工科研究生教材

应用数理统计

吴群英 林 亮
编著

天津大学出版社
TIANJIN UNIVERSITY PRESS

应用数理统计

（工科研究生教材）

吴群英　林　亮　编著

天津大学出版社
Tianjin University Press

内容简介

本书是根据教育部新颁布的“工学硕士研究生应用数理统计课程教学基本要求”编写的。主要内容有：概率论补充知识，数理统计的基本概念，参数估计，假设检验，回归分析，方差分析，正交试验设计法。本书的特点是注重阐明统计思想、统计方法以及统计的实际应用；并通过选材、问题的引入、内容的阐述、例题与习题的配置等环节体现上述特点。

本书可作为高等学校非数学专业的硕士研究生的数理统计课程教材，也可供高年级大学生及从事概率统计研究的科研工作者阅读和参考。

图书在版编目(CIP)数据

应用数理统计/吴群英，林亮编著.—天津：天津大学出版社，2004.9

ISBN 7-5618-2039-9

Ⅰ.应… Ⅱ.①吴… ②林… Ⅲ.数理统计-研究生-教材 Ⅳ.0212

中国版本图书馆 CIP 数据核字(2004)第 095806 号

出版发行 天津大学出版社
出 版 人 杨风和
地　　址 天津市卫津路 92 号天津大学内(邮编：300072)
网　　址 www.tjup.com
电　　话 发行部：022-27403647　邮购部：022-27402742
印　　刷 昌黎太阳红彩色印刷有限责任公司
经　　销 全国各地新华书店
开　　本 185mm×260mm
印　　张 12
字　　数 301 千
版　　次 2004 年 9 月第 1 版
印　　次 2004 年 9 月第 1 次
印　　数 1- 3 000
定　　价 18.00 元

前　言

目前,数理统计的理论和方法已广泛地应用于自然科学、技术科学、社会科学和人文科学的各个方面.随着计算机的发展与普及,数理统计已成为处理信息、进行决策的重要理论和方法.在科学研究中,用数理统计方法从数据中所获得的信息和发现初步规律,往往成为重大科学发现的先导.在理论联系实际的方法中,数理统计是数学最活跃、最重要的方式之一。因此,作为科学和工程技术人员,特别是工科硕士研究生,应该具备数理统计的基本知识.本书是作者根据工科研究生的特点,并结合十多年来在我院进行该门课程教学的经验和体会的基础上经过充实、取舍而形成的.可作为工科硕士研究生应用数理统计课程的教材,也可作为高年级大学生,科学、工程技术人员的参考读物。

编写本教材的主要指导思想有两点:第一,覆盖国家教育部工科研究生数学课程指导小组制定的"工学硕士研究生应用数理统计课程教学基本要求"所规定的内容。注重思想方法的介绍,使学生逐步掌握正确的数理统计的思想方法,因此在阐述某一统计概念方法时,一般先提出问题的实际背景,并较多地采用实际问题和例题的形式,以使学生带着实际问题学习、思考。第二,注重应用性,数理统计是一门应用性很强的学科,它的应用几乎遍及各个领域,成为解决实际问题的重要工具.因此,本教材编入、充实了许多应用性较多的内容,以适应读者解决实际问题的需要.

由于作者水平所限,本教材的讲稿虽经多次使用和修改,但书中一定还存在不少缺点和错误,恳请读者指正.

作者

2004 年 3 月

作者简介:

吴群英　女，博士，教授。研究领域：概率论与数理统计；研究方向：随机过程和极限理论．近几年在《数学学报》、《数学年刊》、《Chinese Journal of Contemporary Mathematics》、《系统科学与数学》、《数学物理学报》、《数学研究与评论》、《应用概率统计》等刊物发表学术论文近40篇。在科学出版社出版专著《广义生灭过程》1部。从事高教数学教学20多年。

目　　录

第1章 概率论补充知识

1.1 概率空间

概率论是数理统计的基础,为了使它们能更好地衔接起来,本章扼要地复习概率论的基本概念、定理与公式,并补充了特征函数等工科概率论中不讲授的内容,以便读者查阅.

在工科概率论部分,已经对古典概型和几何概型这两种特殊类型定义了概率.在古典概型中,要求试验的可能结果是有限个且具有等可能性;对于几何概型,虽然试验的可能结果是无穷多个,但仍要求具有某种等可能性.然而实际问题中还有大量的随机试验,其结果并不属于这两种类型,因此,很有必要对一般的随机现象给出一个明确的概率定义.

