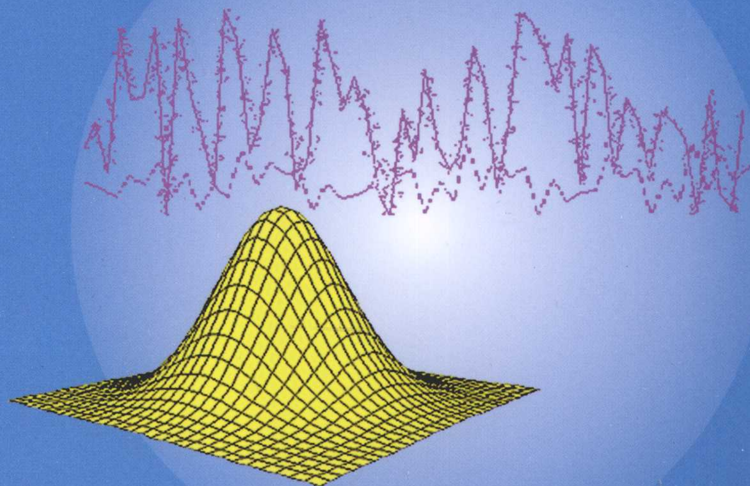
高等院校（非数学类）数学基础课教材

Methods of Applied Statistics

应用统计方法

梁飞豹 吕书龙 薛美玉 刘文丽 ◎ 编著



北京大学出版社
PEKING UNIVERSITY PRESS

高等院校（非数学类）数学基础课教材

应用统计方法

梁飞豹　吕书龙　薛美玉　刘文丽　编著

北京大学出版社
PEKING UNIVERSITY PRESS

图书在版编目(CIP)数据

应用统计方法/梁飞豹等编著.—北京:北京大学出版社,2010.8
ISBN 978-7-301-17646-7

Ⅰ.①应… Ⅱ.①梁… Ⅲ.①应用统计学-高等学校-教材 Ⅳ.①C8

中国版本图书馆 CIP 数据核字(2010)第 160829 号

书　　名:应用统计方法
著作责任者:梁飞豹　吕书龙　薛美玉　刘文丽　编著
责 任 编 辑:刘　勇
封 面 设 计:张　虹
标 准 书 号:ISBN 978-7-301-17646-7/O・0822
出 版 发 行:北京大学出版社
地　　址:北京市海淀区成府路 205 号　100871
网　　址:http://www.pup.cn　电子邮箱:zpup@pup.pku.edu.cn
电　　话:邮购部 62752015　发行部 62750672　理科编辑部 62752021　出版部 62754962
印 刷 者:北京中科印刷有限公司
经 销 者:新华书店
787mm×960mm　16 开本　19.75 印张　400 千字
2010 年 8 月第 1 版　2010 年 8 月第 1 次印刷
定　　价:35.00 元

内 容 简 介

本书是为高等院校非数学类硕士研究生和高年级本科生编写的教材，其主要内容有：概率论基础、数理统计的基本概念与抽样分布、参数估计、假设检验、回归分析、非参数统计初步、方差分析与正交试验设计、多元统计分析. 本书着重介绍各种统计方法的统计思想、问题的背景、应用条件及实际意义，使学生能够对统计方法及其应用有一个系统、全面的了解. 同时我们还为本书配套开发了相应的统计软件(LLLStat)，该软件能方便处理本书中的所有统计方法. 书中各章均配有适量的习题，书末附有详细解答. 本书列入福州大学“211工程”研究生规划教材.

本书可作为高等院校工科、经济、管理、农学、医学等非数学类硕士、博士研究生以及高年级本科生学习统计分析方法(或应用数理统计)课程的教材，也可作为相关学科和工程技术人员的参考书.

为了方便教师教学和学生学习，我们愿意为使用本书的读者免费提供一套统计软件(LLLStat)，需要此软件的读者请与作者联系，电子邮件地址：LLLStat@163.com.

