普通高等教育农业农村部“十三五”规划教材
全国高等农林院校“十三五”规划教材
中华农业科教基金教材建设研究项目 NKJ201502030

概　率　论

第二版

吴清太　方桂英　主编

中国农业出版社

图书在版编目（CIP）数据

概率论 / 吴清太，方桂英主编 . —2 版 . —北京：中国农业出版社，2016. 12（2019. 1 重印）

全国高等农林院校“十三五”规划教材

ISBN 978 - 7 - 109 - 22377 - 6

Ⅰ. ①概… Ⅱ. ①吴… ②方… Ⅲ. ①概率论-高等学校-教材 Ⅳ. ①O211

中国版本图书馆 CIP 数据核字（2016）第 278080 号

中国农业出版社出版

（北京市朝阳区麦子店街 18 号楼）

（邮政编码 100125）

策划编辑 魏明龙

文字编辑 魏明龙

北京万友印刷有限公司印刷 新华书店北京发行所发行

2011 年 7 月第 1 版 2016 年 12 月第 2 版

2019 年 1 月第 2 版北京第 3 次印刷

开本：720mm×960mm 1/16 印张：11

字数：190 千字

定价：22. 00 元

内 容 简 介

本教材为高等农林院校概率论课程教材．全书共有5章：随机事件及其概率、一维随机变量及其分布、多维随机变量及其分布、随机变量的数字特征、大数定律及中心极限定理．附录中还有常用的MATLAB概率统计软件的简介．

本教材是编者经多年教学实践及研究，在不断总结经验的基础上编写而成的．注重随机数学的基本方法及基本思想的渗透，而淡化随机数学理论上的证明与技巧．加强应用随机数学的手段与方法对实际问题的处理能力的培养，达到提高学生对随机问题的认识及解决能力．

本教材可作为高等农林院校概率论课程教学用书，以及相关科技人员的参考书．

编 写 人 员

主　编　吴清太（南京农业大学）

方桂英（江西农业大学）

副主编　陈朝霞（南京农业大学）

胡建根（江西农业大学）

参　编　温阳俊（南京农业大学）

朱烈浪（江西农业大学）

张　梅（南京农业大学）

第二版前言

本教材第一版发行以来各方面反映尚好，同行提出了一些有益的意见和建议，我们在教学中也发现了一些可改进的地方．本次修订的重点放在概念的引入、结论的叙述和解释上，因此在第一版的基础上，修改或补充了书中概念引入和结论解释及其应用的一些例题；同时根据教学的需要补充了求连续型随机变量的函数的方法：微元法；补充了两个相互独立的随机变量的函数的分布，其中一个为连续型，另一个为离散型；增加了三套模拟试题及其答案．

本教材是全国高等农林院校“十三五”规划教材，同时被列入中华农业科教基金教材建设研究项目，在此向关心和支持本教材出版工作的广大教师表示衷心的感谢．

本次修订由吴清太定稿．

由于作者水平有限，书中难免存在不妥之处，恳请读者和使用本教材的教师及专家批评指正．

主　编

2016 年 6 月于南京

注：本教材于 2017 年 12 月 4 日被评为普通高等教育农业农村部“十三五”规划教材（农科（教育）函〔2017〕第 379 号）。

第一版前言

概率是描述随机事件发生的可能性的度量．概率论通过对简单事件的研究，逐步进入复杂随机现象规律的研究，是研究复杂随机现象规律的有效方法和工具．掌握处理随机问题的基本思想与方法是当代大学生必须具备的一种基本能力．

本教材是在教育部高等农业院校理科基础课程教学指导委员会领导下，针对农林院校人才培养目标而为农林院校开设概率论课程编写的．考虑到农林院校学生的特点及培养要求，在编写过程中，遵循教学指导委员会下发的关于本课程的基本要求，同时根据教学改革的趋势，在内容上突出随机数学的基本思想和方法，淡化各种繁琐的技巧，适当降低了难度．考虑到数学建模与数学实验的需要，在附录中增加了目前在工程技术和科学研究中广泛使用的MATLAB软件的介绍．书中列举了大量较为典型、易于接受的例题和实际应用问题，配备了相当数量的习题：习题 A 和习题 B，习题 A 为基本要求题，即一般学生必须掌握的习题，而习题 B 为较高要求题，即立志于要考研究生的学生需要掌握的习题．

本教材由吴清太、方桂英担任主编，陈朝霞、胡建根担任副主编，全书由吴清太统一制定编写教学大纲并统一定稿．参加编写的人员还有温阳俊、朱烈浪、张梅．第 1 章附录、附表 1、附表 2 由吴清太编写，第 2 章 和第 3 章部分图形由吴清太绘制；第 2 章、第 3 章由方桂英、胡建根、朱烈浪编写；第 4 章由陈朝霞编写；第 5 章由温阳俊编写；张梅对第 1 章、第 4 章、第 5 章进行校对．另外，翁利亚、周春俊对全书进行了校对．

本教材是编者在多年从事随机数学的教学，总结经验的基础上编写而成的．由于作者水平有限，书中难免存在不妥之处，请读者不吝指教．

编　者

2011年3月

目　　录

第 1 章　随机事件及其概率

概率论是研究随机现象及其规律的一门学科，是近代数学的重要组成部分. 本章将介绍概率论的基本概念，并进一步讨论事件之间的关系及其运算、概率的定义、概率的确定方法、概率的性质及计算方法等，这些都是我们学习概率论的基础.

1.1　随机事件及其运算

1.1.1　随机现象

人们在生产活动、社会实践和科学试验中，所遇到的自然现象和社会现象，大体分为两类：一类是可以预言其结果，即在保持条件不变的情况下，重复进行试验，其结果总是确定的. 例如，在标准大气压下水加热到 100℃必然沸腾；纯种紫花豌豆的后代一定开紫花；水稻从播种到收割总是经过发芽、育种、长叶、吐穗、扬花、结实几个阶段等，我们称之为**确定性现象**；另一类现象是无法事先断言其结果的，即使在保持条件不变的情况下重复进行试验，其结果也未必相同. 例如，观察某一商店每天来的顾客数与销售商品的数额都不是确定的；观察种子发芽的情况，某粒种子可能发芽，也可能不发芽；某个射手向一目标射击，结果是可能命中，也可能不中；正在放射 α 粒子的放射性物质，每天在同一规定的时间内放射的粒子数，事先无法确定. 我们称这类现象为**随机现象**.

随机现象是广泛存在的. 我国宋代大文学家苏轼有著名的诗句：“人有悲欢离合，月有阴晴圆缺，此事古难全.”这说明人类早就对随机现象的存在有着切身的体验，也记录了人们面对随机现象曾经表现出来的无能为力.

1.1.2　样本空间和随机事件

1. 随机试验

尽管随机现象中出现什么结果不能完全预言，但全部可能结果是已知的. 例如，抛一枚硬币只会有“正面”和“背面”两个可能结果，某一商店每天来的顾客数必定是一个非负整数. 为了叙述方便，我们把对一定条件下的自然现

象和社会现象所进行的观察或实验统称为试验.

若试验具有下列共同特征:

(1) 在相同的条件下可以重复进行;

(2) 每次试验的可能结果不止一个,但是在试验之前可以确定一切可能出现的结果;

(3) 每次试验中有且只有其中的一个结果发生,但试验之前不能准确地预知哪一种结果会出现,

我们称这样的试验为**随机试验**,也简称为**试验**,常用 E 来表示.

为了研究的方便,我们有时也会把具有固定结果的试验,看成是随机试验的极端情形.有时,又需要把几次试验作为一个整体合起来看成一次随机试验,例如,可以把连续掷三次骰子看成是一次随机试验.

在一次试验中,某个结果是否出现具有一定的偶然性.但在多次重复试验中,这些无法准确预知的现象,并不是杂乱无章的,其结果就会出现某种固有规律性.例如,在投掷一枚质地均匀的硬币时,只投掷一次时,投掷的结果是正面还是反面是无法确定的,但当大量重复投掷硬币时,就可以看到出现正面的次数约占总试验次数的一半.又如,某人打靶射击,若射击次数不多,靶上的弹着点似乎是随意分布的,但倘若进行大量的重复射击时,弹着点的分布就逐渐呈现规律性:它们大体上关于靶中心对称,靠近靶心的弹着点密,偏离靶心越远弹着点越稀少,且弹着点落在靶任意指定区域内的次数与射击次数 n 之比(频率)大体上保持稳定,且 n 越大,其频率稳定性就愈加明显,这种在大量重复试验中随机现象所表现出的固有规律,我们通常称之为统计规律.概率论就是揭示和研究随机现象统计规律的一门数学学科.概率论的理论和方法在物理学、医学、生物学等学科以及农业、工业、国防和国民经济等方面都具有极广泛的应用.

2. 样本空间

随机试验的每一个可能发生的结果,称为**样本点**,有时也称为**基本事件**,常用 ω 表示.而所有样本点(或基本事件)组成的集合称为**样本空间**,通常用 Ω 表示.显然 $\omega\in\Omega$.

例 1 观察一粒种子的发芽情况,一次观察就是一次试验,试验的结果为 $\omega_1=$ "发芽",$\omega_2=$ "不发芽",则 $\Omega_1=\{\omega_1, \omega_2\}$.

例 2 服务台在 8:00～9:00 之间到来的顾客数,则 $\Omega_2=\{n\mid n=0, 1, 2, \cdots\}$,可见 Ω_2 包含可列无穷多个样本点.

例 3 在有噪声干扰下,测量某终端电压,则 $\Omega_3=\{x\mid x\in\mathbf{R}\}$,其样本点有不可数无穷多个.

注：在同样的试验条件下，由于试验的考察目的不同，可能选择不同的样本空间，这是初学者必须注意的．

例 4　一枚硬币投掷两次观察出现正面的次数(注意此时投掷两次硬币才算完成一次试验)，$\Omega_4=\{0, 1, 2\}$，其中 0，1，2 分别表示出现正面的次数，共有 3 个样本点．

例 5　一枚硬币投掷两次观察出现正、反面的次序，则 $\Omega_5=\{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}, \mathrm{TT}\}$(其中，H 代表出现正面，T 代表出现反面)，包含 4 个样本点．

需要注意的是：

(1) 样本空间中的元素可以是数也可以不是数．

(2) 从样本空间所含样本点的个数来看，样本空间可以分为有限与无限两类．例如，以上样本空间中 Ω_1，Ω_4，Ω_5 含有有限个样本点，故为有限样本空间，而 Ω_2，Ω_3 中样本点的个数为无限个，为无限样本空间．

3. 随机事件

对于某个随机试验来说，在一次试验中可能出现也可能不出现的结果，这种由部分样本点组成的试验结果称为**随机事件**．而从集合论的观点来说，随机试验 E 的样本空间 Ω 的任一子集就是**随机事件**，简称**事件**，常用大写字母 A，B，C，…表示．在试验中，如果出现某一个样本点 ω，且 $\omega\in A$，则称本次试验事件 A 发生了，反之，若 $\omega\notin A$，则称本次试验事件 A 没有发生．

因为样本空间 Ω 是由所有样本点所组成的，因而在任一次试验中，必然要出现 Ω 中的某一样本点 ω. 也就是在试验中，Ω 必然会发生，所以今后用 Ω 来表示一个**必然事件**．又因为空集 $\varnothing$ 也可以看作是 Ω 的子集，且它不包含任何基本事件，故每次试验中 $\varnothing$ 必定不会发生，故我们称 $\varnothing$ 为**不可能事件**或**空事件**.

例 6　同时抛三枚硬币，顺次记录出现正反面的情况，得样本空间为

$\Omega=\{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{THH}, \mathrm{HTT}, \mathrm{THT}, \mathrm{TTH}, \mathrm{TTT}\}$.

$A=$“反面恰好出现两次”$=\{\mathrm{TTH}, \mathrm{THT}, \mathrm{HTT}\}$，

$B=$“反面至少出现两次”$=\{\mathrm{TTT}, \mathrm{TTH}, \mathrm{THT}, \mathrm{HTT}\}$，

$C=$“正反面出现的次数不相同”$=\Omega$，

$D=$“反面出现的次数多于 3 次”$=\varnothing$.

例 7　投掷一颗骰子，观察其出现的点数，若以“i”表示“掷出 i 点”($i=1, 2, 3, 4, 5, 6$)，则样本空间为 $\Omega=\{1, 2, 3, 4, 5, 6\}$.

$A=$“掷出的点数为偶数点”$=\{2, 4, 6\}$，

$B=$“掷出的点数大于 3”$=\{4, 5, 6\}$，

$C=$“掷出点数为 7 点”$=\varnothing$，$D=$“掷出的点数不超过 6 点”$=\Omega$.

1.1.3　事件之间的关系和运算

一个样本空间 Ω 中，可以有很多的随机事件，由于它们共处于同一个试验中，因而彼此之间是有联系的．我们有必要弄清它们之间的关系，并引进事件间的运算，以便化复杂事件为简单事件，更好地解决相应的概率问题．

1. 事件的包含

设 A，B 是任意两个随机事件，如果事件 A 发生，则必然导致事件 B 发生，也即事件 A 中的任一样本点都属于事件 B，则称事件 A 包含于事件 B(或称事件 B 包含事件 A)，记作 $A\subseteq B$ 或 $B\supseteq A$，它的几何表示如图 1-1 所示．如前面我们提到的例 6 中，有 $A\subseteq B$. 因为不可能事件 $\varnothing$ 不包含任何样本点 ω，故对任一事件 A，我们约定 $\varnothing\subseteq A$.

2. 事件的相等

设 A，B 是两事件，若 $A\subseteq B$ 且 $B\subseteq A$，则称 A 与 B 相等，记作 $A=B$.

3. 事件的并(或事件的和)

事件 A 和 B 中至少有一个发生的事件，称为事件 A 与 B 的和事件，记作 $A\cup B$(或 $A+B$). 从集合的观点来看，事件 A 与 B 的和是由事件 A 与事件 B 中所有样本点组成的，它的几何表示如图 1-2 所示．

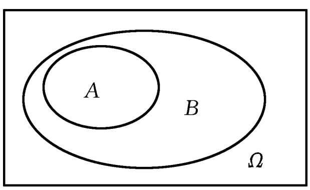


图 1-1

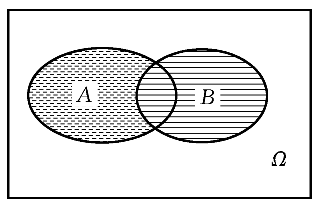


图 1-2

如在掷骰子的试验中，记事件 $A=$“出现奇数点”$=\{1, 3, 5\}$，事件 $B=$“出现的点数不超过 3”$=\{1, 2, 3\}$，则 A 与 B 的和为 $A\cup B=\{1, 2, 3, 5\}$.

事件的和可以推广到更多个事件上去，n 个事件 A_1，A_2，…，A_n 中至少有一个发生的事件，称为事件 A_1，A_2，…，A_n 的和事件，记作 $A_1\cup A_2\cup\cdots\cup A_n=\bigcup\limits_{i=1}^{n}A_i$(或 $A_1+A_2+\cdots+A_n=\sum\limits_{i=1}^{n}A_i$).

事件的和还可以推广到无穷多个事件上去，设 A_1，A_2，…，A_n，…为可列个事件，我们把 A_1，A_2，…，A_n，…中至少有一个发生的事件，称为事件

列 A_1，A_2，…，A_n，…的和事件，记作 $A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots = \bigcup_{i=1}^{\infty} A_i$（或 $A_1 + A_2 + \cdots + A_n + \cdots = \sum_{i=1}^{\infty} A_i$）.

4. 事件的交（积）

事件 A 与 B 同时发生的事件，称为事件 A 和事件 B 的交或积事件，记作 $A \cap B$（或 AB）. 从集合的观点来看，事件 A 和 B 的积是由既属于 A 又属于 B 的样本点组成的集合，它的几何表示如图 1－3 所示 .

如在掷骰子的试验中，记事件 A＝“出现奇数点”＝$\{1, 3, 5\}$，记事件 B＝“出现的点数不超过 3”＝$\{1, 2, 3\}$，则 A 与 B 的交为 $A \cap B = \{1, 3\}$.

积事件可以推广到更多个事件上去，n 个事件 A_1，A_2，…，A_n 同时发生的事件称为事件 A_1，A_2，…，A_n 的积事件，记作 $A_1 \cap A_2 \cap \cdots \cap A_n = \bigcap_{i=1}^{n} A_i$.

事件的积也可以推广到无穷多个事件上去，设 A_1，A_2，…，A_n，…为可列个事件，我们把 A_1，A_2，…，A_n，…同时发生的事件，称为事件列 A_1，A_2，…，A_n，…的积事件，记作 $A_1 \cap A_2 \cap \cdots \cap A_n \cap \cdots = \bigcap_{i=1}^{\infty} A_i$.

5. 互不相容事件（或互斥事件）

若事件 A 与事件 B 不能同时发生，即 $AB = \varnothing$，则称事件 A 与事件 B **互不相容**（或称**互斥**）. 它的几何表示如图 1－4 所示 .

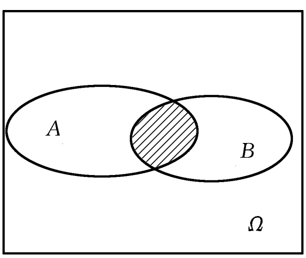


图 1－3

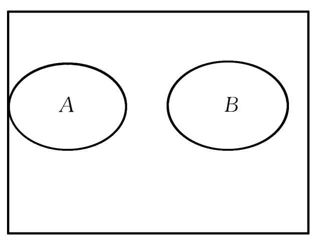


图 1－4

6. 对立事件（或互逆事件）

A 不发生的事件，称为事件 A 的对立事件，亦称为 A 的逆事件，记作 $\overline{A}$. 从集合论的观点来看，A 的对立事件就是由 Ω 中不属于 A 的样本点所组成的集合，即 $\overline{A} = \Omega \setminus A$. 它的几何表示如图 1－5 所示 . 显然：$A\overline{A} = \varnothing$，$A \cup \overline{A} = \Omega$，$\overline{\overline{A}} = A$.

因为 $A\overline{A} = \varnothing$，所以互逆事件一定是互不相容事件，但反之不成立 .

7. 事件的差

事件 A 发生而事件 B 不发生的事件，称为事件 A 与事件 B 的差，记作 $A\setminus B$. 从集合的观点来看，$A\setminus B$ 是由属于 A 但不属于 B 的所有样本点组成的集合 . 它的几何表示如图 1－6 所示 .

例如，在掷骰子的试验中，记事件 A＝“出现奇数点”＝{1，3，5}，记事件 B＝“出现的点数不超过 3”＝{1，2，3}，则 A 与 B 的差为 $A\setminus B=\{5\}$.

由以上的定义可知，我们可以把对事件的分析转化为对集合的分析，利用集合间的运算关系来分析事件之间的关系 . 但是我们要注意学会用概率的语言来描述各种事件，并会用这些运算关系来表示一些事件 .

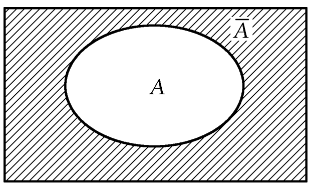


图 1－5

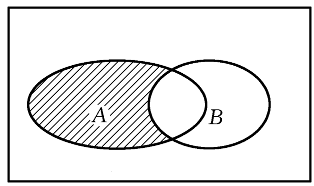


图 1－6

8. 事件的运算性质

（1）**交换律**：

$$A\cup B=B\cup A,\quad A\cap B=B\cap A. \tag{1.1.1}$$

（2）**结合律**：

$$A\cup(B\cup C)=(A\cup B)\cup C,\quad A\cap(B\cap C)=(A\cap B)\cap C. \tag{1.1.2}$$

（3）**分配律**：

$$(A\cup B)\cap C=(A\cap C)\cup(B\cap C), \tag{1.1.3}$$

$$(A\cap B)\cup C=(A\cup C)\cap(B\cup C). \tag{1.1.4}$$

（4）**对偶律（德・摩根公式）**：

$$\overline{A\cup B}=\overline{A}\cap\overline{B},\quad \overline{A\cap B}=\overline{A}\cup\overline{B}. \tag{1.1.5}$$

德・摩根公式可以推广到多个事件的场合：

$$\overline{\bigcup_{i=1}^{n}A_i}=\bigcap_{i=1}^{n}\overline{A_i},\quad \overline{\bigcap_{i=1}^{n}A_i}=\bigcup_{i=1}^{n}\overline{A_i}. \tag{1.1.6}$$

此外，事件的运算还满足以下一些常用的规律性：

$A\cap A=A$，$A\cup A=A$，$A\cap\Omega=A$，$A\cup\varnothing=A$，$A\cap\varnothing=\varnothing$，$A\cup\Omega=\Omega$，$A\setminus A=\varnothing$，$A\setminus B=A\setminus AB=A\overline{B}$，$A\cap B\subseteq A$，$A\subseteq A\cup B$.

例 8　在例 7 中有

$A\cup B=\{2，4，5，6\}$，$A\cap B=\{4，6\}$，$A\setminus B=\{2\}$，$\overline{A\cup B}=\{1，3\}$.

例 9　设 A，B，C 是同一随机试验下的三个任意的事件，试用 A，B，C 的运算式子表示下列各事件．

（1）三个事件恰好有两个发生：$AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC$.

（2）三个事件至少发生一个：$A\cup B\cup C$.

（3）三个事件中至少发生两个：$AB\cup BC\cup AC$.

（4）A 与 B 发生，C 不发生：$AB\overline{C}$ 或 $AB\setminus C$.

（5）A，B，C 都不发生：$\overline{A}\overline{B}\overline{C}$ 或$\overline{A\cup B\cup C}$.

（6）A，B，C 至多发生一个：$\overline{A}\overline{B}\overline{C}\cup A\overline{B}\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\overline{B}C$.

1.2　随机事件的概率定义及其确定方法

在这一节，我们要给出概率的定义及其确定方法，此问题是概率中最基本的问题．简单地讲，概率是随机事件发生的可能性的大小．为此我们先看下面一些经验事实：

1. 随机事件的发生带有偶然性，但随机事件发生的可能性是有大小之分的．例如，口袋中有 5 个围棋子，其中 4 个是黑围棋子，1 个是白围棋子，在口袋中任意取出 1 个棋子，人们的共识是：取出黑棋子的可能性比取出白棋子的可能性大．

2. 随机事件发生的可能性的大小是可以设法度量的，就好比一根木棒有长度一样．例如，抛一枚硬币，出现正面和反面的可能性是相同的，各为 0.5. 足球裁判就用抛硬币的方法让双方的队长选择场地，以示机会均等．

3. 在日常生活中，人们对一些随机事件发生的可能性的大小往往用百分比进行度量．例如，在抽样检验中，任意抽取一个产品可能为合格品，也可能为次品，但产品质量的好坏可以用次品率来度量；购买彩券后可能中奖，也可能不中奖，但中奖的可能性大小可以用中奖率来度量．

在概率论的发展历史上，曾有过概率的古典定义、概率的几何定义和概率的统计定义．这些定义各适合一类随机现象．那么如何给出适合一切随机现象的概率的最一般的定义呢？1900 年数学家希尔伯特提出要建立概率的公理化定义以解决这个问题，即从最少的几条本质特性出发去刻画概率的概念．1933 年前苏联数学家柯尔莫哥洛夫首次提出了概率的公理化定义，这个定义既概括了历史上几种概率定义中的共同特性，又避免了各自的局限性和含混之处，不管什么随机现象，只有满足定义中的三条公理，才能说它是概率．这一公理化

体系迅速获得举世公认，是概率论发展史上的一个里程碑．有了这个公理化定义后，概率论得到了很快的发展．

对于随机事件 A，在一次试验中是否发生具有不确定性，但是在多次重复试验中它的发生却能呈现出一定的规律性，即它出现可能性的大小是可以度量的，我们把度量事件发生可能性大小的数值称为随机事件 A 发生的概率，这一节将从不同角度给出随机事件概率的定义．

1.2.1　概率的公理化定义

定义 1　设 E 是一随机试验，Ω 是 E 的样本空间，对于试验 E 的任意一个随机事件 A，都赋予一个实数 $P(A)$，若集合函数 $P(A)$满足下面三条公理：

（1）**非负性**：对于任一随机事件 A，都有 $P(A)\geqslant 0$；

（2）**规范性**：$P(\Omega)=1$；

（3）**可列可加性**：设事件 A_1，A_2，…，A_n，…两两互不相容，则

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

则称 $P(A)$为事件 A 的概率．

概率的公理化定义没有告诉人们如何去确定概率．历史上在概率的公理化定义出现之前，概率的频率定义、古典定义、几何定义都在一定的场合下，有着各自确定概率的方法，所以在有了概率的公理化定义之后，把它们看作确定概率的方法是恰当的．下面分别讲述确定概率的方法．

1.2.2　随机事件的频率及确定概率的统计方法

定义 2　设在 n 次随机试验中事件 A 发生了 m 次，则称 m 为事件 A 发生的频数，而称

$$f_n(A)=\frac{m}{n} \tag{1.2.1}$$

为 n 次随机试验中事件 A 发生的**频率**．

易知频率具有下列性质：

（1）非负性：即 $0\leqslant f_n(A)\leqslant 1$；

（2）规范性：即若 Ω 是必然事件，则 $f_n(\Omega)=1$；

（3）有限可加性：即若 A，B 互不相容，则 $f_n(A\cup B)=f_n(A)+f_n(B)$.

为了研究在抛掷一枚均匀硬币“出现正面”这一事件发生的规律，历史上一些著名的科学家曾做了大量试验，其部分试验结果见表 1－1.

从表 1－1 可以看出，出现正面的频率虽然随着 n 的不同而不同，但却都

在 0.5 这个数值附近摆动，n 越大，频率在 0.5 附近波动越小．因此，我们可以说，在抛硬币的试验中频率出现的总趋势是随着试验次数的增多而逐渐趋于稳定值 0.5，我们就称 0.5 为投一枚均匀的硬币出现正面的概率．频率的概念简单而直观，容易掌握，因此可以根据频率的基本性质加以提炼与概括，作为确定事件概率的依据和直观背景．

表 1－1

试验者	抛掷次数 n	正面向上的次数 m	正面向上的频率
德·摩根	2408	1061	0.4406
蒲丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

定义 3　在相同的条件下，重复进行 n 次试验，如果随着试验次数的增大，事件 A 出现的频率 $f_n(A)$稳定地在某一常数 p 附近摆动，则称常数 p 为事件 A 发生的概率，记作

$$P(A)=p.$$

这个定义称为概率的**统计定义**．从上面的两个定义可以看出，频率与概率是不相同的，频率随着试验的结果而变化，但是概率却是固定不变的．概率的统计定义有以下的基本性质：

(1) 对于任意给定的事件 A，有 $0\leqslant P(A)\leqslant 1$；

(2) $P(\Omega)=1$，$P(\varnothing)=0$；

(3) 对于 n 个两两互不相容的事件 A_1，…，A_n，有

$$P(A_1\cup A_2\cup\cdots\cup A_n)=P(A_1)+P(A_2)+\cdots+P(A_n).$$

由频率的性质及频率与概率的关系，容易理解上述性质的正确性．

1.2.3　确定概率的古典方法

人们在生活中最早研究的是一类简单的随机试验，比如，足球比赛中扔硬币挑边问题，围棋比赛中猜谁先下棋的问题．这类试验的共同特点是：

(1) 试验的样本空间只含有有限个样本点；

(2) 在一次试验中，每个样本点发生的可能性是相同的．

这两个特点分别称为有限性与等可能性．具备这两个特点的试验是大量存在的，它曾经是概率论中主要研究的对象，人们常把具备有限性与等可能性的试验模型称为古典概率模型，简称古典概型．

定义 4　设随机试验 E 为古典概型，其样本空间为 $\Omega=\{\omega_1，\cdots，\omega_n\}$，对

于任一事件 A，其概率定义为

$$P(A)=\frac{A\text{ 中包含的样本点数}}{\Omega\text{ 中包含的样本点数}}. \tag{1.2.2}$$

若记 $\mu(A)=\{$事件 A 中包含的样本点数$\}$，则 $P(A)=\dfrac{\mu(A)}{\mu(\Omega)}$.

由定义可知，概率 $P(A)$有以下性质：

(1) 非负性：$\forall A\subseteq\Omega$，$P(A)\geqslant 0$；

(2) 规一性：$P(\Omega)=1$；

(3) 设事件 A_1，A_2，…，A_m 互不相容，则 $P\left(\sum\limits_{i=1}^{m}A_i\right)=\sum\limits_{i=1}^{m}P(A_i)$；

(4) 对于每一样本点 $\omega_i(1\leqslant i\leqslant n)$，有

$$P(\{\omega_1\})=P(\{\omega_2\})=\cdots=P(\{\omega_n\})=\frac{1}{n}.$$

证明　只证明(4)，其余的留作课外练习．

设事件 $A_i=\{\omega_i\}$，$1\leqslant i\leqslant n$，则由定义知：$P(A_i)=P(\{\omega_i\})=\dfrac{1}{n}$，所以

$$P(\{\omega_1\})=P(\{\omega_2\})=\cdots=P(\{\omega_n\})=\frac{1}{n}.$$

事实上，“每一个样本点出现等可能”与式(1.2.2)等价，即事件 A 的概率 $P(A)$仅与 A 中所含的样本点数有关，而与它具体含有的是哪些样本点无关.

按照这种定义方式，由古典概率定义求事件的概率就要求出样本空间中所包含的样本点的个数与所求事件包含的样本点的个数．在确定概率的古典方法中大量使用排列与组合公式，下面我们介绍两条计数原理：加法原理与乘法原理.

(1) **加法原理**：如果某件事可由 k 类不同途径之一去完成，在第一类途径中有 m_1 种完成方法，在第二类途径中有 m_2 种完成方法，…，在第 k 类途径中有 m_k 种完成方法，那么完成这件事共有 $m_1+m_2+\cdots+m_k$ 种方法．

(2) **乘法原理**：如果某件事需经 k 个步骤去完成，做第一步有 m_1 种方法，做第二步有 m_2 种方法，…，做第 k 步有 m_k 种方法，那么完成这件事共有 $m_1\times m_2\times\cdots\times m_k$ 种方法．

例 1　从 1，2，3，4，5，6，7，8，9 中随机地取一个数，求这个数为素数的概率．

解　用 A 表示“从 1～9 中取出的数是素数”这一事件，一个样本点等价于从 1～9 这 9 个数中取一个数，所以样本空间的样本点数为 $\mu(\Omega)=9$，而 A 所含的样本点数为 $\mu(A)=4$，由概率的古典定义可得事件 A 发生的概率为

$$P(A)=\frac{\mu(A)}{\mu(\Omega)}=\frac{4}{9}.$$

例 2　有一批橡胶种子，共 2500 粒，其中属于品系 A 的有 250 粒，属于品系 B 的有 1000 粒，属于品系 C 的有 750 粒，属于品系 D 的有 500 粒．现从该批种子中随机抽取一粒，求抽到的种子为品系 A，B，C，D 的概率．

解　由于每粒种子抽到是等可能的，一个样本点就相当于从 2500 粒种子中抽取一粒，所以样本空间的样本点总数为 $\mu(\Omega)=2500$，我们就用 A，B，C，D 分别表示抽出的种子属于品系 A，B，C，D 的种子，则它们所含样本点数分别为 $\mu(A)=250$，$\mu(B)=1000$，$\mu(C)=750$，$\mu(D)=500$，所以抽到的种子分别属于品系 A，B，C，D 的概率为

$$P(A)=\frac{250}{2500}=\frac{1}{10},\ P(B)=\frac{1000}{2500}=\frac{2}{5},\ P(C)=\frac{750}{2500}=\frac{3}{10},\ P(D)=\frac{500}{2500}=\frac{1}{5}.$$

例 3　设有编号分别为 1，2，3 的三个盒子，今将一个红色、一个白色的球放入这三个盒子中，设每个球放入任何一个盒是等可能的，记 $A=\{$编号为 3 的盒子为空盒$\}$，求 $P(A)$.

解　把两个球放进三个盒子中，每个球有 3 种等可能的放法，而两个球的放法互不影响，因此有 9 个可能结果，而事件 A 出现的可能结果为每个球可以等可能地放到编号为 1，2 的盒子中去，即 A 包含 4 种可能的放法，故事件 A 出现的可能性为$\frac{4}{9}$，即

$$P(A)=\frac{4}{9}.$$

例 4　袋中有 10 个球，其中 6 个白球，4 个红球．假设每个球被取出是等可能的，现随机地取出 3 个，试求取出 k 个红球的概率$(0\leqslant k\leqslant 3)$.

解　将球编号，记 6 个白球依次为 1～6，4 个红球为 7～10. 假设每球被取出等可能，这等价于从 10 个号码中任取 3 个，不计次序的每一种取法是等可能的．因而若每次随机地取出 3 个，只观察其号码，不计其次序，则一个样本点等价于从 10 个号码中取其 3，不计次序的一种组合．故样本点数 $\mu(\Omega)=C_{10}^{3}$. 记 $A_k=\{$取 3 个，取出 k 个红球$\}(0\leqslant k\leqslant 3)$. 显然，$A_k$ 发生，当且仅当 3 个球中，需从 7～10 号中取出 $k(0\leqslant k\leqslant 3)$个，且从 1～6 号中取出$(3-k)$个．由乘法原理有：$\mu(A_k)=C_4^k C_6^{3-k}(0\leqslant k\leqslant 3)$，得

$$P(A_k)=\frac{C_4^k C_6^{3-k}}{C_{10}^3}.$$

例 5　一口袋装有 6 个苹果，其中 4 个红元帅，2 个青香蕉．从袋中取苹果两次，每次随机地取一个．考虑两种抽取方式：

（1）**放回抽样**：第一次取一个苹果，观察其种类后放回袋中，搅匀后再取一个苹果．

（2）**不放回抽样**：第一次取出一个苹果后不放回袋中，第二次从剩余的苹果中再取一个．

分别就上面两种方式求：

（1）取到的两个苹果都是红元帅的概率；

（2）取到的两个苹果是相同品种（即两个都为红元帅或都为青香蕉）的概率.

解　从袋中取两个苹果，每一种取法就是一个样本点．设 A=“取到的两个苹果都是红元帅”，B=“取到的两个苹果都为红元帅或青香蕉”．

有放回抽取

$$P(A)=\frac{4^2}{6^2}=\frac{4}{9},\quad P(B)=\frac{4^2+2^2}{6^2}=\frac{5}{9}.$$

无放回抽取

$$P(A)=\frac{C_4^2}{C_6^2}=0.4,\quad P(B)=\frac{C_4^2+C_2^2}{C_6^2}=\frac{7}{15}.$$

*1.2.4　确定概率的几何方法

在古典概率模型中，试验的可能结果是有限的，这就具有非常大的限制．在概率论发展的早期，人们当然就要竭力突破这个限制，尽量扩大自己的研究范围．一般情况下，当试验结果为无限时，会出现一些本质性的困难，使问题的解决变得非常困难．这里我们讨论一种简单的情况，即具有某种“等可能性”的问题．

我们在一个面积为 S_Ω 的区域 Ω 中，等可能地任意投点，如图 1－7 所示．这里等可能的含义是：投向区域 Ω 的任意点的机会都是均等的．也就是：若在区域 Ω 中有任意一个小区域 A，如果它的面积为 S_A，则点落入区域 A 的可能性大小与 S_A 成正比，而与 A 的位置及形状无关．若仍记 A 为“点落入小区域 A”这个事件，由 $P(\Omega)=1$ 可得

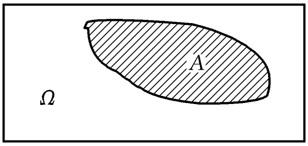


图 1－7

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

这一类概率通常称为**几何概率**．需要注意的是：如果是在一条线段上投点，那么面积应改为长度；如果在一个立方体内投点，则面积应改为体积．几何概率是概率论中所考虑的一种，它满足概率公理化定义的几个性质．

几何概率在现实生活中有许多应用，下面是个实际应用的例子．

例 6(**约会问题**)　甲、乙两人相约在 0 到 T 这段时间内在约定地点会面，先到的人等候另一人，经过时间 $t(t<T)$后离去，设每人在 0 到 T 内到达的时刻是等可能的，且两个人到达的时刻互不牵连，求甲、乙两人能会面的概率．

解　设 x，y 分别表示甲、乙两人到达的时刻，事件 A 表示“两人能会面”，$(x,\ y)$表示平面上的点(图 1－8)，则

$$\Omega=\{(x,\ y)\mid 0\leqslant x\leqslant T,\ 0\leqslant y\leqslant T\},$$

$$A=\{(x,\ y)\mid |x-y|\leqslant t\},$$

所以
$$P(A)=\frac{T^2-(T-t)^2}{T^2}=1-\left(1-\frac{t}{T}\right)^2.$$

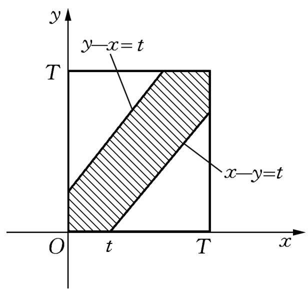


图 1－8

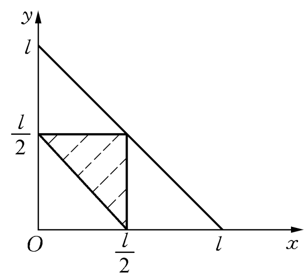


图 1－9

例 7　将一根长为 l 的木棒，任意折成三段，求恰好能构成一个三角形的概率．

解　设折得三段长度分别为 x，y 和 $l-x-y$，那么，样本空间

$$\Omega=\{(x,\ y)\mid 0<x,\ y<l,\ 0<x+y<l\},$$

而随机事件 A：“三段构成三角形”相应的子区域 D 应满足“两边之和大于第三边”的原则，从而得到联立方程组：

$$\begin{cases} l-x-y<x+y, \\ x<(l-x-y)+y, \\ y<(l-x-y)+x, \end{cases}$$

如图 1－9 所示，

$$D=\left\{(x,\ y)\,\middle|\, 0<x<\frac{l}{2},\ 0<y<\frac{l}{2},\ \frac{l}{2}<x+y<l\right\}.$$

由几何概率的定义可得

$$P(A)=\frac{\text{平面区域 } D \text{ 的面积}}{\text{平面区域 } \Omega \text{ 的面积}}=\frac{\frac{1}{2}\times\frac{l}{2}\times\frac{l}{2}}{\frac{1}{2}\times l\times l}=\frac{1}{4}.$$

1.3　概率的性质

根据概率的三条公理，容易得到如下概率的基本性质．

首先，在概率的规范性中说明必然事件 Ω 的概率为 1，那么可想而知，不可能事件 $\varnothing$ 的概率为 0，下面性质正说明这一点．

性质 1　$P(\varnothing)=0$.

证　因为 $\Omega=\Omega\cup\varnothing\cup\varnothing\cup\cdots$，所以

$$\begin{aligned}1&=P(\Omega)=P(\Omega\cup\varnothing\cup\varnothing\cup\cdots)\\&=P(\Omega)+P(\varnothing)+P(\varnothing)+\cdots\\&=1+\lim_{n\to\infty}nP(\varnothing),\end{aligned}$$

从而得到

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

1.3.1　概率的可加性

性质 2　若 A_1，A_2，$\cdots$，A_n 是一组两两互不相容的事件，则

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

证　对 A_1，A_2，$\cdots$，A_n，$\varnothing$，$\varnothing$，$\cdots$应用可列可加性，得

$$\begin{aligned}P\left(\bigcup_{i=1}^{n}A_i\right)&=P(A_1\cup A_2\cup\cdots\cup A_n\cup\varnothing\cup\varnothing\cup\cdots)\\&=P(A_1)+P(A_2)+\cdots+P(A_n)+P(\varnothing)+P(\varnothing)+\cdots,\end{aligned}$$

再由 $P(\varnothing)=0$，得

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

由有限可加性，我们就可以得到下面求对立事件的概率的公式．

性质 3　对于任一事件 A，恒有

$$P(\overline{A})=1-P(A).$$

证明　因为事件 A 与 $\overline{A}$ 互不相容，且 $\Omega=A\cup\overline{A}$. 由概率的规范性与有限可加性可得 $1=P(A)+P(\overline{A})$，由此可得 $P(\overline{A})=1-P(A)$.

有些事件直接考虑较为复杂，而考虑其对立事件则相对简单．对此类问题可以利用性质 3，见下面例子．

例 1　从 5 双不同的鞋子中任意取 4 只，4 只鞋子中至少有 2 只配成一双的概率是多少？

解　设以 A 表示事件“4 只中至少有 2 只配对成双”，则 $P(A)=1-$

$P(\overline{A})$，则 A 的对立事件 $\overline{A}$ 为“4 只鞋子中没有 2 只成双”．现在来求 $\overline{A}$ 中的样本点数：先从 5 双不同的鞋中任取 4 双，并且每双任取一只的不同取法的个数，它共有 $C_5^4\times 2^4$ 种取法，其中 2^4 来自每双取一只可有 2 种取法．而样本空间样本点总数为 10 只中任取 4 只的组合数．所以

$$P(A)=1-P(\overline{A})=1-\frac{C_5^4\times 2^4}{C_{10}^4}=\frac{13}{21}.$$

例 2（**生日问题**）　某大学生宿舍住了 6 个同学，求这 6 个同学中至少有两个同学的生日在同一个月份的概率．

解　认为每个人的生日等可能地出现在 12 个月中的每个月，则样本空间样本点总数为 $\mu(\Omega)=12^6$．用 $\overline{A}$ 来表示 6 个同学的生日月份互不相同，则 $\overline{A}$ 所含的样本点数为 $\mu(\overline{A})=6!C_{12}^6$．要求的概率为

$$P(A)=1-P(\overline{A})=1-\frac{6!C_{12}^6}{12^6}.$$

1.3.2　概率的单调性

性质 4　若 $B\subseteq A$，则 $P(A\setminus B)=P(A)-P(B)$．

证　因为当 $B\subseteq A$ 时，有

$$A=B\cup(A\setminus B),$$

且

$$B\cap(A\setminus B)=\varnothing,$$

由有限可加性有

$$P(A)=P(B)+P(A\setminus B).$$

移项后即得要证的等式．

利用性质 4 我们可以求一些较为复杂的事件的概率．

例 3　口袋中有编号为 1，2，…，n 的 n 只球，从中有放回地任取 m 次，求取出的 m 只球的最大号码为 k 的概率．

解　记事件 A_k 为“取出的 m 只球的最大号码为 k”，如果直接考虑事件 A_k，则比较复杂，因为“最大号码为 k”可以包括取到 1 次 k，取到 2 次 k，…，取到 m 次 k．为此我们记事件 B_i 为“取出的 m 只球的最大号码小于等于 i”，$i=1，2，…，n$，则 B_i 发生只需每次从 1，2，…，i 号球中取出即可，所以由古典概率知

$$P(B_i)=\frac{i^m}{n^m},\quad i=1,2,\cdots,n.$$

又因为 $A_k=B_k\setminus B_{k-1}$，且 $B_{k-1}\subseteq B_k$，由性质 4 得

$$\begin{aligned}P(A_k)&=P(B_k\setminus B_{k-1})=P(B_k)-P(B_{k-1})\\&=\frac{k^m-(k-1)^m}{n^m},\quad k=1,2,\cdots,m.\end{aligned}$$

例如，掷 3 颗骰子，则最大点数为 6 的概率为$\frac{6^3-5^3}{6^3}=0.4213$.

推论(**单调性**)　若 $A\subseteq B$，则 $P(A)\leqslant P(B)$.

容易举例说明：以上推论的逆命题不成立，即由 $P(A)\leqslant P(B)$无法推出 $A\subseteq B$.

性质 5　对任意两个事件 A，B，有

$$P(A\setminus B)=P(A)-P(AB).$$

证明　因为 $A\setminus B=A\setminus AB$，且 $AB\subseteq A$，所以由性质 4 得

$$P(A\setminus B)=P(A\setminus AB)=P(A)-P(AB).$$

1.3.3　概率的加法公式

当事件之间互不相容时，有限可加性或可列可加性给出了求事件和的概率公式．那么对一般的事件(不一定互不相容)，又如何求事件和的概率呢？下面的性质 6 给出了求任意两个事件和的概率加法公式，进一步给出了求任意 n 个事件和的概率加法公式．这些性质在计算概率时是非常有用的．

性质 6　对任意的两个事件 A，B，有

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

证　因为
$$A\cup B=A\cup(B\setminus AB),$$
$$A\cap(B\setminus AB)=\varnothing,$$
所以有
$$P(A\cup B)=P(A)+P(B\setminus AB).$$

又因为 $AB\subseteq B$，从而由性质 4 可得

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

性质 6 推广到任意三个事件有：设 A，B，C 为任意三个事件，则有

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

性质 6 可以用归纳法推广到任意有限个事件．设 A_1，A_2，…，A_n 是 n 个随机事件，则有

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)+\cdots+(-1)^{n-1}P(A_1\cdots A_n).$$

这个公式也称为概率的一般**加法公式**．

例 4　设事件 A，B 的概率分别为$\frac{1}{3}$和$\frac{1}{2}$，求在以下三种情况下的 $P(B\bar{A})$值．

(1) A，B 互斥；(2) $A\subseteq B$；(3) $P(AB)=\frac{1}{8}$.

解　(1) 由于 A 与 B 互斥，则 $B\subseteq\overline{A}$，所以 $B\overline{A}=B$，即得

$$P(B\overline{A})=P(B)=\frac{1}{2}.$$

(2) 当 $A\subseteq B$ 时，$P(B\overline{A})=P(B\setminus A)=P(B)-P(A)=\frac{1}{2}-\frac{1}{3}=\frac{1}{6}$.

(3) 因为 $B\overline{A}=B\setminus AB$，又 $AB\subseteq B$，由概率的性质得

$$P(B\overline{A})=P(B)-P(BA)=\frac{1}{2}-\frac{1}{8}=\frac{3}{8}.$$

例 5　设 A，B，C 是三个事件，且 $P(A)=P(B)=P(C)=\frac{1}{4}$，$P(AC)=P(BC)=\frac{1}{16}$，$P(AB)=0$，求：(1) A，B，C 至少有一个发生的概率；(2) A，B，C 都不发生的概率.

解　由于 $P(AB)=0$，又 $ABC\subseteq AB$，得 $0\leqslant P(ABC)\leqslant P(AB)=0$，所以 $P(ABC)=0$.

$$\begin{aligned}(1)\ P(A\cup B\cup C)&=P(A)+P(B)+P(C)-P(AB)-\\&\quad P(AC)-P(BC)+P(ABC)\\&=\frac{1}{4}+\frac{1}{4}+\frac{1}{4}-0-\frac{1}{16}-\frac{1}{16}+0=\frac{5}{8}.\end{aligned}$$

(2) $P(\overline{A}\overline{B}\overline{C})=1-P(A\cup B\cup C)=1-\frac{5}{8}=\frac{3}{8}$.

例 6　设 A，B，C 是三个事件，证明：$P(AB)+P(AC)-P(BC)\leqslant P(A)$.

证明　
$$\begin{aligned}P(A)&\geqslant P(A(B\cup C))=P(AB\cup AC)\\&=P(AB)+P(AC)-P(ABC)\geqslant P(AB)+P(AC)-P(BC).\end{aligned}$$

例 7　设 A，B 是两个事件，证明：$P(AB)P(A\cup B)\leqslant P(A)P(B)$.

