研究生（非数学类）数学系列规划教材

应用数理统计

陈平　等编著

MATHEMATICAL STATISTICS

机械工业出版社
CHINA MACHINE PRESS

研究生（非数学类）数学系列规划教材

应用数理统计

陈　平　吴诚鸥　刘应安　周圣武　王桂芝　编著
韦博成　主审

机 械 工 业 出 版 社

全书共有10章和1个附录，第1章介绍了概率论与矩阵代数的预备知识；第2章给出数理统计的基本概念；第3章和第4章是参数估计和假设检验，在大学相关内容的基础上作了适当的深化和扩充；第5章介绍多元回归、多项式回归、岭回归及 Logistic 回归等；第6章介绍一元和多元方差分析方法及常见的协方差分析模型；第7章介绍主成分分析与因子分析方法；第8章介绍典型相关分析；第9章给出判别分析与聚类分析方法；第10章介绍各种常见的时间序列分析方法；附录是SAS系统简介。其中第5章至第10章及附录的典型例题除了给出关键的数学模型外，还给出了SAS计算程序，便于自学和应用。本书精心选材，特别注重数理统计与实践的结合。书中详细阐述了如何运用SAS软件系统来分析、研究并解决实际工作中与现代数理统计有关的问题。

本书可作为各类院校的工科专业或其他非数学专业的研究生教材或大学高年级数理统计选修课教材，也可作为数理统计应用工作者的参考书籍。

本书的多媒体课件及各章SAS程序文件可以免费向出版社索取。

图书在版编目（CIP）数据

应用数理统计/陈平等编著．—北京：机械工业出版社，2008.8

研究生（非数学类）数学系列规划教材

ISBN 978-7-111-24890-3

Ⅰ．应…　Ⅱ．陈…　Ⅲ．数理统计—研究生—教材　Ⅳ．0212

中国版本图书馆 CIP 数据核字（2008）第124794号

机械工业出版社（北京市百万庄大街22号　邮政编码100037）

策划编辑：郑　丹　郑　玫　责任编辑：吴志明　版式设计：霍永明

责任校对：陈延翔　封面设计：王伟光　责任印制：杨　曦

三河市宏达印刷有限公司印刷

2008年9月第1版第1次印刷

169mm×239mm·25.25印张·488千字

标准书号：ISBN 978-7-111-24890-3

定价：34.00元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

销售服务热线电话：（010）68326294

购书热线电话：（010）88379639　88379641　88379643

编辑热线电话：（010）88379722

前　言

近几十年来，由于实际问题的需要和概率统计工作者的不懈努力，数理统计无论在理论方面还是在应用方面都得到了快速发展．它的基本理论和方法不仅为数学和概率统计专业所必需，而且也为工程技术、经济金融、农林科学及矿业气象等领域的研究和应用所需要．为适应21世纪形势的发展和社会的需要，信息技术与学科课程的整合已提到教育教学改革“重中之重”的地位，运用信息技术改造和优化传统学科内容是培养新世纪具有创新能力的高素质人才的必然要求．本书结合SAS软件系统介绍了现代数理统计的基本原理、基本方法及其在各个领域的应用．我们的想法是为非数学类专业的研究生编写一本“以方法为主，不追求理论的系统性和完整性，注重实用性和先进性，结构模块化，便于选用”的应用数理统计教材．我们根据近年来讲授这门课程的经验积累，参考国内外有关著作编写了这本教材．

在内容的选择和写作手法上我们力求体现上述思想，主要章节都配备了比较充分的典型例题，在叙述上均按由浅入深、由简到繁、循序渐进的模式展开，除了给出关键的数学模型外，还给出了SAS计算程序，便于自学和应用．本书内容充实、具有很强的实用性，其中部分例子就是作者的研究成果．众所周知，SAS软件是一个功能强、效率高、使用方便且适用于多种操作系统的信息处理和科学计算组合软件系统，已被广泛地应用于政府、科研、教育、生产和商业等不同领域．本书结合SAS软件阐述各类最常用的现代数理统计方法，希望通过该课程的学习，可使学生娴熟地运用SAS软件系统来分析、研究并解决实际工作中与现代数理统计有关的问题．另外，各章习题还配有答案或提示，便于学生检查提高．为了便于教学，作者还编制了简易实用的多媒体课件（建议结合自己的教学特点稍作修改后使用，那将取得更理想的效果）及各章SAS程序文件，从而使应用数理统计的内容直观化、图形化，可以消除同学们对统计的陌生感、抽象感及恐惧感，激活求知欲，增强学好统计、用好统计的信心．本书可以作为工科专业研究生数学公共课教材，也可用于其他非数学专业的研究生数理统计课程教材．

