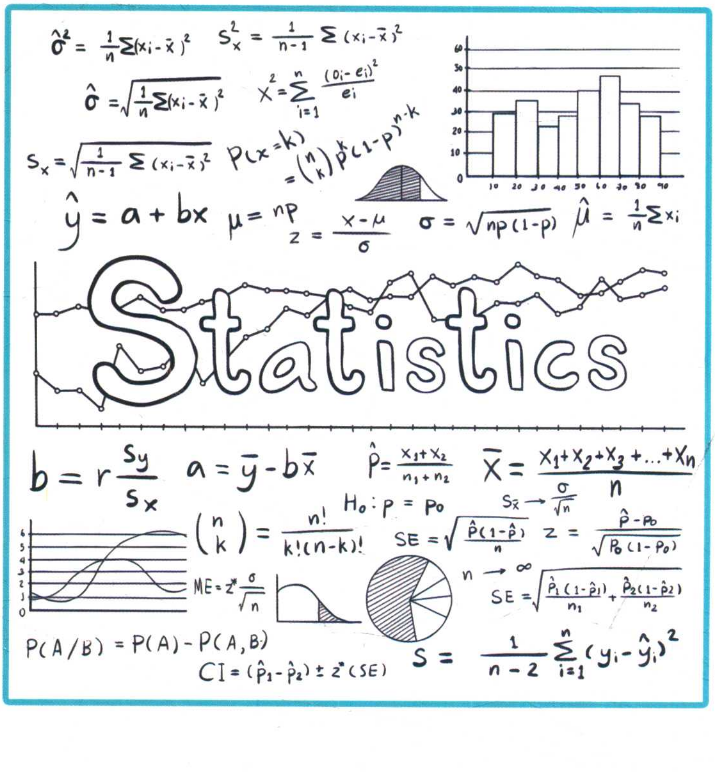
数理统计

MATHEMATICAL STATISTICS

程细玉　程　璟　编著



厦门大学出版社
XIAMEN UNIVERSITY PRESS
国家一级出版社
全国百佳图书出版单位

数理统计

MATHEMATICAL STATISTICS

程细玉　程　璟　编著

厦门大学出版社 国家一级出版社
XIAMEN UNIVERSITY PRESS 全国百佳图书出版单位

图书在版编目(CIP)数据

数理统计/程细玉，程璟编著. —厦门:厦门大学出版社，2016. 8
ISBN 978-7-5615-6103-4

Ⅰ. ①数…　Ⅱ. ①程…　②程…　Ⅲ. ①数理统计　Ⅳ. ①O212

中国版本图书馆 CIP 数据核字(2016)第 153798 号

出版人　蒋东明
责任编辑　陈进才
装帧设计　李嘉彬
责任印制　许克华

出版发行　厦门大学出版社
社　　址　厦门市软件园二期望海路 39 号
邮政编码　361008
总 编 办　0592-2182177　0592-2181406(传真)
营销中心　0592-2184458　0592-2181365
网　　址　http://www.xmupress.com
邮　　箱　xmupress@126.com
印　　刷　厦门市万美兴印刷设计有限公司

开本　720mm×1000mm　1/16
印张　13.5
字数　278 千字
版次　2016 年 8 月第 1 版
印次　2016 年 8 月第 1 次印刷
定价　36.00 元

本书如有印装质量问题请直接寄承印厂调换

厦门大学出版社
微信二维码

厦门大学出版社
微博二维码

内容简介

本书以通俗易懂的语言和严密的逻辑推理，介绍了数理统计的基本概念和统计推断的基本方法，力图兼顾应用性及理论性。全书内容包括相关的概率论理论基础，总体、统计量及抽样分布，点估计，区间估计，参数假设检验，分布假设检验以及非参数统计，每章配有相应的思考练习题，最后附上一些常用统计表以备查找使用。

本书适合经济管理类硕士生、博士生作为经济计量学、时间序列分析等的基础准备课程教材及参考书，同时可以作为本科应用数学专业及统计学专业的教科书，也可以作为其他相关专业的本科生或研究生教材。

