21世纪经济学研究生规划教材

# Advanced Mathematical Statistics

# 高等数理统计

苏良军 ◎编著

北京大学出版社
PEKING UNIVERSITY PRESS

21世纪经济学研究生规划教材

# Advanced Mathematical Statistics

# 高等数理统计

苏良军 ◎编著

北京大学出版社
PEKING UNIVERSITY PRESS

**图书在版编目(CIP)数据**

高等数理统计/苏良军编著. —北京:北京大学出版社,2007.9
(21世纪经济学研究生规划教材)
ISBN 978-7-301-12311-9

Ⅰ.高…　Ⅱ.苏…　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(2007)第080671号

**书　　名:高等数理统计**
著作责任者:苏良军　编著
策 划 编 辑:朱启兵
责 任 编 辑:朱启兵
标 准 书 号:ISBN 978-7-301-12311-9/F·1653
出 版 发 行:北京大学出版社
地　　址:北京市海淀区成府路205号　100871
网　　址:http://www.pup.cn　电子邮箱:em@pup.pku.edu.cn
电　　话:邮购部62752015　发行部62750672　编辑部62752926　出版部62754962
印 刷 者:北京宏伟双华印刷有限公司
经 销 者:新华书店
787毫米×1092毫米　16开本　20印张　449千字
2007年9月第1版　2007年9月第1次印刷
印　　数:0001—4000册
定　　价:38.00元

# 文章有神交有道(代序)

微观经济学、宏观经济学与计量经济学是西方经济学中的三大核心领域,其中计量经济学所提供的各种计量方法,更是任何经济实证研究中不可或缺的分析工具。由于中国的经济正处于快速发展与变动的时期,各种经济现象与数据正好为经济实证研究提供了丰富的素材。英雄既然有用武之地,计量经济学遂为经济学家(乃至统计学家与数学家)所重视,也因此成为经济学领域中的“显学”。

近年来中国的计量经济学教育与研究的进步是极为显著的,而海外中国学者于此功不可没。20 世纪 80 年代中期以后,中国负笈西方学习经济的学生们陆续取得博士学位,许多人在计量理论与方法上推陈出新,而成为望重一方的计量学者。这些海外学者们在回国时带来了各种最新的计量理论知识,进一步提高了计量经济学的水平。而海外学者在学术界的成功,也吸引了更多学生与学者们投入此一领域。

在计量经济学的发展形势看来一片大好的时候,我们也许要问:什么是计量经济学在中国的下一阶段?任何一种学术领域的高度发展当然不能只仰赖海外偶尔传来的养分,而必须持续培养自己的学术土壤,学术才可能真正生根发荣。近年来已有一些学者在学成或在海外工作一段时间后返国,全时全心地投入学术奠基工作。良军正是这批返国投身学术的计量经济学家之一。

计量经济学植基于统计,但又因经济议题和经济数据的特性而有许多不同的发展。所以在为计量经济学扎根的教学上,不仅要重视一般的统计观念与方法,还应当兼顾计量经济学的特殊性与新发展。然而坊间的教材,于此各有偏废,甚至各行其道。因此,大家都期待有一本更完整,更全面的计量经济学入门书。

良军的《高等数理统计》一书结合了统计基础知识与计量理论,一方面呈现了统计学与计量经济学之间的清楚脉络,一方面则涵盖了许多计量经济学中的重要议题。例如,书中除了讨论大家熟悉的统计学中的“极大似然估计”外,也着重介绍了计量经济学中所重视的在错误模型下的“准极大似然估计”以及“广义矩估计”。除此之外,书中也包括了计量经济学的前沿议题,如“轮廓似然估计”、贝叶斯估计中的“马尔可夫链-蒙特卡罗法”和非参数模型中的“经验似然比检验”等。这些都是本书的特殊之处,由此亦可见良军学识之广与他对本书用力之深。而我也相信,这本书必将成为许多有志学习计量经济学的学生们必读之书。

我和良军谊属同门,先后受教于加州大学圣地亚哥分校的 Halbert White 教授。闻道有先后,术业有专攻,良军的学术专业一直走在学术前沿,毕业后三年有成,其表现遂为学界瞩目。而我更敬佩他返国工作的决心,以及他在这么短时间内对中国计量经济学教育与研

究所做出的贡献。良军的成就，我与有荣焉；因为有他以及许多计量学者的努力付出，我对中国计量经济学未来的发展更寄以厚望，也更具信心。

良军出身西北，我则来自东南，两人结识于北京。工部诗云："文章有神交有道"，正足以说明我对良军以及他这本《高等数理统计》的看法。

谨以此代序。

管中闵

丁亥夏月于台湾台北"中央研究院"