作为21世纪的更新教材，本书对于基本理论和方法的表述，从背景引伸到严格定义以及示例解释，均注重使用启发的方法．这些都有助于培养学生提出问题和解决问题的能力，同时，本书不过分探讨理论细节，而更注重统计思想和运用方法的论述，从而尽可能适应非数学专业研究生的需要．书中介绍了较多具有

一定深度和难度的现代数理统计方法，但限于学时和本书的特点，目前国际上非常流行的 Bayes 统计方法没有收录进来，读者如有兴趣可以参见参考文献［1］等资料．全书总计约有 60 学时的内容，教师上课时可以根据各自学校及专业的特点适当取舍．

本书的第 1 章和第 10 章及全书的多媒体课件由陈平编写，第 7 章～第 9 章及附录由吴诚鸥编写，第 2 章和第 6 章由刘应安编写，第 3 章和第 4 章由周圣武编写，第 5 章由王桂芝编写，全书由陈平教授统稿，由韦博成教授审稿．另外，书中的大部分 SAS 程序的编写都得到了吴诚鸥老师的帮助或参考了吴诚鸥老师等编写的有关书籍．书中的部分内容及例子还参考了国内外其他一些图书资料，谨此向有关的作者表示由衷的谢意！

由于作者水平有限，书中难免有不妥与谬误之处，敬请同行专家和广大读者提供宝贵的批评和建议，以使本教材不断得以完善．

东南大学数学系　陈　平
南京信息工程大学数学系　吴诚鸥
南京林业大学应用数学系　刘应安
中国矿业大学理学院　周圣武
南京信息工程大学数学系　王桂芝

目　　录

第 1 章　概率论与矩阵代数预备知识

在工程数学的概率论部分，已经对古典概型和几何概型定义了概率．在古典概型中，要求试验的可能结果是有限个且具有等可能性；对于几何概型，虽然试验的可能结果是无穷多个，但仍要求具有某种等可能性．然而实际问题中还有大量的随机试验，其结果并不属于这两种类型，因此，很有必要对一般的随机现象给出一个明确的概率定义．对于这个问题，经过人们长期探讨，并且随着测度论和积分理论的日益发展，终于在 1933 年由前苏联数学家柯尔莫哥洛夫（Kolmogorov）综合前人的成果，给出了概率论的公理化结构，明确了事件、概率等基本概念，从而使概率论成为一个严谨的数学分支，在此基础上，概率论与数理统计得到了迅速发展．本章主要介绍概率论与矩阵代数的一些预备知识，读者若想了解其中某些较深概念以及定理结论等的细节，可以参见文献［1］、［8］、［20］及［38］等文献资料．

1.1　概率空间

1.1.1　事件域

在给出概率定义之前，先要明确事件的概念．大家知道，事件是样本空间 Ω 的一个子集，但一般并不将 Ω 的一切子集都作为事件，比如在几何概率中就不能将不可度量的子集作为事件．实际上我们只要将具有某些限制而又相当广泛的一类 Ω 的子集作为事件就可以了．为此下面介绍事件域的概念．

定义 1.1　设 Ω 是样本空间，$\mathscr{F}$是由 Ω 的一些子集构成的集类，如果满足以下条件：

（1）$\Omega \in \mathscr{F}$；

（2）若 $A \in \mathscr{F}$，则 $\overline{A} = \Omega \setminus A \in \mathscr{F}$；

（3）若 $A_n \in \mathscr{F}, n = 1,2,\cdots$，则 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{F}$．

则称 $\mathscr{F}$为事件域或 σ 域（有的书上也称为 σ 代数）．（Ω，$\mathscr{F}$）称为可测空间，$\mathscr{F}$中的元素称为事件，Ω 也称为必然事件．

如果 $\mathscr{F}$ 是事件的 σ 域，则它具有下列性质：

（1）空集$\varnothing \in \mathscr{F}$．

这是因为$\varnothing = \overline{\Omega} \in \mathscr{F}$，称事件$\varnothing$为不可能事件.

(2) 若$A_i \in \mathscr{F}, i = 1,2,\cdots$，则$\bigcap\limits_{i=1}^{\infty} A_i \in \mathscr{F}$.

事实上，由德·莫根（De Morgan）公式可知：$\bigcap\limits_{i=1}^{\infty} A_i = \overline{\bigcup\limits_{i=1}^{\infty} \overline{A}_i} \in \mathscr{F}$.

