

## A Course in Parametric Statistics

韦博成编著

高等教育出版社

### 参数统计教程 A Course in Parametric Statistics

韦博成编著

高等教育出版社

#### 内容提要

本书为概率统计专业的研究生教材, 全书共分八章, 比较全面系统地介绍了: 常见 的统计分布, 充分统计量和信息函数, 点估计的基本理论和方法, 假设检验的理论、方 法及其应用, 区间估计及其应用, Bayes统计推断的基本概念和方法等。本书也可作为经 济金融、生物医学、管理科学、工程技术等专业研究生的教学参考书, 还可供相关专业 的大学生、研究生、教师、科技人员和统计工作者参考。

图书在版编目(CIP)数据

参数统计教程/ 韦博成编著. —北京: 高等教育出版 社 , 2006.11

ISBN 7-04-020054-6

I .参... n. 韦... in. 数理统计—高等学校—教 材 IV.0212

中国版本图书馆CIP数据核字(2006)第 120772号

策划编辑杨波 责任编辑高尚华 封面设计李卫青 责任绘图尹文军 版式设计史新薇 责任校对王效珍 责任印制陈伟光

出版发行 社 址 邮政编码 总 机

经 销 印 刷

开 本 印 张 字 数

高等教育出版社 北京市西城区德外大街4号 100011

010 - 58581000

蓝色畅想图书发行有限公司 北京宝旺印务有限公司

购书热线 010 - 58581118 免费咨询 800 — 810 — 0598

网 址 <http://www.hep.edu.cn>

<http://www.hep.com.cn> 网上订购 <http://www.landrac.com>

<http://www.landrac.com.cn> 畅想教育 <http://www.widedu.com>

版 次 2006年11月第1版

印 次 2006年11月第1次印刷 定 价 35.90元

787 X 960 26.25

440 000

1/16

本朽如有缺页、倒页、脱页等质量问题, 请到所购图书销售部门联系调换。 版权所有侵权必究

物料号 20054-00

别

本书为概率统计专业的研究生教材, 比较全面系统地介绍了数理统 计的基本原理、基本方法及其应用。本书也可作为经济金融、生物医 学、管理科学、工程技术等专业研究生的教学参 考书, 还可供相关专业

的大学生、研究生、教师、科技人员和统计工作者参考。阅读本书需要 有数学分析、线性代数, 特别是概率论方面的基础知识, 但不需要测度 论方面的知识。掌握本书的内容, 即可进一步学习统计学其他各分支的 理论与方法, 也可比较顺利地理解其他学科中用到的统计学基 本概念。

本书定位为“中等水平、便于阅读、内容充实、有一定特色”的教材。希望在不很高的起点上，对数理统计的基本理论和方法有比较清楚、深入的阐述，对数理统计的实际背景和应用有适当的介绍。全书共分八章，包罗了通常高等数理统计的主要内容。第一章介绍常见的统计分布及有关问题，也介绍了非中心 $t$ 分布、带有多余参数的指数族分布等内容；由于任何统计问题都涉及统计分布，因此本书单列一章，用较大篇幅介绍这方面的内容，这使仅有本科概率论基本知识的读者，对数理统计的研究对象有更多的了解，以便于后面集中精力去领会数理统计的基本概念和方法。第二章介绍充分统计量和分布族的信息函数，也介绍了Basu定理、Kullback信息等内容；本章内容也相对较多，因为几乎任何统计问题都与充分统计量、样本信息等基本概念有关，本书也单列一章，用较大篇幅比较系统地介绍这些基本概念，这对非数学专业的读者进一步学习数理统计可能更加有用。本书第三至第五章介绍点估计的基本理论和方法，第三章介绍常用的点估计方法，也介绍了不变原理、子集参数的似然等内容，同时也配备了比较丰富的例题与习题；第四章单列一章，以比较初等的方法系统地介绍了“同变估计”及其求解方法；第五章介绍点估计的基本性质，也介绍了广义C-R型不等式等内容。第六章篇幅最大(与点估计的3章相当)，比较全面系统地介绍了假设检验的基本概念、基本理论和基本方法，并附有较多的应用方面的例题与习题；本章还比较详细地介绍了有广泛应用价值的score检验统计量。第七章介绍常用的区间估计方法及其应用，也介绍了单调似然比分布族参数的区间估计方法。第八章以较大的篇幅介绍Bayes统计推断的基本概念和方法，除了常见的参数估计与假设检验的Bayes方法

#### n 前言

以外，本章还从同变原理出发，比较详细地介绍了位置、尺度参数分布族无信息先验分布的选取准则；同时也比较详细地阐述了HPD可信域的基本性质和求解方法。

本书初稿为我校研究生课程讲义，经作者多次讲授、不断修改、逐步形成；在成书过程中又作了全面的充实、加工与修订。厦门大学王海斌博士校阅了全书，并且提出了许多修改意见；香港中文大学博士生周影辉仔细校阅了全书，并且帮助演算了本书习题，提出了许多修改意见；博士生陆建也演算了部分习题；南京农业大学解锋昌副教授为本书

绘制了全部图形；研究生章珏帮助打印了全书初稿，特此表示衷心的感谢！另外，在本书写作过程中，参考了国内外许多图书资料，受益匪浅，一并对这些作者表示衷心的感谢！

本书在写作过程中，自始至终得到高等教育出版社的关心与帮助，特别要感谢高等理工出版中心、数学分社的杨波同志，他们对本书的写作、审定与出版都给予大力的支持与帮助，特此表示衷心的感谢！同时也对审稿先生的厚爱与支持表示衷心的感谢！

由于作者水平有限，难免有不妥与谬误之处，恳请同行专家和广大读者提出批评和建议。

韦博成

2006年6月于东南大学

#### 郑重声明

高等教育出版社依法对本书享有专有出版权。任何未经许可的复制、销售行为均违反《中华人民共和国著作权法》，其行为人将承担相应的民事责任和行政责任，构成犯罪的，将被依法追究刑事责任。为了维护市场秩序，保护读者的合法权益，避免读者误用盗版书造成不良后果，我社将配合行政执法部门和司法机关对违法犯罪的单位和个人给予严厉打击。社会各界人士如发现上述侵权行为，希望及时举报，本社将奖励举报有功人员。

反盗版举报电话：(010) 58581897/58581896/58581879 传真：(010) 82086060

E-mail: dd@hep.com.cn 通信地址：北京市西城区德外大街4号

高等教育出版社打击盗版办公室 邮编：100011

购书请拨打电话：(010)58581118

## 目录

### 第一章统计分布基

#### 础

#### 1 11 随机变量及其分布函 数

#### 1 111 分布函数与分布密 度

1

#### 1 12 1 13 1 14 反函数 及分位

#### 数

#### 3 特征函数和数字特

#### 征

#### 5 经验分布函

#### 数

#### 6

#### 12 常见的离散型分

#### 布

#### 7 13 常见的连续型分

#### 布

#### 13 14 一元非中心 $\chi^2$ 分布及其有关分

#### 布 24

#### 14 1 非中心 $\chi^2$ 分布和非中心 $t$ 分

#### 布 24

#### 142 非中心 $F$ 分布和非中心 $F$ 分

#### 布 27

#### 15 指数族分布 28 15 1 基本定

#### 义

#### 28 152指数族的自然形

#### 式

#### 30 153 带有多余参数的指数

#### 族

#### 33 16 次序统计量的分

#### 布

#### 36 16 1 基本分

#### 布

#### 36 162 均匀分布的次序统计

#### 量

#### 38 163 指数分布的次序统计

#### 量 39 习题

#### —

#### 41

## 第二章充分统计量与样本信

息47 2

1 充分统计量 47 2 1 1 充分统计量的定

义

47

2 12 因子分解定

理

51

2 13 极小充分统计

量

56 22统计量的完备

性

59 2 2 1 分布族的完备

性

60 222统计量的完备

性 62

223指数族统计量的完备

性'

64 2 2 4 Basu 定

理 66

n

目

录

23 分布族的信息函数

2 3 1 Fisher 信

息

68 232 KuIIback - Leibler 信息(K-L 距离)和 Jensen 不等式 75

习题

二

79 第三章点估计基本方法 3 1 统计判决函数

3 1 1 统计判决三要素

3 12 统计判决函数的优良性准则 313 Rao-Blackwell定理

32 无偏估计及其UMRUE和 UMVUE 32 1 基本定义

3 2 2 Lehmann - Scheff6 定理

323 例题

33 极大似然估计

331定义与例题 101 332指数族分布的极大似然估计 107 333 不变原理 109

334子集参数的似然 111 335极大似然估计的迭代算法 113

34 矩方程估计 116

3题

三

第四章最优同

变估计 127 4 1 变换群下的同变估计 127 411 同变性概念 127 412 同变统计判决  
 函数 129 42 平移变换群下位置参数的最优同变估  
 计 133 4 2 1 位置参数分布族的平移变换群 133  
 4 2 2 ' 位置参数的最优同变估计 135  
 423 Pitman 积分公式 138 4-3 相似变换群下尺度参数的最优同变估计 140 431 尺  
 度参数分布族的相似变换群 140 432尺度参数的最优同变估计 142  
 4 3 3 Pitman 积分公式 146 4-4线性变换群下位置尺度参数的最优同变估计 151  
 441位置尺度参数分布族与线性变换  
 群 15]

91 94 100 85 86 86 87 89 90 90

目

录m

442位置尺度参数的最优同变估

计154

4 4 3 Pitman 积分公

式

158 习题

四

第五章点估计的性

质

162 51 C-R不等

式

163

5 1 1 单参数C-R不等

式

164 512等式成立的条

件

168 5 13 Bh 不等

式

170 5 14 多参数C-R不等

式

174

52 广义C-R型不等

式

175 53 估计量的渐近性

质

:179

53 1 随机变量序列的收敛

性 180 532 估计量的相合性和渐近正态

性183

5 33

534 3 题五

矩估计的相合性和渐近正态

性190 极大似然估计的相合性和渐近正态

性191

## 198 第六章参数假设检

验

203 6 1 假设检验的基本概

念

203 6- 1 1 否定域与检验函

数

205 612 两类错误及功效函

数

206

613 Neyman - Pearson准则与一致最优势检

验 208 6 2 Neyman - Pearson 基本引

理

211 6 2 1 Neyman - Pearson 基本引

理 211 622 Neyman - Pearson基本引理应用示

例 217 63单调似然比分布族的单边检

验 222 631单调似然比分布族单边检验的

UMPT 222

632正态分布单参数的单边检

验 230

64单参数指数族分布的双边检

验 232 641 双边检验问题及无偏检

验 232 6- 42 指数族分布的双边检

验

233

643 正态分布单参数的双边检

验 247

65 多参数指数族的检

验

249 651带有多余参数时单参数检验的

UMPUT 251

652 一样本正态总体的检

验

255 653 两样本正态总体的检

验 263

IV

目

录

6 54 两个二项分布总体的比较----- 等价性检

验 267 66 似然比检

验

270 6 6 1 似然比检

验

271 6 62 子集参数的似然比检验及score检

验 276

6 7 拟合优度检

验

281 67 1 拟合优度检验与多项分布检

验282

672 多项分布检验的Pearson定

理 284

673 含参数多项分布的检验及Fisher定理 286

674 应用:列联表及其等价性和独立性检

验 291 习题

六

295 第七章 E 间估

计

305

7 I 置信区间及其枢轴量

法

305 7 1 1 置信区间和置信

限

305 712构造置信域的枢轴量

法

307 713 基于渐近分布的枢轴量

法 313 714单调似然比分布族参数的区间估

计 318

72参数置信域与假设检验的接受

域 323 721 对偶关

系

323 722 一致最准确置信

域

326

73 容忍区间与容忍

限

331 731问题与定

义

331 732容忍上、下限的计

算

333 733应用次序统计量计算容忍

限 336 习题

七 :

337 第八章Bayes统计基

础

343

8- I Bayes统计基本概

念

343 8 1 1 Bayes 统计原

理

344 8 12 先验分布的选取方

法

348

82 Bayes 估

计

356

8 2 1 Bayes 风

险

356 8-22后验期望估

计

359 823 后验极大似然估

计

366 824 Bayes估计的某些性

质

370

8 3 假设检验与区间估计的Bayes方

法375

831 Bayes 假设检

验 375

目录 V

832 Bayes区间估计和HPD可信区

间 383 习题

八

390 参考文

献

399 S弓

I

401



## 第一章统计分布基石出

对于一个随机变量, 其分布函数完全描述了它的概率结构. 但在实际问题中, 分布函数常常是未知的. 数理统计的中心任务, 就是通过样本观察值, 对总体的分布函数以及由此产生的问题进行合理的推断. 统计问题的基本提法是: 给定样本 $\dots$ , 总; 大多情形为独立同分布(即 i. i. d.) 样本, 这时服从未知分布) 或  $\mathcal{P}(\lambda)$ ,  $\lambda$  未知; 如服从正态分布,  $\theta = (\mu, \sigma^2)$ , 或 Poisson 分布  $\mathcal{P}(\lambda)$  等等. 从样本  $X = x_1, \dots, x_n$  出发, 对未知分布或  $\mathcal{F}(\theta)$  进行统计推断(诸如参数估计、假设检验、回归分析等). 因此, 随机变量及其分布函数, 既是统计推断的目的, 也是统计推断的基础. 一切统计问题都离不开随机变量及其分布, 熟练掌握这方面的知识, 对学好数理统计并应用于实践是非常必要和有益的, 读者必须十分熟悉. 关于概率论的基础知识, 读者可参见 王梓坤(1979), 李贤平(1997) 以及严士健, 刘秀芳(1994).

本章第1.1节回顾了分布函数的基本性质; 第1.2、第1.3节分别介绍常见的离散型分布和连续型分布; 第1.4节介绍一元非中心分布及其有关分布; 第1.5节介绍指数族分布以及带有多余参数的指数族分布; 第1.6节介绍次序统计量的分布, 并特别介绍了均匀分布和指数分布次序统计量的性质. 关于常见的统计分布, 可参见方开泰, 许建伦(1987), Zacks(1981), 茆诗松等(1998).

### 1.1 随机变量及其分布函数

1.1.1 分布函数与分布密度 给定  $n$  维随机变量  $X$ , 其相应的概率空间记为  $(\Omega, \mathcal{B}, P)$ . 今设  $X \in \mathbb{R}^1$ , 则一元分布函数定义为

$$F(x) = P(X \leq x) = P\{X \in (-\infty, x]\}.$$

这里  $\mathbb{R}^1$  表示  $\mathbb{R}$  上的 Borel 域;  $P$  表示

中  $f$  表示  $\mathbb{R}_n$  上  $X$  的样本空间  $\Omega$  上对应于  $X$  的概率测度, 即对  $A \in \mathcal{B}$ ,  $P(A) = P\{X \in A\}$  表示  $X$  属于  $A$  的概率. 概率测度  $P$  也常表示为  $P_x$  或  $\mathcal{P}_x$ .

2

## 第一章统计分布基础

熟知, 分布函数  $F$  为定义在  $(-\infty, +\infty)$  上的非降, 右连续函数, 且有  $\lim_{x \rightarrow -\infty} F(x) = 0$ ,  $\lim_{x \rightarrow +\infty} F(x) = 1$ , 并且满足  $F(x) = P(X \leq x)$ ,  $F(x-0) = P(X < x)$ ,  $P(X = x) = F(x) - F(x-0)$ ; 若  $x$  为  $F(x)$  的连续点, 则  $P(X = x) = 0$ . 另外, 单调函数  $F$  可扩张为  $\mathbb{R}^1$  上的测度  $F(-)$ , 且有  $F(A) = P(A)$  (见王梓坤, 1979).

分布函数通常可分为绝对连续型、离散型和奇异型. 若存在密度函数  $f(x)$ , 使  $F(x) = \int_{-\infty}^x f(y) dy$ ,  $F(x) = \int_{-\infty}^x f(y) dy$ , 则称  $F$  为绝对连续

J-

型, 且有  $P(X \in A) = \int_A f(y) dy$ . 若  $F$  连续但不存在密度函数, 则称

以  $F$  为奇异型; 奇异型分布实际中很少见到. 若随机变量  $X$  取有限或

可数个值, 则其分布函数为离散型; 这时  $P(X = x_k) = p_k$ ,  $k =$

$0, \pm 1, \pm 2, \dots$ ,  $\sum_{k=-\infty}^{+\infty} p_k = 1$ . 例如 Poisson 分布:  $P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$ ,  $k =$

$0, 1, 2, \dots$ . 退化分布是离散型分布一个重要特例, 这时  $P(X = 0) = 1$ ,  $P(X \neq 0) = 0$ ; 其分布函数为

$$F(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

其中  $I_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$  为示性函数. 对于一个集合  $A \subset \mathbb{R}^1$ ,  $I_A(x)$

定义为  $I_A(x) = 1$ ,  $I_A(x) = 0$ ,  $x \notin A$ . 因此, 一般情形下离散型分布的分布函数可表示为  $F(x) = \sum_{x_k \leq x} p_k = \sum_{k=-\infty}^{+\infty} I_{(-\infty, x]}(x_k) p_k$  (1.1.1)

离散型分布的分布函数亦可表示为积分形式, 今简要介绍 Lebesgue - Stieltjes 积分如下(可参见 Cramer, 1946 第七章和第九章; Rao, 1973 第二章).

以下介绍  $\mathbb{R}^1$  上的积分,  $\mathbb{R}_n$  上的积分完全类似. 给定  $\mathbb{R}^1$  上的有界函数  $g$  和有限测度  $\mu$

(•), 将一个有穷或无穷区间  $[a, b]$  分划为有限个互不相交的可测集  $A_i (i=1, \dots, n)$  之和. 对任一“分划”, 记  $S_n = \sum_{i=1}^n A_i(x)$   $t=1$  和  $n \rightarrow \infty$   $I = I$

$A_i(x)$ , 其中  $\int (A_i(x))$  为  $A_i$  的测度,  $\mu$  和  $\mu_i$  分别为  $\mu$  和  $\mu_i$   $<X>$

在  $\Delta$  上的下确界与上确界. 若  $\max |M(A_i(x))| \rightarrow 0$  时,  $\mu$  和  $\mu_i$  的极限存在  $1 \leq n$  在且相等, 则其极限称为函数  $g(x)$  在区间  $[a, b]$  上关于测度  $M(\cdot)$  的 Lebesgue - Stieltjes 积分, 并记为  $\int_a^b g(x) d\mu(x)$  或  $\int_a^b g(x) dF(x)$ . 以下介绍与分布函数有关的两种情况.

### 1.1 随机变量及其分布函数

3

(1) 若  $\mu(\cdot)$  为 Lebesgue 测度,  $A_i$  为区间, 则  $(A_i(x)) = A_i$  为区间长度; 若  $g(x) = [g(x)] dx$  因此若  $F(x)$  为绝对连续型分布函数, 存在密度函数  $f(x)$ , 则有

$$F(x) = \int_{-\infty}^x f(y) dy = \int_{-\infty}^x f(y) d\mu(y). \quad (1.1.1)$$

(1.1.2)

(2) 若为计数测度. 设点列  $\{x_k\}$  有可数个值,  $x_k, \dots$ , 计数测度表示  $\mu$  所包含点列  $\{x_k\}$  中点的个数. 因此对区间  $A_i$  有  $\mu(A_i) = 1$ , 若  $A_i$  中包含一个  $x_k$  的点;  $\mu(A_i) = 0$ , 若  $A_i$  不

包含  $x_k$  中的点. 对于定义在点列  $\{x_k\}$  上的函数  $g(x)$ , 则由 Lebesgue - Stieltjes 积分的定义可知  $\int g(x) d\mu(x) = \sum g(x_k) \mu(A_i)$ . 特别, 对于离散型分布的分布函数  $F(x)$ , 由 (1.1.1) 式可知, 若在点列  $\{x_k\}$  上定义计数测度以及  $f(x_k)$

则有

$$F(x) = \sum_{x_k \leq x} f(x_k) \quad (1.1.3)$$

$$f(x) =$$

$$x_k \leq x$$

比较 (1.1.2) 式和 (1.1.3) 式, 通常把  $P(X=x_k) = p_k = f(x_k)$  称为离散型分布的密度函数. 例如 Poisson 分布的密度函数为  $f(x_k) = e^{-\lambda} \lambda^{x_k} / x_k!$ .

因此对于绝对连续型分布和离散型分布, 一个函数的积分, 即数学期望可表示为  $E(X) = \int x f(x) d\mu(x) = \int x f(x) dF(x)$ , 其中  $dF(x) = f(x) dx$ .

特别, 概率  $P(A)$  可表示为  $P(A) = \int_A f(x) d\mu(x) = \int_A dF(x)$ .

另外, Lebesgue - Stieltjes 积分可推广到  $n$  维空间, 可参见 Cramer (1946, 第九章). 特别, 若从 (•) 就取为概率测度  $P(\cdot)$ , 则有

$$P(A) = \int_A dP(x) = \int_A f(x) dP(x). \quad (1.1.4)$$

### 1.1.2 反函数及分位数

在理论与实际问题中, 经常要知道分布函数  $F(x)$  的反函数的值, 即  $F^{-1}(p)$  的值, 其中  $0 < p < 1$ , (如  $p = 0.05, p = 0.95$  等); 通常称为分位数或分位点. 由于  $F(x)$  非降, 不一定连续, 因此  $F^{-1}(p)$  的值可能

不存在或不唯一. 以下定义保证了分位数的存在性和唯一性.

定义 1.1.1 分布函数  $F(x)$  的  $p$  分位数或  $p$  分位点, 定义为  $x_p = \inf \{x: F(x) \geq p\}$ .  $I$

由下确界的定义知  $x_p$  存在且唯一. 也容易证明, 分位数  $x_p$  满足

数, 它有以下重要性质.

定理 1.1.1 函数  $g(x) = E|X - c|$  在中位数  $C = x_{0.5}$  时达到最小值. 证明 设  $F(x)$  为分布函数, 若  $c < x_{0.5}$ , 则

由  $C < X_{0.5}$  及  $X_{0.5}$

的定义知

5

1

贝  $\int_j F(x,y) < p$ ; 若, 则  $F(xz)$

i) 若  $X' < x$

ii)  $F(X_{p-0}) \wedge P \wedge F(X_p)$ ; 若  $x$  为  $F(x)$  的连续点, 则  $F(x_p) = P$ . 常用的分位点有  $X_{0.05}, X_{0.01}, X_{0.9}, X_{0.95}, X_{0.5}$  等,  $X_{0.5}$  通常称为中位

$g(c) = \int_{-\infty}^c x - c \, dF(x) - 0.5$

$(c - X_{0.5}) dF(X_{0.5}) + \int_{X_{0.5}}^c (c - x) dF(x)$

$'C) \setminus [X_{0.5}, +\infty)$

$x - c) dF(x)$

$U_j) - [0.5 - c] dF(X_{0.5}) + \int_{X_{0.5}}^c (c - x) dF(x)$

$- c) dF(X_{0.5}) +$

$\int_{X_{0.5}}^c (c - x) dF(x)$

$[ (X_{0.5} - x) + 2(x - c) - (X_{0.5} - c) ] dF(x) \cdot$

$\int_{-\infty}^c (X_{0.5} - x) dF(x) + 2 \int_{X_{0.5}}^c (x - c) dF(x)$

$+ (X_{0.5} - c) [P(X \leq X_{0.5}) - P(X < X_{0.5})]$ .

$[ \text{多} + \text{和} p(X \leq X_{0.5}) = Z * (X_{0.5} - 0) C$

从而  $g(c) \geq E(X - X_{0.5})$ ; 反之, 若  $c > X_{0.5}$ , 则  $x - c \leq 0$

$c) dF(x)$

$(x - c) dF(x)$

第一章统计分布基础

$- x) dF(x)$

$= L, \quad 0 < [(X_{0.5} - X) + (c - X_{0.5})] dF(x) + \int_{X_{0.5}}^c (c - x) dF(x)$

$J(c, +\infty)$

$-(c - X_{0.5}) dF(X_{0.5}) + \int_{X_{0.5}}^c [(X_{0.5} - x) + 2(c - x) - (c - X_{0.5})]$

$x - X_{0.5} \int_{-\infty}^c dF(x) + 2$

$(X_{0.5}, c]$

$\cdot x) dF(x) + (c - X_{0.5}) [P(X \leq X_{0.5}) - P(X < X_{0.5})]$ .

由  $X_{0.5}$  及  $X_{0.5}$  的定义知  $P(X \leq X_{0.5})$  和  $P(X > X_{0.5})$ , 从而

1.1 随机变量及其分布函数

$g(c) \wedge E(X)$ . 总之, 对于任何  $c$ ,  $g(c) \geq g(X_{0.5})$ , 且当  $c = X_{0.5}$

时等号成立.

1.1.3 特征函数和数字特征

画

分布函数虽然完全描述了一个随机变量的概率结构, 但也存在一些不能令人满意的地方. 比如, 只能保证分布函数是单调非降、有界、右连续的, 而不能保证它是一致连续和绝对连续的. 因此分布函数的分析性质并不完美, 通常引进其他工具作为对分布函数的有益补充.

随机变量  $X$  的数学期望或均值为  $E(X) = M$ ; 方差为  $\text{Var}(X) = \sigma^2$ , 而  $\sigma$  通常称为均方差或标准差. 矩母函数和特征函数分别定义为

$E(e^{tX})$  和平  $E(e^{itX}) = M(t)$ . 常用的性质有

i)  $M(A)(0) = E(X^k)$ ,  $M(t) = 1 + itE(X) + \frac{t^2}{2} E(X^2) + \dots$

ii)  $M(1) = 1 + itE(X)$ ,  $M(t) = 1 + itE(X) + \frac{t^2}{2} E(X^2) + \dots$

iii) 给定随机向量  $X = (X_1, \dots, X_n)$ , 其分量  $X_1, \dots, X_n$  相互独立的充要条件为

$\phi(z_1, \dots, z_n) = \prod_{j=1}^n \phi_j(z_j)$ . 由性质 ii) 可知, 特征函数和矩母函数的 Taylor 展开式的系数

为随

机变量 $X$ 的各阶矩.近年来,特征函数对数的Taylor展开式的系数也经常用到,今介绍如下.

定义1.1.2累积量(cumulant).  $\log^*(t)$ 的Taylor展开式的系数称为累积量,若  $\log^*p(t) = \sum_{r=1}^{\infty} \frac{K_r}{r!} t^r$

则系数 $K_r$ 称为 $r$ 阶累积量或半不变量.随机变量 $X$ 的累积量 $K_r$ 与其各阶矩 $a_r$ 有密切的关系.

其常用性质有 0 累积量与各阶矩可互相表示.其前3阶矩的关系为:

$K_1 = a_1, K_2 = a_2 - a_1^2, K_3 = a_3 - 3a_1a_2 + 2(a_1)^3$ ;  $a_2 = K_2 + (K_1)^2, a_3 = K_3 + 3K_1K_2 + (K_1)^3$ ;

■

ii) 若 $X, Y$ 独立, 则有 $K_r(X+Y) = K_r(X) + K_r(Y)$ ; iii)  $K_r(X+C) = K_r(X), (r \geq 1)$ , 其中 $C$ 为任意常数. 以下数字特征也是统计学中常见的.

随机变量 $X$ 的三阶中心矩与四阶中心矩分别反映了密度函数的偏度与峰度.其定义如下:

偏度系数 $y_3 = a_3/a_2^{3/2}$ ; 峰度系数 $y_4 = a_4/a_2^2 - 3$ ; 其中 $a_r = E(X - EX)^r$

$EX = \mu, \sigma^2 = \text{Var}(X)$ . 容易验证, 若 $X$ 为正态分布, 则有 $y_3 = 0, y_4 = 3$

0. 因此正态分布可作为偏度与峰度的比较标准. 另外, 密度函数的峰

值称为众数(mode). 以下与条件期望有关的公式非常重要, 今后将多次用到.

$E(X) = E_r[E(X|T)]$ ; (1. 1.4)  $\text{Var}(X) = E_r[\text{Var}(X|T)] + \text{Var}_r[E(X|T)]$ . (1. 1.5)

另外, 若 $x = (x_1, \dots, x_n)^T, r = (r_1, \dots, r_n)^T$ 为随机向量, 则其期望与方差定义为  $E(X)$

$\Delta (EX_1, \dots, EX_n, \text{Cov}(X, r))$  ( $n \times m$ );  $a_{ij} = \text{Cov}(x_i, x_j); \text{Var}(X)$  4:

$\text{Cov}(X, J) n \times n$ . 随机向量的期望与方差的常用公式有

i)  $\text{Cov}(x, y) = E[(x - EX)(y - EY)^T] = E(xy^T) - (EX)(EY)^T$ ;  $\text{Var}(X) = E(XX^T) - (EX)(EX)^T$ ;

ii)  $\text{Cov}(J^T V, BY) = 4 \text{Cov}(X, Y) B^T$ ;

iii)  $E||^2 = ||EX||^2 + \text{tr}[\text{Var}(X)]$ ; (1. 1.6)

$X, Y \sim N(\mu, \Sigma)$

iv) 设 $V = 1$ , 则  $\text{Var}(J) =$

$\text{Var}(X)$  1.1.4 经验分布函数

$(\text{Var}(X), \text{Cov}(A, X^2),$

.

易见, 经验分布函数可表示为

$F_n(x)$

$\sum_{j=1}^n I_{T_j} \leq 1$

(1. 1.7)

其中 $I_1, \dots, I_n$ 为独立同分布, 取0, 1两个值的随机变量; 即 $P\{Y_i = 1\} = P\{X \leq x\}$

$= F(x), P(Y_i = 0) = 1 - F(x)$ . 并且有  $E(y_j) = F(x), \text{Var}(y_j) = F(x)[1 - F(x)]$

定理 1. 1. 2 若  $X_1, \dots, X_n$  为 i. i. d. 样本,  $X \sim x$ , 则对  $\forall x$  有 (1)  $F_n(x) \xrightarrow{p} F(x)$  (a. e.);

(2)  $V(F_n(x) - F(x)) \sim V(0, F(x)[1 - F(x)])$ , 其中(a. e.)表示几乎处处收敛,  $l$  表示依分布收敛.

(3)  $\text{Var}(J^2) = J$

经验分布函数是分布函数很好的近似. 考虑独立同分布的样本 $X_1, \dots, X_n$ , 这时 $X$ 服从未知分布 $F(x)$ . 给定 $n$ 表示 $X$ 的个数. 易见, 概率 $F(x) = P(X \leq x)$ 可用频率 $\hat{F}_n(x)$ 来逼近.

定义1.1.3给定独立同分布的样本 $X_1, \dots, X_n$ , 总和 $S_n$ , 经验分布函数 定义为  $F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{(-\infty, x]}(X_i)$ .

|

## 第一章统计分布基础

### 1.2常见的离散型分布

7

证明 由(1.1.7)式,  $F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{(-\infty, x]}(X_i)$

$Y_i$ —根据强大数定律有

$\frac{1}{n} \sum_{i=1}^n I_{(-\infty, x]}(X_i) \xrightarrow{a.s.} F(x)$ , 即

$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n I_{(-\infty, x]}(X_i) = F(x)$

$Y_i = F_n(x) \rightarrow$

$F(x)$ , 且

$\lim_{n \rightarrow \infty} F_n(x) = F(x)$

$= F(x)$  (a.e.).

由中心极限定理有

$\sqrt{n}(F_n(x) - F(x)) \xrightarrow{d} N(0, F(x)(1 - F(x)))$

4=1

$- F(x) \mid \sqrt{n}(F_n(x) - F(x)) \xrightarrow{d} N(0, F(x)(1 - F(x)))$ . I

这个定理表明, 经验分布函数是样本分布函数很好的近似. 更深入 的讨论可参见陈希孺 (1981, 1999).

### 1.2常见的离散型分布

本节介绍若干常见的离散型分布, 其分布函数与密度函数的关系可 参见(1. 1.3)式.

(1) 单点分布(退化分布).

即 $X$ 以概率1取常数 $a$ :  $P(X = a) = 1$ . 其分布函数为 $F(x) = 1 \mid x \geq a \mid$ .

(2) 离散均匀分布 $X \sim U(m)$ .

即 $X$ 等可能性地取 $1, \dots, m$ 之中的任何一个值:  $P(X = i) = 1/m, i = 1, \dots, m$ ;

这相当于从有标号 $1, \dots, m$ 的 $m$ 个球中任取一个球. 其密 度函数为:  $P(x, m) = 1/m \mid x = 1, \dots, m \mid$

$W^x \sim m$  (但 $x$ 取正整数).  $X$ 的期望与方 差分别为  $E(X) = (m + 1)/2, \text{Var}(X) = (m^2 - 1)/12$ .

(3) 两点分布 $X \sim b(1, p)$ .

即 $X$ 仅取0, 1两个值:  $P(X=1)=p, P(X=0)=1-p$ . 实际问题 中,  $X=1$ 表示某事件4成功、发生, 或某产品为次品等情形;  $X=0$ 表 示某事件4 失败、不发生, 或某产品为正品等情形.  $X$ 的密度函数为:

$p(x, p) = p^x (1-p)^{1-x}, x=0, 1$ . 该式可化为指数形式:  $p(x, p) = \exp\{x \log p + (1-x) \log(1-p)\}$

其中 $\log$

n

n 常记为  $h(p)$ , 它表示成功概率与失败概率之比的对数.

## 8 第一章统计分布基础

易见 $0 < p < 1$ 时有  $-\infty < \log p < 0$ , 其变化范围为整个数轴, 应

用上比较方便. 两点分布的特征函数为  $\phi(t) = 1 - p + pe^{it}$ , 并且有  $E(X) = p$ ,

$\text{Var}(X) = p(1 - p)$ . (4)二项分布  $X \sim b(n, p)$ .

$X$  表示 $n$ 次相互独立的试验中事件4 发生(或成功)的次数, 而事 件4发生的概率为

$p$ . 因此  $p(x = 0) = (1-p)^n$ , 且记为

$b(x; n, p)$ .  $X$  的密度函数为

$p(x, p) = \binom{n}{x} p^x (1-p)^{n-x}, x = 0, 1, \dots, n$ .

二项分布有以下主要性质:

i) 重要分解.二项分布又可表示为独立同分布的两点分布 $X_1, \dots, X_n$ 之和.即

$X = \sum_{i=1}^n X_i$ , 第 $i$ 次试验A发生,  $X_i = 0$ , 第 $i$ 次试验A不发生

其中 $X_i \sim B(1, p)$ 为两点分布.由此分解式可推出二项分布的许多基本性质, 例如

ii)  $E(X) = np$ ,  $Var(X) = np(1-p)$ ;

iii) 特征函数 $= [(1-p) + pe^{it}]^n$ ;

iv) 可加性, 即若 $X_1 \sim B(n_1, p), X_2 \sim B(n_2, p)$ , 且独立, 则 $X_1 + X_2 \sim B(n_1 + n_2, p)$ ;

v)  $(X - np)$ 的渐近分布为标准正态分布 $N(0, 1)$  (当 $n \rightarrow +\infty$ ).

以下介绍二项分布的分布函数以 $F(x)$ 的计算公式,  $F(x) = P(X \leq x)$

$= \sum_{k=0}^x \binom{n}{k} p^k (1-p)^{n-k}$

$B(k | n, p)$ , 其中 $[*]$ 表示 $x$ 的整数部分.今记 $B(i | n, p) = \sum_{k=0}^i \binom{n}{k} p^k (1-p)^{n-k}$

$i$

$= 1 - B(n-i | n, p)$ :  $F(x)$ 与不完全函数有密切关系

系.今简要介绍如下.

熟知, 芦函数定义为 $\int_0^x t^{p-1} (1-t)^{q-1} dt = I_p^q(x)$ ; 它可

由 $r$ 函数表示为 $p(p, q) = r(p)r(q)/r(p+q)$ , 其中 $r$ 函数为 $r(p) = \int_0^1 t^{p-1} (1-t)^{p-1} dt$ . 不完全函数定义为:

,

## 1.2 常见的离散型分布 9 (1.2.1)

不完全 $\beta$ 函数有以下主要性质(见习题): i)  $\beta(p, q) + \beta(q, p) = 1$ ;

(1.2.2)

ii)  $\beta(i | n, p) = \beta(n-i+1 | n, 1-p)$ ; (1.2.3)

iii)  $B(i | n, p) = 1 - B(n-i+1 | n, 1-p)$ . (1.2.4)

上式可表示为: $P(X \leq i) = P(Z \leq i)$ , 其中 $Z$ 服从 $\beta$ 分布 $BE(p, q)$ (见后面连续型分布).

(5) 几何分布 $X \sim G(p)$ 或 $Z \sim G(a)$ . ;  $r$ 表示首次成功所需试验次数. 因此有 $P(X = k) = (1-p)^{k-1} p$

$0 < p < 1, p \neq 0, 1, \dots, \infty$ . 直接计算可知

i) 特征函数 $\phi(t) = \sum_{k=1}^{\infty} p(1-p)^{k-1} e^{ikt} = \frac{pe^{it}}{1-(1-p)e^{it}}$ ; 并且有 $E(X) = \frac{1}{p}$ ,

$Var(X) = \frac{1-p}{p^2}$

ii) 几何分布有无记忆性, 即 $P(X \leq n+k | X > n) = P(X \leq k)$ . (6) Pascal 分布  $X \sim PA(r, p)$ .

$X$ 表示取得 $r$ 次成功所需试验次数, 易见 $r = 1$ 时,  $PA(1, p)$ 即为几

何分布. 并且有  $p(x = r+k | X \geq r) = p(x = k)$

$r$ . Pascal分布有以下基本性质:  $r$

i)  $X = \sum_{i=1}^r X_i$ , 其中 $X_1, \dots, X_r$ 为i.i.d.,  $X_i$ 为第 $i-1$ 次成功到第 $i$ 次成功所经历的试验次数, 因此 $X_k \sim PA(1, p)$ 为几何分布;

ii)  $E(X) = r/p, Var(X) = r(1-p)/p^2$ ;

iii) 可加性(包括几何分布). 即若 $X_1 \sim PA(r_1, p), X_2 \sim PA(r_2, p)$ 且独立, 则有 $X_1 + X_2 \sim PA(r_1 + r_2, p)$ .

(7) 负二项分布 $X \sim NB(r, p)$ . 若 $Y \sim PA(r, p)$ , 则称 $X = Y - r$ 服从负二项分布 $NB(r, p)$ .  $J$ 表示

次成功所经历失败的次数.

$P(X = i) = \binom{i+r-1}{r-1} p^r (1-p)^i$

或  $\binom{i+r-1}{i} p^r (1-p)^i$

$Z = 0, 1, 2, \dots$ . 以上概率为 $a(1-p)^{-T}$ 展开式中 $a$ 的第 $i$ 项系数,  $a = 1-p^r$ , 因此称为负二项分布. 它有以下基本性质:

i)  $X \sim NB(r, p)$ , 则 $Y = X + 1 \sim NB(r+1, p)$ ;

## 第一章统计分布基础

ii)  $X = \sum_{k=1}^n X_k$ , 其中  $X_1, \dots, X_n$  为 i.i.d.,  $X_k \sim \text{NB}(1, 6)$ ;  $A=1$

iii) 特征函数  $\phi(t) = \sum_{q=0}^{\infty} \frac{r^q (1-r)^q}{q!} t^q$ ,  $q=1-6$ ; 并且有

$E(X) = \frac{r}{1-r}$ ,  $\text{Var}(X) = \frac{r}{(1-r)^2}$

iv) 可加性. 即若  $X \sim \text{NB}(r, 1)$ ,  $Y \sim \text{NB}(r, 1)$  且独立, 则  $X+Y \sim \text{NB}(r, 2)$

$\text{NB}(r, 2)$

v)  $P(X=i) = \binom{i-1}{r-1} r^r (1-r)^{i-r}$  (其证明类似于 (1.2.3) 式).

(8) 超几何分布  $W$  件产品, 有  $M$  个次品, 不放回抽取  $n$  件产品,  $X$  表示次品的个

数. 因此有  $P(X=i) = \frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}}$

基本性质:

i)  $E(X) = n \frac{M}{N}$ ,  $\text{Var}(X) = n \frac{M}{N} \frac{N-M}{N} \frac{N-n}{N-1}$

ii)  $\frac{M}{N} \rightarrow p$ ,  $\frac{n}{N} \rightarrow 0$ ,  $N \rightarrow \infty$  时,  $X \sim \text{Poisson}(np)$

iii)  $\frac{M}{N} \rightarrow p$ ,  $\frac{n}{N} \rightarrow 0$ ,  $N \rightarrow \infty$  时,  $X \sim \text{Poisson}(np)$

• 它有以下

ii) 若  $n/V$  很小, 则  $P(X=i) \approx \frac{M^i}{i!} e^{-M}$ , 即可化为次品率  $M/N$  不变, 有放回的抽样.

(9) Poisson 分布  $X \sim P(A)$ .

$P(X=i) = \frac{e^{-A} A^i}{i!}$  其密度函数为  $p(x) = \frac{e^{-A} A^x}{x!}$ .

它有以下基本性质:

i) 特征函数  $\phi(t) = e^{-A(1-e^{-t})}$ ,  $E(X) = A$ ,  $\text{Var}(X) = A$

ii) 分布函数  $F(i) = P(X \leq i)$  可由不完全  $\gamma$  函数  $\gamma(A, i+1)$  表示为

$F(i) = 1 - \frac{\Gamma(A)}{\Gamma(A) \Gamma(i+1)} \int_0^A t^i e^{-t} dt$

(1.2.5)

其中  $z$  服从  $\gamma$  分布  $r(i, z+1)$  (见后面连续型分布, 其证明类似于 (1.2.3) 式);

iii) 可加性. 即若  $X_1 \sim P(A_1)$ ,  $X_2 \sim P(A_2)$  且相互独立, 则  $X_1 + X_2 \sim P(A_1 + A_2)$

iv) 密度函数  $p(x, A) = P(X=x)$  为  $x$  的增函数, 若  $x < A$ ; 为  $x$  的减函数, 若  $x > A$

v) 条件分布. 若  $X_1, X_2$  且相互独立, 则  $(X_1, X_2)$  服从二项分布  $b(k, p)$ , 其中  $p = A/(A_1 + A_2)$ ,  $i=1, 2$

vi) 当  $A \rightarrow \infty$  时,  $(X-A)/\sqrt{A}$  的渐近分布为标准正态分布  $N(0, 1)$ .

(10) 多元分布  $X = (X_1, \dots, X_k)^T \sim \text{Af}(1, \gamma)$ ,  $\gamma = (\gamma_1, \dots, \gamma_k)^T$

$X$  和  $\gamma$  为  $A$ : 维向量.

### 1.2 常见的离散型分布

假设有  $A$ : 个事件  $A_1, \dots, A_k$  为样本空间的一个划分, 即这个事

件两两不相容,  $\sum_{i=1}^k P(A_i) = 1$  且有  $\sum_{i=1}^k X_i = 1$

试验中  $A_i$  出现,  $X_i = 1$ , 试验中不出现

$X_i = 0$ .

i=1 易见, 若  $k=2$ , 则  $X$  即为二点分布:  $A$

基本性质:

i) 密度函数  $P(X_1=x_1, \dots, X_k=x_k) = \prod_{i=1}^k p_i^{x_i} (1-p_i)^{1-x_i}$ ,  $x_i \in \{0, 1\}$ ,  $\sum_{i=1}^k x_i = 1$ , 其他为 0;

ii)  $X_i \sim \text{Bernoulli}(p_i)$ ,  $E(X_i) = p_i$ ,  $i=1, \dots, k$ , 且有

$\text{Cov}(X_i, X_j) = -p_i p_j$

$i \neq j$

$i=1, \dots, k$  的

,

;

$H_i$ )期望与方差的向量形式为

$$E(X) = \tau, \text{Var}(X) = \text{diag}(\tau) - \tau\tau^T,$$

其中 $\text{diag}(\tau) = \text{diag}(\tau_1, \dots, \tau_k)$  为 $k$ 阶对角阵.

(1.2.6)

易见, 若 $k=2$ , 则 $X$ 即为二项分布 $B(n, \tau)$ ; 若 $k=1$ , 则

$X$ 为多点分布 $MN(1, \tau)$ . 多项分布有以下基本性质: i) 其密度可表示为

$$p(x) = \frac{n!}{x_1! \cdots x_k!} \tau_1^{x_1} \cdots \tau_k^{x_k}, \quad x_1 + \cdots + x_k = n,$$

ii) 重要分解. 多项分布 $p(x) = p(x_1, \dots, x_k)$ 可表示为独立同分布的多点分布之和. 即

$$p(x) = \sum_{m=1}^n p_m(x_1, \dots, x_k),$$

$X$

$m=1$

0  $r$ . (1.2.7)

其中 $x_1, \dots, x_k$ 为独立同分布的多点分布,  $x_l$  且有  $1$ , 第 $i$ 次试验 $x_l$ 出现,

0, 第 $m$ 次试验 $x_l$ 不出现, 也有  $\sim B(n, \tau_l)$ ;  $i = 1, \dots, k, m = 1,$

$\dots, n$ .

$-77, '$

,

(H) 多项分布 $X = (X_1, \dots, X_k)^T \sim \text{MTV}(a, \tau)$ . 条件及符号与多点分布相同, 但 $X_i$ 表示 $i$ 次试验中 $x_i$ 出现的次数

"  $= 1, \dots, k$ , 因此有

$$P(X_i = x_i, \dots, X_k = x_k) =$$

$$\frac{n!}{x_1! \cdots x_k!} \tau_1^{x_1} \cdots \tau_k^{x_k}, \quad x_1 + \cdots + x_k = n.$$

1

,

$= 4, A_2 = A$ . 多点分布有以下

$$\text{iii) } E(X) = \tau, \text{Var}(X) = \text{diag}(\tau) - \tau\tau^T, \text{ 因为 } X =$$

$X_m \sim \text{扁}(1, \tau)$ .

iv) 特征函数为

$$p(t) = E[e^{it^T X}] = E[e^{it^T X_1}]$$

v) 条件分布(与Poisson分布的关系则).

$; i =$

$\sim \text{脚}(n, \tau)$ , 其中 $\tau = (\tau_1, \dots, \tau_k)^T$ ,  $\tau_i = A_i /$

若 $x_t = p(A_t) f_i =$

$i,$

相互独立,

$k$

表1.2. 1列出了常见的离散型分布(可参见方开泰, 许建伦, 1987; Zacks, 1981; 茆诗松等, 1998; Shao, 1998) ♦

表 1.2.1 常见的离散型分布

$1/m, X=1, \dots, m$

$$\text{又 } e^{it} / m = e^{it} f(1 - e^{it/m}) / [n(1 - e^{it/m})]$$

离散均匀分布 密度函数

特征函数  $k=1$

$U(m)$

期望  $(m+1)/2$  方差  $(m^2-1)/12$

$(1-0) + \text{加} i f$

期望 0



方差  $0(1-0)$   
 两点分布  
 $b(\cdot, 0)$   
 二项分布 密度函数  
 $b(n, 0)$   
 几何分布  
 $Q^{(i-e)}n \sim x=0, 1, 2, \dots, n$   
 特征函数  $[(1-6) + 0e^{i\lambda}]^n$   
 期望  $n0$   
 方差  $n0(1-6)$   
 密度函数  $0(1-ey) \sim x=i, 2, \dots, 00$   
 $C(9)$  期望 方差  
 Pascal分布  
 密度函数 特征函数  
 $(1-e^{-\lambda})^x [1 + \lambda e^{-\lambda}]^n$   
 密度函数 矿(1  
 $x=0, 1$   
 特征函数  
 特征函数  
 $0e^{i\lambda} [1 - (1-0)e^{i\lambda}]^n$   
 $0 \sim$   
 $(i-0)e^{i\lambda}$   
 $] = [i r \{e^{i\lambda} + \dots + 7 r k e^{i\lambda} k\} n.$   
 $k$   
 $,$   
 $x^r$   
 第一章统计分布基础  
 $,$   
 $+ \dots + A = n |$   
 $A; i=1, \dots, k.$   
 $m=1$   
 1.3常见的连续型分布  
 13  
 $PA(r, 0)$   
 期望  $r0 \sim l$   
 方差  $r(i-e)e^{i\lambda}$   
 续表  
 $Xx = 0, 1, 2, -$   
 负二项分布 密度函数 特征函数  
 $(1-p)^r [1 - p e^{i\lambda}]^{-r}$   
 賦  $r, (2)$   
 超几何分布 密度函数  
 Poisson 分布  
 $P(A)$   
 方差 密度函数  
 特征函数  
 $M N-M N-n UN N N-\backslash$

$e^{-Ax} = \sum_{k=0}^{\infty} \frac{(-A)^k x^k}{k!} \exp(-Ax)$   
 多点分布 密度函数  
 $k$   
 $X \geq 1$   
 壓 (1, 77) 多项分布  
 壓 (n, 77)  
 特征函数  $\mathbb{E}[e^{itf}] = \sum_{k=1}^n p_k e^{itf_k}$  期望  $\mathbb{E}$   
 方差  $\text{diag}(\mathbb{E}[X]) - \mathbb{E}[X]\mathbb{E}[X]^T$  密度函数  $n!$

$*$   
 $V = -J I$   
 $;$   
 $\text{矿}[1 - (1 - W)] - r$   
 $\text{冰}_1(1 - 0) \text{re} \sim 2(1 - 0)$   
 期望  $M$   $nN$   
 期望 方差  
 期望  $A$  方差  $A$   
 特征函数  
 $(77)eu, + \dots + irke', k)n$   
 $/ M \setminus N - M \setminus / / N \setminus$   
 $/, \max(0, A/ + z_i - TV)$   
 $\wedge m) / \setminus n) \text{ 矣 } x \text{ 矣 } \min( )$   
 期望  $mr$   
 方差  $\text{ra}[\text{diag}(tt) - ]$

### 1.3 常见的连续型分布

本节介绍若干常见的连续型分布，其分布函数与密度函数的关系可参见(1. L2)式。

$7T, 1 \dots 77:$

14

### 第一章统计分布基础

(1) 均匀分布  $x \sim R_d, e2$ 。其密度函数和分布函数分别为

$I \text{ ②、}$

$0$ , 其他,  $x < \text{沒},,$

$\text{②} < 02,$

$x^{02}.$

, / 卜».

均匀分布有以下基本性质:

i)  $E(X) = (\wedge^j + \wedge^2)/2 = /, l + cr/2$ ;  $\text{Var}(X) = ( )^2/12 = < r^2/12$ ; ii)

$R(0\}, e2) = 0, +(e2 \text{ 即若 } y \sim /?(o, i), \text{ 则又} =$

$A + (\text{么} \sim \text{尺}(A \text{ 為}); \text{反之, 若 } x \sim r(h), \text{ 则 } (x - A)/(6 - h) = y \sim \text{尺}(o, i);$

iii) 若  $X - F(x)$  为连续型分布函数, 贝  $IJ Y = F(X) \sim /?(0, 1); \text{反 之, 若 } r \sim \text{尺}(o, i), Mx = r-, (y)$

这一性质有广泛的应用. 若  $x \sim \text{尸}(\text{幻}, X, , \dots$  人为其  $i. i. d.$  样本,

取次序统计量  $(4)x(1)^x(2) \leq \dots cv(4)$ , 贝  $u r(1) = \text{尺}(x(1)), y(2) =$

为均匀分布  $/?(o, i)$  的次序统计量, 这些次序统计量有很好的性质, 见1.6节.

另外, 根据以上性质可由均匀分布  $7?(0, 1)$  的随机数, 产生分布  $M$  幻的随机数, 见以下例题.

例1.3.1由均匀分布 $U(0,1)$ 的随机数,产生指数分布的随机数.

解 设服从指数分布 $X \sim E(\lambda)$ , 记为 $X \sim E(\lambda)$ . 先求其反函数:

$y = 1 - e^{-\lambda x}$ ,  $x = -\frac{1}{\lambda} \log(1 - y) = F^{-1}(y)$ ,  $0 < y < 1$ . 故由性质iii), 若 $y \sim U(0,1)$ , 则 $X = F^{-1}(Y) = -\frac{1}{\lambda} \log(1 - y) \sim E(\lambda)$ 为指数分布. 因此, 由 $y \sim U(0,1)$ 的随机数可产生指数分布 $X \sim E(\lambda)$ 的随机数. 特别, 由 $x = -\log(1 - y)$ 可产生 $X \sim E(1)$ 的随机数. 一般讲, 常见连续型分布的随机数皆可由均匀分布 $U(0,1)$ 的随机数产生.

1.3.2 指数分布 $X \sim E(\lambda)$ 的密度函数和分布函数分别为

$f(x) = \lambda e^{-\lambda x}$ ,  $F(x) = 1 - e^{-\lambda x}$

$x \geq 0$ ,  $f(x) = 0$ ,  $F(x) = 0$

1.3.3 标准均匀分布 $X \sim U(0,1)$ , 这时 $f(x) = 1$ ,  $F(x) = x$

0,  $x < 0$

1.3.4 另一种常见的形式为 $X \sim M(\mu, \sigma)$ 其中 $\mu$ 为位置参数,  $\sigma$ 为尺度参数;  $f(x) = \frac{1}{\sigma} e^{-\frac{x-\mu}{\sigma}}$

(2) 指数分布 $X \sim E(\lambda)$ . 其密度函数和分布函数分别为

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

$F(x) = 1 - e^{-\lambda x}$

$x \geq 0$ ,  $f(x) = 0$ ,  $F(x) = 0$

1.3.5 常见的连续型分布

1.3.6 指数分布

$f(x) = \lambda e^{-\lambda x}$ ,  $F(x) = 1 - e^{-\lambda x}$ ,  $x \geq 0$

0,  $x < 0$

指数分布有若干常见的形式.  $\lambda = 1$ 时为标准分布:  $X \sim E(1)$ , 其密度函数为 $f(x) = e^{-x}$

$x \geq 0$ ; 记 $\lambda = \frac{1}{\theta}$ , 则 $\theta$ 为尺度参数, 这时 $f(x) = \frac{1}{\theta} e^{-x/\theta}$  ( $x \geq 0$ ).

指数分布有以下基本性质:

i)  $E(X) = \frac{1}{\lambda}$ ,  $Var(X) = \frac{1}{\lambda^2}$ ;

ii) 若 $K \sim U(0,1)$ , 则 $X = -\log(1 - K) = -\log(K)$ ,  $X \sim E(1)$

(例 1.3.1 中  $\lambda = 1$ );

iii) 无记忆性:  $P(X > s + t | X > s) = P(X > t)$ ; iv) 与Poisson分布的关系. 设 $N(t)$ 为 $(0, t]$ 时间内的故障数(或电话数), 通常 $N(t) \sim P(\lambda t)$ , 则首次故障(或电话)出现的时间 $Y = E(X)$ . 因 $P(Y > t) = P(N(t) = 0) = e^{-\lambda t}$ ,  $1 - e^{-\lambda t} = F(t) = F(\lambda t)$ . 或

(3) 带有位置参数的指数分布 $X \sim E(\lambda, \mu)$

若 $X - \mu \sim E(\lambda)$ , 则 $X \sim E(\lambda, \mu)$ 的指数分布, 其密度函数为

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

0,  $x < \mu$

其中 $\mu$ 为位置参数,  $\lambda$ 为尺度参数. 当 $\mu = 0$ 时,  $X \sim E(\lambda)$ . 特别, 当 $\lambda = 1$ 时,  $X \sim M(1, \mu)$  (1, 科), 其密度函数为

$f(x) = e^{-(x-\mu)}$ ,  $x \geq \mu$

0,  $x < \mu$ . 对于分布 $X \sim E(\lambda, \mu)$ ,  $\lambda$ 的期望和方差分别为

$E(X) = \mu + \frac{1}{\lambda}$ ,  $Var(X) = \frac{1}{\lambda^2}$ . (4) 位置尺度参数分布族 $X \sim P(\mu, \sigma)$ 或

如果 $x$ 的分布密度可表示为 $f(x) = \frac{1}{\sigma} e^{-\frac{x-\mu}{\sigma}}$ , 则称 $X$ 服从位置尺度参数分布, 其中 $f(x)$ 为某一密度函数. 若

$\sigma = 1$ , 即 $X \sim E(\lambda, \mu)$ , 则称 $X$ 服从位置参数分布; 若 $\mu = 0$ , 即则称 $X$ 服从尺度参数分布; 若 $\mu = 0$ ,  $\sigma = 1$ , 即 $X \sim E(\lambda)$ , 则称 $X$ 服从标准分布. 易见, 正态分布 $X \sim N(\mu, \sigma^2)$ , 均匀分布 $X \sim U(a, b)$ , 指数分布 $X \sim E(\lambda)$ 等都是位置尺度参数分布族.

容易验证, 若 $X \sim E(\lambda)$ 服从标准分布 $P(0,1)$ ; 反之,  $X \sim E(\lambda)$

,  $\lambda = 1$

, 则 $Y = \lambda X + \mu$ 服从带有位置尺度参数分布 $P(\mu, \sigma)$

故 $P(Y \leq 0) = 0$

$A=a'$   
16

## 第一章统计分布基础

若 $\hat{r} \sim P(0, J)$ , 则 $r =$  服从一般位置尺度参数分布这些转换关系都是经常用到的.

以上定义可推广到更一般的情形. 如果样本 $X = (X_1, \dots, X_J)$ 的密度函数可表示为以下形式, 则称它服从位置尺度参数分布:

其中 $1=(1, \dots, 1)T$ . 类似地, 若则  
 $Y =$

服从标准分布

$a \sim (0, 1)$ ; 反之, 若 $Y \sim P(0, 1)$ , 则 $Y=aX+\hat{l}$ 服从一般位置尺度参数分布  $P(10)$ .

(5) 正态分布 $X \sim V(M, a^2)$ . 其密度函数为

$f(x; \mu, a^2) = \frac{1}{\sqrt{2\pi}a} e^{-\frac{(x-\mu)^2}{2a^2}}$

正态分布是概率统计中最为常见的分布, 又称Gauss分布.  $\mu=0, a=1$  时为标准正态分布, 其密度函数和分布函数常记为 $\phi(x)$ 和 $\Phi(x)$ .

正态分布的常用性质有:

i) 特征函数 $\phi(t) = e^{-\frac{a^2 t^2}{2}}$ , 并且有:

$E(X) = \mu, \text{Var}(X) = a^2, E(X-\mu)^3 = 0, E(X-\mu)^4 = 3a^4$ . 因此正态分布的偏度系数 $\gamma_1=0$ , 峰度系数 $\gamma_2=3$ .

ii)  $\phi(-x) = \phi(x)$ ;  $\Phi(x) = 1 - \Phi(-x)$ . 特别, 若记标准正态分布的 $\alpha$ 分位点为 $z_\alpha$ , 则有一心 另外若 $X \sim V(M, a^2)$ , 则 $X$ 的 $\alpha$ 分位点为 $\mu - z_\alpha a$ .

iii) 若 $X_1, \dots, X_n$  又为独立同分布的标准正态分布, 则其平方和服从  $\chi^2(n)$  分布.

iv) 任意个正态分布的线性组合仍然服从正态分布. (6) 对数正态分布 $X \sim \text{LN}(\mu, a^2)$ .

若 $\ln X = Y \sim N(\mu, a^2)$ ,  $X > 0$ , 则称 $X$ 服从对数正态分布. 其密度 函数为

$f(x; \mu, a^2) = \frac{1}{x\sqrt{2\pi}a} e^{-\frac{(\ln x - \mu)^2}{2a^2}}, x > 0$

注意对数正态分布与正态分布的不同点: 对数正态分布的密度函数曲线 是偏态的, 而且其取值范围仅限于 $x > 0$ ; 其期望和方差分别为

$E(X) = e^{\mu + \frac{a^2}{2}}, \text{Var}(X) = e^{2\mu} (e^{a^2} - 1)$ . (7) t分布 $X \sim t(n)$ .

### 1.3 常见的连续型分布

17

若 $x = r/Z$ ;  $r \sim \chi^2(n)$ ,  $z \sim N(0, 1)$ 且独立, 则 $x$ 服从t分布  $t(n)$ . 其分布密度可表示为

$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$

$f$ 分布为对称分布, 期望和方差分别为 $E(X) = 0, \text{Var}(X) = \frac{n}{n-2} (n > 2)$ ; 且有 $X^2 \sim F(1, n)$ . 另外, 当 $n \rightarrow \infty$  时,  $t(n) \rightarrow N(0, 1)$ .

(8) F 分布  $X \sim F(n, m)$ . 若 $Y \sim \chi^2(n)$ ,  $Z \sim \chi^2(m)$ 且独立, 则 $X$ 服从F分布 $F(n, m)$ .

其期望与自由度 $n$ 无关:  $E[F(n, m)] = \frac{m}{m-2} (m > 2)$ . 另外其分 位点有关系:  $F_\alpha(n, m) = [F_{1-\alpha}(m, n)]^{-1}$ .

(9)  $\Gamma$ 分布 $X \sim \Gamma(a, \lambda)$ 或其密度函数为

$f(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x}, x > 0, \lambda > 0$ .

这是特别重要的分布之一, 其中 $\lambda^{-1} = \theta$ 为尺度参数,  $a$ 为自由度,  $\lambda$  是形状参数. 易见,  $\Gamma(1, \lambda)$ , 即 $\lambda=1$ 时 $\Gamma$ 分布为指数分布 $E(\lambda)$ .

$\Gamma$ 分布有以下基本性质: i)  $\Gamma$ 分布的特征函数以及期望和方差分别为

$\phi(t) = (1 - \frac{it}{\lambda})^{-a}, E(X) = \frac{a}{\lambda}, \text{Var}(X) = \frac{a}{\lambda^2}$

(1.3.2) ii)  $\lambda = 1/2, \dots, n/2$ 时 $\Gamma$ 分布为 $\chi^2$ 分布, 即 $\Gamma(n/2, 1/2) = \chi^2(n)$ .

因此 $X \sim \chi^2(n)$ 分布的特征函数以及期望和方差分别为

$$\phi(t) = (1 - 2it)^{-n/2}, E(X) = n, \text{Var}(X) = 2n$$

$$E(X^2) = n(n+2), E(X^3) = n(n+3), \dots, E(X^k) = n(n+2)\dots(n+k-2) \quad (n \geq k)$$

(1.3.3)

另外, 若  $X \sim r(A, p)$ , 则  $X^2 \sim \chi^2(2p)$ .

iii)  $r(A, p)$  及  $\chi^2(n)$  均有可加性. 若  $X_1, \dots, X_n \sim r(A, p)$  且独立, 则有

$\sum_{i=1}^n X_i \sim r(A, np)$ . 特别, 指数分布也有可加性: 若  $X_1, \dots, X_n$  为  $i.i.d.$

$\text{Exp}(\lambda)$ , 则有  $\sum_{i=1}^n X_i \sim \text{Exp}(\lambda/n)$ ;  $i=1, \dots, n$

$2A \sim \chi^2(2rt)$ .

18

iv) 若  $X \sim \chi^2(n)$ , 则当  $n \rightarrow \infty$  时,  $X$  近似正态分布  $N(\theta, 1)$ .

(10)  $\chi^2$  分布  $X \sim \chi^2(p, q)$ . 其密度函数为  $f(x; p, q) = \frac{1}{\Gamma(p)\Gamma(q)} x^{p+q-1} e^{-x}$

第一章统计分布基础 的渐近分布为标

$\Gamma(p, q)$ . 卢分布是定义在  $[0, 1]$  区间上的连续型分布, 它可用来刻画比例型的数据. 卢分布有以下基本性质:

i) 若  $X \sim \Gamma(p, q)$ , 则  $Y = 1 - X$  标准均匀分布也是冷

分布: 若  $X \sim \Gamma(1, 1)$ , 则  $X \sim U(0, 1)$ ,  $1 - X \sim U(0, 1)$ . ii) 分布的期望和方差分别为

$$E(X) = \frac{p}{p+q}, \text{Var}(X) = \frac{pq}{(p+q)^2(p+q+1)}$$

定理1.3.1 若  $X \sim \Gamma(p, 1)$ ,  $Y \sim \Gamma(q, 1)$  且相互独立, 则  $U = \frac{X}{X+Y}$

$U \sim \Gamma(p, p+q)$  且  $U$  与  $V = X+Y$  独立.

(1.3.4)

证明

$$j = \int_0^1 \int_0^\infty f(u, v) du dv$$

做变换  $(x, y) \rightarrow (u, v)$ , 则有

$$x = uv,$$

$$y = v(1-u),$$

$dx dy = v du dv$ . 从而

则b

这个定理有广泛的应用, 以下推论是其中一部分:

推论1 若  $X \sim \chi^2(n) = \Gamma(n/2, 1/2)$ ,  $Y \sim \Gamma(m/2, 1/2)$  且独立,

$$X \sim \Gamma(n/2, 1/2), Y \sim \Gamma(m/2, 1/2) \Rightarrow X+Y \sim \Gamma((n+m)/2, 1/2)$$

$$X \sim \Gamma(n/2, 1/2) \Rightarrow X \sim \chi^2(n)$$

$v > 0$ ,

1.

$$\Rightarrow v = x + y$$

$$v u = v(1-u)$$

$$f(u, v) = f_X(x) f_Y(y) = \frac{1}{\Gamma(n/2) \Gamma(m/2)} x^{n/2-1} y^{m/2-1} e^{-(x+y)} v$$

$$= \frac{1}{\Gamma(n/2) \Gamma(m/2)} (v u)^{n/2-1} (v(1-u))^{m/2-1} e^{-v} v$$

$$= \frac{1}{\Gamma(n/2) \Gamma(m/2)} v^{n+m/2-1} u^{n/2-1} (1-u)^{m/2-1}$$

$$= C v^{n+m/2-1} u^{n/2-1} (1-u)^{m/2-1} = C v^{(n+m)/2-1} u^{n/2-1} (1-u)^{m/2-1}$$

故  $U \sim \Gamma(n/2, 1/2)$ ,  $V \sim \Gamma((n+m)/2, 1/2)$  且相互独立; 也有  $U \sim \text{Beta}(n/2, m/2)$ .

推论2 若  $B \sim \text{Beta}(p, q)$

$$B \sim \text{Beta}(p, q) \Rightarrow B \sim \Gamma(p, 1) / \Gamma(p+q, 1)$$

1.3 常见的连续型分布

$F(n, m)$ ; 反之, 若  $F \sim F(n, m)$ , 则  $B = F / (F + 1) \sim \text{Beta}(n, m)$  (关于  $F$

分布, 见(8)). ;

推论3 若  $X \sim \chi^2(n)$ ,  $Y \sim \chi^2(m)$

W

$X_{n+1}, \dots, X_{n+m}$  为 i.i.d. 且  $X_1 \sim \text{AT}(0, \sigma^2)$ , 则  $\bar{X}_n \sim \text{BF}(H)$ .

有

(11) Laplace 分布  $X \sim \text{LA}(\mu, \sigma^2)$ . 其密度函数为

19

$Z(T)$  其特征函数为  $\phi(t) = e^{it\mu - \frac{1}{2}\sigma^2 t^2}$ , 且有  $E(X) = \mu$ ;  $\text{Var}(X) = \sigma^2$ .

当  $\mu=0$  时有  $\sigma^2 = 2/\lambda$ ; 并且  $|X| \sim \text{指数分布} E(\lambda/2)$ . (12) Cauchy 分布  $X \sim \text{CA}(\mu, \sigma)$ , 其密度函数为

$$f(x) = \frac{1}{\pi} \frac{\sigma}{(x-\mu)^2 + \sigma^2}$$

其特征函数为  $e^{it\mu - \sigma|t|}$ , Cauchy 分布的特点是各阶矩都不存在. 另外, 直接验证可知, 若  $X_1, X_2$  独立同分布于  $\text{U}(0, 1)$ , 则  $Y = X_1/X_2 \sim \text{G}(0, 1)$ .

(13) Pareto 分布  $X \sim \text{PR}(a, 0)$ . 其密度函数为  $f(x; a, 0) = a\theta^a x^{-(a+1)} I_{\theta < x}$  ( $a > 0; \theta > 0$ ).

其期望和方差分别为

$E(X) = \frac{a}{a-1} \quad (a > 1)$ ,  $\text{Var}(X) = \frac{a}{(a-1)^2(a-2)}$  ( $a > 2$ ).

另外, 若  $X \sim \text{PR}(a, 0)$ , 则有  $K = \log X \sim E(a, \log \theta)$ . (14) 幂函数分布  $X \sim \text{PF}(c, \theta)$ . 其密度函数为

$f(x) = \frac{c}{\theta} x^{c-1} I_{\theta < x}$  ( $c > 0, \theta > 0$ ).

当  $c=1$  时,  $X \sim \text{U}(\theta, \infty)$ ; 当  $\theta=1$  时,  $f(x) = c x^{c-1} I_{0 < x < 1}$  为  $\text{BE}(c, 1)$ . 另有

$E(X) = \frac{1}{c-1} \quad (c > 1)$ ,  $\text{Var}(X) = \frac{1}{(c-1)^2(c+1)}$  ( $c > 2$ ).

(15) Weibull 分布  $X \sim \text{WF}(A, a)$ ,  $a > 0$ .

若  $Y \sim E(1)$ , 则  $X = Y^a$  服从 Weibull 分布  $\text{WF}(A, a)$ . 由于  $y = x^{1/a}$ ,  $dy/dx = \frac{1}{a} x^{1/a-1}$ ,  $K \sim A e^{-A x^{1/a}}$ ,  $y \geq 0$ , 因此 Weibull 分布的密度函数和分布函数分别为  $f(x; A, a) = \frac{a}{A} x^{a-1} e^{-A x^a} I_{x > 0}$ ,  $F(x) = 1 - e^{-A x^a} I_{x > 0}$ .

## 第一章统计分布基础

其中  $a = A^{1/a}$  为尺度参数,  $A$  为形状参数.  $A=1$  时为标准 Weibull 分布:  $W(1, a) \sim \text{WF}(1, a)$ . 另外, 若  $a=1$ , Weibull 分布化为指数分布  $E(A)$ ; 若  $a=2$ , 称为 Rayleigh 分布:

$f(x, A) = 2Ax e^{-Ax^2} I_{x > 0}$ . (16) 极值分布  $X \sim \text{EV}(A, \alpha)$ .

若  $r \sim \text{IT}(A, \alpha)$ , 贝叶斯  $X = -\log y$  服从极值分布  $\text{EV}(X, \alpha)$ . 由于  $y = e^{-x}$ ,  $dy/dx = -e^{-x}$ ,  $f(y) = \frac{1}{y} \frac{dr}{dy} = \frac{1}{y} \frac{dr}{dy} I_{y < 1}$ , 因此极值分布的密度函数和分布函数分别为

$f(x; A, \alpha) = A e^{-x} | -A e^{-x} |^{-\alpha} I_{x > 0}$ ,  $F(x) = \exp \int -A e^{-x} | -A e^{-x} |^{-\alpha} dx$ .  $A=1, \alpha=1$  时为标准极值分布:  $f(x) = e^{-x} | -e^{-x} |^{-1} I_{x > 0}$ ,  $F(x) = \exp \int -e^{-x} | -e^{-x} |^{-1} dx$ .

(17) 寿命分布与生存函数. 在生物医学统计和可靠性问题中经常要考虑寿命的分布. 设  $f$  为某

生物或产品的寿命 ( $t \geq 0$ ),  $S(t) = P(T > t)$  表示寿命大于  $t$  的概率 ( $0 \leq S(t) \leq 1$ ).

$F(t) = 1 - S(t)$  称为生存函数 (survival function);  $h(t) = -S'(t)/S(t)$  称为危险率函数 (hazard function). 生物统计中经常用  $AG(t)$  和  $SU(t) = F(t)$  代替  $f(t)$  和  $F(t)$  进行分析, 它们有以下

基本性质:

0 表示寿命  $f$  大于  $t$ , 但不超过  $t + \Delta t$  的相对概率, 或危险率:

$h(t) = f(t)/S(t)$ ;  $f(t) = A(0e^{-\int_0^t h(x) dx})$ .

iii) 寿命分布的类型依赖于危险率函数的类型, 常见的类型有: 若  $h(t) = A$  (常数速度),

则  $f \sim$  指数分布 / ( $\infty$ )  $= Ae^{-At}$ ;

若  $A(1) = Aa t^{-\alpha}$  (多项式速度), 则  $f \sim$  Weibull 分布

$f(1) = Aa^{-\alpha}, e^{-Ata}$ ;

若  $h(t) = Aae^{-a}$ , (负指数速度), 则  $f \sim$  极值分布

$f(0) = Aa \exp \{ -A(1 - e^{-at}) - at \}$ . (18) 多元正态分布

设  $\mathbf{Y}$  为  $n$  维随机向量, 若  $\Sigma > 0$ , 则  $\mathbf{Y} \sim N(\mu, \Sigma)$  的密度函数为  $f(\mathbf{y}) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp\{ -\frac{1}{2}(\mathbf{y} - \mu)^T \Sigma^{-1} (\mathbf{y} - \mu) \}$ . (1.3.5)

$\lim_{n \rightarrow \infty} a_n = 0$

$A_i$

ii) 危险率函数  $A(t)$  与密度函数  $f(t)$  有以下一一对应的关系:

若  $t > 0$ , 则  $X$  的特征函数定义为

1.3 常见的连续型分布

21

$p(1) = E(e^{\mathbf{Y}}) = \int e^{\mathbf{Y}} dF(\mathbf{Y}) = e^{\mu + \frac{1}{2}\Sigma}$ . (1.3.6) 熟知  $E(X) = \mu$ ,  $\text{Var}(X) = \Sigma$ , 因此多元正态分布的性质由其前两阶矩

$\Sigma$ ,  $\mu$  完全决定. 以下基本性质大多可由其特征函数推出:

i)  $E(X) = \mu$ ,  $\text{Var}(X) = \Sigma = (\sigma_{ij})$ , 若记  $X = (X_1, \dots, X_n)^T$

$= 0$ ,  $E(X_i X_j) = \sigma_{ij} + \mu_i \mu_j$ . ii)  $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b} \sim N(\mathbf{A}\mu + \mathbf{b}, \mathbf{A}\Sigma\mathbf{A}^T)$  特别有

$\mathbf{Y} \sim N(\mu, \Sigma)$  的充要条件是: 对任意  $n$  维向量  $\mathbf{a}$ ,  $\mathbf{a}^T \mathbf{Y}$  为一元正态.

表1.3. 1列出了常见的连续型分布(可参见方开泰, 许建伦, 1987;

Zacks, 1981 ; 茆诗松等, 1998; Shao, 1998).. 表1.3.1常见的连续型分布

均匀分布

以  $\theta_1, \theta_2$

正态分布

对数正态分布  $LN(\mu, \sigma^2)$

$t$  分布  $t(n)$

密度函数

特征函数 期望

方差 密度函数

特征函数

期望 方差

密度函数

期望 方差

密度函数

$\theta, \sigma^2$

$(\theta, \sigma^2)$

$(\theta, \sigma^2)$

$\exp\{ -\frac{1}{2}at^2 \}$

}

$a^2$

$\exp\{ -\frac{1}{2}(\log x / a)^2 \}$

$\exp\{ -\frac{1}{2}(\log x / a)^2 \}$

$\exp\{ -\frac{1}{2}(\log x / a)^2 \}$

$\exp\{ -\frac{1}{2}(\log x / a)^2 \}$

期望  $\theta(n > 1)$

方差  $n/(n-2)$  ( $n>2$ )

22

第一章统计分布基础

F分布

密度函数

$r\Gamma(n/2)m\Gamma(m/2)/\Gamma((n+m)/2) (1-x)^{n/2-1} x^{m/2-1}$

$F(n,m)$  期望

$m/(m-2)$ , ( $m>2$ )

$2m^2/(n+m-2)^2 (m-4)$

指数分布 密度函数

$f(A)$  带有位置参数的指数分布

$E(A, M)$

厂分布  $r(A, p)$

X 分布  $x^2$  (^)

冷分布  $BE(p, q)$

Laplace分布

Cauchy分布

$C_4(x, r)$

| .

$\geq 0$  ],  $A > 0, p > 0$

方差

( $m > 4$ )

$Ae^{-x} \{x^0 \mid \text{特征函数 } (1 - iz/A) \}$  ,

期望

方差  $A^2$  密度函数  $Ae^{-AU-M}/$

特征函数  $e^{(1 - ir/A)}$  期望  $m+a$  ,

方差  $A^2$

密度函数  $L(P)$  特征函数  $(1-U/A)$

期望 方差

密度函数

特征函数 期望

方差 密度函数

期望 方差

$v/X$   $v/X^2$

1 密度函数 巧  $e^{-1/a}$

$1/2$  产、 $1/2$   $1x^0 \mid r(n/2)2n/2$  1

$(1 - 2i\epsilon) \Gamma_n$

$2n$

$\{P, q\} - \{x^p \sim \} - x) f_{l-7} \mid 0 < x^l \mid$

$P = P+Q$

$pq (p+7)^2 (p+ + i)$

.1I 1

特征函数  $e^{(l+W)-1}$  期望

方差  $la$

密度函数

$1a$

$a^2 + (x - /x)^2$



特征函数  $\exp\{i/z\}$  期望 不存在

方差 不存在

续表

### 1.3常见的连续型分布

23

Pareto分布  $PR(a, e)$

幂函数分布  $PF(\theta, c)$

Weibull 分布  $W(A, a)$

多元正态分布

iv) 设  $X =$

$\sim N(A, \Sigma)$  (1.3.6)式可知,  $X$  的特征函数可表示为

$\phi(t) = \exp\{i t^T A + \frac{1}{2} t^T \Sigma t\}$

若令  $t=0$ , 则有  $\phi(0) = \exp\{i \cdot 0\} = 1$ ,

从而

(X

密度函数  $f(x) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\{-\frac{1}{2}(x-A)^T \Sigma^{-1}(x-A)\}$  期望  $A$  方差  $\Sigma$

⑧  $f(x) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\{-\frac{1}{2}(x-A)^T \Sigma^{-1}(x-A)\}$

方差

密度函数  $f(x) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\{-\frac{1}{2}(x-A)^T \Sigma^{-1}(x-A)\}$

期望

4

密度函数  $f(x) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\{-\frac{1}{2}(x-A)^T \Sigma^{-1}(x-A)\}$

期望  $A$ , 方差  $\Sigma$

方差  $\Sigma = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})^T$  密度函数

(f

$I \sim \text{Exp}(\lambda)$

$(\lambda^{-1})$

;则由

特征函数  $\exp\{i t^T A - \frac{1}{2} t^T \Sigma t\}$

期望 方差

j.

$\frac{1}{2} \Sigma$

■

续表

$\sim N(\mu, \Sigma)$ ; 同理  $x_2 \sim N(\mu_2, \Sigma_2)$  即  $X$  的边缘分布仍然为正态分布. 另外, 由上式可知,

$A(1) = P\{X_1, x_2\} = h_1^T h_2 G_2 e^{-\frac{1}{2} x_2^T \Sigma_2^{-1} x_2}$ . 因此  $X_1, X_2$  独立的充要条件为

$\text{Cov}(X_1, X_2) = 0$ ; 即  $X_1, X_2$  不相关. 特别,  $r = x_1$  与  $Z = M$  独立的充要条件为

$\text{Cov}(r, z) = x^T T = 0$ .

V)  $x$  与  $Z = X_2$  独立,  $X_2$  与  $W = X_1$  独立, 且有

$E(z) = \mu_2, \text{Var}(Z) = \Sigma_2$

$-y^T \Sigma_2^{-1} y + \frac{1}{2} y^T \Sigma_2^{-1} T_1$ .

$\sim N(\mu, \Sigma)$

因此  $X \sim N(\mu, \Sigma)$

### 24 第一章统计分布基础

这是因为  $\text{Cov}(X, Z) = \text{Cov}(x_1, x_2 - s_2 a_1 x_1) = s_2 \Sigma_1^{-1} \Sigma_2 = 0$ , 所以  $X, Z$  独立.

Vi) 条件分布  $X_2 | X_1 = x_1$  为正态, 且有

$$E(X_2|X_1) \quad (1.3.7)$$

$\text{Var}(X_2|X_1) = \text{Var}(Z) = \sigma^2 = \text{Var}(Z)$ . (1.3.8) 并且  $A$  与  $X_2 - E(X_2|X_1)$  独立;  $X_2 - E(X_2|X_1)$  与  $X_1$  独立.

为了介绍退化多元正态分布的一个基本性质, 首先简要介绍一下对称矩阵的谱分解. 设  $A$  为  $n \times n$  阶对称矩阵, 其秩为  $r$ ,  $\text{rk}(A) = r$ . 则必存在正交矩阵  $U$ , 使得  $U^T A U = \text{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0) = C$ , 其中  $\lambda_i$  为  $A$  的非零特征根, 其相应的特征向量记为  $u_i = (u_{i1}, \dots, u_{in})^T$ , 并记  $r = (r_1, r_2)$ . 则  $A$  的谱分解可表示为

$$A = U C U^T = \sum_{i=1}^r \lambda_i u_i u_i^T \quad (1.3.9)$$

定理 1.3.2 若  $Y \sim N(p, A)$  为退化正态分布, 即  $\text{rk}(A) = r < n$ , 并设  $A$  的谱分解如 (1.3.9)

所示. 则  $Y$  可表示为  $Y = \mu + r_1 z_1 + r_2 z_2 + \dots + r_r z_r$ , 其中  $z_i$  服从非退化正态分布  $N(0, I)$ ,

$A = \begin{pmatrix} r_1 & 0 & \dots & 0 \\ 0 & r_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_r \end{pmatrix}$ ,  $B = r_l A^{-1}$ .

证明 做变换:  $r(Y - \mu) = Z = (z_1, \dots, z_n)^T$ , 则有  $(r - M)^T r z^T r x z x$

$$+ r_2 z_2^2, Y = \mu + r_1 z_1 + r_2 z_2 + \dots + r_r z_r$$

. 而由 (1.3.9) 式可知,  $z$  的分布为  $N(0, I)$ ,  $r = \text{tr}(A)$ ,

因此  $Z \sim N(0, I)$ ,  $z$  服从退化分布  $p(z_2 = 0) = 1$ . 而  $r \cdot Z = r_l A^{-1} A^{-1} Z = B W$ , 其中  $B = I$ ,

$$I, W = A^{-1} Z \sim N(0, I).$$

#### 1.4 一元非中心 $\chi^2$ 分布及其有关分布

本节主要介绍非中心  $\chi^2$  分布, 它是中心  $\chi^2$  分布的 Poisson 加权和. 由非中心  $\chi^2$  分布可进一步得到非中心  $t$  分布、非中心  $F$  分布以及非中心  $f$  分布.

##### 1.4.1 非中心 $\chi^2$ 分布和非中心 $\chi^2$ 分布

定义 1.4.1 非中心  $\chi^2$  分布  $X \sim \chi^2_r(A, p; 5)$  的密度函数定义为  $y(x; A, F, 6) \triangleq y(\lambda; A, \lambda^2 + 2\lambda p, 6/2) 4u(j)$

#### 1.4 一元非中心 $\chi^2$ 分布及其有关分布

25

$$= y(x; A, \lambda^2 + 2\lambda p) e^{-\lambda^2/2} \quad (1.4.1) \quad y \gg 0 \quad 7^*$$

其中  $r(x; A, p; j)$  为中心  $\chi^2$  分布  $\chi^2_r(A, r; j)$  的密度函数,  $p(j, 5/2)$  为泊松分布  $J \sim P(8/2)$  的离散密度; 如  $(\cdot)$  为计数测度. 特别,  $r(\lambda/2, n/2; 5)$  称为非中心  $\chi^2$  分布, 记为  $\chi^2_r(n, 5)$ , 即  $\chi^2_r(n, a) = r(\lambda/2, n/2; 5)$ .

(1.4.1) 式表明, 非中心  $\chi^2$  分布是中心  $\chi^2$  分布的 Poisson 加权和. 在  $X \sim \chi^2_r(A, p; 5)$  中,  $S$  称为非中心参数. 由定义 (1.4.1) 式可知, 若  $5=0$ , 则非

中心  $\chi^2$  分布  $x \sim \chi^2_r(x \text{ 化为中心 } \chi^2 \text{ 分布, 即 } r(x, \lambda^2; 0) = r(x, p)$ . 非中心  $\chi^2$  分布和非中心  $\chi^2$  分布有以下基本性质:

$=1$ , 即  $y(x; A, 5)$  为密度函数. 证明由定义 (1.4.1) 式可知,

$$y(\lambda; A, M) \text{ 如 } (X) = \int_0^\infty y(\lambda; A, r+y) p(y, 3/2) d/i(j) 4t(x)$$

$$A - A^* (g/2); \quad 777 \text{-----} \quad e^{-x} dx$$

)

其中  $c$  (幻与  $J$  无关,  $a$ , . 巧 易证:  $\lambda \rightarrow 0$  ( $j \rightarrow 0$ ), 因此  $\lambda + j = I$ . )  $i j l a j$

以上幂级数绝对收敛, 可逐项积分:

$$\int_0^\infty y(x; A, \lambda^2 + 2\lambda p) d(\lambda) = \int_0^\infty y(x; A, p + J) p(j, 5/2) dx$$

$$= \int_0^\infty e^{-\lambda^2/2} d\lambda$$

$= \int_0^\infty e^{-\lambda^2/2} d\lambda$  在  $X \sim \chi^2_r(A, p; 5)$  中,  $X$  可视为联合分布  $(X, J)$  中  $X$  的边缘分布, 其中

$J \sim P(5/2)$ ,  $X|J \sim \chi^2_r(A, p + J)$ . (1.4.2) 对于  $\chi^2_r(n, 5)$ ,  $I$  可视为  $(X, J)$  中  $J$  的边缘分布, 其中

$J \sim P(8/2)$ ,  $X|J \sim \chi^2_r(n+2J)$ . (1.4.3) 证明 记  $g(x, j) = y(x; A, r+j) P(j; 5/2)$ , 则由

性质i)可知

$I_0 = 1$ ,

因此  $(X|J) \sim g(x, j)$ . 由假设  $J \sim P(8/2) = p(j, 3/2)$ , 因此  $X|J \sim g(x, j)/r_{00}$

$= I_2^{\lambda} = 0$

$(i/ + j a) \times J dx$ ,

$= I$

## 26 第一章统计分布基础

$\gamma(5/2) = \Gamma(x; A, p+)$ ; 由此可得(1.4.2)式. 又由定义(1.4.1)式可知  $\gamma(x; A, ^, 5)$ .

对于  $X \sim \gamma$ , 则有

$X|J \sim r(l/2, n/2 + J) = r(l/2, (n + 2J)/2) = \gamma_2(l+2, n)$ . iii) 非中心  $J_n$  分布和非中心  $\chi^2$  分布的期望和方差可表示为

$E(X) = \lambda, \text{Var}(X) = 2\lambda$

$E[X^2(n, 5)] = n+5, \text{Var}[X^2(n, 5)] = 2(n+5)$ .

证明 以下仅对非中心  $\chi^2$  分布加以证明. 由  $y|j \sim r(A, r+j)$ ,

从而应用(1.1.4)式和(1.1.5)式有  $E(X) = E_y[E(X|J)] = E[\lambda]$

$\text{Var}(\lambda) = E[\text{Var}(X|J)] + \text{Var}[\lambda] = E[\lambda] + \text{Var}[\lambda]$

$\lambda + 8/2 = \lambda + 4 = \lambda + 4$

$A^2 = 2\lambda$ . iv) 非中心  $\chi^2$  分布和非中心  $\chi^2$  分布的特征函数分别为

$\phi_{JP}(t) = (1 - \lambda t^2)^{-1/2}$

$\phi_{\gamma}(t) = (1 - 2\lambda t^2)^{-1/2} \exp(-\lambda t^2)$

证明 根据  $x|j \sim r(A, p + j)$  以及中心  $\chi^2$  分布的特征函数可得

$E[\exp(itX)] = E[E[\exp(itX)|J]]$

$= E[\exp(-\lambda t^2/2)]$

$= \int_0^\infty \exp(-\lambda t^2/2) p(j) dj$

$= \int_0^\infty \exp(-\lambda t^2/2) \frac{1}{\Gamma(\lambda/2)} \lambda^{\lambda/2} j^{\lambda/2-1} \exp(-\lambda j) dj$

$= \Gamma(\lambda/2) \int_0^\infty \exp(-\lambda t^2/2) \lambda^{\lambda/2} j^{\lambda/2-1} \exp(-\lambda j) dj$

由此即可得到以上第一式, 并可推出第二式.

v) 非中心  $\chi^2$  分布和非中心  $\chi^2$  分布有可加性, 即若  $X \sim \gamma(A, \lambda; ^)$

且独立, 则有  $X + Y \sim \gamma(A + B, \lambda + \mu; ^)$

vi) 若  $X \sim \gamma(M, 1)$ , 则有  $X^2 \sim W, (r=2)$  且独立  $(i=1, \dots, n)$ , 则有  $2X^2/a^2 \sim \chi^2(M)$ ,  $S = \dots$

$\chi^2(8/2)$

$\exp(-\lambda t^2/2)$

$A = \lambda$

$(1, 5), 8 = M^2$

. 因此, 若  $X \sim \gamma$

## 1.4 一元非中心 $\chi^2$ 分布及其有关分布

27

vii) 若  $X \sim N(p, a^2 I_n)$ , 则  $X^T X / a^2 \sim f_2(n, 5)$ ,  $8 = J/x/cr^2$ ; 若  $I \sim V(m, 2)$ , 则  $(X-p)^T (X-p) \sim \chi^2(n, m - x(n, 8))$ ,

$8 = //wV$

viii) 若  $X \sim V(M, a^2 Z)$ , 则  $Y = X^T A X / a^2 \sim \chi^2$  分布的充要条件为

为投影阵 (即  $A^T = A, A^2 = A$ ) 且有  $y \sim \chi^2(r, 5)$ , 其中  $r = \text{rk}(A)$  为  $A$  的秩,  $8 = A/m/a^2$ .

以上性质 vi) - viii) 详见张尧庭、方开泰 (1982).

注 非中心  $\chi^2$  分布的定义有几种, 但最后都殊途同归于相同的特征函数, 因而各种定义相互

等价.非中心/分布也可以从性质vi)出发 加以定义,最后特征函数都一样.

1.4.2 非中心F 分布和非中心/ 分布 由前节定义的非中心; $V_2$ 分布可导出非中心F分布和非中心t分布.

若 $\sim$ 且相互独立,则义服从

非中心F分布 $F(n_1, n_2; \delta)$ .即 $F(\sim, n_2; \delta) / \chi^2_1$

$X \sim \chi^2_1 / n_2$

本性质:

i)若 $X \sim F(n_1, n_2; \delta)$ , 则I 可视为 $Y=(X|J)$ 中X的边缘分布, 其

中  $J \sim P(S/2)$  且  $X|j \sim F(n_1 - 4 - 2j, n_2) n_1 + 2j$ , 即

$n_1$ ,

$+2j \sim F(n_1 + 2j, n_2)$ .

证明 若 $\gamma \sim /(\wedge)$ , 则 $\gamma$ 可视为 $r, =(\wedge, /)$ 中 $\gamma$ 的边缘分 布.其中,  $\gamma \sim P(S/2)$ ,  $X, | /$

$12(\wedge + 2j)$ , 因此

】 $X|j / (n_1 + 2j) \sim n, +2j \sim \chi^2_1$

$a + 2j \sim \chi^2_1 / n_2$ -----.

$n_1$

ii)若 $X \sim F(n_1, n_2; \delta)$ , 则有 $E(X) = \sim$  (与 $n_1, n_2, \delta$ 都有关).  $7ii \sim n_2 - 2$

证明由以上性质i)可得

$E(X) = E; [E(y | J)]$

$= E \int_0^1 F(n_1 + 2j, n_2) \delta^j j^{j-1} dj$

$1 \int_0^1 x / n_1 - 1 \sim X^2 / n_2$

$\sim F(n$

$Y$

; 它有以下基

28

第一章统计分布基础

$= E(\sim \sim + 2, 71 \sim n_2 - 2 \sim H)$

$n_2 + 5 \sim n_2 - 2 \sim T_{ij}$

iii)  $F(n_1, n_2; \delta)$ 的分布函数和分布密度分别为

$F(x; n_1, n_2, \delta) \sim \delta^{5/2} e$

$T|x \sim \sim; n_j + 2j, n_2$ ,

$, n_2 \sim 8) = y; = 0$

证明仍用性质i),

;!

+2.

$= j) \sim j)$

$F(x; n_1, n_2, \delta) = P(X \leq x) = \int_0^1 j = 0$

$5 \sim P(J = j) P(X \leq x$

$r \sim i$

$= s \sim (/ = j) \sim n_j X \sim x \sim H$

$- 5/2 (5/2) y \sim x$

$= 2 e \sim 7 T \sim f \sim + 2 \sim w +$

此即第一式;对 $F(x; n_1, n_2, \delta)$ 求导可得到第二式. 以下简单介绍非中心t分布.若 $Z \sim \chi(n)$ 且相互独立,

则  $T = r / \sqrt{Z^2 / \delta}$  服从非中心\* 分布  $\ll(n, \delta)$ ,  $\delta = M^2$ . 因此, 若  $r = (X + a) /$

$X \sim \chi(n)$  且与 $Z$ 独立,  $a$ 为常数, 则 $T$ 服从非中心 $z$ 分布. 另外•若 $\delta = 0$ , 则 $f(n, 0)$ 为中心 $Z$

分布 $t(n)$ ; 并且有 $T^2 \sim F(1, n; \delta)$ ,  $\delta = /$ . 进一步讨论见方开泰、许建伦(1987).

例 1.5.6 指数分布 $E(X)$ .

$f(x, \theta) = A e^{-A V(x) + \eta(\theta) T(x)}$ , 其中  $T(x) = x$ ,  $Q(\theta) = -A$ ,  $h(x) = I(x)$ ,  $b(\theta) = -\log A$ . 但是, 带有位置参数的指数分布  $E(\lambda, \tau)$ ,  $E(1, m)$  不是指数族分布.

例如  $A, \tau$ ,  $f(x, \theta) = A e^{-\lambda(x - \tau)} = \exp[-A\theta + A/z + \log A] \cdot \tau^{-x}$ , 其中  $\lambda$  与  $\theta$ ,  $m$  都有关, 因而不符合定义 (1.5.1) 式的形式.

以下介绍一个判别指数族分布的一个简单的必要条件.

定义 1.5.2 分布族的支撑集 (support). 若  $X \sim f(x, \theta)$ ,  $\theta \in \Theta$ , 则其支撑集定义为  $S = \{x : f(x, \theta) > 0\}$ . 若  $S$  与  $\theta$  无关, 则称该分布族有共同的支撑集.

易见, 分布族的支撑集可能与参数  $\theta$  有关, 如均匀分布  $X \sim U(0, \theta)$ :  $S = [0, \theta]$ ; 带有位置参数的指数分布  $X \sim E(x, \theta)$ :  $S = [\theta, \infty)$  等. 但是由定义 (1.5.1) 式可知, 指数族分布有共同的支撑集  $S = \mathbb{R}$ .

$U: f(x, \theta) > 0 \mid x \in S$ , 与参数  $\theta$  无关. 因此可得以下重要推论:

推论一个分布族为指数族的必要条件是它有共同的支撑集.

由此可知, 均匀分布  $U(0, \theta)$ ; 带有位置参数的指数分布  $E(x, \theta)$ ,  $E(1, m)$  都不可能是指数族分布. 同理, 双参数的 Pareto 分布以及幂函数分布都不是指数族分布. 值得注意的是: 今后我们得到的许多有关指数族分布的结果, 对于均匀分布以及带有位置参数的指数分布等这些分布就不一定成立.

小结: 常见的指数族分布有: 正态分布、二项分布、多项分布、Poisson 分布、单参数的负二项分布和 Pascal 分布、几何分布、对数正态分布、Rayleigh 分布、逆 Gauss 分布等等.

常见的非指数族分布有: 均匀分布、带有位置参数的指数分布  $E(x, \theta)$ ,  $f(1, \tau/z)$ 、双参数的 Pareto 分布和幂函数分布、极值分布、Weibull 分布、超几何分布、Cauchy 分布、Laplace 分布等等.

### 1.5.2 指数族的自然形式

在指数族的定义 (1.5.1) 式中, 很自然地可以考虑把参数函数  $\eta(\theta)$  的各分量直接看成参数来确定分布密度. 这时有  $f(x, \eta) =$

$A(x) \exp\{\eta(\theta) T(x) - b(\eta)\}$ . 其中  $A(x)$  分布密度的主体就是  $\eta$  与  $T(x)$  的线性组合, 因而应当有良好的性质.

#### 1.5 指数族分布

31

其正式定义如下: 定义 1.5.3 自然形式的指数族. 若  $X \sim f(x, \theta)$  其密度函数可表示为

$f(x, \theta) = h(x) e^{\eta(\theta) T(x) - b(\eta)}$ , (1.5.2) 其中  $\theta = (\theta_1, \dots, \theta_k)$ ,  $T(x) = (T_1(x), \dots, T_k(x))$ , 则称其为自然形式的指数族; 若  $T_1(x), \dots, T_k(x)$  线性无关, 则称指数族为极小、满秩的; 又称

$\eta = (\eta_1, \dots, \eta_k)$  为自然参数空间. (1.5.3)

为分布族的自然参数空间. 根据以上定义, 由于  $J(\eta) \neq 0$ , 因此  $\eta = 0$  等价于  $I$

$e_{\eta} = J(\eta) e^{\eta T(x) - b(\eta)} > 0$  或  $b(\eta) = \log [J(\eta) e^{\eta T(x)} / z(\eta)] < +\infty$ . (1.5.4)

另外, 在定义 (1.5.2) 式中, 若令

数族

族的分布密度的主体部分就是参数  $\theta$  与统计量  $T(x)$  的线性组合, 因而可以预期, 它应当有良好的性质.

好的性质.

定理1.5.1自然参数空间 $\theta$ 必然是 $R'$ 上的凸集, 势函数 $b(\theta)$ 为  $\theta$  上的严凸函数.

证明为证 $\theta$ 为凸集, 对 $\lambda \in (0, 1)$ ,  $\theta_1, \theta_2 \in \theta$ , 要证 $M(1) + (1 - A) \theta_2 \in \theta$ , 良P

$a = J/t(x) \exp \int [Aa(1) + (1 - A) T] T, (x) j d^t(\%) < + \infty$ . 而 $a$ 可表示为根据Holder不等式以及(1.5.3)式有

$a \int j A(x) e^{r(1)} \text{如}(x) f[p^e A] r(1) \text{如}(x) f_A < + \infty$ .

要证 $\theta$ 为 $\theta$ 上的凸函数, 即要证 $b(\lambda \theta_1 + (1 - A) \theta_2)$

:  $Tf(\%) = y$ , , 则可得到更简单的指

$Y - f(y, e)$  由以上定义以及表达式(1.5.2)和(1.5. 5)可知, 自然形式的指数

$A \theta_1 + (1 - A) \theta_2 \in \theta$ . 由(1.5.4)可知  $\theta_1 + (1 - A) \theta_2 \in \theta$

■(4)如(x)}

(1.5.5)

32 第一章统计分布基础

$= \log \{ j u(x) e^{r(4)} A U(\%) e^{2 \theta} \mid X x \}$ .

根据Holder不等式, 上式

$\theta_1 + (1 - A) \theta_2 \in \theta$

$\int \log \{ [p_i(x) e^{r(*)} d^e, (A:)] [J/i(x) e^{2r(x)} d/x(\%)] \}$

$= A \log [J/i(\%) e^{r(t)} d f j L(x) j + (1 - A) \log [J/i(x) e^{2r(x)} d/z, (\%) j]$

$= A \theta_1 + (1 - A) \theta_2$ . I 定理 1.5.2 若  $g(x)$ 在 $x$  上可测, 且  $G(a) =$

$\int g(x) e^{x} d^e(x)$ 存

在, 则 $C$ 在 $\theta$  内部解析. 特别,  $b(\theta)$ 在 $\theta$ 内部解析.

(这个定理实际上就是Laplace变换的解析性质, 其证明可参见陈希 孺, 1981.)

自然形式的指数族的其他性质:

i) 若 $X - f(x, \theta) = h(x) e^{f(\theta) T(x) - b(\theta)}$  则其特征函数由势函数 $b(\theta)$

确定:

$\phi(t) = \exp [b(\theta + i t) - b(\theta)]$ . (1. 5. 6)  $\phi(t) = E [j =$

$e^{i t X} h(x) e^{f(\theta) T(x) - b(\theta)} d^e(x) = \int h(x) e^{(3 + i t) T(x) - b(\theta)} d^e(x)$

ii) 若  $X = h(x) e^{f(\theta) T(x) - b(\theta)}$  则  $T = T(X)$  亦服从指数族 分布:

$T = T(X) - f(t, e) = V(1) \theta_b(4)$  (1.5.7) 证明  $T = T(X)$ 的特征函数为

$\phi(s) = E [e^{b T(x)}] = \int e^{b T(x)} e^{f(\theta) T(x) - b(\theta)} d^e(x)$

$e^{(b + i j) T(x) - b(\theta) + b(e + i i) - b(e) r} = e^{b(\theta + i j) - b(e)}$

证明

由(1.5f.6)式及唯一性可知,  $T = T(X)$ 服从指数族分布, 且有

$E[r(x)] = db(\theta)/d\theta$ ,  $E[r(x)] = A(4)$  (1.5.8)  $V - [刺相 = (L, (1.5.9)$

$h^*(1)$

iii) 若  $X - f(X, e) = l(x) e^{x} - b(t, f)$ , 则有

(4). |

1.5指数族分布

33

$E = \theta(10)$  職. (1.5.10)

证明 由 $\theta(x) \theta_b(4)$ 如(x) = 1, 该式对 $\theta$ 求导可得  $\theta(x) - b(\theta) =$

$\theta_b(x) = 0$ ;

因而得到 $E[T, .(X)] = db(\theta)/d\theta$ ; 继续再求导, 可得 $\text{Var}[T(X)]$ 以及

义的各阶中心矩. 1 由以上结果可知,  $r(x)$ 的各阶中心矩均由势函数 $b(\theta)$ 决定.

(1.5.11)

或

, 幻 “ (y)e<sup>0T</sup>

推论若Y★

E(y)=m = A(6>), Var(y)=b(0).

### 1.5.3 带有多余参数的指数族

”

b(0) , 则有

指数族有若干发展与推广, 以下介绍两种常用的带有多余参数的指数族. (1)带有尺度参数的指数族.

其密度函数可表示为 -

F ~A ”

沒, cr) =exp[

a

( 1. 5. 12)

j\y\Q, (r) =exp{(/)0Ty-6(沒)-c(y,(/))}],小=a\_2,

其中y, 0同为n维向量;0为有兴趣的参数, a或</>常常视为多余参 数, a 是尺度参数.该分布族常记为Y~ED<sup>a</sup>2).

例1.5.7正态分布YKpc'a2), 其分布密度可表示为

识-去/ -\*^-log(27ra2)

= exp ----- . - CT

若取6X =fJL/a\ T. =r, e2 =a-2, T2 = -//2 ;则为前节介绍的一般的 指数分布族.但若把a2视为多余参数,取仏=M, =r; b(6) =M2/2

=W/2; c(y,a) =y/+ylog(211a2);则为(1.5.12)式所定义的带有

尺度参数的指数族. ■ 例1.5.8 r分布r~r(A,p),其中m=e(k)=z, /a, a= 其

密度函数为

该式可改写为

FV(W) +, y>0.

### 34 第一章统计分布基础

) = exp | p[y( -/z-1) - log /z, + log y - p"1 log y + log p - p \_ 1  
log r ( p .

若把p = </> =a-2视为多余参数, 对比(1.5. 12)式, 可取0. =

Ji=y; 6(0)=log/f=-log(-沒!); c(y,a)=-logy+r1logy-logv ^^^logr (p).因此尸分布为由(1.5.12)式所定义的带有尺度参数的

指数族.

例1.5.9 逆Gauss分布Y~IG(fi, a2).其密度函数为

该式可改写为

a h 1 (r -m)2i y2ira2/ 2<7M 7 J

-j^\_2/2 +^\_, -y\_1/2 -^-log(27r(r2/)) =exp

若把<r2视为多余参数, 对比(1.5. 12)式, 取 - (2/i2) -1, =y;

6(0)=-烱=-(-20,)T; c(y,cr)=(2y)\_1+^-log(2irtr2/).则逆

Gauss分布为带有尺度参数的指数族. 带有尺度参数的指数族Y~ED(0, a2)有以下性质:

i)特征函数为^(0

er-2, 特别有

E( Y) = b(0) =//,, Var(K) =cr26(0) =(r2V; ( 1. 5. 13)

对 r 分布:Var( Y) = a2f±2; 对逆 Gauss 分布:Var( Y) = a2 (对 Pois- son 分布:Var(y) =/!).特-征函数的推导与(1.5.6)式完全类似, 从略.



ii) 令  $e = y - z = (C_j,$   
 $E(e_j) = 0; E(n) = r^2 K_y, V = (K_{..}) = 6(0),$   
 $E(\cdot) = a_4 S_{ijk}, E(e, e_{tez}) = tr_4(K_{..} + V_{jt} + V_{kj}) + a_6 \sim u,$   
 $s = W(没)A - \wedge(8)$

vfc deidojdok ' iiU —de.d0jde.d0i  
 iii) 若  $Y_x, y_{Yn}$  为 i.i.d.,  $Y \sim ED(\theta, (r^2))$ , 则  $Y = n^{11}$   
 $f_7?(\wedge, (r^2/n)). (2)$  子集参数情形.  
 $K_{..} \sim$

对于指数族分布 (1.5.2), 有时需要把参数分为两部分: 一部分为 有兴趣参数, 另一部分为  
 多余参数. 这时分布密度可表示为

则可求出其特征函数, 从而有  
 $i = l$

1.5 指数族分布  
 $X = f(x; \theta, (p) = h(x) \exp,$   
 $m$   
 35  
 (1.5. 14)

$0 i U^x) + \wedge^{\cdot}(x) = b(\theta, (p)$   
 其中  $\wedge = (, 6 > m) T$  为有兴趣参数,  $(P = (< P_i, -$  为多余参数,  $" = (U, (x), \dots, \zeta(\%)) T,$   
 $r = (T, (x), \dots, 7; (\%)) t.$

定理 1.5.3 在以上假设下,  $(f/, r)$  的联合分布为  $m$   
 $p(u, t; \theta, < p) h u, /) \exp.$   
 $i = 1$

"和  $r$  的边缘分布分别为  
 $p(u; \theta, (P) h l f i(u) \exp$   
 $p(t \setminus 6, < p) = \sim(1) \exp\{$   
 $+2 - \wedge(没, 妒)',$

$y. 0$  人  
 条件分布  $U \setminus T$  仅与有兴趣参数  $\theta$  有关, 与多余参数无关, 其分布为  
 $P(u | t; \theta, (p) = h * (w, \theta \exp J Z \theta i U i - b : (4) i. (1.5.15) f \ll 1$

证明  $\{T, U\}$  的联合分布可由前面性质 ii), 即 (1.5.7) 式得到; 联合分布积分即得边缘分布:  
 由此可得条件分布为

$P(\wedge \theta; (p) = \int p(u \setminus 7 t; 6, (p) du$   
 姑  $= 6(\theta, 妒) \bullet,$   
 少  $(u, Z) \exp.$   
 $\exp, ; = 1 - b(\theta, W$   
 $= \wedge X 0 \exp J \int (p. 1. - b(0, (p) -$   
 $m$   
 $= h * (u, 0 \exp\{ 2 \theta i u i - b * (6 >)\} i = 1$   
 $(V(没) \log h \theta(t)) \bullet$   
 $i = 1$   
 $\wedge. u. I d z z \bullet$   
 $b(0, < p)$   
 36

## 第一章 统计分布基础

### 1.6 次序统计量的分布

次序统计量是统计学的基本统计量之一, 在参数统计与非参数统计 中都有广泛的应用. 本节

首先介绍次序统计量的基本性质，然后分别介绍均匀分布和指数分布的次序统计量。

给定样本 $X_1, \dots, X_n$ ，若按大小重新排序，其次序统计量定义为

$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ 或记为 $y_1 \leq y_2 \leq \dots \leq y_n$  ( $r = y_{(r)}$ )。特别， $X_{(1)} = \min_{1 \leq j \leq n} X_j$ ， $X_{(n)} = \max_{1 \leq j \leq n} X_j$ ， $R = X_{(n)} - X_{(1)}$ 称为极差。通常假定 $X_1, \dots, X_n$ 独立同分布，且 $X \sim f(x)$ 为绝对连续型分布，因此有 $P(X_{(i)} <$

$y) = 1 - (1 - F(y))^n$ ， $P(X_{(i)} > y) = (1 - F(y))^i$ ， $P(X_{(i)} = y) = 0$ 。

记 $(r, \dots, n) = (X_{(r)}, \dots, X_{(n)})$ ， $(r, \dots, n) = (X_{(r)}, \dots, X_{(n)})$ ，

(i)  $r, \dots, n$ 的分布密度为

(1.6.1) 其分布函数可由不完全 $\beta$ 函数表示为 $I_x(r, n-i+1)$

证明 考虑 $Y_i = X_{(i)}$ 落在微小区间 $(y, y + dy]$ 中的概率，这可表示为

$P\{y < Y_i \leq y + dy\} = F(y + dy) - F(y) = f(y)dy + o(dy)$ ；另一方面，由 $X_{(1)}$ 的定义可得： $P\{r < Y_i \leq y + dy\} = P\{X_{(1)} \leq X_{(n)}$ 中有 $(i-1)$ 个在 $(-\infty, y)$ ，有一个在 $(y, y + dy]$ ；有 $(n-i)$ 个在 $(y + dy, +\infty)$ ；

+  $o(dy)$  | +  $P$ ！其他情形 $I = P_1 + P_2$ ，其中 $P_1, P_2$ 分别表示和式中的第一、

第二项。 $j$ 其他情形)包括诸如 $j$ 在 $(-\infty, y)$ ， $\dots$ ，及中有 $(i-2)$ 个在 $(-\infty, y]$ ；有2个在 $(y, y + dy]$ ；有 $(n-i)$ 个在 $(y + dy, +\infty)$  | 等情形。显然，对

连续型分布有 $P_2 = o(dy)$ ；因 $f(y) = f(y) + o(dy)$ 。而 $f(y)$ 可由多项分布得到：

$f(y) = \frac{n!}{(i-1)!(n-i)!} [F(y)]^{i-1} [1-F(y)]^{n-i}$ 。

(7)  $= 0$ ， $P(X_{(m)} < X_{(2)})$  导次序统计量的基本分布。

1.6.1 基本分布

$X_{(n)} = 1$ 。下面将在此假定下推

比较以上结果，并令 $dy \rightarrow 0$ 可得

1.6次序统计量的分布

37

由 $f_{(i)}(r) = \frac{n!}{(i-1)!(n-i)!} [F(y)]^{i-1} [1-F(y)]^{n-i}$ 出即得  $f_{(i)}(r) = f(y) \frac{n!}{(i-1)!(n-i)!} [F(y)]^{i-1} [1-F(y)]^{n-i}$ 。

$[F(z) - F(y)] y^{n-1} [1-F(z)]^{n-i} | y < z$ 。(1.6.5) 证明 仍然沿用上面的方法。考虑

$(y, y + dy]$ ， $X(y)$ 落在微小矩形 $\{(y, y + dy], (z, z + dz]\}$ 中的概率，这可表示为

$P\{y < X_{(i)} \leq y + dy, z < X_{(j)} \leq z + dz\} = \frac{n!}{(i-1)!(j-1)!(n-i-j)!} [F(y)]^{i-1} [F(z) - F(y)]^{j-1} [1-F(z)]^{n-i-j} dy dz + o(p)$

推论1 最小值和最大值的分布密度函数分别为  $f_{(1)}(y) = n[F(y)]^{n-1}$ ， $f_{(n)}(y) = n[1-F(y)]^{n-1}$ 。(1.6.2)

$f_{(i)}(y) = \frac{n!}{(i-1)!(n-i)!} [F(y)]^{i-1} [1-F(y)]^{n-i}$ 。(1.6.3) 推论2 若 $X \sim U(0, 1)$ ，则 $X_{(1)}$ 服从 $U(0, 1)$ 分布：

$X_{(i)} \sim BE(i, n-i+1)$ 。

(2)  $(X_{(j)}, X_{(i)}) = (y, z) (i < j)$ 的联合分布为  $f_{(i,j)}(y, z) = \frac{n!}{(i-1)!(j-1)!(n-i-j)!} [F(y)]^{i-1} [F(z) - F(y)]^{j-1} [1-F(z)]^{n-i-j}$ 。

(1.6.4)

。

$f_{(i,j)}(y, z) = \frac{n!}{(i-1)!(j-1)!(n-i-j)!} [F(y)]^{i-1} [F(z) - F(y)]^{j-1} [1-F(z)]^{n-i-j}$ 。

$(p = y dy + z dz)$ 。另一方面，以上概率也可表示为 $P$ 中有 $(i-1)$ 个在 $(-\infty, y]$ ；有1个在 $(y, y + dy]$ ；有 $(j-i-1)$ 个在 $(y + dy, z]$ ；有1个在 $(z, z + dz]$ ；有 $(n-j)$ 个在 $(z + dz, +\infty)$  | +  $p$ ！其他情形 |。这一概率可由

多项分布得到：

$D = \frac{n!}{(i-1)!(j-1)!(n-i-j)!} [F(y)]^{i-1} [F(z) - F(y)]^{j-1} [1-F(z)]^{n-i-j}$ 。

$[F(z) - F(y)]^{j-1} [1-F(z)]^{n-i-j} dy dz$ 。比较以上结果，并令 $p \rightarrow 0$ ，即可得到(1.6.5)式。

1

推论1 又(1)和 $t(4)$ 的联合分布密度函数为

$=n(n-1)/(y)/(z)[F(z)-F(1)]^2/\sqrt{y} < z$  |, (1.6.6) 推论2 极差 $R=X(n)-X(1)$ 的分布为

$A(r) = n(n-1) \int_0^1 \int_0^{1-r} [F(t+r) - F(t)]^2 dz \quad (r > 0).$

(1.6.7) 证明由于 $(y, z)$ 的分布已知, 可做变换: $z = z(n) -$

$X(1) = Z - y, T = X(1) = y$ , 则 $Z = r + r$ ,  $Y = T$  变换的Jacobi行列式为 $J = 1$ . 由(1.6.6)式可得

$/0, r) = n(n-1)/(0/(f+r)[F(f+r) - F(t)]^2 \quad (r > 0)$ . 上式对 $t$ 积分, 即得 (1.6.7)式.

38

## 第一章统计分布基础

(3)前 $r$ 个次序统计量 $JV(1), \dots, Y(7)$ 的联合分布为(其推导与(1)、(2)类似, 从略).

$= (1/r!) j/(y_i) \cdot V(y_j) \prod_{j=1}^r (y_j - y_{j-1}) \quad (1.6.8)$

特别,  $x(1), \dots, X(4)$ 的联合分布为

$7 \cdot (1, \dots, y_n) \quad (j_1 < y_2 < \dots < y_n \leq 1) \quad (1.6.9)$

### 1.6.2均匀分布的次序统计量

若 $X \sim F(x)$ 为连续型分布, 则 $U = F(X) \sim \text{尺}(0, 1)$ . 因此若 $X(1) < \dots < X(n)$ , 则

$U(1) = F(X(1)) < \dots < F(X(n)) = U(n)$ 为均匀分布 $\text{尺}(0, 1)$ 的次序统计量.

因此, 考虑 $U(0, 1)$ 的次序统计量, 对一般分布 $F$ 均亦有一定意义. 由(1.6.9)式可知, 若 $U \sim \text{尺}(0, 1)$ , 17(1)名 $\dots C(t/1)$ , 则 $(U(1), \dots,$

$U$  的联合分布为

$(U(1), \dots, U(n)) = n! \int_0^1 \int_0^{1-u_1} \dots \int_0^{1-u_{n-1}} \quad (1.6.10)$

定理1.6.1 均匀分布 $U(0, 1)$ 的次序统计量有以下性质: (1)  $U(1) \sim \text{BE}(k, n-k+1)$ ;

(2)  $U(k) \sim \text{BE}(k, n-k+1)$ 与 $Z$ 无关; (3)  $R = U(n) - U(1) \sim \text{BE}(n-1, 2)$ .

(1.6.10)

(1.6.11)

证明 性质(1)即为(1.6.4)式, 性质(3)可由(2)导出, 下面主要证明性质(2). 为此, 做变换:

$U(1),$

$U(2) = U(1) + Z_1, \quad Z_1 = U(2) - U(1),$

$U(1) = 0$

$U(2) = U(1) + Z_1,$

$U(3) = U(2) + Z_2 + Z_3,$

$\Rightarrow J = 1.$

(心  $= U(1) + U(2) + \dots + U(n)$  其中 $J$ 为变换的Jacobi行列式; 由(1.6.10)式,  $i/$

(1),  $\dots$ , (4)的联合

分布可表示为

$= n! \int_0^1 \int_0^{1-u_1} \dots \int_0^{1-u_{n-1}} \quad (1.6.11)$

$= n! \int_0^1 \int_0^{1-u_1} \dots \int_0^{1-u_{n-1}} \quad (1.6.11)$

$\cdot A_2(u_1, \dots, u_n) = \int_0^1 \int_0^{1-u_1} \dots \int_0^{1-u_{n-1}} \quad (1.6.11)$

$< 1$  }.

$= n! \int_0^1 \int_0^{1-u_1} \dots \int_0^{1-u_{n-1}} \quad (1.6.11)$

因而对任意 $A$ ; 个分量的积分

$i=1$

$z_i < 1$ .

### 1.6 次序统计量的分布

39

值相同，得到的边缘分布也相同。由对称性可推出：

i) 一切  $X_i$  同分布，且与  $(1)$  同分布，即  $X_i \sim BE(1, n)$ 。

ii) 任意  $k$  个  $X_{i_1}, \dots, X_{i_k}$  同分布，且与  $(A_1, \dots, A_k)$  同分布。因而  $X_{i_1} + \dots + X_{i_k}$  与  $Z_k$  同分布，而  $U(1) \sim BE(k, n - k + 1)$ 。

iii)  $U(k+1) - U(i) \sim U(1) - U(i)$ ，因此  $U(1) - U(i) \sim BE(k, n - k + 1)$ 。

另外也可证明：若  $Y_1, \dots, Y_k$  独立，且有  $Y_k \sim BE(k, n - k + 1)$ ，则  $Y_1, \dots, Y_k$  独立，且有  $Y_k \sim BE(k, n - k + 1)$ 。

### 1.6.3 指数分布的次序统计量

设  $X_1, \dots, X_n$  为  $i.i.d.$  样本， $A$  服从指数分布  $M(r, 1)$ ， $A = 1/a$ ，其分布密度和分布函数分别为

$f(x) = re^{-rx}$ ， $F(x) = 1 - e^{-rx}$ 。

(1.6.12) 为了推导与指数分布的次序统计量有关的分布，首先导出其前  $r$  个次序统计量的分布。记  $(X_{(1)}, \dots, X_{(r)}) = (y_1, \dots, y_r)$ ， $r \leq n$ ，则应用公式

(1.6.8) 可得  $(A_1, \dots, A_r)$  的分布密度为

只乃， $\dots$   $(y_1, \dots, y_r)$  的分布密度为

$(e^{-y_1})^{n-r} \dots (e^{-y_r})^{n-r+1}$ 。

由于  $Y_i = X_{(i)}$  为次序统计量，因此有  $f(y_1, \dots, y_r) = r! f(y_1) \dots f(y_r)$ ，上式经过化简可得

$f(y_1, \dots, y_r) = r! r^n e^{-r(y_1 + \dots + y_r)}$ 。

注意，由该式可知， $Y_1, \dots, Y_r$  与  $T_{n-r+1}, \dots, T_n$  同分布。

$Y_1, \dots, Y_r$  是两个重要的统计量。

定理 1.6.2 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本，服从由 (1.6.12) 式给定的指数分布。令

其中

$r$   $(n - r)!$   $(1.6.17)$

$T_{n-r+1} = A_1 + \dots + A_r$  (1.6.18)

40

### 第一章统计分布基础

$5^{-(n-r)} (1 - e^{-y_1})^{n-r} \dots (1 - e^{-y_r})^{n-r+1}$

$i: r$   $(1.6.14)$

(1.6.15)  $x_1, \dots, x_n$   $(1.6.16)$

特别，若  $X_1, \dots, X_n \sim M(1/a, 1)$ ，即  $A = 1/a$ ，则  $A_1, \dots, A_r$  的其他结论都成立。并且有

则  $X(1)$  与  $s_1, \dots, s_r$  独立，且有

$s_1, \dots, s_r$

1

$$s = 2[X(i) - X(1)] = x_{a, \dots}$$

$$= \sum_{j=1}^n j^{-1} \cdot i = 1 \quad i=1$$

证明 易见, 当  $r=z$  时,  $S' = 5$ , 因此只需证明 5, 的有关结果即可. 根据 K, "...",  $r$  的分布 (1.6.18) 式, 可做变换

$$z \setminus = ^{(i)} = nYx, \quad Z_2 = (a-1)(X(2) - X(1)) = (\ll - 1)(72-7),$$

$$LZr = (n-r+1)(X(r) - X(r_t))$$

$$= (n-r+1)(y - y_{r_t})$$

心 4,

$$A \quad z^2 \quad \text{-----} + \quad \text{-----}$$

$$n \quad n-1$$

$$n \quad n-1$$

$$+ \dots + A_{n-r+1}$$

$$j = a(i, \dots, /r) = j_{1 \dots 1} (a_r)! \quad d(z_1, \dots, z_r) \quad n \quad n-1 \quad n-r+1$$

根据以上变换关系式, 特别有

$$Z_1 + Z_2 + \dots + Z_r = Y_1 + Y_2 + \dots + Y_r + (n-r)Y_r = T_{nr} \quad (1.6.19)$$

习题一

41

$$Z_2 + \dots + Z_r = K, \quad + \dots + Y_r + (n-r)Y_r - nY_1 = T_{nr} - nY_1 =$$

$S'$  (注意, 恒有  $r_t = ^{(0)}$ ). 因此可得  $z, z_2, \dots, z_r$  的分布

$$, 2; ) = , \dots, y_r)$$

$$I T_i \quad I \quad 7 \quad (y_i < \dots < y_r \quad (= \pm \exp\{-+( : i + z_2 + \dots + - \text{叫} )$$

$$1 \quad k_0, 1 = 2, \dots, r !$$

$$= [ + \exp \{ - ^{(t)} \} ^{U_i \quad I} ] \quad [ - e^{7 \quad (z, \quad 0 \quad | \quad j.}$$

这一表达式说明:  $Z, Z_2, \dots, A$  独立, 且有

$$-nY \} \sim n/x + r(+, 1) \Rightarrow 6 = X(1) \sim /i + r(1). \quad (1.6.20)$$

同时有  $Z_2, \dots, Z_r$  独立同分布,  $Z_2 \sim r(l/a, l)$ ; 因此  $5. = z_2 + \dots + Z_r(+, r-l)$ .

当  $r = r_i$  时有  $S \sim r(l/(T, n-l))$ .

特别, 若  $A \sim r(l/a, l)$ , 即  $r=0$ , 则由 (1.6.20) 式可知  $x\{\} \sim$

$r(n/a, l)$ . 这时也有  $Z_1 \sim r(l/a, l)$ , 与  $z_2, \dots, z_r$  独立同分布, 因此由  $r$

$$(1.6.19) \text{ 式可得 } T_{nr} = 2 \sum_{i=1}^r y\{i\} + (a-r)^{(f)} = z. + z_2 + \dots + z_r \sim$$

$r(l/a, r)$ , 因此 (1.6.18) 式成立. ■

推论 I(1), 尤(2)-尤(1),  $\dots, -^{(n)} -^{(n-1)}$  相互独立, 并且尤(1)与  $^{(n)}(>A)$  独立, 对任意的“&成立.

注 在可靠性统计和生存分析等问题中,  $S': T_n, r-nX(1)$ , 其中  $r$

$^{(n)} = \sum_{i=1}^n x(o + (^{\sim}r)X(r)$  通常表示“总寿命”.  $1=1$

习题一

1-设', 为的 P 分位数, 如定义 1.1.1 所述. 证明:

(1)  $F(x_f) < p$  的充要条件为  $x' < x_p$ ;  $F^{x'} \wedge p$  的充要条件为  $x'$  为  $V$  是否有  $(W) > P$  的充要条件为  $/ > < ' ?$

$$7 \cdot (2!, \dots$$

(2)  $F(X_p - 0) \wedge p \wedge F(X_p)$ ; 若气为  $F$  (幻的连续点, 则  $F(X_p) = P$ ;

去 [it

$$+ (n-r)n\} \}.$$

## 第一章统计分布基础

(3) 若  $F(xz-0) > p$ , 则  $x' \in p$ . 2. 设随机变量  $X$  和  $Y$  的分位数分别为  $x_p$  和  $y_p$ , 证明:

(1) 若  $X$  服从 Pascal 分布  $PA(0, r)$ ,  $F$  服从负二项分布

则  $x_p - y_p + r = 1$ . (2) 若  $X$  服从  $\Gamma$  分布  $r(1/2, M)$ , 则  $x_p = (r/2) \chi^2(2M)$ ; 其中  $\chi^2(M)$  表示自由度为  $M$  的  $\chi^2$  分布的  $p$  分位数;

(3) 若  $V$  服从极值分布  $\text{EV}(a, A)$ , 则  $x_p = (1/a) (\log A - \log \log 1/(1-p))$ ;

(4) 对于  $F$  分布:  $F(n, m; a) = [F(m, n; 1-a)]^{-1}$ .

\*3. 设  $a$  分位函数定义为

$\text{ptt}(0) = (a - I\{t < 0\}) / (|z| [aZ] > 0 + (1-a)I\{t < 0\})$ ,  $0 < a < 1$ ,  $t \in \mathbb{R}$ .

(1) 设随机变量  $X$  的概率密度函数为  $f(x; a, \theta) = a(1-a) \cdot \exp(-I\{x > \theta\})$ ; 证明:  $X$  的  $a$  分位数为  $\theta$ ;

(2) 设  $F$  为连续型随机变量,  $g(M) = E[p_a(r - Zx)]$  关于一切  $M$  存在, 则当  $M \rightarrow \infty$  时  $g(M)$  达到最小, 其中  $r$  为  $F$  的  $a$  分位数.

4. 设  $r(x)$  为随机变量  $X$  的可测函数,  $E[r(x)] = eY$  在  $e \in (a, b)$

上存在, 记  $S(x) = T(x)$ , 若  $a < r(z) \leq b$ ;  $5(\%) = a$ , 若  $7(\%) < a$ ;  $5(\%) = b$ , 若  $T(x) > b$ . 证明:  $E[S(X) - b] \geq E[T(X) - a]^2$ .

5. 称随机变量  $X$  的分布关于某一点  $\theta$  对称, 若其密度函数  $f(x)$  满足  $f(\theta + x) = f(\theta - x)$ . 证明:

(1)  $X$  的分布关于  $\theta$  对称的充要条件为  $X - \theta$  的分布关于原点对称;  $X$  的分布关于原点对称的充要条件为  $X$  与  $-X$  同分布;

(2)  $E(X - \theta)^{2k} = 0$ , 其中  $k$  为正整数;

(3) 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $A$  服从某一对称分布; 则其样本

均值  $\bar{X} = (X_1 + \dots + X_n)/n$  与样本方差  $S^2 = (1/(n-1)) \sum_{i=1}^n (X_i - \bar{X})^2$  不相关, 即  $\text{Cov}(\bar{X}, S^2) = 0$ .

6. 设随机变量  $X$  的  $r$  阶累积量为  $\mu_r$ , 证明:

(1) 若  $X, Y$  独立, 则有  $K_r(X + Y) = K_r(X) + K_r(Y)$ ;

(2)  $K_r(X + C) = K_r(X) + C^r$  ( $r > 1$ ).

7. 设  $f$  为连续型正值随机变量, 其分布函数为  $F(t)$ ,  $Ff(t) =$

$f(t) \int_t^\infty F(u) du$ , 记  $h(t) = f(t) / (1 - F(t))$ , 通常称  $h(t)$  为危险率函数. 证明 (1)  $h(t)$  表示大于  $t$ , 但不超过  $t + \Delta t$  的相对概率, 或危险率:

习题一

43

(2) 危险率函数  $h(t)$  与密度函数  $f(t)$  有以下一一对应的关系:  $f(t) = h(t) \exp(-\int_0^t h(u) du)$ ,  $H(t) = \int_0^t h(u) du$ .

8. 设随机变量  $X$  服从 Poisson 分布  $P(\lambda)$ , 并已知事件  $U=0$  不可能发生, 求此时  $X$  的分布 (截尾 Poisson 分布) 及其期望和方差.

9. (1) 设  $T$  的期望和方差分别为  $\mu$  和  $\sigma^2$ ,  $(V|r=0)$  服从 Poisson 分布  $P(\lambda t)$ , 求  $TV$  的期望与方差;

(2) 设  $V$  服从 Poisson 分布  $P(\lambda)$ ,  $(X|N=n)$  服从二项分布  $B(n, \theta)$ , 证明  $X$  服从 Poisson 分布  $P(\lambda \theta)$ .

10. 设  $r$  服从  $\Gamma$  分布  $r(A, p)$ ,  $(X|T=t)$  服从 Poisson 分布  $P(t)$ . 证明  $X$  服从负二项分布  $NB(e^p, p)$ , 其中  $\lambda = A/(1-p)$ ; 而  $(r|x = \lambda)$  服从  $\Gamma$  分布  $r(A+1, p)$  (提示: 用特征函数证  $X \sim NB(e^p, p)$ )

\*11. 设X服从二项分布,  $P(X=i)=\binom{n}{i}p^i(1-p)^{n-i}$ . 证明其分布函数  $F(i)=P(X \leq i)$  与不完全

β函数(见(1.2.1)式)有以下关系:

$$F(i) = 1 - \frac{1}{B(i+1, n-i+1)}$$

(提示:对第一式, 令左端= $\frac{1}{B(i+1, n-i+1)}$ , 右端= $\frac{1}{B(i+1, n-i+1)}$ , 易得 $\frac{1}{B(i+1, n-i+1)} = \frac{1}{B(i+1, n-i+1)}$ ;第12,13题的证明类似).

\*12. 若X服从负二项分布  $H(r, p)$  则其分布函数  $F(i)=P(X \leq i)$  可由不完全β函数表示为

$F(i) = 1 - \frac{1}{B(r, i+1)}$  表示为

$F(i) = 1 - \frac{1}{B(r, i+1)}$  其中  $r = (r_1, \dots, r_k)^T$ ,  $T = A^{-1}$ . 15. 设  $\phi(x)$  为标准正态分布的分布函数, 证明:

(1) 若  $X \sim N(a, \sigma^2)$ , 则  $E[\phi(X)] = 0$

证.

$$\frac{d}{dx} \phi(x) = -x \phi(x)$$

14. 若  $X_1, \dots, X_k$  相互独立, 则  $E[\phi(X_1, \dots, X_k)] = 0$

$E[\phi(x)]$  表示为二重积分; 或令  $M = e^{t^T \phi(x)}$ , 在积分号下对M求导, 再对X积分;

(2) 和(3)类似;

(2) 若  $X \sim N(0, 1)$ , 则  $E[\phi(X)] = 3/4$ ;

44 第一章统计分布基础

(3) 若  $Z \sim N(0, 1)$ , 则相关系数  $\rho(X, \phi(X)) = 3/4$ . 16. 设  $X_1, \dots, X_n$  为 i.i.d. 样本, 若  $X_i \sim N(0, 1)$ , 其样本均值  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ , 求  $E[\phi(\bar{X})]$ .

17. 设  $X_1, X_2, X_3$  为 i.i.d. 样本, 且  $\phi = \phi(A)$ .

(1)  $Y = X_1 + X_2$

$X_1, X_2$

$X_1, X_2$

$X_1, X_2$  求  $r$

$r_2, y_3$  的联合密度函数, 并判断  $X_1, X_2, X_3$  是否独立;

(2) 令  $X_1 \sim N(0, 1)$ ,  $X_2 \sim N(0, 1)$ , 求  $(X_1, X_2)$  的联合密度函数. 18. (1) 设  $X \sim N(0, 1)$ , 且  $X$  与  $Y$  独立. 证明:  $X$

与  $X/Y$  独立;

(2) 令  $X_1, X_2$  独立同分布,  $X_i \sim BE(n, 1)$ , 令  $Y = X_1/X_2$ ,  $Z = X_1 + X_2$ , 证明:  $Y$  与  $Z$  独立.

19. 设  $X_1, \dots, X_n$  独立,  $X_i \sim N(0, 1)$ ,

$X_1, X_2, X_3$  立, 且  $X_1, X_2, X_3$

$X_1, X_2, X_3$

证明:  $U_1, \dots, U_m$  相互独立,  $j, Z=1, \dots, n-1$  (提示: 令  $U_m = X_1 + \dots + X_m$ )

:  $X_1, \dots, X_m$

$X_1, X_2, X_3$  作变换  $X_1 = (U_1, \dots, U_m)$ ,  $X_2 = (U_2, \dots, U_m)$ ,  $X_m =$

$HU$

20. 设  $X_1, \dots, X_4$  为 i.i.d. 样本, 若  $X_i \sim N(0, 1)$ , 则  $Z = (X_1^2 + X_2^2 + X_3^2 + X_4^2) \sim \chi^2(4)$ .

21. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,

(1) 设  $X_i$  服从均匀分布  $U(0, 1)$ , 证明:  $T = -2 \sum_{i=1}^n \log(X_i)$  服从  $\chi^2(2n)$ ;

(2) 设  $X_i$  服从指数分布  $Exp(\lambda)$ , 证明:  $T = \sum_{i=1}^n X_i$  服从  $\chi^2(2n)$ ; 11

(3) 设  $X_i$  服从 Weibull 分布, 即  $f(x) = \frac{r}{\lambda} x^{r-1} \exp(-x^r/\lambda)$ , 证明:  $r \sum_{i=1}^n X_i^r$  服从  $\chi^2(2n)$ .  $i=1, \dots, n$ .

22. (1) 设  $X$  服从 Poisson 分布  $P(\lambda)$ , 证明:  $(AT - A)/\sqrt{T}$  的渐近分

习题一

45

布为标准正态分布  $N(0, 1)$  (当  $\lambda \rightarrow \infty$ ) (提示: 用特征函数证明);

(2) 设  $X_i$  服从  $t_2(z_i)$  分布, 证明:  $\bar{X}$  的渐近分布为标准正态分布  $N(0, 1)$  (当  $n \rightarrow \infty$ ).

23. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i$  服从 Pareto 分布  $PR(a, 0)$ , 即  $X_i$  的密度函数为  $f(x) = a x^{-a-1}$ ,  $x > 0$ , 令  $T = \sum_{i=1}^n X_i$ , 证明:  $1/T$

(1)  $2a(\log T - n \log 0')$  服从  $\chi^2(2)$ ; (2)  $X(1)$

24. (1) 若  $X_1, X_2$  独立同分布, 且  $X_i \sim N(0, 1)$ , 则  $Y = \sqrt{X_1^2 + X_2^2}$  服从  $\chi^2(2)$ ;

(2) 设  $X_1, \dots, X_n$  为 i.i.d. 样本, 且  $X_i$  服从  $G(4, 1)$ , 证明:  $\bar{X}$  服从  $N(1, 1/n)$ .

\*25. 设  $X_1, X_2$  独立,  $X_1 \sim N(\mu_1, \sigma_1^2)$ ,  $X_2 \sim N(\mu_2, \sigma_2^2)$ , 证明:  $\bar{X} \sim N(\bar{\mu}, \frac{1}{n}(\sigma_1^2 + \sigma_2^2))$ ,

$\bar{X} \sim N(\bar{\mu}, \frac{1}{n}(\sigma_1^2 + \sigma_2^2))$  (提示: 应用条件期望公式计算特征函数).

26. 设  $X \sim N(\mu, \sigma^2)$ , 求  $X$  的特征函数, 并证明它服从非中心  $\chi^2(2, \lambda)$  分布.

27. 设  $X$  服从非中心  $\chi^2(2, \lambda)$  分布. 证明: 当  $n \rightarrow \infty$  时,  $\sqrt{n}(\bar{X} - 1)$  的渐近分布为标准正态分布  $N(0, 1)$ .

\*28. 设  $X_1 \sim P(\lambda_1)$ ,  $X_2 \sim P(\lambda_2)$ , 二者都服从 Poisson 分布. 证明:

$X_1 + X_2 \sim P(\lambda_1 + \lambda_2)$  为非中心  $\chi^2(2, \lambda)$  分布  $r(A, p; 5)$  的分布函数 (提示: 应用关于  $X_2$  的全概率公式及第13题的结果).

\*29. 设  $X \sim N(\mu, \sigma^2)$ ,  $Y = X/\sigma$ ,  $0 < \lambda < 1$ . 用特征函数的展开证明:  $K$  可视为  $(Y, J)$  中  $Y$  的边缘分布, 其中  $J \sim NB(S, n/2)$ ,  $r|J \sim (n+2J)$ .

30. (1) 设  $X$  服从负二项分布  $NB(r, p)$ . 若  $r$  已知,  $p$  未知, 则为指数族; 若  $p$  已知,  $r$  未知, 则负二项分布不是指数族;

(2) 设  $X$  服从 Laplace 分布  $LA(a, b)$ , 若  $a, b$  都未知, 则  $L4(X, a)$  不是指数族; 若  $b$  为常数,  $a$  未知, 则  $LA(X, a)$  为指数族.

31. 设  $X_1, \dots, X_n$  独立同分布,  $X_i$  服从指数族分布  $ED(\eta)$ , 即  $f(x) = \exp(-bx) - c(y)$ ,  $c(y) = a^2$ .

46

第一章统计分布基础

求  $K$  的特征函数, 并证明:  $Y = \sqrt{n}(\bar{X} - \mu)$

服从指数族分布  $ED(0, \sigma^2/n)$

$i=1, \dots, n$

32. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i$  的密度函数  $f(x)$  关于某一点

对称, 即  $f(x) = f(2a - x)$ . 设该样本的第  $f$  个次序统计量的密度

函数为  $g_f(y)$ , 证明:  $g_f(y) = g_f(2a - y)$ , 对任意的  $y$  及  $f = 1, \dots, n$



都成立.

33. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i$  服从标准指数分布  $E(1)$ . 证明:  $Y = X(n) - \log n$  收敛到标准极值分布  $(y) = \exp\{-e^{-y}\}$ .

34. 设  $X_1, \dots, X_k$  为 i.i.d. 样本,  $X_i \sim U(0, 1)$ ,  $i=1, \dots, k$ , 且各组间也独立. 设  $r(n) = Y_{(n)}$  为第  $Z$  组的最大值,  $V = n^{-1} Y_{(n)}$ . 证明:  $F$  的分布密度为  $i=1$   
 $g(v) = 1(-\log v)^{k-1} \cdot \frac{1}{\Gamma(k)}$

(提示: 先求  $u = -\log v$  的分布密度). 35. 设  $(X_1, \dots, X_n)$  为均匀分布  $U(0, 1)$  的次序统计量, 若记  $K =$

$U_{(m)}/U_{(2)} \dots, [X_{(1)}]^{k-1}, \dots, K_{(1)}, Y_n = U\{n\}$ , 证明: 独立, 且有  $Y_k \xrightarrow{d} BE(k, 1)$   
 $-k y^{k-1} \exp(-y)$ ,  $A_i = 1, \dots, n$ ,

\*36. 设  $X_1, \dots, X_n$  为 i.i.d. 样本, 且  $X_i$  服从均匀分布  $U(0, 1)$ . 证明:

(1)  $(A_1, \dots, A_{n-1})$  与  $(X_1, \dots, X_n)$  独立,  $r=1, \dots, n-1$ ;  $A(r+1) = A(r+1) / A(n-1) A(n-2) \dots A(1)$

(提示: 在 (1) 中求联合分布).

37. 设  $X_1, \dots, X_n$  为 i.i.d. 样本, 且  $X_i$  服从  $(a, M)$  上的均匀分布,  $-\infty < a < M < +\infty$ . 证明:  $X_{(1)}$  与  $Y(1)$  同分布, 其中  $X(1), \dots, X(n)$  是  $X_1, \dots, X_n$  的次序统计量;  $r_1, \dots, r_{n-1}$  为 i.i.d. 样本, 且  $r_i$  服从  $(0, 1)$ . 证明:  $J(n), 7^{\wedge}, \dots$ ,  
 独立

(a

(a

38. 设  $X_1, \dots, X_n$  为 i.i.d. 样本, 且  $X_i \sim T(1/a, 1)$ .

$Y_{(n)}$

,  $x$  上的均匀分布;  $y(1)$  是  $Y_{(1)}$  的次序统计量.

+

(1) 证明:  $e^{-X_{(1)}/a} \sim U(0, 1)$ ;

$X_{(n)}$

(2) 求  $X_{(1)}$  的分布, 其中  $X(1)$  为次序统计量中最小的,  $S =$

(3) 证明:  $J(1), X(2) - X(1), \dots, X(n) - X(n-1)$  相互独立, 并且  $X(1)$  与  $X(n) - X(n-1)$  独立, 对任意的  $i, j$  成立.

## 第二章 充分统计量与样本信息

统计推断都是从样本出发, 推断总体的性质 (如参数估计、假设检验等), 且经常可以归结为参数模型, 如正态分布  $N(\mu, \sigma^2)$ , Poisson 分布  $P(\lambda)$  等等. 一种常见的情况是: 假设样本  $X = (X_1, \dots, X_n)$  的分量独立同分布, 且  $X_i \sim P(\theta)$ , 并以此样本来推断参数  $\theta$  的性质, 从而了解总体的性质. 在具体的推断过程中, 都是通过统计量  $T = T(X)$  来进行的. 通俗地讲, 统计量  $T = T(X)$  是对样本  $X$  的一个“加工”或压缩 (其维数由  $n$  降为  $k$ ), 其目的是为了“去粗取精”, 使之形式更加简单, 使用更加方便. 例如,

如, 样本均值与样本方差  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ ,  $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ ,  $T = (\bar{X}, S^2)$ ; 这是经常用到的统计量, 是一个 2 维向量, 而通常样本容量  $n$  要大得多. 自然要问, 通过压缩或降维以后的统计量  $T(X)$  来推断总体与通过原有样本  $X$  来推断总体, 其效果是否一样? 即是否会损失有用的信息? 如果效果一样, 信息未受到任何损失, 则该统计量就称为充分统计量. 充分统计量是 Fisher 于 1922 年提出来的, 这是统计学中非常重要的基本概念之一, 因为它不损失信息地把  $n$  维样本简化为  $k$  维统计量 (通常  $k \ll n$ ), 在此基础上进行统计推断要简单方便.

得多.因此,今后各章介绍的统计推断方法都是从充分统计量出发的.本章主要介绍这方面的内容,第2.1节介绍充分统计量的定义及其判别法,即因子分解定理;同时还介绍极小充分统计量及其判别法;第2.2节介绍与充分统计量有密切关系的完备性以及完备充分统计量的重要性;第2.3节介绍分布族的信息函数的定义与性质,这也是与充分统计量有密切关系的基本概念,充分统计量就意味着从原有样本转换为该统计量没有任何信息损失.有关本章内容,可参见文献陈希孺(1981,1999),Lehmann(1986),Zacks(1981)等.

## 2.1 充分统计量

2.1.1 充分统计量的定义 下面首先介绍统计量,然后再介绍充分统计量.

### 48 第二章充分统计量与样本信息

给定样本 $x_1, \dots, x_n$ , 记为 $x = (x_1, \dots, x_n)$ .  $X$  为随机向量, 通常表示为  $\theta \in \Theta$ . 其中 $\Omega$ 为样本空间; $\Theta$ 为参数空间; $\mathcal{G}$ 为Borel域上的概率测度: 对 $A \in \mathcal{B}$ , 则 $P(A)$ 表示事件 $A$ 发生的概率. 样本 $X$ 的分布函数和分布密度记为 $F(X, \theta)$ 和 $f(x, \theta)$ , 因而样本分布也常记为 $X \sim F(x, \theta)$ ,  $\theta \in \Theta$ 或 $X \sim (r, \theta)$ ,  $\theta \in \Theta$ . 若把 $f(x, \theta)$ 看成参数 $\theta$ 的函数, 则称其为关于参数 $\theta$ 的似然函数.

今考虑样本  $X = (X_1, \dots, X_n)$  的函数  $T = T(X) = (T_1(X), \dots, T_k(X))$ , 其观察值为  $t = T(x)$ . 可设  $x \in \Omega$  时,  $T(x) = t$ .  $X$  与  $T$  类似, 随机向量  $T = T(X)$  也可表示为  $(T_1, \dots, T_k)$ ,  $\theta \in \Theta$ , 其中 $\mathcal{B}$ 为上的Borel域,  $\mathcal{G}$ 的定义如下:

定义2.1.1  $T = T(X)$ 称为 $X$ 的一个统计量. 若 $t = T(x)$ 为

上的可测函数, 即若 $B \in \mathcal{B}$ , 必有

$P_t(B)$ 定义为  $P_t(B) = P\{T(X) \in B\}$

以上导出测度 $P_t(\cdot)$ 的定义可等价地表示为

$f_t(B) = \int_B f(x, \theta) dP^\theta(x)$ ; 或  $f_t(B) = \int_B f(x, \theta) dP^\theta(x)$

(2.1.1)

其中 $f_t(B)$ 和 $f_t(B)$ 为示性函数, 易见上式对示性函数的线性组合, 即简单函数也成立:

$\sum_{i=1}^n c_i I_{B_i}(x) dP^\theta(x)$

$\sum_{i=1}^n c_i I_{B_i}(x) dP^\theta(x)$ .

由于可测函数可表示为简单函数的极限, 因此在一定条件下对可测函数  $f_t(B)$ 有

$f_t(B) dP_t(B) = \int_B f(x, \theta) dP^\theta(x)$ . (2.1.2)

下面开始介绍充分统计量的定义. 首先从直观上了解其含意. 给定样本  $X = (X_1, \dots, X_n)$  和统计量  $T = T(X) = (T_1(X), \dots, T_k(X))$ , 统计量  $T = T(X)$  是对样本 $X$  的一个加工或压缩. 通俗地讲,  $X$  所包含关于参数 $\theta$ 的信息 $= r(x)$ 包含关于参数 $\theta$ 的信息 $+ r(z)$ 已知后 $x$ 还包含关于参数 $\theta$ 的信息. 若 $T = T(X)$ 为充分统计量, 则后者应当为零, 即 $r(x)$ 已知后,  $x$ 不再包含关于 $\theta$ 的信息. 严格来讲, 就是 $r(z)$ 已知后,  $x$ 的条件分布不再与 $\theta$ 有关. 因此可引出以下定义.

$T = T(X)$ 的导出测度

### 2.1充分统计量

定义2.1.2 给定 $\theta$ ,  $T = T(X)$ 称为充分统计量, 若条件概率 $P\{X \in A | T = t\}$ 与 $\theta$ 无关, 即条件分布 $F(x | t, \theta)$ 或条件密度 $f(x | t; \theta)$ 与 $\theta$ 无关. ■

易见, 若 $P\{X \in A | T = t\}$ 与 $\theta$ 无关, 则 $r(x)$ 已知后,  $X \setminus T$ 不再包含关于 $\theta$ 的信息, 因而 $T = T(X)$ 包含了与 $X$ 一样多的关于参数 $\theta$ 的信息. 反

之, 若 $f(x | t; \theta)$ 与 $\theta$ 有关, 贝 $\int f(x | t; \theta) dP^\theta(x) = 1$ 就是关于 $\theta$ 的一个约束条件, 一个关于 $\theta$ 的信息.

易见,  $X$  本身即为充分统计量, 因此充分统计量必然存在. 判别与

/! 6

(2)为证 $X_1+2X_2$ 不是充分统计量, 只需举一反例, 说明条件概率与A 有关即可..

$$\begin{aligned} &=0, X_2=1 \mid X_1+2X_2=2) \\ P(X_1=0, X_2=1; X_1+2X_2=2) \\ &= P(X_1=0 \mid X_2=1) + P(X_2=0) \\ &=(1+f). \end{aligned}$$

该式与参数A有关, 所以 $X_1+2X_2$ 不是充分统计量. | 例2.1.2设 $X_1, \dots, X_n$ 为i. i. d. 样本, $X_i \sim \text{均匀分布}(0, \theta)$ , 则

$\frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$  该式与参数a 无关, 因此 $\bar{X}$  为充分统计量.

$$\begin{aligned} &I \\ &= \\ &A. \\ &r \\ &= (nA)^{-1} \\ &= \frac{1}{nA} \\ &= \frac{1}{n} \cdot \frac{1}{A} \\ &= \frac{1}{n} \cdot \frac{1}{A} \\ &= \frac{1}{n} \cdot \frac{1}{A} \\ &= \frac{1}{n} \cdot \frac{1}{A} \end{aligned}$$

$T = \sum_{i=1}^n X_i$  为充分统计量.

