查得的临界值 $t_{(n-1)}(\alpha)$ 再代入(8.5.3)式计算 n 值,如此循环,直至(8.5.3)式中两边的 n 值相同或差异很小为止.一般试数次就行了,计算出的 n 值通常不要小于5.

例 8.5.1 某种电子仪器额定电流的总体的方差 $\sigma^2 \approx 200$,经检验确定的置信水平 0.95 的置信区间的长度为 20 mA,问要检查多少台仪器,才能满足这样的要求?

解 取 $S^2=200, \Delta=10$,则由(8.5.4)式计算:

$$n=1+\frac{800}{100}=9.$$

由于 n 值小于30,所以要用"试差法".对于 $\alpha=0.05$,查得临界值 $t_8(0.05)=2.31$,代入(8.5.3)式算得

$$n=1+\frac{200}{100}\times 2.31^2=11.67.$$

以自由度 11 及 $\alpha=0.05$,查得临界值 $t_{11}(0.05)=2.20$,再代入(8.5.3)式,计算得

$$n=1+\frac{200}{100}\times 2.20^2=10.68.$$

以自由度 10 及 $\alpha=0.05$,查得临界值 $t_{10}(0.05)=2.23$,再代入(8.5.3)式,计算得 $n=10.94$.这时确定 $n=11$,不必再"试差"了.

在这个例子中,如果事先给定 $\Delta=5$ mA,α 及 S^2 不变,则由(8.5.4)式计算得

$$n=1+\frac{4\times 200}{25}=33.$$

这时不必采用"试差法"了,就确定样本容量为 $n=33$.其直观意义是:置信区间的长度缩短了,置信水平 $1-\alpha$ 不减小,唯有样本容量 n 增大才行.

例 8.5.2 为了判别某种药剂对高血压的疗效,治疗前与治疗后,血压差数的方差一般为 15 左右,即 $\sigma^2 \approx 15$.假定该药有效,要求在显著性水平 $\alpha=0.05$ 下能辨出误差精度 $\Delta=5$ mmHg

专业课辅导+真题资料+微信公众号：432应统书院

(1 mmHg≈133.3 Pa)的血压差,问需多少个病例的试验?

解 本例作为两个总体的数学期望是否相等的配对比较试验,同例8.2.5那样,变为单个总体情形的方法来确定 n. 用 d 表示治疗前后的血压差,取 $S_d^2=15$, $\Delta_d=5$, $\alpha=0.05$,由(8.5.4)式计算得

$$n=1+\frac{4S_d^2}{\Delta_d^2}=3.4.$$

因 n 值太小,则取 $n=5$ 起,采用"试差法"确定 n:由 $\alpha=0.05$, $n=5$,查得临界值 $t_4(0.05)=2.78$,代入(8.5.3)式计算得 $n=5.6$. 今取 $n=6$,对于 $\alpha=0.05$,查得临界值 $t_5(0.05)=2.57$,再代入(8.5.3)式计算得 $n=4.97$. 由于 n 值取整数,则定 $n=5$,即做5个病例的试验就行了.

现在对于两个总体在不配对比较试验中 n 的确定问题讨论如下:

假定总体 ξ 及 η 分别服从正态 $N(a_1,\sigma)$ 及 $N(a_2,\sigma)$,其中 a_1, a_2 及 σ 都是未知参数,在显著性水平 α 下,要判断假设:

$$H_0: a_1=a_2,\ H_1: a_1\neq a_2.$$

由(8.2.5)式知,(a_1-a_2) 的置信水平为 $1-\alpha$ 的置信区间为:

$$\left[\bar{\xi}-\bar{\eta}-S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\,t_{(n_1+n_2-2)}(\alpha),\right.$$
$$\left.\bar{\xi}-\bar{\eta}+S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\,t_{(n_1+n_2-2)}(\alpha)\right],$$

其中 $S^2=\dfrac{1}{n_1+n_2-2}(n_1S_1^2+n_2S_2^2)$. 置信区间的长度记作 2Δ,则 Δ 为

$$\Delta=S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\,t_{(n_1+n_2-2)}(\alpha). \tag{8.5.5}$$

我们仅就 $n_1=n_2=n$ 的情形作讨论,这时试验精度 Δ 为

$$\Delta=S\sqrt{\frac{2}{n}}\,t_{2(n-1)}(\alpha), \tag{8.5.6}$$

专业课辅导+真题资料+微信公众号：432应统书院

从而有

$$n=\frac{2S^2}{\Delta^2}t_{2(n-1)}^2(\alpha). \tag{8.5.7}$$

对于 $\alpha=0.05$，当 $n>15$ 时，$t_{2(n-1)}(\alpha)\approx 2$，于是

$$n\approx\frac{8S^2}{\Delta^2}. \tag{8.5.8}$$

从(8.5.4)同(8.5.8)作比较看到，对于两个总体的数学期望是否相等的检验问题，在相同的试验精度下，用配对比较试验，将比不配对比较试验所需的样本容量可以减少，因而在试验条件许可的情况下，尽量采用配对比较试验为好.

例 8.5.3 若例 8.5.2 中的 Δ、σ^2 值不变，今用不配对比较试验，试问样本容量 n 应多大？

解 先由(8.5.8)式计算得

$$n=\frac{8\times 15}{5^2}=4.8.$$

取定 $n=5$. 对于显著性水平 $\alpha=0.05$，$n=5$，查得临界值 $t_8(0.05)=2.31$，代入(8.5.7)式，计算得

$$n=\frac{2\times 15}{5^2}\times 2.31^2=6.4.$$

取定 $n=6$，对于显著性水平 $\alpha=0.05$，查得临界值 $t_{10}(0.05)=2.23$，再代入(8.5.7)式，计算得

$$n=\frac{2\times 15}{5^2}\times 2.23^2=6,$$

最后确定 $n=6$. 就是说，不用配对比较试验，那么两个样本是各为 6 个病例.

二、最佳检验中 n 的确定

假设总体 ξ 服从正态 $N(a,\sigma)$，其中 a 为未知参数，σ 为已知，记作 σ_0，提出假设检验问题：

$$H_0: a=a_0,\ H_1: a=a_1,$$

专业课辅导+真题资料+微信公众号：432应统书院

其中 $a_0<a_1$. 在给定犯两类错误的概率 α 及 β 的大小，用最佳检验法判断这个假设时，试问样本容量 n 应多大？

我们知道，对于上述假设检验问题，最佳否定域由下式确定

$$\bar{\xi}\geqslant A. \tag{8.5.9}$$

当原假设 H_0 为真时，样本平均$\bar{\xi}$服从正态 $N\left(a_0,\frac{\sigma_0}{\sqrt{n}}\right)$；若 H_0 不为真，即备选假设 H_1 为真，则样本平均$\bar{\xi}$服从正态 $N\left(a_1,\frac{\sigma_0}{\sqrt{n}}\right)$.

因而有方程组：

$$\begin{cases}\displaystyle\int_A^{\infty}\frac{1}{\sqrt{2\pi}(\sigma_0/\sqrt{n})}e^{-\frac{(\bar{x}-a_0)^2}{2(\sigma_0/\sqrt{n})^2}}d\bar{x}=\alpha;\\ \displaystyle\int_{-\infty}^{A}\frac{1}{\sqrt{2\pi}(\sigma_0/\sqrt{n})}e^{-\frac{(\bar{x}-a_1)^2}{2(\sigma_0/\sqrt{n})^2}}d\bar{x}=\beta.\end{cases} \tag{8.5.10}$$

将样本平均$\bar{\xi}$标准正态化，作变换

$$U=\frac{\bar{\xi}-a}{\sigma_0/\sqrt{n}},$$

U 服从正态 $N(0,1)$，则(8.5.10)方程组等价于

$$\begin{cases}\displaystyle\int_{\frac{A-a_0}{\sigma_0/\sqrt{n}}}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du=\alpha;\\ \displaystyle\int_{-\infty}^{\frac{A-a_1}{\sigma_0/\sqrt{n}}}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du=\beta.\end{cases} \tag{8.5.11}$$

在标准正态分布表中，给定 α，查得数值 $u_\alpha=\frac{A-a_0}{\sigma_0/\sqrt{n}}$，使得

$$P\{U\geqslant u_\alpha\}=\alpha.$$

给定 β，查出 $u_\beta=\frac{A-a_1}{\sigma_0/\sqrt{n}}$，使得

$$P\{U\leqslant u_\beta\}=\beta.$$

专业课辅导+真题资料+微信公众号：432应统书院

于是

$$u_\alpha \cdot \frac{\sigma_0}{\sqrt{n}}+a_0=u_\beta \cdot \frac{\sigma_0}{\sqrt{n}}+a_1, \tag{8.5.12}$$

从而解得

$$\begin{cases} n=\dfrac{(u_\alpha-u_\beta)^2 \cdot \sigma_0^2}{(a_1-a_0)^2}; \\ A=\dfrac{a_1u_\alpha-a_0u_\beta}{u_\alpha-u_\beta}. \end{cases} \tag{8.5.13}$$

举例说明：若 $a_0=0, a_1=1, \sigma_0=1, \alpha=\beta=0.05$，则在标准正态分布表中查得

$$u_\alpha=1.645,\ u_\beta=-1.645.$$

由(8.5.13)式计算得

$$n=10.8,\ A=0.5.$$

就是说，样本容量 $n=11$，最佳否定域为 $\bar{\xi}\geqslant 0.5$.

三、验收抽样方案中 n 的确定

例 8.5.4 今要验收一批水泥：如果这种水泥制成混凝土后，断裂强度为5 000(单位)，验收者希望 100 次中有 95 次被“接收”；若断裂强度为4 600(单位)，验收者希望 100 次中只通过 10 次. 已知混凝土断裂强度的均方差 $\sigma=600$(单位)，试为验收者制定验收抽样方案.

解 这种验收抽样方案，在抽样理论中，称之为单式计量验收方案. 设计这种验收方案时，是以总体 ξ 的数学期望 $E(\xi)=a$ 作为验收的质量指标. 现在，我们在给定犯两种类型错误的概率 α 及 β 下，讨论验收抽样方案的设计：定出样本容量 n 的大小，确定对一批成品“接收”或“拒收”的标准 L.

假定总体 ξ 服从正态 $N(a,\sigma)$，那么样本平均 $\bar{\xi}$ 就服从正态 $N\left(a,\dfrac{\sigma}{\sqrt{n}}\right)$，在本例中 $a_0=5\ 000, a_1=4\ 600, \sigma_0=600$. 就是说断裂强

专业课辅导+真题资料+微信公众号：432应统书院

度达到 a_0 时，成品合格，低于 a_1 时，成品不合格（如图 8.5.1）. α 是成品合格时而作出拒收错误的概率，β 是成品不合格时而作出“接收”错误的概率，在本例是取 $\alpha=0.05$，$\beta=0.10$. 于是有方程组

图 8.5.1

$$\begin{cases}\displaystyle\int_{-\infty}^{L}\frac{1}{\sqrt{2\pi}(\sigma_0/\sqrt{n})}e^{-\frac{(\bar{x}-a_0)^2}{2(\sigma_0/\sqrt{n})^2}}d\bar{x}=0.05; & (8.5.14)\\ \displaystyle\int_{L}^{\infty}\frac{1}{\sqrt{2\pi}(\sigma_0/\sqrt{n})}e^{-\frac{(\bar{x}-a_1)^2}{2(\sigma_0/\sqrt{n})^2}}d\bar{x}=0.10. & (8.5.15)\end{cases}$$

将样本平均 $\bar{\xi}$ 标准正态化，则（8.5.14）式等价于

$$\int_{-\infty}^{\frac{L-a_0}{\sigma_0/\sqrt{n}}}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du=0.05;$$

$$\frac{L-a_0}{\sigma_0/\sqrt{n}}=-1.645, \qquad (8.5.16)$$

（8.5.15）式等价于

$$\int_{\frac{L-a_1}{\sigma_0/\sqrt{n}}}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du=0.10;$$

$$\frac{L-a_1}{\sigma_0/\sqrt{n}}=1.282. \qquad (8.5.17)$$

于是

$$L=a_0-\frac{\sigma_0}{\sqrt{n}}\times1.645=a_1+\frac{\sigma_0}{\sqrt{n}}\times1.282,$$

专业课辅导+真题资料+微信公众号：432应统书院

从而解得

$$n=19,\ L=4\ 775. \tag{8.5.18}$$

就是说，在每批待验的成品中只需随机抽取 19 件进行检查. 若混凝土的断裂强度平均值达到或超过 4 775（单位），便“接收”这批成品，否则就不接收这批成品.

习　　题

1. 某化工原料在处理前后取样分析，测得其含脂率的数据如下表：

处理前	0.19	0.18	0.21	0.30	0.66
	0.42	0.08	0.12	0.30	0.27
处理后	0.19	0.24	1.04	0.08	0.20
	0.12	0.31	0.29	0.13	0.07

假定处理前后的含脂率都服从正态分布，且其标准差 σ 不变. 给定显著性水平 $\alpha=0.05$，问处理前后含脂率的平均值有无显著变化.

2. 今有两台机床加工同一零件，分别取 6 个及 9 个零件测其口径，数据记为 $(x_1,x_2,\cdots,x_6)$ 及 $(y_1,y_2,\cdots,y_9)$，计算得：

$$\sum_{i=1}^{6} x_i = 204.6,\qquad \sum_{i=1}^{6} x_i^2 = 6\ 978.93,$$

$$\sum_{j=1}^{9} y_j = 370.8,\qquad \sum_{j=1}^{9} y_j^2 = 15\ 280.173.$$

假定零件口径 ξ 服从正态分布，给定显著性水平 $\alpha=0.05$，问是否可认为这两台机床加工零件口径的方差无显著差异？

3. 某电话站在一小时内接到电话用户的呼唤次数按每分钟记录得如下表：

呼唤次数	0	1	2	3	4	5	6	≥7
频数	8	16	17	10	6	2	1	0

专业课辅导+真题资料+微信公众号：432应统书院

试问这个分布能否看作为泊松分布？

4. 在某公路上 50 分钟之内，记录每 15 秒钟过路汽车的辆数，得到分布情况如下表：

辆数	0	1	2	3	4	5
频数	92	68	28	11	1	0

试问这个分布能否看作为泊松分布？

5. 在数 $\pi=3.14159\cdots$ 的前 800 位小数中，数字 0、1、2、…、9 出现的次数记录如下表：

数字	0	1	2	3	4	5	6	7	8	9
频数	74	92	83	79	80	73	77	75	76	91

试问这个分布能否看作为均匀分布？

6. 检查产品质量时，每次抽取 10 个产品来检查，共抽取 100 次，记录每 10 个产品中的次品数列如下表：

次品数	0	1	2	3	4	5	6	…	10
频数	35	40	18	5	1	1	0	…	0

试问生产过程中出现次品的概率能否看作是不变的，即次品数 ξ 是否服从二项分布？

7. 某香烟厂生产两种香烟，独立地随机抽取容量大小相同的烟叶标本测其尼古丁含量的毫克数，实验室分别作了六次测定，数据记录如下表：

甲	25	28	23	26	29	22
乙	28	23	30	25	21	27

试问这两种香烟的尼古丁含量有无显著差异？给定显著性水平 $\alpha=0.05$，假定含量服从正态分布并具有公共方差.

8. 为了研究慢性气管炎与吸烟量的关系，调查了 385 人，统计数字由下表所示.

专业课辅导+真题资料+微信公众号：432应统书院

类型 \ 人数 \ 烟量	a 支/日	b 支/日	c 支/日	求和
患病者人数	26	147	37	210
健康者人数	30	123	22	175
求和	56	270	59	385

试问慢性气管炎与吸烟量是否有关？给定显著性水平 $\alpha=0.05$.

*9. 设总体 ξ 服从正态 $N(a,\sigma)$，考察如下检验问题：

$$H_0: a=a_0,\ H_1: a=a_1(\neq a_0).$$

证明当样本容量 n 充分大时，可使犯两类错误的概率任意地小.

10. 设总体 ξ 服从正态 $N(a,2)$，$\xi_1,\xi_2,\cdots,\xi_{16}$ 为其样本，样本平均

$$\bar{\xi}=\frac{1}{16}\sum_{i=1}^{16}\xi_i,$$

考虑如下检验问题：

$$H_0: a=0,\ H_1: a\neq 0.$$

试证下述三个否定域有相同的显著性水平 $\alpha=0.05$：

(1) $2\bar{\xi}\leqslant -1.645$；

(2) $1.50\leqslant 2\bar{\xi}\leqslant 2.125$；

(3) $2\bar{\xi}\leqslant -1.96$ 及 $2\bar{\xi}\geqslant 1.96$.

*11. 设总体 ξ 的密度函数为

$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta}e^{-x/\theta}, & 0<x<\infty,\ 0<\theta<\infty;\\ 0, & \text{其他}.\end{cases}$$

$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，分别试求

(1) $H_0:\theta=2,\ H_1:\theta=4$；

(2) $H_0:\theta=2,\ H_1:\theta=1$

的最佳检验，给定显著性水平 $\alpha=0.05$.

*12. 设总体 ξ 服从伯努利分布 $b(1,p)$，即有：

$$P(x;p)=\begin{cases}C_1^x p^x(1-p)^{1-x}, & x=0,1;\\ 0, & \text{其他},\end{cases}$$

专业课辅导+真题资料+微信公众号：432应统书院

$0<p<1$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，对于足够大的 n，分别试求

(1) $H_0:p=\frac{1}{2}$，　$H_1:p=\frac{1}{3}$；

(2) $H_0:p=\frac{1}{3}$，　$H_1:p=\frac{1}{2}$

在显著性水平 $\alpha=0.05$ 下的最佳检验.

*13. 设总体 ξ 服从参数为 λ 的泊松分布：

$$P(x;\lambda)=\frac{\lambda^x}{x!}e^{-\lambda},\ x=0,1,2,\cdots,$$

$\lambda>0$，$\xi_1,\xi_2,\cdots,\xi_{10}$ 是容量为 10 的样本，试求

$$H_0:\lambda=0.1,\quad H_1:\lambda=1$$

在显著性水平 $\alpha=0.05$ 下的最佳检验.

14. 利用切比雪夫不等式，试说明均匀对称的钱币需抛掷多少次，才能使得样本平均值 $\bar{\xi}$ 落在 0.4 到 0.6 之间的概率至少为 0.9？

15. 对方差 σ^2 为已知的正态总体来说，问需抽取容量 n 为多大的样本，才能使得总体的数学期望 a 的置信水平 $1-\alpha$ 的置信区间的长度不大于 L？

16. 证明：若统计量 F 服从 F_{n_1,n_2} 分布，则 $\frac{1}{F}$ 服从 F_{n_2,n_1} 分布，且临界值之间有如下关系式：

$$F_{n_2,n_1}(1-\alpha)=\frac{1}{F_{n_1,n_2}(\alpha)}.$$

17. 现有抽样所得的 520 间商店获利与否和逃税与否的列联表，试以 χ^2 检验判断获利与逃税两者有无关联，给定 $\alpha=0.05$.

逃税与否 \ 获利与否	获利商店数	未获利商店数	合计
逃税商店数	185	53	238
未逃税商店数	215	67	282
合计	400	120	520

18. 从两个电影制片公司 A，B 生产的电影片中随机地抽出 12 部影片查看每部片长(min)，得到如下数据：

公司 A：102，86，98，109，92；

专业课辅导+真题资料+微信公众号：432应统书院

公司 B:81, 165, 91, 134, 92, 87, 114.

分别以 σ_A^2 和 σ_B^2 记两个公司的影片长的方差,假定影片长服从正态分布,这两个样本相互独立.给定 $\alpha=0.01$,检验假设

$$H_0:\sigma_A^2=\sigma_B^2,\qquad H_1:\sigma_A^2\neq\sigma_B^2.$$

专业课辅导+真题资料+微信公众号：432应统书院

第九章　回归分析与方差分析

回归分析是数理统计学中最常用的统计方法.例如在各种预报预测中,习惯上称预报对象这个随机变量为因变量,称与此有关的非随机变量为自变量(或预报因子),需要建立因变量(或预报对象)与一组自变量(或一组预报因子)之间的数学关系式.又如在寻求生产过程的工艺最优化问题中,先要寻求工艺的最优化区域,在还不完全了解生产过程的物理、化学、生物等原理及经验的条件下,用回归分析来解决生产过程的工艺最优化问题是一个比较有效的数学方法.

对于诸如经验公式的求得、因子分析、产品的质量控制;水文、气象、虫害及地震等方面的预报;自动控制中的数学模型及某些新标准的制定等,回归分析往往是一种行之有效的数学工具.

试验设计是数理统计学的一个重要分支,它的主要内容是讨论如何合理地安排试验以及试验后的数据怎样作统计分析等.系统地设计试验并考虑如何进行统计分析的概念,是由费希尔在他的著作《试验之设计》中首先提出来的.方差分析是试验结果的数据分析中一种常用的统计方法.试验设计开始出现于农田科学实验,第二次世界大战后才逐渐地被应用于工业科学实验中,如今在各个领域中都有其广泛应用.试验设计的内容很丰富,有如均匀设计,平衡不完全区组设计,正交设计等方面,参看书目[4]及[9].在这一章中,我们着重讨论数据分析中的方差分析法.

§9.1　线性回归分析

我们从几个例子谈起.

例 9.1.1　我们知道,一个做等速直线运动的质点,在时刻 t

专业课辅导+真题资料+微信公众号：432应统书院

的位置是 s，一般它可以用公式 $s=\beta_0+\beta_1 t$ 来描写，其中 β_0 是质点在 $t=0$ 时刻的初始位置，β_1 是平均速度. 如果 β_0 及 β_1 是未知的，s 可以被观察，那么我们只要在两个不同的时刻观察质点的位置 s，就可以解出 β_0 及 β_1 的值. 由于种种原因，质点的位置不能被精确地观察，而是带有随机的测量误差. 也就是说 s 不能真正被观察，而观察到的数据是 $y=s+\varepsilon$，其中 ε 是测量（或观察）误差. 于是我们有公式

$$y=s+\varepsilon=\beta_0+\beta_1 t+\varepsilon, \tag{9.1.1}$$

其中 t 是可观察的非随机变量. 对于 ε，我们把它看作是数学期望为零的随机变量，在一般情况下是允许这样做的，或者说这个观点是符合实际情况的. 因此，y 看作为一个随机变量，它是可观察的随机变量，其数学期望为 s，β_0 及 β 是未知参数，通过对 t 及 y 的 n 次观察得 n 组观察值 (t_i, y_i)，$i=1,2,\cdots,n$，即有

$$y_i=\beta_0+\beta_1 t_i+\varepsilon_i \quad (i=1,2,\cdots,n).$$

下面用向量、矩阵的写法表示如下：

$$\boldsymbol{y}=\boldsymbol{X\beta}+\boldsymbol{\varepsilon}, \tag{9.1.2}$$

其中

$$\boldsymbol{y}=\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \boldsymbol{\varepsilon}=\begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix},$$

$$\boldsymbol{X}=\begin{pmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{pmatrix}, \quad \boldsymbol{\beta}=\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix},$$

$$E(\boldsymbol{y})=\boldsymbol{X\beta}.$$

我们利用观察数据 $\boldsymbol{y}$ 及 $\boldsymbol{X}$ 对未知参数 $\boldsymbol{\beta}$ 作出估计.

例 9.1.2 小麦的亩产量记为 y，它与水（x_1）、肥料（x_2）、土质（x_3）、麦种（x_4）、栽培技术（x_5）及管理措施（x_6）等因素有关. 就是

专业课辅导+真题资料+微信公众号：432应统书院

说，小麦亩产量 y 与 x_1,x_2,x_3,x_4,x_5 及 x_6 等都有一定的关系. 但是由于观察(或试验)中总存在随机因素的影响，即使 $x_1,x_2,\cdots,x_6$ 相对固定，小麦的亩产量 y 也不完全相同，因此我们把 y 与 x_1,x_2,x_3,x_4,x_5 及 x_6 的关系分为两部分来研究，即有

$$y=f(x_1,x_2,\cdots,x_6)+\varepsilon, \tag{9.1.3}$$

其中 $f(x_1,x_2,\cdots,x_6)$ 表示 6 个能控制的因素 $x_1,x_2,\cdots,x_6$ 与 y 的关系，$f(x_1,x_2,\cdots,x_6)$ 是普通的函数表达式，是非随机性的部分. ε 表示随机因素对小麦亩产量 y 的影响作用，把 ε 看作数学期望为零的随机变量. 因此，y 为一随机变量，是可观察的随机变量，数学期望为 $f(x_1,x_2,\cdots,x_6)$. 在一般情况下，$f(x_1,x_2,\cdots,x_6)$ 不一定是 $x_1,x_2,\cdots,x_6$ 的线性函数，为处理方便起见，可近似地当作线性的情形，即

$$y=\beta_1x_1+\cdots+\beta_6x_6+\varepsilon. \tag{9.1.4}$$

这样，(9.1.4)式是 y 关于因子 $x_1,x_2,\cdots,x_6$ 的线性函数；当我们对 $(y;x_1,x_2,\cdots,x_6)$ 进行 n 次观察(或试验)后，则(9.1.4)式是未知参数 $\beta_1,\beta_2,\cdots,\beta_6$ 的线性函数.

在这个例子中，有些因子是非数量化的，如小麦种类是非数量化的因子. 对于非数量化的因子，可适当给予数量化，以便于数学上计算分析. 如小麦品种有甲、乙、丙三种，规定 x_4(甲)=1，x_4(乙)=2，x_4(丙)=3. 在某些例子中，有些因子的取值过大而有些因子的取值过小，影响对因子所起作用大小的分析，这时可对各个因子的观察值作标准化处理，然后用标准化后的数值作回归分析计算.

例 9.1.3 在考察某化学反应时，发现其反应速度 y 主要决定于所用某种催化剂的用量 x 及所加的温度 T. 当 x 及 T 改变时，反应速度 y 也跟着改变. 但是由于观察(或测试)及其他条件的影响，当 x 及 T 固定时，所观察到的反应速度也不完全相同，即反应速度的观察值 y 与 x、T 之间存在着关系式

$$y=f(x,T)+\varepsilon, \tag{9.1.5}$$

其中 $f(x,T)$ 是 x 及 T 的普通二元函数，是非随机性的部分；ε 表示

专业课辅导+真题资料+微信公众号：432应统书院

除 x 及 T 之外的对 y 起综合作用的各种次要条件，这些条件或是不可控制的，或是不可观察的. 这些次要条件都归属于“事物之间总存在着随机因素的影响”这个范畴来考虑，我们把 ε 看作是数学期望为零的随机变量. 因此，y 是数学期望为 $f(x,T)$ 的可观察的随机变量，即

$$E(y)=f(x,T).$$

我们通过试验（在 x 及 T 的不同的取值下进行试验），能观察 y 的值，用回归分析法找出 $f(x,T)$ 的具体数学表达式. 有了 $f(x,T)$ 这个数学表达式，对今后实验工作很方便. 例如，通过控制 x 及 T 以达到控制反应速度 y. 为处理方便起见，我们对（9.1.5）中的 $f(x,T)$ 用近似的线性函数来代替，即

$$y=\beta_0+\beta_1 x+\beta_2 T+\varepsilon. \tag{9.1.6}$$

我们通过 n 组观察数据 $(y_i;x_i,T_i)$，$i=1,2,\cdots,n$，对未知参数 β_0、β_1 及 β_2 作出估计.

综上所述，在这一章中，我们着重讨论线性回归分析，归纳出一般的线性模型.

设因变量 y 和自变量 $x_1,x_2,\cdots,x_k$ 之间服从如下线性关系：

$$y=\beta_1 x_1+\beta_2 x_2+\cdots+\beta_k x_k+\varepsilon. \tag{9.1.7}$$

对 y 及 $x_1,x_2,\cdots,x_k$ 同时作了 n 次观察（或试验）之后，得到 n 组数据 $(y_t;x_{t1},x_{t2},\cdots,x_{tk})$ $(t=1,2,\cdots,n)$. 它们满足下述线性关系式：

$$y_t=\beta_1 x_{t1}+\beta_2 x_{t2}+\cdots+\beta_k x_{tk}+\varepsilon_t \quad (t=1,2,\cdots,n) \tag{9.1.8}$$

（从（9.1.8）式起用 t 表示下标）. 用向量、矩阵形式记作

$$\boldsymbol{y}=\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_{n\times 1}, \qquad \boldsymbol{\beta}=\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}_{k\times 1},$$

$$\boldsymbol{X}=\begin{pmatrix} x_{11}\cdots x_{1k} \\ \vdots \qquad\quad \vdots \\ x_{n1}\cdots x_{nk} \end{pmatrix}_{n\times k}, \qquad \boldsymbol{\varepsilon}=\begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}_{n\times 1},$$

专业课辅导+真题资料+微信公众号：432应统书院

则(9.1.8)式可表示为

$$\boldsymbol{y}=\boldsymbol{X\beta}+\boldsymbol{\varepsilon}, \tag{9.1.9}$$

其中 $\boldsymbol{X}$ 是已知的 $n\times k$ 阶常数矩阵；$\boldsymbol{\beta}$ 是 k 维的未知参数向量，$\boldsymbol{\varepsilon}$ 是数学期望为零的 n 维随机向量，$\boldsymbol{y}$ 是已知的 n 维观察向量.

我们对 ε 作如下假定：

$$\left.\begin{aligned} &E(\boldsymbol{\varepsilon})=\boldsymbol{0};\\ &\operatorname{cov}(\boldsymbol{\varepsilon},\boldsymbol{\varepsilon})=\sigma^2\boldsymbol{I}_{n\times n}, \end{aligned}\right\} \tag{9.1.10}$$

其中 σ^2 是未知参数，$\boldsymbol{I}_{n\times n}$ 是单位矩阵，即有：

$$\left.\begin{aligned} &E(\varepsilon_t)=0;\\ &D(\varepsilon_t)=\sigma^2;\\ &\operatorname{cov}(\varepsilon_t,\varepsilon_s)=0,t\neq s \end{aligned}\right\} \tag{9.1.11}$$

$(t=1,2,\cdots,n;s=1,2,\cdots,n)$. 就是说，对随机误差项 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 作这样的假定：无偏性、等方差性、不相关性，这种假定在一般情况下是合理的或允许的.

对于(9.1.9)中 $\boldsymbol{y}$，我们称它是服从线性模型，并且简记作 $(\boldsymbol{y},\boldsymbol{X\beta},\sigma^2\boldsymbol{I}_{n\times n})$. 习惯上，我们说线性模型是指(9.1.9)及(9.1.10)两式组成. 由(9.1.9)及(9.1.10)两式容易算得

$$\left.\begin{aligned} &E(\boldsymbol{y})=\boldsymbol{X\beta};\\ &\operatorname{cov}(\boldsymbol{y},\boldsymbol{y})=\sigma^2\boldsymbol{I}_{n\times n} \end{aligned}\right\} \tag{9.1.12}$$

通常称(9.1.12)式为高斯-马尔可夫线性模型.

若对 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 作具有独立同分布 $N(0,\sigma)$ 的假定，则称(9.1.12)式为正态线性模型.

对于线性回归分析$(\boldsymbol{y},\boldsymbol{X\beta},\sigma^2\boldsymbol{I}_{n\times n})$，所考虑的统计推断问题主要是：

(1) 对未知参数向量 $\boldsymbol{\beta}$ 及 σ^2 进行估计；

(2) 对关于 $\boldsymbol{\beta}$ 的某种假设以及 $\boldsymbol{y}$ 服从线性模型的假设进行检验；

(3) 对 $\boldsymbol{y}$ 进行预测与控制.

专业课辅导+真题资料+微信公众号：432应统书院

在这一章中所作的讨论，总假定 $n>k$ 及常数矩阵 $\boldsymbol{L}$ 是满秩的，其中

$$\boldsymbol{L}=\boldsymbol{X}'\boldsymbol{X}①,$$

$\boldsymbol{L}$ 是 $k\times k$ 矩阵，它的秩 $\operatorname{ran}(\boldsymbol{L})=k$. 关于 $\boldsymbol{L}$ 是降秩的讨论，在此从略.

§9.2 最小二乘法估计

这一节主要讨论线性模型(9.1.9)、(9.1.10)中未知参数 $\beta_1,\beta_2,\cdots,\beta_k$ 及 σ^2 的点估计，即 $(\boldsymbol{y},\boldsymbol{X\beta},\sigma^2\boldsymbol{I}_{n\times n})$ 中的未知参数向量 $\boldsymbol{\beta}$ 及 σ^2 的点估计.

一、参数的最小二乘法估计

我们考虑

$$Q=\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}=(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta}), \tag{9.2.1}$$

亦即

$$Q=\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}=\sum_{t=1}^{n}\varepsilon_t^2=\sum_{t=1}^{n}\left[y_t-\sum_{i=1}^{k}\beta_i x_{ti}\right]^2. \tag{9.2.2}$$

Q 是 n 次观察中误差项 ε_t^2 之和，称 Q 为误差平方和，它反映了 y 与 $\sum\limits_{i=1}^{k}\beta_i x_i$ 之间 n 次观察中总的误差程度.

最小二乘法原理是这样：寻找使得 Q 达到最小值的 $\hat{\beta}_1\ \hat{\beta}_2,\cdots,\hat{\beta}_k$ 作为 $\beta_1,\beta_2,\cdots,\beta_k$ 的点估计. 这时称 $\hat{\beta}_i$ 为参数 β_i 的最小二乘法估计量 $(i=1,2,\cdots,k)$. 由于(9.2.2)中的 Q 是 β_i 的非负的二次函数，求 Q 的最小解是有可能的. 所要求的 $\hat{\beta}_i$ 应满足如下关系式：

$$\sum_{t=1}^{n}\left[y_t-\sum_{i=1}^{k}\hat{\beta}_i x_{ti}\right]^2=\min_{\{\beta_1,\beta_2,\cdots,\beta_k\}}\left\{\sum_{t=1}^{n}\left[y_t-\sum_{i=1}^{k}\beta_i x_{ti}\right]^2\right\}. \tag{9.2.3}$$

① 本书中矩阵 $\boldsymbol{A}$ 的转置用 $\boldsymbol{A}'$ 表示，例如此处以 $\boldsymbol{X}'$ 表示 $\boldsymbol{X}$ 的转置. 以下均同.

专业课辅导+真题资料+微信公众号：432应统书院

我们用微积分求满足(9.2.3)式的解$\hat{\beta}_i(i=1,2,\cdots,k)$. 将(9.2.2)式定义的 Q 对 β_i 分别求偏导数得：

$$\frac{\partial Q}{\partial \beta_i}=-2\sum_{t=1}^{n}\left[y_t-\sum_{j=1}^{k}\beta_j x_{tj}\right]\cdot x_{ti},\quad i=1,2,\cdots,k.$$

令其偏导数等于零，方程组的解用$\hat{\beta}_j$ 表示，即有

$$\sum_{t=1}^{n}y_t x_{ti}-\sum_{t=1}^{n}\sum_{j=1}^{k}\hat{\beta}_j x_{tj}\cdot x_{ti}=0,\quad i=1,2,\cdots,k,$$

或写为

$$\sum_{j=1}^{k}\hat{\beta}_j\left[\sum_{t=1}^{n}x_{tj}\cdot x_{ti}\right]=\sum_{t=1}^{n}y_t x_{ti},\quad i=1,2,\cdots,k. \tag{9.2.4}$$

记

$$\sum_{t=1}^{n}x_{ti}x_{tj}=l_{ij},\quad \sum_{t=1}^{n}x_{ti}y_t=l_{iy},$$

$$\boldsymbol{L}=(l_{ij})_{k\times k}=\boldsymbol{X}'\boldsymbol{X}, \tag{9.2.5}$$

$\boldsymbol{L}$ 为对称的常数矩阵. 又记

$$\boldsymbol{X}'\boldsymbol{y}=(l_{iy})_{k\times 1}, \tag{9.2.6}$$

则(9.2.4)式可写为

$$(\boldsymbol{X}'\boldsymbol{X})\hat{\boldsymbol{\beta}}=\boldsymbol{X}'\boldsymbol{y},$$

即有

$$\boldsymbol{L}\hat{\boldsymbol{\beta}}=\boldsymbol{X}'\boldsymbol{y}. \tag{9.2.7}$$

通常称(9.2.7)式为正规方程，如果 $\boldsymbol{L}$ 是满秩的，即 $|\boldsymbol{L}|\neq 0$，其秩为 k，那么正规方程(9.2.7)式有唯一解，记作

$$\hat{\boldsymbol{\beta}}=\boldsymbol{L}^{-1}\boldsymbol{X}'\boldsymbol{y}, \tag{9.2.8}$$

其中 $\boldsymbol{L}^{-1}=(\boldsymbol{X}'\boldsymbol{X})^{-1}$ 为 $\boldsymbol{L}$ 的唯一的逆矩阵.

* 附注：证明(9.2.8)式中的$\hat{\boldsymbol{\beta}}$ 为线性模型$(\boldsymbol{y},\boldsymbol{X}\boldsymbol{\beta},\sigma^2\boldsymbol{I}_{n\times n})$中未知参数向量$\boldsymbol{\beta}$ 的最小二乘法估计量，亦即$\hat{\beta}_i(i=1,2,\cdots,k)$满足

专业课辅导+真题资料+微信公众号：432应统书院

关系式(9.2.3). 现用矩阵形式来叙述其证明步骤. 从(9.2.1)式知道,对任意$\boldsymbol{\beta}$,

$$Q=(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta}),$$

则有

$$\begin{aligned}(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta})&=[(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})+\boldsymbol{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})]'[(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})\\&\quad+\boldsymbol{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})]\\&=(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})+(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{X}'\boldsymbol{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\\&\quad+(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})'\boldsymbol{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})+(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{X}'(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})\\&\geqslant(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}).\end{aligned}$$

上述证明过程中应用了如下结果:

$$(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{X}'\boldsymbol{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})=[\boldsymbol{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})]'[\boldsymbol{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})]\geqslant\boldsymbol{0},$$

$$\begin{aligned}(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})'\boldsymbol{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})&=(\boldsymbol{y}'\boldsymbol{X}-\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X})(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\\&=(\boldsymbol{y}'\boldsymbol{X}-\boldsymbol{y}'\boldsymbol{X})(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})=\boldsymbol{0}.\end{aligned}$$

至此,在$|\boldsymbol{L}|\neq 0$时,证明了(9.2.8)式中的$\hat{\boldsymbol{\beta}}$是$\boldsymbol{\beta}$的最小二乘法估计量.

在实际工作中,常称

$$\hat{y}=\hat{\beta}_1x_1+\hat{\beta}_2x_2+\cdots+\hat{\beta}_kx_k \tag{9.2.9}$$

为线性回归方程. 在今后工作中,当自变量$x_1,x_2,\cdots,x_k$分别取值$x_{1t},x_{2t},\cdots,x_{kt}$时,由(9.2.9)式算得$\hat{y}_t$值作为$y_t$的估计值,即可用自变量$x_1,x_2,\cdots,x_k$由(9.2.9)式来预测因变量$y$(实际上是对随机变量$y$的数学期望$E(y)$的估计)的取值.

二、最小二乘法估计量的性质

首先,我们对高斯-马尔可夫线性模型,讨论按(9.2.8)式确定$\boldsymbol{\beta}$的最小二乘法估计量$\hat{\boldsymbol{\beta}}$的一些基本性质.

性质1 $\hat{\boldsymbol{\beta}}$是$\boldsymbol{\beta}$的线性无偏估计量.

专业课辅导+真题资料+微信公众号：432应统书院

证 由于$\hat{\boldsymbol{\beta}}=\boldsymbol{L}^{-1}\boldsymbol{X}'\boldsymbol{y}$，每一个$\hat{\beta}_i$都是$y_1,y_2,\cdots,y_n$的线性组合，因而$\hat{\beta}_i$是$\beta_i$的线性估计量，即$\hat{\boldsymbol{\beta}}$为$\boldsymbol{\beta}$的线性估计量.

$$\begin{aligned}E(\hat{\boldsymbol{\beta}})&=E(\boldsymbol{L}^{-1}\boldsymbol{X}'\boldsymbol{y})=\boldsymbol{L}^{-1}\boldsymbol{X}'\cdot E(\boldsymbol{y})\\&=\boldsymbol{L}^{-1}\boldsymbol{X}'\cdot\boldsymbol{X}\boldsymbol{\beta}=\boldsymbol{L}^{-1}\boldsymbol{L}\boldsymbol{\beta}=\boldsymbol{\beta},\end{aligned}\tag{9.2.10}$$

即

$$E(\hat{\beta}_i)=\beta_i\quad(i=1,2,\cdots,k).$$

性质 2 $\hat{\boldsymbol{\beta}}$的协方差矩阵为$\sigma^2\boldsymbol{L}^{-1}$.

证 记$\boldsymbol{B}=\boldsymbol{L}^{-1}\boldsymbol{X}'$，即$\hat{\boldsymbol{\beta}}=\boldsymbol{B}\boldsymbol{y}$，

$$\begin{aligned}\operatorname{cov}(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\beta}})&=E\{[\boldsymbol{B}y-\boldsymbol{B}E(y)][\boldsymbol{B}\boldsymbol{y}-\boldsymbol{B}E(\boldsymbol{y})]'\}\\&=\boldsymbol{B}\cdot E\{[\boldsymbol{y}-E(\boldsymbol{y})][\boldsymbol{y}-E(\boldsymbol{y})]'\}\cdot\boldsymbol{B}'\\&=\boldsymbol{B}\cdot\sigma^2\boldsymbol{I}_{n\times n}\cdot\boldsymbol{B}'\\&=\boldsymbol{L}^{-1}\boldsymbol{X}'\cdot\sigma^2\boldsymbol{I}_{n\times n}\cdot(\boldsymbol{L}^{-1}\boldsymbol{X}')'\\&=\boldsymbol{L}^{-1}\boldsymbol{X}'\cdot\sigma^2\boldsymbol{I}_{n\times n}\cdot\boldsymbol{X}\boldsymbol{L}^{-1}=\sigma^2\boldsymbol{L}^{-1}.\end{aligned}\tag{9.2.11}$$

记

$$\boldsymbol{L}^{-1}=\boldsymbol{C}=(c_{ij})_{k\times k},$$

则有

$$\left.\begin{aligned}&D(\hat{\beta}_i)=\sigma^2c_{ii}\quad(i=1,2,\cdots,k);\\&\operatorname{cov}(\hat{\beta}_i,\hat{\beta}_j)=\sigma^2c_{ij}\quad(i,j=1,2,\cdots,k).\end{aligned}\right\}$$

性质 3 记

$$\boldsymbol{e}=\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}=\boldsymbol{y}-\hat{\boldsymbol{y}},\tag{9.2.12}$$

其中

$$\hat{\boldsymbol{y}}=\boldsymbol{X}\hat{\boldsymbol{\beta}},\tag{9.2.13}$$

称$\boldsymbol{y}$为观察向量，$\hat{\boldsymbol{y}}$为由$\boldsymbol{X}$对$\boldsymbol{y}$的估计向量，$\boldsymbol{e}$为剩余向量，则有：

$$E(\boldsymbol{e})=\boldsymbol{0}，即\ E(e_t)=0,t=1,2,\cdots,n.\tag{9.2.14}$$

$$\operatorname{cov}(\hat{\boldsymbol{\beta}},\boldsymbol{e})=\boldsymbol{0}，即对每一\hat{\beta}_i,i=1,2,\cdots,k,$$

$$\operatorname{cov}(\hat{\beta}_i,e_t)=0,t=1,2,\cdots,n.\tag{9.2.15}$$

专业课辅导+真题资料+微信公众号：432应统书院

证 $E(\boldsymbol{e})=E(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})=E(\boldsymbol{y})-\boldsymbol{X}E(\hat{\boldsymbol{\beta}})$

$$=\boldsymbol{X\beta}-\boldsymbol{X\beta}=\boldsymbol{0},$$

$$\begin{aligned}\operatorname{cov}(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{e}})&=\operatorname{cov}(\hat{\boldsymbol{\beta}},\boldsymbol{y})-\operatorname{cov}(\hat{\boldsymbol{\beta}},\boldsymbol{X}\hat{\boldsymbol{\beta}})\\&=\operatorname{cov}(\boldsymbol{L}^{-1}\boldsymbol{X}'\boldsymbol{y},\boldsymbol{y})-\operatorname{cov}(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\beta}})\boldsymbol{X}'\\&=\boldsymbol{L}^{-1}\boldsymbol{X}'\cdot\sigma^2-\sigma^2\boldsymbol{L}^{-1}\boldsymbol{X}'=\boldsymbol{0}.\end{aligned}$$

*最小二乘法估计的几何解释如下：

$$\begin{aligned}\operatorname{cov}(\boldsymbol{y},\boldsymbol{y})&=\operatorname{cov}(\boldsymbol{X}\hat{\boldsymbol{\beta}}+\boldsymbol{e},\boldsymbol{X}\hat{\boldsymbol{\beta}}+\boldsymbol{e})\\&=\operatorname{cov}(\boldsymbol{X}\hat{\boldsymbol{\beta}},\boldsymbol{X}\hat{\boldsymbol{\beta}})+\operatorname{cov}(\boldsymbol{e},\boldsymbol{e}),\end{aligned}\tag{9.2.16}$$

$$\begin{aligned}\boldsymbol{y}'\boldsymbol{y}&=(\hat{\boldsymbol{y}}+\boldsymbol{e})'(\hat{\boldsymbol{y}}+\boldsymbol{e})\\&=\hat{\boldsymbol{y}}'\hat{\boldsymbol{y}}+\hat{\boldsymbol{y}}'\boldsymbol{e}+\boldsymbol{e}'\hat{\boldsymbol{y}}+\boldsymbol{e}'\boldsymbol{e}\\&=\hat{\boldsymbol{y}}'\hat{\boldsymbol{y}}+\boldsymbol{e}'\boldsymbol{e},\end{aligned}\tag{9.2.17}$$

其中

$$\begin{aligned}\hat{\boldsymbol{y}}'\boldsymbol{e}&=\hat{\boldsymbol{\beta}}'\boldsymbol{X}'(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})\\&=\hat{\boldsymbol{\beta}}'(\boldsymbol{X}'\boldsymbol{y}-\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}})=\boldsymbol{0}.\end{aligned}$$

同理

$$\boldsymbol{e}'\hat{\boldsymbol{y}}=\boldsymbol{0}.$$

称$\sqrt{\boldsymbol{y}'\boldsymbol{y}}$、$\sqrt{\hat{\boldsymbol{y}}'\hat{\boldsymbol{y}}}$及$\sqrt{\boldsymbol{e}'\boldsymbol{e}}$分别为“观察向量”、“估计向量”及“剩余向量”的长度.

由正规方程(9.2.7)知道：

$$\boldsymbol{X}'(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})=\boldsymbol{0},$$

即

$$\boldsymbol{X}'\boldsymbol{e}=\boldsymbol{0},$$

$$\sum_{t=1}^{n}x_{ti}e_t=0\quad(i=1,2,\cdots,k).$$

可见“剩余向量”$\boldsymbol{e}$与k个向量$\boldsymbol{x}_i=(x_{1i},x_{2i},\cdots,x_{ni})$都是正交的，$i=1,2,\cdots,k$. “估计向量”$\hat{\boldsymbol{y}}=\boldsymbol{X}\hat{\boldsymbol{\beta}}$是属于$k$个向量$\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k$所张成的空间，这个空间称之为“估计空间”. 可见，$\boldsymbol{e}$与“估计空间”垂直. 因此，我们可以这样说：“在最小二乘法中，估计向量$\hat{\boldsymbol{y}}=\boldsymbol{X}\hat{\boldsymbol{\beta}}$是

专业课辅导+真题资料+微信公众号：432应统书院

观察向量 y 在估计空间上的垂直投影”(图 9.2.1).

图 9.2.1

性质 4 $E(\boldsymbol{e}'\boldsymbol{e})=(n-k)\sigma^2$，记 $Q_e=\boldsymbol{e}'\boldsymbol{e}$，则 $E\left(\dfrac{Q_e}{n-k}\right)=\sigma^2$，即 $\dfrac{Q_e}{n-k}$ 为 σ^2 的无偏估计量.

***证** 因为

$$Q_e=\boldsymbol{e}'\boldsymbol{e}=(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}), \tag{9.2.18}$$

$$\begin{aligned}
EQ_e &= E(\boldsymbol{e}'\boldsymbol{e})=E[\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{X}\boldsymbol{L}^{-1}\boldsymbol{X}')\boldsymbol{y}] \\
&= E[(\boldsymbol{y}-E\boldsymbol{y})'(\boldsymbol{I}-\boldsymbol{X}\boldsymbol{L}^{-1}\boldsymbol{X}')(\boldsymbol{y}-E\boldsymbol{y})] \\
&\quad +(E\boldsymbol{y})'(\boldsymbol{I}-\boldsymbol{X}\boldsymbol{L}^{-1}\boldsymbol{X}')(E\boldsymbol{y}) \\
&= \sigma^2\operatorname{trace}[\boldsymbol{I}-\boldsymbol{X}\boldsymbol{L}^{-1}\boldsymbol{X}'] \\
&= n\sigma^2-\sigma^2\operatorname{trace}(\boldsymbol{X}\boldsymbol{L}^{-1}\boldsymbol{X}') \\
&= n\sigma^2-\sigma^2\operatorname{trace}(\boldsymbol{L}^{-1}\boldsymbol{X}\boldsymbol{X}') \\
&= n\sigma^2-\sigma^2\operatorname{trace}(\boldsymbol{L}^{-1}\boldsymbol{L}) \\
&= n\sigma^2-k\sigma^2=(n-k)\sigma^2.
\end{aligned}$$

其中 $\operatorname{trace}[\boldsymbol{A}]$ 表示矩阵 $\boldsymbol{A}$ 的迹. 记

$$\hat{\sigma}_e^2=\frac{Q_e}{n-k}, \tag{9.2.19}$$

称 $\hat{\sigma}_e^2$ 为剩余方差，它是 σ^2 的无偏估计量.

性质 5 $\hat{\boldsymbol{\beta}}$ 为 $\boldsymbol{\beta}$ 的最优线性无偏估计量.

专业课辅导+真题资料+微信公众号：432应统书院

*证　$\hat{\boldsymbol{\beta}}=\boldsymbol{L}^{-1}\boldsymbol{X}'\boldsymbol{y}, E\hat{\boldsymbol{\beta}}=\boldsymbol{\beta}$,

$$\operatorname{cov}(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\beta}})=\sigma^2\boldsymbol{L}^{-1}, 0<\sigma^2<\infty,$$

设 $\boldsymbol{T}=\boldsymbol{A}\boldsymbol{y}$ 为 $\boldsymbol{\beta}$ 的任一线性无偏估计，则

$$E\boldsymbol{T}=\boldsymbol{A}\boldsymbol{X}\boldsymbol{\beta},$$

知

$$\boldsymbol{A}\boldsymbol{X}=\boldsymbol{I}_{k\times k},$$

$$\operatorname{cov}(\boldsymbol{T},\boldsymbol{T})=\boldsymbol{A}\operatorname{cov}(\boldsymbol{y},\boldsymbol{y})\boldsymbol{A}'=\sigma^2\boldsymbol{A}\boldsymbol{A}'. \tag{9.2.20}$$

记 $\boldsymbol{D}=\boldsymbol{A}\boldsymbol{A}'-\boldsymbol{L}^{-1}$，我们来证矩阵 $\boldsymbol{D}$ 是对角线元素为非负数的非负定矩阵. 事实上有：

$$\begin{aligned}&(\boldsymbol{A}-\boldsymbol{L}^{-1}\boldsymbol{X}')(\boldsymbol{A}-\boldsymbol{L}^{-1}\boldsymbol{X}')'\\=&\boldsymbol{A}\boldsymbol{A}'-\boldsymbol{A}\boldsymbol{X}\boldsymbol{L}^{-1}-\boldsymbol{L}^{-1}\boldsymbol{X}'\boldsymbol{A}'+\boldsymbol{L}^{-1}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{L}^{-1}\\=&\boldsymbol{A}\boldsymbol{A}'-\boldsymbol{I}\boldsymbol{L}^{-1}-\boldsymbol{L}^{-1}\boldsymbol{I}+\boldsymbol{L}^{-1}\boldsymbol{L}\boldsymbol{L}^{-1}\\=&\boldsymbol{A}\boldsymbol{A}'-\boldsymbol{L}^{-1}=\boldsymbol{D}.\end{aligned}$$

这个性质通常称为高斯-马尔可夫定理.

现在，我们对正态线性模型，进一步讨论 $\boldsymbol{\beta}$ 的最小二乘法估计量 $\hat{\boldsymbol{\beta}}$ 的分布函数问题，并讨论 Q_e 的分布函数问题. Q_e 是误差平方和 Q 在 $\boldsymbol{\beta}$ 用 $\hat{\boldsymbol{\beta}}$ 作估计量时达到的最小值，这可由(9.2.1)、(9.2.3)及(9.2.18)三式看到，今后也称 Q_e 为误差平方和.

性质6　假定 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 为总体 ε 的样本，总体 ε 服从正态 $N(0,\sigma)$，则

(1) $\hat{\boldsymbol{\beta}}$ 与 $\boldsymbol{e}$ 相互独立，$\hat{\boldsymbol{\beta}}$ 与 Q_e 相互独立；

(2) $\hat{\boldsymbol{\beta}}$ 与 $\boldsymbol{e}$ 都服从正态分布，数学期望、协方差分别为

$$E(\hat{\boldsymbol{\beta}})=\boldsymbol{\beta}, \operatorname{cov}(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\beta}})=\sigma^2\boldsymbol{L}^{-1},$$

$$E(\boldsymbol{e})=0, \operatorname{cov}(\boldsymbol{e},\boldsymbol{e})=\sigma^2(\boldsymbol{I}_{n\times n}-\boldsymbol{X}\boldsymbol{L}^{-1}\boldsymbol{X}');$$

(3) $Q_e/\sigma^2\sim\chi^2_{(n-k)}$.　(9.2.21)

*证　由(9.2.8)及(9.2.18)两式知，

$$\hat{\boldsymbol{\beta}}=\boldsymbol{L}^{-1}\boldsymbol{X}'\boldsymbol{y},$$

$$Q_e=(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})=\boldsymbol{y}'[\boldsymbol{I}_{n\times n}-\boldsymbol{X}\boldsymbol{L}^{-1}\boldsymbol{X}']\boldsymbol{y},$$

专业课辅导+真题资料+微信公众号：432应统书院

记

$$B=L^{-1}X',A=I_{n\times n}-XL^{-1}X',$$

则

$$\hat{\boldsymbol{\beta}}=\boldsymbol{B}\boldsymbol{y},\quad Q_e=\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}.$$

(1) $\hat{\boldsymbol{\beta}}$为 $\boldsymbol{y}$ 的线性型(是指每一 β_i 为 $y_1,y_2,\cdots,y_n$ 的线性组合),Q_e 为 $\boldsymbol{y}$ 的半正定型(是指以 $y_1,y_2,\cdots,y_n$ 为变量的半正定型,或称之为非负定型). $\boldsymbol{B}$ 为 $k\times n$ 矩阵,$\boldsymbol{A}$ 为对称的 $n\times n$ 矩阵. 从矩阵代数知道,存在一正交矩阵 $\boldsymbol{H}_{n\times n}$,使得 $\boldsymbol{D}=\boldsymbol{H}'\boldsymbol{A}\boldsymbol{H}$ 为对角线矩阵,记为

$$\boldsymbol{D}=\begin{pmatrix}\boldsymbol{D}_1 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}_{n\times n},$$

其中 $\boldsymbol{D}_1$ 为 $n-k$ 阶的对角线元素非零的对角线矩阵. 同理,有

$$\boldsymbol{F}=\boldsymbol{B}\boldsymbol{H}=\begin{pmatrix}\boldsymbol{F}_{11} & \boldsymbol{F}_{12}\\ \boldsymbol{F}_{21} & \boldsymbol{F}_{22}\end{pmatrix}_{k\times n}.$$

因为

$$\begin{aligned}\boldsymbol{B}\boldsymbol{A}&=(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'[\boldsymbol{I}-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}']\\&=(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{X}'-\boldsymbol{X}')=\boldsymbol{0},\end{aligned}$$

所以

$$\begin{aligned}\boldsymbol{0}&=\boldsymbol{B}\boldsymbol{A}\boldsymbol{H}=\boldsymbol{B}\boldsymbol{H}\boldsymbol{H}'\boldsymbol{A}\boldsymbol{H}=\boldsymbol{F}\boldsymbol{D}\\&=\begin{pmatrix}\boldsymbol{F}_{11} & \boldsymbol{F}_{12}\\ \boldsymbol{F}_{21} & \boldsymbol{F}_{22}\end{pmatrix}\begin{pmatrix}\boldsymbol{D}_1 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}.\end{aligned}$$

推得

$$\boldsymbol{F}_{11}\boldsymbol{D}_1=\boldsymbol{F}_{21}\boldsymbol{D}_1=\boldsymbol{0},\text{即 }\boldsymbol{F}_{11}=\boldsymbol{F}_{21}=\boldsymbol{0}.$$

因而

$$\boldsymbol{F}=\begin{pmatrix}\boldsymbol{0} & \boldsymbol{F}_{12}\\ \boldsymbol{0} & \boldsymbol{F}_{22}\end{pmatrix}=(\boldsymbol{0}\quad \boldsymbol{F}_2),\qquad \boldsymbol{F}_2=\begin{pmatrix}\boldsymbol{F}_{12}\\ \boldsymbol{F}_{22}\end{pmatrix}.$$

$\boldsymbol{H}'$与 $\boldsymbol{H}$ 同为正交矩阵. 今作正交线性变换

$$\boldsymbol{z}=\boldsymbol{H}'\boldsymbol{y}=\begin{pmatrix}\boldsymbol{z}_{(1)}\\ \boldsymbol{z}_{(2)}\end{pmatrix}. \tag{9.2.22}$$

专业课辅导+真题资料+微信公众号：432应统书院

于是 $z_1, z_2, \cdots, z_n$ 相互独立且皆服从正态分布. 由于

$$\boldsymbol{y}=\boldsymbol{H}\boldsymbol{z}, \quad \boldsymbol{y}'=\boldsymbol{z}'\boldsymbol{H}',$$

于是

$$\left.\begin{aligned} Q_e &= \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}=\boldsymbol{z}'\boldsymbol{H}'\boldsymbol{A}\boldsymbol{H}\boldsymbol{z}=\boldsymbol{z}'\boldsymbol{D}\boldsymbol{z}=\boldsymbol{z}'_{(1)}\boldsymbol{D}_1\boldsymbol{z}_{(1)}, \\ \hat{\boldsymbol{\beta}} &= \boldsymbol{B}\boldsymbol{y}=\boldsymbol{B}\boldsymbol{H}\boldsymbol{z}=\boldsymbol{F}\boldsymbol{z}=\boldsymbol{F}_2\boldsymbol{z}_{(2)}, \end{aligned}\right\} \tag{9.2.23}$$

其中 $\boldsymbol{z}_{(1)}$ 是 $n-k$ 维正态随机向量，$\boldsymbol{z}_{(2)}$ 是 k 维正态随机向量，就是说，对于 $z_1, z_2, \cdots, z_n$，Q_e 是前 $n-k$ 个 z_i 的二次型，$\hat{\boldsymbol{\beta}}$ 是后 k 个 z_i 的线性型，因此 Q_e 与 $\hat{\boldsymbol{\beta}}$ 相互独立.

(2) $\hat{\boldsymbol{\beta}}$ 为 $\boldsymbol{y}$ 的线性型，$\boldsymbol{y}$ 为 n 维正态向量，则 $\hat{\boldsymbol{\beta}}$ 为 k 维正态向量. 但要注意，由于 $\operatorname{cov}(\hat{\beta}_i, \hat{\beta}_j)=\sigma^2 c_{ij}$，所以 $\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k$ 之间不一定相互独立.

由(9.2.12)式知

$$\boldsymbol{e}=\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}=\boldsymbol{y}-\boldsymbol{X}\boldsymbol{L}^{-1}\boldsymbol{X}'\boldsymbol{y}=(\boldsymbol{I}_{n\times n}-\boldsymbol{X}\boldsymbol{L}^{-1}\boldsymbol{X}')\boldsymbol{y},$$

所以 $\boldsymbol{e}$ 是 $\boldsymbol{y}$ 的线性型，是 n 维正态向量. 而且

$$\begin{aligned} \operatorname{cov}(\boldsymbol{e}, \boldsymbol{e}) &= \operatorname{cov}[(\boldsymbol{I}_{n\times n}-\boldsymbol{X}\boldsymbol{L}^{-1}\boldsymbol{X}')\boldsymbol{y}, (\boldsymbol{I}_{n\times n}-\boldsymbol{X}\boldsymbol{L}^{-1}\boldsymbol{X}')\boldsymbol{y}] \\ &= \sigma^2(\boldsymbol{I}_{n\times n}-\boldsymbol{X}\boldsymbol{L}^{-1}\boldsymbol{X}'). \end{aligned}$$

(3) 记 $\boldsymbol{A}=\boldsymbol{I}_{n\times n}-\boldsymbol{X}\boldsymbol{L}^{-1}\boldsymbol{X}'$，则有

$$\begin{aligned} &(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{A}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta}) \\ &=\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}-\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{A}\boldsymbol{y}-\boldsymbol{y}'\boldsymbol{A}\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\boldsymbol{\beta}=\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}, \end{aligned}$$

其中

$$\begin{aligned} &\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{A}=\boldsymbol{\beta}'\boldsymbol{X}'-\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{L}^{-1}\boldsymbol{X}'=\boldsymbol{\beta}'\boldsymbol{X}'-\boldsymbol{\beta}'\boldsymbol{L}\boldsymbol{L}^{-1}\boldsymbol{X}'=\boldsymbol{0}, \\ &\boldsymbol{A}\boldsymbol{X}\boldsymbol{\beta}=\boldsymbol{X}\boldsymbol{\beta}-\boldsymbol{X}\boldsymbol{L}^{-1}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}=\boldsymbol{X}\boldsymbol{\beta}-\boldsymbol{X}\boldsymbol{L}^{-1}\boldsymbol{L}\boldsymbol{\beta}=\boldsymbol{0}, \\ &\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\boldsymbol{\beta}=\boldsymbol{0}. \end{aligned}$$

于是

$$\frac{Q_e}{\sigma^2}=\frac{\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}}{\sigma^2}=\frac{1}{\sigma^2}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{A}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})=\frac{\boldsymbol{\varepsilon}'\boldsymbol{A}\boldsymbol{\varepsilon}}{\sigma^2}.$$

我们知道，矩阵 $\boldsymbol{A}$ 的秩为 $n-k$，可选正交矩阵 $\boldsymbol{H}$，使得 $\boldsymbol{D}=\boldsymbol{H}'\boldsymbol{A}\boldsymbol{H}$ 中的 $\boldsymbol{D}_1$ 为 $(n-k)\times(n-k)$ 阶矩阵. 又由 $\boldsymbol{H}$ 为正交矩阵知

专业课辅导+真题资料+微信公众号：432应统书院

$$H'H=HH'=I_{n\times n},$$

于是

$$\begin{aligned}\frac{Q_e}{\sigma^2}&=\frac{\varepsilon'A\varepsilon}{\sigma^2}=\left(\frac{\varepsilon'}{\sigma}\right)H'HAH'H\left(\frac{\varepsilon}{\sigma}\right)\\&=U'HAH'U=U'DU,\end{aligned}$$

其中

$$U=H\left(\frac{\varepsilon}{\sigma}\right),U'=\left(\frac{\varepsilon'}{\sigma}\right)H'.$$

$\left(\frac{\varepsilon}{\sigma}\right)$是每个分量服从正态 $N(0,1)$ 的随机向量. 因而 $U'=(u_1,u_2,\cdots,u_n)$中每个向量也服从正态 $N(0,1)$，而且 $u_1,u_2,\cdots,u_n$ 相互独立. 所以

$$\begin{aligned}\frac{Q_e}{\sigma^2}&=U'DU\\&=\sum_{i=1}^{n-k}u_i^2\sim\chi^2_{(n-k)}.\end{aligned}$$

最后顺便指出，最小二乘法估计量$\hat{\boldsymbol{\beta}}$及$\hat{\sigma}^2=\frac{Q_e}{n}$，也是$\boldsymbol{\beta}$及 σ^2 的极大似然法估计量，因而$\hat{\boldsymbol{\beta}}$及$\hat{\sigma}^2$ 还有极大似然法估计量所具有的性质. 我们知道，$y_1,y_2,\cdots,y_n$ 是相互独立的随机变量，y_t 服从正态 $N\left(\sum_{i=1}^{k}\beta_i x_{ti},\sigma\right)$，$t=1,2,\cdots,n$. 于是 $\beta_1,\beta_2,\cdots,\beta_k$ 及 σ^2 的似然函数为：

$$L(\boldsymbol{\beta},\sigma^2)=\frac{1}{(2\pi\sigma^2)^{n/2}}\exp\left[-\frac{1}{2\sigma^2}\sum_{t=1}^{n}\left(y_t-\sum_{i=1}^{k}\beta_i x_{ti}\right)^2\right],$$

可见，使得 $L(\boldsymbol{\beta},\sigma^2)$ 达到极大值等价于使得(9.2.3)式达到极小值.

§9.3 例　　题

一、讨论三个例题

例 9.3.1　我们考虑例 9.1.1 的表达式

专业课辅导+真题资料+微信公众号：432应统书院

$$y=\beta_0+\beta_1 x+\varepsilon, \tag{9.3.1}$$

$$\begin{cases} y_1=\beta_0+\beta_1 x_1+\varepsilon_1; \\ \vdots \\ y_n=\beta_0+\beta_1 x_n+\varepsilon_n. \end{cases} \tag{9.3.2}$$

记
$$\boldsymbol{y}=\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \boldsymbol{X}=\begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}, \boldsymbol{\varepsilon}=\begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix},$$

则有

$$\boldsymbol{L}=\boldsymbol{X}'\boldsymbol{X}=\begin{pmatrix} n & n\bar{x} \\ n\bar{x} & \sum_{t=1}^{n} x_t^2 \end{pmatrix},$$

$$\boldsymbol{L}^{-1}=\frac{1}{n\sum_{t=1}^{n}(x_t-\bar{x})^2}\begin{pmatrix} \sum_{t=1}^{n} x_t^2 & -n\bar{x} \\ -n\bar{x} & n \end{pmatrix}.$$

其中$\bar{x}=\frac{1}{n}\sum_{t=1}^{n} x_t$. 又记$\bar{y}=\frac{1}{n}\sum_{t=1}^{n} y_t$，因而

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= \boldsymbol{L}^{-1}\boldsymbol{X}'\boldsymbol{y} \\ &= \frac{1}{n\sum_{t=1}^{n}(x_t-\bar{x})^2}\begin{pmatrix} n\bar{y}\sum_{t=1}^{n} x_t^2 - n\bar{x}\sum_{t=1}^{n} x_t y_t \\ n\sum_{t=1}^{n} x_t y_t - n^2\bar{x}\bar{y} \end{pmatrix}, \end{aligned} \tag{9.3.3}$$

由此得到

$$\hat{\beta}_1=\frac{n\sum_{t=1}^{n} x_t y_t - n^2\bar{x}\bar{y}}{n\sum_{t=1}^{n}(x_t-\bar{x})^2}=\frac{\sum_{t=1}^{n} x_t y_t - n\bar{x}\bar{y}}{\sum_{t=1}^{n} x_t^2 - n\bar{x}^2}. \tag{9.3.4}$$

由(9.3.2)可知

$$\hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{x}, \tag{9.3.5}$$

专业课辅导+真题资料+微信公众号：432应统书院

$$
\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{n}(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}) \\
&= \frac{1}{n}\begin{pmatrix} y_1-(\hat{\beta}_0+\hat{\beta}_1 x_1) \\ \vdots \\ y_n-(\hat{\beta}_0+\hat{\beta}_1 x_n) \end{pmatrix}' \begin{pmatrix} y_1-(\hat{\beta}_0+\hat{\beta}_1 x_1) \\ \vdots \\ y_n-(\hat{\beta}_0+\hat{\beta}_1 x_n) \end{pmatrix} \\
&= \frac{1}{n}\begin{pmatrix} (y_1-\bar{y})-\hat{\beta}_1(x_1-\bar{x}) \\ \vdots \\ (y_n-\bar{y})-\hat{\beta}_1(x_n-\bar{x}) \end{pmatrix}' \begin{pmatrix} (y_1-\bar{y})-\hat{\beta}_1(x_1-\bar{x}) \\ \vdots \\ (y_n-\bar{y})-\hat{\beta}_1(x_n-\bar{x}) \end{pmatrix} \\
&= \frac{1}{n}\Big[\sum_{t=1}^{n}(y_t-\bar{y})^2+\hat{\beta}_1^2\sum_{t=1}^{n}(x_t-\bar{x})^2 \\
&\quad -2\hat{\beta}_1\sum_{t=1}^{n}(x_t-\bar{x})(y_t-\bar{y})\Big] \\
&= \frac{1}{n}\Big[\sum_{t=1}^{n}(y_t-\bar{y})^2-\hat{\beta}_1^2\sum_{t=1}^{n}(x_t-\bar{x})^2\Big] \\
&= s_2^2-\hat{\beta}_1^2 s_1^2,
\end{aligned}
\tag{9.3.6}
$$

其中

$$s_1^2=\frac{1}{n}\sum_{t=1}^{n}(x_t-\bar{x})^2,\ s_2^2=\frac{1}{n}\sum_{t=1}^{n}(y_t-\bar{y})^2.$$

则

$$\hat{y}=\hat{\beta}_0+\hat{\beta}_1 x$$

为 y 关于 x 的线性回归方程.

对于正态线性模型,则有

(1) $\hat{\beta}_0$ 服从正态 $N(\beta_0,\sigma_0)$,$\hat{\beta}_1$ 服从正态 $N(\beta_1,\sigma_1)$,其中

$$\sigma_0^2 = D(\hat{\beta}_0) = \sigma^2\sum_{t=1}^{n}x_t^2\Big/n\sum_{t=1}^{n}(x_t-\bar{x})^2 = \sigma^2\left(\frac{1}{n}\sum_{t=1}^{n}x_t^2\right)\Big/ns_1^2,$$

$$\sigma_1^2=D(\hat{\beta}_1)=\sigma^2\Big/\sum_{t=1}^{n}(x_t-\bar{x})^2=\sigma^2/ns_1^2,$$

专业课辅导+真题资料+微信公众号：432应统书院

$$\operatorname{cov}(\hat{\beta}_0,\hat{\beta}_1)=-\sigma^2\bar{x}\Big/\sum_{t=1}^{n}(x_t-\bar{x})^2=-\sigma^2\bar{x}/ns_1^2;$$

(2) $\dfrac{n\hat{\sigma}^2}{\sigma^2}\sim\chi^2_{(n-2)}$, $E\left(\dfrac{n\hat{\sigma}^2}{n-2}\right)=\sigma^2$；记$\dfrac{n\hat{\sigma}^2}{n-2}=\hat{\sigma}_e^2$.