(3) 若$A_i \in \mathscr{F}, i = 1,2,\cdots$，则$\bigcup\limits_{i=1}^{n} A_i \in \mathscr{F}, \bigcap\limits_{i=1}^{n} A_i \in \mathscr{F}$.

这只要在性质（2）和定义1.1的（3）中取$A_{n+1} = A_{n+2} = \cdots = \varnothing$即可.

(4) 若A，$B \in \mathscr{F}$，则$A \setminus B \in \mathscr{F}$.

这是因为$A \setminus B = A \cap \overline{B}$.

由σ域的定义，易证下面诸例中的集类$\mathscr{F}$都是σ域.

例 1.1　$\mathscr{F} = \{\varnothing, \Omega\}$.

例 1.2　$\mathscr{F} = \{\varnothing, A, \overline{A}, \Omega\}$.

例 1.3　$\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$，$\mathscr{F}$是由$\Omega$的一切子集构成的集类.

定义 1.2　设$\Omega = \mathbf{R} = (-\infty, +\infty)$，由所有半无限区间$(-\infty, x)$生成的$\sigma$域（$\sigma$代数）（即包含集族$\{(-\infty, x), x \in \mathbf{R}\}$的最小$\sigma$代数）称为$\mathbf{R}$上的Borel σ代数，记为$\mathscr{B}(\mathbf{R})$，其中的元素称为Borel集合. 类似可定义$\mathbf{R}^n$上的Borel σ代数$\mathscr{B}(\mathbf{R}^n)$.

1.1.2 概率

定义 1.3　设$(\Omega, \mathscr{F})$是可测空间，$P(\cdot)$是定义在$\mathscr{F}$上的实值函数. 如果它满足如下条件：

(1) 对任意$A \in \mathscr{F}$，有$0 \leqslant P(A) \leqslant 1$；

(2) $P(\Omega) = 1$；

(3) 若$A_i \in \mathscr{F}, i = 1, 2, \cdots$，且$A_iA_j = \varnothing, i \neq j$，则有

$P(\bigcup\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} P(A_i)$.　（这条性质称为可列可加性）

则称$P(A)$为事件A的概率，$(\Omega, \mathscr{F}, P)$称为概率空间.

由定义易见，事件的概率有如下性质：

(1) $P(\varnothing) = 0$；

(2) 若$A_i \in \mathscr{F}, i = 1,2,\cdots$，且$A_iA_j = \varnothing, i \neq j$，则有

$P(\bigcup\limits_{i=1}^{n} A_i) = \sum\limits_{i=1}^{n} P(A_i)$；　（这条性质称为有限可加性）

(3) 若$A \in \mathscr{F}$，则有$P(A) = 1 - P(\overline{A})$；

(4) 若A，$B \in \mathscr{F}$，且$B \subset A$，则有

$$P(A \setminus B) = P(A) - P(B),$$

$$P(B) \leqslant P(A);$$

(5) 若 $A, B \in \mathscr{F}$, 则有

$$P(A \cup B) \leqslant P(A) + P(B),$$

$$P(A \cup B) = P(A) + P(B) - P(AB).$$

如果概率空间 $(\Omega, \mathscr{F}, P)$ 的 P-零集（即零概率事件）的每个子集仍为事件，则称为完备的概率空间. 为了避免 P-零集的子集不是事件的情形出现，我们把概率测度完备化. 令 $\mathscr{N}$ 代表 Ω 的所有 P-零集的子集的全体. 由 $\{\mathscr{F}, \mathscr{N}\}$ 生成的 σ 代数（即包含 $\mathscr{F}$ 和 $\mathscr{N}$ 的最小 σ 代数）称为 $\mathscr{F}$ 的完备化，记为 $\widetilde{\mathscr{F}}$. $\widetilde{\mathscr{F}}$ 中的每个集合 F 都可以表示为 $F = A \cup N$，其中 $A \in \mathscr{F}, N \in \mathscr{N}$，且 $A \cap N = \emptyset$. 定义

$$\widetilde{P}(F) = \widetilde{P}(A \cup N) = P(A).$$

这样 P 就被扩张到 $\widetilde{\mathscr{F}}$ 上了. 易证 $\widetilde{P}$ 是 $\widetilde{\mathscr{F}}$ 上的概率测度，集函数 $\widetilde{P}$ 称为 P 的完备化，本书总假定 P 是完备的概率测度.