证明　因为 $A\cup B=AB\cup A\overline{B}\cup\overline{A}B$，且 AB，$A\overline{B}$，$\overline{A}B$ 为三个互不相容的事件，所以

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

从而有　$P(A\cup B)P(AB)=[P(AB)]^2+P(AB)P(A\overline{B})+P(AB)P(\overline{A}B)$，

又　
$$\begin{aligned}P(A)P(B)&=P(AB\cup A\overline{B})P(AB\cup\overline{A}B)\\&=[P(AB)+P(A\overline{B})][P(AB)+P(\overline{A}B)]\\&=[P(AB)]^2+P(AB)P(\overline{A}B)+P(AB)P(A\overline{B})+P(\overline{A}B)P(A\overline{B}),\end{aligned}$$

由概率的非负性知 $P(\overline{A}B)P(A\overline{B})\geqslant 0$，比较上面两式可得

$$P(AB)P(A\cup B)\leqslant P(A)P(B).$$

推论(半可加性)　对任意两个事件 A，B，有

$$P(A\cup B)\leqslant P(A)+P(B).$$

1.4　条件概率及其相关公式

1.4.1　条件概率

在实际问题中，我们往往会遇到在事件 A 已经发生的条件下求事件 B 的概率的情况，这时就有了附加条件．一般说来，它与 B 发生的概率是不相同的，这种概率称为事件 A 发生条件下事件 B 发生的条件概率，记为 $P(B|A)$．下面先看一个例子．

例 1　已知某家庭有 2 个小孩，且至少有一个是女孩，求该家庭有一个男孩，一个女孩的概率．

解　设每个小孩是男孩、女孩是等可能的，则样本空间 $\Omega=\{$(男，男)，(女，男)，(男，女)，(女，女)$\}$，$A=\{$该家庭 2 个小孩中至少有一个女孩$\}$，$B=\{$该家庭 2 个小孩恰好是一个男孩，一个女孩$\}$．用(男，女)表示大的是男孩，小的是女孩．下面我们考虑在事件 A 发生的条件下，事件 B 发生的概率．由于已知事件 A 已经发生，则该试验的所有可能结果为(女，男)，(男，女)，(女，女)．这时，事件 B 是在事件 A 已经发生的条件下只有两种可能结果(女，男)，(男，女)，所以事件 A 发生的条件下，事件 B 发生的概率为

$$P(B|A)=\frac{2}{3}.$$

由上面的例子可知求事件 A 发生的条件下，事件 B 发生的概率，事实上是由于事件 A 的发生，排除了(男，男)发生的可能性，这时样本空间也随之改变为 $\Omega_A=\{$(女，男)，(男，女)，(女，女)$\}$，而在 Ω_A 中事件 B 只含 2 个样本点，故 $P(B|A)=\dfrac{2}{3}$. 这就是条件概率，它与(无条件)概率 $P(B)$是不同的两个概念．

容易知道 $P(A)=\dfrac{3}{4}$，$P(AB)=\dfrac{2}{4}$，$P(B|A)=\dfrac{2}{3}=\dfrac{2/4}{3/4}=\dfrac{P(AB)}{P(A)}$.

这个关系具有一般性，即条件概率是两个无条件概率之商，这就是条件概率的定义．

定义 1　设 A，B 为随机试验 E 的两个随机事件，且 $P(A)>0$，则

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

称为**在事件 A 发生的条件下事件 B 发生的条件概率**，简称**条件概率**．

不难验证，条件概率 $P(B|A)$ 符合概率定义中的三条公理，即

(1) 对于任一事件 B，都有 $P(B|A)\geqslant 0$；

(2) $P(\Omega|A)=1$；

(3) 设 B_1，…，B_n，…是两两互不相容的事件，则有

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

由此，我们很容易证明，条件概率也具有与无条件概率相同的性质，例如，

$$P(\overline{A}|B)=1-P(A|B),$$

$$P(A_1\cup A_2|B)=P(A_1|B)+P(A_2|B)-P(A_1A_2|B).$$

例 2　设 10 件产品中有 4 件不合格，从中任取两件，已知两件中一件是不合格品，求另一件也是不合格品的概率.

解　记事件 A_i 为"第 i 次取出不合格品"，$i=1$，2，B 为"有一件是不合格品"，C 为"另一件也是不合格品"，则 $B=A_1\overline{A}_2\cup\overline{A}_1A_2\cup A_1A_2$，$BC=A_1A_2$. 因为

$$P(B)=P(A_1\overline{A}_2)+P(\overline{A_1}A_2)+P(A_1A_2)=\frac{4\times 6}{10\times 9}+\frac{6\times 4}{10\times 9}+\frac{4\times 3}{10\times 9}=\frac{2}{3},$$

$$P(BC)=P(A_1A_2)=\frac{4\times 3}{10\times 9}=\frac{2}{15},$$

所以根据题意可得

$$P(C|B)=\frac{2/15}{2/3}=\frac{1}{5}.$$

1.4.2　乘法公式

由条件概率的定义可知，对任意两个事件 A 和 B，若 $P(A)>0$，则有

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

此公式称为**乘法公式**.

我们可以将这个公式推广到有限个事件的情况，即：

若 $P(A_1A_2\cdots A_{n-1})>0$，则

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

例 3　设有 100 株玉米，已知病株率 10%，每次检查时从其中任选一株，检查后作上标记不再检查，求第三次才取得健康株的概率.

解　设 $A_i=$"第 i 次取到健康株"，$i=1$，2，3，则所求的概率为 $P(\overline{A}_1\overline{A}_2A_3)$.

$$P(\overline{A}_1)=\frac{10}{100},\ P(\overline{A}_2|\overline{A}_1)=\frac{9}{99},\ P(A_3|\overline{A}_1\overline{A}_2)=\frac{90}{98},$$

所以有
$$P(\overline{A}_1\overline{A}_2A_3)=P(\overline{A}_1)P(\overline{A}_2|\overline{A}_1)P(A_3|\overline{A}_1\overline{A}_2)$$
$$=\frac{10}{100}\times\frac{9}{99}\times\frac{90}{98}=0.0083.$$

例 4　袋中有 10 个广柑，其中 4 个经过化学药品的处理，从口袋中逐个取出．令 $A_k=$“第 k 次取出的广柑是经过化学药品处理过的”，其中 $1\leqslant k\leqslant 10$，求 $P(A_1A_2)$，$P(A_2)$.

解　$P(A_1A_2)=P(A_1)P(A_2|A_1)=\frac{4}{10}\times\frac{3}{9}=\frac{2}{15}$.

由 $A_2=\Omega A_2=A_1A_2+\overline{A}_1A_2$，故
$$\begin{aligned}P(A_2)&=P(A_1A_2+\overline{A}_1A_2)=P(A_1A_2)+P(\overline{A}_1A_2)\\&=P(A_1)P(A_2|A_1)+P(\overline{A}_1)P(A_2|\overline{A}_1)\\&=\frac{2}{15}+\frac{6}{10}\times\frac{4}{9}=\frac{2}{5}.\end{aligned}$$

在求 $P(A_2)$ 的过程中，对 Ω 作如下分解：$\Omega=A_1+\overline{A}_1$，且使 A_1A_2 与 $\overline{A}_1A_2$ 互不相容．这种分解方法体现了以下全概率公式的基本思想．

1.4.3　全概率公式

前面讨论的是直接利用概率的可加性及乘法公式计算简单事件的概率．将复杂问题适当地分解为若干简单问题而逐一解决，是人们常用的工作方法．在解决较复杂的求概率问题时，人们也希望把所涉及的复杂事件分解为简单事件之和．本小节介绍的全概率公式，就是借助于样本空间的分解，把复杂事件先分解为一些互不相容事件之和的形式，再利用概率的加法公式与乘法定理，得到最终要求的复杂事件的概率．

定义 2　设 Ω 是随机试验 E 的样本空间，A_1，…，A_n 为 E 的一组事件，如果

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

（2）$A_1\cup\cdots\cup A_n=\Omega$，

则称 A_1，…，A_n 为样本空间的一个划分，或称 A_1，…，A_n 构成完备事件组．

定理 1　设 Ω 是随机试验 E 的样本空间，B 为 Ω 中的一个事件，A_1，…，A_n 为 Ω 的一个划分，且 $P(A_i)>0$，$i=1,\ 2,\ \cdots,\ n$，则
$$P(B)=\sum_{i=1}^{n}P(A_i)P(B|A_i).\tag{1.4.3}$$

此公式称为全概率公式，它是概率论中一个非常重要的公式，它提供了计算复杂事件概率的一条有效途径，使一个复杂事件的概率计算问题化繁为简．

证明　因为$\bigcup\limits_{i=1}^{n} A_i=\Omega$，所以

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

由于 A_1，…，A_n 两两互不相容，所以 BA_1，…，BA_n 也是两两互不相容，由概率的有限可加性得 $P(B) = \sum\limits_{i=1}^{n} P(BA_i)$. 由于 $P(A_i)>0$，$i=1$，2，…，n，且 $P(BA_i)=P(A_i)P(B|A_i)$，因此，$P(B) = \sum\limits_{i=1}^{n} P(A_i)P(B|A_i)$.

例 5　设一个仓库中共 10 箱同样规格的产品，已知这 10 箱产品中依次有 5 箱、3 箱、2 箱是甲厂、乙厂、丙厂生产的，又甲厂、乙厂、丙厂生产的该种产品次品率依次为$\frac{1}{10}$，$\frac{1}{15}$，$\frac{1}{20}$，现从这 10 箱产品中任取一箱，再从取得的这箱中任取一件产品，求取得正品的概率 .

解　设 B= “取到正品”，A_1 = “取到的这箱产品是甲厂的”，A_2 = “取到的这箱产品是乙厂的”，A_3 = “取到的这箱产品是丙厂的”，由题意得

$$P(A_1)=\frac{5}{10}，P(A_2)=\frac{3}{10}，P(A_3)=\frac{2}{10}，$$

$$P(B|A_1)=\frac{9}{10}，P(B|A_2)=\frac{14}{15}，P(B|A_3)=\frac{19}{20}，$$

由全概率公式得

$$P(B)= \sum_{i=1}^{3} P(A_i)P(B|A_i) = \frac{5}{10}\times\frac{9}{10}+\frac{3}{10}\times\frac{14}{15}+\frac{2}{10}\times\frac{19}{20} = 0.92.$$

例 6　播种时用的一等小麦种子中混有 2%的二等种子，1.5%的三等种子，1%的四等种子 . 如果用一等、二等、三等、四等种子播种后长出来的麦穗含 50 颗以上麦粒的概率分别是 0.5，0.15，0.1，0.05，求这批种子播种后长出来的麦子所结的麦穗含有 50 颗以上麦粒的概率 .

解　记 A= “用这批种子播种后长出来的麦子所结的麦穗具有 50 颗以上的麦粒”，B_1，B_2，B_3，B_4 分别表示播种用的麦种是一等种子、二等种子、三等种子和四等种子，则 $\forall i\neq j(i，j=1，2，3，4)$，$B_iB_j=\varnothing$，$\bigcup\limits_{i=1}^{4} B_i=\Omega$. 而

$$P(B_1)=0.955，P(B_2)=0.02，P(B_3)=0.015，P(B_4)=0.01，$$

$$P(A|B_1)=0.5，P(A|B_2)=0.15，P(A|B_3)=0.1，P(A|B_4)=0.05，$$

所以

$$\begin{aligned}P(A) &= \sum_{i=1}^{4} P(B_i)P(A|B_i)\\ &= 0.955\times0.5+0.02\times0.15+0.015\times0.1+0.01\times0.05\end{aligned}$$

$= 0.4825.$

例 7(**敏感问题调查**)　吸毒会严重影响人的身心健康．而这些都是避着家人进行的，为个人隐私．现在要设计一个调查方案，从调查数据中估计出人群中吸毒人的比率 p.

像这类敏感性问题的调查是社会调查的一类，如人群中参加赌博的比率、经营者中偷税漏税户的比率、学生中阅读黄色书刊和观看黄色影像的比率、考试作弊的比率等．

对敏感性问题的调查方案，关键要使被调查者既愿意作出真实回答又能保守个人秘密．一旦调查方案设计有误，被调查者就会拒绝配合，所得调查数据将失去真实性，经过多年研究和实践，一些心理学家和统计学家设计了一种调查方案，在这个方案中被调查者只需回答以下两个问题中的一个问题，而且只需回答“是”或“否”．

问题 A：你的生日是否在 7 月 1 日之前？

问题 B：你是否吸食过毒品？

这个调查方案看似简单，但为了消除被调查者的顾虑，使被调查者确信他参加这次调查不会泄露个人秘密，在操作上有以下关键点：

(1) 被调查者在没有旁人的情况下，独自一人在一个房间内操作和回答问题.

(2) 被调查者从一个罐子中随机抽一只球，看过颜色后即放回．若抽到白球，则回答问题 A；若抽到红球，则回答问题 B，且罐中只有白球和红球．

被调查者无论回答问题 A 或问题 B，只需在答卷(图 1－10)上认可的方框内打钩，然后把答卷放入一只密封的投票箱内．

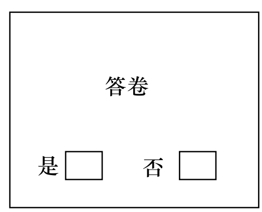


图 1－10

由于这个调查方案使旁人无法知道被调查者回答的是问题 A 还是问题 B，因此可以极大地消除被调查者的顾虑．

接下来是如何分析调查的结果．显然，我们对问题 A 是不感兴趣的．

首先我们设有 n 张答卷(n 较大，譬如 1000 以上)，其中 k 张回答“是”．而我们又无法知道这 n 张答卷中有多少张是回答问题 B 的，同样无法知道 k 张回答“是”的答卷中有多少张是回答问题 B 的，但有两个信息我们是预先知道的，即(1) 在参加调查人数较多的场合，任选一人其生日是 7 月 1 日之前的概率为 0.5.(2) 罐中红球的比率 π 是已知的．现在从这 4 个数据(n，k，0.5，π)去求出 p，因为由全概率公式得

$$P(\text{是})=P(\text{白球})P(\text{是}\mid\text{白球})+P(\text{红球})P(\text{是}\mid\text{红球}),$$

所以将 $P(\text{红球})=\pi$，$P(\text{白球})=1-\pi$，$P(\text{是}\mid\text{白球})=0.5$，$P(\text{是}\mid\text{红球})=p$ 代入上式右边，而上式左边用频率 k/n 代替概率 $P(\text{是})$可得

$$p=\frac{k/n-0.5(1-\pi)}{\pi}.$$

因为我们用频率 k/n 代替了概率 $P(\text{是})$，所以从上式得到的是 p 的估计．

例如，在一次实际调查中，罐中放有红球 30 个，白球 20 个，则 $\pi=0.6$，调查结束后共收到 1845 张有效答卷，其中有 388 张回答“是”，由此可计算得

$$p=\frac{388/1845-0.5\times0.4}{0.6}=0.0172,$$

这表明：约有 1.72％的人吸食过毒品．

由以上几个例子可以看出，全概率公式的形式虽然很简单，但它所代表的“对样本空间进行适当分解”的思想却是十分重要的，具有很高的技巧性．通过对样本空间的分解，我们可以将“较复杂的”事件转化与分解为若干不相容事件的并，从而使问题得以解决．这种事件分解的思想在后面章节仍会不断涉及，它贯穿概率论课程的始终，是概率论的一大特点．

1.4.4　贝叶斯(Bayes)公式

在例 5 中，若已知取得正品，问这件正品来自三个工厂的概率分别为多大？解决这类问题需要用到下面的贝叶斯公式．

定理 2　设 A_1，A_2，…，A_n 为样本空间的一个划分，且 $P(A_i)>0$，B 为样本空间中的任一事件，且 $P(B)>0$，则

$$P(A_i\mid B)=\frac{P(A_i)P(B\mid A_i)}{\sum_{j=1}^{n}P(A_j)P(B\mid A_j)}(i=1,2,\cdots,n).\qquad(1.4.4)$$

该公式由贝叶斯于 1763 年提出，它是观察到事件 B 发生的条件下，寻找导致 B 发生的每一个原因．贝叶斯公式在实际中有很多应用，它可以帮助人们确定某结果发生的最可能的原因．

证明　由条件概率的定义

$$P(A_i\mid B)=\frac{P(A_iB)}{P(B)}.$$

对上式的分子用乘法公式、分母用全概率公式，

$$P(A_iB)=P(A_i)P(B\mid A_i),\ P(B)=\sum_{j=1}^{n}P(A_j)P(B\mid A_j),$$

即得

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

最常用到的贝叶斯公式是当 $P(B)>0$ 时，

$$P(A|B)=\frac{P(A)P(B|A)}{P(A)P(B|A)+P(\overline{A})P(B|\overline{A})}.$$

例 8　发报台分别以概率 0.6 和 0.4 发出信号“1”和“0”，由于通信系统受到干扰，当发出“1”时，收报台分别以概率 0.8 和 0.2 收到信号“1”和“0”，又当发出“0”时，收报台分别以概率 0.9 和 0.1 收到信号“0”和“1”，求当收报台收到“1”时，发报台确实发出“1”的概率，以及收到“0”时，确实发出“0”的概率.

解　设 $A=$“发报台发出信号‘1’”，$B=$“收报台收到信号‘1’”，则

$$P(A)=0.6，P(\overline{A})=0.4，P(B|A)=0.8，$$
$$P(\overline{B}|A)=0.2，P(\overline{B}|\overline{A})=0.9，P(B|\overline{A})=0.1.$$

由贝叶斯公式得

$$P(A|B)=\frac{P(A)P(B|A)}{P(A)P(B|A)+P(\overline{A})P(B|\overline{A})}=\frac{0.6\times0.8}{0.6\times0.8+0.4\times0.1}=\frac{12}{13},$$
$$P(\overline{A}|\overline{B})=\frac{P(\overline{A})P(\overline{B}|\overline{A})}{P(\overline{A})P(\overline{B}|\overline{A})+P(A)P(\overline{B}|A)}=\frac{0.4\times0.9}{0.4\times0.9+0.6\times0.2}=\frac{3}{4}.$$

例 9(疾病普查问题)　一种新方法对某种特定疾病的诊断准确率是 90%(有病被正确诊断和没病被正确诊断的概率都是 90%). 如果群体中这种病的发病率是 0.1%，甲在身体普查中被诊断患有这种病，问甲的确患有这种疾病的概率是多少?

解　设 $A=$“甲患有这种疾病”，$B=$“甲被诊断患这种疾病”. 已知

$$P(A)=0.001，P(B|A)=0.90，$$
$$P(B|\overline{A})=1-P(\overline{B}|\overline{A})=1-0.90=0.10，$$

由贝叶斯公式得

$$P(A|B)=\frac{P(A)P(B|A)}{P(A)P(B|A)+P(\overline{A})P(B|\overline{A})}$$
$$=\frac{0.001\times0.9}{0.001\times0.9+0.999\times0.1}=0.0089<1\%.$$

没有病的概率 $P(\overline{A}|B)=1-P(A|B)=0.9911>99\%$. 造成这个结果的原因是发病率较低和诊断的准确性不够高.

如果用 $P(A)=0.0089$ 代入上式，可计算得 $P(A|B)=7.48\%$，即甲复查时又被诊断有病，则他的确有病的概率将增加到 7.48%.

如果人群的发病率不变，诊断的正确率提高到 99%，可以计算出 $P(A|B)=9.02\%$.

综合上面的讨论可知，条件概率 $P(A|B)$和人群的发病率 $P(A)$以及诊断方法的正确率有关，对发病率很高的疾病进行普查还是有意义的，特别是在诊断的正确性也很高的情况下 .

在贝叶斯公式中，称 $P(A_i)$为 A_i 的先验概率，而 $P(A_i|B)$称为 A_i 的后验概率，贝叶斯公式是专门用于计算后验概率的，也就是通过 B 的发生这个新信息，来对 A_i 的概率作出的修正 . 下面例子很好地说明了这一点 .

例 10　伊索寓言“孩子与狼”讲的是一个小孩每天到山上放羊，山里有狼出没 . 第一天，他在山上喊“狼来了！狼来了!”，山下的村民闻声便去打狼，可到山上，发现狼没来；第二天仍是如此；第三天，狼真来了，可无论小孩怎么喊叫，也没有人来救他，因为前两次他说了谎，人们不再相信他了 .

现在用贝叶斯公式来分析此预言中村民对这个小孩的可信程度是如何下降的.

首先记 B 为事件“小孩说谎”，A 为事件“小孩可信”，不妨设村民过去对小孩的印象为

$$P(A)=0.8,\ P(\overline{A})=0.2. \tag{1.4.5}$$

我们现在用贝叶斯公式来求 $P(A|B)$，即这个小孩说了一次谎后，村民对他可信程度的改变 . 在贝叶斯公式中我们要用到 $P(B|A)$和 $P(B|\overline{A})$，这两个概率的含义分别为可信小孩说谎的可能性和不可信的小孩说谎的可能性 . 在此不妨设

$$P(B|A)=0.1,\ P(B|\overline{A})=0.5.$$

第一次村民上山打狼，发现狼没来，即小孩说了谎，村民根据这个信息，对这个小孩的可信程度改变为(用贝叶斯公式)

$$P(A|B)=\frac{P(A)P(B|A)}{P(A)P(B|A)+P(\overline{A})P(B|\overline{A})}$$

$$=\frac{0.8\times 0.1}{0.8\times 0.1+0.2\times 0.5}=0.444.$$

这表明村民上了一次当后，对这个小孩的可信程度由原来的 0.8 调整为 0.444，也就是(1.4.5)调整为

$$P(A)=0.444,\ P(\overline{A})=0.556. \tag{1.4.6}$$

在此基础上，我们再一次用贝叶斯公式来计算 $P(A|B)$，即这个小孩第二次说谎后，村民对他可信程度的改变为

$$P(A|B)=\frac{0.444\times 0.1}{0.444\times 0.1+0.556\times 0.5}=0.138.$$

这表明村民们经过两次上当，对这个小孩的可信程度已经由 0.8 下降到了 0.138，如此的可信程度，村民听到第三次呼叫时怎么会再上山打狼呢？

1.5 独立性

1.5.1 事件的独立性

1. 两个事件的相互独立

我们已经知道，一般来说，$P(A)$与 $P(A|B)$是不相等的．但是有时事件 B 发生与否并不影响事件A 的发生．例如，口袋里有 10 只球，其中 3 只红球，7 只白球，采用有放回抽样的方式，从口袋中随机地取两次．设事件 B 表示“第一次取到的是红球”，事件 A 表示“第二次取到的是红球”，则显然有 $P(A)=P(A|B)=P(A|\overline{B})$．这时我们称事件 A 与事件B 是相互独立的．

定义 1 设两个随机事件 A，B 满足

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

则称事件 A 与 B 相互独立．

定理 1 当 $0<P(A)<1$ 或 $0<P(B)<1$ 时，事件 A 与 B 相互独立的充要条件是：

$$P(A|B)=P(A)\text{或 } P(A|\overline{B})=P(A),$$

或

$$P(B|A)=P(B)\text{或 } P(B|\overline{A})=P(B). \tag{1.5.1}$$

这也就是说，一个事件的发生与否对另外一个事件的发生并没有影响，这就是事件独立性的含义．

定理 2 若事件 A，B 相互独立，则 A 与$\overline{B}$，$\overline{A}$与 B，$\overline{A}$与$\overline{B}$也相互独立．

证 这里我们只证$\overline{A}$与$\overline{B}$独立(其余的读者自己证明)：

$$\begin{aligned}P(\overline{A}\,\overline{B})&=P(\overline{A\cup B})=1-P(A\cup B)=1-P(A)-P(B)+P(AB)\\&=1-P(A)-P(B)+P(A)P(B)\\&=(1-P(A))(1-P(B))=P(\overline{A})P(\overline{B}),\end{aligned}$$

故$\overline{A}$与$\overline{B}$相互独立．

2. 多个事件的相互独立与两两独立

两个事件的独立性可以推广到多个事件的独立性上去．下面我们来研究三个事件的独立性问题．

定义 2 设事件 A，B，C 满足

$$P(AB)=P(A)P(B), \tag{1.5.2}$$

$$P(AC)=P(A)P(C), \tag{1.5.3}$$

$$P(BC)=P(B)P(C), \tag{1.5.4}$$

$$P(ABC)=P(A)P(B)P(C), \tag{1.5.5}$$

则称 A，B，C **相互独立**，若事件 A，B，C 仅满足前三个条件，则称事件 A，B，C **两两独立**.

根据定义 2 可知，A，B，C 相互独立可推出 A，B，C 两两独立，但反之不成立，下面的例子就说明了这一点.

例 1　一个均匀的正四面体，第一面涂上红色，第二面涂上白色，第三面涂上黄色，第四面同时涂上红、白、黄三种颜色．现在我们以事件 A，B，C 分别表示投一次正四面体出现红，白，黄颜色．因为在正四面体中有两面有红色，所以

$$P(A)=\frac{1}{2}.$$

同理

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

而事件 AB，BC，AC 分别表示投一次正四面体出现红白色、白黄色、红黄色，由于正四面体的第四面有红白黄三种颜色，所以

$$P(AB)=P(BC)=P(AC)=\frac{1}{4},$$

所以定义中的前三个等式成立．事件 ABC 表示投一次正四面体出现红白黄色，所以

$$P(ABC)=\frac{1}{4},$$

但是

$$P(ABC)\neq P(A)P(B)P(C),$$

因此第四个等式不成立，从而事件 A，B，C 不相互独立.

定理 3　若事件 A，B，C 相互独立，则事件 A，B，$\overline{C}$，事件 A，$\overline{B}$，C，事件 $\overline{A}$，B，C，事件 $\overline{A}$，$\overline{B}$，C，事件 $\overline{A}$，B，$\overline{C}$，事件 A，$\overline{B}$，$\overline{C}$，事件 $\overline{A}$，$\overline{B}$，$\overline{C}$ 也相互独立.

例 2　甲、乙两人独立地向同一目标射击，甲击中目标的概率为 0.9，乙击中目标的概率为 0.8，求目标被击中的概率.

解　设 $A=$“甲击中目标”，$B=$“乙击中目标”，则“目标被击中”$=A\cup B$，$P(A)=0.9$，$P(B)=0.8$，并且认为事件 A 与事件 B 是相互独立的，由加法公式可知

$$\begin{aligned}P(A\cup B)&=P(A)+P(B)-P(AB)\\&=0.9+0.8-0.9\times 0.8=0.98.\end{aligned}$$

例 3　甲、乙、丙三人向同一飞机射击，设各人是否命中相互独立，他们的命中率分别为 0.4，0.5，0.7. 若只有一人射中，飞机坠毁的概率为 0.2；若二人同时射中，飞机坠毁的概率为 0.6；若三人同时射中，飞机一定坠毁．求飞机坠毁的概率.

解　飞机坠毁这一事件可看成由以下几个互不相容事件的并组成：因一人击中而坠毁，因二人击中而坠毁，因三人击中而坠毁，故可依此思路进行样本空间的分解．

设 $B=$ “飞机坠毁”，$A_i=$ “恰有 i 个人同时击中”（$i=0，1，2，3$），则

$$\Omega=A_0\cup A_1\cup A_2\cup A_3,$$

记 $C_1=$ “甲射中”，$C_2=$ “乙射中”，$C_3=$ “丙射中”，则

$$A_0=\overline{C}_1\overline{C}_2\overline{C}_3,\ A_1=(C_1\overline{C}_2\overline{C}_3)\cup(\overline{C}_1C_2\overline{C}_3)\cup(\overline{C}_1\overline{C}_2C_3),$$

$$A_2=(\overline{C}_1C_2C_3)\cup(C_1\overline{C}_2C_3)\cup(C_1C_2\overline{C}_3),\ A_3=C_1C_2C_3.$$

由已知数据、概率可加性及乘法公式可计算出

$$P(A_0)=0.6\times0.5\times0.3=0.09,$$

$$P(A_1)=0.4\times0.5\times0.3+0.6\times0.5\times0.3+0.6\times0.5\times0.7=0.36.$$

类似地，可以计算得

$$P(A_2)=0.41,\ P(A_3)=0.14.$$

由题意 $P(B|A_0)=0$，$P(B|A_1)=0.2$，$P(B|A_2)=0.6$，$P(B|A_3)=1$，故

$$\begin{aligned}P(B)&=\sum_{n=0}^{3}P(A_i)P(B|A_i)\\&=0.09\times0+0.36\times0.2+0.41\times0.6+0.14\times1\\&=0.458,\end{aligned}$$

即飞机坠毁的概率为 0.458.

我们可以定义三个以上事件的相互独立性．

定义 3　设有 n 个事件 $A_1，A_2，\cdots，A_n$，如果对任何 $1\leqslant j_1<j_2<\cdots<j_k\leqslant n$ 满足

$$P(A_{j_1}A_{j_2}\cdots A_{j_k})=P(A_{j_1})P(A_{j_2})\cdots P(A_{j_k}),$$

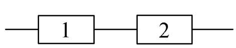
则称 $A_1，A_2，\cdots，A_n$ 相互独立．

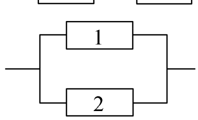
定理 4　设 $A_1，A_2，\cdots，A_n$ 相互独立，则有下面的结果：

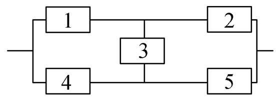
（1）对 $1\leqslant j_1<j_2<\cdots<j_k\leqslant n$，$A_{j_1}，A_{j_2}，\cdots，A_{j_k}$ 相互独立；

（2）用 B_i 表示 A_i 或 $\overline{A}_i$，则 $B_1，B_2，\cdots，B_n$ 相互独立．

例 4　系统由多个元件组成，且所有元件都独立地工作．设每个元件正常工作的概率为 $p=0.9$，试求以下系统的可靠性(即系统正常工作的概率).

（1）串联系统 S_1：

（2）并联系统 S_2：

（3）5 个元件组成的桥式系统 S_3：

解　S_i＝“第 i 个系统能正常工作”，A_i＝“第 i 个元件能正常工作”.

（1）对串联系统而言，“系统能正常工作”相当于“所有元件都能正常工作”，即 $S_1=A_1A_2$，所以

$$P(S_1)=P(A_1A_2)=P(A_1)P(A_2)=p^2=0.81.$$

（2）对并联系统而言，“系统能正常工作”相当于“至少一个元件能正常工作”，即 $S_2=A_1\cup A_2$，所以

$$P(S_2)=P(A_1\cup A_2)=P(A_1)+P(A_2)-P(A_1)P(A_2)=p+p-p^2=0.99.$$

（3）在桥式系统中，第三个元件是关键，我们先用全概率公式得

$$P(S_3)=P(A_3)P(S_3\mid A_3)+P(\overline{A}_3)P(S_3\mid\overline{A}_3).$$

在“第 3 个元件正常工作”的条件下，系统成为先并联后串联系统，如图 1－11 所示，故

$$\begin{aligned}P(S_3\mid A_3)&=P((A_1\cup A_4)(A_2\cup A_5))=P(A_1\cup A_4)P(A_2\cup A_5)\\&=[1-(1-p)^2]^2=0.9801.\end{aligned}$$

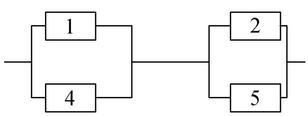


图 1－11　先并后串系统

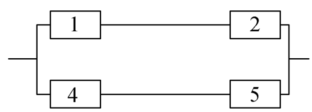


图 1－12　先串后并系统

又因为在“第 3 个元件不正常工作”的条件下，系统成为先串联后并联系统，如图 1－12 所示，所以

$$\begin{aligned}P(S_3\mid\overline{A}_3)&=P(A_1A_2\cup A_4A_5)=1-P(\overline{A_1A_2})P(\overline{A_4A_5})\\&=1-(1-P(A_1)P(A_2))(1-P(A_4)P(A_5))\\&=1-(1-p^2)^2=0.9639.\end{aligned}$$

最后可得

$$\begin{aligned}P(S_3)&=p[1-(1-p)^2]^2+(1-p)[1-(1-p^2)^2]\\&=0.9\times0.9801+0.1\times0.9639=0.9785.\end{aligned}$$

1.5.2　试验的独立性

所谓试验相互独立，就是其中一个试验所得的结果，对其他各试验取得其可能的结果的概率没有影响．下面利用事件的独立性来定义两个或更多个试验的独立性．

定义 4　设有两个试验 E_1 和 E_2，假如试验 E_1 的任一结果(事件)与试验 E_2 的任一结果(事件)都是相互独立的事件，则称**这两个试验相互独立**.

例如，用篮球投一次篮(试验 E_1)出现的结果与掷一颗骰子(试验 E_2)出现的结果之间是相互独立的，所以这两个试验是相互独立的试验.

类似地，可以定义 n 个试验 E_1，E_2，…，E_n 的相互独立性：如果 E_1 的任一结果、E_2 的任一结果，…，E_n 的任一结果都是相互独立的事件，则称**试验 E_1，E_2，…，E_n 相互独立**.

现实生活的很多试验都是在相同条件下重复进行的，且各次试验之间是相互独立的.如果做 n 次试验，它们完全是同一个试验在相同条件下的重复，且在每次试验中事件发生的概率都不依赖于其他各次试验的结果，则称**这种试验为 n 重独立重复试验**.

设试验 E 只有两个可能结果：A 及 $\overline{A}$，则称 E 为**伯努利**(Bernoulli)**试验**.将伯努利试验 E 独立地重复地进行 n 次，则称为 n **重伯努利试验**.

如果已知在一次试验中事件 A 发生的概率为 p，我们考虑某个事件 A 在 n 次重复独立试验中发生 $k(0\leqslant k\leqslant n)$次的概率，可以得到以下的定理.

定理 5　设在一次试验中事件 A 发生的概率为 $p(0<p<1)$，则在 n 次重复独立试验中，事件 A 恰好发生 k 次的概率为

$$P_n(k)=C_n^k p^k(1-p)^{n-k}(k=0,1,2,\cdots,n).\qquad(1.5.6)$$

证　由于在 n 次重复独立试验中事件 A 在指定的 k 个试验中发生，其余 $n-k$ 个试验中不发生的概率为

$$p^k(1-p)^{n-k},$$

且这种指定的方式有 C_n^k 种，因此

$$P_n(k)=C_n^k p^k(1-p)^{n-k}(k=0,1,2,\cdots,n).$$

特别地，

$$P_n(0)+P_n(1)+\cdots+P_n(n)=\sum_{k=0}^{n}C_n^k p^k(1-p)^{n-k}=(p+1-p)^n=1.$$

由于 $C_n^k p^k(1-p)^{n-k}$ 恰好是 $[p+(1-p)]^n$ 的二项展开式中的第 k 项，故此公式又称为二项概率公式.

例 5　设某人打靶，命中率为 0.8，重复射击 5 次，求恰好命中一次的概率.

解　可以认为各次射击是相互独立的，故本试验可以看作 5 重伯努利试验.设 $A=$“恰好命中一次”，则由二项概率公式得

$$P(A)=P_5(1)=C_5^1\times 0.8\times(0.2)^4=0.0064.$$

例 6　一批种子的发芽率为 0.8，试问每穴至少播种几粒种子，才能保证

使每穴不空苗的概率在 0.99 以上？

解　设至少播种 n 粒种子，才能保证使每穴不空苗的概率在 0.99 以上，则可将播 n 粒种子看作 n 重伯努利试验．因为

$$P(\text{至少一粒出苗})\geqslant 0.99 \Leftrightarrow P(\text{没有一粒出苗})<0.01,$$

故要求 n 满足

$$P_n(0)<0.01,\ \text{即}\ 0.2^n<0.01,$$

进而得

$$n>\frac{-2}{\lg 2-1}=2.861,$$

可见每穴至少播种 3 粒，才能保证使每穴不空苗的概率在 0.99 以上．

例 7　明青花(瓷)享有盛誉．设一只青花盘在一年中被失手打破的概率是 0.03.

(1) 计算一只弘治时期(1488—1505 年)的青花麒麟(图案)保留到现在(约 505 年)的概率．

(2) 如果弘治年间生产了 1 万件青花麒麟盘，计算这 1 万件至今都已经被失手打破的概率．

解　(1) 用 A_i 表示该盘在第 i 年没有被打破，则至今没有被打破的概率是

$$\begin{aligned} p &= P(A_1A_2\cdots A_{505})=P(A_1)P(A_2\mid A_1)\cdots P(A_{505}\mid A_1A_2\cdots A_{504}) \\ &=(1-0.03)^{505}=2.09\times 10^{-7}, \end{aligned}$$

则一只弘治时期的青花麒麟盘至今被失手打破的概率

$$q=1-p=0.999999791.$$

(2) 1 万只青花麒麟盘中的每一只至今要么被打破，要么没被打破，且每只是否被打破相互独立，由二项概率公式可得这 1 万件至今都已经被失手打破的概率为

$$P_{10000}(10000)=q^{10000}=0.999999791^{10000}=0.9979.$$

习　题　A

1. 写出下列随机试验的样本空间：

(1) 记录一个小班一次数学考试的平均分数(百分制记分)；

(2) 同时掷两颗骰子，记录两颗骰子的点数之和；

(3) 10 只产品中有 3 只是次品，每次从其中取 1 只，取后不放回，直到 3 只次品都取出为止，记录抽取的次数；

(4) 生产的产品直到得到 5 件正品为止，记录生产产品的总件数；

(5) 测量一辆汽车通过某定点的速度；

(6) 在单位圆内任意取一点，记录它的坐标．

2. 一批产品中有合格品和废品，从中有放回地抽取 3 个产品，设 A_i 表示事件“第 i 次抽得废品”，试用 A_i 的运算表示下列各个事件：

(1) 第一次、第二次中至少有一次抽到废品；

(2) 只有第一次抽到废品；

(3) 三次都抽到废品；

(4) 至少有一次抽到合格品；

(5) 只有两次抽到废品.

3. 已知某设备连接三个管道，事件 A_1，A_2，A_3 分别表示三个管道接通，试用 A_i 的运算表示下列事件：

(1) 设备为发电机，管道用于输送冷水冷却发电机，事件 D 表示发电机正常工作．

(2) 设备为反应锅，管道用于输送三种不同的原料，事件 D 表示反应锅正常工作．

4. 对任意事件 A，B，下列命题中正确的是(　　).

(A) 如果 A，B 互不相容，则 $\overline{A}$，$\overline{B}$ 也互不相容；

(B) 如果 A，B 相容，则 $\overline{A}$，$\overline{B}$ 也相容；

(C) 如果 $\overline{A}$，$\overline{B}$ 互不相容，则 $A \cup B=\Omega$；

(D) 如果 $AB=A$，则 $A \cup B=A$.

5. 一部五卷的文集按任意次序放上书架，则第一卷和第五卷都不在两端的概率为多少？

6. 从一副扑克牌的 13 张黑桃中，一张接一张地有放回地抽取 3 张，求：

(1) 没有同号的概率；

(2) 有同号的概率；

(3) 最多只有两张同号的概率．

7. 设有 10 件产品，其中有 3 件次品，从中任意抽取 5 件，问其中恰有 2 件次品的概率是多少？

8. 一批产品共有 10 个正品 2 个次品，从中任取两次，每次取一个(取后不放回)，求：

(1) 至少取到一个正品的概率；

(2) 第二次取到次品的概率；

(3) 恰有一次取到次品的概率．

9. 把甲、乙、丙三名学生依次随机地分到 5 间宿舍中的任意一间中去，假定每间宿舍最多能住 8 人，试求：

(1) 这 3 名学生住在不同宿舍的概率；

(2) 这 3 名学生至少有 2 名住在同一宿舍中的概率．

10. 有 n 个人，每个人都以同样的概率被分配在 $N(n\leqslant N)$ 间房中的每一间中，求下列事件的概率：

(1) A：某指定 n 间房中各有 1 人；

(2) B：恰有 n 间房，其中各有 1 人；

(3) C：某指定房间中恰有 $m(m\leqslant n)$ 人．

11. 某人有 5 把钥匙，但忘了开房门的是哪一把，逐把试开，问：

(1) 恰好第三次打开房门锁的概率是多少；

(2) 三次内打开的概率是多少；

(3) 如 5 把内有 2 把房门钥匙，三次内打开的概率是多少？

12. 两人约定某日晚 7 点至 8 点在某地会面，试求一人要等另一人半小时以上的概率．

13. 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刻是相等的，如果甲船的停泊时间为 1h，乙船的停泊时间为 2h，求它们中的任何一艘都不需要等候码头空出的概率．

14. 掷三颗骰子，记 $A=$“掷出的点数之和不小于 10”．

(1) 设 $B=$“第一颗骰子出现 1 点”，求 $P(A|B)$；

(2) 设 $C=$“至少有一颗骰子出现 1 点”，求 $P(A|C)$.

15. 已知 $P(A)=\dfrac{1}{4}$，$P(B|A)=\dfrac{1}{3}$，$P(A|B)=\dfrac{1}{2}$，求 $P(A\cup B)$.

16. 从一批次品率为 10% 的 100 件产品中，依次不放回地抽取两件产品，试求：

(1)“两次均取得正品”的概率；

(2)“第二次才取得正品”的概率；

(3)“在已知第一次取得正品的条件下，第二次又取得正品”的概率；

(4) 若题目改为不放回地接连取出 3 个产品，求“第三次才取得正品”的概率．

17. 已知某种动物的生命超过 40 年的概率为 0.8，而超过 50 年的概率为 0.6，现有一只这种动物活了 40 年，则这只动物在 10 年内死亡的概率是多少？

18. 一个学生接连参加同一课程的两次考试．第一次考试及格的概率为 p，若第一次及格则第二次及格的概率为 p，若第一次不及格则第二次及格的概率为 $\dfrac{p}{2}$.

(1) 若至少有一次及格则他能取得某种资格，求他取得该资格的概率；

(2) 若已知他第二次已经及格，求他第一次及格的概率．

19. 玻璃杯成箱出售，每箱 20 只，假设各箱含 0，1，2 只残次品的概率相应为 0.8，0.1 和 0.1，一位顾客欲购一箱玻璃杯，在购买时，售货员随意取一箱，而顾客开箱随机查看 4 只，若无残次品，则买下，否则退回，试求：

(1) 顾客买下该箱的概率；

(2) 在顾客买下的一箱中，确实没有残次品的概率．

20. 步枪 20 支，其中 16 支经过校正，射击命中概率为 0.9，4 支未经过校正，射击命中概率为 0.7. 现从 20 支枪中任取一支射击，结果命中目标，求此支步枪经过校正的概率是多少？

21. 设人群中男人是色盲的概率为 5%，而女人是色盲的概率为 0.25%. 今从人群中任选一人，且发现这人是色盲，求此人是女性的概率．

22. 设甲袋中有 5 个苹果，其中有 2 个是红富士，3 个是红元帅，乙袋中有 4 个苹果，其中 3 个是红富士，1 个是红元帅，丙袋中有 3 个苹果，其中 1 个是红富士，2 个是红元帅，现从甲袋中任取一个苹果放入乙袋中，然后从乙袋中任取一个苹果，求：

(1) 从乙袋中取出的苹果是红元帅的概率；

(2) 将乙袋中取出的苹果放入丙袋，然后从丙袋中任取一个苹果发现是红元帅，从乙袋中取出的苹果是红元帅的概率．

23. 在某乒乓球比赛中，甲、乙两名乒乓球运动员在前 16 名进入前 8 名的淘汰赛中相遇，比赛采取 7 局 4 胜，每局 11 分制．已知甲、乙两名乒乓球运动员进行比赛每局甲胜乙的概率为 0.6(每局比赛的结果互不影响). 求甲淘汰乙顺利进入前 8 强的概率．

24. 设有 3 门高射炮，每门击中飞机的概率都是 0.3，求同时发射一发炮弹而击中飞机的概率．又若一架敌机入侵，若要以 99% 的概率击中它，问至少需要多少门高炮．

25. 设 A，B 是任意两个事件，其中 A 的概率不等于 0 和 1，证明：$P(B|A)=P(B|\overline{A})$ 是事件 A 与 B 独立的充分必要条件．

习　题　B

1. 把 3 棵枫树、4 棵橡树和 5 棵白桦树栽成一行，12 棵的排序是随机的．每一种排列都是等可能的，求没有两棵白桦树相邻的概率．

2. 某大学同窗好友 7 人，临毕业前随机地站成一行照相以作毕业留念，试求以下事件的概率：(1) 甲站在中间；(2) 甲、乙两人不站在两端；(3) 甲、乙、丙三人相邻；(4) 甲、乙、丙三人不相邻；(5) 甲、乙两人之间

有两人；(6) 甲、乙两人相邻但与丙不相邻．

3. 在(0，1)区间内任取两个随机数 x，y，求两数之积小于$\frac{1}{4}$的概率．

4. 证明：对任意三个随机事件 A，B，C，有 $P(AB)+P(AC)-P(BC)\leqslant P(A)$.

5. 设 A，B，C 为三个随机事件，证明：$P(AB)+P(AC)+P(BC)\geqslant P(A)+P(B)+P(C)-1$.

6. 设 A，B，C 为三个随机事件，若 $P(A)=P(B)=P(C)=\frac{1}{2}$，且 $P(ABC)=P(\overline{A}\overline{B}\overline{C})$，证明：$2P(ABC)=P(AB)+P(AC)+P(BC)-\frac{1}{2}$.

7. 设 A，B，C 为三个随机事件，若 $P(A)=a$，$P(B)=2a$，$P(C)=3a$，且 $P(AB)=P(AC)=P(BC)=b$，证明：$a\leqslant\frac{1}{4}$，$b\leqslant\frac{1}{4}$.

8. 某人写了 n 张信笺与 n 个信封，将信笺随便乱装入信封内(每个信封装一张信笺)，求至少有一信匹配的概率是多少．

9. 设有来自三个地区的各 10 名、15 名和 25 名考生的报名表，其中女生的报名表分别为 3 份、7 份和 5 份．随机地取出一个地区的报名表，从中先后抽出两份．

(1) 求先抽到的一份是女生表的概率 p；

(2) 已知后抽到的一份是男生表，求先抽到的一份是女生表的概率 q.

10. 已知 $0<P(B)<1$，$P[(A_1\cup A_2)|B]=P(A_1|B)+P(A_2|B)$，则下列选项成立的是(　　).

(A)$P[(A_1\cup A_2)|\overline{B}]=P(A_1|\overline{B})+P(A_2|\overline{B})$；

(B)$P(A_1B\cup A_2B)=P(A_1B)+P(A_2B)$；

(C)$P(A_1\cup A_2)=P(A_1|B)+P(A_2|B)$；

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

11. 若 $P(A|B)>P(A|\overline{B})$，证明：$P(B|A)>P(B|\overline{A})$.

12. 设根据以往的记录的数据分析，某船只运输的某种物品损坏的情况共有三种：损坏 2%(这一事件记为 A_1)，损坏 10%(事件 A_2)，损坏 90%(事件 A_3)，且已知 $P(A_1)=0.8$，$P(A_2)=0.15$，$P(A_3)=0.05$. 现从已被运输的物品中随机的取 3 件，发现 3 件都是好的(这一事件记为 B)，试求 $P(A_1|B)$，$P(A_2|B)$，$P(A_3|B)$(这里设物品件数很多，取出一件后不影响取后一件是否为好品的概率).

13. 一枚深水炸弹击沉、击伤和击不中一艘潜水艇的概率分别为$\frac{1}{3}$，$\frac{1}{2}$和$\frac{1}{6}$. 设击伤该潜水艇两次也使该潜水艇沉没，求用 4 枚深水炸弹击沉该潜水艇的概率.

14. 如果一危险情况 C 发生时，一电路闭合并发出警报，我们可以借用两个或多个开关并联以改善可靠性. 在 C 发生时这些开关每一个都应闭合，且若至少一个开关闭合了，警报就发出. 如果两个这样的开关并联连接，它们每一个具有 0.96 的可靠性(即在情况 C 发生时闭合的概率)，问这时系统的可靠性(即电路闭合的概率)是多少？如果需要有一个可靠性至少为 0.9999 的系统，则至少需要用多少只开关并联？设各开关闭合与否相互独立.

15. 设 A，B 是任意两个事件，其中 $0<P(A)<1$，$0<P(B)<1$，证明：若 $P(B|A)+P(\overline{B}|\overline{A})=1$，则事件 A 与 B 独立.

