"十二五"普通高等教育规划教材

数理统计

主编　陈仲堂　赵德平　李彦平　潘东升

Mathematical Statistics

国防工業出版社
National Defense Industry Press

“十二五”普通高等教育规划教材

数理统计

主编　陈仲堂　赵德平　李彦平　潘东升
编委　（按姓氏笔画排序）
艾　瑛　刘　丹　闫红梅　孙常春
孙海义　徐启程　隋　英

国防工業出版社
·北京·

内容简介

本书是为适应21世纪的教学模式及现代科技对数理统计的需求，按照国家对工科研究生“数理统计”课程的基本要求编写的.

全书分为7章：概率论的基础知识、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、实用多元统计分析. 各章配有习题，书末附有答案. 除了介绍数理统计的经典理论外，各章还配备了欣赏与提高，对其理论与方法做适当的加深和拓广，以满足学有余力的学生进一步学习的需求. 本书注重体现现代科技的内涵，适量介绍一些近代数理统计理论的概念和方法，如P值检验法、主成分分析法、聚类分析等，附录还介绍了如何用SPSS、Excel、Mathematica等软件处理数理统计问题. 全书论述严谨、行文深入浅出、注重实用性.

本书可作为高等院校理、工、经济、管理、生物等专业的硕士研究生教材，也可作为本科生为了拓宽和加深概率论与数理统计课所学内容的参考书，还可供科技人员和自学者参考.

图书在版编目(CIP)数据

数理统计/陈仲堂等主编. —北京：国防工业出版社，2014.8

“十二五”普通高等教育规划教材

ISBN 978-7-118-09575-3

Ⅰ. ①数… Ⅱ. ①陈… Ⅲ. ①数理统计－高等学校－教材 Ⅳ. ①O212

中国版本图书馆CIP数据核字(2014)第160427号

※

国防工業出版社 出版发行

(北京市海淀区紫竹院南路23号 邮政编码100048)

北京市李史山胶印厂

新华书店经售

*

开本 787×1092 1/16 **印张** 17 **字数** 380千字

2014年8月第1版第1次印刷 **印数** 1—4000册 **定价** 36.00元

国防书店：(010)88540777 发行邮购：(010)88540776

发行传真：(010)88540755 发行业务：(010)88540717

前　　言

“数理统计”是工学、经济学、管理学、生物学等专业研究生的一门重要课程．它既是众多专业的基础，又能直接提供某些实用的数学方法，对提高学生的分析及处理不确定性现象的能力及运用概率统计方法解决实际问题起着重要的作用．随着我国研究生教育的飞速发展，研究生教育的授课模式较以往发生了较大的变化，加之现代科学技术的迅速发展，对“数理统计”这门课提出了更高的要求，以往研究生的教学模式及教学内容已不能完全适应现代科技的需求．如所选用的教材各自为主，缺乏规范；所用的教材内容陈旧，符号、公式与现在的习惯脱节；此外现代科技要求我们增加一些应用环节，并要体现了一些现代科技的内涵．为适应这新情况，我们拟编写此书．

本书是工科类研究生用“数理统计”课程的教材，其目的就是在教材中既要贯彻国家对工科研究生“数理统计”课程教学的基本要求；又要适应工科研究生的特点，侧重实用性，即能在实际问题中灵活应用数理统计知识，解决实际中出现的问题；此外还应体现近代科技的内涵，融入一些近代应用面较广的数理统计方法．

本书力求体现的特色是：

（1）针对工科研究生的特点，以问题为驱动，由直观到抽象、由特殊到一般阐述内容．

对于工科研究生，主要是掌握数理统计的基本概念、基本原理和基本方法，特别是能在实际问题中灵活应用数理统计知识．因此在阐述某一统计概念方法时，我们以问题为驱动，先提出问题的实际背景，通过解决实际问题来引领学生学习概率与统计基本内容，阐述概率与统计的基本思想．在有关材料的处理上，我们着重介绍各种基础的、常用的数理统计方法，特别讲明各种方法的背景、应用条件及数学结论的实际含义，尽量做到由易到难、由具体到抽象，由特殊到一般．