前　言

在将近30年的教学中，笔者深感数理统计是个应用相当广泛的学科，它涵盖了物理学、生物学、医学、经济学、管理学等自然科学与社会科学以及交叉学科的许多方面。由于每个学科对统计学应用的要求不尽相同，故读者在选择适合自己专业的数理统计教学内容学习时会有较大差异。数理统计与一般应用学科的统计学（如体育统计学、卫生统计学、心理统计学、经济统计学、教育统计学）存在着本质的不同。前者是一切统计学的理论基础，它与概率论匹配形成了研究随机现象的逻辑严谨的数学方法学科，具有普遍性的指导意义；后者则是在具体领域针对特定的数据形成自己独特统计推断的应用学科，它往往只是在本专业有着鲜明的应用特色，并不具备普遍性。

笔者从事数量经济学和管理科学工程方面的教学研究已将近30年，主要教学内容涉及可靠性统计、经济计量学、管理统计学、时间序列分析、概率论与数理统计、随机过程，涵盖了专科、本科、硕士、博士层次，专业则主要集中在经济、金融、管理方面。在长期的教学过程中，笔者发现在经济管理类的硕士、博士层面的学习中，学生对定量分析尤其是涉及数理统计方面的定量分析常常难得要领，故我们力图深入浅出地勾勒出数理统计中统计推断的基本思路，使学生能够达到举一反三、触类旁通的效果。

第1章为概率论中一些与数理统计联系特别密切的相关知识，主要内容为数字特征、大数定律及中心极限定理，为后面的矩法估计及大样本理论做铺垫。限于篇幅，并没有把其他的概率论内容包含进来。

第3章的点估计部分中的一致最小方差无偏估计（UMVUE）虽然是数理统计中的重要内容，但在经济模型的参数估计中很难得到，且对于经济管理类的读者属于较难的部分，读者可以根据专业特点自行取舍。

第7章的密度函数估计部分在大部分的数理统计著作中并不是经常出现的，考虑到经济建模中非参数方法是一个重要的部分，本书在非参数统计初步中增加了这部

此为试读，需要完整PDF请访问：www.ertongbook.com

分内容。

本书的一大特点是对假设检验进行了深入讨论，分别就参数假设检验、两类错误、分布假设检验、非参数统计假设检验等问题进行细致而详尽的描述。在进行经济模型的建模分析中，其最重要的部分集中在模型的适应性检验、数据性质分类检验、模型参数显著性检验、模型经济意义检验等方面。毫不夸张地说，假设检验贯穿着整个建模过程，对已有的分析结果进行检验或对正在进行的模型进行各种检验是个不可或缺的过程。在经济计量学的学习过程中，就有着许多著名的检验方法（如D-W检验、Wold检验、Granger 检验、ADF检验、Johansen检验、LM检验、Hausmen检验等），本书针对这个特点对假设检验的思想方法进行了系统的阐述，希冀能够对读者有所启迪。此外，由于受篇幅及专业限制，本书对最佳检验的讨论未尽深入，只是通过第二类错误的维度略为涉及，在许多的重要结论中也没有给出仔细而严格的证明（如中心极限定理、大样本情形等）。想要穷究原理的读者，可以参考一些数学专业统计学相关文献资料。

至于统计中有着重要地位的回归分析，由于在经济计量学中有更多深入的讨论，故本书没有把此部分内容包括进来。在多年的教学经验中，笔者认为本书的内容适合经济管理类硕士生和博士生作为一个学期的学习，也适合本科数学专业和统计学专业数理统计部分的学习。为了方便读者，在每一章都附有一定的思考练习题并在书后附上常用的统计学专用表供查找。

本书的出版得益于华侨大学人文社科基金、华侨大学研究生院教育教学改革项目（16YJG06）、福建省教育厅社会科学研究项目（JA07009S）的资助，谨在此表示感谢。