(3) $\hat{\beta}_0$ 与$\hat{\sigma}^2$ 相互独立，$\hat{\beta}_1$ 与$\hat{\sigma}^2$ 相互独立，于是

$$\frac{(\hat{\beta}_0-\beta_0)/\sigma_0}{\sqrt{\dfrac{n\hat{\sigma}^2}{\sigma^2}\Big/(n-2)}}=\sqrt{n-2}\,\frac{\hat{\beta}_0-\beta_0}{\hat{\sigma}\sqrt{\dfrac{1}{n}\sum_{t=1}^{n}x_t^2}\Big/s_1}\sim t_{(n-2)},\tag{9.3.7}$$

$$\frac{(\hat{\beta}_1-\beta_1)/\sigma_1}{\sqrt{\dfrac{n\hat{\sigma}^2}{\sigma^2}\Big/(n-2)}}=\sqrt{n-2}\,\frac{\hat{\beta}_1-\beta_1}{\hat{\sigma}/s_1}\sim t_{(n-2)}.\tag{9.3.8}$$

上式等价于

$$(n-2)\frac{(\hat{\beta}_1-\beta_1)^2}{\hat{\sigma}^2/s_1^2}=\frac{ns_1^2(\hat{\beta}_1-\beta_1)^2}{n\hat{\sigma}^2/(n-2)}=\frac{[(\hat{\beta}_1-\beta_1)/\sigma_1^2]^2}{\dfrac{n\hat{\sigma}^2}{\sigma^2}\Big/(n-2)}$$

$$\sim F_{(1,n-2)}.\tag{9.3.9}$$

例 9.3.2 现考虑如下线性模型

$$y=\beta_0+\beta_1(x-\bar{x})+\varepsilon,\tag{9.3.10}$$

$$\begin{cases}y_1=\beta_0+\beta_1(x_1-\bar{x})+\varepsilon_1,\\ \vdots\\ y_n=\beta_0+\beta_1(x_n-\bar{x})+\varepsilon_n,\end{cases}$$

其中

$$\bar{x}=\frac{1}{n}\sum_{t=1}^{n}x_t.$$

这个模型的特点是对数据 $x_1,x_2,\cdots,x_n$ 作了中心化的处理. 先获得原来 n 组数据(y_t,x_t)，$t=1,2,\cdots,n$. 求出$\bar{x}=\dfrac{1}{n}\sum\limits_{t=1}^{n}x_t$，然后利

专业课辅导+真题资料+微信公众号：432应统书院

用新的 n 组数据 $(y_t, x_t-\bar{x})$，$t=1,2,\cdots,n$，建立线性回归方程

$$\hat{y}=\hat{\beta}_0+\hat{\beta}_1(x-\bar{x}),$$

$$\boldsymbol{y}=\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \boldsymbol{X}=\begin{pmatrix} 1 & x_1-\bar{x} \\ \vdots & \vdots \\ 1 & x_n-\bar{x} \end{pmatrix},$$

$$\boldsymbol{L}=\boldsymbol{X}'\boldsymbol{X}=\begin{pmatrix} n & 0 \\ 0 & \sum_{t=1}^{n}(x_t-\bar{x})^2 \end{pmatrix},$$

$$\boldsymbol{L}^{-1}=\frac{1}{n\sum_{t=1}^{n}(x_t-\bar{x})^2}\begin{pmatrix} \sum_{t=1}^{n}(x_t-\bar{x})^2 & 0 \\ 0 & n \end{pmatrix}.$$

记

$$s_1^2=\frac{1}{n}\sum_{t=1}^{n}(x_t-\bar{x})^2, s_2^2=\frac{1}{n}\sum_{t=1}^{n}(y_t-\bar{y})^2,$$

因而

$$\hat{\boldsymbol{\beta}}=\boldsymbol{L}^{-1}\boldsymbol{X}'\boldsymbol{y}=\frac{1}{n\sum_{t=1}^{n}(x_t-\bar{x})^2}\begin{pmatrix} n\bar{y}\sum_{t=1}^{n}(x_t-\bar{x})^2 \\ n\sum_{t=1}^{n}(x_t-\bar{x})y_t \end{pmatrix}$$

$$=\frac{1}{ns_1^2}\begin{pmatrix} ns_1^2\bar{y} \\ ns_{12} \end{pmatrix}, \tag{9.3.11}$$

其中

$$\sum_{t=1}^{n}(x_t-\bar{x})y_t=\sum_{t=1}^{n}(x_t-\bar{x})(y_t-\bar{y})=ns_{12}.$$

由此得到

专业课辅导+真题资料+微信公众号：432应统书院

$$\begin{cases}\hat{\beta}_0=\bar{y};\\ \hat{\beta}_1=\dfrac{s_{12}}{s_1^2}=\dfrac{\sum\limits_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum\limits_{i=1}^{n}(x_i-\bar{x})^2},\end{cases}\tag{9.3.12}$$

$$\begin{aligned}\hat{\sigma}^2&=\frac{1}{n}(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2\\&=\frac{1}{n}\sum_{i=1}^{n}[(y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})]^2=s_2^2+\hat{\beta}_1^2s_1^2-2\hat{\beta}_1s_{12}\\&=s_2^2+\hat{\beta}_1^2s_1^2-2\hat{\beta}_1^2s_1^2=s_2^2-\hat{\beta}_1^2s_1^2.\end{aligned}\tag{9.3.13}$$

记

$$R=\frac{s_{12}}{s_1\cdot s_2},\tag{9.3.14}$$

则有

$$\hat{\beta}_1=\frac{s_{12}}{s_1^2}=R\frac{s_2}{s_1};$$

$$\hat{\sigma}^2=s_2^2-\left(R\frac{s_2}{s_1}\right)^2s_1^2=(1-R^2)s_2^2.\tag{9.3.15}$$

对于正态线性模型，则有

(1) $\hat{\beta}_0$ 服从正态 $N(\beta_0,\sigma_0)$，$\hat{\beta}_1$ 服从正态 $N(\beta_1,\sigma_1)$，其中

$$\sigma_0^2=\frac{\sigma^2}{n},\sigma_1^2=\frac{\sigma^2}{\sum\limits_{i=1}^{n}(x_i-\bar{x})^2}=\frac{\sigma^2}{ns_1^2},$$

由 $\mathrm{cov}(\hat{\beta}_0,\hat{\beta}_1)=0$ 知，$\hat{\beta}_0$ 与 $\hat{\beta}_1$ 相互独立.

(2) $\hat{\beta}_0$ 与 $\hat{\sigma}^2$ 相互独立，$\hat{\beta}_1$ 与 $\hat{\sigma}^2$ 相互独立，

$$\frac{n\hat{\sigma}^2}{\sigma^2}\sim\chi^2_{(n-2)},\sqrt{n-2}\,\frac{\hat{\beta}_0-\beta_0}{\hat{\sigma}}\sim t_{(n-2)},\tag{9.3.16}$$

$$\sqrt{n-2}\,s_1\frac{\hat{\beta}_1-\beta_1}{\hat{\sigma}}\sim t_{(n-2)},\tag{9.3.17}$$

专业课辅导+真题资料+微信公众号：432应统书院

$$\frac{ns_1^2(\hat{\beta}_1-\beta_1)^2}{n\hat{\sigma}^2/(n-2)} \sim F_{(1,n-2)}. \tag{9.3.18}$$

(3) 关于 y 的预测问题

由于 $\hat{\beta}_0$ 与 $\hat{\beta}_1$ 相互独立，则 $\hat{y}=\hat{\beta}_0+\hat{\beta}_1(x-\bar{x})$ 服从正态

$$N\left(\beta_0+\beta_1(x-\bar{x}),\sigma\sqrt{\frac{1}{n}+\frac{(x-\bar{x})^2}{ns_1^2}}\right),$$

于是 $y-\hat{y}$ 服从正态 $N\left(0,\sigma\sqrt{1+\frac{1}{n}+\frac{(x-\bar{x})^2}{ns_1^2}}\right)$，其中

$$D(\hat{y})=D(\hat{\beta}_0)+(x-\bar{x})^2D(\hat{\beta}_1)=\sigma^2\left[\frac{1}{n}+\frac{(x-\bar{x})^2}{ns_1^2}\right], \tag{9.3.19}$$

$$\begin{aligned}D(y-\hat{y}) &= D[(\beta_0-\hat{\beta}_0)+(\beta_1-\hat{\beta}_1)(x-\bar{x})+\varepsilon]\\ &=D(\beta_0-\hat{\beta}_0)+(x-\bar{x})^2D(\beta_1-\hat{\beta}_1)+D(\varepsilon)\\ &=\sigma^2\left[1+\frac{1}{n}+\frac{(x-\bar{x})^2}{ns_1^2}\right].\end{aligned} \tag{9.3.20}$$

因此

$$\frac{(y-\hat{y})/\sqrt{D(y-\hat{y})}}{\sqrt{\frac{n\hat{\sigma}^2}{\sigma^2}\Big/(n-2)}}=\frac{y-\hat{y}}{\hat{\sigma}_e\sqrt{1+\frac{1}{n}+\frac{(x-\bar{x})^2}{ns_1^2}}}\sim t_{(n-2)}, \tag{9.3.21}$$

其中 $\hat{\sigma}_e^2=\frac{n\hat{\sigma}^2}{n-2}$. 给定显著性水平 α，查得临界值 $t_{(n-2)}(\alpha)$，使得

$$P\left\{|y-\hat{y}|\leqslant\hat{\sigma}_e\sqrt{1+\frac{1}{n}+\frac{(x-\bar{x})^2}{ns_1^2}}t_{(n-2)}(\alpha)\right\}=1-\alpha, \tag{9.3.22}$$

亦即有

$$P\{\hat{y}_2\leqslant y\leqslant\hat{y}_1\}=1-\alpha, \tag{9.3.23}$$

其中

专业课辅导+真题资料+微信公众号：432应统书院

$$\hat{y}_1=\hat{\beta}_0+\hat{\beta}_1(x-\bar{x})+\hat{\sigma}_e\sqrt{1+\frac{1}{n}+\frac{(x-\bar{x})^2}{ns_1^2}}t_{(n-2)}(\alpha),$$

$$\hat{y}_2=\hat{\beta}_0+\hat{\beta}_1(x-\bar{x})-\hat{\sigma}_e\sqrt{1+\frac{1}{n}+\frac{(x-\bar{x})^2}{ns_1^2}}t_{(n-2)}(\alpha).$$

由上述两条曲线之间围成的区域，称之为 y 的置信水平为 $1-\alpha$ 的置信区域或预测区域.

例 9.3.3 炼钢是个氧化脱碳的过程. 因为钢液含碳量 x 与精炼时间 y 有一定的关系，现将某平炉共炼 34 炉钢记录下来的含碳量 x 与精炼时间 y 的数据列入表 9.3.1，试求 y 与 x 的线性回归方程. 含碳量是指全部炉料熔化完毕时钢液的含碳量，精炼时间是指炉料熔毕至出钢所需要的时间.

表 9.3.1

编　号	含碳量 x/%	精炼时间 y/min	编　号	含碳量 x/%	精炼时间 y/min
1	1.80	200	13	1.90	190
2	1.04	100	14	1.90	210
3	1.34	135	15	1.61	145
4	1.41	125	16	1.65	195
5	2.04	235	17	1.54	150
6	1.50	170	18	1.16	100
7	1.20	125	19	1.23	110
8	1.51	135	20	1.51	180
9	1.47	155	21	1.10	130
10	1.45	165	22	1.08	110
11	1.41	135	23	1.58	130
12	1.44	160	24	1.07	115

专业课辅导+真题资料+微信公众号：432应统书院

续表

编　号	含碳量 x/%	精炼时间 y/min	编　号	含碳量 x/%	精炼时间 y/min
25	1.80	240	30	1.53	145
26	1.27	135	31	1.55	160
27	1.15	120	32	1.77	185
28	1.91	205	33	1.77	205
29	1.90	220	34	1.43	160

解　我们按(9.3.10)式来确定线性回归方程. 当观察数据 $(y_t,x_t)(t=1,2,\cdots,n)$ 的数值比较大时,可通过下述变换而简化计算工作量. 令

$$x_t^*=\frac{x_t-c_1}{d_1},y_t^*=\frac{y_t-c_2}{d_2},t=1,2,\cdots,n,$$

其中 c_1,c_2 及 $d_1\neq0,d_2\neq0$ 为四个常数,则有:

$$x_t=c_1+d_1x_t^*,y_t=c_2+d_2y_t^*,t=1,2,\cdots,n,$$
$$\bar{x}=c_1+d_1\bar{x}^*,\bar{y}=c_2+d_2\bar{y}^*.$$

其中

$$\bar{x}^*=\frac{1}{n}\sum_{t=1}^{n}x_t^*,\bar{y}^*=\frac{1}{n}\sum_{t=1}^{n}y_t^*,$$

并且

$$s_1^2=d_1^2s_1^{*2},s_2^2=d_2^2s_2^{*2},$$

其中

$$s_1^{*2}=\frac{1}{n}\sum_{t=1}^{n}(x_t^*-\bar{x}^*)^2,s_2^{*2}=\frac{1}{n}\sum_{t=1}^{n}(y_t^*-\bar{y}^*)^2.$$

由于

$$\frac{1}{n}\sum_{t=1}^{n}(x_t-\bar{x})(y_t-\bar{y})=\frac{d_1d_2}{n}\sum_{t=1}^{n}(x_t^*-\bar{x}^*)(y_t^*-\bar{y}^*),$$

专业课辅导+真题资料+微信公众号：432应统书院

所以

$$R=R^*.$$

因此，由(y_t^*, x_t^*)，$t=1,2,\cdots,n$，求出$\bar{x}^*$，$\bar{y}^*$，s_1^{*2}，s_2^{*2}，R^*，从而得到$\bar{x}$，$\bar{y}$，s_1^2，s_2^2及R.

实际计算如下（见表9.3.2）.

$\sum_{t=1}^{34} x_t^* = 0.02$，$\bar{x}^* = 0.0006$，则$\bar{x} = 1.5006$；

$\sum_{t=1}^{34} y_t^* = -12$，$\bar{y}^* = -0.353$，则$\bar{y} = 158.235$；

$\sum_{t=1}^{34} (x_t^* - \bar{x}^*)^2 = \sum_{t=1}^{34} x_t^{*2} - 34(\bar{x}^*)^2 \approx 2.5522$；

$\sum_{t=1}^{34} (x_t^* - \bar{x}^*)(y_t^* - \bar{y}^*) = \sum_{t=1}^{34} x_t^* y_t^* - 34\bar{x}^*\bar{y}^* \approx 64.717$；

$\sum_{t=1}^{34} (y_t^* - \bar{y}^*)^2 = \sum_{t=1}^{34} y_t^{*2} - 34(\bar{y}^*)^2 \approx 2003.76$，

则

$$\sum_{t=1}^{34} (x_t - \bar{x})^2 \approx 2.5522, \sum_{t=1}^{34} (x_t - \bar{x})(y_t - \bar{y}) \approx 323.585.$$

表 9.3.2

编　号	x^*	y^*	x^{*2}	y^{*2}	x^*y^*
1	0.30	8	0.0900	64	2.4
2	−0.46	−12	0.2116	144	5.52
3	−0.16	−5	0.0256	25	0.8
4	−0.09	−7	0.0081	49	0.63
5	0.54	15	0.2916	225	8.1
6	0.00	2	0.0000	4	0.0
7	−0.3	−7	0.09	49	2.1
8	0.01	−5	0.0001	25	−0.05
9	−0.03	−1	0.0009	1	0.03

专业课辅导+真题资料+微信公众号：432应统书院

续表

编　号	x^*	y^*	x^{*2}	y^{*2}	x^*y^*
10	− 0.05	1	0.002 5	1	− 0.05
11	− 0.09	− 5	0.008 1	25	0.45
12	− 0.06	0	0.003 6	0	0.0
13	0.40	6	0.160 0	36	2.4
14	0.40	10	0.160 0	100	4.0
15	0.11	− 3	0.012 1	9	− 0.33
16	0.15	7	0.022 5	49	1.05
17	0.04	− 2	0.001 6	4	− 0.08
18	− 0.34	− 12	0.115 6	144	4.08
19	− 0.27	− 10	0.072 9	100	2.7
20	0.01	4	0.000 1	16	0.04
21	− 0.40	− 6	0.160 0	36	2.4
22	− 0.42	− 10	0.176 4	100	4.2
23	0.08	− 6	0.006 4	36	− 0.48
24	− 0.43	− 9	0.184 9	81	3.87
25	0.30	16	0.090 0	256	4.8
26	− 0.23	− 5	0.052 9	25	1.15
27	− 0.35	− 8	0.122 5	64	2.8
28	0.41	9	0.168 1	81	3.69
29	0.40	12	0.160 0	144	4.8
30	0.03	− 3	0.000 9	9	− 0.09
31	0.05	0	0.002 5	0	0
32	0.27	5	0.072 9	25	1.35
33	0.27	9	0.072 9	81	2.43
34	− 0.07	0	0.004 9	0	0

专业课辅导+真题资料+微信公众号：432应统书院

于是

$$\hat{\beta}_0 = \bar{y} = 158.235, \hat{\beta}_1 = \frac{323.585}{2.552\,2} \approx 127.$$

从而得到精炼时间对含碳量的回归方程为：

$$\hat{y} = 158.235 + 127(x - 1.500\,6) = -32.34 + 127x. \tag{9.3.24}$$

在本例中，取 $c_1 = 1.50, d_1 = 1, c_2 = 160, d_2 = 5$.

从(9.3.24)看到，$\hat{\beta}_1$ 值很大，不必对 $H_0: \beta_1 = 0$ 作显著性检验.

若给定显著性水平 $\alpha = 0.05$，按(9.3.23)式求出(9.3.10)式中 y 的区间估计的端点. 下面给出简化的近似计算公式(9.3.25). 我们知道，$y - \hat{y}$ 服从正态 $N\left(0, \sigma\sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{ns_1^2}}\right)$，当 n 很大时且有下述关系式

$$1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{ns_1^2} \approx 1.$$

又 σ 用它的无偏估计量 $\hat{\sigma}_e = \sqrt{\frac{1}{n-2}Q_e}$ 作近似值，那么，

$$P\{|y - \hat{y}| \leqslant 2\hat{\sigma}_e\} \approx 0.95,$$

$$P\{\hat{y} - 2\hat{\sigma}_e \leqslant y \leqslant \hat{y} + 2\hat{\sigma}_e\} \approx 0.95. \tag{9.3.25}$$

$$Q_e = \sum_{t=1}^{n} (y_t - \bar{y})^2 - U = ns_2^2 - \hat{\beta}_1^2 \cdot ns_1^2$$

$$= 50\,094 - 41\,164 = 8\,930,$$

$\hat{\sigma}_e \approx 16.71$. 于是，对于每一 x 值，y 的置信水平为 0.95 的预测区域为下述两条直线之间围成的区域：

$$\begin{cases} y = 1.08 + 127x; \\ y = -66 + 127x. \end{cases} \tag{9.3.26}$$

专业课辅导+真题资料+微信公众号：432应统书院

二、将曲线方程线性化

我们在例 9.3.1 至例 9.3.3 的三个例子中，是先求出它们的线性回归方程，然后对“线性回归方程”作显著性检验. 其实，直观而又简便的办法是用 n 组观察数据 (y_t, x_t) $(t=1,2,\cdots,n)$，在平面上标出 n 个点，根据这 n 个点所呈现出的形状，我们粗略地看看像不像一条直线. 如果很不像一条直线，即使配了线性回归方程，线性模型的假设检验一定是作出被否定的判断. 我们可以根据观察数据的 n 点所呈现出形状，与常见的已知的函数图形作比较，选择一条曲线拟合这 n 个点，它比粗糙地配一条直线更为精确. 下面给出一些常见的、可以通过变量作变换而化成线性回归方程的函数图形及其数学表达式. 在化成线性回归方程之后，就可按最小二乘法估计其参数，从而给出原曲线方程中参数的估计.

1. 双曲线方程：

$$\frac{1}{y}=a+\frac{b}{x}\text{（如图 9.3.1）}.$$

图 9.3.1

2. 幂函数方程：

$$y=ax^b\text{（如图 9.3.2）}.$$

3. 指数曲线方程：

$$y=ae^{bx}\text{（如图 9.3.3）}.$$

4. 指数曲线方程：

专业课辅导+真题资料+微信公众号：432应统书院

图 9.3.2

图 9.3.3

$y=ae^{\frac{b}{x}}$（如图 9.3.4）.

图 9.3.4

专业课辅导+真题资料+微信公众号：432应统书院

5. 对数曲线方程：

$$y=a+b\ln x\text{（如图 9.3.5）}.$$

图 9.3.5

6. S 形曲线方程：

$$y=\frac{1}{a+be^{-x}}\text{（如图 9.3.6）}.$$

图 9.3.6

7. 抛物线方程：

$$y=b_0+b_1x+b_2x^2.$$

以上各曲线方程怎样通过变量作变换化成为线性方程，留作

专业课辅导+真题资料+微信公众号：432应统书院

练习，在此从略. 上述各曲线的图形，选自亨德一书①.

下面举两个例子.

例 9.3.4 在彩色显像中，根据以往的经验，形成染料光学密度 y 与析出银的光学密度 x 之间有下面类型的关系式：

$$y=ae^{-\frac{b}{x}},b>0. \tag{9.3.27}$$

现对 y 及 x 同时作 11 次观察（或试验），获得 11 组数据 (y_t,x_t) $(t=1,2,\cdots,11)$，求出 a 及 b 的估计值. 将这 11 组数据 (y_t,x_t) 列在表 9.3.3 中，并将这 11 个点标在图 9.3.7 上.

表 9.3.3

编号	x	y	$x'=\frac{1}{x}$	$y'=\ln y$	x'^2	$x'y'$
1	0.05	0.10	20.000	−2.303	400.00	−46.06
2	0.06	0.14	16.667	−1.966	277.78	−32.77
3	0.07	0.23	14.286	−1.470	204.08	−21.00
4	0.10	0.37	10.000	−0.994	100.00	−9.94
5	0.14	0.59	7.143	−0.528	51.02	−3.77
6	0.20	0.79	5.000	−0.236	25.00	−1.18
7	0.25	1.00	4.000	0.000	16.00	0
8	0.31	1.12	3.226	0.113	10.41	0.36
9	0.38	1.19	2.632	0.174	6.93	0.46
10	0.43	1.25	2.326	0.223	5.41	0.52
11	0.47	1.29	2.128	0.255	4.53	0.54
$\sum_{t=1}^{11}$	2.46	8.07	87.408	−6.732	1 101.16	−112.84

记

$$y'=\ln y,x'=\frac{1}{x},$$

① A. Hald. Statistical Theory With Engineering Applications. John Wiley & Sons, Inc. (1952).

专业课辅导+真题资料+微信公众号：432应统书院

$$\beta_0=\ln a,\beta_1=-b,$$

则(9.3.27)式变换为：

$$y'=\beta_0+\beta_1 x'. \qquad (9.3.28)$$

我们应用例9.3.1中$\hat{\beta}_0$及$\hat{\beta}_1$的计算公式，利用表9.3.3中的y'及x'的观察数据，在表9.3.3中计算β_0及β_1的估计值，可得：

$$\hat{\beta}_0=\bar{y}'-\hat{\beta}_1\bar{x}',$$

$$\hat{\beta}_1=\frac{\sum_{t=1}^{n}x'_t y'_t-n\bar{x}'\bar{y}'}{\sum_{t=1}^{n}x'^2_t-n\bar{x}'^2}$$

图 9.3.7

实际计算结果如下：

$$\bar{x}'=\frac{1}{11}\sum_{t=1}^{11}x'_t=\frac{87.408}{11}=7.95,$$

$$\bar{y}'=\frac{1}{11}\sum_{t=1}^{11}y'_t=\frac{-6.732}{11}=-0.612,$$

$$\sum_{t=1}^{11}x'_t y'_t-11\bar{x}'\bar{y}'=-112.84+53.49=-59.35,$$

$$\sum_{t=1}^{11}x'^2-11\bar{x}'^2=1\ 101.16-694.56=406.60,$$

$$\hat{\beta}_1=\frac{-59.35}{406.60}\approx-0.146,$$

$$\beta_0=-0.612-(-0.146)\times7.95\approx0.549.$$

因而线性回归方程为

$$\hat{y}'=0.549-0.146x'.$$

由于

$$\hat{a}=e^{\hat{\beta}_0}\approx1.73,$$

$$\hat{b}=-\hat{\beta}_1\approx0.146,$$

专业课辅导+真题资料+微信公众号：432应统书院

所以(9.3.27)式的回归曲线方程为

$$\hat{y}=1.73e^{-\frac{0.146}{x}}. \tag{9.3.29}$$

例 9.3.5 已知某种半成品在生产过程中的废品率 y 与它的某种化学成分 x 有关，现将观察(或试验)得到的一批数据，记录在表 9.3.4 中，并将这 16 对数据 (x_t, y_t) 在平面上标出 16 个点，见图 9.3.8.

表 9.3.4

编号	y	$x_1=x$	$x_2=x^2$	编号	y	$x_1=x$	$x_2=x^2$
1	1.30	34	1 156	9	0.44	40	1 600
2	1.00	36	1 296	10	0.56	41	1 681
3	0.73	37	1 369	11	0.30	42	1 764
4	0.90	38	1 444	12	0.42	43	1 849
5	0.81	39	1 521	13	0.35	43	1 849
6	0.70	39	1 521	14	0.40	45	2 025
7	0.60	39	1 521	15	0.41	47	2 209
8	0.50	40	1 600	16	0.60	48	2 304

y 为废品率(%)，x 为化学成分含量(0.01%).

从表 9.3.4 及图 9.3.8 中都可看到，废品率 y 最初随着化学成分 x 的增加而降低，而当化学成分 x 超过一定值之后，废品率 y 又有所回升. 根据这个规律或散点分布图的形状，可以考虑拟合抛物线

$$\hat{y}=\beta_0+\beta_1 x+\beta_2 x^2. \tag{9.3.30}$$

记 $x_1=x,\quad x_2=x^2,$

图 9.3.8

专业课辅导+真题资料+微信公众号：432应统书院

则(9.3.30)式变换为：

$$\hat{y}=\beta_0+\beta_1 x_1+\beta_2 x_2. \tag{9.3.31}$$

下面我们利用表 9.3.4 中的 16 组数据 $(y_t;x_{1t},x_{2t})$，$t=1,2,\cdots,16$，像表 9.3.3 那样列表计算.

记

$$\boldsymbol{X}=\begin{pmatrix}1 & x_{11} & x_{12}\\ \vdots & \vdots & \vdots\\ 1 & x_{16,1} & x_{16,2}\end{pmatrix}_{16\times 3},$$

$$\boldsymbol{y}=\begin{pmatrix}y_1\\ \vdots\\ y_{16}\end{pmatrix}_{16\times 1},\qquad \boldsymbol{\beta}=\begin{pmatrix}\beta_0\\ \beta_1\\ \beta_2\end{pmatrix}_{3\times 1}.$$

由(9.2.7)及(9.2.8)两式知道，

$$(\boldsymbol{X}'\boldsymbol{X})\hat{\boldsymbol{\beta}}=\boldsymbol{X}'\boldsymbol{y},\qquad \hat{\boldsymbol{\beta}}=(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y},$$

$$\boldsymbol{L}=\boldsymbol{X}'\boldsymbol{X}=\begin{pmatrix}16 & \sum_{t=1}^{16}x_{t1} & \sum_{t=1}^{16}x_{t2}\\ \sum_{t=1}^{16}x_{t1} & \sum_{t=1}^{16}x_{t1}^2 & \sum_{t=1}^{16}x_{t1}x_{t2}\\ \sum_{t=1}^{16}x_{t2} & \sum_{t=1}^{16}x_{t2}x_{t1} & \sum_{t=1}^{16}x_{t2}^2\end{pmatrix},$$

$$\boldsymbol{X}'\boldsymbol{y}=\begin{pmatrix}\sum_{t=1}^{16}y_t\\ \sum_{t=1}^{16}x_{t1}y_t\\ \sum_{t=1}^{16}x_{t2}y_t\end{pmatrix}.$$

专业课辅导+真题资料+微信公众号：432应统书院

记

$$\bar{y}=\frac{1}{16}\sum_{t=1}^{16}y_t,\ \bar{x}_1=\frac{1}{16}\sum_{t=1}^{16}x_{t1},\ \bar{x}_2=\frac{1}{16}\sum_{t=1}^{16}x_{t2},$$

则有

$$\begin{cases}16\hat{\beta}_0+\hat{\beta}_1\cdot 16\bar{x}_1+\hat{\beta}_2\cdot 16\bar{x}_2=16\bar{y};\\ 16\bar{x}_1\hat{\beta}_0+\hat{\beta}_1\sum_{t=1}^{16}x_{t1}^2+\hat{\beta}_2\sum_{t=1}^{16}x_{t1}x_{t2}=\sum_{t=1}^{16}x_{t1}y_t;\\ 16\bar{x}_2\hat{\beta}_0+\hat{\beta}_1\sum_{t=1}^{16}x_{t2}x_{t1}+\hat{\beta}_2\sum_{t=1}^{16}x_{t2}^2=\sum_{t=1}^{16}x_{t2}y_t,\end{cases}$$

其中

$$\hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{x}_1-\hat{\beta}_2\bar{x}_2.$$

$$\begin{cases}\hat{\beta}_1 l_{11}+\hat{\beta}_2 l_{12}=l_{1y};\\ \hat{\beta}_1 l_{21}+\hat{\beta}_2 l_{22}=l_{2y}.\end{cases}$$

$$l_{iy}=\sum_{t=1}^{16}x_{ti}y_t-16\bar{x}_i\cdot\bar{y}\qquad(i=1,2),$$

$$l_{ij}=\sum_{t=1}^{16}x_{ti}x_{tj}-16\bar{x}_i\cdot\bar{x}_j\quad(i,j=1,2).$$

实际计算结果如下：

$$\bar{y}=0.626\ 3,\quad \bar{x}_1=40.687\ 5,$$

$$\bar{x}_2=1\ 669.312\ 5,$$

$$l_{11}=221.44,l_{22}=1\ 513\ 685,$$

$$l_{12}=l_{21}=18\ 283,$$

$$l_{1y}=-11.649,l_{2y}=-923.05,$$

$$\begin{cases}221.44\hat{\beta}_1+18\ 283\hat{\beta}_2=-11.649;\\ 18\ 283\hat{\beta}_1+1\ 513\ 685\hat{\beta}_2=-923.05.\end{cases}$$

由上式解得

$$\hat{\beta}_1=-0.820\ 5,\hat{\beta}_2=0.009\ 301,$$

专业课辅导+真题资料+微信公众号：432应统书院

$$\hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{x}_1-\hat{\beta}_2\bar{x}_2=18.484.$$

从而废品率 y 与化学成分 x 的关系式是：

$$\hat{y}=18.484-0.820\,5x+0.009\,301x^2, \tag{9.3.32}$$

它在

$$x_0=\frac{0.820\,5}{2\times0.009\,301}=44.11$$

时达到极小值，其极小值为

$$\begin{aligned}y_0&=18.484-0.820\,5\times44.11+0.009\,301\times(44.11)^2\\&=0.39.\end{aligned}$$

*§9.4 假设检验

一、线性模型的假设检验

在实际工作中，事先我们并不能断定随机变量 y 与一组变量 $x_1,x_2,\cdots,x_k$ 之间确有线性关系. 在求线性回归方程之前，线性模型(9.1.7)式只是一种假设，尽管这种假设常常不是没有根据的，但在求线性回归方程后，还是需要对这种线性回归方程同实际观察或试验数据拟合的效果进行检验. 可提出原假设

$$H_0:\beta_1=\beta_2=\cdots=\beta_k=0. \tag{9.4.1}$$

如果 H_0 不否定，则表明用线性回归模型(9.1.7)不合适或拟合效果甚差. 现在怎样建立统计量来对 H_0 作显著性检验呢？对于 $i=1,2,\cdots,k$ 中每一个 i，可以对 $H_0:\beta_i=0$ 建立 F 检验统计量作显著性检验.

对于线性模型(9.1.7)式，我们也可导出离差平方和的分解式，即有：

$$\begin{aligned}\boldsymbol{y}'\boldsymbol{y}&=(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}+\boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}+\boldsymbol{X}\hat{\boldsymbol{\beta}})\\&=(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})+\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}}\\&=Q_e+U,\end{aligned} \tag{9.4.2}$$

其中 Q_e 为误差平方和，U 为回归平方和. 现作验证：

专业课辅导+真题资料+微信公众号：432应统书院

$$
\begin{aligned}
U &= \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} = [(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}]'\boldsymbol{X}'\boldsymbol{X}[(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}] \\
&= [\boldsymbol{y}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'][\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}] \\
&= \boldsymbol{y}'[\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}']\boldsymbol{y}, \qquad (9.4.3)
\end{aligned}
$$

$$
\begin{aligned}
Q_e &= (\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}) \\
&= \boldsymbol{y}'[\boldsymbol{I}-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}']\boldsymbol{y}. \qquad (9.4.4)
\end{aligned}
$$

由此可见，将(9.4.3)与(9.4.4)两式相加，即可验证(9.4.2)式成立.

我们知道，y_t 服从正态 $N\left(\sum_{i=1}^{k}\beta_i x_{ti},\sigma\right)(t=1,2,\cdots,n)$，$y_1$，$y_2,\cdots,y_n$ 相互独立，当 H_0 成立时，那么 $y_1,y_2,\cdots,y_n$ 相互独立而且有相同分布 $N(0,\sigma)$，因而

$$
\frac{1}{\sigma^2}\boldsymbol{y}'\boldsymbol{y}=\frac{1}{\sigma^2}\sum_{t=1}^{n}y_t^2\sim\chi_n^2. \qquad (9.4.5)
$$

由§9.2的性质6及定理6.3.3知

$$
\frac{1}{\sigma^2}Q_e=\frac{n\hat{\sigma}^2}{\sigma^2}\sim\chi^2_{(n-k)}. \qquad (9.4.6)
$$

$$
\frac{1}{\sigma^2}U\sim\chi_k^2. \qquad (9.4.7)
$$

Q_e 与 U 相互独立，因而

$$
\frac{U/k}{Q_e/(n-k)}=\frac{(n-k)U}{kQ_e}\sim F_{(k,n-k)}. \qquad (9.4.8)
$$

就是说，(9.4.8)式可作为对(9.4.1)式的 H_0 的检验统计量.

二、回归系数的假设检验

在线性回归模型(9.1.7)式中，并不意味着每个自变量 $x_1,x_2,\cdots,x_k$ 对随机变量 y 的影响作用都是一样重要，可能有的起重要作用，而有的是可有可无的变量. 就是说，有的变量的系数 β 很大，而有的变量的系数 β 很小或为零. 因此，不否定线性模型之后，仍有必要从线性回归模型中，剔除那些次要的、可有可无的自变量，只保留那些起重要作用的自变量，重新建立更为简练的线性回

专业课辅导+真题资料+微信公众号：432应统书院

归方程，以更有利于实际工作. 我们把自变量 x_i 叫作回归因子，其系数 β 称为 y 关于回归因子 x_i 的回归系数. 我们说某个因子 x_i 对 y 的作用不显著，是指原假设

$$H_0:\beta_i=0 \tag{9.4.9}$$

不否定. 因此，对(9.4.9)式作显著性检验，即为挑选重要因子、剔除不显著因子的工作.

我们知道，$\hat{\beta}_i$ 服从正态 $N(\beta_i,\sigma\sqrt{c_{ii}})$，而且 $\hat{\boldsymbol{\beta}}$ 与 Q_e 相互独立，因此当 H_0 成立时，则有

$$\frac{\hat{\beta}_i^2/c_{ii}}{Q_e/(n-k)}=\frac{U_i/1}{Q_e/(n-k)}\sim F_{(1,n-k)}. \tag{9.4.10}$$

于是(9.4.10)式可作为对 H_0 作检验的检验统计量. 记 $U_i=\dfrac{\hat{\beta}_i^2}{c_{ii}}$，$U_i$ 是 y 关于因子 x_i 的偏回归平方和. (9.4.10)式的 F 检验同(8.2.2)式的 t 检验是等价的.

*§9.5 单因子方差分析

方差分析法是对试验(或观察)结果的数据作分析的一种常用的统计方法. 我们在第八章的§8.2中讲到的两个总体的数学期望是否相等的显著性检验，可以称之为单因子二水平的试验，在那里建立了 t 检验法. 在这一节中，我们要讨论单因子多水平的试验，将看到它在实际上是多个总体的数学期望是否相等的显著性检验. 对于这种假设检验问题，在正态总体及方差相同的基本假定下，我们将建立 F 检验法. 这种 F 检验法，是由因子的方差同重复试验的误差项方差之比，建立 F 分布的检验统计量，因而又称作方差分析法.

例 9.5.1 水稻品种比较试验.

设水稻品种有珍珠矮、窄叶青、广进矮、朝阳早、南二矮等五种，现要考察水稻品种对水稻亩产量的影响作用，从中挑选水稻良

专业课辅导+真题资料+微信公众号：432应统书院

种. 为此，可设水稻亩产量为指标，水稻品种为因子，五个品种为该因子所取的五个水平，这是单因子五水平的试验.

例 9.5.2 灯丝的配料方案优选.

某灯泡厂用四种不同配料方案制成的灯丝，生产了四批灯泡. 在每批灯泡中随机地抽取若干个灯泡测其使用寿命（单位：h），数据列于表 9.5.1 中.

表 9.5.1

灯丝 \ 使用寿命 \ 灯泡	1	2	3	4	5	6	7	8
甲	1 600	1 610	1 650	1 680	1 700	1 700	1 780	
乙	1 500	1 640	1 400	1 700	1 750			
丙	1 640	1 550	1 600	1 620	1 640	1 600	1 740	1 800
丁	1 510	1 520	1 530	1 570	1 640	1 680		

试问这四种灯丝生产的灯泡的使用寿命有无显著性的差异？在这里，可设灯泡的使用寿命时数为指标，灯泡的配料为因子，四种配料方案为四水平，这是单因子四水平的试验.

我们将以每一种配料方案制成灯丝后生产的灯泡视为一总体，灯泡使用寿命用 ξ 表示，则 ξ 为一随机变量. 现用 ξ_1,ξ_2,ξ_3 及 ξ_4 分别表示这四批灯泡的使用寿命，即有四个总体：总体 ξ_1，总体 ξ_2，总体 ξ_3 及总体 ξ_4. 若从这四个总体中，分别随机抽取容量为 n_i 的样本 $\xi_{i1},\xi_{i2},\cdots,\xi_{in_i}(i=1,2,3,4)$. 我们将从这四个样本 ξ_{ij} 来推断这四个总体之间有无显著性差异. 现假定总体 ξ_i 服从正态 $N(a_i,\sigma)$，其中 $a_i=E(\xi_i)(i=1,2,\cdots,4)$. 它们有相同的方差 σ^2. 因此，我们要讨论的问题就归结为判断原假设

$$H_0: a_1=a_2=a_3=a_4$$

是否成立，其中 σ^2 是未知参数.

（一）方差分析模型

专业课辅导+真题资料+微信公众号：432应统书院

由例 9.5.2 的讨论，将归纳出单因子多水平重复试验的数学模型及数据结构，如表 9.5.2 所示.

表 9.5.2

试验结果 \ 重复 / 试验号	1	2	…	n_i	平均	方差
1	ξ_{11}	ξ_{12}	…	ξ_{1n_1}	$\bar{\xi}_1$	S_1^2
2	ξ_{21}	ξ_{22}	…	ξ_{2n_2}	$\bar{\xi}_2$	S_2^2
⋮	⋮	⋮		⋮	⋮	⋮
r	ξ_{r1}	ξ_{r2}	…	ξ_{rn_r}	$\bar{\xi}_r$	S_r^2

记

$$\bar{\xi}_i=\frac{1}{n_i}\sum_{j=1}^{n}\xi_{ij},\qquad i=1,2,\cdots,r,$$

$$S_i^2=\frac{1}{n_i}\sum_{j=1}^{n_i}(\xi_{ij}-\bar{\xi}_i)^2,\quad i=1,2,\cdots,r,$$

则$\bar{\xi}_i$ 及 S_i^2 分别为第 i 个总体的样本平均及样本方差. 记

$$\bar{\xi}=\frac{1}{n}\sum_{i=1}^{r}\sum_{j=1}^{n_i}\xi_{ij}=\frac{1}{n}\sum_{i=1}^{r}n_i\bar{\xi}_i,$$

$$S^2=\frac{1}{n}\sum_{i=1}^{r}\sum_{j=1}^{n_i}(\xi_{ij}-\bar{\xi})^2,$$

其中 $n=\sum_{i=1}^{r}n_i$，称$\bar{\xi}$及 S^2 分别为全体样本的平均及方差，简称作总平均及总方差.

1. 基本假定及原假设 H_0

假定 $\xi_1,\xi_2,\cdots,\xi_r$ 相互独立，并假定 ξ_i 服从正态 $N(a_i,\sigma)$，其中 σ 为未知参数，$i=1,2,\cdots,r$.

考察如下原假设的检验问题：

专业课辅导+真题资料+微信公众号：432应统书院

$$H_0: a_1 = a_2 = \cdots = a_r. \tag{9.5.1}$$

2. 离差平方和的分解式

$$\begin{aligned} nS^2 &= \sum_{i=1}^{r} \sum_{j=1}^{n_i} (\xi_{ij} - \bar{\xi})^2 \\ &= \sum_{i=1}^{r} \sum_{j=1}^{n_i} [(\xi_{ij} - \bar{\xi}_i) + (\bar{\xi}_i - \bar{\xi})]^2 \\ &= \sum_{i=1}^{r} \sum_{j=1}^{n_i} (\xi_{ij} - \bar{\xi}_i)^2 + \sum_{i=1}^{r} \sum_{j=1}^{n} (\bar{\xi}_i - \bar{\xi})^2 \\ &= \sum_{i=1}^{r} n_i S_i^2 + \sum_{i=1}^{r} n_i (\bar{\xi}_i - \bar{\xi})^2, \end{aligned}$$

其中

$$\sum_{j=1}^{n_i} (\xi_{ij} - \bar{\xi}_i)(\bar{\xi}_i - \bar{\xi}) = 0, i = 1, 2, \cdots, r.$$

记

$$Q_e = \sum_{i=1}^{r} n_i S_i^2, U_1 = \sum_{i=1}^{r} n_i (\bar{\xi}_i - \bar{\xi})^2, \tag{9.5.2}$$

则有

$$nS^2 = Q_e + U_1. \tag{9.5.3}$$

记 $S_{总}^2 = nS^2$，称 $S_{总}^2$ 为总离差平方和. 称(9.5.3)式为总离差平方和的分解式，也称(9.5.2)式中的 Q_e 为误差项平方和，U_1 为因子的平方和.

3. 建立 F 检验统计量

在(9.5.2)式中，Q_e 是各总体中的样本方差 S_i^2 的线性组合，其中每个 S_i^2 都是 σ^2 的极大似然法估计量，是渐近无偏及渐近优效估计量，Q_e 反映了重复试验中误差的总大小. U_1 由各总体的样本平均 $\bar{\xi}_i$ 与总平均 $\bar{\xi}$ 之间的差异平方所组成，反映了各总体的样本平均之间的差异程度，而样本平均 $\bar{\xi}_i$ 是 a_i 的无偏估计量且是优效

专业课辅导+真题资料+微信公众号：432应统书院

估计量，因此 U_1 在一定程度上可反映原假设 H_0 是否成立. 如果 U_1 显著地大于 Q_e，说明各$\bar{\xi}_i$ 之间的差异大小显著地大于重复试验中误差的总大小，那么 H_0 可能不成立. 这种比较方差大小来判断原假设 H_0 是否成立的方法，就是方差分析法的名称由来. 列出方差分析表 9.5.3.

试问，U_1/Q_e 的比值大到什么程度，可以否定 H_0 呢？我们由第六章中定理 6.3.3 的推论知道，当 H_0 成立时，$\xi_1,\xi_2,\cdots,\xi_r$ 是 r 个相互独立的且都服从正态 $N(a,\sigma)$，因而统计量

$$\frac{U_1/(r-1)}{Q_e/(n-r)} \sim F_{(r-1,n-r)} \tag{9.5.4}$$

可以作为判断 H_0 是否成立的检验统计量.

给定显著性水平 α，怎样选择临界值呢？因为比值较大有可能否定 H_0，所以从右边取单边临界值. 在 F 分布临界值表中，查第一自由度 $r-1$，第二自由度 $n-r$ 对应的临界值 $F_{(r-1,n-r)}(\alpha)$，使得

$$P\{F_{(r-1,n-r)} \geqslant F_{(r-1,n-r)}(\alpha)\} = \alpha. \tag{9.5.5}$$

若按(9.5.4)式由样本观察值计算得 F 检验统计量的观察值 $F_{(r-1,n-r)}$ 大于 $F_{(r-1,n-r)}(\alpha)$，则在显著性水平 α 下否定 H_0，在实际工作中称之为以 $1-\alpha$ 的把握认为这个因子对试验指标的影响是显著的，也就是 a_i 之间的差异是显著的.

表 9.5.3　方差分析表

项目	离差平方和	自由度	方差	F 值
因子	U_1	$r-1$	$U_1/(r-1)$	$\dfrac{U_1/(r-1)}{Q_e/(n-r)}$
误差	Q_e	$n-r$	$Q_e/(n-r)$	
总和	$S^2_{总}$	$n-1$		

（二）例题的计算

例 9.5.2 的实际计算结果见表 9.5.4 和表 9.5.5. 此时 $r=4$，n

专业课辅导+真题资料+微信公众号：432应统书院

$=26.$

表 9.5.4　基本统计量计算表

水平 i	n_i	$\bar{\xi}_i$	S_i^2
1	7	1 674.29	3 238.71
2	5	1 598.00	16 816.00
3	8	1 648.75	5 835.94
4	6	1 575.00	4 091.67

$\bar{\xi}=1\ 628.85$，　$Q_e=177\ 988.51$，

$U_1=39\ 779.19$，　$nS^2=217\ 767.70.$

表 9.5.5　方差分析表

方差来源	平方和	自由度	方差	F 值
因子	39 779.19	3	13 259.73	1.64
误差	177 988.51	22	8 090.39	
总和	217 767.70	25		

给定显著性水平 $\alpha=0.05$，查得临界值 $F_{(3,22)}(0.05)=3.05$. 因为 $1.64<3.05$，所以在置显著性水平 0.05 下不否定 H_0，也就是说这四种灯丝生产的灯泡其平均使用寿命之间没有显著性的差异. 在实际工作中，通常说这个因子对试验指标的影响不显著，灯丝的四种配料方案质量相当，则可选取成本较低的配料方案.

习　题

1. 今有 10 组观测数据由下表给出：

专业课辅导+真题资料+微信公众号：432应统书院

x	0.5	-0.8	0.9	-2.8	6.5
y	-0.3	-1.2	1.1	-3.5	4.6
x	2.3	1.6	5.1	-1.9	-1.5
y	1.8	0.5	3.8	-2.8	0.5

应用线性模型：

$$y = \beta_0 + \beta_1 x + \varepsilon,$$

假定误差项 ε 服从正态 $N(0,\sigma)$，$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_{10}$ 为其样本.

(1) 求 β_0 及 β_1 的最小二乘法估计；

(2) 求 β_1 的置信水平为 0.95 的区间估计；

(3) 在显著性水平 $\alpha=0.05$ 下检验假设 $H_0:\beta_1=0$；

(4) 计算剩余方差 $\hat{\sigma}_e^2$；

(5) 求 $\hat{y}=\beta_0+\beta_1 x$ 的置信水平为 0.95 的预测区域.

2. 在维尼纶醛化试验中，固定其他因素，考虑醛浓度与反应时间对醛化度的关系，试验数据如下表：

醛化度 \ 反应时间 / 甲醛	3	5	7	12	20	30
32.10	17.8	22.9	25.9	29.9	32.9	35.4
33.00	18.2	22.9	25.1	28.6	31.2	34.1
27.60	16.8	20.0	23.6	28.0	30.0	33.1

记醛化度为 y，反应时间为 x_1，甲醛浓度为 x_2，由经验知道，y 与 x_2 成正比，而与 x_1 成反比，并有

$$\hat{y} = b_0 + b_1\frac{1}{x_1} + b_2 x_2.$$

试求 b_0、b_1 及 b_2 的最小二乘法估计值.

3. 物体降落的距离 s 与时间 t 的关系，由下式确定

$$\hat{s} = \beta_0 + \beta_1 t + \beta_2 t^2.$$

专业课辅导+真题资料+微信公众号：432应统书院

若测得数据如下表：

t_i/s	1/30	2/30	3/30	4/30	5/30	6/30	7/30	8/30
s_i/cm	11.86	15.67	20.60	26.69	33.71	41.93	51.13	61.49
t_i/s	9/30	10/30	11/30	12/30	13/30	14/30	15/30	
s_i/cm	72.90	85.44	99.08	113.77	129.54	146.48	165.06	

假定 s_i 服从正态 $N(\beta_0+\beta_1 t_i+\beta_2 t_i^2,\sigma)(i=1,2,\cdots,15)$.

(1) 试确定 β_0、β_1、β_2 及 σ^2 的极大似然估计；

(2) 检验原假设 $H_0:\beta_2=0$；

(3) 给出 s 的置信水平 0.95 的预测区域.

4. 试作变量变换，将 §9.3 之二的 1～7 各曲线方程化为线性方程.

5. 设 ξ_1 与 ξ_2 为独立的随机变量，数学期望分别为 θ 与 2θ，试求 θ 的最小二乘法估计与剩余平方和.

6. 设 $\boldsymbol{y}'=(y_1,y_2,y_3)$ 服从线性模型 $(\boldsymbol{y},\boldsymbol{X\beta},\sigma^2\boldsymbol{I}_3)$，其中

$$E(y_i)=\beta_0+\beta_1 x_i+\beta_2(3x_i^2-2),\qquad i=1,2,3.$$

$x_1=-1,x_2=0,x_3=1$. 试写出矩阵 $\boldsymbol{X}$，并求出 β_0,β_1,β_2 的最小二乘法估计，且证明当 $\beta_2=0$ 时，β_0 与 β_1 的最小二乘法估计量不变.

7. 设

$$\begin{aligned} y_i &= \theta+\varepsilon_i, & i&=1,2,\cdots,m;\\ y_{m+i} &= \theta+\phi+\varepsilon_{m+i}, & i&=1,2,\cdots,m;\\ y_{2m+i} &= \theta-2\phi+\varepsilon_{2m+i}, & i&=1,2,\cdots,n.\end{aligned}$$

假定 ε_i 之间互不相关，且有

$$E(\varepsilon_i)=0,\ D(\varepsilon_i)=\sigma^2,\ i=1,2,\cdots,2m+n.$$

试求 θ 及 ϕ 的最小二乘法估计. 试证当 $m=2n$ 时，$\hat{\theta}$ 与 $\hat{\phi}$ 互不相关.

8. 设

$$\begin{cases} y_1=a+\varepsilon_1;\\ y_2=2a-b+\varepsilon_2;\\ y_3=a+2b+\varepsilon_3,\end{cases}$$

其中 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 相互独立，且有

$$E(\varepsilon_i)=0,\ D(\varepsilon_i)=\sigma^2,\ i=1,2,3.$$

专业课辅导+真题资料+微信公众号：432应统书院

试求 a 及 b 的最小二乘法估计量.

9. 某医院用光电比色计检尿汞时,获得尿汞含量(mg/l)与消光系数读数的数据如下表:

尿汞含量 x_i	2	4	6	8	10
消光系数 y_i	64	138	205	285	360

已知它们之间服从线性模型

$$E(y_i) = \beta_0 + \beta_1 x_i.$$

试求 β_0 和 β_1 的最小二乘法估计,并在显著性水平 $\alpha=0.05$ 下,检验原假设 $H_0: \beta_1=0$ 是否成立.

10. 今有四个物体,按下面方法称,得到数据如下表:

x_1	x_2	x_3	x_4	y
1	1	1	1	20.2
1	-1	1	-1	8.0
1	1	-1	-1	9.2
1	-1	-1	1	1.4

其中 1 表示该物体放在天平左边,-1 表示放在天平的右边,y 是使天平达到平衡时,在右边所加砝码的重量,试估计这四个物体的重量 β,$i=1,2,3,4$.

11. 粮食加工厂用四种不同的方法贮藏粮食,贮藏一段时间后,分别抽样化验,得到粮食含水率如下:

贮藏方法	含水率/%				
A	7.3	8.3	7.6	8.4	8.3
B	5.8	7.4	7.1		
C	8.1	6.4	7.0		
D	7.9	9.0			

专业课辅导+真题资料+微信公众号：432应统书院

检验这四种不同的贮藏方法对粮食的含水率是否有显著影响.

12. 小车产量 x(万辆)与薄钢板的需求量 y(万吨)有关,某汽车厂前五个月份的数据资料如下:

序号 i	x_i	y_i
1	13.98	191.90
2	13.52	199.37
3	12.54	217.19
4	14.91	302.62
5	18.60	303.99

试用线性回归分析:$y=\beta_0+\beta_1 x+\varepsilon$. ε 服从 $N(0,\sigma)$. 求 β_0 及 β_1 的最小二乘法估计;在显著性水平 $\alpha=0.05$ 下,检验假设 $H_0:\beta_1=0$.

13. 某地区对服务性工作人员需求量的预测. 用 x 表示平均国民收入(百美元),用 y 表示每千人口所需要的服务性工作人员数量(个). 现记录了 $n=12$ 组数据资料,试建立指数曲线方程:

$$y=ab^x, \qquad a>0, b>0.$$

注:记 $Z=\lg y, A=\lg a, B=\lg b$. 建立线性回归方程

$$Z=A+Bx+\varepsilon, \varepsilon \text{ 服从 } N(0,\sigma).$$

求 A、B 的最小二乘法估计;在显著性水平 $\alpha=0.05$ 下检验假设 $H_0:B=0$.

序号 i	x_i	y_i
1	1.2	8.0
2	2.0	12.0
3	3.0	10.0
4	4.8	12.5
5	8.3	17.0
6	8.4	21.3

专业课辅导+真题资料+微信公众号：432应统书院

续表

序号 i	x_i	y_i
7	11.5	25.0
8	14.0	47.3
9	14.2	38.6
10	14.8	76.4
11	15.6	97.3
12	16.1	88.0

14. 某地百货商店的销售额(万元)用 y_t 表示，t 为年份序号，记录了 $n=10$ 年的数据资料，试建立(Gomperz)曲线回归方程，预测未来两年的销售额.

t	0	1	2	3	4
y_t	2 239	2 760	3 206	3 417	3 200
t	5	6	7	8	9
y_t	3 308	4 182	4 381	5 610	6 510

$$y_t=ka^{b^t},(k>0,a>0,a\neq 1,b>0,b\neq 1,t>0).$$

试求 k、a 及 b 三个参数的估计值.

注：由时间数列 $\{y_t\}$ 求参数的最小二乘法估计，要用 Q 对 k,a,b 分别求导数. 求解超越方程困难较大.

$$Q(k,a,b)=\sum_{t=1}^{n}[y_t-ka^{b^t}]^2,$$

现用三和值法确定参数的估计值. 记

$$Z_t=\ln y_t,\ K=\ln k,\ A=\ln a,$$

则有

$$Z_t=K+Ab^t.$$

将时间数列 $\{Z_t\}$ 按时间顺序等分三组，分组求和.

$$\mathrm{I}=\sum_{t=0}^{m-1}Z_t=mK+A\sum_{t=0}^{m-1}b^t,$$

专业课辅导+真题资料+微信公众号：432应统书院

$$\text{Ⅱ}=\sum_{t=m}^{2m-1} Z_t = mK + Ab^m \sum_{t=1}^{m-1} b^t,$$

$$\text{Ⅲ}=\sum_{t=2m}^{3m-1} Z_t = mK + Ab^{2m} \sum_{t=0}^{m-1} b^t,$$

$$\text{Ⅱ}-\text{Ⅰ} = A(b^m-1) \sum_{t=0}^{m-1} b^t,$$

$$\text{Ⅲ}-\text{Ⅱ} = Ab^m(b^m-1) \sum_{t=0}^{m-1} b^t,$$

$$b^m = \frac{\text{Ⅲ}-\text{Ⅱ}}{\text{Ⅱ}-\text{Ⅰ}}, \quad A = \frac{b-1}{(b^m-1)^2}[\text{Ⅱ}-\text{Ⅰ}],$$

$$K = \frac{1}{m}\left[1-\left(\frac{b^m-1}{b-1}\right)A\right].$$

15. 柯布-道格拉斯生产函数.

$$Q_t = AK_t^{\alpha}L_t^{\beta},\ A>0, K>0, L>0.$$

认为生产总值 Q_t 同劳力 L_t 及资本 K_t 有关，同技术进步 A 的贡献也有关. 某地制造业记录了 $n=12$ 个年份的数据资料：

年份序号 t	Q_t/万元	K_t/万元	L_t/个
1	144 336	47 729	37 392
2	176 397	46 172	40 818
3	241 911	46 850	48 222
4	311 551	57 929	54 773
5	371 024	74 475	61 045
6	479 744	122 995	63 128
7	568 366	149 694	60 087
8	646 235	186 251	61 570
9	927 999	250 220	76 248
10	1 028 315	295 506	78 070
11	1 135 466	327 464	68 980
12	1 280 389	365 835	81 845

专业课辅导+真题资料+微信公众号：432应统书院

记 $y_t=\ln Q_t, x_{t1}=\ln K_t, x_{t2}=\ln L_t, t=1,2,\cdots,n$

$$\beta_0=\ln A,\ \beta_1=\alpha,\ \beta_2=\beta,$$

则有：
$$y_t=\beta_0+\beta_1 x_{t1}+\beta_2 x_{t2}+\varepsilon_t,\ \varepsilon_t \text{服从} N(0,\sigma).$$

试求 β_0,β_1,β_2 的最小二乘法估计.

专业课辅导+真题资料+微信公众号：432应统书院

第十章　统计决策及贝叶斯统计

前面几章所论述的统计推断(估计、检验、预测等),有许多学者称它为经典统计学.统计推断所使用的信息有两方面:一是总体信息,即总体 ξ 具有精确分布或总体 ξ 具有所属分布族,如正态分布 $N(a,\sigma)$,指数分布族,这给我们带来很重要的信息:二是样本信息,即从总体中随机抽取的样本给我们提供的信息.样本来自总体,样本信息是现场信息,第一手信息,人们希望通过对样本的加工和处理,能较为精确地对总体的某些特征进行统计推断.没有样本就没有统计推断学.从19世纪末期到20世纪上半叶,为创立经典统计学作出重大贡献的有三位统计学家:皮尔逊(Pearson, K. 1857—1936)、费希尔(Fisher, R. A. 1890—1962)、奈曼(Neyman, J.):如今统计学教材都是主要叙述经典统计学的理论与方法.20世纪下半叶,经典统计学在工业、农业、医学、经济、管理、军事等领域里获得了广泛的应用,这又促进了它的发展.

如果在抽样之前即获得样本信息之前,对统计推断的有关统计问题能从实践经验和历史资料获得某些信息,这种信息称之为先验信息.基于上述三种信息(总体信息、样本信息、先验信息)进行的统计推断被称为贝叶斯统计学.如企业的产品不合格率(或不良品率)p 的估计,矩法与极大似然法确定的估计量是样本不合格率 $\frac{d}{n}$(n 为样本容量,d 为其中的不合格个数).一般情况下,企业对产品质量状况作跟踪观测,对产品不合格率的大小积累了大量的资料,如 $0.05<p<0.10$,或者有进一步的信息:

$$P\{p=0.06\}=0.5,\quad P\{p=0.07\}=0.2,$$

$$P\{p=0.08\}=0.2,\quad P\{p=0.09\}=0.1.$$

专业课辅导+真题资料+微信公众号：432应统书院

此时则称 p 有先验分布. 假定参数为随机变量并有先验分布，这是贝叶斯统计学与经典统计学的主要区别.

英国学者贝叶斯(Bayes, T. R.)在1763年发表的论著中提出参数 θ 为随机变量的观点. 随后拉普拉斯(Laplace, P. C. 1749—1827)等人用贝叶斯提出的方法导出一些有意义的结果. 贝叶斯统计学的形成及其发展，是直到二次大战后，瓦尔德(Wald, A. 1902—1950)提出统计决策函数论，引起很多学者对贝叶斯方法研究的兴趣，使得贝叶斯方法在观点、方法和理论上不断地完善. 在这段发展时期，贝叶斯统计在工业、经济、管理等领域内成功的应用，使得贝叶斯统计更为成熟.

经典统计学派与贝叶斯统计学派是统计学中两大主要学派. 这两个学派在有关先验分布这个起点上展开了长期的学术争论. 在这些争论中提出了理论上和实践中许多学术问题，促使双方的许多学者作进一步的研讨，不断做出新的研究成果，从而大大地推动了统计学的进展. 事实上对于许多不同的实际问题，有些运用经典统计推断方法是很有效的，而有些运用贝叶斯统计方法是很合理的. 目前各国出版的统计学教材中，仍是经典统计学占有主要的篇幅. 但是，贝叶斯统计学有许多重要研究成果优于经典统计学.

怎样简述贝叶斯统计的基本内容，值得讨论. 本书上册第一章§1.4条件概率中，讲述了贝叶斯公式，贝叶斯公式指出事后概率 $P(A_i \mid B)$，可通过一系列的"原因概率" $P(B \mid A_j)$ 求得. 第二章§2.3条件分布中，讨论两个随机变量为连续型或离散型的情形，讲述条件概率函数或条件密度函数. 第三章§3.5对条件数学期望作了简要叙述. 上述有关内容，为§10.2贝叶斯统计的讨论，提供了必要的准备，如贝叶斯公式的建立. 关于参数的贝叶斯估计这个概念，是通过统计决策函数的基本概念引入，如损失函数、风险函数等.

本章介绍统计决策论及贝叶斯统计的基础知识. 贝叶斯统计方法在统计推断(点估计与区间估计、参数检验、统计预测等)中

专业课辅导+真题资料+微信公众号：432应统书院

的应用内容很丰富、贝叶斯统计有许多重要研究成果，请参看书目[5]及[7].

§10.1 极大极小估计

一、决策论的基本概念

在这一节里，我们将运用更一般的观点——决策论来考察点估计问题. 为此，我们先介绍决策论的一些名词和概念，如决策函数、损失函数、风险函数等.

设总体 ξ 的分布函数为 $F(x;\theta)$，$\theta\in\Omega$，θ 为未知参数，Ω 为参数空间. $\xi_1,\xi_2,\cdots,\xi_n$ 为其样本，求参数 θ 的点估计，可看作由样本 $\xi_1,\xi_2,\cdots,\xi_n$ 的一次观察值——样本值域 $\mathscr{X}$ 中的一点 $(x_1,x_2,\cdots,x_n)$，对未知参数 θ 采取一种决定. 在数理统计学中，我们称所采取的一个决定为决策，将可能采取的全部决策所组成的集合称为决策空间，记为 $\mathscr{A}$. 因此，对未知参数 θ 或对分布 $F(x;\theta)$ 作出估计，实质上是对样本值域 $\mathscr{X}$ 中的每一点 $(x_1,x_2,\cdots,x_n)$，在决策空间 $\mathscr{A}$ 中寻找一点 $a\in\mathscr{A}$ 与之对应.

在定义 6.1.2 中给出了统计量的定义. 现若将我们选取的统计量 $d(\xi_1,\xi_2,\cdots,\xi_n)$ 作为 θ 或 $F(x;\theta)$ 的估计，则称 $d(\xi_1,\xi_2,\cdots,\xi_n)$ 为决策函数，称 $a=d(x_1,x_2,\cdots,x_n)$ 为一个决策.

例 10.1.1 点估计问题的决策论解释.

设总体 ξ 有分布函数 $F(x;\theta)$，$\theta\in\Omega$，θ 为未知参数，要求由样本 $\xi_1,\xi_2,\cdots,\xi_n$ 建立统计量作为 θ 的点估计量. 这时，每个估计量就是一个决策函数.

我们知道，对于每个具体的统计问题，总存在着许多不同的决策函数可利用，因而就需要建立某些原则，能够使我们评价决策函数的优劣. 为此引入损失函数，记为 $L(\theta,a)$，本章所讲的 $L(\theta,a)$ 是参数 θ 及决策 a 的一个实值非负函数，它表示参数为 θ 时而采取决策 a 所造成的损失. 当然，我们希望当参数为 θ 时，能采取一个“最好”的决策，即使损失 L 达到最小的决策. 但是 θ 的参数值是未知的，此时我们就

专业课辅导+真题资料+微信公众号：432应统书院

无法知道最好的决策是什么. 我们是利用样本建立决策函数 $d(\xi_1,\xi_2,\cdots,\xi_n)$ 来产生决策的，它所对应的损失为 $L(\theta,a)=L[\theta,d(\xi_1,\xi_2,\cdots,\xi_n)]$，所以这个损失是依赖于样本的，因而它是随机变量. 因此，我们还不能运用基于样本的某次观察 $(x_1,x_2,\cdots,x_n)$ 所采取的某个决策所带来的损失 $L[\theta,d(x_1,x_2,\cdots,x_n)]$ 来衡量，而应从总体上来评价，现采用平均损失来作评价. 记

$$R(\theta,d)=E_\theta\{L[\theta,d(\xi_1,\xi_2,\cdots,\xi_n)]\}. \tag{10.1.1}$$

它是损失函数在总体 ξ 的分布函数取参数为 θ 时计算得到的数学期望. 称 $R(\theta,d)$ 为风险函数. 它是参数 θ 及决策函数 $d(\xi_1,\xi_2,\cdots,\xi_n)$ 的函数，它表示参数为 θ 时而采取决策 $d(\xi_1,\xi_2,\cdots,\xi_n)$ 所造成的平均损失.

例 10.1.2 设总体 ξ 服从泊松分布 $P(x,\lambda)$，$\lambda\in\Omega=(0,\infty)$，要求估计未知参数 λ. 如果我们选取损失函数为

$$L(\lambda,d)=(d-\lambda)^2, \tag{10.1.2}$$

即二次损失函数，则对 λ 的任一估计量 $d(\xi_1,\xi_2,\cdots,\xi_n)$，风险函数是

$$R(\lambda,d)=E_\lambda[L(\lambda,d)]=E_\lambda[(d-\lambda)^2].$$

如进一步要求估计量 $d(\xi_1,\xi_2,\cdots,\xi_n)$ 是无偏的，则

$$E_\lambda[d(\xi_1,\xi_2,\cdots,\xi_n)]=\lambda, \tag{10.1.3}$$

$$R(\lambda,d)=E_\lambda[d-E_\lambda(d)]^2=D_\lambda(d). \tag{10.1.4}$$

它是决策函数 $d(\xi_1,\xi_2,\cdots,\xi_n)$ 的方差. 例如取 $d(\xi_1,\xi_2,\cdots,\xi_n)=\bar{\xi}$，那么所对应的风险函数为：

$$R(\lambda,d)=D_\lambda(\bar{\xi})=\frac{\lambda}{n}.$$

若取 $d(\xi_1,\xi_2,\cdots,\xi_n)=\xi_1$，那么所对应的风险函数为：

$$R(\lambda,d)=D_\lambda(\xi_1)=\lambda.$$

显然，当 $n>1$ 时，后者的“风险”比前者大，称 $\bar{\xi}$ 优于 ξ_1.

二、极大极小估计

由上面可看到风险函数提供了一个衡量决策函数优劣的准

专业课辅导+真题资料+微信公众号：432应统书院

则. 当然我们希望选取一个决策函数,使得它的风险尽可能地小.

定义 10.1.1 极大极小估计. 设 $\mathscr{D}$ 是由决策函数为元素组成的集合,若有 $d^*(\xi_1,\xi_2,\cdots,\xi_n)\in\mathscr{D}$,对任一 $d(\xi_1,\xi_2,\cdots,\xi_n)\in\mathscr{D}$,使得

$$\sup_{\theta\in\Omega}\{R(\theta,d^*)\}\leqslant\sup_{\theta\in\Omega}\{R(\theta,d)\},\tag{10.1.5}$$

则称 $d^*(\xi_1,\xi_2,\cdots,\xi_n)$ 为参数 θ 的极大极小估计量.

这个定义的直观背景是很有意思的. 使得最大风险达到最小的决策函数 $d^*(\xi_1,\xi_2,\cdots,\xi_n)$ 是考虑到最不利的情况,并要求最不利的情况尽可能地好的一种策略,这就是通常所说的从最坏处着想,争取最好的结果,是一种稳妥的考虑,也是一种偏于保守的考虑.

例 10.1.3 设总体 ξ 服从两点分布,即有:

$$P\{\xi=x\}=\begin{cases}p, & x=1;\\ 1-p, & x=0,\end{cases}$$

其中 $p\in\Omega=\left\{\frac{1}{4},\frac{1}{2}\right\}$,讨论参数 p 的极大极小估计量.

在本例中参数 p 的决策空间 $\mathscr{A}=\left\{\frac{1}{4},\frac{1}{2}\right\}$. 设损失函数 $L(p,a)$ 由表 10.1.1 给定.

如果我们选取容量为 1 的样本记为 ξ_1,由于 ξ_1 仅取两个可能值及 $\mathscr{A}$ 中只有两个元素,因而决策函数的集合 $\mathscr{D}$ 是由 4 个元素所组成,分别以 d_1,d_2,d_3,d_4 表示,即有表 10.1.2. 它们分别造成的风险 $R(p,d)$ 记录在表 10.1.3 中.