一 样本空间与事件域

给出概率定义之前,先要明确事件的概念.我们知道,事件是样本空间 Ω 的一个子集,但一般并不把 Ω 的一切子集都作为事件,例如在几何概型中就不能把不可度量的子集作为事件.事实上只要把具有某些限制而又相当广泛的一类 Ω 的子集作为事件就够了.为此下面介绍事件域的概念.

定义 1.1.1 设 Ω 是样本空间,$\mathscr{F}$是由 Ω 的一些子集构成的集类,如果满足下列条件:

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

(2)若 $A\in\mathscr{F}$,则 $\bar{A}\in\mathscr{F}$;

(3)若 $A_i\in\mathscr{F},i=1,2,3,\cdots$,则$\bigcup\limits_{i=1}^{\infty}A_i\in\mathscr{F}$.

则称 $\mathscr{F}$为事件域,$\mathscr{F}$中的元素称为事件,Ω 称为必然事件.

一般对满足上述条件的集类称为 σ-域,所以事件域是一个 σ-域,它具有下列性质.

(1)空集$\varnothing\in\mathscr{F}$.

这是因为$\varnothing=\overline{\Omega}\in\mathscr{F}$,我们称事件$\varnothing$为不可能事件.

(2)若 $A_i\in\mathscr{F},\ i=1,2,\cdots$,则$\bigcap\limits_{i=1}^{\infty}A_i\in\mathscr{F}$.事实上,由 A.De Morgan 定理知

$$\bigcap_{i=1}^{\infty}A_i=\overline{\bigcup_{i=1}^{\infty}\bar{A}_i}\in\mathscr{F};$$

(3)若 $A_i\in\mathscr{F},\ i=1,2,\cdots$,则$\bigcup\limits_{i=1}^{n}A_i\in\mathscr{F},\ \bigcap\limits_{i=1}^{n}A_i\in\mathscr{F}$.这只要在定义 1.1.1 的(3)中取 $A_{n+1}=A_{n+2}=\cdots=\varnothing$即得$\bigcup\limits_{i=1}^{n}A_i\in\mathscr{F}$,在性质(2)中取 $A_{n+1}=A_{n+2}=\cdots=\Omega$,即得$\bigcap\limits_{i=1}^{n}A_i\in\mathscr{F}$.

(4)若 $A,B\in\mathscr{F}$,则 $A-B\in\mathscr{F}$.这是因为由性质(3)知,$A-B=A\cap\bar{B}\in\mathscr{F}$.

由 σ-域的定义,容易验证集类$\mathscr{F}_1=\{\varnothing,\Omega\}$,$\mathscr{F}_2=\{\varnothing,A,\bar{A},\Omega\}$以及$\mathscr{F}_3=\{A:A\subseteq\Omega=\{\omega_1,\omega_2\cdots,\omega_n\}\}$都是 σ-域.

二　概率的定义与性质

古典概型和几何概型虽然是两种特殊的类型，但它们的事件的概率都具有共同的性质．在古典概型中，事件的概率具有性质：①对任意事件 A 有 $0\leqslant P(A)\leqslant 1$；②$P(\Omega)=1$；③设事件 $A_1,A_2,\cdots,A_m(m\leqslant n)$互不相容，则 $P(\bigcup\limits_{i=1}^{m}A_i)=\sum\limits_{i=1}^{m}P(A_i)$．在几何概型中，事件的概率除具有上述前两个性质以外，还对可列个互不相容的事件 $A_1,A_2,\cdots$，有 $P(\bigcup\limits_{i=1}^{\infty}A_i)=\sum\limits_{i=1}^{\infty}P(A_i)$．概率的公理化定义，正是用这些性质来作为一般的概率定义．

定义 1.1.2　设 Ω 是给定的样本空间，$\mathscr{F}$是 Ω 中的一个事件域，$P(A)(A\in\mathscr{F})$是定义在 $\mathscr{F}$ 上的一个实值集函数，如果它满足条件：

(1)对任一事件 $A\in\mathscr{F}$，有 $P(A)\geqslant 0$；

(2)$P(\Omega)=1$；

(3)可列可加性　若 $A_i\in\mathscr{F}$，$i=1,2,\cdots$，且 $A_iA_j=\varnothing$，$i\neq j$，则有 $P(\bigcup\limits_{i=1}^{\infty}A_i)=\sum\limits_{i=1}^{\infty}P(A_i)$．

