非参数统计

陈希孺　方兆本　李国英　陶　波　著

上海科学技术出版社

内 容 提 要

本书是对非参数统计的理论和方法作比较严格而系统论述的专著。主要内容有：次序统计量，U统计量，秩统计量的极限理论，秩方法，条件检验与置换检验，密度估计、非参数回归与判别，稳健统计初步等。

本书取材是根据作者在科研和教学中所积累的资料并参考了国外近年出版的一些专著、文献而写成，反映了本学科的现代面貌。有些结果尚属初次发表。

本书可作为高等学校中概率和统计专业的高年级学生及研究生的教材和重要参考书；也可供数理统计研究工作者以及具有相当数学水平的实用统计工作者参考。

非参数统计

陈希孺　方兆本
李国英　陶　波　著

责任编辑　赵序明

上海科学技术出版社出版
（上海瑞金二路450号）

新华书店上海发行所发行　商务印书馆上海印刷厂印刷

开本850×1156 1/32　印张12.75　字数339,000

1989年8月第1版　1989年8月第1次印刷

印数：1—1,900

ISBN 7-5323-0994-0/O·110

定价：8.15元

序言

非参数统计是数理统计学的一个分支，它形成于本世纪四十年代，而在第二次世界大战以后得到迅速发展，至今已成长为一个体系博大、理论精深且富于实用价值的分支，受到数理统计学者和应用工作者的重视.

本书的目的，是对非参数统计的理论和方法作一比较严格而系统的论述. 本书属专著性质，我们的主观愿望在取材上能有一定的广度和深度，以反映本学科的现代面貌；另一方面，我们也希望本书的读者面能尽量广一些，因而在预备知识的要求方面以及在使用的数学工具的深度方面，不能不加以一定的限制，并舍弃一些过于专门的材料。因此，本书也带有"非参数统计引论"的性质.

本书设想的读者对象是高等学校中概率和统计专业的学生、研究生与教师；数理统计研究工作者以及具有相当数学水平的实用统计工作者。本书绝大部分内容所需要的预备知识，是大学数学系的微积分（或比较扎实的非数学系的高等数学）及少量的矩阵知识，以及相当于复旦大学数学系主编的《概率论与数理统计》一书中的概率与统计知识，个别地方用到一点测度论和分析概率论的知识。

写成本书，是根据我们以往在学习、讲授和研究这个分支所积累的一些笔记资料并参考了国外近年出版的一些

专著。1982年夏写成初稿，曾在四川大学的一个讲习班上使用过，在此基础上，我们又分头收集材料，作了充实和加工改写，最后由陈希孺执笔定稿，以利于在表述方式、行文和符号等方面达到统一。本书是集体工作的产物，但在最后定稿中出现的问题，由执笔者负责。由于作者们的水平所限，书中不妥之处，在所难免，希望同行学者、专家及其他读者不吝指正。本书的出版，得到上海科学技术出版社的热心支持，谨在此表示衷心感谢。

作 者

目　录

导　言

本书是一本非参数统计的专著．开宗明义第一件事，就是解释什么是非参数统计和非参数统计方法．不过，这个问题不是三言两语能说清楚的．关于本书在取材上的考虑、各章内容安排及所用符号，也需作说明．故写了这篇导言．其中有的内容，须具备一些非参数统计的基础知识才能理解，不过初次接触时，浏览一过就可以了．

先谈第一个问题：

顾名思义，“非参数统计”是“参数统计”的对立面．所谓参数统计，粗略地说，就是在一般教程中常见的一套基于正态假定的统计方法，如 χ^2、t 和 F 检验、正态线性回归、狭义的多元分析等等．Fraser 的著作[68]有一个副标题，意谓非参数方法是当正态假定被一般假定取代时的统计方法．这种说法作为一个初步的解释是可以的，但作为非参数统计的正式定义则不可以．因为正态分布族固然是最重要的参数分布族，但不是唯一的重要参数分布族．

把以上的说法加以引伸，并适当抽象化，就可以提出下面的说法，以作为划分参数统计问题与非参数统计问题的分界线：如果在一个统计问题中，所假定的总体分布族的数学形式已知，而只包含有限个(通常为数很少)未知的实参数，则这个统计问题是参数性的，否则，就是非参数性的．诸如可靠性统计中估计负指数分布和 Weibull 分布参数的问题，产品抽样验收中对二项分布参数的检验问题等，都是参数统计问题．而确定连续分布的容忍限，一般的两样本检验问题(检验两组样本所来自的总体有相同分布)，对称分布(形式未知)的对称中心的估计及形状未知的回归函数的估计问题等等，则都是非参数统计问题．常见的拟合优度检验问题也是非参数性的，因为在此问题中，理论分布是一个有待检验的假设而

并非假定.