表 10.1.1

$L(p,a)$ (a 值 / p 值)	$a_1=\frac{1}{4}$	$a_2=\frac{1}{2}$
$p_1=\frac{1}{4}$	1	4
$p_2=\frac{1}{2}$	3	2

专业课辅导+真题资料+微信公众号：432应统书院

表 10.1.2

ξ \ $\mathscr{D}$	d_1	d_2	d_3	d_4
0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{2}$
1	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$

表 10.1.3

d_i	$R(p_1,d_i)$	$R(p_2,d_i)$	$\sup R(p_i,d_i)$
d_1	1	3	3
d_2	4	2	4
d_3	$\frac{7}{4}$	$\frac{5}{2}$	$\frac{5}{2}$
d_4	$\frac{13}{4}$	$\frac{5}{2}$	$\frac{13}{4}$

由(10.1.1)式计算得到 $R(p,d)$，例如

$$\begin{aligned}R(p_1,d_3)&=E_{p_1}[L(p_1,d_3)]\\&=L(p_1,a_1)\cdot P_{p_1}\{\xi=0\}+L(p_1,a_2)\cdot P_{p_1}\{\xi=1\}\\&=1\cdot(1-p_1)+4\cdot p_1=\frac{7}{4},\end{aligned}$$

$$\begin{aligned}R(p_2,d_3)&=L(p_2,a_1)\cdot P_{p_2}\{\xi=0\}+L(p_2,a_2)\cdot P_{p_2}\{\xi=1\}\\&=3\cdot(1-p_2)+2\cdot p_2=\frac{5}{2}.\end{aligned}$$

由(10.1.5)式知，参数 p 的极大极小估计为：

$$\hat{p}(\xi_1)=d_3=\begin{cases}\frac{1}{4},\xi_1=0;\\\frac{1}{2},\xi_1=1.\end{cases}\tag{10.1.6}$$

§10.2 贝叶斯统计

前述各节都是在未知参数 θ 作为非随机变量的情况下讨论统

专业课辅导+真题资料+微信公众号：432应统书院

计推断问题. 本节内容将参数 θ 看作为一随机变量,所涉及的随机变量包括复合所得的随机变量,仅讨论连续型和离散型情形,并且用较直观的方法叙述有关内容. 关于条件分布、条件数学期望对于连续型、离散型情形下的严格叙述与论证,请参看上册 §2.3 及 §3.5. 本节关于这些概念的应用与运算,应看作是直观的理解,或是在足够强的条件下进行.

一、贝叶斯估计

设总体 ξ 的密度函数为 $f(x;\theta)$, $\theta\in\Omega$,贝叶斯学派的观点是把参数 θ 看作随机变量 $\boldsymbol{\theta}$,假定 $\boldsymbol{\theta}$ 具有给定的密度函数 $\pi(\theta)$(通常称 $\pi(\theta)$ 为参数 $\boldsymbol{\theta}$ 的先验分布),认为 $f(x;\theta)$ 是 $\boldsymbol{\theta}$ 取值 θ 时 ξ 的条件密度函数 $f(x|\theta)$. 假定二维连续型随机变量 $(\xi,\boldsymbol{\theta})$ 的联合密度函数为 $f(x,\theta)$,它可由下式确定:

$$f(x,\theta)=\pi(\theta)f(x|\theta). \tag{10.2.1}$$

对于决策函数 $d(\xi_1,\xi_2,\cdots,\xi_n)$,由(10.1.1)式给出的风险函数 $R(\theta,d)$ 应写为:

$$R(\theta,d)=E\{L[\boldsymbol{\theta},d(\xi_1,\xi_2,\cdots,\xi_n)]|\boldsymbol{\theta}=\theta\}, \tag{10.2.2}$$

(10.2.2)式右边是随机变量 $\boldsymbol{\theta}$ 取值 θ 时损失函数的条件期望值. 由于 $\boldsymbol{\theta}$ 取值 θ,遵从密度函数 $\pi(\theta)$ 的规律,再取一次平均,可得

$$B(d)=\int_{\Omega}E\{L[\boldsymbol{\theta},d(\xi_1,\xi_2,\cdots,\xi_n)]\mid\boldsymbol{\theta}=\theta\}\pi(\theta)\mathrm{d}\theta, \tag{10.2.3}$$

称 $B(d)$ 为决策函数 $d(\xi_1,\xi_2,\cdots,\xi_n)$ 的贝叶斯风险.

定义 10.2.1(贝叶斯估计) 设总体 ξ 的分布函数 $F(x,\theta)$ 中参数 θ 为随机变量 $\boldsymbol{\theta}$,对于任一决策函数 $d(\xi_1,\xi_2,\cdots,\xi_n)$,若有一决策函数 $d^*(\xi_1,\xi_2,\cdots,\xi_n)$,使得

$$B(d^*)=\min_{d\in\mathscr{D}}\{B(d)\}, \tag{10.2.4}$$

则称 d^* 为参数 θ 的贝叶斯估计量.

从(10.2.4)看到,贝叶斯估计量 $d^*(\xi_1,\xi_2,\cdots,\xi_n)$ 是使得贝叶斯风险 $B(d)$ 达到最小的决策函数. 贝叶斯估计量是依赖于先

专业课辅导+真题资料+微信公众号：432应统书院

验分布 $\pi(\theta)$ 的，也就是说，对应于不同的 $\pi(\theta)$，θ 的贝叶斯估计量是不同的.

下面对 $(\xi,\boldsymbol{\theta})$ 为二维连续型随机变量的情形，建立(10.2.5)式，讨论两个定理. 定理的结论也适用于 ξ 及 $\boldsymbol{\theta}$ 中一方或两者都为离散型的情形，只是要用条件概率函数代替条件密度函数，概率函数代替密度函数的位置，用求和式代替积分式，并相应地推出(10.2.5)式的其他表达式.

$(\xi,\boldsymbol{\theta})$ 的联合密度函数 $f(x,\theta)$ 又可分解为：

$$f(x,\theta)=g(x)h(\theta|x),$$

其中

$$g(x)=\int_{-\infty}^{\infty}f(x,\theta)\mathrm{d}\theta=\int_{-\infty}^{\infty}\pi(\theta)f(x\mid\theta)\mathrm{d}\theta,$$

$$h(\theta\mid x)=\frac{f(x,\theta)}{g(x)}=\frac{\pi(\theta)f(x\mid\theta)}{\int_{-\infty}^{\infty}\pi(\theta)f(x\mid\theta)\mathrm{d}\theta},\tag{10.2.5}$$

当 $g(x)=0$ 或 ∞ 时，令 $h(\theta|x)=\pi(\theta)$，通常称(10.2.5)式为贝叶斯公式($h(\theta|x)$ 为给定 $\xi=x$ 时 $\boldsymbol{\theta}$ 的条件密度函数，称为 $\boldsymbol{\theta}$ 的后验分布). (10.2.3)式可写为：

$$\begin{aligned}B(d)&=\underbrace{\int\cdots\int}_{(n+1)\text{重}}L[\theta,d(x_1,x_2,\cdots,x_n)]f(x_1,x_2,\cdots,x_n\mid\theta)\\&\quad\times\pi(\theta)\mathrm{d}x_1\cdots\mathrm{d}x_n\mathrm{d}\theta\\&=\underbrace{\int\cdots\int}_{n\text{重}}g(x_1,x_2,\cdots,x_n)\left\{\int_{\Omega}L[\theta,d(x_1,x_2,\cdots,x_n)]\right.\\&\quad\left.\times h(\theta|x_1,x_2,\cdots,x_n)\mathrm{d}\theta\right\}\mathrm{d}x_1\cdots\mathrm{d}x_n.\end{aligned}\tag{10.2.6}$$

定理 10.2.1 如果损失函数取二次式：

$$L(\theta,d)=[\theta-d(\xi_1,\xi_2,\cdots,\xi_n)]^2,\tag{10.2.7}$$

则参数 θ 的贝叶斯估计量为：

$$d(\xi_1,\xi_2,\cdots,\xi_n)=E(\boldsymbol{\theta}\mid\xi_1,\xi_2,\cdots,\xi_n)$$

专业课辅导+真题资料+微信公众号：432应统书院

$$= \int_{\Omega} \theta h(\theta \mid \xi_1, \xi_2, \cdots, \xi_n) \mathrm{d}\theta. \quad (10.2.8)$$

证 我们知道，

$$B(d) = \underbrace{\int \cdots \int}_{n重} g(x_1, x_2, \cdots, x_n) \left\{ \int_{\Omega} [\theta - d(x_1, x_2, \cdots, x_n)]^2 \right.$$

$$\left. \times h(\theta \mid x_1, x_2, \cdots, x_n) \mathrm{d}\theta \right\} \mathrm{d}x_1 \cdots \mathrm{d}x_n. \quad (10.2.9)$$

(10.2.9)式对于 $d \in \mathscr{D}$ 中取最小值，几乎处处等价于下面的(10.2.10)式对于 $d \in \mathscr{D}$ 中取最小值：

$$\int_{\Omega} [\theta - d(x_1, x_2, \cdots, x_n)]^2 h(\theta \mid x_1, x_2, \cdots, x_n) \mathrm{d}\theta. \quad (10.2.10)$$

为方便书写，用 $\boldsymbol{X}$ 表示 $(x_1, x_2, \cdots, x_n)$，(10.2.10)式可写为：

$$\begin{aligned}
&\int_{\Omega} [\theta - d(\boldsymbol{X})]^2 h(\theta \mid \boldsymbol{X}) \mathrm{d}\theta \\
&= \int_{\Omega} [\theta - E(\boldsymbol{\theta} \mid \boldsymbol{X}) + E(\boldsymbol{\theta} \mid \boldsymbol{X}) - d(\boldsymbol{X})]^2 h(\theta \mid \boldsymbol{X}) \mathrm{d}\theta \\
&= \int_{\Omega} [\theta - E(\boldsymbol{\theta} \mid \boldsymbol{X})]^2 h(\theta \mid \boldsymbol{X}) \mathrm{d}\theta \\
&\quad + \int_{\Omega} [E(\boldsymbol{\theta} \mid \boldsymbol{X}) - d(\boldsymbol{X})]^2 h(\theta \mid \boldsymbol{X}) \mathrm{d}\theta \\
&\quad + 2\int_{\Omega} [\theta - E(\boldsymbol{\theta} \mid \boldsymbol{X})][E(\boldsymbol{\theta} \mid \boldsymbol{X}) - d(\boldsymbol{X})] h(\theta \mid \boldsymbol{X}) \mathrm{d}\theta,
\end{aligned}$$

其中

$$E(\boldsymbol{\theta} \mid \boldsymbol{X}) = \int_{\Omega} \theta h(\theta \mid \boldsymbol{X}) \mathrm{d}\theta.$$

又

$$\begin{aligned}
&\int_{\Omega} [\theta - E(\boldsymbol{\theta} \mid \boldsymbol{X})][E(\boldsymbol{\theta} \mid \boldsymbol{X}) - d(\boldsymbol{X})] h(\theta \mid \boldsymbol{X}) \mathrm{d}\theta \\
&= \int_{\Omega} [\theta - E(\boldsymbol{\theta} \mid \boldsymbol{X})] h(\theta \mid \boldsymbol{X}) \mathrm{d}\theta \cdot [E(\boldsymbol{\theta} \mid \boldsymbol{X}) - d(\boldsymbol{X})] \\
&= [E(\boldsymbol{\theta} \mid \boldsymbol{X}) - E(\boldsymbol{\theta} \mid \boldsymbol{X})] \cdot [E(\boldsymbol{\theta} \mid \boldsymbol{X}) - d(\boldsymbol{X})] = 0,
\end{aligned}$$

专业课辅导+真题资料+微信公众号：432应统书院

所以

$$\int_{\Omega}[\theta-d(\boldsymbol{X})]^2h(\theta\mid\boldsymbol{X})\mathrm{d}\theta=\int_{\Omega}[\theta-E(\boldsymbol{\theta}\mid\boldsymbol{X})]^2h(\theta\mid\boldsymbol{X})\mathrm{d}\theta$$

$$+\int_{\Omega}[E(\boldsymbol{\theta}\mid\boldsymbol{X})-d(\boldsymbol{X})]^2h(\theta\mid X)\mathrm{d}\theta. \quad (10.2.11)$$

上式等号右边第一项与 $d(\boldsymbol{X})$ 无关，显然当

$$d(\boldsymbol{X})=E(\boldsymbol{\theta}\mid\boldsymbol{X})$$

几乎处处成立时，左边达到最小.

例 10.2.1 设总体 ξ 服从两点分布，即有：

$$P(x,p)=\begin{cases}p, & x=1;\\ 1-p, & x=0,\end{cases}$$

$p\in\Omega=(0,1)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本. 假定损失函数是二次误差，$L(p,d)=(p-d)^2$，并假定参数 p 是随机变量，服从均匀分布，即有：

$$\pi(p)=1,\quad 0<p<1,$$

试求参数 p 的贝叶斯估计量.

解 样本 $(\xi_1,\xi_2,\cdots,\xi_n)$ 的边缘概率函数为：

$$\begin{aligned}g(x_1,x_2,\cdots,x_n)&=\int_0^1 p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}\mathrm{d}p\\&=B\Big(\sum_{i=1}^n x_i+1,n+1-\sum_{i=1}^n x_i\Big)\\&=\frac{\Big[\Big(\sum_{i=1}^n x_i\Big)!\Big]\Big[\Big(n-\sum_{i=1}^n x_i\Big)!\Big]}{(n+1)!},\end{aligned}$$

参照(10.2.5)式求参数 p 的后验分布为：

$$h(p\mid x_1,x_2,\cdots,x_n)=[(n+1)!]\frac{p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}}{\Big[\Big(\sum_{i=1}^n x_i\Big)!\Big]\Big[\Big(n-\sum_{i=1}^n x_i\Big)!\Big]}.$$

由定理 10.2.1 知，参数 p 的条件数学期望为：

专业课辅导+真题资料+微信公众号：432应统书院

$$\hat{p}(x_1,x_2,\cdots,x_n)=\int_0^1 ph(p\mid x_1,x_2,\cdots,x_n)\mathrm{d}p$$

$$=\frac{(n+1)!}{\left[\left(\sum_{i=1}^{n}x_i\right)!\right]\left[\left(n-\sum_{i=1}^{n}x_i\right)!\right]}$$

$$\times\int_0^1 p^{1+\sum_{i=1}^{n}x_i}(1-p)^{n-\sum_{i=1}^{n}x_i}\mathrm{d}p$$

$$=\frac{1+\sum_{i=1}^{n}x_i}{n+2},$$

也即参数 p 的贝叶斯估计量为

$$\hat{p}(\xi_1,\xi_2,\cdots,\xi_n)=\frac{1}{n+2}\left[1+\sum_{i=1}^{n}\xi_i\right]. \qquad (10.2.12)$$

注：在例 7.3.7 中指出，样本平均 $\bar{\xi}=\frac{1}{n}\sum_{i=1}^{n}\xi_i=\frac{d}{n}$（$d$ 为样本不合格率）是参数 p 的最小方差无偏估计. 但是，这估计量 $\hat{p}$ 并非完美无缺. 如从总体中随机抽取 20 个产品检查都为合格品，则 p 的估计值为零；随机抽取 10 个产品检查都为合格品，则 p 的估计值为零；随机抽取 5 个产品检查都为合格品，p 的估计值也为零. 但是例 10.2.1 指出，p 的贝叶斯估计为 $\hat{p}=(d+1)/(n+2)$，在上述三次随机抽样场合，不合格率 p 的估计值分别为 1/22、1/12、1/7. 这表明贝叶斯估计更为合理.

例 10.2.2 设总体 ξ 服从正态 $N(a,1)$，$a\in\Omega=\mathbf{R}_1$，$\xi_1,\xi_2,\cdots,\xi_n$ 为其样本. 假定参数 a 服从正态 $N(0,1)$，又假定损失函数为二次误差 $L(a,d)=(a-d)^2$. 试求参数 a 的贝叶斯估计量.

解 样本的条件密度函数为

$$f(x_1,x_2,\cdots,x_n\mid a)=\frac{1}{(\sqrt{2\pi})^n}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-a)^2\right\},$$

样本和参数 a 的联合密度函数为

$$f(x_1,x_2,\cdots,x_n;a)=f(x_1,x_2,\cdots,x_n\mid a)\pi(a)$$

专业课辅导+真题资料+微信公众号：432应统书院

$$= \frac{1}{(\sqrt{2\pi})^{n+1}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i - a)^2 - \frac{1}{2}a^2\right\}, \tag{10.2.13}$$

样本的边缘密度函数为

$$g(x_1, x_2, \cdots, x_n) = \int_{-\infty}^{\infty} f(x_1, x_2, \cdots, x_n \mid a)\pi(a)\mathrm{d}a$$

$$= \frac{1}{(\sqrt{2\pi})^{n+1}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right\}$$

$$\times \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}[(n+1)a^2 - 2an\bar{x}]\right\}\mathrm{d}a$$

$$= \frac{1}{(\sqrt{2\pi})^{n}} \exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{n} x_i^2 - \frac{n^2\bar{x}^2}{n+1}\right]\right\}$$

$$\times \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{n+1}{2}\left[a - \frac{n\bar{x}}{n+1}\right]^2\right\}\mathrm{d}a$$

$$= \frac{1}{\sqrt{n+1}(\sqrt{2\pi})^{n}} \exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{n} x_i^2 - \frac{n^2\bar{x}^2}{n+1}\right]\right\}.$$

由(10.2.5)式知参数 a 的后验分布为

$$h(a \mid x_1, x_2, \cdots, x_n) = \frac{f(x_1, x_2, \cdots, x_n \mid a)\pi(a)}{g(x_1, x_2, \cdots, x_n)}$$

$$= \frac{\sqrt{n+1}}{\sqrt{2\pi}} \exp\left\{-\frac{n+1}{2}\left[a - \frac{n\bar{x}}{n+1}\right]^2\right\}$$

由定理 10.2.1 知，参数 a 的条件数学期望为：

$$E(a \mid x_1, x_2, \cdots, x_n) = \int_{-\infty}^{\infty} a h(a \mid x_1, x_2, \cdots, x_n)\mathrm{d}a$$

$$= \int_{-\infty}^{\infty} \frac{a}{\sqrt{2\pi}\sigma_0} \exp\left\{-\frac{1}{2\sigma_0^2}\left[a - \frac{n\bar{x}}{n+1}\right]^2\right\}\mathrm{d}a$$

$$= \frac{n\bar{x}}{n+1},$$

其中 $\sigma_0 = \frac{1}{\sqrt{n+1}}$，参数 a 的贝叶斯估计量为

$$\hat{a}(\xi_1, \xi_2, \cdots, \xi_n) = \frac{1}{n+1}\sum_{i=1}^{n} \xi_i. \tag{10.2.14}$$

专业课辅导+真题资料+微信公众号：432应统书院

定理 10.2.2 在线性损失函数

$$L[\theta,d]=\begin{cases}k_0(\theta-d),\theta\geqslant d;\\ k_1(d-\theta),\theta<d\end{cases}\tag{10.2.15}$$

下，θ 的贝叶斯估计量 $\hat{\theta}(\xi_1,\xi_2,\cdots,\xi_n)$ 为 $\boldsymbol{\theta}$ 的后验分布 $h(\theta|\xi_1,\xi_2,\cdots,\xi_n)$ 的 $\lambda=k_0/(k_0+k_1)$ 分位数.

证 在定理 10.2.1 的证明过程中已指出，后验风险最小的估计量同样是贝叶斯风险最小的估计量（几乎处处成立），见(10.2.9)式及(10.2.10)式，简记 $\boldsymbol{X}=(\xi_1,\xi_2,\cdots,\xi_n)$，任一决策函数 $d(\boldsymbol{X})$ 的后验风险为

$$R[d(\boldsymbol{X})|\boldsymbol{X}]=\int_{-\infty}^{\infty}L[\theta,d]h(\theta|\boldsymbol{X})\mathrm{d}\theta$$

$$=k_1\int_{-\infty}^{d}(d-\theta)h(\theta|\boldsymbol{X})\mathrm{d}\theta+k_0\int_{d}^{\infty}(\theta-d)h(\theta|\boldsymbol{X})\mathrm{d}\theta$$

$$=(k_0+k_1)\int_{-\infty}^{d}(d-\theta)h(\theta|\boldsymbol{X})\mathrm{d}\theta+k_0[E(\theta|\boldsymbol{X})-d].$$

$$\frac{\partial R[d(\boldsymbol{X})|\boldsymbol{X}]}{\partial d}=(k_0+k_1)\int_{-\infty}^{d}h(\theta|\boldsymbol{X})\mathrm{d}\theta-k_0.\tag{10.2.16}$$

令(10.2.16)式为零，则有

$$\int_{-\infty}^{d}h(\theta|\boldsymbol{X})\mathrm{d}\theta=\frac{k_0}{k_0+k_1}=\lambda,\tag{10.2.17}$$

使得后验风险达到最小的决策函数 $d^*(\boldsymbol{X})$，由 $\boldsymbol{\theta}$ 的后验分布 $h(\theta|\boldsymbol{X})$ 的 $\lambda=k_0/(k_0+k_1)$ 分位数确定.

(10.2.9)式及(10.2.10)式即为

$$B[d(\boldsymbol{X})]=\int_{\Omega}R[\theta,d(\boldsymbol{X})]\pi(\theta)\mathrm{d}\theta$$

$$=\int_{\Omega}\left\{\int_{X}L[\theta,d(\boldsymbol{X})]f(\boldsymbol{X}|\theta)\mathrm{d}\boldsymbol{X}\right\}\pi(\theta)\mathrm{d}\theta$$

$$=\int_{X}\left\{\int_{\Omega}L[\theta,d(\boldsymbol{X})]h(\theta|\boldsymbol{X})\mathrm{d}\theta\right\}g(\boldsymbol{X})\mathrm{d}\boldsymbol{X}$$

$$=\int_{X}R[d(\boldsymbol{X})|\boldsymbol{X}]g(\boldsymbol{X})\mathrm{d}\boldsymbol{X}.\tag{10.2.18}$$

专业课辅导+真题资料+微信公众号：432应统书院

(10.2.18)式中积分顺序可交换性成立的条件是每个决策函数的贝叶斯风险 $B[d(\boldsymbol{X})]$ 都是有限的. 这个要求在实际中是容易得到满足的,因为贝叶斯风险无穷大的决策函数,对于实际决策者来说是不感兴趣的,理论上意义也不大.

二、区间估计

在例 7.4.1 中讨论了正态总体的数学期望(即均值) a 的置信水平为 $1-\alpha$ 的区间估计. 在方差 $\sigma^2=\sigma_0^2$ 为已知时,由(7.4.2)式确定 a 的区间估计,区间中心为 $\bar{\xi}$,区间长度 $\Delta_1(\boldsymbol{X})=2\sigma_0 u_\alpha/\sqrt{n}$,其中 $\boldsymbol{X}=(\xi_1,\xi_2,\cdots,\xi_n)$,$u_\alpha$ 由标准正态表查得的分位数值(即临界值). 当方差 σ^2 未知时,由(7.4.3)式确定 a 的区间估计,区间中心为 $\bar{\xi}$,区间长度 $\Delta_2(\boldsymbol{X})=2St_{(n-1)}(\alpha)/\sqrt{n-1}$,其中 S^2 为样本方差,$t_{(n-1)}(\alpha)$ 由 t 分布表查得的分位数值. 这是在参数 a 为非随机变量情形建立 a 的置信区间,区间的中心都为样本平均 $\bar{\xi}$,区间的上下置信限由统计量 $\bar{\xi}$ 或 S^2 确定.

当参数 $\boldsymbol{\theta}$ 为随机变量且有先验分布 $\pi(\theta)$ 时,θ 的区间估计则由 $\boldsymbol{\theta}$ 的后验分布 $h(\theta|\boldsymbol{X})$ 建立.

定义 10.2.2 设参数 $\boldsymbol{\theta}$ 的后验分布为 $h(\theta|\boldsymbol{X})$,$\boldsymbol{X}=(\xi_1,\xi_2,\cdots,\xi_n)$,对于给定的概率 $1-\alpha(0<\alpha<1)$,若存在两个统计量 $T_1(\boldsymbol{X})$ 与 $T_2(\boldsymbol{X})$,使得

$$P\{T_1(\boldsymbol{X})\leqslant\theta\leqslant T_2(\boldsymbol{X})\mid\boldsymbol{X}\}\geqslant 1-\alpha,$$

则称随机区间 $[T_1(\boldsymbol{X}),T_2(\boldsymbol{X})]$ 为参数 θ 的置信水平为 $1-\alpha$ 的贝叶斯区间估计,满足

$$P\{\theta\geqslant T_1(\boldsymbol{X})\mid\boldsymbol{X}\}\geqslant 1-\alpha$$

的 $T_1(\boldsymbol{X})$ 称为 θ 的置信水平 $1-\alpha$ 的置信下限. 满足

$$P\{\theta\leqslant T_2(\boldsymbol{X})\mid\boldsymbol{X}\}\geqslant 1-\alpha$$

的 $T_2(\boldsymbol{X})$ 称为 θ 的置信水平 $1-\alpha$ 的置信上限.

例 10.2.3 设总体 ξ 服从 $N(a,\sigma_0)$,$\sigma_0>0$ 为已知值,$\boldsymbol{X}=(\xi_1,\xi_2,\cdots,\xi_n)$ 为其样本,假定参数 a 服从 $N(0,1)$. 试求 a 的置信

专业课辅导+真题资料+微信公众号：432应统书院

水平 $1-\alpha$ 的贝叶斯区间估计.

解 $f(x_1,x_2,\cdots,x_n;a)=\pi(a)f(x_1,x_2,\cdots,x_n|a)$

$$=\frac{1}{\sqrt{2\pi}(\sqrt{2\pi}\sigma_0)^n}\exp\left\{-\frac{1}{2\sigma_0^2}\left[\sum_{i=1}^{n}x_i^2-2an\bar{x}+(n+\sigma_0^2)a^2\right]\right\}. \tag{10.2.19}$$

$$g(x_1,x_2,\cdots,x_n)=\int_{-\infty}^{\infty}f(x_1,x_2,\cdots,x_n;a)\mathrm{d}a$$

$$=\frac{1}{(\sqrt{2\pi}\sigma_0)^n}\exp\left\{-\frac{1}{2\sigma_0^2}\left[\sum_{i=1}^{n}x_i^2-\frac{n^2\bar{x}^2}{n+\sigma_0^2}\right]\right\}\cdot\frac{\sigma_0}{\sqrt{n+\sigma_0^2}}. \tag{10.2.20}$$

由(10.2.5)式知

$$h(a|x_1,x_2,\cdots,x_n)=f(x_1,x_2,\cdots,x_n|a)/g(x_1,x_2,\cdots,x_n)$$

$$=\frac{\sqrt{n+\sigma_0^2}}{\sqrt{2\pi}\sigma_0}\exp\left\{-\frac{n+\sigma_0^2}{2\sigma_0^2}\left[a-\frac{n\bar{x}}{n+\sigma_0}\right]^2\right\}. \tag{10.2.21}$$

于是在给定样本 $\boldsymbol{X}=(\xi_1,\xi_2,\cdots,\xi_n)$ 的条件下,参数 a 的后验分布为 $N\left(\frac{n\bar{\xi}}{n+\sigma_0^2},\frac{\sigma_0}{\sqrt{n+\sigma_0^2}}\right)$.

记

$$U=\frac{a-n\bar{\xi}/(n+\sigma_0^2)}{\sigma_0/\sqrt{n+\sigma_0^2}}, \tag{10.2.22}$$

则 U 服从 $N(0,1)$,且有

$$P\{-u_\alpha\leqslant U\leqslant u_\alpha|\boldsymbol{X}\}=1-\alpha,$$

$$P\left\{\frac{n\bar{\xi}}{n+\sigma_0^2}-\frac{\sigma_0u_\alpha}{\sqrt{n+\sigma_0^2}}\leqslant a\leqslant\frac{n\bar{\xi}}{n+\sigma_0^2}+\frac{\sigma_0u_\alpha}{\sqrt{n+\sigma_0^2}}\Big|\boldsymbol{X}\right\}=1-\alpha, \tag{10.2.23}$$

贝叶斯区间估计的中心是 a 的贝叶斯估计 $\hat{a}=\frac{n\bar{\xi}}{n+\sigma_0^2}$,区间长度用

专业课辅导+真题资料+微信公众号：432应统书院

$\Delta_3(X)$表示，且有

$$\Delta_3(X)=\frac{2\sigma_0 u_\alpha}{\sqrt{n+\sigma_0^2}}<\Delta_1(X)=\frac{2\sigma_0 u_\alpha}{\sqrt{n}}.$$

贝叶斯区间估计的精度比经典的区间估计要高. 这两个区间估计的含义不同. 贝叶斯区间估计的含义是：参数 a 是随机变量，在给定样本 $X=(x_1,x_2,\cdots,x_n)$ 条件下，置信上下限是由统计量的观察值计算得到的数值：

$$T_1(X)=\frac{n\bar{x}}{n+\sigma_0^2}-\frac{\sigma_0 u_\alpha}{\sqrt{n+\sigma_0^2}},T_2(X)=\frac{n\bar{x}}{n+\sigma_0^2}+\frac{\sigma_0 u_\alpha}{\sqrt{n+\sigma_0^2}},$$

a 落在这个区间的概率为 $1-\alpha$.

三、假设检验

在例 8.4.4 中对于正态总体 $N(a,\sigma_0)$，其中 $\sigma_0>0$ 为已知值，讨论了关于 a 的几种情形下假设检验的最佳否定域. 如

$$H_0: a=a_0, H_1: a=a_1(>a_0).$$

运用奈曼-皮尔逊定理，给定显著性水平 α，$0<\alpha<1$，H_0的最佳否定域为

$$\mathscr{X}_0=\{(x_1,x_2,\cdots,x_n): U\geqslant u_\alpha\},$$

其中检验统计量 $U=\sqrt{n}(\bar{\xi}-a_0)/\sigma_0$，$u_\alpha$ 为 $N(0,1)$ 的右侧临界值，即

$$P_{H_0}\{U\geqslant u_\alpha\}=\alpha.$$

假设检验的贝叶斯方法是：不引进检验统计量、两类错判概率及 H_0的否定域. 利用参数 θ 的后验分布 $h(\theta|X)$，分别计算 H_0与 H_1的后验概率 $p_0=h(\theta\in\Omega_0|X)$ 与 $p_1=h(\theta\in\Omega_1|X)$，其中 $X=(x_1,x_2,\cdots,x_n)$ 为样本观察值. 当 $p_0<p_1$ 时，则否定 H_0；当 $p_0>p_1$ 时，不否定 H_0. Ω 为参数空间

$$H_0: \theta\in\Omega_0,\quad H_1: \theta\in\Omega-\Omega_0=\Omega_1.$$

例 10.2.4 设总体 ξ 服从 $N(a,\sigma_0)$，$\sigma_0>0$ 为已知值，考虑参数 a 的简单假设检验问题：

专业课辅导+真题资料+微信公众号：432应统书院

$$H_0: a=3, \qquad H_1: a=5.$$

假定参数 a 有先验信息：

$$\pi\{a_0=3\}=0.8, \qquad \pi\{a_1=5\}=0.2.$$

设充分性统计量$\bar{\xi}$有观察值$\bar{x}=3.8$，试用贝叶斯方法对 H_0是否成立作出统计推断($n=16$).

解

$$f(\bar{x}|a_0)=\frac{\sqrt{n}}{\sigma_0\sqrt{2\pi}}\exp\left\{-\frac{n}{2\sigma_0^2}(\bar{x}-3)^2\right\},$$

$$f(\bar{x}|a_1)=\frac{\sqrt{n}}{\sigma_0\sqrt{2\pi}}\exp\left\{-\frac{n}{2\sigma_0^2}(\bar{x}-5)^2\right\}.$$

$$g(\bar{x})=0.8f(\bar{x}|a_0)+0.2f(\bar{x}|a_1).$$

$$p_0=h(a_0|\bar{x})=0.8f(\bar{x}|a_0)/g(\bar{x}).$$

$$p_1=h(a_1|\bar{x})=0.2f(\bar{x}|a_1)/g(\bar{x}).$$

$$\frac{p_0}{p_1}=\frac{0.8\exp\left\{-\frac{n}{2\sigma_0^2}(\bar{x}-3)^2\right\}}{0.2\exp\left\{-\frac{n}{2\sigma_0^2}(\bar{x}-5)^2\right\}}$$

$$=4\exp\left\{-\frac{2n\bar{x}}{\sigma_0^2}+\frac{8n}{\sigma_0^2}\right\}.$$

假定 $\sigma_0=2$，则有

$$\frac{p_0}{p_1}=4\exp\{-8\bar{x}+32\}=4e^{1.6}=19.812,$$

于是否定 H_0.

例 10.2.5 设总体 ξ 服从 $N(a,\sigma_0)$，$\sigma_0=2$ 为已知值，讨论参数 a 的复合假设检验问题：

$$H_0: a\leqslant 1, H_1: a>1.$$

假定参数 a 有先验分布 $N(0.5,2)$，设从总体 ξ 中随机抽取 $n=10$ 的样本，样本平均$\bar{x}=1.2$，试用贝叶斯方法对 H_0是否成立作出统计推断.

解 设总体 ξ 服从正态分布 $N(a,\sigma_0)$，参数 a 服从正态分布

专业课辅导+真题资料+微信公众号：432应统书院

$N(b,d)$，样本平均$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$，由(10.2.5)式推得参数 a 的后验分布为 $N(a_1,\sigma_1)$，其中

$$a_1=\frac{\bar{x}d^2+b\sigma_0^2/n}{d^2+\sigma_0^2/n},\frac{1}{\sigma_1^2}=\frac{1}{\sigma_0^2/n}+\frac{1}{d^2}. \tag{10.2.24}$$

在 $\sigma_0=2,n=10,\bar{x}=1.2,d=2,b=0.5$ 的情形，则

$$a_1=\frac{1.2\times4+0.5\times4/10}{4+4/10}=1.136,\frac{1}{\sigma_1^2}=\frac{1}{4/10}+\frac{1}{4}=2.75,$$

亦即 a 的后验分布为 $N(1.136,0.603)$，计算后验概率如下：

$$p_0=P\{a\leqslant1\mid\bar{x}\}=\Phi\left(\frac{1-1.136}{0.603}\right)=\Phi(-0.2255)=0.411,$$

$$p_1=P\{a>1\mid\bar{x}\}=1-P\{a\leqslant1\mid\bar{x}\}=0.589,$$

由于 $p_0<p_1$，于是否定 H_0.

*四、共轭先验分布

例 10.2.1 中总体分布为两点分布，设成功概率 p 的先验分布 $\pi(p)$ 是(0,1)上的均匀分布，则 p 的后验分布 $h(p\mid\boldsymbol{X})$ 是 β(贝塔)分布 $Be(\alpha,\beta)$，其中 $\alpha=1+\sum_{i=1}^{n}x_i,\beta=n+1-\sum_{i=1}^{n}x_i$.

例 10.2.2 中总体分布为 $N(a,1)$，设均值 a 的先验分布 $\pi(a)$ 为 $N(0,1)$，则 a 的后验分布 $h(a\mid\boldsymbol{X})$ 为 $N\left(\frac{n\bar{x}}{n+1},\frac{1}{\sqrt{n+1}}\right)$，其中 $n\bar{x}=\sum_{i=1}^{n}x_i$.

在例 10.2.5 中看到，当总体分布为方差已知的正态分布时，则均值 a 的先验分布与后验分布都是正态分布，只是参数不同，函数形式是相同的.

定义 10.2.3 设 θ 是总体 ξ 分布中的参数，$\boldsymbol{X}=(\xi_1,\xi_2,\cdots,\xi_n)$ 表示样本，$\pi(\theta)$ 为 $\boldsymbol{\theta}$ 的先验分布(密度函数或概率函数)，如果 $\boldsymbol{\theta}$ 的后验分布 $h(\theta\mid\boldsymbol{X})$(密度函数或概率函数)与 $\pi(\theta)$ 有相同的函数形式，则称 $\pi(\theta)$ 是 $\boldsymbol{\theta}$ 的共轭先验分布.

专业课辅导+真题资料+微信公众号：432应统书院

参数 θ 的贝叶斯估计 $\hat{\theta}(\boldsymbol{X})$，同总体 ξ 的分布、$\boldsymbol{\theta}$ 的先验分布 $\pi(\theta)$ 及损失函数 $L(\theta,d)$ 都有关. 如何确定 $\boldsymbol{\theta}$ 的先验分布 $\pi(\theta)$，是贝叶斯统计中最困难的、必须解决而又易引起争议的问题. 从实用角度考虑，应充分利用历史上积累的信息（数据、经验）进行分析或模拟，注重采纳专家的经验或见解，综合分析以确定先验分布. 贝叶斯统计学者对这个问题有很多研究成果，如建立"协调性假说"，选取的先验分布 $\pi(\theta)$ 与由它和样本 $\boldsymbol{X}$ 确定的后验分布 $h(\theta|\boldsymbol{X})$ 属于同一类型. 例如，所有正态分布构成一类；所有 Γ 分布构成一类；所有 β 分布构成一类. 许多实践表明，这个假说是符合实际的，运用起来也方便.

例 10.2.6 设总体 ξ 服从泊松分布 $P(\lambda)$，参数 λ 服从 Γ 分布 $\Gamma(\alpha,\beta)$，$\boldsymbol{X}=(\xi_1,\xi_2,\cdots,\xi_n)$ 为样本，试求 λ 的后验分布 $(\lambda>0)$.

解 样本的联合概率函数

$$f(x_1,x_2,\cdots,x_n|\lambda)=\frac{\lambda^{x_1+x_2+\cdots+x_n}}{x_1!\ \cdots\ x_n!}\mathrm{e}^{-n\lambda}\quad(\lambda>0),$$

λ 的先验分布密度为

$$\pi(\lambda)=\frac{\beta^{\alpha}\lambda^{\alpha-1}}{\Gamma(\alpha)}\mathrm{e}^{-\beta\lambda}\quad(\lambda>0),$$

由(10.2.5)式知，λ 的后验分布密度为：

$$h(\lambda|\boldsymbol{X})=\frac{f(x_1,x_2,\cdots,x_n\mid\lambda)\pi(\lambda)}{\int_0^{\infty}f(x_1,x_2,\cdots,x_n\mid\lambda)\pi(\lambda)\mathrm{d}\lambda}$$

$$=\frac{(\beta+n)^{\alpha+n\bar{x}}\lambda^{\alpha+n\bar{x}-1}}{\Gamma(\alpha+n\bar{x})}\mathrm{e}^{-(\beta+n)\lambda}\quad(\lambda>0).\qquad(10.2.25)$$

例 10.2.7 设总体服从两点分布（伯努利分布）.

$$P\{\xi=1\}=p,\qquad P\{\xi=0\}=1-p,\qquad 0<p<1.$$

$\boldsymbol{X}=(\xi_1,\xi_2,\cdots,\xi_n)$ 为样本，设成功概率 p 有先验分布为 β 分布 $Be(\alpha,\beta)$，试求 p 的后验分布.

专业课辅导+真题资料+微信公众号：432应统书院

解 样本 $X=(\xi_1,\xi_2,\cdots,\xi_n)$ 的联合概率函数为：

$$f(x_1,x_2,\cdots,x_n|p)=\prod_{i=1}^{n}p^{x_i}(1-p)^{1-x_i},x_i=1\text{ 或 }0.$$

记 $n\bar{x}=\sum_{i=1}^{n}x_i$，则有：

$$f(x_1,x_2,\cdots,x_n|p)=p^{n\bar{x}}(1-p)^{n-n\bar{x}},0<p<1.$$

p 的先验分布为 β 分布 $Be(\alpha,\beta)$，概率密度为：

$$\pi(p)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1}\quad(0<p<1),$$

由(10.2.5)或知 p 的后验分布为：

$$h(p|X)=\frac{f(x_1,x_2,\cdots,x_n\mid p)\pi(p)}{\int_0^1 f(x_1,x_2,\cdots,x_n\mid p)\pi(p)\mathrm{d}p}$$

$$=\frac{\Gamma(\alpha+\beta+n)}{\Gamma(\alpha+n\bar{x})\Gamma(\beta+n-n\bar{x})}p^{\alpha+n\bar{x}-1}(1-p)^{\beta+n-n\bar{x}-1},\tag{10.2.26}$$

即 p 的后验分布为 β 分布 $Be(\alpha+n\bar{x},\beta+n-n\bar{x})$.

例 10.2.8 设总体 ξ 服从指数分布 $E(\lambda)$，参数 λ 的先验分布为 $\Gamma(\alpha,\beta)$，$X=(\xi_1,\xi_2,\cdots,\xi_n)$ 为样本，试求 λ 的后验分布($\lambda>0$).

解 样本 X 的联合密度函数为：

$$f(x_1,x_2,\cdots,x_n|\lambda)=\lambda^n\mathrm{e}^{-\lambda\sum_{i=1}^{n}x_i},\lambda>0.$$

λ 的先验分布密度为：$\pi(\lambda)=\dfrac{\beta^\alpha\lambda^{\alpha-1}}{\Gamma(\alpha)}\mathrm{e}^{-\beta\lambda}\quad(\lambda>0)$.

由(10.2.5)式知，λ 的后验分布密度为：

$$h(\lambda|X)=\frac{f(x_1,x_2,\cdots,x_n\mid\lambda)\pi(\lambda)}{\int_0^\infty f(x_1,x_2,\cdots,x_n\mid\lambda)\pi(\lambda)\mathrm{d}\lambda}$$

专业课辅导+真题资料+微信公众号：432应统书院

$$=\frac{(\beta+n\bar{x})^{\alpha+n}\lambda^{\alpha+n-1}}{\Gamma(\alpha+n)}e^{-(\beta+\sum_{i=1}^{n}x_i)\lambda}\qquad(\lambda>0).\tag{10.2.27}$$

记 $n\bar{x}=\sum_{i=1}^{n}x_i$，则 λ 的后验分布为 $\Gamma(\alpha+n,\beta+n\bar{x})$. 就是说对于参数 λ 的指数分布，其均值 $E(\xi)=\frac{1}{\lambda}$，对于 λ 而言，Γ 分布是共轭先验分布.

例 10.2.9 设总体 ξ 服从正态分布 $N(a_0,\sigma)$，a_0 为已知，σ 为未知，$\boldsymbol{X}=(\xi_1,\xi_2,\cdots,\xi_n)$ 为样本. 记 $\theta=\frac{1}{\sigma^2}$，设 $\boldsymbol{\theta}$ 的先验分布为 $\Gamma(\alpha,\beta)$，试求 $\boldsymbol{\theta}$ 的后验分布.

解 样本 $\boldsymbol{X}$ 的联合密度函数为：

$$f(x_1,x_2,\cdots,x_n\mid\theta)=\left(\sqrt{\frac{\theta}{2\pi}}\right)^n\exp\left\{-\frac{\theta}{2}\sum_{i=1}^{n}(x_i-a_0)^2\right\}.$$

$\boldsymbol{\theta}$ 的先验分布密度为： $\pi(\theta)=\frac{\beta^\alpha\theta^{\alpha-1}}{\Gamma(\alpha)}e^{-\beta\theta}$，

由(10.2.5)式知，$\boldsymbol{\theta}$ 的后验分布密度为：

$$h(\theta\mid\boldsymbol{X})=\frac{f(x_1,x_2,\cdots,x_n\mid\theta)\pi(\theta)}{\int_0^\infty f(x_1,x_2,\cdots,x_n\mid\theta)\pi(\theta)\mathrm{d}\theta}$$

$$=\left[\beta+\frac{1}{2}\sum_{i=1}^{n}(x_i-a_0)\right]^{\alpha+\frac{n}{2}}\frac{\theta^{\alpha+\frac{n}{2}-1}}{\Gamma\left(\alpha+\frac{n}{2}\right)}\cdot$$

$$\exp\left\{-\left[\beta+\frac{1}{2}\sum_{i=1}^{n}(x_i-a_0)^2\right]\theta\right\},\tag{10.2.28}$$

即 $\boldsymbol{\theta}$ 的后验分布为 $\Gamma\left(\alpha+\frac{n}{2},\beta+\frac{1}{2}\sum_{i=1}^{n}(x_i-a_0)^2\right)$. 就是说对于均值 a 为已知的正态分布，方差 σ^2 的倒数为参数 θ，Γ 分布为共

专业课辅导+真题资料+微信公众号：432应统书院

轭先验分布.

§10.3 应用事例

例 10.3.1 某企业拟一年后生产一种新产品上市，研讨了三种实施方案：

a_1——改建本企业的原有生产线；

a_2——引进一条自动化生产线；

a_3——同另一企业协作组织生产线.

将新产品上市的市场前景状态 $\boldsymbol{\theta}$ 分为四种：θ_1（很好）、θ_2（较好）、θ_3（较差）、θ_4（很差）. 对三种实施方案在市场四种状态下三年内的经济效益（单位：百万元）作了初步评估，提出收益矩阵 $\boldsymbol{Q}=(q_{ij})$ 如下：

$$\boldsymbol{Q}=\begin{pmatrix} 750 & 1\,100 & 500 \\ 500 & 700 & 250 \\ 250 & -500 & 100 \\ -200 & -800 & -50 \end{pmatrix}_{4\times 3}.$$

在给定 $\boldsymbol{Q}$ 矩阵的情况下，试问哪个方案为优？对此，人们提出了几种不同的判断准则.

（一）线性决策方法

1. 悲观准则与乐观准则

悲观准则是指按极小极大准则选优方案：

$$\max_j\{\min_i q_{ij}\}. \tag{10.3.1}$$

悲观准则是指每个方案从最坏处着想（市场状态很差）的保守稳妥准则. 本例有

$$\min_i q_{ij}=\begin{cases} -200, j=1; \\ -800, j=2; \\ -50, \ j=3. \end{cases} \qquad \max_j\{\min_i q_{ij}\}=-50.$$

a_3 方案为优方案. 这种准则较适合资金实力单薄、规模较小、财力经不起冲击的小型企业.

专业课辅导+真题资料+微信公众号：432应统书院

乐观准则是指按极大极大准则选优方案：

$$\max_j \{ \max_i q_{ij} \}. \tag{10.3.2}$$

本例有

$$\max_i q_{ij} = \begin{cases} 750, & j=1; \\ 1\ 100, & j=2; \\ 500, & j=3. \end{cases} \qquad \max_i \{ \max_j q_{ij} \} = 1\ 100.$$

a_2方案为优方案. 乐观准则是收益矩阵 $\boldsymbol{Q}$ 中每个方案选最大元素(即市场前景状态很好),然后比较这几个方案,取它们中的最大元素为最大者,所对应的方案为优方案. 这种准则较适合财力雄厚的大中型企业在开发小项目中搏一搏的大胆策略.

2. 赫维斯的折中准则

赫维斯(Harwicx)认为上述两准则都是极端性准则,提出用权系数 $b(0<b<1)$ 对这两个准则建立线性组合,称之为折中准则:

$$H(a_j) = b \max_i q_{ij} + (1-b) \min_i q_{ij}, \tag{10.3.3}$$

称 $H(a_j)$ 为方案 a_j 的平均收益值. 最大者对应的方案为最优方案.

b 值的选择是折中准则的关键,取决于决策者对产品上市的市场前景状态的判断. b 值的大小反映了决策者对乐观准则的支持程度,直接影响到所选的哪个方案为优.

$$H(a_1) = 750b + (1-b)(-200) = 950b - 200 \overset{\text{def}}{=\!=} H_1;$$

$$H(a_2) = 1\ 100b + (1-b)(-800) = 1\ 900b - 800 \overset{\text{def}}{=\!=} H_2;$$

$$H(a_3) = 500b + (1-b)(-50) = 550b - 50 \overset{\text{def}}{=\!=} H_3.$$

H_1、H_2、H_3 中哪个最大,这同 b 的取值有关. b 限在区间[0,1]范围中有定义,怎样把它划分成三个小区间,才能明确到该选哪个方案为优? 直线 H_1 与直线 H_2 的交点 C 的横坐标 $b_2 = 0.631\ 6$,直线 H_1 与直线 H_3 的交点 B 的横坐标 $b_1 = 0.375$.

当 $0 \leqslant b \leqslant 0.375$ 时,直线 H_3 在上方,a_3 方案为优;

当 $0.375 < b \leqslant 0.631\ 6$ 时,直线 H_1 在上方,a_1 方案为优;

专业课辅导+真题资料+微信公众号：432应统书院

当 $0.631\ 6<b\leqslant 1$ 时，直线 H_2 在上方，a_2 方案为优.

（二）贝叶斯决策方法

贝叶斯决策方法是指市场前景状态 $\boldsymbol{\theta}$ 视为随机变量，如设 $\boldsymbol{\theta}$ 有先验分布 $\pi(\theta)$：

$$\pi(\theta_i)=p_i,0\leqslant p_i\leqslant 1,\ \sum_{i=1}^{4}p_i=1.$$

计算先验期望收益值 $\overline{H}(a_j)$：

$$\overline{H}(a_j)=\sum_{i=1}^{4}p_iq_{ij},j=1,2,3. \tag{10.3.4}$$

按 $\overline{H}(a_j)$ 中最大者对应的方案为最优方案的准则，又称为先验期望准则.

如若 $p_1=0.5,p_2=0.3,p_3=p_4=0.1$，则有

$$\overline{H}(a_1)=530,\quad \overline{H}(a_2)=630,\quad \overline{H}(a_3)=330,$$

于是 a_2 方案为最优方案. 先验期望准则对市场前景状态考虑比较仔细，尽量减少决策失误.

例 10.3.2 某小学对新生作体检，若某孩子在体检中的身高为 110(cm)，试问该孩子的身高 θ 为多少？只有一次观察值，怎样对 θ 作统计推断？现简介贝叶斯统计方法.

（一）贝叶斯估计

基本假定：(1) 该小孩身高测量值 ξ 服从正态分布 $N(\theta,5)$，即标准差 $\sigma_0=5$；(2) 设身高值 $\boldsymbol{\theta}$ 有先验分布 $N(108,3)$，对照(10.2.24)式的记号，即 $b=108,d=3$；(3) 损失函数 $L(\theta,\delta)$ 采用一次式. 依据多次对小孩身高 θ 估计或评价的反应与投诉，发现低估造成的损失为高估造成的损失的两倍，于是建立

$$L(\theta,\delta)=\begin{cases}2(\theta-\delta),\delta\leqslant\theta;\\ \delta-\theta,\delta>\theta,\end{cases} \tag{10.3.5}$$

其中 δ 表示对 θ 的估计或决策.

在上述假定下，试求 θ 的贝叶斯估计 $\hat{\theta}(\boldsymbol{X})$，样本 $\boldsymbol{X}=(\xi_1,$

专业课辅导+真题资料+微信公众号：432应统书院

$\xi_2,\cdots,\xi_n)$，在此 $n=1$.

解 由(10.2.24)式知，$\boldsymbol{\theta}$ 的后验分布为 $N(a_1,\sigma_1)$，其中

$$a_1=\frac{\bar{x}d^2+b\sigma_0^2}{d^2+\sigma_0^2}=\frac{110\times3^2+108\times5^2}{3^2+5^2}=108.53,$$

$$\frac{1}{\sigma_1^2}=\frac{1}{d^2}+\frac{1}{\sigma_0^2}=\frac{1}{9}+\frac{1}{25}=0.151\ 11,\sigma_1=2.572\ 7.$$

运用定理 10.2.2，$k_0=2,k_1=1$，则 $\lambda=\frac{2}{3}$. 在标准正态分布中 $\lambda=\frac{2}{3}$ 分位数值为 $u_\alpha=0.43$，即

$$P\{U\geqslant0.43\}=\frac{2}{3}.$$

于是 $\boldsymbol{\theta}$ 的后验分布 $N(108.53,2.572\ 7)$ 的 $\lambda=\frac{2}{3}$ 分位数为

$$\hat{\theta}=a_1+\sigma_1u_\alpha=109.636\ 3,$$

这就是 θ 的贝叶斯估计值.

（二）贝叶斯决策

假定对 θ 的评估有两个假设：

$$H_1:90\leqslant\theta\leqslant110,\quad H_2:\theta>110,$$

用 d_i 表示接受 H_i 的决策($i=1,2$)，试问该采用 d_1 或 d_2 决策？

（1）建立 d_1 及 d_2 的损失函数

$$L(\theta,d_1)=\begin{cases}0, & 90\leqslant\theta\leqslant110;\\ \theta-110, & \theta>110.\end{cases}\tag{10.3.6}$$

$$L(\theta,d_2)=\begin{cases}110-\theta, & 90\leqslant\theta\leqslant110;\\ 0, & \theta>110.\end{cases}\tag{10.3.7}$$

（2）计算 d_1 及 d_2 的后验期望损失值

$$h(\theta|x_1)=\frac{1}{\sqrt{2\pi}(2.572\ 7)}e^{-(\theta-108.53)^2/2\times2.572\ 7^2},$$

$$R(d_i|x_1)=\int_{-\infty}^{\infty}L(\theta,d_i)h(\theta|x_1)\mathrm{d}\theta.\tag{10.3.8}$$

专业课辅导+真题资料+微信公众号：432应统书院

令 $y=(\theta-108.53)/2.5727$，则 $\theta=108.53+2.5727y$.

$$R(d_1|x_1)=\int_{110}^{\infty}(\theta-110)h(\theta|x_1)\mathrm{d}\theta$$

$$=\frac{1}{\sqrt{2\pi}}\int_{0.5714}^{\infty}(2.5727y-1.47)\mathrm{e}^{-y^2/2}\mathrm{d}y$$

$$=0.4546.$$

$$R(d_2|x_1)=\int_{90}^{110}(110-\theta)h(\theta|x_1)\mathrm{d}\theta$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-7.2025}^{0.5714}(1.47-2.5727y)\mathrm{e}^{-y^2/2}\mathrm{d}y$$

$$=1.9246.$$

按最小后验期望损失准则选优决策，由于 $R(d_1|x_1)$ 小于 $R(d_2|x_1)$，于是接受 H_1 的假设.

例 10.3.3 产品交货检验方案的决策.

某企业的某种产品每 100 件装成一箱，销售部门将每箱产品运交给客户（买方）时，提出两种交货检验方案：

a_1 方案是指每箱全数检验，每件产品的检验费 0.6 元，则每箱检验费为 60 元；

a_2 方案是指每箱全不检查，交货给客户，但对客户作出承诺：若客户发现不合格品，不仅替换为合格品，而且每件不合格品的赔偿费为 12 元，则每箱的赔偿费为 $1\,200\theta$，其中 θ 为每箱产品的不合格率（$0\leqslant\theta\leqslant1$）.

该企业对每箱产品支付的费用函数为

$$W_1(\theta,d)=\begin{cases}60, & d=a_1;\\ 1\,200\theta, & d=a_2.\end{cases}\tag{10.3.9}$$

d 表示决策采用哪个方案. 当 $\theta=\theta_0=0.05$ 时，每箱产品的赔偿费等于每箱产品的检验费都为 60 元. 当 $\theta<\theta_0$ 时，a_2 方案为优；$\theta>\theta_0$ 时，a_1 方案为优. θ 是变量，试问该企业采用 a_1 方案还是 a_2 方案？

（一）先验期望损失准则

专业课辅导+真题资料+微信公众号：432应统书院

(1) $\boldsymbol{\theta}$ 的先验分布 $\pi(\theta)$.

该企业质量检验部门多次记录发现,每箱产品不合格率 θ 在 0.01 至 0.10 之间取值,于是不妨假定 $\boldsymbol{\theta}$ 在[0.01,0.10]区间内服从均匀分布.

(2) a_i方案的损失函数 $L_1(\theta,a_i)$.

由检验费用函数 $W_1(\theta,d)$换算得到 $L_1(\theta,a_i)$:

$$L_1(\theta,a_1)=\begin{cases}60-1\ 200\theta,\theta\leqslant\theta_0=0.05;\\ 0,\qquad \theta>\theta_0.\end{cases}\tag{10.3.10}$$

$$L_1(\theta,a_2)=\begin{cases}0,\qquad \theta\leqslant\theta_0=0.05;\\ 1\ 200\theta-60,\theta>\theta_0.\end{cases}\tag{10.3.11}$$

(3) 先验期望损失 $R_1(a_i)$.

a_i方案的先验期望损失为

$$R_1(a_i)=E\{L_1(\boldsymbol{\theta},a_i)\}=\int_{0.01}^{0.10}L_1(\theta,a_i)\pi(\theta)\mathrm{d}\theta.\tag{10.3.12}$$

于是

$$R_1(a_1)=\frac{1}{0.09}\int_{0.01}^{0.05}(60-1\ 200\theta)\mathrm{d}\theta=10.666\ 7(\text{元/箱}).$$

$$R_1(a_2)=\frac{1}{0.09}\int_{0.05}^{0.10}(1\ 200\theta-60)\mathrm{d}\theta=16.666\ 7(\text{元/箱}).$$

由于 $R_1(a_1)<R_1(a_2)$,即每箱产品的检验费小于每箱产品的赔偿费,于是采用 a_1方案.

上述结论是有客观意义的. 因为 $E(\boldsymbol{\theta})=0.055$ 大于 $\theta_0=0.05$,不敢贸然采用 a_2方案. 当然若每件不合格品的赔偿费由 12 元降低为 10 元,$R_1(a_2)$有可能小于 $R_1(a_1)$.

(二) 后验风险最小准则

设交货检验方案调整为这样:在每箱产品中随机抽取两件进行合格与否的检验,视不合格品个数多少才确定采用 a_1方案或 a_2方案. 现用 ξ 表示其中不合格品个数,则 ξ 服从二项分布 $B(2,\theta)$:

专业课辅导+真题资料+微信公众号：432应统书院

$$f(x|\theta)=C_2^x\theta^x(1-\theta)^{2-x},x=0,1,2.$$

ξ 与 $\boldsymbol{\theta}$ 的联合分布为

$$f(x,\theta)=\frac{1}{0.09}C_2^x\theta^x(1-\theta)^{2-x},0.01\leqslant\theta\leqslant0.10,x=0,1,2.$$

(1) $\boldsymbol{\theta}$ 的后验分布 $h(\theta|x)$.

由(10.2.5)式知，$\boldsymbol{\theta}$ 的后验分布为

$$h(\theta|x)=f(x,\theta)/g(x),x=0,1,2.$$

ξ 的边缘分布为

$$g(x)=\frac{1}{0.09}\int_{0.01}^{0.10}C_2^x\theta^x(1-\theta)^{2-x}d\theta,x=0,1,2.$$

具体数据如下表：

x	0	1	2
$g(x)$	0.893 7	0.120 6	0.003 7

则 $\boldsymbol{\theta}$ 的后验分布分别为：$0.01\leqslant\theta\leqslant0.10$,

$$h(\theta|x=0)=12.432\ 7(1-\theta)^2;$$

$$h(\theta|x=1)=216.590\ 9\ \theta(1-\theta);$$

$$h(\theta|x=2)=3\ 003.003\ \theta^2.$$

(2) 决策函数 $d(x)$.

据 ξ 的取值 x 来决策采用 a_1 方案或 a_2 方案，共能建立八个决策函数 $d_1(x),d_2(x),\cdots,d_8(x)$（列表如表 10.3.1）：

表 10.3.1　决策函数表

x	$d_1(x)$	$d_2(x)$	$d_3(x)$	$d_4(x)$	$d_5(x)$	$d_6(x)$	$d_7(x)$	$d_8(x)$
0	a_1	a_1	a_1	a_1	a_2	a_2	a_2	a_2
1	a_1	a_1	a_2	a_2	a_1	a_1	a_2	a_2
2	a_1	a_2	a_1	a_2	a_1	a_2	a_1	a_2

专业课辅导+真题资料+微信公众号：432应统书院

试问该选用哪个决策函数？

(3) a_i方案的损失函数 $L_2(\theta,a_i)$.

a_1方案每箱产品的检验费仍为60元，每箱产品的赔偿费用要作调整：

$$W_2(\theta,d)=\begin{cases}60, & d=a_1;\\ 1.2+1\ 176\theta, & d=a_2.\end{cases} \tag{10.3.13}$$

当 $\theta=\theta_0=0.05$ 时，每箱产品的费用都为60元.

$$L_2(\theta,a_1)=\begin{cases}58.8-1\ 176\ \theta, & \theta\leqslant\theta_0;\\ 0, & \theta>\theta_0.\end{cases} \tag{10.3.14}$$

$$L_2(\theta,a_2)=\begin{cases}0, & \theta\leqslant\theta_0;\\ 1\ 176\theta-58.8, & \theta>\theta_0.\end{cases} \tag{10.3.15}$$

(4) a_i方案的后验风险 $R_2(a_i|x)$.

$$R_2(a_i|x)=\int_{0.01}^{0.10}L_2(\theta,a_i)h(\theta|x)\mathrm{d}\theta. \tag{10.3.16}$$

其中

$$R_2(a_1|x)=\int_{0.01}^{0.05}(58.8-1\ 176\theta)h(\theta|x)\mathrm{d}\theta;$$

$$R_2(a_2|x)=\int_{0.05}^{0.10}(1\ 176\theta-58.8)h(\theta|x)\mathrm{d}\theta.$$

于是

$$R_2(a_1|x=0)=12.432\ 7\int_{0.01}^{0.05}(58.8-1\ 176\theta)(1-\theta)^2\mathrm{d}\theta;$$

$$R_2(a_2|x=0)=12.432\ 7\int_{0.05}^{0.10}(1\ 176\theta-58.8)(1-\theta)^2\mathrm{d}\theta;$$

$$R_2(a_1|x=1)=216.590\ 9\int_{0.01}^{0.05}(58.8-1\ 176\theta)\theta(1-\theta)\mathrm{d}\theta;$$

$$R_2(a_2|x=1)=216.590\ 9\int_{0.05}^{0.10}(1\ 176\theta-58.8)\theta(1-\theta)\mathrm{d}\theta;$$

$$R_2(a_1|x=2)=3\ 003.003\int_{0.01}^{0.05}(58.8-1\ 176\theta)\theta^2\mathrm{d}\theta;$$

专业课辅导+真题资料+微信公众号：432应统书院

$$R_2(a_2 \mid x=2)=3\ 003.003\int_{0.05}^{0.10}(1\ 176\theta-58.8)\theta^2\mathrm{d}\theta.$$

表 10.3.2　a_i方案的后验风险值表

x	0	1	2
$R_2(a_1 \mid x)$	11.166 9	4.594 5	1.786 9
$R_2(a_2 \mid x)$	15.371 6	13.834 6	31.268 8

从表 10.3.2 看到，不论 x 取何值，$R_2(a_1 \mid x)$ 都小于 $R_2(a_2 \mid x)$，即应选用 a_1 方案，$d_1(x)$ 为最优决策函数.

哪个决策函数为最优，同交货检验的费用函数 $W(\theta,d)$ 及不合格率 θ 都有关.

若 a_2 方案中对客户的承诺调整为：每件不合格品的赔偿费由 12 元下降为 7.5 元，费用函数为

$$W_3(\theta,d)=\begin{cases}60, & d=a_1;\\ 1.2+735\theta, & d=a_2.\end{cases}\tag{10.3.17}$$

通过计算得到表 10.3.3.

表 10.3.3　a_i方案的后验风险值表

x	0	1	2
$R_3(a_1 \mid x)$	48.452 7	12.461 5	7.480 6
$R_3(a_2 \mid x)$	29.028 6	2.693 7	3.853 3

从表 10.3.3 看到，不论 x 取何值，$R_3(a_2 \mid x)$ 都小于 $R_3(a_1 \mid x)$，即采用 a_2 方案，$d_8(x)$ 为最优决策函数.

若 a_2 方案对客户的承诺调整为：每件不合格品的赔偿费由 12 元下降为 10 元，费用函数为

$$W_4(\theta,d)=\begin{cases}60, & d=a_1;\\ 1.2+980\theta, & d=a_2.\end{cases}\tag{10.3.18}$$

则最优决策函数为

专业课辅导+真题资料+微信公众号：432应统书院

$$d_5(x)=\begin{cases}a_2, x=0;\\ a_1, x=1,2.\end{cases}$$

就是说在每箱产品中随机抽取两件作合格与否的检查，当两件都为合格品时，则采用 a_2 方案；否则，应采用 a_1 方案，继续逐件进行检验.

习　题

1. 设总体 ξ 服从正态分布 $N(a,\sigma)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为随机样本，若 a 有先验分布 $N(0,\sigma_0)$，σ_0 为已知数值，试求 a 的后验分布（用充分统计量 $\bar{\xi}$ 来表示）.

2. 设总体 ξ 服从指数分布，即 $f(x)=\lambda e^{-\lambda x}$，$x>0$，$\lambda>0$，损失函数为二次式，$\lambda$ 的先验分布为 $\Gamma(\alpha,\beta)$，即

$$g(\lambda)=\frac{\beta^{-(\alpha+1)}}{\Gamma(\alpha+1)}\lambda^{\alpha}e^{-\frac{\lambda}{\beta}},\lambda>0,\beta>0,\alpha+1>0.$$

试求参数 λ 的贝叶斯估计.

3. 设总体 ξ 服从正态 $N(a,1)$，$-\infty<a<\infty$，ξ_1,ξ_2 为其样本，考虑下述三个估计量

$$\hat{a}_1=d_1(\xi_1,\xi_2)=\frac{2}{5}\xi_1+\frac{3}{5}\xi_2;$$

$$\hat{a}_2=d_2(\xi_1,\xi_2)=\frac{5}{6}\xi_1+\frac{1}{6}\xi_2;$$

$$\hat{a}_3=d_3(\xi_1,\xi_2)=\frac{1}{3}\xi_1+\frac{2}{3}\xi_2.$$

定义损失函数为：$L(a,\hat{a})=3a^2(a-\hat{a})^2$，

试求 $R(a,d_1)$，$R(a,d_2)$，$R(a,d_3)$，问哪个风险为最小？

4. 设总体 ξ 服从 $N(0,\sigma)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为样本，建立 $\hat{\sigma}^2=c\sum_{i=1}^{n}\xi_i^2$，$c$ 为任一实数. 选定损失函数

$$L[\sigma^2,\hat{\sigma}^2]=\left(1-\frac{\hat{\sigma}^2}{\sigma^2}\right)^2,\quad \sigma>0,$$

试求 $\hat{\sigma}^2$ 的风险函数 $R[\sigma^2,\hat{\sigma}^2]$，试问 c 取何值，风险函数达到最小值.

专业课辅导+真题资料+微信公众号：432应统书院

5. 设总体 ξ 服从指数分布：

$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta}e^{-x/\theta}, & x\geqslant 0,\theta>0\\ 0, & x<0,\end{cases}$$

$\xi_1,\xi_2,\cdots,\xi_n$ 为样本，建立 $\hat{\theta}=c\bar{\xi}$，c 为任一实数，$\bar{\xi}$ 为样本均值. 选定损失函数

$$L[\theta,\hat{\theta}]=\left(1-\frac{\hat{\theta}}{\theta}\right)^2,$$

试求 $\hat{\theta}$ 的风险函数 $R[\theta,\hat{\theta}]$，试问 c 取何值，风险函数达到最小值.

6. 设总体 ξ 服从两点分布，$\xi_1,\xi_2,\cdots,\xi_n$ 为样本，

$$P\{\xi=x\}=p^x(1-p)^{1-x},\ x=0,1,\ 0<p<1.$$

p 的先验分布为：$\pi(p)=6p(1-p)$，$0<p<1$. 选定损失函数为 $L(p,\hat{p})=(p-\hat{p})^2$，试求 p 的贝叶斯估计.

7. 设总体 ξ 服从泊松分布，$\xi_1,\xi_2,\cdots,\xi_n$ 为样本，

$$P\{\xi=x\}=\frac{\lambda^x}{x!}e^{-\lambda},\ x=0,1,2,\cdots.\ \lambda>0.$$

λ 的先验分布为：$\pi(\lambda)=e^{-\lambda}$，$\lambda>0$. 选定损失函数为 $L(\lambda,\hat{\lambda})=(\lambda-\hat{\lambda})^2$，试求 λ 的贝叶斯估计.

8. 检查某设备零件的质量状况. 在设备工作时，用手摸零件的温度判别是否正常，温度正常用 $x=1$ 表示，温度异常（发烫）用 $x=0$ 表示. 零件的温度是否正常同零件的质量状况有关. 零件的质量状况有两种可能状态：θ_1 表示好，θ_2 表示坏. 总体 ξ 服从两点分布：

$$P\{\xi=x\}=\begin{cases}p, & x=1,\\ 1-p, & x=0,\end{cases}\quad 0<p<1$$

（注：θ_1 及 θ_2 的含义与 p_1 及 p_2 等同）. 假定 ξ 的概率函数为

$P(x;p)$	$x=1$	$x=0$
θ_1	0.8	0.2
θ_2	0.3	0.7

手摸一次判别零件的质量状况，表明样本容量 $n=1$. 当手摸判别零件的温度是否正常后，假定可能采取的行动有三个：a_1（保留）、a_2（更换）、a_3（修理）. 选定损失函数 $L(\theta,a)$ 为

专业课辅导+真题资料+微信公众号：432应统书院

$L(\theta,a)$	a_1	a_2	a_3
θ_1	0	8	3
θ_2	12	2	6

(1) 试问可能的决策函数 $d(\xi)$ 有哪几个?

(2) 试求决策函数 $d(\xi)$ 的风险函数 $R[\theta,d(\xi)]$.

[提示:$R[\theta,d(\xi)]=E\{L[\theta,d(\xi)]\}$

$=L(\theta,a_1)P_p\{d(\xi)=a_1\}+L(\theta,a_2)P_p\{d(\xi)=a_2\}+L(\theta,a_3)P_p\{d(\xi)=a_3\}$

于是

$R[\theta_1,d(\xi)]=8P_{p_1}\{d(\xi)=a_2\}+3P_{p_1}\{d(\xi)=a_3\}$

$R[\theta_2,d(\xi)]=12P_{p_2}\{d(\xi)=a_1\}+2P_{p_2}\{d(\xi)=a_2\}+6P_{p_2}\{d(\xi)=a_3\}$].

(3) 确定极大极小决策函数.

9. 在第 8 题中,假定 θ 有先验分布:

$$\pi(\theta_1)=0.7,\quad \pi(\theta_2)=0.3,$$

试求贝叶斯决策函数.[提示:决策函数 $d(\xi)$ 的贝叶斯风险为:

$$B[d(\xi)]=\pi(\theta_1)R[\theta_1,d(\xi)]+\pi(\theta_2)R[\theta_2,d(\xi)].$$

贝叶斯风险最小的决策函数称为贝叶斯决策函数.]