（2）兼顾基础与提高，适合因材施教．

本书共分为 7 章，每章均有“内容”“习题”“欣赏与提高”三个板块．每章的“内容”与“习题”，保证了国家对工科研究生“数理统计”课程教学基本要求的全面贯彻实施，是发展共性的素材．“欣赏与提高”是理论与方法做适当的加深和拓广，以满足学有余力或感兴趣的学生进一步学习的需求．

（3）介绍如何用数学软件 Mathematica、SPSS、Excel 等处理数理统计问题，使其实用性、应用性更强．

（4）结合建筑特点，注重理论知识在实际中的应用性，强调用身边生活常识阐述概率与统计的理论，用建筑工程的实例来引领学生对内容的学习．

（5）致力于以近代数学思想、观点和语言处理有关题材（如使公式、符号规范化并与现代习惯一致），并使其内容比传统的相应教材有较大的拓宽、充实、更新和提高，尽量体

现现代科技的内涵．如编入了现在应用性较强的实用多元统计分析一章，介绍了 P 值检验法、主成分分析法、聚类分析法等．

全书论述严谨、行文深入浅出、注重实用性．希望学生能够通过本教材的学习，获得数理统计方面比较系统的知识，了解处理非确定现象一些常用的统计方法，为学生后续课程的学习及工作打下基础．

本书由沈阳建筑大学陈仲堂、赵德平，沈阳大学的李彦平、潘东升主编，各章编写人员如下：闫红梅（第 1 章）、艾瑛（第 2 章，附表）、隋英（第 3 章）、陈仲堂（第 4 章）、李彦平（第 5 章）、孙海义（第 6 章）、徐启程（第 7 章）、孙常春（附录中 SPSS 部分）、刘丹（附录中 Excel Mathematica 部分）．全书由陈仲堂组织、构思及统纂，赵德平、潘东升参与编写组织工作．

本书是陈仲堂教授主持的沈阳建筑大学研究生创新项目“《数理统计》课程教学内容、方法和手段的改革与创新研究”（课题编号：YB2012001）的成果之一．

由于编者水平有限，加之时间仓促，疏漏之处在所难免，恳请有关专家、同行及广大读者批评指正．

陈仲堂

2014 年 5 月

目　　录

第 1 章　概率论的基础知识

概率论是数理统计的理论基础,为了能更好地学习数理统计,本章简要复习概率论的基本概念、定理与公式.

1.1　随机事件及概率

1.1.1　随机现象与随机事件

概率论与数理统计是研究随机现象统计规律性的一门学科. 所谓随机现象是指在一定条件下,可能出现这样的结果,也可能出现那样的结果,而在试验或观察之前不能预知确切的结果. 为了对随机现象的统计规律性进行研究,就需要对随机现象进行重复观察, 我们把对随机现象的观察称为随机试验,简称为试验,记为 E. 随机试验具有下列特点:

(1) 试验可以在相同的条件下重复进行.

(2) 每次试验的可能结果不止一个,并且能事先明确试验的所有可能结果.

(3) 进行一次试验之前不能确定哪一个结果会出现.

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 S,样本空间的元素,即 E 的每个结果,称为样本点, 记为 e.

试验 E 的样本空间 S 的子集称为 E 的随机事件,简称事件. 常用字母 $A,B,\cdots$等表示. 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.

一个样本点组成的单点集,称为基本事件. 样本空间 S 是它自身的子集,它包含所有的样本点,在每次试验中它总是发生的,称为必然事件. 空集$\varnothing$也作为样本空间 S 的子集,它不包含任何样本点,在每次试验中它都不发生,称为不可能事件. 事件是样本空间的一个集合, 故事件之间的关系与运算可按集合之间的关系和运算来处理.

1.1.2　事件的关系与运算

事件间的关系及运算与集合的关系及运算是一致的,为了方便,给出对照表 1－1.