证明 仍然应用(2.1.4)式. 由假设可知:  $n$

$$\begin{aligned} f(x, \theta) &= \frac{1}{n!} \theta^n \exp\left(-\frac{\theta}{n} \sum_{i=1}^n x_i\right) \\ \text{当 } r = U \text{ 寸, } X(1)/\theta &\sim \text{BE}[n, \frac{1}{n}], \text{ 因此} \\ \frac{1}{n} \sum_{i=1}^n x_i &\sim \text{BE}[n, \frac{1}{n}] \text{ 且 } \frac{1}{n} \sum_{i=1}^n x_i \sim \text{BE}[n, \frac{1}{n}] \\ \text{以上两式代入(2.1.4)式可得} \\ \frac{1}{n!} \theta^n \exp\left(-\frac{\theta}{n} \sum_{i=1}^n x_i\right) &= \frac{1}{n!} \theta^n \exp\left(-\frac{\theta}{n} \sum_{i=1}^n x_i\right) \\ \frac{1}{n!} \theta^n \exp\left(-\frac{\theta}{n} \sum_{i=1}^n x_i\right) &> 0. \end{aligned}$$

2.1充分统计量

51

$n$  个独立同分布的随机变量  $X_1, \dots, X_n$  的充分统计量

该式与参数 $\theta$ 无关, 因此 $T = \sum_{i=1}^n X_i$  为充分统计量. 2. 1.2 因子分解定理

Ua f

因子分解定理可简单方便地求出充分统计量, 用途极广. 其严格证明比较复杂, 可参见陈希孺(1981). 以下的推导和证明是一种比较简单的情况.

设统计量 $T=T(X) = (T_1(X), \dots, T_k(X))^T$ ,  $k < n$ , 今补充一个辅助统计量  $W=W(X) = (W_1(X), \dots, W_{n-k}(X))^T$  使  $V \rightarrow Z = (T^T, W^T)^T$  为一一对应(『显然不唯一』). 例如, 若 $T = (T_1, \dots, T_k)^T$ , 可取 $W = (W_1, \dots, W_{n-k})^T$ ; 则显然 $y$ 与 $Z = (T, W)^T$ 一一对应. 假设 $y$ 与 $Z = (T, W)^T$ 之间变换的Jacobi行列式不等于零, 即  $J = \begin{vmatrix} \frac{\partial x_1}{\partial t_1} & \dots & \frac{\partial x_1}{\partial w_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial t_1} & \dots & \frac{\partial x_n}{\partial w_1} \end{vmatrix} \neq 0$

#0, 记  $(r(x), r(x)) = z(x)^T x = x(z)^T$

$X(T, W)$ . 变量之间的函数关系记为  $z = Z(x) = (T(x)^T f_W(X))^T$ ;  $X =$

$X(z) = X(t, w)$ . 由此可得到 $X$ 和 $Z$ 的密度函数之间的关系, 记 $X \sim f(x; \theta)$ ,  $Z = (T, W)$

②, 则有

$$\begin{aligned} f(x; \theta) &= p(T(x) f_W(x); \theta) \\ dx & \end{aligned}$$

(2. 1.5)

看,  $r$  为充分统计量, 即  $X|T$  的分布与  $\theta$  无关. 而  $X|T$  的分布与  $\theta$  无关等价于  $z|T$  的分布与  $\theta$  无关,  $BP(r, JT) \mid T$  的分布与  $\theta$  无关, 也等价于  $(W|T)$  的分布与  $\theta$  无关. 因此有以下引理:

引理2.1.2在上述条件下,  $r(x)$  为充分统计量的充要条件是  $p(w | t; \theta)$  与  $\theta$  无关.

$$p(w, t; \theta) = f(x, t, w; \theta)$$

(2. 1.6)

$d(t, w)$  下面把辅助统计量  $W=W(X)$  应用于充分统计量的研究. 直观上来

证明 由(2. 1.4)和(2. 1.5)式可得

$$p(T(x), IF(x); \theta) \mid T(x) = z \mid$$

$$d(G) dx$$

匯

## 52 第二章充分统计量与样本信息

$$= p(W(x) \mid T(x) = t) \mid d(t, w)/dx \mid$$

因而与  $\theta$  无关的充要条件是  $p(W(x) | T(z); \theta)$  与  $\theta$  无关, 即

$p(w | t)$  与  $\theta$  无关. ■ 定理2.1.1(因子分解定理)  $T = T(X)$  为充分统计量的充要条件是

$f(x; \theta)$  可分解为

$$f(x; \theta) = h(x)g(T(x); \theta) \quad (2. 1. 7) \quad \text{其中 } h(x) \text{ 和 } g(t; \theta) \text{ 都是非负可测函数. 特别, 若}$$

$T=T(X)$  为充分统计

量, 则上式  $g(t; \theta)$  可取为  $r$  的分布密度, 但反之不需要.

证明 为应用引理2.1.2, 取  $Z = (r, JF)$  与  $X$  一一对应, 并设

$$Z = (r, w) \sim p(t, w; \theta). \quad \text{"必要性". 若 } T(X) \text{ 为充分统计量, 要证 (2.1.7) 成立. 由}$$

(2.1.5)

式可得

$$= p(T(x) | W(x); \theta)$$

$$dx$$

$$= P(r; \theta) p(W(x) | T(x); \theta) \mid d(t, w)/dx \mid$$

$$= g(T(x); \theta) h(x),$$

其中  $g(T(x); \theta) = p(T(x); \theta)$  为  $T = T(X)$  的分布密度,  $h(x) =$

$$P(W(x) | T(x) = t) \mid d(t, w)/dx \mid. \text{ 由于 } r(JT) \text{ 为充分统计量, 由引理 2. 1.2 可知, } p(W(x) |$$

$T(x)$  与  $\theta$  无关, 因此  $h(x)$  与  $\theta$  无关,

(2. 1.7) 成立.

"充分性". 若  $g(T(x); \theta) h(x)$  不一定为  $r(X)$

的密度, 要证为充分统计量. 由引理2. 1.2, 只需证明  $p(w | t; \theta)$  与  $\theta$  无关即可. 由(2.

1.6) 式和(2. 1.7) 式可得

$$dx \mid p(t, w; \theta) = f(x; \theta) \mid d(Z, w)$$

$$\mid J \mid$$

$$= g(T(x); \theta) / i(x) \quad \text{其中 } \mid J \mid = \mid dx / a(Z, w) \mid, \text{ 因此条件密度 } p(w | t; \theta) \text{ 可表示为}$$

$$,$$

$$\int p(t, w; \theta) dw$$

$$(尤G, w) \mid \int h(t; \theta) h(X(t, w)) \mid J \mid dw$$

$$= h(X(t, w)) \mid J \mid / \int h(X(t, w)) \mid J \mid dw.$$

该式与  $\theta$  无关, 由引理2.1.2知  $r = r(X)$  为充分统计量. ■ (2.1.7) 式表示: 若  $T(X)$  为充

分统计量, 则  $V$  的分布  $f(x; \theta)$  仅通

2.1充分统计量

过  $r(JT)$  的分布依赖于从

推论(1)  $X$  本身为充分统计量.

(2) 若  $r(x)$  为充分统计量, 并且为  $S(\cdot; V)$  的可测函数,  $g \circ T(X) = \phi(r(X))$ , 则  $S(\cdot)$  为充分统计量.

注意, 充分统计量的可测函数不一定是充分统计量. 例如,  $X$  为充分统计量, 但其可测函数显然不一定是充分统计量.

因子分解定理是判别与推导充分统计量的有力工具, 有非常广泛的应用. 下面举例予以说明. 例 2.1.3 设  $X_1, \dots, X_n$  为 i.i.d. 样本

$\Theta$  或  $A \sim P(A)$ ,

求充分统计量. 解 对两点分布

$n$

$\sum_{i=1}^n X_i$

$\sum_{i=1}^n X_i \sim \text{Bin}(n, \theta)$

$\theta \in (0, 1)$ ,  $f(x; \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$

因此  $\sum_{i=1}^n X_i$  为充分统计量. Poisson 分布类似,  $T = \sum_{i=1}^n X_i$  也是充分

统计量.

例 2.1.4 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim \text{Exp}(\lambda)$  为充分统计量.

$\sum_{i=1}^n X_i$

|

则  $\sum_{i=1}^n X_i$  为充分统计量.

解

$f(x_1, \dots, x_n; \lambda) = \prod_{i=1}^n \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=1}^n x_i}$

$\lambda \in (0, \infty)$ . 由因子分解定理,  $\sum_{i=1}^n X_i$  为充分统计量. I

例 2.1.5 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ , 求充分统计量.

解

$\sum_{i=1}^n X_i$  和  $\sum_{i=1}^n X_i^2$

i) 若  $\sigma^2$  已知, 则  $\sum_{i=1}^n X_i$  为充分统计量; ii) 若  $\mu$  未知 ( $\sigma^2$  已知或未知都一样), 则  $T = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$  为充分统计量. 或等价地 (根据推论),

可取  $T = (\bar{X}, S^2)$  为充分统计量, 其

ii)

54

第二章充分统计量与样本信息

中  $X = (X_1, \dots, X_n)$ ,  $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ . I i-J is 1

例 2.1.6 解

设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim R(0, \theta)$ , 求充分统计量.  $n \geq 1$

$\theta > 0$ ,  $f(x; \theta) = \frac{1}{\theta} e^{-x/\theta}$

$\sum_{i=1}^n X_i$

或  $\sum_{i=1}^n \ln X_i$

$X(1) = \min_{1 \leq i \leq n} X_i$

$X(n) = \max_{1 \leq i \leq n} X_i$

(2.1.8) 因此  $r = (x(1), x(n))$  为充分统计量. 若  $\theta$  已知, 则  $X(1)$  为充分统计量. 特别对  $x$  为充分统计量; 若  $\theta$  未知, 则  $(x(1), x(n))$  为充分统计量.

例 2.1.7 设  $X_1, \dots, X_n$  为 i.i.d. 样本, i)  $X_i \sim \text{Exp}(\lambda)$ ; ii)  $X_i \sim \text{Exp}(\lambda)$

$A = E(X)$ , 求充分统计量.

解 i) 对  $A = 1/\lambda$ ,

$n$

$f(xyz) = f] e'(x, _M)Z | \%f > JU | \gg= 1$   
 $=e^{-nxen} \setminus x(l) = h(x)^{(x(l), /z)}.$

n

注意, 此处用到了  $n7i^{\wedge}Mi=dx(1)>M)$ . 因此  $r=x(1)$  为充分统 1

计量. ii) 对晃  $\sim m + e(a), \setminus n$   
 $/(x$

; A, /z) = \int Ae\_A(W \int i=1

$U(1) \geq \text{蚪} | .$

(2.1.9) 该式仅包含  $x(1)$ ,  $x$ , 因此  $(x(1)J)$  为充分统计量. 或等价地 (根据推

论), 可取  $r; = (x(", \pm)$  不) 或  $r = (x(1), S)$ ,  $5 = \pm(H") = \pm j$ ,

- 为充分统计量, 因为  $r, t; , r$  都是一一对应的, 互为函数. ■ 注意, 对于指数分布  $x$ ,

$\sim/x + E(A)$ ,  $r = (x(1), s)$  为最常用的充分统计

量. 由定理 1.6.2 可知,  $J(1) \sim/z + r(nA J)$ ,  $5 \sim r(A, ^{-1})$  且相互独立. 例 2.1.8 截尾

指数分布. 设  $A, \dots, \text{冬}$  为 i.i.d. 样本,  $\text{久} \sim \wedge^+ r(A, l)$ ; 只观察到前  $r$  个统计量

$(Z(1), \dots, X(r))$  求充分

■

## 2. 1 充分统计量

55

统计量.

解  $(r, , \dots, !;)$  的联合分布在第一章已经求出, 由 (1.6.13) 式可得

加,  $"., /,) = (二" rexp - A 2 \text{冬} + (n-r)yr - n/x \setminus l \setminus yx I. 1=1$

因此  $7 \setminus = (X(1), 7; , J)$  为充分统计量, 其中  $Tnr = ^X(0 + (n - r)X(f). i=H$

应用上经常取其等价形式  $7, = (X(l), S1)$ , 其中  $S. = Tnr - nX(I)$ . 显然  $7;$  与  $\text{丁}-$  对应. I

注意, 由定理 1.6.2 可知,  $X(1)$  与  $S$ , 相互独立, 且有  $X(1) - fl +$

$r(nA, l)$ ,  $S, \sim r(A, r-l)$ . 例 2.1.9 二元正态分布. 设  $\dots, (u,, )$  为 i.i.d. 样本,

$(X, , , yLt2, crj, (j \setminus , p), 0 = (/fj, /X2, CT1 o \setminus , p)$ ; 求充分统计量. 解二元正态分

布的密度可表示为

$f(x, y, \theta)$

$= (- 4 = ^ \setminus 21 TCT, (r2 v1 - p'$

$-1 2(1$

$^2) \}$

$-x \setminus 2 + n(x - /I, )^2$

$\setminus -Mi$

$(T \setminus a2I$

$-1 \_ '1/JL C^{\wedge} \exp 2(1 - /7l: \sim ( i \setminus i=i$

$x \exp$

$V [ /x \ll \sim$

$y(xi - (y. - f) + n(x - ynJCf - /Z2) 4=1$

$+ A( g (y, -y)^2 + \ll (y \sim A2)^2) \}$

(2.1-1.)

因而充分统计量为  $r = (xj, s(x), s(r), s(x, y))$ , 其中  $5(x) = E(x < - J)^2, s(r) = t(r \setminus$

$F)^2, s(x, r) = ^ (x, . - x) (yt. - F)$ . 另外,

$*=j$

$j-i i=i$  若  $P$  未知, 则不管  $\text{川}, M2, g$  是否已知, 都以  $\text{丁}$  为充分统计量;

若  $p = 0$ , 则充分统计量为  $(x, y, s(x), s(y))$ ;

若  $p, a, < r2$  已知, 则充分统计量为  $(^, y)$ . | 例 2.1.10 求指数族分布的充分统计量.

解 (1) 若  $X - f(x^{\wedge}) ^h(X) eQT(e) T(x) - b(e)$ , 则  $r(Z) = (7 \setminus (X), \dots, T \text{人} XW$  为充

分统计量.

(2

)若

$X_n$ 为 i. i. d.样本,  $A$ 、指数族分布

56 第二章充分统计量与样本信息

Wxje即(6)-(10), 则

$X = (x_1, \dots, x_n)$

此时充分统计量为

$T = T(X) = (T_1(X), \dots, T_k(X))$ ,  $I_{\theta} = \sum_{i=1}^k T_i^2(X) / \theta$

注意, 以上结果适用于指数族分布的每一个成员, 诸如: 二项分布、Poisson分布、正态分布、 $\Gamma$ 分布、冷分布等等. 例如对于 $r$ 分布: 若 $X_1, \dots, X_n$ 为i.i.d.样本,  $x_1 \sim r(A, i)$ . 由于

$l(\theta; X) = \sum_{i=1}^n \ln p(x_i; \theta) = n \ln p(x; \theta)$ ,  $U \ln X:$

2.1.3极小充分统计量

给定样本 $X = (X_1, \dots, X_n)$ ,  $T = T(X)$ 是样本的一个

“加工”, 在加工以后, 信息没有损失; 这是统计推断中对统计量的基本要求. 进一步, 还希望这种“加工”越“精致”越好, 即在不损失信息的前提下, 所用的统计量越少越好、越简单越好. 这就是极小充分统计量的涵义.

首先从直观上说明一下统计量的大小概念. 设 $T = T(X) = (T_1(X), \dots, T_k(X))$ , 它是样本 $X = (X_1, \dots, X_n)$ 的函数, 也是对样本的一个

“加工或压缩”, 可认为原来的 $X$ 大, “加工”以后的 $T$ (又)小. 例如  $n$

可认为 $X = (X_1, \dots, X_n)$ 大,  $Z$ 小. 对一般情况, 给

定两个统计量 $T = T(X)$ ,  $r = T^*(X)$ , 若 $V(X)$ 为 $r(X)$ 的一个“加工”, 即 $T^*(X)$

$= V(T(X))$ , 则可认为 $T$ 大,  $V$ 小. 这也可从函数映射观点来了解其大意. 设 $z = r(X)$  e

巧  $C = T^*(x) = e^{J^*}$ , 对函数 $z^* =$

$T = \sum_{i=1}^n t_i$

平(1),

,

因此直观上可认

记则

$=$ );

为, 大, 了 小. 本节介绍的极小充分统计量; 以及第四章介绍的最大不变量都按上述意义来理解.

定义2.1.3称 $T = T^*(X)$ 为极小充分统计量, 若 $V = r(X)$ 为充分统计量; 并且对任一充分统计量 $r = r(x)$ , 都存在 $\phi(\cdot)$ , 使

2.1充分统计量

57

I

的一个“加工”, 因而 $r^*(X)$ 小,  $t(x)$ 大. 以下给出判别、求解极小充分统计量的方法.

引理2.1.3 (1) 设 $T = T(X)$ 和 $r^* = T^*(X)$ 为统计量, 若

由  $T(X_1) = T(X_2)$ 可推出  $r^*(X_1) = r^*(X_2)$ ,  $X_1, X_2 \in \mathcal{X}$ , (2.1.11)

则必存在 $\phi(\cdot)$ , 使  $r^*(x) = \phi(r(x))$ ;

(2)  $r = r(x)$ 为极小充分统计量的充要条件是:  $r^*(x)$ 为充分

统计量, 且对任何充分统计量 $T = T(x)$ 有: 由 $r^*(x_1) = r^*(x_2)$ 可推出  $r^*(x_1) = r^*(x_2)$ ;



$\theta \in \mathcal{X}$  及:

证明(1)定义映射  $\phi(\cdot)$  使得  $V(x), \forall x \in \mathcal{X}$ : 则  $\phi(\cdot)$  确实是一个函数, 因为(2.1.11)式保证了: 对任意的  $T(x_1) = T(x_2)$ , 都有  $\phi(x_1) = \phi(x_2)$  (即在以上定义的映射  $\phi = \phi(\theta)$  中, 不会出现一个  $\theta$  映到两个不同的  $\phi$  的情况).

(2)由(1)可得  $T^*(x) = \phi(T(x))$ ; 由定义可知为极小充分

统计量.  $T^*(x) = \phi(T(x))$ . 由以上定义可知, 极小充分统计量是其他充分统计量  $T(x)$  的函数.

定理 2.1.2 设  $X = (X_1, \dots, X_n) \in \mathcal{X}$

U 画, 对任何

$\theta \in \Theta$ ,  $f(y, \theta)$  与  $\theta$  无关的充要条件为  $r(x) = r^*(x)$ ,

$x, y \in \mathcal{X}$ ; 若似然比  $f(x, \theta) / f(y, \theta)$  与  $\theta$  无关, 则  $r(x) = r(y)$  (必为极小充分统计量).

证明 设  $r = r(x)$  为充分统计量, 要证  $r(X)$  为  $f(x, \theta)$  的函数, 由引理 2.1.3, 只需证明:

由  $T(x_1) = T(x_2)$  可推出  $V(x_1) = V(x_2)$ ;

久,  $X_2 \in \mathcal{X}$ , 今假设  $T(X_1) = T(X_2)$ ;  $X_1, X_2$  由充分统计量的因子分解定理可得

$A_1(x_1) g(T(x_1), \theta) = A_2(x_2) g(T(x_2), \theta)$  且  $A_1(x_1) = A_2(x_2)$  与  $\theta$  无关.

该式与  $\theta$  无关. 由假设条件, 上式与  $\theta$  无关的充要条件是  $T^*(X_1) = T^*(X_2)$ . 因而我们证明了: 由

$r(x_1) = r(x_2)$  可推出  $r^*(x_1) = r^*(x_2)$ ;

$\theta \in \Theta$ : 由引理 2.1.3, 必存在  $\phi(\cdot)$ , 使  $T^*(x) = \phi(T(x))$ , 因而  $r^*(x)$  为极小充分统计量. ■

从下面的例子可以看到, 由该定理经常可以很方便地导出极小充分统计量.

例 2.1.11 设  $X_1, \dots, X_n$  为 i. i. d. 样本,  $(\theta, 1)$ , 求极小充分统计量.

58

第二章充分统计量与样本信息

解

例 2.1.12 充分统计量.

解

$(n \times f(x, \theta))$

$H_r f_e]$

IS 1

$\{1 \leq j \leq n : X_j = 1\}$

$X = (X_1, \dots, X_n)$

1

$\sum_{i=1}^n (x_i - e)^2 e^{ix}$

$f(x, \theta) = \frac{1}{n!} \theta^n e^{-\theta} \prod_{i=1}^n x_i^{x_i-1}$  ----- =  $e^{-\theta} \theta^n$ ,  $e^{-\theta}$ .

似然比与  $\theta$  无关的充要条件是  $x=y$ , 因此  $r(x) = \sum_{i=1}^n x_i$  为极小充分统计量. |

设  $X_1, \dots, X_n$  为 i. i. d. 样本,  $X_i \sim r(A, \theta)$ , 求极小充分

统计量  $T = \sum_{i=1}^n X_i$

$A = \{0, 1\}$  似然比与  $\theta$  无关的充要条件是  $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ ,

$(n \times \sum_{i=1}^n x_i) / n!$

1

为极小充分统计量.

$i=1$

$i=1$

=  $\sum_{i=1}^n x_i$  因此  $r = \sum_{i=1}^n x_i$

|

解

ru 1=1 7

$0 \leq i \leq n-1$

指数族分布. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim \text{Exp}(\theta)$ , 求极大充分统计量.

例 2.1.13

$f(x; \theta) = \frac{1}{\theta} e^{-x/\theta}$ , 求极大充分统计量.

$f(x, \theta) = \frac{1}{\theta^n} e^{-\sum_{i=1}^n x_i/\theta}$

$\theta > 0$

似然比与  $\theta$  无关的充要条件是  $T(X) = \sum_{i=1}^n X_i$

$T$

似然比与  $\theta$  无关的充要条件是  $T(X) = \sum_{i=1}^n X_i$ ; 因此  $T(X) = \sum_{i=1}^n X_i$  为极大充分统计量.

$\sum_{i=1}^n X_i$

注意, 以上结果适用于指数族分布的每一个成员, 诸如: 二项分布、Poisson 分布、正态分布、 $\chi^2$  分布、 $p$  分布等等.

## 2.2 统计量的完备性

例 2.1.14 设  $X_1, \dots, X_n$  为样本,  $X_i \sim \text{Exp}(\theta)$ , 求极大充分统计量.

解

似然比与  $(\theta, M)$  无关的充要条件是  $\sum_{i=1}^n X_i = T$

$= y(1)$ . 因此  $T = \sum_{i=1}^n X_i$

分统计量,  $S = \sum_{i=1}^n X_i$

$= \sum_{i=1}^n X_i$

$A = \sum_{i=1}^n X_i$

$f(x, \theta) = \frac{1}{\theta^n} e^{-\sum_{i=1}^n x_i/\theta}$

$= \frac{1}{\theta^n} e^{-T/\theta}$

$= \frac{1}{\theta^n} e^{-T/\theta}$

$f(x, \theta) = \frac{1}{\theta^n} e^{-\sum_{i=1}^n x_i/\theta}$

$f(y, \theta) = \frac{1}{\theta^n} e^{-\sum_{i=1}^n y_i/\theta}$

$i = 1, \dots, n$

$n$

$\sum_{i=1}^n X_i$

$i = 1$  为极大充分统计量, 或等价地,  $f^*(X) = (X(1), S)$  为极大充分

$n$

$I = 1$

例 2.1.15 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim \text{Exp}(\theta)$ , 求极大充分

分统计量. 解

$f(x, \theta) = \frac{1}{\theta^n} e^{-\sum_{i=1}^n x_i/\theta}$

$f(y, \theta) = \frac{1}{\theta^n} e^{-\sum_{i=1}^n y_i/\theta}$

似然比与  $\theta$  无关的充要条件是  $\sum_{i=1}^n X_i = T$ ,  $\sum_{i=1}^n Y_i = S$ . 因此  $T^* = (X(1), S)$  为极大充分统计量. ■

## 2.2 统计量的完备性

统计量或分布族的完备性与可测函数的数学期望, 即积分有关; 对一族分布, 则与积分变换有关. 例如  $E[h(X)]$ , 可表示为一个积分

变换:  $E[h(X)] = \int h(x) dF(x) = \int h(x) f(x) dx$ . 统计量或分布族的完

备性实际上就是积分变换的唯一性. 本节要讨论的问题是: 若上述积分变换  $\int h(x) f(x) dx = 0$ ,

$\forall h \in \mathcal{H}$ , 是否有  $h(x) = 0$  (a. e.  $P$ )? 或等价地, 若  $\int h(x) f(x) dx = \int h(x) g(x) dx$ ,  $\forall h \in \mathcal{H}$ , 是否有  $h(x) = 0$  (a. e.  $P$ )? 注

意, 以上结论不一定成立, 例如  $X \sim \text{Exp}(\theta)$ ,  $E[h(X)] = 0$  对  $\forall h \in \mathcal{H}$  成立, 但  $h(x) = \sin(x)$

$\neq 0$  (a. e.).

## 60 第二章充分统计量与样本信息

下面首先介绍分布族的完备性与统计量的完备性;然后讨论指数族统计量的完备性;最后介绍Basu定理,该定理揭示了完备充分统计量的重要性质.

## 2.2.1 分布族的完备性

假设或直接表示为  $F = \{f(x, \theta)\}$ ,

其中  $F \setminus x, 0 < f(x, \theta)$ .

定义2.2.1 设  $X \sim f(x, \theta)$ , 称分布族  $F$  为完备的,

若对任何可测函数  $h(x)$ , 由  $E[h(X)] = \int h(x)f(x, \theta)dx = 0$  对任

何  $\theta$  成立, 可推出  $h(x) = 0$  (a. e.  $P_{\theta}$ ). 或等价地, 由  $E[h(X)] = E[h(X)]$  对任何  $\theta \in \Theta$  成立, 可推出  $h(x) = 0$  (a. e.  $P_{\theta}$ ). |

完备性在一定意义上相当于积分变换的唯一性.  $E[h(X)] = \int h(x)f(x, \theta)dx = \int h(x)f(x, \theta)dx$  相当于  $h(x) \in L^1(P_{\theta})$  若由  $E[h(X)] = 0$  可推出

$h(x) = 0$ , 即通常积分变换的唯一性. 因此可由常用积分变换的唯一性来判别分布族的完备性. 常用的积分变换有傅氏变换:

$P(Z) = \int_{-\infty}^{\infty} f(x)e^{itx}dx$ , 即特征函数, 它在  $S(-\infty, \infty)$  上都存在且有

唯一性, 即由  $P(t) = 0$  可推出  $f(x) = 0$ . 另外还有Laplace变换:  $f(x) \rightarrow$

$\hat{f}(s) = \int_{-\infty}^{\infty} f(x)e^{-sx}dx$ , 即矩母函数, 该式在  $s \in \mathbb{R}$  存在, 若至少在  $s \in \mathbb{R}$

$= 0$  的某个邻域(开集)内有定义, 则也有唯一性. 下面举例予以说明. 例2.2.1 分布族  $\{f(x, \theta)\}$  有完备性.

解若有

则对任何  $(A, p)$  有  $\int_A f(x, \theta)dx = 0$ ; 该式左端可视为  $h(x) \in L^1(P_{\theta})$  的拉氏变换, 因此由拉氏变换的唯一性, 可以推出  $h(x) \equiv 0$  (a. e.), 不为零, 即得  $h(x) = 0$  (a. e.). 类似地, 分布族  $\{f(x, \theta)\}$  亦完备. ■

例 2.2.2 正态分布族:  $F = \{f(x, \theta)\}$  的完备性.

解 (1)  $f(x, \theta) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2\theta}}$ , 从  $e^{-\frac{x^2}{2\theta}}$  完备. 因为对任何  $h$ , 由

$$E[h(X)] = \int_{-\infty}^{\infty} h(x)f(x, \theta)dx = 0$$

得

$$\int_{-\infty}^{\infty} h(x)e^{-\frac{x^2}{2\theta}}dx = 0$$

2.2统计量的完备性 61 可以推得

$$\int_{-\infty}^{\infty} h(x)e^{-\frac{x^2}{2\theta}}dx = 0, \quad \theta \in \Theta$$

由拉氏变换唯一性可知:  $h(x) \equiv 0$  (a. e.), 即可得  $h(x) = 0$  (a. e.).

(2)  $f(x, \theta) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2\theta}}$  完备. 与(1)类似.

(3)  $f(x, \theta) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2\theta}}$ ,  $a_2 > 0$  | 不完备. 因为  $h(x) = x$ ,  $E[h(X)] = 0$ , 但  $h(x) \neq 0$  (a. e.).

(4)  $f(x, \theta) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2\theta}}$ ,  $a_2 > 0$  | 不完备. 与(3)类似. (5)  $T = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2\theta}}$ ,  $V/x, < 7$  | 完备. 因为若对任何  $(M, a)$  有

$\int_{-\infty}^{\infty} h(x)f(x, \theta)dx = 0$ , 其中  $K_a(x)$  为正态分布  $N(0, 2)$  的密度函数,

必有  $h(x) \equiv 0$ , 由(1)知  $h(x) = 0$  (a. e.). |

推论 若  $f(x, \theta) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2\theta}}$  完备,  $\theta \in \Theta$ , 则  $f(x, g)$  也完备. 但是反之不一定成立.

因为若对任何  $\theta \in \Theta$ , 由  $E[h(X)] = 0$ , 可推出  $h(x) = 0$  (a. e.), 则对  $(p(\theta) = 0, V(\theta) = 0)$ , 当然有  $h(x) = 0$  (a. e.) (直观上, 越多的使  $E[h(X)] = 0$ , 越容易有  $h(x) = 0$ ). 另外,

在例2.2.2中 成立.

, 丁完备但不完备, 因此以上结论反之不一定

以下介绍判别完备性的其他方法.

例2. 2. 3 二项分布族  $\mathcal{P} = \{ p^n (1-p)^{n-x} \mid 0 \leq x \leq n, 0 < p < 1 \}$  的完备性. 解 若对任何  $\theta$  有  $\sum_{x=0}^n p^n (1-p)^{n-x} \theta^x = 0$ , 即

由此可推出

$\sum_{x=0}^n p^n (1-p)^{n-x} \theta^x = 0$

幻(:)矿  $(1-p)^n \sum_{x=0}^n \binom{n}{x} p^x \theta^{n-x} = 0$ .

n

该式为  $y^{\sum_{x=0}^n \binom{n}{x} p^x \theta^{n-x}}$  的"次多项式, 它对一切  $y \neq 0$  为零, 则其系数必为零, 即  $A(x) = 0$ , 所以  $A(x) = 0, x = 0, 1, \dots, n$ .

例2.2.4均匀分布族:  $\mathcal{P} = \{ \frac{1}{b-a} e^{-\frac{x-a}{b-a}} \mid 0 < a < b < \infty \}$  的完备性. 解若对任何  $\theta$  有  $\int_a^b \frac{1}{b-a} e^{-\frac{x-a}{b-a}} \theta^x dx = 0$ , 则

62

## 第二章充分统计量与样本信息

则  $\int_a^b \frac{1}{b-a} e^{-\frac{x-a}{b-a}} dx = 1$ . 由于  $A(x)$  可测, 其不连续点为零测集. 在

$(a, b)$  的连续点处,  $\frac{d}{dx} \int_a^b \frac{1}{b-a} e^{-\frac{x-a}{b-a}} \theta^x dx$  可导, 因此对任何  $\theta$  有  $\frac{d}{dx} \int_a^b \frac{1}{b-a} e^{-\frac{x-a}{b-a}} \theta^x dx = 0$ , 即  $\frac{d}{dx} \int_a^b \frac{1}{b-a} e^{-\frac{x-a}{b-a}} \theta^x dx = 0, \forall \theta (a. e.)$ , 因此有  $\frac{d}{dx} \int_a^b \frac{1}{b-a} e^{-\frac{x-a}{b-a}} \theta^x dx = 0 (a. e.)$ .

例2. 2. 5 位置参数指数分布族  $\mathcal{P} = \{ \frac{1}{\theta} e^{-\frac{x-\mu}{\theta}} \mid \mu < \infty, \theta > 0 \}$  的完备性.

解 若  $\int_{-\infty}^{\infty} \frac{1}{\theta} e^{-\frac{x-\mu}{\theta}} \theta^x dx = 0$ , 则  $\int_{-\infty}^{\infty} \frac{1}{\theta} e^{-\frac{x-\mu}{\theta}} \theta^x dx = 0, \forall \theta$ .

在  $A(x)$  的连续点处求导可得  $\frac{d}{dx} \int_{-\infty}^{\infty} \frac{1}{\theta} e^{-\frac{x-\mu}{\theta}} \theta^x dx = 0$ , 因此有  $\frac{d}{dx} \int_{-\infty}^{\infty} \frac{1}{\theta} e^{-\frac{x-\mu}{\theta}} \theta^x dx = 0 (a. e.)$ .

推论.  $(M, \mathcal{A}, \mathcal{P})$  完备. 这是因为  $\frac{d}{dx} \int_{-\infty}^{\infty} \frac{1}{\theta} e^{-\frac{x-\mu}{\theta}} \theta^x dx = 0$ . 2. 2.2 统计量的完备性

统计量的完备性与分布族的完备性有显著差异, 下面是一个例子. 例2.2.6 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\mu, \sigma^2)$  且  $E(y)$  存在,

则分布族  $\mathcal{P} = \{ N(\mu, \sigma^2) \mid \mu \in \mathbb{R}, \sigma^2 > 0 \}$  在任何情

况下都不完备.

证明 若取  $h(x) = X_1 - X_2$ , 则  $E(h(X)) = 0$ , 但  $h(x)$  显然不为零. 由此可见, 对最常见的独立同分布样本  $X_1, \dots, X_n$  而言, 其分布族

都是不完备的. 但是并不排除该分布族中有完备的统计量. 例如, 在例 2-2.6 中, 若  $X_i \sim N(\mu, \sigma^2)$ , 则对应于  $A$  或  $X$  的分布族完备 (见例 2.2.2). 由于统计量才是统计推断的出发点, 因此讨论统计量的完备性更有实际意义. 所谓统计量的完备性, 即它所对应分布族的完备性.

定义2.2.2 设  $X_1, \dots, X_n \in \mathcal{P}$ , 统计量  $T = T(X)$  对应的分布族为  $\mathcal{P}_T = \{ F_T : U(z, \theta) \mid \theta \in \Theta \}$

|, 若分布族  $\mathcal{P}_T$  完备, 则称  $T = T(X)$  为完备的统计量. 即由  $E_\theta(h(T)) = 0$  对任何  $\theta$  成立, 可推出  $h(1)$

$= 0 (a. e. \mathcal{P}_T)$ .

例 2. 2.7  $X \sim N(0, \sigma^2)$ ,  $\mathcal{P} = \{ N(0, \sigma^2) \mid \sigma^2 > 0 \}$  不完备, 但是  $T = T(X) = X^2$  完备.

证明由于  $X^2/\sigma^2 \sim \chi^2(1)$ , 因此有

若  $E(h(T)) = 0$ , 则有  $\int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} h(x^2) dx = 0$ , 取  $y = x^2/(2\sigma^2)$ , 则由

拉氏变换唯一性有  $\int_0^\infty t^{1/2} e^{-t} h(t) dt = 0$ , 即  $A(f) = 0 (a. e. \mathcal{P})$ , 因此  $T = X^2$

$X^2$

完备. I

## 2.2 统计量的完备性 63

由定义及以上例题可知, 统计量的完备性有以下性质:

i) 若完备, 则  $\mathcal{P}$  完备, 但反之不真.

ii) 若  $T(X)$  完备, 则  $S(X) = T(X)$  完备, 但反之不真.

iii) 若  $(s, r)$  完备, 则  $r$  完备, 但反之不真.

例 2.2.8  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\mu, \sigma^2)$ , 则  $(\bar{X}, S^2)$  不完备 (见例 2.2.6). 但是  $X \sim N(\mu, \sigma^2/n)$  完备 (见例 2.2.2).

. 同时,  $T$ :

$(X, s)$  也是完备充分统计量 (类似于以下例 2.2.9). | 例 2.2.9  $X_1, \dots, X_n$  独立同分布,  $X_i \sim \text{Exp}(\lambda)$ , 则  $T = (X(1), S)$  完备, 因而为完备的极小充分统计量.

证明 记  $r = (y, s) = (x(1), s)$ , 则  $y, s$  独立, 且有

$f = \lambda e^{-\lambda y} \prod_{i=1}^n \lambda e^{-\lambda x_i} = \lambda^{n+1} e^{-\lambda(y + \sum_{i=1}^n x_i)}$ ,

$S = \sum_{i=1}^n x_i = y + \sum_{i=2}^n x_i$ . 若  $E[h(y, s)] = 0$ , 则有

该式等价于

$\int_0^\infty \int_0^\infty h(y, s) \lambda^{n+1} e^{-\lambda(y + s)} dy ds = 0, \forall \lambda > 0$ .

该式在  $h(y, s)$  中  $y$  的连续点处对  $y$  求导可得  $-\lambda \int_0^\infty h(y, s) e^{-\lambda s} ds = 0, \forall \lambda > 0$ , 由此可得

$\int_0^\infty h(y, s) e^{-\lambda s} ds = 0, \forall \lambda > 0$ . 由拉氏变换唯一性知  $h(y, s) = 0$  (a.e.), 因此

$S = \sum_{i=1}^n X_i$

$- 1)$

$\}$  完备 (类似于例 2.2.7)

$e^{-\lambda \sum_{i=1}^n x_i} = \prod_{i=1}^n e^{-\lambda x_i}$ .  $T = (X(1), S)$  为完备的极小充分统计量. 例 2.2.10  $X_1, \dots, X_n$  独立同分布,  $X_i \sim \text{Exp}(\lambda)$ .

极小充分统计量.

证明  $X(1)/n \sim \text{Exp}(\lambda)$ , 故有

)

| 则  $X(n)$  为完备的

$x$

( $n$

$\lambda, n$ )

.

( $n - 1$ )

若  $E[h(T)] = 0$ , 则  $\int_0^\infty \int_0^\infty h(t, s) \lambda^{n+1} e^{-\lambda(t + s)} dt ds = 0$ , 该式在  $AG$  的连续点处对  $s$  求导可得  $h(t, 0) = 0$ , 所以  $h(t, 0) = 0$  (a.e.), 因此  $T = X(n)$

64

## 第二章充分统计量与样本信息

为完备的极小充分统计量. I 注: 若  $X_1, \dots, X_n$  独立同分布,  $X_i \sim \text{Exp}(\lambda)$ , 则  $r = (x(1), \sum_{i=1}^n x_i)$

为完备的极小充分统计量, 证明详见陈希孺 (1981, p. 78).

例 2.2.11 独立同分布,  $X_i \sim \text{Exp}(\lambda)$ , 则  $r =$

为极小充分统计量, 但不完备. 证明根据均匀分布的性质有  $(X_1, \dots, X_n) \sim \text{Dir}(\alpha_1, \dots, \alpha_k)$ .

则由第一章的定理可知

$r(4) = [(1 - \alpha_1) \prod_{i=1}^n (1 - \alpha_i)]^{n-1}$ . 因此  $E[(X(4) - X(1))] = (n-1)/(n+1)$  对任何  $\theta$  成立, 即  $E[X(n) - X(1)] = (n-1)/(n+1) \neq 0$ , 但显然  $X(n) - X(1) \rightarrow 0$  (a.e.), 因此  $r = (x(1), \sum_{i=1}^n x_i)$  不完备 (直观上看, 对于  $X, 4$  的函数), 由  $(H)$  转变为

$\theta$  相当于参数空间由二维退化为一维, 因而使

$E(M_2)O(r)] = 0$  的参数(4, ^) “不够多”;所以不足以使 $h(t)=0$ 。后面例2. 2. 13的情况也类似)。

另外, 可以证明, 次序统计量 $X(1), \dots, X(n)$ 为完备的充分统计量, 详见陈希孺(1981, p. 82)。

### 2. 2.3 指数族统计量的完备性

定理 2. 2.1 设  $X \sim f(x, \theta) = \frac{1}{Z(\theta)} \exp\{T(x)\eta(\theta)\}$ , 假设  $\theta$  有内点(即亦为 $\theta$ 维集合), 则

(1)  $T = T(X) = (T_1(X), \dots, T_k(X))$  为完备的极小充分统计量; (2) 若 $X_1, \dots, X_n$ 为其 $n$ 个独立同分布的样本, 则

$(T_1(X), \dots, T_k(X))$  为完备的极小充分统计量。

证明 只证明(1)即可。不妨设内点 $\theta_0 = 0$ , 否则可变换到 $\theta = \theta - \theta_0$ 。

氏, 则 $\theta = 0$ 为内点。由第一章的定理可知,

$T(X) = (T_1(X), \dots, T_k(X))$  若 $E_j g(D) = 0, \forall h$  则有

#### 2.2 统计量的完备性

$g(1) A^*(1) e^{i\theta(1)} = 0, \forall \theta$ 。

因此有

$(p(\theta) = g(1) \wedge (C' d f(\theta) = 0, \forall \theta)$

上式在包含 $\theta = 0$ 的某开集内成立, 因而由拉氏变换唯一性知  $g(t)/i^k (t) = 0(a. e.)$ ,

显然  $h^*(\theta \neq 0)$ , 所以必有  $g(1) = 0(a. e.)$ , 因此 $H(X)$ 为完备的极小充分统计量。

由以上定理可推出许多常见分布, 如正态分布、二项分布、Poisson

分布、 $\Gamma$ 分布、冷分布等情形的完备的极小充分统计量。

例2.2. 12 设 $X_1, \dots, X_n$ 为相互独立的样本, 且 $X_i \sim N(\theta, \sigma^2)$ 一切

求完备的极小充分统计量: (1) 若 $\sigma^2 \sim 1/\chi^2(k)$ ; (2) 若 $\theta \sim N(\mu, \sigma^2)$ ; 其中 $\mu \in R, \sigma^2 > 0$ , 都是未知参数。

解 (1)  $X = (X_1, \dots, X_n)$  的联合分布可表示为  $f(x, \theta) = \exp\{-\frac{1}{2\sigma^2} \sum_{i=1}^n x_i^2 - \frac{\theta}{\sigma^2} \sum_{i=1}^n x_i - \frac{n}{2\sigma^2} \theta^2\}$

其中 $\eta(\theta) = D \begin{pmatrix} \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i^2 \end{pmatrix}$

其中  $r_1(x) = -1/(2\sigma^2)$ ;  $r_2(x) = \sum_{i=1}^n x_i^2$ ,  $\eta_1(\theta) =$

$-\theta/(2\sigma^2)$ ;  $b(\theta) = n \log \sigma^2 - \frac{n}{2\sigma^2} \theta^2$ 。因此,  $(\sum_{i=1}^n x_i, \sum_{i=1}^n x_i^2)$  为完备

的极小充分统计量。以下是不符合定理2. 2. 1所列条件的一个反例。

例2.2.13 设 $(U_1), \dots, (U_K)$ 独立同分布,  $(X^Y) \sim N(\theta, \sigma^2)$

且 $\theta \sim N(\mu, \sigma^2)$ 。因此,

;

$Q(\theta) = -1/(2\sigma^2)$ ;  $T_1(x) = \sum_{i=1}^n x_i$ ;  $Q_2(\theta) = \theta/\sigma^2$

为完备的极小充分统计量。(2)类似地,  $X = (X_1, \dots, X_n)$  的联合分布可表示为

65

$M_2,$

,

:

/

4) (即 $p=0$ ), 可化为指数族

$=T(X), T_2=X, T_3=r(7), T_4=y,$

其中  $t(x) = \pm$

$\leq 1 : = 1$

$H_n = 2^{-n} y^n$ —因此  $T = (T(x), T(r), y)$  为完备

66

## 第二章充分统计量与样本均值

的极小充分统计量. 但是若  $(4, 6)$  即  $A = M_2$ ,  $M_r$  仍然为极小充分统计量, 但不是完备的统计量. 因为对任何  $f(x)$ ,  $E f(x - Y) = 0$ , 而显然  $f \neq 0$ . 这一结果与定理 2.2.1 并不矛盾. 因为当

时, 参数空间退化为 3 维, 无内点. 2.2.4 Basu 定理

完备充分统计量与独立性有密切关系, 表现为以下 Basu 定理, 先介绍一个常用概念.

定义 2.2.3 辅助统计量(ancillary statistics). 设  $X \sim \theta$ , 若统计量  $A = A(X)$  的分布与  $\theta$  无关, 则称  $A(X)$  为辅助统计量(即  $A(X)$  中不包含关于  $\theta$  的信息).

例 2.2.14 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\theta, 1)$ , 则

$|$   
 $i=1$  因为与  $\theta$  无关.  $T = X\{n\} - X(1)$  亦为辅助统计量,  
为辅助统计量,

量, 因为  $T = (X\{n\} - X(1)) - (Y(4) - Y(1))$ , 而  $Y \sim N(0, 1)$ , 其分布与  $\theta$  无关. 2.2.2 (Basu 定理) 设  $X \sim \theta$ ,  $A = A(X)$  为完备充分统计量,  $B = B(X)$  为辅助统计量, 则  $A(X)$  与  $B(X)$  独立.

证明 为证  $A(X)$  与  $B(X)$  独立, 只需证  $P_{\theta}(A(X) \in B) = P_{\theta}(A(X) \in B) \cdot P_{\theta}(B(X) \in C)$ , MB. (2.2.1)

首先注意, (2.2.1) 式与  $\theta$  无关. 因为在左端:  $r(x)$  为充分统计量, 因而  $X \setminus T$  分布与  $\theta$  无关, 所以左端概率与  $\theta$  无关; 而由辅助统计量的定义可知, 右端也与  $\theta$  无关. 易见 (2.2.1) 式等价于

$P_{\theta}(A(X) \in B) = P_{\theta}(A(X) \in B) \cdot P_{\theta}(B(X) \in C)$  (2.2.2) 若记  $C = A^{-1}(B) = \{x: A(x) \in B\}$ , 则 (2.2.2) 式可表示为

$P_{\theta}(X \in C \mid T(X) = t) = P_{\theta}(X \in C)$ . (2.2.3) 记该式右端  $P_{\theta}(C) = a$ , 它与  $\theta$  无关. 因此要证 (2.2.3) 式, 即相当于

要证

$E[UC(X) \mid T(X) = z] = a$ . (2.2.4) 为此可应用  $T = T(X)$  的完备性, 记  $h(t) = E[UC(X) \mid T(X) = t] - a$ .

要证 (2.2.4) 式, 即相当于要证  $E[h(T)] = 0$ . 为此, 可计算其期望  $E[h(T)]$

$= E[h(T)] = E[E[h(T) \mid T]] = E[h(T)]$

$= E[h(T)] = 0$ . 由于  $T = T(X)$  为完备充分统计量, 由  $E[h(T)] = 0$ , 可推出  $h(t) = 0$  (a.e.), 因而 (2.2.4) 式成立, 从而 (2.2.1) 式成立. 因此  $A(X)$  与

n

## 2.2 统计量的完备性

67

### 4(独立. ■

例 2.2.15 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\mu, \sigma^2)$ , 则 n

(1)  $X$  与统计量  $S_1 = n^{-1} \sum_{i=1}^n X_i$ ,  $S_2 = n^{-1} \sum_{i=1}^n X_i^2$

$i=1$

$S_3 = \sum_{i=1}^n X_i^3$ ,  $S_4 = \sum_{i=1}^n X_i^4$  独立.

(2) 若  $g(x_1, \dots, x_n)$  满足平移不变性, 即  $g(x_1 + c, \dots, x_n + c) = g(x_1, \dots, x_n)$ ,

$|x_1|, \dots, |x_n|$ ,

$+c) = g(x_1, \dots, x_n)$

$g(x_1, \dots, x_n)$ , 则  $g(X_1, \dots, X_n)$  独立.

证明 任意固定  $a = a_0$  则  $X$  为完备充分统计量. 令  $X_i =$

$Z = 1, \dots, a$ , 则  $K \sim \chi^2(n, d)$ , 其分布与  $M$  无关.  $nn$

(1)  $S_1 = n^{-1} \sum_{i=1}^n (Y_i - \bar{Y})^2 = n^{-1} \sum_{i=1}^n Y_i^2 - \bar{Y}^2 = K(n) - K(1)$ ,  $i=1, \dots, n$   
 $S_4 = \sum_{i=1}^n |Y(i) - Y(j)|$  都是辅助统计量, 因而与完备充分统计量  $i$

独立.  $\dots, X_j - \bar{X}, \dots, -\bar{X}) =$

(2) 由平移不变性,  $g(K, \dots, r_j)$ , 因而其分布与  $n$  无关, 为辅助统计量, 所以与  $i$  独立.

以上独立性对任意的  $A$  成立, 因此对  $x, -n(m, o-2)$  也成立. 例 2.2.16 设  $Z_1, \dots, Z_n$  独立同分布,  $A \sim \chi^2(r(A), l)$ , 则对任意

$f$

的  $f$ ,  $I(1)$  与  $X(7) - X_0$ ,  $Y - \bar{Y}$  独立; 也与  $S = \sum_{i=1}^n Z(x(o - Y(1))) + i=1$   
 $-r) [X(r) - X(1)]$  独立.

证明 对任何固定的  $a = a_0$ , 则  $x(1)$  为完备充分统计量. 若记  $r, \dots =$   
 $h$

$p$  无关, 都是辅助统计量; 因此可以得到以上结论. ■

,  $f = 1, \dots, n$ , 则  $Z$  的分布与  $n$  无关,  $X_i - X_j = Y_i - Y_j$  的分布也与

例 2.2.17 设  $Z_1, \dots, Z_n$  独立同分布,  $AKCM$ , 则对任意的  $Z$ ,

$h X(4)$  与  $Y(1)/Y(7)$ ,  $A/\bar{X}$  独立.

证明 若记  $r = X, f = \bar{X}$ , 则其分布与  $n$  无关, 因而  $X_i/X_j = Y_i/Y_j$  的分

布与  $n$  无关, 都是辅助统计量; 由此可以得到以上结论. ■ Basu 定理在统计中有广泛应用. 对于完备的充分统计量, 今后可根据 Basu 定理得到与之独立的许多统计量, 这在参数估计、假设检验等问题中经常出现. 比较常见的情形之一就是求条件概率或条件期望 (因为不少最优解常与之有关). 例如欲求条件期望  $E_0[W(X) | r(J)]$ , 如果  $r(x)$  为完备的充分统计量, 而  $f(X)$  为辅助统计量, 即其分布与

$n$  无关 (或者经过转换可化为这种情形). 则由于  $p(x)$  与  $r(z)$  独立, 上述条件期望即可简化为无条件期望

68

## 第二章充分统计量与样本信息

### 2.3 分布族的信息函数

分布族的信息函数是与充分统计量有紧密联系的概念. 设  $\theta = (\theta_1, \dots, \theta_k)$ , 若  $T = T(1)$  为充分统计量, 则用  $HZ$  代替原 有样本  $X$  作统计推断不应该有信息损失. 自然要问, 何谓信息? 如何严格加以定义, 并使之符合通常统计推断的直观意义? 诸如: 样本  $X$  与充分统计量  $r(x)$  的“信息”应该相等;  $(\theta_1, \dots, \theta_k)$  的信息应为各分量  $\theta_1, \dots, \theta_k$  所含“信息”之和; 等等. 事实上, 符合若干直观意义的“信息”的定义并不唯一, 通常有三种: (1) Fisher 信息, 这是统计学上应用最广泛的“信息”; (2) Kullback - Leibler 信息 (亦称 K-L 距离), 它在统计中应用相对较少; (3) Shannon 信息 (即熵), 这是通讯中广泛应用的“信息”. 本节主要介绍 Fisher 信息, 同时也简要介绍一下 Kullback - Leibler 信息.

#### 2.3.1 Fisher 信息

首先提出一组正则条件, 这是定义 Fisher 信息所必需的, 也是通常研究统计问题所必需的; 这些条件对于许多常见情形都能满足. 当然也有“非正则”的情形, 届时再予以说明.

定义 2.3.1 设  $\theta$  称为正则分布族 (也称为 Cramer-Rao 正则分布族或简称为 C-R 分布族):

(1) 设  $\theta = (\theta_1, \dots, \theta_k)$ ,  $\theta \in \Theta$ ,  $\Theta$  为  $R^k$  上的开集; 若  $\theta \in \Theta'$ , 则必有  $M \cup \theta \in \Theta$  (即  $\theta$  是可识别的, 不同的  $\theta$  对应于不同的分布).

$= 1$ , 若分布族满足以下条件

,

$\theta \in \Theta$  为  $R^k$  上的开集; 若  $\theta \in \Theta'$ , 则必有  $M \cup \theta \in \Theta$  (即  $\theta$  是可识别的, 不同的  $\theta$  对应于不同的分布).



(2) 分布族的对数似然函数记为  $L(\theta) = L(\theta; x) = \log l(x; \theta)$ ,  $L(\theta, x)$  关于  $\theta$  存在二阶以上的导数, 其前二阶导数记为  $Z(\theta)$  和  $L(\theta)$ , 其分量为  $dL(\theta, x) / d\theta_j$  和  $d^2L(\theta, x) / d\theta_j d\theta_k$ ,  $j, k = 1, \dots, k$ .

(3) 记  $L(\theta; x) = (s_1(x, \theta), \dots, s_k(x, \theta))^T$ ,  $s_i(x, \theta) = dL(\theta, x) / d\theta_i$ .  $S(\theta)$  称为得分函数 (score 函数), 假定它在  $\theta_0$  上存在前二阶矩; 即  $EJ(\theta_0, x) = dL(\theta_0, x) / d\theta_j$  在  $\theta_0$  上存在,  $j = 1, \dots, k$ .

(4)  $F = \{f(x, \theta), \theta \in \Theta\}$  有共同支撑, 即  $\theta_j > 0$  与  $\theta_k$  无关.

(5)  $f(x, \theta)$  关于  $\theta$  的积分和关于  $x$  求导可交换次序, 即关于  $\theta$  可在积分号下求导数.

定义 2.3.2 Fisher 信息. 若  $\theta_0 \in \Theta$  为正则分布族, 则

2.3 分布族的信息函数

$A; (8) \quad I(\theta_0) = -E\{J(\theta_0, X)^2\} = -E\{J(\theta_0, X)J(\theta_0, X)^T\}$

分别称为分布族关于参数  $\theta$  的 Fisher 信息函数和 Fisher 信息阵. 若  $\theta_0 = r(X)$  为正则分布族,  $L(\theta, t) = \log g(e, t)$ , 则  $r(x)$  关于参数  $e$  的 Fisher 信息定义为  $I(\theta_0) = -E\{J(\theta_0, X)^2\}$ .

Fisher 信息有许多良好的性质, 与通常统计推断的直观意义相吻合. 以下首先介绍其基本性质, 然后证明一个重要定理.

(1) score 函数满足  $EJ(\theta_0, X) = 0$ , 并且有  $I(\theta_0) = \text{Var}J(\theta_0, X) = E\{J(\theta_0, X)J(\theta_0, X)^T\} = E\{J(\theta_0, X)Z(\theta_0, X)^T\}$

(2.3.1) 证明 对于正则分布族,  $f(x, \theta)$  关于  $\theta$  的积分和关于  $\theta$  求导可交换次序, 因此有

$= \int \frac{\partial}{\partial \theta_j} f(x, \theta) dx = \frac{\partial}{\partial \theta_j} \int f(x, \theta) dx = 0$ .

因而由 Fisher 信息的定义有

$\text{Cov}(S) = E\{SS^T\} = E\{JJ^T\} = I(\theta_0)$ .

以上公式的矩阵形式为  $I(\theta_0) = \text{Var}X = EXS^T = EXfT$ , 即 (2.3.1) 式. ■

(2)  $I(\theta_0) = -E\{d^2L(\theta_0, X)/d\theta_j d\theta_k\}$ , 即  $I(\theta_0) = -E\{Z(\theta_0, X)Z(\theta_0, X)^T\}$ . 证明直接求导, 并取期望可得  $dL(\theta_0, X) = d \log l(\theta_0, X) = \frac{1}{l(\theta_0, X)} \frac{\partial l(\theta_0, X)}{\partial \theta_j} d\theta_j$ .

$d\theta_j = d\theta_j$

$l(\theta_0, X) = \int l(\theta_0, X) d\theta_j$

9 並二立  $l(\theta_0, X) = \int l(\theta_0, X) d\theta_j$

$\sim f(x, \theta_0) = \int f(x, \theta_0) d\theta_j$

$< E\{f(x, \theta_0) f(x, \theta_0)^T\} = \int f(x, \theta_0) f(x, \theta_0)^T dx$

$f$

$2 d\theta_j d\theta_k \} f(x, \theta_0) = \int f(x, \theta_0) d\theta_j d\theta_k$

$= I(\theta_0)$ . ■

(3) 若  $X_1, \dots, X_n$  独立,  $X = (X_1, \dots, X_n)^T$ , 则  $I_X(e) = \sum_{i=1}^n I_{X_i}(e)$ , 特别, 若  $X_1, \dots, X_n$  独立同分布, 则  $I_X(e) = n I_{X_1}(e)$ .

70

第二章充分统计量与样本信息

证明由  $\theta_0 = \theta_0$  可得  $I = 1$

$n$

$l(\theta_0, X) = \sum_{i=1}^n l(\theta_0, X_i)$ ,  $s(\theta_0, X) = \sum_{i=1}^n s(\theta_0, X_i)$

$n$  因而由独立性可得  $\text{Var}[S(\theta_0)] = \sum_{i=1}^n \text{Var}[s(\theta_0, X_i)]$ , 即  $I_X(\theta_0) = \sum_{i=1}^n I_{X_i}(\theta_0)$ .

$z = \sum_{i=1}^n z_i$

(4) 设  $r =$

i)  $1/TW : \theta_0$  的充要条件是  $r(x)$  为辅助统计量; ii) 若  $T(x)$  为充分统计量, 则  $r(x) = T(x)$

(沒). 证明i)由以上性质(1)可得

■

$r(x)$

则有

$$IT(0)=\text{Var}[S(T,0)] = \text{Var}[d(\cdot/\cdot, (4)=0 \ll E, \{[\wedge^j T]=0$$

$$l=\sim \log^{\wedge} \cdot)=c(1),$$

其中 $c(z)$ 为与0无关的常数, 因此 $r$  的分布密度函数 $g(t,e)$ 与 0无关, 所以 $t=t(x)$ 为辅助统计量. .

ii)若 $r = r(x)$ 为充分统计量, 则由因子分解定理可得 $f(x, G)= g(T(x), 0)h(x)$ , 因此有  $\log/(\%, 沒)=\log g(r(x), 6>) +\log/l(x)$ .该式 对沒求导得到

$$\text{diogf}(x, 0)=\text{alog } g( r(x) ,0、 \sim \text{de } _ \text{de } ,$$

$$\text{即 } L(0,X) =L(\wedge,T) \text{.故由(2.3.1)式可得 } I_x(0) =I_r(0) \text{. } |$$

(5)设②峨为参数变换, 乎为 $q \times l$ 维参数, 则分布族关于参 数平的Fisher信息阵可表示为  
de T de d(p7 d(pT

(2.3.2)

其中 $d_0/\text{dip1} = ( d_0t/ a^{\wedge})A x?$ 为 $A \times 7$ 阶矩阵.(2.3.2)式的分量形式为

$$\wedge \ll b(<p)= 2A)(^)(\text{"/}=1, \dots, \text{女}; a, 6=1, \dots, q) \text{. } o(pa d(pb$$

证明由定义可得

$$k \text{ } 3 \text{ } 3 \text{ } 3 \text{ } 3$$

.

## 2.3分布族的信息函数

=石  $U(10)a$ . ■

上面介绍了 Fisher信息的若干基本性质.特别是性质(3)和性质 (4), 指出了 Fisher信息的基本统计意义, 但是还不完全.以下定理

2.3. 1将要进一步证明:对于任何统计量 $r = r(j)$ 都有心(幻幻入②、, 而且等式成立的充要条件是 $T = T(X)$ 为充分统计量;这就更加完整地 揭示了充分统计量与Fisher信息之间的内在联系.为了证明这一定理, 需要用到得分函数的条件期望的性质.所以下面首先简要介绍条件期望 的一个等价定义, 关于条件期望进一步的内容可参见陈希孺(1981, p. 25—设43).

$t = t(x)$ 为统计量, 则有 其 导出测度为 $P \setminus B) = P_x(r-'(5)) = P_X\{A\}$ .给定可测函数 $\varphi(X)$ , 其 条件期望 $E[\wedge(x)IT(X)=d$ 应得的函数 $m(t)$ , 并且应当满足条件

期望的基本性质(见第一章(1. 1.4)式):

$$E[m(r)] = E (E[\wedge(X) | T] | = E[<p(\wedge)] \text{.}$$

其积分形式为

$$\text{卜(1)dPT}(t)=f<^{\wedge}(x)dPJ(x) \text{. (2.3.3)}$$

该式就是条件期望等价定义的出发点, 具体表述如下: 可测函数 $p(X)$ 的条件期望 $E[<p(X) | T(X) =d$ 定义为上的瓜可

测函数瓜(1), 它满足

$$\text{上 } m(t)dPT(t) = f : (3) (p(x)dPx(x) , \forall BeBr. (2.3.4)$$

在这个定义中, 若令 $B=Z$  就是(2.3.3)式, 因此(2.3.4)式比(2. 3.3) 式更具有一般性, 要求也更高一些.根据(2.3.4)式, 可以进一步定义 条件概率和条件分布.条件概率定义为 $P(A | T(X) =0 = E \setminus IA(X)$

$I T(X) |$  ;条件分布定义为  $F(x|i) 1( -\infty ,x] | T(X) = z$  可以证明:在一定正则条件下,  $m(f)$ 必存在;以上定义的条件概率、条 件分布具有与经典定义一样的性质, 详见陈希孺(1981).

对于本节考虑的正则分布族, 设 $x- |$  ,  $T=T(X)$ 为 统计量,  $t \sim$  ).则 $I$  与 $r$ (幻的得分函数

有以下关系.

引理2.3.1在以上条件下, 统计量 $r(x)$ 的得分函数可表示为 $X$  的得分函数的条件期望  
 $=E\{f'(\log(x)) | T(X)\}$ 叫. (2.3.5) 该式等价于 $S_r(T, e) = E\{S_x(X, e) | T\}$ , 亦可简  
记为 $S_r = E\{S_x | T\}$ .

71

## 72 第二章充分统计量与样本信息

证明 今应用(2.3.4)式来证明(2.3.5), 为此记  $\eta = r(r)$ ,  $\eta = f(x; e)$

要证这两式满足(2.3.5)式, 等价于要证

$$f'(\log(x)) = J; \int f(x, \theta) dP(x), \quad (2.3.6)$$

等价于要证

等价于要证

即要证

如(1)=盖土, <8戶』)叫(幻,

$$\eta(5) = \text{盖尸}: (7^8).$$

而由 $P_j(B)$ 的定义可知 $P_j(B) = \eta^{j-1}(5)$ , 由此倒推即可得到 (2.3.6), 因此(2.3.5)

式成立. ■ 定理 2. 3.1 设  $X \sim f(x; \theta) \theta \in \Theta$ ,  $T(X) = g(t; e)$  都是正则

分布族, 则有

(1)  $X$ 与 $T(X)$ 的Fisher信息之差(即以 $T(X)$ 代替 $T$ 的信息损失)

可表示为

$$I(X, T) = E\{[f'(X, T)]^2\} = E\{[S_x(X, \eta) - S_r(T, \eta)]^2\}$$

$$= \text{Var}\{S_x(X, \eta) - S_r(T, \eta)\} = \text{Var}\{S_x(X, \eta)\} - \text{Var}\{S_r(T, \eta)\}.$$

(2)  $I(X, T)$ 为 $IT(\theta)$ , 且等号成立的充分必要条件为 $r(x)$ 为充分统计量.

证明(1)根据公式(2.3.1)以及第一章(1.1.5)式可得  $W = \text{Var}\{\eta\} = \text{Var}\{S_x(X, \eta)$

$$= E\{[S_x(X, \eta)]^2\} = \text{Var}\{S_x(X, \eta)\} + [E\{S_x(X, \eta)\}]^2. \text{ 由引理2. 3. 1知 } E\{S_x(X, \eta) | T\}$$

$$= S_r(T, \eta), \text{ 因此上式第二项} = \text{Var}\{S_r(T, \eta)\} = IT(\theta), \text{ 代入以上(9)的表达式, 即可以}$$

推出(1)的

第一式. 亦可推出 $I(X, T) = IT(\theta)$ . 再证(1)的第二、三式. 根据Fisher信息的定义, 以上(1)

的第三式可表示为

$$E\{[S_x(X, \eta) - S_r(T, \eta)]^2\}$$

$$= E\{S_x(X, \eta)^2 - 2S_x(X, \eta)S_r(T, \eta) + S_r(T, \eta)^2\}$$

$$= E\{S_x(X, \eta)^2\} - 2E\{S_x(X, \eta)S_r(T, \eta)\} + E\{S_r(T, \eta)^2\}. \text{ 而由引理2. 3. 1可得}$$

$$E\{S_x(X, \eta)S_r(T, \eta)\} = E\{S_x(X, \eta) | T\} S_r(T, \eta) = S_r(T, \eta)^2 = IT(\theta). \text{ 同理 } E\{S_r(T, \eta)^2\} =$$

$$IT(\theta). \text{ 因此有 } E\{S_x(X, \eta)^2\} - 2IT(\theta) + IT(\theta) = E\{S_x(X, \eta)^2\} - IT(\theta)$$

由此即得第二、三式. (2)上面已证明 $I(X, T) = IT(\theta)$ . 若 $T=T(X)$ 为充分统计量, 则由

性质(4)可知 $I(X, T) = IT(\theta)$ . 反之, 若 $I(X, T) = IT(\theta)$ 要证 $r(x)$ 为充分统计量. 由(1)第二

式有 $S_x(X, \eta) = S_r(T, \eta)$ , 即

$$f'(x, \eta) = S_r(T, \eta) \quad \text{即} \quad f'(x, \eta) = S_r(T, \eta)$$

$$f'(x, \eta) = S_r(T, \eta) \quad \text{即} \quad f'(x, \eta) = S_r(T, \eta)$$

该式两端关于 $\eta$ 积分可得

$$\log f(x, \eta) = \log g(T, \eta) + c(x),$$

其中 $c(x)$ 为与 $\eta$ 无关的积分常数, 所以有 $f(x, \eta) = g(T, \eta) e^{c(x)}$  因

此由因子分解定理可知,  $T=T(X)$ 为充分统计量. 以上定理和性质表明, 无论从Fisher信

息观点来了解充分统计量或

者从充分统计量的观点来理解Fisher信息都是合理的. 下面看一些例子.

例2. 3.1 设 $X \sim N(\eta, a^2)$ 为正态分布. 求 $X$ 关于 $\eta = (\mu, a^2)$ 以及  $\theta = (\mu, a^2)$ 的

Fisher 信息.

解由定义可得

$$L(\theta, a^2) = -\log(2\pi\sigma^2) - \frac{1}{2\sigma^2} (z - \theta)^2$$

$$dL/d\theta = (z - \theta)/\sigma^2$$

$$d^2L/d\theta^2 = -1/\sigma^2$$

$$a^2 = \sigma^2$$

$$= \text{Var}(z) = \sigma^2$$

Var

/(x, (T2)= 若又i, ...人独立同分布,  $X' \sim N(\theta, a^2)$ , 则有

$I_x(\theta, a^2) =$  若取参数为  $\theta = (\theta_1, \dots, \theta_k)$ , 通过类似的计算可得

$$Z^T(z, \theta) =$$

$$n/\sigma^2$$

$$0 \quad 2n/\sigma^2$$

$$1/\sigma^2 \quad 0 \quad 0$$

$$n/\sigma^2 \quad 0$$

$$0 \quad n/2\sigma^4$$

■

74

第二章充分统计量与样本信息

例2.3.2 设  $X \sim b(1, \theta)$  为两点分布, 求  $X$  关于  $\theta$  的 Fisher 信息. 解由定义可得

$$x - f(x, \theta) = x - \theta$$

$$L(\theta, x) = \log \theta + (1 - \theta) \log(1 - \theta)$$

$$dL/d\theta = 1/\theta - 1/(1 - \theta)$$

$$d^2L/d\theta^2 = -1/\theta^2 - 1/(1 - \theta)^2$$

$$= -\text{Var}(X) = \theta(1 - \theta)$$

$$1 - \theta$$

$$] \cdot \theta$$

似地, 若  $X \sim b(n, \theta)$ , 也有  $I_x(\theta) = n[\theta(1 - \theta)]^{-1}$ .

例2.3.3 设  $Y$  服从位置尺度参数分布族  $\theta = (\mu, \sigma)$ .

证明: 若  $\sigma$  已知, 则  $I(\mu) = 1/\sigma^2$ ; 若  $\mu$  已知, 则  $I(\sigma) = 2/\sigma^2$ ; 若  $\mu, \sigma$  均未知, 则  $I(\mu, \sigma) = 2/\sigma^2$ .

证明: 若  $\sigma$  已知, 则  $f(y) = \frac{1}{\sigma} \phi(y/\sigma)$ ; 若  $\mu$  已知, 则  $f(y) = \frac{1}{\sigma} \phi(y/\sigma)$ ; 若  $\mu, \sigma$  均未知, 则  $f(y) = \frac{1}{\sigma} \phi(y/\sigma)$ .

证明记  $r = y/\sigma$ , 则  $f$  服从标准分布  $P(0, 1)$  (见第一章), 其分布

与参数  $\theta = (\mu, \sigma)$  无关. 由定义可得  $L(\theta, y) = \log f(y) - \log a$ ,  $y = \sigma r$ .

$$a = \sigma$$

都未知,

$$dL/d\theta = -1/\sigma^2$$

$$d^2L/d\theta^2 = -2/\sigma^3$$

由于  $y = \sigma r$ , 因此所有涉及到  $y$  的期望、方差都与参数  $\theta$  无关. 所以 若  $\sigma$  已知, 则有  $I(\mu) = \text{Var}[dL/d\mu] = 1/\sigma^2$ ; 若  $\mu$  已知, 则有  $I(\sigma) = \text{Var}[dL/d\sigma] = 2/\sigma^2$ ; 若  $\mu, \sigma$  都未知, 则有  $I(\mu, \sigma) = 2/\sigma^2$ . 这些结果可进一步推广到独立同分布的样本以及更一般的样本. |

例2.3.4 设  $r \sim$  多元正态分布  $N(0, A)$ , 其中  $A$  为  $a$  阶单位矩阵. (1) 求  $J$  关于  $\theta$  的 Fisher 信息阵;

(2) 设  $e = g(p)$ , 即  $r = V(g(j), Z_n)$ , 求  $y$  关于  $\theta$  的 Fisher 信息阵.

解 (1) 设  $r = (y_1, \dots, y_a)^T$ ,  $\theta = (\theta_1, \dots, \theta_a)^T$ , 则  $y = (\theta_1, \dots, \theta_a)^T$ ,  $2$ ,

$$U_0(y) =$$

$$2(y^T A^{-1} y - y^T \log(2\pi A)), \text{ 盖 } A = (\sigma_{ij})$$

么  $:-\perp [Z_l \times L_y + 1i.$

## 2.3 分布族的信息函数

75

0,

1,  $i \neq j$ ,

所以  $K(0)$  为单位矩阵. (2) 记  $G = dg(j8)/dT$ , 由变换公式 (2.3.2) 可得

例 2.3.5 指数族分布  $X \sim f(x, \theta) = h(x) \exp\{T(x)\theta - b(\theta)\}$  (1) 求  $X$  关于  $\theta$  的 Fisher 信息阵;

(2) 记  $J = E\{r(X)r(X)^T\}$ , 求  $J$  关于  $\theta$  的 Fisher 信息阵.

解 (1) 其对数似然函数及其导数可表示为  $L(\theta, x) = \theta^T T(x) - b(\theta) + \log h(x)$ ,

告:  $TM = (7) - E[T(X)] \cdot d\theta/d\theta$

因此由 Fisher 信息的定义以及指数族分布的性质可得  $J(\theta) = E\{T(X)T(X)^T - (T(X) - E\{T(X)\})(T(X) - E\{T(X)\})^T\}$

$d^2 b(\theta)/d\theta^2$

$= \text{Cov}(r, r; \theta)$

因此  $J(\theta) = \text{Var} \{r(X)\}$ .

<

$\therefore z = K_n F$

$= \text{Cov}(y, y) = S_{yy} =$

(2) 根据指数族分布的性质以及变换公式 (2.3.2) 可得  $VJ =$

$E\{J^T W\} = J(\theta)$ , 令  $\theta = \theta(\eta)$ ,  $d\theta/d\eta = [J(\eta)]^{-1/2}$ ,

$= d\theta/d\eta$

(8)  $\square(\eta)$

$0 = b'(\eta)$ .

## 2.3.2 Kullback - Leibler 信息(K-L 距离)和 Jensen 不等式

Kullback - Leibler 信息反映了两个密度函数之间的差异, 是一种

“互信息”, 它具有信息和距离的某些性质, 也称为 Kullback - Leibler 距离(K-L 距离), 下面将摘要予以介绍. 另外, 本小节介绍的 Jensen 不

等式在文献中经常用到.

定义 2.3.3 密度函数  $f$  与  $g$  之间的 Kullback-Leibler 信息定

义为

$J(f, g)$

$T'(\eta) = (d\eta/d\theta)^T$

$1 = \square(\eta) - 1'$

## 76 第二章充分统计量与样本信息

$K(f, g) = E\{f(X) \log \frac{f(X)}{g(X)}\}$ ,

其中  $E$  表示对  $f$  求期望 (积分). 特别, 若  $\{f(x, \theta) | \theta \in \Theta\}$ ,  $\theta \in \Theta$ , 则 Kullback-Leibler 信息定义为

$E\{f \log \frac{f}{g}\} = E\{J(\eta, J)\}$  (2.3.7)

其中  $E$  表示对  $f(x, \theta)$  求期望; 对于统计量  $T = T(X)$  则其 Kullback - Leibler 信息定义为

$KT(\theta, \eta)$

注意, 一般以下主要讨论参数分布族, 并简称 为 K-L 距离.

例 2.3.6 对指数族分布  $X \sim h(x) \exp\{T(x)\theta - b(\theta)\}$ , 求  $K(\theta_1, \theta_2)$ . 解根据以上定义以及指数族分布的性质有

从②,  $\pi) = E, (10), X) - L(*, J)]$ ,  $L(\theta, X) = 0J T(x) - b(\theta) + \log h(x)$ ,  
 $K(e, \langle p) = E, |(\wedge - \wedge) t T(X) - [6(4) - 6(4)] I$   
 $= (3 - (p) T b(0) - [6(汐) - 6(妒)]$ . | 为证明K-L距离的性质, 首先介绍三个引理. 事实上, 以下引理 2. 3. 2 的Jensen不等式以及引理2. 3. 3 的信息不等式, 它们本身在统计

中都有非常广泛的应用.

引理2.3.2 (Jensen不等式) 若  $f(\cdot)$  为凸函数, 并假定有关的期望存在, 则有

$E[f(X)] \geq f(E(X))$ . 若  $f(\cdot)$  严凸, 则以上不等式中等号成立的充要条件为;  $r$  服从退化分布.

证明 由  $f(X)$  的凸性可知, 对  $v \in [0, 1]$ , 存在常向量  $C$ , 使得对任意的  $X$ , 有  
 $f(\wedge) \leq v f(\wedge_0) + (1-v) f(X) + C T(X - X_0)$ .

特别取  $x_0 = EX$ , 则有

$f(X) \leq v f(EX) + (1-v) f(X) + C T(X - EX)$ . 两边求期望即得  $E[f(X)] \leq v f(EX) + (1-v) E[f(X)] + C T(E(X) - EX)$ . 若  $f(\cdot)$  严凸, 则对  $\forall x \neq x_0$  (即  $x \neq EX$ ), 以上两个不等式中严格不等号成立; 因此当且仅当  $X = EX$  (a.e.) 时, 等号才能成立.

所以不等式中等号成立的充要条件为;  $r$  服从退化分布. |

推论1 取  $f(x) = -\log(x)$ , 1 或  $f(x) = -\log(x)$  则有

2.3分布族的信息函数

77

$E(f) \geq (E1V,$

$E[-\log K] \geq -\log(EK)$ . (2.3.8)

推论2 (条件Jensen不等式) 以上结果用于  $X | T$  的条件分布, 则有

$E[f(X) | T] \geq f(E(X | T))$ . 若  $f(\cdot)$  严凸, 则等式成立的充要条件为  $X | T$  服从退化分布, 即  $1 =$

$\langle p(r) \text{ (a.e.)}$ . 特别有

$E[-\log K | T] \geq -\log[E(y | T)]$ . (2.3.9)

上式等式成立的充要条件为  $K = \langle p(H \text{ (a.e.)}$ .

引理2.3.3 (信息不等式) 对密度函数  $f(x), g(x)$ , 假定有关的期望存在; 则有

$\int [\log f(x) / g(x)] f(x) dx \geq \int [\log g(u) / f(u)] g(u) du$ . (2. 3. 10)

该不等式中, 当且仅当  $f(x) = g(x)$  (a.e.) 时等式成立. 特别有

(2. 3. 11) 当且仅当  $e = \langle p$  时等式成立. 另外, 该式等价于  $EJL(U) \geq$

$E, [如, 川.$

证明要证(2.3. 10)式, 只需证

$r / (x)$

该式等价于

$- /$

记  $Y = g(X) / f(X)$ , 则由(2.3.8)式有

$Ey[-\log y] \geq -\log[E, r] = -\log$

此即(2.3.12)式, 该式倒推即得(2.3. 10)式. 上式等式成立的充要条件是  $Y = g(X) / f(X)$  服从退化分布, 即  $g(x) / f(x) = c$ ,  $g(x) = cf(x)$  (a.e.), 因此必  $\int c f(x) dx = 1$ , 即  $g(x) = f(x)$  (a.e.). 另外, 由 (2.3. 10) 式即得(2.3. 11)式. ■

Kullback-Leibler信息有以下类似于距离或信息的基本性质. (1)  $K(\theta, \langle p) \geq 0$ , 当且仅当  $\theta = \langle p$  时  $K(\theta, \langle p) = 0$ ;

(2) 独立,  $J = (X^1, \dots, X^J T)$ , 则有

$K(x(\theta) \langle p) = \sum K(x | \theta, \langle p)$ ;

為0,  

$$\int (\hat{p})^n d\mu(x) = 0.$$

[=  
 (2. 3. 12)  

$$-\log 1 = 0.$$

78 第二章充分统计量与样本信息

(3)若 $T = T(X)$ 为辅助统计量, 则  $\theta = 0$ ; 若 $r = T(X)$ 为充分统计量, 则 $K(\theta, W$   
 $= K_X(\theta, p_Y$

证明(1)由定义(2.3.3)以及信息不等式(2.3.11)有 以 $\theta, \langle p \rangle = E J L(\hat{p}, X)$   
 $- L(\langle p, X \rangle]$

$= E J f(\hat{p}, X)] - E, [L(\hat{p}, X)] \geq 0$ , 而且等号成立的充要条件为 $\theta = \langle p$ .

(2)直接由定义可得

$K_X(\theta, (p, =$

$n, (u)$

$\log i=1 n$

$Z K_X(\theta, W. is I$

$n$ 狀 #)

is

(3)若 $T = T(X)$ 为辅助统计量, 则其分布与参数 $\theta$ 和 $p$ 无关, 显

然有 $K_T(e_i(p) = 0$ . 若 $r = r(x)$ 为充分统计量, 则由因子分解定理可得 (10), 汐)叫  
 (10)微}

上面介绍的3个性质中, (1)反映了距离的性质; 而(2)和(3)反映了信息的统计意义, 这与  
 Fisher信息十分相似. 以下定理2.3.2将要进一步证明: 对于任何统计量 $T=T(X)$ 都有

$K_X(\theta, \dots K_T(\theta, \text{平}),$  而且等

式成立的充要条件是 $T = T(X)$ 为充分统计量; 这与Fisher信息的性质 十分相似. 为了证明  
 这一定理, 也需要类似于引理2. 3.1的一个引理.

引理2.3.4设 $X \sim f_H T(xy-g(t; \theta))$ 为正则分布族, 假定有 关的期望存在, 则有  
 1

$T(X) = t$  证明可根据条件期望的定义, (2.3.4)式加以证明. 记

$U). | T(X)$

$\mu(xX(x).$

(2.3. 13)

由条件期望 (2.3. 14)

要证(2. 3. 13)式, 即要证 $z_n(z) = E$ , 的定义(2.3.4)式, 等价于要证

(\_ 根据以上定义, 等价于要证

等价于要证

习题二

79

该式相当于而由 $P; (B)$ 定义可知上式相等; 逐步倒推, 即可得到(2.3.14)式, 因此

(2.3.13)式成立. I 定理2.3.2 设 $X \sim$ ,  $T=T(X)$  都是正则

分布族, 则有

且等式成立的充二是要条件是 $T = T(X)$ 为充分统计量.

$K_X(e, (p) e, (p),$

证明记 $y$

应用条件期望, 上式可表示为

$K_X(\theta, \text{岬}) = E, (E J (-\log y) \setminus T] \setminus. (2.3.15)$

对  $-\log Y$  应用条件 Jensen 不等式 (见 (2.3.9) 式) 得  $E(-\log Y) \geq -\log[E(Y)]$

代入以上 (2.3.15) 式, 并应用引理 2.3.4 的 (2.3.13) 式可得  $E(g^7) \leq E\{g^7\}$

若以上等式成立, 则由引理 2.3.2 的推论 2 可知  $Y \sim T$  服从退化分布, 即该式仅与  $t$  有关. 而对任意的  $(P = p_0, g(\cdot, \cdot), A(x) = \cdot / (\cdot))$ , 可得  $=g(t_0)h(x)$ . 由因子分解定理知  $r$  为充分统计量.

反之, 若  $r$  为充分统计量, 由前面的性质 (3) 知  $KT(e, \mu) : KX(\theta, \mu)$ .

## 习题二

1. 设  $\dots$  为 i.i.d. 样本,  $J \sim r(A, p)$ , 但  $p$  已知. 根据定义直接证明  $T =$

$A$  是充分统计量.

证明:

则有

$= E[-\log g(\cdot)] = E[-\log y]$ .

2. 设  $x', \dots, x_n$  为 i.i.d. 样本,  $g \sim \mathcal{N}(\mu, \sigma^2)$ . 试用定义直接

## 80 第二章充分统计量与样本信息

(1) 当  $m=0$  时,  $r = 2X$  是充分统计量;

(2) 当  $A=1$  时,  $r_2 = X(1)$  是充分统计量; 但  $r$  不是充分统计量; (3) 当  $M, A$  都未知时,  $T_3 = (T_2, T_1)$  是充分统计量.

3. 设  $X, \dots, X_n$  为 i.i.d. 样本,  $X$  服从以下分布, 求相应的充分统计量: (1) 负二项分布:  $X \sim \text{VB}(\theta, r)$ ,  $r$  已知; (2) 离散均匀分布:  $X, m$  未知; (3) 厂分布:  $X \sim r(A, p)$ ; (4) 冷分布:  $\text{BE}(p, f)$

$q$ ; (5) 对数正态分布:  $X \sim \text{LA}(\mu, \sigma^2)$ ; (6) Rayleigh 分布:  $X \sim \text{RA}(A, 2)$ .

4. 设  $X, \dots, X_n$  为 i.i.d. 样本,  $X$  服从以下分布, 求相应的极小充分统计量: I

(1)  $X \sim \text{PR}(a, \theta)$ , Pareto 分布  $a\theta x^{a+i} I_{\{x \geq \theta\}}$ ;

(2)  $X \sim \text{rPF}(\theta, c)$ , 幂函数分布  $\theta \leq x \leq c$ ;

II;

(3)  $X \sim \text{BE}(a, a)$ ,  $a > 0$ , 芦分布.

4.5. 设  $\dots, X_n$  为 i.i.d. 样本,  $X_j \sim \text{V}(\lambda_j; \lambda) = \exp\{-\lambda_j x_j\}$

其中  $\theta = (AC, (r) e = R \times (\theta, +\infty))$ . 试证:  $(\lambda_j; \lambda) = \lambda^j$

为指数

分布族, 并求极小充分统计量,

6. 设  $X, \dots, X_n$  为 i.i.d. 样本,  $X \sim \text{V}(a + pm, a^2)$ , 其中  $\theta <$

$< +\infty, -\infty < a, < +\infty$  是未知参数.  $m, i = 1, \dots, n$  是已知常数. 试求其极小充分统计量.

7. 设  $X, \dots, X_n$  为 i.i.d. 样本,  $X \sim \text{V}(\lambda; \lambda) = h(x)/I^*k$

$2 \exp\{-(a - \lambda x)\}$ , 其中  $-\infty < e \leq e^2 < +\infty$ . 求其极小充分统计量.

统计量.

8. 设  $X, \dots, X_n$  为 i.i.d. 样本, 当  $\lambda$  服从以下分布时, 求相应的

完备充分统计量:

(1)  $X \sim \text{f}(x, \lambda) = 2x \cdot \theta \sim 2I$ ;

(2)  $X \sim \text{f}(x, \lambda) = \theta x \sim 2I$ ;



(3)  $X_j$  服从 Rayleigh 分布 / (即,  $f(x) = 2x e^{-x^2/2}$  ( $x > 0$ ); (4)  $X$  服从 Laplace 分布  $f(x) = e^{-|x|}/2$ .

2a

9. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X \sim \text{Pareto}$  分布 / ( $x; a, 6$ ) =  $a x^a e^{-x}$ ;  $\{a+1\} I\{x\} > 0$  在以下情形下, 求相应的完备充分统计量: (1) 设

习题二

81

已知但  $a$  未知; (2)  $\theta$  已知但  $a$  未知; (3)  $\theta$  和  $a$  都未知. 10. (1) 设  $f(x; \theta) = c(\theta) e^{-\theta x}$  ( $x > 0$ ), 其中  $c(\theta) = \theta$ .

...,  $x_n$  是来自  $f(x; \theta)$  的  $Z$  样本. 求其极小充分统计量; 若  $A$  (或

么) 已知, 证明相应极小充分统计量的完备性. (2) 设  $g(x; \theta) = Z e^{-\theta x}$  ( $x > 0$ ),

$V \in \mathbb{R}$ , 其中  $D(3) =$

$\int_0^\infty f(x) dx < +\infty$  为正则化因子. 若 ..., 是来自  $g(x; \theta)$  的

$Z$  样本. 求极小充分统计量; 并证明其完备性. 11. 证明分布族或统计量的完备性有以下性质:

(1) 若  $T$  完备, 则  $T(X)$  完备, 但反之不真, 其中  $T$  和  $T(X)$  分别为  $X$  和  $T(X)$  的分布族.

(2) 若  $T(X)$  完备, 则  $S(X) = T(X)$  完备, 但反之不真; (3) 若  $T$  完备, 则  $r$  完备, 但反之不真.

\* 12. 判断下列三个命题的正误, 正确的请证明, 错误的给出反例. (1) 当极小充分统计量存在时, 完备的充分统计量一定是极小充分统计量;

(2) 极小充分统计量一定是完备的充分统计量;

(3) 当完备的充分统计量存在时, 极小充分统计量一定是完备的

充分统计量 (提示: 设  $T$  是完备充分统计量,  $r$  是极小充分统计量, 则存在一函数  $h(\cdot)$  使得  $T = h(r)$ ).

13. 证明冷分布族  $\{f(x; a, b) : a > 0, b > 0\}$  的完备性.

14. 设  $X_1, X_2, \dots, X_n$  为相互独立的样本,  $X_i \sim N(0, \sigma^2)$  ( $i = 1, \dots, n$ ), 而  $X \sim N(\mu, \sigma^2)$  ( $\mu \in \mathbb{R}$ ), 求该分布族完备的极小充分统计量, 其中  $\mu \in \mathbb{R}$ ,  $\sigma^2 > 0$ ,  $a, b$  都是未知参数.

\* 15. (1) 试证 Poisson 分布族  $\{P(A) : A \in (0, +\infty)\}$  的完备性 (提示: 利用函数项级数的性质).

(2) 设  $X_1, \dots, X_n$  独立,  $X_i \sim \text{Poisson}(\lambda_i)$  ( $i = 1, \dots, n$ ) 已知, 求完备的极小充分统计量.

16. 设  $x_1, \dots, x_n$  为 i.i.d. 样本,  $x_i \sim N(\mu, \sigma^2)$  ( $\mu > 0$ ), 求极小充分

统计量, 并判别其完备性. 17. 设  $(X_1, \dots, X_n)$  为 i.i.d. 样本

没), 其中  $\rho = \text{Cov}(X_j, X_k)$  ( $j, k = 1, \dots, n$ ). (1) 求极小充分统计量, 并判别其完备性;  $\sim N(0, \sigma^2)$  ( $\sigma^2 > 0$ );

82

第二章充分统计量与样本信息

(2) 证明:

$i = 1, \dots, n$

和  $r^2 = 2K$  都是辅助统计量, 但  $(T; r^2)$  不

是辅助统计量.

18. 设  $X_1, \dots, X_n$  为 i.i.d. 样本;  $Y_1, \dots, Y_m$  为 i.i.d. 样本, 且两总

体独立.在以下情形下,求相应的完备充分统计量:

(1)  $a + rW, i, vt$  (2)  $x\{y,$

$+r(a_2 - i);$  氏);

(3)  $X, \sim 7V(M, y), Y.$

•19.设 $X_1, \dots, X_n$ 为i. i. d.样本,  $X_i \sim R(0, \theta)$ , 若参数空间为 $\theta =$   
但 $\theta = p$ 已知.

$[a, \infty) (a > 0)$ , 证(1)明 $T = X(n)$ 是极小充分统计量, 但不是完备的充分

$< a, \infty$ , 然后证 $\lim_{n \rightarrow \infty} P_n = 0$ .

$(n) = 0$ .

\*20.设 $X_1, \dots, X_n$ 为i. i. d.样本,  $X_i \sim R(0, 36)$ ,  $-\infty < x < \infty$ . 证

明 $r = (x(1), A_r(n))$ 是极小充分统计量, 但不是完备的充分统计量(提示: 证明 $z^{x(1)}/x(n)$ 的分布与 $\theta$ 无关).

21. 设 $X_1, \dots, X_n$ 为i. i. d.样本, 证明以下独立性: (1)若 $X_i \sim \Gamma(\alpha, \lambda)$ ,

$-\infty < Y < \infty$ , 则 $(X(1) - X(1))/$

$(X(4) - X(1)), i=2, \dots, n-1$ 与 $(X(1), V(4))$ 独立;

(2) 若 $X \sim r(A, v)$ ,  $A, p$ 在参数空间中任意, 则 $T = f(A)$ 与 $i=1$

统计量(提示: 取 $h$

$\log x_i$ )

$n$

$2 \log x_i - \log X(1)]$ 独立;  $i=1$

(3) 若 $X \sim M(r(A, 1))$ , 则 $Z$ 独立.

$XY$

22. 设 $X_1, \dots, X_n$ 为i. i. d.样本,  $S^2 =$

$\frac{1}{n} \sum_{i=1}^n X_i^2$ . 证明  $i=1$

$(X_i - \bar{X})^2 / (n-1)$ ,  $T =$

(1) 若 $X \sim N(0, a^2)$ , 则 $11 \text{ IF}(\lambda) = n + i/S$ 与 $r$ 独立;

(2) 若 $X \sim V(x, a^2)$ , 则 $0(X, 5)$ 与 $|X - M_j/S|$ 独立, 其中 $M_j$ 为样本中位数.

23. 设 $X_1, \dots, X_n$ 为i. i. d.样本,  $X_i \sim \Gamma(\alpha, \lambda)$ ,  $\dots$ 为i. i. d.样

$= \sum_{i=1}^n X_i \sim (\theta, i=1, \dots, n) \sim (n-1)$

,  $\dots$

,  $n-2$ 与

$J=1$

习题二

83

与 $r$ 独立;

(2)  $r(x, n) = m$

$t', n-1$

24. 设

为i. i. d.样本,  $X, V(\%1; a) = -$ (位

$m, d)$ , 且两总体独立. 记 $r$ 为 $\dots$ , 无,  $?$ .

本,  $y, 4(M_2,$

$\sum_{i=1}^n X_i (1) F(X^Y) = \frac{1-P}{2} / (m-1) / \frac{1}{2} (1 - X)^2 / (n-1)$

$\alpha x = r^2$ 时, 证明:

$-$

$\sum_{i=1}^n L_i = i$

与 $r$ 独立.

$n \rightarrow \infty$  时

置尺度参数族), 其中  $g(\theta) > 0$  且  $g'(\theta)$  存在 ( $\theta \in R$ ),  $\theta = (M, \sigma^2) \in R \times (0, +\infty)$ . 证明: 其 Fisher 信息阵为

25. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim \text{Poisson}$  分布  $P(A)$ .

(1) 求样本关于参数  $A$  和  $A^2$  的 Fisher 信息;

(2) 求  $I(A)$ , 使  $I(A)$  与参数无关.

26. 设  $X_1, X_2, \dots, X_n$  为相互独立的样本, 在下列情形下求样本关于相应参数的 Fisher 信息阵:

(1)  $X_i \sim P(A)$ , 一切  $i$ , 而  $X_i \sim P(yA)$ ; (2)  $X_i \sim r(l/a, l)$ , 一切  $j$  第  $i$ , 而  $m_j/a, l$ ,

其中  $y > 0$ ,  $A > 0$ ,  $\langle r \rangle > 0$  都是未知参数.

27. 设  $X_1, X_2, \dots, X_n$  为相互独立的样本, 在下列情形下求样本关于相应参数的 Fisher 信息阵:

(1)  $X_i \sim N(0, \sigma^2)$ , 一切  $j$  第  $i$ , 而  $X_i \sim N(0, \sigma^2)$ ; (2)  $X_i \sim N(0, \sigma^2)$ , 一切  $j$  第  $i$ , 而  $X_i \sim N(\mu, \sigma^2)$ ;

(3)  $X_i \sim N(0, \sigma^2)$ , 一切  $n$ , 而  $X_i \sim N(y, \sigma^2)$ ,  $X_n \sim N(0, \sigma^2)$ ,

其中  $y \in R$ ,  $\sigma^2 > 0$ ,  $\sigma^2$  都是未知参数.

• 28. 设  $X \sim N(\theta, \sigma^2)$ ,  $\theta \in R$ , 假设  $\int_{-\infty}^{\infty} f(x) dx = 1$  和  $E(X) = a(\theta)$  且

当  $\theta = 1$  时

#### 84 第二章充分统计量与样本信息

可在积分号下关于  $\theta$  求导数, 并记为  $a(\theta)$ ; 又记  $\text{Var}(X)$  证明: 其 Fisher 信息  $I(\theta)$  满足不等式  $I(\theta) \geq a^2(\theta)/\sigma^2(\theta)$ ; 并且等号成立的充要条件为  $X$  服从指数族分布, 即  $f(x, \theta) = \frac{1}{Z(\theta)} \exp \{xQ(\theta) - b(\theta)\}$  (提示: 利用 Schwarz 不等式和 score 函数  $S(\theta, x)$  与密度函数  $f(x, \theta)$ ,

$f(x, \theta)$  的关系, 即  $S(\theta, x) = d \log f(x, \theta) / d\theta$ ).

#### 第三章点估计基本方法

通过样本  $x = (x_1, \dots, x_n)^T$  对总体的分布函数或与分布函数有关的

量进行统计推断, 通常有两种方法: 一是参数化的方法, 即假定分布函数 (或分布密度) 形式已知, 但含有未知参数, 如正态分布  $N(M, \sigma^2)$ , Poisson 分布  $P(A)$  等. 一种常见的情况是: 假设样本  $X = (X_1, \dots, X_n)^T$  的分量独立同分布, 且  $X_i \sim F(x, \theta)$ ,  $\theta > 0$ . 这时就转化为对未知参数  $\theta$  或其函数  $g(\theta)$  进行推断. 诸如要推断的  $g(\theta)$  为均值  $E(X)$ ; 方差  $\text{Var}(X)$ ; 概率  $p(\theta)$

(或已知); 分位数  $x_p(\theta) = F^{-1}(p)$  等. 另外还有非参数方法, 即对分布函数 (或分布密度) 不作特别假定, 而直接从样本出发对其进行统计推断, 限于篇幅, 本书将不介绍非参数统计推断方法. 参数统计推断通常有两个基本问题, 即参数估计和假设检验, 后者将在第六章予以介绍; 而参数估计又可分为点估计和区间估计两种. 由于区间估计与假设检验有密切关系, 我们将在第七章介绍区间估计; 而在第 3-5 章介绍点估计及其相关问题. 本章首先介绍点估计的基本方法. 顾名思义, 点估计即按照某种优化准则构造一个统计量  $g(X)$ , 当有了样本观察值  $X = (x_1, \dots, x_n)^T$  时, 就用  $g(X)$  作为对未知量  $g(\theta)$  的估计. 例如, 某种砖块的强度  $X$  服从对数正态分布:  $X \sim \text{ZJN}(\mu, \sigma^2)$ , 工程上要求

$P(X \geq 120 \text{ kg}) \geq 0.95$  (或 0.99 等). 可通过抽样对

$\mu$  进行推断, 得到点估计  $\hat{\mu}$ . 若  $\hat{\mu} \geq 0.95$ , 则认为这批砖合格, 否则认为不合格. 类似的问题很多, 因此参数估计是参数统计推断的重要组成部分.

第3.1节首先介绍统计判决函数，这是分析研究统计问题常用的基本观点，可应用于许多统计问题；然后于第3.2-3.4节介绍基本的点估计方法，其中包括一致最小风险无偏估计(UMRUE)、极大似然估计(MLE)以及矩方程估计(MEE)；对于极大似然估计，还介绍了不变原理以及子集参数的似然等内容。此后，第四章将进一步介绍最优同变估计(MREE)，第八章介绍Bayes估计。

86

### 第三章点估计基本方法

#### 3.1统计判决函数

统计判决函数理论是Wald于1950年提出来的，其最初的目的是想建立一套完整的理论，把各种形式的统计问题都归结到该理论中，在统计判决的观点下，应用最优化方法进行统一处理。虽然与其预期的目标有相当距离，但是，统计判决函数的观点已经渗透到统计学的许多领域，对统计学的发展产生了相当大的影响。因此，统计判决的基本内容是学习数理统计的读者所必须掌握的。

##### 3.1.1统计判决三要素

(1) 样本空间和分布族。

即  $I$ ，或  $\mathcal{I}$  这是讨论统计问题的基本空间，前两章已多次用到。

(2) 判决空间和判决函数。记为  $(\mathcal{D}, \mathcal{F})$ ，其中  $\mathcal{D}$  称为判决空间， $\mathcal{F}$  是  $\mathcal{D}$  上的Borel域(今后很

少用到)。任一  $d \in \mathcal{D}$  称为一个判决，表示统计问题的一个解。通常  $\mathcal{D}$  总是样本  $X = x$  的函数  $d = d(x)$ ，称为统计判决函数，简称判决函数。严格来讲，统计判决函数  $d$  应该定义为  $\mathcal{U}$  上的可测函数， $d(X)$  则是一个统计量。例如：假定  $X_1, \dots, X_n$  为  $i.i.d.$  样本， $X \sim N(\theta, 1)$ 。则可考虑以下问题：

i) 点估计问题。这时判决就是数轴上的一个点，判决空间  $\mathcal{D} = (-\infty, +\infty)$ ， $d = d(x)$ 。通常  $d$  总是样本  $X = x$  的函数  $d = d(x)$ 。统计判决函数，例如可取  $d = d(x) = \bar{x}$  等。至于统计判决函数  $d(x)$  的优劣问题，可参见下一小节。

ii) 区间估计问题。这时判决就是一个区间  $d = [a_1, a_2]$ ，判决空间  $\mathcal{D}$  为集合：  
 $\mathcal{D} = \{[a_1, a_2] : a_1 \leq a_2, -\infty < a_1, a_2 < +\infty\}$ 。统计判决函数为

$d(x) = [a_1(x), a_2(x)]$ ，例如可取  $d(x) = [\bar{x} - t_{\alpha/2}(n)/\sqrt{n}, \bar{x} + t_{\alpha/2}(n)/\sqrt{n}]$ ，其中  $t_{\alpha/2}(n)$  为  $i.i.d.$  分布  $z(n)$  的  $\alpha/2$  分位数。

iii) 假设检验问题。 $H_0: \theta \leq \theta_0$  这时判决空间由两个点组成，即  $\mathcal{D} = \{0, 1\}$  其中判决  $d = 0$  表示  $H_0$  成立，判决  $d = 1$  表示  $H_0$  不成立。

亦可取  $\mathcal{D} = (0, 1)$ ，其中判决  $d = 0$  表示  $H_0$  成立，判决  $d = 1$  表示  $H_0$  不成立。

(3) 损失函数和风险函数。

##### 3.1 统计判决函数

87

损失函数为定义在的正值函数  $L(e, d, \theta) \in \mathbb{R}^+$ ，它表示

在参数  $\theta$  下，采取判决  $d$  时给统计问题带来的损失。例如，对于  $\theta$  的点估计，常用的损失函数有  $L_2(\theta, d) = (d - \theta)^2$ ， $L_1(\theta, d) = |d - \theta|$  等。损失函数的选取依赖于理论与实际问题的需要，但通常都要求  $L(e, d)$  为  $d$  的凸函数(通常称为凸损失)。在统计推断中，希望损失  $L(e, d)$  尽量小。对于统计判决函数  $d(x)$ ，其相应的损失函数为  $L(\theta, d(x))$ 。由于  $L(\theta, d(x))$  为统计量，带有随机性，不便于比较损失的大小。因此很自然地取其期望作为比较标准，由此引出以下定义：

定义3.1.1 给定统计判决函数  $d$  和损失函数  $L(\theta, d)$ ，相应的风险函数定义为

$R(\theta, d) = E[L(\theta, d(X))] = \int L(\theta, d(x)) dP_\theta(x)$ 。

$(\theta, \delta)$  表示采取判决  $\delta$  时, 给统计问题带来的平均损失. ■ 例如在估计问题中, 用  $\delta$  估计  $g(\theta)$ . 若取均方损失  $L(\delta, d) = (d - g(\theta))^2$ , 则其相应的风险函数为  $R(\delta, \theta) = E[S(X) - g(\theta)]^2$ , 即为均方误差; 若取绝对损失  $L(\delta, d) = |d - g(\theta)|$ , 则其相应的风险函数为  $R(\delta, \theta) = E|S(X) - g(\theta)|$ , 即为平均绝对误差. 根据统计判决观点, 统计推断所追求的目标就是对于给定的损失函

数  $L(\delta, d)$ , 希望求出统计判决函数  $\delta(x)$ , 使其风险函数  $R(\delta, \theta)$  尽可能小. 显然这一目标是很合理的.

例 3.1.1 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $X_i \sim N(\theta, \sigma^2)$ , 要估计  $\theta$ , 这时  $P = (-\infty, +\infty)$ . 今取均方损失  $L(\delta, \theta) = (S - \theta)^2$ , 并考虑  $\sigma^2$  的估

计及其相应的风险函数. 若取  $\delta(\cdot) = (n-1)^{-1} \sum_{i=1}^n X_i$ , 由于  $\sum_{i=1}^n (X_i - \theta)^2 \sim \chi^2_{n-1}$ , 则有

$R(\delta, \theta) = E[S^2(X) - \theta^2] = \frac{\sigma^2}{n-1}$

$n$   
 $i=1$

$R(\delta, \theta) = E[S^2(X) - \theta^2] = \frac{\sigma^2}{n-1} < R(\delta_0, \theta)$ .  $n+1$

以上结果说明, 对任意的  $\theta \in \Theta$ ,  $\delta(x)$  一致地优于  $\delta_0(x)$ . | 3.1.2 统计判决函数的优良性准则

若取  $S^2(X) = (n+1)^{-1} \sum_{i=1}^n (X_i - \bar{X})^2$ , 则有

$(X_i - \bar{X})^2$ , 则有

对于分布族  $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ , 给定一个统计问题以后, 即

### 第三章点估计基本方法

可确定判决空间同时选取一个合适的损失函数  $L(\delta, d)$ . 这时, 就可以根据风险函数的大小来判断一个判决函数  $\delta$  的优

良性. 具体来讲, 可按照实际需要与可能分为以下几种情形. (1) 一致最优性.

定义 3.1.2 若存在  $\delta^*(x)$  使  $R(\delta^*, \theta) \leq R(\delta, \theta)$  对一切  $\theta \in \Theta$  成立, 则称  $\delta^*(\cdot)$  一致优于或等同于  $\delta(\cdot)$ . 若至少存在一个  $\theta_0 \in \Theta$ , 使  $R(\delta^*, \theta_0) < R(\delta, \theta_0)$ , 则称  $\delta^*(x)$  一致优于  $\delta(x)$ . |

显然, 若  $\delta^*$  一致优于  $\delta$ , 则不必再考虑  $\delta$ . 我们的目标就是设法在判决空间  $\mathcal{P}$  中求最优的判决函数. 以下容许性的定义与一致最优性有密切关系.

定义 3.1.3 容许性. 给定  $L(\delta, d)$ , 对于判决  $\delta(x)$ , 若存在  $\delta^*(x)$  一致优于或等同于  $\delta^*$ , 则称  $\delta$  为不容许的 (因为  $\delta^*$  一致地比  $\delta(x)$  好), 若不存在这样的  $\delta^*$ , 则称  $\delta$  为容许的 (即不存在一致比  $\delta$  好的判决). |

在许多问题中, 要想在原有判决空间  $\mathcal{P}$  以及参数空间  $\Theta$  中求解一致最优的或容许的判决函数往往是困难的, 甚至是不可能的. 通常可放宽条件, 在一定范围内求最优解. 这实际上包含两个方面: 即对判决空间的性状加以限制; 或者对参数空间的性状加以限制.

i) 判决空间. 限制判决函数  $\delta$  的范围, 使  $A = \{\delta: \delta \text{ 满足一定条件}\}$ , 然后在  $A$  中求风险函数最小的  $S(x)$ . 例如可取  $A = \{\delta: \delta \text{ 为 } g(\theta) \text{ 的无偏估计}\}$  (见第 3.2 节); 或者取  $A = \{\delta: \delta \text{ 为 } g(\theta) \text{ 的同$

变估计}\} (见下一章) 等. 而显然有  $V = \{g(\theta) \text{ 的所有估计}\}$  或  $\mathcal{P}$ .

ii) 参数空间. 对判决函数  $\delta$  在参数空间  $\Theta$  中的性状进行一定限制. 这有以下两种常见的情形.

(2) Minimax 准则 (最大最小准则).

定义 3.1.4 设  $\delta$  的风险函数为  $R(\delta, \theta)$ , 它关于  $\theta$  的最大风险为  $M(\delta) = \sup_{\theta \in \Theta} R(\delta, \theta)$ ; 若对一切  $\delta$  有  $M(\delta^*) \leq M(\delta)$ , 则称  $\delta^*$  为统计问题关于损失函数  $L(\delta, d)$  的 Minimax 解 (即先关于  $\theta$  对  $R(\delta, \theta)$  求最大, 再关于  $\delta$  对  $M(\delta)$  求最小). 显而易见, Minimax 准则是比

较保守的，它使风险函数 $R(\theta, \delta)$ 先在参数空间中关于 $\theta$ 求最大；然后再在判决空间中关于 $\delta$ 求最小，因而

而这个解不一定最好。以下Bayes解是使风险函数先在参数空间中关于 $\theta$ 求加权平均；然后再在判决空间中关于 $\delta$ 求最小，因而更合理。

(3) Bayes 准则。

### 3.1 统计判决函数

89

定义3.1.5 若 $\theta$ 有分布 $\theta \sim f(\theta)$ ，并记  $R(S) = \int f(\theta) R(\theta, S) d\theta$ ，

称为Bayes风险。若存在 $\delta^*(\theta)$ ，对一切 $S \in \mathcal{C}$ ， $R(\delta^*(\theta), S) \leq R(\delta, S)$ ，则称 $\delta^*(\theta)$ 为统计问题关于损失函数 $L(\theta, d)$ 和分布 $f(\theta)$ 的Bayes解。

Minimax解和Bayes解都在某种程度上反映了判决函数在参数空间上的整体性质。关于Bayes统计，本书第八章将有更详细的介绍。从统计判决的观点来看，根据Bayes风险最小准则求解显然是合理的。事实

上，关于Bayes统计的争论近年来已经越来越少。3.1.3 Rao - Blackwell 定理

由第二章的讨论可知，充分统计量可以不损失信息地把 $n$ 维样本简化为维数很小统计量，以此为基础进行统计推断比直接从样本出发要简单方便得多。因此，统计推断方法一般都应当从充分统计量出发，通常称为充分性原则。以下Rao - Blackwell定理也印证了这一点，该定理说明：统计推断中的最优解通常都是充分统计量的函数。这个定理可用于一切统计判决问题，并不限于参数估计，对于假设检验以及其他统计问题都是适用的。

定理3.1.1 (Rao - Blackwell) 对于分布族 $\{P_\theta\}$ ，若 $L(\theta, d)$ 为统计判决问题的凸损失函数， $T = T(X)$ 为充分统计量， $\delta$ 为任一统计判决函数，则

$E[L(\theta, \delta) | T] \leq E[L(\theta, \delta^*) | T]$  必优于或等同于 $\delta^*(x)$ 。若 $L(\theta, d)$ 为 $\theta$ 的严凸函数，则 $\delta^*(x)$ 一致

优于 $\delta(x)$ ，而等同于 $\delta^*(x)$ 的充要条件为 $\delta^*(x)$ 是 $r(x)$ 的函数，即 $\delta^*(x) = A(T(x))$ 。

证明 首先，由于 $r(x)$ 为充分统计量，因而 $x | T$ 的分布与 $\theta$ 无关，所以 $d(x)$ 为统计量，且为充分统计量的函数。以下应用条件 Jensen不等式(引理2.3.2的推论2)证明该定理。由定义可得

$E[L(\theta, \delta) | T] = E[L(\theta, \delta^*) | T]$ ，而风险函数 $R(\theta, \delta)$ 可表示为  $R(\theta) = E[L(\theta, \delta)] = E[E[L(\theta, \delta) | T]]$ 。

由条件 Jensen 不等式  $E[L(\theta, \delta) | T] \geq L(\theta, E[\delta | T])$  可得  $E[L(\theta, \delta) | T] \geq E[L(\theta, \delta^*) | T]$ 。

因此有  $R(\theta) \geq E[E[L(\theta, \delta) | T]] \geq E[E[L(\theta, \delta^*) | T]] = E[E[L(\theta, \delta^*) | T]] = R(\theta)$ 。

### 90 第三章点估计基本方法

另外，由引理2.3.2的推论2可知，若 $L(\theta, d)$ 为 $\theta$ 的严凸函数，则以上等式成立的充要条件为 $\delta$ 的分布退化，即 $\delta(x) = \delta^*(x)$ 。Rao - Blackwell定理应用非常广泛，它说明，对任一统计判决，关

于充分统计量取条件期望以后可能会更优。许多统计判决问题的最优解（不一定是估计问题）都可由条件期望来表示。

### 3.2 无偏估计及其UMRUE和UMVUE

给定样本 $A, \dots, X_n$ ，通常记 $X = (X_1, \dots, X_n)$ ， $\theta \in \Theta$ 。要估计未知参数或者它的一个函数 $g(\theta)$ ，并给定凸损失函数 $L(\theta, d)$ 。

以下记 $\theta$ 的估计为 $A(\theta)$ 或 $S(X)$ ，记 $g(\theta)$ 的估计为 $i(A(\theta))$ 或 $g(X)$ 。本节把 $i(x)$ 的范围由 $p = g(\theta)$ 的一切估计缩小为

$A = U(\theta)$ 的一切无偏估计 $I = i(g(X)) : E i(g(X)) = g(\theta)$ 。

定义3.2.1 偏差与无偏估计。若 $g(\theta)$ 的估计为 $i(X)$ ，则称

$E i(X) = g(\theta)$  为无偏估计。

为以尤)的偏差.若对一切 $\theta$ ,  $\text{bias}[i(X)] = 0$ , 则称 $g(X)$ 为 $g(\theta)$ 的无偏估计(UE), 即  
 $E[g(X)] = g(\theta)$ . 例3.2.1若 $X_1, \dots, X_n$ 独立同分布, 且 $X_i \sim N(\theta, \sigma^2)$ , 即  
 $(4), \text{Var}(X_j) = \sigma^2$ .

假定它们存在, 则 $X$ 和 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$   
 计(对任何分布都成立).

证明 主要证 $E(S^2) = \sigma^2$ . 令 $K_i = X_i - \bar{X}$

$Y_i = (X_i - \bar{X})^2$ ,  $S^2 = \frac{1}{n-1} \sum_{i=1}^n Y_i$  由独立同分布假设可得

$E(Y_i) = E[(X_i - \bar{X})^2] = E(X_i^2) - 2\bar{X}E(X_i) + E(\bar{X}^2)$  再由独立性可得

$n$

$n-1$

$E(Y_i) = \sigma^2$  为 $\bar{X}$ 和 $\bar{X}^2$ 的无偏估计

$E(S^2)$

$n$

$\sigma^2 = E(S^2)$

则有  $(Y_i - \sigma^2)$ .

$J$ ).

3. 2 无偏估计及其UMRUE和UMVUE 以下考虑与损失函数和风险函数有关的问题. 设 $g(X)$   
 为 $g(\theta)$ 的估计

$E[\text{MSE}(X)] = E[g(X) - g(\theta)]^2 = E[g(X) - E[g(X)] + E[g(X)] - g(\theta)]^2$

$= E[g(X) - E[g(X)]]^2 + [E[g(X)] - g(\theta)]^2 = \text{Var}(g(X)) + [\text{bias}(g(X))]^2$ .

即估计量的均方误差=方差+偏差<sup>2</sup>. 对于无偏估计类 $A$ , 则 $\text{MSE} = \text{方差}$ ; 因而估计量的均方误差  
 $\text{MSE}$ 最小的充要条件为方差最小.

定义3. 2.2 一致最小风险无偏估计(UMRUE)和一致最小方差无偏估计(UMVUE). 对于一般  
 凸损失函数 $L(\theta, d)$ , 若存在 $g(\theta)$ 的无偏估计 $i(X)$ , 使得对任何其他无偏估计有

$R(\theta, i(X)) \leq R(\theta, g(\theta))$  则称 $i(X)$ 为 $g(\theta)$ 的一致最小风险无偏估计(UMRUE).

对于均方误差,

若

$\text{Var}[i(X)] \leq \text{Var}[g(\theta)]$ , 则称 $i(X)$ 为 $g(\theta)$ 的一致最小方差无偏估计(UMVUE).

注意, 给定分布族, 无偏估计不一定存在. 例3.2.2设 $J \sim N(0, 1)$  则 $g(J) = J^2$ 不存在无偏估计.

计, 通常假定损失函数 $L(\theta, d)$ 为凸函数,  $i(X) = J^2$ . 特别, 若 $L(\theta, d) = (d - g(\theta))^2$ ,

其风险函数即为均方误差, 记为 $\text{MSE}$ , 这时有

解 若 $E[i(X)] = \sigma^2$

其他如,  $X_1, \dots, X_n$ 独立同分布,  $X_i \sim N(\theta, \sigma^2)$  若 $g(\theta) = \theta^2$  也不存在无偏估计.

3. 2. 2 Lehmann - Scheffe 定理

以下考虑如何求解一致最小风险无偏估计(UMRUE)或一致最小方差无偏估计(UMVUE). 统计  
 推断(无论是参数估计、假设检验或其他统计问题)的一个基本思想就是尽量从充分统计量出  
 发. 因为用充分统计量代替原有样本不损失任何信息, 而且通常比原有样本简单, 比非充分  
 统计量性能好, 以下例题说明了这一点.

$X \sim N(0, 1)$ , 则有

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

$\frac{1}{n} \sum_{i=1}^n X_i^2$  令则上式右端 $\rightarrow \infty$ , 而左端 $\rightarrow 0$ 有穷, 这是不可能的. ■

$E[g(X)] =$

$X = 0$

例3.2.3设 $A_1, \dots, A_n$ 为独立同分布样本,  $A_i \sim N(0, \sigma^2)$ , 求的

则

$=EJZ(\hat{\theta},$   
92

### 第三章点估计基本方法

无偏估计.

解 由于 $E(\bar{X})=\theta/2$ ,故可取 $i_1(X)=2\bar{X}$ ,这是一个无偏估计.

这时无不是充分统计量,其均方误差为 $Var(i_1)=\theta^2/3n$ .另外也可从充分统计量 $X(n)$ 出发,由于

$X(n) \sim B(n, \theta/2), E(X(n))=n\theta/2$ ,

因此可取 $i_2(X) = X(n)$ 为无偏估计.其均方误差为  $T^2r$

$MSE(g_2) = Var(g_2) = Var(X(n)) = n\theta(1-\theta/2) = n\theta/2(2-\theta)$

即基于充分统计量的无偏估计 $i_2(X)$ 一致优于 $i_1(X)$ . 上面的例子反映了一个一般的规律.即基于充分统计量的统计推断

通常要优于基于非充分统计量的统计推断.以下从完备充分统计量出发,通过两个引理逐步导出本节的主要定理,即若 $r(x)$ 为完备充分统计量,只要 $\hat{g}$ 为 $g$ 的无偏估计,则它必为一致最小风险无偏估计.

引理3.2.1 (唯一性) 设 $T=T(X)$ 为完备统计量(不必为充分统计量),若 $A(r(x))$ 和 $h(r(x))$ 皆为 $g(\theta)$ 的无偏估计,则必有 $E(T(X)) = E(h(T(X)))$ (a. e.).即若 $g$ 的无偏估计存在,且为 $T(X)$ 的函数,则必唯一(a. e.).

证明 由假设条件可得

$E(h(T(X))) = E(A(r(X))) = g(\theta) = E(h(T(X)))$ ,  $\forall \theta$ .

因此由完备性定义有 $E(h(T(X))) = E(h(T(X)))$ (a. e.).

注 以上证明中用到了 $E_j A(r(x)) = g(\theta)$ 对任 $\theta$ 都成

立,

即 $E(h(T(X)))$ 在 $\theta$ 上“处处无偏”;这也是无偏性的定义中所要求的.

引理3.2.2 (最优性) 设 $g(X)$ 为 $g(\theta)$ 的无偏估计,损失函数 $L(e, d)$ 为凸函数,风险函数为 $R(\theta, g)$ . 设 $T=T(X)$ 为充分统计量(不必为完备的),令

$=E_j g(X) | T] = (P(T(X)))$ , 则 $\hat{g}(T)$ 亦为 $g$ 的无偏估计,且 $\hat{g}(T)$ 优于或等同于 $g(X)$ ,即

$R(\theta, \hat{g}) \leq R(\theta, g)$ . 若 $L(\theta, d)$ 严凸,则上式等号成立的充要条件是: $\hat{g}(X)$ 为 $r(T)$ 的函数,

即 $\hat{g}(X) = h(T(X))$ . 证明由假设条件可知

### 3.2 无偏估计及其UMRUE和 UMVUE 93

$E_j i(X) = E_j [E_j g(X) | T] = E_j g(X) = g(\theta)$ . 因此,  $i(X)$ 为 $g(\theta)$ 的无偏估计,又由定理3.1.1 (即 Rao - Blackwell

定理)知其他结论成立.

推论 条件同上,若 $U \sim d$ , 则有

■

$Var(g) \geq Var(g)$ ,  $\forall \theta$ . 且等式成立的充分必要条件是 $g(X) = h(T(X))$ . I

定理3.2.1 (Lehmann-Scheff6) 给定样本 $X_1, \dots, X_n$  设 $X =$

(1) 设 $\hat{g}$ 为 $g$ 的无偏估计,且 $j(x)$ 为 $T(X)$ 的函数,即 $\hat{g}(X) = A(r(X))$ ,则 $\hat{g}(X)$ 必为 $g$ 的一致最小风险无偏估计.

(2) 设 $\hat{g}(X)$ 为 $g$ 的无偏估计,则 $\hat{g}(X) = E_j g(X) | T]$ 为 $g(\theta)$ 的一致最小风险无偏估计.

(3) 若 $L(\theta, d)$ 为严凸,且的一致最小风险无偏估计存在,则必为 $r=r(X)$ 的函数.

证明 (I) 任给 $g(a)$ 的无偏估计 $g(X)$ ,要证

$R(\theta, g) \geq R(\theta, \hat{g})$ ,  $\forall \theta$ .



,  
 $e^e$ . 考虑  $g(\theta)$  的无偏估计, 损失函数

(戈, ...

$a$ ) 为凸函数,  $t = t(x)$  为完备的充分统计量. 则有

令

$= E_j g(X) | T] = A(T(X))$ . 则由引理 3.2.2 知, 无偏, 且优于  $i(j)$ , 即  $R(e, g^*,$   
 $< R(e$

,

旁),  $VA$  又因  $s(x)$  和  $g(x)$  皆为  $g(\theta)$  的无偏估计, 且为  $r(\theta)$  的函数, 所以由引理 3.2.1 可知,  $i(X) = g(X) (a. e.)$ , 从而有

$= 7(6 >, g^*)^{7(6 < 9, g)}, V < 9, g$ .

(2) 由定义  $i(x) = E_j g(x) | r]$  为  $r$  的函数, 又由于

$= E_j E_j g(X) | r]] = E_j (g(X)) = g(4)$ ,

因此  $j$  为  $g(\theta)$  的无偏估计, 所以由 (1) 可知  $i(x)$  为  $g(\theta)$  的一致最小风险无偏估计.

(3) 设  $g(\theta)$  的一致最小风险无偏估计存在, 记为  $g(X)$ , 要证  $g(x)$  为  $r(\theta; r)$  的函数, 为此令

$g_D = A(n^*) = E_j i(x) | r = 7V]$ , 则由 Rao - Blackwell 定理知

$g_D = R(\theta, E_D(g(T))) = R(e, g)$ . 又由于  $i(x)$  为  $g(\theta)$  的一致最小风险无偏估计, 所以有  $R(e, g) \leq R(\theta,$

$g_D)$ . 因此以上等式成立. 由于  $L(e, d)$  为严凸, 因而  $i(X) | T$  为退化分布, 即

$i(X) = h(T(X))$ .  $\square$

#### 94 第三章点估计基本方法

推论 对于损失函数  $L_2\{e, d\} = (d - g(e))^2$  以上 (1) - (3) 都成立, 且称一致最小风险无偏估计为一致最小方差无偏估计 (这时  $R(\theta, g) = \text{Var}_j g(X)$ ). 对于损失函数  $L_2\{e, d\} = |d - g(e)|$ , 以上 (1)、(2) 成立.

#### 3.2.3 例题

以下通过若干例题, 说明如何根据 Lehmann - Scheffé 定理在各种常

见的分布族中求解  $g(\theta)$  的一致最小风险无偏估计. 当然, 其前提是假设完备充分统计量存在以及一致最小风险无偏估计存在. 根据 Lehmann - Scheffé 定理, 求解  $g(\theta)$  的 UMRUE 有两种方法:

(1) 直接方法. 即找一个完备充分统计量  $r(\theta; v)$  的函数使  $E_j(r(x)) = g(\theta)$ , 则  $i(x) = p(r(x))$  为  $g(\theta)$  的 UMRUE.

(2) 条件期望法. 即取一个完备充分统计量  $T$  以及  $g(\theta)$  的某一个无偏估计, 则  $i(X) = E_j g(X) | T$  为  $g(\theta)$  的 UMRUE. 这时, 关键问题就是求条件期望, 通常比较麻烦.

(1) 直接方法

例 3.2.4 设  $X_1, \dots, X_n$  独立同分布. i)  $X \sim N(1, 6)$ ; ii)  $X \sim$

$P(A)$ ; iii)  $\sim 7V(\wedge, 1)$ , 则  $T$  为完备充分统计量, 容易验证, 它分别为  $\theta < 9$ ,  $A$  和  $\theta$  的 UMRUE. | 例 3.2.5 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim R(\theta, 9)$ . i) 求  $\theta$  和

的 UMRUE; ii) 求  $g(\theta)$  的 UMRUE, 其中  $g(\theta)$  为可导函数. 解 i) 为完备充分统计量, 且  $t = \sum_{i=1}^n X_i + 1$  因此  $\theta = -X(4)$  为  $\theta$  的 UMRUE. 另外经直接计算可得

$n-1$  因此  $f_1 = 5 - \sum_{i=1}^n X_i$  为  $\theta$  的 UMRUE.

ii)  $Y = \sum_{i=1}^n X_i$  的分布密度为  $p(y, \theta) = n \theta^n y^{n-1} e^{-\theta y}$ . 设  $g(\theta)$  的 UMRUE 为  $h(Y)$ , 则由  $E_j h(Y) = g(\theta)$  可得

$\int_0^\infty h(y) n \theta^n y^{n-1} e^{-\theta y} dy = g(\theta)$ ,  $\forall \theta > 0$ ,

即

A  $F \sim U(1)$

3.2 无偏估计及其UMRUE和 UMRUE  $\int_0^1 y^n x h(y) dy = 0 n g(0), V(0)$ .

该式两边对0求导可得

$n \theta^n h(\theta) = 0 n g'(\theta) + n \theta^n \ln(\theta)$ ,

即  $r(\theta) = g(\theta) + \theta^n \ln(\theta)$ , 因此有  $A(y) = g(y) + n \theta^n \ln(y)$ . 例3.2.6 设  $A, \dots$ , 独立同分布,  $X_i \sim N(1, \sigma^2)$ , 求  $a(1) =$

$\ln(1 - \sigma^2)$  的 UMRUE.

解 完备充分统计量为  $r =$

$n$

$i=1$

$E[r(\theta)] = \ln(1 - \sigma^2)$ . 由于  $\ln(1 - e)$  为  $e$  的二次式, 可令  $\hat{T} = aT +$   
 $PT$  用待定系数法求出  $a, P$  使  $E[\hat{T}] = \ln(1 - \sigma^2) + P E(T^2) = 0(1 - \sigma^2)$ .

而由  $T \sim b(n, \theta)$  可得

$E(T) = n \theta, E(T^2) = \text{Var}(T) + (E(T))^2 = n(1 - \theta) + n^2 \theta$ .

代入比较得  $a = (n-1)P = -n_1(n-1) - 1$ , 因此  $\hat{r}$  的 UMRUE 为

例 3.2.7 设  $X_1, \dots, X_n$  相互独立,  $X_i \sim V(0, a^2)$ , 一切求相应参数的 UMRUE:  
i) 若  $\gamma \sim V(y^2)$ ; ii) 若  $\gamma \sim V(0, w^2) (n > 3)$ , 其中  $y \in R, w > 0, a^2$  都是未知参数.

解 根据例2.2.12, 情形i)的完备充分统计量为  $(r, x, \dots)$ , 其中  $r = \sum_{i=1}^n x_i^2$

—而由假设可知,  $E X_i^2 = y^2$ , 故有  $y^2 = X_i^2$ . 而  $E(r - X_i^2) = E r - (n-1) y^2$ , 其中  $r = \sum_{i=1}^n x_i^2 - c r^2 X(n-1)$ , 因此  $a^2 = (n-1)^{-1} \{T X_i\}$ . ii) 与例2.2.12类似, 情形ii)的完备充分统计量为  $(r, \dots)$ , 其中

$=$

$/$

$E[X^{n-3} T; \cdot] = [E X_i^{n-3}] \cdot E[(n-3) T^{1-1}] = w$ , 由此可得(1)的UMRUE 为  $\hat{a}^2 = (r - 3)^{-1} w$ . 疆

例3.2.8 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\mu, \sigma^2)$ , 求以下各参数的UMRUE;

$\mu$

$\sigma^2 \sim b(n, \mu)$ , 要寻找  $\mu$  使

$\mu \sim a Y(n-1)$ . 因此有  $E r$

$4 = (n-1) a^2$ , 所以  $P =$

$7; (n-1)^{-1} \mu$ . 又由假设可知  $E X_i^2 = \mu a^2$ , 为了得到  $w$  的 UMRUE, 必须

由

$(\mu^3) E[(n-3) \mu^2] = a^2$ , 由于与  $7$  相互独立, 因此有

消去参数  $a^2$

分布的公式(见第一章)可得  $E(r/\mu^2) - 1 =$

96

第三章点估计基本方法

$\cdot \sqrt{2} \dots, \sqrt{2} \sqrt{3} \dots, \dots \} \# , cr; 1 \} y, \text{詳}; \text{川} \} tr, ak, i^{\wedge} / cf(n > 2)$

$F_{-1}^-(p)$ , 即  $p$  分位数 ( $p < 1$ ).

解 正态分布的完备充分统计量为

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

;

iv)  $x_p(6)$

经直接计算可得  $\hat{\mu} = \bar{x}, P = S^2$ .

ii)  $M_2, \dots$  的估计应与  $P$  和  $P$  有关. 为简化计算, 可做变换  $K =$

i)  $X_t - M$ , 则有

而  $E(n \cdot S^2) = n \cdot \sigma^2$ , 故有

$$E\left(\frac{1}{n} \sum_{i=1}^n Y_i^2\right) = \sigma^2$$

因此可得  $\sigma^2$  的 UMVUE 为

$$Y_i \sim N(\theta, \sigma^2), X = Y + \epsilon, Y \sim N(\theta, \sigma^2/n).$$

$$E(P) = E(P + 2\sigma^2/z^2) = \sigma^2 + 2\sigma^2/n$$

$$\sigma^2$$

$$= \frac{1}{n} \sum_{i=1}^n Y_i^2$$

$$+ \frac{2\sigma^2}{n} = \frac{1}{n} \sum_{i=1}^n Y_i^2 + \frac{2\sigma^2}{n}$$

. 类似地可得

$$E(P) = E(P + m) = E(P + 3\sigma^2 Z^2/n) = \sigma^2 + 3\sigma^2/n$$

$$3\sigma^2$$

$$S^2, \text{ 无独立, 因而 } E[X^3 - 3n^{-1}(S^2X)] = 3\sigma^3.$$

$$-Q - M^3 = X^3 - nS^2X.$$

iii) 注意, 由  $E(S^2) = \sigma^2$  不能得到  $E(S) = \sigma$ , 需要另行计算. 由于

$$1/n \sum_{i=1}^n Y_i^2$$

$$r(n)$$

$$n-1$$

$$Q^2$$

$$\sim \chi^2_{n-1}$$

$$r$$

$(n-1)$  利用厂积分直接计算  $S_k$  的期望 ( $A$ : 满足  $n+k-1 > 0$  即可) 可得

$$E[(n-1) \int_0^1 t^k f(t) dt] = \frac{1}{n} \int_0^1 t^k f(t) dt$$

$$\sim r^2,$$

$$a > C = \frac{1}{n} \sum_{i=1}^n (Y_i^2 - \sigma^2) = \frac{1}{n} \sum_{i=1}^n Y_i^2 - \sigma^2$$

$$\sim 2\sigma^2$$

特别有  $\sigma^2 = S^2$ ,  $a = \frac{1}{n} \sum_{i=1}^n (Y_i^2 - \sigma^2) = \frac{1}{n} \sum_{i=1}^n Y_i^2 - \sigma^2$ . 因

此有 .,

所以

$$, -2)$$

,

### 3.2 无偏估计及其 UMVUE 和 UMVUE

97

$$a = K(n) S =$$

因此  $\sigma^2$  (通常称为信噪比) 的 UMVUE 为

$$\frac{1}{K(n)} \sum_{i=1}^n Y_i^2$$

iv) 由  $p$  分位数的定义有

$P \sim \chi^2_k$  (叫  $\chi^2_k$ ) 户呼卜 (¥卜 因此有  $(x_p - A)/\sigma = z_p$ , 其中  $z_p$  为标准正态分布的  $p$  分位数, 为

$P$  分位数, 为

. 从而有  $x_p = \sigma^2 + \sigma z_p$ , 因为  $E(x_p) =$

已知, 因此有线性关系  $x_p = \sigma^2 + \sigma z_p$

$E(x_p) = \sigma^2 + \sigma z_p$ . 最后得

$x_p = \sigma^2 + \sigma z_p$ . | 例 3. 2.9 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\theta, \sigma^2)$ , 求以下各

各

参数的 UMVUE: i)  $\theta$ ; ii)  $\sigma^2$ ; iii)  $\sigma^4$ ; iv)  $\sigma^2$ ; v)  $M/\sigma^2$ .

解 完备充分统计量为  $T = (X, S)$ , 其中  $S = (X^2 - J(1))$  且  $X(1)$  与  $S$  独立. 由第一章

的公式有  $u$

$$^{(D+r(nA,l), S \sim r(A,n-1));$$

$$E(X(1))E(S) = \frac{1}{n}$$

$$nA \quad A \quad n-2$$

$$0 \text{ 因为中}(1)-; ru^{ry}=A$$

$$ii) A=(n-2)S-\backslash A^'=(n-1)-JS. \quad iii) \text{ 由独立性以及以上公式可得}$$

$$E[X(1)S'1] = E[X(I)]E[S-,J] =$$

$$\text{因此} A/z = (n-2)X(1)S_{-1} - n$$

$$A \equiv Ag \quad 1 \quad nA \quad n-2 \quad n-2 \quad n(n-2).$$

$$A^2$$

$$iv) E(5'2) = U; -2Xn-3) - \quad ^WP = (n-2)(n-3)5-2.$$

$$v) \text{ 由独立性以及以上公式可得}$$

$$l \setminus n-l \quad z \quad -xu \quad n-1$$

$$E[1(1)S] =$$

$$, E(5'*)$$

98

### 第三章点估计基本方法

$$\text{而 } E(s^2) = \hat{Lg}$$

$$\text{所以最后} < \text{ 可得}$$

$$(2) \text{ 条件期望法}$$

$$E[(a-1)] \quad nA^2,$$

$$n - I$$

$$s^2 = 4, \quad n^2 \quad (n - 1) \quad \blacksquare$$

$$E[(n-1) \sim lX(1)S] = - \sim + - p - \quad A \quad nA$$

有些估计问题，很难用上述直接法求解，下面是一些例子。

例3.2.10 设 $X^1, \dots, X_n$ 独立同分布求以下 $a$ 的UMRUE:

$$= P(\text{尤}1 = \text{々}) = e^A | y$$

(久可理解为(0,0时间内有 $k$ 个故障(或电话呼叫)的概率, $M_0$ 则表示无故障的概率)。

解完备充分统计量为 $T = 2 \sim P(nA)$ , 易见, 由于 $e^A$ 的存

在, 很难求一个函数 $<p(T)$ , 使得 $EA[<jp(T)] = e^A$  (即使 $k=0$ ), 但  $A!$

是可设法先求 $A$ 的某一个无偏估计, 则由定理3.2.1知, 沁:  $E(\hat{I} T)$ 为 $a$ 的UMRUE.

为此, 取

$$\hat{e} = Z(X = (X_1, \dots, X_n)) T: X_1 = k \quad (, \quad \text{门} \cdot I \text{ 为示性函数. 显然有}$$

$$= E[ZU: X] = k\} = PA(X, = k) = \&,$$

即 $\hat{e}$ 为 $\hat{e}$ 的一个无偏估计, 因  $= E(\hat{e} I T)$ 为 $\hat{e}$ 的UMRUE. 剩下 的问题就是求条件期望或

条件概率:

$$n \quad I / (X_1 = c) | T = P(X, = k | 2X, = T$$

$$* = 1$$

由 Poisson分布与二项分布的关系即可求出以上条件概率, 下面直接计

算这一概率. 易见当 $T < k$ 时,  $\text{么} = 0$ ; 当 $T \geq k$ 时则有  $n$

$$m = k, \quad \hat{x} t = T) \quad \text{人} \quad i = 1$$

$$= \rangle (\text{苓} + \dots + \text{冬} = n$$

### 3.2 无偏估计及其UMRUE和 UMVUE

$$P(X, = \text{叫} = T - k \wedge$$

$$p(ix \wedge T)$$

99

例3.2.11 设 $X^1, \dots$ , 及独立同分布,  $X^1 \wedge p = P(X) \wedge f$  (若已知)的UMRUE.

解 完备充分统计量为  $T$ ，但是要求使  $E[\hat{p}(X)] = P(X \leq f)$  很困难，因此可采用条件期望法。为此，取

$$E[\hat{p}(X)] = E[E[\hat{p}(X) | T]] = P(X \leq f) = p.$$

所以  $\hat{p}(X)$  为  $p$  的无偏估计，因而  $p$  的 UMRUE 可表示为

$$P(X) = E(p(X) | T) = P(X \leq f | T).$$

以下根据 Basu 定理求此条件概率。以上条件概率可表示为  $P(X \leq f | T) = P(X \leq f | T = t)$ 。

又为完备充分统计量， $X, -X$  的分布与  $f$  无关： $X, -X \sim N(0, \sigma^2)$ 。

因而为辅助统计量，所以由 Basu 定理可知， $X, -X$  与  $T$  独立，从而以上条件概率可化为无条件概率

$$p(x) = P(X \leq f)$$

其中  $\Phi$  为标准正态分布的分布函数。所以  $p$  的 UMRUE 为  $\Phi(f/\sigma)$ 。

例 3.2.12 设独立同分布， $A \sim r(A, l)$  指数分布，求  $p = P(X \leq f)$  的 UMRUE (若  $p$  表示寿命分布，则  $p$  为寿命小于等于  $f$  的概率)。

解 完备充分统计量为  $T = \sum_{i=1}^n X_i$ ，取

$$\hat{p}(X) = f/T$$

$X_i \sim f(A, n)$ ，类似于前面的

100 第三章点估计基本方法

$$p = p(x) = f/(n+1)$$

则有  $E[\hat{p}(X)] = E[f/T] = P(X \leq f) = p$ ，因此  $\hat{p}(X)$  为  $p$  的无偏估计，并且以  $\hat{p}(X) = f/T$  为  $p$  的 UMRUE。以下根据 Basu 定理

$P(X \leq f | T = t) = P(X \leq f) = p$  为完备充分统计量，而

$$\hat{p}(X) = f/T \sim r(i, i) \text{ 与 } T \sim r(i, n) \text{ 独立}$$

$$n \geq 1$$

求条件概率  $P(X \leq f | T = t)$

则必有  $f < t$ ，因而  $p = f/t$ ；以下考虑的情形，这时有

$$f < t$$

其分布与  $A$  无关，因而为辅助统计量，所以由 Basu 定理可知， $X_i/T$  与

$r$  独立。事实上， $X_i/T$  服从  $r$  分布。因为

$$X_i \sim r(A, 1)$$

$$T = X_1 + \dots + X_n \sim r(A, n)$$

由第一章  $r$  分布定理知  $X_i/T \sim BE(1, n)$ 。因此上面条件概率的表达式可以化为无条件概率

沁叫卜+)。

该式可直接应用  $Pf$  分布积分得到

$$A$$

$$P(X \leq f | T = t) = \int_0^f (y/t)^{n-1} dy$$

$$= f^n / (t^n (n+1)) = 1 - (1 - f/t)^{n+1}$$

综合以上结果， $p = P(X \leq f)$  的 UMRUE 可表示为  $P(X) = 1 - [(1 - f/T)^{n+1}]$ ，

$$f < t, a > 0,$$

$$L$$

$$- \cdot \gg \cdot$$

$$\cdot < - \cdot W$$

$$\cdot * A * \cdot < \cdot J$$

$$* \cdot$$

$T=0$ . 易见, 若  $r =$   
 $i=1$

其中  $a += |* \cdot .H;$   
 $[0, a < 0.$

### 3.3 极大似然估计

极大似然估计 (MLE) 是统计学中最为重要、应用最为广泛的估计方法之一。虽然 Gauss 曾经提出过, 但主要还是 Fisher 的杰出贡献和大力提倡, 使之得到广泛的应用和深入的研究。本节介绍极大似然估计的原理、方法及基本算法。

```
. L
A.- * l 5 A . :
cv. i .
?%: *
>•
?. . . "
<-• ■
;•
f• F
4 A •
.: ■
```

### 3.3 极大似然估计

101

#### 3.3.1 定义与例题

设  $x \sim f(x; \theta)$  极大似然估计在直观上可作如下解释。在观测过程中, 若抽样能抽到则说明此  $X$  出现的可能性最大, 其相应的  $\theta$  应最接近真参数。因此, 当  $X$  取值  $x$  时, 真参数  $\theta$  的估计应取  $\hat{\theta}$ , 它使出现的可能性最大, 即使  $L(\hat{\theta}, x)$  最大:

定义 3.3.1 设  $X \sim f(x; \theta)$  把  $\ln f(x; \theta)$  视为  $\theta$  的函数, 则称它为  $X$  关于  $\theta$  的似然函数,  $L(\theta, x) = \ln f(x; \theta)$  (幻称为对数似然函数。若  $\theta$  满足

则称  $\hat{\theta}$  为  $\theta$  的极大似然估计 (MLE)。注 (1) 若  $c(x) > 0$ , 则使  $f(x; \theta)$  最大的充要条件是使  $c(x)/f(x; \theta)$

最大。因为  $c$  与  $\theta$  无关, 因而  $c(x)f(x; \theta)$  也称为似然函数, 即似然函数可允许相差一个与  $\theta$  无关的“正的常数”。

若分布族  $\{f(x; \theta)\}$

有共同支撑, 则使最大的充

要条件是使  $L(\hat{\theta}; x) = \ln f(\hat{\theta}; x)$  最大, 通常求极大似然估计都从  $L(\hat{\theta}, x) = L(\hat{\theta})$  出发。

(3) 似然方程

$dL(\hat{\theta})/d\theta = 0$  (幻称为对数似然方程)

为求解极大似然估计的必要条件 (方程的解可能是极小值, 也可能出现多峰情形)。但满足上式的解, 也是  $\hat{\theta}$  的一种估计。对常见分布族, 很少出现多峰情形。

(4) 若极大似然估计存在, 则它必为充分统计量的函数, 因为由因子分解定理有  $f(x; \theta) = g(T(x); \theta)h(x)$ , 使该式达到极大的充要条件是使  $g$  达到极大, 而由后者必有。这一性质说明, 极大似然估计与充分性原则是一致的。以下通过若干例题说明求解极大似然估计的基本方法。

例 3.3.1 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\mu, \sigma^2)$ , 求  $\mu, \sigma^2$  的极大似然估计。解相应的密度函数和对数似然函数分别为

$$\theta = \theta(T(x))$$

因此

$$L(t, a^2) \sim L(f_j L, a^2), \quad V/x.$$

$$Z(l, < r^2) \geq \beta(\text{弘}, < r^2), \quad V(A_t, cr).$$

102

第三章点估计基本方法  $-x)^2 + n(x - \bar{x})^2]$ .

$$l(M, a^2)$$

$$-T^{(2)}(2) - \text{志} \quad 2 \text{ 仏}$$

可分两步来求  $M, V$  的极大似然估计. (a) 对任何  $a$ ; 由上式可知,  $l=x$  时, 指数部分方括号内的值最

小, 因而  $l(x; 1, 6T^2)$  最大; (b) 对  $L(A, a^2)$  求导可得

$$= 0, < r^2 = - (x_i - x)^2. \quad da^2 \quad n \quad i=i$$

由于以上的似然方程有唯一解, 直接验证可知  $f(A, a^2)$  关于  $a^2$  的二

阶导数在  $P$  处小于零, 因此

$f$  为  $A(A, a^2)$  的唯一的极大值点, 所以有

$$A(t, cr^2) > L(l, cr^2), \quad V \quad a. \quad \text{又由(a), 对任何 } a \text{ 有}$$

因此  $l, a^2$  的极大似然估计为  $l=1$  和  $9 = S^2 = \frac{1}{n} \sum_{i=1}^n (x_i - l)^2$ .

例3.3.2均匀分布的极大似然估计. 设..., 独立同分布, i)  $X_i \sim R(0, \text{没})$ ; ii)  $X_i \sim R(0, 30)$ ; iii)  $X_1 - \dots - X_n \sim R(a/2, M+a/2)$ . 求相应参数的极大似然估计.

解 先看一般情形, 若  $X' \sim R(a, b)$ , ( $a < b$ ), 则有

$n_i$

$$a, b) = Y \setminus \text{-----} 1 \quad j a \geq x_f \geq 6 \quad |$$

$$* = i \quad o - a$$

$$= 7a \quad 1 \quad v7 \quad ia \text{ 矣 } x(i) \text{ 龙}(4) \quad I. \quad \setminus o - a)$$

i) 相应的密度函数为  $f(x; e) = e^{-n_i \sum_{l=1}^n x(l)^{e+1}}$   $f$  考虑

其最大值. 当  $X = x = (\%, \text{时}, 6 > \text{越小}, \text{则} \text{没}) \text{越大}$ . 但由以上表达式可知, 恒有  $0^x(n)$ , 因此  $0 = x(n)$  时,  $0$  最小  $6 > \text{"最大}$ , 从而  $l(x; < 9)$  最大. 因此  $0 = X\{n\}$  为  $0$  的极大似然估计.

ii) 相应的密度函数为

$$l(\text{太}; \text{没}) = (5) \quad i \quad x(i) \quad x(n) \quad 33 \} = ($$

由以上表达式可知, 要使  $f(x; e)$  尽量大则应要使  $0$  尽量小,  $X(4)/3$ , 因此  $6 > x\{n\}/3$  时,  $l(x; 0)$  最大, 即  $0 = X\{n\}/3$ .

iii) 相应的密度函数为

$f$  越大, 因而  $l(*; ,$

而

### 3.3极大似然估计

103

$= 1/x(1) < x(n)^{e+1} \}$   $= \setminus x(n) - 1 \geq \text{没} \geq \%(1)$ . ( $f(x; e)$  仅取  $0$  或  $1$  两个值, 取  $1$  时最大. 因此  $(4) - 1, x(1)$  时 都可视为  $e$  的极大似然估计, 所以解不唯一 (下一章可求出唯一的最优

同变估计). iv) 相应的密度函数为

$$X(\cdot) \quad X(n) \quad ( \quad + \_y | -$$

要使  $l(X; 0)$  尽量大, 则应使  $< 7$  尽量小, 但是由以上表达式可知,  $CF \&$

满足关系: (a)  $l(1)$  右  $cr/2$ , 取其最小为  $a/2 = \setminus j \quad l - x(1)$ ; (b)  $x(n) - IJL \quad a/2$ , 取其最小为  $a/2 = x(n)$  综合(a)和(b)可知,  $l/z, o \blacksquare$  应满足

联立方程

求解即得

$$-y = x(i), \quad M+y = \sum_{i=1}^n x(i)$$

$$A = \frac{1}{n} \sum_{i=1}^n x(i), \quad \hat{\lambda} = \frac{1}{n} \sum_{i=1}^n \frac{1}{x(i)}$$

例3. 3.3 指数分布的极大似然估计: 设  $X_1, \dots, X_n$  独立同分布, (i)  $x_i \sim r(A, \lambda)$ , (ii)  $x_i \sim \text{Ac}(r(i, \lambda))$ , (iii)  $x_i \sim \text{At}(r(\lambda, i))$ .

由

$$\frac{\partial}{\partial A} \ln L(A, \lambda) = 0 \text{ 可得 } A = \bar{X} \quad (\text{注: 若 } A \sim r(\lambda, i))$$

$$i=1$$

$$T, 1), \text{ 贝 } \{j \in \mathbb{N} \mid j \leq T\}$$

$$= e^{-\lambda} \lambda^j / j!$$

$$= \frac{1}{n} \sum_{i=1}^n \frac{1}{x(i)}$$

$$t \rightarrow \infty, \quad \lambda \rightarrow \infty$$

$$c: \lambda > 0$$

$$\lambda < 0, \quad \lambda = 0$$

$$\lambda = 0$$

$$\lambda = 0$$

$$\lambda = 0$$

$$\lambda = 0$$

解 (i) 相应的密度函数和对数似然函数分别为  $n$

$$l(x; A) = n \ln A - A \sum_{i=1}^n x_i$$

(ii) 相应的密度函数为

$$\lambda$$

$$/$$

$$n \ln A - A \sum_{i=1}^n x_i$$

$$(\lambda; A) = e^{-\lambda} \lambda^j / j!$$

$$($$

$$)$$

$$I \mid x(I)$$

$$)1$$

$$\lambda$$

要使  $l(x; A, \lambda)$  尽量大, 则应使  $M$  尽量大, 即要求  $M$  尽量大, 但是从备  $\lambda(1)$ , 因此当  $\lambda = X(H)$  时  $f$  最大, 因而  $l$  最大, 所以有  $\hat{\lambda} = \bar{X}(1)$ .

