教育部经济管理类核心课程教材

Econometrics

计量经济学

林清泉 主编

中国人民大学出版社

教育部经济管理类核心课程教材

Econometrics

计量经济学

林清泉 主编

中国人民大学出版社

图书在版编目(CIP)数据

计量经济学/林清泉主编.
北京:中国人民大学出版社,2006
教育部经济管理类核心课程教材
ISBN 7-300-07204-6

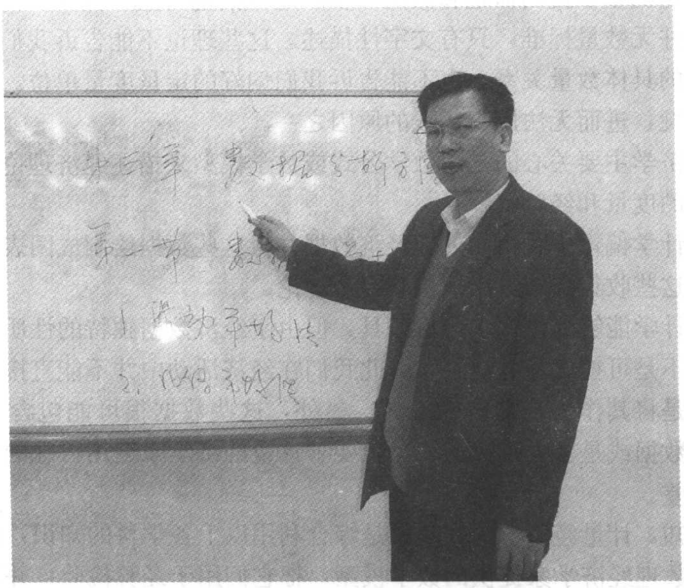
Ⅰ.计…
Ⅱ.林…
Ⅲ.计量经济学-高等学校-教材
Ⅳ.F224.0

中国版本图书馆 CIP 数据核字(2006)第 025269 号

教育部经济管理类核心课程教材
计量经济学
林清泉　主编

出版发行	中国人民大学出版社		
社　　址	北京中关村大街 31 号	**邮政编码**	100080
电　　话	010－62511242(总编室)		010－62511239(出版部)
	010－82501766(邮购部)		010－62514148(门市部)
	010－62515195(发行公司)		010－62515275(盗版举报)
网　　址	http://www.crup.com.cn		
	http://www.ttrnet.com(人大教研网)		
经　　销	新华书店		
印　　刷	三河市新世纪印务有限公司		
开　　本	720×965 毫米 1/16	**版　　次**	2006 年 4 月第 1 版
印　　张	24.5　插页 1	**印　　次**	2006 年 4 月第 1 次印刷
字　　数	448 000	**定　　价**	28.00 元

前 言



林清泉：中国人民大学财政金融学院教授、博士生导师，现任中国人民大学货币金融系副主任。林清泉教授还是德国慕尼黑大学和莱比锡大学的高级访问学者。

一、什么是计量经济学

计量经济学是将经济理论实用化、数量化的实证经济学。它是利用经济理论、数学、统计推断等工具对经济现象进行分析的经济科学的分支，包括模型设计和建立、参数估计和检验以及利用模型进行预测等过程。

计量经济学以经济理论为指导、以事实为依据、以数学为基础、以统计推断为方法、以计算机为手段，对具有随机特征的经济关系与经济活动的数量规律进行研究，并以计量经济模型的建立和应用为核心，对构建于数理经济学基础之上的数学模型提供经验支持，并得出数量结果、预测经济发展趋势、检验经济效果。

二、计量经济学的实用性

从上述定义可知，计量经济学涉及经济理论、数理经济学、经济统计学以及数理统计学等相关学科，但它不同于这些学科，是一门有自己研究方向的独立学科。

经济理论侧重于提出命题和假说，多以定性描述为主。例如，需求法则与偏好理论本身并无数量标准，只有文字性描述。这些理论不能告诉我们需求量与价格变动之间的具体数量关系，也不能告诉我们偏好的定量度量单位。这是许多经济理论被质疑，进而无法用于实践的原因之一。