目　录

第一章　概率论基础 …………………… (1)
§1.1　随机事件及其概率 ………… (1)
一、样本空间与随机事件 …………… (1)
二、事件的概率 ……………………… (3)
三、条件概率与乘法公式 …………… (4)
四、事件的独立性 …………………… (6)
§1.2　随机变量及其分布 ………… (7)
一、随机变量与分布函数 …………… (7)
二、多维随机变量及其分布 ………… (12)
三、随机变量函数的分布 …………… (17)
§1.3　随机变量的数字特征 …… (21)
一、数学期望 ………………………… (21)
二、方差 ……………………………… (23)
三、随机变量"标准化"及矩 ………… (24)
四、常见分布的数学期望与方差 …… (24)
五、协方差与相关系数 ……………… (25)
六、多维随机变量的数字特征 ……… (26)
§1.4　极限定理初步 …………… (27)
一、随机变量序列的收敛性 ………… (27)
二、大数定律 ………………………… (28)
三、中心极限定理 …………………… (29)
习题一 ………………………………… (30)
第二章　数理统计的基本概念与抽样分布 …………………… (32)
§2.1　数理统计的基本概念 …… (33)
一、总体与样本 ……………………… (33)
二、统计量 …………………………… (34)
§2.2　经验分布函数与直方图 … (37)
一、经验分布函数 …………………… (37)
二、直方图 …………………………… (38)
§2.3　常用的概率分布 ………… (40)
一、χ^2 分布 ……………………………… (40)
二、t 分布 ………………………………… (42)
三、F 分布 ………………………………… (43)
四、概率分布的分位点 ……………… (45)
§2.4　抽样分布 …………………… (46)
一、正态总体的抽样分布 …………… (46)
二、非正态总体的一些抽样分布 …… (49)
习题二 ………………………………… (50)
第三章　参数估计 …………………… (52)
§3.1　点估计 ……………………… (52)
一、矩法 ……………………………… (52)
二、极大似然估计法 ………………… (55)
三、贝叶斯估计 ……………………… (60)
§3.2　估计量的评价标准 ……… (64)
一、无偏性 …………………………… (65)
二、有效性 …………………………… (66)
三、一致性 …………………………… (70)
§3.3　区间估计 …………………… (72)
§3.4　正态总体参数的区间估计 ………………………… (73)
一、单总体的情形 …………………… (73)
二、双总体的情形 …………………… (78)
§3.5　非正态总体参数的区间估计 ………………………… (82)
一、指数分布参数的区间估计 ……… (83)
二、0-1 分布参数的区间估计 ……… (83)
§3.6　单侧置信区间 …………… (84)
习题三 ………………………………… (88)
第四章　假设检验 …………………… (91)
§4.1　假设检验的基本概念 …… (91)
一、问题的提出 ……………………… (91)

二、假设检验的基本原理 …………… (92)
三、假设检验的两类错误 …………… (94)
四、假设检验的一般步骤 …………… (95)
§4.2 单个正态总体参数的假设检验 ……………………… (96)
一、单个正态总体均值的假设检验 … (96)
二、单个正态总体方差的假设检验 ……………………… (101)
§4.3 两个正态总体参数的假设检验 ……………………… (103)
一、两个正态总体均值的假设检验 ……………………… (104)
二、两个正态总体方差的假设检验 ……………………… (107)
§4.4 非正态总体参数的假设检验 ……………………… (110)
一、指数分布参数的假设检验 …… (110)
二、0-1 分布参数的假设检验 ……… (111)
习题四 ………………………… (112)

第五章 回归分析 …………………… (114)
§5.1 相关分析 …………………… (114)
一、简单相关系数 ………………… (114)
二、相关系数的检验 ……………… (115)
§5.2 线性回归模型 …………… (116)
一、回归的由来 …………………… (116)
二、回归分析的基本概念 ………… (117)
三、线性回归模型 ………………… (118)
§5.3 最小二乘估计及其性质 ……………………… (119)
一、最小二乘估计 ………………… (119)
二、一元线性回归 ………………… (121)
三、最小二乘估计的性质 ………… (122)
§5.4 回归方程和回归系数的检验 ……………………… (124)
一、复相关系数 …………………… (124)
二、回归方程的 F 检验 …………… (125)
三、回归系数的显著性检验 ……… (126)
§5.5 因变量的预测 …………… (127)
一、点预测 ……………………… (128)
二、区间预测 …………………… (128)
§5.6 自变量的选择与逐步回归 ……………………… (129)
一、自变量选择的准则 ………… (130)
二、选择最优回归方程 ………… (131)
三、逐步回归 …………………… (132)
§5.7 非线性回归 ……………… (134)
一、可线性化的非线性模型 ……… (134)
二、一般的非线性回归模型 ……… (138)
习题五 ………………………… (140)

第六章 非参数统计初步 ………… (143)
§6.1 非参数假设检验 ………… (143)
一、分布函数的拟合检验 ………… (143)
二、列联表的独立性检验 ………… (147)
三、一致性检验 …………………… (150)
四、总体对称中心的检验 ………… (151)
§6.2 非参数回归模型 ………… (155)
一、核密度估计 …………………… (155)
二、非参数回归模型 ……………… (158)
习题六 ………………………… (162)