(in) 相应的密度函数为

$$l(x; A, \lambda) = H[Xe^{-\lambda} \sum_{i=1}^n x_i] = A'e^{-j} \lambda^j / j!, \quad j = 1, 2, \dots$$

(a) 对任何固定的  $A$ , 要使  $l(x; A, \lambda)$  尽量大, 则应使  $e^{-\lambda} \lambda^j$  尽量大,

即要使  $M$  尽量大, 但是  $\lambda(i)$ , 因此当  $\lambda(1)$  时  $f$  最大, 从而  $e^{-\lambda} \lambda^j$  最大, 即  $l(x; A, \lambda)$  最大, 所以有  $\hat{\lambda} = \bar{X}(1)$ ;

$$($$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

$$\lambda \rightarrow \infty$$

## 104 第三章点估计基本方法

(b) 把  $\lambda(1)$  代入  $l(x; A, \lambda)$  的表达式可得  $n$

$$f(x; A, \lambda) = e^{-\lambda} \lambda^j / j!$$

$$l(A, \lambda) = n \ln A - A \sum_{i=1}^n x_i$$

$$s = \sum_{i=1}^n x_i$$

类似地, 亦可求出截尾情况下指数分布的极大似然估计: 若

久, ..., 及为独立同分布样本, 且  $X_1, \dots, X_n$  但是仅观察到前  $r$



个样本,  $(r_1, \dots, r_K) = (X(1), \dots, X(r))$ , 则有  $\hat{\theta} = X(1)$ ,  $a = S/r$ , 其中  
由  $\theta = 0$  可得  $\theta_A$

因此对任何  $(A, I)$ , 有

$l(\theta; A, \hat{\theta}) / l(\theta; A, x) / l(x$

AAA

;  $A$  所以, 以上  $X, A$  为  $A, I$  的极大似然估计. ■

$r$

$51 = 2L[A'] - \log(1) + (\sim r)C^r - X(1) \log(r, 7 \cdot -1).$

特别, 若  $Z=0$ , 则有

$-\log T_n, r > T_n, r - (i) + (\ll \sim$

例 3. 3.4 多项分布  $V = (V_1, \dots, V_k) \sim T(n, \theta)$ , 其中

$k$

$n, \theta = (\theta_1, \dots, \theta_k) \in \Delta_k$ , 求  $\theta$  的极大似然估计.  $i=1$

解相应的密度函数和对数似然函数分别为

$N$

$\theta =$

因为有约束条件  $\theta_i = 1, i=1$

$k$

,  $\theta_i = nA_i; T) = -$

$l(\theta) = 2 \sum_{i=1}^k n_i \log \theta_i + \log(n!) - \log(n! \prod_{i=1}^k \theta_i)$

为  $k$

$l(\theta) = 2 \sum_{i=1}^k n_i \log \theta_i - \log(n! \prod_{i=1}^k \theta_i), i=1$

$dL, \theta_i$

$k$ , 可用 Lagrange 乘子法, 令

-----  $\theta_i = n_i/n, i=1, \dots, k$ .  $I$   $n$

$iZ1$

3.3 极大似然估计

105

$dL, \theta_i$

$G = \theta_i = 0, i=1, \dots, k$

$k, k$

约束条件  $\sum_{i=1}^k \theta_i = 1$  代入上式可得  $l(\theta) = 2 \sum_{i=1}^k n_i \log \theta_i - \log(n! \prod_{i=1}^k \theta_i)$ , 因  $i=1, \dots, k$

此有  $\theta_i = n_i/n$ . 即  $\theta$  的极大似然估计为

$\theta_i = n_i/n, i=1, \dots, k$ .  $I$   $n$

;

$n$ ;

例 3.3.5 Laplace 分布. 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim M(a, b)$ , 求  $M, a$  的极大似然估计.

解 相应的密度函数为  $f(x; a, b) = \frac{1}{b} \exp(-x/b - a/x)$

对任何  $a > 0$ , 求  $a$  使  $l(a)$  最大, 等价于求使  $2 \sum_{i=1}^n |x_i - a|$  最小.  $i=1$

$n$

,

以下求  $\sum_{i=1}^n |x_i - a|$  的最小值点, 为此, 把  $W(a)$  改写为以下形式  $f(a) =$

式并设法去掉绝对值号:

$f(a) =$

$2 \sum_{i=1}^n |x_i - a| = 2 \sum_{i=1}^n (x_i - a) \cdot \text{sgn}(x_i - a) = 2 \sum_{i=1}^n (x_i - a) \cdot \begin{cases} 1 & x_i > a \\ -1 & x_i < a \end{cases}$

当观察值  $X_1, \dots, X_n$  确定后, 考虑  $f(a)$  从  $-\infty$  到  $+\infty$  变化时

何时能达到最小值. 若  $f(a) = 0$ ,  $X(1)$ , 则  $f(a) = 0$

$2 \sum_{i=1}^n (x_i - a) = 0$

—叫为M的线性减函数.同理, 若弘e 打n

\*=| 1=1 - \ x(1)为弘的线性增函

x b ^ A ^

f^'

(x(n), +00), 则<P(M)= Z(从-%(1))

i=1 ^ r

i= 数.再考虑[x(l), x<rv)]时<p(g)的变化情况, 可参见图3.3.1.若从

=2 i = l

[m 一太(1)]+ Z [~)一从]

1

e [X(k) fX(A+l)) (々=1, ..., 沒-1), 则有

kn nx

kn

• \* 5

l-? .

J: -/f

=仏-2><0 + 工X(0 -(n-k)^ i=1 i=t+l

kn

=(2k-n)从- x(/) + \$ x(1). \* = I t = fc + 1

由上式可知, P(M)为#的连续线性函数, 当2k ~ n < 0时, 在区间 [x(n, 戈(\*+i))上斜率为负;当2A?-n》0时, 在区间[x(k) ,x(k+t))上斜率 为正:

Z = Ar + l S. .

f. •

\* . : f

»z- v f: ,

; -J

<- . ?

■?.- V

|

106

### 第三章点估计基本方法

<0, 女<介,

>0,

/ze [%“) , %(“”), 々=l, 2, .“, n-l.

因此当弘从-00到+00时, 由负到正, W/X)由减到增, 以下分7Z

为偶数、奇数来讨论.

(a) n=21为偶数(见图3.3.1右), 则k=I时史'(/z)在[x(J),

x(/+1))上为零, 在 左递减, 在 + 右递增, 乎(M)在(什>, x(/ + I))上都达到最小值, 所以最小值点不唯一.一般/;可取样本中位数

1(. +义(' +1)

= 2\_ = Me.

(b) n=2Z + 1为奇数(见图3.3. 1左), 则对任何/X, 〃(/£)#(), 在[x(1), ~+1))上 </ (M) =21-n= - 1 <0, 在[x(/+J) ,%{/ +2))上 = (21+2) -n=1 >0, 衅=x(, +i)时妒(弘)达至II最4、值, 故/!=A\i+1).

综合(a)、(b)两种情形, M的极大似然估计都是样本中位数財, :

/(%;/z, (r)=(占)exp|一去名 | %, -/I|

- nlog(2cr)

故根据例3. 3. 1和例3.3. 3的论证可知<7的极大似然估计为

<A

$\sum_{i=1}^n V_i = \sum_{i=1}^n V_i$

⊥|A

3.3极大似然估计

“ $\pm g = \pm^i K 1$ .”

3.3.2 指数族分布的极大似然估计

107

指数族分布的密度函数为

$X \sim \pi(x; \theta) = \frac{1}{Z(\theta)} h(x) \exp \{ \eta(\theta)^T T(x) - A(\theta) \}$  (3.3.1)

若 $x_1, \dots, x_n$ 为其独立同分布的样本, 则密度函数为  $P_{\theta} = \prod_{i=1}^n \pi(x_i; \theta)$ ,

仍然为指数族, 因此以下仅考虑单样本情形. 定理3.3.1对于指数族分布(3.3.1)式,  $\theta$ 的似然方程可表示为

$E[T(X)] = \eta(\theta)$  (3.3.2) (或简记为若方程(3.3.2)在 $\theta$ 内部有解, 则解必唯一,

且

为极大似然估计.

证明 由(3.3.1)式可知, 相应的对数似然函数为  $L(\theta) =$

$n \log \pi(x; \theta) = -n A(\theta) + \sum_{i=1}^n \eta(\theta)^T T(x_i) + \sum_{i=1}^n \log h(x_i)$ , 因此有

$\frac{\partial L(\theta)}{\partial \theta} = 0$ . 由

(3.3.3)

由于  $E[T(X)] = \eta(\theta)$ , 故(3.3.3)式即为  $\eta(\theta) = \eta(\theta)$ , 由此即可

推出(3.3.2)式. 又若 $\theta^*$ 满足(3.3.2)式, 且在 $\theta$ 内部, 则有  $L(\theta^*) = 0$ . 另外由(3.3.3)式有

$\frac{\partial^2 L(\theta)}{\partial \theta^2} = -n \text{Cov}(T(X))$

即  $Z(\theta) = -n \text{Cov}(T(X))$ , 由指数族分布的性质可知  $S(\theta)$  在 $\theta$ 内部正定, 从

而在 $\theta$ 内部负定, 因此为  $M(\theta)$  在 $\theta$ 上唯一的最大值点. 注 若  $L(\theta) = 0$  的解在边界上, 则问题通常比较复杂. 定理3.3.2若 $X$ 的分布为以下曲指数族:

$X \sim \pi(x; \theta(p)) = h(x) \exp \{ T(x)^T \eta(p) - A(p) \}$ ,

其中  $\eta = (\eta_1, \dots, \eta_k)$

$T, p \in \mathbb{R}^k$ ,

且  $\eta$  的定义域满

足  $e =$

$\eta, \eta = (\eta_1, \dots, \eta_k)$ . 若  $\eta$  满足方程

“  $\eta = \eta(p)$  ”

证明 仍记指数族分布(3.3.1)的对数似然函数为  $\ell(\theta)$ , 并记

$r(x) = \exp \{ T(x)^T \eta \}$  且设  $(\eta)$  为  $\theta$  的内点, 则  $\eta$  为的极大似然估计.

(3.3.4)

$Z(\eta) = L(\eta(p))$

$\eta = \eta$ , 要证

$i(\eta) \geq 1/(3) \cdot V(\eta)$ .

j

,

108 第三章点估计基本方法

由于  $\eta$  满足方程(3.3.4), 因此亦满足方程(3.3.4). 由定理 3.3.1知  $\eta$  为  $\theta$  的极大似然估计, 即

$L(\eta) = V(\eta)$ .

因此有

$=Z(\theta) \leq L(\theta)$ , 而对任一因  $\theta \in \Theta$ , 所以必有

$L(\theta) \leq L(\hat{\theta})$ ,  $\hat{\theta} = Z(\theta)$ ,  $\forall \theta \in \Theta$ . 所以  $\hat{\theta}$  为  $\theta$  的极大似然估计. ■

注 (3.3.4) 式是求解  $\theta$  的充分条件, 并非关于  $\theta$  的似然方程, 其似然方程为

$P = 0$ , 即  $\sum_{i=1}^n \frac{\partial}{\partial \theta} \ln L(\theta) = 0$

例3.3.6 设  $(X_1, Y_1), \dots, (X_n, Y_n)$  独立同分布,  $(X, Y) \sim$  二元正态分布

$N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ , 求  $\sigma_1^2, \sigma_2^2$  的极大似然估计.

解 样本的联合密度函数为

$$L(\theta) = \prod_{i=1}^n \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(X_i-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(X_i-\mu_1)(Y_i-\mu_2)}{\sigma_1\sigma_2} + \frac{(Y_i-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

这可视为指数族分布, 其中

$T = \left(\sum_{i=1}^n \frac{(X_i-\mu_1)^2}{\sigma_1^2}, \sum_{i=1}^n \frac{(Y_i-\mu_2)^2}{\sigma_2^2}, \sum_{i=1}^n \frac{(X_i-\mu_1)(Y_i-\mu_2)}{\sigma_1\sigma_2}\right)$

$\eta = \left(\frac{1}{\sigma_1^2}, \frac{1}{\sigma_2^2}, -\frac{\rho}{\sigma_1\sigma_2}\right)$

其中  $\theta = (\sigma_1^2, \sigma_2^2, \rho)$ . 由定理3.3.2, 可通过方程  $E(T) = T$  和  $\eta = \eta$  来求  $\sigma_1^2, \sigma_2^2, \rho$ , 即

由此可得

$E\left(\sum_{i=1}^n \frac{(X_i-\mu_1)^2}{\sigma_1^2}\right) = n \frac{\sigma_1^2 + \rho^2 \sigma_2^2}{\sigma_1^2(1-\rho^2)}$

$E\left(\sum_{i=1}^n \frac{(Y_i-\mu_2)^2}{\sigma_2^2}\right) = n \frac{\sigma_2^2 + \rho^2 \sigma_1^2}{\sigma_2^2(1-\rho^2)}$

$E\left(\sum_{i=1}^n \frac{(X_i-\mu_1)(Y_i-\mu_2)}{\sigma_1\sigma_2}\right) = n\rho$

$\frac{1}{\sigma_1^2} = \frac{1}{n} E\left(\sum_{i=1}^n \frac{(X_i-\mu_1)^2}{\sigma_1^2}\right)$

$\frac{1}{\sigma_2^2} = \frac{1}{n} E\left(\sum_{i=1}^n \frac{(Y_i-\mu_2)^2}{\sigma_2^2}\right)$

注 类似地, 可求分布  $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$  中参数的极大似然估计. 例3.3.7 含参数的多项分布. 设

$N = (N_1, N_2, N_3, \dots, N_r)$ ,  $N = n$

$\theta = (\theta_1, \theta_2, \theta_3, \dots, \theta_r)$

3.3 极大似然估计

109

其中  $\theta = (\theta_1, \theta_2, \theta_3, \dots, \theta_r)$ ,  $\theta_i = \frac{1}{n} \frac{N_i}{N}$ ,  $\sum_{i=1}^r \theta_i = 1$ ,  $\theta_i \geq 0$ , 求  $\theta$  的极大似然估计.

解 相应的密度函数为  $p(n, N_1, N_2, N_3, \dots, N_r; \theta) = \frac{n!}{N_1! N_2! N_3! \dots N_r!} \theta_1^{N_1} \theta_2^{N_2} \theta_3^{N_3} \dots \theta_r^{N_r}$

$= \frac{n!}{N_1! N_2! N_3! \dots N_r!} \theta_1^{N_1} \theta_2^{N_2} \theta_3^{N_3} \dots \theta_r^{N_r}$

$= \frac{n!}{N_1! N_2! N_3! \dots N_r!} \exp\left\{\sum_{i=1}^r N_i \ln \theta_i\right\}$

$+ 2n \log(1 - \theta_0)$

为指数族分布, 其中  $T(x) = \sum_{i=1}^r N_i \ln \theta_i$ ,  $\theta = \log \theta$ . 由于  $\theta \sim 1$

$b(n, \theta) = \frac{n!}{N_1! N_2! N_3! \dots N_r!}$ , 因此  $\theta$  满足以下方程:

$\frac{\partial}{\partial \theta_i} \ln L(\theta) = \frac{N_i}{\theta_i} - \frac{1}{1 - \theta_0} = 0$

令代人以上方程可得

$\frac{N_i}{\theta_i} = \frac{1}{1 - \theta_0}$

由此可得  $\theta$  的极大似然估计为  $\hat{\theta}_i = \frac{N_i}{n}$ ,  $i = 1, 2, 3, \dots, r$ . | 3.3.3 不变原理

设  $X \sim \theta$ . 本小节的目的在于说明如何求解一个参数  $\theta$  的函数  $f(\theta)$  的极大似然估计. 主要证明: 在合理的意义下, 若  $\hat{\theta}$  为  $\theta$  的极大似然估计, 则  $f(\hat{\theta})$  为  $f(\theta)$  的极大似然估计. 由第3.2节的讨论

论可知, 对一致最小风险无偏估计, 若  $S$  为  $\theta$  的一致最小风险无偏估计, 则一般  $S$  不是  $f(\theta)$  的一致最小风险无偏估计. 因为  $E(S) = \theta$ , 一般没有  $E(S) = f(\theta)$  (见前面的例子). 但对极大似然估计, 若  $S$  为  $\theta$  的极大似然估计, 则  $S$  为  $f(\theta)$  的极大似然估计; 若  $A$  的极大似然估计为  $X$  而  $\hat{\theta} = A$ , 则必有下面根据极大似然估计的原理加以详细说明.

定义3.3.2 设  $X \sim f(x; \theta)$ ,  $\theta \in \Theta$ . 对于  $f(\theta)$ ,  $\theta \in \Theta$ , 以及  $\theta = g(\eta)$ , 今记  $\eta_1(\theta) = |(\eta: g(\eta) = \theta)|$ , 则  $X$  关于参数  $\eta$  的导出似然定义为

$L^*(\eta) = \max_{\theta \in \Theta} f(\theta) = \max_{\theta \in \Theta} L(\theta)$ .  $\hat{\eta} = \eta_1(\hat{\theta})$

今记  $\eta = \log(\eta; \hat{\theta})$ ,  $\hat{\eta} = (\hat{\eta}_1, \dots, \hat{\eta}_r)$ ,  $i = 1, 2, \dots, r$ ,

W,  $r^k$ . 当  $\theta \neq 0$  时,  $\hat{e}/2$ . 注意, 通常的定义域可能比  $\mathcal{S}$  小, 例如  $\phi(A) = g(\phi) = \hat{e}/2$ , 贝 'J  $6 > e(-\infty, +\infty)$ , 而  $[0, +\infty)$ . 为了说明  $\hat{\theta}$  和极大似然估计之间的关系: 首先必须定义分布族关于参数  $\theta$  的似然 (称为导出似然), 及其极大似然估计.

少的极大似然估计定义为必, 它满足

$K$  中', ...

,  
110

### 第三章点估计基本方法

由于  $g(\theta)$  因此,  $\theta^*$

所以由 LD 的定义可得

$$L^*(A) = \max_{\theta \in \mathcal{S}} f(\theta). \quad \hat{e} f i$$

这个定义可作如下理解. 在空间  $Z^2$  上的一个点  $\phi(A)$ , 它可能对应于  $\theta$

致性 (即都是全局最大值).

定理 3.3.3 设  $X \sim (\mathcal{S}; \phi)$ ,  $\phi \neq 0$ . 若  $\theta$  的极大似然估计为  $\hat{\theta}$  则

$\phi$  关于导出似然的极大似然估计为  $g(\theta)$ .

证明 今记  $g(A) = \theta$ , 根据定义 3.3.2, 我们将证明

$$L^*(\hat{\theta}) = \max_{\theta \in \mathcal{S}} L(\theta), \quad \forall \theta \in \mathcal{S}. \quad (3.3.5) \quad \text{注意: } L^*(\hat{\theta}) = \max_{\theta \in \mathcal{S}} |L(\theta)|, \quad L(\hat{\theta}) = \max_{\theta \in \mathcal{S}} L(\theta).$$

由  $L^*(\hat{\theta})$  定义可知

$|$   
 $\text{cr}_2$ ),  $i=1, 2, \dots$ , ". 求从,  $\text{er}_2$  以及  $a=E(X)$ ,  $t_2=\text{Var}(X)$  的极大似然估计.

解 令  $K^{\log X}$ , 则  $V(y_2)$ ,  $t=1, \dots, n$ . 由正态分布的极大似然估计可得

, 因此必须先对此求最大, 然后再在上

上许多点, 求最大; 这样才能保证 LD 为全局最大, 以保持 LD 与  $f(\hat{\theta})$  的一

$$A^*(\hat{\theta}) = \max_{\theta \in \mathcal{S}} f(\theta).$$

$g(\theta) = \hat{\theta}$

$$L^*(\hat{\theta}) = \max_{\theta \in \mathcal{S}} L(\theta) = \max_{\theta \in \mathcal{S}} L(\theta) = L^*(A), \quad \forall A \in \mathcal{S}.$$

$$: g(\hat{\theta}) = \hat{\theta}$$

$$\hat{\theta} \in \mathcal{S} \cap \mathcal{I}^*$$

{e:

因此 (3.3.5) 式成立, 定理得证.

例 3.3.8 对数正态分布. 若  $A, \dots, X_n$  独立同分布,  $X_i \sim \text{LNW}$ ,

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

$$\hat{\mu} = \bar{Y}, \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (\ln X_i - \bar{Y})^2$$

$r_2 = \text{Var}(X_j) = a^2(e^{a^2} - 1)$ . 由不变原理可得

$$a = \exp\left\{\frac{j}{t} + \frac{\text{cr}_2}{2}\right\}, \quad r_2 = a^2(e^{a^2} - 1) \quad (\text{注意, 很难得到对数正态分布的一致最小风险无偏估计}).$$

| 例 3.3.9 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim \text{TV}(M, a^2)$ ,

$x_p) = p$ . i) 若  $p$  已知, 求  $\theta$  的极大似然估计; ii) 若  $\theta$  已知, 求  $P$  的极大似然估计.

■ -

解 由例 3.3.1 可知  $\theta = X$ ,  $a^2 = S^2$ , 由不变原理得  $a = S$ . 设标准

正态分布的分布函数为  $\Phi(-)$ , 则有

$$P(X \leq a) = \Phi\left(\frac{a - \mu}{\sigma}\right)$$

$$= p.$$

$$\frac{a - \mu}{\sigma} = \Phi^{-1}(p)$$

$$a = \mu + \sigma \Phi^{-1}(p), \quad \text{其中 } \mu = 5,$$

其中  $\mu$  为  $(\cdot)$  的  $p$  分位点, 为已知. 因此  $X_p \sim \mu + a p$ , 由不变原理得

### 3.3 极大似然估计

111

A

AAA </p> Xp=M+azp9 p=01-^~

(本例求解方法比例3.2. 10简单, 但性能不比一致最小风险无偏估计差, 见第五章). I

#### 3. 3.4 子集参数的似然

在前面的例子中, 参数 $\theta$ 经常有好几个分量, 如 $\theta = (\theta_1, \theta_2)$ . 为求

没的极大似然估计, 可固定一个分量, 如 $\theta_1$ , 先求参数 $\theta_2$ 的极大似然估计, 然后再求 $\theta$ 的极大似然估计(例如可参见例3. 3. 1.3. 3. 3. 3.3.5 等). 下面根据极大似然估计的原理讨论一般情形下子集参数的似然函数和极大似然估计.

假设 $X \sim f(x; \theta)$ ,  $\theta = (\theta_1, \theta_2)$ , 其中 $\theta_1$ 为 $k_1$ 维,  $\theta_2$ 为 $k_2$ 维, 其的对数似然函数记为

$l(\theta) = L(\theta) = \log f(x; \theta)$ . 定义3.3.3在以上条件下, 设 $\theta_1$ 任意固定时,  $L(\theta_1)$ 中 $\theta_2$ 的

极大似然估计为 $\hat{\theta}_2(\theta_1)$ , 即

$\hat{\theta}_2(\theta_1) = \arg \max_{\theta_2} L(\theta_1, \theta_2)$ . (3.3.6)

则称 $L(\theta_1)$ 为子集参数 $\theta_1$ 的截面似然(profile likelihood). 注意,  $\hat{\theta}_2$ 应与 $\theta_1$ 有关.

例3.3. 10 设 $X_1, \dots, X_n$ 独立同分布

. 分别固定 $\theta_1$ 和 $\theta_2$ , 求 $\theta$

解 若固定 $\theta_1$ , 则可得到 $\hat{\theta}_2$ , 此估计与 $\theta_1$ 无关(见例3.3.1), 再得到

$\hat{\theta}_1 = \bar{X}$  (尤, -无)2. n

若固定 $\theta_2$ , 先求 $\theta_1$ 的极大似然估计, 则与 $\theta_2$ 有关.

$l(\theta) = -\log(2\pi)^{-n/2} \exp\{-\frac{1}{2}(\theta - \mu)^T \Sigma^{-1}(\theta - \mu)\}$ . (3.3.7)

,  $\theta_1$ 的极大似然估计.

由 $\frac{\partial l}{\partial \theta_1} = 0$ 可得  $\hat{\theta}_1$

$\hat{\theta}_2(\hat{\theta}_1) = \arg \max_{\theta_2} l(\hat{\theta}_1, \theta_2)$

该式与 $X_i$ 有关, 代入(3.3.7)式可得  $L(\hat{\theta}) = l(\hat{\theta}_1, \hat{\theta}_2) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log |\Sigma| - \frac{n}{2} (\hat{\theta} - \mu)^T \Sigma^{-1} (\hat{\theta} - \mu)$

$\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2)$ ,

#### 112 第三章点估计基本方法

为求 $\theta$ 的极大似然估计, 上式对 $\theta$ 求导得 $\frac{\partial l}{\partial \theta} = 0$  把 $\theta_1 = \hat{\theta}_1$ 代入 $\hat{\theta}_2(\theta_1)$ 即得  $\hat{\theta}_2$  的极大似然估计:

$\hat{\theta}_2 = \hat{\theta}_2(\hat{\theta}_1) = \arg \max_{\theta_2} l(\hat{\theta}_1, \theta_2)$

以上例题说明, 两种方法的结果相同, 对于一般情形, 有如下定理: 定理3.3.4设 $\hat{\theta}_2(\theta_1)$ 在 $\theta_1$ 上的极大似然估计 $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2)$ 存在且唯一,

$L(\hat{\theta}) = L(\hat{\theta}_1, \hat{\theta}_2)$  中 $\hat{\theta}_2$ 的极大似然估计为 $\hat{\theta}_2$ , 即 $\hat{\theta}_2$ 满足  $\frac{\partial l}{\partial \theta_2}(\hat{\theta}_1, \hat{\theta}_2) = 0$ . (3.3.8)

又记 $\theta_1 = (\theta_1)$ . 则有(1) $\frac{\partial l}{\partial \theta_1}(\hat{\theta}_1, \hat{\theta}_2) = 0$ ,

(2) $\frac{\partial l}{\partial \theta_2}(\hat{\theta}_1, \hat{\theta}_2) = 0$ , 证明(1)以下证明

$L(\hat{\theta}) \geq L(\theta)$ , 且  $L(\hat{\theta}) > L(\theta)$ .

则由唯一性知 $\hat{\theta}$  由于 $\hat{\theta}_1$  为极大似然估计, 显然有 $L(\hat{\theta}_1) \geq L(\theta_1)$ ,

今证其反面. 由 $\hat{\theta}_2$ 的定义(3.3.8)式知

由(3.3.6)式可知上式等价于

$L(\hat{\theta}_1, \hat{\theta}_2) = L(\hat{\theta}_1, \hat{\theta}_2) > L(\hat{\theta}_1, \theta_2)$ , VR.

又由 $e_2(e_j)$ 的定义知, 对任意固定的 $\theta$ , 以及任意的 $\theta_2$ , 有  
以 $\theta_2$ 为 $\theta$ 代入上面(3.3.9)式有  
(3.3.9)

$\theta_2 \in \{e_j\}$ ,  $\forall \theta \in \Theta$ .

因此有 $\theta_2 \in \{e_j\}$ 多 $a(\theta, \theta_2)$ . 综上所述, 由极大似然估计的唯一性有  
 $\theta_2 = \theta$ .

(2) 由(1)以及 $e_2(e_j)$ 的定义知,  $\theta_2 = e_j = \theta_2^*(\theta)$ . 本小节内容可归纳为以下两点:

(1) 若所讨论的问题只对某一 $\theta$ 个参数 $\theta$ 感兴趣,  $\theta_2$ 看作多余参数, 则可考虑截面似然 $L(\theta, \theta_2)$ 估计等问题中都很有用.

(2) 求参数 $\theta = (\theta, \theta_2)$ 的极大似然估计可以直接求解; 也可考虑应用截面似然. 即先固定 $\theta_2$ , 求出 $\theta$ 的极大似然估计 $\hat{\theta}(\theta_2)$ , 再代入到 $L(\hat{\theta}(\theta_2), \theta_2)$ 求 $\theta_2$ 的极大似然估计 $\hat{\theta}_2$ , 然后再代入 $\hat{\theta}(\hat{\theta}_2)$ 得到 $\hat{\theta}$ 的极大似然估计.

$\hat{\theta} = L(\hat{\theta}, \hat{\theta}_2)$

这在假设检验、区间估计、

### 3.3 极大似然估计

113  
估计 $\theta_2 = \theta_2(\theta)$ . 事实上, 这一方法在前面例题中已多次用到. 3.3.5 极大似然估计的迭代算法

通常的数值计算大多需要使用迭代算法. 以下介绍极大似然估计的常用迭代算法. 事实上, 这也就是非线性规划中求解函数最大值(或最小值)最典型的基本算法, 即 Gauss - Newton 迭代法.

(1) Gauss - Newton 迭代法

设  $X \sim f(x; \theta)$ ,  $L(\theta) = \log f(x; \theta)$ , 则极大似然估计  $\hat{\theta}$  满足以下必要条件(一般函数的最大值或最小值亦然):  
 $\nabla L(\hat{\theta}) = 0$

因此可视 $\theta_0$ 为初值, 设计以下迭代公式:

$\theta_1 = \theta_0 + [-Z(\theta_0)]^{-1} [Z(\theta_0)]$

$\theta_2 = \theta_1 + [-Z(\theta_1)]^{-1} [Z(\theta_1)]$

$\theta_{i+1} = \theta_i + [-Z(\theta_i)]^{-1} [Z(\theta_i)]$ , 其中  $Z(\theta) = \nabla L(\theta)$

直到  $\|\theta_{i+1} - \theta_i\| < \epsilon$ ,  $\epsilon$  为预定的充分小的正数, 如  $\epsilon = 10^{-8}$  等, 则取  $\theta_i$  作为极大似然估计 $\hat{\theta}$ 的近似值.

Gauss - Newton 迭代法有以下几个显著特点:

(a) 在一定正则条件下, 当  $i \rightarrow \infty$  时,  $\theta_{i+1}$  收敛到极大似然估计 $\hat{\theta}$  (b) 迭代过程强烈依赖于初值, 初值选取得当, 迭代收敛快; 反之则收敛慢, 甚至发散.

(c) 迭代过程中, 正定阵  $[-Z(\theta_i)]$  常可用 Fisher 信息阵  $I(\theta_i)$  代替, 不影响收敛性. 另外,  $[-Z(\theta_i)]$  也可用  $-Z(\theta_i) + cI$  或  $[-Z(\theta_i) + cI]$  代替 (其中  $I$  为单位矩阵,  $c$  为合适的常数), 迭代收敛过程类似.

(2) 改进的 Gauss-Newton 迭代法

改进的 Gauss - Newton 迭代法的基本想法是希望  $L(\theta_i)$  能尽快地向

以  $\hat{\theta}$  的最大值  $L(\hat{\theta})$  逼近. 直观上, 应当要求  $L(\theta_i) \rightarrow L(\hat{\theta})$  逐步增加, 直至  $\hat{\theta}$ . 但是上述普通的 Gauss-Newton 迭代法不一定具有单调递增性, 即以下不等式不一定成立:

在点  $\theta_0$  处展开可得

$L(e) = Z(\theta > 0) + Z(\theta = 0)$   
 $\frac{dL}{de}$   
 $\frac{dL}{d\theta}$

114 第三章点估计基本方法  $L(\theta_i + 1) = L(\theta_i + w) > L(\theta_i)$ ,  $D(e) = [$

在改进的 Gauss - Newton 迭代算法中, 改取

的  $A_j$  使得恒成立, 即  $LW$  保持递增性. 以下引理 说明, 存在  $A$ , 使这一递增性恒成立.

引理 3.3.1 若  $-1(\theta) > 0$  (即正定),  $L(\theta) \neq 0$ ,  $D(\theta) = [4(\theta)] - 1 [L(\theta)]$ . 则必存在  $A > 0$ , 使得当  $\delta = e + AD(\theta)$  时有

$L(\theta) = f(\theta + AD(\theta)) > L(\theta)$ .

证明 令  $g(A) = L(\theta + AP(\theta))$ , 则由微分公式有

$L(\theta) - L(\theta) = A - \theta = \theta A + aA = [g(\theta) + a]A$ ,  
 (3.3.10)

其中当时,  $a = 0$ . 为求  $g(A)$  的导数, 把向量和  $D(\theta)$  的第  $r$  个分量用其下标表示, 则以  $\theta$  可表示为

$= L(\theta) y D(\theta) = [L(\theta)]^T [ > 0$ . 因此当  $A$  充分小时必有  $\theta + a > 0$ , 由 (3.3.10) 式可得  $\theta - f(\theta) = [L(\theta) + a]A > 0$ . 即  $L(\theta) = L(\theta + AP(\theta)) > f(\theta)$ , 证毕.

■ 推论 在以上引理中, 若取  $D(\theta) = A(\theta) f(\theta)$ , 其中  $A(\theta)$  为正定 矩阵, 则也存在  $A > 0$ , 使得  $L(\theta) > L(\theta)$  成立; 特别, 取  $D(\theta) =$

$J'(\theta) L(\theta)$ , 上式亦成立.

由引理 3.3.1 可得以下改进的 Gauss - Newton 算法:

(a) 取初值  $e$  计算  $z = -1(\theta)$ , 其中  $p(\theta) = [ - Z(\theta) ] - 1$

取  $A$ . 使

$L(\theta + A D(\theta)) = \max L(\theta + \lambda D(\theta)) > L(\theta)$ .  $0 < A \leq 1$

由引理 3.3.1, 这样的  $A$  必存在.

(b) 取  $\theta = \theta + A D(\theta)$ , 计算  $D(\theta) = D(\theta)$ , 取  $\theta = \theta$

$L(\theta + A_j D_j) = \max f(\theta + AD, \theta) > L(\theta)$ .  $0 < A \leq 1$

(c) 取  $\theta = \theta + A D(\theta)$ , 计算  $D^2(\theta)$ , 求  $\lambda_2$

继续以上步骤, 则必有

$f(\theta) < L(\theta) < L(\theta_2) < \dots < L(\theta_i) < L(\theta_{i+1}) < \dots$

直到  $|L(\theta_{i+1}) - L(\theta_i)| < \epsilon$ ,  $\epsilon$  为预定的充分小的正数, 如  $\epsilon = 10^{-8}$ , 则可认为迭代收敛, 并取  $\theta_{i+1}$  为极大似然估计的近似值. ■

$f = f(\theta)$ ,  $A, 7?(\theta)$  (取适当

### 3.3 极大似然估计

115

例 3.3.11  $X_1, \dots, X_n$  独立同分布,  $X$  服从 Weibull 分布  $PT(A, a)$ , 即

$f(x; A, a) = A a x^{a-1} \exp(-A x^a)$ .

表 3.3.1 为 Zacks (1981) 提供的  $n=50$  的模拟样本, 其真参数为

$a = 1.750$ ,  $A = 1.000$ . 今应用 Gauss - Newton 迭代法 (GN) 和改进的 Gauss - Newton 迭代法 (MGN) 计算参数的极大似然估计. 取初值为  $a = 1.000$ ,  $A = 2.000$ , 其结果见表 3.3.2. 由计算结果可知, 改进的 Gauss - Newton 迭代法比较快, 但是最后结果相同. 从计算结果看,  $a$  的值 1.736 1 比较好; 而  $A$  的值 0.917 6 不够理想, 与真值有一定差距.



```

-
A
?
: . :
. s . * .
"»
:*y . . , : - «
?A- * - * . x > T V /
-<l w
»z'
<-. . • * t * n : 7 <
:# 4^l
.v. ?>/ . * v - c v .
: r i
-?-: z . ' .
i
1
2
3
4
5
6
7 8
2. 972 7
1.449 1
2. 254 2
1.693 9
2. 026 5
1. 806 7
1.942 7 1. 854 9
0. 537 3 0. 945 2 0. 734 2 0. 886 6 0. 797 9 0. 857 3
0. 820 9 0. 844 6
1.448 6
2.201 1
2. 370 3
2. 000 7
1.863 8
1.761 6
1. 738 2 1.736 1
0. 077 2
0.1474
0. 265 2
0. 448 9
0.670 1
0. 847 9
0.911 9 0.917 6
1.244 7 1.6298 1.730 7 1.736 1 1.736 1
0. 951 2 0.9654 0.918 2 0.917 6 0.917 6
0.586 9 0.291 1 0. 929 7 1. 110 6 1.256 4 1. 388 7 1.8924 0. 282 4
0. 441 2 1. 565 4
0. 135 1 1.0519 0. 558 6 0. 181 6 1.056 2 0. 224 5 1.369 4 1. 400 3
1.718 8 0. 366 8
表3. 3.1 0.762 1 0.6389 2.212 3 0.253 0
1. 357 6 0. 630 9 0.869 1 0. 573 9 0. 642 3 1. 681 3

```

Weibull分布模拟样本数据 1.304 7

1.4782

1. 373 3 0. 350 0 0.481 8 1. 314 5 0.906 1 0. 635 3 1. 568 3

0. 411 7

0.948 9 0.2067 0. 486 1 0. 759 2 1.259 1 2. 600 2 0. 539 8 0. 684 6

0.511 4 1. 440 4

a, A, a, A. a. A.

Weibull分布数据的Gauss - Newton迭代法 MGN法

表3. 3.2

Zacks方法 GN法

116 第三章点估计基本方法

续表 Zacks方法 GN法 MGN法

i A, A. a. A, 9 1. 910 1 0. 829 7 1. 736 1 0. 917 6

10 1.874 8 0. 839 2

### 3.4矩方程估计

矩估计是很古老的方法，其理论基础就是独立同分布情形下的大数定律，即观察值的样本平均趋向于总体平均。今具体叙述如下。

设 $X_1, \dots, X_n$ 独立同分布， $X_i \sim \theta$ 。为叙述方便，引

入以下记号： $\mu$ ：总体原点矩

其中 $M_1 = E(X_i)$ ；总体中心矩： $\mu_k = E(X_i - \mu)^k$  其中  $\sigma^2 = \text{Var}(X_i) = \mu_2$ ；

样本原点矩： $\bar{x}_k = \frac{1}{n} \sum_{i=1}^n X_i^k$ 。其中 $\bar{x}_1 = \bar{x}$ ； $\bar{x}_1 = 1$

样本中心矩： $s_k^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{x})^k$ 。  $i=1$

根据独立同分布随机变量序列的大数定律（见李贤平,1997；严士健，刘秀芳,1994），当 $n \rightarrow +\infty$ 时， $\bar{x}_k$ 以概率收敛（几乎处处收敛）到 $\mu_k$ ；

类似地，也有 $s_k^2$ 以概率收敛（几乎处处收敛）到 $\mu_k$ 。因此很自然地可以引入以下矩估计：

$\hat{\mu}_k = \bar{x}_k$

$\hat{\mu}_k = \bar{x}_k = \frac{1}{n} \sum_{i=1}^n X_i^k$

根据矩估计的定义可知， $\hat{\mu}_k$ 是 $\mu_k$ 的无偏估计（ $\mu_k$ 不一定无偏）；并且 $\hat{\mu}_k$ 、 $\mu_k$ 都是 $\theta$ 的相合估计且有渐近正态性（相合，即以概率收敛或几乎处处收敛，见第五章）。

矩估计还可进一步推广到矩方程估计。一般来说，参数 $\theta$ 或其函数 $g(\theta)$ 与总体原点矩和中心矩可能会有一定关系，例如可表示为以 $\theta$ 为

$\mu_1, \dots, \mu_k, \dots, \mu_{k-1}$ ，则可通过把总体原点矩和中心矩替换为样本原点矩和样本中心矩进行估计；就称为矩方程估计。这时 $g(\theta)$ 的矩估计

，

例 3. 4.1 Laplace 分布. 设 $X_1, \dots, X_n$ 独立同分布， $X_i \sim f_4(\theta)$ ，

为 $f_4(\theta) = G(\sim, \dots, a, b; m)$ ：

以下为若干例子。

### 3.4矩方程估计 117

a), 求 $\mu$ ,  $\sigma^2$ 的矩估计。

解 由第一章的公式可知 $M_i = E(X_i^i) = \mu_i$ ，因 $n$

此可 $\bar{x} = \bar{x}_1 = \mu_1$ ,  $\sigma^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{x})^2 = \mu_2 - \mu_1^2$

。这一结果与极大

似然估计有较大差别（见例3.3.5）。例3.4.2 $X_1, \dots, X_n$ 独立同分布， $X_i \sim \text{Poisson}(\lambda)$ ，求 $\lambda, \sigma^2$

的矩估计（或称矩方程估计）。

解 由 $\text{Poisson}$ 分布的性质，可得以下方程

$$M_i = E(\hat{\theta}_i) = a_i = \text{Var}(\hat{\theta}_i) = 1/n \cdot \text{Var}(\theta)$$

则通过把总体原点矩、中心矩替换为样本原点矩、中心矩可得

解方程即得  $\hat{\theta}$  的矩方程估计

注 对于  $r$  分布  $r(A, n)$ ，其一致最小风险无偏估计和极大似然估计都不易求出，因为其完备充分统计量（在应用上不很方便）

例 3.4.3 设  $X_1, \dots, X_n$  独立同分布，求其总体相关系数  $\rho$  的矩估计，其中

根据矩估计的替换原则可得

$$\hat{\rho} = \frac{\text{Cov}(\bar{X}, \bar{Y})}{\sqrt{\text{Var}(\bar{X}) \text{Var}(\bar{Y})}}$$

$$= \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \cdot \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

注意，矩估计有时不一定需要知道总体的分布，以上就是一个例子。另外，矩估计有某些不确定性；例如，若  $X_1, \dots, X_n$  独立同分布，

$$X_i \sim N(\mu, \sigma^2)$$

$$P = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\text{Var}(X) = \sigma^2$$

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}$$

$$= \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \cdot \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

又如，若  $X_1, \dots, X_n$  又独立同分布， $X_i \sim N(\mu, \sigma^2)$

$$E(X) = \mu, \text{Var}(X) = \sigma^2$$

$$E(X^2) = \mu^2 + \sigma^2$$

由  $E(X) = \mu$  和  $\text{Var}(X) = \sigma^2$  有  $E(X^2) = \mu^2 + \sigma^2$ 。因此共 2 个矩可以有以下两种矩

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

估计

$$\hat{\rho} = \frac{\text{Cov}(\bar{X}, \bar{Y})}{\sqrt{\text{Var}(\bar{X}) \text{Var}(\bar{Y})}}$$

例 3.4.4 设  $X_1, \dots, X_n$  独立同分布， $X_i \sim R(H)$ ，求  $\theta$  的矩估计。由均匀分布的性质，可得下方程：

$$E(X) = \frac{1}{2}, \text{Var}(X) = \frac{1}{12}$$

因此

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n X_i$$

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

第三章点估计基本方法

$$E(X) = \frac{1}{2}, \text{Var}(X) = \frac{1}{12}$$

例 3.4.4 方程估计。

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n X_i$$

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

$$\hat{\rho} = \frac{\text{Cov}(\bar{X}, \bar{Y})}{\sqrt{\text{Var}(\bar{X}) \text{Var}(\bar{Y})}}$$

其中  $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ ，联立求解以上方程得

$$\hat{\rho} = \frac{\text{Cov}(\bar{X}, \bar{Y})}{\sqrt{\text{Var}(\bar{X}) \text{Var}(\bar{Y})}}$$

另外，偏度系数  $\gamma_1 = \frac{a_3}{\sigma^3}$  和峰度系数  $\gamma_2 = \frac{a_4}{\sigma^4} - 3$ （其中  $a_k = E(X - EX)^k$ ）的矩估计可直接表示为

$$\hat{\gamma}_1 = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^3}{\left( \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \right)^{3/2}}, \hat{\gamma}_2 = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^4}{\left( \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \right)^2} - 3$$

$$\hat{\rho} = \frac{\text{Cov}(\bar{X}, \bar{Y})}{\sqrt{\text{Var}(\bar{X}) \text{Var}(\bar{Y})}}$$

习题三

1. 设  $X_1, \dots, X_n$  为  $\theta$  的样本，当  $X$  服从以下分布时，求参数  $\theta$  的 UMRUE，其中  $\theta > 0$ 。

(1)  $X$  服从 Rayleigh 分布  $f(x) = 2x e^{-x^2/2}, x > 0$ ; (2)  $X \sim \chi^2(2)$ ;  
 (3)  $X \sim f(x), 0 < x < 1$ .

2. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X$  服从 Poisson 分布  $P(\lambda)$ , 求  $\lambda$  的 UMRUE.

3. 设  $X_1, \dots, X_n$  为 i.i.d. 样本, (1)  $X \sim \Gamma(m, \theta)$ , 求  $\theta$  的 UMRUE;

(2)  $X \sim \Gamma(m, \theta)$ , 求  $\theta$  的 UMVUE. 4. 设  $X_1, \dots, X_n$  为样本,  $X$  服从  $r(A, p)$ ,  $p$  为已知, 求  $A$  的 UMRUE.

(2)  $X \sim$

知, 求  $A$ ,  $A^{-1}$ ,  $A^2$  的 UMRUE.

5. 设  $X_1, \dots, X_n$  为 Z.i.d. 样本,  $X$  服从 Laplace 分布, 即  $X \sim$

习题三

119

$LA(0, a) \sim T^2 e^{-t^2/a^2}$ , 其中  $a > 0$ , 求  $a$ ,  $cr_1, a$  的 UMRUE. 2. 设

6. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X$  服从指数分布  $f(x) = (1/a) e^{-x/a}$ , 求  $M$ ,  $a$ ,  $a$  的 UMRUE.

7. 设  $V$  服从截尾二项分布, 其密度函数为  $p(x) =$

$(1-p)^n \binom{n}{x} p^x = \binom{n}{x} p^x (1-p)^{n-x}, x=0, 1, \dots, n$ . (1) 证明:  $T=X$  关于  $\theta$  为完备的充分统计量;

(2) 求  $T$  的 UMVUE.

8. 设  $A_1, \dots, A_n$  为 i.i.d. 样本,  $X$  服从两点分布, 即  $P(X=1) = \theta, 0 < \theta < 1$ .

(1) 求  $a = \theta m$  的 UMRUE, 其中  $m$  为正数;

(2) 求  $\theta = P(X_1 + X_2 + \dots + X_m = k)$  的 UMRUE, 其中  $0 < \theta < 1$ ;

(3) 求  $\theta = P(X_1 + X_2 + \dots + X_n > 1)$  的 UMRUE.

9. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X$  服从指数分布  $r(A, 1)$ , 求  $e^{-A}$  的 UMRUE.

10. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X$  服从几何分布, 即  $P(X=1) = \theta(1-\theta)^{x-1}, 0 < \theta < 1, x=1, 2, \dots$ .

(1) 求  $\theta$  的 UMRUE;

(2) 求  $\theta^2$  的 UMRUE.

11. 设  $X_1, \dots, X_n$  相互独立,  $X_i \sim \Gamma(1, a_i)$ , 一切  $i=1, \dots, n$ , 而  $X \sim \Gamma(y, a)$ ;  $X_i \sim \Gamma(1, a_i)$ , 求参数  $y, a$  的 UMRUE.

12. 设  $U_1, \dots, U_n$  相互独立,  $X_i \sim \Gamma(1/a, 1)$ , 一切  $i=1, \dots, n$ , 而  $X \sim \Gamma(1/a, 1)$ ,  $y > 0$ , 求  $y, a$  的 UMRUE.

13. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X$  服从 Pareto 分布  $PR(a, \theta)$ , 即密度函数为  $f(x) = \frac{\theta^a}{\Gamma(a)} x^{-a-1}, x > 0$ ,

(1) 若  $\theta$  已知, 求  $a$  的 UMRUE;

(2) 若  $a$  已知, 求  $\theta$  的 UMRUE;

(3) 若  $a, \theta$  都未知, 求  $a$  和  $\theta$  的 UMRUE (提示: 作变换  $K_i = \log X_i, Z = 1, 2, \dots, n$ , 并利用指数分布的性质).

14. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $g(\cdot)$  为  $(0, +\infty)$  上的可微函数.

(1)  $X \sim \Gamma(1, 1)$ , 求  $g(1/x)$  的 UMRUE; 并由此求  $M_2$  的 UMRUE;

(2)  $X \sim \Gamma(1, 1)$ , 求  $g(1/x)$  的 UMRUE; 并由此求  $\theta$  的 UMRUE;

(3)  $X \sim \Gamma(1, 1)$ , 求  $g(1/x)$  的 UMRUE; 并由此求  $\theta, \theta^2$  的 UMRUE.

UMRUE (提示:参见例3.2.5).

\*15. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, 1)$ ,  $-\infty < \mu < \infty$ ,  $\Phi(\cdot)$  为标准正态分布的分布函数. 求  $A(n)$ , 使  $\Phi(A(n))$  为  $\mu$  的 UMRUE. 并由此求  $P(X_1 < a)$  的 UMRUE, 其中  $a$  为给定常数 (提示: 利用第一章习题16的结果).

\*16. 设  $U_1, \dots, U_n$  为 i.i.d. 样本,  $X_i$  服从幂级数分布, 即  $P(X_i = x) = p(x) r / c(x)$ ,  $X_i = 0, 1, 2, \dots$  其中  $p(x)$  为正则化系数, 是已知的,  $\theta$  为未知参数. 求  $\theta r / [c(\theta)]$  的 UMRUE (提示: 记  $P(T = \theta) = p_n(\theta) / [c(\theta)]$ , 其中  $i = 1, \dots, n$ ,  $V$  为正则化系数. 由  $\sum_{i=1}^n P(T = \theta) = 1$  可得  $[c(\theta)] = 1$ ).

17. 设

$(X_1, Y_1), \dots, (X_n, Y_n)$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ,  $Y_i \sim N(\mu, \sigma^2)$ ,  $\mu, \sigma^2$  未知.

(1) 求关于  $y$  的 CSS;

利用此级数展开式).

$r = 0$  为样本,  $(U$

$\dots, (X_n, Y_n)$

) 的分布密度为

(2) 求  $y$  的 UMRUE (提示: 作变换  $x_i = r \cos(\theta_i)$ ,  $y_i = r \sin(\theta_i)$ ;  $i = 1, 2, \dots, n$ ).

18. 设  $y_1, y_2, \dots, y_n$  为 i.i.d. 样本,  $X_i \sim E(a, \lambda)$ ,  $Y_i \sim E(b, \lambda)$ ,  $X_i, Y_i$  为 i.i.d. 样本,  $K, \dots, K_n$  为

$E(\lambda)$ ; 其中  $a > 0$ ,  $b > 0$ , 且  $X_i, Y_i$  相互独立,  $Z = 1, \dots, n$ . (1) 求  $X/Z$  和  $Y/Z$  的 UMRUE;

(2) 当  $c_{X/Z} = a/b$  时, 求  $X/Z$  的 UMRUE.

19. 设  $A_1, \dots, A_n$  为 i.i.d. 样本,  $X_i \sim U(0, 1)$ ,  $Y_i \sim U(0, 1)$ ,  $X_i, Y_i$  为 i.i.d. 样本,  $y$  服从均匀分布  $U(0, 1)$ ; 其中  $a > 0$ ,  $b > 0$ ,

且  $X_i, Y_i$  相互独立,  $f = 1, \dots, m$ ;  $J = 1, \dots, t$ . 当  $a > 1$  时, 试求  $\theta_X / \theta_Y$  的 UMRUE.

20. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ;  $y_1, \dots, y_n$  为

$\mu$  的 UMRUE.

### 习题三

121

i.i.d. 样本,  $Y_i \sim N(\mu, \sigma^2)$ , 且  $X_i, Y_i$  相互独立,  $f = 1, 2, \dots, n$ ,  $y = 1, 2, \dots, a$ . 设  $\mu, \sigma^2$  未知,  $a > 0$ ,  $b > 0$ , 且  $X_i, Y_i$  相互独立,  $Z = 1, \dots, n$ . (1) 求  $X/Z$  和  $Y/Z$  的 UMRUE;

(2) 设  $c_{X/Z} = a/b$  时, 求  $X/Z$  的 UMRUE.

(3) 设  $c_{Y/Z} = b/a$  时, 求  $Y/Z$  的 UMRUE.

21. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $E(X_i) = \mu$ , 记  $M$  的线性估计类为  $\{T(X) : T(X) = \sum_{i=1}^n a_i X_i\}$

(1) 证明:  $\mu$  的无偏估计的充要条件是  $\sum_{i=1}^n a_i = 1$ ;

(2) 在线性无偏估计类中求  $M$  的 UMVUE.

22. 设  $g(X)$  是参数  $g(\theta)$  的一个无偏估计, 证明:  $g(\theta)$  的任一

估计  $i(X)$  可表示为  $i(X) = g(X)$  其中  $g$  是  $\theta$  的某一

无偏估计. 23. 设  $\theta = E[g(X)]$ ,  $ES[g(X)] = \theta$ ,  $Var[g(X)] < +\infty$ , 对一切  $\theta \in \Theta$ ,  $i$  为  $g(\theta)$  的无偏估计.

(1) 证明:  $i(X)$  为  $g(\theta)$  的 UMVUE 的充要条件是  $E[i(X)g(X)] = \theta$ ,  $\forall g(X) \in \Theta$ , 设  $e(\theta)$  (提示: 用反证法);

(2) 若  $i(X)$  为  $g(X)$  的 UMVUE,  $i(X)$  为  $g(X)$  的任一无偏估计, 则必有  $\rho(i(X), g(X)) = 1$ , 其中  $\rho(\cdot, \cdot)$  表示相关系数;

$c_j$  是  $i$ -

24. 设  $X_1, X_2, \dots, X_n$  为 i. i. d. 样本, 当  $X$  的密度函数  $f(x; \theta)$  为下列各种情况时, 分别求  $\theta$  的 MLE.

(1)  $f(x; \theta) = \frac{1}{\Gamma(\theta)} x^{\theta-1} e^{-x}$ , 其中  $\theta > 1$ ; (2)  $y(X_j = 0 \text{ 或 } 1) = \theta^j (1-\theta)^{n-j}$ , 其中  $0 < \theta < 1$ ;

(3) 若  $r_i$  分别为  $g_i(\theta)$  的 UMVUE,  $i = 1, 2, \dots, n$ , 则  $S_K$  为  $g_K(\theta)$  的 UMVUE, 其中  $K$  为常数,  $1 \leq K \leq n$ .

1=1

n

1

(3)  $f(x; \theta) = \frac{1}{\Gamma(\theta)} x^{\theta-1} e^{-x}$ ;

(4)  $\theta$

$a x^{a-1} e^{-x}$ , 其中  $a > 0$ ,

$+ U$ , 其中  $U \sim \text{Exp}(\lambda)$ .

总为 i. i. d. 样本, 当  $X$  的密度函数  $f(x; \theta)$  为下

列各种情况时, 分别求  $\theta$  的 MLE.

25

(1)  $f(x; \theta) = \frac{1}{\Gamma(\theta)} x^{\theta-1} e^{-x}$ , 其中  $\theta > 0$ ;

122

第三章点估计基本方法

$f(x; \theta) = \frac{1}{\Gamma(\theta)} x^{\theta-1} e^{-x}$ ;

$\theta > 0$ ,  $\lambda > 0$ .

(2)

(3)  $f(x; \theta) = \frac{1}{\Gamma(\theta)} x^{\theta-1} e^{-x}$ ;

$f(x; \theta) = \frac{1}{\Gamma(\theta)} x^{\theta-1} e^{-x}$ ;

»./ (01

26. 设  $X_1, X_2, \dots, X_n$  为 i. i. d. 样本, 应用正文中定理 3.3.1 关于指数族分布 MLE 的结果求以下分布中相应参数的 MLE.

(1)  $X_i$  服从正态分布  $N(\mu, \sigma^2)$ ;

(2)  $X_i$  服从 Pascal 分布  $PA(r, \theta)$ , 但  $r$  已知; (3)  $X_i$  服从负二项分布  $NB(r, \theta)$ , 但  $r$  已知; (4)  $X_i$  服从  $r(A, \theta)$ , 但  $p$  已知.

27. 设  $X_1, X_2, \dots, X_n$  相互独立, 在下列情形下求相应参数的 MLE:

(1)  $X_i \sim N(\mu, \sigma^2)$ , 而  $\mu$  未知;

(2)  $X_i \sim N(\mu, \sigma^2)$ , 而  $\sigma^2$  未知;

(3)  $X_i \sim N(\mu, \sigma^2)$ , 而  $\mu, \sigma^2$  均未知;

其中  $y > 0, w > 0, T$  都是未知参数.

28. 设  $X_1, X_2, \dots, X_n$  又相互独立, 在下列情形下求相应参数的 MLE: (1)  $X_i \sim r(A, \theta)$  一切  $i$ , 而  $\theta \sim r(y, \lambda)$ ;

(2)  $X_i \sim r(A, \theta)$  一切  $i$ , 而  $X_i \sim P(y, \lambda)$ ;

(3)  $b(\lambda, \theta)$  一切  $i$ , 而  $X_i \sim b(1, \theta)$ ;

其中  $y > 0, A > 0, 0 < \theta < 1$  为未知参数.

29. 设随机变量  $A$  只能取 4 个值, 每种取值出现的概率分别为  $p_1, p_2, p_3, p_4$

的次数分别为  $V_1, V_2, V_3, V_4$ , 其中  $V_1 + V_2 + V_3 + V_4 = n$ , 求  $\theta$  的 MLE.

30. 设  $X = (X_1, X_2, X_3, X_4)^T$  服从 4 项分布, 即  $X \sim W(n, \pi)$ , 其中  $n$  已知,  $\pi = (a, b, c, d)^T$ , 求  $a, b, c$  的 MLE.

31. 设做了三个相互独立的试验, 每个试验都服从两点分布, 第一个试验共做了  $n_1$  次, 成功的次数为  $X_1$ ; 第二个试验共做了  $n_2$  次, 成功的次数为  $X_2$ ; 第三个试验共做了  $n_3$  次, 成功的次数为  $X_3$ . 设三次试验中成功的概率分别为  $a, a+p, a+2p$ , 求  $a, p$  的 MLE.

f

,  $Y$ ,  $f$ , 其中  $0 < \theta < 1$ . 若观察到  $n$  个样本, 4 个值出现

(1)

$n_1, n_2, n_3, n_4$

32. 设  $TV = (N_i)$  服从多项分布  $MN(n, \pi)$ ;  $\pi = (\pi_1, \dots, \pi_r)$ ;  $j$

-

### 习题三

123

1.  $Y_1, \dots, Y_n$  满足关系:  $Y_i = \pi_i$ ,  $i = 1, \dots, n$ ;  $j = 1, \dots, r$ , 其中  $\pi_j = \pi_1 + \dots + \pi_r = 1$ . 求  $\pi_j$  的 MLE.

33. 设  $A_1, \dots, A_n$  为 i.i.d. 样本,  $A$  的概率密度函数为

$f(x; \theta) = 2x\theta e^{-2x\theta}$ ,  $\theta > 0$ .

(1) 试求具有上述概率密度函数的分布函数的中位数的 MLE; (2) 证明 (1) 中的估计是极小充分统计量.

34. 设  $A_1, \dots, A_n$  为 i.i.d. 样本,  $X$  服从指数分布  $y \sim \text{Exp}(\lambda)$ ,  $P(X \leq a) = a$ .

(1) 若  $a$  已知求  $\lambda$  的 MLE; (2) 若  $\lambda$  已知求  $a$  的 MLE.

35. (1) 设  $X \sim N(\mu, \sigma^2)$ , 其中  $(\mu, \sigma^2)$  未知, 求  $(\mu, \sigma^2)$  的 MLE;

(2) 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $TV(\lambda, 1)$ , 当  $M \in [0, +\infty)$  时, 求  $M$  的 MLE; 当  $M \in (-\infty, 0)$  时, 求  $M$  的 MLE (提示: 对  $X$  的取值情况进行讨论).

36. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X$  服从 Cauchy 分布, 即  $f(x, \theta)$

$f(x, \theta) = \frac{1}{\pi} \frac{\theta}{1 + \theta^2 x^2}$ ,  $\theta > 0$ .

(1) 当  $a = 1$  时, 证明  $\theta$  的 MLE 为  $\sqrt{n}$ ; (2) 当  $n = 2$  时, 证明  $\theta$  的极大似然估计存在且唯一, 但似然方程

有多个根.  $X$  为 i.i.d. 样本,  $X$  的密度函数为

$f(x, \theta) = \frac{1}{\pi} \frac{\theta}{1 + \theta^2 x^2}$ ,  $\theta > 0$ .

设  $X_1, \dots, X_n$  为  $m$  样本,  $X$  服从 Poisson 分布  $P(\theta)$ , 且  $A$  和  $B$  相互独立,  $i = 1, \dots, m$ ,  $j = 1, \dots, n$ , 其中未知参数  $\theta > 0$ .

(1) 利用题中所给的  $n+m$  个样本写出  $\theta$  的似然方程;

(2) 证明 (1) 中的方程有且只有一个允许解.

\*38. 设  $X_1, X_2, \dots, X_n$  为 i.i.d. 样本,  $A$  的分布密度函数为  $f(x, \theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ , 其中  $\mu, \sigma^2$  未知. 证明:  $(\mu, \sigma^2)$  的极大似然估计不存在 (提示: 证当  $m = \infty$  时).

37. 设  $Y_1, Y_2, \dots, Y_n$

### 124 第三章点估计基本方法

$A_1, \dots, A_n$  且  $a_2 \rightarrow 0$  时, 似然函数  $L(x, a_2) \rightarrow \dots$

39. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ,  $i = 1, 2, \dots, n$ , 且当  $j \neq i$  时  $X_i$  与  $X_j$  相互独立. 求  $(\mu, \sigma^2)$  的 MLE.

40. 设  $U_1, \dots, U_n$  为 i.i.d. 样本,  $X = \sum_{i=1}^n U_i$ ; 为

i.i.d. 样本,  $K_i \sim \sqrt{V(x_2, T_2)}$ . 求  $c = M_i^{-2}$  的 MLE; 若样本容量之和  $a + \sqrt{n}$  固定, 久和  $\sqrt{2}$  已知, 如何调整  $n$  和  $\sqrt{n}$  的比例, 使  $c$  的 MLE 的均方误差最小?

41. 设观察样本  $X_1, X_2, \dots$ , 又满足关系  $X_t + r = l, -n$ ,  $X_0 = 0$ . 其中  $u_1, u_2, \dots, u_n$  为 i.i.d. 未知随机变量,  $u' \sim iV(0, (T_2))$ , 求  $\theta$  和  $a_2$  的 MLE.

42. 设  $X_1, X_2, \dots, X_N$  为 i.i.d. 样本,  $X$ : 服从 Poisson 分布  $P(1)$ ,  $A > 0$ . 若仅观察到  $X_1, \dots, X_A$ . 中前个样本  $X_1, X_2, \dots, X_n$  的值, 以  $N$  及后面  $N-n$  个样本的和  $\sum_{i=1}^N X_i$ , 求  $A$  的 MLE.  $i = n+1$

43. 设某种家禽生下的蛋的数目是一个随机变量  $2V$ , 其中  $TV \sim P(A)$ ,  $A$  未知. 若每个蛋能发育成小动物的概率是  $p$ , 且各个蛋能否发育成小动物是彼此独立的; 若记  $\tau$  为发育成小动物的个数.

(1) 求  $(W, A)$  的联合分布;

(2) 设  $(TV_1, \dots, TV_s), (Ns, Ms)$  是来自 (1) 中分布的  $s$  个随机样本, 求参数  $(A, p)$  的 MLE.

44. 设  $X_1, X_2, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim (x, \theta | x > 0)$ , 其中未知参数  $\theta > 0$ . 若仅观察到的值, 以及  $X_i^{(k)} = L(m+1, \dots, i, n)$ ; 试求  $\theta$  的 MLE.

\*45. 设随机变量  $X$  服从超几何分布  $HG(b, N, n)$ , 即

其中  $V, n$  均为非负整数. 求证: 当  $V, n$  固定时, 参数  $b$  具有以下极大似然估计:

$$\hat{b}(X) = \frac{V}{-(7V+1)e^{\frac{1}{n}} - (N+1)GMn}$$

其中  $[\cdot]$  表示整数部分 (提示: 求  $L(b+1, x)/L(b, x)$ , 其中  $L(b, x)$  为关于  $b$  的似然函数).

### 习题三

125

46. 设  $U_1, \dots, U_n$  为 i.i.d. 样本,  $X$  服从几何分布,  $P(X = k) = e^{-k} - e^{-(k+1)}$ ,  $k = 1, 2, \dots$ . 若仅观察到满足以下条件的样本  $r_1, r_2, \dots, r_n$ : 当时,  $r_i = X_i$ ; 当  $X_i^{(r+1)}$  时, 则  $r_i = r+1$ ,  $i = 1, 2, \dots$ . 若  $r$  为某个固定的正整数, 且共有  $m$  个样本 (可设最后  $m$  个样本) 的值为  $r+1$ , 设  $r_n^{(r+1)}$ , 求  $\theta$  的 MLE.

\*47.

MLE, 若

为 i.i.d. 样本,  $(y^* z_j)$  的密度函数为  $f(y, z) = A e^{-Ay} e^{-Bz}$ , 其中  $A > 0, B > 0$ , 求  $(A, B)$  的

(1) 观测到的样本为

(2) 观测到的样本不是  $(y_1, z_1), \dots, (y_n, z_n)$ , 而是  $(y_1, z_1), \dots, (y_n, z_n)$ , 其中

$y_i = \min(y_i, z_i)$ ; 若  $y_i = z_i$ ,  $y_i = z_i$ ; 若  $y_i < z_i$ ,  $y_i = y_i$ ; 若  $y_i > z_i$ ,  $y_i = z_i$ . (提示: 密度函数可表示为  $f(y, z; A, B) = A e^{-Ay} e^{-Bz}$ ,  $y = y_i, z = z_i$ ,  $i = 1, 2, \dots, n$  (提示: 密度函数可表示为  $f(y, z; A, B) = A e^{-Ay} e^{-Bz}$ ,  $y = y_i, z = z_i$ ,  $i = 1, 2, \dots, n$ ).

48. 设  $F$  服从幂级数分布,  $BPP(y = y) = ay^b/c$  (权),  $y = 0, 1, 2, \dots$ , 其中  $\theta > 0$ ,  $(a, b)$  为非负序列且与  $\theta$  无关. 设  $Y_1, \dots, Y_n$  为 i.i.d. 样本,  $K$  服从上述幂级数分布. 证明:  $\theta$  的极大似然估计满足  $P = oc(o)/c(e)$ .

49. 设  $r_1, \dots, r_n$  为 i.i.d. 样本,  $Y_x$  的密度函数为  $f(y) = K e^{-y} g(y)$ , 其中  $g(y) > 0$ ,



$g(\cdot)$  为非负可积实函数.

(1) 证明尺(幻为没的增函数); (2) 试求参数 $\theta$  的 MLE;

)I, 其中

50. 设 $X_1, \dots, X_n$ 为i. i. d.样本, 当戈的密度函数为下列各种情况 时, 求相应参数的矩估计:

(1)  $X \sim \text{Exp}(\theta)$ ,  $\theta > 0$ ; (2)  $X \sim \text{Gamma}(\theta, 1)$ ,  $\theta > 0$ ;

(3)  $X \sim \text{Beta}(\theta, 1)$ ,  $0 < \theta < 1$ ;

(4)  $X \sim \text{Beta}(\theta, \theta)$ ,  $\theta > 0$ ;

(5)  $X \sim \text{Beta}(\theta, 1)$ ,  $\theta > 0$ .

51. 设 $X_1, \dots, X_n$ 为i. i. d.样本, 当戈的密度函数为下列各种情况

(3) 若广的密度函数为 $L(e)g(y|a)$  为常数, 研究类似的问题.

126 第三章点估计基本方法

时, 求相应双参数的矩估计:

(1)  $X \sim \text{Gamma}(\theta, \lambda)$ ; (2)  $X \sim \text{Beta}(\theta, \lambda)$ ;

■  
/ $U^{\theta}$ );

(4)  $X \sim \text{Pareto}(\theta, \lambda)$ :  $A_6 > a/(x, +(9)a+1Z \setminus x\{ > 0 \}$ ;

(5)  $X \sim \text{Pascal}(\theta, r)$ :  $( * j^r(l x) = r, r+1,$   
没 $e(0, l)$ ,  $r$ 为正整数.

52. 设 $X_1, \dots, X_n$ 为i. i. d.样本,  $X$ 的密度函数为 $P\{X' =$   
 $-1 \ n*$

(3)  $X \sim \text{LogNormal}(\mu, \sigma^2)$ :  $\exp(\mu - \frac{\sigma^2}{2})$ ; 广

$\ln(1 - p) \sim k_9$ 女=1,

$X_{s2} + (\tau)^2$ , 其中

2, ..., 其中 $0 < p < 1$ , 证明: $p$ 的矩估计为 $p = 1 -$

又  $sa^{(-x)}2. n 1=1 Tl J=1$

。 1 1 r(In-u)2

第四章最优同变估计

第三章介绍的一致最小风险无偏估计(UMRUE)是在无偏性限制下 求最优解, 本章介绍的最优同变估计或最小风险同变估计(MREE)是在 另一种常见的限制, 即变换群的同变性限制下求最优解. 本章第4.1节 首先介绍同变性、同变估计和最优同变估计的概念, 然后分三节介绍位置尺度参数分布族的最优同变估计. 其中第4.2节介绍平移变换群下位置参数的最优同变估计, 第4.3节介绍相似变换群下尺度参数的最优同变估计, 第4.4节介绍线性变换群下位置尺度参数的最优同变估计. 给定样本 $X_1, \dots, X_n$ , 记 $X = (X_1, \dots, X_n)^T \sim (x, a)$ ,  $\theta \in \Theta$  未知参数没的估计记为 $\hat{\theta}(X)$ 或 $\hat{\theta}(x)$ , 的估计记为 $\hat{a}(X)$ 或 $\hat{a}(x)$ . 本章把#(幻的范围由 $p =$  的一切估计! 缩小为 $A = u(\theta)$ 的一切同变估计1(第三章的UMRUE则考虑 $\Delta$ 为 $g(\theta)$ 的一切无偏估计). 另外, 本章仅限于均方损失函数情形下求最优解, 并且主要考虑参数6本身的MREE, 函数的 MREE限于少数特殊情形. 有关本章内容, 可参见陈希孺(1981, 1999), Lehmann and Casella(1998), Zacks(1981), 茹诗松等(1998).

4.1变换群下的同变估计

4.1.1同变性概念

同变估计, 即在某种变换群下保持同变的估计, 首先看两个例子. 例4.1.1位置参数与平移变换. 设 $X_1, \dots, X_n$ 独立同分布, 戈-

为位置参数, 则4也表示位置. 位置参数应该具有以下性质: 即观察值\的起点变化(即平

移), 则参数的起点也应有相应的变化(平移). 例如温度, 起点由原来的 $\theta_0$ 变为 $\theta_0^\circ\text{K}$ , 则  $M$  都应同时

— 有位置参数的平移变换+ 特别,  $M$ 的估计也应当有同样的

变化, 即/!(幻应满足

$$\int^{\infty} (x_1 + C, \dots, x_n + C) = \int^{\infty} f(x_1, \dots, x_n) + C$$

若记 $1 = (1, \dots, 1)^T, X = (x_1, \dots, x_n)^T$ , 则上式化为

$$\int^{\infty} L(x + c) = \int^{\infty} L(x) + C, \quad \forall c, \quad (4.1.1)$$

改变. 即由观察值的平移变换%

+  $c; i = 1, \dots, n$  相应的也应

128

#### 第四章最优同变估计

因此位置参数 $M$ 的估计 $A$ (幻应满足以上平移同变方程, 即同变条件. 我们可以根据同变方程

(4.1.1)求最优同 $f$  估计. 由于好的估计应 为充分统计量的函数, 本例的充分统计量为

无, 因此可假设 $L(x) = \int^{\infty} f(x)$ , 并根据方程(4.1.1)求出最优的 $\hat{\theta}$ . 由 $x = x + ci$ 可得 $I -$

$J + c$ , 因此由(4.1.1)式及  $L(x) = c p(x)$ 可知  $(p(x + c) = (p(x) + C, \quad \forall c,$

上式对任意的 $c$ 都成立, 可. 取 $c = -i$ , 则有 $(p(0) = (p(x) - X, \text{即} < p(\%) = S + \hat{\theta}) = \hat{\theta}(x)$ . 为

进一步求解, 需要确定损失函数. 本章主要考虑均方

损失, 则可应用均方误差准则 $f$  出 $\hat{\theta}(0)$ 如下

$$MSE = E[\hat{\theta} - \theta]^2 = E[(X - \hat{\theta})^2] = \int^{\infty} (X - \hat{\theta})^2 f(X) dX,$$

显然当 $\hat{\theta}(0) = 0$ 时上式最因此最有

$$\hat{\theta}(X) = X + \hat{\theta}(0) = X.$$

以上讨论可推广到一般情形. 设 $X = (X_1, \dots, X_n)^T, f(x) = (f_1, \dots, f_n)^T$ . 若 $X$ 服从位置参

数分布族

... 对于该分布族, 如果 $\theta$ 的起点改变, 则  $\theta$  也应有相应的改变, 即由观

察值的平移变换 $X \rightarrow X' = X + c, i = 1, \dots, n$ , 相应地也应有位置参数 的平移变

换 $\theta \rightarrow \theta' = \theta + c$ , 并且 $\hat{\theta}$ 的估计 $\hat{\theta}$ 亦应该有同样的变化

$$\hat{\theta}(x + c, \dots, x_n + c) = \hat{\theta}(x) + c, \quad (4.1.2)$$

因此位置参数 $e$ 的估计 $\hat{\theta}$ (幻

应满足以上平移同变方程, 即同变条件.

满足该方程的 $\hat{\theta}$ (幻的全体构成“平移变换群下的同变估计类”, 因此可 在此估计类中求风险

最小的解(MREE).

例4.1.2尺度参数与相似变换. 设 $X_1, \dots, X_n$ 独立同分布,  $X_i \sim N(0, \sigma^2)$ ,  $\sigma$ 为尺度

参数, 则 $\sigma$ 也表示尺度. 它们应满足如下关系: 若 $\sigma$ 的尺度单位改变, 则 $\sigma$  的尺度单位也应

有相应的改变. 例如 $\sigma$ 原来用 $\text{cm}$ 度量, 改为用 $\text{mm}$ 度量, 则 $\sigma$  也应有同样的改变. 即由观察值

$\sim$ 的相似变换 $x_i = kx_i, i = 1, \dots, n$ , 相应地也应有尺度参数的相似 变换 $\sigma \rightarrow \sigma' = k\sigma$ . 特别, 的估

计 $P(X)$ 也应当有同样的变化, 即

$$P(kx_1, \dots, kx_n) = P(x_1, \dots, x_n), \quad k > 0, \quad (4.1.3)$$

(4.1.3)式即为 $\sigma^2$ 的

梦计应该满足的同变方程. 我们可根据此同变方

程求解最优同变估计 $P(X)$ . 同理, 好的估计应为充分统计量的函数,  $n$

本例的充分统计量为 $S^2 =$

$$S^2 = \frac{1}{n} \sum_{i=1}^n x_i^2.$$

因此可设 $P_b): \sigma, \dots,$

$$P_b(x) = \frac{1}{\sigma} f\left(\frac{x}{\sigma}\right),$$

$$f(x) = \frac{1}{\sigma} f\left(\frac{x}{\sigma}\right),$$

$$i=1$$

再根据(4.1.3)式求最优解. 由  $x_i = kx_i; i = 1, \dots, n$  可得  $S^2 \rightarrow k^2 S^2$ ,

因此 $\hat{\theta}(x) = < p(S^2)$ 代入(4.1.3)式有 $< p(k^2 S^2) = k^2 < p(S^2), \quad \forall k > 0$ . 该式

:

#### 4.1 变换群下的同变估计

129

对任意的 $\delta$ 都成立, 可取 $k = s - \delta$ 则有

$$1) = S^{-2}(p(S^2)), \quad p(S^2) = S^2(p(1)) = bS^2,$$

其中 $b = c^{(1)}$ . 因此可根据均方误差最小准则求 $\delta(\cdot)$ , 使 $MSE = E[\delta^2]$ 达到最小. 而

$$MSE = E[(\delta S^2 - c^2)^2] = E[(bS^4 - 2b\delta S^2 + \delta^2)]$$

$= 6(2n + n^2)c^4 - 2bn\delta^2 + c^4$ , 以上 $MSE$ 为 $\delta$ 的二次三项式,  $b = (n+2)^{-1}$ 时 $MSE$ 最小. 因此由 $P(\cdot)$

$$p(S^2) = bS^2, \text{ 最后可得 } p(x)$$

这一估计在 $S^2$ 类型的估计中, 均方误差最小, 因而优于一致最小风险 无偏估计以及极大似然估计(相当于 $\delta = n^{-1}$ ).  $\square$

以上讨论可推广到一般情形. 若 $V$ 服从尺度参数分布族

对于该分布族, 若 $\lambda$ 的尺度单位改变, 则 $a$ 的尺度单位也应有相应的 改变. 即由观察值的相似变换 $x_i \rightarrow \lambda x_i; i = 1, \dots, n$ ,  $a$ 相应的也应有 尺度参数的相似变换 $a \rightarrow \lambda a$ . 特别,  $a$ 的估计 $\hat{a}$ 也应当有同样的变化, 即  $\hat{a} \rightarrow \lambda \hat{a}$  (幻满足

$(r(A: \lambda \cdot), \dots, A: \lambda \cdot) = \lambda (r(x_1, \dots, x_n))$  或  $(t(H)) = \lambda (x)$  因此尺度参数 $a$ 的估计 $\hat{a}$  (幻应满足以上相似同变方程, 即同变条件. 满足该方程的 $\hat{a}$  (幻的全体构成“相似变换群下的同变估计类”, 因此 可在此估计类中求风险最小的解(MREE).

##### 4.1.2 同变统计判决函数

以上通过两个例子初步介绍了在平移变换群和相似变换群下位置参数和尺度参数的同变估计, 以及最优同变估计的求解方法. 这些论述可 推广到更一般的变换群和分布族. 为了从统计判决观点讨论同变性、同 变估计和最优同变估计, 以下首先介绍样本空间的变换群以及参数空间和判决空间的导出群. 在此基础上, 引入同变统计判决函数和同变估计 的定义. 同变统计判决是对统计判决的一种限制, 在此限制下求最优解 (关于某种损失函数), 就得到最优同变判决函数. 本书仅讨论若干常

见情形下的最优同变估计(见本章4.2-4.4节); 至于其他同变统计判 决问题, 例如最优同变检验, 可参看 陈希孺(1981, 1999)、Lehmann (1986)或 Shao(1999)等.

#### 130 第四章最优同变估计

##### (1) 样本空间的变换群和不变分布族

给定样本 $x$  及其概率测度空间 $I \sim \mathcal{I}$  或简记为

$\mathcal{I}$ . 同时给定 $\mathcal{I}$ 上的可测变换群 $G = \{g\}$  其中 $g$ 为 $\mathcal{I}$  上的可测变换:

$x' = gx$  若 $A \in \mathcal{B}_X$ . 则 $0 \leq g(A) \leq 1$ .  $G = \{g\}$ 为变换群, 即恒等变换 $e \in G$ , 且若

$g \in G$ , 则 $g^{-1} \in G$ , 对乘法封闭: 若 $g_1, g_2 \in G$ , 则 $g_1 \circ g_2 \in G$ , 并且满

足结合律:  $g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3$ . 样本空间上常见的变换 群有: 平移变换群、相似变换群、线性变换群以及对称变换群等.

给定样本空间上的变换群, 我们要进一步研究: 参数空间以及判决 空间将会产生何种变化? 特别是, 判决函数将会如何变化? 以下将分别 予以说明. 首先我们介绍参数空间的导出群.

定义4.1.1 给定分布族 $\{P_\theta, \theta \in \Theta\}$ 以及 $\mathcal{I}$ 上的可测变换群  $G = \{g\}$ , 且假设 $\theta \in \Theta$ 必有 $P_\theta \in \mathcal{P}_I$ . 若 $X \sim P_\theta$ , 对任何 $g \in G$ , 都存在使 $Y = gX \sim P_\theta$ , 则称 $\{P_\theta, \theta \in \Theta\}$ 为关于群 $G$ 的不变分布 族; 进而, 可以在 $\Theta$ 上定义一个一一变换 $g: \Theta \rightarrow \Theta$ , 使得 $P_{g(\theta)} = P_\theta$ , 记  $C = \{g\}$ , 则称 $C$ 为群 $C$ 在参数空间上的导出群(见引理4.1.2).  $\square$

以下两个例子说明: 不变分布族是存在的, 但也并非任何分布族都 是不变分布族.

例 4.1.3 设  $X \sim N(0, 1)$ , 设 $e_0 = (-\infty, +\infty)$ , 又 $\lambda = (-\infty, +\infty)$ . 在 $\mathcal{I}$ 上定义平移群

$G = \{g_c : g_c x = x + c, \text{ 则有}$

$y = -X + C \sim N(e + c, 1) = N(0, 1) + c, \theta' = 0 + c.$

即由  $X \sim P_e, W Y = g_c x - P f f.9 \text{ ②、} e + c [0. \text{ 因此 } \{N(\cdot, 1)\}$  (在平移变换  $G$  下为不变分布族. 而且  $G = \{g_c\}$  在  $\theta = (-\infty, +\infty)$  上的导出群亦为平移群:  $I = I \cdot g_c, I : g_c 6 = 0 + c.$

但是, 若取  $G = \{g_k\}$  为相似群:  $g_k x = H(x/c) = kX \sim N(k\theta, k^2).$

显然,  $N(k\theta, k^2)$  不属于分布族因此, 分布族  $\{N(\cdot, 1)\}$  关于相似变换群  $G = \{g_k\}$  不是不变分布族. | 例 4. 1.4 设  $X \sim R(Q, \theta) \theta = (0, +\infty).$  对相似群  $G = \{g_k\}$ ,

$g_k x = f_c x (k > 0)$  有

$Y = g_k X \sim R(Q, k\theta) \sim R(\theta, 6T), \theta f = k e^e.$

因此在相似变换  $G = \{g_k\}$  下为不变分布族, 而且  $G = \{g_k\}$  在  $\theta = (0, +\infty)$  上的导出群亦为相似群:  $G = \{g_k\}$

#### 4.1 变换群下的同变估计

131

但是若考虑平移群  $G = \{g_c\}$   $I : g_c x = x + c$ , 则有  $Y = g_c X = X + c \sim R(c, 3 + c) \sim R(0, 3)$

分布族, 尺度参数分布族关于相似变换群为不变分布族, 详见第2-3 节. 引理4. 1.1 若

$G = \{g_t\}$  为  $t$  上的变换群,  $\{P_\theta, e^\theta\}$  关于群  $G$  为

因此分布族  $\{P_\theta(\cdot)\}$  关于平移群  $G = \{g_c\}$  不是不变分布族. 以上两个例子说明: 一个分布族是否为不变分布族, 与所讨论的变换群有密切的关系. 常见的情形为: 位置参数分布族关于平移变换群为不变

不变分布族, 且相应的导出群为  $G = \{g_t\}$ , 则有

$P_\theta(B) = P_\theta$

$E J A(gX) = E - J A(X). \quad (4.1.5)$

证明由于  $Y = gX - P_\theta = P - \theta f$  因此有  $\sim(B) = \psi(\theta) = P^\theta(reB) = P_e(XGg - lB) = PH$

$E J A(gX) = E - J, [A(y)] = E - J/l(X).$  以上最后一式中  $X \sim P - e$ , 所以(4. 1.5)式成立. |

引理4.1.2  $G = \{g_t\}$  在  $\theta$  上构成与同态群.

证明 易见,  $C$  上的恒等变换  $eX = X \sim P_e$  对应于  $\theta$  上的恒等变换

$e : e\theta = \theta$ . 再证明封闭性. 即若  $g_1 \in G, g_2 \in G$ , 则必有  $g_1 \cdot g_2 \in G$ . 以

下证明  $g_1 \cdot g_2 = g, g \in G$ . 由于  $e \in G, g_2 \in G$ , 因此  $g_1 \cdot g_2 \in G$ . 由不

变分布族定义,  $g_1 \cdot g_2$  必对应于  $g, g \in G$ . 把以上引理应用于  $g_1 - g_2$ , 则有

$: P M g' = P/g_2 - * \cdot g_1 D$  (用于  $g_2$ )

$=$  用于  $\theta$ ),

因此  $i_1 \cdot i_2 = g_1 g_2 \in G$ . 另外若取  $g_2 = g_1^{-1}$  可得  $g_1 \cdot g_1^{-1} X = X = e \in G$ ,

因  $J t t g; 1 = g_7$  存在. 同时, 结合律显然成立; 而且对任一  $g \in G$  都有

$e \in G$  所以  $G = \{g\}$  为与  $G = \{g\}$  同态的群. I (2) 同变统计判决函数和同变估计 今给定样本空间上的变换群  $G$  以及参数空间的导出群  $G$  判决空间

的问题比较复杂, 因为判决空间  $P$  以及损失函数等都依赖于所研究的统计问题, 不能一概而论. 假设检验问题的判决空间与参数估计很不一

样, 即使参数估计, 也各有不同. 例如, 若估计  $\theta$  则显然有  $P = \theta$ ; 若估计  $h(\theta)$ , 则一般有  $V : h(\theta)$ . 因此, 判决空间的导出群也不能一概而论. 以下给出判决空间导出群的定义, 只是假设其存在, 并未证明其存在, 这与参数空间的导出群很不一样.

或  $P^\theta(gA) \quad (4. 1.4)$

定义4.1.2 设分布族  $X \sim P_e f \theta^0$  关于可测变换群  $G = \{g\}$  为不变

## 第四章最优同变估计

分布族,  $\mathcal{G}$  为群  $G$  在参数空间  $\Theta$  上的导出群. 对于给定的判决空间  $P$ , 若 对任何  $g \in G$ ,  $i \in \mathcal{G}$ , 都存在  $P$  上的一一变换且  $1$  构成与  $G$  同态的群, 则称为  $C$  在  $P$  上的导出群, 并记  $g' d = d'$ .  $I$

以下介绍同变统计判决函数和同变估计.

给定统计问题, 设为一个判决, 它必为样本的函数  $d = d(x)$  即为统计判决函数; 相应的损失函数为  $L(\theta, d)$ . 直观上看, 对于统计 判决函数和损失函数, 在以上三个群的作用下应保持协调一致, 或者说

应保持“同变”, 上三个群的作用下, 一方面有  
即

$$x \rightarrow gx, \quad d(x) \rightarrow d(gx) = d(gx)$$

$\theta \rightarrow g\theta$ ,  $d \rightarrow d'$ . 具体对判决函数来说, 在以

而另一方面也有

$$d = d(x) \rightarrow d(gx) = d(gx) = d(gx) \quad \text{因此, 一个保持协调一致的判决函数应满足 } d(gx) = d(x)$$

$= d(x)$ . 类似地, 损失函数也应保持同变. 这可归纳为以下定义:

定义4.1.3 给定  $\mathcal{G}$  上的变换群  $\mathcal{G}'$ , 其相应的导出群为  $\mathcal{G}$  和, 若 统计判决函数  $d$  满足  $d(gx) = d(x)$ , (4.1.6) 则称  $d$  为关于变换群的同变判决函数. 若损失函数  $L(\theta, d)$  满足

$$L(\theta, d) = L(g\theta, d) = L(g\theta, d) \quad (4.1.7) \quad \text{则称 } L(\theta, d) \text{ 为关于变换群的同变损失函数.}$$

例4.1.5 (续例4.1.3) 这时

$(-\infty, +\infty)$ ,  $C$  和  $\mathcal{G}$  都是移变换群. 若所研究的统计问题为估计没

则  $P = \Theta$ . 因而可取  $\theta = d$ , 即若  $J \in P = \Theta$ , 则有  $g; d = g d = d + c$ , 其中  $c$  为  $\theta$  的一个估计, 常记为  $\hat{\theta}$ . 若  $\hat{\theta}$  为同变估计, 应满足 (4.1.6) 式, 即

$$d(gx) = d(x) \Leftrightarrow \hat{\theta}(x+c) = \hat{\theta}(x) + c, \quad \forall c. \quad (4.1.8) \quad \text{该式即为例4.1.1的}$$

(4.1.1) 式. 另外, 若  $L(\theta, d)$  为同变损失函数, 则 应满足

$$L(\theta, d) = L(\theta', d) = L(\theta + c, d + c), \quad \forall c. \quad (4.1.9) \quad \text{取 } c = -\theta \text{ 上式亦成立,}$$

因此同变损失函数应满足条件

$$L(\theta, d) = L(\theta, d - \theta) = p(d - \theta). \quad \text{通常可取二次损失 } p(d - \theta) = (d - \theta)^2 \text{ 或绝对损失 } p(d - \theta) = |d - \theta|. \quad \text{例}$$

引理4.1.3 同变损失函数与同变判决函数应满足

$$L(\theta, d) = L(g\theta, d) = L(g\theta, d) = L(g\theta, d) \quad (4.1.10)$$

若  $\hat{\theta}$  相应的风险函数为  $R(\theta, \hat{\theta}) = E[\hat{\theta}(X)]$ , 则有

$$(\hat{\theta}(x), +\infty), \theta =$$

,

## 4.2 平移变换群下位置参数的最优同变估计

$R(\theta, \hat{\theta}) = R(d - \theta)$  (4.1.11) 证明 由 (4.1.6) 和 (4.1.7) 式可得 (4.1.10) 式, 该式两边取期望可得

$$E[J(\hat{\theta}, d)] = E[J(\hat{\theta}, d - \theta)]. \quad \text{上式左端} = R(\theta, \hat{\theta}), \quad X \sim P_\theta, \quad \text{而右端} = R(g\theta, \hat{\theta}),$$

$\theta' = g\theta$ , 因此右端可表示为  $E[J(\hat{\theta}, d - \theta)] = R(\theta, \hat{\theta})$ , 由此可得 (4.1.11) 式. ■

由以上讨论以及 (4.1.6)–(4.1.11) 可知, 同变统计判决是对统计 判决的一种限制, 在此限制下求最优解 (关于某种损失函数), 就得到 最优同变判决函数. 因此我们引入以下定义:

定义4.1.4 给定  $I$  上的变换群  $C$ , 其相应的导出群为  $\mathcal{G}$  和  $G$ . 设  $A = \{d(x): \text{某一统计问题的}$

全部同变判决函数 $\delta$ ，并设对应于同变损失函数 $L(\theta, d)$ 的风险函数为 $R(\theta, \delta)$ 。若存在 $\delta^*(\theta) \in A$ ，使 $R(\theta, \delta^*) \leq R(\theta, \delta)$ ， $\forall \delta \in A$ ， $\theta \in \Theta$ ，

则称 $\delta^*$ 为统计问题的最小风险同变判决。如果统计问题是参数估计，

则称 $\delta^*(\theta)$ 为最优同变估计，并记为MREE。本章4.2-4.4节将介绍若干最优同变估计(MREE)的求解方法，

以下就同变估计的有关问题再作一些说明。若是估计参数 $\theta$ ，则问题比较简单，此时显然有 $P = \Theta$ 。很自然可

定义 $T$ 中 $g$ 的估计记为 $\hat{\theta}(x)$ ，则同变条件(4.1.6)

式 $\delta(gx) = \delta^*(x)$ 可化为 $\delta(gx) = \delta^*(\theta(x))$ 。但是，若要估计 $h(\theta)$ ，则问

题比较复杂。这时 $V = h(\Theta)$ 记为 $\mathcal{H}$ ，若 $\phi$ 为 $V \rightarrow H$ 上的变换， $d \in H \setminus \mathcal{H}$ 应有 $d = h(\theta)$ ，为保持同变性， $g^*$ 应满足 $d = d'$ ，因此在 $V = h(\Theta)$ 上应有 $d = h(\theta) \rightarrow h(\theta') = d \phi$ ，即，应使 $\phi(h(\theta)) = h(\theta')$ 。进一步的讨论取决于 $\phi$ 的形式，今看一个例子，若 $X \sim p(x, \theta)$ 为尺度参数分布族，相应的相似群和导出群为 $G = \{g_c\}$ 和 $G =$

$\{h_c\}$ ， $g_c x = x + c$ ， $h_c \theta = \theta + c$ 。若要估计 $\theta$ ，贝叶斯估计 $\hat{\theta}(gx) = \hat{\theta}(x)$ 可化为 $\hat{\theta}(gx) = \hat{\theta}(h_c(x))$ ，即 $\hat{\theta}(x) = \hat{\theta}(h_c(x))$ 。若要估计 $\theta$ ，贝叶斯估计 $\hat{\theta}(gx) = \hat{\theta}(x)$ 应满足 $\hat{\theta}(h_c(x)) = \hat{\theta}(x) + c$ ，贝叶斯估计 $\hat{\theta}(h_c(x)) = \hat{\theta}(x) + c$ 的同变估计 $\hat{\theta}(x)$ 应满足 $\hat{\theta}(gx) = \hat{\theta}(x) + c$ 。

4.2.1 位置参数分布族的平移变换群 设 $X \sim p(x, \theta)$ ， $X \sim p(x, \theta)$ ，若即 $\delta(x) =$

$\delta(x)$ ，4.2 平移变换群下位置参数的最优同变估计 $\delta(x)$ 。

，

#### 134 第四章最优同变估计

$p\{x, \theta\} = f(x - \theta) = f(x_1 - \theta, x_2 - \theta, \dots, x_n - \theta)$ ，

其中 $\theta \in \Theta = (-\infty, +\infty)$ ，则称 $X$ 服从位置参数分布族。例如前节例

4.1.3 的 $N(\mu, \sigma^2)$ 就是这种分布族。当 $\sigma^2 = 1$ 时， $X \sim PQ$ 。

称为标准分布，它与参数 $\theta$ 无关。注意，位置参数分布族有以下常用性质：若 $Y = X - \theta \sim PQ$ ；反之，若 $Y \sim PQ$ ，贝 $X = Y + \theta$ 。

本节将要证明：位置参数分布族关于平移变换群为不变分布族；并进一步求出位置参数关于平移变换群的MREE。以下首先介绍样本空间

的平移变换群及其导出群，以及相应的平移同变估计和同变损失函数，为求解MREE作准备。

(1) 样本空间的平移变换群。给定样本空间定义以下平移变换群 $G = \{g_c\}$ ：

$g_c x = x + c$ ， $g_c x_i = x_i + c$ ， $i = 1, \dots, n$ 。

(2) 参数空间的导出群。由于 $Y = g_c X = X + c = (X_1 + c, \dots, X_n + c)$ ，因此 $F$ 的分布为

$Y - f(Y) = f(Y - c) = f(Y - c) - p_0 + c = p_0$ ， $\theta' = \theta + c$ 。  
 $\phi(\theta) = \theta + c$ 。

(3) 判决空间的导出群。首先要确定所研究 $f$ 的统计问题。我们仅

考虑参数 $\theta$ 的估计，这时， $P = \Theta$ ，可定义 $\delta$ ， $\delta = \delta^*$ ，使得 $\delta^* : g : \delta : \delta^* c$ 。特别有 $g^*(x) = \delta^*(x) + c$ 。

(4) 平移同变估计与同变条件。记 $\theta$ 的估计为 $\hat{\theta}(x)$ ，贝叶斯(4.1.6)式的同变条件 $\delta(gx) = \delta^*(x)$ 化为 $\delta(gx) : gKx = g\delta^*(x)$ 。由于 $g_c x = (x_1 + c, \dots, x_n + c) = x + c$ ，因此有 $\delta(x + c) = \delta(x) + c$ ，(4.2.1)该式就是对 $\theta$ 的估计 $\delta$ 的一个限制，以下将在此限制下求最优解。

推论 若 $d(x)$ 为 $\theta$ 的平移同变估计, 则必有  $E(X - \theta) = 0$ ,  
其中 $X \sim P$ . 与 $\theta$ 无关. 在(4.2.1)式中取 $c = -\theta$ 即可得到上式. (5) 平移同变损失函数.  
这时损失函数 $L(\theta, d)$ 必须满足 $L(\theta, d) =$

$W(\theta, d) = L(\theta + c, d + c), \forall c$ . 若取  $c = -\theta$ , 则有

$$L(\theta, d) = L(\theta, d - \theta) = p(d - \theta). \quad (4.2.2)$$

即平移同变损失函数 $L(\theta, d)$ 必为 $(d - \theta)$ 的函数, 以下将在此限制下求

最优解. 通常取均方损失 $L(\theta, d) = (d - \theta)^2$ , 或绝对损失 $L(\theta, d) = |d - \theta|$ .

这仍然是一个位置参数分布族, 所以

而且由上式可知, 其导出群为 $\theta = (-\infty, +\infty)$ 上的平移变换群:

$\{x \mapsto x + c\}$

$I$  为平移不变分布

#### 4.2 平移变换群下位置参数的最优同变估计 135

##### 4.2.2 位置参数的最优同变估计

以下将基于约束条件(4.2.1)–(4.2.2)来求位置参数的MREE. 为此, 首先根据同变条件(4.2.1)引入不变量和最大不变量概念.

引理4.2.1 设 $T(X)$ 和 $T^*(X)$ 都是 $\theta$ 的平移同变估计, 则 $u(X) = T(X) - T^*(X)$ 为平移不变量, 即它满足不变性关系

$u(g_c X) = u(X)$  或  $u(X + c) = u(X)$ . 反之, 若 $u(X)$ 为同变估计,  $a(u(X))$ 为平移不变量, 则 $T^*(X) = a(u(X)) + \theta$ 为 $\theta$ 的同变估计.

证明 由同变条件(4.2.1)式可知

$$u(X + c) = T(X + c) - T^*(X + c)$$

$$= [T(X) + c] - [T^*(X) + c]$$

$= T(X) - T^*(X) = u(X)$ . 因此 $u(X)$ 为平移不变量. 反之若 $u(X)$ 为同变估计,  $a(u(X))$ 为平移不变

量, 则有

$$T^*(X + c) = a(u(X + c)) + \theta = a(u(X) + c) = a(u(X)) + c + \theta = T(X) + c.$$

因此 $T^*(X)$ 为同变估计.

引理4.2.2  $a(u(X))$ 为平移不变量的充要条件是: 存在函数 $\psi(y)$ 使

$$T^*(X) = \psi(u(X)) + \theta, \text{ 其中}$$

$y = y(X) = (X_1 - X_n, \dots, X_2 - X_{n-1})^T$  (4.2.3) 为不变量, 而且不变量 $u(X)$ 的分布仅与标准分布 $P$ 有关, 与 $\theta$ 无关.

证明 必要性. 若 $u(X + c) = u(X)$ 为不变量, 则可取  $c = -\theta$ , 因而有

$$u(X) = u(X) - \theta, \text{ 即 } u(X) = u(X - \theta), \text{ 而 } X - \theta \sim P_\theta, \text{ 其分布与 } \theta \text{ 无关.}$$

充分性. 若  $u(X) = \psi(y)$ , 则有  $y_i = x_i - x_n = (x_i + c) - (x_n + c)$ ,  $y(X + c) = y(X)$ , 所以 $y$ 为平移不变量, 因此 $u(X) = \psi(y)$ 为平移不变量. |

引理4.2.2所描述的性质有一定的共性, 因此引入以下定义.

定义4.2.1  $y = y(X)$ 称为最大(平移)不变量, 若 $y_0$ 为不变量, 且对任一不变量 $u(X)$ , 存在函数 $\psi(\cdot)$ , 使 $u(X) = \psi(y)$

注(1)最大不变量不唯一. 例如在前面证明中, 若取 $c = -x_n$ ,

即

■

则 $T(X) = (x_1 - x_n, \dots, x_{n-1} - x_n)^T$ 亦为最大不变量, 且有 $U(X) = \psi(T(X))$ .

(2) 任一不变量  $u(\theta; n)$  的分布与  $\theta$  无关, 因而为辅助统计量, 因此由 Basu 定理 (见第二章), 它与完备充分统计量独立.

(3) 对于平移变换群,  $X^{(n)} = X(1), \dots, X(n)$  等皆视为同变估计, 而  $X(n) - X(1)$  等皆视为不变量. 因为  $X \sim \theta + c$ , 必有  $X + c \sim \theta + c$ ,

$X^{(n)} - X(1) \sim (\theta + c) - (\theta + c) = 0$ . 引理 4.2.3 设  $\delta(X)$  为  $\theta$  的某一平移同变估计, 则任一平移同变估计

$V(X)$  可表示为

$V(X) = \delta(X) - \theta$ ,  $y = X(2) - X(1)$

证明由引理 4.2.1 和引理 4.2.2 有

$V(X) = \delta(X) - \theta = \delta(X) - d(X) = \delta(X) + u(X) = d(X) - A(y)$ . I

由以上引理 4.2.1-引理 4.2.3 可总结出求解最优同变估计 MREE 的方法如下:

(1) 取某一合适的同变估计  $\delta(X)$ ;

(2) 任一同变估计可表示为  $\delta(X) - \theta$

(3) 对于给定的同变损失函数 (一般为均方损失), 求  $\delta(\cdot)$ , 使风险函数最小.

定理 4.2.1 (Pitman) 设  $d(X)$  为  $\theta$  的某一平移同变估计, 损失函数

其中  $E_{\theta}$  表示对标准分布  $PQ$  取期望,  $r$  由 (4.2.3) 式所示; 并且 (4.2.4) 式的解唯一, 与  $\delta(X)$  的选取无关.

证明 (1) 最优性. 由引理 4.2.3, 可假设  $V(X) = d(X) - \theta$ , 并求  $\delta(\cdot)$ , 使  $W(X)$  的均方误差 MSE 最小, 其中

$MSE = E_{\theta} [W(X) - \theta]^2 = E_{\theta} [\delta(X) - \theta]^2$ . 由同变条件 (4.2.1) 式可知  $\delta(x+c) = \delta(x) + c$ , 取  $c = -\theta$  则有  $\delta(x - \theta) = \delta(x) - \theta$ ; 该式代入上式可得

$MSE = E_{\theta} [\delta(X - \theta) - \theta]^2$ .

由于  $\theta \sim P$ , 且  $f$  的分布仅与  $P$  有关, 因此有

$mse = E_{\theta} [d(X) - \theta]^2$ ,  $X \sim P_{\theta}$ . 为求上式的最小值, 可化为条件期望表达式

$mse = E_{\theta} [E_{\theta} [(d(X) - \theta)^2 | Y]]$ . (4.2.5) 在条件期望中,  $y$  为已知,  $A(y)$  为“常数”,

可先考虑以下函数的最小值问题:

为  $L(e, d)$

则  $\theta$  的最优同变估计可表示为

$= d(X) - E_{\theta} [d(X) | Y]$ , (4.2.4)

4.2 平移变换群下位置参数的最优同变估计

137

$/(A) = E[(X-A)^2 | Y] = E[(X^2 - 2XA + A^2) | Y] = E(X^2 | Y) - 2AE(X | Y) + A^2$ .

$/(A)$  为  $A$  的二次三项式, 且有  $/'(A) = 2A - 2E(X | Y)$ ,  $/''(A) = 2 > 0$ . 因此,  $A = E(X | Y)$  时  $/(A)$  达到最小值. 这一结果应用于 (4.2.5) 式,  $X \sim P_{\theta}$ ,  $a = A(y)$ , 则  $A(y)$

$= E_{\theta} [\delta(X) | Y]$  时 (4.2.5) 达到最小值,

所以 (4.2.4) 式成立. (2) 唯一性. 今设另有一个平移同变估计为  $\hat{\delta}(X)$ , 代入 (4.2.4) 式有

$V(X) = \hat{\delta}(X) - \theta = E_{\theta} [\hat{\delta}(X) - \theta | Y]$ . 今证 (1). 由引理 4.2.3,  $\delta(X)$  可表示为  $\delta(X) = \hat{\delta}(X) - p(y)$ , 其中  $p(\cdot)$  为某一个函数. 该式代入的表达式有  $\delta(X) = \hat{\delta}(X) - E_{\theta} [\hat{\delta}(X) | Y]$

$= d(X) - E_{\theta} [d(X) | Y] - p(y) + p(y) = d(X) - E_{\theta} [d(X) | Y]$ . 1 注意, Pitman 公式 (4.2.4) 中,  $\delta(X)$  和  $\theta$  皆不唯一, 但最后的



解唯一，特别有

推论1  $\theta$ 的最优同变估计可表示为

$r(X) = X, -E_{\theta}(\text{dir}) = X(1) - E_{\theta}(X(1) | Y)$ . (4.2.6) 另外，在公式(4.2.4)

中，主要困难在于计算条件期望，这经常可

通过Basu定理把条件期望化为无条件期望。

推论2 若 $T$ 为完备充分统计量，且 $\xi(X) = (p(T))$ 为 $\leq 9$ 的平移同变

估计，则有

$r(X) = \hat{D} - E_{\theta}[p(n)]$ . (4.2.7) 证明 因为 $r$ 为辅助统计量，与完备充分统计量 $r$ 及

$p(r)$ 独立，

因而有 $E_{\theta}[\hat{(r)} | y] = E_{\theta}[\hat{(r)}]$ ，该式代入(4.2.4)可得  $V(X) = \leq p(T) - E_{\theta}[(P(T) \setminus Y) = \hat{(T)} - E_{\theta}[\hat{(r)}]$ .

推论3 Pitman估计是无偏的。证明

$\text{bias}(d^*) = E W(X) - \theta = E J^{\wedge}(X) - \leq 9 - E_{\theta}(d(X) | K)]$ ,

而  $\theta = y$  的分布也只与 $\rho$ 有关，因此有

$\text{bias}(d^*) = E J^{\wedge}(D) - E_{\theta}(d(x) | F)] = E_{\theta}[d(X) - E_{\theta}(d(x) | r)] = E_{\theta}[\hat{(X)}]$

$- E_{\theta}[E_{\theta}(\hat{(X)} | y)] = 0$ . 我们在第三章已经证明，在无偏估计中，一致最小方差无偏估计

(UMVUE)最好。因此在平移群下，对于均方损失而言， $\theta$ 的最优同变估计(MREE)不会比UMVUE好。从这个观点来看，平移群下位置参数

的MREE意义不是太大。但是通过本节的内容，我们比较系统地介绍

#### 138 第四章最优同变估计

了解最优同变估计的基本思想方法。这些思想方法可用于求解其他变换群下的MREE。下一

节将看到，相似变换群下尺度参数MREE的求解

方法与本节很类似，但有更多的优良性。

例4.2.1 设为独立同分布样本， $\sim yV(M, l)$ ;

ii)  $\sim M+r(l, l)$ ，求 $m$ 的最优同变估计。

» i)可取平移同变估计，它显然满足同变条件+  $+c$ 。无亦为完备充分统计量，所以无与 $r$ 独立。因

此由(4.2.7)式有

$= A(\hat{)} - E J^{\wedge}(KX)] = X - E_{\theta}[X] = x$ . ii)可取 $1(X) = X(1)$ 为平移同变估计，且为完备充分统计量，因此

由(4.2.7)式有

$/T(X) = x(1) - E_{\theta}[X(1)]$ . 而 $y=0$ 时 $X(1) \sim r(n, l)$ ,  $E_{\theta}[X(1)] = 1/n$ , 因此有

$* (\hat{)} - \hat{(i)} - \perp$ . fl

以上结果都与一致最小方差无偏估计一致。4.2.3 Pitman积分公式

记  $\rho(x, -\theta), \dots$ , 由(4.2.4)可知， $(X)$ 可由标准分布 $P_{\theta} = \rho/(x)$ 表示。特别，根据(4.2.6)式，我们可以计算出条件期望 $E_{\theta}C(JK)$ ，从而把 $(X)$ 表示为关于 $\rho/(*)$ 的积分。

#### 定理4.2.2

设则 $\theta$ 的平移最优同变估计可表示为

$$f \text{ ef}(x - \theta) d\theta = \frac{\int_{-\infty}^{\infty} f(x - \theta) d\theta}{\int_{-\infty}^{\infty} f(x - \theta) d\theta} .$$

[  $f(X - \theta) d\theta > J - \theta$

$f \text{ tf}(x - a) dt$

$= \frac{\int_{-\infty}^{\infty} f(x - a) dt}{\int_{-\infty}^{\infty} f(x - a) dt}$  或 谷,  $w =$

[  $f(X - t) dt$

(4.2.8) 证明 在公式 $\theta^{\wedge}(X) = e(x) - E_{\theta}[\xi(x) | r]$ 中，可取 $\theta(x) = \theta$ 和

$r = (x_2 - x_1, \dots, x_n - x_1)^T = (y_2, y_3, \dots, y_j)$   
 因而可对  $V(x) = 1 - E_j j(y)$  直接进行积分. 为求  $E_j x_j y$ , 需要求没=0时的条件分布  $p(\cdot | y)$ , 为此, 可令  $y_1 = x_1$ ? 先求联合分布  $(x_1, y_1, \dots, y_n)^T \sim \cdot, y$ . 从而  $P(x, y) = p(x_1, y) / \int p(x_1, y) dx_1$  以及  $E_0(IY) = \int x \{p(x | IY)$   
 心, 而  $(X_2, r_2, \dots, Y_n)$  的分布可通过下列变换得到:

#### 4.2 平移变换群下位置参数的最优同变估计

139

$$L_{Xn} = + r_n.$$

易见变换的Jacobi行列式  $J=1$ , 则有  $X \sim \theta^{(n)}$ . 无.)  $\sim \rho \theta$ ,

$(m + y_2, \dots, y_i + y_j \sim p(y | y)(n \sim),$

$Y = \int p(y | y) dy + h, \dots, y_j d_7 i \sim p(y)(y = i).$

$J = 00$

因此  $X_1 | y = y | i y \sim p(y_1, y) / p(y)$ , 即

$J = CD$

$E_j x, i r]$

在以上条件期望中,  $r_2 = x_2 - x_1, \dots, r_n = x_n - x_1$  为已知, 可对上述积分进行变换, 使积分有对称的形式. 由于  $u + r_i = w + x_i - x_1 = x_i - (X, -u)$ ,  $i=2, \dots, n$ , 因此若令

则有  $u + y_i = -h, J=2, \dots, 71$ , 从而以上积分可化为

$p(x_1 | y_2, \dots, y_n)$

$f(x_1 + y_2, \dots, x_1 + r_j) / (u, u + y_2, \dots, u + y_n) du$

$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(u, u + y_2, \dots, u + y_n) du$

$f(u, u + y_2, \dots, w + Y_n) du$

$e_0[x, r] = [$

$J = 00$

1- 而冷  $*(x) = x_1 - E_j x_j y$ , 因此有

$b'(x) = \{ / (x -$

$- d) d( / C f \{ X - d) dt \int J - QD$

$J y(x - z_i) d / .$

(尤  $i \sim 0 / (欠 1 - t, X_2 - i, \dots, X_n - () < h$

$[ / (^1 - \dots, \dots,$

$- \dots$

此即(4. 2.8)式, 证毕. 例4.2.2设  $X_1, \dots, X_n$  为独立同分布样本,  $X_i \sim ? (< 9 + 1)$ ,

求没

的最优同变估计.

$\wedge, - 0$  山

■

140

#### 第四章最优同变估计

(注意: 第三章例3.3.2曾经指出, 该分布的极大似然估计不唯一. 另外, 其充分统计量为  $T = (X(1), X(n))$ , 但不是完备的充分统计量, 因而不 易求出UMRUE, 本例题将求出其平移MREE.)

解  $\% = \cdot, \dots, X_n)^T$  的分布为  $nn$

$X = P f f$

$< 9 + 1! = f p | o$  多\一没矣"

- 门 彡

$i x \ 1 \ i = 1$

$= \int_0^\infty \theta \cdot x(1) \cdot (B)^{\theta} + 1 \cdot d\theta$ . 这是一个位置参数分布族, 可用Pitman公式(4.2.8)求出 $\theta$ 的平移最优

同变估计.

$f(x - e_i) = i \cdot e^{x(1)} \cdot (B)^{\theta} + 1 \cdot d\theta$  :  $i \cdot x(n) - 1 \cdot \theta$

$J(x) = J \cdot e^{x(1)} \cdot (B)^{\theta} + 1 \cdot d\theta$

$= J \cdot e^{x(1)} \cdot (B)^{\theta} + 1 \cdot d\theta$

$f \cdot f(x - e_i) = [ \cdot d\theta = x(1) - \theta ] \cdot J \cdot e^{x(1)} \cdot (B)^{\theta} + 1 \cdot d\theta$

以上各式代入(4.2.8)式可得

$m) = \frac{1}{n} \cdot (1 + \frac{1}{n}) \cdot J \cdot e^{x(1)} \cdot (B)^{\theta} + 1 \cdot d\theta$

4.3相似变换群下尺度参数的最优同变估计

4.3.1 尺度参数分布族的相似变换群

一般的尺度参数分布族可表示为  $X \sim \text{Exp}(\theta)$  (伽含),

其中  $\theta = (\theta, +\infty)$ . 当  $t=1$  时,  $X \sim \text{Exp}(\theta)$  为标准分布, 例如  $j/$

$V(\theta, a^2)$  或  $L(\theta, a)$ . 注意, 尺度参数分布族

有以下常用性质: 若  $X = aY - Pa$ .

则  $r = X/a - P$ ; 反之

则

本节将要证明: 尺度参数分布族关于相似变换群为不变分布族; 并进一步求出尺度参数关于相似变换群的MREE. 以下首先介绍样本空间

的相似变换群及其导出群, 以及相应的相似同变估计和同变损失函数, 为求解MREE作准备.

(1) 样本空间的相似变换群. 给定样本空间  $\mathcal{X}$ , 定义以下相似变

4.3相似变换群下尺度参数的最优同变估计

141

变换群  $G = \{g_k \mid :$

$g_k x = kx \mid P \mid g_k X_i = x_i = kx_i, i = 1, \dots, n.$

(2) 参数空间的导出群. 由于  $Y = g_k X = kX$ ,  $\dots$  风  $Y$ , 因此  $Y$  的分布为

$Y = g_k X = kX : pa, \theta = \frac{1}{k} \theta$

这仍然是一个尺度参数分布族, 所以  $X \sim \text{Exp}(a, \theta)$  为相似不变分布  $M$ . 而由上式可知, 其导出群为  $\theta = (\theta, +\infty)$  上的相似变换群:  $G =$

$f \in \mathcal{A}$ ,  $g_k(\theta) = \frac{1}{k} \theta$ .

(3) 判决空间的导出群及参数的同变估计. 首先要确定所研究的

统计问题. 我们考虑两种情形, 即参数  $a$  和  $\theta$  的估计. 这时判决空间、同变条件以及同变损失函数都不一样.

(3a) 关于  $\theta$  的估计. 这时,  $P = (\theta, +\infty)$ , 可定义  $G' = G$ ,  $g' = g_k$  使得  $d(d') = g : d = kd'$

特别有  $g_k^{-1}(\theta) = ka(\theta)$ . 则(4.1.6)式的同变条件  $\delta(gx) = g \cdot \delta(x)$  化为  $a(gkx) = a(x)$ , 即

$a(kx) = ka(x)$ . (4.3.1) 该式就是对  $a$  的估计 (幻的一个限制, 以下将在此限制下求最优解.

特别, 若取则有

$a(x/a) = a(x)$ . (4.3.2) 其中  $X/a \sim P$  该分布与参数  $a$  无关.

相似同变损失函数. 这时损失函数  $L(a, d)$  应满足  $L((r, d)) = f(\theta, (r')) = f(Axr,$

$Ax/)$ ,  $\forall A : > 0$ . 取  $\theta = cr - 1$ , 则有

因此同变损失函数 $L(a, d)$ 必为 $(d/a)$ 的函数，以下将在此限制下求最优解。通常取均方损失

或绝对损失  $L(a, d) = p(d/a) = |1 - d/a|$ 。(3b)关于 $\theta$ 的估计。这时， $P$ 为一切 $\theta$ 的估计组成。易见，当

$a \rightarrow \langle r' = A : a \text{ 时}, ar - (af)r = krar$ ;若 $\theta \in P$ 为 $\theta$ 的估计，则应有  $\int dWd$ ，因此导出群应定义为

$= \{g_k |, d = krd\}$ 。特别，若 $\theta$ 的同变估计记为 $\hat{\theta}(X)$ ，则应有 $J; ar\{X\} = krar(X)$ 。

142

#### 第四章最优同变估计

因此，(4.1.6)式的同变条件 $g(gx) = g \cdot 5(x)$ 化为

$ar(kx) = kr; r(x)$ 。(4.3.3)

该式就是对 $\theta$ 的估计 $P$ （幻的一个限制，以下将在此限制下求最优解。

特别，若取则有 $\hat{a}$

$crr(x/(T)) = ar(x)/ar$ ,

其中 $X/a - P^{\hat{a}}$ 该分布与参数 $a$ 无关。关于同变损失函数 $L((r, d)$ ，应满足 $L(a, d)$

$= L(a', d') = L(ka,$

$Vd)$ ， $\forall k > 0$ 。当 $k = (T-1)$ 时亦成立，因此上式化为

$L\{a, d\} = 1, \hat{a};$ 形)，

即同变损失函数 $L(a, d)$ 必为 $(d/a)$ 的函数，以下将在此限制下求最优解。通常取均方损失或绝对损失  $L((y, d) = p(d/ar) = |1 - d/crr|$ 。4.3.2尺度参数的最优同变估计

以下主要求 $a$ 的最优同变估计( $\theta$ 的情况类似，可作为推论)，其思想方法与位置参数的最优同变估计十分类似，即通过不变量及最大不变量，导出定解条件；而推导定解条件的出发点为同变条件(4.3.1)式和(4.3.2)式，即 $a(kx) = ka(x)$ ，且有 $a(x/a) = a(x)/a$ ，其分布仅与 $\theta$ 有关，与 $a$ 无关。

引理4.3.1若 $\theta = \theta_1(x)$ 和 $\theta_2(x)$ 为 $a$ 的相似同变估计，则 $u(x) = \theta_1(x)/a^2(x)$ 为相似不变量，即 $u(kx) = u(x)$ 。反之，若 $a(x)$ 为相似不变量， $\theta = \theta_1(x)$ 为相似同变估计，则 $a^2(x) = \theta_1(x) u(x)$ 亦为相似同变估计。

证明 由同变条件(4.3.1)可知

$u(kx) = crr\{kx\} ka^x(x) = u(x) \cdot ka^2(x)$

反之，若 $u(x)$ 为相似不变量， $a^2(x)$ 为相似同变估计，则 $a^2(kx) = a^2(x) u(kx) = a^2(x) u(x) = ka^2(x)$ 。

因此 $a^2(x)$ 为相似同变估计。

引理4.3.2  $u(x)$ 为相似不变量的充分必要条件是，存在函数 $\theta$ 使 $u(x) = \theta(a^2(x))$ ，其中 $z = Z(x) = (z_1, z_2, \dots, z_n)$ 且有

4.3相似变换群下尺度参数的最优同变估计

143

$z_i = -1, i=2, \dots, n$  (4.3.4)  $I \cdot I \cdot I$

(这时应要求 $P(1^I = 0) = 0$ )，并且 $a(X)$ 的分布仅与 $\theta$ 有关，与 $\theta$ 无关。

证明必要性。由不变量的定义有

$u(kx) = u(x_1^2, \dots, x_n^2) = u(x_1^2, \dots, x_n^2) \cdot \theta(k^2)$

取 $\theta = 1^I$ ，则有

$\theta(x_1^2, \dots, x_n^2) = \theta(x_1^2, \dots, x_n^2)$

$U(x_1^2, x_2^2, \dots, x_n^2) = U(x_1^2, x_2^2, \dots, x_n^2)$

$\dots, \dots, \dots$

$= U(Z), z_1, z_2, \dots, z_n = \theta(1)$ 。充分性。若 $u(x) = \theta(z)$ ，由(4.3.4)可知， $z$ 显然为不变

量, 因而  $W(\%)$  为不变量. 又因为  $u(U) = u(kx) = u(x/a)$ , 而  $X/a \sim$ , 所以  $\langle \cdot \rangle$  的分布仅与  $U$  有关, 与  $a$  无关. ■

注 (1)  $z = z(x)$  为最大不变量, 因为任一不变量  $u(\cdot; v)$  都可以表示为  $u(\%)$  的函数  $u(\%) = i/r(z(x))$ .

(2) 引理中的 2 与  $\varphi(\cdot)$  都不唯一, 例如以上证明中, 取  $|x_n|^{-1}$ , 则可得到另一最大不变量/与函数  $\langle \cdot \rangle$ .

(3)  $WX$  的分布与  $cr$  无关, 因而为辅助统计量, 所以由 Basu 定理可知, 它必与完备充分统计量独立.

(4) 对于相似变换群,  $|x|$ ,  $|X(1)|$ ,  $|X(n)|$  等皆可视作同变估计

引理 4.3.3 设  $a(x)$  为某一相似同变估计, 则任一相似同变估计  $a^*(\%)$  可表示为

$a^*(x) = a^{-1}(x) i/r(z(x))$ . 证明 由引理 4.3.1 和引理 4.3.2 可得

而

$X(1)/X(4)$  等皆可视作不变量, 因为  $x^k X_i$ , 必有

计,

$x_1^{-1} \dots x_n^{-1} = (kxY)/(kxn)$ .

$a^*(x) = a^{-1}(x)$

$a^{-1}(x) = (T_1(x)u(x) = 0 - J(x)l/f(z(x)))$ .

由以上引理 4.3.1-4.3.3 可归纳出求解最优相似同变估计的方法如下: (1) 取某一特定的同变估计  $\hat{a}(X)$ ;

(2) 令  $a^*(X) = \hat{a}(X) \varphi(Z(X))$ ; (3) 对于给定的同变损失函数 (一般为均方损失), 求  $\varphi(\cdot)$ , 使风

险函数  $R(W)$  达到最小. 以下引理对于求解均方误差的最小值点很方便, 本节和下一节都要用到.

144

引理 4.3.4 记  $MSE = E[a(X) + b(X)]$ ,  $(7) = -E[ : X ]$

时达到最小值.

证明 把  $MS$  化为条件期望的表达式:

第四章最优同变估计 29 则当

$MSE = E[E_j[a(X) + b(X) | Y]]$ . 在条件期望中, 把  $Y$  为“已知常数”, 记为  $A$ , 因此可考虑以下函数的最小值问题:

由于

$f(A) = E[f[a(X) + b(X) | A]] = A^2 E[6^2(X) | Y] + 2AE[a(X)b(X) | K] + E[a^2(X) | K]$ .

$f'(A) = 2AE[6^2(X) | Y] + 2E[a(X)b(X) | K]$ ,  $f'(A) = 2E[6^2(X) | Y] > 0$ ,

所以当  $f'(A) = 0$  时  $f(A)$  达到最小值, 因此当

$= E[a(X)b(X) | X] = E[6^2(X) | Y]$

时  $f(A)$  达到最小值, 即以上  $MSE$  达到最小值. | 定理 4.3.1 (Pitman) 设  $\varphi(X)$  为  $a$  的某一个相似同变估计, 则在

均方损失下,  $a$  的最优同变估计可表示为

$m) | Z] a(X)$ , (4.3.5) | Z]

其中  $E$  表示对标准分布  $\tau$  取期望,  $Z$  的分量如 (4.3.4) 式所示, 并且 (4.3.5) 式的解唯一, 与  $\varphi$  (尤) 的选取无关.

证明 根据引理 4.3.3, 可设  $a^*(X) = a(X)Z$ , 并求  $\varphi(\cdot)$  使  $MSf$  最小, 因此有

服叫  $\tau$  叫  $\varphi(2)_{1\tau}$

由 (4.3.2) 式有  $a(X)/a = a(X/a)$ ,  $X/a$  同时  $Z$  的分布亦仅与  $\tau$  有关, 因此有

$MSf = E[a(X/a) i / (Z - 1)]^2 = E, [S(X) \varphi(Z - 1)]^2$ ,

应用引理4.3.4, 其中  $a(X) = -1$ ,  $b(X) = a(X)$ , 所以当

$\varphi(Z) = m \cdot \mathbb{E}[a^2(X) | Z]$

时达到最小值, 由此即得(4.3.5)式. 再证唯一性, 设  $\hat{a}(X)$  为另一同变估计, 则有

4.3相似变换群下尺度参数的最优同变估计

145

由引理4.3.3可知, 存在某一个  $\varphi(\cdot)$ , 使  $\hat{a}(X)$  可表示为  $\hat{a}(X) = \varphi(X)^{\varphi(Z)}$ , 因而代人上式可得

$\sim \mathbb{E}[a(X) \varphi(Z) | Z]$

$a(X) = \mathbb{E}[W^{\varphi(Z)} | Z] = W^{\varphi(Z)}$

推论1 若  $T: \mathcal{T}(X)$  为完备充分统计量, 且  $\varphi(r)$  是  $a$  的同变估计, 则  $a$  的最优同变估计为  $\mathbb{E}[\varphi(T) | \mathcal{T}(T)]$ . (4.3.6)  $\mathbb{E}[\varphi(T) | \mathcal{T}(T)]$

因为  $Z$  的分布与  $a$  无关, 所以为辅助统计量, 由 Basu 定理,  $\varphi$  与  $Z$  独立, 因而有  $\mathbb{E}[\varphi(T) | Z] = \mathbb{E}[\varphi(T)]$ ,  $\mathbb{E}[\varphi(T) | Z]$

由此可得(4.3.6)式.

(4.3.7)

推论3 设  $\hat{a}(Z)$  为  $a$  的某个同变估计, 则  $a$  的最优同变估计为

推论2  $a^*(X)$  可表示为

$0^{**}(\varphi) = \mathbb{E}[\varphi(X) | T]$

$\mathbb{E}[\varphi(X) | T]$  因为可取  $a(X) = \varphi(X)$ ,  $\varphi$  为  $(T)$  的一个同变估计.

]

$\mathbb{E}[\varphi(X) | Z]$ , (7)

推导过程与(4.3.5)式十分类似, 从略.

$\varphi^*(\varphi)$

(4.3.8)

例4.3.1 设  $X_1, \dots, X_n$  独立同分布,  $\varphi \sim \mathcal{U}$  求  $a(T)$

的最优同变估计.

解. 易见, 分布族为尺度参数分布族. 取  $\varphi = \frac{1}{n}$  则由

有  $T = \sum_{i=1}^n X_i^{\varphi}$ , 因此  $\varphi$  为同变估计, 而且为完备充分统计量, 因此由(4.3.6)式有

$m$ ) 当  $\varphi = \frac{1}{n}$  时,  $X_i \sim r(1, 1)$ ,  $T \sim r(1, n)$ , 因此  $\mathbb{E}[T] = n$ ,  $\mathbb{E}[T^2] =$

$n(n+1)$ . 代人上式可得  $a^*(X) = r/(n+1)$ . 这一结果比极大似然估计及一致最小风险无偏估计都好. ■ 例4.3.2 设  $X_1, \dots, X_n$  总独立同分布,  $X_i \sim R(0, \theta)$  求  $\theta$  的最优同

$\mathbb{E}(nT)$ .  $\mathbb{E}(r^2)$

146

第四章最优同变估计

变估计.

解 取  $\varphi$  为同变估计, 它也是完备充分统计量. 当  $\theta = 1$  时

$X(n) \sim \mathcal{E}(n, 1)$ , 因此由(4.3.6)式有

$\varphi^*(\varphi) = \mathbb{E}[\varphi(X) | T]$

这一结果比极大似然估计及一致最小风险无偏估计都好. ■ 例4.3.3 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim V(0, a^2)$ , 求  $a^2$  的最优

同变估计.

解取  $r =$

$i=1$

因而与  $Z$  独立. 当  $\varphi = 1$  时

$\varphi^*(X) = \mathbb{E}[\varphi(X) | T] = \mathbb{E}[\varphi(X)] = \frac{1}{n+2}$

这一结果比极大似然估计及一致最小风险无偏估计都好。例4.3.4 设 $X_1, \dots, X_n$ 独立同分布,  $X_i$ 服从Rayleigh分布

$f(x) = \frac{x}{a^2} e^{-x^2/(2a^2)}$ ,

$a > 0$

求  $a^2$  的最优同变估计。

解  $f(x, a)$  可表示为

因此为尺度参数分布族。而且  $r=1$  为  $a^2$  的同变估计, 并且是完备

充分统计量。当  $r=i$  时,  $a^2 = 2x_i^2 / e^{-x_i^2/(2a^2)}$ , 因此  $X_i \sim r(i, 1)$ ,  $r \sim r(i, n)$ 。所以由(4.3.8)

式有

$E(T) = (n+1) a^2$

### 4.3.3 Pitman积分公式

由(4.3.5)可知,  $a^2(X)$  可由标准分布  $P, \sim / (\%i, \dots, x_n) = f(x)$  表示。特别, 根据(4.3.7)式, 我们可以计算出条件期望  $E_i(X_j | Z)$  和  $E_i(X_j^2 | Z)$  从而把  $a^2(x)$  表示为关于  $f$  的积分。在第2节的定理

4.2.2 曾经有这方面的结果, 此处将讨论类似的问题, 但是情况要比那里复杂一些。根据

(4.3.7)式, 为了计算出条件期望  $E_i(X_j | Z)$  和  $E_j(X_j^2 | Z)$ , 必须求出  $(X_1, \dots,$

$X_n) | Z, Z^2, \dots, Z_n$  的条件分布  $| X, Z^2, \dots,$

$Z_n$ , 则, 它是  $a^2$  的同变估计, 亦为完备充分统计量,

由(4.3.8)式有

4.3相似变换群下尺度参数的最优同变估计

147

$z_i = \pm 1$  只取两个值, 因此可以分两步求解: 首先求  $(X_1, Z_1, \dots, Z_n)$  的联合分布

$P(X_1, Z_1, \dots, Z_n)$  以及条件分布  $P(X_1 | Z_1, \dots, Z_n)$ , 然后再根据  $z_i = +1$  和  $z_i =$

$-1$  (即  $X_i > 0$  和  $X_i < 0$ ) 两种情况来求条件

分布  $(X_1, Z_1, \dots, Z_n)$ 。

引理4.3.5 假设  $P(X_i = 0) = 0$ , 贝! 当  $a=1$  时,  $(X_1, Z_1, \dots, Z_n)$  的

$Z_n$ , 其中  $A$  与  $Z, \dots,$

条件分布可表示为

证明 首先求出  $X_1, Z_1$ , 作变换

$|x_j| = n^{1/2} / (x_1^2 + \dots + x_n^2) = 1 / \sqrt{u_1^2 + \dots + u_n^2}$  则

$f$  共有  $n+1$  维。但是, 由于

(4.3.9) 的联合分布,  $z_1, \dots, z_n$ 。为此,

$w_1, \dots, w_n = Y_1, \dots, Y_n$ 。

变换的Jacobi行列式的绝对值为  $|J| = |r| \cdot |n|$ , 因此由  $(X_j, \dots, X_n) = f(x_1, x_2, \dots, x_n)$

可得

而  $r = x_1, \dots, (y_1, \dots, y_n)$

$(y_1, y_2, \dots, y_n)^T, \dots, : r_{ij} = X_i Y_j$ ,

因此  $(h_1, \dots,$

以下继续求  $(X_1, Z_1, \dots, Z_n)$  的条件分布  $|z|, z_1, \dots, z_n$ 。这时  $Z_i = X_i / |X_1| = \pm 1$ , 而且  $P$

(见  $|z_1, \dots, z_n|$  已如(4.3.9)式所示。

分布为 由条件分布公式可得(4.3.9)式:

$T = (z_1, \dots,$

$z_n)$  的联合

$z_j T$

$, Z_2, \dots, z_j = f(x_j), x_1, \dots, x_n \in I, x_j$

引理4.3.6 设  $P(X_j > 0 \mid Z_2 = z_2, \dots, Z_n = z_n) = p_x$ , 则有

148 第四章最优同变估计

$P(x_i \mid z_2, \dots, z_n) / \int_I p(x_i \mid z_2, \dots, z_n) dx_i = P(x_i \mid z_2, \dots, z_n) / \int_I p(x_i \mid z_2, \dots, z_n) dx_i$ ,  $A = -$

证明 因为只取两个值:  $Z_i = 1 (X_i > 0)$  或  $Z_i = -1 (X_i < 0)$ , 可按这两种情况分别计算条件分布. 首先考虑  $A = 1$  的情形, 这时

$P(x_1 \mid z_2, \dots, z_n) = \int_I p(x_1 \mid z_2, \dots, z_n) dx_1 = \int_I p(x_1 \mid z_2, \dots, z_n) dx_1$

$P(x_1 \mid z_2, \dots, z_n) = \int_I p(x_1 \mid z_2, \dots, z_n) dx_1 = \int_I p(x_1 \mid z_2, \dots, z_n) dx_1$

$= \int_I p(x_1 \mid z_2, \dots, z_n) dx_1 = \int_I p(x_1 \mid z_2, \dots, z_n) dx_1$

该式对  $A$  求导可得

$P(x_1 \mid z_2, \dots, z_n) = P(x_1 \mid z_2, \dots, z_n) = P(x_1 \mid z_2, \dots, z_n)$

定理4.3.2 设  $X = (X_1, \dots, X_n)^T$  服从尺度参数分布族:  $X \sim p(x, c)$

则  $T$  的最优同变估计可表示为

证明 应用公式(4.3.7),

149

今分别对  $A = 1$  和  $A = -1$  来计算条件期望  $EJ$

$E(x_i \mid Z_i)$ . 当  $A = 1$  时,  $A > 0$ , 这时由引理4.3.6有

$E(x_i \mid Z_i) = \int_I x_i p(x_i \mid z_i) dx_i = \int_I x_i p(x_i \mid z_i) dx_i$

$\int_I x_i p(x_i \mid z_i) dx_i = \int_I x_i p(x_i \mid z_i) dx_i$

$\int_I x_i p(x_i \mid z_i) dx_i = \int_I x_i p(x_i \mid z_i) dx_i$

由引理4.3.5的(4.3.9)式可得  $E(x_i \mid Z_i) = \int_I x_i p(x_i \mid z_i) dx_i$

同理可得

4.3相似变换群下尺度参数的最优同变估计

149

$E(x_i \mid Z_i)$

因此当  $Z_i = 1$  时有

$\int_I x_i p(x_i \mid z_i) dx_i = \int_I x_i p(x_i \mid z_i) dx_i$

$\int_I x_i p(x_i \mid z_i) dx_i = \int_I x_i p(x_i \mid z_i) dx_i$

$\int_I x_i p(x_i \mid z_i) dx_i = \int_I x_i p(x_i \mid z_i) dx_i$

$\int_I x_i p(x_i \mid z_i) dx_i = \int_I x_i p(x_i \mid z_i) dx_i$

$\int_I x_i p(x_i \mid z_i) dx_i = \int_I x_i p(x_i \mid z_i) dx_i$

$\int_I x_i p(x_i \mid z_i) dx_i = \int_I x_i p(x_i \mid z_i) dx_i$

$\int_I x_i p(x_i \mid z_i) dx_i = \int_I x_i p(x_i \mid z_i) dx_i$

$\int_I x_i p(x_i \mid z_i) dx_i = \int_I x_i p(x_i \mid z_i) dx_i$

$\int_I x_i p(x_i \mid z_i) dx_i = \int_I x_i p(x_i \mid z_i) dx_i$

(4.3.11)

$\int_I x_i p(x_i \mid z_i) dx_i = \int_I x_i p(x_i \mid z_i) dx_i$  可设法通过变换化简以上积分 变成对称形式. 由于

$\int_I x_i p(x_i \mid z_i) dx_i = \int_I x_i p(x_i \mid z_i) dx_i$

因此若令  $X_i = t$  则有  $u = X_i/t$  以及  $f(u, z_2, \dots, z_n) =$

所以在积分(4.3.11)式中, 可令  $u = X_i/t$  则有  $du = - (X_i/t^2) dt$  (4.3.

11)式可得

$E(x_i \mid Z_i) = \int_I x_i p(x_i \mid z_i) dx_i$

$\int_I x_i p(x_i \mid z_i) dx_i = \int_I x_i p(x_i \mid z_i) dx_i$

该式代入(4.3.7)式, 并考虑到当  $z_i = X_i/t$  时,  $X_i > 0$ ,  $X_i = X$  即得

(4.3.10)式.



当  $A = -1$  (即  $\alpha < 0$ ) 时, 证明的方法完全类似, 从略. | 推论  $a$  的最优同变估计可表示为

代

证明

$$S^*(X) = \int_{-\infty}^{\infty} a - (r + i) p(X, a) da / \int_{-\infty}^{\infty} (2r + i) p(X, a) da.$$

可取  $I(X, |)$  为  $\alpha$  的一个同变估计, 因此由 (4.3.8) 式有

$$S^*(X)$$

$$E(I(X, |) | Z)$$

$$I(W).$$

#### 150 第四章最优同变估计

剩下的分析和计算与 (4.3.10) 式的证明完全类似. I 例 4.3.5 设  $X_1, \dots, X_n$  总独立同分布,  $X_i \sim N(0, \theta)$ , 求  $\theta$  的最优同变估计.

解  $X_1, \dots, X_n$  的联合分布为

$$p(x_1, \dots, x_n) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\theta}} e^{-x_i^2/(2\theta)}$$

因此由 (4.3.10) 式有

$$(7) = C_6 \int_0^\infty \theta^n e^{-\sum_{i=1}^n x_i^2/(2\theta)} d\theta$$

$$6T(n+2) d\theta) A(n)$$

$$n+2 \int_0^\infty$$

$$J/\int_0^\infty x(1) \leq 0.$$

例 4.3.6 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim \text{Laplace}$  分布;  $f(x) = e^{-|x|+2a}$  求  $a$  的最优同变估计.

解 设  $I_1, \dots, I_n$  的联合分布为  $p(x, \theta) =$

$$A - X \int A \int I, \text{ 则有}$$

$$a^*(X) = \int_{-\infty}^{\infty} a^2 \dots e^{-\sum_{i=1}^n (2a) |x_i|}$$

$$= \int_{-\infty}^{\infty} (2a) e^{-\sum_{i=1}^n (2a) |x_i|}$$

$$\exp\{-\sum_{i=1}^n |x_i| \int_{-\infty}^{\infty} L(2(7) \int I_i = \int$$

记

令  $l/a = t$  则  $da = -l/t^2 dt$  代入上式可得

$$5 - \int_{-\infty}^{\infty} (X) = \int_{-\infty}^{\infty} (t^n e^{-At} dt) : t^n \int_{-\infty}^{\infty} e^{-At} dt$$

$$r(n-4-1)$$

$$\sim r(n+2) A = n^2 \int_{-\infty}^{\infty} (4). \blacksquare$$

例 4.3.7 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim \text{Exp}(\theta)$ ,  $\theta > 0$ , 求  $\theta$  的最优同变估计.

解 易见  $P(X_j = 8) = e^{-8/\theta}$ ,  $\dots$ , 设的联合分布为

$$X_1$$

,

1 为尺度参数分布族,

$$\theta > 0\}.$$

因此由 (4.3.10) 式有

$$p(\theta, \sum_{i=1}^n X_i) = \theta^n e^{-\theta \sum_{i=1}^n X_i} \int_{-\infty}^{\infty} \theta \int_{-\infty}^{\infty} X_i \int_{-\infty}^{\infty} X_i \int_{-\infty}^{\infty} X_i$$

$$0 - \int_{-\infty}^{\infty} (2a) n$$

$$e^{-4da}$$

#### 4.4 线性变换群下位置尺度参数的最优同变估计

151

m)

i-

$$2n-2$$

$$= \int_{-\infty}^{\infty} T(X(1)).$$

$\langle \eta \rangle = 0$  |  
 $0 > Q \setminus A_0$   
 $r^{nx}; 3z(x(1) - 1)$

#### 4.4 线性变换群下位置尺度参数的最优同变估计

位置尺度参数分布族是最常见的分布族，如正态分布； $V(M, a^2)$ 、均匀分布等都是。这时位置参数要考虑平移变换群；尺度参数要考虑相似变换群，合起来就是线性变换群。

##### 4.4.1 位置尺度参数分布族与线性变换群

设  $\mathcal{P} = \{f, \dots\}$  为位置尺度参数分布族

其中

$f \in (-\infty, +\infty)$ ,  $G \in (0, +\infty)$ , 设  $(\mu, \sigma)$   $e_0 = \{( -\infty, +\infty) \times (0, +\infty)\}$ .

当  $\mu=0, \sigma=1$  时,  $X$  称为标准分布. 特别, 若  $Z = (X - \mu)/\sigma$  则  $Z \sim P(0, 1)$ ; 反之, 若  $X \sim P(\mu, \sigma)$ , 则  $X = \sigma Y + \mu$

本节将要证明: 位置尺度参数分布族关于线性变换群为不变分布族, 并进一步求出相应参数关于线性变换群的MREE. 以下首先介绍样本空间的线性变换群及其导出群, 以及相应的同变估计和同变损失函数, 为求解MREE作准备.

(1) 样本空间的线性变换群. 给定样本空间  $X$ , 定义以下线性变换群  $G = \{g[m, k] \mid k > 0\}$ :

$Y = m + kX$  或  $Y_i = m + kX_i, i = 1, \dots, n$ , (4.4.1)

其中  $m$  表示平移变换,  $k$  表示相似变换.

(2) 参数空间的导出群. 今考虑  $Y = g[m, k]X = m + kX$  的分布. 这

一线性变换为

$(\eta, \theta) \rightarrow (\eta', \theta')$

152

#### 第四章最优同变估计

$Y = m + kX$

$Y^2 = m^2 + k^2 X^2$ ,

$\Rightarrow$ ,

$-Y = m + kX$

$X = \frac{Y - m}{k}$

■

,

$Y$  变换的Jacobi行列式的绝对值为  $|J| = k$ . 由于

以上变换关系代入上式可知

$\frac{1}{k} \frac{d}{dx} \frac{1}{k} \frac{d}{dx}$

其中  $\mu' = m + k\mu$ ,  $\sigma' = k\sigma$ , 由此可知  $X$  的分布为

$1/\sigma' \exp\{-(Y - \mu')^2 / (\sigma')^2\}$

$Y = m + kX$

$Y = kA + BZ$  ...

-

j

$(a)^n \setminus a^r$

, ■

$a, j$

$\sim P[\mu, \sigma]$ ,  $\mu = \mu_0 + k\mu_1$ ,  $\sigma' = k\sigma$ .

这说明,  $r$  的分布仍然是位置尺度参数分布族; 即  $p(\cdot)$  在线性变换群

下为不变分布族.这时, 参数空间的导出群可表示为  $g = : =(\rho, \rho)=(m+k\rho, ka)$ , 或者可表示为

$g[m,k]^{\rho} = \rho n \sim ka$ . (4.4.2) 该式表明, 在线性变换群的作用下, 位置参数要进行相应的线性变换. 但是, 尺度参数只需进行相似变换, 这与相似变换群的作用完全一样.

所以可以预料, 尺度参数的同变估计和同变损失函数的性质也与在相似变换群作用下的性质一样.

(3) 判决空间的导出群及参数的同变估计. 首先要确定所研究的统计问题. 我们考虑两种情形, 即参数  $m, \sigma$  的估计以及  $\rho$  的估计. 这时判决空间、同变条件以及同变损失函数都有所不同. 注意, 参数  $a$  和  $\rho$  的估计的同变性质与相似变换情形基本一样.

(3a) 关于  $\rho$  的估计. 这时,  $D = \rho$ , 因此可定义  $\rho$  记

(4.4.2) 式可得  $\rho = m + k d^{\rho}$ ,  $\rho = k d a$

$d =$  表示  $\rho$  的估计, 则有  $G^* = \rho g M$ .

$x, \rho = -4 - 1k$

, ..., ,

$Yx - m$

由  $\rho$  若记  $\rho$  的估计为

4.4 线性变换群下位置尺度参数的最优同变估计  $A(x), J(x)$ , 则有

,  $g[m,k]a(x) = ka(x)$ ,

153

则由上式及 (4.4.1) 式可知, 同变条件  $g(x) = g^*(x)$  用于  $\rho$  的估计有  $\rho = M(m)$

$\rho = g^* m. A M(x) = m + \rho / z(x)$ , 所以  $\rho$  的同变条

件为

$f(t(m+kx)) = 7n + \rho / x(\rho)$ ,  $V(zn, \rho)$ ,  $A: > 0$ . 同理, (5) 的同变条件为

$\rho - (m + kx) = k \rho a(x)$ ,  $V(m, k)$ ,  $k > 0$ . 特别, 若取  $k = a' \setminus m = -\rho / a$ , 则以上两式化为

$\rho / 2(\rho) - f_i - \rho / i l j_{\rho}$

(4.4.3) (4.4.4)

(4.4.5)

而且  $(X - M) / (T - P(0))$  为标准分布, 与参数  $(M, \rho)$  无关. 这些就是求解  $M$ , 同变估计的约束条件.

关于同变损失函数,  $M$  与  $\rho$  的估计要分别讨论. 若  $\rho$  表示  $M$  的估计, 则应有

$L(\rho, cr; \rho) = L(m + k f \rho, ka; m + k d^{\rho})$ ,  $M, m, k, k > 0$ . 若取  $A = cr_{\rho}$ ,  $m = -f_i / a$ , 则上式化为

$f(\rho, ct; c/m)$  叫  $\rho, l$ ;

因此同变损失函数应为  $(4 - \rho) / tr$  的函数. 通常取  $p(\cdot)$  为均方损失, 则有

$Z(\rho, (7; <))$  (4.4.6)

关于  $a$  估计的同变损失函数, 若  $\rho$  表示  $a$  的估计, 则应有  $L(\rho, a; da) = L(m + k f_i, ka; k da)$ ,  $V, m, k, k > 0$ .

仍取  $A = cr_{\rho}$ ,  $m = -\rho / z / cr$ , 则上式化为  $\rho = \rho, 1$

因此同变损失函数应为  $da / a$  的函数, 这与相似变换群的情况完全一样. 通常取  $p(\rho)$  为均方损失

$f(\rho, < 7; \rho) = 1$

(4.4.7)

(3b) 关于  $\rho$  的估计. 这时,  $P$  为一切  $\rho$  组成, 其变换情况与相似变换群完全一样. 易见当  $a^{\rho} a = ka$  时,  $\rho = \rho (\rho = Tc / \rho)$ ; 若  $d \in V$  为  $\rho$  的估计, 则应有  $W = krd$ , 因此导出群应定义为  $\rho = \rho$ ,

## 第四章最优同变估计

记  $U = krd$ . 特别若  $(\cdot)$  的同变估计记为  $a = P(X)$ , 则有  $i[\cdot] :=$

$krar(X)$ . 因此 (4.1.6) 式的同变条件  $g(gx) = g \circ 5(x)$  化为  $ar(ml + kx) = kT(r(x))$ .

关于  $\lambda$  估计的同变损失函数, 若表示  $\lambda$  的估计, 则应有  $f(i, cr; d) = \beta(\lambda', (7';$

$d') = L(m + Ayx, ka; krd), Wm, k$ .

仍取  $A := a_1, m = -fjb/a_9$  则上式化为

$L(\cdot/x, (r; c) = A(0, l; \cdot) = p(\cdot)$ .

因此同变损失函数应为  $d/a$  的函数, 这与相似变换群的情况完全一样. 通常取  $\lambda(\cdot)$  为均方损失

$-\lambda(\cdot) = (1Z^2)_{\cdot}$ .

## 4.4.2 位置尺度参数的最优同变估计

以下主要求  $\lambda$  和  $a$  的最优同变估计 ( $\ll$  的估计类似, 可作为推论), 其思想方法与前两节十分类似, 即通过不变量及最大不变量,

导出定解条件; 而推导定解条件的出发点为同变条件 (4.4.3) 式与 (4.4.4) 式.

引理 4.4.1 设  $\lambda(\cdot)$ ,  $\lambda(\cdot)$ ;  $\lambda(\cdot)$  为线性变换群下  $M$  和  $cr$  的同变估计, 贝!J

$\lambda(\cdot)$ 、 $\lambda(\cdot)$   $o - \lambda(x)$   $a_2(x)$

$\lambda(\cdot)$ 、 $M_i(x)$   $A_2(\lambda)$

$a(x) = \text{-----}, v(x) =$

为不变量. 反之, 若  $a, (\cdot)$ ,  $\lambda(\cdot)$  为同变估计,  $u(x)$ ,  $v(r)$  为不变量, 则

$\lambda(\cdot)$   $A_i(x) + a_1(x)u(x)$ ,  $a_2(x) = 5 - 1(x)v(x)$  为同变估计.

证明 由同变条件 (4.4.3) 式与 (4.4.4) 式可得

$\lambda(\cdot)$ ,  $\lambda(\cdot)$   $+ kx) - \lambda(7nl + kx) u\{ml + kx) = \text{-----}$

$cr, (ml + kx)$

$m + k/l^x) - (th + A: / Z_2(\cdot))$

$k ax(x)$

$A_1(\cdot) - \lambda(\cdot) z, = \text{-----} : \text{-----} = U(X)$ .