10. 设总体 ξ 服从几何分布:

$$P\{\xi=x\}=p(1-p)^{x-1},\ x=1,2,3,\cdots,\ 0<p<1$$

设参数 p 的先验分布为

$$\pi\left(p=\frac{i}{4}\right)=\frac{1}{3},\quad i=1,2,3.$$

设一次抽样得到样本观察值 $x=3$,试求参数 p 的后验分布,并求后验分布的期望与方差.[注:总体 ξ 及参数 p 的先验分布都为离散型,由 $\left(\xi=3,p=\frac{i}{4}\right)$ 的联合概率求出 $P\{\xi=3\}$,从而得出 $P\left\{p=\frac{i}{4}\middle|\xi=3\right\}$,$i=1,2,3$.]

11. 设总体 ξ 服从几何分布:

$$P\{\xi=x\}=p(1-p)^{x},\ x=0,1,2,3,\cdots,\ 0<p<1.$$

设参数 p 的先验分布为(0,1)上的均匀分布:

$$\pi(p)=1,\quad 0<p<1.$$

设一次抽样得到观察值 $x=3$,试求参数 p 的后验分布及其数学期望(也称贝

专业课辅导+真题资料+微信公众号：432应统书院

叶斯估计).

12. 设总体 ξ 服从 $N(a,\sigma)$，$\xi_1,\xi_2,\cdots,\xi_n$ 为样本，建立 σ^2 的三个估计量：

$$d_1(\xi_1,\xi_2,\cdots,\xi_n)=\frac{1}{n-1}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2,$$

$$d_2(\xi_1,\xi_2,\cdots,\xi_n)=\frac{1}{n+1}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2,$$

$$d_3(\xi_1,\xi_2,\cdots,\xi_n)=\frac{1}{n+2}\sum_{i=1}^{n}\xi_i^2.$$

假定损失函数为

$$L[\sigma^2,d(\xi_1,\xi_2,\cdots,\xi_n)]=\left(\frac{d}{\sigma^2}-1\right)^2.$$

试求它们的风险函数.

13. 设总体 ξ 服从两点分布：

$$P\{\xi=x\}=\begin{cases}p, & x=1,\\ 1-p, & x=0.\end{cases}$$

参数 $p\in\Omega=\left\{\frac{1}{4},\frac{1}{2}\right\}$. 样本容量 $n=1$，建立如下四个决策函数：

$$d_1(\xi)=\frac{1}{4},\quad x=0,1$$

$$d_2(\xi)=\begin{cases}\frac{1}{4}, & x=0,\\ \frac{1}{2}, & x=1,\end{cases}$$

$$d_3(\xi)=\begin{cases}\frac{1}{2}, & x=0,\\ \frac{1}{4}, & x=1,\end{cases}$$

$$d_4(\xi)=\frac{1}{2},\quad x=0,1.$$

选定损失函数为

$L(p,a)$	$a=\frac{1}{4}$	$a=\frac{1}{2}$
$p_1=\frac{1}{4}$	0	5
$p_2=\frac{1}{2}$	3	2

专业课辅导+真题资料+微信公众号：432应统书院

试求参数 p 的极大极小决策函数.

14. 在第 13 题中,假定参数 p 有先验分布:

$$\pi\left\{p=\frac{1}{4}\right\}=0.8,\quad \pi\left\{p=\frac{1}{2}\right\}=0.2,$$

试求参数 p 的贝叶斯估计.

15. 设总体 ξ 服从两点分布:

$$P\{\xi=x\}=\begin{cases}p, & x=1,\\ 1-p, & x=0,\end{cases}\quad 0<p<1,$$

设参数 p 在(0,1)上服从均匀分布,损失函数为

$$L(p,d)=\left[\frac{p-d}{p(1-p)}\right]^2.$$

$\xi_1,\xi_2,\cdots,\xi_n$ 为样本,试求 p 的贝叶斯估计.

16. 设总体 ξ 服从指数分布,$\xi_1,\xi_2,\cdots,\xi_n$ 为样本,

$$f(x;\lambda)=\begin{cases}\lambda e^{-\lambda x}, & x\geqslant 0,\lambda>0,\\ 0, & x<0.\end{cases}$$

参数 λ 的先验分布为 $\Gamma(\alpha,\beta)$ 分布:$\lambda<0$ 时,$\pi(\lambda;\alpha,\beta)=0$.

$$\pi(\lambda;\alpha;\beta)=\frac{\lambda^{\alpha}e^{-\lambda/\beta}}{\beta^{\alpha+1}\Gamma(\alpha+1)},\lambda\geqslant 0,\beta>0,\alpha+1>0.$$

选定损失函数 $L(\lambda-\hat{\lambda})^2=(\lambda-\hat{\lambda})^2$,试求 λ 的贝叶斯估计.

17. 设总体 ξ 在 $(0,\theta)$ 上均匀分布,$\xi_1,\xi_2,\cdots,\xi_n$ 为样本,参数 θ 在 $(0,a)$ 上服从均匀分布,$L(\theta,\hat{\theta})=(\theta-\hat{\theta})^2$,试求 θ 的贝叶斯估计.[提示:写出给定 θ 时,样本极大值 $\xi^*_{(n)}$ 的密度函数,求出 $(\theta,\xi^*_{(n)})$ 的联合密度函数.]

18. 设总体 ξ 服从 $N(a,1)$,$\xi_1,\xi_2,\cdots,\xi_{n_1}$ 为样本,总体 η 服从 $N(b,1)$,$\eta_1,\eta_2,\cdots,\eta_{n_2}$ 为样本.设这两个样本相互独立.用 d_0 表示判定"$a\leqslant b$",d_1 表示判定"$a>b$",建立决策函数为

$$\delta(\xi_1,\xi_2,\cdots,\xi_{n_1};\eta_1,\eta_2,\cdots,\eta_{n_2})=\begin{cases}d_0, & 当\bar{\xi}\leqslant\bar{\eta},\\ d_1, & 当\bar{\xi}>\bar{\eta}.\end{cases}$$

$\bar{\xi}$ 及 $\bar{\eta}$ 分别为样本均值.选定损失函数为

专业课辅导+真题资料+微信公众号：432应统书院

$L(\theta,\delta)$	$a\leqslant b$	$a>b$
d_0	0	$a-b$
d_1	$b-a$	0

θ 表示参数 a 及 b. 试求 $\delta(\bar{\xi};\bar{\eta})$ 的风险函数 $R(\theta,\delta)$. [提示：写出 $Z=\bar{\xi}-\bar{\eta}$ 的密度函数 $f(z;\theta)$，分别对于 $a\leqslant b$ 及 $a>b$ 两种场合，计算 $R(\theta,\delta)$，计算结果可用标准正态分布函数 $\Phi(x)$ 表示.]

专业课辅导+真题资料+微信公众号：432应统书院

第十一章　随机过程引论

§11.1　随机过程的概念

一、随机过程的直观背景和定义

在前面几章讨论的随机现象，基本上可由一个或有穷多个随机变量来描述. 在第一章开始，我们还指出，有些随机现象还必须研究它的发展变化过程. 因此，这一类随机现象若仅用一个或有穷多个随机变量去描述，就不能揭示其全部统计规律性，随机过程理论的诞生和发展，就是适应这一客观需要的.

下面举几个例子.

例 11.1.1　考虑某一电话站在正常工作条件下，某时刻 t 以前接到的呼叫次数 ξ，已知它是一个随机变量. 若我们考虑它随时间 t 变化的情形，则必须研究依赖于时间 t 的随机变量 ξ_t. 若以一天 24 小时计，则 t 从 0 变化到 24.

例 11.1.2　在数字通讯中，若传输过程是用数 0 和 1 两个码元通过编码来传递消息. 由于我们事先并不知道传送什么消息，因此，在某一时刻 t 传送的是 0 还是 1，都不能事先预言，因而是一随机变量. 若我们进行长期观察，每隔 1 秒观察一次，则这个随机变量(用 ξ_n 表示)依赖于 n(秒)，其中 $n=0,1,2,\cdots$.

例 11.1.3　在地震勘探工作中，我们通过检波器把混有随机干扰的随时间变动的地层结构信号波记录下来. 如图 11.1.1 所示：在 O 点放炮，在 A 点记录仪把接收到的混有干扰的地震信号波记录下来，我们在相同条件下做了 n 次记录，则得到 n 个彼此有差异的记录. 在时间 t_0 观察它们的信号波的值 ξ_{t_0}，则可以发现它是不规则的，即 ξ_{t_0} 是一个随机变量. 也就是说，混有随机干扰的地

专业课辅导+真题资料+微信公众号：432应统书院

层结构信号波是一个依赖于时间 t 的随机变量 ξ_t.

图 11.1.1

例 11.1.4 在例 11.1.3 中,时间 t 是连续变化的(如在 0 至 5 秒内连续记录信号波). 但在对地震信号波进行数字处理时,则需按一定的时间间隔 Δ(如 1 毫秒)取样. 这时,我们所研究的随机变量则依赖于 $n\Delta$(n 取非负整数). 固定 n_0 时,对应的 $\xi_{n_0\Delta}$ 为一随机变量.

总之,在实际问题中,不仅要求我们研究一个或有穷多个随机变量,而且要求我们研究一族无穷多个随机变量,我们将这样一族随机变量称为随机过程.

下面给出随机过程的正式定义:

定义 11.1.1 设给定概率空间 $(\Omega,\mathscr{F},P)$ 及指标集 T①,对每一 $t\in T$,有定义在 $(\Omega,\mathscr{F},P)$ 上的随机变量 $\xi_t(\omega)$,$\omega\in\Omega$,与之对应. 则称依赖于 t 的一族随机变量 ξ_t 为随机过程,用符号 $\{\xi_t(\omega);t\in T\}$ 表示,如不产生混乱,简写为 $\{\xi_t\}$. 其中指标 t 称为参数或时间,T 称为参数集.

① 一般 T 为如下的指标集:$T=\{0,1,\cdots\}$,或 $T=\{\cdots,-1,0,1,\cdots\}$,或 $T=[a,b]$ 或 $T=\mathbf{R}_1=(-\infty,\infty)$,$T=\mathbf{R}_1^+=[0,\infty)$. 在前面有时把 $\mathbf{R}_1$ 记为 $\mathbf{R}$

专业课辅导+真题资料+微信公众号：432应统书院

这一章将研究能用随机过程描述的随机现象，叙述随机过程最初步的基本概念，至于需要更深入学习的读者，可参看[10]，[14]，[17]，[19].

对于固定的时间 $t(t\in T)$，若随机变量 ξ_t 是离散型的，我们就说随机过程 $\{\xi_t\}$ 有一个离散状态空间，否则就说 $\{\xi_t\}$ 有一个非离散状态空间. 参数集 T 可以是离散的（如 $T=\{1,2,\cdots,n\}$ 或 $T=\{1,2,\cdots\}$）或连续的（如 $T=\{t:t\geqslant 0\}$ 或 $T=\mathbf{R}_1$）. 若参数集 T 是离散的，则称随机过程是一个离散参数随机过程或随机序列；若 T 是连续的，则称随机过程是一个连续参数随机过程.

因此，根据状态空间和参数集 T 的不同情况，可以把随机过程分成四类，列成下表 11.1.1：

表 11.1.1

随机过程 \ 状态空间 / 参数集	离　散	非　离　散
离　散	离散参数链	离散参数随机过程
连　续	连续参数链	连续参数随机过程

把随机过程 $\{\xi_t(\omega);t\in T\}$ 设想为一个作随机运动的质点 M，以 ξ_t 表示 M 在时刻 t 的位置，于是 $\{\xi_t(\omega);t\in T\}$ 描述了 M 所作的随机运动的变化过程，这时，把状态空间的每一点 x 称作状态，$\xi_t=x$ 就可形象地说成“在时刻 t 质点 M 处于状态 x”.

随机过程 $\{\xi_t(\omega);t\in T\}$ 事实上是参数 t 和概率空间中的点 ω 的二元函数. 对每一 $t_0\in T$，$\xi_{t_0}(\omega)$ 是定义在 $(\Omega,\mathscr{F},P)$ 上的随机变量. 对每一 $\omega_0\in\Omega$，$\xi_t(\omega_0)$ 是定义在 T 上的一个普通的函数（实的或复的，也可以是实向量或复向量），称为随机过程 $\{\xi_t\}$ 对应于 ω_0 的一个样本函数，有时也称为轨道或现实. 图 11.1.1 画出了 n 个样本函数. 如果样本函数对几乎所有的 ω（除了概率为 0 以外）为连续（或可微）的，则称样本函数几乎处处连续（或可微）.

专业课辅导+真题资料+微信公众号：432应统书院

随机过程$\{\xi_t\}$的参数t可能不一定是我们通常所指的时间这一物理量，而可能是别的量，也可以是多维的量，例如考虑某一海面的波浪的浪高随时间的变化情况，则这时需用四维的参数来描述（三维坐标x,y,z及时间t）. 对于参数为多维的随机过程$\{\xi_t\}$称为随机场.

二、随机过程的有穷维分布函数族

设$\{\xi_t\}$是一随机过程，对于每一$t\in T$，ξ_t的分布函数为$F_t(x)$，即

$$F_t(x)=P\{\xi_t<x\}.$$

当任意给定$t_1,t_2\in T$时，随机变量ξ_{t_1}和ξ_{t_2}的联合分布函数为$F_{t_1,t_2}(x_1,x_2)$，即

$$F_{t_1,t_2}(x_1,x_2)=P\{\xi_{t_1}<x_1,\xi_{t_2}<x_2\}.$$

一般地，当给定$t_1,t_2,\cdots,t_n\in T$及n时，随机变量$\xi_{t_1},\xi_{t_2},\cdots,\xi_{t_n}$的$n$维联合分布函数为

$$F_{t_1,t_2,\cdots,t_n}(x_1,x_2,\cdots,x_n)=P\{\xi_{t_1}<x_1,\xi_{t_2}<x_2,\cdots,\xi_{t_n}<x_n\}.$$

一切上述分布函数的全体

$$\{F_{t_1,t_2,\cdots,t_n}(x_1,x_2,\cdots,x_n):t_1,\cdots,t_n\in T,n>0\}$$

称为$\{\xi_t\}$的有穷维分布函数族，它描述随机过程$\{\xi_t\}$的统计规律.

如果$\{\xi_t\}$的任一有穷维分布函数作时间平移不改变，即对任意$n>0$及τ，有

$$F_{t_1+\tau,\cdots,t_n+\tau}(x_1,x_2,\cdots,x_n)=F_{t_1,\cdots,t_n}(x_1,x_2,\cdots,x_n),$$

则称$\{\xi_t\}$为强平稳过程（序列）.

读者不难验证，有穷维分布函数族满足下列条件：

（i）对每一$x_i(i=1,2,\cdots,n)$，$F_{t_1,t_2,\cdots,t_n}(x_1,x_2,\cdots,x_n)$非负、不降且左连续.

（ii）$\lim\limits_{x_i\to-\infty}F_{t_1,t_2,\cdots,t_n}(x_1,x_2,\cdots,x_n)=0\quad(i=1,2,\cdots,n)$，

$$\lim_{\substack{x_1\to\infty\\ \vdots\\ x_n\to\infty}}F_{t_1,t_2,\cdots,t_n}(x_1,x_2,\cdots,x_n)=1.$$

专业课辅导+真题资料+微信公众号：432应统书院

(iii) 若 $x_i \leqslant y_i (i=1,2,\cdots,n)$，则

$$F_{t_1,t_2,\cdots,t_n}(y_1,y_2,\cdots,y_n)-\sum_{i=1}^{n}F_{t_1,t_2,\cdots,t_n}(y_1,y_2,\cdots,y_{i-1},x_i,$$

$$y_{i+1},\cdots,y_n)+\sum_{\substack{i,j=1\\i<j}}^{n}F_{t_1,t_2,\cdots,t_n}(y_1,y_2,\cdots,y_{i-1},x_i,y_{i+1},\cdots,y_{j-1},x_j,$$

$$y_{j+1},\cdots,y_n)-\cdots+(-1)^nF_{t_1,t_2,\cdots,t_n}(x_1,x_2,\cdots,x_n)\geqslant 0.$$

以及称为相容性条件的两个条件：

(iv) 对 $(1,2,\cdots,n)$ 的任一排列 $(\alpha_1,\alpha_2,\cdots,\alpha_n)$，有

$$F_{t_1,t_2,\cdots,t_n}(x_1,x_2,\cdots,x_n)=F_{t_{\alpha_1},t_{\alpha_2},\cdots,t_{\alpha_n}}(x_{\alpha_1},x_{\alpha_2},\cdots,x_{\alpha_n}).$$

(v) 若 $m<n$，则

$$F_{t_1,t_2,\cdots,t_m}(x_1,x_2,\cdots,x_m)=\lim_{x_{m+1},\cdots,x_n\to\infty}F_{t_1,t_2,\cdots,t_n}(x_1,x_2,\cdots,x_n).$$

上述命题的逆命题，即所谓随机过程的存在性定理，是一个非常重要且有名的定理，我们把它叙述如下：

设已给参数集 T，对任意 n 及 T 的任意 n 个元素 $t_1,t_2,\cdots,t_n$，有一 n 元函数 $F_{t_1,t_2,\cdots,t_n}(x_1,x_2,\cdots,x_n)$ 与之对应，并且满足上述(i)，(ii)，(iii)，(iv)，(v)五个条件(我们称这族函数为满足相容性条件的有穷维分布函数族)，则必存在一随机过程 $\{\xi_t\}$，使得它的有穷维分布函数族等于所给定的函数族.

对上述定理证明有兴趣的读者可参看[10]或[19].

要确定随机过程的分布函数族，在实际工作中往往是十分困难的，与第三章研究一维或多维随机变量时类似，我们可以转而研究随机过程的某种特征，它在许多实际问题和理论问题中都能很好地满足研究的目的.

(1) $\{\xi_t\}$ 的数学期望与方差

设随机过程 $\{\xi_t\}$ 的一维分布函数为 $F_t(x)$. 仿第三章，随机过程 $\{\xi_t\}$ 的数学期望及方差分别定义为

$$E(\xi_t)=\int_{-\infty}^{\infty}x\mathrm{d}F_t(x) \tag{11.1.1}$$

$\left(当\int_{-\infty}^{\infty}|x|\mathrm{d}F_t(x)<\infty 时\right)$，及

专业课辅导+真题资料+微信公众号：432应统书院

$$D(\xi_t)=\int_{-\infty}^{\infty}[x-E(\xi_t)]^2\mathrm{d}F_t(x) \tag{11.1.2}$$

$\left(当\int_{-\infty}^{\infty}[x-E(\xi_t)]^2\mathrm{d}F_t(x)<\infty\right)$. 一般地说，$E(\xi_t)$ 与 $D(\xi_t)$ 都是 t 的函数.

当 $\{\xi_t\}$ 的状态空间是离散型时，则

$$E(\xi_t)=\sum_i x_i p_i(t), \tag{11.1.3}$$

$$D(\xi_t)=\sum_i[x_i-E(\xi_t)]^2 p_i(t), \tag{11.1.4}$$

其中 $p_i(t)$ 为 ξ_t 的概率函数，即 $p_i(t)=P\{\xi_t=x_i\}$.

若对任一 $t\in T$，ξ_t 为连续型随机变量时，则

$$E(\xi_t)=\int_{-\infty}^{\infty}xf_t(x)\mathrm{d}x, \tag{11.1.5}$$

$$D(\xi_t)=\int_{-\infty}^{\infty}[x-E(\xi_t)]^2 f_t(x)\mathrm{d}x, \tag{11.1.6}$$

其中 $f_t(x)$ 为 ξ_t 的密度函数.

(2) $\{\xi_t\}$ 的相关函数和协方差函数

设随机过程 $\{\xi_t\}$ 的二维分布函数为 $F_{t_1,t_2}(x_1,x_2)$，则 $\{\xi_t\}$ 的相关函数定义为：对任意 $t_1,t_2\in T$，

$$K(t_1,t_2)=E(\xi_{t_1}\xi_{t_2})=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x_1x_2\mathrm{d}F_{t_1,t_2}(x_1,x_2), \tag{11.1.7}$$

而把

$$\begin{aligned}R(t_1,t_2)&=E[\xi_{t_1}-E(\xi_{t_1})][\xi_{t_2}-E(\xi_{t_2})]\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[x_1-E(\xi_{t_1})][x_2-E(\xi_{t_2})]\mathrm{d}F_{t_1,t_2}(x_1,x_2)\end{aligned} \tag{11.1.8}$$

称为 $\{\xi_t\}$ 的协方差函数，有时亦称它为中心化相关函数. 自然，只有当(11.1.7)及(11.1.8)的积分绝对收敛时这两个式子才有定义.

专业课辅导+真题资料+微信公众号：432应统书院

相关函数与协方差函数之间有密切的关系：首先 $R(t_1,t_2)=K(t_1,t_2)-E(\xi_{t_1})E(\xi_{t_2})$；其次，$\{\xi_t\}$ 的协方差函数是随机过程 $\{\xi_t-E(\xi_t)\}$ 的相关函数，而 $\{\xi_t\}$ 的相关函数可考虑作为关于随机过程 $\{\xi_t e^{i\eta}\}$ 的协方差函数，其中 η 为均匀分布于 $(-\pi,\pi)$ 的随机变量，并且与 $\{\xi_t\}$ 独立. 上述结果，读者可自行验证.

相关函数 $K(t_1,t_2)$ 一般说是 t_1,t_2 的二元函数. 可以证明 $K(t_1,t_2)$ 是非负定的. 即对任意 $n(n=1,2,\cdots,)$ 及数 $\alpha_i(i=1,2,\cdots,n)$，

$$\sum_{i,j=1}^{n} K(t_i,t_j)\alpha_i\alpha_j \geqslant 0. \tag{11.1.9}$$

事实上，

$$\sum_{i,j=1}^{n} K(t_i,t_j)\alpha_i\alpha_j = E\left|\sum_{i=1}^{n}\xi_{t_i}\alpha_i\right|^2 \geqslant 0. \tag{11.1.10}$$

上述 $\{\xi_t\}$ 的数学期望、方差、相关函数和协方差函数，分别描述随机过程在随时间 t 变化时的平均数、离散度和相关情况. 对于相关函数，我们还将在平稳过程部分着重讨论.

例 11.1.5 设 $g(t)$ 为如图 11.1.2 所示的以周期为 L 的矩形波，η 为服从两点分布的随机变量，其分布列如下表：

η	1	-1
p_i	$\frac{1}{2}$	$\frac{1}{2}$

图 11.1.2

专业课辅导+真题资料+微信公众号：432应统书院

定义 ξ_t 如下：对任意 $t\in\mathbf{R}_1$，

$$\xi_t=\eta g(t),$$

即 $\{\xi_t\}$ 为具有随机振幅、周期为 L 的矩形波，则有

$$E(\xi_t)=E[\eta g(t)]=g(t)\left[\frac{1}{2}+(-1)\cdot\frac{1}{2}\right]=0,$$

$$D(\xi_t)=E[\eta g(t)]^2=g^2(t)E(\eta^2)=g^2(t);$$

$$\begin{aligned}K(t_1,t_2)&=E[\eta g(t_1)\eta g(t_2)]\\&=g(t_1)g(t_2)E(\eta^2)=g(t_1)g(t_2).\end{aligned}$$

例 11.1.6 设 ξ,η 为相互独立、数学期望为 0、方差为 1 的随机变量，令

$$\xi_t=\xi+\eta t,$$

则

$$E(\xi_t)=E(\xi)+tE(\eta)=0;$$

$$D(\xi_t)=D(\xi+\eta t)=D(\xi)+t^2D(\eta)=1+t^2;$$

$$\begin{aligned}K(t_1,t_2)&=E(\xi_{t_1}\xi_{t_2})\\&=E(\xi^2)+t_1t_2E(\eta^2)+(t_1+t_2)E(\xi)E(\eta)\\&=1+t_1t_2.\end{aligned}$$

例 11.1.7① 随机过程 $\{\xi_t\}$ 定义如下：

考虑"随机点"发生在区间 $(0,t]$ 内的次数，若"随机点"在 $(0,t]$ 内发生的次数是偶数（把 0 也算作偶数），则令 $\xi_t=1$，若为奇数，则令 $\xi_t=-1$，如图 11.1.3 所示. 又设在时间区间 $(t_0,t_0+t]$ 内有 k 个"随机点"发生的概率与 t_0 无关，并且为

$$p_k(t)=\mathrm{e}^{-\lambda t}\frac{(\lambda t)^k}{k!}\quad(t\in\mathbf{R}_1)$$

其中 $\lambda>0,k=0,1,2,\cdots$，则有

$$P\{\text{在}(0,t]\text{内发生偶数次"随机点"}\}=p_0(t)+p_2(t)+\cdots$$

---

① 在此例中所述之"随机点"，在许多概率模型中，可以概括相当广泛的实际背景：电话总机接到的"呼叫"、数字编码通讯中出现的"误码"……都可视为"随机点"，在下一节我们还要进行讨论.

专业课辅导+真题资料+微信公众号：432应统书院

图 11.1.3

$$=e^{-\lambda t}\left\{1+\frac{(\lambda t)^2}{2!}+\frac{(\lambda t)^4}{4!}+\cdots\right\}=e^{-\lambda t}\cosh\lambda t,$$

$$P\{\text{在}(0,t]\text{内发生奇数次“随机点”}\}=p_1(t)+p_3(t)+\cdots$$

$$=e^{-\lambda t}\left\{\lambda t+\frac{(\lambda t)^3}{3!}+\frac{(\lambda t)^5}{5!}+\cdots\right\}=e^{-\lambda t}\sinh\lambda t,$$

即

$$P\{\xi_t=1\}=e^{-\lambda t}\cosh\lambda t;$$

$$P\{\xi_t=-1\}=e^{-\lambda t}\sinh\lambda t.$$

故得

$$\begin{aligned}E(\xi_t)&=1\cdot e^{-\lambda t}\cosh\lambda t+(-1)e^{-\lambda t}\sinh\lambda t\\&=e^{-\lambda t}(\cosh\lambda t-\sinh\lambda t)=e^{-\lambda t}e^{-\lambda t}\\&=e^{-2\lambda t}.\end{aligned}$$

下面求$\{\xi_t\}$的相关函数和方差，首先求ξ_{t_1},ξ_{t_2}的联合概率函数$P\{\xi_{t_1}=x_1,\xi_{t_2}=x_2\}$，其中$t_1,t_2\in\mathbf{R}_1$；$x_i=1$或$-1$，$(i=1,2)$.

因为$P\{\xi_{t_1}=x_1,\xi_{t_2}=x_2\}=P\{\xi_{t_1}=x_1\}P\{\xi_{t_2}=x_2\mid\xi_{t_1}=x_1\}$，因而只需求$P\{\xi_{t_1}=x_1\}$及$P\{\xi_{t_2}=x_2\mid\xi_{t_1}=x_1\}$便可.

设$t_2>t_1$，令$t_2-t_1=r$，则可得

$P\{\xi_{t_2}=1\mid\xi_{t_1}=1\}=P\{$在$(0,t_1]$内“随机点”发生偶数次的条件下，在$(0,t_2]$内仍然发生偶数次$\}=P\{$在$(t_1,t_2]$内“随机点”发生偶数次$\}=P\{$在$(0,r]$内“随机点”发生偶数次$\}=e^{-\lambda r}\cosh\lambda r$，

专业课辅导+真题资料+微信公众号：432应统书院

而

$$P\{\xi_{t_1}=1\}=e^{-\lambda t_1}\cosh\lambda t_1,$$

故

$$P\{\xi_{t_1}=1,\xi_{t_2}=1\}=(e^{-\lambda t_1}\cosh\lambda t_1)(e^{-\lambda r}\cosh\lambda r).$$

同理可得

$$P\{\xi_{t_1}=-1,\xi_{t_2}=-1\}=(e^{-\lambda t_1}\sinh\lambda t_1)(e^{-\lambda r}\cosh\lambda r);$$

$$P\{\xi_{t_1}=-1,\xi_{t_2}=1\}=(e^{-\lambda t_1}\sinh\lambda t_1)(e^{-\lambda r}\sinh\lambda r);$$

$$P\{\xi_{t_1}=1,\xi_{t_2}=-1\}=(e^{-\lambda t_1}\cosh\lambda t_1)(e^{-\lambda r}\sinh\lambda r).$$

这样一来，可求得

$$\begin{aligned}K(t_1,t_2)&=E(\xi_{t_1}\xi_{t_2})\\&=(e^{-\lambda t_1}\cosh\lambda t_1)(e^{-\lambda r}\cosh\lambda r)\\&\quad+(-1)(-1)(e^{-\lambda t_1}\sinh\lambda t_1)(e^{-\lambda r}\cosh\lambda r)\\&\quad+(-1)(e^{-\lambda t_1}\sinh\lambda t_1)(e^{-\lambda r}\sinh\lambda r)\\&\quad+(-1)(e^{-\lambda t_1}\cosh\lambda t_1)(e^{-\lambda r}\sinh\lambda r)\\&=e^{-\lambda(t_1+r)}(\cosh\lambda t_1\cosh\lambda r-\sinh\lambda t_1\sinh\lambda r\\&\quad+\sinh\lambda t_1\cosh\lambda r-\cosh\lambda t_1\sinh\lambda r)\\&=e^{-\lambda(t_1+r)}[\cosh\lambda(r-t_1)-\sinh\lambda(r-t_1)]\\&=e^{-\lambda(t_1+r)}e^{-\lambda(r-t_1)}=e^{-2\lambda r}.\end{aligned}$$

即得

$$K(t_1,t_2)=e^{-2\lambda(t_2-t_1)}.$$

同理可得，当 $t_2<t_1$ 时，

$$K(t_1,t_2)=e^{2\lambda(t_2-t_1)}.$$

因而，对任意 $t_1,t_2\in\mathbf{R}_1$，

$$K(t_1,t_2)=e^{-2\lambda|t_2-t_1|}.$$

由于

$$D(\xi_t)=K(t,t)-[E(\xi_t)]^2,$$

而 $K(t,t)=E(\xi_t^2)=1$，故 $\{\xi_t\}$ 的方差为

$$D(\xi_t)=1-(e^{-2\lambda t})^2=1-e^{-4\lambda t}.$$

专业课辅导+真题资料+微信公众号：432应统书院

前面已说及，随机过程的数学期望与方差一般说来都是 t 的函数，而相关函数则是依赖于 t_1, t_2 的二元函数，但有一些随机过程的数学期望不依赖于 t 而为一常数，例如例 11.1.5 及例 11.1.6；其相关函数只依赖于 t_2-t_1，如例 11.1.7. 即时间的平移不改变上述两个特征的数值. 具有这一特性的随机过程，在研究飞机受湍流影响而产生随机振动、棉纱截面积的随机大小、数字通信中的某些已编码信号等随机现象时起重要作用. 特别对于相关函数具有上述特性的随机过程，称它为弱平稳过程，我们还要着重讨论，详细情况见 §11.6.

§11.2 几类重要的随机过程简介

由于理论研究和应用的不断深入与扩大，各种类型的随机过程很多，讨论得也很深入，我们不可能在这本书中尽述，这里仅介绍在各种自然现象的预测预报、无线电技术、自动控制和通信等方面常用的一些随机过程.

一、独立增量过程（可加过程）

设 $\{\xi_t\}$ 为一随机过程，若对任意正整数 n 及 $t_i \in T, i=1,2,\cdots,n, t_0<t_1<t_2<\cdots<t_n$，随机变量

$$\xi_{t_0}, \xi_{t_1}-\xi_{t_0}, \xi_{t_2}-\xi_{t_1}, \xi_{t_3}-\xi_{t_2}, \cdots, \xi_{t_n}-\xi_{t_{n-1}}$$

相互独立，则称 $\{\xi_t\}$ 为独立增量过程，又称为可加过程，其中 t_0 为初始时刻.

例如：某电话总机在时间区间 $[0,t]$ 内接到的呼叫次数设为 ξ_t，并设 $P\{\xi_0=0\}=1$，则 $\{\xi_t; t\in[0,\infty)\}$ 是一独立增量过程，因为它在不相重叠的时间区间内发生呼叫的次数可以认为是相互独立的.

从定义我们看到，若 $\{\xi_t\}$ 是独立增量过程，则对任意 t，ξ_t 是相互独立随机变量之和. 事实上，对任一时刻 $t, t_0<t_1<\cdots<t_{n-1}<t$，$\xi_t=\xi_{t_0}+\sum_{i=1}^{n}(\xi_i-\xi_{i-1})$.

专业课辅导+真题资料+微信公众号：432应统书院

二、正态随机过程(高斯过程)

设随机过程$\{\xi_t\}$对任意正整数n及$t_1,t_2,\cdots,t_n,t_i\in T(i=1,2,\cdots,n)$,随机变量$\xi_{t_1},\xi_{t_2},\cdots,\xi_{t_n}$的联合分布函数为$n$维正态分布,即有如下的密度函数

$$f(x_1,x_2,\cdots,x_n)=\frac{|\boldsymbol{\Sigma}|^{-\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}}e^{-\frac{1}{2}(x-a)'\boldsymbol{\Sigma}^{-1}(x-a)},\tag{11.2.1}$$

则称$\{\xi_t\}$为正态随机过程,其中

$$\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\boldsymbol{a}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix},\boldsymbol{\Sigma}=\begin{pmatrix}\sigma_{11}&\sigma_{12}&\cdots&\sigma_{1n}\\\sigma_{21}&\sigma_{22}&\cdots&\sigma_{2n}\\\vdots&\vdots&&\vdots\\\sigma_{n1}&\sigma_{n2}&\cdots&\sigma_{nn}\end{pmatrix}.$$

$\boldsymbol{\Sigma}$为对称非负定矩阵,$\boldsymbol{\Sigma}^{-1}$表示$\boldsymbol{\Sigma}$的逆矩阵,$|\boldsymbol{\Sigma}|$表示$\boldsymbol{\Sigma}$的行列式,$(\quad)'$表示向量或矩阵的转置.

特别当$n=2,\boldsymbol{\Sigma}=\begin{pmatrix}\sigma_1^2&r\sigma_1\sigma_2\\r\sigma_1\sigma_2&\sigma_2^2\end{pmatrix}$时,(11.2.1)式即为二维正态密度函数.

在无线电技术研究和应用中,真空管中的散粒效应(shot effect),可以用正态随机过程来描述.

三、维纳过程

设随机过程$\{\xi_t;t\in[t_0,\infty)\}$满足下列条件:

(i) $\{\xi_t\}$是平稳独立增量,即$\{\xi_t\}$是独立增量过程,且对任意$t_1,t_2\in[t_0,\infty)$,$t_1<t_2$,及$h>0$,$\xi_{t_2+h}-\xi_{t_1+h}$有相同分布函数;

(ii) 对任意$t_1,t_2\in[t_0,\infty)$,$\xi_{t_1}-\xi_{t_2}$服从正态

$$N(0,\sqrt{|t_1-t_2|});$$

(iii) $P\{\xi_0=0\}=1$,

则称$\{\xi_t\}$为维纳过程.

维纳过程可以作为质点的布朗运动,以及电子电路中的理论

专业课辅导+真题资料+微信公众号：432应统书院

噪声的一个很好的数学模型. 还可用以研究股票和商品市场中价格变动模型,并且在通讯和控制理论中经常用到它.

维纳过程亦称为维纳-莱维过程或布朗运动过程. 此过程名字的由来是因为 Wiener 在 1923 年用数学方法研究布朗运动(质点在液体或气体中不停地作不规则的一种运动——它由 Brown 所发现)而产生的.

对于维纳过程,我们有:

$$E\xi_t = E(\xi_t - \xi_0) = 0;$$

当 $t_1 < t_2$ 时

$$\begin{aligned} B(t_1, t_2) = E(\xi_{t_1}\xi_{t_2}) &= E[\xi_{t_1}(\xi_{t_2} - \xi_{t_1}) + \xi_{t_1}^2] \\ &= E(\xi_{t_1} - \xi_0)(\xi_{t_2} - \xi_{t_1}) + E\xi_{t_1}^2 \\ &= 0 + E(\xi_{t_1} - \xi_0)^2 = \sigma^2 t_1. \end{aligned}$$

同理可证当 $t_1 \geqslant t_2$ 时有

$$B(t_1, t_2) = \sigma^2 t_2.$$

维纳过程是正态过程. 事实上,若 $\{\xi_t\}$ 是维纳过程,则对任意 n 及 $t_1 < t_2 < \cdots < t_n$,令 $\eta_m = \xi_{t_m} - \xi_{t_{m-1}}$, $m = 1, 2, \cdots, n$,则 $(\eta_1, \eta_2, \cdots, \eta_n)$ 是相互独立且服从正态分布的 n 维正态向量. 由 $\eta_1, \eta_2, \cdots, \eta_n$ 的定义可知:

$$\xi_{t_1} = \eta_1, \xi_{t_2} = \eta_1 + \eta_2, \cdots, \xi_{t_n} = \eta_1 + \eta_2 + \cdots + \eta_n,$$

即 $\xi_{t_1}, \cdots, \xi_{t_n}$ 是 $\eta_1, \eta_2, \cdots, \eta_n$ 的线性组合,对任意 n,知 $(\xi_{t_1}, \xi_{t_2}, \cdots, \xi_{t_n})$ 是 n 维正态向量,由正态随机过程的定义可知 $\{\xi_t\}$ 是正态过程.

四、泊松过程

设随机过程 $\{\xi_t; t \in [t_0, \infty)\}$ $(t_0 \geqslant 0)$ 的状态只取非负整数值,若满足下列三个条件:

(i) $P\{\xi_{t_0} = 0\} = 1$;

(ii) $\{\xi_t\}$ 为平稳独立增量过程;

(iii) 对任意 $t_1, t_2 \in [t_0, \infty)$, $t_1 < t_2$,对应的随机变量的增量

专业课辅导+真题资料+微信公众号：432应统书院

$\xi_{t_2}-\xi_{t_1}$服从数学期望为$\lambda(t_2-t_1)$的泊松分布，即对于$k=0,1,\cdots$有

$$P\{\xi_{t_2}-\xi_{t_1}=k\}=e^{-\lambda(t_2-t_1)}\frac{[\lambda(t_2-t_1)]^k}{k!};$$

$$E(\xi_{t_2}-\xi_{t_1})=\lambda(t_2-t_1);D(\xi_{t_2}-\xi_{t_1})=\lambda(t_2-t_1),$$

则称$\{\xi_t\}$为(齐次)泊松过程.

对于公用事业、通信等许多实际问题，都可以用泊松过程来描述，因为这些实际问题都可满足下面将要叙述的定理11.2.1的四个条件. 例如某电话总机在时间区间$[0,t]$内接到的呼叫次数、某服务窗口排队的人数、数字通信中已编码信号的误码个数等，都能很好地满足所说四个条件，因而都可看成是泊松过程.

五、随机点过程与计数过程

在例11.1.7中，我们提到了"随机点"这一概念，这里我们稍作介绍.

在客观世界中，存在着这样一类随机现象，它们发生的时间、地点或相联系的某些属性，常可归结为某一空间$\mathscr{H}$中的点的随机发生，这种点的发生就构成所谓随机点过程.

例11.2.1 一天中某电话总机接到的呼叫形成一个随机点过程. $\mathscr{H}$为时间区间$[0,24]$，每一次呼叫发生的时间就是一个随机点，这个点过程的一个现实(样本函数)是一时间的序列$\{t_1,t_2,\cdots,t_N\}$，其中$0\leqslant t_1<t_2\cdots<t_N\leqslant 24$，$t_n$表示第$n$次呼叫发生的时间，$N$则表示一天内呼叫发生的次数.

例11.2.2 某一台机器在时刻t_0开始工作，我们在时间区间$[t_0,t_0+T]$内考虑机器发生故障而停止工作这一事件. 假设当机器发生故障后立即修理或换上新的同样的零件后继续工作，则机器在时间区间$[t_0,t_0+T]$内发生故障而停止工作这个事件形成一个随机点过程.

例11.2.3 设从时间t_0开始的长为T的时间区间内考虑某地区发生地震这一事件(在一定的震级范围内，每一次地震的震中、发震时间)，它可以根据各台、站的观察结果记录下来. 由于震

专业课辅导+真题资料+微信公众号：432应统书院

中的广度最多几千米，而震动的持续时间一般以秒计，这相对于我们研究的整个地区和时间范围来说是很小的，所以可以理想化为一点. 这样，每一地震发生这事件，可用一个三维向量(x,y,t)来表示，其中x,y表示震中的经度和纬度，t表时间. 描述地震发生的事件序列$(x_1,y_1,t_1),(x_2,y_2,t_2),\cdots,(x_N,y_N,t_N)$构成一多维随机点过程.

由上所述，随机点过程是一种类型的随机过程，人们关心的首先是代表随机事件发生的时间、地点或状态的点的统计规律，这些统计规律包括如下两个方面：

(1) 计数性质：考虑这种随机的点落在$\mathscr{H}$的某一部分的个数，当$\mathscr{H}$是时间轴$[0,\infty)$时，令N_t表示在$[0,t]$内发生这种点的数目，即随机事件发生的次数，则$\{N_t;t\in[0,\infty)\}$是一个取非负整数值的递增的随机过程，一般称为伴随这随机点过程的计数过程(counting process).

(2) 在一维的情形($\mathscr{H}=\mathbf{R}_1$)，考虑一随机点的发生与相继另一随机点发生的间距，令T_1表示初始时刻到第一个随机点发生的时间间隔，T_n表示第$n-1$个随机点到第n个随机点发生的时间间隔，则$T_1,T_2,\cdots$是一随机序列，$T_n(n=1,2,\cdots)$有时称为点间间距，而把第n个"随机点"发生的时间t_n称为等待时间. 这样等待时间与点间间距有如下关系：

$$T_1=t_1,T_2=t_2-t_1,\cdots,T_n=t_n-t_{n-1},$$

而对任意n，

$$t_n=T_1+T_2+\cdots+T_n.$$

当$T_2,T_3,\cdots$同分布时(可能T_1不同)，重要的问题是考虑$T_n(n=2,3,\cdots)$的分布，即考虑点间间距分布.

随机点过程在早期称为随机事件流，它早期的研究出现于排队论和运筹学等分支中，近年来对随机点过程的基础理论研究和应用正蓬勃开展，这方面的内容可参看[13].

关于计数过程有如下定理：

专业课辅导+真题资料+微信公众号：432应统书院

定理 11.2.1 若计数过程$\{N_t;t\in[t_0,\infty)\}$满足下列条件：

(1) $P\{N_{t_0}=0\}=1$，即以概率为1地在初始时刻处于状态为0；

(2) $\{N_t\}$为平稳独立增量过程；

(3) 对任意$t>t_0$，$0<P\{N_t>0\}<1$. 换句话说，在时间区间$[t_0,t)$内有“随机点”发生的概率为正，但不是一定有“随机点”发生；

(4) 对任意$t\geq t_0$，

$$\lim_{h\to 0}\frac{P\{N_{t+h}-N_t\geq 2\}}{P\{N_{t+h}-N_t=1\}}=0,$$

换句话说，在充分小的时间区间内，至多有一个“随机点”发生，则$\{N_t\}$为齐次泊松过程.

定理 11.2.1 的证明可参看定理 2.1.2 的证明.

上述的计数过程$\{N_t\}$及其相应的等待时间序列$\{t_n\}$有如下列事件的等价关系：对任意$t>0$及$n=1,2,\cdots$有

(1) $\{N_t\leq n\}=\{t_{n+1}>t\}$；

(2) $\{N_t=n\}=\{t_n\leq t\}\cap\{t_{n+1}>t\}$；

(3) $\{N_t<n\}=\{t_n>t\}$.

从而有

$$\begin{aligned}&P\{N_t\leq n\}=P\{t_{n+1}>t\};\\&P\{N_t=n\}=P\{t_n\leq t\}-P\{t_{n+1}\leq t\};\\&P\{N_t=0\}=1-P\{t_1\leq t\};\\&P\{N_t<n\}=P\{t_n>t\}.\end{aligned}\tag{11.2.2}$$

下面讨论满足定理 11.2.1 条件的计数(泊松)过程.

定理 11.2.2 等待时间t_n服从参数为n和λ的Γ分布，即其分布密度函数为

$$f_{t_n}(t)=\begin{cases}\lambda e^{-\lambda t}\dfrac{(\lambda t)^{n-1}}{(n-1)!}, & t>0;\\ 0, & t\leq 0.\end{cases}\tag{11.2.3}$$

证 由(11.2.2)式，并注意到t_n是连续型随机变量，故有

专业课辅导+真题资料+微信公众号：432应统书院

$$1-P\{t_n<t\}=P\{t_n>t\}=P\{N_t<n\}$$
$$=\sum_{k=0}^{n-1}\frac{(\lambda t)^k}{k!}e^{-\lambda t}.$$

因此对 $t>0$，t_n 的分布函数为

$$F_{t_n}(t)=P\{t_n<t\}=P\{t_n\leqslant t\}$$
$$=1-\sum_{k=0}^{n-1}\frac{(\lambda t)^k}{k!}e^{-\lambda t}.$$

显然，当 $t\leqslant 0$ 时，$F_{t_n}(t)=0$. 从而定理 11.2.2 证毕.

读者可求得(或参看本书上册第二章的常用分布表)

$$E(t_n)=\frac{n}{\lambda},\quad D(t_n)=\frac{n}{\lambda^2}.$$

对于点间间距序列 $\{T_n\}$ 有如下定理[13]：

定理 11.2.3 $\{T_n\}$ 是一以 $\frac{1}{\lambda}$ 为数学期望的指数分布，且为独立同分布序列(独立同分布用符号 i. i. d 表示)，其 T_n 的分布密度函数为

$$f_{T_n}(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0;\\ 0, & x\leqslant 0.\end{cases}\tag{11.2.4}$$

例 11.2.4 令 N_t 表示 $[0,t]$ 时间区间内(单位：min)顾客到达某商店的人数，设 $\{N_t\}$ 是泊松过程. 根据历史资料统计分析，顾客到达该商店的强度是每小时平均 30 人. 求两顾客相继到达的时间间隔短于 4 分钟的概率.

解 依题意，我们需要求 $P\{T_n<4\}$. 由假设 $\lambda=\frac{30}{60}=\frac{1}{2}$，按(11.2.4)式得

$$P\{T_n<4\}=\int_0^4\frac{1}{2}e^{-\frac{1}{2}x}dx$$
$$=-e^{-2}+1\approx 0.865.$$

例 11.2.5 设一随机排队系统具有如下性质：

(1)“顾客”源是无穷的，且相互独立地单个到达服务机构，

专业课辅导+真题资料+微信公众号：432应统书院

并且在时间长度 t 内“顾客”到达的个数为 N_t，其所对应的 $\{N_t\}$ 是泊松过程；

（2）服务机构是单“服务台”，各顾客的服务时间 ν_n 是相互独立、服从相同指数分布的随机变量，即

$$f_{\nu_n}(t)=\begin{cases}\mu e^{-\mu t}, & t>0;\\ 0 & t\leqslant 0\end{cases}\quad(\mu>0).$$

（3）排队规则是单队，队的长度无限制，先到先接受“服务”；

（4）“顾客”相继到达的时间与“服务”时间是相互独立的.

试求：

① 平均排队长度 L_s（平均“顾客”数，包括正在接受服务的“顾客”数）；

② 平均等待的排队长度 L_q（等待的“顾客”平均数）；

③ “顾客”逗留时间的平均数 $\overline{W}_s$；

④ “顾客”等待时间的平均数 $\overline{W}_q$.

解 令 $p_n(t)=P\{$在时刻 t 有 n 个“顾客”$\}$.

在前面假设下并注意到定理 11.2.1，我们有：

$P\{$在$[t,t+\Delta t]$内有一个“顾客”到达$\}=\lambda\Delta t+o(\Delta t)$；

$P\{$在$[t,t+\Delta t]$内有一个“顾客”接受“服务”后离去$\}=\mu\Delta t+o(\Delta t)$；

$P\{$在$[t,t+\Delta t]$内有一个“顾客”正在接受“服务”$\}=1-\mu\Delta t+o(\Delta t)$；

$P\{$在$[t,t+\Delta t]$内多于一个“顾客”到达或离去$\}=o(\Delta t)$.

在时刻间隔$[t,t+\Delta t]$中，排队系统有如下四种情况（到达或离去两个以上可忽略）：

$A=\{$在时刻 t 有 n 个“顾客”，在时刻 $t+\Delta t$，“顾客”仍为 n 个，其间没有“顾客”到达或离去$\}$；

$B=\{$在时刻 t 有“顾客”$(n+1)$个，在时刻 $t+\Delta t$ 有 n 个“顾客”，其间无“顾客”到达，但有一个离去$\}$；

专业课辅导+真题资料+微信公众号：432应统书院

$C=\{$在时刻 t 有“顾客”$(n-1)$ 个，在时刻 $t+\Delta t$ 有 n 个“顾客”，其间到达一个“顾客”，而没有“顾客”离去$\}$；

$D=\{$在时刻 t 有 n 个“顾客”，在时刻 $t+\Delta t$ 仍有 n 个“顾客”，其间均有一个“顾客”到达和离去$\}$.

如果略去 $o(\Delta t)$ 不计，则

$$P(A)=p_n(t)(1-\lambda\Delta t)(1-\mu\Delta t);$$

$$P(B)=p_{n+1}(t)(1-\lambda\Delta t)(\mu\Delta t);$$

$$P(C)=p_{n-1}(t)(\lambda\Delta t)(1-\mu\Delta t);$$

$$P(D)=p_n(t)(\lambda\Delta t)(\mu\Delta t).$$

由于 A,B,C,D 是互不相容的，故有

$$p_n(t+\Delta t)=p_n(t)(1-\lambda\Delta t-\mu\Delta t)+p_{n+1}(t)\mu\Delta t$$
$$+p_{n-1}(t)\lambda\Delta t+o(\Delta t),$$

从而

$$\frac{p_n(t+\Delta t)-p_n(t)}{\Delta t}=\lambda p_{n-1}(t)+\mu p_{n+1}(t)-(\lambda+\mu)p_n(t)+\frac{o(\Delta t)}{\Delta t},$$

令 $\Delta t\to 0$，由此得 $n=1,2,\cdots$

$$\frac{\mathrm{d}p_n(t)}{\mathrm{d}t}=\lambda p_{n-1}(t)+\mu p_{n+1}(t)-(\lambda+\mu)p_n(t). \tag{11.2.5}$$

当 $n=0$ 时，只有 A,B 出现，故

$$p_0(t+\Delta t)=p_0(t)(1-\lambda\Delta t)+p_1(1-\lambda\Delta t)(\mu\Delta t),$$

从而有

$$\frac{\mathrm{d}p_0(t)}{\mathrm{d}t}=-\lambda p_0(t)+\mu p_1(t). \tag{11.2.6}$$

下面讨论稳态情形，即 $p_n(t)$ 与 t 无关的情形，这时(11.2.5)及(11.2.6)两式变为

$$\begin{cases}\lambda p_{n-1}+\mu p_{n+1}-(\lambda+\mu)p_n=0;\\ \mu p_1-\lambda p_0=0.\end{cases} \tag{11.2.7}$$

从(11.2.7)式可得，对任意 $n=1,2,\cdots$ 有

专业课辅导+真题资料+微信公众号：432应统书院

$$p_n=\left(\frac{\lambda}{\mu}\right)^n p_0. \tag{11.2.8}$$

令 $\rho=\frac{\lambda}{\mu}$，称 ρ 为服务强度（在通讯领域里称为话务强度），它是刻画服务效率和服务机构的利用程度之一.

下面考虑 $\rho=1$ 的情形（否则等待的“顾客”队列将排至无穷远）. 由于 p_n 是概率分布，故

$$\sum_{n=0}^{\infty} p_n=1. \tag{11.2.9}$$

将(11.2.8)式对 n 相加，并将(11.2.9)式代入得

$$1=\sum_{n=0}^{\infty} p_n=p_0\sum_{n=0}^{\infty}\left(\frac{\lambda}{\mu}\right)^n=p_0\sum_{n=0}^{\infty}\rho^n=p_0\frac{1}{1-\rho},$$

从而有

$$p_0=1-\rho. \tag{11.2.10}$$

将(11.2.10)式代入(11.2.8)式得

$$p_n=\rho^n(1-\rho)\quad(n>0). \tag{11.2.11}$$

从上式可求得答案如下：

① 平均排队长度（即平均“顾客”数）

$$\begin{aligned}L_s&=\sum_{n=0}^{\infty} np_n=\sum_{n=0}^{\infty} n\rho^n(1-\rho)\\&=(\rho+2\rho^2+3\rho^3+\cdots)-(\rho^2+2\rho^3+3\rho^4+\cdots)\\&=\rho+\rho^2+\rho^3+\cdots=\frac{\rho}{1-\rho}=\frac{\lambda}{\mu-\lambda};\end{aligned} \tag{11.2.12}$$

② 平均等待的“顾客”数

$$\begin{aligned}L_q&=\sum_{n=1}^{\infty}(n-1)p_n=\sum_{n=0}^{\infty} np_n-\sum_{n=1}^{\infty} p_n\\&=L_s-\sum_{n=0}^{\infty} p_n+p_0=L_s-1+(1-\rho)\\&=L_s-\rho=\frac{\rho^2}{1-\rho}=\rho\frac{\lambda}{\mu-\lambda}.\end{aligned}$$

专业课辅导+真题资料+微信公众号：432应统书院

关于逗留时间 W_s 可以证明在假设条件下，它服从参数为 $(\mu-\lambda)$ 的指数分布，因此

③ $\overline{W}_s=\dfrac{1}{\mu-\lambda}$；

④ "顾客"等待时间的平均数

$$\overline{W}_q=\overline{W}_s-\frac{1}{\mu}=\frac{1}{\mu-\lambda}-\frac{1}{\mu}=\frac{\rho}{\mu-\lambda}.$$

例 11.2.5 中提到的"顾客"到达"服务"机构这一现象，与我们在前面说及的"随机点"发生是类似的. 因此，那里所说的"顾客"和"服务"，可以是商店里购物的"顾客"和店员提供的服务，也可以是某车站到站的车辆和车站提供的某种服务. 所以，例11.2.5 所得到的那些结果在许多方面获得重要的应用.

例 11.2.6 某医院有一台心电图机，经过一段时间使用后获得统计数据表明，平均每小时有 5 名患者使用该机，且平均占用该机的时间为每人 10 分钟. 根据这些统计数据，院领导是否应该考虑增加设备？是否需要增加坐椅供患者等候时使用？

解 我们假定例 11.2.5 中列举的条件成立，则我们有

$$\hat{\lambda}=5(\text{人/小时}),\quad \hat{\mu}=6(\text{人/小时}).$$

这样有 $\hat{\rho}=\dfrac{\hat{\lambda}}{\hat{\mu}}=\dfrac{5}{6}$，用 $\hat{\rho}=\dfrac{5}{6}$ 作为服务强度 ρ 的估计值，从而得到 (11.2.10) 式中的 $\hat{p}_0=1-\hat{\rho}=\dfrac{1}{6}\approx 17\%$，即患者到来时不用等待的概率约为 17%，也就是说，需要等待的患者约为 83%. 此数值显然偏高，应考虑增加设备.

同理，可算得患者到达时（若设心电图室只有四把坐椅）发现无空椅子可坐的概率为

$$\sum_{n=5}^{\infty}\hat{p}_n=\sum_{n=5}^{\infty}\hat{\rho}^n(1-\hat{\rho})=1-\sum_{n=0}^{4}\hat{\rho}^n(1-\hat{\rho})=\hat{\rho}^5\approx 40\%,$$

即约有 40% 的患者要站着等待，还可根据需要，计算其他指标.

例 11.2.7 最优库存问题.

专业课辅导+真题资料+微信公众号：432应统书院

在建设一个新工厂时，常常需要考虑建造一个存放产品的仓库的容量问题. 设客商在需要某一产品时即向某生产厂订购一种货物，且该生产厂只能一个一个地生产某产品（如该产品有某种特殊要求）. 根据实际计算，一个单位产品存放一个单位时间的费用为 c 元（例如保险费、存贮费等）；缺一个货每单位时间损失 h 元（例如缺货应付赔偿费等）. 现在的问题是确定库存容量 s，使得单位时间内的平均总费用（上述两种费用的总和的平均数）$E(s)$ 达到最小.

解 我们把生产该产品的工厂看作服务机构，每一次订货看作一次"顾客"到来，该产品的需求量看作要求服务的"顾客"数，产品的加工时间看作"顾客"的服务时间. 在例 11.2.5 的假设下，令 p_n 表示生产厂有 n 个订货未交的概率（假定与时间 t 无关），则有

（1）平均缺货数为（注意到（11.2.11）式）

$$\begin{aligned}\sum_{n=s}^{\infty}(n-s)p_n &= \sum_{n=s}^{\infty}np_n - s\sum_{n=s}^{\infty}p_n \\ &= \sum_{n=s}^{\infty}n(1-\rho)\rho^n - s\Big(1-\sum_{n=0}^{s-1}p_n\Big) \\ &= \sum_{n=s}^{\infty}n(1-\rho)\rho^n - s\rho^s;\end{aligned}$$

（2）平均存货数为

$$\begin{aligned}\sum_{n=0}^{s-1}(s-n)p_n &= s\sum_{n=0}^{s-1}p_n - \sum_{n=0}^{s-1}np_n \\ &= s(1-\rho^s) - \sum_{n=0}^{s-1}n(1-\rho)\rho^n.\end{aligned}$$

从上述两个结果可得（注意到（11.2.12）式）

$$\begin{aligned}E(s) &= h\Big[\sum_{n=s}^{\infty}np_n - s\rho^s\Big] + c\Big[s(1-\rho^s) - \sum_{n=0}^{s-1}np_n\Big] \\ &= h\sum_{n=s}^{\infty}np_n - hs\rho^s + cs(1-\rho^s) - c\Big[\sum_{n=0}^{\infty}np_n - \sum_{n=s}^{\infty}np_n\Big]\end{aligned}$$

专业课辅导+真题资料+微信公众号：432应统书院

$$
\begin{aligned}
&= c\left[s(1-\rho^s)-\frac{\rho}{1-\rho}\right]-hs\rho^s+(h+c)\left[\sum_{n=s}^{\infty} np_n\right] \\
&= c\left[s(1-\rho^s)-\frac{\rho}{1-\rho}\right]-hs\rho^s+(h+c)\left[\sum_{n=0}^{\infty} np_n-\sum_{n=0}^{s-1} np_n\right] \\
&= c\left[s(1-\rho^s)-\frac{\rho}{1-\rho}\right]-hs\rho^s+(h+c)\left[\frac{\rho}{1-\rho}-\sum_{n=0}^{s-1} n(1-\rho)\rho^n\right] \\
&= c\left[s-\frac{\rho}{1-\rho}\right]-cs\rho^s-hs\rho^s+(h+c)\left[\frac{\rho}{1-\rho}-\sum_{n=0}^{s-1} n(\rho^n-\rho^{n+1})\right] \\
&= c\left[s-\frac{\rho}{1-\rho}\right]-cs\rho^s-hs\rho^s+(h+c) \\
&\quad\cdot\left[\frac{\rho}{1-\rho}-(\rho+\rho^2+\cdots+\rho^{s-1}-(s-1)\rho^s\right] \\
&= c\left[s-\frac{\rho}{1-\rho}\right]-cs\rho^s-hs\rho^s+(h+c)\left[\frac{\rho}{1-\rho}-\frac{\rho-\rho^s}{1-\rho}+(s-1)\rho^s\right] \\
&= c\left[s-\frac{\rho}{1-\rho}\right]+(h+c)\frac{\rho^{s+1}}{1-\rho} \\
&= cs-\frac{c\rho(1-\rho^s)}{1-\rho}+h\frac{\rho^{s+1}}{1-\rho}.
\end{aligned}
$$

这一问题的解是找一 s^*，使上式 $E(s^*)$ 达到最小.

上述计数过程 $\{N_t\}$ 是一齐次泊松过程，是随机点过程中非常重要的一种. 下面简单介绍随机点过程的另一重要的计数过程，它称为更新（计数）过程，它在诸如例 11.2.2 的许多实际问题中有很好的应用.

考虑一随机点过程 $\{t_n\}$，如果其相继两“随机点”发生的时间间隔 $T_1, T_2, \cdots$ 是独立同分布的正值随机序列，我们称与之对应的整数值（计数）过程 $\{N_t; t\geq 0\}$ 为更新计数过程. 许多作者把 $\{T_n\}$ 称为更新过程.

考虑某一设备（机器、电子元件或部件等），它在使用过程中由于各种原因会损坏，这时立即用同一类型的新设备替换（或叫更新）. 假定相继投入使用的设备工作时间（或者说使用寿命）与

专业课辅导+真题资料+微信公众号：432应统书院

某一随机变量 T 的概率分布相同；并且认为各个设备的工作时间（使用寿命）是相互独立的. 如果我们用 $T_n(n=1,2,\cdots)$ 表示第 n 个设备的使用寿命，则 $\{T_n\}$ 是一更新过程，N_t 是区间 $(0,t]$ 中设备的损坏总数，即我们说的更新计数过程. 更新（计数）过程在理论和应用上有许多有意义的结果，我们不详细叙述，有兴趣的读者可参看[21]、[24].

§11.3 马 氏 链

一、定义及例

定义 11.3.1 设随机序列 $\{\xi_n: n\in\mathbf{N}\}$ 满足下列两条件（这里 $\mathbf{N}$ 一般包含有穷或可列无穷个非负整数）：

(i) 对每一 $n\in\mathbf{N}$，ξ_n 只取非负整数值；

(ii) 对任意非负整数 m,l,k 及任意非负整数 $s_1,s_2,\cdots,s_l(s_1<s_2<\cdots<s_l)$ 及 $i_{s_1},i_{s_2},\cdots,i_{s_l},i_m,i_{m+k}(m>s_l)$，有

$$\begin{aligned}&P\{\xi_{m+k}=i_{m+k}\mid\xi_{s_1}=i_{s_1},\xi_{s_2}=i_{s_2},\cdots,\xi_{s_l}=i_{s_l},\xi_m=i_m\}\\&=P\{\xi_{m+k}=i_{m+k}\mid\xi_m=i_m\}.\end{aligned}\tag{11.3.1}$$

当(11.3.1)式左方有意义时，则称 $\{\xi_n: n\in\mathbf{N}\}$（如无特殊声明，简写为 $\{\xi_n\}$）为马尔可夫链，或简称为马氏链.

在下面讨论马氏链的时候，我们总假设构成条件的事件有正的概率，例如事件 $\{\xi_m=i_m\}$ 应满足 $P\{\xi_m=i_m\}>0$.

马氏链是较早出现的马氏过程中的一类. 它早期是用下述比较直观，但一直沿用到现在的语言叙述的.

考虑一系列随机试验，每一次试验只有有穷个或可列无穷个基本事件 $A_i(i=1,2,\cdots)$ 可能发生，且每次试验将发生而且仅发生这些事件中的一个事件. 若事件 A_i 发生，就说系统处于状态 A_i（或简称为状态 i）. 为方便起见，以 $A_i^{(k)}$ 表示第 k 次试验结果处于状态 A_i 这一事件. 我们还常常用“时刻”这一用语，设想每隔一个单位时间进行一次试验，这样，$A_i^{(k)}$ 就可说成“于时刻 k 系统处于状态 i”

专业课辅导+真题资料+微信公众号：432应统书院

这一事件，而且把 $P\{A_{i+s}^{(k+l)}|A_i^{(k)}\}$ 称为系统在时刻 k 处于状态 i，经过 l 步后系统转移到状态 $i+s$ 的转移概率. 这样一来，(11.3.1)式右边就可称为系统在时刻 m 处于状态 i_m，经 k 步后系统处于 i_{m+k} 的转移概率. (11.3.1)就可直观地叙述为：已知系统现在(时刻 m)处于状态 i_m，那么，将来(时刻 $m+k$)的状态与过去(时刻 $s_1, s_2, \cdots, s_l$)的状态无关. 或简述为：已知系统的"现在"，那么，"将来"与"过去"无关. 这种特性我们称为无后效性或马氏性①.

转移概率 $P\{\xi_{m+k}=j|\xi_m=i\}$ 记为 $p_{ij}(m,k)$②. 即

$$p_{ij}(m,k)=P\{\xi_{m+k}=j|\xi_m=i\} \quad (i,j=0,1,2,\cdots).$$

我们把上述转移概率写成矩阵形式，并用 $\boldsymbol{P}(m,k)$ 表示，即

$$\boldsymbol{P}(m,k)=(p_{ij}(m,k))$$

$$=\begin{pmatrix} p_{00}(m,k) & p_{01}(m,k) & p_{02}(m,k) & \cdots & p_{0n}(m,k) & \cdots \\ p_{10}(m,k) & p_{11}(m,k) & p_{12}(m,k) & \cdots & p_{1n}(m,k) & \cdots \\ p_{20}(m,k) & p_{21}(m,k) & p_{22}(m,k) & \cdots & p_{2n}(m,k) & \cdots \\ \vdots & \vdots & \vdots & & \vdots & \\ p_{n0}(m,k) & p_{n1}(m,k) & p_{n2}(m,k) & \cdots & p_{nn}(m,k) & \cdots \\ \cdots\cdots & & & & & \end{pmatrix}$$

并称 $\boldsymbol{P}(m,k)$ 为 $\{\xi_n\}$ 的 k 步转移概率矩阵，特别称 $\boldsymbol{P}(m,1)$ 为转移概率矩阵.

由 $p_{ij}(m,k)$ 的定义，显然

$$\left.\begin{aligned} &p_{ij}(m,k)\geqslant 0, \\ &\sum_j p_{ij}(m,k)=1, i=0,1,2,\cdots \end{aligned}\right\} \tag{11.3.2}$$

---

① 我们不能把马氏性解释为"将来"决定于"现在"而与"过去"无关，它只说在"知道现在"的条件下，过去的信息对推断将来的概率性质不起作用.

② 对于初学者要特别注意，不要忘记转移概率原意是条件概率，这里用了它的形象说法.

在前面几章，我们用 $P(k|i)$ 表示条件概率，在这一章里，条件概率沿用大多数著作的符号.

专业课辅导+真题资料+微信公众号：432应统书院

定理 11.3.1 设$\{\xi_n\}$为一马氏链，则对任意非负整数k,m,l，有

$$p_{ij}(m,k+l)=\sum_r p_{ir}(m,k)p_{rj}(m+k,l). \tag{11.3.3}$$

或采用矩阵写法，则有

$$\boldsymbol{P}(m,k+l)=\boldsymbol{P}(m,k)\boldsymbol{P}(m+k,l).$$

(11.3.3)式称为切普曼-柯尔莫哥洛夫方程.

证 $p_{ij}(m,k+l)=P\{\xi_{m+k+l}=j|\xi_m=i\}$

$$=\frac{P\{\xi_m=i,\xi_{m+k+l}=j\}}{P\{\xi_m=i\}}$$

$$=\sum_r\frac{P\{\xi_m=i,\xi_{m+k}=r,\xi_{m+k+l}=j\}}{P\{\xi_m=i,\xi_{m+k}=r\}}\cdot\frac{P\{\xi_m=i,\xi_{m+k}=r\}}{P\{\xi_m=i\}}$$

$$=\sum_r P\{\xi_{m+k+l}=j|\xi_m=i,\xi_{m+k}=r\}\cdot P\{\xi_{m+k}=r|\xi_m=i\}.$$

利用马氏性得$P\{\xi_{m+k+l}=j|\xi_m=i,\xi_{m+k}=r\}=p_{rj}(m+k,l)$，而$P\{\xi_{m+k}=r|\xi_m=i\}=p_{ir}(m,k)$，故得

$$p_{ij}(m,k+l)=\sum_r p_{ir}(m,k)p_{rj}(m+k,l).$$

用归纳法并应用切普曼-柯尔莫哥洛夫方程可证得下面公式成立：

$$\begin{cases} p_{ij}(m,k+1)=\sum\limits_{s_1,s_2,\cdots,s_k} p_{is_1}(m,1)p_{s_1s_2}(m+1,1)\cdots \\ \qquad\qquad \cdot p_{s_kj}(m+k,1); \\ p_{ij}(m,k+1)=\sum\limits_r p_{ir}(m,1)p_{rj}(m+1,k) \\ \qquad\qquad =\sum\limits_r p_{ir}(m,k)p_{rj}(m+k,1). \end{cases} \tag{11.3.4}$$

对每一非负整数k，记

$$p_j(k)=P\{\xi_k=j\}\quad(j=0,1,2,\cdots);$$

$$p_j=p_j(0).$$

称$p_j(k)$为绝对概率，特别称p_j为初始概率. 由全概率公式可得

专业课辅导+真题资料+微信公众号：432应统书院

$$p_j(k+1)=\sum_i p_i p_{ij}(0,k+1)$$
$$=\sum_i p_i(k)p_{ij}(k,1). \qquad (11.3.5)$$

(11.3.5)式的前一个等式说明:绝对概率 $p_j(k)$ 可由初始概率和转移概率所决定.

定理 11.3.2 设$\{\xi_n\}$为一马氏链,则对于任意非负整数 $s_1<s_2<\cdots<s_k$,及非负整数 k,有

$$P\{\xi_{s_1}=i_{s_1},\xi_{s_2}=i_{s_2},\cdots,\xi_{s_k}=i_{s_k}\}$$
$$=p_{i_{s_1}}(s_1)\cdot p_{i_{s_1}i_{s_2}}(s_1,s_2-s_1)\cdot\cdots\cdot p_{i_{s_{k-1}}i_{s_k}}(s_{k-1},s_k-s_{k-1}). \qquad (11.3.6)$$

证 由马氏性及概率的乘法公式可得

$$P\{\xi_{s_1}=i_{s_1},\xi_{s_2}=i_{s_2},\cdots,\xi_{s_k}=i_{s_k}\}$$
$$=P\{\xi_{s_1}=i_{s_1}\}\cdot P\{\xi_{s_2}=i_{s_2}|\xi_{s_1}=i_{s_1}\}\cdots$$
$$\cdot P\{\xi_{s_k}=i_{s_k}|\xi_{s_1}=i_{s_1},\xi_{s_2}=i_{s_2},\cdots,\xi_{s_{k-1}}=i_{s_{k-1}}\}$$
$$=P\{\xi_{s_1}=i_{s_1}\}\cdot P\{\xi_{s_2}=i_{s_2}|\xi_{s_1}=i_{s_1}\}\cdots$$
$$\cdot P\{\xi_{s_k}=i_{s_k}|\xi_{s_{k-1}}=i_{s_{k-1}}\}$$
$$=p_{i_{s_1}}(s_1)p_{i_{s_1}i_{s_2}}(s_1,s_2-s_1)\cdot\cdots\cdot p_{i_{s_{k-1}}i_{s_k}}(s_{k-1},s_k-s_{k-1}).$$

从(11.3.5)及(11.3.6)两式可见,对于马氏链,任意有限维概率函数完全由初始概率及转移概率所决定,也就是说,对于马氏链来说,初始概率及转移概率(只牵涉到二维概率函数)就可以完备地描述马氏链的统计特性.

