计量经济学教程

刘振亚 编著

中国人民大学出版社

计量经济学教程

刘振亚　编著

中国人民大学出版社

图书在版编目（CIP）数据

计量经济学教程/刘振亚编著.

北京：中国人民大学出版社，1996

ISBN 7-300-02221-9/F. 659

Ⅰ. 计…

Ⅱ. 刘…

Ⅲ. 计量经济学-高等学校-教材

Ⅳ. F224. 0

中国版本图书馆 CIP 数据核字（96）第 20554 号

计量经济学教程

刘振亚　编著

出　版：中国人民大学出版社

（北京海淀路 175 号　邮码 100872）

发　行：新华书店总店北京发行所

印　刷：河北省涿州市星河印刷厂

开本：1168×850 毫米 1/32　　印张：14.75

1997 年 3 月第 1 版　1997 年 3 月第 1 次印刷

字数：366 000

定价：17.00 元

（图书出现印装问题，本社负责调换）

前　言

计量经济学创立至今，虽然只经过短短几十年时间，却得到了突飞猛进的发展，在经济研究中，日益占有举足轻重的地位。它是以统计为基础，数学方法为手段，经济理论为指导，考察现代经济社会中的各种经济数量关系、预测经济发展趋势、检验经济政策效果的工具。由于它具有很大的实用性，各国政府都不同程度地采用计量经济学方法，建立本国的宏观经济模型进行经济预测，模拟政策效果。

计量经济学的研究和应用在我国起步较晚，有关它的论著多为国外学者所著。更为突出的问题是由于计量经济学偏重于计算，以数学语言代替经济语言的地方很多，这对我国擅长以理论分析为主的经济研究人员来说，学习起来十分吃力。同时，计量经济学内容随着时间的推进而日新月异，尤其是随着70年代和80年代微型计算机的普及，计量经济学研究的重点已转为更多地强调计量经济学的应用问题。国内外许多有关教科书已不能适应现在的要求。

本书作者数年来一直从事计量经济学的教学与研究，在这个过程中参考了近期国外有关书刊，并与国内外专家学者多次进行探讨和研究。在本书的写作过程中，力图在运用数学语言的同时，更多地强调以文字形式加以描述证明，以便使更多的读者能够了解计量经济学的动向。

本书在吸收了国外计量经济学著作的最新成果基础上，系统地介绍了现代计量经济学方法。在结构安排上分：计量经济学的

统计基础（第一章）、计量经济学基础（第二至第十二章，其中第九章系统地介绍了时间序列软件包 TSP）及计量经济学应用（第十三至第二十章）三部分，详细地阐述了计量经济学的基础理论、方法及其在经济实践中的应用。

此外，为了更好地把理论和实际相结合，还加以大量例子配合证明，并在第三部分专门讨论宏观经济模型的建模问题。在内容取舍上也注重吸收国外较为成熟的最新成果。

本书的写作及计量经济课程的开设，始终得到中国人民大学副校长杜厚文教授多方的支持和帮助。作者长期以来在中美经济学培训中心的工作，使作者有机会与国内外的经济专家互相交流、密切合作，形成了令人难忘的、融洽的讨论与研究气氛。在写作过程中直接得到了他们的指导和帮助。他们是：

• 美国纽约州立大学奥尔已尼分校计量经济学教授 K. Lahiri
• 美国堪萨斯大学经济系主任吴得民教授（De－Min Wu）
• 美国哥伦比亚大学经济学教授 H. Watts
• 美国杜克大学计量经济学教授 T. D. Wallance
• 美国加州大学圣地亚哥分校经济学教授 M. Machina

作者在国际货币基金组织工作期间，曾与以下经济学家一起研究和交流：

• 国际货币基金组织亚洲部高级经济学家 M. Fetherston，M. Bell，Hoe. K. Khor，Zhu Liu Hu
• 国际货币基金组织管理部部长助理 A. D. Glotz
• 国际货币基金组织统计部经济学家 Jeffery Taylor
• 世界银行中国局高级经济学家 A. Rajaram，经济研究分析师 T. Rapala

