数 理 统 计 讲 义

(Textbook of Mathematical Statistics)

王兆军　编著

南 开 大 学

数 学 科 学 学 院 统 计 系

天　　津

内容提要

数理统计是一门主要研究如何有效地收集、整理和分析受随机影响的数据，并对所考虑的问题作出科学推断的一门学科，它具有很强的应用性，并且在许多学科中都得到了广泛的应用，且取得了良好的社会和经济效益. 本书主要讲述数理统计的一些基本概念与方法，如几个常用的抽样分布，矩估计、最小方差无偏估计、极大似然估计、最小二乘估计等点估计方法和基于枢轴量法的区间估计，单样本与两样本的显著性检验、优势检验、似然比检验、序贯概率比检验及一些拟合优度检验方法. 另外，本讲义也简介了某些统计模拟方法初步.

本书可作为统计学专业本科教材，也可供其它专业和工程技术人员和应用工作者参考.

Tel: 22+23504709

Fax: 22+23506423

Email: zjwangnk@yahoo.com

Http: www.math.nankai.edu.cn/~zjwang

作 者 简 介

王兆军 南开大学教授、博士生导师. 现任南开大学数学科学学院统计学系系主任；中国现场统计研究会理事；中国概率统计学会理事；中国现场统计研究会生存分析分会理事；天津数学会秘书长；《数理统计与管理》杂志副主编. 1983至1987年于南开大学数学科学学院读本科，并获理学学士学位；1987至1990年于华东师范大学统计系读研究生，并获理学硕士学位；1990年7月到南开大学数学科学学院工作至今，期间于1993至1995年于南开大学数学科学学院读在职博士，并获理学博士学位；1996至1997年在南开大学经济站作博士后.

作 者 序

数理统计是一门应用性非常强的学科，她的历史已有三百多年，即使从K. Pearson (1857–1936)和R. A. Fisher (1890. 2. 17–1962. 7. 29)的工作算起，数理统计的发展也已有一百多年的历史，并且取得了良好的社会和经济效益.

按“不列颠百科全书”上的说法，统计学(Statistics)是“收集和分析数据的科学与艺术”，并且没有标出Mathematical Statistics一词，这是因为在"Statistics"一词的使用上，我们与西方不同. 我们所说的数理统计(Mathematical Statistics)即为西方所说的"Statistics", 其原因是以示与我们这里存在的被视为一门社会科学的统计学加以区别. 在西方，也有Mathematical Statistics的提法，那是特指统计方法的概率–数学理论基础那一部分，可以视为是一种纯粹数学.

我国著名的统计学家陈希孺院士说:“数理统计学是数学的一个分支，它是一门用有效的方法收集和分析带有随机影响的数据的学科，且其目的是解决特定的问题”. 另外，陈希孺院士也认为，数理统计学只是从数量关系的层面上来分析问题，完全不触及问题的专业内涵，且其结论不可混同于因果关系. 因此数理统计方法具有如下两个特点：方法的中立性和工具性以及它不肯定因果关系.

我国另一著名的统计学家茆诗松教授在华东师范大学出版社出版的一套“数理统计丛书”总序中说:“数理统计是一门应用性很强的学科. 它是研究如何有效地收集、整理和分析受随机影响的数据，并对所考虑的问题作出推断或预测，直至为采取决策和行动提供依据和建议的一门学科”.

可见，数理统计是一门研究如何有效的收集和分析受到随机影响数据的学科. 经过过去的研究和发展，数理统计已深入到了多个学科中，可以说，凡是一个实际问题有大量或少量数据，我们都可以利用数理统计方法去分析它、解决它.

随着数理统计学的发展和完善，其研究内容已非常丰富，且形成了多个学科分支，如抽样调查、试验设计、回归分析、多元统计分析、时间序列分析、非参数统计、Bayes方法等等.

本讲义的目的是用较严格的语言，对数理统计学中的一些基础知识作一较为详细的介绍. 在第 1 章中，我们介绍了样本、总体、样本分布、统计量、充分统计量及几个常用抽样分布，如χ^2分布、t分布、F分布等概念和基本性质. 在第 2 章中，我们给出了几个点估计方法，如矩估计、极大似然估计和UMVUE，另外，也给出了Bayes估计的定义. 基于枢轴量法的区间估计和Fisher信任推断方法放在了第 3 章. 在第 4 章中，我们介绍了关于单样本及两样本正态总体的Fisher显著性检验和单参数指数族的显著性检验和似然比检验问题. 基于Neymann-Pearson引理的优势检验和与此序贯概率比检验将在第 5 章讲述. 我们把基于K. Pearson提出的χ^2拟合优度检验和Kolmogorov检验及某些正态性检验方法放到了第 6 章.

在本讲义的编写过程中，作者选择了一些常用的典型统计方法来叙述前人积累起来的统计思想和概念，并注意统计的应用性. 本讲义是为南开大学数学科学学院统计专业本科生的“数理统计”课程编写的，且在南开大学数学科学学院讲授多年. 期间，许多学生对本讲义提出了许多宝贵意见. 另外，本讲义所需课时大约76个学时.

另外，在编写本讲义的过程中，作者参阅了多位作者的书籍，如陈希孺院士的《数理统计

发展简史》、Lahmann教授的《点估计理论》、茆诗松、王静龙及濮晓龙教授等的《数理统计》和《高等数理统计》、复旦编写的《概率论》第二册等. 并且，这些书籍为本讲义的编写提供了许多宝贵的材料，作者在此表示衷心的感谢.

由于作者无论是中文水平还是统计专业知识都很有限，编写之中难免会有不妥之处，欢迎广大读者批评指正.

作者

2006年6月于南开园

目　录

第1章　基本概念

在本课程中用到的一些基本概念将在本章给出，如样本、总体、抽样分布、参数等.

§1.1　引言

在本节，我们将通过几个例子来说明数理统计的几个研究方向，之后再介绍数理统计的几个特点.

§1.1.1　几个例子

在讲解基本概念之前，我们先看几个例子.

例 1.1.1 (抽样抽查(Sampling Survey))　随着高校的扩招和自主招生的推广，每所院校及院系都越来越重视自己的生源情况，也都想出了各种各样的方法以吸引各地优秀的中学生报考自己的学校. 导致这种现象的根本原因就是各个学校都认识到学生质量对学校声誉的重要性. 如果仅从考试成绩来看，大学成绩是否真的与中学成绩存在着很大的相关性？要想回答此问题，我们有必要考虑如下问题:

- 选取哪几所大学进行研究？在选定的大学中选取多少名什么样的大学生？
- 用什么成绩来反映大学和中学的成绩？
- 有了数据后如何进行分析？
- 这种情况对于重点大学与一般大学是否一样？

上面的几个问题中既有如何选取学校和学生的问题，也有如何进行数据分析的问题. 有关学校及学生的选取问题，其中有关抽样内容我们将在抽样调查课程中学习.

例 1.1.2 (估计(Estimation))　如何综合评价一个大学生的学习成绩？

这是一个估计的问题，我们将在本门课程中加以介绍.

例 1.1.3 (假设检验(Hypothesis Test))　在许多体育比赛中，裁判均是通过掷硬币的方法让一方先进行选择，大家之所以接受这种方法是由于我们相信硬币是均匀的，即掷一次硬币出现正面与反面的概率是相同的. 现要验证某硬币的均匀性，我们将这枚硬币投掷 495 次后，正反面分别出现 220 次和 275 次，请问: 我们能说这枚硬币均匀吗？

这个问题的解决方法，就是本课程将要介绍的假设检验中的内容.

例 1.1.4 (试验设计(Design of Experiment))　农民在种地之前，会选择适当的种子. 假如一个农业试验站共有五种小麦可供选择，请问此试验站应向农民提供哪个种子以备来年的播种？

此问题属于试验设计的研究范畴. 试验设计这一学科的产生与发展与R. A. Fisher密不可分. 他1909年入剑桥大学攻读数学和物理，期间受K. Pearson的影响，他的兴趣主要集中在生物学和统计学上. 1914年第一次世界大战爆发后，Fisher也打算投笔从戎，但因视力不好未果. 在此后的5年里他的职业是中学教师，期间他认为农业对于一个国家是非常重要的，于是他曾在一个短时间内经营过一个小型农场，并于1919年由达尔文一位亲戚介绍进入Rothamsted Experimental Station工作. 在此试验站工作期间(1919–1933)，他提出并发展了试验设计这一方向.

例 1.1.5 (质量控制(Quality Control)) *在一个工业产品的设计和生产过程中如何控制其产品质量？如何验证其质量是否合格？又如何改进其产品质量？*

此问题就是质量控制的研究内容，它可分为线外(Off-line)和线内(On-line)质量控制.

例 1.1.6 (时间序列(Time Series) *股票是大家常用的投资工具之一，如把每天的收盘价记录下来，以$\{X_t\}_{t=1}^n$表示过去n天的收盘价，则人们关心的问题就是如何利用这些数据预测未来？*

由于这样的数据是依时间顺次排列的，故它们的研究方法将在时间序列分析这一课程或方向中加以介绍. 当然，类似的与时间有关的数据还有很多，且多与经济、金融等行业有关.

例 1.1.7 (回归(Regression)) *从遗传学的角度看，遗传会把一种性状（如身高）的优势传递给下一代. 如果真是这样的话，我们会看到一代一代的人中，个子很高和很矮的人的比例会日渐升高，而中间部分的比例会日渐下降，但实际上，一代代人的身高却稳定在正态分布，请问如何解释这种情况？*

为解决此问题，F. Galton (1822–1911)花费了十几年的时间，最终此问题得到了园满的解决. Galton是达尔文的近亲表兄，他受达尔文的影响，对遗传学很有兴趣，概率中讲述的Galton板即是他发明的一种产生正态分布的装置（他称这个装置为Quincunx）. 在上述亲子身高问题中，Galton 发现了亲子代间性状遗传中，性状有向中心回归的现象，简言之，高个子的后代平均来说也高些，但不如其亲代那么高，要向平均身高的方向“回归”一些. 其具体验证方法请见陈希孺(2000).

§1.1.2 什么是数理统计

对于数理统计的各种各样的定义，虽无原则性的分岐，但是也很难找到一种说法是完全无懈可击的. 现就这些定义中的一些共性进行一下分析和说明.

1. 必须是受到随机影响的数据，才能成为数理统计学的研究内容.

   “随机性”是数理统计区别其它学科，如数学的一个很重要的一点. 随机性的第一个来源就是试验误差（不是系统误差），第二个来源就是由于研究问题所涉及到的对象太多，故我们只能随机地抽取部分来进行研究所造成的.

2. 如何“有效”地收集数据.

   “有效”有两个方面的含义：一是可以建立一个模型来描述所得数据；二是数据中要尽可能多地包含与研究问题有关的信息，例如想调查某地区共10000农户的经济状况，由

于条件的限制，我们不可能逐户地去调查，现决定从中随机地抽取100户作实际调查，那问题是：

- 100户是否合适？（如何权衡精度与费用）
- 这100户如何去选？（如何抽样以得到更有代表性的数据）

3. 如何“有效”地利用数据.

   获取数据的目的在于提供所研究问题的相关信息，这种信息有时并不是一目了然的，而需要用“有效”的方法去提取或提练，之后再对所研究的问题作出一个结论，这种“结论”在统计上被称为推断(Inference).

   为有效地使用数据进行统计推断，就要涉及到统计中的一些准则，以评价推断的优劣.

§1.1.3 数理统计的应用

数理统计所涉及到的领域是非常广泛的，其原因就是由于随机因素对任何试验的结果都有或多或少的影响.

§1.2 几个基本概念

§1.2.1 样本(sample)和样本分布(distribution)

（一）样本

通过观测或试验而得到的数据就称为样本，又称样品或子样. 如将一物体放在天平上称n次，记录到的数据为$X_1,\dots,X_n$，则它的全体，即$\mathbf{X}=(X_1,\dots,X_n)$就称为样本，其中$X_i$称为第$i$个样本，$n$称为样本大小或容量(Sample Size).

样本分定量(Variable)与定性(Attribute)两种，也有一维与多维之分.

样本可能取值的全体就称为样本空间(Sample Space)，一般记为$\mathcal{X}$.

（二）样本分布

如果仅从应用的角度看，样本就是一组已知的数字，但我们必须注意到，样本是一组受到随机影响的数. 如从概率论的角度看，样本就是随机变量，而我们收集到的具体样本则是这个随机变量的实现或观测值，这既是样本的二重性.

既然样本是随机变量，就有概率分布，于是，这个概率分布就称为样本分布.

例 1.2.1 *假设一批产品共有N个，其中有次品M个，但未知. 现从N中产品中不返回地随机抽取n个产品加以检验，用以估计 $M/N=p$ 的值. 试求样本分布*

对于本问题，假设我们有一个属性指标X，如$X=1$对应着次品，$X=0$对应着合格品. 于是，我们抽取的n个样本可以记为$X_1,\dots,X_n$，其每一个都只能取0 或1. 显然，X_i是一随机变量，且容易验证此时的样本分布为：

$$P\{X_1=x_1,\dots,X_n=x_n\}=\frac{M(M-1)\cdots(M-a+1)}{N(N-1)\cdots(N-a+1)}\frac{(N-M)\cdots(N-M-n+a+1)}{(N-a)\cdots(N-n+1)}, \tag{1.2.1}$$

其中$\sum_{i=1}^n x_i = a$. □

我们回忆一下概率论所学的超几何分布(Hypergeometric Distribution)，知$\sum_{i=1}^n X_i$的分布即为超几何分布，即

$$P\left\{\sum_{i=1}^n X_i = a\right\} = \frac{\begin{pmatrix} M \\ a \end{pmatrix}\begin{pmatrix} N-M \\ n-a \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}.$$

例 1.2.2 *对于上一例子，如果我们考虑有放回抽样，且仍以$X_1, \dots, X_n$记n次抽样得到的结果，即样本，则样本分布如何？*

对于此问题，由于样本之间是独立的，故样本分布为

$$P\{X_1 = x_1, \dots, X_n = x_n\} = \left(\frac{M}{N}\right)^{\sum_{i=1}^n x_i}\left(1-\frac{M}{N}\right)^{\sum_{i=1}^n x_i}. \tag{1.2.2}$$

□

从例1.2.1和例1.2.2可以看出，由于有放回抽样是独立同分布的，而不放回则不是，故不放回比有放回抽样的样本分布要复杂了许多. 但是，当N/n很大时，(1.2.1)与(1.2.2)相差并不太大，故我们可以近似地把不放回抽样当作有放回抽样来处理.

例 1.2.3 *为估计一物体重量μ，现用一架天平将它独立重复测量n次，结果记为$X_1, \dots, X_n$，试确定样本分布.*

对于此问题，为求样本分布，我们可以假设，每次测量都是独立的且是在“相同条件下”进行的，于是，由中心极限定理可知，这n 次测量的结果是独立的且近似服从正态分布$N(\mu, \sigma^2)$. 这样的话，样本分布为

$$f(x_1, \dots, x_n; \mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}\right\}. \tag{1.2.3}$$

□

§1.2.2 总体(Population)与总体分布

总体又称母体，我们常把它理解为“研究问题所涉及的对象的全体的集合”. 总体中的每个元素称为“个体”或单元，从总体中按一定规则抽出一些个体的行为，称为抽样，所抽得的个体称为样本.

对于例1.2.1和例1.2.2而言，由于我们感兴趣的是估计这批产品的次品率，故总体就是这批产品，抽出的n个产品即为样本. 但对于例1.2.3，由于我们感兴趣的是估计一物体的重量，故总体不是现实存在的对象的集合，而只是我们头脑中的抽象：我们把总体理解为“一切可能出现的称量结果的集合”. 当然，不同人可能会给出不同的称量结果的集合，即总体，如区间$[a, b], [0, \infty)$等等. 由于此时的总体已失去了其原来的特性，故我们就仅从样本分布出发研究感兴趣的问题，而不管总体如何.

在有些情况下，从总体出发比从样本具有其更方便之处，于是，我们定义总体分布为当样本容量为1时的样本分布. 之所以定义总体分布，是由于：当我们有n个独立独立同分布的样

本时，如以F记总体分布，故样本$X_1,\ldots,X_n$的分布为$F(x_1)\cdots F(x_n)$, 其写法非常麻烦，此时就不如仅说总体分布为F.

当总体分布为F，而$X_1,\ldots,X_n$为独立同分布的样本时，我们常称$X_1,\ldots,X_n$是从总体F中抽出的简单随机样本或独立同分布的样本(independent identical distribution，简记为iid或IID)，并记

$$X_1,\ldots,X_n \overset{\text{iid}}{\sim} F(x). \tag{1.2.4}$$

若分布F有概率密度函数$f(x)$，则也常记为

$$X_1,\ldots,X_n \overset{\text{iid}}{\sim} f(x). \tag{1.2.5}$$

有时，我们也以一个抽象的记号X来表示所考察的指标，它不是一个样本，而只是一个记号. 我们常把X看成一个随机变量，其分布就是总体分布F. 于是，样本$X_1,\ldots,X_n$是X的观测值，且以

$$X_1,\ldots,X_n \overset{\text{iid}}{\sim} X \tag{1.2.6}$$

表示$X_1,\ldots,X_n$为来自总体X的iid样本.

§1.2.3 样本分布族，参数和参数空间

在统计上，把出现在样本分布中的未知常数称为参数(Parameter)，它可能是一维的，也可能是多维的. 对于例1.2.3，μ就是一个参数，至于σ是否为参数，则取决于实际情况如何.

在某些具体问题中，样本分布中的参数虽然未知，但根据实际情况和该参数的意义，我们可以给出参数所在的大概范围，这个范围就称为参数空间. 对于例1.2.3，如σ未知，则根据它的实际意义，应有$\mu > 0$,$\sigma > 0$，故此时的参数空间可取为

$$\Theta = \{(\mu,\sigma) : \mu > 0, \sigma > 0\}.$$

如果根据以往的经验，我们知道$0 < a \le \mu < b < \infty$，则也可取参数空间为

$$\Theta = \{(\mu,\sigma) : 0 < a \le \mu \le b < \infty, \sigma > 0\}.$$

由于样本分布中包含的未知参数在参数空间中取值，则可能的样本分布就不止一个，而应是一个分布族，于是，我们称之为样本分布族. 样本分布族反映了因我们对所研究问题以及抽样方式的认知，我们能把问题确定到何种程序. 所谓统计推断就是指利用样本推断样本分布族中的未知参数.

在某些实际问题中，其中一些参数可能不是我们感兴趣的，我们称这些参数国讨厌参数.

§1.2.4 统计量

总的来说，为研究一个问题，我们就要收集数据，即样本，之后，通过样本再对问题进行具体的统计推断. 在进行统计推断时，由于样本是一堆杂乱无章的数，故我们必须对它进行必要的加工、整理，以便从中提取那些有利于研究问题的信息出来. 在统计上，把那些凡是能由样本计算出来的量，就称为统计量. 显然，这只是一个定性的定义，而不是一个严密的数学定义. 从这个定义可以看出，一个统计量仅与样本本身有关，而与样本分布或参数没有关系.

下面我们看几个例子.

例 1.2.4 (样本均值(Sample Mean)) 设$X_1,\ldots,X_n$为样本，则

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i \tag{1.2.7}$$

称为样本均值.

例 1.2.5(样本方差(Variance)) 设$X_1,\ldots,X_n$为样本，则

$$S_n^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2 \tag{1.2.8}$$

称为样本方差.

我们注意到，在(1.2.8)式我们定义的样本方差中，其分母为$n-1$，此时我们称之为S_n^2的自由度(degree of freedom)，其解释如下：

- 我们有n个样本$X_1,\ldots,X_n$, 它们都可以自由变化，故其自由度为n，但由于其中一个已用于估计总体均值，故还有$n-1$个可自由变化.
- 从S_n^2的定义不难看出，它是由n个数$X_1-\bar{X},\cdots,X_n-\bar{X}$的平方求和得到的，而这$n$个数有一个约束$\sum_{i=1}^{n}(X_i-\bar{X})=0$，故其自由度只有$n-1$.
- 如果把$\bar{X}$代入$\sum_{i=1}^{n}(X_i-\bar{X})^2$，则可知它是一个如下形式的二次型：$\sum_{ij=1}^{n}a_{ij}X_iX_j$ ($a_{ij}=a_{ji}$)，且不难验证矩阵$A=(a_{ij})$的秩为$n-1$.

例 1.2.6(次序统计量(Order Statistics)) 设$X_1,\ldots,X_n$为样本，把$X_1,\ldots,X_n$由小到大排列成$X_{(1)}\leq\cdots\leq X_{(n)}$，则称$(X_{(1)},\cdots,X_{(n)})$为次序统计量，$X_{(i)}$称为第$i$个次序统计量.

通过次序统计量，我们有如下一些很实用的统计量，如

- 样本p分位数(quantile): 对于给定的$p\in(0,1)$，我们称

$$m_{n,p}=X_{([np])}+(n+1)\left(p-\frac{[np]}{n+1}\right)(X_{([np]+1)}-X_{([np])}), \tag{1.2.9}$$

为此样本的p分位数，其中$[x]$表示不超过x的最大整数. 特别地，样本中位数(median)定义为

$$X_{med}=\begin{cases}X_{((n+1)/2)}, & \text{如果}n\text{是奇数},\\ (X_{(n/2)}+X_{(n/2+1)})/2, & \text{如果}n\text{是偶数}.\end{cases}$$

注 1.2.1 关于样本p的定义还有许多种，如$X_{([np])}$，$X_{([np]+1)}$，和$\inf\{x:F_n(x)\geq p\}$等(F_n的含义见定义1.2.1). 这些定义虽然有所不同，但它们均是一个（或两个）次序统计量，且随着样本容量的加大，它们间的差别并不大($k=np+o(\sqrt{n})$).

- 极值(Extreme Value): 称$X_{(1)}$与$X_{(n)}$为极小与极大值统计量.

- 极差(Range): $R = X_{(n)} - X_{(1)}$.

例 1.2.7 设$X_1, \ldots, X_n$为样本，则称$S_n/\bar{X}$为样本变异系数(Coefficient of Variability).

例 1.2.8 (样本k阶矩) 称$m_k = \frac{1}{n}\sum_{i=1}^n X_i^k$与$\nu_k = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^k$为样本$k$阶原点矩(Moment)和样本$k$阶中心矩.

例 1.2.9 (样本偏度(Skewness)与峰度(Kurtosis)) 称$\sqrt{n}\sum_{i=1}^n (X_i - \bar{X})^3/[\sum_{i=1}^n (X_i - \bar{X})^2]^{3/2}$与$n\sum_{i=1}^n (X_i - \bar{X})^4/[\sum_{i=1}^n (X_i - \bar{X})^2]^2 - 3$为样本偏度与峰度.

§1.2.5 经验分布函数(Empirical Distribution)

定义 1.2.1 设$X_1, \ldots, X_n$为取自总体分布函数为$F(x)$的样本，$X_{(1)} \le \cdots \le X_{(n)}$为其次序统计量，则称

$$F_n(x) = \begin{cases} 0, & x \le X_{(1)}, \\ \frac{k}{n}, & X_{(k)} < x \le X_{(k+1)}, k = 1, \cdots, n-1, \\ 1, & x > X_{(n)} \end{cases} \tag{1.2.10}$$

为样本$X_1, \ldots, X_n$的经验分布函数.

可以证明$F_n(x)$是一个非降的左连续函数，且具有如下性质(当样本是iid时)：

- $F_n(x) \xrightarrow{P} F(x),\ \forall\ x \in R$.
- $F_n(x) \xrightarrow{\text{2nd}} F(x),\ \forall\ x \in R$.
- $P\{\lim_n F_n(x) = F(x)\} = 1,\ \forall\ x \in R$.
- $F_n(x)$ is $AN\left(F(x), \frac{F(x)(1-F(x))}{n}\right),\ \forall\ x \in R$.

从上面性质可知，对于任意给定的实数x，当n充分大时，$F_n(x)$与$F(x)$的值，在概率意义下是非常接近的，下一定理就给出了这种接近程度的大样本度量.

定理 1.2.1 (Glivenko-Cantelli) 对于任意给定的自然数n，设$X_1, \ldots, X_n$为取自总体分布$F(x)$的iid样本，$F_n(x)$为其经验分布函数，记

$$D_n = \sup_x |F_n(x) - F(x)|,$$

则有$P\{\lim_{n\to\infty} D_n = 0\} = 1$.

其证明可见茆诗松、王静龙(1990) P8.

注 1.2.2 For the 1-dimensional case, this result was proved by Glivenko in 1933 for continuous F (*Giorn. Inst. Ital. Attuari*, 4, 92–99) and by Cantelli in 1933 for general F (*Giorn. Inst. Ital. Attuari*, 4, 421–424).

事实上，关于上述统计量D_n，其研究结果相当丰富，我们不加证明地给出如下结论：

定理 1.2.2 (Dvoretzky, Kiefer, and Wolfowitz) 对于iid样本$X_1, \ldots, X_n$, 存在一个不依赖于分布函数F的正常数C，使得

$$P\{D_n > d\} \le Ce^{-2nd^2},\ \forall\ d > 0.$$

上述不等式称为D_n的指数型不等式，它是上述三位作者于1956年证明的(*Ann. Math. Statist.* 27, 642–669). 另外，Kolmogorov于1933年给出了D_n在一维时的分布如下：

定理 1.2.3 (Kolmogorov) Let F be 1-dimensional and continous. Then

$$\lim_{n\to\infty} P\{n^{1/2}D_n \le d\} = 1 - 2\sum_{j=1}^{\infty}(-1)^{j+1}e^{-2j^2d^2},\ \forall\ d > 0.$$

关于D_n的应用，请见§6.4的Kolmogorov检验.

§1.2.6 抽样分布

由于统计量是作为随机变量的样本的函数，故它也有概率分布，于是，我们称统计量的概率分布为该统计量的抽样分布(Sampling Distribution)，这个概念是由R. A. Fisher于1922年提出的.

虽然对于给定的一组样本，用于统计推断的统计量是一个固定的数，但由于样本受到了随机因素的影响，故其推断结果也是随机的，所以，有可能出现这样的情况：一个整体上看来较好的统计推断方法，在个别情况下也可能给出不好的结果. 因此，我们可以用抽样分布来衡量统计推断的好坏.

例如在例1.2.3中，我们可以用样本均值$\bar{X}$来估计正态分布的均值μ，且我们希望二者非常接近. 但由于受到随机的影响，有时二者可能会有一定的差距，为了衡量二者的差距$P\{|\bar{X} - \mu| > c\}$，我们就必须知道统计量$\bar{X}$的抽样分布.

例 1.2.10 设$X_1, \ldots, X_n$为来自$N(\mu, \sigma^2)$的iid样本，则由概率论知识不难求得

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right),$$

更一般地有，

$$T = \sum_{i=1}^{n} c_i X_i \sim N\left(\mu \sum_{i=1}^{n} c_i, \sigma^2 \sum_{i=1}^{n} c_i^2\right).$$

从例1.2.10可能看出，$\bar{X}$的分布与参数μ, σ^2有关，虽然我们在前面说过，统计量仅依赖于样本而与参数无关，这一点并不矛盾. 事实上，如果一个统计量的分布与感兴趣的参数无关，那我们如何利用这个统计量来进行统计推断？故有用的统计量的抽样分布肯定会与参数有关.

例 1.2.11 以T记在例1.2.1和1.2.2中抽出的次品数，求T的抽样分布

解 对于无放回抽样：由于T服从超几何分布，故

$$P\{T=t\}=\frac{\begin{pmatrix} M \\ t \end{pmatrix}\begin{pmatrix} N-M \\ n-t \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}. \tag{1.2.11}$$

对于有放回抽样：由于T服从二项分布$B(n,p)$,其中$p=M/N$，故

$$P\{T=t\}=\begin{pmatrix} n \\ t \end{pmatrix} p^t(1-p)^{n-t},\ t=0,1,\cdots,n. \tag{1.2.12}$$

□

例 1.2.12 (次序统计量) 设$X_1,\ldots,X_n\overset{\text{iid}}{\sim}F(x)$，求次序统计量$X_{(k)}$的抽样分布.

解 实际上，我们有

$$\begin{aligned}
F_k(x) &= P\{X_{(k)}<x\}=P\{\text{在}X_1,\ldots,X_n\text{中至少有}k\text{个小于}x\} \\
&= \sum_{m=k}^{n} P\{\text{在}X_1,\ldots,X_n\text{中恰有}m\text{个小于}x\} \\
&= \sum_{m=k}^{n} \begin{pmatrix} n \\ m \end{pmatrix} F^m(x)(1-F(x))^{n-m},
\end{aligned}$$

即

$$F_k(x)=\sum_{m=k}^{n} \begin{pmatrix} n \\ m \end{pmatrix} F^m(x)(1-F(x))^{n-m}, \tag{1.2.13}$$

另外，不难证明，我们还可以把上式改写成如下积分形式：

$$F_k(x)=k\begin{pmatrix} n \\ k \end{pmatrix}\int_0^{F(x)} t^{k-1}(1-t)^{n-k}\mathrm{d}t. \tag{1.2.14}$$

如总体分布$F(x)$有密度函数$f(x)$，则$X_{(k)}$也有密度，且为

$$f_k(x)=k\begin{pmatrix} n \\ k \end{pmatrix} F^{k-1}(x)(1-F(x))^{n-k}f(x). \tag{1.2.15}$$

特别地，当$k=1$和$k=n$时，我们有极小与极大次序统计量的CDF与PDF如下：

$$\begin{aligned}
F_1(x) &= 1-(1-F(x))^n, & f_1(x) &= n(1-F(x))^{n-1}f(x), \\
F_n(x) &= F^n(x), & f_n(x) &= nF^{n-1}(x)f(x).
\end{aligned}$$

□

§1.3 统计中常用的抽样分布

虽然抽样分布对于统计推断很重要，但在许多情况下，我们并不容易得到统计量的抽样

分布. 然而，当总体为正态时，许多重要统计量的抽样分布却可以很容易得到，这即是本节的内容.

§1.3.1 χ^2分布

定义 1.3.1 (χ^2分布) 设$X_1,\ldots,X_n \overset{\text{iid}}{\sim} N(0,1)$，则称随机变量

$$\xi=\sum_{i=1}^{n}X_i^2 \tag{1.3.1}$$

所服从的分布为自由度n的χ^2分布，记为$\xi\sim\chi^2(n)$.

关于χ^2分布PDF $f(x)$，我们有如下的结论.

定理 1.3.1 由(1.3.1)式所定义的随机变量ξ的PDF为

$$f(x)=\begin{cases}\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}e^{-\frac{x}{2}}x^{\frac{n}{2}-1}, & x>0,\\ 0, & x\le 0,\end{cases} \tag{1.3.2}$$

其中n为自由度.

证明：为求ξ的PDF，我们先求它的CDF $F(x)=P\{\xi<x\}$. 显然，当$x\le 0$时，$F(x)=0$; 当$x>0$时，

$$F(x)=P\left\{\sum_{i=1}^{n}X_i^2<x\right\}. \tag{1.3.3}$$

因为$X_1,\ldots,X_n$相互独立且同分布于$N(0,1)$，故$X_1,\ldots,X_n$的联合PDF为

$$\frac{1}{(2\pi)^{n/2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}x_i^2\right\}.$$

于是，

$$F(x)=\frac{1}{(2\pi)^{n/2}}\int\cdots\int_{\sum_{i=1}^{n}x_i^2<x}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}x_i^2\right\}\mathrm{d}x_1\cdots\mathrm{d}x_n,$$

为求此积分，做如下的球形变换

$$\begin{cases}x_1=\rho\cos\theta_1\cos\theta_2\cdots\cos\theta_{n-1},\\ x_2=\rho\cos\theta_1\cos\theta_2\cdots\sin\theta_{n-1},\\ \cdots\\ x_n=\rho\sin\theta_1,\end{cases}$$

此变换的Jacob矩阵行列式为

$$J=\frac{\partial(x_1,\ldots,x_n)}{\partial(\rho,\theta_1,\ldots,\theta_{n-1})}=\rho^{n-1}D(\theta_1,\ldots,\theta_{n-1}),$$

其中$D(\theta_1,\ldots,\theta_{n-1})$与$\rho$无关. 于是，

$$\begin{aligned}F(x) &= \frac{1}{(2\pi)^{n/2}}\int_0^{\sqrt{x}}\mathrm{d}\rho\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\int_{-\pi}^{\pi}e^{-\rho/2}\rho^{n-1}D(\theta_1,\ldots,\theta_{n-1})\mathrm{d}\theta_1\cdots\mathrm{d}\theta_{n-1}\\ &= C_n\int_0^{\sqrt{x}}e^{-\rho/2}\rho^{n-1}\mathrm{d}\rho,\end{aligned}$$

其中$C_n=\frac{1}{(2\pi)^{n/2}}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\int_{-\pi}^{\pi}D(\theta_1,\ldots,\theta_{n-1})\mathrm{d}\theta_1\cdots\mathrm{d}\theta_{n-1}$.

令$\rho=\sqrt{y}$，则$\mathrm{d}\rho=\frac{\mathrm{d}y}{2\sqrt{y}}$，于是，

$$\begin{aligned}F(x) &= C_n\int_0^x e^{-y/2}y^{(n-1)/2}\frac{1}{2y^{1/2}}\mathrm{d}y\\ &= \frac{C_n}{2}\int_0^x e^{-y/2}y^{n/2-1}\mathrm{d}y.\end{aligned}$$

又因为

$$1=F(+\infty)=\frac{C_n}{2}\int_0^\infty e^{-y/2}y^{n/2-1}\mathrm{d}y,$$

而

$$\int_0^\infty e^{-y}y^{n/2-1}\mathrm{d}y=\Gamma(n/2),$$

故

$$C_n=\frac{1}{2^{\frac{n}{2}-1}\Gamma(\frac{n}{2})}.$$

这样，χ^2分布的CDF为

$$F(x)=\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}\int_0^x e^{-\frac{y}{2}}y^{\frac{n}{2}-1}\mathrm{d}y, \tag{1.3.4}$$

其PDF为

$$f(x)=\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}e^{-\frac{x}{2}}x^{\frac{n}{2}-1},\ \forall\ x>0.$$

□

图1.1给出了自由度分别为$1,5,10,20$时的χ^2分布PDF的曲线.

从图1.1我们可以看到

- 随着n的增大，它的对称性越来越好，峰度越来越小.
- 随着n的增大，其图形越来越像正态分布的概率密度函数.
- 随着n的增大，它的图形越来越向右移动，且尾部越来越大.

例 1.3.1 设$X_1,\ldots,X_n\overset{\text{iid}}{\sim}N(\mu,\sigma^2)$, 其中$\mu$是已知常数, 求统计量

$$\eta=\sum_{i=1}^n(X_i-\mu)^2 \tag{1.3.5}$$

n=5

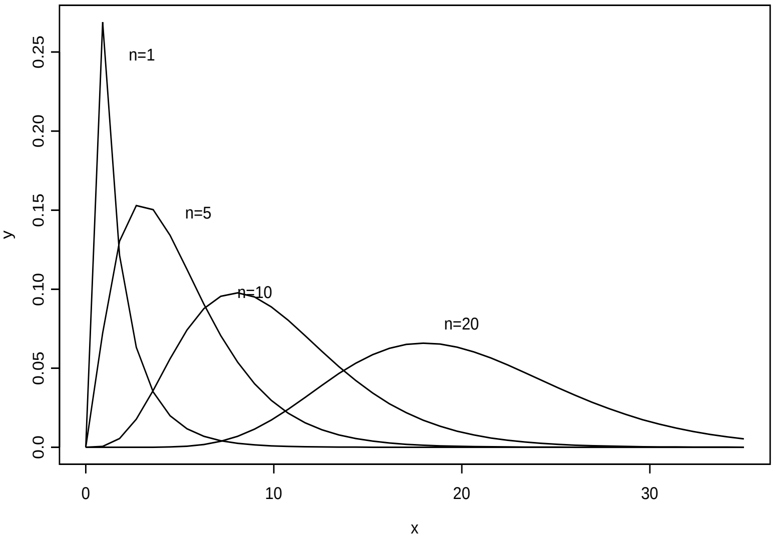

图1.1 χ^2分布的概率密度函数

的分布.

解 如令$Y_i=(X_i-\mu)/\sigma,\ i=1,\cdots,n$，则知$Y_1,\ldots,Y_n\overset{\text{iid}}{\sim}N(0,1)$, 于是由定理1.3.1知

$$\frac{\eta}{\sigma^2}=\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2=\sum_{i=1}^{n}Y_i^2\sim\chi^2(n),$$

于是，由(1.3.2)式可知η的PDF为

$$f(x)=\begin{cases}\frac{1}{2^{\frac{n}{2}}\sigma^n\Gamma(\frac{n}{2})}e^{-\frac{x}{2\sigma^2}}x^{\frac{n}{2}-1}, & x>0,\\ 0, & x\leq 0.\end{cases}\tag{1.3.6}$$

□

比较(1.3.6)与(1.3.2)可知，二者只相差一个因子σ^2，故有时我们也称(1.3.5)式的统计量为Scalled χ^2. 另外，从(1.3.6)可以看到，(1.3.5)式的χ^2分布提供了关于σ^2的有关信息，所以，我们将用它来估计σ^2.

χ^2分布在统计上非常有用，其原因与下面的一些结论有关.

定理 1.3.2 设$\xi\sim\chi^2(n)$，则

(1) ξ的特征函数为$\psi(t)=(1-2it)^{-\frac{n}{2}}$. (1.3.7)

(2) $E\xi=n,\mathrm{Var}\xi=2n$. (1.3.8)

证明：由特征函数的定义有

$$\begin{aligned}\psi(t) &= Ee^{i\xi t} = \int_0^\infty e^{ixt}\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}e^{-\frac{x}{2}}x^{\frac{n}{2}-1}\mathrm{d}x \\ &= \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}\int_0^\infty e^{-(\frac{1}{2}-it)x}x^{\frac{n}{2}-1}\mathrm{d}x \\ &= \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}\frac{\Gamma(n/2)}{(1/2-it)^{n/2}} = (1-2it)^{-n/2}.\end{aligned}$$

于是，由概率论知识知

$$E\xi = \psi'(0)/i = n,\ E\xi^2 = \psi''(0)/i^2 = n(n+2),$$

这样，$\mathrm{Var}\xi = 2n$. □

定理 1.3.3 设$\xi \sim \chi^2(m)$, $\eta \sim \chi^2(n)$, 且ξ,η相互独立，则$\xi+\eta \sim \chi^2(m+n)$.

证明：略.

注 1.3.1 由定理1.3.3知，两个独立的χ^2分布之和仍然是χ^2分布，且其自由度为两者之和，多个之和见下面的推论.

Corollary 1.3.1 设$\xi_i \sim \chi^2(n_i)$, $i=1,2,\cdots,m$，且相互独立，则$\sum_{i=1}^m \xi_i \sim \chi^2(\sum_{i=1}^m n_i)$.

χ^2分布有用的一个重要原因即是下面的定理.

定理 1.3.4 设$X_1,\ldots,X_n \overset{\text{iid}}{\sim} N(\mu,\sigma^2)$，$\bar{X}$和$S_n^2$分别是样本均值和方差，则

- $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$;
- $(n-1)S_n^2/\sigma^2 \sim \chi^2(n-1)$;
- $\bar{X}$与S_n^2独立.

证明：记A为如下的正交阵

$$A = \begin{pmatrix} \frac{1}{\sqrt{n}} & \cdots & \cdots & \cdots & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & \cdots & 0 & 0 \\ \frac{1}{\sqrt{2\cdot 3}} & \frac{1}{\sqrt{2\cdot 3}} & -\frac{2}{\sqrt{2\cdot 3}} & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{\sqrt{n(n-1)}} & \cdots & \frac{1}{\sqrt{n(n-1)}} & \cdots & \frac{1}{\sqrt{n(n-1)}} & -\frac{n-1}{\sqrt{n(n-1)}} \end{pmatrix}.$$

作如下正交变换

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} = A\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix},$$

则有

$$Y_1=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_i=\sqrt{n}\bar{X},\ \sum_{i=1}^{n}Y_i^2=\sum_{i=1}^{n}X_i^2=\sum_{i=1}^{n}(X_i-\bar{X})^2+n\bar{X}^2.$$

综合以上二式，我们有

$$(n-1)S_n^2=\sum_{i=1}^{n}(X_i-\bar{X})^2=\sum_{i=1}^{n}Y_i^2-n\bar{X}^2=\sum_{i=2}^{n}Y_i^2,$$

另外，由概率论知识知道，$Y_1,\ldots,Y_n$也是独立的正态随机变量且

$$EY_1=\sqrt{n}\mu,\ EY_k=0(k\geq 2),\ \mathrm{Var}Y_i=\sigma^2(k=1,\cdots,n),$$

于是，Y_1与$(n-1)S_n^2$独立，即$\bar{X}$与S_n^2独立，且

$$\bar{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right),\ \ (n-1)S_n^2/\sigma^2\sim\chi^2(n-1).$$

□

注 1.3.2 If the mean and variance of random samples from a population are independent, the population must be normal. (See Stuart, A., Ord, J. K. (1987). *Kendall's Advanced Theory of Statistics*. Charles Griffin & Company Limited.) In fact, for the general case the covariance between sample mean and sample variance is given by $\mathrm{Cov}(\bar{X},S_n^2)=\frac{\nu_3}{n}$，where $\nu_k=E(X-E(X))^k$.

从χ^2分布的定义不难看出，χ^2分布是若干个独立的$N(0,1)$随机变量的平方和，而平方和是二次型的一个特殊形式，那一个自然的问题是：一个正态随机向量的二次型的分布如何呢？这即是下面的Cochran定理.

定理 1.3.5 (Cochran) 设$X_1,\ldots,X_n$相互独立，且$X_i\sim N(\mu_i,\sigma^2)$，$i=1,\cdots,n$，令$\mathbf{X}=(X_1,\ldots,X_n)'$. 又设$A_1,\ldots,A_m$是$m$个$n$阶非负定阵，且$A_1+\cdots+A_m=I_n$ (n阶单位阵). 记$\xi_i=\mathbf{X}'A_i\mathbf{X}$，$i=1,\cdots,m$，$\mu=(\mu_1,\ldots,\mu_n)'$. 则

- 如果$\sum_{i=1}^m rk(A_i)=n$，则$\xi_1,\ldots,\xi_m$相互独立；
- 如果$\sum_{i=1}^m rk(A_i)=n$，且$\mu'A_1\mu=0$，则$\xi_1/\sigma^2\sim\chi^2(rk(A_1))$.

证明：记$n_i=rk(A_i)$，$i=1,\cdots,m$. 因为A_i是n阶非负定阵，则存在一个$n\times n_i$阵B_i，使得$A_i=B_iB_i'$. 记$B=(B_1B_2\cdots B_m)$，则由已知条件可知B是一个n阶正交阵.

作如下正交变换：

$$\mathbf{Y}=\begin{pmatrix}Y_1\\ \vdots\\ Y_n\end{pmatrix}=B'\mathbf{X}=\begin{pmatrix}B_1'\mathbf{X}\\ \vdots\\ B_m'\mathbf{X}\end{pmatrix},$$

即

$$B_i'\mathbf{X} = \begin{pmatrix} Y_{n_1+\cdots+n_{i-1}+1} \\ Y_{n_1+\cdots+n_{i-1}+2} \\ \vdots \\ Y_{n_1+\cdots+n_{i-1}+n_i} \end{pmatrix}, \ i = 1, \cdots, m,$$

其中$n_0 = 0$.

由于B是一正交阵，故由概率论知识知道：$Y_1, \ldots, Y_n$独立，且$Y_i \sim N(\beta_i, \sigma^2),\ i = 1, \cdots, n$,其中$\beta = (\beta_1, \ldots, \beta_n)' = B'\mu$.

又因为

$$\xi_i = \mathbf{X}'A_i\mathbf{X} = \mathbf{X}'B_iB_i'\mathbf{X} = (B_i'\mathbf{X})'B_i'\mathbf{X} = \sum_{j=n_1+\cdots n_{i-1}+1}^{n_1+\cdots n_{i-1}+n_i} Y_j^2, \ i = 1, \cdots, m,$$

故$\xi_1, \ldots, \xi_m$相互独立，第一点得证.

下证第二点. 由已知条件知，$\mu'A_1\mu = 0$，即$\mu'B_1B_1'\mu = 0$，于是，$B_1'\mu = \mathbf{0}$, 即$\beta_i = 0,\ i = 1, \cdots, n_1$.

又由于$\xi_1/\sigma^2 = \sum_{i=1}^{n_1} Y_i^2/\sigma^2$，而$Y_1, \ldots, Y_{n_1} \overset{\text{iid}}{\sim} N(0, \sigma^2)$，故$\xi_1/\sigma^2 \sim \chi^2(n_1)$. □

在上面我们所讨论的χ^2分布，均是指若干个独立的标准正态随机变量的平方和. 如不是标准正态，则我们也必须要求其均值为零，否则我们得不到一个χ^2随机变量. 下面，我们将其推广，到得非中心χ^2分布的定义.

定义 1.3.2 (非中心χ^2) 设$X_1, \ldots, X_n$是独立的正态随机变量，$X_i \sim N(\mu_i, 1)$, 且μ_i不全为零，则称

$$\xi = \sum_{i=1}^{n} X_i^2 \tag{1.3.9}$$

为非中心的χ^2随机变量，称它的分布为非中心χ^2分布，自由度为n，非中心参数为$\delta = \sum_{i=1}^{n} \mu_i^2$，并记为$\xi \sim \chi^2(n, \delta)$.

我们注意到，有的书上也取非中心参数为$\delta = (\sum_{i=1}^{n} \mu_i^2)^{1/2}$.

用类似于定理1.3.1的方法，不难求得非中心χ^2分布的PDF为

$$e^{-\delta/2} \sum_{m=0}^{\infty} \frac{(\delta/2)^m}{m!} \chi^2(x; 2m+n), \tag{1.3.10}$$

其中$\chi^2(x; l)$为由(1.3.2)式决定的自由度为l的χ^2分布的PDF. 另外，也不难计算非中心χ^2分布的特征函数为

$$\psi(t) = (1-2it)^{-n/2} \exp\{i\delta^2/(1-2it)\}, \tag{1.3.11}$$

其期望与方差分别为$n + \delta$和$2n + 4\delta$.

有时为了区别由定义1.3.1给出的χ^2分布与由定义1.3.2给出的非中心χ^2的区别，我们也称由定义1.3.1给出的χ^2分布为中心的χ^2分布.

§1.3.2 t分布

在二十世纪初以前，或更具体地说，是在1908年以前，统计学的主要用武之地先是社会统计，尤其是人口统计，后来加入生物统计问题. 在这些问题中的数据一般都是大量的、自然采集的. 所用的方法多以Laplace中心极限定理为依据，总是归结到正态. K. Pearson就是此时统计的统治者. 但到了20世纪，受人工控制的试验条件下所得数据的统计分析问题，日渐引人注意. 此时的数据量一般不大，故那种依赖于中心极限定理的传统方法，开始受到质疑. 这个方向的先驱就是W. S. Gosset (1876–1937)和R. A. Fisher.

Gosset于1899 年作为一名酿酒师进入爱尔兰的都柏林的一家啤酒厂工作，在那里他涉及到有关酿造过程的数据处理问题. 于1906到1907年他到Pearson那里学习统计学，并着重研究少量数据的统计分析问题. 于1908年他在Biometrics杂志上以笔名Student发表了使他名垂统计史册的论文：均值的或然误差. 在这篇文章中，他提出了如下结果：设$X_1, \ldots, X_n$是来自$N(\mu, \sigma^2)$的iid样本，μ, σ^2均未知，则$\frac{\sqrt{n}(\bar{X}-\mu)}{S}$服从自由度为$n-1$的$t$分布，并给出了这种$t$分布的一些分位数值(分位数的概念，我们将在假设检验中加以介绍). 可以说，小样本统计分析由此引起了广大统计科研工作者的重视，虽然Gosset中的证明存在着漏洞. 最早注意到这个问题的是Fisher，并于1922年给出了此问题的完整证明. 另外，我们注意到，当时许多统计学家在Gosset于1937年去世后，尚不知他就是Student. 有关这方面的资料陈希孺(2000)给出了非常详尽的描述.

定义 1.3.3 *设$\xi \sim N(0,1)$, $\eta \sim \chi^2(n)$，且ξ, η相互独立，则称随机变量*

$$T = \frac{\xi}{\sqrt{\eta/n}} \tag{1.3.12}$$

所服从的分布为t分布，n为其自由度，且记为$T \sim t(n)$.

由于t分布是由Student提出的，故有时也称之为Student t分布，其PDF由下面定理给出.

定理 1.3.6 *由(1.3.12)式所定义的$t(n)$分布的PDF为*

$$f(x) = \frac{\Gamma((n+1)/2)}{\Gamma(n/2)\sqrt{n\pi}}(1+x^2/n)^{-\frac{n+1}{2}}. \tag{1.3.13}$$

证明：因为ξ, η相互独立，故可知ξ, η的联合PDF为

$$Ce^{-x^2/2}e^{-y^2/2}y^{n/2-1}, \tag{1.3.14}$$

其中C为不依赖于x, y的常数，但与n有关.

为了求$\xi/\sqrt{\eta/n}$的PDF，我们作如下变换：

$$\begin{cases} \xi = R\sin\Theta, & 0 < R < \infty, \\ \eta = R^2\cos^2\Theta, & -\frac{\pi}{2} < \Theta < \frac{\pi}{2}, \end{cases} \tag{$*1$}$$

由(1.3.14)式及上式，我们可以求得R, Θ的联合PDF为

$$2Ce^{-r^2/2}r^n(\cos\theta)^{n-1}. \tag{$*2$}$$

利用概率论知识，由上式可知，R与Θ是独立的，且Θ的PDF为

$$C_1(\cos\theta)^{n-1}, \tag{$*3$}$$

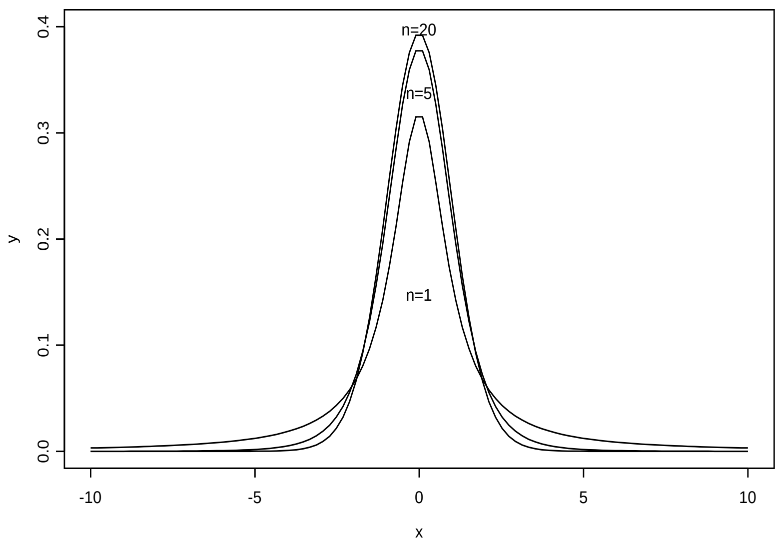

图1.2 t分布的概率密度函数

其中$C_1^{-1} = \int_{-\pi/2}^{\pi/2} (\cos\theta)^{n-1} \mathrm{d}\theta = \frac{\Gamma(1/2)\Gamma(n/2)}{\Gamma((n+1)/2)}$.

又由于$T = \frac{\sin\Theta}{\cos\Theta}\sqrt{n} = \sqrt{n}\tan\Theta$, 如令$t = \sqrt{n}\tan\theta$, 则$\cos\theta = (1+t^2/n)^{-1/2}$, $\mathrm{d}t = \sqrt{n}\sec^2\theta \mathrm{d}\theta = \sqrt{n}(1+t^2/n)\mathrm{d}\theta$, 于是，由(*3)式可得到$t$分布随机变量的PDF为

$$\begin{aligned} f(t) &= C_1 \left[(1+t^2/n)^{-1/2}\right]^{n-1} \frac{1}{\sqrt{n}(1+t^2/n)} \\ &= \frac{\Gamma((n+1)/2)}{\Gamma(n/2)\sqrt{n\pi}}(1+t^2/n)^{-\frac{n+1}{2}}. \end{aligned}$$

□

为了对t分布有一个直观的了解，我们看一看$n = 1, 5, 20$时的PDF曲线，见图1.2.

从图1.2，我们看到

- $t(n)$的PDF关于y轴对称，且$\lim_{|x|\to\infty} f(x) = 0$.
- 随着n的增大，其峰度越来越高，尾部越来直小.
- 由于对于固定的x, 我们有
$$\lim_{n\to\infty}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}} = e^{-x^2/2},$$
故当n很大时，t分布的PDF接近于标准正态分布的PDF. (其常数可用Stirling公式：$n! = \sqrt{2\pi n}n^n e^{-n}$求得.)
- 当$n = 1$时，它是Cauchy分布，故此时期望不存在.

关于t分布的数字特征，有如下的结论.

定理 1.3.7 设$\xi \sim t(n), n > 1$，则 $\forall\ r < n$，$E\xi^r$存在，且有

$$E\xi^r = \begin{cases} 0, & r\text{是奇数}, \\ n^{r/2}\frac{\Gamma((r+1)/2)\Gamma((n-r)/2)}{\Gamma(1/2)\Gamma(n/2)}, & r\text{是偶数}. \end{cases} \tag{1.3.15}$$

证明：略.

由定理1.3.7可知，如果$\xi \sim t(n)$且$n > 2$，则$E\xi = 0, \mathrm{Var}\xi = \frac{n}{n-2}$.

例 1.3.2 设$\xi \sim N(\mu, \sigma^2)$，$\eta/\sigma^2 \sim \chi^2(n)$，且$\xi, \eta$相互独立，则易知

$$T = \frac{\xi - \mu}{\sqrt{\eta/n}} \sim t(n). \tag{1.3.16}$$

在t分布的定义中，我们始终假设分子的正态分布的均值为零，如果非零的话，就是下面的非中心t分布的定义.

定义 1.3.4 (非中心t分布) 设$\xi \sim N(\mu, \sigma^2), \eta/\sigma^2 \sim \chi^2(n)$，且二者独立，$\mu \neq 0$，则称

$$T = \frac{\xi}{\sqrt{\eta/n}}$$

所服从的分布为非中心的t分布，其非中心参数为$\delta = \mu/\sigma$, 自由度为n.

利用定理1.3.6的方法，容易求得非中心t分布的PDF为

$$\frac{n^{n/2}e^{-\delta^2/2}}{\sqrt{\pi}\Gamma(n/2)(n+x^2)^{(n+1)/2}}\sum_{m=0}^{\infty}\Gamma\left(\frac{n+m+1}{2}\right)\left(\frac{\delta^m}{m!}\right)\left(\frac{\sqrt{2}x}{\sqrt{n+x^2}}\right)^m,$$

其期望与方差分别为

$$ET = \delta\frac{\Gamma((n-1)/2)}{\Gamma(n/2)}\sqrt{\frac{n}{2}},\ n > 1$$

$$\mathrm{Var}T = \frac{n(1+\delta^2)}{n-2} - \frac{\delta^2 n}{2}\left(\frac{\Gamma((n-1)/2)}{\Gamma(n/2)}\right)^2,\ n > 2.$$

§1.3.3 F分布

定义 1.3.5 设ξ, η是独立的，自由度分别为m, n的χ^2随机变量，则称随机变量

$$F = \frac{\xi/m}{\eta/n} \tag{1.3.17}$$

所服从的分布为F分布，其自由度为(m, n)，且记为$F \sim F(m, n)$.

从上述定义可以看出，当$n \to \infty$时，统计量mF依概率收敛于一个自由度为m的χ^2随机变量. 同理，当$m \to \infty$时，$\frac{1}{nF}$依概率收敛于自由度为n的随机变量.

F分布的PDF由下面定理给出.

定理 1.3.8 由上面的定义的F分布的PDF为

$$f(x;m,n)=\begin{cases}0, & x<0,\\ \frac{\Gamma((m+n)/2)}{\Gamma(m/2)\Gamma(n/2)}\left(\frac{m}{n}\right)\left(\frac{mx}{n}\right)^{m/2-1}\left(1+\frac{mx}{n}\right)^{-(m+n)/2}, & x>0.\end{cases} \tag{1.3.18}$$

证明：由于$\xi\sim\chi^2(m),\eta\sim\chi^2(n)$，且二者独立，故它们的联合PDF为

$$f(x,y)=\frac{1}{2^{(m+n)/2}\Gamma(m/2)\Gamma(n/2)}e^{-(x+y)/2}x^{m/2-1}y^{n/2-1},\ x>0,\ y>0.$$

作如下变换：

$$\begin{cases}u=x+y,\\ v=\frac{x}{y}\frac{n}{m},\end{cases}$$

则(U,V)的联合PDF为

$$\begin{aligned}g(u,v)&=\frac{1}{2^{(m+n)/2}\Gamma(m/2)\Gamma(n/2)}e^{-u/2}u^{\frac{m+n}{2}-2}\cdot\\ &\quad\left(\frac{m}{n}\right)^{m/2-1}\frac{v^{\frac{m}{2}-1}}{(1+mv/n)^{\frac{m+n}{2}-2}}\frac{m}{n}\frac{u}{(1+mv/n)^2}\\ &=\frac{1}{2^{\frac{m+n}{2}}\Gamma((m+n)/2)}e^{-u/2}u^{\frac{m+n}{2}-2}\cdot\\ &\quad\frac{\Gamma((m+n)/2)(m/n)^{m/2}}{\Gamma(m/2)\Gamma(n/2)}\frac{v^{m/2-1}}{(1+mv/n)^{\frac{m+n}{2}}},\end{aligned}$$

则知U,V相互独立，且$U\sim\chi^2(m+n)$, 而第二部分即为V的PDF，即为所求. □

Corollary 1.3.2 设$\xi\sim\chi^2(m),\eta\sim\chi^2(n)$，且$\xi$与$\eta$相互独立，则$Y=\xi+\eta$与$Z=\xi/\eta$相互独立.

注 1.3.3 设随机变量$X\sim F(m,n)$，则$\frac{1}{X}\sim F(n,m)$.

注 1.3.4 如在(1.3.17)式中取$m=1$，则知$F(1,n)=(t(n))^2$, 即来自自由度为$(1,n)$的F分布的随机变量X与来自自由度为n的t分布的随机变量Y，二者的关系为$X\overset{d}{=}Y^2$.

为了直观地了解F分布的PDF，我们在图1.3中给出了$(m,n)=(10,5),(10,10),(10,100)$时的$F$分布的PDF曲线.

关于F分布的数字特征，我们有如下的定理.

定理 1.3.9 设$\xi\sim F(m,n)$，则对$r>0$有

$$E\xi^r=\left(\frac{n}{m}\right)^r\frac{\Gamma(r+\frac{m}{2})\Gamma(\frac{n}{2}-r)}{\Gamma(m/2)\Gamma(n/2)},\ 2r<n.$$

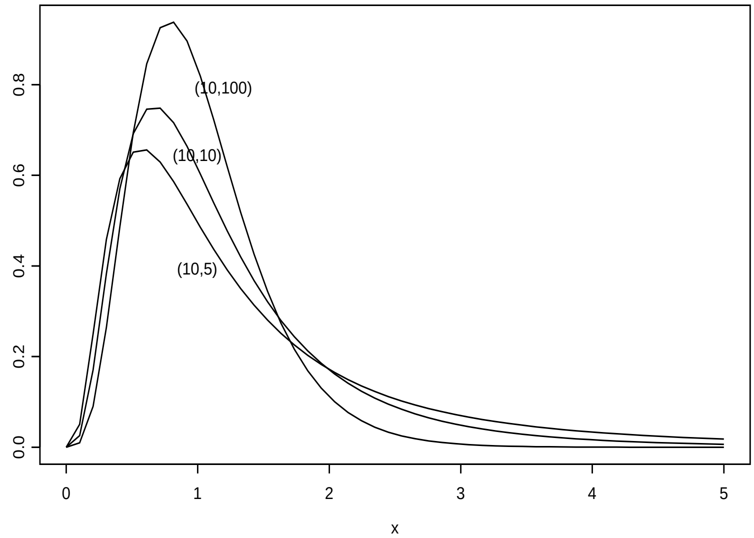

图1.3 F分布的概率密度函数

特别地，其期望与方差分别为

$$E\xi = \frac{n}{n-2}, \quad n > 2,$$
$$\mathrm{Var}\xi = \frac{n^2(2m+2n-4)}{m(n-2)^2(n-4)}, \; n > 4.$$

证明：略.

定理 1.3.10 设$X_1, \ldots, X_n \overset{\text{iid}}{\sim} N(0, \sigma^2)$，$Q_i$，$i = 1, \cdots, k$是秩为$n_i$的关于$X_1, \ldots, X_n$的非负定二次型，且$Q_1 + \cdots + Q_k = \sum_{i=1}^n X_i^2$, $n_1 + \cdots + n_k = n$. 则

$$F_{ij} = \frac{Q_i/n_i}{Q_j/n_j} \sim F(n_i, n_j).$$

证明：此结论容易由Cochran定理和F分布的定义求得，略.

我们注意到，定理1.3.10在方差分析(Analysis of Variance, 简记为ANOVA)中非常有用.

定义 1.3.6 (非中心F分布) 设$\xi \sim \chi^2(m, \delta)$, $\eta \sim \chi^2(n)$, $\delta \neq 0$, 且ξ, η独立, 则称$F = \frac{\xi/m}{\eta/n}$所服从的分布为非中心F分布，其自由度为(m, n), 非中心参数为δ.

由于非中心F分布的PDF比较复杂，且不常用，故我们就不再写出了，如有兴趣，请参见

复旦编《概率论》第二册P38. 我们在这里仅给出其期望与方差：

$$E(F)=\frac{n(m+\delta)}{m(n-2)},\ n>2,$$

$$\mathrm{Var}(F)=\frac{2n^2}{m^2(n-2)^2(n-4)}\left[(m+\delta)^2+(n-2)(m+2\delta)\right],\ n>4.$$

结合前面讲的关于正态总体的有关性质，我们有如下几个非常有用的定理.

定理 1.3.11 设$X_1,\dots,X_n\overset{\text{iid}}{\sim}N(\mu,\sigma^2)$，$\bar{X},S_n^2$分别是样本均值与样本方差，则

$$T=\frac{\sqrt{n}(\bar{X}-\mu)}{S_n}\sim t(n-1). \tag{1.3.19}$$

定理 1.3.12 设$X_1,\dots,X_m\overset{\text{iid}}{\sim}N(\mu_1,\sigma_1^2)$，$Y_1,\dots,Y_n\overset{\text{iid}}{\sim}N(\mu_2,\sigma_2^2)$，且全样本独立，则

$$F=\frac{S_{1m}^2/\sigma_1^2}{S_{2n}^2/\sigma_2^2}\sim F(m-1,n-2), \tag{1.3.20}$$

其中，$S_{1m}^2=\frac{1}{m-1}\sum_{i=1}^m(X_i-\bar{X})^2$，$S_{2n}^2=\frac{1}{n-1}\sum_{i=1}^n(Y_i-\bar{Y})^2$.

特别地，如$\sigma_1=\sigma_2$，则$F=\frac{S_{1m}^2}{S_{2n}^2}\sim F(m-1,n-1)$.

定理 1.3.13 设$X_1,\dots,X_m\overset{\text{iid}}{\sim}N(\mu_1,\sigma^2)$，$Y_1,\dots,Y_n\overset{\text{iid}}{\sim}N(\mu_2,\sigma^2)$，且全样本独立，则

$$T=\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{(m-1)S_{1m}^2+(n-1)S_{2n}^2}}\sqrt{\frac{mn(m+n-2)}{m+n}}\sim t(m+n-2). \tag{1.3.21}$$

§1.3.4 几个常用的分布族

在上一节中，我们给出了几个有用的抽样分布，从其定义可以看出，它们的PDF中仅有自由度，而没有未知参数. 在本节，我们将再给出几个常用的分布族，如Gamma分布族、Beta分布族和指数型分布族.

一、Γ分布族

定义 1.3.7 PDF为

$$f(x;\alpha,\lambda)=\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x},\ x>0 \tag{1.3.22}$$

的分布称为Γ分布，记为$\Gamma(\alpha,\lambda)$，其中$\alpha>0$为形状参数(shape parameter), $\lambda>0$称为尺度参数(scale parameter).

关于$\Gamma(\alpha,\lambda)$，我们注意到

- $\Gamma(\frac{n}{2},\frac{1}{2})=\chi^2(n)$.
- $\Gamma(1,\lambda)$为常用的指数分布$E(\lambda)$.
- 之所以称$\alpha>0$为形状参数，是由于Γ分布的形状取决于α的大小，见图1.4.

在图1.4中我们分别画出了$(\alpha,\lambda)=(0.5,1),(1,1),\ (1.5,2)$和$(4,2)$的$\Gamma$分布的PDF曲线，从中可以看出：

- 当$\alpha\le 1$时，其严格单调下降；
- 当$1<\alpha\le 2$时，它先上凸后下凸，其拐点为$\frac{\alpha-1}{\lambda}$；
- 当$\alpha>2$时，其先下凸，再上凸，最后又下凸，其峰值在$\frac{\alpha-1}{\lambda}$处达到.

另外，也容易算得Γ分布的特征函数为

$$\psi(t)=\left(1-\frac{it}{\lambda}\right)^{-\alpha}, \tag{1.3.23}$$

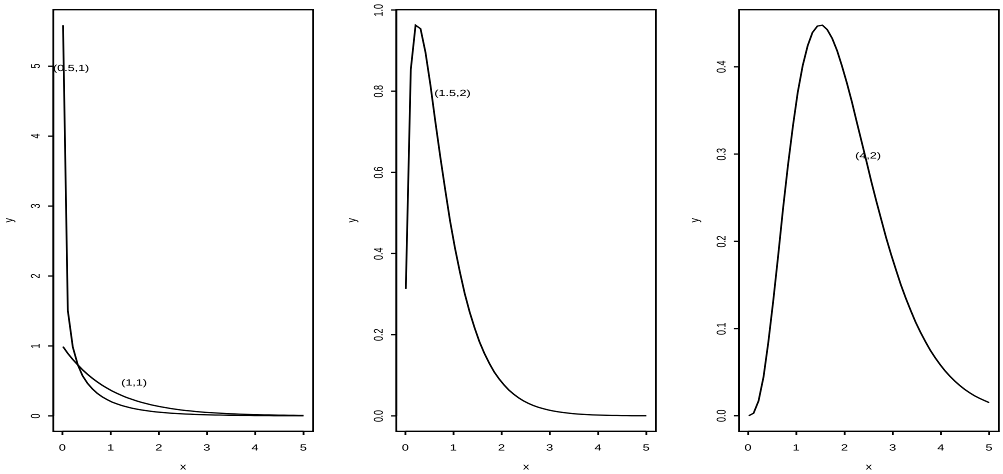

图1.4 Γ分布的概率密度函数曲线

并由此可以求得其k阶矩为

$$E\xi^k=\frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}\frac{1}{\lambda^k}, \tag{1.3.24}$$

特别地，其期望、方差分别为

$$E\xi=\frac{\alpha}{\lambda},\quad \mathrm{Var}\xi=\frac{\alpha}{\lambda^2}.$$

另外，从其特征函数(1.3.23)可以看出，在相同的尺度参数下，Γ分布关于α具有可加性，即下面的结论.

定理 1.3.14 设$\xi\sim\Gamma(\alpha_1,\lambda),\eta\sim\Gamma(\alpha_2,\lambda)$，且$\xi,\eta$相互独立，则

$$\xi+\eta\sim\Gamma(\alpha_1+\alpha_2,\lambda).$$

关于Γ分布的尺度参数，有如下的数乘性质.

定理 1.3.15 设$\xi\sim\Gamma(\alpha\lambda)$，则对于任一实数$k>0$，均有$\eta=\xi/k\sim\Gamma(\alpha,k\lambda)$.

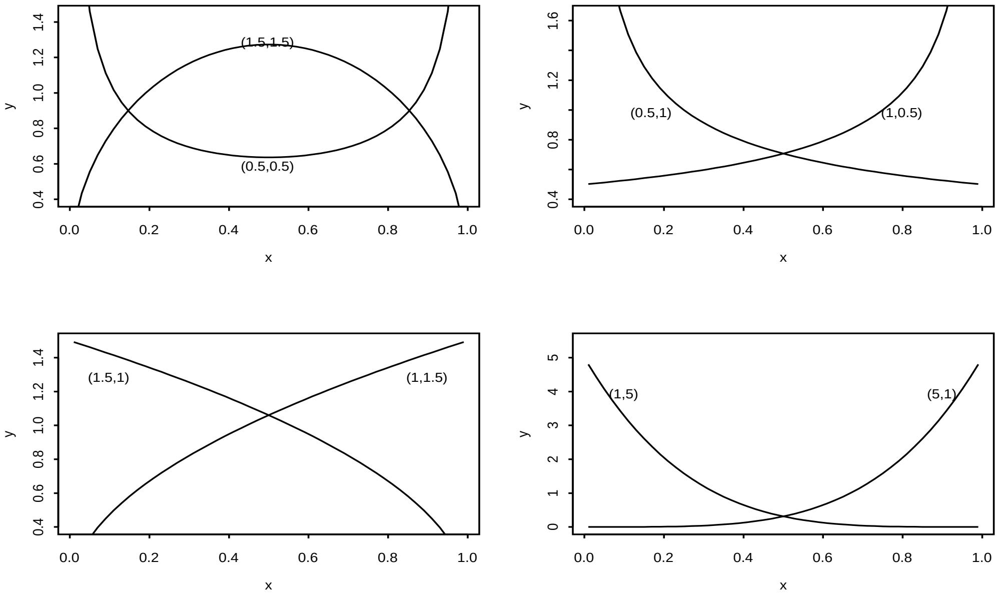

图1.5 β分布的概率密度函数曲线

二、β分布族

定义 1.3.8 定义在$[0,1]$上的具有如下PDF

$$f(x;a,b)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1},\ 0\le x\le 1 \tag{1.3.25}$$

的随机变量ξ称为具有β分布的随机变量，记为$\xi\sim\beta(a,b)$, 其中$a>0,b>0$为两个参数.

当$a=b=1$时，β分布即为$[0,1]$上的均匀分布$U(0,1)$.

关于β分布PDF曲线的形状，请参见图1.5.

关于β分布的数字特征，我们不加证明地如下给出：

$$E\xi=\frac{a}{a+b},\quad \mathrm{Var}\xi=\frac{ab}{(a+b)^2(a+b+1)}.$$

例 1.3.3 设$X_1,\ldots,X_n\overset{\text{iid}}{\sim}U(0,1)$，则由例1.2.12中的(1.2.15)式可知，其第k个次序统计量的PDF为

$$f_k(x)=\frac{n!}{(k-1)!(n-k)!}x^{k-1}(1-x)^{n-k},\ 0\le x\le 1,$$

与(1.3.25)相比知，这即是$\beta(k,n-k+1)$的PDF.

三、Z-分布族

这是R. A. Fisher提出的，它包含着一些著名的分布，有兴趣的同学请参见茆诗松、王静龙(1990).

四、指数型分布族

定义 1.3.9 设$\mathcal{F}=\{f(x;\theta):\theta\in\Theta\}$是某参数分布族，如果$f(x;\theta)$可以表示成

$$f(x;\theta)=c(\theta)\exp\left\{\sum_{i=1}^{k}c_j(\theta)T_j(x)\right\}h(x), \tag{1.3.26}$$

则称此分布族为指数型分布族，其中k为正整数，$c(\theta)>0, h(x)>0$.

为了方便，人们常把上面的指数型密度函数写面如下的典则形式(canonical form):

$$f(x;\eta)=\exp\left\{\sum_{i=1}^{k}\eta_i T_i(x)-a(\eta)\right\}h(x), \tag{1.3.27}$$

其中η称为典则参数.

指数型分布族在统计中有着重要的应用，我们常用的多个分布均为指数型分布族，下面看几个例子.

例 1.3.4 对于正态分布族，因为

$$\begin{aligned}
f(x;\mu,\sigma^2) &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}\\
&= \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{\mu^2}{2\sigma^2}\right\}\exp\left\{-\frac{x^2}{2\sigma^2}+\frac{\mu x}{\sigma^2}\right\},
\end{aligned}$$

如取

$$c(\mu,\sigma)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{\mu^2}{2\sigma^2}\right\}$$

$$c_1(\mu,\sigma)=-\frac{1}{2\sigma^2},\ c_2(\mu,\sigma)=\frac{\mu}{\sigma^2},$$

$$T_1(x)=x^2,\ T_2(x)=x,\ h(x)=1,$$

则正态分布满足定义1.3.9的要求. □

例 1.3.5 对于二项分布$B(n,p)$，由于其分布函数可以写成如下形式:

$$\begin{aligned}
f(x) &= \begin{pmatrix} n \\ x \end{pmatrix} p^x(1-p)^{n-x}\\
&= \begin{pmatrix} n \\ x \end{pmatrix}(1-p)^n\left(\frac{p}{1-p}\right)^x=\begin{pmatrix} n \\ x \end{pmatrix}(1-p)^n\exp\left\{x\ln\frac{p}{1-p}\right\},
\end{aligned}$$

如取

$$x(p)=(1-p)^n,\ c_1(p)=\ln\frac{p}{1-p},\ T_1(x)=x,\ h(x)=\begin{pmatrix} n \\ x \end{pmatrix},$$

则它满足指数型分布族的定义.

注 1.3.5 We note that $c_1(p)=\ln(p/(1-p))$ is called *logit* function. If there are n independent Binomial random variables: $X_i\sim B(n_i,p_i)$, and a covariable Z such that

$$\eta_i=\alpha+\beta Z_i, \tag{1.3.28}$$

this model is called a Logit model, where $\eta_i = \ln(p_i/(1-p_i))$.

例 1.3.6 对于均匀分布族$\{U(0,\theta):\theta>0\}$, 虽然它的PDF为$\frac{1}{\theta}I_{x\in[0,\theta]}$的形式，但它并不是指数型分布，其原因是由于它的支撑集与参数有关.

注 1.3.6 一个PDF $f(x;\theta)$的支撑集(support set)被定义为

$$S=\{x: f(x;\theta)>0\}.$$

注 1.3.7 在指数型分布族的定义1.3.9中，我们必须要求其支撑集与参数无关.

例 1.3.7 对于双参数指数分布，其PDF为

$$f(x;\mu,\sigma)=\frac{1}{\sigma}\exp\left\{-\frac{x-\mu}{\sigma}\right\},\ x>\mu,\ \mu\in R,\ \sigma>0,$$

由于其支撑集也与未知参数μ有关，故它也不是指数型分布族. 但如果μ已知，则它就是单参数指数分布, 属于指数型分布族.

前面我们给出了几个属于指数型分布族的例子，大家可以自己验证：Γ,β, Poisson, 负二项分布等均是指数型分布.

另外，关于指数型分布族，我们有如下一个非常有用的Stein等式：

定理 1.3.16 (Stein 恒等式) 设X的PDF具有(1.3.27)典则形式，且其支撑集为$(-\infty,\infty)$. 设g是满足条件$E|g'(X)|<\infty$的任一可微函数，则

$$E\left\{\left[\frac{h'(X)}{h(X)}+\sum_{i=1}^{k}\eta_i T_i'(X)\right]g(X)\right\}=-Eg'(X). \tag{1.3.28}$$

例 1.3.8 (正态情况下的Stein恒等式) 设$X\sim N(\mu,\sigma^2)$, 则由(1.3.28)知

$$E[g(X)(X-\mu)]=\sigma^2 Eg'(X).$$

如取$g(x)=1$，则得$EX=\mu$. 如取$g(x)=x$，则得$EX^2=\sigma^2+\mu^2$.

§1.4 充分统计量

统计量的引入是为了简化样本的繁杂，但在利用统计量进行统计推断时，一个自然的问题是：我们所用的统计量是否把样本中关于感兴趣问题的信息全部吸收进来了？如果某几个统计量包含了样本中关于感兴趣问题的所有信息，则这几个统计量对将来的统计推断会非常有用，这就是充分统计量的概念，它是R. A. Fisher于1922年正式提出的，而其思想则源于它与天文学家Eddington的有关估计标准差的争论中. 设$X_1,\ldots,X_n$为来自$N(\mu,\sigma^2)$的iid样本，现要估计σ. Fisher主张用样本标准差S_n，而Eddington则主张用如下的平均绝对偏差

$$d=\sqrt{\frac{\pi}{2}}\frac{1}{n}\sum_{i=1}^{n}|X_i-\bar{X}|. \tag{1.4.1}$$

Fisher认为“在S_n中包含了样本中有关σ的全部信息，而d则否.” 这就是充分统计量的思想.

注 1.4.1 虽然(1.4.1)式的统计量d并不是后面所讲的充分统计量，即它没有把样本中关

于σ的全部信息都包含进来，但如果用的准则来衡量的话，此统计量具有非常好的优良性质，如稳健性(Robustness)等.

在给出充分统计量的严格定义之前，我们先看一个例子.

例 1.4.1 某工厂为了解其产品的不合格品率p，检验员抽查了10件产品. 其检查结果是: 除前两件是不合格品外(记$X_1 = X_2 = 1$)，其它的都是合格品(记为$X_i = 0,\ i = 3, \cdots, 10$). 当厂长问及检验结果时，检验员可以有如下两种回答:

- 10件中有两件不合格($\sum_{i=1}^{10} X_i = 2$);
- 前两件不合格($X_1 = X_2 = 1$).

这两种回答所用的统计量有所不同，当然所包含的信息也有所不同. 显然，第二种回答是不能令人满意的，由于它没有包含样本中的关于p的所有信息，而第一种回答则包含了样本中关于p的所有信息. □

由于样本或统计量中关于参数的信息是通过分布来反映的，且我们希望在给定充分统计量后，样本中已没有任何有用信息可用，于是，我们可以用在给定一个统计量T后的样本的条件分布是否与参数有关来反映这个统计是否是充分统计量的定义.

例 1.4.2 (续例1.4.1) 设$X_1, \ldots, X_n$是来自$b(1,p)$的n个iid样本，其中$p \in (0,1), n > 2$. 我们考虑的统计量为$T_1 = \sum_{i=1}^n X_i, T_2 = X_1 + X_2$. 由于此时的样本分布为

$$P\{X_1 = x_1, \cdots, X_n = x_n\} = p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i},$$

且$T_1 \sim B(n,p)$，故在给定$T_1 = t$下样本的条件分布为

$$\begin{aligned} & P\{X_1 = x_1, \cdots, X_n = x_n | T_1 = t\} \\ = & \frac{P\{X_1 = x_1, \cdots, X_{n-1} = x_{n-1}, X_n = t - \sum_{i=1}^{n-1} x_i\}}{P\{T_1 = t\}} \\ = & \frac{p^t(1-p)^{n-t}}{\begin{pmatrix} n \\ t \end{pmatrix} p^t(1-p)^{n-t}} = \begin{pmatrix} n \\ t \end{pmatrix}^{-1}, \end{aligned}$$

它与参数p无关，即这个条件分布中已没有关于p的任何信息了，也就是说，样本中关于p的所有信息都在统计量T_1中了. 但对于统计量T_2而言，由于$T_2 \sim B(2,p)$，故

$$\begin{aligned} & P\{X_1 = x_1, \cdots, X_n = x_n | T_2 = t\} \\ = & \frac{P\{X_1 = x_1, X_2 = t - x_1, X_3 = x_3, \cdots, X_n = x_n\}}{P\{T_2 = t\}} \\ = & \frac{p^{t+\sum_{i=3}^n x_i}(1-p)^{n-t-\sum_{i=3}^n x_i}}{\begin{pmatrix} 2 \\ t \end{pmatrix} p^t(1-p)^{2-t}} \end{aligned}$$

与p有关，故说明T_2没有把样本中关于p的所有信息都包含进来. □

定义 1.4.1 (充分统计量(Sufficient Statistics)) 对于某分布族$\mathcal{F} = \{F_\theta(x) : \theta \in \Theta\}$，$\forall\ F \in \mathcal{F}$，设$X_1, \ldots, X_n$是来自$F$的样本，$T = T(X_1, \cdots, X_n)$是一统计量. 如在给定$T = t$下，

样本$(X_1,\ldots,X_n)$的条件概率分布与总体分布F或参数θ无关，则称统计量T是此分布族$\mathcal{F}$的充分统计量，也称统计量T是参数θ的充分统计量.