则称 $P(A)$是事件 A 的概率(简称为概率)．

至此，我们引进了概率论中三个基本概念：样本空间 Ω，事件域 $\mathscr{F}$和概率 P．它们是描述一个随机试验的三个基本组成部分，把三者结合起来，我们称三元有序总体$(\Omega,\mathscr{F},P)$为概率空间．在实际问题中，如何确定样本空间 Ω，如何选取事件域 $\mathscr{F}$，又如何在 $\mathscr{F}$上定义概率 P，要视具体情况而定，但在一般的理论研究中，我们总是认为它们是预先给定的．本书后面总假设所论及的集合是事件．

设$(\Omega,\mathscr{F},P)$是概率空间，概率 P 有如下性质：

(1)$P(\varnothing)=0$；

(2)有限可加性　若 $A_i\in\mathscr{F}$，$i=1,2,\cdots,n$，且 $A_iA_j=\varnothing$，$i\neq j$，则 $P(\bigcup\limits_{i=1}^{n}A_i)=\sum\limits_{i=1}^{n}P(A_i)$；

(3)设 $A\in\mathscr{F}$，则有 $P(\bar{A})=1-P(A)$；

(4)设 $A\in\mathscr{F}$，$B\in\mathscr{F}$，如果 $A\supset B$，则

$$P(A-B)=P(A)-P(B),\quad P(A)\geqslant P(B);$$

(5)加法公式　设 $A\in\mathscr{F}$，$B\in\mathscr{F}$，则

$$P(A\cup B)=P(A)+P(B)-P(AB),\qquad P(A\cup B)\leqslant P(A)+P(B).$$

这个结论可推广为：若 $A_i\in\mathscr{F}$，$i=1,2,\cdots,n$，则

$$P(\bigcup_{i=1}^{n}A_i)\leqslant\sum_{i=1}^{n}P(A_i).$$

三　条件概率与事件的独立性

1　条件概率

定义 1.1.3　设 A,B 是两个随机事件，且 $P(B)>0$，则称

$$P(A\mid B)=\frac{P(AB)}{A(B)}$$

为事件 B 发生的条件下，事件 A 发生的条件概率．

关于条件概率有如下三个重要公式.

(1)概率的乘法公式 若 $P(B)>0$,则 $P(AB)=P(B)P(A|B)$.

(2)全概率公式 若事件 $A_1,A_2,\cdots,A_n$ 两两互斥,且 $\bigcup_{i=1}^{n}A_i=\Omega$, $P(A_i)>0$, $i=1,2,\cdots,n$,则对任一事件 B 有

$$P(B)=\sum_{i=1}^{n}P(A_i)P(B|A_i).$$

(3)T. Bayes 公式 在全概率公式的条件下,若 $P(B)>0$,则

$$P(A_i|B)=\frac{P(A_i)P(B|A_i)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)},\quad i=1,2,\cdots,n.$$

2 事件的独立性

定义 1.1.4 对两个事件 A,B,若有

$$P(AB)=P(A)P(B),$$

则称 A,B 相互独立(简称独立).

定义 1.1.5 设 $A_1,A_2,\cdots,A_n$ 是 n 个事件,若对任意的 $k(2\leqslant k\leqslant n)$和任意一组 $1\leqslant i_1<i_2<\cdots<i_k\leqslant n$,都有

$$P(A_{i_1}A_{i_2}\cdots A_{i_k})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k})$$

成立,则称 n 个事件 $A_1,A_2,\cdots,A_n$ 独立.

对事件序列 $A_1,A_2,\cdots,A_n,\cdots$,若它们之中的任意有限个事件独立,则称事件序列 $A_1,A_2,\cdots,A_n,\cdots$独立.

事件独立的性质 若 $A_1,A_2,\cdots,A_n$ 独立,则

(1)$A'_1,A'_2,\cdots,A'_n$ 独立,其中 $A'_k=A_k$ 或$\overline{A_k}$;

(2)将事件 $A_1,A_2,\cdots,A_n$ 分成k 组(不重不漏),设 $B_1,B_2,\cdots,B_k$ 分别由第 $1,2,\cdots,k$ 组内的 A_i 经过并、积、差、求余等运算所得,则 $B_1,B_2,\cdots,B_k$ 独立.