16. 设 $P(A)>0$，试证：$P(B|A)\geqslant 1-\dfrac{P(\overline{B})}{P(A)}$.

第 2 章　随机变量及其分布

在前一章的学习中，我们已经讨论了概率论中的一些最基本的概念和一些最简单的概率模型中随机事件的概率求解方法，并给出了概率的公理化定义．这样，我们在遇到一些简单的具体的随机事件时，可以先判断其概型，然后利用已有知识，对其概型特性进行分析，求出概率．但在实际工作中，我们遇到的概率问题往往是错综复杂、千变万化的，并不都是我们熟知的、简单的概型．为更全面研究随机试验结果以揭示随机事件中包含的客观存在的统计规律，我们将随机试验的结果量化，这样就引入了“随机变量”的概念．

引入随机变量这个概念后，可以用随机变量表示随机事件，以便将不同的随机事件的概率及相互关系进行统一处理，而且有助于研究随机事件的本质规律．更为重要的是它极大地丰富了研究对象与范围，它可以解决涉及人类生活中更为广泛的理论与应用的问题．随机变量是概率论中最重要的基本概念之一．

2.1　随机变量

2.1.1　随机变量的定义

通俗地来说，用来表示随机现象结果的变量称为**随机变量**，常用大写字母 X，Y，Z 表示．很多随机事件都可以用随机变量表示，表示时应写明随机变量的含义．而随机变量的含义是人们按需要设置出来的．下面通过一些例子来说明设置是如何进行的．

例 1　很多随机现象的结果本身就是数，把这些数看作某特设变量的取值就可获得随机变量．如掷一颗骰子，可能出现 1，2，3，4，5，6 诸点．若设 $X=$“掷一颗骰子出现的点数”，则 1，2，3，4，5，6 就是随机变量 X 的可能取值，这时

事件“出现 3 点”可用“$X=3$”表示；

事件“出现的点数超过 3 点”可用“$X>3$”表示．

在这个随机现象中，若再设 $Y=$“掷一颗骰子 6 点出现的次数”，Y 是仅取 0 或 1 两个值的随机变量，这是与 X 不同的另一个随机变量．这时，

“$Y=0$”表示事件“没有出现 6 点”；

“$Y=1$”表示事件“出现 6 点”．

上述讨论表明：在同一个随机试验中，不同的设置可获得不同的随机变量，如何设置可按需要进行．

例 2　有些随机现象的结果虽然不是数，但仍可以根据需要设计出有意义的随机变量．如检验一件产品的可能结果有两个：合格品与不合格品．若我们把注意点放在合格品上，则可设置 $X=$“检查一件产品所得的合格品数”，X 是仅取 0 和 1 的随机变量，且

“$X=0$”表示事件“出现不合格品”；

“$X=1$”表示事件“出现合格品”．

若检查 5 件产品，其中不合格品数 Y 是一个随机变量，它的可能取值为 0，1，2，3，4，5 这 6 个值，且

事件“不合格品数不多于 1 件”可用“$Y\leqslant 1$”表示；

“$Y=0$”表示事件“全是合格品”；

“$Y\geqslant 2$”表示事件“至少有两件不合格品”．

下面我们给出随机变量的数学定义．

定义　如果对随机试验的样本空间 Ω 中的每一个样本点 ω，都有一个确定的实数值 $X(\omega)$ 与之对应，则称 X 为**随机变量**，$X(\omega)$ 为随机变量的值．

例 3　设试验为测量车床加工的零件的直径，则有样本点：ω_x 表示“测得零件的直径为 $x$mm”（$a\leqslant x\leqslant b$），样本空间 $\Omega=\{\omega_x \mid a\leqslant x\leqslant b\}$．设 Y 为测得零件的直径，于是有

$$Y(\omega)=x，当\ \omega=\omega_x(a\leqslant x\leqslant b)时．$$

随机变量的取值随试验的结果而定，在试验之前不能预知它取什么值，随机变量不是自变量，它是样本点的函数，这个函数可以是不同的样本点对应不同的实数，也允许多个样本点对应同一个实数．这个函数的自变量(样本点)可以是数，也可以不是数，但因变量一定是实数．由于样本点的出现是随机的，具有一定的概率，所以随机变量的取值有一定的概率．这些性质显示了随机变量与普通函数有着本质的差异．

通过随机变量的引入，我们对随机事件的研究就可以转化到对随机变量的研究上，同时可以用随机变量来描述各种随机现象，这样一来，后者可以利用微积分的手段对随机试验的结果进行广泛的研究和讨论．

2.1.2　随机变量的分类

按照随机变量的可能取值，可以把它们分为两种基本类型，即离散型和非

离散型．

有些随机变量，它可能取到有限个或可列无限多个不相同的值，这样的随机变量称为**离散型随机变量**．

例如，10 个产品有 3 个次品，一次随机抽取 5 个产品，X 表示抽到的产品中所含次品的个数，它只可能取 0，1，2，3 这四个值，它是离散型随机变量；又如，某城市的 120 急救电话台一昼夜收到的呼唤次数 Y 也是离散型随机变量．

而有些随机变量，它所可能取的值充满一个区间，是无法按一定次序一一列举出来的，例如，上述例 3 所指测量车床加工的零件的直径，它是一个**连续型随机变量**．

2.2　离散型随机变量及其分布

2.2.1　离散型随机变量的概率分布律

我们知道，要研究一个离散型随机变量 X 的统计规律，必须且只需知道 X 的所有可能取值以及取每一个可能值的概率．

设离散型随机变量 X 的所有可能取值为 x_1，x_2，…，x_n，…，X 取各个可能值的概率，即事件$\{X=x_i\}$的概率为

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

我们称(2.2.1)式为离散型随机变量 X 的**概率分布**(或**分布律**)．分布律也可以用表格的形式来表示：

X	x_1	x_2	…	x_n	…
p_i	p_1	p_2	…	p_n	…

由概率的性质可知，离散型随机变量 X 的概率分布具有以下性质：

(1) $p_i\geqslant 0$，$i=1$，2，…；(2) $\sum\limits_{i=1}^{\infty} p_i=1$.

性质(2)是由于，如果随机变量 X 只能取得有限个值 x_1，x_2，…，x_n，则随机事件$\{X=x_i\}$($i=1$，2，…，n)构成互不相容的完备事件组，所以 $\sum\limits_{i=1}^{n} p_i=1$；如果随机变量 X 可能取得可列无限个值 x_1，x_2，…，x_n，…，则 $\sum\limits_{i=1}^{\infty} p_i=1$，即级数 $\sum\limits_{i=1}^{\infty} p_i$ 是收敛的，并且它的和等于 1.

例 1　设一汽车在开往目的地的道路上需经过四组信号灯，每组信号灯

以$\frac{1}{2}$的概率允许或禁止汽车通过．以 X 表示汽车首次停下时，它已通过的信号灯的组数(设各组信号灯的工作是相互独立的)，求随机变量 X 的分布律.

解　若以 p 表示每组信号灯禁止汽车通过的概率，则易知 X 的分布律为

X	0	1	2	3	4
p_i	p	$(1-p)p$	$(1-p)^2p$	$(1-p)^3p$	$(1-p)^4$

或写成

$$P\{X=i\}=(1-p)^ip(i=0,1,2,3),\ P\{X=4\}=(1-p)^4.$$

现以 $p=\frac{1}{2}$代入得

X	0	1	2	3	4
p_i	0.5	0.25	0.125	0.0625	0.0625

例 2　袋中有 2 个白球和 3 个黑球，每次从其中任取 1 个球，直至取得白球为止，每次取出的黑球仍放回去，求取球次数的概率分布律．

解　设取球次数是 X，则 X 是随机变量．因为每次取出的黑球仍放回去，所以 X 的可能值是一切正整数，且有

$$P\{X=k\}=\left(\frac{3}{5}\right)^{k-1}\frac{2}{5}=0.4\times0.6^{k-1},\ k=1,2,\cdots.$$

我们可以看出，随机变量 X 取得它的可能值的概率恰为等比数列，所以这种概率分布叫作**几何分布**．几何分布的一般情形是：设离散型随机变量 X 的可能值是一切正整数，且其概率分布律为

$$P\{X=k\}=q^{k-1}p,\ k=1,2,\cdots,$$

其中 $0<p<1$，$p+q=1$. 根据几何级数的收敛性，易知 $\sum_{k=1}^{\infty}pq^{k-1}=\frac{p}{1-q}=1.$

2.2.2　三种常见的离散型随机变量

1. 0—1 **分布**

设随机变量 X 只能取 0 与 1，它的分布律是

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

则称 X 服从 0—1 **分布**或**两点分布**．0—1 分布的分布律也可写成

X	0	1
p_i	$1-p$	p

例 3　200 件产品，190 件是合格的，10 件是不合格的，现从中任取一件，若规定

$$X=\begin{cases}1, & \text{若取得合格产品,}\\ 0, & \text{若取得不合格产品,}\end{cases}$$

则 X 服从参数为 0.95 的 0—1 分布．

0—1 分布是最简单的一种分布类型，它可描述一切只有或只关心两种可能结果的随机事件．比如，产品合格与不合格、新生婴儿是男是女、比赛中的胜与负、电信号的正与负、种子是否发芽等．

2. 二项分布

n 重伯努利试验是一种很重要的数学模型，它有广泛的应用，是研究最多的模型之一．

设随机变量 X 的可能值是 0，1，2，…，n，它的分布律是

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

则称 X 服从参数为 n，p 的**二项分布**，记为 $X\sim B(n,\ p)$.

特别地，当 $n=1$ 时，二项分布化为

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

这就是 0—1 分布．

如果事件 A 在每次试验中发生的概率为 p，随机变量 X 为 n 重伯努利试验中事件 A 发生的次数，则 X 服从二项分布 $B(n,\ p)$.

例如，设一批产品共 N 个，其中有 M 个次品，即次品率 $p=\dfrac{M}{N}$，对这批产品进行放回抽样，即每次任取一个产品，检查其质量后仍放回去，如此连续抽取 n 次，则在被抽查的 n 个产品中的次品数 X 服从二项分布 $B(n,\ p)$.

例 4　已知某种产品的次品率为 0.2，现在从这一大批产品中进行有放回抽样，抽查 20 件，问 20 件产品中恰有 $k(k=0,\ 1,\ \cdots,\ 20)$ 件次品的概率是多少？

解　设随机变量 X 为 20 件产品中所含次品的件数，则 X 服从二项分布 $B(n,\ p)$，其中参数 $n=20$，$p=0.2$. 由二项分布知

$$P\{X=k\}=C_{20}^k(0.2)^k(0.8)^{20-k},\ k=0,\ 1,\ \cdots,\ 20,$$

将计算结果列于表 2－1.

表 2－1

$P\{X=0\}=0.012$	$P\{X=4\}=0.218$	$P\{X=8\}=0.022$
$P\{X=1\}=0.058$	$P\{X=5\}=0.175$	$P\{X=9\}=0.007$
$P\{X=2\}=0.137$	$P\{X=6\}=0.109$	$P\{X=10\}=0.002$
$P\{X=3\}=0.205$	$P\{X=7\}=0.055$	$P\{X=k\}<0.001$，$k\geqslant 11$

为了直观了解，我们作出表 2－1 的图形，如图 2－1 所示．

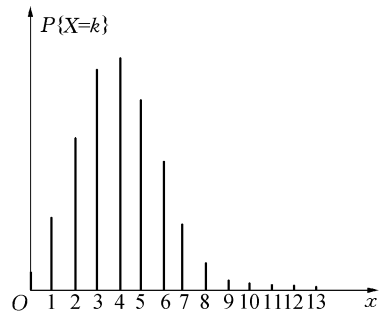


图 2－1

从图 2－1 中可看出，当 k 增加时，概率 $P\{X=k\}$先是随之增加，直至达到最大值(本例中 $k=4$ 时取到最大值)，随之单调减少．

一般地，当 $X\sim B(n,\ p)$时，我们来考虑比值

$$\frac{P\{X=k\}}{P\{X=k-1\}}=\frac{C_n^k p^k q^{n-k}}{C_n^{k-1}p^{k-1}q^{n-k+1}}$$

$$=\frac{(n-k+1)p}{kq}=1+\frac{(n+1)p-k}{kq},$$

因此，

当 $k<(n+1)p$ 时，$P\{X=k\}>P\{X=k-1\}$；

当 $k=(n+1)p$ 时，$P\{X=k\}=P\{X=k-1\}$；

当 $k>(n+1)p$ 时，$P\{X=k\}<P\{X=k-1\}$，

所以，最可能成功次数

$$m_0=\begin{cases} np+p \text{ 或 } np+p-1, & np+p\in\mathbf{Z}^+, \\ [np+p], & np+p\notin\mathbf{Z}^+, \end{cases}$$

我们称 $P\{X=m_0\}$为 $B(n,\ p)$的中心项．

例 4 是不放回抽样，但由于这批产品的总数很大，且抽查产品的数量相对于产品的总数来说又很小，因而可当作放回抽样来处理．这样做会有一些误差，但误差不大．

例 5　某人进行射击训练，每次射中的概率是 0.02，独立射击 400 次，求至少击中 1 次的概率．

解　将每次射击看作一次独立试验，则整个试验可看作一个 400 次的伯努利试验．设击中的次数为 X，则 $X\sim B(400,\ 0.02)$，X 的分布律为

$$P\{X=k\}=C_{400}^k 0.02^k 0.98^{400-k},\ k=0,\ 1,\ 2,\ \cdots,\ 400,$$

则所求概率为

$$P\{X\geqslant 1\}=1-P\{X=0\}=1-0.98^{400}\approx 0.9997.$$

这个例子的实际意义十分有趣．这个射手每次命中的概率只有 0.02，绝不是个天才，但他坚持射击 400 次，则击中目标的概率近似为 1，几乎成为必然事件．这正说明了，“只要功夫深，铁杵磨成针”．不要因为成功的希望小而放弃，只要我们锲而不舍地努力，就一定会到达理想的彼岸．这一道简单的概率题中蕴含着深刻的哲理．

3. 泊松(Poisson)分布

设随机变量 X 的所有可能取值是 0，1，2，…，而取各个值的概率为

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

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

易知，$P\{X=k\}\geqslant 0$，$k=0$，1，2，…，且有

$$\sum_{k=0}^{\infty}P\{X=k\}=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=e^{-\lambda}\cdot e^{\lambda}=1.$$

具有泊松分布的随机变量在实际应用中是很多的．例如，一段时间内电话用户对电话站的呼唤次数、候车的旅客数、某医院在一天内的急诊病人数、一本书一页中的印刷错误数等问题都服从泊松分布．

例 6　已知一本书每页印刷错误的个数 X 服从泊松分布 $P(0.2)$，写出 X 的概率分布律，并求一页上印刷错误数不多于 1 个的概率．

解　随机变量 X 服从泊松分布 $P(0.2)$，则 X 的概率分布律为

$$P\{X=k\}=\frac{0.2^k}{k!}e^{-0.2},\quad k=0,1,2,\cdots,$$

$$P\{X\leqslant 1\}=P\{X=0\}+P\{X=1\}=e^{-0.2}+\frac{0.2}{1!}e^{-0.2}=0.9825.$$

在例 5 中，直接计算 $P\{X=k\}$ 相当麻烦，下面我们介绍一个当 n 相当大，p 很小时二项分布概率的近似公式，这就是二项分布的泊松逼近．

定理(泊松定理)　在 n 重伯努利试验中，记事件 A 在一次试验中发生的概率为 p_n(与 n 有关)，如果当 $n\to\infty$ 时，有 $np_n\to\lambda$，则

$$\lim_{n\to\infty}C_n^k p_n^k(1-p_n)^{n-k}=\frac{\lambda^k}{k!}e^{-\lambda}.$$

证明　记 $np_n=\lambda_n$，即 $p_n=\lambda_n/n$，我们可以得到

$$\begin{aligned}C_n^k p_n^k(1-p_n)^{n-k}&=\frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}\left(\frac{\lambda_n}{n}\right)^k\left(1-\frac{\lambda_n}{n}\right)^{n-k}\\&=\frac{\lambda_n^k}{k!}\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda_n}{n}\right)^{n-k},\end{aligned}$$

对固定的 k 有

$$\lim_{n\to\infty}\lambda_n=\lambda,\ \lim_{n\to\infty}\left(1-\frac{\lambda_n}{n}\right)^{n-k}=e^{-\lambda},\ \lim_{n\to\infty}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)=1,$$

从而

$$\lim_{n\to\infty}C_n^k p_n^k(1-p_n)^{n-k}=\frac{\lambda^k}{k!}e^{-\lambda}$$

对任意的 $k(k=0,1,2,\cdots)$成立．

由于泊松定理是在 $np_n\to\lambda$ 条件下获得的，故在计算二项分布 $B(n,p)$的概率时，当 n 很大，p 很小，而乘积 $\lambda=np$ 大小适中时，可以用泊松分布作近似，即

$$C_n^k p^k(1-p)^{n-k}\approx\frac{\lambda^k}{k!}e^{-\lambda},\ k=0,1,2,\cdots.$$

例 7　已知某批产品共 2000 个，其中 40 个次品，随机地有放回抽取 100 个样品，求样品中次品数 X 的概率分布律．

解　因为次品率 $p=\dfrac{40}{2000}=0.02$，则 $X\sim B(100,0.02)$，

$$P\{X=k\}=C_{100}^k(0.02)^k(0.98)^{100-k},\ k=0,1,\cdots,100. \tag{2.2.2}$$

但因为 $n=100$ 较大，$p=0.02$ 的值较小，则由泊松定理得

$$\lambda=100\times0.02=2,$$

$$P\{X=k\}\approx\frac{2^k}{k!}e^{-2},\ k=0,1,2,\cdots,100. \tag{2.2.3}$$

计算结果列于表 2－2(两种计算结果相当近似)：

表 2－2

X(次品数)	0	1	2	3	4	5	6	…
p_k(按(2.2.2)式)	0.1326	0.2707	0.2734	0.1823	0.0902	0.0353	0.0114	…
p_k(按(2.2.3)式)	0.1353	0.2707	0.2707	0.1804	0.0902	0.0361	0.0120	…

2.3　随机变量的分布函数

对于非离散型随机变量 X，由于其可能值不能一一列举出来，因而就不能像离散型随机变量那样用分布律来描述它．并且，我们通常所遇到的非离散型随机变量取任一指定的实数值的概率都等于零(这一点下节给出证明)．再者，在实际中，对于这样的随机变量，例如，误差 ε，元件的寿命 T 等，我们并不对误差 $\varepsilon=0.01$mm，寿命 $T=1200$h 的概率感兴趣，而是对误差值在某个区间的概率，寿命大于某个数的概率感兴趣．因而我们来研究随机变量的值落在区间$(x_1,x_2]$的概率：$P\{x_1<X\leqslant x_2\}$，由于

$$P\{x_1<X\leqslant x_2\}=P\{X\leqslant x_2\}-P\{X\leqslant x_1\},$$

所以我们只要知道 $P\{X\leqslant x\}$就可以了．为此，引入以下随机变量的分布函数的概念．

定义　设 X 是一个随机变量，x 是任意实数，函数

$$F(x)=P\{X\leqslant x\},\quad -\infty<x<+\infty$$

称为随机变量 X 的**分布函数**．

如果已知随机变量 X 的分布函数 $F(x)$，则随机变量 X 落在区间 $(x_1, x_2]$内的概率为

$$P\{x_1<X\leqslant x_2\}=F(x_2)-F(x_1). \tag{2.3.1}$$

注意：分布函数 $F(x)$是一个普通的函数，$F(x)$在点 x 处的值是随机变量 X 落在区间$(-\infty, x]$内的概率 $P\{X\leqslant x\}$. 分布函数 $F(x)$具有以下性质：

(1) $0\leqslant F(x)\leqslant 1$，且

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

(2) 分布函数 $F(x)$是不减函数．

由公式(2.3.1)知，当 $x_1<x_2$ 时，$F(x_2)-F(x_1)=P\{x_1<X\leqslant x_2\}\geqslant 0$，有 $F(x_1)\leqslant F(x_2)$.

(3) 分布函数 $F(x)$处处右连续(证明略).

例 1　设随机变量 X 的分布律为

X	-1	2	3
P	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

求 X 的分布函数，并求 $P\left\{X\leqslant -\frac{1}{2}\right\}$，$P\left\{-\frac{1}{2}<X\leqslant\frac{1}{2}\right\}$，$P\{2\leqslant X\leqslant 4\}$.

解　随机变量 X 为离散型，它的分布函数为

$$F(x)=\begin{cases}0, & x<-1,\\ P\{X=-1\}, & -1\leqslant x<2,\\ P\{X=-1\}+P\{X=2\}, & 2\leqslant x<3,\\ 1, & x\geqslant 3,\end{cases}$$

即

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

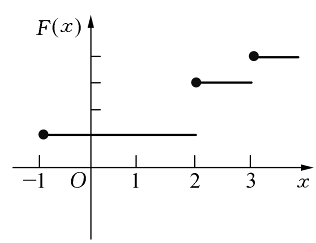


图 2－2

分布函数 $F(x)$的图形是一条阶梯形曲线(图

2-2)，在 $x=-1$，2，3 处分别有跳跃，跳跃值分别为 $\frac{1}{4}$，$\frac{1}{2}$，$\frac{1}{4}$.

$$P\left\{X\leqslant-\frac{1}{2}\right\}=F\left(-\frac{1}{2}\right)=\frac{1}{4},$$

$$P\left\{-\frac{1}{2}<X\leqslant\frac{1}{2}\right\}=F\left(\frac{1}{2}\right)-F\left(-\frac{1}{2}\right)=\frac{1}{4}-\frac{1}{4}=0,$$

$$P\{2\leqslant X\leqslant 4\}=F(4)-F(2)+P\{X=2\}=1-\frac{3}{4}+\frac{1}{2}=\frac{3}{4}.$$

一般地，设离散型随机变量 X 的概率分布律为 $P\{X=x_i\}=p_i$，$i=1$，2，…，则它的分布函数为

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

$F(x)$的图形是一条阶梯形曲线，$F(x)$在任一可能值 x_i 处有跳跃，跳跃值为

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

即 $F(x)$在 $x_i(i=1, 2, \cdots)$处间断，但满足右连续．

例 2　向半径为 1m 的圆形靶射击，击中点落在以靶心为圆心，击中点与靶心距离为半径的圆域的概率与该圆域面积成正比，并且假设不会发生脱靶的情况，设随机变量 X 表示击中点与靶心的距离，试求随机变量 X 的分布函数.

解　X 的分布函数为

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

若 $x<0$，则事件$\{X\leqslant x\}$是不可能事件，于是有 $F(x)=P\{X\leqslant x\}=0$；

若 $x>1$，则事件$\{X\leqslant x\}$是必然事件，于是有 $F(x)=P\{X\leqslant x\}=1$；

若 $0\leqslant x\leqslant 1$，由题意知，$F(x)=P\{X\leqslant x\}=k\pi x^2$，$k$ 为比例系数，为确定 k 的值，取 $x=1$，因为 $P\{0\leqslant X\leqslant 1\}=1^2k\pi$，但 $P\{0\leqslant X\leqslant 1\}=1$，故 $k=\frac{1}{\pi}$. 于是

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

其函数图形是一条连续曲线，如图 2-3 所示．

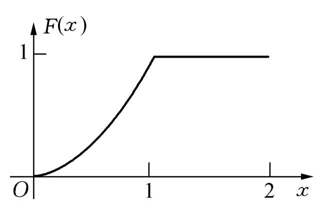


图 2-3

在例 2 中，如果设 $f(t)=\begin{cases}2t, & 0<t<1,\\ 0, & \text{其他},\end{cases}$ 则 X 的分布函数可表示为

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

这就是说，分布函数 $F(x)$等于非负函数 $f(t)$在区间$(-\infty, x]$上的反常积分．在这种情况下，我们称 X 为连续型随机变量．下一节，我们将给出连续型随机变量的一般定义．在实际应用中，我们遇到的基本上是离散型或连续型随机变量，本书主要讨论

这两种随机变量．

2.4　连续型随机变量及其分布

2.4.1　概率密度函数

定义 1　如果对于随机变量 X 的分布函数 $F(x)$，若存在非负函数 $f(x)$，使得对任意实数 x，有

$$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t,\ -\infty<x<+\infty, \tag{2.4.1}$$

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

由(2.4.1)式，根据微积分的知识可知，连续型随机变量 X 的分布函数 $F(x)$ 是连续的．

由定义知道，概率密度 $f(x)$ 具有以下性质：

(1) $f(x)\geqslant 0$；

(2) $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=1$；　　(2.4.2)

(3) 对任意实数 x_1，$x_2(x_1\leqslant x_2)$，有

$$P\{x_1<X\leqslant x_2\}=F(x_2)-F(x_1)=\int_{x_1}^{x_2} f(x)\mathrm{d}x; \tag{2.4.3}$$

(4) 若 $f(x)$ 在点 x 连续，则有 $F'(x)=f(x)$，也就是说，分布函数 $F(x)$ 是概率密度函数 $f(x)$ 的一个原函数．

由性质(2)知道，介于概率密度函数 $f(x)$ 曲线与 x 轴之间的平面图形的面积等于 1(图 2-4)．

由性质(3)知道，连续型随机变量 X 落在区间 $(x_1, x_2]$ 内的概率等于它的概率密度函数 $f(x)$ 在该区间上的积分．几何解释就是：概率 $P\{x_1<X\leqslant x_2\}$ 就是区间 $(x_1, x_2]$ 上概率密度函数 $f(x)$ 之下的曲边梯形的面积(图 2-5)．

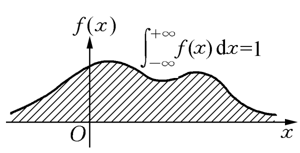


图 2-4

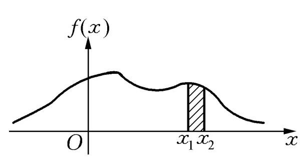


图 2-5

由性质(4)知道，当 $f(x)$ 在点 x 连续时，有

$$f(x)=\lim_{\Delta x\to 0}\frac{F(x+\Delta x)-F(x)}{\Delta x}=\lim_{\Delta x\to 0}\frac{P\{x<X\leqslant x+\Delta x\}}{\Delta x}, \tag{*}$$

从这里，我们看到概率密度的定义与物理学中的线密度的定义相似，这就是称 $f(x)$ 为概率密度的缘故．且由此可得，当 $|\Delta x|$ 很小时，有

$$f(x)\Delta x \approx P\{x < X \leqslant x + \Delta x\},$$

即连续型随机变量 X 落在区间 $(x, x+\Delta x]$ 内的概率近似等于 $f(x)\Delta x$.

需要指出的是，连续型随机变量 X 取任一可能值 x_0 的概率等于零，即

$$P\{X = x_0\} = 0,$$

这是由于 $0 \leqslant P\{X = x_0\} \leqslant P\{x_0 - \Delta x < X \leqslant x_0\} = \int_{x_0-\Delta x}^{x_0} f(x)\mathrm{d}x \to 0$，当 $\Delta x \to 0$ 时，故 $P\{X = x_0\} = 0$. 同时，我们说明，事件 $\{X = x_0\}$ 并非不可能事件，但有 $P\{X = x_0\} = 0$.

据此，在计算连续型随机变量落在某一区间的概率时，可以不必区分该区间是开区间、闭区间还是半开半闭区间．例如，有

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

可见，概率为 0 的事件不一定是不可能事件．

有些随机变量的概率密度函数直接求解较为困难，此时可通过求微小区域上的概率的主要部分(见(＊)式)来求 $f(x)$，这种求解概率密度的方法称作微元法．它是一种重要的方法，利用它可以使一些复杂问题得到简单明了的解决，读者在以后的学习中将会逐步地体会到这一点．

例 1(**柯西分布**)　设连续型随机变量 X 的概率密度为

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

求：(1) 常数 a；(2) X 落在区间 $[-1, 1]$ 内的概率；(3) X 的分布函数．

解　(1) 根据概率密度的性质，有

$$\int_{-\infty}^{+\infty} \frac{a}{1+x^2}\mathrm{d}x = a\pi = 1,$$

由此得

$$a = \frac{1}{\pi},$$

因而，随机变量 X 的概率密度

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

(2) 由公式(2.4.3)，有

$$P\{-1 \leqslant X \leqslant 1\} = \int_{-1}^{1} \frac{1}{\pi(1+x^2)}\mathrm{d}x = \frac{1}{\pi} \times \frac{\pi}{2} = 0.5.$$

(3) 由公式(2.4.1)，有随机变量 X 的分布函数

$$F(x)=\int_{-\infty}^{x}\frac{\mathrm{d}x}{\pi(1+x^2)}=\frac{1}{\pi}\left(\arctan x+\frac{\pi}{2}\right)=\frac{1}{2}+\frac{1}{\pi}\arctan x.$$

例 2　设连续型随机变量 X 的分布函数为

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

求：(1) A，B；(2) X 的密度函数 $f(x)$；(3) $P\left\{X>\frac{1}{3}\right\}$.

解　(1) 因为 $F(x)$为连续函数，所以

$$\lim_{x\to 0^-}F(x)=F(0),\quad \lim_{x\to 1^-}F(x)=F(1),$$

即
$$A=\lim_{x\to 0^-}Ae^{x}=B,\quad B=1-A,$$

从而可得 $A=B=\frac{1}{2}$，因此，随机变量 X 的分布函数为

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

(2) 由连续型随机变量的概率密度与分布函数之间的关系可得

$$f(x)=F'(x)=\begin{cases}\frac{1}{2}e^{x}, & x<0,\\ 0, & 0\leqslant x<1,\\ \frac{1}{2}e^{-(x-1)}, & x\geqslant 1.\end{cases}$$

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

2.4.2　三种常见的连续型随机变量

1. 均匀分布

定义 2　设随机变量 X 的一切可能值充满某一有限区间(a,b)，并且在该区间内的任意点有相同的概率密度，即概率密度 $f(x)$在区间(a,b)内为常数，这种分布称为**均匀分布**(或**等概率分布**).

不难计算在区间(a,b)上服从均匀分布的随机变量 X 的概率密度．事实上，因为在区间(a,b)内概率密度 $f(x)=C$(常数)，而在(a,b)之外为 0，所

以按公式(2.4.2)，有

$$\int_a^b C\mathrm{d}x = C(b-a) = 1,$$

即
$$C=\frac{1}{b-a},$$

于是，我们得到在区间(a,b)上服从均匀分布的随机变量 X 的概率密度

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b,\\ 0, & 其他.\end{cases} \tag{2.4.4}$$

均匀分布含有两个参数 a 及 b，通常记为 $U(a,b)$. 如果随机变量 X 在区间(a,b)上服从均匀分布，则记为 $X\sim U(a,b)$.

由定义可知，在区间(a,b)上服从均匀分布的随机变量 X，落在(a,b)内的任一子区间的概率，只依赖于子区间的长度而与子区间的位置无关. 事实上，对于任一长度为 d 的子区间$(c,c+d)$，$a\leqslant c<c+d\leqslant b$，有

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

由公式(2.4.1)，不难得到在区间(a,b)上服从均匀分布的随机变量 X 的分布函数

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

均匀分布的概率密度及分函数的图形分别如图 2-6 和图 2-7 所示.

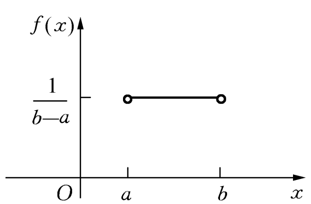


图 2-6

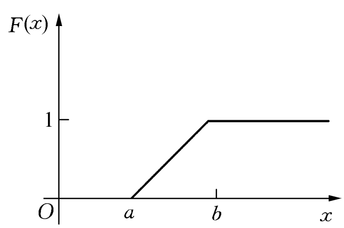


图 2-7

例 3　公共汽车站每隔 5min 有一辆汽车通过，乘客到达汽车站的任一时刻是等可能的，求乘客候车时间不超过 3min 的概率.

解　乘客候车时间记为 X，由题意，随机变量 $X\sim U(0,5)$，则

$$f(x)=\begin{cases}\dfrac{1}{5}, & 0<x<5,\\ 0, & 其他,\end{cases}$$

故有
$$P\{X\leqslant 3\}=\int_0^3\frac{1}{5}\mathrm{d}x=\frac{3}{5}=0.6.$$

2. 指数分布

定义 3　设随机变量 X 的概率密度为
$$f(x)=\begin{cases}\lambda \mathrm{e}^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$
其中 $\lambda>0$ 为常数，则称 X 服从参数为 λ 的**指数分布**，记作 $X\sim E(\lambda)$.

显然 $f(x)\geqslant 0$，且有
$$\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=\int_0^{+\infty}\lambda \mathrm{e}^{-\lambda x}\mathrm{d}x=-\mathrm{e}^{-\lambda x}\Big|_0^{+\infty}=1,$$
按公式(2.4.1)可得，X 的分布函数为
$$F(x)=\begin{cases}1-\mathrm{e}^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

指数分布的随机变量只可能取非负实数，所以指数分布常被用作各种“寿命”分布，譬如，电子元器件的寿命、动物的寿命、电话的通话时间、随机服务系统中的服务时间等都可假定服从指数分布，指数分布在可靠性与排队论中有着广泛的应用．

若随机变量 $X\sim E(\lambda)$，则可以证明 X 具有以下有趣的性质：

对于任意 s，$t>0$，有
$$P\{X>s+t\mid X>s\}=P\{X>t\}.$$
这一性质称为**无记忆性**，如果 X 是某一电子元件的使用寿命，X 服从指数分布，那么指数分布的无记忆性表明已知电子元件已使用了 sh 仍然正常，它总共能使用至少 $(s+t)$h 的条件概率，与从开始使用算起它至少能使用 th 的概率相等．这就是说，电子元件对它已使用 sh 没有记忆．具有这一性质是指数分布有广泛应用的重要原因．

以下例子说明了泊松分布与指数分布的关系．

例 4　很多系统的设备是由于外部冲击而引起失效(或故障)的，假设某设备只受外部冲击才会引起故障，在任何时间区间长度为 t 的时间间隔内由于受着外部冲击而引起故障的次数 $N(t)$ 服从参数为 λt 的泊松分布，则相继两次故障之间的时间间隔 T 服从参数为 λ 的指数分布．

解　由题设可知 $N(t)\sim P(\lambda t)$，即
$$P\{N(t)=k\}=\frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t},\ k=0,1,\cdots.$$

注意到两次故障之间的时间间隔 T 是非负随机变量，且事件 $\{T>t\}$ 说明此设备在 $[0, t]$ 内没有发生故障，即 $\{T>t\}=\{N(t)=0\}$，由此可得

当 $t<0$ 时，有 $F_T(t)=P\{T\leqslant t\}=0$；

当 $t\geqslant 0$ 时，有

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

所以 $T\sim E(\lambda)$，即相继两次故障之间的时间间隔 T 服从参数为 λ 的指数分布 .

3. 正态分布

定义 4　设连续型随机变量 X 的概率密度为

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

其中 μ，$\sigma(-\infty<\mu<+\infty，\sigma>0)$为常数，则称 X 服从参数为 μ，σ 的**正态分布**(或**高斯**(Gauss)**分布**)，记为 $X\sim N(\mu，\sigma^2)$.

在自然现象和社会现象中，大量随机变量都服从或近似服从正态分布 . 例如，一个地区的男性成年人的身高、测量某零件长度的误差、海洋波浪的高度等，因此正态分布是一个重要的分布 .

显然 $f(x)\geqslant 0$，且可以证明 $\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=1$. 事实上，

$$\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t\left(令\frac{x-\mu}{\sigma}=t\right),$$

记 $I=\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t$，则有

$$I^2=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{t^2+u^2}{2}}\mathrm{d}t\mathrm{d}u=\int_0^{2\pi}\int_0^{+\infty}\mathrm{e}^{-\frac{r^2}{2}}r\mathrm{d}r\mathrm{d}\theta=2\pi(用极坐标化累次积分),$$

而 $I>0$，因而有 $I=\sqrt{2\pi}$，即

$$\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t=\sqrt{2\pi},$$

于是

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

正态分布 $N(\mu，\sigma^2)$的概率密度 $f(x)$具有以下性质：

(1) $f(x)$的曲线关于直线 $x=\mu$ 对称(图 2－8)；

(2) 函数 $f(x)$在 $x=\mu$ 处达到最大值 $f(\mu)=\dfrac{1}{\sqrt{2\pi}\sigma}$；

(3) 在 $x=\mu\pm\sigma$ 处有拐点；当 $x\to\pm\infty$时，曲线以 x 轴为其渐近线 .

另外，如果固定 σ，改变 μ 的值，则图形沿着 x 轴平行移动而不改变其形状；如果固定 μ，改变 σ 的值，由于最大值 $f(\mu)=\dfrac{1}{\sqrt{2\pi}\sigma}$，可知当 σ 越小时，图形变得越尖，因而 X 落在 μ 附近的概率越大，如图 2－9 所示 .

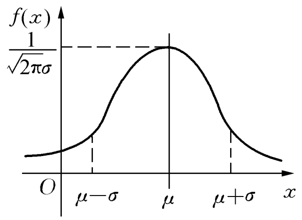


图 2-8

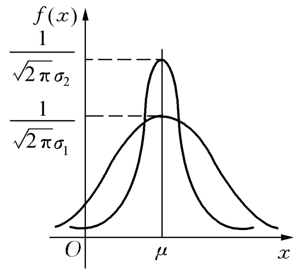


图 2-9

由公式(2.4.1)，正态分布 $N(\mu, \sigma^2)$的分布函数是

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

特别地，当 $\mu=0$，$\sigma=1$ 时，得到的正态分布 $N(0, 1)$，叫作**标准正态分布**，它的概率密度记作

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \quad -\infty<x<+\infty.$$

标准正态分布的分布函数记为 $\Phi(x)$，即有

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

显然，$\Phi(x)$具有下列性质：

(1) $\Phi(0)=0.5$；(2) $\Phi(+\infty)=1$；(3) $\Phi(-x)=1-\Phi(x)$.

由于标准正态分布的分布函数不含任何未知参数，故其值 $\Phi(x)$完全可以算出，附表 2 对 $x\geqslant 0$ 给出了 $\Phi(x)$的值，可供查表用 .

例如，若 $X\sim N(0, 1)$，查表得

$P\{X<0\}=\Phi(0)=0.5$，

$P\{X\leqslant 1.6\}=\Phi(1.6)=0.9452$，

$P\{X<-0.5\}=\Phi(-0.5)=1-\Phi(0.5)=1-0.6915=0.3085$.

为了以后数理统计应用的方便，我们引入上 α 分位点的定义 .

定义 5　设随机变量 $X\sim N(0, 1)$，若 u_α 满足条件

$$P\{X>u_\alpha\}=\alpha, \quad 0<\alpha<1,$$

则称点 u_α 为标准正态分布的**上 α 分位点**(或**上侧临界值**)，如图 2-10 所示 .

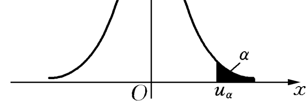


图 2-10

例如，由附表可知

$$u_{0.025}=1.960, \quad u_{0.05}=1.645.$$

现在我们来计算服从一般正态分布 $N(\mu, \sigma^2)$ 的随机变量 X 落在区间 (x_1, x_2) 内的概率，按公式(2.4.3)，我们有

$$P\{x_1 < X < x_2\} = \frac{1}{\sqrt{2\pi}\sigma}\int_{x_1}^{x_2} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x.$$

令 $\frac{x-\mu}{\sigma}=t$，得到

$$P\{x_1 < X < x_2\} = \frac{1}{\sqrt{2\pi}}\int_{\frac{x_1-\mu}{\sigma}}^{\frac{x_2-\mu}{\sigma}} \mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t,$$

则

$$P\{x_1<X<x_2\}=\Phi\left(\frac{x_2-\mu}{\sigma}\right)-\Phi\left(\frac{x_1-\mu}{\sigma}\right). \tag{2.4.5}$$

例 5　设随机变量 $X \sim N(0.5, 4)$，求下列概率：

$$P\{X<0\},\ P\{X\geqslant 5.9\},\ P\{-0.5<x<1.5\}.$$

解　由公式(2.4.5)，我们有

$$P\{X<0\}=\Phi\left(\frac{0-0.5}{2}\right)=\Phi(-0.25)=1-\Phi(0.25)=1-0.5987=0.4013,$$

$$\begin{aligned}P\{X\geqslant 5.9\}&=1-P\{X<5.9\}=1-\Phi\left(\frac{5.9-0.5}{2}\right)=1-\Phi(2.7)\\&=1-0.9965=0.0035,\end{aligned}$$

$$\begin{aligned}P\{-0.5<X<1.5\}&=\Phi\left(\frac{1.5-0.5}{2}\right)-\Phi\left(\frac{-0.5-0.5}{2}\right)\\&=\Phi(0.5)-\Phi(-0.5)=2\Phi(0.5)-1=0.3830.\end{aligned}$$

设随机变量 X 服从正态分布 $N(\mu, \sigma^2)$，由标准正态分布的分布函数表还能得到

$$P\{\mu-\sigma<X<\mu+\sigma\}=\Phi(1)-\Phi(-1)=2\Phi(1)-1=0.6826,$$

$$P\{\mu-2\sigma<X<\mu+2\sigma\}=\Phi(2)-\Phi(-2)=2\Phi(2)-1=0.9544,$$

$$P\{\mu-3\sigma<X<\mu+3\sigma\}=\Phi(3)-\Phi(-3)=2\Phi(3)-1=0.9974.$$

我们看到，尽管正态随机变量的取值范围是$(-\infty, +\infty)$，但它的值落在区间$(\mu-3\sigma, \mu+3\sigma)$内几乎是肯定的事．这就是人们所说的“$3\sigma$”原则．正态分布的 3σ 原则在实际工作中很有用，工业生产上用的控制图和一些产品质量指数(如过程能力指数 C_p，C_{pk})都是根据 3σ 原则制定的．

例 6　从某市南郊某地乘车到北区火车站有两条路可走，第一条路较短，但交通拥挤，所需时间 X 服从分布 $N(50, 10^2)$；第二条路线略长，但意外阻塞较少，所需时间 Y 服从分布 $N(60, 4^2)$.

(1) 若有 70min 可用，问应走哪一条路？

(2) 若只有 65min 可用，又应走哪一条路？

解　(1) $P\{X\leqslant 70\}=\Phi\left(\frac{70-50}{10}\right)=\Phi(2)=0.9772$，

$$P\{Y\leqslant 70\}=\Phi\left(\frac{70-60}{4}\right)=\Phi(2.5)=0.9938,$$

即 $P\{Y\leqslant 70\}>P\{X\leqslant 70\}$，所以此时应选择走第二条路．

(2) $P\{X\leqslant 65\}=\Phi\left(\frac{65-50}{10}\right)=\Phi(1.5)=0.9332$，

$$P\{Y\leqslant 65\}=\Phi\left(\frac{65-60}{4}\right)=\Phi(1.25)=0.8944,$$

即 $P\{X\leqslant 65\}>P\{Y\leqslant 65\}$，所以此时应选择走第一条路．

2.5　随机变量函数的分布

实际问题中，我们往往要讨论随机变量函数的分布，例如，设随机变量 X 是车床加工出来的轴的直径，而轴的横断面的面积 $Y=g(X)=\frac{\pi}{4}X^2$，Y 也是随机变量，我们称 Y 是随机变量 X 的函数．本节，我们将讨论如何由已知随机变量 X 的分布去求得它的函数 $Y=g(X)$(这里 $g(x)$为普通的连续函数)的分布．

2.5.1　离散型随机变量函数的分布

设离散型随机变量 X 的概率分布为

X	x_1	x_2	x_3	…	x_k	…
P	p_1	p_2	p_3	…	p_k	…

$Y=g(X)$为随机变量 X 的函数，当随机变量取得它的某一可能值 x_k 时，随机变量 $Y=g(X)$取 $y_k=g(x_k)$，如果 y_k 的值互不相等，则 $P\{Y=y_k\}=P\{X=x_k\}(k=1，2，\cdots)$，得随机变量 $Y=g(X)$的概率分布为

$Y=g(X)$	$y_1=g(x_1)$	$y_2=g(x_2)$	…	$y_k=g(x_k)$	…
P	p_1	p_2	…	p_k	…

若有 $x_i\neq x_j$ 而 $g(x_i)=g(x_j)(i\neq j)$，则应把那些相等的值分别合并起来，并根据概率的加法公式把对应的概率相加，就可得 $Y=g(X)$的概率分布．

例 1　设随机变量 X 的概率分布律为

X	-1	0	1	2
p_k	0.2	0.3	0.1	0.4

求：(1) $Y=X^2$ 的概率分布律；(2) $Y=2X$ 的概率分布律．

解　(1) 随机变量 Y 的可能值为 0，1，4，且有

$$P\{Y=0\}=P\{X=0\}=0.3,$$
$$P\{Y=1\}=P\{X=\pm 1\}=0.2+0.1=0.3,$$
$$P\{Y=4\}=P\{X=\pm 2\}=0.4+0=0.4,$$

因而随机变量函数 $Y=X^2$ 的分布律为

Y	0	1	4
p_k	0.3	0.3	0.4

(2) 随机变量 Y 的可能值为 -2，0，2，4，且有

$$P\{Y=-2\}=P\{X=-1\}=0.2,$$
$$P\{Y=0\}=P\{X=0\}=0.3,$$
$$P\{Y=2\}=P\{X=1\}=0.1,$$
$$P\{Y=4\}=P\{X=2\}=0.4,$$

因而随机变量函数 $Y=2X$ 的分布律为

Y	-2	0	2	4
p_k	0.2	0.3	0.1	0.4

例 2　设随机变量 X 的概率分布律为

$$P\{X=k\}=\frac{1}{2^k},\ k=1,2,\cdots,$$

求随机变量 $Y=\sin\left(\frac{\pi}{2}X\right)$ 的概率分布律．

解　因为

$$\sin\left(\frac{n\pi}{2}\right)=\begin{cases}-1, & n=4k-1,\\ 1, & n=4k-3,\ k=1,2,\cdots,\\ 0, & n=2k,\end{cases}$$

所以，函数 $Y=\sin\left(\frac{\pi}{2}X\right)$ 只有三个可能值：-1，0，1，且有

$$P\{Y=-1\}=\frac{1}{2^3}+\frac{1}{2^7}+\frac{1}{2^{11}}+\cdots=\frac{1}{8\left(1-\frac{1}{16}\right)}=\frac{2}{15},$$

$$P\{Y=0\}=\frac{1}{2^2}+\frac{1}{2^4}+\frac{1}{2^6}+\cdots=\frac{1}{4\left(1-\frac{1}{4}\right)}=\frac{1}{3},$$

$$P\{Y=1\}=\frac{1}{2}+\frac{1}{2^5}+\frac{1}{2^9}+\cdots=\frac{1}{2\left(1-\frac{1}{16}\right)}=\frac{8}{15},$$

因而随机变量 Y 的概率分布律为

Y	-1	0	1
p_k	$\frac{2}{15}$	$\frac{1}{3}$	$\frac{8}{15}$

2.5.2　连续型随机变量函数的分布

假设函数 $g(x)$ 及其一阶导数在随机变量 X 的一切可能值 x 的区间内是连续的．我们来讨论如何由随机变量 X 的概率密度 $f_X(x)$，得到随机变量 $Y=g(X)$ 的概率密度 $f_Y(y)$．下面介绍两种求 Y 的概率密度 $f_Y(y)$ 方法．

1. 分布函数法

分布函数法就是通过随机变量 $Y=g(X)$，与随机变量 X 的概率密度 $f_X(x)$，先求出随机变量 Y 的分布函数 $F_Y(y)$，再由 $F'_y(y)=f_Y(y)$，计算出随机变量 Y 的概率密度 $f_Y(y)$．

例 3　设随机变量 X 服从区间 $[0，4]$ 上的均匀分布，求随机变量 $Y=2X+1$ 的概率密度．

解　记随机变量 X，Y 的分布函数分别为 $F_X(x)$，$F_Y(y)$，因为 X 的概率密度为

$$f_X(x)=\begin{cases}\frac{1}{4}, & 0<x<4,\\ 0, & \text{其他},\end{cases}$$

则 Y 的分布函数

$$F_Y(y)=P\{Y\leqslant y\}=P\{2X+1\leqslant y\}=P\left\{X\leqslant\frac{y-1}{2}\right\}=F_X\left(\frac{y-1}{2}\right).$$

将分布函数 $F_Y(y)$ 关于 y 求导，得 Y 的概率密度

$$\begin{aligned}f_Y(y)&=F'_X\left(\frac{y-1}{2}\right)\cdot\left(\frac{y-1}{2}\right)'=\frac{1}{2}f_X\left(\frac{y-1}{2}\right)\\&=\begin{cases}\frac{1}{2}\times\frac{1}{4}, & 0<\frac{y-1}{2}<4,\\ 0, & \text{其他}\end{cases}\\&=\begin{cases}\frac{1}{8}, & 1<y<9,\\ 0, & \text{其他}.\end{cases}\end{aligned}$$

例 4　设随机变量 X 服从标准正态分布 $N(0，1)$，求随机变量 $Y=X^2$ 的

概率密度 .

