

参数统计教程

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... II. 韦... III. 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	高等教育出版社	购书热线	010 - 58581118
社 址	北京市西城区德外大街 4 号	免费咨询	800 - 810 - 0598
邮政编码	100011	网 址	http://www.hep.edu.cn
总 机	010 - 58581000		http://www.hep.com.cn
		网上订购	http://www.landaco.com
经 销	蓝色畅想图书发行有限公司		http://www.landaco.com.cn
印 刷	北京宝旺印务有限公司	畅想教育	http://www.widedu.com
开 本	787 × 960 1/16	版 次	2006 年 11 月第 1 版
印 张	26.25	印 次	2006 年 11 月第 1 次印刷
字 数	440 000	定 价	35.90 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 20054-00

前 言

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 Γ 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

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与表示衷心的感谢!

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
1.1 随机变量及其分布函数	1
1.1.1 分布函数与分布密度	1
1.1.2 反函数及分位数	3
1.1.3 特征函数和数字特征	5
1.1.4 经验分布函数	6
1.2 常见的离散型分布	7
1.3 常见的连续型分布	13
1.4 一元非中心 Γ 分布及其有关分布	24
1.4.1 非中心 Γ 分布和非中心 χ^2 分布	24
1.4.2 非中心 F 分布和非中心 t 分布	27
1.5 指数族分布	28
1.5.1 基本定义	28
1.5.2 指数族的自然形式	30
1.5.3 带有多余参数的指数族	33
1.6 次序统计量的分布	36
1.6.1 基本分布	36
1.6.2 均匀分布的次序统计量	38
1.6.3 指数分布的次序统计量	39
习题一	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
2.1.2 因子分解定理	51
2.1.3 极小充分统计量	56
2.2 统计量的完备性	59
2.2.1 分布族的完备性	60
2.2.2 统计量的完备性	62
2.2.3 指数族统计量的完备性	64
2.2.4 Basu 定理	66

2.3 分布族的信息函数	68
2.3.1 Fisher 信息	68
2.3.2 Kullback - 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	85
3.1 统计判决函数	86
3.1.1 统计判决三要素	86
3.1.2 统计判决函数的优良性准则	87
3.1.3 Rao - Blackwell 定理	89
3.2 无偏估计及其 UMRUE 和 UMVUE	90
3.2.1 基本定义	90
3.2.2 Lehmann - Scheffé 定理	91
3.2.3 例题	94
3.3 极大似然估计	100
3.3.1 定义与例题	101
3.3.2 指数族分布的极大似然估计	107
3.3.3 不变原理	109
3.3.4 子集参数的似然	111
3.3.5 极大似然估计的迭代算法	113
3.4 矩方程估计	116
习题三	118
第四章 最优同变估计	127
4.1 变换群下的同变估计	127
4.1.1 同变性概念	127
4.1.2 同变统计判决函数	129
4.2 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
4.2.3 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
4.3.1 尺度参数分布族的相似变换群	140
4.3.2 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
4.4.1 位置尺度参数分布族与线性变换群	151

4.4.2 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	158
第五章 点估计的性质	162
5.1 C-R 不等式	163
5.1.1 单参数 C-R 不等式	164
5.1.2 等式成立的条件	168
5.1.3 Bh 不等式	170
5.1.4 多参数 C-R 不等式	174
5.2 广义 C-R 型不等式	175
5.3 估计量的渐近性质	179
5.3.1 随机变量序列的收敛性	180
5.3.2 估计量的相合性和渐近正态性	183
5.3.3 矩估计的相合性和渐近正态性	190
5.3.4 极大似然估计的相合性和渐近正态性	191
习题五	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
6.1.2 两类错误及功效函数	206
6.1.3 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
6.2.2 Neyman - Pearson 基本引理应用示例	217
6.3 单调似然比分布族的单边检验	222
6.3.1 单调似然比分布族单边检验的 UMPT	222
6.3.2 正态分布单参数的单边检验	230
6.4 单参数指数族分布的双边检验	232
6.4.1 双边检验问题及无偏检验	232
6.4.2 指数族分布的双边检验	233
6.4.3 正态分布单参数的双边检验	247
6.5 多参数指数族的检验	249
6.5.1 带有多余参数时单参数检验的 UMPUT	251
6.5.2 一样本正态总体的检验	255
6.5.3 两样本正态总体的检验	263