第七章 方差分析与正交试验设计 ……………………… (164)
§7.1 单因素方差分析 ………… (164)
一、单因素试验 …………………… (165)
二、提出假设 …………………… (165)
三、统计分析 …………………… (165)
四、例题分析 …………………… (167)
§7.2 双因素方差分析 ………… (169)
一、有交互作用的双因素方差分析 ……………………… (169)
二、无交互作用的双因素方差分析 ……………………… (173)
§7.3 正交试验设计 …………… (176)

一、正交表 …………………………… (176)
二、无交互作用的正交试验 ……… (177)
三、有交互作用的正交试验 ……… (183)
习题七 ………………………………… (187)
第八章 多元统计分析 ……… (190)
§8.1 多维随机变量 ……… (190)
一、多维随机变量 ……………… (190)
二、多元正态分布 ……………… (191)
三、抽样与统计量 ……………… (192)
四、参数估计 …………………… (197)
§8.2 判别分析与聚类分析 …… (198)
一、距离 ………………………… (198)
二、判别分析 …………………… (200)
三、聚类分析 …………………… (207)
§8.3 主成分分析 …………… (212)
一、主成分分析的原理 ………… (212)
二、样本主成分的计算步骤 …… (213)
习题八 ………………………………… (216)
附录Ⅰ 常见分布参数、估计量及数字特征一览表 ………………… (219)
附录Ⅱ 常用分布表 ……………… (220)
附表1 泊松分布表 ……………… (220)
附表2 标准正态分布表 ………… (223)
附表3 t分布的上侧分位点表 ………………………… (225)
附表4 χ^2 分布的上侧分位点表 ………………………… (226)
附表5 F分布的上侧分位点表 ………………………… (227)
附表6 相关系数检验表 ………… (269)
附表7 符号检验表 ……………… (270)
附表8 符号秩和检验表 ………… (271)
附录Ⅲ 正交表 ……………………… (272)
习题解答 ……………………………… (277)
参考文献 ……………………………… (304)

第一章 概率论基础

概率论是数理统计的理论基础，它是研究随机现象统计规律性的一门数学学科，其应用广泛，是从事科技、管理、经济等工作者必备的数学工具. 本章是为以后各章做准备的，主要介绍概率论的基本概念、基本定理以及常用公式，内容包括随机事件及其概率，随机变量及其分布，随机变量的数字特征，极限定理初步等.

§1.1 随机事件及其概率

一、样本空间与随机事件

在自然界及各种社会活动中观察到的各种现象大体上可归结为：确定性现象与不确定性现象. 我们把事前可以预知结果的，在某些确定的条件满足时，某一确定的结果必然会发生的现象称为**确定性现象**. 例如：在一个标准大气压下，水加热到 100 ℃ 时一定沸腾；上抛物体必定落向地面；等等. 这种现象由确定性数学来研究. 而在不确定性现象中，有一类称为随机现象，即：我们把事前能够预知所有可能结果，但在每次试验时不能确定哪一种结果将要出现的现象称为**偶然现象**或**随机现象**. 例如：抛掷一枚质地均匀的硬币，硬币落地后可能是正面(国徽一面)朝上，也可能是反面朝上；新生婴儿可能是男孩，也可能是女孩；等等. 这类现象由概率统计来研究.

1. 随机试验

为了研究随机现象内部隐含的规律性，必须对随机现象进行观测或试验，这种观测或试验统称为**随机试验**，简称为**试验**，记为 E.

随机试验具有如下特点：

(1) 试验可以在相同条件下重复进行；

(2) 每一次试验，可能出现各种不同结果，总共有可能出现的哪几种结果，是可以事先明确知道的；

(3) 每一次试验,实际只出现一种结果,至于实际出现哪一种结果,试验之前是无法预先知道的.

2. 样本空间

在研究随机试验 E 时,首先必须弄清楚这个试验可能出现的所有结果.称每一个可能的结果为**样本点**,一般用小写字母 ω 表示.全体样本点构成的集合称为**样本空间**,一般用大写字母 Ω 表示.

3. 随机事件

在一个随机试验中,可能发生也可能不发生的事情称为**随机事件**,简称**事件**,一般用大写字母 $A,B,C,\cdots$表示.随机事件也可定义为由若干个样本点组成的集合.由一个样本点组成的事件称为**基本事件(不可分事件)**,多于一个样本点组成的事件称为**复合事件(可再分事件)**.每次试验中,一定发生的事情称为**必然事件**,即所有样本点构成的集合,记为 Ω.每次试验中一定不发生的事情称为**不可能事件**,即不含有任何样本点的集合,记为$\varnothing$.