二、齐次马氏链

定义 11.3.2 称马氏链$\{\xi_n\}$为齐次的,如果它的转移概率矩阵 $\boldsymbol{P}(m,1)$ 与 m 无关,记为 $\boldsymbol{P}(1)$. 即对任意非负整数 m,有

$$p_{ij}(m,1)=p_{ij}(1) \quad (i,j=0,1,2,\cdots).$$

由(11.3.4)式的第一个等式可得:

(1) 马氏链$\{\xi_n\}$是齐次的充分必要条件为:任意 k 步转移概率 $p_{ij}(m,k)$ 都与 m 无关. 此时,记 k 步转移概率矩阵为 $\boldsymbol{P}(k)$;

(2) 若$\{\xi_n\}$为齐次马氏链,则任意 k 步转移概率 $p_{ij}(k)$,皆可

专业课辅导+真题资料+微信公众号：432应统书院

由一步转移概率决定. 因此，$p_{ij}(1)$ 有特殊意义.

对齐次马氏链，我们记

$$p_{ij}=p_{ij}(1),$$

$$p_{ij}(k)=P\{\xi_{m+k}=j\mid\xi_m=i\}.$$

这样一来，对于齐次马氏链来说，只需将 $p_{ij}(m,k)$ 改为 $p_{ij}(k)$，前面所叙述的诸命题仍然成立，我们特别写出下列诸式：

$$\begin{cases} p_{ij}\geqslant 0 \quad (i,j=0,1,2,\cdots), \\ \sum\limits_j p_{ij}=1,\text{对任意 } i=0,1,\cdots \text{皆成立}; \end{cases} \tag{11.3.7}$$

$$p_{ij}(k+1)=\sum_{s_1,s_2,\cdots,s_k} p_{is_1}p_{s_1s_2}\cdots p_{s_kj}; \tag{11.3.8}$$

$$p_{ij}(k+l)=\sum_r p_{ir}(k)p_{rj}(l); \tag{11.3.9}$$

$$p_j(k+1)=\sum_i p_ip_{ij}(k+1)=\sum_i p_i(k)p_{ij}. \tag{11.3.10}$$

在(11.3.8)式中，特别当 $k=1$ 时，有

$$p_{ij}(2)=\sum_r p_{ir}p_{rj}.$$

把它写成矩阵的形式为

$$\boldsymbol{P}(2)=[\boldsymbol{P}(1)]^2.$$

用数学归纳法可得

$$\boldsymbol{P}(n)=[\boldsymbol{P}(1)]^n. \tag{11.3.11}$$

例 11.3.1 考虑伯努利试验，每次试验有两种状态：$A_1=A$，$A_2=\bar{A}$. 由于试验是独立的，且 $P(A_1)=p$，$P(A_2)=q$，$q=1-p$. 因此，在第 k 次试验出现 A_i（用 $\xi_k=i$ 表示）的条件下第 $k+1$ 次试验出现 A_j 的条件概率与 k 无关，且等于 $P(A_j)$，这说明伯努利试验构成一个齐次马氏链，并且有

$$p_{11}=p_{21}=p;$$

$$p_{12}=p_{22}=q,$$

其转移概率矩阵为

$$\boldsymbol{P}(1)=\begin{pmatrix} p & q \\ p & q \end{pmatrix},$$

专业课辅导+真题资料+微信公众号：432应统书院

并可算得

$$\boldsymbol{P}(2)=\begin{pmatrix} p & q \\ p & q \end{pmatrix}=\boldsymbol{P}(1).$$

更一般地，有

$$\boldsymbol{P}(n)=\begin{pmatrix} p & q \\ p & q \end{pmatrix}=\boldsymbol{P}(1) \quad (n\geqslant 2).$$

例 11.3.2 考虑所谓“一个分子在两个附着壁之间的随机游动”这一模型. 这是物理学中常常出现的一种现象：一个分子能处在数轴上 $1,2,\cdots,s$ 这 s 点中任一位置. 一旦这分子到达 1 这一点处，它就一直留在那里不动，对 s 点也有这种情况. 将 1 及 s 这两点称为附着壁. 如果在时刻 t 的时候这分子处在点 i 处，其中 $2\leqslant i\leqslant s-1$，那么，它在一个单位时间后转到 $i+1$ 处（称为向前）的概率为 $p(0<p<1)$，转到点 $i-1$ 处（称为向后）的概率为 q，$q=1-p$. 我们规定这分子只能“向前”或“向后”游动一步，并且经过一单位时间，它必须“向前”或“向后”作游动（1 和 s 处除外）. 这样一个随机游动构成一个马氏链，并且是齐次的，其转移概率为：

$$p_{11}=p_{ss}=1.$$

且对于 $i(2\leqslant i\leqslant s-1)$ 有

$$p_{ir}=\begin{cases} p, & \text{当 } r=i+1; \\ q, & \text{当 } r=i-1; \\ 0, & \text{其他}. \end{cases}$$

即转移概率矩阵为如下形式：

$$\boldsymbol{P}(1)=\begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & \cdots & 0 & \cdots & \cdots & 0 \\ q & 0 & p & 0 & \cdots & \cdots & 0 & \cdots & \cdots & 0 \\ 0 & q & 0 & p & 0 & \cdots & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & q & 0 & p \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & \cdots & 1 \end{pmatrix}$$

让我们特别考虑当 $s=3$ 的情形，这时有

专业课辅导+真题资料+微信公众号：432应统书院

$$P(1)=\begin{pmatrix}1 & 0 & 0\\ q & 0 & p\\ 0 & 0 & 1\end{pmatrix}.$$

可得当 $n\geqslant 2$ 时，

$$P(n)=P(1).$$

例 11.3.3 考虑没有附着壁的只有三个状态的随机游动.

类似于例 11.3.2 中的分子随机游动，但不假设有两个附着壁，即设 $p_{11}=q, p_{33}=p$. 其余假设与例 11.3.2 相同. 这时，转移概率矩阵为

$$P(1)=\begin{pmatrix}q & p & 0\\ q & 0 & p\\ 0 & q & p\end{pmatrix}.$$

可求得

$$P(2)=[P(1)]^2=\begin{pmatrix}q^2+pq & pq & p^2\\ q^2 & 2pq & p^2\\ q^2 & pq & p^2+pq\end{pmatrix}$$

$$=\begin{pmatrix}q & pq & p^2\\ q^2 & 2pq & p^2\\ q^2 & pq & p\end{pmatrix}.$$

三、遍历性、最终分布与平稳分布

定义 11.3.3 设 $\{\xi_n\}$ 为齐次马氏链，若对某一状态 j，存在不依赖于 i 的数 π_j，使得

$$\lim_{n\to\infty} p_{ij}(n)=\pi_j, \tag{11.3.12}$$

则称状态 j 是 $\{\xi_n\}$ 的遍历状态. 如果马氏链 $\{\xi_n\}$ 的所有状态都是遍历状态，则称马氏链 $\{\xi_n\}$ 为遍历链. 如果在(11.3.12)式中的 π_j 满足 $\pi_j\geqslant 0(j\in\mathbf{N})$ 且 $\sum\limits_{j\in\mathbf{N}}\pi_j=1$，则称遍历马氏链 $\{\xi_n\}$ 具有最终分布(或称极限分布).

遍历性的直观解释为：不论系统自哪一状态出发，当转移的

专业课辅导+真题资料+微信公众号：432应统书院

“步长”n 充分大时，转移到状态 j 的概率近似于某数 π_j，即 $p_{ij}(n)\approx\pi_j$.

值得注意的是，$\{\pi_j\}$ 不一定是概率分布. 因为转移概率 $p_{ij}(n)$ 满足 $\sum\limits_{j\in\mathbf{N}}p_{ij}(n)=1$. 对遍历链 $\{\xi_n\}$，任取 $\mathbf{N}$ 中的有限子集 N_1 有

$$\sum_{j\in N_1}p_{ij}(n)\leqslant 1.$$

由(11.3.12)式，令上式左端 $n\to\infty$ 时，有

$$\lim_{n\to\infty}\sum_{j\in N_1}p_{ij}(n)=\sum_{j\in N_1}\lim_{n\to\infty}p_{ij}(n)=\sum_{j\in N_1}\pi_j,$$

从而有

$$\sum_{j\in N_1}\pi_j\leqslant 1.$$

由 N_1 的选择任意性. 故有

$$\sum_{j\in\mathbf{N}}\pi_j\leqslant 1.$$

对于马氏链来说，我们有如下定理：

定理 11.3.3 有穷状态的遍历马氏链 $\{\xi_n\}$ 具有最终分布 $\{\pi_j:j\in\mathbf{N}\}$.

解 设 N 为有穷个状态组成，则有

$$\sum_{j\in\mathbf{N}}p_{ij}(n)=1.$$

注意到 N 为有穷集. 因此

$$\sum_{j\in\mathbf{N}}\pi_j=\sum_{j\in\mathbf{N}}\lim_{n\to\infty}p_{ij}(n)=\lim_{n\to\infty}\sum_{j\in\mathbf{N}}p_{ij}(n)=1.$$

从这一定理我们看到，对有穷状态遍历链来说，$\{\pi_j\}$ 是一概率分布，且为遍历链的最终分布.

“最终分布”这一名称是由以下一个事实得到的：不管初始分布 $\{p_i(0):i\in\mathbf{N}\}$ 如何，无条件概率分布 $\{p_i(n):i\in\mathbf{N}\}$ 满足

$$\begin{aligned}\lim_{n\to\infty}p_j(n)&=\lim_{n\to\infty}\sum_{i\in\mathbf{N}}p_i(0)p_{i\cdot j}(n)\\&=\sum_{i\in\mathbf{N}}p_i(0)\left[\lim_{n\to\infty}p_{i\cdot j}(n)\right]\end{aligned}$$

专业课辅导+真题资料+微信公众号：432应统书院

$$=\pi_j.$$

即当 n 充分大时，无条件概率 $p_j(n)\approx\pi_j$.

定理 11.3.4 对有穷 $(s+1)$ 个状态的齐次马氏链 $\{\xi_n\}$，若存在正整数 n_0，使对一切 $i,j=0,1,2,\cdots,s$，有

$$p_{ij}(n_0)>0,$$

则此马氏链 $\{\xi_n\}$ 是遍历的，而且(11.3.12)式中之 $\pi_j(j=0,1,2,\cdots,s)$ 是如下方程组

$$\pi_j=\sum_{i=0}^{s}\pi_i p_{ij}\quad(j=0,1,2,\cdots,s)\qquad(11.3.13)$$

在满足条件：

$$\pi_j\geqslant 0\quad(j=0,1,2,\cdots,s);$$

$$\sum_{j=0}^{s}\pi_j=1$$

下的唯一解.

定理 11.3.4 的证明可参看[17].

例 11.3.4 如例 11.3.3. 假设 $0<p<1$，我们可得：对任意 $i,j=1,2,3$，找到 $n_0=2$，使得

$$p_{ij}(2)>0.$$

由定理 11.3.4 可知，例 11.3.3 所述的随机游动构成的马氏链是遍历的. 即对任意 $i=1,2,3$，

$$\lim_{n\to\infty}p_{ij}(n)=\pi_j\quad(j=1,2,3).$$

且其中 π_j 可根据(11.3.13)式求得如下：由

$$\pi_1=\pi_1 q+\pi_2 q,$$

$$\pi_2=\pi_1 p+\pi_3 q,$$

$$\pi_3=(\pi_2+\pi_3)p$$

及条件

$$\pi_j>0\quad(j=1,2,3);$$

$$\pi_1+\pi_2+\pi_3=1.$$

可得

专业课辅导+真题资料+微信公众号：432应统书院

$$\pi_2=\frac{p}{q}\pi_1;$$

$$\pi_3=\left(\frac{p}{q}\right)^2\pi_1;$$

$$\pi_1\left[1+\frac{p}{q}+\left(\frac{p}{q}\right)^2\right]=1.$$

当 $p=q=\frac{1}{2}$ 时可得，

$$\pi_1=\pi_2=\pi_3=\frac{1}{3}.$$

这表示在极限情况，处于状态 1 或 2 或 3 皆是等可能的. 这个结果很直观，因为此时分子在任一状态“向前”或“向后”的可能性相等 $\left(p=q=\frac{1}{2}\right)$，因而当这分子游动相当长的时间后处于任一状态的可能性也是相同的.

当 $p\neq q$ 时，可得

$$\pi_j=\frac{1-\frac{p}{q}}{1-\left(\frac{p}{q}\right)^3}\left(\frac{p}{q}\right)^{j-1}\quad(j=1,2,3).$$

从上式看到，当 $p>q$ 时，π_j 随 j 的增加而增大；当 $p<q$ 时，则相反. 例如，当 $\frac{p}{q}=2$ 时，

$$\pi_1=\frac{1}{7},\pi_2=\frac{2}{7},\pi_3=\frac{4}{7}.$$

而在 $\frac{p}{q}=\frac{1}{2}$ 时，

$$\pi_1=\frac{4}{7},\pi_2=\frac{2}{7},\pi_3=\frac{1}{7}.$$

上面的结果也是很直观的，当 $p>q$ 时，表示这分子“向前”的可能性大于“向后”的可能性，因而当这分子游动相当长时间后，

专业课辅导+真题资料+微信公众号：432应统书院

处于下一个状态(例如状态3)的可能性比处于前一个状态(例如状态2)的可能性大. 当 $p<q$ 时,则有相反的结果.

定义 11.3.4 设 $\{\xi_n\}$ 为齐次马氏链,一步转移概率为 $\{p_{ij}\}$,若存在数序列 $\{\pi_k\}$,使得

(1) 对任意 k, $\pi_k \geqslant 0$,且 $\sum_{k\in\mathbf{N}} \pi_k = 1$,即 $\{\pi_k\}$ 为一概率分布;

(2) 对任意 $k\in\mathbf{N}$,

$$\pi_k = \sum_{i\in\mathbf{N}} \pi_i p_{ik},$$

则称 $\{\pi_k\}$ 为 $\{\xi_n\}$ 的一个平稳分布.

定理 11.3.5 $\{\pi_k\}$ 是 $\{\xi_n\}$ 的平稳分布的充分必要条件为

(1) $\{\pi_k\}$ 是一概率分布;

(2) 对任一 $k\in\mathbf{N}$,及任一正整数 m,有

$$\pi_k = \sum_{i\in\mathbf{N}} \pi_i p_{ik}(m).$$

证 充分性是显然的. 读者可用归纳法去证明必要性.

对于一个齐次马氏链,何时存在平稳分布? 如存在是否唯一? 对于有穷状态齐次马氏链来说,定理 11.3.3 已给出了一个答案. 即若能找到一个正整数 n_0,使得该马氏链对任意状态 i,j 其转移概率 $p_{ij}(n_0)>0$. 直观地说,系统处于任一状态 i,都存在 n_0,使得经 n_0 步后有可能到达任一状态 j.

例 11.3.4 给出了一个随机游动存在平稳分布的例子.

定理 11.3.4 不仅给出了一个有穷状态齐次马氏链具有遍历性的条件,也给出了该马氏链存在平稳分布和最终分布的条件,而且该最终分布就是平稳分布.

对于一般状态齐次马氏链的讨论,读者可参看[17]、[24].

平稳分布这名称的一个解释是如下一个定理:

定理 11.3.6 设 $\{\pi_k\}$ 是 $\{\xi_n\}$ 的一个平稳分布. 若 $\{\xi_n\}$ 的初始分布 $\{p_k\}$ 为平稳分布 $\{\pi_k\}$,即 $p_k = P\{\xi_0 = k\} = \pi_k$,则 $\{\xi_n\}$ 是强平稳序列.

专业课辅导+真题资料+微信公众号：432应统书院

证 对任意 $k\in\mathbf{N}$，由(11.3.10)式及假设，对任意 $n\geqslant 1$ 有

$$p_k(n)=P\{\xi_n=k\}=\sum_{i\in\mathbf{N}}p_ip_{ik}(n)=\sum_{i\in\mathbf{N}}\pi_ip_{ik}(n)=\pi_k. \tag{11.3.14}$$

另一方面，对任意 $n>0$ 及 $k\geqslant 1$，注意到(11.3.6)式及 $\{\xi_n\}$ 是齐次马氏链假设，由(11.3.14)式可得，对任意 $s_1,s_2,\cdots,s_n$ 有

$$\begin{aligned}&P\{\xi_{s_1+k}=i_{s_1},\cdots,\xi_{s_n+k}=i_{s_n}\}\\&=P\{\xi_{s_1+k}=i_{s_1}\}p_{i_{s_1}\cdot i_{s_2}}(s_2-s_1)\cdots p_{i_{s_{n-1}}\cdot i_{s_n}}(s_n-s_{n-1})\\&=P\{\xi_{s_1}=i_{s_1}\}p_{i_{s_1}\cdot i_{s_2}}(s_2-s_1)\cdots p_{i_{s_{n-1}}\cdot i_{s_n}}(s_n-s_{n-1})\\&=P\{\xi_{s_1}=i_{s_1},\cdots,\xi_{s_n}=i_{s_n}\},\end{aligned}$$

即 $\{\xi_n\}$ 的任意 n 维概率分布作时间平移不改变(在§11.1把这样的随机序列称为强平稳序列). 证毕.

从(11.3.14)式看到，对任意 $n\in\mathbf{N}$，

$$p_k(n)=\pi_k,$$

即在任一时刻 n 系统处于状态 k 的概率不变. 这就是"平稳"的含义.

四、分支过程

考虑一个由个体(这个个体可以是有生命的，如人、动物、细胞或细菌等，也可以是无生命的，如中子、质子那样的物理粒子等)组成的群体. 群体中每一个个体可以产生(繁殖或说"复制")新的同类个体(后代).

我们用 ξ_n 表示群体第 n 代($n=0,1,2,\cdots$)的个体数. 因此，ξ_0 表示群体第零代的个体数(一般假定 $\xi_0\equiv 1$). 第零代的个体称为群体的祖宗. 由祖宗产生的后代组成群体的第一代，由第一代产生的后代称为第二代……对于这样一个繁殖过程，我们假定：

(1) 群体中任一个体繁殖的后代数都为同一个取非负整数值的随机变量 ζ，其概率函数 $P\{\zeta=k\}=p_k,k\geqslant 0$；

(2) 群体中各个体的繁殖是相互独立的.

这样定义的随机序列 $\{\xi_n:n\geqslant 0\}$ 称为带有后代分布 $\{p_k\}$ 的分支过程.

专业课辅导+真题资料+微信公众号：432应统书院

利用本书上册§4.3的母函数方法，可讨论分支过程$\{\xi_n\}$的灭绝概率$P\{\xi_n=0\}$，以及它的数字特征.

例 11.3.5 设分支过程$\{\xi_n\}$中所述的任一个体繁殖的后代数ζ的概率函数为

$$p_k=P\{\zeta=k\}=\begin{cases}1-p, & k=0;\\ p, & k=1;\quad 0<p<1.\\ 0, & \text{其他}\end{cases}$$

可计算得(参考[21])

$$P\{\xi_n=0\}=(1-p)(1+p+p^2+\cdots+p^{n-1})$$

$$=(1-p)\frac{1-p^n}{1-p}=1-p^n\xrightarrow[(n\to\infty)]{}1.$$

上述表明，群体中的任一个体至多只繁殖一个后代，并且有正概率断代(即$1-p>0$)，这样的群体最终要灭绝.

例 11.3.6 设$\{\xi_n\}$为一分支过程，且有$E(\xi_n)=\mu^n$，$n\geqslant 0$. 其中$\mu=E(\zeta)$，即μ为任一个体繁殖后代的平均数. 从而有

$$E(\xi_n)\xrightarrow[(n\to\infty)]{}\begin{cases}0, & \mu<1;\\ 1, & \mu=1;\\ \infty, & \mu>1.\end{cases}$$

上式表示：$\mu<1$，意味着个体繁殖后代的平均数低于上代的死亡数，它将导致群体的灭绝；$\mu>1$意味着个体繁殖后代数平均高于上代的死亡数，它将导致群体的爆炸.

五、销售市场决策中应用的例子

设$\{\xi_n:n\geqslant 0\}$表示n时刻(例如第n年、第n月或第n季度)某一商品的销售状态(例如畅销、滞销等). 如果这一商品的销售状态有如下性质：

(1) 无后效性：n时刻的销售状态如何，只与$(n-1)$时刻有关，而与$(n-1)$时刻以前无关；

(2) 时齐性：销售过程中，从一状态转到另一状态与初始时刻无关.

专业课辅导+真题资料+微信公众号：432应统书院

这样，用以描述销售状态的随机序列$\{\xi_n\}$是一齐次马氏链.

现在考虑如下问题：

经营管理人员为了获得最可能多的盈利，常常通过加强经营管理的各种措施（策略），如提高产品质量、降低产品成本或降价、改善服务态度、加强宣传和推销等来增强竞争能力，提高市场占有率，从而获得更多的盈利. 下面考虑这一问题的数学模型，并给出一个数字例子说明这一问题解决的办法：

设有$k(k=1,2,\cdots,m)$个可能采取的措施，且在第k个措施下，状态转移概率矩阵（设有n个销售状态）为

$$\boldsymbol{P}_k=\begin{pmatrix} {}_kp_{11} & {}_kp_{12} & \cdots & {}_kp_{1n} \\ \vdots & \vdots & & \vdots \\ {}_kp_{n1} & {}_kp_{n2} & \cdots & {}_kp_{nn} \end{pmatrix}$$

令$r_{ij}(k)$表示第k个措施被采取时，从销售状态i转为销售状态j时的盈利（负值表示亏损）. 称矩阵

$$\boldsymbol{R}(k)=\begin{pmatrix} r_{11}(k) & r_{12}(k) & \cdots & r_{1n}(k) \\ \vdots & \vdots & & \vdots \\ r_{n1}(k) & r_{n2}(k) & \cdots & r_{nn}(k) \end{pmatrix}$$

为盈利矩阵.

用$v_i(s)$表示现在状态为i，经s个时刻并选取最优措施下的总期望盈利，则有[17]

$$v_i(s+1)=\max_{1\leqslant k\leqslant m}\left\{\sum_{j=1}^{n}{}_kp_{ij}[r_{ij}(k)+v_j(s)]\right\}. \quad (11.3.15)$$

令$q_i(k)=\sum_{j=1}^{n}{}_kp_{ij}r_{ij}(k)$，它表示现时为状态$i$，取第$k$个措施，经一个单位时刻后的期望盈利. 这样(11.3.15)式变为

$$v_i(s+1)=\max_{1\leqslant k\leqslant m}\left\{q_i(k)+\sum_{j=1}^{n}{}_kp_{ij}v_j(s)\right\}. \quad (11.3.16)$$

引入上述符号后，介绍马尔可夫决策方法和步骤. 为了容易理解，用简单的例子进行叙述，并且把步骤一一列出.

专业课辅导+真题资料+微信公众号：432应统书院

第一步：列出措施表.

首先确定在各种销售状态下采取的措施. 例如，假定销售状态有三种：畅销、一般、滞销. 而在畅销情况下采取两个措施：不刊登广告或刊登广告；在一般销售状态时采取不送货上门或送货上门两个措施；在滞销时采取不降价或降价销售两种措施，把它们列成表 11.3.1.

表 11.3.1

销售状态(i)	(1)畅销	(2)一般	(3)滞销
措施(k)	1. 不刊登广告	1. 不送货上门	1. 不降价
	2. 刊登广告	2. 送货上门	2. 降价

第二步：经过调查研究，列出在各种措施下，不同销售状态下的转移概率和盈利，如表 11.3.2.

表 11.3.2

销售状态	(1)畅销		(2)一般		(3)滞销	
措施	1. 不登广告	2. 登广告	1. 不送货上门	2. 送货上门	1. 不降价	2. 降价
转移概率	${}_1P_{11}=0.4$ ${}_1P_{12}=0.3$ ${}_1P_{13}=0.3$	${}_2P_{11}=0.6$ ${}_2P_{12}=0.3$ ${}_2P_{13}=0.1$	${}_1P_{21}=0$ ${}_1P_{22}=0.8$ ${}_1P_{23}=0.2$	${}_2P_{21}=0.2$ ${}_2P_{22}=0.8$ ${}_2P_{23}=0$	${}_1P_{31}=0$ ${}_1P_{32}=0.1$ ${}_1P_{33}=0.9$	${}_2P_{31}=0.1$ ${}_2P_{32}=0.7$ ${}_2P_{33}=0.2$
盈利/万元	$r_{11}(1)=6$ $r_{12}(1)=4$ $r_{13}(1)=2$	$r_{11}(2)=5$ $r_{12}(2)=3$ $r_{13}(2)=1$	$r_{21}(1)=5$ $r_{22}(1)=3$ $r_{23}(1)=1$	$r_{21}(2)=4$ $r_{22}(2)=2.5$ $r_{23}(2)=0.5$	$r_{31}(1)=5$ $r_{32}(1)=3$ $r_{33}(1)=-1$	$r_{31}(2)=3.5$ $r_{32}(2)=1.5$ $r_{33}(2)=-2.5$

第三步：根据表 11.3.2，应用(11.3.16)式计算有关数字如下：

专业课辅导+真题资料+微信公众号：432应统书院

$$q_1(1)=\sum_{j=1}^{3}{}_1p_{1j}r_{1j}(1)$$
$$=0.4\times6+0.3\times4+0.3\times2$$
$$=4.2;$$
$$q_1(2)=\sum_{j=1}^{3}{}_2p_{1j}r_{1j}(2)$$
$$=0.6\times5+0.3\times3+0.1\times1$$
$$=4.0.$$

同理可算得

$$q_2(1)=2.6;q_2(2)=2.8.$$
$$q_3(1)=-0.6;q_3(2)=0.9.$$

依(11.3.15)式，$v_i(1)=\max\limits_{1\leqslant k\leqslant m}\left\{\sum\limits_{j=1}^{n}{}_kp_{ij}r_{ij}(k)\right\}$. 故有

$$v_1(1)=\max\left\{\sum_{j=1}^{3}{}_1p_{1j}r_{1j}(1),\ \sum_{j=1}^{3}{}_2p_{1j}r_{1j}(2)\right\}$$
$$=\max\{4.2,4.0\}=4.2;$$
$$v_2(1)=\max\{2.6,2.8\}=2.8;$$
$$v_3(1)=\max\{-0.6,0.9\}=0.9.$$

同样根据(11.3.16)式可得

$$v_1(2)=\max\left\{q_1(1)+\sum_{j=1}^{3}{}_2p_{1j}v_j(1),\right.$$
$$\left.q_1(2)+\sum_{j=1}^{3}{}_2p_{1j}v_j(1)\right\}$$
$$=\max\{4.2+(0.4\times4.2+0.3\times2.8+0.3\times0.9),$$
$$4.0+(0.6\times4.2+0.3\times2.8+0.1\times0.9)\}$$
$$=\max\{6.99,7.45\}=7.45;$$
$$v_2(2)=5.88;$$
$$v_3(2)=3.46.$$
$$v_1(3)=10.58;$$
$$v_2(3)=8.99;$$

专业课辅导+真题资料+微信公众号：432应统书院

$v_3(3)=6.45$.

第四步：找出现时状态为 i，经 s 时刻后使得总期望盈利为最大的那个策略（或措施），用 $d_i(s)$ 表示.

例如，现时状态为(1)（即畅销），经一个单位时刻，如果采取策略（措施）为 1（即不刊登广告）可使其期望盈利（$v_1(1)$）最大（为 4.2）. 因此，这时的最优策略（$d_1(1)$）为 1（不刊登广告）. 把上述计算的结果列成表 11.3.3.

表 11.3.3

s	1	2	3	…
$v_1(s)$	4.2	7.45	10.58	…
$v_2(s)$	2.8	5.88	8.99	…
$v_3(s)$	0.9	3.46	6.45	…
$d_1(s)$	1	2	2	…
$d_2(s)$	2	2	2	…
$d_3(s)$	2	2	2	…

从表 11.3.3 看到，若现时状态为畅销，如果采取不刊登广告，经一单位时刻后能获得期望盈利最大（为 4.2）. 但若经两个单位时间，则必须采取刊登广告这一措施，才能获得总期望盈利最大（为 7.45）. 同样，经三个单位时刻时也必须采取刊登广告才能使总期望盈利最大，此时的总期望盈利为 10.58. 其余情况均可由表 11.3.3 看到.

关于马尔可夫决策的理论和方法，可参阅[17]，而有关应用可参阅[24].

§11.4 连续时间马尔可夫链

在 §11.3 里我们讨论的是离散参数马氏链，那里的参数集是可列无穷的. 但我们在 §11.2 里所介绍的维纳过程、泊松过程等的参数集却不是这样. 这就是我们在这一节将要介绍的连续时间

专业课辅导+真题资料+微信公众号：432应统书院

马氏链.

一、定义

定义 11.4.1 设$\{\xi_t:t\in\mathbf{R}_1^+=[0,\infty)\}$为只取非负整数值的实随机过程,若对任意正整数$n$,$t_i\in\mathbf{R}_1^+$,$t_1<t_2<\cdots<t_{n+1}$,以及非负整数$i_1,i_2,\cdots,i_{n+1}$,当$P\{\xi_{t_1}=i_1,\cdots,\xi_{t_n}=i_n\}>0$时,有

$$P\{\xi_{t_{n+1}}=i_{n+1}|\xi_{t_1}=i_1,\cdots,\xi_{t_n}=i_n\}$$
$$=P\{\xi_{t_{n+1}}=i_{n+1}|\xi_{t_n}=i_n\},$$

则称$\{\xi_t:t\in\mathbf{R}_1^+\}$为连续时间马氏链.

令$p_{ij}(t,s)=P\{\xi_s=j|\xi_t=i\}\quad(t,s\in\mathbf{R}_1^+,t<s)$,

则称$p_{ij}(t,s)$为$\{\xi_t:t\in\mathbf{R}_1^+\}$的转移概率函数.

若$p_{ij}(t,s)=p_{ij}(s-t)$,即转移概率函数不依赖于初始时刻,而只依赖于两时刻之差时,则称$\{\xi_t:t\in\mathbf{R}_1^+\}$为齐次连续时间马氏链.下面讨论的马氏链均指这一情形.如不产生混乱,用$\{\xi_t\}$表示齐次连续时间马氏链.

例 11.4.1 设$\{\xi_t\}$是取非负整数值的平稳独立增量过程,$P\{\xi_0=0\}=1$,则$\{\xi_t\}$是齐次连续时间马氏链.

解 事实上,$P\{\xi_{t_{n+1}}=j|\xi_{t_1}=i_1,\cdots,\xi_{t_n}=i_n\}$

$$=\frac{P\{\xi_{t_1}=i_1,\cdots,\xi_{t_n}=i_n,\xi_{t_{n+1}}=j\}}{P\{\xi_{t_1}=i_1,\cdots,\xi_{t_n}=i_n\}}$$

$$=\frac{P\{\xi_{t_1}=i_1,\xi_{t_2}-\xi_{t_1}=i_2-i_1,\cdots,\xi_{t_{n+1}}-\xi_{t_n}=j-i_n\}}{P\{\xi_{t_1}=i_1,\cdots,\xi_{t_n}-\xi_{t_{n-1}}=i_n-i_{n-1}\}}$$

$$=\frac{P\{\xi_{t_1}=i_1\}P\{\xi_{t_2}-\xi_{t_1}=i_2-i_1\}\cdot\cdots\cdot P\{\xi_{t_{n+1}}-\xi_{t_n}=j-i_n\}}{P\{\xi_{t_1}=i_1\}P\{\xi_{t_2}-\xi_{t_1}=i_2-i_1\}\cdot\cdots\cdot P\{\xi_{t_n}-\xi_{t_{n-1}}=i_n-i_{n-1}\}}$$

$$=P\{\xi_{t_{n+1}}-\xi_{t_n}=j-i_n\}$$

$$=P\{\xi_{t_n}=i_n,\xi_{t_{n+1}}-\xi_{t_n}=j-i_n\}/P\{\xi_{t_n}=i_n\}$$

$$=P\{\xi_{t_n}=i_n,\xi_{t_{n+1}}=j\}/P\{\xi_{t_n}=i_n\}$$

$$=P\{\xi_{t_{n+1}}=j|\xi_{t_n}=i_n\}.$$

另一方面,由平稳性可得齐次性.

专业课辅导+真题资料+微信公众号：432应统书院

由此例我们看到§11.2所述的维纳过程、泊松过程均为齐次连续时间马氏链.

齐次连续时间马氏链有如下等价的构造性定义：

定义 11.4.2 设$\{\xi_t: t\in \mathbf{R}_1^+=[0,\infty)\}$为一取非负整数值的实随机过程. 如果它具有以下性质：对任一状态$i\in \mathbf{N}$,

(1) 若随机过程进入状态i后，在这状态的逗留时间为有参数q_i(即其均值为$1/q_i$)的指数分布；

(2) 若随机过程离开状态i,则不管它在状态i的逗留时间有多长和以前经历过的状态，它离开状态i后转移到N中的任一异于i的状态j的概率均为p_{ij},这里$\sum\limits_{\substack{i\neq j\\ j\in \mathbf{N}}} p_{ij}=1$. 证明见[24].

注意到q_i的概率意义($\frac{1}{q_i}$是状态i的逗留时间平均值，$q_i\geqslant 0$),当$q_i=0$时称状态i是吸收的，因为此时逗留平均时间为无穷大，即只要该马氏链进入状态i就永不离开；当$q_i=\infty$时，称状态i为滑过的，因为此时状态i的平均逗留时间为0,因此每当马氏链进入状态i后即马上离开.

对于齐次连续时间马氏链$\{\xi_t: t\in \mathbf{R}_1^+\}$有如下定理[17]：

定理 11.4.1 $\{\xi_t: t\in \mathbf{R}_1^+\}$的转移概率函数$p_{ij}(t)$满足：

(1) 对一切$i\in N$,有

$$\lim_{t\to 0^+}\frac{1-p_{ii}(t)}{t}=q_i \quad (0\leqslant q_i\leqslant \infty),$$

且对任意$t>0$,有

$$0\leqslant \frac{1-p_{ii}(t)}{t}\leqslant q_i;$$

(2) 对任意$i,j\in N, i\neq j$有

$$\lim_{t\to 0^+}\frac{p_{ij}(t)}{t}=q_{ij} \quad (0\leqslant q_{ij}<\infty).$$

由于$p_{ii}(0)=1, p_{ij}(0)=0(i\neq j)$,定理中的条件(1)和(2)中的

专业课辅导+真题资料+微信公众号：432应统书院

q_i(前述的状态 i 的平均逗留时间为$\frac{1}{q_i}$)和 q_{ij}是$\{\xi_t:t\in\mathbf{R}_1^+\}$的转移概率函数 $p_{ii}(t)$和 $p_{ij}(t)$在 $t=0$ 处的右导数. 称 q_i为马氏链在状态 i 的通过强度或通过率,而称 q_{ij}为马氏链的转移强度或转移率,并把 q_{ij}组成的矩阵 $\boldsymbol{Q}=(q_{ij})$称为密度矩阵或速率矩阵,简称 $\boldsymbol{Q}$ 矩阵. 它在研究齐次连续参数马氏链中起重要作用.

二、生灭过程

设$\{\xi_t:t\in\mathbf{R}_1^+\}$是连续时间马氏链,若它的转移概率函数满足下列诸式:

$$\begin{cases} p_{ii+1}(\Delta t)=\lambda_i\Delta t+o(\Delta t) & (i\geqslant 0,\lambda_i\geqslant 0); \\ p_{ii-1}(\Delta t)=\mu_i\Delta t+o(\Delta t) & (i>0,\mu_i=0,\mu_i\geqslant 0); \\ p_{ii}(\Delta t)=1-(\lambda_i+\mu_i)\Delta t+o(\Delta t); \\ p_{ij}(\Delta t)=o(\Delta t) & (|i-j|>1), \end{cases}$$

则称$\{\xi_t:t\in\mathbf{R}_1^+\}$为(齐次)生灭过程.

上述诸式表示,这种过程的状态转移在短时间内仅限于从一个状态向其邻近的另一个状态转移,并且当它处于零状态时,它只可能转移到状态 1. 这一过程可以用来描述某一生物群体的繁殖现象,设 ξ_t表示时刻 t 该群体的个数,则在短时间内群体个数增加 1,或者减少 1,或者不变. 增加 1(多生出一个)的条件概率依赖于原来群体的个数 i,我们用 λ_i表示. 减少 1(死亡一个)的条件概率依赖于原来群体的个数 i,我们用 μ_i表示. 这就是生灭过程称呼的由来.

从生灭过程的定义看到,它是齐次连续时间马氏链. 不难算出其转移强度为 $q_{00}=-\lambda_0,q_{ii}=-(\lambda_i+\mu_i)(i>0),q_{ii+1}=\lambda_i,q_{ii-1}=\mu_i(i>0)$, $q_{ij}=0(|i-j|>1)$,即其 $\boldsymbol{Q}$ 矩阵为

$$\boldsymbol{Q}=\begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \boldsymbol{O} \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & \vdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & \vdots \\ \boldsymbol{O} & \cdots & \cdots & \cdots & \cdots \end{pmatrix},$$

专业课辅导+真题资料+微信公众号：432应统书院

即其 Q 矩阵除主对角线和其两旁的次对角线上的元素外其他元素全为零.

如果生灭过程 $\{\xi_t: t\in \mathbf{R}_1^+\}$ 定义中的 λ_i（称为增殖强度）及 μ_i（称为死亡强度）满足如下关系：

$$\lambda_i = i\lambda \quad (i\in N, \lambda>0);$$
$$\mu_i = i\mu \quad (i\in N, \mu>0).$$

其中 λ、μ 均为常数. 称 $\{\xi_t: t\in \mathbf{R}_1^+\}$ 为线性增长生灭过程，称 λ 为出生率，μ 为死亡率.

当生灭过程 $\{\xi_t: t\in \mathbf{R}_1^+\}$ 定义中的 $\mu_i=0$ $(i\geqslant 0)$，则称 $\{\xi_t: t\in \mathbf{R}_1^+\}$ 为纯生过程；而当 $\lambda_i=0$ $(i\geqslant 0)$，则称 $\{\xi_t: t\in \mathbf{R}_1^+\}$ 为纯死过程. 关于生灭过程的理论和应用，可参看参考书目[24]、[21].

具有更一般的马氏性的过程 $\{\xi_t: t\in T\}$（其中 T 为有穷或无穷区间）称为马氏过程，如果对任意正整数 n，$t_i\in T(i=1,2,\cdots,n)$，$t_1<t_2<\cdots<t_n$，以及任意实数 $x_1,x_2,\cdots,x_{n-2},x,y$，如下等式成立：

$$P\{\xi_{t_n}<y \mid \xi_{t_{n-1}}=x, \xi_{t_{n-2}}=x_{n-2},\cdots,\xi_{t_1}=x_1\}$$
$$=P\{\xi_{t_n}<y \mid \xi_{t_{n-1}}=x\}.$$

这一过程的内容，有兴趣的读者可参阅参考书目[10].

§11.5 均方微积分与随机微分方程

这一节我们引进一些概念与结果，为今后研究随机过程之用.

建立一般函数微积分理论，首先要定义序列收敛的意义. 建立随机过程的微积分理论也是如此. 但随机序列有多种不同的收敛意义：依概率收敛、以概率为 1 收敛、依分布收敛、均方收敛等. 本节着重讲在某些理论与应用中使用较多的一种收敛——均方收敛.

一、随机序列的均方收敛

定义 11.5.1 设 $\{\xi_n: n\in \mathbf{N}_+\}$ 为一实随机序列（其中 $\mathbf{N}_+$ 为全体正整数集，即除去 0 的自然数集，以下同），ξ 为一随机变量（也

专业课辅导+真题资料+微信公众号：432应统书院

可以为一常数)，若$E(\xi_n)^2<\infty$，$E(\xi)^2<\infty$，且

$$\lim_{n\to\infty} E|\xi_n-\xi|^2=0, \tag{11.5.1}$$

则称$\{\xi_n\}$均方收敛于ξ，称ξ为$\{\xi_n\}$的均方极限，记为

$$\underset{n\to\infty}{\text{l. i. m}}\,\xi_n=\xi,$$

或简写为

$$\text{l. i. m}\,\xi_n=\xi.$$

由切比雪夫不等式立即可以得到：若$\{\xi_n\}$均方收敛于ξ，则$\{\xi_n\}$依概率收敛于ξ. 事实上，由切比雪夫不等式：

$$P\{|\xi_n-\xi|\geqslant\varepsilon\}\leqslant\frac{1}{\varepsilon^2}E|\xi_n-\xi|^2,$$

立即得到所要证的结果.

定理 11.5.1 $\{\xi_n\}$均方收敛于ξ的充分必要条件为

$$\lim_{n,m\to\infty} E|\xi_n-\xi_m|^2=0. \tag{11.5.2}$$

证 必要性是明显的，事实上，因为

$$\begin{aligned}|\xi_n-\xi_m|^2&=|\xi_n-\xi+\xi-\xi_m|^2\\&\leqslant 2|\xi_n-\xi|^2+2|\xi_m-\xi|^2,\end{aligned}$$

故

$$\begin{aligned}0\leqslant E|\xi_n-\xi_m|^2&\leqslant 2E|\xi_n-\xi|^2+2E|\xi_m-\xi|^2\\&\longrightarrow 0\ (n,m\to\infty).\end{aligned}$$

充分性的证明可参阅[11]，[25].

定理 11.5.2 设$\{\xi_n\}$，$\{\eta_n\}$，$\{\zeta_n\}$为三个实随机序列，z为随机变量，皆有有限二阶矩. $\{c_n\}$是一常数序列，a，b是两个常数. 令$\xi=\text{l. i. m}\xi_n$，$\eta=\text{l. i. m}\eta_n$，$\zeta=\text{l. i. m}\zeta_n$，$c=\lim\limits_{n\to\infty}c_n$，则

(1) $\text{l. i. m}\,c_n=\lim\limits_{n\to\infty}c_n=c$;

(2) $\text{l. i. m}\,z=z$;

(3) $\text{l. i. m}\,c_nz=cz$;

(4) $\text{l. i. m}(a\xi_n+b\eta_n)=a\xi+b\eta$;

(5) $\lim\limits_{n\to\infty}E(\xi_n)=E(\xi)$，即

专业课辅导+真题资料+微信公众号：432应统书院

$$E(\text{l.i.m}\,\xi_n)=\lim_{n\to\infty}E(\xi_n);$$

(6) $\lim\limits_{n,m\to\infty}E(\xi_n\eta_m)=E(\xi\eta)$，特别有

$$\lim_{n\to\infty}E(\xi_n^2)=E(\xi^2);$$

(7) 若 $E(\xi_n\eta_n)=E(\zeta_n)$，则 $E(\xi\eta)=E(\zeta)$.

证 (1)，(2)，(3)由均方收敛的定义立即可得.(4)的证明则由下面的不等式得到：

$$\begin{aligned}E|a\xi_n+b\eta_n-(a\xi+b\eta)|^2&=E|a(\xi_n-\xi)+b(\eta_n-\eta)|^2\\&\leqslant 2a^2E|\xi_n-\xi|^2+2b^2E|\eta_n-\eta|^2.\end{aligned}$$

(5)的证明由施瓦兹不等式可得.即由

$$\begin{aligned}0\leqslant|E(\xi_n)-E(\xi)|^2&=|E(\xi_n-\xi)|^2\\&\leqslant E|\xi_n-\xi|^2\to 0\quad(n\to\infty),\end{aligned}$$

可得(5).

仿(5)的证明方法可得(6)

(7)的证明如下：

由(5)，(6)及假设可得：

1° $E(\xi_n\eta_n)=E(\zeta_n)\to E(\zeta)$；

2° $E(\xi_n\eta_n)\to E(\xi\eta)$.

由 1°及 2°得 $E(\xi\eta)=E(\zeta)$.

二、随机过程的均方连续

如无特别声明，以下都假定随机过程是实的，其指标集 $T=\mathbf{R}_1$，且有有限二阶矩.

定义 11.5.2 设$\{\xi_t\}$为一随机过程，若对 $t\in\mathbf{R}_1$，

$$\text{l.i.m}_{h\to 0}\xi_{t+h}=\xi_t\left(\text{即}\lim_{h\to 0}E|\xi_{t+h}-\xi_t|^2=0\right),$$

则称$\{\xi_t\}$在 t 点均方连续.若对 $\mathbf{R}_1$ 中的每一点都均方连续，则称$\{\xi_t\}$在 $\mathbf{R}_1$ 上均方连续.

定理 11.5.3 $\{\xi_t\}$在 t 点均方连续的充分必要条件为其相关函数 $K(t_1,t_2)$在点(t,t)连续.

证 先证充分性.因为

专业课辅导+真题资料+微信公众号：432应统书院

$$
\begin{aligned}
E|\xi_{t+h}-\xi_t|^2 &= E(\xi_{t+h}\xi_{t+h})-E(\xi_{t+h}\xi_t)-E(\xi_t\xi_{t+h})+E(\xi_t\xi_t) \\
&= K(t+h,t+h)-K(t+h,t)- \\
&\quad K(t,t+h)+K(t,t)\to 0(h\to 0).
\end{aligned}
$$

故充分性得证.

必要性的证明,可由定理 11.5.2 之(6)得到:

$$
\begin{aligned}
\lim_{h,h'\to 0} K(t+h,t+h') &= \lim_{h,h'\to 0} E(\xi_{t+h}\xi_{t+h'}) \\
&= E(\xi_t\xi_t)=K(t,t).
\end{aligned}
$$

推论 $\{\xi_t\}$的相关函数 $K(t_1,t_2)$在整个平面上每一点(t_1,t_2)连续的充要条件为它在对角线 $t_1=t_2$上连续.

证 必要性是显然的,现证充分性. 由定理 11.5.3 知$\{\xi_t\}$在每一点 $t\in\mathbf{R}_1$均方连续,再由定理 11.5.2 之(6),得

$$
\begin{aligned}
\lim_{h,h'\to 0} K(t_1+h,t_2+h') &= \lim_{h,h'\to 0} E(\xi_{t_1+h}\xi_{t_2+h'}) \\
&= E(\xi_{t_1}\xi_{t_2}) \\
&= K(t_1,t_2).
\end{aligned}
$$

定理 11.5.3 说明随机过程的均方连续性与它的相关函数的连续性是等价的,而定理 11.5.3 的推论则指出了相关函数关于连续性的一个有趣事实:只需要相关函数在对角线 $t_1=t_2$的每一点上连续,则这个相关函数必在整个平面 $\mathbf{R}_1\times\mathbf{R}_1$上连续.

三、随机过程的均方积分

设$\{\xi_t\}$为一随机过程,$a,b\in\mathbf{R}_1$,$a<b$. 将区间$[a,b]$作分划如下:

$$a=t_0<t_1<\cdots<t_n=b.$$

又

$$\Delta=\max_{0\le i\le n-1}\{(t_{i+1}-t_i)\},$$

$$t_i\le t'_i<t_{i+1}(i=0,1,\cdots,n-1).$$

定义 11.5.3 若按照上述记号,和式

$$\sum_{i=0}^{n-1}\xi_{t'_i}(t_{i+1}-t_i)$$

专业课辅导+真题资料+微信公众号：432应统书院

在均方意义下存在有限极限，则说$\{\xi_t\}$在区间$[a,b]$上均方黎曼可积分（或简称均方可积）. 其极限值称为$\{\xi_t\}$在区间$[a,b]$上的均方积分，并以$\int_a^b \xi_t \mathrm{d}t$表示，即

$$\lim_{\Delta\to 0} E\left|\sum_{i=0}^{n-1}\xi_{t_i'}(t_{i+1}-t_i)-\int_a^b \xi_t \mathrm{d}t\right|^2=0,$$

或

$$\underset{\Delta\to 0}{\mathrm{l.i.m}}\sum_{i=0}^{n-1}\xi_{t_i'}(t_{i+1}-t_i)=\int_a^b \xi_t \mathrm{d}t.$$

由定理 11.5.2 之(5)立即得到，若$\{\xi_t\}$在$[a,b]$上均方可积，则有

$$E\left[\int_a^b \xi_t \mathrm{d}t\right]=\int_a^b E(\xi_t)\mathrm{d}t,$$

即求数学期望和求均方积分可调换次序.

定理 11.5.4 随机过程$\{\xi_t\}$在$[a,b]$上均方可积的充分条件为$\{\xi_t\}$的相关函数$K(t_1,t_2)$在矩形$[a,b]\times[a,b]$上依普通意义上黎曼可积.

证 令对区间$[a,b]$的另一分划为

$$a\leqslant \tilde{t}_0<\tilde{t}_1<\cdots<\tilde{t}_m=b.$$

又令

$$\tilde{\Delta}=\max_{0\leqslant k\leqslant m-1}\{(\tilde{t}_{k+1}-\tilde{t}_k)\},$$

$$\tilde{t}_k\leqslant \tilde{t}_k'<\tilde{t}_{k+1}\quad (k=0,1,\cdots,m-1).$$

只需证明，当$\Delta,\tilde{\Delta}\to 0$时，

$$E\left|\sum_{i=0}^{n-1}\xi_{t_i'}(t_{i+1}-t_i)-\sum_{k=0}^{m-1}\xi_{\tilde{t}_k'}(\tilde{t}_{k+1}-\tilde{t}_k)\right|^2\to 0.$$

现因为

$$E\left|\sum_{i=0}^{n-1}\xi_{t_i'}(t_{i+1}-t_i)-\sum_{k=0}^{m-1}\xi_{\tilde{t}_k'}(\tilde{t}_{k+1}-\tilde{t}_k)\right|^2$$

$$=E\left[\sum_{i=0}^{n-1}\xi_{t_i'}(t_{i+1}-t_i)\cdot\sum_{j=0}^{n-1}\xi_{t_j'}(t_{j+1}-t_j)\right]$$

专业课辅导+真题资料+微信公众号：432应统书院

$$+E\left[\sum_{j=0}^{m-1}\xi_{\tilde{t}_j}(\tilde{t}_{j+1}-\tilde{t}_j)\cdot\sum_{k=0}^{m-1}\xi_{\tilde{t}_k}(\tilde{t}_{k+1}-\tilde{t}_k)\right]$$

$$-2E\left[\sum_{i=0}^{n-1}\xi_{t_i}(t_{i+1}-t_i)\cdot\sum_{k=0}^{m-1}\xi_{\tilde{t}_k}(\tilde{t}_{k+1}-\tilde{t}_k)\right]$$

$$=\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}E(\xi_{t_i}\xi_{t_j})(t_{i+1}-t_i)(t_{j+1}-t_j)$$

$$+\sum_{j=0}^{m-1}\sum_{k=0}^{m-1}E(\xi_{\tilde{t}_j}\xi_{\tilde{t}_k})(\tilde{t}_{j+1}-\tilde{t}_j)(\tilde{t}_{k+1}-\tilde{t}_k)$$

$$-2\sum_{i=0}^{n-1}\sum_{k=0}^{m-1}E(\xi_{t_i}\xi_{\tilde{t}_k})(t_{i+1}-t_i)(\tilde{t}_{k+1}-\tilde{t}_k).$$

由此式可见，当 $K(t_1,t_2)$ 于 $[a,b]\times[a,b]$ 上黎曼可积时，上式右边当 $\Delta,\tilde{\Delta}\to 0$ 时趋于 0，再由定理 11.5.1 及定义 11.5.3 可得 $\{\xi_t\}$ 均方可积.

推论 若 $\{\xi_t\}$ 于 $[a,b]$ 均方连续，则必均方可积.

证 由假设及定理 11.5.3 可得 $\{\xi_t\}$ 的相关函数 $K(t_1,t_2)$ 在 $[a,b]\times[a,b]$ 上连续，故 $K(t_1,t_2)$ 于 $[a,b]\times[a,b]$ 上黎曼可积，再由定理 11.5.4 即得证.

对于均方连续于区间 $[a,b]$ 的随机过程，其均方积分有如下性质（由于证明方法与普通数学分析中关于积分性质的证明方法类似，读者可自行补足其证明）：

性质 11.5.1 设随机过程 $\{\xi_t\}$ 与 $\{\eta_t\}$ 皆于 $[a,b]$ 均方连续，则

（1）成立不等式：

$$E\left|\int_a^b\xi_t\mathrm{d}t\right|^2\leqslant M(b-a)^2,$$

其中

$$M=\max_{a\leqslant t\leqslant b}E|\xi_t|^2;$$

（2）对 $a<b<c$，有

$$\int_a^c\xi_t\mathrm{d}t=\int_a^b\xi_t\mathrm{d}t+\int_b^c\xi_t\mathrm{d}t;$$

专业课辅导+真题资料+微信公众号：432应统书院

(3) 对常数 α,β,有

$$\int_a^b(\alpha\xi_t+\beta\eta_t)\mathrm{d}t=\alpha\int_a^b\xi_t\mathrm{d}t+\beta\int_a^b\eta_t\mathrm{d}t.$$

下面我们考虑随机过程$\{\xi_t\}$的不定积分.

若$\{\xi_t\}$于$[a,b]$上均方连续,令

$$\eta_t=\int_a^t\xi_s\mathrm{d}s.$$

称 η_t为 ξ_t的不定积分(在均方意义下). 由上述定理 11.5.4 的推论可知 η_t对任意 $t\in[a,b]$总是存在的. 我们可以证明下面的定理成立.

定理 11.5.5 若随机过程$\{\xi_t\}$于$[a,b]$均方连续,则它的不定积分 η_t所构成的随机过程$\{\eta_t\}$在$[a,b]$上均方连续.

证 由性质 11.5.1 得,

$$\begin{aligned}E|\eta_{t+h}-\eta_t|^2&=E\left|\int_a^{t+h}\xi_s\mathrm{d}s-\int_a^t\xi_s\mathrm{d}s\right|^2\\&=E\left|\int_t^{t+h}\xi_s\mathrm{d}s\right|^2\\&\leqslant h^2\max_{a\leqslant t\leqslant b}E|\xi_t|^2\to0(\text{当 }h\to0).\end{aligned}$$

四、随机过程的均方导数

定义 11.5.4 设$\{\xi_t\}$为一随机过程,若存在随机过程$\{\xi'_t\}$,使得对 $t\in\mathbf{R}_1$,

$$\underset{h\to0}{\mathrm{l.i.m}}\left(\frac{\xi_{t+h}-\xi_t}{h}\right)=\xi'_t,$$

则称$\{\xi_t\}$于 t 点均方可微,且称 ξ'_t 为$\{\xi_t\}$在 t 点的均方导数. 若对每一 $t\in\mathbf{R}_1$,$\{\xi_t\}$都均方可微,则称$\{\xi_t\}$在 $\mathbf{R}_1$上均方可微.

定理 11.5.6 随机过程$\{\xi_t\}$在 $\mathbf{R}_1$上均方可微的充分必要条

专业课辅导+真题资料+微信公众号：432应统书院

件为它的相关函数 $K(t_1,t_2)$ 的广义二阶导数 $\frac{\partial^2 K(t_1,t_2)}{\partial t_1 \partial t_2}$ ① 对于对角线上的每一点 $(t,t)\in \mathbf{R}_1\times\mathbf{R}_1$ 上存在.

证 先证充分性. 因为

$$E\left|\frac{\xi_{t+h}-\xi_t}{h}-\frac{\xi_{t+h'}-\xi_t}{h'}\right|^2$$
$$=E\left|\frac{\xi_{t+h}-\xi_t}{h}\right|^2+E\left|\frac{\xi_{t+h'}-\xi_t}{h'}\right|^2$$
$$-2E\left(\frac{\xi_{t+h}-\xi_t}{h}\cdot\frac{\xi_{t+h'}-\xi_t}{h'}\right).$$

由 $\frac{\partial^2 K(t_1,t_2)}{\partial t_1 \partial t_2}$ 于 (t,t) 存在这一假设得

$$E\left(\frac{\xi_{t+h}-\xi_t}{h}\cdot\frac{\xi_{t+h'}-\xi_t}{h'}\right)$$
$$=\frac{1}{hh'}[K(t+h,t+h')-K(t,t+h')$$
$$-K(t+h,t)+K(t,t)]$$
$$\to\frac{\partial^2 K(t,t)}{\partial t\partial t}\quad(h,h'\to 0).$$

同理可得

$$E\left|\frac{\xi_{t+h}-\xi_t}{h}\right|^2=E\left(\frac{\xi_{t+h}-\xi_t}{h}\right)\left(\frac{\xi_{t+h}-\xi_t}{h}\right)$$
$$\to\frac{\partial^2 K(t,t)}{\partial t\partial t}\quad(h\to 0)$$

及

---

① 若二元函数 $f(x_1,x_2)$ 有如下极限存在:

$$\lim_{\substack{h\to 0\\ h'\to 0}}\frac{f(x_1+h,x_2+h')-f(x_1+h,x_2)-f(x_1,x_2+h')+f(x_1,x_2)}{hh'},$$

则说 $f(x_1,x_2)$ 在点 (x_1,x_2) 的广义二阶导数存在,且称极限为 $f(x_1,x_2)$ 在点 (x_1,x_2) 的广义二阶导数.

专业课辅导+真题资料+微信公众号：432应统书院

$$E\left|\frac{\xi_{t+h'}-\xi_t}{h'}\right|^2=E\left(\frac{\xi_{t+h'}-\xi_t}{h'}\right)\left(\frac{\xi_{t+h'}-\xi_t}{h'}\right)$$

$$\to\frac{\partial^2 K(t,t)}{\partial t\partial t}\quad(h'\to 0).$$

因此可得

$$E\left|\frac{\xi_{t+h}-\xi_t}{h}-\frac{\xi_{t+h'}-\xi_t}{h'}\right|^2\to 0\quad(h,h'\to 0).$$

根据定理 11.5.1 即可得$\{\xi_t\}$于$t\in\mathbf{R}_1$可微.

必要性证明如下：

设$\{\xi_t\}$于$t\in\mathbf{R}_1$的导数为ξ'_t,则由定理 11.5.2 中之(6)有

$$\lim_{h,h'\to 0}E\left(\frac{\xi_{t+h}-\xi_t}{h}\cdot\frac{\xi_{t+h'}-\xi_t}{h'}\right)=E(\xi'_t\xi'_t).$$

由充分性的证明过程知,上式左端等于

$$\lim_{h,h'\to 0}\frac{1}{hh'}[K(t+h,t+h')-K(t,t+h')$$

$$-K(t+h,t)+K(t,t)].$$

因此可得

$$\lim_{h,h'\to 0}\frac{1}{hh'}[K(t+h,t+h')-K(t,t+h')$$

$$-K(t+h,t)+K(t,t)]=E(\xi'_t\xi'_t).$$

即$K(t_1,t_2)$于$(t,t)\in\mathbf{R}_1\times\mathbf{R}_1$的广义二阶导数$\dfrac{\partial^2 K(t,t)}{\partial t\partial t}$存在.

推论 若$\dfrac{\partial^2 K(t_1,t_2)}{\partial t_1\partial t_2}$在对角线点$(t,t)\in\mathbf{R}_1\times\mathbf{R}_1$存在,则对任意$(t_1,t_2)\in\mathbf{R}_1\times\mathbf{R}_1$存在,且

$$\frac{\partial^2 K(t_1,t_2)}{\partial t_1\partial t_2}=E(\xi'_{t_1}\xi'_{t_2}).$$

证 由题设及定理 11.5.6 得知,$\{\xi_t\}$于$t\in\mathbf{R}_1$上可微,即对$t_1,t_2\in\mathbf{R}_1$,有

专业课辅导+真题资料+微信公众号：432应统书院

$$\underset{h\to 0}{\mathrm{l.i.m}}\frac{\xi_{t_1+h}-\xi_{t_1}}{h}=\xi'_{t_1},$$

$$\underset{h'\to 0}{\mathrm{l.i.m}}\frac{\xi_{t_2+h'}-\xi_{t_2}}{h'}=\xi'_{t_2}.$$

再由定理 11.5.2 中之(6)可得

$$E(\xi'_{t_1}\xi'_{t_2})=\lim_{h,h'\to 0}E\left(\frac{\xi_{t_1+h}-\xi_{t_1}}{h}\cdot\frac{\xi_{t_2+h'}-\xi_{t_2}}{h'}\right)$$
$$=\frac{\partial^2 K(t_1,t_2)}{\partial t_1\partial t_2}.$$

均方导数有很多类似普通函数的导数的性质，例如：

性质 11.5.2

(1) 任一随机变量 ξ(可以是常数)的均方导数为 0；

(2) $(\xi_t+\xi)'=\xi'_t$；由此可得，若两个随机过程的均方导数相等，则它们相差一个随机变量(可以是常数).

(3) 若$\{\xi_t\}$与$\{\eta_t\}$于 $t\in\mathbf{R}_1$ 均方可微，α,β 为常数，那么$\{\alpha\xi_t+\beta\eta_t\}$于 $t\in\mathbf{R}_1$ 均方可微，且

$$(\alpha\xi_t+\beta\eta_t)'=\alpha\xi'_t+\beta\eta'_t.$$

定理 11.5.7 若$\{\xi_t\}$于 $\mathbf{R}_1$ 均方可微，则对任意 $t_1,t_2\in\mathbf{R}_1$ 有

(1) $\dfrac{\mathrm{d}}{\mathrm{d}t_1}E(\xi_{t_1})=E(\xi'_{t_1})$；

(2) $\dfrac{\partial K(t_1,t_2)}{\partial t_1}=\dfrac{\partial}{\partial t_1}E(\xi_{t_1}\xi_{t_2})=E(\xi'_{t_1}\xi_{t_2})$；

(3) $\dfrac{\partial^2 K(t_1,t_2)}{\partial t_1\partial t_2}=\dfrac{\partial}{\partial t_1}\dfrac{\partial}{\partial t_2}E(\xi_{t_1}\xi_{t_2})=E(\xi'_{t_1}\xi'_{t_2})$.

证 (1) 由定理 11.5.2 之(5)得

$$E(\xi'_{t_1})=\lim_{h\to 0}E\left(\frac{\xi_{t_1+h}-\xi_{t_1}}{h}\right)=\lim_{h\to 0}\frac{1}{h}[E(\xi_{t_1+h})-E(\xi_{t_1})]$$
$$=[E(\xi_{t_1})]',$$

即

专业课辅导+真题资料+微信公众号：432应统书院

$$\frac{\mathrm{d}}{\mathrm{d}t_1}E(\xi_{t_1})=E(\xi'_{t_1}).$$

(2) 因为

$$\frac{\xi_{t_1+h}\xi_{t_2}-\xi_{t_1}\xi_{t_2}}{h}=\frac{\xi_{t_1+h}-\xi_{t_1}}{h}\xi_{t_2},$$

由定理 11.5.2 之(6)得

$$\lim_{h\to 0}E\left(\frac{\xi_{t_1+h}-\xi_{t_1}}{h}\xi_{t_2}\right)=E(\xi'_{t_1}\xi_{t_2}),$$

而上式左端等于

$$\lim_{h\to 0}\frac{1}{h}[K(t_1+h,t_2)-K(t_1,t_2)].$$

因而得

$$\frac{\partial K(t_1,t_2)}{\partial t_1}=E(\xi'_{t_1}\xi_{t_2}).$$

(3) 的证明同上，请读者自行补足.

定理 11.5.7 说明，当$\{\xi_t\}$在$\mathbf{R}_1$上均方可微时，则求数学期望与求导数可交换次序.

定理 11.5.8 设随机过程$\{\xi_t\}$于区间$[a,b]$上均方连续，则它的不定积分

$$\eta_t=\int_a^t \xi_s\mathrm{d}s,t\in[a,b]$$

在$[a,b]$上均方可微，且

$$\eta'_t=\xi_t.$$

证 由假设及性质 11.5.1 可得

$$\begin{aligned}E\left|\frac{\eta_{t+h}-\eta_t}{h}-\xi_t\right|^2&=E\left|\frac{1}{h}\int_t^{t+h}\xi_s\mathrm{d}s-\xi_t\right|^2\\&=\frac{1}{h^2}E\left|\int_t^{t+h}(\xi_s-\xi_t)\mathrm{d}s\right|^2\\&\leqslant\max_{t\leqslant s\leqslant t+h}E|\xi_s-\xi_t|^2\to 0\quad(h\to 0).\end{aligned}$$

专业课辅导+真题资料+微信公众号：432应统书院

由此得到$\{\eta_t\}$于$[a,b]$均方可微，且对任意$t\in[a,b]$，

$$\eta'_t=\xi_t.$$

推论 设$\{\xi_t\}$于$[a,b]$均方可微，且$\{\xi'_t\}$均方连续，则对任意$t\in[a,b]$，有

$$\xi_t-\xi_a=\int_0^t \xi'_s\mathrm{d}s,$$

特别当$t=b$时，有

$$\xi_b-\xi_a=\int_a^b \xi'_s\mathrm{d}s.$$

证 事实上，由定理11.5.8及性质11.5.1和性质11.5.2即得.

五、随机微分方程

利用数学分析来描述某些自然现象时，通常要建立一个微分方程，并给定初始条件或边界条件. 例如描述一个线性系统输入与输出关系可有如下微分方程及初始条件：

$$a_n y^{(n)}+\cdots+a_1 y'+a_0 y=x(t),$$
$$y(0)=b.$$

但考虑到随机因素的影响，例如初始条件的轻微改变、由测量误差带来的常系数改变或$x(t)$本身是一随机过程（如$x(t)$表示作布朗运动的质点受到随机碰撞）. 这都使得方程的解$y(t)$不可能是一决定性函数而为一随机过程. 这些事实促使建立了近代概率论的一个重要分支——随机微分方程.

考虑一个微分方程

$$a_n\xi_t^{(n)}+a_{n-1}\xi_t^{(n-1)}+\cdots+a_1\xi'_t+a_0\xi_t=\eta_t, \tag{11.5.3}$$

若$a_k(k=0,1,\cdots,n)$是随机变量，$\{\xi_t\}$，$\{\eta_t\}$是随机过程，则称上述微分方程为随机微分方程. 自然，上述之a_k，$\{\xi_t\}$或$\{\eta_t\}$可以有一个或一个以上是普通的常数或函数，假若全部都是普通的常数或函数，则(11.5.3)式就是熟知的微分方程.

随机微分方程可以用来描述一个系统的输入与输出关系，若

专业课辅导+真题资料+微信公众号：432应统书院

把$\{\xi_t\}$看作是一个输入信号(它是随机的),则$\{\eta_t\}$可视为通过该系统的输出.因此,在控制论、滤波及通讯理论方面都需要用(11.5.3)式所表达的随机微分方程来作为重要的数学工具.现在随机微分方程不论在理论方面还是在应用方面都得到了迅速的发展,读者如需较系统的学习,可参看比较基本的教材,如[20],[25].我们在下面只介绍一个简单例子,建立一些概念.

在(11.5.3)式中,我们考虑当$n=1$,a_0及a_1皆为常数的情形,即考虑一阶常系数随机微分方程:

$$a_1\xi'_t+a_0\xi_t=\eta_t \quad t\in[0,\infty), \tag{11.5.4}$$

给定初始条件

$$\xi_0=0(\text{以概率为 }1).$$

下面我们讨论$\{\xi_t\}$与$\{\eta_t\}$之间的一些特征的关系(这里主要介绍数学期望与相关函数),假定它们的二阶矩都存在.

1. $\{\xi_t\}$与$\{\eta_t\}$的数学期望的关系

对(11.5.4)式两边求数学期望,并根据定理11.5.7之(1)得

$$\begin{aligned}E(\eta_t)&=a_1E(\xi'_t)+a_0E(\xi_t)\\&=a_1E'(\xi_t)+a_0E(\xi_t).\end{aligned} \tag{11.5.5}$$

此时,初始条件为

$$E(\xi_0)=0.$$

(11.5.5)式是普通的一阶常微分方程,由熟知的解法,可得$E(\xi_t)$与$E(\eta_t)$的关系.换句话说,我们若能知道输入$\{\xi_t\}$的数学期望,则可通过解(11.5.5)式得到输出$\{\eta_t\}$的数学期望.反之,已知$E(\eta_t)$,可求$E(\xi_t)$.

2. $\{\xi_t\}$与$\{\eta_t\}$的相关函数的关系

设$t_1,t_2\in[0,\infty)$,由(11.5.4)式:

$$a_1\xi'_{t_2}+a_0\xi_{t_2}=\eta_{t_2},$$

两边乘以η_{t_1}得

$$a_1\eta_{t_1}\xi'_{t_2}+a_0\eta_{t_1}\xi_{t_2}=\eta_{t_1}\eta_{t_2},$$

专业课辅导+真题资料+微信公众号：432应统书院

上式两边求数学期望得

$$a_1E(\eta_{t_1}\xi'_{t_2})+a_0E(\eta_{t_1}\xi_{t_2})=E(\eta_{t_1}\eta_{t_2}).$$

由定理 11.5.7 之(2)，上式为

$$a_1\frac{\partial K_{\eta,\xi}(t_1,t_2)}{\partial t_2}+a_0K_{\eta,\xi}(t_1,t_2)=K_\eta(t_1,t_2). \qquad (11.5.6)$$

此时的初始条件为

$$K_{\eta,\xi}(t_1,0)=0.$$

用同样方法，若对(11.5.4)式令 $t=t_1$，且两端乘以 ξ_{t_2}，也可得到类似的常微分方程：

$$a_1\frac{\partial K_\xi(t_1,t_2)}{\partial t_1}+a_0K_\xi(t_1,t_2)=K_{\eta,\xi}(t_1,t_2), \qquad (11.5.7)$$

初始条件为

$$K_\xi(0,t_2)=0.$$

我们联立方程(11.5.6)式及(11.5.7)式，则可解得$\{\xi_t\}$与$\{\eta_t\}$的相关函数之间的关系.

