



全国高等院校21世纪新创规划教材



计量经济学

逢守艳 主 编



中国科学技术出版社
CHINA SCIENCE AND TECHNOLOGY PRESS

全国高等院校 21 世纪新创规划教材

计量经济学

逢守艳 主编

中国科学技术出版社

CHINA SCIENCE AND TECHNOLOGY PRESS

· 北 京 ·

BEIJING

图书在版编目(CIP)数据

计量经济学/逢守艳主编. —北京:中国科学技术出版社,2009.2

全国高等院校 21 世纪新创规规划教材

ISBN 978 - 7 - 5046 - 5384 - 0

I. 计… II. 逢… 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(2009)第 008072 号

自 2006 年 4 月起本社图书封面均贴有防伪标志,未贴防伪标志的为盗版图书。

内 容 提 要

本教材主要介绍经典线性回归模型、可线性化的非线性回归模型;违背经典假定回归模型,如异方差、自相关、多重共线性;模型中的特殊解释变量及联立方程模型等问题。介绍了几种典型的计量经济学应用模型;初步的介绍时间序列模型。同步介绍 Eviews 分析软件。

本书可作为全日制高等学校财经类专业学习经济计量学课程的教材,也可以作为相关的职业培训和自学用书。

中国科学技术出版社出版

北京市海淀区中关村南大街 16 号 邮政编码:100081

策划编辑	林 培 孙卫华	责任校对	林 华
责任编辑	林 培 王 强	责任印制	安利平

电话:010 - 62103210 传真:010 - 62183872

<http://www.kjpbooks.com.cn>

科学普及出版社发行部发行

北京蓝空印刷厂印刷

*

开本:787 毫米×1092 毫米 1/16 印张:13.25 字数:320 千字

2009 年 3 月第 1 版 2009 年 3 月第 1 次印刷 定价:25.00 元

ISBN 978 - 7 - 5046 - 5384 - 0/F · 630

(凡购买本社的图书,如有缺页、倒页、
脱页者,本社发行部负责调换)

前 言

计量经济学是经济学科中的一个重要分支，是宏观经济和微观经济研究的必不可少的重要工具，已在经济学科的课程中获得了与宏观经济学和微观经济学同等重要的地位，是经济学术研究的主流，是国际学术交流的共同语言。

计量经济学的内容体系非常广泛，本教材力求系统介绍经济计量学的基本原理和方法，强调理论与实际工作的结合，注重培养学生解决实际问题的能力。为此，本教材借助经济系统的实际统计数据 and 已有的经济理论构建经济模型，以经济预测和经济结构研究为主线，从模型的设计、估计、检验和应用各个角度，由简单到复杂的阐述其建立各类经济模型的方法。同步介绍 Eviews 分析软件，使学生学完本课程后，不仅能够掌握经济计量学的理论和方法，而且能够学会应用计量经济学解决实际问题。

计量经济学是经济学类的核心课程，也是管理类的各专业的关键课程之一，本书可作为全日制高等学校财经类专业学习经济计量学课程的教材，也可以作为相关的职业培训和自学用书。

参加本书编写的人员有：王涛、宋德军（第一章），张虹敏（第二章），逢守艳（第三、四、六、七、八章），田秀杰（第五、九章），付建华（第十章）。本书由逢守艳任主编，田秀杰任副主编，王涛主审。逢守艳、王涛对书稿进行了修改和完善。