这个划分的准则是合理且明确易行的，在统计文献中多采取这个说法.本书作者也同意这一观点，但认为还不可完全拘泥于文字.例如，在Gauss–Markov线性模型$\{Y=X\beta+e, Ee=0, \mathrm{COV}(e)=\sigma^2 I\}$中，并未对模型中的随机误差向量$e$的分布的具体形式作什么假定，故按上述准则，这模型的统计问题应是非参数性的，但一般并不这样看：在非参数统计著作中，都不把用最小二乘法处理这个模型的理论和方法包括进去．这是因为，在这个模型中，人们感兴趣的是回归系数向量β，它只涉及有限个实参数，且$X\beta$是一个简单的线性形式．又如极值统计，根据极值统计方法中对底分布(参看§1.2(三))并无特定要求一节，可将其归入非参数统计范围，但是，极值分布只有三种简单的参数族，因此，把它列入参数统计范围，也言之成理．我们的结论是：在多数情况下，参数问题和非参数问题的划分是明确的；在少数情况下，则可因人的看法及侧重点的不同而异，这还要考虑到习惯.

除了统计问题有参数与非参数之分以外，在文献中，也常把一定的统计方法划为“参数性”的或“非参数性”的．然而，要对此定出合理的划分标准，则问题更复杂了．大体上说，可以把非参数统计方法理解为“处理非参数统计问题的统计方法”．但取之作为正式定义，仍觉有所不足．因为有些重要的统计方法在参数和非参数统计问题中都有用，矩法就是一个例子．又如最小二乘估计法，在形式上看更适合用于参数问题．但如把Gauss–Markov模型视为一种非参数模型，这个方法的主要应用则在非参数统计中.

本书作者倾向于采取一种较有伸缩性的观点：一种统计方法，如果主要用于非参数性的统计问题(包括其属性可有争议，但习惯上认为是非参数性的统计问题)，则这个方法称为是非参数性的；否则是参数性的．按照这种观点，一个非参数统计方法也可以用于典型的参数统计问题中(反之亦然)，但不是它的主要用武之地．一个典型的例子是秩方法(第四章)，它是非参数统计的主要方法．但秩方法不仅可以而且确实也用于一些典型的参数统计问

题中. 置换检验(第五章)按其方法的性质及 Fisher 引进这种方法的思想,是非参数性的统计方法. 又如,用次序统计量去构造连续分布的容忍区间,是一种非参数方法,但它也可用于正态分布的情况. 不过对后者而言,人们一般使用形如 $\bar{x}\pm cs$ 的区间. 又如,由于考虑到 Gauss-Markov 模型在习惯上并不视为非参数性的,故最小二乘估计法就不宜看作是一种非参数方法.

有的统计方法,在参数和非参数问题中都有重要应用,不能勉强将其划入某一类. 这时,只好就具体情况而言:该方法适于参数应用还是非参数应用. 例如矩估计法,在参数估计中是一个重要方法,但若只假定总体分布有均值,而其他一无所知,则除了用样本均值去估计总体均值外,别无其他良法,这种性质的应用并不限于直接用样本矩估计总体矩,例如,概率密度函数的核估计法(见§6.1),可以认为是一种导源于矩估计思想的方法. 所有这些都可以说是"矩方法的非参数性应用". 又如,二项分布的概率 p 的检验,无疑是一种参数统计方法,但它在非参数性的对比试验模型中也有重要应用(即符号检验,见§3.1),这可以说是二项分布参数检验法的非参数应用.