6.5.4 两个二项分布总体的比较——等价性检验	267
6.6 似然比检验	270
6.6.1 似然比检验	271
6.6.2 子集参数的似然比检验及 score 检验	276
6.7 拟合优度检验	281
6.7.1 拟合优度检验与多项分布检验	282
6.7.2 多项分布检验的 Pearson 定理	284
6.7.3 含参数多项分布的检验及 Fisher 定理	286
6.7.4 应用: 列联表及其等价性和独立性检验	291
习题六	295
第七章 区间估计	305
7.1 置信区间及其枢轴量法	305
7.1.1 置信区间和置信限	305
7.1.2 构造置信域的枢轴量法	307
7.1.3 基于渐近分布的枢轴量法	313
7.1.4 单调似然比分布族参数的区间估计	318
7.2 参数置信域与假设检验的接受域	323
7.2.1 对偶关系	323
7.2.2 一致最准确置信域	326
7.3 容忍区间与容忍限	331
7.3.1 问题与定义	331
7.3.2 容忍上、下限的计算	333
7.3.3 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
8.1 Bayes 统计基本概念	343
8.1.1 Bayes 统计原理	344
8.1.2 先验分布的选取方法	348
8.2 Bayes 估计	356
8.2.1 Bayes 风险	356
8.2.2 后验期望估计	359
8.2.3 后验极大似然估计	366
8.2.4 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
8.3.1 Bayes 假设检验	375

8.3.2 Bayes 区间估计和 HPD 可信区间	383
习题八	390
参考文献	399
索引	401

第一章 统计分布基础

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 $X_1, \dots, X_n$; 大多情形为独立同分布(即 *i. i. d.*)样本,这时 X_1 服从未知分布 $F(x_1)$ 或 $F_\theta(x_1)$, θ 未知;如 X_1 服从正态分布 $N(\mu, \sigma^2)$, $\theta = (\mu, \sigma^2)$, 或 Poisson 分布 $P(\lambda)$ 等等.从样本 $X = (X_1, \dots, X_n)$ 出发,对未知分布 $F(x_1)$ 或 $F_\theta(x_1)$ 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 Γ 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 $X \sim (\mathcal{X}, \mathcal{B}_X, P)$,其中 $\mathcal{X}$ 表示 $\mathbf{R}^n$ 上 X 的样本空间, $X \in \mathcal{X}$; $\mathcal{B}_X$ 表示 $\mathbf{R}^n$ 上的 Borel 域; P 表示 $\mathcal{B}_X$ 上对应于 X 的概率测度,即对 $A \in \mathcal{B}_X$, $P(A) = P(X \in A)$ 表示 X 属于 A 的概率.概率测度 P 也常表示为 P_θ 、 P_X 或 P_θ^X .

今设 $X \in \mathbf{R}^1$,则一元分布函数定义为

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

熟知, 分布函数 $F(x)$ 为定义在 $(-\infty, +\infty)$ 上的非降, 右连续函数, 且有 $F(-\infty)=0$, $F(+\infty)=1$, 并且满足 $F(x+0)=F(x)$, $F(x-0)=P(X<x)$, $P(X=x)=F(x)-F(x-0)$; 若 x 为 $F(x)$ 的连续点, 则 $P(X=x)=0$. 另外, 单调函数 $F(x)$ 可扩张为 $\mathbf{R}^1$ 上的测度 $F(\cdot)$, 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 $F(x)=\int_{-\infty}^x f(y)dy$, $F'(x)=f(x)$, 则称 $F(x)$ 为绝对连续型, 且有 $P(X \in A)=\int_A f(y)dy$. 若 $F(x)$ 连续但不存在密度函数, 则称 $F(x)$ 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 $P(X=x_k)=p_k \triangleq f(x_k)$, $k=0, \pm 1, \pm 2, \dots$, $\sum_k p_k=1$. 例如 Poisson 分布: $P(X=k) \triangleq f(k)=e^{-\lambda} \lambda^k/k! = p_k (k=0, 1, 2, \dots)$. 退化分布是离散型分布一个重要特例, 这时 $P(X=0)=1$ 或 $P(X=a)=1$; 其分布函数为