关于充分统计量，我们有如下几点值得注意的地方或结论：

- 对于连续随机变量，其概率分布为PDF；对于离散随机变量，其概率分布即为CDF.
- 统计量T可以是向量，也可以是实量.
- 如果统计量T是参数θ的充分统计量，且$S(t)$是单值可逆的，则$S(T)$也是θ的充分统计量.

现在看几个例子.

例 1.4.3 设$X_1,\ldots,X_n$是来自几何分布的iid样本，即

$$P\{X=x\}=\theta(1-\theta)^x,\ x=0,1,\cdots,$$

其中$0<\theta<1$，则$T=\sum_{i=1}^n X_i$是充分统计量.

解 由概率论知识知，统计量$T=\sum_{i=1}^n X_i$服从Pascal 分布，即

$$P\{T=t\}=\begin{pmatrix} t+n-1 \\ n-1 \end{pmatrix}\theta^n(1-\theta)^t,\ t=0,1,\cdots.$$

故，当$T=t$时，样本的条件概率分布为

$$\begin{aligned}
& P\{X_1=x_1,\cdots,X_n=x_n|T=t\} \\
= & \frac{P\{X_1=x_1,\cdots,X_n=x_n,T=t\}}{P\{T=t\}} \\
= & \frac{P\{X_1=x_1\}\cdots P\{X_{n-1}=x_{n-1}\}P\{X_n=t-\sum_{i=1}^{n-1}x_i\}}{P\{T=t\}} \\
= & \frac{\theta(1-\theta)^{x_1}\cdots\theta(1-\theta)^{x_{n-1}}\theta(1-\theta)^{t-\sum_{i=1}^{n-1}x_i}}{\begin{pmatrix} t+n-1 \\ n-1 \end{pmatrix}\theta^n(1-\theta)^t}=\left[\begin{pmatrix} t+n-1 \\ n-1 \end{pmatrix}\right]^{-1},
\end{aligned}$$

它与θ无关，故T是θ的充分统计量. □

例 1.4.4 设$X_1,\ldots,X_n$是来自Poisson分布$P(\lambda)$的iid样本，则$T=\sum_{i=1}^n X_i$是λ的充分统计量.

解 由Poisson分布的可加性知，$T=\sum_{i=1}^n X_i\sim P(n\lambda)$，故当$T=t$时，样本的条件概率分布为

$$\begin{aligned}
& P\{X_1=x_1,\ldots,X_n=x_n|T=t\} \\
= & \frac{P\{X_1=x_1\}\cdots P\{X_{n-1}=x_{n-1}\}P\{X_n=t-\sum_{i=1}^{n-1}x_i\}}{P\{T=t\}} \\
= & \frac{\left(\prod_{j=1}^{n-1}\frac{\lambda^{x_j}e^{-\lambda}}{x_j!}\right)\frac{\lambda^{t-\sum_{i=1}^{n-1}x_i}e^{-\lambda}}{(t-\sum_{i=1}^{n-1}x_i)!}}{\frac{(n\lambda)^t}{t!}e^{-n\lambda}} \\
= & \frac{t!}{\prod_{j=1}^{n-1}x_j!(t-\sum_{i=1}^{n-1}x_i)!}\frac{1}{n!},
\end{aligned}$$

与λ无关，故T是充分统计量. □

例 1.4.5 设$X_1,\ldots,X_n$为来自正态总体$N(\mu,1)$的iid样本，则$T=\sum_{i=1}^n X_i$是μ的充分统计量.

解 易知$T\sim N(n\mu,n)$，为求当$T=t$下，样本的条件分布，我们做如下变换：

$$\begin{cases} x_1=x_1,\\ x_2=x_2,\\ \cdots,\\ x_{n-1}=x_{n-1},\\ t=x_1+\cdots+x_n. \end{cases}$$

于是，可以由$(X_1,\ldots,X_n)$的联合PDF求得$(X_1,\ldots,X_{n-1},T)$的联合PDF为

$$f(x_1,\ldots,x_{n-1},t)=(\sqrt{2\pi})^{-n}\exp\left\{-\frac{\sum_{i=1}^{n-1}(x_i-\mu)^2}{2}-\frac{(t-x_1-\cdots-x_{n-1}-\mu)^2}{2}\right\},$$

故当$T=t$时，$X_1,\ldots,X_{n-1}$的条件PDF为

$$\begin{aligned} f(x_1,\ldots,x_{n-1}|T=t)&=\frac{f(x_1,\ldots,x_{n-1},t)}{f_T(t)}\\ &=\frac{(2\pi)^{-n/2}\exp\left\{-\frac{\sum_{i=1}^{n-1}(x_i-\mu)^2}{2}-\frac{[(t-\sum_{i=1}^{n-1}x_i)-\mu]^2}{2}\right\}}{\sqrt{2\pi n}\exp\{-\frac{(t-n\mu)^2}{2n}\}}, \end{aligned}$$

化简后，上式与μ无关，故T是μ的充分统计量. □

在例1.4.5中，仅有一个未知参数μ，如果其方差也是未知的，则利用定义来求充分统计量比较困难. 并且从上面的例子也可以看出，想求取充分统计量，就必须先猜测一个统计量，之后再用定义证明之. 显然，这很不便于利用，于是有如下的因子分解定理.

定理 1.4.1 (因子分解定理) *对于参数分布族*

$$\mathcal{F}=\{f_\theta(x):\theta\in\Theta\},$$

设$X_1,\ldots,X_n$为其一组iid样本，T是一统计量，则T是θ的充分统计量的充要条件是：其样本分布$f_\theta(x_1,\ldots,x_n)$可以分解为如下形式:

$$f_\theta(x_1,\ldots,x_n)=g_\theta(T)\cdot h(x_1,\ldots,x_n), \tag{1.4.2}$$

其中$h(x)$不依赖于参数θ.

证明 其离散情形的证明见茆诗松、王静龙(1990)，一般情形证明见陈希孺、倪国策(1988).

这个定理是R. A. Fisher 于上世纪二十年代提出的，它的严格证明及一般形式则是Halmos与Savage 于1949年给出的. 另外，离散时，(1.4.2)中的f表示CDF；连续时它表示PDF.

例 1.4.6 *求均匀分布$U(0,\theta)$的充分统计量.*

解 此时样本的联合PDF为

$$f_\theta(\mathbf{x})=\begin{cases}\theta^{-n}, & \max(x_1,\ldots,x_n)<\theta,\\ 0, & \text{otherwise.}\end{cases} \tag{1.4.3}$$

如取$T=\max(X_1,\ldots,X_n)$，则(1.4.3)式可改写为

$$f_\theta(\mathbf{x})=\frac{1}{\theta^n}I_{\{0<t<\theta\}},$$

则由因子分解定理可知，$T=X_{(n)}$即为θ的充分统计量. □

例 1.4.7 考虑正态总体$N(\mu,\sigma^2)$的充分统计量.

解 此时简单随机样本$X_1,\ldots,X_n$的联合PDF为

$$\begin{aligned}f(\mathbf{x}) &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n(x_i-\mu)^2\right\}\\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n\exp\left\{-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2\right\}\cdot\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n(x_i-\bar{x})^2\right\},\end{aligned}$$

故由因子分解定理可知$(\bar{X},\sum_{i=1}^n(X_i-\bar{X})^2)$是$(\mu,\sigma^2)$的充分统计量. □

例 1.4.8 考虑$\Gamma(\alpha,\lambda)$的充分统计量.

解 此时随机样本$X_1,\ldots,X_n$的联合PDF为

$$f_{\alpha,\lambda}(\mathbf{x})=\frac{\lambda^{n\alpha}}{\Gamma^n(\alpha)}\left(\prod_{i=1}^n x_i\right)^{\alpha-1}\exp\left\{-\lambda\sum_{i=1}^n x_i\right\}.$$

如定义$T(\mathbf{X})=(T_1,T_2)=(\prod_{i=1}^n X_i,\sum_{i=1}^n X_i)$，且$h(\mathbf{x})=1$，$g_{\alpha,\lambda}(t)=\frac{\lambda^{n\alpha}}{\Gamma^n(\alpha)}t_1^{\alpha-1}e^{-\lambda t_2}$，则由因子分解定理知$T$是$(\alpha,\lambda)$的充分统计量.

特别地，当α已知时，可验证$T_2=\sum_{i=1}^n X_i$是λ的充分统计量；当λ已知时，$T_1=\prod_{i=1}^n X_i$是α的充分统计量. □

例 1.4.9 考虑均匀分布族$U\left(-\frac{1}{2}+\theta,\frac{1}{2}+\theta\right)$的充分统计量.

解 此时样本的联合PDF为

$$f_\theta(\mathbf{x})=I_{\{\frac{1}{2}+\theta<x_{(1)}<x_{(n)}<\frac{1}{2}+\theta\}},$$

故由因子分解定理可知，$(X_{(1)},X_{(n)})$是θ的充分统计量. □

例 1.4.10 考虑Bernoulli分布$b(1,p)$的充分统计量.

解 此时随机样本$X_1,\ldots,X_n$的联合CDF为

$$P\{X_1=x_1,\ldots,X_n=x_n\}=p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i},$$

于是，由因子分解定理可知，如下统计量

$$\begin{aligned}&T_1=(X_1,\ldots,X_n),\\ &T_2=(X_1+X_2,X_3,\ldots,X_n),\\ &\cdots,\\ &T_n=X_1+\cdots+X_n,\end{aligned}$$

均是充分统计量. □

由于次序统计量在许多实际问题中都有应用，并且，许多著名统计量也都是次序统计量的函数，故下面我们看一看次序统计量的充分性.

定理 1.4.2 (次序统计量的充分性) 对于分布放$\mathcal{F}$，设$F\in\mathcal{F}$，$X_1,\ldots,X_n$为来自F的样本，只要$X_1,\ldots,X_n$是iid的，则不论$\mathcal{F}$如何，其次序统计量$X_{(1)},\ldots,X_{(n)}$都是充分的.

证明 略.

显然，在例1.4.10的统计量中，T_n最简单，那是否还存在别的简化统计量是充分的吗？

定义 1.4.2 (极小充分统计量) 设S是分布族$\mathcal{F}$的充分统计量，如对$\mathcal{F}$的任一充分统计量T，均存在函数$f(\cdot)$，使得$S=f(T)$，则称S是此分布族$\mathcal{F}$的极小充分统计量.

我们常用的充分统计量基本都是极小的.

习题一

1. 设$X_1,\ldots,X_n$为来自总体分布$F(x)$的iid样本，记其经验分布函数为$F_n(x)$，试证对于任意给定的$x\in R$，有

$$E(F_n(x))=F(x),\quad \mathrm{Var}(F_n(x))=\frac{1}{n}F(x)(1-F(x)).$$

2. 设$X_1,\ldots,X_n$为来自总体分布$F(x)$的iid样本，记其经验分布函数为$F_n(x)$，试证对于任意给定的$x\in R$，$F_n(x)\xrightarrow{P}F(x)$.

3. 设$X_1,\ldots,X_{n+1}$为来自总体X的iid样本，试证明

$${S^*_{n+1}}^2=\frac{n}{n+1}\left[{S^*_n}^2+\frac{1}{n+1}(X_{n+1}-\bar{X}_n)^2\right].$$

4. 设$X_1,\ldots,X_8$为来自$N(0,1)$的iid样本，$Y=(X_1+\cdots+X_4)^2+(X_5+\cdots+X_8)^2$，试求常数$c$，使$cY$服从$\chi^2$分布.

5. 设$X_1,\ldots,X_4$为来自$N(0,1)$的iid样本，$X=a(X_1-2X_2)^2+b(3X_3-4X_4)^2$，试求$a,b$，使$X$服从$\chi^2$分布.

6. 设$X_1,\ldots,X_n$为来自$N(\mu,\sigma^2)$的iid样本，请问当k为多少时，

$$P\{\bar{X}>\mu+kS_n\}=0.95?$$

7. 设$X_1,\ldots,X_{n+1}$为来自$N(\mu,\sigma^2)$的iid样本，试求统计量

$$T=\frac{X_{n+1}-\bar{X}_n}{S^*_n}\sqrt{\frac{n-1}{n+1}}$$

的抽样分布.

8. 设$X_1,\ldots,X_{m+n}$为来自$N(0,\sigma^2)$的iid样本，试求统计量

$$Z=\sqrt{\frac{m}{n}}\frac{\sum_{i=1}^n X_i}{\sqrt{\sum^{n+m} i=n+1 X_i^2}},\quad F=\frac{m}{n}\frac{\sum_{i=1}^n X_i^2}{\sum_{i=n+1}^{n+m}X_i^2}$$

的抽样分布.

9. 设$X_1,\ldots,X_4$为来自$N(0,\sigma^2)$的iid样本，$U=\frac{\sqrt{3}X_1}{\sqrt{X_2^2+X_3^2+X_4^2}}$，试求$EU$, $\mathrm{Var}U$.

10. 设$X_1,\ldots,X_{n+1}$为来自$N(\mu,\sigma^2)$的iid样本，记$U_i=X_i-\frac{1}{n+1}\sum_{i=1}^{n+1}X_i, i=1,\cdots,n+1$. 试求$V_i$的分布.

11. 设$X_1,\ldots,X_m$为来自$N(\mu_1,\sigma^2)$的iid样本，$Y_1,\ldots,Y_n$为来自$N(\mu_2,\sigma^2)$的iid样本，且全样本独立. 又设a,b为两个常数，$Z=\frac{a(\bar{X}-\mu_1)+b(\bar{Y}-\mu_2)}{\sqrt{(m-1)S_{1m}^2+(n-1)S_{2n}^2}}$，试求常数$c$，使得$cZ$服从$t$分布，并给出其自由度.

12. 设随机变量X服从自由度为n,m的F分布，试求$Y=1/X$的分布，并试明

$$Z=\frac{nX}{m+nX}\sim\beta(n/2,m/2).$$

13. 设$X_1\sim\beta(a_1,b_1)$, $X_2\sim\beta(a_2,b_2)$, $a_1=a_2+b_2$，且X_1与X_2独立，则$Y=X_1X_2\sim\beta(a_2,b_1+b_2)$.

14. 设$X\sim N(0,1),T\sim t(n)$，则存在一个正数t_0，使得

$$P\{|T|\geq t_0\}\geq P\{|X|\geq t_0\}.$$

15. 检验下列分布族是否是指数型分布族：

(1) Poisson分布族；

(2) 双参数指数分布族；

(3) Γ分布族$\Gamma(\alpha,\lambda)$;

(4) Cauchy分布族：$\left\{f(x,\lambda)=\frac{\lambda}{\pi(\lambda^2+x^2)}:\lambda>0\right\}$.

16. 设$X_1,\ldots,X_n$为来自PDF为

$$f(x,\theta)=\theta^{-1}\exp\{-|x|/\theta\},\ x\in R,\theta>0$$

的总体的iid样本，证明$T=\sum_{i=1}^n|X_i|$是θ的充分统计量.

17. 设$X_1,\ldots,X_n$是来自均匀分布族$\{U(\theta_1,\theta_2):-\infty<\theta_1<\theta_2\infty\}$的iid样本，证明：$(X_{(1)},X_{(n)})$是$(\theta_1,\theta_2)$ 的充分统计量.

18. 设$Y_i\sim N(\alpha+\beta x_i,\sigma^2)$, $i=1,\cdots,n$, 且诸Y_i相互独立，其中$\alpha,\beta\in R$, $\sigma>0$是未知参数，$x_1,\ldots,x_n$为已知的常数. 试证明统计量

$$\left(\sum_{i=1}^n Y_i,\sum_{i=1}^n x_iY_i,\sum_{i=1}^n Y_i^2\right)$$

是(α,β,σ^2)的充分统计量.

19. 设$X_1,\ldots,X_n$为来自PDF为

$$f(x,\theta)=\theta/x^2,\ 0<\theta<x<\infty$$

的总体的iid样本，试证明$X_{(1)}$是充分统计量.

20. 设$X_1,\ldots,X_n$为来自两点分布$b(1,p)$的iid样本，试写出$X_1,\ldots,X_n$的联合分布，并指出$X_1+X_2,X_{(n)},X_n+2p,(X_n-X_1)^2$中哪些是统计量，为什么？

21. 设$X_1,\ldots,X_n$为来自概率密度函数为

$$f(x)=\begin{cases}2x, & 0<x<1,\\ 0, & \text{其它},\end{cases}$$

的总体的iid样本，试求$X_{(1)},X_{(n)}$的PDF.

22. 设$X_1,\ldots,X_n$为来自具有PDF为$f(x)$的总体的iid样本，试证明：样本极差$R = X_{(n)} - X_{(1)}$的PDF为

$$f_R(x) = n(n-1)\int_{-\infty}^{\infty} [F(t+x) - F(t)]^{n-2} f(x+t)f(t)\mathrm{d}t,$$

其中F为总体的CDF.

23. 设$X_1,\ldots,X_m$为来自$N(\mu,\sigma_i^2)$的iid样本，$Y_1,\ldots,Y_n$为来自$N(\mu,\sigma_2^2)$的iid样本，且全样本独立，记$\alpha_1 = S_{1m}^2/(S_{1m}^2+S_{2n}^2), \alpha_2 = S_{2n}^2/(S_{1m}^2+S_{2n}^2)$. 试求$Y = \alpha_1\bar{X} + \alpha_2\bar{Y}$的期望.

24. 设$X \sim B(n,p)$, 对$k \geq 1$，请证明

$$P\{X \geq k\} = \sum_{i=k}^{n} C_n^i p^i (1-p)^{n-i} = kC_n^k \int_0^p t^{k-1}(1-t)^{n-k}\mathrm{d}t.$$

25. 设$X_1,\ldots,X_n$为来自总体分布为F的iid样本，$X_{(1)} \leq \cdots \leq X_{(n)}$为其次序统计量. 如果$F(x)$连续，则

$$E[F(X_{(i)}] = \frac{i}{n+1}, \quad \mathrm{Var}[F(X_{(i)}] = \frac{i(n+1-i)}{(n+1)^2(n+2)}.$$

26. 设$X_1,\ldots,X_n$是n个相互独立的随机变量，且$X_i \sim N(0,\sigma_i^2)$. 记

$$Y = \frac{\sum_{i=1}^n X_i/\sigma_i}{\sum_{i=1}^n 1/\sum_{i=1}^n \sigma_i^2}, \quad Z = \sum_{i=1}^n \frac{(X_i - Y)^2}{\sigma_i^2},$$

试证明$Z \sim \chi^2(n-1)$.

27. 设$X_1,\ldots,X_n$为来自正态总体$N(\mu,\sigma^2)$的iid样本，记

$$\tau = \frac{X_1 - \bar{X}}{S}\sqrt{\frac{n}{n-1}},$$

试证明$t = \frac{\tau\sqrt{n-2}}{\sqrt{n-1-\tau^2}} \sim t(n-2)$.

28. 设$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \cdots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$为来自二维正态$N(\mu,\Sigma)$的iid样本，其中$\mu = (\mu_1,\mu_2)'$, $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$. 定义二样本间的相关系数为

$$r(X,Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\left[\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2\right]^{1/2}},$$

试证明：如果$\rho = 0$, 则统计量

$$t = \sqrt{n-2}\frac{r(X,Y)}{\sqrt{1 - r^2(X,Y)}} \sim t(n-2).$$

29. 如果X_i服从非中心$\chi^2(n_i,\delta_i)$分布$(i = 1,\cdots,k)$，且相互独立，则

$$\sum_{i=1}^k X_i \sim \chi^2\left(\sum_{i=1}^k n_i, \sum_{i=1}^k \delta_i\right).$$

30. 证明服从非中心t分布的随机变量的平方服从非中心F分布.

第2章　点估计

在上一章中，我们主要讲述了统计量的几个常用抽样分布及充分统计量. 我们回想一下，引进统计量的目的在于对感兴趣的问题进行统计推断. 而在实际中，我们感兴趣的问题多是以分布族中的未知参数出现的. 当然，在所考虑的分布族中，也可能会有一些未知参数并不是我们感兴趣的，我们称这样的参数为讨厌参数. 从本章开始，我们将讨论感兴趣参数的估计和检验问题.

§2.1　引言

在参数统计中，人们主要从三个方向进行研究：一是参数的估计问题；二是参数的假设检验；三是统计量的抽样分布. 几个常用的抽样分布我们已在上一章介绍过，本章将开始介绍参数的估计问题. 参数估计分点估计(point estimation)和区间估计(interval estimation)，本章将介绍几个常用的点估计方法，如无偏估计(Unbiased Estimation, 简记为UE), 一致最小方差无偏估计(Uniformly Minimum Unbiased Estimation, 简记为UMVUE), 极大似然估计(Maximum Likelihood Estimation, 简记为MLE)等. 之所以引入这些估计方法，是由于我们采取了不同的估计准则，于是，这些准则也将在本章加以介绍.

§2.2　矩估计

矩估计(Moment Estimation)是K. Pearson 于1894年基于某些生物学方面的数据并非来自正态这一发现而提出的，这种估计方法在实际中有着广泛的应用.

定义 2.2.1　对于样本$X_1,\dots,X_n$及任意一正整数k，我们称

$$a_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k,\ \ m_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k, \tag{2.2.1}$$

为样本k阶原点矩和k阶中心矩，其中$\bar{X} = a_1$为样本均值.

从下定义2.2.1可以看出，$m_1 = 0,\ m_2 = \frac{n-1}{n}S_n^2 = \frac{1}{n}(\sum_{i=1}^n X_i^2 - n\bar{X}^2) = a_2 - a_1^2$. 另外，如打开$m_k$成多项式的形式，则知道，样本$k$阶中心矩$m_k$可以由样本$k$阶以下的原点矩线性表出. 同样，样本原点矩$a_k$也可以由样本$k$阶如下的中心矩线性表出.

我们在概率论已经讲过一个随机变量的k阶原点矩和中心矩，为与样本矩做一比较，我们称总体X的k阶原点矩和中心矩分别为

$$\mu_k = EX^k,\ \nu_k = E(X-\mu_1)^k. \tag{2.2.2}$$

显然，样本矩不依赖于总体中的参数，它们是统计量；而总体矩则与分布中的未知参数有关. 另外，由大数定律和中心极限定理可知，样本矩是总体矩的一个很好的估计. 于是，我们可以利用上述想法来估计总体的未知参数，这就是矩估计的基本思想，其具体过程如下：

设$\{F(x,\theta):\theta\in\Theta\}$是总体$X$的分布族，$\theta = (\theta_1,\dots,\theta_m)$是我们感兴趣的有待估计的未知

参数，且假设总体的k阶矩μ_k存在有限. 如令

$$\mu_k(\theta) = a_k,\ k = 1, \cdots, m, \tag{2.2.3}$$

则得到了一个含有m个未知参数的方程组. 如由此方程组可求得θ_i, 记为$\hat{\theta}_i$, $i = 1, \cdots, m$，则称之为θ_i的矩估计.

关于矩估计，我们注意到：

- 在(2.2.3)中，我们也可以通过用样本k阶中心矩作为总体k中心矩的估计而建立相应的方程组. 这样求得的参数估计也称为矩估计. 当然，这两种方法得到的估计可能不相同.
- 方程组(2.2.3)的解是否存在且惟一，要取决于具体问题.
- 在用矩方法求参数估计时，我们尽可能用低阶矩.

下面看几个例子.

例 2.2.1 考虑总体均值与方差的矩估计

解 设$X_1, \ldots, X_n$是一组简单随机样本，且总体二阶矩存在，记$\mu = EX_1, \sigma^2 = \mathrm{Var}X_1$, 则由矩估计法知，其估计方程为

$$\begin{cases} \hat{\mu} = a_1 = \frac{1}{n}\sum_{i=1}^n X_i, \\ \hat{\mu}_2 = \hat{\mu}^2 + \hat{\sigma}^2 = a_2 = \frac{1}{n}\sum_{i=1}^n X_i^2. \end{cases}$$

由此可求得总体均值与方差的矩估计为

$$\begin{cases} \hat{\mu} = \bar{X}, \\ \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2 = \frac{n-1}{n} S_n^2. \end{cases}$$

□

由此例可以看出，总体均值的矩估计是样本均值，而总体方差的矩估计却不是样本方差，而是样本方差的$\frac{n-1}{n}$倍. 当然，当n较大时，这个系数与1非常接近. 为了以后的方便，我们称

$$S_n^{*2} = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2 \tag{2.2.4}$$

为修正的样本方差.

另外，从这个例子还可以看出，上述估计并不要求总体分布具有什么给定的形式.

例 2.2.2 考虑Poisson分布$P(\lambda)$的总体均值的矩估计.

解 对于Poisson分布$P(\lambda)$，由于λ既是总体均值，又是总体方差，故我们有两个矩估计，它们是

$$\hat{\lambda} = a_1 = \bar{X},$$
$$\hat{\lambda} = m_2 = S_n^{*2}.$$

□

由此例可以看出，矩估计不是惟一的. 此时至于选取哪一个作为最终的估计，就要看你选用的准则是什么了.

例 2.2.3 考虑均匀分布$U(0, \theta)$中参数θ的矩估计.

解 因为此时总体期望为$\mu = EX_1 = \int_0^\theta x\frac{1}{\theta}\mathrm{d}x = \frac{\theta}{2}$，故由矩估计法知

$$\hat{\theta} = 2\bar{X}.$$

□

例 2.2.4 *考虑均匀分布$U(\alpha, \beta)$中α, β的矩估计.*

解 因为总体期望与方差分别为

$$\mu = \int_\alpha^\beta \frac{x}{\beta-\alpha}\mathrm{d}x = \frac{\alpha+\beta}{2},$$

$$\nu_2 = \int_\alpha^\beta \frac{1}{\beta-\alpha}\left(x - \frac{\alpha+\beta}{2}\right)^2 \mathrm{d}x = \frac{(\beta-\alpha)^2}{12},$$

故由矩估计法原理有如下的估计方程

$$\begin{cases} \hat{\mu} = \bar{X}, \\ \hat{\nu}_2 = m_2, \end{cases}$$

于是，可求得其矩估计为

$$\begin{cases} \hat{\alpha} = \bar{X} - \sqrt{2S_n^{*2}}, \\ \hat{\beta} = \bar{X} + \sqrt{2S_n^{*2}}. \end{cases}$$

□

关于矩估计的优良性质，我们将在后面介绍.

§2.3 极大似然估计与EM算法

极大似然估计最早是由德国数学家Gauss 在1821 年针对正态分布提出的，之后Fisher于1922年提出了一般分布下的极大似然. 本节将给出极大似然估计的定义及求取某些复杂情况下MLE的一种有效算法—EM算法.

§2.3.1 极大似然估计

在讲极大似然估计之前，我们先看一下其基本思想.

例 2.3.1 *假设在一个罐中放着许多白球和黑球，且二者数目之比为1:3，但不知哪种颜色球多. 现考虑通过返回抽样方法估计罐中黑球数目的比例p.*

解 假设从罐中返回抽样n次，则由概率论知识知道，其中包括黑球个数X的概率分布是二项分布$B(n, p)$，即

$$P(x, p) = P\{X = x|p\} = \begin{pmatrix} n \\ x \end{pmatrix} p^x(1-p)^{n-x}. \tag{2.3.1}$$

由于我们知道p等于1/4或3/4，故我们的目的在于通过X的大小来估计p. 为此，我们先看一个$n = 3$的简单情况. 此时，我们针对不同的x值，计算(2.3.1)式的概率如下：

表2.1 极大似然思想

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

从上表可以看出，当$x=0$时，因为$\frac{27}{64}\gg\frac{1}{64}$，故认为$\hat{p}=\frac{1}{4}$更有可能性发生，因而，$\hat{p}=\frac{1}{4}$比取$\hat{p}=\frac{3}{4}$更合理. 于是，如把$\hat{p}$看成是$x$的函数$\hat{p}(x)$, 则我们得到如下的比较合理的估计：

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

我们在选取上述估计时遵循了如下原理：认为概率最大的事件最有可能发生，即我们选取的估计$\hat{p}(x)$满足如下不等式：

$$P\{X=x|\hat{p}(x)\}\geq P\{X=x|p'\},\ \forall\ p'.$$

这就是我们要讲的极大似然原理的基本思想. □

定义 2.3.1 (似然函数(Likelihood Function) 对于分布族$\{f(x,\theta),\theta\in\Theta\}$，如以$f(\mathbf{x},\theta)$记其$n$个样本的联合概率分布，则对于给定的样本观测值$\mathbf{x}=(x_1,\ldots,x_n)$，我们称$f(\mathbf{x},\theta)$为参数$\theta$的似然函数，简称为似然函数，并记之为

$$L(\theta,\mathbf{x})=f(\mathbf{x},\theta),\ \forall\ \theta\in\Theta, \tag{2.3.2}$$

称$\ln L(\theta,\mathbf{x})$为对数似然函数，记为$l(\theta,\mathbf{x})$或$l(\theta)$.

从上一定义可以看出，似然函数与样本联合概率分布是同一个东西，但二者的含义却不同: 后者是固定参数值为θ下关于样本$\mathbf{x}$的函数，它的取值空间为样本空间$\mathcal{X}$；而似然函数则是固定样本值$\mathbf{x}$下关于参数θ的函数，其在参数空间Θ上取值. 为考察二者的区别，我们不妨把参数θ和样本分别看作"原因"和"结果". 当给定参数后，样本联合分布将告诉我们哪个样本将以多大的概率被观测到；反过来，当有了样本后，似然函数将告诉我们如何最有可能地取参数θ的估计.

"似然"的字面含义就是"看起来象". 下面，我们将利用样本观测值来确定参数看起来最象的值，这就是统计上常说的似然原理.

定义 2.3.2 (MLE) 设$X_1,\ldots,X_n$是来自某概率分布$f(x,\theta)\in\mathcal{F}=\{f(x,\theta):\theta\in\Theta\subseteq R^k\}$的一组样本，如果统计量$\hat{\theta}(\mathbf{X})$满足

$$L(\hat{\theta}(\mathbf{x}),\mathbf{x})=\sup_{\theta\in\Theta}L(\theta,\mathbf{x}), \tag{2.3.3}$$

或等价地满足

$$l(\hat{\theta}(\mathbf{x}),\mathbf{x})=\sup_{\theta\in\Theta}l(\theta,\mathbf{x}), \tag{2.3.4}$$

则称$\hat{\theta}$为θ的MLE.

从上一定义可以看出，求某参数的MLE，就是求(2.3.3)或(2.3.4)的极值问题. 如果似然函

数$L(\theta,\mathbf{x})$关于θ可微，则θ的MLE可以通过求解下面的方程求得：

$$\frac{\partial L(\theta,\mathbf{x})}{\partial\theta_j}=0,\ j=1,\cdots,k, \tag{2.3.5}$$

或等价地有

$$\frac{\partial l(\theta,\mathbf{x})}{\partial\theta_j}=0,\ j=1,\cdots,k. \tag{2.3.6}$$

在统计上，我们称上述方程(2.3.5)和(2.3.6)为似然方程.

例 2.3.2 设$X_1,\dots,X_n$为来自正态总体$N(\mu,\sigma^2)$的iid样本，试求μ和σ^2的MLE.

解 此时的似然函数为

$$L(\mu,\sigma^2;\mathbf{x})=\frac{1}{(\sqrt{2\pi}\sigma)^n}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n(x_i-\mu)^2\right\},$$

其对数似然函数为

$$l(\mu,\sigma^2;\mathbf{x})=-\frac{n}{2}\ln(2\pi)-\frac{n}{2}\ln\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^n(x_i-\mu)^2,$$

两边分别对μ和σ^2求导，再令其为零，则得到如下的似然方程：

$$\begin{cases}\frac{\partial l}{\partial\mu}=\frac{1}{\sigma^2}\sum_{i=1}^n(x_i-\mu)=0,\\ \frac{\partial l}{\partial\sigma^2}=-\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^n(x_i-\mu)^2=0,\end{cases}$$

由此求得其根为

$$\begin{cases}\mu=\bar{x},\\ \sigma^2=s_n^{*2}=\frac{1}{n}\sum_{i=1}^n(x_i-\bar{x})^2,\end{cases}$$

容易验证上述二根的确使其似然达到最大，于是，所求的MLE为

$$\begin{cases}\hat{\mu}=\bar{X},\\ \hat{\sigma}^2=S_n^{*2}.\end{cases}$$

□

比较一下例2.2.1与例2.3.2可知，对于正态总体而言，其均值与方差的矩估计和MLE是一样的.

例 2.3.3 设$X_1,\dots,X_n$为来自均匀分布$U(0,\theta)$的iid样本，试求参数θ的MLE.

解 此时的似然函数为

$$L(\theta;\mathbf{x})=\prod_{i=1}^n f(x_i,\theta)=\begin{cases}\frac{1}{\theta^n}, & 0<\max_{1\le i\le n}x_i<\theta,\\ 0, & \text{其它}.\end{cases}$$

显然，当$\theta=\max_{1\le i\le n}x_i$时，$L(\theta,\mathbf{x})$达到最大，于是$\hat{\theta}=X_{(n)}$是$\theta$的MLE. □

比较一下例2.2.3与例2.3.3可知，对于均匀分布总体$U(0,\theta)$，参数θ的矩估计和MLE是不一样的. 为了对二者作一比较，我们记

$$\hat{\theta}_M=2\bar{X},\quad \hat{\theta}_L=X_{(n)},$$

由概率论知识，我们容易求得它们的一、二矩如下：

$$\begin{array}{ll} E\hat{\theta}_M=\theta, & \mathrm{Var}\hat{\theta}_M=\frac{\theta^2}{3n}, \\ E\hat{\theta}_L=\frac{n}{n+1}\theta, & \mathrm{Var}\hat{\theta}_L=\frac{n\theta^2}{(n+1)^2(n+2)}. \end{array}$$

从二者的期望来看，显然$\hat{\theta}_L$不是以θ为中心，而$\hat{\theta}_M$是. 但是从方差来看，由于$\mathrm{Var}\hat{\theta}_L<\mathrm{Var}\hat{\theta}_M$，故$\hat{\theta}_L$比$\hat{\theta}_M$更集中在其均值附近.

另外，我们注意到，在例2.3.3中，我们不能利用微分法求其MLE，其原因是此时似然函数的支撑集依赖于未知参数θ.

例 2.3.4 设$X_1,\ldots,X_n$为来自$P(\lambda)$的iid样本，试求λ的MLE.

解 此时的似然函数为

$$L(\lambda,\mathbf{x})=\prod_{i=1}^{n}\frac{\lambda^{x_i}e^{-\lambda}}{x_i!}=\frac{\lambda^{\sum_{i=1}^{n}x_i}e^{-n\lambda}}{\prod_{i=1}^{n}x_i!},$$

其对数似然为

$$l(\lambda,\mathbf{x})=-n\lambda+\sum_{i=1}^{n}x_i\ln\lambda-\sum_{i=1}^{n}\ln x_i!,$$

似然方程为

$$\frac{\partial l(\lambda,\mathbf{x})}{\partial\lambda}=-n+\sum_{i=1}^{n}x_i/\lambda=0,$$

由此解得

$$\lambda=\bar{x}.$$

经验证知$\hat{\lambda}=\bar{X}$是λ的MLE. □

由此例可知，Poisson分布参数λ的MLE与例2.2.2中的第一个矩估计一样，并且这个估计也比较符合实际.

例 2.3.5 设$X_1,\ldots,X_n$为来自$U(\theta,\theta+1)$的iid样本，$\theta\in R$, 试求θ的MLE.

解 此时其似然函数为

$$L(\theta,\mathbf{x})=\begin{cases}1, & \theta\le x_{(1)}\le x_{(n)}\le\theta+1,\\ 0, & \text{其它}.\end{cases}$$

由于此时的似然函数在不为零的区域上是常数，故只要θ不超过$x_{(1)}$，而$\theta+1$不小于$x_{(n)}$，其似然函数均达到最大值1，故θ的MLE不止一个，在区间$[X_{(n)}-1,X_{(1)}]$ 中的任何一个均是θ的MLE. □

由此例可知，MLE并不惟一.

例 2.3.6 设$X_1,\ldots,X_n$是来自指数分布$E(\lambda)$的iid样本，试求λ的MLE.

解 我们以$X_{(1)}\le\cdots\le X_{(n)}$记次序统计量. 由次序统计量的知识可知，基于$X_{(1)},\ldots,X_{(r)}$的

似然函数为

$$L(\lambda, x_{(1)}, \ldots, x_{(r)}) = \frac{n!}{(n-r)!}\exp\left\{-\lambda[x_{(1)} + \cdots + x_{(r)} + (n-r)x_{(r)}]\right\}\lambda^r,$$

其对数似然为

$$l(\lambda, x_{(1)}, \ldots, x_{(r)}) = -\lambda[x_{(1)} + \cdots + x_{(r)} + (n-r)x_{(r)}] + \ln\frac{n!}{(n-r)!} + r\ln\lambda,$$

于是，其似然方程为

$$\frac{\partial l}{\partial \lambda} = -(x_{(1)} + \cdots + x_{(r)} + (n-r)x_{(r)}) + \frac{r}{\lambda} = 0,$$

于是求得λ的MLE为

$$\hat{\lambda} = \frac{r}{\sum_{i=1}^r X_{(i)} + (n-r)X_{(r)}}.$$

□

注 2.3.1 从前面的例子可以看出，MLE多是充分统计量的函数，事实上，如果T是充分统计量，则由因子分解定理可知，$f(\mathbf{x};\theta) = g(\theta, T(\mathbf{x}))h(\mathbf{x})$,于是，$l(\theta;\mathbf{x}) = \ln g(\theta, T(\mathbf{x})) + \ln h(\mathbf{x})$,由此可见，为使对数似然达到最大，只需使$g(\theta, T(\mathbf{x}))$达到最大，所以，MLE肯定是充分统计量$T(x)$的函数.

对于表示寿命分布的指数分布$E(\lambda)$，$1/\lambda$表示其平均寿命. 于是，许多实际工作者感兴趣的可能是关于平均寿命$1/\lambda$的估计. 而在上面例子中仅讨论了λ的估计，那我们又如何估计其平均寿命$1/\lambda$呢？见下面注.

注 2.3.2 如$g(\theta)$是1-1映射，且$\hat{\theta}$是θ的MLE，则可以证明$g(\hat{\theta})$也是$g(\theta)$的MLE.

极大似然估计的上述性质称为不变原则，它在许多实际问题中具有很大的应用.

注 2.3.3 对于指数族分布，似然方程(2.3.6)如有解，则必惟一. 对于单参数指数族$f(x;\theta) = \exp\{\theta T(x) + a(\theta)\}h(x)$，其似然方程为

$$\frac{\partial l(\theta;x)}{\partial \theta} = \sum T(x_i) + na'(\theta) = 0. \tag{2.3.7}$$

由于$ET(X) = -a'(\theta), \mathrm{Var}T(X) = -a''(\theta) > 0$(利用$\int f(x;\theta)\mathrm{d}x = 1$，之后两边关于$\theta$求导即可)，故如上式有解，必惟一. 如果上式无解，则其MLE在Θ的边界处达到.

§2.3.2 EM算法

MLE是一种非常有效的参数估计方法，但当分布中有讨厌参数或数据为截尾或缺失时，其MLE的求取是比较困难的. 于是Dempster等人于1977 年提出了EM算法，其含义是把求MLE的过程分两步走，第一步求期望，以便把讨厌的部分去掉，第二步求极大值. 本小节将简单介绍这种非常有用的方法.

一、EM算法

下面我们以一个例子说明EM算法是如何进行的.

例 2.3.7 设一次试验可能有四个结果, 其发生的概率分别为$\frac{1}{2} + \frac{\theta}{4}, \frac{1-\theta}{4}, \frac{1-\theta}{4}, \frac{\theta}{4}$, 其中$\theta \in$

$(0,1)$, 现进行了197 次试验, 四种结果的发生次数分别为125, 18, 20, 34. 试求θ的MLE.

解 由于此时的总体分布为多项分布，故其似然函数正比于

$$p(\theta,y)\propto\left(\frac{1}{2}+\frac{\theta}{4}\right)^{y_1}\left(\frac{1-\theta}{4}\right)^{y_2}\left(\frac{1-\theta}{4}\right)^{y_3}\left(\frac{\theta}{4}\right)^{y_4}\propto(2+\theta)^{y_1}(1-\theta)^{y_2+y_3}\theta^{y_4}, \tag{2.3.8}$$

要由此式求解θ的MLE是比较麻烦的，由于其对数似然方程是一个三次多项式.

但是我们可以通过引入一个变量Z后，使得求解变得比较容易. 现假设第一种结果可以分成两部分，其发生概率分别为$1/2$和$\theta/4$，令Z和Y_1-Z分别表示落入这两部分的次数. 显然，Z是我们人为引入的，它是不可观测的(在文献中称之为latent变量). 也称数据$(\mathbf{Y},Z)$为完全数据(complete data)，而观测到的数据$\mathbf{Y}$称为不完全数据. 此时完全数据的似然函数为

$$p(\theta|\mathbf{y},z)\propto\left(\frac{1}{2}\right)^{z}\left(\frac{\theta}{4}\right)^{y_1-z}\left(\frac{1-\theta}{4}\right)^{y_2}\left(\frac{1-\theta}{4}\right)^{y_3}\left(\frac{\theta}{4}\right)^{y_4}\propto\theta^{y_1-z+y_4}(1-\theta)^{y_2+y_3},$$

其对数似然为

$$l(\theta|z,y)\propto(y_1-z+y_4)\ln\theta+(y_2+y_3)\ln(1-\theta). \tag{2.3.9}$$

如果y,z均已知，则由上式很容易求得θ的MLE(现在是二次方程了)，但遗憾的是，我们仅知道y，而不知道z的值. 但是我们注意到，当y_1及θ已知时，$Z\sim B(y_1,2/(2+\theta))$ (其原因见后面的注2.3.4). 于是，Dempster等人建议如下分两步进行：

E-step: 在给定y及$\theta=\theta^{(i)}$的条件下，求关于安全数据对数似然关于潜在变量Z的期望：

$$Q(\theta|y,\theta^{(i)})=E_Z l(\theta|Z,y); \tag{2.3.10}$$

M-step: 求$Q(\theta|y,\theta^{(i)})$关于θ的最大值$\theta^{(i+1)}$，即

$$\theta^{(i+1)}=\arg\max_{\theta}Q(\theta|y,\theta^{(i)}). \tag{2.3.11}$$

重复(2.3.10)和(2.3.11)式，直至收敛即可得到θ的MLE.

对于本问题，其E步为：

$$\begin{aligned}Q(\theta|y,\theta^{(i)})&\propto[y_1-E_Z(Z|y,\theta=\theta^{(i)})+y_4]\ln\theta+(y_2+y_3)\ln(1-\theta)\\&=[y_1-2y_1/(2+\theta^{(i)})+y_4]\ln\theta+(y_2+y_3)\ln(1-\theta),\end{aligned}$$

其M步即为上式两边关于θ求导，并令其等于0后，可求得

$$\theta^{(i+1)}=\frac{y_1\theta^{(i)}+y_4(\theta^{(i)}+2)}{y_1\theta^{(i)}+(y_2+y_3+y_4)(\theta^{(i)}+2)}=\frac{159\theta^{(i)}+68}{197\theta^{(i)}+144}.$$

如取$\theta^{(0)}=0.5$，则四步迭代后可求得θ的MLE为0.6268. □

注 2.3.4 在$Y_1=y_1$的条件下，$Z\sim b(y_1,2/(2+\theta))$的原因：以$A_1$表示第一种结果出现的事件，$B_1,B_2$分别表示我们所定义的两个事件，$A_1=B_1\cup B_2$. 由定义知它们是独立的，且$P(A_1)=\frac{1}{2}+\frac{\theta}{4},P(B_1)=1/2,P(B_2)=\theta/4$, 则可知$P(B_1|A_1)=P(B_1)/P(A_1)=2/(2+\theta)$.

注 2.3.5 在(2.3.10)式右边的期望是关于Z在$\theta=\theta^{(i)}$的条件下求取的，而其余的参数不变，故左边与$\theta^{(i)}$有关.

注 2.3.6 上述EM算法的收敛性没有问题，但得到的有可能是局部极大值.

关于上述算法的收敛性，我们有如下的结论：

定理 2.3.1 对于上述的EM算法，我们有

$$l_0(\theta^{(i+1)}, y) \geq l_0(\theta^{(i)}, y) \geq 0, \tag{2.3.12}$$

其中$l_0(\theta, y)$为不完全数Y的对数似然函数，即$l_0(\theta, y) = \ln f(y, \theta)$.

证明 为了证明的方便，我们改写前面的记号为：

不完全数据（观测数据）Y的PDF和对数似然分别为$f(y|\theta)$和$l_0(\theta|y)$;

完全数据(Y, Z)的PDF为对数似然分别为$f(y, z|\theta)$和$l(\theta|y, z)$.

于是，由条件PDF公式知,在(Y, θ)已知条件下，Z的PDF为$f(z|y, \theta) = \frac{f(y,z|\theta)}{f(y|\theta)}$, 由此可知$f(y|\theta) = \frac{f(y,z|\theta)}{f(z|y,\theta)}$，两边取对数后有

$$l_0(\theta|y) = l(\theta|y, z) - \ln f(z|y, \theta),$$

对上式两边求Z在$(Y, \theta = \theta^{(i)})$已知条件下的期望，有

$$l_0(\theta|y) = Q(\theta|y, \theta^{(i)}) - \int f(z|y, \theta^{(i)}) \ln f(z|y, \theta) dz, \tag{$*$}$$

在(*)式中，取$\theta = \theta^{(i)}$和$\theta^{(i+1)}$,得到下面二式

$$l_0(\theta^{(i)}|y) = Q(\theta^{(i)}|y, \theta^{(i)}) - \int f(z|y, \theta^{(i)}) \ln f(z|y, \theta^{(i)}) dz, \tag{$*1$}$$

$$l_0(\theta^{(i+1)}|y) = Q(\theta^{(i+1)}|y, \theta^{(i)}) - \int f(z|y, \theta^{(i)}) \ln f(z|y, \theta^{(i+1)}) dz, \tag{$*2$}$$

(*2)−(*1)式得到

$$\begin{aligned} & l_0(\theta^{(i+1)}|y) - l_0(\theta^{(i)}|y) \\ = \quad & Q(\theta^{(i+1)}|y, \theta^{(i)}) - Q(\theta^{(i)}|y, \theta^{(i)}) - \int \ln \frac{f(z|y, \theta^{(i+1)})}{f(z|y, \theta^{(i)})} f(z|y, \theta^{(i)}) dz, \end{aligned}$$

由EM算法的M步知道，$Q(\theta^{(i+1)}|y, \theta^{(i)}) - Q(\theta^{(i)}|y, \theta^{(i)}) \geq 0$，对于第二部分，由Jessen不等式知，

$$\int \ln \frac{f(z|y, \theta^{(i+1)})}{f(z|y, \theta^{(i)})} f(z|y, \theta^{(i)}) dz \leq \ln \int \frac{f(z|y, \theta^{(i+1)})}{f(z|y, \theta^{(i)})} f(z|y, \theta^{(i)}) dz = 0,$$

$(E \ln X \leq \ln EX)$

于是，知道$l_0(\theta^{(i+1)}|y) - l_0(\theta^{(i)}|y) \geq 0$, 即得证. □

此定理告诉我们，通过EM算法得到的序列$\theta^{(n)}$，会使其原数据的似然函数$l_0(\theta^{(n)}|y)$单增.

如以$\hat{\theta}$表示EM算法所得到的估计，则它的渐近方差近似等于下面的观测数据的Fisher信息的倒数(关于Fisher信息的定义请见下节)：

$$I_0^{-1} = \left(-\frac{\partial^2 \ln f(y|\theta)}{\partial \theta^2} \Big|_{\hat{\theta}} \right)^{-1}. \tag{2.3.13}$$

关于EM算法的研究，现已取得了非常丰富的结果，且许多统计软件均给出了EM算法的程序.

二、GEM算法

虽然EM算法已得到了广泛的应用，但在某些情况下，其M步的最大值是难于计算的. 但有时我们可以用一个简单的方法找一个$\theta^{(i+1)}$，使得

$$Q(\theta^{(i+1)}|y,\theta^{(i)}) > Q(\theta^{(i)}|y,\theta^{(i)}), \tag{2.3.14}$$

则称由(2.3.10)和(2.3.14)式求取θ的MLE的方法称广义的EM算法，记为GEM算法.

三、Monte Carlo EM算法

上一小节仅考虑了M步的计算难度. 实际上，对于某些问题，其E步的计算也是比较困难的. 回忆一下(2.3.10)式，它就是求取完全数据的似然函数在$(Y, \theta = \theta^{(i)})$下的条件期望，而前面的随机模拟告诉我们，期望可用样本平均值来估计，于是，就有了本小节的Monte Carlo EM算法：其M步仍然如(2.3.11)式，只是其E步由下面两步组成：

E1 step: 由Z的条件分布$p(z|y,\theta^{(i)})$抽取m个随机数$z_1, \ldots, z_m$;

E2 step: $Q(\theta|y,\theta^{(i)}) = \frac{1}{m}\sum_{k=1}^{m} \ln f(\theta|y, z_k)$.

§2.4 无偏估计与一致最小方差无偏估计

从上一节我们可以看出，对于同一个问题，所用方法不同，得到的估计也可能不同，那我们应如何选择呢？这就要涉及到选择准则问题. 本节将介绍统计中经常使用的无偏性准则和一致最小方差无偏准则.

§2.4.1 无偏估计

定义 2.4.1 (UE) *如果$T(\mathbf{X})$是未知参数θ的函数$g(\theta)$的一个估计量，且满足*

$$E_\theta T(\mathbf{X}) = g(\theta),\ \forall\ \theta \in \Theta,$$

则称T是$g(\theta)$的UE，也称$g(\theta)$是可估的(Estimable)，其中E_θ表示期望是在分布f_θ下进行的.

对于正态总体，我们不难验证，样本均值$\bar{X}$及样本方差S_n^2分别是总体均值与方差的UE (对于非正态总体，这一结论也是正确的)，而总体方差的矩估计及MLE—S_n^{*2}则不是无偏的，这是我们采用S_n^2 作为样本方差定义的一个原因. 但是，我们也注意到，虽然S_n^{*2}不是无偏的，但随着样本容量n的加大，它越来越接近无偏.

定义 2.4.2 (渐近UE) *如果$T(\mathbf{X})$是$g(\theta)$的一个有偏估计量，且满足*

$$\lim_{n\to\infty} E_\theta T(X_1, \ldots, X_n) = g(\theta),\ \forall\ \theta \in \Theta,$$

则称T是$g(\theta)$的渐近UE.

关于无偏估计，我们有如下几个注解：

- 如果$T(\mathbf{X})$是$g(\theta)$的有偏估计，则称$E_\theta T(\mathbf{X}) - g(\theta)$为其偏差(bias).

- 无偏估计是从多次重复的角度而引出的一个概念，从期望的定义不难看出，尽管一次估计，$T(\mathbf{x})$的值不一定恰好等于参数真值$g(\theta)$，但当大量重复使用时，其多次估计的平均值即等于参数值$g(\theta)$.

- 一个参数的无偏估计可能不是惟一的，也可能不存在，也可能不合理.

大偏差通常被视为估计的一种不足，于是，有人提出了缩小偏差的方法. 下面的刀切法就是由Quenouille于1949(*JRSS B*, 11, 18-44)和1956(*Biometrika*, 43, 353–360)提出的，而正式命名则由Tukey于1958年(*Ann. Math. Statist.* 29, 614).

例 2.4.1 (刀切法(Jackknife)) 设$T(\mathbf{X})$是基于样本$\mathbf{X}$=$(X_1,\dots,X_n)'$的关于参数$g(\theta)$的估计量，且满足$ET(\mathbf{X})=g(\theta)+O\left(\frac{1}{n}\right)$. 如以$\mathbf{X}_{(-i)}$表示从样本中删去$X_i$后的向量，则$T(\mathbf{X})$的刀切统计量定义为

$$T_J(\mathbf{X})=nT(\mathbf{X})-\frac{n-1}{n}\sum_{i=1}^{n}T(\mathbf{X}_{(-i)}). \tag{2.4.1}$$

可以证明：由(2.4.1)定义的刀切统计量具有如下性质：

$$ET_J(\mathbf{X})=g(\theta)+O\left(\frac{1}{n^2}\right),$$

并且其方差不会增大. 关于刀切法的详细介绍，请有兴趣的参考Efron (1982). 另外，还有一个与刀切法类似的方法—Bootstrap(自助法)，也是统计中非常有用的一种方法，具体的可见Efron and Tibshirani (1993), 和Shao and Tu (1995).

但是，我们必须注意到，追求偏差压缩可能不是一个好办法，由于有人证明：对于一个$g(\theta)$的有偏估计S，当其偏差趋于0时，其方差$\mathrm{Var}(S)\to\infty$ (见Lehmann and Casella (1998)).

例 2.4.2 对于Poisson分布$P(\lambda)$，$\lambda>0$是未知参数. 现设有一个来自此分布的样本X，令$T(X)=(-1)^X$. 则由于

$$ET(X)=\sum_{k=0}^{\infty}(-1)^k\frac{\lambda^k}{k!}e^{-\lambda}=e^{-2\lambda},$$

故$T(X)$是$e^{-2\lambda}$的一个UE. 但是当X为奇数时，这一估计并不合理.

例 2.4.3 对于Poisson分布$P(\lambda)$，设$X_1\sim P(\lambda)$，试求基于X_1的参数$1/\lambda$的UE.

解 此时$1/\lambda$是不可估的. 事实上，如果存在一个统计量$T(X_1)$是它的UE，则有

$$ET(X_1)=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}T(k)e^{-\lambda}=\frac{1}{\lambda},$$

即

$$\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}T(k)=\frac{e^\lambda}{\lambda}=\frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}.$$

由多项式相等可知，上述等式永远不会成立，于是，参数$1/\lambda$是不可估的. □

前面我们讲，充分统计量具有不变性，那无偏估计具有这种不变性吗？下面的例子将回答这个问题.

例 2.4.4 对于正态总体，样本标准差S_n不是σ的UE.

解 由定理1.3.4可知，此时$(n-1)S_n^2/\sigma^2 \sim \chi^2(n-1)$，故有

$$E\left(\sqrt{n-1}S_n/\sigma\right) = \int_0^\infty \frac{\sqrt{x}}{2^{(n-1)/2}\Gamma((n-1)/2)} e^{-x/2} x^{\frac{n-1}{2}-1} \mathrm{d}x = \cdots = \frac{\sqrt{2}\Gamma(n/2)}{\Gamma((n-1)/2)},$$

故

$$ES_n = \sigma\sqrt{\frac{2}{n-1}}\frac{\Gamma(n/2)}{\Gamma((n-1)/2)}, \tag{2.4.2}$$

所以，样本标准差不是σ的UE. □

上例告诉我们，如果T是θ的UE，则除了g是线性函数之外，$g(T)$并不是$g(\theta)$的UE.

§2.4.2 一致最小均方误差准则

对于均匀分布$U(0,\theta)$，我们在上一节讨论其参数θ的矩估计与MLE的差别. 虽然矩估计是无偏的, MLE是有偏的，但二者的偏差并不大，且矩估计的方差明显大于MLE 的方差，故从实际应用的角度看，MLE应“优”于矩估计，其比较准则就是本节要讲的均方误差准则.

定义 2.4.3 设$X_1,\ldots,X_n$是来自分布族$\mathcal{F}=\{f(x,\theta):\theta\in\Theta\}$中某一分布的样本，$g(\theta)$是一参数函数，以$\mathcal{E}(g)$表示用来估计$g(\theta)$的某些估计量的集合，如果存在一个$T^*\in\mathcal{E}(g)$，使得对任一$T\in\mathcal{E}(g)$均有

$$E_\theta(T^*-g(\theta))^2 \le E_\theta(T-g(\theta))^2,\ \forall\ \theta\in\Theta, \tag{2.4.3}$$

则称T^*为$g(\theta)$的在$\mathcal{E}(g)$中的一致最小均方误差估计，也称满足(2.4.3)式的统计量T^*在均方意义下优于T.

之所以称满足(2.4.3)式的T^*为最小均方误差估计，是由于估计量$T\in\mathcal{E}(g)$的均方误差(Mean Square Error，简记为MSE)的定义为$\mathrm{MSE}(T)=E_\theta[T-g(\theta)]^2$. 当$T$是$g(\theta)$的UE时，其MSE就是它的方差.

例 2.4.5 考虑正态总体$N(\mu,\sigma^2)$中关于σ^2的一致最小均方误差估计

解 在前面我们讲过样本方差S_n^2是σ^2的一个UE，且$\mathrm{Var}S_n^2=\frac{2\sigma^4}{n-1}=\mathrm{MSE}(S_n^2)$. 下面我们在形如$cS_n^2$的估计类中找一个MSE最小的估计量. 事实上，由于

$$\begin{aligned}\mathrm{MSE}(cS_n^2) &= E(cS_n^2-\sigma^2)^2 \\ &= E\left[c(S_n^2-\sigma^2)-\sigma^2(1-c)\right]^2 = \sigma^4\left[\frac{2c^2}{n-1}+(1-c)^2\right].\end{aligned}$$

如令$f(c)=\frac{2c^2}{n-1}+(1-c)^2$，则知它在$c=\frac{n-1}{n+1}$处达到最小. 于是，如取

$$\hat{\sigma}^2 = \frac{1}{n+1}\sum_{i=1}^n (X_i-\bar{X})^2, \tag{2.4.4}$$

则知$\hat{\sigma}^2$不是σ^2的UE，但

$$\frac{2\sigma^4}{n+1} = \mathrm{MSE}(\hat{\sigma}^2) < \mathrm{MSE}(S_n^2) = \frac{2\sigma^4}{n-1}.$$

□

此例说明，无偏准则与均方误差准则是从两个不同角度考察一个估计量好坏的. 但当二者发生矛盾时，更应重视均方误差准则.

我们应注意到，这里所说的一致并不是指在估计类$\mathcal{E}(g)$中一致，而是指式(2.4.3)关于$\theta \in \Theta$一致. 因此，有时一致最小均方误差估计并不存在. 此时，为了找到一个一致最小均方误差估计，通常都是通过来缩小估计类来求得.

§2.4.3 一致最小方差无偏估计

定义 2.4.4 (UMVUE) 对于参数分布族$\mathcal{F} = \{f(x,\theta) : \theta \in \Theta\}$，$g(\theta)$是一可估函数，又设$T^*$是$g(\theta)$的一个UE. 如对于$g(\theta)$的任一UE T，均有

$$\mathrm{Var}_\theta(T^*) \leq \mathrm{Var}_\theta(T),\ \forall\ \theta \in \Theta, \tag{2.4.5}$$

则称T^*是$g(\theta)$的UMVUE.

从UMVUE的定义不难看出，它是一个在无偏估计类中方差最小的估计. 但我们注意到，虽然UMVUE是一个很好的估计，但对于某些分布族或参数，其UMVUE不一定存在(见本节例2.4.5)

为了方便求取UMVUE，我们给出如下一个定理.

定理 2.4.1 设$T(\mathbf{X})$是参数分布族$\mathcal{F} = \{f(x,\theta) : \theta \in \Theta\}$的一个充分统计量，$S(\mathbf{X})$是参数$g(\theta)$的一个UE，则

$$h(T) = E[S(\mathbf{X})|T(\mathbf{X})], \tag{2.4.6}$$

是$g(\theta)$的一个UE，且

$$\mathrm{Var}_\theta(h(T)) \leq \mathrm{Var}_\theta(S(\mathbf{X})), \tag{2.4.7}$$

其中等号成立的充要条件是$P\{S(\mathbf{X}) = h(T)\} = 1$.

证明 因为T是充分统计量，故由充分统计量的定义知，在给定T下，样本$\mathbf{X}$ =$(X_1, \ldots, X_n)$的条件分布与参数θ无关. 因此，由(2.4.6)式定义的$h(T)$是一统计量. 另外，由于

$$E_\theta(h(T)) = E_\theta\left[E(S(\mathbf{X})|T(\mathbf{X}))\right] = E_\theta(S(\mathbf{X})) = g(\theta),\ \forall\ \theta \in \Theta,$$

故$h(T)$也是$g(\theta)$的一个UE.

又因为

$$\begin{aligned}
\mathrm{Var}_\theta(S(\mathbf{X})) &= E_\theta[S(\mathbf{X}) - g(\theta)]^2 \\
&= E_\theta[S(\mathbf{X}) - h(T) + h(T) - g(\theta)]^2 \\
&= E_\theta[S(\mathbf{X}) - h(T)]^2 + \mathrm{Var}_\theta(h(T)) + 2E_\theta[(S(\mathbf{X}) - h(T))(h(T) - g(\theta))],
\end{aligned}$$

又由于

$$\begin{aligned}
& E_\theta[(S(\mathbf{X}) - h(T))(h(T) - g(\theta))] \\
= \ & E_\theta\{E_\theta[(S(\mathbf{X}) - h(T))(h(T) - g(\theta))]|T\} \\
= \ & E_\theta\{(h(T) - g(\theta)) \cdot E_\theta(S(\mathbf{X}) - h(T))|T\} \\
= \ & E_\theta\{(h(T) - g(\theta)) \cdot E_\theta[(S(\mathbf{X}) - h(T))|T]\} \\
= \ & E_\theta\{(h(T) - g(\theta)) \cdot [(E_\theta S(\mathbf{X})\|T) - h(T)]\} = 0,
\end{aligned}$$

故

$$\mathrm{Var}_\theta(S(\mathbf{X})) = E_\theta(S(\mathbf{X}) - h(T))^2 + \mathrm{Var}_\theta(h(T)) \le \mathrm{Var}_\theta(h(T)),\ \forall\ \theta \in \Theta,$$

且等号成立当且仅当$E_\theta(S(\mathbf{X}) - h(T))^2 = 0$，即$P\{S(\mathbf{X}) = h(T)\} = 1$. □

从定理2.4.1可以看出，当我们找到充分统计量后，任一个UE都可以得到改进，并且，为找到UMVUE，我们仅需在基于充分统计量的无偏估计类中去找即可. 但这一定理并没有告诉我们如何才能得到UMVUE. 下一定理将回答这一问题. 为此，我们引入如下两个无偏估计类：

$$U = \{T : E_\theta T = g(\theta), E_\theta T^2 < \infty, \forall\ \theta \in \Theta\}, \tag{2.4.8}$$

$$U_0 = \{T : E_\theta T = 0,\quad E_\theta T^2 < \infty, \forall\ \theta \in \Theta\}, \tag{2.4.9}$$

它们分别表示$g(\theta)$与0的具有二阶矩的无偏估计类.

定理 2.4.2 对于参数分布族$\mathcal{F} = \{f(x,\theta) : \theta \in \Theta\}$，设$g(\theta)$可估. 则估计量$T_0 \in U$是$g(\theta)$的一个UMVUE 的充要条件是

$$E_\theta(\nu T_0) = 0,\ \forall\ \theta \in \Theta,\ \forall\ \nu \in U_0. \tag{2.4.10}$$

证明 先证充分性：反证

设$T_0 \in U$是$g(\theta)$的一个UMVUE，但条件(2.4.10)不成立，即存在$\nu_0 \in U_0$和$\theta_0 \in \Theta$，使得$E_{\theta_0}(\nu_0 T_0) \neq 0$.

因为$\nu_0 \in U_0$，故对于任意的λ，有$T_\lambda = T_0 - \lambda\nu_0 \in U$，而

$$E_{\theta_0} T_\lambda^2 = E_{\theta_0}(T_0 - \lambda\nu_0)^2 = E_{\theta_0} T_0^2 + \lambda^2 E_{\theta_0}\nu_0^2 - 2\lambda E_{\theta_0}(T_0\nu_0).$$

由于$E_{\theta_0}(\nu_0 T_0) \neq 0$，则一定能找到$\lambda_0 (= E_{\theta_0}(\nu_0 T_0)/E_{\theta_0}\nu_0^2)$，使得

$$E_{\theta_0} T_{\lambda_0}^2 < E_{\theta_0} T_0^2,$$

即$\mathrm{Var}_{\theta_0} T_{\lambda_0} < \mathrm{Var}_{\theta_0} T_0$，与$T_0$是一个UMVUE矛盾.

下证必要性.

设T_0满足(2.4.10)式，则对于任一$T \in U$，由于$T - T_0 \in U_0$，则由(2.4.10)知

$$E_\theta[(T - T_0)T_0] = 0,\ \forall\ \theta \in \Theta,$$

即

$$E_\theta T_0^2 = E_\theta(T_0 T),\ \forall\ \theta \in \Theta,$$

于是由Cauchy-Schwartz不等式知，

$$E_\theta T_0^2 \le (E_\theta T_0^2)^{1/2}(E_\theta T^2)^{1/2},\ \forall\ \theta \in \Theta,$$

即

$$E_\theta T_0^2 \le E_\theta T^2,\ \forall\ \theta \in \Theta.$$

又由于T_0, T均是$g(\theta)$ 的UE，且T是任意的，故由上式可知，T_0是$g(\theta)$的UMVUE. □

UMVUE如存在则它惟一吗？下一定理回答这一问题.

定理 2.4.3 如果$g(\theta)$的无偏估计类U非空，则其UMVUE最多有一个.

证明 如果$T \neq T_0$是$g(\theta)$的两个UMVUE，则必有

$$E_\theta T = E_\theta T_0 = g(\theta),\ \mathrm{Var}_\theta T_0 = \mathrm{Var}_\theta T,\ \forall\ \theta \in \Theta.$$

于是，$T-T_0\in U_0$. 因为T是一个UMVUE，故由定理2.4.2知

$$E_\theta[(T-T_0)T]=0,\ \forall\,\theta\in\Theta.$$

又由于T_0也是一个UMVUE，故由定理2.4.2知

$$E_\theta[(T-T_0)T_0]=0,\ \forall\,\theta\in\Theta.$$

综合以上两式，我们有

$$E_\theta(T-T_0)^2=E_\theta[(T-T_0)(T-T_0)]=0,\ \forall\,\theta\in\Theta,$$

故$P_\theta\{T=T_0\}=1,\ \forall\,\theta\in\Theta.$ □

Corollary 2.4.1 如果T_1和T_2分别是$g_1(\theta)$和$g_2(\theta)$的UMVUE, 则对于任给的常数a,b, aT_1+bT_2是$ag_1(\theta)+bg_2(\theta)$ 的UMVUE.

我们先看一个UMVUE不存在的例子，之后再看如何利用上述定理来求UMVUE.

例 2.4.6 设$X_1,\ldots,X_n$是来自一离散分布的iid样本，其分布在$\theta-1,\theta,\theta+1$上的概率均为$1/3$, θ取整数. 则所有的θ的非常数函数都没有UMVUE. 当样本来自均匀分布$U(\theta-1/2,\theta+1/2)$时给出了连续分布的例子.

关于此问题的证明请见Lehmann &CaseHa (1998).

例 2.4.7 设$X_1,\ldots,X_n$为来自正态总体$N(\mu,\sigma^2)$的iid样本，试求μ,σ^2的UMVUE.

解 此时样本的联合PDF为

$$f(\mathbf{x};\mu,\sigma^2)=(2\pi\sigma^2)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n(x_i-\mu)^2\right\},$$

则任给$\nu\in U_0$，有

$$\int\cdots\int\nu(\mathbf{x})\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n(x_i-\mu)^2\right\}\mathrm{d}\mathbf{x}=0,\ \forall\,\mu,\sigma^2. \qquad (*)$$

上式两边关于μ求导，综合(*)式后，得到

$$\int\cdots\int\nu(\mathbf{x})(\sum_{i=1}^n x_i)\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n(x_i-\mu)^2\right\}\mathrm{d}\mathbf{x}=0,\ \forall\,\mu,\sigma^2, \qquad (*1)$$

如取$T(\mathbf{X})=\sum_{i=1}^n X_i/n=\bar{X}$，则上式即为

$$\int\cdots\int\mu(\mathbf{x})T(\mathbf{x})f(\mathbf{x};\mu,\sigma^2)\mathrm{d}\mathbf{x}=0,\ \forall\,\mu,\sigma^2.$$

又由于$T(\mathbf{X})$是μ的UE，故由定理2.4.2知，$T(\mathbf{X})=\bar{X}$是μ的UMVUE.

在(*1)式两边再关于μ求导，综合(*)式后，有

$$\int\cdots\int\mu(\mathbf{x})T^2(\mathbf{x})f(\mathbf{x};\mu,\sigma^2)\mathrm{d}\mathbf{x}=0,\ \forall\,\mu,\sigma^2. \qquad (*2)$$

另外，在(*)式两边关于σ^2求导，有

$$\int\cdots\int\mu(\mathbf{x})(\sum_{i=1}^n(x_i-\mu)^2)f(\mathbf{x};\mu,\sigma^2)\mathrm{d}\mathbf{x}=0,\ \forall\,\mu,\sigma^2. \qquad (*3)$$

综合(*1),(*2)和(*3)式，我们得到

$$\int \cdots \int \mu(\mathbf{x})(\sum_{i=1}^{n}(x_i-\bar{x})^2)f(\mathbf{x};\mu,\sigma^2)\mathrm{d}\mathbf{x}=0,\ \forall\ \mu,\sigma^2. \qquad (*4)$$

又由于$S_n^2=\frac{1}{n-1}\sum_{i=1}^n(X_i-\bar{X})^2$是$\sigma^2$的UE，则由(*4)式和定理2.4.2知，样本方差S_n^2是总体方差σ^2的UMVUE. □

从上例可以看出，由于修正的样本方差S_n^{*2}不是σ^2的UE，故它肯定不是σ^2的UMVUE.

例 2.4.8 设$X_1,\ldots,X_n$为来自指数分布族$E(\lambda)$的iid样本，试求总体均值$1/\lambda$的UMVUE.

解 从定理2.4.1可知，对于某分布族，当其充分统计量存在时，我们只需要在基于充分统计量的无偏估计类中求UMVUE即可. 对于本例，我们可由因子分解定理容易验证$T=\sum_{i=1}^n X_i$是其充分统计量(也可见例1.4.8). 于是，我们将只在充分统计量的无偏估计类求取$1/\lambda$的UMVUE.

由定理1.3.14可知，$T=\sum_{i=1}^n X_i\sim\Gamma(n,\lambda)$，因此，$T$的PDF为

$$f_T(x;\lambda)=\frac{1}{(n-1)!}\lambda^n x^{n-1}e^{-\lambda x},\ x>0,$$

且易证T/n是$1/\lambda$的UE.

$\forall\ \nu(T)\in U_0$，即$\nu(T)$是T的函数且为0的无偏估计，即

$$0=E\nu(T)=\frac{1}{(n-1)!}\int_0^\infty \nu(x)\lambda^n x^{n-1}e^{-\lambda x}\mathrm{d}x,\ \forall\ \lambda>0.$$

上式两边关于λ求导，有

$$\int_0^\infty x\nu(x)x^{n-1}e^{-\lambda x}\mathrm{d}x=0,\ \forall\ \lambda>0,$$

于是，由定理2.4.2知，$T=\bar{X}$是$1/\lambda$的UMVUE. □

例 2.4.9 设$X_1,\ldots,X_n$为来自$U(0,\theta)$的iid样本，试求θ的UMVUE.

解 由例1.4.6可知，$T=X_{(n)}$是θ的充分量，且由例1.2.12知，$X_{(n)}$的PDF为

$$f_T(t;\theta)=nt^{n-1}/\theta^n,\ 0<t<\theta,$$

由此容易验证$\frac{n+1}{n}T$是θ的一个UE，下面验证它也是θ的UMVUE. 为此，任给0的一个UE，$\nu(T)$，即

$$\int_0^\theta \nu(t)nt^{n-1}/\theta^n\mathrm{d}t=0,\ \forall\ \theta>0,$$

两边关于θ求导，得

$$\nu(\theta)=0,\ \forall\ \theta>0,$$

即此时0的无偏估计肯定为0，于是由定理2.4.2可知，$\frac{n+1}{n}X_{(n)}$是θ的UMVUE. 同样，因为$EX_1=\frac{\theta}{2}$，于是可知总体均值$\theta/2$的UMVUE为$\frac{n+1}{2n}X_{(n)}$. □

比例一下例2.4.8与例2.4.9的结果知道，样本均值不一定是总体均值的UMVUE.

§2.5 完备统计量

对于一个统计量$S(\mathbf{X})$，其定义域为$\mathcal{X}$，而其期望$E_\theta S(\mathbf{X})$的定义域是Θ，这就是说，这里的期望E_θ相当于一个从$\mathcal{X}$到Θ的变换. 对于一个变换布言，一个自然问题是：这个变换是1-1的吗？

我们在前面也说过，如果充分统计量存在，则我们可以仅在充分统计量的无偏估计类中去寻找参数的UMVUE. 如果我们知道基于充分统计量的无偏估计类仅有一个元素，则它肯定就是UMVUE. 那何时这个估计类仅有一个元素呢？

为回答上述问题，我们就要引进完备统计量(Complete Statistics)的概念. 它是Lehman和Scheffe于1950年提出的. 有的书上也称之为完全统计量.

§2.5.1 完备性的定义

定义 2.5.1 (完备统计量) 对于参数分布族$\mathcal{F}=\{f(x,\theta):\theta\in\Theta\}$，设$T(\mathbf{X})$为一统计量，如对任何满足条件

$$E_\theta g(T(\mathbf{X}))=0,\ \forall\ \theta\in\Theta, \tag{2.5.1}$$

的统计量$g(T)$，都有

$$P_\theta\{g(T)=0\}=1,\ \forall\ \theta\in\Theta, \tag{2.5.2}$$

则称统计量$T(\mathbf{X})$是完备统计量.

有时完备性的概念是针对分布族而言的，即下面的定义.

定义 2.5.2 (完备分布族) 对于参数分布族$\mathcal{F}=\{f(x,\theta):\theta\in\Theta\}$, 如果对于任一函数$\psi(x)$, 由

$$E_\theta\psi(X)=0,\ \forall\ \theta\in\Theta,$$

总可推出$P_\theta\{\psi(X)=0\}=1$，则称此分布族$\mathcal{F}$是完备的. 如果某统计量的抽样分布族是完备的，则称这个统计量是完备的.

例 2.5.1 考虑二项分布族$\{B(n,p):\ 0<p<1\}$的完备性.

解 设函数$\psi(x)$满足

$$E_p\psi(X)=\sum_{x=0}^{n}\psi(x)\begin{pmatrix} n\\ x\end{pmatrix}p^x(1-p)^{n-x}=0,\ \forall\ 0<p<1.$$

令$\theta=p/(1-p)$, 则$\theta>0$，且上式可以写成

$$\sum_{x=0}^{n}\psi(x)\begin{pmatrix} n\\ x\end{pmatrix}\theta^x=0,\ \forall\ \theta>0.$$

由于上式左边是θ的一个n次多项式，故有$\psi(x)=0,\ x=0,1,\cdots,n$，故二项分布是完备的. □

例 2.5.2 正态分布族$N(0,\sigma^2)$是不完备的.

解 我们注意到，正态分布的PDF是偶函数，故对于任何一个奇函数，比如$\psi(x)=x$，我们均有

$$E_\sigma\psi(X)=0,\ \forall\ \sigma^2>0,$$

但是$P_\sigma\{X=0\}=0$. 这说明此正态分布族是不完备的. □

关于正态总体$N(0,\sigma^2)$，我们在前面讲过，统计量$T_n=\sum_{i=1}^n X_i^2$是σ^2的充分统计量，虽然此分布族是不完备的，但它并不影响充分统计量T_n的完备性. 事实上，由于

$$T_n/\sigma^2\sim\chi^2(n)=\Gamma\left(\frac{n}{2},\frac{1}{2}\right),$$

即$T_n\sim\Gamma(\frac{n}{2},\frac{1}{2\sigma^2})$, 于是，$T_n$的PDF为

$$f(t,\sigma^2)=\frac{1}{(2\sigma^2)^{n/2}\Gamma(n/2)}t^{n/2-1}e^{-t/(2\sigma^2)},\ t>0.$$

如果$\psi(t)$满足$E_\sigma\psi(T)=0,\ \forall\ \sigma^2>0$, 即

$$\int_0^\infty\psi(t)t^{n/2-1}e^{-t/(2\sigma^2)}\mathrm{d}t=0,\ \forall\ \sigma^2>0.$$

上式左边是函数$\psi(t)t^{n/2-1}$的Laplace变换，故由Laplace变换的惟一性知

$$P_\sigma\{\psi(T)T^{n/2-1}=0\}=1,$$

故由定义可知，正态分布族$N(0,\sigma^2)$的充分统计量$T=\sum_{i=1}^n X_i^2$是完备的.

例 2.5.3 对于均匀分布族$U(0,\theta)$，说明其充分统计量$T_n=X_{(n)}$的完备性.

解 由于统计量T_n的PDF为

$$f(t,\theta)=\begin{cases}\frac{nt^{n-1}}{\theta^n}, & 0<t<\theta,\\ 0, & \text{否则},\end{cases}$$

又设$\psi(T)$满足$E_\theta\psi(T)=0,\ \forall\ \theta>0$, 即

$$\int_0^\theta\psi(t)f(t,\theta)\mathrm{d}t=0,\ \forall\ \theta>0,$$

即

$$\int_0^\theta\psi(t)t^{n-1}\mathrm{d}t=0,\ \forall\ \theta>0,$$

两边关于θ求导，有

$$\psi(\theta)\theta^{n-1}=0,\ \forall\ \theta>0,$$

即

$$\psi(\theta)=0,\ \forall\ \theta>0,$$

所以，均匀分布族$U(0,\theta)$的充分统计量$T_n=X_{(n)}$是完备的. □

§2.5.2 完备统计量的应用

为了回答本节刚开始提出的问题，我们看如下定理.

定理 2.5.1 设$X_1,\ldots,X_n$是来自参数分布族$\{f(x,\theta):\theta\in\Theta\}$的iid样本，$T$是$\theta$的充分完备统计量. 如果$g(\theta)$可估，且$S(\mathbf{X})$是$\theta$的一个UE，则$S_0(T)=E[S(\mathbf{X})|T]$是$\theta$的惟一的UMVUE.

证明 设S_1,S_2是θ的任意两个UE，则由定理2.4.1可知，$E(S_1|T),E(S_2|T)$均是θ的UE，即

$$E_\theta[E(S_1|T)]=E_\theta[E(S_2|T)]=\theta,\ \forall\ \theta\in\Theta, \tag{$*$}$$

且

$$\mathrm{Var}_\theta(E(S_1|T))\le \mathrm{Var}_\theta(S_1),\ \forall\ \theta\in\Theta,$$
$$\mathrm{Var}_\theta(E(S_2|T))\le \mathrm{Var}_\theta(S_2),\ \forall\ \theta\in\Theta.$$

另由(*)式可知，$E(S_1|T)-E(S_2|T)$是0的无偏估计，即

$$E_\theta[E(S_1|T)-E(S_2|T)]=0,\ \forall\ \theta\in\Theta,$$

又由于T是完备统计量，故由定义知

$$P_\theta\{E(S_1|T)=E(S_2|T)\}=1,\ \forall\ \theta\in\Theta,$$

这说明在概率1的意义下，$E(S_1|T)$与$E(S_2|T)$相等，又由于S_1与S_2的任意性，故$S_0=E(S|T)$是惟一的UMVUE. □

例 2.5.4 考虑Bernoulli总体$b(1,p)$中参数p的UMVUE.

解 对于此问题，由例1.4.1可知，统计量$T_n=\sum_{i=1}^n X_i$是充分的，而由例2.5.1可知，二项分布族是完备的，而$T_n\sim B(n,p)$，故知统计量T_n是充分完备的. 又由于$\bar X=T_n/n$是p的UE，且是充分完备统计量T_n的函数，故由定理2.5.1知，$\bar X$是p的UMVUE. □

例 2.5.5 对于Poisson分布$P(\lambda)$, 求参数

$$P_\lambda(k)=\frac{\lambda^k}{k!}e^{-\lambda},\ k=0,1,2,\cdots,$$

的UMVUE.

解 由例1.4.1可知，统计量$T_n=\sum_{i=1}^n X_i$是充分的，另外，我们也可以证明(在下一小节)此统计量也是完备的. 于是，由定理2.5.1知，为求$P_\lambda(k)$的UMVUE，仅需找一个依赖于充分完备量T_n的UE即可.