众所周知，一本数十万字的学术论著，它的完成并不是一个单纯的写作过程，其间还有大量繁重细致的劳动。从本书的第一章直至最后一章，我的学生于晓筠、刘纳、Bartha Zsolt、Talos Ga-

bor 和 Hegdus Gabor 等在搜集资料、整理和抄写等多方面给了我许多帮助，他们的友谊和帮助令我终生难忘。

本书的后半部分的写作是我在美国进行博士后研究期间完成的。这里要感谢我的博士后导师，美国普渡大学 George Horwich 教授给了我专门的时间来完成这部分的写作。

希望本书能够使更多的读者应用计量经济学方法来解决中国的实际问题，进一步丰富中国经济的研究内容，加深研究的深度，使中国经济早日腾飞。

最后应指出，由于作者水平有限，本书的缺点与错误在所难免，恳切希望读者批评指正。

刘振亚

于美国西拉菲耶特普渡大学

1995. 4

上篇　计量经济学基础

第一章　计量经济学的统计学基础

在计量经济学的研究中，即使在它最为实用和简明的形式下，对数理统计学基本概念的清晰理解都是必不可少的。这里，假设大多数读者事先都学习过数理统计学，即便如此，我们也意识到，数理统计学在大学的数学基础课教学中，属于比较困难的部分。虽然许多学生，已经顺利地通过了学分考试，但对它的掌握依然是很不完备的。另外，大多数人对于数学公式、数学符号的健忘也提醒我们在进一步讨论计量经济学的专题之前，对一些在本书中经常使用到的数理统计学基本内容必须事先进行一些温习和回顾。

事实上，不懂得数理统计学就不可能学习和研究计量经济学，数理统计学是计量经济学的基础，它为计量经济学的研究提供了唯一而有效的方法。这一点，读者将在今后各章节的学习中深刻地体会到。从某种意义上来说，计量经济学就是使数理统计学在建立经济模型中得以应用的一门学科。正因为如此，在西方各权威性的计量经济学教科书中，对数理统计学基础内容的回顾和复习往往占据极大的篇幅，并以独立、完整而重要的章节出现。

这里，本书将以整个一章的篇幅对数理统计学的基础原理给出一个概略的复习和回顾。应当指出的是，这一回顾是“概略的”，但不是“肤浅的”。通过本章的学习，读者将能从逻辑结构

的层次上，比较完整地掌握数理统计学的基本内容。无论读者事先学过还是没有学过数理统计学，认真学习本章内容都会很有帮助，这不仅由于本章内容完整而自成体系，而且因为它注重于逻辑结构分析，注重于方法的阐述，注重于公式、定义和定理的内在涵义及其相互关系，同时，也注意了紧扣计量经济学这一主题。

为了帮助读者把注意力集中在数理统计学的逻辑结构和思想方法上，而不是数学推导的具体细节上，并且鉴于这不是一部数学教科书，我们对公式、定理往往强调它们的结论、涵义和它所给出的方法，而不强调定理、公式本身的数学证明。

在本章的讨论展开之前，我们首先给出这部分内容的逻辑结构。读者可以在以后各节的学习中随时参阅此表，以便了解自己所阅读的部分在整个数理统计学体系中所处的地位，及其与其它部分的联系。在本章结束时，我们将再次仔细地讨论此表，并以此作为全章小结。