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

或 $F(x)=\Delta(x-a)=I\{x \geq a\}$, 其中 $I\{x \geq a\}$ 为示性函数. 对于一个集合 A , $I_A(x) \triangleq I\{A\}$ 定义为 $I_A(x)=1$, $x \in A$; $I_A(x)=0$, $x \notin A$. 因此, 一般情形下离散型分布的分布函数可表示为

$$F(x) = \sum_k p_k \Delta(x - x_k) = \sum_{x_k \leq x} p_k = \sum_{x_k \leq x} f(x_k). \quad (1.1.1)$$

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 $\mathbf{R}^1$ 上的积分, $\mathbf{R}^n$ 上的积分完全类似. 给定 $\mathbf{R}^1$ 上的有界函数 $g(x)$ 和有限测度 $\mu(\cdot)$, 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 $\Delta x_i (i=1, \dots, n)$ 之和. 对任一“分划”, 记 $S_m = \sum_{i=1}^n m_i \mu(\Delta x_i)$ 和 $S_M = \sum_{i=1}^n M_i \mu(\Delta x_i)$, 其中 $\mu(\Delta x_i)$ 为 Δx_i 的测度, m_i 和 M_i 分别为 $g(x)$ 在 Δx_i 上的下确界与上确界. 若 $\max_{1 \leq i \leq n} \{\mu(\Delta x_i)\} \rightarrow 0$ 时, S_m 和 S_M 的极限存在且相等, 则其极限称为函数 $g(x)$ 在区间 $[a, b]$ 上关于测度 $\mu(\cdot)$ 的 Lebesgue - Stieltjes 积分, 并记为 $\int_a^b g(x) d\mu(x)$ 或 $\int_a^b g(x) \mu(dx)$. 以下介绍与分布函数有关的两种情况.

(1) 若 $\mu(\cdot)$ 为 Lebesgue 测度, Δx_i 为区间, 则 $\mu(\Delta x_i) = \Delta x_i$ 为区间长度; $\int_a^b g(x) d\mu(x) = \int_a^b g(x) dx$. 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

$$F(x) = \int_{-\infty}^x f(y) dy = \int_{-\infty}^x f(y) d\mu(y). \quad (1.1.2)$$

(2) 若 $\mu(\cdot)$ 为计数测度. 设点列 $\{x_k\}$ 有可数个值 $x_1, x_2, \dots, x_k, \dots$, 计数测度 $\mu(A)$ 表示 A 所包含点列 $\{x_k\}$ 中点的个数. 因此对区间 $\Delta x_i \rightarrow 0$ 有 $\mu(\Delta x_i) = 1$, 若 Δx_i 中包含一个 $\{x_k\}$ 中的点; $\mu(\Delta x_i) = 0$, 若 Δx_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 $\int_a^b g(x) d\mu(x) = \sum_{x_k \in [a, b]} g(x_k)$. 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 $f(x_k) = p_k = P(X = x_k)$, 则有

$$F(x) = \sum_{x_k \leq x} f(x_k) = \int_{-\infty}^x f(y) d\mu(y). \quad (1.1.3)$$

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 $f(x_k) = e^{-\lambda} \cdot \lambda^{x_k} / x_k!$.

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

$$E[\phi(X)] = \int \phi(x) f(x) d\mu(x) = \int \phi(x) dF(x),$$

其中 $dF(x) = f(x) d\mu(x)$.

$$\begin{aligned} \text{特别, 概率 } P(A) \text{ 可表示为 } P(A) &= \int_A I_A(x) f(x) d\mu(x) \\ &= \int_A f(x) d\mu(x) = \int_A dF(x). \end{aligned}$$

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 $\mu(\cdot)$ 就取为概率测度 $P(\cdot)$, 则有 $P(A) = \int_A I_A(x) dP(x) = \int_A dP(x)$; 也有 $E[\phi(X)] = \int \phi(x) dP(x)$.

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 $x_p \triangleq \inf \{x; F(x) \geq p\}$.

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

i) 若 $x' < x_p$, 则 $F(x') < p$; 若 $x' \geq x_p$, 则 $F(x') \geq p$;

ii) $F(x_p - 0) \leq p \leq F(x_p)$; 若 x_p 为 $F(x)$ 的连续点, 则 $F(x_p) = p$.

常用的分位点有 $x_{0.05}$, $x_{0.1}$, $x_{0.9}$, $x_{0.95}$, $x_{0.5}$ 等. $x_{0.5}$ 通常称为中位数, 它有以下重要性质.

定理 1.1.1 函数 $g(c) = E|X - c|$ 在中位数 $c = x_{0.5}$ 时达到最小值.