数理经济学主要关心的是用数学公式或数学模型来描述经济理论，而不考虑对经济理论的度量和经验解释。

经济统计学偏重于收集、处理经济数据并将这些数据绘制成图表的形式，它不涉及利用这些收集到的数据来检验经济理论。

数理统计学能够提供许多分析工具，但由于经济数据独特的性质，即许多数据的生成并不是可控试验的结果，因此我们在经济活动中并不能直接运用数理统计学，而只是将其作为一种研究工具。另外，这些数据很可能包含了测量的误差、遗漏了数据或是丢失了数据。这就要求计量经济学家运用特殊的方法来处理这些测量误差。

由上可知，计量经济学的任务就是综合利用以上各学科的知识，依据观测和试验，采用数理经济学家提供的数学模型，把它们用于经验检验，并对大多数经济理论给出经验的、定量的解释。

对于主修经济学和金融专业的学生来说，学习计量经济学有实用性。例如，预测销售量、利息率、货币供给量或是估计商品的需求函数、供给函数以及价格弹性等，均需要应用计量经济学。客观地说，对于经济类专业的学生而言，计量经济学已成为其知识体系中不可或缺的部分。

三、计量经济学方法论

计量经济学发展至今已有多种学派，但是在经济学及有关领域的经验研究中占统治地位的仍属经典方法论。

一般来说，经典的计量经济学方法研究步骤可由下图说明：

理论或假说的陈述

↓

建立数学模型

↓

建立统计或经济计量模型

↓

收集数据 ⟨ 时间序列数据 / 横截面数据

↓

计量经济模型参数的估计

↓

检查模型的准确性；模型的假设检验

↓

检验来自模型的假说

↓

运用模型进行预测

经典计量经济学建模方法

在接下来的章节中，我们将详细介绍每一步骤的具体内容及应用实例。

四、计量经济学的类型及预备知识

计量经济学可划分为两大类：理论计量经济学和应用计量经济学。我们在每一类中均可按经典或贝叶斯（Bayes）的传统方法去探讨这一学科。本书主要介绍经典方法。

理论计量经济学是找出适当的方法去测度由计量经济模型设定的经济关系式，它多采用数理统计学方法。例如，书中广泛使用了最小二乘法，同时说明了利用此方法所涉及的假定、方法的性质；此外，当某一或某些假定不成立时，本书还说明了这些性质将会受到什么影响。

应用计量经济学主要以理论计量经济学为工具，研究经济学或经济活动中的某些特殊领域，如生产函数、投资函数、组合证券理论等等。

本书主要讨论理论计量经济学，首先从基础理论介绍入手，然后逐步放松假设条件，最终建立起完整的计量经济学体系。但本书并不具体讨论特殊领域的计量学应用。

学习计量经济学需要一些基本的预备知识。例如，微积分、矩阵代数、概率

论、统计学和时间序列等。若要达到实际应用的目的，还需要具备相关的计算机技能，比如SAS、SPSS、EViews等回归、统计软件包。通过运用一些价格便宜的计算机软件包，可使初学者易于掌握计量经济学这门课程。

五、全书结构

本书的目的是为金融学、财政学、管理学各专业的本科生和研究生提供一些定量理论、模型和分析方法。阅读本书有助于提高他们将定性分析与定量分析相结合、研究和解决实际问题的能力。本书是作者在2001年、2003年两次修改的原《计量经济学基础》讲义基础上补充修改而成。

第一章至第六章由林清泉编写，第七章由刘志、荣琪整理编写，第八章由王利萍整理编写，第九章由林清泉、李丽萍、荣琪整理编写，第十章由杜敏杰整理编写，第十一章由杨丰整理编写，第十二章由高圣言整理编写。最后，全书由林清泉教授统校。感谢王振、田明明、吴胜姣、邵娟为本书部分内容的录入所付出的辛勤劳动。

在编写过程中，引用了许多学者的研究成果，作者在此表示衷心感谢，并感谢中国人民大学出版社为本书的出版所做的工作。此外，作者还要感谢中国人民大学教务处及财政金融学院的支持与帮助。