解　记随机变量 X，Y 的分布函数分别为 $F_X(x)$，$F_Y(y)$. 因为 X 的概率密度为

$$f_X(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}},\ -\infty<x<+\infty,$$

所以 Y 的分布函数为

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

当 $y\leqslant 0$ 时，$F_Y(y)=P\{Y\leqslant y\}=P\{X^2\leqslant y\}=0$；

当 $y>0$ 时，$F_Y(y)=P\{Y\leqslant y\}=P\{X^2\leqslant y\}=P\left\{-\sqrt{y}\leqslant X\leqslant\sqrt{y}\right\}$

$$=F_X(\sqrt{y})-F_X(-\sqrt{y}).$$

将分布函数 $F_Y(y)$ 对 y 求导，得 Y 的概率密度

$$f_Y(y)=\begin{cases}\dfrac{1}{2\sqrt{y}}f_X(\sqrt{y})+\dfrac{1}{2\sqrt{y}}f_X(-\sqrt{y}), & y>0,\\ 0, & y\leqslant 0\end{cases}$$

$$=\begin{cases}\dfrac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}\mathrm{e}^{-\frac{y}{2}}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

例 5　设随机变量 X 的概率密度函数为

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

求随机变量 $Y=\sin X$ 的概率密度 .

解　设随机变量 $Y=\sin X$ 的分布函数为 $F_Y(y)$，则

$$F_Y(y)=P\{Y\leqslant y\}=P\{\sin X\leqslant y\}.$$

当 $y<0$ 时，$F_Y(y)=0$；

当 $y>1$ 时，$F_Y(y)=1$；

当 $0\leqslant y\leqslant 1$ 时，由图 2 - 11 可知

$$F_Y(y)=P\{\sin X\leqslant y\}=P\{0\leqslant X\leqslant \arcsin y\}+P\{\pi-\arcsin y\leqslant X\leqslant\pi\}$$

$$=\int_0^{\arcsin y}\frac{2x}{\pi^2}\mathrm{d}x+\int_{\pi-\arcsin y}^{\pi}\frac{2x}{\pi^2}\mathrm{d}x$$

$$=\frac{2}{\pi}\arcsin y.$$

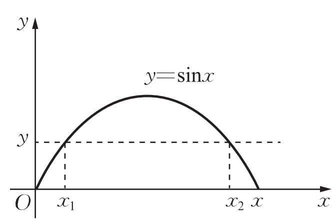


图 2 - 11

所以，随机变量 $Y=\sin X$ 的概率密度函数为

$$f_Y(y)=F'_Y(y)=\begin{cases}\dfrac{2}{\pi\sqrt{1-y^2}}, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

***2. 微元法**

有些连续型随机变量的概率密度函数直接求解较为困难，此时可通过求微小区域$\{x<X<x+\Delta x\}$上的概率的主要部分来求该随机变量 X 的概率密度 $f_X(x)$，这种求解概率密度的方法称作微元法．它是一种重要的方法，利用它可以使一些复杂问题得到简单明了的解决．我们可以通过下面例题来熟悉这种方法．

例 6　设 $X=R\cos\theta$，其中 R 是已知的正常数，而随机变量 θ 在$(0,2\pi)$上服从均匀分布 $U(0,2\pi)$，试求 X 的概率密度．

解　对任意 $x\in(-R,R)$，取 $h>0$ 充分小，使 $x+h\in(-R,R)$，

$$\begin{aligned}&P\{x<X\leqslant x+h\}=P\{x<R\cos\theta\leqslant x+h\}\\&=P\left\{\arccos\frac{x+h}{R}\leqslant\theta<\arccos\frac{x}{R}\right\}+P\left\{2\pi-\arccos\frac{x}{R}<\theta\leqslant 2\pi-\arccos\frac{x+h}{R}\right\}\\&=\frac{1}{2\pi}\left(\arccos\frac{x}{R}-\arccos\frac{x+h}{R}\right)+\frac{1}{2\pi}\left(\arccos\frac{x}{R}-\arccos\frac{x+h}{R}\right)\\&=\frac{1}{\pi}\left(\arccos\frac{x}{R}-\arccos\frac{x+h}{R}\right),\end{aligned}$$

所以

$$f_X(x)=\lim_{h\to 0}\frac{P\{x<X\leqslant x+h\}}{h}=\lim_{h\to 0}\frac{\arccos\dfrac{x}{R}-\arccos\dfrac{x+h}{R}}{\pi h}=\frac{1}{\pi\sqrt{R^2-x^2}},$$

因此，X 的概率密度为

$$f_X(x)=\begin{cases}\dfrac{1}{\pi\sqrt{R^2-x^2}}, & -R<x<R,\\ 0, & \text{其他}.\end{cases}$$

3. 公式法

以下我们仅对 $Y=g(X)$，其中 $g(x)$是严格单调函数，如何由随机变量 X 的概率密度 $f_X(x)$得到随机变量 $Y=g(X)$的概率密度 $f_Y(y)$，最后给出一般的结果．

定理 1　设随机变量 X 具有概率密度 $f_X(x)(-\infty<x<+\infty)$，又设函数 $g(x)$处处可导且恒有 $g'(x)>0$(或恒有 $g'(x)<0$)，则 $Y=g(X)$是连续型随机变量，其概率密度为

$$f_Y(y)=\begin{cases}f_X[h(y)]\,|h'(y)|, & \alpha<y<\beta,\\ 0, & \text{其他},\end{cases}\qquad(2.5.1)$$

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

证明　仅证 $g'(x)>0$ 的情形．设 $g'(x)>0$，此时 $g(x)$ 在 $(-\infty,\ +\infty)$ 内严格单调增加，则它的反函数 $h(y)$ 存在，且在 $(\alpha,\ \beta)$ 内严格单调增加，并且可导．记随机变量 X，Y 的分布函数分别为 $F_X(x)$，$F_Y(y)$.

因为 $Y=g(X)$ 在 $(\alpha,\ \beta)$ 内取值，因而有

当 $y\leqslant\alpha$ 时，$F_Y(y)=P\{Y\leqslant y\}=0$；

当 $y\geqslant\beta$ 时，$F_Y(y)=P\{Y\leqslant y\}=1$.

当 $\alpha<y<\beta$ 时，$F_Y(y)=P\{Y\leqslant y\}=P\{g(X)\leqslant y\}$
$=P\{X\leqslant h(y)\}=F_X[h(y)]$（因为 $g(x)$ 严格单调增加）.

将 $F_Y(y)$ 关于 y 求导数，得

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

同理，如果 $g'(x)<0$，则有

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

综合上述，我们有

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

我们特别指出，若 $f_X(x)$ 在有限区间 $[a,\ b]$ 以外等于零，则只需假设在 $[a,\ b]$ 上恒有 $g'(x)>0$（或 $g'(x)<0$），此时，$\alpha=\min\{g(a+0),\ g(b-0)\}$，$\beta=\max\{g(a+0),\ g(b-0)\}$.

例 7　设随机变量 $X\sim N(\mu,\ \sigma^2)$，求随机变量 $Y=aX+b\ (a\neq0)$ 的概率密度.

解　因为随机变量 X 的概率密度为

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

函数 $y=g(x)=ax+b$，其反函数为 $x=h(y)=\dfrac{y-b}{a}$，且有 $h'(y)=\dfrac{1}{a}$，$\alpha=-\infty$，$\beta=+\infty$，由公式(2.5.1)得，Y 的概率密度

$$f_Y(y)=\frac{1}{|a|}f_X\left(\frac{y-b}{a}\right),\quad -\infty<y<+\infty,$$

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

由例 7 的结果可知，若随机变量 $X\sim N(\mu,\ \sigma^2)$，则 $Y=aX+b\sim N(a\mu+b,\ a^2\sigma^2)$. 特别地，若 $X\sim N(\mu,\ \sigma^2)$，则 $Z=\dfrac{X-\mu}{\sigma}\sim N(0,\ 1)$，若 $X\sim N(0,\ 1)$，则 $Z=\sigma X+\mu\sim N(\mu,\ \sigma^2)$.

例 8　设随机变量 X 服从区间 $[0,\ 4]$ 上的均匀分布，求 $Y=\mathrm{e}^X$ 的概率密度．

解　因为随机变量 X 的概率密度为

$$f_X(x)=\begin{cases}\dfrac{1}{4}, & 0<x<4,\\ 0, & \text{其他},\end{cases}$$

函数 $y=g(x)=\mathrm{e}^x$ 的反函数为 $x=h(y)=\ln y$，且 $h'(y)=\dfrac{1}{y}$,

$$\alpha=\min\{g(0+0),\ g(4-0)\}=\min\{1,\ \mathrm{e}^4\}=1,$$
$$\beta=\max\{g(0+0),\ g(4-0)\}=\max\{1,\ \mathrm{e}^4\}=\mathrm{e}^4,$$

由公式(2.5.1)得，Y 的概率密度

$$f_Y(y)=\begin{cases}\dfrac{1}{4y}, & 1<y<\mathrm{e}^4,\\ 0, & \text{其他}.\end{cases}$$

定理 2　设随机变量 X 具有概率密度 $f_X(x)(a<x<b$，a 可以是 $-\infty$，b 可以是 $+\infty)$，若 $g(x)$在不相重叠的区间 I_1，I_2，…，I_n 上可导且恒有 $g'(x)>0$（或恒有 $g'(x)<0$），反函数分别为 $h_1(y)$，…，$h_n(y)$，则 $Y=g(X)$是连续型随机变量，其概率密度为

$$f_Y(y)=\begin{cases}\displaystyle\sum_{i=1}^{n}f_X[h_i(y)]\,|h'_i(y)|, & y\in g\ \text{的值域},\\ 0, & \text{其他}.\end{cases}\tag{2.5.2}$$

例 9　设随机变量 X 服从标准正态分布 $N(0,\ 1)$，求随机变量 $Y=|X|$ 的概率密度．

解　$y=|x|$的值域为 $D=[0,\ +\infty)$，该函数在不相重叠的区间$(-\infty,\ 0)$和$(0,\ +\infty)$上可导，且当 $x\in(-\infty,\ 0)$时，$y'=-1<0$，当 $x\in(0,\ +\infty)$时，$y'=1>0$. 在$(-\infty,\ 0)$和$(0,\ +\infty)$上的反函数分别为 $x=-y$ 和 $x=y$，由定理 2 可得 $Y=|X|$的概率密度为

$$f_Y(y)=\begin{cases}f_X(-y)\cdot|-1|+f_X(y)\cdot|1|, & y\in[0,\ +\infty),\\ 0, & \text{其他},\end{cases}$$

即

$$f_Y(y)=\begin{cases}\sqrt{\dfrac{2}{\pi}}\,\mathrm{e}^{-\frac{y^2}{2}}, & y\in[0,\ +\infty),\\ 0, & \text{其他}.\end{cases}$$

习　题　A

1. 一个袋子中装有 5 只球，编号分别为 1，2，3，4，5. 从袋中同时取 3 只，以 X 表示取出的 3 只球中的最大号码，写出随机变量 X 的概率分布律．

2. 一批零件中有 9 只合格品和 3 只废品．每次从这批零件中任取 1 只，如果每次取出的废品不再放回去，求在取得合格品以前已取出的废品数 X 的概率分布律．

3. 自动生产线在调整以后出现废品的概率为 p. 生产过程中出现废品时立即重新进行调整，求在两次调整之间生产的合格品数的概率分布律．

4. 已知一电话总机每分钟收到呼唤的次数 X 服从参数为 4 的泊松分布，求下列事件的概率：

(1) 某 1min 恰有 8 次呼唤；

(2) 某 1min 的呼唤次数不多于 1；

(3) 某 1min 的呼唤次数多于 3.

5. 设事件 A 在每一次试验中发生的概率为 0.3，当事件 A 发生不少于 3 次时，指示灯发出信号．

(1) 进行了 5 次重复独立试验，求指示灯发出信号的概率；

(2) 进行了 7 次重复独立试验，求指示灯发出信号的概率．

6. 一个袋子中装有 5 只球，编号为 1，2，3，4，5. 从袋中同时取 3 只，以 X 表示取出的 3 只球中的最大号码，写出随机变量 X 的分布函数．

7. 在区间 $[0, 4]$ 上任意投掷一个质点，以 X 表示这个质点的坐标，设这个质点落在 $[0, 4]$ 中任意小区间内的概率与这个小区间的长度成正比，试求 X 的分布函数．

8. 函数 $F(x)=\dfrac{1}{1+x^2}$ 可否是连续型随机变量 X 的分布函数，如果 X 的可能值充满区间：(1) $(-\infty, +\infty)$；(2) $(-\infty, 0)$？

9. 设连续型随机变量 X 的概率密度为

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

求：(1) 常数 A；(2) X 落在区间 $(0, 0.5)$ 内的概率；(3) X 的分布函数．

10.（拉普拉斯分布）设随机变量 X 的概率密度为

$$f(x)=a\mathrm{e}^{-|x|}, \quad -\infty<x<+\infty,$$

求：(1) 系数 a；(2) 随机变量 X 落在区间 $(0, 1)$ 内的概率；(3) X 的分布函数．

11. 设连续型随机变量 X 的分布函数为

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

求：(1) 常数 A；(2) $P\{X<2\}$ 与 $P\{2<X<3\}$；(3) X 的概率密度.

12. 设随机变量 X 服从标准正态分布 $N(0, 1)$，求下列概率：

(1) $P\{X\leqslant 0\}$；(2) $P\{X>2.5\}$；(3) $P\{|X|<1.68\}$.

13. 设随机变量 X 服从正态分布 $N(1, 2^2)$，求下列概率：

(1) $P\{X<2.2\}$；(2) $P\{-1.6<X<5.8\}$；(3) $P\{|X|<3.5\}$.

14. 已知某种零件直径(单位：mm)服从正态分布 $N(100, 0.6^2)$，规定直径在范围 100 ± 1.2(mm)内为合格品，求这种零件的不合格率.

15. 一工厂生产的某种电子元件的寿命 X(单位：h)服从参数 $\mu=160$，σ 的正态分布，若要求 $P\{120<X\leqslant 200\}\geqslant 0.80$，允许 σ 最大为多少？

16. 设随机变量 X 的概率分布律为

X	-2	-1	0	1	2
p_k	0.1	0.2	0.3	0.2	0.2

求：(1) $Y=X^2$ 的概率分布律；(2) $Y=2X-1$ 的概率分布律.

17. 设随机变量 X 服从区间(1, 2)上的均匀分布，求随机变量 $Y=2X$ 的概率密度.

18. 设随机变量 X 服从指数分布 $E(0.5)$，求随机变量 $Y=X^2$ 的概率密度.

19. 设随机变量 X 服从标准正态分布 $N(0, 1)$，求随机变量 $Y=2X+1$ 的概率密度.

20. 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{A}{1+x^2}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

求：(1) A；(2) 随机变量 $Y=\ln X$ 的概率密度.

21. 设随机变量 X 服从正态分布 $N(0, \sigma^2)$，求随机变量 $Y=|X|$ 的概率密度.

22. 设电流 I 是一个随机变量，它均匀分布在 9～11A 之间，若此电流通过 2Ω 的电阻，在其上消耗的功率 $W=2I^2$，求随机变量 W 的概率密度.

习　题　B

1.(超几何分布)已知某批产品共 100 个，其中 10 个次品，从中随机地抽取 5 个样品，求抽取 5 个样品中次品数 X 的概率分布律，并求 X 的分布函数.

2. 在伯努利试验中，如果每次试验成功的概率为 p.（1）将试验进行到出现 r 次成功为止，求试验次数 Y 的分布律；（2）将试验进行到成功与失败都出现为止，求试验次数 Y 的分布律.

3. 设随机变量 X 服从指数分布 $E(\lambda)$，求随机变量 $Y=\min\{X, 2\}$ 的分布函数.

4. 设随机变量 X 在区间 $(0, 3)$ 上服从均匀分布，求关于变量 a 的方程

$$4a^2+4Xa+X+2=0$$

无实根的概率.

5. 设随机变量 X 服从正态分布 $N(\mu_1, \sigma_1^2)$，随机变量 Y 服从正态分布 $N(\mu_2, \sigma_2^2)$，且有 $P\{|X-\mu_1|<1\}>P\{|Y-\mu_2|<1\}$，则参数 $\sigma_1<\sigma_2$ 是否成立?

6.（伽玛分布）设随机变量 X 的概率密度为

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

其中 $\alpha>0$，$\beta>0$ 都是常数，求系数 A. 当 $\alpha=1$ 时，该分布为什么分布?

7. 设随机变量 X 服从参数为 λ 的指数分布 $E(\lambda)$，试求 $Y=\sqrt[m]{X}+\mu$（$m>0$，$\mu>0$ 为已知常数）的概率分布.

8. 每天某种商品的销售量（单位：件）服从参数为 λ 的泊松分布，随机选取 4 天，其中恰有一天的销售量为 5 件的概率.

9. 设随机变量 X 的概率密度函数为

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

以 Y 表示 X 进行 3 次独立重复观察中事件 $\left\{X\leqslant\dfrac{1}{2}\right\}$ 出现的次数，求概率 $P\{Y=2\}$.

10. 设随机变量 $X\sim U\left(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right)$，求随机变量 $Y=\cos X$ 的概率密度.

11. 某人上班，从家里到单位要经过一交通指示灯，这指示灯有 $\dfrac{1}{3}$ 时间亮红灯，他遇到红灯时要在指示灯旁等待直至绿灯亮，等待时间（单位：s）服从区间 $(0, 30)$ 上的均匀分布，设 X 表示他的等待时间，求 X 的分布函数. 问 X 是否是连续型随机变量? 又是否是离散型随机变量? 请说明理由.

12. 设随机变量 X 服从指数分布 $E(\lambda)$.（1）求随机变量 $Y=[X]$ 的分布律；（2）Y 服从什么分布? 连续型随机变量的函数是否一定还是连续型随机变量?

13. 已知某种昆虫的产卵数 X 服从泊松分布 $P(\lambda)$，而每个卵能孵化成幼虫的概率为 p，且各卵的孵化是相互独立的，试求该昆虫能育成的幼虫数 Y 所服从的概率分布.

第 3 章　多维随机变量及其分布

上一章我们所讨论的随机变量都是一维的，即一个随机变量，也就是说，它们的值仅由一个数来表示．但在实际问题中，对于某些试验结果仅用一个随机变量描述是远远不够的，而需要同时用两个或两个以上的随机变量来描述．例如，考察某地区大学生身体状况时，可用两个随机变量 X，Y 分别描述大学生的身高和体重，我们称(X, Y)为二维随机变量，用(X, Y)描述他们的身体状况．本章仅讨论二维随机变量，至于三维或更多维的情形不难类推．

3.1　二维随机变量

3.1.1　二维随机变量及其分布函数

定义 1　设 E 为一个随机试验，它的样本空间是 $\Omega=\{\omega\}$，设 $X=X(\omega)$和 $Y=Y(\omega)$是定义在 Ω 上的两个随机变量，则称(X, Y)为**二维随机变量**．

注意：二维随机变量(X, Y)的性质不仅与 X 及 Y 有关，还依赖于 X 与 Y 之间的相互关系，我们应将(X, Y)作为一个整体来研究．和一维的情况类似，我们也借助分布函数来研究二维随机变量．

定义 2　设(X, Y)为二维随机变量，对于任意实数 x，y，二元函数

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

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

在几何上，如果将二维随机变量(X, Y)看作是平面上具有随机坐标(X, Y)的点，则分布函数 $F(x, y)$在点(x, y)处的分布函数值，就是随机点(X, Y)落在以点(x, y)为顶点而位于该点左下方的无限矩形区域内的概率(图3－1).

借助于图 3－2 可以计算出随机点(X, Y)落在矩形$\{(x, y)\mid x_1<X\leqslant x_2, y_1<Y\leqslant y_2\}$内的概率为

$$\begin{aligned}&P\{x_1<X\leqslant x_2, y_1<Y\leqslant y_2\}\\&=F(x_2, y_2)-F(x_2, y_1)-F(x_1, y_2)+F(x_1, y_1).\qquad(3.1.1)\end{aligned}$$

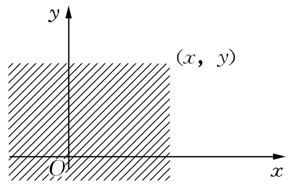


图 3 - 1

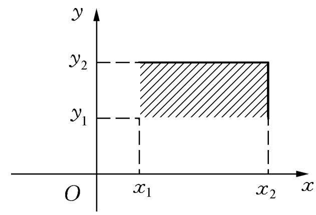


图 3 - 2

二维随机变量的分布函数 $F(x, y)$ 具有以下性质：

(1) 对于任意实数 x，y，有 $0 \leqslant F(x, y) \leqslant 1$，且

对于固定的 y，$F(-\infty, y)=\lim\limits_{x \to -\infty} F(x, y)=0$；

对于固定的 x，$F(x, -\infty)=\lim\limits_{y \to -\infty} F(x, y)=0$.

$$F(-\infty, -\infty)=\lim_{\substack{x \to -\infty \\ y \to -\infty}} F(x, y)=0, \quad F(+\infty, +\infty)=\lim_{\substack{x \to +\infty \\ y \to +\infty}} F(x, y)=1.$$

(2) $F(x, y)$ 是变量 x 和 y 的不减函数，即对于任意固定的 y，当 $x_1<x_2$ 时，$F(x_1, y) \leqslant F(x_2, y)$；同样地，对于任意固定的 x，当 $y_1<y_2$ 时，$F(x, y_1) \leqslant F(x, y_2)$.

(3) $F(x, y)$ 关于 x，y 均右连续，即

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

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

任一二维分布函数 $F(x, y)$ 必须具备上述四条性质，其中性质(4)是二维场合特有的，也是合理的，但性质(4)不能由前三条性质推出，必须单独列出，因为存在这样的二元函数 $G(x, y)$ 满足以上性质(1)、(2)、(3)，但它不满足性质(4)，见下面的例子．

例 1　二元函数

$$G(x, y)=\begin{cases} 0, & x+y<0, \\ 1, & x+y \geqslant 0 \end{cases}$$

满足二维分布函数的性质(1)、(2)、(3)，但它不满足性质(4). 这从 $G(x, y)$ 的定义可看出：若用直线 $x+y=0$ 将平面 xOy 一分为二，则 $G(x, y)$ 在右上半平面($x+y \geqslant 0$)取值为 1，在左下半平面($x+y<0$)取值为 0. $G(x, y)$ 具有单调不减性、有界性和右连续型，但在正方形区域 $\{(x, y) \mid -1 \leqslant x \leqslant 1, -1 \leqslant y \leqslant 1\}$ 的四个顶点上，右上三个顶点位于右上半闭平面，只有左下顶点 $(-1, -1)$ 位于左下半开平面，故有

$$G(1, 1)-G(1, -1)-G(-1, 1)+G(-1, -1)=1-1-1+0=-1<0,$$

所以 $G(x, y)$ 不满足性质(4)，故 $G(x, y)$ 不能成为某二维随机变量的分布函数．

以上关于二维随机变量的讨论不难推广到 $n(n>2)$ 维随机变量的情况．一般地，设 E 为一个随机试验，它的样本空间是 $\Omega=\{\omega\}$，设 $X_1=X_1(\omega)$，$X_2=X_2(\omega)$，…，$X_n=X_n(\omega)$ 是定义在 Ω 上的随机变量，由它们构成一个 n 维向量 $(X_1, X_2, \cdots, X_n)$ 叫作 **n 维随机向量**或 **n 维随机变量**．

对于任意 n 个实数 $x_1, x_2, \cdots, x_n$，n 元函数

$$F(x_1, x_2, \cdots, x_n)=P\{X_1\leqslant x_1, X_2\leqslant x_2, \cdots, X_n\leqslant x_n\}$$

称为 n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 的**分布函数**或随机变量 $X_1, X_2, \cdots, X_n$ 的**联合分布函数**．它具有类似于二维随机变量分布函数的性质．

3.1.2　二维离散型随机变量及其分布

定义 3　如果二维随机变量 (X, Y) 全部可能取到的值是有限对或可列无限多对，则称 (X, Y) 为**二维离散型随机变量**．

显然，当 (X, Y) 为二维离散型随机变量时，X，Y 均为离散型随机变量，对二维离散型随机变量最直接的描述是给出其所有取值及每一对可能取值的概率.

定义 4　设二维离散型随机变量 (X, Y) 的所有可能取值为 (x_i, y_j)，$i, j=1, 2, \cdots$，且

$$P\{X=x_i, Y=y_j\}=p_{ij}, \quad i, j=1, 2, \cdots, \tag{3.1.2}$$

则称(3.1.2)式为二维离散型随机变量 (X, Y) 的**分布律**，或 X，Y 的**联合分布律**.

显然，$p_{ij}\geqslant 0$，$\sum\limits_i\sum\limits_j p_{ij}=1$．为了直观，有时也可用表格来表示 X，Y 的联合分布律，见表 3－1.

表 3－1

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

例 2　一个袋中有 5 个球，其中 3 个红球，2 个白球，每次从中随机抽取一个，抽后不放回，连抽两次，设

$$X=\begin{cases}1, & \text{若第一次抽到红球,}\\ 0, & \text{若第一次抽到白球,}\end{cases}\quad Y=\begin{cases}1, & \text{若第二次抽到红球,}\\ 0, & \text{若第二次抽到白球,}\end{cases}$$

求：(1) 随机变量 X，Y 的联合分布律；(2) 概率 $P\{X\geqslant Y\}$.

解　(1) 由题可知(X, Y)的可能取值为$(0, 0)$，$(0, 1)$，$(1, 0)$，$(1, 1)$，而

$$P\{X=0, Y=0\}=P\{X=0\}P\{Y=0\mid X=0\}=\frac{2}{5}\times\frac{1}{4}=\frac{1}{10}.$$

同理可求得

$$P\{X=0, Y=1\}=\frac{3}{10},\ P\{X=1, Y=0\}=\frac{3}{10},\ P\{X=1, Y=1\}=\frac{3}{10}.$$

于是(X, Y)的分布律为

Y \ X	0	1
0	$\frac{1}{10}$	$\frac{3}{10}$
1	$\frac{3}{10}$	$\frac{3}{10}$

(2) $P\{X\geqslant Y\}=1-P\{X<Y\}=1-P\{X=0, Y=1\}=1-\dfrac{3}{10}=\dfrac{7}{10}.$

例 3　设随机变量 X 在 1，2，3，4 四个整数中等可能地取值，而随机变量 Y 在 $1\sim X$ 中等可能地取一个整数值，求：

(1) 随机变量 X，Y 的联合分布律；(2) 概率 $P\{Y=2\}$.

解　(1) 由乘法公式容易求得(X, Y)的分布律．易知$\{X=i, Y=j\}$的取值情况是 $i=1$，2，3，4，j 取不大于 i 的整数，且

$$P\{X=i, Y=j\}=P\{X=i\}P\{Y=j\mid X=i\}=\frac{1}{4}\times\frac{1}{i},\ i=1, 2, 3, 4,\ j\leqslant i,$$

于是(X, Y)的分布律为

Y \ X	1	2	3	4
1	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{16}$
2	0	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{16}$
3	0	0	$\frac{1}{12}$	$\frac{1}{16}$
4	0	0	0	$\frac{1}{16}$

（2）$P\{Y=2\}=P\{X=2,\ Y=2\}+P\{X=3,\ Y=2\}+P\{X=4,\ Y=2\}$

$$=\frac{1}{8}+\frac{1}{12}+\frac{1}{16}=\frac{13}{48}.$$

设二维离散型随机变量$(X,\ Y)$的分布律为 $P\{X=x_i,\ Y=y_j\}=p_{ij}$，$i$，$j=1,2,\cdots$，由分布函数的定义知，$(X,\ Y)$的分布函数为

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

这里的和式是对满足不等式 $x_i\leqslant x$，$y_j\leqslant y$ 的一切 i，j 来求和的.

3.1.3　二维连续型随机变量的概率密度函数

定义 5　设$(X,\ Y)$为二维随机变量，其分布函数为 $F(x,\ y)$，如果存在一个非负可积二元函数 $f(x,\ y)$，使得对于任意实数 x，y 有

$$F(x,\ y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(x,\ y)\mathrm{d}x\mathrm{d}y,\qquad(3.1.3)$$

则称$(X,\ Y)$为二维连续型随机变量，函数 $f(x,\ y)$称为二维连续型随机变量$(X,\ Y)$的**概率密度函数**，或称为随机变量 X，Y 的**联合概率密度函数**.

概率密度函数 $f(x,\ y)$具有以下性质：

（1）$f(x,\ y)\geqslant 0$；

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

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

$$P\{(X,\ Y)\in D\}=\iint_D f(x,\ y)\mathrm{d}x\mathrm{d}y;\qquad(3.1.4)$$

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

例 4　设二维随机变量$(X,\ Y)$的概率密度为

$$f(x,\ y)=\begin{cases}k\mathrm{e}^{-(3x+y)}, & x>0,\ y>0,\\ 0, & \text{其他},\end{cases}$$

求：（1）常数 k；（2）$(X,\ Y)$的分布函数；（3）概率 $P\{Y\leqslant X\}$.

解　（1）因为$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,\ y)\mathrm{d}x\mathrm{d}y=1$，所以$\int_{0}^{+\infty}\int_{0}^{+\infty}k\mathrm{e}^{-(3x+y)}\mathrm{d}x\mathrm{d}y=1$，则 $k=3$.

（2）由分布函数的定义 $F(x,\ y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(x,\ y)\mathrm{d}x\mathrm{d}y$，得

当 $x>0$，$y>0$ 时，$F(x,\ y)=\int_{0}^{x}\int_{0}^{y}3\mathrm{e}^{-(3u+v)}\mathrm{d}u\mathrm{d}v=(1-\mathrm{e}^{-3x})(1-\mathrm{e}^{-y})$；

当 $x\leqslant 0$ 或 $y\leqslant 0$ 时，$F(x,\ y)=0$.

因此，$(X,\ Y)$的分布函数为

$$F(x,\ y)=\begin{cases}(1-\mathrm{e}^{-3x})(1-\mathrm{e}^{-y}), & x>0,\ y>0,\\ 0, & \text{其他}.\end{cases}$$

(3) $P\{Y\leqslant X\}=\iint\limits_{x>y\geqslant 0}3\mathrm{e}^{-(3x+y)}\mathrm{d}x\mathrm{d}y=\int_0^{+\infty}\int_y^{+\infty}3\mathrm{e}^{-(3x+y)}\mathrm{d}x\mathrm{d}y=\dfrac{1}{4}$.

下面介绍两种常见的二维连续型随机变量.

1. 二维均匀分布

定义 6 设 D 为平面上的有界区域，其面积为 S，且 $S>0$，如果二维随机变量$(X,\ Y)$的概率密度函数为 $f(x,\ y)=\begin{cases}\dfrac{1}{S}, & (x,\ y)\in D,\\ 0, & \text{其他},\end{cases}$则称$(X,\ Y)$服从区域 D 上的均匀分布.

例 5 设$(X,\ Y)$服从 D 上的均匀分布，其中 $D=\{(x,\ y)\mid 0\leqslant y\leqslant x\leqslant 1\}$，求：

(1) $(X,\ Y)$的概率密度函数；

(2) 概率 $P\{X+Y\leqslant 1\}$.

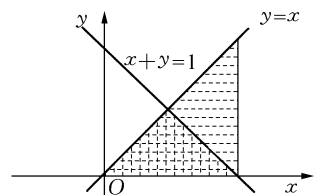


图 3-3

解 (1) 如图 3-3 所示，因为 D 的面积

$$S_D=\frac{1}{2},$$

则$(X,\ Y)$的概率密度函数为

$$f(x,\ y)=\begin{cases}2, & (x,\ y)\in D,\\ 0, & (x,\ y)\notin D.\end{cases}$$

(2) 记 $D_1=\{(x,\ y)\mid x+y\leqslant 1,\ 0\leqslant y\leqslant x\}$，则

$$P\{X+Y\leqslant 1\}=\iint\limits_{x+y\leqslant 1}f(x,\ y)\mathrm{d}x\mathrm{d}y=\iint\limits_{D_1}2\mathrm{d}x\mathrm{d}y=2\times\frac{1}{2}\times\frac{1}{2}\times 1=\frac{1}{2}.$$

2. 二维正态分布

定义 7 若二维随机变量$(X,\ Y)$的概率密度函数为

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

其中，μ_1，μ_2，σ_1^2，σ_2^2，ρ 都是常数，且 $-\infty<\mu_1$，$\mu_2<+\infty$，$\sigma_1>0$，$\sigma_2>0$，$|\rho|<1$，则称$(X,\ Y)$服从**二维正态分布**，记为$(X,\ Y)\sim N(\mu_1,\ \mu_2,\ \sigma_1^2,$

σ_2^2，ρ).

二维正态分布 $N(0, 0, 1, 1, 0)$的密度函数曲面图如图 3-4 所示.

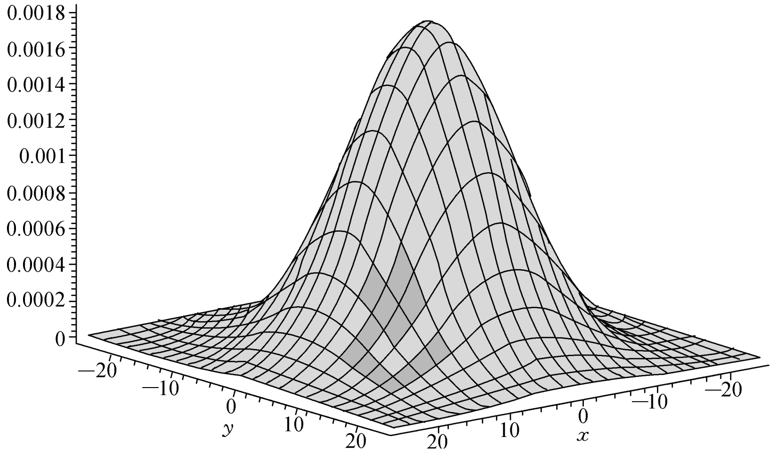


图 3-4

3.2　边缘分布与条件分布

设(X, Y)为二维随机变量，研究其中任何一个变量 X 或 Y，而不管另一个变量取什么值，这样得到的随机变量 X 或 Y 的分布叫作二维随机变量(X, Y)关于 X 或 Y 的**边缘分布**.

3.2.1　离散型随机变量的边缘分布律

定义 1　设(X, Y)为二维离散型随机变量，则其分量 X 或 Y 的分布律称为(X, Y)关于 X 或 Y 的**边缘分布律**.

设(X, Y)的分布律为

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

则关于 X 的边缘分布律是

$$p_{i\cdot} = P\{X = x_i\} = \sum_j p_{ij}, \quad i=1, 2, 3, \cdots; \qquad (3.2.1)$$

关于 Y 的边缘分布律是

$$p_{\cdot j} = P\{Y = y_j\} = \sum_i p_{ij}, \quad j=1, 2, 3, \cdots. \qquad (3.2.2)$$

例 1　已知 $P(A)=\dfrac{1}{4}$，$P(B|A)=\dfrac{1}{3}$，$P(A|B)=\dfrac{1}{2}$，令

$$X=\begin{cases}1, & A\text{ 发生},\\ 0, & A\text{ 不发生},\end{cases}\quad Y=\begin{cases}1, & B\text{ 发生},\\ 0, & B\text{ 不发生},\end{cases}$$

求 X 和 Y 的联合分布律，并写出 X 和 Y 的边缘分布律．

解　$P(AB)=P(A)P(B|A)=\frac{1}{12}$，$P(B)=\frac{P(AB)}{P(A|B)}=\frac{1}{6}$.

$$P\{X=1,\ Y=1\}=P(AB)=\frac{1}{12},$$

$$P\{X=1,\ Y=0\}=P(A\bar{B})=P(A)-P(AB)=\frac{1}{6},$$

$$P\{X=0,\ Y=1\}=P(\bar{A}B)=P(B)-P(AB)=\frac{1}{12},$$

$$\begin{aligned}P\{X=0,\ Y=0\}&=P(\bar{A}\bar{B})=1-P(A\cup B)\\&=1-P(A)-P(B)+P(AB)=\frac{2}{3}.\end{aligned}$$

即$(X,\ Y)$的联合分布律为

X \ Y	0	1	$p_{i\cdot}$
0	$\frac{2}{3}$	$\frac{1}{12}$	$\frac{3}{4}$
1	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{4}$
$p_{\cdot j}$	$\frac{5}{6}$	$\frac{1}{6}$	

在上表中，我们将边缘分布律写在联合分布律表格的边缘上，这就是“边缘分布律”这个名词的来源．

3.2.2　二维连续型随机变量的边缘概率密度

定义 2　设二维连续型随机变量$(X,\ Y)$的联合概率密度函数为 $f(x,\ y)$，则 X 的分布函数为

$$F_X(x)=P\{X\leqslant x,\ Y<+\infty\}=F(x,\ +\infty)=\int_{-\infty}^{x}\left[\int_{-\infty}^{+\infty}f(x,\ y)\mathrm{d}y\right]\mathrm{d}x,$$

所以 X 是连续型随机变量，且其概率密度为

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,\ y)\mathrm{d}y. \tag{3.2.3}$$

同理，Y 也是连续型随机变量，且其概率密度为

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,\ y)\mathrm{d}x. \tag{3.2.4}$$

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

例 2　设(X, Y)服从 $D=\{(x, y)\mid x\geqslant 0, y\geqslant 0, x+y\leqslant 1\}$上的均匀分布，求 X 和 Y 的边缘概率密度.

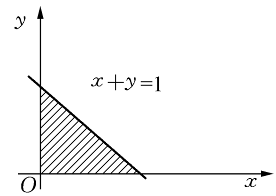


图 3-5

解　如图 3-5 所示，因为(X, Y)的联合概率密度函数为

$$f(x, y)=\begin{cases}2, & (x, y)\in D,\\ 0, & (x, y)\notin D.\end{cases}$$

关于 X 的边缘概率密度为

$$f_X(x)=\int_{-\infty}^{+\infty}f(x, y)\mathrm{d}y=\begin{cases}\int_0^{1-x}2\mathrm{d}y, & 0\leqslant x\leqslant 1,\\ 0, & \text{其他}\end{cases}$$

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

关于 Y 的边缘概率密度为

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x, y)\mathrm{d}x=\begin{cases}\int_0^{1-y}2\mathrm{d}x, & 0\leqslant y\leqslant 1,\\ 0, & \text{其他}\end{cases}$$

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

例 3　设(X, Y)服从二维正态分布 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$，求 X 和 Y 的边缘概率密度.

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

$$\frac{(y-\mu_2)^2}{\sigma_2^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}=\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right)^2-\rho^2\frac{(x-\mu_1)^2}{\sigma_1^2},$$

于是 X 的边缘概率密度为

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

设 $t=\dfrac{1}{\sqrt{1-\rho^2}}\left(\dfrac{y-\mu_2}{\sigma_2}-\rho\dfrac{x-\mu_1}{\sigma_1}\right)$，则利用 $\displaystyle\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t=1$，得

$$f_X(x)=\frac{1}{2\pi\sigma_1}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t=\frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}(-\infty<x<+\infty).$$

同理 Y 的边缘概率密度为

$$f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}\ (-\infty<y<+\infty).$$

从例 3 可看到二维正态分布的两个边缘分布都是正态分布，并且都与参数 ρ 无关，表明两个随机变量的联合概率密度可以唯一确定边缘概率密度，但两个随机变量的边缘概率密度却不一定能确定它们的联合概率密度．

3.2.3　条件分布

设(X, Y)是二维离散型随机变量，其分布律为

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

其关于 X 和 Y 的边缘分布律分别为

$$P\{X=x_i\}=p_{i\cdot}=\sum_j p_{ij}, \ i=1, 2, 3, \cdots,$$

$$P\{Y=y_j\}=p_{\cdot j}=\sum_i p_{ij}, \ j=1, 2, 3, \cdots.$$

当 $p_{\cdot j}>0$ 时，由条件概率公式，有

$$P\{X=x_i|Y=y_j\}=\frac{p_{ij}}{p_{\cdot j}}, \ i=1, 2, \cdots. \tag{3.2.5}$$

显然，上述条件概率具有分布律的性质：

(1) $P\{X=x_i|Y=y_j\}\geqslant 0$；(2) $\sum_i P\{X=x_i|Y=y_j\}=\sum_i \frac{p_{ij}}{p_{\cdot j}}=1$.

称(3.2.5)式为在 $Y=y_j$ 的条件下随机变量 X 的**条件分布律**．

同理，对于固定的 i，若 $p_{i\cdot}>0$，则称

$$P\{Y=y_j|X=x_i\}=\frac{p_{ij}}{p_{i\cdot}}, \ j=1, 2, \cdots \tag{3.2.6}$$

为在 $X=x_i$ 的条件下随机变量 Y 的**条件分布律**．

例 4　设二维离散型随机变量(X, Y)的概率分布律为

Y \ X	0	1	2
-1	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{3}{12}$
1	0	$\frac{1}{12}$	$\frac{1}{12}$
3	$\frac{3}{12}$	$\frac{1}{12}$	0

求：(1) 在 $X=0$ 的条件下 Y 的条件分布律；

(2) 在 $Y=1$ 的条件下 X 的条件分布律．

解　(1) 因为 $P\{X=0\}=\frac{1}{12}+0+\frac{3}{12}=\frac{1}{3}$，所以

$$P\{Y=-1 \mid X=0\}=\frac{P\{X=0,\ Y=-1\}}{P\{X=0\}}=\frac{1/12}{1/3}=\frac{1}{4}.$$

类似可得

$$P\{Y=1 \mid X=0\}=0,\ P\{Y=3 \mid X=0\}=\frac{3}{4}.$$

因此，在 $X=0$ 的条件下 Y 的条件分布律为

$Y \mid X=0$	-1	1	3
P	$\frac{1}{4}$	0	$\frac{3}{4}$

（2）同理，在 $Y=1$ 的条件下 X 的条件分布律为

$X \mid Y=1$	0	1	2
P	0	$\frac{1}{2}$	$\frac{1}{2}$

3.3　随机变量的相互独立性

本节我们将利用两个随机事件相互独立的概念引出两个随机变量相互独立的概念．

定义　设 $F(x,\ y)$，$F_X(x)$和 $F_Y(y)$分别是$(X,\ Y)$，X 和 Y 的分布函数，若对任意实数 x，y，均有

$$P\{X\leqslant x,\ Y\leqslant y\}=P\{X\leqslant x\}\cdot P\{Y\leqslant y\},$$

即

$$F(x,\ y)=F_X(x)\cdot F_Y(y), \tag{3.3.1}$$

则称随机变量 X 与 Y 是**相互独立的**．

当$(X,\ Y)$为离散型随机变量时，X 和 Y 相互独立的充要条件(3.3.1)式等价于：对于$(X,\ Y)$的所有可能取值$(x_i,\ y_j)$，都有

$$P\{X=x_i,\ Y=y_j\}=P\{X=x_i\}\cdot P\{Y=y_j\}. \tag{3.3.2}$$

当$(X,\ Y)$是连续型随机变量时，$f(x,\ y)$，$f_X(x)$和 $f_Y(y)$分别是$(X,\ Y)$，X 和 Y 的概率密度函数，则随机变量 X 和 Y 相互独立的充要条件(3.3.1)式等价于：对一切实数 x 和 y，都有

$$f(x,\ y)=f_X(x)\cdot f_Y(y). \tag{3.3.3}$$

注意：在实际应用中，使用(3.3.2)式或(3.3.3)式要比使用(3.3.1)式更方便．

例 1　设随机变量$(X,\ Y)$的概率分布律为

$Y \backslash X$	-1	0	2
0	$\frac{2}{20}$	$\frac{1}{20}$	$\frac{2}{20}$
1	$\frac{2}{20}$	$\frac{1}{20}$	$\frac{2}{20}$
2	$\frac{4}{20}$	$\frac{2}{20}$	$\frac{4}{20}$

试验证 X 和 Y 的独立性．

证明　由(X, Y)的概率分布可求得 X 和 Y 的概率分布律分别为

X	-1	0	2
P	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{2}{5}$

Y	0	1	2
P	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$

容易验证对一切 $x_i \in \{-1, 0, 2\}$，$i=1, 2, 3$，$y_j \in \{0, 1, 2\}$，$j=1, 2, 3$，均有

$$P\{X=x_i, Y=y_j\}=P\{X=x_i\} \cdot P\{Y=y_j\},$$

所以 X 和 Y 相互独立．

例 2　已知二维随机变量(X, Y)的联合概率密度函数为

$$f(x, y)=\begin{cases} 2\mathrm{e}^{-(x+2y)}, & x>0, y>0, \\ 0, & \text{其他}, \end{cases}$$

证明：随机变量 X 和 Y 相互独立．

证明　当 $x>0$ 时，$f_X(x)=\int_0^{+\infty} 2\mathrm{e}^{-(x+2y)}\mathrm{d}y=2\mathrm{e}^{-x}\int_0^{+\infty}\mathrm{e}^{-2y}\mathrm{d}y=\mathrm{e}^{-x}$；

当 $x \leqslant 0$ 时，$f_X(x)=0$，

则 X 的概率密度函数为

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

同理可得 Y 的概率密度函数为

$$f_Y(y)=\begin{cases} 2\mathrm{e}^{-2y}, & y>0, \\ 0, & y \leqslant 0. \end{cases}$$

因而
$$f(x, y)=f_X(x) \cdot f_Y(y),$$

所以 X 和 Y 相互独立．

例 3　设(X,Y)的联合概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{\pi R^2}, & x^2+y^2\leqslant R^2,\\ 0, & \text{其他},\end{cases}$$

判断 X 和 Y 是否独立．

解　因为 $f_X(x)=\begin{cases}\displaystyle\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}}\frac{1}{\pi R^2}\mathrm{d}y, & |x|\leqslant R,\\ 0, & |x|>R\end{cases}$

$$=\begin{cases}\dfrac{2\sqrt{R^2-x^2}}{\pi R^2}, & |x|\leqslant R,\\ 0, & |x|>R,\end{cases}$$

$$f_Y(y)=\begin{cases}\displaystyle\int_{-\sqrt{R^2-y^2}}^{\sqrt{R^2-y^2}}\frac{1}{\pi R^2}\mathrm{d}x, & |y|\leqslant R,\\ 0, & |y|>R\end{cases}=\begin{cases}\dfrac{2\sqrt{R^2-y^2}}{\pi R^2}, & |y|\leqslant R,\\ 0, & |y|>R,\end{cases}$$

故 $f(x,y)\neq f_X(x)\cdot f_Y(y)$，则 X 和 Y 不相互独立．

例 4　设随机变量 X 和 Y 的概率分布

$$X\sim\begin{pmatrix}-1 & 0 & 1\\ \dfrac{1}{4} & \dfrac{1}{2} & \dfrac{1}{4}\end{pmatrix},\ Y\sim\begin{pmatrix}0 & 1\\ \dfrac{1}{2} & \dfrac{1}{2}\end{pmatrix},$$

而且 $P\{XY=0\}=1$.