证明 设 X 的分布函数为 $F(x)$, 若 $c < x_{0.5}$, 则

$$\begin{aligned} g(c) &= \int_{-\infty}^{+\infty} |x - c| dF(x) \\ &= \int_{(-\infty, c)} (c - x) dF(x) + \left(\int_{[c, x_{0.5})} + \int_{[x_{0.5}, +\infty)} \right) (x - c) dF(x) \\ &= \int_{(-\infty, c)} [(x_{0.5} - x) - (x_{0.5} - c)] dF(x) + \int_{[x_{0.5}, +\infty)} (x - x_{0.5} + x_{0.5} \\ &\quad - c) dF(x) + \int_{[c, x_{0.5})} [(x_{0.5} - x) + 2(x - c) - (x_{0.5} - c)] dF(x) \\ &= \int_{-\infty}^{+\infty} |x - x_{0.5}| dF(x) + 2 \int_{[c, x_{0.5})} (x - c) dF(x) \\ &\quad + (x_{0.5} - c) [P\{X \geq x_{0.5}\} - P\{X < x_{0.5}\}]. \end{aligned}$$

由 $c < x_{0.5}$ 及 $x_{0.5}$ 的定义知 $P\{X \geq x_{0.5}\} \geq \frac{1}{2}$ 和 $P\{X < x_{0.5}\} = F(x_{0.5} - 0) \leq \frac{1}{2}$, 从而 $g(c) \geq E|X - x_{0.5}|$; 反之, 若 $c > x_{0.5}$, 则

$$\begin{aligned} g(c) &= \int_{-\infty}^{+\infty} |x - c| dF(x) = \left(\int_{(-\infty, x_{0.5}]} + \int_{(x_{0.5}, c]} \right) (c - x) dF(x) \\ &\quad + \int_{(c, +\infty)} (x - c) dF(x) \\ &= \int_{(-\infty, x_{0.5}]} [(x_{0.5} - x) + (c - x_{0.5})] dF(x) + \int_{(c, +\infty)} [(x - x_{0.5}) \\ &\quad - (c - x_{0.5})] dF(x) + \int_{(x_{0.5}, c]} [(x - x_{0.5}) + 2(c - x) \\ &\quad - (c - x_{0.5})] dF(x) \\ &= \int_{-\infty}^{+\infty} |x - x_{0.5}| dF(x) + 2 \int_{(x_{0.5}, c]} (c - x) dF(x) \\ &\quad + (c - x_{0.5}) [P\{X \leq x_{0.5}\} - P\{X > x_{0.5}\}]. \end{aligned}$$

由 $c > x_{0.5}$ 及 $x_{0.5}$ 的定义知 $P\{X \leq x_{0.5}\} \geq \frac{1}{2}$ 和 $P\{X > x_{0.5}\} \leq \frac{1}{2}$, 从而

$g(c) \geq E|X - x_{0.5}|$. 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

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 $E(X) = \mu$; 方差为 $\text{Var}(X) = \sigma^2$, 而 σ 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 $M(t) \triangleq E(e^{itX})$ 和 $\varphi(t) \triangleq E(e^{itX}) = M(it)$. 常用的性质有

$$\text{i)} \quad M^{(k)}(0) = E(X^k) \triangleq a_k, \quad \varphi^{(k)}(0) = i^k a_k, \quad E(X^k) = i^{-k} \varphi^{(k)}(0);$$

$$\text{ii)} \quad \varphi(t) = 1 + \sum_{k=1}^{\infty} a_k \frac{(it)^k}{k!}, \quad M(t) = 1 + \sum_{k=1}^{\infty} a_k \frac{t^k}{k!};$$

iii) 给定随机向量 $\xi = (\xi_1, \dots, \xi_n)$, 其分量 $\xi_1, \dots, \xi_n$ 相互独立的充要条件为

$$\varphi_{\xi}(t_1, \dots, t_n) = \varphi_{\xi_1}(t_1) \varphi_{\xi_2}(t_2) \cdots \varphi_{\xi_n}(t_n).$$