# 前　　言

本书是为国内众多商学院与经济管理学院的统计学、经济、金融、管理等专业的研究生与统计工作者精心编写的统计教科书。本书结合国内外统计系、经济系与商学院研究生数理统计教育的现状，系统性地介绍现代统计的基础理论与基本方法，充分反映当代统计的发展，力求做到理论与实际的结合。本书既可作为统计、计量经济及相关专业的教科书，又可作为统计与计量经济工作者必备的理论参考书。

本书包括十五章与一个附录。前三章重点介绍现代统计所需的概率理论、分布理论与渐进理论，第 4 章介绍现代统计的数据降维思想。第 5 章到第 8 章介绍现代统计的各种估计理论，依次包括极大似然估计、准极大似然估计、矩估计与广义矩估计、贝叶斯估计。第 9 章到第 11 章介绍现代统计的假设检验理论，依次包括假设检验的基本理论、参数模型检验、非参数模型检验。第 12 章介绍区间估计的基本理论，第 13 章介绍方差分析，第 14 章介绍回归分析的基本理论，第 15 章介绍回归分析的高级理论与应用(包括结构突变的检验、分块回归、多重共线性、广义最小二乘估计、异方差、工具变量估计等)。我们在附录里简单介绍了本书所需要的矩阵知识。

和国内外其他同类数理统计教材相比，本书主要基于编者在国外多家高校统计系与经济系的学习经历以及在北京大学光华管理学院的执教经历，在参照十几本中英文数理统计及相关学科教科书与数以百计的学术文章后，精心编著而成。独到之处体现在：

(1) 考虑到国内外统计教学的接轨、国内商学院与经济管理学院的数理统计教育需求，系统性地阐述现代数理统计的基础理论与基本方法，并尽量使用商业与经济管理中的案例。全书既可作为研究生水平的数理统计课程讲授，又可作为研究生水平计量经济学的第一门课程讲授。

(2) 编入了当代统计的最新研究成果，如估计理论中的准极大似然估计、广义矩估计与经验似然估计，检验中的经验似然比检验，回归中的内生性问题等。

(3) 本书使用一定数量的高质量习题，既可作教科书，也可以供学生自学使用。为方便学生自学使用，本书的多数习题答案将在编者的个人网页上公布。

由于本书编者长期在国外求学，全书先是用英文写成，再由我的两个研究生俞炀与王芸翻译成中文，俞炀主要负责全书第 1 章到第 8 章的翻译，王芸主要负责全书第 9 章到第 15 章以及附录的翻译。她们在书稿录入与排版方面也做了大量的工作，编者在此对她们的辛勤劳动表示诚挚的谢意！过去三年，编者在北京大学光华管理学院使用了本书的初稿，许多学生对本书的编写提供了有益的帮助，包括陈一侨、王蕾、瞿茜、章永辉等，编者在此一并向

他们表示衷心的感谢！

本书的出版得到国家自然科学基金资助项目(70501001,70601001)的大力资助与北京大学出版社的大力支持，编者感谢林君秀老师和朱启兵老师的鼎力帮助，并对他们的敬业精神深表由衷敬佩。

最后，编者感谢北京大学光华管理学院金赛男老师给本书提供的中肯意见，并特别感谢台湾“中央研究院”的管中闵教授在百忙之中为本书作序。

由于编者与译者水平有限，加之编写时间仓卒，书中错误在所难免，恳请国内同行、专家、学者和读者批评指正。

苏良军

2007年8月

于北京大学光华管理学院商务统计与计量经济系

# 目　　录

# 第 1 章　概率、随机变量与分布

## 1.1 概 率 空 间

概率空间(probability space)由三元参数组$(\Omega,\mathscr{F},\mathbf{P})$所定义，对各参数的描述如下.

首先，$\Omega$ 被称为样本空间(sample space)，其定义如下：

**定义 1.1(样本空间)**　一个随机试验的所有可能结果的集合被称为该随机试验的样本空间.

随机试验(random experiment)指在相同条件下可以重复进行、在每次试验之前不知道准确结果、然而其大量重复试验的结果会呈现出某种规律性的试验. 每一个可能出现的试验结果(outcome)即是样本空间 $\Omega$ 的一个元素 $\omega$.