在撰写过程中，虽力图利用浅白的语言来描述统计学的精髓，本着“知其然，知一些所以然”的思想来给出统计推断方法，但鉴于水平有限，漏缺错误之处难免，欢迎各界专家学者、师生提出批评意见和宝贵建议。

编著者

2016年6月

目　录

第1章 相关的概率论理论基础

§1.1 数字特征

虽然随机变量的分布函数能够完整地描述随机变量的统计特征，但在实际的应用分析中，常出现以下情况，如①人们并不需要知道精确的分布函数形式，只需了解随机变量的某些特征；②随机变量的特性决定了它们无法直接比较大小，只能间接地以某些特征进行比较。例如，以X代表任取一个成年美国男子的身高，Y代表任取一个成年中国男子的身高，两个具体的人的身高直接进行比较并不具备普遍的统计意义。又如，以X 表示某证券组合的收益率，Y表示另一证券组合的收益率，虽然直接的收益率可以比较，但证券组合的规模及品种甚至每个个体的收益率的差异仍然使收益率的直接比较缺乏统计指导意义。总之，随机变量的主要特性是随机性，即不确定性，且由于随机变量的不确定性的研究主要依赖于随机变量的分布函数，然而在实证研究中，分布函数往往无法精确地确定，故常采用随机变量的某些特征来对其进行刻画。随机变量的这些特征虽然不能充分揭示随机变量的内在特征，但能够在应用上起到重要的作用。这些特征由于往往以数的形式出现，故采用数字特征这个概念，一般指的是数学期望、方差、均方差、矩、相关系数等。

定义 1.1.1 设随机变量X的分布函数为$F(x)$，如果X为连续型随机变量，且密度函数为$f(x)$，则称

$$\int_{-\infty}^{+\infty} xf(x)\,\mathrm{d}x$$

为X的数学期望，记为$E(X)$。如果X为离散型随机变量，其分布律为

$$P\{X=a_i\}=p_i,\quad i=1,\ 2,\ \cdots,$$

则称

$$\sum_{i=1}^{+\infty} a_i p_i$$

为X的数学期望，记为$E(X)$。$E(X)$也可直接称为期望或均值。

例 1.1.1 $X \sim c[-\sqrt{3}, \sqrt{3}]$（均匀分布），求$E(X)$。

解：依定义

$$E(X)=\int_{-\sqrt{3}}^{\sqrt{3}} x \cdot \frac{1}{2\sqrt{3}}\,\mathrm{d}x=0。$$

例 1.1.2 $X \sim f(x)=\begin{cases} \frac{1}{2}\mathrm{e}^{-\frac{1}{2}x} & x \geqslant 0 \\ 0 & \text{其他} \end{cases}$，求$E(X)$。

解：

$$\begin{aligned} E(X) &= \int_{0}^{+\infty} x \cdot \frac{1}{2}\mathrm{e}^{-\frac{1}{2}x}\,\mathrm{d}x \\ &= \left(-x\mathrm{e}^{-\frac{1}{2}x}\right)\Big|_{0}^{+\infty} - \int_{0}^{+\infty} 1 \cdot (-\mathrm{e}^{-\frac{1}{2}x})\,\mathrm{d}x = 2。\end{aligned}$$

例 1.1.3 X的分布律为

$$P(X=k)=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}，\quad k=0，1，2，\cdots，$$

λ为未知参数，求$E(X)$。

解：

$$\begin{aligned} E(X) &= \sum_{k=0}^{+\infty} k \cdot \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda} = \sum_{k=1}^{+\infty} k \cdot \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda} \\ &= \lambda \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!}\mathrm{e}^{-\lambda} = \lambda \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda} = \lambda。\end{aligned}$$