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). $\log \varphi(t)$ 的 Taylor 展开式的系数称为累积量, 若

$$\log \varphi(t) = \sum_{r=1}^{\infty} \kappa_r \frac{(it)^r}{r!},$$

则系数 κ_r 称为 r 阶累积量或半不变量. ■

随机变量 X 的累积量 κ_r 与其各阶矩 a_k 有密切的关系. 其常用性质有

i) 累积量与各阶矩可互相表示. 其前 3 阶矩的关系为:

$$\kappa_1 = a_1, \quad \kappa_2 = m_2 = E(X - EX)^2, \quad \kappa_3 = a_3 - 3a_1a_2 + 2(a_1)^3;$$

$$a_1 = \kappa_1, \quad a_2 = \kappa_2 + (\kappa_1)^2, \quad a_3 = \kappa_3 + 3\kappa_1\kappa_2 + (\kappa_1)^3;$$

ii) 若 X, Y 独立, 则有 $\kappa_r(X+Y) = \kappa_r(X) + \kappa_r(Y)$;

iii) $\kappa_r(X+C) = \kappa_r(X)$, ($r > 1$), 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 $\gamma_1 \triangleq \alpha_3/\sigma^3$; 峰度系数 $\gamma_2 \triangleq \alpha_4/\sigma^4 - 3$; 其中 $\alpha_k = E(X -$

$EX)^k$; $\sigma^2 = \text{Var}(X)$. 容易验证, 若 X 为正态分布, 则有 $\gamma_1 = 0$, $\gamma_2 = 0$. 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.

$$E(X) = E_T \{ E(X|T) \}; \quad (1.1.4)$$

$$\text{Var}(X) = E_T \{ \text{Var}(X|T) \} + \text{Var}_T \{ E(X|T) \}. \quad (1.1.5)$$

另外, 若 $X = (X_1, \dots, X_n)^T$, $Y = (Y_1, \dots, Y_m)^T$ 为随机向量, 则其期望与方差定义为 $E(X) \triangleq (EX_1, \dots, EX_n)^T$, $\text{Cov}(X, Y) \triangleq (\sigma_{ij})_{n \times m}$; $\sigma_{ij} = \text{Cov}(X_i, Y_j)$; $\text{Var}(X) \triangleq \text{Cov}(X, X)_{n \times n}$. 随机向量的期望与方差的常用公式有

$$\text{i) } \text{Cov}(X, Y) = E[(X - EX)(Y - EY)^T] = E(XY^T) - (EX)(EY)^T;$$

$$\text{Var}(X) = E(XX^T) - (EX)(EX)^T;$$

$$\text{ii) } \text{Cov}(AX, BY) = A\text{Cov}(X, Y)B^T;$$

$$\text{iii) } E\|X\|^2 = \|EX\|^2 + \text{tr}[\text{Var}(X)]; \quad (1.1.6)$$

$$\text{iv) 设 } X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \text{ 则 } \text{Var}(X) = \begin{pmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) \end{pmatrix}.$$

1.1.4 经验分布函数

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_1 服从未知分布 $F(x) = P(X_1 \leq x)$. 给定 x , 记 ν_n 表示 $X_1, \dots, X_n$ 中 $\leq x$ 的个数. 易见, 概率 $F(x) = P(X_1 \leq x)$ 可用频率 ν_n/n 来逼近.

定义 1.1.3 给定独立同分布的样本 $X_1, \dots, X_n$ 和 x , 经验分布函数定义为 $F_n(x) \triangleq n^{-1}\nu_n$. ■

易见, 经验分布函数可表示为

$$F_n(x) = n^{-1} \sum_{i=1}^n I\{X_i \leq x\} \triangleq n^{-1} \sum_{i=1}^n Y_i, \quad (1.1.7)$$

其中 $Y_1, \dots, Y_n$ 为独立同分布, 取 0, 1 两个值的随机变量; 即 $P(Y_i = 1) = P\{X_i \leq x\} = F(x)$, $P(Y_i = 0) = P\{X_i > x\} = 1 - F(x)$. 并且有 $E(Y_i) = F(x)$, $\text{Var}(Y_i) = F(x)[1 - F(x)]$.

定理 1.1.2 若 $X_1, \dots, X_n$ 为 i. i. d. 样本, $X_1 \sim F(x)$, 则对 $\forall x$ 有

$$(1) F_n(x) \rightarrow F(x) (a. e.);$$

$$(2) \sqrt{n}\{F_n(x) - F(x)\} \xrightarrow{L} N(0, F(x)[1 - F(x)]),$$

其中 (a. e.) 表示几乎处处收敛, $\xrightarrow{L}$ 表示依分布收敛.