**例 1.1(投掷硬币)**　在一个投掷硬币的简单试验中，该次随机试验的样本空间由两个可能出现的试验结果构成，即正面朝上和反面朝上. 令 $H$ 代表正面朝上，$T$ 代表反面朝上，则 $\Omega=\{H,T\}$.

**例 1.2(测量身高)**　从一所大学中随机抽取学生测量身高(假设人的身高不超过 2.4 m)，该次随机试验的样本空间由 0 至 2.4 之间的所有实数构成，即 $\Omega=(0,2.4]$.

参数组$(\Omega,\mathscr{F},\mathbf{P})$中的第二个参数 $\mathscr{F}$ 被称为 $\sigma$-域($\sigma$-field)，它是样本空间 $\Omega$ 的子集的集合，对其更精确的定义如下：

**定义 1.2($\sigma$-域)**　样本空间 $\Omega$ 的子集集合 $\mathscr{F}$ 要成为 $\sigma$-域，必须满足如下性质：

(1) $\Omega\in\mathscr{F}$;

(2) 如果 $E\in\mathscr{F}$，则 $E^c\in\mathscr{F}$;

(3) 如果 $E_1,E_2,\cdots\in\mathscr{F}$，则 $\bigcup_{i=1}^{\infty}E_i\in\mathscr{F}$.

其中 $E^c$ 为 $E$ 在 $\Omega$ 中的补集：$E^c=\Omega-E$.

$E$ 为 $\mathscr{F}$ 中的任一个元素，即样本空间 $\Omega$ 的某个子集，通常称它为一个事件(event)，在文献中一般用 $A,B,C,\cdots$ 或者 $A_1,A_2,A_3,\cdots$ 表示. 样本空间 $\Omega$ 也是 $\sigma$-域 $\mathscr{F}$ 的一个元素. 当 $\Omega$ 的子集集合 $\mathscr{F}$ 满足上述条件(1)和(2)，且其任意有限个子集的并集还在 $\mathscr{F}$ 里时，称之为域.

**注 1**　由于空集 $\varnothing=\Omega^c$，因此 $\varnothing\in\mathscr{F}$;

**注 2**　由于 $\bigcap_{i=1}^{\infty}E_i=\left(\bigcup_{i=1}^{\infty}E_i^c\right)^c$，当 $E_1,E_2,\cdots\in\mathscr{F}$，则 $\bigcap_{i=1}^{\infty}E_i\in\mathscr{F}$;

**注 3**　当 $\Omega$ 为有限集时，$\sigma$-域与域是等价的概念.

**例 1.3(投掷硬币，续 1)**　当 $\Omega=\{H,T\}$，此时 $\mathscr{F}_1=\{\Omega,\varnothing\}$ 是最小的 $\sigma$-域. 而 $\mathscr{F}_1=\{\Omega,\{H\},\{T\},\varnothing\}$ 也是一个 $\sigma$-域.

**例 1.4(波雷尔 $\sigma$-域)**　当 $\Omega$ 为实轴 $\mathbb{R}$ 时，我们可以定义一个 $\sigma$-域 $\mathscr{B}(\mathbb{R})$，它包含所有像

$$(a,b),(a,b],[a,b),[a,b]$$

的集合，这里的 $a$ 和 $b$ 为任意实数. 从 $\sigma$-域的性质可以得到，$\mathscr{B}(\mathbb{R})$包含上述集合的所有并集和交集(可能是无限个). 当定义域为实轴$\mathbb{R}$时，$\mathscr{B}(\mathbb{R})$被称为波雷尔(Borel)$\sigma$-域，它可以视为由$\mathbb{R}$上所有开区间(或闭区间)的集合所生成的$\sigma$-域.

参数组$(\Omega,\mathscr{F},\mathbf{P})$中的第三个参数 $\mathbf{P}$ 是一个定义在 $\sigma$-域 $\mathscr{F}$ 上的实函数.

**定义 1.3(概率)** 定义在 $\sigma$-域 $\mathscr{F}$ 上的实函数 $\mathbf{P}$ 被称为概率或概率测度，它必须满足以下性质：

(a) $\mathbf{P}(E)\geqslant 0$ 对一切 $E\in\mathscr{F}$ 成立；

(b) $\mathbf{P}(\Omega)=1$；

(c) 如果 $E_1,E_2\cdots\in\mathscr{F}$ 两两互斥，即 $E_i\cap E_j=\varnothing$ 对一切 $i\neq j$ 成立，则 $\mathbf{P}(\bigcup\limits_{i=1}^{\infty}E_i)=\sum\limits_{i=1}^{\infty}\mathbf{P}(E_i)$.

上述三条性质通常被视为概率论公理(axioms of probability). 现代概率论把概率看作是满足这三条公理的函数.