定义 1.4 设 $A_n \in \mathscr{F}, n \geqslant 1$. 所有属于无限多个集合 A_n 的 ω 的集合称为集列 $\{A_n\}$ 的上极限，记为 $\limsup\limits_{n \to \infty} A_n$. 可以证明

$$\limsup_{n \to \infty} A_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n.$$

上极限有时也记为 $\{\omega : \omega \in A_n, \text{i.o.}\}$. 集列 $\{A_n\}$ 的下极限定义为

$$\liminf_{n \to \infty} A_n = \{\omega \in \Omega : \exists\, n_0, \forall\, n > n_0, \omega \in A_n\}.$$

下极限有时也记为 $\{\omega : \omega \in A_n, \text{a.a.}\}$. 容易证明 $\liminf\limits_{n \to \infty} A_n = \bigcup\limits_{k=1}^{\infty} \bigcap\limits_{n=k}^{\infty} A_n$.

1.2 随机变量及其分布函数

定义 1.5 设 $(\Omega, \mathscr{F}, P)$ 是（完备的）概率空间，X 是定义在 Ω 上的单值实函数，如果对任一实数 x，$\{\omega : X(\omega) \leqslant x\} \in \mathscr{F}$，则称 $X(\omega)$ 是（$\mathscr{F}$ 上的）随机变量，并称

$$F(x) = P\{\omega : X(\omega) \leqslant x\} = P\{X \leqslant x\}, -\infty < x < +\infty$$

为随机变量 X 的分布函数.

由分布函数的定义可得：$P(a < X \leqslant b) = F(b) - F(a)$. 另外分布函数还有如下性质：

(1) $F(x)$ 是 x 的非降函数；

(2) $0 \leqslant F(x) \leqslant 1, F(-\infty) = 0, F(+\infty) = 1$；

(3) $F(x)$ 在任一点右连续，即 $F(x+0) = F(x)$.

还可以证明，满足上述三条性质的函数，必定是某个随机变量的分布函数.

如果有非负函数 $f(x)$，满足 $F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$，则称 $f(x)$ 为随机变量

X 或其分布函数 $F(x)$ 的密度函数．如果 X 具有密度函数，则称 X 为连续型随机变量；如果 X 最多以正概率取可数多个值，则称 X 为离散型随机变量．在实际问题中，常用的随机变量有两种类型：离散型随机变量和连续型随机变量．

离散型随机变量 X 的概率分布一般用分布律（列）来描述：

$$P\{X = x_k\} = p_k,\quad k = 1,2,\cdots,$$

其分布函数
$$F(x) = \sum_{x_k \leqslant x} p_k.$$

定义 1.6 设 $(\Omega,\mathscr{F},P)$ 是（完备的）概率空间，$X_1,X_2,\cdots,X_n$ 是定义在此概率空间上的 n 个随机变量，称 $(X_1,X_2,\cdots,X_n)$ 为 n 维随机向量．并称

$$F(x_1,x_2,\cdots,x_n) = P\{X_1 \leqslant x_1, X_2 \leqslant x_2,\cdots,X_n \leqslant x_n\},\ x_k \in \mathbf{R},\ 1 \leqslant k \leqslant n$$

为随机向量 $(X_1,X_2,\cdots,X_n)$ 的（联合）分布函数．

定理 1.1 若 $F(x_1,x_2,\cdots,x_n)$ 是联合分布函数，则有

(1) $F(x_1,x_2,\cdots,x_n)$ 对每个变量都是单调非降的；

(2) $F(x_1,x_2,\cdots,x_n)$ 对每个变量都是右连续的；

(3) 对 $i = 1,2,\cdots,n$，有

$$\lim_{x_i \to -\infty} F(x_1,\cdots,x_i,\cdots,x_n) = 0,\quad \lim_{x_1,\cdots,x_n \to +\infty} F(x_1,\cdots,x_n) = 1.$$

若存在非负函数 $f(x_1,\cdots,x_n)$，使得联合分布函数可表示为

$$F(x_1,x_2,\cdots,x_n) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n} f(t_1,\cdots,t_n)\mathrm{d}t_1\cdots\mathrm{d}t_n,$$

则称 $(X_1,X_2,\cdots,X_n)$ 为连续型随机向量，函数 $f(x_1,\cdots,x_n)$ 称为 $(X_1,X_2,\cdots,X_n)$ 的联合密度函数，它有如下性质：

$$f(x_1,\cdots,x_n) \geqslant 0;$$

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f(x_1,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n = 1;$$

对任意的区域 $D \in \mathbf{R}^n$，有

$$P\{(X_1,\cdots,X_n) \in D\} = \int\cdots\int_D f(x_1,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n.$$