表 1－1　事件间的关系及运算与集合的关系及运算对照表

记号	概率论	集合论
S	样本空间,必然事件	全集
$\varnothing$	不可能事件	空集
e	基本事件	元素
A	事件	子集
$\bar{A}=S-A$	A 的对立事件	A 的余集

（续）

记号	概率论	集合论
$A \subset B$	事件 A 发生导致 B 发生(子事件)	A 是 B 的子集
$A=B$	事件 A 与事件 B 相等	A 与 B 的相等
$A \cup B$	事件 A 与事件 B 至少有一个发生(和事件)	A 与 B 的和集
AB	事件 A 与事件 B 同时发生(积事件)	A 与 B 的交集
$A-B$	事件 A 发生而事件 B 不发生(差事件)	A 与 B 的差集
$AB=\varnothing$	事件 A 和事件 B 互不相容(或互斥)	A 与 B 没有相同的元素

并有下列运算性质：

(1) $\varnothing \subset A \subset S$; $A \subset B, B \subset C \Rightarrow A \subset C$.

(2) $A\overline{A}=\varnothing, A \cup \overline{A}=S, A-B=A-AB=A\overline{B}, \overline{\overline{A}}=A$.

(3) 交换律: $A \cup B=B \cup A, AB=BA$.

(4) 结合律: $A \cup (B \cup C)=(A \cup B) \cup C$, $A \cap (B \cap C)=(A \cap B) \cap C$.

(5) 分配律: $A \cap (B \cup C)=(A \cap B) \cup (A \cap C)$, $A \cup (B \cap C)=(A \cup B) \cap (A \cup C)$.

(6) 德摩根律: $\overline{A \cup B}=\overline{A} \cap \overline{B}, \overline{A \cap B}=\overline{A} \cup \overline{B}, \overline{\bigcup_i A_i}=\bigcap_i \overline{A_i}, \overline{\bigcap_i A_i}=\bigcup_i \overline{A_i}$.

1.1.3 概率

1. 概率的公理化定义

定义 1 设 E 是随机试验, S 是它的样本空间,对于 E 的每一个事件 A 赋于一个实数,记为 $P(A)$, 称为事件 A 的概率,如果集合函数 $P(\cdot)$ 满足下列条件:

(1) 非负性:对于每一个事件 A,有 $P(A) \geqslant 0$.

(2) 规范性:对于必然事件 S,有 $P(S)=1$.

(3) 可加性:设 $A_1, A_2, \cdots$ 是两两互不相容的事件,即对于 $i \neq j, A_iA_j=\varnothing, i,j=1,2,\cdots$,有

$$P\left(\bigcup_{i=1}^{\infty} A_i\right)=\sum_{i=1}^{\infty} P(A_i).$$

2. 概率的性质

(1) $P(\varnothing)=0$.

(2) 有限可加性:若 $A_1, A_2, \cdots, A_n$ 为有限个两两不相容的事件,则 $P\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{i=1}^{n} P(A_i)$.

(3) (逆事件的概率): $P(\overline{A})=1-P(A)$.

(4) 减法公式:若 $B \subset A$,则 $P(A-B)=P(A)-P(B)$ 且有 $P(A) \geqslant P(B)$,特别地对任意事件 A,有 $P(A) \leqslant 1$ 成立.

注: $P(A-B)=P(A\overline{B})=P(A)-P(AB)$.

(5) 加法公式:设 A, B 是任意两个事件,则 $P(A \cup B)=P(A)+P(B)-P(AB)$.

上式可以推广到多个事件的情况. 例如,设 A_1, A_2, A_3 为任意三个事件,有

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1A_2) - P(A_1A_3) - P(A_2A_3) + P(A_1A_2A_3).$$

一般地,对于任意 n 个事件 $A_1, A_2, \cdots, A_n$,有

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \leqslant i < j \leqslant n} P(A_iA_j) + \sum_{1 \leqslant i < j < k \leqslant n} P(A_iA_jA_k) - \cdots + (-1)^{n-1} P(A_1A_2\cdots A_n).$$

1.1.4 等可能概型

定义 2 设 E 为一随机试验,若它满足以下两个条件:①试验的结果只有有限个;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,也称古典概型.

定理 1 在古典概型中,设样本空间 S 有 n 个样本点,A 是 S 中事件且 A 中有 k 个样本点,则事件 A 发生的概率为

$$P(A) = \frac{A\text{ 包含的基本事件数}}{S\text{ 中基本事件的总数}} = \frac{k}{n}.$$

1.1.5 条件概率、乘法公式、全概率公式、贝叶斯公式

定义 3 设 A 与 B 是两个随机事件,若 $P(B) > 0$,则称

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

为在事件 B 发生的条件下,事件 A 发生的条件概率.