下面谈一下文献中关于这个问题的一些值得注意的观点:

Walsh 在[172] 中把一个方法的非参数性归结为在应用上的广泛性. 按照这种观点,二项分布参数的检验法是一种非参数方法,因为有许多问题最后都归结到这个检验问题. 本书作者认为,应用的广泛性确实是非参数方法的一个特点(因为它对模型的要求少,应用自然就广了),但以此作为划分标准则不恰当,因为,一种方法在应用上是否"广泛",其看法因人而异,尤其重要的是:划分的标准应当着重于方法的性质,而不在于应用上是否广泛.

在另一个极端,Kendall 等([98])主张,只能讲统计假设的"参数"或"非参数"性,而不能将这些形容词用于检验、统计量等. 换句话说,问题有参数与非参数之分,而统计方法则否. 他们认为,将一个检验称为"参数"或"非参数"这种提法,已造成了不少混乱. 按照这种观点,下面这个两样本问题:$\{X_1, \cdots X_m$ 和 $Y_1, \cdots,$

Y_n 分别是抽自总体分布 F 和 G 的随机样本，要检验假设 $F=G$}可称为非参数性的．而一个具体的检验方法，例如 Смирнов 检验(见 § 4.6)则不能冠之以“参数性”或“非参数性”．本书作者认为：这种看法似乎过于绝对化．因为大多数统计方法的应用有其重点所在．据此对方法的性质(参数或非参数)进行划分，不仅可能而且是有益的．

Conovor 在 [49] 中提出如下的看法：一个统计方法，如果适合下述条件之一，就可以称为*非参数方法*：1．它可用于属性数据；2．它可用于只分别大小次序而不计具体数值的数据；3．它可用于通常的数据(有具体数值的)，其所来自的分布族不能用有限个实参数刻划．我们认为：Conovor 这个提法明确指出了非参数统计方法的主要应用所在．但以此作为定义则似乎欠妥．比方说：最大似然估计与似然比检验可用于属性数据，但这些方法是参数性的．按 Conovor 的观点，最小二乘估计法是非参数方法，因为这方法主要用于 Gauss-Markov 模型，而后者的分布族不能用有限个实参数刻划．因此，我们认为 Conovor 的定义有些失之过宽．

美国统计学家常使用一个被称为“分布无关”(distribution-free)的概念，并理解为“非参数性”的同义语．所谓“分布无关”，一般是指下述情况：设样本 X 的分布 F 属于分布族 $\mathscr{F}$，而 $\mathscr{F}_0\subset\mathscr{F}$．设 $T=T(X)$ 为一统计量，若当 $F\in\mathscr{F}_0$ 时，T 的分布不依赖于 F，则称 T 相对于 $\mathscr{F}_0$ 为“分布无关”的。例如设 $X=(X_1,\cdots,X_n)\sim N(\mu,\sigma^2)$，$\mathscr{F}=\{N(\mu,\sigma^2):-\infty<\mu<\infty,\sigma^2>0\}$，指定 μ_0，而作 t 统计量 $T=\sqrt{n}(\overline{X}-\mu_0)/s$，则对 $\mathscr{F}_0=\{N(\mu_0,\sigma^2):\sigma^2>0\}$ 而言，T 是分布无关的．因此，“分布无关”这个概念并不专用于非参数性的分布族，对参数族也很重要和常见．这个概念主要用于假设检验，$\mathscr{F}_0$ 是原假设，如果检验统计量 T 相对于原假设 $\mathscr{F}_0$ 为分布无关，就可以作出基于 T 的相似检验．一个显著的例子是秩统计量．在一些非参数检验问题中，往往在原假设成立之下，样本 $X_1,\cdots,X_n$ 为独立同分布且分布连续．这时秩统计量相对于原假设下的分布族为“分布无关”的．秩统计量的重要性正在于此．“分布

无关"与"非参数"这两个概念不能混同(前者是指统计量的某种性质;后者是指统计问题和方法的性质). 不过,"分布无关"这个概念更多的是用于非参数分布族之中,且在非参数方法中使用的统计量往往有"分布无关"性. 因此,"分布无关"这个概念确实从一个重要侧面刻划了"非参数性"的含义.