第一节　总体、样本和随机函数

首先，给出三个定义：

定义 1.1　研究对象的全体称为总体或母体，组成总体的每个基本单位称为个体。

注意：(1) 总体根据所包含的个体数目可分为：有限总体和无限总体。

(2) 总体中的每一个体，具有共同的可观察特征。我们把这种特征作为不同总体的区别指标。

(3) 量度同一对象所得到的数据，也构成一种类型的总体。这种总体中的数据之所以不同，是因为测量误差的存在。

(4) 表现总体状况的某一个或某几个数量特征在总体中往往

随概率不同而有不同的取值。

定义 1.2 总体中抽出若干个体而组成的集体，称样本。样本中所含个体的个数，称样本容量。

注意：在进行抽样时，样本的选取必须是随机的，即总体中每一个体有同样的机会被选入样本。

数理统计学的逻辑结构

(1) 总体 (Population) 和样本 (Sample)

如何用一个变量来描述总体——随机变量 (Random Variable) 的引入。

(2) 对总体的描述：随机变量的数字特征

数学期望　$\mu_X = E(X)$ 描述总体的一般水平

方差　$\sigma_X^2 = \mathrm{Var}(X)$ 描述总体的离散程度

(3) 对样本的描述：样本分布的数字特征

样本平均数 $\overline{X}$，描述样本的一般水平

样本方差　S^2，描述样本的离散程度

实际上，$\overline{X}$ 和 S^2 是 $E(X)=\mu_X$ 和 $\mathrm{Var}(X)=\sigma_X^2$ 的无偏估计量。

(4) 总体与样本的连接点：随机变量的分布 (Distribution)

(5) 如何通过这样本数据和样本分布特征来估计总体的数字特征，即 μ 和 σ^2，以及总体中数据生成过程的各种参数。

a. 估计量的特性。

b. 估计的方法：

- 估计方法
  - 点估法 (Point)
    - 矩法(Moment Method)
    - 最大似然估计法(Maximum Likelihood Estimation)
    - 最小二乘估计法(Least-Squares Estimation)
    - 最小 χ^2 法(Minχ^2 Estimation)
  - 区间估计(Confidence Interval Estimation)
    - 单个总体
      - 估计期望 (E(X))
        - 方差已知
          - 总体分布未知
          - 正态总体
          - 一般总体(大样本下)
        - 方差未知
          - 一般总体(大样本下)
          - 正态总体
      - 估计方差(常用小样本下，正态总体)
      - 估计其他参数
    - 两个总体

c. 对估计量的检验——假设检验

它与区间估计本质相同，只是所讨论的出发点不同：

- 假设检验
  - 对总体分布特征的检验
    - 一个正态总体的假设检验
      - 已知方差 σ^2，检验假设 $H_0: \mu=\mu_0$
      - 未知方差 σ^2，检验假设 $H_0: \mu=\mu_0$
      - 未知期望 μ，检验假设 $H_0: \sigma^2=\sigma_0^2$
      - 未知期望 μ，检验假设 $H_0: \sigma^2<\sigma_0^2$
    - 两个正态总体的假设检验
      - 未知 σ_1^2, σ_2^2，但知二者相等，检验假设 $H_0: \mu_1=\mu_2$
      - 未知 μ_1, μ_2，检验假设 $H_0: \sigma_1^2=\sigma_2^2$
      - 未知 μ_1, μ_2，检验假设 $H_0: \sigma_1^2<\sigma_2^2$
    - 总体分布的假设检验：(χ^2 检验法)
      - 总体的离散型 H_0：总体 ξ 的概率函数为 $p_i(\xi=x_i)=p_i, i=1,2,3\cdots\cdots$
      - 总体的连续型 H_0：总体 ξ 有分布密度 $\varphi(x)$
  - 对各种系数、参数估计值的检验

定义 1.3 根据概率的不同而取不同数值的变量，叫做随机变量（Random Variable）。

注意：(1) 一个随机变量具备下列特性：它可以取到许多不同的值，每个值都可能以一个小于或等于 1 的概率被取到。显然，所有这些概率之和为 1。

(2)随机变量以一定的概率取到各种可能值，按其取值情况可以把随机变量分为两类：离散型随机变量和连续型随机变量。离散型随机变量所可能的取值至多为可列个。连续型随机变量所可能的取值可以是整个数轴或数轴上的某个连续区间，如图 1.1 所示。