1.2 随机变(向)量及其分布

一 随机变量及其分布

下面引进随机变量的概念,把对事件的研究转化为对随机变量的研究,由于随机变量是以数量形式来描述随机现象的,因此,它对于理论研究和数学运算都带来很大方便.

设任一随机试验,Ω 是样本空间,如果对每个 $\omega\in\Omega$ 有一实数与之对应,就得到一个定义在 Ω 上的实值函数$\xi(\omega)$.我们不仅关心 $\xi(\omega)$取什么值,更关心它取相应值的概率的大小.例如希望知道集$\{\omega:\xi(\omega)<x\}$的概率,其中 x 是任一实数.以后为简便计,常将$\{\omega:\xi(\omega)<x\}$记为$\{\xi(\omega)<x\}$或$\{\xi<x\}$.因为我们只在事件上定义了概率,讨论$\{\xi(\omega)<x\}$的概率,当然要求$\{\xi(\omega)<x\}$是事件.

定义 1.2.1 设$(\Omega,\mathscr{F},P)$是一概率空间,$\xi(\omega)$是定义在 Ω 上的单值实函数,如果对任一实数 x,$\{\xi(\omega)<x\}\in\mathscr{F}$,则称 $\xi(\omega)$为随机变量,称 $F(x)=P\{\xi(\omega)<x\}(-\infty<x<+\infty)$为

随机变量的分布函数.

由分布函数定义立即得到

$$P\{a \leqslant \xi < b\} = F(b) - F(a).$$

分布函数 $F(x)$ 具有如下基本性质:

(1) $F(x)$ 是 x 的非降函数,即若 $a < b$,则 $F(a) \leqslant F(b)$;

(2) $0 \leqslant F(x) \leqslant 1$, $F(-\infty) = 0$, $F(+\infty) = 1$,其中

$$F(-\infty) = \lim_{x \to -\infty} F(x), \quad F(+\infty) = \lim_{x \to +\infty} F(x);$$

(3) $F(x)$ 在任一点左连续,即 $F(x-0) \triangleq \lim_{y \to x-0} F(y) = F(x)$.

还可以证明,满足上述三个性质的函数,必定是某个随机变量的分布函数.即上三条性质不但是分布函数 $F(x)$ 的必要条件,同时还是 $F(x)$ 成为某随机变量的分布函数的充分条件.

用分布函数还可以进一步表示事件 $\{\xi \leqslant x\}$, $\{\xi = x\}$, $\{\xi > x\}$, $\{\xi \geqslant x\}$, $\{a \leqslant \xi \leqslant b\}$, $\{a < \xi < b\}$, $\{a < \xi \leqslant b\}$ 等的概率.

1 离散型随机变量及其分布

若随机变量的所有可能取值是有限多个或可列无限多个,则称这种随机变量为离散型随机变量.

设离散型随机变量 ξ 的所有可能取值为 $x_1, x_2, \cdots, x_i, \cdots$,而 ξ 取各相应 x_i 的概率为 p_i,即

$$P\{\xi = x_i\} = p_i, \quad i = 1, 2, \cdots.$$

那么离散型随机变量 ξ 的概率分布规律通常用分布列表示成

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

$P(\xi = x_i) = p_i$ 或上表称为离散型随机变量 ξ 的分布列或分布律.

由概率的定义,可知分布列有如下性质:

(1) $p_i \geqslant 0, \quad i = 1, 2, \cdots$;

(2) $\sum_i p_i = 1$.

离散型随机变量的分布函数 $F(x)$ 可用 p_i 表示为

$$F(x) = P\{\xi < x\} = \sum_{x_i < x} P\{\xi = x_i\} = \sum_{x_i < x} p_i.$$

2 连续型随机变量及其分布

定义 1.2.2 若随机变量 ξ 的分布函数 $F(x)$ 能够表示为某个非负可积函数 $p(x)$ 在 $(-\infty, x)$ 上的积分,即

$$F(x) = \int_{-\infty}^{x} p(t)\,\mathrm{d}t,$$

则称 ξ 为连续型随机变量,称 $p(x)$ 为 ξ 的分布密度或密度函数(简称密度).