同理可证  $\lambda(\cdot)$  为不变量. 反之,

,

$\lambda(\cdot)$  为不变量, 则有

$\lambda(x)$

若  $\lambda(x)$  为同变估计, 以  $\lambda$

## 4.4 线性变换群下位置尺度参数的最优同变估计

155

$\lambda(12(ml + kx)) = \lambda(12(ml + kx)) + \lambda(ml + kx')u(ml + kx) + kjjbv(\cdot) + kai$

$(x)u(\cdot)$

$= zn + k/JL_2(x)$ .

因此  $A_2(\cdot)$  为同变估计, 同理可证  $\lambda(\cdot)$  为同变估计. I 引理 4.4.2  $u(x)$ ,  $V(\cdot)$  为线性变换群下不变量的充要条件为

$u(x) = \lambda(f(V))$ ,  $v(x) = (p(V))$ , 其中

$\lambda(\cdot) = (H \cdot) = 'h - x$ ,

(4.4.8) 为最大不变量, 且  $W(\cdot)$  及  $r(x)$  的分布均仅与标准分布有关, 与  $(M, a)$

无关, 即为辅助统计量.

证明 由不变性可知,  $U(\cdot)$   $x_2$

$m + kx_n$ , 若取  $k = l/\lambda^2 - xx$  和  $77z = -x$

, ..., X<sub>n</sub>) = U

I X<sub>2</sub> -

, 无 J = a(m + kx<sub>1</sub>, m + kx<sub>2</sub>, I, 则有

(X<sub>1</sub> - 1)/a = P(0 > 1)

X<sub>y</sub> = X<sub>y</sub>

太3 -

' |X<sub>2</sub> - %| | '

X<sub>n</sub>

%2 - %! 1/

X, %3 - X<sub>j</sub> x<sub>2</sub>

0,

同理可证 v(x) = {p(V)}. 易见 F 为不变量, 因为

- x, (m + kx<sup>2</sup> - (th + A; %!)) V = -----

----- \* x<sub>2</sub> - X<sub>j</sub> (m + kx<sub>2</sub>) - (m<sup>4</sup> - kxx)

..... 1 X<sub>2</sub> - 尤1

"X<sub>1</sub> \ X<sub>2</sub> ~ X<sub>1</sub> 1/

x<sub>2</sub> - xx \ fx<sub>2</sub> - x{

= u(0, V<sub>2</sub>, V<sub>2</sub>V<sub>3</sub>, ..., V<sub>2</sub>V<sub>J</sub>) = <sup>^</sup>(V).

由于 u(幻为任一不变量, 因而 V 为最大不变量. 又由 u(x) = u(zn<sub>1</sub> + kx), 取女 = cr<sub>1</sub>, m

= -/z/a, 则该式化为 u(x) = u((x - 好<sub>1</sub>)/cr), 而

. 所以 u(幻及 t; (x) 的分布都仅与标准分布有关, 与 (M, ) 无关, 即为辅助统计量. 壓

引理 4.4.3 设队 X), 汐(%) 为某同变估计, 则任一同变估计可表示为 A\*(x) = A(X) + (V),

(7\*(%) = a(x)(p(V)).

证明由以上两个引理的结果可得

A\* (x) = M(<sup>^</sup>) + 多" ) ± (2 了") = 4(%) + a(x)i/(V), o-(x)

(7\*(%)

= a(x)(p(V). |

x<sub>2</sub> - x<sub>t</sub>

x<sub>3</sub> - x

X<sub>n</sub> ~ X<sub>1</sub> \ T

由以上引理可归纳出求解 (/x, a) 的最优同变估计的方法:

(1) 取某个特定的同变估计 /1(幻, ±(X);

(2) 令 <sup>^</sup>(X) = /l(X) + a(X)<sup>^</sup>(V), a\*(X) = a(X)<sup>^</sup>(V), 求

少(•) 以及史(•) 使均方误差最小.

X<sub>2</sub> - 尤1

x

x<sub>2</sub> - X | x<sub>2</sub>

• • • , 9

<;

156

#### 第四章最优同变估计

定理 4.4.1 (Pitman) 设 /1(X), ±(X) 为 /f, a 在线性变换群下的

同变估计, 则在均方损失下, M, o■ 的最优同变估计可表示为

(4. 4.9) (4. 4. 10)

"(4) 调

E(0J) [a(X) | r]

并且解唯一, 与 z<sub>1</sub>(x) 及 J(J) 的选取无关.

证明 今取 A\*(X) = A(x) + 5-(x) « A(n, 并求 ☆(•) 使均方误差最 小, 由 (4.4.6) 式, 均方误

差似SE可表示为

d. , 因此有

$MSE = E(0J) [1/T(X)]^2 = E(0J) [A(X) + a(X)(A(V))]^2$ . 应用引理4.3.4, 其中 $a(X) = (X - M)/\sigma$ ,  $6(X) = a(X)$ , 因此当 $W(V)$ 取以下函数时MSE达到最小值:

$E(0J) \frac{g(X)}{V} = E(0J) \frac{f(w)}{V}$ .

该式代入 $V = W$ 中即得(4.4.9)式. 同理取 $a^*(x) = a(x) < p(V)$ , 求 $p(\cdot)$ 使均方误差最小, 由(4.4.7)

式, 均方误差MSE可表示为  $E[(X - M)^2] = \sigma^2$ , 小. (弓午1] $^2$ ,

由于 $(X - M)/\sigma \sim P(0 > 1)$ , 因此有

$AfSE = E(0J) [a(X) - 1]^2 = E(0J) [a(X)V - 1]^2$ . 应用引理4.3.4, 取

$a(X) = -1, b(X) = a(X)$ , 因此当 $W(V)$ 取以下函数时 $AfSE$ 达到最小值:

$W = E(1/t, a$

( $\frac{1}{t} - \frac{1}{a}$ ), 本

2

>

以上用到了(4.4.5)式. 由于 $(X - M)/\sigma \sim P(0 < 1)$

知.  $E(0 > 1) [a^2(X) | V] =$

由此可得(4.4.10)式. 唯一性的证明与前面两个定理完全类似, 从略. 推论1 必 $*$  ( $r$ )

和多 $*$ ( $x$ )可表示为

$E(0J) [X^2 - X_1 | V] = E(0J) [X^2 | V]$

$A(\cdot) = \frac{1}{2} - \frac{1}{2}x^2$

$-x, |$

(4.4.11) (4.4.12)

$\pm(X) | V] = E(0J) [X^2 | V]$

4.4线性变换群下位置尺度参数的最优同变估计

157

因为可取 $1/(X - M)$ ,  $\pm(X - M) = \frac{1}{2} - \frac{1}{2}x^2$  作为 $M$ 和 $f$ 的同变估计, 则得到以上公式.

推论2 若 $5(X)$ 为 $\mathcal{X}$ 的某一同变估计, 则 $\mathcal{X}$ 的最优同变估计为  $3*(X) = d(X) E(0J)$

$[g(X) | V]$

$E(0J) \frac{g(X)}{V} = E(0J) \frac{f(w)}{V}$

推论3 若 $T = T(X)$ 为完备充分统计量, 且 $M = a(H)$ ,  $a(X) =$

$6(T)$ , 则以上定理及推论中的条件期望可改为无条件期望. 因为 $V$ 为辅助统计量, 从而与

$a(D)$ ,  $6(T)$ 独立(Basu定理). 例4.4.1设 $X_1, \dots, X_n$ 独立同分布,  $X_i \sim N(\mu, \sigma^2)$ , 求 $(\mu, \sigma^2)$ 的

最优同变估计.

解 可取 $T = (y, s^2)$ ,  $s^2 =$

$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$  则 $r$ 为完备充分统计

$i = 1$  量, 且 $\bar{x}$ 为 $\mu$ 的同变估计, 因为由 $X_i = \mu + \sigma Z_i$ , 有 $\bar{x} = \mu + \sigma \bar{Z}$ ;  $s^2$

为 $\sigma^2$ 的同变估计, 因为由 $\bar{x} = \mu + \sigma \bar{Z}$ , 有 $s^2 = \sigma^2 \frac{1}{n} \sum_{i=1}^n (Z_i - \bar{Z})^2$ ; 也有 $s^2 \sim \chi^2_{n-1}$ . 因此由定理4.4.1

以及推论3有

$1/T(X) = X$ .

因为 $\mu = 0$ , ( $r = i$ 时,  $s^2 \sim \chi^2_{n-1}$ ), 因此由推论2有

$p^*(x) = E(0J) \frac{g(x)}{V} = \frac{1}{2} - \frac{1}{2}x^2$ ,  $E(0 > 1) (54) n + 1$

本例中 $a_2$ 的估计比极大似然估计以及一致最小方差无偏估计都要好.

要好.

例4.4.2 设  $(M, a)$  的最优同变估计.

解 取  $r = (x(1), s)$ ,  $s =$

独立同分布,  $X: +r(l/a, l)$ , 求  
(A - 尤(1)), 则  $r$  为完备充分统计

1=1 量. 可取  $X(1)$ ,  $S$  分别为  $\mu$  和  $a$  的同变估计:  $\mu(X) = S$ .

注意,  $X(1) \sim +r(a/r, l)$ ,  $S \sim r(l/a, n-1)$ , 且独立. 以上结果代入 (4. 4. 9) 式和 (4. 4. 10) 式可得

$$E(o, i)(U) A* W = \hat{(\cdot)} - 5 E(o, i) (S^2) \\ (X) = X(I) - \\ nn$$

下表给出了  $\mu$  和  $a$  极大似然估计 (MLE)、一致最小方差无偏估计 (UMVUE) 以及最优同变估计 (MREE) 这三种估计的比较, 按均方误差

$$a* (X) = -S. \\ m) \\ E(o, i) (S) = s E(0J) (S^2) \\ 158$$

#### 第四章最优同变估计

最小准则, 最优同变估计都是最好的. 估计

$$a \\ UMVUE \\ MLE \quad MREE \\ Y \\ 1 - \hat{r} S (1) \quad n(n-1) S^{n-1} \\ Y A(l) \\ \_ \_ s \quad n^2 \quad \circ \\ 又(1) \\ Is \quad n \\ T5$$

#### 4.4. 3 Pitman积分公式

与前两节类似, 我们可应用公式 (4. 4.11) 和 (4. 4. 12), 经过直接计算把  $A*$  (幻和表若示为积分形式如下. 定理 4. 4.2

则  $MW$  的最优同变

估计可表示为

$$\mu (幻) = \int \int \int \frac{1}{z(r \sim 3p (X ; /f, a))} d/it \dots \dots \dots , \\ J \quad da \quad a \sim 3p(X) \\ a \sim 2p(Xdp, \\ W: a \sim 3p(X; /jL, a) d/jL$$

注 本定理的证明与定理 4.3.2 的证明十分类似, 只是计算更加复杂一些. 根据公式 (4. 4. 11) 和 (4. 4. 12), 其关键是求出  $(-X,$

$$I \quad V2, v3, -, vn) \text{ 的条件分布. 由于} \\ r/ \quad ^2 \quad -^1 \quad X. \quad -X, \\ z=3' \dots \text{人}$$

因此可先求  $(\wedge, \wedge^2 -X, I \quad V3, v4, \dots, VJ)$  的条件分布. 为此可通过变换先求  $(\wedge, \wedge^2$  的联合分布. 再经过若干计算即得以上公

式, 进一步的细节从略. 在第八章, 我们将从 Bayes 观点出发, 更加快捷地推导出 Pitman 积分公式.

#### 习题四

1. 设  $X, X'', X$ , 为  $i. i. d.$  样本,  $A$  服从位置参数分布族  $f(X1 -$

;  $f(a)d/jL$   
习题四 e). 令

159

$5(hx^3)$

弋,  $X_2$ ,

$X_3 > 0$ ,  $x_3^0$ .

将作为 $\theta$ 的一个估计, 证明: 当损失函数为 $L(\theta - \hat{\theta})$ 时,  $\hat{\theta}(X_1, X_2, X_3)$ 的风险函数与 $\theta$ 无关, 但 $\hat{\theta}(X_1, X_2, X_3)$ 不是 $\theta$ 的同变估计. 2. 设样本 $X = (X_1, \dots, X_n)$ 服从位置参数分布族/ $(\theta - \theta_0, \dots, x_n - \theta_0)$ .

6).  $H(\theta)$ 为 $\theta$ 的同变估计, 取损失函数为 $L(\theta, d) = L(d - \theta)$ . 证明:  $H(\theta)$ 的偏差、方差以及 $H(X)$ 的风险函数都与 $\theta$ 无关.

3. 设样本 $JV = (X_1, \dots, X_n)$ 服从位置参数分布族/ $(\theta - \theta_0, \dots, x_n - \theta_0)$ . 若 $r(X)$ 为关于 $\theta$ 的充分统计量,  $(X)$ 为 $\theta$ 的Pitman估计, 证明:  $(X)$ 是 $r(x)$ 的函数

4. 设 $A_1, \dots, A_n$ 为i.i.d.样本,  $X_i$ 服从均匀分布中 $-\theta$ 到 $\theta$

其中未知. 在均方损失下, 求 $\theta$ 的MREE.

5. 设 $X_1, \dots, X_n$ 为i.i.d.样本,  $X_i$ 的密度函数为 $f(x)$

$x^{\theta}$ , 在均方损失下, 求 $\theta$ 的MREE.

6. 设 $X_1, \dots, X_n$ 为相互独立的样本,  $X_i \sim N(\mu, \sigma^2)$ ,  $Z = 1, \dots, n$ , 其中 $\mu$ 未知,  $\sigma^2$ 已知. 在均方损失下, 求 $\mu$ 的MREE.

7. 设 $X_1, \dots, X_n$ 为i.i.d.样本,  $X_i \sim N(\mu, \sigma^2)$ ;  $Y_1, \dots, Y_m$ 为i.i.d.样本,  $Y_j \sim N(c/\sigma^2, \sigma^2)$ ; 并且合样本独立,  $c \neq 0$ ,  $c \neq \infty$ 已知. 在均方损失下, 求 $\mu$ 的MREE.

8. 设 $X_1, \dots, X_n$ 为i.i.d.样本,  $X_i$ 的密度函数为 $f(x)$ 其中未知, 在均方损失下, 试求 $\mu$ 的MREE.

\*9. 设 $X_1, \dots, X_n$ 为i.i.d.样本,  $X_i$ 服从指数分布/ $(x; \lambda, \theta) =$

$\lambda e^{-\lambda x}$ , 其中 $\lambda$ 已知,  $\theta$ 为未知的位置参数. (1) 在均方损失 $L(\theta, d) = (d - \theta)^2$ 下, 求 $\theta$ 的MREE;

(2) 在绝对损失 $L(\theta, d) = |d - \theta|$ 下, 证明:  $\theta$ 的MREE为 $X(1) - \ln 2 / (n\lambda)$  (提示: 直接用MREE的定义, 求 $L(\theta, d)$ 使最小, (3) 和第 10 题亦类似); \*

(3) 若损失函数为 $L(\theta, d) = 1 - |d - \theta|$ , 其中 $\lambda$ 为已知正数. 证明:  $\theta$ 的MREE为 $V(1) - 4$ .

10. 设 $X_1, \dots, X_n$ 为i.i.d.样本,  $X_i \sim N(\mu, \sigma^2)$ , 其中 $\mu$ 未知,

160

第四章最优同变估计

$c^2 > 0$ 已知. 若损失函数为

$J(\theta, d) = A/(d - \theta) =$

$-a/(JL - d)$ ,  $-d$ , 其中 $a$ ,  $c$ 为正常数, 求 $\theta$ 的MREE.

• 11. (1) 定义函数 $\phi(a) = E[P(J - a)]$ , 其中 $p$ 为 $(-\infty, +\infty)$ 上

的凸函数. 证明: 若 $p$ 为偶函数, 且 $X$ 的密度函数关于 $\theta$ 对称, 则 $\phi(a)$ 在 $a = M$ 处达到最小值 (提示: 对某实数 $c$ , 设 $\phi(a)$ 在 $a + c$ 处取最小值, 先证 $\phi(a + c) = \phi(a - c)$ , 然后用 $P$ 的凸性证 $\phi(a + c) > \phi(a)$ );

(2) 设样本 $JV = (X_1, \dots, X_n)$ 服从位置参数分布族/ $(\theta - A)$ . 取损失函数为 $L(d - \theta)$ , 其中 $L(\theta)$ 为凸的偶函数. 若 $r_Q(x)$ 为 $\theta$ 的同变估计, 其密度函数关于 $\theta$ 对称, 并假设 $r_Q(x)$ 与 $r$ 独立, 其中 $Y = (X_2 - X_1, \dots, X_n - X_1)$ , 则 $Jr_Q(x)$ 为 $\theta$ 的MREE. \*12. 设样本





(3) 当 $y$  都未知时, 求均方损失下 $(M, a)$ 的MREE; (4) 分别求各个估计值的风险函数.

### 第五章点估计的性质

本章讨论点估计的若干性质. 我们知道, 一个未知参数的估计量有很多种, 可参见第三章、第四章. 在理论上或实用上, 什么样的估计量 最好呢? 直观上讲, 应该是估计量与被估计量越接近越好. 但是“接近”两字的含义实际上涉及估计量的优良性标准, 这种优良性标准在文献中有很多讨论, 本书第三章、第四章也已经作了一些介绍. 第三章 曾经介绍统计判决函数以及使风险函数最小的优良性准则, 这是一个很一般性的准则, 其优良性要结合损失函数的选取进行研究. 第三章也介

绍了估计量的无偏性, 它要求估计量与被估计的真参数之间的平均偏差为零, 但是无偏估计也包括 $f$  左右偏差可能都比较大的情况. 所以在第三章和第四章, 我们着重介绍了估计量的均方误差(即平方损失函数)

最小准则, 其中包括UMRUE和MREE. 本章将介绍与估计量的均方误差有密切关系的有效性问题. 另外, 极大似然估计和矩方程估计是从其他统计观点提出的, 与均方误差最小准则无关. 本章将从大样本观点(即考虑样本容量 $n$  趋向 $+\infty$ 时的情形)阐明估计量与被估计量“充分

接近”的意义, 即估计量的相合性和渐近正态性, 并应用于极大似然估计和矩方程估计. 关于有效性问题, 主要研究当样本容量 $n$  给定时, 估计量以 $\theta$ 或 $\theta_0$ 为幻的风险函数(主要是均方误差MSE)可能有多小? 原则上讲, MSE 当然是越小越好, 但事实上是不可能的. 可以证明, 大多数估计量的最

小均方误差不可能无限制的小, 而是有一个下界. 因此, 达到下界的估计量最好, 称为“最有效”. 本章第5.1节和第5.2节主要讨论估计量的均方误差的下界问题, 介绍C-R不等式以及广义的OR型不等式.

关于估计量的渐近性质(或大样本性质), 主要考虑样本容量 $n$  趋向 $+\infty$ 时, 估计量 $\dots, \hat{\theta}_n$ 是否能在某种意义下收敛到真参数 $\theta_0$ , 称为相合性. 再就是渐近正态性, 即考虑 $Z_n = \frac{\hat{\theta}_n - \theta_0}{\sigma/\sqrt{n}}$ 的渐近分布, 对

于很多情况可以证明,  $Z_n$ 趋向于正态分布, 这时称 $\hat{\theta}_n$ 具有渐近正态性. 本章第5.3节将讨论这方面的问题, 并应用于极大似然估计和矩方程估计. 有关本章内容, 可参见陈希孺(1981, 1999), Lehmann (1999), Shao (1998)等.

#### 5.1 C-R不等式

163

#### 5.1 C-R不等式

设  $X = (X_1, \dots, X_n)^T$ ,  $X \sim f(x, \theta)$ , 参数  $\theta \in \Theta$  和  $g(\theta)$  的估计分别记为 $\hat{g}(X)$ 和 $g(X)$ . 第三章曾经证明(见3.2节), 幻的均方误差 可表示为

$MSE(g(X)) = Var(g(X)) + [bias(g(X))]^2$ . 我们要求估计量 $g(X)$ 的均方误差尽量小, 相当于要求其方差与偏差平

方之和尽量小; 当然也应要求方差 $Var(g(X))$ 尽量小. 特别是无偏估计, 由于  $bias(g(X)) = 0$ , 因此  $MSE(g(X)) = Var(g(X))$ , 所以, 要求估

计量的均方误差尽量小, 就等价于要求其方差尽量小. 本章将证明, 不管 $g(X)$ 的形式如何, 其方差 $Var(g(X))$ 都不可能无限制地很小, 而是有一个下界, 即恒有 $Var(g(X)) \geq C$ 对一切 $g(X)$ . 因此能达到方差下界的估计量 $g(X)$ 最好, 称为“有效估计量”. 若 $n \rightarrow \infty$ 时,  $Var(g(X)) \rightarrow C$ , 则称 $g(X)$ 为“渐近有效的”. 以下集中研究估计量 $g(X)$ 的方差下界问题.

本节假定 $H_0: \theta = \theta_0$  为正则分布族(即C-R分布族, 详见第二章), 其基本假定摘要如下:

(1)  $\theta_0$ 有共同支撑, 即  $S = \{x: f(x, \theta_0) > 0\}$  (与  $\theta$  无关); (2)  $L(\theta) = L(\theta_0, x)$

$=\log f(x, \theta)$  关于  $\theta$  连续、可导；

去以  $e$  存在二阶矩，且有  $E[S(X, \theta)] = 0$ ， $E[\{S(X, \theta)\}^2] = I(\theta)$

(4)  $I(\theta) = -E[\partial^2 \log f(X, \theta) / \partial \theta^2]$  为 Fisher 信息阵；

(3)  $f(x, \theta)$  关于  $\theta$  求导数与关于  $x$  求积分可交换次序。以下为常见的两个不是 C-R 分布族的例子：

(1) 均匀分布  $X \sim R(\theta, \theta)$ ：

$1/\theta \leq x \leq \theta$ ， $f(x, \theta) = 1/(\theta - x)$

$\theta < 0$  或  $\theta > 0$  这时与 (2) 有失。另外， $f(x, \theta)$  作为  $\theta$  的函数在  $e^x$  处不连续，因为  $f(x, \theta) = 1/\theta$ ，当  $\theta \rightarrow 0$  时，而  $f(x, \theta) = 0$ ，当  $\theta \rightarrow \infty$  时。

(2) 带位置参数的指数分布  $X \sim \text{Exp}(\theta)$ ，取  $A=1$ ，

$f(x, \theta) = \theta e^{-\theta x}$ ， $I(\theta) = 1/\theta^2$

$f(x, \theta) = e^{-\theta x}$ ， $I(\theta) = 1/\theta^2$

$1/\theta$ ， $X < \infty$ 。

这时  $I(\theta) = [j\theta, +\infty)$ ，与  $\theta$  有关。另外， $f(x, \theta)$  作为  $\theta$  的函数在  $f\theta^x$  处

则有

164 第五章点估计的性质

不连续，因为  $M(\theta) = e^{-\theta}$ ，当  $\theta \rightarrow 0$  时， $I(\theta) = 0$ ，当  $\theta \rightarrow \infty$  时。

5.1.1 单参数 C-R 不等式

本节主要讨论以  $\theta$  或  $\theta$  为无偏估计的情形，这时估计量的方差就是其均方误差。至于有偏

估计  $H(\theta)$  的情形，可作为推论很容易得到。本节的重要结果之一是  $\text{Var}(\hat{\theta}) \geq C-R$  下界  $= I(\theta)$ ，

其中以  $\theta$  为  $\theta$  的无偏估计， $K(\theta)$  为观察样本的 Fisher 信息阵。另外，本节的证明主要应用

得分函数 (即 score 函数)  $S(x, \theta)$  的定义与性质：

$-dL(\theta, X) = \log f(x, \theta) = -L(\theta, X)$

且有

$E[S(X, \theta)] = 0$ ， $\text{Var}[S(X, \theta)] = E[S(X, \theta)^2] = I(\theta)$ 。

引理 5.1.1 设  $\theta$  为 C-R 分布族， $i(\theta)$ ， $g(\theta)$  分别为  $\theta$  和  $\theta$  的无偏估计，且导数  $g'(\theta)$  存在，则有

$\text{Cov}[i(\theta), S(X, \theta)] = E[g(\theta) S(X, \theta)] = g'(\theta)$  (5.1.1)  $\text{Cov}[i(\theta), S(X, \theta)]$

$= 1$ 。(5.1.2) 证明 主要证明 (5.1.1) 式，因为令  $g(\theta) = \theta$  即可由 (5.1.1) 式得

到 (5.1.2) 式。由 score 函数  $S(x, \theta)$  的定义与性质可得  $\text{Cov}[i(\theta), S(X, \theta)] = E[(i(\theta) - E[i(\theta)])(S(X, \theta) - E[S(X, \theta)])]$

$= E[(i(\theta) - E[i(\theta)])(S(X, \theta) - E[S(X, \theta)])]$

$= E[(i(\theta) - E[i(\theta)])(S(X, \theta) - E[S(X, \theta)])]$

$= E[(i(\theta) - E[i(\theta)])(S(X, \theta) - E[S(X, \theta)])]$

$= E[(i(\theta) - E[i(\theta)])(S(X, \theta) - E[S(X, \theta)])]$

$= E[(i(\theta) - E[i(\theta)])(S(X, \theta) - E[S(X, \theta)])]$

推论 若  $i(\theta)$  和  $S(x)$  的偏差分别为  $b(\theta)$  和  $b(\theta)$ ，即  $E[j(\theta)] = g(\theta) + b(\theta)$ ， $E[S(X, \theta)] = E[S(X, \theta)]$ ，

$\text{Cov}[i(\theta), S(X, \theta)] = g'(\theta) + b'(\theta)$ ，

$\text{Cov}[j(\theta), S(X, \theta)] = 1 + b'(\theta)$ 。本节将要证明几个 C-R 不等式，它们都可统一地应用以下引理，

该引理经直接验证即可得到。

$\int_0^\infty W(x, \theta) d\theta = 1$

$\int_0^\infty W(x, \theta) d\theta = 1$

5.1.1 CR 不等式

165

引理 5.1.2 若随机变量  $X$ ， $F$  的二阶矩存在，则 Schwarz 不等式可

表示为

$[\text{Var}(X)][\text{Var}(F)] \geq [\text{Cov}(X, F)]^2$

$\geq [\text{Cov}(X, F)]^2$ ，(5.1.3)

其中

$A = \text{Cov}(\hat{\theta}, y) \text{Var}^{-1}(r)$ , (5.1.4) 且等式成立的充要条件为  $x = \text{常数}$  (a.e.). 特别, (5.1.3) 式等

价于

$\text{Var}(X - Ay) \geq 0$ ,  $A = \text{Cov}(X, F) \text{Var}_F^{-1}(\hat{\theta})$  (5.1.5) 另外, 以上结果对  $v, y$  为向量的情形也成立. I 定理 5.1.1 设  $\theta$  为 C-R 分布族,  $g(\theta)$ ,  $\hat{\theta}(X)$  分别为  $g(\theta)$  和  $\theta$  的无偏估计, 且导数  $g'(\theta)$  (约存在, 则以下不等式成立:  $\text{Var}[\hat{\theta}(X)] \geq [g'(\theta)]^2 / \text{Var}[g(\theta)]$ , (5.1.6)

$\text{Var}[\hat{\theta}(X)] \geq [g'(\theta)]^2 / \text{Var}[g(\theta)]$ , 以上不等式中等式成立 (即方差达到下界) 的充要条件分别为  $\hat{\theta}(X) = g'(\theta) / \text{Var}[g(\theta)]$ ,  $S(x, \theta) = a(\theta) [g'(\theta)]^2$ .

(5.1.7)

(5.1.8)

(5.1.9) 证明 在 Schwarz 不等式 (5.1.3) 中, 以  $y$  取代  $X$ , 以  $S(X, \theta)$  取代  $K$ , 并应用 (5.1.1) 式可得

$[ \int g(\theta) S(x, \theta) dx ]^2 \leq \text{Var}[\hat{\theta}(X)] \int S^2(x, \theta) dx$ .

由于  $\text{Var}[\hat{\theta}(X)] = \text{E}[\hat{\theta}^2(X)] - [g'(\theta)]^2$ , 把它代入上式即得 (5.1.6) 式. 在 (5.1.6) 式中取  $g(\theta) = 0$

即得 (5.1.7) 式. 以上不等式中等式成立的充要条件为  $X - AK = \text{常数}$  (a.e.), 其中  $X = g(\theta)$ ,  $Y = S(X, \theta)$ , 因此  $A = \text{Cov}(g, S) \text{Var}_S^{-1}(S) = g'(\theta) / \text{Var}[g(\theta)]$ , 所以有

$g(x) - g'(\theta) / \text{Var}[g(\theta)] S(x, \theta) = c(\theta)$  (a.e.). 该式两端取期望可得  $\text{E}[g(X)] = g'(\theta) / \text{Var}[g(\theta)]$  因此有

$g(x) - g'(\theta) / \text{Var}[g(\theta)] S(x, \theta) = c(\theta)$  (a.e.). 该式两端取期望可得  $\text{E}[g(X)] = g'(\theta) / \text{Var}[g(\theta)]$  因此有

$g(x) - g'(\theta) / \text{Var}[g(\theta)] S(x, \theta) = c(\theta)$  (a.e.). 该式两端取期望可得  $\text{E}[g(X)] = g'(\theta) / \text{Var}[g(\theta)]$  因此有

$g(x) - g'(\theta) / \text{Var}[g(\theta)] S(x, \theta) = c(\theta)$  (a.e.). 该式两端取期望可得  $\text{E}[g(X)] = g'(\theta) / \text{Var}[g(\theta)]$  因此有

由此即得 (5.1.8) 式, 在 (5.1.8) 式中取  $g(\theta) = 0$  即得 (5.1.9) 式. ■

C-R 不等式是由 Cramer 和 Rao 分别于 1945 年和 1946 年提出的, 不等式有丰富的内涵, 在统计学中有十分重要的意义. C-R 不等式说明, 在 C-R 分布族中,  $g(\theta)$  的任何无偏估计  $\hat{g}(T)$ , 不管形式如何, 其方差, 即均方误差总是有下界的, 其方差下界 (CRLB) 为  $[g'(\theta)]^2 / \text{Var}[g(\theta)]$ . 与估计量无关. 因此方差能达到下界的无偏估计必然是最优的. 另外, 由表达式可知, 方差下界与样本信息成反比, 样本信息越多, 方差下界越小, 因而可能达到的方差, 即可能达到的均方误差越

## 166 第五章点估计的性质

小; 这显然是很合理的. 特别, 参数  $\theta$  的无偏估计  $\hat{\theta}$  可能达到的最小均方误差为  $\text{Var}[\hat{\theta}]$  即 Fisher 信息的逆.

根据以上分析可提出下列有效性的定义: 定义 5.1.1 设  $\hat{\theta}(X)$  为  $g(\theta)$  的无偏估计, 若其方差达到 C-R 下

界, 即

$\text{Var}[\hat{\theta}(X)] = [g'(\theta)]^2 / \text{Var}[g(\theta)]$ , 则称  $\hat{\theta}(X)$  为  $g(\theta)$  的有效无偏估计. 特别, 若  $\text{Var}[\hat{\theta}(X)] = [g'(\theta)]^2 / \text{Var}[g(\theta)]$ , 则称  $\hat{\theta}(X)$  为参数  $\theta$  的有效无偏估计. ■

推论 1 若  $\hat{\theta}(X)$  和  $\hat{\theta}(X)$  分别为  $g(\theta)$  和  $\theta$  的有效无偏估计, 则必为一致最小方差无偏估计, 但反之不一定对 (见下面例 5.1.1 - 5.1.3).

推论 2 若  $X = (X_1, \dots, X_n)$  的分量独立同分布, 而  $\hat{\theta}(X)$  的 Fisher 信息为  $I(\theta)$ , 则有  $\text{Var}[\hat{\theta}(X)] \geq [g'(\theta)]^2 / n I(\theta)$ ,  $n$

$\text{Var}[\hat{\theta}(X)] \geq [g'(\theta)]^2 / n I(\theta)$ .

这个推论说明, 方差下界与样本容量成反比, 样本容量越大, 方差下界越小, 因而可能达到的方差, 即可能达到的均方误差越小; 若样本容量趋向于无穷, 则可能达到的均方误差会趋向于零.

推论 3 若  $g(\theta)$ ,  $\hat{\theta}(X)$  为有偏估计, 偏差分别为  $b(\theta)$ , 则 C-R 不等式可表示为

$\text{Var}[i(X)] \geq [g'(\theta) + 6 : (4)]^2 / - | (8), \quad \text{Var}[\hat{X}] \geq [1 + 6' (8)]^2 / (8).$

若样本独立同分布，则可用  $1(\theta) = n i(\theta)$  代入上式。

把引理5.1.1的推论代入到以上定理的证明中即可得到上式。另

外，此时的C-R不等式与具体的估计量的形式  $g(X)$ ， $g(X)$  有一定关系，因为表达式中有偏差项。

推论4  $g(\theta)$  的无偏估计的方差下界在参数变换下保持不变，即若

有变换  $\eta = e(T)$ ， $g(\theta) = g(\eta)$ ，并记参数  $\eta$  的 Fisher 信息为  $I(\eta)$ ，则有

$\text{CRLB} = [g'(\eta)]^2 / I(\eta) = [g'(\eta)]^2 / (n I(\eta))$  (5.1.10)

证明 由于  $I(\eta) = [g'(\eta)]^2 / \text{Var}(g(\eta))$ ，且  $I(\eta)$  的 Fisher 信息满足关系

5.1 C-R不等式

167

$\text{Var}(g(\eta)) = \text{Var}(g(\theta)) = \text{Var}(g(\eta))$ ，且  $I(\eta) = I(\theta)$ ，因此

I

的无偏估计的C-R下界。

« 首先考虑参数  $\theta$  的无偏估计，即  $g(\theta)$  取  $A = \theta$ ，这时

$\text{Var}(X) = \sigma^2/n$ ，而  $I(\theta) = n/a^2$ ，因此有

$\text{CRLB} = \sigma^2/n = \text{Var}(X)$ ，

即的方差达到C-R下界，因而为有效的无偏估计。但是，的情况不一样，取  $g(a) = a^2$ ，其无偏估计为

$P = \bar{X}^2 - \sigma^2/n$

它是一致最小方差无偏估计，且有  $\text{Var}(P) = 2\sigma^4/(n-1)$ 。而此时  $I(a) = n/a^4$ ，因此

因此

$\text{CRLB} = \sigma^4/n = \text{Var}(P) < \text{Var}(P)$ ，

即  $P$  的方差达不到C-R下界，不是有效的无偏估计，但是可认为是“渐近有效”的，因为

$\text{Var}(P) \rightarrow 0$ ，

$\text{CRLB} \rightarrow 0$ 。

例5.1.2 设  $X_1, \dots, X_n$  独立同分布， $X_i \sim \text{Poisson}(\lambda)$ 。考虑  $\lambda$  和  $e^{-\lambda}$  的无偏估计的C-R下界。

解 (1) 取  $g(\lambda) = \lambda$ ，这时  $\lambda = 1$  为一致最小方差无偏估计，且有  $\text{Var}(\lambda) = \lambda/n$ 。而容易求得  $I(\lambda) = n/\lambda$ ，因此有

$\text{CRLB} = \lambda/n = \text{Var}(\lambda)$ 。

即  $\lambda$  的方差达到C-R下界，因而为有效的无偏估计。(2) 取  $g(\lambda) = e^{-\lambda}$  由第三章的例题可知， $g(X) = (1 - \sigma^2/n)$  为

$e^{-\lambda}$  的一致最小方差无偏估计，其中  $r = \lambda$ ，直接计算可得  $I(\lambda) = n$

$\text{Var}[i(X)] = e^{-2\lambda}(e + 1)$ 。

而C-R下界为

$\text{CRLB} = [g'(\lambda)]^2 / I(\lambda) = e^{-2\lambda} / n$ 。

因此  $S(X) = (1 - \sigma^2/n)$  不是有效的无偏估计；但也是“渐近有

效”的，因为

168

效”的，因为

第五章点估计的性质

$\text{Var}[\hat{X}] \geq \text{CRLB} = 1/n$

根据以上两个例子，我们引进以下渐近有效性的定义：定义5.1.2 若  $i(X)$  为  $g(\theta)$  的无偏估

计, 且有

$\text{CRLB Var}[i(X)]$

则称  $j(X)$  为渐近有效的无偏估计,  $e(g)$  称为估计量  $i(X)$  的效率. ■ 例 5.1.3 (反例) 设

$1, \dots, X$  独立同分布, 均匀分布  $(0, 1)$ ,

②. 这时  $i(0) = j$  不是 C-R 分布族, 取  $g(0) = 0$ , 则  $i(X)$

为  $i$  的一致最小方差无偏估计. 因为所以可计算出 方差

$n(n+2)$  另一方面, 我们也可“形式上地”计算出  $i$  的“Fisher 信息”如下.

$\text{Var}[i(X)] \approx \frac{1}{n I(i)}$

在  $(0, 1)$  上有  $f(x) = j, 0 \leq x \leq 1$ ,  $E[-\log f(x)] = 1$

“形式上地”记  $i(0) = 1/192, 7(0) = n/6 > 2$ , “CRLB” =  $6 > 2/n$ . 这

时  $i(X)$  的方差远远小于形式上的“C-R 下界”:

$\text{Var}[i(X)] \approx \frac{1}{n I(i)}$

$\text{Var}[i(X)] = \frac{1}{n I(i)}$  “CRLB”.  $n\{n+2\} n$

这是因为  $i(0, 0) \neq j$  不是 C-R 分布族, 因而 C-R 不等式 (5.1.7) 不一定成立. 在下面 5.2 节, 我们将介绍广义的 C-R 型方差 inequality, 它可用于 没有共同支撑的分布族, 也包括均匀分布族. ■

### 5.1.2 等式成立的条件

以下进一步研究  $g$  的无偏估计  $i(Z)$  的方差能够达到下界的条件, 即  $j(X)$  为  $i(0)$  的有效的无偏估计的条件. 由以下定理可知, 有效估计的条件十分强设, 从某种意义上讲, C-R 下界可能“偏低”.

定理 5.1.2 常数. 若  $i(I)$  为  $g$  的无偏估计, 且有  $\text{Var}[i(X)] < +\infty$  对一切  $0 \leq \theta \leq 1$ , 则 C-R 不等式中等式对一切  $0 \leq \theta \leq 1$  成立, 即  $i$  为  $g$  的有效的无偏估计的充要条件为  $f(x, \theta)$  服从指数族分布

$f(x, \theta) = h(x) e^{Q(\theta)g(x) - b(\theta)}$

; 为 OR 分布族,  $g(0)$  可导且不为

$f(x, \theta) = h(x) e^{Q(\theta)g(x) - b(\theta)}$ . (5.1.11)

5.1 C-R 不等式

169

证明 由定理 5.1.1 的 (5.1.8) 式可知, C-R 不等式中等式成立的充要条件是

$S(x, \theta)$

$= a(\theta) [g(x)]$  (5.1.12)

$d\theta$

以下证明 (5.1.11) 式与 (5.1.12) 式等价. 易见, 由 (5.1.12) 式导出

(5.1.11) 式需要积分, 而由 (5.1.11) 式导出 (5.1.12) 式需要微分. 以下根据 C-R 分布族的性质说明积分和微分的合理性.

假设 (5.1.12) 式成立, 今取  $X_1, \dots, X_n$  使  $i(x_j) = g(x_j)$ , 因而有  $-S(x_1, \theta) = a(\theta) [g(x_1)]$

由 C-R 条件, 上式左端为  $i$  的连续函数, 因而右端也为  $i$  的连续函数, 所以  $a(\theta)$  可积, 从而对

(5.1.12) 式两端积分可得

$\log f(x_1, \theta) = \log \{h(x_1) e^{Q(\theta)g(x_1) - b(\theta)}\} = Q(\theta)g(x_1) - b(\theta) + \log h(x_1)$ ,

即

因而 (5.1.11) 式成立. 反之, 若 (5.1.11) 式成立, 则有

$\log f(x_1, \theta) = Q(\theta)g(x_1) - b(\theta) + \log h(x_1)$ . (5.1.13)

对  $\theta$  求导有  $\frac{d}{d\theta} \log f(x_1, \theta) = Q(\theta)g(x_1) - b'(\theta) + \frac{d}{d\theta} \log h(x_1)$ . 由

C-R 分布族的性质可知, 上式左端关于  $\theta$  可导, 因而  $Q(\theta)$  可导, 再由 (5.1.13) 知  $b'(\theta) = g(x_1)$  可导. 因此 (5.1.13) 式关于  $\theta$  求导可得

$$s(x, \theta) = Qf(\theta)g(x)$$

该式两端取期望可得

$$0 = (f'(\theta)g(\theta) - f(\theta)g'(\theta)) / Qf(\theta) = Qf(\theta) - g(\theta) / Qf(\theta), \text{ 此即 (5.1.12) 式.}$$

本定理说明, 若用某个无偏估计  $\hat{g}(X)$  估计  $g(\theta)$ , 其方差处处达到 C-R 下界; 只能是指数族分布, 不可能是其他分布

以下定理更进一步缩小了在指数族分布中能够达到 C-R 下界的估计量的范围. 定理 5.1.3 设  $X$  服从指数族分布  $f(x, \theta)$

$g(\theta)$  的无偏估计  $i(x)$  为有效估计的充要条件是

, 则

发(没) =  $\frac{1}{Q} \cdot \frac{1}{3} \cdot g(x) = aT(x) + b$  (5.1.14) 证明 由定理 5.1.1 可知,  $g(X)$  的方差处处达到 C-R 下界的充要

170

第五章点估计的性质

条件为

$$s(x, \theta) = a(\theta) [g(x) - g(\theta)]. \text{ 由于 } f(x, \theta) \text{ 为指数族, 因而有 } S(x, \theta) = Q(\theta)T(x) - b(\theta)$$

可得

该式可化简为

$$Q'(\theta)T(x) - b'(\theta) = a(\theta) [g(x) - g(\theta)].$$

$$g(x) = a(\theta) T(x) \quad (5.1.15)$$

该式对  $\theta$  求导可得  $a'(\theta) T(x) + \frac{1}{3}r(\theta) = 0$ . 取  $\theta = \theta_0$  使  $r(\theta_0) = 0$ ,

则有  $a'(\theta_0) [T(x)] - r(\theta_0) = 0$ , 因而有  $a'(\theta_0) = 0$ , 也有  $\frac{1}{3}r'(\theta_0) = 0$ . 因此  $a(\theta)$

$= a(\theta_0) + \frac{1}{3}r(\theta - \theta_0)$  与  $\theta$  无关, 由 (5.1.15) 式即得

$$= aT(x) + p.$$

所以 (5.1.14) 的第二式成立. 该式两端取期望, 并由  $E[T(X)] =$

$$b'(\theta) / Q'(\theta) \text{ 可得}$$

$$g(\theta) = a(\theta) \cdot \frac{1}{3}r(\theta), \text{ 以 } \theta = \theta_0, \text{ 得}$$

本定理说明, 即使在指数族中, 也只有少数的  $g(\theta)$  及其无偏估计  $g(X)$ , 其方差能够处处达到 C-R 下界.

例 5.1.4 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(0, \sigma^2)$ , 求  $g(\sigma^2)$ , 使其无偏估计处处达到 C-R 下界.

解由假设可知

$$(X_1, \dots, X_n) \sim N_n(0, \sigma^2 I_n) \text{ 为指数族, 其中 } T(x) = x, b(\theta) = \frac{1}{2}n\theta > 0.$$

由定理 5.1.3 可

知, 只有以下  $g(\sigma^2)$  的估计才能够达到 C-R 下界:  $g(\sigma^2) = \frac{1}{Q(\sigma^2)} [b'(\sigma^2)] = \frac{1}{2}n$

此时必有

$$g(\sigma^2) = aT(x) = \frac{1}{2}n = ax + b. \text{ 除此以外, 都不能达到 C-R 下界; 诸如 } g(\sigma^2) = \sigma^2 \text{ 或 } g(\sigma^2) = 1/\sigma^2 \text{ 等等,}$$

都不存在有效的无偏估计. 由以上定理和例题可知, C-R 下界比较低, 应该还可以找到比 C-R 下界更大一些的方差下界. 以下 Bh 不等式及 Bh 下界即为常见的一种.

一种.

### 5.1.3 Bh 不等式

C-R 不等式的证明主要用到引理 5.1.1, 即  $i(u)$  与  $S(x, \theta)$  的关系:

## 5. 1 C-R不等式 171

$\text{Cov}(\hat{g}, EJS) = 0, \text{Var}(S) \neq 0$ , 从而有  
 $\text{Cov}(\hat{g}, S) < \text{Var}(S) \Rightarrow \text{Var}(\hat{g}) < \text{Var}(S)$   
 原则上讲, 上式对任何  $S(X, 0)$  都成立, 若用  $X, \hat{g}$  取代  $S(X^*)$ ,  
 则可得到 C-R 型的不等式:  $(\text{Cov}(\hat{g}, S) \text{Var}(\hat{g}) \text{Var}(S)) \Rightarrow \text{Var}(\hat{g}) \leq \text{Cov}(\hat{g}, S) / \text{Var}(S)$  (5. 1. 16)

关键是  $\text{Cov}(\hat{g}, S(X^*))$  和  $\text{Var}(\hat{g})$  应具有明确的统计意义. 本节以及下一节的许多不等式都基于 (5. 1. 16) 式, 主要是不同的  $S(X)$ , 具有不同的统计意义.

例 5.1.5 设  $f(x, y)$  和  $g(x, y)$  关于  $y$  有  $n$  阶导数  $f^{(n)}(x, y)$  和  $g^{(n)}(x, y)$ . 则有  
 $\text{Var}(\hat{g}) \leq \text{Cov}(\hat{g}, S(X, y)) / \text{Var}(S(X, y))$

其中  $\hat{g} = \text{Var}(\hat{g})$ ,  $S(X, y) = f(x, y) / f(x, 0)$ .

证明 在 (5. 1. 16) 式中取  $S(X, y) = S(X, 0)$ , 则有  
 由假设可得

$$E[S(X, y)] = \int_0^1 f(x, y) dy / \int_0^1 f(x, 0) dy = 0,$$

$$\begin{aligned} & \int_0^1 f(x, y) dy = \int_0^1 f(x, 0) dy = 0 \\ & = -\int_0^1 f(x, y) dy / \int_0^1 f(x, 0) dy = 0 \end{aligned}$$

综合以上结果有  $\text{Var}(\hat{g}) \leq \text{Cov}(\hat{g}, S(X, y)) / \text{Var}(S(X, y))$

为了得到 Bh 不等式, 以下取  $f(x, y)$  的导数所构成的向量, 并记

$$S(X, y) = (S_1(X, y), \dots, S_n(X, y)); x$$

其中  $S_i(x, y) = f_i(x, y) / f(x, 0)$ . 当  $A = 1$  时,  $S(x, 0)$  即为  
 $\hat{g} = \text{Cov}(\hat{g}, S(X, y)) / \text{Var}(S(X, y))$

,  $\hat{g}$  为  $g(x, y)$  的无偏估计. 并

$$\hat{g} = \text{Cov}(\hat{g}, S(X, y)) / \text{Var}(S(X, y))$$

$$\text{Var}(\hat{g}) \leq \text{Cov}(\hat{g}, S(X, y)) / \text{Var}(S(X, y))$$

172

## 第五章点估计的性质

得分函数. 引理 5.1.3 设  $f$  为

偏估计, 且满足以下 Bh 条件:

C-R 分布族,  $i$  为  $g(x, y)$  的无

i)  $f(x, y)$  存在, 且关于  $y$  连续; ii)  $E[S(X, y)] = 0, f = 1, \dots, n$ ;

iii)  $E[S(X, y)]$  可在积分号下关于  $y$  求导数,  $y = 1, \dots, n$  则有

i)  $E[S_i(X, y)] = 0, i = 1, \dots, n$ ;

ii)  $\text{Var}(S(X, y))$  存在, 并记为  $V(y) = \text{Cov}(S(X, y), S(X, y))$ ;

iii)  $\text{Cov}(S(X, y), S(X, y)) = E(g(x, y)) = (g(1), \dots, g(n))$

公丁 (9) | " .

证明 在例 5.1.5 中分别取  $i = 1, \dots, n$ , 即可. 麗

定理 5.1.4 ( Bhattacharya 不等式). 条件同引理 5.1.3, 且假定  $V(y)$  可逆, 则有

$\text{Var}(\hat{g}) \geq \text{PT}(\hat{g}) V(y) \text{PT}(\hat{g})^T \geq \text{Bh}(g)$ , (5. 1. 17) 且等式成立的充要条件为

$$g(x, y) - g(\hat{g}) = \text{DT}(e) V(y)^{-1} (g(x, y) - g(\hat{g})). \quad (5. 1. 18)$$

证明 由于  $f$  和  $g$  都是向量, 应用 Schwarz 不等式 (5. 1.3) 的等价形式 (5. 1.5) 比较方便. 这时可考虑  $\text{Var}(X - AK)$  多 0,



并以 $i(X)$ 取代 $X$ , 以取代 $r$ , 则由引理5.1.2可得  $A = \text{Cov}(X, g) \text{Var}(g)^{-1} = Z^T T(\eta)$

由  $\text{Var}(X - A\eta) = \text{Var}(X) - \text{PT}(\eta) V^{-1}(\eta)^T (X, \eta) \geq 0$  可得  $\text{Var}(X) - \text{Cov}(X, \eta) V^{-1}(\eta) \text{Cov}(\eta, X) \geq 0$

由引理5.1.3可知  $\text{Cov}(g, \eta) = \text{PT}(\eta) \text{Var}(\eta) = V$ , 代入上式可得  $\text{Var}(X) - D^T V^{-1} D \geq 0$ .

由此得到(5.1.17)式. 由引理5.1.2可知, (5.1.17)式等式成立的充要条件为  $X - A\eta = g(\eta) - D^T V^{-1} \xi(x, \eta) : a(\eta) \text{ (a.e.)}$ . 等式两边取期望可得  $a(\eta) = g(\eta)$ , 代入上式, 则等式成立的充要条件化为

### 5.1 C-R不等式

173

$g(x) - g(\eta) = D^T V^{-1} \xi(x, \eta)$ .

此即(5.1.18)式, 证毕. I

通常, 不等式(5.1.17)简称为Bh不等式; 简称为Bh下界. 推论1 对于  $A; = 1, Bh\}$  下界即为C-R下界CRLB.

因为  $\eta=1$  时,  $D\eta = g'(\eta)$ ,  $\text{Var}[y(X, \eta)] = 1$ .

推论2 若  $Ejg(\eta) = g(\eta) + bg(\eta)$  且记  $P = (V \perp x, D, = +6^{\wedge}(\eta), f = 1, \dots)$  则 Bh 不等式化为  $\text{Var}(X) \geq D^T V^{-1} D$ . 下面证明Bh不等式的主要性质, 即Bh下界随A递增, 因而有

$\text{CRLB} = Bh \wedge Bh^2 \dots \langle \text{Var}(X) \rangle$ .

定理5.1.5 在Bh不等式(5.1.17)的右端, 记  $B_k B \wedge 0) =$

$D^T(e) V^{\wedge}(\eta) D \wedge 0)$ , 则当时有  $(\eta)$  人的.

证明 把矩阵  $V_k$  和向量按维数  $Z$  和分块如下:

则有  $V_k = V M D, = Z) z$ , 逆公式

$t = U, V_2 J D_i = (D_2)^T$

在  $B_k = D^T V^{\wedge} D_k$  中, 利用分块求

其中  $\eta - z$  为单位矩阵,  $G = V_n \} V_1 29 H^{\wedge} V_2 2 - V_2 x V - ; V_n^{\wedge}$ , 因此有

$j \ 21 \ y \ 22 \ d \ 2$

$= +(d; g - W) z t, (g t d 1 - d 2)^{\wedge} b \setminus o)$ . I

推论 以上定理中,  $B_k(\eta) = B_t(0)$  的充要条件为  $V_2 1 \ V^{\wedge} D. - D, = 0$ .

例5.1.6 (续例5.1.4) 设  $4, \dots$ , 又独立同分布,  $X_x$  可以证明:  $T(J) = I_2 - n^{-1}$  是  $g(19) = 0_2$  的UMVUE, 其方差达不到 C-R下界, 但是能达到5.9.2下界.

证明 由第三章例3.2.7可知,  $T(X) = X^2 - n^{-1} * g(\eta)$  的 UMVUE  $X * \text{例5.1.4}$  可知, 其方差达不到C-R下界. 经直接计算可得 (可令  $Y = x \sim e$ )

$\text{Var}[T(X)] = \eta^{\wedge} + 4 - n$

今考虑怎么下界. 当  $k=2$  时, 由例5.1.4可知,  $SJ(\eta, 0) = n(x$

第五章点估计的性质 由于  $a/t T(x - 0) \sim 7V(0, 1)$ ,

$^{\circ} V \sim \setminus 0 \ 2n \ i / (2n^2) /$

因此由(5.1.17)式可知

174

$S^2(x, 0) = n^2(x - 0')^2 - n$ ; 而  $D \setminus = (20, 2)$ . 经直接计算可得

$/n^0$

$K=2$

$T.422 \ Bh_7 = Dy; \{D. = - + -$

$nn$  因此  $\text{Var}[r(X)] = Bh_2$ , 即  $7 \setminus X$  的方差达到仙2下界.

### 5.1.4 多参数C-R不等式

今考虑参数  $\theta$  为  $P$  维向量的情形. 设  $X -$

rp,  $\theta = (e^{\eta_1}, \dots, e^{\eta_p})^T$  为 C-R 分布族. 又设  $g(\eta) = (g_1(\eta), \dots, g_p(\eta))^T$  的无偏估计为  $i(\eta) = (i_1(\eta), \dots, i_p(\eta))^T$ , 其方差  $\text{Var} Jg(\eta)$  为  $n \times p$  阶矩阵.

C-R 不等式仍然由 score 函数出发, 这时  $S(\eta, \theta) = (S_1(\eta, \theta), \dots, S_p(\eta, \theta))^T$ , 其中  $K(\eta, \theta) = -\log f(\eta, \theta)$

$\int \frac{\partial}{\partial \eta_j} f(\eta, \theta) d\theta = 0$

$j = 1, \dots, p$

由第二章的结果可知

$E J S(\eta, \theta) = 0, \text{Var} J S(\eta, \theta) = E J S(\eta, \theta) S(\eta, \theta)^T = I(\theta) p \times p.$

引理 5.1.4 设  $\eta$  为 C-R 分布族,  $i(\eta)$  为  $g(\theta)$  的无偏估计, 则有

$\text{Cov} J i(\eta), S(\eta, \theta) = 0 E J i(\eta) S(\eta, \theta)$

$\theta = (\theta_1, \dots, \theta_p)^T$  (5.1.19)

其中

$\text{Cov}(g, S) = E J i(\eta, S) = \frac{\partial}{\partial \eta_j} g(\eta, S), i = 1, \dots, p, j = 1, \dots, p.$

证明 只需对每个分量加以证明即可. 由假设可知

$\int \frac{\partial}{\partial \eta_j} f(\eta, \theta) d\theta = 0$

$E J g(\eta, S) S(\eta, \theta) = 0$

、

机(4)  $p \times p$

$= \frac{\partial}{\partial \eta_j} g(\eta, S) S(\eta, \theta)$

$\eta$

$p$

$\Rightarrow \int \frac{\partial}{\partial \eta_j} f(\eta, \theta) d\theta = 0$  如(4)

/

、!

$= \int \frac{\partial}{\partial \eta_j} g(\eta, S) f(\eta, \theta) d\theta = 0$

,  $\theta \in \Theta$

## 5.2 广义 OR 型不等式

175

定理 5.1.6 设  $\{f(\eta, \theta) | \eta \in \Theta\}$  为 C-R 分布族,  $i(\eta)$  为  $g(\theta)$  的无偏估计, 则有  $\text{Var} J i(\eta) \geq C(6) Z^{-1}(6) G(6)$ , (5.1.20)

且等式成立的充要条件为

$i(\eta) = C(\theta) S(\eta, \theta)$ . (5.1.21)

证明 仍应用 Schwarz 不等式的等价形式 (5.1.5) 式:  $\text{Var}(X - A\eta) \geq 0$ , 并以  $i(\eta)$  取代  $X$ , 以  $S(\eta, \theta)$  取代  $\eta$ , 则由 (5.1.5) 式以及引理 5.1.4 有  $A = \text{Cov}(i, S) \text{Var}(S)^{-1}$

$= G(\theta) r'(\theta)$ ,

$\text{Var}(X - A\eta) = \text{Var} J g(\eta) - G(\theta) (\theta) S(\eta, \theta)$  由此可得

$\text{Var}(i) - \text{Cov}(i, S) r'(\theta) G(\theta) - G(\theta) \text{Cov}(S, i) + \frac{1}{n} \text{Var}(S) \geq 0$ . 由 (5.1.19) 式可得

$\text{Var}(g) - 2GZ^{-1}G + GZ^{-1}G \geq 0$ .

因而有

$\text{Var}(g) \geq G r'(\theta) G$

此即 (5.1.20) 式. 由引理 5.1.2, 该式等式成立的充要条件为

$X - AK = i(\eta) - G(\theta) r'(\theta) S(\eta, \theta) = 0$ . 等式两边取期望得  $a(\eta) = g(\theta)$ , 代入上式即可得到 (5.1.21) 式:

$g(\eta) - g(\theta) = c(\eta) s(\eta, \theta)$ ,  $c(\eta) = G(\theta) r'(\theta)$ . ■

推论 1 若  $i(\eta)$  为  $g(\theta)$  的无偏估计, 则有 推论 2 若  $E J H(\eta) = g(\eta) b(\eta)$ , 则有

VarJiW]渝 &=(&),

J J d0j

推论3 若 $X_j, \dots, X_n$ 独立同分布,  $X$ 的Fisher信息阵为 $i(\theta)$ ,

则有  $-\text{VarJg}(X)]5:n$

$=0(n^{-1})$ ,

对于多参数C-R不等式, 也有类似于定理5. 1.2和5. 1. 3 的结果, 从略, 可参见成平等(1985).

5. 2 广义C-R型不等式 本节主要讨论非C-R分布族情形下, 估计量的方差下界问题. 特

$\text{Vare}[\theta(X)] \sim i^{-1}(\theta) = O(n^{-1})$ . TV

176 第五章点估计的性质

别考虑没有共同支撑情形下有关估计量方差的不等式, 包括 $|\varphi(\theta, \theta_j) - M + r(l, l)f$ 等分布族. 记 $A_\theta = \{x: f(x, \theta) > 0\}$ , 一般应与 $\theta$ 有关,

即当 $\theta$ 变化时, 例如 $\{f(x, \theta) : \theta \in \Theta\}$ ,  $A_\theta = \{x : f(x, \theta) > 0\}$ 就是如此. 在以下讨论中, 假定 $\theta$ 可以与 $\theta$ 有关(即可以不是C-R分布族).

上一节已经指出, C-R型不等式有很多种, 在(5. 1. 16)式中, 不等式对任何的 $\theta$ 都成立. 上一节曾经取 $I f$  或 $\Delta =$

本节将取  $S = s(x, \theta) = f(x, \theta) / f(x, \theta)$ . 它与取

$S = f'(x, \theta) / f(x, \theta)$ 有某些类似之处(见第二章Fisher信息与Kullback信

息的定义与性质). 设 引理5. 2.1

, 时, 有

记 $U(x$

.  $g(X)$ 为 $\theta$ 的无偏估计. 对于  $\theta$ , 命  $U(x, \theta) = g(x) / f(x, \theta)$ , 并假定 $\text{Var}_\theta(S)$ 存在, 则当 $n \rightarrow \infty$  时

证明 类似于(5. 1. 16)式有

$J^2 \text{Var}(g) = \text{Var}(S^2)$ :

以下主要计算

$\text{Cov}(g, S) = E(gS) - E(g)E(S)$ . 由 $U(x, \theta)$ 的定义以及 $U(x, \theta) = g(x) / f(x, \theta)$ 有

$E(S^2) =$

$= \int S^2 f(x, \theta) dx = \int (g^2 / f) f(x, \theta) dx = \int g^2 dx$

(5. 2. 1)

(5.2.2) (5.2.3)

=

$+ L^{-1} \int \frac{d}{dx} (g^2) dx = \int \frac{d}{dx} (g^2) dx = \int \frac{d}{dx} (g^2) dx$

同理有

$E(S^2) = \int (g^2 / f) f(x, \theta) dx$

5.2 广义C-R型不等式

177

$\int g^2(x) / f(x, \theta) dx \geq \int g^2(x) / f(x, \theta) dx$

$\int g^2(x) / f(x, \theta) dx =$

以上结果代入(5.2.3)式可得  $\text{Cov}(g, S) = g(x) - g(x)$ .

以上结果代入(5.2.2)式, 即得(5.2. 1)式.

定理5.2.1 (Chapman, Robbins和Kiefer不等式). 设 $f(x, \theta)$ ,  $g$

$\theta \in \Theta$ ,  $g$ 为 $\theta$ 的无偏估计. 则有  $\text{Var}_\theta[g(X)]$

$[g(x) - g(x)]^2$

$f(x, \theta) - f(x, \theta)$

|

(5.2.5)

$f(x, e)$

证明 根据引理5.2.1, 不等式(5.2.1)对一切 $A \in \mathcal{A}$ 都成立, 因而(5.2.5)式成立.

不等式(5.2.5)简称为CRK不等式, 该式右端称为CRK下

界. 以上定理的意义在于(5.2.5)式可用于非C-R分布族, 特别是无共同支撑

的均匀分布等分布族. (5.2.5)式的主要困难在于计算上确界, 即 CRK 下界, 通常只求近似值, 即比CRK下界稍小一点的下界, 见后面

例子. 推论1 若  $X$  出C-R不等式, 且CRK下界多C-R下界.

,  $g(x) > 0$ ) 为C-R分布族, 则由CRK不等式可推

若  $f(x, e) \geq 0$  为C-R分布族, 则它有共同支撑, 因而  $f(x, e) \geq 0$

证明 有由于(5.2.5)式对任何 $e$ 都成立, 取 $e = 0 + M$ , 则有

上式中令 $A = \{x: f(x, e) > 0\}$ , 则有  $\text{Var} J_g(X) \geq \text{CRK 下界}$

178 第五章点估计的性质 推论2若 $T$ 为充分统计量 $T = T(X)$ 的函数,  $i(j) = i(T)$ ,

且有 $T \sim h(t, \theta)$ , 则有

$[g(4)]$

$h(T, \theta)!$

只需对(5.2.5)式应用因子分解定理即可得到上式. 例5.2.1 设 $X_1, \dots, X_n$ 独立同分

布,  $X_i \sim N(\mu, \sigma^2)$ ,  $r(i, i) = 1$ ,

由于

当 $n \rightarrow \infty$ 时有

$I_n(1) \leq 9I_n, I_n = [0, \infty)$ , 因此有

$\text{Var} J_h(X) \sim \frac{1}{n}$

$\text{Var}$ ,

2

无偏估计的CRK下界.

解 样本分布没有共同支撑:

•  $(x, f(x))$

求 $g(4) = 0$

).

(5.2.6)

$= e^{-J_a}$ ,

$c_0 j = \int_0^1 f(x) dx$

= 而由(5.2.4)式可得

$Z(-1) e^{-1} = \int_0^1 f(x) dx$ ,

CRK下界 $= \sup_{0 \leq t \leq 1} \{ \int_0^t f(x) dx - \int_0^t f(x) dx \}$

$= \int_0^1 f(x) dx$

$1 - U(1)$

$I \setminus \{x\} > 0 \mid j(x(1)) \neq 0 \mid e^{-1}$ .

$\text{Var}(S_n) = E(S_n^2) - (E S_n)^2 = E(S_n^2) - 1$ . 由于 $X(1) \sim N(0, 1)$ ,  $U(1) \sim N(0, 1)$ ,

$u \geq 0$ , 因此有

$E(S_n^2) = \int_0^1 \int_0^1 f(x) f(y) dx dy = \int_0^1 \int_0^1 f(x) f(y) dx dy = \int_0^1 \int_0^1 f(x) f(y) dx dy$ .

因此有 $\text{Var}(S_n) = \int_0^1 \int_0^1 f(x) f(y) dx dy - 1$ . 这些结果代入(5.2.6)式可得

$4 - e^{-1} \text{CRK下界} = \sup_{0 \leq t \leq 1} \{ \int_0^t f(x) dx - \int_0^t f(x) dx \}$

小为 $0.47$ . 特别,  $U(1)$ 的UMVUE为 $4 - 1/n$ , 这时有 $\text{Var}(U(1)) = 1/n^2$ , 其值明显的比CRK下界的近似值大. ■

例5.2.2 设 $X_1, \dots, X_n$ 独立同分布,  $X_i \sim R(0, 1)$ , 求 $g(1) = 1$  无偏估计的CRK下界.

解 样本分布也没有共同支撑: 八X, 的 $= (1/\sqrt{n}) / \sqrt{x(n)} \sqrt{xm}$

$> 0$  I, 么 $= (0, \infty]$ , 并且有  $I =$

, 因此有

4> I

### 5.3估计量的渐近性质

179

因此当时有

因此有

$E(S) = f$

$dy =$

$= \sup \{ (f(x) - e) / \sqrt{2} \cdot \sqrt{n} \}$

0I.

CRK 下界 $= \sup$

$(f(x) - e)^2$

(免  $-1$ 、4>

CRK 下界 $= \sup \{ (f(x) - e)^2 / (2n) \}$

$f(x, e)$

$= \sup \{ (f(x) - e)^2 / (2n) \}$

而由(5.2.4)式可得 $\text{Var}(S) = E(S^2) - (E(S))^2$

$= E(S^2) - f^2$ . 又由

于 $X(1)/n \sim BE(n, 1/n)$ , 因此有  $f(x) / \sqrt{n} \rightarrow 0$

(n)

4)  $11x(n) \sim 3$

$U_i I$

$\sim 0$

$X(n)$

这时不易求出准确最大值, 可取 $4 > -1$  估算其近似值  $n+2$

叫 $\pm C_1$ . -

$-H_7 \cdot \sqrt{2}$

$n+2) \cdot \sqrt{n+2} \sqrt{2}$

$(f(x) - e)^2 / (2n) \rightarrow 0$

特别, 0的UMVUE为 $3 = 1/8$ , 估计量的方差为 $\text{Var}(\hat{\theta})$  其值明显地比CRK下界的近似值大.

### 5.3估计量的渐近性质

1-

切2  $= 0$

$(e^2 - 1) / (2) = 1.5 \sqrt{2}$ .

估计量的渐近性质或大样本性质, 就是考虑当样本容量 $n$  充分大 时, 估计量与被估计量能否按某种意义充分接近. 在数学上, 就是考虑 样本容量 $\rightarrow \infty$ 时估计量的收敛性问题. 设样本为 $X_1, \dots, X_n$ , 并记

$X = (X_1, \dots, X_n)$  (通常假定 $X_1, \dots, X_n$ 独立同分

0n

$e_1 \sim n(n+2)$

9

### 180 第五章点估计的性质

布). 假设 $g(X)$  的估计记为 $g(X) = g(X_1, \dots, X_n) = g_n(x)$ , 类似地,  $\theta$  的估计记为 $\hat{\theta}(1)$ 或 $\hat{\theta}(n)$ . 本节主要考虑相合性, 即 $n \rightarrow \infty$ 时 是否成立, 其中 $\hat{\theta}$ 表示随机变量 $g_n(X)$ 在某种概

率意义下的收敛；渐近正态性，就是考虑随机变量  $Z_n = \sqrt{n}(g_n(x) - g(x))$  的渐近分布是否为正态分布？以下第1小节首先复习介绍数理统计中常用的随机变量序列的收敛性及其有关的性质，进一步的内容可参见李贤平(1997)，陈希孺(1981)，严士健，刘秀芳(1994)等文献；第2小节介绍相合性和渐近正态性的定义与性质；最后两小节则讨论矩估计和极大似然估计的相合性和渐近正态性。

### 5.3.1 随机变量序列的收敛性

记  $I_n$  为随机变量； $b, c$  为常数； $F_n(x)$ ， $f \sim F$  (幻为分布函数。当  $n \rightarrow +\infty$  时，随机变量序列的收敛性通常有以下几种：

(1) 依概率收敛，记作  $P$  (或)：若  $P(|f_n - f| > \varepsilon) \rightarrow 0$  ( $n \rightarrow +\infty$ )， $\forall \varepsilon > 0$ 。

(2)  $r$  阶矩收敛，记作  $M_r$  (或)；特别当  $r=2$  时，记为  $V_2$  (1)4

(3) 几乎处处收敛， $H$  (a.e.) 或 (a.s.)：若  $P\{x: \lim_{n \rightarrow \infty} f_n(x) = f(x)\} = 1$

当  $X \in A$  时有  $(f_n(x) - f(x)) \rightarrow 0$  ( $n \rightarrow +\infty$ )；

，这也可表示为  $P\{|f_n - f| > \varepsilon\} \rightarrow 0$  ( $n \rightarrow +\infty$ )；

，这也可表示为  $P\{|f_n - f| > \varepsilon\} \rightarrow 0$  ( $n \rightarrow +\infty$ )；

$\rightarrow f(x) = 1$ ，即对  $P(A) = 1$

(4) 依分布收敛， $H$  会或若在  $f$  的连续点处有  $F_n(x) \rightarrow F(x)$  ( $n \rightarrow +\infty$ )。

当  $n \rightarrow +\infty$  时，随机变量序列的收敛性有以下常用性质：

(1) (a.e.) 另外也有会。二(9)上特别地，由  $I_n \rightarrow I$   $r=0$  可得  $J^0$ ；当  $r=2$  时，由

$\text{Var}(f_n) \rightarrow 0$  可得  $E f_n^2 \rightarrow 0$ 。更常用的形式为：若  $\text{Var}(f_n) \rightarrow 0$ ，或  $f_n^2 \rightarrow 0$ ，

则有  $H_a$ ，更进一步：若  $H \text{Var}(f_n) \rightarrow 0$ ，则有  $6$  上  $0$ 。

(2) 若或 (a.e.)，函数在  $\perp p(c)$  (或 (a.e.))。

$\% = c$  处连续，则有

(3) 随机变量序列依概率收敛到常数  $c$  的充要条件为依分布收敛到  $\delta_c$  特别， $H_0$  的充要条件为  $\wedge L^0$ 。因为若  $H_0$ ，则  $F_n(x) \rightarrow \wedge^r I x^c$ ；即  $F_n(x) \rightarrow 0$  ( $x < c$ ) 且  $F_n(x) \rightarrow 1$  ( $x > c$ )。因此有

### 5.3 估计量的渐近性质

181

$P\{|f_n - c| > \varepsilon\} \rightarrow 0$  ( $n \rightarrow +\infty$ )；即  $H_0$ 。

(4)  $\wedge n$  (a.e.) 的充分条件 (Borel - Cantelli 引理)。若对  $\forall \varepsilon > 0$ ,

$00$

$\forall \varepsilon > 0$ ， $P\{|f_n - c| > \varepsilon\} < \varepsilon$  (见李贤平, 1997)。特别， $n=1$

$00$   $00II$

若  $Z \text{Var}(f_n)$  收敛，则由切比雪夫不等式知  $P\{|f_n - E f_n| > \varepsilon\} \rightarrow 0$  ( $n \rightarrow +\infty$ )

收敛，因而  $E f_n \rightarrow 0$  (a.e.)；若再有  $f_n \rightarrow a$  或  $f_n \rightarrow a$  (a.e.)。

以下定理对于推导随机序列的渐近分布十分有用。

$m$  么上  $c$ ，则有卷， $d \rightarrow C_L f(c/0)$ 。

特别，若  $\% \rightarrow 0$ ，则  $ma$  去  $o$  律)；若  $\forall n$  (去  $1$  律)。  $d$

证明 我们只证明第一式，即当  $n \rightarrow +\infty$  时有 证明类似，从略。今假定  $\% \sim \text{灰}(\text{幻})$ ， $f(\text{幻})$ 。

已  $\wedge^+ c - F(x-c)$ ， $F(X) - \text{厂}(\%)$ ，且有  $P(\wedge n - c \wedge^+ f T) \rightarrow 0$ ， $\forall \varepsilon > 0$ ；要证  $P$

$+ \wedge n^+ x \rightarrow F(\% \sim c)$ ，其中  $\% - c$  为  $F(\text{幻})$  的连续点。设  $5_{,,} \rightarrow 0$  ( $6_{,,} > 0$ )，对充分大的

$n$  有  $H_n - c \wedge \wedge 5_n$ 。记事件  $\wedge = \{r_j n: c - s < y_n < c + f\}$   $9 A = \wedge r_j n - \wedge n - c \wedge$

$h$  易见  $P(A) \geq 1 - 5_{,,}$ ， $P(A) \wedge 5_n 1$  而在  $\forall n \in A$  上有

定理 5.3.1

$f + 7/n + C$ ， $h$

$+$  时，若

(Slutsky)当

$+c \rightarrow V$ .  $\Rightarrow +c+e$ .  $M_n(f, +c- I D M_n(+7)n^x) I D I A_n(+c+$

因此  $p \wedge n + y n^x \setminus = P \setminus A Q (\wedge n + V n^x) \setminus + P M_n(f n + 77 n^x) \setminus$

(5.3. 1)

), (5. 3. 2)

由(5.3. 1)式有

矣尸  $M_n(f, +c-6: ^x) \setminus + P(A)$  (5.3.3)  $\setminus P \setminus f n + c - e \setminus x ( + 8 n = F n (x - c + e) + 8 n;$

另一方面

$p \in \text{尸} x f \setminus P M_n(f n + " \text{尸} \% ) \setminus \text{多} \text{尸} + c + e ^x \setminus$

$= P u, I - P M_n(\wedge + \epsilon + \wedge ^x))$  (5.3.4)  $\setminus F n (x - c - e) - P(A) \wedge F n (x - c - - 6 n,$

u

, 则么仏上^

$n K + c$ , 其他

182 第五章点估计的性质 综合(5. 3. 3)式和(5. 3. 4)式可得尸 \  $f n + y n^x \}$

$\sim ^ F (x - c)$ , 即  $+ r j n$

九  $f + c$ .

推论 当  $+x$ 时, 若

m

I  $a_n, b_n$  则  $a_n ^ n + b_n$

么  $a f + 6$ .

注以上定理对于随机向量序列也成立, 因为只需对每个分量应用

该定理即可.

例5. 3.1 z分布的渐近正态性. 设  $X, \dots$ , 独立同分布,  $X, \sim$

$/V(0, a^2)$ , 则当  $\leftarrow + \infty$  时有  $z, \setminus /V(0, 1)$ , 其中

证明 因为 例5.3.2), 因此

t

$= \frac{\quad}{n}$  拉  $(X, -J)2/(n - 1)$

n

$\sim /V(0, 1)$ ,  $S^2 = n \cdot \sum (X_i - X)^2 / n^2$  (可参见  $i = I$

所以由Slutsky定理“去1律”知, 当  $+00$ 时,

另外, 若  $X, \dots, J$  独立同分布,  $E(JJ) = 0$ ,  $\text{Var}(X_1) =$ , 则由

中心极限定理有  $V^{\wedge} / a \_ U V(0, 1)$ , 因此以上结果仍然成立. | Slutsky定理的用处十分广

泛, 特别是“去0律”和“去1律”, 后

面经常用到. 以下定理也是常用的形式.

定理5.3.2 当  $+$  时, 设数列  $\setminus - 00$ , 随机变量  $V_n = a_n ^ n -$

$b_n$  函数/(幻在  $x = b$  处存在二阶连续导数, 则有

(1)

(2) 若  $f f(b) \neq 0$ , 则  $a_n = a_n [f(H f(b, \sim \_ L ^ f'(b) Z' (3) 若 /'(6) = 0, f(6) \neq 0$ , 则

$= a J / ($

士斤(4) Z.

证明 主要应用Slutsky定理. 当  $n \rightarrow +\infty$ 时:

(1)  $-6 = a; * [a n (-6)] - ^ 0 - Z = 0$ , 故  $a_n - b$  么0,  $J^b$ .

(2) 心  $= a$ .  $/'(D (么-6)$ , 其中  $L$  在6与么之间, 即  $\setminus L - b \setminus \setminus ^ n - b \setminus$ , 因此由  $H_b$  知么上

6,  $/'(I)$  上  $/'(6)$ , 由

Slutsky定理知  $a_n m h) Z$ .

### 5.3估计量的渐近性质

183

(3)对氏进行二阶Taylor展开可得

氏  $W_n(\hat{\theta}) = W(\theta) + W'(\theta)(\hat{\theta} - \theta) + \frac{1}{2}W''(\theta)(\hat{\theta} - \theta)^2 + o_p(n^{-1/2})$

由专:  $H = W''(\theta)$ , 知  $W''(\theta) = 2\lambda^2 f''(\theta)$

I

最后简要介绍一下随机变量序列的随机阶, 这也是统计学的大样本理论中常用的工具; 其定义和性质与实数序列的阶十分类似.

若当  $n \rightarrow \infty$  时,  $\hat{\theta}_n - \theta = o_p(n^{-1/2})$ , 则记  $\hat{\theta}_n = O_p(n^{-1/2})$ . 特别地, 若  $\hat{\theta}_n - \theta = O_p(n^{-1/2})$ , 则记为  $\hat{\theta}_n = O_p(n^{-1/2})$ . 若  $\hat{\theta}_n - \theta = O_p(n^{-1/2})$ , 则记为  $\hat{\theta}_n = O_p(n^{-1/2})$ .

若对  $\forall \epsilon > 0$ , 存在  $A$ ; 和  $I$  使  $n \geq n_0$  时有  $P(|\hat{\theta}_n - \theta| > \epsilon) \leq A/n$

H

( $K \in \mathbb{N}$  且  $n \geq K$ ). 特别地, 若  $C_n = 1$ , 则记为  $O_p(1)$ ; 若

$C_n = n^{-1/2}$ , 则记为

$O_p(n^{-1/2})$ .

随机变量序列的随机阶有以下基本性质:

(1) 具有与非随机阶类似的性质, 诸如:  $O_p(c_n) = c_n O_p(1)$ ,  $O_p(1) + O_p(1) = O_p(1)$ ;  $O_p(1)O_p(1) = O_p(1)$ ,  $O_p(1) + O_p(1) = O_p(1)$ ;

n

(3) 当  $n \rightarrow \infty$  时, 若  $\hat{\theta}_n = O_p(1)$ .

证明 设  $\hat{\theta}_n \sim \theta$  (幻,  $\downarrow F(x)$ ), 对  $h > 0$ , 必存在  $\hat{\theta}_n$  使

$P(|\hat{\theta}_n - \theta| > h) \leq F(h) - F(0) = \epsilon$ . A•

而  $P(|\hat{\theta}_n - \theta| > h) = F(h) - F(0) = \epsilon$ , 因此必存  $n_0$  或使  $n > n_0$  时有  $P(|\hat{\theta}_n - \theta| > h) \leq \epsilon$ .

#### 5.3.2 估计量的相合性和渐近正态性

设  $X \sim /(\cdot, \cdot, \cdot)$ , 什  $\theta$ ,  $g_n(X) = g(X_1, \dots, X_n)$  为  $g(\theta)$  的估计.

定义5.3.1 当  $n \rightarrow \infty$  时, 若对一切  $\epsilon > 0$  有  $P(|g_n(X) - g(\theta)| > \epsilon) \rightarrow 0$ , 则称

$g_n(X)$  为  $g(\theta)$  的相合 (consistent) 估计 (或弱相合估计). 若  $g_n(X) \rightarrow g(\theta)$  a. e., 则称  $g_n(X)$  为  $g(\theta)$  的强相合估计.

由定义可知, 若以  $g_n(X)$  为相合估计, 则当  $n$  充分大时, 它与  $g(\theta)$  被估计的  $g(\theta)$  可充分“接近”. 研究相合性的主要工具为随机变量序

(2)  $E[g_n(X)] = o(1)$ ,  $E[g_n(X)] = E[g(\theta)]$  和 (2) 的证明从略);

(1)  $E[g_n(X)] = o(n^{-1/2})$

#### 184 第五章点估计的性质

列收敛性的基本性质以及大数定律. 由随机变量序列收敛性的性质 (1) 和 (2) 可得

引理5.3.1 当  $n \rightarrow \infty$  时, 若  $E[g_n(X)] \rightarrow g(\theta)$ , 则  $g_n(X)$  为

$g(\theta)$  的相合估计; 若  $\text{Var}[g_n(X)] \rightarrow 0$ , 并且  $E[g_n(X)] \rightarrow g(\theta)$  或者  $0$

$E[g_n(X)] = g(\theta)$ , 则  $g_n(X)$  为  $g(\theta)$  的相合估计. 若  $\text{Var}[g_n(X)] \rightarrow 0$

收敛, 并且  $E[g_n(X)] \rightarrow g(\theta)$  或者  $E[g_n(X)] = g(\theta)$ , 则  $g_n(X)$  为  $g(\theta)$  的强相合估计. I

引理5.3.2 若  $g_n(X)$  为  $g(\theta)$  的相合 (或强相合) 估计,  $p(y)$  在  $y = g(\theta)$  处连续, 则

$p(g_n(X))$  为  $p(g(\theta))$  的相合 (或强相合) 估计.

例5.3.2 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim /(\mu, \sigma^2)$ ,  $E(X_i) = \mu$ ,  $\text{Var}(X_i) = \sigma^2$

则有;  $\bar{X}$  为  $\mu$  的强相合估计

为  $\mu^2$  的强相合估计.

$n(\bar{X} - \mu) \xrightarrow{d} N(0, \sigma^2)$

2



证明 由独立同分布情形下的强大数定律可知, 当  $n \rightarrow +\infty$  时,  $\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{a.s.} EX$ .  $S_n^2 = \sum_{i=1}^n (X_i - EX)^2$ . 例5.3.3 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $EX = \mu$ ,  $Var(X) = \sigma^2 < \infty$ , 则  $\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{a.s.} EX^2 = \sigma^2 + \mu^2$ . |

例5.3.3 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $EX = \mu$ ,  $Var(X) = \sigma^2 < \infty$ , 则  $\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{a.s.} EX^2 = \sigma^2 + \mu^2$ . |

例5.3.3 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $EX = \mu$ ,  $Var(X) = \sigma^2 < \infty$ , 则  $\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{a.s.} EX^2 = \sigma^2 + \mu^2$ . |

例5.3.3 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $EX = \mu$ ,  $Var(X) = \sigma^2 < \infty$ , 则  $\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{a.s.} EX^2 = \sigma^2 + \mu^2$ . |

例5.3.4 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $EX = \mu$ ,  $Var(X) = \sigma^2 < \infty$ , 则  $\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{a.s.} EX^2 = \sigma^2 + \mu^2$ . |

例5.3.4 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $EX = \mu$ ,  $Var(X) = \sigma^2 < \infty$ , 则  $\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{a.s.} EX^2 = \sigma^2 + \mu^2$ . |

例5.3.4 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $EX = \mu$ ,  $Var(X) = \sigma^2 < \infty$ , 则  $\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{a.s.} EX^2 = \sigma^2 + \mu^2$ . |

例5.3.4 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $EX = \mu$ ,  $Var(X) = \sigma^2 < \infty$ , 则  $\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{a.s.} EX^2 = \sigma^2 + \mu^2$ . |

### 5.3 估计量的渐近性质

证明记  $a_n = \frac{1}{n} \sum_{i=1}^n X_i$ , 则由假设可知

$a_n \xrightarrow{P} \mu$

185

其中  $a_n$ ,  $b_n$ , 分别表示上式右端第一项和第二项. 由假设可知  $E(a_n) = \mu$ ;

$E(b_n) = 0$ . 当  $n \rightarrow +\infty$  而  $m$  固定时,  $Var(a_n) = \frac{\sigma^2}{n} \rightarrow 0$ ;  $Var(b_n) = \frac{\sigma^2}{n} \rightarrow 0$ ;

因而由引理5.3.1可知  $a_n = a_n + b_n \xrightarrow{P} \mu$ , 即  $a_n$  上所以

子为从的相合估计. 当  $n \rightarrow +\infty$  而  $n$  固定时, 由于  $E(b_n) = 0$ ,

$Var(b_n) = \frac{\sigma^2}{n} \rightarrow 0$ ; 因而仍然有  $b_n \xrightarrow{P} 0$ . 但是由于  $n$  固定, 随机变

量  $b_n = \frac{1}{n} \sum_{i=1}^n X_i$  与  $m$  无关, 有确定的分布 (注意  $> 0$ ), 所以由 Slutsky

定理去律可知, 当  $n \rightarrow +\infty$  而  $n$  固定时有  $a_n = a_n + b_n \xrightarrow{P} \mu$ , 因而  $L$  不可能依概率收敛到  $0$ , 所以不是  $\mu$  的相合估计. |

例5.3.5 设  $X_1, \dots, X_n$  独立同分布,  $EX = \mu$ , 证明  $\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{P} EX^2$

为  $\mu$  的相合、强相合估计;  $\frac{1}{n} \sum_{i=1}^n X_i^2$  为  $\mu^2$  的强相合估计;  $\frac{1}{n} \sum_{i=1}^n X_i$  为  $\mu$

$\mu$  的强相合估计, 其中  $S = \sum_{i=1}^n X_i$ . |

率

证明 可根据分布  $X(1) \sim \mu + r(j, i)$  (参见第一章), 直接计算概

$I_n^{(i)} = \frac{1}{n} \sum_{j=1}^n X_j^i$ . 由于  $X(1) \sim (\frac{1}{r} e^{-r(y-\mu)}) / (y^i / z)$ , 因此  $P(I_n^{(1)} - \mu) = [ -e^{-r(y-\mu)} ]$

$= -e^{-r(y-\mu)} \rightarrow 0$  ( $n \rightarrow +\infty$ ).

1n-

因而  $I_n^{(1)} \xrightarrow{P} \mu$ . 另外由上式可知  $2P(I_n^{(1)} - \mu) = 2e^{-r(y-\mu)}$  收  $n=1$  收敛,

由 Borel - Cantelli 引理知  $X_n \xrightarrow{a.s.} \mu$ . 又由大数定律可得

$I_n^{(1)} \xrightarrow{a.s.} \mu$

$\sim \mu$  尤(1)  $\sim EX$   $\sim a.e.$

$\sim a - j_l = cr(a.e.)$ .

由引理5.3.1 知  $a - l^a l(a.e.)$ , 即  $S_n^A = a \sim (a.e.)$ . |

例5.3.6 设  $X_1, \dots, X_n$  独立同分布,  $X \sim P(\lambda)$ , 求  $A$  和  $P(X = 0) = e^{-\lambda}$  的相合估计.

解 由大数定律知  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$  为  $\lambda$  的强相合估计. 对于  $P(X = 0) = e^{-\lambda}$ , 根据引理5.3.2, 显然有  $e^{-\bar{X}} \xrightarrow{a.s.} e^{-\lambda}$ . 另外, 由第三章

的例题可知,  $e_A$  的一致最小方差无偏估计为  $i_{\alpha}(J) = p - \alpha$ , 其中

$T = i_X$ . 由直接计算可知  $\text{Var}[S_{\alpha}(\cdot; V)] = e_{2A}(e + \alpha - 1) \rightarrow 0$ ,  $i = 1$

$E[g_{\alpha}(X)] = e - \alpha$  因此  $g_{\alpha}(X) - E[g_{\alpha}(X)] = e - \alpha$  即  $i_{\alpha}(X)$  为  $e_A$  的

相合估计.  $I$  相合性是对一个估计量很基本的要求, 这表明, 当  $\alpha$  充分大时, 估

计量与被估计量能够充分“接近”. 但是相合性并未涉及估计量的精度, 诸如均方误差, 通常应该要求估计量的方差尽量小. 以下渐近正态性进一步指出了估计量的渐近分布和渐近方差. 很显然, 渐近方差越小, 估计量越好.

定义 5.3.2 渐近正态性. 若存在  $r(0) > 0$  使

, ...,  $\infty$  —  $g(0) \neq 0$   $Z \sim N(0, p(\alpha))$ , (5.3.5) 则称  $g_{\alpha}(X)$  为  $g$  的渐近正态的, 亦称  $g_{\alpha}(X)$  为  $g$  的相合渐近正态 (consistent asymptotic normal, 简称为 CAN) 估计. |

, 由 (5.3.5) 式以及定理 5.3.2 (取  $\alpha = 1$ )  $g_{\alpha}(X) = g(\alpha)$  可

因此若  $g_{\alpha}(X)$  为  $g$  的渐近正态估计, 则必为 (5.3.5) 式中, 以  $g(\alpha)$  为的渐近方差, 因此, 若  $g_{\alpha}(X)$  为 CAN

知

知

$g_{\alpha}(0)$  的相合估计, 可见称  $g_{\alpha}(X)$  为相合渐近正态 (CAN) 估计是合适的.

估计, 其方差的阶必为  $n^{-1}$ . 另外由定义可得

$p_{\alpha}(\alpha) = E[g_{\alpha}(X) - g(\alpha)]^2$  其中  $\alpha$  为标准正态的分布函数. 同时 (5.3.5) 式亦可表示为  $g(X_{\alpha}, \dots, X_n) = g(\alpha) + \sum_{i=1}^n Z_i / n$

其中  $Z_i \sim N(0, F(\alpha))$ . 由该式及 Slutsky 定理, 显然有  $g_{\alpha}(X) \rightarrow g(0)$ .

渐近正态性与中心极限定理有密切关系, 由以下例子可知为什么在定义 (5.3.5) 式中要乘以  $n$ .

例 5.3.7 设  $1, \dots, X_n$  独立同分布, 并设  $E(X_1) = A$ ,  $\text{Var}(X_1) = \sigma^2$

$(\sigma^2, \text{Var}(X_1)) = (\sigma^2, S^2)$

$i = 1$  (1) 又为  $M$  的 CAN, 且有

$(J_1 - J_2)^2$

—

(2)  $S^2$  为  $\sigma^2$  的 CAN, 且有  $A(S^2 - \sigma^2) \xrightarrow{d} N(0, t^2)$ .

$X \rightarrow H$   $s$

. 则有

$\rightarrow 0$  (%).

(5.3.6)

### 5.3 估计量的渐近性质

187

证明 (1) 由中心极限定理有

$\rightarrow E(X)$

,  $\rightarrow -N(0, J)$ .

上式可化为  $a^2$ , 即又为  $\sigma^2$  的 CAN. 另外, 由于  $S^2 = \frac{1}{n} \sum_{i=1}^n X_i^2$

$n \rightarrow \infty$   $X$  (无)  $^2$  为的相合估计, 因此  $S^2/a^2 \rightarrow 1$ , 所以有

由 Slutsky 定理“去 1 律”知  $y/n$

$T Y$

$(0, 1)$ .

(2) 由于  $S^2$  的分布具有平移不变性, 其分布与  $\mu$  无关, 可在  $\mu = 0$  处推导其渐近分布, 以简化计算. 由  $S^2$  的定义,  $S^2 - a^2$  可表示为

$\frac{1}{n} \sum_{i=1}^n (S^2 - a^2)$

由于在 $\theta=0$ 处有 $E(X) = a-2$ ，又由假设可知 $\text{Var}(\hat{\theta}) = T^{-2}$ ，所以由中心极限定理可得  
由于 $X$ 为 $M=0$ 的CAN，因此有

$$\hat{X}^2 = -(a/\hat{X})^2$$

综合以上结果，由Slutsky定理“去0律”可知 $\sqrt{n}(\hat{X}^2 - (a/\hat{X})^2) \xrightarrow{D} N(0, 7V(0, T^{-2}))$ 。

引理5.3.3 设 $g(\theta)$ 为 $\theta$ 的相合渐近正态估计，函数 $p(y)$ 在

$y=g(\theta)$ 处可导，且 $W(g(\theta)) \neq 0$ ，则 $\hat{g}_n(X)$ 为 $g(\theta)$ 的相合渐近正态估计，且有  
 $T_n = \sqrt{n}(p(\hat{g}_n(X)) - p(g(\theta))) \xrightarrow{D} N(0, W(g(\theta)))$ 。(5.3.7)

证明由假设可知， $Z_n = \sqrt{n}(\hat{g}_n(X) - g(\theta)) \xrightarrow{D} N(0, V(\theta))$ 。在定理5.3.2的 $a_n$ 中，取 $a_n =$   
在 $\hat{g}_n(X)$ ， $b = g(\theta)$ 即得

$$\frac{a}{a} \rightarrow 0.$$

188

第五章点估计的性质

$T_n = \sqrt{n}(p(\hat{g}_n(X)) - p(g(\theta))) \xrightarrow{D} N(0, W(g(\theta)))$ 。由此即得(5.3.7)式。

推论 以上结果对向量参数亦成立；即若 $g(\theta) = (g_1(\theta), \dots, g_k(\theta))^T$ ， $\hat{g}_n(X) = (g_1(\hat{X}), \dots, g_k(\hat{X}))^T$ ，且有

$$\sqrt{n}(\hat{g}_n(X) - g(\theta)) \xrightarrow{D} N(0, V(\theta))$$

则  
 $\sqrt{n}(p(\hat{g}_n(X)) - p(g(\theta))) \xrightarrow{D} N(0, W(g(\theta)))$ ， $y=(h_1, \dots, h_n)^T$ ， $p(y)$ 在 $y=g(\theta)$ 处可导，则有

$$(5.3.8)$$

其中 $F(\theta) = d p(y)/dy|_{y=g(\theta)}$ 。特别，若 $I$ 为向量，且有 $\sqrt{n}(\hat{X} - a) \xrightarrow{D} N(0, I)$ ，则有

$\sqrt{n}(\hat{X}^2 - a^2) \xrightarrow{D} N(0, 2aI)$ 。把以上引理用于独立同分布的样本（见例5.3.7），我们即可得到

与 $a^2$ 的可导函数的CAN。例如，对 $\mu(M)=M^2$ ，因为 $\mu'(m)=2m$ ，则由引理5.3.3可得  
 $\sqrt{n}(\hat{X}^2 - a^2) \xrightarrow{D} N(0, 4a^2)$ 。(5.3.9) 因此 $\hat{X}^2$ 为 $M^2$ 的CAN。

例5.3.8 设 $X_1, \dots, X_n$ 独立同分布， $X_i \sim \rho(A)$ 。求 $e^{-A}$ 的CAN。解由中心极限定理可得  
 $\sqrt{n}(\bar{X} - A) \xrightarrow{D} N(0, A)$ 。取 $p(A) = e^{-A}$ ，因为 $p'(A) = -e^{-A}$ ，则由引理5.3.3可得  
 $\sqrt{n}(e^{-\bar{X}} - e^{-A}) \xrightarrow{D} N(0, A e^{-2A})$ 。因此 $e^{-\bar{X}}$ 为 $e^{-A}$ 的CAN。再考虑6.1的一致最小方差无偏估计 $\hat{X}$ ，  
 $\hat{X} = (1 - \hat{\mu})/T = (1 - \hat{\mu})/\sqrt{n}$ ，由于

$\hat{X} - e^{-A} = \sqrt{n}(\hat{X} - e^{-A}) = \sqrt{n}(\bar{X} - A) + \sqrt{n}(\bar{X} - A)$ ，由数学分析的公式可得，当 $n \rightarrow \infty$ 时有 $\sqrt{n}(\hat{X} - e^{-A}) \xrightarrow{D} N(0, A e^{-2A})$ 。所以

上式第2项 $\xrightarrow{D} 0$ ，因此由Slutsky定理“去0律”知  
 $\sqrt{n}(\hat{X} - e^{-A}) \xrightarrow{D} N(0, A e^{-2A})$ 。(5.3.10)

所以 $U = (1 - \hat{\mu})/\sqrt{n}$ 也是 $e^{-A}$ 的CAN。

以下介绍最优渐近正态估计，简称BAN估计。这时其渐近方差应达到C-R下界CRLB。我们知道，(9)的任一估计量 $g_n(X)$ 的方差应满足

5.3估计量的渐近性质

189

$\text{Var}[g_n(X)] \geq \text{CRLB} = G(\theta)^{-1} G'(\theta) G(\theta)$ ，其中 $d g(e)/d \theta$ 上式等价于  
 $\text{Var}[\sqrt{n}(\hat{g}_n(X) - g(\theta))] \xrightarrow{D} N(0, G(\theta))$

今设 $\lim_{n \rightarrow \infty} U(0)/n$ 存在，并记为 $I(\theta)$ （对于独立同分布的样本， $n \rightarrow \infty$

4没）就是单个样本的Fisher信息）。易见，若 $\lim_{n \rightarrow \infty} \text{Var}(X) = \infty$

$G(\theta) \rightarrow G(\theta_0)$  则  $\ln(x)$  的渐近方差达到  $c-r$  下界, 是最理想的情况, 可称为 BAN 估计.

定义 5.3.3 设  $X = \{X_i\}$ ,  $\theta \in \Theta$  其 Fisher 信息阵  $I(\theta)$  满足

$\lim_{n \rightarrow \infty} (I(\theta)/n) = I(\theta)$ , 若  $g(\theta)$  的估计  $(X)$  满足  $n \rightarrow \infty$  时

$(\bar{X}, \dots, X_n) \rightarrow g(\theta) + \sqrt{I(\theta)^{-1}} \cdot N(0, I(\theta))$ , 则称  $\ln(\hat{\theta})$  为  $g(\theta)$  的最优渐近正态 (best asymptotic normal) 估计, 简称

BAN 估计. 特别, 若  $\ln(X) \rightarrow g(\theta)$ , 则称  $\ln(X)$  为  $\theta$  的 BAN 估计. 压

注 若  $X_1, \dots, X_n$  独立同分布, 则  $I(\theta) = nI_1(\theta)$ ,  $I_1(\theta)/n = I_1(\theta)$  为  $X_1$  的 Fisher 信息阵. 另外, 若  $g(\theta)$  与  $\theta$  都是单参数, 则以上定义可表示为

$\ln(\hat{\theta}) - g(\theta) \rightarrow N(0, I(\theta)^{-1})$ .

例 5.3.9 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\mu, \sigma^2)$ , 则  $\mu$  为  $M_2$  的 BAN.

此解 这时  $g(\mu) = \mu$ ,  $g'(\mu) = 1$ ;  $I(\mu) = n/\sigma^2$ ,  $I_1(\mu) = 1/\sigma^2$ , 因

另一方面, 由 (5.3.9) 式可知  $\ln(X) \rightarrow N(\mu, \sigma^2)$ , 因此  $\mu$  为  $M_2$  的 BAN 估计. ■ 例 5.3.10 (续例 5.3.8) 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim P(\lambda)$ , 则

$\ln(\hat{\lambda})$  为  $\lambda$  的 BAN 估计.

解 由于  $g(\lambda) = \ln \lambda$ ,  $g'(\lambda) = 1/\lambda$ ;  $I(\lambda) = n/\lambda^2$ ,  $I_1(\lambda) = 1/\lambda^2$ , 因此

$[g'(\lambda)]^2 I_1(\lambda) = 1/\lambda$ . 由 (5.3.10) 式可知在  $(\ln(X) - \ln \lambda)$  中

$Z \sim N(0, 1/\lambda)$ , 因此  $(1 - \ln \lambda)$  为估计. ■

例 5.3.11 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim B(1, p)$ ,  $0 < p < 1$ , 证明  $\ln(1-X)$  为  $a^2 = \text{Var}(X)$  的 BAN 估计 ( $p = 1/2$ ).

证 由中心极限定理可得

$\ln(X) - \ln p \sim N(0, 1/p^2)$ ,

$a^2 = 1/p^2$

### 5.3估计量的渐近性质

189

$\text{Var}[g_n(X)] \geq \text{CRLB} = G(\theta)' Z^{-1}(\theta) G(\theta)$ , 其中  $d g(\theta)/d \theta$  上式等价于  $\text{Var}[g_n(\theta) \sim \sqrt{n} G(\theta) r(\theta) G(\theta)'] = G(\theta)$

今设  $\lim U(\theta)/n$  存在, 并记为  $I(\theta)$  (对于独立同分布的样本,  $n \rightarrow +\infty$  没) 就是单个样本的Fisher信息). 易见, 若  $\lim \text{Var}(X) = \frac{1}{n I(\theta)}$

$G(\theta) r(\theta) G(\theta)'$  则  $I(\theta)$  的渐近方差达到C-R下界, 是最理想的情况, 可称为BAN估计.

定义5.3.3 设  $X = \{X_i\}$ ,  $\theta \in \Theta$  其Fisher信息阵  $I(\theta)$  满足

$\lim (I(\theta)/n) = I(\theta)$ , 若  $g(\theta)$  的估计  $(X)$  满足  $n \rightarrow +\infty$

(戈, ...,  $X_J - g(\theta) | \sqrt{n} (G(\theta) r(\theta) G(\theta)')^{-1/2} (X - g(\theta))$ ), 则称  $I(\theta)$  为  $g(\theta)$  的最优渐近正态 (best asymptotic normal) 估计, 简称

BAN估计. 特别, 若  $g(\theta) = \theta$  (权), 则称  $(X)$  为  $\theta$  的BAN估计. 压

注 若  $X_1, \dots, X_n$  独立同分布, 则  $I(\theta) = n i(\theta)$ ,  $I(\theta)/n = i(\theta)$  为  $X_1$  的Fisher信息阵. 另外, 若  $g(\theta)$  与  $\theta$  都是单参数, 则以上定义可表示为

$A i(\theta) = g'(\theta) I(\theta) g'(\theta) = 1$  (没)  $\sim (, [ , (没) ] \% - 1 \dots)$ .

例5.3.9 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\mu, \sigma^2)$ , 则  $\mu$  为  $M_2$  的 BAN.

此解这时  $g(\mu) = \mu$ ,  $g'(\mu) = 1$ ;  $I(\mu) = n/\sigma^2$ ,  $i(\mu) = 1/\sigma^2$ , 因

. 另一方面, 由(5.3.9)式可知  $n(X_1 - \mu)^2/\sigma^2 \sim \chi^2_1$ , 因此  $\mu$  为  $M_2$  的 BAN 估计. ■ 例5.3.10 (续例5.3.8) 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim P(\lambda)$ , 则

为  $\lambda$  的 BAN 估计.

解 由于  $g(\lambda) = e^{-\lambda}$ ,  $g'(\lambda) = -e^{-\lambda}$ ;  $I(\lambda) = n/\lambda$ ,  $i(\lambda) = 1/\lambda$ , 因此

$[g'(\lambda)]^2 / i(\lambda) = \lambda$ . 由(5.3.10)式可知在  $(i, (X) - e^{-\lambda})$  i

$Z \sim 2V(0, \lambda e^{-2\lambda})$ , 因此  $(1 - \lambda)$  为 估计. ■

例5.3.11 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim B(1, \theta)$ ,  $0 < \theta < 1$ , 证明  $\theta = 1/(1-X)$  为  $a_2 = \text{Var}(X)$  的BAN估计 ( $\theta \neq 1/2$ ).

证由中心极限定理可得

$\sqrt{n}(\bar{X} - \theta) \sim N(0, \theta(1-\theta))$ ,

$= 4/\theta^2(1-\theta)^2$

190

### 第五章点估计的性质

其中  $a_2 = 6/(1-\theta)^2 = (p(\theta))$ , 则有  $\sqrt{n}(\bar{X} - \theta) \sim N(0, 1/(1-\theta)^2)$ . 取

$\sqrt{n}(\bar{X} - \theta) \sim N(0, 1/(1-\theta)^2)$ , 则由(5.3.7)式可得

$\sqrt{n}(\bar{X} - \theta) \sim N(0, 1/(1-\theta)^2)$ ,

其中  $T_2 = (1 - 2\theta)^2/(1 - \theta)^2$ . 因此  $\theta$  为  $T_2$  的CAN估计. 另夕卜,

$I(\theta) = 1/(\theta(1-\theta))$ ,  $I(\theta) = 1/(\theta(1-\theta))$  L [ 戈 (幻) (幻) =

.

注 若  $\theta = 1/2$ , 则  $I(\theta) = 0$ , 因此以上结果不再成立. 这时

$(1 - 2\theta)^2/(1 - \theta)^2 = T_2$

因此由定义5.3.3知  $\theta$  为  $T_2$  的BAN估计. I

$\sqrt{n}(\bar{X} - \theta) \sim N(0, 1/4)$ ;  $\sqrt{n}(\bar{X} - \theta) \sim N(0, 1/4)$ ,  $Z(3)$  可知

$-4n[\sqrt{n}(\bar{X} - \theta) - 1/4] \sim N(0, 1)$ .

### 5.3.3 矩估计的相合性和渐近正态性

$\theta, 1/4$ . 由定理 5.3.2 的 由于  $2Z \sim N(0, 1)$ , 因此有 |

矩估计就是用样本均值来估计总体均值. 因此可直接引用随机变量 的和式极限定理, 即大数定律和中心极限定理得到矩估计的相合性和渐 近正态性.

(1) 相合性和大数定律 设, 样本为  $X_1, \dots, X_n$ , 及, 为简单起见, 考虑独立同分布情形. 记  $\bar{X}_n =$

$E(\cdot)$

由大数定律可直接得到矩估计的强相合性:

1"

$\bar{X} = \bar{X}_n \xrightarrow{a.s.} E(X_1) = \mu \quad (a.e.).$

$\bar{Y} = \bar{Y}_n \xrightarrow{a.s.} E(Y_1) = \nu \quad (a.e.).$

另外由引理5.3.2可得

定理5.3.3 设  $G(x_1, \dots, x_k; \theta_1, \dots, \theta_r)$  关于各变元连续, 则  $g(X) =$

$G(x_1, \dots, x_k; \theta_1, \dots, \theta_r)$  为  $g(\theta) = G(\theta_1, \dots, \theta_r)$  的强相合估计.

(2) 渐近正态性与中心极限定理

设  $X_1, \dots, X_n$  为随机变量序列, 则在一定正则条件下有

$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2)$

为  $\text{Var}(X_1)$

则中心矩可表示为原点矩的函数  $\mu_k = E(X^k) = \frac{1}{k!} g^{(k)}(\mu)$

:

5.3 估计量的渐近性质

191

若随机序列独立同分布,  $E(X) = \mu$ ,  $\text{Var}(X) = \sigma^2$ , 则有

$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2)$  或  $K/V(0, \text{Var}(X))$ . a

(5.3.11)

因此  $\bar{X}_n$  为  $\mu$  的相合渐近正态估计. 即原点矩的矩估计都是 CAN. 中心矩也有类似的结果, 见下面的定理5.3.4. 但是, 矩估计一般不是 BAN. 另外, 以上结果还可以推广到多元情形. 记

$a = (a_1, \dots, a_r)$ ;

$M = (M_1, \dots, M_r)^T$ , 则有

$\sqrt{n}(\bar{M}_n - M) \xrightarrow{d} N(0, V)$  或  $K/V(0, \text{Cov}(M))$ ,

(5.3.12)

其中  $V$  的元素为  $\text{Cov}(M_i, M_j)$ .

今考虑  $g(\theta) = G(M_1, \dots, M_r; \theta_1, \dots, \theta_r)$  的矩估计, 由于中心矩

$a_1, \dots, a_r$  可表示为原点矩的函数, 因此可认为  $g$  仅为原点矩的函数, 即 (10)

$= h(h_1, \dots, h_r)$ , 其矩估计为  $\hat{g}(\bar{X}) = A(a_1, \dots, a_r)$  则由引理

5.3.3 的推论 (5.3.8) 式 (即向量形式) 有

定理5.3.4 设  $h(x_1, \dots, x_r)$  关于各变元可导, 则  $g(J)$

为  $g(\theta) = G(M_1, \dots, M_r)$  的相合渐近正态估计, 并且有  $\sqrt{n}(\hat{g} - g) \xrightarrow{d} N(0, V)$

其中  $V = (dh/d\theta_1, \dots, dh/d\theta_r)$ .

证明 在  $g(\theta) = G(\theta_1, \dots, \theta_r)$  中, 由 (5.3.12) 可

$\sqrt{n}(\bar{M}_n - M) \xrightarrow{d} N(0, V)$ , 因此由 (5.3.8) 式可得  $\sqrt{n}(\hat{g} - g) \xrightarrow{d} N(0, V)$ . |

推论 若  $S(V)$  和  $g$  均为向量, 类似结果也成立. 5.3.4 极大似然估计的相合性和渐近正态性

极大似然估计有很好的渐近性质, 在一定正则条件下具有强相合性和渐近正态性, 且为

BAN. 但是这些性质的严格证明甚为复杂, 本节

仅概述其大意, 详细讨论可参阅陈希孺 (1981) 等文献. 我们首先介绍似然函数的基本性质, 这些性质在各方面都有广泛的

以上结果用于矩估计  $=$

$A, \theta = \mu, EY_i = h$  则有

$n$

$i=1$

况  $\sim N(0, V)$ ,  $V_j = \text{Var}(r_i)$ .

}

,  $X_n) \in \mathcal{C}R_p$  为 C-R 分布族, 以下

应用. 设  $X = (X$

仅考虑独立同分布情形(独立样本情形类似). 设,  $\dots, X_n$  为独立同分布样本,  $X_i \sim$

$f(x, \theta)$ ,  $l(\theta) = \log f(x, \theta)$ , 则有

192

第五章点估计的性质

$l(\theta) = \sum_{i=1}^n l(x_i, \theta) = n l(\theta)$

$l'(\theta) = \sum_{i=1}^n l'(x_i, \theta) = n l'(\theta)$

$l''(\theta) = \sum_{i=1}^n l''(x_i, \theta) = n l''(\theta)$ ,  $E[l'(\theta)] = 0$ ,

du

$\text{Var}[l'(\theta)] = -K(\theta)$ ,  $n$

其中  $K(\theta)$  为  $A$  的 Fisher 信息. 由于  $X \sim f(x, \theta)$ , 因此有  $i(\theta) = 1$

和式

$n$

$l(\theta) = \sum_{i=1}^n l(x_i, \theta) = n l(\theta)$ . (5.3.13)  $1=1$

(5.3.13) 式是下面推导渐近性质的出发点, 因为它是一个和式, 所以可以引用大数定律

和中心极限定理. 由 (5.3.13) 式有

$E[l'(\theta)] = E[\sum_{i=1}^n l'(x_i, \theta)] = 0$ ,  $\text{Var}[l'(\theta)] = E[l''(\theta)] = -n i(\theta) = -n$ .

同时亦有  $l''(\theta) = -n i(\theta)$ . 以下进一步假设  $H$  存在二阶矩, 并记

$E[l''(\theta)] = -n i(\theta)$ ,  $\text{Var}[l'(\theta)] = n i(\theta)$ . 易见  $a(\theta) = t(\theta) = -a_2(\theta)$ .

下面将基于 (5.3.13) 式应用大数定律和中心极限定理, 由此可以

得到与似然函数有关的许多重要性质.

引理 5.3.4 设  $(X_1, \dots, X_n) \sim f(x, \theta)$ ,  $\theta \in \mathcal{C}R_p$  为 C-R 分布

族, 并设  $X_1, \dots, X_n$  独立同分布,  $A$  的 Fisher 信息为  $i(\theta)$ , 则有 (1)  $n^{-1} l'(\theta) \rightarrow$

$0$  (a.e.), 且有  $l'(\theta) = o_p(n)$ ;

(2)  $n^{-1/2} l''(\theta) \rightarrow N(0, i(\theta))$ , 即  $l''(\theta) = O_p(n^{1/2})$ ,  $= 2, 3, \dots$ , 特

别有

$1/n \sum_{i=1}^n l''(x_i, \theta) \rightarrow -i(\theta)$  (a.e.),  $l''(\theta) = O_p(n^{1/2})$

;

(5.3.14)

(5.3.15)

(3) score 函数  $S(X, \theta) = l'(\theta)$  有以下渐近正态性:

$-p f(\theta) / V(0, i(\theta))$ ,  $l''(\theta) = O_p(n^{1/2})$ ;  $y_n$

(4) 观察信息  $U(\theta)$  与 Fisher 信息  $i(\theta)$  之间有以下重要关系:  $-[l'(\theta) - i(\theta)] \perp$

$V(0, p_2(\theta))$ ,  $-Z(3) = (nT)$ ,

$y/n$

$[-U_0] \rightarrow O_p(n^{3/2})$ , 证明 根据 (5.3.13) 式, 由大数定律可得

$l'(\theta) = O_p(n^{1/2})$  (a.e.).

(5.3.16) (5.3.17)

5.3 估计量的渐近性质 因此结论 (2) 成立. 当  $A = 1$  时,

“

2 时,  $EJ =$

193

$= 0$ , 因此结论 (1) 成立.  $i(\theta) = 1$ , 因此 (5.3.14) 式成立. 对和式

应用中心极限定理 (见 (5.3.11) 式) 可得

$-E_j/(6>, X_1)]\} - L * 7V(0, \text{Var}_j Z((9, X_I)]).$

当  $L(\&) =$

$1=1$  人 $\{\pm$

由于  $E_j Z(\wedge, J_t)] = 0, \text{Var}_j / C(\wedge, \wedge)] = r(\wedge)$ , 因此(5.3.15)式成立.  $n$

对于和式 $-Ue) = f [-z(\wedge, x_f)]$ 应用中心极限定理得  $1^{\wedge}1$

人 $\{\pm \blacksquare$ 名 $[_{7}(\text{没}, A\backslash)] - e * [-_{7}(n)]\}$

$-^{\wedge} / V(0, \text{Var}([_{-?}(H)]).$

由于  $E, [-Z(0, X,)] = i(\wedge) = n - 7(\wedge)$ , 因此有  $1r$

由此也

边同乘

(4)]-

$1 = 0p(n \sim l)$ 可得

该

为OR分布族, 并设 $X_x, -, X_n$ 独立同分布

上的开集. 则似

$-Z(\text{没}) - /(\text{没}[$

$)] - ^{\wedge}(0, 1/2(\wedge)).$

[可得到 $-Z(<9) = Z(6>) + 0p(n^{\wedge})$ , 因此(5.3.16)式成立. 上式两

1

)即可得到 $[-Z(\wedge)] - * = /-, (<9) +$

式两边同乘  $K\}(0) = (? (n -$

0人  $n-, ,$  此即(5.3. 17)式.  $\blacksquare$

推论 若  $L(0)$ 在  $0^{\wedge}0.$ 的某邻域内存在三阶连续导数, 则有  $\pm L(4) = \pm$ 卵.) + /

$JT(0o)A6 > + yA^TB(0o)A19 + || A6 > || 30p(1),$

(5. 3. 18) 其中 $4(\text{没}o) = \text{汀} \vee (\text{权}o), 5(\wedge^0) = n', Z(0o), \wedge^6 = 0 - 6Qy$  且有

$\#(\text{氏})九 7V(0, f(\text{没}.) ), B(0o)l - i(0o).$  证明 $L$ (幻在氏处进行Taylor展开, 并应用

以上引理即得. | 定理5.3.5 (强相合性)设 $1=(4, \dots, XJt \sim / (x, <9), e0CR?$

,

然方程 $=0$ 在 $n^{\wedge} + \infty$ 时必有解 $0n(X)“(X', \dots, X_{\perp}$ 并且是强 相合的. 即对真参数氏 $e0$ 有

$PSo \backslash X; \lim =^{\wedge}0 \} = 1, \text{氏}e0. (5.3.19) n \blacklozenge + \infty$

证明 考虑的一系列闭邻域 $U_m = \backslash 0fz || \textcircled{1} 1 - \textcircled{2} 0 || ^{8m} 98m > 0 1 \gg$ 并有 $5m - *0(rn -$

$> \infty).$  以下证明, 对充分大的 $a, L(0)$ 在邻域 $l/m$

中的最大值不可能在边界 $|| < 9' - \text{氏} || = \infty$ 上达到. 即对充分大的 $n$ , 必

194

第五章点估计的性质

存在 $A$ (幻的最大值点谷人 $x$ ), 以及零测集人 $, ,$  对一切 $x^{\wedge}An^{\wedge} II - 00 II < am.$ 再令  $6m -$

$o,$  则有  $Peo \backslash x: 0n(X^{\wedge}0Q \} = 1.$ 为了证 明这一点, 主要应用信息不等式 $E/\log AXJ]$

$> Ez[\log gCX,)].$ 在氏

处应用信息不等式可得

$Ej \log AX, A)] > EJ \log / (U')], V^{\wedge}V^{\wedge}0. (5.3.20)$

对于 $\backslash \backslash ef - e. II = 8, ", X$

$1'' - Z \log$

、

$\wedge^o) \rightarrow E0o[1 \log / (\wedge^1, \wedge^0)] (a.e. 7\%).$

$i \text{二}$

$-Z \log / dD - Ej \log^{\wedge}U')](a.e. P9o).$



$n_i = i$

由(5.3.20)可知, 必存在  $V$ 。对于  $n > N_m$ , 除了一个零测集 当  $\epsilon \rightarrow 0$  时必有以外,

$-\frac{1}{n} \log \pi_n(X) > -2 \log \pi(X)$ ,  $W$ -氏  $II = 8n$ .  $n \rightarrow \infty$

11

因此当  $p$ -氏  $II = 0$ , 时有  $L(\sim) > L(8)$ ,  $\wedge(U=0)$ . 这说明,  $L$ (幻在邻域  $U_m$  中的最大值不可能在边界  $\partial U_m$  上达到. 因而最大值点  $\theta_n(x) = \hat{\theta}(x)$ , 应在  $U_m$  的内部, 即应满足:

从(4))  $= 0$ ,  $|\hat{\theta}(1) - \theta| < 3m$ ,  $\forall x \in U_m$ .

令  $\epsilon \rightarrow 0$ , 则有  $1 \rightarrow 0$ , 因而  $\theta_n(x) \rightarrow \theta(x)$  对一切  $x \in A = X \cap U_m$  成立. 而  $U_m$  仍然是零测集, 即  $\wedge(U_m) = 0$ , 因  $\forall x \in A$  时有  $\theta_n(x) = 1 = \theta(x)$

久(幻  $\rightarrow \theta(x)$ ), 即(5.3.19)式成立. 以上定理只是说明, 似然方程必有相合解,

但是未能完全证明极大似然估计的强相合性, 关于相合性更深入的讨论可参见陈希孺(1981)等文献. 不过, 以上定理对于若干常见情形还是可用的. 例如, 如果似

然函数是单峰可导的, 则似然方程的解存在唯一, 且为极大似然估计, 因而是强相合的, 若干常见的指数族分布就属于这种情形.

定理5.3.6(渐近正态性) 设  $X = (X_1, \dots, X_n)^T \sim N(\mu, \Sigma)$ ,  $\theta \in \mathbb{R}^p$  分布族, 并设  $\mu_1, \dots, \mu_p$  独立同分布,  $\theta$  为  $\mathbb{R}^p$  上的开集. 假定似然方程  $U(\theta) = 0$  在  $\theta$  时有相合解  $\hat{\theta}_n \rightarrow \theta$ , 且假定  $L(\theta)$  在  $\theta$  中存在且连续, 则  $\hat{\theta}_n$  为  $\theta$  的BAN, 且有

$\sqrt{n}(\hat{\theta}_n - \theta) \rightarrow N(0, I^{-1}(\theta))$ . (5.3.21) 证明 以下仅就  $\theta$  为一维的情形加以证明,  $\theta$

为向量的情形完全类似.

似. 似然方程  $l(\theta) = 0$  在  $\theta$  处进行二阶展开可得

应用强大数定律有

5.3估计量的渐近性质

195

$L(n) = L(\theta_0) + \frac{1}{2} L''(\theta_0) (\hat{\theta}_n - \theta_0)^2 + o_p(1)$ , 其中  $\theta_0 = \theta$  在  $\theta$  和  $\theta_0$  之间. 上式可化简为

$l(\hat{\theta}_n) = -\frac{1}{2} l''(\theta_0)$

$= -\frac{1}{2} l''(\theta_0) (\hat{\theta}_n - \theta_0)^2 + o_p(1)$  (氏)  $= n$

因此有

其中  $v_n$  和  $w_n$  分别代表上式第一项和第二项. 对上式应用引理5.3.4的结果可知:  $v_n \rightarrow 0$

$n^{-1} L(\theta_n) U(\theta_n) \rightarrow 0$ , 又由假设可知

$A \rightarrow \theta_0$ , 因此  $\frac{1}{n} L''(\theta_0) (\hat{\theta}_n - \theta_0)^2 \rightarrow 0$ . 又由引理5.3.4的(5.3.15)式可知  $\hat{\theta}_n \rightarrow \theta_0$

$\hat{\theta}_n \rightarrow \theta_0$ , 因此由 Slutsky 定理可得(5.3.21)式:

根据定义5.3.3,  $\hat{\theta}_n(X)$  为  $\theta$  的BAN. 渐近正态性在理论上、应用上都有重要意义, 可参见第六章和第七章.

章. 以下推论也是很常用的.

推论1 (5.3.21)式对任意的  $\theta$  成立, 因而对任意  $\theta \neq \theta_0$  都成立,

即有  $y/n(\hat{\theta}_n - \theta_0) \rightarrow 0$ , 且有  $\theta_n \rightarrow \theta_0$   $\Rightarrow p(1)$ ; 同时也有

$\hat{\theta}_n \rightarrow \theta_0$ . 推论2 设  $I(\theta)$  为样本  $X$  的 Fisher 信息阵, 则有

$I(\hat{\theta}_n) \rightarrow I(\theta_0)$ ,  $r(\theta_n)(\theta_n - \theta_0) \rightarrow N(0, I(\theta_0))$ .

(5.3.22)  $-E y_i(\theta) H_{ii}(\theta) = H(\theta)$ . (5.3.23)

$(A - \hat{\theta}_n)^T / (A - \hat{\theta}_n) \rightarrow 0$  (5.3.24) 证明 由(5.3.21)式可得  $\sqrt{n}(\hat{\theta}_n - \theta_0) \rightarrow N(0, I(\theta_0))$ ,

由于

$I(\theta_0) = n I(8)$ , 因此该式可表示为  $\sqrt{n}(\hat{\theta}_n - \theta_0) \rightarrow N(0, I(\theta_0))$ , 由此即得(5.3.22)第1式; 而(5.3.22)第2

式可表示为

$\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 - \frac{1}{n} \sum_{i=1}^n (X_i - \mu)(\bar{X} - \mu) + \frac{1}{n} \sum_{i=1}^n (\bar{X} - \mu)^2$

因此由 $\theta_n$ 的相合性以及Slutsky定理去1律可得(5.3.22)第2式. 由(5.3.22)式即可得到(5.3.23)式以及(5.3.24)式. |

推论3  $|\text{Var}(\bar{X})| \rightarrow 0$  当  $n \rightarrow \infty$  时, 且有

$(\bar{X} - \mu) \sqrt{n} \xrightarrow{D} N(0, \sigma^2)$  (5.3.25)

弘(3)(f)] $\Delta$ 认

196 特别对于单参数有

第五章点估计的性质

A

$\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \xrightarrow{P} \sigma^2$  (5.3.26)

$A/\text{Var}(\bar{X})$  证明由推论1可得

$\{ \text{Var}(\bar{X}) \}^{-1/2} (n(\bar{X} - \mu)) \xrightarrow{D} N(0, 1)$

$\rightarrow \sqrt{n}(\bar{X} - \mu) \xrightarrow{D} N(0, \sigma^2)$ . 由此亦可得到(5.3.25)式和(5.3.26)式. 匪

由极大似然估计的渐近正态性还可推出似然比统计量的渐近/

性, 为此, 首先介绍极大似然估计的随机展开, 这一展开式本身也很 有用.

定理5.3.7 条件同定理5.3.6, 则 $\theta_n = \theta_n(X)$ 可展开为(称为随 机展开)

$\theta_n - \theta_0 = [ -L'(\theta_0) ]^{-1} L''(\theta_0) (\theta_n - \theta_0) + o_p(n^{-1/2})$ . (5.3.27)

$\theta_n - \theta_0 = [ -L'(\theta_0) ]^{-1} L''(\theta_0) (\theta_n - \theta_0) + o_p(n^{-1/2})$ , (5.3.28) 证明由  $U(\theta_n) = 0$  可得

$L'(\theta_0) + Z(\theta_n) [ -L'(\theta_0) ]^{-1} L''(\theta_0) (\theta_n - \theta_0) = 0$ ,

其中 $f$ 在 $A$ 与 $\theta_0$ 之间, 余项的第 $f$ 个分量为 $o_p(n^{-1/2})$ ,  $A^j, 2 \leq j \leq k$

为 $A^j = (\theta_n - \theta_0)^j$ 的第 $j$ 个分量. 由上式可得  $\theta_n - \theta_0 = [ -Z'(\theta_0) ]^{-1} L''(\theta_0) (\theta_n - \theta_0) +$

$o_p(n^{-1/2})$  [  $-Z'(\theta_0) ]^{-1} L''(\theta_0) (\theta_n - \theta_0) + o_p(n^{-1/2})$ . (5.3.29) 由引理5.3.4可知

$[ -Z'(\theta_0) ]^{-1} L''(\theta_0) = [ -Z'(\theta_0) ]^{-1} L''(\theta_0) + o_p(n^{-1/2})$ .

由于 $U'(\theta_0) = V'(\theta_0) = 0$ ; 因此由引理5.3.4, (5.3.29)式中的 余项可表示为

余项 $= o_p(n^{-1/2}) [ -Z'(\theta_0) ]^{-1} L''(\theta_0) (\theta_n - \theta_0) + o_p(n^{-1/2})$ .  $\Delta$ 凡  $n \rightarrow \infty$

由此可得(5.3.27)式. 又由引理5.3.4可得  $[ -Z'(\theta_0) ]^{-1} L''(\theta_0) = [ -Z'(\theta_0) ]^{-1} L''(\theta_0) + o_p(n^{-1/2})$  代人(5.3.27)式即得

$\theta_n - \theta_0 = [ -Z'(\theta_0) ]^{-1} L''(\theta_0) (\theta_n - \theta_0) + o_p(n^{-1/2})$

5.3估计量的渐近性质

197

$= [ -Z'(\theta_0) ]^{-1} L''(\theta_0) (\theta_n - \theta_0) + o_p(n^{-1/2})$ . I 最常见的似然比统计量定义为 $LR(\theta) = 2 [ L(\theta) - L(\theta_0) ]$  它实际上是似然比 $f(x|\theta_n)/f(x|\theta_0)$ 对数的两倍:  $LR(\theta) = 2 \log [ f(x|\theta_n)/f(x|\theta_0) ]$ .

这个统计量在大样本假设检验和区间估计中有重要应用, 也是实际中应用最广泛的统计量之一. 其理论基础就是以下极限定理, 通常就称为似

/

$LR(\theta_0) = 2 [ L(\theta_n) - L(\theta_0) ] - \frac{1}{2} X(p)^2$  证明 对 $LR(\theta_0)$ 进行Taylor展开可得

然比统计量的渐近 定理5.3.8条件同定理5.3.6, 则有

其中余项 $= o_p(n^{-1/2})$

将定理 5.3.7 中的(5.3.27)式

性.

$V'(\theta_0) = 2 [ L'(\theta_0) - L'(\theta_0) ] + 2 [ L'(\theta_0) - L'(\theta_0) ] + o_p(n^{-1/2})$ , 余项,

(5.3.30)

代人上式可得

$L'(\theta_0) = 2 [ L'(\theta_0) - L'(\theta_0) ] + 2 [ L'(\theta_0) - L'(\theta_0) ] + o_p(n^{-1/2})$

$- [ L'(\theta_0) - L'(\theta_0) ] + o_p(n^{-1/2}) + o_p(n^{-1/2})$ .

由于  $L(3)(\cdot)/n = O_p(1)$ ,  $\hat{\theta} = O_p(n^{-1/2})$  因此余项  $= O_p(n^{-1/2})$ ; 另有  $U(\hat{\theta}) = W(2)$ , 所以有

$$f'(\theta_0) = Z(\theta_0)' [-Z(\theta_0)]^{-1} 1(\hat{\theta}) + o_p(n^{-1/2}). \quad (5.3.31)$$

由(5.3.17)式可得  $[-Z(\theta_0)]^{-1} = \frac{1}{n} \Sigma (x_i - \bar{x})(x_i - \bar{x})' + O_p(n^{-3/2})$ , 该式代入上式可得  $LR(\hat{\theta}_0)$

$$= \frac{1}{n} \Sigma (x_i - \bar{x})(x_i - \bar{x})' W(2) + W(2)$$

$$= \frac{1}{n} \Sigma (x_i - \bar{x})(x_i - \bar{x})'$$

$$= \frac{1}{n} \Sigma (x_i - \bar{x})(x_i - \bar{x})' + O_p(n^{-1/2})$$

由于  $i(\theta_0) = \frac{1}{n} \Sigma (x_i - \bar{x})(x_i - \bar{x})'$ , 故有  $\frac{1}{n} \Sigma (x_i - \bar{x})(x_i - \bar{x})' \rightarrow \Sigma (x_i - \bar{x})(x_i - \bar{x})'$ , 所以  $LR(\hat{\theta}_0) \rightarrow \Sigma (x_i - \bar{x})(x_i - \bar{x})'$ .

其中

定理5.3.9条件同定理5.3.6, 则似然比统计量有以下等价形式:  $j^2/j^2$

$$SC(\hat{\theta}_0) \sim \chi^2(p), \quad (P), \quad v \in \mathbb{R}. \quad (5.3.33) \quad \text{ira}(\theta_0) \sim \chi^2(p), \quad WDr(e_0) \sim \chi^2(p), \quad v \in \mathbb{R}. \quad (5.3.34)$$

198

第五章点估计的性质 (5.3.35)

$$(5.3.36)$$

$$\Sigma SC'(\theta_0) = \{f(\theta_0)\}$$

dL

$$\leq \frac{1}{n} \Sigma (x_i - \bar{x})(x_i - \bar{x})'$$

$$WD(60)$$

$$WD \setminus \theta^* = (1 - \theta) T[ -l(\theta) ] \{ \theta - \theta^* \}.$$

由(5.3.31)式和(5.3.32)式可得

证明

$$LR(\hat{\theta}_0) = Sr(\hat{\theta}_0) + O_p(n^{-1/2}); \quad f'(\theta_0) = SC(\hat{\theta}_0) + O_p(n^{-1/2}).$$

因而可以得到(5.3.33)式. 而由(5.3.28)式可得  $L(\hat{\theta}_0) = Z(\hat{\theta}_0)' (\hat{\theta}_0 - \theta_0) + o_p(n^{-1/2})$ .

该式代入(5.3.32)式可得

$$LR(\hat{\theta}_0) = \frac{1}{n} \Sigma (x_i - \bar{x})(x_i - \bar{x})' W(2) + W(2).$$

由  $\theta_n$  的相合性以及 Slutsky 定理可以得到(5.3.34)的第一式, 再由(5.3.17)即可得到第二式. 麵

定理5.3.7—定理5.3.9对任意的  $\theta_0$  成立, 因而对任意  $\theta \in \mathbb{R}$  都成立. (5.3.35)式通常称为 score 统计量或 Rao 统计量; 而(5.3.36)式通常称为 Wald 统计量. 注意, (5.3.36)式与(5.3.24)式是一致的. 另外, 定理5.3.8—定理5.3.9可以进一步推广到子集参数的情形, 见第六章.

习题五

1. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i$  的密度函数为  $f(x) = e^{-x} \cdot \exp(-\lambda x)$ , 其中  $x \in \mathbb{R}$ . 求  $\lambda$  的 C-R 下界.

2. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim \text{Rayleigh}$  分布:  $f(x) = \frac{x}{a^2} \exp(-x^2/2a^2)$ ,  $x \geq 0$ .

(1) 求  $\lambda$  的 C-R 下界;

(2) 判断是否存在有效的无偏估计.

3. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i$  的密度函数为

$$f(x) = \frac{1}{A} x^{A-1} \exp(-x^A)$$

$$f(x) = \frac{1}{A} x^{A-1} \exp(-x^A), \quad x \geq 0$$

$$1.0, \quad x \geq 0.$$

(I) 求  $A$  的 Fisher 信息  $I(A)$ ; (2) 求  $h(1)$ , 使  $A(A)$  存在有效无偏估计.

4-设  $\theta_1, \dots, \theta_m$  为  $m$  样本,  $X_t \sim T(A, r)$ , 其中  $r$  已知. 分别求  $A$  和  $1/A$  的 MLE, 并判断其方差是否能达到 C-R 下界.

5. 设随机变量  $X$  的密度函数为  $f(x) = (2\theta)^{-1} \exp\{-x/\theta\}$ ,  $x \geq 0$ , 其

,  
习题五

199

中  $\theta > 0$ . 分别求  $\theta$  和  $(1+\theta)^{-1}$  的 UMVUE, 并检验它们能否达到各自的 C-R 下界 (提示:  $r = |x|$  是关于  $e$  的完备充分统计量).

6. 设  $I_1, \dots, I_n$  为  $i.i.d.$  样本, 服从两点分布, 即  $I_i = 1$  以

,  
并且方差  $\text{Var}(S/n)$  达不到  $(1-p)$  的 C-R 下界.

7. 设  $A_1, \dots, A_n$  为  $i.i.d.$  样本,  $X_i \sim N(0, \sigma^2)$ . 试求  $\sigma^2$  的 UMVUE, 问其能否达到  $\sigma^2$  的无偏估计的 C-R 下界.

8. 设  $\mu_1, \dots, \mu_n$  为  $i.i.d.$  样本,  $X_i \sim N(\mu_i, \sigma^2)$ , 其中  $\mu_i$  是未知参数,  $\sigma^2$  是已知参数.

(1) 求  $\mu_i$  的 UMVUE, 其中  $\sigma^2$  为固定的正数; (2) 设验证 (1) 中所求的 UMVUE 是否能达到  $\mu_i$  的 C-R 下界.

$p$ , 其中  $p \in (0, 1)$ .

9.  $g(\theta)$  的估计, 则有

1). 证明:  $p(1-p)$  的 UMVUE 为  $S/n - (S/n)^2$

$I$  为 C-R 分布族,  $g(\theta)$  在  $\theta$  上可微,  $T(X)$  是

$E_j T(X) + g'(\theta) T(X) = g(\theta)$ , 其中  $g'(\theta) = E_j T(X) = g(\theta)$ .

10. 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $A$  的密度函数为  $f(x)$  且对  $\forall x, \theta > 0$ , 又设  $S/n$  为  $g(\theta)$  的任意一个无偏估计, 证明:

11. 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $X_i \sim N(\mu, \sigma^2)$

已知. 证明  $\bar{X}$  的 MLE 是  $\mu$  的无偏的相合估计. 12. 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $X_i$  的密度函数为

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

考虑  $\theta$  的三个估计:  $T_1 = S/n$ ,  $T_2 = S^2/n^2$ ,  $T_3 = S^3/n^3$

试判断哪些是  $\theta$  的无偏估计, 哪些是  $\theta$  的相合估计.

13. 1 in

其中  $\theta \in (0, 1)$

,  
设  $\theta$  为未知参数, 试找  $\theta$  的一个相合估计.

为  $i.i.d.$  样本,  $X_1$  的密度函数为  $f(x|\theta) =$

13. 1 in

相合估计, 并证明你的结论.

$\theta \in (0, 1)$

$\theta \in (0, 1)$

$\theta \in (0, 1)$

200 第五章点估计的性质

14. 设  $A_1, \dots, A_n$  总相互独立, 其中  $A_i \sim N(\mu_i, \sigma^2)$ ,  $E(X_j) = \mu_j$ ,

$\text{Var}(X_j) = \sigma^2$ ;  $\mu_1, \dots, \mu_n$  为  $i.i.d.$ ,  $E(X_n) = \mu$ ,  $\text{Var}(X_n) = \sigma^2$ . 证

明:  $\bar{X}$  是  $\mu$  的相合估计.

15. 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $A \sim N(0, \sigma^2)$ , 证明:  $S^2/n$  是

$(\sigma^2/n)$

$|X_n|$  是  $\theta$  的相合估计的充要条件为  $A = A_2$ .

16. 设  $K_n, n=1, 2, \dots$  为正态随机变量序列,  $y_n \sim N(a, \sigma^2/n)$ ,  $n=1, 2, \dots$ , 证明:  $\bar{y}_n$  是  $a$  的相合估计的充要条件为  $\sigma^2 > 0$ .

17. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim (1, \theta)$ , 证明:  $T(X) = \sum_{i=1}^n X_i$  是  $\theta$  的相合与强相合估计.

18. 设  $A_1, \dots, A_n$  为 i.i.d. 样本,  $X_i \sim f(x; \theta)$ . 证明:  $\bar{X}_n$  是  $\theta$  的强相合估计, 其中

(1)  $f(x; \theta) = \frac{1}{\Gamma(\theta)} e^{-x} x^{\theta-1}$ ;  $x > 0, \theta > 0$ ;

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

(3)  $f(x; \theta) = \frac{1}{\Gamma(\theta)} e^{-x} x^{\theta-1}$  (提示: 参考例 5.3.5).

,  $\theta > 0$ ; 其分布函数  $F(x; \theta) > 0, x > 0$ .

\* 19. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i$  的分布函数  $F(x)$  为连续函数, 并且有  $F(0) = 1, F(x) < 1, x > 0$ . 证明:

(1)  $\bar{X}_n$  是  $\theta$  的强相合估计;

(2) 若  $F(0) \neq 0$ , 则  $n(F(\bar{X}_n) - F(0)) \xrightarrow{d} Y$ , 其中  $F$  的分布函数为  $F(y) = e^{-y} \int_0^y f(t) dt, y \geq 0$  (提示:  $F(\bar{X}_n) = F(0) + F'(0)(\bar{X}_n - 0) + o_p(\bar{X}_n)$ ), 并证  $n(\bar{X}_n - \theta) \xrightarrow{d} 0$ ).

\* 20. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $A$  服从均匀分布  $U(0, 1)$ ,

其中  $\theta \in R$  未知, 证明:  $\bar{X}_n$  为  $\theta$  的强相合估计 (提示: 证明  $\text{Var}(\bar{X}_n) \rightarrow 0$ ).

21. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i$  服从 Cauchy 分布, 即  $f(x; \theta) =$

$\frac{1}{\pi} \frac{\theta}{x^2 + \theta^2}$ , 其中  $\theta \in R$  未知. 证明:  $\bar{X}_n$  不是  $\theta$  的相合估计, 其中

$\theta \neq 0$ .

22. 设  $S_n^2$  证明:  $\bar{X}_n$  的渐近分布为  $N(\theta, \sigma^2/n)$ .

习题五

201

23. 设  $X_1, \dots, X_n$  独立同分布, 并设  $EX_i = \mu, \text{Var}(X_1) = \sigma^2$ ,

$\text{Var}(\bar{X}_n) = \sigma^2/n, S_n^2 = \sum_{i=1}^n (X_i - \bar{X}_n)^2$ . 求  $W_n$  以及  $a, b$  的 CAN, 并求

它们的渐近分布.

24. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim (1, \theta), \theta > 0$ .

(1) 试基于无偏估计  $g(\theta) = 1/(1 - \theta)$  的一个矩估计  $\hat{\theta}_n(X)$ ;

(2) 求  $\hat{\theta}_n(X)$  的渐近分布, 并证明  $\hat{\theta}_n(X)$  是  $g(\theta)$  的 BAN 估计.

25. 设  $X_1, \dots, X_n$  为 i.i.d. 样本.

(1) 若  $X_i \sim \Gamma(\theta, 1), \theta > 0$ , 证明:  $n(\bar{X}_n - \theta)$  的分布收敛于指数分布  $E(1/e)$ ;

(2) 若  $X_i \sim \Gamma(\theta, \theta), \theta > 0$ , 证明:  $n(\bar{X}_n - \theta)$  的分布收敛于指数分布  $E(1/\theta)$ , 并由此给出  $\theta$  的一个相合估计;

(3) 若  $X_i \sim \Gamma(\theta, \theta), \theta > 0$ , 证明:  $\bar{X}_n$  的分布为  $\Gamma(2n, \theta)$ , 而不是收敛到  $\Gamma(2, \theta)$ .

26. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,

$X_i \sim U(0, 1)$ .

(1) 若  $A \sim U(0, 1)$ , 令  $r_n = (n X_n)^{1/n}$ , 证明:

$r_n \xrightarrow{d} (0, e^2)$ , 其中  $e$  是自然对数的底数 (提示: 令  $Z_n = \ln Y_n =$

$-\ln X_n$ ).

然后用中心极限定理);

(2) 若 $X_i$  推导相应的结果, 并据此给出 $\theta$ 的一个相合估计.

27. 设 $X_1, \dots, X_n$ 为i. i. d.样本,

$A = \langle 9(1-x)^{1-x} \rangle$ , 其中 $0 < x < 1, 0 > 0$ .

(1) 证明:  $T_n = \bar{X}$ 是 $\theta$ 的一个矩估计;  $1-X$

(2) 证明:  $\hat{\theta} = \frac{1}{n} \sum_{i=1}^n X_i$ , 其中 $\theta = 0$ ,

没 $(\frac{1}{2})^2 = \frac{1}{4}$  \*

28. 设 $X_1, \dots, X_n$ 为i. i. d.样本,  $X_i \sim N(\mu, \sigma^2)$ , 即  $E(X_i) = \mu$

$\text{Var}(X_i) = \sigma^2 < \infty$ . 设函数的二阶导数在 $\mu$ 处连续, 且  $f'(\mu) = 0$ .

$(\frac{1}{n} \sum_{i=1}^n X_i)$

$\frac{1}{2\sigma^2}$

(1) 证明:

第五章点估计的性质 且  $n[h(X)]$  的渐近分

12 布为其中  $V = X_1$

;

1—r (2) 证明: 当 $ZZ=y$ 时,  $\langle \frac{1}{n} \sum_{i=1}^n X_i \rangle$  弘(1 其

中 $v \sim$

\*29. 设 $X_1, \dots, X_n$ 为i. i. d.样本,  $X_i \sim A^{\theta}$

计 $6 \ln, 52, \dots$ , 么, 么, 其中 $8 \ln = X_2$  为 $\sigma^2$ 已知条件下#的UMVUE; 71

$\frac{1}{n} \sum_{i=1}^n X_i$

— $(\frac{1}{n} \sum_{i=1}^n X_i^2)$ 为 $\sigma^2$ 未知条件下 $\sigma^2$ 的

UMVUE;  $63n = X_2$  为 $f$  的 MLE;  $S_4n = \max(0, S_{1n})$  为  $L$  的改进. 证明: 当时,

$5 \ln, 52n, 33n$ , 奴有相同的渐近分布, 并且都是#的BAN

估计.

30. 设 $X_1, \dots, X_n$ 为i. i. d.样本,  $X_i$ : 存在前4阶矩. 求下列统计量

的渐近分布:

$(\frac{1}{n} \sum_{i=1}^n X_i^2)$ ;

(2)  $g_2(X) = (X, S_2), S_2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ . '  $n \rightarrow \infty$

31. 设 $X_1, \dots, X_n$ 为i. i. d.样本,  $X_i \sim N(\mu, \sigma^2)$ ,  $\sigma^2$  已知. 设 $p = P(X_i > a)$ 的极大似然估计为人 求的渐近分布.

32. 设 $A_1, \dots, A_n$ 为i. i. d.样本,  $X_i$ 服从两点分布, 即 $X_i \sim (1,$

69, 证明:  $\arcsin \sqrt{\frac{1}{n} \sum_{i=1}^n X_i}$ 具有渐近正态分布, 并且其方差与  $\mu$  参数无关.

,  $\sigma^2$ ). 现有/

的四个估

## 第六章参数假设检验

参数估计与假设检验是统计推断的两个主要组成部分. 但是假设检验所研究的问题及其解决方法与参数估计有很大不同. 假设检验问题是 要对“有关总体分布的某种判断(或假设)”是否成立进行鉴定(或检验), 以便了解总体分布的有关性质. 例如, 某种元件的寿命 $X$ 服从指

假设检验的内容十分丰富, 本章主要介绍其基本理论、基本方法和某

些应用. 第6. 1节通过一个实例说明假设检验所研究的问题以及解决问题

的基本思想方法, 并由此引出否定域、检验函数、功效函数以及两类错误

等Neyman - Pearson理论的基本概念. 第6. 2—6. 5节介绍假设检验的Ney-

man - Pearson理论, 并应用于常见的正态分布和指数族分布的假设检验问

题, 其中包括单边检验、双边检验以及多参数的检验等问题. Neyman - Pearson理论在数

学上很完美，但在应用上有较多的限制.第6. 6—6. 7 节分别介绍具有广泛应用价值的似然比检验和拟合优度检验，特别，第 6. 6 节 还比较详细地介绍了文献中常见的score检验.有关本章内容，可参见陈希孺(1981,1999),陈家鼎等(1993),茆诗松等(1986,1998), 范金城, 吴可法 (2001), 郑忠国(1998), Rao, C. R. (1973), Lehmann(1986) , Shao(1998), Zacks (1981) , Bickel, P. J. and Doksum, K. A. (1977)等文献.

### 6.1假设检验的基本概念

假设 $X \sim (HP, )$ ,  $\theta$ 道 或 $X \sim f(x, \theta)$ ,  $\theta \in \Theta$ .假设检验问题 可表示为 $H_0: \theta \in \Theta_0$ . 其中 $H_0$ 称为原假设,  $H_1$ 称为对立假设.若 $\Theta_0 = \{\theta\}$  则称应用上要求其平均寿命不低于2 000小时, 要通

数分布 过抽样 $X_1, X_2, \dots, X_n$ , 来检验这批元件是否合格.这可归结为一个假设 检验问题, 即研究假设“ $\theta \geq 2000$ ” 是否成立, 如果假设坑成立, 则认为这批元件合格, 否则认为不合格.在实用上, 很多数据分析问题 都能够归结为假设检验问题, 它有极其广泛的应用, 因此构成了统计推 断的重要组成部分.

为简单假设, 其他皆称为复合假设, 即简单假设为 $H_0$

•加②

Qd

### 204 第六章参数假设检验

以下通过一个例子来说明假设检验所研究的问题以及解决问题的基本思想方法.

例 6.1.1 通过观测试验来检验某种新药, 看其疗效是否不低于

0. 75.这可对 $n$  个试验者, 观察其服用新药的疗效.一般可假设  $X_1, \dots, X_n$  独立同分布, 且有

$X_i \sim \text{Bernoulli}(p)$ ,  $p \in [0, 1]$ , 第 $i$ 试验者治愈,  $X_i = 1$ , 未治愈,  $X_i = 0$ .

即 $X$  问题:

广

$X_i = 10$ , 第 $i$ 试验者未愈  $i=1, \dots, n$ ,

,

$\theta$  表示治愈的概率, 则这一问题可化为一个假设检验

$H_0: p \geq 0.75$  vs  $H_1: p < 0.75$ . (6.1.1)

这是一个复合假设检验问题, 以下说明其解法;为明确起见, 不妨设  $n=30$ .

可取 $T = \sum_{i=1}^n X_i$  作为检验统计量.易见,  $T$ 表示治愈  $n$  的人数,  $r$  越大, 表示治愈率 $\theta$  可能越大, 由直接计算可知  $30 \times p(r^k) = g(p)$

$p(r^k) = g(p) = \binom{30}{k} p^k (1-p)^{30-k}$

其中 $f_l(A, 30-A; +1)$ 为不完全卢函数(见第一章), 它是 $\theta$  的增函数, 经查表有

$P(T \geq 27, p=0.75) = 0.05$ ;  $P(T \geq 27, p=0.75) = 0.05$ ;  $P(T \geq 27, p=0.75) = 0.05$ ;

$P(T \geq 27, p=0.75) = 0.05$ ;  $P(T \geq 27, p=0.75) = 0.05$ ;

(6.1.2) 由(6.1.2)式即可对假设检验问题(6.1.1)进行统计推断如下(一定意义上可看作是统计意义下的“反证法”).若成立, 则(6.1.2)式成立, 即 $p \geq 0.75$ 时,  $P(T \geq 27) \leq 0.05$ .这表明:治愈率名0.75的条件下治愈人数多27的概率很小, 约0.05.因此, 当 $p \geq 0.75$ 时, 事件“ $T \geq 27$ ”出现的可能性极小, 几乎不可能.所以, 若真的抽样到  $n=30$ ,  $A=27$ ,  $\dots$ ,  $A=30$ , 使 $T=T(X) = \sum_{i=1}^n X_i \geq 27$ , 则认为 $H_0$ 不成立;因为  $1=1$

愈人数多27的概率很小, 约0.05.因此, 当 $p \geq 0.75$ 时, 事件

“ $T \geq 27$ ”出现的可能性极小, 几乎不可能.所以, 若真的抽样到  $n=30$

$A=27, \dots, A=30$ , 使 $T=T(X) = \sum_{i=1}^n X_i \geq 27$ , 则认为 $H_0$ 不成立;因为  $1=1$

在统计问题中, 认为小概率事件在一次抽样时不可能发生, 这显然是合理的.由此可得到假

设检验问题(6.1.1)的一个解,它表示为如下“否定域”:

$30 \quad 1 \quad R = \{x: T(x) \geq g\}$

若抽样到 $X=x \in R$  则否定原假设 $H_0$ ;若抽样到 $x \notin R$ , 则不否定 $H_0$ .