例 11.5.1 考虑随机微分方程

$$\begin{cases}\xi'_t+\alpha\xi_t=\eta_t, & t\in[0,\infty),\\ \xi_0=0,\end{cases} \qquad (11.5.8)$$

其中 $\alpha>0$，$\{\eta_t\}$为弱平稳过程，且

$$E(\eta_t)=\lambda, B_\eta(\tau)=e^{-2\lambda|\tau|}, \qquad (11.5.9)$$

其中 $\lambda>0$.

由(11.5.5)，(11.5.8)两式及上式可得

$$\begin{cases}E'(\xi_t)+\alpha E(\xi_t)=\lambda,\\ E(\xi_0)=0.\end{cases}$$

解上述微分方程得

$$E(\xi_t)=\frac{\lambda}{\alpha}(1-e^{-\alpha t}).$$

$E(\xi_t)$与$E(\eta_t)$的关系如图 11.5.1 所示.

由(11.5.6)及(11.5.9)两式得

专业课辅导+真题资料+微信公众号：432应统书院

图 11.5.1

$$\frac{\partial K_{\eta,\xi}(t_1,t_2)}{\partial t_2}+\alpha K_{\eta,\xi}(t_1,t_2)=e^{-2\lambda|t_1-t_2|},$$

此时的初始条件由(11.5.8)式可知为

$$K_{\eta,\xi}(t_1,0)=0.$$

因而可解得当 $t_1>t_2$ 时，

$$K_{\eta,\xi}(t_1,t_2)=\frac{1}{2\lambda+\alpha}e^{-2\lambda t_1}[e^{2\lambda t_2}-e^{-\alpha t_2}],$$

将上式代入(11.5.7)式得

$$\frac{\partial K_{\xi}(t_1,t_2)}{\partial t_1}+\alpha K_{\xi}(t_1,t_2)=\frac{1}{2\lambda+\alpha}e^{-2\lambda t_1}[e^{2\lambda t_2}-e^{-\alpha t_2}].$$

此时初始条件为 $K_{\xi}(0,t_2)=0$，由此解得

$$K_{\xi}(t_1,t_2)=\frac{1}{(\alpha+2\lambda)(\alpha-2\lambda)}(e^{2\lambda t_2}-e^{-\alpha t_2})(e^{-2\lambda t_1}-e^{-\alpha t_1}).$$

对于 $t\leqslant t_2$ 时，可类似求解.

3. 随机微分方程的解

下面只给出一个例子.

例 11.5.2 设随机微分方程为

$$\xi'_t=\eta_t,\quad t\in[a,b],\tag{11.5.10}$$

初始条件为

专业课辅导+真题资料+微信公众号：432应统书院

$$\xi_a=\xi_0,$$

其中$\{\eta_t\}$于$[a,b]$均方可积，ξ_0是数学期望有限的随机变量.

解 由定理 11.5.8 的推论得，当$t\in[a,b]$，有

$$\xi_t-\xi_a=\int_a^t \eta_s\mathrm{d}s,$$

由初始条件得

$$\xi_t=\xi_0+\int_a^t \eta_s\mathrm{d}s.$$

$\{\xi_t\}$的数学期望为

$$E(\xi_t)=E(\xi_0)+\int_a^t E(\eta_s)\mathrm{d}s.$$

如若设ξ_0服从正态$N(0,1)$，且$\{\eta_t\}$为例 11.5.1 所述的弱平稳过程，则有

$$E(\xi_t)=\int_a^t \lambda\mathrm{d}s=\lambda(t-a).$$

§11.6 平稳随机过程

一、定义及例

定义 11.6.1 设$\{\xi_t:t\in\mathbf{R}_1\}$为一随机过程，若对任意正整数$n$及实数$r,t_1,t_2,\cdots,t_n$，随机变量$\xi_{t_1},\xi_{t_2},\cdots,\xi_{t_n}$的$n$维分布函数与$\xi_{t_1+r},\xi_{t_2+r},\cdots,\xi_{t_n+r}$的$n$维分布函数相同，即

$$F_{t_1,t_2,\cdots,t_n}(x_1,x_2,\cdots,x_n)=F_{t_1+r,t_2+r,\cdots,t_n+r}(x_1,x_2,\cdots,x_n),\tag{11.6.1}$$

则称$\{\xi_t\}$为强平稳过程. 换句话说，若随机过程$\{\xi_t\}$的任意有限维分布函数沿时间轴作平移时不改变，则称$\{\xi_t\}$为强平稳过程.

定义 11.6.2 若随机过程$\{\xi_t:t\in\mathbf{R}_1\}$满足如下两个条件：

$$\begin{cases}\text{(i) 对任意 } t\in\mathbf{R}_1, E(\xi_t)^2<\infty;\\ \text{(ii) } K(t_1,t_2)=B(t_1-t_2),\end{cases}\tag{11.6.2}$$

则称$\{\xi_t\}$为弱平稳过程.

$K(t_1,t_2)=B(t_1-t_2)$等价于$K(t_1+\tau,t_2+\tau)=K(t_1,t_2)$，或者

专业课辅导+真题资料+微信公众号：432应统书院

$K(t+\tau,t)=B(\tau)$，其中 $\tau\in\mathbf{R}_1$.

显然，若$\{\xi_t\}$为强平稳过程，且 $E|\xi_t|^2<\infty$，则$\{\xi_t\}$必为弱平稳过程，事实上，对任意 t_1,t_2及 $\tau\in\mathbf{R}_1$，

$$\begin{aligned}K(t_1+\tau,t_2+r)&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x_1x_2\mathrm{d}F_{t_1+\tau,t_2+\tau}(x_1,x_2)\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x_1x_2\mathrm{d}F_{t_1,t_2}(x_1,x_2)\\&=K(t_1,t_2).\end{aligned}$$

上述命题的逆命题则不一定成立. 但是，当$\{\xi_t\}$为零均值的正态随机过程时，则强平稳过程必为弱平稳过程，反之亦然. 这从(11.2.1)式可以得到.

在许多问题中，我们还常见到所谓平稳(强或弱)增量过程，它的定义如下：

若对每一 $t,h\in\mathbf{R}_1$，随机过程$\{\xi_t\}$的增量$\{\xi_{t+h}-\xi_t\}$为平稳(强或弱)过程，则称$\{\xi_t\}$为平稳(强或弱)增量过程.

例 11.6.1 考虑一个具有随机相位的余弦波，它由如下定义的随机过程描述，

$$\xi_t=\cos(\lambda t+\theta)\quad(t\in\mathbf{R}_1),$$

其中 λ 是常数，θ 为一个具有有限二阶矩的随机变量，则可求得其数学期望和相关函数如下：

$$\begin{aligned}E(\xi_t)&=E[\cos(\lambda t+\theta)]\\&=E(\cos\lambda t\cos\theta-\sin\lambda t\sin\theta)\\&=\cos\lambda tE(\cos\theta)-\sin\lambda tE(\sin\theta);\end{aligned}$$

$$\begin{aligned}K(t_1,t_2)&=E[\cos(\lambda t_1+\theta)\cos(\lambda t_2+\theta)]\\&=E\left[\frac{1}{2}\cos(\lambda t_1-\lambda t_2)+\frac{1}{2}\cos(\lambda t_1+\lambda t_2+2\theta)\right]\\&=\frac{1}{2}\cos\lambda(t_1-t_2)+\frac{1}{2}[\cos\lambda(t_1+t_2)E(\cos 2\theta)-\\&\quad\sin\lambda(t_1+t_2)E(\sin 2\theta)].\end{aligned}$$

当假定 θ 服从$(-\pi,\pi)$上的均匀分布时，则此时$\{\xi_t\}$为弱平稳

专业课辅导+真题资料+微信公众号：432应统书院

过程. 因为此时 $E(\cos 2\theta)=E(\sin 2\theta)=0$,因此

$$K(t_1,t_2)=\frac{1}{2}\cos\lambda(t_1-t_2).$$

例 11.6.2 考虑一个具有随机振幅的正弦波,它由如下定义的随机过程描述:

$$\xi_t=\xi\cos 2\pi t+\eta\sin 2\pi t \quad (t\in\mathbf{R}_1).$$

其中 ξ,η 为两个随机变量,且满足

$$E(\xi)=E(\eta)=0;$$
$$D(\xi)=D(\eta)=1;$$
$$E(\xi\eta)=0.$$

则可计算得

$$E(\xi_t)=\cos 2\pi tE(\xi)+\sin 2\pi tE(\eta)=0,$$

$$\begin{aligned}K(t_1,t_2)&=\cos 2\pi t_1\cos 2\pi t_2E(\xi^2)+\sin 2\pi t_1\sin 2\pi t_2E(\eta^2)\\&\quad+\cos 2\pi t_1\sin 2\pi t_2E(\xi\eta)+\sin 2\pi t_1\cos 2\pi t_2E(\xi\eta)\\&=\cos 2\pi(t_1-t_2).\end{aligned}$$

由此可得$\{\xi_t\}$是一个弱平稳过程.

二、相关函数

下面如无特别声明,平稳过程均指弱平稳过程而言. 由(11.6.2)式知,平稳过程的相关函数可用 $B(\tau)$ 表示,即

$$B(\tau)=E(\xi_{t+\tau}\xi_t). \tag{11.6.3}$$

平稳过程的相关函数 $B(\tau)$ 有如下性质:

性质 11.6.1

1) $B(0)\geqslant 0$;

2) $B(-\tau)=B(\tau)$;

3) $|B(\tau)|\leqslant B(0)$;

4) $B(\tau)$ 于 $\mathbf{R}_1$ 连续的充分必要条件为 $B(\tau)$ 于 $\tau=0$ 点连续;

5) $B(\tau)$ 是非负定的. 即对任意 $2n$ 个实数 $a_1,a_2,\cdots,a_n$ 及 $\tau_1,\tau_2,\cdots,\tau_n$,有

专业课辅导+真题资料+微信公众号：432应统书院

$$\sum_{i,j=1}^{n} B(\tau_i-\tau_j)a_i a_j \geqslant 0.$$

证 1),2)的证明可由(11.6.3)式立即得到.

现证3).由于

$$0 \leqslant E|\xi_{t+\tau} \pm \xi_t|^2 = 2[B(0) \pm B(\tau)],$$

因此得

$$B(0) \pm B(\tau) \geqslant 0.$$

即得

$$|B(\tau)| \leqslant B(0).$$

4)的证明由下面的计算得到.因为

$$\begin{aligned}
&|B(\tau+\Delta\tau)-B(\tau)|^2 \\
&= |E(\xi_{\tau+\Delta\tau+\tau'}\xi_{\tau'})-E(\xi_{\tau+\Delta\tau+\tau'}\xi_{\Delta\tau+\tau'})|^2 \\
&= |E[\xi_{\tau+\Delta\tau+\tau'}(\xi_{\tau'}-\xi_{\Delta\tau+\tau'})]|^2 \\
&\leqslant E|\xi_{\tau+\Delta\tau+\tau'}|^2 E|\xi_{\tau'}-\xi_{\Delta\tau+\tau'}|^2 \\
&= 2B(0)[B(0)-B(\Delta\tau)] \to 0 \quad (\text{当 } \Delta\tau \to 0).
\end{aligned}$$

因此,充分性得证.必要性是显然的.

5)的证明可从(11.1.9)直接得到.证毕.

读者可以证明,性质11.6.1中之1),2),3)可直接从5)得到.

到现在为止,我们考虑的随机过程均为实的随机过程,在理论研究和应用中,还须考虑复的随机过程.

$\{\xi_t\}$称为复的随机过程是指ξ_t可表为

$$\xi_t = X_t + \mathrm{j}Y_t,$$

其中X_t,Y_t均为实的随机过程.对于复随机过程的数学期望和相关函数定义如下:

$$E(\xi_t) = E(X_t) + \mathrm{j}E(Y_t);$$

$$K(t_1,t_2) = E(\xi_{t_1}\bar{\xi}_{t_2}) \quad (\bar{\xi}_{t_2}\text{表示}\xi_{t_2}\text{的共轭复数}).$$

若复随机过程$\{\xi_t\}$满足以下两式:

专业课辅导+真题资料+微信公众号：432应统书院

1) $E|\xi_t|^2<\infty$；

2) $K(t+\tau,t)=B(\tau)$（一般说，$B(\tau)$是实变复值函数），

则称$\{\xi_t\}$为复的弱平稳过程. 对于复的弱平稳过程，有类似于性质11.6.1的性质. 现只列在下面，证明方法与性质11.6.1证明相仿，这里从略.

性质 11.6.2

1) $B(0)\geqslant 0$；

2) $B(-\tau)=\overline{B(\tau)}$；

3) $|B(\tau)|\leqslant B(0)$；

4) $B(\tau)$于$\mathbf{R}_1$上连续的充分必要条件为$B(\tau)$于$\tau=0$点连续；

5) $B(\tau)$为非负定的，即对任意n个复数$z_1,z_2,\cdots,z_n$及n个实数$\tau_1,\tau_2,\cdots,\tau_n$，有

$$\sum_{i,j=1}^{n}B(\tau_i-\tau_j)z_i\bar{z}_j\geqslant 0.$$

在通信、控制、滤波理论和自然现象中，还必须考虑两个或两个以上的随机过程. 例如在通信中，信号在传输过程中受各种随机干扰的影响，接收端收到的是信号和噪声的叠加，信号和噪声一般来说都是随机过程. 对两个随机过程，我们定义一个描述它们之间相关程度的数字特征，称为互相关函数.

定义 11.6.3 设$\{\xi_t\}$，$\{\eta_t\}$是两个随机过程（复的或实的），称

$$K_{\xi,\eta}(t_1,t_2)=E(\xi_{t_1}\bar{\eta}_{t_2})\quad(t_1,t_2\in\mathbf{R}_1)$$

为$\{\xi_t\}$与$\{\eta_t\}$的互相关函数.

若对任意$t,\tau\in\mathbf{R}_1$，

$$K_{\xi,\eta}(t+\tau,t)=B_{\xi,\eta}(\tau),$$

则称$\{\xi_t\}$与$\{\eta_t\}$平稳相关.

对于平稳相关的两个随机过程，其相关函数$B_{\xi,\eta}(\tau)$有如下两个性质：

专业课辅导+真题资料+微信公众号：432应统书院

1) $\overline{B_{\xi,\eta}(\tau)}=B_{\eta,\xi}(-\tau)$;

2) $|B_{\xi,\eta}(\tau)|^2\leqslant B_{\xi}(0)B_{\eta}(0)$.

性质 11.6.3 设随机过程$\{\xi_t\}$定义如下:

$$\xi_t=\eta_t+\zeta_t,$$

其中$\{\eta_t\}$及$\{\zeta_t\}$为平稳相关的平稳过程(复的或实的),则$\{\xi_t\}$也是平稳过程,且其相关函数为

$$B_{\xi}(\tau)=B_{\eta}(\tau)+B_{\zeta}(\tau)+B_{\eta,\zeta}(\tau)+B_{\zeta,\eta}(\tau),\qquad(11.6.4)$$

若再设$\{\eta_t\}$与$\{\zeta_t\}$正交(若对任意$t_1,t_2\in\mathbf{R}_1$,$E(\eta_{t_1}\overline{\zeta_{t_2}})=0$,则称$\{\eta_t\}$与$\{\zeta_t\}$正交),则(11.6.4)式变为

$$B_{\xi}(\tau)=B_{\eta}(\tau)+B_{\zeta}(\tau).\qquad(11.6.5)$$

性质 11.6.3 的结果可直接从定义出发立即得到.

注:以上考虑了$T=\mathbf{R}_1$的情形.事实上,对$T=[0,\infty)$,$T=\{0,1,2,\cdots\}$,$T=\{0,\pm1,\pm2,\cdots\}$等等,上述定义、定理及性质仍成立,只需把"实数r"分别改为"$r\geqslant0$","非负整数r","整数r"等即可.

三、弱平稳随机过程的功率谱密度

在无线电技术和其他应用问题里,特别是在通信技术中,经常要考虑所谓功率、能量、功率谱密度.我们知道,表示一个信号的函数(例如电压的大小)可以定义它的功率、平均功率等.由于在我们所讨论的问题中,用于描述信号的函数是一个随机过程,我们需要类似地引入上述概念,下面我们只考虑关于功率谱概念.

定义 11.6.4 设$\{\xi_t\}$为一弱平稳过程,$B(\tau)$为其相关函数,当$B(\tau)$的傅氏变换

$$\int_{-\infty}^{\infty}e^{-i\omega\tau}B(\tau)\,d\tau①$$

存在时,我们称它为$\{\xi_t\}$的谱密度,并记为$S(\omega)$.即

① 这里的参数ω不是概率空间$(\Omega,\mathscr{F},P)$中的点,我们仅沿用许多书上的记号而已,学习时应加以注意.

专业课辅导+真题资料+微信公众号：432应统书院

$$S(\omega)=\int_{-\infty}^{\infty} e^{-j\omega\tau}B(\tau)\,d\tau. \tag{11.6.6}$$

$S(\omega)$的意义，我们将在稍后加以解释.

例 11.6.3 设$\{\xi_t\}$为例 11.1.7 所定义的随机过程，已知其相关函数为

$$B(\tau)=e^{-2\lambda|\tau|}.$$

根据(11.6.6)式可算得$\{\xi_t\}$的谱密度为

$$S(\omega)=\int_{-\infty}^{\infty} e^{-2\lambda|\tau|}e^{-j\omega\tau}\,d\tau=\frac{4\lambda}{4\lambda^2+\omega^2}.$$

$\{\xi_t\}$的相关函数与谱密度如图 11.6.1 之(a)及(b).

图 11.6.1

性质 11.6.4 谱密度 $S(\omega)$有如下性质：

1) $S(\omega)$是 ω 的实函数；

2) 当$\{\xi_t\}$为实随机过程时，$S(\omega)$关于 O 点对称，且

$$S(\omega)=2\int_0^{\infty} B(\tau)\cos\omega\tau\,d\tau;$$

3) 若弱平稳过程$\{\eta_t\}$与$\{\zeta_t\}$正交，则由 $\xi_t=\eta_t+\zeta_t$构成的随机过程$\{\xi_t\}$的谱密度为

$$S_\xi(\omega)=S_\eta(\omega)+S_\zeta(\omega),$$

其中 $S_\eta(\omega)$，$S_\zeta(\omega)$分别为$\{\eta_t\}$及$\{\zeta_t\}$的谱密度，皆假设存在.

证 1) 因为$\overline{B(\tau)}=B(-\tau)$，故

专业课辅导+真题资料+微信公众号：432应统书院

$$\overline{S(\omega)}=\int_{-\infty}^{\infty}\overline{B(\tau)}e^{j\omega\tau}d\tau=\int_{-\infty}^{\infty}B(-\tau)e^{j\omega\tau}d\tau$$

$$=\int_{-\infty}^{\infty}B(\tau)e^{-j\omega\tau}d\tau=S(\omega).$$

2）用与1）中同样的证明方法，可得

$$S(-\omega)=S(\omega),$$

并且由1）及 $B(\tau)$ 对称于 O 点可得

$$S(\omega)=\int_{-\infty}^{\infty}B(\tau)e^{-j\omega\tau}d\tau=\int_{-\infty}^{\infty}B(\tau)\cos\omega\tau d\tau$$

$$=2\int_{0}^{\infty}B(\tau)\cos\omega\tau d\tau.$$

3）由假设及(11.6.5)式，即可得所要结果. 证毕.

若 $S(\omega)$ 满足一定的条件$\left(如\int_{-\infty}^{\infty}|S(\omega)|d\omega<\infty\right)$时，则有

$$B(\tau)=\frac{1}{2\pi}\int_{-\infty}^{\infty}S(\omega)e^{j\omega\tau}d\omega, \tag{11.6.7}$$

上式当 $\tau=0$ 时，有

$$E|\xi_t|^2=B(0)=\frac{1}{2\pi}\int_{-\infty}^{\infty}S(\omega)d\omega,$$

我们称 $E|\xi_t|^2$ 为 $\{\xi_t\}$ 的平均功率. 它等于$\frac{1}{2\pi}\int_{-\infty}^{\infty}S(\omega)d\omega$.

类似地，可以对两个平稳相关的弱平稳过程 $\{\xi_t\}$ 与 $\{\eta_t\}$ 定义互谱密度如下：

$$S_{\xi,\eta}(\omega)=\int_{-\infty}^{\infty}B_{\xi,\eta}(\tau)e^{-j\omega\tau}d\tau$$

（当等式右边积分存在时）.

若 $\int_{-\infty}^{\infty}|S_{\xi,\eta}(\omega)|d\omega<\infty$ 时，有

$$B_{\xi,\eta}(\tau)=\frac{1}{2\pi}\int_{-\infty}^{\infty}S_{\xi,\eta}(\omega)e^{j\omega\tau}d\omega.$$

下面我们导出一些结果，说明 $S(\omega)$ 的物理意义

定义 11.6.5 设 $\{\xi_t\}$ 为一弱平稳过程，令

专业课辅导+真题资料+微信公众号：432应统书院

$$\tilde{S}_c(\omega)=\frac{1}{2c}\left|\int_{-c}^{c}\xi_t e^{-j\omega t}dt\right|^{2①}\quad(c>0),$$

若极限

$$\lim_{c\to\infty}\frac{1}{2c}E\left|\int_{-c}^{c}\xi_t e^{-j\omega t}dt\right|^2$$

存在，则称它为$\{\xi_t\}$的功率谱密度，并以$\tilde{S}(\omega)$表示，即

$$\tilde{S}(\omega)=\lim_{c\to\infty}\frac{1}{2c}E\left|\int_{-c}^{c}\xi_t e^{-j\omega t}dt\right|^2. \tag{11.6.8}$$

定理 11.6.1 若弱平稳过程$\{\xi_t\}$均方连续且它的相关函数绝对可积，即若

$$\int_{-\infty}^{\infty}|B(\tau)|d\tau<\infty,$$

则有

$$\tilde{S}(\omega)=S(\omega),$$

并且由(11.6.8)式可知$S(\omega)\geqslant 0$.

证 由定理 11.5.4 的推论可得

$$\begin{aligned}E[\tilde{S}_c(\omega)]&=E\left[\frac{1}{2c}\left|\int_{-c}^{c}\xi_t e^{-j\omega t}dt\right|^2\right]\\&=\frac{1}{2c}\int_{-c}^{c}\int_{-c}^{c}E(\xi_{t_1}\bar{\xi}_{t_2})e^{-j\omega(t_1-t_2)}dt_1dt_2\\&=\frac{1}{2c}\int_{-c}^{c}\int_{-c}^{c}B(t_1-t_2)e^{-j\omega(t_1-t_2)}dt_1dt_2\\&=\frac{1}{2c}\int_{-2c}^{2c}(2c-|\tau|)B(\tau)e^{-j\omega\tau}d\tau\\&=\int_{-2c}^{2c}B(\tau)e^{-j\omega\tau}d\tau-\frac{1}{2c}\int_{-2c}^{2c}|\tau|B(\tau)e^{-j\omega\tau}d\tau.\end{aligned}$$

其中，二重积分化为单积分时，用了变量替换：$t_1-t_2=\tau$，$t_2=t_2$.

---

① 随机积分$\int_{-c}^{c}\xi_t e^{-j\omega t}dt$的定义类似于定义 11.5.3 的均方积分，并且有类似的定理和性质.

专业课辅导+真题资料+微信公众号：432应统书院

若能证明上式右端第二项在假设条件下，当 $c\to\infty$ 时趋于 0，则定理 11.6.1 得证. 下面加以证明：

由假设，$\int_{-\infty}^{\infty}|B(\tau)|\mathrm{d}\tau<\infty$，故任给 $\varepsilon>0$，存在 $N_0>0$，使得当 $c>N_0$ 时，

$$\int_{2c}^{\infty}|B(\tau)|\mathrm{d}\tau\leqslant\int_{2N_0}^{\infty}|B(\tau)|\mathrm{d}\tau<\frac{\varepsilon}{3};$$

$$\int_{-\infty}^{-2c}|B(\tau)|\mathrm{d}\tau\leqslant\int_{-\infty}^{-2N_0}|B(\tau)|\mathrm{d}\tau<\frac{\varepsilon}{3}.$$

将区间 $[-2c,2c]$ 分成三个部分：

$$\Delta_1=[-2c,-2N_0);$$
$$\Delta_2=[-2N_0,2N_0);$$
$$\Delta_3=[2N_0,2c].$$

对于区间 Δ_1，有

$$\begin{aligned}\left|\frac{1}{2c}\int_{-2c}^{-2N_0}|\tau|B(\tau)\mathrm{e}^{-\mathrm{j}\omega\tau}\mathrm{d}\tau\right|&\leqslant\frac{2c}{2c}\int_{-2c}^{-2N_0}|B(\tau)|\mathrm{d}\tau\\&\leqslant\int_{-\infty}^{-2N_0}|B(\tau)|\mathrm{d}\tau\\&<\frac{\varepsilon}{3}.\end{aligned}\tag{11.6.9}$$

同理，对区间 Δ_3，有

$$\left|\frac{1}{2c}\int_{2N_0}^{2c}|\tau|B(\tau)\mathrm{e}^{-\mathrm{j}\omega\tau}\mathrm{d}\tau\right|<\frac{\varepsilon}{3}.\tag{11.6.10}$$

对于区间 Δ_2，有

$$\left|\frac{1}{2c}\int_{-2N_0}^{2N_0}|\tau|B(\tau)\mathrm{e}^{-\mathrm{j}\omega\tau}\mathrm{d}\tau\right|\leqslant\frac{2N_0}{2c}\int_{-2N_0}^{2N_0}|B(\tau)|\mathrm{d}\tau.$$

对于固定的 N_0，由假设，存在 N_1，当 $c>N_1$ 时，上式小于 $\frac{\varepsilon}{3}$，由此及(11.6.9)、(11.6.10)两式，即得所要的结果.

四、遍历性定理

在实际工作中，确定随机过程的特征是很重要的. 例如飞机在

专业课辅导+真题资料+微信公众号：432应统书院

高空飞行受湍流影响产生机翼振动，要确定其振动范围，以决定制造飞机时选取什么材料，这就要考虑机翼在时间变化过程中振动振幅大小（它是一随机过程）的方差. 又如电路中电子的不规则运动——热骚动引起的电位的脉动（通常称为热噪声，它也是随机过程），需要考虑其脉动范围、噪声功率等，也要考虑其方差、相关函数等. 然而，要求出上述这些特征，需要知道随机过程的一维、二维分布函数，而这些分布函数一般在实际问题中是没有给定的. 为了在实际工作中获得这些特征，用类似处理随机变量时的办法，即统计方法：通过大量的观察或试验，由所获得的数据求出它们的估计值. 以地震勘探为例，为了求得地震波振幅$\{\xi_t\}$在t_0时的数学期望$E(\xi_{t_0})$，我们可在相同条件下放n次炮，获得n个记录，如图11.1.1. 把每一记录（一个现实）在t_0时的值记录下来，得到n个数$x_1,x_2,\cdots,x_n$，把这n个数求平均得$\frac{1}{n}\sum_{i=1}^{n}x_i$，将它作为$E(\xi_{t_0})$的估计值. 诚然，试验次数越多，这个估计值就越接近$E(\xi_{t_0})$. 但从生产上看，则希望试验次数越少越好，以节省时间，降低生产费用，等等，特别是对于破坏性试验，更是如此. 能否根据一次试验所获得的一个现实来确定随机过程特征呢？这一节将考虑这个问题.

引理 11.6.1 设$\{\xi_t\}$为一均方连续弱平稳过程，其数学期望$E(\xi_t)$为常数m，则

$$\xi=\frac{1}{2c}\int_{-c}^{c}\xi_t\mathrm{d}t$$

的数学期望和方差分别为：

$$E(\xi)=E(\xi_t)=m,$$

$$\sigma_\xi^2=\frac{1}{4c^2}\int_{-c}^{c}\int_{-c}^{c}R(t_1-t_2)\,\mathrm{d}t_1\mathrm{d}t_2, \tag{11.6.11}$$

其中$R(t_1-t_2)=B(t_1-t_2)-m^2$，$B(\tau)$为$\{\xi_t\}$的相关函数.

证 因为

$$E(\xi)=\frac{1}{2c}E\left(\int_{-c}^{c}\xi_t\mathrm{d}t\right)=\frac{1}{2c}\int_{-c}^{c}E(\xi_t)\,\mathrm{d}t=m;$$

专业课辅导+真题资料+微信公众号：432应统书院

$$E(\xi^2)=\frac{1}{4c^2}E\left(\int_{-c}^{c}\xi_{t_1}\mathrm{d}t_1\cdot\int_{-c}^{c}\xi_{t_2}\mathrm{d}t_2\right)$$

$$=\frac{1}{4c^2}\int_{-c}^{c}\int_{-c}^{c}B(t_1-t_2)\,\mathrm{d}t_1\mathrm{d}t_2.$$

故

$$\sigma_\xi^2=E(\xi^2)-[E(\xi)]^2$$

$$=\frac{1}{4c^2}\int_{-c}^{c}\int_{-c}^{c}B(t_1-t_2)\,\mathrm{d}t_1\mathrm{d}t_2-\frac{1}{4c^2}\left[\int_{-c}^{c}E(\xi_t)\,\mathrm{d}t\right]^2$$

$$=\frac{1}{4c^2}\int_{-c}^{c}\int_{-c}^{c}[B(t_1-t_2)-E(\xi_{t_1})E(\xi_{t_2})]\,\mathrm{d}t_1\mathrm{d}t_2$$

$$=\frac{1}{4c^2}\int_{-c}^{c}\int_{-c}^{c}R(t_1-t_2)\,\mathrm{d}t_1\mathrm{d}t_2.$$

对于(11.6.11)式可以进一步化简，为此令 $t_1-t_2=\tau$，因而有

$$\sigma_\xi^2=\frac{1}{4c^2}\int_{-c}^{c}\int_{-c}^{c}R(t_1-t_2)\,\mathrm{d}t_1\mathrm{d}t_2$$

$$=\frac{1}{4c^2}\int_{-2c}^{2c}(2c-|\tau|)R(\tau)\,\mathrm{d}\tau$$

$$=\frac{1}{2c}\int_{-2c}^{2c}\left(1-\frac{|\tau|}{2c}\right)[B(\tau)-m^2]\,\mathrm{d}\tau. \tag{11.6.12}$$

定理 11.6.2 设$\{\xi_t\}$满足引理 11.6.1 的条件，则

$\underset{c\to\infty}{\mathrm{l.i.m}}\dfrac{1}{2c}\displaystyle\int_{-c}^{c}\xi_t\mathrm{d}t=E(\xi_t)$的充分必要条件为

$$\lim_{c\to\infty}\frac{1}{2c}\int_{-2c}^{2c}\left(1-\frac{|\tau|}{2c}\right)[B(\tau)-m^2]\,\mathrm{d}\tau=0.$$

证 由(11.6.12)式及假设立即得到所要证的结果. 事实上，这是因为

$$E\left|\frac{1}{2c}\int_{-c}^{c}\xi_t\mathrm{d}t-E(\xi_t)\right|^2=\sigma_\xi^2$$

$$=\frac{1}{2c}\int_{-2c}^{2c}\left(1-\frac{|\tau|}{2c}\right)[B(\tau)-m^2]\,\mathrm{d}\tau$$

专业课辅导+真题资料+微信公众号：432应统书院

之故.

满足定理 11.6.2 中 $\underset{c\to\infty}{\mathrm{l.i.m}}\frac{1}{2c}\int_{-c}^{c}\xi_t\mathrm{d}t=E(\xi_t)$ 的随机过程，称为对数学期望具有遍历性，或说是埃尔哥德的.

例 11.6.4 例 11.6.2 的弱平稳过程 $\{\xi_t\}$ 对数学期望是遍历的，事实上，因为其相关函数 $B(\tau)=\cos 2\pi\tau$ 满足定理 11.6.2 的条件，即

$$\lim_{c\to\infty}\frac{1}{2c}\int_{-2c}^{2c}\left(1-\frac{|\tau|}{2c}\right)\cos 2\pi\tau\mathrm{d}\tau$$

$$=\lim_{c\to\infty}\frac{1}{4c^2}\left[\frac{2}{4\pi^2}(1-\cos 4\pi c)\right]=0.$$

§11.7 时间序列与离散鞅

前面所述的随机过程中，有的像 i.i.d(独立同分布)序列. 它描述随机现象在时间变化过程中彼此相互独立；有的像马氏链、强平稳过程，它表现为其概率分布具有某种特性，又如弱平稳过程，它的数字特征对时间平移不改变. 这一节介绍另一些具有某种特性的随机过程.

一、时间序列分析

时间序列分析是近来应用很广泛的一门概率统计分支，它在诸如交通运输、气象、水文、生物学、社会学、经济学和在控制论、通信技术等方面有很好的应用.

我们在观察许多自然现象，如某地气温的变化、某人的体温、某上市公司股票价格等，都可记录到随着时间变化的数据. 我们希望从过去获得的数据来预测未来，或者说，我们希望建立一定的数学模型，由现在和过去预测未来.

在本章一开始，我们指出，一般地，上述现象是一随机过程，其参数空间可以是离散的，也可以是连续型的. 不过它总可以用采样方法使其离散化[17]. 另外，参数也不一定是时间，这一点我们在本

专业课辅导+真题资料+微信公众号：432应统书院

章开始时也指出过. 但大多数表现为时间,故我们在这一节,描述这些现象的都称为随机时间序列,简称为时间序列. 如不产生混乱,我们都把时间序列和时间序列的轨道(现实)或部分轨道均称为时间序列.

在时间序列的理论研究和实际应用中,比较常用的是线性时间序列或一些非线性时间序列模型.

(1) 自回归模型(AR(p))

定义 11.7.1 二阶矩时间序列$\{\xi_t: t\in \mathbf{N}\}$(简记为$\{\xi_t\}$)称为自回归序列或称$\{\xi_t\}$满足自回归模型,如果对任意$t\in\mathbf{N}$,$\{\xi_t\}$满足

$$1^\circ\ \xi_t = \varphi_1\xi_{t-1}+\cdots+\varphi_p\xi_{t-p}+\varepsilon_t = \sum_{i=1}^{p}\varphi_i\xi_{t-i}+\varepsilon_t, \quad (11.7.1)$$

其中$p>0$,$\varphi_p\neq 0$,$\{\varepsilon_t\}$为白噪声,即$E(\varepsilon_t\cdot\varepsilon_s)=0(t\neq s)$,$E\varepsilon_t^2=\sigma_\varepsilon^2>0$,$\varphi_i(i=1,2,\cdots,p)$为常数;

$$2^\circ\ 当\ t<s\ 时,E(\xi_t\varepsilon_s)=0, \quad (11.7.2)$$

称p为自回归模型的阶数. 自回归模型用AR(p)表示.

如果我们把$\{\xi_t\}$看成一个信号,$\{\varepsilon_t\}$是一个干扰,则(11.7.2)式表示信号在某一时刻之前,不受这一时刻以后出现的干扰所影响. 这是很直观的,也是一般应满足的.

例 11.7.1 设一商店某商品月销量ξ_t,满足

$$\xi_t=0.75\,\xi_{t-1}+0.26\,\xi_{t-2}+\varepsilon_t,$$

即第t月商品销售量,线性地依赖它前两个月的销售量,并且靠前一个月比再前一个月的影响大.

以$\xi_{t-k}(k>0)$乘(11.7.1)式两边并求数学期望,注意到(11.7.2)式,可得$\{\xi_t\}$的相关函数(设$\{\xi_t\}$为弱平稳序列)为

$$\begin{aligned}B(k)&=E(\xi_t\cdot\xi_{t-k})\\&=\sum_{i=1}^{p}\varphi_iE(\xi_{t-i}\xi_{t-k})+E(\xi_{t-k}\varepsilon_t)\\&=\sum_{i=1}^{p}\varphi_iB(k-i)\quad(k>0).\end{aligned}$$

专业课辅导+真题资料+微信公众号：432应统书院

将上式改写为

$$\sum_{i=0}^{p} \varphi_i B(k-i)=0 \quad (\varphi_0=-1, k>0). \tag{11.7.3}$$

(11.7.3)式称为 Yule-Walker 方程.

(2) 滑动和模型(MA(q))

定义 11.7.2 设二阶矩时间序列$\{\xi_t\}$称为滑动和序列,或称为$\{\xi_t\}$满足滑动和模型,如果对任一 $t\in\mathbf{N}$,有

1° $\xi_t=\varepsilon_t-\theta_1\varepsilon_{t-1}-\cdots-\theta_q\varepsilon_{t-q}$

$$=-\sum_{i=0}^{q}\theta_i\varepsilon_{t-i}$$

($\theta_0=-1, \theta_i(i=1,2,\cdots,q)$为常数,$\theta_q\neq 0$); (11.7.4)

且

2° 当 $t<s$ 时,$E(\xi_t\cdot\varepsilon_s)=0$. (11.7.5)

用 MA(q)表示上述滑动和模型,q 称为 MA(q)的阶数.

由(11.7.4)式可得

$$E(\xi_{t+k}\xi_t)=\sum_{i,j=0}^{q}\theta_i\theta_j E(\varepsilon_{t+k-i}\varepsilon_{t-j})$$

$$=\begin{cases}\sigma_\varepsilon^2\sum_{i=0}^{q}\theta_i^2, & 当\ k=0;\\ \sigma_\varepsilon^2(-\theta_k+\theta_1\theta_{k+1}+\cdots+\theta_{q+k}\theta_q), & 当\ 1\leqslant k\leqslant q;\\ 0, & 当\ k>q.\end{cases} \tag{11.7.6}$$

从(11.7.6)式看到:$\{\xi_t\}$是弱平稳随机过程,因为它的相关函数与 t 无关;其次它的相关函数"q 步截尾",即 $k>q$ 时,相关函数为零. 注意到相关函数的对称性,在(11.7.6)式中的 k 换为$|k|$时应成立.

(3) 自回归滑动和(混合)模型 ARMA(p,q)

定义 11.7.3 设$\{\xi_t\}$为二阶矩随机序列,若满足

1° 对任意 $t\in\mathbf{N}$,

专业课辅导+真题资料+微信公众号：432应统书院

$$\begin{aligned}&\xi_t-\varphi_1\xi_{t-1}-\cdots-\varphi_p\xi_{t-p}\\&=\varepsilon_t-\theta_1\varepsilon_{t-1}-\cdots-\theta_q\varepsilon_{t-q},\end{aligned}\tag{11.7.7}$$

其中 $\varphi_1,\varphi_2,\cdots,\varphi_p,\theta_1,\theta_2,\cdots\theta_q$ 均为常数；$\varphi_p\neq0,\theta_q\neq0$；

2° 当 $t<s$ 时，$E(\xi_t\varepsilon_s)=0$，　　(11.7.8)

则称 $\{\xi_t\}$ 为自回归滑动和序列，或称 $\{\xi_t\}$ 满足自回归滑动和模型. 记为 ARMA(p,q)，(p,q) 称为模型的阶数.

读者可仿前述方法求得，若 $\{\xi_t\}$ 是平稳序列，且满足 ARMA(p,q) 模型，则其相关函数为

$$B(k)=\begin{cases}\sum\limits_{i=1}^{p}\varphi_iB(k-i) & (k>q);\\ \sum\limits_{i=1}^{p}\varphi_iB(k-i)+E(\xi_{t-k}\varepsilon_t)-\sum\limits_{i=1}^{q}\theta_iE(\xi_{t-k}\varepsilon_{t-j}).\end{cases}\tag{11.7.9}$$

ARMA(p,q) 模型包含了 AR(p) 和 MA(q)，其相关函数较为复杂，我们不打算深入讨论，有兴趣的读者可参阅[17].

(4) 门限自回归模型(TAR)

定义 11.7.4　设 $\{\xi_t\}$ 为一时间序列，若对任一 $t\in\mathbf{N}$.

$$\xi_t=\varphi_0^{(j)}+\sum_{i=1}^{p_j}\varphi_i^{(j)}\xi_{t-i}+\varepsilon_t^{(j)},\text{当 }\xi_{t-d}\in R_j,j=1,2,\cdots,l.\tag{11.7.10}$$

其中 $R_j=(r_{j-1},r_j),j=1,2,\cdots,l-1;r_{j-1}<r_j,r_0=-\infty,R_l=(r_{l-1},\infty)$. (11.7.10)式共有 l 个自回归模型，取哪一个模型，由 ξ_{t-d} 属于哪一个 R_j 决定. 式中 p_j 称为阶数，$\varphi_i^{(j)}$ 称为门限自回归系数，$r_j,j=1,2,\cdots,l-1$ 称为门限值，d 称为门限自回归的延迟参数. $\{\varepsilon_t^{(j)}\}$ 是白噪声序列. 称上述模型为自激门限自回归(SETAR$(d,l,p_1,\cdots,p_l)$)模型. 此模型是一非线性模型，它利用门限将状态空间取值范围分为 l 个区间，在每一区间上用一个线性的 AR(p_j) 模型表示.

二、中国消费与积累的非线性模型

用数量经济学方法研究我国经济的消费积累问题，得到以下

专业课辅导+真题资料+微信公众号：432应统书院

的模型：

$$\begin{cases} Y_t = C_t + I_t; \\ C_t = \gamma_0 + \gamma_1 Y_{t-1} + \gamma_2 C_{t-1}; \\ I_t = \beta_0 + \beta_1 (Y_t - Y_{t-1}) + \beta_2 I_{t-1}. \end{cases}$$

其中 Y_t, C_t 及 I_t 分别表示第 t 年的国民收入、消费与积累.

将上述模型稍作变动后可以看出，它实际上等价于一个闭环线性自回归模型. 其特点是国民收入、消费与积累三者之间呈现一定的线性关系. 用一阶闭环门限自回归模型来拟合我国的消费积累资料时，模型为

$$\begin{cases} C_t = \varphi_0^{(j)} + \varphi_1^{(j)} C_{t-1} + \varphi_2^{(j)} I_{t-1} + \varepsilon_t^{(i)}, \\ I_t = \psi_0^{(j)} + \psi_1^{(j)} C_{t-1} + \psi_2^{(j)} I_{t-1} + \eta_t^{(i)}, \end{cases} \text{当} \frac{I_{t-1}}{I_{t-1} + C_{t-1}} \in R_j \text{时},$$

$j = 1, 2, R_1 = (0, r), R_2 = (r, 1)$. 其中 r 表示积累率的“门限”. 当 r 确定后，模型系数 $\varphi_k^{(j)}, \psi_k^{(j)}$ $(k = 0, 1, 2)$ 可用最小二乘法进行估计，有[17]

$$\hat{\boldsymbol{\Phi}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y},$$

其中

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 & 0 \\ 0 & \boldsymbol{X}_1 \end{pmatrix},$$

$$\boldsymbol{X}_1 = \begin{pmatrix} \boldsymbol{X}_{11} & \boldsymbol{X}_{12} \\ \boldsymbol{X}_{21} & \boldsymbol{X}_{22} \\ \vdots & \vdots \\ \boldsymbol{X}_{N1} & \boldsymbol{X}_{N2} \end{pmatrix},$$

$$\boldsymbol{X}_{kj} = (1, C_{k-1}, I_{k-1})', \quad \left\{ \frac{I_{k-1}}{C_{k-1} + I_{k-1}} \in R_j \right\},$$

$$k = 1, 2, \cdots, N, \quad j = 1, 2.$$

$$\boldsymbol{Y}' = (C_1, C_2, \cdots, C_N, I_1, I_2, \cdots, I_N),$$

$$\boldsymbol{\Phi}' = (\varphi_0^{(1)}, \varphi_1^{(1)}, \cdots, \varphi_2^{(2)}, \psi_0^{(1)}, \psi_1^{(1)}, \cdots, \psi_2^{(2)}).$$

其中 $C_i, I_i (i = 1, 2, \cdots, N)$ 为观察值.

专业课辅导+真题资料+微信公众号：432应统书院

利用《中国统计年鉴》(1983 年)提供的 1952 年至 1982 年国民收入的具体数据①,用最小二乘法估计 Φ,通过优选,得到 $r=0.35$. 参数估计结果如下(单位:亿元):

$$\begin{cases} C_t=-3.37+1.14C_{t-1}-0.13I_{t-1}, \\ I_t=1.72+0.11C_{t-1}+0.81I_{t-1}, \end{cases} \quad 当\frac{I_{t-1}}{C_{t-1}+I_{t-1}}<0.35.$$

$$\begin{cases} C_t=-35.75+0.75C_{t-1}+0.03I_{t-1}, \\ I_t=-86.25-1.01C_{t-1}+3.71I_{t-1}, \end{cases} \quad 当\frac{I_{t-1}}{C_{t-1}+I_{t-1}}\geqslant 0.35.$$

拟合结果表明,积累率高于或低于 35%,将使次年的国民经济有不同的发展趋势. 从拟合结果来看(见表 11.7.1),一些特殊年份(如 1959,1961,1962 年等)误差较大,这反映了这些年份经济发展处于不正常状态.

表 11.7.1

年份	积累率	消费	模型值	残差	积累	模型值	残差
1953	0.231	52.940	49.425	3.515	15.910	17.266	-1.356
1954	0.254	54.130	55.035	-0.905	18.518	20.176	-1.658
1955	0.229	59.649	56.061	3.588	17.741	22.416	-4.675
1956	0.244	65.601	62.461	3.140	21.212	22.365	-1.153
1957	0.249	69.672	68.813	0.859	23.124	25.807	-2.683
1958	0.339	72.590	73.216	-0.626	37.278	27.783	9.495
1959	0.438	69.747	74.740	-4.993	54.355	39.567	14.788
1960	0.396	73.379	72.255	1.124	48.182	45.167	3.015
1961	0.192	67.723	68.638	-0.915	16.144	18.601	-2.457
1962	0.104	70.842	71.883	-1.041	8.261	21.918	-13.657

① 原数据是用当年价格计算的,必须除以综合国民收入指数,使数据具有可比性. 综合国民收入指数(Implicit national imcome deflator)是指当年价格计算的国民收入与用 1952 年可比价格计算的国民收入之比.

专业课辅导+真题资料+微信公众号：432应统书院

续表

年份	积累率	消费	模型值	残差	积累	模型值	残差
1963	0.174	73.739	76.451	-2.712	15.618	15.852	-0.234
1964	0.222	78.532	78.819	-0.287	22.426	22.123	0.303
1965	0.271	82.360	83.421	-1.061	30.613	28.146	2.467
1966	0.306	91.364	86.746	4.618	40.320	35.187	5.133
1967	0.212	95.410	95.785	-0.375	25.805	44.005	-18.200
1968	0.211	92.677	102.260	-9.583	24.858	32.658	-7.800
1969	0.232	102.770	99.260	3.510	31.092	31.604	-0.512
1970	0.329	113.376	109.986	3.390	55.697	37.719	17.978
1971	0.340	118.383	118.952	-0.569	61.159	58.784	2.375
1972	0.315	125.631	123.970	1.661	57.983	63.740	-5.757
1973	0.329	134.917	132.651	2.266	66.164	61.925	4.239
1974	0.323	138.109	142.207	-4.098	66.025	69.534	-3.509
1975	0.338	146.744	145.869	0.875	75.137	69.756	5.381
1976	0.308	152.284	154.564	-2.280	67.964	78.052	-10.088
1977	0.323	156.532	161.805	-5.273	74.804	72.817	1.987
1978	0.365	167.433	165.781	1.652	96.398	78.811	17.587
1979	0.345	187.136	188.544	-1.408	98.981	102.763	-3.782
1980	0.316	207.149	205.948	1.201	95.727	92.504	3.223
1981	0.284	224.166	220.897	3.269	89.150	101.090	-11.940
1982	0.289	242.626	241.165	1.461	99.026	97.543	1.483

三、离散鞅

为了叙述鞅论，必须用到条件数学期望的概念. 关于它，我们曾在上册的§3.5引进了随机变量ξ对随机变量η的条件数学期

专业课辅导+真题资料+微信公众号：432应统书院

望 $E(\xi|\eta)$ 及由定理 3.5.2 所列的诸性质. 这里需要把上述概念及性质推广到随机变量 ξ 对另外 n 个随机变量 $\eta_1,\eta_2,\cdots,\eta_n$ 的条件数学期望 $E(\xi|\eta_1,\eta_2,\cdots,\eta_n)$ 的情形. 对于它,读者完全可仿 §3.5 的办法引入,并且证明具有如下的类似定理 3.5.2 所列的诸性质,在这里不打算详述,只把有关性质列在下面.

条件数学期望 $E(\xi|\eta_1,\eta_2,\cdots,\eta_n)$ 有如下性质:

首先,$E(\xi|\eta_1,\eta_2,\cdots,\eta_n)$ 是依赖于 $\eta_1,\eta_2,\cdots,\eta_n$ 的随机变量,并且

(1) 当 ξ 与 $\eta_1,\eta_2,\cdots,\eta_n$ 独立时,有

$$E(\xi|\eta_1,\eta_2,\cdots,\eta_n)=E\xi;$$

(2) $E[E(\xi|\eta_1,\eta_2,\cdots,\eta_n)]=E\xi$;

(3) $|E(\xi|\eta_1,\eta_2,\cdots,\eta_n)|\leqslant E(|\xi||\eta_1,\eta_2,\cdots,\eta_n)$;

(4) $E[\xi g(\eta_1,\eta_2,\cdots,\eta_n)|\eta_1,\eta_2,\cdots,\eta_n]$
$=g(\eta_1,\eta_2,\cdots,\eta_n)E(\xi|\eta_1,\eta_2,\cdots,\eta_n)$;

(5) $E[g(\eta_1,\eta_2,\cdots,\eta_n)E(\xi|\eta_1,\eta_2,\cdots,\eta_n)]$
$=E[\xi g(\eta_1,\eta_2,\cdots,\eta_n)]$;

(6) $E[g(\eta_1,\eta_2,\cdots,\eta_n)|\eta_1,\eta_2,\cdots,\eta_n]=g(\eta_1,\eta_2,\cdots,\eta_n)$;

(7) $E[E(\xi|\eta_1,\eta_2,\cdots,\eta_{n+1})|\eta_1,\eta_2,\cdots,\eta_n]=E(\xi|\eta_1,\eta_2,\cdots,\eta_n)=E[E(\xi|\eta_1,\eta_2,\cdots,\eta_n)|\eta_1,\eta_2,\cdots,\eta_{n+1}]$;

(8) 对任意常数 a,b,有

$E[(a\xi_1+b\xi_2)|\eta_1,\eta_2,\cdots,\eta_n]=aE(\xi_1|\eta_1,\eta_2,\cdots,\eta_n)+bE(\xi_2|\eta_1,\eta_2,\cdots,\eta_n)$. 特别

$$E(a|\eta_1,\eta_2,\cdots,\eta_n)=a.$$

其中 $g(\cdot)$ 为使得 $E|g|<\infty$ 的随机变量.

下面我们只考虑离散参数鞅(离散鞅)的情形,关于现代鞅论的内容,读者可参阅[23].

定义 11.7.5 若实的随机序列 $\{\xi_n:n\in\mathbf{N}\}$ ($\mathbf{N}$ 表示非负整数集)满足如下性质:对任意 $n\geqslant 0$,有

(1) $E|\xi_n|<\infty$;

专业课辅导+真题资料+微信公众号：432应统书院

(2) $E(\xi_{n+1}\mid\xi_0,\xi_1,\cdots,\xi_n)=\xi_n$,

则称$\{\xi_n\}$为离散鞅序列,如不产生混乱,简称为鞅(Martingales)("Martingale"一词是重复进行赌博直至胜利的一种博弈策略的法兰西语的字头缩写.鞅论首先由 Doob 和 Loévy 在 20 世纪的 30 年代末至 50 年代初创立的).

从定义 11.7.5 的(2),我们看到:如果$\{\xi_n\}$为鞅,则它有某种"无后效"性,即当我们已知时刻 n 以及它以前的值 $\xi_0,\xi_1,\cdots,\xi_n$,那么 $n+1$ 时刻的值 ξ_{n+1} 对 $\xi_0,\xi_1,\cdots,\xi_n$ 的条件数学期望与时刻 n 以前的 $\xi_0,\xi_1,\cdots,\xi_{n-1}$ 无关,并且等于 ξ_n.

鞅的直观背景可作如下解释:

设想"赌徒"(可以是工厂或农村的生产者、商店中的经理、也可以是博弈者等)进行"赌博"(可以是进行某种工业或农业生产、某种商业活动或博弈等),他在第 n 年的"赌本"(可以是生产资金或商业资金等)为 ξ_n,$E(\xi_{n+1}\mid\xi_0,\xi_1,\cdots,\xi_n)$ 表示在已知前 $n+1$ 年的"赌本"$\xi_0,\xi_1,\cdots,\xi_n$ 的条件下,第 $n+1$ 年的平均"赌本".而鞅的性质 $E(\xi_{n+1}\mid\xi_0,\xi_1,\cdots,\xi_n)=\xi_n$ 则表示这种"赌博"使第 $n+1$ 年的平均"赌本"仍为第 n 年的"赌本",这种"赌博"称为"公平"的"赌博".

由定义 11.7.5 并注意到条件数学期望性质(2),可得

性质 11.7.1 若$\{\xi_n\}$为鞅,则对任意 $n\geqslant 0$,有

$$E\xi_n=E\xi_0,$$

即 ξ_n 的数学期望是一常数 $E\xi_0$.

证 $E\xi_{n+1}=E[E(\xi_{n+1}\mid\xi_0,\xi_1,\cdots,\xi_n)]$

$=E(\xi_n)$.

即得到对任意 $n\geqslant 0$,$E\xi_n=E\xi_0$.

性质 11.7.1 表明,若$\{\xi_n\}$是鞅,其"赌本"的平均数在任何时刻都一样.

性质 11.7.2 下列两个命题等价:

1) $\{\xi_n\}$是鞅,

专业课辅导+真题资料+微信公众号：432应统书院

2）对任意非负整数 $m,n(m\geq n)$，有

$$E(\xi_m|\xi_0,\xi_1,\cdots,\xi_n)=\xi_n.$$

证 显然 1）是 2）的特殊情况. 只需证 1）⇒2）.

由 1）知，当 $m=n+1$ 时 2）成立. 现设当 $m=n+k$ 时 2）成立 $(k>1)$. 则有

$$\begin{aligned}&E(\xi_{n+k+1}|\xi_0,\xi_1,\cdots,\xi_n)\\&=E[E(\xi_{n+k+1}|\xi_0,\xi_1,\cdots,\xi_{n+k})|\xi_0,\xi_1,\cdots,\xi_n]\\&=E[\xi_{n+k}|\xi_0,\xi_1,\cdots,\xi_n]\\&=\xi_n.\end{aligned}$$

这表明 $m=n+k+1$ 时 2）成立. 证毕

定义 11.7.6 设 $\{\xi_n:n\in\mathbf{N}\}$ 是一随机序列，若

1）$E|\xi_n|<\infty$；

2）$E(\xi_{n+1}|\xi_0,\xi_1,\cdots,\xi_n)\leq\xi_n$，

则称 $\{\xi_n\}$ 为上鞅.

定义 11.7.7 设 $\{\xi_n:n\in\mathbf{N}\}$ 为一随机序列，若

1）$E|\xi_n|<\infty$；

2）$E(\xi_{n+1}|\xi_0,\xi_1,\cdots,\xi_n)\geq\xi_n$，

则称 $\{\xi_n\}$ 为下鞅.

上述两个定义直观解释为：上鞅表示第 $n+1$ 年的平均"赌本"不多于第 n 年的"赌本"，即具有上鞅性质的"赌博"是没有盈利的，而下鞅恰好相反，它是有盈利的.

显然，常数序列 $\{C_n\}$ $(C_n=C)$ 是上鞅也是下鞅.

性质 11.7.3

1）$\{\xi_n\}$ 为上鞅（或下鞅）的充分必要条件是 $\{-\xi_n\}$ 为下鞅（或上鞅）.

2）$\{\xi_n\}$ 为鞅的充分必要条件是 $\{\xi_n\}$ 既为上鞅也为下鞅.

证 2）是显然的. 下面证明 1）.

由假设 $\{\xi_n\}$ 为上鞅（或下鞅），则有

$$E(\xi_{n+1}|\xi_0,\xi_1,\cdots,\xi_n)\leq(\text{或}\geq)\xi_n.$$

专业课辅导+真题资料+微信公众号：432应统书院

因此

$$E(-\xi_{n+1} \mid \xi_0, \xi_1, \cdots, \xi_n) = -E(\xi_{n+1} \mid \xi_0, \xi_1, \cdots, \xi_n)$$
$$\geqslant (\text{或} \leqslant) -\xi_n,$$

即$\{-\xi_n\}$为下鞅(或上鞅). 反之亦然.

性质 11.7.4 $\{\xi_n\}$为上(下)鞅的充分必要条件为,对任意$m, n(m>n)$有

$$E(\xi_m \mid \xi_0, \xi_1, \cdots, \xi_n) \leqslant (\text{或} \geqslant) \xi_n.$$

证 仿性质 11.7.2 证明.

性质 11.7.5 设$\{\xi_n\}$为上鞅(或下鞅),则对$0 \leqslant k \leqslant n$,有

$$E\xi_n \leqslant (\text{或} \geqslant) E\xi_k \leqslant (\text{或} \geqslant) E\xi_0.$$

证 仿性质 11.7.1 的证明.

上述性质表示,具有上鞅性质的"赌博",其平均"赌本"是不会增加的,而具有下鞅性质的"赌博"恰为相反,它是不会减少的.

专业课辅导+真题资料+微信公众号：432应统书院

第十二章　概率统计在计算方法中的一些应用

§12.1　蒙特卡罗方法与均匀分布随机数

蒙特卡罗方法又称统计试验方法，它是一种采用统计抽样理论近似地求解数学问题或物理问题的方法. 它既可用来研究概率问题，也可用来求解确定性的数学问题. 本章的目的只限于介绍蒙特卡罗方法的基本思想和原理，以及作为这一方法的基础——各种随机变量的模拟.

利用蒙特卡罗方法去解决数学分析问题时，基本的想法是首先建立与描述该问题有相似性的概率模型. 利用这种相似性把这概率模型的某些特征（如随机事件的概率或随机变量的平均值等）与数学分析问题的解答（如积分值、微分方程的解等）联系起来，然后对模型进行随机模拟或统计抽样，再利用所得结果求出这些特征的统计估计值作为原来的分析问题的近似解（如果需要的话还要对解的精度作出估计）. 下面我们将要看到，蒙特卡罗方法的理论基础是概率论中很一般的定理——大数定律，因此这方法的应用范围从原则上说来几乎没有什么限制. 但是这方法的统计特征又决定它和任何一种统计试验一样，如欲获得充分可靠的结果，则必须花费大量的劳动和时间，这就使它在实用上受到一定的限制. 幸好电子数字计算机的普遍应用，大大地减少了这种限制，因此从20世纪50年代以来，蒙特卡罗方法得到了广泛的发展和应用. 特别是随着科学技术的发展，在生产斗争和科学实验过程中，出现了许多复杂难解的问题. 例如核物理中描述质点运动的迁移运动方程；大型系统的可靠性分析；地震波的模拟试验；高维数学问题求

专业课辅导+真题资料+微信公众号：432应统书院

解;多元统计分析;医学、技术中的诊断、识别;大型生产过程中的运筹规划等,用传统的物理试验或数学方法进行处理十分困难,这时进行蒙特卡罗模拟常常能够提供一种有效而可行的解决办法.

把随机模拟方法用于近似数值计算领域已有近百年的历史,早期采用这方法的一个著名例子是所谓蒲丰问题(见例 1.2.9):历史上有人作过一次试验,方法是取 $2a=45$ mm,$2l=36$ mm,投针 5 000 次,有 2 532 次针与平行线相交. 故

$$p=\frac{2\ 532}{5\ 000}=0.506\ 4,$$

$$\pi=\frac{2l}{ap}\approx\frac{36\times 2}{45\times 0.506\ 4}=3.160$$

误差约为 0.55%.

我们再来看一下利用蒙特卡罗方法计算积分的简单例子. 考虑积分

$$I=\int_0^1 f(x)\,\mathrm{d}x.$$

假设当 $x\in[0,1]$ 时 $0\leqslant f(x)\leqslant 1$,这时积分 I 等于由曲线 $y=f(x)$,Ox 轴,Oy 轴以及 $x=1$ 围成的区域 G 的面积(见图12.1.1). 为求此面积,我们设想在正方形 $\{0\leqslant x\leqslant 1,0\leqslant y\leqslant 1\}$ 内随机地投掷一个点,落点的两个坐标是相互独立且有在区间 $[0,1]$ 上的均匀分布(即每点都具有等可能性). 那么,这个点落在曲线 $y=f(x)$ 以下的区域 G 内的概率 p 显然等于这区域的面积. 以 (ξ,η) 表示正方形内任一点的坐标. 如果我们用某种方法得到均匀分布的独立变量 ξ 及 η 的 n 个取样值,对 (ξ,η) 的每一取样值 (ξ_i,η_i) $(i=1,2,\cdots,n)$,检查 η_i 与 $f(\xi_i)$ 是否满足不等式

$$\eta_i<f(\xi_i). \tag{12.1.1}$$

如果(12.1.1)式成立,说明点 (ξ_i,η_i) 落在区域 G 内[如图 12.1.1 中的点 (ξ_1,η_1)],否则落在区域 G 外[如图 12.1.1 中的点 (ξ_2,η_2)]. 设满足不等式(12.1.1)的点数为 m,则由大数定律知当 n

专业课辅导+真题资料+微信公众号：432应统书院

足够大时，频率 $\mu=\frac{m}{n}$ 近似于点落在区域 G 的概率 p，亦即 $\frac{m}{n}$ 是所求积分的近似值.

图 12.1.1

在上面的积分计算中，要求获得在(0,1)上均匀分布的独立随机变量 ξ 与 η 的一系列取样值(即随机变量的现实). 一般地说，利用蒙特卡罗方法作各种类型数值计算时，同样也必须找出模拟随机变量或随机过程的现实，为此就需要所谓随机数. 通常实现一次随机变量或随机过程模拟所需的随机数的数量往往是很大的，简单的问题要几千个，复杂的问题则可能要几十万个或更多. 所以，用蒙特卡罗方法计算问题时，绝大部分的计算工作都是有关随机数的运算.

在区间(0,1)中均匀分布的随机数(简称均匀分布随机数)是最基本和常用的随机数. 这样的随机数可看作是区间(0,1)中均匀分布的随机变量的现实(或者说是样本)，其产生方法有多种. 但是，随着计算机科学与技术的迅速发展和普及，现在人们普遍采用由专门的运算程序在计算机产生的这种随机数. 因为这样产生的数是按照确定性的算法计算出来的，所以它们不是真正的随机数. 然而，如果计算方法选得恰当，它们能够经得起数理统计中的独立性检验和均匀分布检验，因而可以认为它们(很近似地)是相

专业课辅导+真题资料+微信公众号：432应统书院

互独立和均匀分布的，人们把这样产生的数称为伪随机数或拟随机数. 为方便起见，今后我们将其称为随机数或均匀分布随机数.

§12.2 连续型随机变量的一般模拟方法

一、反函数法

本法利用如下的已知结果：若随机变量 η 具有连续分布函数 $F(x)$，而 ξ 是在 $(0,1)$ 中均匀分布的随机变量，它的分布函数是

$$G(x)=\begin{cases}0, & 若\ x\leqslant 0;\\ x, & 若\ x\in(0,1);\\ 1, & 若\ x\geqslant 1,\end{cases}$$

则

$$\eta=F^{-1}(\xi). \tag{12.2.1}$$

这里 F^{-1} 如下定义：

$$F^{-1}(u)=\inf\{x:F(x)>u\},0\leqslant u\leqslant 1. \tag{12.2.2}$$

事实上，我们有

$$P\{F^{-1}(\xi)<x\}=P\{\xi<F(x)\}=G(F(x))=F(x).$$

这方法原则上适用于任意随机变量的模拟. 若进一步设分布函数 F 有密度函数 f，则 η 应满足方程

$$\xi=F(\eta)=\int_{-\infty}^{\eta}f(y)\,\mathrm{d}y.$$

因此，若给定了均匀分布随机数 ξ，则具有给定分布函数 F 的随机数 η 可由求解方程

$$\xi=\int_{-\infty}^{\eta}f(y)\,\mathrm{d}y \tag{12.2.3}$$

得到.

例 12.2.1 对于任意 $a<b$，求在 (a,b) 中均匀分布的随机数.

解 已知在 (a,b) 中均匀分布的随机变量有密度函数

$$f(x)=\begin{cases}0, & 若\ x\leqslant a;\\ (b-a)^{-1}, & 若\ x\in(a,b);\\ 0, & 若\ x\geqslant b.\end{cases}$$

专业课辅导+真题资料+微信公众号：432应统书院

故由(12.2.3)式有

$$\xi=\int_a^\eta \frac{\mathrm{d}y}{b-a}=\frac{\eta-a}{b-a},$$

由此得在(a,b)中均匀分布随机数

$$\eta=(b-a)\xi+a. \tag{12.2.4}$$

例 12.2.2 求具有参数λ的指数分布的随机数η.

参数为λ的指数分布有密度函数

$$f(y)=\lambda \mathrm{e}^{-\lambda y},\quad y\geqslant 0.$$

故(12.2.3)式可写成

$$\xi=\int_0^\eta \lambda \mathrm{e}^{-\lambda y}\mathrm{d}y=1-\mathrm{e}^{-\lambda\eta},$$

由此得

$$\eta=-\frac{1}{\lambda}\ln(1-\xi).$$

因为$1-\xi$和ξ同样是$(0,1)$中均匀分布随机数,故上式可简化为

$$\eta=-\frac{1}{\lambda}\ln \xi. \tag{12.2.5}$$

从理论上讲,反函数法是一种普遍适用的方法,但当反函数难以求出,例如,在不能由积分式(12.2.3)得到x的显式表示时就不宜使用,因而需要改用不必求解(12.2.3)式的方法.下面介绍一种这样的方法.

二、舍选法

设随机变量ζ和η分别有密度函数f和g,如果存在正常数c,使有

$$f(y)\leqslant cg(y),\quad 对所有\ y, \tag{12.2.6}$$

而且我们又有产生具有分布密度g的随机数的算法,则可用如下舍选方法产生具有分布密度f的随机数,其步骤如下:

第一步 独立地产生一个均匀分布随机数ξ和具有分布密度g的随机数η.

第二步 若

专业课辅导+真题资料+微信公众号：432应统书院

$$\xi \leqslant \frac{f(\eta)}{cg(\eta)},$$

则令 $\zeta=\eta$；否则返回第一步.

下面证明这样得到的 ζ 确是具有分布密度 f 的随机数. 令 N 记用上述算法首次产生 ζ 所需的次数，则 N 服从几何分布，即有

$$\begin{aligned} P\{N=n\} &= P\left\{\xi_k > \frac{f(\eta_k)}{cg(\eta_k)}, k=1,2,\cdots,n-1, \xi_n \leqslant \frac{f(\eta_n)}{cg(\eta_n)}\right\} \\ &= pq^{n-1}, \end{aligned} \tag{12.2.7}$$

式中

$$\begin{aligned} p &= 1-q \\ &= P\left\{\xi \leqslant \frac{f(\eta)}{cg(\eta)}\right\} \\ &= \int_{-\infty}^{\infty} P\left\{\xi \leqslant \frac{f(\eta)}{cg(\eta)} \,\middle|\, \eta=y\right\} g(y)\,\mathrm{d}y \\ &= \int_{-\infty}^{\infty} P\left\{\xi \leqslant \frac{f(y)}{cg(y)}\right\} g(y)\,\mathrm{d}y \\ &= \int_{-\infty}^{\infty} \frac{f(y)}{cg(y)} g(y)\,\mathrm{d}y \\ &= \frac{1}{c}, \end{aligned}$$

这里 $\xi_1,\xi_2,\cdots,\xi_n$ 是相互独立的均匀分布随机变量，$\eta_1,\eta_2,\cdots,\eta_n$ 是相互独立的具有相同分布密度 g 的随机变量，而且 $\{\xi_k\}$ 与 $\{\eta_k\}$ 相互独立.

于是由(12.2.7)式知对任意 x 有

$$\begin{aligned} P\{\zeta<x\} &= P\{\eta_N<x\} \\ &= \sum_{n=1}^{\infty} P\{N=n, \eta_n<x\} \\ &= \sum_{n=1}^{\infty} P\{\eta_n<x \mid N=n\} P\{N=n\} \\ &= \sum_{n=1}^{\infty} pq^{n-1} P\{\eta_n<x \mid N=n\} \end{aligned}$$

专业课辅导+真题资料+微信公众号：432应统书院

$$
\begin{aligned}
&= \sum_{n=1}^{\infty} pq^{n-1} \frac{P\left\{\eta_n < x, \xi_n \leqslant \frac{f(\eta_n)}{cg(\eta_n)}\right\}}{P\left\{\xi_n \leqslant \frac{f(\eta_n)}{cg(\eta_n)}\right\}} \\
&= \sum_{n=1}^{\infty} q^{n-1} \int_{-\infty}^{x} P\left\{\xi_n \leqslant \frac{f(\eta_n)}{cg(\eta_n)} \middle| \eta_n = y\right\} g(y)\,\mathrm{d}y \\
&= \sum_{n=1}^{\infty} q^{n-1} \int_{-\infty}^{x} P\left\{\xi_n \leqslant \frac{f(y)}{cg(y)}\right\} g(y)\,\mathrm{d}y \\
&= \sum_{n=1}^{\infty} q^{n-1} \int_{-\infty}^{x} \frac{f(y)}{cg(y)} g(y)\,\mathrm{d}y \\
&= \sum_{n=1}^{\infty} pq^{n-1} \int_{-\infty}^{x} f(y)\,\mathrm{d}y \\
&= \int_{-\infty}^{x} f(y)\,\mathrm{d}y,
\end{aligned}
$$

其中 η_n 独立于 $\eta_1, \eta_2, \cdots, \eta_{n-1}$, $P\left\{\xi_n \leqslant \frac{f(\eta_n)}{cg(\eta_n)}\right\} = p$, $p = \frac{1}{c}$.