在本书的编写过程中，作者参考了部分国内出版的相关书籍，选用了个别案例，引用了一些观点，在此谨向有关参考书的作者表示感谢。

由于作者水平所限，书中难免会有不足之处，我们热忱希望读者提出宝贵意见。

编 者
2009 年 1 月

目 录

第一章 绪 论	1
第一节 计量经济学简介	1
第二节 计量经济模型	3
第三节 计量经济学研究问题的步骤	6
习 题	8
第二章 统计学基础知识	9
第一节 随机变量	9
第二节 常用的随机变量及概率	10
第三节 大数定律与中心极限定理	13
第四节 样本统计量的分布	15
第五节 参数估计与假设检验	17
第六节 相关关系的度量	22
习 题	27
第三章 一元线性回归模型	28
第一节 一元线性回归模型	28
第二节 参数的最小二乘估计	30
第三节 一元线性回归模型的检验	38
第四节 利用回归方程进行预测	43
第五节 案例分析	49
习 题	52
第四章 多元线性回归模型	55
第一节 模型的建立及假定条件	55
第二节 多变量回归模型的最小二乘估计	57
第三节 多元线性回归模型的检验	64
第四节 利用多元线性回归方程进行预测	69
第五节 案例分析	71
习 题	79
第五章 非线性模型	82
第一节 变量间的非线性关系	82
第二节 线性化方法	83
第三节 不可线性化的非线性回归模型的估计	88
第四节 案例分析	88
习 题	92

第六章 违背经典假定回归模型	93
第一节 异方差	93
第二节 自相关	106
第三节 多重共线性	121
习 题	131
第七章 模型中的特殊解释变量	133
第一节 虚拟变量	133
第二节 随机解释变量	141
第三节 滞后变量	145
第四节 案例分析	154
习 题	157
第八章 联立方程模型	159
第一节 联立方程模型的一般问题	159
第二节 联立方程模型的识别	163
第三节 联立方程模型的识别条件	165
第四节 联立方程模型的估计	167
第五节 案例分析	171
习 题	175
第九章 几种典型的计量经济学模型	177
第一节 需求函数模型	177
第二节 消费函数模型	186
第三节 生产函数模型	189
习 题	198
第十章 时间序列分析	199
第一节 几个基本概念	199
第二节 自回归过程	200
第三节 预测	203
第四节 协整理论简介	204
习 题	205
参考文献	206

第一章 绪 论

第一节 计量经济学简介

一、计量经济学的产生与发展

计量经济学一词是1926年挪威的经济学家费瑞希(Ragnar Frish)仿照“生物计量学”一词提出的。并将它定义为经济理论、统计学和数学三者的结合。1930年12月29日由费瑞希和荷兰经济学家丁伯根(Tinbergen)等经济学家在美国成立了国际计量经济学会,并于1933年创办了《Econometrics(计量经济学)》学刊,标志着计量经济学作为一个独立学科正式诞生了。而计量经济学的诞生在某种程度上表明经济学走入了其科学发展的成熟期。从20世纪30年代到今天,尤其是在第二次世界大战以后,计量经济学在西方各国的影响迅速发展。美国著名经济学家萨缪尔森(P. A. Samuelson)曾说:“第二次世界大战以后的经济学是计量经济学的时代”。1969年首届诺贝尔经济学奖授予费瑞希和丁伯根,高度评价他们“开发了经济分析过程的动态模型,并使之实用化”。

在市场经济发达的国家,各个企业、各个部门之间存在着错综复杂的关系,企业要使自己在激烈的市场竞争中生存,必须有可靠的市场预测;政府为干预国民经济运行,更需要及时分析经济动态。企业和政府都十分重视以计量经济方法为基础的关于经济景气、循环周期的研究以及经济政策的模拟、预测和决策分析。而这些经济分析的现实需要,是计量经济学产生的根本原因。

在计量经济学的发展进程中,产生了经济学史上许多著名的具有里程碑性质的理论成果。20世纪20~40年代是计量经济学发展的基础时期,也可以叫做单方程时期。20年代是该学科基本思想形成和产生的关键时期;30年代计量经济学主要用于研究微观经济:如H. 舒尔兹在消费理论与市场行为方面的研究;P. 道格拉斯(Douglas)对边际生产力的研究;J. 丁伯根在景气循环方面的研究,都为计量经济学开拓了新领域。R. 费瑞希以统计学和经济理论为基础来测定需求弹性、边际生产力以及总体经济的稳定性,是一大贡献。40年代计量经济学的重点转向研究宏观经济领域,同时计量经济学家致力于经济理论的模型化与数学化的研究。如T. 哈威勒莫(Hoavelmo), A. 瓦尔德(Wald)将统计推断应用于计量经济学以及50年代, H. 泰尔(Tbeil)发表了二阶段最小二乘法;60年代以后有关分布滞后的新处理方法等。