事实上，由Poisson分布的可加性知，$T_n\sim P(n\lambda)$，且容易验证统计量

$$\psi_k(\mathbf{X})=\begin{cases}1, & X_1=k,\\ 0 & X_1\ne k\end{cases}$$

是$P_\lambda(k)$的一个无偏估计. 于是，由定理2.5.1知，$P_\lambda(k)$的UMVUE为

$$E_\lambda[\psi_k(\mathbf{X})|T_n].$$

为求上述的UMVUE，我们注意到

$$E_\lambda[\psi_k(\mathbf{X})|T=t]=P\{X_1=k|T=t\}=\frac{P\{X_1=k,T=t\}}{P\{T=t\}}=\frac{P\{X_1=k,X_2+\cdots+X_n=t-k\}}{P\{T=t\}},$$

且$X_1 \sim P(\lambda)$与$X_2 + \cdots + X_n \sim P((n-1)\lambda)$独立，则代入上式后，有

$$E_\lambda[\psi_k(\mathbf{X})|T=t] = \begin{pmatrix} t \\ k \end{pmatrix} \left(\frac{1}{n}\right)^k \left(1-\frac{1}{n}\right)^{t-k},\ k=0,1,\cdots.$$

于是，$P_\lambda(k)$的UMVUE为

$$\begin{pmatrix} T_n \\ k \end{pmatrix} \left(\frac{1}{n}\right)^k \left(1-\frac{1}{n}\right)^{T_n-k},\ k=0,1,\cdots,$$

其中$T_n = \sum_{i=1}^n X_i$. □

例 2.5.6 设$X_1, \ldots, X_n$为来自$\Gamma(\alpha, \lambda)$的iid样本，且α已知，试求$\lambda > 0$的UMVUE.

解 由例1.4.8知，$T_n = \sum_{i=1}^n X_i$是λ的充分统计量. 又因为$T_n \sim \Gamma(n\alpha, \lambda)$，故由$\Gamma$分布族的完备性(下一小节将证明)知，它也是完备的. 于是，由定理2.5.1知，为求λ的UMVUE，仅需找一个依赖于T_n的UE即可.

事实上，由于

$$E_\lambda\left(\frac{1}{T}\right) = \int_0^\infty \frac{1}{t}\frac{\lambda^{n\alpha}}{\Gamma(n\alpha)} t^{n\alpha-1} e^{-\lambda t} \mathrm{d}t = \frac{\lambda}{n\alpha-1},$$

则知$(n\alpha-1)/\sum_{i=1}^n X_i = (n\alpha-1)/(n\bar{X})$是$\lambda$的UMVUE. □

§2.5.3 指数型分布族的充分完备性

从上一小节可以看出，充分完备统计量在求参数的UMVUE时非常有用，而统计量的充分性可容易由因子分解定理验证，但完备性的验证就要困难多了. 由于许多我们常用的分布均属于指数型分布族，故本节将考虑指数型分布族的充分完备性问题.

定理 2.5.2 设$\{f(x,\eta): \eta \in \Omega\}$是$k$个参数的标准指数型分布族，即

$$f(x,\eta) = c(\eta) \exp\left\{\sum_{j=1}^k \eta_j T_j(x)\right\} h(x), \tag{2.5.3}$$

其中$\eta = (\eta_1, \ldots, \eta_k)' \in \Omega$, 这里

$$\Omega = \left\{(\eta_1, \ldots, \eta_k): 0 < \int_{\mathcal{X}} \exp\left\{\sum_{j=1}^k \eta_j T_j(x)\right\} h(x)\mathrm{d}x < \infty\right\} \tag{2.5.4}$$

称为自然参数空间. 则

- 统计量$T(\mathbf{X}) = (T_1, \ldots, T_k)$是充分的;
- 如果Ω有内点，则$T(\mathbf{X}) = (T_1, \ldots, T_k)$是完备的.

证明 见陈希孺(1997).

例 2.5.7 考虑Poisson分布族$P(\lambda)$的UMVUE.

解 由Poisson分布的可加性知，统计量$T_n=\sum_{i=1}^n X_i\sim P(n\lambda)$，且其CDF为

$$P\{T=t\}=\frac{(n\lambda)^t}{t!}e^{-n\lambda}=e^{-n\lambda}\exp\{t\ln(n\lambda)\}/t!,\ t=0,1,\cdots.$$

如令$\eta=\ln(n\lambda)$，则知其自然参数空间为$\Omega=R$，它有内点，故由定理2.5.2知，统计量$T=\sum_{i=1}^n X_i$是充分完备的. 另外，由于$\bar{X}=T/n$是λ的UE，故它是λ的UMVUE. □

例 2.5.8 *求Γ分布族$\Gamma(\alpha,\lambda)$的充分完备统计量.*

解 设$X_1,\ldots,X_n$为来自$\Gamma(\alpha,\lambda)$的iid样本，由其联合PDF为

$$f(\mathbf{x};\alpha,\lambda)=\frac{1}{\alpha^\lambda\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}=\frac{1}{\alpha^\lambda\Gamma(\alpha)}\exp\left\{(\alpha-1)\sum_{i=1}^n\ln x_i-\lambda\sum_{i=1}^n x_i\right\},$$

于是，由定理2.5.2可知，统计量$(\sum_{i=1}^n X_i,\sum_{i=1}^n\ln X_i)$ 是(α,λ)的充分完备统计量. □

例 2.5.9 *考虑正态总体$N(\mu,\sigma^2)$中参数的UMVUE.*

解 容易验证正态分布族是指数型分布，且$(\sum_{i=1}^n X_i,\sum_{i=1}^n X_i^2)$是$(\mu,\sigma^2)$的充分完备统计量. 于是，由于样本均值$\bar{X}$，样本方差$S_n^2$分别是$\mu,\sigma^2$的UE，故它们是UMVUE. □

§2.5.4 次序统计量的完备性

关于UMVUE的求取，还有另一种很常用的方法，即U统计量，它是基于次序统计量的充分完备性而得到所需要参数的UMVUE的. 为此，我们在本小节不加证明地给出次序统计量完备性的定理.

定理 2.5.3 (次序统计量的完备性) *设$X_1,\ldots,X_n$是来自分布族$\mathcal{F}$的分布函数为F的iid样本，如果*

(1) *$\mathcal{F}$是凸的;*

(2) *$\forall\ a<b$, 记$S=[a,b)$，由$F(b)-F(a)>0$可导出$P\{X_1<x|X_1\in S\}\in\mathcal{F}$,*

则该样本的次序统计量$X_{(1)},\ldots,X_{(n)}$关于分布族$\mathcal{F}$是完备的.

证明 请参见David (1981).

结合定理1.4.2和定理2.5.3，我们知道，在许多情况下，次序统计量均是充分完备的. 并由可以看到，前面所求取的UMVUE都可以由次序统计量的充分完备性得到证明.

在结束本节前，我们看一个UMVUE不好的例子.

例 2.5.10 *设$X_1,\ldots,X_n$为来自Poisson分布$P(\lambda)$的iid样本，求参数$g(\lambda)=e^{-a\lambda}$的UMVUE, 其中a为已知常数.*

解 设$S(T)$为$g(\lambda)$的UE，由于此时的充分完备量为$T=\sum_{i=1}^n X_i\sim P(n\lambda)$，则它应满足

$$\sum_{x=0}^\infty\frac{S(t)(n\lambda)^t}{t!}=e^{(n-a)\lambda}=\sum_{x=0}^\infty\frac{(n-a)^t(\lambda)^t}{t!},$$

于是，

$$S(T) = (1 - a/n)^T$$

是$g(\lambda)$的UMVUE. 但是当$a > n$这个UMVUE并不合理. (此例子类似于例2.4.2关于UE不合理的例子)

§2.6 信息不等式及有效估计

在前面我们给出了几种求取UMVUE的方法，既然我们得到的UMVUE的方差是最小的，那我们能否给出这个最小方差的一般表达式吗?关于这个问题的解决，Fréchet 于1943年最早把信息不等式引进来(*Rev. Int. Statist.*, 11, 182–205)，后来Darmois(1945), Rao (1945)和Cramér (1946)对此进行了推广和创新，于是，有的书上也称本节的信息不等式为C-R不等式.

§2.6.1 正则分布族与Fisher信息量

本节讨论的内容均是针对满足一定条件的分布族而言的，这即是本小节的正则分布族.

定义 2.6.1 (正则分布) 如果单参数分布族$\mathcal{F} = \{f(x,\theta) : \theta \in \Theta\}$具有如下五个条件:

(1) 参数空间Θ是直线上的开区间(有限、无限或半无限);

(2) 导数$\frac{\partial f(x,\theta)}{\partial \theta}$存在($\forall\, \theta \in \Theta$);

(3) 支撑集与θ无关;

(4) 其PDF $f(x,\theta)$的积分与微分运算可以互换，即

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\int_{-\infty}^{+\infty} f(x,\theta)\mathrm{d}x = \int_{-\infty}^{+\infty} \frac{\mathrm{d}}{\mathrm{d}\theta} f(x,\theta)\mathrm{d}x.$$

(5)

$$I(\theta) = E_\theta\left(\frac{\partial}{\partial\theta}\ln f(X,\theta)\right)^2 \text{ 存在，且} I(\theta) > 0, \tag{2.6.1}$$

则称此分布族为C-R正则分布族，其中条件(1)–(5)也称为正则条件，$T(\theta)$ 称为该分布族的Fisher 信息量(Information).

例 2.6.1 验证Poisson分布族是C-R正则分布族.

解 容易验证正则条件(1)–(4)满足，下证条件(5)成立. 事实上，因为

$$\ln f(x,\lambda) = \ln\left(\frac{\lambda^x}{x!}e^{-\lambda}\right) = x\ln\lambda - \ln x! - \lambda,$$

于是

$$\frac{\partial \ln f(x,\lambda)}{\partial\lambda} = \frac{x}{\lambda} - 1,$$

$$I(\lambda) = E\left(\frac{\partial \ln f(X,\lambda)}{\partial\lambda}\right)^2 = E\left(\frac{X}{\lambda} - 1\right)^2 = \frac{1}{\lambda} > 0,$$

故Poisson分布满足正则条件. □

例 2.6.2 验证正态分布族$\{N(\mu,1):\mu\in R\}$是正则的.

解 容易验证条件(1)–(4)满足，又因为

$$f(x,\mu)=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(x-\mu)^2}{2}\right\},$$

$$\ln f(x,\mu)=-\frac{(x-\mu)^2}{2}-\ln(2\pi)/2,\ \frac{\partial\ln f(x,\mu)}{\partial\mu}=x-\mu,$$

所以，

$$I(\mu)=E_\mu\left(\frac{\partial\ln f(X,\mu)}{\partial\mu}\right)^2=E_\mu(X-\mu)^2=1\neq 0,$$

故正则条件(5)也满足. □

可以验证，我们常用的单参数指数型分布族是正则的. 当然，常用的分布中也有不是正则的，如均匀分布族$\{U(0,\theta):\theta>0\}$.

另外，我们注意到，定义2.6.1是以连续分布为例进行定义的，如果是离散的，也有类似的定义，只是把积分换成求和、把PDF换成CDF即可.

如在(2.6.1)中考虑iid样本的联合PDF，则可以证明

$$E_\theta\left(\frac{\partial}{\partial\theta}\ln f(\mathbf{X},\theta)\right)^2=nI(\theta). \tag{2.6.2}$$

值得注意的是，Fisher信息量$I(\theta)$依赖于我们所选择的参数. 事实上，设$\theta=h(\xi)$, 其中h可微，则X所包含的关于ξ的信息量就是

$$I^*(\xi)=I(h(\xi))[h'(\xi)]^2. \tag{2.6.3}$$

为理解Fisher信息量的意义，我们先考虑一个估计量的方差下界. 设S是$g(\theta)$的一个估计，$\psi(X,\theta)$是任一具有有限二阶矩的函数，则容易证明如下的协方差不等式：

$$\mathrm{Var}(S)\geq\frac{[\mathrm{Cov}(S,\psi)]^2}{\mathrm{Var}(\psi)}. \tag{2.6.4}$$

虽然上一不等式给出了估计量S的方差下界，但对于实际应用没有任何意义，由于此不等式右边包含S. 但若$\mathrm{Cov}(S,\psi)$只与$E_\theta S=g(\theta)$有关，则(2.6.4)式的确提供了无偏估计S的方差的一个下界. 于是，我们有如下结论.

定理 2.6.1 (Blyth定理) $\mathrm{Cov}(S,\psi)$只通过$g(\theta)=E_\theta(S)$依赖于S的充要条件是：$\forall\,\theta$，有

$$\mathrm{Cov}(U,\psi)=0,\ \forall\,U\in U_0, \tag{2.6.5}$$

其中U_0为零的无偏估计类(见(2.4.9)式).

证明 “$\mathrm{Cov}(S,\psi)$只通过$g(\theta)=E_\theta(S)$依赖于S”等价于对任何两个满足条件$E_\theta S_1=E_\theta S_2$的估计$S_1,S_2$，必有$\mathrm{Cov}(S_1,\psi)=\mathrm{Cov}(S_2,\psi)$.

而对于任两个无偏估计S_1,S_2,

$$0=\mathrm{Cov}(S_1,\psi)-\mathrm{Cov}(S_2,\psi)=\mathrm{Cov}(S_1-S_2,\psi)\Longleftrightarrow\mathrm{Cov}(U,\psi)=0,\ \forall\,U\in U_0.$$

□

下面我们看一个满足(2.6.5)式条件的例子，即寻找$\psi(x,\theta)$.

设总体X的PDF为$f(x,\theta)$,且对一切x, $f(x,\theta)>0$. 又设$\theta,\theta+\delta$为两个参数值，且$f(x,\theta)\neq f(x,\theta+\delta)$. 则函数

$$\psi(x,\theta)=\frac{f(x,\theta+\delta)-f(x,\theta)}{f(x,\theta)} \tag{2.6.6}$$

满足定理2.6.1的条件. 这是由于

$$E_\theta(\psi)=\int[f(x,\theta+\delta)-f(x,\theta)]\mathrm{d}x=1-1=0,$$

而 $\forall\ U\in U_0$，我们有

$$\mathrm{Cov}(U,\psi)=E_\theta(U\psi)=E_\theta\left[U\left(\frac{f(x,\theta+\delta)}{f(x,\theta)}-1\right)\right]=E_{\theta+\delta}U-E_\theta U=0.$$

这就是说，对于此分布族$f(x,\theta)$，$\mathrm{Cov}(S,\psi)$只通过$g(\theta)=E_\theta(S)$依赖于S，又由于

$$\mathrm{Cov}(S,\psi)=E_\theta(S\psi)=g(\theta+\delta)-g(\theta),$$

故(2.6.4)式的协方差不等式成为

$$\mathrm{Var}(S)\geq\frac{[g(\theta+\delta)-g(\theta)]^2}{E_\theta[\psi^2(x,\theta)]}=\frac{[g(\theta+\delta)-g(\theta)]^2}{E_\theta\left[\frac{f(x,\theta+\delta)}{f(x,\theta)}-1\right]^2}. \tag{2.6.7}$$

从上面的证明过程可以看出，不等式(2.6.7)对一切满足$f(x,\theta+\delta)\neq f(x,\theta)$的$\delta$都成立，于是，不等式(2.6.7)式的右端可以用其关于δ的极大值代入. 这就是Hammersley-Chapman-Robbins不等式.

如果(2.6.6)式的ψ用下面的函数

$$\frac{f(x,\theta+\delta)-f(x,\theta)}{\delta}\cdot\frac{1}{f(x,\theta)} \tag{2.6.8}$$

代替，则不等式(2.6.7)式仍保持不变. 现令$\delta\to0$，且假定$f(x,\theta)$关于θ可导，于是，我们有

$$\frac{f(x,\theta+\delta)-f(x,\theta)}{\delta}\cdot\frac{1}{f(x,\theta)}\to\frac{\partial f(x,\theta)/\partial\theta}{f(x,\theta)}=\frac{\partial\ln f(x,\theta)}{\partial\theta}.$$

此极限启发我们可以用

$$\psi(x,\theta)=\frac{\partial\ln f(x,\theta)}{\partial\theta}=\frac{\partial f(x,\theta)/\partial\theta}{f(x,\theta)} \tag{2.6.9}$$

作为(2.6.6)式的ψ，而此时的ψ代表PDF $f(x,\theta)$在x处的相对改变率，其平方均值

$$I(\theta)=E_\theta\left[\frac{\partial\ln f(x,\theta)}{\partial\theta}\right]^2$$

即为(2.6.1)式的Fisher信息量. 并且，如把(2.6.8)式的ψ代入(2.6.7)式的Hammersley-Chapman-Robbins不等式右端，且令$\delta\to0$，则可以得到后面的信息不等式.

Efron, and Johnston 于1990年(*Ann. Statist.* 18, 38–62)指出，Fisher信息量与生存分析中非常有用的危险率函数(hazard function)之间有如下关系：

$$I(\theta)=\int_{-\infty}^{\infty}\left(\frac{\partial}{\partial\theta}\ln[h_\theta(x)]\right)^2f(x,\theta)\mathrm{d}x,$$

其中危险率函数$h_\theta(x)$定义为

$$h_\theta(x)=\lim_{\delta\to 0}\frac{P\{x\le X\le x+\delta|X\ge x\}}{\delta}=\frac{f(x,\theta)}{1-F_\theta(x)}.$$

如X是一寿命随机变量，则上述危险率函数表示在活到x的情况下，在x处生存的概率密度.

§2.6.2 信息不等式

本节将讨论正则分布族参数的无偏估计的方差的下界，即著名的信息不等式（许多书上也称之为C-R不等式）.

定理 2.6.2 (信息不等式) 设分布族$\mathcal{F}=\{f(x,\theta):\theta\in\Theta\}$是正则的，可估函数$g(\theta)$在$\Theta$上可微. 又设$X_1,\ldots,X_n$是$n$个来自此分布族的iid样本，$T(\mathbf{X})$是$g(\theta)$的一个UE，且满足积分与微分号可互换的条件，即

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\int T(\mathbf{x})f(\mathbf{x},\theta)\mathrm{d}\mathbf{x}=\int T(\mathbf{x})\frac{\mathrm{d}}{\mathrm{d}\theta}f(\mathbf{x},\theta)\mathrm{d}\mathbf{x},$$

则有

$$\mathrm{Var}_\theta(T(\mathbf{X}))\ge\frac{(g'(\theta))^2}{nI(\theta)},\ \forall\,\theta\in\Theta, \tag{2.6.10}$$

其中$I(\theta)$为$\mathcal{F}$的Fisher信息量. 如果上式等号成立，则存在$c(\theta)\neq 0$，使得

$$\frac{\partial\ln f(\mathbf{X},\theta)}{\partial\theta}=c(\theta)(T(\mathbf{X})-g(\theta)) \tag{2.6.11}$$

以概率1成立.

证明 如$I(\theta)=\infty$或$\mathrm{Var}_\theta(T)=\infty$，则(2.6.10)均显然成立. 故下面假设$I(\theta)<\infty,\mathrm{Var}_\theta(T)<\infty$.

因为

$$\int\frac{\partial\ln f(x,\theta)}{\partial\theta}f(x,\theta)\mathrm{d}x=\int\frac{\partial f(x,\theta)}{\partial\theta}\mathrm{d}x=\frac{\mathrm{d}}{\mathrm{d}\theta}\int f(x,\theta)\mathrm{d}x=0, \tag{$*1$}$$

又因为样本是iid的，且样本分布$f(\mathbf{x},\theta)=\prod_{i=1}^n f(x_i,\theta)$, 故由(*1)式知

$$\int\frac{\partial\ln f(\mathbf{x},\theta)}{\partial\theta}f(\mathbf{x},\theta)\mathrm{d}\mathbf{x}=0,\ \forall\,\theta\in\Theta,$$

于是，对于任意的$g(\theta)$，我们有

$$\int g(\theta)\frac{\partial\ln f(\mathbf{x},\theta)}{\partial\theta}f(\mathbf{x},\theta)\mathrm{d}\mathbf{x}=0,\ \forall\,\theta\in\Theta. \tag{$*2$}$$

又因为

$$\begin{aligned}
g'(\theta) &= \frac{\partial}{\partial\theta}g(\theta) = \frac{\partial}{\partial\theta}E_\theta T(\mathbf{X}) \\
&= \frac{\partial}{\partial\theta}\int T(\mathbf{x})f(\mathbf{x},\theta)\mathrm{d}\mathbf{x} \\
&= \int \frac{\partial}{\partial\theta}T(\mathbf{x})f(\mathbf{x},\theta)\mathrm{d}\mathbf{x} \\
&= \int T(\mathbf{x})\frac{\partial \ln f(\mathbf{x},\theta)}{\partial\theta}f(\mathbf{x},\theta)\mathrm{d}\mathbf{x} \\
&\overset{(*2)}{=} \int (T(\mathbf{x})-g(\theta))\frac{\partial \ln f(\mathbf{x},\theta)}{\partial\theta}f(\mathbf{x},\theta)\mathrm{d}\mathbf{x},\ \forall\ \theta\in\Theta.
\end{aligned}$$

于是，由Cauchy-Schwartz不等式知，对于 $\forall\ \theta\in\Theta$, 有

$$\begin{aligned}
[g'(\theta)]^2 &= \left[\int (T(\mathbf{x})-g(\theta))\frac{\partial \ln f(\mathbf{x},\theta)}{\partial\theta}f(\mathbf{x},\theta)\mathrm{d}\mathbf{x}\right]^2 \\
&\le \int (T(\mathbf{x})-g(\theta))^2 f(\mathbf{x},\theta)\mathrm{d}\mathbf{x}\cdot\int\left(\frac{\partial \ln f(\mathbf{x},\theta)}{\partial\theta}\right)^2 f(\mathbf{x},\theta)\mathrm{d}\mathbf{x} \\
&= \mathrm{Var}_\theta(T)\cdot nI(\theta),
\end{aligned}$$

且由Cauchy-Schwartz 不等式知，上述等号成立的条件为：存在$c(\theta)\neq 0$，使得

$$\frac{\partial \ln f(\mathbf{X},\theta)}{\partial\theta} = c(\theta)(T(\mathbf{X})-g(\theta))$$

以概率1成立. □

关于本定理，我们有如下几点注解:

注 2.6.1 我们称不等式(2.6.10)为信息不等式，也称C-R不等式.

注 2.6.2 在上一定理的证明过程中，我们始终假设样本是iid的. 如果样本不是独立的，上述结论仍成立，只是要把(2.6.10)中的$nI(\theta)$改为$E_\theta\left(\frac{\partial \ln f(\mathbf{X},\theta)}{\partial\theta}\right)^2$.

注 2.6.3 从(2.6.10)可以看出，信息不等式与Fisher信息量有着密切的关系. 从直观上看，一个带有较小方差的估计表明这个估计值有更大机会出现在$g(\theta)$附近. 为了考察$I(\theta)$的含义，我们假设信息不等式的下界可以达到，且$g(\theta)=\theta$，则此时UE的最小方差为$\frac{1}{nI(\theta)}$. 这说明，$nI(\theta)$越大，参数θ可以越精确地估计，其中n是样本容量，这也说明，如以估计量的方差的倒数作为估计量精度的指标，则精度与n成正比，而$I(\theta)$则反映着总体分布的性质. $I(\theta)$越大，说明总体本身提供的信息量越多.

注 2.6.4 从定理的证明不难看出，对于正则分布族，我们有如下结论:

$$E_\theta\left[\frac{\partial}{\partial\theta}\ln f(\mathbf{X},\theta)\right]=0,\quad I(\theta)=\mathrm{Var}_\theta\left[\frac{\partial}{\partial\theta}\ln f(\mathbf{X},\theta)\right]=-E_\theta\left[\frac{\partial^2}{\partial\theta^2}\ln f(\mathbf{X},\theta)\right]. \tag{2.6.12}$$

注 2.6.5 信息不等式也可以从直角三角形的边长关系得到. 设$\mathbf{t},\mathbf{q}$为任意两向量，定

义$\langle \mathbf{t},\mathbf{q}\rangle=\mathbf{t}'\mathbf{q}$, $\langle \mathbf{t},\mathbf{t}\rangle=|\mathbf{t}|^{\mathbf{2}}$.

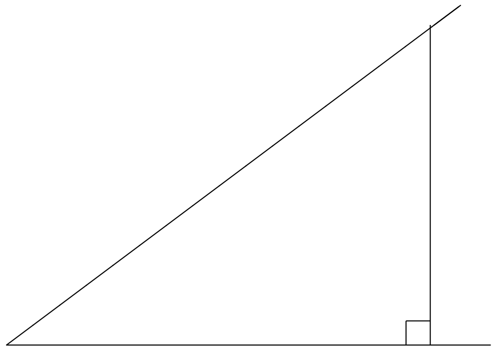

图2.1 信息不等式的几何解释

由图2.1中的三角形可以得到

$$|\mathbf{t}|>|\mathbf{t}|\cos(\mathbf{t},\mathbf{q})=|\mathbf{t}|\cdot\frac{\langle \mathbf{t},\mathbf{q}\rangle}{|\mathbf{t}||\mathbf{q}|}=\frac{\langle \mathbf{t},\mathbf{q}\rangle}{|\mathbf{q}|}.$$

如对于两个随机变量X,Y，定义$\langle X,Y\rangle=\mathrm{Cov}(X,Y)$，则上一不等式即为(2.6.4)式的协方差不等式，由此即可推得信息不等式.

定理2.6.2仅考虑了单参数的情形，而多参数在实际中也很常见，如正态总体$N(\mu,\sigma^2)$. 下面，我们不加证明地给出多参数情形下的信息不等式.

定理 2.6.3 设分布族$\mathcal{F}=\{f(x,\theta):\theta\in\Theta\subset R^k\}$, $X_1,\ldots,X_n$为来自此分布族的iid样本，$T(\mathbf{X})=(T_1(\mathbf{X}),\ldots,T_r(\mathbf{X}))'$是$\mathbf{g}(\theta)=(g_1(\theta),\ldots,g_r(\theta))'$的一个UE，且满足如下条件:

(1) $\frac{\partial}{\partial\theta_i}\int f(\mathbf{x},\theta)\mathrm{d}\mathbf{x}=\int\frac{\partial}{\partial\theta_i}f(\mathbf{x},\theta)\mathrm{d}\mathbf{x},\ \forall\,\theta\in\Theta,\ i=1,\cdots,k,$

$\frac{\partial^2}{\partial\theta_i\partial\theta_j}\int f(\mathbf{x},\theta)\mathrm{d}\mathbf{x}=\int\frac{\partial^2}{\partial\theta_i\partial\theta_j}f(\mathbf{x},\theta)\mathrm{d}\mathbf{x},\ \forall\,\theta\in\Theta,\ i,j=1,\cdots,k,$

(2) $\frac{\partial g_i(\theta)}{\partial\theta_j}$存在，且$\frac{\partial g_i(\theta)}{\partial\theta_i}=\frac{\partial}{\partial\theta_j}\int T_i(\mathbf{x})f(\mathbf{x},\theta)\mathrm{d}\mathbf{x}=\int T_i(\mathbf{x})\frac{\partial f(\mathbf{x},\theta)}{\partial\theta_j}\mathrm{d}\mathbf{x},\ \forall\,\theta\in\Theta,\ i,j=1,\cdots,k.$

记

$$I(\theta)=-E_\theta\left(\frac{\partial^2\ln f(\mathbf{X},\theta)}{\partial\theta_i\partial\theta_j}\right)_{k\times k},\ \Sigma_\theta(T)=(\mathrm{Cov}(T_i,T_j))_{r\times r},\ \Delta=\left(\frac{\partial g_i(\theta)}{\partial\theta_j}\right)_{r\times k},$$

则有

$$\Sigma_\theta(T)\geq\frac{1}{n}\Delta I^{-1}(\theta)\Delta',\ \forall\,\theta\in\Theta. \tag{2.6.13}$$

为了方便，我们常称信息不等式右端的值为相应参数UE的方差的C-R下界.

下面看几个例子.

例 2.6.3 设$X_1,\ldots,X_n$为来自Bernoulli分布$b(1,p)$的iid样本，$p\in(0,1)$. 试求p的UE的方差的C-R下界.

解 容易验证此分布族是正则的，且其Fisher信息量为

$$I(p)=E_p\left(\frac{\partial\ln f(X,p)}{\partial p}\right)^2=\sum_{k=0}^{1}\left(\frac{\partial\ln f(x,p)}{\partial p}\right)^2 f(k,p),$$

其中$f(k,p)=P\{X=k\}=p^k(1-p)^{1-k}$. 于是

$$I(p)=\left(\frac{\partial\ln(1-p)}{\partial p}\right)^2(1-p)+\left(\frac{\partial\ln p}{\partial p}\right)^2 p=\frac{1}{p(1-p)},$$

这样由信息不等式知，p的UE的方差的C-R下界为$\frac{1}{nI(p)}=\frac{p(1-p)}{n}$.

另外，由于$\sum_{i=1}^n X_i\sim B(n,p)$，故$\mathrm{Var}\bar{X}=\frac{p(1-p)}{n}$达到了C-R下界，故可知$\bar{X}$是$p$的UMVUE，这与我们前面讲的例2.5.4的结论一致. □

例 2.6.4 设$X_1,\ldots,X_n$为来自Poisson分布$P(\lambda)$的iid样本，考虑λ的UMVUE.

解 由例2.6.1可知，$P(\lambda)$是正则的，且其Fisher信息量为$I(\lambda)=\frac{1}{\lambda}$，于是$\lambda$的UE的方差的C-R下界为$\frac{\lambda}{n}$. 又由于$\sum_{i=1}^n X_i\sim P(n\lambda)$，则知$\mathrm{Var}\bar{X}=\frac{\lambda}{n}$，达到了C-R下界，故知样本均值$\bar{X}$是$\lambda$的UMVUE. □

例 2.6.5 设$X_1,\ldots,X_n$是来自指数分布$E(\lambda)$的iid样本，考虑$1/\lambda$的UMVUE.

解 容易验证此分布族是正则的，且由于

$$f(x,\lambda)=\lambda e^{-\lambda x},\ln f(x,\lambda)=\ln\lambda-\lambda x,\frac{\partial\ln f(x,\lambda)}{\partial\lambda}=\frac{1}{\lambda}-x,$$

则

$$I(\lambda)=E_\lambda\left(\frac{\partial\ln f(X,\lambda}{\partial\lambda}\right)^2=E_\lambda\left(\frac{1}{\lambda}-X\right)^2=\frac{1}{\lambda^2}.$$

于是，$1/\lambda$的UE的方差的C-R下界为

$$\frac{(g'(\lambda))^2}{nI(\lambda)}=\frac{1}{n\lambda}.$$

因为$X_i\sim E(\lambda)=\Gamma(1,\lambda)$，而$\Gamma$分布关于形状参数具有可加性，所以，$\sum_{i=1}^n X_i\sim\Gamma(n,\lambda)$. 于是，$\mathrm{Var}(\sum_{i=1}^n X_i)=\frac{n}{\lambda^2}$，即$\mathrm{Var}\bar{X}=\frac{1}{n\lambda^2}$达到了C-R下界，故知$\bar{X}$是$1/\lambda$的UMVUE，这与例2.4.7的结果一样. □

例 2.6.6 设$X_1,\ldots,X_n$是来自$N(\mu,\sigma^2)$的iid样本，试考虑μ和σ^2的UE的方差的下界.

解 容易验证此时的正态分布族是正则的，且因为

$$f(x,\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\},\ \ln f(x,\mu,\sigma^2)=-\frac{(x-\mu)^2}{2\sigma^2}-\frac{1}{2}\ln\sigma^2-\frac{1}{2}\ln 2\pi,$$

则

$$\begin{aligned}
&\frac{\partial\ln f(x,\mu,\sigma^2)}{\partial\mu}=\frac{x-\mu}{\sigma^2}, && \frac{\partial^2\ln f(x,\mu,\sigma^2)}{\partial\mu^2}=-\frac{1}{\sigma^2},\\
&\frac{\partial\ln f(x,\mu,\sigma^2)}{\partial\sigma^2}=\frac{(x-\mu)^2}{2\sigma^4}-\frac{1}{2\sigma^2}, && \frac{\partial^2\ln f(x,\mu,\sigma^2)}{\partial(\sigma^2)^2}=-\frac{(x-\mu)^2}{\sigma^6}+\frac{1}{2\sigma^4},\\
&\frac{\partial^2\ln f(x,\mu,\sigma^2)}{\partial\mu\sigma^2}=-\frac{x-\mu}{\sigma^4},
\end{aligned}$$

于是，此时的Fisher信息阵为

$$I(\mu,\sigma^2) = -E_{\mu,\sigma^2}\begin{pmatrix} -\frac{1}{\sigma^2} & -\frac{X-\mu}{\sigma^4} \\ -\frac{X-\mu}{\sigma^4} & \frac{1}{2\sigma^4}-\frac{(X-\mu)^2}{\sigma^6} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{pmatrix},$$

由于可知

$$I(\mu)=1/\sigma^2, I(\sigma^2)=1/2\sigma^4,$$

则μ和σ^2的UE的方差的C-R下界分别为$\sigma^2/n, 2\sigma^4/n$.

因为$\text{Var}\bar{X}=\sigma^2/n$达到了C-R下界，故$\bar{X}$是$\mu$的UMVUE.

另外，由例2.4.6知，样本方差S_n^2是σ^2的UMVUE，但是$\text{Var}(S_n^2) = \frac{2\sigma^4}{n-1} > \frac{2\sigma^4}{n}$，故知$\sigma^2$的UMVUE 并没有达到C-R下界. □

上一个例子说明我们不能用一个UE的方差是否达到C-R下界来判断它是否是UMVUE，并且也说明，定理2.6.2与定理2.6.3给出的C-R下界并不是最大的一个，还有必要进行改进. 另外，我们也有必要对达到与没有达到C-R下界的UE进行一定的区别.

下面我们通过一个例子看一看C-R下界与最小方差究竟有多大的区别.

例 2.6.7 设$X_1,\ldots,X_n$是来自正态总体$N(0,\sigma^2)$的IID样本，其中$\sigma>0$为未知参数. 试求σ^2及σ的UMVUE.

解 对于此分布族，其PDF为$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-x^2/2\sigma^2}$, 故

$$\ln f(x) = -\ln(2\pi\sigma^2)/2 - x^2/2\sigma^2,$$

$$\frac{\partial \ln f(x)}{\partial\sigma^2} = -\frac{1}{2\sigma^2}+\frac{x^2}{2\sigma^4}, \frac{\partial^2 \ln f(x)}{\partial(\sigma^2)^2} = \frac{1}{2\sigma^4}-\frac{x^2}{\sigma^6},$$

于是，其Fisher信息量为

$$I(\sigma^2) = -E\left[\frac{1}{2\sigma^4}-\frac{X^2}{\sigma^6}\right] = \frac{1}{2\sigma^4}.$$

对于参数σ^2，其C-R下界为$2\sigma^4/n$. 对于它的一个估计量$T(X)=\sum_{i=1}^n X_i^2/n$, 因为$ET(X)=\sigma^2$, $\text{Var}T(X)=2\sigma^2/n$，故知它是一个UE且达到了C-R下界，于是它是$\sigma^2$的UMVUE.

对于参数σ，由于$g(x)=\sqrt{x}$, 故它的无偏估计的方差的C-R下界为$\frac{(1/(2\sigma))^2}{n/(2\sigma^2)} = \frac{\sigma^2}{2n}$.

因为$\sum_{i=1}^n X_i^2/\sigma^2 \sim \chi^2(n)$,故可以求得

$$E\left(\sqrt{\sum X_i^2}/\sigma\right) = \frac{\sqrt{2}\Gamma((n+1)/2)}{\Gamma(n/2)} = c_n,$$

于是知$T(X)=c_n^{-1}\sqrt{\sum_{i=1}^n X_i^2}$是$\sigma$的UE，由于它又是充分完备统计量$\sum X_i^2$的函数，故知它是$\sigma$的UMVUE. 另外，其方差为

$$\text{Var}T(X) = ET^2(X)-\sigma^2 = c_n^{-2}E\sum_{i=1}^n X_i^2 - \sigma^2 = (nc_n^{-1}-1)\sigma^2,$$

它与C-R下界是不同的. 对于给定的n，二者的比较见下表.

表2.2 C-R下界与最小方差的比较(例2.6.8)

n	1	2	5	10
$a = nc_n^{-1} - 1$	0.5706	0.2732	0.1405	0.0512
$b = 1/2n$	0.5	0.25	0.1	0.05
b/a	0.8760	0.9149	0.9572	0.9769

从表2.2可以看出，二者的差距随着n的增大而减小,事实上，利用Stirling公式可以证明$2n(nc_n^{-1} - 1) \longrightarrow 1$.

§2.6.3 有效估计

定义 2.6.2 (有效估计及效率) 设$T(\mathbf{X})$是$g(\theta)$的一个UE，则比值

$$e_n = \frac{(g'(\theta))^2/nI(\theta)}{\mathrm{Var}_\theta T(\mathbf{X})} \tag{2.6.14}$$

为$T(\mathbf{X})$的效率. 如果$e_n = 1$, 则称$T(\mathbf{X})$为$g(\theta)$的有效估计. 如$\lim_{n\to\infty} e_n = 1$, 则称$T(\mathbf{X})$为$g(\theta)$的渐近有效估计.

从有效估计的定义可以看出，有效估计的应用很有限制，如分布族应为正则的、待估参数有UE等. 下面我们看一个例子.

例 2.6.8 考虑均匀分布$U(0,\theta)$中参数θ的估计问题，其中$\theta > 0$.

解 此时总体PDF为

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

我们注意到，对于固定的$x > 0$，由于$f(x,\theta)$作为θ的函数在$\theta = x$处不连续，故它关于θ的导数不存在. 于是，此均匀分布族不是正则的，前面讲的C-R下界也不成立. 但如果我们不顾条件，而要利用定理2.6.2的结论，则会导致错误. 事实上，由于

$$\frac{\partial \ln f(x,\theta)}{\partial \theta} = -\frac{1}{\theta},\ \forall\ \theta > 0,$$

故$I(\theta) = 1\theta^2$, 从而C-R下界为θ^2/n.

但是，从例2.4.8可知，θ的UMVUE为$\frac{n+1}{n}X_{(n)}$. 计算一下可知,

$$\mathrm{Var}\left(\frac{n+1}{n}X_{(n)}\right) = \frac{\theta^2}{n(n+2)} < \frac{\theta^2}{n}\text{—C-R下界}.$$

§2.6.4 Bhattacharya下界

从前两节可以看出，有的UMVUE能达到C-R下界，但有的则不可以，这就说明C-R下界可能过低了些，于是，就有了本节的Bhattacharya下界. 下面的定理2.6.4仅是针对于单参数的情形，即$k = 1$.

定理 2.6.4 对于单参数分布族$\mathcal{F} = \{f(x,\theta) : \theta \in \Theta\}$, $g(\theta)$为定义在Θ上的有限实值函数，且$g'(\theta)$在Θ上处处存在. 如成立下列条件

(1) Θ是$\mathcal{R}$中的开子集;

(2) $\partial \ln f(x,\theta)/\partial\theta$对所有的$\theta\in\Theta$均存在;

(3) 其支持集$A=\{x: f(x,\theta)>0\}$与θ无关;

(4) $\int_{\mathcal{X}} \frac{\partial^i f(x,\theta)}{\partial\theta^i}\mathrm{d}x=0,\ i=1,\cdots,m,\theta\in\Theta$;

(5) $\int_{\mathcal{X}} \frac{1}{f(x,\theta)}\left(\frac{\partial^i f(x,\theta)}{\partial\theta^i}\right)^2\mathrm{d}x<\infty,\ i=1,\cdots,m,\theta\in\Theta$;

$T(\mathbf{X})$是$g(\theta)$的UE，其方差有限，且

(6) $\int_{\mathcal{X}} T(x)\left(\frac{\partial^i f(x,\theta)}{\partial\theta^i}\right)\mathrm{d}x=\frac{\partial^i}{\partial\theta^i}\int_{\mathcal{X}} T(x)f(x,\theta)\mathrm{d}x=g^{(i)}(\theta),\ i=1,\cdots,m,\theta\in\Theta$;

记

$$V_{ij}(\theta)=E_\theta\left[\frac{\partial^i \ln f(X,\theta)}{\partial\theta^i}\frac{\partial^j \ln f(X,\theta)}{\partial\theta^j}\right],\ i,j=1,\cdots,m.$$

$$\mathbf{V}(\theta)=(V_{ij}(\theta))_{m\times m},\ \mathbf{g}(\theta)=(g'(\theta),g''(\theta),\cdots,g^{(m)}(\theta))',\theta\in\Theta.$$

则当在$\theta\in\Theta$处$\mathbf{V}(\theta)$的行列式不为0时，必有

$$\mathrm{Var}_\theta(T(X))\geq \mathbf{g}'(\theta)\mathbf{V}^{-1}(\theta)\mathbf{g}(\theta), \tag{2.6.15}$$

(2.6.15)称为Bhattacharya不等式. 显然，C-R不等式为其$m=1$时的特例.

此定理的证明请参见陈希孺(1997). 数理统计引论，科学出版社(P119).

对于不同的m，(2.6.15)式形成一系列的下界，且可以证明，随着m的增大，这些下界是非降的.

§2.7 相合估计

前几节讲的内容，均是针对固定样本容量n而讨论估计的性质. 实际上，估计的大样本性质也是人们非常关心的一个问题，本节则将从大样本角度考虑上面几个估计的优良. 由于点估计的大样本理论非常丰富，而本讲义主要讲述基本统计方法，故本节将只考虑相合估计(Consistent Estimation)和相合渐近正态估计(Consistent Asymptotic Normal，简记为CAN)的基本概念.

§2.7.1 相合估计

定义 2.7.1 设统计量T_n是总体参数$g(\theta)$的估计量.

- 如果当$n\to\infty$时，T_n依概率收敛于$g(\theta)$，即$\forall\ \theta\in\Theta$及$\epsilon>0$，有
$$\lim_{n\to\infty} P\{|T_n-g(\theta)|\geq\epsilon\}=0, \tag{2.7.1}$$
则称T_n是$g(\theta)$的(弱)相合估计.

- 如果当$n\to\infty$时，T_n以概率1收敛于$g(\theta)$，即$\forall\ \theta\in\Theta$，有
$$P\{\lim_{n\to\infty} T_n=g(\theta)\}=1, \tag{2.7.2}$$
则称T_n是$g(\theta)$的强相合估计.

- 如果当$n\to\infty$时，T_n依r阶矩收敛于$g(\theta)$，即$\forall\,\theta\in\Theta$，有
$$\lim_{n\to\infty}E_\theta|T_n-g(\theta)|^r=0,\tag{2.7.3}$$
则称T_n是$g(\theta)$的r阶矩相合估计. 当$r=2$时，称为均方相合估计.

由概率论知识知道，强相合$\Rightarrow$弱相合，r阶矩相合$\Rightarrow$弱相合，反之不成立，且强相合与r阶相合之间没有包含关系.

例 2.7.1 设$X_1,\ldots,X_n$是来自$b(1,p)$ 的iid样本，则由大数定律知
$$\frac{1}{n}\sum_{i=1}^n X_i\xrightarrow{P}p,\ \forall\ 0<p<1,$$
于是，$\bar X$是p的相合估计. □

事实上，只要样本$X_1,\ldots,X_n$是iid的，且其期望$EX_1=\mu$存在，则由Khinchin大数定律知，样本均值$\bar X$就是总体均值μ的相合估计，而与其具体分布无关.

由概率论知识，容易得到如下一些结论.

定理 2.7.1 设$X_1,\ldots,X_n$是来自分布族$\{f(x,\theta):\theta\in\Theta\}$的iid样本，且$E|X_1|^p<\infty$ (p为正整数)，则样本的$k(1\le k\le p)$ 阶原点矩是总体k阶原点矩的相合估计，即
$$a_k=\frac{1}{n}\sum_{i=1}^n X_i^k\xrightarrow{P}\mu_k=EX^k.$$

定理 2.7.2 如果T_n是$g(\theta)$的相合估计，c_n,d_n是两个常数列，且$\lim_n c_n=0,\lim_n d_n=1$，则$d_nT_n+c_n$也是$g(\theta)$的相合估计.

定理 2.7.3 如果T_n是$g(\theta)$的渐近无偏估计，且满足
$$\lim_{n\to\infty}\mathrm{Var}_\theta T_n=0,\ \forall\ \theta\in\Theta,$$
则T_n既是$g(\theta)$的相合估计，也是均方相合估计.

定理 2.7.4 如果T_n是θ的相合估计，$g(x)$在点$x=\theta$处连续，则$g(T_n)$也是$g(\theta)$的相合估计.

由上述定理很容易知道，样本方差是总体方差的相合估计，多数矩估计也是相应参数的相合估计. 关于矩估计的相合性，请参见Serfling (1980) §2.2.

§2.7.2 样本分位数的相合性

由于分位数在数理统计中起着非常重要的作用，而我们在第一章中讲述了样本p分位数的概念，下面看一看其大样本性质如何. 为此，我们先引进总体p位数的概念.

定义 2.7.2 (总体p分位数) 设总体X的分布函数为$F(x)$，对于给定的$p\in(0,1)$，称满足
$$\xi_p=\inf\{x:F(x)\ge p\},\tag{2.7.4}$$
的ξ_p为该总体的p分位数. 特别地，$\xi_{0.5}$称为该总体的中位数.

定理 2.7.5 (样本分位数的Bahadur表示) 以$F_n(x)$表示经验分布函数，ξ_p表示总体的p分

位数，$f(x)$表示总体的PDF. 如果$f(\xi_p)>0$且$f(x)$在ξ_p点连续，则当$n\to\infty$时，有

$$\sqrt{n}\left(m_{n,p}-\xi_p+\frac{F_n(\xi_p)-p}{f(\xi_p)}\right)\xrightarrow{P}0. \tag{2.7.5}$$

此定理的证明请参见陈希孺等(1989), 非参数统计，上海科技出版社.

§2.7.3 极大似然估计的相合性

本小节将讨论MLE的相合性，为此假设$X_1,\ldots,X_n$为来自$f(x;\theta)$的IID样本，为简单起见，我们只考虑单参数的情况，设参数空间Θ是一个开区间，对数似然函数为$l(\theta;x)=\sum_{i=1}^n\ln f(x_i;\theta)$.

定理 2.7.6 如$\ln f(x;\theta)$在Θ上可微，并设$f(x;\theta)$是可识别的，即$\forall\theta\neq\theta',\{x;f(x;\theta)\neq f(x;\theta')\}$不是零测集. 则似然方程在$n\to\infty$时以概率1有解，且此解关于$\theta$是相合的.

证明 因为$f(x;\theta)$是可识别的，故对任意的$\theta'\neq\theta$，由Jensen不等式有

$$E_\theta\left(\ln\frac{f(x;\theta')}{f(x;\theta)}\right)<\ln E_\theta\left(\frac{f(x;\theta')}{f(x;\theta)}\right)=\ln\int f(x;\theta')dx=0.$$

记参数真值为θ_0，任给$\delta>0$，因为Θ是开区间，故$(\theta_0-\delta,\theta_0+\delta)\subset\Theta$，且由上式知

$$E_{\theta_0}\ln\frac{f(x;\theta_0-\delta)}{f(x;\theta_0)}<0,\ E_{\theta_0}\ln\frac{f(x;\theta_0+\delta)}{f(x;\theta_0)}<0.$$

因为$l(\theta;x)=\sum_{i=1}^n\ln f(x_i;\theta)$，故由大数定律有，当$n\to\infty$时

$$\frac{1}{n}(l(\theta_0-\delta;x)-l(\theta_0;x))=\frac{1}{n}\sum\ln\frac{f(x_i;\theta_0-\delta)}{f(x_i;\theta_0)}\xrightarrow{a.s.}E_{\theta_0}\ln\frac{f(x;\theta_0-\delta)}{f(x;\theta_0)}<0,$$

$$\frac{1}{n}(l(\theta_0+\delta;x)-l(\theta_0;x))=\frac{1}{n}\sum\ln\frac{f(x_i;\theta_0+\delta)}{f(x_i;\theta_0)}\xrightarrow{a.s.}E_{\theta_0}\ln\frac{f(x;\theta_0+\delta)}{f(x;\theta_0)}<0.$$

由于$l(\theta;x)$在$(\theta_0-\delta,\theta_0+\delta)$上连续，故当$n$充分大时，它必有一局部最大值，记其为$\hat\theta$. 又由于$l(\theta;x)$可导，于是在最大值点处有$\frac{\partial l}{\partial\theta}|_{\hat\theta}=0$，从而当$n$充分时，似然方程必有解. 又由于$\delta$的任意性及$|\theta_0-\hat\theta|<\delta$, 故可知此最大值解是相合的. □

上一定理给出了似然方程有解且其解是相合的条件. 当然，从证明可以看出，这个解有可能是局部极大值. 当然也存在不相合的极大似然估计的情况，相应的例子请见Lehmann and Casella (1998) P385，及下面的例子.

例 2.7.2 设$X_1,\ldots,X_n$是来自两点分布

$$P\{X=1\}=1-P\{X=0\}=\begin{cases}\theta, & \theta\text{为有理数},\\ 1-\theta, & \theta\text{为无理数}.\end{cases}$$

的IID样本，其中$0<\theta<1$. 试讨论θ的MLE的相合性.

解 记$T_n=\sum X_i$，则其似然函数为

$$L(\theta;x)=\begin{cases}\theta^{T_n}(1-\theta)^{n-T_n}, & \theta\text{为有理数},\\ \theta^{n-T_n}(1-\theta)^{T_n}, & \theta\text{为无理数}.\end{cases}$$

由此可以求出θ的MLE为$\hat{\theta} = T_n/n$. 但是由大数定律知，

$$\frac{T_n}{n} \xrightarrow{P} \begin{cases} \theta, & \theta\text{为有理数}, \\ 1-\theta, & \theta\text{为无理数}, \end{cases}$$

由此可知它的MLE并不是相合的.

§2.7.4 相合渐近正态估计

定义 2.7.3 (CAN) 设T_n是参数$g(\theta)$的相合估计量，如存在与样本容量n有关的定义于参数空间Θ上的函数$\mu(\theta), \sigma_n(\theta)$, 且$\sigma_n(\theta) > 0$，使得当$n \to \infty$时有

$$\frac{T_n - \mu_n(\theta)}{\sigma_n(\theta)} \xrightarrow{\mathcal{L}} N(0,1), \tag{2.7.6}$$

则称T_n为$g(\theta)$的CAN估计，也称T_n渐近正态$N(\mu_n, \sigma_n^2)$, 记为$T_n \sim AN(\mu_n, \sigma_n^2)$.

关于统计量T_n的渐近正态性，由概率论知识知道，其渐近的均值与方差显然不是惟一的. 我们有下面的两个结论：

- If T_n is $AN(\mu_n, \sigma_n^2)$, then also T_n is $AN(\mu_n', {\sigma'}_n^2)$ if and only if
$$\frac{\sigma_n'}{\sigma_n} \to 1, \quad \frac{\mu_n' - \mu_n}{\sigma_n} \to 0.$$
- If T_n is $AN(\mu_n, \sigma_n^2)$, then also $a_nT_n + b_n$ is $AN(\mu_n, \sigma_n^2)$ if and only if
$$\mu_n \to 1, \quad \frac{\mu_n(a_n - 1) + b_n}{\sigma_n} \to 0.$$

利用概率论知识，可以证明样本均值、样本方差、样本标准差都是渐近正态的. 详细的证明请见茆诗松、王静龙、濮晓龙(2000). 下面我们将不加证明地给出样本分位数、矩估计和极大似然估计的渐近相合正态性.

一、样本分位数的相合渐近正态性

定理 2.7.7 设ξ_p表示总体的p分位数，$f(x)$表示总体的PDF. 如果$f(\xi_p) > 0$且$f(x)$在ξ_p点连续，则当$n \to \infty$时，有

$$\sqrt{n}(m_{n,p} - \xi_p) \xrightarrow{\mathcal{L}} N(0, p(1-p)/f^2(\xi_p)). \tag{2.7.7}$$

此定理的证明请参见Serfling (1980).

二、矩估计的相合渐近正态性

设$X_1, \ldots, X_n$为来自某参数分布族$\{f(x,\theta) : \theta \in \Theta\}$的iid样本，且待估参数$g(\theta)$可以表示为总体$k$阶原点矩$\mu_k$的函数，即

$$g(\theta) = G(\mu_1, \ldots, \mu_k).$$

如用样本k阶原点矩估计总体k阶原点矩，则记得到的关于$g(\theta)$的估计为

$$\hat{g}(\mathbf{X}) = G(a_1, \ldots, a_k).$$

再设总体的$2k$阶原点矩存在且函数G关于各变量均一阶连续可微，令

$$B = (b_{ij})_{k\times k} = (\mu_{i+j} - \mu_i\mu_j)_{k\times k},\ \ d_i = \frac{\partial G}{\partial \mu_i},\ i = 1, \cdots, k,$$

并且$\mathbf{d} = (d_1, \ldots, d_k)', \sigma^2 = \mathbf{d}'B\mathbf{d}$.

则在上述假设之下，我们有如下的结论.

定理 2.7.8 在上面条件下，当$n \to \infty$时，有

$$\sqrt{n}\left(\hat{g}(\mathbf{X}) - G(\mu_1, \ldots, \mu_k)\right) \xrightarrow{\mathcal{L}} N(0, \sigma^2). \tag{2.7.8}$$

三、极大似然估计的相合渐近正态性

为了简单明了，下面我们仅给出了总体参数为一维的情形.

定理 2.7.9 设参数空间Θ是开区间，总体PDF $f(x;\theta)$满足

(1) 在参数真值θ_0的邻域内，$\partial^i \ln f/\partial\theta^i$，$i = 1, 2, 3$存在;

(2) 在参数真值θ_0的邻域内,$|\partial^3 \ln f/\partial\theta^3| \le H(x)$，且$EH(X) < \infty$;

(3) 在参数真值θ_0处，

$$E_{\theta_0}\left[\frac{f'(X,\theta_0)}{f(X,\theta_0)}\right] = 0,\ E_{\theta_0}\left[\frac{f''(X,\theta_0)}{f(X,\theta_0)}\right] = 0,\ I(\theta_0) = E_{\theta_0}\left[\frac{f'(X,\theta_0)}{f(X,\theta_0)}\right]^2 > 0.$$

记$\hat{\theta}_n$为当$n \to \infty$时似然方程的相合解，则

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} N(0, I^{-1}(\theta_0)). \tag{2.7.9}$$

证明 由前面定理知，当$n \to \infty$时，似然方程以概率1有解，记其解为$\hat{\theta}_n$,并且$\hat{\theta}_n \xrightarrow{P} \theta_0$. 将$\partial l/\partial\theta$在$\theta_0$处Taylor展开，有

$$\frac{\partial l}{\partial \theta} = \frac{\partial l}{\partial \theta}|_{\theta_0} + \frac{\partial^2 l}{\partial \theta^2}|_{\theta_0}(\theta - \theta_0) + \frac{\partial^3 l}{\partial \theta^3}|_{\theta_1}\frac{(\theta - \theta_0)^2}{2},$$

其中θ_1介于θ与θ_0之间. 在上式中取θ的值为$\hat{\theta}_n$，由于$\hat{\theta}_n$是θ_0 的相合估计，故可知$\theta_1 \xrightarrow{P} \theta_0$，于是

$$0 = \frac{\partial l}{\partial \theta}|_{\hat{\theta}_n} = \frac{\partial l}{\partial \theta}|_{\theta_0} + (\hat{\theta}_n - \theta_0)\left[\frac{\partial^2 l}{\partial \theta^2}|_{\theta_0} + O_p(\hat{\theta}_n - \theta_0)\right],$$

上式的第三项是由条件(2)及上述的相合性得到的. 由此式可以得到

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = \frac{-\sqrt{n}\frac{1}{n}\frac{\partial l}{\partial \theta}|_{\theta_0}}{\frac{1}{n}\left[\frac{\partial^2 l}{\partial \theta^2}|_{\theta_0} + O_p(\hat{\theta}_n - \theta_0)\right]}. \tag{$*$}$$

我们注意到

$$\frac{\partial l}{\partial \theta}|_{\theta_0} = \sum_{i=1}^n \frac{f'(x_i,\theta_0)}{f(x_i,\theta_0)},\ \ \frac{\partial^2 l}{\partial \theta^2}|_{\theta_0} = \sum_{i=1}^n\left[\frac{f''(x_i,\theta_0)}{f(x_i,\theta_0)} - \left(\frac{f'(x_i,\theta_0)}{f(x_i,\theta_0)}\right)^2\right].$$

因为$E_{\theta_0}\left[\frac{f'(X,\theta_0)}{f(X,\theta_0)}\right] = 0, Var_{\theta_0}\left[\frac{f'(X,\theta_0)}{f(X,\theta_0)}\right] = I(\theta_0)$,所以，由中心极限定理有

$$\frac{1}{\sqrt{n}}\frac{\partial l}{\partial \theta}|_{\theta_0} \xrightarrow{\mathcal{L}} N(0, I(\theta_0)). \tag{$*1$}$$

由条件(3)及大数定律知

$$\frac{1}{n}\frac{\partial^2 l}{\partial\theta^2}|_{\theta_0} \xrightarrow{P} -I(\theta_0). \tag{$*2$}$$

结合(*),(*1),(*2)和Slutskey定理，有(2.7.9). □

注 2.7.1 我们注意到，对于不满足上述定理条件的分布族，其MLE的极限分布也有不是正态分布的，如均匀分布$U(0,\theta)$，由定义可知其MLE 是$X_{(n)}$,而由次序统计量的分布不难知道，它的极限分布不是正态的，而是一种极值分布.

注 2.7.2 The meaning of notation $O_p(1)$ and $o_p(1)$

- A sequence of random variables $\{X_n\}$, with respective distribution functions $\{F_n\}$, is said to be bounded in probability if for every $\varepsilon > 0$, there exist M_ε and N_ε such that
$$F_n(M_\varepsilon) - F_n(-M_\varepsilon) > 1 - \varepsilon, \quad \text{for all } n > N_\varepsilon.$$
Then the notation $X_n = O_p(1)$ will be used.

- The notation $U_n = O_p(V_n)$ denotes that the sequence $\{U_n/V_n\}$ is $O_p(1)$.

- The notation $U_n = o_p(V_n)$ denotes that $U_n/V_n \xrightarrow{p} 0$

从(2.7.9)式可以看出，在某些条件下，MLE的极限方差达到了C-R下界，故它是渐近有效的，所以有人也称MLE为有效似然估计.

虽然MLE具有很多的优良性质，但我们仍要注意它存在如下几点不足：

- MLE可能不是“最好”的，如不是UMVUE或均方误差最小的；

  对于来自$U(0,p)$的IID样本，知$\hat{p} = X_{(n)}$是p的惟一的MLE. 但由于$P\{X_{(n)} < p\} = 1$,故它总是低估p. 事实上，由充分完备统计量的性质知，p的UMVUE是$\frac{(n+1)X_{med}}{n}$,$\frac{(n+2)X_{med}}{n+1}$是在形如cX_{med} 的估计中均方误差最小的一个.

- MLE可能不惟一(前面已经讲过)

- MLE依赖于总体的分布函数，如不知样本的分布，则无法求得其感兴趣参数的MLE. 于是，现在研究很热的一个方向即是Empirical Likelihood (Owen, A. B.于1988年提出的.)

§2.8 Bayes估计

有关Bayes估计的详细内容请参见Berger (1985). 本节只是简单地对Bayes估计作一介绍，使大家了解除了经典的估计方法之外，还有Bayes估计.

我们在概率论中学过如下的Bayes公式：

$$P(H_k|A) = \frac{P(H_k)P(A|H_k)}{\sum_i P(H_i)P(A|H_i)}, \ k = 1, 2, \cdots,$$

其中$\cup H_i = \Omega$是必然事件的一个划分. 当时我们称$P(H_k)$为先验(prior)，而$P(H_k|A)$称为后验(posterior). 这个公式是在英国学者Thomas Bayes (1702–1761)于1763年发表的. 从形式上看，它只是条件概率的简单应用，但它却包含了归纳推理的思想，后来的学者把它发展成一种关于统计推断的系统理论与方法，统称为Bayes统计，后来发展成了Bayes学派. 当然，Bayes学派在统计界引起了热烈的争论. 有关Bayes方面的内容，请有兴趣的参见Berger (1985).

Bayes统计的基本观点可以用下面三个假设来归纳：

假设I 设随机变量X有一个PDF $f(x,\theta)$, 其中θ是一个参数，由于对不同的θ值，$f(x,\theta)$对应着不同的PDF，所以，从Bayes观点看，它是在给定θ 值后的一个条件PDF. 因此把$f(x,\theta)$记为$f(x|\theta)$更恰当些，而$f(x|\theta)$中关于θ的信息即为总体信息.

假设II 当给定θ后，从总体$f(x|\theta)$中随机地抽取一组样本$X_1,\ldots,X_n$，该样本中含有大量的关于θ的信息，这就是样本信息. 于是，样本分布$f(\mathbf{x}|\theta)=\prod_{i=1}^n f(x_i|\theta)$即综合了总体与样本信息.

假设III 我们对参数θ已经积累了一些资料，经过分析整理和加工后，可以获得一些关于θ的有用信息，这种信息称为θ的先验信息, 由于θ不是永远固定在一个值上，故从Bayes观点看，未知参数θ也是一个随机变量. 关于θ的分布可以从先验信息中归纳出来，故其分布称为先验分布, 用$\pi(\theta)$表示其PDF.

这样，综合前面几点，我们有样本$X_1,\ldots,X_n$及θ的联合PDF为

$$f(\mathbf{x},\theta)=f(\mathbf{x}|\theta)\pi(\theta). \tag{2.8.1}$$

而我们的目的在于利用上面的联合分布，即三方信息来估计θ. 当有了样本后，我们可以得到θ的条件分布如下：

$$\pi(\theta|\mathbf{x})=\frac{f(\mathbf{x},\theta)}{f(\mathbf{x})}=\frac{f(\mathbf{x}|\theta)\pi(\theta)}{\int f(\mathbf{x}|\theta)\pi(\theta)\mathrm{d}\theta}, \tag{2.8.2}$$

我们称之为θ的后验PDF, 而$f(\mathbf{x})=\int f(\mathbf{x}|\theta)\pi(\theta)\mathrm{d}\theta$是$X_1,\ldots,X_n$的边际分布或样本的无条件分布.

从上式可以看出，参数θ的后验分布归纳了θ的先验信息、总体中关于θ的信息和样本提供的信息，于是，θ的估计就应从其后验分布入手. 常用的估计方法为

(1) 取后验分布$\pi(\theta|\mathbf{x})$的最大值，即众数作为θ的估计，称之为众数型的Bayes估计.

(2) 取后验分布$\pi(\theta|\mathbf{x})$的中位数作为θ的估计，称之为中位数型的Bayes估计.

(3) 取后验分布$\pi(\theta|\mathbf{x})$的期望作为θ的估计，称之为期望型的Bayes估计.

在上述三种常用的估计中，第三种，即期望型的Bayes估计最常用. 以后我们将采用后验期望作为θ的Bayes估计：

$$\hat{\theta}_B=E(\theta|\mathbf{X})=\int\theta\pi(\theta|\mathbf{x})\mathrm{d}\theta. \tag{2.8.3}$$

例 2.8.1 设某产品的次品率为θ，请估计θ.

解 为估计θ，现随机地抽取n件产品进行检验，发现有X件次品，则次品X的条件分布为

$$P\{X=x|\theta\}=\begin{pmatrix} n \\ x \end{pmatrix}\theta^x(1-\theta)^{n-x},\ x=0,1,\cdots,n.$$

如从经典统计来看，次品数X是θ的充分统计量，且其MLE为

$$\hat{\theta}_{ML}=\frac{X}{n}=\frac{\text{样本中次品数}}{\text{样本容量}}.$$

前面我们已经讲过，MLE具有许多优良性质，也是估计中经常利用的. 对于本例，假设我们有如表中的数据，则得到的MLE也在表中给中.

表2.3 MLE与Bayes估计的比较

No.	n	X	$\hat{\theta}_{ML}$	$\hat{\theta}_B$
1	5	5	1	0.875
2	20	20	1	0.955
3	5	0	0	0.143
4	20	0	0	0.045

从表2.3可以看出，针对不同样本容量，其MLE可能是相同的，如抽取的20个样本都是次品，其次品率的MLE为1；而当抽取的5个样本也都是次品时，其次品率的MLE仍为1. 显然，这两种情况所包含的信息是不同的，但MLE所给出的估计值却相同. 对于第3、4种情况，其MLE也有类似的特点. 但是，本表最后一列所给出的Bayes估计则与我们的直观想象相吻合. 为了得到次品率θ的Bayes估计，我们就有必要选择其先验分布，之后再利用前面的方法求取其估计值. 可以说，在应用Bayes方法时，确定未知参数的先验分布是关键. 一般来说，先验分布的确定可以如下进行：

第一步 选一个适应面较广的分布族作先验分布，以便在数据处理上方便一些;

第二步 根据先验信息在第一步选取的分布族中选取一个分布作为先验分布，使它与先验信息符合较好;

第三步 如很难从前面两步选取先验分布，则可用无信息先验，即$\pi(\theta)=1$，或用均匀分布作为先验分布.

在本例中，我们选取β分布族

$$\pi(\theta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1},\ 0\le\theta\le1,\ \alpha>0,\beta>0$$

作为θ的先验分布族. 我们之所以选取上述β分布作为次品率θ的先验分布，其理由为

- 因为$\theta\in[0,1]$，故有必要选用一个在区间[0,1]上取值的分布族.
- β分布中有两个未知参数，故它的适应面较大.
- 当样本的条件分布为$B(n,\theta)$时，可以证明，如取参数θ的先验分布为β分布时，则其后验分布仍为β分布. (我们称这样的先验为共轭先验)

- 使用β分布作为先验已有许多成功的例子.

当取定β分布族作为先验后，就要选取合适的α,β以取定先验分布. 对于先验分布中参数的估计方法常如下进行：

- 假设能从先验信息中较为准确地算出p的先验平均$\bar{\theta}$和先验方差S_θ^2，则我们利用矩法估计α,β，即令

$$\begin{cases} \frac{\alpha}{\alpha+\beta}=\bar{\theta}, \\ \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}=S_\theta^2, \end{cases}$$

由此可求得

$$\begin{cases} \alpha=\frac{(1-\bar{\theta})\bar{\theta}^2}{S_\theta^2}-\bar{\theta}, \\ \beta=\frac{\alpha(1-\bar{\theta})}{\bar{\theta}}. \end{cases}$$

- 假设能从先验信息中找到θ的两个分位数，如10%与50%的分位数$\theta_{0.1},\theta_{0.5}$，则可以通过求解方程

$$\int_0^{\theta_{0.1}}\pi(\theta)\mathrm{d}\theta=0.1,\quad \int_0^{\theta_{0.5}}\pi(\theta)\mathrm{d}\theta=0.5$$

求得α,β的值.

- 如果先验信息很丰富的话，可以用这些先验信息去拟合它的一个分布出来，之后用这个分布作为它的先验分布.

对于本例，我们有

$$\begin{aligned} f(x,\theta) &= P\{X=x|\theta\}\pi(\theta) \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\begin{pmatrix} n \\ x \end{pmatrix}\theta^{\alpha+x-1}(1-\theta)^{\beta+n-x-1},\ x=0,1,\cdots,n, 0<\theta<1, \end{aligned}$$

于是样本边际分布为

$$f(x)=\int_0^1 f(x,\theta)\mathrm{d}\theta=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\alpha+x)\Gamma(\beta+n-x)}{\Gamma(\alpha+\beta+n)}\begin{pmatrix} n \\ x \end{pmatrix},\ x=0,1,\cdots,n.$$

后验分布为

$$\pi(\theta|x)=\frac{f(x,\theta)}{f(x)}=\frac{\Gamma(\alpha+\beta+n)}{\Gamma(\alpha+x)\Gamma(\beta+n-x)}\theta^{\alpha+x-1}(1-\theta)^{\beta+n-x-1},\ 0<x<1,$$

它仍然是β分布$\beta(\alpha+x,\beta+n-x)$，则次品率的期望型Bayes估计为

$$\hat{\theta}_B=\int_0^1\theta\pi(\theta|x)\mathrm{d}\theta=\frac{\alpha+x}{\alpha+\beta+n},$$

它明显地不同于前面讲的MLE $\hat{\theta}_{ML}$. 事实上，

$$\hat{\theta}_B=\frac{\alpha+\beta}{\alpha+\beta+n}\frac{\alpha}{\alpha+\beta}+\frac{n}{\alpha+\beta+n}\frac{x}{n}=\lambda\bar{\theta}+(1-\lambda)\hat{\theta}_{ML},$$

其中$\lambda=\frac{\alpha+\beta}{\alpha+\beta+n}$, $\bar{\theta}=\frac{\alpha}{\alpha+\beta}$是先验期望. 这表明Bayes估计$\hat{\theta}_B$是先验期望与MLE的加权平均，而先验分布中的参数与n则体现了权函数的大小.

对于本例，由于我们对过去的信息一无所知，故可采用无信息先验，即采用(0,1)上的均匀分布，即$\alpha=\beta=1$的β分布. 这样，得到的θ的Bayes估计为

$$\hat{\theta}_B=\frac{X+1}{n+2},$$

这即是我们在前面表中列出的估计值. 比较一下此表中的Bayes估计与MLE，可以看出Bayes估计的合理性. □

§2.9 最小二乘估计

虽然MLE是一个很好的点估计，但由MLE的定义不难看出，它过份地依赖于总体分布. 然而在许多实际问题中，我们可能并不知道总体分布如何，而只是有一组数据，且知道协变量与响应变量之间存在着线性关系. 此时，我们可以对此组数据进行最小二乘分析. 在本节中，我们将介绍最小二乘估计(Least Square Estimation, 简记为LSE)、最优线性无偏估计(Best Linear Unbiased Estimation, 简记为BLUE)和加权最小二乘估计(Weighted LSE.)

§2.9.1 最小二乘估计

最小二乘估计最早出现于十七世纪，是统计中出现最早的方法之一. 它主要是针对如下的Gauss-Markov模型：

$$\begin{cases}\mathbf{y}=X\beta+\varepsilon,\\ E\varepsilon=\mathbf{0},\ \mathrm{Var}\varepsilon=\sigma^2\mathbf{I}_n.\end{cases}$$

其中$\mathbf{y}$为n维观测向量，X为$n\times p$阶设计矩阵，β为p维未知参数，σ^2未知.

定义 2.9.1 (LSE) 对于上面的Gauss-Markov模型，如果

$$(\mathbf{y}-X\hat{\beta})'(\mathbf{y}-X\hat{\beta})=\min_{\beta}(\mathbf{y}-X\beta)'(\mathbf{y}-X\beta), \tag{2.9.1}$$

则称$\hat{\beta}$为β的LSE.

如记$Q(\beta)=(\mathbf{y}-X\beta)'(\mathbf{y}-X\beta)$,则知，求LSE等价于求$Q(\beta)$的最小值. 利用求最值的方法，我们知道求取LSE，即相当于求解$\partial Q(\beta)/\partial\beta=0$的驻点，即解下面的方程：

$$\frac{\partial Q(\beta)}{\partial\beta}=-2X'\mathbf{y}+2X'X\beta=0,$$

此方程称为正规方程(normal equation)，也即

$$X'X\beta=X'\mathbf{y}. \tag{2.9.2}$$

如果设计矩阵X是列满秩的，则上述的正规方程的解惟一，且为

$$\hat{\beta}_{LS}=(X'X)^{-1}X'\mathbf{y}, \tag{2.9.3}$$

这即是β的LSE.

注 2.9.1 当X不是列满秩时，上式中的逆不存在，但可以用广义逆代替.

从(2.9.3)可以看出，LSE是观测向量$\mathbf{y}$的线性函数，这一点在以后的证明过程中非常有用.

定理 2.9.1 对于前面的Gauss-Markov模型，如果设计矩阵X 是列满秩的，且以$\hat{\beta}$ 记由(2.9.3)给出的β的LSE，则

(1) $\hat{\beta}$是β的UE; (如X不是列满秩的，此结论仍然正确)

(2) $\mathrm{Var}\hat{\beta} = \sigma^2(X'X)^{-1}$.

证明 此定理的证明非常容易，留作课后习题，请大家补上.

下面看一个最简单的线性模型的最小二乘估计.

例 2.9.1 考虑线性模型

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,\ i = 1, 2\cdots, n,$$

的LSE，其中$\varepsilon_i \overset{iid}{\sim} (0, \sigma^2)$.

解 对于此问题，根据最小二乘的思想，我们要求$\hat{\beta}_0, \hat{\beta}_1$满足

$$(\hat{\beta}_0, \hat{\beta}_1) = \min_{\beta_0, \beta_1} \arg \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2.$$

利用求最值的方法，容易求得其LSE为

$$\begin{cases} \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}, \\ \hat{\beta}_1 = S_{xy}/S_x, \end{cases} \tag{2.9.4}$$

其中$\bar{x} = \sum_{i=1}^{n} x_i/n, \bar{y} = \sum_{i=1}^{n} y_i/n,\ S_x = \sum_{i=1}^{n}(x_i - \bar{x})^2, S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$. 并且，$\mathrm{Var}\hat{\beta}_0 = \sigma^2 \sum x_i^2/S_x,\ \mathrm{Var}\hat{\beta}_1 = \sigma^2/S_{xy},\ \mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_1) = -\bar{x}/S_x$.

我们注意到，如果$\bar{x} = 0$，则$\mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_1) = 0$, 这就是说，如果把设计点变换后，其均值为0，则得到的截距与斜率的LSE是不相关的. □

§2.9.2 最优线性无偏估计

由(2.9.3)得到的LSE，不仅是无偏的，而且还有一些优良的统计性质，如BLUE等，这即是本节的结论.

定义 2.9.2 (BLUE) 对于前面的Gausss-Markov模型，如果β的一个线性UE $\tilde{\beta} = C\mathbf{y}$ 满足

$$\mathrm{Var}\tilde{\beta} \le \mathrm{Var}(A\mathbf{y}),$$

则称$\tilde{\beta}$是β的BLUE，其中C, A是$p \times n$阶矩阵，且满足$EC\mathbf{y} = EA\mathbf{y} = \beta$.

定理 2.9.2 对于前面的Gauss-Markov模型，如果X是列满秩的，则(2.9.3)式的LSE是β的BLUE.

证明 $\hat{\beta}$的线性无偏性是显然的，下证其方差在线性无偏估计中是最小的. 事实上，设$T = A\mathbf{y}$是β的一个线性UE, 即$EA\mathbf{y} = AX\beta = \beta,\ \forall \beta$, 于是$AX = \mathbf{I}_p$. 另外，其方差为

$$\begin{aligned} \mathrm{Var}A\mathbf{y} &= E[(A\mathbf{y} - \hat{\beta} + \hat{\beta} - \beta)(A\mathbf{y} - \hat{\beta} + \hat{\beta} - \beta)'] \\ &= \mathrm{Var}(A\mathbf{y} - \hat{\beta}) + \mathrm{Var}\hat{\beta} - 2\mathrm{Cov}(A\mathbf{y} - \hat{\beta}, \hat{\beta} - \beta). \end{aligned}$$

注意到$E(A\mathbf{y}-\hat\beta)=0$,故

$$\begin{aligned}\mathrm{Cov}(A\mathbf{y}-\hat\beta,\hat\beta-\beta) &= \mathrm{Cov}[(A-(X'X)^{-1}X')\mathbf{y},(X'X)^{-1}X'\mathbf{y}]\\ &= (A-(X'X)^{-1}X')\mathrm{Var}\mathbf{y}((X'X)^{-1}X')'\\ &= \sigma^2(AX-(X'X)^{-1}X'X)(X'X)^{-1}=0,\end{aligned}$$

于是有$\mathrm{Var}A\mathbf{y}\geq\mathrm{Var}\hat\beta$, 并且从证明过程可以看出，其等号成立的充要条件是$\mathrm{Var}(A\mathbf{y}-\hat\beta)=0$, 即$A\mathbf{y}=\hat\beta, a.s.$ □

由上述定理的证明可以看出，β的LSE是惟一的BLUE. 如果我们感兴趣的是β的线性组合$\mathbf{a}'\beta$，则有下面的定理

定理 2.9.3 对于前面的Gauss-Markov模型，如果X是列满秩的，则$\forall\ \mathbf{a}\in\mathbf{R}^p$, $\mathbf{a}'\hat\beta$是$\mathbf{a}'\beta$ 的惟一的BLUE.

证明 其证明与定理2.9.2的证明类似，留作习题.

注 2.9.2 由定理2.9.3知，$\hat\beta$的任一分量均是其对应参数的BLUE.

注 2.9.3 我们注意到在Gauss-Markov模型中，还有一个参数σ^2仍需估计. 实际上，如记

$$SSE=(\mathbf{y}-X\hat\beta)'(\mathbf{y}-X\hat\beta),$$

它被称为残差平方和，并且$SSE=Q(\hat\beta)$. 如果X是列满秩的，则可以证明

$$\hat\sigma^2=SSE/(n-p), \tag{2.9.5}$$

是σ^2的UE. 另外，由于它是基于β的LSE得到的，故称之为σ^2的LSE.

在上述的讨论中，我们始终没有假设数据的分布形式. 如果已知它是正态的，则可以证明上述LSE是UMVUE.

定理 2.9.4 对于前面的Gauss-Markov模型，如果X是列满秩的，且$\varepsilon\sim N(0,\sigma^2\mathbf{I}_p)$, 则(2.9.3)式中的$\hat\beta$ 及(2.9.5)式中的$\hat\sigma^2$分别是β和σ^2的UMVUE. 另外，$\forall\ \mathbf{a}'\in\mathbf{R}^p$, $\mathbf{a}'\hat\beta$也是$\mathbf{a}'\beta$的UMVUE.

证明 对于此定理的证明，我们概述如下：(1) 先利用指数族分布的性质说明$(\mathbf{y}'\mathbf{y},X'\mathbf{y})$是充分完备统计量(2)利用前面的结论知，(2.9.3)式中的$\hat\beta$及(2.9.5)式中的$\hat\sigma^2$均是上述充分完备统计量的函数且是无偏的，故知它们分别是β和σ^2的UMVUE. □

§2.9.3 加权最小二乘估计

在上一小节的讨论中，我们始终假设各观测间的方差是相等的，但在许多实际问题，每次观测误差的精度可能是不一样的，于是本节将讨论这一问题的LSE.

此时，我们有如下的广义Gauss-Markov模型：

$$\begin{cases}\mathbf{y}=X\beta+\varepsilon,\\ E\varepsilon=\mathbf{0},\ \mathrm{Var}\varepsilon=\sigma^2\mathbf{\Sigma},\end{cases}$$

其中$\mathbf{\Sigma}>\mathbf{0}$为已知正定阵.

对于此模型，我们仍然可以利用(2.9.3)来作为β的估计，并且也是无偏的，但它不是BLUE，这是由于在证明BLUE 过程中要用到观测的方差.

对于此广义的Gauss-Markov模型，由于$\mathbf{\Sigma} > \mathbf{0}$, 故存在$n$阶对称阵$B$，使$\mathbf{\Sigma} = B^2$. 作变换$\mathbf{z} = B^{-1}\mathbf{y}$, $\tilde{X} = B^{-1}X$, 则有如下的Gauss-Markov模型：

$$\begin{cases} \mathbf{z} = \tilde{X}\beta + \varepsilon, \\ E\varepsilon = \mathbf{0},\ \mathrm{Var}\varepsilon = \sigma^2\mathbf{I}_n. \end{cases}$$

这样，由此模型得到β的LSE如下：

$$\tilde{\beta} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'z = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\mathbf{y}, \tag{2.9.6}$$

我们称之为β的加权的或广义的最小二乘估计，并且由定理2.9.3知，它也是BLUE.

在有的书上，(9.9.3)式的LSE也称为普通的LSE(Ordinary LSE,简记为OLSE),以求与GLSE的区别. 另外，最小二乘方法在线性模型上有着广泛的应用，并得到了充分的研究，取得了丰硕的成果,有兴趣的可参见王松桂(1987). 线性模型的理论及其应用，安徽教育出版社.

§2.9.4 线性模型的诊断

虽然最小二乘估计是线性回归中最常用的一种估计方法，并且在前面也讲述了它的某些优良性质，但是在某些情况下它却与实际有着很大的区别. 本节将通过一个例子说明数据对LSE的影响是非常大的.