舍选法的实质就是借助均匀分布随机数 ξ 从一些具有分布密度 g 的随机数 η 中适当地选取一部分(所占比例为 $f(\eta)/cg(\eta)$)作为具有分布密度 f 的随机数 ζ. 这方法首先由冯·诺伊曼(Von Neumann)就 f 是某有限区间 $[a,b]$ 上的有界函数和 g 是 $[a,b]$ 上的均匀分布密度的情形提出来的,这时

$$
g(y) = \begin{cases} \frac{1}{b-a}, & y \in [a,b]; \\ 0, & \text{其他}. \end{cases}
$$

设 B 是 $f(y)$ 的任一个上界,于是有

$$
\frac{f(y)}{g(y)} \leqslant B(b-a), \quad y \in [a,b].
$$

显然,当我们取 $c = B(b-a)$ 时不等式(12.2.6)成立.

因为 N 有参数为 p 的几何分布,故 $EN = 1/p = c$. 这表明按平均来说,利用舍选法产生一个随机数 ζ 需要产生 c 个随机数 η. 因

专业课辅导+真题资料+微信公众号：432应统书院

此，在满足(12.2.6)式的前提下，常数 c 选得愈小愈好.

下面应用舍选法求标准正态分布 $N(0,1)$ 随机数的例子颇有启发性.

例 12.2.3 设 ζ 是 $N(0,1)$ 随机变量，其绝对值 $\theta=|\zeta|$ 有密度函数

$$f(z)=\sqrt{\frac{2}{\pi}}\mathrm{e}^{-\frac{z^2}{2}},\quad z\in[0,\infty). \tag{12.2.8}$$

我们取参数为 1 的指数分布密度

$$g(y)=\mathrm{e}^{-y},\quad y\in[0,\infty)$$

作为舍选法的参照密度. 具有密度函数 $g(y)$ 的随机数 η 的产生方法见例 12.2.2.

注意到

$$\frac{f(y)}{g(y)}=\sqrt{\frac{2\mathrm{e}}{\pi}}\exp\left\{-\frac{(y-1)^2}{2}\right\}\leqslant\sqrt{\frac{2\mathrm{e}}{\pi}},\quad y\in[0,\infty),$$

故可取 $c=\sqrt{\frac{2\mathrm{e}}{\pi}}\approx1.32$. 这时产生随机数 ζ 的步骤如下：

(1) 独立地产生随机数 η 和均匀分布随机数 ξ.

(2) 若

$$\xi\leqslant\exp\{-(\eta-1)^2/2\},$$

或等价地

$$-\ln\xi\leqslant\frac{(\eta-1)^2}{2}, \tag{12.2.9}$$

则令 $\theta=\eta$，并转到下一步(3)，否则返回(1).

(3) 产生均匀分布随机数 ξ_1，并令

$$\zeta=\begin{cases}\theta, & 若\ \xi_1\leqslant\frac{1}{2};\\ -\theta, & 若\ \xi_1>\frac{1}{2}.\end{cases}$$

此时 ζ 即为 $N(0,1)$ 随机数.

因为 $-\ln\xi$ 是参数为 1 的指数分布随机数，故根据(12.2.9)式

专业课辅导+真题资料+微信公众号：432应统书院

可将上面的算法改进为：

(1)′ 分别独立地产生两个参数为 1 的指数分布随机数 η_1 和 η_2.

(2)′ 若 $\eta_2 \geqslant (\eta_1-1)^2/2$，则转到(3)′，否则返回(1)′.

(3)′ 产生一个均匀分布随机数 ξ，并令

$$\zeta=\begin{cases}\eta_1, & 若\ \xi\leqslant\dfrac{1}{2};\\ -\eta_1, & 若\ \xi>\dfrac{1}{2}.\end{cases}$$

此即 $N(0,1)$ 随机数.

如果想要产生正态分布 $N(m,\sigma)$ 随机数 ζ_1，则在获得标准正态分布 $N(0,1)$ 随机数 ζ 后，令 $\zeta_1=\sigma\zeta+m$，即得正态分布 $N(m,\sigma)$ 随机数.

§12.3 连续型随机变量的特殊模拟方法

一、正态分布随机数

除了利用上一节介绍的一般方法外，这里再介绍另外两种方法.

(1) 基于中心极限定理的方法

设 $\xi_1,\xi_2,\cdots,\xi_n$ 是 n 个相互独立的(0,1)中的均匀分布的随机变量. 因为 $E\xi_i=1/2$ 和 $D\xi_i=1/12$，故由中心极限定理知随机变量

$$\eta=\frac{\sum\limits_{i=1}^{n}\xi_i-\dfrac{n}{2}}{\sqrt{\dfrac{n}{12}}} \tag{12.3.1}$$

渐近地有标准正态分布 $N(0,1)$. 为了保证一定的精确度，上式中的 n 应取得足够大(一般情况下取 $n\geqslant 10$ 即可). 又从(12.3.1)式看出，当 $n=12$ 时该式的分母等于 1，这时有最简单的形式.

(2) 坐标变换法

设 ξ_1 和 ξ_2 是两个相互独立的均匀分布随机数. 作变换

专业课辅导+真题资料+微信公众号：432应统书院

$$\begin{cases}\eta_1=(-2\ln\xi_1)^{\frac{1}{2}}\cos(2\pi\xi_2),\\ \eta_2=(-2\ln\xi_1)^{\frac{1}{2}}\sin(2\pi\xi_2),\end{cases}\tag{12.3.2}$$

则 η_1 和 η_2 是两个相互独立的标准正态分布 $N(0,1)$ 随机数. 下面证明这一论断.

易见 (ξ_1,ξ_2) 的联合密度函数是

$$f(x_1,x_2)=\begin{cases}1, & x_1\in(0,1),x_2\in(0,1);\\ 0, & \text{其他}.\end{cases}$$

从(12.3.2)式可解出

$$\begin{cases}\xi_1=\exp\left\{-\dfrac{\eta_1^2+\eta_2^2}{2}\right\},\\ \xi_2=\dfrac{1}{2\pi}\arctan\dfrac{\eta_2}{\eta_1},\end{cases}$$

这变换的雅可比行列式是

$$J=\begin{vmatrix}-y_1\exp\left\{-\dfrac{y_1^2+y_2^2}{2}\right\} & -y_2\exp\left\{-\dfrac{y_1^2+y_2^2}{2}\right\}\\ \dfrac{1}{2\pi}\dfrac{-\dfrac{y_2}{y_1^2}}{1+\left(\dfrac{y_2}{y_1}\right)^2} & \dfrac{1}{2\pi}\dfrac{\dfrac{1}{y_1}}{1+\left(\dfrac{y_2}{y_1}\right)^2}\end{vmatrix}$$

$$=-\frac{1}{2\pi}\exp\left\{-\frac{y_1^2+y_2^2}{2}\right\}$$

$$=-\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{y_1^2}{2}}\right)\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{y_2^2}{2}}\right),$$

故 (η_1,η_2) 的联合密度函数是

$$\begin{aligned}g(y_1,y_2)&=|J|f(x_1(y_1,y_2),x_2(y_1,y_2))\\&=|J|\\&=\frac{1}{\sqrt{2\pi}}e^{-\frac{y_1^2}{2}}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{y_2^2}{2}}.\end{aligned}$$

这表明 η_1 和 η_2 是两个相互独立的标准正态分布随机变量.

专业课辅导+真题资料+微信公众号：432应统书院

根据(12.3.2)式,即可通过均匀分布随机数获得标准正态分布 $N(0,1)$ 随机数 η,而 $\zeta=\sigma\eta+m$ 则是正态分布 $N(m,\sigma)$ 随机数.

二、瑞利分布随机数

通过正态分布随机数可以简便地获得瑞利分布随机数. 瑞利分布的密度函数是

$$f(r)=\begin{cases}\dfrac{r}{\sigma^2}e^{-\frac{r^2}{2\sigma^2}}, & r>0;\\ 0, & r\leqslant 0,\end{cases}$$

其中 $\sigma>0$ 是分布的参数. 因为瑞利分布是具有联合密度函数

$$f(x,y)=\frac{1}{2\pi\sigma^2}\exp\left\{-\frac{x^2+y^2}{2\sigma^2}\right\}$$

的二维正态分布随机向量 (X,Y) 的模 $R=\sqrt{X^2+Y^2}$ 的分布,因而可按如下步骤产生瑞利分布随机数:

(1) 独立地产生两个正态分布 $N(0,\sigma)$ 随机数 ξ 和 η.

(2) 令 $R=\sqrt{X^2+Y^2}$ 即得参数为 σ 的瑞利分布随机数.

三、指数分布随机数

我们在 §12.2 介绍的利用反函数法产生参数为 1 的指数分布随机数需要计算均匀分布随机数的对数,其缺点是计算量大. 下面介绍一种较易计算的产生方法——冯·诺伊曼算法.

设 $\xi_1,\xi_2,\cdots$ 是一列相互独立的均匀分布随机数. 我们定义

$$N=\min\{n:n\geqslant 2,\xi_1\geqslant\xi_2\geqslant\cdots\geqslant\xi_{n-1}<\xi_n\}.$$

于是有

$$\begin{aligned}P\{N>n,\xi_1\leqslant y\}&=\int_0^1 P\{N>n,\xi_1\leqslant y\mid\xi_1=x\}\,dx\\&=\int_0^y P\{N>n\mid\xi_1=x\}\,dx.\end{aligned}\tag{12.3.3}$$

当给定 $\xi_1=x$ 时,

$$\begin{aligned}\{N>n\}&=\{x\geqslant\xi_2\geqslant\xi_3\geqslant\cdots\geqslant\xi_n\}\\&=\{\xi_i\leqslant x,i=2,\cdots,n;\xi_2\geqslant\xi_3\geqslant\cdots\geqslant\xi_n\}.\end{aligned}$$

专业课辅导+真题资料+微信公众号：432应统书院

由$\{\xi_i\}$的相互独立性知对任意给定的x有

$$P(\xi_i \leqslant x, i=2,\cdots,n)=x^{n-1}.$$

又因为$\xi_2,\xi_3,\cdots,\xi_n$是$n-1$个相互独立的在$(0,1)$中均匀分布随机数，按它们取值大小顺序的可能不同排列共有$(n-1)!$种，而满足要求$\xi_2 \geqslant \xi_3 \geqslant \cdots \geqslant \xi_n$的只是其中的一种，故

$$P\{N>n \mid \xi_1=x\}=\frac{x^{n-1}}{(n-1)!},$$

代入(12.3.3)式即得

$$P\{N>n,\xi_1 \leqslant y\}=\int_0^y \frac{x^{n-1}}{(n-1)!}\mathrm{d}x=\frac{y^n}{n!}.$$

据此我们就得到N和ξ_1的联合分布

$$\begin{aligned}P\{N=n,\xi_1 \leqslant y\}&=P\{N>n-1,\xi_1 \leqslant y\}-P\{N>n,\xi_1 \leqslant y\}\\&=\frac{y^{n-1}}{(n-1)!}-\frac{y^n}{n!}.\end{aligned}$$

上式将n对所有偶数求和得

$$\begin{aligned}&P\{N\text{为偶数},\xi_1 \leqslant y\}\\&=y-\frac{y^2}{2!}+\frac{y^3}{3!}-\frac{y^4}{4!}+\cdots\\&=1-\mathrm{e}^{-y}.\end{aligned}\tag{12.3.4}$$

这样一来，我们就有如下算法：

(1) 产生相互独立的均匀分布随机数$\xi_1,\xi_2,\cdots,\xi_N$，这里$N=\min\{n:\xi_1 \geqslant \xi_2 \geqslant \cdots \geqslant \xi_{n-1}<\xi_n\}$.

(2) 若n是偶数，接受这一游程(称这样的游程是成功的)并转至(3)；否则拒绝这一游程(称这样的游程为失败的)并返回(2).

(3) 令η=算法中的失败游程数+成功游程的第一个随机数.

下面证明η是参数为1的指数分布随机数. 首先注意到根据(12.3.4)式知出现成功游程的概率是

$$P\{N\text{为偶数}\}=P\{N\text{为偶数},\xi_1 \leqslant 1\}$$

专业课辅导+真题资料+微信公众号：432应统书院

$$=1-e^{-1}. \tag{12.3.5}$$

故出现失败游程的概率是

$$1-P\{N\text{ 为偶数}\}=e^{-1}. \tag{12.3.6}$$

对任意 $x\in(0,1)$，为使 $\eta>x$，前面 $[x]$ 个游程必须是失败游程（这里 $[x]$ 表示 x 的整数部分），而且随后的一个游程或者是失败的，或者是成功的，但同时要有 $\xi_1>x-[x]$. 因为对任意 $y\in(0,1)$，

$$\begin{aligned}&P\{N\text{ 为偶数},\xi_1>y\}\\&=P\{N\text{ 为偶数}\}-P\{N\text{ 为偶数},\xi_1\leqslant y\}\\&=(1-e^{-1})-(1-e^{-y})\\&=e^{-y}-e^{-1},\end{aligned} \tag{12.3.7}$$

故由上述关于 $\{\eta>x\}$ 的说明及(12.3.5)～(12.3.7)式得

$$\begin{aligned}P\{\eta>x\}&=e^{-[x]}\cdot e^{-1}+e^{-[x]}[e^{-(x-[x])}-e^{-1}]\\&=e^{-[x]}[e^{-1}+e^{-(x-[x])}-e^{-1}]\\&=e^{-x}.\end{aligned}$$

注意在上面的算法中，我们只需确定 $\xi_N>\xi_{N-1}$，而不必利用 ξ_N 的具体值. 因为这时 $(\xi_N-\xi_{N-1})/(1-\xi_{N-1})$ 是 $(0,1)$ 中均匀分布随机数，故可在下一游程加以利用，而省去产生一个均匀分布随机数的计算量.

四、Γ 分布和 χ^2 分布随机数

参数为 (r,λ) 的 Γ 分布 $\Gamma(r,\lambda)$ 的密度函数是

$$f(x)=\frac{1}{\Gamma(r)}(\lambda x)^{r-1}\lambda e^{-\lambda x},\quad x\geqslant 0,$$

其中 r 和 λ 是正常数. 若随机变量 η 服从 $\Gamma(r,\lambda)$ 分布，则 $\tilde{\eta}=\lambda\eta$ 服从 $\Gamma(r,1)$ 分布. 因此，可只讨论 $\Gamma(r,1)$ 分布的随机数. 下面分别就 r 是正整数，是小于 1 或大于 1 的任意实数三种情形来讨论.

情形 1 η 服从 $\Gamma(n,1)$ 分布，其中 n 是任意正整数.

这时 η 是 n 个相互独立的参数为 1 的指数分布随机变量之和. 于是若 $\xi_1,\xi_2,\cdots,\xi_n$ 是 n 个相互独立的均匀分布随机数，则

专业课辅导+真题资料+微信公众号：432应统书院

$$\eta=-\sum_{i=1}^{n}\ln\xi_i=-\ln\left(\prod_{i=1}^{n}\xi_i\right)$$

有 $\Gamma(n,1)$ 分布.

当 n 很大时,这方法的计算量很大,这时亦可用舍选法产生 $\Gamma(n,1)$ 分布随机数. 例如,可取

$$g(x)=\frac{1}{n}e^{-\frac{1}{n}x},\quad x\geqslant 0$$

作为参照的密度函数. 于是有

$$\frac{f(x)}{g(x)}=\frac{n}{\Gamma(n)}x^{n-1}e^{-\left(1-\frac{1}{n}\right)x},\quad x\geqslant 0$$

和

$$\max_{x\geqslant 0}\left[x^{n-1}e^{-\left(1-\frac{1}{n}\right)x}\right]=n^{n-1}e^{-n+1}.$$

于是若取

$$c=\frac{n^n}{\Gamma(n)}e^{l-n},$$

就有

$$f(x)\leqslant cg(x),\quad 所有\ x\geqslant 0.$$

由斯特林(Stirling)公式

$$n!\approx\sqrt{2\pi n}\,n^n e^{-n},$$

得

$$c\approx\sqrt{\frac{n}{2\pi}}e\approx 1.08\sqrt{n}.$$

这表明产生一个 $\Gamma(n,1)$ 随机数平均约需 $1.08\sqrt{n}$ 个指数分布随机数.

对舍选条件

$$\xi\leqslant\frac{f(\eta)}{cg(\eta)},$$

两边取对数并注意到

$$\ln\frac{f(\eta)}{cg(\eta)}=\ln\left[\frac{1}{n^{n-1}}\eta^{n-1}e^{-\left(\frac{n-1}{n}\right)\eta+(n-1)}\right]$$

专业课辅导+真题资料+微信公众号：432应统书院

$$=-(n-1)\left[\ln n-\ln \eta+\frac{1}{n}\eta-1\right]$$

以及当 η 有参数为 $1/n$ 的指数分布时，$\tilde{\eta}=\frac{1}{n}\eta$ 有参数为 1 的指数分布，我们就得到如下算法：

（1）独立地产生均匀分布随机数 ξ_1 和 ξ_2，并令 $\eta_1=-\ln \xi_1$ 和 $\eta_2=-\ln \xi_2$.

（2）若 $\eta_1<(n-1)(\eta_2-\ln \eta_2-1)$，则返回（1）；否则令 $\zeta=n\eta_2$，即为欲求之 $\Gamma(n,1)$ 随机数.

情形 2 η 服从 $\Gamma(r,1)$ 分布，其中 $0<r<1$.

这时 η 有密度函数

$$f(x)=\frac{1}{\Gamma(r)}x^{r-1}\mathrm{e}^{-x},\quad x\geqslant 0.$$

易见有

$$f(x)\leqslant h(x)=\begin{cases}\dfrac{1}{\Gamma(r)}x^{r-1}, & 0<x\leqslant 1;\\[2ex] \dfrac{1}{\Gamma(r)}\mathrm{e}^{-x}, & x>1.\end{cases}$$

因为

$$\int_0^{\infty} h(x)\,\mathrm{d}x=\frac{1}{\Gamma(r)}\left(\frac{1}{r}+\frac{1}{\mathrm{e}}\right),$$

故令

$$g(x)=\frac{r\mathrm{e}}{r+\mathrm{e}}\Gamma(r)h(x),\quad x\geqslant 0,$$

则 $g(x)$ 是一密度函数，对应的分布函数是

$$G(x)=\int_0^x g(y)\,\mathrm{d}y=\begin{cases}\dfrac{1}{b}x^r, & 0<x\leqslant 1;\\[2ex] 1-\dfrac{r}{b}\mathrm{e}^{-x}, & x>1,\end{cases}$$

其中 $b=1-\frac{r}{\mathrm{e}}$. $G(x)$ 的反函数是

专业课辅导+真题资料+微信公众号：432应统书院

$$G^{-1}(u)=\begin{cases}(bu)^{\frac{1}{r}}, & 0<u\leqslant\frac{1}{b};\\ -\ln\left[\frac{b}{r}(1-u)\right], & \frac{1}{b}<u<1.\end{cases}$$

因为

$$c=\max_{x>0}\frac{f(x)}{g(x)}=\frac{r+\mathrm{e}}{r\mathrm{e}}\cdot\frac{1}{\Gamma(r)},$$

故

$$\frac{f(x)}{cg(x)}=\frac{f(x)}{h(x)},$$

从而舍选条件可写为

$$\xi\leqslant\begin{cases}\mathrm{e}^{-\zeta}, & 0<\zeta\leqslant1;\\ \zeta^{r-1}, & \zeta>1,\end{cases}$$

其中 ζ 是分布为 G 的随机数. 于是，我们就得到如下的产生 $\Gamma(r,1)$ 分布随机数的算法：

(1) 产生独立的均匀分布随机数 ξ_1 和 ξ_2.

(2) 若 $b\xi_2\leqslant1$，则转到(3)；否则转到(4).

(3) 令 $\zeta=(b\xi_2)^{\frac{1}{r}}$，若 $\xi_1\leqslant\mathrm{e}^{-\zeta}$，则转到(5)；否则返回(1).

(4) 令 $\zeta=-\ln\left[\frac{b}{r}(1-\xi_2)\right]$，若 $\xi_1\leqslant\zeta^{r-1}$，则转到(5)；否则返回(1).

(5) 令 $\eta=\zeta$ 即为 $\Gamma(r,1)$ 分布随机数.

情形 3 η 服从 $\Gamma(r,1)$ 分布，其中 $r>1$.

这时利用 Γ 分布的如下可加性质：设 η_i 服从 $\Gamma(r_i,1)$ 分布，$i=1,2$，而且 η_1 和 η_2 相互独立，则 $\eta_1+\eta_2$ 有 $\Gamma(r_1+r_2,1)$ 分布. 又因为任意大于 1 的实数 r 恒有表示 $r=[r]+d$，其中 $d=r-[r]$，$0<d<1$. 于是若 η_1 是 $\Gamma([r],1)$ 分布随机数，η_2 是 $\Gamma(d,1)$ 分布随机数，η_1 和 η_2 相互独立，则 $\eta=\eta_1+\eta_2$ 是 $\Gamma(r,1)$ 分布随机数. 然而，η_1 和 η_2 的产生分别在前面的情形 1 和情形 2 已讨论过.

专业课辅导+真题资料+微信公众号：432应统书院

最后，讨论χ^2分布随机数的产生. 设随机变量η_n服从自由度为n的χ^2分布，因为它有表示

$$\eta_n=\sum_{i=1}^{n}\zeta_i^2,$$

这里$\zeta_1,\zeta_2,\cdots,\zeta_n$是相互独立的具有相同的$N(0,1)$分布的随机变量. 另一方面，$\chi_n^2$分布（自由度为$n$的$\chi^2$分布）也就是$\Gamma\left(\frac{n}{2},\frac{1}{2}\right)$分布. 特别地，$\chi_2^2$分布是$\Gamma\left(1,\frac{1}{2}\right)$分布——参数为1/2的指数分布. 故若$\xi$是均匀分布随机数，则$-2\ln\xi$有$\chi_2^2$分布. 利用$\chi^2$分布的可加性知，当$n=2k$为偶数时，$-2\sum_{i=1}^{k}\ln\xi_i$有$\chi_n^2$分布；当$n=2k+1$为奇数时，$-2\sum_{i=1}^{k}\ln\xi_i+\zeta^2$有$\chi_n^2$分布，其中$\xi_1,\xi_2,\cdots,\xi_k$是相互独立的均匀分布随机数，$\zeta$是$N(0,1)$分布随机数.

§12.4 离散型随机变量的模拟

一、一般方法

设离散型随机变量η的分布列为

$$P\{\eta=x_i\}=p_i,\quad i=0,1,2,\cdots \tag{12.4.1}$$

记

$$p^{(n)}=\sum_{i=0}^{n}p_i,\quad n=0,1,2,\cdots,$$

易见数列$\{p^{(n)}\}$将区间$(0,1]$分隔成多个小区间，记

$$I_n=(p^{(n-1)},p^{(n)}],\quad n=1,2,\cdots.$$

产生一个均匀分布随机数ξ，并令

$$\eta=x_n,\text{若 }\xi\in I_n,\quad n=1,2,\cdots.$$

于是由ξ的特性知对$n=1,2,\cdots$有

$$\begin{aligned}P\{\eta=x_n\}&=P\{\xi\in I_n\}\\&=p^{(n)}-p^{(n-1)}\end{aligned}$$

专业课辅导+真题资料+微信公众号：432应统书院

$$=p_n,$$

即 η 确有分布列(12.4.1).

例 12.4.1 几何分布随机数.

设 η 服从参数为 p 的几何分布,即有

$$P\{\eta=i\}=pq^{i-1},i=1,2,\cdots, \tag{12.4.2}$$

式中 $p,q\geqslant 0,p+q=1$. 这时对 $n=1,2,\cdots$ 有

$$p^{(n)}=P\{\eta\leqslant n\}=\sum_{i=1}^{n}p_i=\sum_{i=1}^{n}pq^{i-1}$$

$$=p\left(\sum_{j=0}^{\infty}q^j-\sum_{j=n}^{\infty}q^j\right)$$

$$=p\left(\frac{1}{p}-\frac{q^n}{p}\right)=1-q^n.$$

现产生一均匀分布随机数 ξ,并令

$$\eta=n,\text{若 } 1-q^{n-1}<\xi\leqslant 1-q^n,\quad n=1,2,\cdots, \tag{12.4.3}$$

这等价于令

$$\eta=n,\text{若 } q^n<1-\xi\leqslant q^{n-1},\quad n=1,2,\cdots. \tag{12.4.4}$$

因为若 ξ 是均匀分布随机数,则 $1-\xi$ 亦然,故在(12.4.4)式中可用 ξ 代替 $1-\xi$. 若记 $\zeta=\ln\xi/\ln q$,则亦可取

$$\begin{aligned}\eta&=\min\{j:q^j\leqslant\xi\}\\&=\min\{j:j\geqslant\zeta\}\\&=\begin{cases}[\zeta]+1, & \text{若 } \zeta \text{ 不是整数;}\\ \zeta, & \text{若 } \zeta \text{ 是整数.}\end{cases}\end{aligned}$$

这样得到的 η 就是参数为 p 的几何分布随机数.

例 12.4.2 泊松分布随机数

设 η 服从参数为 λ 的泊松分布,即有

$$p_k\equiv P\{\eta=k\}=\frac{\lambda^k}{k!}e^{-\lambda},\quad k=0,1,2,\cdots.$$

易见有

$$p_{k+1}=\frac{\lambda}{k+1}p_k$$

专业课辅导+真题资料+微信公众号：432应统书院

和

$$p^{(n+1)}=\sum_{k=0}^{n+1}p_k=p^{(n)}+p_{n+1},\quad n=0,1,2,\cdots.$$

于是有如下的产生随机数 η 的算法：

(1) 置 $k:=0,p_k=e^{-\lambda}$ 和 $p^{(k)}=0$.

(2) 产生均匀分布随机数 ξ.

(3) 令 $p_{k+1}=\dfrac{\lambda}{k+1}p_k$ 和 $p^{(k+1)}=p^{(k)}+p_{k+1}$.

(4) 若 $p^{(k)}<\xi\leqslant p^{(k+1)}$，令 $\eta=k+1$；否则令 $k:=k+1$，返回(3).

二、基于伯努利试验概型的方法

由于在具体的计算问题中，我们不一定能获得分布列 $\{p_i\}$ 的简单表达式，这时利用给定分布本身的某些特性往往能得到简便易行的模拟方法. 下面对此作一些介绍.

(A) 二项分布随机数

设 η 服从参数为 p 的二项分布，这时 η 可看作是在 n 次相互独立的“成功”概率为 p 的伯努利试验中出现“成功”的次数. 而对于每一次伯努利试验可用如下简单的方法模拟：产生一均匀分布随机数 ξ，若 $\xi\leqslant p$ 就认为试验“成功”，否则就认为“失败”.

于是，独立地产生 n 个均匀分布随机数 $\xi_1,\xi_2,\cdots,\xi_n$，并令

$$I_j=\begin{cases}1, & 若\ \xi_j\leqslant p;\\ 0, & 否则,\end{cases}$$

则 $\eta=\sum\limits_{j=1}^{n}I_j$ 就是参数为 p 的二项分布随机数.

(B) 几何分布随机数

设 η 服从参数为 p 的几何分布(12.4.2). 因为 η 可看作是一系列相互独立的“成功”概率为 p 的伯努利试验中首次取得“成功”所需要的试验次数. 于是我们有如下的产生随机数 η 的算法：

(1) 置 $k:=1$，产生一均匀分布随机数 ξ_k.

(2) 若 $\xi_k\leqslant p$，令 $\eta=k$，算法终止；若 $\xi_k>p$，转到(3).

专业课辅导+真题资料+微信公众号：432应统书院

(3) 置 $k:=k+1$ 和 $\xi_{k+1}=\beta(\xi_k-p)$，其中 $\beta=\dfrac{1}{1-p}$，返回(2).

注　算法中利用均匀分布随机数 ξ_k 产生均匀分布随机数 ξ_{k+1}，这样做可以节省计算量.

(C) 负二项分布随机数

设 ζ 服从参数为 (r,p) 的负二项分布，即有

$$p_k\equiv P\{\zeta=k\}=C_{k-1}^{r-1}p^r q^{k-r},\quad k=r,r+1,\cdots,$$

其中 r 是任意正整数，p 是任意介于 0 和 1 之间的实数，$q=1-p$.

我们知道，具有如上的负二项分布的随机变量 ζ 可以表为 r 个相互独立且具有相同参数 p 的几何分布随机变量 $\eta_1,\eta_2,\cdots,\eta_r$ 之和. 因而 ζ 可看作是在一系列相互独立的"成功"概率为 p 的伯努利试验中，首次获得 r 次"成功"所需的试验次数. 这样一来，我们有如下产生负二项分布随机数的算法：

(1) 置 $k:=1$，利用前面给出产生几何分布随机数的算法产生参数为 p 的几何分布随机数 η_k.

(2) 若 $k=r$，令 $\zeta=\sum_{i=1}^{r}\eta_i$，算法终止；若 $k<r$，置 $k:=k+1$，返回(1).

易见上述算法获得的 ζ 是具有参数 (r,p) 的负二项分布随机数.

三、其他方法

(A) 二项分布随机数

在利用上面介绍的基于伯努利试验概型的方法产生二项分布 $B(n,p)$ 随机数时，计算量随着 n 的增大而增加，故当 n 很大时，可采用如下能减少计算量的算法. 此算法利用如下事实：当给定均匀分布随机数 $\xi\leqslant p$ 时，ξ 的条件分布是 $(0,p]$ 上的均匀分布，而当给定 $\xi>p$ 时，ξ 的条件分布是 $(p,1)$ 上的均匀分布. 因而可按以下步骤产生二项分布 $B(n,p)$ 随机数 η：

(1) 令 $\alpha=1/p$ 和 $\beta=1/(1-p)$.

专业课辅导+真题资料+微信公众号：432应统书院

(2) 令 $k:=1$,产生一个均匀分布随机数 ξ_k.

(3) 若 $\xi_k \leqslant p$,令 $\xi_k=1$;若 $\xi_k>p$,令 $\xi_k=0$.

(4) 若 $k=n$,算法终止;否则令 $k:=k+1$.

(5) 取 $\xi_{k+1}=\alpha\xi_k$ 或 $\xi_{k+1}=\beta(\xi_k-p)$ 分别视 $\xi_k \leqslant p$ 与否而定,然后转回(3).

按上述步骤产生 $\xi_1,\xi_2,\cdots,\xi_n$ 后,令

$$\xi=\sum_{k=1}^{n}\xi_k,$$

这就是一个二项分布 $B(n,p)$ 随机数.

(B) 泊松分布随机数

在这里我们介绍一个利用齐次泊松过程性质产生泊松分布随机数的方法.

对于任意正整数 n,设 $\xi_1,\xi_2,\cdots,\xi_n$ 是 n 个相互独立的均匀分布随机数,则 $-\ln\xi_1,-\ln\xi_2,\cdots,-\ln\xi_n$ 是 n 个相互独立的参数为 1 的指数分布随机数. 因此,我们可把 $-\sum_{i=1}^{n}\ln\xi_i$ 看作是单位强度的齐次泊松过程的前 n 个事件发生时刻的间距之和,而这个泊松过程在区间 $(0,\lambda]$ 中发生的事件数有参数为 λ 的泊松分布,故

$$\eta=\max\left\{n:-\sum_{i=1}^{n}\ln\xi_i \leqslant \lambda\right\}$$

就是参数为 λ 的泊松分布随机数.

§12.5 随机向量和随机过程的模拟

一、随机向量的模拟

(A) 舍选法

我们在 §12.2 讨论一维随机变量模拟时介绍的舍选法思想也可以应用于多维情形. 下面设随机向量 $(\eta_1,\eta_2,\cdots,\eta_n)$ 的联合密度函数 $f(x_1,x_2,\cdots,x_n)$ 是 n 维长方体 $\prod_{i=1}^{n}(a_i,b_i)=(a_1,b_1)\times$

专业课辅导+真题资料+微信公众号：432应统书院

$(a_2,b_2)\times\cdots\times(a_n,b_n)$上的有界函数，我们选取在这个 n 维长方体上均匀分布的密度函数

$$g(x_1,x_2,\cdots,x_n)=\begin{cases}1\Big/\prod\limits_{i=1}^{n}(b_i-a_i), & (x_1,x_2,\cdots,x_n)\in\prod\limits_{i=1}^{n}(a_i,b_i);\\ 0, & \text{否则}\end{cases}$$

作参照的密度函数，这时存在正常数 c 使不等式

$$f(x_1,x_2,\cdots,x_n)\leqslant c\Big/\prod_{i=1}^{n}(b_i-a_i),\ (x_1,x_2,\cdots,x_n)\in\prod_{i=1}^{n}(a_i,b_i)$$

成立. 为了产生具有密度函数 $f(x_1,x_2,\cdots,x_n)$ 的随机数，首先产生 $n+1$ 个均匀分布随机数 $\xi_1,\xi_2,\cdots,\xi_n$ 和 ξ.

令 $\eta_i=(b_i-a_i)\xi_i+a_i(i=1,2,\cdots,n)$，则当

$$\xi\leqslant\frac{f(\eta_1,\eta_2,\cdots,\eta_n)}{c\Big/\prod\limits_{i=1}^{n}(b_i-a_i)}\tag{12.5.1}$$

时，$(\eta_1,\eta_2,\cdots,\eta_n)$ 即为我们所要求的具有 n 维分布密度 $f(x_1,x_2,\cdots,x_n)$ 的随机数；若(12.5.1)式不成立，则将这组数舍弃，重新如上再产生随机数并决定取舍，……直到获得所要求的随机数为止.

(B) 条件密度法

n 维联合密度函数 $f(x_1,x_2,\cdots,x_n)$ 可以通过条件密度表出如下：

$$\begin{aligned}&f(x_1,x_2,\cdots,x_n)\\&=f(x_n\mid x_1,\cdots,x_{n-1})f(x_{n-1}\mid x_1,\cdots,x_{n-2})\cdots f(x_2\mid x_1)f(x_1).\end{aligned}\tag{12.5.2}$$

根据上式我们可得如下的模拟方法：利用对一维情形介绍的方法，先选出具有一维密度函数 $f(x_1)$ 的随机数 η_1^0. 以 η_1^0 代入 $f(x_2\mid x_1)$ 中的 x_1，于是得到一个确定的一维密度函数 $f(x_2\mid\eta_1^0)$，据此再产生具有这一密度函数的随机数 η_2^0，一般地，若 $\eta_1^0,\eta_2^0,\cdots,\eta_{k-1}^0(k\leqslant n)$ 已确定，则产生一个具有一维密度函数 $f(x_k\mid\eta_1^0,\eta_2^0,\cdots,\eta_{k-1}^0)$ 的随机数 η_k^0. 照此方式进行下去，最后得到的 n 维向量 $(\eta_1^0,\eta_2^0,\cdots,\eta_n^0)$ 就是具有 n 维密度函数 $f(x_1,x_2,\cdots,x_n)$ 的随机数.

专业课辅导+真题资料+微信公众号：432应统书院

例 12.5.1 设(ξ,η)服从二维正态分布$N(m_1,m_2,\sigma_1,\sigma_2,r)$，即有联合密度函数

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\exp\left\{-\frac{1}{2(1-r^2)}\left[\left(\frac{x-m_1}{\sigma_1}\right)^2-\frac{2r(x-m_1)(y-m_2)}{\sigma_1\sigma_2}+\left(\frac{y-m_2}{\sigma_2}\right)^2\right]\right\}.$$

易见，当给定$\eta=y$时，ξ的条件密度函数为

$$f(x|y)=\frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-r^2}}\exp\left\{\frac{(x-m)^2}{2\sigma_1^2(1-r)^2}\right\},$$

其中

$$m=m_1+r\frac{\sigma_1}{\sigma_2}(y-m_2). \tag{12.5.3}$$

于是，可按以下步骤模拟二维正态随机向量(ξ,η)：

(1) 产生正态分布$N(m_2,\sigma_2)$的随机数η.

(2) 产生正态分布$N(m,\sigma)$的随机数ξ，其中m由(12.5.3)式给出，而

$$\sigma^2=\sigma_1^2(1-r^2),$$

则(ξ,η)就是有二维正态密度$f(x,y)$的二维随机数.

二、齐次泊松过程的模拟

设$\{N(t),t\geqslant 0\}$是强度为λ的齐次泊松过程，$\xi_1,\xi_2,\cdots$是相互独立且具有相同参数λ的指数分布随机变量，则有

$$N(t)=\max\{n:S_n=\sum_{i=0}^{n}\xi_i\leqslant t\}, \tag{12.5.4}$$

假定$S_0=\xi_0=0$.

因此，对于给定的$t>0$，齐次泊松过程$\{N(t)\}$可利用(12.5.4)式模拟.

注 若$\xi_1,\xi_2,\cdots,\xi_n,\cdots$是非负的相互独立同分布$F(x)$的随机变量序列，则由(12.5.4)式确定的$\{N(t)\}$是一个间距分布为$F(x)$的更新过程，故(12.5.4)式亦可用来模拟更新过程.

专业课辅导+真题资料+微信公众号：432应统书院

下面介绍的基于给定过程在$(0,t]$中发生的事件数时，这些事件的发生时刻的条件分布的模拟方法更为有效.

我们知道，当给定$N(t)=n$时，在$(0,t]$中的n个事件发生时刻$(S_1,S_2,\cdots,S_n)$和n个相互独立且在$(0,t]$中均匀分布的随机变量的次序统计量有相同分布. 据此有如下的模拟方法：

(1) 对于选定的T，因为$N(T)$有参数为λT的泊松分布，故首先模拟$N(T)$，设模拟结果是$N(T)=n$.

(2) 产生n个独立的均匀分布随机数$\xi_1,\xi_2,\cdots,\xi_n$，于是$T\xi_1$，$T\xi_2,\cdots,T\xi_n$是在$(0,T]$中的n个事件的发生时刻（此时尚未按它们的大小次序排列）. 若以$T\xi_{(1)},T\xi_{(2)},\cdots,T\xi_{(n)}$记$T\xi_1$，$T\xi_2,\cdots,T\xi_n$的次序统计量（即将$T\xi_1,T\xi_2,\cdots,T\xi_n$按从小到大重新排列），则有$S_1=T\xi_{(1)},S_2=T\xi_{(2)},\cdots,S_n=T\xi_{(n)}$.

当n较大时，$\xi_1,\xi_2,\cdots,\xi_n$的排列可按如下方式进行：若随机数ξ_i满足$(k-1)/n\leqslant\xi_i<k/n,k=1,2,\cdots,n$，则把$\xi_i$归入第$[n\xi_i]+1$组，其中$[\cdot]$为取整. 这样把$\xi_1,\xi_2,\cdots,\xi_n$分为若干组后，再在每一组从小到大排列. 最后，把各组归并在一起就得到统一的序. 因为每组中的ξ_i个数不会太多，组内排列就变得容易多了.

三、马尔可夫链的模拟

(A) 离散时间参数情形

设马尔可夫链$\{X_n,n=0,1,2,\cdots\}$的状态空间$E=\{0,1,2,\cdots\}$，一步转移概率矩阵为

$$\boldsymbol{P}=\begin{pmatrix} p_{00} & p_{01} & \cdots & p_{0j} & \cdots \\ p_{10} & p_{11} & \cdots & p_{1j} & \cdots \\ \vdots & \vdots & & \vdots & \\ p_{i0} & p_{i1} & \cdots & p_{ij} & \cdots \\ \vdots & \vdots & & \vdots & \end{pmatrix}$$

初始状态分布$\Pi=(\pi_0,\pi_1,\pi_2,\cdots)$. 令

$$p_i^{(n)}=\sum_{j=0}^{n}p_{ij},\quad i=0,1,2,\cdots$$

专业课辅导+真题资料+微信公众号：432应统书院

和

$$\pi^{(n)}=\sum_{j=0}^{n}\pi_j.$$

模拟马尔可夫链$\{X_n\}$的步骤如下：

（1）确定初始状态X_0.

产生均匀分布随机数ξ_0，若

$$\pi^{(i_0-1)}<\xi_0\leqslant\pi^{(i_0)},$$

则令$X_0=i_0$.

（2）产生均匀分布随机数ξ_1，若

$$p_{i_0}^{(i_1-1)}<\xi_1\leqslant p_{i_0}^{(i_1)},$$

则令$X_1=i_1$.

（3）产生均匀分布随机数ξ_2，若

$$p_{i_1}^{(i_2-1)}<\xi_2\leqslant p_{i_1}^{(i_2)},$$

则令$X_2=i_2$. 如此重复以上做法继续下去，依次确定$X_3=i_3$，$X_4=i_4$，…直到所需的X_n为止. 于是，$\{X_0=i_0,X_1=i_1,X_2=i_2,\cdots,X_n=i_n\}$是马尔可夫链$\{X_n\}$的模拟.

（B）连续时间参数情形

设马尔可夫链$\{X(t),t\geqslant0\}$的状态空间是$E=\{0,1,2,\cdots\}$，其转移强度矩阵是

$$Q=\begin{pmatrix}-q_0 & q_{01} & \cdots & q_{0j} & \cdots\\ q_{10} & -q_1 & \cdots & q_{1j} & \cdots\\ \vdots & \vdots & & \vdots & \\ q_{i0} & q_{i1} & \cdots & q_{ij} & \cdots\\ \vdots & \vdots & & \vdots & \end{pmatrix}$$

初始状态分布是$\Pi=(\pi_0,\pi_1,\pi_2,\cdots)$.

根据连续时间参数马尔可夫链的构造性定义：随机过程$\{X(t),t\geqslant0\}$是齐次马尔可夫链当且仅当它具有如下性质——对任意$i\in E$，(i) 若过程进入i后，在这种状态的逗留时间有参数为

专业课辅导+真题资料+微信公众号：432应统书院

q_i的指数分布；(ii) 若过程离开 i，则不管它在 i 的逗留时间有多长和以前经历过的状态，它离开 i 后将转移到 E 中任一异于 i 的状态 j 的概率是 $p_{ij}(=q_{ij}/q_i)$，$\sum_{j\neq i} p_{ij}=1$. 我们有如下模拟马尔可夫链 $\{X(t),0\leq t\leq T\}$ 的方法，其中 T 是预先任意给定的.

(1) 产生均匀分布随机数 ξ_0 以确定初始状态 $X(0)$：若

$$\sum_{j=0}^{i-1}\pi_j<\xi_0\leq\sum_{j=0}^{i}\pi_j,$$

则令 $X(0)=i$.

(2) 若 i 是吸收状态(即 $q_i=0$)，则算法终止. 否则产生一个参数为 q_i 的指数分布随机数 η_1 作为链 $\{X(t)\}$ 在状态 i 的逗留时间，若 $\eta_1>T$，则令 $X(t)=i$ 对 $0<t\leq T$，算法终止；否则转到(3).

(3) 产生均匀分布随机数 ξ_1 并按概率分布 $\{q_{ij}/q_i,j\neq i\}$ 确定从 i 转到状态 k，并令

$$X(t)=i,\quad 0<t<\eta_1$$

和 $X(\eta_1)=k$，转到(4).

(4) 若 k 是吸收状态，则算法终止. 否则产生一个参数为 q_k 的指数分布随机数 η_2 作为链 $\{X(t)\}$ 在状态 k 的逗留时间，若 $\eta_1+\eta_2>T$，则令 $X(t)=k$ 对 $\eta_1<t\leq T$，算法终止；否则产生均匀分布随机数 ξ_2 并按概率分布 $\{q_{kj}/q_k,j\neq k\}$ 确定从 k 转到状态 l，令

$$X(t)=k,\quad \eta_1<t<\eta_1+\eta_2$$

和 $X(\eta_1+\eta_2)=l$.

按上述做法继续下去，直到对所有 $t\leq T$，$X(t)$ 都能确定为止.

§12.6 定积分的概率计算方法

一、常用的两种算法

在§12.1节中我们已经介绍过一种最简单的积分例子，现在讨论一般的积分形式

$$I=\int_a^b f(x)\,\mathrm{d}x,\tag{12.6.1}$$

专业课辅导+真题资料+微信公众号：432应统书院

积分域(a,b)有限，而且假设$f(x)$是(a,b)上的非负有界可积函数. 设

$$0\leqslant f(x)\leqslant c.$$

易见计算这个积分就是算出图 12.6.1 中有斜线的区域A的面积. 为此，可取n对随机数$(\xi_1,\eta_1),(\xi_2,\eta_2),\cdots,(\xi_n,\eta_n)$，其中$\xi_i(i=1,2,\cdots,n)$是区间$(a,b)$上均匀分布的随机数，$\eta_i(i=1,2,\cdots,n)$是区间$(0,c)$上均匀分布的随机数. 对每一对随机数$(\xi_i,\eta_i)$，考察$\xi_i$和$\eta_i$是否满足不等式

$$f(\xi_i)>\eta_i. \tag{12.6.2}$$

易见，满足不等式(12.6.2)的随机点就是落在区域A内的点，设共有m个，而全部随机点的点数目为n，则不难证明以概率 1 成立下面的近似公式：

$$\int_a^b f(x)\,\mathrm{d}x\approx\frac{m}{n}c(b-a). \tag{12.6.3}$$

事实上，可以把上述算法设想为向矩形$a\leqslant x\leqslant b,0\leqslant y\leqslant c$中独立和均匀地掷点$P_i=(\xi_i,\eta_i)(i=1,2,\cdots,n)$. 设落于$A$中的点数为$m(=m(n))$，则易见每次成功的概率$Q=\dfrac{A}{(b-a)c}$(即图12.6.1中有斜线的区域$A$和矩形$a'abb'$面积之比)，又由强大数定律知以概率为 1 有

$$\lim_{n\to\infty}\frac{m}{n}=Q,$$

由此知当n足够大时近似式(12.6.3)以概率 1 成立.

图 12.6.1

因为n次掷点可看作是进行n次成功概率为Q的伯努利试验，m是其中成功[即点落在A中，亦即不等式(12.6.2)成立]的次数，故m有二项分布

专业课辅导+真题资料+微信公众号：432应统书院

$b(n,Q)$，其平均值为 $E(m)=nQ$，方差为 $D(m)=nQ(1-Q)$. 因此

$$E\left(\frac{m}{n}\right)=\frac{1}{n}E(m)=Q,$$
$$D\left(\frac{m}{n}\right)=\frac{1}{n^2}D(m)=\frac{Q(1-Q)}{n}. \tag{12.6.4}$$

采用这种方法计算积分时，人们自然会问，对于任意给定的正数 ε，如果要求面积之计算值和真值之差

$$\left|\frac{m}{n}c(b-a)-\int_a^b f(x)\,\mathrm{d}x\right|=\left|\frac{m}{n}c(b-a)-Qc(b-a)\right|$$

小于 ε 的概率不小于 $\alpha(0<\alpha<1)$，问 n 应取多大？根据中心极限定理，$\dfrac{m-nQ}{\sqrt{nQ(1-Q)}}$ 渐近地有正态分布 $N(0,1)$. 因此问题化为求 n，使满足

$$P\left\{\left|\frac{m}{n}c(b-a)-Qc(b-a)\right|<\varepsilon\right\}=P\left\{\left|\frac{m}{n}-Q\right|<\frac{\varepsilon}{c(b-a)}\right\}$$
$$\approx\Phi\left(\frac{\varepsilon}{c(b-a)}\sqrt{\frac{n}{Q(1-Q)}}\right)-\Phi\left(-\frac{\varepsilon}{c(b-a)}\sqrt{\frac{n}{Q(1-Q)}}\right)\geqslant\alpha. \tag{12.6.5}$$

其中，把 ε、Q、α 作为已知数，即可求出满足(12.6.5)的最小 n 值. 实用中可以通过查正态分布表找出 $t_\alpha>0$ 使

$$\Phi(t_\alpha)-\Phi(-t_\alpha)\approx\alpha.$$

例如，当 $\alpha=99.7\%$，95.6% 和 68.3% 时，t_α 分别是 3，2 和 1. 找出 t_α 后，解(12.6.5)就变成解

$$\frac{\varepsilon}{c(b-a)}\sqrt{\frac{n}{Q(1-Q)}}=t_\alpha.$$

由此得

$$n\approx\frac{t_\alpha^2c^2(b-a)^2Q(1-Q)}{\varepsilon^2}. \tag{12.6.6}$$

但是，实际上我们并不知道 Q，因此必须首先设法估计 Q. 可用试

专业课辅导+真题资料+微信公众号：432应统书院

算或其他办法得到 Q 的一个估计值 $\hat{Q}$，然后以 $\hat{Q}$ 值代入(12.6.6)式中进行计算.

上述计算方法的思想在前面 §12.1 中计算(0,1)上的积分时已用过，这就是利用随机变量在指定区域内取值的频率来计算积分，我们称之为频率法.

第二种计算积分的方法是利用随机变量的平均值来计算积分(12.6.1). 假设 ξ_i 是一列在区间 (a,b) 上均匀分布的独立随机变量，若函数 $f(x)$ 在区间 (a,b) 上可积，则 $\{f(\xi_i)\}$ 也是一列相互独立同分布的随机变量，算出它们的统计平均值并定义

$$F_n=\frac{1}{n}\sum_{i=1}^{n}f(\xi_i). \tag{12.6.7}$$

于是我们就有如下的近似计算公式

$$\int_a^b f(x)\,\mathrm{d}x\approx(b-a)F_n \quad (\text{记为 } I_n). \tag{12.6.8}$$

事实上，这时有

$$E[f(\xi_i)]=\frac{\int_a^b f(x)\,\mathrm{d}x}{b-a}=\frac{I}{b-a}, \tag{12.6.9}$$

由强大数定律知以概率 1 成立

$$\lim_{n\to\infty}(b-a)\frac{\sum_{i=1}^{n}f(\xi_i)}{n}=(b-a)E[f(\xi_i)]=I.$$

因此，当 n 足够大时，可得近似公式(12.6.8).

我们还不难算出

$$E(I_n)=\frac{(b-a)}{n}\sum_{i=1}^{n}E[f(\xi_i)]=(b-a)E[f(\xi_i)]=I.$$

若 $f(x)$ 在 (a,b) 上平方可积，则方差

$$\begin{aligned}D(I_n)&=\frac{(b-a)^2}{n^2}\sum_{i=1}^{n}D[f(\xi_i)]\\&=\frac{(b-a)^2}{n}\{E[f^2(\xi_i)]-(E[f(\xi_i)])^2\}\end{aligned}$$

专业课辅导+真题资料+微信公众号：432应统书院

$$=\frac{1}{n}\left[(b-a)\int_a^b f^2(x)\,\mathrm{d}x-\left(\int_a^b f(x)\,\mathrm{d}x\right)^2\right]. \qquad (12.6.10)$$

由此并根据中心极限定理就可以按频率法那样讨论精确度和计算次数的问题，这时有近似公式

$$n\approx\frac{1}{\varepsilon^2}t_\alpha^2(b-a)^2\cdot D[f(\xi_i)], \qquad (12.6.11)$$

其中符号 ε，t_α 的意义同前.

例 12.6.1 用蒙特卡罗方法计算积分

$$I=\int_0^1\frac{\mathrm{e}^x}{4}\mathrm{d}x. \qquad (12.6.12)$$

先用频率法计算，为此从均匀分布的随机数表（本例所用随机数表见[26]p. 484）中每次抽取一对数 δ_1 和 δ_2，例如，可以从该表中第 41 个数字开始，挨次抽一个数作为 δ_1，从第 61 个数字开始，挨次抽一个数作为 δ_2，这时要用 10^{-5} 乘以表中的数字[因附表是区间(0,99 999)内均匀分布的随机数，而我们的积分区间是(0,1)]，并四舍五入取到小数点后第三位，然后算出 $f(\delta_1)=\frac{\mathrm{e}^{\delta_1}}{4}$. 将它与 δ_2 比较，若 $f(\delta_1)>\delta_2$，认为试验成功. 设成功的次数为 m，则由(12.6.3)式知成功次数 m 除以总的试验次数 n，所得的商就是积分(12.6.12)式的近似值（这时 $b-a=1$，$c=1$）. 表 12.6.1 列出了试验次数 n 等于 1 到 20 的计算结果 I'_n.

表 12.6.1

n	1	2	3	4	5	6	7	8	9	10
δ_1	0.667	0.993	0.242	0.940	0.610	0.131	0.352	0.646	0.646	0.680
δ_2	0.491	0.182	0.192	0.025	0.557	0.530	0.865	0.105	0.564	0.136
$f(\delta_1)$	0.487	0.675	0.318	0.640	0.460	0.285	0.356	0.477	0.477	0.493
m	0	1	2	3	3	3	3	4	4	5
I'_n	0	0.5	0.667	0.750	0.600	0.500	0.429	0.500	0.444	0.500
I''_n	0.487	0.581	0.493	0.530	0.516	0.478	0.460	0.416	0.423	0.430

专业课辅导+真题资料+微信公众号：432应统书院

续表

n	11	12	13	14	15	16	17	18	19	20
δ_1	0.398	0.339	0.806	0.699	0.984	0.327	0.129	0.146	0.669	0.430
δ_2	0.579	0.541	0.238	0.432	0.033	0.462	0.284	0.161	0.990	0.547
$f(\delta_1)$	0.372	0.351	0.560	0.503	0.669	0.347	0.284	0.289	0.488	0.384
m	5	5	6	7	8	8	8	9	9	9
I'_n	0.455	0.417	0.462	0.500	0.533	0.500	0.471	0.500	0.470	0.450
I''_n	0.424	0.418	0.429	0.434	0.450	0.444	0.434	0.426	0.429	0.427

再用平均值方法来计算，这时只需要抽取随机数 δ_1 及计算 $f(\delta_1)$，然后按公式(12.6.8)算出积分近似值. 表 12.6.1 中也列出了试验次数 n 等于 1 到 20 的计算结果 I''_n.

用分析方法算出积分(12.6.12)式的准确值是

$$I=\int_0^1 \frac{e^x}{4}dx=\frac{1}{4}(e-1)=0.429\,57\cdots\approx 0.430.$$

由此得到当 $n=20$ 时，频率法的计算误差为 4.7%. 平均值法的误差比较小，为 0.7%.

现在进一步问：如果要求以 99.7% 的概率，保证 I_n 作为 I 的近似值能准确到小数点后第三位，即保证

$$|I_n-I|<0.001,$$

那么 n 最低限度应取多大？

对于频率法，由公式(12.6.6)得

$$n\approx\frac{9\times 0.43\times 0.57}{(0.001)^2}=2.205\,9\times 10^6.$$

对于平均值法，因为这时

$$D(I_n)=D\left(\frac{1}{4n}\sum_{i=1}^{n}e^{\xi_i}\right)=\frac{1}{n}\left[\frac{1}{16}\int_0^1 e^{2x}dx-\left(\frac{1}{4}\int_0^1 e^x dx\right)^2\right]$$

$$=\frac{1}{n}\left[\frac{e^2-1}{32}-\frac{(e-1)^2}{6}\right]=\frac{0.015\,1}{n}.$$

专业课辅导+真题资料+微信公众号：432应统书院

因为由(12.6.10)知 $D[f(\xi_i)]=\frac{n}{(b-a)^2}D(I_n)=0.015\ 1$，故由(12.6.11)得 $n\approx\frac{9\times0.015\ 1}{(0.001)^2}=0.135\ 9\times10^6$.

将以上结果比较可看出，平均值法较频率法好，因为它的计算误差较小，或者说在同样精确度要求下，它所需的计算次数较小.

二、重积分的计算

上面介绍的两种算法原则上也适用于多重积分，这是蒙特卡罗方法的一个优点. 现就多重积分的情形简单地介绍一下平均值方法.

以 $\boldsymbol{X}=(x_1,x_2,\cdots,x_m)$ 表示 m 维空间 $\mathbf{R}^m$ 中的点，$f(\boldsymbol{X})$ 是 m 元函数，它在有界区域 D 内可积，记

$$I=\int_D f(\boldsymbol{X})\mathrm{d}\boldsymbol{X}=\int_D\cdots\int f(x_1,x_2,\cdots,x_m)\mathrm{d}x_1\cdots\mathrm{d}x_m, \tag{12.6.13}$$

现计算 I.

取一维区间 (a,b)，使得

$$D\subset\underbrace{(a,b)\times(a,b)\times\cdots\times(a,b)}_{m个}=(a,b)^m.$$

取 m 个相互独立且在 (a,b) 中均匀分布的随机变量 $\xi_1,\xi_2,\cdots,\xi_m$，令 $\xi=(\xi_1,\xi_2,\cdots,\xi_m)$. 可以证明，在 $\xi\in D$ 的条件下，ξ 在 D 中均匀分布，这就是说，有

$$P\{\xi\in V|\xi\in D\}=\frac{|V|}{|D|}. \tag{12.6.14}$$

其中 V 是 D 的任一子域，$|V|$ 和 $|D|$ 分别表示 V 和 D 的 m 维体积. 参看图 12.6.2($m=2$ 的情形).

事实上，这时有

$$P\{\xi\in V|\xi\in D\}$$
$$=\frac{P\{\xi\in V\}}{P\{\xi\in D\}}$$

专业课辅导+真题资料+微信公众号：432应统书院

图 12.6.2

$$=\frac{|V|}{(b-a)^m}\Big/\frac{|D|}{(b-a)^m}$$
$$=\frac{|V|}{|D|}.$$

由此进一步推得

$$E[f(\xi)\mid\xi\in D]=\frac{\int_D f(\boldsymbol{X})\mathrm{d}\boldsymbol{X}}{|D|}=\frac{I}{|D|}. \tag{12.6.15}$$

根据上式，可得到以下计算方法：如同上面的 ξ 那样，构造一列相互独立的 m 维随机向量 $\{\xi^{(i)}\}$，$\xi^{(i)}=(\xi_1^{(i)},\xi_2^{(i)},\cdots,\xi_m^{(i)})$，$\xi^{(i)}$ 在 m 维立方体 $(a,b)^m$ 中均匀分布. 如果 $\xi^{(1)}\in D$，就选出 $\xi^{(1)}$，如果 $\xi^{(1)}\notin D$，则舍弃之. 如此共选出 n 个 $\xi^{(i)}$，记为 $\xi^{(i_1)},\xi^{(i_2)},\cdots,\xi^{(i_n)}$，由(12.6.15)式得到以概率 1 成立

$$|D|\lim_{n\to\infty}\frac{\sum_{k=1}^{n}f(\xi^{(i_k)})}{n}=I,$$

于是我们就可以取

$$I_n=|D|\frac{1}{n}\sum_{k=1}^{n}f(\xi^{(i_k)}) \tag{12.6.16}$$

专业课辅导+真题资料+微信公众号：432应统书院

作为积分 I 的近似值.

最后要指出一点,在实际计算中,最好把分析计算方法和概率计算方法结合起来. 这就是说,在(12.6.13)式中尽可能选用数学分析方法直接算出对某些变元的积分,以降低整个积分的重数,对于剩余的无法直接计算的变元再用平均值法,这样做往往能减少计算量,提高计算效率.

§12.7 某些方程的概率解法

一、线性方程组的求解

我们讨论线性方程组

$$\sum_{k=1}^{n} a_{ik}x_k = b_i \quad (i=1,2,\cdots,n) \tag{12.7.1}$$

的求解. 首先考虑二次型

$$Q=Q(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}\alpha_i\Big(\sum_k a_{ik}x_k-b_i\Big)^2, \tag{12.7.2}$$

这里 $\alpha_i>0$ 是常数. 易见求(12.7.1)式的解等价于求 Q 的最小点 $(x_1^0,x_2^0,\cdots,x_n^0)$. 对于任意常数 $A>0$,易见

$$E:Q(x_1,x_2,\cdots,x_n)\leqslant A$$

是一个 n 维椭球,这个椭球的中心由 $\sum_k a_{ik}x_k-b_i=0\ (i=1,2,\cdots,n)$ 确定,记为 $(x_1^0,x_2^0,\cdots,x_n^0)$. 每个通过中心的 n 维超平面 $x_k=x_k^0$ 都把椭球分成(n 维)体积相等的两部分,我们的解法就是基于这一几何性质,即求 $x_1^0,x_2^0,\cdots,x_n^0$,使椭球 E 属于 $x_k\leqslant x_k^0$ 那部分恰有一半的体积.

作 n 维立方体 $(a,b)^n$,使得椭球 $E\subset(a,b)^n$. 设 $\xi_1,\xi_2,\cdots,\xi_n$ 是 n 个在 (a,b) 上均匀分布的独立随机变量,则 $\xi=(\xi_1,\cdots,\xi_n)$ 是一 n 维随机变量,由上节关于多重积分的讨论知道,在 $\xi\in E$ 的条件下,ξ 在 E 中有均匀分布. 取 ξ 的 l 个独立的样本值 $\xi^{(1)},\xi^{(2)},\cdots,\xi^{(l)}$,如果 $\xi^{(1)}\in E$ 就选取 $\xi^{(1)}$,否则舍弃之. 设共选出 m 个

专业课辅导+真题资料+微信公众号：432应统书院

这样的样本值 $\xi^{(i_1)},\xi^{(i_2)},\cdots,\xi^{(i_m)}(m\leqslant l)$，每一 $\xi^{(i_k)}(k=1,2,\cdots,m)$ 的分量表示为

$$\xi^{(i_k)}=(\xi_1^{(i_k)},\xi_2^{(i_k)},\cdots,\xi_n^{(i_k)}).$$

因为 $\xi^{(i_k)}$ 在 E 中均匀分布，而超平面 $x_j=x_j^0$ 把椭球分成体积相等的两部分，故当 m 充当大时，$\xi^{(i_1)},\xi^{(i_2)},\cdots,\xi^{(i_m)}$ 中应差不多各有一半分别在此超平面的两侧，即这些向量的第 s 个分量 $\xi_s^{(i_1)},\xi_s^{(i_2)},\cdots,\xi_s^{(i_m)}$ 中应约有一半不大于（或不小于）x_s^0. 把这些分量按大小排列如下：

$$\xi_s^{(1)}\leqslant\xi_s^{(2)}\leqslant\cdots\leqslant\xi_s^{(m)},\tag{12.7.3}$$

其中的 $\xi_s^{\left(\left[\frac{m}{2}\right]+1\right)}$ $\left(\left[\dfrac{m}{2}\right]\right.$ 表示 $\dfrac{m}{2}$ 的整数部分$\left.\right)$ 自然应最靠近 x_s^0，因此我们可以取

$$\xi_s^{\left(\left[\frac{m}{2}\right]+1\right)}\approx x_s^0\quad(s=1,2,\cdots,n),\tag{12.7.4}$$

即以 $\xi_1^{\left(\left[\frac{m}{2}\right]+1\right)},\xi_2^{\left(\left[\frac{m}{2}\right]+1\right)},\cdots,\xi_n^{\left(\left[\frac{m}{2}\right]+1\right)}$ 作为方程组(12.7.1)的近似解.

以上是取子样(12.7.3)式的中数作为 x_s^0 的近似值，我们同样也可以取它的平均值作近似，即

$$\frac{1}{m}\sum_{k=1}^{m}\xi_s^{(k)}\equiv\frac{1}{m}\sum_{k=1}^{m}\xi_s^{(i_k)}\approx x_s^0.\tag{12.7.5}$$

现在对一类满足某种条件的线性方程组给出另一种解法.

先把(12.7.1)式改写成

$$\sum_{s=1}^{n}(\delta_{is}-h_{is})x_s=b_i\quad(i=1,2,\cdots,n),\tag{12.7.6}$$

其中

$$\delta_{is}=\begin{cases}1,i=s;\\0,i\neq s.\end{cases}$$

易见(12.7.1)式和(12.7.6)式的系数有如下关系

$$a_{is}=\begin{cases}1-h_{is},i=s;\\-h_{is},i\neq s.\end{cases}$$

专业课辅导+真题资料+微信公众号：432应统书院

如果把(12.7.6)式写成矩阵方程：

$$(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{X}=\boldsymbol{B}, \tag{12.7.7}$$

其中 $\boldsymbol{I}$ 为单位矩阵，$\boldsymbol{H}$，$\boldsymbol{B}$ 是已知的系数矩阵，$\boldsymbol{X}$ 为待求的未知向量. 令

$$\|\boldsymbol{H}\| = \max_i \left(\sum_{s=1}^{n} |h_{is}|\right),$$

我们在 $\|\boldsymbol{H}\|<1$ 的条件下给出方程(12.7.7)式的一种解法. 由条件 $\|\boldsymbol{H}\|<1$ 得

$$\boldsymbol{X}=(\boldsymbol{I}-\boldsymbol{H})^{-1}\boldsymbol{B}=(\boldsymbol{I}+\boldsymbol{H}+\boldsymbol{H}^2+\cdots)\boldsymbol{B}. \tag{12.7.8}$$

于是 $\boldsymbol{X}$ 的第 i 个分量 x_i 为

$$\begin{aligned} x_i &= b_i+(\boldsymbol{HB})_i+(\boldsymbol{H}^2\boldsymbol{B})_i+\cdots \\ &= b_i+\sum_{i_1=1}^{n} h_{ii_1}b_{i_1}+\sum_{i_1=1}^{n}\sum_{i_2=1}^{n} h_{ii_1}h_{i_1i_2}b_{i_2}+\cdots \\ &= b_i+\sum_{k=1}^{\infty}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_k=1}^{n} h_{ii_1}h_{i_1i_2}\cdots h_{i_{k-1}i_k}b_{i_k}. \end{aligned} \tag{12.7.9}$$

现在要定义一个马尔可夫链，使它的某一数字特征恰好等于 x_i. 任取一矩阵 $\boldsymbol{P}=(p_{is})(i,s=1,2,\cdots,n)$，它满足条件

$$p_{is}\geqslant 0; \qquad \sum_s p_{is}\leqslant 1; \tag{12.7.10}$$

$$p_{is}>0, \qquad \text{如果 } h_{is}\neq 0. \tag{12.7.11}$$

令

$$p_i=1-\sum_s p_{is} \quad (\text{如 } b_i\neq 0,\text{则还应取 } p_i>0). \tag{12.7.12}$$

考虑一具有以下性质的马尔可夫链：如果它从状态 i 出发，则它下一步转移到状态 s 的概率为 p_{is}，而永远停止运动的概率是 p_i. 现设它从状态 i 出发，经 k 步后停止运动，则其轨道 J 可记为

$$J: i\to i_1\to i_2\to\cdots\to i_k \quad (k\geqslant 0, i_0=i). \tag{12.7.13}$$

由定义知它有轨道(12.7.13)的概率是

$$p_{ii_1}p_{i_1i_2}\cdots p_{i_{k-1}i_k}p_{i_k}. \tag{12.7.14}$$

专业课辅导+真题资料+微信公众号：432应统书院

又在轨道 J 上定义函数

$$V(J)=\begin{cases}\dfrac{h_{ii_1}}{p_{ii_1}}\cdot\dfrac{h_{i_1i_2}}{p_{i_1i_2}}\cdots\dfrac{h_{i_{k-1}i_k}}{p_{i_{k-1}i_k}}\cdot\dfrac{b_{i_k}}{p_{i_k}}, & \text{如果 } k>0;\\ b_i/p_i, & k=0.\end{cases}\tag{12.7.15}$$

以 E_i 表示从 i 出发时的条件数学期望，则由(12.7.14)，(12.7.15)两式得

$$\begin{aligned}E_i[V(J)]&=\frac{b_i}{p_i}\cdot p_i+\sum_{i_1}\frac{h_{ii_1}b_{i_1}}{p_{ii_1}p_{i_1}}p_{ii_1}p_{i_1}\\&\quad+\sum_{i_1}\sum_{i_2}\frac{h_{ii_1}h_{i_1i_2}b_{i_2}}{p_{ii_1}p_{i_1i_2}p_{i_2}}p_{ii_1}p_{i_1i_2}p_{i_2}+\cdots\\&=b_i+\sum_{k=1}^{\infty}\sum_{i_1}\sum_{i_2}\cdots\sum_{i_k}h_{ii_1}h_{i_1i_2}\cdots h_{i_{k-1}i_k}b_{i_k}.\end{aligned}$$

将此式和(12.7.9)式比较即得

$$x_i=E_i[V(J)].\tag{12.7.16}$$

于是，我们得到如下的计算步骤：

1° 用§12.5所介绍的方法，模拟转移概率矩阵 $\boldsymbol{P}=(p_{is})$ 和初始状态为 i 的马尔可夫链；

2° 独立地造出 n 个形如(12.7.13)的轨道 $J_1,J_2,\cdots,J_n$ 并按(12.7.15)算出 $V(J_1),V(J_2),\cdots,V(J_n)$；

3° 以 $\dfrac{1}{n}\sum\limits_{k=1}^{n}V(J_k)$ 作为 x_i 的估计值.

本法的优点是每个 x_i 可单独求出，而不需要同时算出其他的 x_s. 同时，这方法的原理还可应用于求解某些积分方程.

二、一些偏微分方程的求解

我们首先讨论一个最典型的问题：二维拉普拉斯方程 $\Delta u=0$ $\left(即\dfrac{\partial^2u}{\partial x^2}+\dfrac{\partial^2u}{\partial y^2}=0\right)$ 的狄利克雷问题. 即设有某个区域 D，在其边界 Γ 上给定某个函数 $f(Q)$，试求函数 $u(P)$，使得它在给定区域 D 的内部满足拉普拉斯方程

专业课辅导+真题资料+微信公众号：432应统书院

$$\Delta u=0, \tag{12.7.17}$$

在边界 Γ 上取给定函数值

$$u|_{\Gamma}=f(Q), \tag{12.7.18}$$

其中 $Q\in\Gamma$.