定理 2 设 A 与 B 是两个随机事件,若 $P(B) > 0$,有

$$P(AB) = P(B)P(A \mid B)$$

同理,若 $P(A) > 0$,有

$$P(AB) = P(A)P(B \mid A)$$

上述两式都称为概率乘法公式. 定理 2 称为乘法原理. 它们可以推广如下:

设 $A_1, A_2, \cdots, A_n$ 为 n 个随机事件,且 $P(A_1A_2\cdots A_{n-1}) > 0$,有

$$P(A_1A_2\cdots A_n) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1A_2)\cdots P(A_n \mid A_1\cdots A_{n-1}).$$

定义 4 设 S 为试验 E 的样本空间,$B_1, B_2, \cdots, B_n$ 为 E 的一组事件. 若

(1) $B_iB_j = \varnothing, i \neq j, i, j = 1, 2, \cdots, n.$

(2) $B_1 \cup B_2 \cup \cdots \cup B_n = S.$

则称 $B_1, B_2, \cdots, B_n$ 为样本空间 S 的一个划分.

定理 3 设试验 E 的样本空间为 S,A 为 E 的事件,$B_1, B_2, \cdots, B_n$ 是 S 的一个划分,且 $P(B_i) > 0 (i = 1, 2, \cdots, n)$,则

$$P(A) = P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + \cdots + P(A \mid B_n)P(B_n).$$

上述公式称为全概率公式.

定理 4 设试验 E 的样本空间为 S,A 为 E 的事件,$B_1, B_2, \cdots, B_n$ 是 S 的一个划分,且 $P(A) > 0, P(B_i) > 0 \quad (i = 1, 2, \cdots, n)$,则

$$P(B_i \mid A) = \frac{P(B_iA)}{P(A)} = \frac{P(A \mid B_i)P(B_i)}{\sum_{j=1}^{n} P(A \mid B_j)P(B_j)}, \quad i = 1, 2, \cdots, n.$$

上述公式称为贝叶斯公式.

注:公式中 $P(B_i)$ 和 $P(B_i|A)$ 分别称为原因的前验概率和后验概率.

1.1.6 独立性

定义 5 设 A,B 是两事件,如果满足等式

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

则称事件 A、B 相互独立,简称 A、B 独立.

定理 5 设 A,B 是两事件,且 $P(A)>0$. 若 A,B 相互独立,则 $P(B|A)=P(B)$. 反之亦然.

定理 6 若事件 A 与 B 相互独立,则下列各对事件 A 与 $\overline{B}$,$\overline{A}$ 与 B,$\overline{A}$ 与 $\overline{B}$ 也相互独立.

对于三个或更多个事件,给出下面的定义.

定义 6 设有 n 个事件 $A_1,A_2,\cdots,A_n(n\geqslant 3)$,若对其中任意两个事件 A_i 与 $A_j(1\leqslant i<j\leqslant n)$,有

$$P(A_iA_j) = P(A_i)P(A_j),$$

则称这 n 个事件是两两相互独立的.

定义 7 设有 n 个事件 $A_1,A_2,\cdots,A_n(n\geqslant 3)$,若对其中任意 k 个事件 $A_{i_1},A_{i_2},\cdots,A_{i_k}$ $(2\leqslant k\leqslant n)$,有

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

则称这 n 个事件是相互独立的.

由上述定义,可以得到以下两点推论.

(1) 若事件 $A_1,A_2,\cdots,A_n(n\geqslant 2)$ 相互独立,则其中任意 $k(2\leqslant k\leqslant n)$ 个事件也相互独立的.

(2) 若 n 个事件 $A_1,A_2,\cdots,A_n(n\geqslant 2)$ 相互独立,则将 $A_1,A_2,\cdots,A_n$ 中任意多个事件换成它们的对立事件,所得的 n 个事件仍相互独立.

例 1-1 仪器中有三个元件,它们损坏的概率都是 0.2,并且损坏与否相互独立,当一个元件损坏时,仪器发生故障的概率是 0.25,当两个元件损坏时,仪器发生故障的概率是 0.6,当三个元件损坏时,仪器发生故障的概率是 0.95,当三个元件都不损坏时,仪器不发生故障. 求仪器发生故障的概率.

解 设事件 A 表示仪器故障,B_1,B_2,B_3 分别表示有 1、2、3 个元件损坏,则

$P(B_1) = 3\times 0.2\times 0.8^2 = 0.384, P(B_2) = 3\times 0.2^2\times 0.8 = 0.096, P(B_3) = 0.2^3 = 0.008.$

已知概率 $P(A|B_i), i=1,2,3$ 分别等于 0.25,0.6,0.95,故

$$P(A) = \sum_{i=1}^{3} P(B_i)\cdot P(A|B_i) = 0.384\times 0.25 + 0.096\times 0.6 + 0.008\times 0.95 = 0.1612.$$

1.2 随机变量及其分布

1.2.1 一维随机变量及其分布

1. 随机变量

定义 1 设随机试验的样本空间为 $S=\{e\}$,$X=X(e)$ 是定义在样本空间 S 上的实值

单值函数，称 $X=X(e)$ 为随机变量.

2. 离散型随机变量及其概率分布

有些随机变量，它全部可能取到的值是有限个或可列无限多个，称为离散型随机变量.

定义 2 设离散型随机变量 X 的所有可能取值为 $x_k(k=1,2,\cdots)$，X 取各个可能值的概率 $P\{X=x_k\}$ 为 $p_k,k=1,2,\cdots$，则

$$P\{X=x_k\}=p_k,k=1,2,\cdots$$

称为离散型随机变量 X 的分布律或概率分布. 其中 p_k 满足两个条件：

(1) $p_k\geqslant 0,k=1,2,\cdots$.

(2) $\sum\limits_{k=1}^{\infty}p_k=1$.

分布律也可用表格表示为表 1-2.

表 1-2

X	x_1	x_2	$\cdots$	x_n	$\cdots$
p_i	p_1	p_2	$\cdots$	p_n	$\cdots$

3. 常见的离散型随机变量的分布

(0—1)分布或称两点分布：记为 $b(1,p)$，其分布律为

$$P\{X=k\}=p^k(1-p)^{1-k},k=0,1;0<p<1.$$

二项分布：记为 $b(n,p)$，其分布律为

$$P\{X=k\}=C_n^kp^k(1-p)^{n-k},k=0,1,2,\cdots,n;0<p<1.$$

泊松分布：记为 $\pi(\lambda)$，其分布律为

$$P\{X=k\}=\frac{\lambda^k\mathrm{e}^{-\lambda}}{k!},k=0,1,2,\cdots;\lambda>0.$$

几何分布：记为 $G(p)$，其分布律为

$$P\{X=k\}=p(1-p)^{k-1},k=1,2,\cdots;0<p<1.$$

超几何分布：记为 $H(N,M,n)$，其分布律为

$$P\{X=k\}=\frac{C_M^kC_{N-M}^{n-k}}{C_N^k},k\text{ 为整数};\max\{0,n-N+M\}\leqslant k\leqslant\min(n,M).$$

4. 随机变量的分布函数

定义 3 设 X 是一个随机变量，x 是任意实数，函数

$$F(x)=P(X\leqslant x),\qquad -\infty<x<\infty$$

称为 X 的分布函数.

分布函数的性质：

(1) $F(x)$ 是一个单调不减函数. 若 $x_1<x_2$，则 $F(x_1)\leqslant F(x_2)$.

(2) $0\leqslant F(x)\leqslant 1$，且 $F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0,F(\infty)=\lim\limits_{x\to\infty}F(x)=1$.

(3) $F(x)$ 是右连续的. 即 $\lim\limits_{x\to x_0^+}F(x)=F(x_0)$.

5. 离散型随机变量的分布函数

设离散型随机变量 X 的概率分布见表 1－2.

X 的分布函数为

$$F(x)=P(X\leqslant x)=\sum_{x_i\leqslant x}P(X=x_i)=\sum_{x_i\leqslant x}p_i.$$

6. 连续型随机变量及其概率密度

定义 4 如果对随机变量 X 的分布函数 $F(x)$，存在非负可积函数 $f(x)$，使得对于任意实数 x，有

$$F(x)=P\{X\leqslant x\}=\int_{-\infty}^{x}f(t)\mathrm{d}t,$$

则称 X 为连续型随机变量，$f(x)$ 为 X 的概率密度函数，简称概率密度.

概率密度的性质：

(1) $f(x)\geqslant 0$.

(2) $\int_{-\infty}^{\infty}f(x)\mathrm{d}x=1$.

(3) X 的取值落在任意区间 $(a,b]$ 上的概率为

$$P\{a<X\leqslant b\}=F(b)-F(a)=\int_a^b f(x)\mathrm{d}x.$$

(4) 若 $f(x)$ 在点 x 处连续，则 $F'(x)=f(x)$.

(5) 连续型随机变量 X 取任一指定值 $a(a\in R)$ 的概率为 0.

7. 常用连续型随机变量的分布

1) 均匀分布

若连续型随机变量 X 的概率密度为

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

则称 X 在区间 (a,b) 上服从均匀分布，记为 $X\sim U(a,b)$.

2) 指数分布

若随机变量 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{\theta}\mathrm{e}^{-x/\theta}, & x>0,\\ 0, & \text{其他}.\end{cases}$$

其中 $\theta>0$ 为常数，则称 X 服从参数为 θ 的指数分布.

3) 正态分布

若随机变量 X 的概率密度为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad -\infty<x<\infty.$$

其中 μ 和 $\sigma(\sigma>0)$ 都是常数，则称 X 服从参数为 μ 和 σ^2 的正态分布或高斯(Gauss)分布，记为 $X\sim N(\mu,\sigma^2)$.

特别地，正态分布当 $\mu=0,\sigma=1$ 时称为标准正态分布，此时，其概率密度函数和分布函数常用 $\varphi(x)$ 和 $\Phi(x)$ 表示：

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\qquad \Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt.$$

定理 1 设 $X\sim N(\mu,\sigma^2)$，则 $Y=\frac{X-\mu}{\sigma}\sim N(0,1)$.

标准正态分布表(附表 2)的使用：

(1) 表中给出了 $x>0$ 时 $\Phi(x)$ 的数值，当 $x<0$ 时，利用正态分布的对称性，易见有

$$\Phi(-x)=1-\Phi(x).$$

(2) 若 $X\sim N(0,1)$，则

$$P\{a<X\leqslant b\}=\Phi(b)-\Phi(a).$$

(3) 若 $X\sim N(\mu,\sigma^2)$，则 $Y=\frac{X-\mu}{\sigma}\sim N(0,1)$，故 X 的分布函数

$$F(x)=P\{X\leqslant x\}=P\left\{\frac{X-\mu}{\sigma}\leqslant\frac{x-\mu}{\sigma}\right\}=\Phi\left(\frac{x-\mu}{\sigma}\right),$$

$$P\{a<X\leqslant b\}=P\left\{\frac{a-\mu}{\sigma}<Y\leqslant\frac{b-\mu}{\sigma}\right\}=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right).$$

8. 随机变量的函数的分布

定义 5 如果存在一个函数 $g(X)$，使得随机变量 X,Y 满足

$$Y=g(X),$$

其中 $g(\cdot)$ 是已知的连续函数，则称随机变量 Y 是随机变量 X 的函数.

1) 离散型随机变量的函数的分布

设离散型随机变量 X 的分布律为

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

易见，X 的函数 $Y=g(X)$ 显然还是离散型随机变量.

$Y=g(X)$ 的分布律为

$$P(Y=y_j)=P(g(x)=y_j)=\sum P(X=x_i)g(x_i)=y_j.$$

2) 连续型随机变量的函数的分布

一般地，连续型随机变量的函数不一定是连续型随机变量，但我们主要讨论连续型随机变量的函数还是连续型随机变量的情形.

设已知 X 的分布函数 $F_X(x)$ 或概率密度函数 $f_X(x)$，则随机变量函数 $Y=g(X)$ 的分布函数可按如下方法求得：

$$F_Y(y)=P\{Y\leqslant y\}=P\{g(X)\leqslant y\}=P\{X\in C_y\}=\int_{C_y}f_X(x)dx.$$

其中 $C_y=\{x|g(x)\leqslant y\}$. 再将分布函数 $F_Y(y)$ 关于 y 求导，一般能得到 Y 的概率密度. 特

别当 $g(\cdot)$ 是严格单调函数时可由以下定理写出 Y 的概率密度.

定理 2 设随机变量 X 具有概率密度 $f_X(x), x\in(-\infty,\infty)$,又设 $y=g(x)$ 处处可导且恒有 $g'(x)>0$(或恒有 $g'(x)<0$),则 $Y=g(X)$ 是一个连续型随机变量,其概率密度为

$$f_Y(y) = \begin{cases} f[h(y)]\,|h'(y)|, & \alpha<y<\beta, \\ 0, & \text{其他}. \end{cases}$$

其中 $x=h(y)$ 是 $y=g(x)$ 的反函数,且 $\alpha=\min(g(-\infty),g(\infty))$,$\beta=\max(g(-\infty),g(\infty))$.

1.2.2 多维随机变量及其分布

1. 二维随机变量及其分布函数

定义 6 设随机试验 E 的样本空间为 $S=\{e\}$, $e\in S$ 为样本点,而

$$X=X(e), Y=Y(e)$$

是定义在 S 上的两个随机变量,称 (X,Y) 为定义在 S 上的二维随机变量或二维随机向量.

定义 7 设 (X,Y) 是二维随机变量,对任意实数 x,y,二元函数

$$F(x,y) = P\{(X\leqslant x)\cap(Y\leqslant y)\} \xlongequal{\text{记为}} P\{X\leqslant x, Y\leqslant y\}$$

称为二维随机变量 (X,Y) 的分布函数或称为随机变量 X 和 Y 的联合分布函数.

分布函数 $F(x,y)$ 的性质:

(1) $0\leqslant F(x,y)\leqslant 1$,且对任意固定的 y,$F(-\infty,y)=0$,对任意固定的 x,$F(x,-\infty)=0$,$F(-\infty,-\infty)=0$,$F(\infty,\infty)=1$.

(2) $F(x,y)$ 是变量 x 和 y 的不减函数,即对任意固定的 y,当 $x_2>x_1$,$F(x_2,y)\geqslant F(x_1,y)$;对任意固定的 x,当 $y_2>y_1$,$F(x,y_2)\geqslant F(x,y_1)$.

(3) $F(x^{+0},y)=F(x,y)$,$F(x,y^{+0})=F(x,y)$,即 $F(x,y)$ 关于 x 右连续,关于 y 也右连续.

(4) 对于任意 (x_1,y_1),(x_2,y_2),$x_1<x_2$,$y_1<y_2$,下述不等式成立:

$$F(x_2,y_2)-F(x_2,y_1)+F(x_1,y_1)-F(x_1,y_2)\geqslant 0.$$

2. 二维离散型随机变量及其分布律

定义 8 若二维随机变量 (X,Y) 全部可能取到的不相同的值是有限个或可列无穷多对,则称 (X,Y) 是离散型随机变量.

定义 9 若二维离散型随机变量 (X,Y) 所有可能取的值为 (x_i,y_j) $i,j=1,2,\cdots$,记

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

其中

$$p_{ij}\geqslant 0, \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}p_{ij}=1,$$

称 $P\{X=x_i,Y=y_j\}=p_{ij}(i,j=1,2,\cdots)$ 为二维离散型随机变量 (X,Y) 的分布律,或随机变量 X 与 Y 的联合分布律.

也可用表格来表示 X 与 Y 的联合分布律,如表 1-3 所列。

表 1-3 X 与 Y 的联合分布律

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

其分布函数为

$$F(x,y) = P\{X \leqslant x, Y \leqslant y\} = \sum_{x_i \leqslant x, y_j \leqslant y} p_{ij},$$

其中和式是对一切满足 $x_i \leqslant x, y_j \leqslant y$ 的 i,j 来求和的.

注意:(X,Y) 取值于任何区域 D 上的概率为

$$P\{(X,Y) \in D\} = \sum_{(x_i,y_j) \in D} p_{ij}.$$

3. 二维连续型随机变量及其概率密度

定义 10 设 (X,Y) 为二维随机变量,$F(x,y)$ 为其分布函数,若存在一个非负可积的二元函数 $f(x,y)$,使对任意实数 x,y,有

$$F(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)\,\mathrm{d}u\mathrm{d}v,$$

则称 (X,Y) 为连续型的二维随机变量,函数 $f(x,y)$ 称为二维随机变量 (X,Y) 的概率密度或称为 X 与 Y 的联合概率密度.

概率密度函数 $f(x,y)$ 的性质:

(1) $f(x,y) \geqslant 0$.

(2) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x\mathrm{d}y = F(\infty,\infty) = 1$.

(3) 设 G 是 xOy 平面上的区域,点 (X,Y) 落在 G 内的概率为

$$P\{(x,y) \in G\} = \iint_G f(x,y)\,\mathrm{d}x\mathrm{d}y.$$

(4) 若 $f(x,y)$ 在点 (x,y) 连续,有 $\dfrac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y)$.

4. 边缘分布

设 $F(x,y)$ 为 (X,Y) 的分布函数,关于 X 和 Y 的边缘分布函数分别记为 $F_X(x)$ 和 $F_Y(y)$,有

$$F_X(x) = P\{X \leqslant x\} = P\{X \leqslant x, Y < \infty\} = F(x,\infty),$$
$$F_Y(y) = P\{Y \leqslant y\} = P\{X < \infty, Y \leqslant y\} = F(\infty,y).$$

对于离散型随机变量,有

$$F_X(x) = F(x,\infty) = \sum_{x_i \leqslant x} \sum_{j=1}^{\infty} p_{ij}, F_Y(y) = F(\infty,y) = \sum_{y_j \leqslant y} \sum_{i=1}^{\infty} p_{ij}.$$

离散型随机变量(X,Y)的分量X和Y的分布律分别称为其边缘分布律,分别记为$P_{i\cdot}$和$P_{\cdot j}$.它与联合分布律的关系为

$$p_{i\cdot} = P\{X = x_i\} = \sum_{j=1}^{\infty} P\{X = x_i, Y = y_j\} = \sum_{j=1}^{\infty} p_{ij}, i = 1,2,\cdots,$$

$$p_{\cdot j} = P\{Y = y_j\} = \sum_{i=1}^{\infty} P\{X = x_i, Y = y_j\} = \sum_{i=1}^{\infty} p_{ij}, j = 1,2,\cdots.$$

对于连续型随机变量(X,Y),设它的概率密度为$f(x,y)$,有

$$F_X(x) = F(x,\infty) = \int_{-\infty}^{x} \left[\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y\right]\mathrm{d}x.$$

上式表明:X是连续型随机变量,且其概率密度为

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y.$$

同理,Y是连续型随机变量,且其概率密度为

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x.$$

分别称$f_X(x)$和$f_Y(y)$为(X,Y)关于X和Y的边缘概率密度.

5. 常见的二维连续型随机变量

(1) 二维均匀分布

设G是平面上的有界区域,其面积为A.若二维随机变量(X,Y)具有概率密度

$$f(x,y) = \begin{cases} \dfrac{1}{A}, & (x,y) \in G, \\ 0, & \text{其他}. \end{cases}$$

则称(X,Y)在G上服从均匀分布,常记为$(X,Y) \sim U(G)$.

(2) 二维正态分布

若二维随机变量(X,Y)具有概率密度

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]}.$$

其中$\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$均为常数,且$\sigma_1>0,\sigma_2>0,|\rho|<1$,则称$(X,Y)$服从参数为$\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$的二维正态分布,记为$(X,Y) \sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$.

注:二维正态随机变量的两个边缘分布都是一维正态分布,且都不依赖于参数ρ,亦即对给定的$\mu_1,\mu_2,\sigma_1,\sigma_2$,不同的$\rho$对应不同的二维正态分布,但它们的边缘分布都是相同的,因此仅由关于X和关于Y的边缘分布,一般来说是不能确定二维随机变量X和Y的联合分布的.

6. 条件分布

定义 11 设(X,Y)是二维离散型随机变量,对于固定的j,若$P\{Y=y_j\}>0$,则称