小结：从定义及例子的计算易知，数学期望是随机变量的“平均”。

定义 1.1.2 若随机变量X的密度函数为$f(x)$，则称

$$\int_{-\infty}^{+\infty} x^k f(x)\,\mathrm{d}x$$

为随机变量X的k阶矩，记为$E(X^k)$。同理可定义离散情况下的随机变量X的k阶矩为

$$E(X^k)=\sum_{i=1}^{+\infty} a_i^k p_i。$$

显然，数学期望也是随机变量的一阶矩。

此为试读，需要完整PDF请访问：www.ertongbook.com

定义 1.1.3 设$g(x)$是实值函数，X为随机变量，其密度为$f(x)$。令$Y = g(X)$，则随机变量Y的数学期望$E(Y)$定义为

$$E(Y) = \int_{-\infty}^{+\infty} g(x)f(x)\,\mathrm{d}x,$$

称为随机变量函数的数学期望。同理可定义离散情况下随机变量X的函数$g(X)$的数学期望为

$$\sum_{i=1}^{+\infty} g(a_i)p_i。$$

定义 1.1.4 若随机变量X的分布函数为$F(X)$，当X为连续型随机变量时，其密度函数为$f(x)$，则称

$$\int_{-\infty}^{\infty} [x - E(X)]^2 f(x)\,\mathrm{d}x$$

为X的方差，记为$D(X)$或$Var(X)$。若X为离散型随机变量，分布律为

$$P\{X = a_i\} = p_i,\quad i = 1,\ 2,\ \cdots,$$

则称

$$\sum_{i=1}^{+\infty} [a_i - E(X)]^2 p_i$$

为X的方差，同样记为$D(X)$或$Var(X)$。

方差可理解为随机变量与均值偏差的平均，如果均值体现了随机变量的“集中”趋势，则方差体现了随机变量的“离散”趋势。根据定义1.1.3，方差$D(X)$其实就是$E[X - E(X)]^2$，且有

$$\begin{aligned} D(X) &= E\left\{X^2 - 2XE(X) + [E(X)]^2\right\} \\ &= E(X^2) - 2[E(X)]^2 + [E(X)]^2 = E(X)^2 - [E(X)]^2。\end{aligned}$$

根据定义，对应例1.1.1的X的二阶矩为

$$E(X^2) = \int_{-\sqrt{3}}^{\sqrt{3}} x^2 \cdot \frac{1}{2\sqrt{3}}\,\mathrm{d}x = \frac{1}{2\sqrt{3}} \cdot \frac{1}{3}x^3\Big|_{-\sqrt{3}}^{\sqrt{3}} = \frac{1}{6\sqrt{3}} \times 3\sqrt{3} \times 2 = 1,$$

故$D(X) = E(X^2) - [E(X)]^2 = 1$。

对应例1.1.2的X的二阶矩为

$$\begin{aligned} E(X^2) &= \int_0^{+\infty} x^2 \cdot \frac{1}{2}\mathrm{e}^{-\frac{1}{2}x}\,\mathrm{d}x = \left(-x^2\mathrm{e}^{-\frac{1}{2}x}\right)\Big|_0^{+\infty} + \int_0^{+\infty} 2x \cdot \mathrm{e}^{-\frac{1}{2}x}\,\mathrm{d}x \\ &= 4\int_0^{+\infty} x \cdot \left(\frac{1}{2}\mathrm{e}^{-\frac{1}{2}x}\right)\mathrm{d}x = 4 \times 2 = 8,\end{aligned}$$

故$D(X)=8-2\times 2=4$。

对应例1.1.3的X的二阶矩为

$$\begin{aligned}E(X^2)&=\sum_{k=0}^{+\infty}k^2\cdot\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}=\sum_{k=1}^{+\infty}k^2\cdot\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}\\&=\sum_{k=1}^{+\infty}k(k-1)\cdot\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}+\sum_{k=1}^{+\infty}k\cdot\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}=\lambda+\sum_{k=2}^{+\infty}k(k-1)\cdot\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}\\&=\lambda+\sum_{k=2}^{+\infty}\lambda^2\cdot\frac{\lambda^{k-2}}{(k-2)!}\mathrm{e}^{-\lambda}=\lambda+\lambda^2\sum_{k=2}^{+\infty}\frac{\lambda^{k-2}}{(k-2)!}\mathrm{e}^{-\lambda}\\&=\lambda+\lambda^2\sum_{k=0}^{+\infty}\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}=\lambda+\lambda^2,\end{aligned}$$