本书分四个部分共十二章。第一部分主要介绍计量经济学的基础部分及数据处理的基本方法，它包括第一章、第二章和第三章。第一章简要地介绍了概率论的主要内容。第一章首先介绍了随机事件、随机变量及其分布和随机事件、随机变量的独立性等内容，然后介绍了在经济学、金融学以及保险学中占有重要地位的随机变量数字特征以及它们的应用：随机变量的数学期望、方差和相关系数；同时，第一章还介绍了在保险中经常用到的变异系数概念和切比雪夫不等式(Chebyshev inequality)，在本章的最后还介绍了在经济学、金融学、保险学中有广泛应用的n维正态随机变量的性质、大数定律、中心极限定理。第二章主要介绍了有关矩阵代数的内容，这些内容包括矩阵的定义、矩阵的运算、矩阵的逆、线性方程组的解、矩阵的特征根和特征向量、线性交换、正交变换等基本概念和结论。第三章主要介绍了一些常用数据的平滑技术，如算术平均、加权算术平均法、几何平均法、移动平均法、移动几何平均法，同时还介绍了算术移动平均法和指数平滑法。在介绍了数据的平滑技术之后，本章还介绍了统计推断的两个基本问题——未知参数的估计和假设检验。第三章首先引入了数理统计分析中一些基本概念、抽样分布及几个常用统计量的分布，接着介绍了参数估计问题，其中包括参数估计方法，如矩估计方法、极大似然估计、贝叶斯估计、估计量评估标准，最后介绍了假设检验和方差分析。第二部分主要介绍计量经济学在经典

假设条件下的线性模型，它包括第四章、第五章和第六章。第四章介绍了一元线性回归方法，主要介绍一元线性回归模型、回归系数的确定、回归方程的有效性检验方法，如方差分析法、t 检验和相关检验，最后还介绍了应用线性建立预测和控制模型。第五章主要介绍多元线性回归分析方法，它包括经典假设条件、参数估计方法、估计量的性质、估计量的分布和概率极限的定义及它的性质。这一章还介绍了多元线性回归方程有效性检验方法、多参数区间估计方法及预测方法。第六章介绍了虚拟变量的回归模型，主要内容包括作为解释变量的虚拟变量、作为被解释变量的虚拟变量、虚拟变量回归模型的类型和解释变量个数选取规则、一个定量变量和一个多分定性变量的回归、一个定量变量和多个定性变量的回归、多个定量变量和多个定性变量的回归。其中，虚拟变量的应用包括：应用虚拟变量改变回归直线的截距、应用虚拟变量改变回归直线的斜率、分段线性回归。应用虚拟变量检验回归模型结构的稳定性包括：传统判别结构稳定性的方法及存在的缺陷、应用虚拟变量法比较两个回归方程的结构方法。第三部分主要介绍了不满足经典假设条件下的线性模型，它包括第七章、第八章、第九章和第十章。第七章介绍的内容有多重共线性的概念、产生多重共线性的原因、产生多重共线性的后果、如何发现多重共线性以及处理多重共线性的方法。第八章介绍的内容有：异方差的概念；产生异方差的原因；在经典模型中，当同方差性质无法满足时的 OLS 估计；如何进行检验及校正异方差。第九章介绍自相关分析的内容。在经典回归模型中的干扰项 ε_i 是无自相关或序列相关，这一章进一步分析了这一假定，并介绍了存在自相关情况下的 OLS 估计以及发现自相关的方法。本章的最后还介绍了分布滞后模型。第十章介绍的内容有：结合一些实例讨论了联立方程模型的性质和特点、联立方程模型的变量和表示方法、联立方程模型的识别和估计等几个方面的问题。第四部分介绍了动态的计量经济学模型和应用计量经济学模型，它包括第十一章和第十二章。第十一章介绍了计量经济学的动态模型，主要介绍了时间序列分析方法，它包括时间序列数据的一些特性，如趋势性、季节性和平稳性，并重点介绍了时间序列数据的平稳性、平稳时间序列的分析模型以及非平稳时间序列的分析模型。第十二章介绍了如何应用计量经济学模型，它包括生产、需求和消费等微观经济学的具体问题和宏观经济学中的投资问题，以便学生了解应用计量经济模型的构造理论和方法以及计量经济学的具体应用。

21世纪

目 录

第一章 概率论基础……………………………………………………(1)
第一节 随机现象、随机试验和随机事件……………………………(1)
第二节 事件的频率与概率……………………………………………(4)
第三节 条件概率与事件的独立性……………………………………(7)
第四节 随机变量及其分布……………………………………………(9)
第五节 二维随机变量 ………………………………………………(11)
第六节 随机变量的数字特征 ………………………………………(13)
第七节 随机变量的矩 ………………………………………………(24)

第二章 矩阵代数 ……………………………………………………(30)
第一节 矩阵及其运算 ………………………………………………(30)
第二节 线性方程组 …………………………………………………(38)
第三节 二次型与正交变换 …………………………………………(45)