例 2.9.2 (Anscombe数据) 表2.4中的数据是Anscombe于1973年给出的人为数据(请见*Ann. Statist.* 27, 17-21.)

表2.4 Anscombe数据

No.	$x^{(1)}$	$y^{(1)}$	$y^{(2)}$	$y^{(3)}$	$x^{(2)}$	$y^{(4)}$
1	10.00	8.04	9.14	7.46	8.00	6.58
2	8.00	6.95	8.14	6.77	8.00	5.76
3	13.00	7.58	8.74	12.70	8.00	7.71
4	9.00	8.81	8.77	7.11	8.00	7.71
5	11.00	8.33	9.26	7.81	8.00	8.47
6	14.00	9.96	8.10	8.84	8.00	7.04
7	6.00	7.24	6.13	6.08	8.00	5.25
8	4.00	4.26	3.10	5.39	19.00	12.05
9	12.00	10.08	9.13	8.15	8.00	5.56
10	7.00	4.82	7.26	6.42	8.00	7.91
11	5.00	5.68	4.74	5.73	8.00	6.89

请利用表2.4中的数据分别建立$x^{(1)}$关于$y^{(1)}, y^{(2)}, y^{(3)}$及$x^{(2)}$关于$y^{(4)}$间的线性模型.

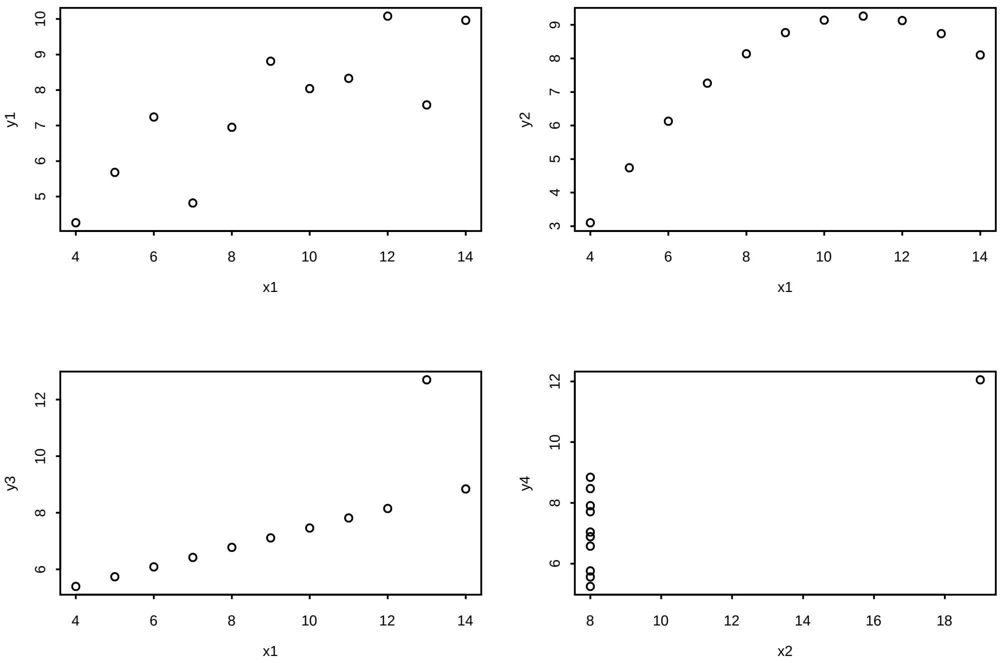

图2.1 Anscombe数据的散点图

解 分别对上述四组变量做如下的一元线性回归：

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i,\ i = 1, 2, \cdots, 11.$$

经过简单计算后知，上述四个模型的参数的LSE估计均为

$$\hat{\beta}_0 = 3.0,\ \hat{\beta}_1 = 0.5,\ \hat{\sigma}^2 = 1.531,\ R^2 = 0.667,$$

其中R^2是复相关系数，是用来衡量模型拟合好坏的一个指标，其具体定义请见有关回归分析的参考书. 从这些指标来看，我们用回归直线$y = 3 + 0.5x$来拟合这四组数据的效果是完全相同的. 但是，从这四组数据的散点图–图2.1可以看出，这四组数据的情况相差非常大. □

例2.9.2说明，在我们对数据进行经典回归分析时，一定要事先对数据做必要的回归诊断(regression diagnostic)，否则会导致错误的结论. 有关回归诊断的内容，请参见韦博成、鲁国斌、史建清(1991).

§2.9.5 一个实例

线性回归主要是处理变量之间线性关系的重要手段，是大家经常使用的一种统计方法，也取得了很好的效益. 本节引述何书元(2006)中的一个实例，以餐读者.

二战期间，德国对法国发动攻势后，英国首相丘吉尔应法国的请求，动用了十几个防空中队对德作战. 由于防空中队的飞机需要在欧洲大陆的机场进行维护，使得空战中英国飞机

损失严重. 这时，法国总理请求英国继续增派十几个中队的飞机，丘吉尔决定同意这一要求. 英国内阁知道此事后，请来统计学家利用线性回归模型对出动飞机与战损飞机的数据进行了统计分析，发现如果飞机的补充率和损失率不变，飞机数量的下降是非常快的："以现在的损失率损失两周，英国在法国的飓风式战斗机就一架也不存在了". 内阁希望丘吉尔收回他的决定. 最后，丘吉尔同意了内阁的要求，并在几天内撤回了在法国的飓风式战机，只留下了三个中队，为以后英国本土的保卫战保留了实力. (读者, 2006.4)

习题二

1. 设总体X的CDF为

$$P\{X=x\}=p(1-p)^{x-1},\ x=1,2,\cdots,$$

$X_1,\ldots,X_n$为来自此总体的iid样本，试求p的矩估计和MLE.

2. 设$X_1,\ldots,X_n$为来自总体分布

$$f(x,\alpha)=\begin{cases}(\alpha+1)x^{\alpha}, & 0<x<1,\\ 0, & \text{其它},\end{cases}$$

的iid样本，试求α的矩估计和MLE.

3. 设$X_1,\ldots,X_n$为来自总体X的iid样本，n个正常数α_i满足$\sum_{i=1}^n\alpha_i=1$. 试证在EX的所有形如$\sum_{i=1}^n\alpha_iX_i$的UE中，以$\bar{X}$为最优. (最优的标准为方差最小)

4. 设$X_1,\ldots,X_n$为来自双指数分布

$$f(x;\theta,\lambda)=\begin{cases}\frac{1}{\lambda}e^{-(x-\theta)/\lambda}, & x\geq\theta,\lambda>0,\\ 0, & \text{其它},\end{cases}$$

的iid样本. 试求θ,λ的MLE.

5. 设$X_1,\ldots,X_n$为来自总体$N(\mu,\sigma^2)$的iid样本，求满足

$$P\{X>a\}=\int_a^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x=0.05$$

的点a的MLE.

6. 设$X_1,\ldots,X_n$为来自均匀分布$U(a,b)$的iid样本，试求a^2的MLE.

7. 设某种产品的寿命X服从指数分布$E(\lambda)$, t_0为给定常数，$X_1,\ldots,X_n$是独立观测到的n件产品的寿命. 求产品在t_0前失效的概率$P\{X\leq t_0\}=1-e^{-\lambda t_0}=g(\lambda)$的UMVUE.

8. 设$X_1,\ldots,X_n$为来自均匀分布$U(0,\theta)$的iid样本，其中$\theta>0$为求知参数.

(1) 试证明$\hat{\theta}_1=\frac{n+1}{n}X_{(n)}$, $\hat{\theta}_2=(n+1)X_{(1)}$均是$\theta$的UE.

(2) 上述两个估计中哪个方差最小？

(3) 试证明$\hat{\theta}_1$是θ的UMVUE.

9. 设$X_1,\ldots,X_n$为来自$N(0,\sigma^2)$的iid样本，试求σ^2的充分完备统计量，σ和$3\sigma^4$的UMVUE，并证明$\frac{1}{n}\sum_{i=1}^n X_i^2$是$\sigma^2$的有效估计和相合估计.

10. 设$a_1,\ldots,a_n$是n个实数，定义函数$h(a)=\sum_{i=1}^n|a_i-a|$. 证明：

(1) 当a为$a_1,\ldots,a_n$的样本中位数时，$h(a)$达到最小值；

(2) 设$X_1,\ldots,X_n$为来自具有PDF为$e^{-|x-\theta|}/2$的总体(此分布称为Laplace分布)的iid样本，求参数θ的矩估计与MLE.

11. (1) 证明

$$f(x,a,\sigma)=(\sqrt{2\pi}\sigma^3)^{-1}(x-a)^2\exp\{-(x-a)^2/2\sigma^2\},\ x\in R$$

作为x的函数是一PDF，其中$a\in R,\sigma>0$为参数.

(2) 设$X_1,\ldots,X_n$为来自此上述总体的iid样本，求a,σ^2的矩估计.

(3) 给出a,σ^2的MLE所满足的方程，并给出一种迭代求解方法.

12. 设X为来自Poisson分布$P(\lambda)$的一个样本，参数λ的先验密度为$\pi(\lambda)=e^{-\lambda}$，$\lambda>0$(当$\lambda\le 0$时，$\pi(\lambda)=0$). 试求$\lambda$的Bayes估计.

13. 设$X_1,\ldots,X_n$为来自均匀分布$U(0,\theta)$的iid样本. 证明

(1) $\hat\theta_1=X_{(n)}+X_{(1)}$是$\theta$的一个UE；

(2) 对适当选择的常数c_n，$\hat\theta_2=c_nX_{(1)}$是θ的UE，但这个估计的方差比另外两个UE$\hat\theta_3=\bar X$和$\hat\theta_4=\frac{n+1}{n}X_{(n)}$都大(除非$n=1$).

14. 设$X_1,\ldots,X_n$为来自PDF为

$$f(x,\theta)=\begin{cases}2\sqrt{\theta/\pi}\exp\{-\theta x^2\}, & x>0\\ 0, & x\le 0,\end{cases}$$

的总体的iid样本，证明：对适当选取的常数c，$\hat\theta=c\sum_{i=1}^n X_i^2/n$是$1/\theta$的UMUVE.

15. (1) 若$\hat\theta_1,\hat\theta_2$是$\theta$的UMVUE，则$(\hat\theta_1+\hat\theta_2)/2$也是.

(2) 如果$\hat\theta$是θ的UMVUE，而$a\ne 0$和b都是已知常数，则$a\hat\theta+b$是$a\theta+b$的UMVUE.

16. 设$X_1,\ldots,X_n$为来自PDF为

$$f(x,\theta)=\begin{cases}e^{-(x-\theta)}, & x>\theta,\\ 0, & \text{其它},\end{cases}$$

的总体的iid样本.

(1) 求θ的矩估计$\hat\theta_M$；

(2) 求θ的MLE $\hat\theta_{ML}$；

(3) $\hat\theta_M$是否是θ的相合估计？为什么？

17. 设$X_1,\ldots,X_n$为来自正态总体$N(\mu,\sigma^2)$ 的iid样本，求常数c使$c\sum_{i=1}^n|X_i-\bar X|$是$\sigma$的UE.

18. 设X_1, X_2为来自PDF为

$$f(x,\theta)=\begin{cases} 3x^2/\theta^3, & 0<x<\theta, \\ 0, & \text{其它}, \end{cases}$$

的总体的iid样本，其中$\theta>0$为参数.

(1) 证明$T_1=\frac{2}{3}(X_1+X_2)$和$T_2=\frac{7}{6}\max\{X_1,X_2\}$是$\theta$的UE；

(2) 计算T_1, T_2的方差，并指出何者更有效;

(3) 证明在均方误差意义下，在形为$T_c=c\max\{X_1,X_2\}$的估计中，$T_{\frac{8}{7}}$最有效.

19. 设$X_1,\ldots,X_n$为来自PDF为

$$f(x,\theta)=\begin{cases} \theta^2 x e^{-\theta x}, & x\geq 0, \\ 0, & x<0, \end{cases}$$

的总体的iid样本.

(1) 求θ的MLE $\hat{\theta}_{ML}$;

(2) 求θ的UMVUE;

(3) $\hat{\theta}_{ML}$是θ的相合估计吗？为什么？

20. 设$X_1,\ldots,X_n$为来自均匀分布$U(\theta-0.5,\theta+0.5)$的iid样本$(n\geq 2)$，$-\infty<\theta<\infty$为未知参数.

(1) 求θ的矩估计$\hat{\theta}_M$;

(2) $\hat{\theta}_M$和$\hat{\theta}_1=(X_{(1)}+X_{(n)})/2$是$\theta$的UE吗？为什么？

(3) $\hat{\theta}_M$和$\hat{\theta}_1$哪个方差较小？为什么？

21. 设$X_1,\ldots,X_n$为来自均匀分布$U(\alpha-\theta,\alpha+\theta)$的iid样本，其中$\alpha\in R,\theta>0$为未知参数，请求$\alpha,\theta$的矩估计.

22. 设$X_1,\ldots,X_n$为来自总体

$$X\sim\begin{pmatrix} -1 & 0 & 2 \\ 2\theta & \theta & 1-3\theta \end{pmatrix}$$

的iid样本，其中$0<\theta<1/3$. 试求参数θ的MLE.

23. 验证(1.4.1)式给出的统计量是σ的UE，并求其效率.

24. 设随机变量X以均等机会按$N(0,1)$分布和按$N(\mu,\sigma^2)$分布取值，其中$\mu\in R,\sigma^2>0$均未知. 此时X的PDF为上述两个正态的平均：

$$f(x,\mu,\sigma^2)=\frac{1}{2}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}+\frac{1}{2}\frac{1}{\sqrt{2\pi}\sigma}e^{-(x-\mu)^2/2\sigma^2}.$$

设$X_1,\ldots,X_n$为来自些混合分布总体的iid样本，试证明μ,σ^2不存在MLE.

25. 在买面包作早点的男、女消费者中，男性购买者的比例p未知，但知道$\frac{1}{2} \le p \le \frac{1}{3}$. 设在70个购买者中发现12个是男性，58个是女性，试求p的MLE. 如果对p没有限制，试求p的MLE.

26. 设$X_1, \dots, X_n$为来自Pareto分布

$$f(x,\theta) = \begin{cases} \frac{\theta}{(1+x)^{\theta+1}}, & 0 < x < \infty, \\ 0, & x \le 0, \end{cases}$$

的iid样本，其中$\theta > 1$为未知参数，试求$1/\theta$的无偏估计.

27. 设$X_1, \dots, X_n$为来自正态总体$N(\mu, \sigma^2)$的iid样本，试求常数d_n，使得$d_n R_n$为σ的无偏估计，其中$R_n = X_{(n)} - X_{(1)}$为极差.

28. 设$X_1, \dots, X_n$为来自对数正态总体

$$f(x,\mu,\sigma^2) = (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2}(\ln x - \mu)^2\right\},\ x > 0$$

的iid样本，其中$\mu \in R, \sigma > 0$为未知参数，试求μ, σ^2的矩估计.

29. 设$\beta = a\theta + b$，$\hat{\theta}$是θ的有效估计，试证明$\hat{\beta} = a\hat{\theta} + b$是$\beta$的有效估计.

30. 设$X_1, \dots, X_n$为来自均匀分布$U(-\theta/2, \theta/2)$的iid样本，其中$\theta > 0$为参数. 试证明$(X_{(1)}, X_{(n)})$是θ 的充分统计量. 它是完备的吗？为什么？

31. 设$X_1, \dots, X_n$为来自两点分布$b(1,p)$的iid样本，其中$p \in (0,1)$为参数. 试求

(1) p^k的UMVUE.

(2) $p^k + (1-p)^{n-k}$的UMVUE，其中k为整数且$0 < k < n$.

32. 设$X_1, \dots, X_n$为来自Poisson分布$P(\lambda)$的iid样本，试求

(1) λ^k的UMVUE.

(2) $P\{X_1 = k\}$的UMVUE，其中k为大于0的整数.

33. 对于线性模型

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i,\ i = 1, \cdots, n$$

其中$\epsilon_1, \dots, \epsilon_n$独立同分布，且$E(\epsilon_i) = 0,\ \mathrm{Var}(\epsilon_i) = \sigma^2$. 证明：$\beta_0, \beta_1$的LSE不相关的充要条件是$\bar{x} = \sum_{i=1}^n x_i/n = 0$.

第3章　区间估计

上一章我们讲述了若干点估计方法，由于点估计只是用一个点去估计未知参数，故在多数情况下会有误差. 于是，人们在实际生活中多用区间估计. 可以说，区间估计是介于估计与检验之间的内容，且区间估计与检验紧密相连，故有的书上也把区间估计看成是检验的一种. 本书为了对应点估计，将在本章专门讲述区间估计. 为了不过多讲述理论上的推导，故本章着重介绍基于枢轴量(Pivotal variable)的区间估计法，单样本与两样本正态总体的置信区间和信任推断法. 至于衡量置信区间的优良性准则，如一致最精确(Uniformly Most Accurate, 简记为UMA)、无偏置信区间和一致最精确无偏(UMAU)置信区间等内容，将不会在本书加以介绍，请有兴趣的读者参见陈希孺(1997)或陈希孺、倪国策(1988)或茆诗松、王静龙、濮晓龙(2006).

§3.1　区间估计中的几个基本概念

我们现在讲述的区间估计理论最早是由美籍波兰人—Neyman在上世纪30年代建立起来的. 在讲述方法之前，本节先介绍几个基本概念.

定义 3.1.1 (区间估计(interval estimation)) 设$X_1,\ldots,X_n$为来自参数分布族$\mathcal{F}=\{f(x,\theta):\theta\in\Theta\}$的样本, θ为一维未知参数. 如果$\hat\theta_L(\mathbf{X}),\hat\theta_U(\mathbf{X})$为两个统计量，且$\hat\theta_L(\mathbf{X})\le\hat\theta_U(\mathbf{X})$，则称随机区间$[\hat\theta_L(\mathbf{X}),\hat\theta_U(\mathbf{X})]$为$\theta$的一个区间估计.

从上述定义可以看出，给出一个未知参数的区间估计并不难，由于只要给出两个有大小关系的统计量即可. 既然是估计，就应有一个好坏的衡量标准. 当参数真值为θ时，我们自然希望随机区间$[\hat\theta_L(\mathbf{X}),\hat\theta_U(\mathbf{X})]$包含$\theta$的概率$P_\theta\{\hat\theta_L(\mathbf{X})\le\theta\le\hat\theta_U(\mathbf{X})\}$越大越好. 这个概率就称为置信度或置信水平(confidence level). 由于这个置信度依赖于参数真值，故我们自然希望对于参数空间Θ中的每一个，其置信水平都很大.

定义 3.1.2 (置信系数) 设随机区间$[\hat\theta_L(\mathbf{X}),\hat\theta_U(\mathbf{X})]$为$\theta$的一个区间估计，则称

$$\inf_{\theta\in\Theta}P_\theta\{\hat\theta_L(\mathbf{X})\le\theta\le\hat\theta_U(\mathbf{X})\}\tag{3.1.1}$$

为该区间估计的置信系数.

注 3.1.1　如果置信度不依赖于未知参数，则置信系数就是置信度.

注 3.1.2　区间估计有时要用开区间或半开半闭区间，但从置信度的角度看，这几种区间估计没有本质的区别.

注 3.1.3　在计算某区间估计的置信度时，我们应知道$\hat\theta_L(\mathbf{X}),\hat\theta_U(\mathbf{X})$的联合分布. 如不知道其联合分布，则得到的置信区间用途不大. 这也是置信区间的构造技巧之所在.

例 3.1.1　设$X_1,\ldots,X_n$为来自$N(\mu,\sigma^2)$的iid样本，其中μ,σ^2未知. 考虑μ的区间估计.

解　前面我们讲过的点估计中，多数是用样本均值估计总体均值. 对于本问题，我们知道，样本均值是μ的一个很好的点估计(矩估计、MLE，UMVUE等). 关于μ的区间估计，一个自然

的想法是基于样本均值，即我们可以用

$$(\bar{X} - kS_n/\sqrt{n}, \bar{X} + kS_n/\sqrt{n})$$

来作为μ的区间估计，其中k是一常数，S_n为样本标准差. 则它的置信度为

$$\alpha_k = P\{\bar{X} - kS_n/\sqrt{n} \le \mu \le \bar{X} + kS_n/\sqrt{n}\} = P\left\{\left|\frac{\sqrt{n}(\bar{X}-\mu)}{S_n}\right| \le k\right\} = P\{|T_n| \le k\},$$

其中

$$T_n = \frac{\sqrt{n}(\bar{X}-\mu)}{S_n} \sim t(n-1).$$

□

从上例可以看出，区间估计的置信度α_k仅依赖于样本容量n与k，且随着k的增大而增大. 如此区间的长度为∞，则其置信度自然会为1，达到最大. 但是这个区间对于θ的估计并没有什么用处. 于是，评价一个区间估计好坏的另一个准则就是其区间估计的长短或称为精确度(当然，关于区间估计精确度的定义还有其它方法，区间长度只是其中一个). 但是从定义可以看出，我们无法使一个区间估计的上述两个准则都达到最优. 现在我们所常用的准则是：先保证置信度达到一定水平后，再尽可能地提高其精度.

定义 3.1.3 (置信区间(Confident Interval)) 设$[\hat{\theta}_L(\mathbf{X}), \hat{\theta}_U(\mathbf{X})]$是参数$\theta$的一个区间估计，如果对给定的$\alpha \in (0,1)$，有

$$P_\theta\{\hat{\theta}_L(\mathbf{X}) \le \theta \le \hat{\theta}_U(\mathbf{X})\} \ge 1-\alpha,\ \forall\ \theta \in \Theta, \tag{3.1.2}$$

则称$[\hat{\theta}_L(\mathbf{X}), \hat{\theta}_U(\mathbf{X})]$为$\theta$的置信水平为$1-\alpha$的置信区间，$\hat{\theta}_L(\mathbf{X}), \hat{\theta}_U(\mathbf{X})$分别称为置信水平为$1-\alpha$的置信区间的置信下限与置信上限或双侧置信上(下)限.

从置信区间的定义可以看出，它可以保证真值落在此区间的概率至少为$1-\alpha$，这与前面讲的区间估计有着很大的不同，故以后我们所说的区间估计均是指上面定义的置信区间.

注 3.1.4 虽然上面定义的置信区间的置信度不小于$1-\alpha$，但在实际中我们均是求取满足

$$P_\theta\{\hat{\theta}_L(\mathbf{X}) \le \theta \le \hat{\theta}_U(\mathbf{X})\} = 1-\alpha,\ \forall\ \theta \in \Theta \tag{3.1.3}$$

的置信区间，有的书上称这种置信区间为同等置信区间，但在不致于引起混淆的时候，我们也简称它为置信区间.

在某些实际问题，人们感兴趣的指标可能是望大(越大越好的指标就称为望大指标)或望小(越小越好的指标就称为望小指标)的，如某产品的寿命就是望大指标，布匹上的疵点数就是望小指标. 对于望大(小)指标，人们关心的是在一定置信水平下的置信下(上)限. 这即是下面的定义.

定义 3.1.4 (单侧置信限) 设$\hat{\theta}_L(\mathbf{X})$与$\hat{\theta}_U(\mathbf{X})$是两上统计量，如对于给定的$\alpha \in (0,1)$，有

$$P_\theta\{\hat{\theta}_L(\mathbf{X}) \le \theta\} \ge 1-\alpha,\ \forall\ \theta \in \Theta, \tag{3.1.3}$$

$$P_\theta\{\hat{\theta}_U(\mathbf{X}) \geq \theta\} \geq 1-\alpha,\ \forall\,\theta \in \Theta, \tag{3.1.4}$$

则分别称$\hat{\theta}_L(\mathbf{X})$与$\hat{\theta}_U(\mathbf{X})$为$\theta$的置信水平为$1-\alpha$的单侧置信下限与上限.

注 3.1.5 一般在实际应用中，我们在(3.1.3)与(3.1.4)的右端取等号，这样的置信限也称为同等置信限.

至于单侧置信限与双侧置信限的关系如下：

定理 3.1.1 设$\hat{\theta}_L(\mathbf{X})$与$\hat{\theta}_U(\mathbf{X})$分别是参数$\theta$的置信水平为$1-\alpha_1$和$1-\alpha_2$的单侧(同等)置信下限与上限，且对于样本$X_1,\ldots,X_n$, 均有$\hat{\theta}_L(\mathbf{X}) \leq \hat{\theta}_U(\mathbf{X})$. 则$[\hat{\theta}_L(\mathbf{X}),\hat{\theta}_U(\mathbf{X})]$是$\theta$的置信水平为$1-(\alpha_1+\alpha_2)$的(同等)置信区间.

证明 因为$\{\hat{\theta}_L(\mathbf{X}) \leq \theta \leq \hat{\theta}_U(\mathbf{X})\}$，$\{\theta \leq \hat{\theta}_L(\mathbf{X})\}$与$\{\theta \geq \hat{\theta}_U(\mathbf{X})\}$是三个互不相容且其并为全事件的三个事件，又因为

$$P_\theta\{\theta < \hat{\theta}_L(\mathbf{X})\} = 1 - P_\theta\{\theta \geq \hat{\theta}_L(\mathbf{X})\} = 1-\alpha_1,$$

$$P_\theta\{\hat{\theta}_U(\mathbf{X}) < \theta\} = 1 - P_\theta\{\theta \leq \hat{\theta}_U(\mathbf{X})\} = 1-\alpha_2,$$

所以

$$P_\theta\{\hat{\theta}_L(\mathbf{X}) \leq \theta \leq \hat{\theta}_U(\mathbf{X})\} = 1-\alpha_1-\alpha_2.$$

□

上面我们仅考虑了单参数的置信区间，而多参数在许多实际问题中也经常遇到. 下面我们给出多参数的置信域的概念.

定义 3.1.5 (置信域(Confident Region)) 设$X_1,\ldots,X_n$是来自参数分布族$\mathcal{F}=\{f(x,\theta):\theta\in\Theta\subset R^k\}$的样本，$\theta=(\theta_1,\ldots,\theta_k)'$. 如果统计量$S(\mathbf{X})$满足

- 对任一样本观测值$\mathbf{x}$, $S(\mathbf{x})$是Θ的一个子集,
- 对给定的$\alpha\in(0,1)$，$P_\theta\{\theta\in S(\mathbf{X})\} \geq 1-\alpha,\ \forall\,\theta\in\Theta$,

则称$S(\mathbf{X})$是θ的置信水平为$1-\alpha$的置信域, 而概率$P_\theta\{\theta\in S(\mathbf{X})\}$在$\Theta$上的上确界就称为置信系数.

在实际应用中，置信域的形状是各式各样的，但一些规则的几何图形，如长方体、球、椭球等应用最多.

§3.2 枢轴量法

求取参数的置信区间的方法有许多，但其中最常用的是枢轴量法，尤其是对于连续型分布族. 本节将主要介绍枢轴量法，至于其它方法可参见茆诗松、王静龙、濮晓龙(2006).

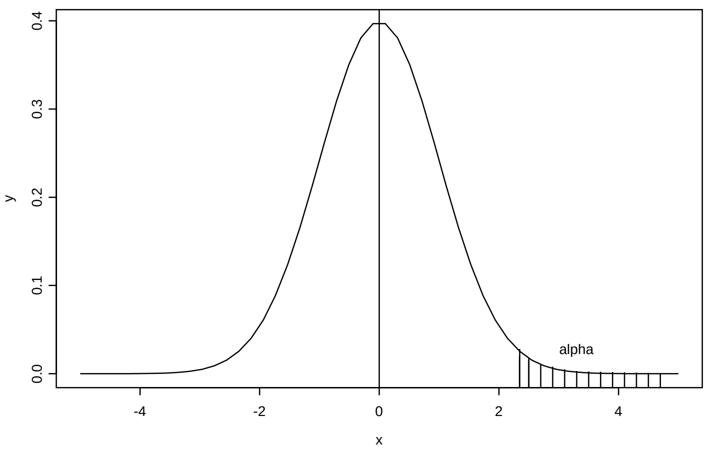

图3.1 标准正态分布的上侧α分位数u_α.

§3.2.1 小样本情况

为了引入用枢轴量法求取参数的置信区间，我们先看一个简单的例子.

例 3.2.1 设$X_1,\ldots,X_n$为来自正态总体$N(\mu,\sigma^2)$的iid样本，其中σ^2已知，μ未知. 求μ的$1-\alpha$的置信区间.

解 由上一章内容知，$\bar{X}$ 是μ的一个很好的点估计，且知道

$$\sqrt{n}(\bar{X}-\mu)/\sigma \sim N(0,1). \tag{3.2.1}$$

为以后的方便，我们始终以$\Phi(x),\phi(x)$分别表示标准正态分布$N(0,1)$的CDF和PDF，且用满足方程

$$\Phi(u_\alpha)=1-\alpha \tag{3.2.2}$$

的u_α表示标准正态分布的上侧α分位数，见图3.1.

虽然定义2.7.2已给出分位数的一般定义，但是我们注意到，那里的定义和(3.2.1)式的定义稍有不同. 事实上，如总体分布$F(x)$是连续且严格单调的，则总体$F(x)$的上侧α分位数ξ_α满足

$$F(\xi_p)=1-\alpha. \tag{3.2.3}$$

为了以后叙述的方便，本书后面均采用上侧分位数的概念. 在不会引起混淆的情况，我们也称上侧分位数为分位数，并且以$\chi^2_\alpha(n),t_\alpha(n),F_\alpha(m,n)$分别表示$\chi^2(n),t(n),F(m,n)$分布的上侧$\alpha$分位数. 关于这些常用分布的分位数，在一般的数理统计书中均有表可查，但应用时请注意是采用定义2.7.2式定义的分位数，还是如本节所定义的上侧分位数.

对于本例，由(3.2.1)式可知

$$P\left\{-u_{\alpha/2}\le \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}\le u_{\alpha/2}\right\}=\Phi(u_{\alpha/2})-\Phi(-u_{\alpha/2})=1-\alpha,$$

即

$$P\left\{\bar{X}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\le\mu\le\bar{X}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\right\}=1-\alpha,$$

这说明

$$\left[\bar{X}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2},\bar{X}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\right]$$

即是μ的置信水平为$1-\alpha$的置信区间. □

从上个例子可以看出，我们在求取某参数的置信区间时，一般是先给一个好的点估计，之后再通过这个点估计的分布求取感兴趣参数的置信区间. 这就是枢轴量的基本想法. 下面，我们把用枢轴量法求取置信区间的步骤总结如下：

❶ 找一个与待估参数$g(\theta)$有关的统计量T. 一般是它的一个很好的点估计(如例3.2.1中的$\bar{X}$).

❷ 设法找出T和$g(\theta)$的某函数$S(T,g(\theta))$，使得$S(T,g(\theta))$的分布$F(x)$与θ无关(如在例3.2.1中的$\sqrt{n}(\bar{X}-\mu)/\sigma$)，$S$就称为枢轴量.

❸ 适当地选取两个常数c,d，使对给定的$\alpha\in(0,1)$，有

$$P_\theta\{c\le S(T,g(\theta))\le d\}=1-\alpha,\tag{3.2.4}$$

即$F(d)-F(c)=1-\alpha$ (一般取$d=F_{\alpha/2}$, $c=F_{1-\alpha/2}$).

❹ 如果能把(3.2.4)中的不等式$c\le S(T,g(\theta))\le d$等价地改写成$\hat{\theta}_L(\mathbf{X})\le g(\theta)\le\hat{\theta}_U(\mathbf{X})$, 其中$\hat{\theta}_L(\mathbf{X}),\hat{\theta}_U(\mathbf{X})$只与$c,d$和$T$有关，而与$\theta$无关. 则$[\hat{\theta}_L(\mathbf{X}),\hat{\theta}_U(\mathbf{X})]$就是$g(\theta)$的置信水平为$1-\alpha$的置信区间.

注 3.2.1 在上述枢轴量方法中，最关键的是第二步寻找枢轴量.

例 3.2.2 设$X_1,\ldots,X_n$为来自正态总体$N(\mu,\sigma^2)$的iid样本，μ,σ^2均未知，试求μ的置信水平为$1-\alpha$ 的置信区间.

解 依枊轴量法的四个步骤，我们如下求取：

❶ 由于$\bar{X}$是μ的一个好的点估计，故我们在第一步中取$T=\bar{X}$.

❷ 由于σ未知，故我们不能取$\sqrt{n}(\bar{X}-\mu)/\sigma$作为枢轴量(虽然它的分布与参数$\mu$无关，但是它与另一未知的讨厌参数$\sigma$有关，故第四步行不通). 此时，既然$\sigma$未知，就用样本标准差估计之，于是，我们可以取$S=\sqrt{n}(\bar{X}-\mu)/S_n$.

❸ 因为$\sqrt{n}(\bar{X}-\mu)/S_n\sim t(n-1)$，故取$c=t_{1-\alpha/2}(n-1)=-t_{\alpha/2}(n-1),d=t_{\alpha/2}(n-1)$.

❹ 因为

$$-t_{\alpha/2}(n-1)\le\sqrt{n}(\bar{X}-\mu)/S_n\le t_{\alpha/2}(n-1)$$

$$\Longleftrightarrow\ \bar{X}-\frac{S_n}{\sqrt{n}}t_{\alpha/2}(n-1)\le\mu\le\bar{X}+\frac{S_n}{\sqrt{n}}t_{\alpha/2}(n-1),$$

故

$$\bar{X} \pm \frac{S_n}{\sqrt{n}} t_{\alpha/2}(n-1)$$

即为μ的置信水平为$1-\alpha$的置信区间. □

对于本例，假设我们在如下的实际问题：为估计一物体重量μ，我们把它放在一架天平上称了5次，记录到的结果如下（单位：克）

$$5.52,\ 5.48,\ 5.64,\ 5.51,\ 5.43.$$

假设此天平没有系统误差，且测量误差服从正态分布$N(0,\sigma^2)$. 则我们可以利用例3.2.2的结果求μ的置信区间.

此时，$\bar{X}=5.516$, $S_n=0.078$，且查表知$t_{0.025}(4)=2.776$. 于是，μ的置信水平为0.95的置信区间为$[5.419, 5.613]$.

我们知道$[5.419, 5.613]$是一个具体的区间，而μ是一个虽未知但其值却是确定的数. 于是，区间$[5.419, 5.613]$或者包含μ，或者不包含μ，二者必居其一. 那置信水平0.95如何解释呢？事实上，置信水平的确切含义是：如果我们把上述试验重复100次，则将得到100个这样的置信水平为0.95的置信区间. 于是，我们有理由相信在这100个区间中，至少有95次确实包含所要估计的参数的真值. 而一旦得到具体的区间，我们就不能再说它有$(1-\alpha)\times 100\%$的可能性包含要估计的参数真值了.

例 3.2.3 设$X_1,\ldots,X_n$为来自正态总体$N(\mu,\sigma^2)$的iid样本，试求σ^2的置信区间.

解 对于此问题，由前面知识知道，样本方差S_n^2是σ^2的一个很好的点估计，且由定理1.3.4知，

$$(n-1)S_n^2/\sigma^2 \sim \chi^2(n-1),$$

于是，我们可取此问题的枢轴量为$(n-1)S_n^2/\sigma^2$. 由此可以得到σ^2的置信水平为$1-\alpha$的置信区间为

$$\left[(n-1)S_n^2/\chi^2_{\alpha/2}(n-1), (n-1)S_n^2/\chi^2_{1-\alpha/2}(n-1)\right].$$

□

例 3.2.4 设$X_1,\ldots,X_n$为来自均匀分布总体$U(0,\theta)$的iid样本，试求θ的置信水平为$1-\alpha$的置信区间.

解 对于此总体，我们在前面讲过，$T=X_{(n)}$ 是θ的充分统计量，也是θ的MLE. 且容易求得$X_{(n)}/\theta$的PDF为

$$f(t,\theta)=nt^{n-1},\ 0<t<1$$

与θ无关，于是我们就取$X_{(n)}/\theta$作为此问题的枢轴量. 对于给定的水平$1-\alpha$，我们可选取c,d满足

$$P_\theta\left\{c \le \frac{X_{(n)}}{\theta} \le d\right\} = d^n - c^n = 1-\alpha,$$

由此得到θ的置信水平为$1-\alpha$的置信区间为$\left[\frac{X_{(n)}}{d}, \frac{X_{(n)}}{c}\right]$.

此区间长度为$X_{(n)}\left(\frac{1}{c}-\frac{1}{d}\right)$，可以证明，当$d=1, c=\sqrt[n]{\alpha}$时，其长度达到最小. 这样，$[X_{(n)}, X_{(n)}/\sqrt[n]{\alpha}]$就是$\theta$的置信水平为$1-\alpha$的最短置信区间. □

例 3.2.5 设某产品的寿命X服从指数分布$E(\lambda)$，试求其平均寿命$1/\lambda$的置信限.

解 对于此问题，其总体的PDF为

$$f(x,\lambda)=\lambda e^{-\lambda x},\ x>0.$$

由于寿命是望大指标，故我们仅需求取其置信下限. 由例1.4.8知，$T=\sum_{i=1}^n X_i$是λ的充分统计量，故且例2.4.7知，T/n也是$1/\lambda$的UMVUE. 另外，由Γ分布的可加性知，$T\sim\Gamma(n,\lambda)$，故知

$$2\lambda T\sim\Gamma(n,1/2)=\chi^2(2n),$$

故我们取$2\lambda T$为此问题的枢轴量，且知

$$P\{2\lambda T\le\chi^2_\alpha(2n)\}=1-\alpha,$$

于是，此产品平均寿命$1/\lambda$的置信水平为$1-\alpha$的置信下限为$\frac{2T}{\chi^2_\alpha(2n)}=\frac{2n\bar{X}}{\chi^2_\alpha(2n)}$.

例 3.2.6 设$X_1,\ldots,X_n$为来自正态总体$N(\mu,\sigma^2)$的iid样本，μ已知. 试求σ^2的置信水平为$1-\alpha$的置信区间.

解 由于此时$\sum_{i=1}^n(X_i-\mu)^2$是σ^2 的充分统计量，且$\sum_{i=1}^n(X_i-\mu)^2/\sigma^2\sim\chi^2(n)$，故它就是本问题的枢轴量. 另外，由于

$$P\left\{\chi^2_{1-\alpha/2}(n)\le\frac{\sum_{i=1}^n(X_i-\mu)^2}{\sigma^2}\le\chi^2_{\alpha/2}(n)\right\}=1-\alpha,$$

故求得σ^2的置信水平为$1-\alpha$的置信区间为

$$\left[\frac{\sum_{i=1}^n(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)}, \frac{\sum_{i=1}^n(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}\right].$$

□

通过上面的例子我们可以看到，枢轴量在上述置信区间的求取过程中起着“轴心的作用”. 但是，我们在前面考虑的例子均是针对连续型总体的，这是由于枢轴量法对于离散型随机变量并不容易操作，其原因在于对于给定的α，一般不存在确切的分位点. 下面，我们将讨论二项和Poisson分布的置信区间的求取方法.

§3.2.2 大样本情况

本小节将从极限分布的角度，用两个例子来说明如何近似求取离散型总体参数的置信区间.

例 3.2.7 设$X_1,\ldots,X_n$为来自Bernoulli总体$b(1,p)$的iid样本，试求p的置信水平为$1-\alpha$的置信区间.

解 对于此问题，由于我们知道$\sum_{i=1}^n X_i$是充分统计量，故枢轴量应与它有关. 又因为$\sum_{i=1}^n X_i \sim B(n,p)$, 故我们不可能由此构造一个满足枢轴量法条件的❷与❹. 但由中心极限定理可知

$$\frac{\sum_{i=1}^n X_i - np}{\sqrt{npq}} \xrightarrow{\mathcal{L}} N(0,1),\ \text{当} n \to \infty,$$

即当n充分大时，我们有

$$P\left\{\frac{T_n - np}{\sqrt{npq}} < x\right\} \approx \Phi(x),$$

且与p无关. 也就是说，当n充分大时，

$$P\left\{-u_{\alpha/2} \le \frac{T_n - np}{\sqrt{npq}} \le u_{\alpha/2}\right\} \approx 1-\alpha. \tag{*}$$

又因为

$$-u_{\alpha/2} \le \frac{T_n - np}{\sqrt{npq}} \le u_{\alpha/2} \Longleftrightarrow n(\bar{X}-p)^2 \le p(1-p)u_{\alpha/2}^2,$$

而上式右端的关于p的二次三项式: $n(\bar{X}-p)^2 - p(1-p)u_{\alpha/2}^2$的判别式

$$\Delta = (-2n\bar{X} - u_{\alpha/2}^2)^2 - 4(n+u_{\alpha/2}^2) = 4nu_{\alpha/2}^2\bar{X}(1-\bar{X}) > 0,$$

且其平方项系数大于零，故可以求得$n(\bar{X}-p)^2 - p(1-p)u_{\alpha/2}^2 = 0$的两个根为

$$p = \frac{n}{n+u_{\alpha/2}^2}\left[\bar{X} + \frac{u_{\alpha/2}^2}{2n} \pm u_{\alpha/2}\sqrt{\frac{\bar{X}(1-\bar{X})}{n} + \frac{u_{\alpha/2}^2}{4n^2}}\right],$$

分别以$\hat{p}_L < \hat{p}_U$记之. 则当$p \in (\hat{p}_L, \hat{p}_U)$时，上述二次三项式小于零. 于是由(*)式知，

$$P\{\hat{p}_L \le p \le \hat{p}_U\} \approx 1-\alpha,$$

即说明$[\hat{p}_L, \hat{p}_U]$是p的置信水平近似$1-\alpha$的置信区间. □

我们注意到，上述的近似置信水平$1-\alpha$的置信区间，当n充分大时，它非常接近水平$1-\alpha$. 在实际中，当$n \ge 30$，就可以认为是充分大了.

例 3.2.8 设$X_1, \ldots, X_n$为来自Poisson分布$P(\lambda)$的iid样本，试求λ的置信水平为$1-\alpha$的置信区间.

解 此时λ的充分统计量为$S_n = \sum_{i=1}^n X_i$, 且$S_n \sim P(n\lambda)$. 由中心极限定理知

$$\frac{S_n - n\lambda}{\sqrt{n\lambda}} \xrightarrow{\mathcal{L}} N(0,1),$$

即当n充分大时，有

$$P\left\{-u_{\alpha/2} \le \frac{S_n - n\lambda}{\sqrt{n\lambda}} \le u_{\alpha/2}\right\} \approx 1-\alpha.$$

因此，用类似于上例中的方法，设$\hat{\lambda}_L, \hat{\lambda}_U$为二次方程

$$(S_n - n\lambda)^2 = n\lambda u_{\alpha/2}^2$$

的根，即

$$\lambda_{L,U} = \bar{X} + u_{\alpha/2}^2 \pm u_{\alpha/2}\sqrt{\frac{u_{\alpha/2}^2}{4n^2} + \frac{\bar{X}}{n}}.$$

于是，λ的置信水平近似为$1-\alpha$的置信区间为$[\hat{\lambda}_L, \hat{\lambda}_U]$. □

由于正态总体的实用性，故我们在结束本节之前，把关于单个正态总体参数的置信区间总结如下：

表3.1 单样本正态总体参数的置信区间

参数情况	μ的置信区间	σ^2的置信区间
σ^2已知	$\left[\bar{X} - u_{\alpha/2}\frac{\sigma}{\sqrt{n}}, X + u_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$	
σ^2未知	$\left[\bar{X} - t_{\alpha/2}(n-1)\frac{S_n}{\sqrt{n}}, X + t_{\alpha/2}(n-1)\frac{S_n}{\sqrt{n}}\right]$	
μ已知		$\left[\frac{\sum_{i=1}^n (X_i-\mu)^2}{\chi_{\alpha/2}^2(n)}, \frac{\sum_{i=1}^n (X_i-\mu)^2}{\chi_{1-\alpha/2}^2(n)}\right]$
μ未知		$\left[\frac{(n-1)S_n^2}{\chi_{\alpha/2}^2(n-1)}, \frac{(n-1)S_n^2}{\chi_{1-\alpha/2}^2(n-1)}\right]$

§3.3 两个正态总体的置信区间

在上一节中，我们讲述了构造连续总体参数的置信区间方法—枢轴量法，并且针对单样本正态总体，给出了其参数的置信区间. 两样本问题也经常出现在许多实际问题中，故本节将考虑两样本正态总体参数的置信区间问题.

设$X_1, \cdots, X_m$和$Y_1, \ldots, Y_n$为分别来自正态总体$N(\mu_1, \sigma_1^2)$和$N(\mu_2, \sigma_2^2)$的样本，且全样本是独立的，其中$\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$为参数.

§3.3.1 Behrens-Fisher问题

我们感兴趣的是求$\delta = \mu_2 - \mu_1$的置信区间. 这个问题是由Behrens于1929年提出的，由于Fisher也讨论过这个问题，故人们习惯称这个问题为Behrens-Fisher问题. 求取δ的置信区间的问题之所以成为一个很有名的问题，是由于当我们对σ_1^2, σ_2^2一无所知时，不知如何求取，虽然，当$\sigma_1^2 = \sigma_2^2 = \sigma^2$未知时，此问题已得到了很好的解决.

设样本均值为$\bar{X} = \sum_{i=1}^m X_i/m, \bar{Y} = \sum_{i=1}^n Y_i/n$, 样本方差为$S_{1m}^2 = \sum_{i=1}^m (X_i - \bar{X})^2/(m-1), S_{2n}^2 = \sum_{i=1}^n (Y_i - \bar{Y})^2/(n-1)$. 下面分几种情况讨论$\delta$的置信区间.

(1) 当$\sigma_1^2 = \sigma_2^2 = \sigma^2$未知时

因为统计量

$$T = \sqrt{\frac{mn(m+n-2)}{m+n}} \frac{\bar{Y} - \bar{X} - \delta}{\sqrt{(m-1)S_{1m}^2 + (n-1)S_{2n}^2}} \sim t(m+n-2), \tag{3.3.1}$$

故可以取它作为枢轴量. 于是，由此可求得δ的置信水平为$1-\alpha$的置信区间为

$$\left[\bar{Y}-\bar{X}\pm t_{\alpha/2}(m+n-2)\sqrt{\frac{m+n}{mn(m+n-2)}}\sqrt{(m-1)S_{1m}^2+(n-1)S_{2n}^2}\right].$$

(2) 当$\theta=\sigma_2^2/\sigma_1^2$已知时

因为统计量

$$T=\sqrt{\frac{mn(m+n-2)}{m\theta+n}}\frac{\bar{Y}-\bar{X}-\delta}{\sqrt{(m-1)S_{1m}^2+(n-1)S_{2n}^2/\theta}}\sim t(m+n-2), \tag{3.3.2}$$

于是，δ的置信水平为$1-\alpha$的置信区间为

$$\left[\bar{Y}-\bar{X}\pm t_{\alpha/2}(m+n-2)\sqrt{\frac{m\theta+n}{mn(m+n-2)}}\sqrt{(m-1)S_{1m}^2+(n-1)S_{2n}^2/\theta}\right].$$

(3) 当$m=n$时

因为$Z_i=Y_i-X_i\sim N(\delta,\sigma_1^2+\sigma_2^2)$, $i=1,\cdots,n$, 且相互独立，于是知$\bar{Z}\sim N(\delta,(\sigma_1^2+\sigma_2^2)/n)$, $\sum_{i=1}^n(Z_i-\bar{Z})^2/(\sigma_1^2+\sigma_2^2)\sim\chi^2(n-1)$. 故有

$$\frac{\sqrt{n(n-1)}(\bar{Z}-\delta)}{\sqrt{\sum_{i=1}^n(Z_i-\bar{Z})^2}}\sim t(n-1), \tag{3.3.3}$$

由此可得δ的置信水平为$1-\alpha$的置信区间为

$$\left[\bar{Z}\pm t_{\alpha/2}(n-1)\frac{\sqrt{\sum_{i=1}^n(Z_i-\bar{Z})^2}}{\sqrt{n(n-1)}}\right],$$

(4) 当m,n都充分大时

因为$S_{1m}^2\xrightarrow{P}\sigma_1^2$, $S_{2n}^2\xrightarrow{P}\sigma_2^2$, 且$\bar{Y}-\bar{X}\sim N(0,\sigma_1^2/m+\sigma_2^2/n)$, 故由Slutskey定理知

$$\tilde{T}=\frac{\bar{Y}-\bar{X}-\delta}{\sqrt{S_{1m}^2/m+S_{2n}^2/n}}\xrightarrow{\mathcal{L}}N(0,1), \tag{3.3.4}$$

于是，δ的置信水平为$1-\alpha$的置信区间为

$$\left[\bar{Y}-\bar{X}\pm u_{\alpha/2}\sqrt{S_{1m}^2/m+S_{2n}^2/n}\right],$$

(5) 一般情况

当m,n不是很大时，(3.3.4)式左边的随机变量既不是t分布，也不是正态分布. Welch于1938年通过研究认为$\tilde{T}$的分布近似地与自由度为r的t分布相同，其自由度的选取可如下考虑：定义

$$T=\frac{\bar{Y}-\bar{X}-\delta}{\sqrt{\sigma_1^2/m+\sigma_2^2/n}},\quad S_{mn}^2=\frac{S_{1m}^2}{m}+\frac{S_{2n}^2}{n},\quad Z=\frac{S^2}{\sigma_1^2/m+\sigma_2^2/n},$$

显然$T \sim N(0,1)$, Z接近χ^2分布，并且(3.3.4)式的$\tilde{T}$可以改写成$\tilde{T} = T/\sqrt{Z}$, 于是，可以认为$\tilde{T}$近似地服从一个t分布. 如果$\tilde{T}$的自由度为r，则可以近似地认为rZ近似地服从$\chi^2(r)$分布，即$\mathrm{Var}Z \doteq 2/r$. 又由Z的定义可求

$$\mathrm{Var}Z = \frac{\frac{2\sigma_1^4}{m^2(m-1)} + \frac{2\sigma_2^4}{n^2(n-1)}}{(\sigma_1^2/m + \sigma_2^2/n)^2},$$

故可得

$$r \doteq \frac{(\sigma_1^2/m + \sigma_2^2/n)^2}{\frac{2\sigma_1^4}{m^2(m-1)} + \frac{2\sigma_2^4}{n^2(n-1)}},$$

但由于σ_i^2是未知的，故我们可用样本方差估计它，这样就得到了r的近似值为

$$r \doteq \frac{S_{mn}^4}{\frac{S_{1m}^4}{m^2(m-1)} + \frac{S_{2n}^4}{n^2(n-1)}}. \tag{3.3.5}$$

由此可得δ的置信水平近似为$1-\alpha$的置信区间为

$$\left[\bar{Y} - \bar{X} \pm t_{\alpha/2}(r)S\right].$$

例 3.3.1 某厂利用两条自动化流水线装蕃茄酱罐头. 现分别从两条流水线上随机地各抽取一组样本容量为6和7的样本，分别称重后得(单位：克)

$$\bar{X} = 10.6, \bar{Y} = 10.1,\ S_{1m}^2 = 0.0125, S_{2n}^2 = 0.01.$$

假设两条流水线上所装蕃茄酱的重量都服从正态分布，其均值分别为μ_1, μ_2，而方差不等. 现求$\mu_1 - \mu_2$的置信水平为0.90的置信区间.

解 对于此问题，我们只能利用(5)的一般情形近似计算之. 因为

$$S_{mn}^2 = \frac{0.0125}{6} + \frac{0.01}{7} = 0.003512,\ S_{mn} = 0.05926,$$

$$l = 0.003512^2 / \left(\frac{0.0125^2}{5 \times 6^2} + \frac{0.01^2}{6 \times 7^2}\right) = 10.209 \approx 10.$$

又因为$t_{\alpha/2}(10) = 1.812$, 于是$\mu_1 - \mu_2$的置信水平为0.9的置信区间为$0.5 \pm 0.05926 \times 1.812 = [0.393, 0.607]$. □

§3.3.2 方差比σ_1^2/σ_2^2的置信区间

因为$(m-1)S_{1m}^2/\sigma_1^2 \sim \chi^2(m-1)$, $(n-1)S_{2n}^2/\sigma_2^2 \sim \chi^2(n-1)$, 且二者是独立的，于是

$$F = \frac{S_{1m}^2/\sigma_1^2}{S_{2n}^2/\sigma_2^2} \sim F(m-1, n-1),$$

可以作为枢轴量，并且

$$P\left\{F_{1-\alpha/2}(m-1,n-1)\le\frac{S_{1m}^2/\sigma_1^2}{S_{2n}^2/\sigma_2^2}\le F_{\alpha/2}(m-1,n-1),\right\}$$

这样，就得到了σ_1^2/σ_2^2的置信区间为

$$\left[\frac{S_{1m}^2/S_{2n}^2}{F_{\alpha/2}(m-1,n-1)},\frac{S_{1m}^2/S_{2n}^2}{F_{1-\alpha/2}(m-1,n-1)}\right].$$

我们把两样本正态总体参数的置信区间总结如下：

表3.2 两样本正态总体参数的置信区间

讨厌参数	待估参数	置信区间
$\sigma_1^2=\sigma_2^2=\sigma^2$未知	$\mu_2-\mu_1$	$\left[\bar{Y}-\bar{X}\pm t_{\alpha/2}(m+n-2)\sqrt{\frac{m+n}{mn(m+n-2)}}(Q_1^2+Q_2^2)\right]$
$\sigma_2^2/\sigma_1^2=\theta$已知	$\mu_2-\mu_1$	$\left[\bar{Y}-\bar{X}\pm t_{\alpha/2}(m+n-2)\sqrt{\frac{m+n}{mn(m+n-2)}}(Q_1^2+Q_2^2/\theta)\right]$
$m=n$时	$\mu_2-\mu_1$	$\left[\bar{Z}\pm t_{\alpha/2}(n-1)\frac{\sqrt{\sum_{i=1}^n(Z_i-\bar{Z})^2}}{\sqrt{n(n-1)}}\right]$
m,n充分大	$\mu_2-\mu_1$	$\left[\bar{Y}-\bar{X}\pm u_{\alpha/2}S\right]$
一般情况	$\mu_2-\mu_1$	$\left[\bar{Y}-\bar{X}\pm t_{\alpha/2}(r)S\right]$
μ_1,μ_2未知	σ_1^2/σ_2^2	$\left[\frac{S_{1m}^2/S_{2n}^2}{F_{\alpha/2}(m-1,n-1)},\frac{S_{1m}^2/S_{2n}^2}{F_{1-\alpha/2}(m-1,n-1)}\right]$

其中$Q_1^2=\sum_{i=1}^m(X_i-\bar{X})^2,Q_2^2=\sum_{i=1}^n(Y_i-\bar{Y})^2,Z_i=Y_i-X_i,S_{mn}^2=S_{1m}^2/m+S_{2n}^2/n.$

§3.4 信仰推断方法

§3.4.1 信仰分布

我们先回忆一下极大似然思想：设$X\sim p(x;\theta)$，当有了样本观测值x后，如果对不同的θ_1,θ_2，$f(x;\theta_1)>f(x;\theta_2)$，则我们认为$x$来自总体$f(x;\theta_1)$的可能性要比来自$f(x;\theta_2)$的可能大，也就是用$\theta_1$估计$\theta$比用$\theta_2$估计$\theta$来的可靠些. 这就是说，当有了样本后，用最有可能出现的θ作为θ的估计. 之后，于三十年代，Fisher把上述似然思想推广而提出了信仰分布(fiducial distribution).

我们先看一个例子. 设有一个样本$X\sim N(\theta,1)$, θ是未知参数. 对于这个样本，我们有$X-\theta\sim N(0,1)$，即$\forall\ t$，有

$$P\{X-\theta<t\}=\Phi(t),\tag{3.4.1}$$

此式也可写成

$$P\{\theta>X-t\}=\Phi(t),\ \text{或}\ P\{\theta<u\}=1-\Phi(X-u).\tag{3.4.2}$$

虽然从(3.4.1)到(3.4.2)，只是写法的不同，但Fisher即认为有着本质的不同. Fisher理解(3.4.2)如下：有了样本后，把θ看成一个随机变量，而上式则给出了θ的分布，这就是θ的信仰分布.

为什么可以把θ看成随机变量呢？按Fisher的思想，在进行抽样得到样本X之前，θ是一个未知数，我们对它一无所知，可以说，它什么值都可能取，但并不知道它取什么值的可能性有多大. 但是在抽样得到X后，虽然θ仍然有可能取各种值，但是对取这些值的可能性我们就有了一定的了解，这就是信仰分布所提供的信息. 因此，在这种情况下，样本X的作用，或样本X所提供的信息，就表现在它破除了我们对θ的完全无知，而以一个概率分布的形式总结了对θ的新认识.

我们注意到，这种信仰分布与Bayes方法的先验分布是不一样的，也可以说有着本质的不同. 但也有一点是相同的，即参数θ的信仰分布与先验分布都无法用频率来解释.

例 3.4.1 设样本$X_1,\ldots,X_n$来自正态总体$\sim N(\mu,1)$, μ是未知参数. 试用信仰分布求取μ的置信区间.

解 由于$\bar{X}$是μ的充分统计量，故我们用$\bar{X}$来考虑此问题. 因为$\bar{X}\sim(\mu,1/n)$,故可记为

$$\bar{X}=\mu+\frac{1}{n}\varepsilon,\ \text{其中}\ \varepsilon\sim N(0,1). \tag{3.4.3}$$

我们改写上述模型如下：

$$\mu=\bar{X}-\frac{1}{n}\varepsilon, \tag{3.4.4}$$

由此得到了μ的信仰分布：$\mu\sim N(\bar{X},1/n)$，由此分布就得到了μ的区间估计：

$$P\left\{\bar{X}-\frac{1}{\sqrt{n}}u_{\alpha/2}\le\mu\le\bar{X}+\frac{1}{\sqrt{n}}u_{\alpha/2}\right\}=1-\alpha.$$

此时，Fisher称$1-\alpha$为信仰系数，$\left[\bar{X}-\frac{1}{\sqrt{n}}u_{\alpha/2},\bar{X}+\frac{1}{\sqrt{n}}u_{\alpha/2}\right]$称为$\mu$的信仰水平为$1-\alpha$的区间估计. □

例3.4.1中由信仰分布得到的区间估计与我们前面得到的置信区间是一致的，但有时二者是不同的.

例 3.4.2 用信仰分布考虑Behrens-Fisher问题

解 注意到，我们有如下的分布：

$$t_1=\frac{\sqrt{m}(\bar{X}-\mu_1)}{S_{1m}}\sim t(m-1),\quad t_2=\frac{\sqrt{n}(\bar{Y}-\mu_2)}{S_{2n}}\sim t(n-1),$$

且t_1与t_2是独立的，于是有

$$\mu_1=\bar{X}-S_{1m}t_1/\sqrt{m},\quad \mu_2=\bar{Y}-S_{2n}t_2/\sqrt{n},$$

即

$$\mu_2-\mu_1=\bar{Y}-\bar{X}+(S_{1m}t_1/\sqrt{m}-S_{2n}t_2/\sqrt{n}),$$

其中$S_{1m}t_1/\sqrt{m}-S_{2n}t_2/\sqrt{n}$是两个相互独立的$t$分布的线性组合，我们把它改写如下：

$$S_{1m}t_1/\sqrt{m}-S_{2n}t_2/\sqrt{n}=r(t_1\cos\theta-t_2\sin\theta),$$

其中$r=\sqrt{S_{1m}^2/m+S_{2n}^2/n}, \cos\theta=\frac{S_{1m}}{r\sqrt{m}}, 0<\theta<\pi/2$.

定义$W=t_1\cos\theta-t_2\sin\theta$，于是，

$$\mu_2-\mu_1=\bar{Y}-\bar{X}+rW, \tag{$*$}$$

其中W的分布仅与θ和t分布的自由度$m-1$和$n-1$有关. 可以证明，W的分布关于原点对称，并且Fisher和Yates已经编制了其分位数表. 如以$w_{\alpha/2}(\theta,m-1,n-1)$记其主侧$\alpha/2$分位数，则由(*)式知$\delta$的信仰水平为$1-\alpha$的区间估计为

$$[\bar{Y}-\bar{X}-rw_{\alpha/2}(\theta,m-1,n-1), \bar{Y}-\bar{X}+rw_{\alpha/2}(\theta,m-1,n-1)].$$

□

由此例可以看出，从信仰分布的角度出发，可以较园满地解决Behrens-Fisher问题. 这也是它受到重视的一个原因. 但由于在寻求信仰分布时，利用不同方法导出的信仰分布有可能不一样，故也引起了大家的质疑，这也是它无法完全被大家接受的一个原因.

§3.4.2 函数模型方法

关于信仰分布的推导，有好几种方法，而其中用建立类似于(3.4.4)的函数模型方法是较为简单的一种，下面简单地介绍一下函数模型方法.

定义 3.4.1 函数模型有三个要素: 取值于可测空间$(\mathcal{X},\mathcal{B}_{\mathcal{X}})$上的观察值$X$、参数$\theta\in\Theta$、误差变量$\varepsilon$, 且$\varepsilon$ 的分布P与θ无关. 如果在这三个量之间存在一个函数关系$X=\theta\circ\varepsilon:(\theta,\varepsilon)\to X$, 使得$\forall\ X\in\mathcal{X}$, 存在$\theta\in\Theta, \varepsilon\sim P$, 使得$X=\theta\circ\varepsilon$. 则称$X=\theta\circ\varepsilon, \varepsilon\sim P$为函数模型，记为$X=\theta\circ\varepsilon$.

例 3.4.3 设样本$X_1,\ldots,X_n$来自$U(0,\theta)$,其中$\theta>0$未知，试求θ的区间估计.

解 由于$X_{(n)}$是θ的充分统计量，故我们将基于它考虑问题. 因为$X_{(n)}$的PDF为$p(y,\theta)=ny^{n-1}/\theta^n, 0\le y\le\theta$,令$\varepsilon=X_{(n)}/\theta$,则$\varepsilon$的PDF为$p(t)=nt^{n-1}, 0\le t\le 1$, 与$\theta$无关. 另由$\varepsilon$的定义有如下的模型：

$$X_{(n)}=\theta\varepsilon,$$

于是，当有了样本后，θ的信仰分布可由$\theta=X_{(n)}/\varepsilon$求得，其PDF为$nX_{(n)}^n/\theta^{n+1}, \theta\ge X_{(n)}$.

由于θ的信仰分布的PDF是θ的严格单减函数，故区间长度尽可能短的角度出发，我们取形如$[X_{(n)},\hat{\theta}_u]$ 的区间作为θ的区间估计，从而由

$$P\{X_{(n)}\le\theta\le\hat{\theta}_u\}=\int_{X_{(n)}}^{\hat{\theta}_u} nX_{(n)}^n/\theta^{n+1}d\theta=1-\alpha,$$

求得$\hat{\theta}_u=X_{(n)}/\sqrt[n]{\alpha}$. 于是$\theta$的信仰水平为$1-\alpha$的区间估计为$\left[X_{(n)}, X_{(n)}/\sqrt[n]{\alpha}\right]$. □

§3.5 容许(tolerance)区间与容许限

此部分内容请参见茆诗松、王静龙(1998).

习题三

1. 假设在105次独立射击中有60次命中目标，试求命中率p的95%的置信区间.

2. 设$X_1, \ldots, X_n$为来自Poisson分布$P(\lambda)$的iid样本，试求λ的置信水平近似为$1-\alpha$的大样本置信区间.

3. 某通讯公司随机抽查了1万个手机短信的字符长度，得到其平均值为$\bar{X}=9$, 样本标准差为$S=18$. 在置信水平95%下，求该公司用户手机短信均值μ的置信区间.

4. 现在南开大学随机调查了1000名学生，发现其中651名同学有手机. 请求我校有手机同学比例p的置信水平为0.95的置信区间.

5. 随机抽取某大学60位正教授的住房面积，得知他们的平均住房面积为$115.4m^2$，标准差是$15.8m^2$. 又随机地调查了该大学60位副教授的住房面积，得知他们的平均住房面积为$89.3m^2$，标准差为$21.3m^2$. 试在置信水平0.95下，

(1) 求该大学正教授平均住房面积的置信区间；

(2) 求该大学副教授平均住房面积的置信区间；

(3) 如认为两个总体的方差等于样本方差时，求该校正副教授住房面积差的置信区间.

6. 科学中的伟大发现往往是由比较年轻的人提出的. 下表是16世纪中叶到20世纪的12个重大科学突破的情况：

科学发现	科学家	年份	年龄
日心说	哥白尼	1543	40
望远镜及天文学的基本定律	伽利略	1600	43
动力学、万有引力、微积分	牛顿	1665	23
电的实质	富兰克林	1746	40
燃烧即氧化	拉瓦锡	1774	31
地球的演变	莱尔	1830	33
进化论	达尔文	1858	49
光的电磁场特性	麦克斯韦尔	1864	33
放射性	居里	1896	34
量子力学	普朗克	1901	43
狭义相对论	爱因斯坦	1905	26
量子力学的数学基础	薛定谔	1926	39

设提出重大科学发现时科学家的年龄服从$N(\mu, \sigma^2)$分布(μ, σ未知)，构造μ的置信水平为0.95的置信区间.

7. 设$X_1, \ldots, X_n$为来自分布为$F(x)$的总体的iid样本，对于固定的$x \in R$，试求总体分布$F(x)$的置信水平近似为$1-\alpha$的大样本置信区间.

8. 假设院领导想考察两位老师给学生的期末考试成绩是否有显著性的区别. 假设本学期共有50位学生选修老师甲的课程，且其期末平均成绩为84.5，标准差为14.5，而选修老师乙所讲授课程的学生人数为115，期末平均成绩为88.3，标准差为6.8. 如假设学生成绩服从正态分布，则请给出甲乙两位老师所教课程的平均成绩的置信水平为0.95的置信区间以及两位老师所教课程平均成绩差的置信水平为0.95的置信区间.

9. 设有来自对数正态分布$LN(\mu,0.8^2)$的iid样本为：4.95，1.73，10.07，9.78，14.15，1.52，11.7，6.62，10.38. 试求

(1) 总体期望$\beta=EX$；

(2) μ的置信水平为0.95的置信区间；

(3) β的置信水平为0.95的置信区间.

10. 为制定高中学生的体育锻炼标准，某区教育局在该区高中随机抽选了30名男生，测验100米跑成绩，结果平均为13.5秒，标准差为0.1秒. 假定高中男生100米跑成绩服从正态分布，试建立全区高中男生100米跑成绩的置信水平为0.95的置信区间.

11. 为估计每个家庭月平均收入，现抽取一定容量的简单随机样本，假设样本标准差为350元. 问为以95%的概率使区间的长度为360元，需要抽取多少家庭？

12. 设$X_1,\ldots,X_n$为来自正态总体$N(\mu,\sigma_0^2)$的iid样本，其中σ_0为已知常数. 又设介于0与1间的三个常数α_1,α_2,α满足：$\alpha_1+\alpha_2=2\alpha$. 又设$\mu$的置信水平$1-\alpha$的置信区间为

$$\left(\bar{X}-u_{\alpha_1}\frac{\sigma_0}{\sqrt{n}},\ \bar{X}+u_{\alpha_2}\frac{\sigma_0}{\sqrt{n}}\right).$$

试证明当$\alpha_1=\alpha_2$时，上述置信区间的长度最小.

13. 设$X_1,\ldots,X_n$和$Y_1,\ldots,Y_m$分别是来自参数为λ_1和λ_2的指数分布的iid样本，试求λ_2/λ_1的置信水平为$1-\alpha$ 的置信区间.

14. 设$X_1,\ldots,X_n$为来自正态总体$N(\mu,16)$的iid样本，为使μ的置信水平为$1-\alpha$的同等置信区间的长度不大于L，样本容量n至少应取多少？

15. 设$X_1,\ldots,X_n$为来自密度函数为

$$f(x,\theta)=\theta/x^2,\ x>\theta>0$$

的iid样本，试求θ的置信水平为$1-\alpha$的置信区间.

第4章　假设检验(I)

假设检验(Hypothesis Test)是由K. Pearson 于二十世纪初提出的，之后由Fisher进行了细化，并最终由Neyman和E. Pearson提出了较完整的假设检验理论. 另外，于上世纪五十年代，Wald发展了统计判决理论.

参数假设检验基本包括显著性检验(Significant test)和优势检验(Most Powerful test)两部分. 本章将先介绍显著性检验，优势检验将在下一章讲述. 另外，关于K. Pearson 早期提出的χ^2拟合优度检验(Goodness-of-fit test)也将在后面介绍.

§4.1　Fisher的显著性检验思想和基本概念

假设检验最早是由K. Pearson于1900年提出的，但Fisher及Neyman和E. Pearson对假设检验理论的发展和完善做出了杰出的贡献. 下面，我们先看一看Fisher的显著性检验思想.

§4.1.1　Fisher显著性检验思想

例 4.1.1 (女士品茶试验) 一种饮料由牛奶与茶按一定比例混合而成，可以先倒茶后牛奶(记为TM)，也可以反过来(记为MT). 某女士声称她可以鉴别是TM还是MT. 为此Fisher设计了如下试验，来检验她的说法是否可信. 现准备8杯饮料，TM与MT各一半，把它们随机地排在一起，让她品尝，并告诉她TM与MT各四杯，然后让她指出哪四杯是TM. 假设她全说对了，你相信她有这个能力吗？

解　对此问题，Fisher先引进一个假设：

$$H: \text{该女士并无鉴别能力.} \tag{4.1.1}$$

当H正确时，她只能随机地指出四杯是TM，而从8杯中选4杯共有$C_8^4=70$种选法，其中仅有一种判断是正确的. 于是，当H正确时，这个女士选对四杯的概率为1/70. 此时，我们必须承认，下述两个情况必发生其一：

1. H不成立，即该女士有一定的鉴别力；
2. 发生了一件概率只有$\frac{1}{70}$ 的事件.

显然，由于第二件事件发生比较稀奇，因而我们有相当的理由认为第一种情况发生了，或者说，该女士4杯全选对时这件事情是一个有利于假设H的显著性证据. 这样的一个推理过程就叫做显著性检验.

如果该女士只说对了三杯，则表面上看，4杯说对了3杯，成绩已经不错，但我们可以计算一下，纯粹出于碰巧而得到这个以至更好成绩的机会有多大，即当H成立时，从70种不同的挑法中，挑出3杯的概率为$17/70=0.243$. 显然发生一个概率为0.243的事件并不稀奇. 因此，试验结果没有提供不利于H的显著性证据. □

那我们凭什么说概率为1/70的事件的发生几乎不可能，而概率为17/70的可能性就大了呢？并且概率小到多小才认为是不可能呢？之所以提出这个问题，是由于此值的大小与最终

的判断结果有着直接的关系. 于是，为了得到一个大家都认可的判决，人们就必须事先指定一个阈值(threshold)α，比如可取$0.01, 0.05, 0.1$等等. 只有当计算所得的概率小于α时，才认为是小概率事件，即结果是显著的，也就是说提供了不利于H的显著性证据. 如在上面的品茶试验中，如取$\alpha = 0.01$，则即使4杯全说对，那我们也不能认为结果显著，这是因为$1/70 > 0.01$. 而如取$\alpha = 0.05$，则说对4杯就可以认为是显著了. 所以，Fisher认为，在做检验之前，一定要先给定好α，而不能在做完试验后再指定α. 否则就是公说公有理，婆说婆有理了. 在后面，我们称这样的α为显著性水平(significant level)，它越小，获得显著性结果越难，即越难拒绝H. Fisher称这样的检验为显著性检验.

虽然Fisher提出的显著性检验在实际得到了广泛的应用，但是对于同一个问题，可以有很多不同的检验方法. 于是，自然就有一个择优的问题，于是Neyman 和E. Pearson提出了优势检验，从而弥补了显著性检验的这一不足.

有关Fisher显著性检验的内容将在本章其它几节中加以介绍. 下面我们先看几个有关假设检验的基本概念.

§4.1.2 基本概念

假设检验在我们的实际生活中多有应用，尤其是在现在的商品经济社会中. 在讲述基本概念前，我们再看一个例子.

例 4.1.2 *某工厂在把产品卖给商店之前，与商店约定：如果此批产品的次品率小于某个给定值，如0.05，则商店接受这批产品，否则拒收. 同时，双方同意从此批产品中随机地抽取n件产品以估计其次品率p. 现以X记抽取n件产品中的次品数，则商店如何根据X的大小来决定是否接受这批产品.*

对于此问题，如我们假设$X \sim B(n,p)$，则上述问题就可以归结成如下的统计问题：设有来自总体$B(n,p)$的样本X，我们如何根据X的观测值对命题

$$H:\ p \le 0.05$$

作出“对”或“不对”的判断.

在统计中，我们把这种需要根据样本去推断其“正确”与否的命题，就称为一个假设或统计假设. 当然，假设可以是关于参数的，也可以是关于分布的.

通过样本对一个假设作出“对”或“不对”的具体判断规则就称为该假设的一个检验. 检验的结果若是肯定该命题，则称接受(accept)这个假设，否则就称为拒绝(reject)该假设. 我们注意到，这里的“接受”或“拒绝”一个假设的行为，只是反映了当事者在给定样本之下对该命题所采取的一种态度、一种自愿行为，而不是从逻辑上或理论上“证明”该命题正确与否. 这是由于我们所应用的样本是随机的，故我们所采取的决策可能是错误的. □

对于例4.1.1，我们可以把假设检验问题总结如下：

设有样本$\mathbf{X}$，取值于样本空间$\mathcal{X}$，且知道样本来自某一个参数分布族$\{F(x,\theta) : \theta \in \Theta\}$，其中$\Theta$为参数空间. 设$\Theta_0 \subset \Theta$，且$\Theta_0 \neq \emptyset$，则命题$H_0 : \theta \in \Theta_0$称为一个假设或零假设(null hypothesis). 如记$\Theta_1 = \Theta - \Theta_0$，则命题$H_1 : \theta \in \Theta_1$称为$H_0$的对立假设或备选假设(alternative hypothesis). 于是，我们感兴趣的假设就是

$$H_0 : \theta \in \Theta_0 \longleftrightarrow H_1 : \theta \in \Theta_1, \tag{4.1.2}$$

我们称此假设的检验问题为假设检验问题.

我们注意到上面的(4.1.2)与Fisher的显著性检验并无非常大的不同，只是这里多了一个备选假设. 然而，也正是有了这个备选假设，也才产生了Neyman和E. Pearson的优势检验. 另外，我们在实际应用中，对零假设的选取要特别加以注意. 比如在例4.1.1的女士品茶试验中，如我们不取(4.1.1)式的假设作为零假设，而取

$$H : \text{该女士有鉴别能力}$$

作为零假设，由于我们无法对“有鉴别能力”进行模型处理，故我们就无法进行上面的统计分析，也就得不到任何结论了.

实际上，对于例4.1.1，我们可以把其零假设写成如下形式：

$$H_0 : P\{X = i\} = \frac{\begin{pmatrix} 4 \\ i \end{pmatrix}\begin{pmatrix} 4 \\ 4-i \end{pmatrix}}{\begin{pmatrix} 8 \\ 4 \end{pmatrix}},$$

其中X表示该女士猜对的杯数，此时的零假设表示X服从超几何分布，它没有一个参数.

对于例4.1.2，参数空间$\Theta = (0,1), \Theta_0 = (0, 0.05]$，就与例4.1.1有所不同，由于假设完全可由参数表示出来.

对于假设(4.1.2)，如果Θ_0只有一个点，则我们称之为简单(simple)零假设，否则就称为复杂(composite)或复合原假设. 同样，对于备选假设也有简单与复合之别. 当H_0为简单假设时，其形式可写成$H_0 : \theta = \theta_0$. 此时的备选假设有两种可能：

$$H_1' : \theta \neq \theta_0, \quad H_1'' : \theta < \theta_0 \text{ 或 } \theta > \theta_0.$$

我们称$H_0 \longleftrightarrow H_1'$为双边的，$H_0 \longleftrightarrow H_1''$为单边的.

另外，对于假设(4.1.2)的检验就是指这样的一个法则或策略：当有了具体的样本后，由该法则或策略就可决定是接受H_0还是否定H_0，即检验就等价于把样本空间$\mathcal{X}$划分成两个互不相交的部分W和$\bar{W}$, 当样本属于$\bar{W}$时，接受H_0; 否则拒绝H_0. 于是，我们称W为该检验的拒绝域，而$\bar{W}$为接受域.

对于例4.1.2，其零假设与备选假设均是复合的，且此检验的拒绝域为$\{X > c\}$, 其中常数c由买卖双方协商决定，当然与前面讲的α有关.

显然，由于样本是随机的，故当我们应用某种检验方法采取决策时，我们可能作出正确的决策：即当$\theta \in \Theta_0$时，我们接受了H_0; 而当$\theta \in \Theta_1$时，我们拒绝了H_0. 但我们也可能犯下面的两种错误：当$\theta \in \Theta_0$时，而样本却落入了拒绝域W，于是我们采取了拒绝H_0的错误决策，我们称这样的错误为第一类错误(type I error)；当$\theta \in \Theta_1$时，而样本却落入了接受域$\bar{W}$，于是我们采取了接受H_0的错误决策，我们称这样的错误为第二类错误(type II error). 具体的可见下表：

决策	H_0为真	H_1为真
接受H_0	正确	Type II
拒绝H_0	Type I	正确

为了方便记忆，我们分别称第一、二类错误为拒真与纳伪.

由于上述两种错误决策受随机样本的影响，故它们也具有随机性，于是，我们定义犯第一、二类错误概率如下：

第一类错误概率：$\alpha = P_\theta\{\mathbf{X} \in W\}$, $\theta \in \Theta_0$，也记为$P\{\mathbf{X} \in W|H_0\}$,

第二类错误概率：$\beta = P_\theta\{\mathbf{X} \in \bar{W}\}$, $\theta \in \Theta_1$，也记为$P\{\mathbf{X} \in \bar{W}|H_1\}$.

既然一个检验都有如上两类错误，那我们能否找到一个检验，使其犯两类错误的概率都尽可能地小呢？实际上，我们做不到这一点. 为了说明其原因，我们先引进如下的势函数(power function)或功效函数的概念.

定义 4.1.1 (势函数) 对于假设(4.1.2)的一个检验方法ψ，其拒绝域记为W，则我们称

$$\beta_\psi(\theta) = P_\theta\{\mathbf{X} \in W\},\ \forall\ \theta \in \Theta, \tag{4.1.3}$$

为此检验的势函数.

从上一定义可以看出，当$\theta \in \Theta_0$时，此检验犯第一类错误的概率等于其势函数$\beta_\psi(\theta)$；而当$\theta \in \Theta_1$时，检验犯第二类错误的概率等于$1 - \beta_\psi(\theta)$.

下面我们通过一个例子说明我们无法使一个检验的第一、二类错误概率都尽可能地小.

例 4.1.3 如何判断厂家欺骗顾客: 如果买到标重是500克的袋糖，其实际重量为400克，我能告厂家欺诈吗？

解 如以μ表示厂家生产此种袋糖的平均重量，则我们要检验的假设为

$$H_0 : \mu = 500 \longleftrightarrow H_1 : \mu < 500.$$

顾客为了打官司，就要搜集证据，假设他又买了n袋糖，称其重量为$X_1, \ldots, X_n$，显然如果$\bar{X} = \sum_{i=1}^n X_i/n$ 远小于500，这场官司就有胜的把握，即当$\bar{X} - 500 < c$时，我们可以拒绝H_0，于是，可取拒绝域为$\{\mathbf{X} : \bar{X} - 500 < c\}$是合理的. 至于如何确定常数$c$，就要依赖于错误概率的选取.

如假设总体服从$N(\mu, \sigma_0^2)$, 且σ_0已知，则此时的检验统计量为$T(X) = \sum X_i/n$, 拒绝域为$W = \{\mathbf{x} : \bar{x} < 500 + c\}$，于是，由定义4.1.1知，这个检验的势函数为$\beta(\mu) = P_\mu\{\bar{X} < 500 + c\}$. 这样其第一、二类错误概率分别为

$$\alpha = P_{\mu=500}\{\bar{X} < 500 + c\} = P_{\mu=500}\left\{\frac{\bar{X} - 500}{\sigma_0/\sqrt{n}} < \frac{c}{\sigma_0/\sqrt{n}}\right\} = \Phi(\sqrt{n}c/\sigma_0), \tag{$*1$}$$

$$\beta = P_\mu\{\bar{X} > 500 + c\} = P_\mu\left\{\frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}} > \frac{500 + c - \mu}{\sigma_0/\sqrt{n}}\right\} = 1 - \Phi\left(\frac{500 + c - \mu}{\sigma_0/\sqrt{n}}\right). \tag{$*2$}$$

比较一下(*1)与(*2)可知，这两个错误概率均依赖于常数c，并且对于固定的样本容量，我们找不到一个c，使得二者均尽可能小. □

从上例我们可以看出：

- 对于固定的样本容量，我们找不到一个检验方法，使得其第一二类错误概率均达到最小；
- 第二类错误概率不易求出，由于它依赖于备选假设中的参数.