**例 1.5(投掷硬币，续 2)** 在一个投掷硬币的随机试验中，样本空间$\Omega=\{H,T\}$. 假设硬币铸造均匀，即 $\mathbf{P}(\{H\})=\mathbf{P}(\{T\})=0.5$.

由概率论公理可以直接推导出很多熟知的概率性质.

**定理 1.1(概率性质)** 已知 $A,A_1$ 及 $A_2$ 为 $\mathscr{F}$ 中的任意子集，则

(1) $\mathbf{P}(A^c)=1-\mathbf{P}(A)$；

(2) $\mathbf{P}(\varnothing)=0$；

(3) (单调性)当 $A_1\subset A_2$ 时，$\mathbf{P}(A_1)\leqslant\mathbf{P}(A_2)$；

(4) $0\leqslant\mathbf{P}(A)\leqslant 1$；

(5) (加法定理)$\mathbf{P}(A_1\cup A_2)=\mathbf{P}(A_1)+\mathbf{P}(A_2)-\mathbf{P}(A_1\cap A_2)$.

在证明(1)时，把样本空间写成 $\Omega=A+A^c$，则 $1=\mathbf{P}(\Omega)=\mathbf{P}(A)+\mathbf{P}(A^c)$；证明(3)时，由于 $A_2=A_1\cup(A_2-A_1)$，则 $\mathbf{P}(A_2)=\mathbf{P}(A_1)+\mathbf{P}(A_2-A_1)\geqslant\mathbf{P}(A_1)$；证明(5)时，由于 $A_1\cup A_2=A_1\cup(A_2\cap A_1^c)$，且 $A_2=(A_2\cap A_1)\cup(A_2\cap A_1^c)$，则$\mathbf{P}(A_1\cup A_2)=\mathbf{P}(A_1)+\mathbf{P}(A_2\cap A_1^c)=\mathbf{P}(A_1)+\mathbf{P}(A_2)-\mathbf{P}(A_1\cap A_2)$.

**例 1.6(庞弗罗尼不等式)** 在计算交集概率时，当一个确切的交集概率值不容易得到甚至无法计算，然而又需要关于概率大小的描述时，庞弗罗尼(Bonferroni)不等式可以发挥重要的作用. 令 $A_1$ 和 $A_2$ 均是概率为 0.95 的事件，则两个事件同时发生的概率范围为

$$\mathbf{P}(A_1\cap A_2)\geqslant\mathbf{P}(A_1)+\mathbf{P}(A_2)-1=0.95+0.95-1=0.90.$$

注意，倘若每个事件自身发生的概率不是足够大，庞弗罗尼边界很有可能只是一个没用的负值.

在本节的末尾，定义互斥事件(mutually exclusive events)为，对一切 $i\neq j$，事件 $A_i,i=1,2,\cdots$满足条件 $A_i\cap A_j=\varnothing$. 当 $\bigcup\limits_{i}^{\infty}A_i=\Omega$ 时，又称它们为完备事件(exhaustive events).

## 1.2 条件概率与独立性

令概率空间为$(\Omega,\mathscr{F},\mathbf{P})$. 以下的讨论均基于该概率空间,在下文不再重复说明.

**定义 1.4(条件概率)** 令 $A,B\in\mathscr{F}$且 $\mathbf{P}(A)>0$,则在事件 $A$ 发生后,事件 $B$ 发生的条件概率为 $\mathbf{P}(B|A)=\mathbf{P}(B\cap A)/\mathbf{P}(A)$.

显然,对每一个固定的 $A$,函数 $\mathbf{Q}(\cdot)=\mathbf{P}(\cdot|A)$满足概率论的三条公理. 容易证明,对任一事件 $E$,$\mathbf{Q}(E)\geqslant 0$ 且 $\mathbf{Q}(\Omega)=1$. 对互斥事件 $E_1,E_2,\cdots$,有$\mathbf{Q}(\bigcup_{i=1}^{\infty}E_i)=\sum_{i=1}^{\infty}\mathbf{Q}(E_i)$. 同时 $\mathbf{Q}$ 作为一个概率,也满足上文推出的所有概率性质. 例如 $\mathbf{Q}(\varnothing)=0$,$\mathbf{Q}(E^c)=1-\mathbf{Q}(E)$,即 $\mathbf{P}(\varnothing|A)=0$,$\mathbf{P}(E^c|A)=1-\mathbf{P}(E|A)$.