(3) 本书中，随机变量用 x、y、ξ、η 等符号表示。

以上我们给出了总体、样本、随机变量三个重要概念，下面我们讨论三者之间的关系。

由上面我们已经了解到，表示总体状况的某一个或某几个数量特征在总体中往往随概率不同而取不同的值（定义 1.1，注意(4)）。显然，对于这样的数量特征，我们用一般的变量是无法加以描述的，能够对之给予描述的是一类特殊的变量，即随机变量（定义 1.3）。

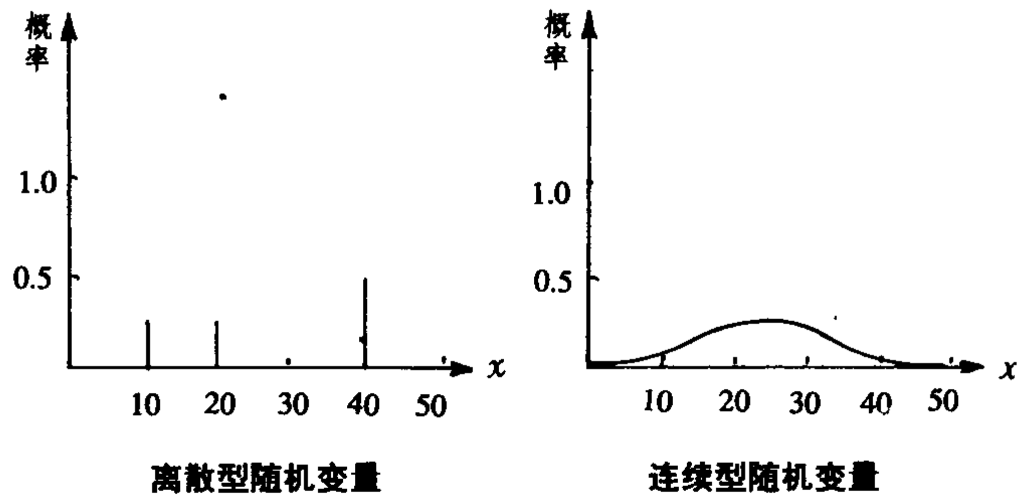

图 1.1

我们对于总体中某一个体所具有的特殊属性往往并不关心，真正感兴趣的是表示总体特征的数量指标。例如，全天生产的 50 000 个灯泡中寿命在 1 000 小时～1 200小时之间，寿命 1 000 小时以下以及 1 200 小时以上的灯泡所占的百分比等。就总体的某一数量特征 ξ 而言，如灯泡的使用寿命，每个个体的取值不一定完全相同，但它是按照一定规律分布的，如 5 万个灯泡中各种寿命灯泡所占的比例是基本稳定的。因此，对于一个总体来说，其每一个数量特征完全符合定义 1.3，是根据概率的不同而取不同值的变量，即总体中每一个数量特征就是一个随机变量。由于我们主要是研究总体的数量特征，所以我们把总体看成是具有若干数量特征的研究对象的全体，可直接用一个随机变量来表示。

可见，所谓总体就是一个随机变量，所谓样本就是 n 个（样本容量为 n）相互独立且与总体有相同分布的随机变量 x_1，…，x_n。每一次具体抽样所得的数据，就是 n 元随机变量的一个观察值（样本值），记为（X_1，…，X_n）。（多元随机变量定义见定义 1.8）。

任何随机变量都有自己的分布，我们已经了解总体可表示为一个随机变量，并具有其自身的分布。而样本就是 n 个相互独立且与总体有相同分布的随机变量 x_1，…，x_n。可见，通过总体的分