既然我们不可能同时控制一个检验的第一、二类错误概率，那我们通常的作法是仅限制第一类错误概率，即我们有如下的定义.

定义 4.1.2 (显著性水平) 对于检验ψ和事先给定的$\alpha \in (0,1)$，如果它满足

$$P_\theta\{\mathbf{X} \in W\} \leq \alpha, \ \forall \ \theta \in \Theta_0, \tag{4.1.4}$$

则称α是检验ψ的水平或显著性水平，也称ψ为显著性水平α的检验.

一般地，我们取$\alpha = 0.01, 0.05, 0.1$这三个值，其原因在于方便实际应用者查表. 在女士品茶试验中，当时我们所讲的阈值就是这里的显著性水平.

另外，当我们只控制一个检验的第一类错误概率时，就称这样的检验为显著性检验. 一般情况下，求取某假设的显著性检验的步骤如下：

❶ 根据实际问题，建立统计假设$H_0 \longleftrightarrow H_1$.

❷ 选取一个合适的统计量$T(\mathbf{X})$，使当H_0成立时，T的分布已知，且与参数θ无关.

❸ 根据H_0及H_1的特点，确定拒绝域W的形状.

❹ 对于给定的显著性水平α，确定拒绝域W.

❺ 由样本观测值$\mathbf{x}$，计算统计量$T(\mathbf{X})$的值$T(\mathbf{x})$，由$T(\mathbf{x})$是否属于W，作出最终判断.

本章后面的内容都将遵循如上步骤求取相应假设的显著性检验.

§4.2 单样本正态总体参数的显著性检验

正态总体是实际中最经常遇到的情况，故本节及下一节将仅考虑单个及两个正态总体参数的显著性检验问题. 在本节我们始终假设$X_1, \ldots, X_n$为来自正态总体$N(\mu, \sigma^2)$的iid样本，且我们感兴趣的是关于μ及σ^2的检验问题. 在讲述方法之前，我们先看一个例子.

例 4.2.1 某糖厂生产某种袋装糖，如果装袋机工作正常，则每袋的标准重量为100(克). 根据以往经验，当生产正常时，其各袋重量的标准差为$\sigma = 1.15$(克). 为了检测此台袋装机是否工作正常，现从中随机地抽取9袋糖，测其重量如下(单位：克)：

$$99.3, 98.7, 100.5, 101.2, 98.3, 99.7, 99.5, 102.1, 100.5.$$

请问此台包装机工作是否正常？

解 对于此问题，当生产正常时，由于每袋糖的重量与标准重量的偏差是由无数个误差造成的，故我们可以认为每袋糖的重量X_i服从正态分布$N(\mu, \sigma^2)$，且X_i间是独立的，其中μ是包

装机现有状态下的每袋糖的平均重量，而$\sigma = 1.15$. 如果$\mu = 100$，就可以认为此包装机工作正常，否则就不正常. 因此，此时的假设为

$$H_0 : \mu = 100 \longleftrightarrow H_1 : \mu \neq 100,$$

根据上节的定义，这是一个双边假设. □

为了对此假设问题进行检验，我们先考虑如下一般情况下的检验问题.

§4.2.1 单样本正态总体均值的检验

设$X_1, \ldots, X_n$为来自$N(\mu, \sigma^2)$的iid样本，现我们感兴趣的是关于其均值μ的如下假设：

$$H_0 : \mu = \mu_0 \longleftrightarrow H_1 : \mu \neq \mu_0, \tag{4.2.1}$$

$$H_0 : \mu = \mu_0 \longleftrightarrow H_1 : \mu < \mu_0, \tag{4.2.2}$$

$$H_0 : \mu = \mu_0 \longleftrightarrow H_1 : \mu > \mu_0, \tag{4.2.3}$$

$$H_0 : \mu \leq \mu_0 \longleftrightarrow H_1 : \mu > \mu_0, \tag{4.2.4}$$

$$H_0 : \mu \geq \mu_0 \longleftrightarrow H_1 : \mu < \mu_0, \tag{4.2.5}$$

其中μ_0是已知的常数. 上述假设有单边的、双边的、简单的及复合的，另外，σ^2的已知与否对上述假设检验问题是有影响的. 于是，我们就σ^2已知与未知两种情况来考虑上述五种假设的检验问题.

一、当$\sigma^2 = \sigma_0^2$已知时

❶ 对于假设(4.2.1)，我们感兴趣的参数在于总体均值μ的大小，而在第二章我们讲过，对于正态总体而言，样本均值$\bar{X}$是μ的一个很好的点估计. 于是，当假设(4.2.1)的H_0成立，即$\mu = \mu_0$时，$\bar{X}$应与μ_0相差不多. 而当H_1成立时，$\bar{X}$与μ_0应相差较大. 这样，我们可以用$|\bar{X} - \mu_0|$的大小来反映假设(4.2.1)，并且当$|\bar{X} - \mu_0| > c$时，我们有理由拒绝H_0, 即认为H_1成立.

为了标准化，因为$\mathrm{Var}(\bar{X} - \mu_0) = \sigma_0^2/n$，故我们可以取此假设的检验统计量为

$$U(\mathbf{X}) = \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma_0}. \tag{4.2.6}$$

根据上述分析，此检验的拒绝域为

$$W = \{\mathbf{x} : |U(\mathbf{x})| > c\}.$$

再根据显著性检验的特点，上述拒绝域中的常数c由其事先给定的显著性水平α确定，即我们要求此检验的第一类错误概率不大于α，也就是说，其常数c满足：

$$P_{H_0}\{|U| > c\} \leq \alpha.$$

又由于当H_0成立时，$U(\mathbf{X}) \sim N(0, 1)$，故上述拒绝域中的常数$c$可取为$c = u_{\alpha/2}$，由于此常数是拒绝与接受零假设的分水岭，故我们称之为检验的临界值(critical value).

总之，我们把此时关于假设(4.2.1)的显著性检验总结如下：

- 检验统计量：$U(\mathbf{X})=\frac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma_0}$,
- 检验的拒绝域：$\{\mathbf{x}:|U(\mathbf{x})|>u_{\alpha/2}\}$,

通过上面分析可知，此检验即为假设(4.2.1)的水平为α的显著性检验，由于此时检验统计量U的零分布是标准正态分布，且是双边的，故我们称之为双边的u检验.

对于例4.2.1，我们可以利用上述检验来处理之. 此时，$\mu_0=100,\sigma_0=1.15$，因为

$$|U|=\left|\frac{\sqrt{9}(99.98-100)}{1.15}\right|=|-0.052|<1.96=u_{0.05/2},$$

故在显著性水平0.05下，我们没有充分的理由拒绝H_0，也就是说，我们不能认为此台包装机工作不正常.

❷ **对于假设**(4.2.2)，由于它与假设(4.2.1)的区别在于备选假设的不同，故我们仍可采用关于假设(4.2.1)的检验统计量，但此时根据其备选假设的特点，我们取其拒绝域为

$$W=\{\mathbf{x}:U(\mathbf{x})<c\}.$$

又由于在H_0下，统计量$U(\mathbf{X})\sim N(0,1)$，故在控制其第一类错误概率为α下，其拒绝域中的常数c可取为$c=-u_\alpha$.

于是，关于假设(4.2.2)的水平为α的显著性检验为：

- 检验统计量：$U(\mathbf{X})=\frac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma_0}$,
- 检验的拒绝域：$\{\mathbf{x}:U(\mathbf{x})<-u_\alpha\}$,

我们称之为单边的u检验.

❸ **对于假设**(4.2.3)，用类似于假设(4.2.2)的方法，可以得到其水平为α的显著性检验为：

- 检验统计量：$U(\mathbf{X})=\frac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma_0}$,
- 检验的拒绝域：$\{\mathbf{x}:U(\mathbf{x})>u_\alpha\}$,

也称之为单边的u检验.

❹ **对于假设**(4.2.4)，由于此时我们感兴趣仍然是均值参数μ，故我们仍可以用上述的检验统计量(4.2.6)，虽然此时它的零分布并不是标准正态$N(0,1)$. 另外，根据备选假设的特点，此检验的拒绝域可取为

$$W=\{\mathbf{X}:U(\mathbf{X})>c\}.$$

为了控制此检验的第一类错误概率为给定的α，我们要求临界值c满足

$$\alpha(\mu)=P_{H_0}\{U(\mathbf{X})>c\}\leq\alpha.$$

事实上，$c = u_\alpha$即满足上式的要求. 实际上，因为当$H_0: \mu \le \mu_0$成立时，样本均值$\bar{X} \sim N(\mu, \sigma_0^2/n)$，故

$$\begin{aligned}
\alpha(\mu) &= P_{H_0}\{U(\mathbf{X}) > u_\alpha\} \\
&= P_{H_0}\left\{\frac{\sqrt{n}(\bar{X}-\mu+\mu-\mu_0)}{\sigma_0} > u_\alpha\right\} \\
&= P_{H_0}\left\{\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma_0} > u_\alpha - \frac{\sqrt{n}(\mu-\mu_0)}{\sigma_0}\right\} \\
&= 1-\Phi\left(u_\alpha - \frac{\sqrt{n}(\mu-\mu_0)}{\sigma_0}\right) \quad (\text{注意到此时}\mu \le \mu_0) \\
&\le 1-\Phi(u_\alpha) = \alpha.
\end{aligned}$$

综上所述，假设(4.2.4)的水平为α的显著性检验同假设(4.2.3)的水平为α的显著性检验.

❺ **对于假设**(4.2.5)，用类似于假设(4.2.4)方法，可以说明，它的水平为α的显著性检验同假设(4.2.2) 的水平为α的显著性检验.

例 4.2.2 假设某种型号三极管的放大倍数服从正态分布，其期望值不小于100(hfe). 为了检验其生产情况是否正常，现从中随机地抽取30只三极管，测得其放大倍数的平均值为95. 请问最近生产的三极管的放大倍数是不是降低了？(假设总体方差$\sigma^2 = 50$)

解 对于此问题，我们感兴趣的假设为(4.2.4)，且$\mu_0 = 100, \sigma_0^2 = 50, n = 30, \bar{X} = 95$. 对于$\alpha = 0.05$，我们有

$$U = \left|\frac{\sqrt{30}(95-100)}{\sqrt{50}}\right| = 3.9 > 1.645 = u_{0.05},$$

故我们有理由拒绝零假设，即认为最近生产的三极管的放大倍数降低了. □

二、当σ^2未知时

在上一段的内容中，我们始终假设正态总体的方差是已知的，但在许多实际问题，总体方差可能是未知的. 于是，本段将考虑当总体方差未知时，关于上述五个假设的显著性检验问题.

对于假设(4.2.1)，由于样本均值$\bar{X}$依然是总体均值的一个好的点估计，故知，当H_0成立时，$\bar{X}$与μ_0应比较接近，而当H_1成立时，二者应相差较大. 但是，此时样本均值$\bar{X}$的零分布$N(\mu_0, \sigma^2/n)$并非完全已知. 另外，我们注意到，样本方差S_n^2是总体方差σ^2的一个好的点估计，且当H_0成立时，有

$$T(\mathbf{X}) = \frac{\sqrt{n}(\bar{X}-\mu_0)}{S_n} \sim t(n-1), \tag{4.2.7}$$

于是，我们可采用它作为检验统计量.

这样，用类似于上一段的方法，我们可以得到此时假设(4.2.1)的水平为α的显著性检验如下：

- 检验统计量：$T(\mathbf{X}) = \frac{\sqrt{n}(\bar{X}-\mu_0)}{S_n}$,
- 检验的拒绝域：$\{\mathbf{x} : |T(\mathbf{x})| > t_{\alpha/2}(n-1)\}$,

由于此时检验统计量T的零分布是t分布，且是双边的，故我们称之为双边的t检验.

对于其余四个假设，我们均可用类似于上一段的方法得到其水平为α的显著性检验，只是这里所用的检验统计量为(4.2.7)式的$T(\mathbf{X})$，且相应的临界值为$t(n-1)$分布的相应分位数.

综上所述，我们把关于单样本正态总体均值的水平为α的显著性检验总结在表4.1中.

表4.1 单样本正态总体均值的显著性检验

方差	假设	检验统计量	拒绝域	名字
$\sigma^2 = \sigma_0^2$	$H_0 : \mu = \mu_0 \longleftrightarrow H_1 : \mu \neq \mu_0$	$U = \frac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma_0}$	$\{\|U\| > u_{\alpha/2}\}$	双边u检验
	$H_0 : \mu = \mu_0 \longleftrightarrow H_1 : \mu < \mu_0$		$\{U < -u_\alpha\}$	单边u检验
	$H_0 : \mu = \mu_0 \longleftrightarrow H_1 : \mu > \mu_0$		$\{U > u_\alpha\}$	单边u检验
	$H_0 : \mu \leq \mu_0 \longleftrightarrow H_1 : \mu > \mu_0$		$\{U > u_\alpha\}$	单边u检验
	$H_0 : \mu \geq \mu_0 \longleftrightarrow H_1 : \mu < \mu_0$		$\{U < -u_\alpha\}$	单边u检验
σ^2未知	$H_0 : \mu = \mu_0 \longleftrightarrow H_1 : \mu \neq \mu_0$	$T = \frac{\sqrt{n}(\bar{X}-\mu_0)}{S_n}$	$\{\|T\| > t_{\alpha/2}(n-1)\}$	双边t检验
	$H_0 : \mu = \mu_0 \longleftrightarrow H_1 : \mu < \mu_0$		$\{T < -t_\alpha(n-1)\}$	单边t检验
	$H_0 : \mu = \mu_0 \longleftrightarrow H_1 : \mu > \mu_0$		$\{T > t_\alpha(n-1)\}$	单边t检验
	$H_0 : \mu \leq \mu_0 \longleftrightarrow H_1 : \mu > \mu_0$		$\{T > t_\alpha(n-1)\}$	单边t检验
	$H_0 : \mu \geq \mu_0 \longleftrightarrow H_1 : \mu < \mu_0$		$\{T < -t_\alpha(n-1)\}$	单边t检验

对于例4.2.1，我们可以利用方差未知时的显著性检验重复做一次，看结果如何. 当然，在实际应用中，我们一定要尽量把已知信息用足.

§4.2.2 单样本正态总体方差的检验

在上一小节的内容中，我们有一部分内容是假设总体方差已知. 由于在许多实际问题，这个已知的方差是利用历史数据估计而来的，故我们有必要对其进行统计检验. 本小节就将介绍单样本正态总体方差的假设检验问题.

假设$X_1, \ldots, X_n$为来自正态总体$N(\mu, \sigma^2)$的iid样本，我们感兴趣的假设为

$$H_0 : \sigma^2 = \sigma_0^2 \longleftrightarrow H_1 : \sigma^2 \neq \sigma_0^2, \tag{4.2.8}$$

$$H_0 : \sigma^2 = \sigma_0^2 \longleftrightarrow H_1 : \sigma^2 < \sigma_0^2, \tag{4.2.9}$$

$$H_0 : \sigma^2 = \sigma_0^2 \longleftrightarrow H_1 : \sigma^2 > \sigma_0^2, \tag{4.2.10}$$

$$H_0 : \sigma^2 \leq \sigma_0^2 \longleftrightarrow H_1 : \sigma^2 > \sigma_0^2, \tag{4.2.11}$$

$$H_0 : \sigma^2 \geq \sigma_0^2 \longleftrightarrow H_1 : \sigma^2 < \sigma_0^2. \tag{4.2.12}$$

类似于均值的检验，我们也分均值已知与未知两种情况来考虑上述问题的显著性检验.

一、当$\mu=\mu_0$已知时

❶ **关于假设**(4.2.8)，由于此时σ^2的一个很好的点估计为$\sum_{i=1}^n(X_i-\mu_0)^2$，且当$H_0$成立时，$\sum_{i=1}^n(X_i-\mu_0)^2/\sigma^2\sim\chi^2(n)$，于是，我们可以选取检验统计量为

$$\chi^2=\frac{\sum_{i=1}^n(X_i-\mu_0)^2}{\sigma_0^2}, \tag{4.2.13}$$

另外，由此时零假设与备选假设的特点可知，其拒绝域为

$$W=\{\chi^2<c_1\}\cup\{\chi^2>c_2\},$$

其中$c_1<c_2$为两个待定的常数，且由检验的显著性水平α来确定，即它们满足

$$P_{H_0}\{\mathbf{X}\in W\}\le\alpha,$$

即

$$P_{H_0}\{\chi^2<c_1\}+P_{H_0}\{\chi^2>c_2\}\le\alpha. \tag{4.2.14}$$

虽然在H_0下，$\chi^2\sim\chi^2(n)$，但我们并不能由(4.2.14)式确定两个未知常数. 于是，一个常用的方法即是如下选取c_1,c_2:

$$P_{H_0}\{\chi^2<c_1\}\le\alpha/2,\ P_{H_0}\{\chi^2>c_2\}\le\alpha/2. \tag{4.2.15}$$

这样，由χ^2的零分布及(4.2.15)式，我们得到满足(4.2.14)式的临界值c_1,c_2如下：

$$c_1=\chi^2_{1-\alpha/2}(n),\quad c_2=\chi^2_{\alpha/2}(n).$$

于是，我们得到假设(4.2.8)的水平为α的一个显著性检验为

- 检验统计量$\chi^2=\frac{\sum_{i=1}^n(X_i-\mu_0)^2}{\sigma_0^2}$.
- 检验的拒绝域：$W=\{\chi^2<\chi^2_{1-\alpha/2}(n)\}\cup\{\chi^2>\chi^2_{\alpha/2}(n)\}$.

❷ **关于假设**(4.2.9)，我们可以利用类似于(4.2.8)的方法，求得其水平为α的显著性检验为

- 检验统计量$\chi^2=\frac{\sum_{i=1}^n(X_i-\mu_0)^2}{\sigma_0^2}$.
- 检验的拒绝域：$W=\{\chi^2<\chi^2_{1-\alpha}(n)\}$.

❸ **关于假设**(4.2.10)，我们也可以利用类似于(4.2.8)的方法，求得水平为α的显著性检验为

- 检验统计量$\chi^2=\frac{\sum_{i=1}^n(X_i-\mu_0)^2}{\sigma_0^2}$.
- 检验的拒绝域：$W=\{\chi^2>\chi^2_{\alpha}(n)\}$.

❹ **关于假设**(4.2.11)，虽然此时的零假设为复合的，但我们感兴趣的仍然是总体方差的大小，故我们仍采用(4.2.13)式的χ^2 统计量作为此假设问题的检验统计量，且根据此时备选假设的特点，其拒绝域为

$$W=\{\chi^2>c\},$$

其中常数c由控制第一类错误概率得到，即

$$P_{H_0}\{\chi^2>c\}\le\alpha.$$

但我们注意到，在H_0下，(4.2.13)式的χ^2 统计量的分布并不是$\chi^2(n)$. 然而，我们有

$$\begin{aligned}\alpha\ &\ge P_{H_0}\{\chi^2>c\}\\ &=P_{H_0}\left\{\tfrac{\sum_{i=1}^n(X_i-\mu_0)^2}{\sigma_0^2}>c\right\}\\ &=P_{H_0}\left\{\tfrac{\sum_{i=1}^n(X_i-\mu_0)^2}{\sigma^2}>\tfrac{c\sigma_0^2}{\sigma^2}\right\},\end{aligned}\tag{4.2.16}$$

由于在H_0成立时，$\sigma^2\le\sigma_0^2$，故只要我们取c满足

$$\alpha\ge P_{H_0}\left\{\frac{\sum_{i=1}^n(X_i-\mu_0)^2}{\sigma^2}>c\right\},\tag{4.2.17}$$

则因为

$$P_{H_0}\left\{\frac{\sum_{i=1}^n(X_i-\mu_0)^2}{\sigma^2}>c\right\}\ge P_{H_0}\left\{\frac{\sum_{i=1}^n(X_i-\mu_0)^2}{\sigma^2}>\frac{c\sigma_0^2}{\sigma^2}\right\},$$

故它一定满足(4.2.16)式.

由于在H_0成立时，(4.2.17)式中的$\frac{\sum_{i=1}^n(X_i-\mu_0)^2}{\sigma^2}\sim\chi^2(n)$，故我们可以取$c=\chi^2_\alpha(n)$. 于是，我们可知，关于假设(4.2.10)的水平为α的显著性也是假设(4.2.11)的水平为α的显著性检验.

❺ **关于假设**(4.2.12)，我们同样可以利用关于假设(4.2.9)的水平为α的显著性检验作此假设的显著性检验.

二、当μ未知时

此时，我们可以用类似于导出μ已知时的检验方法得到五个假设的水平为α的显著性检验. 只是注意到，此时所用的检验统计量为

$$\chi^2=\frac{\sum_{i=1}^n(X_i-\bar X)^2}{\sigma_0^2},\tag{4.2.18}$$

此时各假设的临界值与μ已知时非常类似，只需要把那里的自由度n换成现在的自由度$n-1$即可. 总之，我们把上述两种情况下关于单样本正态总体方差的显著性检验总结见表4.2.

例 4.2.3 某工厂生产某种零件，要求其长度的标准差不得超过0.4cm. 现在某日生产的一批零件中随机地抽取了30件，测量其长度如下:

零件长度X_i	5.1	5.3	5.6	6.2	6.5	6.8	7.0
频数n_i	1	3	7	9	6	3	1

表4.2 单样本正态总体方差的显著性检验

均值	假设	检验统计量	拒绝域
$\mu=\mu_0$	$\sigma^2=\sigma_0^2 \longleftrightarrow \sigma^2\neq\sigma_0^2$	$\chi^2=\frac{\sum_{i=1}^n(X_i-\mu_0)^2}{\sigma_0^2}$	$\{\chi^2<\chi^2_{1-\alpha/2}(n)\}\cup\{\chi^2>\chi^2_{\alpha/2}(n)\}$
	$\sigma^2=\sigma_0^2 \longleftrightarrow \sigma^2<\sigma_0^2$		$\{\chi^2<\chi^2_{1-\alpha}(n)\}$
	$\sigma^2=\sigma_0^2 \longleftrightarrow \sigma^2>\sigma_0^2$		$\{\chi^2>\chi^2_{\alpha}(n)\}$
	$\sigma^2\leq\sigma_0^2 \longleftrightarrow \sigma^2>\sigma_0^2$		$\{\chi^2>\chi^2_{\alpha}(n)\}$
	$\sigma^2\geq\sigma_0^2 \longleftrightarrow \sigma^2<\sigma_0^2$		$\{\chi^2<\chi^2_{1-\alpha}(n)\}$
μ未知	$\sigma^2=\sigma_0^2 \longleftrightarrow \sigma^2\neq\sigma_0^2$	$\chi^2=\frac{\sum_{i=1}^n(X_i-\bar{X})^2}{\sigma_0^2}$	$\{\chi^2<\chi^2_{1-\alpha/2}(n-1)\}\cup\{\chi^2>\chi^2_{\alpha/2}(n-1)\}$
	$\sigma^2=\sigma_0^2 \longleftrightarrow \sigma^2<\sigma_0^2$		$\{\chi^2<\chi^2_{1-\alpha}(n-1)\}$
	$\sigma^2=\sigma_0^2 \longleftrightarrow \sigma^2>\sigma_0^2$		$\{\chi^2>\chi^2_{\alpha}(n-1)\}$
	$\sigma^2\leq\sigma_0^2 \longleftrightarrow \sigma^2>\sigma_0^2$		$\{\chi^2>\chi^2_{\alpha}(n-1)\}$
	$\sigma^2\geq\sigma_0^2 \longleftrightarrow \sigma^2<\sigma_0^2$		$\{\chi^2<\chi^2_{1-\alpha}(n-1)\}$

试问这批零件的长度的标准差较以往有显著的增加吗？(取$\alpha=0.05$，且假设零件的长度服从正态分布$N(\mu,\sigma^2)$)

解 对于此问题，由题义知，需要检验假设为

$$H_0:\sigma^2\leq 0.4^2 \longleftrightarrow H_1:\sigma^2>0.4^2.$$

由于此时不知总体均值的大小，故由表4.2知，由于

$$\chi^2=\frac{\sum_{i=1}^n(X_i-\bar{X})^2}{\sigma_0^2}=\frac{7.987}{0.4^2}=49.918>42.557=\chi^2_{0.05}(29),$$

故我们有理由拒绝H_0，即认为在显著性水平0.05下，这批零件长度的标准差有显著的增大. □

注 4.3.1 从表4.1可以看出，当$\sigma^2=\sigma_0^2$已知时，关于假设

$$H_0:\mu=\mu_0 \longleftrightarrow H_1:\mu\neq\mu_0$$

的检验统计量为

$$U=\frac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma_0},$$

而其拒绝域为

$$\begin{aligned}
& W=\{|U|>u_{\alpha/2}\} \\
\Longleftrightarrow\quad & P_{H_0}\left\{\left|\frac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma_0}\right|>u_{\alpha/2}\right\}=\alpha \\
\Longleftrightarrow\quad & P_{H_0}\left\{\bar{X}-u_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}\leq\mu\leq\bar{X}+u_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}\right\}=1-\alpha.
\end{aligned}$$

而上式中的最后一个式子说明，μ的置信水平为$1-\alpha$的置信区间为

$$\left(\bar{X}-u_{\alpha/2}\frac{\sigma_0}{\sqrt{n}},\bar{X}+u_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}\right).$$

由此可知，置信水平为$1-\alpha$的置信区间与双边的u检验有着异曲同工之妙. 也就是说，我们可以这样理解置信区间：当μ_0落在μ的置信水平为$1-\alpha$的置信区间时，在显著性水平α下，我们没有理由拒绝$H_0:\mu=\mu_0$. 同理，对于单侧置信限，我们也可以用类似于单边的u检验去理解它.

注 4.3.2 从表4.1及表4.2可以看出，我们没有必要考虑五种假设的检验问题，由于其中有两组单边检验完全是相同的，于是，在下一节的两样本正态总体的检验中，我们仅考虑三个假设.

§4.3 两样本正态总体参数的显著性检验

两样本问题在实际中具有广泛的应用，下面我们先看一个例子.

例 4.3.1 *为研究正常成年男女血液中红细胞平均数之差别，现在某地随机地抽取正常成年男子156人，正常成年女子74人，测得男性红细胞平均数为465.13万/mm^3，标准差为54.8万/mm^3，女性红细胞平均数为422.16万/mm^3，样本标准差为49.2万/mm^3. 请问此地区正常成年人的红细胞数与性别有关吗？*$(\alpha=0.01)$

例4.3.1就是一个两样本问题，为了解答此问题，我们先考虑一般情形下的假设检验问题.

设$X_1,\ldots,X_m$为来自正态总体$N(\mu_1,\sigma_1^2)$的iid 样本，$Y_1,\ldots,Y_n$为来自正态总体$N(\mu_2,\sigma_2^2)$的iid 样本，且全样本是独立的.

§4.3.1 两样本正态总体均值的显著性检验

此时，我们感兴趣的假设为

$$H_0:\mu_1=\mu_2\longleftrightarrow H_1:\mu_1\neq\mu_2, \tag{4.3.1}$$

$$H_0:\mu_1\leq\mu_2\longleftrightarrow H_1:\mu_1>\mu_2, \tag{4.3.2}$$

$$H_0:\mu_1\geq\mu_2\longleftrightarrow H_1:\mu_1<\mu_2. \tag{4.3.3}$$

上面的三个假设，考虑的均是两样本正态总体均值. 另外，我们在区间估计一章，考虑过两样本正态总体均值的置信区间问题，由于它们都涉及到两个正态总体，故我们统称之为两样本问题. 如果，σ_1^2,σ_2^2未知，则称其均值差的估计与检验问题为Behrens-Fisher问题. 其原因是由于这是Behrens 于1929年提出的，且Fisher于上世纪30年代曾写过一系列文章讨论过此问题. 关于上述三个假设，我们仅考虑几个特殊情形.

一、当σ_1^2,σ_2^2均已知时

❶ **对于假设**(4.3.1)，由于$\bar{X},\bar{Y}$分别是μ_1,μ_2的一个很好的点估计，故我们有理由用两样本均值差$\bar{X}-\bar{Y}$来反映μ_1与μ_2的区别. 考虑到假设的特点，我们知道，当$|\bar{X}-\bar{Y}|\geq c$时，我们有理由拒绝H_0，其中常数c由此检验的第一类错误概率α 来决定.

由于当H_0成立时，统计量

$$U=\frac{\bar{X}-\bar{Y}}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}\sim N(0,1), \tag{4.3.4}$$

且其拒绝域形如

$$W=\{|U|\geq c\},$$

故为控制其第一类错误概率为α，于是常数c应满足

$$P_{H_0}\{|U|\geq c\}\leq \alpha,$$

这样，由(4.3.4)知，我们选取临界值$c=u_{\alpha/2}$.

综合上述分析，我们得到此时假设(4.3.1)的水平为α的显著性检验为：

- 检验统计量：$U=\frac{\bar{X}-\bar{Y}}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}$,
- 拒绝域：$\{|U|\geq u_{\alpha/2}\}$,

我们称之为两样本的双边u检验.

❷ **对于假设**(4.3.2)，我们可以利用上面的方法得到其水平α的显著性检验为

- 检验统计量：$U=\frac{\bar{X}-\bar{Y}}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}$,
- 拒绝域：$\{U\geq u_{\alpha}\}$,

我们称之为两样本的单边u检验.

我们可以证明，此检验的水平的确为α. 事实上，在H_0成立时，我们有

$$\begin{aligned}
& P_{H_0}\{U\geq u_\alpha\}\\
= \quad & P_{H_0}\left\{\frac{\bar{X}-\bar{Y}}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}\geq u_\alpha\right\}\\
= \quad & P_{H_0}\left\{\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}\geq u_\alpha-\frac{\mu_1-\mu_2}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}\right\}\\
\leq \quad & P_{H_0}\left\{\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}\geq u_\alpha\right\} \quad (H_0:\mu_1\leq\mu_2)\\
= \quad & \alpha. \qquad \text{(因为其中的随机变量为标准正态)}
\end{aligned}$$

❸ **对于假设**(4.3.3)，我们利用类似于关于假设(4.3.2)的方法得到其水平为α的显著性检验为

- 检验统计量：$U=\frac{\bar{X}-\bar{Y}}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}$,
- 拒绝域：$\{U\le -u_\alpha\}$,

我们也称之为两样本的单边u检验.

二、当$\sigma_1^2=\sigma_2^2=\sigma^2$未知时

❶ **对于假设**(4.3.1)，我们仍可以用$\bar{X}-\bar{Y}$来估计$\mu_1-\mu_2$，但是，我们注意到，当H_0成立时，$\bar{X}-\bar{Y}\sim N(0,(m+m)\sigma^2/mn)$，而其中$\sigma^2$未知. 由第二章内容可知，此时我们可以利用总的样本方差

$$S_{mn}^2=\frac{\sum_{i=1}^m(X_i-\bar{X})^2+\sum_{i=1}^n(Y_i-\bar{Y})^2}{m+n-2} \tag{4.3.5}$$

来估计σ^2，且在H_0成立时，$S_{mn}^2/\sigma^2\sim\chi^2(m+n-2)$. 于是，一个自然的检验统计量为

$$T=\frac{(\bar{X}-\bar{Y})/\sigma\sqrt{(m+n)/(mn)}}{\sqrt{(m+n-2)S_{mn}^2/(\sigma^2(m+n-2))}}=\sqrt{\frac{mn}{m+n}}\frac{\bar{X}-\bar{Y}}{S_{mn}}, \tag{4.3.6}$$

且当H_0成立时，$T\sim t(m+n-2)$.

另外，为满足其显著性水平为α，我们可取拒绝域为

$$W=\{|T|\ge t_{\alpha/2}(m+n-2)\}. \tag{4.3.7}$$

综上所述，此时双边假设(4.3.1)的水平α的显著性检验统计量由(4.3.6)式给出，而其拒绝域由(4.3.7)给出. 我们称此检验法为两样本的t检验.

❷ **对于假设**(4.3.2)，我们仍可以利用检验统计量(4.3.6)，且可以证明其拒绝域可取为

$$\{T\le -t_\alpha(m+n-2)\}.$$

另外，我们称此检验法为单边的两样本t检验.

❸ **对于假设**(4.3.3)，我们同样可以利用(4.3.6)给出的检验统计量，且其拒绝域为

$$\{T\ge t_\alpha(m+n-2)\}.$$

它也被称作单边的两样本t检验.

对于例4.3.1中的数据，如假设两总体的方差相等且未知，则我们可以利用本段方法进行显著性检验. 实际上，此时我们感兴趣的是要检验(4.3.1)的双边假设. 由已知条件知：$S_{1m}^2=54.8, S_{2n}^2=49.20$，于是，因为

$$T=\frac{\bar{X}-\bar{Y}}{\sqrt{(m-1)S_{1m}^2+(n-1)S_{2n}^2}}\sqrt{\frac{mn(m+n-2)}{m+n}}=5.7>2.6=t_{0.01/2}(228),$$

故我们有理由拒绝H_0，即认为该地区成年男女性间的红细胞有显著的差别.

三、一般情况

关于一般情况下两样本正态总体均值的显著性检验问题，我们在这里仅不加证明地给出其水平近似为α的显著性检验. 请有兴趣的读者参见陈希孺(1997)和茆诗松、王静龙(1990).

❶ 当m,n都充分大时

对于假设(4.3.1)，(4.3.2)，(4.3.3)，我们采用的统计量均为

$$U=\frac{\bar{X}-\bar{Y}}{\sqrt{\frac{S_{1m}^2}{m}+\frac{S_{2n}^2}{n}}}, \tag{4.3.8}$$

而它们的拒绝域分别为

$$\{|U|\geq u_{\alpha/2}\},\quad \{U>u_\alpha\},\quad \{U<-u_\alpha\}. \tag{4.3.9}$$

❷ 当m,n都不是很大时

对于上述三个假设，此时我们采用的统计量仍为(4.3.8)，只是由于样本比较小，故无法利用大样本近似. 正如我们在讲置信区间时一样，当$\mu_1=\mu_2$时，(4.3.8)式的统计量近似服从自由度为r的t分布，其中

$$r=\frac{S_{mn}^4}{\frac{S_{1m}^4}{m^2(m-1)}+\frac{S_{2n}^4}{n^2(n-1)}}, \tag{4.3.10}$$

它与(3.3.5)一样，其中$S_{mn}^2=\frac{S_{1m}^2}{m}+\frac{S_{2n}^2}{n}$.

另外，对于假设(4.3.1)，(4.3.2)，(4.3.3)的拒绝域，分别为

$$\{|U|\geq t_{\alpha/2}(r)\},\quad \{U>t_\alpha(r)\},\quad \{U<-t_\alpha(r)\}. \tag{4.3.11}$$

综上所述，我们把两样本正态总体均值的显著性检验总结如下：

表4.3 两样本正态总体均值的显著性检验

<table>
<tr><th colspan="2">方差</th><th>假设</th><th>检验统计量</th><th>拒绝域</th></tr>
<tr><td colspan="2" rowspan="3">σ_1^2,σ_2^2已知</td><td>$H_0:\mu_1=\mu_2\longleftrightarrow H_1:\mu_1\neq\mu_2$</td><td rowspan="3">$U=\frac{\bar{X}-\bar{Y}}{\sqrt{\sigma_1^2/m+\sigma_2^2/n}}$</td><td>$\{|U|\geq u_{\alpha/2}\}$</td></tr>
<tr><td>$H_0:\mu_1\leq\mu_2\longleftrightarrow H_1:\mu_1>\mu_2$</td><td>$\{U\leq -u_\alpha\}$</td></tr>
<tr><td>$H_0:\mu_1\geq\mu_2\longleftrightarrow H_1:\mu_1<\mu_2$</td><td>$\{U\geq u_\alpha\}$</td></tr>
<tr><td colspan="2" rowspan="3">$\sigma_1^2=\sigma_2^2$未知</td><td>$H_0:\mu_1=\mu_2\longleftrightarrow H_1:\mu_1\neq\mu_2$</td><td rowspan="3">$T=\sqrt{\frac{mm}{m+n}}\frac{\bar{X}-\bar{Y}}{S_{mn}}$</td><td>$\{|T|\geq t_{\alpha/2}(m+n-2)\}$</td></tr>
<tr><td>$H_0:\mu_1\leq\mu_2\longleftrightarrow H_1:\mu_1>\mu_2$</td><td>$\{T\leq -t_\alpha(m+n-2)\}$</td></tr>
<tr><td>$H_0:\mu_1\geq\mu_2\longleftrightarrow H_1:\mu_1<\mu_2$</td><td>$\{T\geq t_\alpha(m+n-2)\}$</td></tr>
<tr><td rowspan="6">σ_1^2 σ_2^2 未知</td><td rowspan="3">m,n 充分大时</td><td>$H_0:\mu_1=\mu_2\longleftrightarrow H_1:\mu_1\neq\mu_2$</td><td rowspan="3">$U=\frac{\bar{X}-\bar{Y}}{\sqrt{S_{1m}^2/m+S_{2n}^2/n}}$</td><td>$\{|U|\geq u_{\alpha/2}\}$</td></tr>
<tr><td>$H_0:\mu_1\leq\mu_2\longleftrightarrow H_1:\mu_1>\mu_2$</td><td>$\{U\leq -u_\alpha\}$</td></tr>
<tr><td>$H_0:\mu_1\geq\mu_2\longleftrightarrow H_1:\mu_1<\mu_2$</td><td>$\{U\geq u_\alpha\}$</td></tr>
<tr><td rowspan="3">m,n 不都 充分大</td><td>$H_0:\mu_1=\mu_2\longleftrightarrow H_1:\mu_1\neq\mu_2$</td><td rowspan="3">$T=\frac{\bar{X}-\bar{Y}}{\sqrt{S_{1m}^2/m+S_{2n}^2/n}}$</td><td>$\{|T|\geq t_{\alpha/2}(r)\}$</td></tr>
<tr><td>$H_0:\mu_1\leq\mu_2\longleftrightarrow H_1:\mu_1>\mu_2$</td><td>$\{T\leq -t_\alpha(r)\}$</td></tr>
<tr><td>$H_0:\mu_1\geq\mu_2\longleftrightarrow H_1:\mu_1<\mu_2$</td><td>$\{T\geq t_\alpha(r)\}$</td></tr>
</table>

例 4.3.2 为了比较两种矿石的含铁量，现分别从甲乙两矿的矿石中随机地抽取$m=10$及$n=5$个样本，测得其含铁量数据如下:

甲: $\bar{X}=16.01$, $S_{1m}^2=10.8$,

乙: $\bar{Y}=18.98$, $S_{2n}^2=0.27$.

试在水平$\alpha=0.01$下检验甲矿矿石的含铁量不低于乙矿矿石的含铁量.

解 对于此问题，我们有理由假设甲乙两矿的含铁量分别服从正态分布$X\sim N(\mu_1,\sigma_1^2)$, $Y\sim N(\mu_2,\sigma_2^2)$，则要检验的假设为

$$H_0:\mu_1\geq\mu_2\longleftrightarrow H_1:\mu_1<\mu_2.$$

而从样本方差可知，假设$\sigma_1^2=\sigma_2^2$是不合适的，且此时m,n均不大，于是，我们仅能利用上面一般情况下的第二种方法解决本问题.

由数据可计算得$T=-2.79$, 而t分布的自由度为$r=9.88$. 我们可以用自由度为10的t分布分位数计算之，也可以用如下的线性值：

$$t_{0.01}(9.88)=t_{0.01}(9)+(9.88-9)(t_{0.01}(10)-t_{0.01}(9))=2.77,$$

因为$|T|>2.77$，故我们有理由拒绝H_0，即认为在显著性水平$\alpha=0.01$下，甲矿石含铁量低于乙矿. □

§4.3.2 两样本正态总体方差的显著性检验

从上一节内容可以看出，两样本均值的检验方法与总体方差的大小有关. 本节则将考虑两样本方差的检验问题. 此时，我们感兴趣的假设为下面三个：

$$H_0:\sigma_1^2=\sigma_2^2\longleftrightarrow H_1:\sigma_1^2\neq\sigma_2^2, \tag{4.3.12}$$

$$H_0:\sigma_1^2\leq\sigma_2^2\longleftrightarrow H_1:\sigma_1^2>\sigma_2^2, \tag{4.3.13}$$

$$H_0:\sigma_1^2\geq\sigma_2^2\longleftrightarrow H_1:\sigma_1^2<\sigma_2^2. \tag{4.3.14}$$

下面，我们将分均值已知与未知两种情况考虑上述三种假设的显著性检验.

一、当μ_1,μ_2均已知时

❶ **对于假设**(4.3.12)，由于此时σ_1^2 与σ_2^2的合理估计分别为

$${S_{1m}^*}^2=\frac{1}{m}\sum_{i=1}^m(X_i-\mu_1)^2,\quad {S_{2n}^*}^2=\frac{1}{n}\sum_{i=1}^n(Y_i-\mu_2)^2,$$

故一个合理的检验统计量为

$$F=\frac{{S_{1m}^*}^2}{{S_{2n}^*}^2}=\frac{\sum_{i=1}^m(X_i-\mu_1)^2/m}{\sum_{i=1}^n(Y_i-\mu_2)^2/n}, \tag{4.3.15}$$

并且由备选假设的特点知道，此时的拒绝域为

$$\{F \le c_1\} \cup \{F \ge c_2\},\ (c_1 \le c_2)$$

为控制其第一类错误概率不大于α，故我们要求c_1, c_2满足

$$P_{H_0}\{F \le c_1\} + P_{H_0}\{F \ge c_2\} \le \alpha.$$

为了均衡上下两侧，我们通常要求c_1, c_2满足

$$P_{H_0}\{F \le c_1\} \le \frac{\alpha}{2},\quad P_{H_0}\{F \ge c_2\} \le \frac{\alpha}{2}. \tag{4.3.16}$$

因为在H_0成立时，(4.3.15)式的检验统计量$F \sim F(m, n)$，故我们可以取c_1, c_2如下：

$$c_1 = F_{1-\alpha/2}(m, n),\quad c_2 = F_{\alpha/2}(m, n). \tag{4.3.17}$$

于是，我们得到假设(4.3.12)的水平为α的显著性检验为

- 检验统计量：$F = \frac{\sum_{i=1}^m (X_i - \mu_1)^2/m}{\sum_{i=1}^n (Y_i - \mu_2)^2/n}$,
- 拒绝域：$\{F \le F_{1-\alpha/2}(m, n)\} \cup \{F \ge F_{\alpha/2}(m, n)\}$,

我们称之为双边的F检验.

❷ **对于假设**(4.3.13)，我们可以证明，其水平为α的显著性检验为

- 检验统计量：$F = \frac{\sum_{i=1}^m (X_i - \mu_1)^2/m}{\sum_{i=1}^n (Y_i - \mu_2)^2/n}$,
- 拒绝域：$\{F \ge F_\alpha(m, n)\}$,

我们称之为单边的F检验.

❸ **对于假设**(4.3.14)，其水平为α的显著性检验为

- 检验统计量：$F = \frac{\sum_{i=1}^m (X_i - \mu_1)^2/m}{\sum_{i=1}^n (Y_i - \mu_2)^2/n}$,
- 拒绝域：$\{F \le F_{1-\alpha}(m, n)\}$,

我们也称之为单边的F检验.

二、当μ_1, μ_2均未知时

对于上面的三个假设，我们采用的统计量为

$$F = \frac{S_{1m}^2}{S_{2n}^2} = \frac{\sum_{i=1}^m (X_i - \bar{X})^2/(m-1)}{\sum_{i=1}^n (Y_i - \bar{Y})^2/(n-1)}, \tag{4.3.18}$$

其拒绝域只需把上一段中的自由度换成$m-1, n-1$即可.

于是，我们把两样本正态总体方差的显著性总结如下：

表4.4 两样本正态总体方差的显著性检验

讨厌参数	假设	检验统计量	拒绝域
μ_1,μ_2已知	$\sigma_1^2=\sigma_2^2\longleftrightarrow\sigma_1^2\neq\sigma_2^2$	$F=\frac{\sum_{i=1}^m(X_i-\mu_1)^2/m}{\sum_{i=1}^n(Y_i-\mu_2)^2/n}$	$\{F\le F_{1-\alpha/2}(m,n)\}\cup\{F\ge F_{\alpha/2}(m,n)\}$
	$\sigma_1^2\le\sigma_2^2\longleftrightarrow\sigma_1^2>\sigma_2^2$		$\{F\ge F_\alpha(m,n)\}$
	$\sigma_1^2\ge\sigma_2^2\longleftrightarrow\sigma_1^2<\sigma_2^2$		$\{F\le F_{1-\alpha}(m,n)\}$
μ_1,μ_2未知	$\sigma_1^2=\sigma_2^2\longleftrightarrow\sigma_1^2\neq\sigma_2^2$	$F=\frac{\sum_{i=1}^m(X_i-\bar X)^2/(m-1)}{\sum_{i=1}^n(Y_i-\bar Y)^2/(n-1)}$	$\{F\le F_{1-\alpha/2}(m-1,n-1)\}\cup\{F\ge F_{\alpha/2}(m-1,n-1)\}$
	$\sigma_1^2\le\sigma_2^2\longleftrightarrow\sigma_1^2>\sigma_2^2$		$\{F\ge F_\alpha(m-1,n-1)\}$
	$\sigma_1^2\ge\sigma_2^2\longleftrightarrow\sigma_1^2<\sigma_2^2$		$\{F\le F_{1-\alpha}(m-1,n-1)\}$

例 4.3.3 有甲乙两台机床加工不同产品，从此两台机床加工的产品中随机地抽取若干件，测得其产品直径(单位：mm)如下：

甲：30.5，29.8，29.7，20.4，20.1，20.0，19.0，19.9

乙：16.7，17.8，17.5，16.8，16.4，17.6，16.2

假设甲乙两台机床加工产品的直径分别服从正态分布$N(\mu_1,\sigma_1^2),N(\mu_2,\sigma_2^2)$. 问：这两台机床加工的精度有无显著性差异？($\alpha=0.05$)

解 对于此问题，此时的假设为(4.3.12)式的双边假设，且均值未知. 因为

$$F=\frac{\sum_{i=1}^m(X_i-\bar X)^2/(m-1)}{\sum_{i=1}^n(Y_i-\bar Y)^2/(n-1)}=69.13>0.20=\frac{1}{5.12}=\frac{1}{F_{0.025}(6,7)}=F_{0.975}(7,6),$$

故我们有理由认为这两台机床加工的精度有显著的区别. □

例 4.3.4 两厂生产同一产品，其质量指标服从正态分布，且标准规格为均值等于120. 现从甲乙两厂中随机地抽出5件产品，测其指标值为：

甲：119，120，119.2，119.7，119.6

乙：110.5，106.3，122.2，113.8，117.2

请问：根据这些数据我们能说这两厂的产品是否符合规定的规格120？($\alpha=0.05$)

解 对于此问题，我们要检验的假设是

$$H_0:\mu=120\longleftrightarrow H_1:\mu\neq 120,$$

且方差未知. 由数据知

甲：$\bar X=119.5,S=0.4,|T|=|\sqrt{n}(\bar X-120)/S|=2.795>2.776=t_{0.025}(4),$

乙：$\bar X=114,S=6.105,|T|=|\sqrt{n}(\bar X-120)/S|=2.198>2.776=t_{0.025}(4),$

故在水平0.05下，我们有理由认为甲厂产品与规格不相符，而没有发现乙厂有不符合规格的有力证据. □

然而，从例4.3.4的数据知道，甲厂5件产品与120相差都不大，而乙厂5件产品都与规格120相

差较大，这好像与上面的检验结果正了相反，那我们应如何解释它呢？事实上，我们可以从如下几个方面看待此问题：

- 我们从样本标准差上看，甲厂的0.4远低于乙厂的6.105，这表明甲厂产品要比乙厂产品稳定得多.
- 由于甲厂产品的误差很小，故与标准规格120的微小差别都可以被检测出来. 虽然119.5比120 只小0.5，但已经是标准差的一倍多了，故这种差别已经不是随机误差，而有可能是系统误差. 当然，至于这样一个微小差别的实际重要性如何就要另当别论了. 所以说统计上的显著性差异，不一定有现实重要性.
- 从乙厂抽出的数据远低于120，这使我们有理由怀疑该厂产品的平均规格达不到规定的120. 但是，由于该厂产品质量波动太大，故所测得的数据尚不能有把握认为其平均规格确与120有差距. 要进一步分析，就需要有更多的数据.

§4.4 单参数指数型分布族的显著性检验

前两节我们主要讲述了关于正态总体参数的显著性检验问题，而没有涉及到其它分布参数的检验，尤其是例4.1.2的假设如何检验的问题. 本节我们将考虑单参数指数型分布族中的参数的显著性检验问题，其PDF函数为

$$f(x,\theta)=c(\theta)\exp\{\phi(\theta)T(x)\}h(x), \tag{4.4.1}$$

其中$c(\theta)>0, h(x)>0$.

§4.4.1 单参数指数型分布族的性质

定理 4.4.1 对于来自由(4.4.1)表示的单参数指数型分布族的iid样本$X_1,\dots,X_n$, 如果$\phi(\theta)$是θ 的严格增函数，且$\psi(T(\mathbf{X}))$是统计量$T(\mathbf{X})$的非降函数，则$E_\theta[\psi(T)]$ 也是θ 的非降函数.

证明 由于我们可以仅在其支撑集上讨论，故不妨高$f(x,\theta)>0$. $\forall\ \theta_1<\theta_2$，记

$$A=\{\mathbf{x}:f(\mathbf{x},\theta_1)<f(\mathbf{x},\theta_2)\}=\left\{\mathbf{x}:\frac{f(\mathbf{x},\theta_2)}{f(\mathbf{x},\theta_1)}>1\right\},$$

$$B=\{\mathbf{x}:f(\mathbf{x},\theta_1)>f(\mathbf{x},\theta_2)\}=\left\{\mathbf{x}:\frac{f(\mathbf{x},\theta_2)}{f(\mathbf{x},\theta_1)}<1\right\}.$$

由于

$$\frac{f(\mathbf{x},\theta_2)}{f(\mathbf{x},\theta_1)}=\frac{c(\theta_2)}{c(\theta_1)}\exp\{(\phi(\theta_2)-\phi(\theta_1))T(\mathbf{x})\}$$

及ϕ的严格单增性和$c(\theta)>0$，故$\frac{f(\mathbf{x},\theta_2)}{f(\mathbf{x},\theta_1)}$是$T$的严格增函数. 于是，$\forall\ \mathbf{x}_1\in A$, $\forall\ \mathbf{x}_2\in B$, 都有$T(\mathbf{x}_1)>T(\mathbf{x}_2)$.

又由于$\psi(T)$是T的非降函数，故我们有

$$a=\inf_{\mathbf{x}\in A}\psi(T(\mathbf{x}))\geq\sup_{\mathbf{x}\in B}\psi(T(\mathbf{x}))=b,$$

于是，

$$
\begin{aligned}
& \int \psi(T(\mathbf{x}))[f(\mathbf{x},\theta_2)-f(\mathbf{x},\theta_1)]\mathrm{d}\mathbf{x} \\
= \ & \int_A \psi(T(\mathbf{x}))[f(\mathbf{x},\theta_2)-f(\mathbf{x},\theta_1)]\mathrm{d}\mathbf{x} + \int_B \psi(T(\mathbf{x}))[f(\mathbf{x},\theta_2)-f(\mathbf{x},\theta_1)]\mathrm{d}\mathbf{x} \\
\geq \ & a\int_A [f(\mathbf{x},\theta_2)-f(\mathbf{x},\theta_1)]\mathrm{d}\mathbf{x} + b\int_B [f(\mathbf{x},\theta_2)-f(\mathbf{x},\theta_1)]\mathrm{d}\mathbf{x} \quad (\text{在}B\text{上，}f(\mathbf{x},\theta_2)<f(\mathbf{x},\theta_1)) \\
= \ & (a-b)\int_A [f(\mathbf{x},\theta_2)-f(\mathbf{x},\theta_1)]\mathrm{d}\mathbf{x} \geq 0, \quad (\text{注意}f(\mathbf{x},\theta_2)-f(\mathbf{x},\theta_1)\text{在}A,B\text{上积分反号})
\end{aligned}
$$

即

$$E_{\theta_2}\psi(T) \geq E_{\theta_1}\psi(T),$$

由于θ_1,θ_2的任意性，故知$E_\theta\psi(T)$的非降性. □

Corollary 4.4.1 设$X_1,\ldots,X_n$为来自单参数指数型分布族(4.4.1)的iid样本，其中$\phi(\theta)$是θ的严格增函数，则对任意给定的常数c，$P_\theta\{T(\mathbf{X})\geq c\}$和$P_\theta\{T(\mathbf{X})\leq c\}$分别是$\theta$的非降和非增函数.

§4.4.2 单参数指数型分布族的假设检验

设$X_1,\ldots,X_n$为来自单参数指数型分布族(4.4.1)的iid样本，其中$\phi(\theta)$是θ的严格增函数. 此时，我们感兴趣的假设问题为：

$$H_0:\theta\geq\theta_0 \longleftrightarrow H_1:\theta<\theta_0. \tag{4.4.2}$$

由定理4.4.1可知，如取$\psi(x)=x$，则可知$E_\theta T(\mathbf{X})$ 是θ的递增函数. 于是，我们可以取$T(\mathbf{X})$作为上述假设的一个检验统计量(事实上，它也是充分的)，且其拒绝域为$\{T(\mathbf{X})\leq c\}$，而其临界值c由第一类错误概率所决定. 由定理4.4.1的推论可知，$P_\theta\{T(\mathbf{X})\leq c\}$是$\theta$的非增函数，于是，此检验的第一类错误概率的最大值为$P_{\theta_0}\{T(\mathbf{X})\leq c\}$. 这样，我们要求其临界值$c$满足

$$P_{\theta_0}\{T(\mathbf{X})\leq c\}\leq\alpha. \tag{4.4.3}$$

同理，对于另一单边假设

$$H_0:\theta\leq\theta_0 \longleftrightarrow H_1:\theta>\theta_0, \tag{4.4.4}$$

其检验统计量仍为$T(\mathbf{X})$，且拒绝域为$\{T(\mathbf{X})\geq c\}$，其中c满足

$$P_{\theta_0}\{T(\mathbf{X})\geq c\}\leq\alpha. \tag{4.4.5}$$

关于双边假设

$$H_0:\theta=\theta_0 \longleftrightarrow H_1:\theta\neq\theta_0, \tag{4.4.6}$$

其检验统计量仍为$T(\mathbf{X})$，且其拒绝域为$\{T(\mathbf{X})\leq c_1\}\cup\{T(\mathbf{X}\geq c_2\}$，其中$c_1\leq c_2$满足

$$P_{\theta_0}\{T(\mathbf{X})\leq c_1\}\leq\alpha/2,\quad P_{\theta_0}\{T(\mathbf{X})\geq c_2\}\leq\alpha/2. \tag{4.4.7}$$

§4.4.3 Bernoulli分布的假设检验

本节以两点分布为例，细化上一小节的内容.

设$X_1,\ldots,X_n$为来自Bernoulli分布$b(1,p)$的iid样本，则样本分布为

$$\begin{aligned} f(\mathbf{x},p) &= P\{X_1=x_1,\ldots,X_n=x_n\} \\ &= p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i} = (1-p)^n \exp\left\{\ln\frac{p}{1-p}\sum_{i=1}^n x_i\right\}. \end{aligned}$$

显然，它是一个单参数指数型分布，且由于$\phi(p)=\ln\frac{p}{1-p}$是p的严格增函数，故它满足定理4.4.1的条件，于是，我们可用统计量$T(\mathbf{X})=\sum_{i=1}^n X_i$作为检验统计量. 另外，又由于$T(\mathbf{X})\sim B(n,p)$，于是，由定理4.4.4可知关于$p$的显著性检验为

表4.5 Bernoulli总体参数的显著性检验

假设	检验统计量	拒绝域
$H_0: p=p_0 \longleftrightarrow H_1: p\neq p_0$		$\{T\le c_1\}\cup\{T\ge c_2\}$
$H_0: p\le p_0 \longleftrightarrow H_1: p\ge p_0$	$T=\sum_{i=1}^n X_i$	$\{T\ge c_3\}$
$H_0: p\ge p_0 \longleftrightarrow H_1: p\le p_0$		$\{T\le c_4\}$

其中

$$c_1=\sup\left\{k:\sum_{i=0}^k \begin{pmatrix} n \\ i \end{pmatrix} p_0^i(1-p_0)^{n-i}\le\frac{\alpha}{2}\right\},$$

$$c_2=\inf\left\{k:\sum_{i=k}^n \begin{pmatrix} n \\ i \end{pmatrix} p_0^i(1-p_0)^{n-i}\le\frac{\alpha}{2}\right\},$$

$$c_3=\inf\left\{k:\sum_{i=k}^n \begin{pmatrix} n \\ i \end{pmatrix} p_0^i(1-p_0)^{n-i}\le\alpha\right\},$$

$$c_4=\sup\left\{k:\sum_{i=0}^k \begin{pmatrix} n \\ i \end{pmatrix} p_0^i(1-p_0)^{n-i}\le\alpha\right\}.$$

注 4.4.1 对于充分大的n，由于二项分布可由正态分布来近似，故上面的临界值可通过正态分布的大样本近似来得到.

例 4.4.1 (续例4.1.2) 如果双方同意抽取$n=20$件产品进行检验，试确定在水平$\alpha=0.01$下的检验临界值.

解 此时，我们感兴趣的假设为

$$H_0: p\le 0.05 \longleftrightarrow H_1: p>0.05.$$

检验统计量为样本中的次品总数X，且其拒绝域为$\{X \geq c\}$. 由表4.5可知，其临界值c应满足

$$I = \sum_{i=c}^{20} \left(\begin{array}{c} 20 \\ i \end{array} \right) 0.05^i 0.95^{20-i} \leq 0.01.$$

查二项分布表可知：当$c = 4$时，$I = 0.016$; 当$c = 5$时，$I = 0.0026$. 由于0.016与0.01最接近，于是，我们取$c = 4$. □

注 4.4.2 从例4.4.1可知看出，对于离散型分布的参数检验，有时我们无法得到水平严格为α的显著性检验. 此时，如双方坚持要求利用水平严格为α的显著性检验，则可以利用随机化的显著性检验方法(在后面将看到什么是随机化检验).

例 4.4.2 独立地掷一枚硬币495次，得到正面220次，反面275次. 试在$\alpha = 0.05$下检验这枚硬币是否均匀.

解 对于此问题，我们令

$$X_i = \begin{cases} 1, & \text{如第}i\text{ 次出现正面}, \\ 0, & \text{如第}i\text{ 次出现反面}, \end{cases}$$

则知$X_1, \ldots, X_n$为来自$b(1, p)$的iid样本.

另外，此时感兴趣的假设为

$$H_0 : p_0 = 0.5 \longleftrightarrow H_1 : p \neq 0.5.$$

由上面内容可知，此时的检验统计量为$T = \sum_{i=1}^{n} X_i$，且拒绝域为$\{X \leq c_1\} \cup \{X \geq c_2\}$，其中$c_1, c_2$分别是满足

$$\sum_{i=0}^{c_1} \left(\begin{array}{c} 495 \\ i \end{array} \right) \left(\frac{1}{2} \right)^{495} \leq 0.025,$$

$$\sum_{i=c_2}^{495} \left(\begin{array}{c} 495 \\ i \end{array} \right) \left(\frac{1}{2} \right)^{495} \leq 0.025$$

的最大c_1及最小的c_2. 由于此时n很大，故我们可用正态近似计算之.

事实上，上面两个方程可近似写成

$$\Phi \left(\frac{c_1 - 495 \times \frac{1}{2}}{\sqrt{495 \times \frac{1}{2} \times \frac{1}{2}}} \right) \leq 0.025,$$

$$1 - \Phi \left(\frac{c_2 - 495 \times \frac{1}{2}}{\sqrt{495 \times \frac{1}{2} \times \frac{1}{2}}} \right) \leq 0.025.$$

由此求得$c_1 \leq 225.7, c_2 \geq 269.3$, 于是，我们可取$c_1 = 225, c_2 = 270$. 故检验的拒绝域为

$$\{X \leq 225\} \cup \{X \geq 270\}.$$

对于本例，由于正面出现了220次，故大水平0.05下，我们有理由拒绝H_0，即认为此枚硬币不均匀. □

本小节我们仅考虑了二项分布，事实上，我们可以利用指数型分布的显著性检验方法得到关于Poisson分布、指数分布等其它常用分布的显著性检验.

§4.5 似然比检验

我们在前几节讲述的内容均是关于Fisher提出的显著性检验，类似于在估计中存在着多种估计一样，在假设检验中，也有多种有别于显著性检验的方法，如Neyman和E. Pearson于1928年提出的似然比检验，它在假设检验中的地位有如MLE在点估计中的地位. 另外，这也是一种应用非常广的检验方法，且由此方法得到的检验往往具有某种最优性.

我们再分析一下前面讲的显著性检验，它处理问题的思路是：先假设总体服从某参数分布族$f(x,\theta)$，之后根据实际问题提出假设$H_0 \longleftrightarrow H_1$，然后再利用样本及总体信息提出检验统计量及拒绝域，最后通过控制第一类错误概率求得具体的拒绝域. 从表4.1-4.5可以看出，我们考虑的假设均是特定的区间型的，而不是下面的非常一般的假设：

$$H_0: \theta \in \Theta_0 \longleftrightarrow H_1: \theta \in \Theta_1 = \Theta - \Theta_0. \tag{4.5.1}$$

对于有如(4.5.1)的假设，我们很难由显著性检验的步骤确定拒绝域. 但如从似然的角度看，我们则可以得到如下的似然比检验方法.

定义 4.5.1 (似然比统计量) 设$X_1, \ldots, X_n$为来自分布族$\mathcal{F} = \{f(x,\theta): \theta \in \Theta\}$的iid样本，对于感兴趣的假设(4.5.1)，令

$$\lambda(\mathbf{x}) = \frac{\sup_{\theta \in \Theta_0} f(\mathbf{x},\theta)}{\sup_{\theta \in \Theta} f(\mathbf{x},\theta)}. \tag{4.5.2}$$

则我们称统计量$\lambda(\mathbf{X})$为假设(4.5.1)的似然比(Likelihood Ratio)，有时也称之为广义似然比.

从$\lambda(\mathbf{X})$的定义不难看出，如果$\lambda(\mathbf{X})$的值很小，则说明$\theta \in \Theta_0$的可能性要比$\theta \in \Theta$的可能性小，于是，我们有理由认为H_0不成立. 这样，我们有如下的似然比检验.

定义 4.5.2 (似然比检验) 当采用(4.5.2)式的似然比统计量$\lambda(\mathbf{X})$作为假设(4.5.1)的检验统计量，且取其拒绝域为$\{\lambda(\mathbf{x}) \le c\}$时，其中临界值$c$满足

$$P_\theta\{\lambda(\mathbf{X}) \le c\} \le \alpha, \ \forall\, \theta \in \Theta_0, \tag{4.5.3}$$

则称此检验为水平α的似然比检验(Likelihood Ratio Test, 简记为LRT).

注 4.5.1 如果似然比统计量$\lambda(\mathbf{X})$的零分布未知，则我们很难由(4.5.3)式确定似然比检验的临界值. 但如果此时存在一个统计量$T(\mathbf{X})$关于$\lambda(\mathbf{X})$是单调的且它的零分布已知，则我们可以给出一个基于$T(\mathbf{X})$的显著性检验.

例 4.5.1 设$X_1, \ldots, X_n$是来自正态总体$N(\mu,\sigma^2)$的iid样本，μ, σ^2均未知. 试求假设

$$H_0: \mu = \mu_0 \longleftrightarrow H_1: \mu \ne \mu_0$$

的水平为α的检验.

解 此时样本分布为

$$f(\mathbf{x},\theta)=(2\pi\sigma^2)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\right\},$$

其参数空间为

$$\Theta_0=\{(\mu_0,\sigma^2):\sigma^2>0\},\quad \Theta=\{(\mu,\sigma^2):\mu\in R,\ \sigma^2>0\}.$$

利用微分法，我们容易求得

$$\sup_{\theta\in\Theta_0}f(\mathbf{x},\theta)=\left[2\pi\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_0)^2\right]^{-n/2}e^{-n/2},$$

$$\sup_{\theta\in\Theta}f(\mathbf{x},\theta)=\left[2\pi\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{-n/2}e^{-n/2},$$

于是，其似然比统计量为

$$\begin{aligned}\lambda(\mathbf{x}) &= \frac{\sup_{\theta\in\Theta_0}f(\mathbf{x},\theta)}{\sup_{\theta\in\Theta}f(\mathbf{x},\theta)}=\left(\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\mu_0)^2}\right)^{n/2}\\ &= \left[\frac{1}{\frac{\sum_{i=1}^{n}(x_i-\bar{x}+\bar{x}-\mu_0)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}}\right]^{n/2}=\left[\frac{1}{1+n\frac{(\bar{x}-\mu_0)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}}\right]^{n/2}\\ &= \left[1+\frac{T^2}{n-1}\right]^{-n/2},\end{aligned}$$

其中$T=\frac{\sqrt{n}(\bar{x}-\mu_0)}{S_n}$.

另外，从上式可知，此时的似然比统计量与传统的t统计量的平方成反比，于是，两个检验统计量的拒绝域有如下关系：

$$\{\lambda(\mathbf{x})\le c\}\Longleftrightarrow\{|T(\mathbf{x})|\ge d\}.$$

又因为当H_0成立时，$T\sim t(n-1)$，故我们取$d=t_{\alpha/2}(n-1)$即可控制其第一类错误概率不超过α. 由此可见，此时的似然比检验与我们前面讲过的双边t检验完全等价. □

例 4.5.2 设$X_1,\ldots,X_n$为来自具有如下PDF

$$f(x,\mu)=e^{-(x-\mu)},\ x\ge\mu,\ \mu\in R$$

的iid样本，试求假设

$$H_0:\mu=0\longleftrightarrow H_1:\mu\ne 0$$

的水平α的似然比检验.

解 此时的样本分布为

$$f(\mathbf{x},\mu)=e^{-\sum_{i=1}^{n}(x_i-\mu)}I_{(x_{(1)}\geq\mu)},$$

且参数空间为$\Theta=R,\Theta_0=\{0\}$.

于是，可以求得

$$\sup_{\theta\in\Theta_0}f(\mathbf{x},\mu)=e^{-\sum_{i=1}^{n}x_i},\quad \sup_{\theta\in\Theta}f(\mathbf{x},\mu)=e^{-\sum_{i=1}^{n}x_i+nx_{(1)}}.$$

这样，此时的似然比统计量为

$$\lambda(\mathbf{X})=e^{-nX_{(1)}}=e^{-\frac{1}{2}(2nX_{(1)})},$$

它等价于统计量$2nX_{(1)}$, 且其拒绝域为$\{2nX_{(1)}\geq c\}$.

又因为在H_0成立时，$2nX_{(1)}\sim\chi^2(2)$ (计算其PDF即知)，故其临界值$c=\chi^2_\alpha(2)$. 即此时的似然比检验统计量为$2nX_{(1)}$，其拒绝域为$\{2nX_{(1)}\geq\chi^2_\alpha(2)\}$. □

例 4.5.3 设总体分布为

$$P\{X=i\}=p_i,\ i=1,\cdots,r,$$

其中$p_i>0,\sum_{i=1}^{r}p_i=1$. 设有$n$个来自此分布的iid样本$X_1,\ldots,X_n$, 试求

$$H_0:p_i=p_{i0},\ i=1,\cdots,r \tag{4.5.4}$$

的似然比检验，其中p_{i0}已知，且$\sum_{i=1}^{r}p_{i0}=1$.

解 对于此问题，我们注意到，其总体是一个离散分布，且参数空间Θ是$r-1$维，而Θ_0只包含一个点.

另外，为方便，我们以n_i表示样本中取值i的个数，即$\sum_{i=1}^{r}n_i=n$. 此时样本分布为

$$P_\theta\{\mathbf{X}=\mathbf{x}\}=p_1^{n_1}\cdots p_r^{n_r},$$

由此可求得

$$\sup_{\theta\in\Theta}P_\theta\{\mathbf{X}=\mathbf{x}\}=\prod_{i=1}^{r}\left(\frac{n_i}{n}\right)^{n_1},\quad \sup_{\theta\in\Theta_0}P_\theta\{\mathbf{X}=\mathbf{x}\}=\prod_{i=1}^{r}p_{i0}^{n_i}.$$

于是，其似然比统计量为

$$\lambda(\mathbf{X})=\prod_{i=1}^{r}\left(\frac{np_{i0}}{n_i}\right)^{n_i}.$$

另外，由于

$$\begin{aligned}\ln\lambda(\mathbf{X}) &= \sum_{i=1}^{r} n_i \ln\frac{np_{i0}}{n_i}\\ &= -\sum_{i=1}^{r}[np_{i0}+(n_i-np_{i0})]\ln\left(1+\frac{n_i-np_{i0}}{np_{i0}}\right)\\ &\approx -\sum_{i=1}^{r}[np_{i0}+(n_i-np_{i0})]\left[\frac{n_i-np_{i0}}{np_{i0}}-\frac{1}{2}\left(\frac{n_i-np_{i0}}{np_{i0}}\right)^2+o_p(n^{-2})\right]\\ &\approx -\frac{1}{2}\sum_{i=1}^{r}\frac{(n_i-np_{i0})^2}{np_{i0}}+o_p(n^{-1}),\end{aligned}$$

所以，

$$-2\ln\lambda(\mathbf{X})\approx\sum_{i=1}^{r}\frac{(n_i-np_{i0})^2}{np_{i0}}. \tag{4.5.5}$$

又因为$-2\ln\lambda(\mathbf{X})$是$\lambda(\mathbf{X})$的单减函数，故似然比检验等价于用检验统计量$-2\ln\lambda(\mathbf{X})$，且拒绝域为$\{-2\ln\lambda(\mathbf{X})\geq c\}$.

总之，对于此问题，我们可采用检验统计量$\sum_{i=1}^{r}\frac{(n_i-np_{i0})^2}{np_{i0}}$, 且其拒绝域为

$$\left\{(n_1,\ldots,n_r):\sum_{i=1}^{r}\frac{(n_i-np_{i0})^2}{np_{i0}}\geq c\right\}. \tag{4.5.6}$$

□

我们在这里看不出这个统计量有什么特别之处，但是我们在后面的第6章可以看出，这是K. Pearson引进的χ^2拟合优度检验统计量，它在许多实际问题中都得到了广泛的应用.

§4.6 统计与欺骗

虽然统计的应用已越来越广泛，也越来越受到重视，但是，我们在应用统计时一定要注意数据的真实性. 当然，即使数据是真实的，但不同的统计方法也可能得到不同的结论，故在应用统计方法解决实际问题时，一定要全面考虑问题的实际背景，最好是结合专业知识来处理相应的问题.

另外，我们在看一个统计结果时，也要注意其采用的统计方法是否合适. 由于对利用不同方法得到结论的解释可能是不同的. 当然，我们也可以利用统计的这一特点有针对性地解决某些实际问题. 有兴趣的读者请参见Huff (1982).

但是，我们也要注意到，有些骗子有利用大家不懂统计，而利用统计行骗. 下面看一个何书元(2006)给出的例子. (也请见*读者*，2005.22)

一天，乔治在删除垃圾邮件时发现了一个标题：惊人的足球杯预测. 他好奇地打开了它：亲爱的球迷，我们的统计学家已经设计出了准确预测足球比赛的方法，今晚英国足球杯第三场比赛是考文垂队对谢菲尔队，我们以0.95的概率预测考文垂队获胜.

乔治看后一笑，晚上看比赛时，考文垂队果然获胜. 三周后，乔治又收到了那人的邮件：亲爱的球迷，上次我们成功地预测了考文垂队获胜. 今天考文垂队和米德尔斯堡队相遇了，我们以0.95的概率预测米德尔斯堡队获胜. 请你密切关注比赛结果.

考文垂队强于对手，那天晚上却发挥不好，双方打成了1平. 但在加时赛上米德尔斯堡队奇迹般地获胜了. 乔治心中一震. 一周后，那人的电子邮件预测米德尔斯堡队将败给特伦米尔队. 结果果然如此.

接下来的四分之一决赛前，那人的邮件预测特伦米尔队胜陶顿亨队. 结果也是如此，四次预测都成功了. 乔治大吃一惊.

乔治再次收到如下的电子邮件：亲爱的球迷，现在你大概知道了我们的预测的确能够预测比赛的结果. 实际上，我们买断了一位统计学家的研究专利，能够以0.95的概率预测足球比赛的正确结果. 今晚的半决赛中，我们以0.95的概率预测阿森纳队打败伊普斯维队.

乔治是个不信邪的人，他约了几个朋友晚上一起看电视，准备在伊普斯维队获胜后好好地羞辱一下那个家伙. 但是，阿森纳队在比分落后的情况下奋起直追，竟以2：1获胜. 太不可思议了.

第二天，电子邮件又来了：亲爱的球迷，我们已经五次预测成功. 现在希望和你做笔交易，你支付200英镑，把一个月内所关心的比赛和球队告诉我们，我们将以0.95的概率为你预测胜负. 殷切地期望你的合作.

200英镑不是小数目，但是，如果能预测结果，就可以从彩票商手里赚回20万. 乔治心中盘算：如果发邮件的人只是猜测胜负，则5次都猜对的概率仅为$2^{-5}=0.031$, 于是，以0.9687的概率否定他是在猜测. 当然，乔治也怀疑过他们是否和黑社会或某个非法财团有关. 但是，这都和乔治无关，只要赚钱就行了. 于是，乔治支付了200英镑.

实际上，这些骗子先发出8000封电子邮件，一半预测甲胜，一半预测乙胜. 于是有4000人得到正确的预测，另外4000人付之一笑. 第二次只给上次得到成功预测的4000人发电子邮件，以此类推，五次预测以后得到$8000/2^5=250$人. 如果这250人中有100人付钱，就可以骗得20000英镑. 乔治就是这100人中的一个.

有位名人说过：谎话有三种：一般的谎话、糟糕透顶的谎话和统计.

另有一位名人也说过：使我们陷入困境的，倒不是我们不知道的东西，而是我们所知道的不正确的东西.

习题四

1. 在正常情况下，如乘座夏利出租车从南开大学东门到天津火车站需要13元左右. 为检验校内出租车运营是否规范，现随机地乘座其中15辆出租车从东门到火车站（假设路况正常），其平均花费为15.4，标准差为2.3，请在水平0.05下，检验校内出租车运营是否规范.

2. 为研究矽肺病患者肺功能的变化情况，某医院对I，II期矽肺病患者各33名测定其肺活量，得到二者的平均值分别为2710ml和2830ml，其标准差分别为147ml和118ml. 请问：大水平0.05下，第I，II期矽肺病患者的肺活量有无显著差异？

3. 设$X_1,\ldots,X_n$为来自两点分布$b(1,p)$的iid样本，试求假设$H_0:p=p_0$的似然比检验.

4. 设$X_1,\ldots,X_n$为来自正态总体$N(\mu,\sigma^2)$的iid样本，试求假设$H_0:\sigma^2=\sigma_0^2$的似然比检验.

5. 设$X_1,\ldots,X_n$为来自均匀分布$U(0,\theta)$的iid样本，其中参数$\theta>0$. 如对于假设

$$H_0:\theta\geq 2\longleftrightarrow H_1:\theta<2,$$

取检验统计量为最大次序统计量$X_{(n)}$，且拒绝域为$W=\{\mathbf{x}:x_{(n)}\leq 1.5\}$，则求此检验犯第一类错误概率的最大值.

6. 设$X_1,\ldots,X_n$为来自两点分布$b(1,p)$的iid样本.

(1) 试求假设$H_0:p\leq 0.01\longleftrightarrow H_1:p>0.1$的水平为0.05的显著性检验；

(2) 如要求这个检验在$p=0.08$时的第二类错误概率不超过0.1，样本容量n应为多少？

7. 为了检测甲乙两种小麦品种的好坏，现在十块地上同时试种这两个品种. 收获之后测得二者的平均产量分别为670kg和800kg，样本标准差分别为120kg和90kg. 试在水平0.01下，检验这两个品种的小麦有无显著性差异.

8. 下列哪些假设是简单的，哪些是复合的？

(1) X服从正态分布；

(2) X服从正态分布$N(2,0.5^2)$;

(3) $EX=2,\mathrm{Var}X=0.5^2$;

(4) X不服从正态分布；

(5) X服从参数为0.5的两点分布；

(6) X服从参数为0.5的二项分布；

(7) X服从指数分布$E(3)$.

9. 设X是来自均匀分布$U(\theta-0.5,\theta+0.5)$的一个样本，对于假设

$$H_0:\theta\leq 3\longleftrightarrow H_1:\theta\geq 4,$$

构造一个检验法，使其势函数$\beta(\theta)$满足：

$$\beta(\theta)=0,\ 当\theta\leq 3;\ \ \beta(\theta)=0,当\theta\leq 4.$$

10. 设$X_1,\ldots,X_n$为来自指数分布$E(\lambda_1)$的iid样本，$Y_1,\ldots,Y_m$为来自指数分布$E(\lambda_2)$的iid样本，且两组样本独立，其中λ_1,λ_2是的未知的正参数.

(1) 求假设$H_0:\lambda_1=\lambda_2\longleftrightarrow H_1:\lambda_1\neq\lambda_2$的似然比检验；

(2) 证明上述检验法的拒绝域仅依赖于比值$\sum_{i=1}^n X_i/\sum_{i=1}^m Y_i$;

(3) 求统计量$\sum_{i=1}^n X_i/\sum_{i=1}^m Y_i$的零分布.

11. 设$X_1,\ldots,X_m$和$Y_1,\ldots,Y_n$分别来自正态总体$N(\mu_1,\sigma^2)$和$N(\mu_2,\sigma^2)$的iid样本，且全样本独立，其中μ_1,μ_2,σ^2均未知. 请仿照两样本t检验，构造假设$H_0:\mu_1=c\mu_2$的水平α的显著性检验，其中$c\neq 0$为常数.

12. 设$X_1,\ldots,X_n$为来自均匀分布$U(0,\theta)$的iid样本，其中$\theta>0$为参数. 对于给定的$\theta_0>0$，请给出假设$H_0:\theta\leq\theta_0$的水平α的显著性检验.

13. 设$X_1,\ldots,X_n$为来自正态总体$N(\mu,\sigma^2)$的iid样本，其中μ,σ^2为未知参数. 试证明关于假设

$$H_0:\sigma^2=\sigma_0^2\longleftrightarrow H_1:\sigma^2\neq\sigma_0^2$$

的似然比检验就是检验统计量为$\chi^2 = (n-1)S_n^2/\sigma_0^2$的$\chi^2$检验，其中$S_n^2$为样本方差.

14. 在对一种新的流感疫苗进行人体实验时，为实验组的900位志愿者注射了疫苗，在6个月内他们中有9个人得了流感. 为对照组的900位志愿者注射了老疫苗，在6个月内他们中有19个人得到流感. 请问新疫苗是否更有效？

15. 在村A随机调查了90位男村民，其中有45人对现任村委会主任表示满意，又随机调查了100位女村民，有69人对现任村委会主任表示满意. 在显著性水平0.05下，

(1) 能否认为男女村民的态度有明显的差异？

(2) 求村中满意村委会主任的男女村民的比例p_1和p_2的置信水平为95%的置信区间;

(3) 由(2)的区间估计能否认为$p_1 > p_2$？

16. 以46^0仰角发射了9颗库存了1个月的某型号炮弹，射程分别是(单位：km):

$$30.89,\ 31.74,\ 33.82,\ 32.79,\ 31.87,\ 31.79,\ 31.7,\ 32.23,\ 31.85.$$

又以相同的仰角发射了8颗库存了两年的同型号炮弹，其射程分别是(单位：km):

$$32.84,\ 31.46,\ 32.31,\ 31.75,\ 30.15,\ 31.51,\ 31.43,\ 31.74.$$

如果射程都服从正态分布，则在显著性水平0.05下，

(1) 能否认为这两批炮弹射程的标准差有显著性差异？

(2) 在(1)的基础上，能否认为这两批炮弹的平均射程有显著性差异？

第5章　假设检验(II)

第4章的内容主要涉及到了Fisher显著性检验，而显著性检验的最主要特点就是仅考虑如何控制第一类错误概率，而没有思考如何控制检验的第二类错误概率，或在控制第一类错误的前提下，寻找第二类错误概率最小的检验. 本章则将考虑在第一类错误概率为α的检验类中如何求取一个第二类错误概率最小的检验，这方面的内容是由 Neyman和E. Pearson在上世纪二十年代提出的.