关于将统计问题和统计方法划分为"参数"或"非参数"的标准问题,以上讲了作者自己的看法及文献中的一些提法. 但我们并不是要把某一种意见(包括我们自己的意见)作为结论性的观点推荐给读者,而只是想提供一些材料,供读者思考这个问题时参考.

现在转到第二个问题:

由于对非参数统计和非参数方法的含义有不同的理解,故对于非参数统计中应当包含那些题材的问题,看法也就不尽一致了. 本书在取材问题上的指导思想,除了考虑到对非参数统计的理解及材料在理论和应用上的价值外,也还在一定程度上考虑到习惯与叙述上的方便,而不拘泥于一种想法. 例如有关拟合优度的 χ^2 检验和列联表的内容,在性质上可归入非参数统计的范围,但在习惯上则多数是放到其他著作中论述. 所以在本书中不包括这一内容. 次序统计量的有些内容虽然明显地是属于参数统计的范围,但为了照顾到系统性与叙述上的方便,也将其收进本书中,又如第七章中的稳健统计,有一种观点认为这个题材不属于非参数统计,作者也基本上同意这个观点,但考虑到稳健性与非参数的概念有密切联系,并且这方面的材料较新,而在中文文献中介绍得很少,因此将这一题材也收进本书. 下面简介各章的内容.

第一章是次序统计量. 把一组样本 $X_1, \cdots, X_n$ 按大小次序排列为 $X_{(1)} \leqslant \cdots \leqslant X_{(n)}$,后者(或其一部分)就称为次序统计量. 这一章叙述了这种统计量的分布和渐近分布及其在统计问题上的重要应用. 渐近理论方面包含三个内容:样本分位数的渐近正态性;极值($X_{(1)}$ 和 $X_{(n)}$)的渐近分布;次序统计量的线性组合 $\sum_{i=1}^{n} C_{ni} X_{(i)}$ 的渐近正态性. 这些结果有很强的理论和实际意义.

400998

第二章是 U 统计量. 它是 Hoeffding 在 1948 年引进的一类统计量(定义见 § 2.1), 在一些非参数性的估计和检验问题中有用. 这一章介绍了这种统计量的定义、方差的计算与渐近正态性定理, 举了一些例子说明其应用. 也介绍了关于其极限理论的若干深入结果, 但由于它们的统计意义较小, 因此不予详细论证.

第三、四章是基于秩的统计方法. 这是非参数统计的中心内容. 因此在本书中也给予较大的篇幅. 所谓秩, 就是指各样本在其大小比较中所占的位次: 若将样本 $X_1, \cdots, X_n$ 按大小排列为 $X_{(1)}<\cdots<X_{(n)}$, 而 $X_i=X_{(R_i)}$, 则称 X_i 的秩为 R_i, 而 $(R_1, \cdots, R_n)$ 称为秩统计量. 第三章主要讨论线性秩统计量的极限分布. 对同分布的情况作了严格的论证; 对不同分布和结存在的情况则只介绍有关结果 (因为严格的证明过于繁复). 第四章可分为两部分: 前两节是讨论秩检验在两种准则(渐近相对效率与局部最优)下的优良性. 总的结论是: 秩检验与传统的参数检验相比, 处于有利的地位. 本章后四节讨论秩方法在各种统计问题中的应用——估计问题、方差分析模型、独立性检验、Колмогоров 与 Смирнов 检验等. 应当指出: Колмогоров 检验并不是秩方法, 因其与 Смирнов 检验(它是一个秩检验)的密切联系而放在这里.

第五章的主要内容是置换检验, 又称随机化检验, 是 Fisher 1935 年在其名著 [66] 中引进的, 它基于 Fisher 所提出的试验设计三原则之一——随机化原则. 置换检验有重要的理论意义, 因为它把方差分析的理论置于更现实的基础上. 本章通过较多的例子详细分析了条件检验与置换检验的基本思想, 对其大样本理论(包括大样本极限分布及渐近效率)作了严格的处理.