20世纪50~70年代是计量经济学大发展的时期,该时期以联立方程的研究为主要特点。由于计算机的广泛使用,大量复杂的经济计量模型得以建立和应用,促进了计量经济学理论与应用的发展。1950年以库坡曼(Koopman)发表论文“动态经济模型的统计推断”和库坡曼-郝得(Koopman - Hood)发表论文“线性联立经济关系的估计”为标志,计量经济学进入了联立方程模型时代。以联立方程描述一个国家整体的宏观经济活动,最具代表性的人物

是1980年获得诺贝尔经济学奖的美国经济学家克莱因(Klein)教授。其主要研究成果是美国经济波动模型1921~1941年(1950作)和包括20个方程的美国宏观经济模型1928~1950年(1955作),可以说是计量经济学的第二个里程碑。20世纪70年代以后,计量经济学家致力于更大型宏观模型的研究,并从国内向国际模型扩展。在西方最为著名的是联立方程模型是“连接计划”(Link Project),从最初的包括18个国家,7447个方程和3368个外生变量的模型,发展到1987年的78个国家2万多个方程的大型联立模型。而在苏联、东欧一些国家大量编制投入产出模型,并取得了许多有益的成果。

20世纪80年代以后,计量经济学进入了动态建模研究的新阶段。如迪基-福勒(Dickey-Fuller)推出的检验时间序列非平稳性单位根检验方法;韩德瑞(Hendry)的动态建模理论;西姆斯(Sims)的向量自回归模型(VAR)等。1987年恩格尔-格兰杰(Engle-Granger)发表论文“协同积分与误差修正,描述、估计与检验”。该文正式推出了协同积分(Cointegration),即协整的概念。从而把计量经济学的理论研究方法推向了新的阶段。

我国学者自20世纪80年代开始,在计量经济学的理论研究上也作出了一定的贡献,如现代对策论,贝叶斯理论在计量经济学中的应用。这是目前计量经济学研究的一个新课题。

二、计量经济学的定义

计量经济学的创始人之一费瑞希在《计量经济学》杂志的发刊词中有一段话:“用数学方法探讨经济学可以从好几个方面着手,但任何一方面都不能与计量经济学混为一谈。计量经济学与经济统计学绝非是一码事;它也不同于我们所说的一般经济理论,尽管经济理论大部分都具有一定的数量特征;计量经济学也不应视为数学应用于经济学的同义语。经验表明,统计学、经济理论和数学这三者对于真正了解现代经济生活中的数量关系来说,都是必要的。三者结合起来,就有力量,这种结合便构成了计量经济学。”由此可见,计量经济学是以一定的经济理论和实际经济资料为基础,运用数学、统计方法与电脑技术,以建立计量经济模型为主要手段,定量分析研究具有随机性特征的经济变量关系。

计量经济学在长期的发展过程中逐步形成了两个分支:理论计量经济学和应用计量经济学。理论计量经济学是以介绍、研究计量经济学的理论与方法为主要内容,侧重于理论与方法的数学证明与推导,与数理统计联系极为密切。而应用计量经济学则以建立与应用计量经济学模型为主要内容,侧重于建立与应用模型过程中实际问题的处理。

三、计量经济学的特点

计量经济学是定量研究具有随机性特征的经济变量关系的数学模型。注重经济变量的随机性特征,是计量经济学的显著特征。由于实际的经济运行不是在实验室进行的,往往存在一些不确定的随机因素,使得经济变量之间的关系不能表示成精确的函数关系。人们只能在模型中列出对所研究变量起主要影响作用的变量,将不重要的因素和一些不确定因素归并到一个随机变量中,进而来简化变量之间的数学模型。