§5.1　引言

在上一章中，我们给出了有关假设检验的几个定义，如假设、检验、拒绝域、临界值等等. 为了本章叙述的方便，我们再把前面的几个概念简述如下.

针对假设$H_0 \longleftrightarrow H_1$, 一个检验函数或检验法则或检验，就是设法把样本空间$\mathcal{X}$划分为两个互不相交的可测集：

$$\mathcal{X} = W + \bar{W}.$$

当观测值$\mathbf{x} \in W$时，拒绝原假设H_0；当$\mathbf{x} \in \bar{W}$时，就不能拒绝原假设H_0. 其中的W称为拒绝域. 即此时的检验或检验函数被定义成如下的函数.

对于拒绝域$W \subset \mathcal{X}$, 定义

$$\phi(\mathbf{X}) = \begin{cases} 1, & \mathbf{X} \in W, \\ 0, & \mathbf{X} \in \bar{W}, \end{cases}$$

它是W的函数，且仅取0，1两个值. 反之，如果对于一个仅取0，1两个值的函数$\phi(\mathbf{X})$, 则$W = \{\mathbf{X} : \phi(\mathbf{X}) = 1\}$也可作为拒绝域. 这样的函数$\phi(x)$就称为检验函数或检验，其势函数为

$$\beta_\phi(\theta) = P_\theta\{\mathbf{X} \in W\} = E_\theta \phi(\mathbf{X}). \tag{5.1.1}$$

上述定义是我们在上一章中所采用的，但是对于本章要讲的优势检验而言，我们有必要重新定义检验如下：

定义 5.1.1 (检验)　设$\phi(\mathbf{X})$是定义在$\mathcal{X}$上的可测函数，满足$0 \le \phi(\mathbf{x}) \le 1$，则称$\phi(\mathbf{X})$为检验函数，简称检验. 如果$\phi(\mathbf{x})$仅取0，1两值时，则称为非随机化检验，否则，就称为随机化检验，其势函数为$\beta_\phi(\theta) = E_\theta \phi(\mathbf{X})$.

从上述定义可以看出，上一章我们称拒绝域的示性函数为一个检验，而本章则不然：由定义5.1.1所定义的检验可以在$[0, 1]$间取值，它可能不是拒绝域的示性函数.

另外，当检验统计量为连续随机变量时，我们采用的是非随机化检验. 当检验统计量为离散随机变量时，就有可能采用随机化检验. 这一点，我们将在后面接触到.

既然我们用两类错误概率来衡量一个检验的好坏，则我们可以利用其势函数把检验分类.

定义 5.1.2　设$\phi_1(x)$和$\phi_2(x)$是某检验问题$H_0 \longleftrightarrow H_1$的检验函数，如果它们的势函数相同，即

$$E_\theta \phi_1(X) = E_\theta \phi_2(X),\ \forall\ \theta \in \Theta,$$

则称检验函数$\phi_1(x)$和$\phi_2(x)$等价.

前面的知识告诉我们，充分统计量在估计中非常有用，由于它完全包含了样本中关于参数的信息，那它在检验中又如何呢？见下面的定理

定理 5.1.1 设$X_1,\ldots,X_n$是来自分布族$\{P_\theta:\theta\in\Theta\}$的样本. $T(X)$是参数θ的充分统计量. 则对于任意一个检验函数$\phi(x)$, 存在另一个只依赖于$T(x)$的检验函数与它等价.

证明 定义一个新统计量$\psi(t)=E[\phi(X)|T(X)=t]$. 由于$0\le\phi(x)\le 1$，故由条件期望性质知，$\psi(t)\in[0,1]$，故它也是一个检验函数. 另外，又由全期望公式有

$$E_\theta\psi(T(X))=E_\theta[E(\phi(X)|T(X))]=E_\theta\phi(X),$$

故知$\psi(t)$与$\phi(x)$是等价的. □

这个定理告诉我们，当θ的充分统计量存在时，关于此参数的检验问题，我们仅需在由充分统计量构成的检验函数中去寻找就可以了. 这就是假设检验中的“充分性原则”.

定义 5.1.3 (MP检验) 对于参数分布族$\mathcal{F}=\{f(x,\theta):\theta\in\Theta\}$，设我们感兴趣的假设为

$$H_0:\theta=\theta_0\longleftrightarrow H_1:\theta=\theta_1\ (\theta_1\ne\theta_0), \tag{5.1.2}$$

并设有两个水平为α的检验函数$\phi_1(\mathbf{X}),\phi_2(\mathbf{X})$, 即满足

$$E_{\theta_0}\phi_i(\mathbf{X})\le\alpha,\quad i=1,2, \tag{5.1.3}$$

如果

$$\beta_{\phi_1}(\theta_1)\ge\beta_{\phi_2}(\theta_1), \tag{5.1.4}$$

则称检验ϕ_1比ϕ_2有效. 如果检验ϕ_1对于任一个水平小于等于α的检验ϕ_2，(5.1.4)式均成立，则称ϕ_1是假设(5.1.2)的水平α的最优势检验(Most Powerful Test，简记为MPT).

从上述定义可以看出，一个MP检验，不仅控制其第一类错误概率，而且也控制其第二类错误概率，且比一般水平α的显著性检验有效. 那这种好的检验存在吗？请见下节的N-P引理.

§5.2 Neyman-Pearson引理

从1928年开始的大约10年时间里，J. Neyman和E. S. Pearson合作发表了一系列有关假设检验的论文，从而建立了本章将要讲述的优势检验理论，一般我们称之为“Neyman-Pearson”理论. 本节将介绍其中最常用的一个定理，即N-P引理.

定理 5.1.2 (Neyman-Pearson基本引理) 对于参数分布族$\{f(x,\theta):\theta\in\Theta=\{\theta_0,\theta_1\}\}$, 则关于检验问题(5.1.2)，我们有如下结论:

(1) 对给定的$\alpha\in(0,1)$, 存在一个检验函数$\phi(x)$及常数$k\ge 0$, 使得

$$\phi(x)=\begin{cases}1, & f(x;\theta_1)>kf(x;\theta_0),\\ 0, & f(x;\theta_1)<kf(x;\theta_0).\end{cases} \tag{5.2.1}$$

且

$$E_{\theta_0}\phi(X) = \alpha. \tag{5.2.2}$$

(2) 由(5.2.1)和(5.2.2)式确定的检验函数$\phi(x)$是检验问题(5.1.2)的水平为α的MPT.

(3) 如果$\phi'(x)$是此检验问题的水平α的MPT，则一定存在常数$k \geq 0$，使得$\phi'(x)$满足(5.2.1)式. 又如$\phi'(x)$满足$E_{\theta_1}\phi'(X) < 1$，则它也满足(5.2.2)式.

证明　为方便起见，定义$\lambda(x) = f(x;\theta_1)/f(x;\theta_0)$.

(1) 为证明这一点，只须证明形如(5.2.1)式的检验函数存在，且满足(5.2.2)式. 为此，对于任意实数$c \geq 0$，定义

$$h(c) = P_{\theta_0}\{\lambda(X) \leq c\}.$$

由于它是随机变量$\lambda(X)$的分布函数，故它是非降、右连续函数，且$h(\infty) = \lim_{c\to\infty} h(c) = 1$, $h(c) - h(c-0) = P_{\theta_0}\{\lambda(X) = c\}$. 关于$h(0)$的取值有三种可能：

- $h(0) < 1-\alpha$. 此时存在$k \in (0,\infty)$, 使得$h(k-0) \leq 1-\alpha \leq h(k)$. 如$h(k) = 1-\alpha$, 则检验函数

$$\phi(x) = I\left(\lambda(X) > k\right), \tag{$*1$}$$

同时满足(5.2.1)和(5.2.2)式. 如$h(k) > 1-\alpha$, 则检验函数

$$\phi(x) = \begin{cases} 1, & \lambda(x) > k, \\ \frac{h(k)-(1-\alpha)}{h(k)-h(k-0)}, & \lambda(x) = k, \\ 0, & \lambda(x) < k., \end{cases} \tag{$*2$}$$

同时满足(5.2.1)和(5.2.2)式.

- $h(0) = 1-\alpha$. 则在(*1)式中取$k = 0$即可，即

$$\phi(x) = \begin{cases} 1, & \lambda(x) > 0, \\ 0, & \lambda(x) \leq 0, \end{cases}$$

同时满足(5.2.1)和(5.2.2)式.

- $h(0) > 1-\alpha$. 则在(*2)式中即$k = 0$, 即

$$\phi(x) = \begin{cases} 1, & \lambda(x) > 0, \\ \frac{h(0)-(1-\alpha)}{h(0)}, & \lambda(x) = 0, \\ 0, & \lambda(x) < 0., \end{cases}$$

同时满足(5.2.1)和(5.2.2)式.

(2) 为证满足(5.2.1)和(5.2.2)式的$\phi(x)$是(5.1.2)的MPT，设$\phi'(x)$ 是一个水平为α 的检验，即$E_{\theta_0}\phi'(X) \le \alpha$. 由于$\phi(x)$满足(5.2.1)式, 故

$$[\phi(x) - \phi'(x)][f(x;\theta_1) - kf(x;\theta_0)] \ge 0.$$

(当$\lambda(x) < k$时，$\phi(x) = 0 \le \phi'(x)$; 当$\lambda(x) > k$时，$\phi(x) = 1 \ge \phi'(x)$) 于是，

$$\int [\phi(x) - \phi'(x)][f(x;\theta_1) - kf(x;\theta_0)]\mathrm{d}x \ge 0,$$

即

$$E_{\theta_1}\phi(X) - E_{\theta_1}\phi'(X) \ge k(E_{\theta_0}\phi(X) - E_{\theta_0}\phi'(X)) \ge 0,$$

由于$\phi'(x)$的任意性知，$\phi(x)$是MPT.

(3) 设$\phi'(x)$是(5.1.2)的MPT，且由(1)可知，存在检验$\phi(x)$满足(5.2.1)和(5.2.2)式. 定义

$$\begin{aligned} &S^+ = \{x : \phi(x) > \phi'(x)\},\ S^- = \{x : \phi(x) < \phi'(x)\}, \\ &S = S^+ \cup S^-, \qquad S_1 = S \cap \{x : \lambda(x) \ne k\}. \end{aligned} \tag{$*3$}$$

由$\phi(x)$的定义知，$\forall\ x \in S^+, \lambda(x) > k, \forall\ x \in S^-, \lambda(x) < k$，且$\forall\ x \in S_1$, 有

$$\psi(x) = (\phi(x) - \phi'(x))(f(x;\theta_1) - kf(x;\theta_0)) > 0.$$

于是，

$$0 < \int_{S_1} \psi(x)d\mu(x) = E_{\theta_1}\phi(X) - E_{\theta_1}\phi'(X) - k[E_{\theta_0}\phi(X) - E_{\theta_0}\phi'(X)],$$

又因为$E_{\theta_0}\phi(X) = \alpha \ge E_{\theta_0}\phi'(X)$, 故由上式知，$E_{\theta_1}\phi(X) > E_{\theta_1}\phi'(X)$, 这显然与$\phi'(x)$是MPT矛盾. 故可知$S_1$是空集，这就是说当$\phi'(x) \ne \phi(x)$时，必有$\lambda(x) = k$, 而其它都相等，于是$\phi'(x)$也满足(5.2.1)式.

下证，如$E_{\theta_1}\phi'(X) < 1$，则它也满足(5.2.2)式. 此时$\phi'(x)$是水平α的MPT，且满足

$$E_{\theta_1}\phi'(X) = \int_{\mathcal{X}} \phi'(x)f(x;\theta_1)d\mu(x) < 1.$$

假设$E_{\theta_0}\phi'(X) < \alpha$. 令

$$\phi(x) = \min\{1, \phi'(x) + \alpha - E_{\theta_0}\phi'(x)\}.$$

则

$$E_{\theta_0}\phi(x) \le \int (\phi'(x) + \alpha - E_{\theta_0}\phi'(x))f(x;\theta_0)d\mu(x) = \alpha,$$

即是说上式定义的$\phi(x)$是水平α的检验. 另外，由其上述定义可知，$\phi(x) \ge \phi'(x)$ (事实上，如$\phi'(x) = 1$, 则$\phi(x) = 1$; 如$\phi'(x) = 0$, 则$\phi(x) = \alpha - E_{\theta_0}\phi'(X) > 0$), 且等号当且仅当$\phi'(x) = 1$时成立. 由于$E_{\theta_1}\phi'(X) < 1$, 故$P_{\theta_1}\{\phi'(X) = 1\} < 1$, 于是

$$E_{\theta_1}\phi(X) > E_{\theta_1}\phi'(X).$$

这显然与$\phi'(x)$为水平α的MPT检验矛盾，于是$E_{\theta_0}\phi'(X)=\alpha$. □

关于N-P引理，我们注意到：

- 上一定理的证明是构造性，故它可能用来构造形如假设(5.1.2)的MPT;
- 从N-P引理的证明可以看出，MPT是似然比统计量$\lambda(X)=f(X;\theta_1)/f(X;\theta_0)$的函数，这种检验我们也称之为似然比检验;
- 如果似然比$\lambda(X)$的分布是连续的，则假设(5.1.2)的MPT是非随机化的检验; 如果$\lambda(X)$的分布是离散的，则假设(5.1.2)的MPT有可能是随机化的;
- 对于非参数假设，N-P引理仍是正确的.

Corollary 5.2.1 如$\phi(X)$为假设(5.1.2)的水平α的MPT，则必有$\beta_\phi(\theta_1)\geq\alpha$. 如$0<\alpha<1$，且$f(x,\theta_1)\neq f(x,\theta_0)$，则$\beta_\phi(\theta_1)>\alpha$.

证明 对于本推论的第一个结论，只要取一个水平为α的检验$\phi^*(x)\equiv\alpha$即可.

下证第二个结论. 此时，$\alpha\in(0,1), f(x,\theta_0)\neq f(x,\theta_1)$. 如果$\beta_\phi(\theta_1)=\alpha$, 则$\phi'(x)\equiv\alpha$也是引假设水平$\alpha$的MPT. 于是，由N-P引理知，$\phi'(x)\equiv\alpha$应满足(5.2.1). 又因为$\alpha\in(0,1)$, 故由(5.2.1)的形式知，我们必有

$$f(x,\theta_1)=kf(x,\theta_0),\ a.e.$$

又由于$f(x,\theta_1),f(x,\theta_0)$均是PDF，故只有当$k=1$及$f(x,\theta_0)=f(x,\theta_1)$时上式才有可能成立, 但这与已知条件矛盾. □

例 5.2.1 设$X_1,\ldots,X_n$是来自$N(\mu,1)$的IID样本，试考虑假设

$$H_0:\mu=0\longleftrightarrow H_1:\mu=\mu_1(>0) \tag{5.2.3}$$

的水平α的MPT.

解 此时样本的联合PDF为

$$p(x;\mu)=(2\pi)^{-n/2}\exp\left\{-\frac{1}{2}\sum_{i=1}^n(x_i-\mu)^2\right\},$$

由此可求得似然比统计量为

$$\lambda(x)=\exp\{n\mu_1\bar{x}-n\mu_1^2/2\},$$

由此可见，$\lambda(x)$与$\bar{x}$成正比，于是，由N-P引理知，水平α的MPT为

$$\phi(X)=\begin{cases}1, & \bar{X}>k,\\ 0, & \bar{X}<k,\end{cases}$$

并满足$E_{\mu=0}\phi(X)=\alpha$. 又由于在H_0成立时，$\bar{X}\sim N(0,1/n)$, 故上面MPT中的k满足

$$\alpha=P_{H_0}\{\bar{X}>k\}=P_{H_0}\{\sqrt{n}\bar{X}>\sqrt{n}k\}=1-\Phi(\sqrt{n}k),$$

即$k = u_\alpha/\sqrt{n}$.

综上而述，知此假设的水平α的MPT为

$$\phi(X) = \begin{cases} 1, & \bar{X} > u_\alpha/\sqrt{n}, \\ 0, & \bar{X} < u_\alpha/\sqrt{n}. \end{cases} \tag{5.2.4}$$

□

关于例5.2.1，我们有如下几点注解：

- (5.2.4)式的MPT只与α及μ_1的符号有关，而与μ_1的具体数值无关.
- 如果把例5.2.1中的备选假设改为$H_1: \mu = \mu_1 < 0$，则其MPT的拒绝域为

  $$W = \{x : \bar{x} \le -u_\alpha/\sqrt{n}\},$$

  它也仅与μ_1的符号有关，而与其具体数值无关.
- 若上述总体的分布为$N(\mu, \sigma_0^2)$, 其中σ_0^2已知，且我们感兴趣的假设为

  $$H_0: \mu = \mu_0 \longleftrightarrow H_0: \mu = \mu_1(> \mu_0),$$

  则可能通过变换$y_i = (x_i - \mu_0)/\sigma_0$将此假设化为例5.2.1中的假设, 因此其水平为$\alpha$的MPT的拒绝域为

  $$W = \left\{ \frac{\bar{x} - \mu_0}{\sigma_0} \ge \frac{u_\alpha}{\sqrt{n}} \right\}.$$

- 如我们感兴趣的假设为如下的简单对复合假设:

  $$H_0: \mu = 0 \longleftrightarrow H_2: \mu > 0, \tag{5.2.5}$$

  则由于(5.2.4)式的检验$\phi(X)$是假设(5.2.3)的水平α的MPT, 故对于(5.2.3)的任一水平为α的检验$\phi'(X)$，均有

  $$E_{\mu_1}\phi(X) \ge E_{\mu_1}\phi'(X). \tag{$**1$}$$

  我们注意到, (5.2.4)式的$\phi(X)$与μ_1无关, 故它也是(5.2.5)的水平α的一个检验, 且对于(5.2.5)的任一水平为α的检验$\phi''(X)$，由于它也是(5.2.3)的水平为α的一个检验，故由(**1)式知，

  $$E_{\mu_1}\phi(X) \ge E_{\mu_1}\phi''(X),\ \forall\ \mu_0 > 0, \tag{$**2$}$$

  故由(**2)式知，$\phi(X)$也是(5.2.5)的水平α的MPT，关于$\mu > 0$是一致的.

例 5.2.2 设$X_1, \ldots, X_n$为来自均匀分布$U(0, p)$的IID样本，考虑假设

$$H_0: p = 1 \longleftrightarrow H_1: p = p_1(> 1)$$

的水平α的MPT.

解 此时的似然比统计量为

$$\lambda(x) = \frac{\prod_{i=1}^n f(x;p_1)}{\prod_{i=1}^n f(x;p_0)} = \begin{cases} p_1^{-n}, & 0 < x_{(n)} < 1, \\ \infty, & 1 \le x_{(n)} < p_1. \end{cases}$$

于是，$\lambda(X)$的分布函数

$$h(c) = P_{H_0}\{\lambda(X) \le c\}$$

是一个退化分布，即当$c < p_1^{-n}$时，$h(c) = 0$; 当$c \ge p_1^{-n}$时，$h(c) = 1$. 这样根据N-P引理的证明，我们取$k = p_1^{-n}$, 此时有

$$0 = h(k-0) < 1-\alpha < h(k) = 1$$

则由N-P引理知，此时的MPT是如(*2)式的随机化检验，即

$$\phi(x) = \begin{cases} 1, & 1 \le x_{(n)} < p_1, \\ \alpha, & 0 < x_{(n)} < 1, \\ 0, & \text{其它}. \end{cases} \tag{$**$}$$

由于$X_{(n)}$是连续的，故我们也可以取如下的非随机检验(也可以看成是把(**)或的随机部分非随机化)：

$$\phi(x) = \begin{cases} 1, & c \le x_{(n)} < p_1, \\ 0, & 0 < x_{(n)} < c, \end{cases}$$

其中常数c由$E_{H_0}\phi(X) = \alpha$决定. 实际上，当H_0成立时，$X_{(n)}$的PDF为$p(t) = nt^{n-1},\ t \in (0,1)$, 故常数$c$满足

$$\alpha = \int_c^1 nt^{n-1} dt,$$

由此求得$c = \sqrt[n]{1-\alpha}$, 即这个非随机化的MPT为

$$\phi(x) = \begin{cases} 1, & \sqrt[n]{1-\alpha} \le x_{(n)} < p_1, \\ 0, & 0 < x_{(n)} < \sqrt[n]{1-\alpha}. \end{cases} \tag{$***$}$$

我们注意到(**)与(***)式的区别在于把随机化检验非随机化. □

对于例5.2.2，如果我们感兴趣的检验为

$$H_0 : p = p_0 \longleftrightarrow H_1 : p = p_1 (> p_0)$$

则可以通过变换$y_i = x_i/p_0$来把上述问题变成例5.2.2的问题. 此时其水平为α的非随机化MP检验为

$$\phi(x) = \begin{cases} 1, & \sqrt[n]{1-\alpha} \le x_{(n)}/p_0 < p_1/p_0, \\ 0, & 0 < x_{(n)}/p_0 < \sqrt[n]{1-\alpha}. \end{cases}$$

例 5.2.3 设$X_1,\ldots,X_n$为来自Poisson分布$P(\lambda)$的IID样本，求假设

$$H_0:\lambda=1\longleftrightarrow H_1:\lambda=\lambda_1(>1)$$

水平为α的MP检验.

解 此时的似然比统计量为

$$\lambda(x)=\prod f(x_i;\lambda_1)/\prod f(x_i;\lambda_0)=\lambda_1^{\sum x_i}\exp\{-(\lambda_1-1)\},$$

由于它关于$T(x)=\sum x_i$单调上升，故由N-P引理知，其水平为α的MP检验为

$$\phi(x)=\begin{cases}1, & T(x)>k,\\ \delta, & T(x)=k,\\ 0, & T(x)<k,\end{cases}$$

其中k,δ满足$E_{H_0}\phi(X)=\alpha$. 由于当H_0成立时，$T(X)\sim P(n)$，故k满足

$$\alpha=P_{H_0}\{T(X)>k\}+\delta P_{H_0}\{T(X)=k\}=\sum_{i=k+1}^{\infty}\frac{e^{-n}n^i}{i!}+\delta\frac{n^k}{k!}.$$

如果存在整数k_0，使得$\sum_{i=k_0+1}^{\infty}\frac{e^{-n}n^i}{i!}=\alpha$, 则此时的MP检验为

$$\phi(x)=\begin{cases}1, & T(x)>k_0,\\ 0, & T(x)\le k_0.\end{cases}$$

如果存在整数k_0，使得

$$c_1=\sum_{i=k_0+1}^{\infty}\frac{e^{-n}n^i}{i!}<\alpha<\sum_{i=k_0}^{\infty}\frac{e^{-n}n^i}{i!}=c_2,$$

则取$k=k_0,\delta=\frac{\alpha-c_1}{c_2-c_1}$即可. □

例 5.2.4 设$X_1,\ldots,X_n$为来自$b(1,p)$的iid样本，试求假设

$$H_0:p=p_0\longleftrightarrow H_1:p=p_1(>p_0)$$

的水平α的MPT.

解 此时的似然比统计量为

$$\lambda(x)=\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right)^{\sum_{i=1}^n x_i}\left(\frac{1-p_1}{1-p_0}\right)^n,$$

因为$p_1 > p_0$，故$\lambda(x)$是$T=\sum_{i=1}^n x_i$的单增函数. 于是，由N-P引理可知，此假设的水平α的MPT为

$$\phi(X)=\begin{cases} 1, & \sum_{i=1}^n X_i \geq k, \\ 0 & \sum_{i=1}^n X_i < k, \end{cases}$$

其中k满足

$$P_{p_0}\left\{\sum_{i=1}^n X_i \geq k\right\}=\sum_{i=k}^n \begin{pmatrix} n \\ i \end{pmatrix} p_0^i(1-p_0)^{n-i}=\alpha. \tag{$*$}$$

但是，我们注意到，由于$\sum_{i=1}^n X_i$的零分布是二项分布，为离散的，故对于事先给定的α，不一定正好存在一个正整数k使得(*)成立. 一般地，我们只能找到k使得

$$P_{p_0}\left\{\sum_{i=1}^n X_i \geq k\right\} > \alpha > P_{p_0}\left\{\sum_{i=1}^n X_i \geq k+1\right\} \triangleq \alpha_1.$$

由于此时我们找不到一个水平严格为α的MP检验，故如令

$$\delta=\frac{\alpha-\alpha_1}{\begin{pmatrix} n \\ k \end{pmatrix} p_0^k(1-p_0)^{n-n_0}},$$

则所求的水平α的MP检验为

$$\phi(X)=\begin{cases} 1, & \sum_{i=1}^n X_i \geq k, \\ \delta, & \sum_{i=1}^n X_i = k, \\ 0 & \sum_{i=1}^n X_i < k. \end{cases} \tag{5.2.6}$$

□

从前两个例子可以看出，对于某些离散分布，其MP检验有时是随机的. 此时，对于形如(5.2.6)式的随机化检验，我们的检验准则如下：

- 当$\sum_{i=1}^n X_i > k$时，拒绝H_0;
- 当$\sum_{i=1}^n X_i < k$时，接受H_0;
- 当$\sum_{i=1}^n X_i = k$时，先做一个Bernoulli试验$b(1,\delta)$. 如果该试验结果为1，则拒绝H_0；否则就接受H_0.

另外，从这两个例子还可以看出，对于随机检验而言，一个检验可能不是拒绝域的示性函数.

再者，从上面几个例子可以看出，我们得到的MP检验均是充分统计量的函数，这也与定理5.1.1的结论是吻合的.

§5.3 一致最优势检验

上一节给出的N-P引理仅考虑形如(5.1.2)的简单假设对简单假设的最优势检验问题. 但在许多实际问题中，仅考虑简单假设是不够的. 于是，本节将考虑复杂假设的检验问题.

§5.3.1 定义及某些有用的结论

定义 5.3.1 (一致最优势检验) 对于参数分布族$\{f(x,\theta):\theta\in\Theta\}$, 设感兴趣的假设为

$$H_0:\theta\in\Theta_0\longleftrightarrow H_1:\theta\in\Theta_1. \tag{5.3.1}$$

如果对于两个水平为α的检验ϕ_1,ϕ_2，即

$$E_\theta\phi_i\le\alpha,\ \forall\ \theta\in\Theta_0,\ i=1,2, \tag{5.3.2}$$

且满足

$$\beta_{\phi_1}(\theta)=E_\theta\phi_1\ge E_\theta\phi_2=\beta_{\phi_2}(\theta),\ \forall\ \theta\in\Theta_1, \tag{5.3.3}$$

则称检验ϕ_1一致优于检验ϕ_2. 如果存在一个检验ϕ_1, 使对任何水平为α的检验ϕ_2，均有(5.3.3)成立, 则称检验ϕ_1是假设(5.3.1)的水平α 的一致最优势(Uniformly Most Powerful, 简记为UMP)检验.

通过对定义5.1.3与定义5.3.1的比较，我们发现，MP检验仅是UMP检验针对简单假设对简单假设的一种特殊情况. 另外，这里所说的一致性是关于备选假设下的参数空间而言的.

由N-P引理知，对于简单对简单的假设，MP 检验是存在的，另外，通过例5.2.1后面的注不难看出，我们可以通过MP检验来构造某些假设的UMP检验. 但这并不能说对于一般的假设，UMP检验是存在的. 在讲述某些单边与双边假设的UMP检验之前，我们先给出如下一个非常有用的结论.

定理 5.3.1 设有如下三个检验问题

$$H_0:\theta\in\Theta_0\longleftrightarrow H_1:\theta\in\Theta_1, \tag{5.3.4}$$

$$H_0:\theta\in\Theta_{01}\longleftrightarrow H_1:\theta\in\Theta_1, \tag{5.3.5}$$

$$H_0:\theta\in\Theta_0\longleftrightarrow H_1:\theta=\theta_1, \tag{5.3.6}$$

其中$\Theta_{01}\subset\Theta_0,\theta_1\in\Theta_1$，并设$\phi(x)$是检验问题(5.3.4)的水平$\alpha$的检验，则

(1) 如果$\phi(x)$是检验问题(5.3.5)的水平α的UMPT，则$\phi(x)$也是(5.3.4)的水平α的UMPT.

(2) $\phi(x)$是检验问题(5.3.4)的水平α的UMPT的充要条件是，对任意的$\theta_1\in\Theta_1$, $\phi(x)$是检验问题(5.3.6)的水平α的MPT.

证明 证明(1)时，注意到(5.3.4)的水平为α的检验肯定是(5.3.5)的水平α的检验即可. 证明(2)时，注意到UMPT的定义即可. □

定理 5.3.2 设$\phi(x)$是(5.3.4)的水平α的检验，如果对某个$\theta_0\in\Theta_0$及任意一个$\theta_1\in\Theta_1$,$\phi(x)$都是假设

$$H_0:\theta=\theta_0\longleftrightarrow H_1:\theta=\theta_1 \tag{5.3.7}$$

的水平α的MPT，则$\phi(x)$也是(5.3.4)的水平α的UMPT.

证明　此定理是定理5.3.1的推论.

综合定理5.3.2与例5.2.1，我们知道，在例5.2.1中给出的MP检验也是相应假设的UMP，即，对于正态总体$N(\mu,1)$, 假设

$$H_0:\mu=0\longleftrightarrow H_1:\mu>0$$

的水平为α的UMPT为

$$\phi(X)=\begin{cases}1, & \bar{X}>u_\alpha/\sqrt{n},\\ 0, & \bar{X}<u_\alpha/\sqrt{n}.\end{cases}$$

而假设

$$H_0:\mu=0\longleftrightarrow H_1:\mu<0$$

的水平为α的UMPT为

$$\phi(X)=\begin{cases}1, & \bar{X}<-u_\alpha/\sqrt{n},\\ 0, & \bar{X}>-u_\alpha/\sqrt{n}.\end{cases}$$

总而言之，我们注意到：

- 如果简单假设对简单假设的MPT不依赖于备选假设中的参数值，则可适当扩大备选假设成复杂假设;
- 如果简单假设对简单假设的MPT的势函数是单调的，则也可以适当扩大备选假设成复杂假设.

但是，我们要必须注意到，并不是所有的复杂假设对复杂假设的UMPT都是存在的，它的存在性不仅依赖于总体分布，而且还依赖于假设的复杂情况. 本节将只考虑针对单调似然比分布族的两种单边假设和针对指数型分布族的三种双边假设的UMP检验问题.

§5.3.2　单调似然比分布族

为了讨论单边假设的UMP检验，我们先引入如下的单调似然比(Monotone Likelihood Ratio, 简记为MLR)分布族.

定义 5.3.2 (MLR分布族)　一个参数分布族$\{f(x;\theta):\theta\in\Theta\}$称为关于统计量$T(X)$单增(减)似然比分布族，如果它满足

(1) Θ是直线上的一个区间;

(2) 对于$\theta_1\neq\theta_2$, 有$f(x,\theta_1)\neq f(x,\theta_2)$;

(3) 似然比$\lambda(x)=f(x;\theta_2)/f(x;\theta_1)$ $(\theta_2>\theta_1)$是统计量$T(x)$的非降(增)函数.

单增或减似然比分布族统称为MLR分布族.

MLR分布族与单参数指数型分布之间的关系如下：

对于单参数指数型分布

$$f(x;\theta) = c(\theta)\exp\{Q(\theta)T(x)\}h(x),$$

其中$c(\theta) > 0$, 如果$Q(\theta)$是θ的严格单增(减)函数，则它必为关于其充分统计量的单增(减)似然比分布族.

于是，我们常用的分布，如二项分布、负二项分布、Poisson分布、单参数正态分布、指数分布等均是MLR分布族. 另外，虽然均匀分布$U(0,\theta)$并不是单参数指数型分布，但如定义$0/0=\infty$，则它也是关于其充分统计量$X_{(n)}$的MLR分布族.

例 5.3.1 *考察超几何分布族的单调性.*

解 此时的CDF为

$$p(x,m) = \frac{\begin{pmatrix} m \\ x \end{pmatrix}\begin{pmatrix} N-m \\ n-x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}},$$

其中$0 < m < N$为正整数，是未知参数.

由于

$$\frac{p(x,m+1)}{p(x,m)} = \frac{\begin{pmatrix} m+1 \\ x \end{pmatrix}\begin{pmatrix} N-m-1 \\ n-x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}\frac{\begin{pmatrix} N \\ n \end{pmatrix}}{\begin{pmatrix} m \\ x \end{pmatrix}\begin{pmatrix} N-m \\ n-x \end{pmatrix}} = \frac{(m+1)(N-m-n+x)}{(N-m)(m+1-x)}$$

是x的单增函数，故知超几何分布是一单增似然比分布族. □

对于MLR分布族，我们有如下的非常有用的定理.

定理 5.3.3 *如果分布族$\{f(x;\theta):\theta\in\Theta\subset\mathcal{R}\}$是关于统计量$T(x)$非降的MLR分布族，$\psi(t)$是非降(增)函数，则$E_\theta\psi(T(X))$也是$\theta$的一个非降(增)函数.*

证明 设$\psi(t)$是非降的. $\forall\ \theta_1 < \theta_2$, 定义$\lambda(x) = f(x;\theta_2)/f(x;\theta_1)$,

$$A = \{x : f(x;\theta_1) < f(x;\theta_2)\},\quad B = \{x : f(x;\theta_1) > f(x;\theta_2)\}.$$

于是，$\forall\ x_1 \in A,\ x_2 \in B$, 有$\lambda(x_1) > 1 > \lambda(x_2)$, 且由MLR族定义的条件(2)知$A\cup B = \mathcal{X}$.

又由于此分布族关于$T(x)$是单增的，且$\psi(t)$也是非降的，故$\forall\ x_1 \in A,\ x_2 \in B$, $T(x_1) \geq T(x_2) \Rightarrow \psi(T(x_1)) \geq \psi(T(x_2))$.

定义

$$a = \inf\{\psi(T(x)) : x \in A\},\quad b = \sup\{\psi(T(x)) : x \in B\},$$

则由集合A, B的定义知，$a \geq 0 \geq b$. 另外，

$$\begin{aligned}
& E_{\theta_2}\psi(T(X)) - E_{\theta_1}\psi(T(X)) \\
= & \int_{\mathcal{X}} \psi(T(x))(f(x;\theta_2) - f(x;\theta_1))\mathrm{d}\mu(x) \\
\geq & a\int_A (f(x;\theta_2) - f(x;\theta_1))d\mu(x) + b\int_B (f(x;\theta_2) - f(x;\theta_1))\mathrm{d}\mu(x).
\end{aligned}$$

又因为

$$\int_{A\cup B} (f(x;\theta_2) - f(x;\theta_1))\mathrm{d}\mu(x) = 1,$$

故

$$\int_B (f(x;\theta_2) - f(x;\theta_1))\mathrm{d}\mu(x) = -\int_A (f(x;\theta_2) - f(x;\theta_1))\mathrm{d}\mu(x),$$

于是，

$$\begin{aligned}
& E_{\theta_2}\psi(T(X)) - E_{\theta_1}\psi(T(X)) \\
\geq & (a-b)\int_A (f(x;\theta_2) - f(x;\theta_1))\mathrm{d}\mu(x) \geq 0.
\end{aligned}$$

注意到θ_1, θ_2的任意性，故命题得证. □

Corollary 5.3.1 如果分布族$\{f(x;\theta) : \theta \in \Theta \subset \mathcal{R}\}$是关于统计量$T(x)$非降的MLR分布族，则

(1) $E_\theta T(X)$也是θ的一个非降函数;

(2) $P_\theta\{T(X) \leq t_0\}$是$\theta$的一个非增函数.

证明 如在定理5.3.3中取$\psi(t) = t$，则得证(1). 如在定理5.3.3中取$\psi(t) = I_{(t\leq t_0)}$, 则知此时的$\psi(t)$对于任意给定的$t_0$关于$t$是单减的，且$E_\theta\psi(T(X)) = P_\theta\{T(X) \leq t_0\}$, 于是由定理5.3.3可知(2)成立. □

注 5.3.1 上一推论的结论(2)告诉我们，在MLR分布族中，统计量$T(X)$的分布关于θ是单减的. 这一点对求一般情况下的区间估计非常有用，具体的请参见茆诗松、王静龙、濮晓龙(2006).

例 5.3.2 试讨论二项分布的单调性.

解 设$X \sim b(n,p)$, 其中$p \in (0,1)$. 此时其分布函数为

$$P(x;p) = C_n^x p^x(1-p)^{n-x} = (1-p)^n\left(\frac{p}{1-p}\right)^x C_n^x = (1-p)^n \exp\{x\ln(p/(1-p))\}C_n^x,$$

于是，它是指数型分布，且$Q(p) = \ln\frac{p}{1-p}, T(x) = x$. 由于$Q(p)$关于$p$是单增的，故知它是关于$T(x)$的单增MLR 分布族.

另外，由于$T(X) \sim b(n,p)$, 故$E_pT(X) = np$关于p单增，且$P_p\{X \le k\} = \sum_{i=0}^k C_n^i p^i(1-p)^{n-i}$关于$p$是严格单减的. (注意到$\sum_{i=0}^k C_n^i p^i(1-p)^{n-i} = \frac{n!}{(k-1)!(n-k)!}\int_p^1 t^{k-1}(1-t)^{n-k}\mathrm{d}t$, 这是一个不完全Beta积分). 与上一推论的结论是吻合的. □

§5.3.3 单边假设的UMP检验

在本小节，我们将考虑如下单边假设

$$H_0 : \theta \le \theta_0 \longleftrightarrow H_1 : \theta > \theta_0, \tag{5.3.7}$$

$$H_0 : \theta \ge \theta_0 \longleftrightarrow H_1 : \theta < \theta_0 \tag{5.3.8}$$

的UMP检验，并且假设其分布族为MLR分布族.

定理 5.3.4 设单参数分布族$\{f(x;\theta) : \theta \in \Theta \subset \mathcal{R}\}$是关于统计量$T(x)$非降的MLR分布族，则关于单边假设(5.3.7),

(1) 存在水平为α的UMP检验函数

$$\phi(T) = \begin{cases} 1, & T(x) > c, \\ \delta, & T(x) = c, \\ 0, & T(x) < c, \end{cases} \tag{5.3.9}$$

其中常数δ, c由下式决定

$$E_{\theta_0}\phi(T(X)) = \alpha. \tag{5.3.10}$$

(2) 势函数$\beta(\theta) = E_{\theta_0}\phi(T(X))$是非降的，且在集合$\{\theta : 0 < \beta(\theta) < 1\}$严格单增.

(3) 如$\phi'(x)$是假设(5.3.7)的一个检验，且满足$E_{\theta_0}\phi'(X) = \alpha$，则

$$E_\theta\phi(T(X)) \le E_\theta\phi'(X), \ \forall\ \theta < \theta_0.$$

证明 先证(1). $\forall\ \theta_1 > \theta_0$, 先考虑简单对简单的检验问题:

$$H_0 : \theta = \theta_0 \longleftrightarrow H_1 : \theta = \theta_1. \tag{$*1$}$$

由于此时的分布族是关于$T(x)$ 单增的MLR族，故由N-P引理知：对于假设(*1), 存在一个形如(5.3.9) 的MP 检验，且满足(5.3.10)式. 由于(5.3.9)式的$\phi(T(x))$与θ_1的具体数值无关，只与它是否大于θ_0有关，故由定理5.3.1 的结论(2)知，$\phi(T(x)$也是检验问题

$$H_0 : \theta = \theta_0 \longleftrightarrow H_1 : \theta > \theta_0. \tag{$*2$}$$

的水平为α的UMPT.

另外，由于(5.3.9)式的$\phi(T(x)$ 是$T(x)$ 的非降函数，故由定理5.3.3知，其势函数$\beta(\theta) = E_\theta\phi(T(X))$ 也是θ的非降函数. 于是，$\forall\ \theta \le \theta_0$，$\beta(\theta) \le \beta(\theta_0) = \alpha$. 这就是说，$\phi(T(x))$也

是假设(5.3.7)的水平α的检验. 于是由定理5.3.1的结论(1)知, $\phi(T(x))$也是假设(5.3.7)的水平为α的UMPT.

下证(2). $\beta(\theta)$的非降性已经在(1)的证明过程证得，下证其在集合$\{\theta : 0 < \beta(\theta) < 1\}$严格单增.

事实上，由N-P引理知，由(5.3.9)式定义的$\phi(T(x))$也是检验问题

$$H_0 : \theta = \theta_1 \longleftrightarrow H_1 : \theta = \theta_2. \tag{*3}$$

的水平为$\alpha_1 = E_{\theta_1}\phi(T(X))$的MPT，其中$\theta_2 > \theta_1$. 则由推论5.2.1知，当$0 < \alpha_1 < 1$时，

$$E_{\theta_2}\phi(T(X)) > \alpha_1 = E_{\theta_1}\phi(T(X)).$$

而$0 < \alpha_1 < 1$ 是由条件所保证的，再由$\theta_2 > \theta_1$的任意性，结论(2) 得证.

下证(3). $\forall\ \theta_1 < \theta_0$, 记$\alpha' = E_{\theta_1}\phi(T(X))$. 则由N-P引理知，$\phi(T(x))$是假设

$$H_0 : \theta = \theta_1 \longleftrightarrow H_1 : \theta = \theta_0. \tag{*4}$$

的水平为α'的MPT. 如$\phi'(x)$为任一满足

$$E_{\theta_0}\phi'(X) = E_{\theta_0}\phi(T(X)) = \alpha, \tag{*}$$

的检验，则要证$E_{\theta_1}\phi(T(X)) \le E_{\theta_1}\phi'(X)$. 如果它不成立，即$\alpha' = E_{\theta_1}\phi(T(X)) > E_{\theta_1}\phi'(X)$, 这就是说$\phi'(x)$也是(*4)的水平为$\alpha'$的检验，又由(*)式知(达到了MPT的势), $\phi'(x)$也是(*4)的水平为α'的MPT. 又因为$E_{\theta_0}\phi'(X) = \alpha < 1$, 故由N-P引理的结论(3)知，$E_{\theta_1}\phi'(X) = \alpha'$,这与假设矛盾，故结论得证. □

由定理5.3.4，我们知道：

- 满足(5.3.9)和(5.3.10)式的检验$\phi(T(x))$, 不仅是UMPT，即在第一类错误概率为α 的检验类$\{\psi(x) : E_{\theta_0}\psi(X) \le \alpha\}$中，其第二类错误概率最小，而且在所有满足(5.3.10)式的检验类$\{\psi(x) : E_{\theta_0}\psi(X) = \alpha\}$中，其第一类错误概率也是最小的.
- 对于假设
$$(I)\quad H_0 : \theta = \theta_0 \longleftrightarrow H_1 : \theta > \theta_0,$$
定理5.3.4的结论仍然成立.
- 对于假设(5.3.8)
$$(II)\quad H_0 : \theta \ge \theta_0 \longleftrightarrow H_1 : \theta < \theta_0,$$
定理5.3.4的结论仍然成立, 只是(5.3.9)式中的不等号要变号, 结论(3)中的$\theta < \theta_0$要改成$\theta > \theta_0$.

下面，我们看几个例子.

例 5.3.3 设某产品共有N件，其中含有m件次品. 为了估计m，现从中不返回地抽取n件进行检验，设n件中有X件不合格品，则由概率知识知，X服从超几何分布. 试求如下假设

$$H_0 : 0 \le m \le m_0 \longleftrightarrow H_1 : m_0 < m \le N$$

的水平为α的UMPT.

解 由例5.3.1知，超几何分布是关于$T(x)=x$的单增似然比分布族. 则由定理5.3.4知，此检验问题存在水平为α的UMPT，其检验函数为

$$\phi(x)=\begin{cases}1, & x>c,\\ \delta, & x=c,\\ 0, & x<c,\end{cases}$$

其中δ, c由$E_{m_0}\phi(X)=\alpha$决定.

用类似于例5.2.3的方法，我们可以求得其UMPT为

$$\phi(x)=\begin{cases}1, & x>c,\\ \frac{(\alpha-\alpha_1)C_N^n}{C_{m_0}^c C_{N-m_0}^{n-c}}, & x=c,\\ 0, & x<c,\end{cases}$$

其中c满足条件

$$P_{m_0}\{X\geq c\}>\alpha\geq P_{m_0}\{X>c\}=\alpha_1.$$

特别地，当$N=1000, n=100, m_0=10, \alpha=0.05$时，可以算得

$$P(0;10)=0.3470;\ P(1;10)=0.3894;\ P(2;10)=0.1945;\ P(3;10)=0.0569.$$

于是，

$$\sum_{i\geq 3}P(i;10)=1-\sum_{i\leq 2}P(i;10)=0.0695>0.05>0.0122=\sum_{i\geq 4}P(i;10),$$

这样，此时的水平为0.05的UMPT为

$$\phi(x)=\begin{cases}1, & x\geq 4,\\ 0.6643, & x=3,\\ 0, & x\leq 2.\end{cases}$$

□

例 5.3.4 设$X_1,\ldots,X_n$是来自正态总体$N(0,\sigma^2)$的IID样本，试求假设

$$H_0:\sigma^2\leq\sigma_0^2\longleftrightarrow H_1:\sigma^2>\sigma_0^2$$

的水平α的UMPT.

解 此时样本的联合PDF为

$$f(x;\sigma^2)=\exp\left\{-\sum_{i=1}^n x_i^2/2\sigma^2\right\}/\sqrt{2\pi\sigma^2},$$

于是，$Q(\sigma^2) = -1/2\sigma^2$为$\sigma^2$的单增函数，故它是关于$T(x) = \sum x_i^2$的单增似然比分布族. 由定理5.3.4知，此假设的UMPT为

$$\phi(T(x)) = \begin{cases} 1, & T(x) \geq c, \\ 0, & T(x) < c, \end{cases}$$

其中c满足$E_{\sigma_0}\phi(T(X)) = \alpha$. 注意到，当$\sigma^2 = \sigma_0^2$时，$T(X)/\sigma_0^2 \sim \chi^2(n)$, 于是，常数$c$满足

$$\int_{c/\sigma_0^2}^{\infty} \chi^2(x; n)\mathrm{d}x = \alpha,$$

于是，$c = \sigma_0^2\chi_\alpha^2(n)$.

当$\sigma_0 = 1, n = 10, \alpha = 0.1$时，此UMP检验的拒绝域为

$$W = \left\{x : \sum_{i=1}^{n} x_i^2 \geq \sigma_0^2\chi_\alpha^2(n)\right\} = \left\{x : \sum_{i=1}^{10} x_i^2 \geq 15.99\right\}.$$

另外，我们也可以用统计量$T_1(x) = \sum_{i=1}^{10}(x_i - \bar{x})^2$来检验本假设. 如取拒绝域为

$$W_1 = \left\{x : \sum_{i=1}^{n}(x_i - \bar{x})^2 \geq \sigma_0^2\chi_\alpha^2(n-1)\right\} = \left\{x : \sum_{i=1}^{10}(x_i - \bar{x})^2 \geq 14.68\right\},$$

则记此时的检验函数为ϕ_1. 我们不难验证它也是本假设的一个水平为$\alpha = 0.1$的检验.

由于，上述两个检验的势分别为

$$\beta_\phi(\sigma^2) = P\left\{\mathbf{X} : \sum_{i=1}^{10} X_i^2 \geq 15.99\right\} = 1 - \chi^2(15.99/\sigma^2; 10),$$

$$\beta_{\phi_1}(\sigma^2) = P\left\{\mathbf{X} : \sum_{i=1}^{10}(X_i - \bar{X})^2 \geq 14.68\right\} = 1 - \chi^2(14.68/\sigma^2; 9),$$

其中$\chi^2(x; n)$表示自由度为n的χ^2分布随机变量的CDF. 对于不同的σ^2，我们计算如上的势函数，其结果见图5.1.

在图5.1中，实线表示UMP检验ϕ的势函数，而虚线则表示检验ϕ_1的势函数. 从图中可以看出，此时的UMP检验ϕ的势明显优于检验ϕ_1. □

从上个例子可以看出，我们在§4.2.2中讲述的关于单样本正态总体的方差的显著性检验(均值已知时)是MP检验. 我们同样可以证明许多的单样本正态总体的显著性检验都是MP检验.

我们回忆在例4.1.2中，我们讲述了如何利用显著性检验来控制其决策风险的问题. 但是，当时我们仅考虑了第一类错误概率. 事实上，第一类错误概率控制的是企业的风险，而第二类错误概率控制的则是商店的风险. 下面我们通过(n, c)检验方案的例子对此加以说明.

例 5.3.5 ((n, c)方案检验) 对于某批产品，以p记其次品率. 现从中随机地抽取n件产品，以检验如下假设:

$$H_0 : p \leq p_0 \longleftrightarrow H_1 : p \geq p_1, \tag{5.2.7}$$

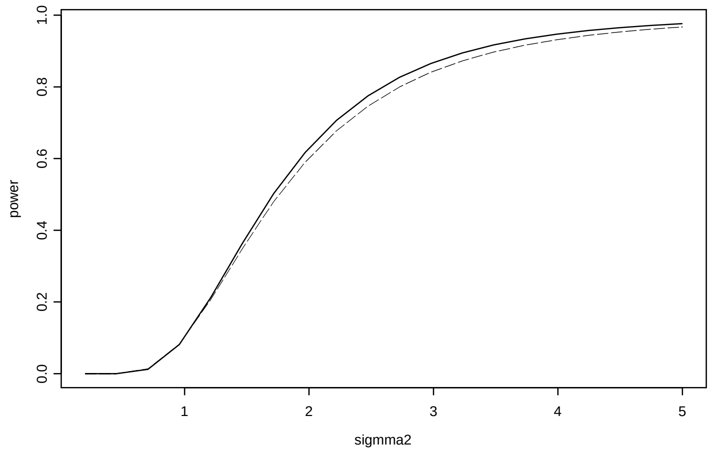

图5.1 例5.3.4中两个检验的势函数的比较

其中$0 < p_0 < p_1 < 1$为两个给定的常数.

解 对于此问题如记$T(\mathbf{X})$为n件产品中的次品数，则由定理5.3.4可知，此时的MP检验(为简单，我们仅考虑非随机检验的情形)为

$$\phi(X) = \begin{cases} 1, & T(\mathbf{X}) > c, \\ 0 & T(\mathbf{X}) \le c. \end{cases}$$

另外，为了既对厂家负责，又对商店负责，我们要求上述检验中的(n, c)满足下面两个条件：

$$\begin{cases} L(p_0; n, c) \ge 1 - \alpha, \\ L(p_1; n, c) \le \beta, \end{cases} \tag{5.3.11}$$

其中$0 < \alpha, \beta < 1$是两个事先给定的常数，函数L如下定义：

$$L(p; n, k) = P\{T(\mathbf{X}) \le k\} = \sum_{i=0}^{k} \begin{pmatrix} n \\ i \end{pmatrix} p^i (1-p)^{n-i}, \tag{5.3.12}$$

上一等式成立的原因为$T \sim B(n, p)$. 另外，我们称满足(5.3.11)式的检验为此假设的(n, c)方案检验. 一般地，我们很难求得满足(5.3.11)式的整数n, c. 但由于二项分布可由正态分布来近似，即

$$L(p; n, c) = P\{T(\mathbf{X}) \le c\} = P\left(\frac{T - np}{\sqrt{npq}} \le \frac{k - np}{\sqrt{npq}}\right) \approx \Phi\left(\frac{k - np}{\sqrt{npq}}\right),$$

故(5.3.11)式近似地等价于

$$\begin{cases} \frac{c - np_0}{\sqrt{np_0 q_0}} \ge u_\alpha, \\ \frac{c - np_1}{\sqrt{np_1 q_1}} \le u_{1-\beta}. \end{cases} \tag{5.3.13}$$

如取$p_0=0.04, p_1=0.1, \alpha=0.05, \beta=0.1$，则由(5.3.13)可求得$n=139, c=9.3$，即如在抽取的139个样本中，当次品数大于9时，此批产品不应出厂，否则可以出厂. □

例 5.3.6 设$X_1,\ldots,X_n$为来自Poisson总体$P(\lambda)$的iid样本，$\lambda>0$是未知参数，请给出假设

$$H_0:\lambda\le\lambda_0\longleftrightarrow H_1:\lambda>\lambda_0$$

的水平α的UMP检验.

解 对于此问题，由于其样本分布为

$$f(\mathbf{X},\lambda)=\frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}=\left(\prod_{i=1}^n x_i!\right)^{-1}e^{-n\lambda}\exp\left\{\sum_{i=1}^n x_i\ln\lambda\right\},$$

故知它是单参数指数型分布，且$T=\sum_{i=1}^n X_i, Q(\lambda)=\ln\lambda$. 另外，由于$Q(\lambda)$是$\lambda$的严格单增函数，则由指数型分布与MLR分布族间的关系和定理5.3.4知，此假设的水平α的UMP检验为

$$\phi(x)=\begin{cases}1, & \sum_{i=1}^n x_i>c,\\ \delta, & \sum_{i=1}^n x_i=c,\\ 0, & \sum_{i=1}^n x_i<c,\end{cases}$$

其中c,δ由$E_{\lambda_0}\phi(\mathbf{X})=\alpha$决定.

事实上，由于$T=\sum_{i=1}^n X_i\sim P(n\lambda)$, 则$c$由下式决定

$$\sum_{k\ge c}\frac{(n\lambda)^k}{k!}e^{-n\lambda}>\alpha\ge\sum_{k\ge c+1}\frac{(n\lambda)^k}{k!}e^{-n\lambda}\triangleq\alpha_1,$$

而

$$\delta=(\alpha-\alpha_1)\frac{c!}{(n\lambda)^c}e^{-n\lambda}.$$

□

例 5.3.7 设$X_1,\ldots,X_n$为来自正态总体$N(\mu,\sigma_0^2)$的iid样本，其中σ_0^2已知，求假设

$$H_0:\mu\le 0\longleftrightarrow H_1:\mu>0$$

的水平α的UMP检验.

解 对于此问题，其样本分布为

$$f(\mathbf{x};\mu)=(2\pi\sigma_0^2)^{-n/2}e^{-n\mu^2/2\sigma_0^2}\exp\{n\mu\bar{x}/\sigma_0^2\}\exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^n x_i^2\right\}.$$

由于它是单参数指数型分布，且因为$Q(\mu)=n\mu/\sigma_0^2$关于μ严格单增，故它是关于统计量$T(\mathbf{X})=\bar{X}$的单增MLR分布族，于是，本假设的水平α的UMP检验为

$$\phi(\mathbf{x})=\begin{cases}1, & T(\mathbf{x})>c,\\ 0, & T(\mathbf{x})<c,\end{cases}$$

其中c满足$E_{\theta_0}\phi(\mathbf{X})=\alpha$.

因为当$H_0:\mu=0$时，$\bar{X}\sim N(0,\sigma_0^2/n)$，故$c=u_\alpha\sigma_0/\sqrt{n}$. □

§5.3.4 双边假设的UMP检验

关于双边假设，我们在上一章的显著性检验中已讨论过许多，但那里的双边仅指假设

$$H_0:\theta=\theta_0\longleftrightarrow H_1:\theta\neq\theta_0. \tag{5.3.14}$$

但是，非常可惜的是，对于这样的双边假设，即使对于正态总体，其UMP检验也是不存在的. 事实上，对于来自$N(\mu,1)$的iid样本$X_1,\ldots,X_n$, 由上例可知，假设

$$H_0:\mu=0\longleftrightarrow H_1:\mu>0$$

的水平α的UMP检验为

$$\phi(\mathbf{x})=\begin{cases}1, & \bar{x}>u_\alpha/\sqrt{n},\\ 0, & \bar{x}<u_\alpha/\sqrt{n}.\end{cases}$$

而当$\mu<0$时，此检验的第二类错误概率为

$$\begin{aligned}\beta_\phi(\mu) &= P_\mu\left\{\sqrt{n}\bar{X}>u_\alpha\right\}\\ &= P_\mu\left\{\sqrt{n}(\bar{X}-\mu)>u_\alpha-\sqrt{n}\mu\right\}\\ &= 1-\Phi(u_\alpha-\sqrt{n}\mu)\longrightarrow 0,\quad \text{当}\mu\to-\infty\text{ 时}.\end{aligned}$$

由此可见，上述关于$H_0:\mu=0\longleftrightarrow H_1:\mu>0$的水平$\alpha$的UMP检验，并不是关于假设$H_0:\mu=0\longleftrightarrow H_1:\mu<0$的一个好的假设. 故对于形如(5.3.14)的双边假设的UMP检验并不存在. 我们同样也可以证明，即使对于MLR分布族，形如(5.3.14)式的UMP检验也不存在.

由于我们在一些实际问题中，感兴趣的双边假设可能形如

$$H_0:\theta\le\theta_1\text{ 或 }\theta\ge\theta_2\longleftrightarrow H_1:\theta_1<\theta<\theta_2, \tag{5.3.15}$$

故我们将在本小节考虑(5.3.15)的双边假设的UMP检验问题.

为求假设(5.3.15)的UMP检验，我们把N-P引理推广如下：

定理 5.3.5 设$f_0(x),f_1(x),\cdots,f_m(x)$是$m+1$个概率分布函数，$\alpha_1,\cdots,\alpha_m$为$m$个给定的常数. 如果存在检验函数$\phi(x)$满足以下条件:

(1) 如果有m个非负常数$k_1,\ldots,k_m$,使得

$$\phi(x)=\begin{cases}1, & f_0(x)>\sum_{i=1}^m k_if_i(x),\\ 0, & f_0(x)<\sum_{i=1}^m k_if_i(x),\end{cases}$$

(2) $\int\phi(x)f_i(x)d\mu(x)=\alpha_i,\ i=1,\cdots,m.$

则对任一满足下式的检验函数$\phi'(x)$,

$$\int \phi'(x)f_i(x)d\mu(x) \leq \alpha_i,\ i=1,\cdots,m, \tag{5.3.16}$$

都有

$$\int \phi(x)f_0(x)d\mu(x) \geq \int \phi'(x)f_0(x)d\mu(x).$$

证明 显然有

$$\left[f_0(x)-\sum_{i=1}^{m}k_if_i(x)\right][\phi(x)-\phi'(x)] \geq 0,$$

于是,

$$\begin{aligned} &\int f_0(x)[\phi(x)-\phi'(x)]\mathrm{d}x \\ \geq\ & \sum_{i=1}^{m}k_i\int f_i(x)[\phi(x)-\phi'(x)]\mathrm{d}x \\ =\ & \sum_{i=1}^{k}k_i\left[\alpha_i-\int\phi'(x)f_i(x)\mathrm{d}x\right] \geq 0. \end{aligned}$$

□

注 5.3.2 比较一下此定理与假设(5.3.15), 发现f_0对应着备选假设下的分布, 而f_i则对应于原假设下不同边界参数值时的分布. 此时本定理只告诉了满足条件(1)(2)的是UMPT, 但并没有说明其存在性及必要性. 事实上, 其存在性及必要性也是成立的, 请参见Lehmann(1986).

注 5.3.3 如果允许上一定理的k_i取负值, 则仅需改动(5.3.16)中的不等号为等号即可(这一结论将在讨论某些假设的UMP无偏检验时用到).

下面的定理将给出单参数指数型分布族假设(5.3.15)的UMPT.

定理 5.3.6 设$X_1,\ldots,X_n$是来自单参数指数型分布

$$f(x;\theta)=c(\theta)\exp\{Q(\theta)T(x)\}h(x)$$

的IID样本, 其中θ为一维实参数, $c(\theta)>0$, $Q(\theta)$是θ的严格单增函数. 则关于双边假设(5.3.15), 存在水平为$\alpha\in(0,1)$的UMPT, 其形式为

$$\phi(T(x))=\begin{cases} 1, & c_1<T(x)<c_2, \\ \delta_i, & T(x)=c_i, i=1,2, \\ 0, & T(x)<c_1 \text{ 或 } T(x)>c_2, \end{cases} \tag{5.3.17}$$

其中常数$c_i,\delta_i,\ i=1,2$由下式确定

$$E_{\theta_1}\phi(T(X))=E_{\theta_2}\phi(T(X))=\alpha. \tag{5.3.18}$$

证明 略(有兴趣的读者请参见茆诗松、王静龙、濮晓龙(2006)).

例 5.3.8 对于正态总体$N(\mu,1)$, 考虑检验问题$H_0:\mu\le\mu_1$ 或$\mu\ge\mu_2\longleftrightarrow H_1:\mu_1<\mu<\mu_2$的UMPT.

解 其样本联合PDF为

$$f(x;\mu)=(2\pi)^{-n/2}\exp\left\{-\sum(x_i-\mu)^2/2\right\}=(2\pi)^{-n/2}e^{-n\mu^2/2}\exp\left\{\mu\sum x_i\right\}e^{-\sum x_i^2/2},$$

它是关于统计量$T(x)=\bar{x}$单增的似然比分布族，且$Q(\mu)=n\mu$, 于是其UMPT为

$$\phi(T)=\begin{cases}1, & c_1\le T\le c_2,\\ 0, & \text{其它},\end{cases}$$

其中c_i满足$E_{\mu_1}\phi(T)=E_{\mu_2}\phi(T)=\alpha$.

由于当$\mu=\mu_i$时，$\bar{X}\sim N(\mu_i,1/n)$, 故c_i满足

$$\Phi(\sqrt{n}(c_2-\mu_1))-\Phi(\sqrt{n}(c_1-\mu_1))=\alpha,$$

$$\Phi(\sqrt{n}(c_2-\mu_2))-\Phi(\sqrt{n}(c_1-\mu_2))=\alpha.$$

□

上面我们讨论了关于假设(5.3.14)和(5.3.15)的UMP检验问题. 针对于双边假设(5.3.15)，自然有如下的一个双边假设

$$H_0:\theta_1\le\theta\le\theta_2\longleftrightarrow H_1:\theta<\theta_1 \text{ 或 } \theta>\theta_2 \tag{5.3.19}$$

的UMP检验问题. 然而，同假设(5.3.14)一样，假设(5.3.19)的UMP检验是不存在的. 事实上，由于UMPT要求的是在满足第一类错误概率小于某个指定的α的检验类中，其第二类错误概率达到最小, 即要求势函数$E_{H_1}\phi(X)$达到最大. 但由于第二类错误概率依赖于备选假设下的分布，且由N-P引理和定理5.3.4 知，UMPT的势函数是单调的（依赖于检验的方向），而假设(5.3.14)和(5.3.19)的备选假设的方向却是双向的，故我们找不到关于这类假设的UMPT. 既然两个双边假设(5.3.14)和(5.3.19)的UMP检验不存在，那是否存在其它的较好的检验吗？这就是我们将在下一节介绍的一致最优势无偏检验.

§5.4 无偏检验和一致最优势无偏检验

如果我们仍控制其第一类错误概率为α, 我们自然会要求其势函数$E_{H_1}\phi(X)\ge\alpha$, 这类的检验就被称为无偏检验(unbiased test).

如果一个检验的势函数在$\Theta_0\cap\Theta_1$相等，则称之为相似检验(similar test).

本节我们将主要考虑双边(5.3.14)和(5.3.19)的最优势无偏检验，为了叙述的方便，我们将上述两个假设重写如下：

$$H_0:\theta=\theta_0\longleftrightarrow H_1:\theta\ne\theta_0, \tag{5.4.1}$$

$$H_0: \theta_1 \le \theta \le \theta_2 \longleftrightarrow H_1: \theta < \theta_1 \text{ 或 } \theta > \theta_2. \tag{5.4.2}$$

§5.4.1 定义

假设我们感兴趣的假设为

$$H_0: \theta \in \Theta_0 \longleftrightarrow H_1: \theta \in \Theta_1.$$

定义 5.4.1 (无偏检验) 设$\phi(x)$是上述假设的一个检验，如果其势函数$E_\theta\phi(X)$满足条件

$$E_\theta\phi(X) \le \alpha,\ \forall\ \theta \in \Theta_0, \quad E_\theta\phi(X) \ge \alpha,\ \forall\ \theta \in \Theta_1,$$

则称之为水平α的无偏检验.

定义 5.4.2 设$\phi(x)$是上述假设的一个检验，如果其势函数$E_\theta\phi(X)$满足

$$E_\theta\phi(X) = \alpha,\ \forall\ \theta \in \Theta_0 \cap \Theta_1,$$

则称之为边界相似检验.

定义 5.4.3 对于上述假设，如果存在一个水平为α的无偏检验$\phi(x)$, 使得对于任何水平为α的无偏检验$\phi'(x)$, 均满足

$$E_\theta\phi(X) \ge E_\theta\phi'(X),\ \forall\ \theta \in \Theta_1,$$

则称检验$\phi(x)$是水平α的一致最优势无偏检验，记为UMPUT.

由上述定义，我们知道它们之间的关系为：

- 如果无偏检验$\phi(x)$的势函数$E_\theta\phi(X)$是连续的，则它一定是相似的;
- 水平为α的UMPT一定是水平为α的无偏检验;
- 如果上述检验问题的所有检验函数的势函数都是连续的，且一个水平为α的检验$\phi(x)$是最优势的相似检验，则它必是水平α的UMPUT;
- MLR分布族的单边假设的UMPT是一致最优势的边界相似检验（定理5.3.4的证明）.

由此可以看出水平为α的检验类Φ_α，水平为α的无偏检验类Φ_α^u, 水平为α的相似检验类Φ_α^s之间的关系为：$\Phi_\alpha^u \subset \Phi_\alpha$; 相似检验不一定是水平$\alpha$的，更不一定是无偏检验, 但当势函数连续时，$\Phi_\alpha^u \subset \Phi_\alpha^s$.

§5.4.2 一致最优势的无偏检验

在有了上一小节的概念之后，我们在本节不加证明地引入几个定理用来求取单参数指数型分布的双边假设的UMPUT.

定理 5.4.1 设$X_1, \ldots, X_n$是来自单参数指数型分布

$$f(x;\theta) = c(\theta)\exp\{Q(\theta)T(x)\}h(x)$$

的IID样本，其中θ为一维实参数，$c(\theta) > 0$, $Q(\theta)$是θ的严格单增函数. 则双边假设(5.4.2)

$$H_0 : \theta_1 \le \theta \le \theta_2 \longleftrightarrow H_1 : \theta < \theta_1 \text{ 或 } \theta > \theta_2,$$

存在水平为$\alpha \in (0,1)$的UMPUT，其检验函数为

$$\phi(T(x)) = \begin{cases} 1, & T(x) < c_1 \text{ 或 } T(x) > c_2, \\ \delta_i, & T(x) = c_i,\ i = 1, 2, \\ 0, & c_1 < T(x) < c_2, \end{cases} \tag{5.4.3}$$

其中常数$\delta_i, c_i,\ i = 1, 2$由下式决定

$$E_{\theta_1}\phi(T(X)) = E_{\theta_2}\phi(T(X)) = \alpha. \tag{5.4.4}$$

证明 请参见茆诗松、王静龙、濮晓龙(2006)P202–204. □

我们注意到此定理的结论，即关于双边假设(5.4.2) 的UMPUT 与双边假设(5.3.15) 的UMPT 非常相似，只是检验函数中的不等号变号而已.

定理 5.4.2 设$X_1, \ldots, X_n$是来自单参数指数型分布

$$f(x;\theta) = c(\theta)\exp\{Q(\theta)T(x)\}h(x)$$

的IID样本，其中θ为一维实参数，$c(\theta) > 0$, $Q(\theta)$是θ的严格单增函数. 则双边假设(5.4.1)

$$H_0 : \theta = \theta_0 \longleftrightarrow H_1 : \theta \ne \theta_0,$$

存在水平为$\alpha \in (0,1)$的UMPUT，其检验函数由(5.4.3)给出，其中常数$\delta_i, c_i,\ i = 1, 2$由下式决定

$$E_{\theta_0}\phi(T(X)) = \alpha,\ E_{\theta_0}[T(X)\phi(T(X))] = \alpha E_{\theta_0}T(X). \tag{5.4.5}$$

证明 参见茆诗松、王静龙、濮晓龙(2006)P204. □

比较定理5.4.1与定理5.4.2知，双边假设(5.4.1)与(5.4.2)的UMPUT的形式是一样的，只是其常数的确定方程不同.

上述几个定理的证明还告诉我们，对于定理中讨论的单参数指数型分布，

- 关于双边假设(5.4.1), (5.4.3)式给出的UMPU检验，就是在所有满足条件(5.4.5)的检验类中的一致最优势检验;
- 关于双边假设(5.4.2), 其UMPUT是一致最优势的边界相似检验;

- 关于双边假设(5.3.15), 其UMPT是一致最优势的边界相似检验.

当统计量$T(X)$的分布是对称时，我们有如下两点有利于计算的注解：

注 5.4.1 对于双边假设(5.4.1), 如果当$\theta=\theta_0$时，$T(X)$的分布关于某常数c对称，则其UMPUT的形式可以化简为

$$\phi(T(x))=\begin{cases}1, & T(x)<c-\delta \text{ 或 } T(x)>c+\delta,\\ \delta_0, & T(x)=c\pm\delta,\\ 0, & c-\delta<T(x)<c+\delta,\end{cases}\tag{5.4.6}$$

其中常数δ,δ_0满足

$$P_{\theta_0}\{T(X)<c-\delta\}+\delta_0 P_{\theta_0}\{T(X)=c-\delta\}=\alpha/2.$$

[其简单证明如下：下面我们始终假设H_0成立，即$\theta=\theta_0$. 此时

$$\begin{aligned}E_{\theta_0}\phi(T) &= P_{\theta_0}\{T<c-\delta\}+P_{\theta_0}\{T>c+\delta\}+\delta_0 P_{\theta_0}\{T=c\pm\delta\}\\ &= 2[P_{\theta_0}\{T(X)<c-\delta\}+\delta_0 P_{\theta_0}\{T(X)=c-\delta\}]=\alpha,\end{aligned}$$

又因为T关于c对称，故$E_{\theta_0}T=c, E_{\theta_0}[T-c]\phi(T)=0$,于是

$$E_{\theta_0}T\phi(T)=E_{\theta_0}[(T-c)\phi(T)]+cE_{\theta_0}\phi(T)=c\alpha=\alpha E_{\theta_0}T,$$

综合以上二式知，(5.4.5)式成立，故得证.]

注 5.4.2 当统计量$T(X)$的分布对称时，关于假设(5.4.2)，其UMPUT可以取为

$$\phi(T(x))=\begin{cases}1, & T(x)<\theta_1-\delta \text{ 或 } T(x)>\theta_2+\delta,\\ \delta_0, & T(x)=\theta_1-\delta \text{ 或 } \theta_2+\delta,\\ 0, & \theta_1-\delta<T(x)<\theta_2+\delta,\end{cases}\tag{5.4.7}$$

其中常数δ,δ_0满足

$$E_{\theta_1}\phi(T(X))=\alpha(\text{或 } E_{\theta_2}\phi(T(X))=\alpha).\tag{5.4.8}$$

注 5.4.3 当统计量$T(X)$的分布对称时，关于假设(5.3.15), 其UMPT可以取为

$$\phi(T(x))=\begin{cases}1, & \theta_1-\delta<T(x)<\theta_2+\delta,\\ \delta_0, & T(x)=\theta_1-\delta \text{ 或 } \theta_2+\delta,\\ 0, & T(x)<\theta_1-\delta \text{ 或 } T(x)>\theta_2+\delta,\end{cases}\tag{5.4.9}$$

其中常数δ, δ_0满足

$$E_{\theta_1}\phi(T(X)) = \alpha(\text{或 } E_{\theta_2}\phi(T(X)) = \alpha). \tag{5.4.10}$$

下面看几个例子.

例 5.4.1 设$X_1, \ldots, X_n$是来自正态总体$N(\mu, 1)$的IID样本, 考虑假设

(1) $H_0: \mu = 0 \longleftrightarrow H_1: \mu \neq 0$;

(2) $H_0: -1 \leq \mu \leq 1 \longleftrightarrow H_1: \mu < -1$ 或 $\mu > 1$;

(3) $H_0: \mu \leq -1$ 或 $\mu \geq 1 \longleftrightarrow H_0: -1 < \mu < 1$

的最优势检验.

解 容易验证此分布族是关于$T(x) = \bar{x}$单增的似然比分布族, 并且$T(X) \sim N(\mu, 1/n)$是连续型的.

(1) 当H_0成立时, $T(X) = \bar{X} \sim N(0, 1/n)$, 于是由上面的注5.4.1知，其水平为$\alpha$的UMPUT形如

$$\phi(T(x)) = \begin{cases} 1, & |\bar{x}| \geq \delta, \\ 0, & |\bar{x}| < \delta, \end{cases}$$

其中δ由下式确定

$$P_{H_0}\{\bar{X} < -\delta\} = \alpha/2,$$

故$\delta = u_{\alpha/2}/\sqrt{n}$.

(2) 关于此双边假设，由定理5.4.1知，其水平为α的UMPUT为

$$\phi(T(x)) = \begin{cases} 1, & \bar{x} < c_1 \text{ 或 } \bar{x} > c_2, \\ 0, & c_1 \leq \bar{x} \leq c_2, \end{cases}$$

其中c_i满足

$$E_{\mu=-1}\phi(\bar{X}) = E_{\mu=1}\phi(\bar{X}) = \alpha,$$

即

$$P_{\mu_1}\{c_1 \leq \bar{X} \leq c_2\} = 1 - \alpha = P_{\mu_2}\{c_1 \leq \bar{X} \leq c_2\},$$

也即

$$P_{\mu_1}\{\sqrt{n}(c_1+1) \leq \sqrt{n}(\bar{X}+1) \leq \sqrt{n}(c_2+1)\} = 1 - \alpha,$$
$$P_{\mu_2}\{\sqrt{n}(c_1-1) \leq \sqrt{n}(\bar{X}-1) \leq \sqrt{n}(c_2-1)\} = 1 - \alpha.$$

即

$$\int_{\sqrt{n}(c_1-1)}^{\sqrt{n}(c_2-1)} \phi(x)\mathrm{d}x = 1 - \alpha = \int_{\sqrt{n}(c_1+1)}^{\sqrt{n}(c_2+1)} \phi(x)\mathrm{d}x,$$

其中$\phi(x)$是标准正态分的PDF.

当然，也可以用注5.4.2中的(5.4.7)式来求取.

(3) 关于此双边假设，其UMPT的拒绝域为(注5.4.3)$\{x: -1-\delta \le \bar{x} \le 1+\delta\}$, 其中$\delta$满足$E_{\mu_1}\phi(T) = P_{\mu_1}\{-1-\delta \le \bar{X} \le 1+\delta\} = \alpha$, 即

$$P_{\mu_1}\{-\sqrt{n}\delta \le \sqrt{n}(\bar{X}+1) \le \sqrt{n}(2+\delta)\} = \Phi(\sqrt{n}(2+\delta)) - \Phi(-\sqrt{n}\delta) = \alpha.$$

由于$\delta > 0$, 故$\Phi(\sqrt{n}(2+\delta)) \approx 1$, 另外由于$\Phi(-\sqrt{n}\delta) = 1-\Phi(\sqrt{n}\delta)$, 故$\delta$近似地满足$\Phi(\sqrt{n}\delta) \approx \alpha$, 即$\delta \approx u_{1-\alpha}/\sqrt{n}$. □

此时的检验方法即是传统的显著性检验中的u检验法.

例 5.4.2 设$X_1, \ldots, X_n$是来自正态分布$N(0, \sigma^2)$的IID样本，考虑检验问题(1) $H_0: \sigma^2 \le \sigma_1^2$ 或 $\sigma^2 \ge \sigma_2^2 \longleftrightarrow H_1: \sigma_1^2 < \sigma^2 < \sigma_2^2$; (2) $H_0: \sigma_1^2 < \sigma^2 < \sigma_2^2 \longleftrightarrow H_1: \sigma^2 \le \sigma_1^2$ 或 $\sigma^2 \ge \sigma_2^2$; (3) $H_0: \sigma^2 = \sigma_0^2 \longleftrightarrow H_1: \sigma^2 \ne \sigma_0^2$的最优势检验.

解 样本的联合PDF为

$$p(x; \sigma^2) = (\sqrt{2\pi\sigma^2})^{-n/2} \exp\{-\sum x_i^2/2\sigma^2\},$$

$Q(\sigma^2) = 1/2\sigma^2$是$\sigma^2$的严格增函数，故它是关于统计量$T(X) = \sum x_i^2$的单增似然比. 并且注意到$T(X)/\sigma^2 \sim \chi^2(n)$是一个连续变量.

对于检验问题(1)，由定理5.3.6知，其水平为α的UMPT为

$$\phi(T) = \begin{cases} 1, & c_1 \le T(x) \le c_2, \\ 0, & \text{其它}, \end{cases}$$

其中常数c_i满足

$$P_{\sigma_1^2}\{c_1 \le T(X) \le c_2\} = P_{\sigma_2^2}\{c_1 \le T(X) \le c_2\} = \alpha,$$

即

$$P_{\sigma_1^2}\{c_1/\sigma_1^2 \le T(X)/\sigma_1^2 \le c_2/\sigma_1^2\} = P_{\sigma_2^2}\{c_1/\sigma_2^2 \le T(X)/\sigma_2^2 \le c_2/\sigma_2^2\} = \alpha,$$

于是c_i由下式决定

$$\int_{c_1/\sigma_1^2}^{c_2/\sigma_1^2} \chi^2(x; n)\mathrm{d}x = \alpha = \int_{c_2/\sigma_2^2}^{c_2/\sigma_2^2} \chi^2(x; n)\mathrm{d}x.$$

对于检验问题(2), 由定理5.4.1知，其水平为α的UMPUT为

$$\phi(t) = \begin{cases} 1, & t \le c_1 \text{ 或 } t \ge c_2, \\ 0, & \text{否则}, \end{cases}$$

其中c_i满足

$$P_{\sigma_1^2}\{T(X)\le c_1\}=P_{\sigma_2^2}\{T(X)\ge c_2\}=\alpha,$$

即

$$\int_{c_1/\sigma_1^2}^{c_2/\sigma_1^2}\chi^2(x;n)\mathrm{d}x=1-\alpha=\int_{c_2/\sigma_2^2}^{c_2/\sigma_2^2}\chi^2(x;n)\mathrm{d}x.$$

对于检验问题(3), 由定理5.4.2知，其水平为α的UMPUT为

$$\phi(t)=\begin{cases}1, & t\le c_1 \text{ 或 } t\ge c_2,\\ 0, & \text{否则},\end{cases}$$

其中常数c_i由

$$E_{\sigma_0^2}\phi(T(X))=\alpha,\quad E_{\sigma_0^2}[T(X)\phi(T(X))]=\alpha E_{\sigma_0^2}T(X), \tag{$*$}$$

确定. 注意到当H_0成立时，$T(X)/\sigma_0^2\sim\chi^2(n)$，于是$E_{\sigma_0^2}T(X)=n\sigma_0^2$. 于是，(*)式的第一部分为

$$\int_{c_1/\sigma_0^2}^{c_2/\sigma_0^2}\chi^2(x;n)\mathrm{d}x=1-\alpha, \tag{$*1$}$$

(*)式的第二部分是

$$\begin{aligned}&E_{\sigma_0^2}[T(X)\phi(T(X))]=\sigma_0^2E_{\sigma_0^2}[T(X)/\sigma_0^2\phi(T(X))]\\ =\ &\sigma_0^2\left[\int_{-\infty}^{c_1/\sigma_0^2}x\chi^2(x;n)\mathrm{d}x+\int_{c_2/\sigma_0^2}^{\infty}x\chi^2(x;n)\mathrm{d}x\right]=\sigma_0^2\left[n-\int_{c_1/\sigma_0^2}^{c_2/\sigma_0^2}x\chi^2(x;n)\mathrm{d}x\right]\\ =\ &\alpha E_{\sigma_0^2}T(X)=n\alpha\sigma_0^2,\end{aligned}$$

即

$$\int_{c_1/\sigma_0^2}^{c_2/\sigma_0^2}x\chi^2(x;n)\mathrm{d}x=n(1-\alpha). \tag{$*2$}$$

这样，常数c_i就由(*1)和(*2)两式确定. □

本例得到的最优势检验即为显著性检验中均值已知时的常用的χ^2检验法.