(1) 求 X 和 Y 的联合分布；

(2) 问 X 和 Y 是否独立：为什么？

解　(1) 由于$\{X=-1,Y=1\}\subset\{XY\neq 0\}$，故

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

所以 $$P\{X=-1,Y=1\}=0,$$

因而， $$P\{X=-1,Y=0\}=P\{X=-1\}-P\{X=-1,Y=1\}=\frac{1}{4}.$$

同理有 $P\{X=1,Y=1\}=0$，

$$P\{X=1,Y=0\}=P\{X=1\}-P\{X=1,Y=1\}=\frac{1}{4},$$

$$P\{X=0,Y=0\}=P\{Y=0\}-P\{X=-1,Y=0\}-P\{X=1,Y=0\}=0,$$

$$P\{X=0,Y=1\}=P\{X=0\}-P\{X=0,Y=0\}=\frac{1}{2},$$

所以(X,Y)的联合分布为

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

(2) 因为 $P\{X=-1, Y=0\}=\frac{1}{4}\neq\frac{1}{4}\times\frac{1}{2}=P\{X=-1\}P\{Y=0\}$，所以 X 和 Y 不相互独立．

由二维正态分布的定义及上节的例 3 易证明，若$(X, Y)\sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$，则 X 和 Y 相互独立的充要条件是参数 $\rho=0$.

以上关于二维随机变量的一些概念，容易推广到 n 维随机变量的情形．若对 n 维随机变量$(X_1, X_2, \cdots, X_n)$的分布函数 $F(x_1, x_2, \cdots, x_n)=P\{X_1\leqslant x_1, X_2\leqslant x_2, \cdots, X_n\leqslant x_n\}$，存在非负函数 $f(x_1, x_2, \cdots, x_n)$，使得对任意实数 $x_1, x_2, \cdots, x_n$ 有

$$F(x_1, x_2, \cdots, x_n)=\int_{-\infty}^{x_n}\int_{-\infty}^{x_{n-1}}\cdots\int_{-\infty}^{x_1}f(x_1, x_2, \cdots, x_n)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n,$$

则称 $f(x_1, x_2, \cdots, x_n)$为$(X_1, X_2, \cdots, X_n)$的概率密度函数．

设$(X_1, X_2, \cdots, X_n)$的分布函数 $F(x_1, x_2, \cdots, x_n)$为已知，则$(X_1, X_2, \cdots, X_n)$的 $k(1\leqslant k<n)$维边缘分布函数就随之确定．例如，$(X_1, X_2, \cdots, X_n)$关于 X_1、关于(X_1, X_2)的边缘分布函数分别为

$$F_{X_1}(x_1)=F(x_1, +\infty, \cdots, +\infty)$$

和

$$F_{X_1,X_2}(x_1, x_2)=F(x_1, x_2, +\infty, \cdots, +\infty).$$

又若 $f(x_1, x_2, \cdots, x_n)$是$(X_1, X_2, \cdots, X_n)$的概率密度，则$(X_1, X_2, \cdots, X_n)$关于 X_1、关于(X_1, X_2)的边缘概率密度分别为

$$f_{X_1}(x_1)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(x_1, x_2, \cdots, x_n)\mathrm{d}x_2\mathrm{d}x_3\cdots\mathrm{d}x_n,$$

$$f_{X_1,X_2}(x_1, x_2)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(x_1, x_2, \cdots, x_n)\mathrm{d}x_3\mathrm{d}x_4\cdots\mathrm{d}x_n.$$

若对所有的实数 $x_1, x_2, \cdots, x_n$ 有

$$F(x_1, x_2, \cdots, x_n)=F_{X_1}(x_1)F_{X_2}(x_2)\cdots F_{X_n}(x_n),$$

则称 $X_1, X_2, \cdots, X_n$ 是相互独立的．

3.4　两个随机变量函数的分布

类似于一维随机变量的函数的分布，我们将由二维随机变量$(X，Y)$的联合分布，讨论如何求得随机变量 X 与 Y 的函数 $Z=g(X，Y)$的分布．本节只讨论几个简单的二维随机变量的函数的分布．

3.4.1　离散型随机变量函数的分布

对于两个离散型随机变量的函数的分布，我们仅通过例子来分析，从中找到解决问题的方法．

例 1　设二维离散型随机变量$(X，Y)$的分布律为

$Y\backslash X$	0	1	2
0	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{8}$
1	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{12}$

求：(1) $Z=X+Y$；(2) $Z=XY$；(3) $Z=\max\{X，Y\}$的分布律．

解　(1) $Z=X+Y$ 可能取值为 0，1，2，3，其分布律为

$Z=X+Y$	0	1	2	3
P	$\frac{1}{4}$	$\frac{5}{12}$	$\frac{1}{4}$	$\frac{1}{12}$

其中
$$P\{Z=0\}=P\{X=0，Y=0\}=\frac{1}{4},$$
$$P\{Z=1\}=P\{X=0，Y=1\}+P\{X=1，Y=0\}=\frac{1}{4}+\frac{1}{6}=\frac{5}{12},$$
其他类似可得．

(2) $Z=XY$ 可能取值为 0，1，2，其分布律为

$Z=XY$	0	1	2
P	$\frac{19}{24}$	$\frac{1}{8}$	$\frac{1}{12}$

(3) $Z=\max\{X，Y\}$的可能取值为 0，1，2，其分布律为

$Z=\max\{X，Y\}$	0	1	2
P	$\frac{1}{4}$	$\frac{13}{24}$	$\frac{5}{24}$

例 2　设随机变量 X 和 Y 相互独立，X 和 Y 的分布律分别为

X	0	1	2
P	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{2}$

Y	0	1	2
P	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

求 $Z=X+Y$ 的分布律．

解　$Z=X+Y$ 的可能取值为 0，1，2，3，4，且

$$P\{Z=0\}=P\{X=0,\ Y=0\}=P\{X=0\}P\{Y=0\}=\frac{1}{8}\times\frac{1}{3}=\frac{1}{24},$$

$$\begin{aligned}P\{Z=1\}&=P\{X=0,\ Y=1\}+P\{X=1,\ Y=0\}\\&=P\{X=0\}P\{Y=1\}+P\{X=1\}P\{Y=0\}\\&=\frac{1}{8}\times\frac{1}{2}+\frac{3}{8}\times\frac{1}{3}=\frac{3}{16},\end{aligned}$$

$$\begin{aligned}P\{Z=2\}&=P\{X=0,\ Y=2\}+P\{X=2,\ Y=0\}+P\{X=1,\ Y=1\}\\&=P\{X=0\}P\{Y=2\}+P\{X=2\}P\{Y=0\}+P\{X=1\}P\{Y=1\}\\&=\frac{1}{8}\times\frac{1}{6}+\frac{1}{2}\times\frac{1}{3}+\frac{3}{8}\times\frac{1}{2}=\frac{3}{8},\end{aligned}$$

$$\begin{aligned}P\{Z=3\}&=P\{X=1,\ Y=2\}+P\{X=2,\ Y=1\}\\&=P\{X=1\}P\{Y=2\}+P\{X=2\}P\{Y=1\}\\&=\frac{3}{8}\times\frac{1}{6}+\frac{1}{2}\times\frac{1}{2}=\frac{5}{16},\end{aligned}$$

$$P\{Z=4\}=P\{X=2,\ Y=2\}=P\{X=2\}P\{Y=2\}=\frac{1}{2}\times\frac{1}{6}=\frac{1}{12},$$

所以 $Z=X+Y$ 分布律为

$Z=X+Y$	0	1	2	3	4
P	$\frac{1}{24}$	$\frac{3}{16}$	$\frac{3}{8}$	$\frac{5}{16}$	$\frac{1}{12}$

3.4.2　连续型随机变量函数的分布

关于两个连续型随机变量的函数的分布，我们仅介绍两种函数情况．

1. 和的分布

设二维随机变量$(X,\ Y)$的联合概率密度函数为 $f(x,\ y)$，先考虑 $Z=$

$X+Y$的分布函数：

$$F_Z(z)=P\{Z\leqslant z\}=P\{X+Y\leqslant z\}=\iint\limits_{x+y\leqslant z}f(x,\ y)\mathrm{d}x\mathrm{d}y,$$

其中二重积分区域是位于直线左下方的半平面(图 3-6)，因此

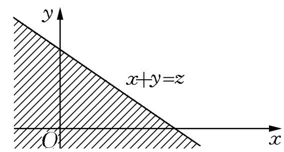


图 3-6

$$F_Z(z)=\int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{z-x}f(x,\ y)\mathrm{d}y,$$

固定 z 和 x，令 $y=u-x$，则

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

因此

$$F_Z(z)=\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{z}f(x,\ u-x)\mathrm{d}u\right]\mathrm{d}x,$$

交换积分次序，得

$$F_Z(z)=\int_{-\infty}^{z}\left[\int_{-\infty}^{+\infty}f(x,\ u-x)\mathrm{d}x\right]\mathrm{d}u,$$

上式对 z 求导，即得 Z 的概率密度函数

$$f_Z(z)=\int_{-\infty}^{+\infty}f(x,\ z-x)\mathrm{d}x, \tag{3.4.1}$$

或

$$f_Z(z)=\int_{-\infty}^{+\infty}f(z-y,\ y)\mathrm{d}y. \tag{3.4.2}$$

如果 X 与 Y 相互独立，设 $f_X(x)$，$f_Y(y)$分别是 X 和 Y 的边缘概率密度，则

$$f_Z(z)=\int_{-\infty}^{+\infty}f_X(x)f_Y(z-x)\mathrm{d}x, \tag{3.4.3}$$

或

$$f_Z(z)=\int_{-\infty}^{+\infty}f_X(z-y)f_Y(y)\mathrm{d}y. \tag{3.4.4}$$

称(3.4.3)式与(3.4.4)式为 $f_X(x)$和 $f_Y(y)$的**卷积公式**.

例 3　设 X 与 Y 是两个相互独立的随机变量，它们都服从标准正态分布 $N(0,\ 1)$，求 $Z=X+Y$ 的概率密度.

解　因为 $X\sim N(0,\ 1)$，$Y\sim N(0,\ 1)$，故

$$f_X(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}},\ x\in(-\infty,\ +\infty),$$

$$f_Y(y)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2}{2}},\ y\in(-\infty,\ +\infty).$$

由(3.4.3)式得 $Z=X+Y$ 的概率密度为

$$f_Z(z)=\int_{-\infty}^{+\infty}f_X(x)f_Y(z-x)\mathrm{d}x=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{e}^{-\frac{(z-x)^2}{2}}\mathrm{d}x$$

$$=\frac{1}{2\pi}e^{-\frac{z^2}{4}}\int_{-\infty}^{+\infty}e^{-\left(x-\frac{z}{2}\right)^2}dx,$$

令 $t=x-\dfrac{z}{2}$，利用 $\int_{-\infty}^{+\infty}e^{-t^2}dt=\sqrt{\pi}$，得

$$f_Z(z)=\frac{1}{\sqrt{2\pi}\sqrt{2}}e^{-\frac{z^2}{2(\sqrt{2})^2}},$$

即 $Z=X+Y$ 服从 $N(0,2)$的正态分布．

一般地，若 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$Z=X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$，还可推广到 n 个独立随机变量的情况．

更一般地，设 $X_1,\cdots,X_n$ 相互独立，$X_i\sim N(\mu_i,\sigma_i^2)$，$i=1,2,\cdots,n$，则 $C_1X_1+C_2X_2+\cdots+C_nX_n\sim N(C_1\mu_1+C_2\mu_2+\cdots+C_n\mu_n,C_1^2\sigma_1^2+C_2^2\sigma_2^2+\cdots+C_n^2\sigma_n^2)$，即有限个相互独立的服从正态分布的随机变量的线性组合仍然服从正态分布．

例 4　已知 X，Y 独立同分布且 $X\sim U(0,1)$，求 $Z=X+Y$ 的概率密度．

解　因为 $X\sim U(0,1)$，$Y\sim U(0,1)$，故

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

由(3.4.3)式得 $Z=X+Y$ 的概率密度为

$$f_Z(z)=\int_{-\infty}^{+\infty}f_X(x)f_Y(z-x)dx,$$

易知，当且仅当$\begin{cases}0<x<1,\\ 0<z-x<1,\end{cases}$也即$\begin{cases}0<x<1,\\ z-1<x<z\end{cases}$时，上式不为零．

参照图 3－7 易得

$$f_Z(z)=\begin{cases}0, & z\leqslant 0 \text{ 或 } z>2,\\ \int_0^z dx, & 0<z\leqslant 1,\\ \int_{z-1}^1 dx, & 1<z\leqslant 2,\end{cases}$$

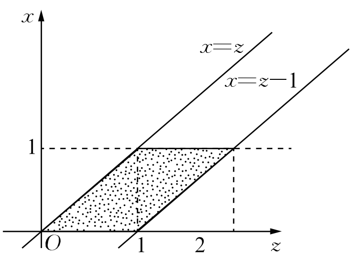


图 3－7

得 $Z=X+Y$ 的概率密度

$$f_Z(z)=\begin{cases}0, & z\leqslant 0 \text{ 或 } z>2,\\ z, & 0<z\leqslant 1,\\ 2-z, & 1<z\leqslant 2.\end{cases}$$

2. $U=\max\{X,Y\}$**与** $V=\min\{X,Y\}$**的分布**

设随机变量 X 与 Y 相互独立，它们的分布函数分别为 $F_X(x)$及 $F_Y(y)$，我们来求 $U=\max\{X,Y\}$与 $V=\min\{X,Y\}$的分布函数．

由于事件“$U=\max\{X, Y\}\leqslant z$”等价于事件“$X$ 与 Y 都不大于 z”，注意到 X 与 Y 的独立性，我们有 $U=\max\{X, Y\}$的分布函数为

$$\begin{aligned} F_{\max}(z) &= P\{\max\{X, Y\}\leqslant z\}=P\{X\leqslant z, Y\leqslant z\} \\ &= P\{X\leqslant z\}P\{Y\leqslant z\}=F_X(z)F_Y(z). \end{aligned}$$

类似可得 $V=\min\{X, Y\}$的分布函数为

$$\begin{aligned} F_{\min}(z) &= P\{\min\{X, Y\}\leqslant z\}=1-P\{\min\{X, Y\}>z\} \\ &= 1-P\{X>z, Y>z\}=1-P\{X>z\}P\{Y>z\} \\ &= 1-[1-P\{X\leqslant z\}][1-P\{Y\leqslant z\}] \\ &= 1-[1-F_X(z)][1-F_Y(z)]. \end{aligned}$$

上述结论不难推广到 n 个相互独立的随机变量的情形．设 X_1，X_2，…，X_n 是 n 个相互独立的随机变量，它们的分布函数分别是 $F_{X_i}(x)(i=1, 2, \cdots, n)$，则 $\max\{X_1, X_2, \cdots, X_n\}$及 $\min\{X_1, X_2, \cdots, X_n\}$的分布函数分别为

$$F_{\max}(z)=F_{X_1}(z)F_{X_2}(z)\cdots F_{X_n}(z), \tag{3.4.5}$$

$$F_{\min}(z)=1-[1-F_{X_1}(z)][1-F_{X_2}(z)]\cdots[1-F_{X_n}(z)]. \tag{3.4.6}$$

特别地，当 X_1，X_2，…，X_n 相互独立且具有相同的分布函数 $F(x)$ 时，有

$$F_{\max}(z)=[F(z)]^n, \tag{3.4.7}$$

$$F_{\min}(z)=1-[1-F(z)]^n. \tag{3.4.8}$$

例 5　某系统由三个独立子部件串联而成，每个部件的寿命都服从相同指数分布，其概率密度函数为

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

求该系统的寿命的概率密度函数．

解　设 X_1，X_2，X_3 分别为三个子部件的寿命，则它们的分布函数都为

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

设该系统的寿命为 Z，则 $Z=\min\{X_1, X_2, X_3\}$，由公式(3.4.8)得 Z 的分布函数为

$$F_Z(z)=\begin{cases}1-[1-F(z)]^3, & z>0, \\ 0, & z\leqslant 0\end{cases}=\begin{cases}1-e^{-3\lambda z}, & z>0, \\ 0, & z\leqslant 0.\end{cases}$$

由此得 Z 的概率密度函数为

$$f(z)=\begin{cases}3\lambda e^{-3\lambda z}, & z>0, \\ 0, & z\leqslant 0.\end{cases}$$

3.4.3 离散型随机变量与连续型随机变量的函数的分布

下面通过一个简单的例子来说明如何来求两个相互独立的随机变量，当其中一个为离散型，而另一个为连续型时，它们的简单函数的分布的求法．

例 6 设随机变量 X 和 Y 相互独立，且随机变量 X 的分布律为 $P\{X=i\}=\frac{1}{3}(i=-1, 0, 1)$，而随机变量 Y 服从标准正态分布 $N(0, 1)$，试求 $Z=X+Y$ 的概率分布．

解 显然 $Z=X+Y$ 是连续型随机变量，下面我们根据分布函数法来求其概率分布．设 $Z=X+Y$ 的分布函数和概率密度函数分别为 $F_Z(z)$和 $f_Z(z)$，则对任意实数 z，

$$
\begin{aligned}
F_Z(z)&=P\{Z\leqslant z\}=P\{X+Y\leqslant z\}\\
&=P\{X+Y\leqslant z|X=-1\}P\{X=-1\}+P\{X+Y\leqslant z|X=0\}P\{X=0\}+\\
&\quad P\{X+Y\leqslant z|X=1\}P\{X=1\}\\
&=\frac{1}{3}\times[P\{Y\leqslant z+1\}+P\{Y\leqslant z\}+P\{Y\leqslant z-1\}]\\
&=\frac{1}{3}\times[\Phi(z+1)+\Phi(z)+\Phi(z-1)],
\end{aligned}
$$

其中 $\Phi(x)$为标准正态分布的分布函数，由此可得 $Z=X+Y$ 的密度函数为

$$f_Z(z)=F_Z'(z)=\frac{1}{3\sqrt{2\pi}}[\mathrm{e}^{-\frac{(z+1)^2}{2}}+\mathrm{e}^{-\frac{z^2}{2}}+\mathrm{e}^{-\frac{(z-1)^2}{2}}].$$

习　题　A

1. 设盒中有 2 个红球，3 个白球，从中每次任取一个球，连取两次．(1) 有放回；(2) 不放回．记 X 与 Y 分别表示第一次与第二次取出的红球数，试写出两种情况的二维随机变量(X, Y)的分布律及 X 与 Y 的边缘分布律．并说明 X 与 Y 是否独立．

2. 设二维随机变量(X, Y)的概率密度为

$$f(x, y)=\begin{cases}k(6-x-y), & 0<x<2,\ 2<y<4,\\ 0, & \text{其他},\end{cases}$$

求：(1) 常数 k；(2) $P\{X<1, Y<3\}$；(3) $P\{X<1.5\}$；(4) $P\{X+Y\leqslant 4\}$.

3. 设二维随机变量(X, Y)的分布函数为

$$F(x, y)=a\left(b+\arctan\frac{x}{2}\right)\left(c+\arctan\frac{y}{3}\right),$$

求：(1) 系数 a，b，c；(2) (X, Y)的概率密度；(3) 关于 X 与 Y 的边缘密度．

4. 一袋中装有 1 个红球，2 个白球，3 个黑球．从中任取 4 个球，以 X 与 Y 分别表示 4 个球中红球及白球的数量，求：

(1) (X, Y)的分布律；

(2) $X=0$ 的条件下 Y 的分布律；

(3) $X=1$ 的条件下 Y 的分布律．

5. 设二维随机变量(X, Y)的分布律为

$X \backslash Y$	1	2	3
1	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{18}$
2	$\frac{1}{3}$	a	b

问 a，b 取何值时，X 与 Y 相互独立？

6. 设随机变量 X 与 Y 的联合分布律为

$X \backslash Y$	0	1
0	$\frac{2}{25}$	b
1	a	$\frac{3}{25}$
2	$\frac{1}{25}$	$\frac{2}{25}$

且 $P\{Y=1 \mid X=0\}=\dfrac{3}{5}$，(1) 求常数 a，b 的值；(2) 当 a，b 取(1)中的值时，X 与 Y 是否独立？为什么？

7. 已知二维随机变量(X, Y)的联合概率密度函数为

$$f(x, y)=\begin{cases} \dfrac{1}{2}\mathrm{e}^{-\left(x+\frac{1}{2}y\right)}, & x>0, y>0, \\ 0, & \text{其他}, \end{cases}$$

(1) 求(X, Y)的分布函数；

(2) 求关于 X 和 Y 的边缘概率密度，说明 X 和 Y 是否相互独立．

8. 设随机变量 X 和 Y 相互独立，且都服从均匀分布 $U(0, 2)$，求随机变量 X 和 Y 之差的绝对值不超过 1 的概率．

9. 设二维随机变量(X, Y)的联合概率密度函数为

$$f(x,\ y)=\begin{cases}A, & x^2\leqslant y\leqslant x,\\ 0, & \text{其他},\end{cases}$$

(1) 求常数 A；

(2) 求关于 X 和 Y 的边缘概率密度，问 X 和 Y 是否相互独立？

(3) 求概率 $P\left\{0<Y<\dfrac{1}{2}\right\}$.

10. 设随机变量 X 和 Y 相互独立，且 X 与 Y 的分布律分别为

X	1	2
P	$\frac{1}{3}$	$\frac{2}{3}$

Y	0	1	2
P	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

求下列关于随机变量 X 和 Y 的函数的分布律：

(1) $Z=X+Y$；(2) $Z=XY$；(3) $Z=\max\{X,\ Y\}$；(4) $Z=\min\{X,\ Y\}$.

11. 设随机变量 X 和 Y 相互独立，且 $X\sim N(0,\ 2)$，$Y\sim N(0,\ 3)$，求 $Z=X+Y$ 的概率密度函数.

12. 设 X 与 Y 是两个相互独立的随机变量，它们的概率密度函数分别为

$$f_X(x)=\begin{cases}1, & 0\leqslant x\leqslant 1,\\ 0, & \text{其他},\end{cases}\quad f_Y(y)=\begin{cases}e^{-y}, & y>0,\\ 0, & \text{其他},\end{cases}$$

求随机变量 $Z=X+Y$ 的概率密度.

13. 电子仪器由 6 个相互独立的部件组成，联结方式如图 3-8 所示，设各个部件的使用寿命 X 服从相同的指数分布，其概率密度为

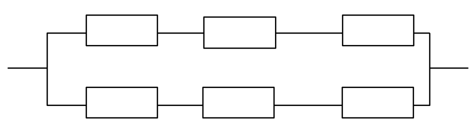


图 3-8

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

求仪器使用寿命的概率密度.

习　题　B

1. 随机变量 X 与 Y 都仅取 1，-1 两个值，且已知

$$P\{X=1\}=\frac{1}{2},\ P\{Y=1\mid X=1\}=P\{Y=-1\mid X=-1\}=\frac{1}{3},$$

(1) 求 $(X,\ Y)$ 的联合分布律；

(2) 求 t 的方程 $t^2+(X+Y)t+(X+Y)=0$ 至少有一个实根的概率.

2. 设某单位班车起点站上乘客人数 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 $p(0<p<1)$. 且各位乘客中途下车与否相互独

立，以 Y 表示在中途下车的乘客数，求：

(1) 在发车时有 n 个乘客的条件下，中途有 m 个人下车的概率；

(2) 二维随机变量 (X, Y) 的分布律；

(3) 关于 Y 的边缘分布律.

3. 设二维随机变量 (X, Y) 的概率密度为

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

求随机变量 $Z=X+Y$ 的概率密度.

4. 设二维随机变量 (X, Y) 的概率密度为

$$f(x, y)=\begin{cases} 2e^{-2x-y}, & x>0,\ y>0, \\ 0, & \text{其他}, \end{cases}$$

求 $Z=\max\{X, Y\}$ 的概率密度.

5. 设随机变量 X 与 Y 相互独立，并且都服从标准正态分布 $N(0, 1)$，求：(1) $Z=X^2+Y^2$ 的概率密度；(2) $P\{Z\geqslant 4\}$.

6. 设二维随机变量 (X, Y) 服从 $D=\{(x, y)\,|\,1\leqslant x\leqslant 3, 1\leqslant x\leqslant 3\}$ 上的均匀分布，求 $Z=|X-Y|$ 的概率密度.

7. 设随机变量 X 与 Y 相互独立，并且 X 为离散型随机变量，它的分布律为 $P\{X=i\}=\dfrac{1}{4}(i=0, 1, 2, 3)$，而随机变量 Y 为连续型随机变量，其概率密度函数为 $f(y)$，求随机变量 $Z=X+Y$ 的概率密度 $f_Z(z)$.

第 4 章　随机变量的数字特征

随机变量的概率分布能够完整地描述随机现象的统计特性，然而，在实际问题中，求随机变量的概率分布往往不是一件容易的事情．就某些实际问题而言，通常也没有必要对随机变量进行详细全面的考察，因而，不要求出它的概率分布．只要知道随机变量的某些综合性的指标，以便更集中更突出地描述随机变量，这些量大多是某种平均值．例如，检查一批棉花质量时，我们既关心棉花纤维的平均长度，又要注意纤维长度与平均长度的偏离程度．平均纤维长度越长，而偏离程度较小，质量就“好”，这里所讲的“平均长度”与“偏离程度”刻画了随机变量在某些方面的重要特征，这些数字特征在理论上和实践上都具有重要的意义，常用的随机变量的数字特征有：数学期望(均值)、方差和矩等．

4.1　随机变量的数学期望

4.1.1　引例

“期望”在我们日常生活中常指心理上的期望，而在概率论中，数学期望来源于历史上一个著名的分赌本问题．

例 1(**分赌本问题**)　17 世纪中叶，一位赌徒向法国数学家帕斯卡(Pascal，1623—1662 年)提出一个使他苦恼长久的分赌本问题：甲、乙两赌徒赌技不相上下，各出赌资 50 法郎，每局中无平局，他们约定，谁先赢三局，则得全部赌本 100 法郎，当甲赢了二局、乙赢了一局时，因故(皇帝召见)要中止赌博．现问这 100 法郎如何分才算公平？

这个问题引起了不少人的兴趣，首先大家都认识到：平均分对甲不公平，全部归甲对乙不公平．合理的分法是：按一定的比例，甲多分些，乙少分些．所以问题的焦点在于：按怎样的比例来分？以下有两种分法：

(1) 甲得 100 法郎中的$\frac{2}{3}$，乙得 100 法郎中的$\frac{1}{3}$．这是基于已赌局数：甲赢了二局、乙赢了一局．

(2) 1654 年帕斯卡提出如下的分法：设想再赌下去，则甲最终所得 X 为

一个随机变量，其可能取值为 0 或 100. 再赌两局必可结束，其结果不外乎以下四种情况之一：

甲甲、甲乙、乙甲、乙乙，

其中“甲乙”表示第一局甲胜第二局乙胜．在这四种情况中有三种可使甲获 100 法郎，只有一种情况(乙乙)下甲获 0 法郎．因为赌技不相上下，所以甲获得 100 法郎的可能性为$\frac{3}{4}$，获 0 法郎的可能性为$\frac{1}{4}$，即 X 的分布律为

X	0	100
P	0.25	0.75

经上述分析，帕斯卡认为，甲的“期望”所得应为：$0\times 0.25+100\times 0.75=75$(法郎). 即甲得 75 法郎，乙得 25 法郎．这种分法不仅考虑了已赌局数，而且还包括了对再赌下去的一种“期望”，它比(1)的分法更合理．

这就是数学期望这个名称的由来，其实这个名称称为“均值”更形象易懂. 对上例而言，也就是再赌下去的话，甲“平均”可以赢得 75 法郎．

下面我们来逐步分析如何由分布来求“均值”．

(1) 算术平均：如果有 n 个数 x_1，x_2，…，x_n，那么求这 n 个数的算术平均当然非常简单，只要将这 n 个数相加后除 n 就可求得．

(2) 加权平均：如果这 n 个数中有相同的，不妨设其中有 n_i 个取值为 x_i，$i=1$，2，…，k. 将其列表为

表 4－1

取值	x_1	x_2	…	x_k
频数	n_1	n_2	…	n_k
频率	$\frac{n_1}{n}$	$\frac{n_2}{n}$	…	$\frac{n_k}{n}$

则其“均值”应为

$$\frac{1}{n}\sum_{i=1}^{k} n_i x_i = \sum_{i=1}^{k}\frac{n_i}{n}x_i,$$

其中“加权”平均的权数$\frac{n_i}{n}$就是出现数值 x_i 的频率，而频率在 n 很大时，就稳定在其概率附近．

(3) 对于一个离散型随机变量 X，若它的可能取值是 x_1，x_2，…，x_n. 若将这 n 个数相加后除 n 作为“均值”，则肯定是不妥的．原因在于 X 取各个值的概率是不同的，概率大的出现的机会就大，则在计算中其权也应该大．而上例分赌

本问题启示我们：用取值的概率作为一种“权数”作加权平均是十分合理的．

经以上分析，我们就可以给出离散型随机变量数学期望的定义．

4.1.2　离散型随机变量的数学期望

定义 1　设离散型随机变量 X 的分布律为

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

若级数 $\sum_{i=1}^{\infty} x_i p_i$ 绝对收敛，则称 $\sum_{i=1}^{\infty} x_i p_i$ 为 X 的**数学期望**，记为 $E(X)$，即

$$E(X)=\sum_{i=1}^{\infty} x_i p_i. \tag{4.1.1}$$

若级数 $\sum_{i=1}^{\infty} x_i p_i$ 不绝对收敛，则称随机变量 X 的数学期望不存在．

例 2　在澳门赌场，有很多人在赌廿一点时顺便押对子．其规则如下：庄家从 6 副(每副 52 张)扑克中随机发给每位参赌人两张．如果参赌人下注 a 元，当得到两张牌是一对时，庄家赔给该参赌人 10 倍，否则输掉该参赌人的赌注. 如果某参赌人每局下注 100 元，求该参赌人在每局中和庄家在每局中各赢钱的期望．

解　用 X，Y 分别表示该参赌人与庄家在一局中的获利，$a=100$，则

$$P\{X=10a\}=\frac{13C_{4\times 6}^{2}}{C_{52\times 6}^{2}}=0.074,\quad P\{X=-a\}=1-0.074,$$

于是，该参赌人每局期望赢

$$E(X)=10a\times 0.074-a\times(1-0.074)=-0.186a=-18.6(\text{元});$$

庄家每局赢 18.6 元，这是因为

$$E(Y)=-10a\times 0.074+a\times(1-0.074)=-E(X).$$

例 3　有甲、乙两个射手，用随机变量 X 和 Y 分别表示他们击中目标的环数，设 X 和 Y 的分布律分别为

X	8	9	10
P	0.1	0.8	0.1

Y	8	9	10
P	0.3	0.2	0.5

试比较两射手射击技术的优劣．

解　从 X 和 Y 的分布律来看，我们无法准确地判断谁的射击技术好，于是我们可以用平均环数来比较两射手的射击技术．

$$E(X)=8\times 0.1+9\times 0.8+10\times 0.1=9(\text{环}),$$

$$E(Y)=8\times0.3+9\times0.2+10\times0.5=9.2(\text{环}),$$

即甲、乙平均每次射击命中的环数分别是 9 环和 9.2 环，$E(Y)>E(X)$，从而可以判断乙的射击技术好．

例 4　在一个人数很多的团体中普查某种疾病，为此要抽验 N 个人的血，可以用两种方法进行．①将每个人的血分别去验，这就需要验 N 次．②按 k 个人一组进行分组，把从 k 个人抽来的血混合在一起进行检验，如果这混合血液呈阴性反应，就说明 k 个人的血液都呈阴性反应，这样 k 个人的血就只需验一次．若呈阳性，则再对这 k 个人的血液分别进行化验．这样 k 个人的血总共要化验$(k+1)$次．假设该疾病的发病率为 p，且人群中各人是否得此病相互独立．试说明当 p 较小时，选取恰当的 k，按第二种方法可减少化验的次数．并说明 k 取什么值时最合适．

解　令 X 为该人群中每个人需要的验血次数，则 X 的分布律为

X	$\frac{1}{k}$	$1+\frac{1}{k}$
P	$(1-p)^k$	$1-(1-p)^k$

所以每人平均验血次数为

$$E(X)=\frac{1}{k}(1-p)^k+\left(1+\frac{1}{k}\right)[1-(1-p)^k]=1-(1-p)^k+\frac{1}{k}.$$

N 个人平均需化验的次数为

$$N\left[1-(1-p)^k+\frac{1}{k}\right],$$

由此可知，只要选取 k 使

$$1-(1-p)^k+\frac{1}{k}<1,$$

则 N 个人平均需化验的次数$<N$，当 p 固定时，我们选取 k 使得

$$L=1-(1-p)^k+\frac{1}{k}$$

小于 1 且取得最小值，这时就能得到最好的分组方法．

例如，$p=0.06$，则当 $k=5$ 时，$L=1-(1-p)^k+\frac{1}{k}$取得最小值 0.466．此时得到最好的分组方法．若 $N=1000$，此时以 $k=5$ 分组，则按第二种方法平均只需化验

$$1000\left[1-(1-0.06)^5+\frac{1}{5}\right]=466(\text{次}),$$

这样平均来说，可以减少 53%的工作量．

我们要注意到，并非所有的随机变量都存在数学期望，下例中的随机变量的数学期望就不存在．

例 5　设随机变量 X 的分布律为

$$P\{X=2^k\}=\frac{1}{2^k},\ k=1,\ 2,\ \cdots,$$

显然

$$\sum_{k=1}^{\infty} p_k = \sum_{k=1}^{\infty} \frac{1}{2^k} = 1,$$

因此它是分布律．但是，

$$\sum_{k=1}^{\infty} |x_k| p_k = \sum_{k=1}^{\infty}\left(2^k \times \frac{1}{2^k}\right) = \infty,$$

因此随机变量 X 的数学期望不存在．

下面我们来计算几种常见离散型随机变量的数学期望．

1. 0—1 分布

例 6　设 $X \sim B(1,\ p)$，求 $E(X)$.

解　X 的分布律为

X	0	1
P	$1-p$	p

则

$$E(X)=0\times(1-p)+1\times p=p.$$

2. 二项分布

例 7　设 $X \sim B(n,\ p)$，求 $E(X)$.

解　X 的分布律为

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

于是

$$\begin{aligned} E(X) &= \sum_{k=0}^{n} kC_n^k p^k(1-p)^{n-k} = \sum_{k=1}^{n} kC_n^k p^k(1-p)^{n-k} \\ &= np\sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1}(1-p)^{(n-1)-(k-1)} \\ &= np(p+1-p)^{n-1} = np. \end{aligned}$$

3. 泊松分布

例 8　设 $X \sim P(\lambda)$，求 $E(X)$.

解　X 的分布律为

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

于是

$$E(X) = \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=1}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \lambda\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda}$$

$$=\lambda\sum_{l=0}^{\infty}\frac{\lambda^{l}}{l!}\mathrm{e}^{-\lambda}=\lambda\times 1=\lambda.$$

4. 几何分布

例 9　设 $X\sim G(p)$，求 $E(X)$.

解　X 的分布律为

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

于是
$$\begin{aligned}E(X)&=\sum_{k=1}^{\infty}k(1-p)^{k-1}p=p\sum_{k=1}^{\infty}k(1-p)^{k-1}\\&=p\left(\sum_{k=1}^{\infty}q^{k}\right)'\Big|_{q=1-p}=p\left(\frac{q}{1-q}\right)'\Bigg|_{q=1-p}\\&=\frac{p}{(1-q)^{2}}\Bigg|_{q=1-p}=\frac{1}{p}.\end{aligned}$$

4.1.3　连续型随机变量的数学期望

对于连续型随机变量 X，设其概率密度为 $f(x)$，X 落入小区间 $[x,\ x+\mathrm{d}x)$ 的概率为

$$P\{x\leqslant X<x+\mathrm{d}x\}=\int_{x}^{x+\mathrm{d}x}f(x)\mathrm{d}x,$$

当 $\mathrm{d}x$ 很小时，$P\{x\leqslant X<x+\mathrm{d}x\}\approx f(x)\mathrm{d}x$. 在大量次试验数中，$X$ 的平均值接近 $\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$，由此我们得到连续型随机变量数学期望的定义．

定义 2　设连续型随机变量 X 的概率密度为 $f(x)$，若积分 $\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$ 绝对收敛，则称 $\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$ 为 X 的**数学期望**，记为 $E(X)$，即

$$E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x. \tag{4.1.2}$$

若积分 $\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$ 不绝对收敛，则称随机变量 X 的数学期望不存在．

例 10　已知随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<0,\\ \dfrac{x^{2}}{4}, & 0\leqslant x<2,\\ 1, & x\geqslant 2,\end{cases}$$

求 $E(X)$.

解　X 的概率密度为

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

则
$$E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\int_0^2\frac{x^2}{2}\mathrm{d}x=\frac{4}{3}.$$

下面我们来计算几种常见的连续型随机变量的数学期望．

1. 均匀分布

例 11　设 $X\sim U(a,b)$，求 $E(X)$.

解　X 的概率密度为

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

于是
$$E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\int_a^b\frac{x}{b-a}\mathrm{d}x=\frac{a+b}{2}.$$

2. 指数分布

例 12　设 $X\sim E(\lambda)$，求 $E(X)$.

解　X 的概率密度为

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

于是
$$\begin{aligned}E(X)&=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\int_0^{+\infty}x\lambda\mathrm{e}^{-\lambda x}\mathrm{d}x\\&=\left[-x\mathrm{e}^{-\lambda x}\right]_0^{+\infty}+\int_0^{+\infty}\mathrm{e}^{-\lambda x}\mathrm{d}x\\&=\left[-\frac{1}{\lambda}\mathrm{e}^{-\lambda x}\right]_0^{+\infty}=\frac{1}{\lambda}.\end{aligned}$$

3. 正态分布

例 13　设 $X\sim N(\mu,\sigma^2)$，求 $E(X)$.

解　X 的概率密度为

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

于是
$$E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{x}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x,$$

令 $\dfrac{x-\mu}{\sigma}=t$，则

$$E(X)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}(\mu+\sigma t)\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t=\frac{\mu}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t=\mu.$$

例 14　设连续型随机变量 X 的概率密度为

$$f(x)=\frac{1}{3}\varphi(x-\mu_1)+\frac{2}{3}\varphi(x-\mu_2),\ -\infty<x<+\infty,$$

其中 $\varphi(x)$ 为标准正态分布的概率密度，求 X 的数学期望 $E(X)$.

解

$$\begin{aligned}E(X)&=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\frac{1}{3}\int_{-\infty}^{+\infty}x\varphi(x-\mu_1)\mathrm{d}x+\frac{2}{3}\int_{-\infty}^{+\infty}x\varphi(x-\mu_2)\mathrm{d}x\\&=\frac{1}{3}\int_{-\infty}^{+\infty}(t+\mu_1)\varphi(t)\mathrm{d}t+\frac{2}{3}\int_{-\infty}^{+\infty}(t+\mu_2)\varphi(t)\mathrm{d}t\\&=\frac{1}{3}\mu_1+\frac{2}{3}\mu_2+\frac{1}{3}\int_{-\infty}^{+\infty}t\varphi(t)\mathrm{d}t+\frac{2}{3}\int_{-\infty}^{+\infty}t\varphi(t)\mathrm{d}t\\&=\frac{1}{3}\mu_1+\frac{2}{3}\mu_2.\end{aligned}$$

4.1.4　随机变量函数的数学期望

在有些问题中，经常需要求随机变量函数的数学期望．这时，我们可以通过下面的定理来计算．

定理 1　设随机变量 Y 是随机变量 X 的函数：$Y=g(X)$（g 是连续函数）.

（1）X 是离散型随机变量，其分布律为

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

若级数 $\sum\limits_{i=1}^{\infty}g(x_i)p_i$ 绝对收敛，则有

$$E(Y)=\sum_{i=1}^{\infty}g(x_i)p_i. \tag{4.1.3}$$

（2）X 是连续型随机变量，其概率密度为 $f(x)$，若积分 $\int_{-\infty}^{+\infty}g(x)f(x)\mathrm{d}x$ 绝对收敛，则有

$$E(Y)=\int_{-\infty}^{+\infty}g(x)f(x)\mathrm{d}x. \tag{4.1.4}$$

例 15　设随机变量 X 的分布律为

X	-1	0	1	2
P	0.3	0.4	0.25	0.05

求 $E(3X^2+1)$.

解　$E(3X^2+1)=4\times0.3+1\times0.4+4\times0.25+13\times0.05=3.25.$

例 16　设 $X\sim U(-1,\ 1)$，求 $Y=\sin\left(\frac{\pi}{2}X\right)$的数学期望．

解　X 的概率密度为

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

则
$$E(Y)=\int_{-\infty}^{+\infty}g(x)f(x)\mathrm{d}x=\frac{1}{2}\int_{-1}^{1}\sin\left(\frac{\pi}{2}x\right)\mathrm{d}x=0.$$

例 17　设 $X\sim N(\mu,\sigma^2)$，求 $E(X^2)$.

解　$E(X^2)=\displaystyle\int_{-\infty}^{+\infty}x^2f(x)\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{x^2}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x$，

令 $\dfrac{x-\mu}{\sigma}=t$，则

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

例 18　设某农产品每周的需求量 X(单位：kg)在区间(10，30)内服从均匀分布，每售出 1kg 可获利 50 元；若供大于求，则需要处理剩余产品，每处理 1kg 亏损 10 元；若供不应求，则可调剂产品，但此时每售出 1kg 仅获利 30 元，问每周应进多少货，才能使商店所获利润的数学期望最大？

解　设每周此农产品的进货量为 y(单位：kg)($10<y<30$)，Y(单位：元)是商店所获利润，则 Y 是随机变量 X 的函数，其函数关系为

$$\begin{aligned}Y=g(X)&=\begin{cases}50X-10(y-X), & 10<X<y,\\ 50y+30(X-y), & y\leqslant X<30\end{cases}\\&=\begin{cases}60X-10y, & 10<X<y,\\ 30X+20y, & y\leqslant X<30.\end{cases}\end{aligned}$$

X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{20}, & 10<x<30,\\ 0, & \text{其他},\end{cases}$$

则
$$E(Y)=\int_{-\infty}^{+\infty}g(x)f(x)\mathrm{d}x=\frac{1}{20}\int_{10}^{30}g(x)\mathrm{d}x$$

$$= \frac{1}{20}\int_{10}^{y}(60x-10y)\mathrm{d}x + \frac{1}{20}\int_{y}^{30}(30x+20y)\mathrm{d}x$$

$$= -\frac{3}{4}y^2 + 35y + 525.$$

令$\frac{\mathrm{d}}{\mathrm{d}y}E(Y)=-\frac{3}{2}y+35=0$，得 $y=\frac{70}{3}$. 由此可知，商店每周应进$\frac{70}{3}$kg 该农产品，才可使利润的数学期望最大 .

定理 1 可以推广到多个随机变量的函数的情况 . 对两个随机变量的函数，我们有

定理 2　设随机变量 Z 是随机变量 X，Y 的函数：$Z=g(X, Y)$（g 是连续函数）.

(1) (X, Y)是离散型随机变量，其联合分布律为

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

若级数 $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} g(x_i, y_j)p_{ij}$ 绝对收敛，则有

$$E(Z) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} g(x_i, y_j)p_{ij}. \tag{4.1.5}$$

(2) (X, Y)是连续型随机变量，其联合概率密度为 $f(x, y)$，若积分

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

绝对收敛，则有

$$E(Z) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x, y)f(x, y)\mathrm{d}x\mathrm{d}y. \tag{4.1.6}$$

特别地，当 $g(X, Y)=X$ 或 $g(X, Y)=Y$ 时，由式(4.1.5)和式(4.1.6)可以分别得到

$$E(X) = \begin{cases} \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} x_i p_{ij}, & (X, Y)\text{ 是离散型随机变量}, \\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x f(x, y)\mathrm{d}x\mathrm{d}y, & (X, Y)\text{ 是连续型随机变量}. \end{cases} \tag{4.1.7}$$

$$E(Y) = \begin{cases} \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} y_j p_{ij}, & (X, Y)\text{ 是离散型随机变量}, \\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y f(x, y)\mathrm{d}x\mathrm{d}y, & (X, Y)\text{ 是连续型随机变量}. \end{cases} \tag{4.1.8}$$

例 19　设随机变量(X, Y)的联合分布律为

$X \backslash Y$	-1	1	2
-1	0.25	0.10	0.30
2	0.15	0.10	0.10

求 $E(Y^2)$，$E(XY)$.

解　因为

(X, Y)	$(-1, -1)$	$(-1, 1)$	$(-1, 2)$	$(2, -1)$	$(2, 1)$	$(2, 2)$
P	0.25	0.10	0.30	0.15	0.10	0.10
Y^2	1	1	4	1	1	4
XY	1	-1	-2	-2	2	4

所以 Y^2，XY 的分布律分别为

Y^2	1	4
P	0.6	0.4

XY	-2	-1	1	2	4
P	0.45	0.10	0.25	0.10	0.10

于是
$$E(Y^2)=1\times 0.6+4\times 0.4=2.2,$$
$$E(XY)=-2\times 0.45+(-1)\times 0.10+1\times 0.25+2\times 0.10+4\times 0.10=-0.15.$$

例 20　设(X, Y)的联合概率密度为
$$f(x, y)=\begin{cases}1, & 0\leqslant x\leqslant 1,\ 0\leqslant y\leqslant 2-2x,\\ 0, & \text{其他},\end{cases}$$
求 $E(X)$，$E(Y)$，$E(XY)$.

解
$$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x, y)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^{2-2x}x\mathrm{d}y$$
$$=\int_0^1 2x(1-x)\mathrm{d}x=\frac{1}{3},$$
$$E(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x, y)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^{2-2x}y\mathrm{d}y$$
$$=\int_0^1 2(1-x)^2\mathrm{d}x=\frac{2}{3},$$
$$E(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x, y)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^{2-2x}xy\mathrm{d}y$$

$$= \int_0^1 2x(1-x)^2 \mathrm{d}x = \frac{1}{6}.$$

4.1.5　数学期望的性质

数学期望具有以下一些重要的性质(设下面等式所涉及的数学期望均存在).

(1) 设 C 是常数，则 $E(C)=C$.

(2) 设 X 是随机变量，C 是常数，则 $E(CX)=CE(X)$.

(3) 设 X，Y 是两个任意的随机变量，则 $E(X+Y)=E(X)+E(Y)$.

一般地，对任意 n 个随机变量 X_1，X_2，…，X_n，有

$$E(X_1+X_2+\cdots+X_n)=E(X_1)+E(X_2)+\cdots+E(X_n).$$

(4) 设 X，Y 是相互独立的随机变量，则 $E(XY)=E(X)E(Y)$.

一般地，若随机变量 X_1，X_2，…，X_n 相互独立，则

$$E(X_1X_2\cdots X_n)=E(X_1)E(X_2)\cdots E(X_n).$$

证　(1)和(2)由读者自己证明，我们证明(3)和(4)，这里我们只对连续型随机变量的情况进行证明，对于离散型随机变量的情况，读者可以仿而证之.

设(X, Y)的联合概率密度为 $f(x, y)$，则

$$\begin{aligned} E(X+Y) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)f(x, y)\mathrm{d}x\mathrm{d}y \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x, y)\mathrm{d}x\mathrm{d}y + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x, y)\mathrm{d}x\mathrm{d}y \\ &= E(X)+E(Y). \end{aligned}$$

对一般的情况，可由数学归纳法推得性质(3).

设(X, Y)的联合概率密度为 $f(x, y)$，其边缘概率密度为 $f_X(x)$，$f_Y(y)$. 因为 X 与 Y 相互独立，则 $f(x, y)=f_X(x)f_Y(y)$，所以

$$\begin{aligned} E(XY) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x, y)\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf_X(x)f_Y(y)\mathrm{d}x\mathrm{d}y \\ &= \left[\int_{-\infty}^{+\infty}xf_X(x)\mathrm{d}x\right]\left[\int_{-\infty}^{+\infty}yf_Y(y)\mathrm{d}y\right] = E(X)E(Y). \end{aligned}$$

对一般的情况，由数学归纳法不难推得性质(4).

例 21　对随机变量 X_1，X_2，X_3 有 $X_1 \sim U(0, 4)$，$X_2 \sim N(0, 3^2)$，$X_3 \sim P(5)$，求 $E(X_1-3X_2+2X_3)$.

解　$E(X_1-3X_2+2X_3)=E(X_1)-3E(X_2)+2E(X_3)$

$$=2-3\times0+2\times5=12.$$

例 22　把数字 1，2，…，10 任意地排成一列，如果数字 k 恰好出现在第

k 个位置上，则称为一个巧合，求巧合个数的数学期望．

解　设巧合个数为随机变量 X，为了便于计算，我们引入随机变量

$$X_k=\begin{cases}1, & \text{数字 } k \text{ 恰好出现在第 } k \text{ 个位置上},\\ 0, & \text{其他},\end{cases}\quad k=1,2,\cdots,10,$$

则 $X=\sum_{k=1}^{10}X_k$，由于数字 k 恰好出现在第 k 个位置上的概率为$\frac{1}{10}$，所以

$$E(X_k)=\frac{1}{10},\ k=1,2,\cdots,10,$$

从而
$$E(X)=E\left(\sum_{k=1}^{10}X_k\right)=\sum_{k=1}^{10}E(X_k)=1.$$

4.2　随机变量的方差

4.2.1　方差的概念

随机变量 X 的数学期望 $E(X)$ 是分布的一种位置特征，它刻画了 X 的取值总在 $E(X)$ 周围波动．但这个位置特征数无法反映出随机变量取值的“波动”大小．而要解决一些与随机数学相关的实际问题，有时我们不仅需要了解随机变量取值的平均程度，还要知道随机变量取值的“波动”大小．下面来看一个例子．

1. 引例

例 1　有甲、乙两个射手，用随机变量 X 和 Y 分别表示他们击中目标的环数，设 X 和 Y 的分布律为

X	8	9	10
P	0.1	0.8	0.1

Y	8	9	10
P	0.3	0.4	0.3

试比较两射手射击技术的优劣．

解　因为
$$E(X)=8\times0.1+9\times0.8+10\times0.1=9,$$
$$E(Y)=8\times0.3+9\times0.4+10\times0.3=9,$$

即甲、乙两射手击中的平均环数都是 9 环，所以从数学期望难以判断两人射击技术的优劣．但由分布律可发现，X 的取值比 Y 的取值更集中于均值 9，这表明甲的射击技术比较稳定，从而我们认为甲的射击技术优于乙．

由此例可见，研究随机变量取值的波动大小是十分重要的．那么如何用数

值来反映随机变量取值的“波动”大小呢？随机变量的取值当然不一定恰好是它的均值 $E(X)$，会有偏离．偏离的量 $X-E(X)$ 有正有负，为了不使正负偏离彼此抵消，我们一般考虑 $[X-E(X)]^2$，而不去考虑数学上难以处理的绝对值 $|X-E(X)|$．因为 $[X-E(X)]^2$ 仍是一个随机变量，所以取其均值 $E[X-E(X)]^2$ 就可以刻画 X 的“波动”程度，这个量称作 X 的方差．

2. 方差的概念与计算公式

定义 1　设 X 为一随机变量，若 $E[X-E(X)]^2$ 存在，则称 $E[X-E(X)]^2$ 为 X 的**方差**，记为 $D(X)$ 或 $\mathrm{Var}(X)$，即

$$D(X)=E[X-E(X)]^2, \tag{4.2.1}$$

并称 $\sqrt{D(X)}$ 为 X 的**均方差**或**标准差**，记为 $\sigma(X)$.

由定义可知，方差与标准差的功能相似，它们都是用来描述随机变量取值的集中与分散程度（即散布大小）的两个特征数．若方差与标准差越小，则随机变量 X 的取值就越集中；方差与标准差越大，随机变量 X 的取值越分散．

方差与标准差之间的差别主要在量纲上，由于标准差与所讨论的随机变量、数学期望有相同的量纲，其加减 $E(X)\pm k\sigma(X)$ 是有意义的（k 为正实数），所以在实际中，人们更乐于选用标准差，但标准差的计算必须通过方差才能算得.

另外，注意到：如果随机变量 X 的数学期望存在，其方差不一定存在；而当 X 的方差存在时，则数学期望必定存在，其原因在于 $|X|\leqslant X^2+1$ 总是成立的．

根据定义，我们有

（1）若 X 是离散型随机变量，其分布律为

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

则

$$D(X)=\sum_{i=1}^{\infty}[x_i-E(X)]^2 p_i. \tag{4.2.2}$$

（2）若 X 是连续型随机变量，其概率密度为 $f(x)$，则

$$D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2 f(x)\mathrm{d}x. \tag{4.2.3}$$

在计算方差 $D(X)$ 时，我们通常由数学期望的性质，用下面的计算公式：

$$D(X)=E(X^2)-[E(X)]^2. \tag{4.2.4}$$

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

例 2　已知随机变量 X 的概率密度为

$$f(x)=\begin{cases} ax^2+bx+c, & 0<x<1, \\ 0, & 其他, \end{cases}$$

且 $E(X)=0.5$，$D(X)=0.05$，求系数 a，b，c.

解　因为 $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=\int_0^1(ax^2+bx+c)\mathrm{d}x=1$，所以 $\frac{1}{3}a+\frac{1}{2}b+c=1$.

又 $E(X)=\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x=\int_0^1 x(ax^2+bx+c)\mathrm{d}x=0.5$，故

$$\frac{1}{4}a+\frac{1}{3}b+\frac{1}{2}c=0.5.$$

再 $D(X)=E(X^2)-[E(X)]^2$，于是 $E(X^2)=0.3$，即

$$E(X^2)=\int_{-\infty}^{+\infty} x^2 f(x)\mathrm{d}x=\int_0^1 x^2(ax^2+bx+c)\mathrm{d}x=0.3,$$

从而 $\frac{1}{5}a+\frac{1}{4}b+\frac{1}{3}c=0.3$.

综上所述，得 $a=-6$，$b=6$，$c=0$.

4.2.2　几种常见随机变量的方差

1. 0—1 **分布**

例 3　设 $X\sim B(1, p)$，求 $D(X)$.

解　X 的分布律为

X	0	1
P	$1-p$	p

于是
$$E(X^2)=0^2\times(1-p)+1^2\times p=p,$$
又因 $E(X)=p$，从而
$$D(X)=E(X^2)-[E(X)]^2=p-p^2=p(1-p).$$

2. 二项分布

例 4　设 $X\sim B(n, p)$，求 $D(X)$.

解　X 的分布律为

$$P\{X=k\}=C_n^k p^k q^{n-k},\ k=0, 1, 2, \cdots, n,\ q=1-p,$$

于是 $E(X^2)=\sum_{k=0}^{n} k^2 C_n^k p^k q^{n-k}=\sum_{k=1}^{n} k(k-1)C_n^k p^k q^{n-k}+\sum_{k=0}^{n} kC_n^k p^k q^{n-k}$

$$=n(n-1)p^2\sum_{k=2}^{n} C_{n-2}^{k-2} p^{k-2} q^{(n-2)-(k-2)}+np$$

$$=n(n-1)p^2(p+q)^{n-2}+np=n(n-1)p^2+np,$$

又因 $E(X)=np$，从而

$$D(X)=E(X^2)-[E(X)]^2=n(n-1)p^2+np-n^2p^2=np(1-p).$$

3. 泊松分布

例 5　设 $X \sim P(\lambda)$，求 $D(X)$.

解　X 的分布律为

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

于是
$$E(X^2)=\sum_{k=0}^{\infty}k^2\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}=\sum_{k=1}^{\infty}k(k-1)\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}+\sum_{k=0}^{\infty}k\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}$$
$$=\lambda^2\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}\mathrm{e}^{-\lambda}+\lambda \xlongequal{\text{令 } l=k-2} \lambda^2\sum_{l=0}^{\infty}\frac{\lambda^l}{l!}\mathrm{e}^{-\lambda}+\lambda=\lambda^2+\lambda,$$

又因 $E(X)=\lambda$，从而

$$D(X)=E(X^2)-[E(X)]^2=\lambda^2+\lambda-\lambda^2=\lambda.$$

4. 几何分布

例 6　设 $X \sim G(p)$，求 $D(X)$.

解　X 的分布律为

$$P\{X=k\}=q^{k-1}p,\ k=1,\ 2,\ \cdots,\ q=1-p,$$

于是
$$E(X^2)=\sum_{k=1}^{\infty}k^2q^{k-1}p=pq\sum_{k=1}^{\infty}k(k-1)q^{k-2}+\sum_{k=1}^{\infty}kq^{k-1}p$$
$$=pq\left(\sum_{k=1}^{\infty}q^k\right)''+\frac{1}{p}=pq\left(\frac{q}{1-q}\right)''+\frac{1}{p}$$
$$=\frac{2pq}{(1-q)^3}+\frac{1}{p}=\frac{2-p}{p^2},$$

又因 $E(X)=\dfrac{1}{p}$，从而

$$D(X)=E(X^2)-[E(X)]^2=\frac{2-p}{p^2}-\frac{1}{p^2}=\frac{1-p}{p^2}.$$

5. 均匀分布

例 7　设 $X \sim U(a,\ b)$，求 $D(X)$.

解　X 的概率密度为

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

于是
$$E(X^2)=\int_{-\infty}^{+\infty}x^2f(x)\mathrm{d}x=\int_a^b\frac{x^2}{b-a}\mathrm{d}x=\frac{a^2+ab+b^2}{3},$$

又因 $E(X)=\dfrac{a+b}{2}$，从而

$$D(X)=E(X^2)-[E(X)]^2=\frac{a^2+ab+b^2}{3}-\frac{(a+b)^2}{4}=\frac{(b-a)^2}{12}.$$

6. 指数分布

例 8　设 $X \sim E(\lambda)$，求 $D(X)$.

解　X 的概率密度为

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

于是

$$E(X^2)=\int_{-\infty}^{+\infty}x^2 f(x)\mathrm{d}x=\int_0^{+\infty}x^2\lambda e^{-\lambda x}\mathrm{d}x$$

$$\xlongequal{\text{令 }\lambda x=t}\frac{1}{\lambda^2}\int_0^{+\infty}t^{3-1}e^{-t}\mathrm{d}t=\frac{1}{\lambda^2}\Gamma(3)=\frac{2}{\lambda^2},$$

又因 $E(X)=\dfrac{1}{\lambda}$，从而

$$D(X)=E(X^2)-[E(X)]^2=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}.$$

7. 正态分布

例 9　设 $X \sim N(\mu,\sigma^2)$，求 $D(X)$.

解　由上一节的例 17 知，$E(X^2)=\mu^2+\sigma^2$，又因 $E(X)=\mu$，从而

$$D(X)=E(X^2)-[E(X)]^2=\mu^2+\sigma^2-\mu^2=\sigma^2.$$

例 10　某人有一笔资金，可投入两个项目：房地产和商业，其收益都与市场状态有关．若把未来市场划分为好、中、差三个等级，其发生的概率分别为 0.2、0.7、0.1. 通过调查，该投资人认为投资于房地产的收益 X(单位：万元)和投资于商业的收益 Y(单位：万元)的分布律分别为

X	11	3	-3
P	0.2	0.7	0.1

Y	6	4	-1
P	0.2	0.7	0.1

试问该投资者如何投资为好?

解　我们先考察数学期望(平均收益)：

$$E(X)=11\times 0.2+3\times 0.7+(-3)\times 0.1=4.0(\text{万元}),$$

$$E(Y)=6\times 0.2+4\times 0.7+(-1)\times 0.1=3.9(\text{万元}),$$

从平均收益看，投资房地产收益大，可比投资商业多收益 0.1 万元．

下面我们再来计算它们各自的方差：

$$D(X)=(11-4)^2\times 0.2+(3-4)^2\times 0.7+(-3-4)^2\times 0.1=15.4,$$

$$D(Y)=(6-3.9)^2\times0.2+(4-3.9)^2\times0.7+(-1-3.9)^2\times0.1=3.29,$$

及标准差

$$\sigma(X)=\sqrt{15.4}=3.92,\ \sigma(Y)=\sqrt{3.29}=1.81.$$

因为标准差越大，则收益的波动越大，从而风险也越大，所以从标准差看，投资房地产的风险比投资商业的风险大一倍多．若收益与风险综合权衡，该投资者还是应该选择投资商业为好，虽然平均收益少 0.1 万元，但风险要小一半以上．

4.2.3　方差的性质

根据方差的定义，可以证明方差具有以下一些重要的性质(设下面等式所涉及的数学期望、方差均存在).

(1) 设 X 是随机变量，则 $D(X)\geqslant0$.

证　①若 X 是离散型随机变量，其分布律为

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

因为 $p_i\geqslant0$，$i=1, 2, \cdots$，所以 $D(X)=\sum_{i=1}^{\infty}[x_i-E(X)]^2p_i\geqslant0$.

②若 X 是连续型随机变量，其概率密度为 $f(x)$，因为 $f(x)\geqslant0$，所以

$$D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2f(x)\mathrm{d}x\geqslant0.$$

(2) 设 C 是常数，则 $D(C)=0$.

证　$D(C)=E[C-E(C)]^2=E[C-C]^2=E(0)=0$.

(3) 设 X 是随机变量，C 是常数，则 $D(X+C)=D(X)$.

证

$$\begin{aligned}D(X+C)&=E[X+C-E(X+C)]^2\\&=E[X+C-E(X)-C]^2\\&=E[X-E(X)]^2=D(X).\end{aligned}$$

(4) 设 X 是随机变量，C 是常数，则 $D(CX)=C^2D(X)$.

证

$$\begin{aligned}D(CX)&=E[CX-E(CX)]^2\\&=E\{C[X-E(X)]\}^2\\&=C^2E[X-E(X)]^2\\&=C^2D(X).\end{aligned}$$

(5) 设 X，Y 是相互独立的随机变量，则 $D(X\pm Y)=D(X)+D(Y)$，

一般地，若随机变量 X_1，X_2，$\cdots$，X_n 相互独立，有

$$D(X_1\pm X_2\pm\cdots\pm X_n)=D(X_1)+D(X_2)+\cdots+D(X_n).$$

证　因为　$D(X\pm Y)=E[(X\pm Y)-E(X\pm Y)]^2$

$$
\begin{aligned}
&=E[(X-E(X))\pm(Y-E(Y))]^2\\
&=E\{[X-E(X)]^2\pm2[X-E(X)][Y-E(Y)]+[Y-E(Y)]^2\}\\
&=E[X-E(X)]^2+E[Y-E(Y)]^2\pm2E\{[X-E(X)][Y-E(Y)]\},
\end{aligned}
$$

又因为

$$
\begin{aligned}
&E\{[X-E(X))][Y-E(Y)]\}\\
&=E[XY-XE(Y)-YE(X)+E(X)E(Y)]\\
&=E(XY)-E(X)E(Y)-E(X)E(Y)+E(X)E(Y)\\
&=E(XY)-E(X)E(Y),
\end{aligned}
$$

而 X 与 Y 相互独立，则 $E(XY)=E(X)E(Y)$，所以 $E\{[X-E(X)][Y-E(Y)]\}=0$，从而 $D(X\pm Y)=D(X)+D(Y)$. 对一般的情况，可由数学归纳法推得 .

(6) $D(X)=0$ 的充分必要条件是 X 以概率 1 取常数 C，即 $P\{X=C\}=1$，其中 $C=E(X)$.

证略 .

例 11　设随机变量 X_1，X_2，X_3 相互独立，且 $X_1\sim U(0,4)$，$X_2\sim N(0,3^2)$，$X_3\sim P(5)$，求 $D(X_1-3X_2+2X_3)$.

解　$D(X_1-3X_2+2X_3)=D(X_1)+9D(X_2)+4D(X_3)$

$$=\frac{4}{3}+9\times9+4\times5=\frac{307}{3}.$$

4.2.4　协方差

1. 协方差的概念

对于二维随机变量(X,Y)，除了知道 X 与 Y 各自的数学期望和方差外，我们还需知道它们的相互关系 . 而协方差正是刻画 X 与 Y 之间相互关系的一个重要数字特征 .

定义 2　设(X,Y)为二维随机变量，若 $E\{[X-E(X)][Y-E(Y)]\}$存在，则称 $E\{[X-E(X)][Y-E(Y)]\}$为 X 与 Y 的**协方差**，记为 $\mathrm{Cov}(X,Y)$，即

$$\mathrm{Cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}.\tag{4.2.5}$$

由本章 4.2.3 节方差的性质(5)的证明，我们易得

$$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y),\tag{4.2.6}$$

常用这一式子来计算协方差 . 特别地，当 X 与 Y 相互独立时，$\mathrm{Cov}(X,Y)=0$.

2. 协方差的性质

根据协方差的定义，容易证明协方差具有以下一些性质 .

(1) $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$.

(2) 对常数 a，$\mathrm{Cov}(X, a)=0$.

(3) 对常数 a，b，$\mathrm{Cov}(aX, bY)=ab\mathrm{Cov}(X, Y)$.

(4) $\mathrm{Cov}(X_1+X_2, Y)=\mathrm{Cov}(X_1, Y)+\mathrm{Cov}(X_2, Y)$.

(5) (**施瓦茨(Schwarz)不等式**)对任意二维随机变量(X, Y)，若 X 与 Y 的方差都存在，则

$$[\mathrm{Cov}(X, Y)]^2 \leqslant D(X)D(Y).$$

由方差的性质(5)的证明过程可知

$$D(X \pm Y)=D(X)+D(Y) \pm 2\mathrm{Cov}(X, Y).$$

例 12　设随机变量 X 与 Y 的方差和协方差分别为 $D(X)=9$，$D(Y)=16$，$\mathrm{Cov}(X, Y)=8$，求 $D(X+Y)$，$D(X-Y)$.

解　$D(X+Y)=D(X)+D(Y)+2\mathrm{Cov}(X, Y)=9+16+2\times 8=41$，

$D(X-Y)=D(X)+D(Y)-2\mathrm{Cov}(X, Y)=9+16-2\times 8=9$.

例 13　设(X, Y)的联合概率密度为

$$f(x, y)=\begin{cases} \dfrac{1}{\pi}, & x^2+y^2 \leqslant 1, \\ 0, & \text{其他}, \end{cases}$$

证明：$\mathrm{Cov}(X, Y)=0$，但 X 与 Y 不相互独立.

证　$$f_X(x)=\int_{-\infty}^{+\infty} f(x, y)\mathrm{d}y=\begin{cases} \displaystyle\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}\mathrm{d}y, & -1 \leqslant x \leqslant 1, \\ 0, & \text{其他} \end{cases}$$

$$=\begin{cases} \dfrac{2\sqrt{1-x^2}}{\pi}, & -1 \leqslant x \leqslant 1, \\ 0, & \text{其他}, \end{cases}$$

$$E(X)=\int_{-\infty}^{+\infty} x f(x)\mathrm{d}x=\int_{-1}^{1} \frac{x\sqrt{1-x^2}}{\pi}\mathrm{d}x=0.$$

同理可得

$$f_Y(y)=\begin{cases} \dfrac{2\sqrt{1-y^2}}{\pi}, & -1 \leqslant y \leqslant 1, \\ 0, & \text{其他}, \end{cases} E(Y)=0.$$

显然 $f(x, y) \neq f_X(x) f_Y(y)$，即 X 与 Y 不相互独立.

又因为　$$E(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x, y)\mathrm{d}x\mathrm{d}y$$

$$=\int_{-1}^{1}\mathrm{d}x\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{xy}{\pi}\mathrm{d}y=0,$$

所以 $\mathrm{Cov}(X, Y)=E(XY)-E(X)E(Y)=0$，得证.

这个例子告诉我们，X 与 Y 相互独立时，一定有 $\mathrm{Cov}(X, Y)=0$；但 $\mathrm{Cov}(X, Y)=0$ 时，X 与 Y 不一定相互独立.

3. 矩

矩是随机变量的一种最广泛的数字特征，我们前面所学的数学期望、方差和协方差都是某种特殊的矩.

定义 3　设 X 与 Y 为随机变量，若

$$E(X^k),\ k=1, 2, \cdots$$

存在，则称它为 X 的 k **阶原点矩**.

若
$$E\{[X-E(X)]^k\},\ k=1, 2, \cdots$$

存在，则称它为 X 的 k **阶中心矩**.

若
$$E(X^kY^l),\ k, l=1, 2, \cdots$$

存在，则称它为 X 与 Y 的 $k+l$ **阶混合原点矩**.

若
$$E\{[X-E(X)]^k[Y-E(Y)]^l\},\ k, l=1, 2, \cdots$$

存在，则称它为 X 与 Y 的 $k+l$ **阶混合中心矩**.

显然，对随机变量 X 与 Y，数学期望 $E(X)$ 是 X 的一阶原点矩，方差 $D(X)$ 是 X 的二阶中心矩，协方差 $\mathrm{Cov}(X, Y)$ 是 X 与 Y 的 $1+1$ 阶混合中心矩.

例 14　设 $X\sim N(0, 1)$，求 X 的三阶和四阶中心矩.

解　$E[X-E(X)]^3=E(X^3)=\int_{-\infty}^{+\infty}\frac{x^3}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx=0$，

$$E[X-E(X)]^4=E(X^4)=\int_{-\infty}^{+\infty}\frac{x^4}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx=2\int_0^{+\infty}\frac{x^4}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx,$$

令 $\frac{x^2}{2}=t$，$E(X^4)=\frac{4}{\sqrt{\pi}}\int_0^{+\infty}t^{\frac{3}{2}}e^{-t}dt=\frac{4}{\sqrt{\pi}}\Gamma\left(\frac{5}{2}\right)=3.$

4.2.5　相关系数

1. 相关系数的概念

对于二维随机变量 (X, Y)，相关系数也是刻画 X 与 Y 之间相互关系的一个重要数字特征.

定义 4　设 (X, Y) 为二维随机变量，若 $D(X)>0$，$D(Y)>0$，$\mathrm{Cov}(X, Y)$ 存在，则称 $\frac{\mathrm{Cov}(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$ 为 X 与 Y 的**相关系数**，记为 ρ_{XY}，即

$$\rho_{XY}=\frac{\mathrm{Cov}(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}}. \tag{4.2.7}$$

相关系数 ρ_{XY} 是一个无量纲的量 .

2. 相关系数的性质

相关系数 ρ_{XY} 描述了 X 与 Y 之间的相互关系，但到底是哪种关系呢？下面我们来推导 ρ_{XY} 的重要性质和说明 ρ_{XY} 的含义 .

当随机变量 X 与 Y 之间有相关关系时，我们先考虑用 X 的线性函数 $aX+b$ 近似表示 Y，然后寻找合适的 a 与 b 使 $\varepsilon=E[Y-(aX+b)]^2$ 达到最小，因为 ε 越小表示 $aX+b$ 与 Y 的近似度越好 . 将 ε 分别关于 a 与 b 求偏导并令其等于零，得

$$\begin{cases}\dfrac{\partial\varepsilon}{\partial a}=2aE(X^2)+2bE(X)-2E(XY)=0,\\ \dfrac{\partial\varepsilon}{\partial b}=2b+2aE(X)-2E(Y)=0,\end{cases}$$

解得
$$a=\frac{\mathrm{Cov}(X,Y)}{D(X)},\quad b=E(Y)-E(X)\frac{\mathrm{Cov}(X,Y)}{D(X)},$$

此时，

$$\varepsilon_{\min}=(1-\rho_{XY}^2)D(Y). \tag{4.2.8}$$

由此，我们得到相关系数 ρ_{XY} 的两个重要性质：

(1) $|\rho_{XY}|\leqslant 1$.

证　由式(4.2.8)与 $E[Y-(aX+b)]^2\geqslant 0$，$D(Y)>0$，知 $1-\rho_{XY}^2\geqslant 0$，即 $|\rho_{XY}|\leqslant 1$.

(2) $|\rho_{XY}|=1$ 的充分必要条件是存在常数 a，b 使 $P\{Y=aX+b\}=1$.

证　必要性：若 $|\rho_{XY}|=1$，由式(4.2.8)知

$$E[Y-(aX+b)]^2=0,$$

又因为　$E[Y-(aX+b)]^2=D[Y-(aX+b)]+\{E[Y-(aX+b)]\}^2$,

所以
$$D[Y-(aX+b)]=0,\ \{E[Y-(aX+b)]\}^2=0,$$

由方差的性质 6 得到

$$P\{Y-(aX+b)=0\}=1,$$

即
$$P\{Y=aX+b\}=1.$$

充分性：若存在常数 a_0，b_0 使

$$P\{Y=a_0X+b_0\}=1,$$

即
$$P\{Y-(a_0X+b_0)=0\}=1,$$

从而
$$E[Y-(a_0X+b_0)]=0,\ D[Y-(a_0X+b_0)]=0,$$

于是　$E[Y-(a_0X+b_0)]^2=D[Y-(a_0X+b_0)]+\{E[Y-(a_0X+b_0)]\}^2=0$,

又由式(4.2.8)知

$$0\leqslant\min_{a,b}E[Y-(aX+b)]^2=(1-\rho_{XY}^2)D(Y)\leqslant E[Y-(a_0X+b_0)]^2=0,$$

故$(1-\rho_{XY}^2)D(Y)=0$，又因为$D(Y)>0$，所以$1-\rho_{XY}^2=0$，即$|\rho_{XY}|=1$.

由式(4.2.8)知，$|\rho_{XY}|$越大，X与Y之间的线性相关程度就越好．特别地，当$|\rho_{XY}|=1$时，由ρ_{XY}的性质(2)表明X与Y之间以概率1存在线性关系；当$\rho_{XY}=0$时，X与Y不存在线性相关关系，这时我们称X与Y**不相关**．

3. 相关与独立性的关系

随机变量X、Y的相互独立与线性相关之间有如下关系：

(1) 若随机变量X与Y相互独立，则$\mathrm{Cov}(X,Y)=0$，从而$\rho_{XY}=0$，即X与Y不相关．

(2) 若随机变量X与Y线性相关，则$\mathrm{Cov}(X,Y)\neq0$，从而$\rho_{XY}\neq0$，即X与Y不独立．

(3) 若随机变量X与Y不相关，则X与Y不一定相互独立．

例 15　设随机变量(X,Y)的联合分布律为

X \ Y	0	1
-1	0.15	0.10
0	0.20	0.30
1	0.15	0.10

求ρ_{XY}.

解　因为X，Y，XY的分布律分别为

X	-1	0	1
P	0.25	0.50	0.25

Y	0	1
P	0.50	0.50

XY	-1	0	1
P	0.10	0.80	0.10

于是
$$E(X)=0,\ E(Y)=0.5,\ E(XY)=0,$$
从而
$$\mathrm{Cov}(X,Y)=0,\ \rho_{XY}=0.$$

由于$\rho_{XY}=0$，所以X与Y不相关．但是$P\{X=-1,Y=0\}\neq P\{X=-1\}P\{Y=0\}$，从而$X$与$Y$不相互独立，其实我们可以发现$Y=X^2$.

例 16　设$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，求$\rho_{XY}$.

解　因为(X,Y)的联合概率密度为

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

X，Y 的边缘概率密度为

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}, \quad -\infty<x<+\infty,$$

$$f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}, \quad -\infty<y<+\infty,$$

所以　$E(X)=\mu_1$，$D(X)=\sigma_1^2$，$E(Y)=\mu_2$，$D(Y)=\sigma_2^2$.

而

$$\begin{aligned}\mathrm{Cov}(X, Y) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_1)(y-\mu_2)f(x, y)\mathrm{d}x\mathrm{d}y \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\mathrm{d}x\int_{-\infty}^{+\infty}(x-\mu_1)(y-\mu_2)\times \\ &\quad \exp\left[\frac{-1}{2(1-\rho^2)}\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right)^2\right]\mathrm{d}y,\end{aligned}$$

令 $\frac{1}{\sqrt{1-\rho^2}}\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right)=t$，$\frac{x-\mu_1}{\sigma_1}=u$，则

$$\begin{aligned}\mathrm{Cov}(X, Y) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(\sigma_1\sigma_2\sqrt{1-\rho^2}\,tu+\rho\sigma_1\sigma_2u^2)e^{-\frac{(u^2+t^2)}{2}}\mathrm{d}t\mathrm{d}u \\ &= \frac{\sigma_1\sigma_2\sqrt{1-\rho^2}}{2\pi}\int_{-\infty}^{+\infty}ue^{-\frac{u^2}{2}}\mathrm{d}u\int_{-\infty}^{+\infty}te^{-\frac{t^2}{2}}\mathrm{d}t+\frac{\rho\sigma_1\sigma_2}{2\pi}\int_{-\infty}^{+\infty}u^2e^{-\frac{u^2}{2}}\mathrm{d}u\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}\mathrm{d}t \\ &= \rho\sigma_1\sigma_2,\end{aligned}$$

于是

$$\rho_{XY}=\frac{\mathrm{Cov}(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=\rho.$$

这表明二维正态随机变量 (X, Y) 的联合概率密度中的参数 ρ 就是 X 与 Y 的相关系数．又由上一章我们知道，X 与 Y 相互独立的充要条件是 $\rho=0$，所以对于二维正态随机变量 (X, Y)，X 与 Y 相互独立的充要条件是 X 与 Y 不相关.

习　题　A

1. 设随机变量 $X\sim P(\lambda)$，且 $E[(X-1)(X-2)]=10$，则 $\lambda=$________.

2. 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}0.2e^{-0.2x}, & x>0, \\ 0, & x\leqslant 0,\end{cases}$$

则 $E(3X+1)=($　　$)$.

(A)15；　　(B)1.6；　　(C)16；　　(D)46.

3. 设 X 是随机变量且 $E(X)=1$，$D(X)=3$，则 $E(2X^2+6)=($　　$)$.

(A)14；　　(B)16；　　(C)8；　　(D)22.

4. 设随机变量 X 与 Y 的概率密度分别为

$$f_X(x)=\begin{cases}2e^{-2x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}\quad f_Y(y)=\begin{cases}3e^{-3y}, & y>0,\\ 0, & y\leqslant 0,\end{cases}$$

求 $E(X+Y)$，$E(2X-Y^2)$.

5. 设随机变量 $X\sim B(n, p)$，且 $E(X)=4$，$D(X)=3.2$，则 $n=$________，$p=$________.

6. 设随机变量 X 的可能取值为 1，2，3，且 $E(X)=2.3$，$D(X)=0.61$，则 X 的分布律为________.

7. 设随机变量 $X\sim U(-1, 2)$，且随机变量

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

则 $D(Y)=$________.

8. 设随机变量 X 与 Y 有 $D(X)=16$，$D(Y)=25$，$\rho_{XY}=0.2$，则 $D(X-Y)=$________.

9. 设随机变量 X 的 $E(X)=\mu$，$D(X)=\sigma^2$，则对任意常数 C，必有(　　).

(A) $E[(X-C)^2]=E(X^2)-C^2$；　　(B) $E[(X-C)^2]\geqslant E[(X-\mu)^2]$；

(C) $E[(X-C)^2]=E[(X-\mu)^2]$；　　(D) $E[(X-C)^2]\leqslant E[(X-\mu)^2]$.

10. 设随机变量 X 的分布律为

X	-1	0	1	2
P	0.15	0.40	0.25	0.20

求 $E(X)$，$D(X)$，$E(-3X+1)$，$D(-3X+1)$.

11. 设连续型随机变量 X 的分布函数为

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

求 A，$E(X)$，$D(X)$，$E(e^X)$.

12. 设随机变量 X 的概率密度为 $f(x)=Ce^{-|x|}$，$-\infty<x<+\infty$，求 C，$E(X)$，$D(X)$.

13. 设某汽车站每天 9:00～10:00,10:00～11:00 都恰有一辆客车到站，但到站的时刻是随机的，且两车到站的时间是相互独立的，其规律为

到站时刻	9:10 10:10	9:30 10:30	9:50 10:50
概率	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

有一旅客 9:20 到汽车站，求他等车时间 X(单位：min)的数学期望和方差．

14. 设随机变量 $X \sim N(2,\ 4)$，$Y \sim N(3,\ 9)$，且 X 与 Y 相互独立，则 $Z=2X-Y+3$ 的概率密度为________．

15. 设随机变量$(X,\ Y)$的联合分布律为

X \ Y	0	1
-1	0.07	0.08
0	0.18	0.32
1	0.15	0.20

则有(　　)．

(A)X 与 Y 不独立；　　(B)X 与 Y 相互独立；

(C)X 与 Y 相关；　　(D)X 与 Y 相互独立且不相关．

16. 设随机变量 X 与 Y 满足 $D(X+Y)=D(X-Y)$，则必有(　　)．

(A)X 与 Y 相互独立；　　(B)$D(X)=0$；

(C)$D(X)\cdot D(Y)=0$；　　(D)X 与 Y 不相关．

17. 将一枚硬币重复抛掷 n 次，若 X 与 Y 分别表示正面向上和反面向上的次数，则 X 与 Y 的相关系数为(　　)．

(A)0；　　(B)-1；　　(C)0.5；　　(D)1.

18. 设随机变量 X 与 Y 有 $D(X)=1$，$D(Y)=4$，$\rho_{XY}=-1$，求 $D(2X-3Y+5)$．

19. 袋中装有 2 只白球及 3 只黑球，现进行无放回地摸球，定义下列随机变量

$$X=\begin{cases}1, & \text{第一次摸出白球,}\\ 0, & \text{第一次摸出黑球,}\end{cases}\quad Y=\begin{cases}1, & \text{第二次摸出白球,}\\ 0, & \text{第二次摸出黑球,}\end{cases}$$

求$(X,\ Y)$的联合分布律及 ρ_{XY}．

20. 设随机变量$(X,\ Y)$的联合概率密度为

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

求 C，$E(X)$，$E(Y)$，ρ_{XY}.

习　题　B

1. 设随机变量 X 与 Y 相互独立，其概率密度分别为

$$f_X(x)=\begin{cases}\dfrac{1}{2}x, & 0<x<2,\\ 0, & \text{其他},\end{cases}\quad f_Y(y)=\begin{cases}\mathrm{e}^{-(y-3)}, & y>3,\\ 0, & \text{其他},\end{cases}$$

则 $E(XY)=$________.

2. 设随机变量 X 与 Y 的相关系数 $\rho_{XY}=0.9$，若 $Z=X-0.4$，则 Y 与 Z 的相关系数为________.

3. 设随机变量 X 与 Y 独立同分布，令 $U=X-Y$，$V=X+Y$，则随机变量 U 与 V 一定有(　　).

(A)U 与 V 不独立；　　(B)U 与 V 相互独立；

(C)U 与 V 不相关；　　(D)U 与 V 相关.

4. 设随机变量 X_1，X_2，…，X_n 独立同分布，且其方差为 $\sigma^2>0$，令 $Y=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$，则有(　　).

(A)$\mathrm{Cov}(X_1,Y)=\dfrac{\sigma^2}{n}$；　　(B)$\mathrm{Cov}(X_1,Y)=\sigma^2$；

(C)$D(X_1+Y)=\dfrac{(n+2)\sigma^2}{n}$；　　(D)$D(X_1-Y)=\dfrac{(n+1)\sigma^2}{n}$.

5. 某人用 5 把钥匙去开门，只有一把能打开，今逐个任取一把试开，假设(1) 打不开的钥匙不放回；(2) 打不开的钥匙放回，求在这两种情况下打开此门所需开门次数 X 的数学期望及方差.

6. 设加工的某种零件的内径 X(单位：mm)服从正态分布 $N(\mu,1)$，内径在 8 与 10 之间的为合格品，其余为不合格品. 已知销售一个零件的利润 Y(单位：元)和该零件的内径 X 有如下关系：

$$Y=\begin{cases}-1, & X<8,\\ 10, & 8\leqslant X\leqslant 10,\\ -2, & X>10,\end{cases}$$

求平均内径 μ 取何值时，销售一个零件所获利润的数学期望最大？

7. 设随机事件 A 与 B 满足 $P(A)=\dfrac{1}{4}$，$P(B|A)=\dfrac{1}{3}$，$P(A|B)=\dfrac{1}{2}$，定

义下列随机变量

$$X=\begin{cases}1, & A\text{ 发生},\\0, & A\text{ 不发生},\end{cases}\quad Y=\begin{cases}1, & B\text{ 发生},\\0, & B\text{ 不发生},\end{cases}$$

求(X, Y)的联合分布律及ρ_{XY}.

8. 设 A 与 B 是两随机事件，定义下列随机变量

$$X=\begin{cases}1, & A\text{ 发生},\\-1, & A\text{ 不发生},\end{cases}\quad Y=\begin{cases}1, & B\text{ 发生},\\-1, & B\text{ 不发生},\end{cases}$$

证明：X 与 Y 不相关的充要条件是 A 与 B 相互独立.

9. 设随机变量(X, Y)的联合概率密度为

$$f(x, y)=\begin{cases}C(x+y), & 0\leqslant x\leqslant 2,\ 0\leqslant y\leqslant 2,\\0, & \text{其他},\end{cases}$$

求 C，$E(X)$，$D(X+Y)$，ρ_{XY}.

10. 设随机变量(X, Y)在区域$\{(x, y)\mid 0\leqslant x\leqslant 2, 0\leqslant y\leqslant 1\}$上服从均匀分布，定义下列随机变量

$$U=\begin{cases}1, & X>Y,\\0, & X\leqslant Y,\end{cases}\quad V=\begin{cases}1, & X>2Y,\\0, & X\leqslant 2Y,\end{cases}$$

求(U, V)的联合分布律及ρ_{UV}.

11. 某商场对某种商品的销售情况作了统计，知顾客对该商品的日需求量 X 服从正态分布 $N(\mu, \sigma^2)$，且日平均销售量为 $\mu=40$(件)，销售机会在 30～50 件之间的概率为 0.5. 若进货不足，则每件利润损失为 70 元；若进货量过大，则因资金积压，每件损失 100 元，求日最优进货量.

第 5 章　大数定律与中心极限定理

随机现象的统计规律性只有在相同条件下进行大量重复试验或观察才呈现出来．要研究大量随机现象，就必须采用极限的方法，大数定律和中心极限定理就是使用极限方法研究大量随机现象统计规律性的．大体讲来，阐明大量重复试验的平均结果具有稳定性的一系列定律都称为大数定律；论证随机变量(试验结果)之和渐进服从某一分布的定理称为中心极限定理．概率论中极限定理的内容是很广泛的，其中最主要的是大数定律和中心极限定理．

5.1　切比雪夫不等式

在讨论极限定理之前，我们需要掌握下面的概率不等式．

定理　设随机变量 X 具有数学期望 $E(X)=\mu$，方差 $D(X)=\sigma^2$，则对于任意正数 ε，不等式

$$P\{|X-\mu|\geqslant\varepsilon\}\leqslant\frac{\sigma^2}{\varepsilon^2}\tag{5.1.1}$$

成立．这一不等式称为**切比雪夫**(Chebyshev)**不等式**．

证　我们只就连续型随机变量的情况来证明．设 X 的概率密度为 $f(x)$，则有

$$\begin{aligned}P\{|X-\mu|\geqslant\varepsilon\}&=\int_{|X-\mu|\geqslant\varepsilon}f(x)\mathrm{d}x\leqslant\int_{|X-\mu|\geqslant\varepsilon}\frac{|X-\mu|^2}{\varepsilon^2}f(x)\mathrm{d}x\\&\leqslant\frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}(x-\mu)^2f(x)\mathrm{d}x=\frac{\sigma^2}{\varepsilon^2}.\end{aligned}$$

切比雪夫不等式还可以写成：

$$P\{|X-\mu|<\varepsilon\}\geqslant1-\frac{\sigma^2}{\varepsilon^2}.\tag{5.1.2}$$

切比雪夫不等式说明，X 的方差越小，事件$\{|X-\mu|<\varepsilon\}$发生的概率就越大，而 X 的取值基本上就集中在它的期望附近．这就进一步说明了方差的意义：方差是一个描述随机变量取值与其平均值偏离程度的数字特征．

另外，切比雪夫不等式给出了在随机变量的分布未知，而只知道 $E(X)$和$D(X)$的情况下估计概率 $P\{|X-\mu|<\varepsilon\}$的界限．例如，式(5.1.2)中取 $\varepsilon=3\sigma$，

得到
$$P\{|X-\mu|<3\sigma\}\geqslant 1-\frac{\sigma^2}{9\sigma^2}=0.8889.$$
这个估计是比较粗糙的，如果已经知道随机变量的分布，那么所求的概率可以确切地计算出来，也就没有必要利用这一不等式来估计了．

例 1　设在每次试验中，事件 A 发生的概率为 0.5. 利用切比雪夫不等式估计：在 1000 次重复独立试验中，事件 A 发生的次数在 400～600 之间的概率.

解　以 X 表示 1000 次重复独立试验中 A 发生的次数，则 $X\sim B(1000, 0.5)$，因此
$$E(X)=np=1000\times 0.5=500,$$
$$D(X)=npq=1000\times 0.5\times 0.5=250.$$
由切比雪夫不等式，得
$$P\{400<X<600\}=P\{|X-500|<100\}\geqslant 1-\frac{250}{100^2}=0.975,$$
即在 1000 次重复独立试验中，事件 A 发生的次数在 400～600 之间的概率至少为 0.975.

例 2　设正常男性成人血液每毫升中含白细胞数为 X，若 $E(X)=7300$，$\sigma=700$，利用切比雪夫不等式估计正常男性成人血液每毫升中含白细胞数在 5200～9400 之间的概率．

解　由题设知 $E(X)=7300$，$D(X)=\sigma^2=700^2$，由切比雪夫不等式，得
$$P\{5200<X<9400\}=P\{|X-7300|<2100\}\geqslant 1-\frac{700^2}{2100^2}=\frac{8}{9}.$$

例 3　随机变量 X 的方差存在，则 $\mathrm{Var}(X)=0$ 的充要条件是存在常数 a 使得 $P\{X=a\}=1$.

证　充分性是显然的，下面证必要性．设 $\mathrm{Var}(X)=0$，这时 $E(X)$存在.

因为
$$\{|X-E(X)|>0\}=\bigcup_{n=1}^{\infty}\left\{|X-E(X)|>\frac{1}{n}\right\},$$
所以有
$$\begin{aligned}P\{|X-E(X)|>0\}&=P\left(\bigcup_{n=1}^{\infty}\left\{|X-E(X)|>\frac{1}{n}\right\}\right)\\&\leqslant\sum_{n=1}^{\infty}P\left\{|X-E(X)|>\frac{1}{n}\right\}\\&\leqslant\sum_{n=1}^{\infty}\frac{\mathrm{Var}(X)}{(1/n)^2}=0,\end{aligned}$$
其中最后一个不等式用到了切比雪夫不等式．再由概率的非负性可得

$$P\{|X-E(X)|>0\}=0,$$

因而有
$$P\{|X-E(X)|=0\}=1,$$

即
$$P\{X=E(X)\}=1,$$

这就证明了结论，且其中的常数 a 就是 $E(X)$.

此外，切比雪夫不等式作为一个理论工具，它的应用是普遍的，在下面大数定律的证明中将要用到它 .

5.2　大数定律

“概率是频率稳定值”，其中“稳定”一词是什么含义？在第 1 章我们从直观上描述稳定性：随机事件 A 的频率 $f_n(A)$ 当重复试验的次数 n 增大时总呈现出稳定性，稳定在某一个常数的附近摆动 . 但如何摆动仍没有说清楚，现在可用大数定律来彻底说清这个问题 .

定理 1（切比雪夫大数定律）　设 $\{X_n\}$ 为相互独立的随机变量序列，若每个 X_n 的方差存在，且有共同的上界，即存在常数 C，满足

$$D(X_n)\leqslant C，n=1，2，\cdots， \tag{5.2.1}$$

则对任意 $\varepsilon>0$，有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E(X_i)\right|\geqslant\varepsilon\right\}=0. \tag{5.2.2}$$

证　记 $Y=\dfrac{1}{n}\sum_{i=1}^{n}X_i$，则 $E(Y)=\dfrac{1}{n}\sum_{i=1}^{n}E(X_i)$，由独立性和(5.2.1)式得

$$D(Y)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)\leqslant\frac{C}{n},$$

再代入切比雪夫不等式得

$$\begin{aligned}0&\leqslant P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E(X_i)\right|\geqslant\varepsilon\right\}\\&=P\{|Y-E(Y)|\geqslant\varepsilon\}\leqslant\frac{D(Y)}{\varepsilon^2}\leqslant\frac{C}{\varepsilon^2 n}\to0，n\to\infty,\end{aligned}$$

即(5.2.2)式成立 .

这一结果在 1866 年被俄国数学家切比雪夫所证明 . 它是关于大数定律的一个较为普遍的结论，一些大数定律的古典结果是它的特例 .

将该定律应用于抽样调查，就会有如下结论：随着样本容量 n 的增加，样本均值将接近于总体均值 . 从而为统计推断中依据样本均值估计总体均值提供了理论依据 .