密度 $p(x)$ 具有性质:

(1) $p(x) \geqslant 0$;

(2)$\int_{-\infty}^{+\infty} p(x)\mathrm{d}x = 1$;

(3)对于 x 轴上的任意区间 S,有

$$P\{\xi \in S\} = \int_S p(x)\mathrm{d}x;$$

(4)对于 $p(x)$ 的连续点 x,有

$$p(x) = \frac{\mathrm{d}F(x)}{\mathrm{d}x} = \lim_{\Delta x \to 0} \frac{P\{x \leqslant \xi < x + \Delta x\}}{\Delta x};$$

从而也有($\Delta x > 0$ 且 Δx 很小)

$$P\{x \leqslant \xi < x + \Delta x\} \approx p(x)\Delta x.$$

为了方便我们引入一个对离散型或连续型两种情况通用的概念——概率函数 $p(x)$:若 ξ 是离散型,则 $p(x)$ 就是 ξ 的分布列;若 ξ 是连续型,则 $p(x)$ 就是 ξ 的分布密度.

3 一些常用的概率分布

离散型:

(1)二项分布 $B(n;p)$ $(0<p<1)$ 概率函数为

$$p(x) = P(\xi = x) = \mathrm{C}_n^x p^x (1-p)^{n-x}, \quad x = 0,1,2,\cdots,n;$$

特别,当 $n=1$ 时,称 $B(1;p)$ 为 0-1 分布;

(2)S.D.Poisson 分布 $P(\lambda)$ $(\lambda > 0)$ 概率函数为

$$p(x) = P\{\xi = x\} = \frac{\lambda^x}{x!}\mathrm{e}^{-\lambda}, \quad x = 0,1,2,\cdots.$$

连续型:

(3)均匀分布 $R_{[a,b]}(a<b)$ 概率函数为

$$p(x) = \begin{cases} \dfrac{1}{b-a}, & a \leqslant x \leqslant b, \\ 0, & \text{其他}; \end{cases}$$

(4)指数分布 $E(\lambda)$ $(\lambda > 0)$ 概率函数为

$$p(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x}, & x > 0, \\ 0, & x \leqslant 0; \end{cases}$$

(5)正态分布 $N(\mu, \sigma^2)$ 概率函数为

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad \sigma > 0.$$

特别,称 $N(0,1)$ 为标准正态分布,它的分布函数记为 $\Phi(x)$,它的数值可以从书末的附表 1 查到.

正态分布概率计算公式 若 $\xi \sim N(\mu, \sigma^2)$,则 $\frac{\xi - \mu}{\sigma} \sim N(0,1)$.如记 ξ 的分布函数为 $F(x)$,则有①$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$;②$P\{a \leqslant \xi < b\} = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$;③$\Phi(-x) = 1 - \Phi(x)$.

二 随机向量及其分布

有些随机试验的结果同时涉及若干个随机变量,我们不但要考察其中各个随机变量的性

质,还要研究它们之间的联系,即要研究随机向量及其分布.

定义 1.2.3 设$(\Omega,\mathscr{F},P)$是一概率空间,$\xi_1(\omega),\xi_2(\omega),\cdots,\xi_n(\omega)$是定义在这个概率空间上的 n 个随机变量,称 $\boldsymbol{\xi}(\omega)=(\xi_1(\omega),\xi_2(\omega),\cdots,\xi_n(\omega))$为 n 维随机向量.

对于 n 个实数$x_1,x_2,\cdots,x_n$ 由于$\{\xi_1<x_1,\xi_2<x_2,\cdots,\xi_n<x_n\}=\bigcap\limits_{i=1}^{n}\{\xi_i<x_i\}\in\mathscr{F}$,所以它的概率是有意义的.

定义 1.2.4 称 n 元函数$F(x_1,x_2,\cdots,x_n)=P\{\xi_1<x_1,\xi_2<x_2,\cdots,\xi_n<x_n\}$为随机向量 $\boldsymbol{\xi}=(\xi_1,\xi_2,\cdots,\xi_n)$的(联合)分布函数.

二维随机向量(ξ_1,ξ_2)的分布函数 $F(x_1,x_2)=P\{\xi_1<x_1,\xi_2<x_2\}$具有如下性质:

(1)对每一个变量 x_1 或 x_2,$F(x_1,x_2)$是左连续、单调非降的函数;

(2)$F(x_1,-\infty)=F(-\infty,x_2)=0,\quad F(+\infty,+\infty)=1$;

(3)当 $a_1\leqslant b_1,a_2\leqslant b_2$ 时,有

$$P\{a_1\leqslant\xi_1<b_1,a_2\leqslant\xi_2<b_2\}=F(b_1,b_2)-F(a_1,b_2)-F(b_1,a_2)+F(a_1,a_2)\geqslant 0.$$

这些性质都可以推广到 n 元分布函数$F(x_1,x_2,\cdots,x_n)$,例如 $F(x_1,x_2,\cdots,x_n)$关于每个变量 x_i 是左连续、单调非降的函数;$F(x_1,x_2,-\infty,\cdots,x_n)=0$;$F(+\infty,\cdots,+\infty)=1$ 等.