例如研究需求函数。经济理论假定某商品的需求量取决于它的价格与代用品价格、消费者的收入和消费者的偏好等因素。然而在实际经济生活中,除了这些因素外,还有其他一些不重要的以及随机因素的影响。例如新产品的发明、职业的改变、气候条件的变化等。另外,人们可能受到谣传、广告的影响。即使市场价格、消费者收入和消费者偏好都不变,商品

的需求量也受到影响。在计量经济学中,这些不重要因素及随机因素的影响也要反映在数学模型中,于是引进一个随机变量 U ,建立下列形式的某商品需求量数学模型:

设: Q ——某一特定商品的需求量;

P_1 ——该商品价格;

P_2 ——相关商品的价格;

Y ——消费者收入;

T ——消费偏好;

U ——影响商品需求量的其他因素和随机因素;

b_i ——需求函数的回归系数(待定参数)。

又假定它们之间是线性关系,则:

$$Q = \beta_0 + \beta_1 P_1 + \beta_2 P_2 + \beta_3 Y + \beta_4 T + U \quad (1.1)$$

该模型可分为两大部分,其一是最后一项 U ,称为随机扰动项;其二是去掉随机扰动项部分,我们称之为数量经济模型。可见计量经济模型,就是在数量经济模型的基础上引入了随机扰动项 U 构成的,这是计量经济模型最显著的特点。通过对随机扰动项 U 的研究,实现了计量经济模型的实用化,我们才能够利用现实统计资料提供的的数据,估计模型参数,给出经济关系的具体数量表示。经济研究也才能真正进入实证的定量分析阶段。

第二节 计量经济模型

计量经济学的任务是用数学模型方法研究客观经济系统的数量关系,也是计量经济学的特点。按照通常的解释,模型(Model)是现实系统的代表。计量经济模型是对现实经济系统的数学抽象。模型不可能复制现实系统的全部属性。因此,我们的任务是要通过理论和实践,将现实经济系统抽象为数学模型。

计量经济模型可以分为两大类:单一方程模型和联立方程模型。单一方程模型用来描述经济领域一个因变量和若干个自变量间的结构关系;联立方程模型用来描述经济领域多个因变量和多个自变量间的结构关系。

计量经济模型是由变量和参数构成的。在计量经济学中经济变量和经济参数都有专门的名称。

一、经济变量

经济变量是一个序列,它包括若干个已知的样本数据。计量经济学就是利用这些变量的样本数据和计量经济学方法,估计数学模型中的待定系数,以确定经济变量之间的结构关系。

(一) 内生变量与外生变量

对一个独立的经济模型(无论由多少个方程组成)来说,变量可以分为两类:内生变量和外生变量。内生变量数值是由模型自身(经济系统)决定的随机变量。外生变量数值是由模型(经济系统)之外因素决定的非随机变量。

$$\begin{cases} C_t = \alpha + \beta Y_t + U_t & (1.2) \\ Y_t = C_t + I_t + G_t & (1.3) \end{cases}$$

在模型

式中: Y_t , C_t , 为内生变量, I_t, G_t 为外生变量。

内生变量影响模型中其他内生变量,但同时又受外生变量影响;而外生变量影响模型中的内生变量,但不受模型中任何变量影响。

(二)解释变量与被解释变量

在计量经济学的模型中,把方程中的自变量称作解释变量,因变量称作被解释变量。在模型(1.2)中 C_t 是被解释变量, Y_t 是解释变量;而在模型(1.3)中是 Y_t 被解释变量。

被解释变量一定是模型的内生变量,而解释变量既包括模型的外生变量又包括模型的内生变量。

(三)滞后变量与前定变量

当内生变量的前期值 Y_{t-1} 作为解释变量时,称为滞后变量。例如收入决定模型中的前期收入 Y_{t-1} ,就是滞后变量。滞后变量显然在求解模型之前是确定的量,因此,通常将外生变量和滞后变量合称为前定变量。