第三章 数据分析方法与参数统计推断 ……………………………(54)
第一节 数据的分析方法 ……………………………………………(54)
第二节 抽样分布 ……………………………………………………(60)
第三节 参数的统计推断 ……………………………………………(67)

第四节　方差分析方法 …………………………………………………… (90)

第四章　一元线性回归……………………………………………………… (100)
第一节　一元线性回归分析……………………………………………… (100)
第二节　线性回归的方差分析…………………………………………… (106)
第三节　t 检验（直接检验法） ………………………………………… (111)
第四节　相关系数及其显著性检验……………………………………… (112)
第五节　回归分析的其他问题…………………………………………… (116)

第五章　多元线性回归……………………………………………………… (123)
第一节　经典多元线性回归模型的概念………………………………… (123)
第二节　最小平方估计…………………………………………………… (126)
第三节　估计量的性质…………………………………………………… (137)
第四节　极大似然估计…………………………………………………… (147)
第五节　假设检验………………………………………………………… (149)
第六节　区间估计………………………………………………………… (159)
第七节　预测……………………………………………………………… (162)

第六章　虚拟变量的回归模型……………………………………………… (168)
第一节　虚拟变量………………………………………………………… (168)
第二节　虚拟变量的应用………………………………………………… (176)

第七章　多重共线性………………………………………………………… (187)
第一节　多重共线性的概念及原因……………………………………… (187)
第二节　多重共线性的后果……………………………………………… (191)
第三节　如何发现多重共线性…………………………………………… (199)
第四节　如何对多重共线性进行补救…………………………………… (201)

第八章　异方差性…………………………………………………………… (207)
第一节　异方差概念……………………………………………………… (207)
第二节　出现异方差时的 OLS 估计 …………………………………… (209)
第三节　异方差的检验…………………………………………………… (211)
第四节　异方差的校正…………………………………………………… (216)

第五节　实例……………………………………………………………………（219）

第九章　自相关分析…………………………………………………………（227）
第一节　自相关及其性质…………………………………………………………（227）
第二节　自相关下的 OLS 估计 ……………………………………………………（235）
第三节　自相关检验………………………………………………………………（236）
第四节　自相关模型的参数估计方法……………………………………………（239）
第五节　广义最小平方估计方法…………………………………………………（243）
第六节　分布滞后模型……………………………………………………………（251）

第十章　联立方程模型…………………………………………………………（266）
第一节　联立方程模型的提出……………………………………………………（266）
第二节　应用 OLS 估计的联立方程偏误 …………………………………………（269）
第三节　联立方程模型的变量和表示方法………………………………………（273）
第四节　联立方程模型的识别……………………………………………………（277）
第五节　间接最小二乘法…………………………………………………………（284）
第六节　工具变量法………………………………………………………………（287）
第七节　二阶段最小二乘法………………………………………………………（290）

第十一章　时间序列分析………………………………………………………（298）
第一节　时间序列数据的特性……………………………………………………（298）
第二节　平稳时间序列分析模型…………………………………………………（304）
第三节　非平稳随机过程与单位根检验…………………………………………（314）
第四节　协整与误差修正模型……………………………………………………（317）

第十二章　应用计量经济模型…………………………………………………（322）
第一节　生产函数…………………………………………………………………（322）
第二节　需求函数…………………………………………………………………（338）
第三节　消费函数…………………………………………………………………（346）
第四节　投资函数模型……………………………………………………………（352）

附表…………………………………………………………………………………（360）
参考文献……………………………………………………………………………（378）

21 世纪

第一章

概率论基础

概率论是经济学的重要数理基础，概率论是数学的一个分支。在经济学领域，许多理论与实践问题都需要运用现代概率论和统计方法，如线性统计推断、估计理论、鞅论、倒向随机微分方程理论等。本章就概率论内容做一些简要介绍。本章首先介绍了随机事件、随机变量及其分布以及随机事件、随机变量的独立性等内容，然后介绍了在经济学、金融学以及保险学中占有重要地位的随机变量数字特征以及它们的应用（如随机变量的数学期望、方差和相关系数），随后介绍了在保险中经常用到的变异系数概念和切比雪夫不等式（Chebyshev inequality），最后介绍了在经济学、金融学、保险学中有广泛应用的 n 维正态随机变量的性质、大数定律、中心极限定理等。