随机向量和一元随机变量一样,常用的也有离散型和连续型两种.

(1)若随机向量 $\boldsymbol{\xi}=(\xi_1,\xi_2,\cdots,\xi_n)$只取有限个或可列个不同的向量值,则称 $\boldsymbol{\xi}$ 为离散型随机向量.

设 $\boldsymbol{\xi}$ 的所有可能值为$(x_{1i_1},x_{2i_2},x_{ni_n})$,$i_1,i_2,\cdots,i_n=1,2,\cdots$,则称概率

$$p_{i_1,i_2,\cdots,i_n}=P(\xi_1=x_{1i_1},\xi_2=x_{2i_2},\cdots,\xi_n=x_{ni_n}),i_1,i_2,\cdots,i_n=1,2,\cdots,$$

为 $\boldsymbol{\xi}$ 的分布列.它具有性质:

$$p_{i_1,i_2,\cdots,i_n}\geqslant 0,\ \sum_{i_1,\cdots,i_n}p_{i_1,i_2,\cdots,i_n}=1.$$

这时,$\boldsymbol{\xi}=(\xi_1,\xi_2,\cdots,\xi_n)$的分布函数为

$$F(x_1,x_2,\cdots,x_n)=\sum_{x_{1i_1}<x_1,\cdots,x_{ni_n}<x_n}p_{i_1,i_2,\cdots,i_n}.$$

(2)若存在非负函数 $p(x_1,x_2,\cdots,x_n)$,使得分布函数可表示为

$$F(x_1,x_2,\cdots,x_n)=\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n}p(y_1,y_2,\cdots,y_n)\mathrm{d}y_1\mathrm{d}y_2\cdots\mathrm{d}y_n,$$

则称 $\boldsymbol{\xi}$ 为连续型随机向量,函数 $p(x_1,x_2,\cdots,x_n)$称为 $\boldsymbol{\xi}$ 的分布密度或密度函数.它满足条件

$$p(x_1,x_2,\cdots,x_n)\geqslant 0,\quad \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}p(x_1,x_2,\cdots,x_n)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n=1.$$

反之,若有满足这两条性质的 n 元函数$p(x_1,x_2,\cdots,x_n)$,则它一定是某一个 n 维随机向量的分布密度.

对 $\mathbf{R}^n$ 上某一区域B,有

$$P\{(\xi_1,\xi_2,\cdots,\xi_n)\in B\}=\int\cdots\int_B p(x_1,x_2,\cdots,x_n)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n.$$

多元正态分布是多元连续型分布中的一个重要分布.

设 $\boldsymbol{B}=(b_{ij})$是 n 阶正定对称矩阵,$\boldsymbol{B}^{-1}=(r_{ij})$是它的逆矩阵,$|\boldsymbol{B}|$表示 $\boldsymbol{B}$ 的行列式,$\boldsymbol{a}=$

$(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$ 是一个实值列向量(其中T表示转置符号),以函数

$$p(x_1,x_2,\cdots,x_n)=\frac{1}{(2\pi)^{n/2}|\boldsymbol{B}|^{1/2}}\exp\left\{-\frac{1}{2}\sum_{i,j=1}^{n}r_{ij}(x_i-a_i)(x_j-a_j)\right\}$$

为密度函数的概率分布,称为 n 元正态分布,简记为 $N(\boldsymbol{a},\boldsymbol{B})$.上式的向量形式为

$$p(\boldsymbol{x})=\frac{1}{(2\pi)^{n/2}|\boldsymbol{B}|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{a})^{\mathrm{T}}\boldsymbol{B}^{-1}(\boldsymbol{x}-\boldsymbol{a})\right\}.$$

其中 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$.