4. 事件间的关系与运算

事件是样本空间的子集,所以事件间的关系与运算同集合间的关系与运算完全一致.

(1) **包含**关系 $A\subset B$:若事件 A 发生,则事件 B 一定发生;

(2) **相等**关系 $A=B$:若 $A\subset B$ 且 $B\subset A$;

(3) **积(交)**事件 $A\cap B$ 或 AB:事件 A 与事件 B 同时发生的事件;

(4) **互不相容**(互斥)关系:$A\cap B=\varnothing$或事件 A 与 B 不能同时发生;

(5) **和(并)**事件 $A\cup B$:事件 A 与 B 至少有一个发生的事件;

(6) **对立(逆)**事件 $\overline{A}$:事件 A 不发生的事件;

(7) **差**事件 $A-B$:事件 A 发生,而事件 B 不发生的事件.

在进行事件运算时,一般是先进行逆的运算,再进行交的运算,最后再进行并或差的运算.事件运算的规律如下:

(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**:$(A\cup B)\cap C=(A\cap C)\cup(B\cap C)$,

$(A\cap B)\cup C=(A\cup C)\cap(B\cup C)$;

(4) **对偶原则**:$\overline{A\cup B}=\overline{A}\cap\overline{B}$,$\overline{A\cap B}=\overline{A}\cup\overline{B}$.

这个原则可推广到任意有限个或可列个事件的情况,即

$$\overline{\bigcup_{i=1}^{n} A_i}=\bigcap_{i=1}^{n}\overline{A_i},\quad \overline{\bigcap_{i=1}^{n} A_i}=\bigcup_{i=1}^{n}\overline{A_i};$$

$$\overline{\bigcup_{i=1}^{\infty} A_i}=\bigcap_{i=1}^{\infty}\overline{A_i},\quad \overline{\bigcap_{i=1}^{\infty} A_i}=\bigcup_{i=1}^{\infty}\overline{A_i}.$$

二、事件的概率

研究随机现象不仅要知道可能出现哪些事件,还要知道各种事件出现的可能性的大小.我们把衡量事件发生的可能性大小的数量指标称为事件的概率,事件 A 的概率用 $P(A)$ 表示.

1. 概率的直观定义

(1) **统计概率**:设在相同条件下对事件 A 重复进行的 n 次试验中,事件 A 出现 m 次,当试验次数 n 充分大时,事件 A 出现的频率 $f_n(A)=m/n$ 的稳定值 p 称为事件 A 的概率,记为 $P(A)$,即

$$p = P(A) \approx f_n(A) = \frac{m}{n}. \tag{1.1.1}$$

(2) **古典概率**:在古典概型试验中,设只有 n 个等可能的基本事件,随机事件 A 包含有 m 个基本事件,则称 m/n 为事件 A 的概率,记为

$$P(A) = \frac{\text{事件 } A \text{ 所包含的基本事件数}}{\text{所有可能的基本事件数}} = \frac{m}{n}, \tag{1.1.2}$$

其中古典概型试验是指具有下列两个特征的随机试验:

特征 A 有限性:试验的所有可能结果只有有限个基本事件.

特征 B 等可能性:每次试验中各基本事件出现的可能性均相同,且任何两个基本事件不可能同时发生.

(3) **几何概率**:在几何概型试验中,样本空间为 Ω,事件 $A\subset\Omega$,则事件 A 的概率为

$$P(A) = \frac{S_A}{S_\Omega} = \frac{A \text{ 的几何度量}}{\Omega \text{ 的几何度量}}, \tag{1.1.3}$$

其中几何度量指长度、面积、体积等,几何概型是指具有下列两个特征的随机试验:

特征 A 有限区域、无限样本点:试验的所有可能结果为无穷多个样本点,但其样本空间 Ω 表现为某一几何区域(直线、平面、三维空间)时为有限区域.

特征 B 等可能性:试验中各基本事件出现的可能性相同.

以上定义的三种概率都具有如下性质:

性质 1(非负性) 对于任意随机事件 A,有 $0\leqslant P(A)\leqslant 1$;

性质 2(规范性) $P(\Omega)=1$;

性质 3(有限可加性) 设 $A_1,A_2,\cdots,A_n$ 是一组两两互不相容的事件,则有

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

但它们都是在特殊情况下给出的事件概率的计算方法,具有明显的局限性,不能作为事件概率的严格定义.

2. 概率的公理化定义

定义 1.1.1(概率的公理化定义) 设随机试验 E 的样本空间为 Ω,对试验 E 的任一随机事件 A,定义实值函数 $P(A)$.若它满足以下三条公理:

公理 1 非负性:$P(A)\geqslant 0$;

公理 2 规范性:$P(\Omega)=1$;

公理 3 可列可加性:对于可列无穷多个两两互不相容的随机事件 $A_1,A_2,\cdots,A_n,\cdots$,有

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

则称 $P(A)$为事件 A 的概率.