故$D(X)=\lambda+\lambda^2-\lambda^2=\lambda$。

定义 1.1.5 随机变量X的密度函数为$f(x)$，则称

$$\int_{-\infty}^{+\infty}[x-E(X)]^k f(x)\,\mathrm{d}x$$

为X的k阶中心矩，记为$E[X-E(X)]^k$。同理可定义离散情形随机变量的k阶中心矩

$$E[X-E(X)]^k=\sum_{k=1}^{+\infty}[a_i-E(X)]^k p_i。$$

易见，方差为随机变量的二阶中心矩。

例 1.1.4 分别计算标准正态分布的一、二、三、四阶中心矩。

解：设随机变量X服从标准正态分布，即

$$X\sim f(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}x^2},\quad -\infty<x<+\infty。$$

显然，由于分布的对称性有

$$E(X)=\int_{-\infty}^{+\infty}xf(x)\,\mathrm{d}x=E(X^3)=\int_{-\infty}^{+\infty}x^3f(x)\,\mathrm{d}x=0。$$

又

$$\begin{aligned}E(X^2)&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}x^2\mathrm{e}^{-\frac{1}{2}x^2}\,\mathrm{d}x=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}(-x)(-x\mathrm{e}^{-\frac{1}{2}x^2})\,\mathrm{d}x\\&=\frac{1}{\sqrt{2\pi}}(-x)\mathrm{e}^{-\frac{1}{2}x^2}\Big|_{-\infty}^{+\infty}+\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{1}{2}x^2}\,\mathrm{d}x=1,\end{aligned}$$

$$
\begin{aligned}
E(X^4) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} x^4 \mathrm{e}^{-\frac{1}{2}x^2}\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}(-x^3)(-x\mathrm{e}^{-\frac{1}{2}x^2})\,\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}}(-x^3)\mathrm{e}^{-\frac{1}{2}x^2}\Big|_{-\infty}^{+\infty} + \frac{3}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} x^2\mathrm{e}^{-\frac{1}{2}x^2}\,\mathrm{d}x = 3,
\end{aligned}
$$

且$E(X)=0$，故标准正态的一、二、三、四阶中心矩分别为0，1，0，3。

应该特别注意，随机变量的矩并不是一定都存在的。

例 1.1.5 随机变量

$$
X \sim f(x) = \frac{2}{\pi}\cdot\frac{1}{1+x^2},\quad 0<x<+\infty。
$$

易知$f(x)>0$，且

$$
\int_0^{+\infty} f(x)\,\mathrm{d}x = \frac{2}{\pi}\int_0^{+\infty}\frac{1}{1+x^2}\,\mathrm{d}x = 1,
$$

故$f(x)$为随机变量的密度函数。但

$$
\begin{aligned}
E(X) &= \int_0^{+\infty} x\cdot\frac{2}{\pi}\cdot\frac{1}{1+x^2}\,\mathrm{d}x = \frac{1}{\pi}\int_0^{+\infty}\frac{\mathrm{d}(1+x^2)}{1+x^2} \\
&= \lim_{A\to+\infty}\frac{1}{\pi}\int_0^{A}\frac{\mathrm{d}(1+x^2)}{1+x^2} = \lim_{A\to+\infty}\frac{1}{\pi}\ln(1+x^2)\Big|_0^{A} \to \infty,
\end{aligned}
$$

故均值不存在，从而$D(X)$也不存在。

例 1.1.6 随机变量

$$
X \sim f(x) = A(1+\frac{x^2}{2})^{-\frac{3}{2}},\quad A = \frac{\Gamma(\frac{3}{2})}{\sqrt{2\pi}} = \frac{\sqrt{2}}{4},\quad -\infty<x<+\infty。
$$