二、经济参数

在计量经济学的模型中,方程中的系数都有一定的经济意义,我们称之为经济参数。经济参数有两类:外生参数和内生参数。外生参数一般是指依据经济法规人为确定的参数,如固定资产折旧率、税率、利息率等等,有些外生参数有时是凭经验估计的。在计量经济模型中,绝大多数参数都是内生参数,内生参数是依据样本观察数据,运用统计方法估计得到的。如何选择估计参数的方法和改进估计参数的方法,这是理论经济学的基本任务。

三、经济数据

经济数据是拟合经济数学模型的原料。在计量经济学中把用来拟合计量经济模型的数据分为三大类:时序数据、截面数据和混合数据。

(一)时序数据

时序数据即时间序列数据。时间序列数据是同一统计指标按时间顺序记录的数据列。在同一数据列中的各个数据必须是同口径的,要求具有可比性。时序数据可以是时期数,也可以是时点数。

(二)截面数据

截面数据是在同一时间,不同统计单位的相同统计指标组成的数据列。与时序数据比较,其区别在于组成数据列的各数据的排列标准不同,时序数据是按时间顺序排列,横截面数据是按统计单位排列。例如,为了研究某一行业各企业产出与投入的关系,我们需要关于同一时间截面上各企业的产出 Q 和劳动投入 L ,资本投入 K 的截面数据。

(三)混合数据

有时为了分析需要,人们常常采用时间序列和横截面数据合并的统计资料。即将若干期的时间序列和每期内的横截面数据合并作为样本数据。时间序列和横截面数据合并的统计资料要求同时具有以上两种统计资料的特征。

例如,表 1-1 为 2002~2006 年我国各地区城乡居民人民币储蓄存款,其样本数据的个

表 1-1 各地区城乡居民人民币储蓄存款(年底余额)

单位:亿元

年 份 地 区	2002	2003	2004	2005	2006
北 京	4389.7	5293.5	6122.4	7477.7	8705.6
天 津	1486.4	1825.3	2116.7	2461.5	2808.1
河 北	4808.3	5457.0	6207.5	7084.0	8014.2
山 西	2307.3	2781.5	3342.3	4119.7	4796.2
内 蒙 古	1137.9	1355.5	1603.9	1973.6	2271.4
辽 宁	4665.0	5434.7	6048.5	6950.2	7701.4
吉 林	1878.5	2161.4	2405.6	2798.1	3107.5
黑 龙 江	2915.7	3342.4	3585.5	4078.6	4373.6
上 海	3891.5	5103.2	6116.1	7665.6	8730.0
江 苏	6276.2	7638.2	8863.1	10581.3	12183.5
浙 江	5212.7	6452.2	7364.1	8746.0	10473.5
安 徽	2047.5	2475.8	2972.4	3508.7	4077.8
福 建	2430.5	2924.7	3322.3	3903.1	4478.3
江 西	1706.6	2015.5	2347.7	2752.9	3151.7
山 东	5803.5	6768.4	7721.5	9035.1	10358.0
河 南	4196.0	4919.1	5607.3	6488.6	7367.4
湖 北	2754.5	3296.5	3860.7	4465.8	5103.6
湖 南	2576.4	3036.5	3483.2	4092.1	4762.3
广 东	11813.3	14061.8	16193.4	19051.4	21584.6
广 西	1733.5	1971.7	2240.1	2561.3	2946.2
海 南	483.5	546.9	615.9	697.6	790.6
重 庆	1582.3	1896.9	2189.7	2545.9	2949.1
四 川	3665.2	4333.8	5019.4	5902.7	6787.7
贵 州	758.7	912.8	1094.6	1350.9	1596.9
云 南	1499.8	1766.5	2052.1	2430.3	2854.9
西 藏	70.4	91.9	107.5	123.1	139.8
陕 西	2108.1	2519.9	2948.4	3534.0	4067.7
甘 肃	1042.4	1217.4	1384.9	1586.7	1825.4
青 海	222.4	260.5	299.3	348.9	406.3
宁 夏	306.8	377.7	425.5	509.5	581.1
新 疆	1137.6	1371.8	1534.7	1816.4	2035.6