设 $F(x_1,x_2,\cdots,x_n)$ 是 $(X_1,X_2,\cdots,X_n)$ 的联合分布函数．$1 \leqslant k_1 < k_2 < \cdots < k_d \leqslant n$，则 $(X_{k_1},X_{k_2},\cdots,X_{k_d})$ 的边际分布 $F_{k_1,\cdots,k_d}(x_{k_1},\cdots,x_{k_d})$ 定义为

$$F_{k_1,\cdots,k_d}(x_{k_1},\cdots,x_{k_d}) = F(+\infty,\cdots,+\infty,x_{k_1},+\infty,\cdots,+\infty,x_{k_d},+\infty,\cdots,+\infty).$$

如果随机向量 $(X_1,X_2,\cdots,X_n)$ 的所有可能取值为 $(x_{1k_1},x_{2k_2},\cdots,x_{nk_n})$，$k_1,k_2,\cdots,k_n = 1,2,\cdots$，则称概率

$$P(X_1 = x_{1k_1},X_2 = x_{2k_2},\cdots,X_n = x_{nk_n}) = p_{k_1,k_2,\cdots,k_n},\ k_1,k_2,\cdots,k_n = 1,2,\cdots$$

为 $(X_1,X_2,\cdots,X_n)$ 的联合分布律（列）．它有如下性质：

$$p_{k_1,k_2,\cdots,k_n} \geqslant 0;$$

$$\sum_{k_1} \cdots \sum_{k_n} p_{k_1,k_2,\cdots,k_n} = 1;$$

这时称 $(X_1,X_2,\cdots,X_n)$ 为 n 元离散型随机向量，$(X_1,X_2,\cdots,X_n)$ 的联合分布函数为

$$F(x_1,x_2,\cdots,x_n) = \sum_{x_{1k_1} \leqslant x_1} \cdots \sum_{x_{nk_n} \leqslant x_n} p_{k_1,k_2,\cdots,k_n}.$$

1.3　随机变量的独立性

定义 1.7　设 $X_1,X_2,\cdots,X_n$ 是定义在概率空间 $(\Omega,\mathscr{F},P)$ 上的 n 个随机变量，若对于任意实数 $x_1,x_2,\cdots,x_n$，有

$$P\{X_1 \leqslant x_1,X_2 \leqslant x_2,\cdots,X_n \leqslant x_n\} = P\{X_1 \leqslant x_1\}P\{X_2 \leqslant x_2\}\cdots P\{X_n \leqslant x_n\}, \tag{1.1}$$

则称 $X_1,X_2,\cdots,X_n$ 是相互独立的.

若设 X_i 的分布函数为 $F_i(x_i)$，$i=1,\cdots,n$，而 $X_1,X_2,\cdots,X_n$ 的联合分布函数为 $F(x_1,x_2,\cdots,x_n)$，则式 (1.1) 等价于

$$F(x_1,x_2,\cdots,x_n) = F_1(x_1)F_2(x_2)\cdots F_n(x_n). \tag{1.2}$$

定理 1.2　设 $X_1,X_2,\cdots,X_n$ 是离散型随机变量，则它们相互独立的充要条件是对任意一组可能值 $x_1,x_2,\cdots,x_n$，有

$$P\{X_1 = x_1,X_2 = x_2,\cdots,X_n = x_n\} = P\{X_1 = x_1\}P\{X_2 = x_2\}\cdots P\{X_n = x_n\}. \tag{1.3}$$

定理 1.3　设 $X_1,X_2,\cdots,X_n$ 是连续型随机变量，$f(x_1,x_2,\cdots,x_n)$ 及 $f_i(x_i),i=1,2,\cdots,n$ 分别为 $X_1,X_2,\cdots,X_n$ 的联合密度函数及 $X_i,i=1,2,\cdots,n$ 的边际密度函数，并设它们处处连续，则 $X_1,X_2,\cdots,X_n$ 相互独立的充要条件是

$$f(x_1,x_2,\cdots,x_n) = f_1(x_1)f_2(x_2)\cdots f_n(x_n). \tag{1.4}$$

定义 1.8　设 $(X_{11},X_{12},\cdots,X_{1n_1}),(X_{21},X_{22},\cdots,X_{2n_2}),\cdots,(X_{k1},X_{k2},\cdots,X_{kn_k})$ 为 k 个随机向量，如果对任意的实数 $x_{11},x_{12},\cdots,x_{1n_1}$，$x_{21},x_{22},\cdots,x_{2n_2},\cdots,x_{k1},x_{k2},\cdots,x_{kn_k}$，有

$$P\Big(\bigcap_{i=1}^{k}\{X_{i1} \leqslant x_{i1},\cdots,X_{in_i} \leqslant x_{in_i}\}\Big) = \prod_{i=1}^{k} P\{X_{i1} \leqslant x_{i1},\cdots,X_{in_i} \leqslant x_{in_i}\}, \tag{1.5}$$

则称这 k 个随机向量是相互独立的.