易知$f(x)>0$，且$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x=1$，

$$
E(X) = \int_{-\infty}^{+\infty} xf(x)\,\mathrm{d}x = A\int_{-\infty}^{+\infty}\frac{x}{(1+\frac{x^2}{2})^{3/2}}\,\mathrm{d}x = 0,
$$

但

$$
\begin{aligned}
E(X^2) \quad &= \quad A\int_{-\infty}^{+\infty} x^2 f(x)\,\mathrm{d}x = 2A\int_0^{\infty}\frac{x^2}{(1+x^2/2)^{3/2}}\,\mathrm{d}x \\
&\xlongequal{\text{令}x=\sqrt{2}y} \quad 4\sqrt{2}A\int_0^{\infty}\frac{y^2}{(1+y^2)^{3/2}}\,\mathrm{d}y,
\end{aligned}
$$

又

$$\int_0^\infty \frac{y^2}{(1+y^2)^{3/2}}\,\mathrm{d}y = \int_0^\infty \frac{1}{(1+y^2)^{1/2}}\,\mathrm{d}y - \int_0^\infty \frac{1}{(1+y^2)^{3/2}}\,\mathrm{d}y,$$

$$\int_0^\infty \frac{1}{(1+y^2)^{3/2}}\,\mathrm{d}y \xlongequal{\text{令}y=\tan z} \int_0^{\pi/2} \frac{1}{\sec^3 z}\sec^2 z\,\mathrm{d}z = \int_0^{\pi/2}\cos z\,\mathrm{d}z = 1,$$

但

$$\begin{aligned}\int_0^\infty \frac{1}{(1+y^2)^{1/2}}\,\mathrm{d}y &= \int_0^{\pi/2}\frac{1}{\sec z}\sec^2 z\,\mathrm{d}z = \int_0^{\pi/2}\frac{1}{\cos z}\,\mathrm{d}z \\ &= \ln\tan(\frac{z}{2}+\frac{\pi}{4})\Big|_0^{\pi/2} \to +\infty,\end{aligned}$$

故X的二阶矩不存在。

对于二维随机变量$(X，Y)$，可定义相应的数字特征。

定义 1.1.6 给定二维随机变量$(X，Y)$，$g(x，y)$为二元实值函数，定义随机变量函数$g(X，Y)$的数学期望为

$$E[g(X，Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x，y)f(x，y)\,\mathrm{d}x\,\mathrm{d}y,$$

其中$f(x，y)$为$(X，Y)$的联合密度函数。此时，

$$E(X) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf(x，y)\,\mathrm{d}x\,\mathrm{d}y。$$

若$(X，Y)$的联合分布律为

$$P(X=a_i，Y=b_j) = P_{ij}，\quad i=1，2，\cdots，\quad j=1，2，\cdots，$$

则

$$E[g(X，Y)] = \sum_{i=1}^{+\infty}\sum_{j=1}^{+\infty} g(a_i，b_j)P_{ij}。$$

定义 1.1.7 随机变量X与随机变量Y的协方差定义为

$$cov(X，Y) = E\left\{[X-E(X)][Y-E(Y)]\right\},$$

相关系数定义为

$$\rho_{XY} = \frac{cov(X，Y)}{\sqrt{D(X)}\sqrt{D(Y)}}。$$

由定义可知:

$$cov(X, Y) = E(XY) - (EX)(EY),$$

当$cov(X, Y) = 0$，即

$$E(XY) = (EX)(EY),$$

称X与Y不相关，此时相关系数也为0。根据定义可知$|\rho_{XY}| \leqslant 1$。事实上，不妨设$E(X) = E(Y) = 0$，则$cov(X, Y) = E(XY)$。考察参数函数