布可以把样本与总体联系起来，从而使我们能够通过对样本特征的研究达到对总体研究的目的，换而言之，总体分布(Distribution）是样本和总体的连接点。

每当提到一个容量为 n 的样本时，常有双重涵义：一是指某一次抽样的具体数值 $(X_1, \cdots, X_n)$；二是泛指一次抽出的可能结果，这时指一个 n 元随机变量 $(x_1, x_2, \cdots, x_n)$。

定义 1.4 设 $(x_1, x_2, \cdots, x_n)$ 为一组样本，函数 $f(x_1, x_2, \cdots, x_n)$ 若不含有未知参数，则称 $f(x_1, x_2, \cdots, x_n)$ 为统计量。

统计量一般是样本的连续函数。由于样本是随机变量，因而它的函数也是随机变量，如 $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$ 就是统计量。

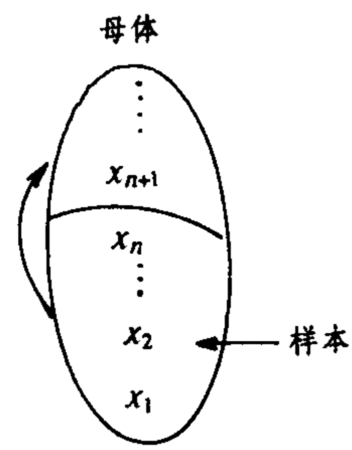

图 1.2

以上我们讨论了总体、样本和随机变量的关系，下面我们讨论总体和样本二者之间的关系。如图 1.2 所示，样本是总体的一小部分，是对总体进行随机抽取后所得到的集合。对于观察者来讲，整个总体的状况是不了解的，观察者所能了解的只是总体的一部分——样本的具体状况，我们所要做的就是通过对这些样本具体状况的研究，来推知整个总体的状况。图 1.2 中的箭头表示了我们的研究方向。既然要通过研究样本来推知总体的状况特性，那么我们通过什么能把样本和总体联系起来，使我们的研究在二者之间得以沟通呢？答案是通过总体及随机变量函数的定义、分类和分布。

下面，我们就以下三个问题展开进行讨论：随机变量的分布、二元随机变量和独立性问题。

一、随机变量的分布

（一）离散型随机变量的分布

定义 1.5 如果随机变量 ξ 只取有限个或可列个可能值，而且 ξ 以确定的概率取到这些值，则称 ξ 为离散型随机变量。

直观上，可将 ξ 的可能取值及相应概率列成下列概率分布表（见表 1—1）。

表 1—1

ξ	x_1	x_2	$\cdots$ x_i $\cdots$
p	p_1	p_2	$\cdots$ p_i $\cdots$

同时，ξ 的概率分布情况也可以用一系列等式表示：

$$p\ (\xi=x_i)\ =p_i \qquad (i=1,\ 2,\ \cdots)$$

此式的涵义为 ξ 取 x_i 的概率为 p_i（$i=1$，2，…），它称为随机变量 ξ 的概率函数（概率分布）。

这里的 x_i（$i=1$，2，…）应包括 ξ 的每一个可能的取值，而且每一个可能的取值都只出现一次。

显然：$\begin{cases} p_i \geqslant 0 \\ \sum p_i = 1 \end{cases}$

所谓离散型随机变量的分布，一般就是指它的概率函数或分布表。

例 1 一批产品的废品率为 5%，从中任取一个进行检验，以随机变量 ξ 来描述这一试验并写出 ξ 的分布。

解 以 $\xi=0$ 表示“产品为合格品”，$\xi=1$ 表示“产品为废品”，分布表如表 1—2 所示。

表 1—2

ξ	0	1
p	0.95	0.05

其概率函数为 $p\ (\xi=0)\ =0.95$，$p\ (\xi=1)\ =0.05$，

或 $p\ (\xi=i)\ =\ (5\%)^i\ (95\%)^{1-i} \qquad (i=0,\ 1)$

例 2 用随机变量 ξ 描述掷一颗骰子的试验。

解 分布表如表 1—3 所示。

表 1—3

ξ	1	2	3	4	5	6
p	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

分布的概率函数：$p(\xi=i)=\frac{1}{6}\quad(i=1,2,3,4,5,6)$

(二) 随机变量的分布函数

定义 1.6 若 ξ 是一个随机变量（可以是离散的，也可以为非离散的），对任何实数 x，令 $F(x)=p(\xi\leqslant x)$，称 $F(x)$ 为随机变量 ξ 的分布函数。