对于随机向量，也有类似于定理 1.2 和定理 1.3 的结论，此外，随机变量或随机向量的独立性还有如下性质：

(1) 若随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，则其中任意 $m(2\leqslant m\leqslant n)$ 个随机变量也相互独立；

(2) 若随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，则它们的 Borel 可测函数 $h_1(X_1),h_2(X_2),\cdots,h_n(X_n)$ 也相互独立；

(3) 若随机向量 $(X_1,X_2,\cdots,X_n)$ 与 $(Y_1,Y_2,\cdots,Y_m)$ 相互独立，则 $(X_1,X_2,\cdots,X_n)$ 的子向量与 $(Y_1,Y_2,\cdots,Y_m)$ 的子向量也相互独立；

(4) 若 $(X_{11},X_{12},\cdots,X_{1n_1}),(X_{21},X_{22},\cdots,X_{2n_2}),\cdots,(X_{k1},X_{k2},\cdots,X_{kn_k})$ 相互独立，则它们的 Borel 可测函数 $h_1(X_{11},X_{12},\cdots,X_{1n_1})$，$h_2(X_{21},X_{22},\cdots,X_{2n_2}),\cdots,h_k(X_{k1},X_{k2},\cdots,X_{kn_k})$ 也是相互独立的.

1.4 随机变(向)量函数的分布

当随机变量 X 的分布已知时，如何求它的函数 $Y=h(X)$ 的分布呢？当随机向量 $(X_1,X_2,\cdots,X_n)$ 的联合分布函数已知时，又如何求 $Y_1=h_1(X_1,\cdots,X_n),Y_2=h_2(X_1,\cdots,X_n),\cdots,Y_n=h_n(X_1,\cdots,X_n)$ 的联合分布函数呢？

下面我们仅就连续型随机变（向）量来说明这个问题.

1.4.1 单个随机变量函数的分布

在大学概率论与数理统计课程中已有如下的结果：

(1) 设 X 为连续型随机变量，其密度函数为 $f(x)$，若函数 $y=h(x)$ 严格单调，其反函数 $x=h^{-1}(y)$ 有连续导数. 则随机变量 $Y=h(X)$ 的密度函数为

$$f_Y(y)=\begin{cases} f[h^{-1}(y)]\,|\,[h^{-1}(y)]'\,|, & \alpha<y<\beta \\ 0, & 其他 \end{cases},$$

其中：$\alpha=\min\{f(-\infty),f(+\infty)\}$，$\beta=\max\{f(-\infty),f(+\infty)\}$.

进一步，可得如下结论：

(2) 若函数 $y=h(x)$ 在两两不交的区间 $I_1,I_2,\cdots$ 上逐段严格单调，其反函数分别为 $h_1^{-1}(y),h_2^{-1}(y)$，…且在各自区间上具有连续导数，则随机变量 $Y=h(X)$ 的密度函数为

$$f[h_1^{-1}(y)]\,\big|[h_1^{-1}(y)]'\big|+f[h_2^{-1}(y)]\,|\,[h_2^{-1}(y)]'\,|+\cdots.$$

1.4.2 单个随机向量函数的分布

设 $(X_1,X_2,\cdots,X_n)$ 是连续型随机向量，联合密度函数为 $f(x_1,x_2,\cdots,x_n)$，则 $Y=h(X_1,\cdots,X_n)$ 的分布函数为

$$G(y)=P(Y\leqslant y)=P\{h(X_1,\cdots,X_n)\leqslant y\}$$

$$= \underset{h(x_1,\cdots,x_n)\leqslant y}{\int\cdots\int} f(x_1,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n.$$

1.4.3　多个随机向量函数的分布

设 $(X_1,X_2,\cdots,X_n)$ 是连续型随机向量，联合密度函数为 $f(x_1,x_2,\cdots,x_n)$，则 $Y_1=h_1(X_1,\cdots,X_n)$，$Y_2=h_2(X_1,\cdots,X_n)$，…，$Y_n=h_n(X_1,\cdots,X_n)$ 的联合分布函数为

$$\begin{aligned} G(y_1,y_2,\cdots,y_n) &= P(Y_1\leqslant y_1,Y_2\leqslant y_2,\cdots,Y_n\leqslant y_n) \\ &= \underset{\substack{h_1(x_1,\cdots,x_n)\leqslant y_1\\ \vdots \\ h_n(x_1,\cdots,x_n)\leqslant y_n}}{\int\cdots\int} f(x_1,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n. \end{aligned}$$