通过条件概率可以方便地定义交集概率,例如 $\mathbf{P}(B\cap A)=\mathbf{P}(A)\mathbf{P}(B|A)$. 设 $A_1,A_2,\cdots$ 为 $\mathscr{F}$ 中的事件,容易证明

$$\mathbf{P}(A_1\cap A_2\cap\cdots\cap A_n)=\mathbf{P}(A_1)\mathbf{P}(A_2\mid A_1)\mathbf{P}(A_3\mid A_1,A_2)\cdots\mathbf{P}(A_n\mid A_1,\cdots,A_{n-1}). \tag{1.1}$$

式(1.1)被称为条件概率的乘法准则(multiplication rule).

设$\{A_i\}$为 $\Omega$ 的一个划分(partition)($\{A_i\}$为互斥且完备事件,即对一切$i\neq j$, $\bigcup_i A_i=\Omega$ 且 $A_i\cap A_j=\varnothing$),$\mathbf{P}(A_i)>0$,$i=1,2,\cdots$,则对任意事件 $B\in\mathscr{F}$,有

$$\mathbf{P}(B)=\sum_{i=1}^{\infty}\mathbf{P}(B\cap A_i)=\sum_{i=1}^{\infty}\mathbf{P}(B\mid A_i)\mathbf{P}(A_i), \tag{1.2}$$

上式在文献中通常被称为全概率公式(law of total probability). 由它可以变形得到如下相关结论.

**定理 1.2(贝叶斯定理)** 设$\{A_i\}$为 $\Omega$ 的一个划分,$B$ 为 $\Omega$ 中任意一个集合,$\mathbf{P}(B)>0$. 则对 $i=1,2,\cdots$,有

$$\mathbf{P}(A_i\mid B)=\frac{\mathbf{P}(B\mid A_i)\mathbf{P}(A_i)}{\sum_{j=1}^{\infty}\mathbf{P}(B\mid A_j)\mathbf{P}(A_j)}.$$

**例 1.7(两个黑球和一个白球)** 一个篮子里有两个黑球和一个白球,令两个人闭着眼睛按顺序从篮子里无放回地分别拿出一个球. 抽到黑球的人会赢得 100 美元的奖金,抽到白球的则什么也不能得到. 分别对这两个人计算他们赢的概率,令事件 $A$=\{第一个人拿到黑球\},事件 $B$=\{第二个人拿到黑球\}. 显然,$\mathbf{P}(A)=2/3$,且

$$\begin{aligned}\mathbf{P}(B)&=\mathbf{P}(B\mid A)\mathbf{P}(A)+\mathbf{P}(B\mid A^c)\mathbf{P}(A^c)\\&=\frac{1}{2}\times\frac{2}{3}+1\times\frac{1}{3}\\&=\frac{2}{3}.\end{aligned}$$

因此,不论抽签的顺序怎样,两人赢得奖金的概率是相等的.

**定义 1.5(独立性)** 令 $A,B\in\mathscr{F}$,如果 $\mathbf{P}(A\cap B)=\mathbf{P}(A)\mathbf{P}(B)$成立,称事件 $A$ 与 $B$ 在统

计意义上独立(两两独立).

与上述独立定义等价的另一种表达方式为:$\mathbf{P}(A|B)=\mathbf{P}(A)$,或者$\mathbf{P}(B|A)=\mathbf{P}(B)$. 其中$\mathbf{P}(B)>0$,或 $\mathbf{P}(A)>0$. 有关独立的性质如下:

**定理 1.3(独立性)** 如果事件 $A$ 与 $B$ 独立,则

(1) 事件 $A$ 与 $B^c$ 独立;

(2) 事件 $A^c$ 与 $B$ 独立;

(3) 事件 $A^c$ 与 $B^c$ 独立.

对任意事件 $A_i$, $i=1,2,\cdots$,当对任意 $k\geqslant 2$,有

$$\mathbf{P}(A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_k})=\prod_{j=1}^{k}\mathbf{P}(A_{i_j}),\tag{1.3}$$

则称事件 $A_i$ 相互独立(mutually independent). 其中$(i_1,\cdots,i_k)$是从$\{1,2,3,\cdots\}$中任意取值的 $k$ 个下标. 注意,相互独立比两两独立(pairwise independence)的约束要强. 当 $\mathbf{P}(A_1\cap A_2\cap A_3)=\mathbf{P}(A_1)\mathbf{P}(A_2)\mathbf{P}(A_3)$时,$A_1$,$A_2$ 和 $A_3$ 不一定相互独立.