当 $n=1$ 时,和以前所述的一元正态分布完全一致;

当 $n=2$ 时,$\boldsymbol{B}=\begin{bmatrix} b_{11} & b_{12} \\ b_{12} & b_{22}\end{bmatrix}$,由于 $\boldsymbol{B}$ 是正定对称矩阵,故 $b_{11}>0,b_{22}>0,b_{12}^2<b_{11}b_{22}$,记 $b_{11}=\sigma_1^2,b_{22}=\sigma_2^2$,则 $b_{12}=r\sigma_1\sigma_2$,其中 $\sigma_1>0,\sigma_2>0,|r|<1$.因此 $\boldsymbol{B}=\begin{bmatrix}\sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix}$.记 $a_1=a$,$a_2=b$,从而得二元正态分布的分布密度为

$$p(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(x-a)^2}{\sigma_1^2}-\frac{2r(x-a)(y-b)}{\sigma_1\sigma_2}+\frac{(y-b)^2}{\sigma_2^2}\right]\right\}. \tag{1.2.1}$$

三 边际分布

我们知道,若 $F(x,y)$ 是二维随机向量 (ξ,η) 的分布函数,则其边际分布函数为

$$F_1(x)=P\{\xi<x\}=F(x,+\infty),\quad F_2(y)=P\{\eta<y\}=F(+\infty,y).$$

若 (ξ,η) 为离散型随机向量,它的分布列

$$P\{\xi=x_i,\eta=y_j\}=p_{ij},i,j=1,2,\cdots,$$

其中 $p_{ij}\geqslant 0$,且 $\sum\limits_{i,j}p_{ij}=1$.

或表示为表1.2.1:

表 1.2.1

ξ \ η	y_1	y_2	$\cdots$	y_j	$\cdots$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\cdots$

记

$$P\{\xi=x_i\}=p_1(x_i),\quad i=1,2,\cdots,P\{\eta=y_j\}=p_2(y_j),\quad j=1,2,\cdots.$$

对于固有的 i,显然有

$$\sum_j p_{ij}=P\{\xi=x_i\}=p_1(x_i);$$

而对于固定的 j 有

$$\sum_i p_{ij}=P\{\eta=y_j\}=p_2(y_j).$$

这里 $p_1(x_i)$与 $p_2(y_i)$分别称为 p_{ij}的边际分布.

若(ξ,η)是连续型随机向量,分布密度为 $p(x,y)$.则边际分布函数为

$$F_1(x)=\int_{-\infty}^{x}\int_{-\infty}^{+\infty}p(u,v)\mathrm{d}v\mathrm{d}u,\quad F_2(y)=\int_{-\infty}^{y}\int_{-\infty}^{+\infty}p(u,v)\mathrm{d}u\mathrm{d}v,$$

从而得边际分布密度为

$$p_1(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y,\quad p_2(y)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x. \tag{1.2.2}$$

对于 n 维随机向量$(\xi_1,\xi_2,\cdots,\xi_n)$也有类似的情况.

例 1.2.1 已知二维随机向量(ξ,η)服从正态分布,分布密度为(1.2.1),求边际分布密度 $p_1(x)$及 $p_2(y)$.

解 由公式(1.2.2)得

$$\begin{aligned}p_1(x)&=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y\\&=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\int_{-\infty}^{+\infty}\exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(x-a)^2}{\sigma_1^2}-\frac{2r(x-a)(y-b)}{\sigma_1\sigma_2}+\frac{(y-b)^2}{\sigma_2^2}\right]\right\}\mathrm{d}y,\end{aligned}$$

令$(x-a)/\sigma_1=u,(y-b)/\sigma_2=v$,得

$$\begin{aligned}p_1(x)&=\frac{1}{2\pi\sigma_1\sqrt{1-r^2}}\int_{-\infty}^{+\infty}\exp\left\{-\frac{1}{2(1-r^2)}[u^2-2ruv+v^2]\right\}\mathrm{d}v\\&=\frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-u^2/2}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi(1-r^2)}}\exp\left\{-\frac{1}{2(1-r^2)}[r^2u^2-2ruv+v^2]\right\}\mathrm{d}v\\&=\frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-u^2/2}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi(1-r^2)}}\mathrm{e}^{-(v-ru)^2/2(1-r^2)}\mathrm{d}v\\&=\frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-u^2/2}=\frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-(x-a)^2/2\sigma_1^2}.\end{aligned}$$

由对称性知

$$p_2(y)=\frac{1}{\sqrt{2\pi}\sigma_2}\mathrm{e}^{-(y-b)^2/2\sigma_2^2}.$$

故知二维正态分布的两个边际分布都是一维正态分布,且 $\xi\sim N(a,\sigma_1^2)$,$\eta\sim N(b,\sigma_2^2)$,这是一个重要的结论.