当然这种推理也有判错的可能,由(6.1.2)式可知,  $H_0$ 成立而 $X=x \in R$  的概率在0.05. I

### 6.1假设检验的基本概念

205

例6.1.1中所用的推理方法可推广到更一般的情形. 6.1.1 否定域与检验函数

假设考虑一般的假设检验问题  $H_0: \theta = \theta_0, H_1: \theta > \theta_0$ .

(6.1.3)

通常可设计一个统计量 $T = T(X)$ ,与 $\theta$ 上的区域 $R$ 对应,使得 $R = \{x: T(x) \in W\}$ ,其常见的形式为 $R = \{x: T(x) \geq c\}$ (见例6.1.1)或 $R = \{x: T(x) \leq c\}$ :

$T(X) \geq c$ ,并有 $1 - \alpha$ 的概率,其中 $\alpha$ 表示 $H_0$ 成立时(即 $\theta = \theta_0$ 时)的概率.通常取 $\alpha$ 充分小,如 $\alpha = 0.05$ 或 $\alpha = 0.01$ 等.这说明,假设 $H_0$ 成立时,事件“ $X \in R$ ”出现的可能性极小( $\alpha$ ).因此,若一次抽样得到 $X = x \in R$ ,则应否定原假设 $H_0$ .

定义6.1.1 对于假设检验问题(6.1.3),  $\theta$ 上的一个区域 $R$ 称为否定域,若抽样到 $X \in R$ 则否定 $H_0$ ;其余集 $R^c$ 称为假设检验问题的接受域.若对任意的 $\theta \in \Theta$ 有 $P_{\theta}(X \in R) \leq \alpha$ ,则称为假设检验问题(6.1.3)的水平为 $\alpha$ 的否定域,  $\alpha$ 称为 $R$ 的水平,

其余集 $R^c$ 亦称为假设检验问题水平为 $1 - \alpha$ 的接受域. I 根据例6.1.1和以上说明,假设检验问题(6.1.3)基于否定域的统计推断如下.若一次抽样得到 $X = x \in R$ ,则否定原假设 $H_0$ . , 因为 $H_0$ 成立

时, 概率 $P_{\theta_0}(X \in R)$ 极小(为 $\alpha$ ).这时犯错误(即判错)的概率为 $\alpha$ .反之, 若云, 则不否定 $H_0$ .常见的否定域的形式为 $R = \{x: T(x) \geq c\}$ ,  $R = \{x: T(x) \leq c\}$ , 以及 $R = \{x: c_1 < T(x) < c_2\}$ 等.为了便于数学上的处理, 否定域可加以推广.因为任一集合与示性函数一一对应, 所以可引入以下定义:

定义6.1.2 对于假设检验问题(6.1.3), 称以下 $\phi(x)$ 为它的一个检验函数:

$\phi(x) = \begin{cases} 1, & x \in R \\ 0, & x \notin R \end{cases}$

, 0,  $x \notin R$ ,

其中 $\phi(x)$ 表示抽样到 $X = x$ 时否定原假设的概率.若对任意的

$\theta \in \Theta$ 有 $E_{\theta}[\phi(X)] \leq \alpha$ , 则称为假设检验问题(6.1.3)的水平为 $\alpha$ 的检验. I

易见, 否定域与检验函数 $\phi(x)$ 是完全等价的.若 $\phi(x) = 1$ (即 $\phi(x) = 1$ ), 则否定 $H_0$ (否定 $H_0$ 的概率为1);若 $\phi(x) = 0$ (即 $\phi(x) = 0$ ), 则不否定 $H_0$ (否定 $H_0$ 的概率为0);且有 $E_{\theta}[\phi(X)] = P_{\theta}(X \in R)$ .检验函数还可以进一步推广到随机化检验.

206

### 第六章参数假设检验

定义6.1.3 对于假设检验问题(6.1.3), 任一满足 $0 \leq \phi(x) \leq 1$ 的函数 $\phi(x)$ 称为一个随机化检验, 其中 $\phi(x)$ 表示:若抽样到 $X = x$ 则以

的概率 $\phi(x)$ 否定原假设 $H_0$ .若对任意的 $\theta \in \Theta$ 有 $E_{\theta}[\phi(X)] \leq \alpha$ , 则称 $\phi(x)$ 为假设检验问题(6.1.3)的水平为 $\alpha$ 的检验. I

随机化检验在实际中用得不多,但在理论上有一定意义.注意,  $E_{\theta}[\phi(X)]$ 表示参数为 $\theta$ 时, 否定 $H_0$ 的平均概率.

我们亦可从统计判决函数的观点出发, 说明假设检验问题的基本概念.

对于假设检验问题(6.1.3), 最终目的只有两个:接受原假设 $H_0$ 以及否定原假设 $H_0$ .因此判决空间一般只包含两个元素, 即 $\mathcal{D} = \{0, 1\}$ , 其中判决 $d = 0$ 表示 $H_0$ 成立, 判决 $d = 1$ 表示 $H_0$ 不成立.关于

统计判决函数由于判决空间 $\mathcal{D}$ 只有0, 1两个值, 因而很自然地取为示性函数



$f_1, X \in \mathcal{X},$   
 $5(x$

$) = I \setminus \{x: x \in \mathcal{X}\} = \mathcal{J} -$

$[0, 1]$ , 这就是定义6.1.2中的否定域与检验函数.在数学上, 如定义6.1.3所述, 亦可取判决空间为 $D = [0, 1]$ , 这时 $0 \sim 1$ 对应于随机化检验  $\phi(x)$ .

### 6.1.2两类错误及功效函数

对于假设检验问题(6.1.3), 给定一个否定域 $R$ 或一个检验函数  $\phi(x)$ , 就是检验问题的一个解.问题是如何判别一个检验或  $\phi(x)$  的好坏?如何求最优解?根据统计判决函数的观点, 就是要求风险函数最小的解.对于假设检验问题, 通常可取 $0 \sim 1$ 损失函数, 即判错时损失为 $I$ , 判对时损失为 $0$ , 具体可表示为

$f_1(\phi) = \int_{\mathcal{X}} I(\phi(x) \neq 1) dP(x) = \int_{\mathcal{X}} \phi(x) dP(x) = 0,$

$[0, 1]$ 且 $d=0$ 或 $d=1$ . 由此容易得到非随机化检验 $\phi(x) = 1$  的风险函数为

$f_1(\phi) = \int_{\mathcal{X}} I(\phi(x) \neq 1) dP(x) = \int_{\mathcal{X}} \phi(x) dP(x) = 0,$

在这一风险函数中, 第一式表示成立但被否定的概率, 即引起的 损失; 第二式表示“付. 不成立但未被否定的概率”, 也是引起的损失; 通常称为犯以上第一、第二两种错误概率的大小. 由于观察值的随机性, 任何检验问题的解都有可能犯这两种错误, 因此引入以下定义:  
 $\alpha = P(\phi(X) = 0 | H_0), \beta = P(\phi(X) = 1 | H_1).$

#### 6.1 假设检验的基本概念

207

定义6.1.4 对于假设检验问题(6.1.3)的一个否定域 $R$ 或一个检验函数, 其第一类错误 $\alpha$ 定义为 $H_0$ 成立时, 否定 $H_0$ 的概率,

其第二类错误 $\beta$ 定义为 $H_1$ 成立时, 不否定 $H_0$ 的概率, 即 $\beta = P(\phi(X) = 0 | H_1)$

即  $\alpha = \int_{\mathcal{X}} \phi(x) dP_0(x)$ , 或  $\beta = \int_{\mathcal{X}} (1 - \phi(x)) dP_1(x)$

尺),  $\alpha > 1$  或  $\beta > 1$  理论上讲, 一个好的检验应该使以上风险函数尽量小, 即两类错误  $\alpha$  和  $\beta$  都尽量小, 但事实上, 当样本容量固定时, 不可能使二者

同时都很小. 首先看一个例子.

例6.1.1 (续) 对于假设检验问题(6.1.1), 取 $R = \{x: T \geq 27\}$  为  $30$

否定域, 即检验问题的一个解, 其中 $r =$  我们来看其相应的两  $\alpha = 1$

类错误. 由以上定义可知,  $\alpha = P(T \geq 27 | H_0) = 0.75$ ; 它表示治愈率 $< 0.75$

的条件下30个病人中治愈人数 $> 27$ 的概率. 由前面的计算可知, 当 $\theta = 0.75$ 时 $\alpha(\theta)$

$< 0.05$ , 而当 $\theta < 0.75$ 时 $\alpha(\theta) > 0.05$  (因为 $\phi(\cdot, 27, 4)$ 为 $\theta$ 的增函数). 因此对否定域 $R$

$= \{x: r \geq 27\}$ 来说, 其第一类错误很小, 都小于 $0.05$ . 但是其第二类错误就很难达到 小于

$0.05$ 的水平. 由以上定义可知,  $\beta(\theta) = P(T < 27 | H_1) = 1 - \alpha(\theta)$ . 它表示治愈率

$> 0.75$ 的条件下30个病人中治愈人数 $< 26$ 的概率, 直观上看, 这一概率并不一定都很小. 由前面的计算可知,

$\beta(\theta) = P(T < 27 | H_1) = 1 - \alpha(\theta) = 1 - \phi(27, 4, \theta)$ , 这是 $\theta$ 的减函数,  $\theta$ 越大

$\beta(\theta)$ 越小. 当 $\theta = 0.75$ 时 $\beta(\theta) = 1 - \alpha(\theta) = 0.25$ , 而当 $\theta$ 从  $0.75$

逐渐增加时,  $\beta(\theta) = P(T < 27 | H_1)$ 从  $0.25$ 逐渐减少, 但也不可能很快减少到 $0.05$ . 例如,

当 $\theta = 0.85$ 时,

$\beta(\theta) = P(T < 27 | H_1) = 1 - \alpha(\theta) = 1 - \phi(27, 4, \theta) \approx 0.58$  还比较大, 与  $0.05$  相差很大 (直观上看, 治

愈率等于 $0.85$ 时, 30个人中治愈人数小于27的可能性 并不一定很小). 因此, 要求第二类错

误 $\beta(\theta)$ 对一切 $\theta > 0.75$ 都很小是不可能的. ■

以下定义假设检验的功效函数 (亦称势函数), 它可以统一第I、第II两类错误, 也可以使

我们进一步看到，同时要求两类错误都尽量小 一般是不可能的。另外，功效函数也便于我们今后讨论假设检验问题的 最优解。

定义6. 1. 5 对于假设检验问题(6. 1. 3)的一个否定域 $R$ 或一个检验函数其功效函数(亦称势函数)定义为

$\beta(\theta) = E[J(X)] = P(X \in R | \theta)$ ，其中 $\beta(\theta)$ 表示当参数取 $\theta$ 时，否定 $H_0$ 的平均概率。

## 208 第六章参数假设检验

功效函数可以统一第I、第II两类错误。例如对否定域来说，当 $\theta = \theta_0$ 时， $\beta(\theta_0) = P(X \in R | \theta_0)$

而当 $\theta \neq \theta_0$ 时， $\beta(\theta) = P(X \in R | \theta)$

由此我们可以从功效函数 $\beta(\theta)$ 看出第I、第II两类错误的变化趋势。易见，要求第一类错误 $\alpha$ 尽量小，就是要求 $\beta(\theta_0)$ 尽量小；而要求第二类错误 $1 - \beta(\theta)$ 尽量小，就是要求 $\beta(\theta)$ 在 $\theta \neq \theta_0$ 上尽量大。显然，这是很困难的，例如在 $\theta_0$ 和 $\theta$ 的边界附近就很困难。

例6. 1. 1(续) 考虑否定域 $R = \{x: T \geq t_0\}$ 的功效函数。由定义可得  $\beta(t) = P(T \geq t_0 | t) = 1 - \Phi\left(\frac{t_0 - t}{\sigma/\sqrt{n}}\right)$

这时有 $\beta(0.75) \approx 0.05$ ，当 $t \rightarrow 0.75$ 时 $\beta(t) \rightarrow 0.05$ 。但是若要 $t > 0.75$ 时  $\beta(t) = 1 - \alpha \approx 0.95$ ，则必须要求 $\beta(t) \approx 0.95$ 。这显然是

很困难的，只有 $t$ 很接近1时才有可

能(见图6. 1. 1)。

根据以上分析，假设检验问题(6. 1. 3)的最优解不可能指望两类错

误同时都很小，只能采取某种妥协方案，这就是有名的Neyman - Pearson准则。其主要精神就是在控制第

一类错误充分小的前提下，要求第 二类错误尽量小。

例6. 1. 1的功效函数图

图6. 1. 1 6. 1. 3 Neyman - Pearson准则与一致最优势检验

为明确起见，对于假设检验问题(6. 1. 3)的一个解，统称为一个检验 $\phi(x)$ ，它可以是示性函数(对应于否定域 $R$ )，也可以是在 $[0, 1]$ 取值的任意一个函数；其功效函数为 $\beta(\theta)$ 。根据Neyman - Pearson准则(简称N-P准则)，首先要确定一个第一类错误的允许水平  $\alpha$ ，其值很小(如 $\alpha = 0.05$ 或 $0.01$ 等等)，根据实际要求决定。在此前提下，寻找第二类错误尽量小(即功效函数 $\beta(\theta)$ 尽可能大)的解。其严

格定义如下：

定义6. 1. 6 一致最优势检验(uniform most powerful test, 简称为

UMPT)。对于假设检验问题(6. 1. 3)和给定的水平 $\alpha$ (通常 $0 < \alpha < 1$ ，且 很小)，若检验 $\phi(x)$ 满足：

$\beta(\theta_0) = E[\phi(X)] = \alpha$ ，对一切 $\theta \in \Theta_0$  (6. 1. 4)

## 6.1假设检验的基本概念

### 209

则称 $\phi(x)$ 为假设检验问题(6. 1. 3)的一个水平为 $\alpha$ 的检验，并记

$\Theta_\alpha = \{ \theta : \beta(\theta) \geq \alpha, \theta \in \Theta \}$  (6. 1. 5)

若存在 $\phi(x)$ 使

$\beta(\theta) \geq \alpha$ ，对一切 $\theta \in \Theta_\alpha$  (6. 1. 6)

则称 $\phi(x)$ 为假设检验问题(6. 1. 3)的水平为 $\alpha$ 的一致最优势检验 (UMPT)。I 根据上述定义，若 $\phi(x)$ 为UMPT，则其第一类错误 $\alpha$ 小于给定水平 $\alpha$ ，第二类错误 $1 - \beta(\theta)$ 在一切水平为 $\alpha$ 的检验中最小(但一般不

可能也有 $n$  ( $<9$ )名 $a$ )。根据 $n$ - $p$ 准则进行统计推断的特点是:两类错误 $I$  (幻与 $n$ (幻不对称;原假设 $H_0$ 与对立假设 $H'$ 不对称。

(1) 若 $x \in T$ ?, 则否定原假设 $H_0: \theta = \theta_0$ , 判定对立假设 $\theta$ ,

成立. 这时理由充分, 因为这时判错属于第一类错误(即 $H_0$ 成立而否定 $A$ ), 而第一类错误 $I(\theta)$ 已控制在 $\alpha$ 的范围, 所以犯错误的概率 很小(通常为 $0.05$ 或 $0.01$ 等)。

(2) 反之, 若则不否定 $H_0: \theta = \theta_0$ 即判为打。成立. 这时理 由并不充分, 因为这时判错属于第二类错误(即 $H_0$ 成立但未否定 $H_0$ ), 若而一定很小, 所以犯错误的概率不一定很小. 例如在例6.1.1中,

出现上述情况的原因就是因为两类错误不可能同时都很小 , 而基于  $N$ - $P$ 准则的妥协方案是控制了第一类错误, 因而保护了原假设, 使其不 被轻易否定. 这时, 若一旦否定原假设(即判定 $H_0$ 成立), 则理由充分 (判错概率小于等于 $\alpha$ ). 因此在实际操作中, 应该把需要有强有力证据 的假设放在对立假设 $H_1$ 上, 一旦否定 $H_0$ , 即判定成立, 则理由充 分, 判错的概率名 $\alpha$ (水平 $\alpha$  可事先设定). 例如在假设检验问题 (6.1.1)中, 检验“药物的疗效是否提高” 至关重要, 判错的概率必须

很小, 因此应把假设 $\theta > 0.75$ 放在对立假设 $H_1$ 上. 这时, 一旦 $H_0$ 被否 定, 即 $H_1$ 成立, 判错的概率将会很小, 我们有充分的理由认为没 $\theta = 0.75$ 成立, 即 “药物的疗效确有提高”。

由以上分析可知, Neyman - Pearson准则是合理并具有实用价值的, 它已得到理论与实际工作者的一致认可. 另外, 由定义6. 1.6可知, 基 于 $N$ - $P$ 准则的一致最优势检验把假设检验问题(6. 1.3)化为一个明确的 数学最优化问题, 即对检验函数在不等式约束(6.1.4)的条件下, 求在 $\theta > 1$ 上达到最大值的解. Neyman和Pearson基于 $N$ - $P$ 准则发展

$<27$  ), 这时不否定 $H_0$ :

$\theta = 0.75$ , 但是实际上判错的 可能性不一定很小, 例如当 $\theta = 0.85$ 时还有 $P(\theta < 27) \ll 0.58$ .

210

## 第六章参数假设检验

了一套严格的数学理论, 本章6.2-6.4节将介绍其大意. 以下首先介

绍检验函数的一个重要性质. 第二章曾经指出, 充分统计量包含了与样本一样多的信息, 而且比

样本简单得多; 因此, 通常的统计推断方法都是从充分统计量出发的. 第三章的Rao - Blackwell定理说明: 统计推断中的最优解通常都是充分 统计- 的函数, 后来几章的结果也证实了这一论断. 同样, 假设检验问

题的最优解也应当是充分统计量的函数, 以下引理也印证了这一点: 引理6. 1.1 (充分性原则) 对于假设检验问题(6.1.3), 设  $T$  = 为充分统计量, 任给一个检验 $\phi(T)$ ,

为充分统计量 $r$  的函数, 并且与 $\phi(T)$ 有相同的功效.

取  $\phi = E\phi(X) | r$ , 则 $\phi$ 为 $r$ 的函数, 且与 $T$ 无关. , 所以 $\phi(T) = \phi(r)$ , 同时有

$E[\phi(T)] = E[\phi(r)]$ ,  $E[\phi(T)] = E[\phi(r)]$ . 因此检验 $\phi(r)$ 为充分统计量的函数, 并且与 $\phi$ 有相同的功效.  $I$

以上引理说明, 若包含随机化检验, 则一致最优势检验可在充分统 计量的函数中寻找.

最后我们结合例6. 1. 1介绍假设检验问题的 $P$ -值, 这是统计文献 中经常出现的一个术语 (可参见Bickel and Doksum, 1977).

在例6. 1. 1中, 假设检验问题(6. 1. 1)的否定域为 $R = \{T(x)$

$> 27\}$ , 其第一类错误为 $I(\theta = 0.75) = P(T > 27 | \theta = 0.75)$ . 今设 某次具体抽样得 $x_1, \dots, x_n$ , 及 $\bar{x} = 28$ , 这时应否定原假设 $H_0$ , 而这一次检验中犯错误的概率应为 $I(\theta = P(T > 28) | \theta = 0.75)$ ,  $0.01$ 就是本次检验的实际水平, 通常称为此次检验

的  $p$ -值。类似的，若此次抽样得  $r(x) = 26$ ，则不否定  $H_0$ ；但是若此时要否定  $H_0$ ，则犯错误的概率为  $I(\pi) = P(r > 26) \approx 0.1$ ， $(\approx 0.75)$ ，也称为此次检验的值。一般情况下，若某次具体抽样得  $r = r(x)$ ，则此次检验的值可定义为  $p = P(r \geq r(x) | H_0)$ ，它表示此次否定  $H_0$  而犯错误的概率。

对于一般的假设检验问题 (6.1.3)，若其否定域为  $C$ ，

$C$  (其他形式的否定域类似)。若某次具体抽样为  $X = x_n$ ， $T = T(x)$ ，则其  $p$ -值定义为  $\hat{p}_n = P(T(X) \geq T(x) | H_0)$  (更准确的，应表示为  $\sup_{\theta \in \Theta_0} P(T(X) \geq T(x) | \theta)$ )，见 Bickel and Doksum, 1977)，它表示根据此次抽样做检验，否定而犯错误的概率，即本次检验的实际水平。在实际应用上，给定检验水平  $\alpha$  以后，可通过值来了解某次检验的具体情况。若  $p$ -值  $\leq \alpha$ ，则否定  $H_0$ ；若  $p$ -值  $> \alpha$ ，则不否定  $H_0$ 。的理由

证明 因为  $\alpha \geq P(T(X) \geq T(x) | H_0)$

## 6.2 Neyman - Pearson 基本引理

211

越充分。反之，若  $p$ -值则不否定  $H_0$ 。； $p$ -值越大，不否定  $H_0$  的理由越充分。

## 6.2 Neyman - Pearson 基本引理

正如前节所述，根据 Neyman - Pearson 准则求解假设检验问题 (6.1.3) 的一致最优势检验，相当于求一个不等式约束下的最优化问题。这个问题并不简单，最优化理论中没有现成的方法，需要 (也应

该) 结合统计学的具体情况加以解决。同时必须注意，假设检验问题 (6.1.3) 实际上是一个很一般的问题。其中原假设和对立假设的空间  $\Theta_0$  和  $\Theta_1$  有很大的任意性；样本分布也有很大的任意性。因此，要想根据定义 6.1.6 求出假设检验问题 (6.1.3) 通用的最优解是不可能

的，只能根据参数空间  $\Theta_0$  和  $\Theta_1$  以及样本分布的具体情况，由简到繁逐步加以解决。Neyman - Pearson 首先考虑了最简单的情况，即原假设  $H_0$  和对立假设  $H_1$  都只有一个状态的简单假设情形： $H_0: \theta = \theta_0$ ， $H_1: \theta = \theta_1$ 。

$\theta_0, \theta_1 \in \Theta$ ，他们得到了最优解，这就是著名的 Neyman - Pearson 基本引理。N-P 基本引理对样本分布没有太多限制，但是若要解决复合假

设检验问题，则必须对参数空间和样本分布进一步加以限制。这时，参数空间  $\Theta$  主要限于一维，样本分布主要限于指数族分布。在这些限制下，以 Neyman - Pearson 基本引理为基础，可以进一步解决若干复合假设检验问题，其中包括单调似然比分布族的单边检验、指数族分布的双边检验及某些多参数的检验问题。这也是 Neyman - Pearson 基本引理的重大贡献所在，可以把简单假设检验问题最优解的结果逐步推广，用以解决若干复合假设检验问题，得到相应的一致最优势检验。

### 6.2.1 Neyman - Pearson 基本引理

假设  $X = (X_1, \dots, X_n)^T \sim f(x; \theta)$ ，或  $X \sim f(x; \theta)$ ，的简单假设检验问题

可表示为  $H_0: \theta = \theta_0$

或

$H_0: \theta \in \Theta_0$ ，

$H_1: \theta \in \Theta_1$ ； $H_1: \theta \in \Theta_1$ 。这两者基本等价，为方便起见，今后都采用以下参数形式：

$H_0: \theta = \theta_0$ ，根据定义 6.1.6，假设检验问题 (6.2.1) 的最优解  $\phi^*(x)$  应该满足：

(6.2.1)

212

第六章参数假设检验 且对一切  $\theta \in \Theta_1$ ， $E[\phi^*(X)] \geq \alpha$ ，有

或等价地，其功效函数应满足

$EJ(X) < a$ , 且对一切  $a$ , 有  
 “对于简单假设检验问题, 其一致最优势检验即为最优势检验 (most powerful test)”.

么 (6) 冰 (2I),

少  $a$ .

powerful test), 简记为MPT. 以下说明如何求解假设检验问题 (6. 2. 1) 的 MPT.  
 似然方法是统计学中应用最多的基本思想方法, 假设检验问题也不例外. 为了求解假设检验问题 (6. 2. 1) 的MPT, 考虑以下似然比:

首先从直观上来看, 若  $H_0$  被否定, 即  $H_1$  成立, 则应该大, 而  $f(x, \theta_0)$  小, 因而似然比  $A(x)$  应该比较大, 所以可取否定域为

而  $c$  可由水平条件  $EJ(X) = P_{\theta_0}\{X \in R\} = a$  决定. 另外, 由以下数值实例亦可了解似然比的意义:

X1234

0 0. 1 00

0. 1 AX, ④、) 0.2 0.7 0 A(x) 2/7 7/2 0

对于以上分布考虑假设检验问题 (6. 2. 1). 若抽样到  $\% = 1$ , 则  $A(\%) = 2/7$  比较小, 应该不否定  $H_0$ . 即  $H_0$  成立; 这时  $f(x, \theta_0) = 0.7$  比  $f(x, \theta_1) = 0.2$  出现的可能性大, 因而判定成立是合理的. 若抽样到  $\% = 2$ , 则  $A(\%) = 7/2$  比较大, 应该否定  $H_0$ . 即  $H_1$  成立; 这时  $f(x, \theta_1) = 0.7$  比  $f(x, \theta_0) = 0.2$  出现的可能性大, 因而判定  $H_1$  成立也是合理的.

若抽样到  $x=3$ , 则  $A(x) = 0$  很小, 因而不否定  $H_0$ . 这时  $f(x, \theta_0) = 0$  为不可能事件, 所以不否定  $H_0$  也是合理的. 若抽样到  $x=4$ , 则  $A(x)$  为无穷大, 这时  $f(x, \theta_1) = 0$  为不可能事件, 所以否定  $H_0$  也是合理的.

根据以上分析, 假设检验问题 (6. 2. 1) 的否定域应该由似然比  $A(x) = J(x, \theta_1) / f(x, \theta_0) > c$  得到. 以下给出严格的理论结果与证明.

定理 6. 2. 1 (Neyman - Pearson 基本引理) 假设  $X \sim f(x, \theta)$  为离散型或连续型密度函数. 则对简单假设检验问题 (6. 2. 1) 有

6. 2 Neyman - Pearson 基本引理

213

(1) 存在性. 设检验函数  $\phi(x)$  由似然比  $A(x)$  定义为

$\phi(x) = \begin{cases} 1 & \text{若 } A(x) > c \\ 0 & \text{若 } A(x) \leq c \end{cases}$ , 则对任意给定的  $0 < a < 1$ , 必存在  $c$  和  $\phi$  使  $E\phi(X) = a$ .

1, 必  $\phi(x) = 1$ ,

(6. 2. 2)

使  $E\phi(X) = a$

(2) 最优性或充分性. 满足 (6. 2. 2) 式的  $\phi(x)$  必为假设检验问题

(6. 2. 1) 的最优势检验.

(3) 唯一性或必要性. 以上  $\phi(x)$  在  $R^+$  上为 (6. 2. 1) 的唯一的

最优势检验, 即若  $\phi^*(x)$  也是 (6. 2. 1) 的最优势检验, 则必有  $\phi(x) = \phi^*(x)$  (a.e.),

$\phi(x) = \begin{cases} 1 & \text{若 } A(x) > c \\ 0 & \text{若 } A(x) \leq c \end{cases}$ . 并且, 若  $E\phi(X) < 1$ , 则有  $E\phi^*(X) =$

证明 (1) 存在性. 以下具体定出  $c$  与  $\phi$  使  $E\phi(X) = a$  成立. 由定义可知,  $c$  与  $y$  应满足

$E\phi(X) = E\begin{cases} 1 & \text{若 } A(X) > c \\ 0 & \text{若 } A(X) \leq c \end{cases} = P\{A(X) > c\} =$

$1 - P\{A(X) \leq c\} = 1 - F(c) = a$ .

设随机变量  $A(X)$  的分布函数为  $F(u)$ , 则上式相当于  $1 - F(c) = a$ ,

即

(6.2.3)

$F(c) = 1 - a + y[F(c) - F(c-0)]$ . 因此  $c$  就取为分布函数  $F(u)$  的  $(1-a)$  分位数, 即  $0 =$  由于似然比  $A(x)$  多  $0$ , 因而  $c = h_a > 0$ . 关于  $y$ , 可分两种情况: 若  $h_a$  为  $F(u)$  的连续点, 则有  $P_{30}(A(X)=c) = [F(c) - F(c-0)] = 0$ ,  $F(u|_a) = 1-a$ , 因而 (6.2.4) 成立 ( $y$  任意); 若  $A_a * F(u)$  的不连续点, 则  $F(u|_a)^+ =$

$\ll$ , 这时在 (6.2.4) 式中, 可取  $c = u_a$ , 且  $y$  满足 (见图 6.2. 1)

$u_a) - (1-a) \sim 0$ .

因此 (6.2.4) 式也成立, 且有  $0$  所以综合以上两种情况都有  $E\ll 0[I] = \ll$ .

(2) 最优性. 要证: 对任一  $\alpha$ , 即  $E\&a(Z) \geq \alpha$ , 都有  $E f_{ll}[小(X)] [;(X)]$ . 由于尺 +  $= jx: A(x) = f(x, 6) / f(x, 0) > c | =$

$-cf(x, 0) > 0$   $f$ , 今考虑函数

$4(x) = [(/)(\ll) - -cf(x, 0)]$ . (6. 2. 5)

(6. 2. 4)

## 214 第六章参数假设检验

当  $x \in R$  时有  $f(x, 0) - rf(x, 0) > 0$ , 又由 (6.2.2) 可知  $\phi = 1$ , 因而

$\ll$  (%)

所以有  $4(\%)^0$ . 类似地, 当  $x \in R$  时有  $f_m < 0$ ,  $(f > (x) = 0$ ,  $(/ > (x) - 各(x) 谷 0$ ; 这时亦有  $A(x)^0$ . 因而有

$f [ < / > (\%) - \$(x) ] \{ . f(x, e) - cf(x^0) \} d / f(x) = f A(x) d / z, (\%)$

$L / (1)$  如 (4) +  $] > (4)$  如 (4) =  $h A(x)^{(\phi > 0)}$ .

(6. 2. 6)

由此式可得

$L [小(x) \_多(\gamma) 1 / (\%, 没 1) 如(\gamma) \geq f^{( / > (\gamma) - 4 > (x) ] cf(x, 0) d / x(x)}$

$= c [E < ? o < / > (\%) - E d o ^ > (X) ] = c [a - J] \ll 0$ . (6. 2. 7) 因此有  $E, i [ ( / > (X) ] ^ E, > [ ^ (X) ]$ .

(3) 唯一性. 要证: 若  $(/)*(\phi)$  也是假设检验问题 (6.2.1) 的最优势 检验, 则必有  $<(\phi = < / > (\%)$ ,  $x \in R + UR$ . 首先, 若  $(/)*(\%)$  亦为最优势 检验, 必有  $EW(X) > E f i i < / > (\%)$ , 反之亦有  $E <, i (/)(X) ^ E f l [ (/)* (X) ]$ , 因而  $E, [ < r a ] = e \sim [必 a]$ . 而在第 (2) 部分的证明中, 取在 (4) ( $\phi$ ), 则上面所有的推导都成立, 因而由 (6. 2.6) 式有

$f [ < / > (\%) - 冷 \cdot " ) ] [ / (x, R) - ^ / ^, 没. ) ] d / x (龙)$

$[ ( / > (\%) - 4 > * (x) ] [ f(x, 0) - cf(x, 0) ] d / i(x) > 0$ .

(6. 2. 8) 该式不等式不能成立, 因为若不等式成立, 则在 (6.2.7) 式中取多 (1) =

$<(\phi$  有

$\phi](X)]$ , 这与  $E t f i [ ( / > * (X) ] = E, J 4) (X) ]$  矛

6. 2 Neyman - Pearson 基本引理 盾. 因此 (6. 2.8) 式等式成立, 即

215

因而有

$I$

$+ U / ? -$

$/ (x)$

$- (/ (X, 没. ) ] 如(\%) = 0$ .

$- (/ > * (\%) ] ) - < / (x, ^0) ] = 0 (a. e. )$ ,  $x \in R +$ .

而在  $/ ? + u / r$  上,  $L / U, A) - < / (\% A) ] \ll 0$ , 因此必有  $( / > (x) - = 0 (a. e. )$ ,  $\% \in ? ? + \setminus J R \sim$ .

$-(f)(x) (a. e. )$ ,  $x \in R + \setminus J R \sim$ . 最后再证, 若  $E 0 j ( / > * (X) < 1$ , 即  $< / > \cdot (\phi =$

则必有  $E, o < / > * (X) = a$ . 由于 (6. 2.8) 式等式成立, 因此有  $E 0 ^ (X) -$

$EJ^*(X) = c[EJ(J) - EJ^*(X)]$ , 由  $(X)] = E, i[</(X)]i$  可得  
 $(\text{幻}-EW(\text{幻}) = c \setminus_a_w(X)] = 0$ . 其中  $c$  不可能为  $0$ , 因为若  $c=0$ , 则  $7?$   
 $+ = jA(x) > c = 0( = ) >$

$0$   $i$ , 由此可得

$[(/ > * (X) ] = )d^{(x)}$   
 $= f + </ > * (x) / (\%, ^!) d / . t(\%) = f + 1 \cdot f(x, e\}) dfJL(X) = 1.$

这与假设  $<1$  矛盾, 因此  $40$ , 从而  $(X)] = 0$ , 即  $E, n[<(X)] = a. |$

注 (i) 唯一性是指在  $x e / r u z r$  上, 恒有  $(/ > \cdot (\%) = \text{小}(\%)$ , 但在  $R^\circ$  上不具有唯一性.  
 即假设检验问题 (6. 2. 1) 可以有不同的最优 势检验, 它们在  $R^\circ$  上可以不相等, 但是在  
 $7TU7T$  上必须相等. 另外, 在许多连续型分布的情形,  $\#$  上测度为零, 即  $/\% . (7?^\circ) = 0$ , 这时  
 上的值可以忽略不计, 因而假设检验问题 (6. 2. 1) 有唯一的, 非随机化的  
 最优势检验 (见后面例题)

$( \text{、} - C1, = ]x:A(x) > c I , .0, x e . R \sim = |\% : A(x) < c |$

(2) 关于零测集以及  $A''$  ) 取值为  $0$  或  $+ \infty$  等特殊情况的说明. 记  $= 1 - \wedge : A x . 0 j = 0 | , A \}$   
 $= |\% : = 0 |$ , 则  $A(\text{幻})$  的取值可以有

以下三种情形, 这三种情形与  $N-P$  引理的结论都是一致的.

i) 若  $x e / 1]$ , 但  $x e 40$ , 则有  $/(\%, R) = 0$ ,  $/(\%, <\%) \neq 0$ , 说明 "  $|$  不 成立,  $H_0$  可能成立;  
 而此时  $A(x) = 0$ , 因而  $x e f A(x) > c$ ; , 则根据

$N-P$  引理应该不否定  $7/$  . , 这与上述结果相吻合.

ii) 若  $x e X, , -$  但  $x e 40$ , 即  $/ (X, \text{久}) \neq 0$ ,  $/ (龙, \text{没} 0) = 0$ , 因而  $\%$  不成

## 216 第六章参数假设检验

立, 可能成立; 而此时  $A(x) = +\infty$ , 因  $f f i j I A(\wedge) > c \setminus$  则 根据  $N-P$  引理应该否  
 定  $//$  . , 这也与上述结果相吻合.

iii) 若  $\%, e$  , 则可取  $A(x) = 0$ , 因而  $x e . R = ] A(x) > c |$  , 则根据  $N-P$  引理不否定  
 仇, 这也是可以接受的.

以下介绍  $N-P$  引理的若干重要推论.

推论1 对任意统计量  $T=T(X)$ , 定义函数  $</ >(\text{幻})$  如下:

$1, (b(x) = . y, .0,$   
 $T(x) > c, \cdot T(x) = c,$   
 $T(x) < c,$

则对任意给定的  $0 < a < 1$ , 必存在和  $0 \leq 1$ , 使  $E . J c / i C X ] = \ll$ . 在以上存在性的证明中, 把  
 $A(\text{幻})$  换成  $7(\%)$ , 则前面的一切推导都  
 成立.

推论2 若  $r = r(x)$  为充分统计量, 且  $. r \sim gG, 0$ ), 则上述定理中  
 的最优势检验可表示为

$1, t e R^* - j f : A(z) > c | , </ > (0 = \blacksquare y, t R^\circ = (z : A(0 = c j , .0, t e R \sim = |$   
 $\wedge : A(0 < c),$

其中  $A(1) = g(n) / gG, \text{么})$  且  $c, y$  由  $E, 0[<Hn]$  可表示为充分统计量的函数.

因为由因子分解定理可得

决定. 即  $A(x)$

$A, . ) = g(\text{心}), \text{仏}) \text{九}(\%) \_g(《, \text{久}) \times f(x, 0a) g(t(x) f60) h(x) g(Mo)'$

推论3 若似然比  $A(\text{幻})$  为某一个统计量  $7 \setminus \text{幻}$  的严增函数:  $A(x) = h(T(x)) \geq h(*)$  严增, 则  $</ >$   
 $>(\text{幻})$  可表示为

$\cdot 1, \%, e 7? + = |\% : r(x) > k \}$ ,

必  $(工) = \text{卜}, x e / ?^\circ = \setminus x : T(x) = k ) , (6. 2.9)$

$0, x \in R \sim = \{ \omega: T(\omega) < k \}$ , 其中  $A: T$  由  $\{ \omega(r) \} = a$  决定.

因为  $\{ \omega: T(\omega) > c \}$  可表示为  $R_+ = \{ \omega: A(x) > c \} = \{ \omega: h(T(x)) > c \} = \{ \omega: T(x) > k \}$ , 其中  $k = h^{-1}(c)$ . | 定理 6.2.2 (无偏性) 若  $H_0$  为假设检验问题 (6.2.1) 的最优势

检验, 则对  $0 < \alpha < 1$  有  $E_{H_0}[I(X)]$  并且, 若  $f(x, \theta) \neq f(x, \theta_0) \neq 0$ , 则  $E_{H_0}[I(X)] = \alpha$ .

证明 取  $\phi(x)$  为  $\theta_0$ , 则  $\phi(x) \in \mathcal{C}$ , 因为  $E_{H_0}[\phi(x)] = \alpha$  因而由必  $\phi(x)$  的

## 6.2 Neyman - Pearson 基本引理

217

最优性知  $E_{H_0}[I(X)] = \alpha$ . 另外, 若  $M$

#

$> 0$ , 则以上不等式等号不成立, 因为若有  $E_{H_0}[I(X)] = \alpha$ ,

则由于  $E_{H_0}[I(X)] = \alpha$ , 因而  $E_{H_0}[I(X) - \alpha]$  也成立, 这说明

$\phi(x) = \alpha$  也是最优优势检验. 因此由唯一性有  $\phi(x) = \alpha$ ,  $\phi(x) \in \mathcal{C}$

U/T. 但是  $0 < \alpha < 1$ , 而似然在  $T$  上为 1, 在  $T$  上为 0, 都不等于

$\alpha$ , 所以  $R_+ \cap R'$  只能是空集, 因而  $R'$  为全空间, 即  $R' = U: \phi(x) = \alpha$

$f(x, \theta_0) = c \}$ . 由此可得  $f(x, \theta_0) = cf(x, \theta_0)$ , 由于  $f(x, \theta_0)$  和  $\phi(x)$  都是密度函数, 因此必有  $c=1$ , 即  $\phi(x) = 1$  当  $x \in R_+$  时, 因此  $\phi(x, \theta_0) = f(x, \theta_0) = 1$ , 这与假设条件  $\phi(x) \neq 1$  矛盾, 所以必有  $E_{H_0}[I(X)] > \alpha$ . | 满足条件  $E_{H_0}[I(X)] = \alpha$  的检验常称为

无偏检验. 易见无偏检验的第二类错误  $1 - E_{H_1}[I(X)] = 1 - \alpha$ , 这对一个好的检验

显然是很起码的要求 (例如一个无偏检验,  $1 - \alpha = 0.05$ , 则其  $D(\phi) \geq 0.95$ ). 本章第 4 节主要讨论无偏检验.

## 6.2.2 Neyman - Pearson 基本引理应用示例

以下举例说明, 如何应用 Neyman - Pearson 基本引理解决假设检验问题. 同时, 由这些例题的解决过程可知, N-P 引理不但可以解决简单假设检验问题, 而且也可以用于解决若干复合假设检验问题.

, 由定理 6.2.1 可知, 求最优优势检验, 主要是确定  $c, y$ , 使之满足 (6.2.3) 式, 即

$\phi(x) = I(A(x) > c) + y I(A(x) = c)$ ,  $0 < \alpha < 1$ . (6.2.10) 如果  $A(X)$  为连续型分布, 则问题化为求

$c$ , 使  $P(A(X) > c) = \alpha$ .

例 6.2.1 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim P(\lambda)$ , 考虑假设检验问题:

(i)  $H_0: \lambda = \lambda_0$  vs  $H_1: \lambda = \lambda_1$ , 但  $\lambda_1 > \lambda_0$ ; (ii)  $H_0: \lambda = \lambda_0$  vs  $H_1: \lambda > \lambda_0$ .

求 (i) 的 MPT, (ii) 的 UMP.

解 对于检验问题 (i), 其似然比表示为

$L(\lambda_1/\lambda_0)$

$= \prod_{i=1}^n \frac{\lambda_1^{x_i} e^{-\lambda_1}}{\lambda_0^{x_i} e^{-\lambda_0}}$

218

## 第六章参数假设检验

在  $\lambda_1 > \lambda_0$  的前提下,  $L(\lambda_1/\lambda_0) = \prod_{i=1}^n \frac{\lambda_1^{x_i} e^{-\lambda_1}}{\lambda_0^{x_i} e^{-\lambda_0}}$  的 (6.2.9) 式有

的严格增函数, 因此由推论 3

$R_+ = \{ \omega: A(x) > c \} = \{ \omega: T(x) > k \}$ , 其中  $T(X) \sim P(n\lambda)$ . 根据 N-P 引理, 要求  $c, y$ , 使之满足

$P_{\lambda_0}(A(X) > c) + y P_{\lambda_0}(A(X) = c) = P_{\lambda_0}(T > k) + y P_{\lambda_0}(T = k) = \alpha$ . 在  $H_0$  成立的条件下,  $T$



$\sim F(nA_0)$  为离散型分布, 可记其分布函数为  $F_0(u)$ , 它可表示为不完全  $\Gamma$  函数 (见第一章).

设其  $1-a$  分位点为  $u_a$ , 则有

$$1 - F_0(u_a) = [F_0(1) - F_0(k-0)] = \llbracket F_0(1) - F_0(A-0) \rrbracket.$$

$$\text{可取 } k_0 = u_{1-a}, y_0 = [F_0(u_{1-a}) - (1-a)] / [F_0(u_{1-a}) - F_0(u_{1-a}-0)],$$

则假设检验问题 (i) 的 MPT 为

$$r_1, \quad \phi(x) = \begin{cases} 1 & \text{若 } T(x) > k_0 \\ 0 & \text{若 } T(x) \leq k_0 \end{cases}$$

$$I_0,$$

,

$$T(x) > k_0,$$

$$T(x) = k_0, \quad (6.2.11)$$

$$T(x) < k_0.$$

可以证明, 这个  $\phi(x)$  也是假设检验问题 (H) 的 UMPT. 事实上, 由 (6.2.11) 式可知,  $\phi(x)$  的取值仅与  $A_0$  有关, 与  $\lambda$  无关, 因此对任意的  $A > A_0$ ,  $\phi(x)$  都是 (1) 的 MPT. 这表明, 对任意的  $A > A_0$ , 只要  $E_j(X)$  都有  $E_4(X) \leq E_{A_i}(X)$ . 因此由定义 6.1.6 可知,

$\phi(x)$  就是假设检验问题 (ii) 的 UMPT.

例 6.2.2 设  $X_1, \dots, X_n$  独立同分布,  $X$ ,

验问题水平为  $\alpha$  ( $0 < \alpha < 1$ ) 的 MPT:

$$h_0: \theta = \theta_0 \text{ 对 } \theta_0 < \theta_1.$$

解分布族的密度函数和似然比为

$$X(1) \sim \theta_1 \quad I'_{jX}(n)$$

' | 求以下假设检

,

$$Z \sim N(0, 1)$$

$$T = \sum_{j=1}^n X_j^2 \sim \chi^2(n)$$

$$T \sim \chi^2(n)$$

$$T \sim \chi^2(n)$$

$$T \sim \chi^2(n)$$

$$T \sim \chi^2(n)$$

$$T \sim \chi^2(n)$$

注意, 虽然  $X$  有连续型分布, 但是  $A(X)$  的分布为离散型, 仅取值  $0, 1$ .

## 6.2 Neyman - Pearson 基本引理

219

$e = (e_0/\theta)$  和  $\theta$  三个值. 以下求  $c$  和  $y$ , 使

$$E_{\theta_0}[\phi(X)] = P_{\theta_0}(X > c) + y P_{\theta_0}(A(J) = c) = \alpha. \quad (6.2.12)$$

以下证明: 当  $0 < \alpha < 1$  时,  $c$  只能取  $A$

(1) 若  $c < A$ , 则有  $P_{\theta_0}(X > c) = 1$ . 因为此时  $X \sim \chi^2(n)$ ,  $P_{\theta_0}(X > c) = 1$ . 因此  $P_{\theta_0}(X > c) = 1$ , 不可能为  $\alpha$ .

(2) 若  $c > A$ , 则有  $P_{\theta_0}(X > c) = 0$ . 因为此时  $X \sim \chi^2(n)$ ,  $P_{\theta_0}(X > c) = 0$ . 因此  $P_{\theta_0}(X > c) = 0$ , 不可能为  $\alpha$ .

因此, 为使 (6.2.12) 式成立, 只能取  $c = A$ . 这时  $P_{\theta_0}(X > c) = P_{\theta_0}(X > A) = 0$ .

$$E_{\theta_0}[\phi(X)] = P_{\theta_0}(X > A) + y P_{\theta_0}(A(J) = A) = 0 + y \cdot 1 = \alpha.$$

代人 (6.2.12) 式有  $E_{\theta_0}[\phi(X)] = 0 + y \cdot 1 = \alpha$ , 所以  $y = \alpha$ , 由此可得最优检验为

$$\phi(x) = \begin{cases} 1 & \text{若 } x > A \\ \alpha & \text{若 } x = A \\ 0 & \text{若 } x < A \end{cases}$$

$\phi(x) = \int_a^x f(t) dt$ ,  $A(x) = \int_a^x f(t) dt$ ,  $0 \leq x \leq 1$ ,  $\phi(1) = 1$

$b$ ,  $A(x)$  或其他.

例6. 2. 3 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\mu, \sigma^2)$ ,  $\sigma^2$  已知. 考虑假设检验问题:

(i)  $H_0: \mu = \mu_0$ ,

(ii)  $H_0: \mu \geq \mu_0$  (iii)  $H_0: \mu \leq \mu_0$

$\mu = \mu_0$ , 但是  $\mu > \mu_0$ ;  $\mu < \mu_0$  ;

$\mu > \mu_0$ .

根据N-P引理: (1) 求(i)的MPT; (2) 求(i)的功效函数  $\phi(M)$ , 并证

明它为  $A$  的增函数; (3) 证明(i)的MPT就是(ii)的UMPT; (4) 证明 (ii)的 UMPT 就是

(iii)的 UMPT.

解 (1) 对于假设检验问题 (i), 其似然比为

$$\frac{L_1}{L_0} = \frac{\exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu_1)^2\right\}}{\exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu_0)^2\right\}}$$

$$A(x) = \frac{\int_{-\infty}^{\infty} \frac{L_1}{L_0} dF(x)}{\int_{-\infty}^{\infty} \frac{L_1}{L_0} dF(x)}$$

在  $A = M_0$  的前提条件下,  $A(x) = T(x) = \frac{\sum_{i=1}^n x_i}{n}$  的严增函数, 因  $i=1$

此有

eXP

220

## 第六章参数假设检验

$$A(x) = \frac{\int_{-\infty}^{\infty} \frac{L_1}{L_0} dF(x)}{\int_{-\infty}^{\infty} \frac{L_1}{L_0} dF(x)}$$

其中  $A$  由水平条件决定. 根据(6.2.10)式, 应求  $h$  使

$$x > k, x < k$$

且有  $\phi(\mu_0) = \alpha$  (注意  $\mu_0$  又  $\mu_0 = 0$ ).

可由  $X$  的分布直接求

$$\mu = \mu_0$$

出,

由于, 因此有

$$1,$$

$$\phi(\mu)$$

$$0,$$

$$I - \text{从 } 0 \text{ 到 } 0$$

因而  $\phi(\mu) = \int_{-\infty}^{\infty} \frac{L_1}{L_0} dF(x)$  等于标准正态分布  $\phi(x)$  的  $(1 - \alpha)$  分位数  $z_{1-\alpha}$ , 为

已知, 因此有  $\mu_0$  最后可得假设检验问题 (i) 的 MPT 为

$$\mu_0 \text{ 即 } [1, x > k] \text{ 或 } [0, x < k] \text{ (6.2.13) } x < k \text{ 即 } \mu_0 = \mu_0 + \frac{1}{\sqrt{n}} z_{1-\alpha}$$

yn

例如, 若  $\alpha = 0.05$ , 则  $\mu_0 = \mu_0 + 1.65 \left( \frac{\sigma}{\sqrt{n}} \right)$ ; 若  $\alpha = 0.01$ , 则

$\mu_0 = \mu_0 + 2.33 \left( \frac{\sigma}{\sqrt{n}} \right)$ . 其统计意义为: 当样本均值  $\bar{X} > k$  即离  $\mu_0$  较远时, 则否定

$H_0: \mu = \mu_0$  且水平条件  $\alpha$  越小,  $\mu_0$  越大.

(2) 根据定义, 功效函数可表示为

$$\phi(\mu) = E\left[\frac{L_1}{L_0}\right] = P\left(\frac{\sum_{i=1}^n X_i}{n} > k\right) = P\left(\bar{X} > k\right) = P\left(\bar{X} - \mu_0 > k - \mu_0\right)$$

$$= P\left(\frac{\sum_{i=1}^n (X_i - \mu_0)}{n} > k - \mu_0\right)$$

此处用到关系式  $1 - \phi(\mu_0) = \phi(-\mu_0)$ . 由上式可知  $\phi(\mu)$  为  $\mu$  的增函数, 且有

$$\phi(\mu) < \phi(\mu_0),$$

$$\phi(\mu) > \phi(\mu_0), \quad (6.2.15)$$

$\mu > \mu_0$ . 我们来看检验的第二类错误. 由上式可知,  $\phi(\mu) > \phi(\mu_0)$  时有  $\phi(\mu) > \alpha$ , 因而  $1 - \phi(\mu) < 1 - \alpha$ , 这与定理6.2.2的结论一致. 再看

$$(6.2.14)$$

## 6.2 Neyman - Pearson 基本引理 221

一个更具体的数值结果, 考虑 $\theta = 0.5$ , 并设  $a = 1$ ,  $n=25$ , 则有  $\beta = (P(0.85) - 0.802, H(\theta)) = 1$

0.198. 该值比0.05大, 但还不算太大. 另外, 由(6.2.14)式可知, 当  $\theta \rightarrow +\infty$  或  $a \rightarrow +\infty$  时有  $(1/\alpha) \rightarrow 1$ , 因而  $n(\alpha) \rightarrow 0$ , 即  $n$  (没) 还是可以充分小的.

(3) 可以证明, 由(6.2.13)式确定的  $\phi(\theta)$  也是假设检验问题(ii)的UMPT, 这在例6.2.1中已有类似的论述. 事实上, 由(6.2.13)式可知,  $\phi(\theta)$  的取值仅与  $t(\theta)$  有关, 与  $M_1$  无关, 因此对任意的  $M_1 \in M_0$ ,  $E_M$

都是(i)的MPT. 这表明, 对任意的  $\phi(x)$  和  $\theta > \theta_0$ ,  $\phi(y)$  小 (幻) 就是假设检验问题(ii)的UMPT.

$\int \phi(x) dP$  都有  $E_M$ )

(4) 以下证明,  $\phi(\theta)$  也是假设检验问题(出)的UMPT. 根据定义 6.1.6, 需要证明两点: 第一,  $\phi(\theta)$  满足水平条件(6.1.4), 即  $E_{\theta_0}(\phi(X)) \leq \alpha$ ,  $\forall \theta \geq \theta_0$ , 这可由(6.2.15)得到; 第二,  $\phi(\theta)$  满足最优性条件(6.1.6), 即

对任意的以及满足水平条件的每 (幻) 都成立, 这在前面实际上已经证明过, 若每 (幻) 满足  $E_{\theta_0}(\phi(X)) = \alpha$ , 则也有

$E_{\theta_0}(\phi(X)) = \alpha$  因此根据(3), 对任意的  $\theta > \theta_0$  都有  $E_{\theta_0}(\phi(X)) = \alpha$ . 所以  $\phi(x)$  也是(iii)的UMPT. |

类似地, 设  $\theta_1, \dots$ , 又独立同分布,  $X' \sim Kf_{JL}(\theta_1)$ ,  $d$  已知. 我们 也可以求出以下假设检验问题的MPT以及UMPT (进一步的讨论见下一节):

(1) (ii)

诉我们如何应用Neyman-Pearson基本引理来解决某些复合假设检验问题.

则有

(in)

$\theta_0: \theta \leq \theta_0$

$\theta_1: \theta > \theta_0$  但是  $\theta_0 < \theta_1$ ;  $\theta_1 < \theta_0$ ;

$\theta_1: \theta \leq \theta_0$ . 以上例题中求解UMPT的方法具有一般性, 可总结为以下引理, 它告

引理6.2.1

考虑以下假设检验问题:

$\theta_0 = \theta_0$ , 但  $\theta_1 \neq \theta_0$ ,  $\theta_1 > \theta_0$  ; ;

$\theta_1: \theta \leq \theta_0$  ; ;

(i) (ii) (iii)

$\theta_0$  g

(1) 若  $\phi(\theta)$  为(i)的MPT, 但  $\phi(\theta)$  与  $\theta$  无关, 则  $\phi(\theta)$  为(ii)的 UMPT;

(2) 若  $\phi(\theta)$  为(ii)的UMPT, 且又为(iii)的水平为  $\alpha$  的检验, 则  $\phi(\theta)$  亦为(Hi)的UMPT. ■

, 只要

因此由定义 6.1.6 可知

222

## 第六章参数假设检验

### 6.3 单调似然比分布族的单边检验

Neyman - Pearson基本引理是关于简单假设检验问题最优解的定理, 要用它来解决复合假设检验问题, 必然受到诸多限制. 本节主要讨论单 参数的单边假设检验问题, 即假定

$H_0$ . 然后逐步讨论以  $H_0$  假设检验问题:

- (i)  $\mu = \mu_0$ , 但  $\sigma^2 > \sigma_0^2$ ; (ii)  $\mu > \mu_0$ ; (iii)  $\sigma^2 > \sigma_0^2$ .

假设检验问题 (iii) 是比较常见的情形, 通常称为单边假设检验问题, 我们的最终目标是要得到它的 UMP 检验, 其解决方法与前一节的例题很类似, 但是在分布族等方面有一定的推广, 所以首先介绍单调似然比分布族.

### 6.3.1 单调似然比分布族单边检验的 UMP 检验

定义 6.3.1 设  $X \sim f(x; \theta)$ ,  $\theta$  为单参数, 若

- (1) 当  $\theta_1 > \theta_2$  时,  $f(x; \theta_1) / f(x; \theta_2) > 1$ ; (2) 当  $\theta_1 < \theta_2$  时, 似然比

单调似然比分布族, 简称为 MLR 分布族. ■

注 以上定义的特点是只要求  $A(\theta)$  为  $\theta$  的非减函数, 不一定是严格增函数 (在定理 6.2.1 的推论 3 中要求  $A(\theta)$  为  $\theta$  的严格增函数), 因而 MLR 包含的分布族更广泛一些, 见下面例 6.3.3.

例 6.3.1 设  $X = (X_1, \dots, X_n)^T$ , 其分量独立同分布, 若  $X_i$  服从  $N(\mu, \sigma^2)$  或  $V(\mu, \sigma^2)$  或  $P(1)$ , 则它们都是 MLR 分布族.

解 以两点分布为例,

$$f(x; \theta) = p^x (1-p)^{n-x}, \quad x = 0, 1, \dots, n$$

当  $\theta_1 > \theta_2$  时, 似然比为

$$= \left( \frac{p_1}{p_2} \right)^x \left( \frac{1-p_1}{1-p_2} \right)^{n-x}$$

这是  $x$  的增函数, 因此为 MLR 分布族, 其他类似, 从略.

$f(x; \theta_1) / f(x; \theta_2) > 1$   
 $A(x) = x$  为  $T(x) = x$  的非减函数, 则称分布族  $\{f(x; \theta) : \theta \in \Theta\}$  关于  $T(x)$  为 MLR 分布族.

### 6.3 单调似然比分布族的单边检验

223

例 6.3.2 指数族分布. 设  $f(x; \theta) = \frac{1}{\Gamma(\theta)} x^{\theta-1} e^{-x}$ , 其中  $\theta > 0$  为  $\theta$  的增函数, 则为 MLR 分布族.

解 当  $\theta_1 > \theta_2$  时, 似然比为

$$A(\theta) = \frac{f(x; \theta_1)}{f(x; \theta_2)} = \frac{\Gamma(\theta_2)}{\Gamma(\theta_1)} x^{\theta_1 - \theta_2} e^{-(\theta_1 - \theta_2)x}$$

这是  $x$  的增函数, 因此为 MLR 分布族.

例 6.3.3 设  $X = (X_1, \dots, X_n)^T$ , 其分量独立同分布,  $X_i \sim U(0, \theta)$ , 没  $\theta$ , 则  $X$  的分布关于  $T(x) = x(n)$  为 MLR 分布族. 解 当  $\theta_1 > \theta_2$  时, 其分布密度和似然比分别为

$$f(x; \theta) = \frac{n!}{(n-1)! \theta^n} x^{n-1} e^{-x/\theta}, \quad 0 \leq x \leq \theta$$

$$f(x; \theta_1) / f(x; \theta_2) = \left( \frac{\theta_2}{\theta_1} \right)^n \left( \frac{x}{\theta_1} \right)^{n-1} e^{-x(\frac{1}{\theta_1} - \frac{1}{\theta_2})}$$

$$= \left( \frac{\theta_2}{\theta_1} \right)^n \left( \frac{x}{\theta_1} \right)^{n-1} e^{-x(\frac{1}{\theta_1} - \frac{1}{\theta_2})}$$

;

其中

其中

### 引理 6.3.1

对于 MLR 分布族

$$f(x; \theta) = \frac{1}{\Gamma(\theta)} x^{\theta-1} e^{-x/\theta}, \quad 0 \leq x \leq \theta$$

,

而  $A(\theta) = U(x(n))$  为  $\theta$  的严格增函数 (一次函数), 而且  $U(x(n))$  为  $x(n)$  的非减函数, 因此  $A(\theta)$  为  $x(n)$  的非减函数, 所以  $X$  的分布关于  $T(x) = x(n)$  为 MLR 分布族. ■

以下针对 MLR 分布族逐步求解假设检验问题 (i) - (iii) 的 MPT 和 UMP.

例 6.3.4 设  $X \sim U(0, \theta)$ , 检验  $H_0: \theta = \theta_0$  对  $H_1: \theta > \theta_0$ .

定义函数 $\phi(x)$ 如下:  $\phi(x) = 1, H(x) > h(k)$ ,

$\phi(x) = 0, T(x) < k$ .

(6.3.1)

则对任意给定的  $0 < \alpha < 1$ , 必存在  $A$ : 和  $\gamma$ , 使  $E[\phi(X)] = \alpha$ .

证明 这就是定理6.2.1的推论1具体应用到MLR分布族的  $H(\cdot)$ .

以下将证明, 上述引理6.3.1中的  $\phi(x)$  (幻即为假设检验问题(i)的MPT和(ii)的UMPT).

为此, 我们首先从直观上说明MLR分布族与N-P引理之间的联系.

易见, 若  $A(x) = h(T(x))$  为  $T$  的严增函数, 则由定理6.2.1的推论3直接可知,  $\phi(x)$  (幻即为假设检验问题(i)的MPT, 且有  $T(x) > I = !A(x) = h(T(x)) > h(k) = c$  (如果  $A(T(x))$  为  $TV$  的非减函数 (这是MLR分布族的特点), 这时情况要复杂一点. 恒是有以下关系

## 224 第六章参数假设检验

(见图 6.3.1):

$\{T(x) > k \mid 3(A(x) - h(T(x)) > h(k) = c\} \subset R^+$ . (6.3.2)

同理也有

$\{T(x) < k \mid 3(A(x) = h(T(x)) < h(k) = c\} \subset R^-$ .

因此, 由引理6.3.1确定的  $\phi(x)$  与由N-P引理确定的最优势检验

必 (幻相比, 它们可能在  $T$  上取相同的值, 因此由唯一性可知

(幻也是假设检验问题(i)的MPT. 以下给出严格证明.

定理 6.3.1 设  $X \sim$

$\theta \in R^1$  为关于  $T = m$  的MLR分布族, 则由(6.3.1)式确定的  $\phi(x)$  为假设检验问题(1)的MPT, (ii)的UMPT.

证明 在(幻中取  $h(k) = c$ ,

则由于  $W$  为非减函数可知 (见图 6.3.1)

$A(x) = h(T(x))$

$c^{h(k)}$

0

图6.3.1 单调似然比分布族  $\{x: T(x) > k \mid jx; T(a; ) > k, A(x) > c\}$

$U\{x: T(x) < k, X(x) = c\}$

同理

$\{x: T(x) < k\} = \{x: T(x) < \gamma, A(x) < c \mid \gamma: T(x) < k, X(x) = c\}$ .

因而可重新表示为

$\phi(x) = 1,$

$0, \phi(x) = 0,$

该式可进一步表示为与N-P引理相吻合的形式:

$\phi(x)$  其中  $y(x)$  (幻为一段函数  $y(x) = 0$ ) =

1,

$T(x) > k, A(x) > c, T(x) > k, A(x)$

$T(x) = \gamma, A(x) = c, T(x) < \gamma, A(x) = c, T(x) < k, A(x) < c.$

$y,$

$A(x) > c,$

■ 1,

$\gamma^{h(k)}, A(x) = c,$

$0, A(x) < c,$

$T(x) > k, A(x) = c,$

(6.3.3)

$T(x) = k, A(x) = c, 0, T(x) < k, A(x) = c.$

因此, 由(6.3.3)式定义的  $\phi(x)$  符合N-P引理中最优势检验的条件 (见(6.2.2)式), 且有  $E[\phi(X)] = \alpha$ , 由唯一性可知  $\phi$  亦为假

### 6.3 单调似然比分布族的单边检验

225

设检验问题 (i) 的 MPT. 又由于  $\phi(x)$  与  $\theta$  无关, 因而由引理 6.2.1 可

知,  $\phi(x)$  也是 (ii) 的 UMPT. I

推论 任给  $\alpha$ , 设  $E[\phi(X)] = \alpha$ , 则  $\phi$  为以下检验 (i) 的水平为  $\alpha$  的 MPT, 检验 (ii) 的水平为  $\alpha$  的 UMPT.

(i)  $H_0: \theta = \theta_0, H_1: \theta = \theta_1$  但

(ii)  $H_0: \theta = \theta_0, H_1: \theta > \theta_0$ . I 这个推论显然是成立的, 因为只需把前面的  $\theta$  看成  $V$  即可. 但是从观念上来看, 这一结果还是有新意的. 在前面的讨论中, 水平  $\alpha$  往往与第一类错误联系在一起, 通常要求很小. 但是从数学推导上来看, 对  $\alpha$  并无此要求, 只要满足  $0 < \alpha < 1$  即可. 这将给今后的数学推导带来较多的方便. 我们后面将根据需要取甚至  $\alpha = 1 - \alpha$  等不同的值. 下面我们要继续证明,  $\phi$  也是单边假设检验问题 (iii) 的 UMPT. 为此首先要求出  $\phi(x)$  的功效函数, 其思路与前一节的例

6.2.3 十分类似.

定理 6.3.2 设为 (6.3.1) 式确定的检验, 其功效函数记为

$\phi$

(1)  $\phi(\theta)$  为  $\theta$  的非减函数, 且在  $\theta_0 < \theta < \theta_1$  上为  $\theta$  的严

增函数;

(2) (1) 为假设检验问题 (iii) 的 UMPT.

证明 (I) 今证明, 对任意的  $H$  有  $\phi(R) \geq \phi(\theta)$ , 即  $E[\phi(X)] \geq E[H(X)]$ , 为此考虑以下假设检验问题:

(1)  $H_0: \theta = \theta_0, H_1: \theta = \theta_1$  但  $\theta_1 > \theta_0$ .

记  $E[\phi(X)] = \alpha$ , 则由定理 6.3.1 的推论可知,  $\phi(x)$  为检验 (i) 的水平为  $\alpha$  的 MPT. 再由定理 6.2.2 关于 MPT 的无偏性可知

则有

$E[\phi(X)] = E[\phi(X)]$ ,

(A)  $\phi(\theta) > \alpha$ . 另外, 根据 MLR 分布族的定义, 当

即  $\theta = \theta_0$  时有  $\phi(x) > \alpha$ , 因此由无偏性定理可知, 若  $\alpha <$

$\phi(\theta_0) < 1$ , 则必有  $E[\phi(X)] > \alpha$ , 即  $\phi$  为  $\alpha$  的严增函数. (2) 定理 6.3.1 已经证明,  $\phi(x)$  是假设检验问题 (ii) 的 UMPT, 根据引理 6.2.1, 只要证明为假设检验问题 (iii) 的水平为  $\alpha$  的检验即可, 即要证明  $0 \leq \theta < \theta_0$  时  $\phi(\theta) \geq \alpha$ . 由于  $\phi$  为  $\theta$  的非减函数, 因此  $0 \leq \theta < \theta_0$  时,  $\phi(\theta) \geq \phi(\theta_0) = \alpha$ , 即  $\phi$  为  $\alpha$  的严增函数.

所以  $\phi$  为 (iii) 的水平为  $\alpha$  的检验, 因而由引理 6.2.1 可知  $\phi(x)$  为 (iii) 的 UMPT. |

226 第六章参数假设检验 例 6.3.4 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\theta, \sigma^2)$ ,  $\theta >$

$\theta_0$ , 求以下

6.3.2 可知以上假设检验问题的 UMPT 为

假设检验问题的 UMPT 及其功效函数

解 由例 6.3.3 可知, 分布族关于  $T = \sum_{i=1}^n X_i$  为 MLR, 所以由定理

1,  $(n) > \theta_0$ , 心  $\phi$ , 其中  $\alpha$  由  $E[\phi(X)] = \alpha$  决定. 注意

所以  $\phi$  应满足

$X(n) \sim N(\theta, \sigma^2)$ ,

$$=P_0(X(n)>k)=1-\alpha$$

$$h_0(X)=$$

:

$$0-H_1: \theta > \theta_0.$$

由此可得  $k = \theta_0$ . 其功效函数为

$$h_1(X) = E_{H_1}(I) : P(X, \dots, > k_0) = -(1-\alpha).$$

易见功效函数  $K(\theta)$  是  $\theta$  的增函数. 例 6.3.5 设  $\dots, X_n$  独立同分布, 指数分布  $r(1/\theta, 1)$ ,  $\theta > 0$ .

(1) 求假设检验问题  $H_0: \theta \leq \theta_0$  对  $H_1: \theta > \theta_0$  的 UMP 检验;

(2) 若仅观察到前  $r$  个样本:  $X(1), X(2), \dots, X(r)$ , 由此求上述假设检验问题的 UMP 检验;

(3) 某电子元件的寿命服从指数分布  $r(1/6, 1)$ , 要求其平均寿命不小于 1000h. 在 10 个样本中, 仅观察到其中前 6 个元件的寿命分别为 896, 950, 1005, 1100, 1150, 1220 (单位为 h). 问这批元件是否合格 (取  $\alpha = 0.05$ )?

解 (1)  $X_1, \dots, X_n$  的联合分布为

$$f(x_1, \dots, x_n) = \exp\{-\sum_{i=1}^n x_i\} / U(1) > 0, T(x) = \sum_{i=1}^n x_i.$$

$n$

分布族关于  $r = \sum_{i=1}^n x_i$  为 MLR, 所以由定理 6.3.2 可知以上假设检验问题  $H_0: \theta \leq \theta_0$  对  $H_1: \theta > \theta_0$  的 UMP 检验为

$$h_1(T) = \begin{cases} 1, & T > k, \\ 0, & T \leq k, \end{cases}$$

其中  $k$  由  $E_{H_0}(T(X)) = \alpha$  决定, 而  $T \sim r(1/\theta, n)$  或  $277.12(270)$ . 因此有

$$1 - \exp(-k/\theta) = \alpha$$

6.3 单调似然比分布族的单边检验

227

$$P_{\theta}(T > k) = P_{\theta_0}(T > k)$$

$$= \alpha.$$

$$\theta_0 < \theta$$
 由此可得  $2k/\theta = 2(2\alpha, 1-\alpha)$ ,  $k_0 = y_{\alpha} \theta_0 X_{2(2\alpha, 1-\alpha)} \cdot$

(2) 由第一章的公式可知,  $X(1), X(2), \dots, X(r)$  的联合分布为

$$n! \cdot 1/r^n \cdot \exp(-\sum_{i=1}^r x_i)$$

其中  $y = \sum_{i=1}^r x_i + (n-r) \cdot j(r)$ . 分布族关于  $T = \sum_{i=1}^r x_i$  为 MLR,

因此检验的 UMP 检验仍然如 (1) 中  $h_1(T)$  所示, 而  $k$  由  $r = T/r = r$

$+ (n-r)X(r) \sim r(1/\alpha, r)$  决定. 这时以上定解条件应改为

$$i=1$$

$$\text{且有 } 2k/\theta_0 = 2Y_{2(2\alpha, 1-\alpha)}, k_0 = y_{\alpha} \theta_0 X_{2(2\alpha, 1-\alpha)} \cdot$$

(3) 该问题是要根据 (2) 的结果来检验  $H_0: \theta \leq 1000$  对  $H_1: \theta > 1000$ , 其中  $\theta_0 = 1000$ ,  $n=10$ ,  $r=6$ . 今取  $\alpha=0.05$ , 根据给定的 6 个数据, 经直接计算可得  $T/r = 11201$ ,  $2T/r < r/\theta_0 = 22.402$ , 而  $Y_{2(0.95, 0.95)} = 21.026$ , 即  $2T/r > Y_{2(0.95, 0.95)}$ . 因此应该否定  $H_0$ . 认为  $H_1$  成立, 所以这批元件是合格的 ( $\alpha=0.05$ ). 曜

例 6.3.6 设  $X_1, \dots, X_n$  独立同分布,  $1 \sim$  两点分布  $B(1, \theta)$ .

(1) 求假设检验问题  $H_0: P \leq P_0$  对  $H_1: P > P_0$  的 UMP 检验及其近似正态否定域;

(2) 某技工以往的生产记录是: 他加工的零件一般有 60% 为一等品. 在晋级考核中, 在他加工的零件中随机抽取 100 个有 69 个一等品. 在水平为  $\alpha=0.05$  和  $0.01$  的条件下检验: 该技工的加工技术是否有显著提高?

$n$

解 (1) 分布族关于  $T = \sum_{i=1}^n X_i$  为 MLR, 因此以上假设检验问题  $H_0: P \leq P_0$  对  $H_1: P > P_0$  的 UMP 检验为

的 UMP 检验为

$f_1, T > k, (x) = I_{10}, T < k,$

其中左由  $= P(T > k) = a$  决定, 而当  $p = p_0$  时  $T \sim b(n, p_0)$ —由第一章的公式可知

228

## 第六章参数假设检验

"。(%) =, —\_!1 斤(0, 1), o

其中  $9_0 = I - P_0$ —由于%(幻为r的严增函数, 因此否定域可表示为  $R_+ = \{UQ(x) > c\}$ , 而  $c$  由  $p_0(t/o(\%) > c) = a$  决定. 根据  $i_0$ 。(幻的渐近止态性, 可取  $C = c_{fl} = ^_{-}$ . 因此检验问题的近似正态否定域为  $7?+ =$

$I(1) > 4-。!。(2)$ 该问题是要根据(1)的结果来检验  $H_0: P=0.6-H\} : p > 0.6.$

由于  $n = 100$  比较大, 可应用近似正态否定域. 这时  $p_0. = 0.6, 9_0 = 0.4,$

$7 = 69$ , 可得  $UQ(x) = (69 - 100 - 0.6)/7100 \cdot 0.6 \cdot 0.4 = 1.837$ . 当水平  $a=0.05$  时,  $95 = 1.65 < \%(%) = 1.837$ , 因此应否定原假设, 即认为该技工的加T. 技术有显著提高. 但是, 若取水平  $a=0.01$ , 则有  $z_{0.99} = 2.326 > \{ /0(x) = 1.837$ , 不能否定原假设, 即不能认为该技工的加T. 技术有显著提高. 这两种结论的差别显然是由于水平条件  $a$  的改变而引起的, 取水平  $a=0.01$ , 就意味着要求有更充分的证据作判断 (即判错的概率更小). 这时, 该技工加工100个零件中有69个一等品, 还不足以证明他的加工技术有显著提高. ■

关于单:边假设检验问题, 很自然地要考虑与检验(川)相对应的另

一种形式 路和方法与 (iii) 十分类似, 不再重复, 以下仅简述其大意与结果.

即下面的 (iii)' . 求解其UMPT的思

与检验 (iii) 类似, 我们可考虑以下假设检验问题及其MPT和 UMPT:

(i)'  $H_0: \theta = 0. -H_1 = \theta, \text{ 但 } 0 < \theta$  (ii)'  $< 0n;$

(iii)'  $: 6 < 0q.$

对于检验 (i)', 根据N-P引理, 其MPT的否定域为

$_f(x, \theta) = A(\%)$

由于氏) 4, 根据MLR的定义6.3. 1有

$\theta_1 < \theta n.$

=

其中  $h(T(x))$  为  $T(x)$  的非减函数, 因而  $A(\%) = h_{-1}(x)$  为  $T_0$  的非增函数, 所以有

$7T=U(x) > c\} = \{ [h(T(x)) - \sim] > c (C_iTV) < k\}$ . 该式与 (6.3.2) 式十分类似, 我们可进一步证明, 假设检验问题 (i)' —

$A(x)$

$'/(x, )' /(\text{无, 没I}).$

$t(x), e_i$

, 没0 ) ]

### 6.3 单调似然比分布族的单边检验

229

(iii)' 的MPT和UMPT的否定域具有  $R_+ = \{T(x) < k\}$  (的形式. 最后可. 得到与定理6.

3. 2 相平行的结果如下:

定理6. 3.3 在定理6.3. 1的假设下, 以下(6.3.4)式确定的 为假设检验问题 (i)' 的MPT, 为 (ii)' 和 (iii)' 的UMPT:

3,  $T(x) < k,$

$= < y \gg r(x) = k, (6.3.4) [0, T(X) > k,$

其中左和7由  $] = a$  决定, 而且相应的功效函数  $pf(e)$  为没的 减函数. I 注意, 单边假设检验问题 (iii) 和 (iii)' 的UMPT的功效函数都是单



调的, 这是一个重要特点. 由于指数族分布包含了许多常见的分布, 同时又是我们下面讨论双

边检验的基本分布族, 因此把指数族分布单边检验的结果归纳如下: 定理6.3.4设

$X=(X_1, \dots, X_n)$  服从指数族分布  $h(x) \cdot$

(8), 其中  $\eta \in R^1$ ,  $\phi(\eta)$  为  $\eta$  的严增函数, 则假设检验问题 (iii) 的UMPT为(6.3.1)式的  $T(x)$ , 假设检验问题 (iii)' 的UMPT为 (6.3.4) 式的  $T(x)$

最后再证明一个关于单边假设检验的常用性质.

定理6.3.5设  $\phi^*(x)$  为假设检验问题 (iii) 的UMPT, 如(6.3.1) 式所示; 又设  $\phi(x)$  为任一满足  $E[\phi(X)] = \alpha$  的检验, 则有

$E[\phi^*(X)] \geq E[\phi(X)]$  (6.3.5) (本定理的意义为:  $\phi^*(x)$  的功效函数在  $\eta_0$  右面最大, 在  $\eta_0$  左面最小(如

图6.3.2所示), 即其  $\beta(\eta)$  和  $n(\eta)$  都是最小.) 证明以下证明采用假设检验问

题中常用的一种方法, 这在定理 6.3.1的推论后面曾经提到过. 考虑

假设检验问题

(i)'  $H_0: \theta = \theta_0 - H_1: \theta = \theta_0^+$  但

1

$\theta_0 > \theta_0^+$  今定义一个检验  $\phi^*(x) = \phi(x)$ , 它可表示为

$\phi(x) = 1, T(x) \leq k, \phi(x) = 0, T(x) > k,$

图6.3.2单边检验UMPT的功效函数

## 230 第六章参数假设检验

且有  $E[\phi^*(X)] = E[\phi(X)] = 1 - \alpha$  因此  $\phi(x)$  为 (i)' 的水平为

$1 - \alpha$  的检验又由于  $\theta_0$  具有定理6.3.3中(6.14) 式的形式, 因此它是 (i)' 的水平为  $1 - \alpha$  的UMPT. 另外, 由  $H_1$  的假设可知,  $E[\phi^*(X)] = 1 - \alpha$ , 因此  $\phi(x)$  也是 (i)' 的水平为  $1 - \alpha$  的检验, 因而有

$E[\phi^*(X)] \geq E[\phi(X)] = 1 - \alpha$ . 由  $E[\phi^*(X)] = E[\phi(X)]$  即可得到

(/

EJ

以上结果对任何  $\theta \in \Theta$  成立, 因此(6.3.5)式成立: ■

## 6.3.2 正态分布单参数的单边检验

以上定理可用于正态分布的假设检验. 因此对于正态分布, 其单参数的单边假设检验问题都可以得到解决. 设  $X_1, \dots, X_n$  独立同分布, 且  $X_i \sim N(\mu, \sigma^2)$ , 则有

(1) 若  $\mu$  已知, 对于以下假设检验问题:

$H_0: \mu = \mu_0 - H_1: \mu > \mu_0$  或  $H_0: \mu = \mu_0 - H_1: \mu < \mu_0$ , 由于分布族关于  $T = \bar{X}$  为MLR, 且  $\bar{X} \sim N(\mu_0, \sigma^2/n)$ , 因此由(6.3.1) 式和(6.3.4) 式可得到相应的UMPT.

(2) 若  $\sigma^2$  已知 (不妨设  $\sigma^2 = 1$ ), 对于以下假设检验问题:

$H_0: \mu = \mu_0 - H_1: \mu > \mu_0$  或  $H_0: \mu = \mu_0 - H_1: \mu < \mu_0$

由于分布族关于  $T = \bar{X}$  为 MLR, 且  $\bar{X} \sim N(\mu_0, 1/n)$

式和(6.3.4) 式也可得到相应的UMPT.

例 6.3.7 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\mu, \sigma^2)$  (1) 求假设检验问题  $H_0: \mu = \mu_0 - H_1: \mu > \mu_0$  的UMPT; (2) 今设  $\tau$  表示某种砖的抗断强度, 要求大于  $30 \text{ N/cm}^2$ . 今对一

批砖测得6 个样品的强度分别为32.56, 29.66, 31.64, 30.00, 31.87, 31.03. 问这批砖是否合格? 其中已知  $d = 1.21$ , 并取  $\alpha = 0.05$ .

解 (1) 容易验证, 相应的分布族关于  $T = \bar{X}$  为MLR, 因此由定理 6.3.2可知, 以上检验的UMPT具有(6.3.1) 式给定的形式, 即

$\phi(x) = 1,$

小(x) = L 0 ,

又由例6.2.3可知, 当心=1。+ (~/人)21\_。时有 $E[1/(X)] = \infty$ 。在应用中, 以上

<#>\* (x)常表示为

$x > k_0 t$

-

.

$x < k_0$ .

$\theta' < e_0$ .

;  $a < a_0$ , 因此由(6.3.1)

4 已知.

### 6.3单调似然比分布族的单边检验 231

1,

0, (2)这相当于考虑以下单边假设检验问题: $H_0: \theta \leq \theta_0 - H_1: \theta > \theta_0$

30, 其UMPT由(1)给出. 这时 $\theta_0 = 6$ ,  $\theta = 30$ ,  $\alpha = 0.05$ ,  $\beta = 0.95 = 1 - 0.05$ . 根据以上测得的样本值可计算出 $\bar{x} = 31.03$ , 代入(6.3.6)式可得 $U(x) = 76(31.03 - 30) /$   
 $= 2.294$  (亦可得到 $k_0 =$

$30 + (1/1721776) - 1.65 = 30.74$ ). 由于  $U(x) = 2.294 > z_{0.95} = 1.65$ , 所

以应否定原假设 $H_0$ ; 认为 $H_1: \theta > 30$ 成立, 即这批砖合格. 注

意, 检验设定的水平为 $\alpha = 0.05$ ; 而本次检验的p-值为 $P_{H_0}(X > 31.03) = P^{(U(X))}$

$> 2.294) = 0.011$ , 即本次判断出错的概率为0.011, 比 $\alpha = 0.05$ 小. 另外, 在假设检验中,

我们把“砖的抗断强度大于

30 N/cm<sup>2</sup>”放在对立假设 $H_1$ 上, 以保证判断出错的概率(即第一类错误)很小, 这是很有必要的. |

例6.3.8 设 $X_1, \dots, X_n$ 独立同分布,  $X_i \sim N(\theta, \sigma^2)$ . 求假设检验问题

$H_0: \theta \leq \theta_0 - H_1: \theta > \theta_0$ 的UMPT及其功效函数.

解当 $\theta > \theta_0$ 时, 似然比为

因此相应的分布族关于 $\theta$ 为MLR, 所以由定理6.3.3可知以上  $1 = 1$

$t/(x) < z_j$

$U(x)$

$1 - X \sim N(\theta, \sigma^2)$

----- (6.3.6)

$f_0$

假设检验问题的UMPT为

$\hat{\theta}(x) = \theta_0$

$T(x) > k_0$

其中 $A$ : 由 $E_{\theta_0}(T) = A$ 决定. 由于 $T/\sigma \sim N(\theta/\sigma, 1)$

$\sigma = 1$

$P_{\theta_0}(T < k_0) = P_{\theta_0}(T/\sigma < k_0/\sigma) = \alpha$ .

因此

由此可得  $A_{\alpha} = \theta_0 + z_{1-\alpha} \sigma$ :  $A_{\alpha} = \theta_0 + z_{1-\alpha} \sigma$   $\sigma = 1$ . 代入以上  $\theta_0 = 6$  即得UMPT. 其功效函数为

人,  $(\theta) = E[1/(X)] = P_{\theta}(T < k_0) = P_{\theta}(T/\sigma < k_0/\sigma) = P(Z < (k_0/\sigma - \theta)/\sigma) = \Phi((k_0/\sigma - \theta)/\sigma)$ .

易见功效函数 $\beta(\theta)$ 是  $\theta$  的减函数. |

232

## 第六章参数假设检验

### 6.4单参数指数族分布的双边检验

上一节我们从简单假设检验的MPT出发，对于单调似然比分布族，得到了两种单边假设检验问题的UMPT。本节将继续讨论双边假设检验问题的UMPT。由下面的讨论可知，双边检验要比单边检验复杂得多，必须对分布族和检验函数进一步加以限制。本节主要讨论指数族分布和无偏检验的一致最优势检验。

#### 6.4.1 双边检验问题及无偏检验

考虑以下双边假设检验问题：设  $(\theta)$  或  $\theta \in \Theta$ ；

(ii)  $\theta \in \Theta_0: \theta < \theta_0$  或  $\theta > \theta_0$ ；

(iii)  $\theta \in \Theta_0: \theta = \theta_0$ 。注意，检验(ii)是(i)的特例，因为在(i)中取  $\theta_0 = \theta_0$  则得到(ii)。另外，检验(ii)和(iii)的是无界的，因而更加复杂。

为了求解这些假设检验问题的一致最优势检验，我们首先考察其功效函数  $\beta(\theta)$  的变化趋势。设检验  $\phi(x)$  的功效函数为  $\beta(\theta)$ 。一个好的检验，其功效函数应该在  $\theta > \theta_0$  上满足  $\beta(\theta) \rightarrow 1$ ，而在  $\theta < \theta_0$  上应尽量小。图6.4.1给出了检验问题(i)–(iii)的功效函数的示意图，前者应该是“两头小中间大”而后者应该是“中间小两头大”。总之，它们的UMPT的功效函数不可能是  $\theta$  的单调函数。特别，在上一节得到的单边检验的

UMPT, BP(6.3.1)式的  $\phi(x)$  以及(6.3.4)式的

$\phi(x)$  都不可能是本节双边检验的UMPT, 因为  $\phi(x)$  的功效函数为增函数，小  $\phi(x)$  的功效函数为减函数，它们不可能有图6.4.1所示的

形式。为了求解以上双边假设检验问题的UMPT，需要对检验函数进一步

加以限制。由图6.4.1可见，若  $\phi(x)$  是一个好的检验，则其功效函数在  $\theta_0$  上的值应该尽量大，显然应该大于  $\alpha$ 。由此可引出以下定义：

定义6.4.1 无偏检验和一致最优无偏检验。对于假设检验问题  $H_0: \theta \in \Theta_0$  对  $H_1: \theta \in \Theta_1$ ，若

则称  $\phi(x)$  为假设检验问题一个水平为  $\alpha$  的无偏检验，并记

$E_{\theta_0}[\phi(X)] \geq \alpha$ ,  $E_{\theta_1}[\phi(X)] \geq \beta(\theta_1)$

#### 6.4单参数指数族分布的双边检验

233

少： $\theta \in \Theta_0: E_{\theta}[\phi(X)] \geq \alpha$ ； $E_{\theta}[\phi(X)] \geq \alpha, \theta \in \Theta_1$ 。

(6.4.1) 则称  $\phi(x)$  为假设检验问题水平为  $\alpha$  的无偏检验类。若  $\phi(x) \in \mathcal{U}$ ，且对一切  $\theta \in \Theta_0$  有

则称  $\phi(x)$  为一致最优无偏检验，记为UMPUT。无偏检验的意义很清楚，其第二类错误满足  $\beta(\theta) = 1 - \alpha$  的矣

1 -  $\alpha$ 。另外，显然有  $\beta(\theta) \geq \alpha$ ， $\beta(\theta) \geq \beta(\theta_1)$  在  $\theta \in \Theta_1$  上。

现在来看双边假设检验问题(i)–(iii)的无偏检验的功效函数(见图6.4.2)。由无偏检验的定义可知，对于检验问题(i)和(ii)，若功效函数  $\beta(\theta)$  关于  $\theta$  连续，则有

$\beta(\theta_0) = E_{\theta_0}[\phi(X)] = \alpha$ 。对于检验问题(iii)，若功效函数  $\beta(\theta)$  关于  $\theta$  连续且可导，则有

$\beta(\theta_0) = E_{\theta_0}[\phi(X)] = \alpha$ ， $\beta'(\theta_0) = 0$ 。这些性质对于下一节讨论的指数族分布的无偏检验都是成立的。

#### (6.4.2) (6.4.3)

6.4.2 指数族分布的双边检验 由于双边假设检验问题的复杂性，目前只能在指数族分布中得到一

234

#### 第六章参数假设检验

图6.4.2 (i)–(iii)无偏检验的功效函数

致最优势检验。因此本节仅限于讨论指数族分布，即假定  $p(x; \theta) = h(x)q(\theta)T(x)$ ， $\theta \in \Theta$ ，

#### (6.4.4)

其中 $Q(\theta)$ 为 $\theta$ 的严增函数.并讨论该分布族的双边假设检验问题(i) — (iii)(见本节开头).

设 $\phi$ 为一个检验函数,通常总假定它可测,因而由第一章的定理可知,其功效函数
$$3(\theta) = E_J(\phi)(X) = \int \phi(x) dF_\theta(x)$$

关于 $\theta$ 连续可导.因此对于无偏检验 $\phi(\theta) \equiv 0$ ;以上(6.4.2)式对检验问题(i)和(ii)成立;(6.4.3)式对检验问题(iii)成立.

本节将证明,对于双边检验(i),存在一致最优势检验(UMPT);而对于双边检验(ii)和(iii),只存在一致最优无偏检验(UMPB).虽然证明比较复杂,但是解决问题的过程与第3节的单边检验还是比较类似的.今简要说明如下:在第3节中,为了求解单边检验“ $H_0: \theta \leq \theta_0$ ”的UMPT,首先考虑简单假设检验 $\theta = \theta_0$ ,但 $\theta > \theta_0$ 的MPT,这可由N-P引理得到;然后再逐步推广,得到单边检验的UMPT.对于本节讨论的双边检验,应该也可以有类似的过程.以第(i)个双边检验为例:为了求解双边检验或没有

K

的umpt

, 首先考虑一个与之有关而最简单的

#### 6.4单参数指数族分布的双边检验

235

检验:

(i)'  $\theta = \theta_0$  :  $\theta = \theta_0$ , 但  $\theta < \theta_0$  (6.4.5)

事实证明,对于指数族分布这是可行的,我们可先求出(i)'的MPT,然后再逐步推广,得到双边检验(i)的UMPT.但是必须注意,求解(i)'的MPT并不简单,与N-P引理中的两点简单假设相比,它的零假设 $H_0$ 中多了一个条件,这就要把N-P引理加以推广.所以必须首先介绍推广的Neyman - Pearson引理.另外,由于后面有好几处要用到推广的N-P引理,因此下面将采用较为一般的形式,读者可从中比较N-P引理与推广的N-P引理之间的异同,从而更深入地了解其意义.

定理6.4.1 (推广的Neyman-Pearson引理) 设 $f_i(x)$ ,  $i=1, 2, 3$  关于可积,  $0 \leq f_i(x) \leq 1$ , 并记

中 $a = \int f_1(x) f_2(x) dx$  (6.4.6)

少:  $\{x: f_1(x) = f_2(x)\}$  (6.4.7)

今定义一个检验函数 $\phi$ 如下:

$\phi(x) = \begin{cases} 1 & \text{if } f_1(x) > c f_2(x) \\ 0 & \text{if } f_1(x) < c f_2(x) \end{cases}$

(6.4.8) 令

$\phi(x) = \begin{cases} 1 & \text{if } f_1(x) < c f_2(x) \\ 0 & \text{if } f_1(x) > c f_2(x) \end{cases}$

(1) 若存在常数 $c_1, c_2$  (不一定多0)及 $y$ , 使得 $\phi(x) \equiv 0$ ; 则对一切 $\theta \in \Theta$ , 即 $\phi(x) \equiv 0$ , 有

$\int f_1(x) dP_\theta(x) \leq \int f_2(x) dP_\theta(x)$  (6.4.9)

(2) 若存在常数 $C_1 > 0, C_2 > 0$ 及 $y$ , 使得 $\phi(x) \equiv 0$ ; 则对一切 $\theta \in \Theta$ , 即

$\int f_1(x) dP_\theta(x) \leq \int f_2(x) dP_\theta(x)$  (6.4.9)式成立.

而且, 以上(1)和(2)中的 $\phi(x)$ 在 $R^+ \cup \infty$ 上都唯一. 证明证明方法与N-P引理完全类似, 记

$\phi(x) = [(\phi(x) - \phi(y)) / (f_1(x) - f_1(y))] \cdot [f_1(x) - f_1(y)]$ , 由(6.4.8)式 $\phi(x)$ 的定义可知, 在 $TU/T$

上,  $A(x) \equiv 0$ , 所以有

$\int [(\phi(x) - \phi(y)) / (f_1(x) - f_1(y))] \cdot [f_1(x) - f_1(y)] dP_\theta(x)$

$$= f A(x)dfc(x) = f 4(\%)d^Lt(x) > 0. \cdot / \wedge JR^ UR-$$

因而有

## 236 第六章参数假设检验

由此可得

$$f \wedge (x) / ; (x) dyct(x) > f \text{ 每}(X) \text{人}(X) \text{如}(X) + \text{余项}, (6.4.10)$$

其中余项为上式后两个积分之和. 对于情形(1), 由于  $\langle / \rangle (x) \in 0; , 4 \rangle (X) E$  中、, 因此有

所以余项等于零, 因此由(6.4.10)式可知(6.4.9)式成立. 对于情形(2), 由于

$C_j > 0, c_2 > 0$ , 且  $\wedge (\%) c(Pa)$ , 则

$$\langle / \rangle (x) / (x) d/i(x) = \langle \bullet, 4 \rangle (x) f_i(x) d/jL(x), i = 1, 2.$$

因而余项为0, 亦可推出(6.4.9)式. 唯一性的证明与N-P引理的证明完全类似, 不再重复.

■ 注 最常用的情形为  $/(\%) = / (x; (?), i = 1, 2, 3$ , 这时  $A = | \%$ :  $/ (x; ^3) > C1 / (x; ^)$

$+ e / (\%; ^2) |$ . 特别, 若  $y; (幻不存在$ , 则  $/ ? + = jx$ :

$/ (x; ^3) > c_2 f(x; e_2) !$ , 即可得到n-p引理. 另外, 在  $0a$  和少: 中经常有  $at = a, 0 < a < l$ ,

$i = 1, 2$ . ■ 现在回到假设检验问题(i)' (见(6.4.5)式), 其最优检验可由推广的N-P引理得到.

引理6.4.1对于指数族分布(6.4.4), 考虑假设检验问题(i)'  $H_0: 6 \geq R$  或  $< 9 = \text{么} -$

$7/1: \text{没} = \text{没} 3$ , 但  $< 03 < 029$  并设  $0 < a < l$ . 设检验函数  $\langle \# \rangle . (\%)$  可表示为

0, 若存在  $y i X^ = h_2$ ), 使

$$k_1 < T(x) < k_2,$$

$$T(x) = k_i (i = 1, 2), T(x) < \text{么} \text{或} T(x) > k_2.$$

(6.4.11)

即  $= E, 2 [(\wedge^1(X)) = a$ , 则  $(t > i(x))$  为假设检验问题(i)'的MPT, 即对一切满足  $E, 3$   $(\wedge)$

每(I) 的每(太) (即4)(x)  $g 0a$  都有

$$E.3 [\langle / \rangle 1(\wedge)] , 6. < 03 < 02f (6.4.12)$$

并且屯(幻在TTu/r上唯一.

证明 对于指数族分布  $f(x, \theta)$  的  $h(x) e_{-T(4)}$ , 记  $= / (x, 6\theta, i = h 2, 3$ . 设企'(x) 由推广的N-P引理的(6.4.8)式所定义, 若

## 6.4单参数指数族分布的双边检验

237

存在  $c, \wedge 0, c_2 \wedge 0$ , 使  $\langle / \rangle , (x) \in$  即  $[(\wedge^1(X)) = a$ , 则由推广的N-P引理的结论(2)可

知,  $(/ > !(\%)$  必使(6.4.12)式成立(见(6.4.9)式). 以下根据(6.4.8)式考虑其  $/ ? +$  可能的形式, 其中

$$R_+ = \setminus x: f(x, \theta_3) > c / (X, ex) + C_j f(\%, \text{没} 2) f$$

$$= (\%: a(h) e_{-}) r(4) SqaWJeM \rangle 7 \blacksquare (7) + c_2 a(6 > 2) \text{户} 2) r(*) \}.$$

由于  $ex < e. < e_2$ . 因此  $(2(< 91) < QCo3) < Q(e_2)$ , 在  $/ ? +$  中用  $a(\wedge) eo(.3) r(\langle \rangle)$  除各式, 可得

$$R_+ = \setminus x: T\{x\} = t, g(\theta = (ije \sim 61t + d2eb2t < 1), (6.4.13) \text{ 其中 } dx) a \sim *(\wedge^3) > d_2$$

$$= c_2 a(\wedge^2) a_{-1}(\wedge^3), -bx = \wedge(\wedge) - Q(\wedge^3) < 0, b_2 = Q(62) - \langle ?(6 > 3) > 0, \text{且有}$$

$$6, > 0, b_2 > 0.$$

“

1), 在(6.4.13)中, 只可能  $d_2 \wedge 0$  (不全为0), 即  $c) 0, c_2 \wedge 0$

为了能使  $e$  少

即  $= a(0 < a <$

(不全为0), 其原因如下:

(1) 若  $d_1 \neq 0, d_2 \neq 0$ , 则 | 惟有  $g(1) < 1$ , 所以  $R^+ = \varnothing$ , 从而  $\langle Mx \rangle = \varnothing$ ,

$\int_{\mathbb{R}^+} f(x) dx = 0$ , 因此只可能有  $\int_{\mathbb{R}^+} f(x) dx = a = 1$ , 这与假设  $0 < a < 1$  矛盾.

(2) 若  $d_1 > 0, d_2 < 0, d_1 + d_2 > 0$  的情形类似, 从略), 则有  $g(t) = -b_1 d_1 e^{-t} + b_2 d_2 e^{2t} < 0$ .

因此  $g(t)$  为  $f$  的减函数, 则由 (6. 4. 13) 式有  $R^+ = \{x: T(x) > t\}$  对某个  $A$ . 因此由定理 6.3.2 可知, 相应的功效函数  $\phi(t)$  为  $\theta$  的严增函数, 所以不可能有  $\phi(t) \equiv 0$  ( $t > 0$ ).

么、 $\phi(t) = \int_{\mathbb{R}^+} f(x) dx = a < 1$ ,  $\phi(0) = 1$  即 结合以上 (1)、(2) 两点可知, 为使  $\phi(t)$  连续

$\phi(t) = \int_{\mathbb{R}^+} f(x) dx = a$ , 只可能  $\phi(t) \equiv 0$ ,  $\phi(t) \equiv 1$  (不全为 0), 即  $c_1 \neq 0, c_2 \neq 0$

$\phi(t)$  (不全为 0), 因此

$\phi(t) = \int_{\mathbb{R}^+} f(x) dx = a$ .

$g''(t)$  (尤) 所以  $g(t)$  为凸函数. 由此可根据 (6. 4. 13) 式进一步确定  $T$  的形式. 图 6.4.3 为凸函数的示意图, 由图可

知, 应存在  $k < k''$

$R^+ = \{x: T(x) = t, g(t) < 1\}$

$= \{x: kx < T(x) < kx^2\}$ .

(6.4. 14) 如果解出  $kx, kx^2$  则上式等价于

使得

238

## 第六章参数假设检验

(6.4.8) 式中的  $T$ , 且使 (6.4.7) 式成立. 综合以上讨论可知, 对于指数族分布, 若要使

(6.4.8) 式确定的  $T$  满

(/即

(6. 4.14) 式, 因而  $\phi(t)$  必有 (6.4.11) 式的形式. 至于  $\phi(t)$  ( $t = 1, 2$ ) 的具体数值, 仍然要由  $\int_{\mathbb{R}^+} f(x) dx$  决定. 解出  $\phi(t)$  ( $t = 1, 2$ ) 以后, 则由推广的 N-P 引理的结论 (2) 可知, 对任何  $\theta(x) \in \Theta_0$ , 即  $\int_{\mathbb{R}^+} f(x) dx = a$ ,  $\int_{\mathbb{R}^+} f(x) dx \geq a$ , 必有 (6.4.12) 式成立. |

由于以上 (6. 4. 11) 式的  $\phi(t)$  与  $\theta$  无关, 因而 (6. 4. 12) 式对任何  $\theta$  的  $\phi(t) < 1$  都成立, 所以有

是 (6.4.7) 式:

$\int_{\mathbb{R}^+} f(x) dx = a$ , 则它等价于

推论 6.4.1 为以下假设检验问题 (i) 的 UMP

② 我们的目标是要求出假设检验问题 (i) 的 UMP, 现在已经知道

么 (幻为 (i) 的 UMP, 下面还需要证明: 为 (i) 的水平为  $a$  的检验, 即有  $\int_{\mathbb{R}^+} f(x) dx = a$ ,  $\int_{\mathbb{R}^+} f(x) dx \geq a$  则根据引理 6.2.1 可知,  $\phi(t)$  为 (i) 的 UMP. 为此, 我们还需要下面的引理 6.4.2. 另外, 这个引理在求解假设检验问题 (ii) 的 UMP 中将起重要作用.

引理 6.4.2 对于指数族分布 (6. 4.4), 考虑假设检验问题

(ii)  $\theta = \theta_0$  或  $\theta = \theta_1$  但  $\theta_0 \neq \theta_1$  或  $\theta_0 = \theta_1$  但  $\theta_0 \neq \theta_1$ .

(i)  $\theta = \theta_0$  或  $\theta = \theta_1$

$\theta = \theta_0$  或  $\theta = \theta_1$ .

|

设检验函数  $\phi(x)$  可表示为

$\phi(x) < k$  或  $\phi(x) > k$ .

小 $2(\alpha) = y_2 i$ ,  $7(\%) = k_i (i=1, 2)$ ,  $o$ ,  $k_x < T(x) < k_2$ .

若存在 $\alpha$ ,  $y_2 i (Z=1, 2)$ 使 $\langle />2(x)g \ 0'' \ 0 < \alpha < 1$ , 即

(6.4. 15)

$[\langle />2(X)] =$

E如尤 $] =$  则对一切满足  $[\wedge^2(X)] = E92[_4 > 2(X)] = a$ 的 $\alpha(x)$  (即去 $(x)e0'J$ 都有

E.4 $[\wedge(\wedge) > E, 4[5 > (\wedge)]$ ,  $v$ 氏 $\langle$ 氏(或E,5 $[(/ > 2(X)] > E f l 5[\wedge(X)]$ ,  $V^{\wedge} >$ 氏),

(6.4.16)

并且 $\langle />2(x)$ 在 $R+ \ U/T$ 上唯一.

证明 以下仅证明“H1:e = 04但氏 $\langle$ 久”的情形, 另一种情形

“%:0 = 么但 $\wedge > V$ 的证明完全类似.记 $Z(x) = / (x, \text{氏})$ ,  $i = 1, 2$ ; 人(幻 $= / (\text{心氏})$ .设

$(/ > 2(x)$ 由推广的N-P引理的(6.4.8)式所定义, 若

“

推广的N-P引理的结论(1)可知,  $(/ > 2(x)$ 必使(6. 4. 16)式成立(见

存在  $q, c_2$ , 使 $(M \text{ 幻 } e \text{ 少}$

,  $J^{\wedge 2}(X)] = \ll$ , 则由

即  $= E$

(6. 4.9)式). 以下根据(6.4.8)式考虑其ZT可能的形式, 其中

$<$

:

#### 6.4单参数指数族分布的双边检验 239

$R+ = \{x: f(x, \theta_4) > C_j / (x, \text{氏}) + x, \theta_2) \mid$

$= (x: a(\alpha) e_{-x}) > c_1 a(\theta_1) e_{QWT(x)} - c_2 a(\wedge^2) e_{Q(\wedge)T(x)}\}$ .

由于  $\langle e' < \theta_2$ , 因此 $(2(\wedge) < QCoj < Q(e_2)$ . • 以下证明, 为使 $(l > 2(x) w a$ , 即

$E^{\wedge}[(/ > 2(X)] = a(i = 1, 2)$ , 必有

$C_1 > 0, c_2 < 0$ .

(1) 若 $C_0, c_2 \neq 0$ , 则 $R+ = \wedge, E J(/ > 2(X)] = 1$ , 这与假设 $\theta <$

$a: < 1$ 矛盾.

(2) 若 $C_0, c_2 > 0$ , 则

$R' = \{x: c_2 a(\text{氏}) e_{(10)2} r(4) < a(\theta_4) e_{QWT(x)} - C_1 a(e_{QWT(x)})\} = (\%:$

$T\{x\} = t, g(z) = \sim 611 + d_2 e^{\sim b_2 t} > 1 \mid$ ,

其中 $dx = c'1a''1(\ )a(\ ) > 0$ ,  $d_2 = -c_21a_1(\text{沒}2)c!a(\text{沒}!) > 0$ ,  $-bx = Q(o_4) - Q(e_2)$

$< 0$ ,  $-b_2 = C(\wedge^i) - \langle 2(\wedge^2) < 0$ ,  $b\} > 0$ ,  $b_2 > 0$ , 因此有

茗'(i) =  $\sim h y t - b_2 d_2 e^{\sim h_2 t} < 0$ , 为减函数, 因此 $= \{x: g(\theta > 1) = U: r(x) < k$ (对某个

女,

由定理6.3.3知,  $/ ? +$ 对应的 $\langle />2(x)$ 的功效函数为P的减函数, 因而不可 能有

$= E, 2[(/ > 2(X)] = a -$

(3) 因此必有 $c, > 0$ , 这时有

$R+ = \{x: c_x a(e\})^{\wedge QWT\{x\}} < a(WWT(x)^{\wedge c_2 a(e_2) e_{QWT\{x\}}}]$

$= jx: T(\%) = t, g(t) = c^{\wedge e-6''} + d_2 e^{6z} < 1(,$

其中  $Q(o_4) < Q(\theta_i) < Q(\theta_2)$ ;  $d_ = c''a-1(\text{氏})a(f-t_j) > 0$ ,  $d_2 = (\text{氏}) \cdot$

$c_2 a(A), -b\} = \langle 2((94) - \langle 2(\theta_j)$ ,  $b_2 = Q(\theta_2)$  在以上 $/ ? +$ 中, 若 $\langle />2^{\wedge 0}$ (即 $c_2 > 0$ ),

则 $gf(t) = +b_2 d_2 e^{b_2 t}$

$< 0$ ,  $gG$ 为减函数. 与(2)类似, 这种情况不可能.

因此必有  $d_2 > 0$ (即  $c_2 < 0$ ), 此时有  $g''(z) = 6? < e^{-6, f + 6^{\wedge 2} e^{ft^2} / > 0$ ,

因而 $g(\theta$ 为凸函数, 参见图6.4.3. 由图可知, 应存在 $k_x < k_2$ , 使得  $R+ = \{x:$

$T(x) = t, g(1) > 1I = \{\%: r(x) < k_i$ 或 $T(x) > k_2 I$ .

(6.4.17) 如果解出  $c$ ,  $k_2$ , 则上式等价于(6.4.8)式中的  $\phi$ , 且使(6.4.7)式成立.

综合以上讨论可知, 若要使(6.4.8)式确定的否定域  $T$  满足

于(6.4.17)式, 因而  $\phi_2(x)$  必有(6.4.15)式的形式. 至于  $k$ .  $\phi_2(x)$  (即 (6.4.7)式)  $\phi_2(x)$  少

即  $E, J[\phi_2(X)] = E, J[\phi_2(X)] = \alpha$ , 则它等价

1, 2)的具体数值, 仍然要由  $E[\phi_2(X)] = E, J[\phi_2(X)] = \alpha$  决定. 解出  $\phi$ ,  $\phi_2(x)$  以后, 则由推广的N-P引理的结论(1)可知, 对任何  $\alpha, \phi_2 > 0, \phi_2 > 0$ .

240

## 第六章参数假设检验

$\phi_2(x)$  即  $E, J[\phi_2(X)] = E, J[\phi_2(X)] = \alpha$ , 必有(6.4.16)式成立. 注意 必须注意, 引理 6.4.2与引理6.4.1有重要的区别. 在引理 6.4.1中, 对任何满足  $E, J[\phi_2(X)] = \alpha$  的  $\phi_2(x)$  (即  $\phi_2(x) \in \mathcal{C}$ ) 都有  $E, J[\phi_2(X)] = \alpha$ . 而在引理6.4.2中, 只是 对任何满足  $E, J[\phi_2(X)] = \alpha$  的  $\phi_2(x)$  (即  $\phi_2(x) \in \mathcal{C}$ ) 才有  $E, J[\phi_2(X)] = \alpha$ . 因此,  $\phi_2(x)$  还不能算是假设检验问题(ii)' 的 MPT, 因为对  $E, J[\phi_2(X)] = \alpha$  的  $\phi_2(x)$  (即  $\phi_2(x) \in \mathcal{C}$ ) 才有  $E, J[\phi_2(X)] = \alpha$ . 引理 6.4.2 未有  $E, J[\phi_2(X)] = \alpha$  的  $\phi_2(x)$  的结论. 但是引理6.4.2对于求解假设

检验问题(ii)的一致最优无偏检验还是可用的, 因为由(6.4.2)式可知, 任何无偏检验  $\phi_2(x)$  都满足  $E, J[\phi_2(X)] = \alpha$ . |

推论1  $\phi_2(x)$  是以下假设检验问题的一致最优无偏检验 (UMP):

(ii)"  $H_0: \theta = \theta_0$  或  $\theta = \theta_1$  或  $\theta = \theta_2$ .

证明 首先, 由(6.4.15)式确定的  $\phi_2(x)$  与  $\theta$  及  $\alpha$  无关, 所以

(6.4.16)式对任意的  $\theta < \theta_0$  或  $\theta > \theta_0$  以及  $\theta > \theta_0$  都成立, 即  $\phi_2(x) \in \mathcal{C}$ .

其次, 容易证明:  $\phi_2(x)$  是(ii)"的无偏检验, 因为若取  $\phi_2(x) \in \mathcal{C}$ , 则

对任何  $\theta < \theta_0$  或  $\theta > \theta_0$  有  $E, J[\phi_2(X)] = \alpha$  为(ii)"的任一无偏检验, 则必有  $E$

另外, 若  $\phi_2(x) \in \mathcal{C}$ ,  $E, J[\phi_2(X)] = \alpha$ ,  $\theta = \theta_0$  或  $\theta = \theta_1$ ;

$\theta > \theta_2$

. 因为对于指数族分布是连

. 因而由以上 引理可知, 以上不等式对(ii)"的任一无偏检验都成立, 所以  $\phi_2(x)$  (是 连续的, 所以  $E, J[\phi_2(X)] = E, J[\phi_2(X)] = \alpha$ , 即  $\phi_2(x) \in \mathcal{C}$ ;

(ii)"的 UMP. | 推论2 记  $\phi_2(x) = 1 - \phi_2(x)$ ,  $\phi_2(x)$  由(6.4.11)式给出, 则  $\phi_2(x)$  为(ii)"水平为  $1 - \alpha$  的 UMP. 证明由(6.4.11)式可知,  $\phi_2(x)$  可表示为

• 1,  $\phi_2(x) = (1 - \alpha)^{-1}$ ,

• 0,

$T(x) < k$  或  $r(r) > k$ ,  $T(x) = k$ ,  $1, 2$ ,

$k < T(x) < k$

且有  $\phi_2(x) = 1 - E, J[\phi_2(X)] = 1 - \alpha$ ,  $1, 2$ , 因此  $\phi_2(x) \in \mathcal{C}$ ;  $\alpha$  且具有引理6.

4.2中(6.4.15)式的形式. 所以由唯一性以及以上 推论1可知,  $\phi_2(x)$  为(ii)"水平为 1 的 UMP. |

有了以上两个引理, 即可求出假设检验问题(i)的 UMP 以及(ii)的 UMP.

### 6.4单参数指数族分布的双边检验

241

定理6.4.2 对假设检验问题  $H_0: \theta = \theta_0$  或  $\theta = \theta_1$  或  $\theta = \theta_2$  若存在  $\theta_0$  和  $\theta_1$  ( $\theta_0 < \theta_1$ ), 使(6.



4. 11)式的 $(\phi, \psi)$ 满足 $E\phi(X) = a (i = 1, 2)$ , 即 $(\phi, \psi) (0 < a < 1)$ , 则检验为 $(i)$ 的UMPT.

证明 根据引理6.4.1的推论以及其后的说明,  $(\phi, \psi)$ 已经是假设检验问题 $(i)$ 的水平为 $a$ 的UMPT, 因此我们的主要任务是要证明 $(\phi, \psi)$ 为假设检验问题 $(i)$ 的水平为 $a$ 的检验. 即要证: 当 $0 \leq \alpha$ 、或 $\alpha = 0.2$ 时有为此可用引理6.4.2的推论2(亦可参见定理6.3.5的证明). 由于 $(\phi, \psi) = 1 - \psi(x)$ 为检验 $(ii)$ 的水平为 $(1 - a)$ 的UMPUT, 因此若取 $\phi(x) = 1 - a$ , 则有

$EJ_1 - \psi(\phi) = EJ(\phi; V) \geq EJ(\phi; T) = 1 - a, 0 < 0.1 \leq \alpha < 0.2$ . 由此即可得到 $0 < \alpha \leq 0.1$ 或 $\alpha = 0.2$ . 另外也有 $a(Z = 1, 2)$ , 因此也 $(\phi, \psi)$ 为假设检验问题 $(i)$ 的水平为 $0.1$ 的检验. 另一方面, 若 $(\phi, \psi)$ 为假设检验问题 $(i)$ 的水平为 $a$ 的检验, 即 $EJ(\phi; X) \geq a, 0.1 \leq \alpha < 0.2$ 或 $\alpha = 0.2$ 外, 则对此 $(\phi, \psi)$ 在 $\alpha = 0.1$ 和 $\alpha = 0.2$ 处也有 $E\phi(X) \geq a$ ,  $E\psi(X) \geq a$ .

即 $\phi(x) \geq 0.1$ , 因此由引理6.4.1的推论可知, 对任何 $0.1 < \alpha < 0.2$ 必有(可参见(6.4.12)式)

$EJ(\phi; X) \geq EJ(\phi; T), 0.1 < \alpha < 0.2$ ,

因此检验 $A$ 为假设检验问题 $(i)$ 的UMPT. |

定理 6.4.3 对于假设检验问题 $(ii)$   $f_0: \phi: \phi \leq \phi_0$

氏, 若存在 $\phi$ 和 $\phi_0 = 1, 2$ , 使(6.4.15)式的 $(\phi, \psi)$ 满足 $E\phi(X) = a (i = 1, 2)$ , 即 $(\phi, \psi) \leq 0; (0 < a < 1)$ , 则检验 $(\phi, \psi)$ 为 $(ii)$ 的UMPUT.

证明 根据引理6.4.2的推论1,  $(\phi, \psi)$ 已经是假设检验问题 $(ii)$ 的水平为 $a$ 的UMPUT, 而且(6.4.16)式成立. 因此我们的主要任务是要证明 $(\phi, \psi)$ 为假设检验问题 $(ii)$ 的水平为 $a$ 的无偏检验. 即要证: 当 $\alpha \leq 0.1$ 时有 $\alpha \leq a$ ; 当 $\alpha < 0.1$ 或 $\alpha = 0.2$ 氏时有 $Ej(\phi; X) \geq a$ .

这可应用类似于证明定理6.4.2的方法. 考虑 $\phi(x) = 1 - \phi(x)$ , 则有

$(\phi, \psi) = 1 - \psi(x), b,$

$kx < T(x) < k_2,$

$r(x) = \phi = 1, 2, r(jt)$  或  $T(x) > k_2$ .

242

## 第六章参数假设检验

则 $(\phi, \psi)$ 具有引理6.4.1中(6.4.11)式的形式, 且有 $E, \phi(X) = 1 - a (i = 1, 2)$ . 因此由引理6.4.1的推论知,  $(\phi, \psi)$ 为检验 $(i)$ 的水平为 $1 - a$ 的UMPT, 即对一切各有

$EJ(\phi; X) \geq EJ(\phi; T), \alpha \leq 0.1$ 或 $\alpha = 0.2$ .

取 $\phi(x) = 1 - a$ 可得 $1 - a,$

$\alpha < 0 < 0.2$ .

$0 < \alpha < 0.2$ ; 另外也有 $[Ej(\phi; X)]$

因此有  $EJ$

$\phi, \psi$

$a(f = 1$

$(\phi, \psi) (X) ]$

. 因此 $\phi(x)$ 为 $(ii)$ 的水平为 $a$ 的检验. 同时, 在引理6.4.2的推论2)

1 中已经证明 的无偏检验.

:

" 所以 $(\phi, \psi)$ 也是 $(ii)$

又) $] \geq a, \alpha < 0.1$ 或 $\alpha = 0.2$

)

另一方面, 若为假设检验问题(ii)的水平为 $\alpha$ 的无偏检验, 则由(6.4.2)式可知 $E\phi(X) = E\phi_2(X) = \alpha$ . 因此由引理6.4.2

的推论1可知, 对任何 $\phi_1, \phi_2$ 必有 $\phi_1(x) \leq \phi_2(x)$ ,  $e^{-\phi_1} \geq e^{-\phi_2}$ .

因此检验 $\phi(X)$ 为假设检验问题(ii)的 UMPUT. | 注 必须注意, 定理6.4.2和定理6.4.3有很重要的区别: 前者由引理6.4.1及其推论得到, 是 UMPT (也是本节和下一节唯一的一个

UMPT); 而后者由引理6.4.2及其推论1得到, 是 UMPUT. 导致这种差别的主要原因是应用了推广的N-P引理的不同结论. 引理6.4.1及其推论是由推广的N-P引理的结论(2)得到 (其中要求 $c_1 > 0, c_2 > 0$ ); 而引理6.4.2及其推论1是由推广的N-P引理的结论(1)得到 (这时不要 求  $c_1 \geq 0, c_2 > 0$ ). |

最后考虑假设检验问题(ni)  $H_0: \theta = \theta_0$  vs  $\theta > \theta_0$ . 它可视为检验

IIH

(ii)的特例(即

(6.4.15)式所示的 $(\phi_1, \phi_2)$ 的形式. 但是定解条件不一样, 在检验(ii)中, 可通过 $E\phi_1(X) = \alpha$ 和 $E\phi_2(X) = \alpha$ 这两个条件来决定

屯(幻中的么,  $y_2, Z=1, 2$ . 而对检验问题(iii), 只能有 $E\phi(X) = \alpha$  这一个定解条件. 因此要设法增加定解条件, 这显然应该从(iii)的功效函数的特点(见图6.4.2)出发, 其主要特点为 $\phi'(\theta) = 0$ . 今归纳为

以下引理:

引理6.4.3 对于指数族分布 $f(x, \theta) = h(x)e^{\theta T(x)}$ , 则 $\phi(\theta)$

的期望可表示为

$E\phi(X) = -\phi'(\theta) - \phi''(\theta)$  若 $\phi(X)$ 为一检验函数, 其功效函数为 $\phi(\theta)$ , 则有

), 因此可以预料, 其最优解应该具有

,

6.4单参数指数族分布的双边检验  $\phi'(4) = 0, (e) \setminus Ee^{X T(X)} - \backslash$   
243

(6.4.19) 对于假设检验问题(iii), 它的任一无偏检验满足 $\phi'(\theta) = 0$ , 其等价条件为

$E, o[\phi(X) r(jr)] = E, o[\phi(X)] E, o[r(y)] = \alpha E, o[r(\theta)]$

(6.4.20) 证明 由于 $\int a(\theta) h(X) e^{\theta T(X)} d\phi(x) = 1$ , 该式对 $\theta$ 求导可得 $\int [a'(\theta) h(X) e^{\theta T(X)} + (2/(6) r(\theta) a(\theta) / l(x) e^{\theta T(X)}) d\phi(x) = 0$ .

(9)/i(x) e\_如(%) = -Q(10) j/(x) a(9) h(x) e\_T(4) 如(x).

由此即可得到(6.4.18)式.

由于功效函数  $\phi(0) = (l(X) a(0) h(X) e^{\theta T(X)}) d\phi(x)$  该式对 $\theta$ 求导可得

陶 = f. [CW(%) r(x) a(8) AWe'm) + a'(8) (Wx) e\_rw] 如(%)

=\_L[t^ (4)w 訓X")

=<2冲[卿跡点缶戶叫

(6.4.18)式代入上式即可得到(6.4.19)式. 若 $\phi(X)$ 为(Hi)的任一无偏检验, 则其功效函数满足 $\phi'(\theta) = 0$ ,

因此可得(6.4.20)式. | 现在考虑如何应用推广的N-P引理求解假设检验问题(iii)的 UMPUT. 由于情况比较特殊, 我们首先作一些直观的说明. 若 $\phi(X)$ 为检验(iii)的任一无偏检验, 则条件 $\phi'(\theta) = E\phi(X) = \alpha$

可表示为

$f$ , 小( $\theta$ )( $x$ )如( $X$ )= $\theta$ ,  $f$ ; ( $\theta=0$ , 没。),  $\theta=0$ . (6.4.21) 条件(6.4.20)式可表示为

$(f(x)r(x))^j = f(x)r(x)/(x^0)dM(x)$  因此若取 $f_2(x)=T(x)f(x, 0Q)$ ,  $a_2=a_0$  则有

#### 244 第六章参数假设检验

$f_2(x)f_2(x)dpL(x)=a_2$ ,  $f_2(x)=T(x)f(x, 0Q)$ ,  $a_2=a_0$ .

(6.4.22) 对于假设检验问题( $H_0: \theta=0$ )  $f_2(x)$  可取 $f_3(x): f(x, Q)$ ,

若 $f_2(x)$ 由推广的N-P引理的(6.4.8)式所定义, 并使之满足(6.4.21)式和(6.4.22)式, 则根据推广的N-P引理的(6.4.9)式, 对任意满足(6.4.21)式和(6.4.22)式的去( $\theta$ )将会有

$E_0[f_2(x)] = (f_2(x)(W(\theta)) \geq J = E_0[f_3(x)]$ .

因此, 根据以上分析得到的 $\theta$ ( $\theta$ )应该就是假设检验问题( $H_0$ )的 UMPUT, 由此可得以下定理:

定理6.4.4考虑满足以下条件的检验函数类:

$\phi \leq 1$  一切小( $\theta$ ):  $E_0[\phi(X)] = \alpha = \alpha!$ ,

$E_0[\phi(X)T(X)] = aE_0[T(X)] = af$   $a_2$  |.

(1) 若存在  $k$ .  $\theta$ 使(6.4.15)式中的  $f_2(x) \leq 5; (0 < a < 1)$ , 则有

$[f_2(x)] \geq E_0[f_3(x)], \forall \theta \neq 0$ .  $\forall f_2(x) \leq 0$ .

(6.4.23) (2) 以上 $f_2(x)$ 为假设检验问题( $H_0$ )的 UMPUT, 即( $H_0$ )的 UMPUT

具有(6.4.15)式的形式, 且满足  $E_j c/$

$h(\theta) = \alpha, E_0[\phi(X)T(X)] = aE_0[T(X)] = af$

(6.4.24) 证明(1)为了应用推广的N-P引理, 今引入以下记号(见以上

$e^0$

(6.4.21)式和(6.4.22)式):

$\phi = a(\theta)h(x)e_{\theta} T(x) f_2(x) = T(x)f(x, \theta)$ .

$Z(\theta) = f(U)$ .

$A(x) = f(x, 0) = a(\theta)h(x)e_{\theta} T(x) ( \forall \theta \neq 0)$ . 若每( $\theta$ )  $e_{\theta}$ , 则有

$h(x), l(x)$ 如( $x$ )  $[f(I)] = a$

$f^2(x)f_2(x)d^2(x) = E_0[f_3(x)r(X)] = a_2$ .

设小( $\theta$ )由推广的N-P引理的(6.4.8)式所定义, 若存在 $q, c_2$ , 使小( $\theta$ )则根据推广的N-P引理的结论(1), 必有(见(6.4.9)式)

$= f_2(x)f_3(x)d^2(x) \geq l^2(x)f_3(x)d^2(x) = E_j^2(X)$ .

,

#### 6.4单参数指数族分布的双边检验

245

即(6.4.23)式成立. 以下根据推广的N-P引理的(6.4.8)式, 考虑 $R_+$ 可能的形式, 其中

$R_+ = \{x: f_3(x) > c_1 f_1(x) + c^2(x)\}$ ,

$= \{x: a(\theta)h(x)e_{\theta} T(x) > c_1 a(\theta_0)h(x)e_{\theta_0} T(x) + c_2 T(x)a(\theta)h(x)e_{\theta} T(x)\}$

且使 $f_2(x) \leq 0$ ; (即 $E_j cMX$ )  $E_j f_2(x)r(X) = af$ . 上式可化简为

$R_+ = \{x: r(x) = t, e^{bt} > d_1 + d_2 t\}$ , 其中  $b = Q(\theta_1) - Q(\theta_0)$ ,

$d_i = a^{-1}(\theta_i)c_i, a(\theta_0)$ ,  $i = 1, 2$ .

今考虑以上 $R_+$ 可能的形式. 指数曲线 $e^{bt}$ 与直线 $d_1 + d_2 t$ 可能有0, 1, 2个交点, 见图6.4.4. 以下证明: 若 $R_+$ 对应的 $f_2(x)$  则 $e^{bt}$ 与 $d_1 + d_2 t$ 必有两个交点, 因为

i) 若  $e_6'$  与  $< +d_2t$  无交点, 则  $7T =$

o 这与假设  $0 < a < l$  矛盾.

$] = a = 1,$

$jx: eAt > (/j + d_2tI = .劣; E^$

ii) 若与  $d\} + d_2t$  只有一个交点, 则

$R+ = \backslash x * T\{x) = t, ebt > dx + d_2t ) = I T(x) < A \mid \text{或} \mid T(x) > k \}$ .

这时相应的功效函数冷 (60为6»的增函数或减函数, 因而  $/3f(e) > 0$  或  $f(\text{没}) < 0$ , 不可能有  $/?'(氏) = 0$ , 因而不可能有  $E, o[</)(X)T(X)] = (/ (该式等价$

图6.4.4检验(iii)否定域的确定

于冷'(氏) = 0, 见引理6.4.3), 因此只可能是最后一种情形, 即  $e_6'$  与  $d、+ d_2t$  有两个交点. 此时

尺+ =  $jx: T(x) = t, ebt > dl + d_2t \backslash = \backslash x: T(x) < kY \text{或} r(\%) > \wedge 2 \}$ . 该式就是 (6.4.17)

式. 如果解出  $A: , , k_2$ , 则等价于 (6.4.8) 式中的  $/ ? +$ , 且使 (6.4.7) 式成立. 因此, 若要使 (6.4.8) 式确定的否定域  $/ ? +$  满足 (6.4.7) 式:  $\text{么}(x) e_0 :$ , 即  $E、$

$o[(MX)] = a, E, o[(/ )^2(X)T(X)] = a \backslash$  则它等价于 (6.4.17) 式, 因而  $</>^2(x)$  必有 (6.4.15) 式的形式. 至于  $k$ .

$y_2iG = l, 2)$  的具体数值, 仍然要由  $<^>^2(x) e_0 ;$ , 即 (6.4.24) 式决定. 解出  $\downarrow$ ,  $yuU = l, 2)$  以后, 此时由推广的  $N-P$  引理的结论 (1) 可知, 对任何  $^(\%) e_0^$ , 必有 (6.4.23) 式成立.

(2) 要证以上  $</>^2(x)$  为假设检验问题 (iii) 的 UMPUT. 首先证明:

246 第六章参数假设检验

$</>^2(x)$  是 ((1)) 的无偏检验, 即  $Ee[(/ )^2(\%)]$  多  $(X, V \neq 00$ . 事实上, 在 (1) 中取每  $(\%) = a$ , 则有杏  $(\%) e$  逢:, 因为  $E^o[^(X)] = a, E<, J^)(X) \cdot$

$7 \backslash X)] = aEajr(X)]$ . 因此由 (6.4.23) 式知  $E, [(/ >^2(X)] E, [^(J)] = a, V$

再证屯 (幻的最优性. 设  $> (幻为 (iii)$  的任一无偏检验, 则有

$E, o[4 > (X) J] = a$ , 再由引理6.4.3可得  $E, o[^(X)T(X)] = aE, o[T(X)]$ , 即  $^> (x) G_0 ;$ ,

因此由本定理的 (1) 可知, 必有 (6.4.23) 式成立. 所以  $</>^2(x)$  为 (iii) 的 UMPUT. | 以上定理的定解条件 (6.4.24) 式在计算上较为复杂, 下面定理给出了一种常见的特殊情况, 把两个定解条件合并为一个简单的定解条件.

定理6.4.5 条件同定理6.4.4. 在 (6.4.24) 式中, 如果当  $\theta = \theta_0$

时,  $Z = T(X)$  的分布关于某个点  $m$ . 对称, 则假设检验问题 (iii) 的 UMPUT 可表示为

$J, 1 \downarrow Mo1)$  女,

$\text{渗} 2(1) = < y, h - \text{弘} o \mid = \text{灸}, (6.4.25)$

o,

证明由于以上定义的  $(/ >^2(\theta$  是 (6.4.15) 式的特例, 因此主要是证明它自然满足条件  $Ef fo[<t>^2(X)T(X)] = aE, o[r(X)J$ . 这样, 以上  $</>^2(\theta$  就满足 (6.4.24) 式. 设  $e = e_0$  时  $r = r(j)$  的分布为  $gG)$ , 由对称性条件知  $E_0 o[T(X)] = / \text{£} 0$  且有

$g(fjL_0 + t) = g(^0 - 0)$ , 因

此  $g(M_0 + 0$  为  $t$  的偶函数. 同理, 由 (6.4.25) 式确定的  $\text{小} 2(1)$  也关于 对称, 因而  $</> > (At_0 + 0$  亦为的偶函数. 所以, 计算  $E do[^(^> r(x)]$  时可变换到对称点, 即令  $\ll = Mo$  则有

$\wedge 00 \wedge 2(^> ^(^> 1 = [ </>^2(1)W(1) \text{山} J * 0 \text{©}$

$(Mo + T) \text{小} 2(/i_0 + 7)g(Ai_0 + r)dT = MoJ - 00 </>^2(Mo + T)g(/Io + T)dr$   
 $1t - \text{弘} 0 \mid < \text{左},$  其中  $k, y$  由条件  $E. . [ \text{小} 2(r)] = a$  决定 (即条件  $Ef lo[</$

$E[T(X)] = \int_0^\infty t f(t) dt$  可去掉。  
 $+ \int_0^\infty t g(t) dt$

#### 6.4 单参数指数族分布的双边检验

247

在以上积分中，第二项为零，因为被积函数为奇函数；而第一项 =  $\int_0^\infty t f(t) dt$   
 $= \int_0^\infty t f(t) dt$ 。因此 (6.4.25) 式的  $\int_0^\infty t f(t) dt$  满足 (6.4.24)

式，所以是假设检验问题 (iii) 的 UMPUT。I 6.4.3 正态分布单参数的双边检验

以上定理可用于正态分布的假设检验，因此对于正态分布，其单参数的双边假设检验问题都可以得到解决。设  $X_1, \dots, X_n$  独立同分布，且  $X_i \sim N(\mu, \sigma^2)$ ，则关于  $\mu$  和  $\sigma^2$  的双边假设检验各有 3 种（见本节开头），共 6 种。为节省篇幅以下通过例题讨论其中 4 种，其他类似。为了得到检验的 UMPUT 或 UMPUT，主要是给出否定域  $T$  的形式及其定解条件。

例 6.4.1 设  $X_1, \dots, X_n$  独立同分布，且  $X_i \sim N(\mu, \sigma^2)$ ，求以下假设检验的 UMPUT 或 UMPUT：

(1)  $\mu = \mu_0$  或  $\mu < \mu_0$  或  $\mu > \mu_0$ ，( $\sigma^2 = \sigma_0^2$  已知)；(2)  $\sigma^2 = \sigma_0^2$ ， $\mu = \mu_0$  或  $\mu < \mu_0$  或  $\mu > \mu_0$ ，( $\sigma^2 = \sigma_0^2$  已知)。

解 (1) 分布族关于  $r(X) = \sum_{i=1}^n (X_i - \mu_0)^2$  为指数族，且  $T(X)$  的

分布连续，因此由定理 6.4.2 知，检验的 UMPUT 可表示为  $1, k, \sum T(x) < k^2$ ，  
 (龙)

其中  $f(n, y)$  表示  $\chi^2(n)$  分布的密度函数，联立求解以上方程可得  $k^2$

$i$ ：的 UMPUT 可表示为

$\sum T(x) =$

$0, T(x) < k^2$  或  $r(x) > k^2$ ，其中  $k^2 = \chi^2_{1-\alpha}(n)$ ，而  $\alpha$  由  $\int_0^\infty f(x) dx = \alpha$ ， $j=1, 2$  决定。

定。即

“从而得到  $\chi^2_{1-\alpha}(n)$ 。但是， $k^2$  的具体计算比较复杂。n

k

(2) 分布族关于  $T = \sum_{i=1}^n X_i^2$  为指数族，因此由定理 6.4.4 知，检验

1

$1, T(x) < k^2$  或  $r(x) > k^2$ ，

$k^2 \in T(x) (k^2,$

0 其中  $k^2 = \chi^2_{1-\alpha}(n)$  或  $T(x) > k^2$ ，而  $k^2$  由 (6.4.24) 式决定，即

，

248

1。[么(又)] =  $\int_0^\infty f(x) dx$ ，式可得

第六章参数假设检验  $\int_0^\infty f(x) dx = 1$ 。因此由以上第一

$\int_0^\infty f(x) dx = 1$

2  $\int_0^\infty f(x) dx = 1 - \alpha$  (TQ

由于  $T(X)/n \sim \chi^2(n)$ ，因此有

$\chi^2_{1-\alpha}(n)$

$\int_0^\infty f(x) dx = 1 - \alpha$ 。

$V(T)$

由于  $E^2(T) = n\sigma^2$ ，因此以上第二式可化为

$E, \sigma^2 [T(X)] = n\sigma^2$ ，或  $E, \sigma^2 [T(X)] = n(1-\alpha)\sigma^2$ ，

即

$\int_0^\infty f(x) dx = 1 - \alpha$

$\int_0^\infty f(x) dx = n(1 - \alpha)$ 。

(6.4.27) 这显然是比较复杂

将联立求解方程(6.4.26)和(6.4.27)可得到 $k$ ,

作为近似, 可根据(6.4.26)式取

$$k \approx \frac{2}{n} \ln \frac{1}{\alpha}$$

$$\approx \frac{2}{n} \ln \frac{1}{\alpha}$$

$$\approx \frac{2}{n} \ln \frac{1}{\alpha}$$

$$n, \ln \frac{1}{\alpha}$$

设 $X_1, \dots, X_n$ 独立同分布, 且 $X_i \sim N(\mu, \sigma^2)$ , 求以下假设(1)  $H_0: \mu = \mu_0$  或  $H_1: \mu > \mu_0$  ( $\sigma^2$  已知);

例6.4.2 设检验的UMP:

$H_0: \mu = \mu_0$ , ( $\sigma^2$  已知).

(3) 某钢铁厂的铁水含碳量的百分比在正常情况下应服从正态分

布 $N(4.55, 0.1082)$ . 为了检测设备维修后生产是否正常, 测试了5个样品的铁水含碳量的百分比, 结果为4.29, 4.39, 4.45, 4.52, 4.54. 问现在生产是否正常(设方差未变, 并取 $\alpha = 0.05$ )?

解 (1) 分布族关于 $T = \bar{X}$ 为指数族, 因此由定理6.4.3知小2(幻

的 $T$ 为 $\bar{X}$   $\Rightarrow \bar{X} < k$  (或 $\bar{X} > k$ ). 而 $k$ , 由 $E[\bar{X}] = \mu_0$ ,  $\mu_0 = 4.55$  决定. 即

(2)  $H_0: \mu = \mu_0$

% 由此可得

$$\frac{\partial}{\partial \mu} \ln L(\mu) = \frac{1}{\sigma^2} (\bar{X} - \mu)$$

$$= 1 - \alpha, \quad i = 1, 2.$$

$$\bar{X} < k \Rightarrow \bar{X} < k \Rightarrow \bar{X} < k \Rightarrow \bar{X} < k$$

$\alpha, z = 1, 2$ . 联立求解以上方程可得到 $k$ , 从而得到 $\phi(\bar{X})$ .

(6.4.26)

### 6.5 多参数指数族的检验

249

(2) 分布族关于 $T = X$ 为指数族, 而且 $X$ 的分布关于 $A$  对称,  $X \sim N(\mu, \sigma^2)$  因此由定理6.4.5可得

$$1 - \alpha = P(\bar{X} < k)$$

$$1 - \alpha = P(\bar{X} < k) = P(\bar{X} < k)$$

其中 $\alpha = P(\bar{X} > k) = \alpha$  决定, 因此有

其中 $\bar{X} > k \Rightarrow \bar{X} > k \Rightarrow \bar{X} > k \Rightarrow \bar{X} > k$ , 因此 $\bar{X} = (\bar{X} - \mu_0) / \sigma$ . 把 $A$ 代入以上小2(x),

最后可表示为

$$[1, \alpha]$$

$$I U(x) |$$

$$\bar{X} > k \Rightarrow \bar{X} > k \Rightarrow \bar{X} > k \Rightarrow \bar{X} > k$$

以上公式与第3节中关于 $\bar{X}$  的单边检验的UMPT的公式十分类似(见例6.3.7).

(3) 问题就是检验  $H_0: \mu = 4.55$  ( $\sigma^2 = 0.1082$ ). 此时 $n = 5$ ,  $\alpha = 0.05$ .

108, 经直接计算得 $\bar{x} = 4.438$ ,  $|\bar{x} - \mu_0| / \sigma =$

$$|4.438 - 4.55| / \sqrt{0.1082/5} = 0.112. \text{ 根据(6.4.28)式, } |z| = 2.319. \text{ 而 } 1 - \alpha/2 = 0.975$$

$= 0.975 \Rightarrow z_{0.975} = 1.96 < 2.319$ , 应否定 $H_0$ . 即认为现在生产不够正常. 检验的 $P$ -值, 即

本次判错的概率为 $P(|U(X)| > 2.319) = 0.02$  (其中 $U(X) \sim N(0, 1)$ ;  $|U(x)| = 2.319$ ).

最后可表示为

### 6.5 多参数指数族的检验

以上我们讨论了单参数的单边检验和双边检验, 但是除了 Poisson、二项等少数分布族外, 大多数常见的分布都不是单参数而是多参数的分布族. 特别是最常见的正态分布, 如果两个参数都是未知的, 并要讨论其假设检验问题. 则前面第3、第4节的结果都不能用. 例如, 在例6.4.2的(3)中所讨论的假设检验问题, 若含碳量的分布不是 $N(4.55, 0.1082)$ , 而是 $N(4.55, \sigma^2)$ , 即 $\sigma^2$ 未知(这可能更合理), 而要讨论同样的问题

Hq :  $Z = 4.55 - 7.55m^4$  (但  $a^2$  未知), 则(6.4.28)式的结果就不能用, 因为  $a$  未知. 直观上, 可用  $cr$  的估计代入(6.4.28)式(这显然是合理的), 但是要证明它是 UMPUT 则颇费周折. 这正是本节所要讨论的问题. 我们首先从直观上看看其困难所在. 设  $\dots$ ,  $X$  独立同分布,  $X \sim K(p, a^2)$  但  $m$  与  $a$  都未知. 考虑以下假设检验问题:

$$H_0: U(X) \leq a/2 \quad H_1: U(X) > a/2, \\ U(X) = a/h \quad (6.4.28)$$

## 第六章参数假设检验

$H_0: \theta = \theta_0$  vs  $H_1: \theta \neq \theta_0$  (但  $a^2$  未知).

我们从参数空间来分析其一致最优势检验的特点. 这时  $\theta = (M^*)$ , 而  $\theta_0 = I(\theta_0)$ ,  $V < r$ ,  $\theta_0$  是二维无穷集, 而两

者又连在一起. 总的来讲, 用 Neyman - Pearson 理论来求解多参数假设检验问题的图 6.5.1 多参数检验的参数空间

一致最优势检验并不是很有成效. 下一节介绍的似然比检验可能更好. 以下仅介绍指数族分布中一种有解的情形.

设  $X = (X_1, \dots, X_n)^T$ , 而  $X \sim K(p, a)$  为指数族分布:  $f(x) = A(x) \exp\{t^T(x)/a\} +$  (6.5.1)

其中  $t$  为一维,  $T(x) = (T_1(x), \dots, T_r(x))^T$ ,  $\theta = (\theta_1, \dots, \theta_r)^T$ ;  $\theta$  称为有兴趣参数,  $p$  称为多余参数. 我们的目的是要研究有兴趣参数  $\theta$  的假设检验问题, 包括第3节研究的单边检验 (iii) 和 (iii)' 以及第4节研究的双边检验 (i) - (iii). 但是多余参数  $p$  是未知的. 这些问题对指数族分布 (6.5.1) 是有解的, 我们首先复习带有多余参数的指数族分布的有关公式, 详见第一章.

(1)  $U = U(X)$  与  $T = T(X)$  的联合分布与边缘分布:  $((T, U) \sim f(t, u; \theta, p)$

$$= h(t, u) \exp\{\theta^T t + g(u; p)\}, \\ U \sim \exp\{\theta^T t - b(\theta, p)\} \quad T \sim \exp\{p^T t - b(p, \theta)\}$$

(2)  $U|T$  的条件分布与多余参数  $p$  无关:  $U|T=t \sim f(u|t; \theta) = h^*(u, t) \exp\{-\theta^T t - b(\theta, p)\}$  (6.5.2)

(6.5.3) 由上式可知, 在条件分布 (6.5.3) 式中, 仅包含有兴趣的参数  $\theta$ , 不包含多余参数  $p$ . 因此, 第3节和第4节的结果可用于该分布族关于

6.5.1 带有多余参数时单参数检验的 UMPUT

我们首先考虑以下单边检验, 说明其求解的基本思想, 其他检验的解法都类似.

(1) - (i):  $H_0: \theta \leq \theta_0$ ,  $p$  任意:  $\theta > \theta_0$  vs  $p$  任意.

## 6.5.1 带有多余参数时单参数检验的 UMPUT

251

没的假设检验问题. 本节的主要结果就是: 分布族 (6.5.2) 关于  $\theta$  的假设检验问题的 UMPUT 可以通过求解条件分布 (6.5.3) 关于  $\theta$  同样的假设检验问题的 UMPUT 得到, 而分布族 (6.5.3) 的 UMPUT 则可由第3节和第4节的相应结果得到.

## 6.5.1 带有多余参数时单参数检验的 UMPUT

我们首先考虑以下单边检验, 说明其求解的基本思想, 其他检验的解法都类似.

(1) - (i):  $H_0: \theta \leq \theta_0$ ,  $p$  任意:  $\theta > \theta_0$  vs  $p$  任意.

(1)-(ii) ;  $p$  任意  $\rightarrow i_0 < p$  任意. 其中“(1)-(i)”表示样本分布为(6.5.2)式, 这一检验的参数空

间为  $\{V < p\}$  和  $\{V = i\}$  ( $i < 9$ , 平),  $V > 0$ ,  $V \neq 0$  ; (1)-(ii)的情况类似. 为了得到检验(1)-(i)和(1)-(ii)的UMPUT, 今考虑基于条件分布(6.5.3)式的类似检验:

(2)-(i):  $H_0: \theta > 0$  (与 $p$ 无关). (2)-(ii):  $H_0$  心陶 (与 $p$ 无关).

其中“(2)-(i)”表示样本分布为条件分布(6.5.3)式, (2)-(ii)的情况类似. 以下主要证明: 检验(2)-(i)的UMPT就是检验(1)-(i)的UMPUT; 检验(2)-(ii)的UMPT就是检验(1)-(ii)的UMPUT. 为此, 首先考虑检验(2)-(i)和(2)-(ii)的UMPT:

引理6.5.1 假设检验问题(2)-(i)在条件分布(6.5.3)式下的UMPT为

$U(x) > k(t)$ ,  
 $\phi(u, t) = y(0, U(x) = k(t))$ , (6.5.4)

其中左(1),  $y(z)$ 由以下(2)-(i)'的UMPT:  
 $1, 0,$   
 $U(x) < k(t),$   
 $|T| = a$  决定. 特别,  $\phi(u, 0)$ 也是  
(2)-(i)':  $\theta > 0$ .

假设检验问题(2)-(ii)在条件分布(6.5.3)式下的UMPT为

$f_1$ ,  $\phi'(u, t) = y(1),$   
 $\theta$ , 其中  $k(1), y(z)$ 由  $E \sim [r] | r]$   
 $\phi(x) < k(1)$ ,  $\phi(x) = k(1)$ ,  $U(x)$   
决定.  
(6.5.5)

证明 直接把定理6.3.1-定理6.3.3用到条件分布(6.5.3)及其检验问题(i)和(ii)即可. |

则有

$[ \phi(7) ] = E(w) ! U W = E(w) [ \phi(t/, r) ] - a = 0, \forall \theta.$   
252

## 第六章参数假设检验

为了证明  $\phi(u, r)$  是(1)-(1)的UMPUT, 现大致分析如下: 由于  $\phi(u, r)$  为(2)-(i)的UMPT, 因此对任何  $\phi(u, t)$ , 若

$E, [ \phi(a, r) | e^e \theta ]$ , 则有  
 $|r|, \theta > 0$  (6.5.6)

注意, 以上结果与 $\theta$ 无关, 即对任何 $\theta$ 都成立. 该式两端取期望即可得到  $\int E(\sim [ \phi(t/, r) ]),$   
 $e) \theta a, \forall \theta$  (6.5.7)

这说明,  $(U, t)$ 可能就是(1)-(i)的UMPUT. 但是必须证明: (a)  $\phi(u, t)$ 也是(I)-(i)的无偏检验; (b) 对任一(1)-(i)的无偏检验  $\phi(u, t)$ , (6.5.6)式也成立 (因为这时才能两端取期望, 从而得到(6.5.7)式). 为此, 首先证明以下引理:

引理6.5.2 设检验

必有,

$r$

特别, 若不  $(u, t)$  为(1)-(i)的无偏检验, 则(6.5.8)式成立.

$E($

$]$

$)$

$w > [$ 如",



)fr

=a

(6.5.8

满足 $E(w)[\phi(t/\tau)] = a$ , 则

证明主要是应用指数族分布中统计量 $T(Z)$ 的完备性. 由假设知

$E(w)[\phi(t/\tau)] = a$ ,  $V^*$ , 考虑函数:

$g(t) = E(w)[\phi(t/\tau) | T=t] - a$ ,  $V^*$ ,

因此有 $E(w)[g(t) | T=t] = 0$ , 而当 $t$ 固定时,  $r(x)$ 关于参数 $p$ 为完备充分统计量(见

(6.5.2)式), 因而由完备性的定义可知 $g(t) = 0$  (a.e.), 即(6.5.8)式成立. 另外, 若义

$(u, t)$ 为(1)-(i)的无偏检验, 则由无偏性的性质可知 $E(\phi(t/\tau)) = a$ ,  $V^*$ , 因此(6.5.8)式成立. |

定理6.5.1对于指数族分布(6.5.1), 在引理6.5.1中由(6.5.4)式确定的 $\phi - W$ 为假设检验问题(1)-(i)的UMPUT; 由(6.5.5)式确定的为假设检验问题(1)-(ii)的UMPUT.

证明 以下仅证明假设检验问题(1)-(i)的UMPUT为小 $\alpha(u, t)$ , (1)-(ii)的情况完全类似. 为此, 要证明: (a)小 $\alpha(u, t)$ 为(1)-(i)

的水平为 $\alpha$ 的无偏检验; (b)  $\phi(u, t)$ 是最优的, 即满足(6.5.7)式. (a)由于 $\phi(u, t)$

为(2)-(i)的UMPT, 因此关于 $T$ 的条件分

布应用定理6.3.1和定理6.3.2的结果可得

对 $\alpha$ ,  $0 < \alpha < 1$ ,

$E(\phi | T=t) = \alpha$ ,  $0 < \alpha < 1$ , 对 $\alpha$ ,  $0 < \alpha < 1$ .

) |  $T=t$  -

6.5 多参数指数族的检验

253

上式在 $(0, p)$ 处取期望可得

$E(w)E(\phi | T=t) = E(\phi) = \alpha$ ,  $0 < \alpha < 1$ ,

因而有

因此 $\phi(u, t)$ 为假设检验问题(1)-(i)的水平为 $\alpha$ 的无偏检验. (b)今设 $\phi(u, t)$ 为检验(1)-(i)的任一无偏检验, 则由引理

6.5.2 知

$m = \alpha$ ,  $E(\phi) = \alpha$ . 又由引理6.5.1可知,  $\phi(u, t)$ 为假设检验问题(2)-(i)'的UMPT, 而

去 $(\alpha, \phi)$ 为(2)-(i)'的水平为 $\alpha$ 的检验, 因此有

$E(\phi) = \alpha$  (与 $p$ 无关).

该式两端在 $(0, p)$ 处取期望可得  $\alpha = \alpha$ ,  $0 < \alpha < 1$ .

综合以上(a), (b)两点可知,  $\phi(u, t)$ 为假设检验问题(1)-(i)的UMPUT.