**例 1.8(相互独立与两两独立)** 假设 $X$ 与 $Y$ 为两两独立的随机变量,其取值为 1 或者 $-1$,取值概率各为 0.5. 令 $Z=XY$,则 $Z$ 与 $X$,$Y$ 分别两两独立,然而 $X$,$Y$,$Z$ 之间并不相互独立.

$$\begin{aligned}\mathbf{P}(X=1,Y=1,Z=1)&=\mathbf{P}(Z=1\mid X=1,Y=1)\mathbf{P}(X=1,Y=1)\\&=\mathbf{P}(X=1,Y=1)=\frac{1}{4}.\end{aligned}$$

而

$$\mathbf{P}(X=1)\mathbf{P}(Y=1)\mathbf{P}(Z=1)=\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}=\frac{1}{8}.$$

下面的例子摘自 Casella 与 Berger(2002,p. 26)的著作.

**例 1.9(相互独立与两两独立)** 该次随机试验的样本空间 $\Omega$ 为三个字母 a,b,c 的三重组合

$$\Omega=\{\text{aaa bbb ccc abc bca cba acb bac cab}\}.$$

令样本空间 $\Omega$ 中的每一元素的发生概率为 1/9. 对 $i=1,2,3$,有

$$A_i=\{\text{每个三重组中的第 } i \text{ 个位置为字母 a}\}.$$

显然 $\mathbf{P}(A_i)=1/3$, $i=1,2,3$. 且

$$\mathbf{P}(A_1\cap A_2)=\mathbf{P}(A_1\cap A_3)=\mathbf{P}(A_2\cap A_3)=\frac{1}{9}.$$

上行说明 $A_i$ 之间是两两独立的. 然而

$$\mathbf{P}(A_1\cap A_2\cap A_3)=\frac{1}{9},$$

而

$$\mathbf{P}(A_1)\mathbf{P}(A_2)\mathbf{P}(A_3)=\frac{1}{27}.$$

因此 $A_i$ 之间并不是相互独立的.

## 1.3 随机变量与分布

对于许多随机试验，与其对概率空间上的原始概率结构进行分析，不如对概率结构进行归纳后的随机变量进行分析来得更为方便.

**定义 1.6(随机变量)** 随机变量 $X$ 可以视为从样本空间 $\Omega$ 到实轴 $\mathbb{R}$ 的一个可测函数.

**注 1** 一个随机变量 $X$ 是从样本空间 $\Omega$ 到实轴 $\mathbb{R}$ 的一个函数，比如它可以对 $\Omega$ 中的每一个可能试验结果指派一个数值.

**注 2** $X$ 是可测度的，即对任意 $A\in\mathscr{B}(\mathbb{R})$，$X^{-1}(A)=\{\omega\in\Omega:X(\omega)\in A\}\in\mathscr{F}$. 实际上，要证明 $X$ 是定义在$(\Omega,\mathscr{F},\mathbf{P})$上的随机变量，只须证明对任意的$x\in\mathbb{R}$，$\{X\leqslant x\}\equiv\{\omega\in\Omega:X(\omega)\leqslant x\}\in\mathscr{F}$.

**例 1.10(投掷硬币，续 3)** 在投掷硬币的随机试验中，我们可以用 $X(H)=1$ 和 $X(T)=0$ 定义随机变量 $X$. 其中 $H$ 和 $T$ 分别表示硬币着地时面朝上或者面朝下.

假设样本空间为 $\Omega=\{\omega_1,\cdots,\omega_n\}$，概率函数为 $\mathbf{P}$，随机变量 $X$ 的取值范围为 $\mathscr{X}=\{x_1,\cdots,x_m\}$. 在 $\mathscr{X}$ 上定义分布(distribution)$\mathrm{P}_X$ 为

$$\mathrm{P}_X(\{X=x_i\})=\mathbf{P}(\{\omega_j\in\Omega:X(\omega_j)=x_i\}).\tag{1.4}$$

因此，当且仅当一个随机试验的输出结果为 $\omega_j\in\Omega$，且 $X(\omega_j)=x_i$ 时，我们会观测到 $X=x_i$. 式(1.4)的左边实际上说明 $\mathrm{P}_X$ 是按照概率 $\mathbf{P}$ 的定义方式在 $\mathscr{X}$ 上定义的一个诱导(induced)概率. 可以证明，$\mathrm{P}_X$ 同样满足概率论公理. 下面给出 $\mathrm{P}_X$ 的一般化定义：

**定义 1.7(分布)** 随机变量 $X$ 的分布 $\mathrm{P}_X$ 是 $X$ 在$(\mathbb{R},\mathscr{B}(\mathbb{R}))$上诱导生成的概率测度. 用公式表达为

$$\mathrm{P}_X(A)=\mathbf{P}(X^{-1}(A))=\mathbf{P}(\omega\mid X(\omega)\in A)\ \text{对任意}\ A\in\mathscr{B}(\mathbb{R})\ \text{成立}.$$

由上式可以得到，$\mathrm{P}_X=\mathbf{P}\circ X^{-1}$. 在不致引起混淆时，经常省略下标 $X$ 用 P 来代替 $\mathrm{P}_X$ 以简化.