这个问题可以化为如下的有限差分问题：给定步长 h，并在所考虑的平面上作正方形网格，我们只研究落在区域 D 上的网格点，网格点分成两类：一类是它的四个相邻网格点位于区域内部或边界 Γ 上，这类点称为内部网格点，在图 12.7.1 中以△标出. 另一类是它的四个相邻网格点不全在区域内部或边界上，这类点称为边界网格点，在图 12.7.1 中以○标出. 在内部网格点 P 上可以建立相应于原来的微分方程组的有限差分方程组

$$u(P)=\frac{1}{4}[u(P_1)+u(P_2)+u(P_3)+u(P_4)], \tag{12.7.19}$$

其中 P_1,P_2,P_3,P_4 是内点 P 的四个相邻网格点. 对于边界网格点 Q 上的取值，则按照某种规则利用给定在区域边界 Γ 上的函数 f 来确定，记为

$$u(Q)=f(Q). \tag{12.7.20}$$

图 12.7.1

现在来研究模拟方程组(12.7.19)及边界条件(12.7.20)的概率模型. 考虑一个质点在所有内部网格点和边界网格点上作随

专业课辅导+真题资料+微信公众号：432应统书院

机游动①:质点可从区域内任一网格点 P 出发,以相等的概率(即各等于 1/4)游动到任一个相邻的网格点,而当它游动到某个内部网格点后,又继续以相同的概率游动到它的四个相邻网格点中的一点,如此下去,一直到质点游动到某一边界点时就停止在那里不再游动,这就产生如何求质点从内网格点 P 出发,最终游动到边界网格点 Q 的概率 $u(P,Q)$ 的问题. 要推出这个概率的明显表示式,一般是复杂和困难的. 但是,我们的问题只需推出有关这概率的某些关系式即可. 同时概率论中已经证明,质点经过有限步游动后最终停在某个边界点的概率等于 1(参看[14]),这保证质点实际上不会无休止地在区域内部游动. 易见事件"质点从 P 游动到 Q"是以下述四个不相容事件之和:"质点从 P 游动到 P_1,然后从 P_1 游动到 Q";"质点从 P 游动到 P_2,然后从 P_2 游动到 Q";"质点从 P 游动到 P_3,然后从 P_3 游动到 Q";"质点从 P 游动到 P_4,然后从 P_4 游动到 Q". 因此,按全概率公式有

$$u(P,Q)=\frac{1}{4}[u(P_1,Q)+u(P_2,Q)+u(P_3,Q)+u(P_4,Q)]. \tag{12.7.21}$$

易见(12.7.21)式和(12.7.19)式是一致的. 此外,由模型本身的性质知概率 $u(P,Q)$ 满足下面的边界条件

$$u(Q',Q)=\begin{cases}1, \text{当 } Q'=Q;\\ 0, \text{当 } Q'\neq Q,\end{cases} \tag{12.7.22}$$

这里 Q' 和 Q 皆为边界网格点.

已经知道,在给定的边界条件下存在满足方程(12.7.19)式的唯一解.

假定模拟质点游动 N 次,每次皆从 P 出发,如果其中质点游动最终到达边界点 Q 的试验次数为 M,则频率 $\frac{M}{N}$ 给出满足边界条

① 这一模型在国外有关文献中常称为"醉汉问题".

专业课辅导+真题资料+微信公众号：432应统书院

件(12.7.22)的方程(12.7.21)的近似解,即 $u(P,Q)\approx\frac{M}{N}$.

如果我们还规定当质点游动到边界网格点 Q 后,我们要付出代价 $f(Q)$,则由 P 出发的质点游动所要付出的代价是一随机变量,我们用 $\xi(P)$ 表之. 设共有 s 个边界网格点 $Q_i(i=1,2,\cdots,s)$. 于是我们为质点游动付出代价 $f(Q_i)$ 的概率为 $u(P,Q_i)$,因而付出代价的平均值为

$$W(P)=E[\xi(P)]=\sum_{i=1}^{s}f(Q_i)u(P,Q_i). \tag{12.7.23}$$

显然 $W(P)$ 依赖于出发点 P,容易验证这函数也满足差分方程

$$W(P)=\frac{1}{4}[W(P_1)+W(P_2)+W(P_3)+W(P_4)], \tag{12.7.24}$$

同时还满足所要求的边界条件. 事实上,在(12.7.23)式中,令 $P=Q$,则由(12.7.22)式可得 $W(Q)=f(Q)$,于是函数 $W(P)$ 就是我们所要求的解.

以上是就二维的情形加以讨论,这方法还可以应用于多维情形和更一般的边界问题. 我们可以证明,如果所考虑的区域为边长等于 r 的 n 维立方体,即变量 $x_1,x_2,\cdots,x_n$ 在 n 维立方体:$0\leqslant x_i\leqslant r(i=1,2,\cdots,n)$ 内取值,则质点从区域内网格点 P 出发游动直至到达边界网格点 Q 为止所走的步数的平均值 $z(P)$ 满足不等式

$$z(P)<r^2/2.$$

这就是说,质点游动的时间不依赖于网格的维数,这是蒙特卡罗方法的一个重要特点,即它较之一般方法更适用于高维的情形.

对于更一般的二阶线性椭圆形方程

$$a(x,y)\frac{\partial^2 u}{\partial x^2}+b(x,y)\frac{\partial^2 u}{\partial x\partial y}+c(x,y)\frac{\partial^2 u}{\partial y^2}+d(x,y)\frac{\partial u}{\partial x}$$

$$+e(x,y)\frac{\partial u}{\partial y}+f(x,y)u=0, \tag{12.7.25}$$

(其中 $ac-b^2>0$),它的解可以利用推广了的随机游动模型,即质点

专业课辅导+真题资料+微信公众号：432应统书院

从某一个网格点游动到另一个网格点的转移概率可以不相等，它依赖于质点当时所在的网格点.

我们还可以研究更一般的边界条件，例如

$$\left.\frac{\partial u}{\partial n}\right|_{\Gamma}=ku|_{\Gamma}+f(Q),$$

这时用来求解的游动模型应当假定质点到达边界点后能以正概率返回区域内继续游动，而不一定停止在该边界点上.

至于其他类型的偏微分方程，其中有一些也能用类似的方法求解.

专业课辅导+真题资料+微信公众号：432应统书院

下册习题答案

第 六 章

1. 0.066 81.

3. $F_{(1,1)}$ 分布.

4. $m=1, k=1/16, \chi_1^2$ 分布.

5. $t_{(15)}$ 分布.

6. $n \geqslant 25$.

8. (1) 0.132; (2) 0.582; (3) 0.291.

9. (1) $e^{-7.2}$; (2) $(1-e^{-4.5})^6$.

11. 0.671.

12. $f_D(y)=\begin{cases}\dfrac{1}{\sqrt{\pi}\sigma}e^{-\frac{y^2}{4\sigma^2}}, & y \geqslant 0;\\ 0, & y<0,\end{cases}$

$$f_n(u)=\frac{1}{\pi\sigma}e^{-\frac{(u-a)^2}{2\sigma^2}}\int_0^{\frac{u-a}{\sigma}} e^{-\frac{x^2}{2}}dx,$$

$$f_1(v)=\frac{2}{\sqrt{2\pi}\sigma}e^{-\frac{(v-a)^2}{2\sigma^2}}-f_n(u).$$

14. $\eta \sim t_{(n-1)}$.

16. $f_n(u)=n\left[u-\left(\theta-\dfrac{1}{2}\right)\right]^{n-1}$;

$$f_1(v)=n\left[\left(\theta+\frac{1}{2}\right)-v\right]^{n-1};$$

$$f(u,v)=n(n-1)(u-v)^{n-2}.$$

第 七 章

1. 提示：$P\{\bar{\xi} \geqslant M\}=0.01, 2n\bar{\xi}/\beta(\alpha+1) \sim \chi_{2n}^2$. $M=0.565$.

专业课辅导+真题资料+微信公众号：432应统书院

2. $1-e^{-\beta M}, M\geqslant 0$.

3. -0.525.

4. (1) 当 $n=2k+1$ 时 $\hat{\theta}=\xi_{k+1}^{*}$；当 $n=2k$ 时，$\xi_k^{*}<\hat{\theta}<\xi_{k+1}^{*}$；

(2) $\hat{\theta}=\dfrac{-1}{n\sum\limits_{i=1}^{n}\ln\xi_i}$；(3) $\hat{\theta}=\max\limits_{1\leqslant i\leqslant n}\xi_i$；(4) $\hat{\theta}=\xi$.

5. $\hat{p}=1-\dfrac{S^2}{\bar{\xi}}, \hat{N}=\bar{\xi}/\hat{p}$.

6. $\hat{a}=\lg\bar{\xi}-\dfrac{\ln 10}{2}\hat{\sigma}^2$，$\hat{\sigma}^2=\dfrac{1}{\ln 10}\ln\dfrac{S^2-\bar{\xi}^2}{\bar{\xi}^2}$.

7. (1) $\hat{\beta}=\dfrac{1}{\bar{T}-t_0}$；(2) $\hat{t}_0=\min\limits_{1\leqslant i\leqslant n}T_i$.

8. (1) $k=n\sqrt{\dfrac{2}{\pi}}$；(2) $k=2(n-1)$.

10. $\hat{a}_2$的方差最小.

11. $c_1=(\sigma_1^2-\rho\,\sigma_1\sigma_2)/(\sigma_1^2+\sigma_2^2-2\rho\,\sigma_1\sigma_2)$；

$c_2=(\sigma_2^2-\rho\,\sigma_1\sigma_2)/(\sigma_1^2+\sigma_2^2-2\rho\,\sigma_1\sigma_2)$.

12. $\hat{a}=\bar{\xi}-\bar{\eta}, n_1=\left[\dfrac{n}{3}\right]$ （[ ]表示取整数部分）.

14. $\dfrac{4}{3}\max\limits_{1\leqslant i\leqslant 3}\xi_i$较 $4\min\limits_{1\leqslant i\leqslant 3}\xi_i$有效.

15. $c_1=\dfrac{1}{3}$，$c_2=\dfrac{2}{3}$.

17. (1) $\dfrac{n_1S_1^2\sigma_2^2(n_2-1)}{n_2S_2^2\sigma_1^2(n_1-1)}\sim F_{(n_1-1,n_2-1)}$.

(2) $\dfrac{(\bar{\xi}-\bar{\eta})-(a_1-a_2)}{\sigma\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ 服从正态 $N(0,1)$；

18. $P\left\{(\bar{\xi}-\bar{\eta})-t_{(n-1)}(0.05)\dfrac{S}{\sqrt{n-1}}<a_1-a_2\right.$

$\left.<(\bar{\xi}-\bar{\eta})+t_{(n-1)}(0.05)\dfrac{S}{\sqrt{n-1}}\right\}=1-\alpha$,

其中 $S^2=S_1^2+S_2^2+2S_{12}$.

专业课辅导+真题资料+微信公众号：432应统书院

19. $(-0.07, 1.23)$；$(0.726, 3.63)$.

20. $\bar{x}=2, S^*=\frac{\sqrt{52}}{3}$；$(0.607, 3.393)$，$(3.415, 17.707)$.

第 八 章

1. 算得 $t_{(10+12-2)}=-0.065$，查表得 $t_{20}(0.05)=2.09$，由于 $|-0.065|<2.09$，所以可认为平均值之间的差异不显著.

2. 算得 $F_{(5,8)}=1.03$，查表得 $F_{(5,8)}(0.05)=3.69$，由于 $1.03<3.69$，所以可认为方差之间无显著性的差异.

3. 算得 $\chi^2_{8-1-1}=1.254$，查表得 $\chi^2_6(0.05)=12.59$，由于 $1.254<12.59$，所以可认为服从泊松分布.

4. 算得 $\chi^2_{6-1-1}=1.551$，查表得 $\chi^2_4(0.10)=7.779$，由于 $1.551<7.779$，所以可认为服从泊松分布.

5. 算得 $\chi^2_{10-1}=5.125$，查表得 $\chi^2_9(0.10)=14.684$，由于 $5.125<14.684$，所以可认为服从均匀分布.

6. 算得 $\chi^2_{11-1-1}=1.1929$，查表得 $\chi^2_9(0.10)=14.684$，由于 $1.1929<14.684$，所以可认为服从二项分布.

10. 当原假设 H_0 成立时，2ζ 服从正态 $N(0,1)$.

11. 记 $\zeta_n=\sum_{i=1}^{n}\xi_i$，当原假设 H_0 成立时，$\zeta_n\sim\chi^2_n$，

(1) $P\{\zeta_n\geqslant(4\ln c+4n\ln 2)\}=0.05$，

(2) $P\{\zeta_n\leqslant(2n\ln 2-2\ln c)\}=0.05$.

给定显著性水平 $\alpha=0.05$ 及样本的容量 n，在 χ^2-分布临界值表中查得 $\chi^2_n(0.05)$ 值，由这个值可确定常数 c 值.

12. 记 $\zeta_n=\sum_{i=1}^{n}\xi_i$，则 ζ_n 服从二项分布 $b(n;p)$，其中 $p=\theta$，则

(1) 当 H_0 成立时，ζ_n 服从 $b\left(n;\frac{1}{2}\right)$，常数 c 由下式确定：

$$P\left\{\zeta_n\leqslant\frac{1}{\ln 2}\left[n\ln\frac{4}{3}-\ln c\right]\right\}=0.05.$$

(2) 当 H_0 成立时，ζ_n 服从 $b\left(n;\frac{1}{3}\right)$，常数 c 由下式确定：

专业课辅导+真题资料+微信公众号：432应统书院

$$P\left\{\zeta_n \geqslant \frac{1}{\ln 2}\left[n\ln\frac{4}{3}+\ln c\right]\right\}=0.05.$$

13. 记 $\zeta_n=\sum_{i=1}^{n}\xi_i$，则 ζ_n 服从参数为 $n\lambda$ 的泊松分布. 当 H_0 成立时，ζ_n 的概率函数为：

$$P\{\zeta_n=x\}=\frac{(0\cdot n)^x}{x!}e^{-0\cdot n}.$$

给定显著性水平 $\alpha=0.05$，常数 c 由下式确定：

$$P\left\{\zeta_n \geqslant \frac{1}{\ln 10}[0.9n+\ln c]\right\}=0.05.$$

14. $n=250$.

15. $n=11$.

第 九 章

1. (1) $\hat{\beta}_0=-0.35$，$\hat{\beta}_1=0.81$.

(2) 在显著性水平 $\alpha=0.05$ 下，否定 H_0.

(3) $\hat{\beta}_{11}=0.73$，$\hat{\beta}_{12}=0.89$.

(4) $\hat{\sigma}_e^2=0.33$.

(5) $\hat{y}_1=\hat{y}-0.2541\sqrt{1+\sum_{i=0}^{1}\sum_{j=0}^{1}c_{ij}x_ix_j}$，

$\hat{y}_2=\hat{y}+0.2541\sqrt{1+\sum_{i=0}^{1}\sum_{j=0}^{1}c_{ij}x_ix_j}$，

其中 $\hat{y}=-0.35+0.81x$，

$$\sum_{i=0}^{1}\sum_{j=0}^{1}c_{ij}x_ix_j=\frac{91.51}{817}-\frac{19.8}{817}x+\frac{10}{817}x^2.$$

2. $\hat{y}=22.44-53.25\frac{1}{x_1}+0.34x_2$.

5. $\hat{\theta}=\frac{1}{5}(\xi_1+2\xi_2)$，$Q_e=\frac{1}{5}(2\xi_1-\xi_2)^2$.

专业课辅导+真题资料+微信公众号：432应统书院

6. $\hat{\beta}=\begin{pmatrix}\frac{1}{3}(y_1+y_2+y_3)\\ -\frac{1}{2}(y_1-y_3)\\ \frac{1}{6}(y_1-2y_2+y_3)\end{pmatrix}$,

当 $\beta_2=0$ 时，则有

$$\hat{\beta}=\begin{pmatrix}\frac{1}{3}(y_1+y_2+y_3)\\ -\frac{1}{2}(y_1-y_2)\end{pmatrix}.$$

7. $\hat{\theta}=\bar{y}$, $\hat{\phi}=\frac{1}{3}[\bar{y}_{(2)}-\bar{y}_{(3)}]$,

其中 $\bar{y}=\frac{1}{5n}\sum_{i=1}^{5n}y_i$,

$\bar{y}_{(2)}=\frac{1}{2n}\sum_{i=1}^{2n}y_{2n+i}$,

$\bar{y}_{(3)}=\frac{1}{n}\sum_{i=1}^{n}y_{4n+i}$.

8. $\hat{a}=\frac{1}{6}(y_1+2y_2+y_3)$, $\hat{b}=\frac{1}{5}(-y_2+2y_3)$.

9. $\hat{\beta}=-11$, $\hat{\beta}_1=37$，否定 H_0.

10. $\hat{\beta}_1=9.7$, $\hat{\beta}_2=5.0$, $\hat{\beta}_3=4.4$, $\hat{\beta}_4=1.1$.

12. $\hat{y}=-22.97+18.08x$,

不否定 $H_0:\beta=0$.

13. $\hat{Z}=0.806\ 3+0.065\ 3x$,

否定 $H_0:B=0$,

$\hat{y}=6.401\ 8\times1.162\ 3^x$.

14. $\hat{y}_t=2\ 049\times1.252^{1.2t}$.

15. $\hat{Q}=0.003\ 6K^{0.492\ 7}L^{1.177\ 7}$.

专业课辅导+真题资料+微信公众号：432应统书院

第 十 章

1. $a_1=\dfrac{\sigma_p^2\bar{\xi}}{\sigma_0^2+1/n},\dfrac{1}{\sigma_1^2}=n+\dfrac{1}{\sigma_0^2}.$

   a 的后验分布为 $N(a_1,\sigma_1)$.

2. $\hat{\lambda}=\dfrac{n+\alpha}{T+\dfrac{1}{\beta}},T=\sum\limits_{i=1}^{n}\xi_i$

3. $R(a,d_1)=\dfrac{39}{25}a^2,R(a,d_2)=\dfrac{13}{6}a^2,$

   $R(a,d_3)=\dfrac{5}{3}a^2.$

4. $C=\dfrac{1}{n+2}.$

5. $C=\dfrac{n}{n+1}.$

6. $\hat{p}=\dfrac{2+\sum\limits_{i=1}^{n}\xi_i}{n+4}.$

7. $\hat{\lambda}=\left(1+\dfrac{1}{n}\right)^{1+\sum\limits_{i=1}^{n}\xi_i}.$

附　表

表 1　χ^2-分布的上侧临界值表

$$P\{\chi_f^2>\chi_\alpha^2\}=\alpha$$

f \ α	0.99	0.98	0.95	0.90	0.80	0.70	0.50	0.30	0.20	0.10	0.05	0.02	0.01	0.001	α \ f
1	0.0^3157	0.0^3628	0.0^2393	0.0158	0.0642	0.148	0.455	1.074	1.642	2.706	3.841	5.412	6.635	10.828	1
2	0.0201	0.0404	0.103	0.211	0.446	0.713	1.386	2.408	3.219	4.605	5.991	7.824	9.210	13.816	2
3	0.115	0.185	0.352	0.584	1.005	1.424	2.366	3.665	4.624	6.251	7.815	9.837	11.345	16.266	3
4	0.297	0.429	0.711	1.064	1.649	2.195	3.357	4.878	5.989	7.779	9.488	11.668	12.277	18.467	4
5	0.554	0.752	1.145	1.610	2.343	3.000	4.351	6.064	7.289	9.236	11.070	13.388	15.068	20.515	5
6	0.872	1.134	1.635	2.204	3.070	3.828	5.348	7.231	8.558	10.645	12.592	15.033	16.812	22.458	6
7	1.239	1.564	2.167	2.833	3.822	4.671	6.346	8.383	9.803	12.017	14.067	16.622	18.475	24.322	7
8	1.646	2.032	2.733	3.490	4.594	5.527	7.344	9.524	11.030	13.362	15.507	18.168	20.090	26.125	8
9	2.088	2.532	3.325	4.168	5.380	6.393	8.343	10.656	12.242	14.684	16.919	19.679	21.666	27.877	9
10	2.558	3.059	3.940	4.865	6.179	7.267	9.342	11.781	13.442	15.987	18.307	21.161	23.209	29.588	10
11	3.053	3.609	4.575	5.578	6.989	8.148	10.341	12.899	14.631	17.275	19.675	22.618	24.725	31.264	11
12	3.571	4.178	5.226	6.304	7.807	9.034	11.340	14.011	15.812	18.549	21.026	24.054	26.217	32.909	12

专业课辅导+真题资料+微信公众号：432应统书院

续表

f \ α	0.99	0.98	0.95	0.90	0.80	0.70	0.50	0.30	0.20	0.10	0.05	0.02	0.01	0.001	α / f
13	4.107	4.765	5.892	7.042	8.634	9.926	12.340	15.119	16.985	19.812	22.362	25.472	27.688	34.528	13
14	4.660	5.368	6.571	7.790	9.467	10.821	13.339	16.222	18.151	21.064	23.685	26.873	29.141	36.123	14
15	5.229	5.985	7.261	8.547	10.307	11.721	14.339	17.322	19.311	22.307	24.996	28.259	30.578	37.697	15
16	5.812	6.614	7.962	9.312	11.152	12.624	15.338	18.418	20.465	23.542	26.296	29.633	32.000	39.252	16
17	6.408	7.255	8.672	10.085	12.002	13.531	16.338	19.511	21.615	24.769	27.587	30.995	33.409	40.790	17
18	7.015	7.906	9.390	10.865	12.857	14.440	17.338	20.601	22.760	25.989	28.869	32.346	34.805	42.312	18
19	7.633	8.567	10.117	11.651	13.716	15.352	18.338	21.689	23.900	27.204	30.144	33.687	36.191	43.820	19
20	8.260	9.237	10.851	12.443	14.578	16.266	19.337	22.775	25.038	28.412	31.410	35.020	37.566	45.315	20
21	8.897	9.915	11.591	13.240	15.445	17.182	20.337	23.858	26.171	29.615	32.671	36.343	38.932	46.797	21
22	9.542	10.600	12.338	14.041	16.314	18.101	21.337	24.939	27.301	30.813	33.924	37.659	40.289	48.268	22
23	10.196	11.293	13.091	14.848	17.187	19.021	22.337	26.018	28.429	32.007	35.172	38.968	41.638	49.728	23
24	10.856	11.992	13.848	15.659	18.062	19.943	23.337	27.096	29.553	33.196	36.415	40.270	42.980	51.179	24
25	11.524	12.697	14.611	16.473	18.940	20.867	24.337	28.172	30.675	34.382	37.652	41.566	44.314	52.618	25
26	12.198	13.409	15.379	17.292	19.820	21.792	25.336	29.246	31.795	35.563	38.885	42.856	45.642	54.052	26
27	12.879	14.125	16.151	18.114	20.703	22.719	26.336	30.319	32.912	36.741	40.113	44.140	46.963	55.476	27
28	13.565	14.847	16.928	18.939	21.588	23.647	27.336	31.391	34.027	37.916	41.337	45.419	48.278	56.893	28
29	14.256	15.574	17.708	19.768	22.475	24.577	28.336	32.461	35.139	39.087	42.557	46.693	49.588	58.301	29
30	14.953	16.306	18.493	20.599	23.364	25.508	29.336	33.530	36.250	40.256	43.773	47.962	50.892	59.703	30

专业课辅导+真题资料+微信公众号：432应统书院

表 2　t–分布的双侧临界值表

$P\{|t|>t_\alpha\}=\alpha$

f \ α	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0.05	0.02	0.01	0.001	α \ f
1	0.158	0.325	0.510	0.727	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	636.619	1
2	0.142	0.289	0.445	0.617	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.598	2
3	0.137	0.277	0.424	0.584	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	12.924	3
4	0.134	0.271	0.414	0.569	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610	4
5	0.132	0.267	0.408	0.559	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	6.859	5
6	0.131	0.265	0.404	0.553	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959	6
7	0.130	0.263	0.402	0.549	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	5.405	7
8	0.130	0.262	0.399	0.546	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041	8
9	0.129	0.261	0.398	0.543	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781	9
10	0.129	0.260	0.397	0.542	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587	10
11	0.129	0.260	0.396	0.540	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437	11
12	0.128	0.259	0.395	0.539	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318	12

专业课辅导+真题资料+微信公众号：432应统书院

续表

f \ α	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0.05	0.02	0.01	0.001	α \ f
13	0.128	0.259	0.394	0.538	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221	13
14	0.128	0.258	0.393	0.537	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140	14
15	0.128	0.258	0.393	0.536	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073	15
16	0.128	0.258	0.392	0.535	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015	16
17	0.128	0.257	0.392	0.534	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965	17
18	0.127	0.257	0.392	0.534	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922	18
19	0.127	0.257	0.391	0.533	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883	19
20	0.127	0.257	0.391	0.533	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850	20
21	0.127	0.257	0.391	0.532	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.819	21
22	0.127	0.256	0.390	0.532	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792	22
23	0.127	0.256	0.390	0.532	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.767	23
24	0.127	0.256	0.390	0.531	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.745	24
25	0.127	0.256	0.390	0.531	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725	25
26	0.127	0.256	0.390	0.531	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707	26
27	0.127	0.256	0.389	0.531	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690	27
28	0.127	0.256	0.389	0.530	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674	28
29	0.127	0.256	0.389	0.530	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659	29
30	0.127	0.256	0.389	0.530	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.646	30
40	0.126	0.255	0.388	0.529	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551	40
60	0.126	0.254	0.387	0.527	0.679	0.848	1.046	1.296	1.671	2.000	2.390	2.660	3.460	60
120	0.126	0.254	0.386	0.526	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617	3.373	120
∞	0.126	0.253	0.385	0.524	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.291	∞

专业课辅导+真题资料+微信公众号：432应统书院

表 3　F 检验的临界值(F_α)表

$$P\{F>F_\alpha\}=\alpha$$

$\alpha=0.10$

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	15	20	30	50	100	200	500	∞	n_2 \ n_1
1	39.9	49.5	53.6	55.8	57.2	58.2	58.9	59.4	59.9	60.2	61.2	61.7	62.3	62.7	63.0	63.2	63.3	63.3	1
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.42	9.44	9.46	9.47	9.48	9.49	9.49	9.49	2
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.20	5.18	5.17	5.15	5.14	5.14	5.14	5.13	3
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.87	3.84	3.82	3.80	3.78	3.77	3.76	3.76	4
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.24	3.21	3.17	3.15	3.13	3.12	3.11	3.10	5
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.87	2.84	2.80	2.77	2.75	2.73	2.73	2.72	6
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.63	2.59	2.56	2.52	2.50	2.48	2.48	2.47	7
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.46	2.42	2.38	2.35	2.32	2.31	2.30	2.29	8
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.34	2.30	2.25	2.22	2.19	2.17	2.17	2.16	9
10	3.28	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.24	2.20	2.16	2.12	2.09	2.07	2.06	2.06	10
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.17	2.12	2.08	2.04	2.00	1.99	1.98	1.97	11

专业课辅导+真题资料+微信公众号：432应统书院

$\alpha = 0.10$

续表

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	15	20	30	50	100	200	500	∞	n_2 \ n_1
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.10	2.06	2.01	1.97	1.94	1.92	1.91	1.90	12
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.05	2.01	1.96	1.92	1.88	1.86	1.85	1.85	13
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.01	1.96	1.91	1.87	1.83	1.82	1.80	1.80	14
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	1.97	1.92	1.87	1.83	1.79	1.77	1.76	1.76	15
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.94	1.89	1.84	1.79	1.76	1.74	1.73	1.72	16
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.91	1.86	1.81	1.76	1.73	1.71	1.69	1.69	17
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.89	1.84	1.78	1.74	1.70	1.68	1.67	1.66	18
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.86	1.81	1.76	1.71	1.67	1.65	1.64	1.63	19
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.84	1.79	1.74	1.69	1.65	1.63	1.62	1.61	20
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.81	1.76	1.70	1.65	1.61	1.59	1.58	1.57	22
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.78	1.73	1.67	1.62	1.58	1.56	1.54	1.53	24
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.76	1.71	1.65	1.59	1.55	1.53	1.51	1.50	26
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.74	1.69	1.63	1.57	1.53	1.50	1.49	1.48	28
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.72	1.67	1.61	1.55	1.51	1.48	1.47	1.46	30
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.66	1.61	1.54	1.48	1.43	1.41	1.39	1.38	40
50	2.81	2.41	2.20	2.06	1.97	1.90	1.84	1.80	1.76	1.73	1.63	1.57	1.50	1.44	1.39	1.36	1.34	1.33	50
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.60	1.54	1.48	1.41	1.36	1.33	1.31	1.29	60
80	2.77	2.37	2.15	2.02	1.92	1.85	1.79	1.75	1.71	1.68	1.57	1.51	1.44	1.38	1.32	1.28	1.26	1.24	80
100	2.76	2.36	2.14	2.00	1.91	1.83	1.78	1.73	1.70	1.66	1.56	1.49	1.42	1.35	1.29	1.26	1.23	1.21	100
200	2.73	2.33	2.11	1.97	1.88	1.80	1.75	1.70	1.66	1.63	1.52	1.46	1.38	1.31	1.24	1.20	1.17	1.14	200
500	2.72	2.31	2.10	1.96	1.86	1.79	1.73	1.68	1.64	1.61	1.50	1.44	1.36	1.28	1.21	1.16	1.12	1.09	500
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.49	1.42	1.34	1.26	1.18	1.13	1.08	1.00	∞

专业课辅导+真题资料+微信公众号：432应统书院

$\alpha = 0.05$

续表

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	n_2 \ n_1
1	161	200	216	225	230	234	237	239	241	242	244	245	246	247	248	1
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	2
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.71	8.69	8.67	8.66	3
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.87	5.84	5.82	5.80	4
5	6.61	5.79	5.41	6.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.64	4.60	4.58	4.56	5
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.96	3.92	3.90	3.87	6
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.53	3.49	3.47	3.44	7
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.24	3.20	3.17	3.15	8
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.03	2.99	2.96	2.94	9
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.86	2.83	2.80	2.77	10
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.74	2.70	2.67	2.65	11
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.64	2.60	2.57	2.54	12
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.55	2.51	2.48	2.46	13
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.48	2.44	2.41	2.39	14
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.42	2.38	2.35	2.33	15
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.37	2.33	2.30	2.28	16
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.33	2.29	2.26	2.23	17
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.29	2.25	2.22	2.19	18
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.26	2.21	2.18	2.16	19
20	4.35	3.49	3.10	2.87	2.21	2.60	2.51	2.45	2.39	2.35	2.28	2.22	2.18	2.15	2.12	20
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.20	2.16	2.12	2.10	21
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.17	2.13	2.10	2.07	22
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.15	2.11	2.07	2.05	23
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.13	2.09	2.05	2.03	24
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.11	2.07	2.04	2.01	25

专业课辅导+真题资料+微信公众号：432应统书院

$\alpha = 0.05$ 续表

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	n_2 \ n_1
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.09	2.05	2.02	1.99	26
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.08	2.04	2.00	1.97	27
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.06	2.02	1.99	1.96	28
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.05	2.01	1.97	1.94	29
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.04	1.99	1.96	1.93	30
32	4.15	3.29	2.90	2.67	2.51	2.40	2.31	2.24	2.19	2.14	2.07	2.01	1.97	1.94	1.91	32
34	4.13	3.28	2.88	2.65	2.49	2.38	2.29	2.23	2.17	2.12	2.05	1.99	1.95	1.92	1.89	34
36	4.11	3.26	2.87	2.63	2.48	2.36	2.28	2.21	2.15	2.11	2.03	1.98	1.93	1.90	1.87	36
38	4.10	3.24	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.02	1.96	1.92	1.88	1.85	38
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.95	1.90	1.87	1.84	40
42	4.07	3.22	2.83	2.59	2.44	2.32	2.24	2.17	2.11	2.06	1.99	1.93	1.89	1.86	1.83	42
44	4.06	3.21	2.82	2.58	2.43	2.31	2.23	2.16	2.10	2.05	1.98	1.92	1.88	1.84	1.81	44
46	4.05	3.20	2.81	2.57	2.42	2.30	2.22	2.15	2.09	2.04	1.97	1.91	1.87	1.83	1.80	46
48	4.04	3.19	2.80	2.57	2.41	2.29	2.21	2.14	2.08	2.03	1.96	1.90	1.86	1.82	1.79	48
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.95	1.89	1.85	1.81	1.78	50
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.86	1.82	1.78	1.75	60
80	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06	2.00	1.95	1.88	1.82	1.77	1.73	1.70	80
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93	1.85	1.79	1.75	1.71	1.68	100
125	3.92	3.07	2.68	2.44	2.29	2.17	2.08	2.01	1.96	1.91	1.83	1.77	1.72	1.69	1.65	125
150	3.90	3.06	2.66	2.43	2.27	2.16	2.07	2.00	1.94	1.89	1.82	1.76	1.71	1.67	1.64	150
200	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93	1.88	1.80	1.74	1.69	1.66	1.62	200
300	3.87	3.03	2.63	2.40	2.24	2.13	2.04	1.97	1.91	1.86	1.78	1.72	1.68	1.64	1.61	300
500	3.86	3.01	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.77	1.71	1.66	1.62	1.59	500
1000	3.85	3.00	2.61	2.38	2.22	2.11	2.02	1.95	1.89	1.84	1.76	1.70	1.65	1.61	1.58	1000
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.69	1.64	1.60	1.57	∞

专业课辅导+真题资料+微信公众号：432应统书院

$\alpha=0.05$

续表

n_2 \ n_1	22	24	26	28	30	35	40	45	50	60	80	100	200	500	∞	n_2 \ n_1
1	249	249	249	250	250	251	251	251	252	252	252	253	254	254	254	1
2	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	2
3	8.65	8.64	8.63	8.62	8.62	8.60	8.59	8.59	8.58	8.57	8.56	8.55	8.54	8.53	8.53	3
4	5.79	5.77	5.76	5.75	5.75	5.73	5.72	5.71	5.70	5.69	5.67	5.66	5.65	5.64	5.63	4
5	4.54	4.53	4.52	4.50	4.50	4.48	4.46	4.45	4.44	4.43	4.41	4.41	4.39	4.37	4.37	5
6	3.86	3.84	3.83	3.82	3.81	3.79	3.77	3.76	3.75	3.74	3.72	3.71	3.69	3.68	3.67	6
7	3.43	3.41	3.40	3.39	3.38	3.36	3.34	3.33	3.32	3.30	3.29	3.27	3.25	3.24	3.23	7
8	3.13	3.12	3.10	3.09	3.08	3.06	3.04	3.03	3.02	3.01	2.99	2.97	2.95	2.94	2.93	8
9	2.92	2.90	2.89	2.87	2.86	2.84	2.83	2.81	2.80	2.79	2.77	2.76	2.73	2.72	2.71	9
10	2.75	2.74	2.72	2.71	2.70	2.68	2.66	2.65	2.64	2.62	2.60	2.59	2.56	2.55	2.54	10
11	2.63	2.61	2.59	2.58	2.57	2.55	2.53	2.52	2.51	2.49	2.47	2.46	2.43	2.42	2.40	11
12	2.52	2.51	2.49	2.48	2.47	2.44	2.43	2.41	2.40	2.38	2.36	2.35	2.32	2.31	2.30	12
13	2.44	2.42	2.41	2.39	2.38	2.36	2.34	2.33	2.31	2.30	2.27	2.26	2.23	2.22	2.21	13
14	2.37	2.35	2.33	2.32	2.31	2.28	2.27	2.25	2.24	2.22	2.20	2.19	2.16	2.14	2.13	14
15	2.31	2.29	2.27	2.26	2.25	2.22	2.20	2.19	2.18	2.16	2.14	2.12	2.10	2.08	2.07	15
16	2.25	2.24	2.22	2.21	2.19	2.17	2.15	2.14	2.12	2.11	2.08	2.07	2.04	2.02	2.01	16
17	2.21	2.19	2.17	2.16	2.15	2.12	2.10	2.09	2.08	2.06	2.03	2.02	1.99	1.97	1.96	17
18	2.17	2.15	2.13	2.12	2.11	2.08	2.06	2.05	2.04	2.02	1.99	1.98	1.95	1.93	1.92	18
19	2.13	2.11	2.10	2.08	2.07	2.05	2.03	2.01	2.00	1.98	1.96	1.94	1.91	1.89	1.88	19
20	2.10	2.08	2.07	2.05	2.04	2.01	1.99	1.98	1.97	1.95	1.92	1.91	1.88	1.86	1.84	20
21	2.07	2.05	2.04	2.02	2.01	1.98	1.96	1.95	1.94	1.92	1.89	1.88	1.84	1.82	1.81	21
22	2.05	2.03	2.01	2.00	1.98	1.96	1.94	1.92	1.91	1.89	1.86	1.85	1.82	1.80	1.78	22
23	2.02	2.00	1.99	1.97	1.96	1.93	1.91	1.90	1.88	1.86	1.84	1.82	1.79	1.77	1.76	23
24	2.00	1.98	1.97	1.95	1.94	1.91	1.89	1.88	1.86	1.84	1.82	1.80	1.77	1.75	1.73	24

专业课辅导+真题资料+微信公众号：432应统书院

$\alpha = 0.05$

续表

n_2 \ n_1	22	24	26	28	30	35	40	45	50	60	80	100	200	500	∞	n_2 \ n_1
25	1.98	1.96	1.95	1.93	1.92	1.89	1.87	1.86	1.84	1.82	1.80	1.78	1.75	1.73	1.71	25
26	1.97	1.95	1.93	1.91	1.90	1.87	1.85	1.84	1.82	1.80	1.78	1.76	1.73	1.71	1.69	26
27	1.95	1.93	1.91	1.90	1.88	1.86	1.84	1.82	1.81	1.79	1.76	1.74	1.71	1.69	1.67	27
28	1.93	1.91	1.90	1.88	1.87	1.84	1.82	1.80	1.79	1.77	1.74	1.73	1.69	1.67	1.65	28
29	1.92	1.90	1.88	1.87	1.85	1.83	1.81	1.79	1.77	1.75	1.73	1.71	1.67	1.65	1.64	29
30	1.91	1.89	1.87	1.85	1.84	1.81	1.79	1.77	1.76	1.74	1.71	1.70	1.66	1.64	1.62	30
32	1.88	1.86	1.85	1.83	1.82	1.79	1.77	1.75	1.74	1.71	1.69	1.67	1.63	1.61	1.59	32
34	1.86	1.84	1.82	1.80	1.80	1.77	1.75	1.73	1.71	1.69	1.66	1.65	1.61	1.59	1.57	34
36	1.85	1.82	1.81	1.79	1.78	1.75	1.73	1.71	1.69	1.67	1.64	1.62	1.59	1.56	1.55	36
38	1.83	1.81	1.79	1.77	1.76	1.73	1.71	1.69	1.68	1.65	1.62	1.61	1.57	1.54	1.53	38
40	1.81	1.79	1.77	1.76	1.74	1.72	1.69	1.67	1.66	1.64	1.61	1.59	1.55	1.53	1.51	40
42	1.80	1.78	1.76	1.74	1.73	1.70	1.68	1.66	1.65	1.62	1.59	1.57	1.53	1.51	1.49	42
44	1.79	1.77	1.75	1.73	1.72	1.69	1.67	1.65	1.63	1.61	1.58	1.56	1.52	1.49	1.48	44
46	1.78	1.76	1.74	1.72	1.71	1.68	1.65	1.64	1.62	1.60	1.57	1.55	1.51	1.48	1.46	46
48	1.77	1.75	1.73	1.71	1.70	1.67	1.64	1.62	1.61	1.59	1.56	1.54	1.49	1.47	1.45	48
50	1.76	1.74	1.72	1.70	1.69	1.66	1.63	1.61	1.60	1.58	1.54	1.52	1.48	1.46	1.44	50
60	1.72	1.70	1.68	1.66	1.65	1.62	1.59	1.57	1.56	1.53	1.50	1.48	1.44	1.41	1.39	60
80	1.68	1.65	1.63	1.62	1.60	1.57	1.54	1.52	1.51	1.48	1.45	1.43	1.38	1.35	1.32	80
100	1.65	1.63	1.61	1.59	1.57	1.54	1.52	1.49	1.48	1.45	1.41	1.39	1.34	1.31	1.28	100
125	1.63	1.60	1.58	1.57	1.55	1.52	1.49	1.47	1.45	1.42	1.39	1.36	1.31	1.27	1.25	125
150	1.61	1.59	1.57	1.55	1.53	1.50	1.48	1.45	1.44	1.41	1.37	1.34	1.29	1.25	1.22	150
200	1.60	1.57	1.55	1.53	1.52	1.48	1.46	1.43	1.41	1.39	1.35	1.32	1.26	1.22	1.19	200
300	1.58	1.55	1.53	1.51	1.50	1.46	1.43	1.41	1.39	1.36	1.32	1.30	1.23	1.19	1.15	300
500	1.56	1.54	1.52	1.50	1.48	1.45	1.42	1.40	1.38	1.34	1.30	1.28	1.21	1.16	1.11	500
1000	1.55	1.53	1.51	1.49	1.47	1.44	1.41	1.38	1.36	1.33	1.29	1.26	1.19	1.13	1.08	1000
∞	1.54	1.52	1.50	1.48	1.46	1.42	1.39	1.37	1.35	1.32	1.27	1.24	1.17	1.11	1.00	∞

专业课辅导+真题资料+微信公众号：432应统书院

$\alpha = 0.01$ 续表

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	n_2 \ n_1
1	405	500	540	563	576	586	593	598	602	606	611	614	617	619	621	1
2	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.4	2
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	26.9	26.8	26.8	26.7	3
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.2	14.2	14.1	14.0	4
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.89	9.77	9.68	9.61	9.55	5
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.60	7.52	7.45	7.40	6
7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.36	6.27	6.21	6.16	7
8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.56	5.48	5.41	5.36	8
9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	5.00	4.92	4.86	4.81	9
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.60	4.52	4.46	4.41	10
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.29	4.21	4.15	4.10	11
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.05	3.97	3.91	3.86	12
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.86	3.78	3.71	3.66	13
14	8.86	6.51	5.56	5.04	4.70	4.46	4.28	4.14	4.03	3.94	3.80	3.70	3.62	3.56	3.51	14
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.56	3.49	3.42	3.37	15
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.45	3.37	3.31	3.26	16
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.35	3.27	3.21	3.16	17
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.27	3.19	3.13	3.08	18
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.19	3.12	3.05	3.00	19
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.13	3.05	2.99	2.94	20
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.07	2.99	2.93	2.88	21
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	3.02	2.94	2.88	2.83	22
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.97	2.89	2.83	2.78	23
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.93	2.85	2.79	2.74	24
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13	2.99	2.89	2.81	2.75	2.70	25

专业课辅导+真题资料+微信公众号：432应统书院

$\alpha = 0.01$ 续表

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	n_2 \ n_1
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.86	2.78	2.72	2.66	26
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.82	2.75	2.68	2.63	27
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.79	2.72	2.65	2.60	28
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.77	2.69	2.63	2.57	29
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.74	2.66	2.60	2.55	30
32	7.50	5.34	4.46	3.97	3.65	3.43	3.26	3.13	3.02	2.93	2.80	2.70	2.62	2.55	2.50	32
34	7.44	5.29	4.42	3.93	3.61	3.39	3.22	3.09	2.98	2.89	2.76	2.66	2.58	2.51	2.46	34
36	7.40	5.25	4.38	3.89	3.57	3.35	3.18	3.05	2.95	2.86	2.72	2.62	2.54	2.48	2.43	36
38	7.35	5.21	4.34	3.86	3.54	3.32	3.15	3.02	2.92	2.83	2.69	2.59	2.51	2.45	2.40	38
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.56	2.48	2.42	2.37	40
42	7.28	5.15	4.29	3.80	3.49	3.27	3.10	2.97	2.86	2.78	2.64	2.54	2.46	2.40	2.34	42
44	7.25	5.12	4.26	3.78	3.47	3.24	3.08	2.95	2.84	2.75	2.62	2.52	2.44	2.37	2.32	44
46	7.22	5.10	4.24	3.76	3.44	3.22	3.06	2.93	2.82	2.73	2.60	2.50	2.42	2.35	2.30	46
48	7.20	5.08	4.22	3.74	3.43	3.20	3.04	2.91	2.80	2.72	2.58	2.48	2.40	2.33	2.28	48
50	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.79	2.70	2.56	2.46	2.38	2.32	2.27	50
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.39	2.31	2.25	2.20	60
80	6.96	4.88	4.04	3.56	3.26	3.04	2.87	2.74	2.64	2.55	2.42	2.31	2.23	2.17	2.12	80
100	6.90	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.59	2.50	2.37	2.26	2.19	2.12	2.07	100
125	6.84	4.78	3.94	3.47	3.17	2.95	2.79	2.66	2.55	2.47	2.33	2.23	2.15	2.08	2.03	125
150	6.81	4.75	3.92	3.45	3.14	2.92	2.76	2.63	2.53	2.44	2.31	2.20	2.12	2.06	2.00	150
200	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50	2.41	2.27	2.17	2.09	2.02	1.97	200
300	6.72	4.68	3.85	3.38	3.08	2.86	2.70	2.57	2.47	2.38	2.24	2.14	2.06	1.99	1.94	300
500	6.69	4.65	3.82	3.36	3.05	2.84	2.68	2.55	2.44	2.36	2.22	2.12	2.04	1.97	1.92	500
1000	6.66	4.63	3.80	3.34	3.04	2.82	2.66	2.53	2.43	2.34	2.20	2.10	2.02	1.95	1.90	1000
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.08	2.00	1.93	1.88	∞

专业课辅导+真题资料+微信公众号：432应统书院

$\alpha = 0.01$ 续表

n_2 \ n_1	22	24	26	28	30	35	40	45	50	60	80	100	200	500	∞	n_2 \ n_1
1	622	623	624	625	626	628	629	630	630	631	633	633	635	636	637	1
2	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	2
3	26.6	26.6	26.6	26.5	26.5	26.5	26.4	26.4	26.4	26.3	26.3	26.2	26.2	26.1	26.1	3
4	14.0	13.9	13.9	13.9	13.8	13.8	13.7	13.7	13.7	13.7	13.6	13.6	13.5	13.5	13.5	4
5	9.51	9.47	9.43	9.40	9.38	9.33	9.29	9.26	9.24	9.20	9.16	9.13	9.08	9.04	9.02	5
6	7.35	7.31	7.28	7.25	7.23	7.18	7.14	7.11	7.09	7.06	7.01	6.99	6.93	6.90	6.88	6
7	6.11	6.07	6.04	6.02	5.99	5.94	5.91	5.88	5.86	5.82	5.78	5.75	5.70	5.67	5.65	7
8	5.32	5.28	5.25	5.22	5.20	5.15	5.12	5.00	5.07	5.03	4.99	4.96	4.91	4.88	4.86	8
9	4.77	4.73	4.70	4.67	4.65	4.60	4.57	4.54	4.52	4.48	4.44	4.42	4.36	4.33	4.31	9
10	4.36	4.33	4.30	4.27	4.25	4.20	4.17	4.14	4.12	4.08	4.04	4.01	3.96	3.93	3.91	10
11	4.06	4.02	3.99	3.96	3.94	3.89	3.86	3.83	3.81	3.78	3.73	3.71	3.66	3.62	3.60	11
12	3.82	3.78	3.75	3.72	3.70	3.65	3.62	3.59	3.57	3.54	3.49	3.47	3.41	3.38	3.36	12
13	3.62	3.59	3.56	3.53	3.51	3.46	3.43	3.40	3.38	3.34	3.30	3.27	3.22	3.19	3.17	13
14	3.46	3.43	3.40	3.37	3.35	2.30	3.27	3.24	3.22	3.18	3.14	3.11	3.06	3.03	3.00	14
15	3.33	3.29	3.26	3.24	3.21	3.17	3.13	3.10	3.08	3.05	3.00	2.98	2.92	2.89	2.87	15
16	3.22	3.18	3.15	3.12	3.10	3.05	3.02	2.99	2.97	2.93	2.89	2.86	2.81	2.78	2.75	16
17	3.12	3.08	3.05	3.03	3.00	2.96	2.92	2.89	2.87	2.83	2.79	2.76	2.71	2.68	2.65	17
18	3.03	3.00	2.97	2.94	2.92	2.87	2.84	2.81	2.78	2.75	2.70	2.68	2.62	2.59	2.57	18
19	2.96	2.92	2.89	2.87	2.84	2.80	2.76	2.73	2.71	2.67	2.63	2.60	2.55	2.51	2.49	19
20	2.90	2.86	2.83	2.80	2.78	2.73	2.69	2.67	2.64	2.61	2.56	2.54	2.48	2.44	2.42	20
21	2.84	2.80	2.77	2.74	2.72	2.67	2.64	2.61	2.58	2.55	2.50	2.48	2.42	2.38	2.36	21
22	2.78	2.75	2.72	2.69	2.67	2.62	2.58	2.55	2.53	2.50	2.45	2.42	2.36	2.33	2.31	22
23	2.74	2.70	2.67	2.64	2.62	2.57	2.54	2.51	2.48	2.45	2.40	2.37	2.32	2.28	2.26	23
24	2.70	2.66	2.63	2.60	2.58	2.53	2.49	2.46	2.44	2.40	2.36	2.33	2.27	2.24	2.21	24

专业课辅导+真题资料+微信公众号：432应统书院

$\alpha=0.01$

续表

n_2 \ n_1	22	24	26	28	30	35	40	45	50	60	80	100	200	500	∞	n_2 \ n_1
25	2.66	2.62	2.59	2.56	2.54	2.49	2.45	2.42	2.40	2.36	2.32	2.29	2.23	2.19	2.17	25
26	2.62	2.58	2.55	2.53	2.50	2.45	2.42	2.39	2.36	2.33	2.28	2.25	2.19	2.16	2.13	26
27	2.59	2.55	2.52	2.49	2.47	2.42	2.38	2.35	2.33	2.29	2.25	2.22	2.16	2.12	2.10	27
28	2.56	2.52	2.49	2.46	2.44	2.39	2.35	2.32	2.30	2.26	2.22	2.19	2.13	2.09	2.06	28
29	2.53	2.49	2.46	2.44	2.41	2.36	2.33	2.30	2.27	2.23	2.19	2.16	2.10	2.06	2.03	29
30	2.51	2.47	2.44	2.41	2.39	2.34	2.30	2.27	2.25	2.21	2.16	2.13	2.07	2.03	2.01	30
32	2.46	2.42	2.39	2.36	2.34	2.29	2.25	2.22	2.20	2.16	2.11	2.08	2.02	1.98	1.96	32
34	2.42	2.38	2.35	2.32	2.30	2.25	2.21	2.18	2.16	2.12	2.07	2.04	1.98	1.94	1.91	34
36	2.38	2.35	2.32	2.29	2.26	2.21	2.17	2.14	2.12	2.08	2.03	2.00	1.94	1.90	1.87	36
38	2.35	2.32	2.28	2.26	2.23	2.18	2.14	2.11	2.09	2.05	2.00	1.97	1.90	1.86	1.84	38
40	2.33	2.29	2.26	2.23	2.20	2.15	2.11	2.08	2.06	2.02	1.97	1.94	1.87	1.83	1.80	40
42	2.30	2.26	2.23	2.20	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.91	1.85	1.80	1.78	42
44	2.28	2.24	2.21	2.18	2.15	2.10	2.06	2.03	2.01	1.97	1.92	1.89	1.82	1.78	1.75	44
46	2.26	2.22	2.19	2.16	2.13	2.08	2.04	2.01	1.99	1.95	1.90	1.86	1.80	1.75	1.73	46
48	2.24	2.20	2.17	2.14	2.12	2.06	2.02	1.99	1.97	1.93	1.88	1.84	1.78	1.73	1.70	48
50	2.22	2.18	2.15	2.12	2.10	2.05	2.01	1.97	1.95	1.91	1.86	1.82	1.76	1.71	1.68	50
60	2.15	2.12	2.08	2.05	2.03	1.98	1.94	1.90	1.88	1.84	1.78	1.75	1.68	1.63	1.60	60
80	2.07	2.03	2.00	1.97	1.94	1.89	1.85	1.81	1.79	1.75	1.69	1.66	1.58	1.53	1.49	80
100	2.02	1.98	1.94	1.92	1.89	1.84	1.80	1.76	1.73	1.69	1.63	1.60	1.52	1.47	1.43	100
125	1.98	1.94	1.91	1.88	1.85	1.80	1.76	1.72	1.69	1.65	1.59	1.55	1.47	1.41	1.37	125
150	1.96	1.92	1.88	1.85	1.83	1.77	1.73	1.69	1.66	1.62	1.56	1.52	1.43	1.38	1.33	150
200	1.93	1.89	1.85	1.82	1.79	1.74	1.69	1.66	1.63	1.58	1.52	1.48	1.39	1.33	1.28	200
300	1.89	1.85	1.82	1.79	1.76	1.71	1.66	1.62	1.59	1.55	1.48	1.44	1.35	1.28	1.22	300
500	1.87	1.83	1.79	1.76	1.74	1.68	1.63	1.60	1.56	1.52	1.45	1.41	1.31	1.23	1.16	500
1000	1.85	1.81	1.77	1.74	1.72	1.66	1.61	1.57	1.54	1.50	1.43	1.38	1.28	1.19	1.11	1000
∞	1.83	1.79	1.76	1.72	1.70	1.64	1.59	1.55	1.52	1.47	1.40	1.36	1.25	1.15	1.00	∞

专业课辅导+真题资料+微信公众号：432应统书院

专业课辅导+真题资料+微信公众号：432应统书院

表 4　检验相关系数 $\rho=0$ 的临界值(r_α)表

$$P(|r|>r_\alpha)=\alpha$$

n \ α	0.10	0.05	0.02	0.01	0.001	α \ n
1	0.98769	0.99692	0.999507	0.999877	0.9999988	1
2	0.90000	0.95000	0.98000	0.99000	0.99900	2
3	0.8054	0.8783	0.93433	0.95873	0.99116	3
4	0.7293	0.8114	0.8822	0.91720	0.97406	4
5	0.6694	0.7545	0.8329	0.8745	0.95074	5
6	0.6215	0.7067	0.7887	0.8343	0.92493	6
7	0.5822	0.6664	0.7498	0.7977	0.8982	7
8	0.5494	0.6319	0.7155	0.7646	0.8721	8
9	0.5214	0.6021	0.6851	0.7348	0.8471	9
10	0.4973	0.5760	0.6581	0.7079	0.8233	10
11	0.4762	0.5529	0.6339	0.6835	0.8010	11
12	0.4575	0.5324	0.6120	0.6614	0.7800	12
13	0.4409	0.5139	0.5923	0.6411	0.7603	13
14	0.4259	0.4973	0.5742	0.6226	0.7420	14
15	0.4124	0.4821	0.5577	0.6055	0.7246	15
16	0.4000	0.4683	0.5425	0.5897	0.7084	16
17	0.3887	0.4555	0.5285	0.5751	0.6932	17
18	0.3783	0.4438	0.5155	0.5614	0.6787	18
19	0.3687	0.4329	0.5034	0.5487	0.6652	19
20	0.3598	0.4227	0.4921	0.5368	0.6524	20
25	0.3233	0.3809	0.4451	0.4869	0.5974	25
30	0.2960	0.3494	0.4093	0.4487	0.5541	30
35	0.2746	0.3246	0.3810	0.4182	0.5189	35
40	0.2573	0.3044	0.3578	0.3932	0.4896	40
45	0.2428	0.2875	0.3384	0.3721	0.4648	45
50	0.2306	0.2732	0.3218	0.3541	0.4433	50
60	0.2108	0.2500	0.2948	0.3248	0.4078	60
70	0.1954	0.2319	0.2737	0.3017	0.3799	70
80	0.1829	0.2172	0.2565	0.2830	0.3568	80
90	0.1726	0.2050	0.2422	0.2673	0.3375	90
100	0.1638	0.1946	0.2301	0.2540	0.3211	100

表 5　随 机 数 表

03	47	43	73	86	36	96	47	36	61	46	98	63	71	62	33	26	16	80	45	60	11	14	10	95
97	74	24	67	62	42	81	14	57	20	42	53	32	37	32	27	07	36	07	51	24	51	79	89	73
16	76	62	27	66	56	50	26	71	07	32	90	79	78	53	13	55	38	58	59	88	97	54	14	10
12	56	85	99	26	96	96	68	27	31	05	03	72	93	15	57	12	10	14	21	88	26	49	81	76
55	59	56	35	64	38	54	82	46	22	31	62	43	09	90	06	18	44	32	53	23	83	01	30	30
16	22	77	94	39	49	54	43	54	82	17	37	93	23	78	87	35	20	96	43	84	26	34	91	64
84	42	17	53	31	57	24	55	06	88	77	04	74	47	67	21	76	33	50	25	83	92	12	06	76
63	01	63	78	59	16	95	55	67	19	98	10	50	71	75	12	86	73	58	07	44	39	52	38	79
33	21	12	34	29	78	64	56	07	82	52	42	07	44	38	15	51	00	13	42	99	66	02	79	54
57	60	86	32	44	09	47	27	96	54	49	17	46	09	62	90	52	84	77	27	08	02	73	43	28
18	18	07	92	45	44	17	16	58	09	79	83	86	19	62	06	76	50	03	10	55	23	64	05	05
26	62	38	97	75	84	16	07	44	99	83	11	46	32	24	20	14	85	88	45	10	93	72	88	71
23	42	40	64	74	82	97	77	77	81	07	45	32	14	08	32	98	94	07	72	93	85	79	10	75
52	36	28	19	95	50	92	26	11	97	00	56	76	31	38	80	22	02	53	53	86	60	42	04	53
37	85	94	35	12	83	39	50	08	30	42	34	07	96	88	54	42	06	87	98	35	85	29	48	39
70	29	17	12	13	40	33	20	38	26	13	89	51	03	74	17	76	37	13	04	07	74	21	19	30
56	62	18	37	35	96	83	50	87	75	97	12	25	93	47	70	33	24	03	54	97	77	46	44	80
99	49	57	22	77	88	42	95	45	72	16	64	36	16	00	04	43	18	66	79	94	77	24	21	90
16	08	15	04	72	33	27	14	34	09	45	59	34	68	49	12	72	07	34	45	99	27	72	95	14
31	16	93	32	43	50	27	89	87	19	20	15	37	00	49	52	85	66	60	44	38	68	88	11	80
68	34	30	13	70	55	74	30	77	40	44	22	78	84	26	04	33	46	09	52	68	07	97	06	57
74	57	25	65	76	59	29	97	68	60	71	91	38	67	54	13	58	18	24	76	15	54	55	95	52
27	42	37	86	53	48	55	90	65	72	96	57	69	36	10	96	46	92	42	45	97	60	49	04	91
00	39	68	29	61	66	37	32	20	30	77	84	57	03	29	10	45	65	04	26	11	04	96	67	24
29	94	98	94	24	68	49	69	10	82	53	75	91	93	30	34	25	20	57	27	40	48	73	51	92

专业课辅导+真题资料+微信公众号：432应统书院

续表

16	90	82	66	59	83	62	64	11	12	67	19	00	71	74	60	47	21	19	68	02	02	37	03	31
11	27	94	75	06	06	09	19	74	66	02	94	37	34	02	76	70	90	30	86	38	45	94	30	38
35	24	10	16	20	33	32	51	26	38	79	78	45	04	91	16	92	53	56	16	02	75	50	95	98
38	23	16	86	38	42	38	97	01	50	87	75	66	81	41	40	01	74	91	62	48	51	84	08	32
31	96	25	91	47	96	44	33	49	13	34	86	82	53	91	00	52	43	48	85	27	55	26	89	62
66	67	40	67	14	64	05	71	95	86	11	05	65	09	68	76	83	20	37	90	57	16	00	11	66
14	90	84	45	11	75	73	88	05	90	52	27	41	14	86	22	98	12	22	08	07	52	74	95	80
68	05	51	18	00	33	96	02	75	19	07	60	62	93	55	59	33	82	43	90	49	37	38	44	59
20	46	78	73	90	97	51	40	14	02	04	02	33	31	08	39	54	16	49	36	47	95	93	13	30
64	19	58	97	79	15	06	15	93	20	01	90	10	75	06	40	78	78	89	62	02	67	74	17	33
05	26	93	70	60	22	35	85	15	13	92	03	51	59	77	59	56	78	06	83	52	91	05	70	74
07	97	10	88	23	09	98	42	99	64	61	71	62	99	15	06	51	29	16	93	58	05	77	09	51
68	71	86	85	85	54	87	66	47	54	73	32	08	11	12	44	95	92	63	16	29	56	24	29	48
26	99	61	65	53	58	37	78	80	70	42	10	50	67	42	32	17	55	85	74	94	44	67	16	94
14	65	52	68	75	87	59	36	22	41	26	78	63	06	55	13	08	27	01	50	15	29	39	39	43
17	53	77	58	71	71	41	61	50	72	12	41	94	96	26	44	95	27	36	99	02	96	74	30	83
90	26	59	21	19	23	52	23	33	12	96	93	02	18	39	07	02	18	36	07	25	99	32	70	23
41	23	52	55	99	31	04	49	69	96	10	47	48	45	88	13	41	43	89	20	97	17	14	49	17
60	20	50	81	69	31	99	73	68	68	35	81	33	03	76	24	30	12	48	60	18	99	10	72	34
91	25	38	05	90	94	58	28	41	36	45	37	59	03	09	90	35	57	29	12	82	62	54	65	60
34	50	57	74	37	98	30	33	00	91	09	77	93	19	82	74	94	80	04	04	45	07	31	66	49
85	22	04	39	43	73	81	53	94	79	33	62	46	86	28	08	31	54	46	31	53	94	13	38	47
09	79	13	77	48	73	82	97	22	21	05	03	27	24	83	72	89	44	05	60	35	80	39	94	83
88	75	80	18	14	22	95	75	42	49	39	32	82	22	49	02	48	07	70	37	16	04	61	67	87
90	96	23	70	00	39	00	03	06	90	55	85	78	38	36	94	37	30	69	32	90	89	00	76	33

专业课辅导+真题资料+微信公众号：432应统书院

续表

53	74	23	99	67	61	32	28	69	84	94	62	67	86	24	98	33	41	19	95	47	53	53	38	09
63	38	06	86	54	99	00	65	26	94	02	82	90	23	07	79	62	67	80	60	75	91	12	81	19
35	30	58	21	46	06	72	17	10	94	25	21	31	75	96	49	28	24	00	49	55	65	79	78	07
63	43	36	82	69	65	51	18	37	88	61	38	44	12	45	32	92	85	88	65	54	34	81	85	35
98	25	37	55	26	01	91	82	81	46	74	71	12	94	97	24	02	71	37	07	03	92	18	66	75
02	63	21	17	69	71	50	80	89	56	38	15	70	11	48	43	40	45	86	98	00	83	26	91	03
64	55	22	21	82	48	22	28	06	00	61	54	13	43	91	82	78	12	23	29	06	66	24	12	27
85	07	26	13	39	01	10	07	82	04	59	63	69	36	03	69	11	15	83	80	13	29	54	19	28
58	54	16	24	15	51	54	44	82	00	62	61	65	04	69	38	18	65	18	97	85	72	13	49	21
34	85	27	84	87	61	48	64	56	26	90	18	48	13	26	37	70	15	42	57	65	65	80	39	07
03	92	18	27	46	57	99	16	96	56	30	33	72	85	22	84	64	38	56	98	99	01	30	98	64
62	93	30	27	59	37	75	41	66	48	86	97	80	61	45	23	53	04	01	63	45	76	08	64	27
08	45	93	15	22	60	21	75	46	91	98	77	27	85	42	28	88	61	08	84	69	62	03	42	73
07	08	55	18	40	45	44	75	13	90	24	94	96	61	02	57	55	66	83	15	73	42	37	11	61
01	85	89	95	66	51	10	19	34	88	15	84	97	19	75	12	76	39	43	78	64	63	91	08	25
72	84	71	14	35	19	11	58	49	26	50	11	17	17	76	86	31	57	20	18	95	60	78	46	75
88	78	28	16	84	13	52	53	94	53	75	45	69	30	96	73	89	65	70	31	99	17	43	48	76
45	17	75	65	57	28	40	19	72	12	25	12	74	75	67	60	40	60	81	19	24	62	01	61	16
96	76	28	12	54	22	01	11	94	25	71	96	16	16	88	68	64	36	74	45	19	59	50	88	92
43	31	67	72	30	24	02	94	08	63	38	32	36	66	02	69	36	38	25	39	48	03	45	15	22
50	44	66	44	21	66	06	58	05	62	68	15	54	35	02	42	35	48	96	32	14	52	41	52	48
22	66	22	15	86	26	63	75	41	99	58	42	36	72	24	58	37	52	18	51	03	37	18	39	11
96	24	40	14	51	23	22	30	88	57	95	67	47	29	83	94	69	40	06	07	18	16	36	78	86
31	73	91	61	19	60	20	72	93	48	98	57	07	23	69	65	95	39	69	58	56	80	30	19	44
78	60	73	99	84	43	89	94	36	45	56	69	47	07	41	90	22	91	07	12	78	35	34	08	72

专业课辅导+真题资料+微信公众号：432应统书院

续表

84	37	90	61	56	70	10	23	98	05	85	11	34	76	60	76	48	45	34	60	01	64	18	39	96
36	67	10	08	23	98	93	35	08	86	99	29	76	29	81	33	34	91	58	98	63	14	52	32	52
07	28	59	07	48	89	64	58	89	75	83	85	62	27	89	80	14	78	56	27	86	63	59	80	02
10	15	83	87	60	79	24	31	66	56	21	48	24	06	93	91	98	94	05	49	01	47	59	38	00
55	19	68	97	65	03	73	52	16	56	00	53	55	90	27	33	42	29	38	87	22	13	88	83	34
53	81	29	13	39	35	01	20	71	34	62	33	74	82	14	53	73	19	09	03	56	54	29	56	93
51	86	32	68	92	33	98	74	66	99	40	14	71	94	58	45	94	19	38	81	14	44	99	81	07
35	91	70	29	13	80	03	54	07	27	96	94	78	32	66	50	95	52	74	33	13	80	55	62	54
37	71	67	95	13	20	02	44	95	94	64	85	04	05	72	01	32	90	76	14	53	89	74	60	41
93	66	13	83	27	92	79	64	64	72	28	54	96	53	84	48	14	52	98	94	56	07	93	89	30
02	96	08	45	65	13	05	00	41	84	93	07	54	72	59	21	45	57	09	77	19	48	56	27	44
49	83	43	48	35	82	88	33	69	96	72	36	04	19	76	47	45	15	18	60	82	11	08	95	97
84	60	71	62	46	40	80	81	30	37	34	39	23	05	38	25	15	35	71	30	88	12	57	21	77
18	17	30	88	71	44	91	14	88	47	89	23	30	63	15	56	34	20	47	89	99	82	93	24	98
79	69	10	61	78	71	32	76	95	62	87	00	22	58	40	92	54	01	75	25	43	11	71	99	31
75	93	36	57	83	56	20	14	82	11	74	21	97	90	65	96	42	68	63	86	74	54	13	26	94
38	30	92	29	03	06	28	81	39	38	62	25	06	84	63	61	29	08	93	67	04	32	92	08	09
51	29	50	10	34	31	57	75	95	80	51	97	02	74	77	76	15	48	49	44	18	55	63	77	09
21	31	38	86	24	37	79	81	53	74	73	24	16	10	33	52	83	90	94	76	70	47	14	54	36
29	01	23	87	88	58	02	39	37	67	42	10	14	20	92	16	55	23	42	45	54	96	09	11	06
95	33	95	22	00	18	74	72	00	18	38	79	58	69	32	81	76	80	26	92	82	80	84	25	39
90	84	60	79	80	24	36	59	87	38	82	07	53	89	35	96	35	23	79	18	05	98	90	07	35
46	40	62	98	82	54	97	20	56	95	15	74	80	08	32	16	46	70	50	80	67	72	16	42	79
20	31	89	03	43	38	46	82	68	72	32	14	82	99	70	80	60	47	18	97	63	49	30	21	30
71	59	73	05	50	08	22	23	71	77	91	01	93	20	49	82	96	59	26	94	66	39	67	98	60

专业课辅导+真题资料+微信公众号：432应统书院

专业课辅导+真题资料+微信公众号：432应统书院

译名对照表[1]

Brown	布朗
Cochran	柯赫伦
David	达威
De Morgan	德·摩根
Dirichlet	狄利克雷
Dugue	杜琪
Fisher	费希尔
Geiger	盖革
Harwicx	赫维斯
Lévy	列维
Monte-Carlo	蒙特卡罗
Neymann	奈曼
Rao	拉奥
Rutherford	卢瑟福
Scheffe	谢菲
Wald	瓦尔德
Wiener	维纳
Гливенко	格列汶科
Von Neumann	冯·诺伊曼

① 凡在上册对照表中出现过的外国人名，下册都未列入.

专业课辅导+真题资料+微信公众号：432应统书院

参考文献

[1] 王梓坤. 概率论基础及其应用. 北京:科学出版社,1976

[2] Fisz M. 概率论及数理统计. 王福保译. 上海:上海科学技术出版社,1962

[3] Cramer H. 统计数学方法. 魏宗舒等译. 上海:上海科学技术出版社,1966

[4] Mann H B. 试验的分析与设计. 张里千等译. 北京:科学出版社,1963

[5] 陈希孺. 数理统计引论. 北京:科学出版社,1981

[6] 陈家鼎等. 数理统计讲义. 北京:高等教育出版社,1993

[7] 茆诗松. 贝叶斯统计学. 北京:中国统计出版社,1997

[8] 杨维权,邓集贤. 概率统计教学参考书. 北京:高等教育出版社,1996

[9] 方开泰. 均匀设计与均匀设计表. 北京:科学出版社,1994

[10] 王梓坤. 随机过程. 第二版. 北京:科学出版社,1978

[11] 中山大学《测度与概率基础编写组》. 测度与概率基础. 广州:广东科学技术出版社,1981

[12] 汪嘉冈. 现代概率论基础. 上海:复旦大学出版社,1988

[13] 斯奈德 D L. 随机点过程. 梁之舜,邓永录译. 北京:人民教育出版社,1980

[14] Perzen E. 随机过程. 邓永录,杨振明译. 北京:高等教育出版社,1987

[15] 徐光辉. 随机服务系统. 北京:科学出版社,1980

[16] 侯振挺,郭青峰. 齐次可列马尔科夫过程. 北京:科学出版社,1978

[17] 邓集贤,许刘俊. 随机过程. 北京:高等教育出版社,1992

[18] 邓集贤,杨维权,许刘俊. 经济预测与决策的数学方法. 广州:中山大学出版社,1986

[19] 基赫曼 И И,斯科罗霍德 A B. 随机过程(Ⅰ). 邓永录,邓集贤,石北源译. 北京:科学出版社,1985

[20] 基赫曼 И И,斯科罗霍德 A B. 随机过程(Ⅲ). 石北源等译. 北京:科

专业课辅导+真题资料+微信公众号：432应统书院

学出版社,1992

[21] 陈家鑫.应用概率论.北京:科学出版社,1992

[22] 安鸿志,陈兆国,杜金观,潘一民.时间序列分析与应用.北京:科学出版社,1983

[23] 胡迪鹤,甘师信.近代鞅论.武汉大学学术丛书,1992

[24] 邓永录.随机模型及其应用.北京:高等教育出版社,1994

[25] Soon T T. Random Differential Equation in Science and Engineering. 1973

[26] [苏]楚耶夫 Ю B 等.军事技术运筹学基础.冷拓等译.北京:国防工业出版社,1976