第一节　随机现象、随机试验和随机事件

一、随机现象

自然界和人类社会发生的现象是多种多样的。有一类现象在一定条件下必然发生，如向上抛一石子必然落地、同性电荷必互相排斥等，这类现象称为确定性现象。在自然界和人类社会还存在着另一类现象。例如，在相同条件下抛同一枚硬币，其结果可能是正面朝上，也可能是反面朝上，而且我们在每次抛掷之前无

法肯定抛掷的结果是什么；单位时间内记录某电话交换台收到用户呼唤的次数，我们在记录之前无法知道这一段时间内收到的呼唤次数。在一定条件下，这类现象可能出现这样的结果，也可能出现那样的结果，而我们在试验或观察之前不可能预知确切的结果。人们经过长期实践并深入研究之后，发现这类现象在大量重复试验或观察下，它的结果却出现某种规律性。例如，多次重复抛掷硬币得到正面朝上的结果大致是一半；重复记录用户的呼唤次数，它按照一定的规律分布，等等。这种在大量重复试验或观察中所显现出的固有规律性，称为统计规律性。

这种在个别试验中其结果显出不确定性，在大量重复试验中其结果又具有统计规律性的现象，称为随机现象。

二、随机试验

在概率论中，试验是一种含义广泛的术语，它包括各种各样的科学试验，如化学试验、物理试验等，甚至对某一事物的某一特征的观察也可认为是一个试验。概率论中所研究的试验具有以下特点：

(1) 在相同条件下试验可以重复进行；

(2) 每次试验的结果具有多种可能性，而且我们在试验之前可以明确试验的所有可能结果；

(3) 在每次试验之前不能准确地预言这次试验将出现哪一种结果。

满足以上三个条件的试验称为随机试验，有时简称为试验。例如：

E_1：抛一颗骰子，观察出现的点数；

E_2：记录电话交换台一分钟内收到的呼唤次数；

E_3：在一批灯泡中任意抽取一只，测试它的寿命；

E_4：记录某地区一昼夜的最高温度和最低温度。

三、随机事件

(一) 样本空间

随机试验 E 的所有可能结果组成的集合称为 E 的样本空间，记为 Ω。样本空间中的元素（即 E 的每个结果）称为样本点。例如，上面四个试验中的样本空间分别为：$\Omega_1=\{1,2,3,4,5,6\}$，$\Omega_2=\{0,1,2,\cdots\}$，$\Omega_3=\{t\mid t\geqslant 0\}$，$\Omega_4=\{(x,y)\mid T_0\leqslant x\leqslant y\leqslant T_1\}$。其中，$x$ 表示该地区的最低温度，y 表示该地区的最高温度，我们同时假设这一地区的温度不会小于 T_0，也不会大于 T_1，样本空间中的元素可以是有限个（Ω_1）、可数无穷多个（Ω_2）和不可数无穷多个（Ω_3，Ω_4）。

（二）随机事件

随机试验 E 的样本空间 Ω 的子集称为 E 的随机事件，简称为事件。在每次试验中，当且仅当这一子集中的一个样本点出现时，称这一事件发生。由一个样本点组成的单点集，称为基本事件。例如，在试验 E_1 中，A 表示掷一枚骰子出现偶数这一事件，则 $A=\{2,4,6\}$。试验中，只要 2、4、6 中任一数出现，则事件就发生了。$\{1\}$，$\{2\}$，…，$\{6\}$ 是这个试验的基本事件。

样本空间 Ω 包含所有的样本点，它是 Ω 自身的子集，在每次试验中必然发生，我们将其称为必然事件。空集 $\varnothing$ 不包含任何样本点，它也是样本空间的子集，它在每次试验中都不会发生，我们将其称为不可能事件。为研究方便，我们把这两个极端事件也称为随机事件。

（三）事件间的关系及事件的运算

设试验 E 的样本空间为 Ω，而 A、B、A_k（$k=1$，2，…）是 Ω 的子集。

（1）事件的包含。如果事件 A 发生必然导致事件 B 发生，则称事件 B 包含事件 A，记为 $A\subset B$。若 $A\subset B$ 且 $B\subset A$，则称事件 A 与事件 B 相等，记为 $A=B$。例如，A 表示某人在五十岁之前死亡；B 表示某人在六十岁之前死亡，则 $A\subset B$。