通过例5.4.2，我们发现，有时最优势检验的拒绝域并没有一个显示解(本例的三个检验问题均没有)，故在某些情况下我们均采用近似的方法来确定. 例如对于例5.4.2中的第三个检验问题，我们可近似地取$c_1=\sigma_0^2\chi^2_{1-\alpha/2}(n),c_2=\sigma_0^2\chi^2_{\alpha/2}(n)$, 此时它肯定能满足(*1)式，但不一定能满足(*2)式.

§5.5 序贯概率比检验

通过前面的内容，我们可以了解到，统计学的基本任务就是通过样本来推断或掌握总体的特性，并且，我们始终假设样本容量是给定的(除例5.3.5之外). 然而，我们知道，样本量越大，由于我们对总体的了解就可能越多，故对总体作出的推断也就越可靠，当然其抽样所需的费用也可能越大. 于是，一个合理的提法是：在保证所得结论有足够可靠性的前提下，样本量应越小越好. 另外，对于某些实际问题，使用固定的样本容量也是没有必要的. 比如，对于例5.3.5中的(n,c)检验方案，如果在未抽到n件时就已经有$c+1$件次品了，则我们就没有必要再继续往下抽样. 这样，在上世纪40年代，人们普遍认识到样本容量不必事先给定，而可以根据抽取的样本来决定何时停止抽样，也就是说样本容量是一个随机变量，这样得到的样本是一个一个地依次得到的，我们称之为序贯样本(Sequential Sample). 序贯分析正是研究如何得到和利用序贯样本进行统计推断的统计学分支(有兴趣的读者请参见陈家鼎(1995)). 最基本的序贯分析方法是序贯分析奠基人A. Wald 于1943至1945年提出的序贯概率比检验(Sequential Probability Ratio Test, 简记为SPRT). 它是为适应第二次世界大战期间美国军火生产中质量检验工作的需要而提出的.

在许多实际问题中，有时我们感兴趣的是想依赖于样本判断总体的分布是$f_1(x)$还是$f_2(x)$，即此时我们感兴趣的假设为

$$H_1: X_1, X_2, \ldots \overset{\text{iid}}{\sim} f_1(x) \longleftrightarrow H_2: X_1, X_2, \ldots \overset{\text{iid}}{\sim} f_2(x). \tag{5.5.1}$$

为检验上述假设，我们可以利用如下的似然比统计量

$$\lambda_n = \frac{\prod_{i=1}^n f_2(x_i)}{\prod_{i=1}^n f_1(x_i)} \tag{5.5.2}$$

进行检验. 前面我们讲过的似然比检验的法则是：对于固定的n，当$\lambda_n \geq c$时，拒绝零假设；否则就接受零假设. 而Wald对此进行了重大改进，从而提出SPRT.

定义 5.5.1 (SPRT) 对于事先给定的两个常数$0 < A < 1 < B < \infty$，我们一个一个地抽取样本，如在抽得$X_1, \ldots, X_{n-1}$后，抽样尚不能停止，则再抽X_n，并计算λ_n. 当$\lambda_n \geq B$时，停止抽样，并拒绝假设H_1；当$\lambda_n \leq A$时，停止抽样，并接受假设H_1；当$A < \lambda_n < B$时，抽取第$n+1$个样本，并计算λ_{n+1}, 照此依次进行. 我们称这样的检验为一个序贯概率比检验，并记为$S(A,B)$. 事实上，其停止抽样时间(简称停时)可以写成

$$\tau^* = \min\{n : \lambda_n \leq A, \text{或}\lambda_n \geq B, n \geq 1\}. \tag{5.5.3}$$

从上述SPRT的定义不难看出，SPRT的关键之处即在于其判决准则. 下面看一个例子.

例 5.5.1 考虑Bernoulli分布$b(1,p)$的关于如下假设

$$H_1: p = p_1 \longleftrightarrow H_2: p = p_2 \tag{5.5.4}$$

的SPRT, 其中$0 < p_1 < p_2 < 1$为二已知常数.

解 因为此时总体的概率分布为

$$f(x,p) = p^x(1-p)^{1-x},\ x = 0, 1,$$

则似然比统计量为

$$\lambda_n = \left(\frac{p_2}{p_1}\right)^{\sum_{i=1}^n x_i}\left(\frac{1-p_2}{1-p_1}\right)^{n-\sum_{i=1}^n x_i},\ n = 1, 2, \cdots,$$

于是，

$$\ln\lambda_n = \sum_{i=1}^n x_i \ln\frac{p_2(1-p_1)}{p_1(1-p_2)} + n\ln\frac{1-p_2}{1-p_1}.$$

为方便，记$T_n = \sum_{i=1}^n x_i$, 且

$$c = -\frac{\ln\frac{1-p_2}{1-p_1}}{\ln\frac{p_2(1-p_1)}{p_1(1-p_2)}},\ d_1 = \frac{\ln B}{\ln\frac{p_2(1-p_1)}{p_1(1-p_2)}},\ d_2 = \frac{\ln A}{\ln\frac{p_2(1-p_1)}{p_1(1-p_2)}}. \tag{5.5.5}$$

于是，对于事先给定的A, B，我们有

$$\lambda_n \geq B \Longleftrightarrow T_n \geq R_n = cn + d_1,$$

$$\lambda_n \leq A \Longleftrightarrow T_n \leq A_n = cn + d_2.$$

则综合上述知，对于两点分布，假设(5.5.4)的SPRT的停时为

$$\tau^* = \min\{n : T_n \leq A_n, \text{或} T_n \geq R_n, n \geq 1\},$$

其图示为图5.2. □

从上例可以看出，SPRT的应用并不复杂，由于它仅依赖于事先给定的两个常数A, B. 那A, B又该如何选择呢？这就与SPRT的性质有关. 关于SPRT，主要有如下几个问题需要研究：

- 是否有限步后一定会停止抽样？即是否成立$P_i\{\tau^* < \infty\} = 1\ (i = 1, 2)$?
- 如何计算其两类错误概率

$$\alpha = P_1\{\tau^* < \infty, \lambda_{\tau^*} \geq B\},\ \beta = P_2\{\tau^* < \infty, \lambda_{\tau^*} \leq A\}. \tag{5.5.6}$$

我们希望找出α, β与常数A, B间的关系，从而根据给定的两类错误概率α, β确定常数A, B.

- 如何计算平均样本容量$E_i\tau^*(\ i = 1, 2)$.
- SPRT有何优良性质？是否存在比SPRT还“好”的检验方法？

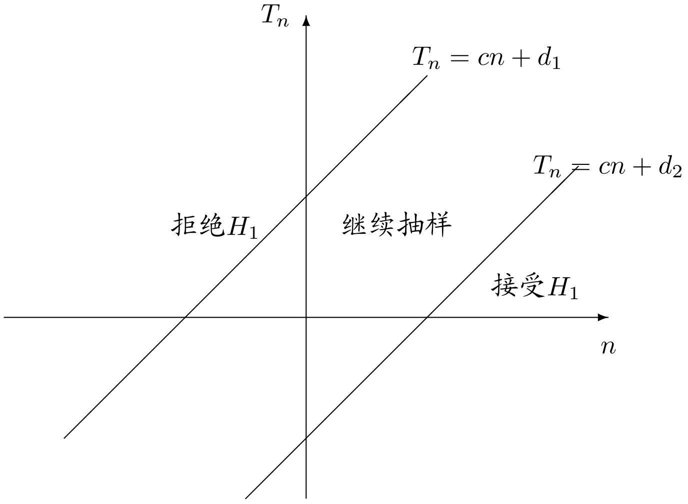

图5.2 例5.5.1的判决准则

下面，我们不加证明地给出一些关于上述问题的相关结论，详细的证明及其它性质请见陈家鼎(1995).

定理 5.5.1 对于一个SPRT $S(A,B)$, 我们有

$$P_i\{\tau^* < \infty\} = 1, \quad E_i e^{\lambda\tau^*} < \infty, \quad E_i(\tau^*)^k < \infty, \tag{5.5.7}$$

其中$i = 1, 2, k \geq 1$为任意的, $\lambda > 0$为给定的常数.

定理 5.5.2 对于SPRT $S(A,B)$, 其临界值与两类错误间的关系为

$$\alpha \leq \frac{1-\beta}{B}, \quad \beta \leq A(1-\alpha). \tag{5.5.8}$$

从(5.5.8)式可知，$\frac{\beta}{1-\alpha} \leq A < B \leq \frac{1-\beta}{\alpha}$. 于是，一个常用的近似公式为

$$A \doteq \frac{\beta}{1-\alpha}, \quad B \doteq \frac{1-\beta}{\alpha}. \tag{5.5.9}$$

从(5.5.9)式可知，对于事先给定的第一二类错误概率，我们可以很容易地由(5.5.9)式确定SPRT的临界值，而不需要了解总体的分布类型. 当然，上面的临界值是近似的，并且下一定理告诉我们两个SPRT $S(A,B)$与$S(\frac{\beta}{1-\alpha}, \frac{1-\beta}{\alpha})$间的关系.

定理 5.5.3 对于事先给定的$\alpha > 0, \beta > 0$ $(\alpha + \beta < 1)$, 设α^*, β^*为SPRT $S(\frac{\beta}{1-\alpha}, \frac{1-\beta}{\alpha})$的第一二类错误概率，则有

$$\alpha^* \leq \frac{\alpha}{1-\beta}, \quad \beta^* \leq \frac{\beta}{1-\alpha}, \quad \alpha^* + \beta^* \leq \alpha + \beta.$$

定理5.5.3说明采用(5.5.9)式临界值的SPRT的两类错误概率之和不会增大.

习题五

1. 设X_1, X_2, X_3为来自两点分布$b(1,p)$的iid样本，对于假设

$$H_0: p = \frac{1}{2} \longleftrightarrow H_1: p = \frac{3}{4},$$

如取一个检验的拒绝域为

$$W = \{(x_1, x_2, x_3): x_1 + x_2 + x_3 \geq 2\},$$

则求此检验的第一、二类错误概率及$p = 3/4$时的势.

2. 设总体X的密度函数可取下面的$f_0(x)$或$f_1(x)$:

$$f_0(x) = \begin{cases} 1, & 0 \leq x \leq 1, \\ 0, & \text{其它}, \end{cases} \qquad f_1(x) = \begin{cases} 2x, & 0 \leq x \leq 1, \\ 0, & \text{其它}, \end{cases}$$

现基于一个样本X_1，考虑如下假设

$$H_0: f(x) = f_0(x) \longleftrightarrow H_1: f(x) = f_1(x),$$

试在水平0.1下，求出使第二类错误概率最小的检验法.

3. 设$X_1, \ldots, X_n$为来自Poisson分布$P(\lambda)$的iid样本，试求假设

$$H_0: \lambda \leq \lambda_0 \longleftrightarrow H_1: \lambda > \lambda_0 \ (\lambda_0\text{已知})$$

的水平为α的UMP检验.

4. 设$X_1, \ldots, X_n$为来自均匀分布$U(0,\theta)$的iid样本，试求假设

$$H_0: \theta \leq \theta_0 \longleftrightarrow H_1: \theta > \theta_0 \ (\theta_0 > 0\text{已知})$$

的水平α的UMP检验.

5. 设$X_1, \ldots, X_n$为来自正态总体$N(\mu, 1)$的iid样本，对于假设

$$H_0: \mu \leq 0 \longleftrightarrow H_1: \mu > 0,$$

(1) 求水平为0.025的UMP检验，并求出其势函数$\beta(\mu)$;

(2) 为了使上述检验在$\mu \geq 0.5$时的势函数$\beta(\mu) \leq 0.9$，样本容量n应至少取多大?

6. 现从某正态总体中随机抽取了9个iid样本，测得其均值为$\bar{X} = 0.4$，样本标准差为$S_n = 1.0$. 在水平0.05和0.01下分别检验假设

(1) $H_0: \mu \leq 0 \longleftrightarrow H_1: \mu > 0$,

(2) $H_0: \mu \geq 0 \longleftrightarrow H_1: \mu < 0$,

并解释检验结果的意义. 如果样本容量增加到25，而样本均值与标准差不变，则再检验上述假设，并裹脚结果的意义.

7. 设$X_1,\ldots,X_n$为来自均匀分布$U(0,\theta)$的iid样本，其中$\theta>0$为参数. 试求假设

(1) $H_0:\theta=1\longleftrightarrow H_1:\theta=2$,

(2) $H_0:\theta=1\longleftrightarrow H_1:\theta=1/2$

的水平为α的MP检验.

8. 设$\alpha\in(0,1)$，且H_0,H_1是两个简单假设. 则ϕ是假设$H_0\longleftrightarrow H_1$的水平$\alpha$ 的MP检验，且$\beta=E_{H_1}\phi(X)<1$，则$1-\phi$是假设$H_1\longleftrightarrow H_0$的水平为$1-\beta$的MP检验.

9. 设$X_1,\ldots,X_n$为来自两点分布$b(1,p)$的iid样本，试求假设

$$H_0:p=1/2\longleftrightarrow H_1:p<1/2$$

的水平α的UMP检验.

10. 设$X_1,\ldots,X_n$为来自Γ分布$\Gamma(\theta,1)$的iid样本，试求假设

$$H_0:p=p_0\longleftrightarrow H_1:p>p_0$$

的水平α的UMP检验.

11. 设$X_1,\ldots,X_n$为来自两点分布$b(1,p)$的iid样本，试求假设

$$H_0:p=p_0\longleftrightarrow H_1:p\neq p_0$$

的水平α的UMPU检验.

12. 设$X_1,\ldots,X_n$是n个来自正态总体的相互独立的样本，且

$$EX_i=\begin{cases}\mu_i, & i=1,\cdots,m,\\ \mu_0, & i=m+1,\cdots,n,\end{cases}\qquad \mathrm{Var}X_i=\sigma^2,\ i=1,\cdots,n.$$

则关于假设

(1) $H_0:\mu_1\le\mu_0\longleftrightarrow H_1:\mu_1\ge\mu_0$;

(2) $H_0:\mu_1=\mu_0\longleftrightarrow H_1:\mu_1\neq\mu_0$,

存在水平α的UMPU检验，且分别为

$$\phi_1(x)=\begin{cases}1, & T(x)<c_1,\\ 0, & 其它,\end{cases}\qquad \phi_2(x)=\begin{cases}1, & |T(x)|>c_2,\\ 0, & 其它,\end{cases}$$

其中

$$T(x)=\frac{\sqrt{n-m}(x_1-\mu_0}{\sqrt{\sum_{i=m+1}^n x_i^2}},$$

且当$\mu_1=\mu_0$时，$T(X)\sim t(n-m)$.

13. 设甲乙各有资本M,N元(M,N均为正整数)，且两人赌博时没有平局. 另设每一局若甲胜，则乙给甲元；若乙胜，则甲给乙一元，且设每局中甲胜的概率为$p(0<p<1)$. 问：如果一局一局地赌下去，甲输光的概率是多少？平均几局后有一方输光？

14. 设$X_1,\ldots,X_n$为来自密度函数为

$$f(x,\sigma)=\begin{cases}\frac{\sigma^3}{\sqrt{2\pi}}\sqrt{x}\exp\{-\sigma^2x/2\}, & x>0,\\ 0, & x\leq 0.\end{cases}$$

的总体的iid样本，其中$\sigma>0$为未知参数. 对于给定的$0<\sigma_0<\sigma_1$，试求假设

$$H_0:\sigma=\sigma_0\longleftrightarrow H_1:\sigma=\sigma_1$$

的序贯概率比检验.

第6章　几个常用的分布检验方法

到目前为止，我们所讨论的估计和检验问题，均假设总体分布属于某个特定的类型(如正态分布)，而对其未知参数或某些数字特征(如期望与方差)进行的. 一个自然的问题是，我们怎么知道总体的分布属于某种类型呢？在许多情况下，我们均是利用样本来判断总体的分布，这就是所谓的分布拟合问题.

K. Pearson 于1900的发表的关于χ^2拟合优度(goodness-of-fit)检验的论文，常被学者们认为是近代数理统计学的发端. 这篇论文对统计学应用中一个常见的重要问题，即一组随机观察数据可否合理地看成是来自一个分布完全已知的总体，提出了一个判定准则. 自此之后，有关拟合优度检验得到了极大的发展，其中一个很有代表性的结果就是Kolmogorov于上世纪三十年代提出的基于经验分布的检验准则. 有关拟合优度检验的详细内容，请参见杨振海(1994).

本章主要介绍某些常见的拟合优度检验，如χ^2拟合优度检验，Kolmogorov检验及几个正态性检验方法.

§6.1　正态概率纸检验法

在计算机不是很发达的时代，利用正态概率纸来检验一组样本是否来自正态这一方法得到了广泛的应用，尤其是在许多实际应用部门. 但随着计算机的发展，这一方法逐渐走出了人们的视线，取而代之的是各个统计软件中的p-p图，q-q图，半正态图等检验方法. 但这些方法的构造都利用了正态检验纸思想，故本节将对正态检验纸检验法作一简介.

设有n个来自分布$F(x)$的观测值$X_1, \ldots, X_n$, 根据以往经验或其它方法，我们对总体分布有一定认识，此时我们感兴趣的为如下的假设

$$H_0 : F(x) \equiv F_0(x), \tag{6.1.1}$$

其中$F_0(x)$为一形式完全已知的分布，但可能包含有未知参数. 现常用的概率纸检验方法有：正态概率纸、Weibull概率纸和对数正态概率纸等，即(6.1.1)中的F_0为这些分布的CDF. 本节将只讲正态概率纸的构造及其应用，其余的构造方法完全类似.

§6.1.1　正态概率纸的构造

所谓正态概率纸(normal probability paper)就是一种具有特殊刻度的坐标纸(见图6.1). 由图6.1可见，正态概率纸是一个网格图，其特点是在横坐标上的划分是等间距的，而关于纵坐标的划分则是不等间距的.

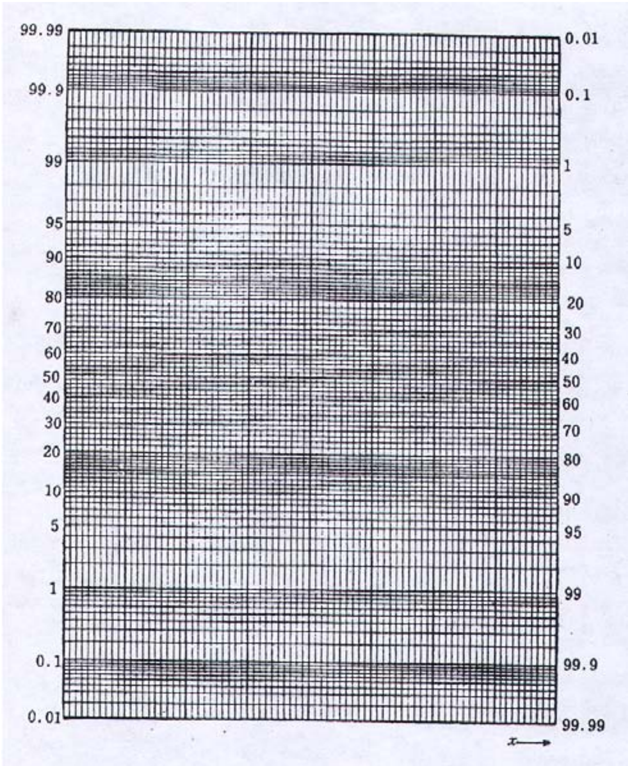

图6.1 正态概率纸

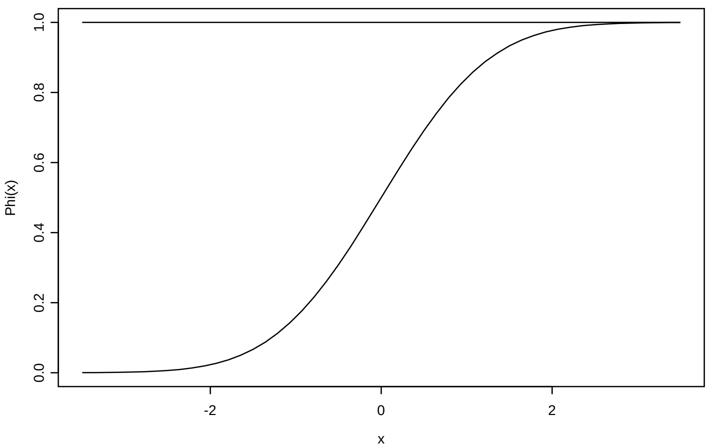

图6.2 标准正态分布函数曲线

我们知道，正态分布的累积分布函数同其它随机变量的CDF一样，是一条介于0与1之间的单增曲线(见图6.2)，虽然不同分布之间的趋势有所不同，但肉眼很难分辨它们之间的区别. 如果能将这些曲线“拉直”，则我们可以很容易地区分它们.

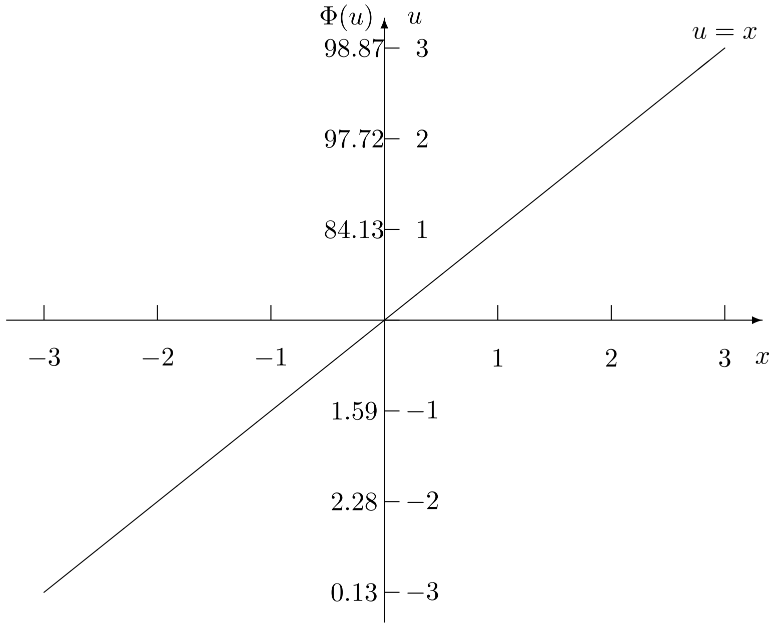

图6.3 变换刻度后的$(x,\Phi(x))$图

一般的正态分布函数

$$y=\Phi\left(\frac{x-\mu}{\sigma}\right) \tag{6.1.2}$$

在我们熟知的直角坐标系的图形如图6.2所示.

显然，在图6.2中y关于x是一曲线，但如果我们做如下变换：

$$\begin{cases} y=\Phi(u), \\ u=\frac{x-\mu}{\sigma}, \end{cases}$$

则知u关于x是线性的，而y与u是一一对应的. 于是，我们可以在(x,u)坐标系上，不改变横坐标x的刻度，而在u轴(即纵坐标)上标上相应$\Phi(u)$的刻度，则得到了u与$\Phi(u)$的一条直线了. 当然，由于$\Phi(u)$是非线性的，故此时纵坐标的刻度是非等距的. 如果$\mu=0,\sigma=1$，则此时的$(x,\Phi(x))$的图形就如图6.3所示的一条直线，只是纵坐标的刻度不均匀而已. 这就是一张正态概率纸.

我们注意到

$$\lim_{u\to-\infty}\Phi(u)=0,\quad \lim_{u\to\infty}\Phi(u)=1,$$

故纵轴刻度为0和1的点在概率纸上无法标出. 所以，大家所用的正态概率纸的纵轴刻度均是从0.01到99.99. 反之，如在正态概率纸上有一条上升的直线，则它就是某正态分布的分布函数的图形.

§6.1.2 正态概率纸的应用

当样本$X_1,\ldots,X_n$来自分布函数$F(x)$时，如果n个点$(X_i,F(X_i))$或$(X_{(i)},F(X_{(i)}))$在正态概

率纸上呈现一条直线时，我们就有充分理由认为此组样本来自正态分布. 但现在的问题是：由于总体分布$F(x)$是未知的，故我们无法将这些点画出来. 但是，我们注意到，如果$\xi \sim F(x)$，且F连续，则$F(\xi) \sim U(0,1)$. 于是，我们有

$$E(F(X_i)) = \frac{i}{n+1},$$

这样，我们可以用$\frac{i}{n+1}$来估计$F(X_{(i)})$的值，且在正态概率纸上画如下点：

$$\left(X_{(i)}, \frac{i}{n+1}\right), \ i = 1, 2, \cdots, n$$

当这n个点在正态概率纸上呈现一条直线时，我们就有理由认为这些数据服从正态分布.

根据这些想法，我们应用正态概率纸检验正态假设的具体步骤如下：

- 把这n个数据$X_1, \ldots, X_n$由小到大排列如下：$X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$.
- 在正态概率纸上画上n个点$\left(X_{(i)}, \frac{i}{n+1}\right)$，$i = 1, 2, \cdots, n$.
- 如果这n个点呈现一条直线，则接受正态性假设.

如果检验之后，有理由认为这组数据服从正态分布时，一个自然的问题就是其两个参数的估计值如何？从正态概率纸的构造来看，当$y = \Phi(u) = 0.5$时，$u = (x-\mu)/\sigma$. 故对应着纵坐标上取值为50%的点的横坐标即为总体均值μ的估计. 类似地，当$u = 1$时，$y = 84.13\%$，故直线上对应着$y = 84.13$的点的横坐标就是$\mu + \sigma$的估计值. 于是，就可以得到μ与σ的估计了. 当然，由于直线的粗糙性，故这种估计也是非常初步的，且无法判断它的优良性.

另外，由于直线的判断也因人而易，故上述正态概率纸检验也是非常粗糙的，且我们无法控制其第一、二类错误概率. 但是在计算机不发达的年代，这种方法由于便于应用，故它在许多实际部门得到了广泛的应用. 然而，随着计算机的高速发展，这种不严格的正态检验纸检验方法已逐步走出了人们的视线. 有关正态性检验的一些精确方法，我们将在本章后面给以简单介绍.

§6.2 Pearson χ^2拟合优度检验

我们注意到，这里要讲述的Pearson χ^2拟合优度检验中的Pearson并不前面的N-P引理中的E. Pearson，他是K. Pearson (1857–1936), E. Pearson的父亲.

K. Pearson提出的χ^2拟合优度检验是针对分类数据的检验而提出的.

§6.2.1 分类数据的χ^2拟合优度检验

我们先看一个在生物中很有名的例子.

例 6.2.1 在19世纪，Mendel按颜色与形状把豌豆分为四类：黄园、青园、黄皱和青皱.

Mendel在他种的$n = 556$个豌豆中，观测到这四类豌豆的个数分别为315, 108, 101, 32. 于是，Mendel判断这四类的比例为9:3:3:1. 那Mendel的这种比例正确吗？

上一例子是关于分类数据的检验问题，它的一般情形为：根据某项指标，总体被分成r类：$A_1, \ldots, A_r$. 此时我们最关心的是关于各类所占的比例的假设

$$H_0: 第i\ 类A_i\ 所占的比例为p_i,\ i = 1, \cdots, r, \tag{6.2.1}$$

其中$\sum_{i=1}^r p_i = 1$. 记$X_1, \ldots, X_n$为从此总体抽出的n个iid样本，且以n_i记这n个样本中属于A_i的样本个数. 由于当H_0成立时，在n个样本中属于A_i类的“理论个数”或“期望个数”为np_i，而我们实际观测到的值为n_i，故，当H_0成立时，n_i与np_i应相差不大. 于是，K. Pearson提出用统计量

$$\chi^2 = \sum_{i=1}^r \frac{(n_i - np_i)^2}{np_i} \tag{6.2.2}$$

来衡量“理论个数”与实际观测值间的差别，并且其拒绝域为$\{\chi^2 \geq c\}$.

我们回忆一下(4.5.5)式，与这里的统计量(6.2.2)有着相同的表达式. 由此可见Pearson提出(6.2.2)的原因. 为了控制上述检验的第一类错误，我们必须知道此检验统计量的零分布，为此，K. Pearson证明了如下定理.

定理 6.2.1 *在H_0成立且p_i均已知时，我们有*

$$\chi^2 \xrightarrow{\mathcal{L}} \chi^2(r-1). \tag{6.2.3}$$

证明 当H_0成立时，由于$(n_1, \ldots, n_r)$服从如下的多项分布:

$$\frac{n!}{n_1! \cdots n_r!} p_1^{n_1} \cdots p_r^{n_r},$$

于是，当$n = 2$时，$n_1 \sim B(n, p_1)$, 且

$$\chi^2 = \frac{(n_1 - np_1)^2}{np_1} + \frac{(n_2 - np_2)^2}{np_2} = \frac{(n_1 - np_1)^2}{np_1 p_2}.$$

而由中心极限定理可知,

$$\frac{n_1 - np_1}{\sqrt{np_1 p_2}} \xrightarrow{\mathcal{L}} N(0, 1),$$

故$\chi^2 \xrightarrow{\mathcal{L}} \chi^2(1)$, 即(6.2.3)式成立. 下证$r > 2$的情形.

当$r > 2$时，我们令

$$Y_i = \frac{n_i - np_i}{\sqrt{np_i}},\ i = 1, 2, \cdots, r,$$

于是$\chi^2 = \sum_{i=1}^r Y_i^2$.

因为$\sum n_i = n, \sum p_i = 1$, 故$\sum_{i=1}^r \sqrt{p_i} Y_i = 0$，即$Y_i$间并不独立.

由于$(n_1,\ldots,n_r)$的特征函数为

$$\begin{aligned}\psi(t_1,\ldots,t_r) &= E\exp\{it_1n_1+\cdots+it_rn_r\}\\ &= \sum_{n_1+\cdots n_r=n}\frac{n!}{n_1!\cdots n_r!}\left(p_1e^{it_1}\right)^{n_1}\cdots\left(p_re^{it_r}\right)^{n_r}\\ &= \left(\sum_{k=1}^r p_ke^{it_k}\right)^n,\end{aligned}$$

所以，$(Y_1,\ldots,Y_r)$的特征函数为

$$\begin{aligned}\psi_y(t_1,\ldots,t_r) &= E\exp\{it_1Y_1+\cdots+it_rY_r\}\\ &= E\exp\left\{i\sum_{k=1}^r\frac{t_k(n_k-np_k)}{\sqrt{np_k}}\right\}\\ &= \exp\left\{-i\sum_{k=1}^r t_k\sqrt{np_k}\right\}\left[\sum_{k=1}^r p_k\exp\left\{\frac{it_k}{\sqrt{np_k}}\right\}\right]^n,\end{aligned}$$

两边取对数后得到

$$\ln\psi_y(t_1,\ldots,t_r)=-i\sum_{k=1}^r t_k\sqrt{np_k}+n\ln\left(\sum_{k=1}^r p_k\exp\left\{\frac{it_k}{\sqrt{np_k}}\right\}\right).$$

由Taylor展开知，

$$e^x=1+x+\frac{x^2}{2}+o(x^2),$$

$$\ln(1+x)=x-\frac{x^2}{2}+o(x^2),$$

故对于充分大的n，我们有

$$\exp\left\{\frac{it_k}{\sqrt{np_k}}\right\}=1+\frac{it_k}{\sqrt{np_k}}-\frac{t_k^2}{2np_k}+o(n^{-1}),$$

$$\sum_{k=1}^r p_k\exp\left\{\frac{it_k}{\sqrt{np_k}}\right\}=1+i\sum_{k=1}^r\frac{t_k\sqrt{p_k}}{\sqrt{n}}-\frac{1}{2n}\sum_{k=1}^r t_k^2+o(n^{-1}),$$

$$\ln\left(\sum_{k=1}^r p_k\exp\left\{\frac{it_k}{\sqrt{np_k}}\right\}\right)=i\sum_{k=1}^r\frac{t_k\sqrt{p_k}}{\sqrt{n}}-\frac{1}{2n}\sum_{k=1}^r t_k^2-\frac{1}{2}\left(i\sum_{k=1}^r\frac{t_k\sqrt{p_k}}{\sqrt{n}}\right)^2+o(n^{-1}),$$

于是

$$\ln\psi_y(t_1,\ldots,t_k)=-\frac{1}{2}\left[\sum_{k=1}^r t_k^2-\left(\sum_{k=1}^r t_k\sqrt{p_k}\right)^2\right]+o(1).$$

作如下的正交变换：

$$\begin{cases} Z_k = \sum_{j=1}^r a_{kj} Y_j,\ k=1,2,\cdots,r-1, \\ Z_r = \sum_{j=1}^r \sqrt{p_j} Y_j. \end{cases}$$

如再令

$$\begin{cases} u_k = \sum_{j=1}^r a_{kj} t_j,\ k=1,2,\cdots,r-1, \\ u_r = \sum_{j=1}^r \sqrt{p_j} t_j, \end{cases}$$

则知

$$\sum_{k=1}^r t_j^2 - \left(\sum_{j=1}^r t_j \sqrt{p_j}\right)^2 = \sum_{k=1}^{r-1} u_k^2,$$

于是，当$n\to\infty$时，$(Z_1,\ldots,Z_r)$的特征函数为

$$\lim_{n\to\infty} \psi_z(u_1,\ldots,u_r) = \exp\left\{-\frac{1}{2}\sum_{k=1}^{r-1} u_k^2\right\},$$

这是$r-1$个独立的标准正态随机向量的特征函数. 由此可以看出，$Z_1,\ldots,Z_{r-1}$是$r-1$个独立的标准正态随机变量，Z_r依概率收敛于0. 于是可得

$$\chi^2 = \sum_{k=1}^r Y_k^2 = \sum_{k=1}^r Z_k^2 \xrightarrow{\mathcal{L}} \chi^2(r-1).$$

□

于是，对于假设(6.2.1)，我们可以采取如下的水平近似为α的显著性检验：

- 检验统计量如(6.2.2)所示，即

$$\chi^2 = \sum_{i=1}^r \frac{(n_i - np_i)^2}{np_i},$$

- 拒绝域为

$$W = \{\chi^2 \geq \chi^2_\alpha(r-1)\}. \tag{6.2.4}$$

这就是K. Pearson提出的最早的一个检验方法，我们称之为Pearson χ^2拟合优度检验.

对于例6.2.1中的数据，我们可以做如下的χ^2拟合优度检验. 注意到，此时

$$n = 556, n_1 = 315, n_2 = 108, n_3 = 101, n_4 = 32,$$

待检验的假设为

$$H_0 : p_1 = \frac{9}{16}, p_2 = \frac{3}{16}, p_3 = \frac{3}{16}, p_4 = \frac{1}{16}.$$

由于

$$\chi^2 = 0.47 < 7.81 = \chi^2_{0.05}(4-1),$$

故在水平0.05下，我们没有理由拒绝H_0，即可以认为Mendel的结论是可接受的.

从以上内容可以看出，对于下面一般的分布假设

$$H_0 : F(x) \equiv F_0(x), \tag{6.2.5}$$

我们仍可利用上述的χ^2拟合优度检验，其中$F_0(x)$为一个完全已知的分布函数(形式及参数均已知).

此时，我们可以把$(-\infty,\infty)$(或样本空间)分成r个互不相交的区间：

$$(-\infty,\infty) = \cup_{i=1}^r I_i = (-\infty, a_1) \cup [a_1, a_2) \cup \cdots \cup [a_{r-1}, \infty), \tag{6.2.6}$$

且以n_i记落在第i个区间I_i内的样本个数, 再记

$$\begin{array}{ll} p_1 = F(a_1), & p_{10} = F_0(a_1), \\ p_2 = F(a_2) - F(a_1), & p_{20} = F_0(a_2) - F_0(a_1), \\ \cdots & \cdots \\ p_r = 1 - F(a_{r-1}), & p_{r0} = 1 - F_0(a_{r-1}), \end{array}$$

则我们可以利用统计量

$$\chi^2 = \sum_{i=1}^r \frac{(n_i - np_{i0})^2}{np_{i0}} \tag{6.2.7}$$

及定理6.2.1的结论来检验假设(6.2.5).

注 6.2.1 虽然上述的χ^2拟合优度检验可以用来检验一般的分布假设(6.2.5)，但通过上面的分析不难看出，此时我们检验的假设仅为

$$H_0 : p_i = p_{i0},\ i = 1, 2, \cdots, r \tag{6.2.8}$$

而不是真正想检验的假设(6.2.5). 显然(6.2.8)与(6.2.5)有着一定的区别. 如果分点选的不是很好，χ^2拟合优度检验可以会把两个有一定差别的分布检验为没有区别，参见图6.4.

注 6.2.2 在一般情形下，分点的选取应保证落在每个区间内的样本点个数不小于5，且总的样本量不应小于30.

注 6.2.3 通过上面分析，我们可以看到，当F_0中含有未知参数时，上述拟合优度检验无法实施.

§6.2.2 带有未知参数的χ^2拟合优度检验

在许多实际问题中，我们感兴趣的假设可能为

$$H_0 : F(x) \equiv F_0(x; \theta_1, \ldots, \theta_k), \tag{6.2.9}$$

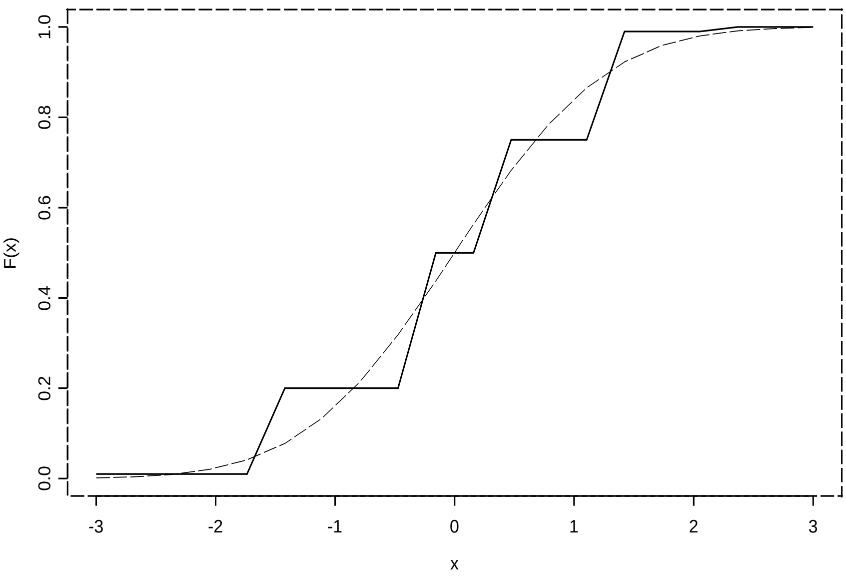

图6.4 针对连续情形的χ^2拟合优度检验

其中$F_0(x;\theta_1,\ldots,\theta_k)$是依赖于$k$个未知参数的形式已知的分布, 如一般的正态分布，二项分布，Poisson分布等. 此时，我们无法计算(6.2.2)或(6.2.7)式的χ^2统计量. Pearson在其1900年的论文中注意到了这个问题，并且认为：用样本估计未知参数后，可以得到$\hat{p}_{i0}$的值，之后可利用下面的统计量

$$\chi^2 = \sum_{i=1}^{r} \frac{(n_i - n\hat{p}_{i0})^2}{n\hat{p}_{i0}} \tag{6.2.10}$$

作为检验统计量，且当H_0成立及$n \to \infty$时，仍然成立$\chi^2 \xrightarrow{\mathcal{L}} \chi^2(r-1)$.

但是，Fisher在上世纪二十年代发现了上述结论是错误的. Fisher指出，当H_0成立时，(6.2.10)式的χ^2统计量在一定条件下有如下的极限零分布

$$\chi^2 = \sum_{i=1}^{r} \frac{(n_i - n\hat{p}_{i0})^2}{n\hat{p}_{i0}} \xrightarrow{\mathcal{L}} \chi^2(r-1-k), \tag{6.2.11}$$

Fisher还指出，并不是所有的参数估计都能满足(6.2.11)式的要求，比如一般情况下的矩估计等，而MLE是近似可行的(请参见陈希孺(1997)). 故对于此类问题，我们将采用MLE先估计未知参数，之后再利用(6.2.11)的结论进行检验.

例 6.2.2 现在某地区随机地抽取1000人, 检查其性别与色盲的数据见表6.1. 按照遗传学规律, 这些数据具有如表6.2的规律. 请问这些数据符合遗传学规律吗？

表6.1 性别与色盲数据

	男	女
正常	442	514
色盲	38	6

表6.2 性别与色盲模型

	男	女
正常	$p/2$	$p^2/2+pq$
色盲	$q/2$	$q^2/2$

解　对于本问题，我们所需检验的假设为

$$H_0: p_1=\frac{p}{2}, p_2=\frac{p^2}{2}+pq=\frac{p(2-p)}{2}, p_3=\frac{1-p}{2}, p_4=\frac{(1-p)^2}{2},$$

其中参数p是未知的，故需要估计. 由于当H_0成立时，数据$(442,514,38,6)$服从一个多项分布，故其似然函数为

$$L(p)\propto\left(\frac{p}{2}\right)^{442}\left(\frac{p(2-p)}{2}\right)^{514}\left(\frac{1-p}{2}\right)^{38}\left(\frac{(1-p)^2}{2}\right)^{6}=p^{956}(2-p)^{514}(1-p)^{50}/2^{1000},$$

于是，由$\partial L(p)/\partial p=0$可求得$p$的MLE为$\hat{p}=0.91$. 由此可算得$\chi^2$的值为

$$\chi^2=\sum_{i=1}^{4}\frac{(n_i-n\hat{p}_i)^2}{n\hat{p}_i}=3.056<5.99=\chi^2_{0.05}(4-1-1),$$

因此，在水平0.05下，我们没有理由拒绝H_0，即认为此地区的数据符合遗传学规律. □

例 6.2.3 某工厂生产一种滚珠，现随机地抽取了50件产品，测得其直径为（单位: mm）:

$$15.0, 15.8, 15.2, 15.1, 15.9, 14.7, 14.8, 15.5, 15.6, 15.3, 15.1, 15.3, 15.0, 15.6, 15.7, 15.8, 14.5,$$

$$14.9, 14.9, 15.2, 15.0, 15.3, 15.6, 15.1, 14.9, 14.2, 14.6, 15.8, 15.2, 15.9, 15.2, 15.0, 14.9, 14.8,$$

$$15.1, 15.5, 15.5, 15.1, 15.1, 15.0, 15.3, 14.7, 14.5, 15.5, 15.0, 14.7, 14.6, 14.2, 14.2, 14.5$$

问: 滚珠直径是否服从正态分布？

解　设滚珠直径为X，其分布函数为$F(x)$，现感兴趣的假设为

$$H_0: F(x)\in\left\{\Phi\left(\frac{x-\mu}{\sigma}\right):\ \mu\in R,\ \sigma>0\right\}.$$

对于此问题，我们首先由数据求得μ,σ^2的MLE为: $\hat{\mu}=15.1,\ \hat{\sigma}^2=0.4325^2$. 另外，由于50个数据的最小值为14.2，最大值为15.9，故我们取分点为

$$a_0=14.05, a_1=14.35, a_2=14.65, a_3=14.95, a_4=15.25, a_5=15.55, a_6=15.85, a_7=16.15$$

由此，可利用公式

$$\hat{p}_i=\Phi\left(\frac{a_i-15.1}{0.4325}\right)-\Phi\left(\frac{a_{i-1}-15.1}{0.4325}\right),\ i=1,\cdots,7$$

求得

$$\hat{p}_1 = 0.0414,\ \hat{p}_2 = 0.1077,\ \hat{p}_3 = 0.2154,$$

$$\hat{p}_4 = 0.2710,\ \hat{p}_5 = 0.2154,\ \hat{p}_6 = 0.1077,\ \hat{p}_7 = 0.0414.$$

于是，由于

$$\chi^2 = 1.7284 < 9.49 = \chi^2_{0.05}(7-1-2),$$

故我们不能拒绝H_0. □

综合前面的内容，我们可以把χ^2拟合优度检验统计量概括如下：

$$\chi^2 = \sum_{i=1}^{r} \frac{(O_i - E_i)^2}{E_i}, \tag{6.2.12}$$

其中O_i, E_i分别表示观测值和期望值.

§6.3 列联表的独立性检验

比较例6.2.2与以前其它例子，我们发现，例6.2.2的数据与众不同. 对于例6.2.2，我们有两个指标（性别、色盲）对这些数据进行分类，而其它例子则只有一个指标值，对于这样的具有几个分类指标的数据表，我们就称之为列联表(Contingency Table). 对于例6.2.2，共有两个分类指标，我们称之为二维列联表. 一般的二维列联表数据如表6.3所示.

表6.3 二维列联表数据

	B_1	B_2	$\cdots$	B_s	Total
A_1	n_{11}	n_{12}	$\cdots$	n_{1s}	$n_{1\cdot}$
A_2	n_{21}	n_{22}	$\cdots$	n_{2s}	$n_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
A_r	n_{r1}	n_{r2}	$\cdots$	n_{rs}	$n_{r\cdot}$
Total	$n_{\cdot 1}$	$n_{\cdot 2}$	$\cdots$	$n_{\cdot s}$	n

其中$n_{i\cdot} = \sum_{j=1}^{s} n_{ij}, n_{\cdot j} = \sum_{i=1}^{r} n_{ij}$. 此时，指标$A, B$分别有$r, s$个水平，且以$n_{ij}$表示在$n$个样本中属于$A_i \cap B_j$的样本个数.

此时我们感兴趣的并不是如在例6.2.2中的分布假设，而是考虑这两个指标之间是否独立，即如下的假设：

$$H_0: \text{指标}A\text{ 与}B\text{ 独立}. \tag{6.3.1}$$

如记

$$p_{ij} = P\{X \in A_i \cap B_j\},\ i = 1, \cdots, r,\ j = 1, \cdots, s, \tag{6.3.2}$$

则这n个样本可以看成是来自多项分布X的样本.

如再记

$$p_{i\cdot} = P\{X \in A_i\},\ i = 1, 2, \cdots, r, \tag{6.3.3}$$

$$p_{\cdot j} = P\{X \in B_j\},\ j = 1, 2, \cdots, s, \tag{6.3.4}$$

则有

$$p_{i\cdot} = \sum_{j=1}^{s} p_{ij},\ p_{\cdot j} = \sum_{i=1}^{r} p_{ij}. \tag{6.3.5}$$

另外，我们还要注意到如下两个约束

$$\sum_{i=1}^{r} p_{i\cdot} = \sum_{j=1}^{s} p_{\cdot j} = 1. \tag{6.3.6}$$

当H_0成立时，应有$p_{ij} = p_{i\cdot}p_{\cdot j}$. 于是，(6.3.1)式的零假设等价于如下假设：

$$H_0 : p_{ij} = p_{i\cdot}p_{\cdot j},\ i = 1, 2, \cdots, r,\ j = 1, 2, \cdots, s. \tag{6.3.7}$$

由于我们可以把上述列联表数据看作是多项分布的样本，故我们可以用χ^2拟合优度检验对其独立性假设(6.3.7)进行显著性检验. 但我们注意到，由于$p_{i\cdot}$和$p_{\cdot j}$是未知的，且有两个约束(6.3.6)，故当H_0成立时共有$r + s - 2$个未知参数. 此时，其未知参数的极大似然估计为

$$\hat{p}_{i\cdot} = \frac{n_{i\cdot}}{n},\ \hat{p}_{\cdot j} = \frac{n_{\cdot j}}{n},\ i = 1, \cdots, r,\ j = 1, \cdots, s.$$

于是，对于二维列联表的χ^2拟合优度检验统计量为

$$\chi^2 = n \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{\left(n_{ij} - \frac{n_{i\cdot}n_{\cdot j}}{n}\right)^2}{n_{i\cdot}n_{\cdot j}}, \tag{6.3.8}$$

且当H_0成立及$n \to \infty$时，我们有

$$\chi^2 \xrightarrow{\mathcal{L}} \chi^2(rs - r - s + 2 - 1) = \chi^2((r-1)(s-1)), \tag{6.3.9}$$

于是，此检验的近似水平为α的检验拒绝域为

$$\{\chi^2 \geq \chi^2_\alpha((r-1)(s-1))\}.$$

对于例6.2.2中的数据，由于$r = s = 2, n = 1000, \chi^2 = 27.4 < 6.64 = \chi^2_{0.01}(1)$, 故我们有理由拒绝$H_0$，即认为色盲与性别之间存在着显著性的相关关系.

注 6.3.1 当分类指标为连续变量时，我们仍然可以先用前面的离散化方法对变量离散后，再用χ^2拟合优度检验对二者的独立性进行检验.

注 6.3.2 当带有未知参数时，Pearson认为与参数已知的极限零分布，而Fisher则从$r = s = 2$的特殊列联表情况指出了此错误.

注6.3.3 我们还可以把上述针对二维列联表的χ^2拟合优度检验推广到高维列联情况，并且在对列联表数据进行分析时，也要注意其维数的重要性.

例 6.3.1 表6.4是1976年-1977年美国佛罗里达洲我凶杀案件中，326个被告的肤色和死刑判决情况:

表6.4 美国某洲的凶杀案二维数据表

被告	死刑 是	死刑 否	合计
白人	19	141	160
黑人	17	149	166
合计	36	290	326

表6.5 美国某洲的凶杀案三维数据表

被告	被害人	死刑 是	死刑 否	死刑比例
白人	白人	19	132	0.126
	黑人	0	9	0.000
黑人	白人	11	52	0.175
	黑人	6	97	0.058

请问死刑与肤色有关吗？

解 从表6.4可以看出，白人被判死刑的比例为$19/160 = 11.9\%$, 而黑人被判死刑的比例为$17/166 = 10.2\%$, 由此可见白人被判死刑的比例高于黑人，说有性别歧视是占不住脚的，并且其似然比统计量$2\ln\lambda = 0.2214$比$\chi^2(1)$的中位数0.455还要小，故我们更无法认为这里有性别歧视.

但是，当我们把被害人的肤色也考虑进去(其数据如表6.5所示), 其情况就大不一样了. 从表6.5可以看出，

- 如果被害人是白人，则白人被告与黑人被告被判死刑的比例分别为12.6%和17.5%;
- 如果被害人是黑人，则白人被告与黑人被告被判死刑的比例分别为0%和5.8%.

从此可以看出，无论被害人是白人还是黑人，黑人被判死刑的比例都比白人被判死刑的比例高. 这就说明死刑判决与肤色有关. 事实上，当我们把上述似然检验应用于这个三维的列联表时，其结论与此吻合. (之所以三维列联表的结论与二维时的不同，其原因是当我们把三维数据看成二维时，会损失掉好多信息) □

§6.4 Kolmogorov检验

概率论的公理化体系是由Kolmogorov于1933年提出的.

我们在§6.2节中讲述的χ^2拟合优度检验具有如图6.4所示的不足，而本节讲述的Kolmogorov检验则可避免其不足，也就是说，本节讲述的Kolmogorov 检验将可以用来很好地检验形如(6.2.5)式的假设.

由于样本经验分布函数$F_n(x)$是$F(x)$的一个很好的估计，故我们知道，当(6.2.5)式的H_0成

立时，$F_n(x)$与$F_0(x)$应相差不大，于是，我们可以用统计量

$$D_n = \sup_x |F_n(x) - F_0(x)|, \tag{6.4.1}$$

来衡量$F(x)$与$F_0(x)$间的差别，且拒绝域为$\{D_n \geq c\}$. 此检验是由Kolmogorov于1933 年提出的，并给出了D_n的极限零分布(见定理1.2.3)，于是，我们称之为Kolmogorov检验. 对于D_n的精确分布，请见下面的定理.

定理 6.4.1 如果$F_0(x)$连续，则当H_0成立时，有

$$P\left\{D_n < \lambda + \frac{1}{2n}\right\} = \begin{cases} 0, & \lambda < 0, \\ \int_{\frac{1}{2n}-\lambda}^{\frac{1}{2n}+\lambda} \int_{\frac{3}{2n}-\lambda}^{\frac{3}{2n}+\lambda} \cdots \int_{\frac{2n-1}{2n}-\lambda}^{\frac{2n-1}{2n}+\lambda} f(\mathbf{x}) \mathrm{d}\mathbf{x}, & 0 \leq \lambda < \frac{2n-1}{2n}, \\ 1, & \lambda \geq \frac{2n-1}{2n}, \end{cases} \tag{6.4.2}$$

其中$f(\mathbf{x}) = \begin{cases} n!, & o < x_1 < \cdots < x_n < 1, \\ 0, & \text{否则}. \end{cases}$

证明 请见陈希孺(1997).

虽然定理6.4.1及定理1.2.3的结论很复杂，但是现已有现成的分位数表可供大家查找.

从D_n的定义可以看出，对于某些连续型分布检验，Kolmogorov检验优于χ^2拟合优度检验. 但是，它的计算看起来比较复杂. 然而，由于$F(x)$及$F_n(x)$均是单增的，且$F_n(x)$是一阶梯函数，于是其上确界仅可能在n个点$x_{(i)}$处达到，即(如果F_0连续)

$$D_n = \max_{1 \leq i \leq n} \left\{F_0(x_{(i)}) - \frac{i-1}{n}, \frac{i}{n} - F_0(x_{(i)})\right\}. \tag{6.4.3}$$

如果F_0是右连续的，则D_n可以写成如下便于计算的公式：

$$D_n = \max_{1 \leq i \leq n} \left\{F_0(x_{(i)} - 0) - \frac{i-1}{n}, \frac{i}{n} - F_0(x_{(i)})\right\}. \tag{6.4.4}$$

例 6.4.1 问在水平$\alpha = 0.1$下是否可以认为下列10个数

0.437, 0.863, 0.034, 0.964, 0.366, 0.469, 0.637, 0.623, 0.804, 0.261

来自于均匀分布$U(0,1)$?

解 对于此组数据，为了便于计算，我们有如下表：

表6.6 例6.4.1的计算表

i	$x_{(i)}$	$F_0(x_{(i)})$	$\frac{i-1}{n}$	$\frac{i}{n}$	δ_i	i	$x_{(i)}$	$F_0(x_{(i)})$	$\frac{i-1}{n}$	$\frac{i}{n}$	δ_i
1	0.034	0.034	0.0	0.1	0.066	6	0.623	0.623	0.5	0.6	0.123
2	0.261	0.261	0.1	0.2	0.161	7	0.637	0.637	0.6	0.7	0.063
3	0.366	0.366	0.2	0.3	0.166	8	0.804	0.804	0.7	0.8	0.104
4	0.437	0.437	0.3	0.4	0.137	9	0.863	0.863	0.8	0.9	0.063
5	0.469	0.469	0.4	0.5	0.069	10	0.964	0.964	0.9	1.0	0.064

其中$\delta_i = \max\left\{F_0(x_{(i)}) - \frac{i-1}{n}, \frac{i}{n} - F_0(x_{(i)})\right\}$. 从表6.6可以看出，本问题的Kolmogorov统计量的值为$D_n = 0.166 < 0.3687 = D_{0.1}(10)$，故我们没有理由拒绝$H_0$，即可以认为这组数是来自$U(0,1)$的样本. □

关于我们前面讲述的两个分布检验方法：χ^2拟合优度检验与Kolmogorov检验，我们有如下几点注解：

注 6.4.1 当F_0为完全已知的连续分布时，Kolmogorov检验优于χ^2拟合优度检验. 但如果F_0是离散的或含有未知参数，则Kolmogorov检验无法应用.

注 6.4.2 当F_0为含有未知参数的连续分布时，如用样本去估计未知参数，则定理1.2.3及定理6.4.1的结论均不成立，故此时我们无法利用Kolmogorov检验，但χ^2拟合优度检验是可以的. (如果F_0是正态分布和指数分布，则有人给出了D_n的分布，有兴趣的请参见茆诗松、王玲玲(1984))

注 6.4.3 对于两样本分布假设检验问题，即设$X_1, \ldots, X_m$和$Y_1, \ldots, Y_n$为分别来自总体F和G 的iid 样本，且全样本独立，则我们可以用统计量

$$D_{m,n} = \sup_x |F_n(x) - G_m(x)|, \tag{6.4.5}$$

进行分布检验，且拒绝域为$\{D_{m,n} \geq c\}$, 其中F_n, G_m 分别表示总体X, Y 的经验分布函数. 由于Simirnov 给出了此统计量的极限分布，故我们称之为Smirnov检验. 由于它是Kolmgorov检验的自然推广，故有的书也统称之为Kolmogorov-Smirnov检验.

注 6.4.4 Kolmogorov检验很难处理高维数据的分布检验，而χ^2拟合优度检验则与一维的类似.

注 6.4.5 如取F_0为$U(0,1)$的分布函数，则Kolmogorov统计量D_n还可作为均匀性的一种度量，我们称这种度量为偏差(discrepancy), 而χ^2拟合优度检验统计则不然.

关于分布检验，还有许多类似于Kolmogorov统计量的许多检验统计量，如

- $\sup_x |w(x)(F_n(x) - F_0(x)|$, 其中$w(x)$为给定的权函数.
- Cramér-Von Mises统计量：$C_n = n\int(F_n(x) - F_0(x))^2 \mathrm{d}F(x)$.

- Anderson-Darling 统计量：$n\int w(F_0(x))(F_n(x)-F_0(x))^2\mathrm{d}F(x)$, 其中$w(t)=[t(1-t)]^{-1/2}$.

关于Kolmogorov检验还有许多更深入的结论，有兴趣的读者请参见Serfling(1980)等有关大样本的书籍.

§6.5 正态性检验

虽然我们在本章第一节讲述了正态分布的正态概率纸检验法，但由于那种方法比较粗糙，故本节将介绍另外两种更有效的常用检验方法. 当然，关于正态性检验现有许多种，有兴趣的读者请参见梁小筠(1997). 另外，经过国际标准化组织(ISO)统计标准分委员会的研究，最后认为Wilk-Shapiro的W检验和D′Agostino的D检验法最好(它们的第二类错误概率最小)，于是它们向世界各国推荐使用这两个正态性检验方法. 我国已把这两个方法列入了国家标准，见GB488–85. 当样本量$n\le 50$时，用W检验；当样本量$n>50$时，用D检验.

为了方便，本节始终假设$X_1,\dots,X_n$为来自总体为F的iid样本，现感兴趣的假设为

$$H_0: F(x)\in\left\{\Phi\left(\frac{x-\mu}{\sigma}\right):\mu\in R,\sigma>0\right\} \tag{6.5.1}$$

§6.5.1 W检验($n\le 50$)

W检验是Wilk和Shapiro于1965年提出的，其检验统计量为

$$W=\frac{\left[\sum_{k=1}^{[n/2]}a_k(n)(X_{(n+1-k)}-X_{(k)})\right]^2}{\sum_{i=1}^n(X_{(i)}-\bar{X})^2}, \tag{6.5.2}$$

其中$X_{(1)}\le\cdots\le X_{(n)}$为次序统计量，$\{a_k(n)\}$是一组给定的与$n$有关的常数，可查表求得.

W检验的拒绝域为

$$W_0=\{W<c\}, \tag{6.5.3}$$

其中临界值c满足$P\{W<c|H_0\}=\alpha$.

此检验的理由大致如下：

设$X_1,\dots,X_n$为来自$N(\mu,\sigma^2)$的iid样本，如记$Y_i=\frac{X_{(i)}-\mu}{\sigma}, m_i=EY_i, \epsilon_i=X_{(i)}-EX_{(i)}$, 则知$m_i$是与$\mu,\sigma$无关的常数，且有

$$X_{(i)}=\mu+\sigma m_i+\epsilon_i. \tag{6.5.4}$$

由于当n较大时，$X_{(i)}\approx EX_{(i)}$，因而ϵ_i较小. 于是n个点$\{(X_{(i)},m_i):i=1,2,\cdots,n\}$近似在一条直线上.

为区别这n个点是否在一条直线上，我们可采用下面的R^2统计量

$$R^2=\left[\frac{\sum_{i=1}^n(X_{(i)}-\bar{X})(m_i-\bar{m})}{\sqrt{\sum_{i=1}^n(X_{(i)}-\bar{X})^2\sum_{i=1}^n(m_i-\bar{m})^2}}\right]^2, \tag{6.5.5}$$

其中$\bar{m} = \sum_{i=1}^{n} m_i/n$.

当H_0成立时，即X_i服从$N(\mu, \sigma^2)$时，(6.5.5)式的R^2应接近1. 于是，当R^2远离1时，我们有理由拒绝H_0.

由于$N(0,1)$是对称分布，所以当H_0成立时，$(Y_1, \ldots, Y_n)$与$(-Y_1, \ldots, -Y_n)$有相同的联合分布，从而Y_k与$-Y_{n+1-k}$也有相同的分布，故$m_k = -m_{n+1-k}, \bar{m} = 0$. 于是，

$$R^2 = \frac{\left[\sum_{i=1}^{n} X_{(i)} m_i\right]^2}{\sum_{i=1}^{n}(X_{(i)} - \bar{X})^2 \sum_{i=1}^{n} m_i^2} = \frac{\left[\sum_{k=1}^{[n/2]} m_{n+1-k}(X_{(n+1-k)} - X_{(k)})\right]^2}{\sum_{i=1}^{n}(X_{(i)} - \bar{X})^2 \sum_{i=1}^{n} m_i^2}, \tag{6.5.6}$$

如取

$$a_k(n) = \frac{m_{n+1-k}}{\left[\sum_{i=1}^{n} m_i^2\right]^{1/2}}, \tag{6.5.7}$$

则知$W = R^2$.

为理解(6.5.6)式，我们定义

$$Q^2 = \sum_{i=1}^{n}(X_{(i)} - \bar{X})^2,\ \hat{\sigma} = \sum_{i=1}^{n} b_i X_{(i)},\ b_i = m_i / \sum_{i=1}^{n} m_i^2,$$

则R^2可以重写成

$$R^2 = \sum_{i=1}^{n} m_i^2 \hat{\sigma}^2 / Q^2.$$

我们注意到，无论样本是否正态，Q^2经修正后都可以成为σ^2的一个UE，而分子$\hat{\sigma}$是样本的线性函数，且当H_0成立时它是σ的UE，于是$\hat{\sigma}^2$也可成为σ^2 的一个估计. 这样，在正态性假设下，二者的差别不应很大. 这就是我们采用R^2来检验正态性假设的一个直观解释. 为了扩大这个差异，Wilk 和Shapiro把上述的$\hat{\sigma}$换为σ的BLUE

$$\hat{\sigma}' = \sum_{i=1}^{n} c_i X_{(i)}, \tag{6.5.8}$$

就得到了W检验统计量.

§6.5.2 D检验($n > 50$)

W检验虽有效，但可惜的是它只适用于容量3–50的样本，其主要原因是，对超过50的样本容量n，我们很难计算其分位数值. 为此，Dagostino提出了D检验.

由W统计量的定义不难看出，如果不计常数因子，它是正态分布的标准差σ的BLUE与样本标准差的比值的平方，即关于σ的两个估计量的比值的平方. 由于BLUE的系数的计算是繁杂的，且随着样本量的增加，这种繁杂在增加. 于是，随着样本量的增加，我们不得不放弃用BLUE来构造检验统计量的打算. 这样，Dagostino建议，在样本容量超过50时，用$\left(i - \frac{n+1}{2}\right)$来

代替BLUE(6.5.8)式中的c_i，即考虑用

$$T = \sum_{i=1}^{n} \left(i - \frac{n+1}{2} \right) X_{(i)}$$

作为σ的一个线性无偏估计. 再考虑到n的阶，我们用统计量

$$D_n = \frac{\sum_{i=1}^{n} \left(i - \frac{n+1}{2} \right) X_{(i)}}{n^{3/2} \sqrt{\sum_{i=1}^{n} (X_i - \bar{X})^2}}, \tag{6.5.9}$$

作为正态性检验统计量，且当它远离1时，我们有理由拒绝正态性假设.

为求取其极限零分布，我们可以证明，当H_0成立时

$$E(D_n) = 0.28209479, \quad \sqrt{\mathrm{Var}(D_n)} = 0.02998598/\sqrt{n},$$

故我们采用如下的标准化随机变量

$$Y_n = \frac{\sqrt{n}(D_n - 0.28209479)}{0.02998598} \tag{6.5.10}$$

作为检验统计量，且其拒绝域为

$$\{Y_n < c_1\} \cup \{Y_n > c_2\},$$

其中$c_1 < c_2$.

关于D检验的分位数现已有表可查.

习题六

1. 请检验下面的20个数据是否来自$N(0,1)$:

−2.4	−2.1	−1.2	−0.7	−0.4	−0.3	−0.2	−0.1	−0.05	−0.02
0.01	0.04	0.15	0.2	0.4	0.8	1.4	2.2	2.8	3.1

2. 在π的前800位小数的数字中，0至9十个数字分别出现了74，92，83，79，80，73，77，75，76和91次. 试在水平0.05下检验这十个数字出现的可能性相等.

3. 在著名的Rutherfort-Chadwik-Ellis实验中，观察了一放射性物质每1/8分钟内放射的α粒子数，他共观察了2612次，结果如下：

粒子数:	0	1	2	3	4	5	6	7	8	9	10	11
频数:	57	203	383	525	532	408	273	139	49	27	10	6

请问在水平0.1下，上述数据是否与Poisson分布相符？

4. 设总体分布关于原点对称，且$X_1,\ldots,X_n$为来自此总体的iid样本，$X_{(1)}\le\cdots\le X_{(n)}$ 为次序统计量，并记

$$E(X_{(i)}=m_i,\ \ \mathrm{Cov}(X_{(i)},X_{(j)})=v_{ij},\ i,j=1,\cdots,n.$$

请证明：

(1)$m_i=-m_{n+1-i}$，且$\sum_{i=1}^n m_i=0$.

(2) $v_{ij}=v_{n+1-i\ n+1-j}$，且$V=(v_{ij})$不仅关于主对角线对称，而且还关于副对角线对称.

5. 证明当$n=3$时，正态性W检验的统计量W的密度函数为

$$\frac{3}{\pi}[w(1-w)]^{-1/2},\ 3/4\le w\le 1.$$

6. 设$X_{(1)}\le\cdots\le X_{(n)}$为来自标准正态总体的次序统计量，$T=\sum_{i=1}^n\left(i-\frac{n+1}{2}\right)X_{(i)}$. 证明

$$E(T)=\frac{n(n-1)}{2\sqrt{\pi}},\ E(T^2)=\frac{n(n-1)}{4}+\frac{n(n-1)(n-2)(n-3)}{4\pi}+\left(\frac{1}{12}+\frac{\sqrt{3}}{2\pi}\right)n(n-1)(n-2).$$

7. 为研究慢性气管炎与吸烟量的关系，现调查了272人，其结果如下：

	吸烟量(支/日)			
是否串病	0–9	10–19	≥ 20	和
是	22	98	25	145
否	22	89	16	127
和	44	187	41	272

请问慢性气管炎与每日的吸烟量有关吗？($\alpha=0.05$)

8. 消费者转会为了解消费者对市场上五种矿泉水的偏好，现随机地调查了1000人，得到如下数据：

品牌	1	2	3	4	5
喜欢的人数	210	212	170	185	223

请问调查结果是否说明消费者对这五种品牌的矿泉水存在着不同的偏好？

第7章　某些统计模拟方法

随着计算机的快速发展，统计计算的作用也越来越大. 同时随着计算速度的提高，某此复杂问题的计算现在也不再是问题，于是就产生了一些新的计算方法或算法. 本章将集中讲述随机数的产生方法及某些常用的统计模拟方法.

§7.1　均匀随机数的产生

有许多分布的随机数的产生都依赖于均匀分布随机数，故本节将看一看均匀分布随机数的产生方法.

线性同余法是产生均匀随机数最常用的方法，其公式为

$$\begin{cases} x_n = (ax_{n-1} + c), \quad (\mathrm{mod}\ M), \\ r_n = x_n/M, n = 1, 2, \cdots. \end{cases} \tag{7.1.1}$$

其中x_0为初值，a为乘子，c为增量，M为模. 当$c = 0$时，称之为同余法；$c > 0$时，称之为混合同余法.

由此式可以看出，随机数的产生依赖于初值及M等，但由于计算机位数的限制，一般M是给定的. 并且从上式可以看出，这样产生的随机数均具有一定的周期性，于是，加大其周期，是研究随机数产生方法的研究课题之一.

一般地，我们$M = 2^L, a = 4\alpha + 1, c = 2\beta + 1, x_0$为任意非负整数，其中$L$为计算机的字长，$\alpha, \beta$为任意的整数.

另外，还有其它一些方法，请参见高惠璇(1995).

§7.2　随机数的产生

随机模拟(simulation)在许多方面都得到了应用，并取得了很好的经济效益. 在随机模拟中，最主要的一点就是要产生符合实际要求的随机数. 由于我们所应用的多是由计算机程序产生的随机数，虽然它具有一些随机的特点，但从严格意义上的讲，它们不是随机的，故我们多称之为伪随机数(pseudo random number). 本节将介绍一些产生伪随机数的基本方法.

§7.2.1　逆变换法

设随机变量X 的分布函数为$F(x)$, 定义

$$F^{-1}(y) = \inf\{x : F(x) \geq y\},\ 0 \leq y \leq 1. \tag{7.2.1}$$

有了逆变换的定义后，就有下面一个在产生随机数时非常有用的定理.

定理 7.2.1 设随机变量U 服从(0,1) 上的均匀分布, 则$X = F^{-1}(U)$ 的分布函数为$F(x)$.

证明 由(7.2.1) 式和均匀分布的分布函数可得

$$\begin{aligned} P(X \le x) &= P\{F^{-1}(U) \le x\} \\ &= P\{U \le F(x)\} = F(x). \end{aligned}$$

□

为了方便，在本章，我们始终假设U是服从(0,1)上均匀分布的随机变量.

例 7.2.1 设$X_1, \cdots, X_n$ 是来自$F(x)$ 的一个样本, 其中$F(x)$是一个已知的函数. 请给出产生$X_{(1)}$的随机数的方法.

解 方法一：

- 产生n个$(0,1)$上均匀分布的随机数$U_1, \ldots, U_n$;
- 做变换：$X_i = F^{-1}(U_i), i = 1, \cdots, n$;
- 计算$X_{(1)} = \min\{X_1, \ldots, X_n\}$；
- 重复上面三步N次，则得到了N个所求的随机数.

方法二：由于我们知道$X_{(1)}$的分布函数为

$$F_1(y) = 1 - [1 - F(y)]^n, \tag{$*.1$}$$

于是，我们可以直接利用上述分布函数求取其随机数，叙述如下：

- 产生N个$(0,1)$上均匀分布的随机数$U_1, \ldots, U_N$;
- 做变换: $Y_i = F_1^{-1}(U_i) = F^{-1}(1 - (1 - U_i)^{1/n}), i = 1, \cdots, N$

上式第二步产生的即为所求. □

注意在求(*.1)式的逆时，我们用到了定理7.2.1的结论.

§7.2.2 合成法

如果一个随机变量X的分布函数比较复杂，我们很难用上小节的逆变换法抽样. 但如果知道条件随机变量$X|Y$及随机变量Y均易于抽样，则可以利用本节的合成法进行. 这种方法是由Butler于1958年提出的. 其抽样步骤为

- 由Y的分布$G(y)$产生其随机数y;
- 由条件分布$F(x|Y = y)$产生的随机数X即为所求.

例 7.2.2 设X 的密度函数为$f(x)=\sum_{i=1}^{n}a_if_i(x)$, 其中$\sum_{i=1}^{n}a_i=1, f_i(x)$ 是密度函数, 给出求取X的随机数的方法.

解 令$a_0=0$,则由合成法可知，其随机数产生的方法为

- 产生(0,1)上均匀分布的随机数u;
- 确定i, 使$\sum_{j=0}^{i-1}a_j<u\le\sum_{j=0}^{i}a_j$;
- 由$f_i(x)$抽取随机数x, 即为所求.

比如，随机变量X的PDF为$f(x)=(1+2x)/6, 0<x<2$, 其分布函数为$F(x)=(x+x^2)/6$, 若用逆变换法抽样, 则要解二次方程, 较为麻烦. 考虑用合成法, 将$f(x)$ 分解成：$f_1(x)=1/2, 0<x<2; f_2(x)=x/2, 0<x<2$, $a_1=1/3, a_2=2/3$, 由合成法及逆变换法, 其具体抽样步骤为：

- 独立地产生$U(0,1)$的随机数u_1,u_2;
- 如果$u_1<1/3$, 则取$x=2u_2$；如$u_1\ge 1/3$，则取$x=2\sqrt{u_2}$.

□

关于随机数的产生，还有许多其它方法，由于现在的统计软件中都有具体分布的随机数的产生程序，故我们就不再多加详细地说明了，可参见Ross (2002)或Rubinstein (1981).

由于随机向量的抽样方法有时会用到，且一般统计软件中没有，故下节将专门看一看随机向量的产生方法.

§7.2.3 随机向量的抽样法

如果随机向量的各分量间的独立的，则它等价于多个一元随机变量的抽样，下面讨论它分量间非独立的情形.

设$X_1,\dots,X_k$ 的概率密度函数为

$$f(x_1,\dots,x_k)=f_1(x_1)f_2(x_2|x_1)\dots f_k(x_k|x_1,\dots,x_{k-1}), \tag{7.2.2}$$

其中$f_1(x_1)$ 为X_1 的边际概率密度函数, $f_i(x_i|x_1,\dots,x_{i-1})$ 为给定$X_1,\dots,X_{i-1}$ 下X_i 的条件概率密度函数. 则其随机向量的产生方法为：

定理 7.2.2 设$U_1,\dots,U_k$ 是独立同分布的$U(0,1)$ 随机变量, $X_1,\dots,X_k$ 是方程

$$\begin{cases} F(X_1)=U_1, \\ F_i(X_i|X_1,\dots,X_{i-1})=U_i, i=2,\dots,k. \end{cases}$$

的解, 其中F_i 是对应于f_i 的分布函数, 则$X_1, \dots, X_k$ 的分布满足(7.2.2).

证明 类似于定理7.2.1，略. □

但注意到，由于(7.2.2)的表示不唯一，而不同的表示有可能得到不同的解，故在应用时要多加注意. 下面看一个例子.

例 7.2.3 设X_1, X_2 的联合密度函数为

$$f(x_1,x_2)=\begin{cases} 6x_1, & \text{如果}x_1+x_2\le 1, x_1\ge 0, x_2\ge 0,\\ 0, & \text{其它}. \end{cases}$$

试求随机向量(X_1,X_2)的随机数.

解 如果将$f(x_1,x_2)$ 写成$f_1(x_1)f_2(x_2|x_1)$ 的形式, 而X_1 的边际分布密度函数为

$$f_1(x_1)=\int_0^{1-x_1} f(x_1,x_2)\mathrm{d}x_2 = 6x_1(1-x_1),\ 0\le x_1\le 1,$$

而给定$X_1=x_1$ 下X_2 的条件分布密度函数为

$$f_2(x_2|x_1)=\frac{f(x_1,x_2)}{f_1(x_1)}=\frac{1}{1-x_1},\ 0\le x_2\le 1-x_1,$$

相应的边际分布函数和条件分布函数分别为

$$\begin{aligned} F_1(x_1) &= \int_0^{x_1} f_1(t)\mathrm{d}t = 3x_1^2-2x_1^3,\ 0\le x_1\le 1,\\ F_2(x_2|x_1) &= \int_0^{x_2} f_2(t|x_1)\mathrm{d}t = x_2(1-x_1)^{-1}, 0\le x_2\le 1-x_1. \end{aligned}$$

于是所求的随机数由方程

$$3X_1^2-2X_1^3=U_1,$$
$$X_2(1-X_1)^{-1}=U_2.$$

来决定.

如果将$f(x_1,x_2)$ 写成$f_2(x_2)f_1(x_1|x_2)$ 的形式, 其中

$$\begin{aligned} f_2(x_2) &= \int_0^{1-x_2} f(x_1,x_2)\mathrm{d}x_1 = 3(1-x_2)^2,\ 0\le x_2\le 0\\ f_1(x_1|x_2) &= \frac{f(x_1,x_2)}{f_2(x_2)} = 2x_1(1-x_1)^{-2},\ 0\le x_1\le x_2. \end{aligned}$$

相应的边际分布函数和条件分布函数分别为

$$\begin{aligned} F_2(x_2) &= \int_0^{x_2} f_2(t)\mathrm{d}t = 1-(1-x_2)^2,\ 0\le x_2\le 1,\\ F_1(x_1|x_2) &= \int_0^{x_1} f_1(t|x_2)\mathrm{d}t = x_1^2(1-x_2)^{-2}. \end{aligned}$$

于是所求的随机数由方程

$$1-(1-X_2)^3=U_1,$$
$$X_1^2(1-X_2)^{-2}=U_2.$$

比较以上两种次序，显然，第二种方法要简单，并且可以得到如下的显示解：

$$X_2=1-(1-U_1)^{1/3},\ X_1=(1-U_1)^{1/3}U_2^{1/2}.$$

□

例 7.2.4 请给出多元正态分布的随机向量的抽样方法.

解 设$\mathbf{X}=(X_1,\ldots,X_k)$ 服从k 维正态分布$N_k(\mu,\Sigma)$, 其密度函数为

$$p(x)=(2\pi)^{-k/2}|\Sigma|^{-1/2}\exp\{-\frac{1}{2}(x-\mu)\Sigma^{-1}(x-\mu)\},$$

此处μ 为已知均值向量, $\Sigma=(\sigma_{ij})_{k\times k}$ 为已知协差阵, $|\Sigma|$ 是其行列式.

我们知道, 因Σ 是正定对称矩阵(此处不讨论奇异情形), 故存在下三角阵

$$C=\begin{bmatrix} c_{11} & 0 & \cdots & 0 \\ c_{21} & c_{22} & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ c_{k1} & c_{k2} & \cdots & c_{kk} \end{bmatrix}$$

使得$\Sigma=CC'$, 于是, 若$\mathbf{Z}=(Z_1,\cdots,Z_k)\sim N_k(0,I_k)$, 则$\mathbf{X}=C\mathbf{Z}+\mu\sim N_k(\mu,\Sigma)$.

事实上，其中的下三角阵C的计算公式为：

$$c_{ij}=\frac{\sigma_{ij}-\sum_{l=1}^{j-1}c_{il}c_{jl}}{(\sigma_{jj}-\sum_{l=1}^{j-1}c_{jl}^2)^{1/2}}.\ i=1,\cdots,k,j=1,\cdots,i.$$

□

§7.3 随机模拟

掌握了随机数的产生方法，则可以利用随机数进行模拟计算了. 统计模拟在许多方面都有应用，我们在这里仅是作一简单介绍.

统计模拟方法也称Monte Carlo方法，它最早产生于18世纪的Buffon投针试验：在1777年, 法国学者Buffon 提出用试验方法求园周率π 的值, 其原理如下: 假设平面上有无数条距离为1 的等矩平行线, 现向该平面随机地投掷一根长度为l $(l\le 1)$ 的针, 则我们可以计算该针与任一

表7.3.1 Buffon 试验的结果

试验者	时间(年)	针长l	投针次数	相交次数	π 估计值
Wolf	1850	0.80	5000	2532	3.15956
Smith	1855	0.60	3204	1218	3.15665
Fox	1884	0.75	1030	489	3.15951
Lazzarini	1925	$0.8\dot{3}$	3408	1808	3.14159292

平行线相交的概率. 此处随机投掷可以这样理解: 针的中心点与最近的平行线间的距离x 均匀地分布在区间[0,1/2] 上, 针与平行线的夹角h (不管相交与否) 均匀地分布在区间$[0,\pi]$上. 于是, 针与线相交的充要条件是$\frac{x}{\sin\psi} \le \frac{l}{2}$, 从而针线相交概率为

$$p\hat{=}P\left\{X \le \frac{l}{2}\sin\psi\right\} = \int_0^{\pi}\int_0^{\frac{l}{2}\sin\psi}\frac{2}{\pi}\mathrm{d}x\mathrm{d}\psi = \frac{2l}{\pi}.$$

由此可见，针与线相交的概率与π有关. 如果我们向这样的平行线上投掷大量的针，则大数定律告诉我们，可用其相交的频率来估计概率，从而得到π的估计值，历史上曾有几位学者做过这样的试验，其结果见表7.3.1.