第六章把几个较近发展起来的, 彼此有关的题材——密度估计、非参数回归与判别——收集在一起作了介绍. 这部分题材较新, 因而还不能说它已经发展得很成熟和定型了, 但考虑到这些内容有较强的实用价值与理论意义, 故将它作为一章写入本书还是合宜的. 本章共分三节: 第一节讨论了两种重要的密度估计法——核估计与近邻估计; 第二节讨论了非参数回归函数的权函数

估计及其两个重要特例——核估计与近邻估计；第三节讨论了非参数判别法，介绍了 Bayes 型的方法与近邻方法．对一些较基本的大样本结果给了详细的证明．与其他各章不同的是：这里不加证明地引证了较多的最近结果，因为这方面还正在发展中，对此有兴趣的读者，可通过这些介绍找到继续研究的线索．

第七章是关于稳健统计．这也是比较新的题材．本章所选择的材料，主要是关于稳健性的基本概念以及较富于实际意义的理论和结果．这方面的材料可认为是比较定型的．

本书所用的符号，大多数在统计文献中是标准的，例如用 E，Var，Cov（或 COV）记随机变量的均值、方差与协方差（协方差阵）．用

$$X_n \xrightarrow{P} X,\ X_n \to X,\ a.s.\ (\text{或} \lim_{n\to\infty} X_n = X,\ a.s.),$$

分别表示当 $n\to\infty$ 时，变量 X_n 依概率或以概率 1 收敛于变量 X，若 X_n 的分布为 F_n，而 F 为一分布函数，则以 $F_n \xrightarrow{\mathscr{L}} F$ 或 $X_n \xrightarrow{\mathscr{L}} F$ 表示 F_n 或 X_n 依分布收敛于 F；以缩写记号“iid.”(independently-identically distributed)表示“独立同分布”．若某一总体的总体分布为 F，而 $X_1, \cdots, X_n$ 为抽自该总体的独立随机样本，则常记为 $X_1, \cdots, X_n \sim F$．有时也用记号“$X\sim F$”表示“变量 X 有分布函数 F”；$X \overset{d}{=} Y$ 表示 X 与 Y 的分布相同；又 I_A 或 $I_A(x)$ 表示集 A 的指示函数，即在 A 上的值为 1 而在 A 外为 0．

本书也涉及少量的矩阵向量运算，按一般的表示法：当将 X 视为向量时，总是看作列向量，向量或矩阵 A 的转置记为 A'，故 X' 表示行向量．向量 a 与 b 的内积记为 $a'b$．向量 a 的长度常记为 $\|a\|$．在任何特定意义下，两点 x 与 y 的距离也常记以 $\|x-y\|$．

第一章

次序统计量

设有样本 $X=(X_1,\cdots,X_n)$. 把 $X_1,\cdots,X_n$ 按由小到大的次序排列为

$$X_{(1)}\leqslant X_{(2)}\leqslant\cdots\leqslant X_{(n)},\tag{1}$$

则 $(X_{(1)},\cdots,X_{(n)})$ 称为样本 X 的次序统计量(order statistics). 习惯上也常把序列(1)的一部分称为次序统计量. 特别, $X_{(i)}$ 常称为第 i 个次序统计量. 在文献中, 有时把 (1) 称为"次序样本"(ordered sample). 如果 $X_1,\cdots,X_n$ 是从总体分布 F 中抽出的 iid. 样本, 则称(1)是从 F 中抽出的(大小为 n 的)次序样本.

次序统计量在统计问题中有广泛的应用, 其理论也有深入的发展. 二十多年来, 出现了不少这方面的专著, 如[53]、[71]、[140]、[83]和[132]. 在一定程度上可以说: 次序统计量的研究已形成数理统计学和概率论的一个分支. 本章的目的就是介绍次序统计量的初步理论及其若干应用. 有兴趣的读者可在本章内容的基础上阅读上面提到的专著.

有一点需要明确一下: 次序统计量既可用于典型的非参数统计问题, 如找连续分布函数的分位数的置信区间; 也可用于典型的参数统计问题, 如用极差(见定义 1.1) 的一适当倍数去估计正态分布的标准差. 因此, 从学科角度看: 不好把次序统计量的理论与方法说成是非参数统计的一部分. 但一些作者在撰写非参数统计的著作时, 却往往把次序统计量这个题目纳入其范围内, 这主要是因为: 如在本章 §1.4 中将看到的, 次序统计量有一系列的重要应用确是属于非参数的性质. 但与次序统计量的专著不同, 在非参数统计著作中介绍次序统计量的应用时, 一般不去深入讨论那种

纯参数性质的应用. 本章的叙述也遵循这个原则.