3. 概率的性质

性质 1 不可能事件的概率为 0,即 $P(\varnothing)=0$.

性质 2(有限可加性) 若随机事件 $A_1,A_2,\cdots,A_n$ 两两互不相容,则

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

性质 3 对任一事件 A,有 $P(A)=1-P(\overline{A})$;

性质 4 对任意的两个事件 A,B, 若 $A\supset B$,则

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

推论 若 $A\supset B$,则 $P(A)\geqslant P(B)$.

性质 5 对任意的两个事件 A,B,有

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

推论 对任意 n 个事件 $A_1,A_2,\cdots,A_n$,有

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

三、条件概率与乘法公式

1. 条件概率

在实际问题中,往往会遇到求在事件 B 已经发生的条件下,事件 A 发生的概率,这时由于附加了条件,它与事件 A 发生的概率 $P(A)$的意义是不同的,我们把这种概率称为条件概率,记为 $P(A|B)$.

定义 1.1.2 设 A,B 是两个事件,且 $P(B)>0$,称

$$P(A|B)=\frac{P(AB)}{P(B)} \tag{1.1.4}$$

为事件 B 发生的条件下事件 A 发生的**条件概率**.

可以验证,条件概率仍然满足概率的 3 条公理,即:

(1) 对于每一个事件 A,有 $P(A|B)\geqslant 0$;

(2) $P(\Omega|B)=1$;

(3) 设 $A_1,A_2,\cdots,A_n,\cdots$是两两互不相容的事件,则有

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

2. 乘法公式

利用条件概率的定义,自然地得到概率的乘法公式.

定理 1.1.1(乘法公式) 设 A,B 为任意事件,若 $P(B)>0$, 则

$$P(AB)=P(B)P(A|B);\tag{1.1.5}$$

若 $P(A)>0$, 则

$$P(AB)=P(A)P(B|A).\tag{1.1.6}$$

实际上,当 $P(B)=0$(或 $P(A)=0$)时,(1.1.5) 式(或(1.1.6)式)也成立. 因为此时有

$$0\leqslant P(AB)\leqslant P(B)=0\quad(\text{或 }0\leqslant P(AB)\leqslant P(A)=0).$$

乘法公式可以推广至多个随机事件的情形.

推论 设有 n 个事件 $A_1,A_2,\cdots,A_n$,则

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)\cdots P(A_n|A_1A_2\cdots A_{n-1}).\tag{1.1.7}$$

3. 全概率公式

定义 1.1.3 若事件组 $A_1,A_2,\cdots,A_n$,满足下面两个条件:

(1) $A_1\cup A_2\cup\cdots\cup A_n=\Omega$;

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

则称事件组 $A_1,A_2,\cdots,A_n$ 为**完备事件组**.

定理 1.1.2(全概率公式) 设 $A_1,A_2,\cdots,A_n$ 是 n 个两两互不相容的事件,$P(A_i)>0$ $(i=1,2,\cdots,n)$且 $B\subset A_1\cup A_2\cup\cdots\cup A_n$,即 B 的发生总是与 $A_1,A_2,\cdots,A_n$ 之一同时发生,则对于事件 B,有

$$P(B)=\sum_{i=1}^{n}P(A_i)P(B|A_i).\tag{1.1.8}$$

特别地,若 $A_1,A_2,\cdots,A_n$ 是完备事件组,且 $P(A_i)>0$ $(i=1,2,\cdots,n)$,则定理结论同样成立.

4. 贝叶斯公式

定理 1.1.3(贝叶斯(Bayes)公式) 设 $A_1,A_2,\cdots,A_n$ 是 n 个两两不相容的事件,$P(A_i)>0$ $(i=1,2,\cdots,n)$且 $B\subset A_1\cup A_2\cup\cdots\cup A_n$,即 B 的发生总是与 $A_1,A_2,\cdots,A_n$ 之一同时发

生，则在 B 已经发生的条件下（$P(B)>0$），A_i 的条件概率为

$$P(A_i|B)=\frac{P(A_i)P(B|A_i)}{\sum_{k=1}^{n}P(A_k)P(B|A_k)}\quad(i=1,2,\cdots,n).\tag{1.1.9}$$

特别地，若 $A_1,A_2,\cdots,A_n$ 是完备事件组，且 $P(A_i)>0\ (i=1,2,\cdots,n)$，则定理结论同样成立.