容易证明 P 是一个概率测度，它满足三条概率论公理.

**注 1** $\mathrm{P}(A)=\mathbf{P}(X^{-1}(A))\geqslant 0$ 对任意 $A\in\mathscr{B}(\mathbb{R})$成立；

**注 2** $\mathrm{P}(\mathbb{R})=\mathbf{P}(\Omega)=1$；

**注 3** 如果 $A_1,A_2,\cdots\in\mathscr{B}(\mathbb{R})$ 为互斥事件，则 $\mathrm{P}(\bigcup_{i=1}^{\infty}A_i)=\mathbf{P}(X^{-1}(\bigcup_{i=1}^{\infty}A_i))=\mathbf{P}(\bigcup_{i=1}^{\infty}X^{-1}(A_i))=\sum_{i=1}^{\infty}\mathbf{P}(X^{-1}(A_i))=\sum_{i=1}^{\infty}\mathrm{P}(A_i)$.

在前面提到的其他概率性质也同样适用于 $\mathrm{P}_X$.

**例 1.11(投掷硬币，续 4)** 考虑上述投掷硬币随机试验中的随机变量 $X$，此时令 $\mathbf{P}(H)=1/3$，$\mathbf{P}(T)=2/3$. $X$ 的分布函数 $\mathrm{P}_X$ 为

$$\mathrm{P}_X(\{0\})=\frac{2}{3},\quad \mathrm{P}_X(\{1\})=\frac{1}{3}.$$

将 $\mathrm{P}_X$ 更准确地表述为，$\mathrm{P}_X(A)=1/3,2/3,1$，其取值分别对应于 $A$ 为$\{1\}$，$\{0\}$或$\{0,1\}$. 如果 $A$ 既不包含 0 也不包含 1，即 $A$ 为$\{\varnothing\}$时，$\mathrm{P}_X(A)=0$.

下面是一个稍微复杂的例子.

**例 1.12(投掷硬币三次)** 考虑一个投掷硬币的随机试验,这次试验连续投三次硬币,定义随机变量 $X$ 为三次投掷中硬币面朝上的次数. $X$ 的所有可能取值以及在样本空间中对应的事件为

| $\omega$ | HHH | HHT | HTH | THH | TTH | THT | HTT | TTT |
|---|---|---|---|---|---|---|---|---|
| $X(\omega)$ | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 0 |

随机变量 $X$ 的取值范围为 $\mathscr{X}=\{0,1,2,3\}$. 由于硬币铸造均匀,且三次投掷互相独立,在 $\mathscr{X}$ 上的生成概率分布为

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $\mathrm{P}_X(X=x)$ | $\frac{1}{8}$ | $\frac{3}{8}$ | $\frac{3}{8}$ | $\frac{1}{8}$ |

比如,$\mathrm{P}_X(X=2)=\mathbf{P}(\{\mathrm{HHT},\mathrm{HTH},\mathrm{THH}\})=3/8$.

对于每一个随机变量 $X$,有一个相对应的累积分布函数(cumulative distribution function).

**定义 1.8(累积分布函数)** 随机变量 $X$ 的累积分布函数(cdf)$F_X$ 定义为

$$F_X(x)=\mathrm{P}_X(X\leqslant x),\quad 其中\ x\in\mathbb{R}.$$

涉及分布函数时,在不至于引起混淆的前提下,经常省略 $F_X$ 的下标 $X$ 而用 $F$ 来表示 $X$ 的分布函数,因此上述分布函数定义式通常记为 $F(x)=\mathbf{P}(X\leqslant x)$,并用 $X\sim F$ 来表示 $X$ 服从分布 $F$. 显然,概率分布 P 由分布函数 $F$ 唯一确定.

**例 1.13(投掷硬币三次,续 1)** 仍然是连续投掷三次硬币的随机试验. 令 $X$ 为三次投掷中硬币面朝上的次数,则 $X$ 的累积分布函数为

$$F_X(x)=\begin{cases}0, & 当\ x<0,\\ \frac{1}{8}, & 当\ 0\leqslant x<1,\\ \frac{1}{2}, & 当\ 1\leqslant x<2,\\ \frac{7}{8}, & 当\ 2\leqslant x<3,\\ 1, & 当\ 3\leqslant x.\end{cases}$$

显然 $F_X(x)$是一个右连续的阶梯函数.