$$Z = E(tX + Y)^2 = t^2E(X^2) + 2tE(XY) + E(Y^2)。$$

由于Z恒非负，故对应关于t的一元二次方程的判别式非正，即

$$4[E(XY)]^2 - 4E(X^2)E(Y^2) \leqslant 0,$$

又

$$D(X) = E(X^2), D(Y) = E(Y^2),$$

有

$$\frac{E(XY)}{D(X)D(Y)} \leqslant 1,$$

故有$|\rho_{XY}| \leqslant 1$。相关系数在统计学及经济计量学上有着很重要的实证意义。

对于多维随机变量$(X_1, X_2, \cdots, X_n)^{\mathrm{T}}$，同样可以定义各种数字特征。但过于繁杂，只定义在实际中比较常用的协方差矩阵。

定义 1.1.8 对于多维随机变量$(X_1, X_2, \cdots, X_n)^{\mathrm{T}}$，称

$$A = \begin{pmatrix} cov(X_1, X_1) & cov(X_1, X_2) & \cdots & cov(X_1, X_n) \\ cov(X_2, X_1) & cov(X_2, X_2) & \cdots & cov(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(X_n, X_1) & cov(X_n, X_2) & \cdots & cov(X_n, X_n) \end{pmatrix}$$

为协方差矩阵。

根据定义，则A为n阶对称方阵，其主对角线元素即为每个随机变量的方差。协方差矩阵在古典线性回归及多元时间序列分析中均有重要的作用。

§1.2 大数定律

在概率论中，我们将概率定义为“频率的极限”，一个事件发生的概率的度量是反复大量试验呈现的某种“稳定”。例如，掷一枚均匀硬币10000次，观察其正、反面出现的次数，如若继续反复试验，可以发现正、反面出现的次数非常接近。又如先观测10万个成年男子的平均身高，再继续观测1万个成年男子的身高，则11万个成年男子的平均身高与10万个成年男子的平均身高差距很小，这种“稳定”就是大数定律的重要背景。

定理 1.2.1 (切比雪夫大数定律) 随机变量X_1，X_2，$\cdots$，X_n相互独立，且数学期望和方差分别相同，即$E(X_i)=\mu$，$D(X_i)=\sigma^2$，$i=1$，2，$\cdots$，n。令

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,$$

则任意的$\varepsilon>0$，有

$$\lim_{n\to\infty}P\{|\overline{X}-\mu|>\varepsilon\}=0$$

或

$$\lim_{n\to\infty}P\{|\overline{X}-\mu|<\varepsilon\}=1。$$

证明：根据切比雪夫不等式，有

$$P\{|\overline{X}-\mu|>\varepsilon\}\leqslant\frac{D(\overline{X})}{\varepsilon^2}=\frac{\frac{1}{n^2}n\sigma^2}{\varepsilon^2}=\frac{\sigma^2}{n\varepsilon^2},$$

故

$$\lim_{n\to\infty}P\{|\overline{X}-\mu|<\varepsilon\}\leqslant\lim_{n\to\infty}\frac{\sigma^2}{n\varepsilon^2}=0,$$

即

$$\lim_{n\to\infty}P\{|\overline{X}-\mu|>\varepsilon\}=0。$$

定理1.2.1虽然不复杂，但蕴含着很多意义。首先$\{|\overline{X}-\mu|>\varepsilon\}$是个随机事件，当$n$较大时，它发生的概率几乎不可能。而$\varepsilon$是个任意正数，如果把$\varepsilon$取得非常小，则该定理表示事件$\overline{X}$与$\mu$ 距离超过ε几乎不可能。换个角度说，$\overline{X}$与μ应该是非常接近的，也就是说随机变量之间的平均与数学期望非常接近。而随机变量的均值为常数（数字特征），故意味着当n较大时，随机变量的平均稳定于一个常数。但定理的条件要求所有随机变量均值相同，方差也相同，又要相互独立，条件好像过于严苛。这是初学者理解的难点，此问题在数理统计里可得到较好的阐释。

此为试读，需要完整PDF请访问：www.ertongbook.com