四、事件的独立性

1. 事件的独立性

事件的独立性是概率论中最重要的概念之一. 所谓两个事件 A 与 B 相互独立，直观上说就是它们互不影响，或者说，事件 A 发生与否不会影响事件 B 发生的可能性，事件 B 发生与否不会影响事件 A 发生的可能性，用数学式子表示就是：

$$P(B|A)=P(B),\quad 且\quad P(A|B)=P(A).$$

但上面两个式子要求 $P(A)>0$ 或 $P(B)>0$. 考虑到更一般的情形，我们给出如下定义：

定义 1.1.4 对任意两个事件 A,B，若有

$$P(AB)=P(A)P(B),\tag{1.1.10}$$

则称事件 A 与 B **相互独立**.

需要强调一点的是，事件的独立性与事件的互不相容是两个完全不同的概念. 实际上，从定义即知，如果两个具有正概率的事件是互不相容的，那么它们一定是不独立的；反之，如果两个具有正概率的事件是相互独立的，那么这两个事件不可能互不相容.

定理 1.1.4 若事件 A 与 B 相互独立，则 A 与 $\bar{B}$，$\bar{A}$ 与 B，$\bar{A}$ 与 $\bar{B}$ 也分别相互独立.

当我们考虑多个事件之间是否相互独立时，除了必须考虑任意两事件之间的相互关系外，还要考虑到多个事件的乘积对其他事件的影响. 基于如此考虑，我们给出下面一般的定义.

定义 1.1.5 若 n 个事件 $A_1,A_2,\cdots,A_n$ 满足

$$P(A_{i_1}A_{i_2}\cdots A_{i_k})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k})$$
$$(1\leqslant i_1<i_2<\cdots<i_k\leqslant n,\ 1<k\leqslant n),\tag{1.1.11}$$

则称 $A_1,A_2,\cdots,A_n$ 相互独立.

2. 伯努利概型

定义 1.1.6 具有以下两个特点的随机试验称为 n 次**伯努利概型试验**或**伯努利概型**：

(1) 在相同条件下，重复 n 次做同一试验，每次试验只有两个可能结果 A 和 $\bar{A}$，且

$$P(A)=p\ (0<p<1),\quad P(\bar{A})=1-p=q;$$

(2) n 次试验是相互独立的（即每次试验结果出现的概率不受其他各次试验结果的影响），或设 A_i 表示“第 i 次试验 A 发生”（$i=1,2,\cdots,n$），则 $A_1,A_2,\cdots,A_n$ 相互独立.

定理 1.1.5 设在 n 次伯努利概型中,每次试验事件 A 发生的概率为 $p(0<p<1)$,则在 n 次试验中,事件 A 恰好发生 k 次的概率为

$$P_n(k)=C_n^k p^k q^{n-k} \quad (k=0,1,2,\cdots,n), \tag{1.1.12}$$

其中 $q=1-p$.

§1.2 随机变量及其分布

为了深入研究和全面掌握随机现象的统计规律,我们将随机试验的结果与实数对应起来,即将随机试验的结果数量化,引入随机变量的概念. 随机变量是概率论中最基本的概念之一,用它描述随机现象是近代概率中最重要的方法,它使概率论从事件及其概率的研究扩大到随机变量及其概率分布的研究. 这样也就可以应用微积分等近代数学工具,使概率论成为真正的一门数学分支.

一、随机变量与分布函数

1. 随机变量

定义 1.2.1 设 E 是随机试验,Ω 是其样本空间. 如果对每个 $\omega\in\Omega$,都有一个确定的实数 $X(\omega)$ 与之对应,则称 Ω 上的实值函数 $X(\omega)$ 为**随机变量**,简记为 X. 通常用大写字母 X,Y,Z,…或希腊字母 ξ,η,…表示随机变量.

从随机试验可能出现的结果来看,随机变量可分为两大类,一类是随机变量 X 的所有可能取值为有限个或可列无穷个值,这种类型的随机变量称为**离散型随机变量**;另一类就是非离散型随机变量. 非离散型随机变量包含的范围很广,情况比较复杂,我们只关注其中最重要也是实际中常遇到的**连续型随机变量**.