累积分布函数 $F(x)$的性质如下.

**定理 1.4(累积分布函数的性质)** 函数 $F$ 为随机变量的累积分布函数,当且仅当它满足以下三个条件:

(1) $\lim\limits_{x\to-\infty}F(x)=0,\lim\limits_{x\to\infty}F(x)=1$;

(2) (单调性)如果 $x\leqslant y$,则 $F(x)\leqslant F(y)$;

(3) (右连续性) $\lim\limits_{x\downarrow y}F(x)=F(y)$对任意 $y$ 成立.

对上述定理的证明见 Casella 与 Berger(2002,p. 31)的著作.

**例 1.14(几何分布)** 假设连续投掷一枚硬币,直到它首次面朝上着地才停止投掷. 令 $p$

为任意一次投掷中硬币面朝上着地的概率，定义随机变量 $X$ 为直到硬币面朝上为止总共的投掷次数，则对任意 $x=1,2,\cdots$，有

$$\mathrm{P}_X(X=x)=(1-p)^{x-1}p.$$

这是由于在 $x$ 次投掷中，在得到一次硬币面朝上之前已经有 $x-1$ 次面朝下的结果，且假设所有的投掷是独立的，因此，对 $x=1,2,\cdots$，有

$$\begin{aligned}F_X(x)&=\mathrm{P}_X(X\leqslant x)=\sum_{i=1}^{x}\mathrm{P}_X(X=i)=\sum_{i=1}^{x}(1-p)^{i-1}p\\&=\frac{1-(1-p)^x}{1-(1-p)}p=1-(1-p)^x.\end{aligned}\tag{1.5}$$

显然，由于对任意 $x<0$，$F_X(x)=0$，因此 $\lim\limits_{x\to-\infty}F_X(x)=0$，$\lim\limits_{x\to\infty}F_X(x)=1$. 而 $F_X(x)$ 的单调性由式(1.5)中的求和形式即能推出. 对任意 $x=1,2,\cdots$，如果 $\varepsilon>0$ 且 $\varepsilon$ 足够小，则 $F(x+\varepsilon)=F(x)$，即

$$\lim_{\varepsilon\downarrow 0}F(x+\varepsilon)=F(x).$$

**例 1.15(逻辑分布)** 用逻辑(logistic)分布函数作为一个连续型分布函数的例子，其分布函数为

$$F(x)=\frac{1}{1+\mathrm{e}^{-x}}.$$

显然 $F(x)$ 满足分布函数的性质. 对 $F(x)$ 求微分可得

$$\frac{\mathrm{d}}{\mathrm{d}x}F(x)=\frac{\mathrm{e}^{-x}}{(1+\mathrm{e}^{-x})^2}>0.$$

即证明了 $F(x)$ 的递增性.

## 1.4 随机向量和联合分布函数

定义随机变量时，我们将其看作一个从样本空间 $\Omega$ 映射到实数轴上的函数，而对由多个随机变量构成的随机向量的定义也是类似的. 设 $X_1,\cdots,X_n$ 为定义在概率空间$(\Omega,\mathscr{F},\mathbf{P})$上的随机变量，定义一个随机向量 $X$ 为

$$X(\omega)=\begin{pmatrix}X_1(\omega)\\ \vdots\\ X_n(\omega)\end{pmatrix}.$$

**定义 1.9(随机向量)** 一个 $n$ 维随机向量是一个从样本空间 $\Omega$ 映射到 $n$ 维欧几里得空间 $\mathbb{R}^n$ 的函数.

因此，一个 $n$ 维随机向量即是由 $n$ 个随机变量构成的一个 $n$ 维数组，随机向量 $X$ 是一个从 $\Omega$ 映射到 $\mathbb{R}^n$ 的可测函数. 在不致引起混淆的情况下，也可以将随机向量称作随机变量.

**例 1.16(投掷硬币，续 5)** 考虑投掷硬币的随机试验. 定义随机变量 $X_1$ 为 $X_1(H)=1$，$X_1(T)=0$；定义随机变量 $X_2$ 为 $X_2(H)=0$，$X_2(T)=1$. 定义随机向量 $X=(X_1,X_2)^{\mathrm{T}}$，其中符号 $a^{\mathrm{T}}$ 表示向量 $a$ 的转置. 显然 $X(H)=(1,0)^{\mathrm{T}}$，$X(T)=(0,1)^{\mathrm{T}}$，即 $X$ 为一个从样本空间 $\Omega=\{H,T\}$ 映射到 $\mathbb{R}^2$ 的函数.