数据来源:《2007 年中国统计年鉴》

数为 $5 \times 31 = 155$ 个。每一行是时间序列数据,每一列又是截面数据,他们统计口径相同,都具有可比性。

四、计量经济模型的作用

计量经济分析的主要作用有三点:

(一) 结构分析

应用计量经济模型对经济变量之间的关系作出定量的度量。它不同于人们通常所说的,诸如产业结构、产品结构、消费结构、投资结构中的结构分析。它研究的是当一个变量或几个变量发生变化时会对其他变量以至经济系统产生什么样的影响,从这个意义上讲,我们所进行的经济系统定量研究工作,说到底,如上面提到的需求函数,假如利用统计资料估计出模型中变量 P 的回归系数等于 -0.12 ,表明当其他因素不变时,商品的价格每提高一个单位,此商品的需求量将减少 0.12 个单位。

(二) 预测未来

预测未来是指应用已建立的计量经济模型求因变量未来一段时期的预测值。如上面的需求函数,假如给出某商品的价格、代用品价格、消费者收入及消费者偏好 2006 年的估计值,就可以求出 2006 年某商品需求 t 预测值,为 2006 年该商品的生产和供给提供可靠的依据。

(三) 政策评价

政策评价是指从许多不同的政策中选择较好的政策予以实行,或者说是研究不同的政策对经济目标所产生的影响的差异。经济数学模型可以起到“经济政策实验室”的作用。尤其是计量经济学模型,揭示了经济系统中变量之间的相互联系,将经济目标作为被解释变量,经济政策作为解释变量,可以很方便的评价各种不同的政策对目标的影响。将计量经济学模型和计算机技术结合起来,可以建成名副其实的“经济政策实验室”。如给出某商品的各种不同价格,分别预测未来各种不同的需求量,由此确定比较合适的商品价格。

第三节 计量经济学研究问题的步骤

计量经济学研究问题包括建立模型和应用模型两个相互关联的基本环节,可分为以下四个工作步骤:

一、建立模型

根据经济理论分析所研究的经济现象,找出经济变量间的因果关系及相互间的联系。把要研究的经济变量作为被解释变量,影响被解释变量的主要因素作为解释变量,影响被解释变量的非主要因素及随机因素归并到随机项,按照它们之间的相互联系的结构关系,建立模型。

二、估计参数

模型建立以后,估计模型的参数,即根据统计资料,选择适当的方法,如最小二乘法,求出模型中的待定系数。参数已经确定,模型中各变量间的相互关系就确定了,模型也就确定了。

三、检验模型

模型确定后,是否符合实际,能否解释经济过程,还需要对模型进行检验。检验包括:经济意义的检验、统计检验、经济计量假设合理性检验和模型的应用检验。

(一)经济意义的检验

经济意义的检验主要检验各个参数是否与经济理论符合,与实际经济问题符合。如某商品的需求量一般应随此商品价格的提高而减少,如果估计出的价格参数是正的,则说明建立的模型不合适或使用的估计方法不合适,或采用的样本数据有问题。

(二)统计意义检验

统计意义检验就是利用数理统计中假设检验的原理,对参数估计的可靠程度,观察数据的拟合程度等内容进行的检验。

(三)经济计量学检验

经济计量学检验是指对各种经济计量的假设的合理性,以及模型总体结构预测功能进行的检验。

(四)模型的应用检验

一般通过经济计量检验的模型还要接受实践的检验如果模型预测的误差较大,就要修改模型,有时可能要重新建立模型。

通过上述检验,才能实际应用,而且可以相对的稳定的使用一定时期。如果以上的检验出现问题,应采取相应的办法予以补救。如改变模型的形式,变换估计方法,重新选取样本数据,修正样本数。

四、模型的应用

单一方程模型主要用于经济预测,联立方程模型除用于经济预测外还用于经济结构分析与经济政策评价。

由于电子计算机的迅速发展,人们将计量经济学的计算方法编制成软件包,不仅可以估计参数,进行统计检验,而且可以进行经济预测,使许多非常复杂的计算问题得以很容易地解决,为计量经济学的学习和应用带来极大的方便。本书中,我们将介绍计量经济分析软件包 Eviews。Eviews 直观、易学,是目前人们最常用的计量经济学软件包之一,本书中的计量经济模型均可用 Eviews 解决。