（2）和事件。在试验中表示事件 A 或事件 B 至少有一个发生的事件称为事件 A 与事件 B 的和事件，记为 $A\cup B$ 或 $A\cup B=\{x\mid x\in A$ 或 $x\in B\}$。例如，A 表示明天下雨，B 表示明天刮风，则 $A\cup B$ 表示“明天既刮风又下雨”、“明天下雨但不刮风”或“明天刮风却不下雨”这三个事件之一发生。

类似的，我们称 $\bigcup\limits_{k=1}^{n}A_k$ 为 n 个事件 A_1，A_2，…，A_n 的和事件；称 $\bigcup\limits_{k=1}^{\infty}A_k$ 为可列个事件 A_1，A_2，…的和事件。

（3）积（交）事件。在试验中表示事件 A 或事件 B 同时发生的事件称为事件 A 与事件 B 的积事件，记为 $A\cap B=\{x|x\in A$ 且 $x\in B\}$。当且仅当 A、B 同时发生时，事件 $A\cap B$ 发生，$A\cap B$ 也记为 AB。例如，A、B 若为（2）中所表示的事件，则 $A\cap B$ 表示“明天既刮风又下雨”。

类似的，我们称 $\bigcap\limits_{k=1}^{n}A_k$ 为 n 个事件 A_1，A_2，…，A_n 的积事件；称 $\bigcap\limits_{k=1}^{\infty}A_k$ 为可列个事件 A_1，A_2，…的积事件。

（4）差事件。我们将试验中表示事件 A 发生、事件 B 不发生的事件称为 A 事件和 B 事件的差事件，记为 $A-B$ 或 $A-B=\{x|x\in A$ 且 $x\notin B\}$。例如，A、B 若为（2）中所表示的事件，则 $A-B$ 表示“明天下雨但不刮风”。

（5）不相容事件。若 $A\cap B=\varnothing$，则称事件 A 与事件 B 是互不相容事件或互

斥事件，也就是事件 A 和事件 B 不能同时发生。基本事件是两两互不相容事件。例如，A 表示本地区干旱，B 表示本地区水灾，则 $A \cap B = \varnothing$，即这两个事件不可能同时发生。

（6）互为逆事件。若 $A \cup B = \Omega$，且 $A \cap B = \varnothing$，则称事件 A 与事件 B 互为逆事件或事件 A 与事件 B 互为对立事件，即每次试验时事件 A、事件 B 中必有一个发生，且仅有一个发生。A 的对立事件记为 $\overline{A}$，$\overline{A} = \Omega - A$。例如，$A$ 表示承保人赔款，B 表示承保人不赔款，则此两个事件为互逆事件。

四、随机事件运算法则

设 A、B、C 为事件，则有：

交换律：$A \cup B = B \cup A$；$A \cap B = B \cap A$　　(1.1.1)

结合律：$A \cup (B \cup C) = (A \cup B) \cup C$　　(1.1.2)

$A \cap (B \cap C) = (A \cap B) \cap C$

分配律：$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$　　(1.1.3)

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

德·摩根定律：$\overline{A \cup B} = \overline{A} \cap \overline{B}$

$\overline{A \cap B} = \overline{A} \cup \overline{B}$　　(1.1.4)

第二节　事件的频率与概率

一、事件的频率

（一）频率的定义

定义 1.1　在相同的条件下进行了 n 次试验，在这 n 次试验中，事件 A 发生的次数记为 n_A，称为事件 A 发生的频数。n_A/n 称为事件 A 发生的频率，并记成 $f_n(A)$。

（二）频率的基本性质

频率的性质：

（1）$0 \leqslant f_n(A) \leqslant 1$；　　(1.2.1)

（2）$f_n(\Omega) = 1$；　　(1.2.2)

（3）若 A_1，A_2，…，A_k 是两两互不相容的事件，则

$f_n(A_1 \cup A_2 \cup \cdots \cup A_k) = f_n(A_1) + f_n(A_2) + \cdots + f_n(A_k)$　　(1.2.3)

由于事件 A 发生的频率是它发生的次数与试验次数之比，其大小表示 A 发生的频繁程度。频率越大，表示事件 A 发生得越频繁，这意味着 A 在一次试验中发生的可能性越大。因而，人们直观的想法是用频率来表示 A 在一次试验中发生的可能性的大小。

历史上有人做过抛硬币的试验，其结果如表 1—1 所示（H 表示出现正面的事件）：

表 1—1　　抛硬币试验结果

试验者	n	n_H	$f_n(H)$
德·摩根	2 048	1 061	0.518 1
蒲丰	4 040	2 048	0.506 9
K. 皮尔逊	12 000	6 019	0.501 6
K. 皮尔逊	24 000	12 012	0.500 5