我们继续考虑在分布(6.5.2)下参数 $\theta$ 的双边检验:

(1)-(iii) (1)-(iv):  $H_0$  (1)-(v):  $H_1 - \theta =$

或②

:  $0 < t$  或  $t > 2$ ,  $V < p$ ;

):  $H_0$

$0 < t < 2$ ,  $V < p$ ;

$E(\phi) = \alpha$

$\geq \alpha$ ,  $e < \infty$ ,  $\alpha < p$ ,

没 $\theta$ ,  $V$ 妒,  $\geq \alpha$ ,  $0 < \phi < 1$ .

$\phi(u, t)$ . 这些检验的UMPUT可由条件分布(6.5.3)下参数 $\theta$ 相应检验的

UMPUT得到, 其证明的基本思想已如上所述. 今归纳为以下定理: 定理6.5.2 对于指数族分

布(6.5. 1 ), 检验(1)-(hi)的

UMPUT 为

其中 $k_M$ 和 $y_G$ 由

检验(1)-(iv)和(1)-(v)的 UMPUT可表示为

$f_1' i/(X) (1)$ 或 $G(\tau) > \tau_2(1)$ ,

1,  $(0 < \tau(1) < \tau_2(0,$

$\tau_2(0, \tau_2(1) = A. (0(i=1,2), (6.5.9) 0, J_7(x) \text{ 或 } U(x) > k_2(t),$

(6. 5. 10)

$\diamond 2(U \gg 0 = 7 * 2 f('), 0,$

$U(x) = k_i(t) (i = 1, 2), k_x (0 < l/(x) < k_2(t).$

$= \ll U = 1, 2)$ 决定.

和 $y_{2l} \cdot (1)$ 由 $E d < M'', r) | r] = a (i=1,2)$

对于检验(1)-(iv),

决定;而对于检验(1)-(v),  $A:$ , (1)和 $y_{2M}$  由以下两式决定:

#### 254 第六章参数假设检验

$E, \circ [ \circ ('', \circ ] = \ll, (6.5.11) E, o [^2(i/, n | r_j = a E j t / | r], (6.5.12)$

证明 关于检验(1)-(iii)和(1)-(iv)的UMPUT, 其证明与定理 6.5. 1的证明十分类似, 故从略. 以下证明: 满足条件(6.5. 10)-

(6.5. 12)的  $\tau_2(u, 0$  为检验(1)-(v)的UMPUT. 为此, 首先考虑条件分布(6.5.3)

下(!)-(v)的相应检验: (2)-(v) :  $H_0: \theta = : \theta^3 0$ (与 $\theta$ 无关) 把定理6.4.4用于

条件分布(6.5.3)下的检验(2)-(v)可知, 满足

(6.5. 10)-(6.5. 12)的  $\tau_2(u, t)$ 为(2)-(v)的 UMPUT. 我们由此进一步证明:  $\tau_2(u, 0$ 也是检验(1)-(V)的UMPUT.

首先由无偏性可知 $E J (/) 2 (C /, T) | 7] \circ \sim$ , 该式两端对 (没4)取期望可得

$E(b) [小2('', \tau)] \geq a, V \text{ 没} \# \text{没} 0, V(p. \text{ 因而} \tau_2(u, t)$ 也是假设检验问题(1)-(V)

的无偏检验.

今考虑假设检验问题(1)-(V)的任一水平为 $a$  的无偏检验 每 $(U, \tau)$ , 以下证明 $\tau_2(u, 0$ 也满足(6.5. 11)式与(6. 5. 12)式, 从而可证 明 $\tau_2(u, t)$ 的最优性.

(a) 由 $\tau_2(u, r)$ 的无偏性可知  $E < e 0, < p) [多('', \tau)] = \ll, V(p.$

因此由引理 6.5.2 知  $E, o [^ ( U f T) | T] = a$ , 即 满足 (6.5. 11)式.

(b) 由无偏性, 应该满足类似于(6.4. 19)-(6. 4.20)式的 公式. 今考虑其功效函数及其导数:

扒 $\circ J$ 多 $(u, t) h(u \exp | 6u + y, \sim b(\theta f(p) ! d/x(x).$

该式对 $e$ 求导并在 $\theta_0$ 处计值可得 (氏, 中)

$J(w - \&, o) ^ (u \gg \theta ^ (w, t) \exp \backslash eou + 2 - \Delta ( \text{权} \theta, \text{史}) \blacksquare$  (如 $(\circ = 0, i=1$

(6. 5. 13)

其中 $b' o$ 表示 $\circ$ , 中、关于 $\theta$ 的导数并在 $\theta_0$ 处计值, 因此由指数族的性 质有  $= E(w) [i/$

$(X)],$ 同时也有  $= \ll,$ 因此由 (6.5. 13)式可得

$E(w) [ "去('', \tau) - a"] = 0, V < p.$

6.5多参数指数族的检验 与引理6.5.2类似, 考虑函数

255

对 $g(r)$ 取期望, 并应用以上两式可得 $E (/W) [g(r)] = \theta ( \backslash / \text{乎}).$ 当 $\theta =$  氏时,  $r(x)$ 关于  $\$$ 为完备充分统计量, 因而由完备性的定义知 $g(0 = \theta(a.e.),$ 从而 $u, 0$ 也满足(6.5.12)式.

(c)由以上证明可知都满足(6.5. 11)和

(6.5.12)式,把定理6.4.4的结论(1)用于条件分布(6.5.3)

可得

$E_j(\cdot) \geq 2(\cdot)$

$n|r]$ , (与

$n \leq E_j(\cdot)$  上式两端对  $(\cdot)$  取期望即得

$<p$  无关).

,

,

(10)) [每  $(\cdot, r)]$ ,  $v^{\wedge}$ .

E 由  $(\cdot, \cdot)$  的任意性以及 UMPUT 的定义可知,  $<^2(u, \theta)$  为假设检验问题 (1) - (V) 的 UMPUT. ■

以上定理可用于正态分布、Poisson 分布、二项分布等常见指数族 分布的假设检验问题.

注 本节结果与第3、第4 节的公式有重要的区别. 在本节带

有多余参数的假设检验问题中, 定解的关键是要计算相应的条件 期望, 而第3、第4 节的单参数检验只需要计算普通的无条件

期望. ■ 6.5.2 一样本正态总体的检验

假设  $1, \dots, n$  独立同分布,  $A \sim N(\mu, \sigma^2)$ , 其中  $\mu$  与  $\sigma^2$  均为

未知参数. 这时关于  $\mu$  和  $\sigma^2$  的检验各有5种, 以  $\mu$  的检验为例列举 如下:

(1)  $H_0: \mu \geq \mu_0$ ;

(ii)  $H_0: \mu \leq \mu_0$ ;

(iii)  $H_0: \mu = \mu_0$ ;

(iv)  $H_0: \mu < \mu_0$  或  $\mu > \mu_0$ ;

(v)  $H_0: \mu = \mu_0$  或  $\mu \neq \mu_0$ .

在以上检验中,  $\mu$  均为未知的多余参数. 类似的,  $\sigma^2$  的检验也

有5种,  $\sigma^2$  为未知的多余参数. 下面将说明, 有关  $\mu$  的检验都有解; 而关于  $\sigma^2$  的检验, 上述

(iii)、(iv) 无解, 其他有解. 根据前一节的 定理, 求解这些检验的 UMPUT 的关键是要计算

相应的条件期望. 对于正态分布, 我们可根据有关的独立性以及 Busa 定理把条件期望 化为

无条件期望. 在  $\mu$  和  $\sigma^2$  的10个检验中, 以下介绍其中几个典

256

## 第六章参数假设检验

型的例子, 其他类似. 首先我们列出正态总体的联合分布密度 如下:

$\exp \{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\}$

(1) 关于  $\mu$  的单边检验 (i)  $H_0: \mu \geq \mu_0$ . 根据 (6.5.14) 式, 这时可取  $\theta = -1/(2\sigma^2)$ ,

$U(x) =$

(6.5.14)

$x]$ ,  $<p =$

$i = 1, \dots, n$  且  $f(x) = x$ , 因此检验 (i) 可转化为 (i)':  $H_0: \mu \geq \mu_0$

$(\theta = -1/(2\sigma^2))$ . 由 (6.5.4) 式可知, 其 UMPUT 为

其中  $AG$  由  $\mu$  决定, 但可以化为

$|T| = a$  决定. 注意  $U(X) = \sum_{i=1}^n x_i$

与  $T = \sum_{i=1}^n x_i$

(6.5.15)

令  $(x) =$

1,  $U(x) > k$ , 0,  $U(x) < k$ ,

"  $(\cdot) = [(\cdot, \cdot)^2 + n\bar{x}^2 = S(\cdot) + nT^2(\cdot)]$ ,  $i = 1$

$n$

$I(x, -xy)$  与  $r=x$  相互独立. 同时, 由于  $\pi$  幻为

其中  $S(x) =$

$i=1$

$S(x)$  的严增函数, 因此  $T$  可简化为

$P\{U(x) > k(t)\} = S(x)$  因此  $E[4](f, T) | r]$  可化为

$|r] : P\{U(X) > k(T) | T\} = P\{s(x) > c(r) | r\} = a$ . 由于  $s(x)$  与  $r$  独立, 因此有

由于  $S(x)/a$

$gpc/a \sim -2$  UMPUT 为

$1 \gg$

$|T] = P\{s(X) > c(r)\} = P\{s(X) > c\} = a$ . 因此有

$> c/t_0) = a$ .

$-oc)9 \ c = a \cdot 2^{(n-1) - a}$ , 最后可得检验 (i) 的

$> o r X^{(-1, 1 - a)}$ ,

$4 > (s) = J$

$0$

根据以上求解方法, 可总结为以下定理.

定理 6.5.3 对于指数族分布 (6.5.1) 的单边检验  $H_0: \theta \leq \theta_0$  :

$S(x) < a Q_x(n-1$

$, 1 - a)$ ,

$S(x) = \int_0^t (-x)^2.$

(6.5.16)

## 6.5 多参数指数族的检验

257

$e > e^x$  若有  $f(x) = s(x, r)$ , 而  $u$  为  $s(x)$  的严增函数, 并且  $S(x)$  与  $T(X)$  相互独立, 则  $J(i)$  的 UMPUT 可表示为

II,  $S(x) > c$ ,

$y$

$0, S(x) < c$ , 其中  $C$  与  $y$  由  $E[S(x)] = a$  决定. I

(2) 关于  $\theta$  的双边检验 (ii)  $H_0: \theta = \theta_0$  :

$d$  此由定理 6.5.2 可知, (ii)' 的 UMPUT 可表示为

检验 (ii) 可转化 (ii)'  $H_0: \theta = \theta_0$

$\theta_0$  从  $\theta_0: -1/(2 < 72)$ , 因

$f(u, t) = [1,$

$U(x) < f_1(1)$  或  $U(x) > k_2(t) < U(x) < k_2(1)$ ,

$,$

$. P$

$,$

$\backslash$

$S(x) = c$ ,

其中  $k_1(4)$ , 由以下定解条件决定: (a)  $E[f_1(x, r)] = a$ ; (b)  $E[f_2(x, r)] = a$ .

$(i, r) | r] = a E[f_1(U, T)]$ .

对于定解条件 (a), 它等价于  $E[J(1 - a) | T] = 1 - a$ , 即.

$P\{U(X) < k_2(T) | T\} = 1 - a$ . 由 (6.5.15) 式可知,  $t/(4)$  为  $S(x)$  的线性严增函数, 因此上式等价于

$P\{U(X) < c_2(n \ln 1 - a)\} = 1 - a$ . 由于  $s(x)$  与  $r(x)$  独立, 上式等价于

$P\{s(X) < c_2\} = 1 - a$ . 由于  $S/d \sim f(n-1)$ , 上式可表示为

f C2/<r0

从而检验(ii)的UMPUT可表示为

$S(x) < c_1$  或  $S(x) > c_2$ ,

$c_1 < S(x) < c_2$ ,

$\int_{c_1}^{c_2} f(y) dy = 1 - \alpha$ .

(6.5.17)

且有  $EaQ[(MS)] = \alpha$ . 对于定解条件(b), 可把  $U = S + nT^2$  代入得

$EaQ[(S + nT^2)] = \alpha$ . 由于  $S(X)$  与  $r(X)$  独立, 上式等价于  
即

$EaQ[S/(S)] + nEaQ[T^2] = \alpha$ .  $EaQ[S/(S)] = \alpha$ .

$S(x) = \int_{c_1}^{c_2} f(x) dx$ , (6.5.18)

## 258 第六章参数假设检验

由于  $S/a; -/(\ll-!)$ ,  $EaQ(S) = (n-1)\alpha$ , 上式化为  $f^{[5(1 />(S))]} = (n-1)\alpha(1 - \alpha)$ .

因此检验(ii)的UMPUT可表示为(6.5.18)式, 其中  $c_1, c_2$  由 (6.5.17) 式和 (6.5.19) 式决定. 联立求解方程(6.5.17)式和 (6.5.19)式可得到  $c_1, c_2$ , 这显然是比较复杂的. 作为近似, 可根据 (6.5.17)式取

$\alpha$

$\alpha = V$

以上求解方法可总结为以下定理.

定理6.5.4 对指数族分布(6.5.1)的检验只.  $\alpha = \alpha_0$  孕②"

若存在线性函数关系  $U(x) = a(i)S(\%) + b(i)$ ,  $a(i) > 0$ , 且在②二②Q处  $S(X)$  与  $r(x)$  独立, 则检验的UMPUT可表示为

$S(x)$  或  $S(x) > c_2$ ,  $c_1/(S) = S(x) = cf(i=1, 2)$ ,

$\alpha, c_j < S(\%) < c_2$ ,

其中  $Q, y: G = 1, 2$  由条件  $EjchS)] = \alpha$  以及  $Ef_{lo}[S(/>(S))] = \alpha$  决定. 特别, 若  $S(V)$  的分布关于某个点  $w$  对称, 则有类似于定理

6.4.5的结果, 即  $\alpha = \int_{c_1}^{c_2} f(x) dx = \alpha$ , 而由  $\alpha$  决定.

关于  $a$  的其他检验可类似地由(6.5.15)式得到, 从略. (3)关于#的单边检验(i)

$H_0: \int X > f i \theta$ . 这时, 有兴趣的参数为  $m$ , 而  $a$  为多余参数. 注意, 在(6.5.14)式

$n$  中, 若取  $\theta = n/x/a^2$ ,  $U(x) = 5$  和  $(P = -1/(2a^2))$ ,  $T(x) =$

$i=1$  验不能简单地转化为  $n^L/a^2 = 6^{\theta} = n f M \theta / a^2$ , 因为其中含有多余

参数  $a$ , 不能应用定理6.5.1的公式, 也不能在此  $\theta = n^L/a^2$  处求期

:

(6.5.20)

(6.5.19)

望和定解. 但是我们可以把检验(i)转化为77.

$> 0$ , 同时把(6.5.14)式转化为

$\theta - H'$ :

, 则检

## 6.5多参数指数族的检验

259

这时可取  $e = n^{L/a^2}$ ,  $U(x) = x$  和  $-1/(2(r^2))$ ,  $T(x) =$

$(x; -$

$1 = 1 - M \theta$ ). 而检验(i)显然等价于检验(i)'  $\theta = 0$ ;  $\alpha > 0$ , 并且在  $\theta = 0$

(即  $M = 0$ ) 处定解. 根据定理6.5.1以及(6.5.4)式, 检验(i)'的UMPUT可表示为

令  $(z, Z) =$

■  
,

$$r(x) > k(t), \quad U(x) < k(t),$$

(6.5.21)

其中 $A:G$ 由 $E$ 。[令 $(\cdot, \cdot) \setminus T = a$ 决定,  $E$ 。表示在 $\theta Q = \theta$ (即处取期望, 因此有 $Eo[W] \setminus T = P\theta \setminus U(X) > k(T) \setminus T = a$ , (6.5.22)

其中 $\langle / (X) = X$ 与 $\cdot = \cdot$ ;  $(\cdot - M_0)$ 不独立。以下设法用Basu定理找  $i=1$

一个统计量 $r(x)$ , 使之成为 $u(x) = x$ 的严增函数, 并且 $I7(x)$ 与 $HI$ 独立, 则以上条件概率可化为无条件概率。由于 $\cdot(\%) = 2(\cdot - i)^2 +$

$n(x - /io)^2$ , 可取

$$1=1$$

$$D = -M_0) ' \cdot (a \sim l) ("_{Mo}) x - : \text{了卵} - /$$

(6.5.23) 以上 $IF$ (幻为 $\mathcal{E}/(\cdot)$ 的严增函数, 并且 $JF(X) \sim r(n-l)$ , 其分布与参数  $\rho$  无关。因为 $r$ 关于参数 $-l/(2cr^2)$ 为完备的充分统计量,  $JF$ 的分布与 $a$  无关, 所以由Basu定理知  $f$  与 $W$ 独立。因此以上(6.5.22)式的

定解条件可化为

$$P. \setminus U > k(T) \setminus T \setminus = P. \setminus W > c \{T\} \setminus T \setminus = P. \setminus W > c \setminus = a. \text{ 由于 } W(X) \wedge z(n-l), \text{ 因此 } c = f(n-l, l - a), \text{ 从而检验 } (i) \setminus \text{即检验 } (i) \text{ 的 UMPUT 由 (6.5.21) 式转化为}$$

杏(识)

$n$

$$(6.5.24)$$

其中 $\langle r^2(\%) = 2(\wedge \sim \wedge)^2 / (n-l)$ 为 $\langle r^2$ 的无偏估计。上式与(6.4.28)  $i=1$

式十分类似, 但是由于此处 $a^2$ 未知, 因而用其无偏估计代替, 所以 识 $(X)$ 服从 $t$ 分布, 而

(6.4.28)式中 $U(X)$ 服从正态分布。

260

## 第六章参数假设检验

(4)关于 $/x$  的双边检验(ii)  $H_0$ : :

仍然从(6.5.20)式出发, 取  $\theta - /io) / a \setminus U(x) = x$ 和 $\cdot =$

$$-l/(2cr^2), \quad T(x) = 2 - A_0) ' . \text{ 检验 } (h) \text{ 显然等价于检验 } (\cdot)' \gg = 1$$

$H_0: \theta = : \wedge \theta$ ; 并且在 $\cdot = \theta$ (即 $/x =$ 从。)处定解。根据定理6.5.2 以及(6.5.10)

式, 检验(ii)'的否定域应为 $R_+ = (x: l/(x) < \wedge(i) \text{ 或 } t/(\%) > k_2\{t\} \setminus$ 。根据前面的讨论,  $\cdot$ (幻为 $U(x) = x$ 的严增函数, 并且  $r(x)$ 与 $r(x)$ 独立, 因此以上否定域可转换为

$$R_+ = \{x: w(x) \text{ 或}$$

, 以上  $a$ 决定,  $BPc = t(n-1, 1 - a/2)$ 。从而检验(ii)', 即检验(ii)的

$\wedge(x) > C_2)$ , 又由于  $\cdot(X)$ 的分布关于 $M$ 。对称, 根据定理6.4.5 否定域可进一步转换为 $+ =$

$$| W(x) I > c I, \text{ 其中 } c \text{ 由 } EQ[\langle />(\cdot)] =$$

UMPUT 为

$$i / 2), \quad _$$

$$10, \setminus W(x) | < z(n-l, l - a/2), \quad \forall a$$

(6.5.25) 例6.4.2 (续)在问题(3)中, 假设 $a^2$ 未知, 考虑检验 $/。$ : =

$$4.55 * - * / f, \quad : / z \# 4.55.$$

解 这时就要应用(6.5.25)式的 $i$ 检验。经直接计算得 $a = 0.1017$  (比原来问题中已知的

$$a = 0.108 \text{ 要小}), x = 4.438, |x - M_0| = |4.438 - 4.55| = 0.112;$$

$$JF(x) = 2.4625. \text{ 此时 } n=5, a=0.05, \text{ 所以 } t(n-1, 1-a/$$

$$2) = z(4, 0.975) = 2.7764 > 2.4625 = W(x). \text{ 因此不能否}$$

定原假设, 即不能认为现在生产不够正常。这个结论与例6.4.2中的结论不一致, 其原因

是:例6. 4.2是在均方差 $\sigma = 0.108$ 保持不变的假定 下进行检验;而现在是在比较小的均方差 $\sigma = 0.1017$ 的条件下进行 检验. ■

注 本小节开头,我们列出了 5种关于参数 $M$ 的检验,其中(i)、(ii)、(v)都可以通过(6. 5. 23)式的识(幻得到其UMPUT.但是,对检

验(iii)和(iv)不可行.今结合检验(iii)简单说明如下:(iii) $H_0$ : 或 $z \leq \sqrt{M} < \sqrt{M}$ .

其 UMPUT 的否定域应为  $\{x: U(x) < k_2(t) \mid U(x) = x\}$ ,它 要满足  $\int_{k_1}^{k_2} f(x) dx = \alpha$ ,  $f=1,2$ . 但是我们很难得到进一步结果.事实上,根据 (6.5.23)式,应取

$T = \sum_{i=1}^n \frac{1}{\sqrt{f}} w_i(x) = -$

,  
t=

## 6.5 多参数指数族的检验

261

易见,在川处『1(1)与『独立,在A 处%(幻与r独立,但是我们 无法确定一个共同的使之在川, 处都与r独立,从而把尺+ =  $1^{\wedge} \cdot (0 < U(x) < k_2(t))$ 转化为  $R_+ = I_c, <$

$JF(x) < c_2 J$ . ■ I 表6.5. 1给出了一样本正态总体常用的假设检验及其检验统

计量.(5)应用示例

今举两个例子说明正态分布假设检验的某些应用.

例6.5.1某厂生产的铜丝一直比较稳定,其折断强度服从正态分 布,均方差为8.今对一批产 品随机抽取10个样本,其折断强度分别 为:578, 512, 570, 568, 572, 570, 570, 572, 596, 654.在  $\alpha =$

0.05水平下检验:这批铜丝的均方差是否合格.

解 要检验付.: ( $\sigma = 8 - \sigma \neq 8$ ,但 $M$ 未知.可用; $t_2$ 检验,根

据(6. 5. 18)式,检验的否定域为 $R_+ = \{x: S(x) > c_2\}$ ,其中

$n$ . 在实用上 $\sigma$ 和 $\sim$ 都是取近似值 $c$ ,

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

$a/2)$ ,  $c_2 = aS(n-1, \alpha/2)$ .本例中,  $a = 64$ ,  $a = 10$ ,  $\alpha = 0.05$ , 因而近似否定域可表示

为 $R_+ = \{x: S(x)/a \leq (9, 0.025) \text{ 或 } (9, 0.975)\}$ .根据以上数据经直接计算可得

$S(x) = 681.6$ ,  $S(x)/a = 10.65$ , 而 $(9, 0.025) = 2.70$ ,  $(9, 0.975)$

$= 19.0$ ,因此  $S(x)$ 不在否 定域.所以没有充分理由否定原假设,即没有充分理由认为这批铜 丝的 均方差不合格. ■

例6. 5. 2 某食品厂生产的一种罐头食品,标准重量为每罐 500克.今对一批产品随机抽取 10个样本,其重量为:495, 510, 505, 498, 503, 492, 502, 512, 497, 506.假 定罐头重 量服从正态分布 $N(\mu, \sigma^2)$ ,在 $\alpha = 0.02$ 水平下检验:这批产品是 否合格.

解 要检验 $H_0: \mu = 500 - \mu \neq 500$ ,但 $\sigma$ 未知.可用 $z$ 检验, 根据(6.5.25)式,检验的 否定域为  $\{x: \sqrt{W(x)} > \sqrt{(n-1, 1 - \alpha/2)}\}$ !本例中,  $\mu_0 = 500$ ,  $a = 10$ ,  $\alpha = 0.$

02,妒(%)由(6.5.23)式给

10

出.根据以上数据经直接计算可得 $\bar{x} = 502$ ,

$i=1 I=0.97$ .而 $r(9, 0.99) = 2.82 > \sqrt{W(x)}$ ,因而观察样本不在

否定域,所以没有充分理由否定原假设,即没有充分理由认为这批罐 头食品不合格. |

(%, ,  $-x)^2 = 380$ ,

262

已知

$a = \hat{\sigma}^2$

# 第六章参数假设检验 表6.5.1 一样本正态总体的假设检验

= "■0

a

未知

(r><r0 cr =rr0

弘o)2 -

T(x)/a1 <^(n) T(x)/ao >^.a/2

或<^(近似)

f ^2/cr/

J x2(n, y)dy = «

J\*)/<

fk2/a^ 2

I X (^,y)dy=1 -a

『(太) -0

W(x) <ta(n-1) | 识(4 1 >g-„/2(^-i)

"0

M=Ao

否定域应满足的条件

"(幻 >^\_a

或 鮮1郎、1

好為弘2

糾以么)]-0[:/, •(\*,)] = 1-a

已知(r^a0

M <^o

M( <M<M2 /z</Z;或(X>/jl2

(T>fT0

a<<70 cr\_<r0

i/(X) <2a 1u(x) | >z、\_a/1

Hx) ~tr-x2(n) <T] 矣o■或cr2 cr<<rt 或a>a2 i=1, 2

P-=Ao

或

弘i ∈T<(T0

<r>cr0 <r=tr0

o■矣rr,或

无解

无解

未知

(r><r2 CT] a

s(x)= s - 5(^)/(T0>^\_a(n-1) 4=1

«)2~d;v2(n-1) S(X)/ffl<Xa( -1) 或<Z/2(近似) .

名a2

=1 - a

或

er, < o- < cr2

A>^o

A <Ao

Mi <M <Mz

M <Mi 或从 >/i2 cr>tr0



$\langle r \rangle = \langle r^2 \rangle / \langle r \rangle$   
 $\text{tr}(\langle r^2 \rangle)$   
 $\text{cr} < (T)$  或  $\text{cr} > (r^2)$   
 $JT(x) =$   
 $VX_n - l(x - j_{\text{ug}}) 7F^{\wedge}$   
 $-i(n-1)$   
 检验统计量  
 $t/(x) : / - \text{弘} 0 <^{\wedge} 0$   
 $\sim 7V(0, l)$   
 $", (x) - / n$   
 $i = 1, 2$   
 $T(x) = 2(\text{龙} i\_ r(x) / <^{\wedge} x^{2i-a(n)})$   
 $S(x) \sim (TiX^2)^{\wedge} \sim 1$   
 $i = 1, 2$   
 广产? 2  
 $y(n-1, y) dy = a$   
 $', X(n-1, y) dy J S \text{ 々?}$   
 $n$   
 $n$

## 6.5 多参数指数族的检验

263

### 6.5.3 两样本正态总体的检验

假设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\mu, \sigma^2)$ , 并且二者也相互独立

• 记  $\mu, \sigma^2$  为未知参数, 则两样本联合分布可表示为

$$f(x, y) = \exp$$

$$Za, 2\sigma^2; a \in \{$$

$$m/jb^2$$

$$a^2 J$$

$$(6.5.26)$$

其中  $\mu$  与  $\sigma^2$  无关, 且有

$$Q_x = 2 \ln 2 (\mu - \bar{x})^2 + n \bar{x}^2, Q_y = 2 \ln 2 (\mu - \bar{y})^2 + n \bar{y}^2 -$$

isj I21 (1) 比较两总体方差的检验

关于比较两个正态总体方差的检验, 它们存在一致最优无偏检验, 其检验的形式可归结为:

(i)  $H_0: \sigma_1^2 = \sigma_2^2$  vs  $H_1: \sigma_1^2 > \sigma_2^2$ .

(ii)  $\sigma_1^2 = \sigma_2^2$  vs  $H_1: \sigma_1^2 > \sigma_2^2$ . 为了比较两个方差, (6.5.26) 式可化为

$$= \exp\{(\text{由一点卜点}(3)0 + \rangle + 5 \text{卜}_\text{—}\} = \exp \{eU + \langle p\{Tx + \langle p^2T^2 + (p^3T^3 - b(0,$$

$$(p)\}, (6.5.27)$$

其中  $\mu = (2(r^2)1 - ( )_\text{—}, \gg " =$

$$= -(20-2)-1, Tx =$$

$$i = l$$

$$"(P^2 = n^{\wedge} / a', T^2 = x; <^{\wedge} 3 = m / L^2 / a^2, A = 7, T = \text{、}$$

$$Q_x + Q$$

因此  $\sigma_1^2 = \sigma_2^2$  等价于  $\mu = 0$ ; 等价于  $\mu = 0$ . 为说明求解方法, 今考虑检验 (i), 该检验相当于

由定理 6.5.1, 其否定域为  $R = \{U(y) > c\}$ , 且满足  $P_{\theta_0}(U > k(T) | T) = \alpha$ , (6.5.28)

类似于 (6.5.20) 式的方法, 以下求函数  $F = F(U, T)$  使之当  $\mu = 0$  时满足: (a)  $F$  为  $t/$

的严增 (或严减) 函数; (b)  $F$  与  $r$  独立; (c)  $F$  的分布易

求. 则 (6.5.28) 式可化简为  $P_{\theta_0}(U > k(T) | T) = P_{\theta_0}(F > c)$  (为此可取

$$\wedge(xty)$$

$$(x_i - \bar{x})^2 / (n - 1) = 1$$

m

$$2 (\bar{x} - \bar{y})^2 / (n - 1)$$

$$(U - nT) / (n - 1) = 7 - U$$

$$j = 1 \quad y_{s1} \cdot$$

$$(6.5.29)$$

其中

264

## 第六章参数假设检验

因为由(6.5.27)式的记号可知

$$U = \sum_{i=1}^n (X_i - \bar{X})^2 = Z^2 - U - mT^2, \quad 7=1 \quad 7=1$$

易见F为1/的严增函数, 而且当 $\rho = 0$ 即 $\rho = a^2$ 时,  $F \sim F(n-1, m-1)$ , 其分布与参数 $\rho_1, \rho_2$ 无关, 而 $T = (T_1, T_2, T_3)$ 为完备充分统计量, 由Basu定理知, F与T独立, 因此(6.5.28)式可化简为

因 $f_{n,c} =$

$$P_W = P_{\theta_0} \{ U > k(T) \mid T = t \} = P_{\theta_0} \{ F > c \mid \rho = a \}, \text{ 所以检验问题 (i) 的 UMPUT 为}$$

$$\phi(F) =$$

$$1, \quad F(x, y) > F(n-1, 7n-1, 1-a), \quad 0, \quad F(x, y) < F(n-1, 7n-1, 1-a).$$

关于检验问题(ii), 可类似地由(6.5.29)式得到UMPUT, 只是定 解条件复杂一些, 从略. 另外, 若 $m, M_2$ 已知, 则仍可由(6.5.27)式 导出(i)和(ii)的UMPUT, 只是F统计量为见表6.5.2.

(2) 比较两总体均值的检验 关于比较两个正态总体均值的检验, 是统计学中有名的 Behrens-

$$(iii) \quad \mu_0 \neq \mu_1 : \mu_i > \mu_2. \quad (iv) \quad H_0 = \mu_1 \mu_2 \neq 0, \quad \mu_1 \neq \mu_2.$$

为了比较两个正态总体的均值, 联合密度(6.5.26)式可重新表

$$(6.5.30)$$

当时, 上式很难化简为类似于(6.5.27)式的形式, 所以无 法得到检验(iii)或(iv)的 UMPUT, 而当 $\rho = a$ 时, 上式则可化 简为

$$f(x, y) = \exp \left\{ -\frac{1}{2} (x^2 + y^2) + \frac{\rho}{2} (x^2 - y^2) \right\},$$

且未知的情形, 有关的检验问题比

Fisher问题, 对于一般的 较复杂(可参见下面(6.5.30)式), 不存在一致最优无偏检验, 在文献 中有很多专门的讨论和求解方法. 下面主要考虑一种常见的可求解的情形, 即但未知的情形. 两个均值比较的检验可归结为

### 6.5多参数指数族的检验

当 $\rho = 0$ , 即 $f_{JL} = f_{JL2}$  时有

$$T = \sum_{i=1}^n (X_i^2 + Y_i^2), \quad U = \sum_{i=1}^n (X_i^2 - Y_i^2),$$

$$265 \quad m$$

$$= 1$$

$$n f_{JL} + \frac{1}{2} (m + n) \ln 2$$

$$71! = n x + m y; \quad (p_2 = -\frac{1}{2}, \quad T_2 = x_j \text{ 为说明求解方法, 今考虑检验 (iv), 它相当于}$$

由于 $f$ 的分布关于原点对称, 因此由定理6.4.5以及定理6.5.2可知, 检验(iv)的否定域为 $T = \{ (x, y) \mid |x| > |y| \}$ , 且满足

$$P_0 \{ |T| > I \} = 1 - \alpha \quad (6.5.31) \text{ 今考虑构造一个与前面 } F \text{ 类似的统计量. 根据}$$

(6.5.30)式, 可取

$$U = \sum_{i=1}^n (X_i^2 - Y_i^2), \quad 2 \sum_{i=1}^n Y_i^2 + \frac{1}{2} (m + n) \ln 2$$

$$n + m \quad n + m$$

(6.5.32) 因此(为a的严增函数, 并且当 $\theta=0$ 时),  $t(X, Y)$ —  
其分布与参数p无关, 而 $T=(T_1, T_2)$ 为完备充分统计量, 由Basu定理知 $t(X, Y)$ 与T独立,  
因此(6.5.31)式简化为

$M < K T) \setminus T \setminus I < C! = 1 - a$  且有  $c = f(n + z_{n-2}, 1 - a/2)$ . 所以检验问题(iv)的UMPUT为  
 $f_l, < p U) = <$   
 10

关于检验问题(iii), 可完全类似地由(6.5.32)式得到其UMPUT.

由(6.5.30)式可知, 若且未知, 则无法找到合适的完备充分统计量 $T(X)$ , 因而无法求出UMPUT. 但是, 有两种常见的特殊情况

$I < Z(n + z_{n-2}, 1 - a/2)$ .  
 $|z(x, y)I > t(n + m - 2, 1 - a/2)$ , , .  
 可以求解:(1)若  $= c$  已知( $c =$  形), 则把  $a_l = c a_l$  代入(6.5.30)式, 根据与上面完全类似的  
 推导, 可 以得到类似的 $f$ 检验统计量;(2)若已知 , 则由(6.5.30)式可以  
 得到正态检验统计量: $U(x_9 y) = (\% - y)A/t$ ,  $t = a \setminus /n + a \setminus /m$ ,  $U(Xf n \sim$   
 $7V(0, 1)$ .

表6.5.2给出了两样本正态总体常用的假设检验及其检验统计量.

$l$  即为前面讨论的4 的情

266

## 第六章参数假设检验

### 表6.5.2 两样本正态总体的假设检验

"o ", 检验统计量 否定域应满足的条件

$< \text{弘} 2 \text{ Ml} > \# 2$

$\wedge 1, \text{ Ml}$

$A1 = \wedge 2 \text{ A}$

, y

)  $> z\{ \_a$

已知 弘1

$Mi < M2$

$\sim (0, 1)$

n zn

n

$y(x_i - \wedge i)^2 / n$

$t/(x, y) \Rightarrow \text{严} - \text{歹}$

J 厂ir

$t/(x$

$[/(%, y) < z a$

1 "(x, y) |  $> z) a/2$

$F(x, y) > Fx \_a(n, m)$

$F(x, y) < Fa(n, m) \quad F(x, y) > FI\_a/2(n, m)$

或  $< Fa/2(a, w)$  (近似)

以, , y) = -----

$Z (\text{人} - \text{弘} 2) 2 / \text{泔} ' \blacksquare \bullet 1$

$- F(nfm)$

未知  $Pl, 2 \text{ Ml} > \wedge 2 \text{ t}(x \text{ v} > - /mn(m+n-2) \_V m+n \text{ X}$

已知

$Mi, \text{ cr}, \wedge a 2 \text{ M2 o} - i = o - 2$

$cr, >(r^2 A <a^2$   
 $(Ti = ^2$   
 $Ml$ 為弘 $^2$   
 $Ml=^2 Ml\sim^2$   
 $x-y$   
 $ys (' \cdot ^)^2+z (人\_刃^2$   
 $\sim t(n+m-2) n$   
 $<Jo(n+m-2) 1\ll(\ll, y)1 >\ll i\sim o/2(^+^{\sim}2)$   
 $科1 <^2$   
 $[(龙, -x)^2/(n - 1)$   
 $未知 0'.介^2 tr, >a^2 F(x, y)= \text{-----} \text{尸}(x, y)>,$   
 $\text{卜}。(n-l, ?n-l)$   
 $Mi, M2$   
 $tr, <a^2 cr, -<r^2$   
 $X (Z; -y)^2/(m - 1) /-I$   
 $\sim \text{尸}(n-l, zn-l)$   
 $<Fa(n F(x, y)>Fl\_a/2(n-ltm-l)$   
 $或 < ^\ll/2 (打- 1, m - 1) (近似)$   
 $B$ 两种方法测量其溶化  
 $tr, =tr^2$   
 $(3)$ 应用示例

例6. 5.3 对冷却到-0.72丈的样品用A,  
 到0 时的潜热, 数据如下:

方法A: 79.98, 80.04, 80.02, 80.04, 80.03, 80.03, 80.04,  
 79.97, 80. 05 , 80. 03 , 80. 02, 80. 00, 80. 02 ; •  
 方法 B: 80.02, 79.94, 79.98, 79. 97, 80.03, 79. 95, 79. 97 ,  
 79. 97.

假设它们服从正态分布, 方差相等. 试检验: 两种测量方法的平均 性能是否等价(取  
 $a=0.01$ ).

$心, y > ti. a(n+zn-2$   
 $)$   
 $)$

### 6.5多参数指数族的检验

解 设两种方法测量的潜热分别记为X和I 并设 $X \sim 2V(M1, a^2)$  和 $Y \sim A^{(\sim c^2, <r^2)}$ . 要检验  
 $H_0 = jLt2*-*ZZ1: fjLi -$ 可用(6.5.32)式的  $t$  检验, 检验的否定域为  $= I (x, y) :$

$11 (x, y) | > ^ (n +zn -2, 1 - a/$

$2) j$  .本例中,  $n=13, m=8$ , 由直接计算可得 $\% = 80.02, y=79.98, Inm$

$t. -\%)^2+ (y>-y)^2//n+m-2=/(0.006912+0.006727)/19$

$/ (x$

$=0.026 79$ . 因此有

$心, y)$

$ffnr(n+m-2, l-a) = t(19, 0.995) = 2.861 < 3.32 = |右(x, y)$

267

$80. 02 \text{ 二 } 79. 98 \text{ 0. } 026 \text{ 79} \sim \sim$

否定原假设, 即两种测量方法的平均性能有显著性差异.

例 6. 5.4 为了比较正常成年男女所含红血球的差异, 对某地区

156名成年男性进行测量, 其红血球的样本均值为465.13(万/mm<sup>2</sup>), 样本均方差为54.80; 对该地区74名成年女性进行测量, 其红血球的样本均值为422.16, 样本均方差为49.20. 试检验: 该地区正常成年男女所含红血球的平均值是否有差异(取 $\alpha = 0.05$ ).

解 设该地区正常成年男女所含红血球数分别记为 $X$ 和 $Y$ , 并设  $X \sim N(\mu_1, \sigma_1^2)$  和  $Y \sim N(\mu_2, \sigma_2^2)$ . 首先要检验  $H_0: \mu_1 = \mu_2$ . 为

此, 可用(6.5.27)式的 $F$ 检验, 检验的否定域为  $F = J(X, Y): F(\%, y) >$

$1 - \alpha/2(n-1, -1)$  或  $< F_{\alpha/2}(n-1, TH-1)$ . 本例中,  $n=156$ ,  $m=74$  并  $wn$

已知  $\sum (X_i - \bar{x})^2 = 156 \times 54.82 = 468474.24$ ,  $\sum (Y_j - \bar{y})^2 = 74 \times 49.22 = 179127.36$ . 由此可得

$F(x, y) = (\sum (X_i - \bar{x})^2 / 155) / (\sum (Y_j - \bar{y})^2 / 73) = 1.25$ .  $* = 1 > = 1$

而  $F_{0.975}(155, 73) \ll 1.5$ ,  $F_{0.025}(155, 73) \ll 0.65$ . 因此观察样本不在否定域, 即不能否定  $H_0: \mu_1 = \mu_2$ . 从而我们可在  $H_0: \mu_1 = \mu_2$  的条件下进一步检验  $H_1: \mu_1 \neq \mu_2$ . 类似的检验例

6.5.3已做过. 以上结果代入(6.5.32)式可得  $t(x, y) = 5.73$ . 而  $t(n+m-2, 1 - \alpha/2) = t(228, 0.975) \ll 1.96 < 5.73 = |t(x, y)|$ . 因此应否定原假设, 即该地区正常成年男女所含红血球的平均值有显著性差异. ■

#### 6.5.4 两个二项分布总体的比较—等价性检验

比较两个总体性质是否存在差异, 在理论上、应用上都是十分重要的问题. 以上讨论了两个正态总体的比较, 下面介绍两个二项总体的比较.

因此应

#### 268 第六章参数假设检验

假设  $X_1 \sim b(n_1, p_1)$ ,  $X_2 \sim b(n_2, p_2)$ , 且独立. 考虑假设检验问题:

(1)  $H_0: p_1 = p_2$  vs  $H_1: p_1 \neq p_2$ . (ii)  $H_0: p_1 = p_2$  vs  $H_1: p_1 > p_2$ .

这种检验有广泛的应用背景, 例如, 在实际问题中,  $G$  可表示第 $Z$ 车间的次品率  $G = 1$ , 2), 要检验两个车间次品率是否相等; 或 $U$ 表示一种新的化验方法的阳性率, 而 $U$ 表示原有标准化验方法的阳性率, 要检验新方法与传统方法是否等价; 等等. 因此检验(ii)也常称为等价性检验.

为了得到检验的UMPUT, 首先必须把  $(X_1, X_2)$  的联合分布表示为指数族分布的形式

$$p(x_1, x_2) = \frac{1}{Z(\theta)} \exp\left\{\sum_{i=1}^n \theta_i x_i\right\}$$

其中  $\theta = (\theta_1, \theta_2)$ ,  $Z(\theta) = \sum_{x_1=0}^{n_1} \sum_{x_2=0}^{n_2} \exp\left\{\sum_{i=1}^n \theta_i x_i\right\}$

其中  $u = x_1 \theta_1 + x_2 \theta_2$

其中  $u = x_1 \theta_1 + x_2 \theta_2$

1, 2). 这是带有多余参数的指数族, 假设检验(i)可转化为

根据定理6.5.1, 检验(i)的UMPUT就是检验(i)'在条件分布  $U | T = X: (X_1 + X_2)$  下的UMPUT. 而这个条件分布为

$p(t; \theta) = P(X_1 + X_2 = t | X_1 + X_2 = t)$  由直接计算可得

$p(X_1 + X_2 = t)$

其中  $t = \min\{n_1, n_2\}$ ,  $Q = \max\{n_1, n_2\}$

$p(u | r; \theta)$

$\exp\{X_1 \theta_1 + X_2 \theta_2 - u\} P(X_1 + X_2 = t)$

$u = X_1 \theta_1 + X_2 \theta_2$

$\gg tZ = X_1 + X_2 \theta = \log(P_2/P_1)$ ,  $g_i = 1 - P_i (i =$

,  $y^* = \max\{0, r\}$  范围内. 由此可得条件分布为

, 因为以上求和应限制在0

## 6.5 多参数指数族的检验

269

其中  $h(u)$  为分母，因此这一条件分布为指

数族分布。检验  $H_0: \theta = \theta_0$  的否定域应为  $\{u = x' \theta : \theta \neq \theta_0\}$ ，其定解条件由  $\theta = \theta_0$  时的分布决定。

而当  $\theta = \theta_0$  时

$$p(u|\theta_0) = P_0(X_j = u | X_1 + X_2 = \dots) = \frac{1}{U} U_1 \quad (6.5.33)$$

为超几何分布，可根据  $P_0(X_1 = 1, X_2 = 1, \dots, X_n = 1) = a$  定解。

例6.5.5 为了观察吸烟与肺部肿瘤的关系，对23只小白鼠进行一年的吸烟试验，结果21只小白鼠得了肺部肿瘤。另外有32只小白鼠进行一年的对比观察，不吸烟，其中得了肺部肿瘤的有19只。试检验

吸烟对肺部肿瘤的影响。

解 假设试验组与对照组小白鼠得肺部肿瘤的概率分别为  $\theta_1$  和  $\theta_2$ 。

则问题化为假设检验问题  $H_0: \theta_1 = \theta_2$ ，其中  $n_1 = 23, n_2 = 32, x_1 = 21, x_2 = 19$ 。由此可根据 (6.5.33) 式，在条件  $t = x_1 + x_2 =$

40 下计算概率

$$P(u|t) = P_0(X_1 = u, X_2 = 40 - u)$$

其部分结果如下表所示：

$$n(u) = t - u, \quad u = 0, 1, \dots, 23.$$

(40)

U	18	19	20	21	22	23	$P(u t)$	0.182	0.42	0.096	0.01	0.033	0.60	0.007	0.39
	0.000	0.91	0.000	0.05											

由以上结果可知， $P_0(X_1 > 19 | t) \approx 0.04$ ， $P_0(X_2 > 20 | t) < 0.01$ 。因此否定域可取  $\{u = x_1 > 19 \mid (水平 \alpha = 0.04) \text{ 或 } T = x_2 > 20 \mid (水平 \alpha = 0.01)\}$ 。由于实际观察值为  $x_1 = 21$ ，因此应否定原假设，即有充

分理由认为  $\theta_1 > \theta_2$  即吸烟对导致肺部肿瘤有显著影响。■ 基于以上公式的检验称为精确条件检验，计算上比较复杂。在实用上，经常采用以下近似正态检验（当样本比较大时）。首先介绍一般情

况，由第三章极大似然估计的性质可知  $p_i = X_i/n_i, i = 1, 2$ ，且有

$$p_i \sim N(\theta_i, \theta_i(1-\theta_i)/n_i), \quad i = 1, 2, \quad y_{Pi} Q_i^{1/2}$$

易见  $p_i \sim P_2$  也有渐近正态性。因为  $\text{Var}(j_{\$1} - p_2) = p_2 g_2 / n_2$ ，因

rr) 270 第六章参数假设检验 此有

九廬  $n$ 。 (6.5.34)

记  $S = p_1 - p_2, r = A - p_2$  则由 Slutsky 定理去 1 律以及以上公式可得

$$\sqrt{n} S \xrightarrow{d} N(0, 1). \quad (6.5.35) \quad \sqrt{n} p + p^2 \sqrt{n}$$

上式表明： $Y = \sqrt{n}(p_1 - p_2) \xrightarrow{d} N(0, a^2)$ ，其中  $a^2 = p_1 q_1 / n_1 + p_2 q_2 / n_2$  视为已知。因此我们可以基于这一正态分布进行关于  $\theta = P_1 - p_2$  的各种检验；诸如  $H_0: \theta > 50, // p: \theta \leq S - H_1$ ，等等（见表 6.5.1 第一栏），但样本值只有  $r = A - p_2$  一个。

对于等价性检验 (i) 与 (ii)，相当于检验问题 (i)  $H_0: \theta = 0$ ；(ii)  $H_0: \theta = 0$  和 (ii)' 好。： $\theta = 0$  戈  $j \neq 0$ ，这时  $\theta = 0$ ，(6.5.35) 式还可进一步化简。以检验 (ii)' 为例，由表 6.5.1 可知，检验的渐近正态否定域为  $R = \{Y/a > z_{\alpha/2}\}$ 。因为当  $\theta = 0$  成立时， $\theta = 0$ ，即  $p_1 = p_2 = p$ ，因而两个总体合为一个总体， $P$  的极大似然估计为  $p = (X_1 + X_2) / (n_1 + n_2)$ 。所以 (6.5.35) 式可化简为

AA

$$U(x_1, x_2) = -\frac{1}{2} \ln \frac{p_1^{x_1} p_2^{x_2}}{p^{x_1+x_2}}$$

$$y/pq (nf' + n2_-,)$$

因此, 检验(ii)', 即检验问题(ii)的近似否定域为  $\{ \sum U(X_i, x_2) \mid \sum > -a/2 \}$ . 类似地, 可得到检验问题(i)的渐近正态否定域为  $\{ \sum U(x_1, x_2) \mid \sum > -a/2 \}$ .

例6.5.5 (续) 今应用(6.5.36)式求解假设检验问题  $H_0: P_1 = P_2$

:  $P_1 > P_2$ , 这相当于检验  $\theta > 0$ . 这时  $\sum X_i/n_j = 21/23 = 0.913$ ;  $p_2 = x_2/n_2 = 19/32 = 0.594$ ; 而当  $\theta = 0$  时有  $\sum (x_i^2/n_i) / (\sum n_i) = 40/55 = 0.727$ . 代入(6.5.36)式可得

$$0.913 - 0.594 / 0.727 \approx 0.273 \sqrt{23 + 32 - 1} = 2.623.$$

而  $z_{0.99} = 2.326 < 2.623 = U(x_1, x_2)$ , 所以应否定原假设(水平  $\alpha = 0.01$ ), 这与精确条件检验的计算结果很相近.

## 6.6 似然比检验

由以上讨论可知, 以Neyman - Pearson理论为基础的一致最优势检验有相当大的局限性, 特别是多参数复合检验, 可解决的问题较少, 限于  $(0, 1)$ . (6.5.36)

### 6.6 似然比检验

271

制较大. 本节介绍的似然比检验, 其最初的出发点与N-P引理类似, 但是在实施上比较简单, 应用范围非常广泛, 是理论上非常重要、应用上最为有效的检验方法之一, 其意义和价值与极大似然估计很类似.

#### 6.6.1 似然比检验

设  $X$  一般假设  $\theta$  为分布族, 考虑假设检验问题

(i)  $\theta \in \Theta_1$ ;  $\theta \in \Theta_0$   $\sum U(\theta) = 0$ . N-P基本引理的出发点是似然比; 类似地, 我们考虑对应于检验(i)的广义似然比如下:

$$A(x) = \sup_{\theta \in \Theta_1} L(\theta, x) / \sup_{\theta \in \Theta_0} L(\theta, x) = (L(\theta_1^*/x) / L(\theta_0^*/x)), \quad (6.6.1)$$

其中  $\theta_1^*$  和  $\theta_0^*$  分别为参数  $\theta$  在  $\Theta_1$  和  $\Theta_0$  上的极大似然估计. 易见  $A(x) \geq 1$ , 其对数似然比为  $LR = 2 \log A(x) = 2 [L(\theta_1^*) - L(\theta_0^*)]$ .

我们来考虑假设检验问题(i)的否定域  $\sum > 0$ . 若  $\theta_0^*$  不成立, 而  $\theta_1^*$  成立, 则当时  $L(\theta_1^*) / L(\theta_0^*)$  应该很小, 而当  $\theta_1^*$  成立时  $L(\theta_1^*) / L(\theta_0^*)$  应该较大. 所以广义似然比  $A(x)$  应该比较大, 因此对于假设检验问题(i), 可定

义其似然比检验的否定域  $\sum T$  与检验函数  $\sum > 0$  分别为  $R = \{x: A(x) > c\} = \{x: LR > k\}$ , (6.6.2)

$$J, \quad x \in R = \{x: A(x) > c\}, \quad \text{炎(太)} = \{x: L(\theta_1^*) / L(\theta_0^*) > c\}, \quad (6.6.3)$$

$$1.0, \quad x \notin R = \{x: A(x) < c\},$$

其中  $c$  与  $y$  的值应该根据具体的假设检验问题, 由条件  $EJ < 0$  来决定.

以上似然比检验既简单又直观, 同时也十分有效. 对于简单假设检验  $\theta = \theta_0$   $A(x) = f(x, \theta_1^*) / f(x, \theta_0^*)$  就是 N-P 引理中的似然比  $A(x)$ . 因此对于简单假设检验问题, 似然比检验(6.6.3)与N-P引理的解完全一致. 另外, 似然比检验对以下检验问题特别有效:

(ii)  $H_0: \theta = \theta_0$   $\theta \in \Theta_0$ .

其似然比统计量可简化为  $A$

$$A(\theta_0^*) = 2 [L(\theta_0^*) - L(\theta_0)] \quad (6.6.4) \quad \text{以上公式是似然比检验中最常用的形式之一. 由第五章定理5.3.8可知,}$$

### 272 第六章参数假设检验

当  $n \rightarrow +\infty$  时  $(\theta)$  的渐近分布为  $(P)$ , 其中  $P$  为参数  $\theta$  的维数. 另外, 第五章定理5.3.9的公式都可用于以上似然比统计量, 那里的公式实际上是下面定理6.6.2的公式(6.6.

14) - (6. 6. 19)的特殊情况.

例6. 6.1设及, ..., 总独立同分布,  $X_i \sim N(\mu_i, \sigma^2)$ , 求以下假设

检验问题的似然比检验: (1)  $H_0: \mu = \mu_0$

(2)  $H_1: \mu = \mu_1$ . 解正态样本的分布密度与对数似然函数分别为

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

$L(\mu) = \prod_{i=1}^n f(x_i, \mu) = \frac{1}{(\sigma \sqrt{2\pi})^n} \exp\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\}$

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

$L(\mu) = \frac{1}{(\sigma \sqrt{2\pi})^n} \exp\{-\frac{n}{2\sigma^2} (\bar{x} - \mu)^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2\}$

(1) 这时  $L(\mu_0) = \frac{1}{(\sigma \sqrt{2\pi})^n} \exp\{-\frac{n}{2\sigma^2} (\bar{x} - \mu_0)^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2\}$

$L(\mu_1) = \frac{1}{(\sigma \sqrt{2\pi})^n} \exp\{-\frac{n}{2\sigma^2} (\bar{x} - \mu_1)^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2\}$

似然比统计量为  $T = \frac{L(\mu_1)}{L(\mu_0)} = \exp\{-\frac{n}{2\sigma^2} (\bar{x} - \mu_1)^2 + \frac{n}{2\sigma^2} (\bar{x} - \mu_0)^2\}$

取  $\mu_0 = 0, \mu_1 = 1$ , 则  $T = \exp\{-\frac{n}{2\sigma^2} (\bar{x} - 1)^2 + \frac{n}{2\sigma^2} \bar{x}^2\} = \exp\{\frac{n}{2\sigma^2} (2\bar{x} - 1)^2\}$

其中

$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$

~

$n - 1$ .

6.6似然比检验

因此  $2T = \frac{L(\mu_1)}{L(\mu_0)} > k \iff \bar{x} > c$ , 其中

$c = \frac{1}{2}$

\*广

$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$

由  $E[\bar{x}] = \mu$  决定, 这一结果与N-P理论的一致最优势检验

~致。八, 且有 (2) 这时  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ ,  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$

$U_0(\bar{x}) = -\frac{n}{2\sigma^2} (\bar{x} - \mu_0)^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2$

为求  $1 + \frac{L(\mu_1)}{L(\mu_0)} > c$ , 把上式表示为  $L(\mu_1) = g(t) = t - n \log t - n + n \log$

$n, t =$

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

王)  $ZJ(\bar{x}) = n \log I + (1) \frac{1}{2} \sum_{i=1}^n (x_i - \bar{x})^2$

$-S(\bar{x}) = -\frac{n}{2\sigma^2} (\bar{x} - \mu_0)^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2$

$i=1$

由于  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ , 因此  $g(\bar{x})$  为凸函数, 从而  $zr$  可表示为

$zr = \frac{1}{2\sigma^2} (\bar{x} - \mu_0)^2 + \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2$

,  $r = \frac{1}{2\sigma^2}$ ,  $T < k$  或  $T > k$ ,

$(\bar{x} - \mu_0)^2 =$

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

其中  $r = \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2$

1),  $k$  由以下条件决定:

$r = k$

$g(\bar{x}) = g(\mu_1)$ , 其中  $g(\bar{x}) = g(\mu_1)$  (见图6.6.1) 相当于

$(\bar{x} - \mu_0)^2 = k$

P理论的一致最优无偏检验略有不同, 但在实用还是取以下近似值:  $k' = \frac{1}{2} (n$

$4 -$

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

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

$r =$

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

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



$\therefore J y - k 5. * -$

$i$

" d

a

$J - \_ . 7 >$

$v : \_ . i . *$

$(n - 1, y) dy = 1 - cr$  及

以上定解条件与N-

图6. 6. 1 以下几个例子都是应用N-P理论难以解决, 但很常见的假设检验问题。

例6. 6.2 设 $X_1, \dots, X_n$ 独立同分布, 且 $X_i \sim \text{Exp}(\lambda)$ , 求以下

假设检验问题的似然比检验: $H_0: \lambda = \lambda_0$  vs  $H_1: \lambda > \lambda_0$  解本问题用N-P理论解决比较困难, 因为它不是指数分布族。

其密度函数为

$f(x) = \lambda e^{-\lambda x}, x \geq 0$

!叫 ...

$E(X) = 1/\lambda$

$(x - 1/\lambda)^2 / \text{Var}(X)$

$k_1, k_2$ 的定解条件

$i$

-

## 274 第六章参数假设检验

设 $X_1, \dots, X_n$ 独立同分布,  $X_i \sim N(\mu, \sigma^2)$ , 记 $\bar{X} = (\bar{X}, S^2)$ , 则其极大似然估计分别为

$\bar{x} = \bar{X}(1), s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{x})^2$ ;  $A = \{ \bar{x} > \bar{x}_0, s^2 < s_0^2 \}$ . 因此有

$A_1 A_2$

其中 $\alpha$ 由

$P_{H_0}(T > k) = \alpha$ 决定, 而 $H_1$ 成立时有

$n \rightarrow \infty$

$n$

$= [i \wedge L - x(1)]$

由于 $A$ 为 $T = n\bar{X}(1)/S^2$ 的严增函数, 因此有

$R = \{ i : A(x) > c \} =$

$\Phi(X) = 1, 0,$

$\alpha > 4$

$T > k, T < k, S'$

$1 \sim r(\pm, 1), S \sim r(+$

$cr$

$\gg -5 - X(2(n-1))$ . 由此可得

$-/2(n-1)$  因此由 $E_o[(/)(J)] = a$ 可确定 $A$ 的值:

$= a$ .

因此有 $(n-1)A = F(2, 2(n-1), 1-a)$ ,  $A = (n-1)^{-1} F(2, 2(n-1), 1-a)$ .

例6.6.3 相关系数的检验. 设 $\dots, (A_i, r_j)$ 独立同分布,

$(r_j, z_2, o-i, o-2, p)$ , 要检验 $\rho$ 与 $K$ 是否相关, 即 $H_0: \rho = 0$  vs  $H_1: \rho > 0$ , 求似然比检验.

解 记 $\theta = (f \in \mathbb{R}^2, a \setminus, a \setminus, p)$ , 则 $S$ 的各分量为

$A_i = \text{又}, "2 = ; cr \setminus = \perp (J. - X)^2, ar \setminus =: \perp V (K. - 7)^2 ; n \ll i n < Tf$

$\Delta = 0, nX(1) \geq 29$

$n-1$ .

因此有 $-\hat{\alpha}(1) - \hat{\alpha}(2)$

er a

$e^{-x}$

,

## 6.6 似然比检验

275

$p = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) / \sqrt{s_X^2 s_Y^2}$

当  $p=0$  时, 除了  $s_Y^2=0$  之外, 么的其他分量与  $6$  相同, 因此由直接计算可得

由此可得  $A(x) =$  示为

其中  $y/n-2p/r(n-2, 1-a/2)$ .

$1/2 \text{Irral} a^2 4^{\sim p}$

$f(x, y, 60)$

$(1-p/2)-B/2$ , 它是  $f$  的严增函数. 因此否定域可表

$\sim p/2 \sim \langle (n-2) \rangle$  (见 Zacks, 1981, p.67). 因此有  $/I = \blacksquare$

例 6.6.4 多项分布的检验. 设  $X = (X_1, \dots, X_k) \sim \text{AfTV}(n, 7r)$ ,

,

求以下假设检验问题的似然比检验:  $7/0 : 77 = 7T0$

$77 = (7TJ - *丑 | : 7T \#7F0$ .

解多项分布的密度为

$/(X; 77) = \text{---} \blacksquare : \text{-----} \dots \%A!$

其极大似然估计分别为  $TT_i = X_i/n$ ,  $7710 = 77-Q (i=1, \dots, k)$ . 式可得

$A(x) = 4=$

$/(欠; 苾0) W UttJ$

$LR(7To)$

$R+ = | A(x) > c | = \{LR > k\}$ .

, 7rJT

代人 (6.6.4)

(6.6.5)

由定理 5.3.5 可知,  $LT(770) \rightarrow y^2(A: -1)$ , 这是因为  $[\% = 1$ , 所以参  $i=1$

数  $7r$  的自由度为  $k-1$ . (6.6.5) 式还可进一步化简, 得到拟合优度统计量, 详见下节.

■ 注似然比检验与  $N-P$  理论的一致最优势检验不同, 一般不能得到其最优性. 但是在一定正则条件下, 它有很好的大样本性质, 详见陈

希孺 (1981, 1999), Lehmann (1986, 1999), Rao (1973), Shao (1998).

276

## 第六章 参数假设检验

### 6.6.2 子集参数的似然比检验及 score 检验

在通常的假设检验问题中, 经常把参数  $\theta$  分成两部分, 为有兴

趣的参数, 并考虑其假设检验问题, 而  $\theta$  为多余参数, 上一节的例子大多亦为这种情形, 今考虑其一般情形.

假设参数  $\theta$  简记为  $\theta = (\theta_1, \theta_2)$ , 其中  $\theta_1$  为  $m$  维,  $\theta_2$  为  $p_2$  维, 考虑假设检验问题:

好  $\theta_1: \theta_1 = \theta_1^0$  (但  $\theta_2$  未知). “  $H_0$

这时  $6 > 0 = f(\theta_1, \theta_2)$ ,  $\forall \theta_2 \in I$ , 并记  $eQ = (\theta_1^0, e_2)$ ,  $(\cdot)$ ,  $e_0 = (\cdot, \&(\cdot))$ ,

其中从  $U \subset H$  为  $\theta_1^0 = \theta_1^0$  时的极大似然估计, 则由第三章的结果可知  $\theta_1^0 = \theta_1^0(4)$ , 且有

$Z(\theta_1^0) = \frac{1}{2} (L(\theta_1^0, \theta_2^0) - L_p(\theta_1^0))$ ,  $1(\theta_1^0) = 1(\theta_1^0, \theta_2^0)$ :  $L$  人  $8'$ . 其中

$LP(\theta_1^0) = L(\theta_1^0, \theta_2^0)$  为截面似然 (见第三章). 因此似然比统计量可表示为

$\Lambda(\theta_1^0) = \frac{L(\theta_1^0) - L(e_0)}{L(\theta_1^0)} = 2 \log \frac{L(\theta_1^0)}{L(e_0)}$  (6.6.6) 并且有  $ZT = \frac{1}{\sqrt{n}} \log \frac{L(\theta_1^0)}{L(e_0)}$

$(\sim) > c$ ).

以下将证明, 在一定条件下有  $LR(\theta_1^0) \rightarrow \chi^2_{(P_1)}$  (当  $n \rightarrow \infty$  时). 为此, 首先回顾第五章

5.3 节关于对数似然函数以及极大似然估计的若干重要性质(见引理5.3.4).

假设 $\theta_1, \dots, \theta_n$  总独立同分布, 且 $x, -f(x, \theta) \in \mathbb{R}^r$  为 $C-r$ 分布族 $\mathcal{A}$ 的Fisher信息阵记为 $K(\theta)$ ,  $X=(X_1, \dots, X_n)^T$ 的对数似然函数及Fisher信息阵记为 $\ell(\theta)$ 和 $I(\theta)$ . 今把有关的量按 $n$ 的阶数分块, 则由

5.3节的结果有

$$\begin{aligned} \ell(\theta) &= \ell_0(\theta) + o_p(n^{-1/2}), \quad L(\theta) \\ &= O_p(n), \\ I(\theta) &= I_0(\theta) + o_p(n), \\ \ell'(\theta) &= (-\ell_1'(\theta) - \ell_2'(\theta) + \ell_3'(\theta))' = o_p(n^{-1/2}), \\ r &\dots \end{aligned}$$

$$\begin{aligned} E[-Z(\theta)] &= I(\theta) = nI_0(\theta) = \\ &= 1 / I_X \\ /L &= L \setminus -O_P(n) \\ \bullet \bullet &1 \end{aligned}$$

$\ell_1(\theta) = V_r$  另外, 对于 $\theta = (\theta_1, \theta_2)$ , 记

$$(\theta) = (\ell_{21} \quad \ell_{22})^{-1}$$

$$, (\theta) = (\theta_1, \theta_2)$$

$$= o_p(n^{-1/2}).$$

$$p'' = \ell_{12} \quad \ell_{21} \quad \ell_{22}^{-1}$$

$$= O(n), \quad I_{12} \setminus$$

#### 6.6似然比检验

277

$$n^{1/2} \lambda >$$

$$, \lambda^2 ( \lambda^2 ) ?$$

则有 $L_1(\theta) = 0$ ,  $L_2(\theta) = 0$ . 又由于 $H_0$ 为 $H$ 时,  $\theta$ 的极大似然估计, 因此有 $L_2(\theta_0) = 0$ .

引理6.6.1 在以上假设条件和符号下, 子集参数的极大似然估计

$$A \subset \mathbb{R}^r \quad A$$

可表示为

证明

$1/p$  其中余项亦为 $p$ 维向量, 其第 $a$ 个分量可表示为 $Z$

$$\Delta A = +O(H^{-1}), \quad A_0^2 = -L^*(\theta_0) L_{21}(\theta_0) A^* + O(B^{-1}).$$

$$(6.6.7) \quad (6.6.8)$$

$$(6.6.9) \cdot$$

$$= O(1) \text{ 以及}$$

极大似然估计的渐近正态性可知, 余项 $= O(1)$ . 由于 $L_2(\theta_0) = 0$ , 因而(6.6.9)式的分块形式可表示为

$$L(\theta) = 0 \text{ 在 } \theta \text{ 处展开可得}$$

$$= L(\theta_0) + t(\theta_0)(\theta - \theta_0) + \text{余项} = 0,$$

叫 其中 $f$ 为 $\theta$ 与 $A$ 之间的一点,

$$2 \quad i \quad j \quad \text{由}$$

$$I(\theta) = [L_{21}$$

该式及下面的式子都在 $\theta$ 处取值.

$$O(1). \quad (6.6.10)$$

$$\text{因此由上式可得 } \theta = L_{21}^{-1} \Delta A + L_{22} A_0^2 +$$

$$O_p(1), \text{ 由于 } \Delta A^2 = O_p(n^{-1/2}) \text{ 所以有 } se_2 = A_0^2 + O_p(n^{-1/2}), \text{ 此即 } (6.6.8) \text{ 式. 又由 } (6.$$

6. 10)式可得

$(\cdot) = -(\mathbf{X}'\mathbf{I}(\cdot))$

由此即可得到(6. 6. 7)式. | 定理6.6.1 在以上假设条件和符号下, 子集参数的似然比统计量

其基于似然比统计量的近似否定域为  $r =$

证明(6. 6. 6)式的LRW)在何处展开有

$LR(3i0) = 2ZT(\wedge^0)A(9+(A^{\wedge})TZ(\wedge^0)A(9 + \text{余项})$

其中余项可表示为  $4-(A)-j_i i \pm ('咚(4)扣城扣响扣城, J \gg 1 j^{\wedge}ikTi$

其中f 为点与A 之间的一点,

有渐近 $X^2$ 分布, 即当 $n \rightarrow +\infty$ 时 $A_7^*(\sim)$ 对于假设检验问题

化:  $1^{(\wedge 10)} > /(?!, 1 - a) (\cdot$

由

$= 0/1)$ 以及极大似然估计

278

## 第六章参数假设检验

的渐近正态性可知, 余项 $= O_p(n^{-1/2})$ . 因此的分块形式可表示为

$LR(0|0) = 2f[A^{\wedge} + 2L^{\wedge 62} + (\cdot)tfu$

$+ 2(\wedge \wedge)tf12\wedge \wedge + (\wedge \wedge)TI22\wedge \wedge + O_p(n_{-}^{1/2})$ .

把引理6.6.1的结果以及 $Lx(eQ) = -(mA + 0/1), L2(0o) = 0,$

代入上式可得

$LjR(U = \sim 2UA)T(Z'') - ' \wedge \wedge + (A \wedge)Tf''A \wedge$

$- 2(A^{\wedge})tL12L; 21L2x + (AA)TLi2L; 21L2\} \Delta^* + (n_{-}) = -2(A^{\wedge})r(f, ,)'IA01 +$

$(A^{\wedge})t(Lh - Li2L; 21L2l)^{\wedge}$

$+ < \backslash (n_{-})$

$= -(A A)T(\text{乙} + O_p(n_{-}^{\wedge})$

$= (\wedge A^{\wedge})t(-nLn) + O_p(n_{-}) \cdot (6.6.11)$

以上L11等都在何处计值. 根据引理5.3.4的结果可知

) ). (6. 6. 12) 同时由引理5. 3.4可知 $n \sim l[-Z(\wedge^o)]^{(\wedge^o)}$ , 因而由极大似然估计的

相合性也有 $n^{-1}l$  (其中 $-Z$ 表示在式处计值. 因此也有

$-nV' - i_D, LR(0i0) = (\wedge A^{\wedge})t(-\text{乙})_{-} + (-nL'') + ap(n_{-})$ .

(6. 6. 13) 由于 $(-nL'') \rightarrow -G''(\wedge) \rightarrow$ , 因此由 Slutsky 定理以及(6.6. 12)式

可知

$(-, ) - \pm (V^{\wedge})''(0人),$

所以由(6. 6. 13)式有  $LR(\text{沒}i0) \perp * \chi^2(Pi) (\wedge \rightarrow + \infty)$ . 鷹

推论 1 对于任何  $R = < 9, o$ , 都有  $LR(0J = 2 \backslash L\{0\} - [(< 9, \wedge(0! - \wedge / (Pi))$ .

推论 2 若 $\text{2}' = 0$ , 即  $6 > 2$  为空集, 则有  $LR(e') = 2 (L(\wedge) - L(l?))$

以上定理和推论都是在独立同分布样本情形下得到的, 但是似然比 统计量的渐近; T2性对于

独立样本以及许多其他情形都是成立的. 进一步的讨论可参见陈希孺(1981, 1999),

Lehmann ( 1986, 1999), Rao (1973)和 Shao (1998)等.

另外, 由公式(6.6.6)可知, 似然比统计量需要同时计算参数没在

因此以上 $LR(U)$ 可表示为

6.6似然比检验

279

$00$ 上以及 $e$ 上的极大似然估计 $e0$ 和 $e$ , 这在一些比较复杂的情形可能会带来不少麻烦. 以下介绍似然比统计量的两种渐近等价的形式. 即 Rao的score检验统计量和Wald检验统计量, 前

者只需要计算 $\theta$ 在 $\theta_0$ 上的极大似然估计;后者只需要计算 $\theta$ 在 $\theta_0$ 上的极大似然估计, 详见 Shao(1998)。

定理6.6.2 似然比统计量 $LR(\theta_0)$ 可由score函数总(6)和Fisher 信息阵 $I(\theta_0)$ 表示为以下score统计量的形式:

$$= \frac{S(\theta_0)}{\sqrt{I(\theta_0)}}, \quad LR(\theta_0) = \frac{S(\theta_0)^2}{I(\theta_0)}, \quad S(\theta_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial \log f(X_i; \theta)}{\partial \theta} \bigg|_{\theta=\theta_0} \quad (6.6.15)$$

其中

又由(6.6.7)式可知 $\theta_0 = L^* + \theta_0/H - 1$ , 以上统计量都是在 $\theta_0$ 处计值, 因此有

$\theta_0 = G(\theta_0)$ . 此即(6.6.15)式. 又由第五章5.3节的结果可知

$$-Z(\theta_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial \log f(X_i; \theta)}{\partial \theta} \bigg|_{\theta=\theta_0} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial \log f(X_i; \theta)}{\partial \theta} \bigg|_{\theta=\theta_0} + o_p(n^{-1/2}).$$

(6.6.16) 因此有 $-Z(\theta_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial \log f(X_i; \theta)}{\partial \theta} \bigg|_{\theta=\theta_0} + o_p(n^{-1/2})$ , 该式代入以上 $LR(\theta_0)$ 表达式有

$LR(\theta_0) = \frac{(-Z(\theta_0))^2}{I(\theta_0)} = \frac{1}{n} \sum_{i=1}^n \left( \frac{\partial \log f(X_i; \theta)}{\partial \theta} \bigg|_{\theta=\theta_0} \right)^2 + o_p(1)$ . 由此即得(6.6.14)式. | 以上

(6.6.14)式和(6.6.15)式统称为score检验统计量, 是 Rao于 1947年提出的, 所以也称Rao统计量. score检验统计量在现代文献中有广泛的应用, 因为它只需要计算零假设成立时, 即 $\theta$ 在 $\theta_0$ 上的极大似然估计, 这在很多情况下它比通常的似然比统计量(6.6.6)式

更加方便. 注意, 若 $\theta_0 = \theta$ , 即 $\theta_0$ 为真值, 则以上公式与第五章的定理 5.3.9一致. 同时, 以上公式对任意的 $\theta_0$ 成立, 因此对任意的 $R$ 成立. 另外, 以下 $ITT$ 和  $IFD'$ 常称为Wald检验统计量, 是 Wald于 1943年提出的, 也是似然比统计量(6.6.6)式的一种变形, 只需要计算 $\theta$ 在 $\theta_0$ 上的极大似然估计反

证明 由定理6.6.1的证明中的(6.6.11)式可得

推论 似然比统计量 $LR(\theta_0)$ 有以下等价形式:

280

## 第六章参数假设检验

$W(\theta_0) = \frac{1}{n} \sum_{i=1}^n \left( \frac{\partial \log f(X_i; \theta)}{\partial \theta} \bigg|_{\theta=\theta_0} \right)^2$ ;

(6.6.17)

$W(\theta_0) = \frac{1}{n} \sum_{i=1}^n \left( \frac{\partial \log f(X_i; \theta)}{\partial \theta} \bigg|_{\theta=\theta_0} \right)^2$ .

(6.6.18)

证明 由(6.6.11)式以及Slutsky定理可得(6.6.18)式;再由(6.6.18)式以及(6.6.16)式即可得到(6.6.17)式. |

例6.6.5 方差齐性检验. 设 $X_1, \dots, X_n$ 为相互独立的样本. (1)若 $X_i \sim N(\mu, \sigma^2)$ , 求以下假设检验

问题的score检验统计量:  $H_0: \sigma^2 = 1$ ;  $H_1: \sigma^2 \neq 1$ ; (2)在(1)的假设下, 观测到一组数据为-0.68, 0.18, -10.07, -5.01, 11.09, 9.93, -9.67, 28.77, -1.95, -15.56, 8.05, 试检验其方差齐性.

解 (1)记 $e_i = X_i - \bar{X}$ ,  $e_i^2 = (X_i - \bar{X})^2$ , 则以上检验相当于 $H_0: \sigma^2 = 1$  vs  $H_1: \sigma^2 \neq 1$ , 可应用公式(6.6.14)求出score检验统计量. 由正态性假设可得样本的分布密度与对数似然函数分别为

当成立时, 即(1)=1时, 参数的极大似然估计为 $\mu = (\bar{X}, S^2)$ ,  $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$

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

通过对 $\mu$ 求导可得

孔 1 1 2  $(dL/d\mu) = 0$  (7), UV了(1-x)山

$dL/d\mu = 0$  2  $\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = 1$  了 汐

,

由此可得到  $\theta_0 = (1, \hat{\theta})$  处的 Fisher 信息阵为

$2 \text{ cTo } n$

、

$/(A) =$  以上结果代入 score 检验统计量的公式 (6.6.14) 可得

### 6.7 拟合优度检验

281

$sc, =$  。二少-況]  $\sim$  这时  $n = 11$ , 直接计算可得  $d = 143.685$ 。

(6.6.19) 由 (6.6.1.9) 式可得

(2)

$sc$  的值 ( $Z = 1, \dots, 11$ ) , 如下表所示:

-0.68  $sc$ . 0.547

-9.67  $sc$ , 0.067

0.18 0.550 28.77 12.465

-10.07 0.048 - 1.95 0.521

-5.01 0.375 -15.56 0.258

11.09 0.011 8.05 0.166

9.93 0.054

在上表中,  $SC_8$  的值特别大, 为 12.465, 而  $/(1, 0.99) = 6.635$ , 因此,  $\sim$  对应的假设

$H_0: a = 1$  被否定, 所以这一组数据具有非齐性的方差。同时, 由于其他义对应的假设

$H_0: c_0 = 1$  都未被否定; 因而可认为久有异于其他  $X$ , 的方差。 |

### 6.7 拟合优度检验

给定一组数据, 要验证它是否来自某一分布, 诸如正态分布、指数分布、Poisson 分布、多项分布等分布, 这就是拟合优度检验问题, 主要研究给定的样本是否服从某一既定分布。设

$X_1, \dots$  为独立同分布样本, 拟合优度检验为

(x) 解决以  $f$  上检验问题常用的方法是构造一个拟合优度检验统计量  $D =$

的分布不是  $F_0(x)$ 。

(i)  $\sim F_0$  其中 ( $\sim$  为已知分布。这一检验问题也可理解为 “久的分布是否可用 ( $\sim$  来拟合?”

。), 且当  $\sim \sim \sim$ 。 ( $\sim$  时的分布已知, 诸如

/

$/( \sim, \dots, \dots$  因此可取一个阈值  $Z)_a$ , 使  $\sim$ 。 ( $\sim$  成立时有

分布等。

$PF_q(D \wedge D_a) = a$ 。 (6.7.1)  $a = 0.01$  或  $0.05$  等值。由此可构造否定域为

$/ ? = \{ \% : / \} (\%, \dots, F_0) \wedge D_a \mid$ , (6.7.2) 其中  $Z > a$  为已知, 诸如  $/(?, 1 - a)$  等。

拟合优度则可理解为以上检验的  $p$ -值, 例如  $X$  的具体一组抽

样为  $X_1, \dots, X_n$ , 可计算出  $D_0 = P(X_1, \dots$

, 若  $D_0 \wedge D_a$  则否定  $H_0$ 。; 若  $D_0 < D_a$ , 则不否定  $H_0$ 。而本次抽样的

$P$ -值为

$P(Z)_0 = PF_0(D(X_1, \dots, X_n; r_0) > D_0 \mid$ ,

易见

,

### 282 第六章参数假设检验

因此, 若  $p(D_0) \wedge a$ , 则否定  $H_0$ 。; 若  $p(P_0) > a$ , 则不否定  $H_0$ 。从拟合

的观点来看, 也可作如下直观解释: 即  $P(D_0)$  越小, 用 ( $\sim$  来拟合  $A$

的分布越不合适;  $P(P_0)$  越大, 则用  $h$  ( $\sim$  来拟合  $A$  的分布越合适。因此,  $p(Z > a)$  可理解为

用 $\chi^2$ 来拟合 $X$ 的分布的合适程度，即拟合优度。

拟合优度检验与多项分布以及多项分布的检验有密切关系。本节的主要内容就是关于多项分布的检验及其应用。

### 6.7.1 拟合优度检验与多项分布检验

例6.7.1 孟德尔豌豆杂交试验。在此试验中，孟德尔同时考虑豌豆的颜色和形状，颜色有绿色、黄色，形状有圆形光滑和皱皮之分，经杂交以后，根据孟德尔遗传学理论，其后代的比例应为9:3:3:1。对杂交试验的豌豆进行了556次观测，其结果如下表所示，要检验以上9:3:

3:1 律是否成立。

黄-圆 黄-皱 绿-圆 绿-皱 理论比例 9 3 3 1

观测值 315 108 101 32

这可看作一个多项分布的检验问题，共有4个状态：黄-圆、黄-

皱、绿-圆、绿-皱；做了  $n=556$ 次试验，检验其概率是否为 $\frac{9}{16}, \frac{3}{16}, \frac{3}{16}, \frac{1}{16}$ ，就是

17, 16?记  $a=556$ ,  $\hat{V} = (\hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4)^T = (315/556, 108/556, 101/556, 32/556)^T$  要检验

$H_0: \hat{V} \sim \text{Mult}(556; \text{备} A \text{ 各占一})$  7V的分布不是

$9/16, 3/16, 3/16, 1/16$

$9/16, 3/16, 3/16, 1/16$

5W 其解法见下一小节。

以下例子说明，多项分布检验可用于很广泛的情形，并不仅限于多项分布本身的检验。

MN 556

$I^A$ .

例6.7.2 设 $X_1, \dots, X_n$ 独立同分布，考虑检验

,

的分布不是  $F_0(x)$  (其中  $F_0$  已知)。

(i)  $H_0: X_j \sim F_0(x)$  这一检验可近似化为多项分布的检验，通常的做法如下：把数轴  $(-\infty, \infty)$  划分为 $k$ 个区间， $(-\infty, \infty) = A_1 \cup A_2 \cup \dots \cup A_k$ ，其中 $n_i =$

$I : X$

1

### 6.7拟合优度检验 283

$(-\infty, \infty]$ ,  $A = (\sim, a_2], \dots, A_k = (\sim_{k-1}, \infty)$ ，并记 $a_0 = -\infty$ ,  $a_k = \infty$ 。

由于 $F_0$ 为已知分布，可求出

$p_{0i} = P\{F_0(X) \in A_i\} = F_0(a_i) - F_0(a_{i-1}), i=1, \dots, k. (6.7.3)$

同时可统计出观察值 $X_1, \dots, X_n$ 落入区间 $A_i$ 中的样本个数 $n_i$ ，如下表所示。

$\begin{matrix} 1 & \dots & i & \dots & k \\ n_1 & \dots & n_i & \dots & n_k \end{matrix}$   $\begin{matrix} p_{01} & p_{02} & \dots & p_{0k} \end{matrix}$

则 $N = (n_1, \dots, n_k)$ 服从多项分布，并可进行以下检验：(1)  $H_0: N \sim$

$MN(n; p_{01}, \dots, p_{0k})$   $H_1: N$ 不服从 $MN(n; p_{01}, \dots, p_{0k})$ 。

检验(i)可视为上述检验(i)的一个近似检验。因为若 $X \sim F_0(x)$ 成立，则必有 $\sqrt{n}(\hat{F}_n - F_0)/\sqrt{V(n; p_{01}, \dots, p_{0k})} \xrightarrow{d} N(0, 1)$ ，所以若检验(i)被否定，即 $\sqrt{n}(\hat{F}_n - F_0)/\sqrt{V(n; p_{01}, \dots, p_{0k})} \not\xrightarrow{d} N(0, 1)$ ，则 $X \sim F_0(x)$ 不可能成立。这种检验方法在实际问题中经常被采用(注意，若(i)的 $H_0$ 未被否定，只能说明我们没有充分理由否定假设 $X \sim F_0(x)$ )。画

例6.7.3 含参数情形的检验，问题同例6.7.2，但假设 $X \sim F_0(x) = F(x, \theta)$ ， $\theta$ 是未知参数；以正态为例，若 $F(X|\theta) = N(\mu, \sigma^2)$ ，其中 $\sigma^2$ 未知；则检验(i)化为

(ii)  $H_0: X \sim N(\mu, \sigma^2)$   $H_1: X$ 不服从 $N(\mu, \sigma^2)$ 。

这一检验可用来验证样本 $X_1, \dots, X_n$ 是否来自某一正态分布，即检验 $A$

的分布是否属于正态分布族。这一检验化为多项分布检验的做法与例6.7.2一样，但是

(6.7.3)式化为

$Pf_0(X' e /, ) = p_0, (M, a_2), i = i, \dots, k.$

即其中含有未知参数 $\theta$ . 因此检验 $(i)$ 可转化为以下检验  $(ii)' H_0 : N - MN(n; p_01 (/x, a_2), \dots, p_{0i} (弘, tr_2))$ .

这是一个含参数的多项分布检验, 其解法见6.7.3. | 因此, 多项分布的检验可以广泛用于其他分布的拟合优度检验

问题.

本节的主要结果为拟合优度检验的/ 检验统计量, 都是基于渐近 分布得到的. 而多项分布的渐近性质是其基础, 以下首先介绍这一性 质. 关于多项分布的基本公式可参见第一章.

引理6.7.1设 $V = (\%, \dots, 7V, )$ 服从多项分布 $MN(n, n)$ ,  $77 = (7T, , \dots, 77\backslash)$   $f$ , 则当  $n \rightarrow + \infty$  时有

即

$1n$

$-(r-Er) \sim 7V(0, Var(X))$ ,  $y_n y=1$

$T_{,,} = y/n - n7r) - ^{-} * TV(0, JS)$ ,  $X-g-7T77T$ .

284

## 第六章参数假设检验

(1) (2)

1

$n n''$  记  $R$  为以下统计量:

;

$\perp(jV-jut) \rightarrow 0(a.e.)$  或  $\perp(\lambda: -JW, ) \rightarrow 0(a.e.) i=1, \dots, k.$

由(6.7.4)式  $Y_n$ 的定义可得“  $y_n = g^{Tn}$

$A \setminus -mr\}$

$E(r) = 0;$

$yV- -nir1$

$Nk \sim ^{\wedge} T /$

(6. 7.4)

(6.7.5)

$Ylt =$

则有 其中

个多点分布之和, 即

$2V= Z \lambda, f=(\lambda, \dots, \lambda), ;=1, \dots, \lambda, (6.7.7) ;=i$

其中 $x_1, x_2, \dots, ;r$ 独立同分布,  $x_1 \sim mtv(i, 77)$ 为多点分布, 且有  $E(X_1) = 7r$ ,  $Var(X_1) = ^{-} g-7777T$ .

$1n$

(1) 根据独立同分布情形的大数定律可知 $\perp y (x_j -$

$n jT\{$

$0(a.e.)$ , 由(6.7.7)式及  $E(T) = 77$  即得 $\perp(TV-柳) \rightarrow 0(a. e.)$ .  $n$

(2) 根据独立同分布情形的中心极限定理可知

$7V(0, A)$ ,  $A=Ik-棘 T$ ,

$\lambda=(7^{\wedge}7, \dots, a/^T)T=茗_{1/2}开订, , 茗=diag(77'' \dots, TrJ. (6.7.6)$  证明 由于/

$V=(/V, , ", A\setminus)T$ 服从多项分布 $W(n, 7r)$ , 它可视为 $n$

$(0, +為4) = /V(0, A)$ .

其中 $4 =发 \cdot 了(茗_{7777T})茗 \cdot + =Ik -艸飞. I$

## 6·7. 2 多项分布检验的Pearson定理

假设观察值为/V= 要检验它是否服从某一多项分



布，前一小节的例6.7.1和例6.7.2都是这种情形。因此我们考虑 检验：

(i)  $H_0: N \sim MN(71, 7T) / 1:N$ 不服从 $MN(n, rr)$ ，其中 $7r$ 为已知。

exj)→

#### 6.7拟合优度检验

285

为了得到检验问题的否定域，必须寻找一个统计量，使得当 $H_0$ 成立时该统计量有方便的分

布。由引理6.7.1可知，若  $N \sim MN(n, rr)$ ，

则 $rn \sim V(0, 4)$ 。但是下面将看到，这是一个多维的退化正态分布，使用上不方便。很自然会想到会有渐近 $\chi^2$ 分布，并由此可得到检验的否定域。下面给出具体的证明。

引理6.7.2 若 $Y \sim N(QJA)^*k$  维正态分布，其中 $A$ 为投影阵，秩为 $r$ ， $a$  的谱分解为 $a =$

$r \text{diag}(1/r, 0) r^T$ ，其中 $r = (r_1, \dots, r_2)$ 为正交矩

阵，阶。 $z = r^T r$ 可分解为 $Z^T = (Z_1^T, Z_2^T)$ ，其中  $z_1 \sim V(0, 1)$ ，

$z_2$ 为退化分布 $P(Z_2=0)=1$ ，且有

$K = r_1 z_1^T + r_2 z_2^T = r_1 z_1^T + 0$  (6.7.8)

证明 可参见定理i.3.i。由假设可知， $z = r^T r \sim V(0, r^T A r)$ ，

而 $r^T A r = \text{diag}(7r, 0)$ ，因此 $z$ 的第一分量 $z_1 \sim V(0, 1)$ ，而 $z$ 的第二分

量为退化分布 $p(z_2=0)=1$ ，又由于 $r = r z = r_1 z_1 + r_2 z_2$ ，即得 (6.7.8) 式。

定理6.7.1 (Pearson)对于给定的概率向量 $\pi$ ，若  $T \sim V(n, \pi \pi^T)$ ，则有

(6.7.9)  $-a \mid$ 。

$A = -n n A^2$ ，

$I \sim$

$, = i \setminus y / r_i T T_i$  )

因此假设检验问题(i)的否定域可表示为 $R_+ = \{K_n$  证明由引理6.7.1可知

-----  $-X$  (女-1),  $*+00$ 。

>

$\pi = 6$  坨， $rn \sim (0, A)$ 。容易验证， $\pi = 1k - \pi'$  为投影阵，因为 $4T = A$ ， $A^2 = d - \pi'$  ) ( $Ik -$

$i=1$

$\text{tr}(A) = k - h$  设 $A$ 的谱分解为 $A = r \text{diag}(Z_1^T, 0) r^T$ ，并作变换 $A = r^T y n$ ，则由引理

6.7.2 可知  $+ r_2 z_2^T, z_j$  )，其中  $z \setminus, Ik -'$  )， $Z_2 \xrightarrow{P} 0 (n \rightarrow + \infty)$ 。因此有

$K_n = r_1 r_n = (r_2 z_n)^T (r_2 z_n) = z : z_n = + z_j z_2^T$ 。

由于 $Z_2 \xrightarrow{P} 0, Z_1^T (n \rightarrow + \infty)$ ，因此由Slutsky定理可得 (6.7.9) 式。■

例6.7.1 (续)本问题的检验前面已经给出，其中 $n = 556$ ， $N : (m ; V4) T = (315, 108,$

$101, 32) t, p$  (备各吾，占)， $k=4$ 。

#  $T$ ) = 忍 - 种 \ 注意

$k$

$= 1$  ) • 因此 $A$ 的秩为 $\text{rk}(\pi) =$

代人Pearson公式(6.7.9)式可得

286

#### 第六章参数假设检验

火， $\chi^2(3)$ ，而 $\chi^2(3, 0.95) = 7.81$  尺， $= 0.47$ ，因此 $2V$ 不属于否定域，从而不能

否定 $H_0$ 。即不能否定9:3:3:1律。另外，可算出相应的 拟合优度，即 $P$ -值，它等于 $P1/$

$(3) > 0.47$   $1 - 0.90$ ，拟合优度很大，所以可认为9:3:3:1律十分可信。■

例 6.7.4 对某地区男性公民的疾病调查中发现，在一段时间内男 性公民因“无先兆突发性心脏病”而死亡的人数为65人。若按星期几 分类，死亡人数结果如下表所示。试检验：该地区的男性公民因“无 先兆突发性心脏病”而死亡是否与星期几有关( $\alpha = 0.01$ )。

星期一 死亡人数 22 7

三 四 五 六 日 6 13 5 4 6

解 记  $71=65$ ,  $7V = (W, 2V_2, 2V_3, 2V_4, /V_5, /V_6, iV_7)T =$

$(22, 7, 6, 13, 5, 4, 6)T$ , 问题就是要检验

$H_0: W \sim MA^{(65; 1/7, 1/7, 1/7, 1/7, 1/7, 1/7, 1/7)} : /V$  不服从该分布.

其中  $77 = (1/7, 1/7, 1/7, 1/7, 1/7, 1/7, 1/7)T$ ,  $k=7$ . 代入 Pearson 公式 (6.7.9) 式可得  $Kn = 26.76$ , 而  $\chi^2_{(6)}(0.99) = 16.812 < Kn = 26.76$ . 因此  $/V$  属于否定域, 应否定

原假设  $H_0$ , 即认为该地区的男性公民因“无

先兆突发性心脏病”而死亡这一事件的发生, 在一星期中不是等可能性的 (由上表可知, 假日后的第一天, 即星期一的死亡率特别高).

### 6.7.3 含参数多项分布的检验及 Fisher 定理

与例 6.7.3 类似, 考虑以下假设检验问题:

(ii)  $/V \sim W(n, 7r) - 771 : /V$  不服从该分布.

其中  $N = (N^1, \dots, N^k)T$ ,  $TT(0) = (7Ti(0))$  参数  $\theta = (\theta^1, \dots, \theta^k)T$  未知, 且  $p < k$ . 这类检验比前一小节讨论的,  $77$  已知时的检验 (i) 有

更广泛的应用, 也更加符合实际情况 (见下面例题).

首先提醒读者注意, 对于任意固定的  $\theta$  以及  $77 = 77(\theta)$ , 前一小节中有关多项分布的性质都成立, 例如, 若  $N(\theta)$ , 则对任意固定的  $\theta$  有

$/v = \sum_{r=1}^k r, r \sim \text{觀}(1, \text{訂}(\theta))$ ,

#### 6.7 拟合优度检验

287

$E(/V) = \sum_{r=1}^k r, \text{Var}(/V) = nS(\theta) = n \sum_{r=1}^k r^2 g_r(\theta) - (n \sum_{r=1}^k r g_r(\theta))^2$ . 特别, 引理 6.7.1

和 6.7.2 在  $77(\theta)$  处都成立, 即当  $n \rightarrow +\infty$  时有

$\frac{1}{n} \sum_{r=1}^k N_r - m_r(\theta) \rightarrow 0 (a.e.)$ ,  $V_0$ ,

其中  $A(\theta) = (A^1(\theta), \dots, A^k(\theta))T$ . 同时也有

但是, 在检验 (ii) 中, 参数  $\theta$  是未知的, 很自然的想法是用  $\theta$  的某个估计代替. 以下将证明, 若坑成立, 则有

— “、T

$N^k$ 、

$Y_n(\theta) = \sqrt{n} V(\theta, A(\theta))$

,  $\sqrt{n} V' - m r^e)$

义  $2(\text{人} \text{戶} - 1)$ ,  $n \rightarrow +\infty$

其中  $\hat{\theta}$  为参数  $\theta$  的极大似然估计. 该式与 Pearson 公式 (6.7.9) 式相比,  $\chi^2$  分布的自由度减少了  $p$  个, 这是因为参数  $77$  变为多了  $p$  个

约束. 以上公式的证明显然与参数  $\theta$  的极大似然估计有密切关系, 因此我

们首先要了解含参数多项分布  $N \sim MN(n, 7r(\theta))$  中, 参数  $\theta$  的极大似然估计的基本性质. 下面首先介绍这方面的内容.

若  $N \sim W(n, 7r(\theta))$ , 则  $/V$  关于参数  $\theta$  的密度函数和对数似然函数可分别表示为

$P(n; \theta) = \frac{n!}{71! \dots JI!} \prod_{r=1}^k \frac{7r^{\theta_r}}{7r!} \exp\{-7T(\theta)\}$ ,  $71, \dots, JI$ .

(6.7.10)

其中  $C$  为与  $\theta$  无关的常数. 参数  $\theta$  的极大似然估计  $\hat{\theta}$  的性质与对数似然函数  $l(\theta)$  及其前几阶导数有关. 由 (6.7.10) 式可知, 这些都涉及到  $7T(\sim)$  的导数, 今记

$^L l \backslash \text{den}$

引理 6.7.3 矩阵  $D(\theta)$ ,  $B(\theta)$  在任何  $\theta$  处满足关系:  $l r D = 0$ ,  $^L l r D = 0$ ,  $AB = B$ ,  $APB = PB$ .

其中  $1 = (1, \dots, 1)^T$ ,  $P_b = B(B^T B)^{-1} B^T$  是  $B$  的投影阵.

## 288 第六章参数假设检验

证明由  $I(\theta) = 1$  对  $\theta$  求导可得

$i=1, \dots, p$  由此即得  $r(\theta) = 0$ . 由  $b$  的定义有

$$i/jTB = i//Tg \sim D = 1 \quad fD = 0,$$

$$AB = (Ik - \hat{g})g - yD = g - \hat{D} = By$$

$APb \sim lBy = 5(BtB)^{-1} Bt = Ps$ . | 引理 6.7.4 对于多项分布  $\Pi_n$  (的), 假设  $\theta$  的参数空间  $\Theta$  为  $R^p$  上的开集;  $\eta$  在参数空间  $\Theta$  上关于  $\theta$  存在三阶以上连续偏导数;  $B$  为列满秩矩阵.  $\theta$  的极大似然估计记为  $\hat{\theta}$  且假定其真参数

$\theta_0$  为  $\Theta$  的内点, 则有

$$(1) \quad \hat{\theta}(\theta) = ST(a) y_n(a) \eta V(0, 5t(\theta) S(\theta)); \quad y/n$$

$$(2) \quad \eta = BT(e) B(e) + op(o_p(-1/2)), \quad e[-Z(\theta)] = (\eta) S(\theta), \quad \pm z/3)(\theta) = ap(1);$$

$Tl$

$A$

$$13L1 * \eta J dTT, \quad n, 7ri$$

代人上式有

$$2k \quad dTT \text{ 人 } \theta,$$

$$(3) \quad A\theta = \theta - \theta_0 = \eta[(B^T B)^{-1} B^T Y_j \theta = f\theta + op(n^{-1})]. \quad y_n$$

证明 由  $A(\theta)$  的公式 (6.7.10) 式对  $\theta$  求导可得

$$(6.7.11)$$

$=$

$$I \quad d\theta a \quad 1 \quad 1 \quad d\theta a$$

$$\Delta a, \dots) = 42$$

其中  $\hat{\theta}$  和  $K$  分别为  $b(\theta)$  和  $y_n(w)$  的分量, 因此  $\pm L(e) = bt(\theta) y_n(\theta)$ .  $y_n$

而由 (6.7.5) 式及引理 6.7.3 可得  $\hat{\theta}(\theta) Y_n(\theta) - \eta N(0, BrT(e) A(\theta) B(\theta)^T)$

$$\eta N(Q, BT(e) B(\theta)),$$

卵上入上

$-I$

$:$

$\perp$

## 6.7 拟合优度检验

289

因此可得结论 (1). 同理对上式继续对  $\theta$  求导, 通过类似的计算与推导 可得 (2), 细节从略.

为证明 (3), 对  $\theta = \theta_0$  在  $\theta_0$  处展开可得

$$K(e) = r: (\hat{\theta}) r n(\hat{\theta}) =$$

$$- n^7 T i(3) \blacksquare$$

$$L(\hat{\theta}_0) + Z(6 > 0) A < 9 + \text{余项} = 0,$$

$$(6.7.12)$$

其中余项为  $P$  维向量, 其第  $a$  个分量可表示为  $1pp$

$$丁 Z Z n \sim 1 \quad i=1 \quad ;=1$$

$$y/n^e j,$$

其中  $\theta_0$  为  $\Theta$  与  $\theta_0$  之间的一点, 由极大似然估计的渐近正态性可知, 余项  $= O_p(1/n)$ . 利用以上

结果以及 (1)、(2), (6.7.12) 式可表示为

$$1 \cdot Z(\hat{\theta}_0)$$

$$\sim L(\theta_0) + \eta A < 9 + op(n^{-1/2}) = 0,$$

$$4^{\wedge} n$$

$$STr \quad -(BTB + op(n^{-1})) v^{\wedge} A^{\wedge} + op(n^{-1}) = 0,$$

(BTB)  $77TA19=BTyn +\%(n-+)$ . 以上各式均在氏处计值, 由此即得(6.7.11)式:  
 $y/n$

定理6.7.2 (Fisher)在引理6.7.4的条件下, 有

(6.7.13) 证明 为方便起见, 今记 $X^Y$ 表示随机变量 $X$ 和 $r$ 有相同的渐近分布, 并记  $77, = 7Tf( ? )$ ,  $77, = \bullet 77,(\%)$ , 则由 Slutsky 定理有  
 $Kn(6) = f f [况- (10)\(<\%)] - [n7r/6) -TiTr/j\%)]$

$X^{\otimes} a(\text{公}, -\%) \blacksquare$

$^{\wedge}, .(^{\wedge}0)]2 \text{订} i (4)$

$=S \ 1$

4

$i=1$

$k \ rTV.$

$E(^{\wedge}-^{\wedge})2^{(Kn-t/n)T(K.-t7J), (6.7.14) \ i=1$

其中 $y_{,,}$ 和 $u_n$ 的分量分别为

$V -n7r\ll(a_{.}) \ rr \ _J$

$1ni \ \text{厂. 丁}, Uni=$

$Vn7ri() \ y/niTi(0Q)$

由引理6.7.1可知,  $Yn \ ^Af(0, A(^{\wedge}o))$ , 以下主要推导%的渐近分 布,  $Uni$ 可展开为

$\blacksquare \ \text{二}$

1

290

第六章参数假设检验

$7(2\text{卵}, \text{' } 1 \text{么} \text{dead}0b \ i \ [n^6a$

$=-$

由矩阵 $B$  的定义可知, 上式的向量形式可表示为  $Un : \text{仰冰})$ ,  $0 \text{人} nS$

$= B(BJB) -]BTYn+0p(n'^{\wedge})$  因此由(6.7.14)式可知

$Kn(e)(yn -pBrn)\Yn -pByn)=$  由于 $y_{,,} \text{LN}(0, A(0o))$ , 由引理6.7.2的(6.7.8)式可知 $Ya^{\wedge}rtZt9$  代入上式有

(6.7.15) 其中 $Z, 4 = r\}(/ \text{厂} \text{匕})rr$  以下证明矩阵4为投影阵,

且秩为 $k-P-1$ , 显然有 $at=a$ , 另外  $W=r, -r^{\wedge}A-pjr, .$

. 这些结果代人A2

$42 = r^{\wedge}(/t-pB)(A-APB)r1 = r:(/fc-pB)(r1r;r1 -pBrj =rl(ik-pB)dh-pB)rl =$   
 儿

由引理6.7.3可知 $r, r^{\wedge} =4$ ,  $AB = B$ ,  $APB =PB$

可得

因此4为投影阵, 其秩为  $rk(4)=tr(4)=trirj'r, -TTi^{\wedge})$

$=trIA-1-^{\wedge}PBI =A:-1-tr(PB)=k-l-p$ . 因此由(6.7.15)式以及正态向量二次型的性质(见1.4节)可知 $K(J)=$

$zMi$

例6.7.5

t

I). I Poisson分布检验.观察了共1 008个细胞单位, 并测试

2(々

每个细胞单位所含白血球的个数, 要检验它是否服从Poisson分布.数 据见下表:

k01234567

nk 64 171 239 220 155 83 46 20 6 3 0 1 1 008

其中 $A'$ 表示细胞单位含白血球的个数, 久表示1 008个观测单位中, 含 &个白血球的细胞单

位的个数.

$PB(e_0)Y_n(0_0).$

8 9 10 11

总数

$-pB)y.$

6.7拟合优度检验

291

解 这相当于观测了  $X_1, X_2, \dots, X_{1008}$  个独立同分布样本,  $i$  表示第  $i$  个细胞单位含有白血球的个数, 取值为  $0, 1, 2, \dots, 11, \dots$  等等; 久则表示观测样本中取值为  $i$  的样本个数; 问题为

(i)  $\sim$  某一-Poisson 分布  $P(A * X_i)$  不服从 Poisson 分布.

问题可化为含参数  $A$  的多项分布检验问题, 与例6.7.3类似, 把 1008个观测值分为10个组, 其中  $7_i =$  (含有  $Z$  个白血球  $1=0, 1, \dots, 8$ ,  $I_g =$  |含有多9个白血球|, 并记  $\%_i$  表示观测值中属于  $i$  的个数, 则有  $N_i = n_i, i=0, \dots, 8; N_9=4$ . 若  $X_i \sim P(A)$ , 则有

$U_i = PA(X_i = 0) = e^{-A} A^{7_i} / 7_i!$ ,  $i=0, 1, \dots, 8$ .

$779(A) = \prod_{i=0}^8 \frac{A^{7_i}}{7_i!} = 1$

记  $7V = (7V_0, 7V_1, \dots, 7V_9)^T$ ,  $77(A) = (77_0(A), 77_1(A), \dots, 77_9(A))^T$ , 则可

考虑以下多项分布的检验:

(i)'  $H_0: N \sim MN(1008, 77(A))$ ;  $7V$  不服从  $MN(1008, 77(A))$ .

对此检验, 可应用Fisher定理, 其中  $n = 1008$ ,  $k = 10$ ,  $p = 1$ ,  $Kn(A) = Wh-1)^2(8)$ , 即

尺,  $(\chi^2) = I / \chi^2$

$N_t - n_i T_j(A) \sim \chi^2(8)$ ,  $(A)$

$\sim \chi^2(8)$ .

$A$  的计算比较复杂, 通常可取原来样本  $1, \dots, 1008$ ,  $X_i \sim P(A)$  中  $A$  的极大似然估计作为近似. 这时有  $A = 2.82$ , 由此可算出  $Kn(A) = 2.614$ , 而  $(8, 0.95) = 15.51$ , 因此不能否定  $7/$ . 其拟合优度, 即  $P$ -值  $= P(\chi^2(8) > 2.614) = 0.95$  很大, 因而开. 可信, 所以亦可认为

可信. ■ 6.7.4 应用: 列联表及其等价性和独立性检验

多项分布的检验对于列联表有重要的应用, 以下通过两个例子说明其意义.

设又有  $r$  个状态  $4, \dots, 人$ ;  $\wedge$  有 5 个状态  $B, \dots$ , 则  $(久, 久)$  有  $ns$  个状态, 假设进行了  $\langle$  次试验, 出现的次数为  $\wedge$ , 概率为  $\sim$ , 并假设

$z$ :

记  $V = (V_1, \dots, V_s)^T$

,  $7T = (7r_1, \dots, 7r_s)^T$

服从多项分布  $N = v$  (称为  $2$  维  $r \times s$  列联表:

$2 \times 2$   $N_u = n - 1 = 1$   $\neq 1$

$j=1, \dots, s)$  为  $k=r \times s$  维向量, 则  $V/V$  的观测值排列起来就是一个列联表,

把

二

熏

一

、扁

292

第六章参数假设检验

方-

"22(苐22)

々.2(苐.2) n

N、 、 n2'

B2 ..

N'2 .. . N't N12 ..

Nrl ... Nrt

基于多项分布的列联表有非常广泛的实际应用和丰富的统计推断问题，以下介绍其中两个常见的实例。

例6. 7.6 等价性检验.假设对  $n$  个病人志愿者用两种方法进行化验(例如化验肝炎等).应用新方法(例如为验尿)和标准化方法(例如为验血)进行化验，其结果为阳性、阴性的人数 $n_{ij}$ 和概率  $p_{ij}$ 列于下表：

标准方法 +和

新 + (苐")

N'.h、.)

$\sum_{i=1}^2 (\sum_{j=1}^2) \text{ 法 和 } i(7T \cdot !)$

" 2 . (7T2 . )

其中 $n_{11}$ 表示化验后新方法为阳性，标准方法亦为阳性的人数； $n_{12}$ 表示化验后新方法为阳性，标准方法为阴性的人数；其他类似。 $\sim$ ：表示新方法化验为阳性，标准方法化验亦为阳性的概率； $p_{12}$ 表示新方法 化验为阳性，标准方法化验为阴性的概率；而  $\sum_{j=1}^2 p_{1j} = p_{1\cdot}$  则表示新方法化验为阳性的概率.同理 $p_{\cdot 1} = p_{\cdot 1} + p_{\cdot 2}$ 表示标准方法化验为阳性的概率，其他类似.等价性检验问题就是要检验： $p_{11}/p_{12} = p_{21}/p_{22}$ ，新方法化验和标准方法化验效果是否等价，因此问题可化为

由于  $p_{11}/p_{12} = p_{21}/p_{22}$ ， $p_{11}p_{22} = p_{12}p_{21}$  所以上述检验问题等价于

$\frac{p_{11}p_{22}}{p_{12}p_{21}} = 1$

$h_1$ ：贫12六开21. (6.7.17) 本问题为一个 $2 \times 2=4$ 格列联表， $N = \sum_{i,j} n_{ij}$

$MN(\cdot, w, 77)$ 为四项分布， $77=(7F11, rrl2, 7r21, tt22)T$ .应用 Fisher定理可得

定理6. 7. 3 等价性检验(6. 7. 16)或(6. 7. 17)的检验统计量为

$\chi^2 = \sum_{i,j} \frac{(n_{ij} - \frac{n_{i\cdot} n_{\cdot j}}{N})^2}{\frac{n_{i\cdot} n_{\cdot j}}{N}}$ ，因此 $77 \sim 77.1$  相当于 $77,2 = 772,,$

(6. 7. 16)

## 6.7 拟合优度检验

293

其否定域为 $77+ = \sum_{i,j} (n_{ij} - \frac{n_{i\cdot} n_{\cdot j}}{N})^2 / \frac{n_{i\cdot} n_{\cdot j}}{N}$ .上式通常称为统计量.

证明为了求解以上检验问题，可应用Fisher定理.为此，定义

$77(\text{权})$ 如下：取 $7TU = \sum_{i,j} \frac{(n_{ij} - \frac{n_{i\cdot} n_{\cdot j}}{N})^2}{\frac{n_{i\cdot} n_{\cdot j}}{N}}$ ，则 $\chi^2 = \sum_{i,j} \frac{(n_{ij} - \frac{n_{i\cdot} n_{\cdot j}}{N})^2}{\frac{n_{i\cdot} n_{\cdot j}}{N}}$ 在好。成立时 $\chi^2_{ij} = \sum_{i,j} \frac{(n_{ij} - \frac{n_{i\cdot} n_{\cdot j}}{N})^2}{\frac{n_{i\cdot} n_{\cdot j}}{N}}$  (沒)

为： $7T'' = \sum_{i,j} \frac{(n_{ij} - \frac{n_{i\cdot} n_{\cdot j}}{N})^2}{\frac{n_{i\cdot} n_{\cdot j}}{N}}$ ， $772] = 0.29$   $7722 = 1 - R - 2$ 氏，即  $k = 4$ ， $p = 2$ ，

由 Fisher定理有

$A = \sum_{i,j} \frac{(n_{ij} - \frac{n_{i\cdot} n_{\cdot j}}{N})^2}{\frac{n_{i\cdot} n_{\cdot j}}{N}} \sim \chi^2_{k-1}$ ，填H  $\sim 2$ (灸

$\chi^2_{k-1}$   $J=1$

(6. 7. 19) 对于以上多项分布，可求出 $H_0$ 成立时(即 $7712 = 7721$ 时)参数 $\theta$ 的极大似然估计为(见第三章例题) A

因而有

$\theta$

X

$= 7Tn = -\sum_{i,j} \frac{n_{ij} \ln n_{ij}}{n_{i\cdot}}$   $Tt = 2^{TL}$

$yn = 77t$

$y(\chi^2)$

$ti) = \sum_{i,j} \frac{(n_{ij} - \frac{n_{i\cdot} n_{\cdot j}}{N})^2}{\frac{n_{i\cdot} n_{\cdot j}}{N}}$

$jVjJ - A$

$N, 2 + N2l$

$AA/V|| AAN+TV'' ^1(^)=01=-, \sim \textcircled{2}, -^--$

开 $21(^)=02, 7T22(3)=1-A-202=-. n$

这些结果代入上述(6. 7. 19)式, 经简单计算即可得到

$Kn(0) \pm Z=(4-\sim)2 -/(I). (?v12+/v21)$

例6. 7. 6 (续) 对300名志愿者用两种方法进行化验, 其观测结果 为: $/V''=211, /$   
 $V12=19, N2l=7, N22=63$ , 这些数值代人 统计量 得 $7=5.23$ , 而 $/ (1, 0.95) =3.84$ , 因  
此否定 $7/.$ , 即说明两种检验方

法不等价. ■ 例6.7.7 独立性检验.要检验某种感冒药, 其疗效是否与年龄有  
显著尽

较差 23 18 14

总和 109 100 91 300

$W(1).$

(6. 7. 18)

关, 对300个人进行调查, 其疗效与年龄的关系如下表所示: 少年4, 成年a2 老年么

总和 58 38 32 128 一般 b 2 28 44 45 117

55

294

第六章参数假设检验

本问题为 $3 \times 3 = 9$ 格列联表, 假设年龄 $X$ 分为少年、成年、老年 (即4, , 42, 43), 疗效 $r$ 分为  
显著、一般、较差 (即 $U_2, B3$ ), 而 $(Xe A^Y^B.)$ 的人数为 $/V, y$ , 概率为 $77y (Z=1, 2, 3;$   
 $j=1, 2, 3)$ , 因此 $/V=$

$(^.)$ 服从9项分布 $/V \sim M/V(300, 7t)$ , 若 $X,$  独立, 即年龄与疗效无

关, 则应有

$770=P(XeAMyGBy) =P(XeAf.)P(yeBy), i = 1, 2, 3; ; = 1, 2, 3.$

这种独立性检验可推广到一般情形.我们考虑本节开头讨论的2 维  $r \times s$ 列联表, 这时 $X$ 有个状  
态,  $F$ 有 $B、, ". , Bs$ 个状态, 要进 行独立性检验:

$=P(X^Ai)P(Y^Bj) 9 Z = 1, \dots, r; y = 1, -, s. (6. 7. 20)$

3S

由于  $P(XeAf) =P( U (^e4, , reB.))= \text{工} 77, =^.,$  同理  $P(Ye ./=* j=i$

T

$Bj) \text{工} \sim .$ 因此, 以上独立性检验可化为

$1=1 , ^0 =7T, . tTTij^TTi. n , j9 i = 1$

应用Fisher定理可得

$, r; J = 1, ". , S.$

(6. 7. 21 )

定理6. 7. 4 独立性检验(6. 7. 20)或(6. 7. 21)的检验统计量为

$1r N2 \setminus$

$K^=n[Z^2 (1), f=(r'1)(5_1)*$

(6. 7. 22)

其否定域为 $/T = \rangle /((, -1)0-1), 1 -\langle \rangle \}$ . 证明为了求解这一检验问题, 可设法应用  
Fisher定理.为此, 根

据以上关系可定义一个函数 $7T(0)$ , 取 $0=(77, . , 7T2., \dots, \text{町}(r_u.; 77., ,$

$77 \dots .(.-1))T \gg$  其维数 $p=r+s-2$ , 这时有

$7T; . =0"i=1, \dots, r-1, 7Tr. =1- \text{-----} 0r_-, ,$

TT.j=Gr\_、...\_/\*=1, ..., S-1, 7T.c =1 -7T.)-----一开.(:-"• 若^成立 , 则有

$y^{(j)} = 77, .77.;$ ,  $I = 1, \dots, f$ ;  $J = 1, \dots, S$ , 而的计算可直接从多项分布(6. 7. 10)式得到:

.N

$77 \cdot = \text{丁}$ ,  $77-j = \sim 7 \sim$ ,  $f = 1, \dots, r$ ;  $\blacksquare / = 1, \dots, \langle 5$ ,

SF

其中 $v, . = 2$  义,  $N., 2 \sim$ , 因此当坑成立时有 $= 7 = 1 i =: 1$

习题六

295

NN

-^A这些结果代人公式(6.7. 13)式可得 nn

定理6. 7. 5 尺(幻, 其中

在一定正则条件下, 久 $\rightarrow 0(a.e.)$ , 且有

$x \rightarrow 0$ ,

$x < 0$ .

$\sim ) = z^2 - 1 = 1$  ;=1

$\vdash AM \text{ snn}$

义2(1),

$n \text{ ----- } nnJ$

其中自由度 $t=k-p-1$ ,  $k=rs$ ,  $p=r+s-2$ , 可得 $Z=rs +2-1$

$(r-l)(5 \sim l)$ . 上式经化简即可得到(6.7.22)式. ■ 例6. 7.7 (续)把公式(6. 7. 22)

式用于本例可得 $Kn(0) = 13.59$ ,

而  $f = (3 - 1)(3 - 1) = 4$ ,  $x^2(4, 0.95) = 9.49 < Kn = 13.59$ , 所以应否定 尽, 即

“疗效与年龄独立”这一假设不成立, 认为该感冒药的疗效与 年龄有关. I

最后我们简单介绍一下Kolmogonov拟合优度检验统计量, 通常认 为它是最精密的检验统计量, 有关结果的证明比较复杂, 以下仅介绍其 基本公式, 详见陈希孺( 1981, 1999).

设 $A, \dots$ , 及为独立同分布样本, 要检验它们是否服从某个已知的 分布函数 $F_0(x)$ :

"0: 尤 $I \sim F_0(\text{文}) \rightarrow * "1$  : 义1不服从 $f'_0(x)$ .

检验的出发点是经验分布函数 $F_n(x)$ , 其中 $F_n(x) = n^{-1} \sum_{i=1}^n I_{\{X_i \leq x\}}$  中 $I_{\{X_i \leq x\}}$ 的个数 $I$ . 由第一章的定理1. 1.2可知, 当 $n \rightarrow + \infty$ 时,  $F_n(x) \rightarrow F_0(x)(a.e.)$ . 因此Kolmogonov取以下统计量来度量样本分布与 $F_0(x)$

拟合的优劣程度:

$D_n = \sup_{0 \leq x \leq 1} |F_n(x) - F_0(x)|$

反映了用 $F_0(x)$ 拟合 $F_n(x)$ 的总偏差.

, ..., ^

以上假设检验问题的否定域为 $R_+ = \{D_n > A_{1-\alpha}\}$ , 其中 $A_{1-\alpha}$ 为分布 尺(幻的 $1 - \alpha$ 分位数(函数 $K(x)$ 有表可查).

习题六 1. 在以下假设检验问题中, 求检验 $\alpha$  (幻的第一、, 第二类错误以及

296

第六章参数假设检验

在义情形下的功效:

(1) 设 $X_1, X_2, X_3$ 为 $i. i. d.$ 样本,  $X_i$  服从两点分布 $b(1, \theta)$ . 考虑

假设检验问题= 检验 $H_0: \theta = 0.5$ 的否定域为 $R_+ =$

$U, +A +X_3^2 \mid$ ;

(2) 设 $X_1, X_2$ 相互独立, 都服从Poisson分布 $P(\lambda)$ . 考虑假设检验



问题  $H_0: A = y \sim //1: A = 2$ . 检验  $</>(\%)$  的否定域为  $尺+ = 1\%$ ,  $+X^2$  多 3 !.

2. 考虑简单假设检验问题  $H_0: \theta = 30^\circ$   $H_1: \theta \neq 30^\circ$ . 对于以下分布密度在一样本情况下, 求检验问题水平为  $\alpha$  ( $0 < \alpha < 1$ ) 的 MPT:

(1)  $f(x; \theta) = 2\theta^{-1} 2(\theta - x) / \theta^2$ ,  $0 \leq x < \theta$

(2)  $f(x; \theta) = 2[快x + (1 - \theta)(1 - x)] / <没0 \leq 1$ . 3. 考虑简单假设检验问题

$H_0: f(x) = f_0(x) - H_1: f(x) = f_1(x)$ . 对于以下分布密度人(幻和  $Z(x)$ ), 在一样本情况下, 求检验问题水平为  $\alpha$

( $0 < \alpha < 1$ ) 的 MPT:

(1)  $f_0(x) = I_{[0, 1]}$ ,  $f_1(x) = 2x I_{[0, 1]}$ ; (2)  $f_0(x) = I_{[0, 1]}$ ,  $f_1(x) = 4x I_{[0, 1]}$

$I_{[0, 1/2]}$

4. 设  $X_1, \dots, X_n$  为  $i. i. d.$  样本, 求以下假设检验问题水平为  $\alpha$  ( $0 < \alpha < 1$ ) 的 MPT:  $H_0: \theta \geq 1 - H_1: \theta < 1$ . (1)  $X$ , 服从均匀分布  $U(0, 1)$ ; (2)  $X$ , 服从冷分布  $BE(\theta, 1)$ .

5. 设  $X_1, \dots, X_n$  为  $i. i. d.$  样本, 考虑以下假设检验问题  $H_0: \theta \geq 1 - H_1: \theta < 1$ . 求检验问题水平为  $\alpha$  ( $0 < \alpha < 1$ ) 的 MPT

氏  $- "H_1: \theta = 0$

(1) 如  $"\theta) = e_{-}(v) / \wedge x^8$ ,  $\theta' < e0$ ; (2)  $f(x; \theta) = 2e^{2x-3} I_{[0, \infty)}$ ,  $e \in \mathbb{R}$ .

6. 设  $X_1, \dots, X_n$  为  $i. i. d.$  样本,  $\theta \sim r(A, p)$ , 其中  $p > 0$  已知,  $A > 0$  未知.

(1) 求假设检验问题  $H_0: A = A_0$   $H_1: A \neq A_0$  ( $A_0 > 0$ ) 的一个水平为  $\alpha$  的 MPT;

(2) 求 (1) 中的检验函数冷 (又) 的功效函数氏 ( $A$ ), 并证明其为  $A$  的增函数;

(3) 证明 (1) 中所得的 MPT 是以下假设检验问题的 UMPT: (a)  $H_0: A = A_0$   $H_1: A > A_0$ ;

(b)  $H_0: A \leq A_0$   $H_1: A > A_0$ . )  $j: f(x) = f_j(x)$ .

7. 考虑简单假设检验问题  $H_0: f(x) = f_0(x)$

习题六

297

(1) 证明: 以下  $4 > x$  为上述检验问题水平为  $\alpha = 0$  的 MPT, 并且

在  $R^+$  上唯一,

$f_l, e/?+ = \{x: f_Q(x) = 0, /i(\%) \neq 0\}$ ,

$</>i(\%) = j \gg x \&R^+ = \{x: f_Q(x) = 0, /i(\%) = 0\}$ ,  $0 < y < 1$ ,  $l_0, xg/?$

$_ = \{x: f_0(x) \neq 0\}$ ;

(2) 证明: 以下小 2 (幻为上述检验问题水平为  $\alpha = 1$  的 MPT, 并且 在  $尺+ U / T$  上唯一,

1,

$x \in R^+$  其中  $A = \{x: f_0(x) \neq 0\}$ ,  $B = \{x: f_1(x) \neq 0\}$ ,  $0 < y < 1$ .

8. 设  $\phi(x)$  为  $N-P$  基本引理导出的 MPT (见定理 6. 2. 1 的 (6.2.2) 式), 其水平为  $\alpha$  的否定域为  $R^+ = \{x: \phi(x) > c\}$ ; 且满足  $P, \phi(x) > c\} = \alpha$ , 其中  $c$  可看作  $\alpha$  的函数  $c = c(\alpha)$ .

(1) 证明:  $c(\alpha)$  为  $\alpha$  的非增函数, 即若  $\alpha_1 < \alpha_2$ , 则  $c(\alpha_1) \geq c(\alpha_2)$ ;

(2) 若  $\phi(x)$  的分布函数连续, 且  $\alpha_1 < \alpha_2$ , 证明: 当水平  $\alpha = \alpha_1$  时,  $\phi(x)$  犯第二类错误的概率不小于水平  $\alpha = \alpha_2$  时  $\phi(x)$  犯第二类错误的概率.

9. 对于简单假设检验问题  $H_0: \theta = \theta_0 - H_1: \theta \neq \theta_0$  以及水平  $\alpha$  ( $0 < \alpha < 1$ ). 设  $\phi(x)$  为上述检验问题水平为  $\alpha$  的 MPT, 且  $E[\phi(X)] < 1$ . 证明: 检验  $1 - \phi(x)$  是假设检验问题:  $\theta = \theta_0$  的一个水平为  $1 - \alpha$  的 MPT.

10. 证明以下分布族为单调似然比分布族:

(1)  $A \sim M(\lambda)$ :  $\lambda \in R^+$  未知,  $\lambda > 0$  已知;

(2)  $U(0, 0 + 1) : 6 > eR$ ;  
 (3)  $U(x) : \wedge gR I$ , 其中  $f_e(x) = c(0)h^{\wedge}I$  且  $A(\wedge)$  和  $c(\wedge)$  都是正函数,  $a(\wedge)$  和  $6(\wedge)$  都是  $6$  的非降函数.

$(/ > 2(x) = y$   
 $\%e/?+ -A\JB, -$

11. 对习题5 中的两个小题的分布族:

(1) 证明其分布族为单调似然比分布族;  
 (2) 对于简单假设检验问题  $H_0 : \theta = \theta_0$  个水平为  $\alpha$  的非随机化的MPT;  
 :沒=久(久 <沒。), 求一

,

,

(3) 对于假设检验问题  $U \leq \theta_d : \theta > \theta_0$ , 求一个水平为  $\alpha$  的 UMPT.

## 298 第六章参数假设检验

12. 设  $X_1, \dots, X_n$  为 i. i. d. 样本,  $A \sim (x_1; \wedge)$ ,  $0 < \wedge < \infty$ . 在下列情况下, 基于  $T = \sum_{i=1}^n X_i$  分布 (见第一章习题), 求假设检验问题  $H_0 : \theta = \theta_0$  对  $H_1 : \theta > \theta_0$  水平为  $\alpha$  的 UMPT:

(1)  $\theta = \theta_0$  对  $\theta > \theta_0$ ; (2)  $f(X_i; \theta) = (9\theta)^{-1} (1 - \theta x_i)^{9-1}$ ,  $0 < \theta < 1$ ;  
 (3)

$\theta = \theta_0$  对  $\theta > \theta_0$ , 其中  $\theta_0 > 0$  未知,  
 $c > 0$  已知.

13. 求以下单边假设检验问题水平为  $\alpha$  的 UMPT:

(1)  $X$  服从超几何分布  $A(x; n, V, M)$ , 求检验  $H_0 : M \leq M_0$  对  $H_1 : M > M_0$  的 UMPT;

(2)  $X$  服从负二项分布  $NB(r, \theta)$ , 求检验  $H_0 : \theta \leq \theta_0$  对  $H_1 : \theta > \theta_0$  的 UMPT.

14. 设  $X_1, \dots, X_n$  为 i. i. d. 样本,  $X_i$  服从两点分布  $b(1, p)$ .

(1) 求检验  $H_0 : p \leq 0.02$  对  $H_1 : p > 0.02$  水平为  $\alpha = 0.05$  的 UMPT;

若要求这个检验在  $p = 0.08$  时犯第二类错误的概率不超过  $0.1$ ,  $n$  要取多大?

(2) 某厂生产的轴套出厂标准为次品率不超过  $2\%$ , 现从一批产品中随机抽取  $400$  只轴套, 发现次品  $12$  只. 在水平为  $\alpha = 0.05$  的条件下, 是否应允许这批产品出厂?

15. 设  $X_1, \dots, X_n$  为 i. i. d. 样本,  $X_i$  表示寿命且服从指数分布

$r(y, i)$ . 现仅观测到前  $r$  个寿命的值:  $(x_1, \dots, x_r)$ ,  $r \leq n$ . 求  $X_{r+1}, \dots, X_n$  的联合分布.

(1) 对于假设检验问题  $H_0 : \theta \leq \theta_0$  对  $H_1 : \theta > \theta_0$ , 求其 UMPT 的功效函数;

(2) 设某元件的寿命服从指数分布  $r(y, \lambda)$ . 今对  $12$  个元件观测

其寿命, 仅测得前  $6$  个元件的失效时间 (小时) 为  $1380, 1510, 1650, 1760, 2100, 2320$ . 在水平为  $\alpha = 0.05$  的条件下检验: 元件的平均寿命是否在  $2000$  小时以上?

\*16.  $X_1, \dots, X_n$  为 i. i. d. 样本,  $n \geq 2$ ,  $X_i \sim \text{Exp}(\theta)$ . (1) 求  $X(1)$  和  $X(n)$  的联合分布;

(2) 证明:  $\wedge^*(X(1), X(n))$

$X(1) < 1 - \alpha$  且  $X(n) < 1$ , 去或

## 习题六

299

是假设检验问题:  $\theta > \theta_0$  的水平为  $\alpha$  的 UMPT (提示: 仿照例 6.2.3).

\*17. 假设  $J \in CR$  | 关于  $T = T(X)$  为单调似然比分布族. 考虑假设检验问题  $H_0 : \theta \leq \theta_0$  对

$H_1 : \theta > \theta_0$ , 其水平为  $\alpha$  的 umpt 的否定域为  $\{r(X) > c(\theta_0)\}$ , 且满足  $P_{\theta_0}(r(X) > c(\theta_0)) = \alpha$ .

$>c(0_0)\} = a$ ; 该式对任意的 $\delta$ 都成立. 若把 $C(\delta)$ 看成 $\theta$ 的函数, 证明: $c(\delta)$ 是 $\theta$ 的增函数 (提示: 利用UMPT的无偏性).

18. 设 $X_1, \dots, X_n$ 为i. i. d. 样本,  $X_i \sim N(\theta, 1)$ ,  $\theta \in \mathbb{R}$ . 证明: 对于假设检验问题

或

(1)  $H_0: \theta = 0$

(2)  $H_0: \theta = 0$  vs  $H_1: \theta > 0$  求其水平为 $\alpha$ 的UMPUT.

19. 设 $X_1, \dots, X_n$ 为i. i. d. 样本,  $X_i \sim N(\theta, 1)$ . 对于假设检验问题

$H_0: \theta = 0$

vs  $H_1: \theta > 0$  求其水平为 $\alpha$ 的UMPUT.

;

$\theta = 0$  vs  $H_1: \theta > 0$ . 证明: 其水平为 $\alpha$ 的

UMPUT,  $\theta = 0$  vs  $H_1: \theta > 0$ , 其他

$\theta = 0$

求其水平为 $\alpha$ 的umput

并给出 $\delta$ 的定解条件.

20. 设 $X_1, \dots, X_n$ 为i. i. d. 样本,  $X_i \sim N(\theta, 1)$ . 证明以下假设检

验问题的UMPT不存在: (1)  $H_0: \theta = 0$  vs  $H_1: \theta > 0$

(2)  $H_0: \theta = 0$  vs  $H_1: \theta > 0$  (提示: 利用单边检验中功效

函数的单调性).

$\theta = 0$ 成立.

21. 设 $X_1, \dots, X_n$ 为i. i. d. 样本,

求其水平为 $\alpha$ 的UMPUT.

22. 对于指数族分布, 设 $X_1, \dots, X_n$ 为i. i. d. 样本,  $X_i \sim \exp(\theta)$ ,  $\theta > 0$ . 证明: 以下假设检

验问题的UMPT不存在: (1)  $H_0: \theta = 1$  vs  $H_1: \theta > 1$  (提示: 利用单边检验中功效函数的单调性).

(2)  $H_0: \theta = 1$  vs  $H_1: \theta > 1$  求其水平为 $\alpha$ 的umput

并给出 $\delta$ 的定解条件.

23. 设 $f(x)$ 和 $g(x)$ 为两个已知的密度函数, 又设 $A$ 的密度函数

为 $f(x)$ 和 $g(x)$ 的凸组合, 即 $f(x) + (1 - \alpha)g(x)$ ,  $\alpha \in [0, 1]$ . 证明:

对于假设检验问题 $H_0: \theta = 0$  vs  $H_1: \theta > 0$ ,  $\theta > 0$ . 证明: 以下假设检验问题的UMPT不存在: (1)  $H_0: \theta = 0$  vs  $H_1: \theta > 0$

(2)  $H_0: \theta = 0$  vs  $H_1: \theta > 0$

$\theta = 0$

第六章参数假设检验

1. 设 $X_1, \dots, X_n$ 为i. i. d. 样本,  $X_i \sim N(\theta, 1)$ ,  $\theta \in \mathbb{R}$ . 证明: 对于假设检验问题

$H_0: \theta = 0$  vs  $H_1: \theta > 0$  求其水平为 $\alpha$ 的UMPUT.

(提示: 对 $\theta$ 分情况讨论, 把 $\theta$ 分为 $\theta < 0$ ,  $\theta = 0$ 以及 $\theta > 0$ 三种情况证明多 E以).

2. 设 $X_1, \dots, X_n$ 为i. i. d. 样本, 求假设检验问题

$H_0: \theta = 0$  vs  $H_1: \theta > 0$  求其水平为 $\alpha$ 的umput:

(1)  $\theta = 0$  vs  $H_1: \theta > 0$  已知;

(2)  $\theta = 0$  vs  $H_1: \theta > 0$ ,  $\theta > 0$ ,  $c > 0$ , 即 $f(X_j; c) = c \exp(-cx) \sim \{c + i\} I(X_j \leq c)$ ,  $c > 0$ 已知 (提示: 参见第24题中UMPT的形式).

26. 某饮料厂生产的瓶装果汁重量服从正态分布, 要求出厂重量的标准差(即 $\sigma$ )在0.02公斤之内. 今抽取28个样本, 计算出样本方差为

0.001公斤, 问这批瓶装果汁在水平 $\alpha = 0.05$ 下是否合格?

27. 某品牌香烟尼古丁含量服从正态分布, 合格品要求尼古丁平均含量不超过17.5毫克. 今对一批产品随机抽取8支样品, 算出其尼古丁含量的平均值为18.6毫克, 标准差为1.4毫克. 在水平 $\alpha=0.01$ 下, 检验这批香烟是否合格.

28. 为了比较甲、乙两个砖厂生产砖块的平均强度, 在两厂抽取样品, 测得强度分别为74, 65, 72, 69和75, 78, 74, 76, 72. 假设两厂砖块强度分别服从正态分布  $N(\mu_1, \sigma_1^2)$  和  $N(\mu_2, \sigma_2^2)$ . 试检验: 两厂生产砖块的平均强度是否有显著性差异 (取  $\alpha=0.05$ ).

29. 为了比较温度对于针织品断裂强度的影响, 分别在70 t 和 t 下对某针织品测试其断裂强度, 具体数据如下

80

70 X. 时: 20.5, 18.8, 19.8, 20.9, 21.5, 19.5, 21.0, 21.2; 80 X. 时: 17.7, 20.3, 20.0, 18.8, 19.0, 20.1, 20.2, 19.1.

在正态性假设下: (1) 通过假设检验说明: 可认为两组数据相应分布的方差相等; (2) 在此基础上检验: 温度对于针织品断裂强度是否有显著性差异 (取  $\alpha=0.05$ ).

30. 为了比较A, B两种同类药物副作用的大小, 对服用A种药的92人进行调查, 其中7人有不良反映; 对服用B种药的78人进行调查, 其中5人有不良反映. 试检验: A, B两种药物副作用的大小是否有显著性差异 (取  $\alpha=0.01$ ).

31. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ,  $\mu, \sigma^2$  都未知. 求假设检验问题:  $H_0: \mu \leq \mu_0$  的水平为  $\alpha$  的 UMPUT 的功效

:

习题六

301

函数.

32. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ;  $Y_1, \dots, Y_m$  为

i.i.d. 样本,  $Y_i \sim N(\mu, \sigma^2)$ , 且两总体独立. 求下列假设检验问题水平为  $\alpha$  的 UMPUT:

(i)  $H_0: \mu_1 = \mu_2$ ; (ii)  $H_0: \mu_1 \geq \mu_2$ ;

At] 和 (iii)  $H_0: \mu_1 > \mu_2$ ,  $c$  为已知常数.

33. 设  $X_1, \dots, X_L$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ,  $\mu, \sigma^2$  未知. 若  $\mu$  已知但  $\sigma^2$  未知, 求下列假设检验问题水平为  $\alpha$  的 UMPUT:

(i)  $H_0: \sigma^2 \leq \sigma_0^2$  vs  $H_1: \sigma^2 > \sigma_0^2$  和 (ii)  $H_0: \sigma^2 = \sigma_0^2$  vs  $H_1: \sigma^2 \neq \sigma_0^2$ .

34. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ;  $Y_1, \dots, Y_m$  为

$N(\mu, \sigma^2)$  和 (ii)  $H_0: \mu = \mu_0$  vs  $H_1: \mu \neq \mu_0$ ;

35. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ;  $Y_1, \dots, Y_m$  为

$N(\mu, \sigma^2)$  和 (ii)  $H_0: \mu = \mu_0$  vs  $H_1: \mu \neq \mu_0$ ;

36. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ;  $Y_1, \dots, Y_m$  为  $N(\mu, \sigma^2)$  和 (ii)  $H_0: \mu = \mu_0$  vs  $H_1: \mu \neq \mu_0$ ;

37. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ;  $Y_1, \dots, Y_m$  为  $N(\mu, \sigma^2)$  和 (ii)  $H_0: \mu = \mu_0$  vs  $H_1: \mu \neq \mu_0$ ;

38.

39. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ;  $Y_1, \dots, Y_m$  为  $N(\mu, \sigma^2)$  和 (ii)  $H_0: \mu = \mu_0$  vs  $H_1: \mu \neq \mu_0$ ;

40. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ;  $Y_1, \dots, Y_m$  为  $N(\mu, \sigma^2)$  和 (ii)  $H_0: \mu = \mu_0$  vs  $H_1: \mu \neq \mu_0$ ;

41. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ;  $Y_1, \dots, Y_m$  为  $N(\mu, \sigma^2)$  和 (ii)  $H_0: \mu = \mu_0$  vs  $H_1: \mu \neq \mu_0$ ;

42. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ;  $Y_1, \dots, Y_m$  为  $N(\mu, \sigma^2)$  和 (ii)  $H_0: \mu = \mu_0$  vs  $H_1: \mu \neq \mu_0$ ;

43. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ;  $Y_1, \dots, Y_m$  为  $N(\mu, \sigma^2)$  和 (ii)  $H_0: \mu = \mu_0$  vs  $H_1: \mu \neq \mu_0$ ;

44. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ;  $Y_1, \dots, Y_m$  为  $N(\mu, \sigma^2)$  和 (ii)  $H_0: \mu = \mu_0$  vs  $H_1: \mu \neq \mu_0$ ;

i. i. d. 样本

(1) 求下列假设检验问题水平为  $\alpha$  的 UMPUT:  $6 < 9$ 。《打1:

检验的处理方法)。(36. 设  $X^1, \dots, X_n$  为 i. i. d. 样本,  $X_j \sim r$   
 $^{ye-^{'ZU.^0}}$  | , 其中  $cr, />$  都未知。

» 即  $/(\alpha, p) =$

(1) 对于假设检验问题  $H_0: a \leq a_0$ , 证明: 存在一个

否定域为  $\{x: > g(FT\%.)\}$  的 UMPUT, 其中  $g$  为某一函数;  $\gg, f=i$

(2) 对于假设检验问题 (i) :  $p \leq p_0$  和 (ii) 好。:  $/> =$

$\tau)2,$

302

## 第六章参数假设检验

$P_0 - H_1: P \leq p_0$ , 试证: 存在否定域与  $n$  有关的 umput-  $i=l$

37. 设  $A \sim P(A_1)$ ,  $X_2 \sim P(A_2)$ ,  $X_1$  和  $X_2$  独立. 求假设检验问题  $//0 : A_1 \leq A_2$   
 $^H]$  :  $A_1 < A_2$  的一个水平为  $\alpha$  的 UMPUT.

38. 设  $X_1, \dots, X_n$  为 i. i. d. 样本,  $X_1, X_2, X_3$  相互独立. 对于假设检验问题证明: 存在一个  
水平为  $\alpha$  的 UMPUT.

\*39. 设  $X_1, \dots, X_n$  为 i. i. d. 样本,  $X_1, \dots, X_n$  且两总体独立.

i. i. d. 样本,  $r$ , 且两总体独立.

(1) 设  $P_1, p_2$  已知, 对于假设检验问题 (i) 片。

$<72$  和 (ii)  $H_0: a \leq a_0$ ,  $= < r_2 \sim / Z_1: a$

可用分布确定;

(2) 如果在 (1) 中,  $P_1, p_2$  都未知, 对假设检验问题 (i) 和 (ii), 证

明: 存在 UMPUT, 并描述它们的一般形式;

(3) 假设  $a_0 = a_0$  但未知, 对假设检验问题 (i)  $//0; P_1 \leq p_2 - //$ , :

$P_1 > P_1$  和 (ii)  $H_1: P_1 \leq P_2$ , 证明: 存在 UMPUT, 并描述它们的一般形式。

40. 设  $X_1, \dots, X_n$  为 i. i. d. 样本,  $X_1, \dots, X_n$

验问题水平为  $\alpha$  的似然比检验:

(1) 当  $i=0$  时,  $//0 : (T = a_0 - ^{H_1} : a \neq a_0$ ; (2) 当  $z$  未知时,  $H_0: a = (T_0 * -$   
 $*//, : a \neq t_0$ ; (3) 当  $<7$  已知时,  $H_0: f_j L 0$ ;

求以下假设检

1

$A > \#0-2$ , 试证: 存在 UMPUT, 且其否定域

$f_i = ^$

41. 设  $X_1, \dots, X_n$  为 i. i. d. 样本,  $X_1, \dots, X_n$  服从均匀分布  $R(Q, 0)$ ,  $e > 0$ .

求假设检验问题  $H_0: 0 = \sim - H_1: 0 \neq 0$  水平为  $\alpha$  的似然比检验。

42. 设  $X_1, \dots, X_n$  为 i. i. d. 样本,  $X_1, \dots, X_n \sim PR, a, e$ , 即  $= a_6 ax; \{0l+\}$  ) I

$\backslash xx^e !$ , 其中  $a, <9$  都未知. 求假设检验问题 77。:  $a = 1-$

的一个水平为  $\alpha$  的似然比检验。

43. 设  $X$  是来自密度函数为  $f(x) = 2e^{-x}I_{[0, \infty)}$  的总体

的一个容量为 1 的样本, 其中  $0 < \theta$  未知. 求假设检验问题  $f_0: \theta = \theta_0 : <9 \#$  氏的一个水平为  $\alpha$  的似然比检验。

44. 设  $X_1, \dots, X_n$  为相互独立的样本,  $X_j \sim N(0, a^2)$ ,  $y \neq Z$ , 求以下假设检验问题水  
平为  $\alpha$  的似然比检验统计量:

(4) 当  $\theta$  未知时,  $H_0$ :

.

$K, \dots, )$ ; 为

#### 习题六

303

(1)  $X \sim N(y, a^2)$ ,  $H_0: y = 0$ ;  $H_1: y \neq 0$ ;

(2)  $X \sim \chi^2(k)$ ,  $H_0: k = 1$ ;  $H_1: k \neq 1$  (提示: 利用 F 分布与  $\chi^2$  分布的关系, 由  $\chi^2$  分布确定否定域, 第45题的(2)、第48题和第49题的(3)也类似).

45. 设  $U_1, \dots, U_n$  为相互独立的样本,  $U_i \sim \chi^2(k, a^2)$ ,  $a > 0$ ,  $n$ , 而  $X_i = U_i / V(y, a^2)$ ;  $X_n \sim \chi^2(k, a^2)$ . 求以下假设检验问题水平为  $\alpha$  的似然比检验和 score 检验统计量:

(1)  $H_0: y = 0$ ;  $H_1: y \neq 0$ ;

$H_0: a = 1$ ;  $H_1: a \neq 1$ ;

$H_0: y = 0, a = 1$ ;  $H_1: y \neq 0, \text{或 } a \neq 1$ .

设  $U_1, \dots, U_n$  为相互独立的样本, 并考虑假设检验问题  $H_0: y = 1$ ;  $H_1: y \neq 1$ . 在以下情形下求假设检验问题的 score 检验统计量:

(1)  $X_i = U_i / V(y, 1)$ , 但  $X_i \sim \chi^2(k, y)$ ;

(2)  $X_i = P(A_i)$ ,  $j = 1, \dots, n$ , 但  $X_i \sim P(y)$ .

47. 证明: 在定理 6.2 中的 score 检验统计量  $S_C(\hat{\theta})$  可表示为  $L_p(\hat{\theta}) / L_p(\theta)$ , 其中  $L_p(h) = L(\hat{\theta}^2(h))$ ,  $L_p(\theta)$  和  $L_p(\hat{\theta})$  为  $L_p(\theta)$  关于  $\theta$  的前二阶导数.

48. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim T(k, 1)$ ; 设  $Y_1, \dots, Y_m$  为 i.i.d. 样本,  $Y_i \sim T(k, 1)$ , 且两样本独立. 求以下假设检验问题水平为  $\alpha$  的似然比检验:  $H_0: k = 1$ ;  $H_1: k \neq 1$ .

49. 设  $X_1, \dots, X_n$  为 i.i.d. 样本, 及  $X_i \sim \chi^2(k, M)$ ;  $M > 0$ ,  $k > 0$ , 求

(2)

(3) 46.

"d. 样本,  $Y_i$  水平为  $\alpha$  的似然比检验:

50. 设随机变量  $U_1, \dots, U_n$  满足关系  $U_i = f(U_{i-1})$ ,  $f = 1, \dots, n$ ,  $X_0 = 0$ . 其中 " $U_1, U_2, \dots, U_n$ " 为 i.i.d. 随机变量,  $U_i \sim \chi^2(k, a^2)$ . 求以下假设检验问题水平为  $\alpha$  的似然比检验:  $H_0: \theta = 0$ ;  $H_1: \theta \neq 0$ .

51. 对某中学学生每周看电视的小时数进行调查, 情况如下表所示:

(1)  $H_0: M_1 = M_2$  (2)  $H_0: \sigma_1^2 = \sigma_2^2$

, 且两样本独立. 求以下假设检验问题水平为  $\alpha$  的似然比检验:  $H_0: \mu_1 = \mu_2$ ;  $H_1: \mu_1 \neq \mu_2$ ;

其中  $e_r = a^2 = tr$  未知 (3)  $H_0: a = 1$ ;  $H_1: a \neq 1$ , 其中  $\theta = 1$ ,  $\theta^2$  未知.

;

304

#### 第六章参数假设检验

每周观看时数 去年比例 今年人数

<3 0.18

[3, 6) 0.30

[6, 9) 0.26

[9, 13) 0.20 0.06

患者人数 健康者人数 总人数

吸烟量(支/日) 总人数 [0, 9] [10, 19] >20

22 98 25 145 22 89 16 127 44 187 41 272

10 30 20 30 10

试检验: 从今年100人的调查结果来看, 今年的比例与去年的比例是否相同 ( $\alpha = 0.05$ )?

52. 某工厂近5年共发生63次事故, 若按星期几分类, 事故发生

在星期一至星期六的次数如下表所示. 判别一下, 事故的发生与星期几 是否有关( $\alpha=0.05$ )?

星期- 四五

事故次数 9 10 11 8 13 12

53-为了研究慢性气管炎与吸烟量的关系, 共调查了 272人, 结果 如下表所示. 试检验: 慢性气管炎与日吸烟量是否有关( $\alpha:=0.05$ ).

54. 某实验室在2 608个相等时间单位内观察了一种放射性物质所 释放出来的 $\alpha$ -粒子的个数, 结果如下表所示, 其中频数 $\nu$ 表示在 2 608个时间单位中释放出个粒子的时间单位的个数. 试验证: 该放

射性物质所释放出来的 $\alpha$ -粒子的个数服从Poisson分布( $\alpha = 0.05$ ).

粒子数 $x$  0 1 2 3 4 5 6 7 8 9 10 11 12 13 总数 频数 \ 57 203 383 525  
532 408 273 139 45 27 10 4 2 0 2 608

## 第七章区间估计

参数的区间估计与点估计一样, 是参数估计的重要方法, 在某些具 体问题中可能比点估计更有实用价值. 给定参数分布族 $X \sim$

$\theta$ , 参数的点估计是通过一个统计量以及某些优良性准则(诸如均 方误差最小准则等)给出未知参数的估计. 例如要估计, 点估 计得到的是样本的某一个函数 $g(\hat{x})$ (根据某种估计方法), 在实用上就 把 $g(\hat{x})$ 当作 $g(\theta)$ 的近似值. 但是, 尽管有均方误差最小等许多判别点 估计精度的优良性准则, 我们还是无法知道估计值与真值究竟相差多 少, 区间估计在一定程度上解决了这个问题. 区间估计是通过两个统计 量以及覆盖概率(也会有适当的优良性准则)给出未知参数的估计, 这 时得到的是样本的两个函数 $g_1(\hat{x})$ 和 $g_2(\hat{x})$ , 并使覆盖概率

达到一定水平(如多95%等等). 在实用上就认为 $g_1(\hat{x}) \leq g(\theta) \leq g_2(\hat{x})$

在内 都是在覆盖概率的一定水平下得到的, 因此这个区间通常称为置信区 间. 点估计与区间估计是参数估计的两个组成部分, 是相互联系、相辅 相成的两种估计方法. 但是它们也各有其特定的概念与问题. 本章第 7.1节介绍区间估计的概念与问题及其基本的求解方法, 即枢轴量法, 同时也介绍了单调似然比分布族参数的区间估计方法, 该方法可用于某 些离散型分布的区间估计; 第7.2节介绍区间估计与假设检验的内在联 系, 并在此基础上进一步介绍置信区间的优良性准则; 第7.3节介绍与 置信区间有密切关系的容忍区间和容忍限. 有关本章内容, 可参见陈希 孺(1981, 1999), Lehmann(1986), Shao(1998), Zacks(1981)等.