设 $t_1=h_1(x_1,\cdots,x_n)$，$t_2=h_2(x_1,\cdots,x_n)$，…，$t_n=h_n(x_1,\cdots,x_n)$ 存在唯一一组反函数：$x_1=h_1^{-1}(t_1,\cdots,t_n)$，$x_2=h_2^{-1}(t_1,\cdots,t_n)$，…，$x_n=h_n^{-1}(t_1,\cdots,t_n)$；且 $t_i=h_i(x_1,\cdots,x_n)$ 及 $x_i=h_i^{-1}(t_1,\cdots,t_n),i=1,\cdots,n$ 连续并具有连续偏导数，则由积分变换可得 $(Y_1,Y_2,\cdots,Y_n)$ 的联合分布函数为

$$\begin{aligned} G(y_1,y_2,\cdots,y_n) &= \underset{\substack{h_1(x_1,\cdots,x_n)\leqslant y_1\\ \vdots \\ h_n(x_1,\cdots,x_n)\leqslant y_n}}{\int\cdots\int} f(x_1,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n \\ &= \int_{-\infty}^{y_1}\cdots\int_{-\infty}^{y_n} f[h_1^{-1}(t_1,\cdots,t_n),\cdots,h_n^{-1}(t_1,\cdots,t_n)]\,|J|\,\mathrm{d}t_1\cdots\mathrm{d}t_n, \end{aligned}$$

其中 J 为坐标变换的雅可比行列式：

$$J=\begin{vmatrix} \dfrac{\partial x_1}{\partial t_1} & \cdots & \dfrac{\partial x_1}{\partial t_n} \\ \vdots & & \vdots \\ \dfrac{\partial x_n}{\partial t_1} & \cdots & \dfrac{\partial x_n}{\partial t_n} \end{vmatrix},$$

这样就可以得到随机向量 $(Y_1,Y_2,\cdots,Y_n)$ 的联合密度函数为

$$g(y_1,y_2,\cdots,y_n)=\begin{cases} f[h_1^{-1}(y_1,\cdots,y_n),\cdots,h_n^{-1}(y_1,\cdots,y_n)]\,|J|, \\ \qquad 若(y_1,\cdots,y_n)属于(h_1,\cdots,h_n)的值域 \\ 0,\quad 其他 \end{cases}$$

其中
$$J=\begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \cdots & \dfrac{\partial x_1}{\partial y_n} \\ \vdots & & \vdots \\ \dfrac{\partial x_n}{\partial y_1} & \cdots & \dfrac{\partial x_n}{\partial y_n} \end{vmatrix}.$$

与前面类似，若 $y_1 = h_1(x_1,\cdots,x_n)$，$y_2 = h_2(x_1,\cdots,x_n)$，…，$y_n = h_n(x_1,\cdots,x_n)$ 在两两不交的区域 $D_1, D_2, \cdots$ 中分别存在单值反函数，并满足一定的条件，则随机向量 $(Y_1, Y_2,\cdots, Y_n)$ 的联合密度函数为

$$f[h_{11}^{-1}(y_1,\cdots,y_n),\cdots,h_{1n}^{-1}(y_1,\cdots,y_n)]\,|J_1| +$$
$$f[h_{21}^{-1}(y_1,\cdots,y_n),\cdots,h_{2n}^{-1}(y_1,\cdots,y_n)]\,|J_2| + \cdots$$

例 1.4 设二元随机向量 (X, Y) 的联合密度函数为

$$f(x,y) = \frac{1}{2\pi}\mathrm{e}^{-\frac{1}{2}(x^2+y^2)},\ -\infty < x < +\infty,\ -\infty < y < +\infty,$$

试求 $U = X^2 + Y^2$ 与 $V = X/Y$ 的联合密度函数，并判断 U 与 V 是否相互独立.