习 题

1. 试述什么是计量经济学？它的研究对象、方法又如何？
2. 试述经济计量学与经济理论、数理经济学、经济统计学和数理统计学的区别。
3. 计量经济学中应用的数据是怎样进行分类的？试分别举出时间序列数据、横截面数据、混合数据、虚拟变量数据的实例。
4. 在建模时，一般都需要哪些过程？模型建好之后，要进行哪些必要的检验？
5. 为了减少对外国石油供给的依赖，政府正考虑对汽油收取联邦税。假设福特公司雇佣你分析税收增加对需求量的影响，你将如何向公司提出建议？
6. 假设美国总统正考虑对进口钢材征收关税以保护国内钢铁行业的利益。作为总统经济顾问，你的建议是什么？你如何展开经济计量研究以分析征收关税的影响？

第二章 统计学基础知识

统计学为计量经济学提供了方法,要学好计量经济学就必须掌握统计学有关的理论和方法。本章将对计量经济分析中,使用频率很高的统计知识,进行初步的学习。

第一节 随机变量

一、随机变量的概念

随机变量是一次随机试验的结果的数值性描述。随机变量一般用希腊 ξ, ζ, η 字母或大写的英文字母 X, Y, Z 等表示。

例如:

(1)一个射手对目标进行射击,击中目标记为 1 分,未击中目标记为 0 分。如果用 X 表示射手在一次射击中的得分,则它是一个随机变量, X 可以取 0 和 1 两个可能值。

(2)考查“掷硬币”这一试验,它有两个可能结果:“出现正面”用“1”表示,“出现反面”用“0”表示。如果用 X 表示试验的结果,则可以取 0 和 1 两个可能值。

(3)单位面积上稻谷的产量 X 是一个随机变量,它可以取一个区间内的一切实数值,即 $X \in [0, T]$, T 为某一个常数。

(4)一个沿着数轴进行随机运动的质点,它在数轴上的位置 X 是一个随机变量,可以取任何实数,即 $X \in (-\infty, +\infty)$ 。

随机变量按其取值情况可以分为两类:

1. 离散型随机变量

随机变量的可能值是有限个或可列个。描述离散型随机变量的取每一个值的概率情况用概率分布。上例中的(1)和(2)。

2. 连续型随机变量

随机变量的可能值充满一个或多个区间。描述连续型随机型随机变量在任一个区间取值的概率情况用密度函数。上例中的(3)和(4)。

二、随机变量的数字特征

在很多情况下,人们并不需要全面地考察随机变量的变化情况,而只要知道随机变量的一些数字特征。例如,在测量某零件长度时,由于种种偶然因素的影响,零件长度的测量结果是一个随机变量。但是,人们一般关心的是这个零件的平均长度和零件长度的离散程度。平均长度和零件长度的离散程度等都是零件长度这个随机变量的数字特征。常用的随机变量的数字特征有数学期望、方差。

(一) 随机变量的数学期望

1. 离散型随机变量的数学期望

定义 2.1: 设 X 为离散型随机变量, 其分布律为 $P\{x = x_i\} = p_i, i = 1, 2, \dots, n$ 如果级数 $\sum_i x_i p_i$ 绝对收敛, 则此级数为 X 的数学期望(或均值)既为 $E(X)$, 即

$$E(X) = \sum x p \quad (2.1)$$

意义: $E(X)$ 表示 X 取值的(加权)平均值。

2. 连续型随机变量的数学期望

定义 2.2: 设 X 为连续型随机变量, 其概率密度为 $f(X)$, 如果积分 $\int_{-\infty}^{+\infty} f(x) dx$ 绝对收敛, 则称此积分为 X 的数学期望(或均值)即为 $E(X)$, 即