### 7.1置信区间及其枢轴量法

给定参数分布族本节主要讨论参数  $\theta$  的区间 估计问题, 因为函数 $g(\theta)$ 可看成一个新参数  $\eta = g(\theta)$  的区间估计问题; 方法完全类似. 另外, 若无特别说明, 也假定参数 $\theta$ 是一维的.

#### 7.1.1置信区间和置信限

定义7.1.1 若统计量 $T(X)$ 和 $g(\theta)$ 满足以下关系:

的值在区间 $[g_1(\hat{x}), g_2(\hat{x})]$

(以相当大的覆盖概率). 由于区间估计

306

## 第七章区间估计

设  $\{g_1(x), g_2(x) \mid 1 - \alpha \leq P(g_1(X) \leq g_2(X) \mid \theta) \leq 1 - \alpha, \forall \theta \in \Theta\}$  为参数 $\theta$ 的一个水平为 $1 - \alpha$ 的置信区间, 称 $g_1(x)$  和 $g_2(x)$  分别为 $\theta$ 的水平为 $1 - \alpha$  的置信上限和置信下限(亦称置信上、下 界), 若它们满足

$P(g_1(X) \leq g(\theta) \leq g_2(X) \mid \theta) \geq 1 - \alpha, \forall \theta \in \Theta$ .

以上定义亦可简记为  $P(g_1(X) \leq g(\theta) \leq g_2(X) \mid \theta) \geq 1 - \alpha$ ;

和 $P(g_1(X) \leq g(\theta) \leq g_2(X) \mid \theta) \geq 1 - \alpha$ . 在(7.1.1)式中,  $\theta$ 表示真参数,  $g(X)$ 和

义)表示其上、下限,该式表明,区间以不小于 $1-\alpha$  的概率覆盖真参数 $\theta$ 实用上就认为易见, 在一定的覆盖概 率下,覆盖范围(即平均区间长度)越小,区间估计的精度越高,区间估计的优良性问题将在下一节有;所讨论.关于置信上、下限,它们分别 相当于(7.1.1)式中 $-\infty$ 和 $+\infty$ 的情形.置信上、下限在实用上 也很重要,例如:若参数为材料强度或元件寿命,只需要考虑其值不低 于多少,即置信下限;而若参数为次品率或发病率,则只需要考虑其值

不高于多少,即置信上限.

若 $l=g(\theta)$ 为 $\theta$ 的严增函数,则由(7. 1. 1)式可得

$P\{U \leq g(\theta) \leq V\} = 1-\alpha, \forall \theta \in \Theta. (7. 1.4)$  由此可得以下常用的引理:

引理7.1.1 若 $g(\theta)$ 为 $\theta$ 的严增函数,  $[f(X), g(X)]$ 为参数 $\theta$

的一个水平为 $1-\alpha$  的置信区间, 则 $[g(\theta), g(X)]$ 为参数 $\theta$ 的一个水平为 $1-\alpha$  的置信区间.参数的置信上、下限亦有类似的结果.

类似于(7. 1. 1)式,我们也可定义多维参数 $\theta$ 的置信域.

定义7.1.2 若  $S(T)$ 为 $R^n$ 中的一个区域,使得  $P_{\theta} \{T \in S\} \geq 1-\alpha$ ,则称为参数 $\theta$ 的一个水平为 $1-\alpha$ 的置

信域.

关于多维参数的置信域,以下简单性质常常是有用的: . 引理7.1.2 若 $S_1(X)$ 为参数 $\theta$ 的一个水平为 $1-\alpha_1$  的置信域( $\alpha_1 =$

$\alpha_1, \alpha_2$ ), 则 $S_1(X) \cap S_2(X)$ 为参数 $\theta$ 的一个水平为 $1-\alpha_1-\alpha_2$ 的 置信域.

证明 记 $S_i^c(X)$ 为 $S_i(X)$ 的对立事件( $i = 1, 2$ ).则由定义可知  $P_{\theta} \{X \in S_i^c(X)\} \leq \alpha_i$  ( $i=1, 2$ ), 因此根据概率公式有

$P_{\theta} \{X \in S_1^c(X) \cup S_2^c(X)\} \leq \alpha_1 + \alpha_2$

( $\alpha_1 = \alpha_2 = \alpha$ )

(7. 1.2) (7. 1. 3)

## 7.1置信区间及其枢轴量法

307

$=1 - P_{\theta} \{X: (\theta - e(X), \theta + e(X)) \cap S_1(X) \cap S_2(X) \neq \emptyset\}$

$=1 - P_{\theta} \{X: (e(X), e(X)) \cap S_1(X) \cap S_2(X) \neq \emptyset\} \geq 1 - [P_{\theta} \{X: \theta \leq e(X)\} + P_{\theta} \{X: \theta \geq e(X)\}]$

( $\alpha_1 = \alpha_2 = \alpha$ )

I

的个数不一定是 2, 也可以是  $A: (Bickel \text{ and Doksum, } 1977 ; \text{Shao, } 1998 )$ , 但本章主要考虑一维单参数的置信区间和置信上、下限.置信上、下限 与置信区间有如下关系.

引理7.1.3 若 $e(X)$ 为 $\theta$ 的水平为  $1-\alpha$ 和 $1-\alpha$ 的置信下、上限, 则 $[e(X), e(X)]$ 为 $\theta$ 的水平为 $1-\alpha$ 的置信区间.

证明由定义可知

$(-\infty, e(X)) \cup (e(X), +\infty) \subset S_1^c(X) \cup S_2^c(X)$

而

$P_{\theta} \{X: (e(X), e(X)) \cap S_1(X) \cap S_2(X) \neq \emptyset\} = 1 - P_{\theta} \{X: \theta \leq e(X) \text{ or } \theta \geq e(X)\}$

$= 1 - P_{\theta} \{X: \theta \leq e(X) \text{ or } \theta \geq e(X)\} \geq 1 - [P_{\theta} \{X: \theta \leq e(X)\} + P_{\theta} \{X: \theta \geq e(X)\}] \geq 1 - \alpha - \alpha = 1 - 2\alpha$

构造置信区间的方法通常有以下几种:(1)从某一点估计出发的枢 轴量法;(2)大样本近似的枢轴量法;(3)通过假设检验的接受域构造 置信域.

## 7.1.2构造置信域的枢轴量法

首先通过一个例子来说明其意义.



例7.1.1 设 $X_1, \dots, X_n$ 为独立同分布样本,  $X_i \sim N(\mu, \sigma^2)$ , 求 $\mu$ 的水平为 $1-\alpha$ 的置信区间.

解 可从 $\mu$ 的估计 $\bar{X}$ 出发建立 $\bar{X}$ 与 $\sigma^2$ 的关系, 取  $G(\bar{X} - \mu, \sigma^2) = 1 - \alpha$ , 则函数 $G(\bar{X} - \mu, \sigma^2)$ 包含了 $\mu$ 与 $\sigma^2$ , 且其分布为 $Z(n-1)$ , 与 $\sigma^2$ 无关. 简记  $Z(n-1)$ , 则有  $1 - \alpha = P\{ -Z(n-1) \leq \bar{X} - \mu \leq Z(n-1) \}$ . 注意, 在引理7.1.2中, 参数 $\mu$ 不一定是一维, 可以是 $n$ 维, ②:

P 该式可改写为

$$A = 1 - \alpha.$$

### 308 第七章区间估计

因此可得 $\mu$ 的水平为 $1-\alpha$ 的置信区间为  $[\bar{X} - Z(n-1)\sigma/\sqrt{n}, \bar{X} + Z(n-1)\sigma/\sqrt{n}]$ , 其中

类似地可求出置信上、下限, 今求 $M(\bar{X})$ :

$$P\{ -Z(n-1) \leq \bar{X} - \mu \leq Z(n-1) \} = 1 - \alpha,$$

该式可改写为

$$\text{故置信下限为 } \bar{X} - Z(n-1)\sigma/\sqrt{n}, \text{ 上限为 } \bar{X} + Z(n-1)\sigma/\sqrt{n}.$$

另外, 若参数 $\sigma^2$ 已知, 则在 $C(X, M)$ 中用 $\sigma^2$ 代替 $\sigma^2$ 亦可得到  $\mu$ 的置信区间. 这时

服从标准正态分布, 因而有  $[\bar{X} - Z(n-1)\sigma/\sqrt{n}, \bar{X} + Z(n-1)\sigma/\sqrt{n}]$ .

]

枢轴量法的一般原则可归纳为

(1) 从 $\theta$ 的某一合适估计出发, 构造一个函数, 即枢轴量 $c(X, \theta)$ , 使其分布已知, 且与 $\theta$ 无关;

(2) 由  $P\{ -Z(n-1) \leq c(X, \theta) \leq Z(n-1) \} = 1 - \alpha$  解出

$$P\{ -Z(n-1) \leq c(X, \theta) \leq Z(n-1) \} = 1 - \alpha,$$

则可得到 $\theta$ 的置信区间  $[f(X), g(X)]$ . 置信上、下限则可通过 $G(JT, \theta)$ 的单边不等式反解得到, 以下再看几个例子.

例7.1.2 设  $X_1, \dots, X_n$ 为独立同分布样本,  $X_i \sim N(\mu, \sigma^2)$ , 求

$\sigma^2$ ,  $\mu$  的水平为 $1-\alpha$ 的置信区间以及 $\sigma^2$ 的置信上限.

解 与上例类似, 可从 $\sigma^2$ 的一个估计出发. 在 $P(J) = (n-1) \cdot$

$S(X) = (n-1) \sum_{i=1}^n (X_i - \bar{X})^2$ 中, 由于 $S(X)/\sigma^2$ 服从  $(n-1)$ , 因  $t=1$

此可取枢轴量为

$$G(X, \sigma^2)$$

$\sim;$

$a$

$$n - 1 \cdot J - a.$$

该式可化为

$a$

### 7.1置信区间及其枢轴量法

309

P

$$S(2) \sim \sigma^2 \cdot S(2) / 2t\alpha$$

«• (7.1.5)

X

$n-1$

,  $1 - \alpha = P\{ -Z(n-1) \leq \bar{X} - \mu \leq Z(n-1) \}$ , 了

因此可得 $\sigma^2$ 的水平为 $1-\alpha$ 的置信区间为  $[a^2(X), b^2(X)]$ , 其中

$$f^2(X) = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2, \quad p(x) = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2$$

斗

由于 $P\{ a^2(X) \leq \sigma^2 \leq b^2(X) \} = 1 - \alpha$ 等同于 $P\{ a(X) \leq \sigma \leq b(X) \} = 1 - \alpha$ , 因此 $a$ 的水平为 $1$

$1-\alpha$  的置信区间为  $[a(X), a(X)]$ . 同时, 根据类似于

(7.1.5) 的公式可得  $a^2$  的水平为  $1-\alpha$  的置信上限为  $(r^2(X) \wedge S(X)/x\{n-1, a\})$ .

另外, 若从已知为  $\sim$ , 则可取枢轴量为  $G(X, (r^2) \wedge T(X)/a^2 \sim/(n))$ , 其中

$T(X) = \sum_{i=1}^n (X_i - M)^2$ ; 详见例 7.1.7.  $|isI$

根据以上两个例子以及引理 7.1.2, 我们可以求出正态样本中参数  $(M, a^2)$  的联合置信域.

关于正态总体情形下参数置信域的求解问题, 大多可以通过假设检验问题的对偶接受域得到, 这将在下一节进行比较系统地讨论.

例 7.1.3 设  $X_1, \dots, X_n$  为独立同分布样本, 且  $X_1$  求没 的水平为  $1-\alpha$  的置信区间.

解 由  $3(X) = X(n)$ , 可取  $G(X, 0) = X(n)/6 - BE(n, 1)$ , 其分布与

0 无关, 且有 由此可取  $0 < a < 6 \wedge 1$  :

而  $a, 6$  可由以下积分决定

$\int_{ntn-1}^{di=6n-an} 1-a.$

因此, 由满足  $an - a$  的  $a, 6$  都可以得到没 的水平为  $1-\alpha$  的置信区间其中

$V_V$  没 (7)  $= -^1, 5(2) = J 21.$

—ba 对于本问题, 我们还可以在  $6n - an = 1 - a$  的约束条件下, 求使区间长

$11a$

$2$

$"-1, 1 -7$

$BE(n, 1) \sim ntn-1 f 1 J. = 1 -a,$

$310$

第七章区间估计

度最小的解, 即求  $-IT'$  的最小值点. 而直接验证可知,  $a'^* =$

$[-(1 - :)' - b_l]$  是  $t$  的减函数, 所以  $6=1, a'1 - b \sim l$

最小, 因此没 的水平为  $1-\alpha$  的置信区间可取为  $I$  例 7.1.4 设  $X_1, \dots, X_n$  为独立同分布样

本,  $\sim M + r(l/a,$

1), 求参数  $a$  和的水平为  $1-\alpha$  的置信区间以及它们的联合置信域. 解  $a$  和  $\#$  的点估计都与完备

充分统计量  $r = (x(1), s(x))$  有关,

其中  $s(x) = (\sum_{i=1}^n x_i - Ao)$  与  $x(1)$  独立, 并有  $i=1$

$x(1) \sim /x + r(\text{合}, 1), s(x) \sim r(\text{去}, \text{打}-1).$

$7$

$Pa\{/(2n-2, \text{号卜去}-f)\} = 1-a.$

$P\{z | 2S(X)/\wedge^2 2n-2, l-\text{爹})-2\} = 1-a.$  因此可得  $tr$  的水平为  $1-\alpha$  的置信区间为  $C(T(a)$

$= [(\wedge), 0-(\wedge)],$

因此可取  $G\}(X \wedge = -S(X) \wedge (2n-2),$  其分布与参数无关. 由此可得  $a$

其中  $-$

为求  $M$  的置信区间, 由  $I(1)$  的分布可知,  $2n(X(1) - /x)/a - /(2),$  且与  $S(X)$  独立. 因此

有

$GAX$  小)

$J(\text{又}(1) - M)/2$

$\sim 2$

$-S(X)/(2n-2)$

$a$

$n(n-1)(X(1) \sim /jl) - F(2, 2n-2). S(X)$

$a(X) = \frac{\dots}{\dots}, a(X) =$

$- / \text{卜}-2, 1-\text{幻} ?(2n-2, \wedge j$

$-.$

由此即可得到M的置信区间.今记 $F_{\alpha/2} = F(2, 2n-2; \alpha/2)$ ;  $F_{1-\alpha/2} = F(2, 2n-2; 1-\alpha/2)$ , 则有

$$P\left\{ \frac{a}{2} \leq S(X) \leq \frac{1}{1-\alpha} F_{1-\alpha/2} \right\} = 1-\alpha$$

因此可得M的水平为 $1-\alpha$ 的置信区间为 $Q(a) = [0(X), JZ(X)]$ , 其中

$$J(1) = \dots = J(1) -;$$

$$A(X) = X(I) - n(n.p F_{\alpha/25}(\cdot) -$$

7. 1 置信区间及其枢轴量法

311

根据引理7. 1.2可得 $\mu$ 和 $\sigma^2$ 的水平为 $1-\alpha$ 的联合置信域为 $C(a/2) \cap C(a/2)$ . I

例7.1.5  $y = (y_1, \dots, y_n)$  服从多元正态分布求参数 $\mu$ 的置信域.

解 由于  $y \sim N(\mu, \Sigma)$ ,  $I(y - \mu)^T \Sigma^{-1} (y - \mu) \leq \chi^2_{n-1}(1-\alpha)$  因此有  $|\mu - \bar{y}| \leq \sqrt{\frac{1}{n} \chi^2_{n-1}(1-\alpha)}$ .

由此可得 $n$ 维参数 $\mu$ 的置信域为

$$S(\mu) = \{ \mu : |\mu - \bar{y}| \leq \sqrt{\frac{1}{n} \chi^2_{n-1}(1-\alpha)} \},$$

即 $\mu$ 的置信域为以 $\bar{y}$ 为中心、以 $\sqrt{\frac{1}{n} \chi^2_{n-1}(1-\alpha)}$ 为半径的球. ■ 由于置信区间具体给出了未知参数的上、下限以及覆盖概率, 因此

它在很多情况下比点估计更加实用, 今举几个例子予以说明.

例 7.1.6 某厂生产的滚珠, 其直径可认为服从正态分布. 今从一批产品中抽取6个, 测得直径为14.70, 15.21, 14.90, 14.91, 15.32, 15.32 (毫米). 试估计这批产品直径的平均值; 并按两种情况求直径平

均值的置信区间: (1) 方差已知为0.05; (2) 方差未知 (取置信水平为 0.95).

6

解 样本均值为 $\bar{x} = 15.06$ . 可根据例7.1.1的公式求均值的置信区间, 其中,  $x_1, \dots, x_6$ 独立同分布,  $X \sim N(\mu, \sigma^2)$ ;  $n=6$ . (1) 若方差已知为0.05, 则 $\sigma^2 = 0.2236$ ,  $z_{0.975} = 1.96$ . 因此有

$$M = 15.06 - 0.2236 \times 1.96 / \sqrt{6} = 14.88; \quad f = 15.06 + 0.2236 \times 1.96 / \sqrt{6} = 15.24; \text{ 所以置信区间为 } [14.88, 15.24].$$

(2) 若方差未知, 直接计算与查表可得 $t_{0.025}(5) = 2.5706$ . 因此有 $M = 15.06 - 2.5706 \times \sqrt{0.05/6} = 14.76$ ;  $f = 15.06 + 2.5706 \times \sqrt{0.05/6} = 15.36$ ; 所以置信区间为 $[14.76, 15.36]$ . 这一置信区间的覆盖范围比方差已知情况要大一点, 这显然是合理的. ■ 例 7. 1. 7 为了解某型号

测量仪的精度, 用此仪器对一根长度为30 mm

的标准金属棒进行了6次测量, 其结果为30.1, 29.9, 29.8, 30.3, 30.2, 29.6. 设测量值服从正态分布 $N(\mu, \sigma^2)$ , 求均方差的点估计及其水平为0.95的置信区间.

解 记 $T(X) = (X_1, \dots, X_n)$ , 则 $a^2 = T(X) / n = 0.058$  ( $n=6$ ,  $\sum_{i=1}^6 x_i^2 = 174.0$ );  $a = 0.24$ . 可参照例7. 1.2的公式求解的置信区间. 当 $\sigma^2$ 已知时,  $T(X) / a^2 \sim \chi^2_{n-1}$ , 因此类似于(7.1.4)式有

312

第七章区间估计

$$F_{1-\alpha/2} = 1/\alpha.$$

1? 卜

1-f) 斗

f).

$$f_{2W} = \frac{1}{2} \Gamma\left(\frac{n}{2}\right) \left(\frac{1}{2}\right)^{n/2} \exp\left(-\frac{1}{2}x\right) \Gamma\left(\frac{n}{2}\right) \left(\frac{1}{2}\right)^{n/2} \exp\left(-\frac{1}{2}x\right)$$

在本问题中,  $x_1, \dots, x_6$ 独立同分布,  $X \sim N(0, \sigma^2)$ ;  $T(X) = 0.35$ . 经查表可得/

$(6, 0.975) = 14.4494$ ,  $^2(6, 0.025) = 1.2375$ . 因此有  
 $f_2 = 0.35/14.4494 = 0.0242$ ,  $a = 0.15$ ;  $a_2 = 0.35/1.2375 = 0.2828$ ,  
 $7 = 0.53$ . 因此  $r$  的水平为  $0.95$  的置信区间为  $[0.15, 0.53]$ . |

例7. 1.8 (续例6.3.5) 某电子元件的寿命服从指数分布  $r(l/\lambda, 1)$ , 在10个独立同分布的样本中, 仅观察到其中前6个元件的寿命分别为 896, 950, 1005, 1100, 1150, 1220(小时).

(1) 问这批元件的平均寿命不低于多少?

(2) 设  $p = P(X, > 1100)$  为元件寿命不低于1100小时的概率, 求  $P$  的置信区间.(置信水平取为  $1 - \alpha = 0.95$ ).

解(1) 本问题可看作例6.3.5的继续, 要求  $\theta$  的置信下限. 由于  $r$

$T_n.r/r$ , 其中  $rn < r = 2 + (n-r)X(z) \sim r(l/\lambda, r)$ . 因此可取枢轴  $1 = 1$

量为  $G(X, e) = 2rn r/\lambda^2(2r)$ . 由此可得  $P\{e^{Tnr} \leq 2 \dots, 1 - \alpha\} = 1 - \alpha$ ,

因此  $a$  的置信下限为  $f(J) = 27 \dots 2(2r, l_1)$ . 对于上述数据,  $n =$

10,  $r = 6$ ,  $a = 0.05$ , 经直接计算可得  $2Tnr = 22402$ , 而  $2(12, 0.95) = 21.026$ , 由此可得  $= 2Tnr/Y_2(12, 0.95) = 1065.44$ . 即这批元件

的平均寿命不低于1065.44(置信水平为0.95), 这与例6.3.5中断言元件的平均寿命大于1000小时的结论是一致的, 但是本例给出了平均寿命更具体的置信下限以及覆盖概率.

(2) 由于  $p = P(X > 1100) = e^{-1100/\lambda}$  为  $\lambda$  的严增函数, 所以根据引理7.1.1, 其  $\lambda$  信区间可由参数  $\theta$  的置信区间得到. 设  $\theta$  的置信区间为  $[l, u]$ , 则  $[f_l, f_u] = [e^{-1100/\lambda}, e^{-1100/\lambda}]$  为参数  $p$  的同一水平的置信区间.

类似于(1)的讨论可得  $\theta$  的置信区间为  $[0] = [27 \dots 2(2r, 1 - \alpha/2), 27 \dots 2(2r, \alpha/2)]$ .

“

,

## 7.1 置信区间及其枢轴量法

313

而  $/(12, 0.025) = 4.404$ ,  $/(12, 0.975) = 23.337$ , 由此可得  $\lambda =$

959.93,  $\theta = 5086.73$ . 因此有

$[f, g] = [e^{-1100/\lambda}, e^{-1100/\lambda}] = [e^{-L, 46}, e^{-2, 6}] = [0.32, 0.81]$ . |

### 7. 1. 3 基于渐近分布的枢轴量法

在许多情形下, 枢轴量  $c(x, \theta)$  的精确分布不易求得, 这时可考虑采用渐近分布, 特别可采用渐近正态分布和渐近  $\chi^2$  分布. 另外, 对离散型分布(诸如二项分布、Poisson分布等), 其分布函数和分位数的计算比较麻烦, 其参数的区间估计也常用基于渐近分布的枢轴量法.

(1) 渐近正态置信域

设  $\dots$ , 又独立同分布, 由第五章的定理5.3.6可知, 在一定正

则条件下, 参数  $\theta$  的极大似然估计  $\hat{\theta}$  具有渐近正态性, 即似 “ 其中  $K$  为一个样本的 Fisher 信息阵, 因此可取枢轴量为

$G(X, \hat{\theta}) = \sqrt{n}(\hat{\theta} - \theta) / \sqrt{I(\theta)}$ . (7. 1.6)

这时有  $G(X, \hat{\theta}) \xrightarrow{d} N(0, 1)$ , 其中  $I(\theta)$  为样本关于  $\theta$  的 Fisher 信息阵,  $\theta$  为  $P$  维. 同时根据定理5.3.6的推论3还可取

$C, (X, \hat{\theta}) = \{ \text{Var}(\hat{\theta}) \}^{-1/2} (\hat{\theta} - \theta)$ , (7. 1.7)

其渐近分布亦为标准正态, 与  $\theta$  无关. 今记  $\text{Var}(\hat{\theta}) = J$ ,  $i$  为其相合估计, 则由相合性以及 Slutsky 定理的去1律可得

$G(X, \hat{\theta}) = S \cdot (\hat{\theta} - \theta) / \sqrt{V(\theta, p)}$ . (7. 1.8)

特别, 若  $\theta$  为1维单参数, 并记  $\text{Var}(\hat{\theta}) = a^2$ ,  $\hat{a}$  为其相合估计, 则有  $A$

$$= 1 - a.$$

其中为标准正态分布的  $1 - \alpha$  分位点. 由此可得参数  $\theta$  的渐近正态 置信域为  $\theta \pm z_{\alpha/2} \sigma_{\theta}$ . (7. 1. 10)

同理, 亦可得到单边的置信上、下限. (7.1.10) 式有广泛的应用, 下面看两个例子.

例7.1.9 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim b(1, p)$ . (1) 求  $p$  的置信区间; (2) 应用: 某化学溶液配制过程中, 原有方案的成功率为

、  
k

### 314 第七章区间估计

70%, 现设计了一种新方案, 40次试验中有34次成功, 求新方案成功率的置信区间和置信下限(取置信水平为  $1 - \alpha = 0.95$ ); 新方案是否优于原有方案?

解 (1)  $p = X/n$ , 且有  $\text{Var}(f) = p(1-p)/n$ ,  $g=1$

$-p$ , 由《渐近正

态性或中心极限定理都可得

由此可得

即

$$G(X, P) = -$$

$$y \text{ Var}(p)$$

$$p\} z_{\alpha/2} \sqrt{p(1-p)}$$

$$y/pq/n$$

$$= 1 -$$

$$/n p(1$$

$$-p)$$

$$P\{t\{n(X-p)^2 / p(1-p)\} \leq \chi^2_{1-\alpha}\} = 1 - \alpha.$$

这是  $p$  的一个二次三项式, 可经过反解, 得到  $p$  的上下限  $\hat{p} \pm z_{\alpha/2} \sqrt{\hat{p}(1-\hat{p})/n}$ . 但在实用上, 经常可采用

(7. 1. 10) 式. 由于  $f = pq/n$  是  $o(1)$  的相合估计, 由 (7. 1. 10) 式即可得到  $p$  的置信区间为

$$\hat{p} \pm z_{\alpha/2} \sqrt{\hat{p}(1-\hat{p})/n},$$

其中  $j = 1, 2$ ;

(2) 本问题中,  $n=40$ ,  $X=34/40$ . 因此  $p=X=0.85$ ,  $q=0.15$ ,

.

$$\hat{p}^2 = 0.85 \times 0.15/40 = 0.0031875$$

而  $z_{0.975} = 1.96$ , 所以有  $p=0.85 \pm 1.96 \sqrt{0.0031875}$

. 另外, 由 (7.1.10) 式可得  $p$  的单边置信下限为由于  $f=0.85$ ,  $g=0.056$ ,  $z_{0.95}$

$=1.645$ , 因此  $p=0.757$ . 以上结果都说明, 新方案的成功率比原有方案高, 因而优于原有

方案. ■ 例7.1.10 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim P(\lambda)$ . (1) 求  $\lambda$  的置信区间和置信下限; (2) 应用: 在例6.7.5中, 求各个细胞单位所含白

血球平均个数的置信区间和置信下限(取置信水平为  $1 - \alpha = 0.95$ ). 解 (1)  $\lambda = \bar{X}$ , 且有

$\text{Var}(\bar{X}) = \lambda/n$ . 由于  $\bar{X} = \lambda$  具有渐近正态性, 而且  $\bar{X}$  为  $\lambda$  的相合估计, 因此直接由 (7.

1. 10) 式可以得到  $\lambda$  的置信区间为  $[\bar{X} - z_{\alpha/2} \sqrt{\bar{X}/n}, \bar{X} + z_{\alpha/2} \sqrt{\bar{X}/n}]$ , 其中  $\bar{X} = \bar{X}$ ,  $a = \lambda/n$ . 而  $\lambda$

的单边置信下限为  $\lambda = \bar{X}$

(2) 本问题中,  $n=1008$ ,  $\bar{X}=2.82$ ,  $a = 2.82/1008 = 0.0028$ ,

$$1.96 \sqrt{2.82/1008} = 0.11, \quad p=0.85 \pm 0.11 = 0.96, \quad \text{置信区间为 } [0.74, 0.96]$$

. 而  $\lambda$  的置信下限为  $\lambda = \bar{X} - z_{0.95} \sqrt{\bar{X}/n} = 2.82 - 1.645 \times 0.053 = 2.73$ . 本问题中, 由于样本容量  $n$  很大, 所以  $\lambda$  的置信区间和置信下限与其点估计很接近, 这

$-0.075 = 1.96$ . 由此可以得到A的置信区间为  $[2.72, 2.92]$

### 7.1 置信区间及其枢轴量法

315

也是合理的. 以上两个例题还可推广到两样本情形, 用于两个总体的比较. 例

如, 若  $X_i \sim P(\lambda_i), i = 1, 2$ , 可根据渐近正态性求解比例之差  $p_1 - p_2$

的置信区间和置信下限, 见第六章(6.5.35)式. Poisson分布亦

类似, 以下是两个正态总体比较的例子.

例7. 1. 11 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\mu, \sigma^2)$ ;  $Y_1, \dots, Y_m$

且

解 (1) 由假设可知  $S^2$  与  $\mu$  独立, 且有

独立同分布,  $K$ ,

区间(此即第六章提到的Behrens - Fisher问题); (2) 应用: 在例6.5.4 中, 求男女所含红血球的平均值之差的置信区间(取置信水平为0.95).

(1) 求  $\mu_1 - \mu_2$  的渐近正态置信

■

由于  $\hat{\mu}_1$  和  $\hat{\mu}_2$  分别为  $\mu_1$  和  $\mu_2$  的相合估计, 其中  $\hat{\mu}_1 =$

$\bar{X}, V(\bar{X} - \bar{X})^2 = S^2/n$ ;  $S^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2$ . 因此由Slutsky定理可得  $\sqrt{n}(\bar{X} - \mu_1) \rightarrow$

$N(0, \sigma^2)$

若记  $a = \mu_1 - \mu_2$ ; 区间为

$\bar{a} \pm t_{\alpha/2} \sqrt{S^2/n}$

则由(7. 1. 10)式可得  $\mu_1 - \mu_2$  的渐近正态置信

区间  $[\bar{a} - t_{\alpha/2} \sqrt{S^2/n}, \bar{a} + t_{\alpha/2} \sqrt{S^2/n}]$ . (2) 在例6.5.4中,  $n = 156, m = 74$ , 其样本容量都比梦大, 可应

用渐近正态置信区间. 由例6.5.4的数据可得  $\bar{x} = 465.13, \bar{y} = 422.16; S^2 =$

$42.97$ . 另有  $\bar{s}^2 = 54.8, S^2 = 49.2$ , 因此

$a = \bar{x} - \bar{y} = 465.13 - 422.16 = 42.97$ .

而  $t_{0.975} = 1.96, (t \times \sqrt{42.97/156 + 49.2/74}) = 14.13$ , 由此可以得到  $\mu_1 - \mu_2$  的渐近正态置信

区间为  $[28.84, 57.10]$ . 这一结果说明, 该地区男性所含红血球的平均值显著地高于女

性, 这与例6. 5.4的结论是一致的, 但更加具体. ■

(2) 似然置信域 似然置信域与前面介绍的渐近正态置信域是相通的, 在大多数情况

下, 渐近正态置信域直接应用极大似然估计的渐/近正态性(即定理

性(即定理5. 3. 8, 其实质就是渐近正态统计量的“平方”.

5. 3.6); 而似然置信域则应用似然比统计量的渐近

和定理5.3.9)

第七章区间估计 在定理5.3.8中, (5.3.3-0)式对任意的  $\theta$  成立, 因此有

权. (7.1.11)

$P_{\theta} \{LR(\theta) \leq a\} = 1 - \alpha$ . (7. 1. 12) 由此即可反解得到  $\{e^{S(X)} \leq 1 - \alpha$ ,

因此  $s(x)$  即为  $\theta$  的水平为  $1 - \alpha$  的似然置信域. 同时, 由定理5.3.8和定理5.3.9可知, 似然

比统计量的等价形式score统计量  $SC(\theta) = (\hat{\mu} - \mu_0)' / \sqrt{I(\mu_0)}$  (见 (5.3.35)式) 和

Wald统计量  $WD(\theta) = (\hat{\mu} - \mu_0)' / \sqrt{I(\hat{\mu})}$  (见 (5.3. 36)式) 也具有渐近  $\chi^2$  性, 因此在

(7. 1. 11)式和(7. 1. 12)式中可以用它们取代同样可以得到似然置信域; 当然表达形

式可能会

其中

316

$LR(\theta) = 2 \{Z(\hat{\mu})\}$  这时  $LR(\theta)$  就可以视为枢轴量, 由上式可得

有所差异. 另外, 我们在第六章曾经介绍了子集参数的似然比统计量  $LR(\theta)$

解

样本  $X = (X_1, \dots, X_n)$  的分布密度和对数似然函数分别为  $f(x; A) = e^{-AL}$

/ 性 (见 (6.6.6) 式). 同时也介绍了 score 统计量  $S(\hat{\theta})$  (见

的渐近 (6.6.14) 式) 和 Wald 统计量 (见 (6.6.17) 式) 的渐近

/ 性. 它

们都可以用来构造子集参数的似然置信域. 沿用 6.6 节的符号, 记  $\theta_0 = \theta_0(H)$ , 则由定理

6.6.1 的推论 1 有

$$= 2 \{ L(\theta_0) - L(\hat{\theta}) \} / I(\theta_0) \xrightarrow{d} \chi^2_1 \quad (7.1.13)$$
 这时  $LR(\theta_0)$  就可以视为枢轴量, 由上式可得

$$P\{|S(\hat{\theta})| \leq \sqrt{2 \ln(1/\alpha)}\} = 1 - \alpha. \quad (7.1.14)$$

由此即可反解得到  $P\{|S(\hat{\theta})| \leq \sqrt{2 \ln(1/\alpha)}\} = 1 - \alpha$ , 因此  $S(\hat{\theta})$  即为  $\theta$  的水平

为  $1 - \alpha$  的似然置信域. 类似地, 若在 (7.1.13) 式和 (7.1.14) 式中用 score 统计量

$S(\hat{\theta})$  或 Wald 统计量  $WD(\hat{\theta})$  取代, 亦可得到子

集参数的似然置信域; 其表达形式亦可能会有所差异. 进一步的讨论可参见 Shao (1998).

例 7.1.12 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $X_i \sim \text{Poisson}(\lambda)$ . 求  $\lambda$  的似然置信区间, 以及基于 score 统计量和 Wald 统计量的置信区间.

$$l(\lambda) = \sum_{i=1}^n \log f(X_i; \lambda) = n \log \lambda - \sum_{i=1}^n X_i.$$

$\lambda$ . 由此可得  $dL/d\lambda = T/\lambda - n$ ,  $I(\lambda) = \text{Var}(dL/d\lambda) = n/\lambda$ ,

7.1 置信区间及其枢轴量法 317

且有  $A = T/n$ . 由此可得似然比统计量为

$$LR(\lambda) = 2 \{ l(A) - l(\lambda) \} = 2 [n \log(T/n\lambda) - (T - n\lambda)].$$

类似地, 根据 (5.3.35) 式和 (5.3.36) 式可得  $SC(\lambda) = (n/\lambda) (A - \lambda)^2$ ;

$WD(\lambda) = (n/\lambda) (A - \lambda)^2$ . 这些结果代入 (7.1.12) 式, 并经过反解, 可以得到  $\lambda$  的置信区间. 在这一例题中,  $WD(\lambda)$  产生的置信区间与例 7.1.10 中的渐近正态置信区间十分相似, 而另外两个反解起来都比较麻烦. 一般来说, 对于

一维单参数, 大多采用渐近正态置信区间比较方便. 似然置信域更多地用于多参数情形.

例 7.1.13 设  $Y_1, \dots, Y_n$  其中

$(Z(W), \dots, Y(W))$  为已知函数, 求  $\theta$  的似然置信域, 以及基于 score 统计量和 Wald 统计量的置信域.

解由于  $r$  服从正态分布, 因此有

$$f(y) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2} \log(2\pi TCT^2)\right\} \exp\left\{-\frac{1}{2} \frac{y^2}{TCT^2}\right\} \quad (7.1.15)$$

其中  $S(\theta) = e^{T(\hat{\theta})} e(\hat{\theta})$ ,  $e(\theta) = Y - f(\theta)$ . 记  $\theta$  的极大似然估计为  $\hat{\theta}$ , 则由 (7.1.11) 式有

$$LR(\hat{\theta}) = 2 \{ l(\hat{\theta}) - l(\theta) \} = 2 \{ T \log \frac{f(\hat{\theta})}{f(\theta)} - Y(\hat{\theta} - \theta) \}. \quad (7.1.16)$$

由 (7.1.12) 式可得

$$P\{|S(\hat{\theta})| \leq \sqrt{2 \ln(1/\alpha)}\} = 1 - \alpha. \quad (7.1.17)$$

此即  $\theta$  应该满足的关系式, 由此即可确定  $\theta$  的置信域. 若  $\alpha$  未知, 则由 Slutsky 定理的推论 1 可知, 在 (7.1.16) 式中用  $\hat{\alpha}$  的相合估计  $\hat{\alpha}$  来代

替

$$(\hat{\alpha}) = \alpha - S(\hat{\theta}) \sqrt{2 \ln(1/\alpha)}$$

, 其渐近分布不变. 以, 由  $LR(\hat{\theta})$  确定的置信域可表示为

我们亦可求出基于 score 统计量  $SC(\hat{\theta})$  和 Wald 统计量  $WD(\hat{\theta})$  的置信域. 由 (7.1.15) 式可得

纪  $L = \sqrt{v(e)}v(0)9 a$

其中  $V(0) = df(0)/d\theta$  因此有  $1(8) = \text{Var}(\text{况/卵})^a 2VT(e)V(0)$ .

从而由 (5. 3. 35) 式和 (5. 3. 36) 式可得

T

糊  $= (\text{尝}) \text{广}(8) (\text{尝}) \text{夕}(4) \text{以咖}(8) \text{九} Z(P)$

$\text{ira}(4) = \sim i(0 - \text{erLv}\backslash o)v(0) - [(e - e)^x 2Cp),$

,

318

## 第七章区间估计

其中  $p_v(e) = v(e) [v_1'(e)v(e)rlvr(<0)]$ . 因此类似于 (7.1.17) 式可得  $Pe \setminus 6 - eT(\wedge)Pv(\wedge)e(0) \wedge crV(p, l - a) = 1 - a$ .

$Pe - 0)T[VT(e)V(0') - 0) a1X(P, 1 \sim a) 1 = 1 - a$ .

f

本小节介绍一个定理, 它可用于单调似然比分布族中参数的区间估计; 特别可用于某些离散型分布参数的区间估计. 本小节内容可参见 Shao(1998), 茆诗松等(1998) 及陈家鼎等(1993).

定理 7.1.1 假设  $U \sim / (K9)$ ,  $<9e0CR |$ , 统计量  $T \equiv T\{X、$  的分布函数记为  $Fr(i; 6 >)$ .

(1) 若对任意固定的  $G \text{ FA}h0'$  和心  $(/-0; 6 >)$  为  $<9$  的减函数 (见图 7. 1. 1 左), 并记  $\text{权}(1) = \sup | \}$ ,  $6\{t.\} = \inf j^: Fr(i-0; \wedge) 1-a |$ , (7.1.18)

则  $i(r)$  和  $yr$  分别为  $\theta$  的水平为  $1 - a$  的置信上、下限; (2) 若对任意固定的  $f$ , 在  $0 < F \wedge (9) < 1$  的范围内,  $FT(t; 0)$  和  $FT(t-0; e)$  为  $\theta$  的严格减函数, 并且处处连续, 则  $\theta$  的水平为  $1 - a$  的

置信上、下限  $e = e(T)$  和  $f =$  分别满足方程

$FT(z; \wedge) = a$ ,  $FT\{t-0; 3\} = 1-a$  (7.1.19)

并且方程的解唯一. 而  $\theta$  的水平为  $1 - a$  的置信区间  $[\wedge, \wedge] = [6(T), 0(T)]$  应满足方程  $FT\{t\backslash 6\} Fr(\ll \sim 0; 0) = 1 -$  (7.1.20)

证明 (1) 以下证明主要用到分位数的定义与性质, 可参见第一章及习题一. 由于  $FT(t; 0)$  为  $6 >$  的减函数, 因此由 (7. 1. 18) 式可知 (见图 7.1.1),  $Pe = P0 \setminus Fr(T; \&) \wedge a f$ . 由分位数的性质可知, 任一分布函数  $F(x)$ ,  $F(x')$  等价于  $x' \gg xp$ , 其中  $\setminus$  为  $F(x)$  的  $p$  分位数, 因此, 若记  $ta$  为  $Fr(t; \wedge)$  的  $a$  分位数, 则有

$pe = P0 \setminus FT(T; 0) \wedge a | = P0 \setminus T \wedge tQ \} = 1 - FT(ta - a$ .

同理, 若  $a_2$  未知, 则可用  $a_2$  的相合估计 变. 由此即可以得到基于 score 统计量和 Wald 统计量的置信域分别为  $C_2(Y) = \setminus e: e \setminus 6) Pv(\wedge) e(\wedge) \wedge a Y(pJ-a) |$  和  $C_3(y) = H(6-a)T \bullet$

$[VT(0)V(\wedge)](\wedge-0)CaY(p, l-a))$ ; 其中  $f$  为  $a_2$  的某一相合估计. I

### 7.1.4 单调似然比分布族参数的区间估计

来代替  $a_2$ , 其渐近分布不

#### 7.1 置信区间及其枢轴量法

319

另一方面, 由于  $FT(t - 0; 0)$  为的减函数, 因此由 (7.1.18) 式可知 (见图 7.1.1),  $Po | = Pe \setminus Ft(T-0; 0) - a$ . 由分位数的性质可知, 任一分布函数  $F(x)$ , 若  $x'$  矣则有  $F(x' - 0) < F(xp -$

$0) \wedge p$ . 因此, 若记为  $FT(t; 0)$  的  $1 - a$  分位数, 则事件 " $Fr(T-0; 幻 \setminus 1-a$ " 包含事件 " $7 \wedge 1/$ , 因此有

$P0 \setminus e \wedge em \setminus = Pe \setminus FT(T-0; 0) \wedge l-a \setminus \wedge Po \{T \wedge_a \} = Fr(\wedge_{ft}; \wedge) \wedge l-a$ .

(2) 若  $FT(t; 0)$  和  $GG-0; 幻$  为  $\theta$  的严格减函数, 并且处处连续, 则 (7.1.18) 式等价于 (7.1. 19) 式, 并且方程的解唯一, 因而的水平 为  $1-a$  的置信上、下限满足 (7. 1. 19)



式.根据引理7.1.3,  $\theta$ 的水平为  $1 - \alpha$ 的置信区间满足(7.1.20)式. |

下面定理7.1.2将证明:在一定条件下, 单调似然比分布族(见定义6.3.1)存在相应的分布函数, 能够满足定理7.1.1的要求, 首先看两个例子.

例7.1.14 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim \text{Exp}(\lambda)$ . 求  $M$  的水平为  $1 - \alpha$  的置信区间.

解 容易验证: 该分布族关于完备充分统计量  $T = \sum_{i=1}^n X_i$  为单调似然比分布族(见第六章习题), 而  $r = x(1)/n + r(n, 1)$ ; 所以  $r = x(1)$  的分布函数为  $Fr(t; M) = [1 - e^{-n(t-1)}]^n$ . 显然,  $FT(t)$  为  $M$  的严格减函数, 处处连续, 并且有  $Fr(t-0; M) = FT(t; M)$ . 因此可根据 (7.1.19) 式和 (7.1.20) 式求解的置信上、下限和置信区间. 例如,

根据(7.1.20)式可得

$$\begin{aligned} 1 - \alpha &= [1 - e^{-n(t-1)}]^n \\ 1 - \alpha^{1/n} &= 1 - e^{-n(t-1)} \\ 20 \end{aligned}$$

## 第七章区间估计

由此即可解出  $M$  的水平为  $1 - \alpha$  的置信区间为  $k(X)$ , 其中

$$k(X) = [X(1) - \frac{1}{n} \log(\alpha/2), X(1) + \frac{1}{n} \log(1 - \alpha/2)]$$

例如, 若  $n=10, \alpha=0.05$ , 则  $[e(X), \text{弘}(\alpha)] = [X(1) - 0.3689, X(1) - 0.0025]$ . 根据枢轴量法所得的置信区间为  $[X(1) - \frac{1}{n} \log(1 - \alpha/2), X(1) - \frac{1}{n} \log(\alpha/2)]$ ,

当  $n=10, \alpha=0.05$  时,  $\frac{1}{n} \log(1 - \alpha/2) = 0.3689, \frac{1}{n} \log(\alpha/2) = 0.0025$ .

二者非常接近. 定理7.1.1的主要价值在于它能够应用于若干常见的离散型分布的

区间估计, 例如, Poisson 分布、二项分布、负二项分布等.

例7.1.15 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim \text{Poisson}(\lambda)$ .

求  $A$  的水平为  $1 - \alpha$  的置信上、下限和置信区间(见范金城, 吴可法, 2001).

解 容易验证: 分布族关于完备充分统计量为单调似然比  $i=1$  分布族(见第六章), 而  $r \sim p(nA)$ . 由第一章的公式可知,  $r$  的分布函数  $Fr(MA)$  可表示为不完全  $\gamma$  函数的形式(见(1.2.5)式):

显然 离散型分布, 因此有

$$P(T \leq t) = \sum_{k=0}^t \frac{e^{-\lambda} \lambda^k}{k!} \quad (7.1.21)$$

, 为  $A$  的严格减函数, 并且处处连续. 由于  $T \sim p(nA)$  为

$$P(T \leq t) = P(A \leq \frac{t}{n}) = P(T \leq t) = \sum_{k=0}^t \frac{e^{-\lambda} \lambda^k}{k!}$$

(7.1.20) 式求解  $A$  的置信上、下限和置信区间. 例如, 根据(7.1.19)式的第二式,  $A$  的置信下限应满

$$Fr(t-0; X) = 1 - \alpha, \text{ 代入 (7.1.21) 式可得}$$

$$FT(t-1) = \sum_{k=0}^{t-1} \frac{e^{-\lambda} \lambda^k}{k!} \quad (7.1.22)$$

该式被积函数对应于一个  $r$  分布的密度函数, 即  $z \sim r(i, 0, r(0))$ , 因此(7.1.22)式等价于

$$\int_0^t \frac{dx}{r(x)} = \sum_{k=0}^{t-1} \frac{1}{r(k)} = 1 - \alpha \quad (7.1.22)$$

由于  $2Z \sim r(l/2, 0, Y_2(2t))$ , 因此由  $P(2Z > 2nA) = 1 - \alpha$ , 所以  $2^A = \frac{1}{2} \log(1 - \alpha)$ , 即  $A$  的水平为  $1 - \alpha$  的置信下限为

$$ACT) = (2n)^{-1/2} \log(1 - \alpha)$$

## 7.1 置信区间及其枢轴量法

321

根据(7.1.19)式的第一式,  $A$  的置信上限应满足  $FT(t; A) = \alpha$ , 代入(7.1.21)式可得

$FT(t; \lambda) = \int_0^t e^{-\lambda x} \lambda dx / r(t+1) = a.$

该式被积函数也对应于一个  $r$  分布的密度函数, 这时  $\sim r(1+1)$ ,  $2Z' \sim / (2i+2)$ , 因此上式等价于

$dx/r(f+1) = P(>nA) = P(2Zr > 2nA) = a.$

所以  $2nA = r(2r+2; 1-a)$ , 即  $A$  的水平为  $1-a$  的置信上限为  $= (2n) - y(2T+2; 1-a)$ . 根据引理 7.1.3,  $A$  的水平为  $1-a$  的置信区间为  $[(2n) - y(2T; a/2); (2n) - y(2T+2; 1-a/2)]$ . I

最后证明单调似然比分布族的一个重要性质, 即其相应统计量  $r(x)$  的分布函数为参数  $\theta$  的减函数, 因而可以应用定理 7.1.1 来求解参数的置信上、下限和置信区间. 事实上, 可以得到一个更一般的结果.

定理 7.1.2 假设  $yV(\cdot, \theta)$ ,  $\theta \in CR$  | 关于  $T=T(X)$  为单调似然比分布族.

(1) 设  $g(\cdot)$  为  $f$  的非减函数, 则  $g(\cdot) = E_j(A(n^j))$  为  $\theta$  的非减函数;

(2) 设统计量  $T=T(X)$  的分布函数记为  $F_t(j, \theta)$ ,

则对任意固定的  $z$ ,  $F_r(t_0; \theta)$  和  $F_r(z, -\theta; \cdot)$  为  $\theta$  的减函数. 证明 (I) 设  $\theta \in CR$  则

$A(x) = f(x; \theta_2) / f(x; \theta_1)$  为  $r=r(x)$  的非减函数, 即  $A(x) = /l(T(x); \theta_1, \theta_2)$ ,  $h$  为  $T = T(x)$  的非减函数. 要证

尽  $(A) W(\cdot; \theta) \in E, 2[(A(r(j))) = g(\cdot^2)]$ . 为此, 记  $R+ = jx: /(\theta, \cdot) > /(\cdot, \cdot) |$ ,  $R- = |x: f(x; \theta_2) < /(\cdot, \cdot)$ , 则有  $A(xz) > 1, \forall x'$

$A(xz) A(x)$ , 与  $A(x') > 1$ ,  $A(x) < 1$  矛盾. 记  $a = \sup_{T \in R-} A(x)$ ,  $b$

$= \inf_{x \in R+} A(x)$ , 则  $b \geq a$ . 以下积分在  $R^0 = |x: /(\cdot, \cdot) = f(x; \theta_2) |$  上为零, 因此有

客  $(h) = \int f(x; \theta_2) / f(x; \theta_1) dx$

U.

因此若  $x' \in T$ ,  $x \in R-$ , 则必有  $T(xf) >$

$e/T$ ;  $A(\cdot) < 1$ ,  $r(x); \text{汐}(\cdot) \sim (r(x))$ , 因为若  $T(x) \wedge T(x)f$  则由于  $A(x)$  为  $r=r(x)$  的非减函数, 应有  $\wedge / i(r(x); A, \cdot)$ , 即

322

## 第七章区间估计

I 少  $(H(\cdot) / (L(\cdot)) \cdot / (太, 4))$  如  $(x)$

$a \in [ / (x, \theta_2) ] d/z(x)$

$= (b - a) [f(x, \theta_2)]$  如  $(x)$  多  $\theta$ .

上面的推导用到了  $[ / (\cdot, \cdot^2) \text{ 如 } (\cdot) = -f + [f(x, \theta_2) - f(x, \theta_1)] d^{\wedge}(\cdot)$ , 这是因为  $\cdot) (Vt(\cdot) + f [f(x, \theta_2) - f(x, \theta_1)] d^{\wedge}(\cdot)$

(2) 在 (1) 中, 取  $\theta = /ba \in CR$ , 则它是  $\theta$  的非减函数, 因而  $E_j(A(r(X))) = P^m$

$> h$  为  $\theta$  的非减函数, 所以  $FT(t; \theta) = 1 - Pe(T(X) > \theta)$  为  $\theta$  的非增函数. 同理, 取

$(A'(z))$ , 则它也是  $\theta$  的非减函数, 因而  $E, [ \text{汐}'(7 \setminus \cdot) ] = / \theta(r(\cdot) X)$  为  $\theta$  的非减函数,

所以  $FT(t; -\theta; \cdot) = Pe(T(X) < \theta) = 1 - Pe(T(X) \wedge t)$  为  $\theta$  的非增函数. |

另外, 在定理 7.1.1 中, 若假设  $FT(t; \theta)$  和  $\wedge(\cdot; \theta)$  为  $\theta$  的增函数, 也应该有完全平行的结果, 只是其应用比较少.

定理 7.1.3 假设  $U \sim /(\cdot)^{\wedge} \theta \in CR$ , 统计量  $T=T(X)$  的分布函数记为  $\wedge(M)$ .

(1) 若对任意固定的  $G, FT(t; \theta)$  和  $\wedge(\cdot; \theta)$  为  $\theta$  的增函数 (见图 7.1.1 右), 并记

$\theta(t) = \inf_{\theta: FT(t; \theta) \leq \alpha}, \theta(0) = \sup_{\theta: FT(t; \theta) \leq 1 - \alpha},$

则 $\theta(t)$ 和 $\theta(0)$ 分别为 $\theta$ 的水平为 $1 - \alpha$ 的置信上、下限;

(2) 若对任意固定的 $t$ 在 $0 < FT(t; \theta) < 1$  的范围内,  $FT(t; \theta)$ 和  $FrU(0; \theta)$ 为 $\theta$ 的严格增函数并且处处连续, 则 $\theta$ 的水平为 $1 - \alpha$ 的置

信上、下限 $e = e(t)$ 和 $f = f(t)$ 分别满足方程  $Fr(z; \theta) = \alpha, FT(t; \theta) = 1 - \alpha$

并且方程的解唯一. 而 $e$ 的水平为 $1 - \alpha$ 的置信区间 $J = [\theta(t), \theta(0)]$ 应满足方程  $-|Fr(Fr(i; \theta)) = 1$

$= f. [/($

$-V, ^2$

$))$

$]d/z(x) = 0.$

## 7.2 参数置信域与假设检验的接受域

323

### 7.2参数置信域与假设检验的接受域

读者可能已经发现, 上节似然置信域中用到的似然比统计量就是第

六章似然比检验中用到的统计量(见(6.6.4)式与(6.6.6)式). 事实上,

参数的置信域与参数的假设检验本来就存在内在联系. 例如在 (7.1.11)式中, 由我们由

$\{LR(6) - \alpha\} = 1 - \alpha$  可解出置信

域; 另一方面, 问题的否定域, 其余集就是接受域. 因此, 由假设 检验问题的接受域可以得到参数的置信域; 反之, 由参数的置信域也可 得到假设检验问题的接受域. 本节首先介绍根据假设检验的接受域构造 参数置信域的方法, 然后通过假设检验的最优性准则引导出参数置信域的最优性准则, 即一致最准确置信域(UMA). 为简单起见, 以下的讨论主要限于非随机化检验, 因而主要适用于连续型分布的情形. 随机化 检验与参数置信域的关系比较复杂, 实际中很少用到, 可参见Shao (1998).

#### 7.2.1对偶关系

例7. 2.1 设 $X_1, \dots, X_n$ 为独立同分布样本, 且有 $X_i$ , 考虑假设检验问题

" $H_0: \theta = \theta_0$  vs  $H_1: \theta > \theta_0$ . 由于 $X_1, \dots, X_n$ 的联合分布为指数族分布

$f(x, \theta) =$

因此关于 $\theta = \theta_0$ 为单调似然比分布族, 所以否定域为 $R = \{t: \sum_{i=1}^n$

$-a\} = \alpha$ 也可得到假设检验问

的否定域满足

因而接受域满足

$n\}$ . 由于  $\theta = \theta_0$ , 因此关于  $H_0$

$\theta_0, 1 - \alpha$

324

## 第七章区间估计

注意, 该式对任意的 $e = e_0$ 都成立, 因而有

$[2T [ \sim \theta$

该式可以反解为

这说明, 没的置信下界为 $0W = 2T/Y^2(2n, 1 - \alpha)$ . I 以上例子的论证可以推广到一般情形, 假设检验的接受域与参数置

信域之间有以下对偶关系: (1) 由假设检验的接受域得到参数的置信域. 考虑假设检验问题已知其否定域为 $(\theta)$ , 接受域为 $A(6) J(6)$ , 且满足

$\sim I$ 尤 $e \times (\theta) ! \gg 1 - \alpha$ , 若该式对任意成立, 且 $Ue4(\theta_0)$  | 可反解为  $|\theta_0 eS(\%) i$ , 则有  $\backslash XeA(\theta) \backslash = / \% l^eS(X) \}$ 多 $1 - \alpha$ . 因此  $|X: ^eS(X) ($ 为 $\theta$ 的水平为 $1 - \alpha$ 的置信域.

(2) 由参数置信域得到假设检验的否定域与检验函数. 设  $|X: \theta eS(X) |$ 为没的水平为 $1$  的置

信域, 若  $\{x \in \mathcal{X} \mid P_0(x) \leq \alpha\}$  可反解为  $\{x \in \mathcal{X} \mid P_0(x) > 1 - \alpha\}$ . 该式对任意的  $0 < \alpha < 1$  亦

成立, 因此有

$$P_0(\mathcal{X} \mid \theta_0) = P_0(\mathcal{X} \mid \theta_0) > 1 - \alpha. \quad P_0(\mathcal{X} \mid \theta_0) \leq 1 - \alpha$$

因此集合  $\mathcal{X}(\alpha)$  可以认为是某一假设检验问题:  $H_0: \theta = \theta_0$  的否定域, 而其相应的检验函数可表示为

$$\phi(x, \theta_0) = \begin{cases} 1, & x \in \mathcal{X}(\alpha) \\ 0, & x \notin \mathcal{X}(\alpha) \end{cases}$$

在第六章, 我们研究了许多常见分布有关参数的假设检验问题, 因此根据以上对偶关系, 我们也就平行地解决了相应参数的置信域问题. 这也就大大地节省了篇幅, 所以通常都是把区间估计放在假设检验之后予以介绍(当然, 由置信域亦可得到相应假设检验问题的否定域). 我们 还要特别强调以下两类重要情况:

(1) 正态总体情形下参数的置信域. 第六章的表 6.5.1 和表 6.5.2 总结了正态总体情形下参数假设检验问题的检验统计量以及否定域, 根据对偶关系, 这两个表也可用于求解参数的置信域. 事实上, 由表中的

$$2(n-1, 1-\alpha) = 2(n-1, \alpha)$$

## 7.2 参数置信域与假设检验的接受域

325

检验统计量即可得到枢轴量; 而由表中的否定域即可得到接受域, 从而 得到置信域. 因此我们根据假设检验的接受域与参数置信域之间的对偶 关系就一举解决了正态总体情形下参数的置信域问题(因此我们在

7.1.2 没有系统介绍正态总体情形下参数置信域的求解方法). (2) 我们在第六章 6.6 节介绍了似然比检验, 以及相应的似然比 统计量和子集参数的似然比统计量. 因此由上述对偶关系可以得到相应

的置信域, 这就是 7.1.3 节介绍的似然置信域. ); 例 7.2.2 设  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\mu, \sigma^2)$ . (1) 求  $\mu$  的置信区间; (2) 应用: 某厂 有两条袋装果酱生产线, 今从第一条生产线随机抽取 12 个样品, 测得重量 平均值为 10.6 (克), 标准差为 1.55 (克); 从第二条生产线随机抽取 17 个样品, 测得重量平均值为 9.5 (克), 标准差为 2.17 (克). 假设重量服从正态 分布, 方差相等, 求平均值之差的置信区间(置信水平为 0.95).

解 (1) 考虑则由 6.5.3 类似的讨论可知 (见 (6.5.32) 式)

$$T = \frac{\bar{X} - \mu}{\sqrt{S^2/n}} \sim t(n-1)$$

其否定域为  $|T| > t_{\alpha/2}(n-1)$ , 接受域为  $|T| \leq t_{\alpha/2}(n-1)$

(, 且满足

由此可

其中  $z = x - \mu$ ,

$$np \left( \frac{z}{\sqrt{S^2/n}} \right)^2 = 1 - \alpha,$$

2)  $n + m$

$$4 = \left( \frac{z}{\sqrt{S^2/n}} \right)^2 = \frac{(X - \mu)^2}{S^2/n}$$

而  $c$  可取为  $t(n+m-2)$  分布的分位点, 即  $c = t_{\alpha/2}(n+m-2)$ , 1 得  $S$  的置信区间为

$$z \sim \sqrt{\frac{S^2}{n+m-2}} \cdot t_{\alpha/2}(n+m-2)$$

,

(7.2.1) (2) 由假设可知,  $n=12$ ,  $m=17$ , 因此自由度为  $n+m-2=27$ ;

$t_{0.025}(27) = 1.89$ ,  $t_{0.025}(27) = 1.89$ . 又由假设可知,  $\bar{X} = 10.6$ ,  $\bar{Y} = 9.5$ ;  $S_1^2 = 1.552$ ,  $S_2^2 = 2.172$ . 由于  $S_1^2/S_2^2 = 0.71$ ,  $t_{0.025}(27) = 1.89$ , 因此  $S^2 = \frac{(1.552 \times 11 + 2.172 \times 16)}{27} = 1.016$ ,

$S_{\alpha} = 10.08$ . 而  $27, 0.975) = 2.0518$ , 由  
326

## 第七章 区间估计

(7.2.1) 可得  $8 = h - j_{112}$  的置信区间为  $[-0.401, 2.601]$ . 7.2.2 一致最准确置信域

由于置信域与假设检验的接受域存在对偶关系, 因而可以通过假设检验的优良性准则来描述置信域的优良性.

定义 7.2.1 参数  $\theta$  的水平为  $1 - \alpha$  的置信域  $S$  (幻称为一致最准确置信域 (UMA)), 若对任一水平为  $1 - \alpha$  的置信域  $S^*(\%)$  有

$\wedge S(X) \subset (X);$ , 一切  $\theta$  关 2. (7.2.2) 在 (7.2.2) 式中,  $\theta$  表示真参数,  $V$  表示非真参数. 因此该定义的

意义为: 一致最准确置信域  $S(X)$  包含非真参数  $V$  的概率最小.

定义 7.2.2 称  $S$  (幻为  $\theta$  的一个水平为  $1 - \alpha$  的无偏置信域, 若它满足

$| \wedge^{1 - \alpha}, eS(X) | \wedge^{1 - \alpha}$ . (7.2.3) 若对所有的无偏置信域  $S$  (幻,  $5^*(\%)$ , 它们都满足 (7.2.2) 式, 则称

$5$  (幻为一致最准确无偏置信域 (UMA)). (7.2.3) 式表示:  $S(a;)$  包含真参数的概率大于等于  $1 - \alpha$ , 包含非

真参数的概率小于等于  $1 - \alpha$ .

定义 7.2.1 和定义 7.2.2 实际上与假设检验有密切关系, 我们可以

根据置信域与假设检验的接受域之间的对偶关系导出它们最优解之间的对偶关系. 为此, 考虑假设检验问题

$= \text{氏} - * (\wedge \theta)$ . (7.2.4) 假设上述检验问题的一个水平为  $1 - \alpha$  的检验函数为

$[1, \wedge g_4(6 > \theta)],$

$\text{少}(\text{龙}A) = \{1.0$

$, \wedge e A(\wedge \theta),$

(7.2.5)

其中其否定域为  $A$  (氏), 接受域为  $4$  (氏), 其对偶置信域表示为  $S(x) \setminus$ . 检验的功效函数为  $UW = EJ\theta(X \wedge \theta)J$ . (7.2.6)

假设  $(x, \theta_0)$  为另外一个检验函数, 其相应的量分别记为  $(\%) 4^*(\text{沒.})$ ,  $\setminus \theta^S X x \setminus$  和  $\text{氏}$  其形式与 (7.2.5) 式和 (7.2.6) 式类似. 定理 7.2.1 在以上条件下, 置信域  $S$  (幻为一致最准确置信域的

充要条件为相应的检验函数  $\text{小}(X, \text{么})$  为假设检验问题 (7.2.4) 的一致最优势检验, 且对  $V$   $\%$  成立.

证明以下根据置信域与假设检验的接受域之间的对偶关系分两步予以证明.

,

## 7.2 参数置信域与假设检验的接受域

327

(1) 必要性. 设  $UeS(\%)$  为一致最准确置信域, 要证明相应的  $\text{心}(10)(\theta)$  为一致最优的功效函数, 即对检验 (7.2.4) 的任一检验函数  $\text{</>}^*(\%,$

$e\theta$ ) 及其相应的功效函数  $\text{氏.}(x., \theta) (\theta)$  有

$heoy(\theta) \text{外} \cdot (10) W, V\text{祥} 2\theta.$

根据 UMA 的定义 (7.2.2) 式以及对偶关系可得  $Pe \setminus \theta^S(X) \wedge P\theta WeS^*(X) |, |$

$Xe4(6 >') \text{f} | Xe4^*(6 >') \}$ ,

$-7\% \text{沒}') | \text{名} 1 - / \% UeA^*( < 9') I,$



$$= \int_{-\infty}^{\infty} \frac{1}{\sigma} e^{-\frac{|x-\theta|}{\sigma}} dx = 2$$

此即(7.2.9)式, 证毕.  $\square$

推论1 若  $f(x)$  为的水平为  $1-\alpha$  的一致最准确置信区间, 则对任一水平为  $1-\alpha$  的置信区间  $J^*(x)$  有

推论2 定理7.2.2及推论1的结论对一致最准确无偏置信域亦

$EJ^*(x)$  即一致最准确置信区间的平均长度最小.

"

.

## 7.2 参数置信域与假设检验的接受域

329

I

域.但是也有类似的结果, 今以置信下限为例予以说明(置信上限的情况完全类似).根据定义7.2.1, 水平为  $1-\alpha$  的一致最准确置信下限

$f(x)$  应满足  $P[f(x) \leq \theta] = 1-\alpha$  以及

$P[f(x) \leq \theta] = 1-\alpha$  以及  $f(x) \leq \theta$  且  $f(x) \leq \theta$  且  $f(x) \leq \theta$  (7.2.10)

其中  $(x)$  为  $\theta$  的任一水平为  $1-\alpha$  的置信下限 ( $f$  表示非真参数).直观上, UMA 置信下限应该比其他置信下限更靠近真参数(见图7.2.1), 这可表述为

$\int_{-\infty}^{\infty} f(x) dF(x) \leq \int_{-\infty}^{\infty} g(x) dF(x)$

图7.2.1 UMA 置信下限

定理7.2.3 设  $X \sim P(f(x))$  为连续型分布,  $f(x)$  为  $\theta$  的水平为  $1-\alpha$  的一致最准确置信下限,  $\theta^*(x)$  为  $\theta$  的任一水平为  $1$  的置信下限, 并假设  $f(x)$  和  $\theta^*(x)$  都有连续型分布且期望存在. 今记  $a = \theta$ , 若

(7.2.11)

; 它们的分布

函数分别记为  $F(u)$  和  $G(w)$ . 由定义可知  $F(u) = 0$ , 若  $u < 0$ ; 对

$u > 0$  有

$F(u) = P[f(x) \leq u] = P[f(x) \leq \theta] = 1 - P[f(x) > \theta] = 1 - P[e^{-u/\theta} > 0] = 1 - P[e^{-u/\theta} > 0]$ . 若记  $\theta' < 0$ , 则上式

可表示为

$1 - F(u) = P[f(x) > u] = P[f(x) > \theta] = 1 - P[f(x) > \theta] = 1 - P[f(x) > \theta]$ .

同理, 对  $F(u)$  也有:

$1 - F(u) = P[f(x) > u] = P[f(x) > \theta] = 1 - P[f(x) > \theta] = 1 - P[f(x) > \theta]$

$1 - F(u) = P[f(x) > u] = P[f(x) > \theta] = 1 - P[f(x) > \theta] = 1 - P[f(x) > \theta]$

(7.2.12) 根据概率论中的常用公式(见王梓坤, 1979; 李贤平, 1997), 正值随机变量  $U$  的数学期望可表示为  $E(U) = \int_0^{\infty} [1 - F(u)] du$  同理也有  $E(U^*) = \int_0^{\infty} [1 - F(u)] du$ . 因此(7.2.12)式两边取积分可得

成立.

注意, 以上推论1 对置信上、下限不成立, 因为它们对应于无界区

$a > 0$ ;  $a = 0$ , 若  $a \leq 0$ ; 则  $J$  必有

$EJ = \int_{-\infty}^{\infty} f(x) dx + \int_{-\infty}^{\infty} g(x) dx$

, 证明 今记  $h = k - f(x)$ ,  $h = [0 - f(x)] +$

$= \int_0^{\infty} \{ \theta - G(x) \} dx$ .

$EJ^* = \int_0^{\infty} [1 - F(u)] du = \int_0^{\infty} [1 - F(u)] du = E f(U^*)$ .

330

## 第七章区间估计

由此即可得到(7.2. 11)式.  $\theta \in I$  推论 设 $\theta(X)$ 为 $\theta$ 的水平为 $1-\alpha$ 的一致最准确置信上限; $e(x)$

为 $\theta$ 的任一水平为 $1-\alpha$ 的置信上限, 并假设 $\theta(X)$ 和 $e(X)$ 都有连续型分布且期望存在, 则必有

$$E[e(X) - \theta] + E[\theta - \theta(X)] \geq \alpha, \quad V[e(X)] \leq \alpha E[e(X)]$$

在第六章, 我们对许多常见分布得到了一致最优势检验或一致最优无偏检验, 因此根据定理7.2. 1和定理7.2.2, 它们相应的对偶置信域必然是一致最准确置信域或一致最准确无偏置信域, 并且有最小的平均区间长度或平均体积. 特别, 对于正态分布, 第六章的表6.5.1和表

6.5.2总结了正态总体情形下参数假设检验问题的UMPT或 UMPUT, 由此我们根据这两个表以及对偶关系即可一举得到相应参数的UMA或 UMAU置信域. 例如在例7.2.2中,  $\theta$ 的置信区间必为UMAU置信区

间, 并且其平均区间长度在 $\theta$ 的无偏置信区间中最短. 另外, 根据单调似然比分布族得到的单边检验的UMPT, 由对偶关系即可得到相应参数的UMA置信上、下限.

例7.2.3 (续例7.2.1) 设 $X_1, \dots, X_n$ 为独立同分布样本,  $X_i \sim r(l/\theta, 1)$ . 求 $\theta$ 的水平为 $1-\alpha$ 的UMA置信下限和UMAU置信区间.

解 由于 $X_1, \dots, X_n$ 的联合分布为指数族分布,  $T = T(X) = \sum_{i=1}^n X_i$ ,

为其完备充分统计量, 因此对例7.2.1的单边检验, 那里由 $r$ 给出的否定域及其检验函数必为UMPT, 因而由此得到的对偶置信下限必然为UMA置信下限.

至于 $\theta$ 的UMAU置信区间, 则要考虑双边检验:  $H_0: \theta = \theta_0$  vs  $H_1: \theta \neq \theta_0$ . 由定理6.4.4可知, 该检验的UMPUT 所对应的否定域为

$\{T < k_1 \text{ 或 } T > k_2\}$ , 其相应的接受域为 $A(\theta_0) = \{k_1 \leq T \leq k_2\}$ . 由于  $2T/\theta_0 \sim \chi^2_{2n}$  因此接受域可表示为

其中常数 $A$ , 由定理6.4.4的(6.4.24)式决定, 即  $A = \chi^2_{2n, 1-\alpha/2}$ ,  $E[f(X)] = 1 - \alpha$ ,

$E[f(X)] = \int_{A(\theta_0)} f(X) dP_{\theta_0}(X) = 1 - \alpha$ . 为便于积分, 这两式亦表示为

$$E[f(X)] = 1 - \alpha,$$

$$f(X) = \frac{1}{2n} \left( \frac{2T}{\theta_0} \right)^{n-1} e^{-\frac{2T}{\theta_0}} \quad (7.2.14)$$

由于 $2T/\theta_0 \sim \chi^2_{2n}$ ,  $f(X)$ 在 $\chi^2_{2n}$ 上的值为1, 因此根据

(7.2.13)和(7.2.14)式,  $c_1, c_2$ 的定解条件可表示为

### 7.3 容忍区间与容忍限

331

J

$\int_{c_1}^{c_2} f(y) dy = 1 - \alpha$

,

$\int_{c_1}^{c_2} f(y) dy = 1 - \alpha$

$$\int_{c_1}^{c_2} f(y) dy = 1 - \alpha. \quad (7.2.15)$$

(7.2.15)

由于(7.2.13)-(7.2.15)式对任意的 $\theta_0$ 都成立, 因此 $\theta$ 的UMAU置信区间为

$$C(X) = \{ \theta : -c_1 \leq \frac{2T}{\theta} \leq c_2 \}$$

其中常数 $c_1, c_2$ 由(7.2.15)式决定. 显然, 为了得到UMAU置信区间,  $c_2$ 的求解比较麻烦, 通常实用上还是取为  $[e^{-c_1/2}, e^{c_2/2}]$

$[e^{-c_1/2}, e^{c_2/2}]$ , 这相当于由水平为  $1 - \alpha/2$  的 UMA置信上、下限构成. ■

以上例7.2.3 反映的情况有一定代表性. 根据假设检验的接受域与参数置信域的对偶关



系，我们可以很容易地从单边检验的UMPT得到 UMA置信上、下限。但是，由双边检验的定理6.4.4及其定解条件 (6.4.24)式求解UMAU置信区间则通常比较麻烦。实用上则可取水平为 $1 - \alpha/2$ 的UMA置信上、下限构成置信区间(见引理7.1.3)。另外，如果能用定理6.4.5，即 $T(X)$ 的分布关于某个点对称，则能很容易地从UMPUT得到UMAU置信区间，见例7.2.2。

### 7.3容忍区间与容忍限

容忍区间和容忍限与置信区间和置信限有密切的关系，其计算也依赖于置信区间和置信限的相应方法，但两者研究的问题有原则的区别。

#### 7.3.1问题与定义

设 $X_1, \dots, X_n$ 为独立同分布样本，且有 $X_i \sim T(\mu, \sigma)$ 。本节不是考虑参数 $\theta$ 的置信区间，而是考虑随机变量 $\hat{\theta}$ 的“置信区间”，称之为容忍区间(tolerance interval)，即希望求 $f = P\{T(\hat{\theta}) \in [T, T']\} \geq 1 - \alpha$ ，使

$P\{T(\hat{\theta}) \in [T, T']\} \geq 1 - \alpha$ 。(7.3.1) 以下先举两个例子，然后再介绍其严格定义。

例7.3.1 设某种机床主轴长度为100 cm，允许误差为 $\pm 0.5$  cm。生产中要求99%以上的产品达到以上要求，即轴长在(99.5, 100.5)区间内。今对一批主轴，测试了 $n$ 根轴长，其值为 $X_1, \dots, X_n$ ，假设为独立同分布样本，问这批主轴是否合格？

332

#### 第七章区间估计

解 解决这一问题的一种方法就是确定一个上下界 $T = T(\hat{\theta})$ 和 $T' = T'(\hat{\theta})$ 使得 $P\{X_i \in [T, T'] | \hat{\theta}\} \geq 99\%$ ，若 $(T, T') \subset (99.5, 100.5)$ ，则说明以上轴长合格，这就可以归结为求解容忍区间

的问题，具体计算见下一小节。■ 例7.3.2 设砖块的强度 $X$ 服从对数正态分布，要求产品强度 $P(X \geq f) \geq 99\%$ (如 $f = 120$ )。今对一批产品，测试了 $n$ 个砖块的强度为 $X_1, \dots, X_n$ ，问这批砖块是否合格？

解 解决这一问题的一种方法就是确定一个下界 $I = I(X_1, \dots, X_n)$ ，使得 $m \in [f, +\infty) | I \geq f\} \geq 99\%$ 。若 $I \geq f$ ，则说明这一批砖块合格，这就可以归结为求解容忍下限的问题。■ 下面考虑容忍区间与容忍上、下限的定义。为简单起见，考虑连续型分布的情形。在(7.3.1)式中， $X_i$ 的分布未知或依赖于未知参数；而 $f(X_1, \dots, X_n)$ 和 $T(X_1, \dots, X_n)$ 的形式和分布都未知，因此(7.3.1)式左端的计算十分困难。可考虑一个基本等价的形式。易见，若 $z$ 和 $\gamma$ 为常数，则(7.3.1)式可表示为

$P\{U \in [r, f] | F, f\} = F(f) - p$ 。(7.3.2) 但实际上 $I(X)$ 和 $f(X)$ 为随机变量， $F$ 和 $F_e(T)$ 亦为随机变量，所以“ $I \geq f$ ”为随机事件，因此我们只能要求(7.3.2)式以概率为1成立，即对充小的 $y$ 有

$P\{I \geq f\} \geq 1 - \alpha$ 。(7.3.3) 该式与(7.3.1)式的意义基本一致。由此可以得到以下定义：

定义7.3.1 设 $X_1, \dots, X_n$ 为独立同分布样本， $X_i \sim T(\mu, \sigma)$ 。若 $I = T(X)$ ， $T = f(X)$ 满足(7.3.3)式，则称 $[I, T]$ 为分布 $F_3(X)$ 的一个水平为 $(1 - \alpha, 1 - \gamma)$ 的容忍区间；若 $f(X_1, \dots, X_n)$ 和 $I(X_1, \dots, X_n)$ 满足以下关系式，则称它们为分布的水平为 $(1 - \alpha, 1 - \gamma)$ 的容忍上限和容忍下限：

$P\{I \geq f\} \geq 1 - \alpha$ 。(7.3.4)

$P\{T \leq f\} \geq 1 - \gamma$ 。即 $P\{I \geq f\} \geq 1 - \alpha$ 。

(7.3.5) 在这个定义中，(7.3.3)式表示 $X_1 \in (T, T')$ 以很大的概率成立。在(7.3.3)式中，若 $f = -\infty$ ，则得到(7.3.4)式，该式表示 $X \in (-\infty, T)$ 以很大的概率成立。同理，在(7.3.3)式中，若 $T = +\infty$ ，则得到(7.3.5)式，该式表示 $X, G(P, \infty)$

以很大的概率成立。|

计算容忍区间和容忍上、下限的主要方法为：(1) 容忍区间可由容

### 7.3 容忍区间与容忍限

333

忍上、下限得到；(2) 容忍上、下限可由分布 $F_{\theta}(x_j)$ 分位数的置信上、

下限得到。下面分别予以介绍。

定理7.3.1 若 $f(X)$ 和 $g(X)$ 分别为分布 $F_{\theta}(x)$ 的水平为

$(1-\alpha, 1-\beta)$ 的容忍上、下限，贝 $IJ[F(X), G(X)]$ 为其水平为 $(1-\alpha, 1-\beta)$ 的容忍区间。

证明 今以 $A$ 表示事件“ $\bar{C}_m \geq \alpha$ ”； $B$ 表示事件“ $\bar{C}_n > 1-\beta$ ”； $C$ 表示事件“ $\bar{C}_n - \bar{C}_m \leq 1-\alpha$ ”，则由定义7.3.1可知

$P(A \cap B) > 1-\alpha$ ， $P(B) \geq 1-\beta$ 。(7.3.6)

而要证 $P(C) \geq 1-\alpha$ 。由以上定义可知，若事件 $S$ 同时成立，则必有 $\bar{C}_n - \bar{C}_m \leq 1-\alpha$ ，

即 $C$ 成立，因此 $C \supset A \cap B$ ， $P(C) \geq P(A \cap B)$ 。由此可得

$P(C) \geq P(A \cap B) \geq 1-\alpha$ 。 $P(C) \geq 1-\alpha$ 。由此可得

$P(C) \geq P(A \cap B) \geq 1-\alpha$ 。 $P(C) \geq 1-\alpha$ 。由此可得

$P(C) \geq P(A \cap B) \geq 1-\alpha$ 。 $P(C) \geq 1-\alpha$ 。由此可得

### 7.3.2 容忍上、下限的计算

下面的定理说明， $F_{\theta}(x)$ 的容忍上、下限可由其分位数的置信上、下限得到。

定理7.3.2 设 $x_1, \dots, x_n$ 为独立同分布样本， $A$ 服从连续型分布 $F_{\theta}(x)$ ，则 $\bar{C}_n(X)$ 为其水

平为 $(1-\alpha, 1-\beta)$ 的容忍上限的充要条件为 $\bar{C}_n(X)$ 是 $x_{\alpha}^{(n)}$ 的水平为 $1-\alpha$ 的置信上

限； $\bar{C}_m(X)$ 为其水平为 $(1-\alpha, 1-\beta)$ 的容忍下限的充要条件为 $\bar{C}_m(X)$ 是 $x_{\beta}^{(m)}$ 的水平为 $1-\beta$ 的

置信下限。其中 $\alpha$ 和 $\beta$ 分别为分布 $r(x)$ 的 $1-\alpha$ 分位数和 $\beta$ 分位数。

证明 可直接从容忍上、下限以及分位数的定义来推导。 $\bar{C}_n(X)$ 的水平为 $(1-\alpha, 1-\beta)$ 的容忍上限的定义及其等价

形式为

$P(\bar{C}_n \leq x_{\alpha}^{(n)}) \geq 1-\alpha$

334 第七章区间估计

$\Leftrightarrow P(\bar{C}_n \leq x_{\alpha}^{(n)}) \geq 1-\alpha$ 。该式说明 $\bar{C}_n(X)$ 为参数的水平为 $1-\alpha$ 的置信上限。

同理， $\bar{C}_m(X)$ 为 $F_{\theta}(x)$ 的水平为 $(1-\alpha, 1-\beta)$ 的容忍下限的定义及其等价形式为

$P(\bar{C}_m \geq x_{\beta}^{(m)}) \geq 1-\beta$ 。该式说明 $\bar{C}_m(X)$ 为参数 $\theta$ 的置信下限，证毕。|

为了求解分位数和 $\theta$ 的置信上、下限，通常需要根据给定的分布，导出它们与参数 $\theta$ 之

间的关系，然后再根据这些关系求解。

例7.3.3 设 $x_1, \dots, x_n$ 为i.i.d.样本， $A \sim M(r, l)$ 。求此分布水平为 $(1-\alpha, 1-\beta)$ 的

容忍上、下限和容忍区间。

解 根据定理7.3.2，只需分别求出分位数和 $\theta$ 的置信上、下限即可。首先，由假设条件可

求出分位数与参数 $m$ 之间的关系。由于 $2(X) \sim \chi^2(2)$ ，因此 $P(X \leq x_p) = \alpha$ 可化为

$P(2(X) \leq 2(x_p)) = \alpha$ ； $2(x_p) = \chi^2_{\alpha}(2)$ 。

1 因此有 $x_p = \chi^2_{\alpha}(2)/2$ 。由于 $\theta$ 为 $M$ 的线性严增函数，因此由引理

7.1.1，由 $M$ 的置信下限即可得到 $\theta$ 的置信下限。而 $m$ 的置信下限可类似于例7.1.4得到。由

于 $X(1) \sim x + r(n, l)$ ， $2n(X(1) - x) \sim \chi^2(2)$ ，因此有

$P(2n(X(1) - x) \leq \chi^2_{\alpha}(2)) = 1-\alpha$ 。

由此可得 $P(X(1) \leq x + \chi^2_{\alpha}(2)/(2n)) = 1-\alpha$ 。所以 $\theta$ 的水平为 $1-\alpha$ 的

置信下限可表示为  $M = \bar{X} - z_{1-\alpha/2} \sqrt{\frac{s^2}{n}}$ ; 当置信水平为  $1-\alpha$  时, 置信下限为  $A = \bar{X} - z_{1-\alpha/2} \sqrt{\frac{s^2}{n}}$ . 因而分布  $M \sim N(\mu, \sigma^2/n)$  的水平为  $(1-\alpha, 1-\alpha)$  的容忍下限为  $I(x) = X(1) - z_{1-\alpha/2} \sqrt{\frac{s^2}{n}}$ .

类似地, 水平为  $(1-\alpha, 1-\alpha)$  的容忍上限为  $f(X) = X(n) - z_{1-\alpha/2} \sqrt{\frac{s^2}{n}}$ .

### 7.3 容忍区间与容忍限

335

$-z_{1-\alpha/2} \sqrt{\frac{s^2}{n}}$ . 由定理 7.3.1, 水平为  $(1-\alpha, 1-\alpha)$  的容忍区间为  $(\bar{X} - z_{1-\alpha/2} \sqrt{\frac{s^2}{n}}, \bar{X} + z_{1-\alpha/2} \sqrt{\frac{s^2}{n}})$ .  
尤(1)  $z_{1-\alpha/2} \sqrt{\frac{s^2}{n}}$  为  $(1-\alpha, 1-\alpha)$  的容忍限.

例 7.3.4 设  $X_1, \dots, X_n$  为独立同分布样本, 且有  $X_i \sim N(\mu, \sigma^2)$ . (1) 求该分布水平为  $(1-\alpha, 1-\alpha)$  的容忍上、下限和容忍区间; (2) 应用: 某厂生产乐器用镍合金线, 今从一批产品中随机抽取 10 个样品,

测得其抗拉强度为

10 512, 10 623, 10 668, 10 554, 10 776, 10 717, 10 557, 10 581, 10 666, 10 670. 求该镍合金线抗拉强度的容忍下限 (设置信水平为  $(0.95, 0.95)$ ).

解 (1) 根据定理 7.3.2, 我们先求人的置信下限. 为此, 首先

求出  $\bar{X}$  与参数  $(\mu, \sigma^2)$  之间的关系. 由假设可知  $P(\bar{X} - z_{1-\alpha/2} \sqrt{\frac{s^2}{n}} \leq \mu) = 1-\alpha$ . 由于  $(\bar{X} - \mu) / \sqrt{\frac{s^2}{n}} \sim N(0, 1)$ , 因此有  $(\bar{X} - \mu) / \sqrt{\frac{s^2}{n}} = z_{1-\alpha/2}$  其中  $z_{1-\alpha/2}$  为标准正态分布的  $1-\alpha/2$  分位数, 为已知. 因此可根据以上关系以及枢轴量法求  $\mu$  的置信下限. 由于

$\bar{X} - \mu$

可取枢轴量为

代入上式可得  $c(10) = 7 = z_{1-\alpha/2}$

$s \sqrt{\frac{1}{n}}$

~

数. 记  $t(n-1, 1-\alpha/2)$  为此非中心  $Z$  分布的  $1-\alpha/2$  分位数, 则有  $-t(n-1, 1-\alpha/2) \leq \bar{X} - \mu \leq t(n-1, 1-\alpha/2)$ .

$1)/(n-1)$

$-t(n-1, 1-\alpha/2) \sqrt{\frac{s^2}{n}}$ .

即  $G(X, 0)$  服从非中心  $f$  分布, 与参数  $\theta$  无关; 其中  $w$  为非中心参

336

### 第七章 区间估计

上式  $\bar{X} - z_{1-\alpha/2} \sqrt{\frac{s^2}{n}}$  常记为  $A = A(n, 1-\alpha/2, y)$ , 因此  $\bar{X}$  的置信下

限, 即正态分布的容忍下限可表示为

$\bar{X} - z_{1-\alpha/2} \sqrt{\frac{s^2}{n}}$

$T = \bar{X} - AS, A = A(n, 1-\alpha/2, y) = z_{1-\alpha/2} \sqrt{\frac{s^2}{n}}$

. 注意,  $A(n, 1-\alpha/2, y)$

$y$  有表可查, 例如, 可参见茆诗松, 王静龙 (1986).

(2) 本问题中,  $n = 10$ , 置信水平为  $(0.95, 0.95)$ , 查表可得  $A =$

2.91. 经简单计算可得  $\bar{X} = 10 632.4, S^2 = 6 738.77, S = 82.09$ . 这些数据代人

(7.3.7) 式可得  $T(X) = \bar{X} - AS = 10 632.4 - 2.91 \times 82.09 = 10 393.52$ .

因此这批镍合金线的抗拉强度不低于 10 393.52.

### 7.3.3 应用次序统计量计算容忍限

设  $X_1, \dots, X_n$  为独立同分布样本,  $X_i$  为连续型分布,  $f(X)$  和  $F(X)$  分别为其水平为  $(1-\alpha, 1-\alpha)$  的容忍上、下限. 由定理 7.3.2 可知,  $T(X)$  为  $X$  的水平为  $1-\alpha$  的置信上限;  $I(X)$  为  $X$  的水平为  $1-\alpha$  的置信下限.

1-y的置信下限.而分位数 $u_{1-\alpha}$ 和 $u_{\alpha}$ 的置信上、下限常常可通过次序统计量求出.设样本的次序统计量为 $X(1) \leq X(2) \leq \dots \leq X(n)$ , 直观上看, 在次序统计量中, 可取右端最小的人 $X(1)$ 为 $1-\alpha$ 的置信上限;而取介左端最大的 $X(n)$ 为 $\alpha$ 的置信下限.下面将证明这一结论.

平),  $t_{\alpha/2, n-1}$

图7.3.1 应用次序统计量计算容忍限

为此, 记  $Y(i) = F(X(i))$ ,  $i=1, \dots, n$ , 则根据第一章的讨论有

$Y(i) \sim \text{Beta}(i, n-i+1)$ . (7.3.8) 另外, 不完全函数记为 $\Gamma(x)$ , 且有(见第一章)

$\Gamma(x) = \Gamma(x+1)/x$ . (7.3.9) 定理7.3.3 设为独立同分布样本,  $X_1, \dots, X_n$  为连续 $\sim F(x)$ .

类似 $\chi^2$  可得 $T = X + A(n, \alpha/2, y)S$ , 而水平为 $(1-\alpha/2, 1-y)$ 的容忍区

间为 $[X - A(n, 1-\alpha/2, y/2)S, X + A(n, \alpha/2, y/2)S]$

型分布, 则其水平为 $(1-\alpha, 1-y)$ 的容忍上、下限及容忍区间满足以下(7.3.7)

习题七

337

关系:

(1)  $T(X) = X(j)$ ,  $j = \min\{j: 1 - y \leq F(X(j))\}$ ;

(2)  $T(X) = X(i)$ ,  $i = \max\{i: F(X(i)) \leq 1 - y\}$ ; (3)  $[T(X), T(X)] = [X(j), X(i)]$ , 其中 $Z, \alpha$ 应满足关系

$1 - \alpha + i - j \geq 1 - y$ . 证明(1)由定理7.3.2可知为分布 $F(X)$ 的容忍上限的充要

条件为为 $A_4$ 的置信上限), 其等价条件为

$1 - P\{X(j) \leq x\} \geq 1 - y$

$\Leftrightarrow 1 - 1 - \alpha + j - 1 \geq 1 - y$  (由(7.3.8)式)  $\Leftrightarrow 1 - \alpha + j - 1 \geq 1 - y$  (由(7.3.9)式).

取满足上式最小的人 $j$ 并记为 $j$ , 则有 $T(X) = X(j)$ .

(2) 类似地,  $T(X)$ 为水平为 $1 - y$ 的置信下限的等价条件为

$P\{X(i) \leq x\} \geq 1 - y$

$(i, n-i+1) \geq 1 - y$ . 取满足上式最大的 $i$ 并记为 $i$ , 则 $T(X) = X(i)$ .

(3) 由定义可知, 若 $[X(j), X(i)]$ 为 $F(X)$ 的水平为 $(1-\alpha, 1-y)$ 的容忍区间, 则其等价条件为

$\Leftrightarrow 1 - \alpha + i - j \geq 1 - y \Leftrightarrow 1 - \alpha + i - j \geq 1 - y \Leftrightarrow 1 - \alpha + i - j \geq 1 - y$

以上推导用到了性质 $K(i) \sim \text{Beta}(i, n-i+1)$  (见第一章).

习题七

1. 设 $X_1, \dots, X_n$ 为i.i.d.样本,  $X \sim E(\lambda, 1) = M(r(\lambda, 1))$ , 试用枢

轴量法求 $M$ 的一个水平为 $1 - \alpha$ 的置信区间.

2. 设 $\dots$ 人为 $d$ .样本,  $X \sim T(A, r)$ , 但 $p$ 已知. 试用枢轴

量法求 $A$ 的一个水平为 $1 - \alpha$ 的置信区间和置信上、下限. 3. 设 $X_1, \dots, X_n$ 为i.i.d.样本, 试用枢轴量法求以下分布中相应

参数水平为 $1 - y$ 的置信区间.

(1) 设 $X$ 服从冷分布即 $f(x) = \lambda e^{-\lambda x}$ ;

$p = 1 - y$

$\{F(X)\} \sim \text{Beta}(1, \lambda)$

$-y$

## 第七章区间估计

(2) 设  $X$  服从 Weibull 分布, 即  $f(x) = \frac{a}{\Gamma(a)} x^{a-1} \exp(-x^a)$ ,  $x > 0$ , 但  $a$  已知;

(3)  $X$  服从 Pareto 分布  $PR(a, 0)$ , 即  $f(x) = \frac{a}{\Gamma(a)} x^{-a-1}$ ,  $x > 0$ ,  $a$  已知或  $a$  已知.

4. 设  $X_1, \dots, X_n$  为 i.i.d. 样本. (1) 若  $X \sim V(6^2)$  ( $\sigma^2 > 0$ ), 找一个枢轴量并用此枢轴量构造  $\sigma^2$  的一个水平为 1 的置信区间; (2) 若  $X_1 \sim V(0, \sigma^2)$  ( $\sigma^2 > 0$ ), 求  $\sigma^2$  的一个水平为 1 -  $\alpha$  的置信区间.

5. 设  $X_1, \dots, X_n$  为样本,  $X \sim N(\mu, \sigma^2)$ , 其中  $\sigma^2 > 0$  已知,  $\mu > 0$  未知. 试找一个枢轴量, 并用此枢轴量构造  $\mu$  的一个水平为 1 的置信区间.

\*6. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $A$  服从均匀分布的同变估计为

7. 从某自动车床加工的一批零件中随机抽取 10 个, 测得其长度 (毫米) 分别为 12.01, 12.15, 12.08, 12.09, 12.16, 12.03, 12.06, 12.13, 12.07, 12.11.

假定长度服从正态分布, 置信水平取为 0.95. (1) 求平均长度的置信区间; (2) 该自动车床加工的精度  $\sigma$  在什么范围内?

8. 假设某元件的寿命服从指数分布.

(1) 从一批产品中随机抽取 9 个, 测得其寿命 (小时) 分别为 150, 450, 500, 530, 600, 650, 700, 830, 910.

问这批产品的平均寿命不低于多少 (置信水平为 0.95)? 又设  $p$  为这批产品中寿命大于 600 小时的比例, 求  $p$  的置信区间;

(2) 从一批产品中随机抽取 20 个, 测得其前 10 个的寿命 (小时) 分别为 500, 920, 1380, 1510, 1650, 1760, 2100, 2320, 2350, 2900. 求这批产品平均寿命的置信区间 (置信水平为 0.95).

9. 设  $X_1, \dots, X_n$  及  $Y_1, \dots, Y_m$  为 i.i.d. 样本,  $X \sim N(\mu, \sigma^2)$ ,  $Y \sim N(\mu, \sigma^2)$ , 且两总体独立. (1) 求  $p = \frac{\sigma^2}{\sigma^2 + \sigma_0^2}$  的置信区间; 试构造  $\mu$  的一个水平为 1 -  $\alpha$  的置信区间. (提示:  $Y = X + Z$ ,  $Z \sim N(0, \sigma_0^2)$ )

(2) 求  $T = \frac{Y}{X}$  的密度函数).

## 习题七

(2) 设甲、乙两位化验员独立地对某聚合物的含氮量用相同方法各测量 10 次, 其样本方差分别为 0.5419 和 0.6050. 假定测量值服从

正态分布, 求方差比  $p = \frac{\sigma_1^2}{\sigma_2^2}$  的置信区间 (置信水平为 0.95).

设  $X_1, \dots, X_n$  为 i.i.d. 样本,

$Y_1, \dots, Y_m$  为

10.

i.i.d. 样本,  $Y \sim N(\mu, \sigma^2)$ ; 且两总体独立. 若已知, 求

$\mu$  的水平为 1 -  $\alpha$  的置信区间. 11. 设  $X_1, \dots, X_n$  为 i.i.d. 样本

$Y_1, \dots, Y_m$  为, 且两总体独立. 若  $m = n$ , 求  $d = \frac{\sigma_1^2}{\sigma_2^2}$  水平为 1 -  $\alpha$  的置信区间. 若  $m = kn$  ( $k$  为已知正整数), 能否求解同样

i.i.d. 样本,  $Y \sim N(\mu, \sigma^2)$

$\mu$ ,  $\sigma^2$ )

问题?

12. 某养猪场分别用标准配方饲料和新配方的饲料喂养两组猪，一段时间后，在两组猪中各随机抽取10头，测得其体重增加值分别为  
新配方:36.0, 32.7, 39.2, 37.6, 32.0, 40.2, 34.4, 30.7, 36.4, 37.2.  
标准配方:35.2, 30.0, 36.5, 38.1, 29.4, 36.0, 31.3, 31.6, 31.1, 34.0.

假设两组猪体重的增加值都服从正态分布，求体重平均增加值之差的置信区间(置信水平为0.95).

13. 设 $X = (X_1, X_2)$ ,  $i=1, 2, \dots, n$ 为i.i.d.样本,  $X \sim N(\theta, \Sigma)$ ;  $P$ , 令 $\theta = \theta_0 + \delta$ ,  $\delta \neq 0$ . 求 $\theta$ 的一个水平为 $1-\alpha$ 的置信区间(通常称为Fieller区间)(提示:令 $Y = \theta_0 + \delta$ ,  $i=1, 2, \dots, n$ , 以 $K_1(\hat{Y}), \dots, K_n(\hat{Y})$ 为基础构造枢轴量).

•14. 设 $\mu_1, \dots, \mu_n$ 为i.i.d.样本,  $X_i \sim N(\mu_i, \sigma^2)$ ;  $\mu_1, \dots, \mu_n$ 为i.i.d.样本,  $Y_i \sim N(\mu_i, \sigma^2)$ , 且两总体独立. 求 $\mu_1/\mu_2$ 的一个水平为 $1-\alpha$ 的置信区间(提示:证 $y^2/x^2$ 为枢轴量).

15. 设 $f(A, \lambda) = r(A\lambda)$ ,  $i=1, 2$ , 且 $X_1, X_2$ 独立.

(1) 令 $\theta = A_1/A_2$ , 证明: $T = \frac{X_1^2}{X_2^2}$ 是一个枢轴量, 并用它构造 $\theta$ 的一个水平为 $1-\alpha$ 的置信区间;

(2) 令 $\theta = (A_1, A_2)$ , 证明: $T = \frac{X_1^2 + X_2^2}{X_1^2 + X_2^2}$ 是一个枢轴量, 并用它构造 $\theta$ 的一个水平为 $1-\alpha$ 的置信区域.

16. 设 $X_1, \dots, X_n$ 为i.i.d.样本,  $X_i \sim r(A, \lambda)$ ,  $A > 0$ ;  $K_1, \dots, K_m$ 为i.i.d.样本,  $Y_i \sim r(\lambda, \mu_i)$ ,  $\mu_i > 0$ , 且两总体独立. 试基于统计量 $r$

$n/m$ 构造 $\theta = (1/\mu)$ 的水平为 $1-\alpha$ 的置信上界.

1

弋尽

340

第七章区间估计 17. 若 $\theta_i$ 为参数 $\theta_i$ 的一个水平为 $1-\alpha$ 的置信域" $=1, \dots, n$ , 则只 $\theta_i = \theta_i$ "

其中 $a_1, \dots, a_n$ .

为参数 $\theta_1, \dots, \theta_n$ 的一个水平为 $1-\alpha$ 的置信域,

$i=1, \dots, n$ 设 $X_1, \dots, X_n$ 为i.i.d.样本,  $X_i \sim P(\lambda_i)$ ,  $Y_1, \dots, Y_m$ 为

18. (1

i.i.d.样本,  $Y_i \sim P(\lambda_i)$ . 求 $\theta = \lambda_1/\lambda_2$ 的渐近正态置信区间;

(2) 设 $X_1, \dots, X_n$ 为i.i.d.样本,  $X_i \sim N(\mu_1, \sigma_1^2)$ ;  $Y_1, \dots, Y_m$ 为i.i.d.样本,  $Y_i \sim N(\mu_2, \sigma_2^2)$ . 求 $\theta = \mu_1/\mu_2$ 的渐近正态置信区间;

(3) 某厂对职工出勤率进行调查, 甲车间随机抽取50人, 其中一年全勤的有40人; 乙车间随机抽取40人, 其中一年全勤的有35人. 求甲、乙两车间职工全勤率之差的区间估计(置信水平为0.95).

19. 设 $X_1, \dots, X_n$ 为i.i.d.样本,  $X_i \sim E(1/\theta) \sim r(1/\theta, 1)$ (指数分布).

(1) 证明: $T = \frac{X_1}{X_2}$ 是一渐近枢轴量(即其渐近分布与 $\theta$ 无关), 并用 $T$ 构造 $\theta$ 的一个水平为 $1-\alpha$ 的渐近置信区间;

$X = \theta X$

个水平为 $1-\alpha$ 的渐近置信区间;

(3) 求 $\theta$ 的基于似然比统计量、score统计量以及Wald统计量的水平为 $1-\alpha$ 的渐近置信区间.

20. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ,  $0 < \mu < 1$ . 求  $\mu$  的基

于似然比统计量、score 统计量以及 Wald 统计量的水平为  $1 - \alpha$  的渐近 置信区间.

21. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim V(\lambda, \sigma^2)$ . 求  $\lambda$  的基于似然比统计量、score 统计量以及 Wald 统计量的水平为  $1 - \alpha$  的渐近置信域.

22. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i$  服从均匀分布  $U(0, 1)$ , 应用定理 7.1.1 求  $\theta$  的水平为  $1 - \alpha$  的置信区间.

\*23. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ,  $0 < \mu < 1$ . 应用定理 7.1.1 证明:  $\theta$  的水平为  $1 - \alpha$  的置信区间可表示为其中

(2) 证明:  $t_2(x; \theta)$

是渐近枢轴量, 并用  $r_2$  构造  $\theta$  的一

T

②-

-----  
'  $r + (n - r + 1)F(2n - 2r + 2, 2r; 1 - \alpha/2)$  f  
? $(r + 1)F(2r + 2, 2n - 2r; 1 - \alpha/2)$  " $(n - T) + (T + 1)F(2r + 2, 2n - 2T; 1 - \alpha/2)$  \*  
2) \*

习题七

341

$r = 2$  x: (提示: 参见第一章,  $r$  的分布函数可由不完全  $\beta$  函数表示, 而  $\lambda = 1$

分布可由  $F$  分布表示). \*24. 设  $r$  服从负二项分布  $NB(\theta, r)$ ,  $0 < \theta < 1$ ,  $r$  已知. 应用定理

7.1.1 证明:  $\theta$  的水平为  $1 - \alpha$  的置信区间可表示为  $[g_j]$ , 其中

$n = \frac{r}{r + rF(2r, 2T; 1 - \alpha/2)}$   $r^{n + (T + 1)F(2r + 2, 2n - 2T; 1 - \alpha/2)}$   
 $rF(2r, 2T; 1 - \alpha/2)$   
 $r + rF(2r, 2T; 1 - \alpha/2)$  \*

25. 证明定理 7.1.3 的结论. \*26. 设  $f$  为一密度函数, 它在  $\mathbb{R}$  上非零, 在  $\mathbb{R}$  以外

$= 1 - \alpha$  的区间  $[a, b]$  中, 当取  $a = x_*$ , 使  $\int_{x_*}^b f(x) dx = 1 - \alpha$  的  $a, b$

时, 所得区间长度最短;

(2) 当  $f$  在  $[x_*, x_J]$  上严格递增时, 证明与 (1) 类似的结论; 并

用 (1) 或 (2) 来证明例 7.1.3 中所得的区间  $[X(n), a^*X(n)]$  在所有满足例 7.1.3 要求的区间中是最短的.

27. 设  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ . 试在  $M$  的水平为  $1 - \alpha$  的置信区间类  $[Y$

$(1) + e, X(1) + (/)]$  中, 找一个长度最短的置信区间.

\*28. (1) 设  $t/(X)$  是一个取正值的统计量. 是一个枢轴量. 其密  $U$

,

等于零

(1) 若  $f$  在  $[x_*, x_J]$  上严格递减, 证明: 在所有满足  $f: f(x) dx$

$_{-} < X + + 00$ .

度函数  $f$  在  $\mathbb{R}$  上是单峰的 (即当时,  $f$  不降; 当时,  $f(X)$  不增若). 考虑的如下置信区间

类:  $\int_{a^*}^b f(x) dx = 1 - \alpha$

(2) 考虑例 7.1.1, 即  $X_1, \dots, X_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ , 其中  $\sigma^2$  都未知; 试在  $\mathcal{C}$  的置信区间类:  $j [无 - 6SZ^{\wedge}, X - aS^{\wedge}]$ :

$\int_{a^*}^b f(x) dx = 1 - \alpha$  中找一个长度最短的置信区间, 其中  $S =$

$\int_{a^*}^b f(x) dx = 1 - \alpha$ , 且  $a^* < x_*$ . 名  $6^*$

$aI \gg$

则  $[T - b^*U, T - a^*U]$  是  $\mathcal{C}$  中长度最短的区间 (提示: 因  $\mathcal{C}$  中任一区间的长度都是  $(b - a)U$ , 所以

只需证若  $a < b$  且  $b - a < b - a^*$ , 则有

$\int_{a^*}^b f(x) dx < 1 - \alpha$ );

## 第七章 区间估计

$(n-1)^{-1} S^2$  ( $X \sim N(\mu, \sigma^2)$ ).

29. 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $X_i \sim N(\mu, \sigma^2)$ ,  $\mu$  和  $\sigma^2$  都未知. (1) 求的水平为  $1-\alpha$  的 UMAU 置信上、下限以及 UMAU 置信

区间;

(2) 证明: (1) 中的置信区间可由反解例 6.6.1 中关于  $\mu$  的似然比

检验的接受域而得到;

(3) 求  $\sigma^2$  的水平为  $1-\alpha$  的 UMAU 置信上、下限以及 UMAU 置信

区间. 30. 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $X_i \sim \text{Exp}(\theta)$ , 其中  $\theta > 0$  未知. 求  $\theta$  的水平为  $1-\alpha$  的 UMAU 置信区间.

31. 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $X_i \sim r(\lambda, 1)$  为寿命分布. 若仅

观测到  $(y_1, \dots, y_r)$  其中  $y_j = (x_1, \dots, x_r)$  这前  $r$  个寿命的值, 求  $\lambda$  的水平为  $1-\alpha$  的 UMA 置信下界和 UMAU 置信区间.

32. 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $X_i \sim \text{Exp}(\theta)$ . (1) 求  $\theta$  的

水平为  $1-\alpha$  的 UMA 置信下限; (2) 证明:  $[J(\theta), a' + X(\theta)]$  为  $\theta$  的水平为  $1-\alpha$  的 UMA 置信区间.

(提示: 见第六章 24 题)

33. 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $X_i \sim N(\mu, \sigma^2)$ , 求  $\mu$  的水平为

$(1-\alpha)$  的容忍上、下限和容忍区间. (1) 设  $\mu$  未知,  $\sigma^2 = B$  已知; (2)  $y = \mu/x$ . 已知,  $a > 0$

未知; (3) 设棉纱的断裂负荷服从正态分布,

从  $a$  未知. 现从一批棉纱中随机抽取 12 段测得其断裂负荷 (单位为  $100_N$ ) 为

228.6, 232.7, 238.8, 317.2, 315.8, 275.1, 222.2, 236.7, 224.7, 251.2, 210.4, 270.7. 设置信水平为  $(0.95, 0.95)$ , 已知

$A(12, 0.95, 0.95) = 2.74$ , 求棉纱断裂负荷的容忍下限.

34. 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $X_i$  服从均匀分布  $U(0, 3)$ ,  $\theta > 0$ , 求该分布水平为  $(1-\alpha, 1-\alpha)$  的容忍上、下限和容忍区间.

35. 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $X_i$  服从 Weibull 分布  $W(\lambda, a)$ ,

其中  $a > 0$  已知,  $\lambda > 0$  未知. 求该分布水平为  $(1-\alpha, 1-\alpha)$  的容忍上、下限和容忍区间.

36. 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $X_i \sim \text{BE}(\theta, 1)$  -

1. 求该分布水平为  $(1-\alpha, 1-\alpha)$  的容忍上、下限和容忍区间.

=加广 7 三

## 第八章 Bayes 统计基础

Bayes 统计, 特别是 Bayes 统计计算, 近年来取得重大进展, 是当今统计学发展最快的分支之一. 虽然 Bayes 统计推断曾经有很多争论, 但是现在这些争论已经不是主流, 大家更看重的是 Bayes 方法的成效和应用价值, 而把这种方法在哲学方面的问题先搁在一边. 近年来, 一般理论与应用工作者都承认, Bayes 统计方法可用于统计学的几乎所有分支, 而且通常效果都不错, 同时 Bayes 方法也能解决一些经典统计方法所不能解决的问题. 另外, Bayes 统计方法往往比较简单直接, 它把各种统计问题都归结为后验分布的计算问题, 很受应用工作者的欢迎. 因此, Bayes 统计已成为当今统计学的重要组成部分.

本章主要介绍 Bayes 统计推断的基础知识. 第 8.1 节首先介绍 Bayes 统计的基本概念和原理, 包括先验分布、后验分布以及 Bayes 统计推断方法, 本节也比较详细地介绍了无信息先验分布的选取方法; 第 8.2 节介绍 Bayes 估计方法及其性质; 第 8.3 节介绍假设检验和区间估计的 Bayes 方法, 特别, 本节比较详细地介绍了 HPD 可信域的基本性质和求解方法. 关于 Bayes 统计推断更进一步的内容可参见茆诗松 (1999), 郑忠国 (1998), 范金城,



吴可法(2001)，Lehmann (1985, 1998)，Berger(1985)，Shao(1998)等文献。

### 8.1 Bayes统计基本概念

且有

前，对于未知参数 $\theta$ 一般应当有所了解，即积累了一些关于参数 $\theta$ 的“先验信息”，对于未知参数的统计推断应该既考虑观察样本的信息，也考虑参数的先验信息。这种先验信息用数理统计的语言表述出来，就是未知参数 $\theta$ 为一随机变量，有一定的先验分布，其分布密度为 $K(\theta)$ 。这种观点在某些场合是合理的。例如，某厂通过观察样本 $I$ 了解一种产品在每天的次品率 $\theta$ ， $X$ 通常服从二项分布。就这一天而言， $\theta$ 为一确定的数，但是，由于工厂每天在生产，对次品率逐日波动的情况有所了解，

Bayes统计的基本观点就是要充分利用先验信息，并综合样本信息，然后进行Bayes统计推断。具体来讲，设观察样本为 $X =$

$R_n$ ,

Bayes观点认为，人们在观察到样本 $X$ 之

344

### 第八章Bayes统计基础

因此，在估计当天次品率 $\theta$ 时，适当地参考过去次品率的波动情况是十分合理的。从长期来看，次品率 $\theta$ 每天不一样，可以看成是一个随机变量，某一天的次品率则可看成随机变量的取值。按照Bayes观点，可以给次品率 $\theta$ 一个先验分布，从而综合样本分布得到其后验分布，并基于后验分布进行统计推断。但是，也有许多情况，把未知参数看成随机变量并不是很合理。例如，通过取样来估计某铁矿的含铁量 $\theta$ ，通过抽样调查来估计某人在选举中的得票率 $\theta$ 等，这时 $\theta$ 只能看作一个确定的数。因此，按照Bayes观点，一律把未知参数都看成随机变量，并赋予一定的先验分布，这种观点尚无充分论据。不过正如上面所述，Bayes方法的成效和应用价值已经没有什么争论；Bayes方法作为一种有力的工具已经得到充分的认可。另外，从统计判决的观点来看，损失函数同时在样本空间和参数空间取平均，得到Bayes风险（见第三章定义3.1.5），然后根据Bayes风险最小准则求解，这显然是合理的。

#### 8.1.1 Bayes统计原理

Bayes统计推断主要可归纳为以下三点：

- (1) 参数 $\theta$ 为随机变量，有先验分布，记为 $K(\theta)$ ，它集中了关于未知参数的先验信息。
- (2) 认为样本分布为参数已知时样本 $X$ 的条件分布，即 $p(x|\theta)$ ，它表示参数给定时，样本 $X$ 的分布规律。

- (3) Bayes统计推断的出发点为参数 $\theta$ 的后验分布，即 $X$ 已知时的条件分布 $P(\theta|X)$ ，因为它综合了先验分布 $K(\theta)$ 以及样本分布 $f(x, \theta)$ 所提供的关于参数 $\theta$ 的全部信息。后验分布是一切Bayes统计推断的出发点，因而计算后验分布就成为Bayes统计的主要任务，其分布密度记为 $\pi(\theta|X)$ 。

引理8.1.1设

后验分布，即 $X$ 已知时的条件分布 $\pi(\theta|X)$ 可表示为

$$\pi(\theta|X) = \frac{f(X|\theta)K(\theta)}{\int f(X|\theta)K(\theta)d\theta} \quad (8.1.1)$$

证明的联合分布可表示为 $(X, \theta) \sim p(x, \theta) = p(x)p(\theta|x)$ 。

根据Bayes观点，其中 $p(\theta) = \pi(\theta) \int p(x|\theta)dx = f(x, \theta)$ ， $p(\theta|x) = \pi(\theta|x)$ ，因此有

$$\pi(\theta|x) = \frac{p(\theta)p(x|\theta)}{\int p(\theta)p(x|\theta)d\theta} = \frac{p(\theta)f(x, \theta)}{\int p(\theta)f(x, \theta)d\theta}$$

,

0, 0的先验分布为 $\pi_0(\theta)$ , 则 $\pi_0(\theta)$ 的

## 8.1 Bayes统计基本概念

345

边缘分布 $p(x)$ 可表示为

$$p(x) = \int \pi_0(\theta) f(x|\theta) d\theta.$$

由此即可得到(8.1.1)式, 其中 $c(x) = p(x)$ .

注 (8.1.1)式有很丰富的含义, 今摘要说明如下:

(1) 首先, 在(8.1.1)式中,  $c(x) = p(x)$ 可看成积分常数, 有时

不需要进行计算, 而可以由以下关系式决定:  $\int c(x) f(x|\theta) d\theta = 1$ ,

可参见下面的例题. 在本章许多场合都会出现代表积分常数的 $C(\theta)$ 或 $C$ , 它们在不同场合可能代表不同的值, 但意义大体相同, 为了简单起见, 我们不再仔细地加以区分. 因此在很多问题中, (8.1.1)式也常常可以略去常数而表示为

$\pi_0(\theta|x) \propto \pi_0(\theta) f(x|\theta)$ . (8.1.2) (2) 由(8.1.1)式可知, 先验密度的选取亦可相差一个常数, 即 $\pi_0(\theta)$ 亦可取为广义密度 $k\pi_0(\theta)$ 不变, 因为在

(8.1.1)式的分子分母中常数 $k$ 可消去. 因此, 先验密度 $\pi_0(\theta)$ 亦可表示为诸如

$\pi_0(\theta)$ 或者等形式. 另外, 若为 $\theta$ 上的

非负可积函数, 即 $\int \pi_0(\theta) d\theta < \infty$ , 并在(8.1.1)式中用 $w(\theta)$ 取代先验分布 $\pi_0(\theta)$ , 则

$\pi_0(\theta|x)$ 仍然为一个后验密度函数, 因为 $\pi_0(\theta)$ 可表示为 $k\pi_0(\theta)$ ;  $=k\pi_0(\theta)$ . 更进一步, 若 $\int \pi_0(\theta) d\theta = \infty$ , 但

$\int \pi_0(\theta) d\theta = 1$

是1, 则(8.1.1)式仍然有意义, 仍然为密度函

数. 对于这些情况,  $\pi_0(\theta)$ 通常称为广义先验密度; 当 $\int \pi_0(\theta) d\theta = \infty$

时, 则称为非正常的(improper)先验密度. 而 $\int \pi_0(\theta) d\theta = 1$ , 则称为

正常的(proper)先验密度.

(3) 如上所述, 在推导后验分布时, 可以相差一个常数, 因此若 $X$ 的分布密度 $f(x)$ 可表示为 $f(x) = kg(x)$ ,  $A$ 为常数, 则称 $g(x)$ 为 $Z$ 分布的核. 这个名字在Bayes统计中常常用到, 可参见下面的例子. ■

例8.1.1 设 $X_1, \dots, X_n$ 为独立同分布样本, 且 $X_i \sim \text{Bernoulli}(p)$ , 求的后验

密度假设 $p$ 的先验分布分别为 $\pi(p) \propto \exp(-np)$ .  $n$

解 样本分布为 $f(x) = \prod_{i=1}^n \pi(x_i|p) = p^{\sum x_i} (1-p)^{n-\sum x_i}$ , 其中

$I$

346

## 第八章 Bayes统计基础

$T = \sum_{i=1}^n X_i$ . 易见, 若把 $\theta$ 看成变量, 则上式为冷分布 $BE(T+1, n-T+1) \propto$  的核.

i)  $\pi_0(p) = 1$  代入(8.1.1)式可得

$$\pi_0(p|x) = C(x) \cdot 1 \cdot \pi_0(p) (1-p)^{n-T}.$$

$$I \quad c(1) = \frac{1}{\int_0^1 (1-p)^{n-T} dp} = \frac{1}{\frac{1}{n-T+1} (1-p)^{n-T+1} \Big|_0^1} = (n-T+1)$$

因而上式为分布, 即 $\pi_0(p|x) \sim \text{Beta}(T+1, n-T+1)$ ,

由于随机变量为 $n-T+1$ , 这时 $c(p)$ 可由该分布的常数决定, 即

$$c(p) = \frac{1}{\int_0^1 (1-p)^{n-T} dp} = (n-T+1)$$

ii)  $\pi_0(p) = c \exp(-np)$  代入(8.1.1)式可得

$$\pi_0(p|x) \propto c \exp(-np) (1-p)^{n-T} = c \exp(-np) (1-p)^{n-T}.$$

$$\pi_0(p|x) = c(x) \exp(-np) (1-p)^{n-T}.$$

因此有  $e^X \sim BE(T+p, n-T+q)$ . 鷹

由前面各章的内容可知, 充分统计量在统计推断中起着十分重要的作用, 以下引理说明, Bayes准则与充分性准则是相容的. 这个引理也是由(8.1.1)式得到的.

引理8.1.2 设  $X \sim f(X|\theta)$ ,  $\theta$  的先验分布为  $\pi(\theta)$ . 若  $T(X)$  为充分统计量, 其分布为  $g(t, \theta)$ , 则  $\theta$  的后验分布  $\pi(\theta|X)$  可表示为

$$\pi(\theta|X) = \frac{g(T(X)|\theta)\pi(\theta)}{\int g(T(X)|\theta)\pi(\theta)d\theta}$$

(8.1.3) 反之, 若  $\theta$  的后验分布  $\pi(\theta|X)$  可表示为  $\theta$  和某个  $\pi(\theta)$  的函数:

$$\pi(\theta|X) = \pi(\theta)h(X), \text{ 则 } T = T(X) \text{ 为充分统计量.}$$

证明 若  $T=T(X)$  为充分统计量, 则由因子分解定理可知  $f(x|\theta) =$

$$g(T(x)|\theta)h(x), \text{ 且有}$$

$$\pi(\theta|X) = \pi(\theta)h(X), \text{ 的 } h(X)$$

$$\int \pi(\theta)g(T(x)|\theta)h(x)d\theta = \pi(\theta)g(T(x)|\theta)$$

$= C(X)\pi(\theta)g(T(x)|\theta)$ . 此即(8.1.3)式, 其中  $C'(X) = \int \pi(\theta)g(T(x)|\theta)d\theta$  (10) 反之, 若  $\theta$  的后验分

布可表示为  $\pi(\theta|X) = \pi(\theta)h(X)$ , 则由(8.1.1)式可得  $\pi(\theta|X) = \pi(\theta)h(X)$  (权,  $T(x)$ )  $= c(x)\pi(\theta)f(x|\theta)$ . 因此有

$$\pi(\theta|X) = [c'(X)\pi(\theta)] [f(x|\theta)] h(x)g(T(x)|\theta).$$

## 8.1 Bayes统计基本概念

347

由因子分解定理可知,  $T = T(X)$  为充分统计量.

(8.1.3)式有广泛的应用, 该式说明, 我们也可以从充分统计量

$r = r(x)$  的分布  $g(r|\theta)$  (或者它的核) 出发, 根据(8.1.3)式计算后验分布, 这往往比从样本分布出发简单, 可参见下面的例子. 另外,

引理8.1.1后面的注释对于(8.1.3)式亦适用., 例8.1.2 ..., 氦为独立同分布样本

为正态分

布,  $a, \sigma^2$  已知, 求  $\mu$  的后验分布, 假设  $\mu$  的先验分布分别为 i)  $\mu$  服从广义均匀分布, 即  $\pi(\mu) \propto 1$ ; ii)  $\pi(\mu) \sim N(\mu_0, \sigma_0^2)$  为正态分布, 其中

$\mu_0, \sigma_0^2$  为已知.  $x$

$$i) \pi(\mu) = 1 \text{ 代入(8.1.3)式可得 } \pi(\mu|x) = c(x) \cdot \exp|$$

因此  $\pi(\mu|x) \sim N(x, a^2/n)$  为正态分布, 其中  $c(x) = 1/n$ .

ii) 下面证明: 若  $\pi(\mu) \sim N(\mu_0, \sigma_0^2)$ , 则  $\pi(\mu|x)$  也服从一个正态分布  $N(a, \sigma^2/n)$ , 其中

$\sigma^2 = \sigma_0^2 + \sigma^2/n$  为充分统计量,

解 当  $\sigma^2$  已知时, 因此可以基于无的分布, 根据(8.1.3)式求后验分布.

~

:

n-

~x+

aa0 21 1

1

$$a = \frac{\mu_0 + \frac{\sigma^2}{\sigma_0^2} x}{1 + \frac{\sigma^2}{\sigma_0^2}} \quad (8.1.4) \quad n1 \quad p n1$$

$$2 + 2 \quad 2 + 2 \quad a < r0 \quad a < r0$$

由于  $\pi(\mu) \sim N(\mu_0, \sigma_0^2)$ ,  $XV(x, z) \sim N(\mu, (\sigma^2/n))$ , 因此有

对上式的主要部分进行配方和化简可得

■

$$T) CTq + a / TI$$

348 因此联合分布、边缘分布和后验分布分别为

第八章 Bayes统计基础

$P(X)$

$I h(x私 J -Q0$

$(Mo-x)^2 1 (fTo +a^2/n)J$

(8. 1.6)

$=\exp\{-^{(/x_a)^2H J (浩:知}\}$

$7T(/xIx) 尸(\%)$

=

$_a)^2|- (8.1.7)$

$v^/XP卜了$

■  $1\_exP($

$i 2v j$

以上后验分布有很好的统计意义，今说明如下。在正态分布 $yV(M, a^2)$ 中，方差 $a^2$ 的倒数

$a^2$ 称为精度， $a$  越小，精度越高。在本题 中，参数 $A$  的先验均值为 $A_0$ ，先验精度为

$cr; 2$ ，而样本均值 $\bar{y} \sim N(\bar{y}, a^2/n)$ ，其精度为 $(a^2/n)^{-1}=n/a^2$ ，因此(8.1.4)式说明，后

验 分布的精度为先验分布的精度与样本均值分布的精度之和；而后验 分布的均值为先验均值

$P$  与样本均值 $\bar{y}$  关于精度的加权平均。 ■

### 8.1.2 先验分布的选取方法

选取先验分布的方法多种多样，也有不少争论，但不管哪种选取方 法，最后还要看Bayes统计推断的效果，以下介绍两类最常用的选取方 法，即共轭先验分布和无信息先验分布。

(1)共轭先验分布 即取先验分布 $\pi(\theta)$ ，使得后验分布 $\pi(\theta|X)$ 与  $\pi(\theta)$ 属于同一类分布族。直观上看，该准则使先验分布与后验分布 保持协调一致，这也是合理的。例如 $\pi(\theta)$ ， $\pi(\theta|X)$ 同为正态分布， 或同为分布等等。前面的两个例子都是这种情形。共轭先验分布在 实际中应用很广，下面的例子都是常见的共轭先验分布。

例8.1.3 及...，及为独立同分布样本，且有 $X_i \sim P(A)$ ，设 $A$ 的先验分布为 $\pi(A) \sim r(a, p)$ ，则其后验分布也是 $\Gamma$ 分布。

解...，总的联合分布可表示为

$f(x, A) = e^{-A} A^c(x) e^{-nAAr}$ ,  $T = \sum_{i=1}^n X_i$

易见，若把 $A$ 看成变量，则上式包含了 $\Gamma$ 分布的核，而 $A \sim \pi(A) = e^{-A} A^{p-1}$ ，因此由

(8.1.1)式有

$\pi(A|x) = c(x) \pi(A) / f(x, A) = e^{-A} A^{p-1} / e^{-nAAr} = e^{-A(n+r)} A^{p-1}$ ，即

$\pi(A|x) \sim r(n+a, T+p)$ ，与 $\pi(A)$ 同属于 $\Gamma$ 分布族。 厘

例8.1.4  $X_1, \dots, X_n$ 为独立同分布样本，且有 $X_i \sim E(A) = r(A, 1)$ ；设  $A$ 的先验分布为 $\pi(A) \sim r(a, p)$ ，则其后验分布也是 $\Gamma$ 分布。

### 8.1 Bayes统计基本概念

349

解  $X_1, \dots, X_n$ 的联合分布可表示为

$y(\%, A) = A^p e^{-Ar}$ ,  $T = \sum_{i=1}^n X_i$

因此由(8. 1. 1)式有

$\pi(A|x) = C(\%) e^{-f(A)p} A^{p-1} = c(x) A^{n+p-1} e^{-(n+a)A}$

即 $\pi(A|x) \sim r(a+r, T+p)$ ，与 $\pi(A)$ 同属 $\Gamma$ 分布族。由以上两个例子都可以看出，在样本分布 $f(x, A)$ 中，若把 $A$ 看成

变量，则它们都包含了 $\Gamma$ 分布的核(例如在例8.1.3中为 $r(n, r + i)$ )，而且两个 $\Gamma$ 分布的核相乘还是 $\Gamma$ 分布的核，因此可以取 $\Gamma$ 分布为共轭 先验分布。这种情况很常见，例如在例

8.1.1中, 若把  $\theta$  看成变量,

包含了冷分布的核, 而且两个冷分布的核相乘还是  $\theta$  分布的核, 因此也可以取  $\theta$  分布为共轭先验分布.

以下介绍逆  $r$  分布及其共轭先验分布. 设  $x \sim r(A, p)$ , 则称  $y = x^{-1}$  服从逆  $r$  分布, 并记为  $r^{-*}(A, p)$ ,

根据变换公式,  $r$  的分布密度为

而  $e^{-y}$  (7) 注意, 若  $f$  为  $r$  分布  $r^{-*}(A, p)$ , 则其数学期望为  $E(r) = E(a^{-1}) =$

在例 8.1.4 中, 若  $X_i \sim E(a^{-1}) = r^{-*}(a^{-1}, 1)$ . 容易验证, 若  $a$  的先验分布为逆  $r$  分布: 即  $a \sim r^{-*}(A, p)$  则其后验分布也服从逆  $r$  分布.

例 8.1.5 设  $X_1, \dots, X_n$  为独立同分布样本, 且有  $X_i \sim N(0, \sigma^2)$ ,

$a = \sigma^2$  服从逆  $r$  分布 其后验分布也是逆  $r$  分布.

解样本分布为

$x) = c(\theta) e^{-\frac{1}{2\theta} \sum x_i^2}$

即  $T(a | X) \sim r^{-1}(A + 5, y + p)$ , 与  $r(a)$  同属逆  $r$  分布族, 因此逆  $r$

其中  $S =$

含了逆  $r$  分布  $r^{-1}(S, f^{-1})$  的核. 因此由 (8.1.1) 式以及 (8.1.8) 式有

$\sum_{i=1}^n$

$I \{y^0\}$ . (8.1.8)

即其精度  $a^{-2}$  服从  $r$  分布), 则

匪

易见, 若把  $\theta$  看成变量, 由 (8.1.8) 可知,  $f(\theta, a)$  包

350 第八章 Bayes 统计基础

分布为参数  $a = a$  的共轭先验分布. 其后验期望为  $E(a | \theta) = (A + S) / (y + p - 1) \cdot$

例 8.1.5 和例 8.1.2 合在一起, 可推广到  $4 - \sqrt{V(\theta, a^2)}$  的一般情

形, 若  $a \sim a^2$  服从逆  $r$  分布, 而  $\theta \sim N(0, \sigma^2)$  服从正态分布, 则  $(a, \theta)$  的后验分布也是逆  $r$ -正态分布 (见习题).

例 8.1.6 设  $X_1, \dots, X_n$  为独立同分布样本, 且有  $X_i \sim P(0, \theta)$ , 设  $\theta$  的先验分布为

Pareto 分布  $PR(a, \theta)$ ,  $BPit(\theta) = a\theta^a e^{-(a+1)\theta}$   $I$

$\theta \in I$ , 则其后验分布也是 Pareto 分布.

解  $\pi(\theta | x) = \pi(\theta) \prod_{i=1}^n f(x_i | \theta) \propto \theta^{-a} e^{-(a+1)\theta} \prod_{i=1}^n \theta^{-1} e^{-\theta x_i}$ , 因此有

$77(\theta | x) = c(\theta) \theta^{-(a+1)} e^{-(a+1)\theta} \prod_{i=1}^n \theta^{-1} e^{-\theta x_i} = c(\theta) \theta^{-(n+a+1)} e^{-\theta \sum_{i=1}^n x_i - (a+1)\theta}$ ,  
u

其中  $\theta > 0 = \max \{x_i\}$  因此

例 8.1.7 设  $X_1, \dots, X_n$  为独立同分布样本, 且有  $X_i \sim PR(a, \theta)$  的,

若  $\theta > 0$  已知,  $a > 0$  的先验分布为  $77(a) \sim r(a, p)$ , 则其后验分布也是  $r$  分布.

解 由于  $f(x_i | \theta) = a\theta^a [1 - e^{-\theta x_i}]^{a-1} e^{-\theta x_i}$ ;  $X_i \sim r(a, p)$ , 因此由 (8.1.1) 式有

$77(a | x) = e(X) e^{-aap} \cdot a^n \prod_{i=1}^n [1 - e^{-\theta x_i}]^{a-1} e^{-\theta x_i} \propto \theta^{an+p-1} e^{-a\theta \sum_{i=1}^n x_i - a\theta}$

即  $77(a | x) \sim r(r + a, n + p)$ , 与  $77(r)$  同属  $r$  分布族, 其中  $T = \sum_{i=1}^n x_i$

$2 \cdot \sum_{i=1}^n x_i$ .  $I \theta > 0$ .

以上两个例子的结果可以推广到幂函数分布  $x \sim pf(c, \theta)$ : 即

$f(x; c, \theta) = -; I(c > 0; \theta > 0)$ .  $\theta$

可以证明: 若  $c$  已知, 则  $\theta$  的共轭先验分布为 Pareto 分布; 若  $\theta$  已知, 则  $c$  的共轭先验分布为  $r$  分布 (见本章习题).

例8.1.8 设  $X = (J_1, \dots, J_n)$  服从指数族分布:  $\pi(\theta) =$  为共轭先验分布.

也服从 Pareto 分布  $PR(n+a, \theta^0)$ . I

### 8.1 Bayes统计基本概念

351

ii)把以上结果用于两点分布,即令  $\dots$  人为样本,  $\theta \sim \text{Bernoulli}(\theta)$ . 解 i)由(8.1.1)式有

$$\pi(\theta | X) = C(X) \pi(\theta) / \pi(X)$$

$$= c(\theta) \exp \left\{ \sum_{j=1}^n (a + T(\theta)) \theta^j - (n+1) \theta \right\} / \pi(\theta). \quad (8.1.9)$$

因此后验分布也具有形式  $\pi(\theta | X) \propto \theta^{a+T(X)} (1-\theta)^{n+1-T(X)}$ , 其中  $a = \theta^0$ ,  $n = n+1$ . 因此为共轭先验分布.

n

h)样本分布为指数族分布  $f(x, \theta) = \prod_{i=1}^n f(x_i, \theta)$

设  $r(x) =$

$$(\theta) = \log \left[ \theta^a (1-\theta)^{n+1} \right], \quad b(\theta) =$$

$i=1$

$-n \log(1-\theta)$  因此可取共轭先验分布为

$$\pi(\theta) = c(a, n) \exp \{ a \log \theta + (n+1) \log(1-\theta) \} = c(a, n) \theta^a (1-\theta)^{n+1}$$

由此式可得  $\pi(\theta | X) \propto \theta^{a+T(X)} (1-\theta)^{n+1-T(X)}$

由此式可得  $\pi(\theta | X) \propto \theta^{a+T(X)} (1-\theta)^{n+1-T(X)}$ . 即  $\pi(\theta)$  服从冷分布 (应当要求  $a > -1$ ,  $n > -1$ ), 由例8.1.1可知, 后验分布也服从冷分布. 另外, 由(8.1.9)也容易看出, 后验分布可表示为

$$\pi(\theta | X) \propto \exp \{ (a+T(X)) \log \theta + (n+1) \log(1-\theta) \}.$$

因此  $\pi(\theta | X)$  服从分布  $BE(p, q)$ , 其中  $p = T(X) + a + 1$ ,  $q = (n+1) - T(X)$

$-a$  (应当要求  $p > 0, q > 0$ ). 另外, 本例的结果亦可用于 Poisson, 正态等分布; 同时, 本例的结果亦可到  $\theta$  和  $T(a)$  为向量的情形 (见习题). 进一步的讨论可参见 Lehmann and Casella (1998).

表8.1.1总结了常见的共轭先验分布.

表8.1.1常见的共轭先验分布 总体分布参数

二项分布 参数  $\theta$  负二项分布 参数  $e$

共轭先验分布 冷分布

$\gamma$  分布  $\Gamma$  分布  $\rho$  分布 逆  $\Gamma$  分布 正态分布 逆  $\chi^2$  分布

Pareto分布  $\Gamma$  分布

Poisson 分布 指数分布

指数分布 正态分布 (方差已知) 正态分布 (均值已知) 均匀分布 Pareto分布 ( $\theta$  已知)

均值  $A$  均值倒数  $A$  均值  $\theta = A^{-1}$

$$a = a_0$$

a

352

## 第八章 Bayes统计基础

(2) 无信息先验分布 如果对  $\theta$  的先验情况知之甚少, 没有什么先验信息可以利用, 一个很直观的假定就是同等无知原则, 即认为  $\theta$  的取值机会均等, 各向同性; 用随机变量的语言来表述就是  $\theta$  在某个区间或全空间服从均匀分布, 即  $\pi(\theta) = c$  为常数. 一般可取  $\pi(\theta) \propto 1$ , 这时  $\pi(\theta)$  可视为广义密度 (见引理8.1.1后面的注释), 这在前面例8.1.2中已经用过. 同等无知原则是 Bayes 首先提出来的, 因此也称为 Bayes 假设.

但是, 无信息先验分布的内涵比同等无知原则要广泛很多, 它又称为非主观先验分布. 因为尽管我们在主观上对  $\theta$  的先验信息知之甚少, 但是, 在  $\pi(\theta) = f(\theta) \pi(\theta)$  中,  $f(\theta)$  作

为样本 $X$ 的分布; $\theta$ 作为一个随机变量,它们应该与概率统计中的规律协调一致.由此也可推出若干先验分布的选取原则和方法,下面以位置尺度参数分布族为例予以说明,请读者注意,下面要用到第四章的有关符号及公式.

i) 设 $X=(X_1, \dots, X_n)$ 服从位置参数分布 $\pi(y-\theta/1)$ ,其中 $\theta = \theta_0 + C$ ,  $V_C$ .今假设 $\theta$ 的先验分布为 $\pi(\theta)$ ,以下根据位置参数分布族的特性导出 $\pi(\theta)$ 应当满足的条件.设 $\theta$ 的先验分布为 $\pi(\theta)$ ,一方面,作为位置参数, $\theta$ 和 $\theta_0$ 应有相同的先验分布,即 $\pi(\theta) = \pi(\theta_0)$ ;另一方面,作为随机变量, $\theta$ 的分布 $\pi(\theta)$ 与 $\theta_0$ 的分布 $\pi(\theta_0)$ 应满足关系: $\pi(\theta_0) = |d\theta/d\theta_0| \cdot \pi(\theta)$ ;因为 $\theta = \theta_0 + C$ ,所以由 $d\theta/d\theta_0 = 1$ 可得 $\pi(\theta) = \pi(\theta_0 + C)$ ,  $V_C$ .因此综合

合这两个方面的结果可得 $\pi(\theta) = \pi(\theta_0 + C)$ ,  $M_C$ .上式对 $C = -\infty$ 也成立,所以 $\pi(\theta) = \pi(\theta_0) = \text{常数}$ ,这表明:位置参数分布族的无信息先验分布为 $\pi(\theta) \propto 1$ ;这与前述Bayes假设一致.

ii) 设 $Y=(Y_1, \dots, Y_n)$ 服从尺度参数分布 $\pi(y/a)^k$ 也服从尺度参数分布 $\pi(y/a)^k$ ,其中 $V=ka$ ,  $V_A > 0$ .今假设 $a$ 的先验分布为 $\pi(a)$ ,以下根据尺度参数分布族的特性导出 $\pi(a)$ 应当满足的条件.设 $a$ 的先验分布为 $\pi(a)$ ,一方面,作为尺度参数, $a$ 和 $a_0$ 应有相同的先验分布,即 $\pi(a) = \pi(a_0)$ ;另一方面,作为随机变量, $a$ 的分布 $\pi(a)$ 与 $a_0$ 的分布 $\pi(a_0)$ 应满足关系: $\pi(a) = |da/da_0| \cdot \pi(a_0)$ ;因为 $a = ka_0$ ,所以由 $da/da_0 = k$ 可得 $\pi(a) = k\pi(a_0)$ ,  $V_A > 0$ .因此综合

这两个方面的结果可得 $\pi(a) = \pi(a_0)$ ,  $V_A > 0$ .该式对 $a_0$ 也成立,所以 $\pi(a)$ 这表明:尺度参数分布族的无信息先验分布为 $\pi(a) \propto 1/a$ 这也是非正常的先验分布.

iii) 设 $X = (X_1, \dots, X_n)$ 服从位置尺度参数分布

则 $Y^k X^m$ 也服从位置尺度参数分布 $\pi(a) \cdot \pi(y - \theta/1) \pi(a)$ (见第四章),

## 8.1 Bayes统计基本概念

353

其中 $\theta = \theta_0 + C$ ;  $V_C$ ,  $k > 0$ ,并记 $\theta = (\theta_0, C)$ .今假设 $(\theta, a)$ 的先验分布为 $\pi(\theta, a)$ ,以下根据位置尺度参数分布族的特性导出 $\pi(\theta, a)$ 应当满足的条件.设 $\theta = (\theta_0, C)$ 的先验分布为 $\pi(\theta_0, C)$ ,一方面,作为位置尺度参数, $\theta$ 和 $\theta_0$ 应有相同的先验分布,即 $\pi(\theta) = \pi(\theta_0)$ ;另一方面,作为随机变量, $\theta$ 的分布 $\pi(\theta)$ 与 $\theta_0$ 的分布 $\pi(\theta_0)$ 应满足关系: $\pi(\theta) = |d\theta/d\theta_0| \cdot \pi(\theta_0)$ ,而由 $d\theta/d\theta_0 = 1$ 此综合这两个方面的结果可得 $\pi(\theta) = \pi(\theta_0)$ ,  $V_C > 0$ .

上式对 $k = a^{-1}$ ,  $m = -fx/a$ 也成立,所以 $\pi(\theta, a) = \pi(\theta, 1)$ ,这表明:位置尺度参数分布族的无信息先验分布为 $\pi(\theta, a) \propto 1/a^2$ ,这也是非正常的先验分布.

另外,在上述位置尺度参数分布族中,若 $a$ 已知,则它相当于位置参数分布,可取 $\pi(\theta) \propto 1$ ;若 $\theta$ 已知,则它相当于尺度参数分布,可取 $\pi(a) \propto 1/a$ ;若 $\theta$ 和 $a$ 都未知,也可结合前面i), ii)的情形,取

$\pi(\theta, a)$

$\pi(\theta, a) = k_1 \pi(\theta) \pi(a)$

$\pi(\theta, a) = k_2 \pi(\theta) \pi(a)$

.

因

$\pi(\theta, a) = \pi(\theta) \pi(a)$ ,而 $\pi(\theta) \propto 1$ ;  $\pi(a) \propto 1/a$

位置尺度参数分布族的无信息先验分布,求出相应的后验分布:i)  $\theta = 1$ ; ii)  $\theta = 0$ ; iii)  $\theta$ 和 $a$ 都未知.

解 i) 当  $a = 1$  时,  $X, \sim^+ r(i, i)$  服从位置参数分布, 可取  $\pi(M) \propto 1$ , 且有  

$$p_{\pi}(X, M) = e^{-n} \int_0^{\infty} \lambda^n x^{n-1} e^{-\lambda x} d\lambda = \frac{1}{\Gamma(n)} x^{n-1} e^{-x}$$
  
 直接计算可得,  $c(X) = e^{-x}$ .  
 ii) 当  $M=0$  时,  $X$  服从尺度参数分布, 可取  $\pi(a) \propto a^{-1}$ , 且有  

$$p_{\pi}(x, a) = \frac{1}{a} e^{-x/a} = \frac{1}{a} e^{-r/a} \quad T = Y/x, \quad 0 < T < \infty$$
  
 因此  $\pi(a | T)$  服从逆  $r$  分布  $r_{-1}(r, n)$ .  
 iii) 和  $\gamma$  都未知时,  $X_j + r(1, q)$  服从位置尺度参数分布, 可

因此有  

$$\pi(a) \propto a^{-1}$$

例 8.1.9 设  $x_1, \dots, x_n$  为独立同分布样本,

354

## 第八章 Bayes 统计基础

取  $T(0) \sim a^{-2}$ . 这时充分统计量为  $T = (S, X(1))$ , 其中  $S = \sum_{i=1}^n x_i$ ,  $X(1) = \min\{x_1, \dots, x_n\}$ , 且二者相互独立. 根据充分性原则, 由 (8.1.3) 式可得,  $\pi(0 | x) \propto$   
 $r^{n-1} g(5, x(1); \gamma)$ , 其中

$I(1) \propto r$

$g(5, x(1); \gamma) = c \cdot \frac{1}{r} e^{-x(1)/r} \cdot e^{-\gamma x(1)}$

因此,  $(M, \gamma)$  的后验联合分布可表示为  $\pi(M, \gamma | x) = c(\gamma) \frac{1}{r} e^{-s/r} e^{-\gamma x(1)}$

$\sim (x(1))^{-1}$   
 (8.1.10) 直接积分可得,  $c(x) = n! / r^n$ . 根据上式可求出  $M$  和  $a$  的边缘分布

如下:

$2n$

$\pi(T | x) = \int_0^{\infty} \pi(T, \gamma | x) d\gamma = \frac{1}{\Gamma(n)} \frac{1}{T^n} e^{-x(1)/T}$   
 从  $X(1) \sim S + n(x(1))^{-1/I}$  ]  $n+$

$\pi(r |$

$\pi(r |$

$\gamma) = \frac{1}{\Gamma(n)} \frac{1}{r^n} e^{-s/r} e^{-\gamma x(1)}$ , =

$n \cdot$  各

$-(n+1) \cdot s / (r \Gamma(n+1))$

$e^{-x(1)/r} \pi(r |$

其中  $X$  的后验分布比较复杂;  $a$  的后验分布为逆  $r$  分布  $r_{-1}(S, n)$ . 另外,  $M$  的变化范围也可限制在  $a > 0$  (如可靠性等问题), 可得到类似的结果.  $I$

以下 Jeffreys 准则可认为是上述位置、尺度参数分布族无信息先验分布的进一步推广.

(3) Jeffreys 准则 Jeffreys 准则也是一种无信息先验分布, 该准则认为, 参数  $\theta$  作为随机变量, 其分布  $\pi(\theta)$  在变换  $V = V(\theta)$  下,  $\theta$  和  $V$  的分布密度应满足随机变量的变换公式, 即

$\pi(V) = \pi(\theta) |J|$  (8.1.11)

直接验证可以证明, 若取  $\pi(\theta) = c |J|^{-1/2}$ , 则上述变换公式成立, 其

中  $|J|$  为样本关于参数的 Fisher 信息阵的行列式. 这可由 Fisher 信息阵的变换公式 (见第二章 (2.3.2) 式) 得到



$x(1) = 1 -$

$1/(8)I = \dots$ , ("没))I. 因此有  $|7(0) - 7(77(0))|$  肉却WT|, 即  $7T(9) = (4) +$  满足初2 de1

## 8.1 Bayes统计基本概念

355

(8.1.11)式. 注意, 根据Jeffreys准则得到的先验分布可能是广义密度

甚至是非正常的先验密度. 经直接计算可以证明: 对于位置参数分布族  $f(x - \theta)$ ,

$I(\theta) = 1/\sigma^2$ , 因此  $77(1)$ ; 对于尺度参数分布  $a^{-1}f(x/a)$

$1/(cr) \log cr - 2$ , 因此  $(\hat{\theta}) = -1$

,  $1/(\hat{\theta}) = -4$ , 因此  $n(\theta) = -2$  (见习题).

所以, 根据Jeffreys准则得到的先验分布与前面得到的无信息先验分

布是一致的.

例8. 1. 10 设  $X_1, \dots, X_n$  为独立同分布样本, 且有  $X_i$ ,

求Jeffreys准则下的先验分布: i)  $\theta < r$  已知; ii)  $\theta/x$  已知; iii)  $cr$  和  $ic$  都未知. 解由第二章Fisher信息阵的定义与例题可得, 样本关于参数

(7)的Fisher信息阵为  $\dots$   $n/a^2 \theta \dots$

$1/(\hat{\theta}) = 2 \cdot \dots 2n/a^2$

当  $\theta$  已知时,  $1/(\hat{\theta}) = n/\theta^2$  与参数  $\theta$  无关, 因此  $77(\hat{\theta}) = 1$ ; 当  $\theta$  已知 时  $1/(\hat{\theta}) =$

$2n/a^2$ , 因此  $77((r) \wedge 1/a$ ; 当  $(m, \dots)$  都看成未知参数时,

$-2n/a^2$  因此  $7r(\hat{\theta}) = 1/\theta - 2$

$1/(\hat{\theta}, \theta) =$

中都是常用的, 不过都是非正常的广义先验密度. ■

例8. 1. 11 设  $X_1, \dots, X_n$  为独立同分布样本, 求Jeffreys准则下的 先验分布, 若

i)  $J, -6(1, 6)$ ; ii)  $X, \sim P(A)$ ;  $\sim E(1/(r))$ .

解 i) 由第二章公式知  $1/(\hat{\theta}) =$  此  $77(\hat{\theta}) = \dots (1-\theta) - \dots 5f(\pm, \pm)$ , 这也是先验共轭分布, 由例8. 1.1可知, 其

; 对于位置尺度参数分布族

. 这些先验分布在实际问题

后验分布为  $TT(\theta|x) = BE(7+y,$

$t =$ .

ii) 由第二章公式知  $1/(\hat{A}) = n/A$ , 因此  $\hat{A} = A^*$ , 根据与例 8.1.3类似的计算可知, 其后验分布为  $77(A|x) = c(x)e^{-nAArA^*}$ , 因此  $vr(a|x) \sim r(a, r + \pm)$ ,  $r = x \dots$

Hi) 样本分布为  $f(x|a) = a^{-n}e^{-T/a}$  其中  $T = \sum_{i=1}^n X_i$ . 直

因此  $7r(a|x)$  幻为逆厂分布  $f^*(7>)$ .

接计算可得  $1/(\hat{a}) = n/a^2 - c(x)a^{-n}e^{-T/a} \sim i \dots cca'u + "e$

i 因此所以后验分布为  $7r(a|x) =$

356

## 第八章Bayes统计基础

### 8. 2 Bayes 估计

在 Bayes统计中, 比较有争议的问题是如何理解参数  $\theta$  为随机变量, 以及如何决定其先验分布. 但是从统计判决观点来求Bayes解并无不妥之处, 在数学上是合理的. 事实上, 在第三章第3.1节, 我们已经介绍过Bayes风险的概念. 以下首先介绍Bayes风险与Bayes后验风险, 及其基本性质; 然后在此基础上介绍两种常用的Bayes估计, 即基于均方损失的后验期望估计以及基于0-1损失的后验极大似然估计; 最后 简要介绍Bayes估计的一些性质.