由于随机变量 X 的取值具有随机性,对随机变量 X 而言,$\{X>x\}$,$\{X=x\}$,$\{X\leqslant x\}$,$\{a<X\leqslant b\}$ 等都表示随机事件,其概率相应地简记为 $P(X>x)$,$P(X=x)$,$P(X\leqslant x)$,$P(a<X\leqslant b)$. 注意到这些事件都可以通过形如 $\{X\leqslant x\}$ 的事件来表示. 如

$$\{X<x\}=\bigcup_{k=1}^{\infty}\left\{X\leqslant x-\frac{1}{k}\right\};$$

$$\{X=x\}=\{X\leqslant x\}-\{X<x\}=\{X\leqslant x\}-\bigcup_{k=1}^{\infty}\left\{X\leqslant x-\frac{1}{k}\right\};$$

$$\{a<X\leqslant b\}=\{X\leqslant b\}-\{X\leqslant a\};$$

……

故我们只需考虑 $\{X\leqslant x\}$ 这种事件的概率 $P(X\leqslant x)$ 即可,于是引入分布函数的定义.

2. 分布函数

定义 1.2.2 设 X 是随机变量,对任意实数 x,令

$$F(x)=P(X\leqslant x),\quad x\in(-\infty,+\infty),$$

则称函数 $F(x)$ 为随机变量 X 的**分布函数**.

分布函数是一个定义在全体实数上的一个普通实函数，同时分布函数也具有明确的概率意义：对任意实数 x，$F(x)$ 就是随机变量 X 落在区间 $(-\infty, x]$ 上的概率.

设 X 为随机变量，我们所关注的随机事件都可以通过形如 $\{X \leqslant x\}$ 事件来表示，所以任何随机事件的概率都可以用分布函数 $F(x)$ 来表示，如：

$$P(X < x) = F(x-0); \tag{1.2.1}$$

$$P(X = x) = F(x) - F(x-0); \tag{1.2.2}$$

$$P(x_1 < X \leqslant x_2) = F(x_2) - F(x_1); \tag{1.2.3}$$

$$P(x_1 \leqslant X \leqslant x_2) = F(x_2) - F(x_1 - 0); \tag{1.2.4}$$

$$P(x_1 < X < x_2) = F(x_2 - 0) - F(x_1); \tag{1.2.5}$$

等等（$F(x-0)$ 表示分布函数 $F(x)$ 在 x 点的左极限）.

由定义 1.2.2 可得出分布函数 $F(x)$ 具有如下基本性质：

(1) $F(x)$ 是单调不减函数，即 $x_1 < x_2$ 时，有 $F(x_1) \leqslant F(x_2)$；

(2) $F(x)$ 非负有界，即 $0 \leqslant F(x) \leqslant 1$ $(-\infty < x < +\infty)$，且

$$F(-\infty) = \lim_{x \to -\infty} F(x) = 0,\ F(+\infty) = \lim_{x \to +\infty} F(x) = 1;$$

(3) $F(x)$ 是右连续函数，即 $F(x+0) = F(x)$.

反过来可以证明，任给一个满足性质(1)，(2)，(3)的实值函数 $F(x)$，它必是某个随机变量的分布函数. 所以，性质(1)，(2)，(3)是 $F(x)$ 成为某个随机变量的分布函数的充分必要条件. 顺便指出，即使随机变量 X 和 Y 的分布函数相同（称为 X 与 Y 同分布），也不能误认为 $X = Y$，这时 X 与 Y 有可能是意义完全不同的两个随机变量.

3. 离散型随机变量及其分布

定义 1.2.3 设离散型随机变量 X 的所有可能取值为 $x_k (k=1,2,\cdots)$，X 取各个可能值的概率为

$$P(X = x_k) = p_k \quad (k = 1,2,\cdots), \tag{1.2.6}$$

则称(1.2.6)式为随机变量 X 的**概率分布**或**分布律**.

X 的分布律(1.2.6)也可写成如下的表格形式，称之为 X 的**分布列**：

表 1.2.1

X	x_1	x_2	$\cdots$	x_k	$\cdots$
P	p_1	p_2	$\cdots$	p_k	$\cdots$

表 1.2.1 直观地表明，对于离散型随机变量要全面掌握它的统计规律性，必须知道它所有可能的取值及取每一个可能值的概率. 根据概率的性质，易知 $p_k (k=1,2,\cdots)$ 满足下面两个性质：

(1) $p_k \geqslant 0$, $k=1,2,\cdots$; (1.2.7)

(2) $\sum_k p_k = 1.$ (1.2.8)

反之,任给有限或可列无穷多个满足(1.2.7),(1.2.8)式的实数 $p_k(k=1,2,\cdots)$,必是某个离散型随机变量 X 的分布律.