从表7.3.1可以看出，其精度不是很高，其原因是试验次数不够(为达到指定的精度，可以用中心极限定理估计出需要的样本容量). 虽然上述投针试验在实际中不易大量实施，但如果利用计算机，则可以很容易做到关于π的精确估计.

从Buffon投针试验可以看出，许多实际问题都与积分的近似计算有关，于是，下面我们仅从积分的近似计算角度出发，给出几种随机模拟方法. 设我们感兴趣的积分为

$$\theta = \int_a^b f(x)\mathrm{d}x, \tag{7.3.1}$$

其中$0 \le f(x) \le M,\ \forall\ x \in [a,b], M$为给定的正数.

§7.3.1 随机投点法

由于(7.3.1)式的积分是曲线$f(x)$在区间$[a,b]$内的面积，故积分的近似计算问题就是面积的估计问题. 为此定义$\Omega = \{(x,y) : a \le x \le b, 0 \le y \le M\}$，$(X,Y)$ 是在Ω 上均匀分布的二维随机变量, 其联合密度函数为

$$f(x,y) = \frac{1}{M(b-a)}I_{(a\le x\le b, 0\le y\le M)},$$

由此可求得(X,Y)落在Ω内的概率为$p = \int\int_{(x,y)\in\Omega} f(x,y)\mathrm{d}x\mathrm{d}y = \theta/(M(b-a))$. 如用频率估计概率，则我们有如下的$\theta$的一个估计值

$$\hat{\theta}_1 = M(b-a)\frac{n_0}{n}, \tag{7.3.2}$$

其中n为由(X,Y)产生的随机数的个数，n_0为落入区域Ω内的个数. 这种随机模拟方法就称为随机投点法.

另外，由于$n_0\sim b(n,p)$，故知此估计的方差为

$$\mathrm{Var}\hat{\theta}_1=\frac{M^2(b-a)^2}{n^2}\mathrm{Var}(n_0)=\frac{\theta}{n}[M(b-a)-\theta]. \tag{7.3.3}$$

若用估计的标准差来衡量其精度, 则此种估计精度的阶为$n^{-1/2}$.

§7.3.2 样本平均值法

如设$g(x)$为某随机变量X的PDF，且其取值在(a,b)上，于是，(7.3.1)式的积分可以写成

$$\theta=\int_a^b\frac{f(x)}{g(x)}g(x)\mathrm{d}x=E\left[\frac{f(X)}{g(X)}\right], \tag{7.3.4}$$

这样，由矩估计的想法，我们可以用样本矩来估计θ，这便是样本平均值法的思想.

特别地，如果a,b均有限，可取$g(x)=1/(b-a)$，它是$U(a,b)$的PDF，则$\theta=(b-a)Ef(X)$. 如设$X_1,\ldots,X_n$为来自$U(a,b)$ 的随机数，则θ的一个估计为

$$\hat{\theta}_2=\frac{b-a}{n}\sum_{i=1}^n f(X_i). \tag{7.3.5}$$

显然，$\hat{\theta}_2$是θ的UE，且其方差为

$$\mathrm{Var}\hat{\theta}_2=\frac{1}{n}\mathrm{Var}f(X)=\frac{b-a}{n}\left[\int_a^b f^2(x)\mathrm{d}x-\theta^2\right]. \tag{7.3.6}$$

虽然样本平均值法的精度的阶也为$n^{-1/2}$, 但可以证明它的标准差不大于随机投点法. 由于(7.3.5)式的估计是取$g(x)$ 为均匀分布的PDF而得到的，则一个问题是：是否可以找到一个更好的$g(\cdot)$，以使得估计

$$\hat{\theta}_g=\frac{1}{n}\sum_{i=1}^n\frac{f(X_i)}{g(X_i)}, \tag{7.3.7}$$

的精度更好呢？其中$X_1,\ldots,X_n$为来自$g(x)$的独立随机数.

事实上，因为$\mathrm{Var}\hat{\theta}_g=\frac{1}{n}\left[E\left(\frac{f(X)}{g(X)}\right)^2-\theta^2\right]$，故当$f(x)>0$时，取$g(x)=f(x)/\theta$，则有$\mathrm{Var}\hat{\theta}_g=0$. 但由于$\theta$是未知的，故知上述方法不可行. 然而，这个结果告诉我们，如果选一个与$f(x)$非常接近的PDF作为$g(x)$，则知此时的估计精度会很好(这种方法也称重要抽样法). 看下面一个例子.

例 7.3.1 *考虑$\theta=\int_0^1 e^x\mathrm{d}x$的近似计算.*

解 如用(7.3.5)式估计θ，则知其方差为$\mathrm{Var}\hat{\theta}_2=\frac{0.242}{n}$.

下面看一看重要抽样方法的应用，其主要想法就是要寻找一个与$f(x)=e^x$相近的概率密度函数$g(\cdot)$. 由于$f(x)=1+x+\frac{x^2}{x}+\cdots$, 故我们取$g(x)=\frac{2(1+x)}{3}$, 它是$(0,1)$上一个PDF. 现设$X_1,\ldots,X_n$ 为来自$g(x)$的随机数，则可取θ的估计为

$$\hat{\theta}_g=\frac{1}{n}\sum_{i=1}^{n}\frac{f(X_i)}{g(X_i)}=\frac{3}{2n}\sum_{i=1}^{n}\frac{e^{X_i}}{1+X_i},$$

另外，也可以求得其方差为$\mathrm{Var}\hat{\theta}_g=\frac{0.0269}{n}$, 它远小于$\mathrm{Var}\hat{\theta}_2$. □

再分析一下估计式(7.3.5), 它是在假设在区间$(0,1)$内均匀抽样而得到的. 显然，在区间$(0,1)$内的不同点对此积分的贡献是不同的，于是我们可以通过适当地划分区间，而得到一个更好的估计，这就是分层抽样方法的基本思想. 其具体做法为：首先把样本空间(a,b) 分成一些小区间$D_1,\cdots,D_m$, 且诸D_i 不交, $\cup D_i=(a,b)$, 然后在各小区间内的抽样数由其贡献大小决定, 亦即, 定义

$$p_i=\int_{D_i}f(x)\mathrm{d}x,$$

则, D_i 内的抽样数应与p_i 成正比. 如此, 对θ 贡献大的D_i 抽样多, 以提高抽样效率.

下设$a=0,b=1$，且把$[0,1]$ 分成m个小区间, 各区间端点记为a_i, $0=a_0<a_1<\cdots<a_m=1$,则

$$\theta=\int_0^1 f(x)\mathrm{d}x=\sum_{i=1}^{m}\int_{a_{i-1}}^{a_i}f(x)\mathrm{d}x\hat{=}\sum_{i=1}^{m}I_i,$$

之后用样本平均值法求取I_i的估计值$\hat{I}_i=\frac{l_i}{n_i}\sum_{j=1}^{n_i}f(X_{ij})$, 其中$l_i=a_i-a_{i-1},X_{i1},\ldots,X_{in_i}$为来自$U(a_{i-1},a_i)$的随机数, $\sum n_i=n$. 于是得到θ的估计为

$$\hat{\theta}_3=\sum_{i=1}^{m}\hat{I}_i, \tag{7.3.8}$$

其方差可以求得为

$$\mathrm{Var}\hat{\theta}_3=\sum_{i=1}^{m}\frac{l_i^2}{n_i}\sigma_i^2,\quad 其中\ \sigma_i^2=\int_{a_{i-1}}^{a_i}\frac{f^2(x)}{l_i}\mathrm{d}x-\left(\frac{I_i}{l_i}\right)^2. \tag{7.3.9}$$

从(7.3.9)可以看出，分层抽样法的精度与区间划分有关，也与样本量的分配有关，看下例.

例 7.3.2 (续例7.3.1) 如将区间分为$[0,0.5)$和$[0.5,1]$两个区间，则考虑$\hat{\theta}_3$的估计精度.

解 此时可求得$I_1=\sqrt{e}-1, I_2=e-\sqrt{e}$, $\sigma_1^2=(e-1)-4(\sqrt{e}-1)^2=0.03492$, $\sigma_2^2(e^2-e)-4(e-\sqrt{e})^2=0.09493$, 且其方差为$\mathrm{Var}\hat{\theta}_3=\frac{0.5^2}{n_1}\sigma_1^2+\frac{0.5^2}{n-n_1}\sigma_2^2$, 由此可知，当$\frac{n_1}{n}=\frac{\sigma_1}{\sigma_1+\sigma_2}=0.37753$时方差达到最小，且其最小方差为$0.06125/n$, 它小于$\mathrm{Var}\hat{\theta}_1$，而大于$\mathrm{Var}\hat{\theta}_2$. □

如果将上述分法继续下去，则可以得到一个比第二种方法来得小的方差.

对于固定的区间划分，如诸σ_i^2已知，则可以证明在n固定时，取

$$n_i = \frac{n l_i \sigma_i}{\sum_{i=1}^m l_i \sigma_i}, \tag{7.3.10}$$

能使(7.3.8)式的$\hat{\theta}_3$的方差达到最小值$(\sum_{i=1}^m l_i \sigma_i)^2/n$.

但我们注意到，实际中(7.3.10)式中σ_i是未知的. 然而，即使我们取简单的分配$n_i = \frac{l_i}{\sum l_i} n = \frac{l_i n}{b-a}$, 也有$\mathrm{Var}(\hat{\theta}_3) \le \mathrm{Var}(\hat{\theta}_2)$

关于随机模拟还有许多研究内容有讲述，如，减小方差的技术、MCMC方法等. 请有兴趣的读者参见Ross (2005).

第8章　某些常用分布的分位数表

附表1：准正态分布累积分布函数值$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\mathrm{d}x$.

附表2：t分布的分位数表$\int_{c}^{\infty}t(x;n)\mathrm{d}x=\alpha$.

附表3：χ^2分布的分位数表 $\int_{c}^{\infty}\chi^2(x;n)\mathrm{d}x=\alpha$.

附表4：F分布的分位数表 $\int_{c}^{\infty}f(x;m,n)\mathrm{d}x=\alpha(\alpha=0.2)$.

附表5：F分布的分位数表 $\int_{c}^{\infty}f(x;m,n)\mathrm{d}x=\alpha(\alpha=0.1)$.

附表6：F分布的分位数表 $\int_{c}^{\infty}f(x;m,n)\mathrm{d}x=\alpha(\alpha=0.05)$.

附表7：F分布的分位数表 $\int_{c}^{\infty}f(x;m,n)\mathrm{d}x=\alpha(\alpha=0.025)$.

附表8：F分布的分位数表 $\int_{c}^{\infty}f(x;m,n)\mathrm{d}x=\alpha(\alpha=0.01)$.

附表9：Kolmogorov检验统计量的分位数表$P\{D_n\geq D_\alpha(n)\}=\alpha$.

附表10：W检验统计量的系数表$(a_i(n))$

附表10(续)：W检验统计量的系数表$(a_i(n))$

附表11：W检验的下侧分位数表$P\{W_n\leq W_\alpha(n)\}=\alpha$

附表12：D检验的分位数表$P\{D\geq D_\alpha\}=\alpha$

附表13：二项分布表$Q(k;n,p)=\sum_{i=k}^{n}C_n^ip^i(1-p)^{n-i}(n=5,10,15,20)$.

附表13(续)：二项分布表$Q(k;n,p)=\sum_{i=k}^{n}C_n^ip^i(1-p)^{n-i}(n=25,30)$.

附表13仅给出了$p\leq 0.5$的分布值. 对于$p>0.5$的分布值，我们可以利用下面公式

$$Q(k;n,p)=1-Q(n-k+1,n,1-p)$$

得到.

本讲义所给的13个附表中的表9, 11和表12来自梁小筠(1997)，其余各表均由Fortran 程序包IMSL 计算而得. 但是我们注意到，本讲义表10的结果是由IMSL 求得的精确解(与模拟方法所得结果一致)，与梁小筠(1997)给出的附表有一定的区别.

另外，如读者感兴趣二元随机变量的相关性检验，则可以利用习题一的第28题的结论.

如有读者对其它一些表感兴趣，则请参考中国科学院数学研究所概率统计室(1979)及本讲义所列的部分参考文献.

附表1 标准正态分布累积分布函数值$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\mathrm{d}x$.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.10	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.20	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.30	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.40	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.50	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.60	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.70	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.80	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.90	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.00	0.84134	0.84375	0.84614	0.84850	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.10	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.20	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.30	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.40	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.50	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.60	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.70	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.80	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.90	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.00	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.10	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.20	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.30	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.40	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.50	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.60	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.70	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.80	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.90	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.00	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99897	0.99900
3.10	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.20	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.30	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.40	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.50	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983
3.60	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.70	0.99989	0.99990	0.99990	0.99990	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.80	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.90	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997

附表2 t分布的分位数表$\int_c^\infty t(x;n)\mathrm{d}x=\alpha$.

n	α								n
	0.3	0.2	0.1	0.05	0.025	0.01	0.005	0.001	
1	0.727	1.376	3.078	6.314	12.706	31.821	63.657	318.313	1
2	0.617	1.061	1.886	2.920	4.303	6.965	9.925	22.327	2
3	0.584	0.978	1.638	2.353	3.182	4.541	5.841	10.215	3
4	0.569	0.941	1.533	2.132	2.776	3.747	4.604	7.173	4
5	0.559	0.920	1.476	2.015	2.571	3.365	4.032	5.893	5
6	0.553	0.906	1.440	1.943	2.447	3.143	3.707	5.208	6
7	0.549	0.896	1.415	1.895	2.365	2.998	3.499	4.785	7
8	0.546	0.889	1.397	1.860	2.306	2.896	3.355	4.501	8
9	0.543	0.883	1.383	1.833	2.262	2.821	3.250	4.297	9
10	0.542	0.879	1.372	1.812	2.228	2.764	3.169	4.144	10
11	0.540	0.876	1.363	1.796	2.201	2.718	3.106	4.025	11
12	0.539	0.873	1.356	1.782	2.179	2.681	3.055	3.930	12
13	0.538	0.870	1.350	1.771	2.160	2.650	3.012	3.852	13
14	0.537	0.868	1.345	1.761	2.145	2.624	2.977	3.787	14
15	0.536	0.866	1.341	1.753	2.131	2.602	2.947	3.733	15
16	0.535	0.865	1.337	1.746	2.120	2.583	2.921	3.686	16
17	0.534	0.863	1.333	1.740	2.110	2.567	2.898	3.646	17
18	0.534	0.862	1.330	1.734	2.101	2.552	2.878	3.610	18
19	0.533	0.861	1.328	1.729	2.093	2.539	2.861	3.579	19
20	0.533	0.860	1.325	1.725	2.086	2.528	2.845	3.552	20
21	0.532	0.859	1.323	1.721	2.080	2.518	2.831	3.527	21
22	0.532	0.858	1.321	1.717	2.074	2.508	2.819	3.505	22
23	0.532	0.858	1.319	1.714	2.069	2.500	2.807	3.485	23
24	0.531	0.857	1.318	1.711	2.064	2.492	2.797	3.467	24
25	0.531	0.856	1.316	1.708	2.060	2.485	2.787	3.450	25
26	0.531	0.856	1.315	1.706	2.056	2.479	2.779	3.435	26
27	0.531	0.855	1.314	1.703	2.052	2.473	2.771	3.421	27
28	0.530	0.855	1.313	1.701	2.048	2.467	2.763	3.408	28
29	0.530	0.854	1.311	1.699	2.045	2.462	2.756	3.396	29
30	0.530	0.854	1.310	1.697	2.042	2.457	2.750	3.385	30
35	0.529	0.852	1.306	1.690	2.030	2.438	2.724	3.340	35
40	0.529	0.851	1.303	1.684	2.021	2.423	2.704	3.307	40
45	0.528	0.850	1.301	1.679	2.014	2.412	2.690	3.281	45
50	0.528	0.849	1.299	1.676	2.009	2.403	2.678	3.261	50
100	0.526	0.845	1.290	1.660	1.984	2.364	2.626	3.174	100
150	0.526	0.844	1.287	1.655	1.976	2.351	2.609	3.145	150
200	0.525	0.843	1.286	1.653	1.972	2.345	2.601	3.131	200
∞	0.524	0.842	1.282	1.645	1.960	2.326	2.576	3.090	∞

附表3 χ^2分布的分位数表$\int_c^\infty \chi^2(x;n)\mathrm{d}x = \alpha$.

n	α										n
	0.995	0.990	0.975	0.950	0.900	0.100	0.050	0.025	0.010	0.005	
1	0.00	0.00	0.00	0.00	0.02	2.71	3.84	5.02	6.64	7.88	1
2	0.01	0.02	0.05	0.10	0.21	4.61	5.99	7.38	9.21	10.60	2
3	0.07	0.11	0.22	0.35	0.58	6.25	7.81	9.35	11.35	12.84	3
4	0.21	0.30	0.48	0.71	1.06	7.78	9.49	11.14	13.28	14.86	4
5	0.41	0.55	0.83	1.15	1.61	9.24	11.07	12.83	15.09	16.75	5
6	0.68	0.87	1.24	1.64	2.20	10.64	12.59	14.45	16.81	18.55	6
7	0.99	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48	20.28	7
8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09	21.96	8
9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67	23.59	9
10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21	25.19	10
11	2.60	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.73	26.76	11
12	3.07	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22	28.30	12
13	3.57	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69	29.82	13
14	4.07	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14	31.32	14
15	4.60	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58	32.80	15
16	5.14	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00	34.27	16
17	5.70	6.41	7.56	8.67	10.09	24.77	27.59	30.19	33.41	35.72	17
18	6.26	7.01	8.23	9.39	10.86	25.99	28.87	31.53	34.81	37.16	18
19	6.84	7.63	8.91	10.12	11.65	27.20	30.14	32.85	36.19	38.58	19
20	7.43	8.26	9.59	10.85	12.44	28.41	31.41	34.17	37.57	40.00	20
21	8.03	8.90	10.28	11.59	13.24	29.62	32.67	35.48	38.93	41.40	21
22	8.64	9.54	10.98	12.34	14.04	30.81	33.92	36.78	40.29	42.80	22
23	9.26	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64	44.18	23
24	9.89	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56	24
25	10.52	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31	46.93	25
26	11.16	12.20	13.84	15.38	17.29	35.56	38.89	41.92	45.64	48.29	26
27	11.81	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96	49.64	27
28	12.46	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28	50.99	28
29	13.12	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59	52.34	29
30	13.79	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67	30
31	14.46	15.66	17.54	19.28	21.43	41.42	44.99	48.23	52.19	55.00	31
32	15.13	16.36	18.29	20.07	22.27	42.58	46.19	49.48	53.49	56.33	32
33	15.82	17.07	19.05	20.87	23.11	43.75	47.40	50.73	54.78	57.65	33
34	16.50	17.79	19.81	21.66	23.95	44.90	48.60	51.97	56.06	58.96	34
35	17.19	18.51	20.57	22.47	24.80	46.06	49.80	53.20	57.34	60.27	35
36	17.89	19.23	21.34	23.27	25.64	47.21	51.00	54.44	58.62	61.58	36
37	18.59	19.96	22.11	24.07	26.49	48.36	52.19	55.67	59.89	62.88	37
38	19.29	20.69	22.88	24.88	27.34	49.51	53.38	56.90	61.16	64.18	38
39	20.00	21.43	23.65	25.70	28.20	50.66	54.57	58.12	62.43	65.48	39
40	20.71	22.16	24.43	26.51	29.05	51.81	55.76	59.34	63.69	66.77	40

附表4 F分布的分位数表$\int_c^\infty f(x;m,n)\mathrm{d}x=\alpha$. ($\alpha=0.2$)

	n															
m	1	2	3	4	5	6	7	8	9	10	15	20	30	50	100	∞
1	9.5	3.56	2.68	2.35	2.18	2.07	2.00	1.95	1.91	1.88	1.80	1.76	1.72	1.69	1.66	1.64
2	12.0	4.00	2.89	2.47	2.26	2.13	2.04	1.98	1.93	1.90	1.80	1.75	1.70	1.66	1.64	1.61
3	13.1	4.16	2.94	2.48	2.25	2.11	2.02	1.95	1.90	1.86	1.75	1.70	1.64	1.60	1.58	1.55
4	13.6	4.24	2.96	2.48	2.24	2.09	1.99	1.92	1.87	1.83	1.71	1.65	1.60	1.56	1.53	1.50
5	14.0	4.28	2.97	2.48	2.23	2.08	1.97	1.90	1.85	1.80	1.68	1.62	1.57	1.52	1.49	1.46
6	14.3	4.32	2.97	2.47	2.22	2.06	1.96	1.88	1.83	1.78	1.66	1.60	1.54	1.49	1.46	1.43
7	14.4	4.34	2.97	2.47	2.21	2.05	1.94	1.87	1.81	1.77	1.64	1.58	1.52	1.47	1.43	1.40
8	14.6	4.36	2.98	2.47	2.20	2.04	1.93	1.86	1.80	1.75	1.62	1.56	1.50	1.45	1.41	1.38
9	14.7	4.37	2.98	2.46	2.20	2.03	1.93	1.85	1.79	1.74	1.61	1.54	1.48	1.43	1.40	1.36
10	14.8	4.38	2.98	2.46	2.19	2.03	1.92	1.84	1.78	1.73	1.60	1.53	1.47	1.42	1.38	1.34
11	14.8	4.39	2.98	2.46	2.19	2.02	1.91	1.83	1.77	1.72	1.59	1.52	1.46	1.41	1.37	1.33
12	14.9	4.40	2.98	2.46	2.18	2.02	1.91	1.83	1.76	1.72	1.58	1.51	1.45	1.39	1.36	1.32
13	15.0	4.40	2.98	2.45	2.18	2.01	1.90	1.82	1.76	1.71	1.57	1.50	1.44	1.38	1.35	1.31
14	15.0	4.41	2.98	2.45	2.18	2.01	1.90	1.82	1.75	1.70	1.56	1.50	1.43	1.38	1.34	1.30
15	15.0	4.42	2.98	2.45	2.18	2.01	1.89	1.81	1.75	1.70	1.56	1.49	1.42	1.37	1.33	1.29
16	15.1	4.42	2.98	2.45	2.17	2.00	1.89	1.81	1.74	1.70	1.55	1.48	1.42	1.36	1.32	1.28
17	15.1	4.42	2.98	2.45	2.17	2.00	1.89	1.80	1.74	1.69	1.55	1.48	1.41	1.35	1.31	1.27
18	15.1	4.43	2.98	2.45	2.17	2.00	1.88	1.80	1.74	1.69	1.54	1.47	1.40	1.35	1.31	1.26
19	15.1	4.43	2.98	2.45	2.17	2.00	1.88	1.80	1.73	1.69	1.54	1.47	1.40	1.34	1.30	1.26
20	15.2	4.43	2.98	2.44	2.17	2.00	1.88	1.80	1.73	1.68	1.54	1.47	1.39	1.34	1.30	1.25
21	15.2	4.43	2.98	2.44	2.16	1.99	1.88	1.79	1.73	1.68	1.53	1.46	1.39	1.33	1.29	1.25
22	15.2	4.44	2.98	2.44	2.16	1.99	1.88	1.79	1.73	1.68	1.53	1.46	1.39	1.33	1.29	1.24
23	15.2	4.44	2.98	2.44	2.16	1.99	1.87	1.79	1.73	1.67	1.53	1.46	1.38	1.33	1.28	1.24
24	15.2	4.44	2.98	2.44	2.16	1.99	1.87	1.79	1.72	1.67	1.53	1.45	1.38	1.32	1.28	1.23
25	15.3	4.44	2.98	2.44	2.16	1.99	1.87	1.79	1.72	1.67	1.52	1.45	1.38	1.32	1.27	1.23
26	15.3	4.44	2.98	2.44	2.16	1.99	1.87	1.78	1.72	1.67	1.52	1.45	1.37	1.31	1.27	1.22
27	15.3	4.44	2.98	2.44	2.16	1.99	1.87	1.78	1.72	1.67	1.52	1.44	1.37	1.31	1.27	1.22
28	15.3	4.45	2.98	2.44	2.16	1.98	1.87	1.78	1.72	1.67	1.52	1.44	1.37	1.31	1.26	1.22
29	15.3	4.45	2.98	2.44	2.16	1.98	1.87	1.78	1.72	1.66	1.51	1.44	1.37	1.31	1.26	1.21
30	15.3	4.45	2.98	2.44	2.16	1.98	1.86	1.78	1.71	1.66	1.51	1.44	1.36	1.30	1.26	1.21
35	15.3	4.45	2.98	2.44	2.15	1.98	1.86	1.77	1.71	1.66	1.51	1.43	1.35	1.29	1.24	1.19
40	15.4	4.46	2.98	2.44	2.15	1.98	1.86	1.77	1.70	1.65	1.50	1.42	1.35	1.28	1.23	1.18
45	15.4	4.46	2.98	2.44	2.15	1.97	1.85	1.77	1.70	1.65	1.50	1.42	1.34	1.28	1.23	1.17
50	15.4	4.46	2.98	2.43	2.15	1.97	1.85	1.76	1.70	1.65	1.49	1.41	1.34	1.27	1.22	1.16
60	15.4	4.46	2.98	2.43	2.15	1.97	1.85	1.76	1.69	1.64	1.49	1.41	1.33	1.26	1.21	1.15
70	15.5	4.47	2.98	2.43	2.14	1.97	1.85	1.76	1.69	1.64	1.48	1.40	1.32	1.25	1.20	1.14
80	15.5	4.47	2.98	2.43	2.14	1.97	1.84	1.76	1.69	1.64	1.48	1.40	1.32	1.25	1.19	1.13
100	15.5	4.47	2.98	2.43	2.14	1.96	1.84	1.75	1.69	1.63	1.47	1.39	1.31	1.24	1.18	1.12
150	15.5	4.47	2.98	2.43	2.14	1.96	1.84	1.75	1.68	1.63	1.47	1.39	1.30	1.23	1.17	1.10
200	15.5	4.48	2.98	2.43	2.14	1.96	1.84	1.75	1.68	1.63	1.46	1.38	1.30	1.22	1.16	1.08
∞	15.6	4.48	2.98	2.43	2.13	1.95	1.83	1.74	1.67	1.62	1.46	1.37	1.28	1.21	1.14	1.00

附表5 F分布的分位数表$\int_c^\infty f(x;m,n)\mathrm{d}x=\alpha$. ($\alpha=0.1$)

m \ n	1	2	3	4	5	6	7	8	9	10	15	20	30	50	100	∞
1	39.86	8.53	5.54	4.54	4.06	3.78	3.59	3.46	3.36	3.29	3.07	2.97	2.88	2.81	2.76	2.71
2	49.50	9.00	5.46	4.32	3.78	3.46	3.26	3.11	3.01	2.92	2.70	2.59	2.49	2.41	2.36	2.30
3	53.59	9.16	5.39	4.19	3.62	3.29	3.07	2.92	2.81	2.73	2.49	2.38	2.28	2.20	2.14	2.08
4	55.83	9.24	5.34	4.11	3.52	3.18	2.96	2.81	2.69	2.61	2.36	2.25	2.14	2.06	2.00	1.94
5	57.24	9.29	5.31	4.05	3.45	3.11	2.88	2.73	2.61	2.52	2.27	2.16	2.05	1.97	1.91	1.85
6	58.20	9.33	5.28	4.01	3.40	3.05	2.83	2.67	2.55	2.46	2.21	2.09	1.98	1.90	1.83	1.77
7	58.91	9.35	5.27	3.98	3.37	3.01	2.78	2.62	2.51	2.41	2.16	2.04	1.93	1.84	1.78	1.72
8	59.44	9.37	5.25	3.95	3.34	2.98	2.75	2.59	2.47	2.38	2.12	2.00	1.88	1.80	1.73	1.67
9	59.86	9.38	5.24	3.94	3.32	2.96	2.72	2.56	2.44	2.35	2.09	1.96	1.85	1.76	1.69	1.63
10	60.19	9.39	5.23	3.92	3.30	2.94	2.70	2.54	2.42	2.32	2.06	1.94	1.82	1.73	1.66	1.60
11	60.47	9.40	5.22	3.91	3.28	2.92	2.68	2.52	2.40	2.30	2.04	1.91	1.79	1.70	1.64	1.57
12	60.71	9.41	5.22	3.90	3.27	2.90	2.67	2.50	2.38	2.28	2.02	1.89	1.77	1.68	1.61	1.55
13	60.90	9.41	5.21	3.89	3.26	2.89	2.65	2.49	2.36	2.27	2.00	1.87	1.75	1.66	1.59	1.52
14	61.07	9.42	5.20	3.88	3.25	2.88	2.64	2.48	2.35	2.26	1.99	1.86	1.74	1.64	1.57	1.50
15	61.22	9.42	5.20	3.87	3.24	2.87	2.63	2.46	2.34	2.24	1.97	1.84	1.72	1.63	1.56	1.49
16	61.35	9.43	5.20	3.86	3.23	2.86	2.62	2.45	2.33	2.23	1.96	1.83	1.71	1.61	1.54	1.47
17	61.46	9.43	5.19	3.86	3.22	2.85	2.61	2.45	2.32	2.22	1.95	1.82	1.70	1.60	1.53	1.46
18	61.57	9.44	5.19	3.85	3.22	2.85	2.61	2.44	2.31	2.22	1.94	1.81	1.69	1.59	1.52	1.44
19	61.66	9.44	5.19	3.85	3.21	2.84	2.60	2.43	2.30	2.21	1.93	1.80	1.68	1.58	1.50	1.43
20	61.74	9.44	5.18	3.84	3.21	2.84	2.59	2.42	2.30	2.20	1.92	1.79	1.67	1.57	1.49	1.42
21	61.81	9.44	5.18	3.84	3.20	2.83	2.59	2.42	2.29	2.19	1.92	1.79	1.66	1.56	1.48	1.41
22	61.88	9.45	5.18	3.84	3.20	2.83	2.58	2.41	2.29	2.19	1.91	1.78	1.65	1.55	1.48	1.40
23	61.95	9.45	5.18	3.83	3.19	2.82	2.58	2.41	2.28	2.18	1.90	1.77	1.64	1.54	1.47	1.39
24	62.00	9.45	5.18	3.83	3.19	2.82	2.58	2.40	2.28	2.18	1.90	1.77	1.64	1.54	1.46	1.38
25	62.05	9.45	5.17	3.83	3.19	2.81	2.57	2.40	2.27	2.17	1.89	1.76	1.63	1.53	1.45	1.38
26	62.10	9.45	5.17	3.83	3.18	2.81	2.57	2.40	2.27	2.17	1.89	1.76	1.63	1.52	1.45	1.37
27	62.15	9.45	5.17	3.82	3.18	2.81	2.56	2.39	2.26	2.17	1.88	1.75	1.62	1.52	1.44	1.36
28	62.19	9.46	5.17	3.82	3.18	2.81	2.56	2.39	2.26	2.16	1.88	1.75	1.62	1.51	1.43	1.35
29	62.23	9.46	5.17	3.82	3.18	2.80	2.56	2.39	2.26	2.16	1.88	1.74	1.61	1.51	1.43	1.35
30	62.26	9.46	5.17	3.82	3.17	2.80	2.56	2.38	2.25	2.16	1.87	1.74	1.61	1.50	1.42	1.34
35	62.42	9.46	5.16	3.81	3.16	2.79	2.54	2.37	2.24	2.14	1.86	1.72	1.59	1.48	1.40	1.32
40	62.53	9.47	5.16	3.80	3.16	2.78	2.54	2.36	2.23	2.13	1.85	1.71	1.57	1.46	1.38	1.30
45	62.62	9.47	5.16	3.80	3.15	2.77	2.53	2.35	2.22	2.12	1.84	1.70	1.56	1.45	1.37	1.28
50	62.69	9.47	5.15	3.80	3.15	2.77	2.52	2.35	2.22	2.12	1.83	1.69	1.55	1.44	1.35	1.26
60	62.79	9.47	5.15	3.79	3.14	2.76	2.51	2.34	2.21	2.11	1.82	1.68	1.54	1.42	1.34	1.24
70	62.87	9.48	5.15	3.79	3.14	2.76	2.51	2.33	2.20	2.10	1.81	1.67	1.53	1.41	1.32	1.22
80	62.93	9.48	5.15	3.78	3.13	2.75	2.50	2.33	2.20	2.09	1.80	1.66	1.52	1.40	1.31	1.21
100	63.01	9.48	5.14	3.78	3.13	2.75	2.50	2.32	2.19	2.09	1.79	1.65	1.51	1.39	1.29	1.18
150	63.11	9.48	5.14	3.77	3.12	2.74	2.49	2.31	2.18	2.08	1.78	1.64	1.49	1.37	1.27	1.15
200	63.17	9.49	5.14	3.77	3.12	2.73	2.48	2.31	2.17	2.07	1.77	1.63	1.48	1.36	1.26	1.13
∞	63.33	9.49	5.13	3.76	3.10	2.72	2.47	2.29	2.16	2.06	1.76	1.61	1.46	1.33	1.21	1.00

附表6 F分布的分位数表$\int_c^{\infty} f(x;m,n)\mathrm{d}x=\alpha$. ($\alpha=0.05$)

m	n: 1	2	3	4	5	6	7	8	9	10	15	20	30	50	100	∞
1	161.45	18.51	10.13	7.71	6.61	5.99	5.59	5.32	5.12	4.96	4.54	4.35	4.17	4.03	3.94	3.84
2	199.50	19.00	9.55	6.94	5.79	5.14	4.74	4.46	4.26	4.10	3.68	3.49	3.32	3.18	3.09	3.00
3	215.71	19.16	9.28	6.59	5.41	4.76	4.35	4.07	3.86	3.71	3.29	3.10	2.92	2.79	2.70	2.60
4	224.58	19.25	9.12	6.39	5.19	4.53	4.12	3.84	3.63	3.48	3.06	2.87	2.69	2.56	2.46	2.37
5	230.16	19.30	9.01	6.26	5.05	4.39	3.97	3.69	3.48	3.33	2.90	2.71	2.53	2.40	2.31	2.21
6	233.99	19.33	8.94	6.16	4.95	4.28	3.87	3.58	3.37	3.22	2.79	2.60	2.42	2.29	2.19	2.10
7	236.77	19.35	8.89	6.09	4.88	4.21	3.79	3.50	3.29	3.14	2.71	2.51	2.33	2.20	2.10	2.01
8	238.88	19.37	8.85	6.04	4.82	4.15	3.73	3.44	3.23	3.07	2.64	2.45	2.27	2.13	2.03	1.94
9	240.54	19.38	8.81	6.00	4.77	4.10	3.68	3.39	3.18	3.02	2.59	2.39	2.21	2.07	1.97	1.88
10	241.88	19.40	8.79	5.96	4.74	4.06	3.64	3.35	3.14	2.98	2.54	2.35	2.16	2.03	1.93	1.83
11	242.98	19.40	8.76	5.94	4.70	4.03	3.60	3.31	3.10	2.94	2.51	2.31	2.13	1.99	1.89	1.79
12	243.91	19.41	8.74	5.91	4.68	4.00	3.57	3.28	3.07	2.91	2.48	2.28	2.09	1.95	1.85	1.75
13	244.69	19.42	8.73	5.89	4.66	3.98	3.55	3.26	3.05	2.89	2.45	2.25	2.06	1.92	1.82	1.72
14	245.36	19.42	8.71	5.87	4.64	3.96	3.53	3.24	3.03	2.86	2.42	2.22	2.04	1.89	1.79	1.69
15	245.95	19.43	8.70	5.86	4.62	3.94	3.51	3.22	3.01	2.85	2.40	2.20	2.01	1.87	1.77	1.67
16	246.46	19.43	8.69	5.84	4.60	3.92	3.49	3.20	2.99	2.83	2.38	2.18	1.99	1.85	1.75	1.64
17	246.92	19.44	8.68	5.83	4.59	3.91	3.48	3.19	2.97	2.81	2.37	2.17	1.98	1.83	1.73	1.62
18	247.32	19.44	8.67	5.82	4.58	3.90	3.47	3.17	2.96	2.80	2.35	2.15	1.96	1.81	1.71	1.60
19	247.69	19.44	8.67	5.81	4.57	3.88	3.46	3.16	2.95	2.79	2.34	2.14	1.95	1.80	1.69	1.59
20	248.01	19.45	8.66	5.80	4.56	3.87	3.44	3.15	2.94	2.77	2.33	2.12	1.93	1.78	1.68	1.57
21	248.31	19.45	8.65	5.79	4.55	3.86	3.43	3.14	2.93	2.76	2.32	2.11	1.92	1.77	1.66	1.56
22	248.58	19.45	8.65	5.79	4.54	3.86	3.43	3.13	2.92	2.75	2.31	2.10	1.91	1.76	1.65	1.54
23	248.83	19.45	8.64	5.78	4.53	3.85	3.42	3.12	2.91	2.75	2.30	2.09	1.90	1.75	1.64	1.53
24	249.05	19.45	8.64	5.77	4.53	3.84	3.41	3.12	2.90	2.74	2.29	2.08	1.89	1.74	1.63	1.52
25	249.26	19.46	8.63	5.77	4.52	3.83	3.40	3.11	2.89	2.73	2.28	2.07	1.88	1.73	1.62	1.51
26	249.45	19.46	8.63	5.76	4.52	3.83	3.40	3.10	2.89	2.72	2.27	2.07	1.87	1.72	1.61	1.50
27	249.63	19.46	8.63	5.76	4.51	3.82	3.39	3.10	2.88	2.72	2.27	2.06	1.86	1.71	1.60	1.49
28	249.80	19.46	8.62	5.75	4.50	3.82	3.39	3.09	2.87	2.71	2.26	2.05	1.85	1.70	1.59	1.48
29	249.95	19.46	8.62	5.75	4.50	3.81	3.38	3.08	2.87	2.70	2.25	2.05	1.85	1.69	1.58	1.47
30	250.10	19.46	8.62	5.75	4.50	3.81	3.38	3.08	2.86	2.70	2.25	2.04	1.84	1.69	1.57	1.46
35	250.69	19.47	8.60	5.73	4.48	3.79	3.36	3.06	2.84	2.68	2.22	2.01	1.81	1.66	1.54	1.42
40	251.14	19.47	8.59	5.72	4.46	3.77	3.34	3.04	2.83	2.66	2.20	1.99	1.79	1.63	1.52	1.39
45	251.49	19.47	8.59	5.71	4.45	3.76	3.33	3.03	2.81	2.65	2.19	1.98	1.77	1.61	1.49	1.37
50	251.77	19.48	8.58	5.70	4.44	3.75	3.32	3.02	2.80	2.64	2.18	1.97	1.76	1.60	1.48	1.35
60	252.20	19.48	8.57	5.69	4.43	3.74	3.30	3.01	2.79	2.62	2.16	1.95	1.74	1.58	1.45	1.32
70	252.50	19.48	8.57	5.68	4.42	3.73	3.29	2.99	2.78	2.61	2.15	1.93	1.72	1.56	1.43	1.29
80	252.72	19.48	8.56	5.67	4.41	3.72	3.29	2.99	2.77	2.60	2.14	1.92	1.71	1.54	1.41	1.27
100	253.04	19.49	8.55	5.66	4.41	3.71	3.27	2.97	2.76	2.59	2.12	1.91	1.70	1.52	1.39	1.24
150	253.46	19.49	8.54	5.65	4.39	3.70	3.26	2.96	2.74	2.57	2.10	1.89	1.67	1.50	1.36	1.20
200	253.68	19.49	8.54	5.65	4.39	3.69	3.25	2.95	2.73	2.56	2.10	1.88	1.66	1.48	1.34	1.17
∞	254.31	19.50	8.53	5.63	4.36	3.67	3.23	2.93	2.71	2.54	2.07	1.84	1.62	1.44	1.28	1.00

附表7 F分布的分位数表$\int_c^{\infty} f(x;m,n)\mathrm{d}x = \alpha$. ($\alpha = 0.025$)

	n															
m	1	2	3	4	5	6	7	8	9	10	15	20	30	50	100	∞
1	647.8	38.51	17.44	12.22	10.01	8.81	8.07	7.57	7.21	6.94	6.20	5.87	5.57	5.34	5.18	5.02
2	799.5	39.00	16.04	10.65	8.43	7.26	6.54	6.06	5.71	5.46	4.77	4.46	4.18	3.97	3.83	3.69
3	864.2	39.17	15.44	9.98	7.76	6.60	5.89	5.42	5.08	4.83	4.15	3.86	3.59	3.39	3.25	3.12
4	899.6	39.25	15.10	9.60	7.39	6.23	5.52	5.05	4.72	4.47	3.80	3.51	3.25	3.05	2.92	2.79
5	921.8	39.30	14.88	9.36	7.15	5.99	5.29	4.82	4.48	4.24	3.58	3.29	3.03	2.83	2.70	2.57
6	937.1	39.33	14.73	9.20	6.98	5.82	5.12	4.65	4.32	4.07	3.41	3.13	2.87	2.67	2.54	2.41
7	948.2	39.36	14.62	9.07	6.85	5.70	4.99	4.53	4.20	3.95	3.29	3.01	2.75	2.55	2.42	2.29
8	956.7	39.37	14.54	8.98	6.76	5.60	4.90	4.43	4.10	3.85	3.20	2.91	2.65	2.46	2.32	2.19
9	963.3	39.39	14.47	8.90	6.68	5.52	4.82	4.36	4.03	3.78	3.12	2.84	2.57	2.38	2.24	2.11
10	968.6	39.40	14.42	8.84	6.62	5.46	4.76	4.30	3.96	3.72	3.06	2.77	2.51	2.32	2.18	2.05
11	973.0	39.41	14.37	8.79	6.57	5.41	4.71	4.24	3.91	3.66	3.01	2.72	2.46	2.26	2.12	1.99
12	976.7	39.41	14.34	8.75	6.52	5.37	4.67	4.20	3.87	3.62	2.96	2.68	2.41	2.22	2.08	1.94
13	979.8	39.42	14.30	8.72	6.49	5.33	4.63	4.16	3.83	3.58	2.92	2.64	2.37	2.18	2.04	1.90
14	982.5	39.43	14.28	8.68	6.46	5.30	4.60	4.13	3.80	3.55	2.89	2.60	2.34	2.14	2.00	1.87
15	984.9	39.43	14.25	8.66	6.43	5.27	4.57	4.10	3.77	3.52	2.86	2.57	2.31	2.11	1.97	1.83
16	986.9	39.44	14.23	8.63	6.40	5.24	4.54	4.08	3.74	3.50	2.84	2.55	2.28	2.08	1.94	1.80
17	988.7	39.44	14.21	8.61	6.38	5.22	4.52	4.05	3.72	3.47	2.81	2.52	2.26	2.06	1.91	1.78
18	990.4	39.44	14.20	8.59	6.36	5.20	4.50	4.03	3.70	3.45	2.79	2.50	2.23	2.03	1.89	1.75
19	991.8	39.45	14.18	8.58	6.34	5.18	4.48	4.02	3.68	3.44	2.77	2.48	2.21	2.01	1.87	1.73
20	993.1	39.45	14.17	8.56	6.33	5.17	4.47	4.00	3.67	3.42	2.76	2.46	2.20	1.99	1.85	1.71
21	994.3	39.45	14.16	8.55	6.31	5.15	4.45	3.98	3.65	3.40	2.74	2.45	2.18	1.98	1.83	1.69
22	995.4	39.45	14.14	8.53	6.30	5.14	4.44	3.97	3.64	3.39	2.73	2.43	2.16	1.96	1.81	1.67
23	996.3	39.45	14.13	8.52	6.29	5.13	4.43	3.96	3.63	3.38	2.71	2.42	2.15	1.95	1.80	1.66
24	997.3	39.46	14.12	8.51	6.28	5.12	4.41	3.95	3.61	3.37	2.70	2.41	2.14	1.93	1.78	1.64
25	998.1	39.46	14.12	8.50	6.27	5.11	4.40	3.94	3.60	3.35	2.69	2.40	2.12	1.92	1.77	1.63
26	998.9	39.46	14.11	8.49	6.26	5.10	4.39	3.93	3.59	3.34	2.68	2.39	2.11	1.91	1.76	1.61
27	999.6	39.46	14.10	8.48	6.25	5.09	4.39	3.92	3.58	3.34	2.67	2.38	2.10	1.90	1.75	1.60
28	1000.2	39.46	14.09	8.48	6.24	5.08	4.38	3.91	3.58	3.33	2.66	2.37	2.09	1.89	1.74	1.59
29	1000.8	39.46	14.09	8.47	6.23	5.07	4.37	3.90	3.57	3.32	2.65	2.36	2.08	1.88	1.72	1.58
30	1001.4	39.46	14.08	8.46	6.23	5.07	4.36	3.89	3.56	3.31	2.64	2.35	2.07	1.87	1.71	1.57
35	1003.8	39.47	14.06	8.43	6.20	5.04	4.33	3.86	3.53	3.28	2.61	2.31	2.04	1.83	1.67	1.52
40	1005.6	39.47	14.04	8.41	6.18	5.01	4.31	3.84	3.51	3.26	2.59	2.29	2.01	1.80	1.64	1.48
45	1007.0	39.48	14.02	8.39	6.16	4.99	4.29	3.82	3.49	3.24	2.56	2.27	1.99	1.77	1.61	1.45
50	1008.1	39.48	14.01	8.38	6.14	4.98	4.28	3.81	3.47	3.22	2.55	2.25	1.97	1.75	1.59	1.43
60	1009.8	39.48	13.99	8.36	6.12	4.96	4.25	3.78	3.45	3.20	2.52	2.22	1.94	1.72	1.56	1.39
70	1011.0	39.48	13.98	8.35	6.11	4.94	4.24	3.77	3.43	3.18	2.51	2.20	1.92	1.70	1.53	1.36
80	1011.9	39.49	13.97	8.33	6.10	4.93	4.23	3.76	3.42	3.17	2.49	2.19	1.90	1.68	1.51	1.33
100	1013.2	39.49	13.96	8.32	6.08	4.92	4.21	3.74	3.40	3.15	2.47	2.17	1.88	1.66	1.48	1.30
150	1014.9	39.49	13.94	8.30	6.06	4.89	4.19	3.72	3.38	3.13	2.45	2.14	1.85	1.62	1.44	1.24
200	1015.7	39.49	13.93	8.29	6.05	4.88	4.18	3.70	3.37	3.12	2.44	2.13	1.84	1.60	1.42	1.21
∞	1018.3	39.50	13.90	8.26	6.02	4.85	4.14	3.67	3.33	3.08	2.40	2.09	1.79	1.55	1.35	1.00

附表8 F分布的分位数表$\int_c^{\infty} f(x;m,n)\mathrm{d}x = \alpha.$ $(\alpha = 0.01)$

	n															
m	1	2	3	4	5	6	7	8	9	10	15	20	30	50	100	∞
1	4052.2	98.50	34.12	21.20	16.26	13.75	12.25	11.26	10.56	10.04	8.68	8.10	7.56	7.17	6.90	6.64
2	4999.5	99.00	30.82	18.00	13.27	10.92	9.55	8.65	8.02	7.56	6.36	5.85	5.39	5.06	4.82	4.61
3	5403.4	99.17	29.46	16.69	12.06	9.78	8.45	7.59	6.99	6.55	5.42	4.94	4.51	4.20	3.98	3.78
4	5624.6	99.25	28.71	15.98	11.39	9.15	7.85	7.01	6.42	5.99	4.89	4.43	4.02	3.72	3.51	3.32
5	5763.7	99.30	28.24	15.52	10.97	8.75	7.46	6.63	6.06	5.64	4.56	4.10	3.70	3.41	3.21	3.02
6	5859.0	99.33	27.91	15.21	10.67	8.47	7.19	6.37	5.80	5.39	4.32	3.87	3.47	3.19	2.99	2.80
7	5928.4	99.36	27.67	14.98	10.46	8.26	6.99	6.18	5.61	5.20	4.14	3.70	3.30	3.02	2.82	2.64
8	5981.1	99.37	27.49	14.80	10.29	8.10	6.84	6.03	5.47	5.06	4.00	3.56	3.17	2.89	2.69	2.51
9	6022.5	99.39	27.35	14.66	10.16	7.98	6.72	5.91	5.35	4.94	3.89	3.46	3.07	2.78	2.59	2.41
10	6055.9	99.40	27.23	14.55	10.05	7.87	6.62	5.81	5.26	4.85	3.80	3.37	2.98	2.70	2.50	2.32
11	6083.3	99.41	27.13	14.45	9.96	7.79	6.54	5.73	5.18	4.77	3.73	3.29	2.91	2.63	2.43	2.25
12	6106.3	99.42	27.05	14.37	9.89	7.72	6.47	5.67	5.11	4.71	3.67	3.23	2.84	2.56	2.37	2.18
13	6125.9	99.42	26.98	14.31	9.82	7.66	6.41	5.61	5.05	4.65	3.61	3.18	2.79	2.51	2.31	2.13
14	6142.7	99.43	26.92	14.25	9.77	7.60	6.36	5.56	5.01	4.60	3.56	3.13	2.74	2.46	2.27	2.08
15	6157.3	99.43	26.87	14.20	9.72	7.56	6.31	5.52	4.96	4.56	3.52	3.09	2.70	2.42	2.22	2.04
16	6170.1	99.44	26.83	14.15	9.68	7.52	6.28	5.48	4.92	4.52	3.49	3.05	2.66	2.38	2.19	2.00
17	6181.4	99.44	26.79	14.11	9.64	7.48	6.24	5.44	4.89	4.49	3.45	3.02	2.63	2.35	2.15	1.97
18	6191.5	99.44	26.75	14.08	9.61	7.45	6.21	5.41	4.86	4.46	3.42	2.99	2.60	2.32	2.12	1.93
19	6200.6	99.45	26.72	14.05	9.58	7.42	6.18	5.38	4.83	4.43	3.40	2.96	2.57	2.29	2.09	1.90
20	6208.7	99.45	26.69	14.02	9.55	7.40	6.16	5.36	4.81	4.41	3.37	2.94	2.55	2.27	2.07	1.88
21	6216.1	99.45	26.66	13.99	9.53	7.37	6.13	5.34	4.79	4.38	3.35	2.92	2.53	2.24	2.04	1.85
22	6222.9	99.45	26.64	13.97	9.51	7.35	6.11	5.32	4.77	4.36	3.33	2.90	2.51	2.22	2.02	1.83
23	6229.0	99.46	26.62	13.95	9.49	7.33	6.09	5.30	4.75	4.34	3.31	2.88	2.49	2.20	2.00	1.81
24	6234.6	99.46	26.60	13.93	9.47	7.31	6.07	5.28	4.73	4.33	3.29	2.86	2.47	2.18	1.98	1.79
25	6239.8	99.46	26.58	13.91	9.45	7.30	6.06	5.26	4.71	4.31	3.28	2.84	2.45	2.17	1.97	1.77
26	6244.6	99.46	26.56	13.89	9.43	7.28	6.04	5.25	4.70	4.30	3.26	2.83	2.44	2.15	1.95	1.76
27	6249.1	99.46	26.55	13.88	9.42	7.27	6.03	5.23	4.68	4.28	3.25	2.81	2.42	2.14	1.93	1.74
28	6253.2	99.46	26.53	13.86	9.40	7.25	6.02	5.22	4.67	4.27	3.24	2.80	2.41	2.12	1.92	1.72
29	6257.1	99.46	26.52	13.85	9.39	7.24	6.00	5.21	4.66	4.26	3.23	2.79	2.40	2.11	1.91	1.71
30	6260.7	99.47	26.50	13.84	9.38	7.23	5.99	5.20	4.65	4.25	3.21	2.78	2.39	2.10	1.89	1.70
35	6275.6	99.47	26.45	13.79	9.33	7.18	5.94	5.15	4.60	4.20	3.17	2.73	2.34	2.05	1.84	1.64
40	6286.8	99.47	26.41	13.75	9.29	7.14	5.91	5.12	4.57	4.17	3.13	2.69	2.30	2.01	1.80	1.59
45	6295.5	99.48	26.38	13.71	9.26	7.11	5.88	5.09	4.54	4.14	3.10	2.67	2.27	1.97	1.76	1.55
50	6302.5	99.48	26.35	13.69	9.24	7.09	5.86	5.07	4.52	4.12	3.08	2.64	2.25	1.95	1.74	1.52
60	6313.0	99.48	26.32	13.65	9.20	7.06	5.82	5.03	4.48	4.08	3.05	2.61	2.21	1.91	1.69	1.47
70	6320.6	99.48	26.29	13.63	9.18	7.03	5.80	5.01	4.46	4.06	3.02	2.58	2.18	1.88	1.66	1.43
80	6326.2	99.49	26.27	13.61	9.16	7.01	5.78	4.99	4.44	4.04	3.00	2.56	2.16	1.86	1.63	1.40
100	6334.1	99.49	26.24	13.58	9.13	6.99	5.75	4.96	4.41	4.01	2.98	2.54	2.13	1.82	1.60	1.36
150	6344.7	99.49	26.20	13.54	9.09	6.95	5.72	4.93	4.38	3.98	2.94	2.50	2.09	1.78	1.55	1.29
200	6350.0	99.49	26.18	13.52	9.08	6.93	5.70	4.91	4.36	3.96	2.92	2.48	2.07	1.76	1.52	1.25
∞	6365.9	99.50	26.13	13.46	9.02	6.88	5.65	4.86	4.31	3.91	2.87	2.42	2.01	1.68	1.43	1.00

附表9 Kolmogorov检验分位数表$P\{D_n \geq D_\alpha(n)\} = \alpha$.

α	0.2	0.1	0.05	0.02	0.01
n=1	0.9000	0.9500	0.9750	0.9900	0.9950
2	0.6838	0.7764	0.8419	0.9000	0.9293
3	0.5648	0.6360	0.7076	0.7846	0.8290
4	0.4927	0.5652	0.6239	0.6889	0.7342
5	0.4470	0.5095	0.5633	0.6271	0.6685
6	0.4104	0.4680	0.5193	0.5774	0.6166
7	0.3815	0.4361	0.4834	0.5384	0.5758
8	0.3583	0.4096	0.4543	0.5065	0.5418
9	0.3391	0.3875	0.4300	0.4796	0.5133
10	0.3226	0.3687	0.4093	0.4566	0.4889
11	0.3083	0.3524	0.3912	0.4367	0.4677
12	0.2958	0.3382	0.3754	0.4192	0.4491
13	0.2847	0.3255	0.3614	0.4036	0.4325
14	0.2748	0.3142	0.3489	0.3897	0.4176
15	0.2659	0.3040	0.3376	0.3771	0.4042
16	0.2578	0.2947	0.3273	0.3657	0.3920
17	0.2504	0.2863	0.3180	0.3553	0.3809
18	0.2436	0.2785	0.3094	0.3457	0.3706
19	0.2374	0.2714	0.3014	0.3369	0.3612
20	0.2316	0.2647	0.2941	0.3287	0.3524
21	0.2262	0.2586	0.2872	0.3210	0.3443
22	0.2212	0.2528	0.2809	0.3139	0.3367
23	0.2165	0.2475	0.2749	0.3073	0.3295
24	0.2121	0.2424	0.2693	0.3010	0.3229
25	0.2079	0.2377	0.2640	0.2952	0.3166
26	0.2040	0.2332	0.2591	0.2896	0.3106
27	0.2003	0.2290	0.2544	0.2844	0.3050
28	0.1968	0.2250	0.2499	0.2794	0.2997
29	0.1935	0.2212	0.2457	0.2747	0.2947
30	0.1903	0.2176	0.2417	0.2702	0.2899
31	0.1873	0.2141	0.2379	0.2660	0.2853
32	0.1845	0.2109	0.2342	0.2619	0.2809
33	0.1817	0.2077	0.2308	0.2580	0.2768
34	0.1791	0.2047	0.2274	0.2543	0.2728
35	0.1766	0.2019	0.2243	0.2507	0.2690
36	0.1742	0.1991	0.2212	0.2473	0.2653
37	0.1719	0.1965	0.2183	0.2440	0.2618
38	0.1697	0.1939	0.2154	0.2409	0.2584
39	0.1675	0.1915	0.2127	0.2379	0.2552
40	0.1655	0.1891	0.2101	0.2349	0.2521
45	0.1562	0.1786	0.1984	0.2218	0.2380
50	0.1484	0.1696	0.1884	0.2107	0.2260
70	0.1259	0.1438	0.1598	0.1786	0.1917
100	0.1056	0.1207	0.1840	0.1499	0.1608
	$1.07/\sqrt{n}$	$1.22/\sqrt{n}$	$1.36/\sqrt{n}$	$1.52/\sqrt{n}$	$1.63/\sqrt{n}$

附表10 W检验统计量的系数表($a_i(n)$)

n	1	2	3	4	5	6	7	8	9	10
$i=1$	—	—	0.7071	0.6794	0.6506	0.6246	0.6015	0.5812	0.5631	0.5470
2	—	—	—	0.1960	0.2769	0.3163	0.3369	0.3479	0.3535	0.3559
3	—	—	—	—	—	0.0993	0.1569	0.1930	0.2169	0.2332
4	—	—	—	—	—	—	—	0.0623	0.1041	0.1336
5	—	—	—	—	—	—	—	—	—	0.0436

n	11	12	13	14	15	16	17	18	19	20
$i=1$	0.5324	0.5192	0.5071	0.4960	0.4857	0.4763	0.4674	0.4592	0.4514	0.4442
2	0.3564	0.3555	0.3539	0.3517	0.3492	0.3465	0.3436	0.3407	0.3377	0.3348
3	0.2446	0.2526	0.2584	0.2624	0.2652	0.2671	0.2682	0.2689	0.2691	0.2690
4	0.1550	0.1711	0.1833	0.1927	0.2000	0.2058	0.2104	0.2140	0.2168	0.2190
5	0.0755	0.0995	0.1181	0.1327	0.1443	0.1537	0.1614	0.1677	0.1729	0.1773
6	—	0.0327	0.0579	0.0778	0.0938	0.1069	0.1176	0.1265	0.1340	0.1404
7	—	—	—	0.0257	0.0463	0.0630	0.0769	0.0885	0.0983	0.1066
8	—	—	—	—	—	0.0208	0.0380	0.0524	0.0646	0.0749
9	—	—	—	—	—	—	—	0.0174	0.0320	0.0445
10	—	—	—	—	—	—	—	—	—	0.0147

n	21	22	23	24	25	26	27	28	29	30
$i=1$	0.4373	0.4308	0.4247	0.4189	0.4133	0.4080	0.4030	0.3982	0.3936	0.3891
2	0.3319	0.3290	0.3261	0.3233	0.3206	0.3179	0.3153	0.3127	0.3102	0.3078
3	0.2686	0.2681	0.2673	0.2665	0.2656	0.2645	0.2635	0.2623	0.2612	0.2600
4	0.2208	0.2221	0.2231	0.2239	0.2243	0.2247	0.2248	0.2248	0.2247	0.2245
5	0.1809	0.1839	0.1864	0.1886	0.1903	0.1918	0.1930	0.1940	0.1948	0.1955
6	0.1458	0.1504	0.1543	0.1577	0.1607	0.1632	0.1654	0.1673	0.1690	0.1704
7	0.1138	0.1199	0.1253	0.1299	0.1339	0.1375	0.1406	0.1434	0.1458	0.1479
8	0.0838	0.0915	0.0982	0.1041	0.1092	0.1138	0.1178	0.1214	0.1246	0.1274
9	0.0552	0.0645	0.0726	0.0797	0.0859	0.0915	0.0964	0.1008	0.1047	0.1083
10	0.0274	0.0383	0.0479	0.0563	0.0637	0.0702	0.0760	0.0813	0.0859	0.0902
11	—	0.0127	0.0238	0.0335	0.0421	0.0497	0.0564	0.0625	0.0679	0.0728
12	—	—	—	0.0111	0.0209	0.0296	0.0373	0.0443	0.0505	0.0561
13	—	—	—	—	—	0.0098	0.0186	0.0264	0.0334	0.0398
14	—	—	—	—	—	—	—	0.0088	0.0167	0.0238
15	—	—	—	—	—	—	—	—	—	0.0079

n	31	32	33	34	35	36	37	38	39	40
$i=1$	0.3849	0.3808	0.3769	0.3731	0.3694	0.3659	0.3625	0.3592	0.3560	0.3530
2	0.3054	0.3031	0.3008	0.2986	0.2964	0.2943	0.2923	0.2902	0.2883	0.2864
3	0.2588	0.2576	0.2563	0.2551	0.2539	0.2526	0.2514	0.2502	0.2490	0.2478
4	0.2242	0.2239	0.2235	0.2230	0.2225	0.2219	0.2214	0.2208	0.2201	0.2195
5	0.1960	0.1964	0.1966	0.1968	0.1969	0.1970	0.1969	0.1969	0.1967	0.1966

附表10(续) W检验统计量的系数表($a_i(n)$)

n	31	32	33	34	35	36	37	38	39	40
$i=6$	0.1716	0.1727	0.1736	0.1743	0.1750	0.1756	0.1760	0.1764	0.1767	0.1770
7	0.1498	0.1515	0.1530	0.1543	0.1555	0.1566	0.1575	0.1583	0.1590	0.1596
8	0.1300	0.1322	0.1343	0.1361	0.1378	0.1393	0.1406	0.1418	0.1429	0.1439
9	0.1114	0.1143	0.1169	0.1193	0.1214	0.1233	0.1251	0.1267	0.1281	0.1295
10	0.0940	0.0974	0.1005	0.1034	0.1060	0.1083	0.1105	0.1125	0.1143	0.1160
11	0.0773	0.0813	0.0850	0.0883	0.0913	0.0941	0.0967	0.0990	0.1012	0.1032
12	0.0612	0.0658	0.0700	0.0738	0.0773	0.0805	0.0835	0.0862	0.0887	0.0910
13	0.0455	0.0507	0.0555	0.0598	0.0638	0.0674	0.0708	0.0739	0.0767	0.0794
14	0.0302	0.0360	0.0413	0.0462	0.0506	0.0547	0.0585	0.0619	0.0651	0.0681
15	0.0150	0.0215	0.0274	0.0328	0.0377	0.0423	0.0464	0.0503	0.0538	0.0571
16	—	0.0072	0.0137	0.0196	0.0250	0.0300	0.0346	0.0389	0.0428	0.0464
17	—	—	—	0.0065	0.0125	0.0180	0.0230	0.0276	0.0319	0.0359
18	—	—	—	—	—	0.0060	0.0115	0.0165	0.0212	0.0256
19	—	—	—	—	—	—	—	0.0055	0.0106	0.0153
20	—	—	—	—	—	—	—	—	—	0.0051

n	41	42	43	44	45	46	47	48	49	50
$i=1$	0.3500	0.3471	0.3442	0.3415	0.3389	0.3363	0.3338	0.3313	0.3289	0.3266
2	0.2845	0.2827	0.2809	0.2791	0.2774	0.2757	0.2741	0.2725	0.2709	0.2694
3	0.2466	0.2454	0.2443	0.2431	0.2420	0.2409	0.2397	0.2387	0.2376	0.2365
4	0.2188	0.2182	0.2175	0.2168	0.2161	0.2154	0.2147	0.2140	0.2133	0.2126
5	0.1963	0.1961	0.1958	0.1955	0.1952	0.1948	0.1945	0.1941	0.1937	0.1933
6	0.1771	0.1773	0.1774	0.1774	0.1774	0.1774	0.1773	0.1772	0.1771	0.1769
7	0.1602	0.1606	0.1611	0.1614	0.1617	0.1620	0.1622	0.1623	0.1625	0.1626
8	0.1448	0.1456	0.1463	0.1470	0.1476	0.1481	0.1485	0.1489	0.1493	0.1496
9	0.1307	0.1318	0.1328	0.1337	0.1346	0.1353	0.1360	0.1367	0.1373	0.1378
10	0.1175	0.1189	0.1202	0.1214	0.1225	0.1235	0.1244	0.1253	0.1261	0.1268
11	0.1050	0.1067	0.1083	0.1098	0.1111	0.1123	0.1135	0.1146	0.1156	0.1165
12	0.0932	0.0952	0.0970	0.0987	0.1003	0.1018	0.1032	0.1044	0.1056	0.1068
13	0.0818	0.0841	0.0862	0.0882	0.0900	0.0917	0.0933	0.0948	0.0962	0.0975
14	0.0708	0.0734	0.0758	0.0780	0.0801	0.0820	0.0838	0.0855	0.0871	0.0886
15	0.0602	0.0630	0.0657	0.0682	0.0705	0.0726	0.0746	0.0765	0.0783	0.0800
16	0.0498	0.0529	0.0559	0.0586	0.0611	0.0635	0.0658	0.0679	0.0698	0.0717
17	0.0396	0.0430	0.0462	0.0492	0.0520	0.0546	0.0571	0.0594	0.0616	0.0636
18	0.0296	0.0333	0.0368	0.0401	0.0431	0.0459	0.0486	0.0511	0.0535	0.0557
19	0.0197	0.0237	0.0275	0.0310	0.0343	0.0374	0.0403	0.0430	0.0456	0.0480
20	0.0098	0.0142	0.0183	0.0221	0.0256	0.0290	0.0321	0.0350	0.0378	0.0404
21	—	0.0047	0.0091	0.0132	0.0171	0.0206	0.0240	0.0271	0.0301	0.0329
22	—	—	—	0.0044	0.0085	0.0124	0.0160	0.0193	0.0225	0.0255
23	—	—	—	—	—	0.0041	0.0080	0.0116	0.0150	0.0182
24	—	—	—	—	—	—	—	0.0039	0.0075	0.0109
25	—	—	—	—	—	—	—	—	—	0.0036

附表11 W检验的下侧分位数表$P\{W_n \le W_\alpha(n)\} = \alpha$

α	0.01	0.05	0.10	α	0.01	0.05	0.10
				$n=26$	0.891	0.920	0.933
				27	0.894	0.923	0.935
$n=3$	0.753	0.767	0.789	28	0.896	0.924	0.936
4	0.687	0.748	0.792	29	0.898	0.926	0.937
5	0.686	0.762	0.806	30	0.900	0.927	0.939
6	0.731	0.788	0.826	31	0.902	0.929	0.940
7	0.730	0.803	0.838	32	0.904	0.930	0.941
8	0.749	0.818	0.851	33	0.906	0.931	0.942
9	0.764	0.829	0.859	34	0.908	0.933	0.943
10	0.781	0.842	0.869	35	0.910	0.934	0.944
11	0.792	0.850	0.876	36	0.912	0.935	0.945
12	0.805	0.859	0.886	37	0.914	0.936	0.946
13	0.814	0.866	0.889	38	0.916	0.938	0.947
14	0.825	0.874	0.895	39	0.917	0.939	0.948
15	0.835	0.881	0.901	40	0.919	0.940	0.949
16	0.844	0.887	0.906	41	0.920	0.941	0.950
17	0.851	0.892	0.910	42	0.922	0.942	0.951
18	0.858	0.897	0.914	43	0.923	0.943	0.951
19	0.863	0.901	0.917	44	0.924	0.944	0.952
20	0.868	0.905	0.920	45	0.926	0.945	0.953
21	0.873	0.908	0.923	46	0.927	0.945	0.953
22	0.878	0.911	0.926	47	0.928	0.946	0.954
23	0.881	0.914	0.928	48	0.929	0.947	0.954
24	0.884	0.916	0.930	49	0.929	0.947	0.955
25	0.888	0.918	0.931	50	0.930	0.947	0.955

附表12 D检验的分位数表$P\{D \geq D_\alpha\} = \alpha$

α	0.995	0.975	0.95	0.05	0.025	0.005
n =50	−3.91	−2.74	−2.21	0.937	1.06	1.24
60	−3.81	−2.68	−2.17	0.997	1.13	1.34
70	−3.73	−2.64	−2.14	1.05	1.19	1.42
80	−3.67	−2.60	−2.11	1.08	1.24	1.48
90	−3.61	−2.57	−2.09	1.12	1.28	1.54
100	−3.57	−2.54	−2.07	1.14	1.31	1.59
150	−3.41	−2.45	−2.00	1.23	1.42	1.75
200	−3.30	−2.39	−1.96	1.29	1.50	1.85
250	−3.23	−2.35	−1.93	1.33	1.55	1.93
300	−3.17	−2.32	−1.91	1.36	1.53	1.98
350	−3.13	−2.29	−1.89	1.38	1.61	2.03
400	−3.90	−2.27	−1.87	1.40	1.63	2.06
450	−3.06	−2.25	−1.86	1.41	1.65	2.09
500	−3.04	−2.24	−1.85	1.42	1.67	2.11
550	−3.02	−2.23	−1.84	1.43	1.68	2.14
600	−3.00	−2.22	−1.83	1.44	1.69	2.15
650	−2.98	−2.21	−1.83	1.45	1.70	2.17
700	−2.97	−2.20	−1.82	1.46	1.71	2.18
750	−2.96	−2.19	−1.81	1.47	1.72	2.20
800	−2.94	−2.18	−1.81	1.47	1.73	2.21
850	−2.93	−2.18	−1.80	1.48	1.74	2.22
900	−2.92	−2.17	−1.80	1.48	1.74	2.23
950	−2.91	−2.16	−1.80	1.49	1.75	2.24
1000	−2.91	−2.16	−1.79	1.49	1.75	2.25

附表13 二项分布表$Q(k;n,p)=\sum_{i=k}^{n}C_n^i p^i(1-p)^{n-i}$.

n	k	0.01	0.02	0.04	0.06	0.08	0.1	0.2	0.3	0.4	0.5
	5						0.00001	0.00032	0.00243	0.01024	0.03125
	4			0.00001	0.00006	0.00019	0.00046	0.00672	0.03078	0.08704	0.18750
5	3	0.00001	0.00008	0.00060	0.00197	0.00453	0.00856	0.05792	0.16308	0.31744	0.50000
	2	0.00098	0.00384	0.01476	0.03187	0.05436	0.08146	0.26272	0.47178	0.66304	0.81250
	1	0.04901	0.09608	0.18463	0.26610	0.34092	0.40951	0.67232	0.83193	0.92224	0.96875
	10								0.00001	0.00010	0.00098
	9								0.00014	0.00168	0.01074
	8							0.00008	0.00159	0.01229	0.05469
	7						0.00001	0.00086	0.01059	0.05476	0.17188
10	6				0.00001	0.00004	0.00015	0.00637	0.04735	0.16624	0.37695
	5			0.00002	0.00015	0.00059	0.00163	0.03279	0.15027	0.36690	0.62305
	4		0.00003	0.00044	0.00203	0.00580	0.01280	0.12087	0.35039	0.61772	0.82812
	3	0.00011	0.00086	0.00621	0.01884	0.04008	0.07019	0.32220	0.61722	0.83271	0.94531
	2	0.00427	0.01618	0.05815	0.11759	0.18788	0.26390	0.62419	0.85069	0.95364	0.98926
	1	0.09562	0.18293	0.33517	0.46138	0.56561	0.65132	0.89263	0.97175	0.99395	0.99902
	15										0.00003
	14									0.00003	0.00049
	13								0.00001	0.00028	0.00369
	12								0.00009	0.00193	0.01758
	11							0.00001	0.00067	0.00935	0.05923
	10							0.00011	0.00365	0.03383	0.15088
	9							0.00078	0.01524	0.09505	0.30362
15	8					0.00001	0.00003	0.00424	0.05001	0.21310	0.50000
	7				0.00001	0.00008	0.00031	0.01806	0.13114	0.39019	0.69638
	6			0.00002	0.00015	0.00070	0.00225	0.06105	0.27838	0.59678	0.84912
	5		0.00001	0.00022	0.00140	0.00497	0.01272	0.16423	0.48451	0.78272	0.94077
	4	0.00001	0.00018	0.00245	0.01036	0.02731	0.05556	0.35184	0.70313	0.90950	0.98242
	3	0.00042	0.00304	0.02029	0.05713	0.11297	0.18406	0.60198	0.87317	0.97289	0.99631
	2	0.00963	0.03534	0.11911	0.22624	0.34027	0.45096	0.83287	0.96473	0.99483	0.99951
	1	0.13994	0.26143	0.45791	0.60471	0.71370	0.79411	0.96482	0.99525	0.99953	0.99997
	20										0.00000
	19										0.00002
	18									0.00001	0.00020
	17									0.00005	0.00129
	16								0.00001	0.00032	0.00591
	15								0.00004	0.00161	0.02069
	14								0.00026	0.00647	0.05766
	13							0.00002	0.00128	0.02103	0.13159
	12							0.00010	0.00514	0.05653	0.25172
20	11							0.00056	0.01714	0.12752	0.41190
	10						0.00001	0.00259	0.04796	0.24466	0.58810
	9					0.00001	0.00006	0.00998	0.11333	0.40440	0.74828
	8				0.00001	0.00009	0.00042	0.03214	0.22773	0.58411	0.86841
	7			0.00001	0.00011	0.00064	0.00239	0.08669	0.39199	0.74999	0.94234
	6			0.00010	0.00087	0.00380	0.01125	0.19579	0.58363	0.87440	0.97931
	5		0.00004	0.00096	0.00563	0.01834	0.04317	0.37035	0.76249	0.94905	0.99409
	4	0.00004	0.00060	0.00741	0.02897	0.07062	0.13295	0.58855	0.89291	0.98404	0.99871
	3	0.00100	0.00707	0.04386	0.11497	0.21205	0.32307	0.79392	0.96452	0.99639	0.99980
	2	0.01686	0.05990	0.18966	0.33955	0.48314	0.60825	0.93082	0.99236	0.99948	0.99998
	1	0.18209	0.33239	0.55800	0.70989	0.81131	0.87842	0.98847	0.99920	0.99996	1.00000

附表13(续) 二项分布表$Q(k;n,p)=\sum_{i=k}^{n}C_n^i p^i(1-p)^{n-i}$.

n	k	0.01	0.02	0.04	0.06	0.08	0.1	0.2	0.3	0.4	0.5
25	24										0.00000
	23										0.00001
	22										0.00008
	21									0.00001	0.00046
	20									0.00005	0.00204
	19									0.00028	0.00732
	18								0.00002	0.00121	0.02164
	17								0.00010	0.00433	0.05388
	16								0.00045	0.01317	0.11476
	15							0.00001	0.00178	0.03439	0.21218
	14							0.00008	0.00599	0.07780	0.34502
	13							0.00037	0.01747	0.15377	0.50000
	12							0.00154	0.04425	0.26772	0.65498
	11						0.00001	0.00555	0.09780	0.41422	0.78782
	10					0.00001	0.00008	0.01733	0.18944	0.57538	0.88524
	9				0.00001	0.00008	0.00046	0.04677	0.32307	0.72647	0.94612
	8				0.00007	0.00052	0.00226	0.10912	0.48815	0.84645	0.97836
	7			0.00004	0.00051	0.00277	0.00948	0.21996	0.65935	0.92643	0.99268
	6		0.00001	0.00038	0.00306	0.01229	0.03340	0.38331	0.80651	0.97064	0.99796
	5		0.00012	0.00278	0.01505	0.04514	0.09799	0.57933	0.90953	0.99053	0.99954
	4	0.00011	0.00145	0.01652	0.05976	0.13509	0.23641	0.76601	0.96676	0.99763	0.99992
	3	0.00195	0.01324	0.07648	0.18711	0.32317	0.46291	0.90177	0.99104	0.99957	0.99999
	2	0.02576	0.08865	0.26419	0.44734	0.60528	0.72879	0.97261	0.99843	0.99995	1.00000
	1	0.22218	0.39653	0.63960	0.78709	0.87564	0.92821	0.99622	0.99987	1.00000	1.00000
30	27										0.00000
	26										0.00003
	25										0.00016
	24									0.00001	0.00072
	23									0.00005	0.00261
	22									0.00022	0.00806
	21								0.00001	0.00086	0.02139
	20								0.00004	0.00285	0.04937
	19								0.00016	0.00830	0.10024
	18								0.00063	0.02124	0.18080
	17							0.00001	0.00212	0.04811	0.29233
	16							0.00005	0.00637	0.09706	0.42777
	15							0.00023	0.01694	0.17537	0.57223
	14							0.00090	0.04005	0.28550	0.70767
	13							0.00311	0.08447	0.42153	0.81920
	12						0.00002	0.00949	0.15932	0.56891	0.89976
	11					0.00001	0.00009	0.02562	0.26963	0.70853	0.95063
	10				0.00001	0.00007	0.00045	0.06109	0.41119	0.82371	0.97861
	9				0.00005	0.00041	0.00202	0.12865	0.56848	0.90599	0.99194
	8			0.00002	0.00030	0.00197	0.00778	0.23921	0.71862	0.95648	0.99739
	7			0.00015	0.00167	0.00825	0.02583	0.39303	0.84048	0.98282	0.99928
	6		0.00003	0.00106	0.00795	0.02929	0.07319	0.57249	0.92341	0.99434	0.99984
	5	0.00001	0.00030	0.00632	0.03154	0.08736	0.17549	0.74477	0.96985	0.99849	0.99997
	4	0.00022	0.00289	0.03059	0.10262	0.21579	0.35256	0.87729	0.99068	0.99969	1.00000
	3	0.00332	0.02172	0.11690	0.26760	0.43460	0.58865	0.95582	0.99789	0.99995	1.00000
	2	0.03615	0.12055	0.33882	0.54453	0.70421	0.81631	0.98948	0.99969	1.00000	1.00000
	1	0.26030	0.45452	0.70614	0.84374	0.91803	0.95761	0.99876	0.99998	1.00000	1.00000

参考文献

[1] Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd Edition, New York: Springer-Verlag. (有中译本)

[2] 陈家鼎(1995). 序贯分析. 北京大学出版社.

[3] 陈家鼎、孙山泽、李东风、刘力平(2006). 数理统计学讲义，高等教育出版社.

[4] 陈希孺(1997). 数理统计引论, 中国科学出版社.

[5] 陈希孺(2000). 数理统计发展简史，湖南教育出版社.

[6] 陈希孺、倪国策(1988). 数理统计教程. 上海科技出版社.

[7] David, H. A. (1981). *Order Statistics*, 2nd Edition, Academic Press.

[8] Efron, B. (1982). *The Jachhnife, the bootstrap, and other Resampling Plans*, Volume 38 of CBMS-NSF Regional Conference Series in Applied Mathematics, Philadelphia:SIAM.

[9] Efron, B., and Thbshirani, R. J. (1993). *An Introduction to the Bootstrap*, London: Chapman & Hall.

[10] 高惠璇(1995). 统计计算，北京大学出版社

[11] 何书元(2006). 概率论与数理统计，高等教育出版社.

[12] Huff, D. (1982). *How to Lie with Statistics*, W. W. Norton & Company. (沈恩杰、马世宽、马安、吴世农译(1989)，中国统计出版社.)

[13] 梁小筠(1997). 正态性检验，中国统计出版社.

[14] Lehmann, E. L. (1986). *Testing Statistical Hypotheses*, 2nd Edition, New York: John Wiley & Sons.

[15] Lehmann, E. L. and CaseHa, G. (1998). *Theory of Point Estimation*, 2nd Edition, Springer-Verlag (郑忠国、蒋建成、童行伟译(2004)，中国统计出版社) .

[16] 茆诗松、王静龙(1990). 数理统计，华东师范大学出版社.

[17] 茆诗松、王静龙、濮晓龙(2006). 高等数理统计，第二版，高等教育出版社.

[18] 茆诗松、王玲玲(1984). 可靠性统计，华东师范大学出版社.

[19] Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd Edition, New York: John Wiley & Sons. (有中译本)

[20] Ross, S. M. (2002). *Simulation*, 3rd Edition, Academic Press. (王兆军、陈广雷、邹长亮译(2006)，人民邮电出版社)

[21] Rubinstein, R. Y. (1981). *Simulation and Monte Carlo Methods*, New York: John Wiley & Sons.

[22] Santner, T. J. and Duffy, D. E. (1990). *The Statistical Analysis of Discrete Data*, New York: Springer-Verlag.

[23] Serfling, R. J. (1980). *Approximation Theorems of Mat hematical Statistics*, New York: John Wiley & Sons.

[24] Shao, J. and Tu, D. (1995). *The Jackknife and the Bootstrap*. New York: Springer-Verlag.

[25] 王静龙、梁小筠(2006). 离散数据分析方法，华东师大出版社.

[26] 韦博成、鲁国斌、史建清(1991). 统计诊断引论, 东南大学出版社.

[27] 杨振海(1994). 拟合优度检验, 安徽教育出版社.

[28] 中国科学院数学研究所概率统计室编(1979). 常用数理统计表, 科学出版社.

索　引