解 令 $\begin{cases} u = x^2 + y^2 \\ v = \dfrac{x}{y} \end{cases}$，当 $u > 0$ 时，它的反函数有两个分支：

$$\begin{cases} x = v\sqrt{\dfrac{u}{1+v^2}} \\ y = \sqrt{\dfrac{u}{1+v^2}} \end{cases},\quad \begin{cases} x = -v\sqrt{\dfrac{u}{1+v^2}} \\ y = -\sqrt{\dfrac{u}{1+v^2}} \end{cases}.$$

这两个变换的雅可比行列式的绝对值均为

$$|J| = \left\|\begin{matrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{matrix}\right\| = \left\|\begin{matrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{matrix}\right\|^{-1} = \left\|\begin{matrix} 2x & 2y \\ \dfrac{1}{y} & -\dfrac{x}{y^2} \end{matrix}\right\|^{-1}$$
$$= \left[2\left(1+\frac{x^2}{y^2}\right)\right]^{-1} = \frac{1}{2\ (1+v^2)},$$

从而 (U, V) 的联合密度函数为

$$g(u,v) = 2\frac{1}{2\pi}\mathrm{e}^{-\frac{1}{2}u} \times \frac{1}{2(1+v^2)} = \frac{1}{2\pi}\mathrm{e}^{-\frac{1}{2}u} \times \frac{1}{1+v^2},\ u > 0,\ -\infty < v < +\infty,$$

而关于 U 的边际密度函数为

$$g_U(u) = \int_{-\infty}^{+\infty} g(u,v)\mathrm{d}v = \frac{1}{2}\mathrm{e}^{-\frac{u}{2}},\ u > 0,$$

关于 V 的边际密度函数为

$$g_V(v) = \int_0^{+\infty} g(u,v)\mathrm{d}u = \frac{1}{\pi(1+v^2)},\ -\infty < v < +\infty.$$

注意到 $g(u,v) = g_U(u) \times g_V(v)$，所以随机变量 U 与 V 相互独立.

1.5　黎曼—斯蒂尔切斯（Riemann-Stieltjes）积分

定义 1.9　设 $h(x), F(x)$ 为有限区间 $[a,b]$ 上的实值函数，$a = x_0 < x_1 < \cdots < x_n = b$ 为 $[a,b]$ 的一个分割，令

$$\Delta F(x_i) = F(x_i) - F(x_{i-1}),$$
$$\xi_i \in [x_{i-1}, x_i], 1 \leqslant i \leqslant n,$$
$$\lambda = \max_{1 \leqslant i \leqslant n}(x_i - x_{i-1}),$$

如果当 $\lambda \to 0$ 时，极限 $\lim\limits_{\lambda \to 0} \sum\limits_{i=1}^{n} h(\xi_i)\Delta F(x_i)$ 存在，且与分割的选择以及 $\xi_i \in [x_{i-1}, x_i]$ 的取法无关，则称该极限为函数 $h(x)$ 关于 $F(x)$ 在 $[a,b]$ 上的黎曼-斯蒂尔切斯积分，简称 R-S 积分，记为

$$\int_a^b h(x)\mathrm{d}F(x) = \lim_{\lambda \to 0} \sum_{i=1}^{n} h(\xi_i)\Delta F(x_i). \tag{1.6}$$

易见，当 $F(x) = x$ 时，式（1.6）即为黎曼积分；当 $F'(x) = f(x)$ 存在时，式（1.6）成为黎曼积分：$\int_a^b h(x)f(x)\mathrm{d}x$.

关于 R-S 积分的存在条件，我们给出一个充分条件：若函数 $h(x)$ 连续，$F(x)$ 单调，则 R-S 积分（1.6）存在. 为了后面的需要，再将 R-S 积分的定义推广到无限区间上：

$$\int_a^{+\infty} h(x)\mathrm{d}F(x) \xlongequal{\text{def}} \lim_{b \to +\infty}\int_a^b h(x)\mathrm{d}F(x),$$
$$\int_{-\infty}^{b} h(x)\mathrm{d}F(x) \xlongequal{\text{def}} \lim_{a \to -\infty}\int_a^b h(x)\mathrm{d}F(x),$$
$$\int_{-\infty}^{+\infty} h(x)\mathrm{d}F(x) \xlongequal{\text{def}} \lim_{\substack{b \to +\infty \\ a \to -\infty}}\int_a^b h(x)\mathrm{d}F(x).$$

易见 R-S 阶积分具有和黎曼积分类似的性质，例如

（1）积分的线性性质：

$$\int_a^b [\alpha h_1(x) \pm \beta h_2(x)]\mathrm{d}F(x) = \alpha\int_a^b h_1(x)\mathrm{d}F(x) \pm \beta\int_a^b h_2(x)\mathrm{d}F(x);$$

（2）积分对于区间的可加性：

$$\int_a^b h(x)\mathrm{d}F(x) = \int_a^c h(x)\mathrm{d}F(x) + \int_c^b h(x)\mathrm{d}F(x);$$

（3）$\int_a^b \mathrm{d}F(x) = F(b) - F(a)$,

其中 a，b 均可为有限数或无穷大. 另外，R-S 积分与黎曼积分不同的是