#### 8.2.1 Bayes 风险

关于统计判决函数及损失函数、风险函数等概念，可参见第三章第 3.1 节，以下摘要回顾一些基本概念及符号。样本空间可表示为  $\Omega$  (I,  $\mathcal{F}$ )，其中  $\mathcal{P}$  带测度的参数空间可表示为  $\Theta$  (0,  $\mathcal{G}$ ,  $\mu$ )，其中  $d\pi(\theta) = \pi(\theta)d\theta$ ；按照本节的 Bayes 观点， $\pi(\theta)$  就是  $\theta$  的先验分布。统计问题的判决空间可表示为  $\mathcal{D}$ 。V 为一个判决，统计判决函数的风险函数为

$$R(\pi, \delta) = E_{\pi}[L(\theta, \delta(X))] = \int_{\Omega} L(\theta, \delta(x)) \pi(d\theta) \quad (8.2.1)$$

(8.2.1) 统计问题就是要在一定条件下，求出使  $R(\pi, \delta)$  达到最小的解。在参数

因此很自然地应考虑  $L(\theta, \delta(x))$  在  $\theta$  上的平均，从而得

空间  $\Theta$  上有测度的情形下，即在  $\mu$  的情形下，

$\delta$  中仍然包

含随机变量 到 Bayes 风险

I

,

$$R(\pi, \delta) = \int_{\Omega} L(\theta, \delta(x)) \pi(d\theta) \quad (8.2.2)$$

表示损失函数  $L(\theta, \delta(x))$  在样本空间  $\Omega$  和参数空间  $\Theta$  上的整体平均，而 Bayes 解就是要求出使  $R(\pi, \delta)$  最小的解。为了突出其重要性，我们把定义 3.1.5 归纳如下：

定义 8.2.1 对于给定的统计判决问题，设其损失函数为  $L(\theta, \delta)$ ，参数  $\theta$  的先验分布为  $\pi(\theta)$ 。若存在判决  $\delta^*(x)$ ，使对一切  $\theta$  (幻有) 则称  $\delta^*(x)$  为统计判决问题的 Bayes 解；若  $\pi$  的为广义先验分布，则称  $\delta^*(x)$  为统计判决问题的广义 Bayes 解；对于参数

8.2 Bayes 估计

357

估计问题的 Bayes 解或广义 Bayes 解，则称为 Bayes 估计或广义 Bayes

估计。综合 (8.2.1) 式和 (8.2.2) 式，Bayes 风险  $R(\pi, \delta)$  可表示为

$$R(\pi, \delta) = E_{\pi}[L(\theta, \delta(x))] \quad (8.2.3)$$

$$L(\theta, \delta(x)) \pi(d\theta) f(x) dx \quad (8.2.3)$$

其中  $E[\cdot]$  表示对  $(X, \pi)$  的联合分布求期望。

注 Bayes 解不但与损失函数  $L(\theta, \delta)$  的选取有关，也与当初先验分

布  $\pi(\theta)$  的选取有关，这两者都对 Bayes 解的结果有重要影响。特别是，当损失函数确定以后，先验分布  $\pi(\theta)$  的选取仍然可以千变万化，因而可得到多种多样的 Bayes 解 (包括 Bayes 估计)，这是 Bayes 统计的重要特点之一。

截至目前为止，我们实际上只涉及统计判决观点，下面的后验风险才是 Bayes 统计所特有的，也是最重要的，因为 Bayes 统计推断都是从后验分布出发的。

(义，

根据第一节的讨论可知， $(X, \pi)$  的联合分布可表示为

没)

$\pi(X) \pi(\theta|X)$ ，为后验分布。由于 (8.2.3) 式中的

其中  $\pi(\theta)$  为  $X$  的边缘分布 被积函数非负，因此积分可以交换次序，并可表示为

,

$$R(\pi, \delta) = \int_{\Omega} L(\theta, \delta(x)) \pi(d\theta) \quad (8.2.4)$$

(8.2.4)

该式方括号中的积分称为后验风险，定义如下。

定义 8.2.2 损失函数  $L(\theta, \delta(x))$  关于后验分布  $\pi(\theta|X)$  的加权平

均称为后验风险，\* 它定义为

$$r(\delta|x) = \int_{\Theta} L(\theta, \delta(x)) \pi(d\theta) \quad (8.2.5)$$

后验风险与 Bayes 风险有非常密切的关系，由 (8.2.4) 式和 (8.2.5) 式可得

$$r(x) = \int_{\mathcal{X}} L(e, d) p(x) dy(x). \quad (8.2.6)$$

由这个关系式可以看出，若一个统计判决函数使后验风险 $r(x)$ 达到最小，则它也会使Bayes风险 $R(\delta)$ 达到最小。事实上，我们有以下重要定理。

定理8.2.1 设 $X \sim f(x, \theta)$ ,

统计判决问题的损失函数为 $L(e, d)$ ，参数 $\theta$ 的先验分布为 $\pi(\theta)$ 。若 $r(x)$ 和 $\pi(\theta)$ 在 $\mathcal{X}$ 上都存

358

## 第八章 Bayes统计基础

在有限的最小值(对几乎所有的 $\theta$ )，则它们有相同的解；若损失函数 $L(e, d)$ 在 $\mathcal{X}$ 上为严凸函数，则该统计判决问题的Bayes解几乎处处唯一。

证明 今记 $r^*(x) = \min_{\delta} r(\delta, x)$ ， $r^*(x) = \min_{\delta} \int_{\mathcal{X}} L(e, d) p(x) dy(x)$ 。

则由假设可知，对任何 $\delta(x)$ 有 $r(\delta(x), x) \geq r^*(x)$ ，该式两边积分，并应用(8.2.6)式可得 $R(\delta) \geq \int_{\mathcal{X}} r^*(x) p(x) d\mu(x) = R^*$ 。

该式对任何 $\delta(x)$ 成立，因此根据定义， $\delta^*(x)$ 为判决问题的Bayes解。反之，若 $\delta^*(x)$ 为给定统计判决问题的Bayes解，则对任何 $\delta(x)$

即 $r(\delta(x), x) \geq r^*(x)$ ，应用积分关系式(8.2.6)可得 $R(\delta) - R(\delta^*) = \int_{\mathcal{X}} (r(\delta(x), x) - r^*(x)) p(x) d\mu(x) \geq 0$ 。

(8.2.7)

但是由于 $r^*(x)$ 使后验风险 $r(\delta^*(x), x)$ 达到最小，所以有 $r(\delta^*(x), x) = r^*(x)$ 。

因此以上积分必为零；被积函数也为零，即 $r(\delta(x), x) = r^*(x)$ 。

另外，若损失函数 $L(e, d)$ 在 $\mathcal{X}$ 上关于 $\delta(x)$ 为严凸函数，则由(8.2.5)式可知，其积分作为 $\delta(x)$ 的函数，仍然关于 $\delta(x)$ 为

严凸函数(对几乎所有的 $\theta$ )。因为由凸函数的定义可得 $L(\lambda e_1 + (1-\lambda)e_2, d) \leq \lambda L(e_1, d) + (1-\lambda)L(e_2, d)$ ，

其中 $0 < \lambda < 1$ 。该式两边关于 $\pi(\theta)$ 积分，由(8.2.5)式可知，不等式关于 $\delta(x)$ 仍然成立：

$R(\lambda \delta_1 + (1-\lambda)\delta_2) \leq \lambda R(\delta_1) + (1-\lambda)R(\delta_2)$ 。

因而 $R(\delta)$ 为 $\delta(x)$ 的严凸函数。则根据凸函数的性质可知， $r^*(x)$ 的最小值存在必唯一(对几乎所有的 $\theta$ )。

以上定理说明，一个统计判决问题的Bayes解就是使后验风险达到最小的解，这与Bayes统计推断的出发点为后验分布是一致的。注意，以上定理不但可用于参数估计问题，也可用于假设检验以及其他统计判决问题。

该式对 $\delta^*(x)$ 也成立，因此有 $\int_{\mathcal{X}} (r(\delta^*(x), x) - r^*(x)) p(x) d\mu(x) = 0$ 。

8.2 Bayes估计

359

推论 若 $\delta^*(x)$ 使下式达到最小，则为Bayes解： $\delta^*(x) = \arg \min_{\delta} \int_{\mathcal{X}} L(e, d) p(x) d\mu(x)$ 。

因为由(8.1.1)式知， $r(\delta(x), x) = \int_{\mathcal{X}} L(e, d) p(x) d\mu(x)$ ，该式代入

(8.2.5)式后，后验风险 $R(\delta)$ 与上式仅相差一个常数 $c(x)$ ，因而

有相同的最小值点。以下应用定理8.2.1，分别考虑损失函数为均方损失和0-1损失

时的Bayes估计。8.2.1 均方损失估计

8.2.1 均方损失估计 8.2.1 均方损失估计 8.2.1 均方损失估计

8.2.1 均方损失估计 8.2.1 均方损失估计 8.2.1 均方损失估计

8.2.1 均方损失估计 8.2.1 均方损失估计 8.2.1 均方损失估计

8.2.1 均方损失估计 8.2.1 均方损失估计 8.2.1 均方损失估计

(8.2.9) (8.2.8)式是最常用的损失函数;(8.2.9)式可用于尺度参数的估计(见第四章),或某些特殊问题.

定理8.2.2 设 $X \sim f(x, \theta)$ ,  $\theta$ 的先验分布为若损

失函数为  $L(e, d) = A(\theta[d - g(\theta)]^2)$ , 则  $g(\theta)$ 唯一的 Bayes 估计为  $\hat{g} =$

$E[AM|g(\theta)] = \int \lambda(4) \rho(8) \delta(8) d\theta$

(8.2.10)

$g(e) \int r(e) f(x, \theta) d\theta$

$i \lambda X = E[g(\theta)] = \frac{\int \lambda(4) \rho(8) \delta(8) d\theta}{\int \lambda(4) \rho(8) d\theta}$ . (8.2.11)

$f(77(\theta))/(X, \theta) \propto J(\theta)$

证明设 $g(\theta)$ 的某一估计为 $\hat{g} = \hat{g}(x)$ , 则由(8.2.5)式及(8.2.9)式可知, 其 Bayes后验风险为

$R(\hat{g}) = \int [A(\theta)[\hat{g}(x) - g(\theta)]^2 \pi(\theta|x) d\theta$

其中

$E[A(\theta)] = \int A(\theta) \pi(\theta|x) d\theta$

若损失函数为 $L(\theta, d) = [\hat{g} - g(\theta)]^2$ (即上面 $A(\theta) = 1$ ), 则有

$a$

$A(\theta) \pi(\theta|x) d\theta = E[A(\theta) | X] > 0,$

$= \int [52(x) - 28(x)g(\theta) + g^2(\theta)] \pi(\theta|x) d\theta = 32(x) + 65(x) + c,$

360

## 第八章 Bayes统计基础

$b = -2 \int \hat{g}(\theta) g(\theta) \pi(\theta|x) d\theta = -2E[A(\theta)g(\theta) | X],$

$c = \int A(\theta) g^2(\theta) \pi(\theta|x) d\theta.$

$R(\hat{g})$ 作为 $\hat{g}$ 的二次三项式, 当 $\hat{g}(x) = -b/2a$ 时,  $R(\hat{g}|x)$ 达到最小值, 而且解唯一, 因此得到(8.2.10)式. 在(8.2.10)式中取 $X(6) = 1$ , 则有 $E[g(\theta) | X] = \int g(\theta) \pi(\theta|x) d\theta$ ; 再根据后验分布的公式(8.1.1)式计算后验期望, 即可得到(8.2.10)和(8.2.11)的第二式. ■

以上定理无论在理论上还是在应用上都有非常重要的意义, 该定理表明: 在均方损失下, 任何参数估计问题都可归结为后验分布的数学期望(即积分)的计算问题. 从原理上来讲, 只要给出先验分布, 算出后验分布, 即可求出参数估计, 不存在理论上的困难. 对于常见的简单分布和待估参数, 常常可得到显式解; 对于某些比较复杂的情形, 若得不到显式解, 则可归结为数值积分问题.

由于只需计算常见分布的数学期望, 因此本节介绍的后验期望估计通常比第三章和第四章介绍的传统估计方法更加简单直接(某些复杂的情况除外). 在上一节, 我们介绍了若干先验分布的选取方法, 并对于样本分布为二项分布、Poisson分布、正态分布、指数分布、均匀分布等常见分布的情形得到了相应的后验分布, 这实际上也就得到相应参数的Bayes估计, 因为只需再求一下数学期望即可. 例如对正态分布, 例

8.1.2中得到了均值的后验正态分布(72已知时), 因而也就自然地得到了它的Bayes估计见(8.1.4)式, 同时那里也对后验均值

意义作了说明; 同理, 由例8.1.5可以得到 $\sigma^2 = a^2$ 的Bayes估计(72已知时); 如果得到 $a^2$ 和 $M$ 的联合后验分布, 也就能得到它们的Bayes估计(见习题). 其他分布的情况也类似.

如上所述, 根据上一节的结果, 我们对不少常见分布实际上已经得到了相应参数的Bayes估计, 因此在下面的例题中, 我们略去与上一节重复的部分, 而更加注重于说明Bayes估计的意义和性质. 另外, 在以下例题中, 如无特别申明, 都是在二次损失(8.2.8)式下求Bayes解, 即应用公式(8.2.11)式, 求待估参数 $g(\theta)$ 的后验期望估计.

例8.2.1 设 $X_1, \dots, X_n$ 为独立同分布样本, 且有 $X \sim f(x, \theta)$ 求 $\theta$ 的 Bayes估计, 假设 $\theta$ 的先验

分布为(1)  $\theta \sim \pi(\theta, 1)$ ; (2)  $\theta \sim BE(p, q)$ .

解 (1)由例8.1.1 可知  $\pi(x) = BE(T+1, n-T+1)$

$T: X$ . 因此由(8.2.11)式可得

$\pi = I \cdot$

## 8.2 Bayes 估计

361

$e_B = E[\theta|x]$

$n+2$

这一估计有很好的统计意义, 它与极大似然估计  $\theta_M = T/n$  相差很小:  $e_B - \theta_M = O_p(n^{-1/2})$ .

因此,  $e_B$  是  $\theta$  的相合估计, 同时也是渐近无偏的 (即当  $n \rightarrow +\infty$  时其偏差趋向于零). 但是, 以上 Bayes 估计在直观上比极大似然估计更加符合实际情况. 例如, 若  $\theta = P(X_i=1)$  表示次品率或命中率, 若本次观测  $T=0$ , 则  $e_B = 1/(n+2)$ , 表示  $\theta$  很小, 而  $\theta=0$  表示不可能事件; 显然  $e_B$  更合理. 若  $T=n$ , 则  $e_B = n/(n+2)$  表示事件出现的可能性很大, 而  $\theta=1$  则认为是必然事件, 显然  $e_B$  更合理一些. 另外, 若本次观测  $r = n = 10$ , 则有  $e_B = 11/12$ ;  $T = n = 100$ , 则  $e_B = 101/102$ , 直观上看, 它说明“10发10中”与“100发100中”对评价命中率是有差别的.

(2)由例8.1.1 可知  $\pi(\theta|x) = \frac{1}{n+1} \binom{n}{T} \theta^T (1-\theta)^{n-T}$  因此有

$= E(\theta^k | X) = \frac{1}{n+1} \frac{\Gamma(n+1) \Gamma(k)}{\Gamma(n-k+1) \Gamma(k+1)}$  (8.2.12)

该式也有较好的统计解释. 由假设  $\theta \sim BE(p, q)$ , 其先验均值为  $p/(p+q)$ . 这可理解为:  $p+q$  次试验中有  $p$  次成功; 对于样本,  $n$  次试验中有  $T$  次成功, 因此综合先验信息与样本信息, 可理解为:  $n+p+q$  次试验中有  $T+p$  次成功, 因此得到后验均值为 (8.2.12) 式. 另外 (8.2.12) 式还可以表示为

$e_B = \frac{T+p}{n+p+q}$

因此  $e_B$  表示样本均值与先验均值的加权平均. | 例8.2.2 设  $X_1, \dots, X_n$  为独立同分布样本, 且有  $X_i \sim \exp(A) = r(A, 1)$ ; 设  $A$  的先验分布为  $\pi$  分布,  $A \sim \pi(A) \sim r(a, p)$ , 求  $g(A) = E(A | X_1, \dots, X_n)$  的 Bayes 估计 (若  $E(A)$  表示寿命分布, 则  $g(A)$  可视为元件寿命大于  $t$  的概率).

解 由例8.1.4 可知,  $A$  的后验分布为  $\pi(A | x) \sim r(a+T, n+p)$ ,

$T = \sum_{i=1}^n X_i$ , 因此由 (8.2.11) 式可得  $i = 1$

$= \int_0^\infty e^{-At} \pi(A | x) dA = \frac{C e^{-(a+T)t}}{(a+T)^{n+p+1}}$

该式直接积分可得

$\pi(A | x) = \frac{(a+T)^{n+p}}{\Gamma(n+p)} e^{-(a+T)A}$

$\pi(A | x) = \frac{(a+T)^{n+p}}{\Gamma(n+p)} e^{-(a+T)A}$

该式直接积分可得

$\pi(A | x) = \frac{(a+T)^{n+p}}{\Gamma(n+p)} e^{-(a+T)A}$

$\pi(A | x) = \frac{(a+T)^{n+p}}{\Gamma(n+p)} e^{-(a+T)A}$

由于  $E(X_j | A) = 1/A$ , 经简单计算可得  $i$  因此

$r(n+p)$

)

362

## 第八章 Bayes 统计基础

是  $g(A)$  的相合估计. ■ 例8.2.3 设  $X_1, \dots, X_n$  为独立同分布样本,  $\pi$  表示寿命分布, 并

且只观测到前  $r$  个寿命值:  $y$

望估计: i)  $\theta \sim M(r, 1)$ ,  $M$  服从无信息先验分布,  $BP = r/(i) \theta C1$ ;

ii)  $JV. 1$ ,  $\theta$  服从逆  $r$  先验分布  $\pi^*(a, p)$ .

解 i) 当  $X_i \sim \pi(x | r(i, i))$  时, 由第一章的公式可知,  $\pi = (\pi_1, \dots, \pi_r)$  的分布可表示为

$= C \exp \left[ -\sum_{i=1}^r \frac{1}{\theta_i} \right]$ , 其中  $T_{nr} = r_1 + \dots + r_r$  由于  $\pi(M) \propto \theta^r$ , 因此有

$\pi = \dots$

求相应参数的后验期

$$\pi(\theta | y) = c(y) \pi(\theta) / \int c(y) \pi(\theta) d\theta$$

1,

\

$$\pi(\theta) = C \int \pi(\theta | y) d\theta = r, \dots, y_x = x(1) \cdot$$

\* 2.1.1 这一结果与UMRUE完全一致. 如果 $\theta$ 的变化范围限制在 $\theta > 0$ , 也可得到相应的估计, 但结果略有不同.

ii) 当 $X_1, \dots, X_n \sim r(y, 1)$ 时, 由第一章的公式可知,  $y = (y_1, \dots, y_n)$ 的分布可表示为

$$f(y, \theta) = \frac{1}{(n-r)!}$$

$$(n-r)!$$

由于 $\pi(\theta) \sim r_1(a, p)$ , 因此后验分布可表示为

$$\pi(\theta | y) = c(y) \pi(\theta)^{(p+1)} e^{-\theta} \cdot \frac{1}{\Gamma(p+1)} = c(y) e^{-(p+r+1)\theta} \theta^p$$

$$\sim r_1(\theta, p+r+1, p+r+1). \text{ 因此由逆 } r \text{ 分布的期望公式可得}$$

其中 $e(y) = ne^n$ ?

由 $E(K) = E(m|K)$ , 经直接积分可得

$$a + T + a = E(a | y) = \dots$$

$p+r-1$  该式亦可表示为先验均值与样本均值的加权平均,

$$n p - a \cdot$$

$$2.1.1 \int \pi(\theta) d\theta = 1$$

如果取无信息先验分布, 或者取(8.2.9)式所示的损失函数, 即 $L(\theta, d) = \dots$  可得到类似的结果. 歷

例8.2.4 设 $X_1, \dots, X_n$  为独立同分布样本,  $X_i$  服从均匀分布  $U(0, 1)$ . 损失函数取为 $L(\theta, d) = |d - \theta|$ , 并取无信息先验分布

,

$$\pi(\theta) = 1$$

## 8.2 Bayes 估计

363

$\pi(\theta) \propto \theta^{-1}$ , 求 $\theta$ 的 Bayes 估计. 解后验分布可表示为

$$\pi(\theta | y) \propto \theta^{-1}$$

因此 $\theta$ 服从Pareto分布 $PR(n, x(n))$ , 且有 $c(a) = nx(n)$ . 根据公式(8.2.10)式,

$$g(\theta) := \int \pi(\theta | y) d\theta = \int \theta^{-1} d\theta = \dots$$

$$= \frac{1}{n} \int \theta^{-1} d\theta = \dots$$

直接计算可知, 若 $y$ 服从Pareto分布 $PR(a)$ , 则有 $E(Y) = a$ ,  $E(y-2) = a$  因此有 $E(\hat{\theta}) = a$ . 这些结果代入上式可得

这一结果与第四章同变估计的结果(见例4.3.2)完全一致, 所取的损失函数也相同. |

例8.2.5 设 $X = (X_1, \dots, X_n)^T$  服从指数族分布:  $f(x, \theta) = h(\theta) \cdot \exp \{ \sum_{i=1}^n x_i \theta_i - n \psi(\theta) \}$  的先验分布为 $\pi(\theta)$ , 求 $\theta$ 的Bayes估计.

解 由(8.1.1)式,  $\theta$ 的后验分布可表示为

$$\pi(\theta | x) \propto \pi(\theta) f(x, \theta)$$

在这个表达式中,  $\theta$ 为随机变量的值; 若把 $(\theta_1, \dots, \theta_k)^T$  看成参数, 则 可重新表示为

$$\pi(\theta | x) = h(\theta) \exp \{ \sum_{i=1}^k \theta_i x_i - J(\theta) \}$$

其中  $h(\theta) = \pi(\theta) e^{-J(\theta)}$ ,  $-b(x) = \log [c(x)/i(x)]$ . 因此后验分布为指数族分布, 其后验期望可由指数族分布的有关公式得到(见第一章

$$(1.5.11) \text{ 式}), \text{ 即 } E(\theta | x) = -b'(x), \text{ 因此有}$$

$b(x) = \dots$  该公式的进一步的讨论可参见Lehmann and Casella(1998). 歷

$$T(a|\%) = c(x)A(x)7r(^)\exp.$$

f=1

]. (8. 2. 13)

364

## 第八章 Bayes统计基础

由前面的不少例子可以看出, Bayes估计常常是渐近无偏的(亦可 见本章习题), 但是以下定理说明, 它们不可能是严格无偏的.

定理8. 2.3 设 $X \sim$  先验分布, 考虑 $g(\theta)$ 在均方误差损失(8.2.8)下的Bayes估计  $\hat{S}(X)$ , 并假设 $E[g^2(\hat{\theta})] < \infty$ . 则 $\hat{S}(X)$ 不可能既是Bayes估计, 又是 无偏估计, 除非其Bayes风险  $R_n(\theta)$ 为零, 即

$$E[\hat{S}(X) - g(\theta)]^2 = 0. \quad (8.2. 14)$$

其中 $E$  表示对 $(X, \theta)$ 的联合分布求期望(见(8. 2.3)式).

证明 今考虑 $g(\theta)$ 的无偏估计及Bayes估计的定义与性质.. 若  $\hat{S}(X)$ 为 $g(\theta)$ 的无偏估计, 则有

$$E[\hat{S}(X)] = g(\theta).$$

由于 $\theta$ 具有非广义的先验分布 $\pi(\theta)$ , 因此对于 $(X, \theta)$ 的联合分布可应 用条件期望的性质:

$$\begin{aligned} E[\hat{S}(X)g(\theta)] &= E[E[\hat{S}(X)g(\theta) | X]] \\ &= E[g(\theta)E[\hat{S}(X) | X]] = E[g^2(\theta)]. \end{aligned} \quad (8.2. 15)$$

若 $\hat{S}(\theta)$ 为 $g(\theta)$ 的Bayes估计, 则 $\hat{S}(X) = E[g(\theta) | X]$ . 由条件期望的

性质有

$$\begin{aligned} E[\hat{S}(X)g(\theta)] &= E[E[g(\theta) | X]g(\theta)] \\ &= E[E[g(\theta) | X]E[g(\theta) | X]] = E[\hat{S}^2(X)]. \end{aligned}$$

而 $\hat{S}(X)$ 的Bayes风险 $R_n(\theta)$ 可表示为

$$\begin{aligned} R_n(\theta) &= E[\hat{S}(X) - g(\theta)]^2 \\ &= E[\hat{S}^2(X)] - 2E[\hat{S}(X)g(\theta)] + E[g^2(\theta)] \\ &= E[\hat{S}^2(X)] - 2E[g^2(\theta)] + E[g^2(\theta)] = E[\hat{S}^2(X)] - E[g^2(\theta)]. \end{aligned} \quad (8. 2. 16)$$

$R_n(\theta) \geq 0$ , 具有正常的(即非广义的)

因此, 若 $\hat{S}(X)$ 既是Bayes估计, 又是无偏估计, 则! ( 8. 2. 15 )式和 (8.2.16)式代人上式可得 $R_n(\theta) = 0$ , 即(8.2.14)式成立. 因此由以上定理可知, 若 $\hat{S}(X)$ 是 $g(\theta)$ 在均方误差损失下的Bayes

估计, 其 Bayes风 险  $> 0$ , 则它不可能是无偏的;反之, 若  $\hat{S}(X)$  是 $g(\theta)$ 的无偏估计, 且在均方误差损失下的Bayes风险 $> 0$ , 则 它不可能是Bayes估计. 这一定理可推广到有偏估计的情形, 见本章 习题.

例8.2.6 设 $X_1, \dots, X_n$  为独立同分布样本,  $E(X_i) = \mu$ , 而 $\text{Var}(X_i) = \sigma^2$

已知. 则在均方误差损失下, 对于任何先验分布, 无都不可能是 $\mu$ 的 Bayes估计.

解 由定义可知, 对于任意给定的都有 $E[(X - \mu)^2] = \text{Var}(X) = \sigma^2/n$ , 因此对于任何先验分布  $\pi(\theta)$ 都有

### 8. 2 Bayes 估计

365

$$E[(X - \mu)^2] = E[E[(X - \mu)^2 | \theta]] = E[\sigma^2/n] = \sigma^2/n > 0.$$

因此由定理8. 2.3可知, 又不可能是 $\mu$ 的Bayes估计. ■ 但是, 定理8.2.3对广义先验分布不一定成立. 例如, 在例8.1.2

的i) 中, 要估计 $g(\mu) = \mu$ , 但 $\mu$ 服从广义均匀分布:  $\pi(\mu) = 1, \mu \in (-\infty, \infty)$  (这时条件 $E[\mu^2 | Y]$   $< \infty$ 也不成立). 由例8. 1.2可知, 其 后验分布为 $\pi(\mu | Y) \propto \exp(-\mu^2/2\sigma^2)$

$\propto \exp(-\mu^2/2\sigma^2)$ 因此有 $E(\mu | Y) = \mu$ , 这时又既是

$\mu$ 的无偏估计, 又是广义的Bayes估计. 以下Pitman估计也是广义Bayes估计, 其中第(1)个估计, 即位置

后验极大似然估计与第三章介绍的极大似然估计有很多类似之处，并且包含普通的极大似然估计为其特例。但是，两者的出发点完全不一样，后验极大似然估计就是0-1损失下的



Bayes估计. 设  $X \sim f(x, \theta)$ ,  $\theta \in \Theta$ ,  $\Theta \subset \mathbb{R}^n$  的先验分布为  $\pi(\theta)$ 、并要求  $g(\theta)$  的Bayes估计. 为了简单起见, 考虑  $\theta$  的估计, 并取损失函数为

这时称为后验极大似然估计.

$a \sim 3p(X$

$\|d - \theta\| \leq \epsilon$ , 其中  $\epsilon$  为充分小的实数. 其直观意义为: 若一个判决 (即参数  $\theta$  的估计) 与真参数的距离较大, 则损失为1; 若与真参数的距离很小, 则 损失函数为零, (8.2.17) 简称为0-1损失.

$\|d - \theta\| \leq \epsilon$  E,

(8.2. 17)

,  $\epsilon$ ),  $\pi$  的先验分布为在  $\Theta$ -

定理8.2.4 设  $X \sim f(x$

损失(8. 2. 17)下, 若  $f$  可以充分小, 则  $\theta$  的Bayes估计为使  $\int_{\Theta} f(x|\theta) d\theta$  达到 极大值的估计, 即  $M$  满足

$\int_{\Theta} f(x|\theta) d\theta = \max_{\theta \in \Theta} \int_{\Theta} f(x|\theta) d\theta$ . (8. 2. 18) 6b &

1

8. 2

即

Bayes 估计

证明 设  $\theta$  的估计为  $d = d(x)$ , 则其后验风险为

367

$\int_{\Theta} f(x|\theta) d\theta = \int_{\Theta} f(x|\theta) d\theta, \int_{\Theta} f(x|\theta) d\theta = \int_{\Theta} f(x|\theta) d\theta$

$\int_{\Theta} f(x|\theta) d\theta = \int_{\Theta} f(x|\theta) d\theta$

$\int_{\Theta} f(x|\theta) d\theta = \int_{\Theta} f(x|\theta) d\theta = \int_{\Theta} f(x|\theta) d\theta$ . 'S(x)-s

大, 也等价于使以下积分最大:

$\int_{\Theta} f(x|\theta) d\theta = \int_{\Theta} f(x|\theta) d\theta$

$\int_{\Theta} f(x|\theta) d\theta = \int_{\Theta} f(x|\theta) d\theta$

由积分中值定理可得

$R_n(g) = \int_{\Theta} f(x|\theta) d\theta, \int_{\Theta} f(x|\theta) d\theta = \int_{\Theta} f(x|\theta) d\theta$ .

因此以  $\theta$  应使以上  $\int_{\Theta} f(x|\theta) d\theta$  达到最大, 当时有  $\theta^*(x)$ , 因此  $\theta^*(x)$  应使  $\int_{\Theta} f(x|\theta) d\theta$  达到最大. 所以, 在0-1损失(8.2.17)下,  $\theta^*(x)$  使  $\int_{\Theta} f(x|\theta) d\theta$  达到最小等价于使  $\int_{\Theta} f(x|\theta) d\theta$  达到最大, 即  $\theta^*(x)$  应满足

(8.2. 18)式. 推论 若  $\pi(\theta)$  的先验分布为  $\pi(\theta)$ , 则相应的后验极大似然估计即为通常的极大似然估计.

证明 应使  $\int_{\Theta} f(x|\theta) d\theta = \int_{\Theta} f(x|\theta) d\theta$  达到最大, 当  $\int_{\Theta} f(x|\theta) d\theta = \int_{\Theta} f(x|\theta) d\theta$

时, 等价于使  $\int_{\Theta} f(x|\theta) d\theta = \int_{\Theta} f(x|\theta) d\theta$  达到最大, 这就是通常的极大似然估计. ■ 若  $\pi(\theta)$  有共同的支撑集, 则  $\ln \pi(\theta)$  称为后验对数 似然函数;  $\ln \pi(\theta)$  也使  $\int_{\Theta} f(x|\theta) d\theta$  达到最大. 后验极大似然估计与通常的极大 似然估计有十分相似的性质, 因为两者都是使某个目标函数达到最大值 的解; 其目标函数也很相似: 都是随机变量的密度函数或者它们的对 数.

因此, 第三章和第五章所介绍的极大似然估计的大多数性质对后验

极大似然估计也成立 (当然也需要一定的正则条件), 例如, 不变原理, 子集参数的似然, Gauss - Newton迭代法等等, 对后验极大似然估计也 适用.

例8.2.8 设为独立同分布样本, 且有  $X_i \sim \text{叫}$ , ,

$\mu$  已知. 若  $M(\mu, d)$ , 求  $M$  的后验极大似然估计. 解由假设可知

$\int_{\Theta} f(x|\theta) d\theta = \int_{\Theta} f(x|\theta) d\theta$

$\int_{\Theta} f(x|\theta) d\theta = \int_{\Theta} f(x|\theta) d\theta$

$\int_{\Theta} f(x|\theta) d\theta = \int_{\Theta} f(x|\theta) d\theta$

"

要求6(%),使  $|x|$  最小, 相当于求5(%),使上式第二项积分最

$I(X)$

其后验对数似然函数及其导数分别为

$Z(p|x)$

$21n$

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

$+ \log c(x).$

368

第八章Bayes统计基础

$dL/d\theta$

该式为零可解出  $M$  为

$M = \frac{1}{n} \sum_{i=1}^n x_i$

$0-Q$

$\sum_{i=1}^n x_i \sim$

$N(a, \sigma^2)$

由(8.1.4)式可知, 其结果与后验期望估计  $k$  完全相同. ■

例8.2.9 设  $x_1, \dots, x_n$  为独立同分布样本, 求相应参数的后验极大似然估计: i)  $X$ .

$(1, \theta)$ , 并设  $\theta \sim P(A)$ , 并设  $A \sim \Gamma(A, p)$ .

解 i) 由例8.1.1可知,  $\Gamma(\theta|x) \sim BE(p+\sum x_i, n+\sum 1)$ , 因此有  $\frac{d}{d\theta} \Gamma(\theta|x)$

$= c(x) \theta^{p+\sum x_i-1} e^{-\theta(n+\sum 1)}$

$\frac{d}{d\theta} \Gamma(\theta|x) = (p+\sum x_i-1) \theta^{p+\sum x_i-2} e^{-\theta(n+\sum 1)} - (n+\sum 1) \theta^{p+\sum x_i-1} e^{-\theta(n+\sum 1)}$

$c(x)$ . 直接求导可得

$n + p + \sum x_i - 1$

$\theta^{p+\sum x_i-2} e^{-\theta(n+\sum 1)}$

由例8.2.1可知, 该式与  $\theta$  相差无几, 因而也是相合的, 渐近无偏的. ii) 由

$\Gamma(\theta|x) \sim \Gamma(\sum x_i + 1, n + \sum 1)$ , 直接计算可知

$\frac{d}{d\theta} \Gamma(\theta|x) = -(\sum x_i + 1) \theta^{\sum x_i} e^{-\theta(n+\sum 1)} + (n+\sum 1) \theta^{\sum x_i+1} e^{-\theta(n+\sum 1)}$

(而

$\theta^{\sum x_i+1} e^{-\theta(n+\sum 1)}$

由例8.2.2可知, 该式与  $\theta$  相差无几, 因而也是相合的, 渐近无偏的.

例8.2.10 设  $X_1, \dots, X_n$  为独立同分布样本,  $X_i \sim r(\theta, 1)$ ,

其中  $\theta = \frac{1}{a}$  都未知, 在无信息先验分布下, 求  $a$  的后验极大似然估计.

解 由例8.1.9可知, 当  $\theta$  和  $\frac{1}{a}$  都未知时, 可取无信息先验分布  $\pi(a, \theta) \propto a^{-2}$ , 因此后验分布可表示为(见(8.1.10)式)

$\pi(a, \theta|x) = c(x) a^{-j} e^{-\frac{1}{a} \sum x_i} / \int_0^\infty \int_0^\infty a^{-j} e^{-\frac{1}{a} \sum x_i} da d\theta$

如上所述, 得到后验似然函数以后, 求解后验极大似然估计的方法与求解普通极大似然估计的方法完全类似. 由上式可知, 对任意固定的  $a$ , 若  $\theta$  越大, 则  $\pi(\theta|x, a)$  越大; 而  $\theta$  必须满足, 因此  $\theta = \sum x_i$  时

. 把  $\theta = \sum x_i$  代入  $\pi(\theta|x, a)$ , 并取

$\frac{d}{da} \pi(\theta|x, a)$  最大, 所以有  $\frac{1}{a} = \sum x_i$

而  $\theta = \sum x_i$

8.2 Bayes 估计 对数可得

$Z(\theta), \text{ or } I(\theta) = -\frac{1}{n} \log \pi(\theta|x)$

369

由此直接求导可得  $\frac{d}{d\theta} Z(\theta) = \frac{1}{n} \sum x_i$  (普通的极大似然估计为  $\sum x_i/n$ ). I 例8.2.11 设  $F \sim N(X, \sigma^2)$

$Z$ , 其中  $K$  为  $z$  维观测向量,  $X$  为  $n \times p$  阶已知矩阵,  $\theta$  为  $p$  维未知参数向量. 假设  $a$  已知;  $\theta$  的先验分布

为  $\pi(\theta) \sim 1/V(\theta, \tau^2)$ ，求  $\theta$  的后验极大似然估计。解由假设可知

$$P \sim \pi(\theta) \exp(-\frac{1}{2}(\theta - \tau)^2 / \tau^2)$$

因此的后验分布可表示为

$$\pi(\theta|y) = c(y) \pi(\theta) \exp(-\frac{1}{2}(\theta - \tau)^2 / \tau^2) \\ QW = (\theta - \tau) / (\tau^2 + 1) + (\tau - \theta) / (\tau^2 + 1)$$

要求  $A$  使  $\pi(\theta|y)$  最大，相当于求  $\theta$  使  $QW$  最小，该式直接求导可得

$$\frac{d}{d\theta} QW = -\frac{1}{\tau^2 + 1} = 0,$$

$\theta = \tau$ 。该式通常称为广义岭估计，若  $\tau = 0$ ，则上式化为普通的最小二乘估计

$\hat{\theta} = (X^T X)^{-1} X^T Y$ 。注意， $\tau = 0$  相当于精度为 0，即服从无信息先验分布： $\pi(\theta) \propto 1$ （见 (8.2.19) 式）。另外，若  $\tau^2$  未知但服从逆  $\chi^2$  分布，也可得到类似的结果。

以上我们介绍了基于均方损失和  $0-1$  损失情形下的 Bayes 估计。定理 8.2.1 也可用于其他损失函数情形，以下定理考虑了绝对损失下的 Bayes 估计。

证明 根据定理 8.2.1，Bayes 估计就是使后验风险达到最小的解。给定  $\theta$  的一个估计

$\delta(x)$ ，它对应于绝对损失的后验风险为

$$R(\delta|x) = \int |\delta(x) - \theta| \pi(\theta|x) d\theta = E(|\delta(x) - \theta| | x) \quad \text{根据定理 1.1.1, } \delta(x) \text{ 等于 } \pi(\theta|x) \text{ 的中位数时, } R(\delta|x) = \int |\delta(x) - \theta| \pi(\theta|x) d\theta$$

$\theta$ ， $\theta$  的先验分布为  $\pi(\theta)$ 。若损

定理 8.2.5 设

失函数为绝对损失  $f(\delta, \theta) = |\delta - \theta|$ ，则  $\delta$  的 Bayes 估计为后验分布  $\pi(\theta|x)$  的中位数。

370

## 第八章 Bayes 统计基础

$E(|\delta - \theta| | x)$  达到最小值。因此定理的结论成立。■ 8.2.4 Bayes 估计的某些性质

本小节主要介绍：如何通过 Bayes 估计求解 Minimax 估计的方法，其中也涉及容许估计。由第三章的定义可知，Minimax 估计以及容许估计的求解通常都很困难。虽然无论从形式上，还是从内容上来看，它们与 Bayes 估计都没有什么关系，但是，由于在 Bayes 估计中，先验分布的选取方法千变万化、多种多样，因此有可能通过选取某些特定的先验分布而得到 Minimax 估计或容许估计。事实上，可考虑一般统计判决问题的 Bayes 解与 Minimax 解以及容许性之间的关系。

首先考虑 Bayes 解与 Minimax 解之间的关系。设  $X \sim \pi(x, \theta)$ ，统计判决问题（包括估计问题）的损失函数为  $L(\theta, d)$ ，判决函数以  $\delta$  对应的风险函数为  $R(\delta)$

$\pi(\theta)$  的 Bayes 解为  $\delta^*(\pi)$ ，其相应的 Bayes 风险为  $R^*(\pi)$ 。由定义可知，Minimax 解是在峰值中求最小；而 Bayes 解是在平均值中求最小，所以前者的风险值一般应该比较大。因此，只有某些特定的先验分布，才能使 Bayes 解也是 Minimax 解，有的著作称之为最不利的先验分布（Lehmann and Casella, 1998；郑忠国, 1998）。但是，由于先验分布选取方法的多样性，这种先验分布还是有可能存在的，以下两个定理是比较常见的情形，也是当前寻求 Minimax 估计的主要方法（陈希孺, 1981, 1999）。

定理 8.2.6 设  $X \sim \pi(x, \theta)$ ， $\theta \in \Theta$ ，统计判决问题的损失函数为  $L(\theta, d)$ ，则有

i) 对于  $\theta$  的任意正常的（即非广义的）先验分布  $\pi(\theta)$  都

，8)，并设  $\theta$  的先验分布为  $\pi(\theta)$

。对于损失函数  $L(\theta, d)$ ，它使  $\max_{\theta} R(\delta^*(\pi), \theta)$  达到最小

其相应的Minimax解记为 $\delta_5(\cdot)$  (见第三章), 其相应的Minimax风险为 $r_5$ 又记对应于先验分布

有似 ii)若存在正常的先验分布 $\pi(\theta)$ ,使得则对

应于此先验分布的Bayes解也是Minimax解; Hi)若存在正常的先验分布 $\pi(\theta)$ , 使得其Bayes解心(幻的风险

函数)为常数, 则此Bayes解也是Minimax解.

证明 i)由定义可知, 对于任意的判决 $\delta(\cdot)$ , 都有  $R(\delta, \pi) \geq R(\delta_5, \pi)$ ,

由于先验分布是非广义的, 因此两边关于 $\pi(\theta)$ 积分可得

$$\int R(\delta, \pi) \pi(\theta) d\theta = R(\delta, \pi) \geq R(\delta_5, \pi) = \int R(\delta_5, \pi) \pi(\theta) d\theta = R(\delta_5, \pi)$$

(心);

## 8.2 Bayes 估计

371

由于上式对任意的判决 $\delta$ 都成立, 所以有)成 $(S:)$ • ii) 若有  $R(\delta_5, \pi) = R(\delta, \pi)$  则由(8.2.20)式有

$R(\delta_5, \pi) = R(\delta, \pi)$ , 该式对任意的判决 $\delta(x)$ 都成立, 这说明是Minimax解.

iii) 若Bayes解 $\delta_5(\cdot)$  (幻的风险函数 $R(\delta_5, \pi)$ 为常数, 则由于先验分布是非广义的, 所以有 $R(\delta_5, \pi) = R(\delta, \pi)$ , 因此由ii)可知  $\delta_5(x)$ 是 Minimax 解. ■ 例8.2.12 设

$X_1, \dots, X_n$  为独立同分布样本, 且有 $X_i \sim N(\theta, 1)$ ,

在均方损失下求 $\theta$ 的 Minimax估计.

解 由例8.2.1可知, 若取 $\theta$ 的先验分布为 $\theta \sim N(\mu, \tau^2)$ , 则在均

方损失下 $\theta$ 的Bayes估计为

$$\hat{\theta} = E(\theta | X) = \frac{\tau^2 \bar{x} + \mu}{1 + \tau^2}$$

n

其中 $r = \frac{\tau^2}{1 + \tau^2}$ . 由此可直接计算出该估计的风险函数 $R(\hat{\theta}, \pi)$ 为  $i=1$

$$R(\hat{\theta}, \pi) = E(\hat{\theta} - \theta)^2$$

$$= \frac{\tau^2}{1 + \tau^2} [1 - \tau^2] + \frac{\tau^4}{(1 + \tau^2)^2} [p(1 - \tau^2) - q\tau^2]^2, \quad (8.2.21) \quad (n + p + q) \sim$$

上式对任意的 $p, q$ 都成立, 我们可设法选取适当的 $p, q$ , 使 $R(\hat{\theta}, \pi)$ 为常数, 即与 $\theta$ 无关, 则由定理8.2.6的iii)可知, 相应的Bayes估计为Minimax估计. 为此, 只需使(8.2.21)式中 $\tau^2$ 与 $\theta$ 的系数为零即可,

即 $(p - q)\tau^2 = n$ ;  $2p(p + q) = n$ . 由此可得 $p = q = \sqrt{n}/2$ , 这时 $r$  (没, 么)为 常数, 因而相应的Bayes估计为Minimax估计. 所以 $\hat{\theta}$ 的Minimax估计为

$$\hat{\theta} = \frac{\bar{x}}{1 + \tau^2} = \frac{\bar{x}}{2}$$

以上定理的应用范围比较窄, 因为通常不易找到风险函数为常数的 Bayes解, 而下面定理的应用范围要广泛得多.

定理8.2.7 设 $X_1, \dots, X_n$  统计判决问题的损失函数为  $L(e, d)$ . 若有一列正常的先验分布, 其Bayes解为 $\delta_k(x)$  ( $k = 1, 2, \dots$ ), 并且其相应Bayes风险 $r_k$ 的极限存在, 即  $\lim_{k \rightarrow \infty} r_k = r$ .

$k \rightarrow \infty$

则有

i) 若存在判决 $\delta^*(\cdot)$ , 使 $R(\delta^*, \pi) = r$ , 则 $\delta^*(x)$ 为Minimax解; ii) 若 $\delta^*(x)$ 的风险函数 $R(\delta^*, \pi)$ 为常数, 且等于 $r$ , 则 $\delta^*(\cdot)$ 为

372

## 第八章 Bayes统计基础

Minimax 解. 证明i)对于任意的判决函数 $\delta(\cdot)$ , 由于 $r_k$ 为一列正常的先

验分布, 因此由(8.2.20)式可得 $R_n(\delta_k) = R(\delta_k, \pi_k)$ , 所以也有  $R(\delta_k) \geq r$ ; 再由假设可得

$M(\theta^*) \leq M(\theta) \leq V(\theta)$ , 因此  $S^*(x)$  是 Minimax 解.

ii) 若判决  $S^*(x)$  满足  $E[L(\theta, S^*(x))] = r$ , 则也有  $E[L(\theta, S^*(x))] = r$ , 因此由 i) 知  $S^*(x)$  是 Minimax 解.

定理 8.2.7 常常可用于以下情况. 若  $S^*(x)$  为  $\theta$  的关于广义先验分布  $\pi(\theta)$  的广义 Bayes 估计, 其风险函数为常数  $r$ . 在某些情况下, 可构造一系列正常的先验分布  $\pi_n(\theta)$  ( $n = 1, 2, \dots$ ), 使其相应 Bayes 估计的风险函数的极限等于  $r$ . 这样, 根据定理 8.2.7 的 ii),

$S^*(x)$  为  $\theta$  的 Minimax 估计. 以下两个例子都是这种情形. 例 8.2.13 设  $X_1, \dots, X_n$  为独立同分布样本, 且有  $X_i \sim N(\theta, 1)$ .

证明: 在均方损失下,  $X$  为  $\theta$  的 Minimax 估计.

证明 首先必须注意, 在均方损失下,  $\theta$  的风险函数为  $R(\theta, \hat{\theta}) = E[(\hat{\theta} - \theta)^2]$ .

虽然其值为常数, 但是正如例 8.2.6 所述, 又不是  $\theta$  的 Bayes 估计 (而是广义 Bayes 估计,  $\pi(\theta) \equiv 1$ ); 因而不能应用定理 8.2.6 的结论 (这种情况比较常见, 下一个例子也是).

为了证明  $\hat{\theta} = \bar{X}$  是  $\theta$  的 Minimax 估计, 通常可选取  $\theta$  的一系列正常的先验分布, 使其相应 Bayes 估计的风险函数的极限等于  $R(\theta, \bar{X})$ , 则由定理 8.2.7 可知,  $\bar{X}$  为

$\theta$  的 Minimax 估计. 由例 8.1.2 可知,  $\bar{X}$  的后验分布为一系列正态分布:  $\pi_n(\theta | x) \sim N(\frac{n\bar{x} + \theta_0}{n+1}, \frac{1}{n+1})$ .

为此取  $\theta$  的一系列先验分布为  $\pi_n(\theta) \sim N(\theta_0, \frac{1}{n})$ .

其中  $\theta_0 = \frac{1}{2}$ ,  $\frac{1}{n} \rightarrow 0$  当  $n \rightarrow \infty$ .

由此可得  $\theta$  的一系列 Bayes 估计为  $S_n(\bar{X}) = \frac{n\bar{X} + \theta_0}{n+1}$ , 其相应的风险函数为

$R(\theta, S_n(\bar{X})) = E[(S_n(\bar{X}) - \theta)^2]$ .

由于  $\pi_n(\theta) \sim N(\theta_0, \frac{1}{n})$ , 上式关于  $N(\theta_0, \frac{1}{n})$  求期望, 可得相应的 Bayes 风险为

$R(\theta, S_n(\bar{X})) = \frac{1}{n+1} + \frac{n}{n+1} R(\theta, \bar{X})$ .

由此可得  $\theta$  的一系列 Bayes 估计为  $S_n(\bar{X}) = \frac{n\bar{X} + \theta_0}{n+1}$ , 其相应的风险函数为

$R(\theta, S_n(\bar{X})) = E[(S_n(\bar{X}) - \theta)^2] = \frac{1}{n+1} + \frac{n}{n+1} R(\theta, \bar{X})$ .

由于  $\pi_n(\theta) \sim N(\theta_0, \frac{1}{n})$ , 上式关于  $N(\theta_0, \frac{1}{n})$  求期望, 可得相应的 Bayes 风险为

$R(\theta, S_n(\bar{X})) = \frac{1}{n+1} + \frac{n}{n+1} R(\theta, \bar{X})$ .

所以由定理 8.2.7 可知,  $\bar{X}$  为  $\theta$  的 Minimax 估计. | 例 8.2.14 设  $X_1, \dots, X_n$  为独立同分布样本,  $X_i$  服从均匀分布  $U(0, 1)$ .

证明  $\hat{\theta} = \frac{n+1}{n+2} \bar{X}$  为  $\theta$  的 Minimax 估计. 例 8.2.14 设  $X_1, \dots, X_n$  为独立同分布样本,  $X_i$  服从均匀分布  $U(0, 1)$ .

证明  $\hat{\theta} = \frac{n+1}{n+2} \bar{X}$  为  $\theta$  的 Minimax 估计.

8.2 Bayes 估计

373

尺 (0, 1). 损失函数取为  $L(\theta, \hat{\theta}) = [(\hat{\theta} - \theta)^2]$ , 证明:  $\hat{\theta} = \frac{n+1}{n+2} \bar{X}$  为  $\theta$  的 Minimax 估计.

解 由例 8.2.4 可知,  $\pi(\theta)$  为无信息先验分布  $\pi(\theta) \propto 1$  下的  $\theta$  的 Bayes 估计. 这是一个广义 Bayes 估计. 该估计的风险函数  $R(\hat{\theta}, \theta)$  为

$R(\hat{\theta}, \theta) = E[(\hat{\theta} - \theta)^2] = \frac{1}{n+2} + \frac{n}{n+2} R(\theta, \bar{X})$ .

由例 8.2.4 可知,  $\pi(\theta)$  为无信息先验分布  $\pi(\theta) \propto 1$  下的  $\theta$  的 Bayes 估计. 这是一个广义 Bayes 估计. 该估计的风险函数  $R(\hat{\theta}, \theta)$  为

$R(\hat{\theta}, \theta) = E[(\hat{\theta} - \theta)^2] = \frac{1}{n+2} + \frac{n}{n+2} R(\theta, \bar{X})$ .

由例 8.2.4 可知,  $\pi(\theta)$  为无信息先验分布  $\pi(\theta) \propto 1$  下的  $\theta$  的 Bayes 估计. 这是一个广义 Bayes 估计. 该估计的风险函数  $R(\hat{\theta}, \theta)$  为

(8.2.23)

该式的计算用到了  $y = X(n)/\Gamma(n) = \Gamma(n) \int_0^1 x^{n-1} (1-x)^{n-1} dx$ . 因此  $\hat{\theta} = \frac{n+1}{n+2} \bar{X}$  的风险函数为  $R(\hat{\theta}, \theta) = \frac{1}{n+2} + \frac{n}{n+2} R(\theta, \bar{X})$ .

常数. 为了证明  $\hat{\theta}$  是  $\theta$  的 Minimax 估计, 可取  $\theta$  的一系列先验分布为 Pareto 分布  $PR(k, \frac{1}{k})$ , 即  $\pi_k(\theta) \propto \theta^{k-1} e^{-\theta}$ ,  $k = 1, 2, \dots$ . 这是一列正常的先验分布 (当  $k \rightarrow \infty$  时, 这些分布可看做无信息先验分布的一个逼近). 由例 8.1.6 可知,  $\theta$  的后验分布也是一列 Pareto 分布:  $\pi_k(\theta | x) \propto \theta^{k-1} e^{-\theta}$ , 其中  $n_k = n + k - 1$ ,  $a_k = \sum_{j=1}^n x_j$ .

我们要考虑  $k \rightarrow \infty$  时,  $\theta$  的相应 Bayes 估计的风险函数, 这时  $k \rightarrow \infty$  时, 因此有  $\pi_k(\theta | x) \propto \theta^{k-1} e^{-\theta}$ .

这个后验分布与例 8.2.4 的后验分布在形式上完全类似 (那里是  $\pi(\theta | x) \propto \theta^{n-1} e^{-\theta}$ ).

这个后验分布与例 8.2.4 的后验分布在形式上完全类似 (那里是  $\pi(\theta | x) \propto \theta^{n-1} e^{-\theta}$ ).

这个后验分布与例 8.2.4 的后验分布在形式上完全类似 (那里是  $\pi(\theta | x) \propto \theta^{n-1} e^{-\theta}$ ).

$77(\wedge 1\%) - ra(n, x(n))$ ), 因此, 根据完全类似的推导可知, 相应的Bayes 估计为 (即在例 8. 2. 4 中把  $a$  换成  $n$ )

与 (8. 2. 23) 式类似, 其相应的风险函数  $R(\theta, \theta_k)$  为

$$= E|1 - \frac{1}{n+1}| \quad (8.2.24)$$

由于  $Y=X(n)/\theta \sim BE(n+1, \theta)$ , 因此  $R(\theta, \theta_k)$  与  $\theta$  无关, 所以  $\theta$  的 Bayes 风险  $Q$  也就是  $R(\theta, \theta_k)$  (注意,  $PR(k \sim k \sim 1)$  是正常的先验分布). 又由

于当  $k \rightarrow \theta$  时,  $n_k = n + k \sim 1 - \theta^n$  (注意,  $n$  固定), 所以比较 (8. 2. 23) 式和 (8. 2. 24) 式可知

$$R(\theta, \theta_k) - R(\theta, \theta) = \frac{1}{n+1} = \frac{1}{n+1}$$

因此由定理 8.2.7 可知,  $\hat{X}(n)$  为 Bayes 估计. |

Minimax 解虽然比较保守, 但是这种最大风险最小化的策略在某些领域还是很有用的 (茆诗松, 1999). 求 Minimax 解通常比较困难, 我们不能指望对一大类问题得到一般的求解方法, 只能逐个加以解决 (Lehmann and Casella, 1998). 目前来看, Bayes 方法是求 Minimax 解

oo ).

374

## 第八章 Bayes 统计基础

的主要工具, 特别是 Minimax 估计, 大多通过以上两个定理获得.

最后简要介绍 Bayes 解的容许性. 由于 Bayes 解和 Minimax 解都在

某种程度上反映了判决问题在参数空间的整体优良性, 因而它们有可能是容许的 (范金城, 吴可法, 2001).

定理 8.2.8 设若以  $\delta$  为统计问题关于损失函数  $L(\delta, \theta)$  和先验分布  $\pi(\theta)$  在判决空间  $D$  上唯一的 Bayes 解, 则它必是容许的.

证明 用反证法. 若  $\delta$  不是容许的, 则根据定义 (见第三章), 必存在  $\delta'(x)$ , 使得  $R(\delta, \theta) > R(\delta', \theta)$  对一切  $\theta \in \Theta$  成立, 而且至少对一个  $\theta_0$  有  $R(\delta, \theta_0) < R(\delta', \theta_0)$  (该式说明  $Sf(x)$  不同于  $\delta(x)$ ). 以上不等式两边关于先验分布  $\pi(\theta)$  积分可得

$$R(\delta) = \int R(\delta, \theta) \pi(\theta) d\theta > \int R(\delta', \theta) \pi(\theta) d\theta = R(\delta').$$

(8. 2. 25) 这说明,  $\delta'$  也是一个 Bayes 解, 与唯一性的假设矛盾. 因此  $\delta$  是容许的.

| 注 类似的反证法也可证明: 若  $\delta$  为统计问题关于损失函数  $L(\delta, \theta)$  唯一的 Minimax

解, 则它也是容许的, 这只需把 (8.2.25) 式的

取平均 (积分) 改为取峰值即可. 定理 8.2.9 设  $X \sim f(x, \theta)$ ,  $\theta \in \Theta$ , 以  $\delta$  为统计问题关于

损失函数  $L(\delta, \theta)$  和先验分布  $\pi(\theta)$  的 Bayes 解, 若  $\pi(\theta)$  在  $\Theta$  上处处  $> 0$ ; 风险函数

$R(\delta, \theta)$  对任意的  $\delta$  都是  $\theta$  的连续函数, 且  $Sf(\theta)$  的 Bayes 风险有

限, 则是容许的. 证明仍用反证法. 若  $\delta$  不是容许的, 则必存在  $\delta'(\theta)$ , 使得

$R(\delta, \theta) > R(\delta', \theta)$  对一切  $\theta \in \Theta$  成立, 而  $\delta$  至少对一个  $\theta_0$  有

$R(\delta, \theta_0) < R(\delta', \theta_0)$ . 由于  $R(\delta, \theta)$  及充分小的正数  $\epsilon > 0$ , 使得

$R(\delta, \theta) < R(\delta', \theta) + \epsilon$  在  $\theta_0$  的一个邻域  $U$  上成立. 由以上不等式可得

$$\begin{aligned} R(\delta) &= \int R(\delta, \theta) \pi(\theta) d\theta < \int R(\delta', \theta) \pi(\theta) d\theta + \epsilon \int \pi(\theta) d\theta \\ &= R(\delta') + \epsilon < R(\delta) \end{aligned}$$

,

6)

8) 和  $7(1?$ ,

:久(8)-eP(0 e TVJ .

### 8.3假设检验与区间估计的Bayes方法

375

由假设可知  $P(6eNe) > 0$ , 因此, 这与  $5(x)$  是 Bayes 解的假设矛盾. 所以  $S(\text{幻})$  必须是容许的. I

由定理8.2.1和定理8.2.2可知, 许多常见的Bayes估计(诸如在均方损失(8.2.25)式下的Bayes估计等)都有唯一性, 因而它们是容许的. 另外, 不少常见的共轭先验分布能满足定理8.2.9的条件, 因而相应的Bayes解是容许的. 以上定理既说明了 Bayes估计的优良性, 同时也提供了一种证明容许性的有效方法. 容许性是一个纯理论的问题, 迄今还未见关于应用成果方面的报道, 进一步的文献可参见Lehmann and Casella(1998); 陈希孺(1981 J999).

### 8.3 假设检验与区间估计的Bayes方法

对于假设检验与区间估计问题, Bayes统计推断的出发点仍然是后验分布, 但是其基本思想比较简单直观.

#### 8.3.1 Bayes假设检验

给定样本 $X$ , 且有 $X \sim f(x, \theta)$ , ② 考虑假设检验问题

—  $(\theta_0 \cup \theta_1) = (9. (8.3.1)$

在第六章, 我们系统地介绍了常见分布的各种假设检验问题及其求解方法. 其基本点就是要确定一个检验统计量, 其分布在零假设时已知, 由此即可进一步得到否定域与检验函数. 构造检验统计量是经典统计解决假设检验问题最重要的步骤, 同时也是最困难的部分. 但是, 假设检验的Bayes方法就不需要检验统计量, 而是从后验分布出发, 通过直接计算后验概率导出否定域与检验函数.

假设给定 $\theta$ 的先验分布 $K(\theta)$ , 并计算出其后验分布为 $e \sim \pi(\theta|x)$ . 由此可以得到参数落在 $\theta_0$ 和 $\theta_1$ 的后验概率:

$$\pi_0(X) = \int_{\theta_0} \pi(\theta|x) d\theta = P\{\theta \in \theta_0 | x\}, \quad (8.3.2)$$

$$\pi_1(x)$$

$$\pi_1(\theta|x) d\theta = P\{\theta \in \theta_1 | x\}. \quad (8.3.3)$$

从直观上看,  $\pi_0(x)$ 和 $\pi_1(x)$ 分别表示参数 $\theta$ 属于 $\theta_0$ 和 $\theta_1$ 的后验概率, 因此若 $\pi_0(x) > \pi_1(x)$ , 则表示 $\theta \in \theta_0$ 的可能性更大, 因而否定

所以可取否定域与检验函数为

$$\pi_0(x) > \pi_1(x) \quad (8.3.4)$$

376

### 第八章Bayes统计基础 (8.3.5)

其中 $\pi_1$ 表示 $\pi_0$ 的余集.

例8.3.1 设 $X_1, \dots, X_n$ 为独立同分布样本, 且 $X_i \sim B(1, \theta)$ . 假设

没的先验分布为 $\pi(\theta) \sim U(0, 1)$ , 考虑以下假设检验问题的否定域 $H_0$ :  $\theta \leq 1/2$  —

$H_1: \theta > 1/2$ .

解 由例8.1.1可知, 没的后验分布为 $\pi(r+1, n-T+1)$ ,  $n$

其中 $r = \sum_{i=1}^n X_i$  因此后验概率( $\pi_0$ 和 $\pi_1$ )可由不完全 $\beta$ 函数表示  $\pi_0 = I_{1/2}(n-r+1, r+1)$

$$\pi_0(X) = P(\theta \leq 1/2 | X) = \int_0^{1/2} \pi(\theta|x) d\theta = \int_0^{1/2} \frac{1}{B(n-r+1, r+1)} \theta^{n-r} (1-\theta)^r d\theta$$

因此否定域为 $\pi_1 = 1 - \pi_0$

例如, 当 $n=5$ , 观察值 $T=0, 1, 2, 3, 4, 5$ 时, 我们计算出后验概率  $\pi_0(x)$  的值分别

为 0.015, 0.109, 0.344, 0.656, 0.891, 0.984. 因此

时就否定原假设, 这显然是合理的. I 下面我们从统计判决函数的观点出发, 根据定理8.

2.1来证明上述

直观论断的正确性。首先考虑检验问题的判决函数和损失函数。对于假设检验问题(8.3.1)，假设其否定域为 $T$ （不一定是(8.3.4)的形式），检验函数 $\phi(x)$ 如(8.3.5)式所示。检验问题的判决函数只取两个值： $\phi = 1$ 表示 $H_1$ 成立， $\phi = 0$ 表示 $H_0$ 成立，通常取 $d_1 = 1$ 表示 $H_1$ 成立， $d_0 = 0$ 表示 $H_0$ 成立，因此判决函数就等同于非随机化的检验函数 $\phi(x) = I_T(x) = \begin{cases} 1, & x \in T \\ 0, & x \in T^c \end{cases}$ 。

这时，否定域 $T$ ，检验函数 $\phi$ 以及判决函数 $d$ 三者统计意义基本相同。对于给定的判决函数 $\phi(x)$ ，通常取0-1损失函数，即判错时损失为1，判对时损失为0，具体可表示为 $L(\phi, x) = \begin{cases} 1, & \phi(x) \neq 1 \\ 0, & \phi(x) = 1 \end{cases}$ 。

(8.3.7) 根据Bayes观点，我们必须求出对应于以上检验函数和损失函数的后验风险，并找出使后验风险达到最小值的解，即 Bayes解。

引理8.3.1 对于假设检验问题(8.3.1)的任一判决函数 $\phi$ ，假设相应的否定域为 $T$ ，则在0-1损失下，其后验风险可表示为

$$R(\phi | x) = \int_0^1 L(\phi, x) d\pi(x) \\ = \int_0^1 \begin{cases} 1, & \phi(x) \neq 1 \\ 0, & \phi(x) = 1 \end{cases} d\pi(x) \\ = \int_{T^c} \pi(x) d\pi(x) \\ (8.3.8)$$

### 8.3 假设检验与区间估计的Bayes方法

377

其中 $\pi(x) = P(\theta = 0 | \phi = 0, 1)$ 如(8.3.2)和(8.3.3)所示。证明 根据(8.3.7)式的定义， $\int_0^1 L(\phi, x) d\pi(x)$ 可表示为

$$R(\phi | x) = \int_0^1 L(\phi, x) \pi(x) d\pi(x) + \int_0^1 L(\phi, x) (1 - \pi(x)) d\pi(x) \cdot \pi_0$$

记上式第一项为 $J_1(\phi)$ ，第二项为 $J_2(\phi)$ ，则积分主要计算 $L(\phi, x) = 1$

的部分。首先计算 $J_1(\phi)$ ，这时 $\phi(x) = 1$ ，因而由

$$(8.3.7) \text{ 式有 } L(\phi, x) = 1,$$

$$R_1(\phi) = \int_0^1 L(\phi, x) \pi(x) d\pi(x) = \int_0^1 \pi(x) d\pi(x) = \pi_0$$

同理可得

$$J_2(\phi) =$$

$$0,$$

$$J_2(\phi) = \int_0^1 L(\phi, x) (1 - \pi(x)) d\pi(x) =$$

$$\int_0^1 \pi(x) d\pi(x) =$$

两式合并即可得到(8.3.8)式。这个引理说明，对于给定的假设检验问题(8.3.1)，尽管其判决函

数 $\phi(x)$ 可能有多种多样的形式，但是后验风险只能取两个值： $R_1(\phi)$ 和 $R_2(\phi)$ 。它们由 $\pi_0$ 和 $\pi_1$ 决定。这给求Bayes解，即后验风险最小的解

提供了方便。直观上，Bayes解的后验风险每次都应该取两个值 $R_1(\phi)$ 和 $R_2(\phi)$ 中较小者。

今具体考虑检验问题(8.3.1)的Bayes解可能的形式，具体分析如下。首先，若 $\phi(x)$ 使(8.3.8)式中的达到最小值，其相应的否定域和检验函数记为

$$\phi^*(x) = \begin{cases} 1, & x \in R^* \\ 0, & x \in R^c \end{cases}$$

$$(8.3.9)$$

， $\pi_0, \pi_1$ ，因此，求Bayes解 $\phi^*(x)$ 主要就是求否定域 $T$ 的形式；其次，由(8.3.8)

式可知，对于任意固定的 $\phi$ ，任何判决 $\phi$ （包括 $\phi^*$ ）只能取两个值 $R_1(\phi)$ 和 $R_2(\phi)$ 。

若 $\phi(x)$ 使 $\int_0^1 \pi(x) d\pi(x)$ 成立，则 $R_1(\phi) = \int_0^1 \pi(x) d\pi(x)$ 或

值中较小者，而当 $\phi(x) = \phi^*(x)$ ， $R_1(\phi) = R_1(\phi^*)$ 或





有

$f$  十少旧兮卜卜<sup><0</sup>

域类似 分布决定.

$=\{x: a > 0 \mid = \{x: n(x - \theta_0) > (\theta_0 - f_j L)/r^2\}$ . 易见, 这个否定域与根据 Neyman-Pearson理论(MLR检验)得到的否定

, 但是此处常数  $k = -f_j L/r^2$ ,

由先验 ■ 以下例题说明, 对于简单假设问题, Bayes解与Neyman-Pearson理论得到的解是一致的.

例8.3.3 设对于简单假设检验问题  $H_0: \theta =$

$\theta_0$

$U = \theta$  若取先验分布为

$\pi(\theta), e = e_0, \pi(\theta) =$

$(8-3-15)$  叹=氏,  $\pi(\theta > t | x) = c(x) \pi_r / (\pi_r + c(x))$  ( $i =$

则其Bayes否定域可表示为

其中  $c = \pi_r / \pi_r$ .

证明 由假设可知  $H$ . 则有

$\theta, 1)$

$\pi(\theta, X) = \pi(\theta | X) = \pi(\theta) \pi_r / (\pi_r + c(x))$

( $i = 0, 1$ ).

因此代入(8.3.11)式有

$\pi(\theta | X) = \pi(\theta) \pi_r / (\pi_r + c(x))$  因此根据(8.3.11)式, 简单假设检验问题的Bayes否定域为

$\pi^* = \{x: \pi(\theta | X) > 1\}$  由此即可得到(8.3.15)式.

$\pi(\theta | X) = \pi(\theta) \pi_r / (\pi_r + c(x))$

,

$\pi(\theta | X) = \pi(\theta) \pi_r / (\pi_r + c(x))$ ,

$\pi(\theta | X) = \pi(\theta) \pi_r / (\pi_r + c(x))$ ,

(8.3.16)

380 第八章 Bayes统计基础

在不少文献中(如Berger 1985, Shao 1998, 茆诗松1999), 经常提到 假设检验问题的 Bayes因子这一名字, 我们首先结合简单假设检验问题 说明其统计意义. 由(8.3.16)式可得

$\pi(\theta | X) = \pi(\theta) \pi_r / (\pi_r + c(x))$

该式右端为<sup>^</sup>假设下和<sup>0</sup>假设下的似然比, 其值越大, 则成

立的可能性越大, 因此它反映了样本对假设<sup>0</sup>支持的程度. (8.3.17)式左端为<sup>0</sup>假设下和<sup>1</sup>假设下后验概率比与先验概率比 之间的比值, 它应该与右端有相同的含义, 即反映了样本对原假设

支持的程度.

对于一般假设检验问题(8.3.1), 我们得不到象(8.3.17)式那样简

单明了的表达式, 但是该式左端仍然有类似的含义, 可以推广到一般情形, 称为Bayes因子, 定义如下:

定义8.3.1 设 $X \sim f(x | \theta)$   $\theta$ 的先验分布为 $\pi(\theta)$ , 其相应的后验分布为 $\pi(\theta | x)$ . 对于假设检验问题(8.3.1), 其Bayes因子 定义为

$B(x) = \pi(\theta | x) / \pi(\theta)$

$t$

$= \pi(\theta | x) / \pi(\theta)$ ,

$= \pi(\theta | x) / \pi(\theta), i = 0, 1.$

$\pi(x) = \frac{f(x|\theta_1)}{f(x|\theta_1) + f(x|\theta_0)}$ ;  $\pi(x) = \frac{f(x|\theta_1)}{f(x|\theta_1) + f(x|\theta_0)}$ , 77

;

(8.3.18) 在这个定义中, 后验概率比  $\pi(x|\theta_1)$  综合了先验信息与样本信息的影响, 其值越大, 表明好。成立的可能性越大; 它除以先验概率

比  $\pi(x|\theta_1)$ , 以后, 就抵消了一部分先验信息的影响, 因而就较多地突出了样本, 即数据的影响。因此不少学者认为, Bayes因子反映了样本对原假设  $H_0$  支持的程度。另外, 若把

(8.3.18)式表示为  $\pi(x|\theta_1)/\pi(x|\theta_0) = B(x)$  (式(8.3.18)), 由此式亦可看出:  $B(x)$  越大, 则  $H_1$  成立的可能性越大。但是, Bayes检验的主要任务还是根据定理8.3.1 以及(8.3.10) - (8.3.12)式求出检验问题的后验概率和否定域。

在 Bayes检验中, 需要特别予以关注的是以下常见的单点双边假设检验问题:

$H_0: \theta \geq \theta_0$  vs  $H_1: \theta < \theta_0$ . (8.3.19) 对于这个检验, 其先验分布不能简单地套用前面两节所介绍的若干连续型分布。因为连续型分布在一个点  $\theta_0$  处的概率恒为零, 从而导致后验概率亦为零, 无法进行有效的检验。因此对于检验(8.3.19), 通常应在  $\theta_0$  处赋予一定概率, 以便得到非零的后验概率。所以, 其先验分布的一般形式应当为

### 8.3 假设检验与区间估计的Bayes方法

381

其中 因此有,

$\pi(\theta_0) =$ ,

$f(\theta_0)$ ,

$\theta_0 = \theta_0$ ,

(8.3.20)

$U \sim \text{Unif}(\theta_0, \theta_1)$ ,  $\pi(\theta_0) = \pi(\theta_1)$ ,

其中  $0 < \pi(\theta_0) < 1$  表示在  $\theta = \theta_0$  处的先验权重, 在  $\theta \in (\theta_0, \theta_1)$  上,  $\pi(\theta) = 1 - \pi(\theta_0)$

示先验权重, 为一个连续型或离散型先验分布。因此, 若  $I \sim f(x, \theta)$ ,  $\theta \in \Theta$ , 则后验分布可表示为

$\pi(\theta|x) = \frac{\pi(\theta)f(x|\theta)}{\int \pi(\theta)f(x|\theta)d\theta}$ ,

$\pi(\theta_0|x) = \frac{\pi(\theta_0)f(x|\theta_0)}{\pi(\theta_0)f(x|\theta_0) + \int_{\theta_0}^{\theta_1} \pi(\theta)f(x|\theta)d\theta}$ .

$m(x) := \frac{\pi(\theta_0)f(x|\theta_0)}{\pi(\theta_0)f(x|\theta_0) + \int_{\theta_0}^{\theta_1} \pi(\theta)f(x|\theta)d\theta}$

由此可得检验(8.3.19)的后验概率和交比  $K$ 。以及 Bayes因子分别为

$\pi(\theta_0|x) = \frac{\pi(\theta_0)f(x|\theta_0)}{\pi(\theta_0)f(x|\theta_0) + \int_{\theta_0}^{\theta_1} \pi(\theta)f(x|\theta)d\theta}$   $c(x) = m(x) =$

$0 =$

$z = \frac{\pi(\theta_0)f(x|\theta_0)}{\pi(\theta_0)f(x|\theta_0) + \int_{\theta_0}^{\theta_1} \pi(\theta)f(x|\theta)d\theta}$ ,  $\pi(\theta_0) = \pi(\theta_1)$ .

$\pi(\theta_0|x) = \frac{\pi(\theta_0)f(x|\theta_0)}{\pi(\theta_0)f(x|\theta_0) + \int_{\theta_0}^{\theta_1} \pi(\theta)f(x|\theta)d\theta}$

(8.3.22)

(8.3.23)

开 of  $(x, \theta_0)$  由以上公式可知, 在假设检验问题(8.3.19)的求解过程中, 要考

虑到  $\theta = \theta_0$  处权重的影响, 而在  $\theta \in (\theta_0, \theta_1)$  处则要扣除权重的影响(即  $\pi(\theta) = 1 - \pi(\theta_0)$ ), 因而要分别进行计算。至于先验分布  $\pi(\theta)$  若是连续型分布, 则由它得到的后验分布与前两节介绍的情况类似, 因为  $\theta = \theta_0$  处的概率为零。

另外, 根据充分性准则(见(8.1.3)式), 以上公式中  $X \sim f(x, \theta)$  的分布密度可以换成充分统计量  $T = T(X)$  的分布密度, 结果不变。

例8.3.4 设  $A_1, \dots, A_n$  为独立同分布样本, 且  $X_i \sim \text{Bernoulli}(p)$ . 求以下假设检验问题的 Bayes解:  $H_0: p = 1/2$  vs  $H_1: p \neq 1/2$ . 假设  $p$  的先验分布如(8.3.20)所示, 其中

$\pi(1/2) = \pi_0$ , 在  $\theta \in (1/2, 1)$  处先验分布  $\pi(\theta)$  为均匀分布  $\pi(\theta) = 2(1 - \pi_0)$ .

解 由假设可知,  $\pi_0 = 1/2$ ,  $\pi(\theta) = 2(1 - \pi_0)$  其中  $T = \sum_{i=1}^n X_i$ .

因而  $\pi(X, \theta) = (1/2)^n$ . 当  $\theta = 1/2$  时,  $\pi$  为均匀分布  $\pi(\theta, 1)$ , 由例 8.1.1 可知, 的后验分布为分布  $BE(T+1, n-r+1)$ , 因此有

$$\pi(\theta) = \int_0^1 \pi(\theta, t) dt = \int_0^1 \frac{1}{B(T+1, n-r+1)} t^{T+1} (1-t)^{n-r} dt = \frac{1}{B(T+1, n-r+1)} \frac{1}{T+1} \frac{1}{n-r+1} \frac{1}{B(T, n-r)} = \frac{1}{B(T+1, n-r+1)}$$

382

## 第八章 Bayes 统计基础

$$r(r+1)\cdots(r+n-1) = \frac{(n+r-1)!}{r!}$$

$$A(4) = 1 - \frac{1}{2} \frac{1}{(n-1)!} \frac{1}{2} \frac{1}{(n-2)!} \cdots \frac{1}{2} \frac{1}{(n-r)!} = \frac{1}{2^n} \frac{1}{(n-r)!}$$

例如, 若取  $\theta = 1/2$ ,  $n = 6$ , 当观察值  $X = 0, 1, 2, 3, 4, 5, 6$  时,  $P(X \leq 3)$  的值分别为 0.143, 0.524, 0.610, 0.457, 0.390, 0.256, 0.143. 因此  $r = 0, 1$ , 或  $r = 5, 6$  时就否定原假设, 即  $r$  取值太大或太小则  $\theta = 1/2$  不成立, 这显然是合理的. 另外,  $T = 3$  时 (即成功与失败的次数相等), Bayes 因子  $B(3) = 2.185$ , 这表明, 样本强烈支持原假设  $\theta = 1/2$ . 这也显然是合理的.

例 8.3.5 设  $X_1, \dots, X_n$  为独立同分布样本, 且有  $X_i \sim N(\theta, \sigma^2)$   $f(x)$  已知, 求以下假设检验问题的 Bayes 解:  $H_0: \theta = \theta_0$  vs  $H_1: \theta \neq \theta_0$ . 假设  $\pi(\theta)$  的先验分布如 (8.3.20) 所示, 其中  $\pi(\theta) = \frac{1}{\sigma} \exp\left(-\frac{1}{\sigma} |\theta - \theta_0|\right)$  (其中  $\pi$  为无信息先验; 关于  $\theta$ ), 若  $\theta_0$  接近于  $\theta$ , 把先验均值设置为  $\theta_0$  是合适的).

解 i) 当  $\sigma^2$  已知时,  $r = r(x) =$  无充分统计量,  $x \sim N(\theta, \sigma^2/n)$ . 因此可用  $T = \bar{X}$  的分布代替  $X$  的分布. 由 (8.3.23) 式可得

由于交叉比  $K(\theta)$  为  $\exp\left(-\frac{1}{2\sigma^2} (\theta - \theta_0)^2\right)$  的严增函数, 因此否定域可表示为

$x \sim \theta_0$   
这个否定域与经典的结果一致, 但是此处常数  $c$  与先验分布有关.  $H_0$  这时  $\pi(\theta_0, c)$ , 由 (8.3.23) 式可得

$$R' = \{K(\theta) > 1\} = \left\{ \frac{1}{\sigma} \exp\left(-\frac{1}{\sigma} |\theta - \theta_0|\right) > 1 \right\}$$

尺  $\pi(\theta)$  (龙)  
 $S(\theta) = \int_0^\infty \exp\left(-\frac{1}{2\sigma^2} (\theta - \theta_0)^2\right) \frac{1}{\sigma} \exp\left(-\frac{1}{\sigma} |\theta - \theta_0|\right) d\theta = \frac{1}{\sigma} \int_0^\infty \exp\left(-\frac{1}{2\sigma^2} (\theta - \theta_0)^2 - \frac{1}{\sigma} |\theta - \theta_0|\right) d\theta$

$$R' = \{K(\theta) > 1\} = \left\{ \frac{1}{\sigma} \exp\left(-\frac{1}{\sigma} |\theta - \theta_0|\right) > 1 \right\}$$

$$S(\theta) = \int_0^\infty \exp\left(-\frac{1}{2\sigma^2} (\theta - \theta_0)^2 - \frac{1}{\sigma} |\theta - \theta_0|\right) d\theta = \frac{1}{\sigma} \int_0^\infty \exp\left(-\frac{1}{2\sigma^2} (\theta - \theta_0)^2 - \frac{1}{\sigma} |\theta - \theta_0|\right) d\theta$$

$\exp\left(-\frac{1}{2\sigma^2} (\theta - \theta_0)^2 - \frac{1}{\sigma} |\theta - \theta_0|\right)$  由于  $g(\theta) \sim TV(\theta)$  (氏, 为连续型分布, 因此上式分子就是一般的边缘

## 8.3 假设检验与区间估计的 Bayes 方法

383

分布, 这在例 8.1.2 中已经给出, (8.1.6) 式的结果代入上式分子可得  $(\theta - \theta_0)^2$

$$\pi(\theta) = \frac{1}{\sigma} \exp\left(-\frac{1}{\sigma} |\theta - \theta_0|\right) = \frac{1}{\sigma} \exp\left(-\frac{1}{\sigma} |\theta - \theta_0|\right) = \frac{1}{\sigma} \exp\left(-\frac{1}{\sigma} |\theta - \theta_0|\right)$$

由于交叉比  $A:10(\%)$  为  $Z = \frac{\bar{X} - \theta_0}{\sigma/\sqrt{n}}$  的严增函数, 因此否定域可表示为

$$R^* = \{K(\theta) > 1\} = \left\{ \frac{1}{\sigma} \exp\left(-\frac{1}{\sigma} |\theta - \theta_0|\right) > 1 \right\}$$

## 8.3.2 Bayes 区间估计和 HPD 可信区间

设  $X \sim f(x, \theta)$ ,  $\theta \in \Theta$ ,  $\theta$  的先验分布为  $\pi(\theta)$ . 对于区间估计, Bayes 方法比经典方法简单直接, 不要求枢轴量以及枢轴量的分布. 因为  $\theta$  是随机变量, 有了后验分布  $\pi(\theta|x)$  以

后,就很容易算出它落在一个区间(或区域)的概率,从而得到区间估计.

定义8.3.2 若区域 $C(x)$ 满足条件

$$P\{0 \leq C(x) \leq 1 - \alpha\} = 1 - \alpha. \quad (8.3.24)$$

则称 $C(x)$ 为参数 $\theta$ 的一个水平为 $1 - \alpha$ 的Bayes置信域,或称为可信域(credible region).特别,若 $C(x) = [\hat{\theta}(x), \bar{\theta}(x)]$ ,则称 $[\hat{\theta}(x), \bar{\theta}(x)]$ 为水平为 $1 - \alpha$ 的Bayes置信区间,或可信区间;若 $C(x) = (-\infty, \bar{\theta}(x)]$ 和 $[\hat{\theta}(x), \infty)$ ,则称 $\bar{\theta}(x)$ 和 $\hat{\theta}(x)$ 分别为 $\theta$ 的水平为 $1 - \alpha$ 的Bayes置信上限和Bayes置信下限. ■

与参数置信域的定义类似,若 $T(\theta|x)$ 为连续型分布,则以上定义中的不等式可改为等式,以便于计算.另外,实用上大多考虑一维单参数的情形.

例8.3.6 设 $X_1, \dots, X_n$ 为独立同分布样本,且有 $X_i \sim N(\theta, \sigma^2)$ ,  $\sigma^2$ 已知.假设 $\theta$ 的先验分布为 $e^{-ir}(e) \sim N(\mu, \tau^2)$ ,求参数 $\theta$ 的Bayes置信区间.  
解 由例8.1.2的结果可知 $\theta \sim N(a, v^2)$ ,其中 $a$ 和 $v^2$ 如

$$a = \frac{\tau^2 \mu + \sigma^2 \sum_{i=1}^n x_i}{\tau^2 + n\sigma^2}, \quad v^2 = \frac{\tau^2 \sigma^2}{\tau^2 + n\sigma^2}$$

(无-氏)

384

## 第八章Bayes统计基础

(8.1.4)所示.因此可化为标准正态分布:由此可得

$$Q_a = \frac{a - \mu}{v}$$

从而即可反解得到

$$[\hat{\theta}(x), \bar{\theta}(x)] = [a - z_{\alpha/2} v, a + z_{\alpha/2} v], \quad (8.3.25)$$

其中, $z_{\alpha/2}$ 为标准正态分布的 $1 - \alpha/2$ 分位数.这是以后验期望为中心的一个区间,与经典结果不同.例如:若 $\mu = 10, \sigma^2 = 100, \tau^2 = 15$ ,并设 $n = 1, X_1 = 115$ ,则由(8.3.25)式可得水平为95%的可信区间为 $(94.08, 126.70)$ .而相应的经典置信区间为 $(95.4, 134.6)$ .但是,由(8.1.4)式可知,若 $\tau^2 \rightarrow \infty$ (即 $\theta$ 退化为广,均匀分布),则 $a = x, v =$

$\sigma/\sqrt{n}$ ,这时(8.3.25)与经典解一致. ■ 例8.3.7 设某镇每周火灾发生的次数服从Poisson分布 $P(A)$ ,今观测到:连续5周火灾发生的次数分别为 $0, 1, 1, 0, 0$ .假设样本独立同分布, $A$ 服从无信息先验分布 $\pi(A)$ 求平均次数 $A$ 的水平为90%的可信区间.又若观测值为 $0, 1, 1, 0, 1$ ,其结果如何?

解 对于本问题, $n = 5, (x_1, \dots, x_5) = (0, 1, 1, 0, 0)$ ,  
 $x^2$ .与例8.1.3类似, $A$ 的后验分布可表示为  $i=1$   
 $\pi(A|x) = c(x)\pi(A)f(x,A) = \frac{e^{-A} A^{\sum x_i}}{\sum_{j=0}^{\infty} \frac{e^{-A} A^j}{j!}} = \frac{e^{-A} A^{\sum x_i}}{e^{-A} \sum_{j=0}^{\infty} \frac{A^j}{j!}} = \frac{e^{-A} A^{\sum x_i}}{e^{-A} e^A} = e^{-A} A^{\sum x_i}$  即 $\pi(A|x)$ 服从 $\Gamma(\sum x_i + 1, 1)$ 分布,式中 $c = 1$ .因此由 $\Gamma$ 分布与 $\chi^2$ 分布的关系有  $10A$  由此可得  
 $P\{1/4 \leq 10A \leq 1/0.95\} = P\{0.711 \leq 10A \leq 9.488\} = 0.9$ .

因此, $A$ 的水平为90%的可信区间为 $(0.07, 0.95)$ .又若观测值为 $0, 5$   
 $1, 1, 0, 1$ ,则  $r = 2, \sum x_i = 3, 10A \sim \chi^2(6)$ ,  $A$ 的水平为90%的可信  
区间为 $(0.16, 1.26)$ .这些结果都是比较合理的.  $I$  与参数的置信域类似,满足条件  
(8.3.24)的可信域通常有很多,我们应该从其中寻找最优的.通常就是体积或区间长度最小的可信域,

在Bayes统计中,经常采用HPD可信域,其定义如下:

定义8.3.3 若区域 $C(x)$ 满足条件(8.3.24),并且存在 $W(\alpha) > 0$ ,使 $C(x)$ 可表示为

$$C(x) = \{x: \pi(x) \geq W(\alpha)\}. \quad (8.3.26)$$

### 8.3 假设检验与区间估计的Bayes方法

385

则称 $C(\theta)$ 为参数 $\theta$ 的水平为 $1 - \alpha$ 的最大后验密度(highest posterior density)可信域, 简称为水平为 $1 - \alpha$ 的HPD可信域. ■

HPD可信域有许多特点, 今摘要概述如下.

(1) 首先,  $C(\theta)$ 必然是有界的区域, 因为由(8.3.26)式可知,

$\int_{C(\theta)} \pi(\theta | x) d\theta = 1 - \alpha$ ; 若  $C(\theta)$  无界, 则积分  $\int_{C(\theta)} \pi(\theta | x) d\theta = 1$  必为无穷. 特别, HPD可信区间都是有限的区间.

(2) 由于HPD可信域集中了密度尽量大的点, 因而它应该有最小的体积或区间长度, 这可表述为以下定理:

定理8.3.2 在同等水平的可信域中, HPD可信域具有最小的体

积. 即若  $P\{\theta \in C(x) | X\} = P\{\theta \in C^*(x) | X\} = 1 - \alpha < 1$ , 其中  $C(x)$  为HPD可信域, 则必有  $V(C(x)) \leq V(C^*(x))$ .

证明 设在 $\theta$ 上, 对应于集合 $C(\theta)$ 与 $C^*(\theta)$ 的示性函数分别记为  $I_C(\theta)$ 与 $I_{C^*}(\theta)$ , 即  $I_C(\theta) = \begin{cases} 1 & \theta \in C \\ 0 & \theta \notin C \end{cases}$

$I_{C^*}(\theta) = \begin{cases} 1 & \theta \in C^* \\ 0 & \theta \notin C^* \end{cases}$ . 则要证  $\int I_C(\theta) \pi(\theta | x) d\theta \leq \int I_{C^*}(\theta) \pi(\theta | x) d\theta$

$C(x) \mid \int I_{C^*}(\theta) \pi(\theta | x) d\theta = 1 - \alpha$

$\pi(\theta | x) \geq \pi(\theta^* | x)$  并且水平条件  $P\{\theta \in C(x) | X\} = P\{\theta \in C^*(x) | X\} = 1 - \alpha$

等价于  $\int_{C(x)} \pi(\theta | x) d\theta = \int_{C^*(x)} \pi(\theta | x) d\theta = 1 - \alpha$ . 以下证明类  $J_0$  似于Neyman - Pearson基本引理的证明方法. 考虑函数  $A(\theta) = \int_{C(x)} \pi(\theta | x) d\theta - \int_{C^*(x)} \pi(\theta | x) d\theta$

及其积分:

$\int_{C(x)} \pi(\theta | x) d\theta - \int_{C^*(x)} \pi(\theta | x) d\theta = \int_{C(x)} \pi(\theta | x) d\theta - \int_{C^*(x)} \pi(\theta | x) d\theta$  (8.3.27)

则当  $\theta \in C(x)$  时,  $\pi(\theta | x) \geq \pi(\theta^* | x)$ ; 当  $\theta \in C^*(x)$  时,  $\pi(\theta | x) \leq \pi(\theta^* | x)$ . 因而  $A(\theta) \geq 0$ ; 而当  $\theta \in C(x)$  时,  $\pi(\theta | x) \geq \pi(\theta^* | x)$ ; 当  $\theta \in C^*(x)$  时,  $\pi(\theta | x) \leq \pi(\theta^* | x)$ . 也有  $A(\theta) \leq 0$ .

因此恒有积分  $\int A(\theta) \pi(\theta | x) d\theta = 0$ , 从而由(8.3.27)式可得

$\int_{C(x)} \pi(\theta | x) d\theta = \int_{C^*(x)} \pi(\theta | x) d\theta = 1 - \alpha$

(8.3.28) 由假设条件可知,  $\int_{C(x)} \pi(\theta | x) d\theta = \int_{C^*(x)} \pi(\theta | x) d\theta = 1 - \alpha$

$\pi(\theta | x) \geq \pi(\theta^* | x)$ , 因此由(8.3.28)式可得  $\int_{C(x)} \pi(\theta | x) d\theta = \int_{C^*(x)} \pi(\theta | x) d\theta = 1 - \alpha$ , 即

$\int_{C(x)} \pi(\theta | x) d\theta = \int_{C^*(x)} \pi(\theta | x) d\theta = 1 - \alpha$ , 因此有  $\int_{C(x)} \pi(\theta | x) d\theta = \int_{C^*(x)} \pi(\theta | x) d\theta = 1 - \alpha$

386

### 第八章 Bayes统计基础

(3) 若 $\theta$ 为一维单参数, 则 $\pi(\theta | x)$ 对应于平面上一条曲线, 可用水平线 $\pi(\theta | x) = c$ 把密度曲线截为上下两部分, 则上面部分对应的 $\theta$ 就是HPD可信区间, 见图8.3.1和8.3.2. 特别, 若 $\pi(\theta | x)$ 的密度曲

线为单峰的(这是很常见的情形), 并且关于某一个点 $M(\theta)$ 对称, 则此水平线截取的HPD可信区间必然关于 $M(\theta)$ 对称, 并具有形式  $[x - c(a), x + c(a)]$ , 其中 $c(a)$ 由(8.3.24)式决定. 上面例8.3.6的(8.3.25)式就是这种情形, 因而是HPD可信区间. 若密度曲线为单峰非对称的, 则HPD可信区间应满足

$\int_{C(x)} \pi(\theta | x) d\theta = 1 - \alpha$ ,  $\pi(\theta | x) = \pi(\theta^* | x)$ : (8.3.29)

例8.3.8 设 $X_1, \dots, X_n$ 为独立同分布样本, 且 $X_i \sim \text{BE}(1, \theta)$ , 求参

数 $\theta$ 的水平为 $1 - \alpha$ 的HPD可信区间: (1)  $\theta$ 的先验分布为冷分布  $\text{BE}(p, q)$ ; (2)  $\theta$ 的先验分布为均匀分布  $\pi(\theta) = 1$ .

解 (1) 由例8.1.1可知,  $\theta$ 的后验分布为  $\text{BE}(p+T, q+n-T)$ , 其中 $J; X$ , 因而后验密度可表示为  $\pi(\theta | x) = \frac{1}{B(p+T, q+n-T)} \theta^{p+T-1} (1-\theta)^{q+n-T-1}$

$77(4) X) = C(X)r + r_1(1 - \textcircled{2})''$

易见,  $77(\wedge | x)$  为  $[0, 1]$  区间上的一条曲线, 由  $77(\wedge | \%)$  的一阶导数可以看出, 对于不同的  $p, q, n, r$  的取值, 该曲线可能为以下4种情况

之一(见图 8. 3. 1、8. 3. 2 及 Lehmann 1985):

1];

ii)  $77(0 | \%)$  在  $[0, 1]$  上为没的减函数, 则 HPD可信区间为  $[0, a_2(x)]$ ;

8.3假设检验与区间估计的Bayes方法

387

图 8.3.2 单峰和单谷函数的HPD可信区间

iii)  $77(\wedge 1\%)$  在  $[0, 1]$  上先增后减, 则HPD可信区间为  $[a_1(x), a_2(x)]$ ;

iv) 在  $[0, 1]$  上先减后增, 则HPD可信区间为  $[0, 七(\text{幻})]$  和  $[a_2(x), 1]$ .

其中  $a_1(x)$  或  $a_2(x)$  由 (8.3.24) 式决定, 并且在 iii), iv) 两种情形,  $a_1(x)$  和  $a_2(x)$  同时还应满足  $77(f1 | \%) | x) = 77(a_2(x) | \%)$  (见 (8. 3. 29) 式).

(2) 对于  $0 \sim \text{尺}(0, 1)$  的情形, 即  $p = 1, g = 1$ , 这时  $77(6 | \%) = c(x)eT(i \text{ 没的HPD町信区间与 } 71, r \text{ 的取值有关. 若 } r = n, \text{ 则 } 7T(\wedge I X) = (n + 1)6 >''$ ,  $77(191\%)$  在  $[0, 1]$  上为 (9 的增函数, 其 HPD 可

信区间为  $[a_1(x), 1]$ . 而  $a_1(x)$  由

$77\blacksquare(\text{没} | \%) = a$  决定, 因此

0

$f1 | (x) = a''77$ . 若  $T=0$ , 则  $7T(\wedge I \%) = (n+1)(1-0Y$ ,  $7T(e \setminus X)$  在  $[0,$

门上为  $e$  的减函数, 其HPD可信区间为  $[0, a_2(\text{幻})]$ ,  $a_2(x)$  由

$\text{广}2(*) J0$

$7T, (0 \setminus x) = c(x)n^{-1}(1 \text{ 该式表明, 当 } 0 < - \text{时, } \setminus 71/ n$

$7T(\text{汐}1\%)$  为没的增函数; 当  $0 > - \text{时, } 7T(e \setminus x)$  为 0 的减函数; 而当  $e = - nn$

时,  $n(0 \setminus x)$  达到其峰值. 这就是上述第 iii) 种情形, 因此HPD可信区 间为  $[a_1(x),$

$a_2(x)]$ , 其中  $a_1(x)$  或 七(4) 由 (8. 3. 29) 式决定. |

例8.3.9 设  $\text{弋}, \dots, X$ . 为独立同分布样本, 且有  $\sim 7V(M, a_2)$ .

假设  $(/Z, a)$  服从无信息先验分布  $77(/1, cr) CC(T_1, \text{求参数})11$  和  $a$  的水平 为  $1 - a$  的HPD可信区间.

$77(0 \setminus x) dG \sim -a$  决定, 因此  $a_2(x) = 1 - a \sim$ . 若  $0 < * < n$ , 则

解 由假设可知,  $(/u, a)$  的后验分布为

388

第八章Bayes统计基础

$| \%) = c(x)a \exp$

(8.3.30). 首先求  $M$  的边缘分布, 这时  $77(fJL | X) = 77(/X, 67 | \%) d < T$ . 记

$S_2 = Z - x)^2$ , 并令  $y = cr_2$ , 则  $7r(/z, o - | x) dcr$  可表示为  $i=1$

$-3/2 \int j - x)^2 + y ay.$

$f^{\circ\circ}$

$7r(/z | x) = \int 77(/x, a | \%) dtr = c(x)[n(/i - x)^2 + S_2]$

该式町进一步化简, 令

$S/ \text{扣} - 1$

$)da = c(x)y(n+, )/2 \exp$

由此根据厂函数的积分可得

$7T(/x, (r | a;$

其中  $a = SX/n^T$ , 则有  $n(M - \%)^2 = z^2 S_2 / (n - 1)$ , 该式代人以上  $7r(/x | \%)$ , 并把  $S_2$  纳入  $c(\text{幻}$

可得

$$2^{-n/2} \pi^{-1/2} = C(X) I^{1/2} \quad (8.3.31)$$

这说明, 给定  $X = x$ ,  $t = (pL - x)/a$  服从自由度为  $n-1$  的  $t$  分布(见第一章). 因此有

$$\int_{-\infty}^{\infty} \pi(t|x) dt = 1 - \alpha.$$

易见  $\pi(t|x)$  的分布关于  $Z=0$  对称,  $\pi(t|x)$  的分布关于  $A=x$  对称, 因此根据前面第(3)点的论证以及上式可知,  $M$  的水平为  $1-\alpha$  的 HPD 可信区间为

$$[M, \pi] =$$

这一结果与经典方法所得到的结果在形式上完全一致, 但是, 基本假设与

推导过程完全不一样. 在(8.3.31)式中, 给定  $X = x$ ,  $t$  是随机变量, 而在第七章的经典方法中,  $m$  为参数,  $I$  是随机变量.

再求  $f$  的边缘分布, 这时  $\pi(a|\%) = \int \pi(M, a|\%) dM$ . 在  $J=0$

(8.3.30) 式中,  $\pi(t|z, cr|x)$  关于  $m$  为正态分布, 其均值为  $0$ , 方差为  $\sigma^2/n$ , 因此有

$$\int_{-\infty}^{\infty} \pi(t|z, cr|x) dt =$$

$$2^{-n/2} \pi^{-1/2}$$

$$(8.3.31)$$

### 8.3 假设检验与区间估计的 Bayes 方法

389

$$7F(z, tr|\%) d/i = c(x) \exp(-x^2/2)$$

(8.3.32) 这时  $\pi(cr|x)$  作为  $a$  的函数, 是一个单峰函数. 因为若令  $Z(a|\%) = \log$

$$7r(cr|\%), \text{ 则有 } l\{a|x\} = \log c(\%) - n \log a - S^2/(2a^2), \quad Z_z(a|x) =$$

$$na^{-3} [-a^2]. \text{ 这说明, } Z(a|\%) \text{ 及 } \pi(a|\%) \text{ 为先增后减的单峰函数, 且}$$

在  $a^2 = -1$  时达到峰值. 这属于例 8.3.8 的第 iii) 种情况,  $cr$  的 HPD 可信

区间可表示为  $[R(\%), a^2(x)]$ , 其中  $\hat{a}(x)$  和  $a^2(x)$  满足(8.3.29)式, 显然, 其求解过程比较复杂.

我们亦可求的一般的 Bayes 置信区间. 事实上, 可以证明, 龙给

$n$  分布. 由(8.3.32)式可

定时,  $W = 2(A-I)^2 A t^2$  的条件分布为  $\chi^2_{1,1}$

知, 后验密度  $\pi(a|x)$  可表示为  $7F(a|x) = c(x') w^{n/2} e^{-1/2 w}$ , 若记  $M$  的密度为

$$(p(w|I)) \text{ 则有}$$

$$(p(w|x) = 7r(\chi^2_1|x) a w = c(x) w^{n/2} e^{-1/2 w} = c(x) w^{(n-3)/2} e^{-w/2}.$$

$$n$$

这说明,  $X$  给定时,  $w = I(X - x)^2/a^2 = S^2/a^2$  服从自由度为  $z-1$  的  $\chi^2_1$

$Y$  分布(见第一章). 这与第七章例 7.1.2 的枢轴量  $G(X, a^2)$  十分类似, 因此有

$-a^2$ . 因此可得  $a^2$  的 Bayes 置信区间为  $[a^2(x), a^2(x)]$ , 其中

这一结果与经典方法所得到的结果在形式上完全一致, 但是, 基本假设

与推导过程完全不一样. 在本例的  $\pi(x) = S^2/a^2$  中, 给  $i=1$

定  $X = x$ ,  $0$  是随机变量; 而在例 7.1.2 中,  $a$  为参数,  $X$  是随机变量. 另外, 以上 Bayes 置信区间并不是 HPD 可信区间, 因为它们不满足(8.3.29)式.

(4) 在文献中, HPD 可信域还有另一种定义(例如, 可参见茆诗松, 1999), 我们就权且称之为定义 8.3.3', 其定义如下: 若区域  $C'$  幻

390

## 第八章 Bayes 统计基础

$C_0$  满足条件(8.3.24), 并且对任何  $\hat{e}C(x)$  及  $\theta$ 、宅有  $7F(\theta|\%)$

$\pi(\theta|\%)$ , 则称  $C_f(x)$  为参数  $\theta$  的水平为  $1-\alpha$  的 HPD 可信域. 这两

种定义虽然不完全等价, 但是差别很小. 易见, 若  $C_h$  满足定义 8.3.3, 则必满足定义



8.3.3', 因为由(8.3.26)式可知, 若  $0 \in C(x) \cap \emptyset$ . 则有  $77(6 \text{ I } \emptyset) < k(a)^{7r(0 \setminus x)}$ . 反之, 若  $C'(x)$  满足定义 8.3.3', 则可记  $W(x)$  时  $77(\text{叫幻的下确界为 } Wa)$ , 对此, 可根据(8.3.26)式定义相应的  $C(\%)$ . 则对任何  $0 \in C'(x)$  有  $77(0 \mid \text{幻多 } \wedge(a))$ , 因而  $0 \in C(x)$ , 即  $C(x) \subset C'(x)$ . 事实上,  $C_f(x)$  与  $C(\text{幻})$  的差别主要在集合  $= \setminus 6: 7r(0 \setminus x) = A: (a) \mid$  上, 因为若  $77(\wedge \mid x) < k(a)$ , 则显然  $0 \in C(x)$ ,  $0 \in C_f(x)$ ; 若  $77(0 \mid \%) > k(a)$ , 则有  $0 \notin C(x)$ ,  $0 \notin C_f(x)$ . 在  $R^\circ = \{e: 7T(e \mid x) - k(a) \mid\}$  上,  $C \setminus x$  与  $C(\text{幻})$  可能有差别, 见图 8.3.3 左, 其中  $C(\text{幻}) = [aU], 6(x)]$ , 而  $C'\{x\}$  可以具有  $C_r(x) = [af(x), br(x)]$  的形式. 但是若, 为零测集, 则  $C'(x)$  与  $C(x)$  没有差别, 见图 8.3.3 右. 所以在实际应用上, 这两种定义差别不大. 另外, 在实际应用上也不一定非要求 HPD 可信域, 大多数情形下求解水平为 90% 或 95% 的 Bayes 置信区间即可.

图 8.3.3 HPD 可信区间的两种定义

#### 习题八

1. 设  $x_1, \dots, x_n$  为  $i.i.d.$  样本, 在给定  $x$  的分布和参数先验分布的情况下, 求相应的后验分布:

(1)  $X_i \sim E(a)$ , 并设  $\theta$  的先验分布为指数分布  $E(a)$ , 但  $a$  已知, 求的后验分布;

(2)  $X_j \sim f(A; f)$ ,  $p$  已知,  $A \sim r(a, p)$ , 求  $A$  的后验分布;

#### 习题八

391

(3)  $x_i \sim f(x_i | \theta) = 2x_i^{2\theta-1}$ , 并设  $\theta \sim e^{-\theta^2}$ , 求  $\theta$  的后验分布.

2. 证明下列先验分布是共轭先验分布:

(1) 设  $x_1, \dots, x_n$  为  $i.i.d.$  样本,  $X_i$  服从两点分布  $b(1, \theta)$ , 其中  $\theta$  已知, 样本容量  $n$  未知,  $n$  的先验分布为  $P(A)$ ;

(2) 设  $x_1, \dots, x_n$  为  $i.i.d.$  样本,  $X_i$  服从指数分布  $E(1/a) = r(i < r, i)$ ,  $a$  的先验分布为逆  $r$  分布  $r^-(a, p)$ ;

(3) 设  $x_1, \dots, x_n$  为  $i.i.d.$  样本, 服从负二项分布  $NB(\theta, r)$ ,  $r$  已知,  $\theta$  未知; 其先验分布为  $BE(\gamma, g)$ ;

(4) 设  $x_1, \dots, x_n$  为  $i.i.d.$  样本,  $x_i \sim \Gamma(\theta, 1)$  并设  $\theta \sim \Gamma(M, \gamma)$ , 其中  $M$  和  $\gamma$  已知.

3. 设  $x_1, \dots, x_n$  为  $i.i.d.$  样本,  $X_i$  服从幂函数分布  $A \sim PF(c, \gamma)$ , 即  $f(x) = cx^{\gamma-1}e^{-cx}$  ( $c > 0; \gamma > 0$ ). 证明: 若  $c$  已知, 则  $\gamma$  的共轭先验分布为 Pareto 分布; 若  $\gamma$  已知, 则  $c$  的共轭先验分布为  $\Gamma$  分布.

4. 设  $x_1, \dots, x_n$  为样本,  $X_i \sim N(\mu, \sigma^2)$  且  $\mu$  和  $\sigma^2$  都未知. 又设  $\sigma^2$  的先验分布为  $77(6) \sim r(a, p)$ ; 在  $a$  给定条件下,  $\mu$  的先验分布为  $77(1 \mid a) \sim$

(1) 求  $(\mu, \sigma^2)$  的联合后验分布;

(2) 求  $\sigma^2$  的后验边缘分布, 以及给定  $\sigma^2$  条件下  $\mu$  的后验边缘分布. 5. 设  $x_1, \dots, x_n$  为  $i.i.d.$  样本,  $x_i \sim V(\gamma, 1, r^2)$ , 其中  $\gamma$  和  $r^2 = a^2$  未知, 又设  $\theta = (p, \gamma)$  的先验密度  $\pi(\theta) = \gamma^{-1}$ . 证明:

(1)  $\theta = (\gamma, 1)$  的后验密度函数可表示为  $77(\theta \mid X) = 77(\gamma \mid X, 1)$ .

$77(1 \mid X)$ , 其中  $\gamma \sim \Gamma(\gamma, 1)$ ,  $t \sim \Gamma(t, 1)$ ,  $t \sim \Gamma(t, 1)$  服从逆  $r$  分布  $\Gamma_i(s, n')$ , 其中  $n' = \gamma, 3 = \gamma - 2 (X_t - X)^2$ ;

(2)  $S$  的后验边缘分布服从逆厂分布  $f^{-1}(s, nr)$ ; .

(3) 的后验边缘密度是其中  $\theta = - (X: -X)^2$ ,

$f(\cdot)$  是自由度为  $n-1$  的分布的密度函数.

6. 设  $X = (\theta_1, \dots, \theta_k)^T$  服从多项分布  $MN(n, \theta)$  即  $f(x; \theta) =$

$\frac{n!}{x_1! \dots x_k!} \theta_1^{x_1} \dots \theta_k^{x_k}$ , 其中  $\theta_1, \dots, \theta_k$  为参数,  $\theta_1 = 1, \theta_2 = \dots = \theta_k = 0$  的先验分布为 Dirichlet 分布, 即为

392

$f(\theta) = \frac{\Gamma(n+1)}{\Gamma(a_1) \dots \Gamma(a_k)} \theta_1^{a_1-1} \dots \theta_k^{a_k-1}$

$t=1$

$k$

其中  $a_i > 0, i=1, \dots, k$ ;

$i=1$

## 第八章 Bayes 统计基础

$\theta_1^{a_1} \dots \theta_k^{a_k}$ .

$a = (a_1, \dots, a_k)^T$ , 并把这一

分布记为  $Z(a)$ . 证明: 19 的后验分布服从 Dirichlet 分布  $D(a+x)$ .

7. 设  $X = (X_1, \dots, X_n)^T$  服从指数族分布  $f(x; \theta) = \frac{1}{Z(\theta)} \exp(\eta^T x - \psi(\theta))$

$f(x) = \frac{1}{Z(\theta)} \exp(\eta^T x - \psi(\theta))$

$f_k$   $i=1$

验分布;

ii) 把以上结果用于双参数的分布和厂分布.

8. 设  $X_1, \dots, X_n$  为 i. i. d. 样本, 由 Jeffreys 准则确定以下参数的先验分布:

(1)  $X_i \sim r(A; p)$ , 但  $p$  已知;  $A$  未知;

(2)  $X_i \sim E(1/\theta)$  为寿命分布, 但只观测到前  $r$  个寿命的值:  $y =$

$X(1), \dots, X(r)$ ;

(3) 但  $r$  已知,  $\theta$  未知;

(4)  $X = (X_1, \dots, X_k)^T$  服从第 6 题的多项分布  $MN(n, \theta)$ .

9. 根据 Jeffreys 准则证明:

(1) 若  $X = (X_1, \dots, X_k)^T$  服从位置参数分布则  $f(\theta) \propto 1$ ;

(2) 若  $X = (X_1, \dots, X_k)^T$  服从尺度参数分布  $a \sim n/(\theta/a)$ , 则  $W(\theta) \propto 1/a$ ;

(3) 若  $X = (X_1, \dots, X_k)^T$  服从位置尺度参数分布族  $a \sim y((x < M)/\theta)$  贝  $I(\theta) \propto 1/\theta^2$ .

W. 设  $X_1, \dots, X_n$  为 i. i. d. 样本, 在均方损失下, 求相应参数的 Bayes 估计:

(1)  $X_i \sim \text{Pareto}(\alpha, \theta)$ ; (2)  $X_i \sim \text{Gamma}(\alpha, \theta)$ ; (3)  $X_i \sim \text{Gamma}(\alpha, \theta)$ ,  $\alpha$  已知,  $\theta$  服从逆厂分布  $r^{-1}(\alpha, p)$ ;

(4)  $X_i \sim \text{Gamma}(\alpha, \theta)$ ,  $\alpha$  服从厂分布  $r(\alpha, p)$ . 11. 设  $X_1, \dots, X_n$  为 i. i. d. 样本,  $X_i \sim$

$X_i \sim \text{Gamma}(\alpha, \theta)$ ,  $\alpha$  服从厂分布  $r(\alpha, p)$ .

(1) 若  $\theta = H$   $A$  的先验分布为  $f(\theta) = e^{-\theta}$ ,  $\theta > 0$ , 在均方损失下,

i) 证明:  $t(\theta) = c(\alpha \ln \theta) \exp$ .

$a \sim \text{Gamma}(\alpha, \theta)$ , 为  $\theta$  的共扼先

## 习题八

393

求  $u$  的 Bayes 估计;

(2) 若  $y$  已知,  $\theta \sim r^{-1}(\alpha, p)$ , 求  $\theta$  的后验期望估计;

(3) 条件同 (2), 在损失函数  $L(e, d)$  下, 求  $\theta$  的 Bayes

估计.

12. 设  $x_1, \dots, x_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ ,  $\mu$  和  $\sigma^2$  都未知.

设  $S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$  的先验分布为  $\pi(\sigma^2) \sim \tau(a, p)$ ; 在  $a^2$  给定条件下,  $\sigma^2$  的先

验分布为  $\tau(\sigma^2 | a^2) \sim \tau(a, p)$ . 在均方损失下, 求  $\sigma^2$  的 Bayes 估计.

13. 设  $x_1, \dots, x_n$  为 i.i.d. 样本,  $J_i \sim U(1)$ ,  $A \sim \tau(a, p)$ . (1) 在均方损失下, 求  $g(A) = A; 0 \leq A \leq 1$  的 Bayes 估计;

(2) 在损失函数  $L(A, J) = \sum_{i=1}^n (A - J_i)^2$  下, 求  $A$  的 Bayes 估计及其后

$A$

验风险和 Bayes 风险.

14. 设  $x_1, \dots, x_n$  为 i.i.d. 样本,  $x_i \sim b(i, P)$ ,  $0 < P < 1$ , 的先验

分布  $\pi(P) \sim \tau(0, 1)$ . 在损失函数  $L(d, p) = P(1 - P)$  下, 求  $P$  的 Bayes

估计, 并求其 Bayes 风险.

15. 设  $x_1, \dots, x_n$  为 i.i.d. 样本,  $X_i$  服从均匀分布

$U(0, \theta)$ . 在无信息先验分布和均方损失下, 求  $\theta$  的 Bayes 估计.

16. 设  $x_1, \dots, x_n$  为 i.i.d. 样本,  $X_i \sim N(\mu, \sigma^2)$ . 在无信息先验

分布下, 求相应参数的后验期望估计: i)  $\mu = 1$ ,  $\sigma^2$  未知; ii)  $\mu = 0$ ,  $\sigma^2$  未知; iii)  $\mu = \sigma^2$  都未知.

17. 设  $x_1, \dots, x_n$  为 i.i.d. 样本,  $X_i \sim U(0, 1)$  表示寿命分布,  $E(X_i)$

$= 0.5$  为平均寿命. 若只观测到前  $r$  个寿命值:  $K = (x_1, \dots, x_r)$ , 取损失函数为

$L(e, d) = [(J - \theta)/\theta]^2$ , 求  $\theta$  的 Bayes 估计. 若  $\theta$  的先验分布为: (1) 逆 Gamma 分布  $\tau_1(a, p)$ ;

(2)  $\tau_2(\theta) \propto \theta^{-1}$  (由 Jeffreys 准则确定的先验分布).

18. 设  $x_1, \dots, x_n$  为 i.i.d. 样本, 损失函数为均方损失, 证明: 以下参数的 Bayes 估计是渐近无偏的和相合的:

(1)  $X_i \sim b(k, e)$ ,  $e \in (0, 1)$ ,  $e$  的先验分布  $\pi(e) \sim BE(p, q)$ ; (2)  $X_i \sim U(0, \theta)$ ,  $\theta > 0$ ,  $\theta$  的先验分布  $\pi(\theta) \sim PR(a, b)$   $- a < \theta < b$ .

=

21. 设  $(\mu, \sigma^2)$ ,  $\mu > 0$  未知,  $a^2$  已知,  $f_j$  的先验分布为  $\pi_j(\mu)$  证明: 在均方损失下,  $M$  的 Bayes 估计为:  $\hat{\mu} = \bar{x} +$

其中  $\phi$  和  $\psi$  分别是标准正态的密度函数和分

394

## 第八章 Bayes 统计基础

$I \propto \theta^{e_0} I_1$ , 且  $e_0 < x(n)$ ;

(3)  $X_i \sim P(A)$ ,  $A$  的先验分布  $\pi(A) \sim \tau(a, p)$ .

19. 设  $x_1, \dots, x_n$  为 i.i.d. 样本,  $X_i \sim b(1, p)$ ,  $0 < p < 1$ .

(1) 当  $p \sim BE(a, 1/3)$  时, 在均方损失下, 求  $g(p) = P(1 - p)$  的 Bayes 估计, 并讨论其渐近无偏性和相合性;

(2) 设  $p$  服从广义先验分布  $\pi(p) = p^a (1 - p)^b$ ,  $0 < p < 1$ , 在均方误差损失下, 求  $P$  和  $g(p)$  的 Bayes 估计, 并讨论其渐近无偏性和相合性.

20. 设  $x_1, \dots, x_n$  为  $r$  次样本,  $X_i$  服从负二项分布  $NB(p, r)$ ,  $r > 0$  已知; 又设

(1) 在均方损失下, 求  $p$  和  $p_1$  的 Bayes 估计; (2) 证明 (1) 中得到的 Bayes 估计是相合的. 布函数.

22. 设  $X_i$  服从几何分布  $G(p)$ ,  $0 < p < 1$ ,  $p \sim \tau(p)$  当损失函数为  $L(p, d) = (d - p)^2/p$  时证明:  $p$  的 Bayes 估计为

$$[(1-p)x + p]^{-1} dp$$

$$P = 1 - \int_0^1 [(1-p)x + p]^{-1} dx, \quad 0 < p < 1, \quad 1, 2, \dots$$

$$(1-p)^{-1} \int_0^1 [(1-p)x + p]^{-1} dx$$

23. 设  $X_1, \dots, X_n$  为 i. i. d. 样本,  $X_i$  的先验分布为逆 Gamma 分布  $e^{-r/\lambda}(a, p)$ , 证明: 在均方损失下,  $\theta$  的 Bayes 估计为:

$$\hat{\theta} = \frac{a + \sum_{i=1}^n X_i}{n+2a}$$

其中  $n = \sum_{i=1}^n 1_{\{X_i > 0\}}$  为样本最大值 (提示: 利用 Gamma 分布与 Gamma 分布的关系).

24. (1) 设  $Y_i (X_i)$  是  $g_i(\theta)$  在均方损失下的 Bayes 估计,  $Z = 1, \dots, p$ .  $pP$

证明:  $q_{SjX}$  是  $2 \text{ Cigi}(\theta)$  在均方损失下的 Bayes 估计;  $i=1, \dots, l$

(2) 设  $X_1, \dots, X_n$  为 i. i. d. 样本,  $X_i \sim PR(a, \theta_0) - a^{\theta_0 x}; \{a^{\theta_0}\} \in \mathbb{R}^+$

多氏  $f$ , 其中  $\theta_0$  已知,  $a \sim r(a, p)$ , 在均方损失下求  $g(a) = a + a^2$  的 Bayes 估计.

\*25. (1) 设  $\gamma(\theta)$  为正常的先验分布,  $3(\theta)$  是  $g(\theta)$  在均方误差损失

习题八

395

下的 Bayes 估计, 估计的偏差是  $6(\theta)$ , 即  $E[5(X) - 1] = g(6) + b(6)$ . 假设  $Eg_2(\theta)$

$< \infty$ , 则  $5(\theta)$  的 Bayes 风险  $Rv(d)$  可表示为

$$Rn(S) = E[5(X) - g(6)]^2 = \int \hat{g}(\theta) b(\theta) n(\theta) d\theta, \quad \text{其中 } \hat{g}(\theta) \text{ 表示对 } (X, \theta) \text{ 的联合分布求期望 (提示: 参考定理 8.2.3 的证明).}$$

(2) 设  $X_1, \dots, X_n$  为 i. i. d. 样本,  $X_i$  服从 Poisson 分布  $P(A)$ , 设入

的先验分布为  $\gamma(A) \sim r(a, p)$ , 通过直接计算验证以上公式.

26. 设  $X_1, \dots, X_n$  为 i. i. d. 样本,  $X_i \sim P(A)$ ,  $A \sim r(a, p)$ . 设  $K_1, \dots, K_n$  为 i. i. d. 样本

且与  $X_i$  独立,  $Y_i \sim P(b_i)$ ,  $b_i \sim r(b, q)$ , 且

与  $A$  独立. 在均方损失下, 求  $p = X/fJL$  的 Bayes 估计.

27. 设  $X_1, X_2 \sim r(\theta, \lambda)$ , 且  $A$  与  $X_2$  独立;  $\theta \sim$

$KD$ , 且  $\lambda$  与  $\theta$  独立; 又设  $\lambda$  是  $\theta$ . 在均方损失下基于  $X_1, X_2$  的 Bayes 估计,  $i=1, 2$ . 证明:

(1)  $S_1$  是  $A$  在均方损失下基于  $X_2$  的 Bayes 估计; (2)  $S_2$  是  $\lambda$  在均方损

失下基于  $(A, X_2)$  的 Bayes 估计.

28. 设  $\lambda \sim r(\lambda, \theta)$ ,  $\theta \sim r(\theta, \lambda)$ ,  $Z=1, 2$ ;  $e^{-7r(\theta)} = \int_0^1 f(\theta) d\theta$ , 其中

$f(\theta)$  是  $\theta$  上的密度函数, 对任意给定的  $\theta$ ,

是  $A$  上的密度函数, 若在  $\theta$  给定条件下,  $h(\theta) = g(e^{i\theta})$  在均方损失下的 Bayes 估计是

$8(X_2)$ , 则  $g(\theta)$  在均方损失下的

Bayes 估计是  $8(X)$ , 且满足  $5(X) = \int_0^1 8(X_2) n(\theta | X) d\theta$ , 其中  $J\theta$

$f(\theta | X)$  是  $\theta$  的后验分布密度函数. 把以上结果用于本章习题 4, 求

$(\hat{\theta} \text{ 及 } g(\theta, S)) = \int_0^1 \lambda^2$  在均方误差下的 Bayes 估计. 29. 在均方损失函数  $L(\lambda d; We)$  下

证明:  $g(\theta)$  的 Bayes

解的后验风险为  $\text{Var}(g(\theta) | X)$ , 其中  $\text{Var}(\cdot | X)$  表示关于后验分布  $r(e | X)$  的方差.

30. 设  $X_1, \dots, X_n$  为 i. i. d. 样本, 求相应参数的后验极大似然估计:

(1)  $X_i \sim f(x | \theta) = 2\theta^2 x^{-3}, \theta > 0, \theta$  服从 Pareto 分布  $P_2(a, \lambda)$ ;

(2)  $X_i \sim A^2 x e^{-Ax}, x > 0, A$  服从 Gamma 分布  $r(a, p)$ ;

(3)  $X_i \sim r(l/\theta; i)$ ,  $p$  已知,  $a$  服从逆 Gamma 分布  $r^-(a, p)$ ;

(4)  $V(x_i; c) = cx; \sum_{j=1}^n X_j < 1 (c > 0)$ .  $c$  的先验分布为  $r$  分布  $r(a, p)$ .

31. 设  $X_1, \dots, X_n$  为 i. i. d. 样本,  $X_i \sim NC(\mu, \sigma^2)$ ,  $\mu, \sigma^2$  已知

396

第八章 Bayes 统计基础

(1) 设  $c$  的先验分布求  $a$  的后验极大似然估计;

估计;

(2) 记  $5 = c^2$ , 设  $5$  的先验分布  $\pi(5)$

求  $6$  的后验极大似然 (3) 设  $S$  的先验分布  $\pi(3) = \pi_l(a, P)$ , 求  $5$  的后验极大似然

估计. 32. 设  $X_1, \dots, X_n$  为独立同分布样本, 记  $S = \sum_{i=1}^n X_i$ , 又设  $\pi(\theta)$  求  $\theta$  的后验极大似然估计.

33. 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $X_i$  服从幂函数分布  $\pi(x; \theta) = c x^{\theta-1} e^{-\theta x}$  ( $c > 0; \theta > 0$ ). 求相应参数的后验极大似然估计: i) 若  $\theta$  已知,  $c$  的先验分布为  $\pi(c)$ ; ii) 若  $c$  已知,  $\theta$  的先验分布为 Pareto 分布  $\pi(\theta)$ .

34. 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $X_i \sim \text{Exp}(\lambda)$ . 若只观测到前

$r$  个寿命值:  $y_1 = x(1), \dots, y_r = x(r)$ . 在以下三种无信息先验分布下, 求相应参数的后验极大似然估计: i)  $\lambda = 1$ , 从未知; ii)  $\lambda = \theta$ ,  $\theta$  未知; iii)  $\theta = (c, \lambda)$  都未知.

• 35. 设  $X \sim f(x; \theta)$ ,  $Q$  氏  $\theta$  的先验分布为  $K(\theta)$ , 对  $g(\theta)$  的估计, 取损失函数为  $L(\theta, d) = (g(\theta) - d)^2$ ; 求  $g(\theta)$  的 Bayes 估计.

36. 设  $X_1, \dots, X_n$  为独立同分布样本, 且有  $X_i \sim N(\mu, \sigma^2)$ , 损失函数为  $L(\theta, d)$  证明:

$\hat{\mu}$  为  $\mu$  的 Minimax 估计, 其中  $T = \sum_{i=1}^n X_i$ .

为  $\mu$ .

样本  $X_1, \dots, X_n$  其中  $M$ ,  $\lambda$  都未

37. 设  $X_1, \dots, X_n$  为  $i.i.d.$  样本,  $X_i \sim \text{Exp}(\lambda)$

均匀分

布). 证明: 在均方损失下,  $\hat{\lambda} = (1/n + \sum_{i=1}^n X_i)/2$  为  $\lambda$  的 Minimax 估计 (提示: 取一系列先验分布为  $(1/k, \text{在 } [0, k] \text{ 上的均匀分布})$ ).

38. 设  $X_1, \dots, X_n$  为独立同分布样本,  $X_i \sim r(y, l)$ . 取损失为  $L(\theta, d) = [(d - \theta)/\theta]^2$

证明: 在此损失函数下, 证明:  $T/(n + 1)$  为  $\theta$  的

Minimax 估计, 其中  $r = \text{布: } r \sim l(k^{-1} l^k - 1)$ .

$i = \backslash$

$X_i$  (提示: 取一系列先验分布为逆 Gamma 分

$\theta$

#### 习题八

397

39.  $X_1, \dots, X_n$  为独立同分布样本, 求以下单边假设检验问题的 Bayes 解:  $H_0: \theta \leq \theta_0$  vs  $H_1: \theta > \theta_0$ . (1)  $X_i \sim N(\theta, \sigma^2)$ , 但  $\sigma^2$  已知,  $\theta$  服从无信息先验分布  $\pi(\theta) \propto 1$ ; (2)  $X_i$

$\sim \text{Poisson}(\theta)$  为 Poisson 分布,  $\theta$  的先验分布  $\pi(\theta) \sim r(a, p)$  为  $\Gamma$  分布.

40. 假设  $X$  服从二项分布  $B(5, \theta)$ ,  $\theta$  的先验分布为  $BE(1, 9)$ , 若观察值为  $x = 0$ ,  $H_0: \theta \leq 0.1$ , 求检验问题的后验概率和交比.

41. 假设  $X$  服从正态分布  $N(\mu, \sigma^2)$ , 参数都未知,  $\mu, \sigma^2$  服从无信息先验分布

$\pi(\mu, \sigma^2) \propto \sigma^{-2}$ , 其“d. 观察值为  $(1.2, 1.6, 1.3, 1.4, 1.4)$ ,  $\hat{\theta}: \theta > 1$ , 求检验问题的后验概率和交比.

42. 假设某公共汽车站的候车时间  $X$  服从均匀分布  $U(0, \theta)$ , 近日观察到的候车时间分别为 10, 3, 2, 5, 14 分钟. 考虑假设检验问题  $H_0: \theta \leq 15$  vs  $H_1: \theta > 15$ , 假设  $\theta$  的先验分布为 Pareto 分布  $\pi(\theta)$ , 求

检验问题设的后验概率和交比.

6. 的先验分布为  $\pi(\theta) \propto \theta^{-2}$ . 对于假设检验问

$H_0: \theta \leq \theta_0$  vs  $H_1: \theta > \theta_0$ , 取损失函数为  $C_1$ ,  $C_2$  且  $g(x) = 1$ ,

$L(e, g(x)) = C_1$ ,  $C_2$ ,  $\theta \leq \theta_0$  且  $g(x) = 0$ ,

$0, 0 \in \text{且 } 3(x) = 0 \text{ 或 } 5(x) = 1.$

证明: 以上假设检验问题Bayes解的否定域满足  $\int_0^1 \pi(x) dx > C / (C_0 + C_1)$ , 其中  $\pi(x) = \int_0^1 \pi(\theta|x) d\theta$  (提示: 参考定理8.3.1和引理8.3.1).

44. 对于假设检验问题(8.3.19), 其先验分布如(8.3.20)所示. 证明:

. 题付0:

:

43没

(1)

爪1(4) .

$= [1 + 10(\theta)]^*$ ;

(2) 叫1德K,

其中  $\pi(\theta)$  为在  $\theta = 1$  上的最大值点.

(3) 对于例8.3.5的检验ii), 记  $Z = \sqrt{n}(\bar{X} - \theta)/\sigma$ , 证明:  $\phi(x) = [1 + \exp(-x^2/2)]^{-1}$ .

45-设  $A_1, \dots, A_n$  为 i. i. d. 样本,  $X_i$  服从 Poisson 分布  $P(\theta)$ ,  $\theta$  的先验分布  $\pi(\theta)$  为  $r$  分布  $r(a, \lambda)$ , 求  $\theta$  的一个水平为  $1 - \alpha$  的 Bayes 置

398 第八章 Bayes 统计基础

信区间.

46. 设  $X_1, \dots, X_n$  为 i. i. d. 样本,

(1) 当  $\theta$  已知, 且  $\pi(\theta) = 1$  时, 求  $\theta$  的水平为  $1 - \alpha$  的 Bayes 区间估计;

(2) 当  $\theta$  未知,  $\pi(\theta) = \theta^2$  时, 求  $\theta$  的水平为  $1 - \alpha$  的 Bayes 区间估计;

(3) 记  $\theta = \tau^2$ ,  $\theta \in (0, 1)$ , 设  $\theta$  的先验分布  $\pi(\theta) = 1$ , 求  $\theta$  的水平为  $1 - \alpha$  的 Bayes 区间估计.

47. 车间生产出大批零件, 要检测其次品率. 今从中抽出6件, 假

设次品数  $X$  服从二项分布  $B(6, \theta)$ ,  $\theta$  的先验分布为  $B(\lambda, 10)$ , 若观察值为  $x = 0$ , 求的水平为95%的HPD可信区间.

48. 用仪器观测某星体的质量  $\theta$ , 其5次 i. i. d. 观察值为 1.2, 1.6, 1.3, 1.4, 1.4 质量单位. 假设观察值  $X$  服从正态分布  $N(\theta, \sigma^2)$ , 参数  $\theta, \sigma^2$  都未知, 且  $\theta, \sigma^2$  服从无信息先验分布  $\pi(\theta, \sigma^2) \propto 1/\sigma^2$ . 求  $\theta$  的水平为

90%的HPD可信区间.

参考文献

1. 陈希孺. 数理统计引论. 北京: 科学出版社, 1981

2. 陈希孺. 高等数理统计. 合肥: 中国科技大学出版社, 1999

3. 陈家鼎, 孙山泽, 李东风. 数理统计讲义. 北京: 高等教育出版社, 1993

4. 方开泰, 许建伦. 统计分布. 北京: 科学出版社, 1987

5. 范金城, 吴可法. 统计推断导引. 北京: 科学出版社, 2001

6. 李贤平. 概率论基础. 北京: 高等教育出版社, 1997

7. 茆诗松, 王静龙. 数理统计. 上海: 华东师范大学出版社, 1986 8. 茆诗松, 王静龙, 濮晓龙. 高等数理统计. 北京: 高等教育出版社, 1998

9. 茆诗松. 贝叶斯统计. 北京: 中国统计出版社, 1999

10. 王梓坤. 概率论基础及其应用. 北京: 科学出版社, 1979

H. 严士健, 刘秀芳. 测度与概率. 北京: 北京师范大学出版社, 1994

12. 张尧庭, 方开泰. 多元统计分析引论. 北京: 科学出版社, 1983 13. 郑忠国. 高等统计

学.北京:北京大学出版社, 1998

14. Berger J 0. Statistical Decision Theory and Bayesian Analysis. New

York: Springer – Verlag, 1985(贾乃光译, 北京:中国统计出版社,1998) 15.

Bickel P J, Doksum K A. Mathematical Statistics. San Francisco:

Holden Day, 1977(李泽慧等译, 兰州:兰州大学出版社,2004)

16, Cramer H. Mathematical Methods of Statistics. Princeton: Prince  
ton University Press, 1946 (魏宗舒等译, 上海:上海科学技术出版社,  
1966)

17. Lehmann E L. Testing of Statistical Hypothesis. New York :  
Spring

er – Verlag, 1986

18. Lehmann E L. Elements of Large – Sample Theory. New York:  
Springer – Verlag, 1999

19. Lehmann E L, Casella G. Theory of Point Estimation. New York:  
Springer – Verlag, 1998(郑忠国等译, 北京:中国统计出版社, 2005 )

20. Rao C R. Linear Statistical Inference and Its Applications. New  
400 参考文献

York: Wiley, 1973

21. Shao J. Mathematical Statistics. New York: Springer – Verlag,

1998 22. Zacks, S. Parametric Statistical Inference. New York:  
Pergamon

Press, 1981

索引

B

BAN 估计(或 BAN): 5.3.2; 5.3.3; 5.3.4

Basu定理:2.2.4; 3.2.3; 4.2.2; 4.3.2; 4.4.2; 6.5.2; 6.5.3 Bayes: 3.  
1.2;第八章

Bayes 风险:3. 1.2; 8. 2. 1 ;

Bayes 估计:8. 2. 1 ; 8.2.4

Bayes假设检验:8. 3. 1

Bayes 解:3.1.2; 8. 2. 1

Bayes区间估计:8. 3. 2

Bayes 因子:8. 3. 1

3分布:1.2; 1.3; 1.5.1; 1.6.2; 3.2.3; 4.3.2; 6.2.2; 7.1.2; 8. 1. 2;  
8.2.2; 8.2.3; 8.2.4; 8. 3

Bh 不等式(或 Bhattacharya 不等式)i 5. 1. 2; 5. 1.3

边缘分布:1. 4; 8. 1. 1

变换群:4.1.1; 4.1.2; 4.2.1; 4.2.2; 4.3.1; 4.3.2; 4.4.1; 4.4.2 标准差  
(或均方差):1. 1. 1; 6.5.2; 7. 1.2; 7.2. 1

标准分布:1. 3; 4. 2; 4. 3; 4.4

标准正态分布:1.3; 3. 2.3; 3. 3.3; 5. 3.4; 6.2.2; 7.1.3; 8. 3. 1 不变分  
布族:4.1.2; 4.2. 1; 4.3. 1; 4.4. 1

不变原理:3. 3. 3

不容许的:3. 1.2; 8. 2.4

CAN 估计(或 CAN): 5.3.2; 5.3.3

Cauchy 分布:1. 3

X 分布:1.3; 1.4; 5.3.4; 6.3.2; 6.4.3; 6.5.2; 6.6; 6.7.2;  
6.7.3; 7. 1. 2; 7. 1. 3; 7. 1.4; 7.2.2; 7.3.2; 8. 3.2  
C-R 不等式:5. 1; 5.2  
C-R 分布族(或正则分布族):2.3.1; 5. 1. 1; 5. 1.4; 5.2; 5. 3.4; 6.6  
8.2.4  
402  
索引  
C-R下界(或方差下界):5. 1; 5.2; 5.3.2 C-R 型不等式:5. 1.3; 5.2 CRK不等  
式:5.2  
CRK 下界:5.2  
测度:1.1.1; 2.1.1; 2.3.1; 7.2.2  
超几何分布:1.2  
尺度参数:1.3; 1.5.3; 4. 1 ; 4. 3; 4.4; 8. 1.2 尺度参数分布族:4. 1. 1;  
4.3; 8. 1. 2 次序.统计量:1.6; 2. 1.2; 2. 2. 2; 7.3.3  
充分统计量:2. 1; 2.2.4; 2. 3; 3.1.3; 3.2.2; 4.3.2; 4.4.2; 6.1.3;  
8.1.1  
充分性原则:3. 1. 3; 3. 3. 1; 6.1.3; 8. 1. 1 D  
大样本性质:5.3; 6.5.4; 6.6 单边检验:6.3  
单参数:5. 1. 1; 6. 3; 6.4 单调似然比分布族:6.3; 7. 1.4 等价性检验:6.5.4;  
6.7.4 第二类错误:6.1.2; 6.1.3; 6.2; 6.3.1 第一乘错误:6. 1.2; 6. 1.3; 6.  
3. 1  
3.3. 1 ; 4.2.2;  
迭代:3.3.5  
独立性检验:6.7.4  
对立假设:6. 1. 1  
对偶关系:7.2.1  
对数似然函数:2.3.1; 3. 3. 1 ; 5. 3.4; 6.6; 6. 7.3; 7.1.3; 8. 2.3 对数正  
态分布:1.3; 3.3.3  
多参数:5. 1.4; 6.5; 6.6  
多点分布:1.2; 6. 7. 1  
多项分布:1.2; 3.3.1;3.3.2; 6.6.1; 6.7 多余参数:1.5.3; 3.3.4; 6.6.2;  
7.1.3 多元正态分布:1. 3; 2. 3. 1 ; 6. 7.2; 6.7.3; 7. 1. 2  
E  
二项分布(包括两点分布):1.2; 2. 1.2; 2. 2. 1 ; 2.3.1; 3.2.3;  
索引  
403  
3.3.1; 6.3. 1; 6.5.4; 7.1.3; 8. 1.2; 8.2.2; 8.2.3; 8.2.4; 8.3 二元正  
态分布:1.5.3; 2. 1.2; 2.2.3; 3.3.2; 6. 6. 1  
F分布:1.3; 1.4.2; 6.5.3; 6.6.1; 7.1.1  
F检验:6.5.3; 6.6.1  
Fisher信息:2.3.1; 5.1.1; 5.1.4; 5.3.2; 5.3.4; 6.6.2; 7.1.3 Fubini 定  
理:7.2.2  
方差齐性检验:6. 6.2  
方差下界:5. 1; 5.2; 5.3.2  
非正常先验分布:8. 1. 1; 8. 1.2  
非中心分布:1-4



非中心 $\chi^2$ 分布:1.4.2; 7.3.2

非主观先验分布:8. 1.2

分位数(或分位点):1.1.2; 3.2.3; 3.3.3; 6.2.2; 7.1.4; 7.3.2; 7.3.3 风险函数:3.1; 3.2.1; 3.2.2; 4.1.2; 6.1.2; 8.2.1; 8.2.4 否定域:6.1; 6.2.1; 6.6.1; 6.7; 7.2.1; 8.3.1 负二项分布:1.2; 1.5. 1; 8. 1.2

辅助统计量:2. 2.4; 4. 2. 2. ; 4.3.2; 4. 4.2

$\Gamma$ 分布(或 Gamma 分布):1.3; 1.4.1; 1.5.1; 1.5.3; 2.1.2; 2. 1.3; 3.4; 7.1.4; 8. 1.2; 8.2.2; 8. 2.3

Gauss – Newton 迭代法(或 G–N 迭代法):3.3.5; 8.2.3 功效函数(或功效): 6. 1.2; 6. 1. 3; 6. 3. 1 ; 6.4. 1; 6.4.2 共轭先验分布(或共轭分布):8. 1.2

广义 Bayes 估计:8. 2. 1 ; 8.2.2; 8. 2.4

广义 Bayes 解:8. 2. 1 广义C–R型不等式:5.2 广义先验分布(或广义先验密度):8. 1.2;

H

HPD可信区间(或 HPD可信域):8.3.2 后验分布(或后验密度):第八章 后验风险:8.2.1; 8.2.2; 8.2.3; 8.3.1

8.2.2; 8. 2.4

404

索引

后验概率:8.3 后验极大似然估计:8. 3. 2 后验期望估计:8.2.2; 8.2.4

Jeffreys 准则:8. 1. 2

Jensen 不等式:2.3.2; 3. 1. 3

几何分布:1.2

计数测度:1. 1. 1; 1. 1–4 极大似然估计:3.3; 5.3.4; 极小充分统计量:2. 1.3; 2. 2.3 极值分布:1.3; 1.5. 1 检验函数:6.1.1; 6.6.1; 7.2; 8.3.1

检验统计量:6.5.2; 6.5.3; 6.6.2; 6. 7 ; 渐近分布:1.2; 1. 3; 5. 3; 6.7; 7. 1.3 渐近有效:5. 1. 1

渐近正态性:1. 1.4; 5.3; 6. 3. 1; 6.5.4; 交比:8. 3. 1

6. 6;

6. 7. 3 ;

8. 2. 3

7. 2. 1 ;

6. 7. 1 ;

接受域:6. 1. 1; 7.2

截面似然:3.3.4; 6. 6.2

截尾指数分布:2.1.2; 3.3.1; 6.3.1; 7.1.2; 8.2.2 经验分布函数:1. 1. 1; 6.7.4

矩估计(或矩方程估计):3.4; 5.3.3

绝对损失:3. 1. 1; 3.2.2; 8.2.3 均方损失(或平方损失):3.1.1; 4.2.1; 4.2.2; 4.3.1; 4.3.2;

4.4. 1; 4.4.2; 8.2. 1; 8.2.2

均方误差(MSE): 3.1.1; 3.2.1; 4.1.1; 4.2.2; 4.3.2; 4.4.2; 5.1.1 均匀分布:1.3; 1.6.2; 2. 1; 2. 2.1 ; 2. 2.2; 2.2.4; 3. 2.3; 3. 3. 1 ; 4.1.2; 4.3.2; 4.3.3; 5. 1. 1; 5.2; 6.2.2; 6.3. 1; 7.1.2; 8. 1; 8.2.2; 8. 2.4; 8.3

K–L 距离(或 Kullback – Leibler 距离):2. 3. 2 可测函数:1.5.2; 2. 1. 1; 2. 2. 1 ; 2.3, 1

8. 3. 1  
6.7.2;  
7. 1.3

索引

405

L

Laplace 分布:1.3; 3.3.1; 3. 4 ; 4. 3. 3 Lebesgue – Stieltjes 积分:1. 1

Lehmann – Scheffe 定理:3. 2. 2 ; 累积量:1.1.3 离散型分布:1.2; 6.2; 7. 1.4

连续型分布:1. 3; 1. 6. 1 ; 6.2; 两点分布:见二项分布

两个二项分布总体的比较:6. 5.4 两类错误:6. 1.2; 6. 1.3 两样本正态总体的检验:6. 5.3 列联表:6.7.4

0 –1 损失函数:6. 1.2; 8.2.3; 8. 3. 1 零假设(或原假设) :6. 1. 1

M

minimax准则(及minimax估计):3.1.2; 8.2.4 MLE:见极大似然估计 MLR:见单调似然比分布族 MPT:见最优势检验 MREE:见最小风险同变估计

MSE:见均方误差 幂函数分布:1.3; 8. 1.2

N

Neyman – Pearson 准则:6. 1. 3 Neyman–Pearson基本引理:6.2; 6.4.2; 8.3.1; 8.3.2

拟合优度检验:6.7

逆尸分布:8. 1. 2; 8.2.2; 8.2.3 逆 Gauss 分布:1. 5. 3

P

Pareto分布:1.3; 1.5.1; 8.1.2; 8.2.2; 8.2.4 Pascal 分布:1.2; 1. 5. 1 3. 2. 3

7. 2; 7.3.2; 7.3.3

406 索引

Pitman估计:4.2.2; 4.3.2; 4.4.2; 8.2.2

Pitman 积分公式:4.2.3; 4.3.3; 4. 4. 3 ; 8. 2. 2

Poisson 分布:1.2; 1.5. 1; 2. 1.2; 2. 1. 3; 2.2.3; 3.2.3; 3. 3. 1 ; 3.4; 5. 1. 1; 5. 3. 2; 6. 2. 2; 6. 3. 1 ; 6. 7. 3 ; 7. 1. 3; 7. 1.4; 8. 1.2; 8.2.2; 8. 2. 3; 8. 3. 2

p-值:6.1.3; 6.3.2; 6.4.3; 6.7 判决函数(或统计判决函数):3.1; 4.1.2; 6.1.1; 6.1.2; 8.2.1; 8.2.4 偏差:3.2. 1; 5. 1. 1

谱分解:1.3; 6. 7.2

Q

强相合估计:5.3.2; 5.3.3; 5.3.4 区间估计:7. 1; 7. 2. 1; 8.3.2

群 :见变换群

Rao – Blackwell 定理:3. 1.3; Rao统计量:见 score统计量 Rayleigh 分布:1.3; 4. 3. 2 容忍区间与容忍限:7.3 容许性:3. 1.2; 8. 2.4

S

Schwarz不等式:5.1.1; 5.1.3; 5.1.4

score函数:2.3.1; 5.1.1; 5.1.4; 5.3.4; 6.6.2

score 检验:6. 6. 2

score 统计量(或 Rao 统计量):5. 3. 4; 6. 6. 2; 7. 1. 3

Slutsky 定理:5. 3.1; 5.3.2; 5.3.4; 6.5.4; 6.6.2; 6.7.2; 6.7.3; 7.1.3

示性函数:1. 1. 1 ; 3.2.3; 6. 1. 1; 7.2.2; 8.3.2

寿命分布(或寿命数据):1.3; 2.1.2; 3.2.3; 3. 3. 1; 6.3.1; 7. 1.2; 8.2.2

枢轴量法及枢轴量:7. 1.2; 7. 1.3; 8.3.2 似然比:2.1.3; 6.2.1; 6.3.1;

6.6.1 似然比检验:6.6

似然比统计量:5. 3.4; 6. 6; 7. 1. 3

R

3. 2. 2

索引

407

似然函数(亦见对数似然函数):2.1.1; 3.3.1; 3.3.4 似然置信域:7. 1.3

随机化检验:6. 1. 1; 6. 1.3

损失函数:3.1; 3.2.1; 3.2.2; 4.1; 4.2.1; 4.2.2; 4.3.1; .4.3.2;

4.4. 1; 4.4.2; 6. 1.2; 8.2; 8.3. 1

T

f分布:1.3; 1.4.2; 5.3.1; 6.5.2; 6.5.3; 7.1.2; 7.2.1; 7.3.2; 8.3.2

【检验:5.3.1; 6.5.2; 6.5.3

条件期望:1. 1. 3; 2.3.1; 3. 1. 3; 3.2.3; 4.2.2; 4.2.3; 4.3.2;

4.3.3; 4. 4.2 统计判决函数:见判决函数

同变估计:第四章

同变判决函数:4. 1.2

同变损失函数:4.2.1; 4.2.2; 4.3.1; 4.3.2; 4.4.1; 4.4.2

投影阵:1.4.2; 6.7.2; 6. 7.3

凸函数:1.5.2; 2.3.2; 3. 1. 1; 3. 1.3; 3.2.2; 8.2. 1 凸损失函数:3. 1. 1; 3. 1.3; 3.2.2

退化分布:1.2; 2. 3.2; 3. 1.3; 3.2.2 U

UMA:见一致最准确置信域 UMPT:见一致最优势检验 UMPUT:见一致最优无偏检验 UMRUE:见一致最小风险无偏估计 UMVUE:见一致最小方差无偏估计

W

Wald检验统计量:6. 6. 2 Wald统计量:5.3.4; 6.6.2; 7.1.3 位置参数:1.3; 4. 1; 4.2; 4,4; 8.1.2 位置参数分布族:4. 1. 1; 4.2; 8.1.2 无偏估计:3.2; 5.1; 5.2; 8.2.2 无偏检验:6. 4. 1 ; 6.4.2; 6.5. 1 无偏置信域:7. 2. 2

408

索引

X

先验分布(或先验密度):第八章

先验信息:8. 1. 1

相合性和相合估计:5.3.2; 5. 3. 3; 5.3.4; 8. 2. 2 相合渐近正态估计(CAN估计):5.3.2; 5.3.3 相关系数的估计:3.3.2

相关系数的检验:6. 6. 1 信息不等式:2.3.2; 5.3.4

Y

一致最小方差无偏估计(UMVUE): 3.2; 5. 1. 1; 5. 1.3 一致最小风险无偏估计(UMRUE): 3.2 一致最优势检验(UMPT): 6. 1.3; 6.2; 6.3; 6. 4.2 一致最优无偏检验(UMPUT): 6.4; 6.5 一致最准确置信域(UMA): 7.2.2

因子分解定理:2.1.2; 2. 3. 1; 2.3.2; 3.3.1; 5,2; 8. 1. 1

有兴趣的参数:6.6.2 原点矩:3.4; 5.3.3 原假设(或零假设):6. 1. 1

Z

正态分布:1.3; 1.5; 2.1.2; 2. 1.3; 2.2; 2.3.1; 3.2.3; 3.3; 4. 1;  
4.2.2; 4.3.2; 4.4.2; 5. 1; 5.3; 6.2,2; 6.3; 6.4.3; 6.5.2; 6.5.3;  
6.6; 7.1.2; 7. 1.3; 7.3.2; 8. 1.2; 8.2.2; 8.2.3; 8.2.4; 8.3  
正则分布族:见C-R分布族 置信区间:7. 1; 7.2; 8.3.2 置信上下限:7. 1; 7.2;  
7.3.2; 7.3.3 置信水平:7. 1; 7. 2; 8.3.2 置信域:7. 1; 7.2  
指数分布:1.3; 1.5. 1; 1.6.3; 2. 1.2;  
4.2.2; 4.3.2; 4.4.2; 6.3. 1; 6.6.1; 7. 1.2; 7. 1.4; 7.3.2; 8. 1.2;  
8.2.2; 8.2,3  
指数族分布:1.5; 2. 1.2; 2. 1.3; 2. 2. 3 ; 2.3; 3.3.2; 5.1.2; 6.3.1;  
6.4.2; 6.5.1; 8.1.2; 8.2.2  
2.2.2; 3.2.3;  
3. 3. 1 ;

索引

409

中心矩:3.4; 5.3.3 中位数:1. 1. 2; 3. 3. 1; 8. 2.3

子集参数:3.3.4; 6. 6. 2; 7. 1.3 子集参数的似然函数:3.3.4; 6. 6.2; 最不利的先验分布:8. 2.4

最大风险:3. 1.2; 8. 2.4

最大后验密度(HPD): 8.3.2 最小风险同变估计(MREE): 第四章 最优渐近正态估计(BAN估计):5.3.2; 5.3.3; 5.3.4  
7. 1.3

ISBN 7-04-020054-6 9787040200546

9> 定价35.90元