定理 2（伯努利大数定律）　设 S_n 是 n 次独立重复试验中事件 A 发生的次数，p 是事件 A 在每次试验中发生的概率，则对任意 $\varepsilon>0$，有

$$\lim_{n\to\infty}P\left\{\left|\frac{S_n}{n}-p\right|\geqslant\varepsilon\right\}=0. \tag{5.2.3}$$

证　设 $X_i=\begin{cases}1, & A\text{ 在第 } i \text{ 次试验中发生},\\ 0, & A\text{ 在第 } i \text{ 次试验中不发生},\end{cases}$ $i=1, 2, \cdots, n$，则 $S_n=\sum_{i=1}^{n}X_i$，其中 X_i，$i=1, 2, \cdots, n$ 相互独立，且都服从以 p 为参数的 0—1 分布，因而 $E(X_i)=p=E\left(\frac{S_n}{n}\right)$，$D(X_i)=pq$，$q=1-p(i=1, 2, \cdots, n)$，即 X_i 的方差存在，并有共同的上界 pq，由切比雪夫大数定律即得(5.2.3)式．

该定律是切比雪夫大数定律的特例，其含义是，当 n 足够大时，事件 A 出现的频率将几乎接近于其发生的概率 p，它以严格的形式表达了频率的稳定性，这也为我们在实际应用中用频率去估计概率提供了理论依据．

在参数估计中，用样本矩及其函数估计相应的总体矩及其函数，其理论依据即在于此．

例　设$\{X_n\}$是相互独立的随机变量序列，且

X_n	-3^n	0	3^n
p_n	$\frac{1}{3^{2n+2}}$	$1-\frac{2}{3^{2n+2}}$	$\frac{1}{3^{2n+2}}$

试证$\{X_n\}$服从大数定律．

证　由题意可知

$$E(X_n)=0,\ D(X_n)=\frac{2}{9},\ n=1, 2, \cdots,$$

由切比雪夫大数定律得随机变量序列$\{X_n\}$服从大数定律．

5.3　中心极限定理

在客观实际中有许多随机变量，它们是由大量的相互独立的随机因素的综合影响所形成的，而其中每一个别因素在总的因素中所起的作用都是微小的．这种随机变量往往近似地服从正态分布．这些事实究其原因是什么？另一方面，对一串随机变量序列$\{X_n\}$，对于其前 n 项和 $Y_n=\sum_{k=1}^{n}X_k$ 当 n 很大时求 Y_n 的分布是件复杂的事情，自然希望在一定条件下 Y_n 的极限分布能够存在，这是很有意义的工作．本节要介绍的中心极限定理就是研究 Y_n 在什么条件下，

它的极限分布恰是正态分布．为使问题的提法更具一般化和应用性，这一节我们感兴趣的问题是：随机变量序列$\{X_n\}$的前 n 项和 Y_n 经标准化后，即 $Z_n=\dfrac{Y_n-E(Y_n)}{\sqrt{D(Y_n)}}$在什么条件下，它们的极限分布是标准正态分布．

历史上，关于这个论题的第一个结果是由法国数学家棣莫弗(De Moivre)于 1730 年对伯努利试验情形证明的．在此后的 200 多年当中，有关独立随机变量和的极限分布的讨论，一直是概率论研究的一个中心，故称作中心极限定理．下面只介绍两个常用的中心极限定理．

定理 1(**林德伯格—列维**(Lindeberg－Levy)**定理**)　设随机变量 X_1，X_2，…，X_n，…相互独立，服从同一分布，且具有数学期望和方差：$E(X_k)=\mu$，$D(X_k)=\sigma^2$，$k=1$，2，…，则随机变量之和 $\sum\limits_{k=1}^{n}X_k$ 的标准化变量

$$Y_n=\frac{\sum\limits_{k=1}^{n}X_k-E\left(\sum\limits_{k=1}^{n}X_k\right)}{\sqrt{D\left(\sum\limits_{k=1}^{n}X_k\right)}}=\frac{\sum\limits_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma} \tag{5.3.1}$$

的分布函数 $F_n(x)$对于任意 x 满足

$$\begin{aligned}\lim_{n\to\infty}F_n(x)&=\lim_{n\to\infty}P\{Y_n\leqslant x\}=\lim_{n\to\infty}P\left\{\frac{\sum\limits_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}\leqslant x\right\}\\&=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2}\mathrm{d}t=\Phi(x).\end{aligned} \tag{5.3.2}$$

该定理也称为独立同分布的中心极限定理．

证明略．

这个定理说明 n 个独立同分布的随机变量之和，当 n 充分大时，近似服从正态分布，也就是说当 n 充分大时，不管 X_k 服从什么样的分布，只要其数学期望和方差存在，就可以利用正态分布对 $\sum\limits_{k=1}^{n}X_k$ 作理论分析或实际计算．

将(5.3.2)式改写成

$$\begin{aligned}Y_n&=\frac{\sum\limits_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}=\frac{\frac{1}{n}\sum\limits_{k=1}^{n}X_k-\mu}{\sigma/\sqrt{n}}\\&=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\stackrel{\cdot}{\sim}N(0,1)\text{ 或 }\overline{X}\stackrel{\cdot}{\sim}N\left(\mu,\frac{\sigma^2}{n}\right),\end{aligned} \tag{5.3.3}$$

这是独立同分布中心极限定理结果的另一个形式．这就是说，均值为 μ，方差

为 σ^2 的独立同分布的随机变量 X_1，X_2，…，X_n 的算术平均 $\overline{X}=\frac{1}{n}\sum_{k=1}^{n}X_k$，当 n 充分大时近似地服从均值为 μ，方差为 σ^2/n 的正态分布．这一结果是数理统计中大样本统计推断的基础．

定理 2(棣莫弗—拉普拉斯(De Moivre - Laplace)**极限定理)**　设随机变量 S_n 服从参数为 n，$p(0<p<1)$的二项分布，则对于任意的 x，有

$$\lim_{n\to\infty}P\left\{\frac{S_n-np}{\sqrt{np(1-p)}}\leqslant x\right\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt=\Phi(x). \tag{5.3.4}$$

证　S_n 可以分解为 n 个相互独立且服从同一 0—1 分布的诸随机变量 X_1，X_2，…，X_n 之和，即有 $S_n=\sum_{k=1}^{n}X_k$，其中 $X_k(k=1,2,\cdots,n)$的分布律为 $P\{X_k=i\}=p^i(1-p)^{1-i}$，$i=0,1$．因为 $E(X_k)=p$，$D(X_k)=p(1-p)$，$k=1,2,\cdots,n$，由定理 1 得

$$\lim_{n\to\infty}P\left\{\frac{S_n-np}{\sqrt{np(1-p)}}\leqslant x\right\}=\lim_{n\to\infty}P\left\{\frac{\sum_{k=1}^{n}X_k-np}{\sqrt{np(1-p)}}\leqslant x\right\}$$

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

这个定理表明，正态分布是二项分布的极限分布．当 n 充分大时，我们可以利用(5.3.4)式来计算二项分布的概率．下面举几个关于中心极限定理应用的例子．

例 1　设一批产品的强度服从期望为 14，方差为 4 的分布，每箱装有该产品 100 件，问：

(1) 每箱产品的平均强度超过 14.5 的概率是多少?

(2) 每箱产品的平均强度在 14～14.5 之间的概率是多少?

解　设 X_k 是第 k 件产品的强度，则 $E(X_k)=14$，$D(X_k)=4$，$k=1,2,\cdots,100$，记 $\overline{X}=\frac{1}{100}\sum_{k=1}^{100}X_k$，则由定理 1 得

$$Y_n=\frac{\sum_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}=\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma}=\frac{\overline{X}-14}{0.2}\stackrel{\cdot}{\sim}N(0,1),$$

于是

(1) $P\{\overline{X}>14.5\}=P\left\{\frac{\overline{X}-14}{0.2}>\frac{14.5-14}{0.2}\right\}\approx1-\Phi(2.5)=0.0062.$

（2）$P\{14<\overline{X}<14.5\}=P\left\{\dfrac{14-14}{0.2}<\dfrac{\overline{X}-14}{0.2}<\dfrac{14.5-14}{0.2}\right\}$

$$\approx\Phi(2.5)-\Phi(0)=0.4938.$$

例 2　现有一批种子，其中良种占$\dfrac{1}{6}$. 今从其中任选 6000 粒，试问：在这批种子中，良种所占的比例与$\dfrac{1}{6}$之差小于 1％的概率是多少？

解　选一粒种子可以看成一次伯努利试验．设 X 表示 6000 粒种子中的良种数，则有 $X\sim B\left(6000,\dfrac{1}{6}\right)$，由定理 2，有

$$P\left\{\left|\frac{X}{6000}-\frac{1}{6}\right|<0.01\right\}=P\left\{\frac{|X-6000\times1/6|}{\sqrt{6000\times1/6\times5/6}}<\frac{0.01\times\sqrt{6000}}{\sqrt{1/6\times5/6}}\right\}$$
$$\approx2\Phi(2.078)-1=0.9624.$$

下列关系式在解决一些实际问题时很有用．

当 n 较大时，有

$$P\left\{\left|\frac{S_n}{n}-p\right|<\varepsilon\right\}=P\{|S_n-np|<n\varepsilon\}=P\left\{-\varepsilon\sqrt{\frac{n}{pq}}<\frac{S_n-np}{\sqrt{npq}}<\varepsilon\sqrt{\frac{n}{pq}}\right\}$$
$$\approx2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right)-1. \tag{5.3.5}$$

第一类问题是已知 n，p，ε，求概率 $P\left\{\left|\dfrac{S_n}{n}-p\right|<\varepsilon\right\}$. 利用(5.3.5)式，并查正态分布函数 $\Phi(x)$的数值表即可解决．这类问题在二项分布计算中常常遇到，如例 2.

第二类问题是欲使$\dfrac{S_n}{n}$与 p 的差异不大于定数ε 的概率不小于预先给定的数 β，问至少需要做多少次试验？这时只需要求出满足式

$$2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right)-1\geqslant\beta$$

的最小的 n，这也可以通过查表求得．

第三类问题是已知 n，β，求 ε. 这类问题是在进行误差估计时提出来的．解法如下：先找 x_β，使 $2\Phi(x_\beta)-1=\beta$. 这时，$\varepsilon=x_\beta\sqrt{\dfrac{pq}{n}}$即为所求．若 p 未知，可利用 $pq\leqslant\dfrac{1}{4}$，则有估计式：$\varepsilon\leqslant\dfrac{x_\beta}{2\sqrt{n}}$.

例 3　（1）一个复杂系统由 100 个相互独立的元件组成．在系统运行期间，每个元件损坏的概率为 0.1. 为了使系统正常运行，至少要 85 个以上的元

件工作，求系统的可靠度，即正常运行的概率.(2) 上述系统若由 n 个相互独立的元件组成，而且又要求至少有 80%的元件工作才能使整个系统正常工作，问 n 至少为多少时才能保证系统的可靠度为 0.95?

解　(1) 设 X 为系统正常运行时完好的元件个数，则

$$X\sim B(100,\ 0.9),\ E(X)=100\times0.9=90,\ D(X)=100\times0.9\times0.1=9.$$

由定理 2，得

$$P\{X>85\}=1-P\{X\leqslant85\}=1-P\left\{\frac{X-90}{\sqrt{9}}\leqslant\frac{85-90}{\sqrt{9}}\right\}$$
$$\approx1-\Phi\left(-\frac{5}{3}\right)=0.952.$$

(2) $X\sim B(n,\ 0.9)$，$E(X)=0.9n$，$D(X)=0.09n$，n 应满足 $P\{X\geqslant 0.8n\}=0.95$，而

$$P\{X\geqslant0.8n\}=P\left\{\frac{X-0.9n}{0.3\sqrt{n}}\geqslant\frac{0.8n-0.9n}{0.3\sqrt{n}}\right\}=P\left\{\frac{X-0.9n}{0.3\sqrt{n}}\geqslant\frac{-\sqrt{n}}{3}\right\}$$
$$\approx1-\Phi\left(-\frac{\sqrt{n}}{3}\right)=\Phi\left(\frac{\sqrt{n}}{3}\right),$$

所以 $\Phi\left(\frac{\sqrt{n}}{3}\right)=0.95$，查表得 $\frac{\sqrt{n}}{3}=1.65$，$n=24.5$，取 $n=25$.

例 4　某电站供应 10000 户居民用电，设在高峰时每户用电的概率为 0.8，且各户用电量多少是相互独立的，求：

(1) 同一时刻有 8100 户以上用电的概率.

(2) 若每户用电功率为 1000W，则电站至少需要多少电功率才能保证以 0.975 的概率供应居民用电?

解　(1) 设随机变量 Y_n 表示在 10000 户中同时用电的用户，则 $Y_n\sim B(10000,\ 0.8)$，于是

$$np=10000\times0.8=8000,\quad \sqrt{np(1-p)}=\sqrt{10000\times0.8\times0.2}=40,$$

所求概率为

$$P\{8100\leqslant Y_n\leqslant10000\}=P\left\{2.5\leqslant\frac{Y_n-np}{\sqrt{np(1-p)}}\leqslant50\right\}\approx\Phi(50)-\Phi(2.5)$$
$$=1-0.9938=0.0062.$$

(2) 若每户用电功率为 1000W，则 Y_n 户用电功率为 $1000Y_n$W. 设供电站功率为 QW，则由题意得

$$P\{1000Y_n\leqslant Q\}=P\left\{Y_n\leqslant\frac{Q}{1000}\right\}=P\left\{\frac{Y_n-np}{\sqrt{np(1-p)}}\leqslant\frac{Q/1000-8000}{40}\right\}$$
$$\approx\Phi\left(\frac{Q/1000-8000}{40}\right)=0.975,$$

查表可知 $\Phi(1.96)=0.975$，故$\dfrac{Q/1000-8000}{40}=1.96$，$Q=8078400$，所以，电站供电功率不应少于 8078.4kW.

习　题　A

1. 设随机变量 X 服从参数为 λ 的泊松分布，$E(X)=6$，则 $P\{3<X<9\}\geqslant$ ________.

2. 设随机变量 X 和 Y 的数学期望分别是 2 和 -2，方差分别是 1 和 4，而相关系数为 0.5，则根据切比雪夫不等式有 $P\{|X+Y|\geqslant 6\}\leqslant$ ________.

3. 设 X_1，…，X_n，…是相互独立的随机变量序列，且它们都服从参数为 λ 的指数分布，则 $\lim\limits_{n\to\infty}P\left\{\dfrac{\lambda\sum\limits_{i=1}^{n}X_i-n}{\sqrt{n}}\leqslant x\right\}=$ ________.

4. 设 X_1，X_2，…，X_n，…是相互独立的随机变量序列，X_n 服从参数为 n 的指数分布($n=1$，2，…)，则下列随机变量序列不服从切比雪夫大数定律的是(　　).

(A)X_1，…，X_n，…；　　(B)X_1，2^2X_2，…，n^2X_n，…；

(C)X_1，$\dfrac{X_2}{2}$，…，$\dfrac{X_n}{n}$，…；　　(D)X_1，$2X_2$，…，nX_n，….

5. 在天平上重复称量一重为 a 的物品，假设每次称量的结果相互独立，且都服从正态分布 $N(a,0.2^2)$. 若以 $\overline{X}_n$ 表示 n 次称量结果的算术平均值，则为使 $P\{|\overline{X}_n-a|<0.1\}>0.95$，$n$ 的最小值应不小于自然数(　　).

(A)2；　　(B)4；　　(C)10；　　(D)16.

6. 一船舶在某海区航行，已知每遭受一次波浪的冲击，纵摇角大于 3°的概率为 $p=\dfrac{1}{3}$，若船舶遭受了 90000 次波浪冲击，问其中有 29500～30500 次纵摇角大于 3°的概率是多少？

7. 一保险公司有 10000 人参加人寿保险，每人每年付 12 元保险费，在一年内一个人死亡的概率为 0.006，死亡时，其家属可向保险公司领得 1000 元，试问：

(1) 保险公司亏本的概率有多大？

(2) 保险公司年利润为零的概率是多少？

(3) 保险公司年利润不少于 60000 元的概率是多少？

8. 已知在某十字路口，一周事故发生的数学期望为 2.2，标准差为 1.4.

(1) 以 $\overline{X}$ 表示一年(以 52 周计)内，此十字路口事故发生数的算术平均，试用中心极限定理求 $\overline{X}$ 的近似分布，并求 $P\{\overline{X}<2\}$.

(2) 求一年内事故发生数小于 100 的概率．

习　题　B

1. 设 X_1，X_2，…，X_{1000} 是相互独立的随机变量序列，且 $X_i \sim B(1, p)$，$i=1, 2, \cdots, 1000$，则下列各项不正确的是(　　).

(A) $\dfrac{1}{1000}\sum\limits_{i=1}^{1000} X_i \approx p$；

(B) $\sum\limits_{i=1}^{1000} X_i \sim B(1000, p)$；

(C) $P\left\{a < \sum\limits_{i=1}^{1000} X_i < b\right\} \approx \Phi(b) - \Phi(a)$；

(D) $P\left\{a < \sum\limits_{i=1}^{1000} X_i < b\right\} \approx \Phi\left(\dfrac{b-1000p}{\sqrt{1000pq}}\right) - \Phi\left(\dfrac{a-1000p}{\sqrt{1000pq}}\right)$.

2. 设 X，Y 是两个相互独立的随机变量，则下列说法中正确的是(　　).

(A)当已知 X 与 Y 的分布时，对于随机变量 $X+Y$ 可使用切比雪夫不等式进行概率估计；

(B)当 X 与 Y 的期望和方差都存在时，可用切比雪夫不等式估计 $X+Y$ 落在任意区间 (a, b) 内的概率；

(C)当 X 与 Y 的期望和方差都存在时，可用切比雪夫不等式估计 $X+Y$ 落在对称区间 $(-a, a)$ 内的概率(常数 $a>0$)；

(D)当 X 与 Y 的期望和方差都存在时，可用切比雪夫不等式估计 $X+Y$ 落在区间 $(E(X)+E(Y)-a, E(X)+E(Y)+a)$ 内的概率(常数 $a>0$).

3. 计算机在进行加法运算时，每个加数取整数(取最为接近于它的整数)，设所有的整数误差是相互独立的，且它们都在 $[-0.5, 0.5]$ 上服从均匀分布．

(1) 将 1500 个数相加，误差总和的绝对值超过 15 的概率是多少？

(2) 最多多少个数加在一起可使误差总和的绝对值小于 10 的概率不小于 90%？

4. 设某种电子器件的使用寿命服从参数 $\lambda=0.1$(单位：h^{-1})的指数分布，其使用情况是第一个损坏第二个立即使用，第二个损坏第三个立即使用，依此类推．已知每个器件为 a 元，那么在年计划中一年至少需要多少元才能有 95% 的概率保证够用(假定一年有 306 个工作日，每个工作日为 8h)？

5. 假定生男孩的概率近似等于 0.515，求 10000 个婴孩中男孩不多于女孩的概率．

模　拟　试　题

模拟试卷一

一、填空题(每题 3 分，共 21 分)

1. 袋中有三个白球两个红球，从袋中任取两个球，则取得一个白球一个红球的概率为________.

2. 一批产品的次品率为 4%，正品中一等品率为 75%，现从这批产品中任意取一件，则恰好取到一等品的概率是________.

3. 设随机变量 X 服从指数分布，且已知随机变量 X 的平方的数学期望为 $E(X^2)=10000$，则随机变量 X 的期望 $E(X)=$________.

4. 设 A、B 是两个相互独立的事件，已知 $P(A)=0.3$，$P(B)=0.2$，则 $P(A\cup B)=$________.

5. 设随机变量 X 的密度函数为 $f(x)=\begin{cases}1-|x|, & |x|\leqslant 1,\\ 0, & \text{其他},\end{cases}$ 则在 $0\leqslant x<1$ 时，X 的分布函数为 $F(x)=$________.

6. 设连续型随机变量 X 的分布函数为

$$F(x)=\begin{cases}\dfrac{1}{3}\mathrm{e}^x, & x<0,\\ \dfrac{1}{3}(x+1), & 0\leqslant x<2,\\ 1, & x\geqslant 2,\end{cases}$$

则当 $x<0$ 时，其密度函数 $f(x)=$________.

7. 随机变量 X 在(0，5)上服从均匀分布 $U(0,5)$，则方程 $4t^2+4Xt+X+2=0$ 有实根的概率为________.

二、选择题(每题 3 分，共 21 分)

8. 在 10～99 的所有两位数中，任取一个数，求这个数能被 2 或 3 整除的概率是(　　).

(A) $\frac{1}{6}$；　(B) $\frac{1}{3}$；　(C) $\frac{2}{3}$；　(D) $\frac{5}{6}$.

9. 设当事件 A、B 同时发生时必导致事件 C 发生，则（　　）.

(A)$P(AB)=P(C)$；　　(B)$P(AB)\geqslant P(C)$；

(C)$P(C)\geqslant P(A)+P(B)-1$；　　(D)$P(C)\leqslant P(A)+P(B)-1$.

10. 在一等品占 20% 的一批产品中进行重复抽样检查，取 6 件进行检查，则恰有三件是一等品的概率为（　　）.

(A)0.2^3；　　(B)$0.2^3\times 0.8^3$；

(C)$0.2^3\times 20$；　　(D)$20\times 0.2^3\times 0.8^3$.

11. 打靶 3 发，事件 A_i 表示“击中 i 发”$i=0, 1, 2, 3$，则事件 $A=A_0\cup A_1\cup A_2$ 表示（　　）.

(A)至少击中 2 发；　　(B)至多击中 2 发；

(C)恰好击中 2 发；　　(D)3 发都击中.

12. 设连续型随机变量 ξ 的概率密度函数为

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

且 $P\left\{\xi>\frac{1}{2}\right\}=\frac{13}{20}$，则（　　）.

(A)$a=\frac{3}{5}$，$b=\frac{6}{5}$；　　(B)$a=\frac{6}{5}$，$b=\frac{3}{5}$；

(C)$a=2$，$b=1$；　　(D)$a=1$，$b=2$.

13. 离散型随机变量 ξ 的概率分布为 $P\{\xi=k\}=b\lambda^k (k=1, 2, \cdots)$ 的充分必要条件是（　　）.

(A)$b>0$ 且 $0<\lambda<1$；　　(B)$b=1-\lambda$ 且 $0<\lambda<1$；

(C)$b=\frac{1}{\lambda}-1$ 且 $\lambda<1$；　　(D)$\lambda=\frac{1}{1+b}$ 且 $b>0$.

14. 设随机变量 X 的概率密度函数为 $f(x)=\frac{1}{\sqrt{\pi}}e^{-(x^2-2x+1)}$，则 X 的均值和方差分别为（　　）.

(A)$E(X)=-1$，$D(X)=1$；　　(B)$E(X)=1$，$D(X)=\frac{1}{2}$；

(C)$E(X)=\frac{1}{2}$，$D(X)=\frac{1}{\sqrt{2}}$；　　(D)$E(X)=-1$，$D(X)=\frac{1}{2}$.

三、计算题（每题 10 分，共 50 分）

15. 临床诊断记录表明，利用某种试验检查癌症具有如下的效果：对癌症患者进行试验结果呈阳性反应者占 95%，对非癌症患者进行试验结果呈阴性反应

者占 96%. 现在用这种试验对某市居民进行癌症普查，如果该市癌症患者数约占居民总数的 0.4%，求试验结果呈阳性反应的被检查者确实患有癌症的概率.

16. 已知随机变量 X 的分布律为

X	-2	-1	0	1	2	3
概率	$4a$	$1/12$	$3a$	a	$10a$	$4a$

求 $Y=X^2$ 的分布律.

17. 设随机变量 X 服从标准正态分布 $N(0, 1)$，求 $Y=X^2$ 的密度函数.

18. 盒子中装有 3 只黑球、2 只红球、2 只白球，在其中任取 4 只，以 X，Y 分别表示取出的黑球数和红球数，求：

(1) (X, Y)的联合分布；(2) X 和 Y 的边缘分布；(3) $E(XY)$.

19. 游客乘电梯从底层到电视塔顶层观光，电梯于每个整点的第 5min、25min 和 55min 从底层起行. 假设一乘客在早上 8 点的第 X 分钟到达底层候梯处，且 X 在$(0, 60)$上均匀分布，求该游客等候时间的数学期望.

四、证明题(本题 8 分)

20. 设 A，B，C 表示 3 个随机事件，若 $P(A)=a$，$P(B)=2a$，$P(C)=3a$，且 $P(AB)=P(AC)=P(BC)=b$，证明：$b\leqslant a\leqslant \frac{1}{4}$.

模 拟 试 卷 二

一、填空题(每题 3 分，共 30 分)

1. 设 A、B、C 为三个随机事件，则“三个事件中至少有一个不发生”可表示为________.

2. 某油漆公司发出 17 桶油漆，其中白漆 10 桶，黑漆 4 桶，红漆 3 桶，在搬运过程中所有标签脱落，交货人随意将这些油漆发给顾客，则一个订货为白漆 4 桶，黑漆 3 桶，红漆 2 桶的顾客，能按所订颜色如数得到订货的概率为________.

3. 设 $P(A)=\frac{1}{4}$，$P(B|A)=\frac{1}{3}$，$P(A|B)=\frac{1}{2}$，则 $P(B)=$________，$P(A\cup B)=$________.

4. 将 3 只球随机地放入 4 个杯子中，以 X 表示杯子中球的最大个数，则 $P\{X=2\}=$________.

5. 设甲、乙两名篮球运动员投篮命中率分别为 0.7 和 0.6，每个人投 3

次，则两人进球数相等的概率为________.

6. 设随机变量 X 服从均值为 2，方差为 σ^2 的正态分布（即 $X \sim N(2, \sigma^2)$），且 $P\{2<X<4\}=0.3$，则 $P\{X<0\}=$________.

7. 一电话总机每分钟收到呼唤的次数服从泊松分布，已知每分钟有 3 次呼唤与每分钟有 2 次呼唤的概率相等，则某一分钟恰有 8 次呼唤的概率为________.（只用式子表示）

8. 设随机变量 X 的分布函数为 $F(x)=\begin{cases}1-\mathrm{e}^{-0.2(x-1)}, & x>1, \\ 0, & x\leqslant 1,\end{cases}$ 则 X 的方差 $D(X)=$________.

9. 设随机变量 (X, Y) 的联合分布律为

(X, Y)	(1, 1)	(1, 2)	(1, 3)	(2, 1)	(2, 2)	(2, 3)
概率	$\frac{1}{3}$	a	b	$\frac{1}{6}$	$\frac{1}{18}$	$\frac{1}{9}$

则 $a=$________，$b=$________时，随机变量 X 和 Y 相互独立.

10. 某车间生产的圆盘的直径在区间 (a, b) 上服从均匀分布，则圆盘面积的数学期望为________.

二、选择题（每题 3 分，共 15 分）

11. 随机事件 A、B 适合 $B\subseteq A$，则以下各式中错误的是（　　）.

(A) $P(A\cup B)=P(A)$；　　(B) $P(B|A)=P(B)$；

(C) $P(\overline{A}\cap\overline{B})=P(\overline{A})$；　　(D) $P(B)\leqslant P(A)$.

12. 设 $P(A\cup B)=a$，$P(\overline{A})=b$，$P(\overline{B})=c$，则 $P(A\overline{B})=$（　　）.

(A) $(a+c)c$；　(B) $a+c-1$；　(C) $a+b-c$；　(D) $(1-b)c$.

13. 设 $X\sim N(4, 4)$，且 c 满足 $P\{x>c\}=P\{x\leqslant c\}$，则 $c=$（　　）.

(A) 1；　(B) 2；　(C) 4；　(D) $\frac{1}{3}$.

14. 设随机变量 X 的密度函数为

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

则 $P\{X<1.5\}=$（　　）.

(A) 0.875；　　(B) 0.75；

(C) $\int_0^{1.5}(2-x)\,\mathrm{d}x$；　　(D) $\int_1^{1.5}(2-x)\,\mathrm{d}x$.

15. 利用切比雪夫不等式确定，要保证投掷一枚均匀的硬币“正面向上”的频率在 0.4 与 0.6 之间的概率不小于 90%，则至少需要投掷硬币的次数为(　　).

(A)100；　(B) 200；　(C) 250；　(D)300.

三、计算题(每题 10 分，共 50 分)

16. 对飞机进行独立射击三次，第一次射击命中率为 0.4，第二次射击命中率为 0.5，第三次射击命中率为 0.7，飞机击中一次而被击落的概率为 0.2，飞机击中二次而被击落的概率为 0.6，若被击中三次则飞机必然被击落，求：

(1) 射击三次而击落飞机的概率；

(2) 在射击三次飞机被击落的条件下，第一次射击没有击中飞机的概率.

17. 设 X 为离散型随机变量，其分布函数为

$$F(x)=\begin{cases}1-2^{-k}, & k-1\leqslant x<k,\ k=1,2,\cdots,\\ 0, & x<0,\end{cases}$$

求：(1) X 的概率分布律；(2) $E(2^{\frac{X}{2}})$.

18. 设连续型随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\dfrac{A}{1+x^2}, & |x|\leqslant 1,\\ 0, & \text{其他},\end{cases}$$

求：(1) 常数 A 的值；(2) $Y=\dfrac{1}{|X|}$的概率密度函数.

19. 袋中有 8 只白球，3 只红球，4 只黑球，从中任取 2 只，用 X，Y 表示取出的白球数和红球数，求：

(1) (X, Y)的联合分布；(2) X 和 Y 的边缘分布；(3) $P\{X+Y\leqslant 1\}$.

20. 一工厂生产某种设备的寿命 X(单位：年)服从指数分布，概率密度为

$$f(x)=\begin{cases}\dfrac{1}{4}e^{-\frac{x}{4}}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

工厂规定，出售的设备若在售出一年之内损坏可予以调换. 若工厂售出一台设备赢利 150 元，调换一台设备厂方需花费 200 元，试求厂方出售一台设备净利润的期望值.

四、证明题(本题 5 分)

21. 若离散型随机变量 X 的分布律为

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

则称 X 服从参数为 p 的几何分布 $G(p)$，证明取值为正整数的随机变量 X 服从几何分布的充分必要条件是对任意正整数 k，l，都有 $P\{X>k+l\mid X>k\}=P\{X>l\}$.

模 拟 试 卷 三

一、填空题(每题 3 分，共 21 分)

1. 某年级来了 15 名新生，其中有 3 名优秀运动员，现将这些新生任意平均地分到 3 个班中去，则每个班各分到一名优秀运动员的概率为________.

2. 甲、乙两人各自独立地向同一目标射击一次，结果发现目标被击中，设甲、乙射中目标的概率分别为 0.75，0.6，则是甲击中目标的概率为______.

3. 设随机变量 X 服从指数分布，且已知随机变量 X 的方差为 $D(X)=10000$，则随机变量 X 的期望 $E(X)=$________.

4. 掷两颗骰子，已知两颗骰子的点数之和为 7，则其中有一颗的点数为 1 的概率为________.

5. 设随机变量 X 的密度函数为 $f(x)=a\mathrm{e}^{-|x|}$，$-\infty<x<+\infty$，则 $x>0$ 时，X 的分布函数为 $F(x)=$________.

6. 设连续型随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<-a,\\ A+B\arcsin\left(\dfrac{x}{a}\right), & -a\leqslant x<a,\\ 1, & x\geqslant a,\end{cases}$$

则 X 的密度函数 $f(x)=$________.

7. 设随机变量 X 的分布函数为

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

则 $P\{X=0\}=$________.

二、选择题(每题 3 分，共 21 分)

8. 将两封信随机地投入 4 个邮筒中，则未向前面两个邮筒投信的概率为(　　).

(A)$\frac{2^2}{4^2}$；　(B)$\frac{C_2^1}{C_4^2}$；　(C)$\frac{2!}{A_4^2}$；　(D)$\frac{2!}{4!}$.

9. 设 A 和 B 是任意两个概率不为零的不相容事件，则下列结论中肯定正确的是(　　).

(A)$\overline{A}$与$\overline{B}$不相容；　(B)$\overline{A}$与$\overline{B}$相容；

(C)$P(AB)=P(A)P(B)$；　(D)$P(A-B)=P(A)$.

10. 每次射击命中目标的概率为$\frac{3}{4}$，某人连续射击直到命中为止，则射击次数为 3 的概率是(　　).

(A)$\left(\frac{3}{4}\right)^3$；　(B)$\left(\frac{3}{4}\right)^2\times\frac{1}{4}$；(C)$\left(\frac{1}{4}\right)^2\times\frac{3}{4}$；(D)$C_4^2\left(\frac{1}{4}\right)^2\ \frac{3}{4}$.

11. 设随机变量 X 服从参数为 λ 的泊松分布 $P(\lambda)$，且已知 $E(X+1)(X-3)=3$，则 λ 的值为(　　).

(A)$\lambda=3$；　(B)$\lambda=-2$；

(C)$\lambda=3$ 或 $\lambda=-2$；　(D)$\lambda=1$.

12. 下列各函数中是随机变量的分布函数的为(　　).

(A)$F_1(x)=\frac{1}{1+x^2}$，$-\infty<x<+\infty$；

(B)$F_2(x)=\mathrm{e}^x-\mathrm{e}^{-x}$，$-\infty<x<+\infty$；

(C)$F_3(x)=\mathrm{e}^{-x}$，$-\infty<x<+\infty$；

(D)$F_4(x)=\frac{1}{2}+\frac{1}{\pi}\arctan x$，$-\infty<x<+\infty$.

13. 离散型随机变量 ξ 的概率分布为 $P\{\xi=k\}=c\left(\frac{1}{3}\right)^k(k=1，2，\cdots)$的充分必要条件是(　　) .

(A)$c=2$；　(B)$c=1$；　(C)$c=\frac{1}{2}$；　(D)$c=\frac{3}{2}$.

14. 设随机变量 $X\sim B(n，p)$，且 $E(X)=2.4$，$D(X)=1.44$，则(　　).

(A)$n=6$，$p=0.6$；　(B)$n=8$，$p=0.3$；

(C)$n=6$，$p=0.4$；　(D)$n=24$，$p=0.1$.

三、计算题(每题 10 分，共 50 分)

15. 在一袋麦种中，其中一等麦种占 80%，二等麦种占 18%，三等麦种占 2%，已知一、二、三等麦种的发芽率分别为 0.8，0.5，0.2.(1) 现从袋中任取一粒麦种，求它发芽的概率；(2) 从袋中任取一粒麦种，播种后未发芽，求它是一等种子的概率 .

16. 设随机变量 X 的分布律为 $P\{X=k\}=\frac{1}{2^k}(k=1,2,\cdots)$，求 $Y=\sin\left(\frac{\pi}{2}X\right)$的概率分布．

17. 设随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\frac{A}{1+x^2}, & x>0,\\ 0, & \text{其他},\end{cases}$$

试求：(1) 常数 A；(2) $Y=\frac{1}{X}$的概率密度．

18. 将一枚硬币掷 3 次，以 X 表示前 2 次中出现正面的次数，以 Y 表示 3 次中出现的正面的次数，求：

(1) X 和 Y 的联合分布律；

(2) 关于 X 的边缘分布律，关于 Y 的边缘分布律．

19. 甲、乙两人在同样的条件下打靶，他们的命中情况的概率分布为

X(环)	5	6	7	8	9	10
$P_{甲}$	0.05	0.05	0.2	0.1	0.2	0.4
$P_{乙}$	0.1	0.1	0.1	0.2	0.1	0.4

试评定甲、乙两人射击技术的优劣．

四、证明题(本题 8 分)

20. 设 A，B 为两个随机事件，且 $0<P(A)<1$，$0<P(B)<1$，$P(B|A)+P(\overline{B}|\overline{A})=1$，则事件 A 和 B 相互独立．

附录　MATLAB 在概率方面的应用简介

MATLAB 是美国 MathWorks 公司在 20 世纪 80 年代中期推出的数学软件，优秀的数值计算能力和卓越的数据可视化能力使其很快在数学软件中脱颖而出．随着版本的不断升级，它在数值计算及符号计算功能上得到了进一步完善．MATLAB 已经发展成为多学科、多种工作平台的功能强大的大型软件．在欧美等高校，MATLAB 已经成为线性代数、自动控制理论、概率论及数理统计、数字信号处理、时间序列分析、动态系统仿真等高级课程的基本教学工具，是攻读学位的大学生、硕士生、博士生必须掌握的基本技能．

MATLAB 的主要特点是：

(1) 有高性能数值计算的高级算法，特别适合矩阵代数领域．

(2) 有大量事先定义的数学函数，并且有很强的用户自定义函数的能力．

(3) 有强大的绘图功能以及具有教育、科学和艺术学的图解和可视化的二维、三维图．

(4) 基于 HTML 的完整的帮助功能．

(5) 适合个人应用的强有力的面向矩阵(向量)的高级程序设计语言．

(6) 与其他语言编写的程序结合和输入、输出格式化数据的能力．

(7) 有在多个应用领域解决难题的工具箱．

以下我们主要对 MATLAB 在概率方面的内容做一些介绍．

1. 用 MATLAB 计算组合数、计算古典概率、验证概率的频率定义

(1) 计算组合数 C_n^k 时，使用语句 nchoosek(n，k)．例如，计算 C_{10}^4 时，输入

N＝nchoosek(10，4)(回车得到)

N＝210

(2) 计算阶乘数 n! 时，使用语句 factorial(n)．例如，计算 8! 时，输入

N＝factorial(8)(回车得到)

N＝40320

(3) 计算概率 $p_1=C_4^2C_6^2/C_{10}^4$，$p_2=1-6!\ C_{365}^6/365^6$ 时，分别输入

P1＝nchoosek(4，2) * nchoosek(6，2)/ nchoosek(10，4)(回车得到)

P1＝0.4286

P2＝1－factorial(6) * nchoosek(365，6)/365^6(回车得到)

P2＝0.0405

(4) 在计算机上将均匀的骰子投掷 400 次，计算各面出现的频率时，输入

x＝unidrnd(6，1，400)；

y＝tabulate(x)(计算各面出现的频率)(回车得到)

Value	Count	Percent
1.0000	61.0000	15.2500
2.0000	63.0000	15.7500
3.0000	66.0000	16.5000
4.0000	80.0000	20.0000
5.0000	65.0000	16.2500
6.0000	65.0000	16.2500

2. 用 MATLAB 产生随机数

(1) 产生 6 个参数为 10 和 0.5 的二项分布的随机数时，输入

R＝binornd(10，0.5，1，6)(回车可得)

R＝5　6　3　5　6　5

(2) 产生 3×6 个参数为 3.5 的泊松分布的随机数时，输入

R＝poissrnd(3.5，3，6)(回车可得)

R＝

5	5	2	1	0	1
5	6	1	4	4	3
4	5	3	6	3	3

(3) 产生 3×6 个均匀分布 $U(-2, 3)$的随机数时，输入

R＝unifrnd(－2，3，3，6)(回车可得)

R＝

2.6913	－1.4056	0.7841	－0.8404	1.9636	2.6110
－0.2843	－1.1549	0.4279	0.3933	－1.0350	－1.9337
0.8148	－0.6055	2.7611	0.6326	2.5480	1.8377

(4) 产生 3×8 个均匀分布 $N(2, 3^2)$的随机数时，输入

R＝normrnd(2，3，3，8)(回车可得)

R＝

0.7023	2.8630	5.5675	2.5239	0.2351	2.3418	1.7131	－2.0085
－2.9968	－1.4394	1.8871	1.4399	8.5496	5.2003	－0.4970	4.1430
2.3760	5.5727	2.9819	4.1774	1.5908	2.1778	2.8832	6.8707

MATLAB 自带的一些常用的分布的随机数产生函数见表 1，调用格式与

上面相同.

表1　随机数产生函数表

函数名	调用形式	注　释
Unifrnd	unifrnd(A，B，m，n)	(A, B)上均匀分布(连续)随机数
Unidrnd	unidrnd(N，m，n)	均匀分布(离散)随机数
Exprnd	exprnd(lambda，m，n)	参数为λ的指数分布随机数
Normrnd	normrnd(mu，sigma，m，n)	参数为μ，σ的正态分布随机数
chi2rnd	chi2rnd(N，m，n)	自由度为N的卡方分布随机数
Trnd	trnd(N，m，n)	自由度为N的t分布随机数
Frnd	frnd(N_1，N_2，m，n)	第一自由度为N_1，第二自由度为N_2的F分布随机数
gamrnd	gamrnd(A，B，m，n)	参数为A，B的γ分布随机数
betarnd	betarnd(A，B，m，n)	参数为A，B的β分布随机数
binornd	binornd(N，p，m，n)	参数为N，p的二项分布随机数
geornd	geornd(p，m，n)	参数为p的几何分布随机数
hygernd	hygernd(M，K，N，m，n)	参数为M，K，N的超几何分布随机数
Poissrnd	poissrnd(lambda，m，n)	参数为λ的泊松分布随机数

3. 用MATLAB计算一些常用分布的分布律或概率密度

(1) 已知$X \sim B(10, 0.6)$，计算概率$P\{X=4\}$时，输入

P＝binopdf(4，10，0.6)(回车可得)

P＝0.1115

(2) 已知$X \sim N(0, 1)$，求X的密度函数在3.4处的密度函数值$f(3.4)$时，输入

f＝normpdf(3.4，0，1)(回车可得)

f＝0.0012

MATLAB自带的一些常用分布的分布律或概率密度函数见表2，调用格式与上面相同.

表2　计算常用分布的概率密度函数值函数表

函数名	调用形式	注　释
Unifpdf	unifpdf(x，a，b)	(a, b)上均匀分布(连续)的概率密度在$X=x$处的函数值
unidpdf	Unidpdf(x，n)	均匀分布(离散)的概率密度函数值
Exppdf	exppdf(x，lambda)	参数为λ的指数分布的概率密度函数值
normpdf	normpdf(x，mu，sigma)	参数为μ，σ的正态分布的概率密度函数值
chi2pdf	chi2pdf(x，n)	自由度为n的卡方分布的概率密度函数值

（续）

函数名	调用形式	注　释
Tpdf	tpdf(x，n)	自由度为 n 的 t 分布的概率密度函数值
Fpdf	fpdf(x，n_1，n_2)	第一自由度为 n_1，第二自由度为 n_2 的 F 分布的概率密度函数值
gampdf	gampdf(x，a，b)	参数为 a，b 的 γ 分布的概率密度函数值
betapdf	betapdf(x，a，b)	参数为 a，b 的 β 分布的概率密度函数值
binopdf	binopdf(x，n，p)	参数为 n，p 的二项分布的概率密度函数值
geopdf	geopdf(x，p)	参数为 p 的几何分布的概率密度函数值
hygepdf	hygepdf(x，M，K，N)	参数为 M，K，N 的超几何分布的概率密度函数值
poisspdf	poisspdf(x，lambda)	参数为 λ 的泊松分布的概率密度函数值

4. 用 MATLAB 计算一些常用分布的分布函数值以及分布函数的反函数值

(1) 已知 $X\sim P(0.6)$，计算概率 $F(4)=P\{X\leqslant 4\}$时，输入

P＝poisscdf(4，0.6)(回车可得)

P＝0.9996

(2) 已知 $X\sim N(-2,3)$，求率 $F(0)=P\{X\leqslant 0\}$时，输入

F＝normcdf(0，－2，sqrt(3))(回车可得)

F＝0.8759

MATLAB 自带的一些常用分布的分布函数见表 3，调用格式与上面相同．

表 3　常用分布的分布函数值函数表

函数名	调用形式	注　释
unifcdf	unifcdf(x，a，b)	(a,b)上均匀分布(连续)的累积分布函数值 $F(x)=P\{X\leqslant x\}$
unidcdf	unidcdf(x，n)	均匀分布(离散)的累积分布函数值 $F(x)=P\{X\leqslant x\}$
expcdf	expcdf(x，lambda)	参数为 λ 的指数分布的累积分布函数值 $F(x)=P\{X\leqslant x\}$
normcdf	normcdf(x，mu，sigma)	参数为 μ，σ 的正态分布的累积分布函数值 $F(x)=P\{X\leqslant x\}$
chi2cdf	chi2cdf(x，n)	自由度为 n 的卡方分布的累积分布函数值 $F(x)=P\{X\leqslant x\}$
tcdf	tcdf(x，n)	自由度为 n 的 t 分布的累积分布函数值 $F(x)=P\{X\leqslant x\}$
fcdf	fcdf(x，n_1，n_2)	第一自由度为 n_1，第二自由度为 n_2 的 F 分布的累积分布函数值
gamcdf	gamcdf(x，a，b)	参数为 a，b 的 γ 分布的累积分布函数值 $F(x)=P\{X\leqslant x\}$
betacdf	betacdf(x，a，b)	参数为 a，b 的 β 分布的累积分布函数值 $F(x)=P\{X\leqslant x\}$
binocdf	binocdf(x，n，p)	参数为 n，p 的二项分布的累积分布函数值 $F(x)=P\{X\leqslant x\}$
geocdf	geocdf(x，p)	参数为 p 的几何分布的累积分布函数值 $F(x)=P\{X\leqslant x\}$
hygecdf	hygecdf(x，M，K，N)	参数为 M，K，N 的超几何分布的累积分布函数值
poisscdf	poisscdf(x，lambda)	参数为 λ 的泊松分布的累积分布函数值 $F(x)=P\{X\leqslant x\}$

(3) 已知 $X\sim N(0,1)$，求满足 $\Phi(x)=P\{X\leqslant x\}=0.975$ 的 x 时，输入

x＝norminv(0.975，0，1)(回车可得)

x＝1.9600

(4) 已知 $X\sim E(0.01)$，求满足 $F(x)=P\{X\leqslant x\}=0.628x$ 时，输入

x＝expinv(0.628，0.01)(回车可得)

x＝0.0099

MATLAB 自带的一些常用分布的分布函数的反函数见表 4，调用格式与上面相同．

表 4　常用临界值函数表

函数名	调用形式	注　释
unifinv	x＝unifinv(p，a，b)	均匀分布(连续)的逆累积分布函数($P=P\{X\leqslant x\}$，求 x)
unidinv	x＝unidinv(p，n)	均匀分布(离散)的逆累积分布函数，x 为临界值
expinv	x＝expinv(p，lambda)	指数分布的逆累积分布函数
norminv	x＝Norminv(x，mu，sigma)	正态分布的逆累积分布函数
chi2inv	x＝chi2inv(x，n)	卡方分布的逆累积分布函数
tinv	x＝tinv(x，n)	t 分布的累积分布函数
finv	x＝finv(x，n_1，n_2)	F 分布的逆累积分布函数
gaminv	x＝gaminv(x，a，b)	γ 分布的逆累积分布函数
betainv	x＝betainv(x，a，b)	β 分布的逆累积分布函数
binoinv	x＝binoinv(x，n，p)	二项分布的逆累积分布函数
geoinv	x＝geoinv(x，p)	几何分布的逆累积分布函数
hygeinv	x＝hygeinv(x，M，K，N)	超几何分布的逆累积分布函数
poissinv	x＝poissinv(x，lambda)	泊松分布的逆累积分布函数

5. 用 MATLAB 绘制密度函数曲线或分布函数曲线

设 $X\sim N(0,1)$，$Y\sim N(2,3)$，求 X，Y 的密度函数图像．

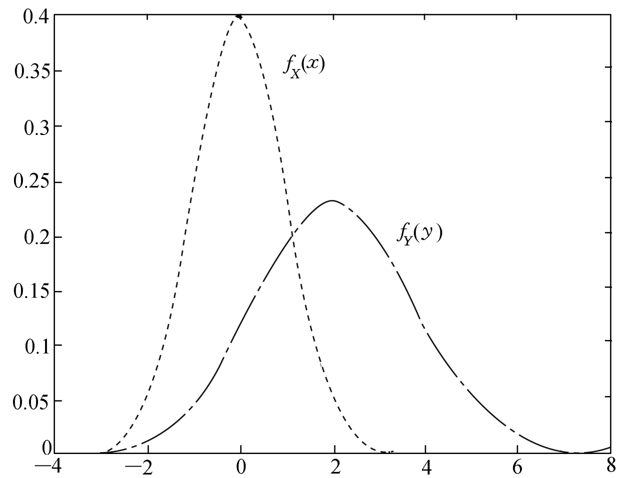
X，Y 的密度函数为

$$f_X(x)=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{x^2}{2}\right\},\ f_Y(y)=\frac{1}{\sqrt{6\pi}}\exp\left\{-\frac{(y-2)^2}{6}\right\}.$$

在命令行窗口输入命令行：

```
fplot(' normpdf(x，0，1)'，[－4，8]，' b－')
hold on
fplot(' normpdf(x，2，sqrt(3))'，[－4，8]，' r：')
hold off
```

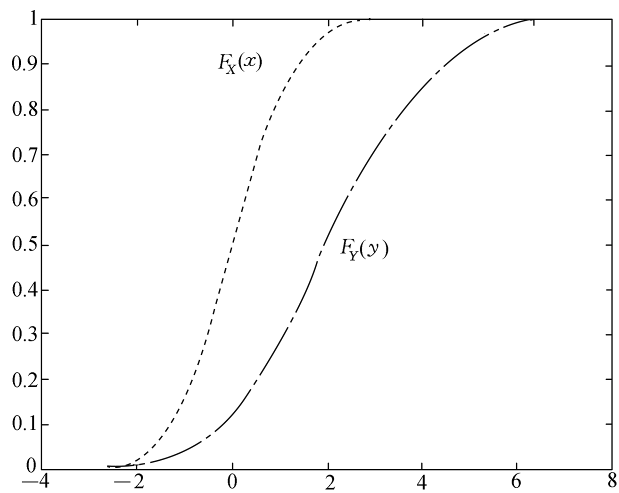
则在相应的图像窗口得密度函数图像为



在命令行窗口输入命令行：

```
fplot('normcdf(x，0，1)'，[−4，8]，'b—')
hold on
fplot('normcdf(x，2，sqrt(3))'，[−4，8]，'r：')
hold off
```

则在相应的图像窗口得分布函数图像为



6. 用 MATLAB 由分布律或概率密度函数计算均值或方差

(1) 利用 sum 函数(求和函数)计算离散型随机变量的数学期望：

例 1 设随机变量 X 的分布律为

X	-2	-1	0	1	2
P	0.3	0.1	0.2	0.1	0.3

求 $E(X)$，$E(X^2-1)$.