$F(x)$，即事件"$\xi\leqslant x$"的概率，是 x 的一个实函数。对任意实数 $x_1<x_2$，有：$p(x_1<\xi\leqslant x_2)=p(\xi\leqslant x_2)-p(\xi\leqslant x_1)=F(x_2)-F(x_1)$。可见，若已知 ξ 的分布函数 $F(x)$，就能知道 ξ 在任一个区间上取值的概率。所以，它完整地描述了随机变量的变化情况。

$F(x)$ 的性质：

(1) 对一切 $x\in(-\infty,+\infty)$，$0\leqslant F(x)\leqslant 1$。

(2) $F(x)$ 为不减函数。

(3) $F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0$

$F(+\infty)=\lim\limits_{x\to+\infty}F(x)=1$

(4) $F(x)$ 至多有可列个间断点。且在间断点上右连续。

例 3 求例 1 中的分布函数。

解 $F(x)=p(\xi\leqslant x)=\begin{cases}0 & x<0\\ 0.95 & 0\leqslant x<1\\ 1 & x\geqslant 1\end{cases}$

例 4 求例 2 中 ξ 的分布函数 $F(x)$。

解 $F(x)=p(\xi\leqslant x)=\begin{cases}0 & x\leqslant 1\\ 1/6 & 1\leqslant x<2\\ 2/6 & 2\leqslant x<3\\ 3/6 & 3\leqslant x<4\\ 4/6 & 4\leqslant x<5\\ 5/6 & 5\leqslant x<6\\ 1 & x\geqslant 6\end{cases}$

分布函数与概率函数满足下列关系 $F(x)=\sum_{x_i\leqslant x}p_i$

（三）连续型随机变量的分布。

定义1.7 对于任何实数x，如果随机变量ξ的分布函数$F(x)$可以写成

$$F(x)=\int_{-\infty}^{x}\psi(t)\,dt$$

其中$\psi(x)\geqslant 0$，则称ξ为连续型随机变量，称$\psi(x)$为ξ的概率分布密度函数，也常写为$\xi\sim\psi(x)$。

$\psi(x)$的性质：(1) $\psi(x)\geqslant 0$

(2) $\int_{-\infty}^{+\infty}\psi(x)\,dx=1$

显然$p(a<\xi\leqslant b)=\int_a^b\psi(x)\,dx$，并且在$\psi(x)$的连续点上，有$F'(x)=\psi(x)$。

下面我们来看$\psi(x)$的涵义：

$\because F'(x)=\psi(x)$

$$\therefore \psi(x)=\lim_{\triangle x\to 0}\frac{F(x+\triangle x)-F(x)}{\triangle x}$$

$$=\lim_{\triangle X\to 0}\frac{p(x<\xi<x+\triangle x)}{\triangle x}$$

这表明，$\psi(x)$不是ξ取值x的概率，而是在x点概率分布的密集程度。但是$\psi(x)$的大小能反映出ξ在x附近取值的概率大小。（在本书中，我们有时直接把$\psi(x)$看成ξ取值x的概率。当然，这是不严谨的。）

例5 若ξ有密度函数

$$\psi(x)=\begin{cases}\lambda & a\leqslant x\leqslant b \qquad (a<b)\\ 0 & \text{其它}\end{cases}$$

则称ξ服从区间$[a,b]$上的均匀分布。试求$F(x)$。

解 $\because \int_{-\infty}^{\infty}\psi(x)\,dx=\int_a^b\lambda dx=\lambda(b-a)=1$