结果表明，当 n 逐渐增大时，频率 $f_n(H)$ 逐渐稳定于一个常数。对于每一个事件 A 都有这样一个客观存在的常数与之相对应。这种"频率稳定性"就是我们通常所说的统计规律性，它揭示了隐藏在随机现象中的规律性。因此，用频率稳定值来表示事件发生的可能性大小是合适的。

二、事件的概率及其性质

我们从频率的稳定性和频率的性质得到启发，给出如下表征事件发生可能性大小的概率定义。

（一）概率的定义

定义 1.2　设 E 是随机试验，Ω 是它的样本空间。对于 E 的每一个事件 A 赋予一个实数，记为 $P(A)$，称为事件 A 的概率，如果集合函数 $P(A)$ 满足下列条件：

（1）对于每个事件 A，有 $P(A)\geqslant 0$；　　(1.2.4)

（2）对于必然事件 Ω，有 $P(\Omega)=1$；　　(1.2.5)

（3）设 A_1，A_2，…是两两不相容的事件，即对于 $i\neq j$，$A_iA_j=\varnothing$，i，$j=1$，2，…，则有

$$P(A_1\cup A_2\cup\cdots)=P(A_1)+P(A_2)+\cdots \tag{1.2.6}$$

此式称为概率的可列可加性。

（二）概率的性质

概率的性质：

（1）$P(\varnothing)=0$　　(1.2.7)

（2）若 A_1，A_2，…，A_n 是两两互不相容的事件，则有

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n) \tag{1.2.8}$$

（3）设 A、B 是两个事件，若 $A \subset B$，则有 $P(A) \leqslant P(B)$，更有 $0 \leqslant P(A) \leqslant 1$

（4）对于任一事件 A，有 $P(\overline{A}) = 1 - P(A)$ (1.2.9)

（5）对于任意两事件 A、B，有 $P(A \cup B) = P(A) + P(B) - P(AB)$。此性质推广到任意的 n 个事件 A_1，A_2，…，A_n 之和，则有

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \leqslant i < j \leqslant n} P(A_i A_j) + \sum_{1 \leqslant i < j < k \leqslant n} P(A_i A_j A_k) + \cdots + (-1)^{n-1} P(A_1 A_2 \cdots A_n) \tag{1.2.10}$$

三、古典概型

若试验 E 具有如下特点：①试验的样本空间只包含有限个元素，②试验中每个基本事件发生的可能性相同，则称这种试验为等可能概型或古典概型。

设 A 为等可能概型中的事件，则 A 的概率计算公式为：

$$P(A) = \frac{A\text{ 中包含的基本事件数}}{\Omega\text{ 中基本事件的总数}} \tag{1.2.11}$$

例 1.1 一批产品共有 100 件，其中有 5 件次品，从中任取 8 件，求所取 8 件中至多有 2 件是次品的概率？

解：产品共有 100 件，从中取 8 件，一共有 C_{100}^{8} 种取法。

以事件 A 表示"所取 8 件中至多有 2 件是次品"这一事件，则 A 所包含的基本事件为"所取 8 件中没有次品"、"有一件次品"、"有两件次品"这三部分基本事件数的总和：

$$C_{95}^{8} \cdot C_5^0 + C_{95}^{7} \cdot C_5^1 + C_{95}^{6} \cdot C_5^2$$

式中第一项为全取到正品的基本事件数，第二项为取到一件次品的基本事件数，第三项为取到两件次品的基本事件数，于是所求概率为：

$$P(A) = \frac{C_{95}^{8} \cdot C_5^0 + C_{95}^{7} \cdot C_5^1 + C_{95}^{6} \cdot C_5^2}{C_{100}^{8}} \approx 0.997$$

例 1.2 全班有 25 人，设每人的生日在 365 天中的可能性是均等的。问此 25 人生日都不同的概率是多少？其中至少有两人生日相同的概率是多少？

解：设 A 表示"25 人的生日都不相同"这一事件，则 $\overline{A}$ 为"至少有两人的生日相同"。由假设 25 人都有可能在 365 天中的任一天出生，故样本空间 Ω 中的基本事件总数为 $365 \times 365 \times \cdots \times 365 = 365^{25}$。25 人生日都不相同的排列数为