四 条件分布

在同时考察若干个随机变量时,常要考虑条件概率分布,即已知某些随机变量取得某值的条件下,另一些随机变量取值的条件概率.在此书中就二维随机向量作讨论,多维的情况类似.

设(ξ,η)是离散型随机向量,它的联合分布为 $p_{ij}=P(\xi=x_i,\eta=y_i)$.若已知 $P\{\xi=x_i\}>0$,则在 $\xi=x_i$ 下,事件$\{\eta=y_j\}$的条件概率为

$$P\{\eta=y_j \mid \xi=x_i\}=\frac{P\{\xi=x_i,\eta=y_j\}}{P\{\xi=x_i\}}=\frac{p_{ij}}{\sum_j p_{ij}},$$

称它为随机变量 η 关于随机变量 ξ 的条件分布.

同样,当 $P\{\eta=y_j\}>0$ 时,

$$P\{\xi=x_i \mid \eta=y_j\}=\frac{P\{\xi=x_i,\eta=y_j\}}{P\{\eta=y_j\}}=\frac{p_{ij}}{\sum_i p_{ij}},$$

称它为随机变量 ξ 关于随机变量 η 的条件分布.

定义 1.2.5 设(ξ,η)为连续型随机向量,它的联合分布密度为 $p(x,y)$,且 $p_1(x)>0$,则称

$$F(y\mid x)\triangleq P\{\eta<y\mid\xi=x\}=\frac{\int_{-\infty}^{y}p(x,v)\mathrm{d}v}{p_1(x)}=\int_{-\infty}^{y}\frac{p(x,v)}{p_1(x)}\mathrm{d}v$$

为在 $\xi=x$ 下 η 的条件分布函数.

记在 $\xi=x$ 给定的条件下,η 的条件分布密度为$p(y\mid x)$,则由上式得

$$p(y\mid x)=\frac{p(x,y)}{p_1(x)}\quad(p_1(x)>0).\tag{1.2.3}$$

同理可得在给定 $\eta=y$ 的条件下,ξ 的条件分布密度为

$$p(x\mid y)=\frac{p(x,y)}{p_2(y)}\quad(p_2(y)>0).\tag{1.2.4}$$

例 1.2.2 设随机向量(ξ,η)服从正态分布 $N(\boldsymbol{a},\boldsymbol{B})$,求 $p(y\mid x)$.

解 由式(1.2.1),例 1.2.1 以及式(1.2.3),有

$$\begin{aligned}
p(y\mid x)&=\frac{p(x,y)}{p_1(x)}\\
&=\frac{1}{\sigma_2\sqrt{2\pi}\sqrt{1-r^2}}\exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(x-a)^2}{\sigma_1^2}-\frac{2r(x-a)(y-b)}{\sigma_1\sigma_2}+\frac{(y-b)^2}{\sigma_2^2}\right]+\frac{(x-a)^2}{2\sigma_1^2}\right\}\\
&=\frac{1}{\sigma_2\sqrt{2\pi}\sqrt{1-r^2}}\exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(x-a)^2r^2}{\sigma_1^2}-\frac{2r(x-a)(y-b)}{\sigma_1\sigma_2}+\frac{(y-b)^2}{\sigma_2^2}\right]\right\}\\
&=\frac{1}{\sigma_2\sqrt{2\pi}\sqrt{1-r^2}}\exp\left\{-\frac{1}{2(1-r^2)}\left(\frac{y-b}{\sigma_2}-r\frac{x-a}{\sigma_1}\right)^2\right\}\\
&=\frac{1}{\sigma_2\sqrt{2\pi}\sqrt{1-r^2}}\exp\left\{-\frac{1}{2\sigma_2^2(1-r^2)}\left[y-b-r\frac{\sigma_2}{\sigma_1}(x-a)\right]^2\right\}.
\end{aligned}$$

可见,在 $\xi=x$ 的条件下,η 的条件分布仍然是正态分布

$$N\left(b+r\frac{\sigma_2}{\sigma_1}(x-a),\sigma_2^2(1-r^2)\right).$$

同样由式(1.1.2),可求得 $p(x\mid y)$有类似的结论.

1.3 随机变量的独立性

定义 1.3.1 设 $\xi_1,\xi_2,\cdots,\xi_n$ 是定义在概率空间$(\Omega,\mathscr{F},P)$上的 n 个随机变量,若对于任意实数 $x_1,x_2,\cdots,x_n$,有