$\therefore \lambda=\frac{1}{b-a}$

$\because F(x)=\int_{-\infty}^{x}\psi(t)\,dt$

$$\therefore F(x)\Rightarrow\begin{cases}0 & x<a\\ \dfrac{x-a}{b-a} & a\leqslant x<b\\ 1 & x\geqslant b\end{cases}$$

（四）分布函数、概率函数、密度函数三者的关系

分布函数既适用于离散型随机变量，也可适用于连续型随机变量，是描述各种类型随机变量变化规律的最一般的共同形式。但是由于它不够直观，所以往往不常用。对于离散型随机变量，概率函数的描述既简单又直观；对于连续型随机变量，其密度函数 $\psi(x)$ 的大小能反映出 ξ 在 x 附近取值的概率大小，从而要比分布函数的描述直观。因而在实际应用中，我们一般用概率函数和密度函数来分别对离散型和连续型的随机变量进行描述。

二、二元随机变量

定义 1.8 如果每次试验的结果对应着一组确定的实数（ξ_1，…，ξ_n），它们是随试验结果不同而变化的 n 个随机变量。如果对任何一组实数 x_1，…，x_n，事件"$\xi_1\leqslant x_1$，$\xi_2\leqslant x_2$，…，$\xi_n\leqslant x_n$"有着确定的概率，则称 n 个随机变量的总体（ξ_1，…，ξ_n）为一个 n 元随机变量。

定义 1.9 称 n 元函数 $F(x_1, \cdots, x_n)=p(\xi_1\leqslant x_1, \cdots, \xi_n\leqslant x_n)$，$(x_1, \cdots, x_n)\in R^n$ 为 n 元随机变量的分布函数。

定义 1.10 如果二元随机变量（ξ，η）所有可能的取值为有限或可列个，并且以确定的概率取各个不同数值，则称（ξ，η）为二元离散型随机变量。

为了直观，可以把（ξ，η）所有可能的取值及相应概率列成下表，称为（ξ，η）的概率分布表如表 1—4 所示。

表 1—4

$\xi \backslash \eta$	y_1	y_2	…	y_j	…
x_1	p_{11}	p_{12}	…	p_{1j}	…
x_2	p_{21}	p_{22}	…	p_{2j}	…
…	…	…	…	…	…
x_i	p_{i1}	p_{i2}	…	p_{ij}	…
…	…	…	…	…	…

为了简单，也可以用一系列等式表示离散型二元随机变量（ξ，η）的概率分布，即：

$$p\ (\xi=x_i,\ \eta=y_j)\ =p_{ij}\quad (i,\ j=1,\ 2,\ \cdots)$$

上式也称为ξ与η的联合分布。显然$p_{ij}\geqslant 0$，i，$j=1$，2，…，并且$\sum\limits_i \sum\limits_j p_{ij}=1$。

例 6 同一品种的 5 个产品中，有两个正品，每次从中取一个检验质量，不放回地抽取，连续两次。令“$\xi_i=0$”表示第i次取到正品，而“$\xi_i=1$”为第i次取到次品（$i=1$，2），写出（ξ_1，ξ_2）的分布。

解 $p\ (\xi_1=0,\ \xi_2=0)\ =p\ (\xi_1=0)\ \cdot p\ (\xi_2=0|\xi_1=0)$

$$=\frac{2}{5}\times\frac{1}{4}=\frac{1}{10}$$

$$p\ (\xi_1=0,\ \xi_2=1)\ =\frac{2}{5}\times\frac{3}{4}=\frac{3}{10}$$

$$p\ (\xi_1=1,\ \xi_2=0)\ =\frac{3}{5}\times\frac{2}{4}=\frac{3}{10}$$

$$p\ (\xi_1=1,\ \xi_2=1)\ =\frac{3}{5}\times\frac{2}{4}=\frac{3}{10}$$

列成分布表如表 1—5 所示。