一般地,对于概率分布为式(1.2.6)的离散型随机变量 X,由概率的可加性知 X 的分布函数为

$$F(x) = P(X \leqslant x) = \sum_{x_k \leqslant x} P(X = x_k) = \sum_{x_k \leqslant x} p_k \quad (-\infty < x < +\infty). \tag{1.2.9}$$

4. 连续型随机变量及其分布

定义 1.2.4 设随机变量 X 的分布函数为 $F(x)$,若存在非负实函数 $f(x)$,使得对任意的实数 x,都有

$$F(x) = P(X \leqslant x) = \int_{-\infty}^{x} f(t)\mathrm{d}t, \tag{1.2.10}$$

则称 X 为**连续型随机变量**,其中 $f(x)$称为 X 的**概率密度函数**,简称**概率密度**或**分布密度**.

概率密度函数 $f(x)$具有下列性质:

(1) $f(x) \geqslant 0$; (1.2.11)

(2) $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1$; (1.2.12)

(3) 对任何实数 $a,b(a<b)$,有

$$P(a < X \leqslant b) = F(b) - F(a) = \int_a^b f(x)\mathrm{d}x; \tag{1.2.13}$$

(4) 对任意实数 x,$P(X=x)=0$;

(5) $F(x)$是连续函数,且在 $f(x)$的连续点 x 处有

$$F'(x) = f(x). \tag{1.2.14}$$

5. 常见随机变量及其分布

(1) **0-1 分布**(离散型):若随机变量 X 只可能取 0 和 1 两个值,其概率分布为

$$P(X = 1) = p, \quad P(X = 0) = 1 - p \quad (0 < p < 1), \tag{1.2.15}$$

则称 X 服从参数为 p 的 0-1 分布.

(2) **二项分布**(离散型):若随机变量 X 的概率分布为

$$P_n(k) = P(X = k) = \mathrm{C}_n^k p^k q^{n-k} \quad (k = 0,1,2,\cdots,n),$$
$$0 < q = 1 - p < 1, \tag{1.2.16}$$

则称 X 服从参数为 n,p 的二项分布,记为 $X \sim B(n,p)$. 当 $n=1$ 时,有

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

这就是 0-1 分布,故 0-1 分布是二项分布在 $n=1$ 时的特例.

(3) **泊松分布**(离散型):若随机变量 X 的概率分布为

$$P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}\quad(k=0,1,2,\cdots),\tag{1.2.17}$$

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

虽然泊松分布本身是一种非常重要的分布,但有趣的是,历史上它却是作为二项分布的近似.在计算二项分布 $B(n,p)$ 的概率 $P(X=k)=C_n^k p^k q^{n-k}$ 时,若 n 很大,p 较小,可用 $e^{-\lambda}\frac{\lambda^k}{k!}$ 近似代替 $C_n^k p^k(1-p)^{n-k}$,其中 $\lambda=np$,从而得到以下近似公式

$$C_n^k p^k(1-p)^{n-k}\approx e^{-\lambda}\frac{\lambda^k}{k!}.\tag{1.2.18}$$

实际应用时,当 $n\geqslant 10,p\leqslant 0.1$ 时就可采用上述近似公式计算.当 $n\geqslant 20,p\leqslant 0.05$ 时,近似效果就相当好了.

(4) **几何分布**(离散型):在独立试验序列中,若一次伯努利试验中某事件 A 发生的概率为 $P(A)=p(0<p<1)$,只要事件 A 不发生,试验就不断地重复下去,直到事件 A 发生试验才停止.设随机变量 X 为直到事件 A 发生为止所需的试验次数,则 X 的概率分布为

$$P(X=k)=(1-p)^{k-1}p\quad(k=1,2,\cdots),\tag{1.2.19}$$

则称随机变量 X 服从以 p 为参数的几何分布,记为 $X\sim G(p)$.

(5) **超几何分布**(离散型):设 N 个元素分为两类,有 M 个属于第一类,$N-M$ 个属于第二类.现在从中不重复抽取 n 个,其中包含的第一类元素的个数 X 的分布律为

$$P(X=k)=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}\quad(k=0,1,\cdots,l),\tag{1.2.20}$$

其中 $n\leqslant N,M\leqslant N,l=\min\{n,M\}$,$n,N,M$ 均为正整数,则称随机变量 X 服从参数为 N,M,n 的超几何分布,记为 $X\sim H(N,M,n)$.

(6) **均匀分布**(连续型):如果随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a\leqslant x\leqslant b,\\ 0, & 其他,\end{cases}\tag{1.2.21}$$

则称 X 服从区间 $[a,b]$ 上的均匀分布,记为 $X\sim U[a,b]$,其相应的分布函数为

$$F(x)=\begin{cases}0, & x<a,\\ \dfrac{x-a}{b-a}, & a\leqslant x<b,\\ 1, & x\geqslant b.\end{cases}\tag{1.2.22}$$

(7) **指数分布**(连续型):如果随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}\tag{1.2.23}$$