在 MATLAB 编辑器中建立 M 文件如下：

```
X=[-2  -1  0  1  2];
p=[0.3  0.1  0.2  0.1  0.3];
EX=sum(X.*p)
Y=X.^2-1
EY=sum(Y.*p)
```

运行后结果如下：

```
EX=0
Y=3  0  -1  0  3
EY=1.6000
```

(2) 利用 int 函数(求积分函数)计算连续型随机变量或其函数的数学期望：

例 2　设随机变量 X 服从参数为 0.01 的指数分布，即 X 的密度函数为

$$f(x)=\begin{cases}0.01\exp\{-0.01x\}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

求 $E(X)$，$E(X^2-1)$.

在 MATLAB 编辑器中建立 M 文件如下：

```
syms  x
f=0.01*exp(-0.01*x);
t=x*f;
EX=int(t, 0, inf)
y=(x^2-1)*f;
EY=int(y, 0, inf)
```

运行后结果如下：

```
EX=100
EY=19999
```

(3) 利用专有函数求常见分布的期望和方差：

例 3　设随机变量 X 服从区间(0，4)上的均匀分布 $U(0, 4)$，求其均值 $E(X)$和方差 $D(X)$.

在 MATLAB 编辑器中输入：

[EX，DX]＝unifstat(0，4)(回车可得)

EX＝2(均值)

DX＝1.3333(方差)

例4　设随机变量 X 服从参数为－2，3的正态分布 $N(-2，3)$，求其均值 $E(X)$ 和方差 $D(X)$.

在MATLAB编辑器中输入：

[EX，DX]＝normstat(－2，sqrt(3))(回车可得)

EX＝－2(均值)

DX＝3(方差)

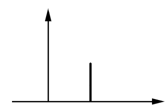
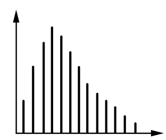
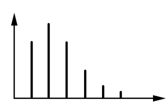
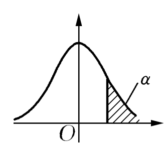
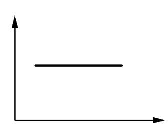
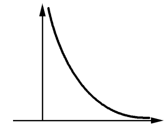
MATLAB自带的一些常用分布的数学期望和方差的专有函数见表5，调用格式与上面相同.

表5　常见分布的均值和方差的专有函数表

函数名	调用形式	注　释
unifstat	[M，V]＝unifstat(a，b)	均匀分布(连续)的期望和方差，M 为期望，V 为方差
unidstat	[M，V]＝unidstat(n)	均匀分布(离散)的期望和方差
expstat	[M，V]＝expstat(p，lambda)	指数分布的期望和方差
normstat	[M，V]＝normstat(mu，sigma)	正态分布的期望和方差
chi2stat	[M，V]＝chi2stat(x，n)	卡方分布的期望和方差
tstat	[M，V]＝tstat(n)	t 分布的期望和方差
fstat	[M，V]＝fstat(n_1，n_2)	F 分布的期望和方差
gamstat	[M，V]＝gamstat(a，b)	γ 分布的期望和方差
betastat	[M，V]＝betastat(a，b)	β 分布的期望和方差
Binostat	[M，V]＝binostat(n，p)	二项分布的期望和方差
Geostat	[M，V]＝geostat(p)	几何分布的期望和方差
hygestat	[M，V]＝hygestat(M，K，N)	超几何分布的期望和方差
Poisstat	[M，V]＝poisstat(lambda)	泊松分布的期望和方差

对于MATLAB在概率论或概率统计方面的其他应用我们就不过多的介绍了，有兴趣的同学可以参看MATLAB数学手册.

附表 1　几种常见的概率分布表

概率分布	概率与密度函数 $p(x)$	数学期望	方差	图形
伯努利分布 两点分布	$p_k=\begin{cases} q, & k=0, \\ p, & k=1, \end{cases}$ $0<p<1$，$q=1-p$.	p	pq	
二项分布 $b(k;n,p)$	$b(k;n,p)=C_n^k p^k q^{n-k}$， $k=0,1,\cdots,n$； $1<p<1$，$q=1-p$.	np	npq	
泊松分布 $p(k,\lambda)$	$p(k,\lambda)=\dfrac{\lambda^k}{k!}e^{-\lambda}$， $k=0,1,2,\cdots$，常数 $\lambda>0$.	λ	λ	
正态分布 $N(\mu,\sigma^2)$	$p(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{\sigma^2}}$， $-\infty<x<+\infty$，常数 μ，$\sigma>0$.	μ	σ^2	
均匀分布 $U(a,b)$	$p(x)=\begin{cases} \dfrac{1}{b-a}, & a<x<b, \\ 0, & 其他, \end{cases}$ 常数 $a<b$.	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$	
指数分布	$p(x)=\begin{cases} \lambda e^{-\lambda x}, & x>0, \\ 0, & x\leqslant 0, \end{cases}$ 常数 $\lambda>0$.	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$	

附表 2　标准正态分布表

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

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	d.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
x	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
3	0.9987	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

习 题 参 考 答 案

第1章　习题A

1. （1）$\Omega=\left\{\frac{0}{n},\ \frac{1}{n},\ \cdots,\ \frac{100n}{n}\right\}$；（2）$\Omega=\{2,\ 3,\ \cdots,\ 11,\ 12\}$；

（3）$\Omega=\{3,\ 4,\ \cdots,\ 9,\ 10\}$；（4）$\Omega=\{5,\ 6,\ \cdots\}$；

（5）$\Omega=\{v\,|\,v>0\}$（不考虑汽车的运动方向），

$\Omega=\{(v\cos\theta,\ v\sin\theta)\,|\,v>0,\ 0\leqslant\theta<2\pi\}$（考虑汽车的运动方向）；

（6）$\Omega=\{(x,\ y)\,|\,x^2+y^2<1,\ x,\ y\in\mathbf{R}\}$.

2. （1）$A_1\cup A_2$；（2）$A_1\overline{A}_2\overline{A}_3$；（3）$A_1A_2A_3$；

（4）$\overline{A}_1\cup\overline{A}_2\cup\overline{A}_3$ 或$\overline{A_1A_2A_3}$；（5）$A_1A_2\overline{A}_3\cup A_1\overline{A}_2A_3\cup\overline{A}_1A_2A_3$.

3. （1）$D=A_1\cup A_2\cup A_3$；（2）$D=A_1\cap A_2\cap A_3$.

4. (C).

5. $\frac{P_3^2P_3^3}{P_5^5}=\frac{3}{10}$.

6. （1）$\frac{132}{169}$；（2）$\frac{37}{169}$；（3）$\frac{168}{169}$.

7. $\frac{5}{12}$.

8. （1）$\frac{65}{66}$；（2）$\frac{1}{6}$；（3）$\frac{10}{33}$.

9. （1）$\frac{12}{25}$；（2）$\frac{13}{25}$.

10. （1）$P(A)=\frac{n!}{N^n}$；（2）$P(B)=\frac{n!\ C_N^n}{N^n}=\frac{N!}{N^n(N-n)!}$；

（3）$P(C)=\frac{C_n^m(N-1)^{n-m}}{N^n}=C_n^m\left(\frac{1}{N}\right)^m\left(\frac{N-1}{N}\right)^{n-m}$.

11. （1）$\frac{1}{5}$；（2）$\frac{3}{5}$；（3）$\frac{9}{10}$.

12. $\frac{1}{4}$.

13. 0.879.

14. （1）$\frac{5}{18}$；（2）$\frac{30}{91}$.

15. $\frac{1}{3}$.

16. （1）$\frac{89}{110}$；（2）$\frac{1}{11}$；（3）$\frac{89}{99}$；（4）$\frac{9}{1078}$.

17. $\frac{1}{4}$.

18. （1）$\frac{p(3-p)}{2}$；（2）$\frac{2p}{p+1}$.

19. （1）0.943；（2）0.848.

20. $\frac{36}{43}$.

21. 假设男女人数相同，0.0476.

22. （1）0.32；（2）$\frac{12}{29}$.

23. 0.7102.

24. 0.657；$n \geqslant 12.9114$，取 $n=13$.

25. 略.

第 1 章　习题 B

1. $\frac{7}{99}$.

2. （1）$\frac{1}{7}$；（2）$\frac{10}{21}$；（3）$\frac{1}{7}$；（4）$\frac{2}{7}$；（5）$\frac{4}{21}$；（6）$\frac{4}{21}$.

3. $\frac{1}{4}(1+\ln 4)$.

4～7. 略.

8. $1-\frac{1}{2!}+\frac{1}{3!}+\cdots+(-1)^{n-1} \cdot \frac{1}{n!}$.

9. （1）$p=\frac{29}{90}$；（2）$q=\frac{20}{61}$.

10. (B).

11. 略.

12. 0.8731；0.1268；0.0001.

13. $1-\frac{1}{6^4}-C_4^3 \times \frac{1}{6^3} \times \frac{1}{2}=\frac{1283}{1296}$.

14. 0.9984，$n \geqslant 2.8614$，取 $n=3$.

15～16. 略.

第 2 章　习题 A

1.

X	3	4	5
p_k	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{3}{5}$

2.

X	0	1	2	3
p_k	$\frac{3}{4}$	$\frac{9}{44}$	$\frac{9}{220}$	$\frac{1}{220}$

3.

X	0	1	2	…	n	…
p_k	p	pq	pq^2	…	pq^n	…

其中 $p+q=1$.

4. (1) 0.0298；(2) 0.09158；(3) 0.5665.

5. (1) $P\{X \geqslant 3\} = \sum_{k=3}^{5} C_5^k (0.3)^k (0.7)^{5-k} = 0.1631$；

(2) $P\{Y \geqslant 3\} = 1 - P\{Y < 3\} = 1 - \sum_{k=0}^{2} C_7^k (0.3)^k (0.7)^{7-k} = 0.3529$.

6. $$F(x)=\begin{cases} 0, & x<3, \\ \frac{1}{10}, & 3\leqslant x<4, \\ \frac{4}{10}, & 4\leqslant x<5, \\ 1, & x\geqslant 5. \end{cases}$$

7. $$F(x)=\begin{cases} 0, & x<0, \\ \frac{x}{4}, & 0\leqslant x<4, \\ 1, & x\geqslant 4. \end{cases}$$

8. (1) 不可以；(2) 可以，取 $F(x)=\begin{cases} \frac{1}{1+x^2}, & x\leqslant 0, \\ 1, & x>0. \end{cases}$

9. (1) $A=2$；(2) $P\{0<X<0.5\}=0.25$；

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

10. (1) $a=\frac{1}{2}$；(2) $P\{0<X<1\}\approx 0.316$；

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

11. (1) $A=1$；(2) $\ln 2$，$1-\ln 2$；(3) $f(x)=\begin{cases}\frac{1}{x}, & 1<x<e,\\ 0, & \text{其他}.\end{cases}$

12. (1) 0.5；(2) 0.0062；(3) 0.9070.

13. (1) 0.7257；(2) 0.895；(3) 0.8822.

14. 4.56%.

15. $\sigma=31.25$.

16. (1)

Y	0	1	4
p_k	0.3	0.4	0.3

(2)

Y	-5	-3	-1	1	3
p_k	0.1	0.2	0.3	0.2	0.2

17. $f_Y(y)=\begin{cases}\frac{1}{2}, & 2<y<4,\\ 0, & \text{其他}.\end{cases}$

18. $f_Y(y)=\begin{cases}\frac{1}{4\sqrt{y}}e^{-\frac{\sqrt{y}}{2}}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$

19. $f_Y(y)=\frac{1}{2\sqrt{2\pi}}e^{-\frac{(y-1)^2}{8}}$，$-\infty<y<+\infty$.

20. (1) $A=\frac{2}{\pi}$；(2) $f_Y(y)=\frac{2e^{y}}{\pi(1+e^{2y})}$，$-\infty<y<+\infty$.

21. $f_Y(y)=\begin{cases}\frac{1}{\sigma}\sqrt{\frac{2}{\pi}}e^{-\frac{y^2}{2\sigma^2}}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$

22. $f_W(t)=\begin{cases}\dfrac{1}{8}\left(\dfrac{2}{t}\right)^{\frac{1}{2}}, & 162<t<242,\\ 0, & \text{其他}.\end{cases}$

第 2 章　习题 B

1.

X	0	1	2	3	4	5
P	0.5838	0.3394	0.0702	0.0064	0.0002	0

$$F(x)=\begin{cases}0, & x<0,\\ 0.5838, & 0\leqslant x<1,\\ 0.9232, & 1\leqslant x<2,\\ 0.9934, & 2\leqslant x<3,\\ 0.9998, & 3\leqslant x<4,\\ 1, & x\geqslant 4.\end{cases}$$

2. （1）$P\{Y=k\}=C_{k-1}^{r-1}p^{r-1}q^{k-r}p$，$k=r$，$r+1$，$r+2$，…（$0<p<1$，$p+q=1$）；

（2）$P\{Y=k\}=pq^{k-1}+p^{k-1}q$，$k=2$，3，….

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

4. $p=\dfrac{2}{3}$.

5. 成立.

6. $A=\dfrac{\beta^{\alpha}}{\Gamma(\alpha)}$，当 $\alpha=1$ 时，该分布为指数分布.

7. $f_Y(y)=\begin{cases}\lambda m(y-\mu)^{m-1}\mathrm{e}^{-\lambda(y-\mu)^m}, & y>\mu,\\ 0, & y\leqslant\mu.\end{cases}$

8. X(每天的销售量)服从 $P(\lambda)$，$p=P\{X=5\}=\dfrac{\lambda^5\mathrm{e}^{-\lambda}}{5!}$；$Y$ 为销售量为 5 件的天数，Y 服从二项分布 $B(4, p)$，$P\{Y=1\}=C_4^1\dfrac{\lambda^5\mathrm{e}^{-\lambda}}{5!}\left(1-\dfrac{\lambda^5\mathrm{e}^{-\lambda}}{5!}\right)^3$.

9. $P\{Y=2\}=\dfrac{9}{64}$.

10. $f_Y(y)=\begin{cases}\dfrac{2}{\pi\sqrt{1-y^2}}, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$

11. $F(x)=\begin{cases} 0, & x<0, \\ \dfrac{2}{3}+\dfrac{x}{90}, & 0\leqslant x<30, \\ 1, & x\geqslant 30. \end{cases}$ X既不是连续型随机变量也不是离散型随机变量．

12. （1）$P\{Y=n\}=\mathrm{e}^{-n\lambda}(1-\mathrm{e}^{-\lambda})$，$n=0$，1，2，…；

（2）Y是离散型随机变量，服从几何分布．不一定．

13. $Y\sim P(p\lambda)$.

第3章　习题A

1. （1）

$Y \backslash X$	0	1
0	$\frac{9}{25}$	$\frac{6}{25}$
1	$\frac{6}{25}$	$\frac{4}{25}$

X	0	1
P	$\frac{3}{5}$	$\frac{2}{5}$

Y	0	1
P	$\frac{3}{5}$	$\frac{2}{5}$

X与Y独立．

（2）

$Y \backslash X$	0	1
0	$\frac{6}{20}$	$\frac{6}{20}$
1	$\frac{6}{20}$	$\frac{2}{20}$

X	0	1
P	$\frac{3}{5}$	$\frac{2}{5}$

Y	0	1
P	$\frac{3}{5}$	$\frac{2}{5}$

X与Y不独立．

2. （1）$\dfrac{1}{8}$；（2）$\dfrac{3}{8}$；（3）$\dfrac{27}{32}$；（4）$\dfrac{2}{3}$.

3. （1）$a=\dfrac{1}{\pi^2}$，$b=c=\dfrac{\pi}{2}$；

(2) $f(x, y)=\dfrac{6}{\pi^2(x^2+4)(y^2+9)}$;

(3) $f_X(x)=\dfrac{2}{\pi(x^2+4)}$, $f_Y(y)=\dfrac{3}{\pi(y^2+9)}$, X 与 Y 独立.

4. (1)

X \ Y	0	1	2
0	0	$\frac{2}{15}$	$\frac{3}{15}$
1	$\frac{1}{15}$	$\frac{6}{15}$	$\frac{3}{15}$

(2) $P\{Y=0 \mid X=0\}=0$, $P\{Y=1 \mid X=0\}=\dfrac{2}{5}$, $P\{Y=2 \mid X=0\}=\dfrac{3}{5}$;

(3) $P\{Y=0 \mid X=1\}=\dfrac{1}{10}$, $P\{Y=1 \mid X=1\}=\dfrac{6}{10}$, $P\{Y=2 \mid X=1\}=\dfrac{3}{10}$.

5. $a=\dfrac{2}{9}$, $b=\dfrac{1}{9}$.

6. (1) $a=\dfrac{14}{25}$, $b=\dfrac{3}{25}$;

(2) 当 $a=\dfrac{14}{25}$, $b=\dfrac{3}{25}$ 时,可求得 $P\{X=0\}=\dfrac{5}{25}$, $P\{Y=0\}=\dfrac{17}{25}$,易见

$$P\{X=0, Y=0\}=\frac{2}{25}\neq P\{X=0\}P\{Y=0\},$$

因此,X 与 Y 不独立.

7. (1) $F(x, y)=\begin{cases}(1-\mathrm{e}^{-x})(1-\mathrm{e}^{-\frac{1}{2}y}), & x>0, y>0,\\ 0, & \text{其他};\end{cases}$

(2) $f_X(x)=\begin{cases}\mathrm{e}^{-x}, & x>0,\\ 0, & \text{其他},\end{cases}$ $f_Y(y)=\begin{cases}\dfrac{1}{2}\mathrm{e}^{-\frac{1}{2}y}, & y>0,\\ 0, & \text{其他},\end{cases}$ X 与 Y 独立.

8. $\dfrac{3}{4}$.

9. (1) 6;

(2) $f_X(x)=\begin{cases}6(x-x^2), & 0<x<1,\\ 0, & \text{其他},\end{cases}$ $f_Y(y)=\begin{cases}6(\sqrt{y}-y), & 0<y<1,\\ 0, & \text{其他},\end{cases}$

X 与 Y 不独立;

(3) $\sqrt{2}-\dfrac{3}{4}$.

10. (1)

$Z=X+Y$	1	2	3	4
P	$\frac{1}{12}$	$\frac{1}{3}$	$\frac{5}{12}$	$\frac{1}{6}$

(2)

$Z=XY$	0	1	2	4
P	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{5}{12}$	$\frac{1}{6}$

(3)

$Z=\max\{X, Y\}$	1	2
P	$\frac{1}{4}$	$\frac{3}{4}$

(4)

$Z=\min\{X, Y\}$	0	1	2
P	$\frac{1}{4}$	$\frac{7}{12}$	$\frac{1}{6}$

11. $f_Z(z)=\dfrac{1}{\sqrt{10\pi}}e^{-\frac{z^2}{10}}$.

12. $f_Z(z)=\begin{cases} 1-e^{-z}, & 0\leqslant z\leqslant 1, \\ (e-1)e^{-z}, & z\geqslant 1, \\ 0, & \text{其他}. \end{cases}$

13. $f_Z(z)=\begin{cases} 6\lambda e^{-3\lambda z}(1-e^{-3\lambda z}), & z>0, \\ 0, & z\leqslant 0. \end{cases}$

第3章　习题B

1. (1)

$X \backslash Y$	-1	1
-1	$\frac{1}{6}$	$\frac{1}{3}$
1	$\frac{1}{3}$	$\frac{1}{6}$

(2) $\dfrac{5}{6}$.

2. (1) $P\{X=n\}=\dfrac{\lambda^n}{n!}e^{-\lambda}$, $n=0, 1, 2, \cdots$.

$P\{Y=m|X=n\}=C_n^m p^m(1-p)^{n-m}$, $0\leqslant m\leqslant n$, $n=0, 1, 2, \cdots$.

(2) $P\{X=n, Y=m\}=P\{X=n\}P\{Y=m|X=n\}$

$$=C_n^m p^m(1-p)^{n-m}\frac{e^{-\lambda}}{n!}\lambda^n,$$

$$0\leqslant m\leqslant n,\ n=0,1,2,\cdots;$$

(3) $P\{Y=m\}=\sum\limits_{n=m}^{\infty}P\{X=n, Y=m\}=\dfrac{(\lambda p)^m e^{-\lambda p}}{m!}$.

3. $f_Z(z)=\begin{cases} z^2, & 0<z<1, \\ 2z-z^2, & 1\leqslant z<2, \\ 0, & \text{其他}. \end{cases}$

4. $f_Z(z)=\begin{cases} e^{-z}+2e^{-2z}-3e^{-3z}, & z>0, \\ 0, & z\leqslant 0. \end{cases}$

5. (1) $f_Z(z)=\begin{cases} \dfrac{1}{2}e^{-\frac{1}{2}z}, & z>0, \\ 0, & z\leqslant 0; \end{cases}$ (2) e^{-2}.

6. $f_Z(z)=\begin{cases} \dfrac{1}{2}(2-z), & 0<z<2, \\ 0, & \text{其他}. \end{cases}$

7. $f_Z(z)=\dfrac{1}{4}[f(z)+f(z-1)+f(z-2)+f(z-3)]$.

第4章　习题A

1. $\lambda=4$.　2. (C).　3. (A).　4. $E(X+Y)=\dfrac{5}{6}$, $E(2X-Y^2)=\dfrac{7}{9}$.

5. $n=20$, $p=0.2$

6.

X	1	2	3
P	0.2	0.3	0.5

7. $D(Y)=\dfrac{8}{9}$.　8. $D(X-Y)=33$.　9. (B).

10. $E(X)=0.5$, $D(X)=0.95$, $E(-3X+1)=-0.5$, $D(-3X+1)=8.55$.

11. $A=\dfrac{1}{2}$, $E(X)=1$, $D(X)=\dfrac{1}{3}$, $E(e^X)=\dfrac{e^2-1}{2}$.

12. $C=\dfrac{1}{2}$, $E(X)=0$, $D(X)=2$.

13.

X	10	30	50	70	90
P	$\frac{2}{4}$	$\frac{1}{4}$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

$E(X)=30(\text{min})$，$D(X)=650(\text{min}^2)$.

14. $f_Z(z)=\dfrac{1}{5\sqrt{2\pi}}e^{-\frac{(z-4)^2}{50}}$，$-\infty<z<+\infty$.

15. (A).　16. (D).　17. (B).

18. $D(2X-3Y+5)=64$.

19. (X, Y)的联合分布律为

$X \backslash Y$	0	1
0	0.3	0.3
1	0.3	0.1

$\rho_{XY}=-\dfrac{1}{4}$.

20. $C=6$，$E(X)=\dfrac{1}{2}$，$E(Y)=\dfrac{3}{4}$，$\rho_{XY}=\dfrac{\sqrt{3}}{3}$.

第 4 章　习题 B

1. $E(XY)=\dfrac{16}{3}$.　2. 0.9.　3. (C).　4. (A).

5. (1) $E(X)=3$，$D(X)=2$；(2) $E(X)=5$，$D(X)=20$.

6. 平均内径 $\mu=9+\dfrac{1}{2}\ln\dfrac{11}{12}$时，销售一个零件所获利润的数学期望最大.

7. (X, Y)的联合分布律为

$X \backslash Y$	0	1
0	$\frac{8}{12}$	$\frac{1}{12}$
1	$\frac{2}{12}$	$\frac{1}{12}$

$\rho_{XY}=\dfrac{\sqrt{15}}{15}$.

8. 因为
$$E(X)=2P(A)-1,\ E(Y)=2P(B)-1,$$
$$E(XY)=4P(AB)-2P(A)-2P(B)+1,$$
于是
$$\text{Cov}(X, Y)=4P(AB)-4P(A)P(B),$$
所以 $\text{Cov}(X, Y)=0$ 当且仅当 $P(AB)=P(A)P(B)$，即 X 与 Y 不相关的充要

条件是 A 与 B 相互独立.

9. $C=\frac{1}{8}$，$E(X)=\frac{7}{6}$，$D(X+Y)=\frac{5}{9}$，$\rho_{XY}=-\frac{1}{11}$.

10. (U, V)的联合分布律为

$U \backslash V$	0	1
0	$\frac{1}{4}$	0
1	$\frac{1}{4}$	$\frac{1}{2}$

$\rho_{UV}=\frac{\sqrt{3}}{3}$.

11. 日最优进货量 37 件.

第 5 章　习题 A

1. $\frac{1}{3}$.　2. $\frac{7}{36}$.　3. $\Phi(x)$.　4. (B).　5. (D).　6. 0.9995.

7. (1) 0；(2) 0；(3) 0.5.

8. (1) $\overline{X}\sim N(2.2, 1.4^2/52)$，$P\{\overline{X}<2\}=0.1515$；(2) 0.077.

第 5 章　习题 B

1. (C).　2. (D).　3. (1) 0.1802；(2) 444.　4. $272a$.

5. 0.0013.

模 拟 试 卷 一

一、填空题

1. $\frac{3}{5}$.　2. 72%.　3. 100.　4. 0.44.　5. $\frac{-x^2+2x+1}{2}$.　6. $\frac{1}{3}e^x$.　7. $\frac{3}{5}$.

二、选择题

8. (C).　9. (C).　10. (D).　11. (B).　12. (A).　13. (D).　14. (B).

三、计算题

15. **解**　设事件 $A=\{$试验结果呈阳性反应$\}$，事件 $B=\{$被检查者确实患有癌症$\}$，则按题意有

$$P(B)=0.004，P(A|B)=0.95，P(A|\overline{B})=0.96，$$

由此可知：

$$P(\bar{B})=0.996,\ P(\bar{A}|B)=0.05,\ P(A|\bar{B})=0.04,$$

于是，由全概率公式得

$$P(A)=0.004\times0.95+0.996\times0.04=0.04364.$$

按贝叶斯公式得

$$P(B|A)=\frac{P(B)P(A|B)}{P(A)}=\frac{0.004\times0.95}{0.04364}=0.0871.$$

16. **解**　先求 a，利用 $\sum\limits_k p_k=1$，可得

$$4a+\frac{1}{12}+3a+a+10a+4a=1\Rightarrow a=\frac{1}{24},$$

所以

$$X\sim\begin{pmatrix}-2 & -1 & 0 & 1 & 2 & 3\\ \frac{1}{6} & \frac{1}{12} & \frac{1}{8} & \frac{1}{24} & \frac{5}{12} & \frac{1}{6}\end{pmatrix}.$$

$Y=X^2$ 的可能取值为 0，1，4，9，

$P\{Y=0\}=P\{X=0\}=\frac{1}{8}$，

$P\{Y=1\}=P\{X=-1\}+P\{X=1\}=\frac{1}{12}+\frac{1}{24}=\frac{1}{8}$，

$P\{Y=4\}=P\{X=-2\}+P\{X=2\}=\frac{1}{6}+\frac{5}{12}=\frac{7}{12}$，

$P\{Y=9\}=P\{X=3\}=\frac{1}{6}$，

所以

$$Y\sim\begin{pmatrix}0 & 1 & 4 & 9\\ \frac{1}{8} & \frac{1}{8} & \frac{7}{12} & \frac{1}{6}\end{pmatrix}.$$

17. **解**　设 $Y=X^2$ 的分布函数为 $F_Y(y)$，密度函数为 $f_Y(y)$.

当 $y<0$ 时，

$$F_Y(y)=P\{Y\leqslant y\}=P\{X^2\leqslant y\}=0;$$

当 $y>0$ 时，

$$\begin{aligned}F_Y(y)&=P\{Y\leqslant y\}=P\{X^2\leqslant y\}=P\{-\sqrt{y}\leqslant X\leqslant\sqrt{y}\}\\&=\int_{-\sqrt{y}}^{\sqrt{y}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx,\end{aligned}$$

所以 $Y=X^2$ 的密度函数为

$$f_Y(y)=F'_Y(y)=\begin{cases}\frac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}e^{-\frac{y}{2}}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

18. (1)、(2) (X, Y)的联合分布律和边缘分布律为

$Y \backslash X$	0	1	2	3	$P_{\cdot j}=P\{Y=j\}$
0	0	0	$\frac{3}{35}$	$\frac{2}{35}$	$\frac{1}{7}$
1	0	$\frac{6}{35}$	$\frac{12}{35}$	$\frac{2}{35}$	$\frac{4}{7}$
2	$\frac{1}{35}$	$\frac{6}{35}$	$\frac{3}{35}$	0	$\frac{2}{7}$
$P_{i\cdot}=P\{X=i\}$	$\frac{1}{35}$	$\frac{12}{35}$	$\frac{18}{35}$	$\frac{4}{35}$	

(3) $$E(XY)=\sum_{i=0}^{3}\sum_{j=0}^{2} i\times j\times p_{ij}$$
$$=1\times 1\times\frac{6}{35}+1\times 2\times\frac{6}{35}+2\times 1\times\frac{12}{35}+2\times 2\times\frac{3}{35}+3\times 1\times\frac{2}{35}=\frac{12}{7}.$$

19. **解**　设 Y 为该乘客的等待时间(单位：min)，则

$$Y=\begin{cases}5-X, & 0\leqslant X\leqslant 5,\\ 25-X, & 5<X\leqslant 25,\\ 55-X, & 25<X\leqslant 55,\\ 65-X, & 55<X\leqslant 60,\end{cases}$$

$$E(Y)=\frac{1}{60}\left[\int_0^5(5-x)\mathrm{d}x+\int_5^{25}(25-x)\mathrm{d}x+\int_{25}^{55}(55-x)\mathrm{d}x+\int_{55}^{60}(65-x)\mathrm{d}x\right]$$
$$=\frac{35}{3}.$$

四、证明题

20. **证**　由于 $1\geqslant P(B\cup C)=P(B)+P(C)-P(BC)$
$$=2a+3a-P(AB)=5a-P(AB),$$
$$b=P(AB)\leqslant P(A)=a,$$
所以
$$1\geqslant 5a-a=4a\Rightarrow a\leqslant\frac{1}{4},$$
又由上面的证明知 $b\leqslant a$，故 $b\leqslant\frac{1}{4}$.

模 拟 试 卷 二

一、填空题

1. $\overline{A}\cup\overline{B}\cup\overline{C}$.　　2. $\frac{C_{10}^4 C_4^3 C_3^2}{C_{17}^9}=\frac{252}{2431}$.　　3. $\frac{1}{6}$，$\frac{1}{3}$.　　4. $\frac{9}{16}$.

5. $\sum_{k=0}^{3}(C_3^k)^2 0.42^k\times 0.12^{3-k}=0.32076$. 　6. 0.2. 　7. $\frac{3^8}{8!}e^{-3}$.

8. 25. 　9. $a=\frac{2}{9}$，$b=\frac{1}{9}$. 　10. $\frac{\pi(b^2+ab+a^2)}{12}$.

二、选择题

11.(B). 12.(B). 13.(C). 14.(A). 15.(C).

三、计算题

16. **解** (1) 飞机坠毁这一事件可看成由以下几个互不相容事件的并组成：因一个人击中而坠毁，因二人击中而坠毁，因三人击中而坠毁，故可依此思路进行样本空间的分解．设 B=飞机坠毁，A_i=恰有 i 个人同时击中($i=0$，1，2，3)，则 $\Omega=A_0+A_1+A_2+A_3$，记 C_1=甲射中，C_2=乙射中，C_3=丙射中，则

$$A_0=\bar{C}_1\bar{C}_2\bar{C}_3，A_1=(C_1\bar{C}_2\bar{C}_3)\cup(\bar{C}_1C_2\bar{C}_3)\cup(\bar{C}_1\bar{C}_2C_3)，$$
$$A_2=(\bar{C}_1C_2C_3)\cup(C_1\bar{C}_2C_3)\cup(C_1C_2\bar{C}_3)，A_3=C_1C_2C_3，$$

由已知数据、概率可加性及乘法公式可计算出：

$P(A_0)=0.6\times0.5\times0.3=0.09$，

$P(A_1)=0.4\times0.5\times0.3+0.6\times0.5\times0.3+0.6\times0.5\times0.7=0.36$，

$P(A_2)=0.41$，$P(A_3)=0.14$，

由题意$P(B|A_0)=0$，$P(B|A_1)=0.2$，$P(B|A_2)=0.6$，$P(B|A_3)=1$，

所以
$$\begin{aligned}P(B)&=\sum_{n=0}^{3}P(A_i)P(B|A_i)\\&=0.09\times0+0.36\times0.2+0.41\times0.6+0.14\times1\\&=0.458.\end{aligned}$$

(2) $P(\bar{C}_1A_0)=P(A_0)=0.09$，

$P(\bar{C}_1A_1)=P\{(\bar{C}_1C_2\bar{C}_3)\cup(\bar{C}_1\bar{C}_2C_3)\}=0.3$，

$P(\bar{C}_1A_2)=P(\bar{C}_1C_2C_3)=0.21$，$P(\bar{C}_1A_3)=P(\varnothing)=0$，

$$\begin{aligned}P(\bar{C}_1B)&=\sum_{n=0}^{3}P(\bar{C}_1A_i)P(B|A_i\bar{C}_1)\\&=0.09\times0+0.3\times0.2+0.21\times0.6+0\times1=0.186，\end{aligned}$$

$P(\bar{C}_1|B)=\frac{P(\bar{C}_1B)}{P(B)}=\frac{93}{229}$.

17. **解** (1)
$$\begin{aligned}P\{X=k\}&=P\{X\leqslant k\}-P\{X<k\}=F(k)-\lim_{x\to k^-}F(x)\\&=(1-2^{-(k+1)})-(1-2^{-k})\\&=2^{-(k+1)}=\frac{1}{2^{k+1}}(k=0，1，2，\cdots)；\end{aligned}$$

(2) $E(2^{\frac{X}{2}}) = \sum_{k=0}^{\infty} 2^{\frac{k}{2}} \frac{1}{2^{k+1}} = \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^{\frac{k}{2}+1} = \frac{\frac{1}{2}}{1-\sqrt{\frac{1}{2}}} = \frac{1}{2-\sqrt{2}}$.

18. **解**　(1) 由 $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1$，即 $\int_{-1}^{1} \frac{A}{1+x^2}\mathrm{d}x = 1 \Rightarrow A\arctan x\big|_{-1}^{1} = 1$，得 $A=\frac{2}{\pi}$.

(2) $F_Y(y)=P\{Y\leqslant y\}=P\left\{\frac{1}{|X|}\leqslant y\right\}$，

当 $y\leqslant 0$ 时，$F_Y(y)=0$；

当 $0<\frac{1}{y}<1$，即 $y>1$ 时，

$$F_Y(y) = P\left\{|X| \geqslant \frac{1}{y}\right\} = 1 - P\left\{|X| < \frac{1}{y}\right\}$$
$$= 1 - \int_{-\frac{1}{y}}^{\frac{1}{y}} \frac{2}{\pi(1+x^2)}\mathrm{d}x = 1 - \frac{4}{\pi}\arctan\frac{1}{y};$$

当 $\frac{1}{y}\geqslant 1$，即 $0<y\leqslant 1$ 时，

$$F_Y(y) = P\left\{|X| \geqslant \frac{1}{y}\right\} = 1 - P\left\{|X| < \frac{1}{y}\right\}$$
$$= 1 - \int_{-1}^{1} \frac{2}{\pi(1+x^2)}\mathrm{d}x = 1 - 1 = 0.$$

所以随机变量 Y 的密度函数为

$$f_Y(y)=F_Y'(y)=\begin{cases}\frac{4}{\pi(1+y^2)}, & y>1,\\ 0, & y\leqslant 1.\end{cases}$$

19. **解**　(1)、(2) (X,Y)的联合分布律，X 和 Y 的边缘分布律为

$X \backslash Y$	0	1	2	$P_{i\cdot}=P\{X=i\}$
0	$\frac{2}{35}$	$\frac{4}{35}$	$\frac{1}{35}$	$\frac{7}{35}$
1	$\frac{32}{105}$	$\frac{8}{35}$	0	$\frac{56}{105}$
2	$\frac{28}{105}$	0	0	$\frac{28}{105}$
$P_{\cdot j}=P\{Y=j\}$	$\frac{22}{35}$	$\frac{12}{35}$	$\frac{1}{35}$	

(3) $P\{X+Y\leqslant 1\}=P\{X=0, Y=0\}+P\{X=0, Y=1\}+P\{X=1, Y=0\}$

$$=\frac{50}{105}=\frac{10}{21}.$$

20. **解** 设 Y 为工厂出售一台设备的净利润(单位：元)，则

$$Y=g(X)=\begin{cases}150, & X>1,\\ -200, & X\leqslant 1,\end{cases}$$

因此
$$\begin{aligned}EY&=E[g(X)]=150\times P\{X>1\}-200\times P\{X\leqslant 1\}\\&=150\times\int_1^{+\infty}\frac{1}{4}\mathrm{e}^{-\frac{x}{4}}\mathrm{d}x-200\times\int_0^1\frac{1}{4}\mathrm{e}^{-\frac{x}{4}}\mathrm{d}x\\&=350\mathrm{e}^{-\frac{1}{4}}-200=72.58.\end{aligned}$$

四、证明题

21. **证** 必要性 若 $X\sim G(p)$，则 $\forall k, l\in\mathbf{Z}^+$，

$$\begin{aligned}P\{X>k+l\mid X>k\}&=\frac{P\{X>k+l, X>k\}}{P\{X>k\}}=\frac{P\{X>k+l\}}{P\{X>k\}}\\&=\frac{\sum\limits_{i=k+l+1}^{\infty}q^{i-1}p}{\sum\limits_{i=k+1}^{\infty}q^{i-1}p}=\frac{q^{k+l}\sum\limits_{i=1}^{\infty}q^{i-1}p}{q^k\sum\limits_{i=1}^{\infty}q^{i-1}p}\\&=q^l\sum_{i=1}^{\infty}q^{i-1}p=\sum_{i=l+1}^{\infty}q^{i-1}p=P\{X>l\}.\end{aligned}$$

充分性 若 $\forall k, l\in\mathbf{Z}^+$，$P\{X>k+l\mid X>k\}=P\{X>l\}$，令 $l=1$，则有

$$\forall k\in\mathbf{Z}^+,\ P\{X>k+1\mid X>k\}=\frac{P\{X>k+1\}}{P\{X>k\}}=P\{X>1\}\xlongequal{\text{令}}q,$$

则得 $P\{X>k\}=qP\{X>k-1\}=q^2P\{X>k-2\}=\cdots=q^k$，$\forall k\in\mathbf{Z}^+$，

所以
$$\begin{aligned}P\{X=k\}&=P\{X>k-1\}-P\{X>k\}=q^{k-1}-q^k=q^{k-1}(1-q)\\&=q^{k-1}p\,(p=1-q),\ k=1,2,\cdots,\end{aligned}$$

即 X 服从几何分布.

模 拟 试 卷 三

一、填空题

1. $\dfrac{3!\,C_{12}^4C_8^4}{C_{15}^5C_{10}^5}$.　　2. $\dfrac{1}{3}$.　　3. 100.　　4. $\dfrac{1}{3}$.　　5. $1-\dfrac{1}{2}\mathrm{e}^{-x}$.

6. $f(x)=\begin{cases}\dfrac{1}{\pi\sqrt{a^2-x^2}}, & |x|<a,\\ 0, & \text{其他}.\end{cases}$　　7. $\dfrac{1}{2}$.

二、选择题

8.（A）.　9.（D）.　10.（C）.　11.（A）.　12.（D）.　13.（A）.　14.（C）.

三、计算题

15. **解**　设 B＝麦种发芽，A_1，A_2，A_3 分别表示麦种为一等、二等、三等麦种，则 $\Omega=A_1+A_2+A_3$.

（1）由已知数据及概率可加性及乘法公式可计算出：

$$P(A_1)=0.80,\ P(A_2)=0.18,\ P(A_3)=0.02,$$

由题意 $$P(B|A_1)=0.8,\ P(B|A_2)=0.5,\ P(B|A_3)=0.2,$$

所以 $$\begin{aligned}P(B) &= \sum_{n=0}^{3} P(A_i)P(B|A_i) \\ &= 0.80\times 0.8+0.18\times 0.5+0.02\times 0.2=0.734.\end{aligned}$$

（2）$$\begin{aligned}P(A_1|\bar{B}) &= \frac{P(A_1\bar{B})}{P(\bar{B})}=\frac{P(A_1)P(\bar{B}|A_1)}{1-P(B)} \\ &= \frac{0.8\times 0.2}{1-0.734}=0.6015\left(\text{或}=\frac{80}{133}\right).\end{aligned}$$

16. **解**　$Y=\sin\left(\dfrac{\pi}{2}X\right)$的可能取值为$-1$，0，1，

$$\begin{aligned}P\{Y=-1\} &= P\left\{\sin\left(\frac{\pi}{2}X\right)=-1\right\}=P\left(\bigcup_{k=0}^{\infty}\{X=4k+3\}\right) \\ &= \sum_{k=0}^{\infty}P\{X=4k+3\}=\sum_{k=0}^{\infty}\frac{1}{2^{4k+3}}=\frac{2}{15},\end{aligned}$$

$$\begin{aligned}P\{Y=0\} &= P\left\{\sin\left(\frac{\pi}{2}X\right)=0\right\}=P\left(\bigcup_{k=1}^{\infty}\{X=2k\}\right) \\ &= \sum_{k=1}^{\infty}P\{X=2k\}=\sum_{k=1}^{\infty}\frac{1}{2^{2k}}=\frac{1}{3},\end{aligned}$$

$$\begin{aligned}P\{Y=1\} &= P\left\{\sin\left(\frac{\pi}{2}X\right)=1\right\}=P\left(\bigcup_{k=0}^{\infty}\{X=4k+1\}\right) \\ &= \sum_{k=0}^{\infty}P\{X=4k+1\}=\sum_{k=0}^{\infty}\frac{1}{2^{4k+1}}=\frac{8}{15},\end{aligned}$$

所以 $$Y\sim\begin{pmatrix}-1 & 0 & 1 \\ \dfrac{2}{15} & \dfrac{1}{3} & \dfrac{8}{15}\end{pmatrix}.$$

17. **解**　（1）由密度函数的性质有

$$1=\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=A\int_{0}^{+\infty}\frac{1}{1+x^2}\mathrm{d}x=A\cdot\arctan x\Big|_0^{+\infty}\Rightarrow A=\frac{2}{\pi}.$$

（2）设 $Y=\dfrac{1}{X}$的分布函数为 $F_Y(y)$，密度函数为 $f_Y(y)$.

当 $y<0$ 时，

$$F_Y(y)=P\{Y\leqslant y\}=P\left\{\frac{1}{X}\leqslant y\right\}=0;$$

当 $y>0$ 时，

$$F_Y(y)=P\{Y\leqslant y\}=P\left\{\frac{1}{X}\leqslant y\right\}=1-\int_0^{\frac{1}{y}}\frac{2}{\pi}\frac{1}{1+x^2}\mathrm{d}x,$$

所以 $Y=\dfrac{1}{X}$的密度函数为

$$f_Y(y)=F_Y'(y)=\begin{cases}-\dfrac{2}{\pi}\dfrac{1}{1+(1/y)^2}\cdot\left(-\dfrac{1}{y^2}\right)=\dfrac{2}{\pi(1+y^2)}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

18. **解** $P\{X=0,\ Y=0\}=P\{X=0\}P\{Y=0\mid X=0\}=\dfrac{1}{2^2}\times\dfrac{1}{2}=\dfrac{1}{8}$,

$P\{X=0,\ Y=1\}=P\{X=0\}P\{Y=1\mid X=0\}=\dfrac{1}{2^2}\times\dfrac{1}{2}=\dfrac{1}{8}$,

$P\{X=0,\ Y=j\}=P\{X=0\}P\{Y=j\mid X=0\}=\dfrac{1}{2^2}\times 0=0$，$j=2$，3，

$P\{X=1,\ Y=1\}=P\{X=1\}P\{Y=1\mid X=1\}=C_2^1\dfrac{1}{2^2}\times\dfrac{1}{2}=\dfrac{1}{4}$,

$P\{X=1,\ Y=2\}=P\{X=1\}P\{Y=2\mid X=1\}=C_2^1\dfrac{1}{2^2}\times\dfrac{1}{2}=\dfrac{1}{4}$,

$P\{X=1,\ Y=j\}=P\{X=1\}P\{Y=j\mid X=1\}=\dfrac{1}{2^2}\times 0=0$，$j=0$，3，

$P\{X=2,\ Y=2\}=P\{X=2\}P\{Y=2\mid X=2\}=\dfrac{1}{2^2}\times\dfrac{1}{2}=\dfrac{1}{8}$,

$P\{X=2,\ Y=3\}=P\{X=2\}P\{Y=3\mid X=2\}=\dfrac{1}{2^2}\times\dfrac{1}{2}=\dfrac{1}{8}$,

$P\{X=2,\ Y=j\}=P\{X=2\}P\{Y=j\mid X=2\}=\dfrac{1}{2^2}\times 0=0$，$j=0$，1，

所以，$(X，Y)$的联合分布律和边缘分布律为

X \ Y	0	1	2	3	$P_{i\cdot}=P\{X=i\}$
0	$\frac{1}{8}$	$\frac{1}{8}$	0	0	$\frac{1}{4}$
1	0	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{2}$
2	0	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$
$P_{\cdot j}=P\{Y=j\}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	

19. **解**　用 X_1 和 X_2 分别表示甲、乙命中的环数，则甲、乙平均每枪命中的环数为

$$E(X_1)=5\times0.05+6\times0.05+7\times0.2+8\times0.1+9\times0.2+10\times0.4=8.55(\text{环}),$$

$$E(X_2)=5\times0.1+6\times0.1+7\times0.1+8\times0.2+9\times0.1+10\times0.4=8.3(\text{环}),$$

由于 $E(X_1)>E(X_2)$，所以甲的射击技术优于乙．

四、证明题

20. **证明**

$$\begin{aligned}
&P(B|A)+P(\overline{B}|\overline{A})=1\\
\Leftrightarrow\ &P(B|A)=1-P(\overline{B}|\overline{A})\\
\Leftrightarrow\ &P(B|A)=P(B|\overline{A})\\
\Leftrightarrow\ &\frac{P(AB)}{P(A)}=\frac{P(\overline{A}B)}{P(\overline{A})}=\frac{P(B)-P(AB)}{1-P(A)}\\
\Leftrightarrow\ &P(AB)-P(A)P(AB)=P(A)P(B)-P(A)P(AB)\\
\Leftrightarrow\ &P(AB)=P(A)P(B)\\
\Leftrightarrow\ &\text{事件 } A \text{ 与 } B \text{ 相互独立}.
\end{aligned}$$

参 考 文 献

陈萍，等．2006. 概率与统计．第 2 版．北京：科学出版社．
陈薇．1998. 概率论与数理统计．北京：中国农业大学出版社．
杜忠复，崔文善，雷鸣．2009. 概率论与数理统计．北京：中国农业大学出版社．
冯予，陈萍．2005. 概率论与数理统计．北京：国防工业出版社．
何书元．2006. 概率论与数理统计．北京：高等教育出版社．
茆诗松，程依明，濮晓龙．2011. 概率论与数理统计教程．第 2 版．北京：高等教育出版社．
蒲俊，吉家锋，伊良忠．2002. Matlab 6.0 数学手册．上海：浦东电子出版社．
盛骤，谢式千，潘承毅．2008. 概率论与数理统计．第 4 版．北京：高等教育出版社．
吴坚．2002. 应用概率统计．北京：高等教育出版社．
杨振明．2008. 概率论．第 2 版．北京：科学出版社．
叶尔骅，张德平．2005. 概率论与随机过程．北京：科学出版社．