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx \quad (2.2)$$

3. 数学期望的性质

- (1) 若 C 为常数, 则 $E(C) = C$
 - (2) 若 C 为常数, X 为随机变量, 则 $E(CX) = CE(X)$
 - (3) 设 X, Y 为任意两个随机变量, 则 $E(X \pm Y) = E(X) \pm E(Y)$
 - (4) 如果 X, Y 相互独立, 则有 $E(XY) = E(X)E(Y)$
- 性质(3)、(4)可以推广为 n 个随机变量。

(二) 随机变量的方差

1. 方差定义

定义 2.3: 设 X 为随机变量, 如果 $E[X - E(X)]^2$ 存在, 则称之为 X 的方差, 记为 $D(X)$, 即

$$D(X) = E[X - E(X)]^2 \quad (2.3)$$

2. 方差意义

方差是度量随机变量变异性的量。

3. 方差公式

$$\text{Var}(X) = E(X^2) - [E(X)]^2 \quad (2.4)$$

4. 方差的性质

- (1) 若 C 为常数, 则 $D(C) = 0$
- (2) 若 C 为常数, X 为随机变量, 则 $D(CX) = C^2 D(X)$
- (3) 如果 X, Y 相互独立, 则有 $D(X \pm Y) = D(X) + D(Y)$

第二节 常用的随机变量及概率

一、正态概率分布

正态分布在随机变量分布中占有特别重要的地位。在自然现象和社会现象中, 有许多

是可以用来描述。由正态分布出发,导出了一系列重要的抽样分布,如 t 分布、 χ^2 分布、 F 分布等。

若随机变量 X 的概率密度为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2.5)$$

式中: μ —— 均值;

σ —— 标准差;

π —— 3.1415;

e —— 2.71828。

则称 X 服从参数为 μ 、 σ^2 的正态分布,记为 $X \sim N(\mu, \sigma^2)$ 。正态分布的图形如图 2-1。

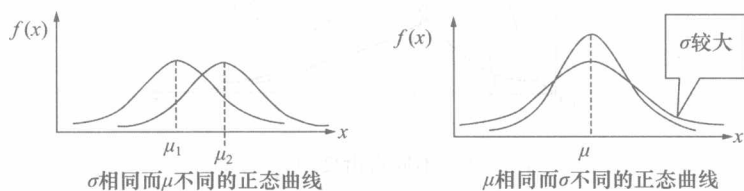


图 2-1 正态分布

正态曲线性质:

(1) 它是单峰曲线,它的最大值应在 $x = \mu$, 最大值为 $\frac{1}{\sqrt{2\pi}\sigma}$ 。

(2) 曲线以 $x = \mu$ 为对称轴,沿不同的水平方向单调下降,当 $x \rightarrow \pm\infty$ 时,曲线以 X 轴为其渐近线。

(3) 曲线在 $x = \mu \pm \sigma$ 处有拐点,拐点离对称轴有同样的距离,曲线在拐点处改变运行方向。

(4) 正态曲线的形状由方差决定, σ 越大,曲线越平滑; σ 越小,曲线陡峭。

(5) 正态曲线位置由数学期望 μ 决定。

当总体均值 $\mu = 0$, 方差 $\sigma^2 = 1$ 时,正态分布称为标准正态分布,记为 $N(0,1)$ 。若 $X \sim N(0,1)$, 则 X 在任意区间的概率都可查表求值。

任意一个正态分布的随机变量,我们都可以把它转换成服从标准正态分布的随机变量。转换的工具是标准正态变换:

$$Z = \frac{x - \mu}{\sigma} \quad (2.6)$$

二、 χ^2 分布

定义 2.4: $X_1, X_2, \dots, X_n$ 相互独立,且服从标准正态分布 $N(0,1)$ 令

$$\chi^2 = X_1^2 + X_2^2 + \dots + X_n^2 \quad (2.7)$$

则 χ^2 服从自由度为 n 的 χ^2 分布,计作 $\chi^2 \sim \chi^2(n)$ 。

χ^2 分布的概率密度为: