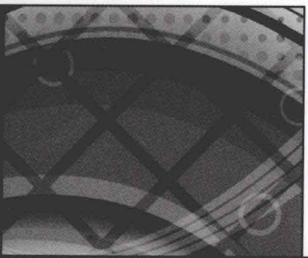
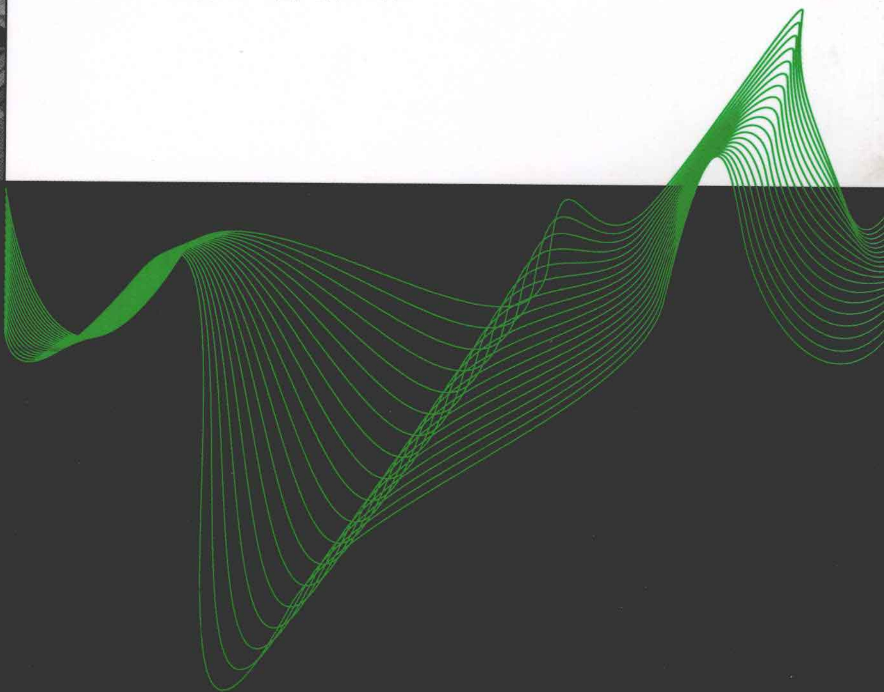
大学数学系列教材

概率论与数理统计



◆ 大连理工大学数学科学学院
◆ 冯敬海　王晓光　鲁大伟　编



高等教育出版社
HIGHER EDUCATION PRESS

大学数学系列教材

概率论与数理统计

Gailülun yu Shuli Tongji

大连理工大学数学科学学院
冯敬海　王晓光　鲁大伟　编

高等教育出版社·北京
HIGHER EDUCATION PRESS　BEIJING

内容提要

本书以概率论与数理统计教学基本要求为依据，参考国内外主流教材编写而成。内容简练明确，同时注重理论分析与实际应用。主要讲授概率论的基本概念、随机变量及其分布、二维随机变量及其分布、随机变量的数字特征、大数定律与中心极限定理、数理统计的基本概念、参数的点估计及其优良性、参数的区间估计与假设检验等内容。

本书可作为高等院校各专业（数学专业除外）概率论与数理统计课程的教材，也可供工程技术人员参考。

图书在版编目(CIP)数据

概率论与数理统计／冯敬海,王晓光,鲁大伟编. --北京:高等教育出版社, 2012. 8

ISBN 978-7-04-035542-0

Ⅰ.①概… Ⅱ.①冯…②王…③鲁… Ⅲ.①概率论-高等学校-教材②数理统计-高等学校-教材 Ⅳ.①O21

中国版本图书馆 CIP 数据核字(2012)第 181115 号

策划编辑 李 茜　责任编辑 李晓鹏 杨 帆　封面设计 李卫青　版式设计 杜微言
插图绘制 尹文军　责任校对 王 雨　责任印制 毛斯璐

出版发行	高等教育出版社	咨询电话	400－810－0598
社　　址	北京市西城区德外大街 4 号	网　　址	http://www.hep.edu.cn
邮政编码	100120		http://www.hep.com.cn
印　　刷	国防工业出版社印刷厂	网上订购	http://www.landraco.com
开　　本	787mm×960mm 1/16		http://www.landraco.com.cn
印　　张	13	版　　次	2012 年 8 月第 1 版
字　　数	230 千字	印　　次	2012 年 8 月第 1 次印刷
购书热线	010－58581118	定　　价	19.50 元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换

物 料 号　35542－00

前　　言

本书以概率论与数理统计教学基本要求为依据，参考国内外主流教材编写而成，可供高等院校各专业（数学专业除外）使用，也可供工程技术人员参考。

概率论是研究随机现象内在规律的一门科学。它已广泛地应用于经济、金融、保险、工程技术、军事和工农业等各个领域，已成为理工、经管等各类专业本科生的必修课。本书由两部分组成：第一部分包括概率论的基本概念、随机变量及其分布、二维随机变量及其分布、随机变量的数字特征、大数定律与中心极限定理等内容。第二部分包括数理统计的基本概念、参数的点估计及其优良性、参数的区间估计与假设检验等内容。

本书的主要任务是帮助读者了解概率论与数理统计的基本概念，熟悉概率论与数理统计的思维方式，学会分析与解决实际问题的基本方法。在总结了多年教学经验的基础上，本书在编写过程中主要注意突出以下特点：

1. 对学生不容易理解的概念进行重点讲解，必要时提供了多种理解方式。

2. 对正态总体参数的置信区间与假设检验问题进行了全新的处理。

3. 习题量较大，每节后配有习题，每章后配有复习题，同时每章的最后都增加了综合例题，重点讲解较难的例题。

教师如果需要讲授本书全部内容，需 48 学时，如果只讲授概率论部分（即前 5 章），只需 32 学时。

虽然编者在教学过程中积累了一些教学经验，但由于水平所限，不足之处在所难免，恳请读者能够给予批评指正。

编　者

2012 年 6 月于创新园

郑重声明

目　　录

第 1 章　概率论的基本概念

在自然界和人类的生活活动中,人们会遇到各种各样的现象,这些现象按照结果可以分成两类:必然现象与随机现象.

所谓必然现象是指在一定的条件下某特定结果一定会出现的现象.例如“在一个大气压下,纯水加热到 100℃必然会沸腾”,“以一定速度一定角度抛出一物体,在该物体落地前的任何一个时刻,其高度可以精确计算出一个唯一的结果,落点也是唯一确定的”.总之,对于必然现象,当条件满足时,其结果是可以预测的.

除了必然现象之外,还有一类现象,其结果是无法事先预知的,我们称之为随机现象.例如“掷一颗骰子出现的点数”是无法预知的,“某股票明天的收盘价格”也是无法预知的,“从 52 张(无大小王)扑克牌中任取一张,其字码与花色”是不能事先知道的.与必然现象不同,对于随机现象,尽管条件不变,但每次的结果都是不可预测的.

概率论之所以能够成为一门重要的数学学科,是因为它所研究的自然现象与其他的数学学科有本质区别.读者在此之前所学的所有数学学科都是研究必然性现象的,而概率论则是研究随机现象的,确切地说,是研究随机现象中的必然规律.

本章主要介绍概率论的基本知识以及一些简单应用,包括随机事件及事件之间的关系和运算、概率的定义、古典概型及几何概型、条件概率和全概率公式、事件的独立性与伯努利试验.这些基本概念与基本知识是学习概率论的基础,对理解整个概率论的内容是至关重要的.

§1.1　随机事件及其运算

1.1.1　随机试验(随机现象)与随机事件

研究随机现象,首先要对研究对象进行观察,所以在给出随机现象(或随机试验)的定义之前,我们先看几个简单的例子:

E_1:掷一枚硬币,观察出现正面还是反面;

E_2:掷一颗骰子,观察出现的点数;

E_3:掷两颗骰子,观察出现的点数;

E_4:掷两颗骰子,观察出现的点数之和;

E_5:任取一个灯泡,观察其寿命(单位:h);

E_6:观察某商店每天到达的顾客人数.

上述试验(或现象)具有两个共同特征:每个试验的所有可能结果都是试验之前已知的,但是在试验之前并不能预知哪个结果会出现.比如骰子的所有点数是1,2,3,4,5,6,但掷之前并不知道哪个点数会出现;每天到达商店的顾客人数可能是$\{0,1,2,\cdots\}$中的任何一个数,但不可能预知到底有几个人.

定义 1.1.1　一个现象E如果具有以下特征,我们就称该现象为一个**随机现象(或随机试验)**:

(1) 该试验可在相同条件下重复地进行;

(2) 所有可能出现的结果是已知的;

(3) 试验之前不可预知哪个结果会出现.

以$\Omega=\{\omega\}$表示随机试验E的所有可能结果组成的集合,并称之为随机试验E对应的**样本空间**.Ω的元素称为**样本点**,即样本点就是可能结果.

例 1.1.1　随机试验E_1到E_6对应的样本空间分别为:

$\Omega_1=\{$正面,反面$\}$;　　$\Omega_2=\{1,2,3,4,5,6\}$;

$\Omega_3=\{(i,j):i,j=1,2,\cdots,6\}$;　　$\Omega_4=\{2,3,\cdots,12\}$;

$\Omega_5=[0,+\infty)$;　　$\Omega_6=\{0,1,2,\cdots\}$.

由随机试验E_3与E_4对应的样本空间Ω_3与Ω_4的区别可以看出,虽然是相同的试验,但由于观察的目的不一样,对应的样本空间也是不一样的.

样本空间是一个集合,但是,一般情况下,样本空间的子集对我们来讲更加重要.比如,赌徒在用骰子进行赌博时,更加关心出现的点数是“大”还是“小”,“大”就是4或5或6,可以表示成$A=\{4,5,6\}$,“小”就是1或2或3,可以表示成$B=\{1,2,3\}$.掷一颗骰子,可能“大”也可能“小”,或者说“大”和“小”都是可能发生也可能不发生的事件.在随机现象中,这种可能发生也可能不发生的事件是我们主要研究的对象.由上面的例子可以看出$A=\{4,5,6\}$与$B=\{1,2,3\}$都是样本空间$\Omega_2=\{1,2,3,4,5,6\}$的子集,于是我们有下面的定义:

定义 1.1.2　一般地,我们将随机试验E对应的样本空间Ω的子集称为随机试验的**随机事件**,简称事件.事件一般用$A,B,C,\cdots$表示.

随机事件是可能发生也可能不发生的事件,**如果在试验中,属于某个事件A的样本点(即可能结果)出现了,我们就称该事件发生了**.比如,骰子出现的点数为2,那么“小”这个随机事件就发生了.

我们知道样本空间Ω作为一个集合,它是自己的子集,也是一个随机事件,由于它包含了所有的样本点,故每次试验它必然发生,所以我们把样本空间Ω

称为**必然事件**.另外,空集$\varnothing$作为样本空间Ω的子集,也是一个事件,只不过它不包含任何样本点,所以它在每次试验中都不可能发生,于是我们把空集$\varnothing$称为**不可能事件**.比如买彩票,如果你把所有的组合全买,那么你必然中奖;反之,如果你一注也不买,那你不可能中奖.

1.1.2 事件间的关系与运算

实际问题中遇到的随机事件往往是比较复杂的,这就需要我们将较复杂的事件"分解"成一些简单事件的组合.由前面的定义,样本空间可以看作全集,事件是其子集,因此就可以用集合之间的关系和运算来描述事件之间的关系和运算.

1. 事件的包含与相等

若事件A的每一个样本点都包含在事件B中,则称事件B包含事件A,记为$A\subset B$或$B\supset A$(图 1-1).这时,若事件A发生必导致事件B发生.若$A\subset B$且$B\subset A$,则称事件A与B相等,记为$A=B$.

2. 和事件

设A与B为两个随机事件,事件$A\cup B$称为A与B的和事件.有时$A\cup B$也记为$A+B$(图 1-2).用概率论的语言来说,$A\cup B$发生的充要条件为A与B至少有一个发生.类似地,如果$A_1,A_2,\cdots,A_n$为n个事件,那么$\bigcup\limits_{i=1}^{n}A_i$表示$A_1,A_2,\cdots,A_n$中至少有一个发生.

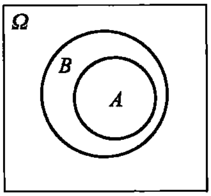


图 1-1

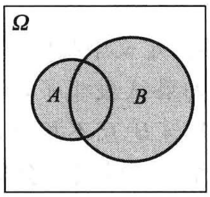


图 1-2

3. 积事件

设A与B为两个随机事件,事件$A\cap B$称为A与B的积事件.$A\cap B$经常写成AB(图 1-3).用概率论的语言来说,$A\cap B$发生的充要条件为A与B都(或同时)发生.类似地,如果$A_1,A_2,\cdots,A_n$为n个事件,那么$\bigcap\limits_{i=1}^{n}A_i$表示$A_1,A_2,\cdots,A_n$都发生.

4. 差事件

设A与B为两个随机事件,$A-B$为A与B的差事件(图 1-4).用概率论的语言来说,$A-B$发生的充要条件为A发生而B不发生.

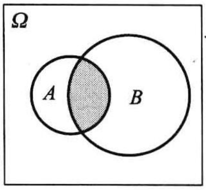


图 1-3

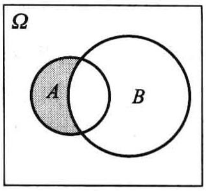


图 1-4

5. 补事件

设 A 为随机事件，我们称 $\bar{A}=\Omega-A$ 为事件 A 的补事件(图 1-5). 一般地我们称 A 与 $\bar{A}$ 互补，或互逆.

注意 $A-B=A-AB=A\bar{B}$.

6. 互不相容

若 $AB=\varnothing$，我们称 A 与 B **互不相容**，或互斥(图 1-6).

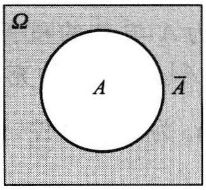


图 1-5

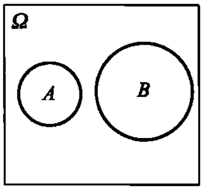


图 1-6

从互不相容的定义中我们可以看出，$AB=\varnothing$ 的意思是 A 与 B 不可能同时发生，也就是说如果 A 发生了，B 就不会发生；反之，如果 B 发生了，A 就不会发生. 注意，如果 A 与 B 互不相容，并不是说 A 与 B 没有关系，而是有很大的关系，因为其中一个事件的发生会限制另一个事件的发生.

另外，事件的运算与子集的运算一样，必须遵循一些原则：

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

(2) 结合律：$A\cup(B\cup C)=(A\cup B)\cup C$，$(A\cap B)\cap C=A\cap(B\cap C)$；

(3) 分配律：$B\cap(\bigcup\limits_{i=1}^{n}A_i)=\bigcup\limits_{i=1}^{n}(A_iB)$；$B\cup(\bigcap\limits_{i=1}^{n}A_i)=\bigcap\limits_{i=1}^{n}(A_i\cup B)$；

(4) 德摩根律：$\overline{\bigcup\limits_{i=1}^{n}A_i}=\bigcap\limits_{i=1}^{n}\overline{A_i}$；$\overline{\bigcap\limits_{i=1}^{n}A_i}=\bigcup\limits_{i=1}^{n}\overline{A_i}$.

例 1.1.2　设有三个人各购买了一注福利彩票，以 A 表示"第一个人中奖"，B 表示"第二个人中奖"，C 表示"第三个人中奖". 试用 A,B,C 表示下列事件：

(1) 至少有一个人中奖；

(2) 恰有一个人中奖；

(3) 至多有一个人中奖.

解 (1) $A\cup B\cup C$.

(2) 恰有一个人中奖是指其中有一个人中奖而另外两个人没中奖，即

$$A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C.$$

(3) 至多有一个人中奖是指没有人中奖或恰有一个人中奖，所以

$$\bar{A}\bar{B}\bar{C}\cup A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C.$$

例 1.1.3 靶子由 10 个同心圆组成，半径分别为 $r_1,r_2,\cdots,r_{10}$，且 $r_1<r_2<\cdots<r_{10}$，以事件 A_k 表示命中点在半径为 r_k 的圆内，叙述下列事件的意义：

(1) $\bigcup\limits_{k=1}^{6}A_k$；(2) $\bigcap\limits_{k=1}^{8}A_k$；(3) $\bar{A}_1A_2$.

解 (1) 命中点在半径为 r_6 的圆域内；

(2) 命中点在半径为 r_1 的圆域内；

(3) 命中点在内径为 r_1，外径为 r_2 的圆环域内.

数学家简介 德摩根(De Morgan，1806—1871)，英国数学家，1823—1827 年间入读剑桥大学三一学院．1828 年，他的老师皮科克等人推荐他任伦敦大学学院数学教授一职．1865 年，他积极筹备伦敦数学会，1866 年担任第一任会长. 德·摩根主要在分析学、代数学、数学史及逻辑学等方面作出重要的贡献. 他的工作对 19 世纪的数学具有相当的影响力.

习 题

1. 设 A,B,C 为三个事件，用 A,B,C 的运算关系表示下列各事件：

(1) B 发生，A,C 都不发生；
(2) A,B,C 都发生；
(3) A,B,C 至少有一个发生；
(4) A,B,C 都不发生；
(5) A,B,C 中不多于一个发生；
(6) A,B,C 中不多于两个发生；
(7) A,B,C 中至少有两个发生.

2. 在区间[0,1]上任取一数，记 $A=\left\{x\,\middle|\,\frac{1}{4}<x\leqslant\frac{1}{2}\right\}$，$B=\left\{x\,\middle|\,\frac{1}{3}\leqslant x<1\right\}$，求下列事件的表达式：(1)$A\cup B$；(2)$AB$；(3)$\bar{A}B$；(4)$A\cup\bar{B}$.

3. 某射手向一目标射击三次，令 A_i＝{第 i 次射击命中目标}($i=1,2,3$)，B＝{三次射击中至少命中 1 次}，C_j＝{三次射击中恰好命中 j 次}($j=0,1,2,3$). 用 A_1,A_2,A_3 表示 B 和 C_j.

4. 记录某电话机在一天内的呼叫次数，设 A_k＝{至少 k 次呼叫}，试分别叙述事件 $\bar{A}_k$，A_k-A_{k+1}，$\bigcup\limits_{k=1}^{n}A_k$，$\bigcap\limits_{k=0}^{n}A_k$ 的含义.

§ 1.2 概率的定义及其基本性质

1.2.1 频率与概率

一个随机试验有多个可能结果，但是各种结果出现的机会并不一样. 就是说，如果将试验进行很多次，有些结果出现的次数明显要多，有些则很少，它们具有统计规律性. 比如，26 个英文字母在英语的使用中出现的频率(即可能性)是不一样的，如果键盘把出现频率较高的一些字母排在便于手指碰到的地方，那么这样的键盘在打字时速度就会比较快. 为了统一地描述这种规律性，下面我们给出一个定量的刻画.

设 E 是随机试验，Ω 是它的样本空间，A 为 E 的一个事件. 将试验重复进行 n 次，其中事件 A 发生了 n_A 次，则称比值$\frac{n_A}{n}$为 A 发生的频率，记作 $f_n(A)=\frac{n_A}{n}$. "频率"的大小反映了事件 A 在试验中发生的频繁程度. 容易看出，频率具有以下性质：

(1) 非负性：$f_n(A)\geqslant 0$；

(2) 归一性：$f_n(\Omega)=1$；

(3) 有限可加性：若 $A_1,A_2,\cdots,A_k$ 为两两互不相容的事件，则

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

例如，将"抛硬币"的随机试验重复进行 n 次，记 $A=\{$正面向上$\}$，事件 A 发生的次数记作 n_A，则事件 A 发生的频率为 $f_n(A)=\frac{n_A}{n}$. 当 n 较小时，频率 $f_n(A)$在 0 与 1 之间随机波动，其幅度比较大，但随着 n 的增大，频率将呈现出稳定性. 一些著名的统计学家曾进行过大量抛掷硬币的试验，所得结果如表 1－1 所示.

表 1－1

试验者	投掷次数	出现正面的次数	出现正面的频率
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
维尼	30 000	14 994	0.499 8

从表 1-1 中可以看出,随着投掷次数的不断增加,出现正面的频率越来越接近$\frac{1}{2}$,这说明出现正面的概率为$\frac{1}{2}$.

1.2.2 概率的公理化定义

在上面,事件的概率是用频率来定义的,称为概率的统计学定义.但在实际问题中,我们不可能对每一个事件都进行大量的试验.一个事件 A 的概率 $P(A)$ 不是定义出来的,也不是凭空捏造出来的,它是现实存在的.因此,为了理论分析与实际应用的需要,从频率的三条性质出发定义概率,就得到了概率的公理化定义.

定义 设 Ω 为随机试验 E 的样本空间,如果对于任意事件 $A\subset\Omega$,有一个实数 $P(A)$与之对应,且满足:

(1) 非负性:$P(A)\geqslant 0$;

(2) 归一性:$P(\Omega)=1$;

(3) 可加性:设 $A_1,A_2,\cdots$是一列两两互不相容的事件,则有

$$P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i), \qquad (1.2.1)$$

则称 $P(A)$为事件 A 的概率.

从定义中可以看出,概率 P 是一个映射,它将 Ω 的任何一个子集(事件)A 映成一个实数 $P(A)$,但要遵循一定的规则.另外,读者应该知道,只有事件才能有概率,其他任何事物都没有概率.

上述定义是柯尔莫哥洛夫于 1933 年给出的,在此之前许多人将概率论视为伪科学而拒不接受.在柯尔莫哥洛夫给出概率的定义之后,概率论才发展成为一门科学,并应用日益广泛,渗透到各个工程领域及其他应用领域,已成为目前每个工程技术人员与理论研究工作者必不可少的工具.

由定义的三条公理,可以得到概率的一些基本性质如下:

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

证明 由于$\varnothing=\varnothing\cup\varnothing\cup\cdots$,所以,由可加性可得

$$P(\varnothing)=P(\varnothing)+P(\varnothing)+\cdots,$$

所以 $P(\varnothing)=0$.

(2) **有限可加性** 设 $A_1,A_2,\cdots,A_n$ 是 n 个两两互不相容的事件,则有

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i). \qquad (1.2.2)$$

证明 取 $A_{n+1}=A_{n+2}=\cdots=\varnothing$,由可加性可得

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

(3) 如果 $A \subset B$,则 $P(A) \leqslant P(B)$.

证明　因为 $B = A \cup B\bar{A}$,且 A 与 $B\bar{A}$ 互不相容,所以由有限可加性得

$$P(B) = P(A) + P(B\bar{A}),$$

又因为 $P(B\bar{A}) \geqslant 0$,所以 $P(A) \leqslant P(B)$.

(4) $P(A) + P(\bar{A}) = 1$.

证明　因为 A 与 $\bar{A}$ 互不相容且 $\Omega = A \cup \bar{A}$,所以

$$1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A}).$$

(5) **减法公式**

$$P(A - B) = P(A\bar{B}) = P(A) - P(AB). \tag{1.2.3}$$

证明　因为 $A = AB \cup A\bar{B}$,且 AB 与 $A\bar{B}$ 互不相容,所以

$$P(A) = P(AB \cup A\bar{B}) = P(AB) + P(A\bar{B}),$$

即 $P(A\bar{B}) = P(A) - P(AB)$.

(6) **加法公式**

$$P(A \cup B) = P(A) + P(B) - P(AB). \tag{1.2.4}$$

证明　因为 $A \cup B = B \cup A\bar{B}$,且 B 与 $A\bar{B}$ 互不相容,所以

$$P(A \cup B) = P(B \cup A\bar{B}) = P(B) + P(A\bar{B}) = P(B) + P(A) - P(AB),$$

其中加法公式可以推广到多个事件的情形,比如

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC).$$

证明留给读者.

例 1.2.1　设事件 A 与 B 的概率分别为 $\frac{1}{2}$ 与 $\frac{1}{3}$,试求下列三种情况下 $P(A-B)$ 的值:(1) $AB = \varnothing$;(2) $B \subset A$;(3) $P(AB) = \frac{1}{4}$.

解　(1) 因为 $P(AB) = 0$,所以 $P(A-B) = P(A) - P(AB) = P(A) = \frac{1}{2}$;

(2) 因为 $B \subset A$,所以 $P(A-B) = P(A) - P(AB) = P(A) - P(B) = \frac{1}{6}$;

(3) $P(A-B)=P(A)-P(AB)=\frac{1}{2}-\frac{1}{4}=\frac{1}{4}$.

例 1.2.2 已知 $P(A)=0.5, P(B)=0.4, P(A-B)=0.3$，求 $P(\bar{A}\cup\bar{B})$ 与 $P(A\cup\bar{B})$.

解 由 $P(A-B)=P(A)-P(AB)=0.3$ 及 $P(A)=0.5$ 知，$P(AB)=0.2$，所以

$$P(\bar{A}\cup\bar{B})=P(\overline{AB})=1-P(AB)=0.8,$$

$$\begin{aligned}P(A\cup\bar{B})&=P(\overline{\bar{A}B})=1-P(\bar{A}B)=1-P(B)+P(AB)\\&=1-0.4+0.2=0.8.\end{aligned}$$

例 1.2.3 设 A,B,C 为三个事件，且 $P(A)=P(B)=P(C)=\frac{1}{4}, P(AB)=P(BC)=\frac{1}{8}, P(AC)=0$，求 A,B,C 都不发生的概率.

解 A,B,C 都不发生可以表示为 $\bar{A}\bar{B}\bar{C}$，所以 $P(\bar{A}\bar{B}\bar{C})=1-P(A\cup B\cup C)$，又因为 $P(AC)=0$，所以 $P(ABC)=0$，于是

$$\begin{aligned}P(A\cup B\cup C)&=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)\\&=\frac{1}{4}+\frac{1}{4}+\frac{1}{4}-\frac{1}{8}-\frac{1}{8}=\frac{1}{2},\end{aligned}$$

所以 $P(\bar{A}\bar{B}\bar{C})=\frac{1}{2}$.

以上几个例题都是给定某些事件的概率，去求与这些事件有关的另一些事件的概率，这并不是“真正”的求概率，在下一节我们就介绍如何去求“真正的概率”.

数学家简介 柯尔莫哥洛夫(Kolmogorov,1903—1987)，20 世纪苏联最杰出的数学家，也是 20 世纪世界上最有影响的数学家之一. 他的研究遍及数学的所有领域，做出许多开创性的贡献. 母亲出身于贵族家庭，在他出生后 10 天去世. 1920 年，他高中毕业进入莫斯科大学，先学习冶金，后来转学数学，并决心以数学为终身职业. 大学三年级时就发表了论文，表现出卓越的数学才能．1935 年获得苏联首批博士学位．1931 年起，他担任莫斯科大学教授，并指导研究生. 1933 年担任莫斯科大学数学力学研究所所长，创建了概率论、数理统计、数理逻辑、概率统计方法等教研室，先后教过数学分析、常微分方程、复变函数论、概率论、数理逻辑和信息论等课程．1939 年当选为苏联科学院院士、主席团委员和数学研究所所长．1954 年担任莫斯科大学数学力学系主任．1966 年当选为苏联教

育科学院院士.柯尔莫哥洛夫热爱生活,兴趣广泛,喜欢旅行、滑雪、诗歌、美术和建筑.他十分谦虚,从不夸耀自己的成就和荣誉.他淡泊名利,不看重金钱,他把奖金捐给学校图书馆,并且不去领取高达10万美元的沃尔夫奖.他是一位具有高尚道德品质和崇高的无私奉献精神的科学巨人.

习　题

1. 设 $P(A)=x$, $P(B)=y$, $P(AB)=z$,用 x,y,z 表示下列事件的概率:$P(\bar{A}\cup\bar{B})$, $P(\bar{A}B)$, $P(\bar{A}\cup B)$, $P(A\bar{B})$.

2. 设 $P(A)=0.8$, $P(B)=0.5$,问:

(1) 在什么条件下 $P(AB)$ 取最大值,最大值是多少?

(2) 在什么条件下 $P(AB)$ 取最小值,最小值是多少?

3. 若 $A\supset B$, $A\supset C$,且 $P(A)=0.9$, $P(\bar{B}\cup\bar{C})=0.8$,求 $P(A-BC)$.

4. 设 $P(A)=0.4$, $P(B)=0.5$, $P(A\cup B)=0.7$,求 $P(A-B)$ 及 $P(\bar{A}\cup B)$.

§1.3　等可能概型(古典概型与几何概型)

等可能概型是指在这类随机试验中,每个样本点发生的可能性都相等.特别是在概率论发展的初期,最早研究的概率是诸如"掷骰子、抛硬币、抽扑克"等问题,它们都是等可能概型,也是概率论的最简单的应用之一.

1.3.1　古典概型

定义　如果一个随机试验具有以下两个特点:

(1) 样本空间中的样本点总数是有限的;

(2) 每个样本点出现的可能性相同,

则称该试验对应的概率模型为古典概型.

古典概型是一类非常常见的随机现象,比如掷骰子、抛硬币、抽扑克、抽签等,都是古典概型.

假设一个古典概型的样本空间中共有 n 个样本点 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$,由于每个样本点出现的可能性都一样,那么必有

$$P(\{\omega_i\})=\frac{1}{n}.$$

如果事件 $A\subset\Omega$ 中有 k 个样本点 $\Omega=\{\omega_{i_1},\omega_{i_2},\cdots,\omega_{i_k}\}$,则

$$P(A)=P(\{\omega_{i_1},\omega_{i_2},\cdots,\omega_{i_k}\})=\bigcup_{j=1}^{k}P(\{\omega_{i_j}\})=\frac{k}{n},$$

即在古典概型中，事件 A 的概率等于 A 中样本点数除以样本点总数.

古典概型是人类研究的最早的一类概率模型，此类模型与赌博的关系比较密切，也是概率论中最吸引人的一类模型. 利用古典概型的计算公式求概率看起来很简单，只需将样本点总数与 A 中样本点数求出即可，然而事情并非那么简单. 可以说古典概型是最难的一类模型，对于一些比较复杂的问题，人们往往不知道如何求出样本点总数与 A 中样本点数. 下面我们通过几个例子来说明古典概型求概率的方法.

例 1.3.1 设有 6 个白球和 4 个红球混合后装入袋中，从这 10 个球中任取 5 个，

(1) 在有放回的情形下，求这 5 个球中恰有 3 个白球的概率；

(2) 在不放回的情形下，求这 5 个球中恰有 3 个白球的概率；

(3) 在不放回取球下，求第 3 个球为白球的概率.

解 (1) 设事件 $A=\{$这 5 个球中恰有 3 个白球$\}$. 先求样本点总数，由于是有放回的，所以每次取球都是从 10 个球中任取一个，连续取 5 次，总的取法(即样本点总数)为

$$10\cdot10\cdot10\cdot10\cdot10=10^5.$$

下面求事件 A 中的样本数，要求这 5 个球中恰有 3 个为白球，但 5 个球中哪 3 个是白球是不能确定的，前 3 个是白球可以，后 3 个是白球也可以，中间 3 个是白球也可以，等等. 共有 $C_5^3=10$ 种不同的情况. 如果前 3 个是白球，后 2 个必是红球，这种情况下取法有 $6\cdot6\cdot6\cdot4\cdot4=6^3\cdot4^2$ 种，同理可以求出其他几种情况下的取法也是 $6^3\cdot4^2$ 种，所以 A 中的样本点数为 $C_5^3\cdot6^3\cdot4^2$，所以

$$P(A)=\frac{C_5^3\cdot6^3\cdot4^2}{10^5}=C_5^3\cdot0.6^3\cdot0.4^2.$$

(2) 设事件 $B=\{$这 5 个球中恰有 3 个白球$\}$. 还是先求样本点总数，此时是不放回地取球，从 10 个球中取出 5 个即可，故样本点总数为 C_{10}^5. 下面求样本点数，要求这 5 个球中有 3 个白球 2 个红球，显然这 3 个白球是从 6 个白球中取出的，2 个红球是从那个红球中取出的，所以 B 中样本点数为 $C_6^3\cdot C_4^2$，所以此时

$$P(B)=\frac{C_6^3\cdot C_4^2}{C_{10}^5}.$$

(3) 设事件 $C=\{$第 3 个球为白球$\}$. 先求样本点总数，此时要考虑次序(为何要考虑次序，见下文总结). 从 10 个球中取出 5 个，一个一个地取，不放回，总的取法即样本总点数为

$$10 \cdot 9 \cdot 8 \cdot 7 \cdot 6.$$

C 中样本点数：事件 C 要求第 3 个为白球，其他的是什么颜色无关紧要. 我们可以先取出一个白球“预留”给第 3 次就可以了，先取一个白球的取法为 6 种，其他 4 个球的取法为 $9 \cdot 8 \cdot 7 \cdot 6$，所以 C 中样本点数为 $6 \cdot 9 \cdot 8 \cdot 7 \cdot 6 = \frac{6}{10}\mathrm{C}_{10}^{5} \cdot 5!$，故得

$$P(C) = \frac{\frac{6}{10}\mathrm{C}_{10}^{5} \cdot 5!}{\mathrm{C}_{10}^{5} \cdot 5!} = \frac{6}{10}.$$

通过此例，读者可以看出，我们在计算三个样本点总数时采用的方法都不一样. 为什么会这样？

设有 N 个不同的球，要从中任取 n 个球($n \leqslant N$)：

第一种取法：“可重复”或者“有放回”，其含义是每次取一个球，观察后放回，再取下一个. 在这种方法下，实际上不是取了 n 个球，而是取了 n 次，两次不同的取球可能取到同一个球. 这种情况下样本点总数为 N^n.

第二种取法：“不放回”但与次序无关. 这种取法不需要考虑次序，也就是只要拿出 n 个球即可，此时样本点总数为 C_N^n.

第三种取法：“不放回”且与次序有关，其含义是每次取一个球，但不放回，连续取 n 个. 这样取球自然有一个排序，即次序. 这种情况下样本点总数为 $N(N-1)\cdots(N-n+1) = \mathrm{P}_N^n = \mathrm{C}_N^n \cdot n!$

例 1.3.1 中求三个事件 A,B,C 对应的样本总点数时，使用的模型与以上三种取法一一对应.

例 1.3.2(分房问题)　将 n 个人随机分到 N 个房间中去($n \leqslant N$)，每个人分到哪个房间是等可能的，且假设每个房间可容纳的人数没有限制，求：

(1) 某个指定的房间(比如第一个房间)恰有 m 个人的概率($m \leqslant n$)；

(2) 每两个人都不在同一个房间的概率.

解　(1) 设 $A=\{$某个指定的房间恰有 m 个人$\}$，由于每一个人都有 N 种分法，分完 n 个人才算结束，所以样本点总数为 N^n. 对于 A 中样本点数，房间是固定的，但哪 m 个人分到此房间是不确定的，也就是说哪 m 个人分到此房间都可以，那么我们就先从 n 个人中选出 m 个人来，让这 m 个人分到此房间，剩下的 $n-m$ 个人分到其他的 $N-1$ 个房间中. 所以事件 A 中的样本点数为 $\mathrm{C}_n^m \cdot (N-1)^{n-m}$，故得

$$P(A) = \frac{\mathrm{C}_n^m \cdot (N-1)^{n-m}}{N^n}.$$

(2) 设 $B=\{$每两个人都不在同一个房间$\}$. 样本点总数仍为 N^n，B 中样本点

数为 $N(N-1)\cdots(N-n+1)=\mathrm{C}_N^n\cdot n!$,故得

$$P(B)=\frac{\mathrm{C}_N^n\cdot n!}{N^n}.$$

分房问题的一个直接应用是下面的生日问题.

例 1.3.3 求 m 个人中至少有两人生日相同的概率($m\leqslant 365$).

解 设事件 $A=\{m$ 个人中至少有两人生日相同$\}$,如果将 365 天看作是 365 个房间,那么生日问题就是分房问题,比如某个人的生日是 3 月 18 日,就相当于把他分到 318 房间.注意到事件 A 的逆事件 $\bar{A}=\{$任何两个人生日都不一样$\}$,它相当于例 1.3.2 中的事件 B,所以

$$P(A)=1-\frac{\mathrm{C}_{365}^m\cdot m!}{365^m}.$$

经过计算可得如下结果:$m=23$ 时,$P(A)=0.5$;$m=50$ 时,$P(A)=0.97$;$m=64$ 时,$P(A)=0.997$.

也就是说 23 个人中至少有两个人生日相同的概率为 0.5,50 个人中几乎可以肯定至少有两个人生日相同.这个结果让人吃惊,与人们的直觉相差很大.直观感觉与事实有可能差别很大.

例 1.3.4(抽签问题) 设 15 个人要去看电影,但只有 7 张电影票,于是进行抽签决定谁去看电影.求第 5 个人抽到电影票的概率.

解 设事件 $A=\{$第 5 个人抽到电影票$\}$.这是一个不放回抽样,且题意与次序有关,所以样本点总数为

$$15\cdot 14\cdot\cdots\cdot 2\cdot 1=15!.$$

事件 A 要求第 5 个人抽到电影票.与例 1.3.1 的(3)一样,我们可以先任取一张电影票"预留"给第 5 个人,其他 14 个人任意抽取,那么 A 的样本点总数为 $7\cdot 14\cdot 13\cdot\cdots\cdot 2\cdot 1=7\cdot 14!$,所以

$$P(A)=\frac{7\cdot 14!}{15!}=\frac{7}{15}.$$

几乎每个人都有抽签的经历,也有许多人认为先抽与后抽是不一样的,然而此例说明,抽到"好签"的概率与先后次序无关.这一结果同样令人"不可理解",因为有不少人会说:"如果第一个人抽到电影票,那么第二个人抽到的概率就是 $\frac{6}{14}$,而不是 $\frac{7}{15}$;如果第一个人没抽到电影票,第二个人抽到的概率就是 $\frac{7}{14}$,也不是 $\frac{7}{15}$".需要注意的是,此处的 $\frac{6}{14}$ 与 $\frac{7}{14}$ 并不是第二个人抽到电影票的概率,它们是已知第一个人的结果的条件下,第二个人抽到电影票的概率,属于 1.4 节要介

绍的条件概率. 那么,要怎样理解抽签与次序无关这件事情呢? 我们可以这样理解:第一个人抽完后,不要看结果,第二个人抽完后也不看结果,等到最后一个人抽完后,这 15 个人一起把抽到的签亮出来,这就相当于把这 15 个签平均分给这 15 个人,所以每个人抽到电影票的概率应该是一样的. 也就是说,抽签问题看起来是有次序的,实际上是无次序的.

有了抽签的原则,许多问题可以变得很简单. 比如,设有 7 个红球 8 个白球共 15 个球放在一起,不小心丢了其中的 4 个球,现从剩下的 11 个球中任取一个,此球为白球的概率为多少? 丢失的 4 个球可以认为被前 4 个人拿走了,从剩下的 11 个球中任取一个相当于第 5 个人取球,利用抽签原则,可知此球为白球的概率为$\frac{7}{15}$.

例 1.3.5(分组法) 将 n 个不同的球分成 k 个不同的组,使得这 k 个组各有 $n_1, n_2, \cdots, n_k$ 个球,$n_1+n_2+\cdots+n_k=n$. 问:共有多少种分法?

解 显然同一组内的物体是不用考虑次序的,且取法是不放回的. 所以,从 n 个球中取出 n_1 个组成第一组有 $C_n^{n_1}$ 种取法,再从 $n-n_1$ 个球中取出 n_2 个组成第二组有 $C_{n-n_1}^{n_2}$ 种取法,……,最后剩下的必是 n_k 个球,正好组成第 k 组. 所以总的取法为

$$C_n^{n_1} C_{n-n_1}^{n_2} \cdots C_{n-n_1-n_2-\cdots-n_{k-2}}^{n_{k-1}} = \frac{n!}{n_1!\,n_2!\cdots n_k!}.$$

此结果说明,若要将 n 个不同的物体分成 k 个不同的组,各组分别有 n_1, $n_2, \cdots, n_k$ 个物体,那么总的分法就是物体总个数的阶乘除以每组中物体个数阶乘的乘积.

例 1.3.6 求在桥牌比赛(无大小王)中 4 张"A"落入同一个人手中的概率.

解 桥牌比赛中是将 52 张扑克(无大小王)平均分配给 4 个人,每人 13 张. 于是由分组法可知,样本点总数为$\frac{52!}{(13!)^4}$;为了保证 4 张"A"落入同一个人手中,只需先选择一人,有 C_4^1 种选法,将 4 张"A"分给此人,然后将剩下的 48 张扑克分成 4 组,其中得到 4 张"A"的人只需要 9 张,其余三人每人 13 张,有$\frac{48!}{9!\ (13!)^3}$种分法,所以符合条件的分法共有$\frac{C_4^1 \cdot 48!}{9!\ (13!)^3}$,最后得所求概率为$\frac{4 \cdot 48! \cdot 13!}{9! \cdot 52!} \approx 0.010\,56$.

1.3.2 几何概型

本节将考察等可能概型的另一类情形,它与古典概型的区别在于此时样本

点总数不是有限多个,而是无限多个. 当然此时已不能像古典概型那样通过求样本点总数来求概率了.

我们在一个面积为 S_Ω 的平面区域 Ω 中等可能地任意投点,这里"等可能"的确切含义是:对于区域 Ω 中的任意一个面积为 S_A 的小区域 A,点落入区域 A 中的可能性的大小与 S_A 成正比,而与区域 A 的位置以及形状无关. 如果"点落入区域 A"这个随机事件仍用 A 来表示,由题意可设 $P(A)=kS_A$,k 为比例系数,则由 $P(\Omega)=1$ 可得 $k=\dfrac{1}{S_\Omega}$,所以

$$P(A)=\frac{S_A}{S_\Omega}.$$

这一类概率通常称作**几何概率**. 请注意,如果是在一个线段上投点,那么应该将面积换成长度;如果是在一个三维区域内投点,应该将面积换成体积,并以此类推.

例 1.3.7　如果在一个 $5\times10^4\ \mathrm{km}^2$ 的海域里有表面积达 $40\ \mathrm{km}^2$ 的大陆架储藏着石油,假如在这海域里随意选定一点钻探,问钻到石油的概率是多少?

解　由于选点的随机性,可以认为该海域中各点被选中的可能性是一样的,因而这是一个几何概型. 所以所求概率为 $\dfrac{40}{50\ 000}=0.000\ 8$.

例 1.3.8(会面问题)　两人相约 7 点到 8 点在某地会面,先到者等候另一人 20 min,过时就离开,试求这两个人能会面的概率.

解　以 x,y 分别表示两个人的到达时刻(单位:min),则会面的充要条件为

$$|x-y|\leqslant 20.$$

这是一个几何概率问题,样本点的全体是边长为 60 的正方形里面的点,能会面的点为区域 $|x-y|\leqslant20$(图 1-7). 所以概率为 $\dfrac{60^2-40^2}{60^2}=\dfrac{5}{9}$.

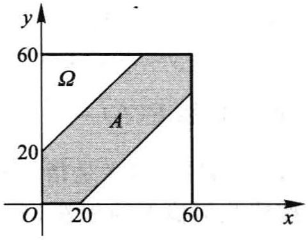


图 1-7

习　　题

1. 将 3 个球随机地放入 4 个杯子中,求杯子中球的最大个数分别为 1,2,3 的概率.

2. 在 1 000 个产品中有 200 个次品、800 个正品. 从中任取 100 个,求(1) 恰有 50 个次品的概率;(2) 至少有 3 个次品的概率.

3. 将 10 本书任意放到书架上,求其中仅有的 3 本外文书恰好排在一起的概率.

4. 设6位同学每位都等可能地进入10间教室的任何一间自习.求下列事件的概率：

(1) 某指定教室有两位同学；

(2) 6位同学所在的教室都不相同；

(3) 至少有2位同学在同一个教室.

5. 从5双不同的鞋子中任取4只，求：(1) 这4只鞋子至少有一双的概率；(2) 都不成双的概率.

6. 两人相约在早晨7点到8点之间在某处见面，先到者等候10 min，过时不候.求两人能见面的概率.

7. 在区间$(-1,1)$内随机地取两个实数r,s，求这两个实数使方程$x^2-2rx+s=0$有实根的概率.

8. 随机地向半圆$0<y<\sqrt{2ax-x^2}$(a为正常数)内掷一点，点落在半圆内任何区域的概率与区域的面积成正比，试求原点和该点的连线与x轴的夹角小于$\pi/4$的概率.

§1.4　条件概率

1.4.1　条件概率的定义

到目前为止，我们考虑的都是随机事件A的概率$P(A)$.在实际问题中，我们往往还要考虑另外一类概率，即在已知"事件B发生"的条件下，事件A的概率.此类概率我们称之为条件概率，记为$P(A|B)$.一般情况下，$P(A)$与$P(A|B)$是不相同的.下面先看两个小例子.

例1.4.1　从区间$[0,1]$上任取两点，记为x,y.令$A=\{(x,y):y\geqslant x^2\}$，$B=\{(x,y):y<x\}$.试计算$P(A)$与$P(A|B)$.

解　显然$P(A)=1-P(\bar{A})=1-\int_0^1 x^2\mathrm{d}x=\frac{2}{3}$.

$P(A|B)$的含义是指已知$\{y<x\}$已经发生，在此条件下求事件A发生的概率.按照几何概型的做法，此概率应是在B中满足A的要求的那部分面积(即AB的面积)除以B的面积.

$$|AB|=\int_0^1(x-x^2)\mathrm{d}x=\frac{1}{2}-\frac{1}{3}=\frac{1}{6},\ |B|=\frac{1}{2},$$

所以$P(A|B)=\frac{1}{3}$.

可以看到$P(A)\neq P(A|B)$.同时我们也发现对此题等式$P(A|B)=\frac{P(AB)}{P(B)}$成立.

例1.4.2　设有一对夫妻，他们有两个孩子(假设生男生女概率相等)，已知

其中一个为女孩，求这对夫妻有男孩的概率.

解 由题意知所求概率为条件概率. 设 $B=\{$其中一个为女孩$\}$，$A=\{$这对夫妻有男孩$\}$. 样本空间为 $\Omega=\{$(男，男)，(男，女)，(女，男)，(女，女)$\}$，由于事件 B 已经发生，所以样本点变成了三个：

(男，女)，(女，男)，(女，女).

在这三个样本点中，满足事件 A 要求的有两个样本点

(男，女)，(女，男).

于是根据古典概型的做法可知 $P(A|B)=\frac{2}{3}$.

另外我们可以很容易地求出 $P(AB)=\frac{1}{2}$，$P(B)=\frac{3}{4}$. 读者容易发现，等式 $P(A|B)=\frac{P(AB)}{P(B)}$仍然成立.

一般地，在已知事件 B 发生的条件下，事件 A 发生的概率称为条件概率，其定义如下：

定义 1.4.1 设 A 与 B 为两个事件，且 $P(B)>0$. 那么在已知“事件 B 发生”的条件下，事件 A 的条件概率 $P(A|B)$定义为

$$P(A|B)=\frac{P(AB)}{P(B)}. \tag{1.4.1}$$

例 1.4.3 设有 7 个红球，5 个白球，从中任取两次，每次一球不放回. 令 $A=\{$第一次为红球$\}$，$B=\{$第二次为红球$\}$. 试求 $P(A|B)$与 $P(B|A)$.

解 由抽签原则可知 $P(A)=P(B)=\frac{7}{12}$，又 $P(AB)=\frac{7\cdot 6}{12\cdot 11}=\frac{7}{22}$. 所以

$$P(A|B)=\frac{P(AB)}{P(B)}=\frac{6}{11},P(B|A)=\frac{P(AB)}{P(A)}=\frac{6}{11}.$$

不难验证，条件概率满足下列性质：

(1) 非负性：对任意事件 A，有 $P(A|B)\geqslant 0$；

(2) 归一性：$P(\Omega|B)=1$；

(3) 可列可加性：对任意的一列两两互不相容的事件 A_n，$n=1,2,\cdots$，有

$$P\left(\bigcup_{n=1}^{\infty}A_n\,|B\right)=\sum_{n=1}^{\infty}P(A_n\,|B).$$

1.4.2 乘法公式

如果 $P(A)>0$，$P(B)>0$，根据条件概率的定义可得

$$P(AB)=P(A)P(B|A),P(AB)=P(B)P(A|B). \tag{1.4.2}$$

以上两式均称为概率的乘法公式.

容易验证,对多个事件的乘法公式也成立:

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

其中 $P(A_1A_2\cdots A_{n-1})>0$.

下面,我们用乘法公式再一次验证抽签原则的正确性.

例 1.4.4(抽签问题续) 某超市举办有奖销售活动,投放 n 张奖券只有 1 张可以中奖.每位顾客可随机抽取一张,求第 k 位顾客中奖的概率($1\leqslant k\leqslant n$).

解 令事件 $A=\{$前 $k-1$ 位顾客未中奖而第 k 位顾客中奖$\}$,$A_i=\{$第 i 位顾客未中奖$\}$($i=1,2,\cdots,k-1$),$A_k=\{$第 k 位顾客中奖$\}$,则

$$\begin{aligned}P(A)&= P(A_1A_2\cdots A_{k-1}A_k) = P(A_1)P(A_2\mid A_1)P(A_3\mid A_1A_2)\cdots P(A_k\mid A_1A_2\cdots A_{k-1})\\&=\frac{n-1}{n}\times\frac{n-2}{n-1}\times\frac{n-3}{n-2}\times\cdots\times\frac{n-k+1}{n-k+2}\times\frac{1}{n-k+1}\\&=\frac{1}{n}.\end{aligned}$$

1.4.3 全概率公式与贝叶斯公式

当读者看到“全概率公式”时,一定觉得这个公式肯定很“厉害”,能把所有概率都求出来!其实不然,全概率公式确实功能强大,但绝不可能求出所有的概率.它的主要作用是计算我们在前面从来没有碰到过的一类(复杂)概率,这类概率在现实生活中非常常见,它们基本上只能用全概率公式来计算.为了介绍全概率公式,我们先从一个例子入手.

例 1.4.5 设有甲、乙两个盒子,甲中有 7 个红球 3 个白球,乙中有 5 个红球 6 个白球.先从甲盒中任取一球放入乙中,再从乙盒中任取一球,试求从乙盒中取出的是红球的概率.

这是一个我们以前从未遇到过的问题.由题目的叙述,我们可以看到,这个随机试验由先后两个阶段组成,第二阶段的结果受到第一阶段结果的影响,然而第一阶段的结果是未知的,所求概率又是第二阶段结果之概率.换言之,由甲中放入乙中球的颜色未知,所以乙中原有的 5 红 6 白这种组合发生了变化,可能变成 6 红 6 白,也可能是 5 红 7 白,这样就使得概率的计算变得无从下手.

为了求出概率,我们换一个角度来考虑.尽管不知道从甲中放入乙中的是一个什么颜色的球,但我们肯定知道这个球要么是红球,要么是白球.如果设 $A_1=\{$从甲中取得红球$\}$,$A_2=\{$从甲中取得白球$\}$,那么 $A_1A_2=\varnothing$,且 $A_1\cup A_2=\Omega$.又设 $B=\{$从乙中取得红球$\}$,于是

$$P(B)= P(B\cap\Omega) = P(B\cap(A_1\cup A_2)) = P(A_1B\cup A_2B)$$

$$= P(A_1B) + P(A_2B)$$
$$= P(A_1)P(B|A_1) + P(A_2)P(B|A_2),$$

上式右端的几个概率都是很简单的，易知

$$P(A_1) = \frac{7}{10}, P(A_2) = \frac{3}{10}, P(B|A_1) = \frac{6}{12}, P(B|A_2) = \frac{5}{12}.$$

所以 $P(B)=\frac{7}{10}\cdot\frac{6}{12}+\frac{3}{10}\cdot\frac{5}{12}=\frac{19}{40}$.

在这个问题的计算过程中，我们成功地把一个复杂的概率问题分解成几个易于计算的步骤，其中最关键的就是 A_1 与 A_2 的引入，而性质 $A_1A_2=\varnothing$ 与 $A_1\cup A_2=\Omega$ 起到了至关重要的作用. 一般地，我们给出如下的定义：

定义 1.4.2 设 $A_1,A_2,\cdots,A_n$ 为一组事件，如果

(1) $A_iA_j=\varnothing, \forall i\neq j$;

(2) $\bigcup\limits_{i=1}^{n}A_i=\Omega$,

则称 $A_1,A_2,\cdots,A_n$ 为 Ω 的一个划分(分割，完备事件组).

读者应当知道，组成一个划分的事件可以是无限多个.

事实上，例 1.4.5 的成功解决完全是利用了下面的全概率公式：

定理 1.4.1 设 $A_1,A_2,\cdots,A_n$ 为 Ω 的一个划分，则对任意的事件 B，有

$$P(B) = \sum_{i=1}^{n}P(A_i)P(B|A_i). \tag{1.4.3}$$

如果划分为 $A_1,A_2,\cdots$，则全概率公式为

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

证明 由于 $A_1,A_2,\cdots,A_n$ 为 Ω 的一个划分(图 1-8)，所以

$$B = B\cap\Omega = B\cap\left(\bigcup_{i=1}^{n}A_i\right)=\bigcup_{i=1}^{n}(A_iB),$$

且 $A_1B,A_2B,\cdots,A_nB$ 两两互不相容，所以由概率的可加性知

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

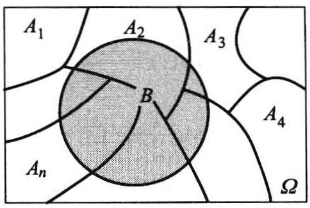


图 1-8

注意，在使用全概率公式的时候，关键是正确选择划分 $A_1,A_2,\cdots,A_n$. 如果选择不当，要么求不出结果，要么结果错误. 例 1.4.5 是一个典型的使用全概率来计算的问题，在此问题中，我们用 A_1，A_2 作为划分，需要注意的是，A_1，A_2 完全属于第一阶段的事件，它们囊括了第一阶段的所有两种

结果.如果将原题中“从甲中任取一球放入乙中”改成“从甲中任取两球放入乙中”,那么以上的 A_1,A_2 就不再是一个划分了.因为第一阶段有三种不同的结果:$A_i=\{$这两球中有 i 个红球$\}$,$i=0,1,2$.此时 A_0,A_1,A_2 才是一个划分,它包含了第一阶段的所有可能结果.

综上所述,使用全概率公式的题型大致具有下述特征:

(1) 随机试验可以分为两个相互影响的阶段;

(2) 第一阶段的所有可能结果已知;

(3) 所求概率为第二阶段结果之概率.

此类问题解决的关键在于划分的选取,划分的选取方法为利用第一阶段的所有可能结果来构造.

例 1.4.6 设有甲、乙两个盒子,甲中有7个红球3个白球,乙中有5个红球6个白球.先从甲盒中任取两球放入乙中,再从乙盒中任取一球,试求从乙盒中取出的是红球的概率.

解 设 $A_i=\{$这两球中有 i 个红球$\}$,$i=0,1,2$;则 A_0,A_1,A_2 为划分,又 $B=\{$从乙中取得红球$\}$,于是

$$\begin{aligned}P(B)&=P(A_0)P(B|A_0)+P(A_1)P(B|A_1)+P(A_2)P(B|A_2)\\&=\frac{C_7^0C_3^2}{C_{10}^2}\frac{5}{13}+\frac{C_7^1C_3^1}{C_{10}^2}\frac{6}{13}+\frac{C_7^2C_3^0}{C_{10}^2}\frac{7}{13}\\&=\frac{32}{65}.\end{aligned}$$

例 1.4.7 设有3旧6新9个乒乓球,第一次比赛任取3只,用后放回,第二次比赛任取3只.

(1) 求这3个球全是新球的概率.

(2) 若已知第二次比赛的3个乒乓球全新,求第一次比赛的3个乒乓球为二新一旧的概率.

解 (1) 第一次比赛任取3个乒乓球,这3个乒乓球共有4种不同情况:$A_i=\{$第一次3个乒乓球中有 i 个新球$\}$,$i=0,1,2,3$.

$B=\{$第2次的3个乒乓球中有 i 个新球$\}$,则 $P(B)=\sum\limits_{i=0}^{3}P(A_i)P(B|A_i)$.

由于 $P(A_i)=\dfrac{C_6^iC_3^{3-i}}{C_9^3}$,$P(B|A_i)=\dfrac{C_{6-i}^3}{C_9^3}$,所以

$$P(B)=\frac{1}{(C_9^3)^2}(C_3^3\cdot C_6^3+C_3^2\cdot C_6^1\cdot C_5^3+C_3^1\cdot C_6^2\cdot C_4^3+C_6^3\cdot C_3^3)=\frac{25}{441}.$$

(2) 利用条件概率的定义以及乘法公式,容易得到

$$P(A_2|B)=\frac{P(A_2B)}{P(B)}=\frac{P(A_2)P(B|A_2)}{P(B)}=\frac{9}{20}.$$

例 1.4.8 设有甲、乙两个盒子,甲中有 7 个红球 3 个白球,乙中有 5 个红球 6 个白球. 先从甲盒中任取两球放入乙中,再从乙盒中任取一球. 若已知从乙盒中取出的是红球,试求从甲盒中取出的是一个红球一个白球的概率.

此问题与例 1.4.6 不同,此题是已知第二阶段的结果,反求第一阶段的概率,且求的是条件概率.

解 设 $A_i=\{$这两球中有 i 个红球$\}$,$i=0,1,2$;$B=\{$从乙中取得红球$\}$,所求概率可表示为 $P(A_1|B)$,由例 1.4.6 知 $P(B)=\frac{32}{65}$,所以

$$P(A_1|B)=\frac{P(A_1B)}{P(B)}=\frac{P(A_1)P(B|A_1)}{P(B)}=\frac{\frac{14}{65}}{\frac{32}{65}}=\frac{7}{16}.$$

例 1.4.8 及其解答过程具有一定的普遍性,实际上我们可以给出下面的定理:

定理 1.4.2(贝叶斯公式) 设 $A_1,A_2,\cdots,A_n$ 是 Ω 的一个划分,如果 $P(A_k)>0$,$k=1,2,\cdots,n$,则对任意事件 B,只要 $P(B)>0$,就有

$$P(A_k|B)=\frac{P(A_k)P(B|A_k)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)}. \tag{1.4.4}$$

该定理的证明很简单,利用乘法公式 $P(A_kB)=P(A_k)P(B|A_k)$,以及全概率公式 $P(B)=\sum_{i=1}^{n}P(A_i)P(B|A_i)$ 可知

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

贝叶斯公式在概率论和数理统计中有着多方面的应用. 假定 $A_1,A_2,\cdots,A_n$ 是导致试验结果的"原因",$P(A_i)$称为**先验概率**,它们反映了各种"原因"发生的可能性大小,一般是以往经验的总结,在试验之前是已知的. 如果试验产生了结果 B,这个信息将有助于探讨结果发生的"原因". 条件概率 $P(A_i|B)$称为**后验概率**,它反映了各个原因对这个结果的贡献. 比如,对例 1.4.6,可以求出三个后验概率分别为 $P(A_0|B)=\frac{5}{96}$;$P(A_1|B)=\frac{42}{96}$;$P(A_2|B)=\frac{49}{96}$,此结果说明,在 B 发生的条件下,A_1 与 A_2 发生的可能性较大,A_0 发生的可能性较小.

数学家简介 贝叶斯(Bayes,1702—1763),英国数学家.1702 年出生于伦敦,做过神父.1742 年成为英国皇家学会会员,1763 年 4 月 7 日逝世. 贝叶斯在数学方面主要研究概率论. 他首先将归纳推理法用于概率论基础理论,并创立了

贝叶斯统计理论，对于统计决策函数、统计推断、统计的估算等做出了贡献. 贝叶斯所采用的许多术语被沿用至今.

习 题

1. 已知 $P(A)=0.25, P(B|A)=0.3, P(A|B)=0.5$，求 $P(A\cup B)$.

2. 设 $P(A)=P(B)=0.4, P(A|B)=0.3$，求 $P(\bar{A}|\bar{B})$.

3. 设 $P(\bar{A})=0.3, P(B)=0.4, P(A\bar{B})=0.5$，求 $P(B|A\cup\bar{B})$.

4. 证明：如果 $P(A|B)=1$，则 $P(\bar{B}|\bar{A})=1$.

5. 设第一个盒子中装有5只白球4只红球，第二个盒子中装有4只白球6只红球. 先从第一个盒子中任取2只球放入第二个盒子，再从第二个盒子中任取一球. 求此球为红球的概率.

6. 设第一个盒子中装有5只白球4只红球3只黑球，第二个盒子中装有3只白球4只红球5只黑球. 独立地分别在两个盒子中各取一球. 求：

(1) 至少有一只白球的概率；

(2) 有一只白球一只黑球的概率；

(3) 已知至少有一只白球，求有一只白球一只黑球的概率.

7. 空战中甲机先向乙机开火，击落乙机的概率为0.2；若乙机未被击落，就进行还击，击落甲机的概率为0.3；若甲机未被击落，再进攻乙机，击落乙机的概率为0.4. 求在这几个回合中，(1) 甲机被击落的概率；(2) 乙机被击落的概率.

8. 设有来自三个地区的各10名、15名和25名考生的报名表. 其中女生的报名表依次为3份、7份和15份. 现任取一个地区的报名表，从中先后抽出2份.

(1) 求先抽到的一份是女生表的概率；

(2) 已知后抽到的是男生表，求先抽到的一份是女生表的概率.

9. 一枚深水炸弹击沉一艘潜水艇的概率为0.3，击伤的概率为0.5，击不中的概率为0.2. 假设击伤两次也会导致潜水艇下沉. 求施放4枚深水炸弹能击沉潜水艇的概率.

10. 用“胎甲蛋白法”普查癌症，已知确有癌症者，查出为阳性的概率为0.95；未患癌症者，查出为阴性的概率为0.95. 一人生活在低发病区，该地区癌症发病率为0.000 5；另一人生活在高发病区，该地区癌症发病率为0.01. 若两人均查出为阳性，问他们真正患有癌症的概率各为多少？

§1.5 独立性与伯努利试验

1.5.1 事件的相互独立性

在上一节中，我们曾讲到乘法公式：若 $P(A)>0$，则 $P(AB)=P(A)P(B|A)$.

这说明一般的 $P(AB)=P(A)P(B)$ 并不成立. 也就是说,一般情况下,$P(B)$ 与 $P(B|A)$ 不一定相等,即事件 A 对事件 B 的概率是有影响的. 那么什么时候 $P(B)$ 与 $P(B|A)$ 相等呢? 为此,我们先给出独立性的定义.

定义 1.5.1 设 A 与 B 为两个事件,如果等式成立:

$$P(AB) = P(A)P(B), \tag{1.5.1}$$

则称事件 A 与 B 相互独立.

若 A 与 B 相互独立,由前面的讨论立刻可知若 $P(A)>0$,$P(B)>0$,那么必有 $P(A)=P(A|B)$,$P(B)=P(B|A)$. 即事件 A 发生与否对事件 B 的概率没有影响(注意,并不是 A 对 B 没有影响,是 A 对 B 的概率没有影响). 读者应该牢记判断 A 与 B 是否独立的唯一原则是等式 $P(AB)=P(A)P(B)$ 是否成立,不可用其他原则.

性质 1.5.1 设 $P(A)>0$,$P(B)>0$,且 A 与 B 相互独立,则 $AB\neq\varnothing$.

证明 因为 A 与 B 相互独立,且 $P(A)>0$,$P(B)>0$,所以 $P(AB)=P(A).P(B)>0$,故 $AB\neq\varnothing$.

该性质说明两个概率大于零的事件如果是相互独立的,那么它们一定是相交的. 其逆否命题也一定成立.

性质 1.5.2 若 $P(A)>0$,$P(B)>0$,且 $AB=\varnothing$,则 A 与 B 必不独立.

即互不相容的两个概率大于零的事件不可能相互独立.

性质 1.5.3 若事件 A 与 A 独立,则 $P(A)=0$ 或 1.

由于 A 与 A 独立,所以 $P(AA)=P(A)P(A)$,即 $P(AA)=P(A)^2$,故 $P(A)=0$ 或 1.

注意"$P(A)=0$ 或 1"与"$A=\varnothing$ 或 Ω"是两种不同的说法,由后者可得前者,但前者得不到后者,即由 $P(A)=0$,得不到 A 为不可能事件 $\varnothing$.

性质 1.5.4 在 (A,B),$(\bar{A},B)$,$(A,\bar{B})$,$(\bar{A},\bar{B})$ 这四对事件中,如果有一对独立,则另外三对也独立.

证明 这里只证明由 A 与 B 独立可得 $\bar{A}$ 与 B,其余留给读者练习.

由 A 与 B 独立,知 $P(AB)=P(A)P(B)$. 所以 $P(\bar{A}B)=P(B)-P(AB)=P(B)-P(A)P(B)=P(B)(1-P(A))=P(\bar{A})P(B)$,即 $\bar{A}$ 与 B 独立.

例 1.5.1 掷一均匀硬币两次,令 $A_1=\{$第一次为正面$\}$,$A_2=\{$第一次为反面$\}$,$A_3=\{$正反面各一次$\}$. 试判断 A_1,A_2,A_3 任何两个事件是否独立.

解 由于 $P(A_1)=\frac{1}{2}$,$P(A_2)=\frac{1}{2}$,$P(A_3)=\frac{2}{4}=\frac{1}{2}$,且 $P(A_1A_2)=\frac{1}{4}$,$P(A_1A_3)=\frac{1}{4}$,$P(A_2A_3)=\frac{1}{4}$. 所以,A_1,A_2,A_3 中任何两个事件都相互独立.

我们已经知道了两个事件相互独立的定义与基本性质，那么多个事件的相互独立是什么样的？

定义 1.5.2 设 $A_1,A_2,\cdots,A_n$ 为 n 个事件，从中任取的 k 个事件($2\leqslant k\leqslant n$) $A_{i_1},A_{i_2},\cdots,A_{i_k}$ 都满足

$$P(A_{i_1}A_{i_2}\cdots A_{i_k})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k}), \tag{1.5.2}$$

则称事件 $A_1,A_2,\cdots,A_n$ 相互独立.

如果 $A_1,A_2,\cdots,A_n$ 中任何两个事件都相互独立，即

$$P(A_iA_j)=P(A_i)P(A_j)\quad (i,j=1,2,\cdots,n;i\neq j),$$

则称 $A_1,A_2,\cdots,A_n$ 两两相互独立.

由定义可知，判断 n 个事件相互独立是一件极其复杂的事情. 在理论证明及实际问题中，n 个事件相互独立往往作为已知条件. 这样就给我们创造了便利条件，比如，如果已知 $A_1,A_2,\cdots,A_8$ 相互独立. 那么

$$P(\bar{A}_1A_2\bar{A}_5\bar{A}_7A_8)=P(\bar{A}_1)P(A_2)P(\bar{A}_5)P(\bar{A}_7)P(A_8).$$

同时读者也容易看出由 n 个事件相互独立可得两两独立，反之不成立.

例 1.5.2 由例 1.5.1 可知 A_1,A_2,A_3 是两两独立的，但由于 $P(A_1A_2A_3)=P(\varnothing)=0$，即 $P(A_1A_2A_3)\neq P(A_1)P(A_2)P(A_3)$，所以 A_1,A_2,A_3 不是相互独立的.

例 1.5.3 设事件 A,B,C 相互独立，试判断 $A-B$ 与 $\bar{C}$ 是否相互独立？

解 由于 A,B,C 相互独立，所以

$$\begin{aligned}P((A-B)\bar{C})&=P(A\overline{B\bar{C}})=P(A)P(\bar{B})P(\bar{C})\\&=P(A\bar{B})P(\bar{C})=P(A-B)P(\bar{C}).\end{aligned}$$

$A-B$ 与 $\bar{C}$ 相互独立.

1.5.2 n 重伯努利试验

下面我们用独立性来研究一类应用非常广泛的试验模型——n 重伯努利试验.

定义 1.5.3 如果试验 E 只有两个可能的结果：A 及 $\bar{A}$，并且 $P(A)=p$，$P(\bar{A})=1-p=q$(其中 $0<p<1$). 把试验 E 独立地重复做 n 次构成一个试验，用 E^n 表示，这个试验称为 **n 重独立重复试验或 n 重伯努利试验**.

将一相同硬币连续抛掷 n 次，向一目标连续射击 n 次，这样的试验均可认为是 n 重伯努利试验.

例 1.5.4 一名射手向目标连续射击 5 次，已知每次命中率均为 $p(0<p<1)$，且每次命中与否相互独立，求恰好命中三次的概率.

解 令 $A_i=\{$第 i 次命中$\}$，$i=1,2,3,4,5$. 由题意知 A_1,A_2,A_3,A_4,A_5 相互独立，又令 $B=\{$恰好命中三次$\}$. 则 B 发生的充要条件为 A_1,A_2,A_3,A_4,A_5 中恰好有三个发生而另两个不发生，共有 $C_5^3=10$ 种不同的情况，分别为

$$A_1A_2A_3\overline{A}_4\overline{A}_5,A_1A_2\overline{A}_3A_4\overline{A}_5,A_1\overline{A}_2A_3A_4\overline{A}_5,\overline{A}_1A_2A_3A_4\overline{A}_5,$$

$$A_1A_2\overline{A}_3\overline{A}_4A_5,A_1\overline{A}_2A_3\overline{A}_4A_5,\overline{A}_1A_2A_3\overline{A}_4A_5,A_1\overline{A}_2\overline{A}_3A_4A_5,$$

$$\overline{A}_1A_2\overline{A}_3A_4A_5,\overline{A}_1\overline{A}_2A_3A_4A_5,$$

利用 A_1,A_2,A_3,A_4,A_5 的相互独立性，可得每种情况发生的概率均为 $p^3(1-p)^2$，所以

$$P(B)=C_5^3p^3q^2,$$

其中 $q=1-p$.

该结果可以推广如下：在 n 次独立重复试验中，如果事件 A 在每次试验中发生的概率均为 p，那么 A 在这 n 次试验中恰好发生 k 次的概率为

$$C_n^kp^kq^{n-k},\ k=0,1,2,\cdots,n,$$

其中 $q=1-p$.

例 1.5.5 某学生平时根本不学习，期末参加考试，考题均为单项选择题，且每题 4 个选项. 该生只好瞎猜(每题任选一个选项). 若考题数量为 5 道，求此人能及格的概率. 如果考题数量分别变成 10 道与 20 道时，此人能及格的概率又是多少？

解 由题意知，此人做一道选择题相当于做一次试验，所以这是伯努利试验. 如果只有 5 道题，那么此人至少要猜对三道才能及格. 其概率为

$$C_5^3\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right)^2+C_5^4\left(\frac{1}{4}\right)^4\left(\frac{3}{4}\right)^1+C_5^5\left(\frac{1}{4}\right)^5\left(\frac{3}{4}\right)^0=\frac{105}{1\ 024}=0.103\ 5.$$

当题目数量为 10 道时，其概率为

$$\sum_{i=6}^{10}C_{10}^i\left(\frac{1}{4}\right)^i\left(\frac{3}{4}\right)^{10-i}\approx 0.01.$$

当题目数量为 20 道时，其概率为

$$\sum_{i=12}^{20}C_{20}^i\left(\frac{1}{4}\right)^i\left(\frac{3}{4}\right)^{20-i}\approx 0.000\ 9.$$

数学家简介 近代科学史上，最著名的科学家家族可能要算伯努利家族了. 伯努利家族是瑞士的一个曾产生过 11 位科学家的家族. 其中著名的有雅可比·伯努利、雅可比的弟弟约翰·伯努利、约翰的次子丹尼尔·伯努利等.

雅可比·伯努利(Jacobi Bernoulli，1654—1705)，是伯努利家族中重要的一员，卓越的数学家. 青年时曾学习神学，1676 年开始到荷兰、德国、法国旅行. 回

国后于1687年到1705年在巴塞尔大学任教.雅可比同莱布尼茨共同协作,对微积分的发展做出了出色的贡献,为常微分方程的积分法奠定了充分的理论基础,他创立了变分法.雅可比还是概率论的早期研究者.许多概率论的术语都是以他的名字命名的.丹尼尔·伯努利(Daniel Bernoulli,1700—1782)由于受到家庭的影响,从小对自然科学的各个领域有着极大兴趣.1716—1717年在巴塞尔大学学医,1718—1719年在海德堡大学学习哲学,1719—1720年又在斯特拉斯堡大学学习伦理学,此后专攻数学,1721年获得了医科大学学位,1725—1732年丹尼尔·伯努利在圣彼得堡科学院担任数学教师,1733—1750年他担任了巴塞尔大学的解剖学、植物学教授,1750年丹尼尔又任物理学教授和哲学教授,同年被选为英国皇家学会会员.丹尼尔是伯努利家庭中成就最大的科学家.在数学方面,丹尼尔的研究涉及代数、概率论、微积分、级数理论、微分方程等多学科的内容,取得了重大成就.在物理学方面,丹尼尔所取得的成功是惊人的,其中对流体力学和气体动力学的研究尤为突出.1738年出版的《流体力学》一书是他的代表作,书中根据能量守恒定律解决了流体的流动理论,提出了著名的伯努利定理,这是流体力学的重要基本定理之一.

习　题

1. 设事件 A,B 相互独立,且 $P(A\bar{B})=\frac{1}{4}$,$P(\bar{A}B)=\frac{1}{6}$,求 $P(AB)$.

2. 设 A,C 相互独立,B,C 相互独立,A,B 互不相容,证明 $A\cup B$ 与 C 相互独立.

3. 设 $0<P(B)<1$,证明 A,B 相互独立的充要条件为 $P(A\mid B)+P(\bar{A}\mid\bar{B})=1$.

4. 设两个相互独立的随机事件 A 和 B 都不发生的概率为 $\frac{1}{9}$,A 发生而 B 不发生的概率与 B 发生而 A 不发生的概率相等,试求 $P(A)$.

5. 设两两相互独立的三事件 A,B,C 满足条件:$ABC=\varnothing$,$P(A)=P(B)=P(C)<0.5$,且已知 $P(A\cup B\cup C)=9/16$,试求 $P(A)$.

6. 某射手对同一目标独立地进行4次射击,若至少命中一次的概率为 $\frac{80}{81}$,试求该射手的命中率.

7. 甲、乙、丙三人独立地向同一飞机射击,设击中的概率分别为0.4,0.5,0.7.如果只有一人击中,则飞机被击落的概率为0.2;如果有两人击中,则飞机被击落的概率为0.6;如果三人都击中,则飞机一定被击落.求飞机被击落的概率.

8. 已知每门高射炮击中敌机的概率都是0.3.现有一架敌机入侵,问需令多少门炮同时开火才能以0.99的概率击中敌机?

9. 某厂生产的仪器中一次检验合格的占60%,其余的需要重新调试.经调试的产品中有

80%经检验合格，而20%会被判定为不合格而不能出厂．现该厂生产了200台仪器，求下列事件的概率：

(1) 全部仪器都能出厂；

(2) 恰有10台不合格；

(3) 至少有2台不合格.

§1.6 综合例题

例 1.6.1 甲、乙两人投掷均匀硬币，其中甲投掷 $n+1$ 次，乙投掷 n 次．求“甲投出正面的次数大于乙投出正面的次数”这一事件的概率.

解 令$甲_正$＝甲投出的正面次数，$甲_反$＝甲投出的反面次数；$乙_正$＝乙投出的正面次数，$乙_反$＝乙投出的反面次数.

于是所求事件的概率为 $P(甲_正>乙_正)$，另一方面，显然有

$$\Omega-(甲_正>乙_正)=(甲_正\leqslant 乙_正)=(甲_反>乙_反),$$

因为硬币是均匀的，由对称性知

$$P(甲_正>乙_正)=P(甲_反>乙_反),$$

由此即得

$$P(甲_正>乙_正)=\frac{1}{2}.$$

例 1.6.2 17世纪上半叶，赌徒德·梅尔向数学家帕斯卡(Pascal)提出了一个问题：一颗骰子投4次至少得到一个六点与两颗骰子投24次至少得到一个双六，这两个随机事件中哪一个发生的可能性更大？

解 令 $A=\{$一颗骰子投4次至少得到一个六点$\}$，$B=\{$两颗骰子投24次至少得到一个双六$\}$.

读者应该知道，A 与 B 这两个随机事件分属于不同的样本空间，A 对应的样本空间是一颗骰子投4次的所有可能结果，样本点总数为 6^4；B 对应的样本空间是两颗骰子投24次的所有可能结果，样本点总数为36^{24}．下求 A 与 B 中的样本点数．读者可能已经注意到 A 与 B 中的样本点数，但 $\bar{A}$ 与 $\bar{B}$ 中的样本点数比较好计算．$\bar{A}$ 表示一颗骰子投4次，六点都不出现，$\bar{A}$ 中的样本点数为 5^4；$\bar{B}$ 表示两颗骰子投24次，双六一次也不出现，$\bar{B}$ 中的样本点数为35^{24}．我们可以得到：

$$P(A)=1-\frac{5^4}{6^4}\approx 0.517\ 7,P(B)=1-\frac{35^{24}}{36^{24}}\approx 0.491\ 4.$$

所以一颗骰子投4次至少得到一个六点发生的可能性大于0.5，两颗骰子投24次至少得到一个双六发生的可能性小于0.5.

例 1.6.3(Buffon 投针问题)　在平面上画满间距为 a 的平行直线，向该平面随机投掷一枚长度为 $l\ (l<a)$ 的针，试求针与直线相交的概率.

解　以 x 表示针的中点到最近的一条平行线的距离，φ 表示针与平行线的夹角. 针与平行线的置位关系见图 1-9.

显然有 $0\leqslant x\leqslant\frac{a}{2}$，$0\leqslant\varphi\leqslant\pi$. 以 G 表示边长为 $\frac{a}{2}$ 及 π 的长方形(图 1-10). 为使针与平行线相交，必须 $x\leqslant\frac{l}{2}\sin\varphi$，满足这个关系的区域记为 A. 故所求概率为

$$p=\frac{0.5\int_0^\pi l\sin\varphi\mathrm{d}\varphi}{0.5a\pi}=\frac{2l}{a\pi}.$$

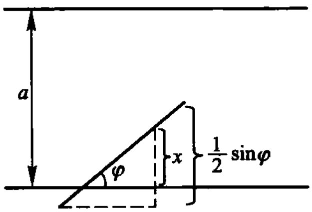


图 1-9

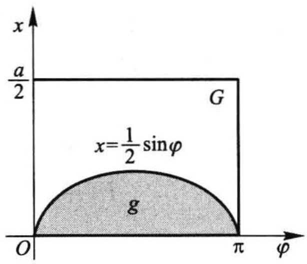


图 1-10

例 1.6.4(Polyá 罐子模型)　设罐中有 a 个红球 b 个黑球，随机取出一个，把原球放回，并加进与抽出球同色的球 c 个；再摸第二次，这样下去共摸了 n 次. 试证明第 n 次取球时取出红球的概率为 $\frac{a}{a+b}$.

解　设 $A_k=\{$第 k 次取球时取出红球$\}$，$k\geqslant1$. 用归纳法来证明此题. 首先，对 $k=1$，$P(A_1)=\frac{a}{a+b}$. 假设对 $k=n-1$ 结论成立，下证对 $k=n$ 结论也成立. 首先 A_1 与 $\bar{A}_1$ 为 Ω 的一个划分，其次应注意到 $P(A_n|A_1)$ 相当于在已知罐中有 $a+c$ 个红球及 b 个黑球的情况下，第 $n-1$ 次取出红球的概率. 由假设知 $P(A_n|A_1)=\frac{a+c}{a+b+c}$，同样可求出 $P(A_n|\bar{A}_1)=\frac{a}{a+b+c}$，所以

$$\begin{aligned}P(A_n)&=P(A_1)P(A_n|A_1)+P(\bar{A}_1)P(A_n|\bar{A}_1)\\&=\frac{a}{a+b}\frac{a+c}{a+b+c}+\frac{b}{a+b}\frac{a}{a+b+c}\\&=\frac{a}{a+b},\end{aligned}$$

因此,结论对一切 n 都成立.

例 1.6.5 甲乙二人比赛下棋,每局胜者得一分,甲在每局比赛中胜的概率为 α,乙在每局比赛中胜的概率为 $\beta(\alpha+\beta=1)$. 独立地进行比赛,直到有一人超过对方 2 分就停止,多得 2 分者胜. 求甲最终获胜的概率.

解 设 $A_i=\{$前两局中甲胜 i 局$\}$,$i=0,1,2$,则 A_0,A_1,A_2 为 Ω 的一个划分,又设 $B=\{$甲最终获胜$\}$,则

$$P(B)=P(A_0)P(B|A_0)+P(A_1)P(B|A_1)+P(A_2)P(B|A_2).$$

由题意知 $P(B|A_0)=0,P(B|A_1)=P(B),P(B|A_2)=1$,所以

$$P(B)=P(A_1)P(B)+P(A_2).$$

容易求出 $P(A_1)=2\alpha\beta,P(A_2)=\alpha^2$,故得

$$P(B)=\frac{\alpha^2}{\alpha^2+\beta^2}.$$

复 习 题 1

1. 某城市的电话号码由 8 位数组成,除首位数非 0 外,其余位数可以是 0,1,2,…,9 中的任一个数,求电话号码是由完全不同的数字组成的概率.

2. 将长度为 a 的木棒分成三段,求这三段可以构成一个三角形的概率.

3. 甲、乙两艘轮船驶向一个不能同时停泊两艘轮船的码头,它们在一昼夜内到达的时间是等可能的,如果甲船停泊的时间是 1 h,乙船停泊的时间是 2 h,求它们中任何一艘都不需要等候码头空出的概率.

4. 一口袋中有 6 个红球 4 个白球,从袋中任取一球观察颜色后放回袋中,再从袋中任取一球,设每次取球时袋中每个球被取中的可能性相同,求:

(1) 两次都取到红球的概率;

(2) 一次取到红球,一次取到白球的概率;

(3) 第二次取到红球的概率.

5. 一批灯泡共 100 只,次品率为 10%,不放回地抽取 3 次,每次 1 只,求第三次才取到合格品的概率.

6. 有三台车床,加工零件之比为 1∶2∶3,第一台车床的次品率为 0.05,第二台为 0.03,第三台为 0.02,将三台车床加工好的零件混合在一起,从中随机抽取一只,求:

(1) 所取零件为次品的概率;

(2) 已知取出的是次品,求它是第二台车床加工的概率.

7. 设事件 A,B 相互独立,且 $P(A\bar{B})=\frac{1}{3}$,$P(\bar{A}B)=\frac{1}{6}$,求 $P(AB)$.

8. 甲、乙两人独立地对同一目标射击 1 次,命中率分别为 0.6 和 0.5,现已知目标被射中,求它是乙射中的概率.

9. 某人给 5 位同学各写了一封信,并写好信封,然后随机地在每一信封里装入一封信,求下列事件的概率:

(1) 只有一封信装对;

(2) 没有一封信装对.

10. 设 $A_1,A_2,\cdots,A_n$ 为 Ω 的一个划分,且 $P(A_i)=p_i>0$,求事件 A_i 比 A_j 先发生的概率.

11. 设有 4 名战士每人在晚上睡觉之前都把枪放在门口,早晨紧急集合,每人随意拿一把枪参加集合,求至少有一个人拿到自己的枪的概率.如果是 n 名战士,概率又是多少?

12. 设某种微生物恰好繁殖 n 个幼虫的概率为 $\frac{6^n}{n!}e^{-6}$,每个幼虫能够成长为成虫的概率为 0.5,且每个幼虫能否成长为成虫是相互独立的,求恰有 m 个幼虫能成长为成虫的概率.

第 2 章　随机变量及其分布

随机变量是概率论最重要的概念之一，引入随机变量可将微积分中的理论和方法应用到概率论中. 随机变量可以说是古典概率向现代概率发展的里程碑.

本章中首先给出随机变量的概念及分布函数，然后介绍两类最重要的随机变量：离散型随机变量与连续型随机变量，包括离散型随机变量的概念及其分布列. 连续型随机变量的概念及其概率密度函数. 最后介绍随机变量的函数的分布.

§2.1　随机变量及其分布函数

2.1.1　随机变量

在第一章中我们研究了事件的概率以及概率论的基本概念. 大家可能已经发现，许多随机试验的结果都与实数有联系，比如下面几个例子：

例 2.1.1　n 重独立重复试验中事件 A 在每次试验中发生的概率均为 p. 在此试验中，我们主要关心事件 A 发生的次数，它是随机变量. 如果用 X 表示 A 发生的次数，那么 X 的取值是随机的，其可能的取值为 $0,1,2,\cdots,n$.

例 2.1.2　掷一枚硬币两次，样本空间为 $\Omega=\{\omega_{00},\omega_{01},\omega_{10},\omega_{11}\}$，0 表示反面，1 表示正面. 那么两次中出现正面的次数是不确定的，我们以 X 表示正面出现的次数，则有：$X(\omega_{00})=0,X(\omega_{01})=1,X(\omega_{10})=1,X(\omega_{11})=2$，0,1,2 为 X 的所有可能取值.

例 2.1.3　人有 10 个手指，人的指纹能确定身份，每个手指的指纹大致有两种，"斗"与"箕". 从某地区任取一人，若以 X 表示此人手指上斗的个数，那么 X 的取值是不确定的，但它的所有可能取值为 0、1、2、…、10. "斗"的个数因人而异，与遗传基因有关. 此时样本空间 Ω 就是该地区全体人，每个人都是一个样本点. $X(\omega)$就表示样本点 ω 对应的这个人的指纹中"斗"的个数.

下面给出一个重要概念——随机变量的定义.

定义 2.1.1　设 Ω 为一个样本空间，若对任意 $\omega\in\Omega$，都有一个实数 $X(\omega)$与之对应，则称 $X(\omega)$为一个**随机变量**，并简记为 X. 一般地随机变量用大写字母 X、Y、Z、…表示.

上面 3 例中的 X 都满足此定义.

下面再举几个随机变量的例子.

例 2.1.4　(1) 任取一个中国人,以 X 表示其寿命,此时 Ω 是全体中国人,而 X 的取值范围可定为 $[0,150]$.

(2) 记录某人每天接电话的次数,记录 100 天,以 X 表示他接电话次数. 此时,这 100 天的每一天都是一个样本点 $\omega_i, i=1,2,\cdots,100$. 而 $X(\omega_i)$ 表示第 i 天接到的电话个数,X 的可能取值为 $0,1,2,\cdots$.

2.1.2　随机变量的分布函数

由随机变量的定义可知,随机变量 X 就是一个由样本空间到实轴 $\mathbf{R}$ 的映射,即 $X: \Omega \to \mathbf{R}$. 它把每个样本点 $\omega \in \Omega$ 都变成一个实数 $X(\omega)$. 这不由使人想起在高等数学中函数 $f(x)$ 的定义. 严格地说,$f(x)$ 是一个由实轴到实轴的映射,它把每个实数 x 变成另一个实数. 但是,尽管随机变量 $X(\omega)$ 与 $f(x)$ 有诸多相似之处,但在处理方法上却有天壤之别. 大家知道,在高等数学中,我们先研究 $f(x)$ 的极限,然后是连续性、可导性等. 但对随机变量 $X(\omega)$ 却无法这么做,甚至于我们根本无法定义 $X(\omega)$ 的极限,比如:从 100 人中任取一个人,以 X 表示此人身高,此时 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_{100}\}$,每个 ω_i 都表示一个人. $X(\omega_i)$ 当然就是第 i 个人的身高,那么 X 在 ω_i 处的极限如何定义? 如果仿照 $f(x)$ 在 x_0 处的极限定义,应该是

$$\lim_{\omega \to \omega_1} X(\omega).$$

但是 ω_1 是一个人,$\omega \to \omega_1$ 是什么意思? 无意义! 总之,我们对一个随机变量 X 无法定义极限,更不用说什么连续性与可导性了. 另外,X 在某个样本点处取值大小并不重要. 那么我们研究什么呢? 我们主要研究随机变量 X 的概率性质. 由于 $\{X \leqslant a\}=\{\omega \in \Omega: X(\omega) \leqslant a\}$ 为 Ω 的子集,所以 $\{X \leqslant a\}$ 为事件,它就是一定有概率,记为 $P(X \leqslant a)$. 这样的概率对随机变量 X 来说是至关重要的,比如,以 X 表示 100 人中任一个人的身高(cm). 如果已知 $P(X \leqslant 180)=0.88, P(X \leqslant 155)=0.17$. 那么,这 100 人中大约有 12 个身高超过 180 cm,完全可以组成一支篮球队,而身高小于 155 cm 的人大约有 17 人. 易知,当 a 变化时,概率 $P(X \leqslant a)$ 也会随之变化,我们给出如下定义:

定义 2.1.2　设 X 为一个随机变量,我们称

$$F(x)=P(X \leqslant x), \quad x \in \mathbf{R} \tag{2.1.1}$$

为随机变量的**分布函数**.

如果把随机变量 X 看作是实轴 $\mathbf{R}$ 上的一个随机点,$F(x)$ 是指随机点 X 落

在固定点 x 左方(含点 x)的概率.

每个随机变量都对应一个唯一分布函数,并且一维随机变量的概率特性完全由它的分布函数来决定.所以分布函数对于研究随机变量是至关重要的,因为它包含了随机变量几乎一切概率信息.所有关于随机变量 X 的概率计算都可借助其分布函数来完成,比如

$$P(a<X\leqslant b)=P(X\leqslant b)-P(X\leqslant a)=F(b)-F(a). \qquad (2.1.2)$$

分布函数的基本性质:

(1) $F(x)$为 x 的右连续函数,即对 $\forall x\in\mathbf{R}$, $\lim\limits_{h\to 0^+}F(x+h)=F(x)$.(证略)

(2) $F(x)$为 x 的单调不减函数.

(3) $F(-\infty)=0, F(+\infty)=1$.

证明 若 $x_1<x_2$,则 $F(x_2)-F(x_1)=P(X\leqslant x_2)-P(X\leqslant x_1)=P(x_1\leqslant X\leqslant x_2)\geqslant 0$,

$$F(-\infty)=\lim_{x\to-\infty}F(x)=\lim_{x\to-\infty}P(X\leqslant x)=P(X\leqslant -\infty)=0,$$

$$F(+\infty)=\lim_{x\to+\infty}F(x)=\lim_{x\to+\infty}P(X\leqslant x)=P(X\leqslant +\infty)=1.$$

以上三条性质是所有的分布函数都必须满足的.反之,若一个函数 $g(x)$满足上述三条性质,那么必存在一个随机变量,使 g 为其分布函数.利用分布函数的定义与性质可以证明:

$$P(a\leqslant X\leqslant b)=F(b)-F(a-0),$$

$$P(a<X<b)=F(b-0)-F(a),$$

$$P(a\leqslant X<b)=F(b-0)-F(a-0).$$

例 2.1.5 设 $F_1(x),F_2(x),\cdots,F_n(x)$为 n 个分布函数,那么当 $a_i\geqslant 0, i=1,2,\cdots,n$, $\sum\limits_{i=1}^{n}a_i=1$ 时, $\sum\limits_{i=1}^{n}a_iF_i(x)$ 是分布函数.

例 2.1.6 掷一枚均匀硬币 3 次,以 X 表示出现正面的次数,求 X 的分布函数.

解 由题知 X 的可能取值为 0,1,2,3.所以,

当 $x<0$ 时,$F(x)=P(X\leqslant x)=P(\varnothing)=0$;

当 $0\leqslant x<1$ 时,$F(x)=P(X\leqslant x)=P(X=0)=\dfrac{1}{8}$;

当 $1\leqslant x<2$ 时,
$$\begin{aligned}F(x)&=P(X\leqslant x)=P(\{X=0\}\cup\{X=1\})\\&=P(X=0)+P(X=1)\\&=\frac{1}{8}+\frac{3}{8}=\frac{1}{2};\end{aligned}$$

当 $2\leqslant x<3$ 时，$F(x)=P(X\leqslant x)=\frac{1}{8}+\frac{3}{8}+\frac{3}{8}=\frac{7}{8}$；

当 $x\geqslant 3$ 时，$F(x)=P(X\leqslant x)=1$.

所以

$$F(x)=\begin{cases}0, & x<0,\\ \frac{1}{8}, & 0\leqslant x<1,\\ \frac{1}{2}, & 1\leqslant x<2,\\ \frac{7}{8}, & 2\leqslant x<3,\\ 1, & x\geqslant 3.\end{cases}$$

例 2.1.7 在曲线 $y=2x-x^2$ 与 x 轴所围区域中等可能地投点，以 X 表示该点到 y 轴的距离，求 X 的分布函数.

解 首先由题知，X 的取值范围为区间 $[0,2]$. 所以，当 $x<0$ 时，$F(x)=0$；当 $x\geqslant 2$ 时，$F(x)=1$；当 $0\leqslant x<2$ 时，事件 $\{X\leqslant x\}$ 对应区域 D_x. 故

$$F(x)=P(X\leqslant x)=\frac{D_x\text{ 的面积}}{G\text{ 的面积}}=\frac{\int_0^x(2t-t^2)\mathrm{d}t}{\int_0^2(2t-t^2)\mathrm{d}t}=\frac{3}{4}\left(x^2-\frac{1}{3}x^3\right).$$

于是

$$F(x)=\begin{cases}0, & x<0,\\ \frac{3}{4}\left(x^2-\frac{1}{3}x^3\right), & 0\leqslant x<2,\\ 1, & x\geqslant 2.\end{cases}$$

例 2.1.6 中的分布函数是一个阶梯函数，而例 2.1.7 中的分布函数却是一个连续函数，而且例 2.1.7 中的分布函数 $F(x)$ 还可以写成如下形式

$$F(x)=\int_{-\infty}^x f(t)\mathrm{d}t,\quad x\in\mathbf{R}.$$

其中 $f(t)=\begin{cases}\frac{3}{4}(2t-t^2), & 0\leqslant t<2,\\ 0, & \text{其他}.\end{cases}$

一个随机变量的概率特性由它的分布函数来决定，通常在实际应用中，主要研究下面将要介绍的离散型与连续型这两类随机变量.

习 题

1. 已知随机变量 X 的分布函数为

$$F(x)=\begin{cases}1-e^{-x}, & x\geqslant 0,\\ 0, & x<0,\end{cases}$$

求 $P(X\leqslant 2)$，$P(0<X\leqslant 3)$，$P(X>\ln 2)$.

2. 设随机变量 X 的分布函数为

$$F(x)=A+B\arctan x,\quad -\infty<x<+\infty,$$

求常数 A,B.

§2.2 离散型随机变量

2.2.1 分布列及其性质

定义 2.2.1 如果随机变量 X 的所有可能取值为一列离散的点(可以编号)，$x_1,x_2,\cdots,x_i,\cdots$，则称 X 为一个离散型随机变量，并称概率

$$P(X=x_i)=p_i,\quad i=1,2,\cdots \tag{2.2.1}$$

为 X 的分布列或分布律. 分布列也常常可以写成下面的形式：

$$X\sim\begin{pmatrix}x_1 & x_2 & x_3 & \cdots\\ p_1 & p_2 & p_3 & \cdots\end{pmatrix}.$$

例 2.2.1 例 2.1.2 中的 X 就是一个离散型随机变量，可能值为 0,1,2,3，不难求出它的分布列为

$$X\sim\begin{pmatrix}0 & 1 & 2 & 3\\ \dfrac{1}{8} & \dfrac{3}{8} & \dfrac{3}{8} & \dfrac{1}{8}\end{pmatrix}.$$

由例 2.1.2 与例 2.2.1 可以看出：一个离散随机变量的分布函数与其分布列等价，但对于离散型随机变量来说，研究分布列远比研究分布函数要简单. 我们已经知道分布函数对一个随机变量来说至关重要，那么分布列对于一个离散随机变量来说就是第一重要的. 在 2.1 节中，我们已经知道例 2.1.2 中的分布函数是一个阶梯函数，其实我们有更加一般的结论：**离散型随机变量的分布函数必为阶梯函数，反之，分布函数为阶梯函数的随机变量必为离散型随机变量.**

易知，离散型随机变量的分布列必满足下述性质：

(1) 非负性：$p_i\geqslant 0(i=1,2,\cdots)$；

(2) 归一性：$\sum_{i=1}^{\infty} p_i = 1$.

下面是几个求离散随机变量分布列的例子.

例 2.2.2 设在 n 重独立重复试验中，事件 A 在每次试验中发生的概率为 P，以 X 表示事件 A 在这 n 次试验中发生的次数，求 X 的分布列.

解 由题知，X 的可能取值为 $0,1,\cdots,n$. 事件 $\{X=k\}$ 表示在 n 重独立重复试验中 A 恰发生 k 次，所以

$$P(X=k)=\mathrm{C}_n^k p^k q^{n-k}, \quad k=0,1,2,\cdots,n.$$

例 2.2.3 设某射手向一目标独立地进行连续射击，每次命中的概率都是 p，以 X 表示首次命中时的射击次数，求 X 的分布列.

解 X 的可能取值为 $1,2,\cdots$，事件 $\{X=k\}$ 表示在前 k 次射击中，第 k 次命中前 $k-1$ 次都没有命中，所以

$$P(X=k)=pq^{k-1}, \quad k=1,2,\cdots.$$

例 2.2.4 设某射手向一目标独立地进行连续射击，每次命中的概率都是 p，以 X 表示第三次命中时的射击次数，求 X 的分布列.

解 X 的可能取值为 $3,4,\cdots$，事件 $\{X=k\}$ 表示在前 k 次射击中，第 k 次命中了，前 $k-1$ 次中有两次命中，另外的 $k-3$ 次没有命中，所以

$$P(X=k)=\mathrm{C}_{k-1}^2 p^3 q^{k-3}, \quad k=3,4,\cdots.$$

例 2.2.5 设一口袋中共有 N 个球，其中有 M 个红球，现从口袋中任取 n 个球，以 X 表示这 n 个球中红球的个数，求 X 的分布列.

解 X 的可能取值为 $0,1,\cdots,\min\{M,n\}$，事件 $\{X=k\}$ 表示取出的这 n 个球中恰有 k 个红球，其概率为

$$P(X=k)=\frac{\mathrm{C}_M^k \mathrm{C}_{N-M}^{n-k}}{\mathrm{C}_N^n}, \quad k=0,1,\cdots,\min\{M,n\}.$$

例 2.2.6 设某随机变量 X 的分布列为

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

其中 A 为非负常数，求 A.

解 由分布列的归一性可得

$$\sum_{k=0}^{\infty} P(X=k) = \sum_{k=0}^{\infty} A\frac{\lambda^k}{k!} = 1,$$

又因为 $\mathrm{e}^{\lambda} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$，所以可得 $A=\mathrm{e}^{-\lambda}$.

2.2.2 常见的离散型随机变量

如前所述,分布列对研究离散随机变量来说是至关重要的,也就是说,一旦知道一个离散随机变量的分布列,就相当于完全掌握了该离散型随机变量.实际应用中经常出现的离散随机变量并不太多,主要有以下几个.

1. 二项分布

定义 2.2.2 设 X 为一个离散型随机变量,若 X 的分布列为

$$P(X=k)=C_n^k p^k q^{n-k},\quad k=0,1,\cdots,n. \tag{2.2.2}$$

其中,$0<p<1,q=1-p$,则称 X 服从参数为 n 与 p 的二项分布,记为 $X\sim B(n,p)$.

由例 2.2.2 可知,二项分布 $B(n,p)$ 就是在 n 重独立重复试验中事件 A 发生的概率,p 为 A 在每次试验中发生的次数.习惯上,我们将参数为 1 与 p 的二项分布 $B(1,p)$ 称为 0−1 分布(或两点分布),它表示在一次试验中事件 A 发生的概率,它是最简单的随机变量之一.但 0−1 分布有特殊的用途,一般来说,两状态的现象可以用 0−1 分布来描述.比如,人的生与死、电闸的开与关、是与否等问题.另外,我们也可以用 0−1 分布来构造一些比较复杂且难以处理的随机变量.

例 2.2.7 根据以往资料,患有某种疾病的人死亡率为 0.002,试求 2 000 名患者中死亡人数大于 8 的概率.

解 以 X 表示这 2 000 名患者中的死亡人数,每观察一名患者就相当于做了一次试验,观察这 2 000 名患者就相当于做了 2 000 次独立重复试验,所以由题意可知随机变量 X 服从参数为 2 000 与 0.002 的二项分布 $B(2\,000,0.002)$.死亡人数大于 8 就是 $\{X>8\}$,所以

$$P(X>8)=1-P(X\leqslant 8)=1-\sum_{k=0}^{8}C_{2\,000}^k\,0.002^k\,0.998^{2\,000-k}.$$

大家可以看到,上面这个概率的计算会非常麻烦,我们将使用下面的泊松逼近定理来计算这个概率.

2. 泊松(Poisson)分布

定义 2.2.3 设 X 为一个离散型随机变量,若 X 的分布列为

$$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda},\quad k=0,1,\cdots, \tag{2.2.3}$$

其中,$\lambda>0$,则称 X 服从参数为 λ 的泊松分布,记为 $X\sim P(\lambda)$.

泊松分布是一种理想化的分布,许多自然现象及稀有事件均可用泊松分布来描述,比如一个人每天接电话的次数,某地区每年发生的洪水次数、遭受的台风次数、某高校每年的自杀人数,某商店每天 8:00—10:00 到达的顾客人数以及

各种稀有事件发生的次数等,这类随机现象用泊松分布来描述比较恰当、逼真.同时泊松分布与二项分布具有密切的联系:

泊松逼近定理 设 $X_n \sim B(n, p_n)$,常数 $\lambda>0$,如果 $\lim\limits_{n\to\infty} np_n=\lambda$,则有

$$\lim_{n\to\infty} \mathrm{C}_n^k p_n^k (1-p_n)^{n-k}=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda} \quad (k=0,1,\cdots). \tag{2.2.4}$$

证明
$$\begin{aligned}\lim_{n\to\infty} \mathrm{C}_n^k p_n^k (1-p_n)^{n-k}&=\lim_{n\to\infty} \mathrm{C}_n^k \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}\\&=\lim_{n\to\infty} \frac{\lambda^k}{k!}\cdot\frac{n(n-1)\cdots(n-k+1)}{n^k}\cdot\left(1-\frac{\lambda}{n}\right)^n\cdot\left(1-\frac{\lambda}{n}\right)^{-k}\\&=\frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}.\end{aligned}$$

上面用到了重要极限 $\lim\limits_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n=\lim\limits_{n\to\infty}\left[\left(1+\left(-\frac{\lambda}{n}\right)\right)\right]^{-\frac{n}{\lambda}\cdot(-\lambda)}=\mathrm{e}^{-\lambda}$.

该定理说明,当 n 很大时,二项分布 $B(n,p)$ 与泊松分布 $P(\lambda)(\lambda=np)$ 几乎一样.它不仅说明了泊松分布与二项分布之间的联系,同时对二项分布的概率计算提供了很大的帮助.利用该定理以及附录中的泊松分布表,读者容易求出例 2.2.7 中的概率为 0.021 368.

例 2.2.8 设某工厂产品的次品率为 0.02,从该厂生产的一大批产品中随机抽取 100 件进行检测,求:

(1) 恰有 2 件次品的概率;

(2) 次品数不超过 2 件的概率.

解 由于产品非常多,无论是有放回抽样还是不放回抽样,都可以作为有放回抽样来处理.若以 X 表示 100 件产品中的次品数,那么随机变量 X 服从二项分布 $B(100,0.02)$.故得

(1) $P(X=2)=\mathrm{C}_{100}^2\cdot 0.02^2\times 0.98^{98}\approx\frac{2^2}{2!}\mathrm{e}^{-2}\approx 0.270\ 7$.

(2) $$\begin{aligned}P(X\leqslant 2)&=0.98^{100}+\mathrm{C}_{100}^1\cdot 0.02\times 0.98^{99}+\mathrm{C}_{100}^2\cdot 0.02^2\times 0.98^{98}\\&\approx\mathrm{e}^{-2}+\frac{2^1}{1!}\mathrm{e}^{-2}+\frac{2^2}{2!}\mathrm{e}^{-2}=5\mathrm{e}^{-2}\approx 0.676\ 7.\end{aligned}$$

3. 几何分布

定义 2.2.4 设 X 为一个离散型随机变量,若 X 的分布列为

$$P(X=k)=pq^{k-1}, \quad k=1,2,\cdots, \tag{2.2.5}$$

其中,$p>0$,$q=1-p$,则称 X 服从参数为 p 的几何分布,记为 $X\sim G(p)$.

由例 2.2.3 可知,首次命中的射击次数就服从参数为 p 的几何分布,反之,如果 X 服从参数为 p 的几何分布,那么就可以将 X 看作是首次命中的射击次

数,其中参数 p 是每次射击命中的概率.

例 2.2.9(几何分布的无记忆性) 设 X 服从参数为 p 的几何分布,那么对于任何正整数 n,m,都有

$$P(X>n+m|X>m)=P(X>n). \tag{2.2.6}$$

解 因为 $P(X>n)=\sum\limits_{k=n+1}^{\infty}P(X=k)=\sum\limits_{k=n+1}^{\infty}pq^{k-1}=q^n$,所以

$$P(X>m)=q^m,\quad P(X>n+m)=q^{n+m},$$

于是有

$$P(X>n+m|X>m)=\frac{P(X>n+m)}{P(X>m)}=q^n=P(X>n).$$

$\{X>n\}$表示前 n 次射击没有命中,那么 $P(X>n+m|X>m)=P(X>n)$就是说,在已知前 m 次没有命中的条件下,再射击 n 次也没有命中的概率等于前 n 次射击没有命中的概率,相当于前 m 次没有命中的信息被遗忘了.

*4. 超几何分布

定义 2.2.5 设 X 为一个离散型随机变量,若 X 的分布列为

$$P(X=k)=\frac{C_M^kC_{N-M}^{n-k}}{C_N^n},\quad k=0,1,\cdots,\min\{M,n\}, \tag{2.2.7}$$

其中,$N>M,N>n$,则称 X 服从参数为 n,M,N 的超几何分布,记为 $X\sim H(n,M,N)$.

例 2.2.5 中的随机变量就服从超几何分布.

数学家简介 泊松(Poisson,1781—1840),法国数学家、物理学家和力学家.泊松在数学方面贡献很多,最突出的是 1837 年提出概率论中的泊松分布,这一分布在公用事业、放射性现象等许多方面都有应用.他还研究过定积分、傅里叶级数、数学物理方程等.除泊松分布外,还有许多数学名词是以他名字命名的,如泊松积分、泊松求和公式、泊松方程、泊松定理,等等.

习 题

1. 一袋中装有 5 只球,编号为 1,2,3,4,5.在袋中同时取出 3 只球,以 X 表示取出的 3 只球中的最大号码,求随机变量 X 的分布列.

2. 一批产品共 20 个,其中有 4 个次品.从中抽取 6 个产品,以 X 表示次品的个数,分别在有放回与不放回两种情形下求 X 的分布列.

3. 设离散型随机变量 X 的分布列为 $P(X=k)=\frac{a}{2k+1}(k=0,1,2,3)$,求:(1) 常数 a;(2) $P(X<2)$.

4. 在相同条件下独立地进行 5 次射击，每次射击时击中目标的概率为 0.6，以 X 表示击中目标的次数. 求：(1) X 的分布列；(2) 恰有两次击中的概率；(3) 至少击中两次的概率.

5. 设随机变量 $X \sim P(\lambda)$，问 k 为何值时，概率 $P(X=k)$ 最大？最大值为多少？

6. 从学校乘汽车到火车站的途中有 3 个交通岗，假设在各个交通岗遇到红灯的事件是相互独立的，并且概率都为 $\frac{1}{4}$，设 X 为途中遇到红灯的次数，求：(1) X 的分布列；(2) 至多遇到一次红灯的概率.

7. 设随机变量 $X \sim B(2,p)$，随机变量 $Y \sim B(3,p)$. 已知 $P(X \geqslant 1)=\frac{5}{9}$，求 $P(Y \geqslant 1)$.

8. 设每分钟通过某交叉路口的汽车流量 X 服从泊松分布，且已知在一分钟内无车辆通过与恰有一辆车通过的概率相同，求在一分钟内至少有两辆车通过的概率.

9. 甲、乙两人投篮，投中的概率分别为 0.6、0.7，今各投 3 次，求：

(1) 两人投中次数相等的概率；

(2) 甲比乙投中次数多的概率.

10. 已知在 5 重伯努里试验中成功的次数 X 满足 $P\{X=1\}=P\{X=2\}$，求概率 $P\{X=4\}$.

11. 同时掷两枚骰子，直到一枚骰子出现 6 点为止，求抛掷次数 X 的分布列.

12. 设有 80 台相同的机器，每台机器的工作是相互独立的，发射故障的概率都是 0.01，且一台机器的故障只能由一人维修. 下面有两种配备维修工人的方法，第一种是由 4 人维护，每人 20 台；第二种是由 3 人共同维护 80 台机器. 问哪种方法更好？

§2.3 连续型随机变量

2.3.1 连续型随机变量的定义与密度函数

在例 2.1.3 中，随机变量 X 的分布函数 $F(x)$ 不是一个阶梯函数，所以它不是一个离散随机变量，且可以写成 $F(x)=\int_{-\infty}^{x} f(t)\,\mathrm{d}t$，这类随机变量我们称之为连续型随机变量.

定义 2.3.1 设 X 为一个随机变量，如果存在一个可积函数 $f(x)$ 使得 X 的分布函数 $F(x)$ 满足

$$F(x)=\int_{-\infty}^{x} f(t)\,\mathrm{d}t, \quad x \in \mathbf{R}, \tag{2.3.1}$$

则称 X 为一个连续型随机变量，并称 $f(x)$ 为 X 的概率密度函数，简称密度函数（或概率密度）.

容易看出，密度函数 $f(x)$ 具有以下性质：

(1) 非负性：$f(x) \geqslant 0$；

(2) 归一性：$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1$.

注 (1) 连续型随机变量的密度函数相当于离散型随机变量的分布列. 连续型随机变量可由密度函数唯一确定，知道了密度函数就可以完全掌握其对应的连续型随机变量. 另外，连续型随机变量的分布函数必为连续函数. 但是有反例可以说明，分布函数为连续函数的随机变量不一定是连续型随机变量.

(2) 由归一性知，介于密度曲线 $y=f(x)$ 与 x 轴之间的图形的面积为 1.

(3) 在密度函数 $f(x)$ 的连续点 x_0 处有 $F'(x_0)=f(x_0)$，即密度函数为分布函数的导数.

(4) 如果 X 为连续型随机变量，那么

$$P(a < X \leqslant b) = \int_a^b f(x)\mathrm{d}x. \tag{2.3.2}$$

事实上，$P(a<X\leqslant b)=P(X\leqslant b)-P(X\leqslant a)=F(b)-F(a)$

$$= \int_{-\infty}^b f(x)\mathrm{d}x - \int_{-\infty}^a f(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x .$$

(2.3.2)式虽然简单，但非常重要，因为从理论上来讲，所有连续型随机变量求概率都要使用它. 由图 2－1 知，X 落在区间 $(a,b]$ 的概率 $P(a<X\leqslant b)$ 等于区间 $(a,b]$ 上曲线 $y=f(x)$ 与 x 轴之间的图形的面积.

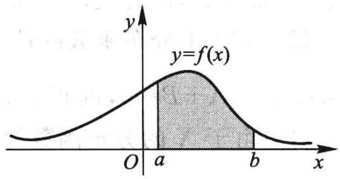


图 2－1

下面换个角度来理解概率密度的意义. 假设在区间 $[a,b]$ 上分布着一些质量，在每个点处质量密度为 $f(x)$，则区间 $[a,b]$ 的总质量为

$$\int_a^b f(x)\mathrm{d}x . \tag{2.3.3}$$

比较(2.3.2)式与(2.3.3)式，我们完全可以把密度函数类比为质量密度函数，把求概率类比为求质量.

另外，由图 2－1 知，对任何常数 c，如果 X 为连续型随机变量，就有

$$P(X=c)=0. \tag{2.3.4}$$

事实上，设 X 的分布函数为 $F(x)$，$\Delta x>0$，则有

$$0 \leqslant P(X = c) \leqslant P(c < X \leqslant c + \Delta x)$$

$$= \int_c^{c+\Delta x} f(x)\mathrm{d}x = f(\xi) \cdot \Delta x , c < \xi < c + \Delta x ,$$

令 $\Delta x \to 0$，可得 $P(X=c)=0$.

这说明，对于连续型随机变量，一般讨论的是 X 落在某区间内的概率，并且计算随机变量落在该区间概率时，无需区分该区间是开区间还是闭区间或半开半闭区间. 即

$$\begin{aligned} P(a < X < b) &= P(a < X \leqslant b) = P(a \leqslant X \leqslant b) \\ &= P(a \leqslant X < b) = \int_a^b f(x)\mathrm{d}x . \end{aligned}$$

还需要指出，虽然 $P(X=c)=0$，但并不意味着 $\{X=c\}$ 一定是不可能事件. 也就是说，若 $A=\varnothing$，则有 $P(A)=0$；反之若 $P(A)=0$，并不能推出 $A=\varnothing$. 连续型随机变量取个别值的概率为 0，这是它与离散型随机变量完全不同的一个重要区别.

例 2.3.1 例 2.1.3 中随机变量 X 的密度函数为

$$f(x)=\begin{cases} \dfrac{3}{4}(2x-x^2), & 0\leqslant x<2, \\ 0, & \text{其他}. \end{cases}$$

例 2.3.2 设连续型随机变量 X 的分布函数为

$$F(x)=\begin{cases} A+B\mathrm{e}^{-x}, & x>0, \\ 0, & x\leqslant 0. \end{cases}$$

求：(1) 常数 A,B 的值；(2) 密度函数 $f(x)$ 的表达式.

解 (1) 由分布函数的性质 $\lim\limits_{x\to+\infty}F(x)=1$，立刻得到 $A=1$. 又由 $\lim\limits_{x\to 0^+}F(x)=F(0)$，可得 $A+B=0$，进而有 $B=-1$.

(2) 由于 X 的分布函数为

$$F(x)=\begin{cases} 1-\mathrm{e}^{-x}, & x>0, \\ 0, & x\leqslant 0. \end{cases}$$

因此 X 的密度函数为

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

需要注意，此处的区间 $(0,+\infty)$ 对于连续型随机变量的运算很重要. 关于连续型随机变量 X 的一切概率运算（例如积分）都必须限制在该区间之内. 另外，该区间可以写成开区间，也可以写成闭区间或半开半闭区间. 不同的写法对题目的计算结果没有影响.

例 2.3.3 设随机变量 X 的概率密度为

$$f(x)=\begin{cases} x, & 0\leqslant x<1, \\ 2-x, & 1\leqslant x<2, \\ 0, & \text{其他}. \end{cases}$$

(1) 求 X 的分布函数 $F(x)$；(2) 求概率 $P(-1<X<1.5)$.

解 (1) 当 $x<0$ 时，$F(x)=0$.

当 $0\leqslant x<1$ 时，$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t=\int_{-\infty}^{0}f(t)\mathrm{d}t+\int_{0}^{x}f(t)\mathrm{d}t=\int_{0}^{x}t\mathrm{d}t=\dfrac{x^2}{2}$；

当 $1\leqslant x<2$ 时，

$$
\begin{aligned}
F(x)&=\int_{-\infty}^{x}f(t)\mathrm{d}t\\
&=\int_{-\infty}^{0}f(t)\mathrm{d}t+\int_{0}^{1}f(t)\mathrm{d}t+\int_{1}^{x}f(t)\mathrm{d}t\\
&=\int_{0}^{1}t\mathrm{d}t+\int_{1}^{x}(2-t)\mathrm{d}t\\
&=\frac{1}{2}+2x-\frac{x^2}{2}-\frac{3}{2}\\
&=-\frac{x^2}{2}+2x-1.
\end{aligned}
$$

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

故

$$
F(x)=\begin{cases}0, & x<0,\\ \dfrac{x^2}{2}, & 0\leqslant x<1,\\ -\dfrac{x^2}{2}+2x-1, & 1\leqslant x<2,\\ 1, & x\geqslant 2.\end{cases}
$$

(2) $P(-1<X<1.5)=\int_{0}^{1.5}f(x)\mathrm{d}x=\int_{0}^{1}x\mathrm{d}x+\int_{1}^{1.5}(2-x)\mathrm{d}x=\dfrac{7}{8}$，或者

$P(-1<X<1.5)=F(1.5)-F(-1)=\dfrac{7}{8}$.

由例 2.3.2 和例 2.3.3 可以知道，对于连续型随机变量，知道了分布函数就可以求得其概率密度函数；反之，知道了概率密度函数也可以得到分布函数. 但研究概率密度函数比研究分布函数要简单.

2.3.2 常见的连续型随机变量

我们知道，随机变量的密度函数决定了连续型随机变量的所有特征，下面用密度函数定义几种常见的连续型随机变量.

1. 均匀分布

定义 2.3.2 若随机变量 X 的密度函数为

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b,\\ 0, & 其他,\end{cases} \tag{2.3.5}$$

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

易知，$f(x)\geqslant 0$，且 $\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=1$. 又对于$[a,b]$的子区间$(c,d)$，

$$P(c<X\leqslant d)=\int_c^d f(x)\mathrm{d}x=\int_c^d \frac{1}{b-a}\mathrm{d}x=\frac{d-c}{b-a}.$$

这表明随机变量 X 落在(a,b)内任何等长度的子区间内的可能性都是相同的，而与子区间的位置无关. 如果从区间$[a,b]$上等可能地任取一点，记为 X，那么这个随机变量就在区间(a,b)上服从均匀分布.

均匀分布的情形在实际问题中经常可以见到，例如，在数值计算中由“四舍五入”最后一位数字引起的随机误差，在刻度器上把零头数化为整分度时发生的随机误差，在每隔一段时间有一辆公共汽车通过的汽车停车站上乘客候车的时间等.

容易得到服从均匀分布的随机变量的分布函数为

$$F(x)=\begin{cases}0, & x<a,\\ \dfrac{x-a}{b-a}, & a\leqslant x<b,\\ 1, & x\geqslant b.\end{cases} \tag{2.3.6}$$

均匀分布随机变量的密度函数 $f(x)$和分布函数 $F(x)$的图像如图 2-2 所示.

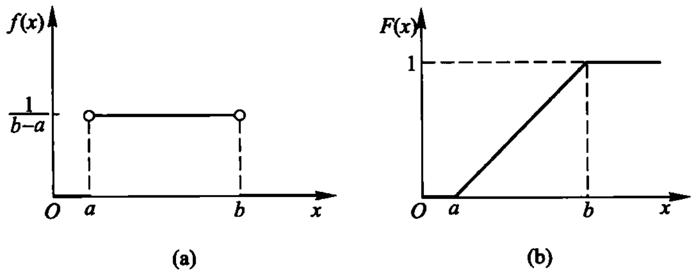


图 2-2

例 2.3.4　在区间$(-1,1)$上任取一点，记为 X，求方程 $t^2-3Xt+1=0$ 有实根的概率.

解　据题意知 X 在区间$(-1,1)$上服从均匀分布，其密度函数为

$$f(x)=\begin{cases}\dfrac{1}{2}, & -1<x<1,\\ 0, & \text{其他},\end{cases}$$

令判别式 $9X^2-4\geqslant 0$，得 $|X|\geqslant\dfrac{2}{3}$，即当 $X\leqslant -\dfrac{2}{3}$ 或 $X\geqslant\dfrac{2}{3}$ 时，方程 $t^2-3Xt+1=0$ 有实根，于是

$$P\left(|X|\geqslant\frac{2}{3}\right)=P\left(X\leqslant -\frac{2}{3}\right)+P\left(X\geqslant\frac{2}{3}\right)=\int_{-1}^{-\frac{2}{3}}\frac{1}{2}\mathrm{d}x+\int_{\frac{2}{3}}^{1}\frac{1}{2}\mathrm{d}x=\frac{1}{3}.$$

2. 指数分布

指数分布是较常见的一种连续型分布，在系统工程和可靠性理论中应用较广，通常用它描述电子元件的寿命及等待时间等指标.

定义 2.3.3 若随机变量 X 的密度函数为

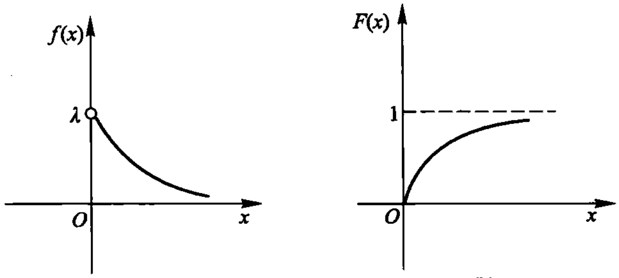
$$f(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}\tag{2.3.7}$$

其中，参数 $\lambda>0$，则称 X 服从参数为 λ 的指数分布，记为 $X\sim E(\lambda)$.

易得到 X 的分布函数为

$$F(x)=\begin{cases}1-\mathrm{e}^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}\tag{2.3.8}$$

密度函数和分布函数的图像如图 2 - 3 所示.



(a) (b)

图 2 - 3

例 2.3.5 假设一种投影仪的寿命 X(单位:h)服从参数 $\lambda=\dfrac{1}{2\,000}$ 的指数分布，(1) 任取一台这种投影仪，求其能正常使用 500 h 的概率；(2) 若一台投影仪已经使用了 500 h，求它至少还能使用 500 h 的概率.

解 (1) $P(X>500)=\displaystyle\int_{500}^{+\infty}\frac{1}{2\,000}\mathrm{e}^{-\frac{1}{2\,000}x}\mathrm{d}x=\mathrm{e}^{-\frac{1}{4}}$.

(2) 由条件概率计算公式，所求概率为

$$P(X>1\,000\mid X>500)=\frac{P(X>1\,000,X>500)}{P(X>500)}=\frac{P(X>1\,000)}{P(X>500)}$$
$$=\frac{e^{-\frac{1\,000}{2\,000}}}{e^{-\frac{500}{2\,000}}}=e^{-\frac{1}{4}}.$$

注意到 $P(X>500)=e^{-\frac{1}{4}}$，所以

$$P(X>1\,000\mid X>500)=P(X>500).$$

这个等式提示我们指数分布的一个重要性质——**“无记忆性”**，即若 $X\sim E(\lambda)$，则对任意 $t,s>0$ 有

$$P(X>t+s\mid X>s)=P(X>t). \tag{2.3.9}$$

一般来讲，指数随机变量往往用来描述寿命类的分布，或者用来描述等待时间的分布. 前面的几何分布也具有无记忆性，其实几何分布和指数分布描述的现象是类似的，几何分布可以看作是第一次命中的等待时间，而指数分布是等待某个事件首次发生的等待时间.

例 2.3.6　设顾客在某银行的窗口等待服务的时间 X(单位:min)服从参数 $\lambda=\frac{1}{5}$ 的指数分布. 某顾客在窗口等待服务，若超过 15 min 他就离开. 他一个月要去银行 4 次，以 Y 表示一个月内他未等到服务而离开窗口的次数. 求离散型随机变量 Y 的分布列.

解　随机变量 Y 的可能取值为 0,1,2,3,4，他每一次去银行就相当于做了一次独立试验，去了 4 次相当于做了 4 次独立重复试验，所以 Y 服从二项分布，每一次离开的概率为

$$p=P(X>15)=e^{-\frac{15}{5}}=e^{-3},$$

所以

$$Y\sim B(4,e^{-3}).$$

3. 正态分布

定义 2.3.4　若连续型随机变量 X 的密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},x\in\mathbf{R}, \tag{2.3.10}$$

其中，参数 $\mu\in\mathbf{R},\sigma>0$，则称 X 服从参数为 μ,σ^2 的正态分布，记为 $X\sim N(\mu,\sigma^2)$.

正态分布是概率论和数理统计中最常见也是最重要的分布之一. 正态分布也称为高斯分布，德国数学家高斯在研究误差理论时曾用它来描述误差. 我们生活中的很多指标，如身高、体重、学生成绩、炮弹落点的分布等都可以用它来描述. 正态分布具有许多其他分布都不具有的优良性质. 正态分布密度函数 $f(x)$ 的图像称为正态曲线，如图 2－4 所示.

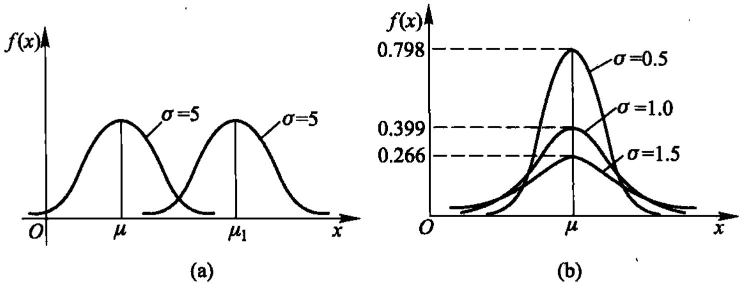


图 2-4

正态曲线 $y=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ 随着 μ 和 σ 的取值不同而变化. μ 决定了图像的中心位置,正态曲线以 $x=\mu$ 为对称轴. 当 σ 减小时,$f(x)$ 的最大值 $f(\mu)=\frac{1}{\sqrt{2\pi}\sigma}$ 增大,此时图像变得尖细,两侧尾部趋于 0 速度加快. 反之,当 σ 增大时,图像趋于扁平,两侧尾部趋于 0 速度变慢,因此,σ 称为形状参数.

正态分布的分布函数为

$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}\mathrm{d}t. \tag{2.3.11}$$

因为 $F(x)$ 不是初等函数,所以只能以积分的形式表示,这样就给计算分布函数值或事件概率带来困难. 不过所有的正态分布 $N(\mu,\sigma^2)$ 都可以经过一个简单的变换变成所谓的**标准正态分布**(参见定理 2.3.1). 于是,从理论上来讲,研究正态分布只需研究标准正态分布即可.

我们将 $\mu=0,\sigma^2=1$ 时的正态分布 $N(0,1)$ 称为标准正态分布. 记它的密度函数为

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \quad (x\in\mathbf{R}). \tag{2.3.12}$$

此时,它的分布函数为

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}\mathrm{d}t. \tag{2.3.13}$$

标准正态分布的密度函数和分布函数的图像如图 2-5 所示.

显然,标准正态分布的密度函数的图像关于 y 轴对称,且

$$\Phi(0)=\frac{1}{2},\quad \Phi(-x)=1-\Phi(x).$$

在本书的附录中,列有标准正态分布 $\Phi(x)$ 的函数表,常用数据可以从表中直接查到,例如 $\Phi(1.96)=0.975$. 由此可推算出 $\Phi(-1.96)=1-\Phi(1.96)=0.025$.

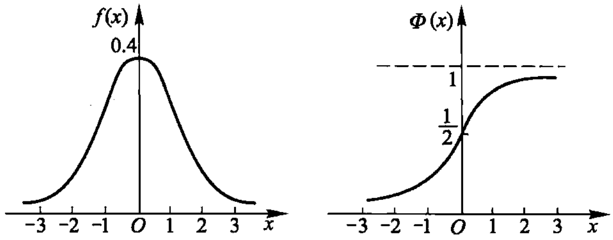


图 2-5

借助于分布函数值表，就可以通过公式

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

方便地计算服从标准正态分布的随机变量落入区间(a,b)的概率.

例 2.3.7 假设随机变量 $X\sim N(0,1)$，试求：(1) $P(X\leqslant 1.5)$；(2) $P(X>2.36)$；(3) $P(X\leqslant -2.1)$；(4) $P(-0.5<X\leqslant 0.8)$.

解 (1) 查表可知 $P(X\leqslant 1.5)=\Phi(1.5)=0.933\,2$.

(2) $P(X>2.36)=1-\Phi(2.36)=1-0.990\,9=0.009\,1$.

(3) $P(X\leqslant -2.1)=\Phi(-2.1)=1-\Phi(2.1)=1-0.982\,1=0.017\,9$.

(4)
$$\begin{aligned}P(-0.5<X\leqslant 0.8)&=\Phi(0.8)-\Phi(-0.5)\\&=\Phi(0.8)-[1-\Phi(0.5)]\\&=0.788\,1-1+0.691\,5\\&=0.479\,6.\end{aligned}$$

例 2.3.8 假设随机变量 $X\sim N(0,1)$，试求 a 与 b 的值：

(1) 使 $P(X>a)=0.05$；(2) 使 $P(X\leqslant b)=0.025$.

解 (1) 由 $P(X>a)=1-\Phi(a)=0.05$ 知

$$\Phi(a)=0.95,$$

查表知

$$a=1.65.$$

(2) 因为正态分布表中的函数值均不小于 0.5，因而欲求 b，使 $P(X\leqslant b)=\Phi(b)=0.025$，可转为求 $\Phi(-b)=1-\Phi(b)=0.975$，查表知

$$-b=1.96,$$

所以

$$b=-1.96.$$

标准正态分布与一般正态分布有什么关系呢？能否利用标准正态分布的一些性质解决一般正态分布的问题呢？回答是肯定的，对一般正态分布 $X\sim$

$N(\mu,\sigma^2)$，可以通过一个线性变换，将之化为标准正态分布.

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

证明 $Z=\frac{X-\mu}{\sigma}$的分布函数为

$$P(Z\leqslant z)=P\left(\frac{X-\mu}{\sigma}\leqslant z\right)=P(X\leqslant\sigma z+\mu)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\sigma z+\mu}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x\ ,$$

令$\frac{x-\mu}{\sigma}=t$，得

$$P(Z\leqslant z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t=\Phi(z)\ ,$$

可知

$$\frac{X-\mu}{\sigma}\sim N(0,1).$$

这样就可以通过查标准正态分布表来求一般的正态分布的概率了. 若 $X\sim N(\mu,\sigma^2)$，则

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

$$P(a<X\leqslant b)=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right). \tag{2.3.14}$$

例 2.3.9 假设随机变量 $X\sim N(1,4)$，试求：

(1) $P(X\leqslant2.5)$；(2) $P(0.5<X\leqslant2.4)$；(3) $P(X>1)$.

解 (1) $P(X\leqslant2.5)=F(2.5)=\Phi\left(\frac{2.5-1}{2}\right)=\Phi(0.75)=0.773\,4$.

(2) $$\begin{aligned}P(0.5<X\leqslant2.4)&=\Phi\left(\frac{2.4-1}{2}\right)-\Phi\left(\frac{0.5-1}{2}\right)\\&=\Phi(0.7)-\Phi(-0.25)\\&=\Phi(0.7)-[1-\Phi(0.25)]\\&=0.758\,0-1+0.598\,7\\&=0.356\,7.\end{aligned}$$

(3) $P(X>1)=1-P(X\leqslant1)=1-\Phi\left(\frac{1-1}{2}\right)=1-\frac{1}{2}=\frac{1}{2}$.

另外，对于一般的正态分布，我们有下面的结论：

$$P(|X-\mu|<\sigma)=2\Phi(1)-1=0.682\,6,$$
$$P(|X-\mu|<2\sigma)=2\Phi(2)-1=0.954\,4,$$
$$P(|X-\mu|<3\sigma)=2\Phi(3)-1=0.997\,4.$$

这组数值说明，服从正态分布的随机变量的取值范围虽然很广，是整个实数域，但又比较集中，落在以μ为中心3σ为半径的区间以外的概率不足0.3%，这就是著名的“3σ原则”(图2-6).

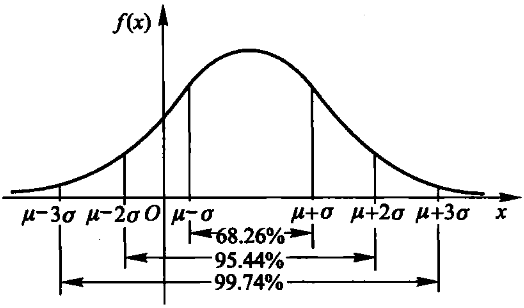


图2-6

例2.3.10 某单位招聘155人，按考试成绩录用，共有526人报名. 假设报名者的考试成绩$X\sim N(\mu,\sigma^2)$，已知90分以上的12人，60分以下的83人. 若从高分到低分依次录用，某人成绩为78分，问此人能否被录用？

解 需先根据题设，求出参数μ和σ，由题设知

$$P(X>90)=\frac{12}{526}=0.022\,8,\quad P(X<60)=\frac{83}{526}=0.157\,8,$$

于是

$$\Phi\left(\frac{90-\mu}{\sigma}\right)=1-0.022\,8=0.977\,2=\Phi(2),$$

$$\Phi\left(\frac{60-\mu}{\sigma}\right)=0.157\,8=\Phi(-1).$$

解方程组$\begin{cases}\dfrac{90-\mu}{\sigma}=2,\\[2mm] \dfrac{60-\mu}{\sigma}=-1,\end{cases}$ 得$\mu=70,\sigma=10$. 即

$$X\sim N(70,10^2).$$

设录用分数线为a，那么$P(X\geqslant a)=\dfrac{155}{526}=0.294\,7$，所以

$$\Phi\left(\frac{a-70}{10}\right)=1-0.294\,7=0.705\,3=\Phi(0.54).$$

解方程$\dfrac{a-70}{10}=0.54$，得$a=75.4$. 由于$78>75.4$，所以这个人能被录用.

*4. Γ函数与Γ分布

先给出 Γ 函数的定义.

定义 2.3.5　设 $\alpha>0$,由如下的广义积分定义的 α 的函数

$$\Gamma(\alpha)=\int_0^{\infty}x^{\alpha-1}\mathrm{e}^{-x}\mathrm{d}x \tag{2.3.15}$$

称为 **Γ(Gamma)函数**. 可以证明,在 $\alpha>0$ 时,积分是绝对收敛的.

对于一般的 $\alpha>0$,这个积分积不出来,但有以下性质:

(1) $\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$;　(2.3.16)

(2) $\Gamma(1)=\int_0^{\infty}\mathrm{e}^{-x}\mathrm{d}x=1,\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$.　(2.3.17)

利用这两个性质,可以得到

$$\Gamma(n+1)=\int_0^{\infty}x^n\mathrm{e}^{-x}\mathrm{d}x=n!. \tag{2.3.18}$$

定义 2.3.6　如果随机变量 X 的密度函数为

$$f(x)=\begin{cases}\dfrac{\beta^{\alpha}x^{\alpha-1}}{\Gamma(\alpha)}\mathrm{e}^{-\beta x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}\quad \alpha>0,\beta>0, \tag{2.3.19}$$

则称 X 服从参数为 α,β 的 **Γ 分布**. 当参数 $\alpha=1$ 时,Γ 分布就是指数分布.

数学家简介　高斯(Gauss,1777—1855),德国数学家、物理学家、天文学家和大地测量学家. 11 岁时发现了二项式定理,17 岁时发明了二次互反律,18 岁时发明了正 17 边形的尺规作图法,解决了 2 000 多年来悬而未决的难题,他也视此为生平得意之作,还交代要把正 17 边形刻在他的墓碑上. 他发现了质数分布定理、算术平均、几何平均. 21 岁大学毕业,22 岁获博士学位. 1804 年被选为英国皇家学会会员. 高斯有"数学王子"、"数学家之王"的美称,被认为是人类有史以来"最伟大的四位数学家之一"(阿基米德、牛顿、高斯、欧拉). 人们还称赞高斯是"人类的骄傲". 天才、早熟、高产、创造力不衰、……,人类智力领域的几乎所有褒奖之词,对于高斯都不过分.

高斯的数学研究遍及数学的所有领域,在数论、代数学、非欧几何、复变函数和微分几何等方面都做出了开创性的贡献. 高斯开辟了许多新的数学领域,从最抽象的代数数论到内蕴几何学,都留下了他的足迹. 从研究风格、方法乃至所取得的具体成就方面,他都是 18—19 世纪之交的中坚人物. 如果我们把 18 世纪的数学家想象为一系列的高山峻岭,那么最后一个令人肃然起敬的巅峰就是高斯;如果把 19 世纪的数学家想象为一条条江河,那么其源头就是高斯. 图 2-7 为印在德国马克上的高斯头像.

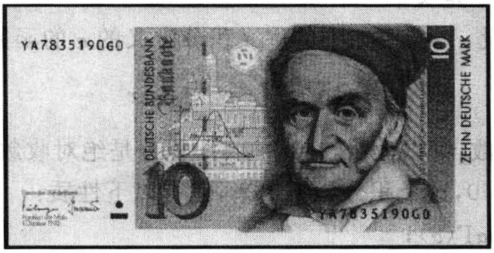


图 2-7

习 题

1. 已知随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<1,\\ \ln x, & 1\leqslant x<\mathrm{e},\\ 1, & x\geqslant \mathrm{e}.\end{cases}$$

(1) 求 $P(X\leqslant 2)$,$P(0<X\leqslant 3)$,$P(X>\sqrt{\mathrm{e}})$;(2) 求密度函数 $f(x)$.

2. 设连续型随机变量 X 的分布函数为

$$F(x)=\begin{cases}a+b\mathrm{e}^{-x^2}, & x\geqslant 0,\\ 0, & x<0,\end{cases}$$

(1) 求常数 a 和 b;(2) 求随机变量 X 的概率密度函数.

3. 设随机变量 X 的密度函数为

$$f(x)=\frac{1}{2}\mathrm{e}^{-|x|},\quad -\infty<x<\infty,$$

求随机变量 X 的分布函数.

4. 设随机变量 X 在区间[0,2]内取值,且对于每个 $a\in[0,2]$,X 落入$[0,a]$内的概率与 a^2 成正比.试求:(1) X 的分布函数;(2) X 的密度函数;(3) X 落在[0,1]内的概率.

5. 设随机变量 $Y\sim U(a,5)$,且方程 $x^2+Yx+\frac{3Y}{4}+1=0$ 没有实根的概率为 0.25,试求常数 a.

6. 设随机变量 X 的密度函数为

$$f(x)=\begin{cases}kx^2-x+1, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$$

(1) 求常数 k;(2) 求随机变量 X 的分布函数.

7. 某种电脑显示器的使用寿命 X(单位:1 000 h)服从参数为 $\lambda=0.02$ 的指数分布.生产厂家承诺:购买者使用一年内显示器损坏将免费予以更换.

(1) 假设用户一般每年使用电脑 2 000 h,求厂家免费更换显示器的概率;

(2) 显示器至少可以使用 10 000 h 的概率是多少?

(3) 已知某台显示器已经使用了 10 000 h,求其至少还能使用 10 000 h 的概率.

8. 设 $X \sim N(0,1)$,查表求下列概率:$P(X<2.2)$,$P(X>1.76)$,$P(X<-1.79)$,$P(|X|<1.55)$.

9. 假设随机变量 $X \sim N(108,9)$,试求:

(1) 常数 a,使 $P(X \leqslant a)=0.9$;

(2) 常数 b,使 $P(|X-b|>b)=0.1$.

10. 若 $X \sim N(2,\sigma^2)$,且 $P(2<X<4)=0.3$,求 $P(X<0)$.

11. 设随机变量 X 在$[2,5]$上服从均匀分布.现对 X 进行三次独立观测,求至少有两次的观测值大于 3 的概率.

12. 某人乘汽车去火车站乘火车,有两条路可走.第一条路程较短但交通拥挤,所需时间 X 服从 $N(40,10^2)$(单位:min);第二条路程较长,但阻塞少,所需时间 X 服从 $N(50,4^2)$.

(1) 若动身时离火车开车只有 1 h,问应走哪条路能乘上火车的把握大些?

(2) 又若离火车开车时间只有 45 min,问应走哪条路赶上火车把握大些?

13. 设成年男子的身高服从正态分布 $N(170,10^2)$(单位:cm).

(1) 求成年男子身高大于 160 cm 的概率;

(2) 公共汽车的车门应设计多高,才能使成年男子上车时碰头的概率不大于 5%?

(3) 在这个设计之下,求 100 个成年男子上车时,至少有 2 个人碰头的概率.

§ 2.4 随机变量函数的分布

对于很多问题需要考虑随机变量函数的分布.例如,在一些试验中,所关心的随机变量往往不能直接测量到,而它是某个能直接测量的随机变量的函数.看下面的问题:

(1) 在交流电中,相角 θ 是一个随机变量,$\theta \sim U\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$,求电压 $V=A\sin\theta$ 的分布,其中 A 是一个已知的正常数.

(2) 通过(0,1)点的任意直线与 x 轴的夹角 θ 是一个随机变量,$\theta \sim U(0,\pi)$,直线在 x 轴的截距 X 是关于 θ 的函数,$X=-\dfrac{1}{\tan\theta}=-\cot\theta$,求 X 的分布.

现在我们把这类问题统一描述一下:设 X 为一个随机变量,分布已知.$g(x)$ 为一个一元函数,令 $Y=g(X)$.由于 X 是随机变量,因此 Y 也是一个随机变量.一般来讲,若 X 是离散型随机变量,Y 也是离散型随机变量;若 X 是连续型随机变量,Y 也是连续型随机变量.现在的问题是,如何由已知随机变量 X 的分布,去求它的函数 $Y=g(X)$ 的概率分布.下面就 X 是离散型随机变量和连续型随机变量两种情况,分别加以讨论.

2.4.1　离散型随机变量函数的分布列

若 X 是离散型随机变量，其分布列为

$$X\sim\begin{pmatrix} x_1 & x_2 & x_3 & \cdots \\ p_1 & p_2 & p_3 & \cdots \end{pmatrix},$$

$Y=g(X)$，显然 Y 也是离散型随机变量，则 Y 的分布列为

$$Y\sim\begin{pmatrix} g(x_1) & g(x_2) & g(x_3) & \cdots \\ p_1 & p_2 & p_3 & \cdots \end{pmatrix}.$$

这里需要注意的是，若 $g(x_i)$ 的值中有相等的，则应把那些相等的值分别合并，同时把对应的概率 p_i 相加.

例 2.4.1　已知 X 的分布列为 $X\sim\begin{pmatrix} -1 & 0 & 1 & 2 \\ 0.1 & 0.2 & 0.3 & 0.4, \end{pmatrix}$，求：(1) $Y=2X+1$的分布律；(2) $Y=X^2$ 的分布律.

解　由 X 的分布列可得

X	-1	0	1	2
$Y=2X+1$	-1	1	3	5
$Y=X^2$	1	0	1	4
P	0.1	0.2	0.3	0.4

于是 Y 的分布列为

$$(1)\ Y\sim\begin{pmatrix} -1 & 1 & 3 & 5 \\ 0.1 & 0.2 & 0.3 & 0.4 \end{pmatrix};\ (2)\ Y\sim\begin{pmatrix} 0 & 1 & 4 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}.$$

2.4.2　连续型随机变量函数的分布

若 X 为连续型随机变量，已知 X 的密度函数为 $f(x)$，$Y=g(X)$，若 Y 也为连续型随机变量，则求 Y 的密度函数的一般步骤为：

(1) 根据 $Y=g(X)$，由 X 的取值范围，确定 Y 的取值范围 D.

(2) 在 D 内求 Y 的分布函数 $F_Y(y)$，此分布函数是一个积分函数形式.

(3) 对 Y 的分布函数 $F_Y(y)$ 求导，即得 Y 的密度函数.

例 2.4.2　设连续型随机变量 X 的密度函数为

$$f(x)=\begin{cases} 2x, & 0<x<1, \\ 0, & \text{其他}, \end{cases}$$

求随机变量 $Y=X^2+1$ 的密度函数.

解 Y 的取值范围是 $[1,2]$，所以当 $y<1$ 时，$F_Y(y)=0$；当 $y\geqslant 2$ 时，$F_Y(y)=1$；当 $1\leqslant y<2$ 时，

$$F_Y(y)=P(Y\leqslant y)=P(X^2+1\leqslant y)=P(-\sqrt{y-1}\leqslant X\leqslant \sqrt{y-1}).$$
$$=\int_0^{\sqrt{y-1}} 2x\mathrm{d}x=y-1.$$

将 $F_Y(y)$ 对 y 求导，得 Y 的密度函数为 $f_Y(y)=\begin{cases}1, & 1<y<2,\\ 0, & \text{其他}.\end{cases}$

例 2.4.3 设随机变量 X 服从区间 $[-1,2]$ 上的均匀分布，求 $Y=X^2$ 的密度函数.

解 随机变量 Y 的取值范围为 $[0,4]$，所以，当 $y<0$ 时，$F_Y(y)=0$；当 $y\geqslant 4$ 时，$F_Y(y)=1$；

当 $0\leqslant y<1$ 时，

$$F_Y(y)=P(Y\leqslant y)=P(X^2\leqslant y)=P(-\sqrt{y}\leqslant X\leqslant \sqrt{y})$$
$$=\int_{-\sqrt{y}}^{\sqrt{y}}\frac{1}{3}\mathrm{d}x=\frac{2}{3}\sqrt{y}\ ;$$

当 $1\leqslant y<4$ 时，

$$F_Y(y)=P(Y\leqslant y)=P(X^2\leqslant y)=P(-\sqrt{y}\leqslant X\leqslant \sqrt{y})$$
$$=\int_{-1}^{\sqrt{y}}\frac{1}{3}\mathrm{d}x=\frac{1}{3}(\sqrt{y}+1)\ .$$

所以随机变量 Y 的密度函数为

$$f_Y(y)=\begin{cases}\dfrac{1}{3\sqrt{y}}, & 0<y<1,\\ \dfrac{1}{6\sqrt{y}}, & 1<y<4,\\ 0, & \text{其他}.\end{cases}$$

例 2.4.4 设随机变量 X 服从参数为 4 的指数分布，试求 $Y=\min\{2,X\}$ 的分布函数.

解 随机变量 Y 的取值范围为 $[0,2]$，所以，当 $y<0$ 时，$F_Y(y)=0$；当 $y\geqslant 2$ 时，$F_Y(y)=1$；当 $0\leqslant y<2$ 时，

$$F_Y(y)=P(Y\leqslant y)=P(\min\{2,X\}\leqslant y).$$

注意到 $0\leqslant y<2$ 时，$\{\min\{2,X\}\leqslant y\}=\{X\leqslant y\}$，所以当 $0\leqslant y<2$ 时，$F_Y(y)=P(X\leqslant y)=\int_0^y\frac{1}{4}\mathrm{e}^{-\frac{x}{4}}\mathrm{d}x=1-\mathrm{e}^{-\frac{y}{4}}$. 于是 Y 的分布函数为

$$F_Y(y)=\begin{cases}0, & y<0,\\ 1-\mathrm{e}^{-\frac{y}{4}}, & 0\leqslant y<2,\\ 1, & y\geqslant 2.\end{cases}$$

注意，对本题来讲，随机变量 Y 不是连续型随机变量，当然也不是离散型随机变量. 我们将此类随机变量称为混合型随机变量，混合型随机变量既没有密度函数，也没有分布列，但是肯定有分布函数，所以对这类随机变量我们只能通过分布函数去研究.

例 2.4.5 设随机变量 $X\sim N(\mu,\sigma^2)$，试求 $Y=\mathrm{e}^X$ 的密度函数.

解 显然随机变量 Y 的取值范围为 $[0,+\infty)$，所以，当 $y<0$ 时，$F_Y(y)=0$；当 $y\geqslant 0$ 时，

$$\begin{aligned}F_Y(y)&=P(Y\leqslant y)=P(\mathrm{e}^X\leqslant y)=P(X\leqslant \ln y)\\&=\int_{-\infty}^{\ln y}\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]\mathrm{d}x,\end{aligned}$$

两边求导可得 Y 的密度函数为

$$f_Y(y)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\sigma y}\exp\left[-\dfrac{(\ln y-\mu)^2}{2\sigma^2}\right], & y>0,\\ 0, & \text{其他}.\end{cases}$$

今后，我们称形如例 2.4.5 中 Y 的密度函数为对数正态的密度函数. 如果 $y=g(x)$ 是一个严格单调且有一阶连续导数的函数，则随机变量的函数 $Y=g(X)$ 的密度函数有如下结论：

定理 设连续型随机变量 X 的密度函数为 $f_X(x)$，$y=g(x)$ 是一严格单调函数，且具有一阶连续导数，$x=h(y)$ 是 $y=g(x)$ 的反函数，则 $Y=g(X)$ 的密度函数为

$$f_Y(y)=f_X(h(y))\,|h'(y)|.$$

证明 当 $g(x)$ 为单调增加函数时，可得 $h'(y)>0$，且

$$\begin{aligned}F_Y(y)&=P(Y\leqslant y)=P(g(X)\leqslant y)\\&=P(X\leqslant h(y))\\&=\int_{-\infty}^{h(y)}f_X(x)\mathrm{d}x.\end{aligned}$$

对上式求导，得到 Y 的密度函数为 $f_Y(y)=f_X(h(y))h'(y)$.

当 $g(x)$ 为单调递减函数时，可得 $h'(y)<0$，且

$$\begin{aligned}F_Y(y)&=P(Y\leqslant y)=P(g(X)\leqslant y)\\&=P(X\geqslant h(y))\end{aligned}$$

$$=\int_{h(y)}^{+\infty} f_X(x)\mathrm{d}x.$$

对 y 求导，得 Y 的密度函数为 $f_Y(y)=f_X(h(y))(-h'(y))$.

综上，$f_Y(y)=f_X(h(y))\,|h'(y)|$.

例 2.4.6 设随机变量 $X\sim N(\mu,\sigma^2)$. $Y=aX+b(a\neq 0)$，则

$$Y\sim N(a\mu+b,(a\sigma)^2).$$

证明 $y=ax+b$ 为一个单调函数，且具有一阶连续导数，解得 $x=h(y)=\frac{y-b}{a}$，$h'(y)=\frac{1}{a}$. 由定理 2.4.1 得

$$f_Y(y)=f_X(h(y))\,|h'(y)|=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}}\left|\frac{1}{a}\right|$$

$$=\frac{1}{\sqrt{2\pi}\,|a|\,\sigma}\mathrm{e}^{-\frac{[y-(a\mu+b)]^2}{2(a\sigma)^2}}\quad(-\infty<y<+\infty),$$

因此 $Y\sim N(a\mu+b,(a\sigma)^2)$.

本例表明，服从正态分布的随机变量的线性函数也服从正态分布.

习 题

1. 已知 X 的分布列为 $P(X=k)=2^{-k}$，$k=1,2,\cdots$. 求 $Y=\sin\frac{\pi}{2}X$ 的分布列.

2. 设随机变量 X 服从参数 $\lambda=1$ 的泊松分布，记随机变量

$$Y=\begin{cases}0, & X\leqslant 1,\\ 1, & X>1,\end{cases}$$

试求随机变量 Y 的分布列.

3. 设 X 的密度函数为

$$f(x)=\begin{cases}3x^2, & 0<x<1,\\ 0, & \text{其他},\end{cases}$$

求以下随机变量的密度函数：(1) $2X$；(2) $-X+1$；(3) X^2.

4. 设随机变量 X 服从区间[0,1]上的均匀分布.

(1) 求随机变量 $Y=\mathrm{e}^{2X}$ 的密度函数；

(2) 求随机变量 $Y=-\ln X$ 的密度函数.

5. 设随机变量 $X\sim N(0,1)$，

(1) 求随机变量 $Y=\mathrm{e}^X$ 的密度函数；

(2) 求随机变量 $Y=X^2+1$ 的密度函数；

(3) 求随机变量 $Y=|X|$ 的密度函数.

6. 设随机变量 X 服从参数为 $\lambda=2$ 的指数分布，求 $Y=1-\mathrm{e}^{-2X}$ 的密度函数.

7. 设随机变量 X 的密度函数为

$$f(x)=\frac{1}{\pi(1+x^2)},\quad -\infty<x<+\infty,$$

求 $Y=\tan X$ 的密度函数.

§2.5 综合例题

例 2.5.1 设随机变量 X 服从参数为λ 的泊松分布,试求 X 取偶数的概率.

解 由于

$$P(X\text{ 为偶数})=\sum_{k=0}^{\infty}P(X=2k)=\sum_{k=0}^{\infty}\frac{\lambda^{2k}}{(2k)!}\mathrm{e}^{-\lambda},$$

$$P(X\text{ 为奇数})=\sum_{k=0}^{\infty}P(X=2k+1)=\sum_{k=0}^{\infty}\frac{\lambda^{2k+1}}{(2k+1)!}\mathrm{e}^{-\lambda},$$

所以

$$\begin{aligned}P(X\text{ 为偶数})-P(X\text{ 为奇数})&=\sum_{k=0}^{\infty}\frac{\lambda^{2k}}{(2k)!}\mathrm{e}^{-\lambda}-\sum_{k=0}^{\infty}\frac{\lambda^{2k+1}}{(2k+1)!}\mathrm{e}^{-\lambda}\\&=\sum_{k=0}^{\infty}\frac{(-\lambda)^k}{(k)!}\mathrm{e}^{-\lambda}\\&=\mathrm{e}^{-2\lambda}.\end{aligned}$$

又 $P(X$ 为偶数$)+P(X$ 为奇数$)=1$,所以

$$P(X\text{ 为偶数})=\frac{1}{2}(1+\mathrm{e}^{-2\lambda}).$$

例 2.5.2 设随机变量 X 的绝对值不大于 1,且 $P(X=-1)=\frac{1}{8}$,$P(X=1)=\frac{1}{4}$. 而在$\{-1<X<1\}$出现的条件下,X 在$(-1,1)$内任一子区间上取值的条件概率与该子区间长度成正比,试求 X 的分布函数 $F(x)$.

解 显然当 $x<-1$ 时 $F(x)=0$,而 $x\geqslant1$ 时 $F(x)=1$. 由题意可知,对 $x\in(-1,1)$,

$$P(-1<X\leqslant x\mid-1<X<1)=k(x+1),$$

其中 k 为比例系数. 注意到 $P(-1<X<1\mid-1<X<1)=k(1+1)=1$,所以 $k=\frac{1}{2}$. 又 $P(-1<X<1)=1-\frac{1}{8}-\frac{1}{4}=\frac{5}{8}$,于是可得对 $x\in(-1,1)$,

$$\begin{aligned}F(x)&=P(X\leqslant x)=P(X\leqslant-1)+P(-1<X\leqslant x)\\&=P(X=-1)+\frac{P(-1<X\leqslant x)}{P(-1<X<1)}P(-1<X<1)\end{aligned}$$

$$=\frac{1}{8}+\frac{5}{16}(x+1).$$

所以 X 的分布函数为

$$F(x)=\begin{cases}0, & x<-1,\\ \frac{5}{16}(x+1)+\frac{1}{8}, & -1\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$$

例 2.5.3 假设一大型设备在任何长为 t 的时间内发生故障的次数 $N(t)$ 服从参数为 λt 的泊松分布.

(1) 求相继两次故障之间时间间隔 T 的概率密度;

(2) 求在设备已经无故障工作 8 h 的情形下,再无故障运行 8 h 的概率 Q.

解 (1) 当 $t<0$ 时,$F_T(t)=P(T\leqslant t)=0$. 当 $t\geqslant 0$ 时,事件 $\{T>t\}$ 与 $\{N(t)=0\}$ 等价,因为 $\{T>t\}$ 表示从上次故障之后的长为 t 的时间内没有发生故障,即是说在这个长为 t 的时间内发生故障次数 $N(t)=0$. 所以

$$F_T(t)=P(T\leqslant t)=1-P(T>t)=1-P(N(t)=0)=1-\mathrm{e}^{-\lambda t},$$

即

$$F_T(t)=\begin{cases}1-\mathrm{e}^{-\lambda t}, & t\geqslant 0,\\ 0, & t<0,\end{cases}$$

即间隔时间 T 服从参数为 λ 的指数分布.

(2) 所求概率为 $Q=P(T>16\mid T>8)=P(T>16)/P(T>8)=\dfrac{\mathrm{e}^{-16\lambda}}{\mathrm{e}^{-8\lambda}}=\mathrm{e}^{-8\lambda}$.

这实际上是指数分布的无记忆性.

例 2.5.4 设 $F(x)$ 为连续型随机变量 X 的密度函数,且严格单调增加,求 $Y=F(X)$ 的密度函数.

解 由于 $F(x)$ 为分布函数,所以随机变量 Y 的取值范围为$[0,1]$.

当 $y<0$ 时,$F_Y(y)=0$;当 $y\geqslant 1$ 时,$F_Y(y)=1$; 当 $0\leqslant y<1$ 时,由于 $F(x)$ 严格单调增加,所以 $F(x)$ 的反函数存在,于是

$$F_Y(y)=P(Y\leqslant y)=P(F(X)\leqslant y)=P(X\leqslant F^{-1}(y))=F(F^{-1}(y))=y.$$

所以 Y 的密度函数为

$$f_Y(y)=\begin{cases}1, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

例 2.5.5 设随机变量 X 服从区间$[0,1]$上的均匀分布,试求一个严格单调增加的函数 $g(x)$,使得 $Y=g(X)$ 服从参数为 2 的指数分布.

解 由于 Y 服从指数分布,所以当 $y\geqslant 0$ 时,

$$F_Y(y) = P(Y \leqslant y) = \int_0^y 2\mathrm{e}^{-2x}\mathrm{d}x = 1 - \mathrm{e}^{-2y},$$

又

$$F_Y(y)=P(Y\leqslant y)=P(g(X)\leqslant y)$$
$$= P(X \leqslant g^{-1}(y)) = \int_0^{g^{-1}(y)} 1\mathrm{d}x = g^{-1}(y).$$

所以

$$g^{-1}(y)=1-\mathrm{e}^{-2y}.$$

解此函数方程可得

$$g(x)=-\frac{1}{2}\ln(1-x),\quad 0<x<1.$$

复 习 题 2

1. 进行某种试验,成功的概率为$\frac{3}{4}$,失败的概率为$\frac{1}{4}$. 以 X 表示试验首次成功所需试验的次数,试写出 X 的分布列,并计算 X 取偶数的概率.

2. 设随机变量 $X\sim B(n,p)$,求 X 取偶数的概率.

3. 设三次独立试验中,事件 A 出现的概率相等.若已知 A 至少出现一次的概率为$\frac{19}{27}$,求 A 在一次试验中出现的概率.

4. 甲、乙两人独立地轮流投篮,直至有一人投中为止.甲先投,已知甲、乙的命中率分别为 0.4,0.5.以 X,Y 表示甲、乙的投篮次数,求 X,Y 的分布列.

5. 已知

$$F(x)=\begin{cases} 0, & x<0, \\ x+\frac{1}{2}, & 0\leqslant x<\frac{1}{2}, \\ 1, & x\geqslant\frac{1}{2}. \end{cases}$$

(1) 证明 $F(x)$是一个分布函数;(2) 如果 $F(x)$是随机变量 X 的分布函数,那么随机变量 X 是离散型随机变量还是连续型随机变量? 为什么?

6. 设随机变量 $X\sim N(\mu,\sigma^2)$,问:当 σ 取何值时,X 落入区间$(\mathrm{e},\mathrm{e}^2)$的概率最大?

7. 设随机变量 X 服从参数为 $\ln 3$ 的指数分布,求 $Y=[X]+1$ 的分布列.其中$[x]$为向下取整函数.

8. 如果随机变量 X 与 $-X$ 具有相同的分布函数,称随机变量 X 是对称的.如果 X 是对称的,证明 X 的分布函数满足 $F(x)+F(-x)=1$,且如果 X 又是连续型随机变量,那么的密度函数必为偶函数.

9. 设在一段时间内进入某一商店的顾客人数 X 服从参数为 λ 的泊松分布,每个顾客购

买某种物品的概率为 p,并且各个顾客是否购买该种物品相互独立,求进入商店的顾客购买这种物品的人数 Y 的分布列.

10. 某地抽样调查结果表明,考生的外语成绩(百分制)近似服从正态分布,平均成绩为72分,96分以上的占考生总数的2.3%,试求考生的外语成绩在60分至84分之间的概率.

第3章　二维随机变量及其分布

在实际问题中，除了讨论一个随机变量的情形，许多随机现象需要用两个或者更多个随机变量来描述，这就需要我们研究多维随机变量的问题. 多维随机变量的性质不仅与每一个随机变量有关，而且还与它们之间的相互联系有关，研究多维随机变量不仅要研究各个随机变量的性质，而且还要研究他们之间的联系. 为了简明起见，本章只介绍二维随机变量. 二维随机变量是一维随机变量的延伸，与一维随机变量相比情况要复杂得多.

本章首先介绍二维随机变量的概念，以及二维随机变量的联合分布函数. 类似于第二章的内容，本章中只涉及两类最重要的随机变量：二维离散型随机变量和二维连续型随机变量.

在二维离散型随机变量内容中，主要介绍二维随机变量的联合分布律. 在二维连续型随机变量内容中，主要介绍二维随机变量的联合密度函数，最后介绍二维随机变量的函数的分布. 把一维随机变量推广到二维随机变量会产生一些新问题，把二维随机变量推广到三维以至于 n 维，很多研究可以类推，结果可以自然推广.

§3.1　二维随机变量的联合分布与边际分布

3.1.1　二维随机变量的联合分布函数及其性质

设 X 和 Y 为两个随机变量，则称有序数组 (X,Y) 为**二维随机变量**.

一维随机变量就是直线 $\mathbf{R}$ 上的一个随机点，那么二维随机变量就是平面 $\mathbf{R}^2$ 上的一个二维的随机点. 在研究一维随机变量的时候，我们已经知道，一维随机变量的所有概率特性完全由它的分布函数决定. 二维随机变量的研究与一维随机变量非常类似，它的所有概率特性完全由它的联合分布函数决定.

定义　设二维随机变量 (X,Y)，对任意实数 x,y，二元函数

$$F(x,y)=P(X\leqslant x,Y\leqslant y),\tag{3.1.1}$$

称为 (X,Y) 的**联合分布函数**. 其中 $\{X\leqslant x,Y\leqslant y\}$ 为 $\{X\leqslant x\}\cap\{Y\leqslant y\}$ 的简写形式.

从几何上看，(X,Y) 表示平面直角坐标系中随机点的坐标，设 (x,y) 表示坐标系中的任一点，那么分布函数 $F(x,y)$ 在 (x,y) 处的函数值表示随机点落在以

(x,y)为顶点的左下方无穷矩形域上的概率(图 3－1).

由分布函数的几何意义可以得出，对任何 $x_1 \leqslant x_2, y_1 \leqslant y_2$，有

$$P(x_1 < X \leqslant x_2, y_1 < Y \leqslant y_2) = F(x_2,y_2) - F(x_1,y_2) - F(x_2,y_1) + F(x_1,y_1),$$

它表示随机点落在区域 D 内的概率(图 3－2).

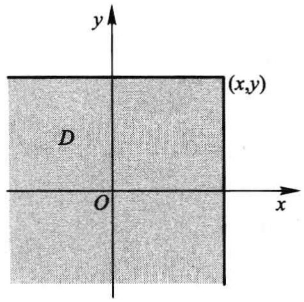


图 3－1

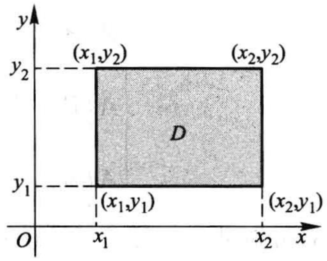


图 3－2

注意：$F(x,y)$是一个普通的二元函数，其定义域为 $\mathbf{R}^2$.

可以证明分布函数具有以下的性质：

(1) $F(x,y)$对 x 或 y 都是不减函数，即对任意 y，若 $x_1 \leqslant x_2$，则 $F(x_1,y) \leqslant F(x_2,y)$，对任意 x，若 $y_1 \leqslant y_2$，则 $F(x,y_1) \leqslant F(x,y_2)$.

(2) 对任意的 x,y，

$$F(-\infty,y)=0,\quad F(x,-\infty)=0,$$
$$F(-\infty,-\infty)=0,\quad F(+\infty,+\infty)=1.$$

(3) $F(x,y)$分别对 x,y 右连续，即有

$$F(x+0,y)=F(x,y),\quad F(x,y+0)=F(x,y).$$

(4) (矩形法则)对任何 $x_1 \leqslant x_2, y_1 \leqslant y_2$，有

$$F(x_2,y_2) - F(x_1,y_2) - F(x_2,y_1) + F(x_1,y_1) \geqslant 0. \tag{3.1.2}$$

例 3.1.1　已知二元函数 $F(x,y)=\begin{cases}1, & 2x+y \geqslant 0,\\ 0, & 2x+y<0,\end{cases}$ 证明 $F(x,y)$不是联合分布函数.

证明　显然 $F(x,y)$对 x 或 y 都是右连续的不减函数，且对任意的 x,y，

$$F(-\infty,y)=0,\quad F(x,-\infty)=0,$$
$$F(-\infty,-\infty)=0,\quad F(+\infty,+\infty)=1.$$

但是，$F(x,y)$不满足矩形法则，比如取 $x_1=-1, x_2=1, y_1=0, y_2=3$，则有

$$F(x_2,y_2) - F(x_2,y_1) - F(x_1,y_2) + F(x_1,y_1)$$

$$= F(1,3) - F(1,0) - F(-1,3) + F(-1,0) = -1 < 0,$$

所以 $F(x,y)$ 不是分布函数.

例 3.1.2　已知 $F(x,y)=A(B+\arctan x)(C+\arctan y)(x,y\in\mathbf{R})$ 为二维随机变量 (X,Y) 的联合分布函数，求常数 A,B,C.

解　由联合分布函数的性质 $F(-\infty,y)=0, F(x,-\infty)=0, F(+\infty,+\infty)=1$，得到联立方程组

$$\begin{cases} A(B+\arctan x)\left(C-\dfrac{\pi}{2}\right)=0, & (1)\\ A\left(B-\dfrac{\pi}{2}\right)(C+\arctan y)=0, & (2)\\ A\left(B+\dfrac{\pi}{2}\right)\left(C+\dfrac{\pi}{2}\right)=1, & (3)\end{cases}$$

由方程(3)可知 $A\neq 0$，所以对任意的 x 都有

$$(B+\arctan x)\left(C-\frac{\pi}{2}\right)=0,$$

于是 $C=\dfrac{\pi}{2}$，同理可得 $B=\dfrac{\pi}{2}$，所以 $A=\dfrac{1}{\pi^2}$.

3.1.2　边际分布函数

在研究二维随机变量 (X,Y) 的过程中，为了避免将单个随机变量 X 或 Y 的分布与联合分布混为一谈，往往将单个随机变量 X 或 Y 的分布称为边际分布，即 X 的分布函数 $F_X(x)$ 称为 X 的**边际分布函数**，Y 的分布函数 $F_Y(y)$ 称为 Y 的**边际分布函数**. (X,Y) 的联合分布函数与 X,Y 的边际分布函数有如下关系

$$F_X(x)=F(x,+\infty),\quad F_Y(y)=F(+\infty,y). \tag{3.1.3}$$

比如，在例 3.1.2 中 X 的边际分布函数为

$$F_X(x)=F(x,+\infty)=\frac{1}{\pi}\left(\frac{\pi}{2}+\arctan x\right),\quad x\in\mathbf{R}.$$

Y 的边际分布函数为

$$F_Y(y)=F(+\infty,y)=\frac{1}{\pi}\left(\frac{\pi}{2}+\arctan y\right),\quad y\in\mathbf{R}.$$

读者需要注意的是，由联合分布函数必然可求出边际分布函数，但是即使两个边际分布函数都已知，一般来说也是求不出联合分布函数的.

虽然二维随机变量的概率性质完全由它的联合分布函数决定，但是联合分

布函数非常不便于实际应用,它只在理论上比较有用.因此,在处理实际问题时,我们通常只考虑二维离散与二维连续两类随机变量.

习 题

1. 设二维随机变量(X,Y)的联合分布函数为

$$F(x,y)=\begin{cases}\sin x\sin y, & 0\leqslant x\leqslant\frac{\pi}{2},0\leqslant y\leqslant\frac{\pi}{2},\\ 0, & \text{其他}.\end{cases}$$

(1) 求边际分布函数 $F_X(x)$ 与 $F_Y(y)$;(2) 求 $P\left(0<X\leqslant\frac{\pi}{4},\frac{\pi}{6}<Y\leqslant\frac{\pi}{3}\right)$.

2. 已知二元函数 $F(x,y)=\begin{cases}1, & x+3y\geqslant 0,\\ 0, & x+3y<0,\end{cases}$ $F(x,y)$能否成为一个联合分布函数?

§3.2 二维离散型随机变量

如果 X 与 Y 是两个一维离散型随机变量,称(X,Y)为**二维离散型随机变量**.由于 X 与 Y 的可能取值都是至多可数个,所以(X,Y)的可能取值也是至多可数个.我们已经知道一维离散型随机变量的概率特性是由它的分布列来决定的,类似地,二维离散型随机变量的概率特性由它的联合分布列来决定.

定义 3.2.1 设(X,Y)为二维离散型随机变量,且 X 的可能取值记为 $x_1,x_2,\cdots,Y$ 的可能取值记为 $y_1,y_2,\cdots$,称

$$p_{ij}=P(X=x_i,Y=y_j),\quad i,j=1,2,\cdots \tag{3.2.1}$$

为二维离散型随机变量(X,Y)的**联合分布列**或**联合分布律**.

读者容易验证,p_{ij}具有以下性质:

(1) 非负性:$p_{ij}\geqslant 0(i,j=1,2,\cdots)$;

(2) 归一性:$\sum\limits_i\sum\limits_j p_{ij}=1$.

(X,Y)的联合分布列与(X,Y)的联合分布函数二者是等价的:已知(X,Y)的联合分布列为 $p_{ij}=P(X=x_i,Y=y_j),i,j=1,2,\cdots$,可以求出$(X,Y)$的联合分布函数为

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

反之,由联合分布函数也可以求出联合分布列.对于二维离散型随机变量的联合分布列,使用下面的表格会更加方便.

$X \backslash Y$	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}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	

3.2.1 离散型随机变量的边际分布

设(X,Y)为二维离散型随机变量,其联合分布律为

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

则称

$$P(X=x_i)=P(X=x_i,Y<+\infty),\quad i=1,2,\cdots$$

为 X 的**边际分布列**,记作 $p_{i\cdot}$. 同理称

$$P(Y=y_j)=P(X<+\infty,Y=y_j),\quad j=1,2,\cdots$$

为 Y 的**边际分布列**,记作 $p_{\cdot j}$.

根据边际分布列的定义,可以得到

$$\begin{aligned}p_{i\cdot}&=P(X=x_i)=P(X=x_i,Y<+\infty)\\&=P(X=x_i,Y=y_1)+P(X=x_i,Y=y_2)+\cdots+P(X=x_i,Y=y_j)+\cdots\\&=p_{i1}+p_{i2}+\cdots+p_{ij}+\cdots=\sum_j p_{ij},\end{aligned}$$

即

$$p_{i\cdot}=\sum_j p_{ij},\quad i=1,2,\cdots. \tag{3.2.2}$$

同理

$$p_{\cdot j}=\sum_i p_{ij},\quad j=1,2,\cdots. \tag{3.2.3}$$

可见将(X,Y)的联合分布律的表格形式中的第 i 行各数相加,即得 $p_{i\cdot}$,将第 j 列各数相加,即得 $p_{\cdot j}$. 如下表所示:

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$	$p_{i\cdot}$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$	$p_{1\cdot}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$	$p_{2\cdot}$

续表

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$	$p_{i\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$	$p_{i\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
$p_{\cdot j}$	$p_{\cdot 1}$	$p_{\cdot 2}$	$\cdots$	$p_{\cdot j}$	$\cdots$	1

例 3.2.1 已知 X 和 Y 的边际分布列均为 $\begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$，且 $P(XY=0)=1$，试求(X,Y)的联合分布律.

解 由于 $P(XY=0)=1$，所以 $P(XY\neq 0)=0$，又因为

$$\{X=-1,Y=-1\}\subset\{XY\neq 0\},$$

得

$$P(X=-1,Y=-1)=0.$$

同理可得 $P(X=-1,Y=1)=P(X=1,Y=-1)=P(X=1,Y=1)=0$. 利用联合分布列和边际分布列的关系可得$(X,Y)$的联合分布律为

$X \backslash Y$	-1	0	1	$p_{i\cdot}$
-1	0	$\frac{1}{4}$	0	$\frac{1}{4}$
0	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
1	0	$\frac{1}{4}$	0	$\frac{1}{4}$
$p_{\cdot j}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

例 3.2.2 某射手向一目标独立地连续射击，每次命中的概率都是 p，以 X 表示第二次命中时的射击次数，以 Y 表示第三次命中时的射击次数. 求(X,Y)的联合分布列，并利用联合分布列求出 X 和 Y 的边际分布列.

解 求(X,Y)的联合分布列就是求$\{X=m,Y=n\}(m<n)$的概率，$\{X=m,Y=n\}$表示在前 n 次射击中，第 m 次命中了第二次，第 n 次命中了第三次，而在前 $m-1$ 次中有一次命中，所以

$$P(X=m,Y=n)=\mathrm{C}_{m-1}^{1}p^3q^{n-3},$$

其中 $n=3,4,\cdots;m=2,3,\cdots,n-1$. X 的边际分布列为

$$P(X=m)=\sum_{n=m+1}^{\infty}C_{m-1}^{1}p^{3}q^{n-3}=C_{m-1}^{1}p^{3}\sum_{n=m+1}^{\infty}q^{n-3}=C_{m-1}^{1}p^{2}q^{m-2},\quad m=2,3,\cdots.$$

Y 的边际分布列为

$$P(Y=n)=\sum_{m=2}^{n-1}C_{m-1}^{1}p^{3}q^{n-3}=p^{3}q^{n-3}\sum_{m=2}^{n-1}C_{m-1}^{1}=C_{n-1}^{2}p^{3}q^{n-3},\quad n=3,4,\cdots.$$

例 3.2.3 一批产品共有 5 件,其中 3 件正品 2 件次品,从中任取一件后不放回,然后再从中任取一件. 设每次抽取时,每件产品被取到的概率相同,分别以 X 与 Y 表示两次取到的正品个数. 试分别在有放回和不放回两种情况下,求(X,Y)的联合分布律和边际分布列.

解 (X,Y)可能取的值为$(0,0)$,$(0,1)$,$(1,0)$和$(1,1)$.

(1) 由乘法公式知,在有放回抽取的情况下,

$$P(X=0,Y=0)=\frac{2}{5}\times\frac{2}{5}=\frac{4}{25}=0.16,$$

$$P(X=0,Y=1)=\frac{2}{5}\times\frac{3}{5}=\frac{6}{25}=0.24,$$

$$P(X=1,Y=0)=\frac{3}{5}\times\frac{2}{5}=\frac{6}{25}=0.24,$$

$$P(X=1,Y=1)=\frac{3}{5}\times\frac{3}{5}=\frac{9}{25}=0.36.$$

(2) 在无放回抽取的情况下,

$$P(X=0,Y=0)=\frac{2}{5}\times\frac{1}{4}=\frac{1}{10}=0.1,$$

$$P(X=0,Y=1)=\frac{2}{5}\times\frac{3}{4}=\frac{3}{10}=0.3,$$

$$P(X=1,Y=0)=\frac{3}{5}\times\frac{2}{4}=\frac{3}{10}=0.3,$$

$$P(X=1,Y=1)=\frac{3}{5}\times\frac{2}{4}=\frac{3}{10}=0.3.$$

于是得到在有放回抽取下,(X,Y)的联合分布律为

X \ Y	0	1	$p_{i\cdot}$
0	0.16	0.24	0.4
1	0.24	0.36	0.6
$p_{\cdot j}$	0.4	0.6	1

(X,Y)的边际分布列为

X	0	1
$p_{i\cdot}$	0.4	0.6

Y	0	1
$p_{\cdot j}$	0.4	0.6

在不放回抽取下，(X,Y)的联合分布律为

X \ Y	0	1	$p_{i\cdot}$
0	0.1	0.3	0.4
1	0.3	0.3	0.6
$p_{\cdot j}$	0.4	0.6	1

(X,Y)的边际分布列为

X	0	1
$p_{i\cdot}$	0.4	0.6

Y	0	1
$p_{\cdot j}$	0.4	0.6

这个例子说明，“有放回”和“不放回”时(X,Y)具有不同的分布律，但它们相应的边际分布列却是一样的. 这一事实表明，虽然二维随机变量的联合分布律完全决定了两个边缘分布，但反过来边际分布列却不能决定(X,Y)的联合分布律.

另外，这个例子又再一次证明了抽签的公平性，即无论是“有放回”还是“不放回”，每次抽中的概率都是一样的.

3.2.2 二维离散型随机变量的独立性

在第一章中我们已经知道，两个事件 A 与 B 相互独立是指积事件的概率等于概率乘积，即

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

如果(X,Y)为二维离散型随机变量，那么事件$\{X=x_i\}$与$\{Y=y_j\}$相互独立就是

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j).$$

那么 X 和 Y 是否可以定义独立性呢？事实上，我们有下面的定义.

定义 3.2.2 设(X,Y)为二维离散型随机变量，X 与 Y 的可能取值分别为 $x_1,x_2,\cdots$与 $y_1,y_2,\cdots$，如果对任意的 $i,j=1,2,\cdots$，都有

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j), \tag{3.2.4}$$

则称 X 与 Y 是**相互独立**的.

在例3.2.3中的有放回情形下，X与Y是相互独立的；在不放回情形下，X与Y是不独立的.

例3.2.4　试判断例3.2.2中的X与Y是否独立.

解　因为

$$\begin{aligned}P(X=m,Y=n)&=C_{m-1}^{1}p^{3}q^{n-3}\neq C_{m-1}^{1}p^{2}q^{m-2}\cdot C_{n-1}^{2}p^{3}q^{n-3}\\&=P(X=m)P(Y=n),\end{aligned}$$

所以X与Y是不独立的.

3.2.3　二维离散型随机变量的条件分布列

在第一章中，我们曾经研究过两个事件的条件概率. 同样地，我们可以定义二维离散型随机变量的条件概率.

定义3.2.3　设(X,Y)为二维离散型随机变量，X与Y的可能取值分别为$x_1,x_2,\cdots$与$y_1,y_2,\cdots$，对任意的$i,j=1,2,\cdots$，称

$$p_{i|j}\stackrel{\Delta}{=}P(X=x_i\,|\,Y=y_j)=\frac{P(X=x_i,Y=y_j)}{P(Y=y_j)}\tag{3.2.5}$$

为已知$\{Y=y_j\}$的条件下X的分布列，简称条件分布列. 类似地定义已知$\{X=x_i\}$的条件下Y的分布列为

$$p_{j|i}\stackrel{\Delta}{=}P(Y=y_j\,|\,X=x_i)=\frac{P(X=x_i,Y=y_j)}{P(X=x_i)}.\tag{3.2.6}$$

例3.2.5　试求例3.2.2中的X与Y的条件分布列.

解　已知$\{Y=n\}$的条件下X的分布列为

$$P(X=m\,|\,Y=n)=\frac{P(X=m,Y=n)}{P(Y=n)}=\frac{C_{m-1}^{1}p^{3}q^{n-3}}{C_{n-1}^{2}p^{3}q^{n-3}}=\frac{C_{m-1}^{1}}{C_{n-1}^{2}},$$

其中$n=3,4,\cdots;m=2,3,\cdots,n-1$.

已知$\{X=m\}$的条件下Y的分布列为

$$P(Y=n\,|\,X=m)=\frac{P(X=m,Y=n)}{P(X=m)}=\frac{C_{m-1}^{1}p^{3}q^{n-3}}{C_{m-1}^{1}p^{2}q^{m-2}}=pq^{n-m-1},$$

其中$n=3,4,\cdots;m=2,3,\cdots,n-1$.

例3.2.6　设在某医院每天出生的婴儿总数X服从参数为14的泊松分布，其中每个初生婴儿为男婴的概率为0.51，以Y表示每天在该医院出生的男婴个数. (1) 求(X,Y)的联合分布列；(2) 求Y的边际分布列.

解　(1) 由题意知在已知$\{X=n\}$的条件下，男婴个数Y服从参数为n，0.51的二项分布，所以

$$P(Y=m\,|\,X=n)=C_{n}^{m}\,0.51^{m}\,0.49^{n-m},\quad m=0,1,\cdots,n.$$

又 $P(X=n)=\dfrac{14^n}{n!}e^{-14}, n=0,1,\cdots$,所以$(X,Y)$的联合分布列为

$$\begin{aligned}P(X=n,Y=m)&=P(X=n)P(Y=m\mid X=n)\\&=\frac{14^n}{n!}e^{-14}C_n^m\ 0.51^m\ 0.49^{n-m}\\&=\frac{7.14^m\ 6.86^{n-m}}{m!(n-m)!}e^{-14},\end{aligned}$$

其中 $n=0,1,\cdots;m=0,1,\cdots,n$.

$$\begin{aligned}(2)\ P(Y=m)&=\sum_{n=m}^{\infty}P(X=n,Y=m)=\sum_{n=m}^{\infty}\frac{7.14^m\ 6.86^{n-m}}{m!(n-m)!}e^{-14}\\&=\frac{7.14^m}{m!}\sum_{n=m}^{\infty}\frac{6.86^{n-m}}{(n-m)!}e^{-14}=\frac{7.14^m}{m!}e^{-7.14},\quad m=0,1,\cdots.\end{aligned}$$

即 Y 服从参数为 7.14 的泊松分布.

本例中使用的等式

$$P(X=n,Y=m)=P(X=n)P(Y=m\mid X=n)$$

对于一般的二维离散型随机变量可以写成

$$P(X=x_i,Y=y_j)=P(X=x_i)P(Y=y_j\mid X=x_i).\qquad(3.2.7)$$

可以称之为二维离散随机变量的乘法公式.

数学家简介 欧拉(Euler,1707—1783),瑞士数学家和物理学家,被称为历史上最伟大的四位数学家之一,他是把微积分应用于物理学的先驱者之一.欧拉是有史以来遗产最多的数学家,他的全集共计 75 卷.欧拉实际上支配了 18 世纪的数学,对于当时的新发明微积分,他推导出了很多结果.在他生命的最后 7 年中,欧拉的双目完全失明,尽管如此,他还是以惊人的速度产出了生平一半的著作.

习 题

1. 一口袋中有四个球,它们依次标有数字 1,2,2,3.从这袋中任取一球后,不放回袋中,再从袋中任取一球.设每次取球时,袋中每个球被取到的可能性相同.以 X,Y 分别记第一二次取得的球上面标有的数字,求(X,Y)的联合分布律及 $P(X=Y)$.

2. 盒中装有 3 个黑球、2 个红球、2 个白球,从中任取 4 个,以 X 表示取到的黑球数,以 Y 表示取到的白球数,求(X,Y)的联合分布律和边际分布列.

3. 将一个硬币抛掷 3 次,以 X 表示 3 次中出现正面的次数,以 Y 表示出现正面次数与反面次数之差的绝对值,求(X,Y)的联合分布律和边际分布列.

4. 设(X,Y)的联合分布律为

$X \backslash Y$	-1	0	1
0	0.1	0.2	α
1	β	0.1	0.2

且 $P(X+Y=1)=0.4$,求(1) α,β;(2) $P(X+Y<1)$;(3) $P(X^2Y^2=1)$.

5. 箱子中有 12 件产品,其中 2 件是次品.从中任取 2 次,每次取 1 件,定义随机变量 X,Y 如下:

$$X=\begin{cases}1, & \text{第 1 次取出的是次品,}\\ 0, & \text{第 1 次取出的是正品.}\end{cases} \qquad Y=\begin{cases}1, & \text{第 2 次取出的是次品,}\\ 0, & \text{第 2 次取出的是正品.}\end{cases}$$

试就下面两种情况写出(X,Y)的联合分布律和边际分布列,并判断 X,Y 是否独立? (1) 有放回抽取;(2) 无放回抽取.

6. 已知(X,Y)的联合分布律为

$X \backslash Y$	0	1
0	0.4	a
1	b	0.1

若事件$\{X=0\}$和$\{X+Y=1\}$互相独立,求 a,b.

7. 袋中有五个号码 1,2,3,4,5,从中任取三个,记这三个号码中最小的号码为 X,最大的号码为 Y.

(1) 求 X 与 Y 的联合分布律;

(2) X 与 Y 是否相互独立?

8. 设某班车起点站上客人数 X 服从参数为$\lambda(\lambda>0)$的泊松分布,每位乘客在中途下车的概率为 $p(0<p<1)$,且中途下车与否相互独立,以 Y 表示在中途下车的人数,求:(1) 在发车时有 n 个乘客的条件下,中途有 m 人下车的概率;(2) 二维随机变量(X,Y)的联合分布律.

§3.3 二维连续型随机变量

一维连续型随机变量的概率特性是由它的密度函数来决定的,类似地,二维连续型随机变量的概率特性由它的联合密度来决定.

定义 3.3.1 设 $F(x,y)$为二维随机变量(X,Y)的联合分布函数,若存在非负函数 $f(x,y)$,使得对于任意的 $x,y\in\mathbf{R}$,有

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\mathrm{d}v\mathrm{d}u, \tag{3.3.1}$$

则称(X,Y)为**二维连续型随机变量**，并称 $f(x,y)$为(X,Y)的**联合概率密度函数**，简称**联合密度**(或**概率密度**).

联合密度 $f(x,y)$具有以下性质：

(1) 非负性：$f(x,y)\geqslant 0\ (x,y\in\mathbf{R})$；

(2) 归一性：$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y=1.$ (3.3.2)

由联合密度函数的定义还可以得到如下性质：

(1) $F(x,y)$是二元连续函数；

(2) 在 $f(x,y)$的连续点(x,y)处有

$$\frac{\partial^2 F(x,y)}{\partial x\partial y}=f(x,y). \tag{3.3.3}$$

注　(1) 二维连续型随机变量的定义与二维离散型随机变量的定义截然不同.如果 X 与 Y 都是一维离散型随机变量，那么(X,Y)就是二维离散型随机变量；但是，如果 X 与 Y 都是一维连续型随机变量，(X,Y)不一定是二维连续型随机变量.因为在定义中要求(X,Y)的联合分布函数 $F(x,y)$在 $f(x,y)$的连续点处混合偏导存在，然而，对于任意两个一维连续型随机变量 X 与 Y，我们都不能保证 $F(x,y)$是二维连续的.反之，如果(X,Y)是二维连续型随机变量，那么 X 与 Y 肯定都是一维连续型随机变量，因为由第一节的知识，我们知道 X 的分布函数为

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

所以 X 具有密度函数

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y,\quad x\in\mathbf{R}.$$

同样地可以得到 Y 的密度函数为

$$f_Y(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x,\quad y\in\mathbf{R}.$$

所以 X 与 Y 必为一维连续型随机变量.

(2) 我们曾经将一维连续型随机变量的密度函数比喻成质量线密度，那么二维连续型随机变量的联合密度就相当于质量面密度.

(3) 由于一维连续型随机变量 X 为直线 $\mathbf{R}$ 上的一个随机点，所以在求 X 的概率的时候都是求 X 落在某个区间(a,b)内的概率，就是在区间(a,b)上关于 X 的密度函数 $f(x)$进行一重积分

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

现在二维连续型随机变量(X,Y)作为平面 $\mathbf{R}^2$ 上的一个二维的随机点，所以求(X,Y)的概率就是求(X,Y)落在某个平面区域 G 中的概率，这个概率当然应该是在区域 G 上关于(X,Y)的联合密度函数 $f(x,y)$进行二重积分

$$P((X,Y) \in G) = \iint_G f(x,y)\mathrm{d}x\mathrm{d}y. \tag{3.3.4}$$

这个公式非常重要，可以说，它是所有二维连续型随机变量求概率的唯一的公式.

3.3.1 二维连续型随机变量的边际密度

定义 3.3.2 设(X,Y)为二维连续型随机变量，联合密度函数为 $f(x,y)$，称

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y, \quad x \in \mathbf{R}, \tag{3.3.5}$$

为 X 的边际密度函数，称

$$f_Y(x) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x, \quad y \in \mathbf{R}, \tag{3.3.6}$$

为 Y 为边际密度函数.

例 3.3.1 设二维连续随机变量(X,Y)的联合密度函数为

$$f(x,y) = \begin{cases} Axy, & 0 < x < 1, x^2 < y < 1, \\ 0, & \text{其他}. \end{cases}$$

其中常数 $A<0$. 试求：(1) 常数 A；(2) X 与 Y 的边际密度函数 $f_X(x)$与 $f_Y(y)$；(3) $P(X>Y)$.

解 将(X,Y)的联合密度函数的非零区域画在平面 $\mathbf{R}^2$ 上，如图 3－3 所示.

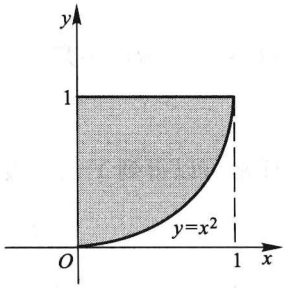


图 3－3

(1) 利用联合密度函数的归一性可得

$$\int_0^1 \int_{x^2}^1 Axy\,\mathrm{d}x\mathrm{d}y = 1,$$

由于 $\int_0^1 \int_{x^2}^1 xy\,\mathrm{d}x\mathrm{d}y = \dfrac{1}{6}$，所以 $A=6$.

(2) 随机变量 X 的取值范围是$[0,1]$，所有当 $0<x<1$时，

$$f_X(x) = \int_{x^2}^1 6xy\,\mathrm{d}y = 3(x - x^5),$$

于是

$$f_X(x) = \begin{cases} 3(x - x^5), & 0 < x < 1, \\ 0, & \text{其他}. \end{cases}$$

随机变量 Y 的取值范围是[0,1],所有当 $0<y<1$ 时,

$$f_Y(y)=\int_0^{\sqrt{y}}6xy\,\mathrm{d}x=3y^2,$$

于是

$$f_Y(y)=\begin{cases}3y^2, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

(3) $\{X>Y\}$表示的区域应该是曲线 $y=x^2$ 与直线 $y=x$ 围成的区域,所以

$$P(X>Y)=\int_0^1\int_{x^2}^{x}6xy\,\mathrm{d}x\mathrm{d}y=\int_0^1 3x(x^2-x^4)\mathrm{d}x=\frac{3}{12}.$$

下面介绍两个常用的二维连续型随机变量的分布:

1. 二维均匀分布

设 D 为平面有界闭区域,其面积为 S_D,若密度函数为

$$f(x,y)=\begin{cases}\dfrac{1}{S_D}, & (x,y)\in D,\\ 0, & (x,y)\notin D,\end{cases}\tag{3.3.7}$$

则称二维随机变量(X,Y)服从 D 上的**二维均匀分布**.

若 G 为 D 的子区域,面积为 S_G,则由二维随机变量求概率的公式得

$$P((X,Y)\in G)=\iint_G f(x,y)\mathrm{d}\sigma=\frac{1}{S_D}\iint_G\mathrm{d}\sigma=\frac{S_G}{S_D}.$$

这表明二维均匀分布随机变量(X,Y)落入 D 内任意子区域 G 内的概率只与 G 的面积有关,与 G 的形状及位置没有关系,这就是“均匀”一词的体现,与第一章介绍的几何概型是一致的.

例 3.3.2　设平面区域 D 由曲线 $y=\dfrac{1}{x^2}$ 及直线 $y=0,x=1$ 所围成,二维随机变量(X,Y)在 D 上服从均匀分布.求(1) (X,Y)的联合密度函数;(2) X 与 Y 的边际密度函数 $f_X(x)$与 $f_Y(y)$.

解　平面区域 D 如图 3-4 所示.

(1) 首先计算出 D 的面积 $S_D=\int_1^{\infty}\dfrac{1}{x^2}\mathrm{d}x=1$. 则随机变量$(X,Y)$的联合密度函数为

$$f(x,y)=\begin{cases}1, & (x,y)\in D,\\ 0, & (x,y)\notin D.\end{cases}$$

(2) 随机变量 X 的取值范围是$[1,+\infty)$,所以当 $1<x<\infty$时,

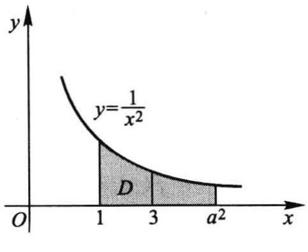


图 3-4

$$f_X(x)=\int_0^{x^{-2}}1\mathrm{d}y=\frac{1}{x^2},$$

于是

$$f_X(x)=\begin{cases}\dfrac{1}{x^2}, & x>1,\\ 0, & \text{其他}.\end{cases}$$

随机变量 Y 的取值范围是$[0,1]$，所以当 $0<y<1$ 时，

$$f_Y(y)=\int_1^{\frac{1}{\sqrt{y}}}1\mathrm{d}x=\frac{1}{\sqrt{y}}-1,$$

于是

$$f_Y(y)=\begin{cases}\dfrac{1}{\sqrt{y}}-1, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

2. 二维正态分布

若随机变量(X,Y)的联合密度函数为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\cdot\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}\quad(x,y\in\mathbf{R}),\qquad(3.3.8)$$

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**(其中 $\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$ 为常数，且有 $\sigma_1>0,\sigma_1>0,|\rho|<1$). 记为$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$.

特殊地，当 $\mu_1=\mu_2=0,\sigma_1=\sigma_2=1$ 时

$$f(x,y)=\frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)\right].$$

更特殊地，当 $\rho=0$ 时，则有

$$f(x,y)=\frac{1}{2\pi}\exp\left[-\frac{1}{2}(x^2+y^2)\right].$$

正态分布是最常见，也是最有用的分布，在后续课程中有重要的应用.

例 3.3.3　求二维正态分布的边缘分布.

解　设随机变量(X,Y)服从二维正态分布，即$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，利用变量替换 $t=\dfrac{y-\mu_2}{\sigma_2}$，得

$$\begin{aligned}f_X(x)&=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y\\&=\int_{-\infty}^{+\infty}\frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)t}{\sigma_1}+t^2\right]\mathrm{d}t\right.\end{aligned}$$

$$=\int_{-\infty}^{+\infty}\frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[(1-\rho^2+\rho^2)\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)t}{\sigma_1}+t^2\right]\mathrm{d}t\right.$$

$$=\frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\cdot\sqrt{1-\rho^2}}\cdot\exp\left\{-\frac{1}{2(1-\rho^2)}\left[t-\frac{\rho(x-\mu_1)}{\sigma_1}\right]^2\right\}\mathrm{d}t$$

$$=\frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\quad(x\in\mathbf{R}).$$

类似地，$f_Y(y)=\dfrac{1}{\sqrt{2\pi}\sigma_2}\mathrm{e}^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}\quad(y\in\mathbf{R}).$

这表明 $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，即二维正态分布的两个边缘分布都是一维正态分布. 此外还注意到，两个边缘密度都不含参数 ρ，这意味着具有相同参数 $\mu_1,\mu_2,\sigma_1,\sigma_2$，但是 ρ 值不同的二维正态分布具有相同的边缘分布函数. 这一事实再一次说明，对二维连续型随机变量而言，仅有 X 和 Y 的边缘分布一般不能决定 X 和 Y 的联合分布.

3.3.2 二维连续型随机变量的独立性

下面是二维连续型随机变量相互独立的定义.

定义 3.3.3 设 $F(x,y)$ 及 $F_X(x)$，$F_Y(y)$ 分别是二维连续型随机变量 (X,Y) 的分布函数和边缘分布函数，若对于所有的 x,y 有

$$f(x,y)=f_X(x)f_Y(y)\tag{3.3.9}$$

或

$$F(x,y)=F_X(x)\cdot F_Y(y),\tag{3.3.10}$$

则称**随机变量 X 和 Y 是相互独立的**.

易知(3.3.9)式与(3.3.10)式是等价的.

例 3.3.1 与例 3.3.2 中的 X 与 Y 都是不独立的. 对于二维正态分布，利用例 3.3.3 的结果可得如下的结论：

性质 设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则 X 与 Y 相互独立的充要条件是 $\rho=0$.

例 3.3.4 设 X 和 Y 是两个相互独立的随机变量，X 在 $(0,1)$ 上服从均匀分布，Y 服从参数为 $\frac{1}{2}$ 的指数分布. 求关于 a 的二次方程 $a^2+2Xa+Y=0$ 有实根的概率.

解 $f_X(x)=\begin{cases}1, & 0<x<1,\\ 0, & 其他,\end{cases}$ 又 X 与 Y 相互独立，故 X 与 Y 的联合密度函

数为

$$f(x,y)=f_X(x)\cdot f_Y(y)$$
$$=\begin{cases}\frac{1}{2}e^{-\frac{y}{2}}, & 0<x<1,y>0,\\ 0, & \text{其他}.\end{cases}$$

因{方程有实根}={判别式 $\Delta^2=4X^2-4Y\geqslant 0$}=$\{X^2\geqslant Y\}$，所求概率为随机点落入图 3-5 所示阴影部分的概率. 则

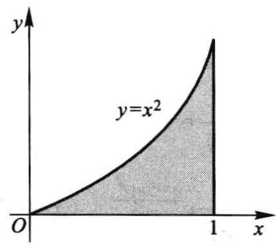


图 3-5

$$P(\text{方程有实根})=P(Y\leqslant X^2)=\iint\limits_{y\leqslant x^2}f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^{x^2}\frac{1}{2}e^{-\frac{y}{2}}\mathrm{d}y$$
$$=\int_0^1\mathrm{d}x\left(-e^{-\frac{y}{2}}\right)\Big|_0^{x^2}=\int_0^1(1-e^{-\frac{x^2}{2}})\mathrm{d}x$$
$$=1-\int_0^1 e^{-\frac{x^2}{2}}\mathrm{d}x=1-\sqrt{2\pi}(\Phi(1)-\Phi(0)).$$

3.3.3 二维连续型随机变量的条件密度

在上一节中，我们曾经讨论过二维离散型随机变量的条件分布列，那么对于二维连续型随机变量，我们是否可以类似地定义条件密度函数呢？大家知道，如果(X,Y)是二维连续型随机变量，X 和 Y 都是一维连续型随机变量，那么对任意的 x,y，有

$$P(X=x)=P(Y=y)=0,$$

所以$\frac{P(X=x,Y=y)}{P(X=x)}$与$\frac{P(X=x,Y=y)}{P(Y=y)}$毫无意义！注意到这两个式子都是 0 比 0，所以我们可以借用洛比达法则的思想，先来定义条件分布函数.

定义 3.3.4 设(X,Y)是二维连续型随机变量，对 $x,y\in\mathbf{R}$，称

$$F_{X|Y}(x|y)=\lim_{\varepsilon\to 0^+}P(X\leqslant x,y-\varepsilon<Y\leqslant y+\varepsilon)$$
$$=\lim_{\varepsilon\to 0^+}\frac{P(X\leqslant x,y-\varepsilon<Y\leqslant y+\varepsilon)}{P(y-\varepsilon<Y\leqslant y+\varepsilon)}$$

为给定 $Y=y$ 的条件下 X 的**条件分布函数**. 类似地，称

$$F_{Y|X}(y|x)=\lim_{\varepsilon\to 0^+}P(Y\leqslant y,x-\varepsilon<X\leqslant x+\varepsilon)$$
$$=\lim_{\varepsilon\to 0^+}\frac{P(Y\leqslant y,x-\varepsilon<X\leqslant x+\varepsilon)}{P(x-\varepsilon<X\leqslant x+\varepsilon)}$$

为给定 $X=x$ 的条件下 Y 的**条件分布函数**.

利用上述定义，可知

$$F_{X|Y}(x|y)=\lim_{\varepsilon\to 0^+}\frac{F(x,y+\varepsilon)-F(x,y-\varepsilon)}{F_Y(y+\varepsilon)-F_Y(y-\varepsilon)}=\frac{\dfrac{\partial F(x,y)}{\partial y}}{\dfrac{\partial F_Y(y)}{\partial y}}$$

$$=\frac{1}{f_Y(y)}\int_{-\infty}^{x}f(u,y)\mathrm{d}u=\int_{-\infty}^{x}\frac{f(u,y)}{f_Y(y)}\mathrm{d}u.$$

类似地可得

$$F_{Y|X}(y|x)=\int_{-\infty}^{y}\frac{f(x,u)}{f_X(x)}\mathrm{d}u.$$

于是，按照密度函数的定义，我们可以给出条件密度的定义.

定义 3.3.5 设二维连续型随机变量(X,Y)的联合密度为 $f(x,y)$，对于给定的 $Y=y$，如果 $f_Y(y)>0$，则称

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)} \tag{3.3.11}$$

为给定 $Y=y$ 的条件下 X 的**条件密度函数**；对于给定的 $X=x$，如果 $f_X(x)>0$，则称

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)} \tag{3.3.12}$$

为给定 $X=x$ 的条件下 Y 的**条件密度函数**.

给定 $Y=a$ 的条件下 X 的条件密度函数 $f_{X|Y}(x|a)$可以理解为将二维随机点(X,Y)限制在平行于 x 轴的直线 $y=a$ 上之后，这条直线上的质量线密度.

例 3.3.5 已知二维连续随机变量(X,Y)的联合密度为 $f(x,y)=\begin{cases}9\mathrm{e}^{-3y}, & 0<x<y,\\ 0, & \text{其他},\end{cases}$ 试求条件密度 $f_{X|Y}(x|y)$与 $f_{Y|X}(y|x)$.

解 (X,Y)的联合密度的非零区域为第一象限内对角线的上方，如图 3-6 所示.

容易求得 X 的边际密度为

$$f_X(x)=\begin{cases}3\mathrm{e}^{-3x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

Y 的边际密度为

$$f_Y(y)=\begin{cases}9y\mathrm{e}^{-3y}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

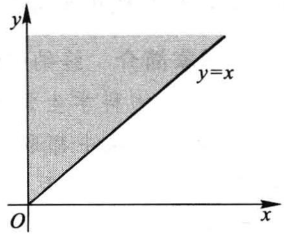


图 3-6

所以

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}=\begin{cases}\dfrac{1}{y}, & 0<x<y,\\ 0, & \text{其他};\end{cases}$$

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=\begin{cases}3\mathrm{e}^{-3(y-x)}, & 0<x<y,\\ 0, & \text{其他}.\end{cases}$$

其中,在给定 $Y=y$ 的条件下 X 服从区间$[0,y]$上的均匀分布.

例 3.3.6 已知连续型随机变量 $X\sim E(\theta)$,且对 $x>0$,当 $X=x$ 时,$Y\sim E(x)$.设 $a>0$,求概率 $P(XY<a)$.

解 为了求概率 $P(XY<a)$,必须求(X,Y)的联合密度 $f(x,y)$.由题意知给定 $X=x$ 时,Y 的条件密度为 $f_{Y|X}(y|x)=\begin{cases}x\mathrm{e}^{-xy}, & y>0,\\ 0, & y\leqslant 0,\end{cases}$ 又 $f_X(x)=\begin{cases}\theta\mathrm{e}^{-\theta x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$ 所以(X,Y)的联合密度为

$$f(x,y)=f_X(x)f_{Y|X}(y|x)=\begin{cases}\theta x\mathrm{e}^{-\theta x-xy}, & x>0,y>0,\\ 0, & \text{其他},\end{cases}$$

概率 $P(XY<a)$的积分区域就是第一象限中双曲线 $xy=a$ 的下方,故

$$\begin{aligned}P(XY<a)&=\int_0^{\infty}\mathrm{d}x\int_0^{\frac{a}{x}}\theta x\mathrm{e}^{-\theta x-xy}\mathrm{d}y=\int_0^{\infty}\theta\mathrm{e}^{-\theta x}\mathrm{d}x\int_0^{\frac{a}{x}}x\mathrm{e}^{-xy}\mathrm{d}y\\&=\int_0^{\infty}\theta\mathrm{e}^{-\theta x}\mathrm{d}x(1-\mathrm{e}^{-a})=1-\mathrm{e}^{-a}.\end{aligned}$$

本例中出现的公式

$$f(x,y)=f_X(x)f_{Y|X}(y|x) \tag{3.3.13}$$

称为二维连续型随机变量的乘法公式.

利用例 3.3.3,读者可以证明,如果$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$,那么在给定 $Y=y$ 的条件下 X 服从正态分布 $N\left(\mu_1+\rho\dfrac{\sigma_1}{\sigma_2}(y-\mu_2),\sigma_1^2(1-\rho^2)\right)$,在给定 $X=x$ 的条件下 Y 服从正态分布 $N\left(\mu_2+\rho\dfrac{\sigma_2}{\sigma_1}(x-\mu_1),\sigma_2^2(1-\rho^2)\right)$.

数学家简介 维纳(Wiener,1894—1964),美国数学家,控制论的创始人.维纳在其 50 年的科学生涯中,先后涉足哲学、数学、物理学和工程学,最后转向生物学,在各个领域中都取得了丰硕成果,称得上是恩格斯颂扬过的本世纪多才多艺和学识渊博的科学巨人.他一生发表论文 240 多篇,著作 14 本.由他首先研究的一类随机过程——维纳过程至今仍然是概率论的核心问题.

习　题

1. 设(X,Y)的联合密度函数为 $f(x,y)=\begin{cases} ke^{-2x-y}, & x>0,y>0, \\ 0, & \text{其他}. \end{cases}$ 求：

(1) 常数 k；(2) $P(Y\leqslant X)$；(3) $P(X+Y\leqslant 1)$；(4) $F(x,y)$；(5) $P(Y=X)$.

2. 设二维随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases} 4.8y(2-x), & 0\leqslant x\leqslant 1, 0\leqslant y\leqslant x, \\ 0, & \text{其他}, \end{cases}$$

求边缘概率密度.

3. 设(X,Y)的联合密度函数为 $f(x,y)=\begin{cases} c(R-\sqrt{x^2+y^2}), & x^2+y^2<R^2, \\ 0, & \text{其他}. \end{cases}$ (1) 求常数 c；(2) 求(X,Y)落在圆 $x^2+y^2\leqslant r^2\,(0<r<R)$内的概率.

4. 已知(X,Y)在 $D=\{(x,y)\mid x^2+y^2\leqslant 1\}$上服从均匀分布. 求(1) (X,Y)的密度函数；(2) $P(X+Y\leqslant 1)$；(3) X 与 Y 是否独立？

5. 设随机变量 X 与 Y 相互独立，且均服从区间$[0,3]$上的均匀分布，求 $P(\max\{X,Y\}<1)$.

6. 设随机变量 X,Y 相互独立，X 服从$[0,0.2]$上的均匀分布，Y 服从参数为 5 的指数分布，求(X,Y)的联合密度函数及 $P(X\geqslant Y)$.

7. 设二维随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases} \dfrac{1}{2x^2y}, & 1<x<\infty, \dfrac{1}{x}<y<x, \\ 0, & \text{其他}, \end{cases}$$

求条件概率密度 $f_{X|Y}(x\mid y)$及 $f_{Y|X}(y\mid x)$.

§ 3.4　二维随机变量函数的分布

设(X,Y)为二维随机变量，$g(x,y)$为一个二元函数，由此可得到二维随机变量(X,Y)的函数 $Z=g(X,Y)$，这是一个一维随机变量. 本节讨论二维随机变量函数的分布问题，即已知(X,Y)的联合分布，如何求函数 $Z=g(X,Y)$的分布. 首先讨论(X,Y)为二维离散型随机变量的情形.

3.4.1　二维离散型随机变量函数的分布

例 3.4.1(泊松分布的可加性)　已知 X 与 Y 相互独立，且分别服从参数为 λ_1 和 λ_2 的泊松分布. 证明：$Z=X+Y$ 服从参数为 $\lambda_1+\lambda_2$ 的泊松分布.

证明　$Z=X+Y$ 的可能取值为 0,1,2,⋯. 对任何非负整数 k，有

$$P(Z=k)=P(X+Y=k)=\sum_{i=0}^{k}P(X=i,Y=k-i)$$
$$=\sum_{i=0}^{k}\frac{\lambda_1^i}{i!}\mathrm{e}^{-\lambda_1}\frac{\lambda_2^{k-i}}{(k-i)!}\mathrm{e}^{-\lambda_2}=\frac{\mathrm{e}^{-(\lambda_1+\lambda_2)}}{k!}\sum_{i=0}^{k}\frac{k!}{i!(k-i)!}\lambda_1^i\lambda_2^{k-i}$$
$$=\frac{(\lambda_1+\lambda_2)^k}{k!}\mathrm{e}^{-(\lambda_1+\lambda_2)},$$

即 $Z=X+Y\sim P(\lambda_1+\lambda_2)$.

例 3.4.2 已知 X 与 Y 相互独立,且都服从参数为 p 的几何分布.求 $Z=X+Y$的分布列.

解 $Z=X+Y$ 的可能取值为 2,3,….对任何非负整数 $k\geqslant 2$,有

$$P(Z=k)=P(X+Y=k)=\sum_{i=1}^{k-1}P(X=i,Y=k-i)$$
$$=\sum_{i=1}^{k-1}P(X=i)P(Y=k-i)=\sum_{i=1}^{k-1}p(1-p)^{i-1}p(1-p)^{k-i-1}$$
$$=p^2\sum_{i=1}^{k-1}(1-p)^{k-2}=(k-1)p^2(1-p)^{k-2}.$$

3.4.2 二维连续型随机变量函数的分布

设(X,Y)为二维连续型随机变量,其联合密度函数为 $f(x,y)$,$g(x,y)$是一个已知的连续函数,由此得到的二维随机变量 $Z=g(X,Y)$仍是连续型的随机变量.为求其密度函数 $f_Z(z)$,可先求 Z 的分布函数 $F_Z(z)$,然后对 $F_Z(z)$求导数,则得到随机变量 Z 的密度函数 $f_Z(z)=F_Z'(z)$.

一般地,Z 的分布函数为

$$F_Z(z)=P(Z\leqslant z)=P(g(X,Y)\leqslant z)=P((X,Y)\in D_z)$$
$$=\iint_{D_z}f(x,y)\mathrm{d}x\mathrm{d}y,\tag{3.4.1}$$

其中区域 $D_z=\{(x,y)\mid g(x,y)\leqslant z\}$,然后对 $F_Z(z)$求导就可得随机变量 Z 的密度函数.

从上面的推导过程中可以看到:理论上显然可以计算任意的 $Z=g(X,Y)$的密度函数.一定要先求 Z 的取值范围,但最关键的是要在密度函数 $f(x,y)$的非零区域中去找准区域 $D_z=\{(x,y)\mid g(x,y)\leqslant z\}$.

例 3.4.3 设随机变量 X 与 Y 相互独立,且 $X\sim E(\alpha)$,$Y\sim E(\beta)$.试求下列随机变量的密度函数

(1) $Z_1=X+Y$;(2) $Z_2=\dfrac{Y}{X}$;(3) $Z_3=\dfrac{Y}{X+Y}$.

解 因为 X 与 Y 相互独立,所以(X,Y)的联合密度为

$$f(x,y)=f_X(x)f_Y(y)=\begin{cases}\alpha\beta e^{-\alpha x-\beta y}, & x>0,y>0,\\ 0, & \text{其他}.\end{cases}$$

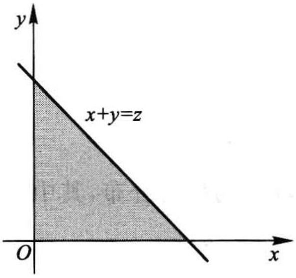
(1) 显然 Z_1 的取值范围为$[0,+\infty)$,所以(图 3-7),当 $z<0$ 时,$F_{Z_1}(z)=0$.
当 $z\geqslant 0$ 时,

$$\begin{aligned}F_{Z_1}(z)&=P(Z_1\leqslant z)=P(X+Y\leqslant z)\\&=\int_0^z dx\int_0^{z-x}\alpha\beta e^{-\alpha x-\beta y}dy\\&=1-e^{-\alpha z}-\frac{\alpha}{\alpha-\beta}(e^{-\beta z}-e^{-\alpha z}).\end{aligned}$$

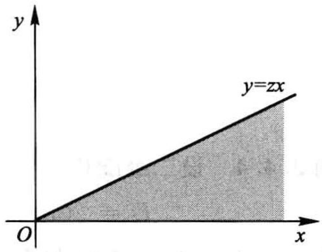
所以

$$f_{Z_1}(z)=\begin{cases}\dfrac{\alpha\beta}{\alpha-\beta}(e^{-\beta z}-e^{-\alpha z}), & z\geqslant 0,\\ 0, & z<0.\end{cases}$$

(2) Z_2 的取值范围为$[0,+\infty)$,所以(图 3-8),当 $z<0$ 时,$F_{Z_2}(z)=0$.



(积分区域为第一象限中直线
$x+y=z$ 的下方)
图 3-7

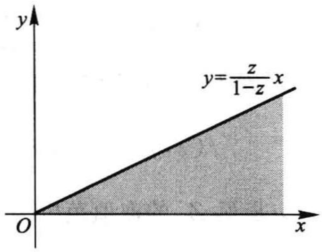


(积分区域为第一象限中直线
$y=zx$ 的下方)
图 3-8

当 $z\geqslant 0$ 时,

$$\begin{aligned}F_{Z_2}(z)&=P(Z_2\leqslant z)=P\left(\frac{Y}{X}\leqslant z\right)\\&=\int_0^{\infty}dy\int_{\frac{y}{z}}^{\infty}\alpha\beta e^{-\alpha x-\beta y}dx=\int_0^{\infty}\beta e^{-\beta y}dy\int_{\frac{y}{z}}^{\infty}\alpha e^{-\alpha x}dx\\&=\int_0^{\infty}\beta e^{-\beta y}dy e^{-\alpha\frac{y}{z}}=\int_0^{\infty}\beta e^{-y\left(\beta+\frac{\alpha}{z}\right)}dy\end{aligned}$$

$$= \frac{\beta z}{\alpha + \beta z}.$$



（积分区域为第一象限中直线 $y=\frac{z}{1-z}x$ 的下方）

图 3-9

所以

$$f_{Z_2}(z) = \begin{cases} \frac{\alpha\beta}{(\alpha+\beta z)^2}, & z \geqslant 0, \\ 0, & z < 0. \end{cases}$$

(3) Z_3 的取值范围为[0,1]，所以（图 3-9），当 $z<0$ 时，$F_{Z_3}(z)=0$；当 $z\geqslant 1$ 时，$F_{Z_3}(z)=1$.

当 $0\leqslant z<1$ 时，

$$\begin{aligned} F_{Z_3}(z) &= P(Z_3 \leqslant z) = P\left(\frac{Y}{X+Y} \leqslant z\right) \\ &= \int_0^{\infty} \mathrm{d}y \int_{\frac{(1-z)y}{z}}^{\infty} \alpha\beta \mathrm{e}^{-\alpha x-\beta y} \mathrm{d}x = \int_0^{\infty} \beta \mathrm{e}^{-\beta y} \mathrm{d}y \int_{\frac{(1-z)y}{z}}^{\infty} \alpha \mathrm{e}^{-\alpha x} \mathrm{d}x \\ &= \int_0^{\infty} \beta \mathrm{e}^{-\beta y} \mathrm{d}y \mathrm{e}^{-\alpha\frac{(1-z)y}{z}} = \int_0^{\infty} \beta \mathrm{e}^{-y\left(\beta+\alpha\frac{1-z}{z}\right)} \mathrm{d}y \\ &= \frac{\beta z}{\beta z + \alpha(1-z)}. \end{aligned}$$

所以

$$f_{Z_3}(z) = \begin{cases} \frac{\alpha\beta}{[\beta z + \alpha(1-z)]^2}, & 0 \leqslant z < 1, \\ 0, & \text{其他}. \end{cases}$$

例 3.4.4 设二维随机变量 (X,Y) 服从区域 D 上的均匀分布，其中

$$D = \{(x,y) \mid 0 < x < 2, 0 < y < 2\},$$

求 $Z=X-Y$ 的分布函数及密度函数.

解 (X,Y) 的联合密度函数为

$$f(x,y) = \begin{cases} \frac{1}{4}, & 0 < x < 2, 0 < y < 2, \\ 0, & \text{其他}. \end{cases}$$

因为 $Z=X-Y$，所以 Z 的取值范围为 $[-2,2]$，所以当 $z<-2$ 时，$F_Z(z)=0$（图 3-10(a)）；当 $z\geqslant 2$ 时，$F_Z(z)=1$（图 3-10(d)）.

当 $-2\leqslant z<0$ 时，积分区域见图 3-10(b)，

$$F_Z(z) = \iint\limits_{D_z} \frac{1}{4} \mathrm{d}x\mathrm{d}y = \frac{1}{4} S_{D_z'} = \frac{1}{4} \times \frac{1}{2}(2+z)^2 = \frac{1}{8}(2+z)^2;$$

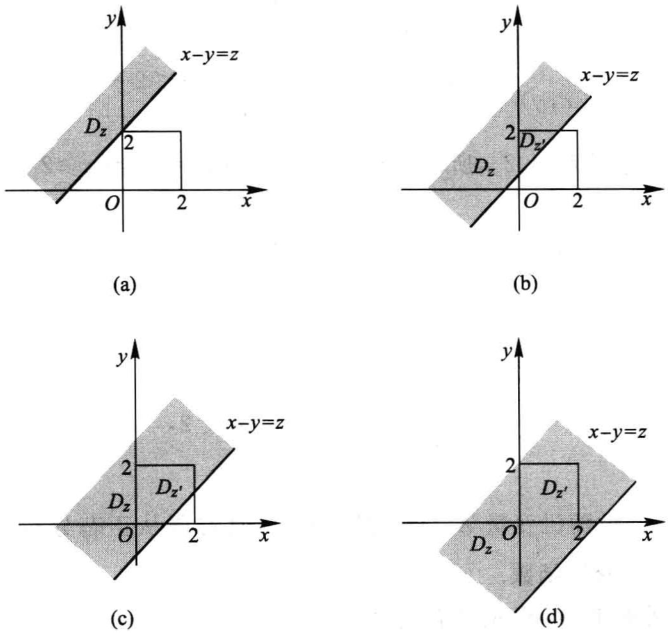


图 3－10

当 $0\leqslant z<2$ 时，积分区域见图 3－10(c)，

$$F_Z(z)=\iint\limits_{D_z}\frac{1}{4}\mathrm{d}x\mathrm{d}y=\frac{1}{4}S_{D_{z'}}=\frac{1}{4}\left[4-\frac{1}{2}(2-z)^2\right]=1-\frac{1}{8}(2-z)^2;$$

于是

$$F_Z(z)=\begin{cases}0, & z<-2,\\ \dfrac{1}{8}(2+z)^2, & -2\leqslant z<0,\\ 1-\dfrac{1}{8}(2-z)^2, & 0\leqslant z<2,\\ 1, & z\geqslant 2.\end{cases}$$

所以

$$f_Z(z)=F'_Z(z)=\begin{cases}\dfrac{1}{4}(2+z), & -2<z<0,\\ \dfrac{1}{4}(2-z), & 0\leqslant z<2,\\ 0, & \text{其他}.\end{cases}$$

例 3.4.5 设(X,Y)服从区域$\{(x,y):0<x<1,0<y<2-2x\}$上的二维均

匀分布(图 3-11),求 $Z=Y+2X$ 的密度函数.

解 (X,Y)的联合密度函数为

$$f(x,y)=\begin{cases}1, & 0<x<1,0<y<2-2x,\\ 0, & \text{其他}.\end{cases}$$

Z 的取值范围为$[0,2]$,所以,当 $z<0$ 时,$F_Z(z)=0$;当 $z\geqslant 2$ 时,$F_Z(z)=1$.
当 $0\leqslant z<2$ 时,

$$\begin{aligned}F_Z(z)&=P(Z\leqslant z)=P(Y+2X\leqslant z)\\&=\frac{z^2}{4}.\end{aligned}$$

所以

$$f_Z(z)=\begin{cases}\dfrac{z}{2}, & 0\leqslant z<2,\\ 0, & \text{其他}.\end{cases}$$

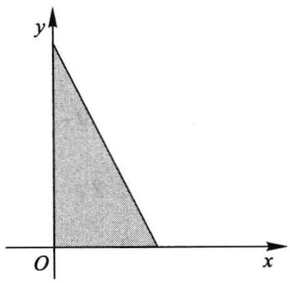
例 3.4.6 设随机变量 X 与 Y 相互独立,且都服从正态分布 $N(0,\sigma^2)$,求 $Z=\sqrt{X^2+Y^2}$的密度函数.

解 由题意知(X,Y)的联合密度为

$$f(x,y)=\frac{1}{2\pi\sigma^2}\exp\left\{-\frac{x^2+y^2}{2\sigma^2}\right\},x,y\in\mathbf{R}.$$

显然 Z 的取值范围为$[0,+\infty)$,所以,

当 $z<0$ 时,$F_Z(z)=0$(积分区域为圆 $x^2+y^2=z^2$ 的内部)(图 3-12);



(积分区域为第一象限中直线 $y+2x=z$ 的下方)

图 3-11

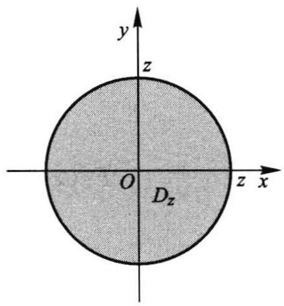


图 3-12

当 $z\geqslant 0$ 时,$F_Z(z)=P(Z\leqslant z)=P(X^2+Y^2\leqslant z^2)$

$$=\iint\limits_{x^2+y^2<z^2}\frac{1}{2\pi\sigma^2}\exp\left\{-\frac{x^2+y^2}{2\sigma^2}\right\}\mathrm{d}x\mathrm{d}y$$

$$= \int_0^{2\pi} \mathrm{d}\theta \int_0^z \frac{r}{2\pi\sigma^2} \exp\left\{-\frac{r^2}{2\sigma^2}\right\} \mathrm{d}r$$

$$= 1 - \exp\left\{-\frac{z^2}{2\sigma^2}\right\}.$$

所以

$$f_Z(z) = \begin{cases} \dfrac{z}{\sigma^2} \exp\left\{-\dfrac{z^2}{2\sigma^2}\right\}, & 0 \leqslant z, \\ 0, & z < 0. \end{cases}$$

通常称此分布为瑞利分布.

3.4.3 极大极小分布

例 3.4.7 设随机变量 $X_1, X_2, \cdots, X_n$ 相互独立,且 X_i 的分布函数为 $F_{X_i}(x)$, $i=1,2,\cdots$. 令 $Y=\max\{X_1, X_2, \cdots, X_n\}$, $Z=\min\{X_1, X_2, \cdots, X_n\}$. 那么有

(1) Y 的分布函数为 $F_{\max}(x) = \prod\limits_{i=1}^{n} F_{X_i}(x)$;

(2) Z 的分布函数为 $F_{\min}(x) = 1 - \prod\limits_{i=1}^{n} (1 - F_{X_i}(x))$.

证明

$$\begin{aligned} F_{\max}(x) &= P(Y \leqslant x) = P(\max\{X_1, X_2, \cdots, X_n\} \leqslant x) \\ &= P(X_1 \leqslant x, X_2 \leqslant x, \cdots, X_n \leqslant x) \\ &= \prod_{i=1}^{n} P(X_i \leqslant x) = \prod_{i=1}^{n} F_{X_i}(x). \\ F_{\min}(x) &= 1 - P(Z > x) = 1 - P(\min\{X_1, X_2, \cdots, X_n\} > x) \\ &= 1 - P(X_1 > x, X_2 > x, \cdots, X_n > x) \\ &= 1 - \prod_{i=1}^{n} P(X_i > x) = 1 - \prod_{i=1}^{n} (1 - F_{X_i}(x)). \end{aligned}$$

特别地,如果 $X_1, X_2, \cdots, X_n$ 具有相同的分布函数 $F(x)$,则

$$F_{\max}(x) = F(x)^n, \quad F_{\min}(x) = 1 - (1 - F(x))^n.$$

例 3.4.8 设随机变量 $X_1, X_2, \cdots, X_n$ 相互独立,且 X_i 都服从区间[0,1]上的均匀分布,求 $Y = \max\{X_1, X_2, \cdots, X_n\}$ 与 $Z = \min\{X_1, X_2, \cdots, X_n\}$ 的密度函数.

解 X_i 的分布函数均为

$$F(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leqslant x < 1, \\ 1, & x \geqslant 1. \end{cases}$$

所以由例 3.4.7 知

$$F_{\max}(x)=F(x)^n=\begin{cases}0, & x<0,\\ x^n, & 0\leqslant x<1,\\ 1, & x\geqslant 1,\end{cases}$$

$$F_{\min}(x)=1-(1-F(x))^n=\begin{cases}0, & x<0,\\ 1-(1-x)^n, & 0\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$$

得

$$F_{\max}(x)=\begin{cases}nx^{n-1}, & 0\leqslant x<1,\\ 0, & \text{其他},\end{cases}$$

$$F_{\min}(x)=\begin{cases}n(1-x)^{n-1}, & 0\leqslant x<1,\\ 0, & \text{其他}.\end{cases}$$

习　　题

1. 设(X,Y)的联合分布律为

X \ Y	−1	1	2
−1	0.1	0.1	0.1
2	0.2	0.3	0.2

(1) 求 $X+2Y$ 的分布律；

(2) 求 X^2Y 的分布律；

(3) 求 $\min(X,Y)$的分布律.

2. 设(X,Y)的联合密度函数为

$$f(x,y)=\begin{cases}e^{-y}, & 0\leqslant x\leqslant 1, y\geqslant 0,\\ 0, & \text{其他}.\end{cases}$$

(1) 问 X,Y 是否独立？

(2) 求 $Z=2X+Y$ 的密度函数 $f_Z(z)$；

(3) 求 $P(Z>3)$.

3. 设 X 服从(0,1)上的均匀分布，Y 服从参数$\lambda=1$ 的指数分布，X,Y 相互独立，$Z=X+Y$，求 Z 的概率密度函数.

4. 设 X 和 Y 分别表示两个不同电子器件的寿命(单位:h)，并设 X 和 Y 相互独立，且服从同一分布，其概率密度为

$$f(x)=\begin{cases}\dfrac{1\ 000}{x^2}, & x>1\ 000,\\ 0, & \text{其他}.\end{cases}$$

求 $Z=\dfrac{X}{Y}$的概率密度.

5. 设某种型号的电子管的寿命(单位:h)近似地服从 $N(160,20^2)$分布,随机地选取 4 只,求其中至少有一只寿命小于 180 的概率.

6. 已知 X 与 Y 相互独立,且分别服从参数为 λ_1 和 λ_2 的泊松分布,求 $P(X=k\mid X+Y=n)$,其中 $k\leqslant n$.

§3.5 综合例题

例 3.5.1 设随机变量 X 与 Y 相互独立,且分别服从参数为 p_1 与 p_2 的几何分布,试求 $Z=\min\{X,Y\}$的分布列.

解 先求$\{Z\geqslant n\}$的概率,其中 $n\geqslant 1$.

$$\begin{aligned}P(Z\geqslant n)&=P(X\geqslant n,Y\geqslant n)=P(X\geqslant n)P(Y\geqslant n)\\&=\sum_{k=n}^{\infty}P(X=k)\sum_{k=n}^{\infty}P(Y=k)\\&=\sum_{k=n}^{\infty}p_1(1-p_1)^{k-1}\sum_{k=n}^{\infty}p_2(1-p_2)^{k-1}\\&=(1-p_1)^{n-1}(1-p_2)^{n-1}.\end{aligned}$$

所以

$$P(Z=n)=P(Z\geqslant n)-P(Z\geqslant n+1)=p(1-p)^{n-1},$$

其中 $p=p_1+p_2-p_1p_2$,Z 也服从几何分布.

例 3.5.2 设离散型随机变量 X 的分布列为 $P(X=i)=\dfrac{1}{3}$,$i=1,2,3$,连续型随机变量 $Y\sim U(0,1)$,且 X 与 Y 相互独立,求 $Z=X+Y$ 的密度函数.

解 本题是一个离散型随机变量加上一个连续型随机变量的问题,处理方法与以往的方法都不一样.

由题意可知 Z 的取值范围为$[1,4]$,所以当 $z<1$ 时,$F_Z(z)=0$;当 $z\geqslant 4$ 时,$F_Z(z)=1$;当 $1\leqslant z<4$ 时,

$$F_Z(z)=P(Z\leqslant z)=P(X+Y\leqslant z).$$

由于 X 的可能取值为 1,2,3,所以 $A_i=\{X=i\}$,$i=1,2,3$ 为样本空间的一个划分,于是对事件$\{X+Y\leqslant z\}$用全概率公式可得

$$
\begin{aligned}
F_Z(z) = P(X+Y \leqslant z) &= \sum_{i=1}^{3} P(X=i)P(X+Y \leqslant z \mid X=i) \\
&= \frac{1}{3}\sum_{i=1}^{3} P(Y \leqslant z-i \mid X=i) \\
&= \frac{1}{3}\sum_{i=1}^{3} P(Y \leqslant z-i) \\
&= \frac{1}{3}(F_Y(z-1)+F_Y(z-2)+F_Y(z-3)).
\end{aligned}
$$

所以

$$
f_Z(z) = \frac{1}{3}(f_Y(z-1)+f_Y(z-2)+f_Y(z-3)).
$$

注意到 $f_Y(y)=\begin{cases}1, & 0<y<1, \\ 0, & 其他,\end{cases}$ 经过简单的讨论可得

$$
f_Z(z) = \begin{cases}\dfrac{1}{3}, & 1<z<4, \\ 0, & 其他.\end{cases}
$$

即 Z 服从区间 $[1,4]$ 上的均匀分布.

一般地,可以证明一个离散型随机变量与一个连续型随机变量的和,在二者独立的情况下仍然是连续型随机变量.这一点可能与许多人的感觉相反.

例 3.5.3(正态分布的可加性) 设随机变量 X 与 Y 相互独立,且 $X\sim N(\mu_1,\sigma_1^2)$,$Y\sim N(\mu_2,\sigma_2^2)$,则 $Z=X+Y$ 服从正态分布 $N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$.

证明 由于 X 与 Y 相互独立,所以联合密度为

$$
f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2}\exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right\},\ x,y \in \mathbf{R}.
$$

$Z=X+Y$ 的分布函数为

$$
F_Z(z) = P(X+Y \leqslant z) = \int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{z-x} f(x,y)\mathrm{d}y,
$$

所以 $Z=X+Y$ 的密度函数为

$$
f_Z(z) = \int_{-\infty}^{+\infty} f(x,z-x)\mathrm{d}x = \int_{-\infty}^{+\infty}\frac{1}{2\pi\sigma_1\sigma_2}\exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}-\frac{(z-x-\mu_2)^2}{2\sigma_2^2}\right\}\mathrm{d}x,
$$

令 $s=x-\mu_1$,$b=z-\mu_1-\mu_2$,则

$$
\begin{aligned}
\frac{(x-\mu_1)}{\sigma_1^2}+\frac{(z-x-\mu_2)}{\sigma_2^2} = \frac{s^2}{\sigma_1^2}+\frac{(b-s)^2}{\sigma_2^2} &= \frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2}s^2-\frac{2bs}{\sigma_2^2}+\frac{b^2}{\sigma_2^2} \\
&= \left(\frac{\sqrt{\sigma_1^2+\sigma_2^2}}{\sigma_1\sigma_2}s-\frac{\sigma_1 b}{\sigma_2\sqrt{\sigma_1^2+\sigma_2^2}}\right)^2+\frac{b^2}{\sigma_1^2\sigma_2^2}.
\end{aligned}
$$

再令 $t=\dfrac{\sqrt{\sigma_1^2+\sigma_2^2}}{\sigma_1\sigma_2}s-\dfrac{\sigma_1 b}{\sigma_2\sqrt{\sigma_1^2+\sigma_2^2}}$,那么

$$f_Z(z)=\int_{-\infty}^{+\infty}f(x,z-x)\mathrm{d}x=\frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}}\exp\left\{-\frac{b^2}{2(\sigma_1^2+\sigma_2^2)}\right\}$$

$$\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{t^2}{2}\right\}\mathrm{d}t$$

$$=\frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}}\exp\left\{-\frac{(z-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}\right\}.$$

所以 Z 服从正态分布 $N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$.

其实我们可以证明更加一般的结论.

正态分布的可加性续 设随机变量 $X_1,X_2,\cdots,X_n$ 相互独立,且 $X_i\sim N(\mu_i,\sigma_i^2)$,$i=1,2,\cdots,n$,那么对于任意常数 a_i,$i=1,2,\cdots,n$ 及常数 c,有

$$\sum_{i=1}^n a_iX_i+c\sim N\left(\sum_{i=1}^n a_i\mu_i+c,\sum_{i=1}^n a_i^2\sigma_i^2\right).$$

复 习 题 3

1. 设 X 与 Y 是相互独立的随机变量,且 $X\sim B(n,p)$ 与 $Y\sim B(m,p)$. 证明 $X+Y\sim B(n+m,p)$.

2. 设(X,Y)的联合密度为 $f(x,y)=\begin{cases}\dfrac{1}{4}(1+xy), & -1<x,y<1,\\ 0, & 其他,\end{cases}$ 证明 X 与 Y 不独立,但是 X^2 与 Y^2 独立.

3. 设(X,Y)的联合密度为 $f(x,y)=\dfrac{1}{2\pi}\exp\left\{-\dfrac{x^2+y^2}{2}\right\}(1+\sin x\sin y)$,$x,y\in\mathbf{R}$,证明:

(1) X 与 Y 都服从标准正态分布,但(X,Y)不是二维正态分布;

(2) X 与 Y 不独立,但是 $|X|$ 与 $|Y|^2$ 独立.

4. 设 $X_1,X_2,\cdots,X_n$ 相互独立,且具有相同的分布函数 $F(x)$和密度函数 $f(x)$,试求 $Z=\max\{X_1,X_2,\cdots,X_n\}-\min\{X_1,X_2,\cdots,X_n\}$的密度函数.

5. 已知(X,Y)的联合分布列为

$$P(X=n,Y=m)=\frac{(p\lambda)^m(q\lambda)^{n-m}}{m!(n-m)!}\mathrm{e}^{-\lambda},$$

其中 $\lambda>0,q=1-p,n=0,1,2,\cdots,m=0,1,\cdots,n$. (1) 求 $Z=X-Y$ 的分布列;(2) 证明 Y 与 Z 相互独立.

6. 设 $X\sim U(0,1)$,且对 $0<x<1$,当 $X=x$ 时,$Y\sim U(0,x)$,(1) 求 $P(X>2Y)$;(2) 求 $Z=X-Y$ 的密度函数.

7. 设随机变量 X 和 Y 独立,其中 X 的概率分布为 $X\sim\begin{pmatrix}1 & 2\\ 0.3 & 0.7\end{pmatrix}$,而 Y 的概率密度为 $f(y)$,求随机变量 $Z=Y-X^2$ 的概率密度.

8. 设随机变量 X 和 Y 独立,其中 X 的概率分布为 $X\sim\begin{pmatrix}1 & 0\\ 0.3 & 0.7\end{pmatrix}$,而 $Y\sim U(0,2)$,求随机变量 $Z=XY$ 的分布函数.

第 4 章　随机变量的数字特征

前面学习的随机变量分布函数，是对随机变量概率性质的完整刻画，它可以全面描述随机变量的统计规律. 但在许多实际问题中，求随机变量的分布函数并不容易. 另一方面，有时也并不需要知道随机变量的分布函数，更关心的是它的某些特征，这些特征就是随机变量的数字特征. 例如考察某个班级学生的学习成绩时，通常只需知道该班的平均成绩和学习成绩的分散程度，就可对该班的学习状况做出较客观的判断. 本章将介绍随机变量常用的数字特征——数学期望、方差、协方差、相关系数等. 随机变量的数字特征是概率论的核心内容之一.

§4.1　随机变量的数学期望

4.1.1　离散型随机变量的数学期望

随机变量的数学期望是从求数据的平均数概念提炼推广而来的. 设以数据集{10,15,30,10,10,15,20,10,20,10}作为总体，则其平均数

$$\begin{aligned}\mu &= \frac{10+15+30+10+10+15+20+10+20+10}{10}\\ &= \frac{10\times 5+15\times 2+20\times 2+30\times 1}{10}\\ &= 10\times\frac{5}{10}+15\times\frac{2}{10}+20\times\frac{2}{10}+30\times\frac{1}{10}\\ &= 15,\end{aligned}$$

可将上式概括成简单的公式 $\mu=\sum_i x_i f_i$，其中 $\{x_i\}$ 为总体所有可能的不同值，上例中即为 10,15,20,30；f_i 为 x_i 在总体中的权重，也是 x_i 在总体中出现的频率，在上例中即为 $\frac{5}{10},\frac{2}{10},\frac{2}{10},\frac{1}{10}$. 平均数 μ 所描述的是数据集的中心位置.

将对求数据平均数的方法，引申到随机变量上，并注意到频率的稳定性，得知一般随机变量的“平均数”，应是随机变量所有可能的取值与其相应的概率乘积之和，这就是随机变量的数学期望. 下面分离散型和连续型两种情况讨论.

定义 4.1.1　设离散型随机变量 X 的分布律为

$$P\{X=x_k\}=p_k, k=1,2,3\cdots,$$

若级数 $\sum_{k=1}^{\infty} x_k p_k$ 绝对收敛，则称

$$EX = \sum_{k=1}^{+\infty} x_k p_k \tag{4.1.1}$$

为随机变量 X 的**数学期望**. 若级数 $\sum_{k=1}^{+\infty} |x_k| p_k$ 不收敛，则称 X 的数学期望不存在. 数学期望也称为**平均值**、**均值**.

数学期望相当于物理学中重心的概念. 显然，如果随机变量 X 只取有限个值，那么 EX 必存在，但如果 X 的取值是无限个，那就不一定了.

例 4.1.1　设随机变量 X 的分布列为

$$P\{X=k\} = \frac{6}{\pi^2}\frac{1}{k^2}, k=1,2,\cdots,$$

因为

$$\sum_{k=1}^{+\infty} k \frac{6}{\pi^2}\frac{1}{k^2} = \frac{6}{\pi^2}\sum_{k=1}^{+\infty}\frac{1}{k} = +\infty,$$

所以这个随机变量的期望不存在.

例 4.1.2　设 X 服从参数为 n,p 的二项分布，求 X 的数学期望 $E(X)$.

解　X 的分布律为 $P\{X=k\}=C_n^k p^k (1-p)^{n-k}, k=0,1,\cdots,n$，所以 X 的数学期望为

$$\begin{aligned} EX &= \sum_{k=0}^{n} k C_n^k p^k (1-p)^{n-k} = \sum_{k=1}^{n} n C_{n-1}^{k-1} p^k (1-p)^{n-k} \\ &= np\sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} (1-p)^{n-1-(k-1)} = np\,(p+1-p)^{n-1} \\ &= np. \end{aligned}$$

例 4.1.3　设 X 服从参数为 λ 的泊松分布，求 X 的数学期望 $E(X)$.

解　X 的分布律为 $P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda}, k=0,1,2,\cdots$，$X$ 的数学期望为

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

由此可知，泊松分布的参数 λ 就是它的数学期望.

例 4.1.4　设 X 服从参数为 p 的几何分布，求 X 的数学期望 $E(X)$.

解　X 的分布律为 $PX=k=p\,(1-p)^{k-1}, k=0,1,\cdots$，所以 X 的数学期望为

$$EX=\sum_{k=1}^{+\infty}kpq^{k-1}=p\left(\sum_{k=0}^{+\infty}q^k\right)'=p\left(\frac{1}{1-q}\right)'$$
$$=p\frac{1}{(1-q)^2}=\frac{1}{p}.$$

4.1.2 连续型随机变量的数学期望

如果随机变量 X 是连续型随机变量,X 的期望应该如何定义呢?在第二章介绍一维连续随机变量的时候,就已经知道,连续随机变量就相当于直线 $\mathbf{R}$ 上的质量分布.现在,既然随机变量的期望相当于重心,那么按照用质量线密度求重心的公式,我们就有下面的定义.

定义 4.1.2 设连续型随机变量 X 的密度函数为 $f(x)$,若积分 $\int_{-\infty}^{+\infty}|x|f(x)\mathrm{d}x$ 收敛,则称

$$E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x \tag{4.1.2}$$

为随机变量 X 的**数学期望**.如果 $\int_{-\infty}^{+\infty}|x|f(x)\mathrm{d}x=+\infty$,则称 X 的**数学期望不存在**.

例 4.1.5 设 X 服从均匀分布 $U(a,b)$,求 $E(X)$.

解 由题设可知,X 的概率密度为

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

于是

$$E(X)=\int_a^b\frac{1}{b-a}x\mathrm{d}x=\frac{a+b}{2}.$$

这个结果是显然的.因为 X 在$[a,b]$上均匀分布,它取值的平均值当然应该在区间$[a,b]$的中间,也就是$\frac{a+b}{2}$.

当然,连续随机变量的期望也可能不存在,比如:

例 4.1.6 设随机变量 X 的密度函数为 $f(x)=\frac{1}{\pi(1+x^2)}$,$x\in\mathbf{R}$,因为

$$\int_{-\infty}^{+\infty}|x|f(x)\mathrm{d}x=\int_{-\infty}^{+\infty}|x|\frac{1}{\pi(1+x^2)}\mathrm{d}x=\frac{1}{\pi}\int_0^{+\infty}\frac{1}{1+x^2}\mathrm{d}x^2$$
$$=\frac{1}{\pi}\ln(1+x^2)\Big|_0^{+\infty}=+\infty,$$

所以这个随机变量的期望不存在.

例 4.1.7　设 X 服从参数为 λ 的指数分布，求 X 的数学期望 $E(X)$.

解　X 的概率密度函数为 $f(x)=\begin{cases}\lambda e^{-\lambda x}, x>0,\\ 0, 其他,\end{cases}$

故由定义有

$$E(X)=\int_0^{+\infty} x\lambda e^{-\lambda x}dx=-\int_0^{+\infty} x de^{-\lambda x}=\int_0^{+\infty} e^{-\lambda x}dx=\frac{1}{\lambda}.$$

指数分布是最有用的"寿命分布"之一．由上述计算可知，一个元器件的寿命分布如果是参数为 λ 的指数分布，则它的平均寿命为$\frac{1}{\lambda}$. 如果某种元器件的平均寿命是10^k 小时，则相应的 $\lambda=10^{-k}$，在电子工业中，人们就称该产品是"k 级"产品．由此可知，k 越大，则产品的平均寿命越长，使用也就越可靠.

例 4.1.8　设随机变量 X 服从正态分布 $N(\mu,\sigma^2)$，求 X 的数学期望 $E(X)$.

解　由定义有

$$E(X)=\int_{-\infty}^{+\infty} x\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx,$$

令 $z=\frac{x-\mu}{\sigma}$，则

$$\begin{aligned}E(X)&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}(\sigma z+\mu)e^{-\frac{z^2}{2}}dz\\&=\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} ze^{-\frac{z^2}{2}}dz+\frac{\mu}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}}dz=\mu.\end{aligned}$$

由以上可知，正态分布 $N(\mu,\sigma^2)$中的参数 μ 恰好是正态分布的随机变量的数学期望.

4.1.3　随机变量函数的数学期望

在理论推导和现实问题研究中，经常需要计算随机变量函数的数学期望，例如 X^2，$|X|$，e^X 等的数学期望．我们可以先求出它们的分布律或密度函数，再求数学期望，但是那样做显得太繁琐．事实上，有如下结论可简化计算.

定理 4.1.1　设 Y 是随机变量 X 的函数，$Y=g(X)$($g(x)$是连续函数).

(1) 若 X 是离散型随机变量，它的分布律为 $P\{X=x_k\}=p_k(k=1,2,3,\cdots)$，且 $\sum\limits_{k=1}^{+\infty}g(x_k)p_k$ 绝对收敛，则随机变量 $Y=g(X)$的数学期望存在，为

$$E(Y)=E[g(x)]=\sum_{k=1}^{+\infty}g(x_k)p_k. \tag{4.1.3}$$

(2) 若 X 是连续型随机变量，它的概率密度为 $f(x)$，且$\int_{-\infty}^{+\infty}g(x)f(x)dx$ 绝

对收敛,则有

$$E(Y)=E[g(X)]=\int_{-\infty}^{+\infty}g(x)f(x)\mathrm{d}x. \tag{4.1.4}$$

特别地,称随机变量 X 的 k 次方的数学期望 $E(X^k)$ 为 X 的 k 阶原点矩,称 $X-E(X)$ 的 k 次方的期望 $E\{[X-E(X)]^k\}$ 为 X 的 k 阶中心矩.

该定理的证明超出了本书的范围,故从略.

例 4.1.9 设 $X\sim P(5)$,求 $E(3^X)$ 与 $E\left(\frac{1}{X+1}\right)$.

解 因为 3^X 与 $\frac{1}{X+1}$ 均为 X 的函数,且 X 为离散型随机变量. 所以

$$E(3^X)=\sum_{k=0}^{+\infty}3^k\frac{5^k}{k!}\mathrm{e}^{-5}=\mathrm{e}^{-5}\sum_{k=0}^{+\infty}\frac{15^k}{k!}=\mathrm{e}^{-5}\mathrm{e}^{15}=\mathrm{e}^{10},$$

$$E\left(\frac{1}{X+1}\right)=\sum_{k=0}^{+\infty}\frac{1}{k+1}\cdot\frac{5^k}{k!}\mathrm{e}^{-5}=\frac{1}{5}\sum_{k=0}^{+\infty}\frac{5^{k+1}}{(k+1)!}\mathrm{e}^{-5}$$

$$=\frac{\mathrm{e}^{-5}}{5}\sum_{k=1}^{+\infty}\frac{5^k}{k!}=\frac{\mathrm{e}^{-5}}{5}(\mathrm{e}^5-1)=\frac{1}{5}(1-\mathrm{e}^{-5}).$$

例 4.1.10 对球体的直径作近似测量,设测量值均匀分布在区间 $[a,b]$ 内,求球体体积的数学期望.

解 设随机变量 X 表示球的直径,Y 表示球的体积. 依题意,X 的概率密度为

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

球体积 $Y=\frac{1}{6}\pi X^3$,则有

$$E(Y)=E\left(\frac{1}{6}\pi X^3\right)=\int_a^b\frac{1}{6}\pi x^3\frac{1}{b-a}\mathrm{d}x$$

$$=\frac{\pi}{6(b-a)}\int_a^b x^3\mathrm{d}x=\frac{\pi}{24}(a+b)(a^2+b^2).$$

4.1.4 二维随机变量函数的数学期望

定理 4.1.1 可以推广到二维随机变量函数的情形.

定理 4.1.2 设 Z 是随机变量 X,Y 的函数,$Z=g(X,Y)$(g 为连续函数),那么 Z 也是一个随机变量.

(1) 当 (X,Y) 为二维离散型随机变量,其分布律为 $P\{X=x_i,Y=y_j\}=p_{ij}$,

$i,j=1,2,\cdots$. 若 $\sum\limits_{i=1}^{+\infty}\sum\limits_{j=1}^{+\infty}|g(x_i,y_j)|p_{ij}<+\infty$，则随机变量 $Z=g(X,Y)$ 的数学期望存在，且

$$E(Z)=E[g(X,Y)]=\sum_{i=1}^{+\infty}\sum_{j=1}^{+\infty}g(x_i,y_j)p_{ij}. \tag{4.1.5}$$

(2) 当 (X,Y) 为二维连续型随机变量，联合密度为 $f(x,y)$，若 $\int_{-\infty}^{+\infty}\mathrm{d}x\cdot\int_{-\infty}^{+\infty}|g(x,y)|f(x,y)\mathrm{d}y<+\infty$，则随机变量 $Z=g(X,Y)$ 的数学期望存在，且

$$E(Z)=E(g(X,Y))=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y \tag{4.1.6}$$

由定理 4.1.2 可知：

若 (X,Y) 为离散型随机变量，其概率分布为

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

则

$$E(X)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}x_ip_{ij},\quad E(Y)=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}y_jp_{ij}.$$

若 (X,Y) 为连续型随机变量，有联合概率密度函数 $f(x,y)$，且 X 与 Y 都是连续型随机变量，则有

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

以上结论还可以推广到 n 维随机变量.

例 4.1.11　设二维随机变量 (X,Y) 的联合分布律为

$X \backslash Y$	0	1	2	$p_{i\cdot}$
-1	$\frac{1}{9}$	$\frac{2}{9}$	0	$\frac{1}{3}$
2	$\frac{3}{9}$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{2}{3}$
$p_{\cdot j}$	$\frac{4}{9}$	$\frac{3}{9}$	$\frac{2}{9}$	1

试求 $E(X),E(X^2),E(Y),E(Y^2),E(XY),E(X+3)^Y$.

解　X 和 Y 的边缘分布律见上表，那么

$$E(X)=-1\times\frac{1}{3}+2\times\frac{2}{3}=1,$$

$$E(X^2)=(-1)^2\times\frac{1}{3}+2^2\times\frac{2}{3}=3,$$

$$E(Y)=0\times\frac{4}{9}+1\times\frac{3}{9}+2\times\frac{2}{9}=\frac{7}{9},$$

$$E(Y^2)=0^2\times\frac{4}{9}+1^2\times\frac{3}{9}+2^2\times\frac{2}{9}=\frac{11}{9},$$

$$\begin{aligned}E(XY)&=(-1)\times0\times\frac{1}{9}+(-1)\times1\times\frac{2}{9}+(-1)\times2\times0+2\times0\times\frac{3}{9}\\&\quad+2\times1\times\frac{1}{9}+2\times2\times\frac{2}{9}\\&=\frac{8}{9},\end{aligned}$$

$$\begin{aligned}E(X+3)^Y&=(-1+3)^0\times\frac{1}{9}+(-1+3)^1\times\frac{2}{9}+(-1+3)^2\times0\\&\quad+(2+3)^0\times\frac{3}{9}+(2+3)^1\times\frac{1}{9}+(2+3)^2\times\frac{2}{9}\\&=7.\end{aligned}$$

例 4.1.12 设随机变量(X,Y)的联合密度为 $f(x,y)=\begin{cases}e^{-y}, & 0<x<y<\infty,\\ 0, & 其他,\end{cases}$ 求 $E(X)$,$E(X^2)$,$E(Y)$,$E(Y^2)$,$E(XY)$.

解 (X,Y)的联合密度的非零区域为第一象限中分角线的上方,

所以,

$$E(X)=\int_0^{+\infty}dx\int_x^{+\infty}xe^{-y}dy=\int_0^{+\infty}xdx\int_x^{+\infty}e^{-y}dy=\int_0^{+\infty}xe^{-x}\,dx=1,$$

$$E(X^2)=\int_0^{+\infty}dx\int_x^{+\infty}x^2e^{-y}dy=\int_0^{+\infty}x^2dx\int_x^{+\infty}e^{-y}dy=\int_0^{+\infty}x^2e^{-x}dx=2,$$

$$E(Y)=\int_0^{+\infty}dy\int_0^y ye^{-y}dx=\int_0^{+\infty}y^2e^{-y}dy=2,$$

$$E(Y^2)=\int_0^{+\infty}dy\int_0^y y^2e^{-y}dx=\int_0^{+\infty}y^3e^{-y}dy=6,$$

$$E(XY)=\int_0^{+\infty}dy\int_0^y xye^{-y}dx=\frac{1}{2}\int_0^{+\infty}y^3e^{-y}dy=3.$$

4.1.5 数学期望的性质

(1) $E(c)=c$,其中 c 为常数.

常数 c 可以看成只取一个值 c 的离散型随机变量,且取 c 的概率为 1,于是

$$E(c)=c\times1=c.$$

(2) $E(cX)=cE(X)$.

不妨设 X 为连续型随机变量,密度为 $f(x)$,cX 为随机变量 X 的函数,所以

$$E(cX)=\int_{-\infty}^{+\infty} cx\cdot f(x)\mathrm{d}x = c\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x = cE(X).$$

(3) $E(X_1+X_2)=E(X_1)+E(X_2)$.

证明 不妨设(X_1,X_2)为二维连续型随机变量,联合密度为 $f(x,y)$,所以

$$\begin{aligned}E(X_1+X_2)&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)f(x,y)\mathrm{d}x\mathrm{d}y\\&=\int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{+\infty}xf(x,y)\mathrm{d}y+\int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{+\infty}yf(x,y)\mathrm{d}y\\&=E(X_1)+E(X_2).\end{aligned}$$

推论 设 $X_1,X_2,\cdots,X_n$ 为 n 个随机变量($n\geqslant 2$),则

$$E\Big(\sum_{i=1}^{n}X_i\Big)=\sum_{i=1}^{n}E(X_i). \tag{4.1.7}$$

(4) 若随机变量 X 与 Y 相互独立,则

$$E(XY)=E(X)E(Y). \tag{4.1.8}$$

证明 不妨设(X,Y)为二维连续型随机变量,由于 X 和 Y 相互独立,所以(X,Y)的联合密度为 $f(x,y)=f_X(x)f_Y(y)$,于是

$$\begin{aligned}E(XY)&=\int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{+\infty}xyf(x,y)\mathrm{d}y\\&=\int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{+\infty}xyf_X(x)f_Y(y)\mathrm{d}y\\&=\int_{-\infty}^{+\infty}xf_X(x)\mathrm{d}x\int_{-\infty}^{+\infty}yf_Y(y)\mathrm{d}y\\&=E(X)E(Y).\end{aligned}$$

性质(4)可以推广为:若 X 和 Y 相互独立,$g(X)$与$h(Y)$分别为 X 和 Y 的函数,则

$$E[g(X)h(Y)]=E[g(X)]\cdot E[h(Y)].$$

例 4.1.13 设 X 服从二项分布 $B(n,p)$,求 $E(X)$.

解 考虑 n 重伯努利试验,若令

$$X_k=\begin{cases}0,\text{若第 } k \text{ 次试验不成功},\\1,\text{若第 } k \text{ 次试验成功}.\end{cases}$$

则有 $X=\sum_{k=1}^{n}X_k$,再由数学期望性质(3)的推论,有 $E(X)=\sum_{k=1}^{n}E(X_k)$.

因为每个 X_k 均服从相同的伯努利分布,数学期望为 $E(X_k)=p$,从而得到 $E(X)=np$.

将此题与例 4.1.3 比较会发现,如果不用数学期望的性质,而用 X 的分布律直接按定义来计算,要复杂得多.

习 题

1. 已知甲乙两箱中装有同种产品，其中甲箱中装有 3 件合格品和 3 件次品，乙箱中仅有 3 件合格品. 从甲箱中任取 3 件产品放入乙箱后，求乙箱中次品件数 X 的数学期望.

2. 设随机变量 X 有分布律

X	-2	-1	0	1
P	0.2	0.3	0.4	0.1

求 $E(X)$，$E(X+1)$，$E(2X^2+1)$.

3. 设随机变量 X 的分布律为

X	1	0	-1
P	p_1	p_2	p_3

且已知 $E(X)=0.1$，$E(X^2)=0.9$，求 p_1，p_2，p_3.

4. 设随机变量 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases}k, & 0<x<1, 0<y<x,\\ 0, & \text{其他}.\end{cases}$$

试确定常数 k，并求 $E(XY)$.

5. 设随机变量 X 在区间 $[a,b]$ 上服从均匀分布，$E(X)=1$，$E(X^2)=4$，试求 a 和 b.

6. 假设由自动线加工的某种零件的内径 X(mm) 服从正态分布 $N(\mu,1)$，内径小于 10 或大于 12 为不合格品，其余为合格品. 销售每件合格品获利，销售每件不合格品亏损，已知销售利润 T(元) 与销售零件的内径 X 有如下关系

$$T=\begin{cases}-1, & X<10,\\ 20, & 10\leqslant X\leqslant 12,\\ -5, & X>12.\end{cases}$$

问：平均直径 μ 取何值时，销售一个零件的平均利润最大？

7. 某产品的次数率为 0.1，检验员每天检验 4 次，每次随机地取 10 件产品进行检验. 设各产品是否为次品是相互独立的，如发现其中的次品数多于 1，就去调整设备，以 X 表示一天中调整设备的次数，试求 $E(X)$.

8. 一工厂生产的某种设备的寿命 X (年)服从参数为 $\frac{1}{4}$ 的指数分布. 工厂规定，出售的设备在售出一年之内损坏可予以调换. 若工厂售出一台设备盈利 100 元，调换一台设备厂方需花费 300 元，求工厂出售一台净盈利的数学期望.

§4.2 方　差

4.2.1 随机变量的方差

数学期望描述了随机变量取值的“平均”，但有时仅知道这个平均值还不够.

比如，有A，B两名射手，他们每次射击命中的环数分别为X，Y，已知X，Y的分布律分别为：

表4-1　A　射　手

X	8	9	10
P	0.2	0.6	0.2

表4-2　B　射　手

Y	8	9	10
P	0.1	0.8	0.2

容易计算出$E(X)=E(Y)=9$(环)，即二人的平均成绩相同. 这样仅从均值的角度分不出谁的射击技术更高，故还需考虑其他的因素. 通常的想法是：还要考虑谁的射击技术更加稳定些. 这时用命中的环数X与它的平均值$E(X)$之间的离差$|X-E(X)|$的均值$E[|X-E(X)|]$来度量，即$E[|X-E(X)|]$愈小，表明X的值愈集中于$E(X)$的附近，技术稳定；$E[|X-E(X)|]$愈大，表明X的值很分散，技术不稳定. 但由于$E[|X-E(X)|]$带有绝对值，运算不便，故采用X与$E(X)$的离差$|X-E(X)|$的平方平均值$E\{[X-E(X)]^2\}$来度量随机变量X取值的分散程度.

定义　设X为随机变量，若$E\{[X-E(X)]^2\}$存在，则称$E\{[X-E(X)]^2\}$为X的方差，记作$D(X)$，即

$$D(X)=E\{[X-E(X)]^2\}. \tag{4.2.1}$$

称$\sqrt{D(X)}$为X标准差(或者均方差)，记作σ_X，即$\sigma_X=\sqrt{D(X)}$.

方差$D(X)$是随机变量X的取值相对于均值偏离程度的一种度量.

由定义知方差实际上是随机变量X的函数$g(X)=[X-E(X)]^2$的数学期望.

对于离散型随机变量，有$D(X)=\sum\limits_i[x_i-E(X)]^2p_i$，式中$P\{X=x_i\}=p_i(i=1,2,\cdots)$是$X$的分布律.

对连续型随机变量X，若$f(x)$是它的概率密度函数，则，

$$D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2f(x)\mathrm{d}x.$$

在实际计算中，用定义计算方差不是很方便，而是常用下面的方法.

$$\begin{aligned}D(X)&=E[X-E(X)]^2=E\{X^2-2X\cdot E(X)+[E(X)]^2\}\\&=E(X^2)-2E(X)\cdot E(X)+E[E(X)]^2\\&=E(X^2)-2[E(X)]^2+[E(X)]^2\\&=E(X^2)-[E(X)]^2,\end{aligned}$$

即

$$D(X)=E(X^2)-[E(X)]^2. \tag{4.2.2}$$

例 4.2.1　随机变量 X 的分布律如下：

X	-1	0	$\frac{1}{2}$	1	2
P	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{4}$

求 $E(X),E(X^2),E(2X-3),D(X)$.

解　$E(X)=(-1)\times\frac{1}{3}+0\times\frac{1}{6}+\frac{1}{2}\times\frac{1}{6}+1\times\frac{1}{12}+2\times\frac{1}{4}=\frac{1}{3}$,

$$E(X^2)=(-1)^2\times\frac{1}{3}+0^2\times\frac{1}{6}+\left(\frac{1}{2}\right)^2\times\frac{1}{6}+1^2\times\frac{1}{12}+2^2\times\frac{1}{4}=\frac{35}{24},$$

$$E(2X-3)=2E(X)-3=2\times\frac{1}{3}-3=-\frac{7}{3},$$

$$D(X)=E(X^2)-[E(X)]^2=\frac{35}{24}-\left(\frac{1}{3}\right)^2=\frac{97}{72}.$$

例 4.2.2　设随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\dfrac{3}{(x+1)^4}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

求 $E(X)$及 $D(X)$.

解　$E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\int_0^{+\infty}\frac{3x}{(x+1)^4}\mathrm{d}x=\int_0^{+\infty}\frac{3}{(x+1)^3}\mathrm{d}x-\int_0^{+\infty}\frac{3}{(x+1)^4}\mathrm{d}x=\frac{1}{2}$,

又

$$E(X^2)=\int_0^{+\infty}\frac{3x^2}{(x+1)^4}\mathrm{d}x=\int_0^{+\infty}\frac{3(x-1)}{(x+1)^3}\mathrm{d}x+\int_0^{+\infty}\frac{3}{(x+1)^4}\mathrm{d}x=1,$$

故

$$D(X)=E(X^2)-[E(X)]^2=\frac{3}{4}.$$

4.2.2 方差的性质

由于方差是用数学期望定义的，故由数学期望的性质很容易推得方差的一些重要性质.

设随机变量 X 与 Y 的方差存在：

(1) 若 c 为常数，则 $D(c)=0$；

(2) 若 a,b 为常数，X 为随机变量，则

$$D(aX+b)=a^2D(X); \tag{4.2.3}$$

(3) 若随机变量 X 和 Y 相互独立，则

$$D(X+Y)=D(X)+D(Y). \tag{4.2.4}$$

性质(3)还可以推广如下：

如果随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，则

$$D\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}D(X_i). \tag{4.2.5}$$

下面证明性质(2)和(3).

证明 (2) $$\begin{aligned}D(aX+b)&=E\{[aX+b-E(aX+b)]^2\}\\&=E\{[aX-aE(X)]^2\}\\&=a^2E\{[X-E(X)]^2\}\\&=a^2D(X).\end{aligned}$$

(3) 由于 X 与 Y 相互独立，可知 $E(XY)=E(X)\cdot E(Y)$，所以

$$\begin{aligned}D(X+Y)&=E[(X+Y)^2]-[E(X+Y)]^2\\&=E(X^2+2XY+Y^2)-\{[E(X)]^2+2E(X)E(Y)+[E(Y)]^2\}\\&=D(X)+D(Y)+2E(XY)-2E(X)E(Y)\\&=D(X)+D(Y).\end{aligned}$$

4.2.3 常见分布的随机变量的期望与方差

1. 0—1分布

设 X 服从0—1分布，则它的分布列为 $X\sim\begin{pmatrix}0 & 1\\1-p & p\end{pmatrix}$，所以

$$E(X)=p,E(X^2)=0^2\times(1-p)+1^2\times p=p,$$
$$D(X)=E(X^2)-[E(X)]^2=p-p^2=p(1-p).$$

2. 泊松分布

设 X 服从参数为 λ 的泊松分布，我们已经知道 $E(X)=\lambda$，又

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

所以

$$D(X) = E(X^2) - [E(X)]^2 = \lambda.$$

由此可以看出，泊松分布中的参数 λ，既是相应随机变量 X 的数学期望，又是它的方差，泊松分布可由他的数学期望或方差唯一决定.

3. 几何分布

X 的分布律为 $P\{X=k\}=p(1-p)^{n-1}, k=0,1,\cdots$，由前知 $EX=\dfrac{1}{p}$，又

$$\begin{aligned}
E(X^2) &= \sum_{k=1}^{+\infty} k^2 pq^{k-1} = \sum_{k=1}^{+\infty} k(k-1)pq^{k-1} + \sum_{k=1}^{+\infty} kpq^{k-1} \\
&= q\sum_{k=1}^{+\infty} k(k-1)pq^{k-2} + \frac{1}{p} \\
&= pq\left(\sum_{k=0}^{+\infty} q^k\right)'' + \frac{1}{p},
\end{aligned}$$

由于 $\left(\sum\limits_{k=0}^{+\infty} q^k\right)'' = \left(\dfrac{1}{1-q}\right)'' = \dfrac{2}{p^3}$，所以 $E(X^2) = \dfrac{2q+p}{p^2} = \dfrac{q+1}{p^2}$，于是

$$DX = E(X^2) - (EX)^2 = \frac{q}{p^2}.$$

4. 均匀分布

设 X 服从均匀分布 $U(a,b)$，$E(X)=\dfrac{a+b}{2}$，而

$$E(X^2) = \int_a^b \frac{1}{b-a} x^2 \mathrm{d}x = \frac{a^2+ab+b^2}{3},$$

故

$$D(X) = E(X^2) - [E(X)]^2 = \frac{(b-a)^2}{12}.$$

5. 指数分布

设 X 服从参数为λ 的指数分布，$E(X)=\frac{1}{\lambda}$，而

$$E(X^2)=\int_0^{+\infty}x^2\lambda e^{-\lambda x}dx=-\int_0^{+\infty}x^2 de^{-\lambda x}=\int_0^{+\infty}2xe^{-\lambda x}dx=\frac{2}{\lambda^2},$$

故

$$D(X)=\frac{2}{\lambda^2}-\left(\frac{1}{\lambda}\right)^2=\frac{1}{\lambda^2}.$$

可见，指数分布也由期望或方差唯一决定.

6. 正态分布

设随机变量 X 服从正态分布 $N(\mu,\sigma^2)$，则

$$D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2f(x)dx=\int_{-\infty}^{+\infty}(x-\mu)^2\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx,$$

令 $z=\frac{x-\mu}{\sigma}$，则 $D(X)=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}z^2e^{-\frac{z^2}{2}}dz$

$$=\frac{\sigma^2}{\sqrt{2\pi}}\left\{\left[-ze^{-\frac{z^2}{2}}\right]_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty}e^{-\frac{z^2}{2}}dz\right\}$$

$$=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{z^2}{2}}dz=\sigma^2 .$$

由此可知，正态分布 $N(\mu,\sigma^2)$中的参数 μ,σ^2，分别表示相应随机变量的数学期望和方差，于是正态分布由他的数学期望和方差唯一决定.

习　　题

1. 随机变量 X 的分布函数为

$$F(x)=\begin{cases}1-\frac{1}{x^3}, & x\geqslant 1,\\ 0, & x<1,\end{cases}$$

求 X 的数学期望及其方差.

2. 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}x, & 0<x\leqslant 1,\\ 2-x, & 1<x\leqslant 2,\\ 0, & \text{其他},\end{cases}$$

求 $E(X)$和 $D(X)$.

3. 已知随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}ax^2+bx+c, & 0\leqslant x\leqslant 1,\\ 0, & \text{其他},\end{cases}$$

又已知 $E(X)=0.5, D(X)=0.15$，求 a,b,c.

4. 设随机变量 X 的概率密度为

$$f(x)=\begin{cases} cx\mathrm{e}^{-k^2x^2}, & x\geq 0, \\ 0, & x<0, \end{cases}$$

求(1) 常数数 c;(2) EX;(3) DX.

5. 设二维随机变量(X,Y)联合分布律为

X \ Y	1	2	3
1	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{18}$
2	$\frac{1}{3}$	$\frac{1}{9}$	$\frac{2}{9}$

求 $E(X),D(X),E(Y),D(Y)$.

6. 设随机变量 X 服从参数为 2 的指数分布，试求 $E(\mathrm{e}^{-3X})$与 $D(\mathrm{e}^{-3X})$.

7. 设(X,Y)联合概率密度为

$$f(x,y)=\begin{cases} 24(1-x)y, & 0<x<1, 0<y<x, \\ 0, & \text{其他}, \end{cases}$$

求 $E(X),D(X),E(Y),D(Y)$.

§4.3 协方差和相关系数

4.3.1 协方差

对二维随机变量(X,Y)来说，期望 $E(X),E(Y)$只反映了 X 和 Y 各自的平均值，$D(X),D(Y)$反映的是 X,Y 各自与均值的偏离程度，它们对 X 与 Y 之间的相互联系不提供任何信息.我们自然希望有一个数字特征能够在一定程度上反映 X 与 Y 之间的联系.

由数学期望的性质可以得到，当 X 与 Y 相互独立时，有

$$\begin{aligned} & E\{[X-E(X)][Y-E(Y)]\} \\ = & E[XY-Y\cdot E(X)-X\cdot E(Y)+E(X)E(Y)] \\ = & E(XY)-E(X)E(Y)=0. \end{aligned}$$

换言之，当 $E\{[X-E(X)][Y-E(Y)]\}\neq 0$ 时，X 与 Y 肯定不独立，这说明 $E\{[X-E(X)][Y-E(Y)]\}$的数值大小在一定程度上反映了 X 与 Y 相互间的联系.

定义 4.3.1 设(X,Y)为二维随机变量，若 $E\{[X-E(X)][Y-E(Y)]\}$存

在，则称其为随机变量 X 和 Y 的**协方差**，记为 $\mathrm{Cov}(X,Y)$. 即

$$\mathrm{Cov}(X,Y) = E\{[X-E(X)][Y-E(Y)]\}. \tag{4.3.1}$$

在实际计算协方差时，经常使用下面的实用公式：

$$\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y). \tag{4.3.2}$$

读者可以很容易证明式(4.3.1)与式(4.3.2)是等价的.

下面介绍一下协方差的基本性质：

(1) 对称性：$\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$.

(2) $\mathrm{Cov}(X,X)=D(X)$.

(3) 若 a,b 为常数，则 $\mathrm{Cov}(aX,bY)=ab\cdot\mathrm{Cov}(X,Y)$.

(4) 若 X 与 Y 相互独立，则 $\mathrm{Cov}(X,Y)=0$.

(5) $\mathrm{Cov}(X_1+X_2,Y)=\mathrm{Cov}(X_1,Y)+\mathrm{Cov}(X_2,Y)$.

(6) 随机变量和的方差与协方差的关系：

$$D(aX+bY) = a^2D(X)+b^2D(Y)+2ab\mathrm{Cov}(X,Y). \tag{4.3.3}$$

下面仅证明性质(5)和性质(6).

证明　(5)
$$\begin{aligned}\mathrm{Cov}(X_1+X_2,Y)&=E[(X_1+X_2)Y]-E(X_1+X_2)E(Y)\\&=E(X_1Y)+E(X_2Y)-E(X_1)E(Y)-E(X_2)E(Y)\\&=[E(X_1Y)-E(X_1)E(Y)]+[E(X_2Y)-\\&\quad E(X_2)E(Y)]\\&=\mathrm{Cov}(X_1,Y)+\mathrm{Cov}(X_2,Y).\end{aligned}$$

(6) 利用协方差的性质得

$$\begin{aligned}D(aX+bY)&=\mathrm{Cov}(aX+bY,aX+bY)\\&=\mathrm{Cov}(aX,aX)+\mathrm{Cov}(aX,bY)+\mathrm{Cov}(bY,aX)+\mathrm{Cov}(bY,bY)\\&=a^2D(X)+b^2D(Y)+2ab\mathrm{Cov}(X,Y).\end{aligned}$$

性质(5)可以推广为：

设 $Y,X_1,X_2,\cdots,X_n$ 都是随机变量，则

$$\mathrm{Cov}\left(\sum_{i=1}^n X_i,Y\right) = \sum_{i=1}^n \mathrm{Cov}(X_i,Y). \tag{4.3.4}$$

例 4.3.1　已知二维随机变量(X,Y)的联合密度函数为

$$f(x,y) = \begin{cases}2, 0<x<y, 0<y<1,\\ 0, \text{其他}.\end{cases}$$

求：(1) $E(X),E(Y),D(X),D(Y),\mathrm{Cov}(X,Y)$；(2) $D(2X-Y+5)$.

解　(1) $$E(X) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf(x,y)\mathrm{d}x\mathrm{d}y = 2\int_0^1 x\mathrm{d}x\int_x^1\mathrm{d}y = \frac{1}{3}.$$

$$E(Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} yf(x,y)\mathrm{d}x\mathrm{d}y = 2\int_0^1 \mathrm{d}x\int_x^1 y\mathrm{d}y = \frac{2}{3}.$$

又 $$E(X^2)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}x^2f(x,y)\mathrm{d}x\mathrm{d}y=2\int_0^1 x^2\mathrm{d}x\int_x^1\mathrm{d}y=\frac{1}{6},$$

$$E(Y^2)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}y^2f(x,y)\mathrm{d}x\mathrm{d}y=2\int_0^1 y^2\mathrm{d}y\int_0^y\mathrm{d}x=\frac{1}{2},$$

$$E(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)\mathrm{d}x\mathrm{d}y=2\int_0^1 x\mathrm{d}x\int_x^1 y\mathrm{d}y=\frac{1}{4},$$

故 $$D(X)=E(X^2)-[E(X)]^2=\frac{1}{18},$$

$$D(Y)=E(Y^2)-[E(Y)]^2=\frac{1}{18},$$

$$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=\frac{1}{36}.$$

(2) $D(2X-Y+5)=D(2X-Y)=4D(X)+D(Y)-4\mathrm{Cov}(X,Y)=\frac{1}{6}$.

4.3.2 相关系数的定义与性质

用协方差来刻画两个随机变量之间关系，有时显得并不很清晰、直观. 例如，随机变量 X 与 Y 有某种线性关系时，$2X$ 与 $2Y$ 也应该有完全相同的线性关系. 但 $2X$ 与 $2Y$ 之间的协方差 $\mathrm{Cov}(2X,2Y)=4\mathrm{Cov}(X,Y)$，是 X 与 Y 之间的协方差 $\mathrm{Cov}(X,Y)$ 的 4 倍，这容易引起误解. 可见仅仅用协方差来描述 X 与 Y 的关系并不完整，为解决此问题，下面引入相关系数的概念.

定义 4.3.2 设 (X,Y) 为二维随机变量，且 $D(X)>0$，$D(Y)>0$，则称

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}} \tag{4.3.5}$$

为随机变量 X 和 Y 的相关系数.

关于相关系数，有如下重要性质：

(1) $|\rho_{XY}|\leqslant 1$.

(2) 当 $\rho_{XY}=1$ 时，则存在正常数 a 与实数 b 使 $Y=aX+b$.

证明 (1) 对任意实数 λ，

$$D(\lambda X+Y)=\lambda^2D(X)+2\lambda\mathrm{Cov}(X,Y)+D(Y)\geqslant 0.$$

这是关于 λ 的一元二次函数，由于其恒大于等于零，故其判别式应小于等于 0，即

$$4\mathrm{Cov}^2(X,Y)-4D(X)D(Y)\leqslant 0,$$

故有 $$|\rho_{X,Y}|\leqslant 1.$$

(2) 由 $\rho_{XY}=1$ 知，$\mathrm{Cov}(X,Y)=\sqrt{D(X)}\sqrt{D(Y)}$，所以

$$D[\sqrt{D(X)}Y-\sqrt{D(Y)}X]=D(X)D(Y)-2\sqrt{D(X)D(Y)}\cdot\mathrm{Cov}(X,Y)$$

$$+D(Y)\cdot D(X)$$
$$=D(X)D(Y)-2D(X)\cdot D(Y)+D(X)\cdot D(Y)$$
$$=0.$$

当一个随机变量的方差为零时，可以认为这个随机变量就是一个常数，所以 $\sqrt{D(X)}Y-\sqrt{D(Y)}X$ 就是一个常数. 又由于常数的数学期望为常数本身，因此

$$\sqrt{D(X)}Y-\sqrt{D(Y)}X=E[\sqrt{D(X)}Y-\sqrt{D(Y)}X]$$
$$=E(Y)\sqrt{D(X)}-E(X)\sqrt{D(Y)},$$

即
$$\frac{Y-E(Y)}{\sqrt{D(Y)}}=\frac{X-E(X)}{\sqrt{D(X)}}.$$

所以
$$Y=\frac{\sqrt{D(Y)}}{\sqrt{D(X)}}X+E(Y)-\frac{\sqrt{D(Y)}}{\sqrt{D(X)}}E(X).$$

令　$a=\frac{\sqrt{D(Y)}}{\sqrt{D(X)}}>0$，$b=E(Y)-\frac{\sqrt{D(Y)}}{\sqrt{D(X)}}E(X)$，则有

$$Y=aX+b.$$

当 $\rho_{XY}=1$ 时，称 X 和 Y **正线性相关**，即存在常数 $a>0$ 与实数 b，使

$$Y=aX+b.$$

同理，当 $\rho_{XY}=-1$ 时，称 X 和 Y **负线性相关**，即存在常数 $a<0$ 与实数 b，使

$$Y=aX+b.$$

特别地，如果 $\rho_{XY}=0$，则称 X 和 Y **不相关**.

值得注意的是，当 X 和 Y 相互独立时 $\rho_{XY}=0$，即 X 和 Y 不相关；反之却不一定成立. 由相关系数的性质(2)可以看出，$|\rho|$ 越接近 1，说明 X 和 Y 的线性关系越强，$|\rho|$ 越接近 0，说明 X 和 Y 的线性关系越弱. 那么如果 X 和 Y 互不相关，就可以认为 X 和 Y 没有线性关系，但 X 和 Y 可能有其他的非线性关系，而 X 和 Y 独立是指 X 和 Y 任何关系都没有. 可见“不相关”是比“独立”要弱的一个概念.

例 4.3.2　设二维正态随机变量 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$，求 $\mathrm{Cov}(X,Y)$ 和相关系数 ρ_{XY}.

解　因 $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，所以 $E(X)=\mu_1$，$E(Y)=\mu_2$.

$$\mathrm{Cov}(X,Y)=E((X-E(X))(Y-E(Y)))$$
$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_1)(y-\mu_2)f(x,y)\mathrm{d}x\mathrm{d}y$$
$$=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_1)(y-\mu_2)\cdot$$

$$\mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]}\mathrm{d}x\mathrm{d}y$$

$$=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_1)(y-\mu_2)\cdot$$

$$\mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left(\frac{x-\mu_1}{\sigma_1}-\rho\frac{y-\mu_2}{\sigma_2}\right)^2-\frac{1}{2}\left(\frac{y-\mu_2}{\sigma_2}\right)^2}\mathrm{d}x\mathrm{d}y$$

$$=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}(y-\mu_2)\mathrm{e}^{-\frac{1}{2}\left(\frac{y-\mu_2}{\sigma_2}\right)^2}\cdot$$

$$\left(\int_{-\infty}^{+\infty}(x-\mu_1)\mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left(\frac{x-\mu_1}{\sigma_1}-\rho\frac{y-\mu_2}{\sigma_2}\right)^2}\mathrm{d}x\right)\mathrm{d}y.$$

令 $u=\frac{1}{\sqrt{1-\rho^2}}\left(\frac{x-\mu_1}{\sigma_1}-\rho\frac{x-\mu_2}{\sigma_2}\right), v=\frac{y-\mu_2}{\sigma_2}$，得

$$\mathrm{Cov}(X,Y)=\frac{\sigma_1\sigma_2}{2\pi}\int_{-\infty}^{+\infty}v\mathrm{e}^{-\frac{v^2}{2}}\left(\int_{-\infty}^{+\infty}(\rho v+u\sqrt{1-\rho^2})\mathrm{e}^{-\frac{u^2}{2}}\mathrm{d}u\right)\mathrm{d}v$$

$$=\frac{\sigma_1\sigma_2}{2\pi}\int_{-\infty}^{+\infty}v\mathrm{e}^{-\frac{v^2}{2}}(\rho v\sqrt{2\pi}+0)\mathrm{d}v$$

$$=\frac{\sigma_1\sigma_2}{2\pi}\rho\int_{-\infty}^{+\infty}v^2\mathrm{e}^{-\frac{v^2}{2}}\mathrm{d}v=\sigma_1\sigma_2\rho,$$

于是，X 与 Y 的相关系数 $\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=\rho$.

通过本例可以清楚地了解二维正态分布中 5 个参数 $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$ 和 ρ 的含义. 在第 3 章中，我们已经知道如果 (X,Y) 服从二维正态分布，那么 X 与 Y 相互独立的充要条件是 $\rho=0$，再结合例 4.3.2 可得

定理 4.3.1 如果 (X,Y) 服从二维正态分布，那么 X 与 Y 相互独立的充要条件是 X 与 Y 互不相关.

关于二维正态分布我们还有一个重要的结论：

定理 4.3.2 设二维随机变量 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$，令 $Z=aX+bY$，$W=cX+dY$，其中 a,b,c,d 为任意常数，则 (Z,W) 也服从二维正态分布.

该定理的证明远超出本书的范围，我们只能补充二点说明：

(1) 如果 (X,Y) 服从二维正态分布，那么 X 与 Y 的线性组合 $Z=aX+bY$ 肯定是一维正态分布. 因为 (Z,W) 服从二维正态分布，所以其边际分布是一维分布.

(2) 如果 (X,Y) 服从二维正态分布，且 X 与 Y 的两个线性组合 $Z=aX+bY$ 与 $W=cX+dY$ 互不相关，那么 Z 与 W 必然独立. 结合定理 4.3.1 与定理 4.3.2 即可得出.

例 4.3.3（随机变量的标准化） 设随机变量 X 的期望和方差都存在，称

$$X^{*}=\frac{X-EX}{\sqrt{DX}} \tag{4.3.6}$$

为 X 的标准化随机变量.证明：

(1) 对任何一个随机变量有 $EX^{*}=0, DX^{*}=1$；

(2) 设 X 与 Y 的期望和方差都存在，则 $\rho_{X,Y}=\rho_{X^{*},Y^{*}}=EX^{*}Y^{*}$.

证明 (1)是显然的；关于(2)，因为

$$\rho_{X,Y}=\frac{\operatorname{Cov}(X,Y)}{\sqrt{DX}\sqrt{DY}}=\frac{E[(X-EX)(Y-EY)]}{\sqrt{DX}\sqrt{DY}}=EX^{*}Y^{*},$$

又 $EX^{*}=EY^{*}=0, DX^{*}=DY^{*}=1$，得

$$\rho_{X^{*},Y^{*}}=\frac{E[(X^{*}-EX^{*})(Y^{*}-EY^{*})]}{\sqrt{DX^{*}}\sqrt{DY^{*}}}=EX^{*}Y^{*},$$

所以(2)成立.

例 4.3.4 设 X 为连续型随机变量，密度函数为 $f(x)=\frac{3}{2}x^{2}(-1<x<1)$，随机变量 $Y=X^{2}$.

(1) 求 $E(X), D(X)$?

(2) 求 $\operatorname{Cov}(X,Y)$，X 与 Y 是否相关?

(3) X 与 Y 是否相互独立，为什么?

解 (1) 因为

$$E(X)=\int_{-1}^{1}x\cdot\frac{3}{2}x^{2}\,\mathrm{d}x=0,\ E(X^{2})=\int_{-1}^{1}x^{2}\cdot\frac{3}{2}x^{2}\,\mathrm{d}x=3\int_{0}^{1}x^{4}\,\mathrm{d}x=\frac{3}{5},$$

所以 $D(X)=E(X^{2})-[E(X)]^{2}=\frac{3}{5}$.

(2) 因为

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

所以

$$\operatorname{Cov}(X,Y)=E(XY)-E(X)E(Y)=0.$$

故 X 与 Y 不相关.

(3) 用反证法讨论.

假设 X 与 Y 相互独立，那么必有

$$E(X^{2}Y)=E(X^{2})\cdot E(Y).$$

而

$$E(X^{2}Y)=E(X^{4})=\int_{-1}^{1}x^{4}\cdot\frac{3}{2}x^{2}\,\mathrm{d}x=3\int_{0}^{1}x^{6}\,\mathrm{d}x=\frac{3}{7},$$

$$E(X^2)\cdot E(Y)=[E(X^2)]^2=\left(\frac{3}{5}\right)^2=\frac{9}{25}\neq\frac{3}{7}.$$

这就导致了矛盾,所以 X 与 Y 不相互独立.

这个具体的例子说明,当两个随机变量不相关时,它们并不一定相互独立,还可能存在其他的函数关系,例如在本例中 $Y=X^2$.

习 题

1. 某箱装有 100 件产品,其中一,二和三等品分别为 80,10 和 10 件,现在从中随机抽取一件,记

$$X_i=\begin{cases}1,\text{若抽到 } i \text{ 等品}(i=1,2,3),\\0,\text{其他}.\end{cases}$$

试求随机变量 X_1 与 X_2 的相关系数.

2. 设随机变量(X,Y)的联合分布律为

$X \backslash Y$	0	1
0	0.3	0.2
1	0.4	0.1

求 $E(X)$,$E(Y)$,$D(X)$,$D(Y)$,$\mathrm{Cov}(X,Y)$,$\rho_{X,Y}$.

3. 设二维随机变量(X,Y)在以(0,0),(0,1),(1,0)为顶点的三角形区域上服从均匀分布,求协方差 $\mathrm{Cov}(X,Y)$和相关系数 $\rho_{X,Y}$.

4. 设(X,Y)的概率密度为

$$f(x,y)=\begin{cases}\frac{1}{2}\sin(x+y),0\leqslant x\leqslant\frac{\pi}{2},0\leqslant y\leqslant\frac{\pi}{2},\\0,\qquad\qquad\text{其他}.\end{cases}$$

求协方差 $\mathrm{Cov}(X,Y)$和相关系数 $\rho_{X,Y}$.

5. 假设随机变量 X 和 Y 的数学期望都等于 1,方差都等于 2,X 和 Y 的相关系数为 0.25,求随机变量 $U=X+2Y$ 和 $V=X-2Y$ 的相关系数.

6. 已知随机变量 X 和 Y 满足 $DX=2$,$DY=3$,$\mathrm{Cov}(X,Y)=1$,求 $\mathrm{Cov}(2X-3Y+1,X+4Y-3)$.

7. 设随机变量 X 和 Y 的相关系数为 0.5,$E(X)=E(Y)=0$,$E(X^2)=E(Y^2)=2$,求 $E(X+Y)^2$.

8. 设二维随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{\pi}, & x^2+y^2\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

试验证 X 和 Y 是不相关的，但 X 和 Y 不是相互独立的.

9. 随机变量 X 和 Y 的相关系数为 0.9，若 $Z=X-0.4$，求 Y 和 Z 的相关系数.

§4.4　其他数字特征

4.4.1　矩

定义 4.4.1　设 X 为一个随机变量，如果 $E|X|^k<+\infty$，其中 $k=1,2,\cdots$，则称 EX^k 为随机变量 X 的 k 阶**原点矩**，称 $E(X-EX)^k$ 为 X 的 k 阶**中心矩**.

例 4.4.1　设随机变量 $X\sim e(\theta)$，求 X 的 k 阶原点矩.

解　$EX^k=\int_0^{+\infty}x^k\theta\mathrm{e}^{-\theta x}\,\mathrm{d}x=\dfrac{1}{\theta^k}\int_0^{+\infty}(\theta x)^k\mathrm{e}^{-\theta x}\,\mathrm{d}(\theta x)=\dfrac{k!}{\theta^k}$.

例 4.4.2　设随机变量 $X\sim N(0,\sigma^2)$，求 X 的 k 阶原点矩.

解　由于 X 的密度函数为偶函数，所以 X 的奇数阶矩均为零；所以只需求偶数阶矩.

$$EX^{2k}=\int_{-\infty}^{\infty}x^{2k}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\mathrm{d}x=2\int_0^{\infty}x^{2k}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\mathrm{d}x\quad(\text{令 }x=\sqrt{2}\sigma\sqrt{t}).$$

$$=\frac{2^k\sigma^{2k}}{\sqrt{\pi}}\int_0^{\infty}t^{k-\frac{1}{2}}\mathrm{e}^{-t}\mathrm{d}t=\frac{2^k\sigma^{2k}}{\sqrt{\pi}}\Gamma\left(k+\frac{1}{2}\right).$$

利用性质 $\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$ 及 $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$，可得

$$\Gamma\left(k+\frac{1}{2}\right)=\left(k-\frac{1}{2}\right)\left(k-\frac{3}{2}\right)\cdots\frac{1}{2}\Gamma\left(\frac{1}{2}\right)=\frac{(2k-1)!!}{2^k}\sqrt{\pi},$$

所以

$$EX^{2k}=\sigma^{2k}(2k-1)!!.$$

4.4.2　协方差矩阵

定义 4.4.2　设 $X_1,X_2,\cdots,X_n$ 为 n 个随机变量，令 $\sigma_{ij}=\mathrm{Cov}(X_i,X_j)$，其中 $i,j=1,2,\cdots,n$，称矩阵

$$\boldsymbol{\Sigma}=(\sigma_{ij})_{n\times n}$$

为 $X_1,X_2,\cdots,X_n$ 的**协方差矩阵**.

协方差矩阵的性质：

(1) 协方差矩阵为对称矩阵.

这一点由 $\sigma_{ij}=\mathrm{Cov}(X_i,X_j)=\mathrm{Cov}(X_j,X_i)=\sigma_{ji}$ 即可得出.

(2) 协方差矩阵为半正定矩阵. 即,对任意的常数 $a_1,a_2,\cdots,a_n$,都有

$$\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\alpha}\geqslant 0,$$

其中 n 维向量 $\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$.

习 题

1. 设 $X\sim U[0,1]$,求 X 的 k 阶原点矩与 k 阶中心距.

2. 已知二维随机变量(X,Y)的协方差矩阵为 $\begin{bmatrix}1 & 1\\ 1 & 4\end{bmatrix}$,试求 $Z_1=2X-Y$ 和 $Z_2=X-2Y$ 的相关系数.

3. 已知 $X_1,X_2,\cdots,X_n$ 的协方差矩阵 $\boldsymbol{\Sigma}=(\sigma_{ij})_{n\times n}$.

(1) 对任意的常数 $a_1,a_2,\cdots,a_n$,令 $Y=\sum\limits_{i=1}^{n}a_iX_i$,证明 $DY=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\alpha}$,其中 n 维向量 $\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$.

(2) 对任意的常数 $a_1,a_2,\cdots,a_n$ 与 $b_1,b_2,\cdots,b_n$,令 $Y=\sum\limits_{i=1}^{n}a_iX_i$, $Z=\sum\limits_{i=1}^{n}b_iX_i$. 证明 $\mathrm{Cov}(Y,Z)=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\beta}$,其中 n 维向量 $\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$,$\boldsymbol{\beta}=(b_1,b_2,\cdots,b_n)^{\mathrm{T}}$.

§ 4.5 综 合 例 题

例 4.5.1 设$(X,Y)\sim N(1,1,2^2,3^2,0.5)$,试求以下问题.

(1) $\mathrm{Cov}(X,Y)$,$E(XY)$,$\mathrm{Cov}(X,2X-3Y)$,$D(X-2Y)$;

(2) 求 k,使得 $X+kY$ 与 $X-Y$ 相互独立;

(3) 求 $Z=X-3Y$ 的分布;

(4) 求概率 $P(3X-2Y>7)$.

解 (1) $\mathrm{Cov}(X,Y)=\rho\sigma_1\sigma_2=0.5\cdot 2\cdot 3=3$,

$E(XY)=\mathrm{Cov}(X,Y)+EX\cdot EY=3+1\cdot 1=4$,

$\mathrm{Cov}(X,2X-3Y)=2DX-3\mathrm{Cov}(X,Y)=2\cdot 4-3\cdot 3=-1$,

$D(X-2Y)=DX+4DY-4\mathrm{Cov}(X,Y)=4+4\cdot 9-4\cdot 3=28$.

(2) 因为(X,Y)为二维正态分布,所以 $X+kY$ 与 $X-Y$ 相互独立等价于互不相关. 又

$$\mathrm{Cov}(X+kY,X-Y)=DX-kDY+(k-1)\mathrm{Cov}(X,Y)=1-6k,$$

所以当 $k=\frac{1}{6}$时,$X+kY$ 与 $X-Y$ 相互独立.

(3) 由(X,Y)为二维正态分布可知$Z=X-3Y$服从一维正态分布,又

$$EZ=E(X-3Y)=-2,D(X-3Y)=DX+9DY-6\mathrm{Cov}(X,Y)=67,$$

所以$Z\sim N(-2,67)$.

(4) 仿照(3)的做法,可得$3X-2Y\sim N(1,36)$,所以

$$P(3X-2Y>7)=P\left(\frac{3X-2Y-1}{6}>\frac{7-1}{6}\right)=1-\Phi(1)$$
$$=1-0.841\,3=0.158\,7.$$

例 4.5.2　从区间$[0,1]$上任取n个点,求最大点与最小点之间距离的数学期望.

解　设$X_1,X_2,\cdots,X_n$为这n个点,则$X_1,X_2,\cdots,X_n$相互独立,且都服从区间$[0,1]$上的均匀分布.令

$$Y=\max\{X_1,X_2,\cdots,X_n\},\quad Z=\min\{X_1,X_2,\cdots,X_n\},$$

则Y与Z分别为最大点与最小点,距离就是$Y-Z$.由例3.4.8知Y与Z的密度函数分别为

$$f_{\max}(x)=\begin{cases}nx^{n-1}, & 0\leqslant x<1,\\ 0, & \text{其他}\end{cases}\text{与 } f_{\min}(x)=\begin{cases}n(1-x)^{n-1}, & 0\leqslant x<1,\\ 0, & \text{其他},\end{cases}$$

所以$EY=\int_0^1 nx^n\mathrm{d}x=\frac{n}{n+1}$,　$EZ=\int_0^1 nx(1-x)^{n-1}\mathrm{d}x=\frac{1}{n+1}$,于是

$$E(Y-Z)=EY-EZ=\frac{n-1}{n+1}.$$

例 4.5.3　设$(X,Y)\sim N(1,1,2^2,3^2,0.5)$,求$E\max\{X,Y\}$.

解　由于

$$\max\{X,Y\}=\frac{1}{2}(X+Y+|X-Y|),$$

且X与Y的期望是已知的,所以关键是求$|X-Y|$的期望,为此令$Z=X-Y$.因为(X,Y)为二维正态分布,所以$Z=X-Y$为一维正态分布.又

$$EZ=E(X-Y)=0,\quad DZ=D(X-Y)=DX+DY-2\mathrm{Cov}(X,Y)=7,$$

故$Z\sim N(0,7)$,有

$$E|X-Y|=E|Z|=\int_{-\infty}^{\infty}|z|\frac{1}{\sqrt{14\pi}}\mathrm{e}^{-\frac{z^2}{14}}\mathrm{d}z$$
$$=\frac{14}{\sqrt{14\pi}}\int_0^{\infty}\mathrm{e}^{-\frac{z^2}{14}}\mathrm{d}\frac{z^2}{14}=\frac{14}{\sqrt{14\pi}}.$$

所以

$$E\max\{X,Y\}=\frac{1}{2}(EX+EY+E|X-Y|)$$

$$= \frac{1}{2}\left(1+1+\sqrt{\frac{14}{\pi}}\right)$$
$$= 1+\sqrt{\frac{7}{2\pi}}.$$

例 4.5.4 将 n 个球随机地放入 N 个盒子中，假设每个球落入各个盒子是等可能的，求有球盒子数的数学期望.

解 令

$$X_i = \begin{cases} 1, & \text{第 } i \text{ 个盒子有球}, \\ 0, & \text{第 } i \text{ 个盒子没球}, \end{cases} i = 1,2,\cdots,n,$$

则 $X = \sum\limits_{i=1}^{n} X_i$ 就表示有球盒子数. 由于 X_i 为 0—1 分布，所以

$$EX_i = P\{X_i = 1\} = 1 - P\{X_i = 0\} = 1 - \frac{(N-1)^n}{N^n},$$

所以

$$EX = E\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} EX_i = n\left(1-\left(1-\frac{1}{N}\right)^n\right).$$

如果想通过求 $X = \sum\limits_{i=1}^{n} X_i$ 的分布列来求期望的话，那将是非常麻烦的一件事情，学有余力的同学可以考虑一下.

例 4.5.5 将 n 个编号为 $1,2,\cdots,n$ 的球随机地放入编号为 $1,2,\cdots,n$ 的 n 个盒子中，一个盒子只能装一个球，如果第 i 号球正好放入了第 i 号盒子，称为一个配对. 求配对数的期望与方差.

解 令

$$X_i = \begin{cases} 1, & \text{第 } i \text{ 个盒子配对}, \\ 0, & \text{第 } i \text{ 个盒子没配对}, \end{cases} i = 1,2,\cdots,n,$$

则 $X = \sum\limits_{i=1}^{n} X_i$ 就表示配对数. 注意 X_i 为 0—1 分布，且 $EX_i = \frac{1}{n}$，$DX_i = \frac{n-1}{n^2}$.

所以

$$EX = E\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} EX_i = n\,\frac{1}{n} = 1.$$

此处 $X_1, X_2, \cdots, X_n$ 是不独立的，所以有

$$DX = D\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} DX_i + \sum_{i\neq j} \mathrm{Cov}(X_i, X_j),$$

求 DX 的关键就是求 $\mathrm{Cov}(X_i, X_j) = EX_iX_j - EX_iEX_j$，注意到 X_iX_j 也是 0—1 分布，且 $X_iX_j = 1$ 表示第 i 号盒子与第 j 号盒子都配对，所以 $P\{X_iX_j = 1\} =$

$\frac{1}{n(n-1)}$，有

$$\mathrm{Cov}(X_i,X_j)=EX_iX_j-EX_iEX_j=P\{X_iX_j=1\}-\frac{1}{n^2}=\frac{1}{n^2(n-1)}.$$

于是

$$\begin{aligned}DX=D\sum_{i=1}^{n}X_i&=\sum_{i=1}^{n}DX_i+\sum_{i\neq j}\mathrm{Cov}(X_i,X_j)\\&=n\frac{n-1}{n^2}+2\mathrm{C}_n^2\frac{1}{n^2(n-1)}=1.\end{aligned}$$

复 习 题 4

1. 假设随机变量 X 和 Y 相互独立，并且都在区间(0,1)上服从均匀分布，试求随机变量 $Z=|X-Y|$ 的数学期望.

2. 设两个随机变量 X 和 Y 相互独立，且都服从均值为 0，方差为 1/2 的正态分布，求随机变量 $|X-Y|$ 的方差.

3. 将一枚硬币重复掷 n 次，以 X 和 Y 分别表示正面向上和反面向上的次数. 试求 X 和 Y 的相关系数 ρ_{XY}.

4. 已知随机变量 X 和 Y 分别服从正态分布 $N(1,3^2)$ 和 $N(0,4^2)$，且 X 和 Y 的相关系数 $\rho=0.5$，设 $Z=\frac{X}{3}+\frac{Y}{2}$.

(1) 求 Z 的数学期望 EZ 和方差 DZ；

(2) 求 X 与 Z 的相关系数 ρ_{XZ}；

(3) 问 X 与 Z 是否相互独立，为什么？

5. 对于两事件 A 和 B 满足 $0<P(A)<1$，$0<P(B)<1$，定义

$$X=\begin{cases}1, & \text{若 } A \text{ 发生},\\ 0, & \text{否则},\end{cases}\quad Y=\begin{cases}1, & \text{若 } B \text{ 发生},\\ 0, & \text{否则}.\end{cases}$$

证明 X 和 Y 相互独立的充要条件是 X 和 Y 互不相关.

6. 设 X 为只取非负整数值的离散随机变量，证明 $EX=\sum_{n=1}^{+\infty}P(X\geqslant n)$.

7. 对于两个随机变量 X 和 Y，若 EX^2，EY^2 都存在，证明：

$$(E(XY))^2\leqslant EX^2EY^2.$$

(这一不等式称为柯西许瓦兹不等式.)

8. 设随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，且 $EX_i=\mu$，$DX_i=\sigma^2$，$i=1,2,\cdots,n$，令 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 及 $Y_i=X_i-\bar{X}$，$i=1,2,\cdots,n$.

(1) 求 $E\bar{X}$，$D\bar{X}$；

(2) 证明 $S^2=\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\bar{X}^2\right)$，并求 ES^2；

(3) 求 $\mathrm{Cov}(X_i,\bar{X}),i=1,2,\cdots,n$；

(4) 求 $DY_i,i=1,2,\cdots,n$；

(5) 求 $\mathrm{Cov}(Y_i,Y_j),i\neq j$；

(6) 求 $D(Y_i-Y_j),i\neq j$.

第 5 章　大数定律与中心极限定理

在第 1 章概率的统计定义中，我们看到一个事件发生的频率具有稳定性，即一个随机事件发生的频率，在某一定值附近摆动，也就是在某种收敛意义下频率逼近该随机变量的概率. 同时，人们在实践中也发现，大量测量值的算术平均也具有稳定性. 例如某学校有上万名学生，如果随意观察一名学生的身高 X，则 X 与全校学生的平均身高 h 可能相差较大；若随机观察 10 名学生的身高，其平均值与 h 接近的机会就较大；若随机观察 100 名学生的身高，则其平均值就与 h 更为接近. 本节介绍的大数定理将从理论上概括和论证此类现象，即随机变量频率的稳定性，以及 n 个随机变量的平均值的稳定性.

§5.1　大 数 定 律

5.1.1　切比雪夫不等式

首先介绍一个重要的不等式——切比雪夫不等式.

定理 5.1.1　设随机变量 X 具有期望 $E(X)=\mu$，方差 $D(X)=\sigma^2$，则对于任意的正数 ε，有

$$P\{|X-\mu|\geqslant\varepsilon\}\leqslant\frac{\sigma^2}{\varepsilon^2}. \tag{5.1.1}$$

下面仅就连续型随机变量的情况加以讨论.

证明　设 X 为连续型随机变量，密度函数为 $f(x)$，则

$$\begin{aligned}P\{|X-\mu|\geqslant\varepsilon\}&=\int_{|x-\mu|\geqslant\varepsilon}f(x)\mathrm{d}x\leqslant\int_{|x-\mu|\geqslant\varepsilon}\frac{|x-\mu|^2}{\varepsilon^2}f(x)\mathrm{d}x\\&\leqslant\frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}(x-u)^2f(x)\mathrm{d}x=\frac{\sigma^2}{\varepsilon^2}\end{aligned}$$

切比雪夫不等式也可以写为

$$P\{|X-\mu|<\varepsilon\}\geqslant1-\frac{\sigma^2}{\varepsilon^2}.$$

切比雪夫不等式说明，X 的方差越小，则事件$\{|X-\mu|<\varepsilon\}$发生的概率就越大，即 X 取的值越集中于它的期望 μ 附近.

切比雪夫不等式常用来求在随机变量分布未知，只知其期望和方差的情况

下，对事件$\{|X-E(X)\geqslant\varepsilon|\}$概率的下限估计. 这个结论有着很实际的意义：人们在进行精密测量时，为了减少随机误差，往往重复测量多次，得到若干实测值$X_1,X_2,\cdots,X_n$，然后用其平均值$\frac{1}{n}\sum_{i=1}^{n}X_i$来代替其真值$\mu$.

例 5.1.1 设电站供电网有 10 000 盏电灯，夜晚每一盏灯开灯的概率都是 0.7，而假定开、关时间彼此独立，估计夜晚同时开着的灯数在 6 800 盏与 7 200 盏之间的概率.

解 设 X 表示在夜晚同时开着的灯的数目，它服从参数为 $n=10\ 000,p=0.7$ 的二项分布，若要准确计算，应该用伯努利公式：

$$P\{6\ 800<X<7\ 200\}=\sum_{k=6\ 801}^{7\ 199}C_{10\ 000}^{k}\times 0.7^{k}\times 0.3^{10\ 000-k}.$$

如果用切比雪夫不等式估计：

$$E(X)=np=10\ 000\times 0.7=7\ 000,$$

$$D(X)=npq=10\ 000\times 0.7\times 0.3=2\ 100,$$

$$P\{6\ 800<X<7\ 200\}=P\{|X-7\ 000|<200\}\geqslant 1-\frac{2\ 100}{200^2}\approx 0.95.$$

例 5.1.2 设随机变量 $X\sim P(2),Y\sim e\left(\frac{1}{2}\right)$，且 $\rho_{X,Y}=\frac{1}{3}$，试估计概率 $P(|X-Y|>5)$.

解 令 $Z=X-Y$，则

$$EZ=EX-EY=2-2=0,$$

$$\begin{aligned}DZ&=D(X-Y)=DX+DY-2\mathrm{Cov}(X,Y)\\&=DX+DY-2\sqrt{DXDY}\rho_{X,Y}\\&=2+4-2\sqrt{8}\,\frac{1}{3}=6-\frac{4}{3}\sqrt{2}.\end{aligned}$$

所以由切比雪夫不等式可得

$$P\{|X-Y|>5\}\leqslant\frac{D(X-Y)}{25}=\frac{18-4\sqrt{2}}{75}.$$

5.1.2 大数定律

我们已经知道，一个事件 A 发生的频率具有稳定性. 当实验次数 n 增大时，其频率 $f_n(A)$接近于某个常数，即事件 A 的概率 $P(A)$. 这似乎告诉我们，$P(A)$可以看作 $f_n(A)$在某种意义下的“极限”. 为此，引入依概率收敛的定义.

定义 设 $X_1,X_2,\cdots,X_n,\cdots$是一个随机变量序列，$a$ 是一个常数，若对于任意正数 ε，有

$$\lim_{n\to+\infty} P\{|X_n - a| \geqslant \varepsilon\} = 0, \tag{5.1.2}$$

或

$$\lim_{n\to+\infty} P\{|X_n - a| < \varepsilon\} = 1,$$

则称序列 $X_1, X_2, \cdots, X_n, \cdots$ **依概率收敛**于 a，记为 $X_n \xrightarrow{P} a$.

$X_n \xrightarrow{P} a$ 的直观解释是：对任意的 $\varepsilon>0$，当 n 充分大时，“X_n 与 a 的偏差大于 ε”这一事件 $\{|X_n-a|>\varepsilon\}$ 发生的概率很小（收敛于 0），或 X_n 发生的概率越来越接近于 a. 但是这里的极限是在概率意义下的极限，和高等数学中的极限有很大不同，不能当做 $\lim\limits_{n\to+\infty} P(X_n)=a$.

定理 5.1.2（切比雪夫大数定律） 设 $X_1, X_2, \cdots, X_n, \cdots$ 是相互独立的随机变量序列，每一随机变量都存在有限的方差，且一致有界，即存在常数 C，使 $D(X_i) \leqslant C, i=1,2,\cdots$，则对任意的 $\varepsilon>0$，有

$$\lim_{n\to+\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i)\right| \geqslant \varepsilon\right\} = 0 \tag{5.1.3}$$

证明 由切比雪夫不等式知：$\forall \varepsilon>0$，有

$$0 \leqslant P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i)\right| \geqslant \varepsilon\right\} \leqslant \frac{1}{\varepsilon^2} D\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

$$= \frac{\sum\limits_{i=1}^{n} D(X_i)}{n^2\varepsilon^2} \leqslant \frac{nC}{n^2\varepsilon^2} = \frac{C}{n\varepsilon^2} \to 0 \quad (n \to +\infty).$$

该定理表明：当 n 很大时，随机变量 $X_1, \cdots, X_n$ 的算术平均值 $\frac{1}{n}\sum\limits_{i=1}^{n} X_i$ 接近于其数学期望 $E\left(\frac{1}{n}\sum\limits_{i=1}^{n} X_i\right)$，这种接近是在概率意义下的接近. 通俗地说，在定理的条件下，n 个相互独立的随机变量算术平均值，在 n 无限增加时将几乎变成一个常数.

定理 5.1.3（伯努利大数定律） 设 n_A 是 n 重相互独立重复试验（即伯努利试验）中事件 A 出现的次数，而 p $(0<p<1)$ 是事件 A 在每次试验中出现的概率，则对任意的正数 $\varepsilon>0$，有

$$\lim_{n\to+\infty} P\left\{\left|\frac{n_A}{n} - p\right| \geqslant \varepsilon\right\} = 0 \quad 或 \quad \lim_{n\to+\infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1$$

证明 令

$$X_i = \begin{cases} 1, 第\ i\ 次试验中\ A\ 出现, \\ 0, 第\ i\ 次试验中\ A\ 不出现, \end{cases} \quad i = 1,2,\cdots,n,$$

则 $n_A=\sum\limits_{i=1}^{n}X_i$，然后应用切比雪夫大数定律即可.

伯努利大数定律表明：随着 n 的增大，事件 A 发生的频率与其概率 p 的偏差 $\left|\frac{n_A}{n}-p\right|$ 大于预先给定的精度 ε 的可能性越来越小，小到可以忽略不计，这就是频率稳定于概率的含义，或者说频率依概率收敛于概率. 在实际应用中，当试验次数很大时，便可以用事件发生的频率来代替事件的概率. 该定理的证明也说明了伯努利大数定律只是切比雪夫大数定律的一个特例.

本节最后再介绍辛钦大数定律.

定理 5.1.4(辛钦大数定律) 设随机变量 $X_1,\cdots,X_n,\cdots$ 为一系列独立同分布的随机变量，且具有数学期望 $E(X_i)=\mu(i=1,2,\cdots)$，则对任意正数 $\varepsilon>0$，有

$$\lim_{n\to+\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right|>\varepsilon\right\}=0\ ,\left(\text{即}\frac{1}{n}\sum_{i=1}^{n}X_i\text{ 以概率收敛于 }\mu\right).$$

定理证明超出本书范围，略.

伯努利大数定律是辛钦大数定律的特殊情况. 辛钦大数定律与切比雪夫大数定律的一个区别在于定理的条件不同，辛钦大数定律对方差没要求，但它要求 $X_1,X_2,\cdots,X_n,\cdots$ 是同分布的；切比雪夫大数定律对方差有要求，但它不要求同分布.

例 5.1.3 设 $X_1,X_2,\cdots,X_n,\cdots$ 相互独立，且均服从参数为 λ 的泊松分布，问当 $n\to+\infty$ 时，$\left\{\frac{1}{2n}\sum\limits_{i=1}^{n}X_i^2\right\}_{n=1}^{+\infty}$ 依概率收敛到什么？

解 令 $Y_i=\frac{1}{2}X_i^2(i=1,2,\cdots)$，则 $Y_1,Y_2,\cdots$ 相互独立且具有同分布. 由辛钦大数定律知，$\frac{1}{n}\sum\limits_{i=1}^{n}Y_i$ 依概率收敛到 $E(Y_i)=\frac{1}{2}(\lambda+\lambda^2)$，即 $\left\{\frac{1}{2n}\sum\limits_{i=1}^{n}X_i^2\right\}_{n=1}^{+\infty}$ 依概率收敛到 $\frac{1}{2}(\lambda+\lambda^2)$.

习 题

1. 利用切比雪夫不等式估计随机变量与其数学期望之差大于 3 倍标准差的概率.

2. 设随机变量 X 的方差为 2，试根据切比雪夫不等式估计 $P\{|X-E(X)|\geqslant 2\}$ 之值.

3. 在每次试验中，事件 A 发生的概率为 0.5，利用切比雪夫不等式估计：在1 000次试验中，事件 A 发生次数在 400～600 之间的概率.

4. 设随机变量 X 和 Y 的数学期望都是 2，方差分别为 1 和 4，而相关系数为 0.5. 试根据

切比雪夫不等式给出 $P\{|X-Y|\geqslant 6\}$ 的估计.

§5.2　中心极限定理

正态分布在随机变量的各种分布中，占有非常重要的地位.在客观实际中，有许多被研究的随机变量，可以表示为大量相互独立的随机变量的和，而其中个别随机变量在总的影响中，所起的作用都是微小的.这些个别随机变量的个数无限增加时，它们的和趋于正态分布，这种现象就是中心极限定理的客观背景.

一般地，我们把在某种条件下，使得随机变量序列的极限分布是正态分布的结果，统称为中心极限定理.

下面介绍两个常用的中心极限定理.

定理5.2.1(独立同分布的中心极限定理)　设 $X_1,\cdots,X_n,\cdots$ 是相互独立同分布的随机变量序列，且 $E(X_i)=\mu$，$D(X_i)=\sigma^2$，$i=1,2,\cdots$，则对 $\forall x\in\mathbf{R}$，有

$$\lim_{n\to+\infty}P\left\{\frac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}\leqslant x\right\}=\Phi(x).\tag{5.2.1}$$

证明略.

注意到 $\dfrac{\sum_{i=1}^{n}X_i-n\mu}{\sigma\sqrt{n}}=\left(\sum_{i=1}^{n}X_i\right)^*$，于是有下式成立

$$\lim_{n\to\infty}P\left(\left(\sum_{i=1}^{n}X_i\right)^*\leqslant x\right)=\Phi(x),\tag{5.2.2}$$

其中 $\Phi(x)$ 为标准正态分布的分布函数.

在一般情况下，很难求出 n 个随机变量之和 $\sum_{i=1}^{n}X_i$ 的分布函数.该定理表明，当 n 充分大时，可以通过 $\Phi(x)$ 给出其近似分布.这样就可以利用正态分布对 $\sum_{i=1}^{n}X_i$ 作理论分析或作实际计算，其好处是明显的.该定理也可改写为：对任意的 $a<b$，有 $\lim\limits_{n\to\infty}P\left\{a<\dfrac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}\leqslant b\right\}=\Phi(b)-\Phi(a)$.

上述结果表明，无论 X_i 服从何种分布，只要 $D(X_i)<\infty$，那么当 n 较大时，随机变量 $\dfrac{1}{\sigma\sqrt{n}}\left(\sum_{i=1}^{n}X_i-n\mu\right)$ 就近似服从标准正态分布 $N(0,1)$.这一事实确立

了标准正态分布 $N(0,1)$在概率论中的中心地位."极限"是指 n 越大越好,故称为中心极限定理.中心极限定理的核心之处在于随机变量的标准化:当 n 很大时,$\sum_{i=1}^{n} X_i$ 的标准化随机变量就可以认为是标准正态分布.

例 5.2.1 某餐厅每天接待 400 名顾客,设每位顾客的消费额(元)服从(20,100)上的均匀分布,且顾客的消费额是相互独立的,求:

(1) 该餐厅每天的平均营业额;

(2) 该餐厅每天的营业额在平均营业额±760 元内的概率.

解 记 X_i 为第 i 位顾客的消费额,则 X_i 服从(20,100)上的均匀分布,且 $E(X_i)=\frac{100+20}{2}=60, D(X_i)=\frac{(100-20)^2}{12}=\frac{1\ 600}{3}$,则该餐厅每天的营业额为 $Y=\sum_{i=1}^{400} X_i$.

(1) 给餐厅每天的平均营业额为 $E(Y)=\sum_{i=1}^{400} E(X_i)=24\ 000$ (元).

(2) 利用独立同分布的中心极限定理知,

$$P\{-760<Y-24\ 000<760\}\approx 2\Phi\left(\frac{760}{\sqrt{400\times\frac{1\ 600}{3}}}\right)-1$$
$$=2\Phi(1.645)-1=0.90.$$

例 5.2.2 设有 2 500 人参见了某保险公司的人寿保险,在一年中每个人死亡的概率为 0.002,每个人交保费 120 元,若在一年内死亡,保险公司赔付20 000 元,问:

(1) 保险公司亏本的概率是多少?

(2) 保险公司至少获利 100 000 元的概率是多少?

解 以 X 表示这 2 500 人中的死亡人数.令

$$X_i=\begin{cases}1,\text{第 } i \text{ 个人死亡},\\ 0,\text{否则},\end{cases}\quad i=1,2,\cdots,2\ 500,$$

则 $X_i \sim B(1,0.002)$,且 $X=\sum_{i=1}^{2\ 500} X_i \sim B(2\ 500,0.002)$.

保险公司共收人 120×2 500=300 000 元,所以当死亡人数 $X=\sum_{i=1}^{2\ 500} X_i$ 大于等于 15 人时,保险公司亏本,于是

(1) $P\{X\geqslant 15\}=P\left\{\frac{X-EX}{\sqrt{DX}}\geqslant\frac{15-EX}{\sqrt{DX}}\right\}$

$$=P\left\{\frac{X-5}{\sqrt{4.99}}>\frac{10}{\sqrt{4.99}}\right\}\approx 1-\Phi\left(\frac{10}{\sqrt{4.99}}\right)$$
$$=1-0.999\,996=0.000\,004.$$

(2) 当死亡人数 $X=\sum_{i=1}^{2\,500}X_i$ 小于等于 10 人时，保险公司至少获利 100 000 元，于是

$$P\{X\leqslant 10\}=P\left\{\frac{X-EX}{\sqrt{DX}}\leqslant\frac{10-EX}{\sqrt{DX}}\right\}$$
$$=P\left\{\frac{X-5}{\sqrt{4.99}}\leqslant\frac{5}{\sqrt{4.99}}\right\}\approx\Phi\left(\frac{5}{\sqrt{4.99}}\right)=0.987\,1.$$

下面介绍另一个中心极限定理，它是独立同分布中心极限定理的特例.

定理 5.2.2(棣莫弗-拉普拉斯中心极限定理)　设随机变量 X 表示 n 重伯努利试验中事件 A 发生的次数，$p(0<p<1)$是事件 A 在每次试验中出现的概率，则对任意实数 x，有

$$\lim_{n\to+\infty}P\left\{\frac{X-np}{\sqrt{np(1-p)}}\leqslant x\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t=\Phi(x).$$

证明留作练习.

这个定理表明，正态分布是二项分布的极限分布，即当 n 充分大时，$\frac{X-np}{\sqrt{np(1-p)}}$近似地服从标准正态分布 $N(0,1)$. 可用下面的方法计算二项分布的概率：

$$P\{a<X\leqslant b\}\approx\Phi\left(\frac{b-np}{\sqrt{np(1-p)}}\right)-\Phi\left(\frac{a-np}{\sqrt{np(1-p)}}\right)$$

例 5.2.3　某公司的 200 名员工参加一种资格证书考试，按往年经验，该考试通过率为 0.8，试计算这 200 名员工至少有 150 人考试通过的概率.

解　令

$$X_i=\begin{cases}1,\text{第 } i \text{ 人通过考试},\\0,\text{第 } i \text{ 人未通过考试},\end{cases}\quad i=1,2,\cdots,200,$$

由题设知，$P\{X_i=1\}=0.8$，$np=200\times 0.8=160$，$np(1-p)=32$，$\sum_{i=1}^{200}X_i$ 是考试通过人数.

依定理 5.2.2 有，$\dfrac{\sum_{i=1}^{200}X_i-160}{\sqrt{32}}\sim N(0,1)$，

$$P\left\{\sum_{i=1}^{200}X_i \geqslant 150\right\} = P\left\{\frac{\sum_{i=1}^{200}X_i - 160}{\sqrt{32}} \geqslant \frac{150-160}{\sqrt{32}}\right\}$$
$$= P\left\{\frac{\sum_{i=1}^{200}X_i - 160}{\sqrt{32}} \geqslant -1.77\right\}$$
$$\approx 1-\Phi(-1.77) = \Phi(1.77) = 0.96,$$

即至少有 150 名员工通过这种资格证书考试的概率为 0.96.

本节介绍的中心极限定理证明了，在很一般条件下，n 个随机变量的和当 $n\to\infty$ 时的极限分布是正态分布. 这样，在数理统计中许多复杂的随机变量分布，可以用正态分布去近似，借助于正态分布的完美结论，可以获得既实用又简单的统计分析.

习　题

1. 重复投掷硬币 100 次，设每次出现正面的概率均为 0.5，问正面出现次数小于 61 大于 50 的概率是多少？

2. 从区间[0,1]任取 100 个数，求这 100 个数的和介于 47.5～52.5 之间的概率.

3. 设供电站供应某地区 1 000 户居民用电，各户用电情况相互独立，已知每户每日用电量(单位：度)在[0,20]上服从均匀分布，求：

(1) 这 1 000 户居民每日用电量超过 10 100 度的概率；

(2) 要以 0.99 的概率保证该地区居民用电量的需要，问供电站每天至少要向该地区供应多少度电？

4. 某厂生产的螺丝钉的不合格品率为 0.01，问一盒子中应至少装多少只螺丝钉才能保证其中含有的 100 只合格品的概率不小于 0.95？

5. 一生产线上加工成箱零件，每箱平均重量为 50 kg，标准差为 5 kg，假设承运这批产品的汽车的最大载重量为 5 t，试利用中心极限定理说明该车最多可以装多少箱，才能以概率 99.7%保障不超载？

6. 设某单位有 260 部内线电话，每部电话有 4%的时间要使用外线，且不同的内线电话是否使用外线相互独立，问该单位至少需要多少条外线，才能保证每部电话在使用外线时，可以打通的概率是 95%？

7. 证明极限：$\lim\limits_{n\to+\infty}\sum\limits_{k=0}^{n}\frac{n^k}{k!}e^{-n} = \frac{1}{2}$.

第 6 章　数理统计的基本概念

从本章起，我们开始讲述数理统计．数理统计是以概率论为理论基础，应用非常广泛的一个数学分支．

在概率论中，我们研究的随机变量，其分布大都是已知的．在这一前提下，随机现象的统计规律性完全可以描述．但是对于太多的实际问题，一个随机变量的概率分布往往是不知道的，或者虽知道其分布的类型，但不知道其中的参数．比如，某工厂生产成批的电视机显像管，显像管的寿命服从什么分布？这是不知道的．如果凭以往的经验，假设显像管的寿命服从指数分布 $e(\lambda)$，但是其中的参数 λ 却是未知的．怎么才能估计出一个随机变量分布中的参数呢？这类问题属于参数估计问题，这是数理统计所研究的最重要的问题之一．

假设电视机显像管批量生产的质量标准是平均使用寿命为 1 200 h，标准差为 300 h．某电视机厂宣称其生产的显像管质量大大超过规定的标准．为了进行验证，随机抽取了 100 件为样本，测得平均使用寿命 1 245 h．能否说明该厂的显像管质量显著地高于规定的标准．这类问题与上面问题不同，它需要在两种假设：接受或拒绝厂家的说法中选一个．这类问题属于假设检验问题，这也是数理统计所研究的最重要的问题之一．

数理统计的内容相当丰富，本书只介绍抽样分布、参数估计和假设检验．

在本章中，首先介绍数理统计最基本的三个概念：总体、样本、统计量，随后介绍使用最多的四个随机变量分布：标准正态分布、χ^2 分布、t 分布、F 分布，并介绍正态总体下的六个抽样分布，最后研究随机变量的上 α 分位点性质．

§6.1　总体、样本、统计量

在一个统计问题中，研究对象的全体称为**总体**，其中每个成员称为**个体**．如在研究一批电视机显像管的质量时，该批电视机显像管就是一个总体，其中的每个显像管就是个体．在统计研究中，人们关心的是个体的某个或某些数量指标的分布情况，这时所有个体的数量指标的全体就是总体．由于个体的出现是随机的，因此相应的数量指标的出现也是随机的，故这种数量指标是一个随机变量（或随机向量）．而这个随机变量的分布，就是该指标在总体中的分布．总体可用一个随机变量 X 及其分布 F 来描述．如在研究电视机显像管的质量时，人们关心的数量指标是寿命 X，那么就用 X 或 X 的分布 F 表示总体．

为了推断总体分布及其各种特征，就必须从该总体中按一定法则，抽取若干个体进行观测或试验，以获得有关总体的信息. 这一过程称为“抽样”，所抽取的部分个体称为**样本**，抽取的个体的个数称为**样本容量**. 常用的抽样方法分为无放回抽样和有放回抽样两种. 当样本容量 n 与个体总数 N 相比很小时，两种抽样方法可近似地看作是等价的. 在抽取样本之前，由于获得的个体是随机得到的，其相应的数量指标也是随机的，通常用 n 个随机变量 $X_1, X_2, \cdots, X_n$ 来表示，称其为样本. 一旦个体被抽出且观察或试验结束，就得到 n 个试验数据 $x_1, x_2, \cdots, x_n$. 这 n 个数据称为**样本观测值**（或样本值）. 我们约定用大写英文字母表示样本，用小写英文字母表示样本观测值. 例如，为了研究某批电视机显像管的寿命，决定从中抽取 10 个进行试验，这样就获得了一个容量为 10 的样本 $X_1, X_2, \cdots, X_{10}$，对这 10 个显像管进行寿命检测，就得到样本观测值 $x_1, x_2, \cdots, x_{10}$.

最常用的抽样方法为“**简单随机抽样**”，它应满足：

(1) **代表性**. 总体中每个个体都有同等机会被抽入样本，即可以认为样本 $X_1, X_2, \cdots, X_n$ 中的每个 $X_i(i=1,2,\cdots,n)$ 都与总体 X 有相同的分布.

(2) **独立性**. 样本中每个个体的取值并不影响其他个体的取值，这意味着 $X_1, X_2, \cdots, X_n$ 相互独立.

由简单随机抽样所得的样本 $X_1, X_2, \cdots, X_n$，称为**简单随机样本**，在不引起混淆的情况下，也可简称为样本. 今后，除非特别说明，样本即指简单随机样本.

设总体 X 的分布函数为 $F(x)$，概率密度函数为 $f(x)$，则样本 $(X_1, X_2, \cdots, X_n)$ 的联合分布函数为

$$F(x_1, x_2, \cdots, x_n) = F(x_1)F(x_2)\cdots F(x_n) = \prod_{i=1}^{n} F(x_i),$$

样本 $(X_1, X_2, \cdots, X_n)$ 的联合概率密度函数为

$$f(x_1, x_2, \cdots, x_n) = f(x_1)f(x_2)\cdots f(x_n) = \prod_{i=1}^{n} f(x_i).$$

样本是总体的反映，但是样本观测值往往不能直接用于提取总体的信息，通常要通过加工、整理，针对不同问题，构造样本的适当函数，并利用这些样本函数进行统计推断，进而获得所关心的信息.

定义 6.1.1 设 $X_1, X_2, \cdots, X_n$ 是总体 X 的一个样本，若 $T=T(X_1, X_2, \cdots, X_n)$ 是样本的函数，且不含任何未知参数，则称 T 为一个**统计量**.

按定义 6.1.1，统计量 $T=T(X_1, X_2, \cdots, X_n)$ 是一个随机变量. 从理论上讲，可以通过样本的分布来确定 T 的分布，称其为抽样分布. 由于样本的分布与未知的总体分布 $F(\cdot)$ 有关，因此抽样分布也与 $F(\cdot)$ 有关. 若 $x_1, x_2, \cdots, x_n$ 是样本观测值，则称 $T(x_1, x_2, \cdots, x_n)$ 是 T 的观测值. 以下介绍一些常见的统计量.

定义 6.1.2 设 $X_1,X_2,\cdots,X_n$ 是从总体 X 中抽取的样本,称统计量

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

为**样本均值**;

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

为**样本方差**;

$$S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2}$$

为**样本标准差**;

$$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k\quad(k=1,2,\cdots)$$

为**样本 k 阶原点矩**;

$$X_{(n)}=\max\{X_1,X_2,\cdots,X_n\}$$

为**极大次序统计量**;

$$X_{(1)}=\min\{X_1,X_2,\cdots,X_n\}$$

为**极小次序统计量**.

一般来讲,用样本均值 $\bar{X}$ 的观测值 $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$,来近似估计总体均值 $\mu=E(X)$;用样本方差 S^2 的观测值 $s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$,来近似估计总体方差 $\sigma^2=D(X)$;用样本 k 阶原点矩 A_k 的观测值,来近似估计总体 k 阶矩 $E(X^k)$.

§6.2 常用统计量的分布

统计量是样本的函数,也是一个随机变量,统计量的分布称为**抽样分布**. 如果能够求出统计量的概率分布,就会对问题的研究非常有用. 但是一般情况下,求出统计量的分布是一件非常困难的事情. 但如果总体是正态分布,问题会变得相对比较简单. 下面简单介绍几个非常重要且常用统计量的分布.

1. 标准正态分布

关于标准正态分布 $N(0,1)$,在概率论中已经进行了详尽的讨论,在此不再多言.

2. χ^2 分布

定义 6.2.1 设 $X_1,X_2,\cdots,X_n$ 相互独立,且均服从标准正态分布 $N(0,1)$,

则称

$$\chi^2 = X_1^2 + X_2^2 + \cdots + X_n^2 \tag{6.2.1}$$

服从自由度为 n 的 χ^2 分布，记为 $\chi^2(n)$，其密度函数为

$$f(x) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\dfrac{n}{2}\right)} e^{-\frac{x}{2}} x^{\frac{n}{2}-1}, & x > 0, \\ 0, & x \leqslant 0, \end{cases}$$

其密度函数图形如图 6－1 所示.

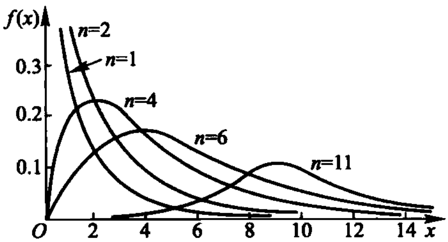


图 6－1

χ^2 分布有如下性质：

(1) (χ^2 分布的可加性)若 $X \sim \chi^2(n)$，$Y \sim \chi^2(m)$，且 X 与 Y 相互独立，则

$$X + Y \sim \chi^2(n+m). \tag{6.2.2}$$

此性质可有定义直接得出.

(2) 若 $X \sim \chi^2(n)$，则有

$$E[\chi^2(n)] = n, D[\chi^2(n)] = 2n. \tag{6.2.3}$$

证明 因为 $X_i \sim N(0,1)$，所以有

$$EX_i = 0,\ EX_i^2 = DX_i + (EX_i)^2 = 1$$

$$\begin{aligned} EX_i^4 &= \int_{-\infty}^{+\infty} x^4 \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} dx = -\int_{-\infty}^{+\infty} x^3 \frac{1}{\sqrt{2\pi}} d\left(\exp\left\{-\frac{x^2}{2}\right\}\right) \\ &= -x^3 \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} \Bigg|_{-\infty}^{+\infty} + 3\int_{-\infty}^{+\infty} x^2 \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} dx \\ &= 3DX_i = 3, \end{aligned}$$

于是

$$DX_i^2 = EX_i^4 - (EX_i^2)^2 = 3 - 1 = 2.$$

因为 $X_1, X_2, \cdots, X_n$ 相互独立，所以

$$E\chi^2 = EX_1^2 + EX_2^2 + \cdots + EX_n^2 = n,$$

$$D\chi^2 = DX_1^2 + DX_2^2 + \cdots + DX_n^2 = 2n.$$

3. t 分布

定义 6.2.2 设 $X\sim N(0,1)$，$Y\sim\chi^2(n)$，且 X 与 Y 相互独立，则称

$$t=\frac{X}{\sqrt{Y/n}} \tag{6.2.4}$$

服从**自由度**为 n 的 **t 分布**，记为 $t(n)$. 其密度函数为

$$f(x)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}\quad(-\infty<x<+\infty).$$

其密度函数图形如图 6-2 所示.

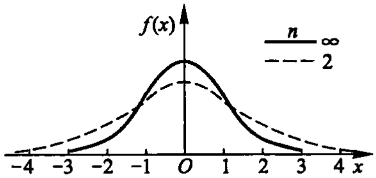


图 6-2

t 分布的密度函数是偶函数，且由于

$$\lim_{n\to+\infty}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}=\mathrm{e}^{-\frac{x^2}{2}},$$

又可以证明下面的极限成立

$$\lim_{n\to+\infty}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}=\frac{1}{\sqrt{2\pi}},$$

因此，有

$$\lim_{n\to+\infty}f(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}$$

这表明当 n 充分大时，自由度为 n 的 t 分布可以近似地看成是标准正态分布，一般地，当 $n\geqslant45$ 时，就可以将 t 分布作为标准正态分布.

4. F 分布

定义 6.2.3 设 $X\sim\chi^2(n)$，$Y\sim\chi^2(m)$，且 X 与 Y 相互独立，则称

$$F=\frac{X/n}{Y/m} \tag{6.2.5}$$

服从自由度为 n,m 的 F 分布，记为 $F(n,m)$. 其密度函数为

$$f(x)=\begin{cases}\dfrac{\Gamma\left(\dfrac{m+n}{2}\right)}{\Gamma\left(\dfrac{m}{2}\right)\Gamma\left(\dfrac{n}{2}\right)}m^{\frac{m}{2}}n^{\frac{n}{2}}x^{\frac{n}{2}-1}(m+nx)^{-\frac{m+n}{2}}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

其密度函数图形如图 6－3 所示.

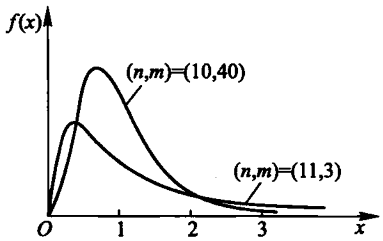


图 6－3

F 分布有如下性质：

(1) 若 $X\sim F(n,m)$，则$\dfrac{1}{X}\sim F(m,n)$.

(2) 若 $t\sim t(n)$，则 $t^2\sim F(1,n)$.

性质(1)可由定义直接得出，性质(2)作为练习.

例 设总体 $X\sim N(0,1)$，$X_1,X_2,\cdots,X_n$ 是简单随机样本，试问下列统计量各服从什么分布？

(1) $X_1^2+X_2^2$，(2) $\dfrac{X_1-X_2}{\sqrt{X_3^2+X_4^2}}$，(3) $\dfrac{X_2}{|X_6|}$，(4) $\dfrac{\left(\dfrac{n}{3}-1\right)\sum\limits_{i=1}^{3}X_i^2}{\sum\limits_{i=4}^{n}X_i^2}$.

解 (1) 因为 $X_i\sim N(0,1)(i=1,2,\cdots,n)$，且相互独立，所以

$$X_1^2+X_2^2\sim\chi^2(2).$$

(2) 因为

$$X_1-X_2\sim N(0,2),$$

所以

$$\frac{X_1-X_2}{\sqrt{2}}\sim N(0,1),$$

$$X_3^2+X_4^2\sim\chi^2(2),$$

并且 X_1-X_2 与 $X_3^2+X_4^2$ 相互独立，所以

$$\frac{\frac{X_1 - X_2}{\sqrt{2}}}{\sqrt{(X_3^2 + X_4^2)/2}} = \frac{X_1 - X_2}{\sqrt{X_3^2 + X_4^2}} \sim t(2).$$

(3) 因为 $X_2 \sim N(0,1)$，$X_6^2 \sim \chi^2(1)$，且相互独立，所以

$$\frac{X_2}{\sqrt{X_6^2/1}} = \frac{X_2}{|X_6|} \sim t(1).$$

(4) 因为 $\sum_{i=1}^{3} X_i^2 \sim \chi^2(3)$，$\sum_{i=4}^{n} X_i^2 \sim \chi^2(n-3)$，且相互独立，所以

$$\frac{\sum_{i=1}^{3} X_i^2/3}{\sum_{i=4}^{n} X_i^2/(n-3)} = \left(\frac{n-3}{3}\right)\frac{\sum_{i=1}^{3} X_i^2}{\sum_{i=4}^{n} X_i^2} \sim F(3,n-3).$$

习　　题

1. 设 $X_1, X_2, \cdots, X_6$ 是来自 $(0,\theta)$ 上的均匀分布的样本，$\theta>0$ 未知. 指出下列样本函数中哪些是统计量，哪些不是？为什么？

(1) $T_1 = \max(X_1, \cdots, X_6)$；(2) $T_2 = X_6 - \theta$；(3) $T_3 = X_6 - E(X_1)$.

2. 证明：样本方差 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right)$.

3. 设样本的一组观测值是：0.5，1，0.7，0.6，1，1，写出样本均值、样本方差和标准差.

4. 设总体 X 服从泊松分布 $P(\lambda)$，$\bar{X}$ 是容量为 n 的样本的均值，求 $E(\bar{X})$，$D(\bar{X})$.

5. 设总体 X 服从均匀分布 $U(-1,1)$，$\bar{X}$ 是容量为 n 的样本的均值，求 $E(\bar{X})$，$D(\bar{X})$.

6. 已知 $X \sim t(n)$，证明：$X^2 \sim F(1,n)$.

7. 利用 F 分布的性质，证明：$F_{1-\alpha}(n_1,n_2) = \frac{1}{F_\alpha(n_2,n_1)}$.

§6.3　正态总体的抽样分布

正态分布 $N(\mu,\sigma^2)$ 在概率论中处于中心地位，同样地，在数理统计中，正态分布的作用仍然是举足轻重的. 一般来讲，如果总体服从正态分布，那么关于该总体的几乎所有统计问题都比较简单. 因此，我们有必要对总体服从正态分布时统计量的分布进行研究.

下面给出几个在数理统计中常用到的有关抽样分布的结论.

定理 6.3.1（单正态总体的抽样分布定理）　设总体 $X \sim N(\mu,\sigma^2)$，X_1，

$X_2,\cdots,X_n$ 为总体 X 的简单随机样本.样本均值 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，样本方差 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$，则有

(1) $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$；

(2) $\frac{n-1}{\sigma^2}S^2\sim\chi^2(n-1)$，且 $\bar{X}$ 与 S^2 相互独立；

(3) $\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$.

证明 只证(1)与(3)，(2)的证明超出本书范围，略.

(1) $X_i(i=1,2,\cdots,n)$相互独立，所以由正态分布的可加性知 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 也为正态分布，且 $\bar{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right)$，利用正态分布的标准化性质即得

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1).$$

(3) 由(2)中知 $\bar{X}$ 与 S^2 相互独立，知$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$与$\frac{n-1}{\sigma^2}S^2\sim\chi^2(n-1)$也相互独立.故由 t 分布的定义得

$$\frac{\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{n-1}{\sigma^2}S^2/(n-1)}}=\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1).$$

下面看一下**双正态总体**的抽样分布.

一般来说，当涉及比较两个量的大小、优劣，或两个指标的好坏时要用到双总体.比如甲、乙两个工厂都生产汽车轮胎，每个工厂都声称自己的轮胎质量比另一个工厂的好，为了比较这两个工厂所生长的轮胎质量，需要分别从两工厂生产的轮胎中抽取一部分轮胎.假设从甲厂中抽取 26 个，从乙厂中抽取 34 个，将这些轮胎随机地装到一些汽车上，记录下每个轮胎的寿命，可得甲厂 26 个轮胎的寿命 $x_1,\cdots,x_{26}$，乙厂 34 个轮胎寿命 $y_1,\cdots,y_{34}$.使用这两组数据就可以判断出甲乙两个工厂哪个工厂生产的轮胎质量好.

双总体在医药行业使用得非常多，因为要经常衡量两种治疗同一种疾病药物的治疗效果.比如：为了比较甲乙两种安眠药的治疗效果，经过抽样调查得平

均延长睡眠时间 $\bar{X}_{甲}=2.5\ \mathrm{h}$，$\bar{X}_{乙}=1.2\ \mathrm{h}$，$s_{甲}=3.2\ \mathrm{h}$，$s_{乙}=0.4\ \mathrm{h}$，此时不能单纯的比较均值，$\bar{X}_{甲}>\bar{X}_{乙}$ 并不能说明甲药比乙药好. 大家知道方差是稳定性的体现，$s_{甲}$ 太大了，也就是说病人本来想睡觉，但吃了甲药之后反倒不想睡觉了；或者吃甲药后一睡不醒. 这种药没人敢吃，反观乙药，就比较稳定，每个人吃了都能有一定的效果.

对于双正态总体的抽样分布，我们有如下结果.

定理 6.3.2(双正态总体的抽样分布定理) 设总体 $X\sim N(\mu_1,\sigma_1^2)$ 与总体 $Y\sim N(\mu_2,\sigma_2^2)$ 相互独立，$X_1,X_2,\cdots,X_n$ 与 $Y_1,Y_2,\cdots,Y_m$ 分别为总体 X 与总体 Y 的简单随机样本. 以 $\bar{X},\bar{Y},S_1^2,S_2^2$ 分别表示两样本的样本均值与样本方差，则有

(1) $\dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n}+\dfrac{\sigma_2^2}{m}}}\sim N(0,1)$；

(2) $\dfrac{S_1^2}{S_2^2}\cdot\dfrac{\sigma_2^2}{\sigma_1^2}\sim F(n-1,m-1)$；

(3) 若 $\sigma_1^2=\sigma_2^2$，则

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_\omega\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}}\sim t(n+m-2),$$

其中

$$S_\omega=\sqrt{\frac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}}.$$

证明 (1) 由于 $\bar{X}\sim N\left(\mu_1,\dfrac{\sigma_1^2}{n}\right)$，$\bar{Y}\sim N\left(\mu_2,\dfrac{\sigma_2^2}{m}\right)$，且二者独立，所以利用正态分布的可加性得

$$\bar{X}-\bar{Y}\sim N\left(\mu_1-\mu_2,\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}\right),$$

将上式标准化即可得(1)式.

(2) 由定理 6.3.1 知 $\dfrac{n-1}{\sigma_1^2}S_1^2\sim\chi^2(n-1)$，$\dfrac{n-1}{\sigma_2^2}S_2^2\sim\chi^2(m-1)$，且二者独立，利用 F 分布的定义可知(2)成立.

(3) 如果 $\sigma_1^2=\sigma_2^2$，那么

$$\frac{n-1}{\sigma_1^2}S_1^2+\frac{m-1}{\sigma_1^2}S_2^2\sim\chi^2(n+m-2),$$

又由(1)知

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sigma_1\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim N(0,1),$$

于是利用 t 分布的定义可得(3).

例 6.3.1 设总体 $X \sim N(52, 6.3^2)$，$X_1, X_2, \cdots, X_{36}$ 为简单随机样本，$\bar{X}$ 为样本均值，求(1) $\bar{X}$ 的数学期望与方差；(2) $P\{50.8<\bar{X}<53.8\}$.

解 (1) 已知 $X \sim N(52, 6.3^2)$，得，

$$\bar{X} \sim N\left(52, \frac{6.3^2}{36}\right),$$

所以

$$E(\bar{X}) = 52, D(\bar{X}) = \frac{6.3^2}{36} = 1.1025.$$

(2) 由 $\bar{X} \sim N\left(52, \frac{6.3^2}{36}\right)$，知

$$\frac{\bar{X}-52}{6.3/6} \sim N(0,1),$$

所以

$$\begin{aligned} P\{50.8 < \bar{X} < 53.8\} &= P\left\{\frac{50.8-52}{6.3/6} < \frac{\bar{X}-52}{6.3/6} < \frac{53.8-52}{6.3/6}\right\} \\ &= \Phi(1.714) - \Phi(-1.143) \\ &= 0.9564 - (1-0.8729) = 0.8293. \end{aligned}$$

例 6.3.2 设 $X_1, X_2, \cdots, X_n, X_{n+1}$ 是正态总体 $X \sim N(\mu, \sigma^2)$ 的样本，记

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2.$$

试证明统计量 $U = \sqrt{\frac{n-1}{n+1}}\frac{X_{n+1}-\bar{X}_n}{S_n}$ 服从 $t(n-1)$ 分布.

证明 由题意及抽样本分布定理知 X_{n+1}，$\bar{X}_n$ 与 S_n^2 三者相互独立.

由于 $\bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$，$X_{n+1} \sim N(\mu, \sigma^2)$，所以 $X_{n+1} - \bar{X}_n \sim N\left(0, \frac{n+1}{n}\sigma^2\right)$，于是

$$\sqrt{\frac{n}{n+1}}\frac{X_{n+1}-\bar{X}_n}{\sigma} \sim N(0,1),$$

又 $\frac{n}{\sigma^2}S_n^2 \sim \chi^2(n-1)$，且 $X_{n+1} - \bar{X}_n$ 与 S_n^2 独立，所以

$$\frac{\sqrt{\frac{n}{n+1}}\frac{X_{n+1}-\bar{X}_n}{\sigma}}{\sqrt{\frac{n}{\sigma^2}\frac{S_n^2}{n-1}}}=\sqrt{\frac{n-1}{n+1}}\frac{X_{n+1}-\bar{X}_n}{S_n}\sim t(n-1).$$

习 题

1. 设 $X_1,X_2,\cdots,X_n$ 是来自总体 $N(\mu,\sigma^2)$ 的样本，其样本均值与样本方差分别为 $\bar{X},S^2$，求 $E(\bar{X}+S^2)$.

2. 设总体 $X\sim N(25,2^2)$，$X_1,X_2,\cdots,X_{16}$ 为简单随机样本，$\bar{X}$ 为样本均值，求(1) $\bar{X}$ 的数学期望与方差；(2) $P\{|\bar{X}-25|\leqslant 0.3\}$.

3. 从正态总体 $N(4.2,5^2)$ 中抽取容量为 n 的样本，若要求其样本均值位于区间(2.2, 6.2)内的概率不小于 0.95，则样本容量 n 至少取多大？

4. 设两个总体 X,Y 都服从 $N(20,3)$，今分别从两总体中抽得容量为 10 和 15 的相互独立的样本，求 $P\{|\bar{X}-\bar{Y}|>0.3\}$.

§6.4 抽样分布的上 α 分位点

分位点是数理统计学中的一个重要概念，特别是在参数的区间估计与假设检验中都是必不可少的，同时在金融、经济领域中也能起到非常关键的作用.

定义 6.4.1 设随机变量 $Z\sim N(0,1)$，若对 $\alpha\in(0,1)$，实数 z_α 满足

$$P\{Z>z_\alpha\}=\alpha,$$

则称点 z_α 为标准正态分布的上 α 分位点.(图 6-4)

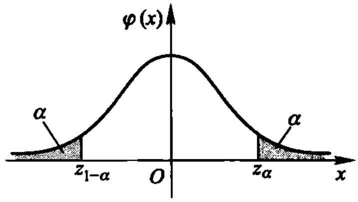


图 6-4

由于标准正态分布的密度函数为偶函数，可知 $z_{1-\alpha}=-z_\alpha$.

在本书附录中，给出了常见的几个统计量的分布表，对于给定的 α 值，可以查到所要求的上 α 分位点值.

例 6.4.1 给定 $\alpha=0.025$ 与 $\alpha=0.05$，查表求 z_α 和 $z_{1-\alpha}$ 值.

解　反查正态分布表(附表3)，得

$$z_{0.025} = 1.96, z_{0.975} = -z_{0.025} = -1.96 .$$

$$z_{0.05} = 1.645, z_{0.95} = -z_{0.05} = -1.645 .$$

定义 6.4.2　设随机变量 $\chi^2 \sim \chi^2(n)$，若对 $\alpha \in (0,1)$，实数 $\chi^2_\alpha(n)$ 满足

$$P\{\chi^2 > \chi^2_\alpha(n)\} = \alpha,$$

则称点 $\chi^2_\alpha(n)$ 为 $\chi^2(n)$ 的上 α 分位点. χ^2 分布的上 α 分位点如图 6－5 所示. 易知

$$P\{\chi^2(n) \leqslant \chi^2_{1-\alpha}(n)\} = \alpha.$$

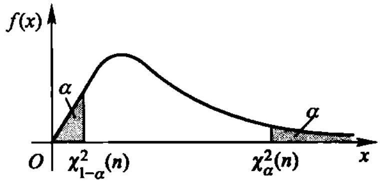


图 6－5

例 6.4.2　给定 $\alpha=0.05$，查表求 $\chi^2_\alpha(10)$ 和 $\chi^2_{1-\alpha}(10)$ 值.

解　由 χ^2 分布表(附表4)可得

$$\chi^2_{0.05}(10) = 18.307, \chi^2_{0.95}(10) = 3.940.$$

定义 6.4.3　设随机变量 $t \sim t(n)$，若对 $\alpha \in (0,1)$，实数 $t_\alpha(n)$ 满足

$$P\{t > t_\alpha(n)\} = \alpha,$$

则称点 $t_\alpha(n)$ 为 $t(n)$ 的上 α 分位点. t 分布的上 α 分位点如图 6－6 所示. 类似于标准正态分布，有

$$t_{1-\alpha}(n) = -t_\alpha(n).$$

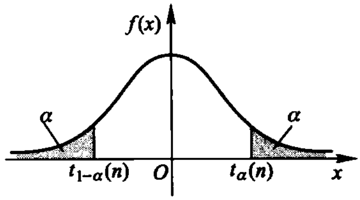


图 6－6

例 6.4.3　给定 $\alpha=0.01$，查表求 $t_\alpha(12)$ 和 $t_{1-\alpha}(12)$ 值.

解　由 t 分布表(附表5)可得

$$t_{0.01}(12) = 2.681, t_{0.99}(12) = -2.681.$$

定义 6.4.4　设随机变量 $F \sim F(n,m)$，若对 $\alpha \in (0,1)$，实数 $F_\alpha(n,m)$ 满足

$$P\{F > F_\alpha(n,m)\} = \alpha,$$

则称点 $F_\alpha(n,m)$ 为 $F(n,m)$ 的上 α 分位点. F 分布的上 α 分位点如图 6－7 所示.

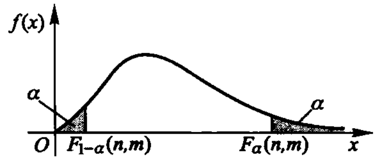


图 6－7

利用 F 分布的性质,读者容易证明

$$F_{1-\alpha}(n,m)=\frac{1}{F_\alpha(m,n)}.$$

例 6.4.4　给定 $\alpha=0.01$,查表求 $F_\alpha(10,8)$ 和 $F_{1-\alpha}(10,8)$.

解　查 F 分布表(附表 6)可知,

$$F_{0.01}(10,8)=5.81,$$

但对 $F_{0.99}(10,8)$ 的情况表中查不到,即只能查到右侧尾部的分位点值. 由

$$F_{1-\alpha}(n,m)=\frac{1}{F_\alpha(m,n)},$$

所以

$$F_{0.99}(10,8)=\frac{1}{F_{0.01}(8,10)}=\frac{1}{5.06}=0.197\ 62.$$

综合以上四个定义,可以看出它们的形式是一样的,可以统一使用下面的定义.

定义 6.4.5　设 Y 为一个连续型随机变量,若对 $\alpha\in(0,1)$,实数 Y_α 满足

$$P\{Y>Y_\alpha\}=\alpha,$$

则称点 Y_α 为 Y 的**上 α 分位点**. 读者容易得到下面的几个等式.

(1) $P\{Y>Y_\alpha\}=\alpha$,

(2) $P\{Y<Y_{1-\alpha}\}=\alpha$,

(3) $P\{Y<Y_{1-\alpha/2}$ 或 $Y>Y_{\alpha/2}\}=\alpha$,

(4) $P\{Y<Y_\alpha\}=1-\alpha$,

(5) $P\{Y>Y_{1-\alpha}\}=1-\alpha$,

(6) $P\{Y_{1-\alpha/2}<Y<Y_{\alpha/2}\}=1-\alpha$.

我们将使用前三个等式处理假设检验问题,使用后三个等式处理区间估计问题.

习　题

1. 给定 $\alpha=0.03$,查表求 z_α 和 $z_{1-\alpha}$.

2. 给定 $\alpha=0.025$,查表求 $\chi_{\alpha}^2(10)$和 $\chi_{1-\alpha}^2(10)$.

3. 给定 $\alpha=0.005$,查表求 $t_{\alpha}(12)$和 $t_{1-\alpha}(12)$.

4. 给定 $\alpha=0.025$,查表求 $F_{\alpha}(10,8)$和 $F_{1-\alpha}(10,8)$.

复 习 题 6

1. 设总体 $X\sim N(0,\sigma^2)$, $X_1,\cdots,X_{10},\cdots,X_{15}$ 为总体的一个样本. 求 $Y=\dfrac{X_1^2+X_2^2+\cdots+X_{10}^2}{2(X_{11}^2+X_{12}^2+\cdots+X_{15}^2)}$ 的分布.

2. 设 $X_1,X_2,\cdots,X_n,X_{n+1},\cdots,X_{n+m}$ 为总体 $X\sim N(0,\sigma^2)$的样本. 求常数 a 与 b,使得 $a\sum\limits_{i=1}^{n}X_i^2+b\left(\sum\limits_{i=n+1}^{n+m}X_i\right)^2$ 服从 χ^2 分布,并求自由度.

3. 设总体 $X\sim N(\mu,4^2)$,$X_1,X_2,\cdots,X_{10}$是来自总体 X 的简单随机样本,S^2 是样本方差. 已知 $P\{S^2>a\}=0.1$,求 a.

4. 设总体 $X\sim N(\mu_1,\sigma^2)$,总体 $Y\sim N(\mu_2,\sigma^2)$,$X_1,X_2,\cdots,X_{n_1}$ 和 $Y_1,Y_2,\cdots,Y_{n_2}$ 分别来自总体 X 和 Y 的简单随机样本,求 $\sum\limits_{i=1}^{n_1}(X_i-\bar{X})^2+\sum\limits_{j=1}^{n_2}(Y_j-\bar{Y})^2$ 的数学期望.

5. 设总体 X 的概率密度为 $f(x)=\dfrac{1}{2}\mathrm{e}^{-|x|}$,$-\infty<x<+\infty$,$X_1,X_2,\cdots,X_n$为总体 X 的简单随机样本,其样本方差为 S^2,求 $E(S^2)$.

6. 设总体 $X\sim N(\mu,\sigma^2)$,$X_1,X_2,\cdots,X_{2n}(n\geqslant 2)$是总体 X 的一个样本,$\bar{X}=\dfrac{1}{2n}\sum\limits_{i=1}^{2n}X_i$,令 $Y=\sum\limits_{i=1}^{n}(X_i+X_{n+i}-2\bar{X})^2$,求 EY.

7. 设总体 $X\sim N(\mu_1,\sigma^2)$与总体 $Y\sim N(\mu_2,\sigma^2)$相互独立,$X_1,X_2,\cdots,X_n$ 与 $Y_1,Y_2,\cdots,Y_m$ 分别为 X 与 Y 的样本,α,β为常数,证明

$$T=\frac{\alpha(\bar{X}-\mu_1)+\beta(\bar{Y}-\mu_2)}{\sqrt{\dfrac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}}\sqrt{\dfrac{\alpha^2}{n}+\dfrac{\beta^2}{m}}}$$

服从自由度为 $n+m-2$ 的 t 分布.

第7章 参数的点估计及其优良性

数理统计的主要任务是通过样本信息来推断总体的信息，即统计推断工作.总体的信息是由总体分布来刻画的，在实际问题中，可以根据问题本身的背景，确定该随机现象的总体所具有的分布类型.但总体中某些参数往往是未知的，例如，以 X 表示某地区居民的身高，并假设 X 服从正态分布 $N(\mu,\sigma^2)$，那么如何求出参数 μ 和 σ^2 就十分重要了.一般来说，μ 和 σ^2 不可能精确求出，为此需要通过从总体中抽取的样本对其进行估计，从而对总体作出推断，这类问题称为参数估计问题.

参数估计是统计推断的基本问题之一，方法上大体上有两类：点估计与区间估计.本章专门介绍参数点估计的概念与方法.

§7.1 点 估 计

点估计即是通过样本值求出总体参数的一个具体估计量或估计值.要得出参数的估计值，首先是得到参数的估计量.具体做法如下：

设总体 X 的分布函数 $F(x;\theta)$ 形式已知，其中含有一个未知参数 θ.为了估计参数 θ，首先从总体 X 中抽取样本 $X_1,X_2,\cdots,X_n$，然后按照一定的方法构造合适的统计量 $g(X_1,X_2,\cdots,X_n)$ 作为 θ 的估计量，记为 $\hat{\theta}=g(X_1,X_2,\cdots,X_n)$.代入样本观测值 $x_1,x_2,\cdots,x_n$，即得到 θ 的估计值 $\hat{\theta}=g(x_1,x_2,\cdots,x_n)$.

θ 的估计量和估计值统称为 θ 的点估计.下面介绍两种应用广泛的点估计方法.

7.1.1 矩估计法

矩估计法是英国数学家 K.皮尔逊在 19 世纪末 20 世纪初提出的，其基本思想是替换原理，即用样本矩替换同阶总体矩.

设总体 X 的分布为 $F(x;\theta)$，θ 为待估参数，$X_1,X_2,\cdots,X_n$ 为来自总体的样本.如果总体 X 的数学期望 $E(X)$ 存在，那么一般来说 $E(X)$ 应为 θ 的函数 $h(\theta)$.由于 $X_1,X_2,\cdots,X_n$ 相互独立且与总体同分布，则由大数定律知，$n\to+\infty$ 时，样本均值 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 依概率收敛于总体均值 $E(X)=h(\theta)$，于是可令

$$EX = \bar{X}, \tag{7.1.1}$$

即

$$h(\theta) = \bar{X},$$

再解此方程求出 θ 即可.

上述做法可以看作是用样本一阶矩 $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ 完成对总体一阶矩 $E(X)$ 的估计.

例 7.1.1 设总体 X 的分布律为

X	1	2	3
p_k	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

其中 $0<\theta<1$ 为未知参数.若现已取得样本值为 $x_1=1, x_2=2, x_3=3$,试求 θ 的矩估计值.

解 因已取得样本值为 $x_1=1, x_2=2, x_3=3$,由总体 X 的分布律可得

$$E(X) = \theta^2 + 2 \cdot 2\theta(1-\theta) + 3 \cdot (1-\theta)^2 = 3 - 2\theta,$$

于是令 $E(X)=3-2\theta=\bar{X}$,可得 θ 的矩估计量为

$$\hat{\theta} = \frac{3-\bar{X}}{2}.$$

根据已取得的样本观测值,可知

$$\bar{x} = \frac{1+2+3}{3} = 2,$$

从而得到 θ 的矩估计值为

$$\hat{\theta} = \frac{3-2}{2} = \frac{1}{2}.$$

例 7.1.2 设总体 X 的概率密度函数为

$$f(x;\theta) = \begin{cases} (\theta+1)x^{\theta}, & 0<x<1, \\ 0, & \text{其他}. \end{cases}$$

其中 $\theta>-1$ 为未知参数. $X_1, X_2, \cdots, X_n$ 为来自总体 X 的一个样本,求 θ 的矩估计量.

解 由于

$$E(X) = \int_{-\infty}^{+\infty} xf(x;\theta)\mathrm{d}x = \int_0^1 x(\theta+1)x^{\theta}\mathrm{d}x = \int_0^1 (\theta+1)x^{\theta+1}\mathrm{d}x = \frac{\theta+1}{\theta+2},$$

则令 $E(X)=\bar{X}$,解得 θ 的矩估计量为

$$\hat{\theta} = \frac{1-2\bar{X}}{\bar{X}-1}.$$

如果总体 X 的未知参数多于一个，假设有 k 个：$\theta_1,\theta_2,\cdots,\theta_k$，该如何估计这些参数呢？此时假设总体 X 的前 k 阶矩 $E(X^i)(i=1,2,\cdots,k)$ 都存在，则由大数定律知，各阶样本矩依概率收敛于同阶总体矩，于是令各阶样本矩与同阶总体矩相等，即

$$E(X^i) = A_i = \frac{1}{n}\sum_{j=1}^{n} x_j^i (i = 1,2,\cdots,k). \tag{7.1.2}$$

由于 $E(X^i)$ 都是 $\theta_1,\theta_2,\cdots,\theta_k$ 的函数，因此从上面 k 个方程中求出参数 $\theta_1,\theta_2,\cdots,\theta_k$，即可得到其对应的矩估计量.

例 7.1.3　设总体 $X\sim U(a,b)$，其中 $a<b$ 为未知参数. $X_1,X_2,\cdots,X_n$ 为来自总体 X 的一个样本，求 a 和 b 的矩估计量.

解　由于 $X\sim U(a,b)$，则

$$E(X) = \frac{a+b}{2}, E(X^2) = D(X) + [E(X)]^2 = \frac{(b-a)^2}{12} + \left(\frac{a+b}{2}\right)^2.$$

令

$$\begin{cases} E(X) = \bar{X}, \\ E(X^2) = A_2, \end{cases}$$

即

$$\begin{cases} \dfrac{a+b}{2} = \bar{X}, \\ \dfrac{(b-a)^2}{12} + \left(\dfrac{a+b}{2}\right)^2 = \dfrac{1}{n}\sum_{i=1}^{n} X_i^2, \end{cases}$$

从而解得 a 和 b 的矩估计量为

$$\begin{cases} \hat{a} = \bar{X} - \sqrt{\dfrac{3}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2}, \\ \hat{b} = \bar{X} + \sqrt{\dfrac{3}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2}. \end{cases}$$

例 7.1.4　设总体为 X，总体均值 $E(X)=\mu$ 和总体方差 $D(X)=\sigma^2$ 存在. $X_1,X_2,\cdots,X_n$ 为来自总体 X 的一个样本，求 μ 和 σ^2 的矩估计量.

解　由

$$\begin{cases} E(X) = \bar{X}, \\ E(X^2) = D(X) + [E(X)]^2 = A_2, \end{cases}$$

即

$$\begin{cases}\mu = \bar{X}, \\ \sigma^2 + \mu^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^2,\end{cases}$$

解得 μ 和 σ^2 的矩估计量为

$$\begin{cases}\hat{\mu} = \bar{X}, \\ \hat{\sigma}^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^2 - \bar{X}^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n} (X_i - \bar{X})^2.\end{cases}$$

矩估计方法直观而又简单，适用性广，特别是估计总体数字特征时直接用到的仅仅是总体的原点矩，而无需知道总体分布的具体形式. 但矩估计法也有缺点，它要求总体矩存在，否则不能使用；此外，矩估计法只利用了矩的信息，而没有充分利用分布对参数所提供的信息. 即便如此，矩估计法还是一种很常用的有效的点估计方法.

7.1.2 最大似然估计法

最大似然估计本意是最大可能性估计，也叫做极大似然估计，是点估计的另一种方法. 它最早是由德国数学家高斯(C. F. Gauss)于 1821 年首先提出. 英国统计学家费希尔(R. A. Fisher)于 1822 年重新提出此概念，并证明了最大似然估计法的一些性质，使得该方法得到了很大的发展.

最大似然估计法是建立在最大似然原理的基础之上. 最大似然原理的直观方法是：设一个随机试验有若干个可能的结果 $A_1, A_2, \cdots, A_n$，若在一次试验中，结果 A_k 出现，则一般认为试验对 A_k 出现有利，即 A_k 出现的概率较大. 这里用到了“概率最大的事件最可能出现”的直观想法. 下面用一个例子说明最大似然估计的思想方法.

假设一个服从离散型分布的总体 X，不妨设 $X \sim B(4, p)$，其中参数 p 未知. 现抽取容量为 3 的样本 X_1, X_2, X_3，如果出现的样本观测值为 1,2,1，此时 p 的取值如何估计比较合理?

考虑这样的问题：出现的样本观测值为什么是 1,2,1，而不是另外一组数 x_1, x_2, x_3? 设事件 $A = \{X_1 = 1, X_2 = 2, X_3 = 1\}$，事件 $B = \{X_1 = x_1, X_2 = x_2, X_3 = x_3\}$. 应用概率论的思想，大概率事件发生的可能性当然比小概率事件发生的可能性大. 现在事件 A 发生了，则可以认为 A 发生的概率比较大，即

$$\begin{aligned} P(A) &= P\{X_1 = 1, X_2 = 2, X_3 = 1\} = P\{X_1 = 1\}P\{X_2 = 2\}P\{X_3 = 1\} \\ &= 96p^4\ (1-p)^8 \end{aligned}$$

应该比较大. 换句话说，p 的取值应该使得 $96p^4(1-p)^8$ 较大才对. 通过计算可知，当 $p=\frac{1}{3}$时，$96p^4(1-p)^8$ 取得最大值. 所以有理由认为 $p=\frac{1}{3}$有利于事件 A 出现，所以 p 的应该在$\frac{1}{3}$左右比较合理.

为了介绍最大似然估计，首先要引进一个概念.

定义 7.1.1　设 $X_1,X_2,\cdots,X_n$ 为来自总体 X 的简单随机样本，$x_1,x_2,\cdots,x_n$ 为样本观测值. 称

$$L(\theta)=\prod_{i=1}^{n}p(x_i,\theta), \tag{7.1.3}$$

为参数 θ 的**似然函数**. 其中，当总体 X 为离散型随机变量时，$p(x_i,\theta)$表示 X 的分布列 $P\{X=x_i\}=p(x_i,\theta)$；当总体 X 为连续性型随机变量时，$p(x_i,\theta)$表示 X 的密度函数函数 $f(x,\theta)$在 x_i 出的取值 $f(x_i,\theta)=p(x_i,\theta)$.

参数 θ 的似然函数 $L(\theta)$实际上就是样本 $X_1,X_2,\cdots,X_n$ 恰好取观测值 $x_1,x_2,\cdots,x_n$(或其邻域)的概率. 如果总体 X 为离散型随机变量时，

$$\begin{aligned}L(\theta)&=P\{X_1=x_1,X_2=x_2,\cdots,X_n=x_n\}\\&=P\{X_1=x_1\}P\{X_2=x_2\}\cdots P\{X_n=x_n\}\\&=\prod_{i=1}^{n}p(x_i,\theta).\end{aligned}$$

如果总体 X 为连续性型随机变量，由于当 Δx_i 非常小时

$$\begin{aligned}P\left\{x_i-\frac{\Delta x_i}{2}<X_i<x_i+\frac{\Delta x_i}{2}\right\}&=P\left\{x_i-\frac{\Delta x_i}{2}<X<x_i+\frac{\Delta x_i}{2}\right\}\\&=\int_{x_i-\frac{\Delta x_i}{2}}^{x_i+\frac{\Delta x_i}{2}}f(x,\theta)\mathrm{d}x\approx f(x_i,\theta)\Delta x_i,\end{aligned}$$

于是

$$\begin{aligned}&P\left\{x_1-\frac{\Delta x_1}{2}<X_1<x_1+\frac{\Delta x_1}{2},x_2-\frac{\Delta x_2}{2}<X_2<x_2+\frac{\Delta x_2}{2},\cdots,x_n-\right.\\&\left.\frac{\Delta x_n}{2}<X_n<x_n+\frac{\Delta x_n}{2}\right\}\\&=\prod_{i=1}^{n}P\left\{x_i-\frac{\Delta x_i}{2}<X_i<x_i+\frac{\Delta x_i}{2}\right\}\approx\prod_{i=1}^{n}f(x_i,\theta)\Delta x_i\\&=L(\theta)\prod_{i=1}^{n}\Delta x_i.\end{aligned}$$

所以 $L(\theta)$表示样本 $X_1,X_2,\cdots,X_n$ 落在观测值 $x_1,x_2,\cdots,x_n$ 附近的概率.

利用“概率最大的事件最可能出现”的想法，结合上面的分析，我们就可以得

到最大似然估计的定义如下.

定义 7.1.2 设 $L(\theta)=\prod\limits_{i=1}^{n}p(x_i,\theta)$ 为参数 θ 的似然函数,若存在一个只与样本观测值 $x_1,x_2,\cdots,x_n$ 有关的实数 $\hat{\theta}(x_1,x_2,\cdots,x_n)$,使得

$$L(\hat{\theta})=\max L(\theta), \tag{7.1.4}$$

则称 $\hat{\theta}(x_1,x_2,\cdots,x_n)$为参数 θ 的**最大似然估计值**. 称 $\hat{\theta}(X_1,X_2,\cdots,X_n)$为参数 θ 的**最大似然估计量**.

由上可知,所谓最大似然估计是指通过求似然函数 $L(\theta)$的最大(或极大)值点来估计参数 θ 的一种方法. 另外,最大似然估计对总体中位置参数的个数没有要求,可以求一个未知参数的最大似然估计,也可以一次求出多个未知参数的最大似然估计. 需要注意的是似然函数 $L(\theta)$不一定有极大值点,但是必然有最大值点,所以对有些问题,利用求驻点的方法求最大似然估计可能失效.

例 7.1.5 求例 7.1.1 中 θ 的最大似然估计值.

解 由 X 的分布列知,θ 的似然函数为

$$L(\theta)=\prod_{i=1}^{3}p(x_i;\theta)=\theta^2\cdot 2\theta(1-\theta)\cdot(1-\theta)^2=2\theta^3(1-\theta)^3,$$

求解

$$\frac{\mathrm{d}}{\mathrm{d}\theta}L(\theta)=6\theta^2(1-\theta)^2(1-2\theta)=0,$$

得 θ 的最大似然估计值为 $\hat{\theta}=\dfrac{1}{2}$.

由于似然函数 $L(\theta)$为多个函数乘积的形式,当 $L(\theta)$比较复杂时,通过直接求导数得到最(极)大值点比较困难. 为简化运算,可以考虑先对 $L(\theta)$取对数,得到**对数似然函数** $\ln L(\theta)$. 由于 $L(\theta)$和 $\ln L(\theta)$具有相同的最(极)大值点,因此求对数似然函数的最(极)大值点即可得到 θ 的最大似然估计.

例 7.1.6 设总体 $X\sim P(\lambda)$,$\lambda>0$ 为未知参数,$X_1,X_2,\cdots,X_n$ 为取自总体的一个样本,求 λ 的最大似然估计量.

解 由于 $X\sim P(\lambda)$,故 λ 的似然函数为

$$L(\lambda)=\prod_{i=1}^{n}\frac{\lambda^{x_i}\mathrm{e}^{-\lambda}}{x_i!}=\frac{\lambda^{\sum\limits_{i=1}^{n}x_i}\mathrm{e}^{-n\lambda}}{\prod\limits_{i=1}^{n}x_i!},$$

取对数似然函数

$$\ln L(\lambda)=\sum_{i=1}^{n}x_i\ln\lambda-n\lambda-\sum_{i=1}^{n}\ln x_i!,$$

求解

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\ln L(\lambda)=\frac{1}{\lambda}\sum_{i=1}^{n}x_i-n=0,$$

解得

$$\lambda=\frac{1}{n}\sum_{i=1}^{n}x_i,$$

因此 λ 的最大似然估计量为

$$\hat{\lambda}=\frac{1}{n}\sum_{i=1}^{n}X_i=\bar{X}.$$

例 7.1.7 在例 7.1.2 中，假设已取得样本观测值 $x_1,x_2,\cdots,x_n$，求 θ 的最大似然估计值.

解 由于总体 X 的概率密度函数为

$$f(x;\theta)=\begin{cases}(\theta+1)x^{\theta}, & 0<x<1,\\ 0, & \text{其他},\end{cases}$$

所以 θ 的似然函数为

$$L(\theta)=\prod_{i=1}^{n}(\theta+1)x_i^{\theta}=(\theta+1)^n\left(\prod_{i=1}^{n}x_i\right)^{\theta},$$

则对数似然函数为

$$\ln L(\theta)=n\ln(1+\theta)+\theta\sum_{i=1}^{n}\ln x_i,$$

求解

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta)=\frac{n}{1+\theta}+\sum_{i=1}^{n}\ln x_i=0,$$

得 θ 的最大似然估计值为

$$\hat{\theta}=-\frac{n}{\sum\limits_{i=1}^{n}\ln x_i}-1.$$

例 7.1.8 设总体 $X\sim N(\mu,\sigma^2)$，$X_1,X_2,\cdots,X_n$ 为来自总体的一个样本，求未知参数 μ 和 σ^2 的最大似然估计量.

解 总体 X 的概率密度函数

$$f(x;\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}},x\in\mathbf{R}.$$

所以 μ 和 σ^2 的似然函数为

$$L(\mu,\sigma^2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x_i-\mu)^2}{2\sigma^2}}=(2\pi\sigma^2)^{-\frac{n}{2}}\mathrm{e}^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{n}(x_i-\mu)^2},$$

取对数，有

$$\ln L(\mu,\sigma^2)=-\frac{n}{2}\ln 2\pi-\frac{n}{2}\ln\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2,$$

求解

$$\begin{cases}\dfrac{\partial\ln L(\mu,\sigma^2)}{\partial\mu}=\dfrac{1}{\sigma^2}\displaystyle\sum_{i=1}^{n}(x_i-\mu)=0,\\ \dfrac{\partial\ln L(\mu,\sigma^2)}{\partial\sigma^2}=-\dfrac{n}{2\sigma^2}+\dfrac{1}{2\sigma^4}\displaystyle\sum_{i=1}^{n}(x_i-\mu)^2=0,\end{cases}$$

得

$$\begin{cases}\mu=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}x_i,\\ \sigma^2=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i-\mu)^2,\end{cases}$$

从而可得 μ 和 σ^2 的最大似然估计量为

$$\begin{cases}\hat{\mu}=\bar{X},\\ \hat{\sigma}^2=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(X_i-\bar{X})^2.\end{cases}$$

可见，它们与相应的矩估计量是相同的.

例 7.1.9 设总体 X 服从区间$[a,b]$上的均匀分布，$X_1,X_2,\cdots,X_n$ 为来自总体的一个样本，求参数 a,b 的最大似然估计量.

解 参数 a,b 的似然函数为

$$L(a,b)=\sum_{i=1}^{n}f(x_i,a,b)=\frac{1}{(b-a)^n}.$$

显然，当 a 越大且 b 越小时，$L(a,b)$越大，另一方面，由于对 $i=1,2,\cdots,n$，有

$$a\leqslant x_i\leqslant b,$$

所以

$$a\leqslant\min\{x_i\}=\max\{x_i\}\leqslant b.$$

于是，当 $a=\min\{x_i\}=x_{(1)}$，$b=\max\{x_i\}=x_{(n)}$时，$L(a,b)$最大，得参数 a,b 的最大似然估计量为

$$\hat{a}=X_{(1)},\quad\hat{b}=X_{(n)}.$$

本题的做法与前面的做法不一样，驻点法失效. 建议读者比较练习题 3 中 4 个小题做法有何不同.

习　　题

1. 设总体 $X \sim e(\lambda)$，求 λ 的矩估计量. 如果测得容量为 10 的样本观测值分别为

134　106　125　115　130　120　110　108　105　115

求 λ 的矩估计值.

2. 设总体 X 具有分布律

X	1	2	3
P	θ	θ	$1-2\theta$

其中 $\theta>0$ 未知，求 θ 的矩估计量和最大似然估计量. 并根据取得的样本观测值 1,1,2,3,1,2,2,3,2,1，求 θ 的矩估计值.

3. 设 $x_1,x_2,\cdots,x_n$ 为来自总体 X 的一组样本观测值，按要求对下列各题中的总体参数 θ 进行估计：

(1) $f(x;\theta)=\begin{cases}\sqrt{\theta}x^{\sqrt{\theta}-1}, & 0\leqslant x\leqslant 1,\theta>0,\\ 0, & \text{其他},\end{cases}$ 求 θ 的矩估计量和最大似然估计量；

(2) $f(x)=\begin{cases}\dfrac{x}{\theta}e^{-\frac{x^2}{2\theta}}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$ 求 θ 的矩估计量和最大似然估计量；

(3) $f(x;\theta)=\begin{cases}e^{-(x-\theta)}, & x>\theta,\\ 0, & \text{其他},\end{cases}$ 求 θ 的矩估计量和最大似然估计量.

4. 设 $X_1,X_2,\cdots,X_n$ 为取自总体 X 的一个样本，总体 X 服从参数为 p 的几何分布，即

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

其中 p 未知且 $0<p<1$，求 p 的最大似然估计量.

5. 设总体 X 的概率密度为 $f(x;\sigma)=\dfrac{1}{2\sigma}e^{-\frac{|x|}{\sigma}}(-\infty<x<+\infty)$，其中 $\sigma>0$ 未知，设 $X_1,X_2,\cdots,X_n$ 为取自这个总体的一个样本. 求 σ 的最大似然估计量.

§7.2　点估计优良性的评定标准

由前面的一些例子可以看到，虽然总体分布中的参数是确定的，但对同一参数采取不同的估计法，可能得到不同的估计量. 从参数估计本身来看，原则上任何统计量都可以作为未知参数的估计量. 那么不同的估计量中哪一个更好？如何做出选择，一定要有标准. 通常采用的标准有三个：无偏性、有效性、一致性.

7.2.1　无偏性

参数的估计量是一个统计量，对于不同的样本值所求得的参数估计值，一般

是不相同的，所以估计量也是一个随机变量. 因此要确定一个估计量的优劣，就不能仅仅依赖于某一次试验的结果来衡量，而是希望这个估计量在多次试验的结果中，在待估参数的附近随机摆动，并使得这个估计量的平均值恰好就是待估的参数，由此引出无偏性的标准.

定义 7.2.1 若参数 θ 的估计量 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$ 满足

$$E(\hat{\theta}) = \theta, \tag{7.2.1}$$

则称 $\hat{\theta}$ 为 θ 的一个无偏估计量，否则就称为有偏估计量.

例 7.2.1 设 $X_1,X_2,\cdots,X_n$ 为总体 X 的一个样本，已知 $E(X)=\mu$，$D(X)=\sigma^2$.

(1) 证明：样本均值 $\bar{X}$ 和样本方差 S^2 分别是 μ 和 σ^2 的无偏估计量.

(2) 判定 $\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 是否是 σ^2 的无偏估计量.

(1) **证明** 因为

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)= \frac{1}{n}\sum_{i=1}^{n}E(X_i) = \frac{1}{n}\cdot n\mu = \mu,$$

$$\begin{aligned}E(S^2) &= E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2\right]\\
&= \frac{1}{n-1}E\left(\sum_{i=1}^{n}X_i^2 - n\bar{X}^2\right)\\
&= \frac{1}{n-1}\left(\sum_{i=1}^{n}E(X_i^2) - nE(\bar{X}^2)\right)\\
&= \frac{1}{n-1}\left(\sum_{i=1}^{n}(\mu^2+\sigma^2) - n[D(\bar{X})+E^2(\bar{X})]\right)\\
&= \frac{1}{n-1}(n\mu^2+n\sigma^2-\sigma^2-n\mu^2) = \sigma^2,\end{aligned}$$

故样本均值 $\bar{X}$ 和样本方差 S^2 分别是 μ 和 σ^2 的无偏估计量.

(2) **解** 因为

$$E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2\right] = \frac{1}{n}\sum_{i=1}^{n}[E(X_i^2)-E(\bar{X}^2)]$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left[(\sigma^2+\mu^2)-\left(\frac{\sigma^2}{n}+\mu^2\right)\right]$$

$$= \frac{n-1}{n}\sigma^2 \neq \sigma^2,$$

所以 $\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 不是 σ^2 的无偏估计量.

显然,如果将 $\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 乘个系数 $\frac{n}{n-1}$,就修改成一个无偏估计量,即 $\frac{n}{n-1}\cdot\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 是 σ^2 的无偏估计量.这就是为什么用 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 定义样本方差的原因.

例 7.2.2　设 X_1,X_2,X_3 是来自总体 X 的一个样本,问下列估计量哪一个是总体均值 μ 的无偏估计量?

$$\hat{\mu}_1=\frac{1}{6}X_1+\frac{1}{3}X_2+\frac{1}{2}X_3,$$

$$\hat{\mu}_2=\frac{2}{5}X_1+\frac{2}{5}X_2+\frac{1}{5}X_3,$$

$$\hat{\mu}_3=\frac{1}{3}X_1+\frac{2}{9}X_2+\frac{1}{7}X_3.$$

解　因为

$$E(\hat{\mu}_1)=E\left(\frac{1}{6}X_1+\frac{1}{3}X_2+\frac{1}{2}X_3\right)=\frac{1}{6}E(X_1)+\frac{1}{3}E(X_2)+\frac{1}{2}E(X_3)$$
$$=\frac{1}{6}\mu+\frac{1}{3}\mu+\frac{1}{2}\mu=\mu,$$

同理

$$E(\hat{\mu}_2)=\frac{2}{5}E(X_1)+\frac{2}{5}E(X_2)+\frac{1}{5}E(X_3)=\mu,$$

$$E(\hat{\mu}_3)=\frac{1}{3}E(X_1)+\frac{2}{9}E(X_2)+\frac{1}{7}E(X_3)=\frac{44}{63}\mu\neq\mu.$$

因此 $\hat{\mu}_1,\hat{\mu}_2$ 是总体均值 μ 的无偏估计量,而 $\hat{\mu}_3$ 不是总体均值 μ 的无偏估计量.

例 7.2.3　设总体 $X\sim N(\mu,\sigma^2)$,其中 μ,σ^2 未知. $X_1,X_2,\cdots,X_6$ 为来自总体 X 的一个样本,试确定常数 C,使 $CY=C[(X_1-X_2)^2+(X_3-X_4)^2+(X_5-X_6)^2]$ 是 σ^2 的无偏估计量.

解　由于

$$\begin{aligned}E[(X_1-X_2)^2]&=D(X_1-X_2)+E^2(X_1-X_2)\\&=D(X_1)+D(X_2)+[E(X_1)-E(X_2)]^2\\&=\sigma^2+\sigma^2+(\mu-\mu)^2=2\sigma^2,\end{aligned}$$

同理可得

$$E[(X_3-X_4)^2]=E[(X_5-X_6)^2]=2\sigma^2,$$

所以

$$
\begin{aligned}
&E\{C[(X_1-X_2)^2+(X_3-X_4)^2+(X_5-X_6)^2]\}\\
&=C\{E[(X_1-X_2)^2]+E[(X_3-X_4)^2]+E[(X_5-X_6)^2]\}\\
&=6C\sigma^2.
\end{aligned}
$$

因此,若 $CY=C[(X_1-X_2)^2+(X_3-X_4)^2+(X_5-X_6)^2]$是 σ^2 的无偏估计量,则

$$E(CY)=6C\sigma^2=\sigma^2,$$

由此解得 $C=\frac{1}{6}$.

无偏性是对估计量的基本要求,它具有系统误差为零的特点. 在例 7.2.2 中,我们看到对于总体均值 μ,可以有不同的无偏估计量$\hat{\mu}_1$和$\hat{\mu}_2$. 那么哪一个无偏估计量更好?

直观上说,如果$\hat{\theta}_1$和$\hat{\theta}_2$都是 θ 的无偏估计量,其取值都在 θ 周围波动. 如果$\hat{\theta}_1$的取值比$\hat{\theta}_2$的取值更集中地聚集在 θ 的邻近,则认为用$\hat{\theta}_1$来估计 θ 比$\hat{\theta}_2$更好些. 由于方差是随机变量取值与其数学期望偏离程度的度量,所以无偏估计以方差小者为好. 由此引出估计量有效性的概念.

7.2.2 有效性

定义 7.2.2 设$\hat{\theta}_1$和$\hat{\theta}_2$都是参数 θ 的无偏估计量,如果

$$D(\hat{\theta}_1)<D(\hat{\theta}_2), \tag{7.2.2}$$

则称$\hat{\theta}_1$比$\hat{\theta}_2$有效.

有效性的定义指明,在期望相等的条件下,方差小者估计的效果更好.

例 7.2.4 试判断在例 7.2.2 中,总体均值 μ 的无偏估计量$\hat{\mu}_1$和$\hat{\mu}_2$哪一个更有效.

解 比较 $D(\hat{\mu}_1)$和 $D(\hat{\mu}_2)$的大小即可. 因为

$$D(\hat{\mu}_1)=D\left(\frac{1}{6}X_1+\frac{1}{3}X_2+\frac{1}{2}X_3\right)=\frac{1}{36}\sigma^2+\frac{1}{9}\sigma^2+\frac{1}{4}\sigma^2=\frac{7}{18}\sigma^2,$$

$$D(\hat{\mu}_2)=D\left(\frac{2}{5}X_1+\frac{2}{5}X_2+\frac{1}{5}X_3\right)=\frac{4}{25}\sigma^2+\frac{4}{25}\sigma^2+\frac{1}{25}\sigma^2=\frac{9}{25}\sigma^2,$$

则 $D(\hat{\mu}_1)>D(\hat{\mu}_2)$,故总体均值 μ 的无偏估计量$\hat{\mu}_2$比无偏估计量$\hat{\mu}_1$更有效.

7.2.3 一致性(相合性)

容易看出,估计量 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$与样本容量 n 有关, 为明确起见,不妨将其记为$\hat{\theta}_n$. 一般来说,当 n 越大,$\hat{\theta}_n$的取值与 θ 的误差应越小. 即当 n 充分大

时,估计量$\hat{\theta}_n$的取值应稳定在参数θ的一个充分小的邻域内,于是就有一致性的标准.

定义 7.2.3　设$\hat{\theta}_n=\hat{\theta}(X_1,X_2,\cdots,X_n)$是$\theta$的一个估计量,若对于任意的$\varepsilon>0$,有

$$\lim_{n\to+\infty}P\{|\hat{\theta}_n-\theta|<\varepsilon\}=1, \tag{7.2.3}$$

则称$\hat{\theta}_n$是θ的**一致估计量**(或称**相合估计量**).

例 7.2.5　已知$X_1,X_2,\cdots,X_n$是来自总体X的一个样本,且$E(X)=\mu$.证明:样本均值$\overline{X}$是总体均值μ的一致估计量.

证明　因为样本$X_1,X_2,\cdots,X_n$相互独立且与总体X同分布,则有

$$E(X_i)=\mu,i=1,2,\cdots,n.$$

由切比雪夫大数定律可知

$$\lim_{n\to+\infty}P\{|\overline{X}-\mu|<\varepsilon\}=\lim_{n\to+\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right|<\varepsilon\right\}=1,$$

故$\overline{X}$是总体均值μ的一致估计量.

对于估计量,我们当然希望寻求它的一致估计量,但是这要求样本容量必须相当大,往往难以做到.因此在实际问题中,更多的是使用无偏性和有效性这两个标准.

习　题

1. 设$X_1,X_2,\cdots,X_n$为取自总体X的一个样本.已知期望$E(X)=0$,而方差$D(X)=\sigma^2$是未知参数,试确定k,使$T=k\sum_{i=1}^{n}X_i^2$是σ^2的无偏估计量.

2. 设$X_1,X_2,\cdots,X_n$是取自总体X的样本,$EX=\mu,DX=\sigma^2$,$\hat{\sigma}^2=k\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2$,问$k$为何值时$\hat{\sigma}^2$为$\sigma^2$的无偏估计.

3. 设$\hat{\theta}$是θ的无偏估计量,且有$D(\hat{\theta})>0$,试证:$\hat{\theta}^2$不是θ^2的无偏估计量.

4. 设$X_1,X_2,\cdots,X_n$为来自总体$X\sim N(\mu,\sigma^2)$的样本,在下列μ的无偏估计量中,最有效的是哪一个?

(1) $\frac{1}{2}X_1+\frac{1}{3}X_2+\frac{1}{6}X_3$;

(2) $\frac{1}{2}X_1+\frac{1}{4}X_2+\frac{1}{4}X_3$;

(3) $\frac{1}{3}X_1+\frac{1}{3}X_2+\frac{1}{3}X_3$.

5. 若$\hat{\theta}_1$和$\hat{\theta}_2$是θ的两个相互独立的无偏估计量，且$D(\hat{\theta}_1)=2D(\hat{\theta}_2)$，问常数$a$和$b$满足什么条件，才能使$a\hat{\theta}_1+b\hat{\theta}_2$是$\theta$的无偏估计量？$a$和$b$取何值时，$\theta$的无偏估计量$a\hat{\theta}_1+b\hat{\theta}_2$最有效？

§7.3 综 合 例 题

例 7.3.1 设$X_1,X_2,\cdots,X_n$是取自总体X的样本，且$EX=\mu,DX=\sigma^2$.

(1) 证明：如果$a_i(i=1,2,\cdots,n)$满足$\sum_{i=1}^n a_i=1$，则$\sum_{i=1}^n a_iX_i$是μ的无偏估计量.

(2) 如果$\sum_{i=1}^n a_iX_i$是μ的无偏估计量，求a_i，使得$\sum_{i=1}^n a_iX_i$的方差最小.

解 (1) 因为

$$E\Big(\sum_{i=1}^n a_iX_i\Big)=\sum_{i=1}^n a_iEX_i=\sum_{i=1}^n a_i\mu=\mu\sum_{i=1}^n a_i,$$

所以当$\sum_{i=1}^n a_i=1$时，$\sum_{i=1}^n a_iX_i$是μ的无偏估计量.

(2) 利用$X_1,X_2,\cdots,X_n$的独立同分布性，可得

$$D\Big(\sum_{i=1}^n a_iX_i\Big)=\sum_{i=1}^n a_i^2DX_i=\sum_{i=1}^n a_i^2\sigma^2=\sigma^2\sum_{i=1}^n a_i^2.$$

为了求$\sigma^2\sum_{i=1}^n a_i^2$的最小值，应用拉格朗日乘数法，为此构造目标函数

$$F=\sum_{i=1}^n a_i^2+\lambda\Big(\sum_{i=1}^n a_i-1\Big).$$

令$\frac{\partial F}{\partial a_i}=2a_i+\lambda=0,i=1,2,\cdots,n$，及$\sum_{i=1}^n a_i=1$，可以得到当$a_1=a_2=\cdots=a_n=\frac{1}{n}$时，$\sum_{i=1}^n a_iX_i$的方差最小，即在所有形为$\sum_{i=1}^n a_iX_i$的$\mu$的无偏估计量中，样本均值$\bar{X}$的方差最小，且最小值为$\frac{\sigma^2}{n}$.

例 7.3.2 设总体X服从指数分布，其概率密度函数为

$$f(x;\theta)=\begin{cases}\frac{1}{\theta}e^{-\frac{x}{\theta}}, & x>0,\\ 0, & 其他,\end{cases}$$

其中 $\theta>0$ 未知，$X_1,X_2,\cdots,X_n$ 为来自总体 X 的一个样本.

(1) 证明：样本均值 $\bar{X}$ 和 $n\min\{X_i\}$ 都是 θ 的无偏估计量.

(2) 证明：当 $n>1$ 时，$\bar{X}$ 比 $n\min\{X_i\}$ 更有效.

证明 (1) 因为 $E\bar{X}=EX=\theta$，所以样本均值 $\bar{X}$ 是 θ 的无偏估计量.

因为 $\min\{X_i\}$ 的密度函数为

$$f(x)=\begin{cases}\dfrac{n}{\theta}e^{-\frac{n}{\theta}x}, & x>0,\\ 0, & \text{其他},\end{cases}$$

所以可得

$$E\Big(n\min\{X_i\}\Big)=\int_0^{\infty}nx\,\frac{n}{\theta}e^{-\frac{n}{\theta}x}\,dx=\theta,$$

于是 $n\min\{X_i\}$ 也是 θ 的无偏估计量.

(2) 由于 $D\bar{X}=\dfrac{1}{n}DX=\dfrac{\theta}{n}$，又利用 $\min\{X_i\}$ 的密度函数容易求出

$$D\Big(n\min\{X_i\}\Big)=\theta^2.$$

所以 $\bar{X}$ 比 $n\min\{X_i\}$ 更有效.

例 7.3.3 已知总体 X 是一个离散型随机变量，其分布列为

$$P\{X=k\}=a_k\frac{\theta^k}{g(\theta)},\quad k=0,1,2,\cdots$$

且对某些 k，可能有 $a_k=0$，$g(\theta)$ 有连续的导函数. 设 $X_1,X_2,\cdots,X_n$ 为来自总体 X 的一个样本.

(1) 证明：$EX=\dfrac{\theta g'(\theta)}{g(\theta)}$；

(2) 证明：θ 的最大似然估计量为方程

$$\bar{X}=\frac{\theta g'(\theta)}{g(\theta)}=EX$$

的一个根，即此时 θ 的最大似然估计与矩估计是一样的.

证明 (1) 因为 $P\{X=k\}=a_k\dfrac{\theta^k}{g(\theta)}$，$k=0,1,2,\cdots$，所以有

$$\sum_{k=0}^{\infty}a_k\frac{\theta^k}{g(\theta)}=1,$$

即

$$g(\theta) = \sum_{k=0}^{\infty} a_k \theta^k,$$

两边求导

$$g'(\theta) = \sum_{k=0}^{\infty} k a_k \theta^{k-1} = \frac{1}{\theta} \sum_{k=0}^{\infty} k a_k \theta^k.$$

利用数学期望的定义可知 $EX = \sum_{k=0}^{\infty} k a_k \frac{\theta^k}{g(\theta)} = \frac{\theta}{g(\theta)} \sum_{k=1}^{\infty} k a_k \theta^{k-1}$，所以有

$$EX = \frac{\theta g'(\theta)}{g(\theta)}.$$

(2) θ 的似然函数为

$$L(\theta) = \prod_{i=1}^{n} a_{x_i} \frac{\theta^{x_i}}{g(\theta)},$$

对数似然函数为

$$\ln L(\theta) = \sum_{i=1}^{n} (\ln a_{x_i} + x_i \ln \theta - \ln g(\theta)),$$

令 $\frac{\partial \ln L(\theta)}{\partial \theta} = 0$，可得 $\frac{1}{\theta} \sum_{i=1}^{n} x_i = n \frac{g'(\theta)}{g(\theta)}$，即

$$\bar{x} = \frac{\theta g'(\theta)}{g(\theta)}.$$

所以 θ 的最大似然估计量为方程

$$\bar{X} = \frac{\theta g'(\theta)}{g(\theta)} = EX$$

的一个根.

复 习 题 7

1. 设总体 X 的密度函数

$$f(x,\theta) = \begin{cases} \frac{2}{\theta^2}(\theta - x), & 0 < x < \theta, \\ 0, & \text{其他}. \end{cases}$$

$X_1, X_2, \cdots, X_n$ 为其样本，试求参数 θ 的矩法估计.

2. 设总体 X 的密度函数为

$$f(x,\theta) = \begin{cases} \frac{6x}{\theta^3}(\theta - x), & 0 < x < \theta, \\ 0, & \text{其他}. \end{cases}$$

$X_1,X_2,\cdots,X_n$ 为总体 X 的一个样本.

(1) 求 θ 的矩估计量;

(2) 求 $D(\hat{\theta})$.

3. 设某种电子元件的使用寿命 X 的概率密度函数为

$$f(x,\theta)=\begin{cases}2e^{-2(x-\theta)}, & x>\theta,\\ 0, & x\leqslant\theta.\end{cases}$$

其中 $\theta(\theta>0)$为未知参数,又设 $x_1,x_2,\cdots,x_n$ 是总体 X 的一组样本观察值,求 θ 的最大似然估计值.

4. 设总体 X 的概率分布为

X	0	1	2	3
p_i	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中 $\theta\left(0<\theta<\dfrac{1}{2}\right)$是未知参数,利用总体的如下样本值3,1,3,0,3,1,2,3,求 θ 的矩估计值和最大似然估计值.

5. 设总体 X 的分布函数为

$$F(x,\alpha,\beta)=\begin{cases}1-\dfrac{\alpha^\beta}{x^\beta}, & x>\alpha,\\ 0, & x\leqslant\alpha.\end{cases}$$

其中未知参数 $\beta>1,\alpha>0$,设 $X_1,X_2,\cdots,X_n$ 为来自总体 X 的样本

(1) 当 $\alpha=1$ 时,求 β 的矩估计量;

(2) 当 $\alpha=1$ 时,求 β 的最大似然估计量;

(3) 当 $\beta=2$ 时,求 α 的最大似然估计量.

6. 设总体 X 的概率密度为

$$f(x,\theta)=\begin{cases}\theta, & 0<x<1,\\ 1-\theta, & 1\leqslant x<2,\\ 0, & \text{其他}.\end{cases}$$

其中 θ 是未知参数$(0<\theta<1)$, $X_1,X_2,\cdots,X_n$ 为来自总体 X 的简单随机样本,记 N 为样本值 $x_1,x_2,\cdots,x_n$ 中小于1的个数.

(1) 求 θ 的矩估计;

(2) 求 θ 的最大似然估计.

第 8 章　参数的区间估计与假设检验

从总体中取得样本观测值后，通过点估计的方法，我们可以求出未知参数 θ 的估计值，但是由于样本是随机的，所以我们无法知道这个估计值与参数的真值之间的误差到底有多大，这时如果能知道误差的范围那也是一个不错的选择. 为了弥补点估计的不足，本章将首先讨论区间估计的概念与方法.

§8.1　区 间 估 计

在实际应用时，由于参数不可能精确求出，所以人们往往更加关注该参数大致的取值范围. 例如，一些小的电器设备包装盒里，都会有使用说明书. 一般来说，使用说明书上注明的额定电流不是一个数，常常是“额定电流：1A±0.05A”这种格式. 这里有 0.05A 的偏差，它表明当额定电流(参数 θ)在 0.95A 与 1.05A 之间时，电器会处于正常状态，否则可能出问题. 这里以区间[0.95A，1.05A]给出额定电流 θ 的取值范围，要求额定电流“大致”在 0.95A 与 1.05A 之间，这种方法称为 θ 的区间估计. 这里有几个问题需要澄清：

(1) “大致”在 0.95A 与 1.05A 之间中的“大致”是什么意思?

(2) 额定电流 θ 有无可能在区间[0.95A，1.05A]之外?

(3) 0.95A 与 1.05A 是如何求出的?

额定电流 θ 是未知的，并且具有随机性的特点，它完全有可能在[0.95A，1.05A]之外取值. 那么，这个可能性有多大? 实际问题中，当然希望额定电流值在[0.95A，1.05A]上的可能性尽可能地大，而落在这个区间之外的可能性尽可能地小. 一般来说，如果没有特殊说明，要求区间[0.95A，1.05A]能包含额定电流 θ 的概率为 0.95，即有 95%的把握保证区间[0.95A，1.05A]包含额定电流 θ，而落在区间[0.95A，1.05A]之外的可能性是 0.05. 这里将区间[0.95A，1.05A]称为额定电流 θ 的置信区间.

需要说明的是，用区间给出未知参数 θ 的取值范围，也要依赖于样本信息. 也就是说，必须利用样本 $X_1,X_2,\cdots,X_n$ 来求置信区间. 由于置信区间完全由区间的端点决定，因此置信区间的两个端点是由样本 $X_1,X_2,\cdots,X_n$ 确定的两个统计量.

下面先给出置信区间的定义.

定义　设总体 X 的分布函数为 $F(x;\theta)$，其中 θ 为未知参数，$X_1,X_2,\cdots,X_n$

为来自总体的简单随机样本. 对于给定的 $\alpha \in (0,1)$，如果由样本确定的两个统计量 $T_1(X_1, X_2, \cdots, X_n)$ 和 $T_2(X_1, X_2, \cdots, X_n)$ 满足

$$P\{T_1 \leqslant \theta \leqslant T_2\} = 1 - \alpha, \tag{8.1.1}$$

则称随机区间 $[T_1, T_2]$ 是参数 θ 的置信度(或置信水平)为 $1-\alpha$ 的置信区间.

如果统计量 $T_1(X_1, X_2, \cdots, X_n)$ 满足

$$P\{\theta \leqslant T_1\} = 1 - \alpha(\text{或 } P\{\theta \geqslant T_1\} = 1 - \alpha), \tag{8.1.2}$$

则称 T_1 是参数 θ 的置信度为 $1-\alpha$ 的单侧置信上限(或单侧置信下限).

这里有几点说明：

(1) 所谓置信度就是指可信度，代表用区间 $[T_1, T_2]$ 估计参数 θ 的可靠程度. 例如置信度 $1-\alpha=0.95$，它表示如果反复抽样多次(假设各次得到的样本容量相等，均为 n)，每个样本都可以确定一个区间 $[T_1, T_2]$，这样的区间要么包含 θ 的真值，要么不包含 θ 的真值. 理论上来说，在这些区间中，包含 θ 的真值的区间约占 $100(1-\alpha)\%$，不包含 θ 的真值的区间仅占 $100\alpha\%$. 例如，如果反复抽样 100 次，可以得到 100 个区间 $[T_1, T_2]$，若 $\alpha=0.05$，那么理论上讲其中大约有 95 个区间包含 θ 的真值，而不包含 θ 的真值的区间大约仅有 5 个. 一般来说，在求置信区间之前必须先给出置信度. 如果没有明确指出，约定置信度取 0.95.

(2) 置信区间的长度 $T_2 - T_1$ 代表估计的精度，我们希望它越小越好. 而同时还希望区间 $[T_1, T_2]$ 包含 θ 的真值的概率越大越好，也就是置信度越高越好. 但这两方面是相互制约的. 置信度越高，置信区间长度越大；置信度降低，置信区间的长度就会减少. 因此，奈曼(Neyman)提出的原则是，通常优先考虑置信度，即在满足可信度 $P\{T_1 \leqslant \theta \leqslant T_2\} = 1 - \alpha$ 的前提下，再去求给定的置信度下精度最高的置信区间. 否则，只有增加样本容量，才能解决.

(3) 如果总体是一般的分布，那么参数的区间估计是非常复杂的，很少有人去研究. 如果总体是正态分布，那么参数的区间估计会变得比较简单，在这一章中，我们也只研究正态总体参数的置信区间问题.

求未知参数 θ 的置信区间，就是找出满足条件的区间端点 T_1 和 T_2，这两个统计量不但受到置信度的制约，同时也依赖抽样分布的情形. 抽样分布不同，则结果就存在差别. 那么，如何确定 T_1 和 T_2，来得到置信区间 $[T_1, T_2]$？由于我们所遇到的总体在大多数情况下服从正态分布，或是近似服从正态分布，因而主要讨论正态总体下的总体参数的置信区间.

对于给定的置信度 $1-\alpha$，参数 θ 的区间估计的基本思想是：

① 从正态总体的六个已知的抽样分布(单总体、双总体各 3 个)中选一个只含有待估参数 θ 而不包含其他未知参数的抽样分布，记为 Y.

② 由于 Y 一定是四大分布($N(0,1), \chi^2, t, F$)中的一个，所以它的分位点

$Y_{\alpha/2}$,$Y_{1-\alpha/2}$能从表中查到,且必然满足

$$P\{Y_{1-\alpha/2} < Y < Y_{\alpha/2}\} = 1-\alpha. \tag{8.1.3}$$

③ 由于Y中含有待估参数θ,所以可以从不等式$Y_{1-\alpha/2}<Y<Y_{\alpha/2}$中解出$\theta$,得到等价的不等式$T_1<\theta<T_2$.

由于$P\{T_1\leqslant\theta\leqslant T_2\}=1-\alpha$,且$T_1$,$T_2$中没有其他未知数,是统计量,所以$[T_1,T_2]$就是$\theta$的置信度为$1-\alpha$的置信区间.

当然,如果在②中使用$P\{Y\leqslant Y_\alpha\}=1-\alpha$或$P\{Y\geqslant Y_{1-\alpha}\}=1-\alpha$,就可以求出$\theta$的单侧置信限.

例 8.1.1 某地区的磁场强度$X\sim N(\mu,20^2)$,现从该地区取36个点,测得样本均值为$\bar{x}=61.1$,在0.95的置信度下求该地区平均磁场强度μ的置信区间.

解 本题是单总体问题,且是求μ的置信区间,由于σ已知,所以利用$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$可得:

$$P\left\{z_{1-\alpha/2} \leqslant \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leqslant z_{\alpha/2}\right\} = 1-\alpha,$$

解不等式得

$$P\left\{\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leqslant \mu \leqslant \bar{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right\} = 1-\alpha,$$

所以μ的置信度为$1-\alpha$的置信区间为

$$\left[\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right].$$

已知样本容量$n=36$,样本均值为$\bar{x}=61.1$,总体方差$\sigma^2=20^2$,$\alpha=0.05$,代入数据得平均磁场强度μ的置信区间为

$$[54.6,67.6].$$

例 8.1.2 已知某种灯泡的寿命服从正态分布,现从一批灯泡中随机抽取16只作为样本,测得其平均使用寿命为1 490 h,样本标准差为25.4 h,在0.95的置信度下求这批灯泡平均使用寿命μ的置信区间.

解 本题是单总体问题,求μ的置信区间,由于σ未知,所以利用$\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$可得:

$$P\left\{t_{1-\alpha/2}(n-1) \leqslant \frac{\bar{X}-\mu}{S/\sqrt{n}} \leqslant t_{\alpha/2}(n-1)\right\} = 1-\alpha,$$

解不等式μ的置信度为$1-\alpha$的置信区间为

$$\left[\bar{X}-t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}},\bar{X}-t_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right].$$

将 $n=16,\bar{x}=1\ 490,s=25.4,\alpha=0.05$ 代入得 μ 的置信区间为

$$[1\ 476.46,\quad 1\ 503.54].$$

例 8.1.3 设某车间生产的滚珠的直径 $X\sim N(\mu,\sigma^2)$，现从某日生产的滚珠中抽取 9 个，测得样本方差为 $s^2=0.25^2$，在 0.95 的置信度下求总体方差的 σ^2 的置信区间.

解 本题是单总体问题，求 σ^2 的置信区间，由于 μ 未知，所以利用

$$\frac{n-1}{\sigma^2}S^2\sim\chi^2(n-1),$$

得

$$P\left\{\chi^2_{1-\frac{\alpha}{2}}(n-1)\leqslant\frac{n-1}{\sigma^2}S^2\leqslant\chi^2_{\frac{\alpha}{2}}(n-1)\right\}=1-\alpha,$$

从而得到 σ^2 的置信区间为

$$\left[\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)},\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right].$$

将 $n=9,s^2=0.25^2,\chi^2_{0.025}(8)=17.535,\chi^2_{0.975}(8)=2.180$ 代入，得 σ^2 的置信区间为

$$[0.03,0.23].$$

例 8.1.4 某车间用两台型号相同的机器生产同一种产品，已知机器 A 生产的产品长度 $X\sim N(\mu_1,1)$，机器 B 生产的产品长度 $Y\sim N(\mu_2,1)$. 为了比较两台机器生产的产品的长度，现从 A 生产的产品中抽取 10 件，测得样本均值 $\bar{x}=49.83$ cm. 从 B 的产品中抽取 15 件，测得样本均值 $\bar{y}=50.24$ cm. 在 0.99 的置信度下求两总体均值差 $\mu_1-\mu_2$ 的置信区间.

解 本题是双总体问题，求 $\mu_1-\mu_2$ 的置信区间，且两个总体方差 $\sigma_1^2=\sigma_2^2=1$ 已知，所以利用

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim N(0,1),$$

得

$$P\left\{z_{1-\frac{\alpha}{2}}\leqslant\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\leqslant z_{\frac{\alpha}{2}}\right\}=1-\alpha,$$

从而得到 σ^2 的置信区间为

$$\left[(\bar{X}-\bar{Y})-z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},(\bar{X}-\bar{Y})+z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right].$$

将 $n_1=10,n_2=15,\bar{x}=49.83,\bar{y}=50.24,z_{0.005}=2.58$,代入得 $\mu_1-\mu_2$ 的置信区间为

$$[-1.46,0.64].$$

例 8.1.5 某车间用两台型号相同的机器生产同一种产品,机器 A 生产的产品长度 $X\sim N(\mu_1,\sigma^2)$,机器 B 生产的产品长度 $Y\sim N(\mu_2,\sigma^2)$. 为了比较两台机器生产的产品的长度,现从 A 的产品中抽取 10 件,测得样本均值 $\bar{x}=49.83$ cm,样本标准差 $s_1=1.09$ cm. 从 B 的产品中抽取 15 件,测得样本均值 $\bar{y}=50.24$ cm,样本标准差 $s_2=1.18$ cm. 在 0.99 的置信度下,求两总体均值差 $\mu_1-\mu_2$ 的置信区间.

解 本题是双总体问题,求 $\mu_1-\mu_2$ 的置信区间,且两个总体方差相等但未知,所以利用

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_\omega\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t(n_1+n_2-2)$$

得

$$P\left\{t_{1-\alpha/2}(n_1+n_2-2)\leqslant\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_\omega\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\leqslant t_{\alpha/2}(n_1+n_2-2)\right\}=1-\alpha,$$

所以 $\mu_1-\mu_2$ 的置信区间为

$$\left[\bar{X}-\bar{Y}-t_{\alpha/2}(n_1+n_2-2)S_\omega\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\right.$$
$$\left.\bar{X}-\bar{Y}+t_{\alpha/2}(n_1+n_2-2)S_\omega\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right].$$

将 $n_1=10,n_2=15,\bar{x}=49.83,s_1=1.09,\bar{y}=50.24,s_2=1.18,t_{0.005}(23)=2.8073$ 代入得 $\mu_1-\mu_2$ 的置信区间为

$$[-1.72,0.9].$$

例 8.1.6 在 0.95 的置信度下,求例 7.1.5 中两总体方差比 $\frac{\sigma_1^2}{\sigma_2^2}$ 的置信区间.

解 本题为双总体问题,求 $\frac{\sigma_1^2}{\sigma_2^2}$ 的置信区间,于是利用

$$\frac{S_1^2}{S_2^2}\cdot\frac{\sigma_2^2}{\sigma_1^2}\sim F(n_1-1,n_2-1),$$

得

$$P\left\{F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)\leqslant\frac{S_1^2}{S_2^2}\cdot\frac{\sigma_2^2}{\sigma_1^2}\leqslant F_{\frac{\alpha}{2}}(n_1-1,n_2-1)\right\}=1-\alpha,$$

因此得到$\frac{\sigma_1^2}{\sigma_2^2}$的置信区间为

$$\left[\frac{S_1^2}{S_2^2}\frac{1}{F_{\frac{\alpha}{2}}(n_1-1,n_2-1)},\quad\frac{S_1^2}{S_2^2}\frac{1}{F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)}\right].$$

将 $n_1=10,n_2=15,\frac{s_1^2}{s_2^2}=0.853,F_{0.025}(9,14)=3.21,F_{0.025}(14,9)=3.77,$

$F_{0.975}(9,14)=\frac{1}{F_{0.025}(14,9)}=\frac{1}{3.77}$代入得$\frac{\sigma_1^2}{\sigma_2^2}$的置信区间为

$$[0.27,3.22].$$

例 8.1.7 为估计制造某种产品所需的单件平均工作时间(h),现制造5件,得所需工时如下:

$$10.5,\ 11,\ 11.2,\ 12.5,\ 12.8$$

设制造单件产品所需工作时间 $X\sim N(\mu,\sigma^2)$,试求 μ 的 0.95 的单侧置信下限和方差 σ^2 的 0.95 的单侧置信上限.

解 (1) 本题是单总体问题,求 μ 的单侧置信下限而 σ^2 未知,所以利用

$$\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$$

得

$$P\left\{t_{1-\alpha}(n-1)\leqslant\frac{\bar{X}-\mu}{S/\sqrt{n}}\right\}=1-\alpha,$$

即

$$P\{\mu\leqslant\bar{X}+t_\alpha(n-1)S/\sqrt{n}\}=1-\alpha,$$

所以 μ 的单侧置信下限为

$$\bar{X}+t_\alpha(n-1)\frac{S}{\sqrt{n}}.$$

将 $\bar{x}=11.6,s^2=0.995,n=5,t_{0.05}(4)=2.1318$ 代入得 $\bar{x}+t_\alpha(n-1)\frac{s}{\sqrt{n}}=12.55$,即制造单件产品所需工作时间最少为 12.55 h.

(2) 此时,由

$$\frac{n-1}{\sigma^2}S^2\sim\chi^2(n-1),$$

得

$$P\left\{\frac{(n-1)S^2}{\sigma^2} \geqslant \chi_\alpha^2(n-1)\right\}=1-\alpha,$$

即

$$P\left\{\sigma^2 \leqslant \frac{(n-1)S^2}{\chi_\alpha^2(n-1)}\right\}=1-\alpha,$$

所以 σ^2 的单侧置信上限为$\dfrac{(n-1)S^2}{\chi_\alpha^2(n-1)}$. 将 $s^2=0.995, n=5, \chi_{0.05}^2(4)=9.488$ 代入得 σ^2 的 0.95 的单侧置信上限为 0.419.

习　题

1. 某车间生产的一批圆形纽扣的直径 $X\sim N(\mu,0.05)$，现从中随机抽取 6 个，量得平均直径 $\bar{x}=14.95$ mm. 在 0.95 的置信度下求这批纽扣平均直径 μ 的置信区间.

2. 一批袋装大米质量 $X\sim N(\mu,0.62^2)$，现从中随机抽取 10 袋称得质量(单位:kg)为

50.6　50.8　49.5　50.5　50.4　49.7　51.2　49.3　50.6　51.2

求这批袋装大米平均质量 μ 在 0.99 的置信度下的置信区间.

3. 设 $X_1, X_2, \cdots, X_n$ 取自总体 $N(\mu,1)$ 的一个样本，置信度 $1-\alpha=0.95$，问样本容量 n 多大时才能使抽样误差(即置信区间半径)不超过 0.2.

4. 设 0.50, 1.25, 0.80, 2.00 是来自总体 X 的简单随机样本值，已知 $Y=\ln X\sim N(\mu,1)$.

(1) 求 X 的数学期望 $b=E(X)$；

(2) 求 μ 的置信度为 0.95 的置信区间；

(3) 利用上述结果求 b 的置信度为 0.95 的置信区间.

5. 从一批火箭推力装置中抽取 10 个进行试验，测得样本平均燃烧时间为 51.8 s，样本标准差为 1.5 s，设燃烧时间服从正态分布 $N(\mu,\sigma^2)$，求总体均值 μ 的置信度为 0.99 的置信区间.

6. 设某种清漆的干燥时间服从正态分布 $N(\mu,\sigma^2)$，现抽取 9 个样品，测得其干燥时间(单位:h)如下：

6.0　5.7　5.8　6.5　7.0　6.3　5.6　6.1　5.0

在下述条件下求总体均值 μ 的置信度为 0.95 的置信区间：

(1) 若由以往经验知 $\sigma=0.6$；

(2) 若 σ^2 未知.

7. 某人实测他从家到办公室的上班路上所花时间(单位:min)如下：

9.95　10.05　10.20　10.25　9.88　10.10　10.10　10.15　10.12

根据经验，上班路上所花时间服从正态分布 $N(\mu,\sigma^2)$，求：

(1) 总体均值 μ 的置信度为 0.99 的置信区间；

(2) 总体方差 σ^2 的置信度为 0.95 的置信区间.

8. 设某种金属丝长度 $X\sim N(\mu,\sigma^2)$，现从一批这种金属丝中随机抽测 9 根，测得其长度

数据如下(单位:mm)

1532 1297 1647 1356 1435 1483 1574 1517 1463

试估计该批金属丝长度方差 σ^2 的置信度为 0.95 的置信区间.

9. 欲比较甲、乙两种棉花品种的优劣,现假设用它们纺出的棉纱强度 X,Y 分别服从正态分布 $N(\mu_1,2.18^2)$和 $N(\mu_2,1.76^2)$. 试验者从甲、乙这两种棉纱中分别抽取样本容量为 200 和 100 的样本,得样本均值 $\bar{x}=5.32$ 和 $\bar{y}=5.76$,分别在 0.95 和 0.99 的置信度下,求两总体均值差 $\mu_1-\mu_2$ 的置信区间.

10. 如果用 A 种饲料喂牛,牛的增重 $X\sim N(\mu_1,\sigma^2)$;如果用 B 种饲料喂牛,牛的增重 $Y\sim N(\mu_2,\sigma^2)$. 现分别用 A、B 两种饲料各喂牛 10 头,经一个周期后,测得牛的增重(单位:kg)如下:

A 种饲料:20 24 32 31 28 17 25 19 24 30

B 种饲料:27 29 27 38 38 27 35 29 31 36

在 0.95 的置信度下,求 A、B 两种饲料喂牛平均增重的差值 $\mu_1-\mu_2$ 的置信区间.

11. 为了估计磷肥对某种农作物的增产作用,现选取 20 块条件大致相同的土地,其中 10 块不施磷肥,另 10 块施用磷肥,测得亩产量(单位:kg)如下:

不施磷肥:620 270 650 600 630 580 570 600 600 580

施用磷肥:560 590 560 570 580 570 600 550 570 550

设农作物的亩产量服从正态分布.

(1) 若方差相同,求平均亩产量之差 $\mu_1-\mu_2$ 的置信度为 0.99 的置信区间;

(2) 求方差比的置信度为 0.95 的置信区间.

12. 设用两种不同的方法冶炼某种金属材料,分别抽样测试其杂质含量(单位:%),得到如下数据:

原冶炼方法:26.9 22.3 27.2 25.1 22.8 24.2 30.2 25.7 26.1

新冶炼方法:22.6 24.3 23.4 22.5 21.9 20.6 20.6 23.5

假设两种冶炼方法的杂质含量 X,Y 都服从正态分布,且方差 σ_1^2 和 σ_2^2 均未知,求方差比 $\frac{\sigma_1^2}{\sigma_2^2}$ 的置信度为 0.90 的置信区间.

§8.2 假设检验

在前一节中,我们讨论了总体中未知参数的区间估计问题,其前提是总体的分布形式已知. 统计推断还有另外一类重要问题,即在总体的分布未知,或只知其形式而不知道它的参数的情况下,推断其分布或参数. 根据专业的知识或以往的实践经验,人们预先对总体参数存在一定的“认识”,但不确定这种“认识”是否正确. 为此,提出某些关于总体分布或总体参数的假设,然后根据样本对所提出的假设做出判断,是接受还是拒绝. 这就是本章所讨论的另一种统计推断的方

法——假设检验.

假设检验是统计推断的另一项重要内容,它与参数估计类似,都是建立在抽样分布理论基础之上的,但角度不同.参数估计是利用样本信息推断未知的总体参数,而假设检验则是先对总体参数提出一个假设值,然后利用样本信息判断这一假设是否成立.假设检验在许多领域中都有应用,具有十分重要的地位.

本节主要介绍假设检验的基本概念和基本原理,着重解决总体为正态分布时的参数假设检验问题.

8.2.1 假设检验问题的提法

首先通过两个例子,来说明假设检验的一般提法.

例 8.2.1 某车间用一台自动包装机包装葡萄糖,额定标准每袋净重 0.5 kg,设包装机包装的葡萄糖每袋质量 $X \sim N(\mu, 0.015^2)$.某日,从包装机包装的葡萄糖中随机抽取 9 袋,称得净重分别为

0.497 0.506 0.518 0.524 0.498 0.511 0.520 0.515 0.512

问当日该包装机工作是否正常?

例 8.2.2 某手机生产厂家在其宣传广告中,声称其生产的某品牌手机其待机的平均时间至少为 71.5 h.质检部门检查了该厂生产的这种手机 6 部,得到待机时间分别为

69 68 72 70 66 75

假设手机的待机时间 $X \sim N(\mu, \sigma^2)$,由样本数据能否判断该广告有欺骗消费者之嫌?

上面这两个例子所代表的问题很普遍,它们有一个共同点:首先都对总体的参数存在一定的"认识"(称之为"假设").如例 8.2.1 中的"额定标准每袋净重 0.5 kg",例 8.2.2 中"待机的平均时间至少为 71.5 h".其次需要利用样本观测值去判断这种"假设"是否成立.

这里所说的"认识",是一种设想.研究者在检验一种新理论时,通常要先提出一种自己认为正确的看法,即假设.至于它是否真的正确,在建立假设之前并不知道,需要通过样本信息验证该"假设"是否正确.这一过程称之为假设检验.

假设检验的过程如何实现呢?下面具体来分析例 8.2.1 的信息,以便得出假设检验的一般步骤.

在例 8.2.1 中,已知总体方差 $\sigma^2 = 0.015^2$,而 μ 未知,要判断的是该日包装机是否正常工作.根据题意,所谓包装机正常工作,即包装出的葡萄糖每袋净重的均值是否等于 0.5 kg.显然,如果均值等于 0.5 kg,就是正常工作,并用 $\mu = \mu_0 = 0.5$ 的形式提出假设,通常记为 $H_0: \mu = 0.5$,称之为原假设.这里"H"指

的是英文单词 Hypothesis(假设)的首字母.但这个假设不一定是正确的,还需要考虑另外一个方面,即"工作不正常",也就是 X 的均值不等于 0.5 kg.由此,同时提出原假设的对立假设 $H_1: \mu\neq 0.5$,称之为备择假设.这样,要解决例 8.2.1 的问题,首先要提出原假设和备择假设:

$$H_0: \mu=0.5, H_1: \mu\neq 0.5.$$

这里 H_0 和 H_1 有且仅有一个是正确的.如果 $H_0:\mu=0.5$ 成立,这就表明样本均值 $\bar{X}$ 与 μ_0 的偏差不应过大,也就是事件 $|\bar{X}-\mu_0|\leqslant a$ 应是一个大概率事件,在一次抽样试验中应能正常发生,这里 a 表示 $\bar{X}$ 与 μ_0 的偏差.反之,$|\bar{X}-\mu_0|>a$ 在 H_0 为真时是小概率事件,即它应该满足

$$P\{|\bar{X}-\mu_0|>a\}=\alpha, \tag{8.2.1}$$

其中 α 是一个小的正数($0<\alpha<1$),称其为显著性水平.由于此时 $\bar{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right)$,因而(8.2.1)式就变成了

$$P\left\{\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right|>z_{\alpha/2}\right\}=\alpha, \tag{8.2.2}$$

那么 $\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right|>z_{\alpha/2}$ 就是小概率事件,而小概率事件在一次抽样试验中几乎不发生.在本题中,如果取 $\alpha=0.05$,则 $z_{0.025}=1.96$.又知 $n=9$,$\sigma=0.015$,$\mu_0=0.5$,应用样本数据还可以算得 $\bar{x}=0.511$.将这些数据代入式(8.2.2)中,有

$$\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}=2.2>1.96=z_{0.025}.$$

这说明小概率事件发生了,也就是说包装机工作不正常了.因此在这种情况下,应拒绝 H_0 接受 H_1.

8.2.2　双侧检验与单侧检验

在上面的例子中备择假设 H_1 表示只要 $\mu>0.5$ 或 $\mu<0.5$ 中有一个成立,就可以拒绝 H_0,因而形如

$$H_0: \theta=\theta_0,\ H_1: \theta\neq\theta_0 \tag{8.2.3}$$

的假设检验称为双侧检验,但在许多情况下有如下形式的单侧检验:

形如

$$H_0: \theta=\theta_0(\text{或 } H_0: \theta\leqslant\theta_0),\ H_1: \theta>\theta_0 \tag{8.2.4}$$

的假设检验称为右侧检验;

形如

$$H_0: \theta = \theta_0 (\text{或 } H_0: \theta \geqslant \theta_0), \ H_1: \theta < \theta_0 \tag{8.2.5}$$

的假设检验称为左侧检验.

上面的分析过程包含了反证法的思想,但是与真正的反证法是有区别的,只能说是一种基于小概率原则的反证法.

“小概率事件原则”是说概率很小的事件在一次试验中不会发生.

8.2.3 两类错误

假设检验是根据小概率事件原则来判断的,但是大家都知道小概率事件也是可能发生的,也就是说我们可能犯错误.具体来讲,所犯错误主要有两类:

(1) 第一类错误(弃真) 当 H_0 实际上为真时,检验结果却是拒绝 H_0,称之为弃真.犯第一类错误的概率就是显著性水平 α,即

$$P\{\text{拒绝 } H_0 \mid H_0 \text{ 为真}\} = \alpha.$$

(2) 第一类错误(采伪) 当 H_0 实际上为不真时,检验结果却是接受 H_0,称之为采伪.犯第二类错误的概率通常记为 β,即

$$P\{\text{接受 } H_0 \mid H_0 \text{ 不真}\} = \beta.$$

我们当然不希望犯错误,或者犯错误的概率越小越好,但是犯这两类错误的概率不可能同时减小,通常的做法是在给定犯第一类错误的概率 α 的条件下,尽量减小犯第二类错误的概率 β.

8.2.4 正态总体假设检验的基本步骤

对假设检验我们仍然主要研究正态总体参数的假设检验问题.所以由前面的讨论可以得出假设检验的基本步骤(双侧检验):

(1) 根据实际问题提出原假设 H_0 和备择假设接受 H_1;

(2) 从正态总体的六个已知的抽样分布(单总体、双总体各 3 个)中选一个只含有待检验参数 $\theta=\theta_0$ 而不包含其他未知参数的抽样分布,记为 Y,并称之为检验统计量.

(3) 由于 Y 一定是四大分布($N(0,1)$,χ^2,t,F)中的一个,所以它的分位点 $Y_{\alpha/2}$,$Y_{1-\alpha/2}$ 能从表中查到,且必然满足

$$P\{Y < Y_{1-\alpha/2} \text{ 或 } Y > Y_{\alpha/2}\} = \alpha, \tag{8.2.6}$$

那么 $Y<Y_{1-\alpha/2}$ 或 $Y>Y_{\alpha/2}$ 就是小概率事件,今后称之为拒绝域.

(4) 由于 Y 中没有未知参数 θ,所以将已知数据代入即可把 Y 求出.

(5) 如果 $Y<Y_{1-\alpha/2}$ 或 $Y>Y_{\alpha/2}$ 有一个成立,就拒绝 H_0,否则接受 H_0.

如果是左侧检验或右侧检验,那么只需要将 $P\{Y<Y_{1-\alpha/2} \text{ 或 } Y>Y_{\alpha/2}\}=\alpha$

换成

$$P\{Y < Y_{1-\alpha}\} = \alpha \text{或} P\{Y > Y_{\alpha}\} = \alpha. \tag{8.2.7}$$

此时拒绝域分别为 $Y<Y_{1-\alpha}$或 $Y>Y_{\alpha}$.

下面针对正态总体参数的假设检验进行具体讨论.

例 8.2.3　根据过去大量资料显示，某工厂产品的使用寿命(单位：h) $X\sim N(1\ 020,100^2)$. 现从最近生产的一批产品中随机抽取 16 件，测得样本平均寿命为 1 080 小时. 试在 0.05 的显著性水平下，判断这批产品的使用寿命是否有显著提高?

解　根据题意提出假设为

$$H_0: \mu\leqslant 1\ 020, H_1: \mu>1\ 020.$$

这是单正态总体对 μ 的右侧检验问题，且 σ^2 已知，所以应使用统计量

$$Z=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1),$$

得

$$P\{Z>z_{\alpha}\}=\alpha,$$

故拒绝域为 $Z>z_{\alpha}$.

将 $n=16$, $\bar{x}=1\ 080$, $\mu_0=1\ 020$, $\sigma=100$，代入可得

$$z=\frac{1\ 080-1\ 020}{100/\sqrt{16}}=2.4>1.645=z_{0.05},$$

因此应该拒绝 H_0 接受 H_1. 也就是说在 0.05 的显著性水平下，这批产品的使用寿命有显著提高.

例 8.2.4　在 0.05 的显著性水平下，解答例 8.2.2 问题.

解　要检验“该广告是否有欺骗消费者之嫌”，则由题设，提出假设：

$$H_0: \mu=71.5,\quad H_1: \mu<71.5.$$

由于 σ^2 未知，故取检验统计量

$$t=\frac{\bar{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1).$$

得

$$P\{t<t_{1-\alpha}(n-1)\}=\alpha,$$

拒绝域为 $t<t_{1-\alpha}(n-1)$.

将 $n=6$, $\bar{x}=\frac{\sum_{i=1}^{6}x_i}{6}=70$, $s=\sqrt{\frac{\sum_{i=1}^{6}(\bar{x}-x_i)^2}{5}}=3.16$, $\mu_0=71.5$ 代入得

$$t=\frac{70-71.5}{3.16/\sqrt{6}}=-1.16>-t_{0.05}(5)=-2.015,$$

故接受 H_0，即在 0.05 的显著性水平下，不能认为该广告有欺骗消费者之嫌.

例 8.2.5 设某车间生产的滚珠的直径 $X\sim N(\mu,\sigma^2)$，现从某日生产的滚珠中抽取 9 个，测得样本方差为 $s^2=0.25^2$，在显著性水平 0.05 下可否认为总体方差的 $\sigma^2=0.36^2$.

解 由题设，提出假设：

$$H_0:\sigma^2=0.36^2, H_1:\sigma^2\neq 0.36^2.$$

使用统计量

$$\chi^2=\frac{n-1}{\sigma^2}S^2\sim\chi^2(n-1),$$

得

$$P\{\chi^2<\chi^2_{1-\alpha/2}(n-1) \text{ 或 } \chi^2>\chi^2_{\alpha/2}(n-1)\}=\alpha,$$

拒绝域为 $\chi^2<\chi^2_{1-\alpha/2}(n-1)$ 或 $\chi^2>\chi^2_{\alpha/2}(n-1)$.

将 $n=9$，$s^2=0.25^2$ 代入，得 $\chi^2\approx 3.858$，而 $\chi^2_{0.025}(8)=17.535$，$\chi^2_{0.975}(8)=2.180$，所以没有落在拒绝域里面，应接受原假设，认为总体方差的 $\sigma^2=0.36^2$.

例 8.2.6 某公司有甲、乙两个分厂，公司首脑认为甲分厂工人的生产效率要高于乙分厂工人的生产效率. 现从甲分厂随机抽取 10 名工人生产某种工件，测得平均所用时间 $\bar{x}=20$ min，样本标准差 $s_1=4$ min；从乙分厂随机抽取 20 名工人完成相同的工作，测得平均所用时间 $\bar{y}=24$ min，样本标准差 $s_2=5$ min. 假设工人生产这种工件所用时间均服从正态分布，且总体方差相等. 在 0.005 的显著性水平下，是否可以认为公司首脑的判断是可靠的？

解 提出假设如下：

$$H_0:\mu_1=\mu_2, H_1:\mu_1<\mu_2.$$

使用检验统计量

$$t=\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_\omega\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t(n_1+n_2-2),$$

得

$$P\{t<t_{1-\alpha}(n_1+n_2-2)\}=\alpha,$$

拒绝域为 $t<t_{1-\alpha}(n_1+n_2-2)$.

由将 $n_1=10, \bar{x}=20, n_2=20, \bar{y}=24, s_1=4, s_2=5$ 代入得

$$t=\frac{(20-24)-0}{\sqrt{\frac{22.11}{10}+\frac{22.11}{20}}}=-2.196>-t_{0.01}(28)=-2.47.$$

故接受 H_0，即在 0.01 的显著性水平下，可以认为公司首脑的判断是不可靠的.

例 8.2.7 一家房地产开发公司准备购进一批灯泡，有两个供货商可供选择. 这两家供货商的灯泡平均使用寿命差别不大，价格也很相近，考虑的主要因素就是灯泡使用寿命的方差大小. 如果方差相同，就选择距离较近一家供货商进货. 为此，公司管理人员对两家供货商提供的样品进行了检测，得知供货商甲提供的 20 个样品的方差 $S_1^2=3\ 675.46$，供货商乙提供的 15 个样品的方差 $S_2^2=2\ 431.43$. 试在 0.05 的显著性水平下，检验两家供货商的灯泡的使用寿命的方差是否有显著性差异？

解 提出假设如下：

$$H_0:\frac{\sigma_1^2}{\sigma_2^2}=1,H_1:\frac{\sigma_1^2}{\sigma_2^2}\neq 1\quad(\text{或 } H_0:\sigma_1^2=\sigma_2^2,H_1:\sigma_1^2\neq\sigma_2^2).$$

使用检验统计量

$$F=\frac{S_1^2}{S_2^2}\sim F(n_1-1,n_2-1).$$

得

$$P\{F<F_{1-\alpha/2}(n_1-1,n_2-1)\text{ 或 }F>F_{\alpha/2}(n_1-1,n_2-1)\}=\alpha,$$

拒绝域为 $F<F_{1-\alpha/2}(n_1-1,n_2-1)$ 或 $F>F_{\alpha/2}(n_1-1,n_2-1)$.

将 $s_1^2=3\ 675.46$，$s_2^2=2\ 431.43$ 代入得

$$F=\frac{3\ 675.46}{2\ 431.43}=1.51.$$

而 $F_{0.025}(14,19)=2.62$，$F_{0.025}(19,14)=2.84$，故 $F_{0.975}(19,14)=\frac{1}{F_{0.025}(14,19)}=\frac{1}{2.62}=0.38$. 由于 $0.38<1.51<2.84$，故接受 H_0 即在 0.05 的显著性水平下，可以认为两家供货商的灯泡的使用寿命的方差没有显著性差异.

习 题

1. 按照过去的铸造法，某厂所造的零件强度的平均值是 52.1 g/mm^2，标准差为 1.6 g/mm^2. 为降低成本，该厂改变了铸造方法，从按新方法生产的产品中抽取了 9 个样品，测得其强度平均为 52.9 g/mm^2. 假设零件的强度服从正态分布，试在 0.05 的显著性水平下，判断新的铸造方法是否提升了零件的强度，即检验总体均值是否变大？

2. 已知某炼铁厂生产的铁水的含量碳量服从正态分布 $N(4.55,0.11^2)$. 现测试 9 炉铁

水，其平均含碳量为 4.484. 如果方差没有变化，可否认为现在生产的铁水的含碳量仍为 4.55，取显著性水平 $\alpha=0.05$？

3. 一种汽车配件的长度要求为 12 cm，高于或低于该标准都被认为是不合格的. 现对一个配件提供商提供的 10 个样品进行了检测，测得样本均值 $\bar{x}=11.89$ cm，样本标准差 $s=0.493\ 2$ cm. 假定这种汽车配件的长度服从正态分布，在 0.05 的显著性水平下，检验该供货商提供的配件是否符合要求？

4. 测定某溶液中的水分，得到 10 个测定值，经计算 $\bar{x}=5.2\%$，$s^2=0.037^2$，设溶液中的水分含量 $X\sim N(\mu,\sigma^2)$，σ^2 和 μ 未知，在 0.05 的显著性水平下，该溶液中水分含量均值 μ 是否超过 5%？

5. 从某厂生产的电子元件中随机抽取 25 个进行寿命测试，根据测得数据算得(单位：h) $\bar{x}=100$，$\sum\limits_{i=1}^{25}x_i^2=4.9\times10^5$. 已知这种电子元件的使用寿命服从 $N(\mu,\sigma^2)$，且出厂标准为使用寿命必须达到 90 h 以上. 在 0.05 的显著性水平下，检验该厂生产的电子元件是否符合标准？

6. 随机地从一批外径为 1 cm 的钢珠中抽取 10 只，测试其屈服强度(单位：kg)得平均值 $\bar{x}=2\ 200$，$s=220$，已知钢珠的屈服强度 $X\sim N(\mu,\sigma^2)$.

(1) 求总体均值 μ 置信度为 0.95 的置信区间；

(2) 在 0.05 的显著性水平下，检验总体均值 μ 是否等于 2 000？

(3) 若设 $X\sim N(\mu,200^2)$，在 0.05 的显著性水平下，检验 X 的方差 σ^2 是否有显著提高？

7. 有甲、乙两个品种的作物，分别各用 10 块地试种，根据收集到的数据得到平均产量结果分别为 $\bar{x}=30.97$ 和 $\bar{y}=21.97$. 已知这两种作物的产量分别服从正态分布 $N(\mu_1,27)$ 和 $N(\mu_2,12)$，问在 0.01 的显著性水平下，这两种品种的平均产量是否有显著性差异？

8. 在甲、乙两个居民区分别抽取 8 户和 10 户调查每月煤气用量(m^3)，计算得样本均值分别为 $\bar{x}_1=7.56$，$\bar{x}_2=6.02$. 根据以往经验，两区居民煤气用量近似服从正态分布，相互独立，且两总体标准差 $\sigma_1=\sigma_2=1.1$. 在 0.05 的显著性水平下，判断甲区居民煤气用量是否高于乙区？

9. 甲、乙两台机床同时加工某种的零件，已知两台机床加工的零件的直径均服从正态分布，并且方差相同. 现从甲加工的零件中随机抽取 8 件，测得其平均直径为 19.925 cm，样本方差为 0.216 4 cm^2. 从乙加工的零件中抽取 7 件，测得其平均直径为 20.643 cm，样本方差为 0.272 9 cm^2. 在 0.05 的显著性水平下，是否能够显示甲机床加工的零件的直径要小于乙机床加工的零件直径？

10. 随机地挑选 20 位失眠者分别服用甲、乙两种安眠药，记录下他们睡眠的延长时间(单位：h)，分别得到数据 $x_1,\cdots,x_{10}$ 和 $y_1,\cdots,y_{10}$，并由此算得 $\bar{x}=4$，$s_1^2=0.001$，$\bar{y}=4.04$，$s_2^2=0.004$. 设甲、乙两种安眠药的延长时间均服从正态分布，且方差相等. 在 0.05 的显著性水平下，判断两种安眠药的疗效是否相同？

11. 有两台机器生产的金属部件，分别在两台机器所生产的部件中各取一容量 $n_1=60$，$n_2=40$ 的样本，测得部件质量(单位：kg)的样本方差分别为 $s_1^2=15.46$，$s_2^2=9.66$. 设两样本相

互独立，两总体分别服从 $N(\mu_1,\sigma_1^2)$，$N(\mu_2,\sigma_2^2)$分布，其中 $\mu_1,\sigma_1^2,\mu_2,\sigma_2^2$ 均未知. 试在 0.05 的显著性水平下，检验如下假设：

$$H_0: \sigma_1^2 = \sigma_2^2, H_1: \sigma_1^2 > \sigma_2^2.$$

12. 甲、乙两个铸造厂生产同一种铸件，假定两厂的铸件重量都服从正态分布，现从两厂的铸件中各抽取若干个，分别测得重量如下(单位:kg)：

甲厂:93.3　92.1　94.7　90.1　95.6　90.0　94.7

乙厂:95.6　94.9　96.2　95.8　95.1　96.3

取显著性水平 $\alpha=0.05$，检验甲厂铸件重量的方差与乙厂的铸件重量的方差是否存在显著性差异？

附　　表

附表 1　几种常用的概率分布

分布	参数	分布律或概率密度	数学期望	方差
0-1 分布	$0<p<1$	$P\{X=k\}=p^k(1-p)^{1-k}$ $k=0,1$	p	$p(1-p)$
二项分布	$n\geqslant 1$ $0<p<1$	$P\{X=k\}=\binom{n}{k}p^k(1-p)^{n-k}$ $k=0,1,\cdots,n$	np	$np(1-p)$
负二项分布	$r\geqslant 1$ $0<p<1$	$P\{X=k\}=\binom{k-1}{r-1}p^r(1-p)^{k-r}$ $k=r,r+1,\cdots$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$
几何分布	$0<p<1$	$P\{X=k\}=p(1-p)^{k-1}$ $k=1,2,\cdots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
超几何分布	N,M,n $(n\leqslant M)$	$P\{X=k\}=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$ $k=0,1,\cdots,n$	$\frac{nM}{N}$	$\frac{nM}{N}\left(1-\frac{M}{N}\right)\left(\frac{N-n}{N-1}\right)$

续表

分布	参数	分布律或概率密度	数学期望	方差
泊松分布	$\lambda>0$	$P\{X=k\}=\frac{\lambda^k e^{-\lambda}}{k!}$ $k=0,1,\cdots$	λ	λ
均匀分布	$a<b$	$f(x)=\begin{cases}\frac{1}{b-a}, & a<x<b\\ 0, & \text{其他}\end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
正态分布	μ $\sigma>0$	$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2
Γ 分布	$\alpha>0$ $\beta>0$	$f(x)=\begin{cases}\frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}, & x>0\\ 0, & \text{其他}\end{cases}$	$\alpha\beta$	$\alpha\beta^2$
指数分数	$\theta>0$	$f(x)=\begin{cases}\frac{1}{\theta}e^{-x/\theta}, & x>0\\ 0, & \text{其他}\end{cases}$	θ	θ^2
χ^2 分布	$n\geqslant1$	$f(x)=\begin{cases}\frac{1}{2^{n/2}\Gamma(n/2)}x^{n/2-1}e^{-x/2}, & x>0\\ 0, & \text{其他}\end{cases}$	n	$2n$
威布尔分布	$\eta>0$ $\beta>0$	$f(x)=\begin{cases}\frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1}e^{-(\frac{x}{\eta})^\beta}, & x>0\\ 0, & \text{其他}\end{cases}$	$\eta\Gamma\left(\frac{1}{\beta}+1\right)$	$\eta^2\left\{\Gamma\left(\frac{2}{\beta}+1\right)-\left[\Gamma\left(\frac{1}{\beta}+1\right)\right]^2\right\}$

续表

分布	参数	分布律或概率密度	数学期望	方差
瑞利分布	$\sigma>0$	$f(x)=\begin{cases}\dfrac{x}{\sigma^2}\mathrm{e}^{-x^2/(2\sigma^2)}, & x>0\\ 0, & 其他\end{cases}$	$\sqrt{\dfrac{\pi}{2}}\sigma$	$\dfrac{4-\pi}{2}\sigma^2$
β分布	$\alpha>0$ $\beta>0$	$f(x)=\begin{cases}\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}, & 0<x<1\\ 0, & 其他\end{cases}$	$\dfrac{\alpha}{\alpha+\beta}$	$\dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
对数正态分布	μ $\sigma>0$	$f(x)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\sigma x}\mathrm{e}^{\frac{-(\ln x-\mu)^2}{2\sigma^2}}, & x>0\\ 0, & 其他\end{cases}$	$\mathrm{e}^{\mu+\frac{\sigma^2}{2}}$	$\mathrm{e}^{2\mu+\sigma^2}(\mathrm{e}^{\sigma^2}-1)$
柯西分布	α $\lambda>0$	$f(x)=\dfrac{1}{\pi}\dfrac{1}{\lambda^2+(x-a)^2}$	不存在	不存在
t分布	$n\geqslant1$	$f(x)=\dfrac{\Gamma\left(\dfrac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma(n/2)}\left(1+\dfrac{x^2}{n}\right)^{-(n+1)/2}$	0	$\dfrac{n}{n-2}, n>2$
F分布	n_1, n_2	$f(x)=\begin{cases}\dfrac{\Gamma[(n_1+n_2)/2]}{\Gamma(n_1/2)\Gamma(n_2/2)}\left(\dfrac{n_1}{n_2}\right)\left(\dfrac{n_1}{n_2}x\right)^{(n_1+n_2)/2}\\ \cdot\left(1+\dfrac{n_1}{n_2}x\right)^{-(n_1+n_2)/2}, & x>0\\ 0, & 其他\end{cases}$	$\dfrac{n_2}{n_2-2}$ $n_2>2$	$\dfrac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)}$ $n_2>4$

附表 2 泊松分布表

设 $X \sim P(\lambda)$，给出概率

$$P\{X \geqslant x\} = \sum_{r=x}^{\infty} \frac{e^{-\lambda}\lambda^r}{r!}$$

x	$\lambda=0.2$	$\lambda=0.3$	$\lambda=0.4$	$\lambda=0.5$	$\lambda=0.6$
0	1.000 000 0	1.000 000 0	1.000 000 0	1.000 000 0	1.000 000 0
1	0.181 269 2	0.259 181 8	0.329 680 0	0.323 469	0.451 188
2	0.017 523 1	0.036 936 3	0.061 551 9	0.090 204	0.121 901
3	0.001 148 5	0.003 599 5	0.007 926 3	0.014 388	0.023 115
4	0.000 056 8	0.000 265 8	0.000 776 3	0.001 752	0.003 358
5	0.000 002 3	0.000 015 8	0.000 061 2	0.000 172	0.000 394
6	0.000 000 1	0.000 000 8	0.000 004 0	0.000 014	0.000 039
7			0.000 000 2	0.000 001	0.000 003

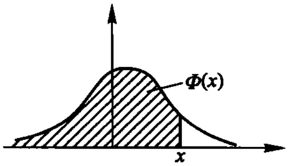
x	$\lambda=0.7$	$\lambda=0.8$	$\lambda=0.9$	$\lambda=1.0$	$\lambda=1.2$
0	1.000 000 0	1.000 000 0	1.000 000 0	1.000 000 0	1.000 000 0
1	0.503 415	0.550 671	0.593 430	0.632 121	0.698 806
2	0.155 805	0.191 208	0.227 518	0.264 241	0.337 373
3	0.034 142	0.047 423	0.062 857	0.080 301	0.120 513
4	0.005 753	0.009 080	0.013 459	0.018 988	0.033 769
5	0.000 786	0.001 411	0.002 344	0.003 660	0.007 746
6	0.000 090	0.000 184	0.000 343	0.000 594	0.001 500
7	0.000 009	0.000 021	0.000 043	0.000 083	0.000 251
8	0.000 001	0.000 002	0.000 005	0.000 010	0.000 037
9				0.000 001	0.000 005
10					0.000 001

x	$\lambda=1.4$	$\lambda=1.6$	$\lambda=1.8$		
0	1.000 000	1.000 000	1.000 000		
1	0.753 403	0.798 103	0.834 701		
2	0.408 167	0.475 069	0.537 163		
3	0.166 502	0.216 642	0.269 379		
4	0.053 725	0.078 813	0.108 708		
5	0.014 253	0.023 682	0.036 407		
6	0.003 201	0.006 040	0.010 378		
7	0.000 622	0.001 336	0.002 569		
8	0.000 107	0.000 260	0.000 562		
9	0.000 016	0.000 045	0.000 110		
10	0.000 002	0.000 007	0.000 019		
11		0.000 001	0.000 003		

续表

x	$\lambda=2.5$	$\lambda=3.0$	$\lambda=3.5$	$\lambda=4.0$	$\lambda=4.5$	$\lambda=5.0$
0	1.000 000	1.000 000	1.000 000	1.000 000	1.000 000	1.000 000
1	0.917 915	0.950 213	0.969 803	0.981 684	0.988 891	0.993 262
2	0.712 703	0.800 852	0.864 112	0.908 422	0.938 901	0.959 572
3	0.456 187	0.576 810	0.679 153	0.761 897	0.826 422	0.875 348
4	0.242 424	0.352 768	0.463 367	0.566 530	0.657 704	0.734 974
5	0.108 822	0.184 737	0.274 555	0.371 163	0.467 896	0.559 507
6	0.042 021	0.083 918	0.142 386	0.214 870	0.297 070	0.384 039
7	0.014 187	0.033 509	0.065 288	0.110 674	0.168 949	0.237 817
8	0.004 247	0.011 905	0.026 739	0.051 134	0.086 586	0.133 372
9	0.001 140	0.003 803	0.009 874	0.021 368	0.040 257	0.068 094
10	0.000 277	0.001 102	0.003 315	0.008 132	0.017 093	0.031 828
11	0.000 062	0.000 292	0.001 019	0.002 840	0.006 669	0.013 695
12	0.000 013	0.000 071	0.000 289	0.000 915	0.002 404	0.005 453
13	0.000 002	0.000 016	0.000 076	0.000 274	0.000 805	0.002 019
14		0.000 003	0.000 019	0.000 076	0.000 252	0.000 698
15		0.000 001	0.000 004	0.000 020	0.000 074	0.000 226
16			0.000 001	0.000 005	0.000 020	0.000 069
17				0.000 001	0.000 005	0.000 020
18					0.000 001	0.000 005
19						0.000 001

附表 3 标准正态分布表



$$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-u^2/2}\,\mathrm{d}u$$

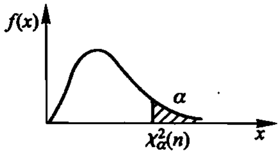
x	0	1	2	3	4	5	6	7	8	9
0.0	0.500 0	0.504 0	0.508 0	0.512 0	0.516 0	0.519 9	0.523 9	0.527 9	0.531 9	0.535 9
0.1	0.539 8	0.543 8	0.547 8	0.551 7	0.555 7	0.559 6	0.563 6	0.567 5	0.571 4	0.575 3
0.2	0.579 3	0.583 2	0.587 1	0.591 0	0.594 8	0.598 7	0.602 6	0.606 4	0.610 3	0.614 1
0.3	0.617 9	0.621 7	0.625 5	0.629 3	0.633 1	0.636 8	0.640 6	0.644 3	0.648 0	0.651 7
0.4	0.655 4	0.659 1	0.662 8	0.666 4	0.670 0	0.673 6	0.677 2	0.680 8	0.684 4	0.687 9
0.5	0.691 5	0.695 0	0.698 5	0.701 9	0.705 4	0.708 8	0.712 3	0.715 7	0.719 0	0.722 4
0.6	0.725 7	0.729 1	0.732 4	0.735 7	0.738 9	0.742 2	0.745 4	0.748 6	0.751 7	0.754 9
0.7	0.758 0	0.761 1	0.764 2	0.767 3	0.770 3	0.773 4	0.776 4	0.779 4	0.782 3	0.785 2
0.8	0.788 1	0.791 0	0.793 9	0.796 7	0.799 5	0.802 3	0.805 1	0.807 8	0.810 6	0.813 3
0.9	0.815 9	0.818 6	0.821 2	0.823 8	0.826 4	0.828 9	0.831 5	0.834 0	0.836 5	0.838 9
1.0	0.841 3	0.843 8	0.846 1	0.848 5	0.850 8	0.853 1	0.855 4	0.857 7	0.859 9	0.862 1
1.1	0.864 3	0.866 5	0.868 6	0.870 8	0.872 9	0.874 9	0.877 0	0.879 0	0.881 0	0.883 0
1.2	0.884 9	0.886 9	0.888 8	0.890 7	0.892 5	0.894 4	0.896 2	0.898 0	0.899 7	0.901 5
1.3	0.903 2	0.904 9	0.906 6	0.908 2	0.909 9	0.911 5	0.913 1	0.914 7	0.916 2	0.917 7
1.4	0.919 2	0.920 7	0.922 2	0.923 6	0.925 1	0.926 5	0.927 8	0.929 2	0.930 6	0.931 9
1.5	0.933 2	0.934 5	0.935 7	0.937 0	0.938 2	0.939 4	0.940 6	0.941 8	0.943 0	0.944 1
1.6	0.945 2	0.946 3	0.947 4	0.948 4	0.949 5	0.950 5	0.951 5	0.952 5	0.953 5	0.954 5
1.7	0.955 4	0.956 4	0.957 3	0.958 2	0.959 1	0.959 9	0.960 8	0.961 6	0.962 5	0.963 3
1.8	0.964 1	0.964 8	0.965 6	0.966 4	0.967 1	0.967 8	0.968 6	0.969 3	0.970 0	0.970 6
1.9	0.971 3	0.971 9	0.972 6	0.973 2	0.973 8	0.974 4	0.975 0	0.975 6	0.976 2	0.976 7
2.0	0.977 2	0.977 8	0.978 3	0.978 8	0.979 3	0.979 8	0.980 3	0.980 8	0.981 2	0.981 7
2.1	0.982 1	0.982 6	0.983 0	0.983 4	0.983 8	0.984 2	0.984 6	0.985 0	0.985 4	0.985 7
2.2	0.986 1	0.986 4	0.986 8	0.987 1	0.987 4	0.987 8	0.988 1	0.988 4	0.988 7	0.989 0
2.3	0.989 3	0.989 6	0.989 8	0.990 1	0.990 4	0.990 6	0.990 9	0.991 1	0.991 3	0.991 6
2.4	0.991 8	0.992 0	0.992 2	0.992 5	0.992 7	0.992 9	0.993 1	0.993 2	0.993 4	0.993 6

续表

x	0	1	2	3	4	5	6	7	8	9
2.5	0.993 8	0.994 0	0.994 1	0.994 3	0.994 5	0.994 6	0.994 8	0.994 9	0.995 1	0.995 2
2.6	0.995 3	0.995 5	0.995 6	0.995 7	0.995 9	0.996 0	0.996 1	0.996 2	0.996 3	0.996 4
2.7	0.996 5	0.996 6	0.996 7	0.996 8	0.996 9	0.997 0	0.997 1	0.997 2	0.997 3	0.997 4
2.8	0.997 4	0.997 5	0.997 6	0.997 7	0.997 7	0.997 8	0.997 9	0.997 9	0.998 0	0.998 1
2.9	0.998 1	0.998 2	0.998 2	0.998 3	0.998 4	0.998 4	0.998 5	0.998 5	0.998 6	0.998 6
3.0	0.998 7	0.999 0	0.999 3	0.999 5	0.999 7	0.969 8	0.999 8	0.999 9	0.999 9	1.000 0

注：表中末行系函数值 $\Phi(3.0),\Phi(3.1),\cdots,\Phi(3.9)$

附表4 χ^2 分布表



$P\{\chi^2(n)>\chi^2_\alpha(n)\}=\alpha$

n	α=0.995	0.99	0.975	0.95	0.90	0.75
1	—	—	0.001	0.004	0.016	0.102
2	0.010	0.020	0.051	0.103	0.211	0.575
3	0.072	0.115	0.216	0.352	0.584	1.213
4	0.207	0.297	0.484	0.711	1.064	1.923
5	0.412	0.554	0.831	1.145	1.610	2.675
6	0.676	0.872	1.237	1.635	2.204	3.455
7	0.989	1.239	1.690	2.167	2.833	4.255
8	1.344	1.646	2.180	2.733	3.490	5.071
9	1.735	2.088	2.700	3.325	4.168	5.899
10	2.156	2.558	3.247	3.940	4.865	6.737
11	2.603	3.053	3.816	4.575	5.578	7.584
12	3.074	3.571	4.404	5.226	6.304	8.438
13	3.565	4.107	5.009	5.892	7.042	9.299
14	4.075	2.660	5.629	6.571	7.790	10.165
15	4.601	5.229	6.262	7.261	8.547	11.037
16	5.142	5.812	6.908	7.962	9.312	11.912
17	5.697	6.408	7.564	8.672	10.085	12.792
18	6.265	7.015	8.231	9.390	10.865	13.675
19	6.844	7.633	8.907	10.117	11.651	14.562
20	7.434	8.260	9.591	10.851	12.443	15.452
21	8.034	8.897	10.283	11.591	13.240	16.344
22	8.643	9.542	10.982	12.338	14.042	17.240
23	9.260	10.196	11.689	13.091	14.848	18.137
24	9.886	10.856	12.401	13.848	15.659	19.037
25	10.520	11.524	13.120	14.611	16.473	19.939

续表

n	α=0.995	0.99	0.975	0.95	0.90	0.75
26	11.160	12.198	13.844	15.379	17.292	20.843
27	11.808	12.879	14.573	16.151	18.114	21.749
28	12.461	13.565	15.308	16.928	18.939	22.657
29	13.121	14.257	16.047	17.708	19.768	23.567
30	13.787	14.954	16.791	18.493	20.599	24.478
31	14.458	15.655	17.539	19.281	21.434	25.390
32	15.134	16.362	18.291	20.072	22.271	26.304
33	15.815	17.074	19.047	20.807	23.110	27.219
34	16.501	17.789	19.806	21.664	23.952	28.136
35	17.192	18.509	20.569	22.465	24.797	29.054
36	17.887	19.233	21.336	23.269	25.613	29.973
37	18.586	19.960	22.106	24.075	26.492	30.893
38	19.289	20.691	22.878	24.884	27.343	31.815
39	19.996	21.426	23.654	25.695	28.196	32.737
40	20.707	22.164	24.433	26.509	29.051	33.660
41	21.421	22.906	25.215	27.326	29.907	34.585
42	22.138	23.650	25.999	28.144	30.765	35.510
43	22.859	24.398	26.785	28.965	31.625	36.430
44	23.584	25.143	27.575	29.787	32.487	37.363
45	24.311	25.901	28.366	30.612	33.350	38.291

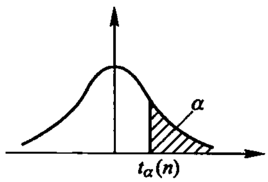
n	α=0.25	0.10	0.05	0.025	0.01	0.005
1	1.323	2.706	3.841	5.024	6.635	7.879
2	2.773	4.605	5.991	7.378	9.210	10.597
3	4.108	6.251	7.815	9.348	11.345	12.838
4	5.385	7.779	9.488	11.143	13.277	14.860
5	6.626	9.236	11.071	12.833	15.086	16.750
6	7.841	10.645	12.592	14.449	16.812	18.548
7	9.037	12.017	14.067	16.013	18.475	20.278
8	10.219	13.362	15.507	17.535	20.090	21.955
9	11.389	14.684	16.919	19.023	21.666	23.589
10	12.549	15.987	18.307	20.483	23.209	25.188

续表

n	$\alpha=0.25$	0.10	0.05	0.025	0.01	0.005
11	13.701	17.275	19.675	21.920	24.725	26.757
12	14.845	18.549	21.026	23.337	26.217	28.299
13	15.984	19.812	22.362	24.736	27.688	29.819
14	17.117	21.064	23.685	26.119	29.141	31.319
15	18.245	22.307	24.996	27.488	30.578	32.801
16	19.369	23.542	26.296	28.845	32.000	34.267
17	20.489	24.769	27.587	30.191	33.409	35.718
18	21.605	25.989	28.869	31.526	34.805	37.156
19	22.718	27.204	30.144	32.852	36.191	38.582
20	23.828	28.412	31.410	34.170	37.566	39.997
21	24.935	29.615	32.671	35.479	38.932	41.401
22	26.039	30.813	33.924	36.781	40.289	42.796
23	27.141	32.007	35.172	38.076	41.638	44.181
24	28.241	33.196	36.415	39.364	42.980	45.559
25	29.339	34.382	37.652	40.646	44.314	46.928
26	30.435	35.563	38.885	41.923	45.642	48.290
27	31.528	36.741	40.113	43.194	46.963	49.645
28	32.620	37.916	41.337	44.461	48.278	50.993
29	33.711	39.087	42.557	45.722	49.588	52.336
30	34.800	40.256	43.773	46.979	50.892	53.672
31	35.887	41.422	44.985	48.232	52.191	55.003
32	36.973	42.585	46.194	49.480	53.486	56.328
33	38.053	43.745	47.400	50.725	54.776	57.648
34	39.141	44.903	48.602	51.966	56.061	58.964
35	40.223	46.059	49.802	53.203	57.342	60.275
36	41.304	47.212	50.998	54.437	58.619	61.581
37	42.383	48.363	52.192	55.668	59.892	62.883
38	43.462	49.513	53.384	56.896	61.162	64.181
39	44.539	50.660	54.572	58.120	62.428	65.476
40	45.616	51.805	55.758	59.342	63.691	66.766
41	46.692	52.949	53.942	60.561	64.950	68.053
42	47.766	54.090	58.124	61.777	66.206	69.336
43	48.840	55.230	59.304	62.990	67.459	70.606
44	49.913	56.369	60.481	64.201	68.710	71.893
45	50.985	57.505	61.656	65.410	69.957	73.166

附表5 t 分布表

$P\{t(n)>t_\alpha(n)\}=\alpha$

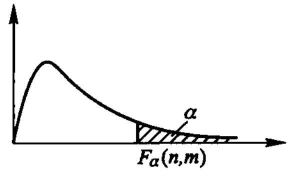


n	$\alpha=0.25$	0.10	0.05	0.025	0.01	0.005
1	1.000 0	3.077 7	6.313 8	12.706 2	31.820 7	63.657 4
2	0.816 5	1.885 6	2.920 0	4.302 7	6.964 6	9.924 8
3	0.764 9	1.637 7	2.353 4	3.182 4	4.540 7	5.840 9
4	0.740 7	1.533 2	2.131 8	2.776 4	3.746 9	4.604 1
5	0.726 7	1.475 9	2.015 0	2.570 6	3.364 9	4.032 2
6	0.717 6	1.439 8	1.943 2	2.446 9	3.142 7	3.707 4
7	0.711 1	1.414 9	1.894 6	2.364 6	2.998 0	3.499 5
8	0.706 4	1.396 8	1.859 5	2.306 0	2.896 5	3.355 4
9	0.702 7	1.383 0	1.833 1	2.262 2	2.821 4	3.249 8
10	0.699 8	1.372 2	1.812 5	2.228 1	2.763 8	3.169 3
11	0.697 4	1.363 4	1.795 9	2.201 0	2.718 1	3.105 8
12	0.695 5	1.356 2	1.782 3	2.178 8	2.681 0	3.054 5
13	0.693 8	1.350 2	1.770 9	2.160 4	2.650 3	3.012 3
14	0.692 4	1.345 0	1.761 3	2.144 8	2.624 5	2.976 8
15	0.691 2	1.340 6	1.753 1	2.131 5	2.602 5	2.946 7
16	0.690 1	1.336 8	1.745 9	2.119 9	2.583 5	2.920 8
17	0.689 2	1.333 4	1.739 6	2.109 8	2.566 9	2.898 2
18	0.688 4	1.330 4	1.734 1	2.100 9	2.552 4	2.878 4
19	0.687 6	1.327 7	1.729 1	2.093 0	2.539 5	2.860 9
20	0.687 0	1.325 3	1.724 7	2.086 0	2.528 0	2.845 3
21	0.686 4	1.323 2	1.720 7	2.079 6	2.517 7	2.831 4
22	0.685 8	1.321 2	1.717 1	2.073 9	2.508 3	2.818 8
23	0.685 3	1.319 5	1.713 9	2.068 7	2.499 9	2.807 3
24	0.684 8	1.317 8	1.710 9	2.063 9	2.492 2	2.796 9
25	0.684 4	1.316 3	1.708 1	2.059 5	2.485 1	2.787 4

续表

n	α=0.25	0.10	0.05	0.025	0.01	0.005
26	0.684 0	1.315 0	1.705 8	2.055 5	2.478 6	2.778 7
27	0.683 7	1.313 7	1.703 3	2.051 8	2.472 7	2.770 7
28	0.683 4	1.312 5	1.701 1	2.048 4	2.467 1	2.763 3
29	0.683 0	1.311 4	1.699 1	2.045 2	2.462 0	2.756 4
30	0.682 8	1.310 4	1.697 3	2.042 3	2.457 3	2.750 0
31	0.682 5	1.309 5	1.695 5	2.039 5	2.452 8	2.744 0
32	0.682 2	1.308 6	1.693 9	2.036 9	2.448 7	2.738 5
33	0.682 0	1.307 7	1.692 4	2.034 5	2.444 8	2.733 3
34	0.681 8	1.307 0	1.690 9	2.032 2	2.441 1	2.728 4
35	0.681 6	0.306 2	1.689 6	2.030 1	2.437 7	2.723 8
36	0.681 4	1.305 5	1.688 3	2.028 1	2.434 5	2.719 5
37	0.681 2	1.304 9	1.687 1	2.026 2	2.431 4	2.715 4
38	0.681 0	1.304 2	1.686 0	2.024 4	2.428 6	2.711 6
39	0.680 8	1.303 6	1.684 9	2.022 7	2.425 8	2.707 9
40	0.680 7	1.303 1	1.683 9	2.021 1	2.423 3	2.704 5
41	0.680 5	1.302 5	1.682 9	2.019 5	2.420 8	2.701 2
42	0.680 4	1.302 0	1.682 0	2.018 1	2.418 5	2.698 1
43	0.680 2	1.301 6	1.681 1	2.016 7	2.416 3	2.695 1
44	0.680 1	1.301 1	1.680 2	2.015 4	2.414 1	2.692 3
45	0.680 0	1.300 6	1.679 4	2.014 1	2.412 1	2.680 6

附表6 F 分布表



$$P\{F(n,m)>F_\alpha(n,m)\}=\alpha$$

$\alpha=0.10$

m \ n	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.06	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80

续表

$\alpha=0.10$

m \ n	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.00	2.03	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00

续表

$\alpha=0.05$

m \ n	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.80	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96

续表

$\alpha=0.05$

m \ n	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.93
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.53	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

$\alpha=0.025$

续表

$m \backslash n$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	647.8	799.5	864.2	899.6	921.8	937.1	943.2	956.7	963.3	368.6	976.7	984.9	933.1	997.2	1 001	1 006	1 010	1 014	1 018
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49	39.50
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95	13.90
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31	8.26
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73	3.67
9	7.21	5.71	5.08	4.72	4.48	4.23	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.45	3.39	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.72	2.66	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.61	2.55	2.49
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.70	2.64	2.59	2.52	2.46	2.40
16	6.12	4.69	4.08	3.73	3.5	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.56	2.50	2.44	2.38	2.32	2.25

$\alpha=0.025$

续表

n / m	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.32	2.26	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.45	2.39	2.38	2.27	2.20	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.58	2.42	2.37	2.31	2.25	2.18	2.11	2.04
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.33	2.27	2.21	2.14	2.08	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.47	2.36	2.30	2.24	2.18	2.11	2.04	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.07	1.94
25	5.69	4.20	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41	2.30	2.24	2.18	2.12	2.05	1.98	1.91
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65	2.59	2.49	2.39	1.28	2.22	2.16	2.09	2.03	1.95	1.88
27	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63	2.57	2.47	2.36	1.25	2.19	2.13	2.07	2.00	1.93	1.85
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.98	1.91	1.83
29	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.32	2.21	2.15	2.09	2.03	1.96	1.89	1.81
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	3.20	2.14	2.07	2.01	1.94	1.87	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72	1.64
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58	1.48
120	5.15	3.08	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.58	1.43	1.31
∞	5.02	3.60	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.64	1.57	1.48	1.39	1.27	1.00

续表

$\alpha=0.01$

m \ n	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	1 052	4 999.5	5 403	5 625	5 764	5 859	5 928	5 982	6 022	6 056	6 106	6 157	6 209	6 235	6 261	6 287	6 313	6 339	6 366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	24.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.50	13.40
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.35	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65

续表

$\alpha=0.01$

m \ n	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.55	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.71	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

续表

$\alpha=0.005$

m \ n	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	16 211	20 000	21 615	22 500	23 056	23 437	23 715	23 925	24 091	24 224	24 426	24 630	24 836	24 940	25 044	25 148	25 253	25 359	25 465
2	198.5	199.0	199.2	199.2	199.3	199.3	199.4	199.4	199.4	199.4	199.4	199.4	199.4	199.5	199.5	199.5	199.5	199.5	199.5
3	55.55	49.80	47.47	46.19	45.39	44.84	44.43	44.13	43.88	43.69	43.39	43.08	42.78	42.62	42.47	42.31	42.15	41.99	41.83
4	31.33	26.28	24.26	23.15	22.46	21.97	21.62	21.35	21.14	20.97	20.70	20.44	20.17	20.03	19.89	19.75	19.61	19.47	19.32
5	22.78	18.31	16.53	15.56	14.94	14.51	14.20	13.96	13.77	13.62	13.38	13.15	12.90	12.78	12.66	12.53	12.40	12.27	12.14
6	18.63	14.54	12.92	12.03	11.46	11.07	10.79	10.57	10.39	10.25	10.03	9.81	9.59	9.47	9.36	9.24	9.12	9.00	8.88
7	16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.51	8.38	8.18	7.97	7.75	7.65	7.53	7.42	7.31	7.19	7.08
8	14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.34	7.21	7.01	6.81	6.61	6.50	6.40	6.29	6.18	6.06	5.95
9	13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.54	6.42	6.23	6.03	5.83	5.73	5.62	5.52	5.41	5.30	5.19
10	12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97	5.85	5.66	5.47	5.27	5.17	5.07	4.97	4.86	4.75	4.64
11	12.23	8.91	7.60	6.88	6.42	6.10	5.86	5.68	5.54	5.42	5.24	5.05	4.86	4.76	4.65	4.55	4.44	4.34	4.23
12	11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20	5.09	4.91	4.72	4.53	4.43	4.33	4.23	4.12	4.01	3.90
13	11.37	8.19	6.93	6.23	5.79	5.48	5.25	5.08	4.94	4.82	4.64	4.46	4.27	4.17	4.07	3.97	3.87	3.76	3.65
14	11.06	7.92	6.68	6.00	5.56	5.26	5.03	4.86	4.72	4.60	4.43	4.25	4.06	3.96	3.86	3.76	3.66	3.55	3.44
15	10.80	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54	4.42	4.25	4.07	3.88	3.79	3.69	3.58	3.48	3.37	3.26
16	10.58	7.51	6.30	5.64	5.21	4.91	4.69	4.52	4.38	4.27	4.10	3.92	3.73	3.64	3.54	3.44	3.33	3.22	3.11
17	10.38	7.35	6.16	5.50	5.07	4.78	4.56	4.39	4.25	4.14	3.97	3.79	3.61	3.51	3.41	3.31	3.21	3.10	2.98

续表

$\alpha=0.005$

m \ n	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
18	10.22	7.21	6.03	5.37	4.96	4.66	4.44	4.28	4.14	4.03	3.86	3.68	3.50	3.40	3.30	3.20	3.10	2.99	2.87
19	10.07	7.09	5.92	5.27	4.85	4.56	4.34	4.18	4.04	3.93	3.76	3.59	3.40	3.31	3.21	3.11	3.00	2.89	2.78
20	9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96	3.85	3.68	3.50	3.32	3.22	3.12	3.02	2.92	2.81	2.69
21	9.83	6.89	5.73	5.09	4.68	4.39	4.18	4.01	3.88	3.77	3.60	3.43	3.24	3.15	3.05	2.95	2.84	2.73	2.61
22	9.73	6.81	5.65	5.02	4.61	4.32	4.11	3.94	3.81	3.70	3.54	3.36	3.18	3.08	2.98	2.88	2.77	2.66	2.55
23	9.63	6.73	5.58	4.95	4.54	4.26	4.05	3.88	3.75	3.64	3.47	3.30	3.12	3.02	2.92	2.82	2.71	2.60	2.48
24	9.55	6.66	5.52	4.89	4.49	4.20	3.99	3.83	3.69	3.59	3.42	3.25	3.06	2.97	2.87	2.77	2.66	2.55	2.43
25	9.48	6.60	5.46	4.84	4.43	4.15	3.94	3.78	3.64	3.54	3.37	3.20	3.01	2.92	2.82	2.72	2.61	2.50	2.38
26	9.41	6.54	5.41	4.79	4.38	4.10	3.89	3.73	3.60	3.49	3.38	3.15	2.97	2.87	2.77	2.67	2.56	2.45	2.33
27	9.34	6.49	5.36	4.74	4.34	4.06	3.85	3.69	3.56	3.45	3.28	3.11	2.93	2.83	2.73	2.63	2.52	2.41	2.29
28	9.28	6.44	5.32	4.70	4.30	4.02	3.81	3.65	3.52	3.41	3.25	3.07	2.89	2.79	2.69	2.59	2.48	2.37	2.25
29	9.23	6.40	5.28	4.66	4.26	3.98	3.77	3.61	3.48	3.38	3.21	3.04	2.86	2.76	2.66	2.56	2.45	2.33	2.21
30	9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45	3.34	3.18	3.01	2.82	2.73	2.63	2.52	2.42	2.30	2.18
40	8.83	6.07	4.98	4.37	3.99	3.71	3.51	3.35	3.22	3.12	2.95	2.78	2.60	2.50	2.40	2.30	2.18	2.06	1.93
60	8.49	5.79	4.73	4.14	3.76	3.49	3.29	3.13	3.01	2.90	2.74	2.57	2.39	2.29	2.19	2.08	1.96	1.82	1.69
120	8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81	2.71	2.54	2.37	2.19	2.09	1.98	1.87	1.75	1.61	1.41
∞	7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62	2.52	2.36	2.19	2.00	1.90	1.79	1.67	1.53	1.36	1.00

续表

$\alpha=0.001$

m \ n	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4 053+	5 000+	5 404+	5 625+	5 764+	5 859+	5 929+	5 981+	6 023+	6 056+	6 107+	6 158+	6 209+	6 235+	6 261+	6 287+	6 313+	6 340+	6 366+
2	998.5	999.0	999.2	999.2	999.3	999.3	999.4	999.4	999.4	999.4	999.4	999.4	999.4	999.5	999.5	999.5	999.5	999.5	999.5
3	167.0	148.5	141.1	137.1	134.6	132.8	131.6	130.6	129.9	129.2	128.3	127.4	126.4	125.9	125.4	125.0	124.5	124.0	123.5
4	74.14	61.25	53.18	53.44	51.71	50.53	49.66	49.00	48.47	48.05	47.41	46.76	46.10	45.77	45.43	45.09	44.75	44.40	44.05
5	47.18	37.12	33.20	31.09	29.75	28.84	28.16	27.64	27.24	26.92	26.42	25.91	25.39	25.14	24.87	24.60	24.33	24.06	23.79
6	35.51	27.00	23.70	21.92	20.81	20.03	19.46	19.03	18.69	18.41	17.99	17.56	17.12	16.89	16.67	16.44	16.21	15.99	15.75
7	29.25	21.69	18.77	17.19	16.21	15.52	15.02	14.63	14.33	14.08	13.71	13.32	12.93	12.73	12.53	12.33	12.12	11.91	11.70
8	25.42	18.49	15.83	14.39	13.49	12.86	12.40	12.04	11.77	11.54	11.19	10.84	10.18	10.30	10.11	9.92	9.73	9.53	9.33
9	22.86	16.39	13.90	12.56	11.71	11.13	10.70	10.37	10.11	9.89	9.57	9.24	8.90	8.72	8.55	8.37	8.19	8.00	7.81
10	21.04	14.91	12.55	11.28	10.48	9.92	9.52	9.20	8.96	8.75	8.45	8.13	7.80	7.64	7.47	7.30	1.12	6.94	6.67
11	19.69	13.81	11.56	10.35	9.58	9.05	8.66	8.35	8.12	7.92	7.63	7.32	7.01	6.85	6.68	6.52	6.35	6.17	6.00
12	18.64	12.97	10.80	9.63	8.89	8.38	8.00	7.71	7.48	7.29	7.00	6.71	6.40	6.25	6.09	5.93	5.76	5.59	5.42
13	17.81	12.31	10.21	9.07	8.35	7.86	7.49	7.21	6.98	6.80	6.52	6.23	5.93	5.78	5.63	5.47	5.30	5.14	4.97
14	17.14	11.78	9.73	8.62	7.92	7.43	7.08	6.80	6.58	6.40	6.13	5.85	5.56	5.41	5.25	5.10	4.94	4.77	4.60
15	16.59	11.34	9.34	8.25	7.57	7.09	6.74	6.47	6.26	6.08	5.81	5.54	5.25	5.10	4.95	4.80	4.64	4.47	4.31
16	16.12	10.97	9.00	7.94	7.27	6.81	6.46	6.19	5.98	5.81	5.55	5.27	4.99	4.85	4.70	4.54	4.39	4.23	4.06
17	15.72	10.66	8.73	7.68	7.02	7.56	6.22	5.96	5.75	5.58	5.32	5.05	4.78	4.63	4.48	4.33	4.18	4.02	3.85

$\alpha=0.001$ 续表

m \ n	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
18	15.38	10.39	8.49	7.46	6.81	6.35	6.02	5.76	5.56	5.39	5.13	4.87	4.59	4.45	4.30	4.15	4.00	3.84	3.67
19	15.08	10.16	8.28	7.26	6.62	6.18	5.85	5.59	5.39	5.22	4.97	4.70	4.43	4.29	4.14	3.99	3.84	3.68	3.51
20	14.82	9.95	8.10	7.10	6.46	6.02	5.69	5.44	5.24	5.08	4.82	4.56	4.29	4.15	4.00	3.86	3.70	3.54	3.38
21	14.59	9.77	7.94	6.95	6.32	5.88	5.56	5.31	5.11	4.95	4.70	4.44	4.17	4.03	3.88	3.74	3.58	3.42	3.26
22	14.38	9.61	7.80	6.81	6.19	5.76	5.44	5.19	4.99	4.83	4.58	4.33	4.06	3.92	3.78	3.63	3.48	3.32	3.15
23	14.19	9.47	7.67	6.69	6.08	5.65	5.33	5.09	4.89	4.73	4.48	4.23	3.96	3.82	3.68	3.53	3.38	3.22	3.05
24	14.03	9.34	7.55	6.59	5.98	5.55	5.23	4.99	4.80	4.64	4.39	4.14	3.87	3.74	3.59	3.45	3.29	3.14	2.97
25	13.88	9.22	7.45	6.49	5.88	5.46	5.15	4.91	4.71	4.56	4.31	4.06	3.79	3.66	3.52	3.37	3.22	3.06	2.89
26	13.74	9.12	7.36	6.41	5.80	5.38	5.07	4.83	4.64	4.48	4.24	3.99	3.72	3.59	3.44	3.30	3.15	2.99	2.82
27	13.61	9.02	7.27	6.33	5.73	5.31	5.00	4.76	4.57	4.41	4.17	3.92	3.66	3.52	3.38	3.23	3.08	2.92	2.75
28	13.50	8.93	7.19	6.25	5.66	5.24	4.93	4.69	4.50	4.35	4.11	3.86	3.60	3.46	3.32	3.18	3.02	2.86	2.69
29	13.39	8.85	7.12	6.19	5.59	5.18	4.87	4.64	4.45	4.29	4.05	3.80	3.54	3.41	3.27	3.12	2.97	2.81	2.54
30	13.29	8.77	7.05	6.12	5.53	5.12	4.82	4.58	4.39	4.24	4.00	3.75	3.49	3.36	3.22	3.07	2.92	2.76	2.59
40	12.61	8.25	6.60	5.70	5.13	4.73	4.44	4.21	4.02	3.87	3.64	3.40	3.15	3.01	2.87	2.73	2.57	2.41	2.23
60	11.97	7.76	6.17	5.31	4.76	4.37	4.09	3.87	3.69	3.54	3.31	3.08	2.83	2.69	2.55	2.41	2.25	2.08	1.89
120	11.38	7.32	5.79	4.95	4.42	4.04	3.77	3.55	3.38	3.24	3.02	2.78	2.53	2.40	2.26	2.11	1.95	1.76	1.54
∞	10.83	6.91	5.42	4.62	4.10	3.74	3.47	3.27	3.10	2.96	2.74	2.51	2.27	2.12	1.99	1.84	1.66	1.45	1.00

+表示要将所列数乘以100

大学数学系列教材

高等数学（上册）

高等数学（下册）

线性代数与解析几何

概率论与数理统计

复变函数与积分变换

ISBN 978-7-04-035542-0

定价 19.50 元