本章共分四节: §1.1叙述有关次序统计量分布的若干基本事实; §1.2是关于次序统计量的渐近分布的几个重要定理; §1.3讨论这种统计量的充分性与完全性; 最后在§1.4中, 介绍次序统计量的一些重要应用.

§1.1 基本分布

在应用上, 最常见的情况是: $X_1, \cdots, X_n$ 是从一个有分布 F 的总体中抽出的 iid. 样本. 本节就在这个假定下来讨论问题.

(一) $X_{(r)}$的分布

以 F_r 记 $X_{(r)}$的分布函数, 依定义有

$$
\begin{aligned}
F_r(x) &= P(X_{(r)}<x) = P(X_1, \cdots, X_n \text{ 中至少有 } r \text{ 个小于 } x) \\
&= \sum_{j=r}^{n} P(X_1, \cdots, X_n \text{ 中恰有 } j \text{ 个小于 } x) \\
&= \sum_{j=r}^{n} \binom{n}{j} F^j(x)(1-F(x))^{n-j} \\
&= \frac{n!}{(r-1)!(n-r)!}\int_0^{F(x)} t^{r-1}(1-t)^{n-r}\,dt. \qquad (2)
\end{aligned}
$$

(2)中的最后一式是基于恒等式

$$
\sum_{j=r}^{n}\binom{n}{j}p^j(1-p)^{n-j} = \frac{n!}{(r-1)!(n-r)!}\int_0^p t^{r-1}(1-t)^{n-r}dt \qquad (3)
$$

$$
(r=1, \cdots, n,\ 0\leqslant p\leqslant 1).
$$

(3)式不难用下法证明: 首先, 当 $p=0$ 时, 两边皆为0; 其次, (3)式左边对 p 求导, 再经过简单的整理, 就得到

$$
[n!/(r-1)!(n-r)!]p^{r-1}(1-p)^{n-r}.
$$

若分布 F 有概率密度函数 f, 则由(2)易知 $F_r(x)$ 也有概率密度函数 $f_r(x)$, 且

$$
f_r(x) = \frac{n!}{(r-1)!(n-r)!}F^{r-1}(x)(1-F(x))^{n-r}f(x). \qquad (4)
$$

有两个重要情况对应于 $r=n$ 和 $r=1$. $X_{(n)}$和 $X_{(1)}$ 分别是样本 $X_1, \cdots, X_n$ 中的极大和极小值,统称为极值. 由(2)与(4)知

$$F_n(x)=F^n(x), \qquad f_n(x)=nF^{n-1}(x)f(x). \tag{5}$$

$$F_1(x)=1-(1-F(x))^n, \quad f_1(x)=n(1-F(x))^{n-1}f(x). \tag{6}$$

(二) 多个次序统计量的联合分布

设 $1\leqslant r<s\leqslant n$, 要求$(X_{(r)}, X_{(s)})$的联合分布函数 $F_{rs}(x, y)$及密度函数 $f_{rs}(x,y)$, 方法可仿照上一段: 为要 $X_{(r)}<x$ 且 $X_{(s)}<y$, 需要在 $X_1, \cdots, X_n$ 中恰有 i 个小于 x, 有 $j-i$ 个小于 y, 但不小于 x, 有 $n-j$ 个不小于 y, 这里 $r\leqslant i\leqslant j, s\leqslant j\leqslant n$. 由此, 利用多项分布, 易得

$$F_{rs}(x,y)=\begin{cases}\displaystyle\sum_{j=s}^{n}\sum_{i=r}^{j}\frac{n!}{i!\,(j-i)!\,(n-j)!}F^i(x)\,[F(y)-F(x)]^{j-i} \\ \qquad\times[1-F(y)]^{n-j}, \quad x<y; \\ F_s(y), \quad x\geqslant y.\end{cases} \tag{7}$$

再通过求偏导数, 得到 $f_{rs}(x,y)$为

$$f_{rs}(x,y)=\begin{cases}\dfrac{n!}{(r-1)!\,(s-r-1)!\,(n-s)!}F^{r-1}(x)\,[F(y) \\ \qquad -F(x)]^{s-r-1}[1-F(y)]^{n-s}f(x)\,f(y), x<y; \\ 0, x\geqslant y.\end{cases} \tag{8}$$

这样的推导在理论上是严格的, 但略繁琐些. 一般说, 仅 $f_{rs}(x, y)$ 有较大的意义, 它可直接推导如下: 因为总有 $X_{(s)}\geqslant X_{(r)}$, 在 $x\geqslant y$ 时必有 $f_{rs}(x,y)=0$. 若 $x<y$, 则为了使 $X_{(r)}$落在$(x, x+\Delta x)$内, $X_{(s)}$ 落在$(y, y+\Delta y)$内, 需要在$(-\infty, x]$内有 $X_1, \cdots, X_n$ 中的 $r-1$ 个, 在$(x, x+\Delta x)$内有一个, 在$[x+\Delta x, y]$内有 $s-r-1$ 个, 在$(y, y+\Delta y)$内有一个, 而在$[y+\Delta y, \infty)$内有 $n-s$ 个. 另外的情况是: 在$(x, x+\Delta x)$或$(y, y+\Delta y)$中包含不止一个. 这种情况的概率是 $\Delta x\Delta y$ 的高价无穷小, 由多项分布得

$$f_{rs}(x,y)\,\Delta x\Delta y\approx P(x<X_{(r)}\leqslant x+\Delta x, y<X_{(s)}<y+\Delta y)$$

$$= \frac{n!}{(r-1)!1!(s-r-1)!1!(n-s)!} F^{r-1}(x)$$
$$\times [F(y)-F(x)]^{s-r-1}[1-F(y)]^{n-s}$$
$$\times f(x)f(y)\Delta x\Delta y + 0(\Delta x\Delta y).$$

两边除以 $\Delta x\Delta y$, 再令 $\Delta x\to 0$, $\Delta y\to 0$, 即得(8)式.

有一个重要情况是 $r=1$, $s=n$, 有

$$f_{1n}(x,y)=\begin{cases} n(n-1)(F(y)-F(x))^{n-2}f(x)f(y), & x<y;\\ 0, & x\geqslant y. \end{cases} \tag{9}$$

仿此可求出任意个次序统计量的联合分布函数与密度函数, 但这些没有多少实用意义, 故不列出了. 有一个值得注意的情况是: 全体 $(X_{(1)},\cdots,X_{(n)})$ 的密度函数 $f_{12\cdots n}(x_1,\cdots,x_n)$ 是:

$$f_{12\cdots n}(x_1,\cdots,x_n)=\begin{cases} n!f(x_1)\cdots f(x_n), & \text{当 } x_1<\cdots<x_n;\\ 0, & \text{其他}. \end{cases} \tag{10}$$

(三) 总体分布为 (0, 1) 内均匀分布的情况

当总体分布 F 为 $(0,1)$ 均匀分布 $R(0,1)$, 即有密度函数

$$f(x)=I_{(0,1)}(x)$$

时, 上两段的公式取很简单的形式. 例如(4)式成为

$$f_r(x)=\frac{n!}{(r-1)!(n-r)!}x^{r-1}(1-x)^{n-r}I_{(0,1)}(x). \tag{11}$$

(5)和(6)有相应的形式. 又(8)式成为

$$f_{rs}(x,y)=\begin{cases} \dfrac{n!}{(r-1)!(s-r-1)!(n-s)!}x^{r-1}(y-x)^{s-r-1}\\ \quad\times(1-y)^{n-s}, \quad 0<x<y<1;\\ 0, \text{其他}. \end{cases} \tag{12}$$

(9)式有相应的形式. 又全体 $(X_{(1)},\cdots,X_{(n)})$ 的密度函数为

$$f_{12\cdots n}(x_1,\cdots,x_n)=\begin{cases} n!, & 0<x_1<\cdots<x_n<1;\\ 0, & \text{其他}. \end{cases} \tag{13}$$

这个情况的重要性是基于下面的简单定理: