公共课系列

2001年研究生入学考试应试指导丛书

8

2001

策划：北京大学研究生院

2001

Entrance Exams for MD

研究生入学考试

# 概率论与数理统计

姚孟臣 编著

北京大学出版社

# 内容简介

本书是经济类硕士研究生入学考试科目“概率论与数理统计”复习指导书.本书作者多年来参加有关经济类考研数学试卷的命题、阅卷和考研辅导班的教学，亲临考研第一线，具有丰富的教学与应试经验.作者把他们的教学经验以及潜心研究试卷命题的体会加以细化、归纳和总结，整理成书奉献给广大读者，旨在提高考研者的考试成绩与命中率.

本书紧扣数学三、四的考试大纲，贴切考试实践，内容丰富.全书共分八章.内容包括：随机事件和概率，随机变量及其分布，多维随机变量，随机变量的数字特征，大数定律和中心极限定理，数理统计的基本概念，参数估计及假设检验等.本书结构新颖，每一章按照：考试要求，复习要点(重要定义、定理及公式)，典型例题分析，练习题四部分编写.本书概念叙述简洁，解题思路清晰，对典型例题从多侧面、不同角度、用多种解法进行讲解，注重对考生基本概念的理解、多种类型基础题目的训练和综合解题能力的培养，是考研者较好的复习指导书和良师益友.

本书可作为经济类硕士研究生入学考试数学考试科目“概率论与数理统计”的复习指导书，也可作为理工科考研者的复习参考书.对于在校的经济类及管理类大学生、大专生及自学考试者，本书也是一本较好的学习用书.

2001 年研究生入学考试应试指导丛书

# 概率论与数理统计

姚孟臣　编著

北京大学出版社
·北　　京·

**图书在版编目(CIP)数据**

概率论与数理统计/姚孟臣编著. 一北京:北京大学出版社,2000.4

(2001年研究生入学考试应试指导丛书)

ISBN 7-301-04480-1

Ⅰ.概… Ⅱ.姚… Ⅲ.①概率论-研究生-入学考试-自学参考资料 ②数理统计-研究生-入学考试-自学参考资料 Ⅳ.021

中国版本图书馆CIP数据核字(2000)第03826号

**书　　名:概率论与数理统计**

著作责任者:姚孟臣　编著

责 任 编 辑:刘　勇

标 准 书 号:ISBN 7-301-04480-1/G·564

出　版　者:北京大学出版社

地　　址:北京市海淀区中关村北京大学校内　100871

网　　址:http://cbs.pku.edu.cn/cbs.htm

电　　话:出版部62752015　发行部62754140　理科编辑部62752021

电 子 信 箱:zpup@pup.pku.edu.cn

排　版　者:北京高新特公司照排中心

印　刷　者:北京飞达印刷厂

发　行　者:北京大学出版社

经　销　者:新华书店

850×1168　32开本　10.125印张　255千字

2000年4月第1版　2000年4月第1次印刷

定　　价:14.50元

# 前　言

为了帮助有志攻读硕士研究生的广大考生能在较短的时间内全面、系统地复习有关的数学内容，我们根据教育部最近制定的“全国硕士研究生入学统一考试数学考试大纲”的有关要求，结合我们多年参加经济类数学考研及有关考试的命题、阅卷及辅导的经验，编写了这套经济类数学《2001 年研究生入学考试应试指导丛书》，其中包括复习指导书：《微积分》、《线性代数》、《概率论与数理统计》以及《2001 年研究生入学考试数学模拟试卷（经济学类）》共 4 册，供广大考生选用。

本套应试指导丛书的每一章由以下四部分构成：

**一、考试要求**——编写这部分的目的是使广大考生明确每一章考的内容是什么，掌握到什么程度就可以了。在编写过程中，根据我们多年以来参加有关命题的经验把考试大纲所要求的内容加以细化、归纳和总结，使广大考生能够正确把握考试要求，这是区别于其他考研辅导书的一大特点。

**二、重要定义、定理及公式**——这部分根据考试大纲的要求，将概念、定理和公式、方法进行了简明扼要地叙述、归纳和总结；精选了各种典型的例题并作详细的解答，使得考生能够在较短的时间内对重点、难点、热点问题进行复习，全面、系统地掌握所需要的知识，在考试时能够拿得出、用得上。

**三、典型例题分析**——这部分根据考试大纲要求的题型进行了分类、归纳，总结了各种题型解题的方法及技巧，开阔了考生的解题思路，使所学的知识能够融会贯通，并迅速提高考生的综合解题能力。

**四、练习题**——每章的最后部分精选了适量的练习题，全部

习题都附有答案或提示，这些题目作为考生巩固所学知识、复习有关内容时使用，这有利于提高考生分析问题和解决问题的能力。

《2001年研究生入学考试数学模拟试卷(经济学类)》一书由两部分组成：一部分是为了检查前一阶段考生复习效果、更好地提高考生的应试能力而设计编制的全真模拟试题共十二份，其中数学三、四各6份；另一部分是1999年和2000年数学三、四考研试题及解答。

我们是在深入研究了历年考研试卷的结构、知识点及难度的分布，并紧密结合我们命题实践、阅卷过程中常见问题及在全国各大城市"考研辅导班"辅导的经验来编好每一道题。因此，每一份试卷都从不同角度选择了具有多种风格的题目，基本上涵盖了全部命题思路，能够达到实际考试效果。这样有利于广大考生检验自己复习的效果，更加全面系统地掌握所需知识，迅速提高综合解题能力。为了方便考生了解近两年有关数学考试的题目类型、知识点及难度的分布情况，我们给出了1999年和2000年数学三、四考研试题及解答，仅供参考。

本书不仅是经济类硕士研究生入学考试数学考试科目应试者的一本复习用书，同时对于参加理工类考研及在校的经济类和管理类本科生、大专生、参加自学考试考生也是一本较好的学习用书和参考书。

由于编者水平有限，加之时间比较仓促，书中难免有错误和疏漏之处，恳请读者批评指正。

编　者

2000年4月20日

# 目　录

# 第一章 随机事件和概率

## 一、考试要求

1. 了解样本空间的概念，理解随机事件的概念，重点掌握事件的关系与运算；

2. 理解概率与条件概率的概念，掌握概率的基本性质，会运用古典概型计算有关的概率；

3. 掌握概率的加法公式、乘法公式、全概公式及贝叶斯公式；

4. 理解事件独立性的概念，掌握利用事件独立性计算有关概率的各种方法；理解独立重复试验的概念，掌握运用二项概型计算有关的概率的各种方法.

## 二、复习要点

### （一）重要概念及性质

**定义 1.1**（随机现象） 在一定条件下，具有多种可能发生的结果的现象称为**随机现象**. 这类现象的一个共同点是：事先不能预言多种可能结果中究竟出现哪一种.

**定义 1.2**（随机试验） 为了叙述方便，我们把对随机现象进行的一次观测或一次实验统称为它的一个试验. 如果这个试验满足下面的两个条件：

（1）在相同的条件下可以重复进行；

（2）试验都有哪些可能的结果是明确的，但每次试验的具体结果在试验前是无法得知的，

那么就称它是一个**随机试验**，简称为**试验**，一般用字母 $E$ 表示.

**定义 1.3**（样本空间） 在随机试验中，每一个可能出现的不

可分解的最简单的结果称为随机试验的**基本事件**或**样本点**,用 $\omega$ 表示;而由全体基本事件构成的集合称为**基本事件空间**或**样本空间**,记为 $\Omega$.

**例 1** 设 $E_1$ 为从 10 件产品(其中 2 件次品,8 件正品)之中任取 3 件,观察其中次品的件数. 记 $\omega_i$ 为恰有 $i$ 件次品($i=0,1,2$),于是 $\Omega=\{\omega_0,\omega_1,\omega_2\}$.

**例 2** 设 $E_2$ 为在相同条件下接连不断地向一个目标射击,直到击中目标为止,观察射击次数. 记 $\omega_i$ 为射击 $i$ 次($i=1,2,\cdots$),于是 $\Omega=\{\omega_1,\omega_2,\cdots\}$.

**例 3** 设 $E_3$ 为某地铁站每隔 5 分钟有一列车通过,乘客对于列车通过该站的时间完全不知道,观察乘客候车的时间. 记乘客的候车时间为 $\omega$. 显然有 $\omega\in[0,5)$,即 $\Omega=[0,5)$.

通过上面的几个例子可以看出,随机试验大体可以分成只有有限个可能结果的(如 $E_1$);有可列个可能结果的(如 $E_2$)和有不可列个可能结果的(如 $E_3$)这样三种情况.

应该说明的是,一个随机试验中样本点个数的确定都是相对试验目的而言的. 另外,一个随机试验的条件有的是人为的,有的是客观存在的. 在后一种情况下,每当试验条件实现时,人们便会观测到一个结果 $\omega$. 虽然我们无法事先准确地说出试验的结果,但是能够指出它出现的范围 $\Omega$. 因此,我们所讨论的随机试验是有着十分广泛的含意的.

**例 4** 写出下列随机试验的样本空间 $\Omega$:

(1) 同时掷两枚骰子,记录两枚骰子点数之和;

(2) 10 件产品中有 3 件是次品,每次从中取 1 件,取出后不再放回,直到 3 件次品全部取出为止,记录抽取的次数;

(3) 生产某种产品直到得到 10 件正品,记录生产产品的总件数;

(4) 将一尺之棰折成三段,观察各段的长度.

**解** (1) $\Omega=\{2,3,\cdots,12\}$;

(2) $\Omega=\{3,4,\cdots,10\}$;

(3) $\Omega=\{10,11,\cdots\}$;

(4) 设 $x,y,z$ 分别表示第一段、第二段、第三段的长度,有

$\Omega=\{(x,y,z)\mid x>0,y>0,z>0,x+y+z=1\}$.

**定义 1.4**(随机事件) 所谓**随机事件**是样本空间 $\Omega$ 的一个子集,随机事件简称为**事件**,用字母 $A,B,C$ 等表示. 因此,某个事件 $A$ 发生当且仅当这个子集中的一个样本点 $\omega$ 发生,记为 $\omega\in A$.

在每次试验中必定要发生的事件称为**必然事件**,记作 $\Omega$. 在每次试验中必定不会发生的事件称为**不可能事件**,记为 $\varnothing$. 我们知道,必然事件 $\Omega$ 与不可能事件 $\varnothing$ 都不是随机事件. 因为作为试验的结果,它们都是确定性的,并不具有随机性. 但是为了今后讨论问题方便,我们也将它们当作随机事件来处理.

**定义 1.5**(事件的关系)

(1) 包含

设 $A,B$ 为两个事件. 如果 $A$ 中的每一个样本点都属于 $B$,那么称事件 $B$ **包含**事件 $A$,或称事件 $A$ 包含于事件 $B$,记为 $A\subset B$ 或 $B\supset A$. 这就是说,在一次试验中,如果事件 $A$ 发生必然导致事件 $B$ 发生.

(2) 等价

如果 $A\supset B$ 与 $B\supset A$ 同时成立,那么称事件 $A$ 与事件 $B$ **等价**或**相等**,记为 $A=B$. 这就是说,在一次试验中,等价的两个事件同时发生或同时不发生,因此可以把它们看成是一样的.

(3) 互斥(互不相容)

设 $A,B$ 为两个事件. 如果 $A\cdot B$[①] $=\varnothing$,那么称事件 $A$ 与 $B$ 是**互不相容**的(或**互斥**的). 这就是说,在一次试验中事件 $A$ 与事

---

① 这里出现的 $A\cdot B$(及后面出现 $AB$)表示事件 $A$ 与 $B$ 的交,在第 5 页才给出了它的定义. 本书为了便于考生复习,集中分类归纳叙述了一些概念或术语,这就可能导致某些概念或术语提前引出,而它们的定义是滞后给出的. 请读者在阅读本书时能正确理解作者的用意.

件 $B$ 不可能同时发生.

事件的互不相容关系也可以推广到多于两个事件的情形. 即，如果 $A_1 \cdot A_2 \cdot \cdots \cdot A_n=\varnothing$，则称事件 $A_1,A_2,\cdots,A_n$ 是互斥的. 如果 $A_i \cdot A_j=\varnothing(i\neq j;i,j=1,2,\cdots,n)$，这时我们又称 $A_1,A_2,\cdots,A_n$ 是**两两互斥**的. 注意，如果 $n$ 个事件两两互斥，那么这 $n$ 个事件之间一定互斥；反之不真.

(4) 独立

设 $A,B$ 是某一随机试验的任意两个随机事件. 称 $A$ 与 $B$ 是**相互独立**的，如果

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

这就是在 $A$ 与 $B$ 独立的情况下事件 $A$ 与 $B$ 乘积的概率公式. 可见事件 $A$ 与 $B$ 相互独立是建立在概率基础上事件之间的一种关系. 所谓事件 $A$ 与 $B$ 相互独立就是指其中一个事件发生与否不影响另一个事件发生的可能性. 类似地当 $P(B)\neq 0$ 时，$A$ 与 $B$ 相互独立也可以用

$$P(A|B)=P(A)$$

来定义.

由两个随机事件相互独立的定义，我们可以得到：若事件 $A$ 与 $B$ 相互独立，则 $\overline{A}$ 与 $B$，$A$ 与 $\overline{B}$，$\overline{A}$ 与 $\overline{B}$ 也相互独立.

如果事件 $A,B,C$ 满足

$$\begin{cases}P(AB)=P(A)P(B),\\P(BC)=P(B)P(C),\\P(AC)=P(A)P(C),\\P(ABC)=P(A)P(B)P(C),\end{cases}$$

则称事件 $A,B,C$ 相互独立.

注意，事件 $A,B,C$ 相互独立与事件 $A,B,C$ 两两独立不同，两两独立是指上述四个式子中前三个式子成立. 因此，相互独立一定是两两独立，但反之不一定.

对于 $n$ 个事件的独立性，我们有

设 $A_1,A_2,\cdots,A_n$ 为 $n$ 个事件. 如果对于所有可能的组合 $1\leqslant i<j<k<\cdots\leqslant n$ 下列各式同时成立：

$$\begin{cases}P(A_iA_j)=P(A_i)P(A_j),\\P(A_iA_jA_k)=P(A_i)P(A_j)P(A_k),\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\P(A_1A_2\cdots A_n)=P(A_1)P(A_2)\cdots P(A_n),\end{cases}$$

那么称 $A_1,A_2,\cdots,A_n$ 是相互独立的.

**定义 1.6**（事件的运算）

（1）并

设 $A,B$ 为两个事件. 我们把至少属于 $A$ 或 $B$ 中一个的所有样本点构成的集合称为事件 $A$ 与 $B$ 的**并**或**和**，记为 $A\cup B$ 或 $A+B$. 这就是说，事件 $A\cup B$ 表示在一次试验中，事件 $A$ 与 $B$ 至少有一个发生.

我们用 $A_1+A_2+\cdots+A_n\xlongequal{\text{def}}\bigcup\limits_{i=1}^{n}A_i$ 表示事件 $A_1,A_2,\cdots,A_n$ 中至少有一个发生；进而用 $A_1+A_2+\cdots+A_n+\cdots\xlongequal{\text{def}}\bigcup\limits_{i=1}^{\infty}A_i$ 表示 $A_1,A_2,\cdots$ 中至少有一个发生.

（2）交

设 $A,B$ 为两个事件. 我们把同时属于 $A$ 及 $B$ 的所有样本点构成的集合称为事件 $A$ 与 $B$ 的**交**或**积**，记为 $A\cap B$ 或 $A\cdot B$，有时也简记为 $AB$. 这就是说，事件 $A\cap B$ 表示在一次试验中，事件 $A$ 与 $B$ 同时发生.

我们用 $A_1\cdot A_2\cdot\cdots\cdot A_n\xlongequal{\text{def}}\bigcap\limits_{i=1}^{n}A_i$ 表示事件 $A_1,A_2,\cdots,A_n$ 同时发生；进而用 $A_1\cdot A_2\cdot\cdots\cdot A_n\cdots\xlongequal{\text{def}}\bigcap\limits_{i=1}^{\infty}A_i$ 表示 $A_1,A_2,\cdots$ 同时发生.

（3）逆（对立）

对于事件 $A$，我们把不包含在 $A$ 中的所有样本点构成的集合

称为事件 $A$ 的**逆**(或 $A$ 的**对立事件**),记为 $\overline{A}$. 这就是说,事件 $\overline{A}$ 表示在一次试验中事件 $A$ 不发生. 我们规定它是事件的基本运算之一.

在一次试验中,事件 $A$ 与 $\overline{A}$ 不会同时发生(即 $A\cdot\overline{A}=\varnothing$,称它们具有**互斥性**),而且 $A$ 与 $\overline{A}$ 至少有一个发生(即 $A+\overline{A}=\Omega$,称它们具有**完全性**). 这就是说,事件 $A$ 与 $\overline{A}$ 满足:

$$\begin{cases} A\cdot\overline{A}=\varnothing, \\ A+\overline{A}=\Omega. \end{cases}$$

根据上面的基本运算定义,不难验证事件之间的运算满足以下的一些规律:

1) $A+B=B+A$ (加法交换律);

2) $A+(B+C)=(A+B)+C$ (加法结合律);

3) $A+A=A$;

4) $A+\overline{A}=\Omega$;

5) $A+\Omega=\Omega$;

6) $A+\varnothing=A$;

7) $A\cdot B=B\cdot A$ (乘法交换律);

8) $(AB)C=A(BC)$ (乘法结合律);

9) $A\cdot A=A$;

10) $A\cdot\overline{A}=\varnothing$;

11) $A\cdot\Omega=A$;

12) $A\cdot\varnothing=\varnothing$;

13) $A(B+C)=AB+AC$ (分配律);

14) $A+BC=(A+B)(A+C)$ (分配律);

15) $\overline{A+B}=\overline{A}\cdot\overline{B}$ (对偶原理(1));

16) $\overline{A\cdot B}=\overline{A}+\overline{B}$ (对偶原理(2)).

有了事件的三种基本运算我们就可以定义事件的其他一些运算. 例如,我们称事件 $A\overline{B}$ 为事件 $A$ 与 $B$ 的**差**,记为 $A-B$. 可见,事件 $A-B$ 是由包含于 $A$ 而不包含于 $B$ 的所有样本点构成的集

合.

**例 5** 设 $A,B,C$ 是三个随机事件. 试用 $A,B,C$ 表示下列各事件：

(1) 恰有 $A$ 发生； (2) $A$ 和 $B$ 都发生而 $C$ 不发生；

(3) 所有这三个事件都发生；(4) $A,B,C$ 至少有一个发生；

(5) 至少有两个事件发生； (6) 恰有一个事件发生；

(7) 恰有两个事件发生； (8) 不多于一个事件发生；

(9) 不多于两个事件发生； (10) 三个事件都不发生.

**解** (1) $A\overline{B}\,\overline{C}$；(2) $AB\overline{C}$；(3) $ABC$；(4) $A+B+C$；

(5) $AB+BC+CA$；(6) $A\overline{B}\,\overline{C}+\overline{A}B\overline{C}+\overline{A}\,\overline{B}C$；

(7) $AB\overline{C}+A\overline{B}C+\overline{A}BC$；

(8) $\overline{AB+BC+CA}$；(9) $\overline{ABC}$；(10) $\overline{A}\,\overline{B}\,\overline{C}$.

**例 6** 设某工人连续生产了四个零件，$A_i$ 表示他生产的第 $i$ 个零件是正品($i=1,2,3,4$)，试用 $A_i$ 表示下列各事件：

(1) 没有一个是次品； (2) 至少有一个是次品；

(3) 只有一个是次品； (4) 至少有三个不是次品；

(5) 恰好有三个是次品； (6) 至多有一个是次品.

**解** (1) $A_1A_2A_3A_4$；(2) $\overline{A_1A_2A_3A_4}$；

(3) $\overline{A_1}A_2A_3A_4+A_1\overline{A_2}A_3A_4+A_1A_2\overline{A_3}A_4+A_1A_2A_3\overline{A_4}$；

(4) $A_1A_2A_3\overline{A_4}+A_1A_2\overline{A_3}A_4+A_1\overline{A_2}A_3A_4+\overline{A_1}A_2A_3A_4+A_1A_2A_3A_4$；

(5) $A_1\overline{A_2}\,\overline{A_3}\,\overline{A_4}+\overline{A_1}A_2\overline{A_3}\,\overline{A_4}+\overline{A_1}\,\overline{A_2}A_3\overline{A_4}+\overline{A_1}\,\overline{A_2}\,\overline{A_3}A_4$；

(6) $A_1A_2A_3A_4+\overline{A_1}A_2A_3A_4+A_1\overline{A_2}A_3A_4+A_1A_2\overline{A_3}A_4+A_1A_2A_3\overline{A_4}$.

**例 7** 下列各式说明 $A$ 与 $B$ 之间具有何种包含关系？

(1) $AB=A$； (2) $A+B=A$.

**解** (1) 因为"$AB=A$"与"$AB\subset A$ 且 $A\subset AB$"是等价的，由 $A\subset AB$ 可以推出 $A\subset A$ 且 $A\subset B$，因此有 $A\subset B$.

(2) 因为“$A+B=A$”与“$A+B\subset A$ 且 $A\subset A+B$”是等价的，由 $A+B\subset A$ 可以推出 $A\subset A$ 且 $B\subset A$，因此有 $B\subset A$.

**定义 1.7(概率的公理化定义)** 设 $E$ 是一个随机试验，$\Omega$ 为它的样本空间，以 $E$ 中所有的随机事件组成的集合为定义域，定义一个函数 $P(A)$(其中 $A$ 为任一随机事件)，且 $P(A)$满足以下三条公理，则称函数 $P(A)$为事件 $A$ 的**概率**.

**公理 1**(非负性) $0\leqslant P(A)\leqslant 1$.

**公理 2**(规范性) $P(\Omega)=1$.

**公理 3**(可列可加性) 若 $A_1,A_2,\cdots,A_n,\cdots$两两互斥，则

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

利用概率的公理化定义，可导出概率的下列一些性质：

1) $P(\varnothing)=0$；

2) 若 $A_1,A_2,\cdots,A_n$ 为 $n$ 个互不相容事件，则

$$P(A_1\cup A_2\cup\cdots\cup A_n)=\sum_{i=1}^{n}P(A_i),$$

此称为概率的有限可加性；

3) 对任意事件 $A$，$P(\overline{A})=1-P(A)$，$P(A)=1-P(\overline{A})$；

4) 对事件 $A,B$，若 $A\supset B$，则

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

5) 对任意两个事件 $A,B$，有

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

该公式称为**加法公式**. 利用归纳法可推出 $n$ 个事件的加法公式：

对任意 $n$ 个事件 $A_1,A_2,\cdots,A_n$，有

$$P(A_1\cup A_2\cup\cdots\cup A_n)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)$$
$$+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)-\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n).$$

这些性质将帮助我们解决概率计算问题.

**定义 1.8(概率的统计定义)** 在一组不变的条件 $S$ 下，独立

地重复作 $n$ 次试验. 设 $\mu$ 是 $n$ 次试验中事件 $A$ 发生的次数，当试验次数 $n$ 很大时，如果 $A$ 的频率 $f_n(A)=\frac{\mu}{n}$ 稳定地在某一数值 $p$ 附近摆动；而且一般说来随着试验次数的增多，这种摆动的幅度会越来越小，则称数值 $p$ 为事件 $A$ 在条件组 $S$ 下发生的**概率**，记作

$$P(A)=p.$$

**定义 1.9（概率的古典定义）** 我们把具有特性：

1）试验的结果是有限个；

2）每个结果出现的可能性是相同的

随机试验称为**古典概型随机试验**，就是说，在我们所讨论的基本事件空间 $\Omega$ 中，基本事件 $\omega_1,\omega_2,\cdots,\omega_n$ 是有限个，并且

$$P(\omega_1)=P(\omega_2)=\cdots=P(\omega_n).$$

设古典概型随机试验的基本事件空间由 $n$ 个基本事件组成，即 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$. 如果事件 $A$ 是由上述 $n$ 个事件中的 $m$ 个组成，则称事件 $A$ 发生的概率为

$$P(A)=\frac{m}{n}. \tag{1.1}$$

所谓**古典概型**就是利用关系式(1.1)来讨论事件发生的概率的数学模型.

**例 8** 一个口袋装有 10 个外形相同的球，其中 6 个是白球，4 个是红球. "无放回"地从袋中取出 3 个球，求下述诸事件发生的概率（所谓"无放回"是指，第一次取一个球，不再把这个球放回袋中，再去取另一个球）：

(1) $A_1=\{$没有红球$\}$；　(2) $A_2=\{$恰有两个红球$\}$；

(3) $A_3=\{$至少有两个红球$\}$；　(4) $A_4=\{$至多有两个红球$\}$；

(5) $A_5=\{$至少有一个白球$\}$；　(6) $A_6=\{$颜色相同的球$\}$.

**解** 设 $A=\{$任取三个球$\}$，其基本事件空间中基本事件的个数（即从 10 个球中任取 3 个的"一般组合"数）

$$n=C_{10}^3=120.$$

(1) $A_1$ 是由上面 120 个基本事件中的

$$m_1 = C_6^3 C_4^0 = 20$$

个组成.这里的 $C_6^3 C_4^0$ 是从 6 个白球中任取 3 个,从 4 个红球中取出 0 个(即不取红球的“两类不同元素的组合”数).根据关系式(1.1),有

$$P(A_1) = \frac{m_1}{n} = \frac{20}{120} = \frac{1}{6}.$$

(2) $A_2$ 是由基本事件中的

$$m_2 = C_6^1 C_4^2 = 36$$

个组成.根据关系式(1.1),有

$$P(A_2) = \frac{m_2}{n} = \frac{36}{120} = \frac{3}{10}.$$

(3) $A_3$ 是由基本事件中的

$$m_3 = C_6^1 C_4^2 + C_6^0 C_4^3 = 40$$

个组成.根据关系式(1.1),有

$$P(A_3) = \frac{m_3}{n} = \frac{40}{120} = \frac{1}{3}.$$

(4) $A_4$ 是由基本事件中的

$$m_4 = C_6^3 C_4^0 + C_6^2 C_4^1 + C_6^1 C_4^2 = 116$$

个组成.根据关系式(1.1),有

$$P(A_4) = \frac{m_4}{n} = \frac{116}{120} = \frac{29}{30}.$$

(5) $A_5$ 是由基本事件中的

$$m_5 = C_6^1 C_4^2 + C_6^2 C_4^1 + C_6^3 C_4^0 = 116$$

个组成.根据关系式(1.1),有

$$P(A_5) = \frac{m_5}{n} = \frac{116}{120} = \frac{29}{30}.$$

(6) $A_6$ 是由基本事件中的

$$m_6 = C_6^3 C_4^0 + C_6^0 C_4^3 = 24$$

个组成.根据关系式(1.1),有

$$P(A_6)=\frac{m_6}{n}=\frac{24}{120}=\frac{1}{5}.$$

**例 9** 在例 8 的条件下,"有放回"地从袋中取出 3 个球,求例 8 中诸事件发生的概率(所谓"有放回"是指,第一次取一个球,记录下这个球的颜色后,再把这个球放回袋中,然后再去任取一个球).

**解** 显然有放回的抽取是一个可重复的排列问题,于是基本事件的个数

$$n=10^3=1000.$$

(1) $A_1$ 是由上面 1000 个基本事件中的

$$m_1=6^3=216$$

个组成.根据关系式(1.1),有

$$P(A_1)=\frac{m_1}{n}=\frac{216}{1000}=0.216.$$

(2) $A_2$ 是由基本事件中的

$$m_2=3\times 6\times 4^2=288$$

个组成.根据关系式(1.1),有

$$P(A_2)=\frac{m_2}{n}=\frac{288}{1000}=0.288.$$

(3) $A_3$ 是由基本事件中的

$$m_3=3\times 6\times 4^2+4^3=352$$

个组成.根据关系式(1.1),有

$$P(A_3)=\frac{m_3}{n}=\frac{352}{1000}=0.352.$$

(4) $A_4$ 是由基本事件中的

$$m_4=3\times 6\times 4^2+3\times 6^2\times 4+6^3=936$$

个组成.根据关系式(1.1),有

$$P(A_4)=\frac{m_4}{n}=\frac{936}{1000}=0.936.$$

(5) $A_5$ 是由基本事件中的

$$m_5 = 6^3 + 3 \times 6^2 \times 4 + 3 \times 6 \times 4^2 = 936$$

个组成.根据关系式(1.1),有

$$P(A_5) = \frac{m_5}{n} = \frac{936}{1000} = 0.936.$$

(6) $A_6$ 是由基本事件中的

$$m_6 = 6^3 + 4^3 = 280$$

个组成.根据关系式(1.1),有

$$P(A_6) = \frac{m_6}{n} = \frac{280}{1000} = 0.28.$$

**例 10** 把 10 本书随意放在书架上,求其中指定的 5 本书放在一起的概率.

**解** 基本事件总数 $n=10!$,有利于将指定的 5 本书放在一起的基本事件个数 $m=6!\,5!$(其中 6! 是指 5 本书当作一个元素进行全排列的总数,5! 是 5 本书相互之间进行全排列的总数),故

$$P(A) = \frac{6!5!}{10!} = \frac{1}{42}.$$

**例 11** 将 $k$ 个不同的球随机地放入 $N$ 个盒子中去($k \leqslant N$).假设每个盒子能容纳的球数不限,试求下列事件的概率:

(1) 指定的 $k$ 个盒子中各有一球(事件 $A$);

(2) 恰有 $k$ 个盒子,其中各有一球(事件 $B$).

**解** 将 $k$ 个球放入 $N$ 个盒子中去的不同放法作为不同的基本事件,因每只球都有 $N$ 个盒子可供放入,故共有 $N \times N \times \cdots \times N = N^k$ 种不同放法,即基本事件总数 $n=N^k$.

(1) 对事件 $A$,其不同放法相当于这 $k$ 个球在这 $k$ 个指定位置的全排列 $k!$,即 $m=k!$,故

$$P(A) = \frac{m}{n} = \frac{k!}{N^k}.$$

(2) 对事件 $B$,首先在 $N$ 个盒子中任选 $k$ 个盒子出来,其不同选法有 $C_N^k$ 种;再在选定的 $k$ 个盒子中各放一球,其不同放法由(1)知为 $k!$ 种,故有利于事件 $B$ 的不同放法共有 $C_N^k \cdot k!$ 种,即 $m$

$=C_N^k \cdot k! = A_N^k$,所以

$$P(B)=\frac{m}{n}=\frac{A_N^k}{N^k}.$$

**例 12** 考虑一元二次方程 $x^2+Bx+C=0$,其中 $B$,$C$ 分别是将一枚色子(骰子)接连掷两次先后出现的点数,求该方程有实根的概率 $p$ 和有重根的概率 $q$.

**解** 一枚色子(骰子)掷两次其基本事件总数 $n=6^2=36$,方程有实根的充要条件是 $B^2-4C\geqslant 0$ 或 $C\leqslant B^2/4$. 易见

| $B$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 使 $C\leqslant\frac{B^2}{4}$ 的基本事件数 | 0 | 1 | 2 | 4 | 6 | 6 |
| 使 $C=\frac{B^2}{4}$ 的基本事件数 | 0 | 1 | 0 | 1 | 0 | 0 |

所以

$$p=\frac{1+2+4+6+6}{36}=\frac{19}{36},\quad q=\frac{1+1}{36}=\frac{1}{18}.$$

**定义 1.10(概率的几何定义)** 我们把具有特性:

1) 试验的结果是无限且不可列的;

2) 每个结果出现的可能性是均匀的

随机试验称为**几何型随机试验**. 在几何型随机试验中,我们是通过几何度量(长度、面积、体积等)来计算事件出现的可能性.

设 $E$ 为几何型的随机试验,其基本事件空间中的所有基本事件可以用一个有界区域来描述,而其中一部分区域可以表示事件 $A$ 所包含的基本事件,则称事件 $A$ 发生的概率为

$$P(A)=\frac{L(A)}{L(\Omega)},\tag{1.2}$$

其中 $L(\Omega)$ 与 $L(A)$ 分别为 $\Omega$ 与 $A$ 的**几何度量**.

所谓**几何概型**就是利用关系式(1.2)来讨论事件发生的概率的数学模型.

注意，上述事件 $A$ 的概率 $P(A)$ 只与 $L(A)$ 有关，而与 $L(A)$ 对应区域的位置及形状无关.

**例 13** 某地铁每隔五分钟有一列车通过，在乘客对列车通过该站时间完全不知道的情况下，求每一个乘客到站等车时间不多于 2 分钟的概率.

**解** 设 $A=\{$每一个乘客等车时间不多于 2 分钟$\}$. 由于乘客可以在接连两列车之间的任何一个时刻到达车站，因此每一乘客到达站台时刻 $t$ 可以看成是均匀地出现在长为 5 分钟的时间区间上的一个随机点，即 $\Omega=[0,5)$. 又设前一列车在时刻 $T_1$ 开出，后一列车在时刻 $T_2$ 到达，线段 $T_1T_2$ 长为 5（见图 1.1），即 $L(\Omega)=5$；$T_0$ 是 $T_1T_2$ 上一点，且 $T_0T_2$ 长为 2. 显然，乘客只有在 $T_0$ 之后到达（即只有 $t$ 落在线段 $T_0T_2$ 上），等车时间才不会多于 2 分钟，即 $L(A)=2$. 因此

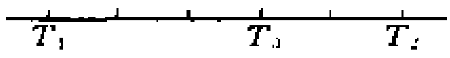


图 1.1

$$P(A)=\frac{L(A)}{L(\Omega)}=\frac{2}{5}.$$

**例 14**（会面问题） 甲乙两艘轮船驶向一个不能同时停泊两艘轮船的码头，它们在一昼夜内到达的时间是等可能的，如果甲船的停泊时间是一小时，乙船停泊的时间是两小时，求它们中任何一艘都不需要等候码头空出的概率是多少？

**解** 这是一个几何概型问题. 设 $A=\{$它们中任何一艘都不需要等候码头空出$\}$. 又设甲乙两船到达的时刻分别是 $x,y$，则 $0\leqslant x\leqslant 24$，$0\leqslant y\leqslant 24$. 由题意，若甲先到，则乙必须晚 1 小时到达，即 $y\geqslant 1+x$，若乙先到，则甲必须晚 2 小时到达，即 $x\geqslant y+2$. 由图1.2可知：$L(\Omega)$ 是由 $x=0$，$x=24$，$y=0$，$y=24$ 所围图形面积 $S=24^2$，而

$$L(A)=S_1+S_2=\frac{1}{2}(24-1)^2+\frac{1}{2}(24-2)^2=\frac{1}{2}(23^2+22^2).$$

所以 $$P(A)=\frac{L(A)}{L(\Omega)}=\frac{S_1+S_2}{S}=\frac{1013}{1152}\approx 0.8793.$$

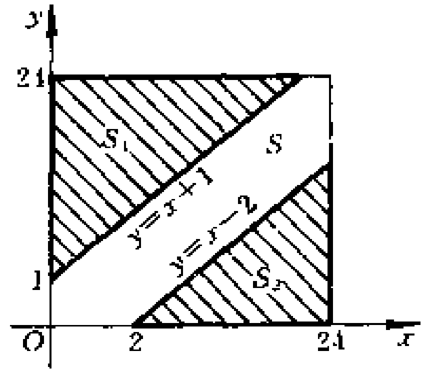



图 1.2

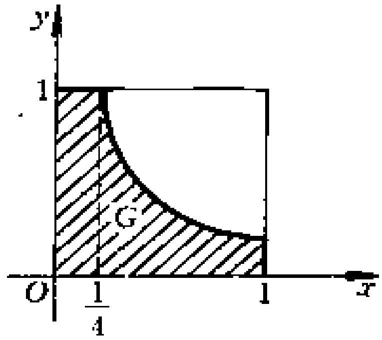



图 1.3

**例 15** 从区间(0,1)内任取两个数,求这两个数的积小于 1/4 的概率.

**解** 设 $A=\{$这两个数的积小于 $1/4\}$,以 $x,y$ 表示从(0,1)内任取的两个数,那末样本空间所对应的区域为(见图 1.3):

$$S=\{(x,y):0<x<1,\ 0<y<1\},$$

$S$ 的面积$=1$. 事件 $A$ 所包含的区域为

$$G=\left\{(x,y):xy<\frac{1}{4},\ 0<x<1,\ 0<y<1\right\},$$

区域 $G$ 的面积为

$$\int_{\frac{1}{4}}^{1}\frac{1}{4x}\mathrm{d}x+\frac{1}{4}=\frac{1}{2}\ln 2+\frac{1}{4},$$

于是所求的概率为

$$P(A)=\frac{G\text{ 的面积}}{S\text{ 的面积}}=\frac{1}{2}\ln 2+\frac{1}{4}.$$

**定义 1.11(条件概率)** 在一般情况下,我们所讨论的事件 $B$ 的概率 $P_S(B)$,都是指在一组不变条件 $S$ 下事件 $B$ 发生的概率(但是为了叙述简练,一般不再提及条件组 $S$,而把 $P_S(B)$简记为 $P(B)$).

设 $A$、$B$ 是条件 $S$ 下的两个随机事件,且$P(A)\neq 0$. 则把在 $A$ 发生的前提下 $B$ 发生的概率称为**条件概率**,记作 $P(B|A)$,读作在 $A$ 发生的条件下事件 $B$ 的概率,其定义为

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

**例 16** 在 100 个圆柱形零件中有 95 件长度合格，有 93 件直径合格，有 90 件两个指标都合格. 从中任取一件(这就是条件 $S$)讨论在长度合格的前提下，直径也合格的概率.

**解** 设 $A$=\{任取一件，长度合格\}，$B$=\{任取一件，直径合格\}，$AB$=\{任取一件，长度与直径都合格\}. 根据古典概型，在条件 $S$ 下，基本事件的总数

$$n=C_{100}^{1}.$$

事件 $A$ 与 $B$ 所包含的基本事件个数分别为

$$m_A=C_{95}^{1},\quad m_B=C_{93}^{1}.$$

$AB$ 所包含的基本事件个数为

$$m_{AB}=C_{90}^{1}.$$

所以在长度合格的情况下直径也合格的零件概率 $P(B|A)$ 为

$$P(B|A)=\frac{P(AB)}{P(A)}=\frac{90/100}{95/100}=\frac{90}{95}.$$

**例 17** 设随机事件 $B$ 是 $A$ 的子事件，已知 $P(A)=1/4$，$P(B)=1/6$，求 $P(B|A)$.

**解** 因为 $B\subset A$，所以 $P(B)=P(AB)$. 因此，

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

**例 18** 某牌号的电视机使用到 3 万小时的概率为 0.6，使用到 5 万小时的概率为 0.24，一台电视机已使用到 3 万小时，求这台电视机使用到 5 万小时的概率.

**解** 设 $A$=\{使用到 3 万小时\}，$B$=\{使用到 5 万小时\}，于是 $P(A)=0.6$，$P(AB)=P(B)=0.24$，则

$$P(B|A)=\frac{P(AB)}{P(A)}=0.4.$$

**定义 1.12(完备事件组)** 如果事件组 $A_1,A_2,\cdots,A_n$ 满足

1) $\sum_{i=1}^{n} A_i = \Omega$ 且 $P(A_i) > 0 \ (i=1,2,\cdots,n)$；

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

则称之为**完备事件组**.

**定义 1.13(独立重复试验)** 在实际问题中，我们常常要做多次试验条件完全相同(即可以看成是一个试验的多次重复)并且都是相互独立(即每次试验中的随机事件的概率不依赖于其他各次试验的结果)的试验. 我们称这种类型的试验为**独立重复试验**.

(二) 重要定理及公式

**公式 1.1(加法公式)** 设 $A,B$ 为任意两个随机事件，则

$$P(A+B) = P(A) + P(B) - P(AB). \qquad (1.3)$$

上面的加法公式可以推广到有限多个事件的情况，例如对于三个事件 $A_1,A_2,A_3$，我们有

$$P(A_1 + A_2 + A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1A_2) - P(A_2A_3) - P(A_3A_1) + P(A_1A_2A_3).$$

**推论 1** 设 $A$ 为任意随机事件，则

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

**推论 2** 设 $A,B$ 为两个任意的随机事件，若 $A \subset B$，则

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

由于 $P(B-A) \geqslant 0$，根据推论 2 可以推得：当 $A \subset B$ 时，

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

**例 19** 一批产品共有 100 件，其中 90 件是合格品，10 件是次品，从这批产品中任取 3 件，求其中有次品的概率.

**解 方法 1** 设 $A=\{$有次品$\}$，$A_i=\{$有 $i$ 件次品$\}$，$i=1,2,3$. 故 $A=A_1 \cup A_2 \cup A_3$，并且 $A_1,A_2,A_3$ 是两两互斥的，由概率的古典定义，我们有

$$P(A_1) = \frac{C_{10}^1 \cdot C_{90}^2}{C_{100}^3} = 0.24768,$$

$$P(A_2) = \frac{C_{10}^2 \cdot C_{90}^1}{C_{100}^3} = 0.02504,$$

$$P(A_3)=\frac{C_{10}^3}{C_{100}^3}=0.00074.$$

于是

$$P(A)=P(A_1)+P(A_2)+P(A_3)=0.2735.$$

**方法 2** 由于事件 $A$ 的对立事件 $\overline{A}=\{$取出的 3 件产品全是合格品$\}$,故

$$P(\overline{A})=\frac{C_{90}^3}{C_{100}^3}=0.7265.$$

由推论 1,

$$P(A)=1-P(\overline{A})=1-0.7265=0.2735.$$

**例 20** 某一企业与甲乙两公司签订某物资长期供货关系的合同,由以前的统计得知,甲公司按时供货的概率为 0.9,乙公司能按时供货的概率为 0.75,两公司都能按时供货的概率为 0.7,求至少有一公司能按时供货的概率.

**解** 分别用 $A,B$ 表示甲乙两公司按时供货的事件,由题意,$A,B$ 为非互斥事件,由(1.3),我们有

$$\begin{aligned}P(A\cup B)&=P(A)+P(B)-P(AB)\\&=0.9+0.75-0.7=0.95.\end{aligned}$$

故至少有一公司能按时供货的概率为 0.95.

**例 21** 在某城市中,共发行三种报纸 $A,B,C$. 在这城市的居民中,订购 $A$ 的占 45%,订购 $B$ 的占 35%,订购 $C$ 的占 30%,同时订购 $A$ 及 $B$ 的占 10%,同时订购 $A$ 及 $C$ 的占 8%,同时订购 $B$ 及 $C$ 的占 5%,同时订购 $A,B,C$ 的占 3%. 试求下列百分率:

(1) 只订购 $A$ 的; (2) 只订购 $A$ 及 $B$ 的;

(3) 只订购一种报纸的; (4) 正好订购两种报纸的;

(5) 至少订购一种报纸的; (6) 不订购任何报纸的.

**解**

(1) $P\{$只订购 $A$ 的$\}=P(A\overline{B}\overline{C})=P(A-(B\cup C))$

$=P(A-A(B\cup C))=P(A)-P(AB\cup AC)$

$$=P(A)-P(AB)-P(AC)+P(ABC)$$

$$=0.45-0.10-0.08+0.03$$

$$=0.30=30\%.$$

(2) $P\{$只订购 $A$ 及 $B$ 的$\}=P(AB\overline{C})=P(AB-ABC)$

$$=P(AB)-P(ABC)=0.10-0.03$$

$$=0.07=7\%.$$

(3) $P\{$只订购 $B$ 的$\}=P(B\overline{A}\overline{C})=P(B-B(A\cup C))$

$$=P(B)-P(BA\cup BC)$$

$$=P(B)-P(BA)-P(BC)+P(ABC)$$

$$=0.35-0.10-0.05+0.03=0.23.$$

$P\{$只订购 $C$ 的$\}=P(C\overline{A}\overline{B})=P(C-C(A\cup B))$

$$=P(C)-P(AC\cup BC)$$

$$=P(C)-P(AC)-P(BC)+P(ABC)$$

$$=0.30-0.08-0.05+0.03=0.20.$$

所以 $P\{$只订购一种报纸的$\}=P\{$只订购 $A$ 的$\}+P\{$只订购 $B$ 的$\}+P\{$只订购 $C$ 的$\}=0.30+0.23+0.20=0.73=73\%.$

(4) $P\{$正好订两种报纸的$\}=P\{AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC\}$

$$=P(AB\overline{C})+P(A\overline{B}C)+P(\overline{A}BC)$$

$$=P(AB-ABC)+P(AC-ABC)+P(BC-ABC)$$

$$=P(AB)-P(ABC)+P(AC)-P(ABC)+P(BC)-P(ABC)$$

$$=0.10+0.08+0.05-0.03\times 3$$

$$=0.14=14\%.$$

(5) $P\{$至少订购一种报纸的$\}=P(A\cup B\cup C)$

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

$$=0.45+0.35+0.30-0.10-0.08-0.05+0.03$$

$=0.90=90\%.$

(6) $P\{$不订购任何报纸的$\}=1-P\{$至少订购一种报纸的$\}$

$=1-0.90=0.10=10\%.$

**公式 1.2(乘法公式)** 两个事件 $A$ 与 $B$ 的积的概率等于事件 $A$ 的概率乘以在 $A$ 发生的前提下 $B$ 发生的概率，即

$$P(AB)=P(A)P(B|A) \quad (P(A)>0). \tag{1.4}$$

同理有

$$P(AB)=P(B)P(A|B) \quad (P(B)>0).$$

上述的计算公式可以推广到有限多个事件的情形，例如对于三个事件 $A_1,A_2,A_3$(若 $P(A_1)>0,P(A_1A_2)>0$)有

$$P(A_1A_2A_3)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2).$$

**例 22** 在 100 件产品中有 5 件是不合格的，无放回地抽取两件，问第一次取到正品而第二次取到次品的概率是多少?

**解** 设事件

$A=\{$第一次取到正品$\}$, $B=\{$第二次取到次品$\}$,

用古典概型方法可求出

$$P(A)=\frac{95}{100}\neq 0.$$

由于第一次取到正品后不放回，那么第二次是在 99 件中(不合格品仍是 5 件)任取一件，所以

$$P(B|A)=\frac{5}{99}.$$

由公式(1.4),

$$P(AB)=P(A)P(B|A)=\frac{95}{100}\cdot\frac{5}{99}=\frac{19}{396}.$$

**例 23(抓阄问题)** 五个人抓一个有物之阄，求第二个人抓到的概率.

**解** 这是一个乘法公式的问题. 设 $A_i=\{$第 $i$ 个人抓到有物之阄$\}$($i=1,2,3,4,5$)，有

$$A_2 = A_2\Omega = A_2(A_1 + \overline{A}_1)$$
$$= A_1A_2 + \overline{A}_1A_2 = \varnothing + \overline{A}_1A_2 = \overline{A}_1A_2.$$

根据事件相同，对应概率相等有

$$P(A_2) = P(\overline{A}_1A_2) = P(\overline{A}_1)P(A_2 \mid \overline{A}_1).$$

又因为 $P(A_1)=\frac{1}{5}$，$P(\overline{A}_1)=\frac{4}{5}$，$P(A_2\mid\overline{A}_1)=\frac{1}{4}$，所以

$$P(A_2)=\frac{4}{5}\times\frac{1}{4}=\frac{1}{5}.$$

在实际应用中，常常是根据问题的具体情况，按照独立性的直观意义来判定事件的独立性. 例如，连续两次抛掷一枚硬币，事件{第一次出现正面}与{第二次出现正面}是相互独立的.

**例 24** 一盒螺钉共有 20 个，其中 19 个是合格的，另一盒螺母也有 20 个，其中 18 个是合格的. 现从两盒中各取一个螺钉和螺母，求两个都是合格品的概率.

**解** 令

$A$ = {任取一个，螺钉合格}，　$B$ = {任取一个，螺母合格}.

显然 $A$ 与 $B$ 是相互独立的，并且有

$$P(A) = \frac{C_{19}^1}{C_{20}^1} = \frac{19}{20},\quad P(B) = \frac{C_{18}^1}{C_{20}^1} = \frac{9}{10}.$$

由公式

$$P(AB) = P(A)P(B) = \frac{19}{20}\times\frac{9}{10} = \frac{171}{200}.$$

**例 25** 设 10 件产品中有 4 件不合格品，从中任取两件，已知所取的两件产品中有一件是不合格品，求另一件也是不合格品的概率.

**解** 设 $A$ 表示所取两件产品中有一件是不合格品的事件，$B$ 表示另一件也是不合格品的事件，则 $AB$ 表示所取两件产品都是不合格品的事件.

$$P(A) = [C_4^1C_6^1 + C_4^2]/C_{10}^2,\quad P(AB) = C_4^2/C_{10}^2,$$

故所求概率为

$$P(B|A)=\frac{P(AB)}{P(A)}=\frac{C_4^2/C_{10}^2}{[C_4^2+C_4^1C_6^1]/C_{10}^2}$$

$$=\frac{C_4^2}{C_4^2+C_4^1C_6^1}=\frac{6}{30}=\frac{1}{5}.$$

**例 26** 某种动物由出生活到 10 岁的概率为 0.8，活到 12 岁的概率为 0.56，问现年 10 岁的这种动物活到 12 岁的概率是多少？

**解** $A=\{$活到 10 岁以上$\}$，$B=\{$活到 12 岁以上$\}$，显然 $B\subset A$. 因为 $P(A)=0.8$，$P(B)=0.56$，又 $B\subset A$，$AB=B$，$P(AB)=P(B)=0.56$，所以所求概率

$$P(B|A)=\frac{P(AB)}{P(A)}=\frac{P(B)}{P(A)}=\frac{0.56}{0.8}=0.7.$$

**例 27** 某厂的产品中有 4%的废品，在 100 件合格品中有 75 件一等品，试求在该厂中任取一件产品是一等品的概率.

**解** $A=\{$任取一件产品是合格品$\}$，$B=\{$任取一件产品是一等品$\}$，显然 $B\subset A$.

$$P(A)=1-P(\overline{A})=96\%,\quad P(B|A)=75\%,$$

$$P(B)=P(AB)=P(A)P(B|A)=0.96\times 0.75=0.72.$$

**例 28** 用高射炮射击飞机. 如果每门高射炮击中飞机的概率是 0.6，试问：(1) 用两门高射炮分别射击一次击中飞机的概率是多少？(2) 若有一架敌机入侵，需要多少架高射炮同时射击才能以 99%的概率命中敌机？

**解** (1) 令

$$B_i=\{\text{第 } i \text{ 门高射炮击中敌机}\}\ (i=1,2),$$

$$A=\{\text{击中敌机}\}.$$

在同时射击时，$B_1$ 与 $B_2$ 可以看成是互相独立的，从而 $\overline{B}_1$，$\overline{B}_2$ 也是相互独立的，且有

$$P(B_1)=P(B_2)=0.6,$$

$$P(\overline{B}_1)=P(\overline{B}_2)=1-P(B_1)=0.4.$$

故 $P(A)=1-P(\overline{A})=1-P(\overline{B}_1\overline{B}_2)=1-P(\overline{B}_1)P(\overline{B}_2)$
$=1-0.4^2=0.84.$

(2) 令 $n$ 是以 99%的概率击中敌机所需高射炮的门数，由上面讨论可知，

$$99\%=1-0.4^n \quad 即 \quad 0.4^n=0.01,$$

亦即
$$n=\frac{\lg 0.01}{\lg 0.4}=\frac{-2}{-0.3979}\approx 5.026.$$

因此若有一架敌机入侵，至少需要配置 6 门高射炮方能以 99%的把握击中它.

**定理 1.1（全概率公式）** 设 $A_1,A_2,\cdots,A_n$ 是一个完备事件组，则对任一事件 $B$，有

$$P(B)=\sum_{i=1}^{n}P(A_i)P(B|A_i). \tag{1.5}$$

上式称之为**全概率公式**.

全概率公式是概率的加法公式和乘法公式的综合. 利用这个公式可以从较简单事件的概率推算出复杂事件的概率，即把复杂事件 $B$ 分解成简单事件之和的形式：$\bigcup_{i=1}^{n}A_iB$. 只要求出互不相容的事件 $A_iB$ 的概率

$$P(A_iB)=P(A_i)P(B|A_i) \quad (i=1,2,\cdots,n),$$

就可以得到事件 $B$ 的概率.

**例 29** 三人同时向一架飞机射击. 设三人都射不中的概率为 0.09，三人中只有一人射中的概率为 0.36，三人中恰有两人射中的概率为 0.41，三人同时射中的概率为 0.14. 又设无人射中，飞机不会坠毁；只有一人击中飞机坠毁的概率为 0.2；两人击中飞机坠毁的概率为 0.6；三人射中飞机一定坠毁. 求三人同时向飞机射击一次飞机坠毁的概率.

**解** 令 $B=\{飞机坠毁\}$，

$A_0=\{三人都射不中\}$，　$A_1=\{只有一人射中\}$，

$A_2=\{恰有两人射中\}$，　$A_3=\{三人同时射中\}$.

显然有$\sum_{i=0}^{3}A_i=\Omega$,且$A_iA_j=\varnothing(i\neq j;i,j=0,1,2,3)$.

由题设可知

$$P(B|A_0)=0,\qquad P(B|A_1)=0.2,$$
$$P(B|A_2)=0.6,\quad P(B|A_3)=1,$$

并且

$$P(A_0)=0.09,\quad P(A_1)=0.36,$$
$$P(A_2)=0.41,\quad P(A_3)=0.14.$$

利用全概率公式可以得到

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

**例30** 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0.1.一顾客欲买下一箱玻璃杯,在购买时,售货员随意取出一箱,而顾客开箱随意察看其中的4只,若无残次品,则买下该箱玻璃杯,否则退回.试求

(1) 顾客买下该箱的概率$\alpha$;

(2) 在顾客买下的一箱中,确实没有残次品的概率$\beta$.

**解** 设$A_i$={一箱中含有$i$只残次品}($i=0,1,2$),$B$={顾客买下所察看的一箱},则

$$P(A_0)=0.8,\quad P(A_1)=P(A_2)=0.1,$$
$$P(B|A_0)=1,\quad P(B|A_1)=\frac{C_{19}^4}{C_{20}^4}=\frac{4}{5},$$
$$P(B|A_2)=\frac{C_{18}^4}{C_{20}^4}=\frac{12}{19},$$

于是

$$\begin{aligned}\alpha=P(B)&=\sum_{i=0}^{2}P(A_i)P(B|A_i)\\&=0.8\times1+0.1\times\frac{4}{5}+0.1\times\frac{12}{19}=0.943,\end{aligned}$$

$$\beta = P(A_0|B) = \frac{P(A_0B)}{P(B)} = \frac{P(A_0)P(B|A_0)}{\alpha}$$

$$= \frac{0.8 \times 1}{0.943} \approx 0.85.$$

**定理 1.2**(贝叶斯(Bayes)公式) 设 $A_1, A_2, \cdots, A_n$ 是一个完备事件组,则对任一事件$B(P(B)>0)$,在事件 $B$ 已发生的条件下事件 $A_i$ 发生的概率为

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{n} P(A_j)P(B|A_j)} \quad (1.6)$$

$$(i = 1, 2, \cdots, n).$$

上式称之为**贝叶斯公式**.

贝叶斯公式主要用于在已知某事件 $B$ 发生的条件下,来判断 $B$ 是伴随着 $A_1, A_2, \cdots, A_n$ 中的哪一个事件发生的情况下而发生的,即要求知道 $B$ 发生的条件下某个原因 $A_i$ 发生的概率,这就是条件概率 $P(A_i|B)$,所以这个公式又称为**原因概率公式**.

**例 31** 设某人从外地赶来参加紧急会议. 他乘火车、轮船、汽车或飞机来的概率分别是 3/10, 1/5, 1/10 及 2/5, 如果他乘飞机来,不会迟到;而乘火车、轮船或汽车来迟到的概率分别为 1/4, 1/3, 1/12. 此人迟到,试推断他是怎样来的?

**解** 令 $A_1=\{$乘火车$\}$, $A_2=\{$乘轮船$\}$, $A_3=\{$乘汽车$\}$, $A_4=\{$乘飞机$\}$, $B=\{$迟到$\}$. 按题意有:

$$P(A_1) = \frac{3}{10}, \quad P(A_2) = \frac{1}{5},$$

$$P(A_3) = \frac{1}{10}, \quad P(A_4) = \frac{2}{5}, \quad P(B|A_1) = \frac{1}{4},$$

$$P(B|A_2) = \frac{1}{3}, P(B|A_3) = \frac{1}{12}, \quad P(B|A_4) = 0.$$

将这些数值代入贝叶斯公式

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{4} P(A_j)P(B|A_j)} \ (i = 1,2,3,4)$$

得到 $$P(A_1|B)=\frac{1}{2},\quad P(A_2|B)=\frac{4}{9},$$
$$P(A_3|B)=\frac{1}{18},\quad P(A_4|B)=0.$$
由上述计算结果可以推断出此人迟到乘火车的可能性最大.

**例 32** 一个工厂有甲、乙、丙三个车间生产同一种产品，这三个车间的产量分别占总产量的 25%，35%，40%. 如果每个车间成品中的次品占其产量的 5%，4%，2%，求在从全厂该产品中抽出一个是次品的条件下它恰好是甲车间生产出来的概率.

**解** 设事件 $A_1,A_2,A_3$ 分别是所抽出的产品是甲、乙、丙车间生产的，事件 $B$ 是抽到一个次品，则

$$P(A_1)=\frac{25}{100},\qquad P(A_2)=\frac{35}{100},\qquad P(A_3)=\frac{40}{100},$$
$$P(B|A_1)=\frac{5}{100},\quad P(B|A_2)=\frac{4}{100},\quad P(B|A_3)=\frac{2}{100}.$$
应用全概率公式，有
$$P(B)=\sum_{i=1}^{3}P(A_i)P(B|A_i)$$
$$=\frac{25}{100}\cdot\frac{5}{100}+\frac{35}{100}\cdot\frac{4}{100}+\frac{40}{100}\cdot\frac{2}{100}=\frac{345}{10000}.$$
再应用贝叶斯公式，有
$$P(A_1|B)=\frac{P(A_1)P(B|A_1)}{\sum_{k=1}^{3}P(A_k)P(B|A_k)}=\frac{\frac{25}{100}\cdot\frac{5}{100}}{\frac{345}{10000}}=\frac{125}{345}=\frac{25}{69}.$$
利用全概率公式和贝叶斯公式计算概率的关键是找满足全概率公式中条件的事件组，即完备事件组 $A_1,A_2,\cdots,A_n$. 要掌握以下要点：

1）事件 $B$ 必须伴随着 $n$ 个互不相容事件 $A_1,A_2,\cdots,A_n$ 之一发生，$B$ 的概率就可用全概率公式计算.

2）如果我们已知事件 $B$ 发生了，求事件 $A_j(j=1,2,\cdots,n)$的

概率,则应使用贝叶斯公式. 这里用贝叶斯公式计算的是条件概率 $P(A_j|B)(j=1,\cdots,n)$.

**定理 1.3**(伯努利(Bernoulli)概型)　在单次试验中事件 $A$ 发生的概率为 $p(0<p<1)$,则在 $n$ 次独立重复试验中

$$P\{A \text{ 发生 } k \text{ 次}\} = C_n^k p^k(1-p)^{n-k}(k=0,1,2,\cdots,n). \tag{1.7}$$

所谓**伯努利概型**就是利用关系式(1.7)来讨论事件概率的数学模型. 伯努利概型又称为**独立试验序列概型**(或**二项概型**). 在这个概型中基本事件的概率可以直接计算出来. 与古典概型不同的是伯努利概型的基本事件不一定是等概的.

**例 33**　一批产品废品率为 10%,每次抽取一个,观察后再放回去,独立地重复 5 次,求 5 次观察中恰有 2 次是废品的概率.

**解**　设 $A$ 表示"1 次观察中出现废品",$B$ 表示"5 次观察中出现 2 次废品". 由题意可知:$P(A)=0.1$,这样我们就有

$$P(B)=C_5^2(P(A))^2(1-P(A))^{5-2} = 10\times 0.1^2\times 0.9^3 = 0.0729.$$

**例 34**　某车间有 10 台用电各为 7.5 千瓦的机床,如果每台机床使用情况是相互独立的,且每台机床平均每小时开动 12 分钟,问全部机床用电超过 48 千瓦的概率为多少?

**解**　设事件 $A$ 为一台机床工作,则

$$P(A)=\frac{12}{60}=\frac{1}{5},\quad P(\overline{A})=\frac{4}{5}.$$

故由二项概型知某一时刻恰有 $k$ 台机床工作的概率为

$$C_{10}^k\left(\frac{1}{5}\right)^k\left(\frac{4}{5}\right)^{10-k},\quad k=0,1,\cdots,10.$$

此外,48 千瓦可供 6 台机床同时工作,故事件"用电超过 48 千瓦"就等价于事件"有 7 台或 7 台以上的机床在工作". 因此,用电超过 48 千瓦的概率为

$$\sum_{k=7}^{10}C_{10}^k\left(\frac{1}{5}\right)^k\left(\frac{4}{5}\right)^{10-k}.$$

(三) 重要方法

**方法 1.1(加法法则)** 设完成一件事有 $n$ 类方法，只要选择任何一类中的一种方法，这件事就可以完成. 若第一类方法有 $m_1$ 种，第二类方法有 $m_2$ 种，…，第 $n$ 类方法有 $m_n$ 种，并且这 $m_1+m_2+\cdots+m_n$ 种方法里，任何两种方法都不相同，则完成这件事就有 $m_1+m_2+\cdots+m_n$ 种方法.

**方法 1.2(乘法法则)** 设完成一件事有 $n$ 个步骤，第一步有 $m_1$ 种方法，第二步有 $m_2$ 种方法，…，第 $n$ 步有 $m_n$ 种方法，并且完成这件事必须经过每一步，则完成这件事共有 $m_1\times m_2\times\cdots\times m_n$ 种方法.

**方法 1.3(排列)** 从 $n$ 个不同元素中，每次取出 $m$ 个元素，按照一定顺序排成一列，称为从 $n$ 个元素中每次取出 $m$ 个元素的排列.

(1) 元素可重复的排列

从 $n$ 个不同元素中，有放回地逐一取出 $m$ 个元素进行排列(简称为**可重复排列**)，共有 $n^m$ 种不同的排列.

(2) 元素不可重复的排列

从 $n$ 个不同元素中，无放回地取出 $m$ 个($m\leqslant n$)元素进行排列(简称为**选排列**)共有

$$n(n-1)\cdots(n-m+1)=\frac{n!}{(n-m)!}$$

种不同的排列. 选排列的种数用 $A_n^m$(或 $P_n^m$)表示，即

$$A_n^m=\frac{n!}{(n-m)!}.$$

特别地，当 $m=n$ 时的排列(简称为**全排列**)共有

$$n(n-1)(n-2)\cdot\cdots\cdot3\cdot2\cdot1=n!$$

种不同排列. 全排列的种数用 $P_n$(或 $A_n^n$)表示，即

$$P_n=n!,$$

并规定 $0!=1$.

**例 35** 从 10 件产品(其中 2 件次品,8 件正品)中每次取 1 件观测后放回,共取 3 次(以后简称为有放回地取 3 件).求这 3 件产品中

(1) 恰有 2 件次品的排列数;

(2) 至多有 1 件次品的排列数.

**解** (1) 这是一个可重复的排列问题,可知其排列数为 $3\times 2^2\times 8=96$;

(2)"至多有 1 件次品"可分为"恰有 1 件次品"与"没有次品"两种情况,可以得到其排列数为 $3\times 2\times 8^2+8^3=896$.

**方法 1.4(组合)** 从 $n$ 个不同元素中,每次取出 $m$ 个元素不考虑其先后顺序作为一组,称为从 $n$ 个元素中每次取出 $m$ 个元素的组合.

(1) 一般组合

从 $n$ 个不同元素中取出 $m$ 个元素的组合(简称为**一般组合**)共有

$$\frac{n(n-1)\cdots(n-m+1)}{m!}=\frac{n!}{m!(n-m)!}$$

种不同的组合.一般组合的组合种数用 $C_n^m\left(\text{或}\binom{n}{m}\right)$ 表示,即

$$C_n^m=\frac{n!}{m!(n-m)!},$$

并且规定 $C_n^0=1$. 不难看出

$$C_n^m=\frac{A_n^m}{P_m}.$$

组合数的两个性质:

$$C_n^m=C_n^{n-m},\quad C_n^m=C_{n-1}^m+C_{n-1}^{m-1}.$$

(2) 不同类元素的组合

从不同的 $k$ 类元素中,取出 $m$ 个元素.从第 1 类 $n_1$ 个不同元素中取出 $m_1$ 个,从第 2 类 $n_2$ 个不同的元素中取出 $m_2$ 个,…,从第 $k$ 类 $n_k$ 个不同的元素中取出 $m_k$ 个,并且 $n_i\geqslant m_i>0(i=1,2,\cdots,k)$

（简称为**不同类元素的组合**），共有

$$C_{n_1}^{m_1}C_{n_2}^{m_2}\cdots C_{n_k}^{m_k}=\prod_{i=1}^{k}C_{n_i}^{m_i}$$

种不同取法.

**例 36** 从 10 件产品（其中 2 件次品，8 件正品）之中任取 3 件，求这 3 件产品中

(1) 恰有 2 件次品的组合数；

(2) 至多有 1 件次品的组合数.

**解** (1) 这是一个不同类元素组合问题，其组合数为 $C_2^2C_8^1=8$；

(2) "至多有 1 件次品"可分为"恰有 1 件次品"与"没有次品"两种情况，可以得到其组合数为 $C_2^1C_8^2+C_2^0C_8^3=112$.

## 三、典型例题分析

### （一）填空题

1. 设在一次试验中，事件 $A$ 发生的概率为 $p$. 现进行 $n$ 次独立试验，则 $A$ 至少发生一次的概率为________，而事件 $A$ 至多发生一次的概率为________.

答案是：$1-(1-p)^n$，$(1-p)^n+np(1-p)^{n-1}$.

**分析** 设 $B_i$ 代表"在 $n$ 次独立试验中，事件 $A$ 发生 $i$ 次" $(i=0,1,\cdots,n)$，则有

$$P(B_0)=(1-p)^n,\quad P(B_1)=np(1-p)^{n-1},$$

因此，$A$ 至少发生一次的概率为 $1-(1-p)^n$；而事件 $A$ 至多发生一次的概率为 $(1-p)^n+np(1-p)^{n-1}$.

2. 三个箱子，第一个箱子中有 4 个黑球、1 个白球，第二个箱子中有 3 个黑球、3 个白球，第三个箱子有 3 个黑球、5 个白球. 现随机地取一个箱子，再从这个箱子中取出 1 个球，这个球为白球的概率等于________. 已知取出的球是白球，此球属于第二个箱子的概率为________.

答案是：$\frac{53}{120}$，$\frac{20}{53}$.

**分析** 用 $A_i$ 代表"取第 $i$ 只箱子"，$i=1,2,3$；用 $B$ 代表"取出的球是白球". 由全概率公式

$$P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+P(A_3)P(B|A_3)$$
$$=\frac{1}{3}\cdot\frac{1}{5}+\frac{1}{3}\cdot\frac{3}{6}+\frac{1}{3}\cdot\frac{5}{8}=\frac{53}{120}.$$

由贝叶斯公式

$$P(A_2|B)=\frac{P(A_2)P(B|A_2)}{P(B)}=\frac{\frac{1}{3}\cdot\frac{3}{6}}{\frac{53}{120}}=\frac{20}{53}.$$

3. 设三次独立试验中，事件 $A$ 出现的概率相等. 若已知 $A$ 至少出现一次的概率等于 19/27，则事件 $A$ 在一次试验中出现的概率为________.

答案是：$\frac{1}{3}$.

**分析** 设事件 $A$ 在一次试验中出现的概率为 $p(0<p<1)$，则有 $1-(1-p)^3=\frac{19}{27}$，从而解得 $p=\frac{1}{3}$.

4. 已知随机事件 $A$ 的概率 $P(A)=0.5$，随机事件 $B$ 的概率 $P(B)=0.6$ 及条件概率 $P(B|A)=0.8$，则和事件 $A\cup B$ 的概率 $P(A\cup B)=$________.

答案是：0.7.

**分析** 
$$P(A\cup B)=P(A)+P(B)-P(AB)$$
$$=P(A)+P(B)-P(A)P(B|A)$$
$$=0.5+0.6-0.5\times0.8=0.7.$$

5. 甲、乙两人独立地对同一目标射击一次，其命中率分别为 0.6 和 0.5. 现已知目标被命中，则它是甲射中的概率为________.

答案是：0.75.

**分析** 用 $A$ 代表事件"甲命中目标"，$B$ 代表事件"乙命中目标"，则 $A\cup B$ 代表事件"目标被命中"，且 $P(A\cup B)=P(A)+$

$P(B)-P(AB)=P(A)+P(B)-P(A)P(B)=0.5+0.6-0.5\times 0.6=0.8$,所求概率为

$$P(A|A\cup B)=\frac{P(A)}{P(A\cup B)}=\frac{0.6}{0.8}=0.75.$$

6. 设随机事件 $A,B$ 及其和事件 $A\cup B$ 的概率分别是 0.4, 0.3 和 0.6. 若 $\overline{B}$ 表示 $B$ 的对立事件,那么积事件 $A\overline{B}$ 的概率 $P(A\overline{B})=$________.

答案是: 0.3.

**分析** $P(AB)=P(A)+P(B)-P(A\cup B)$

$$=0.4+0.3-0.6=0.1.$$

因为 $A=A(B+\overline{B})=AB+A\overline{B}$,故

$$P(A\overline{B})=P(A)-P(AB)=0.4-0.1=0.3.$$

7. 已知 $P(A)=P(B)=P(C)=\frac{1}{4}$, $P(AB)=0$, $P(AC)=P(BC)=\frac{1}{8}$,则事件 $A$、$B$、$C$ 全不发生的概率为________.

答案是: $\frac{3}{8}$.

**分析** 由 $ABC\subset AB$, $P(AB)=0$ 得 $P(ABC)=0$,所求事件概率为

$$\begin{aligned}P(\overline{A}\cdot\overline{B}\cdot\overline{C})&=(\overline{A\cup B\cup C})=1-P(A\cup B\cup C)\\&=1-\{P(A)+P(B)+P(C)-P(AB)\\&\quad -P(AC)-P(BC)+P(ABC)\}\\&=\frac{3}{8}.\end{aligned}$$

8. 一批产品共有 10 个正品和 2 个次品,任意抽取两次,每次抽一个,抽出后不再放回,则第二次抽出的是次品的概率为________.

答案是: $\frac{1}{6}$.

**分析** 用 $A_i$ 代表事件"第 $i$ 次抽次品", $i=1,2$,则所求概率为

$$P(A_2)=P(A_1)P(A_2|A_1)+P(\bar{A}_1)P(A_2|\bar{A}_1)$$
$$=\frac{2}{12}\cdot\frac{1}{11}+\frac{10}{12}\cdot\frac{2}{11}=\frac{1}{6}.$$

9. 已知 $A$、$B$ 两个事件满足条件 $P(AB)=P(\overline{A}\overline{B})$，且 $P(A)=p$，则 $P(B)=$________.

答案是：$1-p$.

**分析** 由 $P(AB)=P(\overline{A}\overline{B})=P(\overline{A\cup B})=1-P(A\cup B)$
$$=1-[P(A)+P(B)-P(AB)]$$
$$=1-P(A)-P(B)+P(AB)$$
得
$$P(B)=1-P(A)=1-p.$$

10. 设工厂 $A$ 和工厂 $B$ 的次品率分别为 1%和 2%，现从由 $A$ 和 $B$ 的产品分别占 60%和 40%的一批产品中随机抽取一件，发现是次品，则该次品属 $A$ 生产的概率是________.

答案是：$\frac{3}{7}$.

**分析** 用 $A$ 和 $B$ 分别代表产品是工厂 $A$ 和工厂 $B$ 生产的，$C$ 代表产品是次品，则所求概率为
$$P(A|C)=\frac{P(A)P(C|A)}{P(A)P(C|A)+P(B)P(C|B)}$$
$$=\frac{\frac{60}{100}\cdot\frac{1}{100}}{\frac{60}{100}\cdot\frac{1}{100}+\frac{40}{100}\cdot\frac{2}{100}}=\frac{3}{7}.$$

11. 在区间(0,1)中随机地取两个数，则事件“两数之和小于 $\frac{6}{5}$”的概率为________.

答案是：$\frac{17}{25}$.

**分析** 用 $X$ 和 $Y$ 分别表示随机抽取的两个数，则
$$0<X<1,\quad 0<Y<1.$$
$X$,$Y$ 取值的所有可能结果（即样本点全体）对应的集合为以 1 为

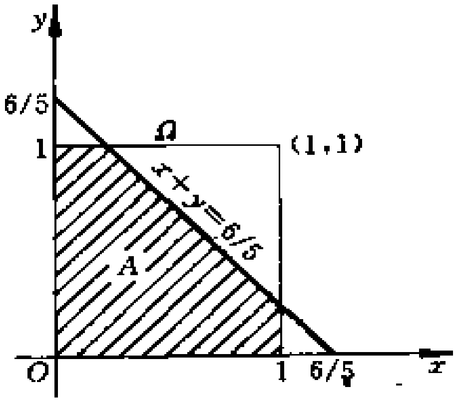


图 1.4

边长的正方形 $\Omega$，其面积为 1，事件"$X+Y\leqslant\frac{6}{5}$"对应图 1.4 中阴影部分 $A$，$A$ 的面积为

$$1-\frac{1}{2}\left(\frac{4}{5}\right)^2=\frac{17}{25}.$$

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

答案是：$\frac{1}{2}+\frac{1}{\pi}$.

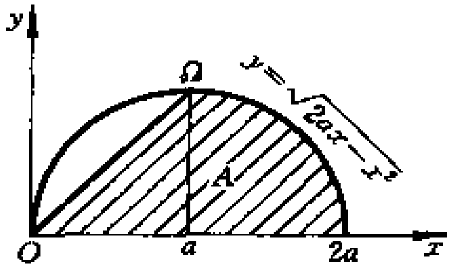


图 1.5

**分析** 半圆 $0<y<\sqrt{2ax-x^2}$ 也即样本空间 $\Omega$ 的面积为 $m(\Omega)=\frac{1}{2}\pi a^2$，所求事件对图 1.5 中阴影部分即区域 $A$ 的面积为 $m(A)=\frac{1}{2}a^2+\frac{\pi}{4}a^2$，故得所求事件概率为

$$P(A)=\frac{m(A)}{m(\Omega)}=\frac{\frac{1}{2}a^2+\frac{\pi}{4}a^2}{\frac{1}{2}\pi a^2}=\frac{1}{2}+\frac{1}{\pi}.$$

13. 袋中有 50 个乒乓球，其中 20 个是黄球，30 个是白球，今有两人依次随机地从袋中各取一球，取后不放回，则第二个人取得黄球的概率是________.

答案是：$\frac{2}{5}$.

**分析** 设 $A_i$ 代表"第 $i$ 个人取得黄球"，$i=1,2$；有

$$A_2=A_2\Omega=(A_1+\overline{A}_1)A_2=A_1A_2+\overline{A}_1A_2.$$

因此，

$$P(A_2)=P(A_1A_2+\overline{A}_1A_2)$$
$$=P(A_1)P(A_2|A_1)+P(\overline{A}_1)P(A_2|\overline{A}_1)$$
$$=\frac{2}{5}\times\frac{19}{49}+\frac{3}{5}\times\frac{20}{49}=\frac{2}{5}.$$

14. 设两两相互独立的三事件 $A,B$ 和 $C$ 满足条件：$ABC=\varnothing$，$P(A)=P(B)=P(C)<\frac{1}{2}$，且已知 $P(A\cup B\cup C)=\frac{9}{16}$，则 $P(A)=$________.

答案是：$\frac{1}{4}$.

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

由于 $ABC=\varnothing$，$P(A)=P(B)=P(C)$，并且 $A,B,C$ 两两独立. 因此，

$$P(A\cup B\cup C)=3P(A)-3(P(A))^2=\frac{9}{16}.$$

解得 $P(A)=\frac{1}{4}$ 或 $\frac{3}{4}$. 又由于 $P(A)<\frac{1}{2}$，所以 $P(A)=\frac{1}{4}$.

15. 假设 $P(A)=0.4$，$P(A\cup B)=0.7$，那么

(1) 若 $A$ 与 $B$ 互不相容，则 $P(B)=$________；

(2) 若 $A$ 与 $B$ 相互独立，则 $P(B)=$________.

答案是：0.3，0.5.

**分析** (1) $P(B)=P(A\cup B)-P(A)=0.7-0.4=0.3$.

(2) 由 
$$P(A\cup B)=P(A)+P(B)-P(AB)$$
$$=P(A)+P(B)-P(A)P(B),$$

得
$$P(B)=[P(A\cup B)-P(A)]/[1-P(A)]$$
$$=[0.7-0.4]/[1-0.4]=0.5.$$

16. 一射手对同一目标独立地进行四次射击，若至少命中一次的概率为 $\frac{80}{81}$，则该射手的命中率为________.

答案是：$\frac{2}{3}$.

**分析** 设命中率为 $p(0<p<1)$，则至少命中一次概率为 $1-(1-p)^4$，由 $1-(1-p)^4=\frac{80}{81}$，解得 $p=\frac{2}{3}$.

17. 设 $A,B$ 为随机事件，$P(A)=0.7$，$P(A-B)=0.3$，则 $P(\overline{AB})=$________.

答案是：0.6.

**分析** 由 $P(A-B)=P(A-AB)=P(A)-P(AB)=0.3$ 得

$$P(AB)=P(A)-0.3=0.7-0.3=0.4,$$

故 $P(\overline{AB})=0.6$.

18. 假设一批产品中一、二、三等品各占 60%，30%，10%，从中随意取出一件，结果不是三等品，则取到的是一等品的概率为________.

答案是：$\frac{2}{3}$.

**分析** 用 $A_i$ 代表事件"取出的产品为第 $i$ 等品"，$i=1,2,3$. 则 $A_1,A_2,A_3$ 互不相容，所求概率为

$$P(A_1|A_1\cup A_2)=\frac{P(A_1)}{P(A_1\cup A_2)}=\frac{P(A_1)}{P(A_1)+P(A_2)}$$

$$=\frac{0.6}{0.6+0.3}=\frac{2}{3}.$$

19. 一实习生用同一台机器接连独立地制造 3 个同种零件，第 $i$ 个零件是不合格品的概率 $p_i=\frac{1}{i+1}$ $(i=1,2,3)$，以 $X$ 表示 3 个零件中合格品的个数，则 $P\{X=2\}=$________.

答案是：$\frac{11}{24}$.

**分析** 用 $A_i$ 表示事件"第 $i$ 个零件是合格品"，则 $P(\overline{A}_i)=\frac{1}{i+1}$，$P(A_i)=1-\frac{1}{i+1}=\frac{i}{i+1}$，所求概率

$$P\{X=2\}=P(\overline{A}_1A_2A_3)+P(A_1\overline{A}_2A_3)+P(A_1A_2\overline{A}_3)$$

$$= P(\overline{A}_1)P(A_2)P(A_3) + P(A_1)P(\overline{A}_2)P(A_3)$$
$$+ P(A_1)P(A_2)P(\overline{A}_3)$$
$$= \frac{1}{2}\cdot\frac{2}{3}\cdot\frac{3}{4} + \frac{1}{2}\cdot\frac{1}{3}\cdot\frac{3}{4} + \frac{1}{2}\cdot\frac{2}{3}\cdot\frac{1}{4} = \frac{11}{24}.$$

20. 设 10 件产品中有 4 件不合格品，从中任取两件，已知两件中有一件是不合格品，则另一件也是不合格品的概率为________.

答案是：$\frac{1}{5}$.

**分析** 用 $A,B$ 分别代表取出的第 1 和第 2 件为正品，则所求概率为

$$P(\overline{A}\cdot\overline{B}|\overline{A}\cup\overline{B}) = \frac{P(\overline{A}\cdot\overline{B})}{P(\overline{A}\cup\overline{B})} = \frac{P(\overline{A}\cdot\overline{B})}{1-P(AB)}$$

$$= \frac{\frac{A_4^2}{A_{10}^2}}{\left[1-\frac{A_6^2}{A_{10}^2}\right]} = \frac{\frac{4\times3}{10\times9}}{\left[1-\frac{6\times5}{10\times9}\right]} = \frac{1}{5}.$$

21. 设 $A,B$ 是任意两个随机事件，则 $P\{(\overline{A}+B)(A+B)(\overline{A}+\overline{B})(A+\overline{B})\}=$________.

答案是：0.

**分析** $(\overline{A}+B)(A+B)(\overline{A}+\overline{B})(A+\overline{B})$

$$=[(\overline{A}+B)(A+B)]\cdot[\overline{(\overline{A}+B)(A+B)}]=\varnothing.$$

**(二) 选择题**

1. 若二事件 $A$ 和 $B$ 同时出现的概率 $P(AB)=0$，则(　　).

(A) $A$ 和 $B$ 不相容(相斥)；　(B) $AB$ 是不可能事件；

(C) $AB$ 未必是不可能事件；(D) $P(A)=0$ 或 $P(B)=0$.

答案是：C.

2. 对于任意二事件 $A$ 和 $B$，有 $P(A-B)=$(　　).

(A) $P(A)-P(B)$；　(B) $P(A)-P(B)+P(AB)$；

(C) $P(A)-P(AB)$；(D) $P(A)+P(\overline{B})+P(\overline{B})-P(A\overline{B})$.

答案是：C.

3. 以 $A$ 表示事件"甲种产品畅销，乙种产品滞销"，则其对应事件 $\overline{A}$ 为(　　).

(A)"甲种产品滞销，乙种产品畅销"；

(B)"甲、乙两种产品均畅销"；　　(C)"甲种产品滞销"；

(D)"甲种产品滞销或乙种产品畅销".

答案是：D.

**分析**　用 $A_1$ 表示甲产品畅销，$A_2$ 表示乙产品畅销，则 $A=A_1\overline{A}_2$，从而 $\overline{A}=\overline{A_1\overline{A}_2}=\overline{A}_1\cup A_2$.

4. 设 $A,B$ 为两随机事件，且 $B\subset A$，则下列式子正确的是(　　).

(A) $P(A+B)=P(A)$；　(B) $P(AB)=P(A)$；

(C) $P(B|A)=P(B)$；　(D) $P(B-A)=P(B)-P(A)$.

答案是：A.

**分析**　若 $B\subset A$，则 $A+B=A, AB=B$.

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

答案是：D.

**分析**　因为 $A$ 与 $B$ 不相容(即 $AB=\varnothing$)，所以

$$A-B=A-AB=A.$$

6. 两个随机变量$X$和$Y$，若$E(XY)=E(X)\cdot E(Y)$，则(　　).

(A) $D(XY)=D(X)\cdot D(Y)$；(B) $D(X+Y)=D(X)+D(Y)$；

(C) $X$ 和 $Y$ 独立；　　(D) $X$ 和 $Y$ 不独立.

答案是：B.

**分析**　$X$ 与 $Y$ 独立$\Rightarrow X$ 与 $Y$ 互不相关；

$$E(XY)=E(X)\cdot E(Y)\Leftrightarrow X \text{ 与 } Y \text{ 互不相关} \Leftrightarrow D(X+Y)=D(X)+D(Y).$$

7. 设当事件 $A$ 与 $B$ 同时发生时，事件 $C$ 必发生，则(　　).

(A) $P(C)\leqslant P(A)+P(B)-1$；

(B) $P(C)\geqslant P(A)+P(B)-1$；

(C) $P(C)=P(AB)$；　　　(D) $P(C)=P(A\cup B)$.

答案是：B.

**分析**　$P(C)\geqslant P(AB)=P(A)+P(B)-P(A\cup B)$

$$\geqslant P(A)+P(B)-1.$$

8. 设 $0<P(A)<1$，$0<P(B)<1$，$P(A|B)+P(\overline{A}|\overline{B})=1$，则(　　).

(A) 事件 $A$ 和 $B$ 互斥；　　(B) 事件 $A$ 和 $B$ 对立；

(C) 事件 $A$ 和 $B$ 不独立；　(D) 事件 $A$ 和 $B$ 相互独立.

答案是：D.

**分析**

$$P(A|B)+P(\overline{A}|\overline{B})=1$$
$$\Rightarrow P(A|B)=1-P(\overline{A}|\overline{B})=P(A|\overline{B})$$
$$\Rightarrow \frac{P(AB)}{P(B)}=\frac{P(A\overline{B})}{P(\overline{B})}=\frac{P(A)-P(AB)}{1-P(B)}$$
$$\Rightarrow P(AB)[1-P(B)]=[P(A)-P(AB)]P(B)$$
$$\Rightarrow P(AB)=P(A)P(B).$$

9. 已知 $0<P(B)<1$ 且 $P\{(A_1+A_2)|B\}=P(A_1|B)+P(A_2|B)$，则下列选项成立的是(　　).

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

(B) $P(A_1B+A_2B)=P(A_1B)+P(A_2B)$；

(C) $P(A_1+A_2)=P(A_1|B)+P(A_2|B)$；

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

答案是：B.

**分析**　将 $P\{(A_1+A_2)|B\}=P(A_1|B)+P(A_2|B)$ 两边同乘以 $P(B)$ 即得(B)式.

10. 设 $A,B$ 为任意两个事件且 $A\subset B$，$P(B)>0$，则下列选项必然成立的是(　　).

(A) $P(A)<P(A|B)$;　　　(B) $P(A)\leqslant P(A|B)$;

(C) $P(A)>P(A|B)$;　　　(D) $P(A)\geqslant P(A|B)$.

答案是：B.

**分析** $P(A)=P(AB)=P(B)P(A|B)\leqslant P(A|B)$.

11. 设 $A$、$B$ 是两个随机事件，且 $0<P(A)<1, P(B)>0$, $P(B|A)=P(B|\overline{A})$，则必有(　　).

(A) $P(A|B)=P(\overline{A}|B)$;　　(B) $P(A|B)\neq P(\overline{A}|B)$;

(C) $P(AB)=P(A)P(B)$;　　(D) $P(AB)\neq P(A)P(B)$.

答案是：C.

**分析** 由全概率公式及题设有

$$\begin{aligned}P(B)&=P(A)P(B|A)+P(\overline{A})P(B|\overline{A})\\&=[P(A)+P(\overline{A})]P(B|A)=P(B|A),\end{aligned}$$

于是 $P(AB)=P(A)P(B|A)=P(A)P(B)$.

（三）解答题

1. 设 $A, B$ 为任意两个事件，求证

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

**证** 左 $=1-P(\overline{AB})=1-P(\overline{A}\cup\overline{B})$

$=1-[P(\overline{A})+P(\overline{B})-P(\overline{A}\overline{B})]=$ 右.

2. 设 $P(A)=p, P(B)=q, P(AB)=r$，求下列各事件的概率：$P(\overline{A}\cup\overline{B})$, $P(\overline{A}B)$, $P(\overline{A}\cup B)$, $P(\overline{A}\,\overline{B})$.

**解** $P(\overline{A}\cup\overline{B})=P(\overline{AB})=1-P(AB)=1-r$;

$P(\overline{A}B)=P(B-AB)=P(B)-P(AB)=q-r$;

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

$=1-P(A)+P(B)-P(\overline{A}B)=1-p+r$;

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

$=1-[P(A)+P(B)-P(AB)]=1-p-q+r$.

3. 从一副扑克牌的 13 张梅花中，有放回地取 3 次，求三张都不同号的概率.

**解** 这是一个古典概型问题. 设 $A=\{$三张都不同号$\}$. 由题

意，有 $n=13^3$，$m=P_{13}^3$，则

$$P(A)=\frac{m}{n}=\frac{132}{169}.$$

4. 从 5 副不同的手套中任取 4 只，求这 4 只都不配对的概率.

**解** 这是一个古典概型问题. 设 $A=\{$这 4 只都不配对$\}$. 考虑用组合数及乘法原理计算，由题意，有

$$n=C_{10}^4,\quad m=C_5^4\cdot C_2^1\cdot C_2^1\cdot C_2^1\cdot C_2^1,$$

所以
$$P(A)=\frac{8}{21}.$$

5. 一批产品共 100 件，对产品进行不放回抽样检查，整批产品不合格的条件是：在被检查的 5 件产品中至少有一件是废品. 如果该批产品中有 5 件是废品，求该批产品因不合格而被拒绝接收的概率.

**解 方法 1** 设事件 $A_i$ 为被检查的第 $i$ 件产品是废品（$i=1,2,3,4,5$），设事件 $B$ 为该批产品被拒绝接收，则

$$B=A_1\cup A_2\cup A_3\cup A_4\cup A_5.$$

于是

$$\begin{aligned}P(B)&=1-P(\overline{A}_1\overline{A}_2\overline{A}_3\overline{A}_4\overline{A}_5)\\&=1-P(\overline{A}_1)P(\overline{A}_2|\overline{A}_1)P(\overline{A}_3|\overline{A}_1\overline{A}_2)\\&\quad\times P(\overline{A}_4|\overline{A}_1\overline{A}_2\overline{A}_3)P(\overline{A}_5|\overline{A}_1\overline{A}_2\overline{A}_3\overline{A}_4).\end{aligned}$$

而

$$P(\overline{A}_1)=1-P(A_1)=\frac{95}{100},\quad P(\overline{A}_2|\overline{A}_1)=\frac{94}{99},$$

$$P(\overline{A}_3|\overline{A}_1\overline{A}_2)=\frac{93}{98},$$

$$P(\overline{A}_4|\overline{A}_1\overline{A}_2\overline{A}_3)=\frac{92}{97},\quad P(\overline{A}_5|\overline{A}_1\overline{A}_2\overline{A}_3\overline{A}_4)=\frac{91}{96},$$

因此
$$P(B)=1-\frac{95}{100}\cdot\frac{94}{99}\cdot\frac{93}{98}\cdot\frac{92}{97}\cdot\frac{91}{96}\approx 0.23.$$

**方法 2** 本题用古典概型也可以计算，先求 $P(\overline{B})$，有

$$P(\overline{B})=\frac{P_{95}^{5}}{P_{100}^{5}}\quad\left(\text{或}\frac{C_{95}^{5}}{C_{100}^{5}}\right),$$

再由 $P(B)=1-P(\overline{B})\approx 0.23$ 即可.

6. 设事件 $AB$ 发生,则事件 $C$ 一定发生. 证明

$$P(A)+P(B)-P(C)\leqslant 1.$$

**证** 由概率基本性质,因为 $AB\subset C$,有 $P(AB)\leqslant P(C)$. 考虑到

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

以及 $P(AB)+P(A\overline{B})+P(\overline{A}B)=1-P(\overline{A}\,\overline{B})$有

$$\begin{aligned}
&P(A)+P(B)-P(C)\\
&\leqslant[P(AB)+P(A\overline{B})]+[P(AB)+P(\overline{A}B)]-P(AB)\\
&=P(AB)+P(A\overline{B})+P(\overline{A}B)\\
&=1-P(\overline{A}\,\overline{B})\leqslant 1.
\end{aligned}$$

7. 袋内放有两个伍分的,三个贰分和五个壹分的钱币,任取其中五个,求钱额总数超过壹角的概率.

**解** 这是一个古典概型问题. 设 $A=\{$取 5 个钱币钱额超 1 角$\}$,于是有

$$n=C_{10}^{5}.$$

由题意可知,当取二个五分币,其余的三个可任取,其种数为:

$$C_2^2C_3^3+C_2^2C_3^2C_5^1+C_2^2C_3^1C_5^2+C_2^2C_5^3.$$

而当取一个 5 分币,2 分币至少要取 2 个,其种数为:

$$C_2^1C_3^3C_5^1+C_2^1C_3^2C_5^2.$$

因此有利于事件 $A$ 的基本事件总数:

$$\begin{aligned}
m=&C_2^2C_3^3+C_2^2C_3^2C_5^1+C_2^2C_3^1C_5^2\\
&+C_2^2C_5^3+C_2^1C_3^3C_5^1+C_2^1C_3^2C_5^2=126,
\end{aligned}$$

故
$$P(A)=\frac{126}{C_{10}^{5}}=\frac{1}{2}.$$

8. 设有甲、乙两名射手轮流独立地对同一目标射击,甲的命中率为 $p_1$,乙的命中率为 $p_2$. 甲先射,谁先命中谁得胜,试分别求

甲获胜的概率和乙获胜的概率.

**解** 以 $A$ 表示甲获胜，$A_{2k+1}$表示前 $2k$ 次均未命中，第 $2k+1$ 次甲命中($k=0,1,2,\cdots$)，那么

$$A=\bigcup_{k=0}^{\infty}A_{2k+1},$$

由于 $A_1,A_3,A_5,\cdots$两两互不相容，故甲获胜的概率为

$$P(A)=P\Big(\bigcup_{k=0}^{\infty}A_{2k+1}\Big)=\sum_{k=0}^{\infty}P(A_{2k+1})=\sum_{k=0}^{\infty}(1-p_1)^k(1-p_2)^kp_1$$

$$=\frac{p_1}{1-(1-p_1)(1-p_2)}=\frac{p_1}{p_1+p_2-p_1p_2}.$$

同理，以 $B$ 表示乙获胜，$B_{2k}$表示前 $2k-1$ 次均未命中，第 $2k$ 次乙命中($k=1,2,\cdots$)，有

$$P(B)=P\Big(\bigcup_{k=1}^{\infty}B_{2k}\Big)=\sum_{k=1}^{\infty}P(B_{2k})=\sum_{k=1}^{\infty}(1-p_1)^k(1-p_2)^{k-1}p_2$$

$$=\frac{(1-p_1)p_2}{1-(1-p_1)(1-p_2)}=\frac{(1-p_1)p_2}{p_1+p_2-p_1p_2}.$$

9.（配对问题） 某人写了 $n$ 封不同的信，欲寄往 $n$ 个不同的地址. 现将这 $n$ 封信随意地插入 $n$ 只具有不同通信地址的信封里，求至少有一封信插对信封的概率.

**解** 令 $A=\{$至少有一封信插对信封$\}$，$A_i=\{$第 $i$ 封信插对信封$\}$ ($i=1,2,\cdots,n$)，则 $A=\bigcup\limits_{i=1}^{n}A_i$. 再由 $n$ 个事件的加法定理

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

易知，$n$ 封信插入 $n$ 只不同地址的信封，共有 $n!$ 种插法. 而有 $k$ 封信插对信封意味着这 $k$ 封信插入固定的 $k$ 只信封后，其他 $n-k$ 封信可随意插入剩下的 $n-k$ 只信封中，这将有$(n-k)!$ 种插法. 故

$$P(A_{i1}A_{i2}\cdots A_{ik}) = \frac{(n-k)!}{n!},$$

而在和式 $\sum\limits_{i_1<i_2<\cdots<i_k} P(A_{i1}A_{i2}\cdots A_{ik})$中共有 $C_n^k=\frac{n!}{k!\,(n-k)!}$项，故

$$\sum_{i_1<i_2<\cdots<i_k} P(A_{i1}A_{i2}\cdots A_{ik}) = C_n^k\frac{(n-k)!}{n!} = \frac{1}{k!},$$

则

$$P(A) = 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n-1}\cdot\frac{1}{n!},$$

这即为所求的概率．当 $n$ 较大时，在 $e^x$ 的级数展开式中

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots,$$

以 $x=-1$ 代入其中得

$$e^{-1} = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n\frac{1}{n!} + \cdots,$$

由此，我们近似地有

$$P(A) \approx 1 - e^{-1} \approx 0.632.$$

10．在第一个箱中有 10 个球，其中 8 个是白的；在第二个箱中有 20 个球，其中 4 个是白的．现从每个箱中任取一球，然后从这两球中任取一球，取到白球的概率是多少？

**解**　这是一个全概率公式问题．

**方法 1**　设 $C$：取到白球；$A_i$：从第 $i$ 个箱中取到白球($i=1,2$)，有

$B_0=\{$取到黑黑$\}=\overline{A}_1\overline{A}_2$，$P(B_0)=(2/10)\times(16/20)=4/25$；

$B_1=\{$取到白黑$\}=A_1A_2$，$P(B_1)=(8/10)\times(16/20)=16/25$；

$B_2=\{$取到黑白$\}=\overline{A}_1A_2$，$P(B_2)=(2/10)\times(4/20)=1/25$；

$B_3=\{$取到白白$\}=A_1A_2$，$P(B_3)=(8/10)\times(4/20)=4/25$.

$P(C|B_0)=0$，$P(C|B_1)=1/2$，$P(C|B_2)=1/2$，$P(C|B_3)=1$.

因此　$P(C)=0\times(4/25)+(1/2)\times(16/25)$

$+(1/2)\times(1/25)+1\times(4/25)=1/2.$

**方法 2** 设 $A_i$：已取出的球来自第 $i$ 个箱；$P(A_i)=1/2,(i=1,2)$. $B$：取到白球；则 $P(B|A_1)=8/10$；$P(B|A_2)=4/20$；由全概率公式有

$$
\begin{aligned}
P(B) &= P(A_1)P(B|A_1) + P(A_2)P(B|A_2) \\
&= (1/2) \times (8/10) + (1/2) \times (4/20) = 1/2.
\end{aligned}
$$

11. 有一批产品，其中正品有 $n$ 个，次品有 $m$ 个，先从这批产品中任意取出 $l$ 个（不知其中的次品数），然后再从剩下的产品中任取一个恰为正品的概率是多少？

**解** 这是一个全概率公式问题.

设 $A_k$：前 $l$ 次中恰有 $k$ 个正品，$k=q,q+1,q+2,\cdots,p$，其中

$$q=\max(l-m,0),\quad p=\min(n,l).$$

又设 $B$：第 $l+1$ 个恰为正品，从而有

$$A_q+A_{q+1}+A_{q+2}+\cdots+A_p=\Omega,$$

$$P(A_k)=\frac{C_n^k C_m^{l-k}}{C_{m+n}^l},$$

而
$$P(B|A_k)=\frac{C_{n-k}^1}{C_{m+n-l}^1}=\frac{n-k}{m+n-l},$$

由全概率公式，得到

$$P(B)=\sum_{k=q}^{p}P(A_k)P(B|A_k)=\frac{n}{m+n}.$$

12. 在对某厂的产品进行重复抽样检查时，从抽取的 200 件中发现有 4 件次品，问能否相信该厂产品的次品率不超过0.005.

**解** 如果该厂产品的次品率为 0.005，由二项概型可知，这 200 件样品中出现大于或等于 4 件次品的概率为

$$
\begin{aligned}
P_{200}(\mu \geqslant 4) &= 1 - P_{200}(\mu < 4) \\
&= 1 - \sum_{k=0}^{3} C_{200}^k (0.005)^k (1-0.005)^{200-k} \\
&\approx 0.0190.
\end{aligned}
$$

而当次品率小于 0.005 时，这个概率还要小，这说明在我们进行的

一次抽取（一共抽取 200 个样品）的试验中，一个小概率的事件竟发生了. 因此，我们可以说该厂产品的次品率不超过 0.005 是不可信的.

13. 设平面区域 $D_1$ 是由 $x=1,y=0,y=x$ 所围成，今向 $D_1$ 内随机地投入 10 个点，求这 10 个点中至少有 2 个点落在由曲线 $y=x^2$ 与 $y=x$ 所围成的区域 $D$ 内的概率.

**解** 设 $A$ 表示"任投一点落在区域 $D$ 内"，则有

$$P(A)=\frac{L(A)}{L(\Omega)}=\frac{\frac{1}{2}-\frac{1}{3}}{\frac{1}{2}}=\frac{1}{3}=p.$$

于是，由二项概型可知这 10 个点中至少有 2 个点落在区域 $D$ 内的概率为

$$\begin{aligned}P_{10}(\mu\geqslant 2)&=1-P_{10}(\mu=0)-P_{10}(\mu=1)\\&=1-C_{10}^{0}p(1-p)^{10}-C_{10}^{1}p^{1}(1-p)^{9}\\&=1-\left(\frac{2}{3}\right)^{10}-C_{10}^{1}\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^{9}.\end{aligned}$$

14. 某类电灯泡使用时数在 1000 个小时以上的概率为 0.2，求三个灯泡在使用 1000 小时以后最多只坏一个的概率.

**解** 利用二项概型，有

$$\begin{aligned}P_3\{\mu\leqslant 1\}&=P_3\{\mu=0\}+P_3\{\mu=1\}\\&=C_3^0(0.8)^0(0.2)^3+C_3^1(0.8)^1(0.2)^2\\&=0.104.\end{aligned}$$

15. 假设有两箱同种零件：第一箱内装 50 件，其中 10 件一等品；第二箱内装 30 件，其中 18 件一等品. 现从两箱中随意挑出一箱，然后从该箱中先后随机取两个零件（取出的零件均不放回）. 试求：

(1) 先取出的零件是一等品的概率 $p$；

(2) 在先取出的零件是一等品的条件下，第二次取出的零件仍然是一等品的条件概率 $q$.

**解** 引进下列事件：

$H_i=\{$被挑出的是第 $i$ 箱$\}$， $i=1,2$；

$A_j=\{$第 $j$ 次取出的零件是一等品$\}$， $j=1,2$，

那么，由题设知

$$P(H_1)=P(H_2)=\frac{1}{2};\ P(A_1|H_1)=\frac{1}{5},\ P(A_1|H_2)=\frac{3}{5}.$$

(1) 由全概率公式

$$p=P(A_1)=P(H_1)P(A_1|H_1)+P(H_2)P(A_1|H_2)$$
$$=\frac{1}{2}\cdot\frac{1}{5}+\frac{1}{2}\cdot\frac{3}{5}=\frac{2}{5}.$$

(2) 由条件概率的定义和全概率公式

$$q=P(A_2|A_1)=\frac{P(A_1A_2)}{P(A_1)}$$
$$=\frac{1}{P(A_1)}[P(H_1)P(A_1A_2|H_1)+P(H_2)P(A_1A_2|H_2)]$$
$$=\frac{5}{2}\left[\frac{1}{2}\cdot\frac{10\times 9}{50\times 49}+\frac{1}{2}\cdot\frac{18\times 17}{30\times 29}\right]$$
$$=\frac{1}{4}\left[\frac{9}{49}+\frac{51}{29}\right]=0.48557.$$

16. 从 0,1,2,…,9 等十个数字中任意选出三个不同的数字，试求下列事件的概率：

$A_1=\{$三个数字中不含 0 和 5$\}$；

$A_2=\{$三个数字中不含 0 或 5$\}$；

$A_3=\{$三个数字中含 0 但不含 5$\}$.

**解** $P(A_1)=\dfrac{C_8^3}{C_{10}^3}=\dfrac{7}{15}$， $P(A_2)=\dfrac{2C_9^3-C_8^3}{C_{10}^3}=\dfrac{14}{15}$，

$$P(A_3)=\frac{C_9^2-C_8^1}{C_{10}^3}=\frac{7}{30}\ \text{或}\ P(A_3)=\frac{C_8^2}{C_{10}^3}=\frac{7}{30}.$$

17. 假设一厂家生产的每台仪器，以概率 0.70 可以直接出厂；以概率 0.30 需进一步调试，经调试后以概率 0.80 可以出厂；以概率 0.20 定为不合格品不能出厂. 现该厂新生产了 $n(n\geqslant 2)$台

仪器(假设各台仪器的生产过程相互独立),求

(1) 全部能出厂的概率 $\alpha$;

(2) 其中恰好有两件不能出厂的概率 $\beta$;

(3) 其中至少有两件不能出厂的概率 $\theta$.

**解** 对于新生产的每台仪器,引进事件 $A=\{$仪器需进一步调试$\}$,$B=\{$仪器可以出厂$\}$,则任一仪器可出厂概率为

$$P(B)=P(A)P(B|A)+P(\overline{A})P(B|\overline{A}))$$
$$=0.3\times0.8+0.7\times1=0.94.$$

用 $X$ 代表所生产的 $n$ 台仪器中能出厂的台数,则 $X$ 为 $n$ 次独立试验(仪器出厂)的次数,服从参数为$(n,0.94)$的二项分布,因此

$$\alpha=P\{X=n\}=0.94^n,$$
$$\beta=P\{X=n-2\}=C_n^2(0.94)^{n-2}(0.06)^2,$$
$$\theta=P\{X\leqslant n-2\}=1-P\{X=n-1\}-P\{X=n\}$$
$$=1-n\cdot(0.94)^{n-1}(0.06)-0.94^n.$$

18. 甲袋中放有 5 只红球,10 只白球;乙袋中放有 5 只白球,10 只红球.今先从甲袋中任取一球放入乙袋,然后从乙袋中任取一球放回甲袋.求再从甲袋中任取两球全是红球的概率.

**解** $A_i=\{$从甲袋中任取一球放入乙袋,再从乙袋任取一球放回甲袋后甲袋中含有 $i$ 只红球$\}$, $i=4,5,6$.

$$B=\{\text{最后从甲袋中任取两球全是红球}\}.$$

$P(A_4)=P($先从甲袋中取出红球,后从乙袋中取出白球$)$

$$=\frac{5}{15}\cdot\frac{5}{16}=\frac{5}{48},$$

$P(A_5)=P($先从甲袋中取出红球,后从乙袋中取出红球$)$

$+P($先从甲袋中取出白球,后从乙袋中取出白球$)$

$$=\frac{5}{15}\cdot\frac{11}{16}+\frac{10}{15}\times\frac{6}{16}=\frac{23}{48},$$

$P(A_6)=P($先从甲袋中取出白球,后从乙袋中取出红球$)$

$$=\frac{10}{15}\cdot\frac{5}{16}=\frac{20}{48},$$

$$P(B|A_4)=C_4^2/C_{15}^2=\frac{6}{105},\ P(B|A_5)=C_5^2/C_{15}^2=\frac{10}{105},$$

$$P(B|A_6)=C_6^2/C_{15}^2=\frac{15}{105}.$$

于是由全概率公式得所求概率为

$$\begin{aligned}P(B)&=\sum_{i=4}^{6}P(A_i)P(B|A_i)\\&=\frac{5}{48}\cdot\frac{6}{105}+\frac{23}{48}\cdot\frac{10}{105}+\frac{20}{48}\cdot\frac{15}{105}=\frac{1}{9}.\end{aligned}$$

19. 设来自三个地区的各 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 $q$.

**解** 设 $H_i=\{$报名表是第 $i$ 区考生的$\}$ $(i=1,2,3)$，$A_j=\{$第 $j$ 次抽到的报名表是男生表$\}$ $(j=1,2)$，则

$$P(H_1)=P(H_2)=P(H_3)=\frac{1}{3};$$

$$P(A_1|H_1)=\frac{7}{10},\ P(A_1|H_2)=\frac{8}{15},\ P(A_1|H_3)=\frac{20}{25}.$$

(1) $$\begin{aligned}p=P(\overline{A}_1)&=\sum_{i=1}^{3}P(H_i)P(\overline{A}_1|H_i)\\&=\frac{1}{3}\left(\frac{3}{10}+\frac{7}{15}+\frac{5}{25}\right)=\frac{29}{90}.\end{aligned}$$

(2) 由全概率公式得

$$\begin{aligned}P(A_2|H_1)&=P(A_2(A_1\cup\overline{A}_1)|H_1)\\&=P(A_1A_2|H_1)+P(\overline{A}_1A_2|H_1)\\&=P(A_1|H_1)P(A_2|A_1H_1)\\&\quad+P(\overline{A}_1|H_1)P(A_2|\overline{A}_1H_1)\\&=\frac{7}{10}\cdot\frac{6}{9}+\frac{3}{10}\cdot\frac{7}{9}=\frac{63}{90}=\frac{7}{10},\end{aligned}$$

$$P(A_2|H_2)=P(A_1|H_2)P(A_2|A_1H_2)$$
$$+P(\overline{A}_1|H_2)P(A_2|\overline{A}_1H_2)$$
$$=\frac{8}{15}\cdot\frac{7}{14}+\frac{7}{15}\cdot\frac{8}{14}=\frac{8}{15},$$
$$P(A_2|H_3)=P(A_1|H_3)P(A_2|A_1H_3)$$
$$+P(\overline{A}_1|H_3)P(A_2|\overline{A}_1H_3)$$
$$=\frac{20}{25}\cdot\frac{19}{24}+\frac{5}{25}\cdot\frac{20}{24}=\frac{20}{25},$$
$$P(\overline{A}_1A_2|H_1)=P(\overline{A}_1|H_1)P(A_2|\overline{A}_1H_1)$$
$$=\frac{3}{10}\cdot\frac{7}{9}=\frac{7}{30},$$
$$P(\overline{A}_1A_2|H_2)=\frac{7}{15}\cdot\frac{8}{14}=\frac{8}{30},$$
$$P(\overline{A}_1A_2|H_3)=\frac{5}{25}\cdot\frac{20}{24}=\frac{5}{30},$$
$$P(A_2)=\sum_{i=1}^{3}P(H_i)P(A_2|H_i)=\frac{1}{3}\left[\frac{7}{10}+\frac{8}{15}+\frac{20}{25}\right]=\frac{61}{90},$$
$$P(\overline{A}_1A_2)=\sum_{i=1}^{3}P(H_i)P(\overline{A}_1A_2|H_i)=\frac{1}{3}\left[\frac{7}{30}+\frac{8}{30}+\frac{5}{30}\right]=\frac{2}{9},$$
因此 $$q=p(\overline{A}_1|A_2)=\frac{P(\overline{A}_1A_2)}{P(A_2)}=\frac{2}{9}\Big/\frac{61}{90}=\frac{20}{61}.$$

20. 设 $0<P(A)<1,0<P(B)<1,P(A|B)+P(\overline{A}|\overline{B})=1$. 问 $A$ 与 $B$ 是否独立?

**解** 因为 $P(A|B)+P(\overline{A}|\overline{B})=1$, 所以 $P(A|B)=1-P(\overline{A}|\overline{B})=P(A|\overline{B})$,即

$$\frac{P(AB)}{P(B)}=\frac{P(A\overline{B})}{P(\overline{B})},\quad P(AB)[1-P(B)]=P(B)P(A\overline{B}),$$
$$P(AB)=P(B)[P(AB)+P(A\overline{B})]=P(B)P(AB\cup A\overline{B})$$
$$=P(A(B\cup\overline{B}))P(B)=P(A)P(B).$$
故 $A$ 与 $B$ 相互独立.

21. 设甲、乙都有 $n$ 个硬币,全部掷完后分别计算掷出的正面

数，求甲、乙两人掷出的正面数相等的概率.

**解** 甲、乙二人掷出 $k$ 个正面的概率均为

$$P_n(k)=C_n^k\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{n-k}=C_n^k\left(\frac{1}{2}\right)^n,$$

由于甲、乙两人是独立投掷的，所以两人掷出正面数相等的概率为

$$\sum_{k=0}^{n}P_n(k)P_n(k)=\sum_{k=0}^{n}C_n^k\left(\frac{1}{2}\right)^nC_n^k\left(\frac{1}{2}\right)^n=\frac{1}{2^{2n}}\sum_{k=0}^{n}C_n^kC_n^{n-k}=C_{2n}^n/2^{2n}.$$

22. 设一昆虫产 $i$ 个卵的概率为 $\frac{\lambda^i}{i!}e^{-\lambda}(i=0,1,\cdots)$. 而每个卵能孵化为成虫的概率为 $p$，且各卵的孵化是相互独立的，试求这昆虫的下一代有 $k$ 只的概率.

**解** 引进事件：$B_i=\{$昆虫产 $i$ 个卵$\}(i=0,1,\cdots)$，$A=\{$昆虫的下一代有 $k$ 只$\}$，依题意：

$$P(B_i)=\frac{\lambda^ie^{-\lambda}}{i!},\quad i=0,1,\cdots,$$

$$P(A|B_i)=C_i^kp^k(1-p)^{i-k},\quad i=k,k+1,\cdots,$$

由全概率公式知

$$\begin{aligned}P(A)&=\sum_{i=k}^{\infty}P(B_i)P(A|B_i)=\sum_{i=k}^{\infty}\frac{\lambda^ie^{-\lambda}}{i!}C_i^kp^k(1-p)^{i-k}\\&=\sum_{i=k}^{\infty}\frac{\lambda^ie^{-\lambda}}{i!}\cdot\frac{i!}{k!(i-k)!}p^k(1-p)^{i-k}\\&=\frac{\lambda^kp^ke^{-\lambda}}{k!}\sum_{i=k}^{\infty}\frac{[(1-p)\lambda]^{i-k}}{(i-k)!}\\&\xlongequal{令 t=i-k}\frac{\lambda^kp^ke^{-\lambda}}{k!}\sum_{t=0}^{\infty}\frac{[(1-p)\lambda]^t}{t!}\\&=\frac{\lambda^kp^ke^{-\lambda}}{k!}e^{(1-p)\lambda}=\frac{(\lambda p)^k}{k!}e^{-\lambda p}.\end{aligned}$$

## 四、练习题

1. 写出下列各随机试验的样本空间：

(1) 甲、乙二人下棋一局，观察棋赛结果；

(2) 同时掷两颗骰子，记录其出现的点数；

(3) 某寻呼台单位时间内接到的寻呼次数；

(4) 任取一只灯泡，观察其寿命.

2. 证明：(1) $A-B=A\overline{B}$；　(2) $A-B=A-AB$.

3. 化简：(1) $(A\cup B)(A\cup\overline{B})$；　(2) $\overline{(\overline{AB}\cup C)\overline{AC}}$.

4. 已知 $A,B$ 两事件满足 $P(AB)=P(\overline{AB})$，且 $P(A)=p$，求 $P(B)$.

5. 从 0,1,…,9 十个数字中每次任取一个数，然后放回，共取 5 次，试求下列事件的概率：$A_1=\{5$ 个数字各不相同$\}$，$A_2=\{5$ 个数字中不含 0 和 $1\}$，$A_3=\{5$ 个数字中 1 恰好出现两次$\}$.

6. 甲、乙、丙三人独立地破译一种密码，他们能译出的概率分别是$\frac{1}{5}$，$\frac{1}{3}$，$\frac{1}{4}$，求能把这种密码译出的概率.

7. 设有十个数字 0,1,…,9，从中任取两个数字，求其和大于 10 的概率.

8. 在无线电通讯中，由于随机干扰，当发出信号为“·”时，收到信号为“·”，“—”及“不清”的概率分别为 0.7，0.1 及 0.2；当发出信号为“—”时，收到信号为“—”，“·”及“不清”的概率分别为 0.9，0 及 0.1. 如果在整个发报过程中，信号“·”及“—”出现的概率分别为 0.6 及 0.4. 求收到的信号分别为“·”和“—”的概率；当收到信号“不清”时，求发出信号分别是“·”及“—”的概率.

9. 设有一架长机两架僚机飞往某目的地进行轰炸，由于只有长机装有导航设备，因此僚机不能单独到达目的地，并在飞行途中要经过敌方高射炮防御地带，每机被击落的概率为 0.2，到达目的地后，各机独立轰炸，每机炸中目标的概率为 0.3，求目标被炸中的概率.

10. 设甲、乙两篮球运动员投篮命中率分别为 0.7 及 0.6，每人投篮 3 次，求两人进球数相等的概率.

## 习题答案与提示

**1.** (1) $\Omega_1=\{$甲胜乙负,乙胜甲负,和局$\}$.

(2) $\Omega_2=\{(i,j):\ i,j=1,2,\cdots,6\}$.

(3) $\Omega_3=\{0,1,2,\cdots\}$.　　(4) $\Omega_4=\{t:\ t\geqslant 0\}$.

**2.** 提示　(1) 设 $\omega\in A-B$,则 $\omega\in A$ 且 $\omega\bar{\in} B$,从而 $\omega\in A$ 且 $\omega\in\overline{B}$,故 $\omega\in A\overline{B}$,于是 $A-B\subset A\overline{B}$. 设 $\omega\in A\overline{B}$,那末 $\omega\in A$ 且 $\omega\in\overline{B}$,也即 $\omega\in A$ 且 $\omega\bar{\in} B$,故 $\omega\in A-B$,于是 $A\overline{B}\subset A-B$. 从而 $A-B=A\overline{B}$.

(2) 设 $\omega\in A-B$,那末 $\omega\in A$ 且 $\omega\bar{\in} B$,由 $\omega\bar{\in} B$ 知 $\omega\bar{\in} AB$,也即 $\omega\in A$ 且 $\omega\bar{\in} AB$,所以 $\omega\in A-AB$,于是 $A-B\subset A-AB$. 反之若 $\omega\in A-AB$,那末 $\omega\in A$ 且 $\omega\bar{\in} AB$,即 $\omega\bar{\in} B$,故有 $\omega\in A-B$,于是 $A-AB\subset A-B$. 从而 $A-B=A-AB$.

**3.** 提示　(1)
$$\begin{aligned}(A\cup B)(A\cup\overline{B})&=AA\cup A\overline{B}\cup BA\cup B\overline{B}\\&=A\cup A\overline{B}\cup BA\cup\varnothing=A\cup A\overline{B}\cup AB\\&=A\cup A(B\cup\overline{B})=A\cup A=A.\end{aligned}$$

(2) 多次运用对偶原理,便得到:
$$\begin{aligned}\overline{(\overline{A}\,\overline{B}\cup C)\,\overline{AC}}&=\overline{\overline{A}\,\overline{B}\cup C}\cup AC=\overline{\overline{A}\,\overline{B}}\,\overline{C}\cup AC\\&=(A\cup B)\overline{C}\cup AC=A\overline{C}\cup B\overline{C}\cup AC\\&=A\overline{C}\cup AC\cup B\overline{C}=A(\overline{C}\cup C)\cup B\overline{C}\\&=A\cup B\overline{C}.\end{aligned}$$

**4.** $P(B)=1-p$.

**5.** $P(A_1)=\dfrac{P_{10}^5}{10^5}$ (或 0.3024); $P(A_2)=\dfrac{8^5}{10^5}$ (或 0.32768);

$P(A_3)=\dfrac{C_5^2\times 9^3}{10^5}$ (或 0.0729).

**6.** $\dfrac{3}{5}$.　　**7.** $\dfrac{16}{45}$.

**8.** 0.42;0.42;0.75;0.25.　　**9.** 0.477.　　**10.** 0.32.

# 第二章　随机变量及其分布

## 一、考试要求

1. 理解随机变量及其分布的概念；

2. 掌握几种常见分布，会计算与随机变量有关的概率；

3. 理解分布函数的定义及性质，掌握随机变量函数的分布.

## 二、复习要点

### （一）重要概念及性质

**定义 2.1（随机变量）**　在条件 $S$ 下，随机试验的每一个可能的结果 $\omega$ 都用一个实数 $X=X(\omega)$ 来表示，且实数 $X$ 满足

1）$X$ 是由 $\omega$ 唯一确定；

2）对于任意给定的实数 $x$，事件 $\{X\leqslant x\}$ 都是有概率的，则称 $X$ 为**一随机变量**，一般用英文大写字母 $X,Y,Z$ 等表示.

随机变量一般可分为离散型和非离散型两大类. 非离散型又可分为连续型和混合型.

（1）离散型随机变量

若随机变量 $X$ 的取值为有限个或可列个（即称为至多可列个），则称 $X$ 为**离散型随机变量**.

设离散型随机变量 $X$ 的取值为 $x_1,x_2,\cdots,x_k,\cdots$，$X$ 取各个可能值的概率为

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

我们称上式为离散型随机变量 $X$ 的**概率分布**或**分布律**. 随机变量 $X$ 的概率分布可以用列表的形式给出：

| $X$ | $x_1$ | $x_2$ | $\cdots$ | $x_k$ | $\cdots$ |
|---|---|---|---|---|---|
| $P(X=x_k)$ | $p_1$ | $p_2$ | $\cdots$ | $p_k$ | $\cdots$ |

这种表格被称为 $X$ 的**分布列**；它还可以用矩阵的形式来表示：

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_k & \cdots \\ p_1 & p_2 & \cdots & p_k & \cdots \end{pmatrix},$$

它被称为 $X$ 的**分布阵**. 这里 $p_k$ 有下列性质：

1）$p_k \geqslant 0$，$k=1,2,\cdots$；

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

（2）连续型随机变量

对于随机变量 $X$，如果存在非负可积函数 $p(x)(-\infty<x<+\infty)$，使得 $X$ 取值于任一区间$(a,b)$的概率为

$$P\{a<X<b\}=\int_a^b p(x)\mathrm{d}x,$$

则称 $X$ 为**连续型随机变量**；并称 $p(x)$为 $X$ 的**分布密度函数**，有时简称为**分布密度**.

分布密度 $p(x)$有下列性质：

1）$p(x) \geqslant 0, -\infty<x<+\infty$；

2）$\int_{-\infty}^{+\infty} p(x)\mathrm{d}x = P\{-\infty<x<+\infty\}=P(\Omega)=1$.

**定义 2.2（分布函数）** 设 $X$ 为一随机变量，$x$ 是任意实数，称函数

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

为 $X$ 的**分布函数**.

分布函数是一个以全体实数为其定义域，以事件$\{\omega \mid -\infty< X(\omega) \leqslant x\}$的概率为函数值的一个实值函数. 分布函数 $F(x)$具有以下的基本性质：

1）$0 \leqslant F(x) \leqslant 1$；

2）$F(x)$是非减函数；

3) $F(x)$是右连续的；

4) $\lim\limits_{x \to -\infty} F(x)=0$，$\lim\limits_{x \to +\infty} F(x)=1$.

由分布函数的定义，显然有：

$$P\{a < X \leqslant b\} = P\{X \leqslant b\} - P\{X \leqslant a\} = F(b) - F(a),$$

$$P\{X < x_0\} = \lim_{x \to x_0 - 0} P\{X \leqslant x\} = F(x_0 - 0),$$

$$P\{X = x_0\} = P\{X \leqslant x_0\} - P\{X < x_0\} = F(x_0) - F(x_0 - 0).$$

当 $P\{X=x_0\}=0$ 时，$x_0$ 为 $F(x)$的连续点，当 $P\{X=x_0\}\neq 0$ 时，$x_0$ 为 $F(x)$的间断点，$P\{X=x_0\}$为 $X$ 在点 $x_0$ 的跳跃值.

$$P\{X > x_0\} = 1 - P\{X \leqslant x_0\} = 1 - F(x_0),$$

$$P\{X \geqslant x_0\} = 1 - P\{X < x_0\} = 1 - F(x_0 - 0).$$

(1) 离散型随机变量的分布函数

设离散型随机变量 $X$ 的概率分布为

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

由分布函数的定义可知，其分布函数

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

即 $F(x)$是 $X$ 取小于或等于 $x$ 的所有可能值的概率之和. 对于 $X$ 的取值为 $x_1 < x_2 < \cdots < x_n < \cdots$ 时，其分布函数可以写成分段函数形式

$$F(x) = \begin{cases} 0, & x < x_1, \\ p_1, & x_1 \leqslant x < x_2, \\ p_1 + p_2, & x_2 \leqslant x < x_3, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots. \end{cases}$$

不难看出，这个分段函数的分点就是在 $X$ 的概率分布中取正概率的点.

**例 1** 设 $X$ 的分布为：

$$P\{X=1\}=p,\quad P\{X=0\}=1-p,$$

则

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

**例 2** 一袋中装有 5 只球，编号为 1,2,3,4,5. 在袋中同时取 3 只球，用 $X$ 表示取出的 3 只球中的最大号码数，求 $X$ 的分布函数.

**解** $X$ 的可能取值为 3,4,5.

$$P\{X=3\}=C_2^2/C_5^3=\frac{1}{10},\quad P\{X=4\}=C_3^2/C_5^3=\frac{3}{10},$$

$$P\{X=5\}=C_4^2/C_5^3=\frac{6}{10},$$

即 $X$ 的分布列为

$$\begin{bmatrix}3 & 4 & 5\\ \frac{1}{10} & \frac{3}{10} & \frac{6}{10}\end{bmatrix},$$

从而 $X$ 的分布函数为

$$F(x)=\begin{cases}0, & x<3;\\ \frac{1}{10}, & 3\leqslant x<4;\\ \frac{4}{10}, & 4\leqslant x<5;\\ 1, & x\geqslant 5.\end{cases}$$

(2) 连续型随机变量的分布函数

设连续型随机变量 $X$ 的分布密度为

$$X\sim p(x),$$

则其分布函数为

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

即 $F(x)$是 $p(x)$在区间$(-\infty,x]$上的积分值. 在一定条件(例如 $p(x)$最多有有限多个间断点)下,有

$$p(x)=F'(x).$$

**例 3** 设

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

则

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

**例 4** 设连续型随机变量 $X$ 的分布函数为

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

试求:(1) 系数 $A$;(2) $X$ 落在$\left(-1,\dfrac{1}{2}\right)$及$\left(\dfrac{1}{3},2\right)$内的概率;(3) $X$ 的分布密度.

**解** (1) 由于 $F(x)$的连续性,有 $\lim\limits_{x\to 1-0}F(x)=F(1)$,即 $\lim\limits_{x\to 1-0}Ax^2=1$,得到 $A=1$. 于是

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

(2) $P\left\{-1<X<\dfrac{1}{2}\right\}=F\left(\dfrac{1}{2}\right)-F(-1)=\left(\dfrac{1}{2}\right)^2-0=\dfrac{1}{4}$;

$$P\left\{\frac{1}{3}<X<2\right\}=F(2)-F\left(\frac{1}{3}\right)=1-\left(\frac{1}{3}\right)^2=\frac{8}{9}.$$

(3) $p(x)=F'(x)=\begin{cases}2x, & 0\leqslant x<1,\\ 0, & \text{其他}.\end{cases}$

**例 5** 设 $X$ 的分布函数为

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

求(1) $B$ 的值； (2) $P\{-2<X<1\}$.

**解** 由 $F(x)$的连续性有 $F(0+0)=F(0)$,得 $1+B=0$,$B=-1$,即

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

$$\begin{aligned}P\{-2<X<1\}&=F(1)-F(-2)\\&=(1-e^{-1})-0=1-e^{-1}.\end{aligned}$$

**例 6** 对某目标进行射击,直至击中为止.如果每次射击的命中率为 $p$,求射击次数 $X$ 的分布列和分布函数.

**解**

$$\begin{aligned}P\{X=k\}&=\underbrace{(1-p)(1-p)\cdots(1-p)}_{k-1\text{个}}p\\&=(1-p)^{k-1}p=q^{k-1}p\\&\quad(k=1,2,\cdots;q=1-p),\end{aligned}$$

或写成

$$X\sim\begin{pmatrix}1 & 2 & 3 & \cdots & k & \cdots\\ p & pq & pq^2 & \cdots & pq^{k-1} & \cdots\end{pmatrix}.$$

当 $x<1$ 时,

$$F(x)=P\{X\leqslant x\}=0;$$

当 $1\leqslant x<2$ 时,

$$F(x)=P\{X\leqslant x\}=P\{X=1\}=p=1-q;$$

当 $2\leqslant x<3$ 时,

$$\begin{aligned}F(x)&=P\{X\leqslant x\}=P\{X=1\}+P\{X=2\}\\&=p+pq=1-q^2;\end{aligned}$$

......................................................

当 $k\leqslant x<k+1$ 时,

$$\begin{aligned}F(x)&=P\{X\leqslant x\}=P\{X=1\}+\cdots+P\{X=k\}\\&=p+pq+\cdots+pq^{k-1}=p(1+q+\cdots+q^{k-1})\end{aligned}$$

$$= p \cdot \frac{1-q^k}{1-q} = 1 - q^k;$$

……………………………………………………………………

故分布函数为

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

其中 $q=1-p$，$[x]$表示对 $x$ 取整，即取小于或等于 $x$ 的最大整数.

**例 7** 下列函数是否为分布函数？若是，则判断是哪种类型随机变量的分布函数.

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

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

(3) $F(x)=\begin{cases} 0, & x<0, \\ \sin x, & 0\leqslant x<\frac{\pi}{2}, \\ 1, & x\geqslant \frac{\pi}{2}; \end{cases}$

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

**解** (1) 由题设，$F(x)$在$(-\infty,+\infty)$上单调不减，右连续，并有 $F(-\infty)=\lim\limits_{x\to-\infty} F(x)=0$，$F(+\infty)=\lim\limits_{x\to+\infty} F(x)=1$，所以 $F(x)$是某一随机变量 $X$ 的分布函数. 因 $F(x)$是一阶梯函数，所以 $X$ 是离散型随机变量，其概率分布为：

$$P\{X=-2\}=\frac{1}{2}, \quad P\{X=0\}=\frac{1}{2}.$$

(2) 因 $F(x)$在$\left(\frac{\pi}{2},\pi\right)$上单调下降,所以 $F(x)$不可能是分布函数.

(3) 因为 $F(x)$在$(-\infty,+\infty)$上连续,单调不减,且有 $F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0$,$F(+\infty)=\lim\limits_{x\to+\infty}F(x)=1$,所以 $F(x)$是某一随机变量 $X$ 的分布函数.

设非负函数

$$p(x)=\begin{cases}0, & x<0 \text{ 或 } x\geqslant\frac{\pi}{2},\\ \cos x, & 0\leqslant x<\frac{\pi}{2},\end{cases}$$

则因为

$$\int_{-\infty}^{x}p(t)\mathrm{d}t=\begin{cases}0, & x<0,\\ \sin x, & 0\leqslant x<\frac{\pi}{2},\\ 1, & x\geqslant\frac{\pi}{2},\end{cases}$$

即$\int_{-\infty}^{x}p(t)\mathrm{d}t=F(x)$,所以 $F(x)$是连续型随机变量 $X$ 的分布函数.

(4) 因为 $F(x)$在$(-\infty,+\infty)$上单调不减,右连续,且有 $F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0$,$F(+\infty)=\lim\limits_{x\to+\infty}F(x)=1$,所以 $F(x)$是某一随机变量 $X$ 的分布函数.

又因为 $F(x)$在点 $x=0$ 处不连续,故 $X$ 不是连续型的随机变量,$F(x)$又不是阶梯函数,故 $X$ 也不是离散型的随机变量.

**定义 2.3**(常见分布)

1. 离散型随机变量

(1) 0-1 分布

设随机变量 $X$ 的分布为

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

则称 $X$ 服从参数为 $p$ 的**两点分布**，两点分布又称为 **0-1 分布**，记为 $X \sim B(1,p)$.

凡是只有两个基本事件的随机试验都可以确定一个服从两点分布的随机变量.

(2) 二项分布

设随机变量 $X$ 的分布为

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

$$(k=0,1,2,\cdots,n;\ 0<p<1,q=1-p),$$

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

利用二项式定理，容易验证二项分布的概率值 $p_k$ 满足：

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

一般地，在 $n$ 重伯努利试验中，事件 $A$ 恰发生 $k(0\leqslant k\leqslant n)$ 次的概率为：

$$P_n\{\mu=k\}=C_n^k p^k q^{n-k},\quad k=0,1,2,\cdots,n.$$

用 $X$ 表示 $n$ 重伯努利试验中事件 $A$ 发生的次数，则 $X \sim B(n,p)$.

(3) 超几何分布

设随机变量 $X$ 的分布为

$$P\{X=k\}=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}\quad (k=0,1,2,\cdots,l),$$

其中 $n<N-M,l=\min(M,n)$，则称 $X$ 服从**超几何分布**，记为 $H(n,M,N)$.

利用组合性质：$\sum\limits_{k=0}^{n} C_M^k C_{N-M}^{n-k}=C_N^n$，可以证明

$$\sum_{k=0}^{n} P\{X=k\}=1.$$

(4) 泊松(Poisson)分布

设随机变量 $X$ 的分布为

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

则称 $X$ 服从参数为 $\lambda$ 的**泊松分布**，记为 $X\sim P(\lambda)$.

利用 $e^x$ 的幂级数展开式，容易验证泊松分布的概率值 $p_k$ 满足：

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

(5) 几何分布

设随机变量 $X$ 的分布为

$$P\{X=k\}=pq^{k-1}$$

$$(k=1,2,\cdots,n,\cdots;\ 0<p<1,\ q=1-p),$$

则称 $X$ 服从参数为 $p$ 的**几何分布**，记为 $X\sim G(p)$.

利用几何级数求和公式容易验证几何分布的概率值 $p_k$ 满足：

$$\sum_k p_k=\sum_{k=1}^{\infty}pq^{k-1}=p\frac{1}{1-q}=1.$$

一般地，在伯努利试验中，事件 $A$ 首次出现在第 $k$ 次的概率为

$$p_k=pq^{k-1},\ k=1,2,\cdots,$$

通常称 $k$ 为事件 $A$ 的首发生次数. 如用 $X$ 表示事件 $A$ 的首发生次数，则 $X$ 服从几何分布.

2. 连续型随机变量

(1) 均匀分布

设随机变量 $X$ 的分布密度函数为

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

则称 $X$ 服从参数为 $a,b$ 的**均匀分布**，记为 $X\sim U(a,b)$.

如果随机变量 $X\sim U(a,b)$，那么对于任意的 $c,d$ $(a\leqslant c<d\leqslant b)$，按概率密度定义，有

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

上式表明，$X$ 在$(a,b)$(即有正概率密度区间)中任一个小区间上取值的概率与该区间的长度成正比，而与该小区间的位置无关，并且不难看出

$$\int_{-\infty}^{+\infty} p(x)\mathrm{d}x = \int_{-\infty}^{a} 0\mathrm{d}x + \int_a^b \frac{1}{b-a}\mathrm{d}x + \int_b^{+\infty} 0\mathrm{d}x = 1.$$

(2) 指数分布

设随机变量 $X$ 的分布密度函数为

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

则称 $X$ 服从参数为 $\lambda$ 的**指数分布**，记为 $X \sim \Gamma(1,\lambda)$或 $\mathrm{e}(\lambda)$. 不难看出

$$\int_{-\infty}^{+\infty} p(x)\mathrm{d}x = \int_{-\infty}^{0} 0\mathrm{d}x + \int_0^{+\infty} \lambda \mathrm{e}^{-\lambda x}\mathrm{d}x = 1.$$

(3) 正态分布

设随机变量 $X$ 的分布密度函数为

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

其中 $\mu,\sigma$ 为常数且 $\sigma>0$，则称 $X$ 服从参数为 $\mu,\sigma^2$ 的**正态分布**，记为 $X \sim N(\mu,\sigma^2)$. 若 $X$ 近似服从正态分布，简记为 $X \overset{\cdot}{\sim} N(\mu,\sigma^2)$.

特别地，称 $\mu=0,\sigma^2=1$ 的正态分布为**标准正态分布**，记为 $X \sim N(0,1)$，其密度函数为

$$p(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}.$$

在微积分中，我们证明了

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

由此，只要作变换$\frac{x-\mu}{\sigma}=t$，不难验证一般的正态分布密度也满足

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

（二）重要定理及公式

**定理 2.1（二项定理）** 若当 $N\to\infty$时，$\frac{M}{N}\to p$（$n,k$ 不变），则

$$\frac{C_M^k C_{N-M}^{n-k}}{C_N^n} \to C_n^k p^k (1-p)^{n-k} \quad (N \to \infty).$$

可见，超几何分布的极限分布为二项分布.

**定理 2.2（泊松定理）** 若当 $n\to\infty$时，$np\to\lambda>0$，则

$$C_n^k p^k (1-p)^{n-k} \to \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda} \quad (n \to \infty),$$

其中 $k=0,1,2,\cdots,n,\cdots$.

可见，二项分布的极限分布为泊松分布，因此，当 $n$ 很大与 $p$ 很小时，可以用泊松分布来作二项分布的近似计算.

（三）重要方法

**方法 2.1**（计算有关事件的概率）

（1）公式法

离散型随机变量 $X$ 的取值虽然是“不确定的”，但是它具有一定的“概率分布”. 概率分布不仅明确地给出了 $X$ 在点 $x_i$（以后称为正概率点）处的概率，而且对于任意实数 $a<b$，事件 $\{a\leqslant X\leqslant b\}$ 发生的概率都可以由分布算出. 这是因为事件

$$\{a \leqslant X \leqslant b\} = \bigcup_{a\leqslant x_i\leqslant b}\{X = x_i\},$$

于是由概率的可加性有

$$P\{a \leqslant X \leqslant b\} = \sum_{a\leqslant x_i\leqslant b} P\{X = x_i\}.$$

一般来说，对于实数集 $R$ 中任一个区间 $D$，都有

$$P\{X \in D\} = \sum_{x_i\in D} P\{X = x_i\}.$$

**例 8** 设某射手每次射击打中目标的概率为 0.5，现在连续射击 10 次，求击中目标的次数 $X$ 的概率分布. 又设至少命中 3 次才

可以参加下一步的考核,求此射手不能参加考核的概率.

**解** 这是一个 10 重伯努利试验,击中目标的次数 $X$ 的可能取值为 0,1,2,…,10,利用二项概型可求得

$$P\{X=k\}=C_{10}^{k}0.5^{k}0.5^{10-k},\quad k=0,1,2,\cdots,10.$$

设 $A=\{$此射手不能参加考核$\}$,有

$$P(A)=P\{X\leqslant 2\}=\sum_{k=0}^{2}P\{X=k\}$$

$$=\sum_{k=0}^{2}C_{10}^{k}0.5^{k}0.5^{10-k}=0.0546875.$$

与离散型随机变量类似,对于实数集 $\boldsymbol{R}$ 中任一区间 $D$,事件 $\{X\in D\}$的概率都可以由分布密度算出:

$$P\{X\in D\}=\int_{D}p(x)\mathrm{d}x,$$

其中 $p(x)$为一可求积函数.

**例 9** 设随机变量 $X$ 的分布密度函数为

$$p(x)=\begin{cases}Cx, & 0\leqslant x\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

求(1) 常数 $C$;(2) $P\{0.3\leqslant X\leqslant 0.7\}$;(3) $P\{-0.5\leqslant X<0.5\}$.

**解** (1) 由 $p(x)$的性质,有

$$1=\int_{-\infty}^{+\infty}p(x)\mathrm{d}x=\int_{0}^{1}Cx\mathrm{d}x=C\cdot\frac{x^2}{2}\Big|_0^1=\frac{1}{2}C,$$

因此,$C=2$.

(2) $P\{0.3\leqslant X\leqslant 0.7\}=\int_{0.3}^{0.7}2x\mathrm{d}x=x^2\Big|_{0.3}^{0.7}=0.4.$

(3) $P\{-0.5\leqslant X\leqslant 0.5\}=\int_{-0.5}^{0}0\mathrm{d}x+\int_{0}^{0.5}2x\mathrm{d}x=x^2\Big|_{0}^{0.5}=0.25.$

**例 10** 设 $X\sim p(x)=\begin{cases}2\mathrm{e}^{-2x}, & x\geqslant 0,\\ 0, & x<0,\end{cases}$求

(1) $P\{-1\leqslant X\leqslant 4\}$; (2) $P\{X<-3\}$; (3) $P\{X\geqslant -10\}$.

**解** (1) $P\{-1\leqslant X\leqslant 4\}=\int_{-1}^{4}p(x)\mathrm{d}x=\int_{-1}^{0}0\mathrm{d}x+\int_{0}^{4}2\mathrm{e}^{-2x}\mathrm{d}x$

$$=1-e^{-8};$$

(2) $P\{X<-3\}=\int_{-\infty}^{-3}p(x)\mathrm{d}x=\int_{-\infty}^{-3}0\mathrm{d}x=0;$

(3) $P\{X\geqslant-10\}=\int_{-10}^{+\infty}p(x)\mathrm{d}x=\int_{-10}^{0}0\mathrm{d}x+\int_{0}^{+\infty}2e^{-2x}\mathrm{d}x=1.$

(2) 查表法(正态分布数值表)

现在我们来讨论如何计算服从正态分布的随机变量在任一区间上取值的概率。正态分布是最常用的分布，为计算方便，人们已经编制了 $\Phi(x)$值的表(见书后附表 1)，其中

$$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\mathrm{d}t\ (x\geqslant 0).$$

由 $\Phi(x)$的性质可知，表中只需列出 $x\geqslant 0$ 的值.

下面用实例来说明，对于服从正态分布或服从标准正态分布的随机变量均能利用 $x\geqslant 0$ 时 $\Phi(x)$的值来计算其取值于任一区间的概率.

**例 11** 设 $X\sim N(0,1)$，求 $P\{X<2.35\}$，$P\{X<-1.25\}$以及 $P\{|X|<1.55\}$.

**解** $P\{X<2.35\}=\Phi(2.35)\xlongequal{\text{查表}}0.9906;$

$$\begin{aligned}P\{X<-1.25\}&=\Phi(-1.25)=1-\Phi(1.25)\\&=1-0.8944=0.1056;\\P\{|X|<1.55\}&=P\{-1.55<X<1.55\}\\&=\Phi(1.55)-\Phi(-1.55)\\&=2\Phi(1.55)-1=2\times0.9394-1\\&=0.8788.\end{aligned}$$

**例 12** 设 $X\sim N(1,2^2)$，求 $P\{0<X\leqslant 5\}$.

对于服从非标准正态分布 $N(\mu,\sigma^2)$的随机变量，我们只需进行积分变换，有

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

$$\xlongequal{\text{令 } t=\frac{x-\mu}{\sigma}} \int_{\frac{\alpha-\mu}{\sigma}}^{\frac{\beta-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$= \Phi\left(\frac{\beta-\mu}{\sigma}\right) - \Phi\left(\frac{\alpha-\mu}{\sigma}\right).$$

查表即可求出此值.

**解** 这里 $\mu=1, \sigma=2, \beta=5, \alpha=0$,有

$$\frac{\beta-\mu}{\sigma}=2, \quad \frac{\alpha-\mu}{\sigma}=-0.5.$$

于是

$$\begin{aligned} P\{0<X\leqslant 5\} &= \Phi(2)-\Phi(-0.5) \\ &= \Phi(2)-[1-\Phi(0.5)] \\ &= \Phi(2)+\Phi(0.5)-1 \\ &= 0.9772+0.6915-1=0.6687. \end{aligned}$$

**方法 2.2**(随机变量函数的分布的求法)

1. 离散型随机变量函数的分布

设 $X$ 是离散型随机变量,其概率分布为

| $X$ | $x_1$ | $x_2$ | … | $x_n$ | … |
|---|---|---|---|---|---|
| $P(X=x_i)$ | $p_1$ | $p_2$ | … | $p_n$ | … |

记 $y_i=f(x_i)(i=1,2,\cdots)$. 如果 $f(x_i)$的值全都不相等,那么 $Y$ 的概率分布为

| $Y$ | $y_1$ | $y_2$ | … | $y_n$ | … |
|---|---|---|---|---|---|
| $P(Y=y_i)$ | $p_1$ | $p_2$ | … | $p_n$ | … |

但是,如果 $f(x_i)$的值中有相等的,那么就把那些相等的值分别合并,并根据概率加法公式把相应的概率相加,便得到 $Y$ 的分布.

**例 13** 设随机变量 $X$ 的分布为

| $X$ | $-2$ | $-1$ | $0$ | $1$ | $2$ |
|---|---|---|---|---|---|
| $P(X=x_i)$ | $\frac{1}{5}$ | $\frac{1}{5}$ | $\frac{1}{5}$ | $\frac{1}{10}$ | $\frac{3}{10}$ |

求 $Y=X^2+1$ 的概率分布.

**解** 由 $y_i=x_i^2+1(i=1,2,\cdots,5)$及 $X$ 的分布,得到

| $x^2+1$ | $(-2)^2+1$ | $(-1)^2+1$ | $0^2+1$ | $1^2+1$ | $2^2+1$ |
|---|---|---|---|---|---|
| $P(X=x_i)$ | $\frac{1}{5}$ | $\frac{1}{5}$ | $\frac{1}{5}$ | $\frac{1}{10}$ | $\frac{3}{10}$ |

把 $f(x_i)=x_i^2+1$ 相同的值合并起来,并把相应的概率相加,便得到 $Y$ 的分布,即

$$P\{Y=5\}=P\{X=-2\}+P\{X=2\}=\frac{1}{2},$$

$$P\{Y=2\}=P\{X=-1\}+P\{X=1\}=\frac{3}{10},$$

$$P\{Y=1\}=P\{X=0\}=\frac{1}{5}.$$

所以

| $Y$ | $5$ | $2$ | $1$ |
|---|---|---|---|
| $P(Y=y_i)$ | $\frac{1}{2}$ | $\frac{3}{10}$ | $\frac{1}{5}$ |

2. 连续型随机变量函数的分布

(1) 定义法

设 $X$ 是连续型随机变量,其分布密度函数为 $p(x)$. 对于给定的一个其导函数是连续的函数 $f(x)$,我们用分布函数的定义导出 $Y=f(X)$的分布.

为了讨论方便,对于 $X$ 有正概率密度的区间上的一切 $x$,令

$$\alpha=\min_x\{f(x)\},\quad \beta=\max_x\{f(x)\}.$$

于是，对于 $\alpha>-\infty,\beta<+\infty$ 情形，有：

当 $y<\alpha$ 时，$\{f(X)\leqslant y\}$ 是一个不可能事件，故 $F(y)=P\{f(X)\leqslant y\}=0$；而当 $y\geqslant\beta$ 时，$\{f(X)\leqslant y\}$ 是一个必然事件，故 $F(y)=P\{f(X)\leqslant y\}=1$. 这样，我们可设 $Y$ 的分布函数为

$$F(y)=\begin{cases}0, & y\leqslant\alpha,\\ *, & \alpha<y<\beta,\\ 1, & y\geqslant\beta.\end{cases}$$

对于 $\alpha=-\infty$ 或 $\beta=+\infty$ 的情形，只要去掉相应区间上 $F(y)$ 的表达式即可. 这里我们只需讨论 $\alpha<y<\beta$ 的情形，根据分布函数的定义有

$$\begin{aligned}* &= P\{Y\leqslant y\}=P\{f(X)\leqslant y\}\\ &= P\{X\in D_y\}=\int_{D_y}p(x)\mathrm{d}x,\end{aligned}$$

其中 $D_y=\{x\mid f(x)\leqslant y\}$，即 $D_y$ 是由满足 $f(x)\leqslant y$ 的所有 $x$ 组成的集合，它可由 $y$ 的值及 $f(x)$ 的函数形式解出. 根据 $p(y)=F'(y)$，并考虑到常数的导数为 0，于是 $Y$ 的分布密度为

$$p(y)=\begin{cases}\left[\int_{D_y}p(x)\mathrm{d}x\right]_y', & \alpha<y<\beta,\\ 0, & \text{其他}.\end{cases}$$

**例 14** 对一圆片直径进行测量，其值在[5,6]上均匀分布，求圆片面积的概率分布密度.

**解** 设圆片直径的测量值为 $X$，面积为 $Y$，则有 $Y=\frac{\pi}{4}X^2$. 按已知条件，$X$ 的分布密度为

$$p(x)=\begin{cases}1, & x\in[5,6],\\ 0, & \text{其他}.\end{cases}$$

对于函数 $y=\frac{\pi}{4}x^2$，当 $x\in[5,6]$ 时

$$\alpha=\min\left\{\frac{\pi}{4}x^2\right\}=\frac{25}{4}\pi,\quad \beta=\max\left\{\frac{\pi}{4}x^2\right\}=\frac{36}{4}\pi=9\pi.$$

于是

$$F(y)=\begin{cases}0, & y\leqslant 25\pi/4,\\ *, & 25\pi/4<y<9\pi,\\ 1, & y\geqslant 9\pi.\end{cases}$$

当 $25\pi/4<y<9\pi$ 时

$$F(y)=P\{Y\leqslant y\}=P\left\{\frac{\pi X^2}{4}\leqslant y\right\}=P\left\{X\leqslant\sqrt{\frac{4}{\pi}y}\right\}$$

$$=\int_{-\infty}^{\sqrt{\frac{4}{\pi}y}}p(x)\mathrm{d}x=\int_{-\infty}^{5}0\mathrm{d}x+\int_{5}^{\sqrt{\frac{4}{\pi}y}}1\mathrm{d}x=\sqrt{\frac{4}{\pi}y}-5.$$

由于

$$p(y)=F'(y)=\left(\sqrt{\frac{4}{\pi}y}-5\right)'=\frac{1}{\sqrt{\pi y}},$$

故随机变量 $Y$ 的分布密度函数为

$$p(y)=\begin{cases}\dfrac{1}{\sqrt{\pi y}}, & 25\pi/4<y<9\pi,\\ 0, & \text{其他}.\end{cases}$$

(2) 公式法

利用上述方法可以推出，当函数 $y=f(x)$ 为单调函数时，随机变量 $Y$ 的分布密度可由下面的公式得到

$$p(y)=\begin{cases}p_X(f^{-1}(y))\cdot|(f^{-1}(y))'|, & \alpha<y<\beta,\\ 0, & \text{其他},\end{cases}$$

其中 $f^{-1}(y)$ 为 $f(x)$ 的反函数，$p_X(x)$ 为随机变量 $X$ 的分布密度函数.

在例 14 中

$$f^{-1}(y)=\sqrt{\frac{4}{\pi}y},\quad (f^{-1}(y))'=\left(\sqrt{\frac{4y}{\pi}}\right)'=\frac{1}{\sqrt{\pi y}},$$

而当 $25\pi/4<y<9\pi$ 时，$5<x<6$，有

$$p_X\left(\sqrt{\frac{4y}{\pi}}\right)=p_X(x)=1.$$

由公式可得到 $Y$ 的分布密度函数

$$p(y)=\begin{cases}1\cdot\dfrac{1}{\sqrt{\pi y}}, & 25\pi/4<y<9\pi,\\ 0, & \text{其他}\end{cases}$$

$$=\begin{cases}\dfrac{1}{\sqrt{\pi y}}, & 25\pi/4<y<9\pi,\\ 0, & \text{其他}.\end{cases}$$

当函数 $y=f(x)$不是单调函数时，也可利用上述公式求出随机变量 $Y$ 的分布密度，即

$$p(y)=\sum_i p_i(y),$$

其中 $p_i(y)$是 $y=f(x)$的第 $i$ 个单调可微子区间上的分布密度.

**例 15** 设随机变量 $X$ 的密度函数为

$$p(x)=\begin{cases}0, & x<0,\\ x^3e^{-x^2}, & x\geqslant 0.\end{cases}$$

试求：(1) $Y_1=2X+3$；(2) $Y_2=X^2$；(3) $Y_3=\ln X$ 的密度函数.

**解** (1) $Y_1=2X+3$，于是 $y=2x+3$，其反函数为 $x=\dfrac{y-3}{2}$，于是 $x'=\dfrac{1}{2}$，故

$$p_1(y)=\begin{cases}\dfrac{1}{2}\left(\dfrac{y-3}{2}\right)^3e^{-\left(\frac{y-3}{2}\right)^2}, & y\geqslant 3,\\ 0, & y<3.\end{cases}$$

(2) $Y_2=X^2$，于是 $y=x^2$，其反函数分别为 $x_1=\sqrt{y}$ 或 $x_2=-\sqrt{y}$，于是

$$x_1'=\frac{1}{2\sqrt{y}},\quad x_2'=-\frac{1}{2\sqrt{y}}.$$

故 $$p_2(y)=p(\sqrt{y})(\sqrt{y})_y'+p(-\sqrt{y})\left|(-\sqrt{y})_y'\right|$$

$$=\frac{1}{2\sqrt{y}}(\sqrt{y})^3e^{-(\sqrt{y})^2}+0\times\frac{1}{2\sqrt{y}}.$$

由此得
$$p_2(y)=\begin{cases}\frac{1}{2}ye^{-y}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

(3) $Y_3=\ln X$，于是 $y=\ln x$，其反函数为 $x=e^y$，于是 $x'=e^y$，故
$$p_3(y)=p(e^y)e^y=e^{4y}e^{-e^{-2y}}\quad(-\infty<y<+\infty).$$

## 三、典型例题分析

### （一）填空题

1. 若随机变量 $\xi$ 在(1,6)上服从均匀分布，则方程 $x^2+\xi x+1=0$ 有实根的概率是______.

答案是：0.8.

**分析**
$$\begin{aligned}P\{x^2+\xi x+1=0\text{ 有实根}\}&=P\{\xi^2-4\geqslant 0\}\\&=P\{|\xi|\geqslant 2\}=P\{6>\xi\geqslant 2\}\\&=\int_2^6\frac{1}{5}du=\frac{4}{5}=0.8.\end{aligned}$$

2. 已知连续随机变量 $X$ 的概率密度函数为
$$f(x)=\frac{1}{\sqrt{\pi}}e^{-x^2+2x-1},$$
则 $X$ 的数学期望为______；$X$ 的方差为______.

答案是：$1,\frac{1}{2}$.

**分析** 将 $f(x)$ 改写为
$$f(x)=\frac{1}{\sqrt{2\pi}\cdot\sqrt{\frac{1}{2}}}\cdot\exp\left\{-\frac{(x-1)^2}{2(\sqrt{1/2})^2}\right\},$$
可见 $X$ 服从正态分布 $N\left(1,\frac{1}{2}\right)$，所以 $E(X)=1,D(X)=\frac{1}{2}$.

3. 设随机变量 $X$ 服从均值为 10，均方差为 0.02 的正态分布. 已知 $\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du$，$\Phi(2.5)=0.9938$，则 $X$ 落在区间

(9.95,10.05)内的概率为______.

答案是：0.9876.

**分析** $P\{9.95<X<10.05\}$

$$=P\left\{\frac{9.95-10}{0.02}<\frac{X-10}{0.02}<\frac{10.05-10}{0.02}\right\}$$
$$=\Phi\left(\frac{10.05-10}{0.02}\right)-\Phi\left(\frac{9.95-10}{0.02}\right)$$
$$=\Phi(2.5)-\Phi(-2.5)$$
$$=\Phi(2.5)-[1-\Phi(2.5)]$$
$$=2\Phi(2.5)-1=0.9876.$$

4. 已知随机变量 $X$ 的概率密度函数

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

则 $X$ 的概率分布函数 $F(x)=$______.

答案是：$\begin{cases}\frac{1}{2}e^{x}, & x<0,\\ 1-\frac{1}{2}e^{-x}, & x\geqslant 0.\end{cases}$

5. 已知离散型随机变量 $X$ 服从参数为 2 的泊松(Poisson)分布，即 $P\{X=k\}=\frac{2^k e^{-2}}{k!}$，$k=0,1,2,\cdots$，则随机变量 $Z=3X-2$ 的数学期望 $E(Z)=$______.

答案是：4.

**分析** $E(Z)=3E(X)-2=3\times2-2=4.$

6. 设随机变量 $X$ 服从参数为 1 的指数分布，则数学期望 $E\{X+e^{-2X}\}=$______.

答案是：$\frac{4}{3}$.

**分析** $E\{X+e^{-2X}\}=\int_0^{+\infty}xe^{-x}dx+\int_0^{+\infty}e^{-2x}e^{-x}dx$

$$=1+\frac{1}{3}=\frac{4}{3}.$$

7. 设随机变量 $X$ 服从(0,2)上的均匀分布，则随机变量 $Y=$

$X^2$ 在(0,4)内概率分布密度 $p_Y(y)=$______.

答案是：$\dfrac{1}{4\sqrt{y}}$.

**分析** $y=x^2(0<x<2)$的反函数 $x=\sqrt{y}(0<y<4)$.

$$p_Y(y)=p_X(\sqrt{y})\cdot|(\sqrt{y})'|=\frac{1}{2\sqrt{y}}p_X(\sqrt{y})$$

$$=\frac{1}{2\sqrt{y}}\cdot\frac{1}{2},\quad 0<\sqrt{y}<2,$$

即
$$f_Y(y)=\frac{1}{4\sqrt{y}},\quad 0<y<4.$$

8. 设 $X$ 表示 10 次独立重复射击命中目标的次数，每次射中目标的概率为 0.4，则 $X^2$ 的数学期望 $E(X^2)=$______.

答案是：18.4.

**分析** $X\sim B(10,0.4)$，$E(X)=10\times0.4=4$，

$D(X)=10\times0.4\times0.6=2.4$，

$E(X^2)=(E(X))^2+D(X)=4^2+2.4=18.4$.

9. 若随机变量 $X$ 服从均值为 2，方差为 $\sigma^2$ 的正态分布，且 $P\{2<X<4\}=0.3$，则 $P\{X<0\}=$______.

答案是：0.2.

**分析** 由于 $X$ 的密度函数关于 $X=2$ 为轴对称，故

$$P\{X<2\}=P\{X>2\}=0.5,$$
$$P\{0<X<2\}=P\{2<X<4\}=0.3,$$

从而

$$\begin{aligned}P\{X<0\}&=P\{X<2\}-P\{0\leqslant X<2\}\\&=P\{X<2\}-P\{0<X<2\}\\&=0.5-0.3=0.2.\end{aligned}$$

10. 设随机变量的分布函数为

$$F(x)=\begin{cases}0, & 若\ x<0,\\ A\sin x, & 若\ 0\leqslant x\leqslant\pi/2,\\ 1, & 若\ x>\pi/2,\end{cases}$$

则 $A=$______，$P\left\{|x|<\frac{\pi}{6}\right\}=$______.

答案是：1，$\frac{1}{2}$.

**分析** 由 $F(x)=P\{X\leqslant x\}$ 右连续，即 $F\left(\frac{\pi}{2}+0\right)=F\left(\frac{\pi}{2}\right)$，得到 $A=1$. 而

$$P\left\{|X|<\frac{\pi}{6}\right\}=P\left\{-\frac{\pi}{6}<X<\frac{\pi}{6}\right\}$$

$$=F\left(\frac{\pi}{6}\right)-F\left(-\frac{\pi}{6}\right)=\sin\frac{\pi}{6}-0=\frac{1}{2}.$$

11. 设随机变量 $X$ 的分布函数为

$$F(x)=P\{X\leqslant x\}=\begin{cases}0, & \text{若 } x<-1,\\ 0.4, & \text{若 } -1\leqslant x<1,\\ 0.8, & \text{若 } 1\leqslant x<3,\\ 1, & \text{若 } x\geqslant 3,\end{cases}$$

则 $X$ 的概率分布为__________.

答案是：

| $x$ | $-1$ | $1$ | $3$ |
|---|---|---|---|
| $P\{X=x\}$ | 0.4 | 0.4 | 0.2 |

**分析** 由公式 $P\{X=x_0\}=F(x_0)-F(x_0-0)$ 算出

$$P\{X=-1\}=0.4-0=0.4,$$

$$P\{X=1\}=0.8-0.4=0.4,$$

$$P\{X=3\}=1-0.8=0.2.$$

12. 设随机变量 $X$ 的概率密度为

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

以 $Y$ 表示对 $X$ 的三次独立重复观察中事件 $\left\{X\leqslant\frac{1}{2}\right\}$ 出现的次数，则 $P\{Y=2\}=$______.

答案是：$\frac{9}{64}$.

**分析**　由于 $Y\sim B(3,p)$，其中

$$p=P\left\{X\leqslant\frac{1}{2}\right\}=\int_0^{\frac{1}{2}}2x\mathrm{d}x=\frac{1}{4},$$

故　$$P\{Y=2\}=C_3^2\left(\frac{1}{4}\right)^2\cdot\left(\frac{3}{4}\right)=\frac{9}{64}.$$

13. 设随机变量 $X$ 服从参数为 $(2,p)$ 的二项分布，随机变量 $Y$ 服从参数为 $(3,p)$ 的二项分布. 若 $P\{X\geqslant1\}=\frac{5}{9}$，则 $P\{Y\geqslant1\}=$ ______.

答案是：$\frac{19}{27}$.

**分析**　由于 $P\{X=0\}=1-P\{X\geqslant1\}=1-\frac{5}{9}=\frac{4}{9}$，故由 $p\{X=0\}=C_2^0p^0q^2=q^2=\frac{4}{9}$，得 $q=\frac{2}{3}$. 从而

$$P\{Y\geqslant1\}=1-P\{Y=0\}=1-C_3^0p^0q^3=1-q^3$$
$$=1-\left(\frac{2}{3}\right)^3=\frac{19}{27}.$$

14. 一实习生用一台机器接连独立地制造 3 个同种零件，第 $i$ 个零件是不合格品的概率 $p_i=\frac{1}{i+1}(i=1,2,3)$，以 $X$ 表示 3 个零件中合格品的个数，则 $P\{X=2\}=$ ______.

答案是：$\frac{11}{24}$.

**分析**　设 $B_i$ 为"第 $i$ 个零件为合格品"有

$$P(\overline{B}_1)=\frac{1}{2},\quad P(\overline{B}_2)=\frac{1}{3},\quad P(\overline{B}_3)=\frac{1}{4};$$

而事件 $\{X=2\}=\overline{B}_1B_2B_3+B_1\overline{B}_2B_3+B_1B_2\overline{B}_3$，于是

$$P\{X=2\}=\frac{1}{2}\times\frac{2}{3}\times\frac{3}{4}+\frac{1}{2}\times\frac{1}{3}\times\frac{3}{4}+\frac{1}{2}\times\frac{2}{3}\times\frac{1}{4}$$
$$=\frac{11}{24}.$$

（二）选择题

1. 设随机变量 $X$ 的密度函数为 $p(x)$，且 $p(-x)=p(x)$，$F(x)$是 $X$ 的分布函数，则对任意实数 $a$，有（　　）.

(A) $F(-a)=1-\int_0^a p(x)\mathrm{d}x$； (B) $F(-a)=\frac{1}{2}-\int_0^a p(x)\mathrm{d}x$；

(C) $F(-a)=F(a)$； (D) $F(-a)=2F(a)-1$.

答案是：B.

**分析** 由 $p(-x)=p(x)$，有

$$\int_{-\infty}^0 p(x)\mathrm{d}x=\int_0^{-\infty} p(x)\mathrm{d}x=\frac{1}{2}$$

和

$$\int_0^{-a} p(x)\mathrm{d}x=-\int_0^a p(-t)\mathrm{d}t=-\int_0^a p(t)\mathrm{d}t=-\int_0^a p(x)\mathrm{d}x.$$

所以

$$\begin{aligned}F(-a)&=\int_{-\infty}^{-a} p(x)\mathrm{d}x=\int_{-\infty}^0 p(x)\mathrm{d}x+\int_0^{-a} p(x)\mathrm{d}x\\&=\frac{1}{2}-\int_0^a p(x)\mathrm{d}x.\end{aligned}$$

2. 设随机变量 $X$ 与 $Y$ 均服从正态分布，$X\sim N(\mu,4^2)$，$Y\sim N(\mu,5^2)$，记 $p_1=P\{X\leqslant\mu-4\}$，$p_2=P\{y\geqslant\mu+5\}$，则（　　）.

(A) 对任何实数 $\mu$，都有 $p_1=p_2$；

(B) 对任何实数 $\mu$，都有 $p_1<p_2$；

(C) 只对 $\mu$ 的个别值，才有 $p_1=p_2$；

(D) 对任何实数 $\mu$，都有 $p_1>p_2$.

答案是：A.

**分析** 用 $\Phi$ 代表标准正态分布 $N(0,1)$的分布函数，有

$$p_1=P\left\{\frac{X-\mu}{4}\leqslant-1\right\}=\Phi(-1),$$

$$p_2=P\left\{\frac{Y-\mu}{5}\geqslant1\right\}=1-P\left\{\frac{Y-\mu}{5}<1\right\}=1-\Phi(1),$$

由于 $\Phi(-1)=1-\Phi(1)$，所以 $p_1=p_2$.

3. 设随机变量 $X$ 服从正态分布 $N(\mu,\sigma^2)$，则随 $\sigma$ 的增大，概率 $P\{|X-\mu|<\sigma\}$(　　).

(A) 单调增大；　　(B) 单调减小；

(C) 保持不变；　　(D) 增减不定.

答案是：C.

**分析**　$P\{|X-\mu|<\sigma\}=P\left\{\left|\dfrac{X-\mu}{\sigma}\right|<1\right\}$

$$=\Phi(1)-\Phi(-1)=2\Phi(1)-1,$$

其中 $\Phi$ 表示 $N(0,1)$的分布函数.

4. 设两个相互独立的随机变量 $X$ 和 $Y$ 分别服从正态分布 $N(0,1)$和 $N(1,1)$，则(　　).

(A) $P\{X+Y\leqslant 0\}=\dfrac{1}{2}$；　　(B) $P\{X+Y\leqslant 1\}=\dfrac{1}{2}$；

(C) $P\{X-Y\leqslant 0\}=\dfrac{1}{2}$；　　(D) $P\{X-Y\leqslant 1\}=\dfrac{1}{2}$.

答案是：B.

**分析**　由于 $X$ 与 $Y$ 相互独立，则 $Z=X+Y$ 仍为正态分布，即 $Z\sim N(\mu,\sigma^2)$，并且

$$\mu=E(X+Y)=E(X)+E(Y)=1,$$
$$\sigma^2=D(X+Y)=D(X)+D(Y)=2,$$

可知，$P\{X+Y\leqslant 1\}=P\{Z\leqslant 1\}=\dfrac{1}{2}$.

5. 假设随机变量 $X$ 服从指数分布，则随机变量 $Y=\min\{X,2\}$的分布函数(　　).

(A) 是连续函数；　　(B) 至少有两个间断点；

(C) 是阶梯函数；　　(D) 恰好有一个间断点.

答案是：D.

**分析**　这里的 $Y$ 是一个离散与连续混合型的随机变量. $Y$ 的分布函数为：

$$F(y)=\begin{cases}1-\mathrm{e}^{-\lambda y}, & y<2,\\ 1, & y\geqslant 2,\end{cases}$$

可见，$y=2$ 是它的一个间断点，且只有这一个.

（三）解答题

1. 设随机变量 $X$ 服从二项分布，其概率分布 $P\{X=k\}=C_n^k p^k q^{n-k}$，$k=0,1,\cdots,n$，$q=1-p$，问 $k$ 为何值时能使 $P\{X=k\}$ 最大.

**解** 因为对 $0<p<1$，有

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

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

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，如果 $(n+1)p$ 是整数，则当 $k_0=(n+1)p-1$ 或 $k_0=(n+1)p$ 时，$P\{X=k_0\}$ 最大；如果 $(n+1)p$ 不是整数，则当 $k_0=[(n+1)p]$ 时，$P\{X=k_0\}$ 最大.

通常我们称使 $P\{X=k\}$ 达最大的 $k_0$ 为最可能出现次数.

2. 假设随机变量 $X$ 在区间 $(1,2)$ 上服从均匀分布，试求随机变量 $Y=e^{2X}$ 的概率密度 $p_2(y)$.

**解** $X$ 的密度函数为

$$p_1(x)=\begin{cases}1, & 1<x<2,\\ 0, & \text{其他}.\end{cases}$$

记 $F(y)=P\{Y\leqslant y\}$ 为 $Y$ 的分布函数，则有

$$F(y)=P\{Y\leqslant y\}=P\{e^{2x}\leqslant y\}=\begin{cases}0, & y\leqslant e^2,\\ \int_1^{\frac{1}{2}\ln y}dx, & e^2<y<e^4,\\ 1, & y\geqslant e^4,\end{cases}$$

因此

$$p_2(y)=F'(y)=\begin{cases}0, & y\leqslant e^2,\\ \dfrac{1}{2y}, & e^2<y<e^4,\\ 0, & y\geqslant e^4,\end{cases}$$

故
$$p_2(y)=\begin{cases}\dfrac{1}{2y}, & e^2<y<e^4,\\ 0, & \text{其他}.\end{cases}$$

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

**解** $X$ 的密度函数为

$$p(x)=\begin{cases}\dfrac{1}{3}, & 2\leqslant x\leqslant 5,\\ 0, & \text{其他}.\end{cases}$$

记 $A=\{X>3\}$，则

$$P(A)=P\{X>3\}=\int_3^5\frac{1}{3}\mathrm{d}x=\frac{2}{3}.$$

用 $\xi$ 表示三次独立观测中观测值大于 3 的次数，则 $\xi$ 服从参数为 $n=3,p=\dfrac{2}{3}$ 的二项分布，故所求概率为

$$\begin{aligned}P\{\xi\geqslant 2\}&=P\{\xi=2\}+P\{\xi=3\}\\&=C_3^2\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)+C_3^3\cdot\left(\frac{2}{3}\right)^3=\frac{20}{27}.\end{aligned}$$

4. 某仪器装有三只独立工作的同型号电子元件，其寿命(单位：小时)都服从同一指数分布，分布密度为

$$p(x)=\begin{cases}\dfrac{1}{600}e^{-\frac{x}{600}}, & \text{若 } x>0,\\ 0, & \text{若 } x\leqslant 0.\end{cases}$$

试求：在仪器使用的最初 200 小时内，至少有一只电子元件损坏的概率 $\alpha$.

**解** 以 $X_i(i=1,2,3)$ 表示第 $i$ 只元件寿命，以 $A_i(i=1,2,3)$ 表示事件“在仪器使用最初 200 小时内，第 $i$ 只元件损坏”，则

$$P(\overline{A}_i)=P\{X_i>200\}=\int_{200}^{+\infty}\frac{1}{600}e^{-\frac{x}{600}}dx=e^{-\frac{1}{3}}.$$

所求概率为

$$\begin{aligned}\alpha&=P(A_1\cup A_2\cup A_3)=1-P(\overline{A}_1\cdot\overline{A}_2\cdot\overline{A}_3)\\&=1-P(\overline{A}_1)\cdot P(\overline{A}_2)\cdot P(\overline{A}_3)\\&=1-(e^{-\frac{1}{3}})^3=1-e^{-1}.\end{aligned}$$

5. 已知随机变量 $X$ 的概率密度为

$$p(x)=Ae^{-|x|},\quad -\infty<x<+\infty.$$

试求(1) $A$；(2) $P\{0<X<1\}$；(3) $X$ 的分布函数.

**解** (1) 由于 $\int_{-\infty}^{+\infty}p(x)dx=\int_{-\infty}^{+\infty}Ae^{-|x|}dx=2A\int_0^{+\infty}e^{-x}dx=1$，

即 $2A=1,A=\frac{1}{2}$. 所以 $p(x)=\frac{1}{2}e^{-|x|}$.

(2) $P\{0<X<1\}=\int_0^1\frac{1}{2}e^{-x}dx=\frac{1-e^{-1}}{2}$.

(3) $F(x)=\int_{-\infty}^{x}\frac{1}{2}e^{-|t|}dt$.

当 $x<0$ 时，

$$F(x)=\frac{1}{2}\int_{-\infty}^{x}e^{t}dt=\frac{1}{2}e^{x};$$

当 $x\geqslant0$ 时，

$$\begin{aligned}F(x)&=\frac{1}{2}\int_{-\infty}^{x}e^{-|t|}dt=\frac{1}{2}\int_{-\infty}^{0}e^{t}dt+\frac{1}{2}\int_0^{x}e^{-t}dt\\&=1-\frac{1}{2}e^{-x}.\end{aligned}$$

所以

$$F(x)=\begin{cases}\frac{1}{2}e^{x}, & x<0,\\ 1-\frac{1}{2}e^{-x}, & x\geqslant0.\end{cases}$$

6. 设随机变量 $X$ 的分布函数为：

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

求(1) 系数 $A$ 与 $B$；(2) $X$ 落在$(-1,1)$内的概率；(3) $X$ 的分布

密度.

**解** (1) 由 $F(-\infty)=0, F(+\infty)=1$ 可知

$$\begin{cases} A+B\left(-\dfrac{\pi}{2}\right)=0, \\ A+B\left(\dfrac{\pi}{2}\right)=1, \end{cases}$$

解得 $A=\frac{1}{2}$, $B=\frac{1}{\pi}$,所以

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

(2) $P\{-1<X<1\}=F(1)-F(-1)$

$$=\left(\frac{1}{2}+\frac{1}{\pi}\arctan 1\right)-\left(\frac{1}{2}+\frac{1}{\pi}\arctan(-1)\right)$$

$$=\frac{1}{2}+\frac{1}{\pi}\cdot\frac{\pi}{4}-\left(\frac{1}{2}+\frac{1}{\pi}\left(-\frac{\pi}{4}\right)\right)=\frac{1}{2}.$$

(3) 分布密度

$$p(x)=F'(x)=\left(\frac{1}{2}+\frac{1}{\pi}\arctan x\right)'=\frac{1}{\pi(1+x^2)},$$

$$-\infty<x<+\infty.$$

7. 假设测量的随机误差 $X\sim N(0,10^2)$,试求在 100 次独立重复测量中,至少有三次测量误差的绝对值大于 19.6 的概率 $\alpha$,并利用泊松分布求出 $\alpha$ 的近似值.(要求小数点后取两位有效数字).

附表

| $\lambda$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | … |
|---|---|---|---|---|---|---|---|---|
| $e^{-\lambda}$ | 0.368 | 0.135 | 0.050 | 0.018 | 0.007 | 0.002 | 0.001 | … |

**解** 每次测量误差的绝对值大于 19.6 的概率

$$p=P\{|X|>19.6\}=P\left\{\frac{|X|}{10}>1.96\right\}=0.05.$$

设 $\mu$ 为 100 次独立重复试验中事件 $\{|X|>19.6\}$ 出现的次数,$\mu$ 服从参数为 $n=100, p=0.05$ 的二项分布,所求概率

$$\alpha=P\{\mu\geqslant 3\}=1-P\{\mu<3\}$$

$$
\begin{aligned}
&= 1 - P\{\mu = 0\} - P\{\mu = 1\} - P\{\mu = 2\} \\
&= 1 - (0.95)^3 - 100 \times 0.05 \times 0.95^{99} \\
&\quad - \frac{100 \times 99}{2} \times 0.05^2 \times 0.95^{98}.
\end{aligned}
$$

由泊松定理，$\mu$ 近似服从参数为 $\lambda = np = 100 \times 0.05 = 5$ 的泊松分布，从而

$$
\begin{aligned}
\alpha &\approx 1 - e^{-\lambda} - \lambda e^{-\lambda} - \frac{\lambda^2}{2} e^{-\lambda} = 1 - e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2}\right) \\
&= 1 - 0.007 \times (1 + 5 + 12.5) \approx 0.87.
\end{aligned}
$$

8. 假设随机变量 $X$ 的概率密度为

$$
p(x) = \begin{cases} 2x, & \text{若 } 0 < x < 1, \\ 0, & \text{若不然}. \end{cases}
$$

现在对 $X$ 进行 $n$ 次独立重复观测，以 $V_n$ 表示观测值不大于 0.1 的次数. 试求随机变量 $V_n$ 的概率分布.

**解** 事件"观测值不大于 0.1"的概率为

$$
p = P\{X \leqslant 0.1\} = \int_{-\infty}^{0.1} p(x)\mathrm{d}x = 2\int_0^{0.1} x\mathrm{d}x = 0.01.
$$

$V_n$ 服从参数为 $(n, p)$ 的二项分布：

$$
P\{V_n = m\} = C_n^m (0.01)^m (0.99)^{n-m},\ m = 0, 1, 2, \cdots, n.
$$

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

(1) 求相继两次故障之间时间间隔 $T$ 的概率分布；

(2) 求在设备已无故障工作 8 小时的情形下，再无故障运行 8 小时的概率 $Q$.

**解** (1) 当 $t < 0$ 时，由于 $T$ 是非负随机变量，

$$
F(t) = P\{T \leqslant t\} = 0.
$$

当 $t \geqslant 0$ 时，由于事件 $\{T > t\}$ 与 $\{N(t) = 0\}$ 等价，

$$
\begin{aligned}
F(t) &= P\{T \leqslant t\} = 1 - P\{T > t\} \\
&= 1 - P\{N(t) = 0\} = 1 - e^{\lambda t}.
\end{aligned}
$$

于是，$T$ 服从参数为 $\lambda$ 的指数分布

(2) $Q=P\{T\geqslant 16 \mid T\geqslant 8\}=\dfrac{P\{T\geqslant 16,T\geqslant 8\}}{P\{T\geqslant 8\}}=\dfrac{P\{T\geqslant 16\}}{P\{T\geqslant 8\}}$

$=\dfrac{e^{-16\lambda}}{e^{-8\lambda}}=e^{-8\lambda}$.

10. 假设随机变量 $X$ 服从参数为 2 的指数分布，证明：$Y=1-e^{-2x}$在区间(0,1)上服从均匀分布.

**解** 由题意知 $X$ 的分布函数 $F(x)=\begin{cases}1-e^{-2x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$设 $G(y)=P\{Y\leqslant y\}$为 $Y$ 的分布函数，由于 $X>0$，有 $0<Y=1-e^{-2x}<1$，易得：

(1) 当 $y\leqslant 0$ 时，$G(y)\equiv 0$；

(2) 当 $y\geqslant 1$ 时，$G(y)\equiv 1$；

(3) 当 $0<y<1$ 时，

$$\begin{aligned}G(y)&=P\{Y\leqslant y\}=P\{1-e^{-2X}\leqslant y\}\\&=P\{e^{-2X}\geqslant 1-y\}=P\left\{X\leqslant -\frac{1}{2}\ln(1-y)\right\}\\&=F\left(-\frac{1}{2}\ln(1-y)\right)=y.\end{aligned}$$

总之有

$$G(y)=\begin{cases}0, & 若\ y\leqslant 0,\\ y, & 若\ 0<y<1,\\ 1, & 若\ y>1.\end{cases}$$

所以 $Y$ 在区间(0,1)上服从均匀分布.

11. 设随机变量 $X$ 服从正态分布 $N(60,3^2)$，求分点 $x_1,x_2$，使 $X$ 分别落在$(-\infty,x_1)$，$(x_1,x_2)$，$(x_2,+\infty)$的概率之比为 3:4:5.

**解** 因为 $X\sim N(60,3^2)$，所以

$$P\left\{\frac{X-60}{3}<\frac{x_1-60}{3}\right\}=P\{X<x_1\}=\frac{3}{3+4+5}=0.25,$$

即 $\Phi\left(\dfrac{x_1-60}{3}\right)=0.25$. 查标准正态分布表，得$\dfrac{x_1-60}{3}=-0.675$，

于是 $$x_1=-0.675\times3+60=57.975.$$

同理

$$P\left\{\frac{X-60}{3}<\frac{x_2-60}{3}\right\}=P(X<x_2)=\frac{3+4}{3+4+5}=0.5833,$$

查表得 $\frac{x_2-60}{3}=0.21$，于是

$$x_2=0.21\times3+60=60.63.$$

12. 假设一电路装有三个同种电气元件，其工作状态相互独立，且无故障工作时间都服从参数为 $\lambda>0$ 的指数分布. 当三个元件都无故障时，电路正常工作，否则整个电路不能正常工作. 试求电路正常工作的时间 $T$ 的概率分布.

**解** 以 $X_i(i=1,2,3)$ 表示第 $i$ 个元件无故障的时间，则 $X_1$，$X_2$，$X_3$ 相互独立同分布，其分布函数为

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

设 $G(t)$ 是 $T$ 的分布函数，当 $t\leqslant 0$ 时，$G(t)=0$；当 $t>0$ 时，

$$\begin{aligned}G(t)&=P\{T\leqslant t\}=1-P\{T>t\}\\&=1-P\{X_1>t,X_2>t,X_3>t\}\\&=1-P\{X_1>t\}\cdot P\{X_2>t\}\cdot P\{X_3>t\}\\&=1-[1-F(t)]^3\\&=1-e^{-3\lambda t},\end{aligned}$$

所以 $$G(t)=\begin{cases}1-e^{-3\lambda t}, & t>0,\\ 0, & t=0.\end{cases}$$

$T$ 服从参数为 $3\lambda$ 的指数分布.

13. 假设随机变量 $X$ 的绝对值不大于 1；$P\{X=-1\}=\frac{1}{8}$，$P\{X=1\}=\frac{1}{4}$；在事件 $\{-1<X<1\}$ 出现的条件下，$X$ 在 $(-1,1)$ 内在任一子区间上取值的条件概率与该子区间长度成正比. 试求：

(1) $X$ 的分布函数 $F(x)=P\{X\leqslant x\}$；

(2) $X$ 取负值的概率 $p$.

**解** (1) 据已知,有 $x<-1$ 时,$F(x)=0$;$x\geqslant 1$ 时,$F(x)=1$. 以下考虑 $-1<x<1$ 时的情形. 由于

$$\begin{aligned}1&=P\{|X|\leqslant 1\}\\&=P\{X=-1\}+P\{-1<X<1\}+P\{X=1\},\end{aligned}$$

故
$$P\{-1<X<1\}=1-\frac{1}{8}-\frac{1}{4}=\frac{5}{8}.$$

另据条件,有

$$P\{-1<X\leqslant x\mid -1<X<1\}=\frac{1}{2}(x+1).$$

于是,对于 $-1<x<1$,有 $(-1,x]\subset(-1,1)$,因此

$$\begin{aligned}P\{-1<X\leqslant x\}&=P\{-1<X\leqslant x,\ -1<X<1\}\\&=P\{-1<X<1\}\\&\quad\cdot P\{-1<X\leqslant x\mid -1<X<1\}\\&=\frac{5}{8}\times\frac{1}{2}(x+1)=\frac{5}{16}(x+1),\end{aligned}$$

$$\begin{aligned}F(x)&=P\{X\leqslant -1\}+P\{-1<X\leqslant x\}\\&=\frac{1}{8}+\frac{5}{16}(x+1)=\frac{5x+7}{16}.\end{aligned}$$

综上,有
$$F(x)=\begin{cases}0, & x<-1,\\(5x+7)/16, & -1\leqslant x<1,\\1, & x\geqslant 1.\end{cases}$$

14. 设随机变量 $X$ 的概率密度函数为 $p_X(x)=\dfrac{1}{\pi(1+x^2)}$,求随机变量 $Y=1-\sqrt[3]{X}$ 的概率密度函数 $p_Y(y)$.

**解** $Y$ 的分布函数

$$\begin{aligned}F_Y(y)&=P\{Y<y\}=P\{1-\sqrt[3]{X}<y\}=P\{\sqrt[3]{X}>1-y\}\\&=P\{X>(1-y)^3\}=\int_{(1-y)^3}^{+\infty}\frac{\mathrm{d}x}{\pi(1+x^2)}\\&=\frac{1}{\pi}\arctan x\Big|_{(1-y)^3}^{+\infty}=\frac{1}{\pi}\left[\frac{\pi}{2}-\arctan(1-y)^3\right],\end{aligned}$$

因此,$Y$ 的概率密度函数为

$$p_Y(y)=\frac{\mathrm{d}}{\mathrm{d}y}F_Y(y)=\frac{3}{\pi}\cdot\frac{(1-y)^2}{1+(1-y)^6}.$$

15. 从一批有 13 个正品和 2 个次品的产品中任意取 3 个，求抽得的次品数 $X$ 的分布列和分布函数，并求 $P\left\{\frac{1}{2}<X\leqslant\frac{5}{2}\right\}$.

**解** 先求 $X$ 的分布列，$X$ 的所有可能取值为 0，1，2，由古典概型的概率计算公式知

$$P\{X=0\}=\frac{C_{13}^3}{C_{15}^3}=\frac{22}{35},\quad P\{X=1\}=\frac{C_2^1C_{13}^2}{C_{15}^2}=\frac{12}{35},$$

$$P\{X=2\}=\frac{C_2^2C_{13}^1}{C_{15}^3}=\frac{1}{35}.$$

故 $X$ 的分布律为：

| $x_i$ | 0 | 1 | 2 |
|---|---|---|---|
| $p_i$ | $\frac{22}{35}$ | $\frac{12}{35}$ | $\frac{1}{35}$ |

为了求 $X$ 的分布函数 $F(x)$，我们将 $(-\infty,+\infty)$ 分成 $(-\infty,0)$，$[0,1)$，$[1,2)$，$[2,+\infty)$ 四个区间.

当 $x<0$ 时，$F(x)=P\{X\leqslant x\}=0$；

当 $0\leqslant x<1$ 时，$F(x)=P\{X=0\}=\frac{22}{35}$；

当 $1\leqslant x<2$ 时，$F(x)=P\{X=0\}+P\{X=1\}=\frac{34}{35}$；

当 $x\geqslant2$ 时，$F(x)=P\{X=0\}+P\{X=1\}+P\{X=2\}=1.$

综上有 $X$ 的分布函数为

$$F(x)=\begin{cases}0, & x<0,\\ \frac{22}{35}, & 0\leqslant x<1,\\ \frac{34}{35}, & 1\leqslant x<2,\\ 1, & x\geqslant 2.\end{cases}$$

由分布函数可求出

$$P\left\{\frac{1}{2}<X\leqslant\frac{5}{2}\right\}=F\left(\frac{5}{2}\right)-F\left(\frac{1}{2}\right)=1-\frac{22}{35}=\frac{13}{35}.$$

16. 设连续型随机变量 $X$ 的分布函数

$$F(x)=\begin{cases}A+Be^{-\frac{x^2}{2}}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

求系数 $A$ 和 $B$

**解** 由 $\lim\limits_{x\to+\infty}F(x)=1$，知 $A=1$. 再由 $F(x)$在 $x=0$ 处的右连续性知

$$0=\lim_{x\to 0}F(x)=\lim_{x\to 0}(A+Be^{-\frac{x^2}{2}})=A+B.$$

故 $$B=-A=-1.$$

17. 设随机变量 $X$ 的分布密度函数为

$$p(x)=\begin{cases}\dfrac{A}{\sqrt{1-x^2}}, & |x|<1,\\ 0, & |x|\geqslant 1.\end{cases}$$

试求：(1) 系数 $A$； (2) $P\left\{|X|<\frac{1}{2}\right\}$； (3) 分布函数 $F(x)$.

**解** (1) 由 $p(x)$的性质有

$$1=\int_{-\infty}^{+\infty}p(x)dx=\int_{-1}^{1}\frac{A}{\sqrt{1-x^2}}dx=2A\arcsin x\Big|_0^1$$

$$=2A\cdot\frac{\pi}{2}=\pi A.$$

所以 $A=\frac{1}{\pi}$.

(2) $P\left\{|X|<\frac{1}{2}\right\}=\int_{-1/2}^{1/2}\frac{1}{\pi}\frac{1}{\sqrt{1-x^2}}dx=\frac{2}{\pi}\arcsin x\Big|_0^{\frac{1}{2}}=\frac{1}{3}$.

(3) $F(x)=\int_{-\infty}^{x}p(t)dt$.

当 $x<-1$ 时，$F(x)=\int_{-\infty}^{x}0dt=0$；

当 $-1\leqslant x<1$ 时,

$$F(x)=\int_{-1}^{x}\frac{1}{\pi}\frac{1}{\sqrt{1-t^2}}\mathrm{d}t=\frac{1}{\pi}\arcsin t\Big|_{-1}^{x}$$

$$=\frac{1}{\pi}\arcsin x+\frac{1}{2};$$

当 $x\geqslant 1$ 时,

$$F(x)=\int_{-1}^{1}\frac{1}{\pi}\frac{1}{\sqrt{1-t^2}}\mathrm{d}t=\frac{1}{\pi}\arcsin t\Big|_{-1}^{1}=1.$$

故 $X$ 分布函数为

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

18. 设随机变量 $X$ 的分布函数为

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

求 $X$ 的分布列.

**解** $P\{X=-1\}=F(-1)-F(-1-0)=\frac{1}{4}-0=\frac{1}{4}$,

$P\{X=0\}=F(0)-F(0-0)=\frac{3}{4}-\frac{1}{4}=\frac{1}{2}$,

$P\{X=1\}=F(1)-F(1-0)=1-\frac{3}{4}=\frac{1}{4}$.

故 $X$ 的分布列为

| $x_i$ | $-1$ | $0$ | $1$ |
|---|---|---|---|
| $p_i$ | $\frac{1}{4}$ | $\frac{1}{2}$ | $\frac{1}{4}$ |

19. 公共汽车车门的高度,是按男子与车门碰头的机会在

0.01 以下来设计的，设男子身高 $X$ 服从 $\mu=168$ cm，$\sigma=7$ cm 的正态分布，即 $X\sim N(168,7^2)$，问车门的高度应如何确定？

**解** 若车门的高度为 $h$cm，由题意：

$$P\{X\geqslant h\}\leqslant 0.01 \quad \text{或} \quad P\{X<h\}\geqslant 0.99.$$

由于 $X\sim N(168,7^2)$，因此

$$P\{X<h\}=\Phi\left(\frac{h-168}{7}\right)\geqslant 0.99.$$

由查表可知

$$\Phi(2.33)\approx 0.9901>0.99.$$

即有
$$\frac{h-168}{7}=2.33.$$

于是
$$h=168+7\times 2.33=184.31.$$

故车门的高度为184.31 cm 时，男子与车门碰头的机会不超过 0.01.

20. 某地抽样调查，考生的英语成绩(按百分制计算，近似服从正态分布)，平均成绩为 72 分，96 分以上的占考生总数的 2.3%，试求考生的英语成绩在 60 分到 84 分之间的概率？

**解** 此题是关于考生的英语成绩的，由题意它是近似服从正态分布的，因此它的概率密度曲线是一条中间高两边低呈对称分布的钟形曲线，我们在讲正态分布的概念时，已提到参数 $\mu$ 在实际上体现了随机变量取值在概率意义下的平均值. 在此题中，我们把它当成已知条件，即 $\mu$ 为平均成绩 72. 因此，我们设 $X$ 为考生的英语成绩，由题意知

$$X\sim N(\mu,\sigma^2),$$

其中 $\mu=72$，现确定 $\sigma$. 由题设，

$$P\{X\geqslant 96\}=0.023,$$

由正态分布的查表法，把非标准正态分布转换成标准正态分布，即有

$$P\left\{\frac{X-\mu}{\sigma}\geqslant\frac{96-72}{\sigma}\right\}=0.023,$$

故

$$1-\Phi\left(\frac{24}{\sigma}\right)=0.023,\quad \Phi\left(\frac{24}{\sigma}\right)=0.977.$$

查表可知 $\frac{24}{\sigma}=2$. 所以 $\sigma=12$. 因此,

$$X \sim N(72,12^2).$$

所求的概率为

$$\begin{aligned}P\{60\leqslant X\leqslant 84\} &= P\left\{\frac{60-72}{12}\leqslant\frac{X-\mu}{\sigma}\leqslant\frac{84-72}{12}\right\}\\ &= P\left\{-1\leqslant\frac{X-\mu}{\sigma}\leqslant 1\right\}=\Phi(1)-\Phi(-1)\\ &= 2\Phi(1)-1=0.6826.\end{aligned}$$

21. 某地 1987 年全国高校统考物理成绩 $X$ 服从正态分布 $N(42,6^2)$,如考生得 48 分,求有多少考生名列该考生之后?

**解** 由 $X\sim N(42,6^2)$,则 $Y=\frac{X-42}{6}\sim N(0,1)$,因此,

$$\begin{aligned}P\{X>48\} &= P\{Y>1\}=1-P\{Y\leqslant 1\}\\ &= 1-\Phi(1)=0.16.\end{aligned}$$

这说明有 16%的考生成绩超过 48 分,因而有 84%的考生,名列得 48 分的考生之后.

22. 某企业准备通过招聘考试招收 300 名职工,其中正式工 280 人,临时工 20 人;报考的人数是 1657 人,考试满分是 400 分.考试后得知,考试总平均成绩,即 $\mu=166$ 分,360 分以上的高分考生 31 人.某考生 $B$ 得 256 分,问他能否被录取?能否被聘为正式工?

**解** 分两步来解答.*第一步* 预测最低分数线.设最低分数线为 $x_1$,考生成绩为 $\xi$,则对一次成功的考试来说,$\xi$ 服从正态分布,由题意可知:

$$\xi\sim N(166,\sigma^2),$$

这样,$\eta=\frac{\xi-166}{\sigma}\sim N(0,1)$,因为高于 360 分的考生的频率是

$\frac{31}{1657}$,故

$$P\{\xi > 360\} = P\left\{\eta > \frac{360-166}{\sigma}\right\} \approx \frac{31}{1657},$$

因此 $$P\left\{\eta \leqslant \frac{360-166}{\sigma}\right\} \approx 1-\frac{31}{1657} \approx 0.981.$$

查表可知

$$\frac{360-166}{\sigma} \approx 2.08,$$

即 $$\sigma \approx 93.$$

故 $\xi \sim N(166, 93^2)$. 因为最低分数线的确定应使录取的考生的频率等于$\frac{300}{1657}$,即

$$P\left\{\eta > \frac{x_1-166}{93}\right\} \approx \frac{300}{1657},$$

所以

$$P\left\{\eta \leqslant \frac{x_1-166}{93}\right\} \approx 1-\frac{300}{1657} \approx 0.819.$$

查表得$\frac{x_1-166}{93} \approx 0.91$. 由此求得 $x_1 \approx 251$. 也就是说,最低分数线是 251 分.

*第二步* 预测 $B$ 的考试名次,这样就可以确定他是否能被录取. 在 $\xi=256$ 分时,由查表可知:

$$P\left\{\eta \leqslant \frac{x-166}{93}\right\} = P\left\{\eta \leqslant \frac{256-166}{93}\right\} \approx 0.835.$$

这样

$$P\left\{\eta > \frac{256-166}{93}\right\} \approx 1-0.8315 = 0.1685.$$

这表明,考试成绩高于 256 分的频率是 0.1685,也就是成绩高于考生 $B$ 的人数大约占总考生 16.85%,所以名次排在考生 $B$ 之前的考生人数约有:

$$1657 \times 16.85\% \approx 280.$$

即考生 $B$ 大约排在 281 名.

由于一共招收 300 名,故考生 $B$ 可以被录取,但正式工只招 280 名,而 $281>280$,故考生 $B$ 被录为临时工的可能性很大.

23. 已知某批建筑材料的强度 $X$ 服从 $N(200,18^2)$,现从中任取一件时,求

(1) 取得这件材料的强度不低于 180 的概率;

(2) 如果所用的材料要求以 99%的概率保证强度不低于 150,问这批材料是否符合这个要求.

**解** (1) 
$$\begin{aligned}P\{X\geqslant 180\}&=1-P\{X<180\}\\&=1-\Phi\left(\frac{180-200}{18}\right)=\Phi(1.11)\\&=0.8665.\end{aligned}$$

(2) 
$$\begin{aligned}P\{X\geqslant 150\}&=1-P\{X<150\}\\&=1-\Phi\left(\frac{150-200}{18}\right)=1-\Phi(-2.78)\\&=\Phi(2.78)=0.9973.\end{aligned}$$

即从这批材料中任取一件,以概率 99.73%(大于 99%)保证强度不低于 150,故这批材料符合所提出的要求.

24. 如果电源电压在不超过 200 V,在 200~240 V 之间和超过 240 V 三种情况下,某种电子元件损坏的概率分别为 0.1, 0.001 和 0.2,并假设电源电压 $X$ 服从 $N(220,25^2)$,求(1)该电子元件损坏的概率 $p_1$;(2) 该电子元件损坏时,电源电压在 220 V~240 V 的概率 $p_2$.

**解** 设 $A_1=\{$电压不超过 200 V$\}$,$A_2=\{$电压在 200 V~240 V$\}$,$A_3=\{$电压超过 240 V$\}$,$B=\{$电子元件损坏$\}$. 由题意知: $X\sim N(220,25^2)$,故

$$\begin{aligned}P\{A_1\}&=P\{X<200\}=P\left\{\frac{X-220}{25}<\frac{200-220}{25}\right\}\\&=\Phi(-0.8)=0.212.\end{aligned}$$

$$P\{A_2\}=P\{200\leqslant X\leqslant 240\}=\Phi(0.8)-\Phi(-0.8)$$

$= 0.576.$

$P\{A_3\} = P\{X > 240\} = 1 - 0.212 - 0.576 = 0.212.$

(1) 由题意,

$P\{B|A_1\}=0.1,\quad P\{B|A_2\}=0.001,\quad P\{B|A_3\}=0.2.$

再由全概率公式,

$$p_1 = P(B) = \sum_{i=1}^{3} P(A_i)P(B|A_i) = 0.0642.$$

(2) 由贝叶斯公式:

$$p_2 = P(A_2|B) = \frac{P(A_2)P(B|A_2)}{P(B)} \approx 0.009.$$

25. 设 $X$ 的分布列为

$$\begin{bmatrix} -1 & 0 & 1 & 2 & \frac{5}{2} \\ \frac{1}{5} & \frac{1}{10} & \frac{1}{10} & \frac{3}{10} & \frac{3}{10} \end{bmatrix},$$

试求:(1) $2X$ 的分布列;(2) $X^2$ 的分布列.

**解** 先根据 $X$ 的分布列,列出下表:

| $p_k$ | $\frac{1}{5}$ | $\frac{1}{10}$ | $\frac{1}{10}$ | $\frac{3}{10}$ | $\frac{3}{10}$ |
|---|---|---|---|---|---|
| $x_k$ | $-1$ | 0 | 1 | 2 | $\frac{5}{2}$ |
| $2x_k$ | $-2$ | 0 | 2 | 4 | 5 |
| $x_k^2$ | 1 | 0 | 1 | 4 | $\frac{25}{4}$ |

(1) 由于 $2x_k(k=1,2,3,4,5)$的值全不等,所以 $2X$ 的分布列为

$$\begin{bmatrix} -2 & 0 & 2 & 4 & 5 \\ \frac{1}{5} & \frac{1}{10} & \frac{1}{10} & \frac{3}{10} & \frac{3}{10} \end{bmatrix}.$$

(2) 由于 $x_k^2(k=1,2,3,4,5)$中值 1 出现了两次,

$$P\{X^2=1\}=P\{X=-1\}+P\{X=1\}=\frac{1}{5}+\frac{1}{10}=\frac{3}{10},$$

所以 $X^2$ 的分布列为

$$\begin{bmatrix} 0 & 1 & 4 & 25/4 \\ \dfrac{1}{10} & \dfrac{3}{10} & \dfrac{3}{10} & \dfrac{3}{10} \end{bmatrix}.$$

26. 设 $X$ 的概率分布为

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

求 $Y=\sin\left(\dfrac{\pi}{2}X\right)$ 的概率分布.

**解** 因为

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

所以,$Y=\sin\left(\dfrac{\pi}{2}X\right)$ 只有 3 个可能取值 $-1,0,1$,而取这些值的概率分别为

$$\begin{aligned}P\{Y=-1\}&=P\{X=3\}+P\{X=7\}+P\{X=11\}+\cdots\\&=\frac{1}{2^3}+\frac{1}{2^7}+\frac{1}{2^{11}}+\cdots=\frac{1}{8}\cdot\frac{1}{1-\frac{1}{16}}=\frac{2}{15},\end{aligned}$$

$$\begin{aligned}P\{Y=0\}&=P\{X=2\}+P\{X=4\}+P\{X=6\}+\cdots\\&=\frac{1}{2^2}+\frac{1}{2^4}+\frac{1}{2^6}+\cdots=\frac{1}{4}\cdot\frac{1}{1-\frac{1}{4}}=\frac{1}{3},\end{aligned}$$

$$\begin{aligned}P\{Y=1\}&=P\{X=1\}+P\{X=5\}+P\{X=9\}+\cdots\\&=\frac{1}{2}+\frac{1}{2^5}+\frac{1}{2^9}+\cdots=\frac{1}{2}\cdot\frac{1}{1-\frac{1}{16}}=\frac{8}{15}.\end{aligned}$$

于是,$Y=\sin\left(\dfrac{\pi}{2}X\right)$ 的分布列为

$$\begin{pmatrix} -1 & 0 & 1 \\ \dfrac{2}{15} & \dfrac{1}{3} & \dfrac{8}{15} \end{pmatrix}.$$

27. 设随机变量 $X \sim U\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$，求随机变量 $Y=\sin X$ 的分布密度 $p_2(y)$.

**解** $X$ 的分布密度函数为

$$p_1(x)=\begin{cases} \dfrac{1}{\pi}, & x \in \left(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right), \\ 0, & \text{其他}. \end{cases}$$

因为 $y=\sin x$ 在 $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ 内单调增加，所以存在反函数

$$x=\arcsin y,$$

其导数为

$$x_y'=\frac{1}{\sqrt{1-y^2}}.$$

利用公式求出 $Y$ 的分布密度函数，首先计算

$$\alpha=\min_{-\frac{\pi}{2}\leqslant x\leqslant\frac{\pi}{2}}\{\sin x\}=-1, \quad \beta=\max_{-\frac{\pi}{2}\leqslant x\leqslant\frac{\pi}{2}}\{\sin x\}=1.$$

于是

$$p_2(y)=\begin{cases} p_1(f^{-1}(y)) \cdot |x_y'|, & -1<y<1, \\ 0, & \text{其他} \end{cases}$$

$$=\begin{cases} \dfrac{1}{\pi} \cdot \dfrac{1}{\sqrt{1-y^2}}, & -1<y<1, \\ 0, & \text{其他}. \end{cases}$$

28. 设随机变量 $X \sim N(0,1)$，求随机变量 $Y=X^2$ 的分布密度 $p_2(y)$.

**解** $X$ 的分布密度函数为

$$p_1(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \quad (-\infty<x<+\infty).$$

因为 $y=x^2$ 在 $(-\infty,+\infty)$ 内不是单调的. 于是将这个区间分成

$(-\infty,0)$和$(0,+\infty)$两个单调区间. 在$(-\infty,0)$其反函数为$x=f_1^{-1}(y)=-\sqrt{y}$,在$(0,+\infty)$内为$x=f_2^{-1}(y)=\sqrt{y}$.

由公式

$$p_2(y)=p_1[f_1^{-1}(y)]\cdot|(f_1^{-1}(y))'|+p_1[f_2^{-1}(y)]\cdot|(f_2^{-1}(y))'|,$$

以及 $$(f_1^{-1}(y))'=\frac{-1}{2\sqrt{y}},\quad (f_2^{-1}(y))'=\frac{1}{2\sqrt{y}},$$

$$p_1[f_1^{-1}(y)]=p_1[f_2^{-1}(y)]=\frac{1}{\sqrt{2\pi}}e^{-\frac{y}{2}}$$

有 $$p_2(y)=\frac{1}{\sqrt{2\pi y}}e^{-\frac{y}{2}}.$$

因为$y=x^2(-\infty<x<+\infty)$,所以$0<y<+\infty$,因此,在区间$(0,+\infty)$内有分布密度函数

$$p_2(y)=\frac{1}{\sqrt{2\pi y}}e^{-\frac{y}{2}},$$

在其他区间,则$p_2(y)=0$.

故

$$p_2(y)=\begin{cases}\dfrac{1}{\sqrt{2\pi y}}e^{-\frac{y}{2}}, & y>0,\\ 0, & 其他.\end{cases}$$

29. 设随机变量$X$在$(1,6)$上服从均匀分布,求方程$x^2+Xx+1=0$有实根的概率.

**解** 因为$X$在$(1,6)$上服从均匀分布,故$X$的分布密度为

$$p(x)=\begin{cases}\dfrac{1}{5}, & x\in(1,6),\\ 0, & 其他.\end{cases}$$

方程$x^2+Xx+1=0$有实根的条件是:

$$\Delta=X^2-4\geqslant 0,$$

解得$X\leqslant-2$或$X\geqslant 2$. 舍去$X\leqslant -2$. 最后得$2\leqslant X<6$. 因之所求

概率为

$$P\{2 \leqslant X < 6\} = \int_2^6 \frac{1}{5}\mathrm{d}x = 0.8.$$

30. 设随机变量 $X$ 的概率密度为

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

求随机变量 $Y=\mathrm{e}^X$ 的概率密度 $p_Y(y)$.

**解** $Y$ 的分布函数 $F_Y(y)=P\{Y\leqslant y\}=P\{\mathrm{e}^X\leqslant y\}$.

当 $y<1$ 时，$F_Y(y)=0$；

当 $y\geqslant 1$ 时，$F_Y(y)=P\{X\leqslant \ln y\}=\int_0^{\ln y}\mathrm{e}^{-x}\mathrm{d}x$.

因此 $Y$ 的概率密度为

$$p_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y}F_Y(y) = \begin{cases} 0, & y < 1, \\ \frac{1}{y^2}, & y \geqslant 1. \end{cases}$$

31. 设随机变量 $X$ 服从标准正态分布，即 $X\sim N(0,1)$，试求

(1) $Y=2X^2+1$；　　(2) $Z=|X|$的密度函数.

**解** 因为 $X\sim N(0,1)$，所以

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

(1) $Y=2X^2+1$，$Y$ 的分布函数 $F_Y(y)=P\{2X^2+1\leqslant y\}$.

当 $y\leqslant 1$ 时，$F_Y(y)=0$；当 $y>1$ 时，

$$\begin{aligned} F_Y(y) &= P\{2X^2+1\leqslant y\} \\ &= P\left\{-\sqrt{\frac{y-1}{2}} < X < \sqrt{\frac{y-1}{2}}\right\} \\ &= \int_{-\sqrt{\frac{y-1}{2}}}^{\sqrt{\frac{y-1}{2}}} \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x = \frac{2}{\sqrt{2\pi}}\int_0^{\sqrt{\frac{y-1}{2}}}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x, \end{aligned}$$

即

$$F_Y(y)=\begin{cases}0, & y\leqslant 1,\\ \dfrac{2}{\sqrt{2\pi}}\displaystyle\int_0^{\sqrt{\frac{y-1}{2}}}e^{-\frac{x^2}{2}}dx, & y>1.\end{cases}$$

于是 $Y$ 的概率密度函数为

$$p_Y(y)=F_Y'(y)=\begin{cases}\dfrac{1}{2\sqrt{\pi(y-1)}}e^{-\frac{y-1}{4}}, & y>1,\\ 0, & y\leqslant 1.\end{cases}$$

(2) $Z=|X|$ 的分布函数

$$F_Z(z)=P\{Z\leqslant z\}=P\{|X|\leqslant z\}.$$

当 $z<0$ 时，$F_Z(z)=0$；当 $z\geqslant 0$ 时，

$$F_Z(z)=P\{|X|\leqslant z\}=P\{-z\leqslant X\leqslant z\}$$
$$=\int_{-z}^{z}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx=\frac{2}{\sqrt{2\pi}}\int_0^z e^{-\frac{x^2}{2}}dx.$$

于是，$Z$ 的概率密度函数为

$$p_Z(z)=F_Z'(z)=\begin{cases}\sqrt{\dfrac{2}{\pi}}e^{-\frac{1}{2}z^2}, & z\geqslant 0,\\ 0, & z<0.\end{cases}$$

32. 设点随机地落在单位圆周(圆心位于原点)上，并且对弧长是均匀分布的，求这点的横坐标的密度函数.

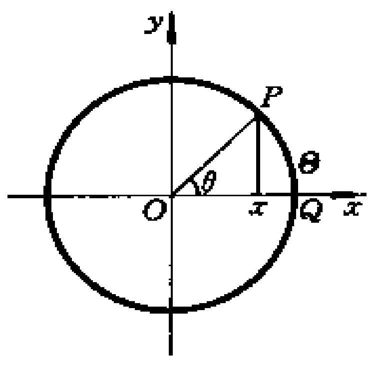



图 2.1

**解** 据题设，落在圆周上的点 $P$ 与定点 $Q$(见图 2.1)之间的弧长 $\Theta$ 的密度函数为

$$p_\Theta(\theta)=\begin{cases}\dfrac{1}{2\pi}, & 0\leqslant\theta<2\pi,\\ 0, & 其他,\end{cases}$$

试求点 $P$ 的横坐标 $X=\cos\Theta$ 的密度函数.

因为 $x=\cos\theta(0\leqslant\theta<2\pi)$ 不是单调

函数，使 $\cos\theta\leqslant x$ 成立的 $\theta$ 满足

$$\arccos x \leqslant \theta \leqslant 2\pi - \arccos x.$$

于是，对 $-1\leqslant x\leqslant 1$，有

$$F_X(x)=P\{X\leqslant x\}=P\{\cos\Theta\leqslant x\}=\int_{\cos\theta\leqslant x}p_\Theta(\theta)\mathrm{d}\theta$$

$$=\int_{\arccos x}^{2\pi-\arccos x}\frac{1}{2\pi}\mathrm{d}\theta=1-\frac{1}{\pi}\arccos x;$$

对 $x<-1$，有

$$F_X(x)=P\{X\leqslant x\}=P\{\cos\Theta\leqslant x\}=P(\varnothing)=0;$$

对 $x>1$，有

$$F_X(x)=P\{X\leqslant x\}=P\{\cos\Theta\leqslant x\}=P(\Omega)=1,$$

即

$$F(x)=\begin{cases}0, & x<-1,\\ 1-\dfrac{1}{\pi}\arccos x, & -1\leqslant x\leqslant 1,\\ 1, & x>1.\end{cases}$$

所以 $X$ 的密度函数为

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

33. 设随机变量 $X$ 具有连续的分布函数 $F(x)$，求 $Y=F(X)$ 的分布密度函数.

**解** 由于 $F(x)$ 为 $x$ 的连续分布函数，可知 $F(-\infty)=0$，$F(+\infty)=1$. 因为 $F(x)$ 是单调递增函数，所以 $F^{-1}(y)$ 存在（单调函数有单值反函数存在），因而有

$$F_Y(y)\xlongequal{\text{def}}P\{Y\leqslant y\}=\begin{cases}0, & y<0,\\ *, & 0\leqslant y<1,\\ 1, & y\geqslant 1.\end{cases}$$

当 $0\leqslant y<1$ 时，

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

代入 $F_Y(y)$ 表达式有

$$F_Y(y) = \begin{cases} 0, & y < 0, \\ y, & 0 \leqslant y < 1, \\ 1, & y \geqslant 1. \end{cases}$$

因此 $Y$ 的分布密度函数为 $p(y) = \begin{cases} 1, & 0 \leqslant y \leqslant 1, \\ 0, & 其他, \end{cases}$ 即

$$Y \sim U(0,1).$$

## 四、练习题

1. 一汽车沿着街道行驶，需通过 3 个均设有红、绿信号灯的路口. 每个信号灯为红或绿与其他信号灯为红或绿相互独立，且红绿两种信号显示的时间相等，以 $X$ 表示该汽车首次遇到红灯前已通过路口的个数，求 $X$ 的概率分布.

2. 盒中有形状与功率均相同的 10 个灯泡，其中 7 个是螺口的，3 个是卡口的，现要用 1 个螺口灯泡. 从盒中任取一个，若为卡口的就不再放回去，求取到螺口灯泡前已取出的卡口灯泡数 $X$ 的概率分布.

3. 一批产品中有 10 件正品和 3 件次品，现从中一件一件地抽取，设每次抽取时所面对的各件产品被取到的可能性相等. 在下列两种情况下，分别求出直到取到正品时为止所需的抽取次数 $X$ 的概率分布：

(1) 每次抽出的产品都不放回这批产品中；

(2) 每次抽出的产品经检验后又放回这批产品中，再取下一件产品.

4. 一口袋中有红、白、黄色球各 5 个，现从中任取 4 个，用 $X$ 表示取到的白球个数，求 $X$ 的概率分布.

5. 为了保证设备正常工作，需要配备适量的维修工人. 现有

同一类型的设备 300 台，各台设备工作相互独立，发生故障的概率都是 0.01. 在通常情况下，一台设备的故障可以由一个工人处理.

(1) 问至少需要配备多少工人，才能保证设备发生故障时不能及时维修的概率不超过 0.01?

(2) 若由一个工人负责维修 20 台设备，求设备发生故障时不能及时维修的概率.

6. 设随机变量 $X$ 服从泊松分布，且已知 $P\{X=1\}=P\{X=2\}$，求 $P\{X=4\}$.

7. 设随机变量 $X$ 的密度函数为 $p(x)=\dfrac{A}{e^x+e^{-x}}$. 求

(1) 常数 $A$；(2) $P\left\{0<X<\dfrac{1}{2}\ln 3\right\}$；(3) 分布函数 $F(x)$.

8. 设连续型随机变量 $X$ 的分布函数为

$$F(x)=\begin{cases}A+Be^{-x^2/2}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

求 (i) 常数 $A,B$；(ii) $X$ 的概率密度函数；(iii) $X$ 的取值落在区间 $(1,2)$ 内的概率.

9. 测量到某一目标的距离时出现的随机误差 $X$(以 m 计)服从 $N(10,20^2)$. 求：

(1) 在一次测量中 $X$ 的绝对值不超过 15 m 的概率；

(2) 在 3 次测量中至少有一次 $X$ 的绝对值不超过 15 m 的概率.

10. 设随机变量的密度函数为

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

求分布函数 $F(x)$.

11. 设随机变量 $X$ 的密度函数为

$$p(x)=\begin{cases}\dfrac{A}{\sqrt{1-x^2}}, & |x|<1,\\ 0, & |x|\geqslant 1,\end{cases}$$

求(1) 常数 $A$； (2) $P\left\{|X|<\frac{1}{2}\right\}$； (3) 分布函数 $F(x)$.

12. 设轰炸机向敌方某铁路投弹，炸弹落弹点与铁路的距离 $X$ 的密度函数为

$$p(x)=\begin{cases}\dfrac{100-|x|}{10000}, & |x|\leqslant 100,\\ 0, & |x|>100.\end{cases}$$

如炸弹落在铁路两旁 40 m 以内，将招致铁路交通的破坏，现投弹 3 颗，求敌方铁路交通受到破坏的概率.

13. 设 $X\sim N(\mu,\sigma^2)$，求 $Y=e^X$ 的密度函数.

14. 设随机变量 $X$ 在区间(1,2)上服从均匀分布，求 $Y=e^{2X}$ 的密度函数.

15. 设随机变量 $X$ 具有下列概率分布：

| $X$ | $-2$ | $-1$ | $0$ | $1$ | $2$ |
|---|---|---|---|---|---|
| $p_k$ | $\frac{1}{5}$ | $\frac{1}{6}$ | $\frac{1}{5}$ | $\frac{1}{15}$ | $\frac{11}{30}$ |

试求随机变量 $Y=X^2$ 的概率分布.

16. 设随机变量 $X$ 的密度函数为

$$p(x)=\begin{cases}0, & x<0,\\ x^3e^{-x^2}, & x\geqslant 0,\end{cases}$$

求(1) $Y_1=2X+3$；(2) $Y_2=X^2$；(3) $Y_3=\ln X$ 的密度函数.

17. 设随机变量 $X$ 在$(0,2\pi)$内服从均匀分布，求随机变量 $Y=\cos X$ 的分布密度函数.

18. 设随机变量 $X$ 服从指数分布：

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

试求下列随机变量的密度：(1) $\sqrt{X}$；(2) $\frac{1}{\lambda}\ln X$；(3) $e^{-X}$.

# 习题答案与提示

**1.**

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $p$ | $\frac{1}{2}$ | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{8}$ |

**2.**

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $p$ | $\frac{7}{10}$ | $\frac{7}{30}$ | $\frac{7}{120}$ | $\frac{1}{120}$ |

**3.** (1)

| $X$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $p$ | $\frac{10}{13}$ | $\frac{5}{26}$ | $\frac{5}{143}$ | $\frac{1}{286}$ |

(2) $P\{X=k\}=\left(\frac{3}{13}\right)^{k-1}\frac{10}{13}$,其中 $k=1,2,\cdots$.

**4.**

| $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P$ | $\frac{2}{13}$ | $\frac{40}{91}$ | $\frac{30}{91}$ | $\frac{20}{273}$ | $\frac{1}{273}$ |

**5.** 8; 0.0175.

**6.** $\frac{2}{3}e^{-2}$.　　**7.** $\frac{2}{\pi}$; $\frac{1}{6}$; $F(x)=\frac{2}{\pi}\arctan e^{x}$.

**8.** (1) 1, −1;

(2) $p(x)=\begin{cases} xe^{-\frac{1}{2}x^2}, & x>0, \\ 0, & x\leqslant 0. \end{cases}$

(3) $e^{-\frac{1}{2}}-e^{-1}=0.4712$.

**9.** 0.4931; 0.8698.

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

11. (1) $\dfrac{1}{\pi}$; (2) $\dfrac{1}{3}$;

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

12. 0.9533.

13. $Y=e^X$ 的密度函数为

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

这一分布称为对数正态分布.

14. $Y=e^{2X}$的密度函数为

$$p_Y(y)=\begin{cases}p_X\left(\dfrac{1}{2}\ln y\right)\left|\left(\dfrac{1}{2}\ln y\right)'\right|, & e^2<y<e^4, \\ 0, & \text{其他},\end{cases}$$

$$=\begin{cases}\dfrac{1}{2y}, & e^2<y<e^4, \\ 0, & \text{其他}.\end{cases}$$

15.

| $Y$ | 0 | 1 | 4 |
|---|---|---|---|
| $p$ | $\frac{1}{5}$ | $\frac{7}{30}$ | $\frac{17}{30}$ |

16. (1) $p_{Y_1}(y)=\dfrac{1}{2}\left(\dfrac{y-3}{2}\right)^3 e^{-\left(\frac{y-3}{2}\right)^2}(y\geqslant 3)$;

(2) $p_{Y_2}(y)=\frac{1}{2}ye^{-y}(y>0)$;

(3) $p_{Y_3}(y)=e^{4y}e^{-e^{2y}}(-\infty<y<+\infty)$.

**17.** $p_Y(y)=\frac{1}{\pi\sqrt{1-y^2}}\ (|y|<1)$.

**18.** (1) $2\lambda ye^{-\lambda y^2}\ (y>0)$;

(2) $\lambda^2e^{\lambda[y-e^{\lambda y}]}\ (-\infty<y<+\infty)$;

(3) $\lambda y^{\lambda-1}\ (0<y<1)$.

# 第三章　多维随机变量

## 一、考试要求

1. 理解二维随机变量联合概率分布及其性质，会利用二维随机变量联合概率分布计算有关的概率；

2. 理解二维随机变量的边缘分布和条件分布，会利用二维随机变量联合概率分布计算有关的概率；

3. 理解随机变量的独立性，掌握二维随机变量独立的条件；

4. 了解二维正态分布，掌握二维均匀分布，会计算与随机变量有关的概率；

5. 会求两个独立随机变量的简单函数的分布.

## 二、复习要点

### （一）重要概念及性质

**定义 3.1**（$n$ 维随机变量）　我们把 $n$ 个随机变量 $X_1,X_2,\cdots,X_n$ 作为一个整体来考察称为一个 **$n$ 维随机变量**或 **$n$ 维随机向量**，记为 $\xi=(X_1,X_2,\cdots,X_n)$，其中 $X_i$ 称为 $\xi$ 的第 $i$ 个分量.

本章主要讨论二维随机向量及其“联合”分布. 我们用 $\xi=(X,Y)$ 表示一般二维随机向量，$X$ 为第一个分量，$Y$ 为第二个分量. 与一维随机变量的情形相类似，对于二维随机向量，我们也只讨论离散型和连续型两大类.

（1）二维离散型随机变量联合概率分布及其性质

如果二维随机向量 $(X,Y)$ 的所有可能取值为至多可列个有序对 $(x,y)$ 时，则称 $\xi$ 为**离散型随机向量**.

设 $\xi=(X,Y)$ 的所有可能取值为 $(x_i,y_j)(i,j=1,2,\cdots)$，且事

件$\{\xi=(x_i,y_j)\}$的概率为 $p_{ij}$，称

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

为 $\xi=(X,Y)$的**分布律**或称为 $X$ 和 $Y$ 的**联合分布律**. 联合分布有时也用下面的概率分布表来表示：

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

这里 $p_{ij}$具有下面两个性质：

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

2) $\sum\limits_i\sum\limits_j p_{ij}=1$.

(2) 二维连续型随机变量联合概率密度及其性质

对于二维随机向量 $\xi=(X,Y)$，如果存在非负函数 $p(x,y)$ $(-\infty<x<+\infty,-\infty<y<+\infty)$，使对任意一个其邻边分别平行于坐标轴的矩形区域 $D$，即 $D=\{(x,y)\,|\,a<x<b,c<y<d\}$有

$$P\{(X,Y)\in D\}=\iint\limits_D p(x,y)\mathrm{d}x\mathrm{d}y,$$

则称 $\xi$ 为**连续型随机向量**；并称 $p(x,y)$为 $\xi=(X,Y)$的**分布密度**或称为 $X$ 和 $Y$ 的**联合分布密度**.

分布密度 $p(x,y)$具有下面两个性质：

1) $p(x,y)\geqslant 0$；

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

下面介绍两种常见的连续型随机向量的分布：

(1) 均匀分布

设二维随机向量$(X,Y)$的分布密度函数为

$$p(x,y)=\begin{cases}\dfrac{1}{S_D}, & (x,y)\in D,\\ 0, & \text{其他},\end{cases}$$

其中$S_D$为区域$D$的面积. 则称$(X,Y)$服从$D$上的**均匀分布**,记为$(X,Y)\sim U(D)$.

(2) 正态分布

设随机向量$(X,Y)$的分布密度函数为

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

其中$\mu_1,\mu_2,\sigma_1>0,\sigma_2>0,|\rho|<1$是5个参数. 则称$(X,Y)$服从**二维正态分布**,记为$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$.

**例1** 设$(X,Y)$的联合分布密度为

$$p(x,y)=\begin{cases}Ce^{-(x+y)}, & x\geqslant 0,y\geqslant 0,\\ 0, & \text{其他}.\end{cases}$$

试求:(1) 常数$C$;(2) $P\{0<X<1,0<Y<1\}$.

**解** (1) 由$p(x,y)$的性质,有

$$1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x\mathrm{d}y=\int_0^{+\infty}\int_0^{+\infty}Ce^{-(x+y)}\mathrm{d}x\mathrm{d}y$$

$$=C\cdot\int_0^{+\infty}e^{-x}\mathrm{d}x\cdot\int_0^{+\infty}e^{-y}\mathrm{d}y=C,$$

即$C=1$.

(2) 令$D=\{(x,y)\mid 0<x<1,0<y<1\}$,有

$$P\{0<X<1,0<Y<1\}=P\{(X,Y)\in D\}$$

$$=\iint_D p(x,y)\mathrm{d}x\mathrm{d}y=\iint_D e^{-(x+y)}\mathrm{d}x\mathrm{d}y$$

$$=\int_0^1 e^{-x}\mathrm{d}x\int_0^1 e^{-y}\mathrm{d}y=\left(1-\frac{1}{e}\right)^2.$$

1. 边缘分布

对于二维随机向量$(X,Y)$，称其分量$X$（或$Y$）的分布为$(X,Y)$的关于$X$（或$Y$）的**边缘分布**.

当$(X,Y)$为离散型，并且其联合分布律为

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

则$X$的边缘分布为

$$P\{X=x_i\}=\sum_j p_{ij}\quad (i=1,2,\cdots);$$

$Y$的边缘分布为

$$P\{Y=y_j\}=\sum_i p_{ij}\quad (j=1,2,\cdots).$$

**例 2** 设二维随机向量$(X,Y)$共有六个取正概率的点，它们是：$(1,-1)$，$(2,-1)$，$(2,0)$，$(2,2)$，$(3,1)$，$(3,2)$，并且$(X,Y)$取得它们的概率相同，则$(X,Y)$的联合分布为：

$$P\{(X,Y)=(1,-1)\}=\frac{1}{6};\quad P\{(X,Y)=(2,-1)\}=\frac{1}{6};$$

$$P\{(X,Y)=(2,0)\}=\frac{1}{6};\quad P\{(X,Y)=(2,2)\}=\frac{1}{6};$$

$$P\{(X,Y)=(3,1)\}=\frac{1}{6};\quad P\{(X,Y)=(3,2)\}=\frac{1}{6}.$$

其概率分布表为

| $X$ \ $Y$ | $-1$ | $0$ | $1$ | $2$ |
|---|---|---|---|---|
| 1 | $\frac{1}{6}$ | 0 | 0 | 0 |
| 2 | $\frac{1}{6}$ | $\frac{1}{6}$ | 0 | $\frac{1}{6}$ |
| 3 | 0 | 0 | $\frac{1}{6}$ | $\frac{1}{6}$ |

而$X$和$Y$的边缘分布为

$$P\{X=1\}=\sum_{j=1}^{4}p_{1j}=\frac{1}{6}+0+0+0=\frac{1}{6},$$

即$(X,Y)$的分布表中第一行的数值之和. 类似地，有

$$P\{X=2\}=\frac{1}{2},\quad P\{X=3\}=\frac{1}{3}.$$

于是

| $X$ | 1 | 2 | 3 |
|---|---|---|---|
| $P(X=x_k)$ | $\frac{1}{6}$ | $\frac{1}{2}$ | $\frac{1}{3}$ |

同样有

| $Y$ | $-1$ | 0 | 1 | 2 |
|---|---|---|---|---|
| $P(Y=y_k)$ | $\frac{1}{3}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{3}$ |

当$(X,Y)$为连续型随机向量，并且其联合分布密度为$p(x,y)$，则$X$和$Y$的边缘分布密度为

$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y,\quad p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x.$$

例如，例1中的$X$和$Y$的边缘分布为

$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y=\begin{cases}\int_{-\infty}^{+\infty}\mathrm{e}^{-(x+y)}\mathrm{d}y, & x\geqslant 0,\\ \int_{-\infty}^{+\infty}0\mathrm{d}y, & x<0\end{cases}$$

$$=\begin{cases}\mathrm{e}^{-x}, & x\geqslant 0,\\ 0, & x<0.\end{cases}$$

即$X\sim\Gamma(1,1)$；同理$Y\sim\Gamma(1,1)$. 又如，当$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$时，可以推出$X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$. 即二维正态分布的边缘分布是一维正态分布.

**例3** 设二维随机变量$(X,Y)\sim U(D)$，其中$D$：$\begin{cases}0\leqslant x\leqslant 1,\\ 0\leqslant y\leqslant x.\end{cases}$求出$X$与$Y$的边缘分布密度$p_1(x)$与$p_2(y)$.

**解** 由均匀分布的定义

$$p(x,y)=\begin{cases}\dfrac{1}{S_D}, & (x,y)\in D,\\ 0, & \text{其他}.\end{cases}$$

考虑到 $S_D=\dfrac{1}{2}$，因此

$$p(x,y)=\begin{cases}2, & (x,y)\in D,\\ 0, & \text{其他}.\end{cases}$$

为了求出 $X$ 的边缘分布，我们分两种情况来讨论：

(1) 当 $x<0$ 或 $x>1$ 时，$p(x,y)\equiv 0$，于是

$$p_1(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y=0;$$

(2) 当 $0\leqslant x\leqslant 1$ 时，只有 $0\leqslant y\leqslant x$ 时，$p(x,y)=2$，于是

$$\begin{aligned}p_1(x)&=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y\\&=\int_{-\infty}^{0}0\mathrm{d}y+\int_0^x 2\mathrm{d}y+\int_x^{+\infty}0\mathrm{d}y=2x.\end{aligned}$$

所以

$$p_1(x)=\begin{cases}2x, & 0\leqslant x\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

同理，可求出 $Y$ 的边缘分布

$$p_2(y)=\begin{cases}2(1-y), & 0\leqslant y\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

2. 条件分布

当二维随机向量 $(X,Y)$ 为离散型，并且其联合分布律为

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

**则在已知 $Y=y_j$ 的条件下，$X$ 取值的条件分布为**

$$P\{X=x_i|Y=y_j\}=\frac{p_{ij}}{p_j};$$

**在已知 $X=x_i$ 的条件下，$Y$ 取值的条件分布为**

$$P\{Y=y_j|X=x_i\}=\frac{p_{ij}}{p_i},$$

其中 $p_i, p_j$ 分别为 $X$、$Y$ 的边缘分布.

当二维随机向量$(X,Y)$为连续型随机向量，并且其联合分布密度为 $p(x,y)$，则在已知 $Y=y$ 的条件下，$X$ 的条件分布密度为

$$p(x|y)=\frac{p(x,y)}{p_2(y)};$$

在已知 $X=x$ 的条件下，$Y$ 的条件分布密度为

$$p(y|x)=\frac{p(x,y)}{p_1(x)},$$

其中 $p_1(x)>0, p_2(y)>0$ 分别为 $X$、$Y$ 的边缘分布密度.

**例 4** 设二维随机变量$(X,Y)$的联合分布密度为

$$p(x,y)=\begin{cases}1, & |y|\leqslant x, 0\leqslant x\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

求 $X$ 与 $Y$ 的条件分布密度 $p(x|y)$与 $p(y|x)$.

**解** 首先求出 $X$、$Y$ 的边缘分布

$$p_1(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y=\begin{cases}\int_{-x}^{x}\mathrm{d}y=2x, & 0<x<1,\\ 0, & \text{其他};\end{cases}$$

$$\begin{aligned}p_2(y)&=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x\\ &=\begin{cases}\int_{y}^{1}\mathrm{d}x=1-y, & 0\leqslant y<1,\\ \int_{-y}^{1}\mathrm{d}x=1+y, & -1<y<0,\\ 0, & \text{其他}\end{cases}\\ &=\begin{cases}1-|y|, & |y|<1,\\ 0, & \text{其他}.\end{cases}\end{aligned}$$

于是条件分布密度分别为

$$p(x|y)=\frac{p(x,y)}{p_2(y)}=\begin{cases}\dfrac{1}{1-|y|}, & |y|<x<1,\\ 0, & \text{其他};\end{cases}$$

$$p(y|x)=\frac{p(x,y)}{p_1(x)}=\begin{cases}\frac{1}{2x}, & |y|<x,\\ 0, & \text{其他}.\end{cases}$$

**定义 3.3**(二维随机变量联合分布函数)　设$(X,Y)$为二维随机变量,对于任意实数$x,y$,二元函数

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

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**.

随机变量$(X,Y)$的分布函数具有以下几个性质:

1) $F(x,y)$分别对$x$和$y$是非减的,即

当$x_2>x_1$时,有$F(x_2,y)\geqslant F(x_1,y)$;

当$y_2>y_1$时,有$F(x,y_2)\geqslant F(x,y_1)$.

2) $F(x,y)$分别对$x$和$y$是右连续的,即

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

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

4) 对于任意$a_1\leqslant b_1,a_2\leqslant b_2$,有

$$F(b_1,b_2)-F(a_1,b_2)-F(b_1,a_2)+F(a_1,a_2)\geqslant 0.$$

我们常利用性质2),3)来确定分布函数$F(x,y)$中的常数.

对于二维随机变量$(X,Y)$,我们有

1) 若$p(x,y)$在点$(x,y)$连续,则

$$\frac{\partial^2F(x,y)}{\partial x\partial y}=p(x,y).$$

2) 设$D$是$Oxy$平面上的一个区域,则点$(X,Y)$落在$D$内的概率为

$$P\{(X,Y)\in D\}=\iint_D p(x,y)\mathrm{d}\sigma.$$

**例 5**　设二维随机变量$(X,Y)$的联合分布函数为

$$F(x,y)=\begin{cases}C-3^{-x}-3^{-y}+3^{-x-y}, & x\geqslant 0, y\geqslant 0,\\ 0, & \text{其他}.\end{cases}$$

求(1) 常数 $C$;(2) 分布密度 $p(x,y)$.

**解** (1) 由性质 $F(+\infty,+\infty)=1$,得到 $C=1$.

(2) 由公式:$p(x,y)=\dfrac{\partial^2 F}{\partial x\partial y}$,有

$$\frac{\partial F}{\partial x}=3^{-x}\ln 3-3^{-x-y}\ln 3,$$

$$\frac{\partial^2 F}{\partial x\partial y}=\frac{\partial}{\partial y}(3^{-x}\ln 3-3^{-x-y}\ln 3)=3^{-x-y}(\ln 3)^2.$$

故

$$p(x,y)=\begin{cases}3^{-x-y}(\ln 3)^2, & x\geqslant 0, y\geqslant 0,\\ 0, & \text{其他}.\end{cases}$$

**定义 3.4(随机变量的独立性)** 设 $X,Y$ 是两个随机变量.如果对于任意的 $a<b, c<d$,事件$\{a<X<b\}$与$\{c<Y<d\}$相互独立,则称随机变量 $X$ 与 $Y$ 是**相互独立的**.

对于离散型随机变量,可以证明:当 $X,Y$ 的分布律分别为 $P\{X=x_i\}, i=1,2,\cdots; P\{Y=y_j\}, j=1,2,\cdots$时,则 $X$ 与 $Y$ 相互独立的充要条件是:对一切 $i,j$ 有

$$P\{X=x_i, Y=y_j\}=P\{X=x_i\}P\{Y=y_j\}.$$

**例 6** 若 $X,Y$ 的取值均为 1,2,3,4,并且事件$\{X=i,Y=j\}$ $(i,j=1,2,3,4)$的概率都相等.求$(X,Y)$的联合分布律、$X$ 和 $Y$ 的边缘分布,并讨论它们的独立性.

**解** 由 $p_{ij}$性质$\sum\limits_{i=1}^{4}\sum\limits_{j=1}^{4}p_{ij}=1$,按题意式中的 16 个 $p_{ij}$是相等的,因此,$(X,Y)$的联合分布为

$$p_{ij}=P\{X=i,Y=j\}=\frac{1}{16}\quad(i,j=1,2,3,4).$$

容易看出 $X$ 和 $Y$ 的边缘分布为

$$P\{X=i\}=\sum_{j=1}^{4}p_{ij}=\sum_{j=1}^{4}\frac{1}{16}=\frac{1}{4}\quad(i=1,2,3,4),$$

$$P\{Y=j\}=\frac{1}{4}\quad(j=1,2,3,4).$$

又由于对于一切 $i,j=1,2,3,4$ 有

$$P\{X=i,Y=j\}=\frac{1}{16}=P\{X=i\}\cdot P\{Y=j\},$$

可知 $X$ 和 $Y$ 是相互独立的.

如例 2 中的 $X$ 和 $Y$ 不独立,这是因为

$$P\{X=1\}P\{Y=-1\}=\frac{1}{18},\quad P\{X=1,Y=-1\}=\frac{1}{6}$$

的缘故.

对于二维连续型随机变量,可以证明:当 $X,Y$ 的分布密度分别是 $p_1(x),p_2(y)$时,则 $X$ 与 $Y$ 相互独立的**充要条件**是:二元函数

$$p_1(x)p_2(y)$$

为随机向量$(X,Y)$的联合分布密度 $p(x,y)$,即

$$p(x,y)=p_1(x)p_2(y).$$

这是概率统计中的一个重要结论. 在前面的讨论中,我们知道联合密度确定了边缘密度,而边缘密度一般来说是不能确定联合密度的. 然而这个结论告诉我们,当 $X,Y$ 独立时,两个边缘密度的乘积就是它们的联合密度,这就是说,只要 $X,Y$ 是独立的,那么边缘密度也能确定联合密度. 例如,设 $X\sim N(\mu_1,\sigma_1^2)$,$Y\sim N(\mu_2,\sigma_2^2)$,且 $X$ 与 $Y$ 相互独立,则$(X,Y)$的联合密度为

$$\begin{aligned}p(x,y)&=p_1(x)p_2(y)\\&=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\cdot\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}\\&=\frac{1}{2\pi\sigma_1\sigma_2}e^{-\frac{1}{2}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]}.\end{aligned}$$

我们把这一结果与前面的二元正态分布密度比较,不难发现:当 $X,Y$ 独立时,$X,Y$ 的联合密度 $p(x,y)=p(x)p(y)$恰好是二元正

态分布密度中 $\rho=0$ 时的特殊情况，由此我们给出了另一个重要结论：

若 $(X,Y)$ 服从二元正态分布，即 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则 $X$ 与 $Y$ 相互独立的充要条件是：$\rho=0$.

可见 $\rho$ 是二元正态分布密度函数中的一个重要参数.

如例 1 中的 $X$ 和 $Y$ 是相互独立的，这是因为当 $x\geqslant 0, y\geqslant 0$ 时，$p(x,y)=\mathrm{e}^{-(x+y)}=\mathrm{e}^{-x}\cdot\mathrm{e}^{-y}=p_1(x)\cdot p_2(y)$；而 $x<0$ 或 $y<0$ 时，$p(x,y)=0=p_1(x)\cdot p_2(y)$ 的缘故.

又如例 3 中的 $X$ 与 $Y$ 是不独立的，这是因为当 $(x,y)\in D$ 时，$p_1(x)p_2(x)=4x(1-y)\neq 2=p(x,y)$ 的缘故.

**定义 3.5**（独立同分布的随机变量） 称随机变量 $X_1,X_2,\cdots,X_n$ 是相互独立的，如果对于任意的 $a_i<b_i(i=1,2,\cdots,n)$，事件 $\{a_1<X_1<b_1\},\{a_2<X_2<b_2\},\cdots,\{a_n<X_n<b_n\}$ 相互独立. 此时，若所有的 $X_1,X_2,\cdots,X_n$ 都有共同的分布，则说 $X_1,X_2,\cdots,X_n$ 是**独立同分布的随机变量**.

（二）重要定理及公式

**定理 3.1（和的分布）** 设二维随机向量 $(X,Y)$ 的联合分布密度为 $p(x,y)$，随机变量 $Z=X+Y$，则 $Z$ 的分布密度为：

$$p_Z(z)=\int_{-\infty}^{+\infty}p(x,z-x)\mathrm{d}x.$$

特别，当 $X$ 和 $Y$ 相互独立时，有

$$p_Z(z)=\int_{-\infty}^{+\infty}p_X(x)p_Y(z-x)\mathrm{d}x.$$

利用上述公式，可以证明：若 $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，并且 $X$ 与 $Y$ 相互独立，则

$$X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2).$$

**例 7** 设两个独立的随机变量 $X$ 与 $Y$ 的概率分布为：

| $X$ | 1 | 3 |
|---|---|---|
| $p_i$ | 0.3 | 0.7 |

| $Y$ | 2 | 4 |
|---|---|---|
| $p_j$ | 0.6 | 0.4 |

求随机变量 $Z=X+Y$ 的分布.

**解** 因为 $X$ 与 $Y$ 相互独立,于是有

$$p_{ij}=p_i\cdot p_j.$$

$(X,Y)$的联合分布为

| X \ Y | 2 | 4 |
|---|---|---|
| 1 | 0.18 | 0.12 |
| 3 | 0.42 | 0.28 |

因此

| $p_{ij}$ | $(X,Y)$ | $Z=X+Y$ |
|---|---|---|
| 0.18 | (1,2) | 3 |
| 0.12 | (1,4) | 5 |
| 0.42 | (3,2) | 5 |
| 0.28 | (3,4) | 7 |

故 $Z=X+Y$ 的分布为

| $Z$ | 3 | 5 | 7 |
|---|---|---|---|
| $p_k$ | 0.18 | 0.54 | 0.28 |

**例 8** 设 $X$ 和 $Y$ 是两个相互独立的随机变量,且 $X\sim U(0,1)$,$Y\sim\Gamma(1,1)$,求 $Z=X+Y$ 的分布密度函数 $p_3(z)$.

**解** 由 $X\sim U(0,1)$,$Y\sim\Gamma(1,1)$,有

$$p_1(x)=\begin{cases}1, & 0\leqslant x\leqslant 1,\\ 0, & \text{其他},\end{cases}$$

$$p_2(y)=\begin{cases}\mathrm{e}^{-y}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

因为 $X$ 与 $Y$ 相互独立,所以$(X,Y)$的联合分布密度函数为

$$p(x,y)=p_1(x)p_2(y)$$
$$=\begin{cases}\mathrm{e}^{-y}, & 0\leqslant x\leqslant 1, y>0,\\ 0, & \text{其他}.\end{cases}$$

要使 $p(x,y)>0$，即 $p_1(x)>0, p_2(y)>0$，应满足 $0\leqslant x\leqslant 1$ 同时 $y>0$，考虑到 $z=x+y$，于是由

$$\begin{cases}0\leqslant x\leqslant 1,\\ y>0,\end{cases}\Rightarrow\begin{cases}0\leqslant x\leqslant 1,\\ z-x>0,\end{cases}\Rightarrow\begin{cases}0\leqslant x\leqslant 1,\\ x<z.\end{cases}\tag{3.1}$$

下面分三种情况讨论：

(1) 当 $z>1$ 时，(3.1)式合并为 $0\leqslant x\leqslant 1$，于是

$$p_3(z)=\int_{-\infty}^{+\infty}p_1(x)p_2(z-x)\mathrm{d}x=\int_0^1 1\cdot\mathrm{e}^{-(z-x)}\mathrm{d}x$$
$$=(\mathrm{e}-1)\mathrm{e}^{-z};$$

(2) 当 $0<z\leqslant 1$ 时，(3.1)式合并为 $0\leqslant x<z$，于是

$$p_3(z)=\int_{-\infty}^{+\infty}p_1(x)p_2(z-x)\mathrm{d}x=\int_0^z 1\cdot\mathrm{e}^{-(z-x)}\mathrm{d}x$$
$$=1-\mathrm{e}^{-z};$$

(3) 当 $z\leqslant 0$ 时，(3.1)式发生矛盾，因此，$p(x,y)=0$，于是

$$p_3(z)=\int_{-\infty}^{+\infty}0\mathrm{d}x=0.$$

故 $Z$ 的分布密度函数为

$$p_3(z)=\begin{cases}(\mathrm{e}-1)\mathrm{e}^{-z}, & z>1,\\ 1-\mathrm{e}^{-z}, & 0<z\leqslant 1,\\ 0, & z\leqslant 0.\end{cases}$$

**定理 3.2(正态分布)** 对于独立同 $N(\mu,\sigma^2)$ 分布的 $X_1,X_2,\cdots,X_n$，可以证明有下面三个重要结论：

1) 设 $S=\sum\limits_{i=1}^{n}X_i$，则 $S\sim N(n\mu,n\sigma^2)$；

2) 设 $\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$，则 $\overline{X}\sim N\left(\mu,\dfrac{\sigma^2}{n}\right)$；

3) 设 $U=\dfrac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}}$，则 $U\sim N(0,1)$.

**定理 3.3（$\chi^2$ 分布）** 设 $n$ 个随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，且服从标准正态分布，可以证明它们的平方和

$$\sum_{i=1}^{n}X_i^2 \xlongequal{\text{def}} \chi^2$$

的分布密度为

$$p_n(u)=\begin{cases}\dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\dfrac{n}{2}\right)}u^{\frac{n}{2}-1}\mathrm{e}^{-\frac{u}{2}}, & u\geqslant 0,\\ 0, & u<0.\end{cases}$$

这时我们称随机变量 $\chi^2=\sum\limits_{i=1}^{n}X_i^2$ 服从自由度为 $n$ 的 **$\chi^2$ 分布**，记为 $\chi^2\sim\chi^2(n)$，其中 $\Gamma\left(\dfrac{n}{2}\right)=\int_0^{+\infty}x^{\frac{n}{2}-1}\mathrm{e}^{-x}\mathrm{d}x$.

所谓自由度是指独立正态随机变量的个数，它是随机变量分布中的一个重要参数.

$\chi^2$ 分布密度函数 $p_n(u)$ 的图形与 $n$ 有关，对于不同的自由度 $n$，$p_n(u)$ 的图形各异. 当 $n$ 增大时，其图形逐渐接近于正态分布（见图 3.1）.

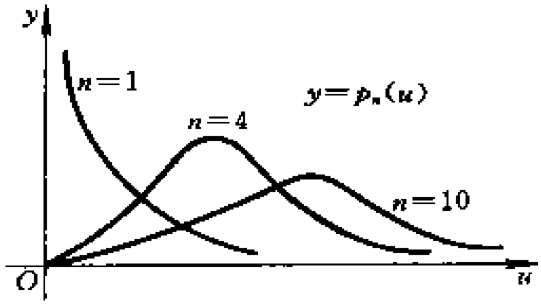



图 3.1

**定理 3.4（$t$-分布）** 设 $X,Y$ 是两个相互独立的随机变量，且

$$X\sim N(0,1),\quad Y\sim\chi^2(n),$$

可以证明函数

$$t=\frac{X}{\sqrt{Y/n}}$$

的概率密度为

$$p(t)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}\quad(-\infty<t<+\infty).$$

这时我们称随机变量 $t=\dfrac{X}{\sqrt{Y/n}}$ 服从自由度为 $n$ 的 **$t$ 分布**，记为 $t\sim t(n)$.

如图 3.2 所示，$t$ 分布的密度函数图形关于 $t=0$ 是对称的，其图形类似于标准正态分布的图形. $t$ 分布的图形与 $n$ 有关，对于很小的 $n$，$t$ 分布与标准正态分布相差很大；当 $n$ 增大时，$t$ 分布的图形渐渐接近于标准正态分布的图形；当 $n\geqslant 50$ 时，可以用标准正态分布代替 $t$ 分布.

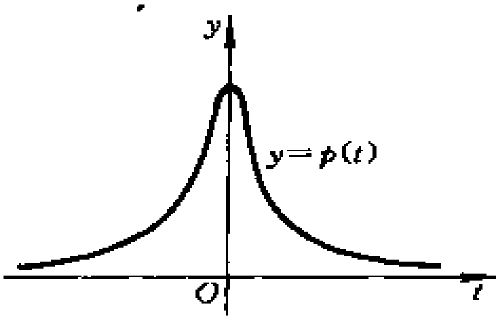



图 3.2

**定理 3.5（$f$-分布）** 设 $X\sim\chi^2(n_1)$，$Y\sim\chi^2(n_2)$，且 $X$ 与 $Y$ 独立，可以证明 $f=\dfrac{X/n_1}{Y/n_2}$ 的概率密度函数为

$$p(y)=\begin{cases}\dfrac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\left(\dfrac{n_1}{n_2}\right)^{\frac{n_1}{2}}y^{\frac{n_1}{2}-1}\left(1+\dfrac{n_1}{n_2}y\right)^{-\frac{n_1+n_2}{2}}, & y\geqslant 0,\\ 0, & y<0.\end{cases}$$

这时我们称随机变量 $f$ 服从第一个自由度为 $n_1$，第二个自由度为 $n_2$ 的 $f$ 分布，记为 $f\sim f(n_1,n_2)$. $p(y)$的图形与 $n_1,n_2$ 有关，如图 3.3 所示.

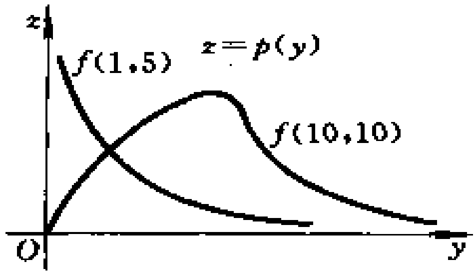



图 3.3

（三）重要方法

**方法 3.1**（判断随机变量的独立性） 设 $\xi=(X,Y)$的分布密度函数为 $p(x,y)$，则 $X$ 与 $Y$ 相互独立的充分必要条件是 $p(x,y)$ 可分离变量，即

$$p(x,y)=g(x)\cdot h(y).$$

**证** "⇒"必要性 若 $X$ 与 $Y$ 相互独立，记它们的分布分别为 $p_1(x),p_2(y)$，由独立性，可知

$$p(x,y)=p_1(x)\cdot p_2(y),$$

则取 $g(x)=p_1(x),h(y)=p_2(y)$即可.

"⇐"充分性 若 $p(x,y)$可分离变量，即

$$p(x,y)=g(x)\cdot h(y).$$

由于 $p(x,y)\geqslant 0$，可知 $g(x)$与 $h(y)$同号，不妨假设它们恒为正.

记 $\int_{-\infty}^{+\infty}g(x)\mathrm{d}x=a>0$，由联合分布密度性质：

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}\sigma=1$$

有 $\int_{-\infty}^{+\infty}h(y)\mathrm{d}y=\frac{1}{a}$. 令

$$p_1(x)=\frac{g(x)}{a},\quad p_2(y)=ah(y),$$

则 $p_1(x)\geqslant 0,p_2(y)\geqslant 0$，且

$$\int_{-\infty}^{+\infty} p_1(x)\mathrm{d}x = 1, \quad \int_{-\infty}^{+\infty} p_2(y)\mathrm{d}y = 1.$$

所以 $p_1(x),p_2(y)$分别为 $X,Y$ 的边缘分布密度,且

$$p(x,y) = p_1(x)p_2(y).$$

因此,$X$ 与 $Y$ 是相互独立的.

**例 9** 设二维随机向量 $\xi=(X,Y)$的分布密度函数为

$$p(x,y) = \begin{cases} Axy^2, & 0 \leqslant x \leqslant 2, 0 \leqslant y \leqslant 1, \\ 0, & \text{其他}. \end{cases}$$

(1) 确定常数 $A$;

(2) 求 $X,Y$ 的边缘分布密度函数 $p_1(x),p_2(y)$;

(3) 讨论 $X$ 与 $Y$ 的独立性.

**解** (1) 由$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)\mathrm{d}\sigma=1$,即

$$\int_0^2 \mathrm{d}x\int_0^1 Axy^2 = 1, \Rightarrow A = \frac{3}{2}.$$

(2) 由 $p_1(x)=\int_{-\infty}^{+\infty} p(x,y)\mathrm{d}y$,分情况讨论:

当 $x<0$ 或 $x>2$ 时, $p_1(x)=\int_{-\infty}^{+\infty} 0\mathrm{d}y=0$;

当 $0\leqslant x\leqslant 2$ 时, $p_1(x)=\int_{-\infty}^{+\infty} p(x,y)\mathrm{d}y=\int_0^1 \frac{3}{2}xy^2\mathrm{d}y=\frac{x}{2}$.

所以

$$p_1(x)=\begin{cases} \frac{x}{2}, & 0\leqslant x\leqslant 2, \\ 0, & \text{其他}. \end{cases}$$

同理,可求出

$$p_2(y) = \begin{cases} 3y^2, & 0 \leqslant y \leqslant 1, \\ 0, & \text{其他}. \end{cases}$$

(3) 由于 $p_1(x) \cdot p_2(y)=p(x,y)$,因此,$X$ 与 $Y$ 相互独立.

本题也可以使用其他方法讨论.

由于 $p(x,y)$可以拆成 $f_1(x) \cdot f_2(y)$,其中

$$f_1(x)=\begin{cases}ax, & 0\leqslant x\leqslant 2,\\0, & \text{其他},\end{cases}\qquad f_2(y)=\begin{cases}by^2, & 0\leqslant y\leqslant 1,\\0, & \text{其他}.\end{cases}$$

由$\int_{-\infty}^{+\infty}f_1(x)\mathrm{d}x=1$， 故 $a=\frac{1}{2}$. 同理 $b=3$. 因此 $A=\frac{3}{2}$.

这时 $p(x,y)=p_1(x)\cdot p_2(y)$，其中

$$p_1(x)=\begin{cases}\frac{x}{2}, & 0\leqslant x\leqslant 2,\\0, & \text{其他},\end{cases}\qquad p_2(y)=\begin{cases}3y^2, & 0\leqslant y\leqslant 1,\\0, & \text{其他}.\end{cases}$$

可见，$X$ 与 $Y$ 是相互独立的.

## 三、典型例题分析

### (一) 填空题

1. 设相互独立的两个随机变量 $X,Y$ 具有同一分布律，且 $X$ 的分布律为：

| $X$ | 0 | 1 |
|---|---|---|
| $P$ | $\frac{1}{2}$ | $\frac{1}{2}$ |

则随机变量 $Z=\max\{X,Y\}$ 的分布律为：______.

答案是：

| $Z$ | 0 | 1 |
|---|---|---|
| $P$ | $\frac{1}{4}$ | $\frac{3}{4}$ |

**分析** $P\{Z=0\}=P\{X=0,Y=0\}=P\{X=0\}P\{Y=0\}$

$$=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}.$$

$$P\{Z=1\}=1-P\{Z=0\}=\frac{3}{4}.$$

2. 设 $X$ 和 $Y$ 为两个随机变量，且

$$P\{X\geqslant 0,Y\geqslant 0\}=\frac{3}{7},\quad P\{X\geqslant 0\}=P\{Y\geqslant 0\}=\frac{4}{7},$$

则 $P\{\max(X,Y)\geqslant 0\}=$______.

答案是：$\frac{5}{7}$.

**分析** 记 $A=\{X\geqslant 0\}, B=\{Y\geqslant 0\}$，则

$$\{\max(X,Y)\geqslant 0\}=A\cup B,\quad \{X\geqslant 0,Y\geqslant 0\}=AB,$$

从而

$$\begin{aligned}P\{\max(X,Y)\geqslant 0\}&=P(A\cup B)=P(A)+P(B)-P(AB)\\&=P\{X\geqslant 0\}+P\{Y\geqslant 0\}-P\{X\geqslant 0,Y\geqslant 0\}\\&=\frac{4}{7}+\frac{4}{7}-\frac{3}{7}=\frac{5}{7}.\end{aligned}$$

3. 设 $\xi,\eta$ 是两个相互独立且均服从正态分布 $N\left(0,\left(\frac{1}{\sqrt{2}}\right)^2\right)$ 的随机变量，则随机变量 $|\xi-\eta|$ 的数学期望 $E(|\xi-\eta|)=$______.

答案是：$\sqrt{\frac{2}{\pi}}$.

**分析** 记 $Z=\xi-\eta$，$Z$ 为正态随机变量的线性组合，仍然服从正态分布，

$$E(Z)=E(\xi)-E(\eta)=0,$$

$$D(Z)=D(\xi)+D(\eta)=\frac{1}{2}+\frac{1}{2}=1,$$

于是 $Z\sim N(0,1)$. 从而

$$\begin{aligned}E(|\xi-\eta|)&=E(|Z|)=\int_{-\infty}^{+\infty}|z|\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz\\&=\sqrt{\frac{2}{\pi}}\int_0^{+\infty}z\cdot e^{\frac{z^2}{2}}dz=\sqrt{\frac{2}{\pi}}\left[-e^{-\frac{z^2}{2}}\right]_0^{+\infty}=\sqrt{\frac{2}{\pi}}.\end{aligned}$$

4. 已知随机变量 $X\sim N(-3,1)$，$Y\sim N(2,1)$，且 $X,Y$ 相互独立，设随机变量 $Z=X-2Y+7$，则 $Z\sim$______.

答案是：$N(0,5)$.

**分析** $Z$ 为正态随机变量的线性组合，仍然服从正态分布，且

$$E(Z)=E(X)-2E(Y)+7=-3-2\cdot 2+7=0,$$

$$D(Z)=D(X)+(-2)^2D(Y)=1+4\cdot 1=5,$$

故 $Z\sim N(0,5)$.

5. 设平面区域 $D$ 由曲线 $y=\frac{1}{x}$ 及直线 $y=0,x=1,x=e^2$ 所围成，二维随机变量 $(X,Y)$ 在区域 $D$ 上服从均匀分布，则 $(X,Y)$ 关于 $X$ 的边缘概率密度在 $x=2$ 处的值为______.

答案是：$\frac{1}{4}$.

**分析** 区域 $D$ 的面积为 $A=\int_1^{e^2}\frac{1}{x}dx=2$，由题设知 $(X,Y)$ 的概率密度为

$$p(x,y)=\begin{cases}\frac{1}{A}=\frac{1}{2}, & (x,y)\in D,\\ 0, & \text{其他}.\end{cases}$$

$(X,Y)$ 关于 $X$ 的边缘密度为 $p_X(x)=\int_{-\infty}^{+\infty}p(x,y)dy$，于是

$$p_X(2)=\int_{-\infty}^{+\infty}p(2,y)dy=\int_0^{\frac{1}{2}}\frac{1}{2}dy=\frac{1}{4}.$$

（二）选择题

1. 设两个随机变量 $X$ 与 $Y$ 相互独立且同分布：

$$P\{X=-1\}=P\{Y=-1\}=\frac{1}{2},$$

$$P\{X=1\}=P\{Y=1\}=\frac{1}{2},$$

则下列各式中成立的是（　　）.

(A) $P\{X=Y\}=\frac{1}{2}$；　　(B) $P\{X=Y\}=1$；

(C) $P\{X+Y=0\}=\frac{1}{4}$；　　(D) $P\{XY=1\}=\frac{1}{4}$.

答案是：A.

**分析** 由 $X$ 与 $Y$ 相互独立，并且都具有下面的分布：

| $x_i$ | $-1$ | $1$ |
|---|---|---|
| $p_i$ | $\frac{1}{2}$ | $\frac{1}{2}$ |

而

$$\begin{aligned}P\{X=Y\}&=P\{(X=-1,Y=-1)\cup(X=1,Y=1)\}\\&=P\{X=-1\}P\{Y=-1\}+P\{X=1\}P\{Y=1\}\\&=\frac{1}{2}\times\frac{1}{2}+\frac{1}{2}\times\frac{1}{2}=\frac{1}{4}.\end{aligned}$$

因此，选择 A.

2. 设 $F_1(x)$ 与 $F_2(x)$ 分别为随机变量 $X_1$ 与 $X_2$ 的分布函数. 为使 $F(x)=aF_1(x)-bF_2(x)$ 是某一随机变量的分布函数，在下列给定的各组数值中应取（　　）.

(A) $a=\frac{3}{5},b=-\frac{2}{5}$；　　(B) $a=\frac{2}{3},b=\frac{2}{3}$；

(C) $a=-\frac{1}{2},b=\frac{3}{2}$；　　(D) $a=\frac{1}{2},b=-\frac{3}{2}$.

答案是：A.

**分析**　因为分布函数 $F_1(x)$，$F_2(x)$ 为非减，为使 $F(x)=aF_1(x)-bF_2(x)$ 也为非减，应排除选项(B)和(C). 又因 $F_1(+\infty)=F_2(+\infty)=1$，为使 $F(+\infty)=1$. 因此，选择 A.

3. 设随机变量 $X$ 和 $Y$ 相互独立，其概率分布为

| $m$ | $-1$ | $1$ |
|---|---|---|
| $P\{X=m\}$ | $\frac{1}{2}$ | $\frac{1}{2}$ |

| $m$ | $-1$ | $1$ |
|---|---|---|
| $P\{Y=m\}$ | $\frac{1}{2}$ | $\frac{1}{2}$ |

则下列式子正确的是（　　）.

(A) $X=Y$；　　(B) $P\{X=Y\}=0$；

(C) $P\{X=Y\}=\frac{1}{2}$；　　　　　(D) $P\{X=Y\}=1$.

答案是：C.

**分析** $P\{X=Y\}=P\{X=-1,Y=-1\}+P\{X=1,Y=1\}$

$$=P\{X=-1\}P\{Y=-1\}+P\{X=1\}P\{Y=1\}=\frac{1}{2}.$$

因此，选择 C.

4. 设随机变量

$$X_i\sim\begin{pmatrix}-1 & 0 & 1\\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4}\end{pmatrix}\quad(i=1,2),$$

且满足 $P\{X_1X_2=0\}=1$，则 $P\{X_1=X_2\}$等于(　　).

(A) 0；　　(B) $\frac{1}{4}$；　　(C) $\frac{1}{2}$；　　(D) 1.

答案是：A.

**分析**　下面给出$(X_1,X_2)$的联合分布：

| $X_1$ \ $X_2$ | $-1$ | 0 | 1 | $p_{i\cdot}$ |
|---|---|---|---|---|
| $-1$ | 0 | $\frac{1}{4}$ | 0 | $\frac{1}{4}$ |
| 0 | $\frac{1}{4}$ | 0 | $\frac{1}{4}$ | $\frac{1}{2}$ |
| 1 | 0 | $\frac{1}{4}$ | 0 | $\frac{1}{4}$ |
| $p_{\cdot j}$ | $\frac{1}{4}$ | $\frac{1}{2}$ | $\frac{1}{4}$ | 1 |

可见，$P\{X_1=X_2\}=0$，因此，选择 A.

（三）解答题

1. 把一枚均匀硬币连掷 3 次，以 $X$ 表示三次中正面向上的次数，$Y$ 表示三次中正面向上的次数与反面向上的次数差的绝对值，试求$(X,Y)$的联合分布律和边缘分布律.

**解**　一均匀硬币连掷三次，将等可能地出现下面 8 种情形：

出现三次正面：正正正；

出现二次正面：正正反，正反正，反正正；

出现一次正面：正反反，反正反，反反正；

不出现正面：反反反.

则$(X,Y)$的可能取值分别为$(3,3)$，$(2,1)$，$(1,1)$和$(0,3)$，且

$$P\{X=3,Y=3\}=\frac{1}{8},\quad P\{X=2,Y=1\}=\frac{3}{8},$$

$$P\{X=1,Y=1\}=\frac{3}{8},\quad P\{X=0,Y=3\}=\frac{1}{8}.$$

故$(X,Y)$的联合分布律与边缘分布律为：

| $X$ \ $Y$ | 1 | 3 | $p_{i\cdot}$ |
|---|---|---|---|
| 0 | 0 | $\frac{1}{8}$ | $\frac{1}{8}$ |
| 1 | $\frac{3}{8}$ | 0 | $\frac{3}{8}$ |
| 2 | $\frac{3}{8}$ | 0 | $\frac{3}{8}$ |
| 3 | 0 | $\frac{1}{8}$ | $\frac{1}{8}$ |
| $p_{\cdot j}$ | $\frac{3}{4}$ | $\frac{1}{4}$ | 1 |

2. 10 件产品中有 2 件一级品，7 件二级品，1 件次品. 从中任取 3 件，用 $X$ 表示其中的一级品数，用 $Y$ 表示其中的二级品数，求$(X,Y)$的联合分布律与边缘分布律.

**解** $X$ 的可能取值为 0，1，2；$Y$ 的可能取值为 0，1，2，3. 注意到 $2\leqslant X+Y\leqslant 3$，故有

$$\begin{aligned}P\{X=0,Y=0\}&=P\{X=0,Y=1\}=P\{X=1,Y=0\}\\&=P\{X=1,Y=3\}=P\{X=2,Y=2\}\\&=P\{X=2,Y=3\}=0,\end{aligned}$$

$$P\{X=0,Y=2\}=C_2^0C_7^2C_1^1/C_{10}^3=\frac{21}{120},$$

$$P\{X=0,Y=3\}=C_2^0C_7^3/C_{10}^3=\frac{35}{120},$$

$$P\{X=1,Y=1\}=C_2^1C_7^1C_1^1/C_{10}^3=\frac{14}{120},$$

$$P\{X=1,Y=2\}=C_2^1C_7^2/C_{10}^3=\frac{42}{120},$$

$$P\{X=2,Y=0\}=C_2^2C_7^0C_1^1/C_{10}^3=\frac{1}{120},$$

$$P\{X=2,Y=1\}=C_2^2C_7^1/C_{10}^3=\frac{7}{120},$$

故$(X,Y)$的联合分布律和边缘分布律为：

| $X \backslash Y$ | 0 | 1 | 2 | 3 | $p_{i\cdot}$ |
|---|---|---|---|---|---|
| 0 | 0 | 0 | $\frac{21}{120}$ | $\frac{35}{120}$ | $\frac{7}{15}$ |
| 1 | 0 | $\frac{14}{120}$ | $\frac{42}{120}$ | 0 | $\frac{7}{15}$ |
| 2 | $\frac{1}{120}$ | $\frac{7}{120}$ | 0 | 0 | $\frac{1}{15}$ |
| $p_{\cdot j}$ | $\frac{1}{120}$ | $\frac{21}{120}$ | $\frac{63}{120}$ | $\frac{35}{120}$ | 1 |

3. 设随机变量$X,Y$相互独立，其概率密度函数分别为

$$p_X(x)=\begin{cases}1, & 0\leqslant x\leqslant 1,\\ 0, & \text{其他},\end{cases}\qquad p_Y(y)=\begin{cases}e^{-y}, & y>0,\\ 0, & y\leqslant 0,\end{cases}$$

求随机变量$Z=2X+Y$的概率密度函数.

**解** **方法1** 由于$X$与$Y$相互独立，所以$(X,Y)$的概率密度函数为

$$p(x,y)=p_X(x)\cdot p_Y(y)=\begin{cases}e^{-y}, & 0\leqslant x\leqslant 1, y>0,\\ 0, & \text{其他}.\end{cases}$$

因此，$Z$的分布函数为

$$F_Z(z)=P\{2X+Y<z\}=\iint\limits_{2x+y<z} p_X(x)\cdot p_Y(y)\mathrm{d}x\mathrm{d}y$$

$$
=\begin{cases}0, & z\leqslant 0,\\ \int_0^{\frac{z}{2}}(1-e^{2x-z})dx, & 0<z\leqslant 2,\\ \int_0^1(1-e^{2x-z})dx, & z>2,\end{cases}
$$

所以，$Z$ 的概率密度函数为

$$
p_Z(z)=F_Z'(z)=\begin{cases}0, & z\leqslant 0,\\ \frac{1}{2}(1-e^{-z}), & 0<z\leqslant 2,\\ \frac{1}{2}(e^2-1)e^{-z}, & z>2.\end{cases}
$$

**方法 2** 由于 $X$ 与 $Y$ 相互独立，所以 $Z$ 的概率密度函数为

$$
p_Z(z)=\int_{-\infty}^{+\infty}p_X(x)p_Y(z-2x)dx=\int_0^1 p_Y(z-2x)dx
$$

$$
=\begin{cases}0, & z\leqslant 0,\\ \int_0^{\frac{z}{2}}e^{-(z-2x)}dx, & 0<z\leqslant 2,\\ \int_0^1 e^{-(z-2x)}dx, & z>2\end{cases}
$$

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

4. 设二维随机变量$(X,Y)$的概率密度为

$$
p(x,y)=\begin{cases}2e^{-(x+2y)}, & x>0,y>0,\\ 0, & \text{其他},\end{cases}
$$

求随机变量 $Z=X+2Y$ 的分布函数.

**解** $F_Z(z)=P\{Z\leqslant z\}=P\{X+2Y\leqslant z\}=\iint\limits_{x+2y\leqslant z}p(x,y)dxdy.$

当 $z\leqslant 0$ 时，$P\{Z\leqslant z\}=0$；当 $z>0$ 时，

$$P\{Z\leqslant z\}=\int_0^z \mathrm{d}x\int_0^{\frac{z-x}{2}} 2\mathrm{e}^{-(x+2y)}\mathrm{d}y=\int_0^z \mathrm{e}^{-x}\mathrm{d}x\int_0^{\frac{z-x}{2}} 2\mathrm{e}^{-2y}\mathrm{d}y$$

$$=\int_0^z(\mathrm{e}^{-x}-\mathrm{e}^{-z})\mathrm{d}x=1-\mathrm{e}^{-z}-z\mathrm{e}^{-z}.$$

所以 $Z=X+2Y$ 的分布函数为

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

5. 设随机变量 $X$ 以概率 1 取值 0，而 $Y$ 是任意的随机变量，证明 $X$ 与 $Y$ 相互独立.

**证** $X$ 的分布函数为

$$F_1(x)=\begin{cases}0, & \text{当 } x<0 \text{ 时},\\ 1, & \text{当 } x\geqslant 0 \text{ 时}.\end{cases}$$

设 $Y$ 的分布函数为 $F_2(y)$，$(X,Y)$的分布函数为 $F(x,y)$，则当 $x<0$ 时，对任意的 $y$ 有

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

$$=P(\varnothing\cap\{Y\leqslant y\})=P(\varnothing)=0=F_1(x)F_2(y).$$

当 $x\geqslant 0$ 时，对任意的 $y$ 有

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

$$=F_2(y)=F_1(x)F_2(y).$$

因此对任意的 $x,y$ 均有

$$F(x,y)=F_1(x)F_2(y),$$

即 $X$ 与 $Y$ 相互独立.

6. 已知随机变量 $X$ 和 $Y$ 的联合概率密度为

$$p(x,y)=\begin{cases}4xy, & \text{若 } 0\leqslant x\leqslant 1,0\leqslant y\leqslant 1,\\ 0, & \text{其他},\end{cases}$$

求 $X$ 和 $Y$ 的联合分布函数 $F(x,y)$.

**解** 1) 当 $x<0$ 或 $y<0$ 时，有

$$F(x,y)=P\{X\leqslant x,Y\leqslant y\}=0.$$

2）当 $x>1$ 且 $y>1$ 时，有 $F(x,y)=1$.

3）当 $0\leqslant x\leqslant 1$ 且 $0\leqslant y\leqslant 1$ 时，有

$$F(x,y)=\int_0^x\int_0^y 4uv\mathrm{d}u\mathrm{d}v=x^2y^2.$$

4）当 $0\leqslant x\leqslant 1$，且 $y>1$ 时，有

$$F(x,y)=\int_0^x\int_0^1 4uv\mathrm{d}u\mathrm{d}v=x^2.$$

5）当 $x>1$ 且 $0\leqslant y\leqslant 1$ 时，有

$$F(x,y)=\int_0^1\int_0^y 4uv\mathrm{d}u\mathrm{d}v=y^2.$$

故 $X$ 和 $Y$ 的联合分布函数为

$$F(x,y)=\begin{cases}0, & x<0\text{ 或 }y<0,\\ x^2y^2, & 0\leqslant x\leqslant 1,0\leqslant y\leqslant 1,\\ x^2, & 0\leqslant x\leqslant 1,y>1,\\ y^2, & x>1,0\leqslant y\leqslant 1,\\ 1, & x>1,y>1.\end{cases}$$

7. 设二维随机变量$(X,Y)$的概率密度为

$$p(x,y)=\begin{cases}\mathrm{e}^{-y}, & 0<x<y,\\ 0, & \text{其他}.\end{cases}$$

(1) 求 $X$ 的密度 $p_X(x)$；

(2) 求概率 $P\{X+Y\leqslant 1\}$.

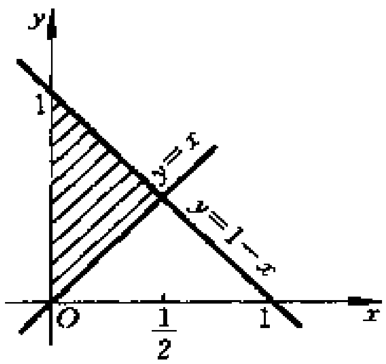


图 3.4

**解**

(1) $$p_X(x)=\begin{cases}\int_x^{+\infty}\mathrm{e}^{-y}\mathrm{d}y, & x>0,\\ 0, & x\leqslant 0\end{cases}=\begin{cases}\mathrm{e}^{-x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

(2) $$P\{X+Y\leqslant 1\}=\iint\limits_{x+y\leqslant 1}p(x,y)\mathrm{d}x\mathrm{d}y=\int_0^{\frac{1}{2}}\mathrm{d}x\int_x^{1-x}\mathrm{e}^{-y}\mathrm{d}y$$

$$=-\int_0^{\frac{1}{2}}[\mathrm{e}^{-(1-x)}-\mathrm{e}^{-x}]\mathrm{d}x=1+\mathrm{e}^{-1}-2\mathrm{e}^{-\frac{1}{2}}.$$

这里二重积分中用到的积分区域如图 3.4 所示.

8. 设 $X$ 与 $Y$ 的密度函数为

$$p(x,y)=\begin{cases}1, & |y|<x,0<x<1,\\ 0, & 其他.\end{cases}$$

求

(1) 边缘密度 $p_X(x),p_Y(y)$;

(2) 条件密度 $p_{X|y}(x),p_{Y|x}(y)$;

(3) $P\left\{X>\frac{1}{2}\middle|Y>0\right\}$.

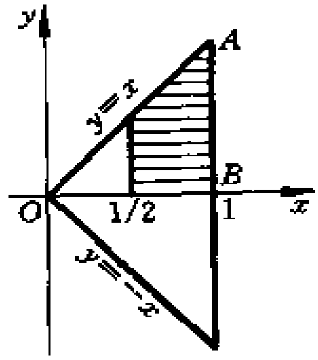



图 3.5

**解** (1) 如图 3.5 有:

$$p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x$$

$$=\begin{cases}\int_y^1\mathrm{d}x=1-y, & 0\leqslant y<1,\\ \int_{-y}^1\mathrm{d}x=1+y, & -1<y<0,\\ 0, & 其他\end{cases}$$

$$=\begin{cases}1-|y|, & |y|<1,\\ 0, & 其他.\end{cases}$$

$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y=\begin{cases}\int_{-x}^{x}\mathrm{d}y=2x, & 0<x<1,\\ 0, & 其他.\end{cases}$$

(2) 当 $|y|<1$ 时,

$$p_{X|y}(x)=\frac{p(x,y)}{p_Y(y)}=\begin{cases}\dfrac{1}{1-|y|}, & |y|<x<1,\\ 0, & 其他.\end{cases}$$

当 $0<x<1$ 时,

$$p_{Y|x}(y)=\frac{p(x,y)}{p_X(x)}=\begin{cases}\dfrac{1}{2x}, & |y|<x,\\ 0, & 其他.\end{cases}$$

(3) $P\left\{X>\frac{1}{2}\middle|Y>0\right\}=\dfrac{P\left\{X>\frac{1}{2},Y>0\right\}}{P\{Y>0\}}$

$$=\frac{\text{图中阴影面积}}{\triangle AOB\text{ 的面积}}=\frac{3}{4}.$$

9. 设$(X,Y)$的联合密度函数

$$p(x,y)=\begin{cases}Ae^{-(2x+3y)}, & x>0,y>0,\\ 0, & \text{其他}.\end{cases}$$

试求(1) 常数 $A$;(2) $X$ 的边缘密度;(3) $P\{X+Y<2\}$;(4) 条件密度 $p_{X|y}(x)$;(5) $P\{X<2|Y<1\}$.

**解** (1) 因为

$$1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x\mathrm{d}y=A\int_0^{+\infty}e^{-2x}\mathrm{d}x\int_0^{+\infty}e^{-3y}\mathrm{d}y=\frac{A}{6},$$

所以 $A=6$.

(2) $$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y=\int_0^{+\infty}6e^{-(2x+3y)}\mathrm{d}y$$

$$=6e^{-2x}\int_0^{+\infty}e^{-3y}\mathrm{d}y=2e^{-2x}\quad(x>0).$$

(3) $$P\{X+Y<2\}=6\int_0^2\int_0^{2-x}e^{-(2x+3y)}\mathrm{d}y\mathrm{d}x$$

$$=6\int_0^2e^{-2x}\left[\int_0^{2-x}e^{-3y}\mathrm{d}y\right]\mathrm{d}x$$

$$=6\int_0^2e^{-2x}\cdot\frac{1}{3}(1-e^{3x-6})\mathrm{d}x$$

$$=1-3e^{-4}+2e^{-6}.$$

(4) 因为

$$p_Y(y)=\int_0^{+\infty}6e^{-(2x+3y)}\mathrm{d}x$$

$$=6e^{-3y}\int_0^{+\infty}e^{-2x}\mathrm{d}x=3e^{-3y}\quad(y>0),$$

所以当 $y>0$ 时,

$$p_{X|y}(x)=\frac{p(x,y)}{p_Y(y)}=\begin{cases}\dfrac{6e^{-(2x+3y)}}{3e^{-3y}}, & x>0,\\ 0, & x\leqslant 0\end{cases}$$

$$=\begin{cases}2e^{-2x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

(5) 因为 $P\{X<2|Y<1\}=\dfrac{P\{X<2,Y<1\}}{P\{Y<1\}}$，而

$$P\{X<2,Y<1\}=\int_0^2\int_0^1 6e^{-(2x+3y)}dydx$$

$$=6\int_0^2 e^{-2x}dx\int_0^1 e^{-3y}dy=(1-e^{-4})(1-e^{-3}),$$

$$P\{Y<1\}=\int_0^1 3e^{-3y}dy=1-e^{-3},$$

所以

$$P\{X<2/Y<1\}=\frac{(1-e^{-4})(1-e^{-3})}{1-e^{-3}}=1-e^{-4}.$$

10. 假设随机变量 $X_1,X_2,X_3,X_4$ 相互独立，且同分布 $P\{X_i=0\}=0.6,P\{X_i=1\}=0.4$ $(i=1,2,3,4)$，求行列式

$$X=\begin{vmatrix} X_1 & X_2 \\ X_3 & X_4 \end{vmatrix}$$

的概率分布.

**解** 记 $Y_1=X_1X_4,Y_2=X_2X_3$，则 $X=Y_1-Y_2$，且 $Y_1$ 和 $Y_2$ 独立同分布：

$$P\{Y_1=1\}=P\{Y_2=1\}=P\{X_2=1,X_3=1\}=0.16,$$

$$P\{Y_1=0\}=P\{Y_2=0\}=1-0.16=0.84.$$

随机变量 $X=Y_1-Y_2$ 有三个可能值：$-1,0,1$.

$$P\{X=-1\}=P\{Y_1=0,Y_2=1\}$$

$$=0.84\times0.16=0.1344,$$

$$P\{X=1\}=P\{Y_1=1,Y_2=0\}$$

$$=0.16\times0.84=0.1344,$$

$$P\{X=0\}=1-2\times0.1344=0.7312.$$

于是，行列式 $X$ 的概率分布为

$$X\sim\begin{bmatrix} -1 & 0 & 1 \\ 0.1344 & 0.7312 & 0.1344 \end{bmatrix}.$$

11. 一电子仪器由两个部件构成，以 $X$ 和 $Y$ 分别表示两个部

件的寿命(单位：千小时)．已知 $X$ 和 $Y$ 的联合分布函数为

$$F(x,y)=\begin{cases}1-e^{-0.5x}-e^{-0.5y}+e^{-0.5(x+y)}, & \text{若 } x\geqslant 0, y\geqslant 0,\\ 0, & \text{其他}.\end{cases}$$

(1) 问 $X$ 和 $Y$ 是否独立？

(2) 求两个部件的寿命都超过 100 小时的概率 $\alpha$.

**解　方法 1**　(1) $X$ 和 $Y$ 的分布函数分别为：

$$F_X(x)=F(x,+\infty)=\begin{cases}1-e^{-0.5x}, & \text{若 } x\geqslant 0,\\ 0, & \text{若 } x<0;\end{cases}$$

$$F_Y(y)=F(+\infty,y)=\begin{cases}1-e^{-0.5y}, & \text{若 } y\geqslant 0,\\ 0, & \text{若 } y<0.\end{cases}$$

由于 $F(x,y)=F_X(x)\cdot F_Y(y)$，可知 $X$ 和 $Y$ 独立.

(2) $\alpha=P\{X>0.1,Y>0.1\}=P\{X>0.1\}\cdot P\{Y>0.1\}$

$$=[1-F_X(0.1)]\cdot[1-F_Y(0.1)]=e^{-0.05}\cdot e^{-0.05}$$

$$=e^{-0.1}.$$

**方法 2**　(1) 以 $f(x,y)$，$f_1(x)$和 $f_2(y)$分别代表$(X,Y)$，$X$ 和 $Y$ 的概率密度，有

$$f(x,y)=\frac{\partial F^2(x,y)}{\partial x\partial y}=\begin{cases}0.25e^{-0.5(x+y)}, & \text{若 } x\geqslant 0, y\geqslant 0,\\ 0, & \text{其他};\end{cases}$$

$$f_1(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\begin{cases}0.5e^{-0.5x}, & \text{若 } x\geqslant 0,\\ 0, & \text{若 } x<0;\end{cases}$$

$$f_2(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\begin{cases}0.5e^{-0.5y}, & \text{若 } y\geqslant 0,\\ 0, & \text{若 } y<0.\end{cases}$$

由于 $f(x,y)=f_1(x)\cdot f_2(y)$，所以 $X$ 和 $Y$ 独立.

(2) $\alpha=P\{X>0.1,Y>0.1\}=\int_{0.1}^{+\infty}\int_{0.1}^{+\infty}0.25e^{-0.25(x+y)}\mathrm{d}x\mathrm{d}y$

$$=\int_{0.1}^{+\infty}0.5e^{-0.5x}\mathrm{d}x\cdot\int_{0.1}^{+\infty}0.5e^{-0.5y}\mathrm{d}y=e^{-0.1}.$$

12. 设随机变量$(X,Y)$的密度为

$$p(x,y)=\begin{cases}x^2+\frac{1}{3}xy, & 0\leqslant x\leqslant 1,\quad 0\leqslant y\leqslant 2,\\ 0, & \text{其他}.\end{cases}$$

试求：(1) $(X,Y)$的分布函数；

(2) $(X,Y)$的边缘分布密度；

(3) $(X,Y)$的条件密度；

(4) 概率 $P\{X+Y>1\}$，$P\{Y>X\}$及 $P\left\{Y<\frac{1}{2}\middle|X<\frac{1}{2}\right\}$.

**解** (1) 当 $x<0$ 或 $y<0$ 时，因为 $p(x,y)=0$，所以 $F(x,y)=0$；

当 $0\leqslant x\leqslant 1,0\leqslant y\leqslant 2$ 时，

$$F(x,y)=\int_0^x\int_0^y\left(u^2+\frac{1}{3}uv\right)\mathrm{d}u\mathrm{d}v=\frac{1}{3}x^3y+\frac{1}{12}x^2y^2;$$

当 $0\leqslant x\leqslant 1,y>2$ 时，

$$\begin{aligned}F(x,y)&=\int_{-\infty}^x\int_{-\infty}^y p(u,v)\mathrm{d}u\mathrm{d}v=\int_0^x\int_0^2\left(u^2+\frac{1}{3}uv\right)\mathrm{d}v\mathrm{d}u\\&=\frac{1}{3}(2x+1)x^2;\end{aligned}$$

当 $x>1,0\leqslant y\leqslant 2$ 时，

$$\begin{aligned}F(x,y)&=\int_{-\infty}^x\int_{-\infty}^y p(u,v)\mathrm{d}u\mathrm{d}v=\int_0^1\int_0^y\left(x^2+\frac{1}{3}xv\right)\mathrm{d}v\mathrm{d}x\\&=\frac{1}{12}(4+y)y;\end{aligned}$$

当 $x>1,y>2$ 时，

$$F(x,y)=\int_0^1\int_0^2\left(x^2+\frac{1}{3}xy\right)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^2\left(x^2+\frac{1}{3}xy\right)\mathrm{d}y=1.$$

综上所述，分布函数为

$$F(x,y)=\begin{cases}0, & x<0\text{ 或 }y<0,\\ \frac{1}{3}x^2y\left(x+\frac{y}{4}\right), & 0\leqslant x\leqslant 1,0\leqslant y\leqslant 2,\\ \frac{1}{3}x^2(2x+1), & 0\leqslant x\leqslant 1,y>2,\\ \frac{1}{12}y(4+y), & 0>1,0\leqslant y\leqslant 2,\\ 1, & x>1,y>2.\end{cases}$$

(2) 当 $0\leqslant x\leqslant 1$ 时，

$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y=\int_0^2\left(x^2+\frac{1}{3}xy\right)\mathrm{d}y=2x^2+\frac{2}{3}x;$$

当 $0\leqslant y\leqslant 2$ 时，

$$p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x=\int_0^1\left(x^2+\frac{x}{3}y\right)\mathrm{d}x=\frac{1}{3}+\frac{1}{6}y.$$

(3) 当 $0\leqslant x\leqslant 1,0\leqslant y\leqslant 2$ 时，

$$p_{X|y}(x)=\frac{p(x,y)}{p_Y(y)}=\frac{6x^2+2xy}{2+y},$$

$$p_{Y|x}(y)=\frac{p(x,y)}{p_X(x)}=\frac{3x+y}{6x+2}.$$

(4) $$P\{X+Y>1\}=\iint\limits_{x+y>1}p(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\int_0^1\mathrm{d}x\int_{1-x}^2\left(x^2+\frac{xy}{3}\right)\mathrm{d}y=\frac{65}{72},$$

$$P\{Y>X\}=\iint\limits_{y>x}\left(x^2+\frac{1}{3}xy\right)\mathrm{d}x\mathrm{d}y$$

$$=\int_0^1\mathrm{d}x\int_x^2\left(x^2+\frac{1}{3}xy\right)\mathrm{d}y=\frac{17}{24},$$

$$P\left\{Y<\frac{1}{2}\middle|X<\frac{1}{2}\right\}=\frac{P\left\{X<\frac{1}{2},Y<\frac{1}{2}\right\}}{P\left\{X<\frac{1}{2}\right\}}$$

$$=\frac{F\left(\frac{1}{2},\frac{1}{2}\right)}{F_X\left(\frac{1}{2}\right)}=\frac{\frac{1}{3}x^2y\left(x+\frac{y}{4}\right)\Big|_{\left(\frac{1}{2},\frac{1}{2}\right)}}{\int_0^{\frac{1}{2}}p_X(x)\mathrm{d}x}=\frac{5}{32}.$$

13. 甲、乙两人独立地各进行两次射击，假设甲的命中率为 0.2，乙的命中率为 0.5，以 $X$ 和 $Y$ 分别表示甲和乙的命中次数. 试求 $X$ 和 $Y$ 的联合概率分布.

**解** $X$ 服从参数为 $n=2,p=0.2$ 的二项分布，$Y$ 服从参数为 $n=2,p=0.5$ 的二项分布，它们的概率分布分别为：

| $x$ | 0 | 1 | 2 |
|---|---|---|---|
| $P\{X=x\}$ | 0.64 | 0.32 | 0.04 |

| $y$ | 0 | 1 | 2 |
|---|---|---|---|
| $P\{Y=y\}$ | 0.25 | 0.5 | 0.25 |

由 $X$ 和 $Y$ 的独立性知 $X$ 和 $Y$ 的联合概率分布为：

| $Y \backslash X$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0.16 | 0.08 | 0.01 |
| 1 | 0.32 | 0.16 | 0.02 |
| 2 | 0.16 | 0.08 | 0.01 |

14. 设随机变量 $X$ 与 $Y$ 相互独立，下表列出了二维随机变量 $(X,Y)$ 联合分布律及关于 $X$ 和关于 $Y$ 的边缘分布律中的部分数值，试将其余数值填入表中的空白处.

| $X \backslash Y$ | $y_1$ | $y_2$ | $y_3$ | $P\{X=x_i\}=p_{i\cdot}$ |
|---|---|---|---|---|
| $x_1$ | | $\frac{1}{8}$ | | |
| $x_2$ | $\frac{1}{8}$ | | | |
| $P\{Y=y_j\}=p_{\cdot j}$ | $\frac{1}{6}$ | | | 1 |

**解** 由于

$$P\{X=x_1,Y=y_1\}=P\{Y=y_1\}-P\{X=x_2,Y=y_1\}$$

$$=\frac{1}{6}-\frac{1}{8}=\frac{1}{24},$$

考虑到 $X$ 与 $Y$ 相互独立，有

$$P\{X=x_1\}P\{Y=y_1\}=P\{X=x_1,Y=y_1\},$$

所以 $$P\{X=x_1\}=\frac{\frac{1}{24}}{\frac{1}{6}}=\frac{1}{4}.$$

同理，可以导出其他数值. 故 $XY$ 的联合分布律为：

| $X \backslash Y$ | $y_1$ | $y_2$ | $y_3$ | $P\{X=x_i\}=p_{i\cdot}$ |
|---|---|---|---|---|
| $x_1$ | $\frac{1}{24}$ | $\frac{1}{8}$ | $\frac{1}{12}$ | $\frac{1}{4}$ |
| $x_2$ | $\frac{1}{8}$ | $\frac{3}{8}$ | $\frac{1}{4}$ | $\frac{3}{4}$ |
| $P\{Y=y_j\}=p_{\cdot j}$ | $\frac{1}{6}$ | $\frac{1}{2}$ | $\frac{1}{3}$ | 1 |

15. 假设二维随机变量$(X,Y)$在矩形 $G=\{(x,y)\mid 0\leqslant x\leqslant 2,0\leqslant y\leqslant 1\}$（见图 3.6）上服从均匀分布. 记

$$U=\begin{cases}0, & 若\ X\leqslant Y,\\ 1, & 若\ X>Y;\end{cases}\quad V=\begin{cases}0, & 若\ X\leqslant 2Y,\\ 1, & 若\ X>2Y.\end{cases}$$

(1) 求 $U$ 和 $V$ 的联合分布；

(2) 求 $U$ 和 $V$ 的相关系数 $r$.

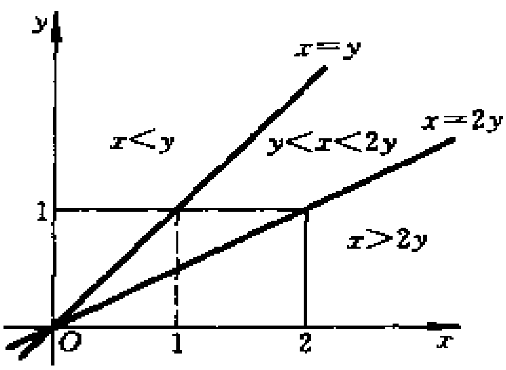



图 3.6

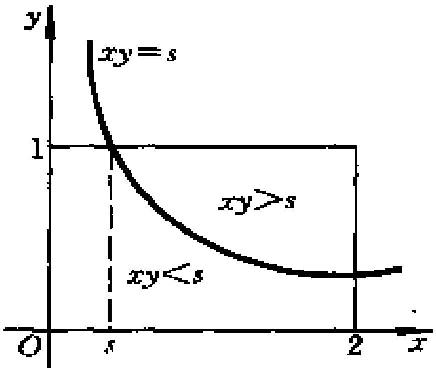



图 3.7

**解** 由题设可知

$$P\{X\leqslant Y\}=\frac{1}{4},\quad P\{X>2Y\}=\frac{1}{2},$$

$$P\{Y < X \leqslant 2Y\} = \frac{1}{4}.$$

(1) $(U,V)$有四个可能值：$(0,0),(0,1),(1,0),(1,1)$.

$$P\{U=0,V=0\} = P\{X \leqslant Y, X \leqslant 2Y\}$$
$$= P\{X \leqslant Y\} = \frac{1}{4};$$
$$P\{U=0,V=1\} = P\{X \leqslant Y, X > 2Y\} = 0;$$
$$P\{U=1,V=0\} = P\{X > Y, X \leqslant 2Y\}$$
$$= P\{Y < X \leqslant 2Y\} = \frac{1}{4};$$
$$P\{U=1,V=1\} = 1 - \left(\frac{1}{4} + \frac{1}{4}\right) = \frac{1}{2}.$$

(2) 由以上可见 $UV$ 以及 $U$ 和 $V$ 的分布为

$$UV \sim \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}; \quad U \sim \begin{bmatrix} 0 & 1 \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}, \quad V \sim \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

于是，有

$$E(U)=\frac{3}{4},\ D(U)=\frac{3}{16};\ E(V)=\frac{1}{2},\ D(V)=\frac{1}{4};\ E(UV)=\frac{1}{2};$$

$$\operatorname{cov}(U,V) = E(UV) - E(U)\cdot E(V) = \frac{1}{8};$$

$$r = \frac{\operatorname{cov}(U,V)}{\sqrt{D(U)\cdot D(V)}} = \frac{1}{\sqrt{3}}.$$

16. 设二维随机变量$(X,Y)$在矩形 $G=\{(x,y)\,|\,0\leqslant x\leqslant 2, 0\leqslant y\leqslant 1\}$（见图 3.7）上服从均匀分布，试求边长为 $X$ 和 $Y$ 的矩形面积 $S$ 的概率密度 $p(s)$.

**解** 二维随机变量$(X,Y)$的概率密度为

$$p(x,y) = \begin{cases} \frac{1}{2}, & 若(x,y) \in G, \\ 0, & 若(x,y) \overline{\in} G. \end{cases}$$

设 $F(s)=P\{S\leqslant s\}$为 $S$ 的分布函数，则

当 $s\leqslant 0$ 时，$F(s)=0$；当 $s\geqslant 2$ 时，$F(s)=1$.

现在，设 $0<s<2$. 曲线 $xy=s$ 与矩形 $G$ 的上边交于点 $(s,1)$；位于曲线 $xy=s$ 上方的点满足 $xy>s$，位于下方的点满足 $xy<s$，于是

$$
\begin{aligned}
F(s) &= P\{S\leqslant s\} = P\{XY\leqslant s\} = 1-P\{XY>s\} \\
&= 1-\iint\limits_{xy>s}\frac{1}{2}\mathrm{d}x\mathrm{d}y = 1-\frac{1}{2}\int_s^2\mathrm{d}x\int_{\frac{s}{x}}^1\mathrm{d}y \\
&= \frac{s}{2}(1+\ln 2-\ln s).
\end{aligned}
$$

由此得

$$
p(s)=\begin{cases}\dfrac{1}{2}(\ln 2-\ln s), & 若\ 0<s<2, \\ 0, & 若\ s\leqslant 0\ 或\ s\geqslant 2.\end{cases}
$$

17. 已知随机变量 $X_1$ 和 $X_2$ 的概率分布

$$
X_1\sim\begin{pmatrix}-1 & 0 & 1 \\ \dfrac{1}{4} & \dfrac{1}{2} & \dfrac{1}{4}\end{pmatrix},\quad X_2\sim\begin{pmatrix}0 & 1 \\ \dfrac{1}{2} & \dfrac{1}{2}\end{pmatrix},
$$

而且 $P\{X_1X_2=0\}=1$.

(1) 求 $X_1$ 和 $X_2$ 的联合分布；

(2) 问 $X_1$ 和 $X_2$ 是否独立？为什么？

**解　方法 1**　(1) 由 $P\{X_1X_2=0\}=1$，可见

$$
P\{X_1=-1,X_2=1\}=P\{X_1=1,X_2=1\}=0.
$$

易见

$$
P\{X_1=-1,X_2=0\}=P\{X_1=-1\}=\frac{1}{4};
$$

$$
P\{X_1=0,X_2=1\}=P\{X_2=1\}=\frac{1}{2};
$$

$$
P\{X_1=1,X_2=0\}=P\{X_1=1\}=\frac{1}{4};
$$

$$
P\{X_1=0,X_2=0\}=1-\left(\frac{1}{4}+\frac{1}{2}+\frac{1}{4}\right)=0.
$$

于是，得 $X_1$ 和 $X_2$ 的联合分布

| $X_2$ \ $X_1$ | $-1$ | $0$ | $1$ | $p_{i\cdot}$ |
|---|---|---|---|---|
| $0$ | $\frac{1}{4}$ | $0$ | $\frac{1}{4}$ | $\frac{1}{2}$ |
| $1$ | $0$ | $\frac{1}{2}$ | $0$ | $\frac{1}{2}$ |
| $p_{\cdot j}$ | $\frac{1}{4}$ | $\frac{1}{2}$ | $\frac{1}{4}$ | $1$ |

(2) 由以上结果，可见

$$P\{X_1 = 0, X_2 = 0\} = 0,$$

$$P\{X_1 = 0\}P\{X_2 = 0\} = \frac{1}{4} \neq 0,$$

于是，$X_1$ 和 $X_2$ 不独立.

**方法 2** 由 $P\{X_1X_2=0\}=1$，可见

$$P\{X_1 = -1, X_2 = 1\} = P\{X_1 = 1, X_2 = 1\} = 0.$$

因此，$X_1$ 和 $X_2$ 的联合分布有如下结构：

| $X_2$ \ $X_1$ | $-1$ | $0$ | $1$ | $p_{i\cdot}$ |
|---|---|---|---|---|
| $0$ | $p_{11}$ | $p_{21}$ | $p_{31}$ | $\frac{1}{2}$ |
| $1$ | $0$ | $p_{22}$ | $0$ | $\frac{1}{2}$ |
| $p_{\cdot j}$ | $\frac{1}{4}$ | $\frac{1}{2}$ | $\frac{1}{4}$ | $1$ |

于是，由上表易见 $X_1$ 和 $X_2$ 的联合分布为：

| $X_2$ \ $X_1$ | $-1$ | $0$ | $1$ | $p_{i\cdot}$ |
|---|---|---|---|---|
| $0$ | $\frac{1}{4}$ | $0$ | $\frac{1}{4}$ | $\frac{1}{2}$ |
| $1$ | $0$ | $\frac{1}{2}$ | $0$ | $\frac{1}{2}$ |
| $p_{\cdot j}$ | $\frac{1}{4}$ | $\frac{1}{2}$ | $\frac{1}{4}$ | $1$ |

(2) 同方法 1.

18. 袋中有 2 个白球和 3 个黑球，现从中依次摸出两球，设

$$X=\begin{cases}1,\text{第一次摸出白球},\\0,\text{第一次摸出黑球},\end{cases}\quad Y=\begin{cases}1,\text{第二次摸出白球},\\0,\text{第二次摸出黑球}.\end{cases}$$

试采用无放回摸球方式，求$(X,Y)$的联合分布和边缘分布.

**解** 利用古典概型的概率计算公式容易求得

$$p_{00}=p\{X=0,Y=0\}=\frac{A_3^2}{A_5^2}=\frac{3}{10},$$

$$p_{01}=p\{X=0,Y=1\}=\frac{3\times 2}{A_5^2}=\frac{3}{10},$$

$$p_{10}=p\{X=1,Y=0\}=\frac{2\times 3}{A_5^2}=\frac{3}{10},$$

$$p_{11}=p\{X=1,Y=1\}=\frac{2!}{A_5^2}=\frac{1}{10}.$$

由此可得 $X$ 的边缘分布为

$$P\{X=0\}=p_{00}+p_{01}=\frac{3}{10}+\frac{3}{10}=\frac{3}{5},$$

$$P\{X=1\}=p_{10}+p_{11}=\frac{3}{10}+\frac{1}{10}=\frac{2}{5};$$

$Y$ 的边缘分布为

$$P\{Y=0\}=p_{00}+p_{10}=\frac{3}{10}+\frac{3}{10}=\frac{3}{5},$$

$$P\{Y=1\}=p_{01}+p_{11}=\frac{3}{10}+\frac{1}{10}=\frac{2}{5}.$$

故$(X,Y)$的联合分布为

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

19. 设$(X,Y)$的分布律如下

| $X \backslash Y$ | 1 | 2 | 3 |
|---|---|---|---|
| 1 | $\frac{1}{6}$ | $\frac{1}{9}$ | $\frac{1}{18}$ |
| 2 | $\frac{1}{3}$ | $\alpha$ | $\beta$ |

问$\alpha$、$\beta$为何值时，$X$与$Y$相互独立.

**解** 由题意可知，$X$与$Y$的边缘分布律分别为

| $X$ | 1 | 2 |
|---|---|---|
| $p_i$ | $\frac{1}{3}$ | $\frac{1}{3}+\alpha+\beta$ |

和

| $Y$ | 1 | 2 | 3 |
|---|---|---|---|
| $p_j$ | $\frac{1}{2}$ | $\frac{1}{9}+\alpha$ | $\frac{1}{18}+\beta$ |

若$X$与$Y$相互独立，则必有$p_{ij}=p_i\cdot p_j,i=1,2,j=1,2,3$. 于是由

$$p_{12}=P\{X=1,Y=2\},\quad p_{13}=P\{X=1,Y=3\},$$

即

$$\frac{1}{9}=\frac{1}{3}\left(\frac{1}{9}+\alpha\right)\quad 及\quad \frac{1}{18}=\frac{1}{3}\left(\frac{1}{18}+\beta\right).$$

解之可得$\alpha=\frac{2}{9},\beta=\frac{1}{9}$. 此时有

$$p_{ij}=p_i\cdot p_j\quad (\forall\ i=1,2,\ j=1,2,3)\ 都成立,$$

即$X$与$Y$相互独立.

20. 设随机变量$(X,Y)$的概率密度为

$$p(x,y)=\begin{cases}ce^{-(3x+4y)}, & x>0,y>0,\\ 0, & 其他.\end{cases}$$

试求：(1) 常数 $c$；(2) 联合分布函数 $F(x,y)$；(3) 讨论 $X$ 与 $Y$ 的独立性.

**解** (1) 由 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x\mathrm{d}y=1$，有

$$1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x\mathrm{d}y=\int_0^{+\infty}\int_0^{+\infty}ce^{-(3x+4y)}\mathrm{d}x\mathrm{d}y$$
$$=c\int_0^{+\infty}\mathrm{e}^{-3x}\mathrm{d}x\int_0^{+\infty}\mathrm{e}^{-4y}\mathrm{d}y=\frac{c}{12}.$$

故 $c=12$.

(2) 设 $(X,Y)$ 的分布函数为 $F(x,y)$，则

当 $x>0,y>0$ 时，

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}p(x,y)\mathrm{d}x\mathrm{d}y=\int_0^x\int_0^y 12\mathrm{e}^{-(3x+4y)}\mathrm{d}x\mathrm{d}y$$
$$=(1-\mathrm{e}^{-3x})(1-\mathrm{e}^{-4y})；$$

当 $x,y$ 为其他情形时，$F(x,y)=0$.

故

$$F(x,y)=\begin{cases}(1-\mathrm{e}^{-3x})(1-\mathrm{e}^{-4y}), & x>0,y>0,\\ 0, & \text{其他}.\end{cases}$$

(3) $X$ 与 $Y$ 的边缘分布为 $p_1(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y$.

当 $x>0$ 时，$p_1(x)=\int_0^{+\infty}12\mathrm{e}^{-(3x+4y)}\mathrm{d}y=3\mathrm{e}^{-3x}$；

当 $x\leqslant 0$ 时，$p_1(x)=\int_{-\infty}^{+\infty}0\mathrm{d}y=0$.

因此

$$p_1(x)=\begin{cases}3\mathrm{e}^{-3x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

同理 $$p_2(y)=\begin{cases}4\mathrm{e}^{-y}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

由于 $p(x,y)=p_1(x)p_2(y)$，故 $X$ 与 $Y$ 是相互独立的.

21. 设 $(X,Y)$ 具有联合密度函数

$$p(x,y)=\begin{cases}\dfrac{1+xy}{4}, & |x|<1,|y|<1,\\ 0, & \text{其他}.\end{cases}$$

试证 $X$ 与 $Y$ 不独立，但 $X^2$ 与 $Y^2$ 是相互独立的.

**证** 当 $|x|<1$ 时，

$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y=\int_{-1}^{1}\frac{1+xy}{4}\mathrm{d}y=\frac{1}{2},$$

当 $|x|\geqslant 1$ 时，$p_X(x)=0$，即

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

同理

$$p_Y(y)=\begin{cases}\dfrac{1}{2}, & |y|<1,\\ 0, & \text{其他}.\end{cases}$$

当 $0<|x|<1,0<|y|<1$ 时，$p(x,y)\neq p_X(x)p_Y(y)$，所以 $X$ 与 $Y$ 不独立.

现通过分布函数来证 $X^2$ 与 $Y^2$ 独立. $X^2$ 的分布函数记为 $F_1(x)$，则当 $0\leqslant x<1$ 时，有

$$F_1(x)=P\{X^2\leqslant x\}=P\{-\sqrt{x}\leqslant X\leqslant\sqrt{x}\}$$
$$=\int_{-\sqrt{x}}^{\sqrt{x}}\frac{1}{2}\mathrm{d}x=\sqrt{x}.$$

同理可求得 $Y^2$ 的分布函数 $F_2(y)$，于是，有

$$F_1(x)=\begin{cases}0, & x<0;\\ \sqrt{x}, & 0\leqslant x<1;\\ 1, & x\geqslant 1,\end{cases}\quad F_2(y)=\begin{cases}0, & y<0;\\ \sqrt{y}, & 0\leqslant y<1;\\ 1, & y\geqslant 1.\end{cases}$$

$(X^2,Y^2)$ 的分布函数记为 $F_3(x,y)$，则

当 $x<0$ 或 $y<0$ 时，$F_3(x,y)=0$；

当 $0\leqslant x<1,y\geqslant 1$ 时，

$$F_3(x,y)=P\{X^2\leqslant x,Y^2\leqslant y\}=P\{X^2\leqslant x\}=\sqrt{x}.$$

同理，当 $0 \leqslant y < 1, x \geqslant 1$ 时，$F_3(x,y) = \sqrt{y}$；

当 $0 \leqslant x < 1, 0 \leqslant y < 1$ 时，

$$
\begin{aligned}
F_3(x,y) &= P\{X^2 \leqslant x, Y^2 \leqslant y\} \\
&= P\{-\sqrt{x} \leqslant X \leqslant \sqrt{x}, -\sqrt{y} \leqslant Y \leqslant \sqrt{y}\} \\
&= \int_{-\sqrt{x}}^{\sqrt{x}} \mathrm{d}s \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1+st}{4} \mathrm{d}t = \sqrt{xy};
\end{aligned}
$$

当 $x \geqslant 1, y \geqslant 1$ 时，

$$
\begin{aligned}
F_3(x,y) &= P\{X^2 \leqslant x, Y^2 \leqslant y\} \\
&= \int_{-1}^{1}\int_{-1}^{1} \frac{1+xy}{4} \mathrm{d}x\mathrm{d}y = 1.
\end{aligned}
$$

综合起来得

$$
F_3(x,y) = \begin{cases} 0, & x < 0 \text{ 或 } y < 0, \\ \sqrt{x}, & 0 \leqslant x < 1, y \geqslant 1, \\ \sqrt{y}, & 0 \leqslant y < 1, x \geqslant 1, \\ \sqrt{xy}, & 0 \leqslant x < 1, 0 \leqslant y < 1, \\ 1, & x \geqslant 1, y \geqslant 1. \end{cases}
$$

不难验证 $F_3(x,y) = F_1(x)F_2(y)$ 对所有的 $x, y$ 都成立，所以 $X^2$ 与 $Y^2$ 独立.

22. 设 $X$ 与 $Y$ 为两个相互独立的随机变量，它们的分布密度分别为

$$
p_1(x) = \begin{cases} \mathrm{e}^{-x}, & x > 0, \\ 0, & x \leqslant 0; \end{cases} \quad p_2(y) = \begin{cases} \mathrm{e}^{-y}, & y > 0, \\ 0, & y \leqslant 0. \end{cases}
$$

试求：(1) $(X,Y)$ 的分布密度；(2) $P(X \leqslant 1 \mid y > 0)$.

**解** (1) 因为 $X$ 与 $Y$ 相互独立，所以 $(X,Y)$ 的分布密度为

$$
\begin{aligned}
p(x,y) &= p_1(x)p_2(y) \\
&= \begin{cases} \mathrm{e}^{-(x+y)}, & x > 0, y > 0, \\ 0, & \text{其他}. \end{cases}
\end{aligned}
$$

(2) $P(X \leqslant 1 \mid y > 0) = P(X \leqslant 1) = \int_{-\infty}^{1} p_1(x) \mathrm{d}x$

$$=\int_{-\infty}^{0}0\mathrm{d}x+\int_{0}^{1}\mathrm{e}^{-x}\mathrm{d}x=1-\mathrm{e}^{-1}.$$

23. 设 $\xi=(X,Y)\sim F(x,y)=A\left(B+\arctan\frac{x}{2}\right)\times\left(C+\arctan\frac{y}{3}\right)$. 求

(1) 常数 $A$、$B$、$C$ 之值;

(2) $\xi=(X,Y)$的分布密度函数 $p(x,y)$;

(3) $X$ 与 $Y$ 的边缘分布密度 $p_1(x)$、$p_2(y)$.

**解** (1) 由$F(+\infty,+\infty)=A\left(B+\frac{\pi}{2}\right)\left(C+\frac{\pi}{2}\right)=1$,

$$F(-\infty,+\infty)=A\left(B-\frac{\pi}{2}\right)\left(C+\frac{\pi}{2}\right)=0,$$

$$F(+\infty,-\infty)=A\left(B+\frac{\pi}{2}\right)\left(C-\frac{\pi}{2}\right)=0,$$

可导出 $A=\frac{1}{\pi^2}$, $B=C=\frac{\pi}{2}$.

(2) $p(x,y)=F''_{xy}(x,y)=\frac{1}{\pi^2}\frac{\frac{1}{2}}{1+\left(\frac{x}{2}\right)^2}\cdot\frac{\frac{1}{3}}{1+\left(\frac{y}{3}\right)^2}$

$$=\frac{6}{\pi^2}\cdot\frac{1}{4+x^2}\cdot\frac{1}{9+y^2}.$$

(3) 由 $p(x,y)=f_1(x)\cdot f_2(y)$,其中

$$f_1(x)=\frac{2}{\pi}\frac{1}{4+x^2}\ (-\infty<x<+\infty),$$

$$f_2(y)=\frac{3}{\pi}\frac{1}{9+y^2}\ (-\infty<y<+\infty).$$

考虑到 $\int_{-\infty}^{+\infty}f_1(x)\mathrm{d}x=1$, $\int_{-\infty}^{+\infty}f_2(y)\mathrm{d}y=1$,故

$$p_1(x)=\frac{2}{\pi}\frac{1}{4+x^2},\quad p_2(y)=\frac{3}{\pi}\frac{1}{9+y^2}.$$

24. 设随机向量 $\xi=(X,Y)$的联合分布密度函数为

$$p(x,y)=\begin{cases}C(x+y), & 0\leqslant y\leqslant x\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

(1) 求常数 $C$；(2) 求 $X$、$Y$ 的边缘分布密度函数 $p_1(x)$、$p_2(y)$；(3) 讨论 $X$ 与 $Y$ 的独立性；(4) 计算 $P\{X+Y\leqslant 1\}$.

**解** (1) 由于$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}\sigma=1$，即

$$\int_0^1\mathrm{d}x\int_0^x C(x+y)\mathrm{d}y=1.$$

可导出 $C=2$.

(2) 当 $x<0$ 或 $x>1$ 时，$p_1(x)=0$；当 $0\leqslant x\leqslant 1$ 时，

$$p_1(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y=\int_0^x 2(x+y)\mathrm{d}y=3x^2.$$

因此
$$p_1(x)=\begin{cases}3x^2, & 0\leqslant x\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

同理

$$p_2(y)=\begin{cases}1+2y-3y^2, & 0\leqslant y\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

(3) 由于 $p_1(x)\cdot p_2(y)\neq p(x,y)$，故 $X$ 与 $Y$ 不独立.

(4) $$P\{X+Y\leqslant 1\}=\iint\limits_{\substack{x+y\leqslant 1\\ 0\leqslant y\leqslant x\leqslant 1}}2(x+y)\mathrm{d}\sigma$$

$$=\int_0^{\frac{1}{2}}\mathrm{d}y\int_y^{1-y}2(x+y)\mathrm{d}x=\frac{1}{3}.$$

25. 设随机变量 $X\sim p_1(x)$，$Y\sim p_2(y)$，且

$$p(x,y)=p_1(x)\cdot p_2(y)+h(x,y)$$
$$(-\infty<x<+\infty,\ -\infty<y<+\infty)$$

为二维随机向量 $\xi=(X,Y)$ 的联合分布密度函数. 试证

(1) $h(x,y)\geqslant -p_1(x)p_2(y)$

$$(-\infty<x<+\infty,\ -\infty<y<+\infty);$$

(2) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}h(x,y)\mathrm{d}\sigma=0$.

**证** (1) 由 $p(x,y)$ 的性质：$p(x,y)\geqslant 0$，有

$$p_1(x)\cdot p_2(y)+h(x,y)\geqslant 0,$$

即

$$h(x,y) \geqslant -p_1(x) \cdot p_2(y)$$

$$(-\infty < x < +\infty, -\infty < y < +\infty).$$

(2) 由于$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)\mathrm{d}\sigma = 1$，有

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[p_1(x) \cdot p_2(y) + h(x,y)]\mathrm{d}\sigma = 1.$$

于是有

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(x,y)\mathrm{d}\sigma = 1 - \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p_1(x)p_2(y)\mathrm{d}x\mathrm{d}y$$

$$= 1 - \int_{-\infty}^{+\infty} p_1(x)\mathrm{d}x \cdot \int_{-\infty}^{+\infty} p_2(y)\mathrm{d}y = 1 - 1 = 0.$$

26. 已知随机变量$(X,Y)$的概率密度为

$$p(x,y) = A\mathrm{e}^{-ax^2+bxy-cy^2},$$

问在什么条件下 $X$ 和 $Y$ 相互独立.

**解** 先求 $X$ 和 $Y$ 的边缘分布密度：

$$p_1(x) = \int_{-\infty}^{+\infty} p(x,y)\mathrm{d}y = \int_{-\infty}^{+\infty} A\mathrm{e}^{-ax^2+bxy-cy^2}\mathrm{d}y$$

$$= A\mathrm{e}^{-ax^2}\mathrm{e}^{(bx/(2\sqrt{c}))^2}\int_{-\infty}^{+\infty}\mathrm{e}^{-(\sqrt{c}y-bx/(2\sqrt{c}))^2}\mathrm{d}y$$

$$= \frac{A}{\sqrt{c}}\mathrm{e}^{-ax^2+(bx/(2\sqrt{c}))^2}\int_{-\infty}^{+\infty}\mathrm{e}^{-t^2}\mathrm{d}t$$

$$\left(令 t = \sqrt{c}y - \frac{bx}{2\sqrt{c}}\right)$$

$$= A\sqrt{\frac{\pi}{c}}\mathrm{e}^{-\left(a-\frac{b^2}{4c}\right)x^2}, \quad -\infty < x < +\infty.$$

类似地

$$p_2(y) = A\sqrt{\frac{\pi}{a}}\mathrm{e}^{-\left(c-\frac{b^2}{4a}\right)y^2}, \quad -\infty < y < +\infty.$$

为了满足独立性条件

$$p(x,y) = p_1(x) \cdot p_2(y),$$

即有

$$Ae^{-ax^2+bxy-cy^2} = A^2 \frac{\pi}{\sqrt{ac}} e^{-\left(a-\frac{b^2}{4c}\right)x^2-\left(c-\frac{b^2}{4a}\right)y^2},$$

通过比较系数，得到

$$b = 0, \quad \sqrt{ac} = A\pi.$$

27. 设 $X$ 与 $Y$ 相互独立，它们分别服从参数为 $\lambda_1,\lambda_2$ 的泊松分布，试证 $Z=X+Y$ 服从参数为 $\lambda_1+\lambda_2$ 的泊松分布.

**证** 因为

$$P\{X = i\} = \frac{\lambda_1^i}{i!}e^{-\lambda_1} \quad (i = 0,1,2,\cdots),$$

$$P\{Y = j\} = \frac{\lambda_2^j}{j!}e^{-\lambda_2} \quad (j = 0,1,2,\cdots),$$

于是

$$\begin{aligned}
P\{Z = k\} &= \sum_{i=0}^{k} P\{X = i\}P\{Y = k-i\} \\
&= \sum_{i=0}^{k} \frac{\lambda_1^i}{i!}e^{-\lambda_1} \cdot \frac{\lambda_2^{k-i}}{(k-i)!}e^{-\lambda_2} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{k!}\sum_{i=0}^{k} \frac{k!}{i!(k-i)!}\lambda_1^i\lambda_2^{k-i} \\
&= \frac{(\lambda_1+\lambda_2)^k}{k!}e^{-(\lambda_1+\lambda_2)} \quad (k = 0,1,2,\cdots),
\end{aligned}$$

即 $Z$ 服从参数为 $\lambda_1+\lambda_2$ 的泊松分布.

28. 设随机向量 $(X,Y)$ 的分布密度函数为

$$p(x,y) = \begin{cases} 3x, & 0 < x < 1, 0 < y < x, \\ 0, & \text{其他}, \end{cases}$$

求 $Z=X-Y$ 的分布密度函数.

**解** 当 $z<0$ 时，$F_Z(z)=0$；

当 $0\leqslant z<1$ 时，

$$F_Z(z)=\iint\limits_{x-y\leqslant z} p(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\iint\limits_{D_1}3x\mathrm{d}x\mathrm{d}y+\iint\limits_{D_2}3x\mathrm{d}x\mathrm{d}y$$

(积分区域 $D_1$ 与 $D_2$ 见图 3.8)

$$=\int_0^z\mathrm{d}x\int_0^x 3x\mathrm{d}y+\int_z^1\mathrm{d}x\int_{x-z}^x 3x\mathrm{d}y$$

$$=\int_0^z 3x^2\mathrm{d}x+\int_z^1 3xz\mathrm{d}x=\frac{3}{2}z-\frac{1}{2}z^3;$$

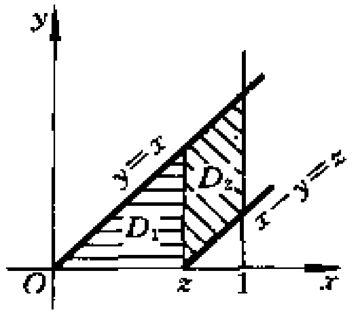



图 3.8

当 $z\geqslant1$ 时，

$$F_Z(z)=\int_0^1\mathrm{d}x\int_0^x 3x\mathrm{d}y=1.$$

综上所述，得

$$F_Z(z)=\begin{cases}0, & z<0,\\ \dfrac{3}{2}z-\dfrac{1}{2}z^3, & 0\leqslant z<1,\\ 1 & z\geqslant1.\end{cases}$$

故 $Z=X-Y$ 的分布密度函数为

$$p_Z(z)=F'_Z(z)=\begin{cases}\dfrac{3}{2}(1-z^2), & 0\leqslant z<1,\\ 0, & 其他.\end{cases}$$

29. 设 $X$、$Y$ 相互独立，服从相同的分布 $U(0,1)$，求 $Z=X/Y$ 的分布密度函数 $p_3(z)$.

**解** 由题意有

$$p_1(x)=\begin{cases}1, & 0\leqslant x\leqslant1,\\ 0, & 其他;\end{cases}\qquad p_2(y)=\begin{cases}1, & 0\leqslant y\leqslant1,\\ 0, & 其他.\end{cases}$$

由于 $X$ 与 $Y$ 相互独立，故 $(X,Y)$ 的联合分布密度函数为

$$p(x,y)=p_1(x)\cdot p_2(y)$$

$$=\begin{cases}1, & 0\leqslant x\leqslant1, 0\leqslant y\leqslant1,\\ 0, & 其他.\end{cases}$$

分三种情况讨论：

(1) 当 $z \geqslant 1$ 时，

$$F_3(z)=P(Z\leqslant z)=\iint\limits_{\frac{x}{y}\leqslant z} p(x,y)\mathrm{d}\sigma=\int_0^1 \mathrm{d}x\int_{\frac{x}{z}}^1 \mathrm{d}y=1-\frac{1}{2z},$$

$$p_3(z)=\frac{1}{2z^2};$$

(2) 当 $0\leqslant z<1$ 时，

$$F_3(z)=\iint\limits_{\frac{x}{y}\leqslant z} p(x,y)\mathrm{d}\sigma=\int_0^1 \mathrm{d}y\int_0^{yz}\mathrm{d}x=\frac{z}{2},$$

$$p_3(z)=\frac{1}{2};$$

(3) 当 $z<0$ 时，$p(x,y)=0$，因此 $F_3(z)=0, p_3(z)=0$.

于是，我们有

$$p_3(z)=\begin{cases}0, & z<0,\\ \dfrac{1}{2}, & 0\leqslant z<1,\\ \dfrac{1}{2z^2}, & z\geqslant 1.\end{cases}$$

30. 设 $X$、$Y$ 相互独立服从相同的分布 $N(0,1)$，求 $Z=\sqrt{X^2+Y^2}$ 的分布密度函数 $p_3(z)$.

**解** 由题意有

$$p_1(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}, \quad -\infty<x<+\infty,$$

$$p_2(y)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2}{2}}, \quad -\infty<y<+\infty.$$

由于 $X$ 与 $Y$ 相互独立，故 $(X,Y)$ 的联合分布密度函数为

$$p(x,y)=\frac{1}{2\pi}\mathrm{e}^{-\frac{x^2+y^2}{2}}, \quad (x,y)\in R^2.$$

由分布函数的定义，有

$$F_3(z)=P\{Z\leqslant z\}=P\{\sqrt{X^2+Y^2}\leqslant z\}$$
$$=\iint\limits_{\sqrt{x^2+y^2}\leqslant z}\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}d\sigma.$$

当 $z>0$ 时，$F_3(z)=\int_0^{2\pi}d\theta\int_0^z\frac{1}{2\pi}e^{-\frac{r^2}{2}}rdr=\int_0^z re^{-\frac{r^2}{2}}dr$，

$$p_3(z)=ze^{-\frac{z^2}{2}};$$

当 $z\leqslant 0$ 时，$F_3(z)=0$，$p_3(z)=0$.

于是，我们有

$$p_3(z)=\begin{cases}ze^{-\frac{z^2}{2}}, & z>0,\\ 0, & z\leqslant 0.\end{cases}$$

上式称之为**瑞利(Rayleigh)分布**.

31. 设 $X$ 与 $Y$ 相互独立，且均在区间[0,1]上服从均匀分布，求 $Z=X+Y$ 的密度函数.

**解　方法 1**　$X$ 与 $Y$ 的密度函数分别为

$$p_X(x)=\begin{cases}1, & 0\leqslant x\leqslant 1,\\ 0, & 其他,\end{cases}\qquad p_Y(y)=\begin{cases}1, & 0\leqslant y\leqslant 1,\\ 0, & 其他.\end{cases}$$

因为 $X$ 与 $Y$ 相互独立，所以

$$(X,Y)\sim p(x,y)=\begin{cases}1, & 0\leqslant x\leqslant 1, 0\leqslant y\leqslant 1,\\ 0, & 其他.\end{cases}$$

$$Z=X+Y\text{ 的分布函数 }F_Z(z)=\iint\limits_{x+y\leqslant z}p(x,y)dxdy.$$

当 $z<0$ 时，区域 $x+y\leqslant z$ 与正方形区域 $0\leqslant x\leqslant 1, 0\leqslant y\leqslant 1$ 没有公共部分，故 $F_Z(z)=0$；

当 $0\leqslant z<1$ 时，区域 $x+y\leqslant z$ 与正方形区域 $0\leqslant x\leqslant 1, 0\leqslant y\leqslant 1$ 的公共部分如图 3.9 所示的三角形区域，所以

$$F_Z(z)=\int_0^z dx\int_0^{z-x}dy=\frac{1}{2}z^2;$$

当 $1\leqslant z\leqslant 2$ 时，区域 $x+y\leqslant z$ 与正方形区域 $0\leqslant x\leqslant 1, 0\leqslant y\leqslant$

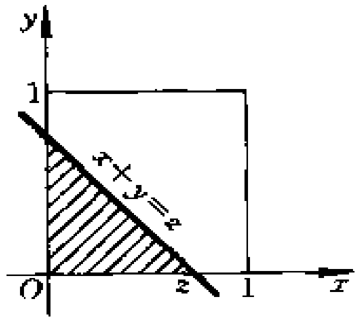



图 3.9

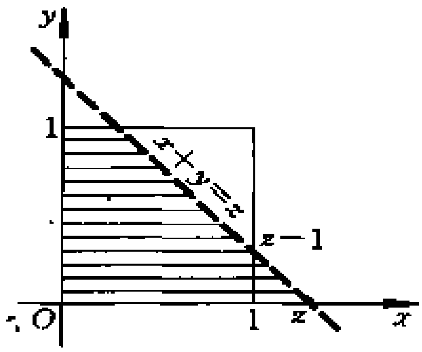



图 3.10

1 的公共部分如图 3.10 所示，所以

$$F_Z(z)=1-\frac{1}{2}[1-(z-1)]^2$$

$$=1-\frac{1}{2}(2-z)^2=2z-1-\frac{1}{2}z^2;$$

当 $z\geqslant 2$ 时，区域 $x+y\leqslant z$ 包含整个正方形区域 $0\leqslant x\leqslant 1, 0\leqslant y\leqslant 1$，所以 $F_Z(z)=1$.

综合起来，有

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

于是 $Z=X+Y$ 的密度函数为

$$p_Z(z)=F'_Z(z)=\begin{cases}z, & 0\leqslant z<1,\\ 2-z, & 1\leqslant z<2,\\ 0, & \text{其他}.\end{cases}$$

**方法 2** 因为 $X$ 与 $Y$ 独立，由公式

$$p_Z(z)=\int_{-\infty}^{+\infty}p_X(z-y)p_Y(y)\mathrm{d}y$$

可分以下几种情况讨论.

当 $z<0$ 时，$z-y$ 与 $y$ 至少有一个为负，被积函数为 0，所以 $p_Z(z)=0$；

当 $0\leqslant z<1$ 时，欲使被积函数非零，必须有 $0\leqslant z-y\leqslant 1, 0\leqslant y\leqslant 1$，从而 $0\leqslant y\leqslant z$，所以

$$p_Z(z)=\int_0^z \mathrm{d}y=z;$$

当 $1\leqslant z<2$ 时，欲使被积函数非零，必须有 $0\leqslant z-y\leqslant 1, 0\leqslant y\leqslant 1$，从而有 $z-1\leqslant y<1$，所以

$$p_Z(z)=\int_{z-1}^1 \mathrm{d}y=2-z;$$

当 $z\geqslant 2$ 时，$0\leqslant y\leqslant 1, 0\leqslant z-y\leqslant 1$ 不可能同时满足，故 $p_Z(z)=0$.

综合起来，有

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

32. 设 $X$ 与 $Y$ 相互独立，它们都服从指数分布，即

$$p_X(x)=\begin{cases} \lambda e^{-\lambda x}, & x\geqslant 0, \\ 0, & x<0, \end{cases} \quad p_Y(y)=\begin{cases} \lambda e^{-\lambda y}, & y\geqslant 0, \\ 0, & y<0. \end{cases}$$

求 $Z=\dfrac{X}{Y}$ 的密度函数.

**解** $Z=\dfrac{X}{Y}$ 的分布函数为

$$\begin{aligned} F_Z(z)&=P\{Z\leqslant z\}=P\left\{\frac{X}{Y}\leqslant z\right\} \\ &=\iint\limits_{\frac{x}{y}\leqslant z} p_X(x)p_Y(y)\mathrm{d}x\mathrm{d}y \\ &=\int_0^{+\infty}\mathrm{d}y\int_{-\infty}^{yz} p_X(x)p_Y(y)\mathrm{d}x \\ &\quad+\int_{-\infty}^0\mathrm{d}y\int_{yz}^{+\infty} p_X(x)p_Y(y)\mathrm{d}x, \end{aligned}$$

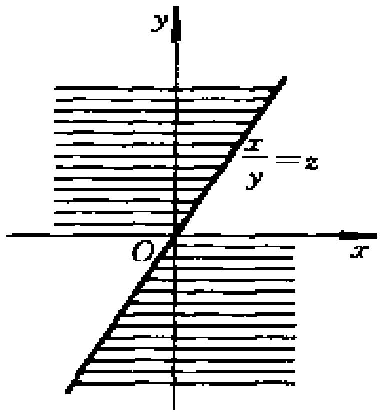



图 3.11

于是,$Z=X/Y$ 的密度函数为

$$p_Z(z)=F_Z'(z)=\int_0^{+\infty}p_X(yz)p_Y(y)y\mathrm{d}y-\int_{-\infty}^{0}p_X(yz)p_Y(y)y\mathrm{d}y$$

$$=\int_{-\infty}^{+\infty}p_X(yz)p_Y(y)|y|\mathrm{d}y;$$

当 $z>0$ 时,$p_Z(z)=\lambda^2\int_0^{+\infty}\mathrm{e}^{-\lambda y(1+z)}y\mathrm{d}y=\dfrac{1}{(1+z)^2}$;

当 $z\leqslant 0$ 时,$p_Z(z)=0$.

综合起来,有

$$p_Z(z)=\begin{cases}\dfrac{1}{(1+z)^2}, & z>0,\\ 0, & z\leqslant 0.\end{cases}$$

33. 设 $X$ 与 $Y$ 相互独立,它们都服从标准正态分布 $N(0,1)$,求 $Z=X^2+Y^2$ 的密度函数.

**解** 当 $z>0$ 时,

$$F_Z(z)=P\{Z\leqslant z\}=P\{X^2+Y^2\leqslant Z\}$$

$$=\iint\limits_{x^2+y^2\leqslant z}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2}{2}}\mathrm{d}x\mathrm{d}y$$

$$=\int_0^{2\pi}\mathrm{d}\theta\int_0^{\sqrt{z}}\frac{1}{2\pi}\mathrm{e}^{-\frac{r^2}{2}}r\mathrm{d}r$$

$$=1-\mathrm{e}^{-z/2};$$

当 $z\leqslant 0$ 时,

$$F_Z(z)=P\{Z\leqslant z\}=P\{X^2+Y^2\leqslant z\}=0.$$

于是

$$F_Z(z)=\begin{cases}1-\mathrm{e}^{-z/2}, & z>0,\\ 0, & z\leqslant 0.\end{cases}$$

从而,$Z$ 的密度函数为

$$p_Z(z)=F_Z'(z)=\begin{cases}\dfrac{1}{2}\mathrm{e}^{-z/2}, & z>0,\\ 0, & z\leqslant 0.\end{cases}$$

34. 在区间[0,1]上随机地投掷两点，试求这两点间距离的密度函数.

**解** 设 $X$ 和 $Y$ 分别表示两投点的坐标，那么$(X,Y)$服从二维均匀分布 $U[0,1;0,1]$，其联合密度函数为

$$p_{X,Y}(x,y)=\begin{cases}1, & 0\leqslant x\leqslant 1, 0\leqslant y\leqslant 1;\\ 0, & \text{其他}.\end{cases}$$

令 $Z=|X-Y|$，则 $Z$ 的分布函数为

$$F_Z(z)=P\{Z\leqslant z\}=P\{|X-Y|\leqslant z\}$$
$$=\iint\limits_{|x-y|\leqslant z} p_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y.$$

当 $z<0$ 时，$F_Z(z)=0$；

当 $0\leqslant z<1$ 时，区域 $|x-y|\leqslant z$ 与正方形区域 $0\leqslant x\leqslant 1, 0\leqslant y\leqslant 1$ 的公共部分如图 3.12 所示，并用 $S$ 表示其面积，则

$$S=1-(1-z)^2=2z-z^2.$$

于是

$$F_Z(z)=\iint\limits_{S}\mathrm{d}x\mathrm{d}y=2z-z^2.$$

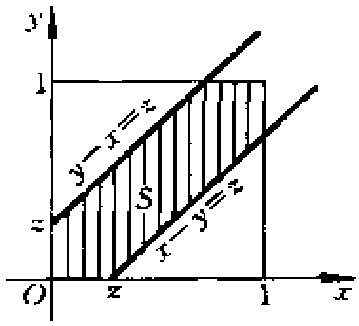


图 3.12

当 $z\geqslant 1$ 时，区域 $|x-y|\leqslant z$ 包含整个正方形区域：

$$0\leqslant x\leqslant 1,\quad 0\leqslant y\leqslant 1,$$

于是

$$F_Z(z)=\int_0^1\int_0^1\mathrm{d}x\mathrm{d}y=1.$$

综上所述，有

$$F_Z(z)=\begin{cases}0, & z<0,\\ 2z-z^2, & 0\leqslant z<1,\\ 1, & z\geqslant 1,\end{cases}$$

从而，$Z=|X-Y|$的密度函数为

$$p_Z(z) = F'_Z(z) = \begin{cases} 2(1-z), & 0 \leqslant z < 1, \\ 0, & \text{其他}. \end{cases}$$

35. 设随机变量 $X$ 与 $Y$ 相互独立，它们的分布函数分别为 $F_X(x)$ 与 $F_Y(y)$. 求 (1) $Z_1 = \max(X,Y)$ 的分布函数；(2) $Z_2 = \min(X,Y)$ 的分布函数.

**解** (1) $Z_1$ 的分布函数

$$\begin{aligned} F_{Z_1}(z) &= P\{Z_1 \leqslant z\} = P\{\max(X,Y) \leqslant z\} \\ &= P\{X \leqslant z, Y \leqslant z\} \\ &= P\{Z \leqslant z\}P\{Y \leqslant z\} = F_X(z)F_Y(z). \end{aligned}$$

(2) $Z_2$ 的分布函数

$$\begin{aligned} F_{Z_2}(z) &= P\{Z_2 \leqslant z\} = P\{\min(X,Y) \leqslant z\} \\ &= 1 - P\{\min(X,Y) > z\} \\ &= 1 - P\{X > z, Y > z\} \\ &= 1 - P\{X > 2\}P\{Y > z\} \\ &= 1 - [1 - P\{X \leqslant z\}][1 - P\{Y \leqslant z\}] \\ &= 1 - [1 - F_X(z)][1 - F_Y(z)]. \end{aligned}$$

特别，若 $X$ 与 $Y$ 独立同分布，分布函数为 $F(x)$，则

$Z_1$ 的分布函数 $F_1(z) = [F(z)]^2$；

$Z_2$ 的分布函数 $F_2(z) = 1 - [1 - F(z)]^2$.

此结论可以推广，若 $X_1, X_2, \cdots, X_n$ 独立同分布，分布函数为 $F(x)$，则

$Z_1 = \max\{X_1, X_2, \cdots, X_n\}$ 的分布函数 $F_1(z) = [F(z)]^n$；

$Z_2 = \min\{X_1, X_2, \cdots, X_n\}$ 的分布函数

$$F_2(z) = 1 - [1 - F(z)]^n.$$

## 四、练习题

1. 设二维随机变量 $(X,Y)$ 的联合分布函数为

$$F(x,y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{-(x+y)}, & \text{当 } x \geqslant 0, y \geqslant 0, \\ 0, & \text{其他}, \end{cases}$$

求二维随机变量$(X,Y)$的联合密度 $p(x,y)$.

2. 设$(X,Y)$服从二维正态分布,其概率密度函数为

$$p(x,y)=\frac{1}{2\pi 10^2}e^{-\frac{1}{2}\left(\frac{x^2}{10^2}+\frac{y^2}{10^2}\right)}$$

$$(-\infty<x<+\infty,\ -\infty<y<+\infty),$$

求 $P\{X<Y\}$.

3. 设二维随机变量$(X,Y)$在区域 $D$ 上服从均匀分布,其中 $D$ 为 $x$ 轴,$y$ 轴及直线 $y=2x+1$ 所围成的三角形区域. 求$(X,Y)$的联合密度函数及边缘密度函数.

4. 设二维随机变量$(X,Y)$的密度函数为

$$p(x,y)=\begin{cases}C(R-\sqrt{x^2+y^2}), & x^2+y^2\leqslant R^2,\\ 0, & \text{其他},\end{cases}$$

求(1) 常数 $C$;(2) 当 $R=2$ 时,二维随机变量$(X,Y)$在以原点为圆心,$r=1$ 为半径的圆域内的概率.

5. 设$(X,Y)$的联合密度函数为

$$p(x,y)=\begin{cases}\dfrac{x^{m-1}(y-x)^{n-1}}{\Gamma(m)\Gamma(n)}e^{-y}, & y\geqslant x>0,\\ 0, & \text{其他},\end{cases}$$

其中 $m>0,n>0$. 求边缘密度函数.

6. 箱中装有 12 只开关,其中有 2 只次品,从中任取两次,每次取一只. 现定义随机变量 $X,Y$ 如下:

$$X=\begin{cases}1, & \text{第一次取出正品},\\ 0, & \text{第一次取出次品};\end{cases}\quad Y=\begin{cases}1, & \text{第二次取出正品},\\ 0, & \text{第二次取出次品}.\end{cases}$$

试分别由(1) 放回抽取;(2) 不放回抽取两种情况,求出$(X,Y)$的联合概率分布.

7. 设 $X$ 服从均匀分布 $U[a,b]$$(a\geqslant 0)$,$Y$ 服从指数分布,即

$$p_X(x)=\begin{cases}\dfrac{1}{b-a}, & a\leqslant x\leqslant b,\\ 0, & \text{其他};\end{cases}\quad p_Y(y)=\begin{cases}\lambda e^{-\lambda y}, & y\geqslant 0,\\ 0, & y<0,\end{cases}$$

且 $X$ 与 $Y$ 相互独立. 试求

(1) $(X,Y)$的联合密度函数;

(2) $P\{Y\leqslant X\}$.

8. 设二维随机变量$(X,Y)$的密度函数

$$p(x,y)=A \quad (x^2+y^2<r^2),$$

试求常数 $A$ 和边缘分布密度函数,并确定 $X$ 与 $Y$ 的相互独立性.

9. 设$(X,Y)$的分布律为:

| $X \backslash Y$ | $-1$ | $1$ | $2$ |
|---|---|---|---|
| $-2$ | $\frac{1}{12}$ | $\frac{2}{12}$ | $\frac{2}{12}$ |
| $-1$ | $\frac{1}{12}$ | $\frac{1}{12}$ | $0$ |
| $0$ | $\frac{2}{12}$ | $\frac{1}{12}$ | $\frac{2}{12}$ |

求 $Z=X+Y$ 的分布律.

10. 设 $X$、$Y$ 是相互独立且服从同一分布的两个随机变量,已知 $X$ 的分布律为 $P\{X=i\}=\frac{1}{3}(i=1,2,3)$. 求 $Z=X+Y$ 的分布律.

11. 设$(X,Y)$服从二维均匀分布 $U[-1,1;-1,1]$,其联合密度

$$p(x,y)=\begin{cases}\frac{1}{4}, & -1\leqslant x\leqslant 1,\ -1\leqslant y\leqslant 1,\\ 0, & \text{其他},\end{cases}$$

求 $Z=XY$ 的密度函数.

12. 设$(X,Y)$的联合密度函数为

$$p(x,y)=\begin{cases}e^{-(x+y)}, & x>0,y>0,\\ 0, & \text{其他},\end{cases}$$

试求 $Z=\frac{X+Y}{2}$的密度函数.

13. 设某种商品一周的需要量为随机变量，其密度函数为

$$p(t)=\begin{cases}te^{-t}, & t>0,\\ 0, & t\leqslant 0.\end{cases}$$

又各周的需要量是相互独立的，试求(1) 两周的需要量的密度函数；(2) 三周的需要量的密度函数.

## 习题答案与提示

**1.** $p(x,y)=\begin{cases}e^{-x-y}, & x\geqslant 0, y\geqslant 0,\\ 0, & \text{其他}.\end{cases}$　　**2.** $\frac{1}{2}$.

**3.** $p(x,y)=\begin{cases}4, & (x,y)\in D,\\ 0, & \text{其他},\end{cases}$

$$p_X(x)=\begin{cases}8x+4, & -\frac{1}{2}<x<0,\\ 0, & \text{其他};\end{cases}$$

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

**4.** $\frac{3}{\pi R^3}$, $\frac{1}{2}$.

**5.** $p_1(x)=\begin{cases}\frac{x^{m-1}}{\Gamma(m)}e^{-x}, & x>0,\\ 0, & x\leqslant 0;\end{cases}$

$$p_2(y)=\begin{cases}\frac{y^{n+m-1}}{\Gamma(m+n)}e^{-y}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

**6.** (1)

| Y \ X | 0 | 1 |
|---|---|---|
| 0 | 25/36 | 5/36 |
| 1 | 5/36 | 1/35 |

(2)

| $Y$ \ $X$ | 0 | 1 |
|---|---|---|
| 0 | 15/22 | 5/33 |
| 1 | 5/33 | 1/66 |

7. (1) $p(x,y)=\begin{cases}\dfrac{\lambda}{b-a}e^{-\lambda y}, & a\leqslant x\leqslant b, y>0,\\ 0, & \text{其他};\end{cases}$

(2) $P(Y\leqslant X)=1+\dfrac{1}{\lambda(b-a)}(e^{-\lambda b}-e^{-\lambda a})$.

8. $A=\dfrac{1}{\pi r^2}\quad (x^2+y^2<r^2)$;

$$p_X(x)=\begin{cases}\dfrac{2}{\pi r^2}\sqrt{r^2-x^2}, & |x|<r,\\ 0, & |x|\geqslant r;\end{cases}$$

$$p_Y(y)=\begin{cases}\dfrac{2}{\pi r^2}\sqrt{r^2-y^2}, & |y|<r,\\ 0, & |y|\geqslant r.\end{cases}$$

不独立.

9.

| $Z$ | $-3$ | $-2$ | $-1$ | 0 | 1 | 2 |
|---|---|---|---|---|---|---|
| $p$ | $\frac{1}{12}$ | $\frac{1}{12}$ | $\frac{4}{12}$ | $\frac{3}{12}$ | $\frac{1}{12}$ | $\frac{2}{12}$ |

10.

| $Z$ | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| $p$ | $\frac{1}{9}$ | $\frac{2}{9}$ | $\frac{3}{9}$ | $\frac{2}{9}$ | $\frac{1}{9}$ |

11. $p_Z(z)=-\dfrac{1}{2}\ln|z|\quad (0<|z|<1)$.

12. $p_Z(z)=4ze^{-2z}\quad (z>0)$.

13. (1) $p_1(z)=\dfrac{1}{6}z^3e^{-z}\quad (z>0)$;

(2) $p_2(z)=\dfrac{1}{5!}z^5e^{-z}\quad (z>0)$.

# 第四章　随机变量的数字特征

## 一、考试要求

1. 理解随机变量的数字特征的概念，会运用数字特征的基本性质计算具体分布的数字特征；

2. 理解一维、二维随机变量函数的数学期望公式，会运用公式计算有关数学期望及方差；

3. 掌握常见几种分布的数字特征.

## 二、复习要点

### （一）重要概念及性质

**定义 4.1（数学期望）**

（1）离散型

设离散型随机变量 $X$ 的概率分布为

| $X$ | $x_1$ | $x_2$ | … | $x_n$ |
| --- | --- | --- | --- | --- |
| $p(X=x_i)$ | $p_1$ | $p_2$ | … | $p_n$ |

则称 $\sum_{i=1}^{n} x_i p_i$ 为 $X$ 的**数学期望**或**均值**，记作 $E(X)$.

当 $X$ 的可能取值 $x_i$ 为可列个时，则 $E(X) \xlongequal{\text{def}} \sum_{i=1}^{\infty} x_i p_i$，这时要求 $\sum_{i=1}^{\infty} |x_i| p_i < +\infty$，以保证和式 $\sum_{i=1}^{\infty} x_i p_i$ 的值不随和式中各项次序的改变而改变.

对于二维离散型随机向量 $\xi=(X,Y)$，如果其概率分布为 $p_{ij}$，

则其数学期望定义为 $E(\xi)\xlongequal{\text{def}}(E(X),E(Y))$，其中

$$E(X)=\sum_i\sum_j x_i p_{ij},\quad E(Y)=\sum_i\sum_j y_i p_{ij},$$

这里要求上述的级数都是绝对收敛的；

(2) 连续型

若连续型随机变量 $X$ 的密度函数为 $p(x)$，并且 $\int_{-\infty}^{+\infty}|x|p(x)\mathrm{d}x<+\infty$，则称 $\int_{-\infty}^{+\infty}xp(x)\mathrm{d}x$ 为 $X$ 的**数学期望**，记作 $E(X)$，即

$$E(X)=\int_{-\infty}^{+\infty}xp(x)\mathrm{d}x.$$

对于二维连续型随机向量 $\xi=(X,Y)$，如果其密度函数为 $p(x,y)$，则其数学期望定义为 $E(\xi)\xlongequal{\text{def}}(E(X),E(Y))$，其中

$$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xp(x,y)\mathrm{d}\sigma,\quad E(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yp(x,y)\mathrm{d}\sigma,$$

这里要求上述的积分都是绝对收敛的.

由定义可知，数学期望是加权平均数这一概念在随机变量中的推广，它反映了随机变量取值的平均水平. 其统计意义就是对随机变量进行长期观测或大量观测所得数值的理论平均数.

**数学期望的性质**

**性质 1** 常量 $C$ 的数学期望等于它自己，即

$$E(C)=C.$$

**性质 2** 常量 $C$ 与随机变量 $X$ 乘积的数学期望，等于常量 $C$ 与这个随机变量的数学期望的积，即

$$E(CX)=CE(X).$$

**性质 3** 随机变量和的数学期望，等于随机变量数学期望的和，即

$$E(X+Y)=E(X)+E(Y).$$

**推论** 有限个随机变量和的数学期望，等于它们各自数学期望的和，即

$$E\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}E(X_i).$$

**性质 4** 设随机变量 $X$ 与 $Y$ 相互独立，则它们乘积的数学期望等于它们数学期望的积，即

$$E(X\cdot Y)=E(X)\cdot E(Y).$$

**推论** 有限个相互独立的随机变量乘积的数学期望，等于它们各自数学期望的积，即

$$E\left(\prod_{i=1}^{n}X_i\right)=\prod_{i=1}^{n}E(X_i).$$

**例 1** 袋中有 5 个乒乓球，编号为 1，2，3，4，5，从中任取 3 个．以 $X$ 表示取出的 3 个球中的最大编号，求 $E(X)$．

**解** 由题意可知

$$X\sim\begin{pmatrix}3 & 4 & 5\\ \frac{1}{10} & \frac{3}{10} & \frac{6}{10}\end{pmatrix},$$

因此

$$E(X)=3\times\frac{1}{10}+4\times\frac{3}{10}+5\times\frac{6}{10}=\frac{9}{2}.$$

**例 2** 设随机变量 $X$ 的分布密度函数为

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

求 $E(X)$．

**解** 由定义，有

$$E(X)=\int_{-\infty}^{+\infty}xp(x)\mathrm{d}x=\int_0^1 2x(1-x)\mathrm{d}x=\frac{1}{3}.$$

**定义 4.2（方差）** 设离散随机变量的分布律为

| $X$ | $x_1$ | $x_2$ | $\cdots$ | $x_n$ | $\cdots$ |
|---|---|---|---|---|---|
| $P(X=x_i)$ | $p_1$ | $p_2$ | $\cdots$ | $p_n$ | $\cdots$ |

若$\sum\limits_{i=1}^{\infty}(x_i-E(X))^2p_i<+\infty$，则称此级数的和为 $X$ 的**方差**，记为 $D(X)$，即

$$D(X)=\sum_{i=1}^{\infty}(x_i-E(X))^2p_i.$$

对于连续型随机变量的方差有以下的定义.

设连续型随机变量 $X$ 的分布密度函数为 $p(x)$，若$\int_{-\infty}^{+\infty}(x-E(X))^2p(x)\mathrm{d}x<+\infty$，则称此无穷积分值为 $X$ 的**方差**，记为 $D(X)$，即

$$D(X)=\int_{-\infty}^{+\infty}(x-E(X))^2p(x)\mathrm{d}x.$$

由方差的定义和数学期望的性质，有

$$D(X)=E(X^2)-(E(X))^2.$$

这就是说，要计算随机变量 $X$ 的方差，在求出 $E(X)$后，再根据随机变量函数的数学期望公式算出 $E(X^2)$即可.

根据方差的定义显然有 $D(X)\geqslant 0$，我们称方差的算术根$\sqrt{D(X)}$为随机变量 $X$ 的**标准差**(**或均方差**). 这样，随机变量的标准差、数学期望与随机变量本身有相同的计量单位.

对于二维连续型随机向量 $\xi=(X,Y)$，如果其密度函数为 $p(x,y)$，则其方差定义为 $D(\xi)\xlongequal{\text{def}}(D(X),D(Y))$，其中

$$D(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x-E(X)]^2p(x,y)\mathrm{d}x\mathrm{d}y,$$

$$D(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[y-E(Y)]^2p(x,y)\mathrm{d}x\mathrm{d}y.$$

离散型的情况也有类似的定义，这里不再叙述了.

**方差的性质**

**性质 1** 常量 $C$ 的方差等于零，即

$$D(C)=0.$$

**性质 2** 随机变量 $X$ 与常量 $C$ 的和的方差，等于这个随机变

量的方差，即

$$D(X+C)=D(X).$$

**性质 3** 常量 $C$ 与随机变量 $X$ 乘积的方差，等于这个常量的平方与随机变量的方差的积，即

$$D(CX)=C^2D(X).$$

**性质 4** 设随机变量 $X$ 与 $Y$ 相互独立，则它们和的方差，等于它们的方差的和，即

$$D(X+Y)=D(X)+D(Y).$$

**推论** 有限个相互独立的随机变量和的方差，等于它们各自方差的和，即

$$D\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}D(X_i).$$

**性质 5** 对于一般的随机变量 $X$ 与 $Y$，则

$$\begin{aligned}D(X\pm Y)=&D(X)+D(Y)\\&\pm 2E[(X-E(X))(Y-E(Y))].\end{aligned}$$

**例 3** 设随机变量 $X$ 的概率密度为

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

求 $E(X)$ 及 $D(X)$.

**解** $E(X)=\int_{-\infty}^{+\infty}xp(x)\mathrm{d}x=\int_{-\infty}^{+\infty}x\,\frac{1}{2}e^{-|x|}\mathrm{d}x=0,$

$$\begin{aligned}D(X)&=\int_{-\infty}^{+\infty}[x-E(X)]^2p(x)\mathrm{d}x=\int_{-\infty}^{+\infty}x^2\,\frac{1}{2}e^{-|x|}\mathrm{d}x\\&=\int_{0}^{+\infty}x^2e^{-x}\mathrm{d}x=-e^{-x}(x^2+2x+2)\Big|_{0}^{+\infty}=2.\end{aligned}$$

**例 4** 已知随机变量 $X$ 的分布函数

$$F(x)=\begin{cases}0, & x\leqslant 0,\\ \dfrac{x}{4}, & 0<x\leqslant 4,\\ 1, & x>4.\end{cases}$$

求 $E(X)$，$D(X)$.

**解** 随机变量 $X$ 的分布密度为

$$p(x)=F'(x)=\begin{cases}\frac{1}{4}, & 0<x\leqslant 4,\\ 0, & \text{其他}.\end{cases}$$

$$E(X)=\int_{-\infty}^{+\infty}xp(x)\mathrm{d}x=\int_0^4\frac{x}{4}\mathrm{d}x=\frac{x^2}{8}\Big|_0^4=2,$$

$$E(X^2)=\int_{-\infty}^{+\infty}x^2p(x)\mathrm{d}x=\int_0^4\frac{x^2}{4}\mathrm{d}x=\frac{1}{4}\times\frac{1}{3}x^3\Big|_0^4=\frac{16}{3},$$

$$D(X)=E(X^2)-[E(X)]^2=\frac{16}{3}-2^2=\frac{4}{3}.$$

**定义 4.3(协方差)** 对于随机变量 $X$ 与 $Y$,称

$$\sigma_{XY}=E[(X-E(X))(Y-E(Y))]$$

为 $X$ 与 $Y$ 的**协方差**或**相关矩**,记为 $\sigma_{XY}$或 $\mathrm{cov}(X,Y)$.

与记号 $\sigma_{XY}$相对应,$X$ 与 $Y$ 的方差 $D(X)$与 $D(Y)$也可分别记为 $\sigma_{XX}$与 $\sigma_{YY}$.

**协方差的性质**

**性质 1** $\mathrm{cov}(X,Y)=\mathrm{cov}(Y,X)$.

**性质 2** $\mathrm{cov}(aX,bY)=ab\,\mathrm{cov}(X,Y)$.

**性质 3** $\mathrm{cov}(X_1+X_2,Y)=\mathrm{cov}(X_1,Y)+\mathrm{cov}(X_2,Y)$.

**性质 4** $\mathrm{cov}(X,Y)=E(XY)-(E(X))(E(Y))$.

由方差的性质可知,如果随机变量 $X$ 与 $Y$ 相互独立,则 $\sigma_{XY}=0$. 但是 $\sigma_{XY}=0$ 并不能保证 $X$ 与 $Y$ 相互独立.

**定义 4.4(相关系数)** 对于随机变量 $X$ 与 $Y$,如果 $D(X)>0$,$D(Y)>0$,则称

$$\frac{\sigma_{XY}}{\sqrt{D(X)}\sqrt{D(Y)}}$$

为 $X$ 与 $Y$ 的**相关系数**,记作 $\rho_{XY}$(有时可简记为 $\rho$).

例如,在$(X,Y)\sim(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$的情况下,可以推出

$$\rho_{XY}=\rho,$$

即二维正态分布的第五个参数 $\rho$ 就是相关系数.

相关系数具有以下两个性质：

**性质 1** 任意两个随机变量 $X$ 与 $Y$ 的相关系数 $\rho_{XY}$ 满足：

$$|\rho_{XY}| \leqslant 1.$$

若 $\rho \neq 0$，称 $X$ 与 $Y$ 是**相关的**；若 $\rho = 0$，称 $X$ 与 $Y$ **不相关**.

**性质 2** $|\rho_{XY}| = 1$ 的充要条件是随机变量 $X$ 与 $Y$ 之间以概率为 1 地存在着线性关系，即存在常数 $a$ 与 $b$，使

$$P\{Y = aX + b\} = 1.$$

相关系数 $\rho_{XY}$ 的实际意义是：它刻画了 $X,Y$ 之间线性关系的近似程度. 一般说来，$|\rho|$ 越接近于 1，$X$ 与 $Y$ 越近似地有线性关系. 要注意的是，当 $X,Y$ 之间有很密切的曲线关系时，$|\rho|$ 的数值也可能很小. 例如，$X$ 服从 $N(0,1)$，$Y = X^2$，此时 $Y$ 与 $X$ 有很密切的曲线关系，但是 $\rho_{XY} = 0$.

**定义 4.5（矩）**

(1) 原点矩

对于正整数 $k$，称随机变量 $X$ 的 $k$ 次幂的数学期望为 $X$ 的 $k$ **阶原点矩**，记为 $\nu_k$，即

$$\nu_k = E(X^k) \quad (k = 1,2,\cdots).$$

于是，我们有

$$\nu_k = \begin{cases} \sum\limits_i x_i^k p_i, & \text{当 } X \text{ 为离散型时}, \\ \int_{-\infty}^{+\infty} x^k p(x)\mathrm{d}x, & \text{当 } X \text{ 为连续型时}. \end{cases}$$

(2) 中心矩

对于正整数 $k$，称随机变量 $X$ 与 $E(X)$ 差的 $k$ 次幂的数学期望为 $X$ 的 $k$ 阶**中心矩**，记为 $\mu_k$，即

$$\mu_k = E(X - E(X))^k \quad (k = 1,2,\cdots).$$

于是，我们有

$$\mu_k = \begin{cases} \sum\limits_i (x_i - E(X))^k p_i, & \text{当 } X \text{ 为离散型时}, \\ \int_{-\infty}^{+\infty} (x - E(X))^k p(x)\mathrm{d}x, & \text{当 } X \text{ 为连续型时}. \end{cases}$$

由上述定义可知：

1）$X$ 的一阶原点矩就是 $X$ 的数学期望，即

$$\nu_1 = E(X);$$

2）$X$ 的二阶中心矩就是 $X$ 的方差，即

$$\mu_2 = D(X);$$

3）$X$ 的一阶中心矩为零，即

$$\mu_1 = E(X - E(X)) = 0;$$

4）$X$ 的二阶中心矩可用原点矩来表示，即

$$\mu_2 = \nu_2 - \nu_1^2.$$

(3) 混合矩

对于随机变量 $X$ 与 $Y$，如果有 $E(X^kY^l)$存在，则称之为 $X$ 与 $Y$ 的 $k+l$ 阶混合原点矩，记为 $\nu_{kl}$，即

$$\nu_{kl} = E(X^kY^l);$$

如果有 $E[(X-E(X))^k(Y-E(Y))^l]$存在，则称之为 $X$ 与 $Y$ 的 $k+l$ 阶混合中心矩，记为 $\mu_{kl}$，即

$$\mu_{kl} = E[(X - E(X))^k(Y - E(Y))^l].$$

由上述定义可知：

1）$X$ 与 $Y$ 的一阶混合原点矩有两个，它们分别是 $X$ 与 $Y$ 的数学期望，即

$$\nu_{10} = E(X), \quad \nu_{01} = E(Y);$$

2）在 $X$ 与 $Y$ 的二阶混合中心矩中有两个它们分别是 $X$ 与 $Y$ 的方差，即

$$\mu_{20} = D(X), \quad \mu_{02} = D(Y);$$

3）$X$ 与 $Y$ 的二阶混合中心矩 $\mu_{11}$就是 $X$ 与 $Y$ 的协方差，即

$$\mu_{11} = E[(X - E(X))(Y - E(Y))].$$

**定义 4.6（协方阵与相关阵）**

对于随机向量$(X,Y)$，我们把矩阵

$$\Sigma = \begin{pmatrix} \sigma_{XX} & \sigma_{XY} \\ \sigma_{YX} & \sigma_{YY} \end{pmatrix}$$

称为 $X,Y$ 的**协方阵**；把矩阵

$$R=\begin{bmatrix}\rho_{XX} & \rho_{XY}\\ \rho_{YX} & \rho_{YY}\end{bmatrix}$$

称为 $X,Y$ 的**相关阵**.

这样二元正态分布密度就可由矩阵表示为

$$p(x,y)=(2\pi)^{-\frac{2}{2}}|\Sigma|^{-\frac{1}{2}}\exp\{(X-\mu)'\Sigma^{-1}(X-\mu)\},$$

其中 $(X-\mu)'=(x-\mu_1,y-\mu_2)$,

$$\Sigma=\begin{bmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{bmatrix},$$

$|\Sigma|$ 为 $\Sigma$ 的行列式；$\Sigma^{-1}$ 为 $\Sigma$ 的逆矩阵.

于是我们可以把二维正态分布简记为

$$\xi\sim N(\mu,\Sigma),\quad 其中\quad \mu=(\mu_1,\mu_2)'.$$

特别地，二维标准正态分布可由记号 $N(0,I)$ 表示，其中 $I$ 为单位矩阵.

（二）重要定理及公式

**定理 4.1（表示性定理）**

（1）一维表示性定理

若随机变量 $X$ 的概率分布已确知，则随机变量函数 $f(X)$ 的数学期望为

$$E[f(X)]=\begin{cases}\sum\limits_{i=1}^{\infty}f(x_i)p_i, & 当\ X\ 为离散型时,\\ \int_{-\infty}^{+\infty}f(x)p(x)\mathrm{d}x, & 当\ X\ 为连续型时,\end{cases}$$

这里要求上述的级数与积分都是绝对收敛的.

**例 5** 设随机变量 $X$ 的分布密度函数为

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

求 $Y=e^{-2X}$ 的数学期望.

**解** 由定义，有

$$E(Y)=E(\mathrm{e}^{-2X})=\int_{-\infty}^{+\infty}\mathrm{e}^{-2x}p(x)\mathrm{d}x$$

$$=\int_{0}^{+\infty}\mathrm{e}^{-2x}\mathrm{e}^{-x}\mathrm{d}x=\frac{1}{3}.$$

(2) 二维表示性定理

对于二维连续型随机向量，类似地可以证明：若 $\xi=(X,Y)$ 的密度函数为 $p(x,y)$，则随机向量的函数 $f(X,Y)$ 的数学期望为

$$E[f(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)p(x,y)\mathrm{d}x\mathrm{d}y,$$

这里要求上述积分是绝对收敛的. 对于离散型的情况也有类似的定义，我们不再叙述了.

**定理 4.2(一维常见分布的数字特征)**

1. 离散型随机变量

(1) 0-1 分布

设 $X\sim B(1,p)$，即

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

则

$$E(X)=p,$$

$$D(X)=E[X-E(X)]^2$$

$$=(1-p)^2\cdot p+(0-p)^2\cdot q=pq.$$

(2) 二项分布

设 $X\sim B(n,p)$，即

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

$$(0<p<1,q=1-p;k=0,1,2,\cdots,n),$$

则

$$E(X)=\sum_{k=0}^{n}kC_n^k p^k q^{n-k}=\sum_{k=1}^{n}k\frac{n!}{k!(n-k)!}p^k q^{n-k}=np.$$

根据数学期望的公式可以算出

$$E(X^2)=n(n-1)p^2+np,$$

所以

$$D(X)=E(X^2)-(E(X))^2=npq.$$

(3) 超几何分布

设 $X\sim H(n,M,N)$,即

$$P\{X=k\}=\frac{C_M^kC_{N-M}^{n-k}}{C_N^n}$$

($n\leqslant N-M,k=0,1,2,\cdots,l$; $l=\min(M,n)$). 利用等式

$$\sum_{k=0}^{l}\frac{C_M^kC_{N-M}^{n-k}}{C_N^n}=1$$

不难推出

$$E(X)=\frac{nM}{N},\quad D(X)=\frac{n(N-n)(N-M)M}{N^2(N-1)}.$$

(4) 泊松分布

设 $X\sim P(\lambda)$,即

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

则

$$E(X)=\sum_{k=0}^{\infty}k\frac{\lambda^ke^{-\lambda}}{k!}=\lambda,$$

$$D(X)=E(X^2)-(E(X))^2=\lambda^2+\lambda-\lambda^2=\lambda.$$

(5) 几何分布

设 $X\sim G(p)$,即

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

则

$$E(X)=\sum_{k=1}^{\infty}kpq^{k-1}=\frac{1}{p}.$$

根据数学期望的公式可以算出

$$E(X^2)=\frac{2-p}{p^2},$$

所以

$$D(X)=E(X^2)-(E(X))^2=\frac{1-p}{p^2}.$$

2. 连续型随机变量

(1) 均匀分布

设 $X\sim U(a,b)$,则

$$E(X)=\int_{-\infty}^{+\infty}xp(x)\mathrm{d}x=\int_a^b x\frac{1}{b-a}\mathrm{d}x=\frac{a+b}{2}.$$

由 $E(X^2)=\frac{a^2+ab+b^2}{3}$,可以推出

$$D(X)=E(X)^2-(E(X))^2=\frac{(b-a)^2}{12}.$$

(2) 指数分布

设 $X\sim\Gamma(1,\lambda)$,则

$$E(X)=\int_{-\infty}^{+\infty}xp(x)\mathrm{d}x=\int_0^{+\infty}x\lambda \mathrm{e}^{-\lambda x}\mathrm{d}x=\frac{1}{\lambda}.$$

由 $E(X^2)=\frac{2}{\lambda^2}$,可以推出

$$D(X)=E(X^2)-(E(X))^2=\frac{1}{\lambda^2}.$$

(3) 正态分布

设 $X\sim N(\mu,\sigma^2)$,则

$$E(X)=\int_{-\infty}^{+\infty}xp(x)\mathrm{d}x=\int_{-\infty}^{+\infty}x\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x=\mu.$$

由方差的定义可以直接算出

$$\begin{aligned}D(X)&=\int_{-\infty}^{+\infty}(x-\mu)^2\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x\\&\xlongequal{令 t=\frac{x-\mu}{\sigma}}\int_{-\infty}^{+\infty}\frac{\sigma^2}{\sqrt{2\pi}}t^2\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t\\&=\frac{\sigma^2}{\sqrt{2\pi}}\left(-t\mathrm{e}^{-t^2}\Big|_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t\right)\end{aligned}$$

$$= 0 + \sigma^2\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt = \sigma^2.$$

**定理 4.3**(二维常见分布的数字特征)

(1) 正态分布

若 $\xi \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$,则

$$E(\xi) = (\mu_1, \mu_2), \quad D(\xi) = (\sigma_1^2, \sigma_2^2).$$

(2) 均匀分布

若 $\xi \sim U(D)$,其中 $D = \{(x,y) \mid a \leqslant x \leqslant b, c \leqslant y \leqslant d\}$,则

$$E(\xi) = \left(\frac{a+b}{2}, \frac{c+d}{2}\right), \quad D(\xi) = \left(\frac{(b-a)^2}{12}, \frac{(d-c)^2}{12}\right).$$

**定理 4.4**(协方差等于零的充要条件)

(1) $\mathrm{cov}(X,Y) = 0 \Leftrightarrow E(XY) = (E(X))(E(Y))$.

(2) $\mathrm{cov}(X,Y) = 0 \Leftrightarrow D(X+Y) = D(X) + D(Y)$.

## 三、典型例题分析

### (一) 填空题

1. 设随机变量 $X_1, X_2, X_3$ 相互独立,其中 $X_1$ 在$[0,6]$上服从均匀分布,$X_2$ 服从正态分布 $N(0,2^2)$,$X_3$ 服从参数为 $\lambda = 3$ 的泊松分布.记 $Y = X_1 - 2X_2 + 3X_3$,则 $D(Y) =$ ________.

答案是:46.

**分析** $D(Y) = D(X_1) + (-2)^2 D(X_2) + 3^2 D(X_3)$

$$= \frac{6^2}{12} + 4 \cdot 2^2 + 9 \cdot 3 = 46.$$

2. 设 $X$ 是一个随机变量,其概率密度为

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

则方差 $DX =$ ________.

答案是:$\frac{1}{6}$.

**分析** $E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$

$$=\int_{-1}^{0}x(1+x)\mathrm{d}x+\int_{0}^{1}x(1-x)\mathrm{d}x=0,$$

$$E(X^2)=\int_{-\infty}^{+\infty}x^2f(x)\mathrm{d}x$$

$$=\int_{-1}^{0}x^2(1+x)\mathrm{d}x+\int_{0}^{1}x^2(1-x)\mathrm{d}x=\frac{1}{6}.$$

所以 $$D(X)=E(X^2)-(EX)^2=\frac{1}{6}.$$

3. 设一次试验成功的概率为 $p$，进行 100 次独立重复试验，当 $p=$ ________ 时，成功次数的标准差的值最大，其最大值为 ________.

答案是：$\frac{1}{2}$；5.

**分析** 若 $X$ 满足二项分布，则 $D(X)=np(1-p)$，

$$\frac{\mathrm{d}D(X)}{\mathrm{d}p}=n(1-p)-np=n(1-2p)=0,\quad p=\frac{1}{2},$$

$$\left.\frac{\mathrm{d}^2D(X)}{\mathrm{d}p^2}\right|_{p=\frac{1}{2}}=-2n<0.$$

故 $p=\frac{1}{2}$是方差最大值点，也是标准差的最大值点，方差最大值为

$$\left.np(1-p)\right|_{p=\frac{1}{2}}=100\times\frac{1}{2}\times\frac{1}{2}=25,$$

从而标准差最大值为 $\sqrt{25}=5$.

4. 设随机变量 $X_{ij}(i,j=1,2,\cdots,n;n\geqslant2)$独立同分布，$E(X_{ij})=2$，则行列式

$$Y=\begin{vmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nn} \end{vmatrix}$$

的数学期望 $E(Y)=$ ________.

答案是：0.

**分析**　由 $n$ 阶行列式的定义，注意到

$$E(X_{11}\cdot X_{22}\cdot\cdots\cdot X_{nn})=E(X_{11})E(X_{22})\cdots E(X_{nn})=2^n,$$

因此 $E(Y)=0$.

5. 设随机变量 $X$ 服从参数为 $\lambda$ 的泊松(Poisson)分布，且已知 $E[(X-1)(X-2)]=1$，则 $\lambda=$________.

答案是：1.

**分析**　$E(X)=\lambda, D(X)=\lambda, E(X^2)=\lambda+\lambda^2$,

$$\begin{aligned}E[(X-1)(X-2)]&=E[X^2-3X+2]\\&=\lambda^2+\lambda-3\lambda+2=1,\end{aligned}$$

解得 $\lambda=1$.

6. 设 $X$ 的均值、方差都存在，且 $D(X)\neq 0$，并且

$$Y=\frac{X-E(X)}{\sqrt{D(X)}},$$

则 $E(X)=$________，$D(Y)=$________.

答案是：0，1.

**分析**

$$\begin{aligned}E(Y)&=E\left(\frac{X-E(X)}{\sqrt{D(X)}}\right)=\frac{1}{\sqrt{D(X)}}E(X-E(X))\\&=\frac{1}{\sqrt{D(X)}}(E(X)-E(X))=0;\\D(Y)&=D\left(\frac{X-E(X)}{\sqrt{D(X)}}\right)=\frac{1}{D(X)}D(X-E(X))\\&=\frac{1}{D(X)}[D(X)+D(-E(X))]\\&=\frac{D(X)}{D(X)}=1.\end{aligned}$$

**(二) 选择题**

1. 设两个相互独立的随机变量 $X$ 和 $Y$ 的方差分别为 4 和 2，则随机变量 $3X-2Y$ 的方差是(　　).

(A) 8；　　(B) 16；　　(C) 28；　　(D) 44.

答案是：D.

**分析**　因为 $X$ 与 $Y$ 是相互独立的，由方差的性质，有

$$D(3X-2Y)=3^2D(X)+2^2D(Y)=9\times4+4\times2=44.$$

2. 设随机变量 $X$ 和 $Y$ 独立同分布，记 $U=X-Y, V=X+Y$，则随机变量 $U$ 与 $V$ 必然(　　).

(A) 不独立；　　(B) 独立；

(C) 相关系数不为零；　　(D) 相关系数为零.

答案是：D.

**分析**　由于

$$\begin{aligned}E(UV)&=E[(X+Y)(X-Y)]=E[X^2-Y^2]\\&=E(X^2)-E(Y^2)=0=E(U)\cdot E(V),\end{aligned}$$

所以 $U$ 与 $V$ 互不相关，(D)必然成立. 当 $X$ 与 $Y$ 为正态随机变量时，$U$ 与 $V$ 也为正态随机变量，$U$ 与 $V$ 独立. 但若取 $P\{X=0\}=P\{Y=0\}=\frac{1}{2}, P\{X=1\}=P\{Y=1\}=\frac{1}{2}$，则由

$$P\{U=-1\}=P\{X=0,Y=1\}=\frac{1}{4},$$

$$P\{V=2\}=P\{X=1,Y=1\}=\frac{1}{4},$$

而 $P\{U=-1,V=2\}=0$ 知 $U$ 与 $V$ 不独立，说明(A)与(B)都不一定成立，故只有(D)必然成立.

3. 已知随机变量 $X$ 服从二项分布，且 $E(X)=2.4, D(X)=1.44$，则二项分布的参数 $n, p$ 的值为(　　).

(A) $n=4, p=0.6$；　　(B) $n=6, p=0.4$；

(C) $n=8, p=0.3$；　　(D) $n=24, p=0.1$.

答案是：B.

**分析**　由 $E(X)=np, D(X)=np(1-p)$ 得方程组

$$np=2.4,\quad np(1-p)=1.44.$$

解方程组即得 $n=6,\ p=0.4$.

4. 设 $X$ 是一随机变量，$E(X)=\mu, D(X)=\sigma^2(\mu,\sigma>0$ 常数），则对任意常数 $C$，必有（　　）.

(A) $E(X-C)^2=EX^2-C^2$； (B) $E(X-C)^2=E(X-\mu)^2$；

(C) $E(X-C)^2<E(X-\mu)^2$； (D) $E(X-C)^2\geqslant E(X-\mu)^2$.

答案是：D.

**分析** 对于任意常数 $C$，

$$\varphi(C)=E(X-C)^2=E(X^2-2CX+C^2)$$
$$=E(X^2)-2\mu C+C^2$$

是 $C$ 的二次函数，易求得函数 $\varphi(C)$ 的最小值为 $\varphi(\mu)$，即 $\varphi(C)\geqslant\varphi(\mu)$，也就是 $E(X-C)^2\geqslant E(X-\mu)^2$.

5. 设随机变量 $X$ 和 $Y$ 的方差存在且不等于 0，则 $D(X+Y)=D(X)+D(Y)$ 是 $X$ 和 $Y$（　　）.

(A) 不相关的充分条件，但不是必要条件；

(B) 独立的充分条件，但不是必要条件；

(C) 不相关的充分必要条件； (D) 独立的充分必要条件.

答案是：C.

**分析** $\sigma_{XY}=0 \Leftrightarrow D(X+Y)=DX+DY$. 因此，选择 C.

（三）解答题

1. 假设随机变量 $X$ 和 $Y$ 在圆域 $x^2+y^2\leqslant r^2$ 上服从联合均匀分布.

(1) 求 $X$ 和 $Y$ 的相关系数 $\rho$； (2) 问 $X$ 和 $Y$ 是否独立?

**解** (1) $X$ 和 $Y$ 的联合密度为

$$p(x,y)=\begin{cases}\dfrac{1}{\pi r^2}, & 若\ x^2+y^2\leqslant r^2,\\ 0, & 其他.\end{cases}$$

$X$ 的密度为

$$p_1(x)=\begin{cases}\dfrac{1}{\pi r^2}\displaystyle\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}}\mathrm{d}y=\dfrac{2}{\pi r^2}\sqrt{r^2-x^2}, & -r\leqslant x\leqslant r,\\ 0, & 其他.\end{cases}$$

$Y$ 的密度为

$$p_2(y)=\begin{cases}\dfrac{2}{\pi r^2}\displaystyle\int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}}\mathrm{d}x=\dfrac{2}{\pi r^2}\sqrt{r^2-y^2}, & -r\leqslant y\leqslant r,\\ 0, & \text{其他}.\end{cases}$$

$$E(X)=\int_{-r}^{r}x\cdot\frac{2}{\pi r^2}\sqrt{r^2-x^2}\mathrm{d}x=0,$$

$$E(Y)=\int_{-r}^{r}y\cdot\frac{2}{\pi}\sqrt{r^2-y^2}\mathrm{d}y=0.$$

$$\mathrm{cov}(X,Y)=E(XY)=\iint\limits_{x^2+y^2\leqslant r^2}xy\cdot\frac{1}{\pi r^2}\mathrm{d}x\mathrm{d}y=0.$$

于是,$X$ 和 $Y$ 的相关系数 $\rho=0$.

(2) 由于 $p(x,y)\neq p_1(x)\cdot p_2(y)$,可见 $X$ 与 $Y$ 不独立.

2. 设排球队 $A$ 与 $B$ 进行比赛,若有一队胜 3 场,则比赛结束.假定 $A$ 在每场比赛中获胜的概率 $p=\dfrac{1}{2}$,试求比赛场数 $X$ 的数学期望.

**解** $X$ 的可能取值为 3,4,5.

若以 3 场结束比赛,或 $A$ 全胜,或 $B$ 全胜,此时概率为 $p^3+q^3=\left(\dfrac{1}{2}\right)^3+\left(\dfrac{1}{2}\right)^3=\dfrac{1}{4}$,即 $P(X=3)=\dfrac{1}{4}$.这里 $q$ 为 $B$ 在每场比赛中获胜的概率,$q=1-p=\dfrac{1}{2}$.

若以 4 场结束比赛,则或 $A$ 在第 4 场取胜,或 $B$ 在第 4 场取胜,故 $A$ 胜的概率为 $pC_3^2p^2q=\dfrac{3}{16}$,同样 $B$ 获胜的概率为 $qC_3^2q^2p=\dfrac{3}{16}$, $P\{X=4\}=\dfrac{3}{16}+\dfrac{3}{16}=\dfrac{3}{8}$.

若以 5 场结束比赛,则或 $A$ 在第 5 场取胜或 $B$ 在第 5 场取胜,

$$P\{X=5\}=pC_4^2p^2q^2+qC_4^2q^2p^2=2\times6\times\left(\frac{1}{2}\right)^5=\frac{3}{8}.$$

故 $X$ 的分布律为:

| $X$ | 3 | 4 | 5 |
|---|---|---|---|
| $p$ | $\frac{1}{4}$ | $\frac{3}{8}$ | $\frac{3}{8}$ |

由此得

$$E(X)=3\times\frac{1}{4}+4\times\frac{3}{8}+5\times\frac{3}{8}=\frac{33}{8}=4.125(\text{场}).$$

3. 一汽车沿一街道行驶，需要通过三个均设有红绿信号灯的路口，每个信号灯为红或绿与其他信号灯为红或绿相互独立，且红绿两种信号显示的时间相等. 以 $X$ 表示该汽车首次遇到红灯前已通过的路口的个数.

(1) 求 $X$ 的概率分布；　　(2) 求 $E\left(\frac{1}{1+X}\right)$.

**解**　(1) $X$ 的可能值为 0,1,2,3. 以 $A_i$ 表示事件“汽车在第 $i$ 个路口首次遇到红灯”，则 $P(A_i)=P(\overline{A}_i)=\frac{1}{2}$，$i=1,2,3$，且 $A_1$，$A_2$，$A_3$ 相互独立.

$$P\{X=0\}=P(A_1)=\frac{1}{2},$$

$$P\{X=1\}=P(\overline{A}_1\cdot A_2)=P(\overline{A}_1)\cdot P(A_2)=\frac{1}{4},$$

$$P\{X=2\}=P(\overline{A}_1\cdot\overline{A}_2\cdot A_3)=P(\overline{A}_1)P(\overline{A}_2)P(A_3)=\frac{1}{8},$$

$$P\{X=3\}=P(\overline{A}_i\cdot\overline{A}_2\cdot\overline{A}_3)=P(\overline{A}_1)P(\overline{A}_2)P(\overline{A}_3)=\frac{1}{8}.$$

(2) $E\left(\frac{1}{1+X}\right)=1\times\frac{1}{2}+\frac{1}{2}\times\frac{1}{4}+\frac{1}{3}\times\frac{1}{8}+\frac{1}{4}\times\frac{1}{8}=\frac{67}{96}$.

4. 对某目标进行射击，直到击中为止. 如果每次命中率为 $p$，求射击次数的数学期望及方差.

**解**　设射击次数为随机变量 $X$，其分布律为：

| $X=k$ | 1 | 2 | 3 | … | $n$ | … |
|---|---|---|---|---|---|---|
| $P\{X=k\}$ | $p$ | $pq$ | $pq^2$ | … | $pq^{n-1}$ | … |

$$E(X)=1\times p+2\times pq+3pq^2+\cdots+npq^{n-1}+\cdots$$

$$=\sum_{k=1}^{\infty}kpq^{k-1}=p\sum_{k=1}^{\infty}kq^{k-1}=p\sum_{k=1}^{\infty}\frac{\mathrm{d}}{\mathrm{d}q}(q^k)$$

$$=p\frac{\mathrm{d}}{\mathrm{d}q}\Big(\sum_{k=1}^{\infty}q^k\Big)=p\frac{\mathrm{d}}{\mathrm{d}q}\Big(\frac{q}{1-q}\Big)$$

$$=p\cdot\frac{1-q+q}{(1-q)^2}=\frac{p}{(1-q)^2}=\frac{1}{p},$$

$$E(X^2)=\sum_{k=1}^{\infty}k^2p(X=k)=\sum_{k=1}^{\infty}k^2pq^{k-1}$$

$$=\sum_{k=1}^{\infty}[k(k-1)+k]pq^{k-1}$$

$$=\sum_{k=2}^{\infty}k(k-1)pq^{k-1}+\sum_{k=1}^{\infty}kpq^{k-1}$$

$$=pq\sum_{k=2}^{\infty}k(k-1)q^{k-2}+E(X)$$

$$=pq\sum_{k=2}^{\infty}\frac{\mathrm{d}^2}{\mathrm{d}q^2}(q^k)+\frac{1}{p}=pq\frac{\mathrm{d}^2}{\mathrm{d}q^2}\Big(\sum_{k=2}^{\infty}q^k\Big)+\frac{1}{p}$$

$$=pq\frac{\mathrm{d}^2}{\mathrm{d}q^2}\Big(\frac{q^2}{1-q}\Big)+\frac{1}{p}=pq\frac{\mathrm{d}}{\mathrm{d}q}\Big(\frac{2q-q^2}{(1-q)^2}\Big)+\frac{1}{p}$$

$$=pq\cdot\frac{2}{(1-q)^3}+\frac{1}{p}=\frac{2q}{p^2}+\frac{1}{p}=\frac{2-p}{p^2},$$

$$D(X)=E(X^2)-[E(X)]^2=\frac{2-p}{p^2}-\frac{1}{p^2}=\frac{1-p}{p^2}.$$

5. 某人用 $n$ 把钥匙去开门，只有一把能打开，今逐个任取一把试开，求打开此门所需开门次数 $X$ 的数学期望及方差.

**解** 打不开门的钥匙不放回的情况下，所需开门次数 $X$ 的可能值为 $1,2,\cdots,n$. 注意到 $X=k$ 意味着第一次到第 $k-1$ 次均未能打开，第 $k$ 次才打开，$P\{X=k\}=\frac{n-1}{n}\cdot\frac{n-2}{n-1}\cdot\cdots\cdot\frac{n-k}{n-k+1}\cdot\frac{1}{n-k}=\frac{1}{n}$，$k=1,2,\cdots,n$. 于是，随机变量 $X$ 的分布律为

| $X$ | 1 | 2 | 3 | … | $n$ |
|---|---|---|---|---|---|
| $p$ | $\frac{1}{n}$ | $\frac{1}{n}$ | $\frac{1}{n}$ | … | $\frac{1}{n}$ |

$$E(X)=\sum_{k=1}^{n}k\cdot\frac{1}{n}=\frac{1}{n}(1+2+\cdots+n)=\frac{n+1}{2},$$

$$\begin{aligned}E(X^2)&=\sum_{k=1}^{n}k^2\cdot\frac{1}{n}=\frac{1}{n}(1^2+2^2+\cdots+n^2)\\&=\frac{1}{n}\cdot\frac{1}{6}n(n+1)(2n+1)=\frac{(n+1)(2n+1)}{6},\end{aligned}$$

$$\begin{aligned}D(X)&=E(X^2)-[E(X)]^2\\&=\frac{(n+1)(2n+1)}{6}-\left(\frac{n+1}{2}\right)^2\\&=\frac{1}{12}(n+1)(n-1).\end{aligned}$$

6. 假设有十只同种电器元件，其中有两只废品. 装配仪器时，从这批元件中任取一只，如是废品，则扔掉重新任取一只；如仍是废品，则扔掉再取一只. 试求在取到正品之前，已取出的废品只数的分布、数学期望和方差.

**解** 用 $X$ 代表在取到正品之前已取出的废品数，$X$ 只可能取三个值：0，1，2.

(1) 分布

$$P\{X=0\}=\frac{8}{10}=\frac{4}{5},\quad P\{X=1\}=\frac{2}{10}\cdot\frac{8}{9}=\frac{8}{45},$$

$$P\{X=2\}=\frac{2}{10}\cdot\frac{1}{9}\cdot\frac{8}{8}=\frac{1}{45}.$$

(2) 数学期望

$$E(X)=0\times\frac{4}{5}+1\times\frac{8}{45}+2\times\frac{1}{45}=\frac{10}{45}=\frac{2}{9}.$$

(3) 方差

$$E(X^2)=0^2\times\frac{4}{5}+1^2\times\frac{8}{45}+2^2\times\frac{1}{45}=\frac{4}{15},$$

$$D(X)=E(X^2)-(E(X))^2=\frac{88}{405}.$$

7. 设 $X$ 服从自由度为 $n$ 的 $\chi^2$ 分布，它的密度函数为

$$p(x)=\begin{cases}\dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}x^{\frac{n}{2}-1}\mathrm{e}^{-\frac{x}{2}}, & x\geqslant 0,\\ 0, & x<0,\end{cases}$$

其中 $n$ 为正整数，求 $E(X)$ 及 $D(X)$.

**解** 由数学期望的定义得：

$$E(X)=\int_{-\infty}^{+\infty}xp(x)\mathrm{d}x=\int_0^{+\infty}\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}x^{\frac{n}{2}}\mathrm{e}^{-x/2}\mathrm{d}x$$

$$\xlongequal{\text{令 } y=\frac{x}{2}}\int_0^{+\infty}\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}(2y)^{\frac{n}{2}}\mathrm{e}^{-y}\cdot 2\mathrm{d}y$$

$$=\frac{2}{\Gamma\left(\frac{n}{2}\right)}\int_0^{+\infty}y^{\frac{n}{2}}\mathrm{e}^{-y}\mathrm{d}y$$

$$=\frac{2}{\Gamma\left(\frac{n}{2}\right)}\Gamma\left(\frac{n}{2}+1\right)=\frac{2}{\Gamma\left(\frac{n}{2}\right)}\cdot\frac{n}{2}\Gamma\left(\frac{n}{2}\right)=n,$$

$$E(X^2)=\int_{-\infty}^{+\infty}x^2p(x)\mathrm{d}x=\int_0^{+\infty}x^2\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}x^{\frac{n}{2}-1}\mathrm{e}^{-\frac{x}{2}}\mathrm{d}x$$

$$\xlongequal{y=\frac{x}{2}}\int_0^{+\infty}\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}(2y)^{\frac{n}{2}+1}\mathrm{e}^{-y}\cdot 2\mathrm{d}y$$

$$=\frac{4}{\Gamma\left(\frac{n}{2}\right)}\int_0^{+\infty}y^{\frac{n}{2}+1}\mathrm{e}^{-y}\mathrm{d}y$$

$$= \frac{4}{\Gamma\left(\frac{n}{2}\right)}\Gamma\left(\frac{n}{2}+2\right) = 4\left(\frac{n}{2}+1\right)\cdot\frac{n}{2}$$

$$= n(n+2),$$

$$D(X) = E(X^2) - (E(X))^2 = n(n+2) - n^2 = 2n.$$

8. 已知随机变量 $X$ 和 $Y$ 的联合密度为

$$p(x,y) = \begin{cases} e^{-(x+y)}, & 若\ 0 < x < +\infty, 0 < y < +\infty, \\ 0, & 其他, \end{cases}$$

试求：(1) $P\{X<Y\}$；　(2) $E(XY)$.

**解**　(1) $P\{X<Y\} = \iint\limits_{x<y} p(x,y)\mathrm{d}x\mathrm{d}y = \int_0^{+\infty}\mathrm{d}y\int_0^y e^{-(x+y)}\mathrm{d}x$

$$= \int_0^{+\infty} e^{-y}(1-e^{-y})\mathrm{d}y = \frac{1}{2}.$$

(2) $E(XY) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyp(x,y)\mathrm{d}x\mathrm{d}y$

$$= \int_0^{+\infty}\int_0^{+\infty} xye^{-(x+y)}\mathrm{d}x\mathrm{d}y$$

$$= \int_0^{+\infty} xe^{-x}\mathrm{d}x\int_0^{+\infty} ye^{-y}\mathrm{d}y = 1.$$

9. 一台设备由三大部件构成，在设备运转中各部件需要调整的概率相应为 0.10，0.20 和 0.30. 假设各部件的状态相互独立，以 $X$ 表示同时需要调整的部件数，试求 $X$ 的概率分布. 数学期望 $E(X)$和方差 $D(X)$.

**解**　设 $A_i$＝{部件 $i$ 需要调整} $(i=1,2,3)$，

$$P(A_1) = 0.10,\quad P(A_2) = 0.20,\quad P(A_3) = 0.30.$$

$X$ 可能取值 0,1,2,3. 由于 $A_1, A_2, A_3$ 相互独立，

$$P\{X=0\} = P(\overline{A}_1\overline{A}_2\overline{A}_3) = 0.9\times0.8\times0.7 = 0.504,$$

$$P\{X=1\} = P(A_1\overline{A}_2\overline{A}_3) + P(\overline{A}_1A_2\overline{A}_3) + P(\overline{A}_1\overline{A}_2A_3)$$

$$= 0.1\times0.8\times0.7 + 0.9\times0.2\times0.7$$

$$+ 0.9\times0.8\times0.3$$

$$= 0.398,$$

$$P\{X=2\}=P(A_1A_2\overline{A}_3)+P(A_1\overline{A}_2A_3)+P(\overline{A}_1A_2A_3)$$
$$=0.1\times0.2\times0.7+0.1\times0.8\times0.3$$
$$+0.9\times0.2\times0.3$$
$$=0.092,$$
$$P(X=3)=P(A_1A_2A_3)=0.1\times0.2\times0.3=0.006.$$

于是

$$X\sim\begin{bmatrix}0 & 1 & 2 & 3\\0.504 & 0.398 & 0.092 & 0.006\end{bmatrix},$$

$$E(X)=1\times0.398+2\times0.092+3\times0.006=0.6,$$
$$D(X)=E(X^2)-(E(X))^2=1\times0.398+4\times0.092$$
$$+9\times0.006-(0.6)^2=0.46.$$

10. 已知离散型随机变量 $X$ 的概率分布为：

$P\{X=1\}=0.2$，$P\{X=2\}=0.3$，$P\{X=3\}=0.5$.

(1) 写出 $X$ 的分布函数 $F(x)$；

(2) 求 $X$ 的数学期望和方差.

**解** (1) $F(x)=P\{X\leqslant x\}=\begin{cases}0, & x<1,\\0.2, & 1\leqslant x<2,\\0.5, & 2\leqslant x<3,\\1, & x\geqslant3,\end{cases}$

或

$$F(x)=P\{X<x\}=\begin{cases}0, & x\leqslant1,\\0.2, & 1<x\leqslant2,\\0.5, & 2<x\leqslant3,\\1, & x>3.\end{cases}$$

(2) $E(X)=1\times0.2+2\times0.3+3\times0.5=2.3$,

$E(X^2)=1\times0.2+4\times0.3+9\times0.5=5.9$,

$D(X)=E(X^2)-(E(X))^2=5.9-5.29=0.61$.

11. 已知随机变量 $Y$ 的概率密度为

$$p(y)=\begin{cases}\dfrac{y}{a^2}e^{-\frac{y^2}{2a^2}}, & y>0,\\0, & y\leqslant0,\end{cases}$$

求随机变量 $Z=\frac{1}{Y}$ 的数学期望 $E(Z)$.

**解** 由数学期望的定义得：

$$E(Z)=E\left(\frac{1}{Y}\right)=\int_{-\infty}^{+\infty}\frac{1}{y}p(y)\mathrm{d}y=\int_{0}^{+\infty}\frac{1}{a^2}\mathrm{e}^{-\frac{y^2}{2a^2}}\mathrm{d}y$$

$$=\frac{1}{2a^2}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{y^2}{2a^2}}\mathrm{d}y=\frac{\sqrt{2\pi}a}{2a^2}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}a}\mathrm{e}^{-\frac{y^2}{2a^2}}\mathrm{d}y=\frac{\sqrt{2\pi}}{2a}.$$

12. 已知随机变量 $X$ 与 $Y$ 的联合概率分布为：

| $(x,y)$ | (0,0) | (0,1) | (1,0) | (1,1) | (2,0) | (2,1) |
|---|---|---|---|---|---|---|
| $P\{X=x,Y=y\}$ | 0.10 | 0.15 | 0.25 | 0.20 | 0.15 | 0.15 |

求 (1) $X+Y$ 的概率分布； (2) $Z=\sin\frac{\pi(X+Y)}{2}$ 的数学期望；(3) $X$ 与 $Y$ 的相关系数 $\rho$.

**解** 将 $X$ 与 $Y$ 的概率分布写成概率分布表形式：

| $X$ \ $Y$ | 0 | 1 |
|---|---|---|
| 0 | 0.10 | 0.15 |
| 1 | 0.25 | 0.20 |
| 2 | 0.15 | 0.15 |

(1) $X+Y$ 的概率分布为：

| $S$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P\{X+Y=S\}$ | 0.10 | 0.40 | 0.35 | 0.15 |

(2) $$E\left[\sin\frac{\pi(X+Y)}{2}\right]=\sin0^\circ\times0.10+\sin\frac{\pi}{2}\times0.40+\sin\pi\times0.35+\sin\frac{3}{2}\pi\times0.15$$
$$=1\times0.40-1\times0.15=0.25.$$

(3)

| $X=k$ | 0 | 1 | 2 |
|---|---|---|---|
| $P\{X=k\}$ | 0.25 | 0.45 | 0.30 |

,

| $Y=l$ | 0 | 1 |
|---|---|---|
| $P\{Y=l\}$ | 0.5 | 0.5 |

,

$$E(X) = 0 \times 0.25 + 1 \times 0.45 + 2 \times 0.30 = 1.05,$$

$$E(X^2) = 0^2 \times 0.25 + 1^2 \times 0.45 + 2^2 \times 0.30 = 1.65,$$

$$D(X) = E(X^2) - (EX)^2 = 0.5475,$$

$$E(Y) = 0.5, D(Y) = 0.5 \times 0.5 = 0.25,$$

$$\begin{aligned} E(XY) = {} & 0 \times 0 \times 0.10 + 0 \times 1 \times 0.15 \\ & + 1 \times 0 \times 0.25 + 1 \times 1 \times 0.2 \\ & + 2 \times 0 \times 0.15 + 2 \times 1 \times 0.15 = 0.5, \end{aligned}$$

$$\begin{aligned} \operatorname{cov}(X,Y) &= E(XY) - E(X)E(Y) \\ &= 0.5 - 1.05 \times 0.5 = -0.025, \end{aligned}$$

$$\rho = \frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)}\ \sqrt{D(Y)}} = \frac{-0.025}{\sqrt{0.5475}\ \sqrt{0.25}} = -0.0676.$$

13. 设 $\xi,\eta$ 是相互独立且服从同一分布的两个随机变量，已知 $\xi$ 的分布律为 $P(\xi=i)=\frac{1}{3}, i=1,2,3$，又设 $X=\max\{\xi,\eta\}, Y=\min(\xi,\eta)$.

(1) 写出二维随机变量 $(X,Y)$ 的分布律：

| $Y$ \ $X$ | 1 | 2 | 3 |
|---|---|---|---|
| 1 | | | |
| 2 | | | |
| 3 | | | |

(2) 求随机变量 $X$ 的数学期望 $E(X)$.

**解** (1) 由于总有 $Y\leqslant X$,故

① $P\{X=i,Y=j\}=0$,当 $i<j$ 时.

② $P\{X=i,Y=i\}=P\{\xi=i,\eta=i\}=P\{\xi=i\}P\{\eta=i\}$

$$=\frac{1}{3}\cdot\frac{1}{3}=\frac{1}{9},\ i=1,2,3.$$

③ $P\{X=i,Y=j\}=P\{\xi=i,\eta=j\}+P\{\xi=j,\eta=i\}$

$$=\frac{1}{9}+\frac{1}{9}=\frac{2}{9},\text{当 } i>j \text{ 时}.$$

| $Y$ \ $X$ | 1 | 2 | 3 |
|---|---|---|---|
| 1 | 1/9 | 2/9 | $\frac{2}{9}$ |
| 2 | 0 | 1/9 | 2/9 |
| 3 | 0 | 0 | 1/9 |

(2) $X$ 的分布律为

| $X$ | 1 | 2 | 3 |
|---|---|---|---|
| $P_i$ | $\frac{1}{9}$ | $\frac{3}{9}$ | $\frac{5}{9}$ |

$$E(X)=1\cdot\frac{1}{9}+2\cdot\frac{3}{9}+3\cdot\frac{5}{9}=\frac{22}{9}.$$

14. 从学校乘汽车到火车站的途中有 3 个交通岗,假设在各个交通岗遇到红灯的事件是相互独立的,并且概率都是 $\frac{2}{5}$. 设 $X$ 为途中遇到红灯的次数,求随机变量 $X$ 的分布律、分布函数和数学期望.

**解** 显然 $X$ 服从二项分布 $B\left(3,\frac{2}{5}\right)$,$X$ 的可能值为 0,1,2,3;其概率分别为

$$P\{X=0\}=C_3^0\left(\frac{2}{5}\right)^0\left(1-\frac{2}{5}\right)^3=\frac{27}{125},$$

$$P\{X=1\}=C_3^1\left(\frac{2}{5}\right)\left(1-\frac{2}{5}\right)^2=\frac{54}{125},$$

$$P\{X=2\}=C_3^2\left(\frac{2}{5}\right)^2\left(1-\frac{2}{5}\right)=\frac{36}{125},$$

$$P\{X=3\}=C_3^3\left(\frac{2}{5}\right)^3\left(1-\frac{2}{5}\right)^0=\frac{8}{125}.$$

即 $X$ 的分布律为

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P$ | $\frac{27}{125}$ | $\frac{54}{125}$ | $\frac{36}{125}$ | $\frac{8}{125}$ |

据上表，可得 $X$ 的分布函数为

$$F(x)=P\{X\leqslant x\}=\begin{cases}0, & x<0,\\ \frac{27}{125}, & 0\leqslant x<1,\\ \frac{81}{125}, & 1\leqslant x<2,\\ \frac{117}{125}, & 2\leqslant x<3,\\ 1, & x\geqslant 3.\end{cases}$$

$X$ 的数学期望为

$$E(X)=0\times\frac{27}{125}+1\times\frac{54}{125}+2\times\frac{36}{125}+3\times\frac{8}{125}=\frac{6}{5}.$$

$\left(或：E(X)=np=3\times\frac{2}{5}=\frac{6}{5}\right)$.

15. 设 $X$ 为只取非负整数的随机变量，证明

$$E(X)=\sum_{k=1}^{\infty}P\{X\geqslant k\}.$$

**证**

$$\sum_{k=1}^{\infty}P\{X\geqslant k\}=\sum_{k=1}^{\infty}\sum_{j=k}^{\infty}P\{X=j\}$$

$$\begin{aligned}
&=P\{X=1\}+P\{X=2\}+P\{X=3\}+P\{X=4\}+\cdots\\
&\qquad\qquad\quad+P\{X=2\}+P\{X=3\}+P\{X=4\}+\cdots\\
&\qquad\qquad\qquad\qquad\qquad+P\{X=3\}+P\{X=4\}+\cdots\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad+P\{X=4\}+\cdots\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\vdots\\
&=1\cdot P\{X=1\}+2P\{X=2\}+3P\{X=3\}+4P\{X=4\}+\cdots\\
&=\sum_{k=0}^{\infty}kP\{X=k\}=E(X).
\end{aligned}$$

16. 设常数 $a$ 与 $b$ 为随机变量 $X$ 的一切可能取值中的最小值与最大值，$D(X)$ 为 $X$ 的方差，则 $D(X)\leqslant\left(\frac{b-a}{2}\right)^2$.

**证** 因为 $a\leqslant X\leqslant b$，所以 $-\frac{b-a}{2}\leqslant X-\frac{a+b}{2}\leqslant\frac{b-a}{2}$，即

$$\left(X-\frac{a+b}{2}\right)^2\leqslant\left(\frac{b-a}{2}\right)^2,$$

于是

$$E\left(X-\frac{a+b}{2}\right)^2\leqslant\left(\frac{b-a}{2}\right)^2.$$

从而

$$\begin{aligned}
D(X)&=E[(X-E(X))^2]\\
&=E\left[\left(X-\frac{a+b}{2}\right)+\left(\frac{a+b}{2}-E(X)\right)\right]^2\\
&=E\left(X-\frac{a+b}{2}\right)^2+2\left(\frac{a+b}{2}-E(X)\right)E\left(X-\frac{a+b}{2}\right)\\
&\quad+\left(\frac{a+b}{2}-E(X)\right)^2\\
&=E\left(X-\frac{a+b}{2}\right)^2-\left(\frac{a+b}{2}-E(X)\right)^2\\
&\leqslant E\left(X-\frac{a+b}{2}\right)^2\leqslant\left(\frac{b-a}{2}\right)^2.
\end{aligned}$$

17. 假设随机变量 $Y$ 服从参数为 $\lambda=1$ 的指数分布，随机变量

$$X_k = \begin{cases} 0, & 若 Y \leqslant k, \\ 1, & 若 Y > k. \end{cases} \quad (k = 1,2)$$

(1) 求 $X_1$ 和 $X_2$ 的联合概率分布；

(2) 求 $E(X_1+X_2)$.

**解** (1) 随机变量 $Y$ 的分布函数为

$$F(y) = \int_{-\infty}^{y} f(y)\mathrm{d}y = \begin{cases} \int_0^y \mathrm{e}^{-y}\mathrm{d}y = 1 - \mathrm{e}^{-y}, & y > 0, \\ 0, & y \leqslant 0. \end{cases}$$

二维随机变量 $(X_1, X_2)$ 的所有可能取值为：(0,0)、(0,1)、(1,0)、(1,1)，且其概率分别为

$$P\{X_1 = 0, X_2 = 0\} = P\{Y \leqslant 1, Y \leqslant 2\}$$
$$= P\{Y \leqslant 1\} = F(1) = 1 - \mathrm{e}^{-1};$$
$$P\{X_1 = 0, X_2 = 1\} = P\{Y \leqslant 1, Y > 2\} = 0;$$
$$P\{X_1 = 1, X_2 = 0\} = P\{Y > 1, Y \leqslant 2\} = P\{1 < Y \leqslant 2\}$$
$$= F(2) - F(1) = \mathrm{e}^{-1} - \mathrm{e}^{-2};$$
$$P\{X_1 = 1, X_2 = 1\} = P\{Y > 1, Y > 2\} = P\{Y > 2\}$$
$$= \int_2^{+\infty} \mathrm{e}^{-y}\mathrm{d}y = \mathrm{e}^{-2}.$$

可得 $(X_1, X_2)$ 的联合分布律为

| $X_2$ \ $p$ \ $X_1$ | 0 | 1 |
|---|---|---|
| 0 | $1-\mathrm{e}^{-1}$ | $\mathrm{e}^{-1}-\mathrm{e}^{-2}$ |
| 1 | 0 | $\mathrm{e}^{-2}$ |

(2) 由于 $X_k(k=1,2)$ 服从(0-1)分布，且

$$P\{X_k = 0\} = P\{Y \leqslant k\} = F(k) = 1 - \mathrm{e}^{-k},$$
$$P\{X_k = 1\} = P\{Y > k\} = 1 - P\{Y \leqslant k\} = \mathrm{e}^{-k}.$$

故得

$$E(X_k) = 0 \cdot (1 - \mathrm{e}^{-k}) + 1 \cdot \mathrm{e}^{-k} = \mathrm{e}^{-k} \quad (k = 1,2).$$

因此

$$E(X_1+X_2)=EX_1+EX_2=\mathrm{e}^{-1}+\mathrm{e}^{-2}.$$

18. 设随机变量 $X$ 和 $Y$ 同分布，$X$ 的概率密度为

$$p(x)=\begin{cases}\dfrac{3}{8}x^2, & 0<x<2,\\ 0, & 其他.\end{cases}$$

(1) 已知事件 $A=\{X>a\}$和 $B=\{Y>a\}$独立，且 $P(A\cup B)=\dfrac{3}{4}$，求常数 $a$；

(2) 求$\dfrac{1}{X^2}$的数学期望.

**解** (1) 由条件知 $P(A)=P(B)$，$P(AB)=P(A)P(B)$，

$$\begin{aligned}P(A\cup B)&=P(A)+P(B)-P(AB)\\&=2P(A)-[P(A)]^2=\frac{3}{4}.\end{aligned}$$

由此得 $P(A)=\dfrac{1}{2}$ 并且知 $0<a<2$. 由于

$$P\{X>a\}=\int_a^{+\infty}p(x)\mathrm{d}x=\int_a^2\frac{3}{8}x^2\mathrm{d}x=\frac{x^3}{8}\Big|_a^2=1-\frac{1}{8}a^3.$$

从而有 $1-\dfrac{1}{8}a^3=\dfrac{1}{2}$，于得得 $a=\sqrt[3]{4}$.

(2) $E\left(\dfrac{1}{X^2}\right)=\displaystyle\int_{-\infty}^{+\infty}\frac{1}{x^2}p(x)\mathrm{d}x=\int_0^2\frac{1}{x^2}\cdot\frac{3}{8}x^2\mathrm{d}x=\frac{3}{8}\int_0^2\mathrm{d}x=\frac{3}{4}$.

19. 设随机变量 $X$ 和 $Y$ 独立，都在区间$[1,3]$上服从均匀分布；引进事件 $A=\{X\leqslant a\}$，$B=\{Y>a\}$.

(1) 已知 $P(A\cup B)=\dfrac{7}{9}$，求常数 $a$；

(2) 求$\dfrac{1}{X}$的数学期望.

**解** (1) 设 $p=P(A)$. 由 $X$ 与 $Y$ 同分布，可知

$$P(\overline{B})=P\{Y\leqslant a\}=P\{X\leqslant a\}=P(A)=p,$$

$$P(B)=1-p.$$

由 $$P(A\cup B)=P(A)+P(B)-P(A)P(B)$$

$$= p + (1-p) - p(1-p)$$

$$= p^2 - p + 1 = \frac{7}{9},$$

得 $p_1 = \frac{1}{3}$，$p_2 = \frac{2}{3}$. 于是 $a$ 有两个值：

由 $\frac{a-1}{2} = p_1$ 得 $a_1 = 1 + \frac{2}{3} = \frac{5}{3}$；

由 $\frac{a-1}{2} = p_2$ 得 $a_2 = 1 + \frac{4}{3} = \frac{7}{3}$.

(2) $E\left(\frac{1}{X}\right) = \int_{-\infty}^{+\infty} \frac{1}{x} p(x) \mathrm{d}x = \frac{1}{2}\int_1^3 \frac{1}{x} \mathrm{d}x = \frac{1}{2}\ln 3$.

20. 设二维随机变量$(X,Y)$在区域 $D$(见图 4.1)：$0<x<1$，$|y|<x$ 内服从均匀分布，求关于 $X$ 的边缘概率密度函数及随机变量 $Z=2X+1$ 的方差 $D(Z)$.

**解** $(X,Y)$的联合概率密度函数是

$$p(x,y) = \begin{cases} 1, & 0 < x < 1, |y| < x, \\ 0, & \text{其他}, \end{cases}$$

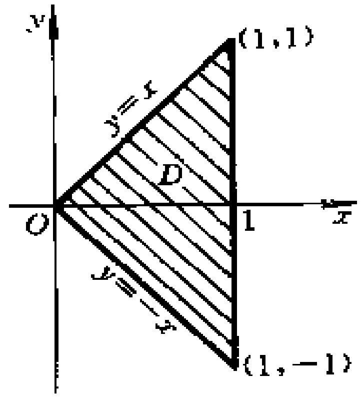



图 4.1

$X$ 的边缘概率密度是

$$p_X(x) = \int_{-\infty}^{+\infty} p(x,y)\mathrm{d}y = \begin{cases} 2x, & 0<x<1, \\ 0, & \text{其他}. \end{cases}$$

$$E(X) = \int_{-\infty}^{+\infty} x p_X(x) \mathrm{d}x = \int_0^1 x \cdot 2x \mathrm{d}x = \frac{2}{3}x^3 \Big|_0^1 = \frac{2}{3},$$

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 p(x) \mathrm{d}x = \int_0^1 x^2 \cdot 2x \mathrm{d}x = \frac{1}{2}x^4 \Big|_0^1 = \frac{1}{2},$$

$$D(X) = E(X^2) - [E(X)]^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18},$$

$$D(Z) = 2^2 \cdot D(X) = \frac{2}{9}.$$

21. 设随机变量 $X$ 与 $Y$ 独立，$X$ 服从正态分布 $N(\mu,\sigma^2)$，$Y$ 服从$[-\pi,\pi]$上的均匀分布，求 $Z=X+Y$ 的概率分布密度（计算结果用标准正态分布函数 $\Phi$ 表示，其中 $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt$）.

**解** $X$ 和 $Y$ 的概率密度分别为

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

$$p_Y(y)=\begin{cases}\frac{1}{2\pi}, & -\pi\leqslant x\leqslant\pi,\\ 0, & \text{其他}.\end{cases}$$

由于 $X$ 与 $Y$ 独立，可用卷积公式求 $Z=X+Y$ 的概率密度，注意到 $p_Y(y)$ 仅在$[-\pi,\pi]$上才取非零值，所以 $Z$ 的概率密度函数为

$$p_Z(z)=\int_{-\infty}^{+\infty}p_X(z-y)p_Y(y)dy=\int_{-\pi}^{\pi}p_X(z-y)p_Y(y)dy$$

$$=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(z-y-\mu)^2}{2\sigma^2}}dy.$$

令 $t=\frac{z-y-\mu}{\sigma}$，则有

$$p_Z(z)=\frac{1}{2\pi}\int_{\frac{z-\pi-\mu}{\sigma}}^{\frac{z+\pi-\mu}{\sigma}}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

$$=\frac{1}{2\pi}\left[\Phi\left(\frac{z+\pi-\mu}{\sigma}\right)-\Phi\left(\frac{z-\pi-\mu}{\sigma}\right)\right].$$

22. 设随机变量 $x$ 的概率分布密度为

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

(1) 求 $X$ 的数学期望 $E(X)$和方差 $D(X)$；

(2) 求 $X$ 与$|X|$的协方差，并问 $X$ 与$|X|$是否不相关？

(3) 问 $X$ 与$|X|$是否相互独立？为什么？

**解** (1) $E(X)=\int_{-\infty}^{+\infty}xp(x)dx=0$,

$$D(X)=\int_{-\infty}^{+\infty}x^2p(x)dx=\int_{0}^{+\infty}x^2e^{-x}dx=2.$$

(2) $\mathrm{cov}(X,|X|)=E(X|X|)-E(X)\cdot E(|X|)=E(X|X|)$

$$=\int_{-\infty}^{+\infty}x|x|\cdot p(x)\mathrm{d}x=0,$$

所以 $X$ 与 $|X|$ 互不相关.

(3) 对于任意给定的 $0<x_0<+\infty$，事件 $\{|X|<x_0\}$ 包含在事件 $\{X<x_0\}$ 内，故有

$$0<P\{|X|<x_0\}\leqslant P\{X<x_0\}<1,$$

从而

$$P\{X<x_0,|X|<x_0\}$$
$$=P\{|X|<x_0\}>P\{|X|<x_0\}\cdot P\{X<x_0\}.$$

因此，$X$ 与 $|X|$ 不独立.

23. 已知随机变量 $(X,Y)$ 服从二维正态分布，并且 $X$ 和 $Y$ 分别服从正态分布 $N(1,3^2)$ 和 $N(0,4^2)$，$X$ 与 $Y$ 的相关系数 $\rho_{XY}=-\frac{1}{2}$. 设 $Z=\frac{X}{3}+\frac{Y}{2}$，

(1) 求 $Z$ 的数学期望 $E(Z)$ 和方差 $D(Z)$；

(2) 求 $X$ 与 $Z$ 的相关系数 $\rho_{XZ}$；

(3) 问 $X$ 与 $Z$ 是否相互独立？为什么？

**解** (1) $E(Z)=\frac{1}{3}E(X)+\frac{1}{2}E(Y)=\frac{1}{3}$. 注意由 $D(X)=9$，$D(Y)=16$，$\mathrm{cov}(X,Y)=\rho_{XY}\sqrt{D(X)}\sqrt{D(Y)}=(-1/2)\times3\times4=-6$，有

$$D(Z)=\left(\frac{1}{3}\right)^2\cdot D(X)+\left(\frac{1}{2}\right)^2\cdot D(Y)+2\cdot\frac{1}{3}\cdot\frac{1}{2}\mathrm{cov}(X,Y)$$
$$=\frac{1}{9}D(X)+\frac{1}{4}D(Y)+\frac{1}{3}\mathrm{cov}(X,Y)=1+4-2=3.$$

(2) $\mathrm{cov}(X,Z)=\mathrm{cov}\left(X,\frac{X}{3}\right)+\mathrm{cov}\left(X,\frac{Y}{2}\right)$

$$=\frac{1}{3}\mathrm{cov}(X,X)+\frac{1}{2}\mathrm{cov}(X,Y).$$

注意 $\mathrm{cov}(X,X)=D(X)=9$，$\mathrm{cov}(X,Y)=-6$，有

$$\operatorname{cov}(X,Z)=\frac{1}{3}\cdot 9+\frac{1}{2}(-6)=3-3=0,$$

所以 $$\rho_{XZ}=\frac{\operatorname{cov}(X,Z)}{\sqrt{D(X)}\cdot\sqrt{D(Z)}}=0.$$

(3) 因为 $Z$ 是正态随机变量 $X$ 与 $Y$ 的线性组合，故 $Z$ 也是正态随机变量，又因为 $\rho_{XZ}=0$，所以 $X$ 与 $Z$ 相互独立.

24. 假设由自动线加工的某种零件的内径 $X$(毫米)服从正态分布 $N(\mu,1)$，内径小于 10 或大于 12 为不合格品，其余为合格品. 销售每件合格品获利，销售每件不合格品亏损，已知销售利润 $T$(单位：元)与销售零件的内径 $X$ 有如下关系：

$$T=\begin{cases}-1, & 若 X<10, \\ 20, & 若 10\leqslant X\leqslant 12, \\ -5, & 若 X>12.\end{cases}$$

问平均内径 $\mu$ 取何值时，销售一个零件的平均利润最大？

**解** 平均利润

$$\begin{aligned}E(T)&=20P\{10\leqslant X\leqslant 12\}-P\{X<10\}-5P\{X>12\}\\&=20[\Phi(12-\mu)-\Phi(10-\mu)]\\&\quad-\Phi(10-\mu)-5[1-\Phi(12-\mu)]\\&=25\Phi(12-\mu)-21\Phi(10-\mu)-5,\end{aligned}$$

$$\frac{\mathrm{d}}{\mathrm{d}\mu}E(T)=-25p(12-\mu)+21p(10-\mu).$$

(其中 $\Phi(x)$和 $p(x)$分别为标准正态分布函数和标准正态密度函数)令上式为 0 得

$$\frac{-25}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(12-\mu)^2}{2}}+\frac{21}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(10-\mu)^2}{12}}=0,$$

即 $$25\mathrm{e}^{-\frac{(12-\mu)^2}{2}}=21\mathrm{e}^{-\frac{(10-\mu)^2}{2}}.$$

解此方程得

$$\mu=11-\frac{1}{2}\ln\frac{25}{21}\approx 10.9.$$

由此知当 $\mu=10.9$ 毫米时，平均利润最大.

25. 假设一部机器在一天内发生故障的概率为 0.2，机器发生故障时全天停止工作. 若一周 5 个工作日里无故障，可获利润 10 万元；发生一次故障仍可获利润 5 万元；发生二次故障所获利润 0 元；发生三次或三次以上故障就要亏损 2 万元，求一周内期望利润是多少？

**解** 以 $X$ 表示一周 5 天内机器发生故障天数，则 $X$ 服从参数为(5,0.2)的二项分布：

$$P\{X=k\}=C_5^k 0.2^k 0.8^{5-k} \quad (k=0,1,2,3,4,5),$$

$$P\{X=0\}=0.8^5=0.32768,$$

$$P\{X=1\}=C_5^1(0.2)(0.8)^4=0.4096,$$

$$P\{X=2\}=C_5^2(0.2)^2(0.8)^3=0.2048,$$

$$P\{X\geqslant 3\}=1-P\{X=0\}-P\{X=1\}-P\{X=2\} = 0.05792.$$

以 $Y$ 表示所获利润，则

$$Y=g(X)=\begin{cases}10, & \text{若 } X=0,\\ 5, & \text{若 } X=1,\\ 0, & \text{若 } X=2,\\ -2, & \text{若 } X\geqslant 3.\end{cases}$$

$$\begin{aligned}E(Y)&=10\times 0.32768+5\times 0.4096\\&\quad+0\times 0.2048-2\times 0.05792\\&=5.216(\text{万元}).\end{aligned}$$

26. 一商店经销某种商品，每周进货的数量 $X$ 与顾客对该种商品的需求量 $Y$ 是相互独立的随机变量，且都服从区间[10,20]上的均匀分布，商店每售出一单位商品可得利润 1000 元；若需求量超过了进货量，商品可从其他商店调剂供应，这时每单位商品获利润为 500 元. 试求此商店经销该种商品每周所得利润的期望值.

**解** 设 $Z$ 表示商店每周所得的利润，则

$$Z=\begin{cases}1000Y, & Y\leqslant X,\\ 1000X+500(Y-X)=500(X+Y), & Y>X.\end{cases}$$

由于 $X$ 与 $Y$ 的联合概率密度为：

$$p(x,y)=\begin{cases}\dfrac{1}{100}, & 10\leqslant x\leqslant 20,10\leqslant y\leqslant 20,\\ 0, & 其他,\end{cases}$$

所以

$$\begin{aligned}E(Z)&=\iint\limits_{D_1}1000y\times\frac{1}{100}\mathrm{d}x\mathrm{d}y+\iint\limits_{D_2}500(x+y)\times\frac{1}{100}\mathrm{d}x\mathrm{d}y\\&=10\int_{10}^{20}\mathrm{d}y\int_{y}^{20}y\mathrm{d}x+5\int_{10}^{20}\mathrm{d}y\int_{10}^{y}(x+y)\mathrm{d}x\\&=10\int_{10}^{20}y(20-y)\mathrm{d}y+5\int_{10}^{20}\left(\frac{3}{2}y^2-10y-50\right)\mathrm{d}y\\&=\frac{20000}{3}+5\times1500\approx14166.67(元),\end{aligned}$$

这里二重积分中的积分区域 $D_1,D_2$ 如图 4.2 所示.

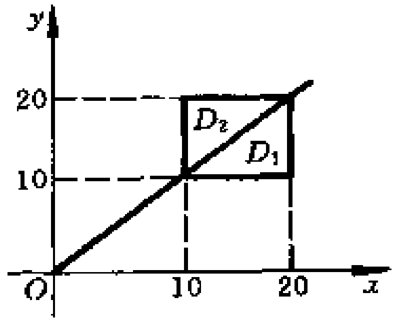



图 4.2

27. 设某种商品每周的需求量 $X$ 是服从区间[10,30]上均匀分布的随机变量，而经销商店进货数量为区间[10,30]中的某一整数，商店每销售一单位商品可获利 500 元；若供大于求则削价处理，每处理 1 单位商品亏损 100 元；若供不应求，则可从外部调剂供应，此时每 1 单位商品仅获利 300 元，为使商店所获利润期望值不少于 9280 元，试确定最少进货量.

**解** 设进货数量为 $a$，则利润为

$$M(a)=\begin{cases}500a+(X-a)300, & a<X\leqslant 30,\\ 500X-(a-X)100, & 10\leqslant X\leqslant a\end{cases}$$

$$= \begin{cases} 300X + 200a, & a < X \leqslant 30, \\ 600X - 100a, & 10 \leqslant X \leqslant a, \end{cases}$$

期望利润

$$E(M(a)) = \int_{10}^{30} \frac{1}{20} \cdot M(a)\mathrm{d}x$$
$$= \frac{1}{20}\int_{10}^{a} (600x - 100a)\mathrm{d}x + \frac{1}{20}\int_{a}^{30} (300x + 200a)\mathrm{d}x$$
$$= \frac{1}{20}\left(600 \cdot \frac{x^2}{2} - 100ax\right)\bigg|_{10}^{a} + \frac{1}{20}\left(300 \cdot \frac{x^2}{2} + 200ax\right)\bigg|_{a}^{30}$$
$$= -7.5a^2 + 350a + 5250.$$

依题意有$-7.5a^2+350a+5250 \geqslant 9280$，即 $7.5a^2-350a+4030 \leqslant 0$. 解得

$$20\frac{2}{3} \leqslant a \leqslant 26.$$

故利润期望值不少于 9280 元的最少进货量为 21 单位.

28. 某箱装有 100 件产品，其中一、二和三等品分别为 80、10 和 10 件，现在从中随机抽取一件，记

$$X_i = \begin{cases} 1, & \text{若抽到 } i \text{ 等品}, \\ 0, & \text{其他}. \end{cases} \quad (i = 1,2,3)$$

试求 (1) 随机变量 $X_1$ 与 $X_2$ 的联合分布；(2) 随机变量 $X_1$ 与 $X_2$ 的相关系数 $\rho$.

**解** (1) 设事件 $A_i$=“抽到 $i$ 等品”($i=1,2,3$)，由题意知 $A_1$，$A_2$，$A_3$ 两两互不相容，且

$$P(A_1) = 0.8, \quad P(A_2) = P(A_3) = 0.1.$$

易见

$$P\{X_1 = 0, X_2 = 0\} = P(A_3) = 0.1,$$
$$P\{X_1 = 0, X_2 = 1\} = P(A_2) = 0.1,$$
$$P\{X_1 = 1, X_2 = 0\} = P(A_1) = 0.8,$$
$$P\{X_1 = 1, X_2 = 1\} = P(\varnothing) = 0;$$

(2) $E(X_1)=0.8$, $E(X_2)=0.1$,

$$D(X_1)=0.8\times 0.2=0.16,\quad D(X_2)=0.1\times 0.9=0.09,$$

$$\begin{aligned}E(X_1X_2)&=0\times 0\times 0.1+0\times 1\times 0.1+1\times 0\times 0.8\\&\quad+1\times 1\times 0=0,\end{aligned}$$

$$\begin{aligned}\operatorname{cov}(X_1,X_2)&=E(X_1X_2)-E(X_1)\cdot E(X_2)\\&=0-0.8\times 0.1=-0.08,\end{aligned}$$

$$\rho=\frac{\operatorname{cov}(X_1,X_2)}{\sqrt{D(X_1)\cdot D(X_2)}}=\frac{-0.08}{\sqrt{0.16\times 0.09}}=-\frac{2}{3}.$$

29. 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 5 分钟、25 分钟和 55 分钟从底层起行. 假设一游客在早八点的第 $X$ 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.

**解** 已知 $X$ 在[0,60]上服从均匀分布，其密度为

$$X\sim p(x)=\begin{cases}\dfrac{1}{60}, & \text{若 } 0\leqslant x\leqslant 60,\\ 0, & \text{其他}.\end{cases}$$

设 $Y$ 是游客等候电梯的时间（单位：分），则

$$Y=g(X)=\begin{cases}5-X, & \text{若 } 0<X\leqslant 5,\\ 25-X, & \text{若 } 5<X\leqslant 25,\\ 55-X, & \text{若 } 25<X\leqslant 55,\\ 60-X+5, & \text{若 } 55<X\leqslant 60.\end{cases}$$

因此

$$\begin{aligned}E(Y)=E(g(X))&=\int_{-\infty}^{+\infty}g(x)\cdot p(x)\mathrm{d}x=\frac{1}{60}\int_0^{60}g(x)\mathrm{d}x\\&=\frac{1}{60}\Big[\int_0^5(5-x)\mathrm{d}x+\int_5^{25}(25-x)\mathrm{d}x\\&\quad+\int_{25}^{55}(55-x)\mathrm{d}x+\int_{55}^{60}(65-x)\mathrm{d}x\Big]\\&=\frac{1}{60}[12.5+200+450+37.5]=11.67.\end{aligned}$$

30. 两台同样自动记录仪，每台无故障工作的时间服从参数为 5 的指数分布；首先开动其中一台，当其发生故障时停用而另一台自行开动. 试求两台记录仪无故障工作的总时间 $T$ 的概率密度 $p(t)$、数学期望和方差.

**解** 以 $X_1$ 和 $X_2$ 表示先后开动的记录仪无故障工作的时间，则 $T=X_1+X_2$. 由条件知 $X_i(i=1,2)$的概率密度为

$$p_i(x)=\begin{cases}5\mathrm{e}^{-5x}, & 若\ x>0,\\ 0, & 若\ x\leqslant 0.\end{cases}$$

两台仪器无故障工作时间 $X_1$ 和 $X_2$ 显然相互独立.

利用两个独立随机变量和的密度公式求 $T$ 的概率密度. 对于 $t>0$，有

$$\begin{aligned}p(t)&=\int_{-\infty}^{+\infty}p_1(x)p_2(t-x)\mathrm{d}x=25\int_0^t\mathrm{e}^{-5x}\mathrm{e}^{-5(t-x)}\mathrm{d}x\\&=25\mathrm{e}^{-5t}\int_0^t\mathrm{d}x=25t\mathrm{e}^{-5t}.\end{aligned}$$

当 $t\leqslant 0$ 时，显然 $p(t)=0$. 于是，得

$$p(t)=\begin{cases}25t\mathrm{e}^{-5t}, & 若\ t>0,\\ 0, & 若\ t\leqslant 0.\end{cases}$$

由于 $X_i$ 服从参数为 $\lambda=5$ 的指数分布，故

$$E(X_i)=\frac{1}{5};\quad D(X_i)=\frac{1}{25}\quad(i=1,2).$$

因此，有

$$E(T)=E(X_1+X_2)=E(X_1)+E(X_2)=\frac{2}{5}.$$

又由于 $X_1$ 和 $X_2$ 相互独立，可见

$$D(T)=D(X_1+X_2)=D(X_1)+D(X_2)=\frac{2}{25}.$$

31. 设两个随机变量 $X$,$Y$ 相互独立，且都服从均值为 0、方差为$\frac{1}{2}$的正态分布，求随机变量$|X-Y|$的方差.

**解** 令 $Z=X-Y$. 由于

$$X \sim N\left(0, \left(\frac{1}{\sqrt{2}}\right)^2\right), \quad Y \sim N\left(0, \left(\frac{1}{\sqrt{2}}\right)^2\right),$$

且 $X$ 和 $Y$ 相互独立,故 $Z \sim N(0,1)$. 因为

$$D(|X-Y|) = D(|Z|) = E(|Z|^2) - [E(|Z|)]^2$$
$$= E(Z^2) - [E(|Z|)]^2,$$

而

$$E(Z^2) = D(Z) = 1,$$

$$E(|Z|) = \int_{-\infty}^{+\infty} |z| \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} z e^{-\frac{z^2}{2}} dz = \sqrt{\frac{2}{\pi}},$$

所以 $$D(|X-Y|) = 1 - \frac{2}{\pi}.$$

32. 设二维离散随机变量$(X,Y)$的联合分布律为

| $X \backslash Y$ | $-1$ | $1$ |
|---|---|---|
| $-1$ | $\frac{1}{6}$ | $\frac{1}{4}$ |
| $1$ | $\frac{1}{4}$ | $\frac{1}{3}$ |

求 $E(X^3+Y^3)$.

**解** $E(X^3+Y^3) = [(-1)^3+(-1)^3] \times \frac{1}{6} + [(-1)^3+1^3] \times \frac{1}{4}$
$+ [1^3+(-1)^3] \times \frac{1}{4} + [1^3+1^3] \times \frac{1}{3} = \frac{1}{3}.$

33. 设随机变量 $X$ 与 $Y$ 的联合分布律如下表所示:

| $X \backslash Y$ | $-1$ | $0$ | $1$ |
|---|---|---|---|
| $-1$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{8}$ |
| $0$ | $\frac{1}{8}$ | $0$ | $\frac{1}{8}$ |
| $1$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{8}$ |

试证明 $X$ 与 $Y$ 不相关,但不相互独立.

**证** $X$ 与 $Y$ 的边缘分布律分别为

$$X \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{3}{8} & \frac{2}{8} & \frac{3}{8} \end{pmatrix}, \quad Y \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{3}{8} & \frac{2}{8} & \frac{3}{8} \end{pmatrix},$$

$$E(X) = -1 \times \frac{3}{8} + 0 \times \frac{2}{8} + 1 \times \frac{3}{8} = 0,$$

$$\begin{aligned} E(XY) = &(-1)\cdot(-1)\cdot\frac{1}{8} + (-1)\cdot 0 \cdot \frac{1}{8} \\ &+ (-1)\cdot 1 \cdot \frac{1}{8} + 0 \cdot (-1) \cdot \frac{1}{8} \\ &+ 0 \cdot 1 \cdot \frac{1}{8} + 1 \cdot (-1) \cdot \frac{1}{8} \\ &+ 1 \cdot 0 \cdot \frac{1}{8} + 1 \cdot 1 \cdot \frac{1}{8} = 0, \end{aligned}$$

$$\operatorname{cov}(X,Y) = E(XY) - E(X)E(Y) = 0,$$

故 $X$ 与 $Y$ 不相关. 但

$$P\{X=i, Y=j\} \neq P\{X=i\}P\{Y=j\}, \quad i,j = -1,0,1.$$

所以 $X$ 与 $Y$ 不相互独立.

34. 对球的直径作近似测量,设其值均匀分布于区间$[a,b]$内,求球的体积的期望值.

**解** 设球的直径 $X$,则 $X$ 在$[a,b]$内均匀分布,故其密度函数

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

又设球的体积为 $Y$,则 $Y = \frac{1}{6}\pi x^3$. 故

$$\begin{aligned} E(Y) &= \int_a^b \frac{1}{b-a} \cdot \frac{1}{6}\pi x^3 \mathrm{d}x = \frac{\pi(b^4 - a^4)}{24(b-a)} \\ &= \frac{\pi}{24}(b+a)(b^2+a^2). \end{aligned}$$

35. 设随机变量$(X,Y)$服从二维正态分布,其密度函数为

$$p(x,y)=\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)},$$

求随机变量 $Z=\sqrt{X^2+Y^2}$ 的数学期望及方差.

**解** $$E(Z)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\sqrt{x^2+y^2}e^{-\frac{x^2+y^2}{2}}dxdy$$

$$=\frac{1}{2\pi}\int_0^{2\pi}d\theta\int_0^{+\infty}r^2e^{-\frac{r^2}{2}}dr$$

$$=\frac{1}{2\pi}\cdot 2\pi\left[-re^{-\frac{r^2}{2}}\Big|_0^{+\infty}+\int_0^{+\infty}e^{-\frac{r^2}{2}}dr\right]$$

$$=\int_0^{+\infty}e^{-\frac{r^2}{2}}dr=\sqrt{\frac{\pi}{2}},$$

$$E(Z^2)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x^2+y^2)e^{-\frac{x^2+y^2}{2}}dxdy$$

$$=\frac{1}{2\pi}\int_0^{2\pi}d\theta\int_0^{+\infty}r^3e^{-\frac{r^2}{2}}dr$$

$$=\frac{1}{2\pi}\cdot 2\pi\left[-r^2e^{-\frac{r^2}{2}}\Big|_0^{+\infty}+2\int_0^{+\infty}re^{-\frac{r^2}{2}}dr\right]$$

$$=2\left[-e^{-\frac{r^2}{2}}\Big|_0^{+\infty}\right]=2.$$

因此 $$D(Z)=E(Z^2)-(E(Z))^2=2-\frac{\pi}{2}.$$

36. 设 $\xi(X,Y)$ 的联合密度函数为

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

(1) 判别 $X,Y$ 是否相互独立,是否相关;

(2) 求 $E(\xi),D(\xi),D(X+Y)$.

**解** (1) 首先求出 $X$ 的边缘分布密度

$$p_1(x)=\int_{-\infty}^{+\infty}(x,y)dy.$$

当 $x<0$ 或 $x>1$ 时,$p_1(x)=0$;当 $0\leqslant x\leqslant 1$ 时,

$$p_1(x)=\int_0^1(2-x-y)\mathrm{d}y=\frac{3}{2}-x.$$

因此
$$p_1(x)=\begin{cases}\frac{3}{2}-x, & 0\leqslant x\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

同理,可以求出 $Y$ 的边缘分布密度

$$p_2(y)=\begin{cases}\frac{3}{2}-y, & 0\leqslant y\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

由于

$$p_1(x)p_2(y)=\begin{cases}\left(\frac{3}{2}-x\right)\left(\frac{3}{2}-y\right), & 0\leqslant x\leqslant 1,0\leqslant y\leqslant 1,\\ 0, & \text{其他},\end{cases}$$
$$\neq p(x,y),$$

所以 $X,Y$ 不相互独立.

$$E(X)=\int_{-\infty}^{+\infty}xp_1(x)\mathrm{d}x=\int_0^1x\left(\frac{3}{2}-x\right)\mathrm{d}x=\frac{5}{12},$$

$$E(Y)=\int_{-\infty}^{+\infty}yp_2(y)\mathrm{d}x=\int_0^1y\left(\frac{3}{2}-y\right)\mathrm{d}y=\frac{5}{12},$$

$$E(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyp(x,y)\mathrm{d}x\mathrm{d}y$$
$$=\int_0^1\mathrm{d}x\int_0^1xy(2-x-y)\mathrm{d}y=\frac{1}{6},$$

$$D(X)=E(X^2)-[E(X)]^2$$
$$=\int_0^1x^2\left(\frac{3}{2}-x\right)\mathrm{d}x-\left(\frac{5}{12}\right)^2=\frac{11}{144}.$$

同理 $D(Y)=\frac{11}{144}$. 由于

$$\rho_{XY}=\frac{E(XY)-E(X)E(Y)}{\sqrt{D(X)}/\sqrt{D(Y)}}=\frac{\frac{1}{6}-\left(\frac{5}{12}\right)^2}{\frac{11}{144}}=-\frac{1}{11}\neq 0,$$

所以 $X$ 与 $Y$ 相关.

(2) $E(\xi)=(E(X),E(Y))=\left(\frac{5}{12},\frac{5}{12}\right)$,

$$D(\xi)=(D(X),D(Y))=\left(\frac{11}{144},\frac{11}{144}\right),$$

$$\begin{aligned}D(X+Y)&=D(X)+D(Y)+2\sigma_{XY}\\&=D(X)+D(Y)+2[E(XY)-E(X)E(Y)]\\&=\frac{11}{144}+\frac{11}{144}+2\left(\frac{1}{6}-\frac{5}{12}\times\frac{5}{12}\right)=\frac{5}{36}.\end{aligned}$$

37. 设随机变量 $X\sim\begin{pmatrix}1 & 0\\ p & q\end{pmatrix}$,当 $p$ 为何值时,$D(X)$取得最大值?

**解** 由 $E(X)=p,E(X^2)=p$,有 $D(X)=p-p^2$. 由$(D(X))'_p=1-2p$,令$(D(X))'_p=0$ 得到 $p=\frac{1}{2}$. 又由于$(D(X))''_p\Big|_{p=\frac{1}{2}}=-2<0$,因此,当 $p=\frac{1}{2}$时,$D(X)$取得最大值.

38. 设随机变量 $X\sim N(1,2),Y\sim N(2,4)$,且 $X$ 与 $Y$ 相互独立,求 $Z=2X+Y-3$ 的分布密度函数 $p(z)$.

**解** 由题意 $E(Z)=2E(X)+E(Y)-3=1$, $D(Z)=4D(X)+D(Y)=12$. 因为 $Z$ 为 $X$ 与 $Y$ 的线性组合,且 $X,Y$ 相互独立,所以 $Z\sim N(\mu,\sigma^2)$,其中 $\mu=1,\sigma^2=12$. 因此

$$p(z)=\frac{1}{\sqrt{2\pi}\sqrt{12}}e^{-\frac{(z-1)^2}{2\times 12}}=\frac{1}{2\sqrt{6\pi}}e^{-\frac{(z-1)^2}{24}}$$
$$(-\infty<z<+\infty).$$

39. 设 $X\sim p(x)=\frac{1}{\pi}\frac{1}{1+x^2}$ (柯西分布),求 $E(X),D(X)$.

**解** 由于

$$\begin{aligned}\int_{-A}^{A}|x|\frac{1}{\pi}\frac{1}{1+x^2}dx&=\frac{2}{2\pi}\int_0^A\frac{d(1+x^2)}{1+x^2}\\&=\frac{1}{\pi}\ln(1+x^2)\Big|_0^A=\frac{1}{\pi}\ln(1+A^2),\end{aligned}$$

而 $$\int_{-\infty}^{+\infty}|x|\frac{1}{\pi}\frac{1}{1+x^2}dx=\lim_{A\to+\infty}\frac{1}{\pi}\ln(1+A^2)=+\infty,$$

由数学期望定义，可知 $E(X)$ 不存在，因而 $D(X)$ 也不存在.

40. 假定在国际市场上每年对我国某种出口商品的需求量 $X \sim U(2000,4000)$. 设每售出此商品 1 t，可为国家挣得外汇 3 万元；但是若销售不出而囤积在仓库中，则每吨需花保养费 1 万元. 问需要组织多少货源，才能使国家的收益最大？

**解** 设准备出口商品量为 $x_0$，收益为 $Y$，则

$$Y=\begin{cases}3x_0, & X\geqslant x_0,\\ 3X-(x_0-X), & X<x_0,\end{cases}$$

而

$$p_1(x)=\begin{cases}\dfrac{1}{2000}, & x\in[2000,4000],\\ 0, & \text{其他},\end{cases}$$

于是

$$\begin{aligned}E(Y)&=\int_{-\infty}^{+\infty}f(x)p_1(x)\mathrm{d}x\\ &=\int_{-\infty}^{2000}f(x)\cdot 0\mathrm{d}x+\int_{2000}^{x_0}(4x-x_0)\frac{1}{2000}\mathrm{d}x\\ &\quad+\int_{x_0}^{4000}3x_0\frac{1}{2000}\mathrm{d}x+\int_{4000}^{+\infty}f(x)\cdot 0\mathrm{d}x\\ &=\frac{1}{1000}(-x_0^2+7000x_0-4\times 10^6).\end{aligned}$$

令 $[E(Y)]'_{x_0}=-2x_0+7000=0$，有 $x_0=3500$，并且考虑到

$$[E(Y)]''_{x_0}\Big|_{x_0=3500}=-2<0,$$

因此，当组织 3500 t 货源时，才能使收益最大.

41. 设随机变量 $X$ 服从 $[0,2\pi]$ 上的均匀分布：

$$p_X(x)=\begin{cases}\dfrac{1}{2\pi}, & 0\leqslant x\leqslant 2\pi,\\ 0, & \text{其他},\end{cases}$$

又 $Y=\sin X$，$Z=\sin(X+a)$，其中 $a\in[0,2\pi]$ 为常数，试求 $\rho_{YZ}$，并

讨论 $Y$ 与 $Z$ 的相关性及独立性.

**解** 根据题意有：

$$E(Y)=E[\sin X]=\int_0^{2\pi}\sin x\cdot\frac{1}{2\pi}\mathrm{d}x=-\frac{1}{2\pi}\cos x\Big|_0^{2\pi}=0,$$

$$E(Z)=E[\sin(X+a)]=\int_0^{2\pi}\sin(x+a)\cdot\frac{1}{2\pi}\mathrm{d}x$$

$$=-\frac{1}{2\pi}\cos(x+a)\Big|_0^{2\pi}=0,$$

$$E(Y^2)=E(\sin^2X)=\int_0^{2\pi}\sin^2x\cdot\frac{1}{2\pi}\mathrm{d}x$$

$$=\frac{1}{4\pi}\int_0^{2\pi}(1-\cos 2x)\mathrm{d}x$$

$$=\frac{1}{4\pi}\left(x-\frac{1}{2}\sin 2x\right)\Big|_0^{2\pi}=\frac{1}{2},$$

$$D(Y)=E(Y^2)-[E(Y)]^2=\frac{1}{2},$$

$$E(Z^2)=E[\sin^2(X+a)]=\int_0^{2\pi}\sin^2(x+a)\cdot\frac{1}{2\pi}\mathrm{d}x$$

$$=\frac{1}{4\pi}\int_0^{2\pi}(1-\cos 2(x+a))\mathrm{d}x=\frac{1}{2},$$

$$D(Z)=E(Z^2)-[E(Z)]^2=\frac{1}{2},$$

$$E(YZ)=E[\sin X\sin(X+a)]$$

$$=\int_0^{2\pi}\sin x\sin(x+a)\cdot\frac{1}{2\pi}\mathrm{d}x$$

$$=\frac{1}{4\pi}\int_0^{2\pi}[\cos a-\cos(2x+a)]\mathrm{d}x=\frac{1}{2}\cos a,$$

$$\mathrm{cov}(Y,Z)=E(YZ)-E(Y)E(Z)=\frac{1}{2}\cos a,$$

则

$$\rho_{YZ}=\frac{\mathrm{cov}(Y,Z)}{\sqrt{D(Y)}\sqrt{D(Z)}}=\frac{\frac{1}{2}\cos a}{\sqrt{\frac{1}{2}}\sqrt{\frac{1}{2}}}=\cos a.$$

当 $a=0,\pi,2\pi$ 时，$|\rho_{YZ}|=1$，$Y$ 与 $Z$ 以概率 1 存在线性关系.

事实上，当 $\rho_{YZ}=1$，即 $a=0,2\pi$ 时，

$$Z=\sin(X+a)=\sin X=Y,$$

而 $\rho_{YZ}=-1$，即 $a=\pi$ 时，

$$Z=\sin(X+\pi)=-\sin X=-Y,$$

即 $Y,Z$ 确实线性相关.

当 $a=\frac{\pi}{2},\frac{3}{2}\pi$ 时，$\rho_{YZ}=0$，即 $Y$ 与 $Z$ 不相关. 但此时 $Y$ 与 $Z$ 并不相互独立，因为这时有

$$Y^2+Z^2=\sin^2X+\sin^2(X+a)=\sin^2X+\cos^2X=1,$$

这表明，虽然当 $a=\frac{\pi}{2}$ 或 $\frac{3}{2}\pi$ 时，$Y$ 与 $Z$ 间不存在线性关系，却有上述函数关系，故此时 $Y$ 与 $Z$ 并不相互独立.

42. 设随机向量 $\xi=(X,Y)$ 的分布密度为

$$p(x,y)=\begin{cases}\frac{1}{4}\sin x\sin y, & 0\leqslant x\leqslant\pi,0\leqslant y\leqslant\pi,\\ 0, & \text{其他}.\end{cases}$$

求 $E(\xi),D(\xi),\rho_{XY}$.

**解　方法 1**　利用表示性定理

$$E[f(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)p(x,y)\mathrm{d}\sigma$$

有

$$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xp(x,y)\mathrm{d}\sigma=\frac{\pi}{2}.$$

同理 $E(Y)=\frac{\pi}{2}$. 而

$$D(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-EX)^2p(x,y)\mathrm{d}\sigma=\frac{\pi^2}{4}-2.$$

同理 $D(Y)=\frac{\pi^2}{4}-2$. 因此

$$E(\xi)=\left(\frac{\pi}{2},\frac{\pi}{2}\right),\quad D(\xi)=\left(\frac{\pi^2}{4}-2,\frac{\pi^2}{4}-2\right),$$

$$\sigma_{XY}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-E(X))(y-E(Y))p(x,y)\mathrm{d}\sigma=0.$$

因此

$$\rho_{XY}=\frac{\sigma_{XY}}{\sqrt{DX}\sqrt{DY}}=0.$$

**方法 2** 因为 $p(x,y)=p_1(x)\cdot p_2(y)$，其中

$$p_1(x)=\begin{cases}\frac{1}{2}\sin x, & 0\leqslant x\leqslant\pi,\\ 0, & 其他;\end{cases}\qquad p_2(y)=\begin{cases}\frac{1}{2}\sin y, & 0\leqslant y\leqslant\pi,\\ 0, & 其他.\end{cases}$$

所以，$X$ 与 $Y$ 相互独立，因此 $\rho_{XY}=0$.

$$E(X)=\int_{-\infty}^{+\infty}xp_1(x)\mathrm{d}x=\frac{\pi}{2},$$

$$E(X^2)=\int_{-\infty}^{+\infty}x^2p_1(x)\mathrm{d}x=\frac{\pi^2}{2}-2.$$

同理 $E(Y)=\frac{\pi}{2}$，$E(Y^2)=\frac{\pi^2}{2}-2$.

$$D(X)=E(X^2)-(E(X))^2=\frac{\pi^2}{2}-2-\frac{\pi^2}{4}=\frac{\pi^2}{4}-2.$$

同理 $D(Y)=\frac{\pi^2}{4}-2$. 因此

$$E(\xi)=\left(\frac{\pi}{2},\frac{\pi}{2}\right),\quad D(\xi)=\left(\frac{\pi^2}{4}-2,\frac{\pi^2}{4}-2\right).$$

43. 设二维随机向量 $\xi=(X,Y)$ 的分布密度函数为

$$p(x,y)=\begin{cases}2x\mathrm{e}^{-(y-5)}, & 0\leqslant x\leqslant 1, y>5,\\ 0, & 其他.\end{cases}$$

讨论 $X$ 与 $Y$ 的独立性，并计算 $E(XY)$.

**解** 因为 $p(x,y)=p_1(x)\cdot p_2(y)$，其中

$$p_1(x)=\begin{cases}2x, & 0\leqslant x\leqslant 1,\\ 0, & 其他;\end{cases}\qquad p_2(y)=\begin{cases}\mathrm{e}^{-(y-5)}, & y\geqslant 5,\\ 0, & 其他,\end{cases}$$

所以，$X$ 与 $Y$ 相互独立. 由于

$$E(X)=\int_{-\infty}^{+\infty}xp_1(x)\mathrm{d}x=\int_0^1 2x^2\mathrm{d}x=\frac{2}{3},$$

$$E(Y)=\int_{-\infty}^{+\infty} y p_2(y)\mathrm{d}y=\int_5^{+\infty} y\mathrm{e}^{-(y-5)}\mathrm{d}y=6,$$

因此

$$E(XY)=(E(X))(E(Y))=\frac{2}{3}\times 6=4.$$

44. 设随机变量 $X\sim\Gamma(1,1)$，随机变量 $Y_k=\begin{cases}0, & 若 X\leqslant k,\\ 1, & 若 X>k\end{cases}$ $(k=1,2)$. 求

(1) $\xi=(Y_1,Y_2)$的分布；

(2) $Y_1$ 与 $Y_2$ 的边缘分布，并讨论它们的独立性；

(3) $E(Y_1+Y_2)$.

**解** 由于 $Y_1$ 的取值为 0,1；$Y_2$ 的取值为 0,1，因此 $\xi=(Y_1,Y_2)$是取值为

$$(0,0),(0,1),(1,0),(1,1)$$

的离散型随机向量，其概率分布为

$$\begin{aligned}P\{Y_1=0,Y_2=0\}&=P\{X\leqslant 1,X\leqslant 2\}\\&=P\{X\leqslant 1\}=F_X(1)=1-\mathrm{e}^{-1}.\end{aligned}$$

同理 $P\{Y_1=0,Y_2=1\}=P\{X\leqslant 1,X>2\}=P(\varnothing)=0$；

$$\begin{aligned}P\{Y_1=1,Y_2=0\}&=P\{X>1,X\leqslant 2\}\\&=P\{1<X\leqslant 2\}=F(2)-F(1)\\&=\mathrm{e}^{-1}-\mathrm{e}^{-2};\end{aligned}$$

$$\begin{aligned}P\{Y_1=1,Y_2=1\}&=P\{X>1,X>2\}\\&=P\{X>2\}=1-P\{X\leqslant 2\}\\&=1-F(2)=1-(1-\mathrm{e}^{-2})=\mathrm{e}^{-2}.\end{aligned}$$

故 $\xi=(Y_1,Y_2)$的分布为

| $Y_1$ \ $Y_2$ | 0 | 1 |
|---|---|---|
| 0 | $1-\mathrm{e}^{-1}$ | 0 |
| 1 | $\mathrm{e}^{-1}-\mathrm{e}^{-2}$ | $\mathrm{e}^{-2}$ |

(2) 由 $\xi$ 的分布，我们有

$$Y_1 \sim \begin{pmatrix} 0 & 1 \\ 1-e^{-1} & e^{-1} \end{pmatrix},$$

即 $Y_1 \sim B(1, e^{-1})$.

$$Y_2 \sim \begin{pmatrix} 0 & 1 \\ 1-e^{-2} & e^{-2} \end{pmatrix},$$

即 $Y_2 \sim B(1, e^{-2})$. 又因为

$$0 = P\{Y_1=0, Y_2=1\} \neq P\{Y_1=0\}P\{Y_2=1\} = (1-e^{-1})e^{-2},$$

所以，$Y_1$ 与 $Y_2$ 不独立.

(3) 因为 $E(Y_1)=e^{-1}, E(Y_2)=e^{-2}$，所以

$$E(Y_1+Y_2)=e^{-1}+e^{-2}.$$

45. 设 $A$ 与 $B$ 是随机试验 $E$ 的两个事件且 $P(A)>0, P(B)>0$，又设随机变量 $X, Y$ 为

$$X = \begin{cases} 1, & 若 A 发生, \\ 0, & 若 A 不发生; \end{cases} \quad Y = \begin{cases} 1, & 若 B 发生, \\ 0, & 若 B 不发生. \end{cases}$$

证明：若 $\rho_{XY}=0$，则 $X$ 与 $Y$ 必定相互独立.

**证** 由题设可知 $X$ 与 $Y$ 的分布律分别为

| $x_i$ | 1 | 0 |
|---|---|---|
| $p_i$ | $P(A)$ | $P(\overline{A})$ |

| $y_j$ | 1 | 0 |
|---|---|---|
| $p_j$ | $P(B)$ | $P(\overline{B})$ |

$Z=XY$ 的分布律为

| $z_k$ | 1 | 0 |
|---|---|---|
| $p_k$ | $P(AB)$ | $1-P(AB)$ |

$$E(X)=P(A), \quad E(Y)=P(B), \quad E(Z)=P(AB).$$

由于

$$\rho_{XY}=\frac{E(XY)-E(X)E(Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=0,$$

因此,$E(XY)=E(X)\cdot E(Y)$,即 $P(AB)=P(A)P(B)$.可知事件 $A$ 与 $B$ 相互独立,因之 $A$ 与 $\overline{B}$,$\overline{A}$ 与 $B$,$\overline{A}$ 与 $\overline{B}$ 均相互独立,故有

$$P\{X=1,Y=1\}=P(AB)=P(A)P(B)=P\{X=1\}P\{Y=1\},$$
$$P\{X=1,Y=0\}=P(A\overline{B})=P(A)P(\overline{B})=P\{X=1\}P\{Y=0\},$$
$$P\{X=0,Y=1\}=P(\overline{A}B)=P(\overline{A})P(B)=P\{X=0\}P\{Y=1\},$$
$$P\{X=0,Y=0\}=P(\overline{A}\,\overline{B})=P(\overline{A})P(\overline{B})=P\{X=0\}P\{Y=0\}.$$

可见,$X$ 与 $Y$ 相互独立.

46. 设 $\xi=(X,Y)$服从在 $D$ 上的均匀分布,其中 $D$ 为 $x$ 轴、$y$ 轴及直线 $x+\frac{y}{2}=1$ 所围成的三角形区域,求 $E(\xi)$和 $D(\xi)$.

**解** 由二维均匀分布的定义,有

$$p(x,y)=\begin{cases}\frac{1}{S_D}, & (x,y)\in D,\\ 0, & \text{其他},\end{cases}$$

其中 $S_D=1$. 于是,二维随机变量$(X,Y)$的联合分布密度为

$$p(x,y)=\begin{cases}1, & (x,y)\in D,\\ 0, & \text{其他}.\end{cases}$$

$X$ 与 $Y$ 的边缘分布密度分别为

$$p_1(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y=\begin{cases}2-2x, & 0<x<1,\\ 0, & \text{其他};\end{cases}$$

$$p_2(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x=\begin{cases}1-\frac{y}{2}, & 0<y<2,\\ 0, & \text{其他}.\end{cases}$$

于是,有

$$E(X)=\int_{-\infty}^{+\infty}xp_1(x)\mathrm{d}x=\int_0^1 x(2-2x)\mathrm{d}x$$
$$=\left(x^2-\frac{2}{3}x^3\right)\Big|_0^1=\frac{1}{3},$$

$$E(X^2)=\int_{-\infty}^{+\infty}x^2p_1(x)\mathrm{d}x=\int_0^1 x^2(2-2x)\mathrm{d}x$$

$$=\left(\frac{2}{3}x^3-\frac{1}{2}x^4\right)\Big|_0^1=\frac{1}{6},$$

$$D(X)=E(X^2)-(E(X))^2=\frac{1}{6}-\left(\frac{1}{3}\right)^2=\frac{1}{18};$$

$$E(Y)=\int_{-\infty}^{+\infty}yp_2(y)\mathrm{d}y=\int_{-\infty}^{+\infty}y\left(1-\frac{y}{2}\right)\mathrm{d}y$$

$$=\int_0^2 y\left(1-\frac{y}{2}\right)\mathrm{d}y=\left(\frac{1}{2}y^2-\frac{1}{6}y^3\right)\Big|_0^2=\frac{2}{3},$$

$$E(Y^2)=\int_{-\infty}^{+\infty}y^2p_2(y)\mathrm{d}y\int_0^2 y^2\left(1-\frac{y}{2}\right)\mathrm{d}y$$

$$=\left(\frac{1}{3}y^3-\frac{1}{8}y^4\right)\Big|_0^2=\frac{2}{3},$$

$$D(Y)=EY^2-(EY)^2=\frac{2}{3}-\left(\frac{2}{3}\right)^2=\frac{2}{9}.$$

因此，

$$E(\xi)=(E(X),E(Y))=\left(\frac{1}{3},\frac{2}{3}\right),$$

$$D(\xi)=(D(X),D(Y))=\left(\frac{1}{18},\frac{2}{9}\right).$$

47. 设 $(X,Y)$ 服从区域 $D$(见图 4.3)上的均匀分布，这里 $D$ 是由 $x$ 轴、$y$ 轴及直线 $x+y+1=0$ 所围成的区域，求 $X$ 与 $Y$ 的相关系数 $\rho$.

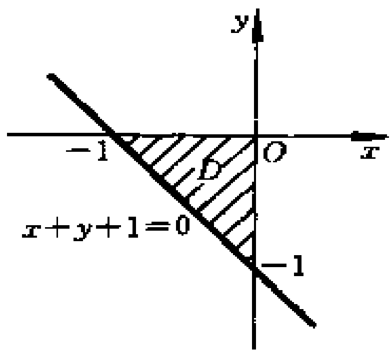



图 4.3

**解** 因为 $D$ 的面积为 $\frac{1}{2}$，故 $X$ 与 $Y$ 的联合密度函数为

$$p(x,y)=\begin{cases}2, & (x,y)\in D,\\ 0, & 其他,\end{cases}$$

于是

$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y=\int_{-1-x}^{0}2\mathrm{d}y$$

$$=2(1+x)\quad(-1<x<0),$$

$$p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x=\int_{-1-y}^{0}2\mathrm{d}x$$
$$=2(1+y)\quad(-1<y<0),$$

即

$$p_X(x)=\begin{cases}2(1+x), & -1<x<0,\\ 0, & \text{其他},\end{cases}$$
$$p_Y(y)=\begin{cases}2(1+y), & -1<y<0,\\ 0, & \text{其他}.\end{cases}$$

由此有

$$E(X)=\int_{-1}^{0}x\cdot 2(1+x)\mathrm{d}x$$
$$=\int_{-1}^{0}(2x+2x^2)\mathrm{d}x=\left(x^2+\frac{2}{3}x^3\right)\Big|_{-1}^{0}$$
$$=-1+\frac{2}{3}=-\frac{1}{3},$$
$$E(X^2)=\int_{-1}^{0}(2x^2+2x^3)\mathrm{d}x=\left(\frac{2}{3}x^3+\frac{x^4}{2}\right)\Big|_{-1}^{0}$$
$$=\frac{2}{3}-\frac{1}{2}=\frac{1}{6},$$
$$D(X)=E(X^2)-(E(X))^2=\frac{1}{6}-\frac{1}{9}=\frac{1}{18}.$$

同理 $$E(Y)=-\frac{1}{3},\quad D(Y)=\frac{1}{18}.$$

进而得到

$$E(XY)=\iint_D xyp(x,y)\mathrm{d}x\mathrm{d}y=\int_{-1}^{0}2x\mathrm{d}x\int_{-1-x}^{0}y\mathrm{d}y$$
$$=\int_{-1}^{0}2x\cdot\frac{y^2}{2}\Big|_{y=-1-x}^{y=0}\mathrm{d}x=-\int_{-1}^{0}x(1+x)^2\mathrm{d}x$$
$$=-\int_{-1}^{0}(x+2x^2+x^3)\mathrm{d}x$$
$$=-\left(\frac{x^2}{2}+\frac{2}{3}x^3+\frac{x^4}{4}\right)\Big|_{-1}^{0}$$

$$= \frac{1}{2} - \frac{2}{3} + \frac{1}{4} = \frac{1}{12},$$

$$\mathrm{cov}(X,Y) = E(XY) - E(X)E(Y)$$

$$= \frac{1}{12} - \left(-\frac{1}{3}\right)\left(-\frac{1}{3}\right) = -\frac{1}{36},$$

从而

$$\rho = \frac{\mathrm{cov}(X,Y)}{\sqrt{D(X)}\ \sqrt{D(Y)}} = \frac{-\frac{1}{36}}{\sqrt{\frac{1}{18}}\sqrt{\frac{1}{18}}} = -\frac{1}{2}.$$

48. 设$(X,Y)$的联合密度函数为

$$p(x,y) = \frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}},$$

求 $Z=\sqrt{X^2+Y^2}$的数学期望与方差.

**解** 根据题意有

$$E(\sqrt{X^2+Y^2}) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\sqrt{x^2+y^2}\,p(x,y)\mathrm{d}x\mathrm{d}y$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\sqrt{x^2+y^2}\,\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}\mathrm{d}x\mathrm{d}y \quad \left(令\begin{cases}x=\rho\cos\theta\\ y=\rho\sin\theta\end{cases}\right)$$

$$= \int_0^{2\pi}\mathrm{d}\theta\int_0^{+\infty}\frac{\rho}{2\pi\sigma^2}e^{-\frac{\rho^2}{2\sigma^2}}\rho\mathrm{d}\rho = \frac{1}{\sigma^2}\int_0^{+\infty}\rho^2 e^{-\frac{\rho^2}{2\sigma^2}}\mathrm{d}\rho$$

$$= -\int_0^{+\infty}\rho e^{-\frac{\rho^2}{2\sigma^2}}\mathrm{d}\left(-\frac{\rho^2}{2\sigma^2}\right) = -\rho e^{-\frac{\rho^2}{2\sigma^2}}\Big|_0^{+\infty} + \int_0^{+\infty}e^{-\frac{\rho^2}{2\sigma^2}}\mathrm{d}\rho$$

$$= 0 + \frac{\sqrt{2\pi}\sigma}{2}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{\rho^2}{2\sigma^2}}\mathrm{d}\rho = \frac{\sqrt{2\pi}}{2}\sigma.$$

$$E((\sqrt{X^2+Y^2})^2) = E(X^2+Y^2)$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x^2+y^2)\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}\mathrm{d}x\mathrm{d}y$$

$$= \int_0^{2\pi}\mathrm{d}\theta\int_0^{+\infty}\frac{\rho^2}{2\pi\sigma^2}e^{-\frac{\rho^2}{2\sigma^2}}\rho\mathrm{d}\rho = \int_0^{+\infty}\frac{\rho^3}{\sigma^2}e^{-\frac{\rho^2}{2\sigma^2}}\mathrm{d}\rho$$

$$\xlongequal{\text{令 } t=\frac{\rho^2}{2\sigma^2}} 2\sigma^2\int_0^{+\infty} t\mathrm{e}^{-t}\mathrm{d}t = 2\sigma^2\Gamma(2) = 2\sigma^2.$$

于是

$$\begin{aligned}D(\sqrt{X^2+Y^2}) &= E[(\sqrt{X^2+Y^2})^2] - [E(\sqrt{X^2+Y^2})]^2 \\ &= 2\sigma^2 - \left(\frac{\sqrt{2\pi}}{2}\sigma\right)^2 = \left(2-\frac{\pi}{2}\right)\sigma^2.\end{aligned}$$

49. 设随机变量 $X_1,X_2,\cdots,X_{m+n}(m<n)$ 相互独立且服从同一分布,并且有有限的方差,试求随机变量 $X=X_1+X_2+\cdots+X_n$ 和 $Y=X_{m+1}+X_{m+2}+\cdots+X_{m+n}$ 的相关系数 $\rho_{XY}$.

**解** 由题意可设 $D(X_i)=\sigma^2(i=1,2,\cdots,m+n)$,则

$$D(X) = D\Big(\sum_{i=1}^{n}X_i\Big) = \sum_{i=1}^{n}D(X_i) = n\sigma^2,$$

$$D(Y) = D\Big(\sum_{i=m+1}^{m+n}X_i\Big) = \sum_{i=m+1}^{m+n}D(X_i) = n\sigma^2,$$

于是

$$\begin{aligned}\mathrm{cov}(X,Y) &= E(XY) - E(X)E(Y) \\ &= E[(X_1+X_2+\cdots+X_n) \\ &\quad \cdot (X_{m+1}+X_{m+2}+\cdots+X_{m+n})] \\ &\quad - E[X_1+X_2+\cdots+X_n] \\ &\quad \cdot E[X_{m+1}+X_{m+2}+\cdots+X_{m+n}] \\ &= E\Big(\sum_{k=m+1}^{n}X_k^2 + \sum_{i\neq j}X_iX_j\Big) - \sum_{k=m+1}^{n}[E(X_k)]^2 \\ &\quad - \sum_{i\neq j}E(X_i)E(X_j) \\ &= \sum_{k=m+1}^{n}E(X_k^2) + \sum_{i\neq j}E(X_i)E(X_j) \\ &\quad - \sum_{k=m+1}^{n}[E(X_k)]^2 - \sum_{i\neq j}E(X_i)E(X_j)\end{aligned}$$

$$= \sum_{k=m+1}^{n} \{E(X_k^2) - [E(X_k)]^2\}$$

$$= \sum_{k=m+1}^{n} D(X_k) = (n-m)\sigma^2,$$

从而随机变量 $X$ 与 $Y$ 的相关系数

$$\rho_{XY} = \frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)}\ \sqrt{D(Y)}} = \frac{(n-m)\sigma^2}{\sqrt{n\sigma^2}\ \sqrt{n\sigma^2}} = \frac{n-m}{n}.$$

50. 设随机变量 $X$ 与 $Y$ 相互独立且都服从正态分布 $N(\mu,\sigma^2)$，试求：

(1) $Z_1=\alpha X+\beta Y$ 和 $Z_2=\alpha X-\beta Y$ 的相关系数 $\rho_{Z_1Z_2}$，其中 $\alpha,\beta$ 为任意常数；

(2) $E[\max(X,Y)]$和 $E[\min(X,Y)]$.

**解** (1) 由题意可知 $E(X)=E(Y)=\mu, D(X)=D(Y)=\sigma^2$，于是

$$E(Z_1)= E[\alpha X + \beta Y] = \alpha E(X) + \beta E(Y) = \mu(\alpha + \beta);$$

$$E(Z_2)= E[\alpha X - \beta Y] = \alpha E(X) - \beta E(Y) = \mu(\alpha - \beta);$$

$$D(Z_1)= D[\alpha X + \beta Y] = \alpha^2 D(X) + \beta^2 D(Y) = \sigma^2(\alpha^2 + \beta^2);$$

$$D(Z_2)= D[\alpha X - \beta Y] = \alpha^2 D(X) + \beta^2 D(Y) = \sigma^2(\alpha^2 + \beta^2);$$

$$\begin{aligned}E(Z_1Z_2) &= E[(\alpha X + \beta Y)(\alpha X - \beta Y)]\\ &= E[\alpha^2X^2 - \beta^2Y^2] = \alpha^2E(X^2) - \beta^2E(Y^2)\\ &= \alpha^2\{D(X) + [E(X)]^2\} - \beta^2\{D(Y) + [E(Y)]^2\}\\ &= \alpha^2(\sigma^2 + \mu^2) - \beta^2(\sigma^2 + \mu^2) = (\alpha^2 - \beta^2)(\sigma^2 + \mu^2);\end{aligned}$$

$$\begin{aligned}\operatorname{cov}(Z_1,Z_2) &= E(Z_1Z_2) - E(Z_1)E(Z_2)\\ &= (\alpha^2 - \beta^2)(\sigma^2 + \mu^2) - \mu(\alpha + \beta)\mu(\alpha - \beta)\\ &= (\alpha^2 - \beta^2)(\sigma^2 + \mu^2) - \mu^2(\alpha^2 - \beta^2)\\ &= (\alpha^2 - \beta^2)\sigma^2,\end{aligned}$$

从而得

$$\rho_{Z_1Z_2}=\frac{\operatorname{cov}(Z_1,Z_2)}{\sqrt{D(Z_1)}\sqrt{D(Z_2)}}=\frac{(\alpha^2-\beta^2)\sigma^2}{\sqrt{\sigma^2(\alpha^2+\beta^2)}\ \sqrt{\sigma^2(\alpha^2+\beta^2)}}$$

$$=\frac{\alpha^2-\beta^2}{\alpha^2+\beta^2}.$$

(2) 由题意可知$(X,Y)$的联合密度函数

$$p(x,y)=\frac{1}{2\pi\sigma^2}e^{-\frac{1}{2\sigma^2}[(x-\mu)^2+(y-\mu)^2]},$$

$$E[\max(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\max(x,y)p(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{x}xp(x,y)\mathrm{d}y+\int_{-\infty}^{+\infty}\mathrm{d}x\int_{x}^{+\infty}yp(x,y)\mathrm{d}y$$

$$=\int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{x}xp(x,y)\mathrm{d}y-\mu\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x\mathrm{d}y$$

$$+\int_{-\infty}^{+\infty}\mathrm{d}x\int_{x}^{+\infty}yp(x,y)\mathrm{d}y+\mu\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{x}xp(x,y)\mathrm{d}y$$

$$-\mu\left[\int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{x}p(x,y)\mathrm{d}y+\int_{-\infty}^{+\infty}\mathrm{d}x\int_{x}^{+\infty}p(x,y)\mathrm{d}y\right]$$

$$+\int_{-\infty}^{+\infty}\mathrm{d}x\int_{x}^{+\infty}yp(x,y)\mathrm{d}y$$

$$+\mu\left(\text{因为}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x\mathrm{d}y=1\right)$$

$$=\int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{x}(x-\mu)p(x,y)\mathrm{d}y$$

$$+\int_{-\infty}^{+\infty}\mathrm{d}x\int_{x}^{+\infty}(y-\mu)p(x,y)\mathrm{d}y+\mu$$

$$=\int_{-\infty}^{+\infty}\mathrm{d}y\int_{y}^{+\infty}(x-\mu)p(x,y)\mathrm{d}x$$

$$+\int_{-\infty}^{+\infty}\mathrm{d}y\int_{y}^{+\infty}(x-\mu)p(x,y)\mathrm{d}x+\mu$$

$$=2\int_{-\infty}^{+\infty}\mathrm{d}y\int_{y}^{+\infty}(x-\mu)\frac{1}{2\pi\sigma^2}e^{-\frac{1}{2\sigma^2}[(x-\mu)^2+(y-\mu)^2]}\mathrm{d}x+\mu$$

$$= \frac{1}{\pi\sigma^2}\int_{-\infty}^{+\infty} e^{-\frac{1}{2\sigma^2}(y-\mu)^2} dy \int_y^{+\infty} (x-\mu) e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx + \mu$$

$$= \frac{1}{\pi}\int_{-\infty}^{+\infty} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy \int_y^{+\infty} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} d\left(\frac{1}{2\sigma^2}(x-\mu)^2\right) + \mu$$

$$= \frac{1}{\pi}\int_{-\infty}^{+\infty} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \left[-e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\right]_{x=y}^{x=+\infty} dy + \mu$$

$$= \frac{1}{\pi}\int_{-\infty}^{+\infty} e^{-\frac{(y-\mu)^2}{\sigma^2}} dy + \mu$$

$$\xlongequal{\text{令 } t = \frac{y-\mu}{\sigma}} \frac{\sigma}{\pi}\int_{-\infty}^{+\infty} e^{-t^2} dt + \mu$$

$$= \frac{\sigma}{\pi}\sqrt{\pi}\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sqrt{\frac{1}{2}}} e^{-\frac{t^2}{2\cdot\frac{1}{2}}} dt + \mu$$

$$= \frac{\sigma}{\sqrt{\pi}} + \mu,$$

$$E[\min(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \min(x,y) p(x,y) dx dy$$

$$= \int_{-\infty}^{+\infty} dy \int_{-\infty}^{y} x p(x,y) dx + \int_{-\infty}^{+\infty} dy \int_y^{+\infty} y p(x,y) dx$$

$$= \int_{-\infty}^{+\infty} dy \int_{-\infty}^{y} x p(x,y) dx - \mu\left[\int_{-\infty}^{+\infty} dy \int_{-\infty}^{y} p(x,y) dx\right.$$

$$\left. + \int_{-\infty}^{+\infty} dy \int_y^{+\infty} p(x,y) dx\right] + \int_{-\infty}^{+\infty} dy \int_y^{+\infty} y p(x,y) dx + \mu$$

$$= \int_{-\infty}^{+\infty} dy \int_{-\infty}^{y} (x-\mu) p(x,y) dx$$

$$+ \int_{-\infty}^{+\infty} dy \int_y^{+\infty} (y-\mu) p(x,y) dx + \mu$$

$$= \int_{-\infty}^{+\infty} dx \int_{-\infty}^{x} (y-\mu) p(x,y) dy$$

$$+ \int_{-\infty}^{+\infty} dx \int_{-\infty}^{x} (y-\mu) p(x,y) dy + \mu$$

$$= 2\int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{x} (y-\mu) p(x,y) \mathrm{d}y + \mu$$

$$= 2\int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{x} (y-\mu) \frac{1}{2\pi\sigma^2} \mathrm{e}^{-\frac{1}{2\sigma^2}[(x-\mu)^2+(y-\mu)^2]} \mathrm{d}y + \mu$$

$$= \frac{1}{\pi\sigma^2}\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{1}{2\sigma^2}(x-\mu)^2} \cdot \left[-\sigma^2 \mathrm{e}^{-\frac{1}{2\sigma^2}(y-\mu)^2}\right]\Big|_{-\infty}^{x} \mathrm{d}x + \mu$$

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

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

$$= -\frac{\sigma}{\pi}\sqrt{\pi}\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sqrt{\frac{1}{2}}} \mathrm{e}^{-t^2} \mathrm{d}t + \mu$$

$$= \mu - \frac{\sigma}{\sqrt{\pi}}.$$

51. 设随机变量 $X$ 和 $Y$ 相互独立，试证

$$D(XY) = D(X)D(Y) + [E(X)]^2 D(Y) + [E(Y)]^2 D(X).$$

**证** 因为随机变量 $X$ 与 $Y$ 独立，所以

$$E(XY) = E(X)E(Y), \quad E(X^2Y^2) = E(X^2)E(Y^2),$$

于是有

$$\begin{aligned} D(XY) &= E(X^2Y^2) - [E(XY)]^2 \\ &= E(X^2)E(Y^2) - [E(X)]^2[E(Y)]^2. \end{aligned}$$

又因为

$$D(X) = E(X^2) - [E(X)]^2, \quad D(Y) = E(Y^2) - [E(Y)]^2,$$

所以

$$E(X^2) = D(X) + [E(X)]^2, \quad E(Y^2) = D(Y) + [E(Y)]^2,$$

从而

$$\begin{aligned} D(XY) &= \{D(X) + [E(X)]^2\}E(Y^2) - [E(X)]^2[E(Y)]^2 \\ &= D(X)E(Y^2) + [E(X)]^2\{E(Y^2) - [E(Y)]^2\} \\ &= D(X)\{D(Y) + [E(Y)]^2\} + [E(X)]^2 D(Y) \end{aligned}$$

$$= D(X)D(Y) + [E(X)]^2D(Y) + [E(Y)]^2D(X).$$

52. 设随机向量$(X,Y)$的联合密度函数为

$$p(x,y)=\begin{cases}A\sin(x,y), & 0\leqslant x\leqslant \dfrac{\pi}{2},0\leqslant y\leqslant \dfrac{\pi}{2},\\ 0, & \text{其他},\end{cases}$$

求 (1) 系数$A$； (2) 数学期望$E(X),E(Y)$； (3) 方差$D(X)$，$D(Y)$； (4) 协方差$\operatorname{cov}(X,Y)$和相关系数$\rho_{XY}$.

**解** (1) 因为$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x\mathrm{d}y=1$，所以有

$$\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}}A\sin(x+y)\mathrm{d}x\mathrm{d}y = A\int_0^{\frac{\pi}{2}}[-\cos(x+y)]_0^{\frac{\pi}{2}}\mathrm{d}x$$

$$= A\int_0^{\frac{\pi}{2}}\left[\cos x-\cos\left(x+\frac{\pi}{2}\right)\right]\mathrm{d}x$$

$$= A\left[\sin x-\sin\left(\frac{\pi}{2}+x\right)\right]_0^{\frac{\pi}{2}} = 2A = 1,$$

从而$A=\dfrac{1}{2}$.

(2) $$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xp(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\frac{1}{2}\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}}x\sin(x+y)\mathrm{d}x\mathrm{d}y$$

$$=\frac{1}{2}\int_0^{\frac{\pi}{2}}x[-\cos(x+y)]_0^{\frac{\pi}{2}}\mathrm{d}x$$

$$=\frac{1}{2}\int_0^{\frac{\pi}{2}}x\left[\cos x-\cos\left(x+\frac{\pi}{2}\right)\right]\mathrm{d}x$$

$$=\frac{1}{2}\left[x\sin x-x\sin\left(x+\frac{\pi}{2}\right)\right]\Bigg|_0^{\frac{\pi}{2}}$$

$$-\frac{1}{2}\int_0^{\frac{\pi}{2}}\left[\sin x-\sin\left(x+\frac{\pi}{2}\right)\right]\mathrm{d}x$$

$$= \frac{\pi}{4} + \frac{1}{2}\left[\cos x - \cos\left(x + \frac{\pi}{2}\right)\right]\Big|_0^{\frac{\pi}{2}} = \frac{\pi}{4}.$$

同理 $$E(Y) = E(X) = \frac{\pi}{4}.$$

(3)

$$\begin{aligned}
D(X) &= E[X - E(X)]^2 \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x - E(X))^2 p(x,y)\mathrm{d}x\mathrm{d}y \\
&= \frac{1}{2}\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}}\left(x - \frac{\pi}{4}\right)^2 \sin(x+y)\mathrm{d}y\mathrm{d}x \\
&= \frac{1}{2}\int_0^{\frac{\pi}{2}}\left(x - \frac{\pi}{4}\right)^2 [-\cos(x+y)]\Big|_0^{\frac{\pi}{2}}\mathrm{d}x \\
&= \frac{1}{2}\int_0^{\frac{\pi}{2}}\left(x - \frac{\pi}{4}\right)^2 (\sin x + \cos x)\mathrm{d}x \\
&= \frac{1}{2}\int_0^{\frac{\pi}{2}}\left(x - \frac{\pi}{4}\right)^2 (\mathrm{d}\sin x - \mathrm{d}\cos x) \\
&= \frac{1}{2}\left(x - \frac{\pi}{4}\right)^2 (\sin x - \cos x)\Big|_0^{\frac{\pi}{2}} \\
&\quad - \int_0^{\frac{\pi}{2}}\left(x - \frac{\pi}{4}\right)(\sin x - \cos x)\mathrm{d}x \\
&= \frac{\pi^2}{16} + \int_0^{\frac{\pi}{2}}\left(x - \frac{\pi}{4}\right)\mathrm{d}(\sin x + \cos x) \\
&= \frac{\pi^2}{16} + \left(x - \frac{\pi}{4}\right)(\sin x + \cos x)\Big|_0^{\frac{\pi}{2}} \\
&\quad - \int_0^{\frac{\pi}{2}}(\sin x + \cos x)\mathrm{d}x \\
&= \frac{\pi^2}{16} + \frac{\pi}{2} - (\sin x - \cos x)\Big|_0^{\frac{\pi}{2}} \\
&= \frac{\pi^2}{16} + \frac{\pi}{2} - 2.
\end{aligned}$$

同理 $$D(Y)=D(X)=\frac{\pi^2}{16}+\frac{\pi}{2}-2.$$

(4)

$$\mathrm{cov}(X,Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x-E(X)][y-E(Y)]p(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\frac{1}{2}\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}}\left(x-\frac{\pi}{4}\right)\left(y-\frac{\pi}{4}\right)\sin(x+y)\mathrm{d}x\mathrm{d}y$$

$$=\frac{1}{2}\int_0^{\frac{\pi}{2}}\left(x-\frac{\pi}{4}\right)\left[\int_0^{\frac{\pi}{2}}\left(y-\frac{\pi}{4}\right)\mathrm{d}[-\cos(x+y)]\right]\mathrm{d}x$$

$$=\frac{1}{2}\int_0^{\frac{\pi}{2}}\left(x-\frac{\pi}{4}\right)\left[-\left(y-\frac{\pi}{4}\right)\cos(x+y)\right]\Big|_0^{\frac{\pi}{2}}\mathrm{d}x$$

$$+\frac{1}{2}\int_0^{\frac{\pi}{2}}\left(x-\frac{\pi}{4}\right)\int_0^{\frac{\pi}{2}}\cos(x+y)\mathrm{d}y\mathrm{d}x$$

$$=\frac{1}{2}\int_0^{\frac{\pi}{2}}\left(x-\frac{\pi}{4}\right)\left[-\frac{\pi}{4}\cos\left(x+\frac{\pi}{2}\right)-\frac{\pi}{4}\cos x\right]\mathrm{d}x$$

$$+\frac{1}{2}\int_0^{\frac{\pi}{2}}\left(x-\frac{\pi}{4}\right)\sin(x+y)\Big|_0^{\frac{\pi}{2}}\mathrm{d}x$$

$$=-\frac{\pi}{8}\int_0^{\frac{\pi}{2}}\left(x-\frac{\pi}{4}\right)\mathrm{d}\left[\sin\left(x+\frac{\pi}{2}\right)+\sin x\right]$$

$$+\frac{1}{2}\int_0^{\frac{\pi}{2}}\left(x-\frac{\pi}{4}\right)\left[\sin\left(x+\frac{\pi}{2}\right)-\sin x\right]\mathrm{d}x$$

$$=-\frac{\pi}{8}\left(x-\frac{\pi}{4}\right)\left[\sin\left(x+\frac{\pi}{2}\right)+\sin x\right]\Big|_0^{\frac{\pi}{2}}$$

$$+\frac{\pi}{8}\int_0^{\frac{\pi}{2}}[\sin x+\cos x]\mathrm{d}x$$

$$+\frac{1}{2}\int_0^{\frac{\pi}{2}}\left(x-\frac{\pi}{4}\right)\mathrm{d}(\sin x+\cos x)$$

$$=-\frac{\pi}{8}\cdot\frac{\pi}{4}+\frac{\pi}{8}\left(-\frac{\pi}{4}\right)+\frac{\pi}{8}[\sin x-\cos x]\Big|_0^{\frac{\pi}{2}}$$

$$+\frac{1}{2}\left(x-\frac{\pi}{4}\right)(\sin x+\cos x)\Big|_0^{\frac{\pi}{2}}$$

$$-\frac{1}{2}\int_0^{\frac{\pi}{2}}(\sin x+\cos x)\mathrm{d}x$$

$$=-\frac{\pi^2}{16}+\frac{\pi}{8}\cdot 2+\frac{1}{2}\left(\frac{\pi}{4}+\frac{\pi}{4}\right)$$

$$-\frac{1}{2}(\sin x-\cos x)\Big|_0^{\frac{\pi}{2}}$$

$$=-\frac{\pi^2}{16}+\frac{\pi}{2}-1.$$

$$\rho_{XY}=\frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=\frac{-\frac{\pi^2}{16}+\frac{\pi}{2}-1}{\frac{\pi^2}{16}+\frac{\pi}{2}-2}.$$

## 四、练习题

1. 设 $X$ 表示 10 次独立重复射击中命中目标的次数，每次射中目标的概率为 0.4，求 $X^2$ 的数学期望.

2. 设 $X$、$Y$ 是两个相互独立且均服从正态分布 $N\left(0,\left(\frac{1}{\sqrt{2}}\right)^2\right)$ 的随机变量，求 $E(|X-Y|)$.

3. 设随机变量 $X$ 与 $Y$ 同分布，$X$ 的概率密度为

$$f(x)=\begin{cases}\frac{3}{8}x^2, & 0<x<2,\\ 0, & \text{其他}.\end{cases}$$

(1) 已知事件 $A=\{X>a\}$ 和 $B=\{Y>a\}$ 独立，且 $P(A\cup B)=\frac{3}{4}$，求常数 $a$；

(2) 求 $\frac{1}{X^2}$ 的数学期望.

4. 设 $\xi,\eta$ 是相互独立且服从同一分布的两个随机变量，已知

$\xi$的分布律为 $P\{\xi=i\}=\frac{1}{3}, i=1,2,3$. 又设 $X=\max(\xi,\eta), Y=\min(\xi,\eta)$.

(1) 写出二维随机变量$(X,Y)$的分布律；

(2) 求随机变量 $X$ 的数学期望 $E(X)$.

5. 设随机变量 $X$ 服从 Laplace 分布，其密度函数为

$$p(x)=\frac{1}{2\lambda}e^{-\frac{|x-\mu|}{\lambda}} \quad (\lambda>0 \text{ 为常数}),$$

试求 $E(X), D(X)$.

6. 设二维随机变量$(X,Y)$在区域 $D$：$0<x<y<1$ 内服从均匀分布.

(1) 写出$(X,Y)$的联合密度函数；

(2) 求出 $X$ 和 $Y$ 的边缘概率密度函数；

(3) 求相关系数 $\rho$.

7. 随机向量$(X,Y)$在上半单位圆 $D=\{(x,y): x^2+y^2\leqslant 1, y\geqslant 0\}$上均匀分布，求相关系数 $\rho$，并问 $X$ 与 $Y$ 是否独立?

8. 若 $X_1$ 服从参数 $\lambda=\frac{1}{2}$的指数分布，$X_2$ 的密度函数为

$$p_2(X)=\begin{cases} Cxe^{-\frac{1}{2}x}, & x>0, \\ 0, & x\leqslant 0, \end{cases}$$

试计算 $C$ 与 $E(X_2)$的值.

9. 设随机变量 $X$ 的密度函数为

$$p(X)=\begin{cases} \dfrac{x^n e^{-x}}{n!}, & x\geqslant 0, \\ 0, & x<0, \end{cases}$$

求 $E(X), D(X)$.

10. 设随机变量 $X$ 的概率密度函数为

$$p(x)=\begin{cases} ax^2+bx+c, & 0<x<1, \\ 0, & \text{其他}. \end{cases}$$

已知 $E(X)=0.5, D(X)=0.15$，求系数 $a,b,c$.

11. 一辆送客汽车，载有 20 位乘客从起点站开出，沿途有 10 个车站可以下车，若到达一个车站，没有乘客下车就不停车. 设每位乘客在每一个车站下车是等可能的，试求汽车平均停车次数.

12. 已知 $D(X)=25, D(Y)=36, \rho_{XY}=0.4$，试求 $D(X+Y)$ 和 $D(X-Y)$.

13. 设随机向量 $(X,Y)$ 的概率密度为

$$p(X,Y)=\begin{cases}6xy, & 0<x<1, 0<y<2(1-x),\\ 0, & \text{其他}.\end{cases}$$

求 $E(X), D(Y)$ 及 $E(XY)$.

14. 已知二维随机变量 $(X,Y)$ 服从二维正态分布，$E(X)=E(Y)=0, D(X)=16, D(Y)=25, \operatorname{cov}(X,Y)=12$，求 $(X,Y)$ 的密度函数 $p(x,y)$.

15. 设 $(X,Y)$ 的密度函数为

$$p(x,y)=\begin{cases}e^{y}, & 0<x<y<1,\\ 0, & \text{其他},\end{cases}$$

求相关系数 $\rho$.

## 习题答案与提示

**1.** 18.4. **2.** $\sqrt{2/\pi}$. **3.** (1) $a=\sqrt[3]{4}$； (2) $E\left(\frac{1}{X^2}\right)=\frac{3}{4}$.

**4.** (1)

| $X$ \ $Y$ | 1 | 2 | 3 |
|---|---|---|---|
| 1 | $\frac{1}{9}$ | 0 | 0 |
| 2 | $\frac{2}{9}$ | $\frac{1}{9}$ | 0 |
| 3 | $\frac{2}{9}$ | $\frac{2}{9}$ | $\frac{1}{9}$ |

(2) $E(X)=\frac{22}{9}$.

5. $E(X)=\mu$; $D(X)=2\lambda^2$.

6. (1) $p(x,y)=\begin{cases}2, & (x,y)\in D,\\ 0, & \text{其他};\end{cases}$

(2) $p_1(x)=\begin{cases}2(1-x), & (0<x<1),\\ 0, & \text{其他};\end{cases}$

$p_2(y)=\begin{cases}2y, & 0<y<1,\\ 0, & \text{其他};\end{cases}$

(3) $\rho=\dfrac{1}{2}$.

7. $\rho=0$; 不独立.

8. $C=\dfrac{1}{4}$, $E(X_2)=4$.  9. $E(X)=D(X)=n+1$.

10. $a=12, b=-12, c=3$.  11. $10(1-(0.9)^{20})$.

12. $D(X+Y)=85$; $D(X-Y)=37$.

13. $E(X)=\dfrac{2}{5}$; $D(Y)=\dfrac{4}{25}$; $E(XY)=\dfrac{14}{5}$.

14. $p(x,y)=\dfrac{1}{32\pi}e^{-\frac{25}{32}\left[\frac{x^2}{16}-\frac{3xy}{50}+\frac{y^2}{25}\right]}$.

15. $\rho=\sqrt{\dfrac{3[3-(e-1)^2]}{2+4e-3e^2}}$.

# 第五章　大数定律和中心极限定理

## 一、考试要求

1. 了解切比雪夫不等式；

2. 了解大数定律成立的条件及结论；

3. 了解中心极限定理应用的条件及结论；并会利用相关的定理近似计算有关事件的概率.

## 二、复习要点

### （一）重要概念及性质

**定义 5.1**（独立同分布的随机变量列）　如果对于任何 $n\geqslant 1$，$X_1,X_2,\cdots,X_n$ 是相互独立的，那么称随机变量列 $X_1,X_2,\cdots,X_n,\cdots$ 是相互独立的. 此时，若所有的 $X_i$ 都有共同的分布，则称 $X_1,X_2,\cdots,X_n,\cdots$ **是独立同分布的随机变量列.**

**定义 5.2**（依概率收敛）　设 $X_n(n=1,2,\cdots)$为随机变量列，若存在随机变量 $X$，对于任意 $\varepsilon>0$，有

$$\lim_{n\to\infty}P\{|X_n-X|\geqslant\varepsilon\}=0 \quad 或 \quad \lim_{n\to\infty}P\{|X_n-X|<\varepsilon\}=1,$$

则称随机变量列$\{X_n\}$**依概率收敛**于随机变量 $X$，并用下面符号表示：

$$X_n\xrightarrow{P}X \quad 或 \quad \lim_{n\to\infty}X_n=X \quad (P).$$

**定义 5.3**（大数定律）　设$\{X_n\}$为一随机变量列，并且 $E(X_n)$存在，令

$$\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_n.$$

若

$$\lim_{n\to\infty}[\overline{X}_n - E(\overline{X}_n)] = 0 \quad (P),$$

则称随机变量列$\{X_n\}$服从**大数定律**.

**定义 5.4（中心极限定理）** 设$\{X_n\}$为随机变量列，并且$E(X_k)=a_k<+\infty, D(X_n)=\sigma_k^2<+\infty$ $(k=1,2,\cdots)$. 令

$$B_n^2 = \sum_{k=1}^{n} D(X_k), \quad Y_n = \sum_{k=1}^{n} \frac{X_n - a_k}{B_n} \quad (n = 1,2,\cdots).$$

若对于$x\in(-\infty,+\infty)$一致地有

$$\lim_{n\to\infty} P(Y_n \leqslant x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt,$$

则称$\{X_n\}$服从**中心极限定理**.

（二）重要定理及公式

**定理 5.1（切比雪夫不等式）** 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$\varepsilon$，有下列**切比雪夫**不等式

$$P\{|X-\mu| \geqslant \varepsilon\} \leqslant \frac{\sigma^2}{\varepsilon^2}.$$

**定理 5.2（切比雪夫大数定律）** 设随机变量$X_1,X_2,\cdots$相互独立，均具有有限方差，且被同一常数$C$所界：$D(X_i)<C(i=1,2,\cdots)$，则对于任意的正数$\varepsilon$，有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i)\right| < \varepsilon\right\} = 1.$$

特殊情形：若$X_1,X_2,\cdots$具有相同的数学期望$E(X_i)=\mu$，则上式成为

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| < \varepsilon\right\} = 1.$$

**定理 5.3（伯努利大数定律）** 设$n_A$是$n$次独立试验中事件$A$发生的次数，$p$是事件$A$在每次试验中发生的概率，则对于任意的正数$\varepsilon$，有

$$\lim_{n\to\infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1.$$

**定理 5.4**(辛钦大数定律) 设 $X_1,X_2,\cdots,X_n,\cdots$是相互独立同分布的随机变量序列,且 $E(X_n)=\mu$,则

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right|<\epsilon\right\}=1.$$

**例 1** 设$\{X_k\}$为相互独立且同分布的随机变量序列,并且 $X_k$ 的概率分布为

$$P\{X_k=2^{i-2\ln i}\}=2^{-i}\quad(i=1,2,\cdots),$$

试证$\{X_k\}$服从大数定律.

**证** 因为

$$E(X_k)=\sum_{i=1}^{\infty}2^{i-2\ln i}\cdot 2^{-i}=\sum_{i=1}^{\infty}\frac{1}{2^{2\ln i}}=\sum_{i=1}^{\infty}\frac{1}{4^{\ln i}}=\sum_{i=1}^{\infty}\frac{1}{i^{\ln 4}}<+\infty$$

$(k=1,2,\cdots)$,由辛钦大数定律可知$\{X_k\}$服从大数定律.

**例 2** 设$\{X_k\}$为相互独立的随机变量序列,且

$$X_k\sim\begin{pmatrix}-2^k & 0 & 2^k\\ \frac{1}{2^{2k+1}} & 1-\frac{1}{2^{2k}} & \frac{1}{2^{2k+1}}\end{pmatrix},\quad k=1,2,\cdots,$$

试证$\{X_k\}$服从大数定律.

**证** 由题意可知

$$E(X_k)=(-2^k)\times\frac{1}{2^{2k+1}}+0\times\left(1-\frac{1}{2^{2k}}\right)+2^k\cdot\frac{1}{2^{2k+1}}=0,$$

$$D(X_k)=E(X_k^2)=(-2^k)^2\times\frac{1}{2^{2k+1}}+(2^k)^2\cdot\frac{1}{2^{2k+1}}=1,$$

$k=1,2,\cdots$. 由切比雪夫大数定律可知随机变量序列$\{X_k\}$服从大数定律.

**定理 5.5**(列维-林德伯格定理) 设随机变量 $X_1,X_2,\cdots$相互独立,服从同一分布,且具有相同的数学期望和方差: $E(X_k)=\mu$, $D(X_k)=\sigma^2\neq 0(k=1,2,\cdots)$,则随机变量

$$Y_n=\frac{\sum_{k=1}^{n}X_k-n\mu}{\sqrt{n}\,\sigma}$$

的分布函数 $F_n(x)$ 对任意的实数 $x$，有

$$\lim_{n\to\infty}F_n(x)=\lim_{n\to\infty}P\left\{\frac{\sum_{k=1}^{n}X_k-n\mu}{\sqrt{n}\,\sigma}\leqslant x\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt.$$

此定理也称为独立同分布的**中心极限定律**.

**定理 5.6**（棣莫弗-拉普拉斯） 设随机变量 $X_1,X_2,\cdots$ 均为具有参数 $n,p(0<p<1)$ 的二项分布，则对于任意实数 $x$，有

$$\lim_{n\to\infty}P\left\{\frac{X_n-np}{\sqrt{np(1-p)}}\leqslant x\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt.$$

**例 3** 某车间有 200 台车床，在生产时间内由于需要检修，调换刀具，变换位置，调换工件等常需停车. 设开工率为 0.6，并设每台车床的工作是独立的，且在开工时需电力 1 kW，问应供应该车间多少瓦电力，才能以 99.9% 的概率保证该车间不会因供电不足而影响生产.

**解** 首先，我们把对每台车床的观察作为一次试验，若观察到车床正在工作作为事件 $A$，而由题意可知 $A$ 发生的概率为 0.6. 由于 200 台车床是独立地工作，因此这是一个重复独立试验. 把某时刻正在工作的车床数记为 $S_n$，并把第 $i$ 台车床开工用 $\{X_i=1\}$ 表示，停工用 $\{X_i=0\}$ 表示，则 $X_i\sim B(1,0.6)(i=1,2,\cdots,200)$，且 $S_n=X_1+X_2+\cdots+X_{200}$，显然有 $0\leqslant S_n\leqslant 200$. 现在问题是要求 $m$，使

$$P\{S_n\leqslant m\}=\sum_{k=0}^{m}C_{200}^{k}(0.6)^k(0.4)^{200-k}\geqslant 0.999. \tag{5.1}$$

由于

$$\begin{aligned}P\{S_n\leqslant m\}&=P\{0\leqslant S_n\leqslant m\}\\&=P\left\{\frac{0-200\times0.6}{\sqrt{200\times0.6\times0.4}}\leqslant\frac{S_n-200\times0.6}{\sqrt{200\times0.6\times0.4}}\right.\\&\qquad\left.\leqslant\frac{m-200\times0.6}{\sqrt{200\times0.6\times0.4}}\right\},\end{aligned}$$

然后根据中心极限定理，

$$P\{S_n \leqslant m\} = \sum_{k=0}^{m} C_{200}^{k}(0.6)^k(0.4)^{200-k}$$

$$= P\left\{\frac{-120}{\sqrt{48}} \leqslant \frac{S_n - 120}{\sqrt{48}} \leqslant \frac{m-120}{\sqrt{48}}\right\}$$

$$\approx \int_{\frac{-120}{(48)^{1/2}}}^{\frac{m-120}{(48)^{1/2}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$$

$$= \Phi\left(\frac{m-120}{6.928}\right) - \Phi(-17.32)$$

$$= \Phi\left(\frac{m-120}{6.928}\right) \geqslant 0.999.$$

当 $\Phi\left(\frac{m-120}{6.928}\right)=0.999$ 时，得 $\frac{m-120}{6.928}=3.1$，所以 $m=141.4768$，取 $m$ 为 142.

这结果表示，$P\{S_n \leqslant 142\} \geqslant 0.999$. 可见，若供电 142 kW，则由于供电不足而影响生产的可能性小于 0.001. 即在 8 h 的工作时间中可能少于 $8\times60\times0.001=0.48$ min 的时间会受影响，这在一般情况下是允许的. 当然不同的生产单位有不同的要求. 这只要改变(5.1)式右端的概率值即可.

**例 4** 假设 $X_1, X_2, \cdots, X_n$ 是来自总体 $X$ 的简单随机样本，已知 $E(X^k)=\alpha_k(k=1,2,3,4)$，证明当 $n$ 充分大时，随机变量 $Z_n=\frac{1}{n}\sum_{i=1}^{n}X_i^2$ 近似服从正态分布，并指出其分布参数.

**证** 依题意 $X_1, X_2, \cdots, X_n$ 独立同分布，可见 $X_1^2, X_2^2, \cdots, X_n^2$ 也独立同分布. 由 $E(X^k)=\alpha_k(k=1,2,3,4)$ 有

$$E(X_i^2) = \alpha_2,$$

$$D(X_i^2) = E(X_i^4) - [E(X_i^2)]^2 = \alpha_4 - \alpha_2^2,$$

$$E(Z_n) = E\left[\frac{1}{n}\sum_{i=1}^{n}X_i^2\right] = \frac{1}{n}\sum_{i=1}^{n}E(X_i^2) = \alpha_2,$$

$$D(Z_n) = D\left[\frac{1}{n}\sum_{i=1}^{n}X_i^2\right] = \frac{1}{n^2}\sum_{i=1}^{n}D(X_i^2)$$

$$=\frac{1}{n}(\alpha_4-\alpha_2^2),$$

因此,根据中心极限定理 $U_n=\dfrac{Z_n-\alpha_2}{\sqrt{(\alpha_4-\alpha_2^2)/n}}$ 的极限分布是标准正态分布,即当 $n$ 充分大时,$Z_n$ 近似服从 $N\left(\alpha_2,\frac{1}{n}(\alpha_4-\alpha_2^2)\right)$.

## 三、典型例题分析

**解答题**

1. 利用切比雪夫不等式估计随机变量与其数学期望之差大于 3 倍标准差的概率.

**解** 由切比雪夫不等式

$$P\{|X-E(X)|>\varepsilon\}\leqslant\frac{D(X)}{\varepsilon^2}.$$

令 $\varepsilon=3\sqrt{D(X)}$,有

$$P\{|X-E(X)|>3\sqrt{D(X)}\}\leqslant\frac{D(X)}{(3\sqrt{D(X)})^2}=\frac{1}{9}.$$

2. 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 次独立试验中,事件 $A$ 发生的次数在 400~600 之间的概率.

**分析** 利用切比雪夫不等式估计某事件的概率,要做以下工作:(1) 选择随机变量 $X$;(2) 求出 $E(X)$,$D(X)$;(3) 根据题意确定 $\varepsilon$;(4) 利用切比雪夫不等式对概率进行估计.

**解** 设 $X$ 表示在 1000 次独立试验中事件 $A$ 发生的次数,则 $X\sim B(1000,0.5)$,且

$$E(X)=np=1000\times0.5=500,$$

$$D(X)=npq=1000\times0.5\times0.5=250.$$

要估计事件$\{400<X<600\}$的概率,可改写如下形式:

$$\begin{aligned}\{400<X<600\}&=\{400-500<X-500<600-500\}\\&=\{-100<X-E(X)<100\}\end{aligned}$$

$$= \{|X - E(X)| < 100\},$$

在切比雪夫不等式中取 $\varepsilon=100$，则有

$$P\{400 < X < 600\} = P\{|X - E(X)| < 100\}$$

$$\geqslant 1 - \frac{D(X)}{100^2} = 1 - \frac{250}{10000} = \frac{39}{40}.$$

3. 某工厂有 400 台同类机器，各台机器发生故障的概率都是 0.02. 假设各台机器工作是相互独立的，试求机器出故障的台数不少于 2 的概率.

**解** 设机器出故障的台数为 $X$，则

$$X \sim B(400, 0.02).$$

经过计算有 $E(X)=8$，$D(X)=7.84$，$\sqrt{D(X)}=2.8$. 由中心极限定理，有

$$P\{X \geqslant 2\} = 1 - P\{X < 2\} = 1 - P\left\{\frac{X-8}{2.8} \leqslant \frac{2-8}{2.8}\right\}$$

$$= 1 - \Phi(-2.1429) \approx 0.9842.$$

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

［附表］

| $x$ | 0 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 |
|---|---|---|---|---|---|---|---|
| $\Phi(x)$ | 0.500 | 0.692 | 0.841 | 0.933 | 0.977 | 0.994 | 0.999 |

表中 $\Phi(x)$ 是标准正态分布函数.

**解** 设 $X$ 为考生的外语成绩，由题设 $X \sim N(\mu, \sigma^2)$，其中 $\mu=72$. 现在求 $\sigma^2$. 由条件知

$$0.023 = P\{X \geqslant 96\} = P\left\{\frac{X-\mu}{\sigma} \geqslant \frac{96-72}{\sigma}\right\} = 1 - \Phi\left(\frac{24}{\sigma}\right),$$

从而 $\Phi\left(\frac{24}{\sigma}\right)=0.977$. 由 $\Phi(x)$ 的数值表，可见 $\frac{24}{\sigma}=2$，因此 $\sigma=12$，这样 $X \sim N(72, 12^2)$. 故所求概率为：

$$Q\{60 \leqslant X \leqslant 84\} = P\left\{\frac{60-72}{12} \leqslant \frac{X-\mu}{\sigma} \leqslant \frac{84-72}{12}\right\}$$
$$= P\left\{-1 \leqslant \frac{X-\mu}{\sigma} \leqslant 1\right\}$$
$$= \Phi(1) - \Phi(-1) = 2\Phi(1) - 1$$
$$= 2 \times 0.841 - 1 = 0.682.$$

5. 试利用(1) 切比雪夫不等式;(2) 中心极限定理分别确定投掷一枚均匀硬币的次数,使得出现"正面向上"的频率在 0.4 到 0.6 之间的概率不小于 0.9.

**解** 设 $X$ 表示投掷一枚均匀硬币 $n$ 次"正面向上"的次数,则 $X \sim B(n,0.5)$,$E(X)=np=0.5n$,$D(X)=np(1-p)=0.25n$.

(1) $P\left\{0.4<\frac{X}{n}<0.6\right\}=P\{0.4n<X<0.6n\}$
$$=P\{-0.1n<X-0.5n<0.1n\}$$
$$=P\{|X-0.5n|<0.1n\}$$
$$\geqslant 1-\frac{D(X)}{(0.1n)^2}=1-\frac{0.25n}{0.01n^2}=1-\frac{25}{n}\geqslant 0.9,$$

由此得 $$\frac{25}{n}\leqslant 0.1, \quad n\geqslant 250.$$

(2) $P\left\{0.4<\frac{X}{n}<0.6\right\}=P\{0.4n<X<0.6n\}$
$$=\Phi\left(\frac{0.6n-0.5n}{\sqrt{0.25n}}\right)-\Phi\left(\frac{0.4n-0.5n}{\sqrt{0.25n}}\right)$$
$$=\Phi(0.2\sqrt{n})-\Phi(-0.2\sqrt{n})$$
$$=2\Phi(0.2\sqrt{n})-1\geqslant 0.9,$$

由此得 $$\Phi(0.2\sqrt{n})\geqslant 0.95.$$

查表得 $0.2\sqrt{n}\geqslant 1.645$,$n\geqslant 67.65$,取 $n=68$.

6. 某保险公司多年的统计资料表明,在索赔户中被盗索赔户占 20%,以 $X$ 表示在随意抽查的 100 个索赔户中因被盗向保险公司索赔的户数

(1) 写出 $X$ 的概率分布;

(2) 利用棣莫弗-拉普拉斯定理，求被盗索赔户不少于 14 户且不多于 30 户的概率的近似值.

[附表] 设 $\Phi(x)$是标准正态分布函数

| $x$ | 0 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 |
|---|---|---|---|---|---|---|---|
| $\Phi(x)$ | 0.500 | 0.692 | 0.841 | 0.933 | 0.977 | 0.994 | 0.999 |

**解** (1) $X$ 服从二项分布，参数 $n=100, p=0.2$,

$$P\{X=k\}=C_{100}^{k}0.2^{k}0.8^{100-k},\quad k=0,1,\cdots,100.$$

(2) $E(X)=np=20, D(X)=np(1-p)=16$. 根据棣莫弗-拉普拉斯定理，

$$\begin{aligned}P\{14\leqslant X\leqslant 30\}&=P\left\{\frac{14-20}{4}\leqslant\frac{X-20}{4}\leqslant\frac{30-20}{4}\right\}\\&=P\left\{-1.5\leqslant\frac{X-20}{4}\leqslant 2.5\right\}\\&\approx\Phi(2.5)-\Phi(-1.5)\\&=\Phi(2.5)-[1-\Phi(1.5)]\\&=\Phi(2.5)+\Phi(1.5)-1\\&=0.994+0.933-1=0.927.\end{aligned}$$

7. 设有 1000 人独立行动，每个人能够按时进入掩蔽体的概率为 0.9. 以 95%概率估计，在一次行动中：(1) 至少有多少人能够进入掩蔽体；(2) 至多有多少人能进入掩蔽体.

**解** 用 $X_i$ 表示第 $i$ 人能够按时进入掩蔽体($i=1,2,\cdots,1000$)，令

$$S_n=X_1+X_2+\cdots+X_{1000}.$$

(1) 设至少有 $m$ 人能进入掩蔽体，要求 $P\{m\leqslant S_n\leqslant 1000\}\geqslant 0.95$. 事件

$$\{m\leqslant S_n\}=\left\{\frac{m-1000\times 0.9}{\sqrt{1000\times 0.9\times 0.1}}\leqslant\frac{S_n-900}{\sqrt{90}}\right\}.$$

令 $\dfrac{S_n-900}{\sqrt{90}}=Y$，显然 $Y\dot{\sim}N(0,1)$. 令 $\dfrac{m-900}{\sqrt{90}}=a$，根据中心极限

定理，有

$$P\{m \leqslant S_n\} = P\{Y \geqslant a\} = 1 - P\{Y < a\}$$
$$= 1 - \Phi(a) = 0.95.$$

查正态分布数值表，得 $a = -1.65$，即 $\dfrac{m-900}{\sqrt{90}} = -1.65$. 故 $m = 900 - 15.65 = 884.35 \approx 884$ 人.

(2) 用同样方法，可求出至多有 916 人进入掩蔽体

8. (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 0.10，又知为使系统正常运行，至少必须有 85 个元件工作，求系统的可靠度(即正常运行的概率)；(2) 上述系统假如有 $n$ 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运行，问 $n$ 至少为多大时才能保证系统的可靠度为 0.95?

**解** (1) 设

$$X_k = \begin{cases} 1, & \text{第 } k \text{ 个元件正常工作}, \\ 0, & \text{第 } k \text{ 个元件损坏}, \end{cases}$$

又设 $X$ 为系统正常运行时完好的元件数，于是 $X = \sum_{i=1}^{100} X_i$. 由题设可知 $X_k (k=1,2,\cdots,100)$ 服从 0-1 分布，

$$X = \sum_{i=1}^{100} X_i \sim B(100, 0.9),$$

于是

$$E(X) = np = 100 \times 0.9 = 90,$$
$$D(X) = npq = 100 \times 0.9 \times 0.1 = 9.$$

故所求概率为

$$P\{X > 85\} = 1 - P\{X \leqslant 85\} = 1 - \Phi\left(\frac{85-90}{\sqrt{9}}\right)$$
$$= 1 - \Phi\left(-\frac{5}{3}\right) = \Phi\left(\frac{5}{3}\right) = \Phi(1.67) = 0.9525.$$

(2) $P\{X \geqslant 0.8n\} = 0.95$，$E(X) = 0.9n$，$D(X) = 0.09n$，

而

$$P\{X \geqslant 0.8n\} = 1 - \Phi\left(\frac{0.8n - 0.9n}{\sqrt{0.09n}}\right)$$
$$= 1 - \Phi\left(\frac{-0.1n}{0.3\sqrt{n}}\right) = \Phi\left(\frac{\sqrt{n}}{3}\right),$$

故 $\Phi\left(\frac{\sqrt{n}}{3}\right)=0.95$，查表得 $\frac{\sqrt{n}}{3}=1.645$，$n=24.35$，取 $n=25$.

9. 计算机做加法运算时，要对每个加数取整(即取最接近它的整数)，设所有的取整误差是相互独立的，且它们都服从均匀分布 $U[-0.5,0.5]$，如果将 1500 个数相加，求误差总和的绝对值超过 15 的概率.

**解** 以 $X_k$ 表示第 $k$ 个数的取整误差($k=1,2,\cdots,1500$)，它们相互独立，且都服从 $U[-0.5,0.5]$，有

$$E(X_k) = 0, \quad D(X_k) = \frac{1}{12}.$$

由列维-林德伯格定理，$\sqrt{\frac{12}{1500}}\sum_{k=1}^{1500}X_k$ 渐近服从 $N(0,1)$ 分布，于是

$$P\left\{\left|\sum_{k=1}^{1500}X_k\right| > 15\right\} = 2P\left\{\sqrt{\frac{12}{1500}}\sum_{k=1}^{1500}X_k > \sqrt{\frac{12}{1500}} \times 15\right\}$$
$$\approx 2[1 - \Phi(1.34)]$$
$$= 2[1 - 0.9099] = 0.1802.$$

10. 试证当 $n\to\infty$ 时，$\mathrm{e}^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}\to\frac{1}{2}$.

**证** 设 $X_k(k=1,2,\cdots)$ 为服从具有参数 $\lambda=1$ 的泊松分布，并且 $\{X_k\}$ 相互独立，由列维-林德伯格定理知

$$P\left\{\frac{1}{\sqrt{n}}\sum_{k=1}^{n}(X_k - 1) \leqslant 0\right\} \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t = \frac{1}{2}(n \to \infty),$$

即 $P\left\{\sum_{k=1}^{n}X_k\leqslant n\right\}\to\frac{1}{2}\ (n\to\infty)$，而 $\sum_{k=1}^{n}X_k$ 服从具有参数 $n$ 的泊松分布，故有

$$P\left\{\sum_{k=1}^{n}X_k \leqslant n\right\} = \sum_{k=1}^{n}\frac{n^k e^{-n}}{k!} = e^{-n}\sum_{k=1}^{n}\frac{n^k}{k!},$$

于是
$$e^{-n}\sum_{k=1}^{n}\frac{n^k}{k!} \to \frac{1}{2} \quad (n \to \infty).$$

## 四、练习题

1. 现有一批种子，其中良种占$\frac{1}{6}$，今任取 6000 粒种子，试以 0.99 的概率推断，在这 6000 粒种子中良种所占的比例与$\frac{1}{6}$的差是多少？这时相应的良种粒数在哪个范围内？

2. 某单位有 200 台电话分机，每架分机有 5%的时间要使用外线通话. 假定每架分机是否使用外线是相互独立的，问该单位总机要安装多少外线，才能以 90%以上的概率保证分机使用外线时不等待.

3. 一部件包含 10 个部分，每部分的长度是个随机变量，它们相互独立且具有同一分布，其数学期望为 2(mm)，标准差为 0.05(mm)，规定部件总长度为 20±0.1(mm)时产品合格，试求产品合格的概率.

4. 某保险公司有 10000 人参加人身保险，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约多大？

(2) 保险公司每年的利润大于 40000 元的概率为多少？

5. 甲、乙两电影院在竞争 1000 名观众. 假定每个观众任选一个影院且观众间的选择彼此独立，问每个影院至少要设多少个座位，才能保证因缺少座位而使观众离去的概率小于 1%？

## 习题答案与提示

**1.** 0.0124；925 粒至 1075 粒之间.　　**2.** 14.

**3.** 0.4714.　　**4.** (1) 0；(2) 0.9952.　　**5.** 537 个.

# 第六章　数理统计的基本概念

## 一、考试要求

1. 理解总体与简单随机样本，样本函数与统计量的概念；

2. 了解 $\chi^2$ 分布、$t$ 分布、$F$ 分布的定义及性质，了解分位数(临界值)的概念并会查有关分布的临界值表；

3. 理解样本均值、样本方差及样本矩的概念，了解经验分布函数.

## 二、复习要点

### (一) 重要概念及性质

**定义 6.1(总体)**　在数理统计中，常把被考察对象的某一个(或多个)指标的全体称为**总体(或母体)**；而把总体中的每一个单元称为**样品(或个体)**. 在以后的讨论中，我们总是把总体看成一个具有分布的随机变量(或随机向量).

**定义 6.2(简单随机样本)**　我们把从总体中抽取的部分样品 $x_1, x_2, \cdots, x_n$ 称为**样本**. 样本中所含的样品数称为**样本容量**. 在一般情况下，总是把样本看成是 $n$ 个相互独立的且与总体有相同分布的随机变量①. 这样的样本称为**简单随机样本**(以后我们只讨论这种样本).

---

① 一般的随机变量用大写字母 $X$ 或 $Y$ 表示，此处样本 $x_1, x_2, \cdots, x_n$ 按通常写法应为 $X_1, X_2, \cdots, X_n$. 但是在一次抽取后样本 $X_1, X_2, \cdots, X_n$ 就是 $n$ 个具体的值 $x_1, x_2, \cdots, x_n$，称这 $n$ 个值为**样本值**. 为叙述简练，我们对样本与样本值所使用的符号不再加以区别，即我们赋予 $x_1, x_2, \cdots, x_n$ 有双重意义：在泛指任一次抽取的结果时，$x_1, x_2, \cdots, x_n$ 表示 $n$ 个**随机变量**(样本)；在具体的一次抽取之后，$x_1, x_2, \cdots, x_n$ 表示 $n$ 个具体的**数值**(样本值).

**定义 6.3（样本函数与统计量）** 设 $x_1,x_2,\cdots,x_n$ 为总体的一个样本，称

$$\varphi=\varphi(x_1,x_2,\cdots,x_n)$$

为样本函数，其中 $\varphi$ 为一个连续函数. 如果 $\varphi$ 中不包含任何未知参数，则称 $\varphi(x_1,x_2,\cdots,x_n)$ 为一个**统计量**.

**定义 6.4（样本均值）** 设 $x_1,x_2,\cdots,x_n$ 为总体的一个样本，称

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$$

为**样本均值**.

**定义 6.5（样本方差）** 设 $x_1,x_2,\cdots,x_n$ 为总体的一个样本，称

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

为**样本方差**.

**例 1** 设 $X_1,X_2,\cdots,X_{25}$ 相互独立且都服从 $N(3,10^2)$ 分布，求 $P\{0<\overline{X}<6,\ 57.70<S^2<151.73\}$.

**解** 因 $\overline{X}\sim N\left(3,\frac{10^2}{25}\right)$，$\frac{24S^2}{10^2}\sim\chi^2(24)$，且 $\overline{X}$ 与 $S^2$ 独立. 所以

$$\begin{aligned}&P\{0<\overline{X}<6,\ 57.70<S^2<151.73\}\\&\quad=P\{0<\overline{X}<6\}\cdot P\{57.70<S^2<151.73\}.\end{aligned}$$

而

$$\begin{aligned}P\{0<\overline{X}<6\}&=\Phi\left(\frac{6-3}{2}\right)-\Phi\left(\frac{-3}{2}\right)\\&=2\Phi\left(\frac{3}{2}\right)-1=2\Phi(1.5)-1\\&=2\times0.9332-1=0.8664,\end{aligned}$$

$$\begin{aligned}P\{57.70<S^2<151.73\}&=P\left\{13.848<\frac{24S^2}{100}<36.415\right\}\\&=P\left\{\frac{24S^2}{100}>13.848\right\}-P\left\{\frac{24S^2}{100}>36.415\right\}\\&=0.95-0.05=0.90,\end{aligned}$$

于是 $P\{0<\overline{X}<6,57.7<S^2<151.73\}=0.8664\times0.90\approx0.78$.

**定义 6.6(样本矩)** 设 $x_1,x_2,\cdots,x_n$ 为总体的一个样本，则对应的样本 **$k$ 阶原点矩**、**样本的 $k$ 阶中心矩**分别为

$$\hat{\nu}_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k, \quad \hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^k,$$

其中 $k=1,2,\cdots$. 此外还分别称样本的一阶原点矩

$$\bar{x} = \hat{\nu}_1 = \frac{1}{n}\sum_{i=1}^{n} x_i$$

及样本的二阶中心矩

$$\tilde{S}^2 = \hat{\mu}_2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \hat{\nu}_2 - (\hat{\nu}_1)^2$$

为**样本均值**与**样本方差**.

对于两个总体 $X$ 与 $Y$，类似地可通过其样本矩给出**样本协方差**与**样本相关系数**

$$S_{XY} = \hat{\mu}_{11} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}),$$

$$r_{XY} = \hat{\rho} = \frac{\hat{\mu}_{11}}{\sqrt{\hat{\mu}_{20}}\sqrt{\hat{\mu}_{02}}} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}.$$

**定义 6.7(经验分布函数)** 设总体 $X$ 的 $n$ 个样本值可以按大小次序排列成：

$$x_1 \leqslant x_2 \leqslant \cdots \leqslant x_n.$$

如果 $x_k \leqslant x < x_{k+1}$，则不大于 $x$ 的样本值的频率为 $\frac{k}{n}$. 因而函数

$$F_n(x) = \begin{cases} 0, & x < x_1, \\ \dfrac{k}{n}, & x_k \leqslant x < x_{k+1}, \\ 1, & x \geqslant x_n \end{cases}$$

$$(k = 1,2,\cdots,n-1)$$

与事件 $\{X \leqslant x\}$ 在 $n$ 次重复独立试验中的频率是相同的. 我们称 $F_n(x)$ 为**样本的分布函数**或**经验分布函数**.

（二）重要定理及公式

**定理 6.1**（正态分布） 设 $x_1,x_2,\cdots,x_n$ 为来自正态总体 $N(\mu,\sigma^2)$ 的一个样本，则样本函数

$$u \xlongequal{\text{def}} \frac{\bar{x}-\mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1).$$

**定理 6.2**（$t$-分布） 设 $x_1,x_2,\cdots,x_n$ 为来自正态总体 $N(\mu,\sigma^2)$ 的一个样本，则样本函数

$$t \xlongequal{\text{def}} \frac{\bar{x}-\mu}{\sqrt{\frac{S^2}{n}}} \sim t(n-1),$$

其中 $t(n-1)$ 表示自由度为 $n-1$ 的 $t$ 分布.

**定理 6.3**（$\chi^2$ 分布） 设 $x_1,x_2,\cdots,x_n$ 为来自正态总体 $N(\mu,\sigma^2)$ 的一个样本，则样本函数

$$w \xlongequal{\text{def}} \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1),$$

其中 $\chi^2(n-1)$ 表示自由度为 $n-1$ 的 $\chi^2$ 分布.

**定理 6.4**（$f$-分布） 设 $x_1,x_2,\cdots,x_{n_1}$ 为来自正态总体 $N(\mu_1,\sigma_1^2)$ 的一个样本，而 $y_1,y_2,\cdots,y_{n_2}$ 为来自正态总体 $N(\mu_2,\sigma_2^2)$ 的一个样本，则样本函数

$$F \xlongequal{\text{def}} \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim f(n_1-1,n_2-1),$$

其中 $S_1^2=\dfrac{1}{n_1-1}\sum\limits_{i=1}^{n_1}(x_i-\bar{x})^2$， $S_2^2=\dfrac{1}{n_2-1}\sum\limits_{i=1}^{n_2}(y_i-\bar{y})^2$；

$f(n_1-1,n_2-1)$ 表示第一自由度为 $n_1-1$，第二自由度为 $n_2-1$ 的 $f$ 分布.

**例 2** 设 $X_1,X_2,X_3,X_4$ 是来自正态总体 $N(0,2^2)$ 的简单随机样本，当 $a,b$ 为何值时，统计量 $X=a(X_1-2X_2)^2+b(3X_3-4X_4)^2$ 服从 $\chi^2$ 分布，其自由度是多少？

**解** 依题意 $X_1,X_2,X_3,X_4$ 独立同分布，

$$E(X_1-2X_2)=0,$$
$$D(X_1-2X_2)=D(X_1)+4D(X_2)=20,$$
$$E(3X_3-4X_4)=0,$$
$$D(3X_3-4X_4)=9D(X_3)+16D(X_4)=100.$$

从而

$$\frac{X_1-2X_2}{\sqrt{20}}\sim N(0,1),\quad \frac{3X_3-4X_4}{10}\sim N(0,1),$$

又 $X_1-2X_2$ 与 $3X_3-4X_4$ 独立，于是

$$\frac{(X_1-2X_2)^2}{20}+\frac{(3X_3-4X_4)^2}{100}\sim\chi^2(2),$$

所以，当 $a=\frac{1}{20},b=\frac{1}{100}$时，统计量 $X$ 服从 $\chi^2$ 分布，其自由度为 2.

**例 3** 设 $X_1,X_2,\cdots,X_n,X_{n+1},\cdots,X_{n+m}$是分布为 $N(0,\sigma^2)$的正态总体容量为 $n+m$ 的样本，试求下列统计量的概率分布：

$$(1)\ Y_1=\frac{\sqrt{m}\sum\limits_{i=1}^{n}X_1}{\sqrt{n}\sqrt{\sum\limits_{i=n+1}^{n+m}X_i^2}};\qquad (2)\ Y_2=\frac{m\sum\limits_{i=1}^{n}X_i^2}{n\sum\limits_{i=n+1}^{n+m}X_i^2}.$$

**解** (1) 因为 $X_1,\cdots,X_n,X_{n+1},\cdots,X_{n+m}$独立同分布，

$$X_i\sim N(0,\sigma^2),\quad \frac{X_i}{\sigma}\sim N(0,1),\quad i=1,2,\cdots,n+m,$$

所以
$$\sum_{i=1}^{n}X_i\sim N(0,n\sigma^2),\quad \frac{\sum\limits_{i=1}^{n}X_i}{\sqrt{n}\,\sigma}\sim N(0,1),$$

$$\frac{\sum\limits_{i=n+1}^{n+m}X_i^2}{\sigma^2}=\sum_{i=n+1}^{n+m}\left(\frac{X_i}{\sigma}\right)^2\sim\chi^2(m),$$

且$\sum\limits_{i=1}^{n}X_i$ 与 $\sum\limits_{i=n+1}^{n+m}X_i^2$ 相互独立. 由 $t$ 分布的定义知

$$Y_1 = \frac{\sqrt{m}\sum\limits_{i=1}^{n} X_i}{\sqrt{n}\sqrt{\sum\limits_{i=n+1}^{n+m} X_i^2}} = \frac{\sum\limits_{i=1}^{n} X_i / \sqrt{n}\,\sigma}{\sqrt{\dfrac{\sum\limits_{i=n+1}^{n+m} X_i^2}{\sigma^2} \Big/ m}} \sim t(m).$$

(2) 因 $\dfrac{\sum\limits_{i=1}^{n} X_i^2}{\sigma^2} = \sum\limits_{i=1}^{n}\left(\dfrac{X_i}{\sigma}\right)^2 \sim \chi^2(n)$，$\dfrac{\sum\limits_{i=n+1}^{n+m} X_i^2}{\sigma^2} = \sum\limits_{i=n+1}^{n+m}\left(\dfrac{X_i}{\sigma}\right)^2 \sim \chi^2(m)$，且两者独立，于是由 $F$ 分布之定义知

$$Y_2 = \frac{m\sum\limits_{i=1}^{n} X_i^2}{n\sum\limits_{i=n+1}^{n+m} X_i^2} = \frac{\dfrac{\sum\limits_{i=1}^{n} X_i^2}{\sigma^2} \Big/ n}{\dfrac{\sum\limits_{i=n+1}^{n+m} X_i^2}{\sigma^2} \Big/ m} \sim F(n,m).$$

## 三、典型例题分析

### (一) 填空题

1. 设随机变量 $X$ 和 $Y$ 相互独立且都服从正态分布 $N(0,3^2)$. $X_1,\cdots,X_9$ 和 $Y_1,\cdots,Y_9$ 分别是来自总体 $X$ 和 $Y$ 的随机样本，则统计量 $U=\dfrac{X_1+\cdots+X_9}{\sqrt{Y_1^2+\cdots+Y_9^2}}$ 服从________分布，参数为________.

答案是：$t$，9.

**分析** 由于 $\dfrac{Y_i}{3} \sim N(0,1)$，故

$$Y = \sum_{i=1}^{9}\left(\frac{Y_i}{3}\right)^2 = \frac{1}{9}\sum_{i=1}^{9} Y_i^2 \sim \chi^2(9).$$

再由 $\overline{X}=\dfrac{1}{9}\sum\limits_{i=1}^{9} X_i \sim N(0,1)$，据 $t$ 分布的定义，有

$$U = \frac{\overline{X}}{\sqrt{Y/9}} \sim t(9).$$

2. 设 $X_1, X_2, X_3, X_4$ 是来自正态总体 $N(0, 2^2)$ 的简单随机样本，$X = a(X_1 - 2X_2)^2 + b(3X_3 - 4X_4)^2$，则当 $a=$ ________，$b=$ ________ 时，统计量 $X$ 服从 $\chi^2$ 分布，其自由度为________.

答案是：$\frac{1}{20}, \frac{1}{100}, 2$.

**分析** 根据 $\chi^2$ 分布的性质：若 $X_i \sim N(0,1)$，则 $X_1^2 + \cdots + X_n^2 \sim \chi^2(n)$，现在要求是 $\sqrt{a}(X_1 - 2X_2) \sim N(0,1)$ 及 $\sqrt{b}(3X_3 - 4X_4) \sim N(0,1)$，由

$$\sqrt{a}(X_1 - 2X_2) = \sqrt{a}X_1 - 2\sqrt{a}X_2$$
$$\sim N(0, (\sqrt{a} \cdot 2)^2 + (2\sqrt{a} \cdot 2)^2) = N(0, 20a)$$

知 $20a = 1$，即 $a = \frac{1}{20}$. 由

$$\sqrt{b}(3X_3 - 4X_4) \sim N(0, (3\sqrt{b} \cdot 2)^2 + (4\sqrt{b} \cdot)^2)$$
$$= N(0, 100b)$$

知 $100b = 1$，即 $b = \frac{1}{100}$，显然这时自由度 $n = 2$.

**(二) 选择题**

1. 设 $n$ 个随机变量 $X_1, X_2, \cdots, X_n$ 独立同分布，

$$D(X_1) = \sigma^2,\ \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,\ S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2,$$

则(    ).

(A) $S$ 是 $\sigma$ 的无偏估计量；(B) $S$ 是 $\sigma$ 的最大似然估计量；

(C) $S$ 是 $\sigma$ 的相合估计量(即一致估计量)；

(D) $S$ 与 $\overline{X}$ 相互独立.

答案是：C.

**分析** 这是一个判断型题目，一般来说不用进行任何计算. 我们知道，总体 $X$ 无论是什么分布，只要 $E(X)$，$D(X)$ 存在，$S^2$ 是 $\sigma^2$ 的无偏估计量，即 $E(S^2) = \sigma^2$，而 $S$ 不是 $\sigma$ 的无偏估计量. 因此，(A)是不成立的.

设 $X \sim N(\mu, \sigma^2)$，令

$$\tilde{S}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2, \quad \tilde{S} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

可见 $\tilde{S}^2$ 是 $\sigma^2$ 的最大似然估计量，$\tilde{S}$ 也是 $\sigma$ 的最大似然估计量，但 $S^2$ 和 $S$ 不是，而本题又要求对任意总体. 因此，(B)是不成立的.

对于任何总体，只要 $E(X)$，$D(X)$存在，一切样本矩和样本矩的连续函数都是相应总体数的相合估计量. 因此，$S^2$ 和 $S$ 是 $\sigma^2$ 和 $\sigma$ 的相合估计量. 本题应选择 C.

由于只有正态总体 $S^2$ 与 $\overline{X}$ 才相互独立，但题目中是任意总体，因此，$S$ 与 $\overline{X}$ 相互独立不成立.

2. 设 $X_1, X_2, \cdots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$的简单随机样本，$\overline{X}$ 是样本均值，记

$$S_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2, \quad S_2^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2,$$

$$S_3^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \mu)^2, \quad S_4^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2,$$

则服从自由度为 $n-1$ 的 $t$ 分布的随机变量是(　　).

(A) $t = \dfrac{\overline{X}-\mu}{S_1/\sqrt{n-1}}$；　　(B) $t = \dfrac{\overline{X}-\mu}{S_2/\sqrt{n-1}}$；

(C) $t = \dfrac{\overline{X}-\mu}{S_3/\sqrt{n}}$；　　(D) $t = \dfrac{\overline{X}-\mu}{S_4/\sqrt{n}}$.

答案是：B.

(三) 解答题

1. 假设随机变量 $X_1, X_2, \cdots, X_n$ 是来自总体 $X$ 的简单随机样本；已知 $E(X^k) = \alpha_k (k=1,2,3,4)$. 证明当 $n$ 充分大时，随机变量 $Z_n = \dfrac{1}{n}\sum_{i=1}^{n} X_i^2$ 近似服从正态分布，并指出其分布参数.

证　依题意 $X_1, X_2, \cdots, X_n$ 独立同分布，可知 $X_1^2, X_2^2, \cdots, X_n^2$ 也独立同分布且有

$$E(X_i^2) = \alpha_2, \quad D(X_i^2) = E(X_i^4) - (EX_i^2)^2 = \alpha_4 - \alpha_2^2.$$

由中心极限定理

$$V_n=\frac{\sum_{i=1}^{n}X_i^2-n\alpha_2}{\sqrt{n(\alpha_4-\alpha_2^2)}}=\frac{\frac{1}{n}\sum_{i=1}^{n}X_i^2-\alpha_2}{\sqrt{(\alpha_4-\alpha_2^2)/n}}=\frac{Z_n-\alpha_2}{\sqrt{(\alpha_4-\alpha_2^2)/n}}$$

的极限分布是标准正态分布，所以当 $n$ 充分大时 $V_n$ 近似服从标准正态分布，从而 $Z_n=\sqrt{\frac{\alpha_4-\alpha_2^2}{n}}\cdot V_n+\alpha_2$ 近似服从参数为 $\mu=\alpha_2,\sigma^2=\frac{\alpha_4-\alpha_2^2}{n}$ 的正态分布.

2. 设 $X_1,X_2,\cdots,X_9$ 是来自正态总体 $X$ 的简单随机样本，

$$Y_1=\frac{1}{6}(X_1+\cdots+X_6),\quad Y_2=\frac{1}{3}(X_7+X_8+X_9),$$

$$S^2=\frac{1}{2}\sum_{i=7}^{9}(X_i-Y_2)^2,\qquad Z=\frac{\sqrt{2}\,(Y_1-Y_2)}{S},$$

证明统计量 $Z$ 服从自由度为 2 的 $t$ 分布.

**证** 记 $D(X)=\sigma^2$(未知). 易见 $E(Y_1)=E(Y_2)$，$D(Y_1)=\sigma^2/6$，$D(Y_2)=\sigma^2/3$. 由于 $Y_1$ 和 $Y_2$ 独立，可见 $E(Y_1-Y_2)=0$，

$$D(Y_1-Y_2)=\frac{\sigma^2}{6}+\frac{\sigma^2}{3}=\frac{\sigma^2}{2}.$$

从而
$$U=\frac{Y_1-Y_2}{\sigma/\sqrt{2}}\sim N(0,1).$$

由正态总体样本方差的性质，知

$$\chi^2=\frac{2S^2}{\sigma^2}$$

服从自由度为 2 的 $\chi^2$ 分布.

由于 $Y_1$ 与 $Y_2$，$Y_1$ 与 $S^2$ 独立，以及 $Y_2$ 与 $S^2$ 独立，可见 $Y_1-Y_2$ 与 $S^2$ 独立.

于是，由服从 $t$ 分布随机变量的结构，知

$$Z=\frac{\sqrt{2}\,(Y_1-Y_2)}{S}=\frac{U}{\sqrt{\chi^2/2}}$$

服从自由度为 2 的 $t$ 分布.

3. 设总体 $X$ 服从正态分布,$\bar{x}$ 与 $S^2$ 分别为样本均值和样本方差,又设 $x_{n+1}\sim N(\mu,\sigma^2)$ 且 $x_{n+1}$ 与 $x_1,x_2,\cdots,x_n$ 相互独立,求统计量

$$T=\frac{x_{n+1}-\bar{x}}{S}\sqrt{\frac{n}{n+1}}$$

的分布.

**解** 因为 $x_{n+1}\sim N(\mu,\sigma^2)$, $\bar{x}\sim N\left(\mu,\frac{\sigma^2}{n}\right)$,所以

$$x_{n+1}-\bar{x}\sim N\left(0,\frac{n+1}{n}\sigma^2\right).$$

因此

$$\frac{x_{n+1}-\bar{x}}{\sigma\sqrt{\frac{n}{n+1}}}\sim N(0,1).$$

又因 $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$,且 $S^2$ 与 $x_{n+1}-\bar{x}$ 独立,故

$$T=\frac{\dfrac{x_{n+1}-\bar{x}}{\sigma\sqrt{\frac{n}{n+1}}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2(n-1)}}}=\frac{x_{n+1}-\bar{x}}{S}\sqrt{\frac{n}{n+1}}\sim t(n-1),$$

即 $T$ 服从 $n-1$ 个自由度的 $t$ 分布.

4. 设 $X_1,\cdots,X_{10};Y_1,\cdots,Y_{15}$ 相互独立且都为 $N(20,(\sqrt{3})^2)$ 分布,求 $P\{|\bar{X}-\bar{Y}|>0.3\}$.

**解** 因为 $X_1,\cdots,X_{10};Y_1,\cdots,Y_{15}$ 独立且都服从 $N(20,3)$ 分布,所以

$$\bar{X}=\frac{1}{10}\sum_{i=1}^{10}X_i\sim N\left(20,\frac{3}{10}\right),\quad \bar{Y}=\frac{1}{15}\sum_{i=1}^{15}Y_i\sim N\left(20,\frac{3}{15}\right),$$

且两者独立,于是

$$\bar{X}-\bar{Y}\sim N\left(0,\frac{1}{2}\right),$$

$$P\{|\overline{X}-\overline{Y}|>0.3\}=P\left\{\frac{|\overline{X}-\overline{Y}|}{\sqrt{1/2}}>\frac{0.3}{\sqrt{1/2}}\right\}$$
$$=P\left\{\frac{|\overline{X}-\overline{Y}|}{\sqrt{1/2}}>0.3\sqrt{2}\right\}$$
$$=1-P\left\{\frac{|\overline{X}-\overline{Y}|}{\sqrt{1/2}}<0.42\right\}$$
$$=2[1-\Phi(0.42)]$$
$$=2[1-0.6628]=0.6744.$$

5. 若随机变量 $X$ 具有自由度为 $n_1,n_2$ 的 $F$ 分布，求证

(1) $Y=\frac{1}{X}$具有自由度为 $n_2,n_1$ 的 $F$ 分布；

(2) 并由此证明 $F_{1-\alpha}(n_1,n_2)=\frac{1}{F_\alpha(n_2,n_1)}$.

**证** (1) 因 $X\sim F(n_1,n_2)$，故可设 $X=\frac{U/n_1}{V/n_2}$，其中 $U\sim\chi^2(n_1)$，$V\sim\chi^2(n_2)$，且 $U$ 与 $V$ 相互独立，于是$\frac{1}{X}=\frac{V/n_2}{U/n_1}$，由 $F$ 分布之定义知 $Y=\frac{1}{X}\sim F(n_2,n_1)$.

(2) 由上侧 $\alpha$ 分位数的定义知：

$$P\{X\geqslant F_{1-\alpha}(n_1,n_2)\}=1-\alpha,$$
$$P\left\{\frac{1}{X}\leqslant\frac{1}{F_{1-\alpha}(n_1,n_2)}\right\}=1-\alpha,$$

即

$$P\left\{Y\leqslant\frac{1}{F_{1-\alpha}(n_1,n_2)}\right\}=1-\alpha,$$
$$1-P\left\{Y>\frac{1}{F_{1-\alpha}(n_1,n_2)}\right\}=1-\alpha,$$
$$P\left\{Y>\frac{1}{F_{1-\alpha}(n_1,n_2)}\right\}=\alpha,$$

而 $P\{Y\geqslant F_\alpha(n_2,n_1)\}=\alpha$. 又 $Y$ 为连续型随机变量，故

$$P\left\{Y\geqslant\frac{1}{F_{1-\alpha}(n_1,n_2)}\right\}=\alpha.$$

从而 $F_\alpha(n_2,n_1)=\dfrac{1}{F_{1-\alpha}(n_1,n_2)}$，即 $F_{1-\alpha}(n_1,n_2)=\dfrac{1}{F_\alpha(n_2,n_1)}$.

6. 求证：若 $X\sim t(n)$，则 $X^2\sim F(1,n)$.

**证** 因 $X\sim t(n)$，由 $t$ 分布之定义，故可设 $X=\dfrac{U}{\sqrt{V/n}}$，其中 $U\sim N(0,1)$，$V\sim\chi^2(n)$，且 $U$ 与 $V$ 独立，于是 $U^2\sim\chi^2(1)$，$X^2=\dfrac{U^2}{V/n}=\dfrac{U^2/1}{V/n}$，由 $F$ 分布之定义知 $X^2\sim F(1,n)$.

## 四、练习题

1. 设 $X_1,X_2,\cdots,X_n$ 是分布为 $N(\mu,\sigma^2)$ 的正态总体的一个样本，求 $Y=\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i-\mu)^2$ 的概率分布.

2. 设总体 $X$ 具有正态分布 $N(0,1)$，从中取一容量为 6 的样本 $(X_1,X_2,X_3,X_4,X_5,X_6)$. 又设 $Y=(X_1+X_2+X_3)^2+(X_4+X_5+X_6)^2$. 试决定常数 $C$，使得随机变量 $CY$ 服从 $\chi^2$ 分布.

3. 设 $X_1,\cdots,X_m$ 和 $Y_1,\cdots,Y_n$ 分别是从 $N(\mu_1,\sigma^2)$ 和 $N(\mu_2,\sigma^2)$ 的总体中抽取的独立随机样本，$\overline{X}$ 和 $\overline{Y}$ 分别表示 $X$ 和 $Y$ 的样本均值，$S_{1m}^2$ 和 $S_{2n}^2$ 分别表示 $X$ 和 $Y$ 的样本方差. $\alpha$ 和 $\beta$ 是两个固定的实数. 试求

$$\frac{\alpha(\overline{X}-\mu_1)+\beta(\overline{Y}-\mu_2)}{\sqrt{\dfrac{(m-1)S_{1m}^2+(n-1)S_{2n}^2}{m+n-2}}\sqrt{\dfrac{\alpha^2}{m}+\dfrac{\beta^2}{n}}}$$

的概率分布.

### 习题答案与提示

**1.** $\chi^2(n)$.  **2.** $\dfrac{1}{3}$.  **3.** $t(m+n-2)$.

# 第七章　参数估计

## 一、考试要求

1. 理解参数点估计、估计量与估计值的概念；

2. 掌握参数矩估计法(一阶、二阶矩)和最大似然估计法；

3. 了解估计量的优良性；并会验证估计量的无偏性；

4. 了解区间估计的概念，会求单正态总体均值与方差的置信区间及双正态总体均值差与方差比的置信区间.

## 二、复习要点

### (一) 重要概念及性质

**定义 7.1(估计量)**　设总体 $X$ 的分布函数 $F(x;\theta)$ 的形式已知，其中 $\theta$ 为一个未知参数，又设 $x_1,x_2,\cdots,x_n$ 为总体 $X$ 的一个样本. 我们构造一个统计量 $K=K(x_1,x_2,\cdots,x_n)$ 作为参数 $\theta$ 的估计，称统计量 $K$ 为参数 $\theta$ 的一个**估计量**. 当 $x_1,x_2,\cdots,x_n$ 为一组样本值时，则 $\hat{K}=K(x_1,x_2,\cdots,x_n)$ 就是 $\theta$ 的一个点估计值.

**定义 7.2(无偏性)**　设 $\hat{\theta}=\hat{\theta}(x_1,x_2,\cdots,x_n)$ 为未知参数 $\theta$ 的估计量. 若 $E(\hat{\theta})=\theta(\forall\theta\in\Theta)$，则称 $\hat{\theta}$ 为 $\theta$ 的**无偏估计量**.

**定义 7.3(有效性)**　设 $\hat{\theta}_1=\hat{\theta}_1(x_1,x_2,\cdots,x_n)$ 和 $\hat{\theta}_2=\hat{\theta}_2(x_1,x_2,\cdots,x_n)$ 是未知参数 $\theta$ 的两个无偏估计量. 若 $D(\hat{\theta}_1)<D(\hat{\theta}_2)(\forall\theta\in\Theta)$，则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ **有效**.

**例 1**　设 $x_1,x_2,\cdots,x_n$ 是总体的一个样本. 试证

(1) $\hat{\mu}_1=\frac{1}{5}x_1+\frac{3}{10}x_2+\frac{1}{2}x_3$；

(2) $\hat{\mu}_2=\frac{1}{3}x_1+\frac{1}{4}x_2+\frac{5}{12}x_3$;

(3) $\hat{\mu}_3=\frac{1}{3}x_1+\frac{3}{4}x_2-\frac{1}{12}x_3$

都是总体均值 $\mu$ 的无偏估计,并比较哪一个最有效.

**证** 因为 $E(x_i)=\mu \quad (i=1,2,\cdots,n)$,所以

$$E(\hat{\mu}_1)=\frac{1}{5}E(x_1)+\frac{3}{10}E(x_2)+\frac{1}{2}E(x_3)$$
$$=\left(\frac{1}{5}+\frac{3}{10}+\frac{1}{2}\right)\mu=\mu.$$

同理 $E(\hat{\mu}_2)=\mu$, $E(\hat{\mu}_3)=\mu$. 即$\hat{\mu}_1,\hat{\mu}_2,\hat{\mu}_3$ 都是总体均值 $\mu$ 的无偏估计. 又因为

$$D(\hat{\mu}_1)=\frac{1}{25}D(x_1)+\frac{9}{100}D(x_2)+\frac{1}{4}D(x_3)$$
$$=\frac{38}{100}D(X)=\frac{19}{50}D(X),$$

同理

$$D(\hat{\mu}_2)=\frac{50}{144}D(X)=\frac{25}{72}D(X),$$
$$D(\hat{\mu}_3)=\frac{98}{144}D(X)=\frac{49}{72}D(X).$$

由于 $D(\hat{\mu}_2)<D(\hat{\mu}_1)$, $D(\hat{\mu}_2)<D(\hat{\mu}_3)$,故$\hat{\mu}_2$ 最有效.

**定义 7.4(相合性)** 设$\hat{\theta}_n$ 是 $\theta$ 的一串估计量,如果对于 意的正数 $\varepsilon$,都有

$$\lim_{n\to\infty}P\{|\hat{\theta}_n-\theta|>\varepsilon\}=0,$$

则称$\hat{\theta}_n$ 为 $\theta$ 的**相合估计量(或一致估计量)**.

**定义 7.5(置信区间与置信度)** 设总体 $X$ 含有一个待估的未知参数 $\theta$. 如果我们从样本 $x_1,x_2,\cdots,x_n$ 出发,找出两个统计量 $\theta_1=\theta_1(x_1,x_2,\cdots,x_n)$与 $\theta_2=\theta_2(x_1,x_2,\cdots,x_n)$ $(\theta_1<\theta_2)$,使得区间$[\theta_1,\theta_2]$以 $1-\alpha(0<\alpha<1)$的概率包含这个待估参数 $\theta$,即

$$P\{\theta_1\leqslant\theta\leqslant\theta_2\}=1-\alpha,$$

那么称区间$[\theta_1,\theta_2]$为 $\theta$ 的**置信区间**,$1-\alpha$ 为该区间的**置信度**(或

置信水平).

（二）重要定理及公式

**定理 7.1** 设总体 $X$ 的均值 $E(X)$ 存在，则其样本均值 $\bar{x}$ 是 $E(X)$ 的一个无偏估计量，即 $E(\bar{x})=E(X)$.

**证** 由于样本中每一个 $x_i$ 都与总体 $X$ 具有相同的分布，因此有 $E(x_i)=E(X)$ $(i=1,2,\cdots,n)$. 于是

$$E(\bar{x})=E\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(x_i)$$

$$=\frac{1}{n}\sum_{i=1}^{n}E(X)=\frac{1}{n}\cdot nE(X)=E(X).$$

**定理 7.2** 设总体 $X$ 的均值 $E(X)$、方差 $D(X)$ 都存在，则其样本均值 $\bar{x}$ 的方差存在，并且

$$D(\bar{x})=\frac{1}{n}D(X).$$

**证** $$D(\bar{x})=\frac{1}{n^2}(D(X_1)+D(X_2)+\cdots+D(X_n))$$

$$=\frac{1}{n}D(X).$$

**定理 7.3** 设总体 $X$ 的均值 $E(X)$ 与方差 $D(X)$ 都存在，则对于其样本方差 $\tilde{S}^2$，有

$$E(\tilde{S}^2)=\frac{n-1}{n}D(X).$$

**证** 考虑到

$$D(\bar{x})=D\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(x_i)=\frac{1}{n^2}\cdot nD(X)=\frac{1}{n}D(X),$$

$$\sum_{i=1}^{n}(x_i-\bar{x})^2=\sum_{i=1}^{n}(x_i^2-2x_i\bar{x}+\bar{x}^2)=\sum_{i=1}^{n}x_i^2-2\bar{x}\sum_{i=1}^{n}x_i+n\bar{x}^2$$

$$=\sum_{i=1}^{n}x_i^2-2\bar{x}n\bar{x}+n\bar{x}^2=\sum_{i=1}^{n}x_i^2-n\bar{x}^2;$$

以及 $$E(X^2)=D(X)+(E(X))^2,$$

$$E(\bar{x}^2)=D(\bar{x})+(E(\bar{x}))^2.$$

于是，我们有

$$E(\tilde{S}^2)=E\left[\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]=\frac{1}{n}E\left(\sum_{i=1}^{n}x_i^2-n\bar{x}^2\right)$$

$$=\frac{1}{n}\sum_{i=1}^{n}E(x_i^2)-E(\bar{x}^2)=E(X^2)-E(\bar{x}^2)$$

$$=D(X)+(E(X))^2-[D(\bar{x})+(E(\bar{x}))^2]$$

$$=D(X)+(E(X))^2-\frac{1}{n}D(X)-(E(X))^2$$

$$=\frac{n-1}{n}D(X).$$

（三）重要方法

**方法 7.1（矩估计法）** 所谓**矩估计法**就是利用样本各阶原点矩与相应的总体矩，来建立估计量应满足的方程，从而求出未知参数估计量的方法．

设总体 $X$ 的分布中包含有未知参数 $\theta_1,\theta_2,\cdots,\theta_m$，则其分布函数可以表成 $F(x;\theta_1,\theta_2,\cdots,\theta_m)$．显然它的 $k$ 阶原点矩 $v_k=E(X^k)$ $(k=1,2,\cdots,m)$ 中也包含了未知参数 $\theta_1,\theta_2,\cdots,\theta_m$，即 $v_k=v_k(\theta_1,\theta_2,\cdots,\theta_m)$．又设 $x_1,x_2,\cdots,x_n$ 为总体 $X$ 的 $n$ 个样本值，其样本的 $k$ 阶原点矩为

$$\hat{v}_k=\frac{1}{n}\sum_{i=1}^{n}x_i^k \quad (k=1,2,\cdots,m).$$

这样，我们按照"当参数等于其估计量时，总体矩等于相应的样本矩"的原则建立方程，即有

$$\begin{cases} v_1(\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_m)=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i, \\ v_2(\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_m)=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i^2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ v_m(\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_m)=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i^m. \end{cases}$$

由上面的 $m$ 个方程中，解出的 $m$ 个未知参数 $(\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_m)$ 即为参数 $(\theta_1,\theta_2,\cdots,\theta_m)$ 的**矩估计量**.

**例 2** 设总体 $X\sim U(a,b)$，求对 $a,b$ 的矩估计量.

**解** 我们知道，$E(X)=\frac{1}{2}(a+b)$，$D(X)=\frac{1}{12}(b-a)^2$. 由方程组

$$\begin{cases}\bar{x}=\dfrac{1}{2}(a+b),\\ \tilde{S}^2=\dfrac{1}{12}(b-a)^2\end{cases}$$

解得
$$\begin{cases}\hat{a}=\bar{x}-\sqrt{3}\,\tilde{S},\\ \hat{b}=\bar{x}+\sqrt{3}\,\tilde{S},\end{cases}$$

其中
$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i,\quad \tilde{S}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}.$$

**例 3** 设总体 $X\sim N(\mu,\sigma^2)$，求对 $\mu,\sigma^2$ 的矩估计量.

**解** 我们知道，$E(X)=\mu$，$D(X)=\sigma^2$. 由矩估计法可直接得到

$$\begin{cases}\hat{\mu}=\bar{x},\\ \hat{\sigma}^2=\tilde{S}^2.\end{cases}$$

**例 4** 求总体均值 $\mu$ 与方差 $\sigma^2$ 的矩估计量.

**解** 设 $X_1,X_2,\cdots,X_n$ 是总体 $X$ 的样本，又设总体的二阶矩存在，则由

$$\mu_1=E(X)=\mu,\mu_2=E(X^2)=D(X)+(E(X))^2=\sigma^2+\mu^2,$$

令

$$\hat{\mu}_1=\hat{\mu}=A_1=\bar{X},\quad \hat{\mu}_2=\hat{\sigma}^2+\hat{\mu}^2=A_2=\frac{1}{n}\sum_{i=1}^{n}X_i^2,$$

解得

$$\hat{\mu}=\bar{X},$$

$$\hat{\sigma}^2=\hat{\mu}_2-\hat{\mu}^2=\frac{1}{n}\sum_{i=1}^{n}X_i^2-\bar{X}^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = S_n^2.$$

从所得结果看出，这两个矩估计量的表达式与总体分布无关.

**方法 7.2（最大似然估计法）** 设总体 $X$ 的分布密度为 $p(x;\theta_1,\theta_2,\cdots,\theta_m)$，其中 $\theta_1,\theta_2,\cdots,\theta_m$ 为未知参数. 又设 $x_1,x_2,\cdots,x_n$ 为总体 $X$ 的一个样本，称

$$L_n(\theta_1,\theta_2,\cdots,\theta_m) = \prod_{i=1}^{n} p(x_i;\theta_1,\theta_2,\cdots,\theta_m)$$

为样本的**似然函数**，简记为 $L_n$. 我们把使 $L_n$ 达到最大的 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_m$ 分别作为 $\theta_1,\theta_2,\cdots,\theta_m$ 的估计量的方法称为**最大似然估计法**.

由于 $\ln x$ 是一个递增函数，所以 $L_n$ 与 $\ln L_n$ 同时达到最大值. 我们称

$$\left.\frac{\partial \ln L_n}{\partial \theta_i}\right|_{\theta_i=\hat{\theta}_i} = 0, \quad i = 1,2,\cdots,m$$

为似然方程. 由多元微分学可知，由似然方程可以求出 $\hat{\theta}_i = \hat{\theta}_i(x_1, x_2,\cdots,x_n)(i=1,2,\cdots,m)$ 为 $\theta_i$ 的最大似然估计量.

容易看出，使得 $L_n$ 达到最大的 $\hat{\theta}_i$ 也可以使这组样本值出现的可能性最大.

**例 5** 设总体 $X\sim N(\mu,\sigma^2)$，求对 $\mu,\sigma^2$ 的最大似然估计量.

**解** 我们知道，$\mu$ 和 $\sigma^2$ 的似然函数为

$$L_n(\mu,\sigma^2) = \prod_{i=1}^{n}\left[\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}\right] = \frac{1}{(\sqrt{2\pi}\sigma)^n}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}.$$

似然方程为

$$\begin{cases} \left.\dfrac{\partial \ln L_n(\mu,\sigma^2)}{\partial \mu}\right|_{\substack{\mu=\hat{\mu}\\ \sigma^2=\hat{\sigma}^2}} = \dfrac{1}{\hat{\sigma}^2}\displaystyle\sum_{i=1}^{n}(x_i - \hat{\mu}) = 0, \\ \left.\dfrac{\partial \ln L_n(\mu,\sigma^2)}{\partial \sigma^2}\right|_{\substack{\mu=\hat{\mu}\\ \sigma^2=\hat{\sigma}^2}} = -\dfrac{n}{2\hat{\sigma}^2} + \dfrac{1}{2\hat{\sigma}^4}\displaystyle\sum_{i=1}^{n}(x_i - \hat{\mu})^2 = 0. \end{cases}$$

解得

$$\hat{\mu}=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{x},\quad \hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2=\tilde{S}^2.$$

可见，对于正态分布的参数 $\mu$ 和 $\sigma^2$ 来说，最大似然估计量与矩估计量完全相同. 但是对其他一些分布它们并不一样. 一般来说，用矩法估计参数较为方便，但当样本容量较大时，矩估计量的精度不如最大似然估计量高. 因此，最大似然法用得较为普遍.

**例 6** 设在 $n$ 次独立试验中事件 $A$ 发生 $k$ 次，求事件 $A$ 发生的概率 $p$ 的矩估计量与最大似然估计量.

**解** 令 $X$ 表示在一次随机试验中事件 $A$ 发生的次数，即

$$X=\begin{cases}1, & \text{若 } A \text{ 发生},\\ 0, & \text{若 } A \text{ 不发生},\end{cases}$$

则 $P(X=1)=P(A)=p$，$P(X=0)=P(\bar{A})=1-p$. 重复 $n$ 次试验，设 $X_i$ 为第 $i$ 次试验中事件 $A$ 发生的次数($i=1,\cdots,n$)，则 $k=\sum_{i=1}^{n}X_i$. 因为 $E(X)=p$，故得 $p$ 的矩估计量为

$$\hat{p}=\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i=\frac{k}{n}.$$

设$(x_1,x_2,\cdots,x_n)$是相应于样本$(X_1,X_2,\cdots,X_n)$的观察值，$X\sim B(1,p)$，其分布律为

$$P(X=x)=p^x(1-p)^{1-x},\quad x=1,2.$$

故似然函数为

$$L(p)=\prod_{i=1}^{n}p\{X=x_i\}=\prod_{i=1}^{n}p^{x_i}(1-p)^{1-x_i}$$

$$=p^{\sum x_i}(1-p)^{n-\sum x_i},$$

而

$$\ln L(p)=\sum x_i\ln p+(n-\sum x_i)\ln(1-p),$$

令

$$\frac{\mathrm{d}\ln L(p)}{\mathrm{d}p}=\frac{\sum x_i}{p}-\frac{n-\sum x_i}{1-p}=0,$$

解得 $p$ 的最大似然估计值为 $\hat{p}=\frac{1}{n}\sum x_i=\bar{x}$，$p$ 的最大似然估计量为 $\hat{p}=\frac{1}{n}\sum X_i=\bar{X}$.

**方法 7.3**（区间估计的查表方法）

1. 均值的区间估计

设 $x_1,x_2,\cdots,x_n$ 为来自总体 $X\sim N(\mu,\sigma^2)$ 的一个样本. 在置信度为 $1-\alpha$ 下，我们来确定 $\mu$ 的置信区间 $[\theta_1,\theta_2]$.

(1) 已知方差，估计均值

设方差 $\sigma^2=\sigma_0^2$，其中 $\sigma_0^2$ 为已知数. 我们知道 $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ 是 $\mu$ 的一个点估计，并且知道包含未知参数 $\mu$ 的样本函数

$$u=\frac{\bar{x}-\mu}{\sigma_0/\sqrt{n}}$$

是服从标准正态分布的. 因此，对于给定的置信度 $1-\alpha$，查正态分布数值表（见附表 1），找出两个临界值 $\lambda_1$ 与 $\lambda_2$，使得

$$P\{\lambda_1\leqslant u\leqslant\lambda_2\}=1-\alpha.$$

满足上式的临界值 $\lambda_1,\lambda_2$ 由表中可以找出无穷多组. 不过一般我们总是取成对称区间 $[-\lambda,\lambda]$，使

$$P\{|u|\leqslant\lambda\}=1-\alpha,$$

即

$$P\left\{-\lambda\leqslant\frac{\bar{x}-\mu}{\sigma_0/\sqrt{n}}\leqslant\lambda\right\}=1-\alpha.$$

将本书所附的正态分布数值表的构造

$$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt\quad(x\geqslant 0)$$

（即图 7.1 中的阴影部分）与 $P\{|u|\leqslant\lambda\}=1-\alpha$（即图 7.2 中的阴影部分）比较，不难看出，确定 $\lambda$ 之值的方法是查 $\Phi(\lambda)=1-\frac{\alpha}{2}$. 找出 $\lambda$ 的值以后把它代入不等式

$$-\lambda \leqslant \frac{(\bar{x}-\mu)\sqrt{n}}{\sigma_0} \leqslant \lambda,$$

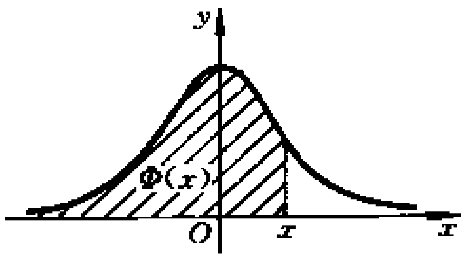



图 7.1

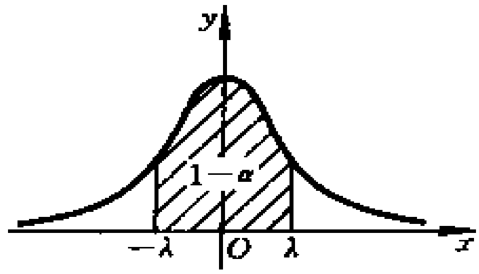



图 7.2

推得
$$\bar{x}-\lambda \frac{\sigma_0}{\sqrt{n}} \leqslant \mu \leqslant \bar{x}+\lambda \frac{\sigma_0}{\sqrt{n}}.$$
这就是说，随机区间
$$\left[\bar{x}-\lambda \frac{\sigma_0}{\sqrt{n}}, \bar{x}+\lambda \frac{\sigma_0}{\sqrt{n}}\right] \tag{7.1}$$
以 $1-\alpha$ 的概率包含 $\mu$.

**例 7** 已知幼儿的身高在正常情况下服从正态分布.现从某一幼儿园 5 岁至 6 岁的幼儿中随机地抽查了 9 人，其高度(以 cm 为单位)分别为 115,120,131,115,109,115,115,105,110.假设 5 至 6 岁幼儿身高总体的标准差 $\sigma_0=7$，在置信度为 95%的条件下，试求出总体均值 $\mu$ 的置信区间.

**解** 已知 $\sigma_0=7, n=9, \alpha=0.05$.由样本值算得
$$\bar{x}=\frac{1}{9}(115+120+\cdots+110)=115.$$
查正态分布数值表得到临界值 $\lambda=1.96$.由(7.1)式 $\mu$ 在置信度为 95%下的置信区间为
$$\left[115-1.96 \frac{7}{\sqrt{9}}, 115+1.96 \frac{7}{\sqrt{9}}\right],$$
即[110.43,119.57].

(2) 未知方差，估计均值

设 $x_1, x_2, \cdots, x_n$ 为来自总体 $N(\mu, \sigma^2)$ 的一个样本，由于 $\sigma^2$ 是

未知的，不能再选取样本函数 $u$. 这时可用样本方差

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

来代替 $\sigma^2$，而选取样本函数

$$t = \frac{\bar{x} - \mu}{S/\sqrt{n}}.$$

由前面的讨论已经知道，随机变量 $t$ 服从 $n-1$ 个自由度的 $t$ 分布. 因此，对于给定的置信度 $1-\alpha$，查 $t$ 分布临界值表（见附表 2），找出两个临界值 $\lambda_1$ 与 $\lambda_2$，使得

$$P\{\lambda_1 \leqslant t \leqslant \lambda_2\} = 1 - \alpha.$$

与前面讨论类似，我们仍然取成对称区间$[-\lambda,\lambda]$，使得

$$P\{|t| \leqslant \lambda\} = 1 - \alpha,$$

即
$$P\left\{-\lambda \leqslant \frac{\bar{x}-\mu}{S/\sqrt{n}} \leqslant \lambda\right\} = 1-\alpha.$$

将本书所附的 $t$ 分布临界值表的构造

$$P\{|t| > \lambda\} = \alpha$$

（即图 7.3 中的阴影部分）与 $P\{|t| \leqslant \lambda\} = 1-\alpha$（即图 7.4 中的阴影部分）比较，不难看出，确定 $\lambda$ 之值的方法是查 $t(n-1,\alpha)$表找出 $\lambda$ 之值，其中 $n$ 是样本容量，$n-1$ 是表中的自由度.

这样我们把 $\lambda$ 代入不等式

$$-\lambda \leqslant \frac{\bar{x}-\mu}{S/\sqrt{n}} \leqslant \lambda,$$

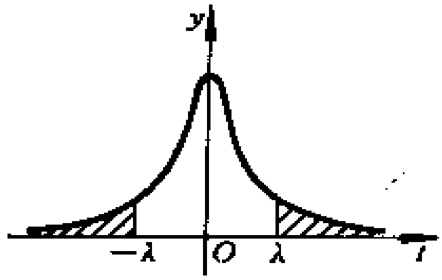



图 7.3

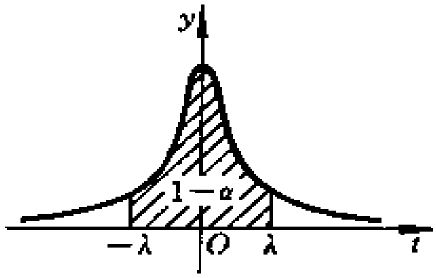



图 7.4

推得

$$\bar{x}-\lambda\frac{S}{\sqrt{n}}\leqslant\mu\leqslant\bar{x}+\lambda\frac{S}{\sqrt{n}}.$$

这就是说，随机区间

$$\left[\bar{x}-\lambda\frac{S}{\sqrt{n}},\ \bar{x}+\lambda\frac{S}{\sqrt{n}}\right] \tag{7.2}$$

以 $1-\alpha$ 的概率包含 $\mu$.

**例 8** 用某种仪器间接测量温度，重复测量 7 次，测得温度分别为 120.0，113.4，111.2，114.5，112.0，112.9，113.6(℃). 设温度 $X\sim N(\mu,\sigma^2)$，在置信度为 95%的条件下，试求出温度的真值所在的范围.

**解** 设 $\mu$ 为温度的真值，$X$ 为测量值，在仪器没有系统偏差情况下，即 $E(X)=\mu$ 时，重复测量 7 次，得到 $X$ 的 7 个样本值. 问题就是在未知方差(即仪器的精度)的情况下，找出 $\mu$ 的置信区间.

已知 $n=7$，$\alpha=0.05$，由样本值算得

$$\bar{x}=112.8,\quad S^2=1.29.$$

查 $t(6,0.05)$ 得到临界值 $\lambda=2.447$. 由(7.2)式 $\mu$ 在置信度为 95%下的置信区间为

$$\left[112.8-2.447\sqrt{\frac{1.29}{7}},\quad 112.8+2.447\sqrt{\frac{1.29}{7}}\right],$$

即[111.75，113.85].

2. 方差的区间估计

设 $x_1,x_2,\cdots,x_n$ 为来自总体 $N(\mu,\sigma^2)$ 的一个样本，我们知道 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$ 是 $\sigma^2$ 的一个点估计，并且知道包含未知参数 $\sigma^2$ 的样本函数

$$w=\frac{(n-1)S^2}{\sigma^2}$$

是服从 $n-1$ 个自由度的 $\chi^2$ 分布. 因此，对于给定的置信度 $1-\alpha$，

查 $\chi^2$ 分布临界值表(见附表 3),找出两个临界值 $\lambda_1$ 与 $\lambda_2$,使得

$$P\{\lambda_1 \leqslant w \leqslant \lambda_2\} = 1 - \alpha.$$

由于 $\chi^2$ 分布不具有对称性,因此通常采取使得概率对称的区间,即

$$P\{w < \lambda_1\} = P\{w > \lambda_2\} = \frac{\alpha}{2}.$$

于是有

$$P\left\{\lambda_1 \leqslant \frac{(n-1)S^2}{\sigma^2} \leqslant \lambda_2\right\} = 1 - \alpha.$$

将本书所附的 $\chi^2$ 分布临界值表的构造

$$P\{\chi^2 > \lambda\} = \alpha$$

(即图 7.5 中的阴影部分)与 $P\{\lambda_1 \leqslant w \leqslant \lambda_2\} = 1-\alpha$(即图 7.6 中的阴影部分)比较,不难看出,确定 $\lambda$ 之值的方法是查 $\chi^2(n-1,\alpha/2)$ 找出 $\lambda_2$,而查 $\chi^2\left(n-1,1-\frac{\alpha}{2}\right)$ 找出 $\lambda_1$ 来,其中 $n$ 是样本容量,$n-1$ 是表中的自由度.

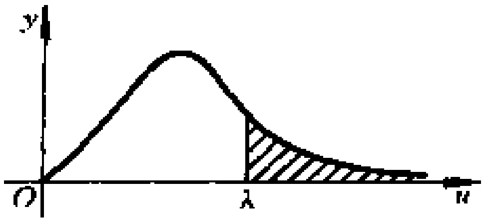



图 7.5

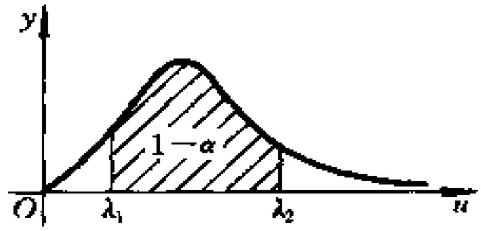



图 7.6

这样我们把 $\lambda_1$ 与 $\lambda_2$ 代入不等式

$$\lambda_1 \leqslant \frac{(n-1)S^2}{\sigma^2} \leqslant \lambda_2,$$

有

$$\frac{1}{\lambda_1} \geqslant \frac{\sigma^2}{(n-1)S^2} \geqslant \frac{1}{\lambda_2}.$$

推得

$$\frac{(n-1)S^2}{\lambda_2} \leqslant \sigma^2 \leqslant \frac{(n-1)S^2}{\lambda_1}.$$

这就是说，随机区间

$$\left[\frac{(n-1)S^2}{\lambda_2}, \frac{(n-1)S^2}{\lambda_1}\right] \tag{7.3}$$

以 $1-\alpha$ 的概率包含 $\sigma^2$. 而随机区间

$$\left[\sqrt{\frac{n-1}{\lambda_2}}S, \sqrt{\frac{n-1}{\lambda_1}}S\right]$$

以 $1-\alpha$ 概率包含 $\sigma$.

**例 9** 设某自动车床加工的零件其长度 $X \sim N(\mu,\sigma^2)$，今抽查 16 个零件，测得长度(以 mm 为单位)如下：

12.15，12.12，12.01，12.08，

12.09，12.16，12.03，12.01，

12.06，12.13，12.07，12.11，

12.08，12.01，12.03，12.06.

在置信度为 95%的条件下，试求出总体方差 $\sigma^2$ 的置信区间.

**解** 已知 $n=16$，$\alpha=0.05$，由样本值算得 $S^2=0.00244$. 查 $\chi^2(15,0.975)$ 得到 $\lambda_1=6.26$，查 $\chi^2(15,0.025)$ 得到 $\lambda_2=27.5$. 由(7.3)式 $\sigma^2$ 的置信度为 95%下的置信区间为

$$\left[\frac{15\times 0.00244}{27.5}, \frac{15\times 0.00244}{6.26}\right],$$

即[0.0013,0.0058].

3. 双正态总体的均值差与方差比的区间估计

设 $\overline{X}_1$ 和 $S_1^2$ 是来自正态总体 $N(\mu_1,\sigma_1^2)$ 的容量为 $n_1$ 的样本均值和样本方差；$\overline{X}_2$ 和 $S_2^2$ 是来自正态总体 $N(\mu_2,\sigma_2^2)$ 的容量为 $n_2$ 的样本均值和样本方差，且设这两个样本相互独立.

(1) 当 $\sigma_1^2,\sigma_2^2$ 都为已知时，均值差 $\mu_1-\mu_2$ 的 $1-\alpha$ 置信区间为

$$\left(\overline{X}_1-\overline{X}_2-\lambda\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}, \overline{X}_1-\overline{X}_2+\lambda\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right),$$

其中临界值 $\lambda$ 是由 $\Phi(\lambda)=1-\frac{\alpha}{2}$ 得到的.

(2) 当 $\sigma_1^2,\sigma_2^2$ 都为未知时，只要 $n_1,n_2$ 充分大，$\mu_1-\mu_2$ 的 $1-\alpha$ 置

信区间近似为

$$\left(\overline{X}_1-\overline{X}_2-\lambda\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}},\ \overline{X}_1-\overline{X}_2+\lambda\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\right),$$

其中临界值 $\lambda$ 是由 $\Phi(\lambda)=1-\frac{\alpha}{2}$ 得到的.

(3) 当 $\sigma_1^2=\sigma_2^2=\sigma^2$ 但 $\sigma^2$ 为未知时，$\mu_1-\mu_2$ 的 $1-\alpha$ 置信区间为

$$\left(\overline{X}_1-\overline{X}_2-\lambda S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\ \overline{X}_1-\overline{X}_2+\lambda S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right),$$

其中

$$S_w^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}.$$

而临界值 $\lambda$ 是查 $t(n_1+n_2-2,\alpha)$ 表所得到的.

(4) 当两个正态总体的参数均为未知时，方差比 $\sigma_1^2/\sigma_2^2$ 的 $1-\alpha$ 置信区间为

$$\left(\frac{S_1^2}{S_2^2}\frac{1}{\lambda_1},\ \frac{S_1^2}{S_2^2}\lambda_2\right),$$

其中临界值 $\lambda_1$ 是查 $f\left(n_1-1,n_2-1\ \frac{\alpha}{2}\right)$ 表所得到的，而 $\lambda_2$ 是查 $f\left(n_2-1,n_1-1,\frac{\alpha}{2}\right)$ 表所得到的.

## 三、典型例题分析

### (一) 填空题

1. 设总体 $X$ 的方差为 1，根据来自 $X$ 的容量为 100 的简单随机样本，测得样本均值为 5，则 $X$ 的数学期望的置信度近似等于 0.95 的置信区间为________.

答案是：[4.804,5.196].

**分析** 这是一个已知方差，估计均值问题. 由于 $1-\alpha=0.95$，查 $\Phi(x)=1-\frac{\alpha}{2}=0.975$. 得到 $\lambda=1.96$，因此，置信区间为

$$\left[5-1.96\frac{1}{\sqrt{100}},\ 5+1.96\frac{1}{\sqrt{100}}\right],$$

即[4.804,5.196].

2. 设由来自正态总体 $X\sim N(\mu,0.9^2)$容量为 9 的简单随机样本得样本均值 $\overline{X}=5$,则未知参数 $\mu$ 的置信度为 0.95 的置信区间是________.

答案是:[4.412,5.588].

**分析** 这是一个已知方差,估计均值问题. 由于 $1-\alpha=0.95$,查 $\Phi(x)=1-\frac{\alpha}{2}=0.975$,得到 $\lambda=1.96$,因此,置信区间为

$$\left[5-1.96\frac{0.9}{\sqrt{9}},\ 5+1.96\frac{0.9}{\sqrt{9}}\right],$$

即[4.412,5.588].

3. 在天平上重复称量一重为 $a$ 的物品,假设各次称量结果相互独立且同服从正态分布 $N(a,0.2^2)$. 若以$\overline{X}_n$ 表示 $n$ 次称量结果的算术平均值,则为使

$$P\{|\overline{X}_n-a|<0.1\}\geqslant 0.95,$$

$n$ 的最小值应不小于自然数________.

答案是:16.

**分析** 这是一个已知方差,估计均值问题. 由

$$P\{|\bar{x}-a|<0.1\}=P\left\{\left|\frac{\bar{x}-a}{0.2/\sqrt{n}}\right|<\frac{0.1}{0.2/\sqrt{n}}\right\}=0.95$$

查 $\Phi(x)=1-\frac{\alpha}{2}=0.975$,得到 $x=1.96$,即

$$\frac{\sqrt{n}}{2}=1.96,$$

解得 $n=4\times1.96^2$. 故 $n$ 的最小值应不小于自然数 16.

(二) 解答题

1. 从正态总体 $N(3.4,6^2)$中抽取容量为 $n$ 的样本,如果要求其样本均值位于区间(1.4,5.4)内的概率不小于 0.95,问样本容

量 $n$ 至少应取多大？

附表：标准正态分布表

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

| $z$ | 1.28 | 1.645 | 1.96 | 2.33 |
|---|---|---|---|---|
| $\Phi(z)$ | 0.900 | 0.950 | 0.975 | 0.990 |

**解** 以 $\overline{X}$ 表示该样本均值，则

$$\frac{\overline{X}-3.4}{6}\sqrt{n}\sim N(0,1).$$

从而有

$$P\{1.4<\overline{X}<5.4\}=P\{-2<\overline{X}-3.4<2\},$$

$$P\{|\overline{X}-3.4|<2\}=P\left\{\frac{|\overline{X}-3.4|}{6}\sqrt{n}<\frac{2\sqrt{n}}{6}\right\}$$

$$=2\Phi\left(\frac{\sqrt{n}}{3}\right)-1\geqslant 0.95,$$

故 $\Phi\left(\frac{\sqrt{n}}{3}\right)\geqslant 0.975$. 由此得 $\frac{\sqrt{n}}{3}\geqslant 1.96$，即

$$n\geqslant(1.96\times 3)^2\approx 34.57,$$

所以 $n$ 至少应取 35.

2. 设总体 $X$ 的概率密度为

$$f(x)=\begin{cases}\frac{6x}{\theta^3}(\theta-x), & 0<x<\theta,\\ 0, & \text{其他},\end{cases}$$

$X_1,X_2,\cdots,X_n$ 是取自总体 $X$ 的简单随机样本.

(1) 求 $\theta$ 的矩估计量 $\hat{\theta}$；

(2) 求 $\hat{\theta}$ 的方差 $D(\hat{\theta})$.

**解** (1) $E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\int_0^{\theta}\frac{6x^2}{\theta^3}(\theta-x)\mathrm{d}x=\frac{\theta}{2}$.

记 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，令 $\frac{\theta}{2}=\overline{X}$，得 $\theta$ 的矩估计量为 $\hat{\theta}=2\overline{X}$.

(2) 由于

$$E(X^2)=\int_{-\infty}^{+\infty}x^2f(x)\mathrm{d}x=\int_0^\theta\frac{6x^3}{\theta^3}(\theta-x)\mathrm{d}x=\frac{6\theta^2}{20},$$

$$D(X)=E(X^2)-[E(X)]^2=\frac{6\theta^2}{20}-\left(\frac{\theta}{2}\right)^2=\frac{\theta^2}{20},$$

所以$\hat{\theta}=2\overline{X}$ 的方差为

$$D(\hat{\theta})=D(2\overline{X})=4D(\overline{X})=\frac{4}{n}D(X)=\frac{\theta^2}{5n}.$$

3. 设$\hat{\theta}$是参数$\theta$的无偏估计量，且$D(\hat{\theta})>0$. 证明：$\hat{\theta}^2$ 不是 $\theta^2$ 的无偏估计量.

**证** 由公式$E(X^2)=D(X)+(E(X))^2$，有

$$E(\hat{\theta})^2=D(\hat{\theta})+(E(\hat{\theta}))^2=D(\hat{\theta})+\theta^2>\theta^2.$$

因此$\hat{\theta}^2$ 不是 $\theta^2$ 的无偏估计量.

4. 设 $x_1,x_2,\cdots,x_n$ 是来自正态分布 $N(\mu,\sigma^2)$的一个样本，适当选择常数 $c$，使 $c\sum\limits_{i=1}^{n-1}(x_{i+1}-x_i)^2$ 为 $\sigma^2$ 的无偏估计.

**解** **方法 1** 由于

$$\begin{aligned}\sigma^2&=E\left[c\sum_{i=1}^{n-1}(x_{i+1}-x_i)^2\right]=c\sum_{i=1}^{n-1}E(x_{i+1}^2-2x_{i+1}x_i+x_i^2)\\&=c\sum_{i=1}^{n-1}[E(x_{i+1}^2)-2E((x_{i+1})(x_i))+E(x_i^2)]\\&=c\sum_{i=1}^{n-1}2[E(X^2)-(E(X))^2]=2c(n-1)\sigma^2,\end{aligned}$$

因此

$$c=\frac{1}{2(n-1)}.$$

**方法 2** 令 $Y=x_{i+1}-x_i$，则 $Y\sim N(0,2\sigma^2)$. 由

$$\begin{aligned}E\left(c\sum_{i=1}^{n-1}Y^2\right)&=c\sum_{i=1}^{n-1}E(Y^2)=c(n-1)(D(Y)+(E(Y))^2)\\&=c(n-1)2\sigma^2=\sigma^2,\end{aligned}$$

因此

$$c=\frac{1}{2(n-1)}.$$

5. 设 $x_1,x_2,\cdots,x_n$ 是取自总体 $X\sim N(\mu,\sigma^2)$的样本,试证

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

是 $\sigma^2$ 的相合估计量.

**证** 由于 $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$,并且有

$$E(S^2)=\sigma^2,\quad D(S^2)=\frac{\sigma^4}{(n-1)^2}2(n-1)=\frac{2\sigma^4}{n-1}.$$

根据切比雪夫不等式有:

$$P\{|S^2-\sigma^2|<\varepsilon\}\geqslant 1-\frac{D(S^2)}{\varepsilon^2}=1-\frac{2\sigma^4}{(n-1)\varepsilon^2},$$

即得:

$$\lim_{n\to\infty}P\{|S^2-\sigma^2|<\varepsilon\}=1.$$

所以 $S^2$ 是 $\sigma^2$ 的相合估计量.

6. 设总体 $X$ 的概率密度为

$$p(x;\lambda)=\begin{cases}\lambda\alpha x^{\alpha-1}e^{-\lambda x^{\alpha}}, & 若\ x>0,\\ 0, & 若\ x\leqslant 0,\end{cases}$$

其中 $\lambda>0$ 是未知参数,$\alpha>0$ 是已知常数,试根据来自总体 $X$ 的简单随机样本 $x_1,x_2,\cdots,x_n$,求 $\lambda$ 的最大似然估计量$\hat{\lambda}$.

**解** 似然函数为

$$L(x_1,x_2,\cdots,x_n;\lambda)=\prod_{i=1}^{n}p(x_i,\lambda)=(\lambda\alpha)^n e^{-\lambda\sum_{i=1}^{n}x_i^{\alpha}}\prod_{i=1}^{n}x_i^{\alpha-1}.$$

由对数似然方程

$$\frac{\partial\ln L}{\partial\lambda}=\frac{n}{\lambda}-\sum_{i=1}^{n}x_i^{\alpha}=0$$

得 $\lambda$ 的最大似然估计量为

$$\hat{\lambda}=n\Big/\sum_{i=1}^{n}x_i^{\alpha}.$$

7. 设总体 $X$ 的分布为

$$p(x)=\begin{cases}(\theta+1)x^{\theta}, & 0<x<1,\\ 0, & 其他,\end{cases}$$

其中 $\theta>-1$ 是未知参数，$x_1,x_2,\cdots,x_n$ 为来自总体 $X$ 的一个容量为 $n$ 的简单随机样本．分别用矩法及最大似然法估计 $\theta$．

**解** （1）矩法．因为

$$E(X)=\int_{-\infty}^{+\infty}xp(x)\mathrm{d}x=\int_0^1(\theta+1)x^{\theta+1}\mathrm{d}x=\frac{\theta+1}{\theta+2}.$$

又因为 $E(X)=\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$，有 $\frac{\theta+1}{\theta+2}=\bar{x}$．解之 $\hat{\theta}=\frac{2\bar{x}-1}{1-\bar{x}}$．

（2）最大似然法．由题意有

$$L(\theta)=\begin{cases}(\theta+1)^n\left(\prod_{i=1}^{n}x_i\right)^{\theta}, & 0<x_i<1,\\ 0, & \text{其他}\end{cases}$$

$$(i=1,2,\cdots,n).$$

当 $0<x_i<1$ 时，$L(\theta)>0$，

$$\ln L(\theta)=n\ln(\theta+1)+\theta\sum_{i=1}^{n}\ln x_i,$$

$$\frac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta}=\frac{n}{\theta+1}+\sum_{i=1}^{n}\ln x_i.$$

令 $\frac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta}=0$ 便得到 $\theta$ 的最大似然估计量

$$\hat{\theta}=-1-\frac{n}{\sum_{i=1}^{n}\ln x_i}=-1-\frac{1}{\frac{1}{n}\sum_{i=1}^{n}\ln x_i}.$$

8. 设 $X\sim U(0,\theta)$，$\theta>0$，求 $\theta$ 的最大似然估计量及矩估计量．

**解** （1）由于 $X\sim U(0,\theta)$，

$$p(x;\theta)=\begin{cases}\frac{1}{\theta}, & 0\leqslant x\leqslant\theta,\\ 0, & \text{其他},\end{cases}$$

有 $$L(\theta)=\prod_{i=1}^{n}p(x_i;\theta)=\begin{cases}\frac{1}{\theta^n}, & 0\leqslant x_i\leqslant\theta,\\ 0, & \text{其他}.\end{cases}\quad(i=1,2,\cdots,n)$$

又因为 $\theta>0$，所以 $L(\theta)$ 随着 $\theta$ 减小而增大．但 $\theta\geqslant\max\limits_{1\leqslant i\leqslant n}\{x_i\}$，故取

$$\hat{\theta} = \max_{1\leqslant i\leqslant n}\{x_i\}$$

为 $\theta$ 的最大似然估计量.

(2) $E(X)=\dfrac{\theta}{2}$, $\bar{x}=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i$. 由 $\dfrac{\theta}{2}=\bar{x}$,故 $\hat{\theta}=2\bar{x}$ 为 $\theta$ 的矩估计量.

9. 为了估计灯泡使用时数的均值 $\mu$ 和标准差 $\sigma$,共测试了 10 个灯泡,得 $\bar{x}=1500\text{h}$,$S=20\text{h}$. 如果已知灯泡使用时数是服从正态分布的,求出 $\mu$ 和 $\sigma$ 的置信区间(置信度为 0.95).

**解** 先求 $\mu$ 的置信区间.

(1) 选择样本函数

$$t=\frac{\bar{x}-\mu}{\sqrt{\dfrac{S^2}{n}}}\sim t(n-1);$$

(2) 查 $t$ 分布表,找临界值. 由 $t(9,0.05)$得到 $\lambda=2.262$;

(3) 置信区间为

$$\left[\bar{x}-\lambda\sqrt{\frac{S^2}{n}},\ \bar{x}+\lambda\sqrt{\frac{S^2}{n}}\right],$$

将具体值代入,得到

$$\mu\in[1485.7,\ 1514.3].$$

再求 $\sigma$ 的置信区间.

(1) 选择样本函数

$$w=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1);$$

(2) 查 $\chi^2$ 分布表,找临界值 $\lambda_1$,$\lambda_2$.

由 $\chi^2(9,0.025)$得到 $\lambda_2=19.0$,又由 $\chi^2(9,0.975)$得到 $\lambda_1=2.70$;

(3) 置信区间为

$$\left[\sqrt{\frac{n-1}{\lambda_2}}S,\ \sqrt{\frac{n-1}{\lambda_1}}S\right],$$

将具体值代入，得到

$$\sigma \in [13.8,\ 36.5].$$

10. 设总体 $X$ 在$[\theta_1,\theta_2]$上服从均匀分布，$\theta_1,\theta_2$ 未知. 试由样本 $X_1,X_2,\cdots,X_n$，求 $\theta_1$ 与 $\theta_2$ 的矩估计量与最大似然估计量.

**解** (1) $\mu_1=E(X)=\dfrac{\theta_1+\theta_2}{2}$,

$$\mu_2=E(X^2)=D(X)+(E(X))^2$$
$$=\frac{(\theta_2-\theta_1)^2}{12}+\frac{(\theta_1+\theta_2)^2}{4}.$$

令

$$\hat{\mu}_1=\overline{X},\quad \hat{\mu}_2=\frac{1}{n}\sum_{i=1}^{n}X_i^2,$$

即

$$\overline{X}=\frac{1}{2}(\hat{\theta}_1+\hat{\theta}_2),\quad \frac{1}{n}\sum_{i=1}^{n}X_i^2=\frac{(\hat{\theta}_2-\hat{\theta}_1)^2}{12}+\overline{X}^2,$$

$$\hat{\theta}_1+\hat{\theta}_2=2\overline{X},$$

$$\frac{(\hat{\theta}_2-\hat{\theta}_1)^2}{12}=\frac{1}{n}\sum X_i^2-\overline{X}^2=\frac{1}{n}\sum(X_i-\overline{X})^2=S_n^2,$$

$$\hat{\theta}_2-\hat{\theta}_1=2\sqrt{3}S_n,$$

解得 $\theta_1$ 与 $\theta_2$ 的矩估计量为

$$\hat{\theta}_1=\overline{X}-\sqrt{3}S_n,\quad \hat{\theta}_2=\overline{X}+\sqrt{3}S_n.$$

(2) 由 $X$ 的密度函数

$$f(x;\theta_1,\theta_2)=\begin{cases}\dfrac{1}{\theta_2-\theta_1}, & \theta_1\leqslant x\leqslant\theta_2,\\ 0, & \text{其他}\end{cases}$$

知 $\theta_1$ 与 $\theta_2$ 的似然函数为

$$L(\theta_1,\theta_2)=\begin{cases}\left(\dfrac{1}{\theta_2-\theta_1}\right)^n, & \theta_1\leqslant x_i\leqslant\theta_2,\ i=1,2,\cdots,n,\\ 0, & \text{其他},\end{cases}$$

似然方程为

$$\begin{cases}\dfrac{\partial \ln L}{\partial \theta_1}=\dfrac{n}{\theta_2-\theta_1}=0,\\ \dfrac{\partial \ln L}{\partial \theta_2}=\dfrac{-n}{\theta_2-\theta_1}=0,\end{cases}$$

从似然方程不可能解得 $\theta_1$ 与 $\theta_2$ 的最大似然估计量. 现在我们可直接从最大似然原理出发来确定 $\theta_1$ 与 $\theta_2$ 的最大似然估计量. 我们知道,欲使似然函数非零,必须要求:

$$\theta_1 \leqslant x_i \leqslant \theta_2,\quad i=1,2,\cdots,n.$$

从而

$$\theta_1 \leqslant x_1^* = \min\{x_1,x_2,\cdots,x_n\},$$

$$\theta_2 \geqslant x_n^* = \max\{x_1,x_2,\cdots,x_n\}.$$

由于

$$L(\theta_1,\theta_2)=\left(\frac{1}{\theta_2-\theta_1}\right)^n \leqslant \left(\frac{1}{x_n^*-x_1^*}\right)^n,$$

今取

$$\hat{\theta}_1 = x_1^* = \min\{x_1,x_2,\cdots,x_n\},$$

$$\hat{\theta}_2 = x_n^* = \max\{x_1,x_2,\cdots,x_n\},$$

则有

$$L(\theta_1,\theta_2) \leqslant L(\hat{\theta}_1,\hat{\theta}_2).$$

故得 $\theta_1$ 与 $\theta_2$ 的最大似然估计量分别为

$$\hat{\theta}_1 = \min\{X_1,X_2,\cdots,X_n\},$$

$$\hat{\theta}_2 = \max\{X_1,X_2,\cdots,X_n\}.$$

11. 设总体 $X$ 的密度函数为

$$f(x)=\begin{cases}\dfrac{\beta^k}{(k-1)!}x^{k-1}e^{-\beta x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

其中 $k$ 是已知的正整数,$\beta$ 为未知参数. 试由样本 $X_1,X_2,\cdots,X_n$ 求 $\beta$ 的矩估计量与最大似然估计量.

**解** (1) $E(X)=\int_0^{+\infty}\frac{\beta^k}{(k-1)!}x^k e^{-\beta x}dx$ (令 $\beta x=y$)

$$=\frac{1}{\beta(k-1)!}\int_0^{+\infty}y^k e^{-y}dy=\frac{\Gamma(k+1)}{\beta(k-1)!}$$

$$=\frac{k!}{\beta(k-1)!}=\frac{k}{\beta},$$

令$\frac{k}{\beta}=\overline{X}$,解得 $\beta$ 的矩估计量为 $\hat{\beta}=\frac{k}{\overline{X}}$.

(2) 似然函数

$$L(\beta)=\prod_{i=1}^{n}f(x_i)$$

$$=\begin{cases}\left(\frac{1}{(k-1)!}\right)^n\beta^{nk}\left(\prod_{i=1}^{n}x_i\right)^{k-1}e^{-\beta\sum_{i=1}^{n}x_i}, & x_i>0,\\ 0, & \text{其他},\end{cases}$$

$$i=1,2,\cdots,n.$$

当 $x_i>0$ $(i=1,2,\cdots,n)$时,$L>0$,

$$\ln L(\beta)=-n\ln(k-1)!+nk\ln\beta$$

$$+(k-1)\ln\left(\prod_{i=1}^{n}x_i\right)-\beta\sum_{i=1}^{n}x_i.$$

由

$$\frac{d\ln L}{d\beta}=\frac{nk}{\beta}-\sum_{i=1}^{n}x_i=0,$$

解得 $\beta$ 的最大似然估计值为$\hat{\beta}=\frac{nk}{\sum_{i=1}^{n}x_i}=\frac{k}{\overline{x}}$,从而得 $\beta$ 的最大似然估计量为

$$\hat{\beta}=\frac{k}{\overline{X}}.$$

12.(钓鱼问题) 设湖中有鱼 $N$ 条,现钓出 $r$ 条,做上记号后放回湖中.一段时间后,再钓出 $s$ 条(设 $s\geqslant r$),结果其中有 $t$ 条($0\leqslant t\leqslant r$)标有记号.试根据此种信息,估计湖中鱼数 $N$ 的值.

**解** 这是个典型的统计估值问题.钓出 $s$ 条,其中标有记号的

鱼数应是个随机变量，记为 $X$. 显然 $X$ 只可能取 $0,1,\cdots,r$ 这 $r+1$ 个值，现 $X=t$，且

$$P\{X=t\}=C_r^t C_{N-r}^{s-t}/C_N^s=L(t,N),$$

其中 $N$ 为未知参数. 今钓出 $s$ 条即已出现 $t$ 条，则我们认为 $N$ 应该使得 $P\{X=t\}$ 最大，即取 $\hat{N}$，使得 $L(t,\hat{N})=\max\limits_N L(t,N)$. 为具体决定 $N$，我们考虑比值

$$\begin{aligned}R(t,N)&=\frac{L(t,N)}{L(t,N-1)}=\frac{C_r^t C_{N-r}^{s-t}}{C_N^s}\Big/\frac{C_r^t C_{N-1-r}^{s-t}}{C_{N-1}^s}\\&=\frac{C_{N-r}^{s-t}C_{N-1}^s}{C_{N-1-r}^{s-t}C_N^s}=\frac{(N-r)(N-s)}{N(N-r-s+t)}\\&=\frac{N^2-Nr-Ns+rs}{N^2-Nr-Ns+Nt}.\end{aligned}$$

从上式可见，当 $rs<Nt$ 时，$R(t,N)<1$，当 $rs>Nt$ 时，$R(t,N)>1$. 故当 $N<\frac{rs}{t}$ 时，$L(t,N)$ 是 $N$ 的上升函数；当 $N>\frac{rs}{t}$ 时，$L(t,N)$ 是 $N$ 的下降函数. 由于 $N$ 是正整数，故取 $\hat{N}=\left[\frac{rs}{t}\right]$ $\left(\frac{rs}{t}\text{的整数部分}\right)$ 作为 $N$ 的估计.

从直观上看，湖中有标记的鱼的比例和钓出的 $s$ 条鱼中有标记的鱼所占的比例似应相一致，即 $r:N=t:s$，因而有 $\hat{N}\approx\frac{rs}{t}$. 上面的估计正好与此直观的结果相符，而这也说明了最大似然估计的合理性.

13. 铅的密度测量值是服从正态分布的. 如果测量 16 次，算得 $\bar{x}=2.705$，$S=0.029$，试求铅的密度置信度为 95%的置信区间.

**解** 这里方差 $\sigma^2$ 未知. $1-\alpha=95\%$，$\alpha=0.05$，$n=16$，自由度 $n-1=15$，查 $t$ 分布表得 $t_{\alpha/2}(15)=2.131$，从而有

置信下限：$\bar{x}-t_{\alpha/2}(n-1)\dfrac{S}{\sqrt{n}}=2.705-2.131\times\dfrac{0.029}{\sqrt{16}}$

$$=2.690,$$

置信上限：$\bar{x}+t_{\alpha/2}(n-1)\dfrac{S}{\sqrt{n}}=2.705+2.131\times\dfrac{0.029}{\sqrt{16}}$

$$=2.720,$$

故铅的密度的置信度为95%的置信区间是(2.690, 2.720).

**评注** 铅的密度等于它的测量值的平均值

14. 设炮弹速度服从正态分布，取9发炮弹做试验，得样本方差$S^2=11(\mathrm{m/s})^2$，分别求炮弹速度方差$\sigma^2$和标准差$\sigma$的置信度为90%的置信区间.

**解** 据题意$n=9$，自由度$n-1=8$，$1-\alpha=0.90$，$\alpha=0.10$，查$\chi^2$分布表得$\chi^2_{\alpha/2}(8)=15.507$，$\chi^2_{1-\alpha/2}(8)=2.733$. 从而$\sigma^2$的

置信下限：$\dfrac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}=\dfrac{8\times 11}{15.507}=5.675$，

置信上限：$\dfrac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}=\dfrac{8\times 11}{2.733}=32.199$，

故$\sigma^2$的置信区间是(5.675, 32.199)，而$\sigma$的置信区间是(2.382, 5.674).

15. 设$\hat{\theta}_1$，$\hat{\theta}_2$是参数$\theta$的二个相互独立的无偏估计量，且设$D(\hat{\theta}_1)=2D(\hat{\theta}_2)$. 找出常数$k_1$，$k_2$，使$k_1\hat{\theta}_1+k_2\hat{\theta}_2$也是$\theta$的无偏估计量，并且使它在所有这样形状的估计量中方差最小.

**解** 因为$E(\hat{\theta}_1)=E(\hat{\theta}_2)=\theta$，所以

$$E(k_1\hat{\theta}_1+k_2\hat{\theta}_2)=(k_1+k_2)\theta,$$

欲使$E(k_1\hat{\theta}_1+k_2\hat{\theta}_2)=\theta$，只须$k_1+k_2=1$. 又因为$\hat{\theta}_1$与$\hat{\theta}_2$相互独立，$D(\hat{\theta}_1)=2D(\hat{\theta}_2)$，故

$$\begin{aligned}D(k_1\hat{\theta}_1+k_2\hat{\theta}_2)&=k_1^2D(\hat{\theta}_1)+k_2^2D(\hat{\theta}_2)\\&=(2k_1^2+k_2^2)D(\hat{\theta}_2).\end{aligned}$$

欲使$D(k_1\hat{\theta}_1+k_2\hat{\theta}_2)$为最小，只须$S=2k_1^2+k_2^2=2k_1^2+(1-k_1)^2$为最小.

$$\frac{\mathrm{d}S}{\mathrm{d}k_1} = 4k_1 - 2(1-k_1) = 0, \quad 6k_1 - 2 = 0,$$

得
$$k_1 = \frac{1}{3}, \quad k_2 = \frac{2}{3}.$$

16. 设 $X$ 在 $(0,\theta]$ 上均匀分布，$X_1, X_2, \cdots, X_n$ 是来自 $X$ 的样本. 令 $X_{(n)} = \max\{X_1, X_2, \cdots, X_n\}$，$X_{(1)} = \min\{X_1, X_2, \cdots, X_n\}$，试证 $\hat{\theta}_1 = \frac{n+1}{n} X_{(n)}$，$\hat{\theta}_2 = (n+1)X_{(1)}$ 都是 $\theta$ 的无偏估计量，并且 $\hat{\theta}_1$ 较 $\hat{\theta}_2$ 有效.

**解** $X$ 的密度函数与分布函数分别是

$$p(x) = \begin{cases} \frac{1}{\theta}, & 0 < x \leqslant \theta, \\ 0, & \text{其他}, \end{cases} \quad F(x) = \begin{cases} 0, & x < 0, \\ \frac{x}{\theta}, & 0 \leqslant x < \theta, \\ 1, & x \geqslant \theta, \end{cases}$$

可知[1] $X_{(n)}$ 的分布函数为 $F_n(x) = [F(x)]^n$，$X_{(1)}$ 的分布函数为 $F_1(x) = 1 - [1-F(x)]^n$，从而它们的密度函数分别为

$$p_n(x) = F_n'(x) = n[F(x)]^{n-1}p(x)$$
$$= \begin{cases} \frac{n}{\theta^n}x^{n-1}, & 0 < x \leqslant \theta, \\ 0, & \text{其他}, \end{cases}$$

$$p_1(x) = F_1'(x) = n[1-F(x)]^{n-1}p(x)$$
$$= \begin{cases} \frac{n}{\theta}\left(1 - \frac{x}{\theta}\right)^{n-1}, & 0 < x \leqslant \theta, \\ 0, & \text{其他}, \end{cases}$$

故

$$E(X_{(n)}) = \int_0^\theta x \cdot \frac{nx^{n-1}}{\theta^n}\mathrm{d}x = \frac{n\theta}{n+1},$$

$$E(X_{(1)}) = \int_0^\theta x \cdot \frac{n}{\theta}\left(1 - \frac{x}{\theta}\right)^{n-1}\mathrm{d}x$$

---

① 见浙江大学数学系高等数学教研组编《概率论与数理统计》一书，p92～94.

$$\xlongequal{y=\frac{x}{\theta}} n\theta\int_0^1 y(1-y)^{n-1}\mathrm{d}y$$

$$= n\theta \mathrm{B}(2,n) = n\theta\frac{\Gamma(2)\Gamma(n)}{\Gamma(n+2)}$$

$$= n\theta\frac{1!(n-1)!}{(n+1)!} = \frac{\theta}{n+1},$$

$$E(X_{(n)}^2) = \int_0^\theta x^2\frac{nx^{n-1}}{\theta^n}\mathrm{d}x = \int_0^\theta\frac{nx^{n+1}}{\theta^n}\mathrm{d}x = \frac{n\theta^2}{n+2},$$

$$E(X_{(1)}^2) = \int_0^\theta x^2\cdot\frac{n}{\theta}\left(1-\frac{x}{\theta}\right)^{n-1}\mathrm{d}x$$

$$\xlongequal{y=\frac{x}{\theta}} n\theta^2\int_0^1 y^2(1-y)^{n-1}\mathrm{d}y$$

$$= n\theta^2\mathrm{B}(3,n) = n\theta^2\frac{\Gamma(3)\Gamma(n)}{\Gamma(n+3)}$$

$$= n\theta^2\frac{2!(n-1)!}{(n+2)!} = \frac{2\theta^2}{(n+1)(n+2)},$$

$$D(X_{(n)}) = E[X_{(n)}^2] - [E(X_{(n)})]^2$$

$$= \frac{n\theta^2}{n+2} - \frac{n^2\theta^2}{(n+1)^2} = \frac{n\theta^2}{(n+2)(n+1)^2},$$

$$D(X_{(1)}) = E[X_{(1)}^2] - [E(X_{(1)})]^2$$

$$= \frac{2\theta^2}{(n+1)(n+2)} - \frac{\theta^2}{(n+1)^2}$$

$$= \frac{n\theta^2}{(n+2)(n+1)^2},$$

于是

$$E(\hat{\theta}_1) = E\left[\frac{n+1}{n}X_{(n)}\right] = \frac{n+1}{n}E(X_{(n)})$$

$$= \frac{n+1}{n}\cdot\frac{n}{n+1}\theta = \theta,$$

$$E(\hat{\theta}_2) = E[(n+1)X_{(1)}] = (n+1)E(X_{(1)})$$

$$= (n+1)\cdot\frac{\theta}{n+1} = \theta,$$

即$\hat{\theta}_1$与$\hat{\theta}_2$均是$\theta$的无偏估计. 又

$$D(\hat{\theta}_1) = D\left[\frac{n+1}{n}X_{(n)}\right] = \frac{(n+1)^2}{n^2}D(X_{(n)})$$

$$= \frac{(n+1)^2}{n^2}\cdot\frac{n\theta^2}{(n+2)(n+1)^2} = \frac{\theta^2}{n(n+2)},$$

$$D(\hat{\theta}_2) = D[(n+1)X_{(1)}] = (n+1)^2 D(X_{(1)})$$

$$= (n+1)^2\frac{n\theta^2}{(n+2)(n+1)^2} = \frac{n\theta^2}{n+2},$$

可见 $D(\hat{\theta}_1)\leqslant D(\hat{\theta}_2)$. 当 $n\geqslant 2$ 时, $D(\hat{\theta}_1)<D(\hat{\theta}_2)$, 即$\hat{\theta}_1$较$\hat{\theta}_2$有效.

## 四、练习题

1. 设总体 $X$ 服从泊松分布 $P(\lambda)$, $X_1, X_2, \cdots, X_n$ 是 $X$ 的一个简单随机样本,试求未知参数 $\lambda$ 的矩估计量与最大似然估计量.

2. 设总体 $X$ 的密度函数为

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

由样本 $X_1, X_2, \cdots, X_n$ 求 $\sigma$ 的最大似然估计量;并问所得估计量是否为 $\sigma$ 的无偏估计量.

3. 设总体 $X$ 的密度函数为

$$f(x) = \begin{cases} \sqrt{\theta}\,x^{\sqrt{\theta}-1}, & 0\leqslant x\leqslant 1, \\ 0, & \text{其他}, \end{cases} \quad (\theta > 0)$$

由样本 $X_1, X_2, \cdots, X_n$ 求 $\theta$ 的矩估计量与最大似然估计量.

4. 设总体 $X$ 的密度函数为

$$f(x) = \begin{cases} \frac{1}{\theta}e^{-(x-\mu)/\theta}, & x\geqslant \mu, \\ 0, & \text{其他}, \end{cases} \quad (\theta > 0)$$

由样本 $X_1, X_2, \cdots, X_n$ 求未知参数 $\mu, \theta$ 的矩估计量与最大似然估计量.

5. 设 $X_1,X_2,\cdots,X_n$ 是数学期望 $\mu$ 为已知的正态总体的一个样本，试用最大似然法求 $\sigma^2$ 的估计量 $\hat{\sigma}^2$，并验证此估计量是无偏的和一致的.

6. 设分别从总体 $N(\mu_1,\sigma^2)$ 和 $N(\mu_2,\sigma^2)$ 中抽取容量为 $n_1,n_2$ 的两独立样本. 其样本方差分别为 $S_1^2,S_2^2$，试证：对任意常数 $a$、$b$($a+b=1$)，$T=aS_1^2+bS_2^2$ 都是 $\sigma^2$ 的无偏估计，并确定常数 $a$、$b$ 使 $D(T)$ 达到最小.

7. 对于方差 $\sigma^2$ 为已知的正态总体，问需抽取容量 $n$ 为多大的样本，才能使总体均值 $\mu$、置信度为 $1-\alpha$ 的置信区间的长度不大于 $L$？

## 习题答案与提示

**1.** $\hat{\lambda}_1=\overline{X}$，$\hat{\lambda}_2=\overline{X}$. **2.** $\hat{\sigma}=\dfrac{1}{n}\sum\limits_{i=1}^{n}|X_i|$，它是无偏估计.

**3.** $\hat{\theta}_1=\left(\dfrac{\overline{X}}{1-\overline{X}}\right)^2$，$\hat{\theta}_2=\left(\dfrac{n}{\sum\limits_{i=1}^{n}\ln X_i}\right)^2$.

**4.** $\hat{\mu}_1=\overline{X}-\sqrt{\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}$，$\hat{\theta}_1=\sqrt{\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}$.

$\hat{\mu}_2=X_{(1)}=\min\{X_i\}$，$\hat{\theta}_2=\overline{X}-X_{(1)}$.

**5.** $\hat{\sigma}^2=\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-\mu)^2$.

**6.** $a=\dfrac{n_1-1}{n_1+n_2-2}$，$b=\dfrac{n_2-1}{n_1+n_2-2}$. **7.** $n\geqslant 4u_{\alpha/2}^2\sigma^2/L^2$.

# 第八章 假设检验

## 一、考试要求

1. 理解假设检验的基本思想，了解假设检验的可能产生的两类错误，掌握假设检验的基本步骤；

2. 了解单正态总体的均值和方差的假设检验；双正态总体的均值和方差的假设检验.

## 二、复习要点

### (一) 重要概念及性质

**定义 8.1(统计假设)** 我们把关于总体(分布，特征，相互关系等)的论断称为**统计假设**，记作 $H$. 例如

(1) 对某一总体 $X$ 的分布提出某种假设，如 $H$: $X$ 服从正态分布；或 $H$: $X$ 服从二项分布等等；

(2) 对于总体 $X$ 的分布参数提出某种假设，如 $H$: $\mu=\mu_0$；或 $H$: $\mu\leqslant\mu_0$；或 $H$: $\sigma^2=\sigma_0^2$；或 $H$: $\sigma^2\leqslant\sigma_0^2$ 等等(其中 $\mu_0,\sigma_0^2$ 是已知数，$\mu,\sigma^2$ 是未知参数)；

(3) 对于两个总体 $X$ 与 $Y$ 提出某种假设，$H$: $X,Y$ 具有相同的分布；$H$: $X,Y$ 相互独立等等.

这种对总体分布的未知参数或分布类型提出的假设称为**原假设或零假设**，用 $H_0$ 表示. 当对某个问题提出原假设 $H_0$ 时，事实上也同时给出了另外一个假设，称为**备择假设**或**对立假设**，用 $H_1$ 表示.

统计假设一般可以分成参数假设与非参数假设两种. **参数假设**是指在总体分布类型已知的情况下，关于未知参数的各种统计

假设;**非参数假设**是指在总体分布类型不确知或完全未知的情况下,关于它的各种统计假设.

**定义 8.2(检验)** 不论在哪种统计检验中,所谓对 $H_0$ 进行检验,就是建立一个准则来考核样本,如样本值满足该准则我们就接受 $H_0$,否则就拒绝 $H_0$. 我们称这种准则为**检验准则**,或简称为**检验**.

由于一个样本值或者满足准则或者不满足准则,而没有其他可能,所以一个检验准则本质上就是将样本可能取值的集合 $D$(统称为样本空间)划分成两个部分 $V$ 与 $\overline{V}$,即

$$V \cap \overline{V} = \varnothing, \quad V \cup \overline{V} = D.$$

检验方法如下:当样本值 $(x_1, x_2, \cdots, x_n) \in V$ 时,认为假设 $H_0$ 不成立,从而否定 $H_0$(如 $H_1$ 存在则判其成立,因此接受 $H_1$);相反,当 $(x_1, x_2, \cdots, x_n) \overline{\in} V$ 即 $(x_1, x_2, \cdots, x_n) \in \overline{V}$ 时,认为 $H_0$ 成立,从而接受 $H_0$(如 $H_1$ 存在则判其不成立,因此否定 $H_1$). 通常我们称 $V$ 为 $H_0$ **否定域**,$\overline{V}$ 为 $H_0$ 的**接受域**.

**定义 8.3(两类错误与检验水平)** 如果我们给出了某个检验准则,也就是给出了 $D$ 的一个划分 $V$ 与 $\overline{V}$. 由于样本本身是具有随机性的,因此当我们通过样本进行判断时,还是有可能犯以下两类错误的:

1) 当 $H_0$ 为真时,而样本值却落入了 $V$,按照我们规定的检验法则,应当否定 $H_0$. 这时,我们把客观上 $H_0$ 成立判为 $H_0$ 不成立(即否定了真实的假设),称这种错误为"**以真当假**"的错误或第一类错误,记 $\tilde{\alpha}$ 为犯此类错误的概率,即

$$P\{否定\ H_0 | H_0\ 为真\} = \tilde{\alpha};$$

2) 当 $H_1$ 为真时,而样本值却落入了 $\overline{V}$,按照我们规定的检验法则,应当接受 $H_0$. 这时,我们把客观上 $H_0$ 不成立判为 $H_0$ 成立(即接受了不真实的假设),称这种错误为"**以假当真**"的错误或第二类错误,记 $\tilde{\beta}$ 为犯此类错误的概率,即

$$P\{接受\ H_0 | H_1\ 为真\} = \tilde{\beta}.$$

在显著性检验中，我们把允许犯第一类错误的上界 $\alpha$ 称为显著性水平或**检验水平**.

**定义 8.4(双边检验与单边检验)**

如前所述，建立统计假设的检验准则本质上是要确定否定域 $V$. 我们将会看到在多数情况下，一个好的统计检验准则，其否定域可以通过某个检验统计量 $K=K(x_1,x_2,\cdots,x_n)$ 来描述，即否定域 $V$ 可表示为

$$V=\{(x_1,x_2,\cdots,x_n)\mid K(x_1,x_2,\cdots,x_n)\in R_\alpha\}.$$

即 $(x_1,x_2,\cdots,x_n)\in V$ 与 $K(x_1,x_2,\cdots,x_n)\in R_\alpha$ 是等价的. 这里我们也称 $R_\alpha$ 为**否定域**，$\overline{R}_\alpha$ 为**相容域**. 于是有

$$\begin{aligned}P\{K\in R_\alpha\mid H_0\text{ 为真}\}&=P\{(x_1,x_2,\cdots,x_n)\in V\mid H_0\text{ 为真}\}\\&=\tilde{\alpha},\\P\{K\in \overline{R}_\alpha\mid H_1\text{ 为真}\}&=P\{(x_1,x_2,\cdots,x_n)\in \overline{V}\mid H_1\text{ 为真}\}\\&=\tilde{\beta}.\end{aligned}$$

这样一来，我们便可以根据样本值来计算统计量 $K$ 之值 $\hat{K}$，作出等价的判断：当 $\hat{K}\in R_\alpha$ 时，我们就否定 $H_0$；当 $\hat{K}\in\overline{R}_\alpha$ 时，我们就接受 $H_0$.

在上面的讨论中否定域 $R_\alpha$ 常以下面两种形式给出：一种是

$$R_\alpha=\{x\mid -\infty<x<\lambda_1\text{ 或 }\lambda_2<x<+\infty\},$$

我们把否定域是这种形式的检验叫做**双边检验**；另一种是

$$R_\alpha=\{x\mid \lambda<x<+\infty\},$$

或

$$R_\alpha=\{x\mid -\infty<x<\lambda\},$$

我们把否定域是这种形式的检验叫做**单边检验**.

**定义 8.5(小概率原理)** 假设检验的统计思想是：概率很小的事件在一次试验中可以认为基本上是不会发生的，即**小概率原理**.

为了检验一个假设 $H_0$ 是否成立，我们就先假定 $H_0$ 是成立

的.如果根据这个假定导致了一个不合理的事件发生,那就表明原来的假定 $H_0$ 是不正确的,我们**拒绝接受** $H_0$;如果由此没有导出不合理的现象,则不能拒绝接受 $H_0$,我们称 $H_0$ 是**相容**的.

这里所说的小概率事件就是事件 $\{K\in R_\alpha\}$,其概率就是检验水平 $\alpha$,通常我们取 $\alpha=0.05$,有时也取 0.01 或 0.10.

(二) 重要方法

**方法 8.1**(单正态总体假设检验的基本步骤)

1. 已知方差,检验均值

关于一元正态总体当方差已知时期望的检验程序:

(1) 提出零假设,$H_0$: $\mu=\mu_0$(或 $\mu\leqslant\mu_0$);

(2) 由样本值 $x_1,x_2,\cdots,x_n$ 计算统计量

$$U=\frac{\bar{x}-\mu_0}{\sqrt{\sigma_0^2/n}}$$

的数值$\hat{U}$;

(3) 对于检验水平 $\alpha$ 查标准正态分布表 $\Phi(\lambda)=1-\frac{\alpha}{2}$(或 $\Phi(\lambda)=1-\alpha$)得到临界值 $\lambda$;

(4) 将$\hat{U}$ 与 $\lambda$ 进行比较,作出判断:当 $|\hat{U}|>\lambda$(或$\hat{U}>\lambda$)时否定 $H_0$,否则认为 $H_0$ 相容.

**例 1** 已知滚珠直径服从正态分布.现随机地从一批滚珠中抽取 6 个,测得其直径(单位为 mm)为 14.70,15.21,14.90,14.91,15.32,15.32.假设滚珠直径总体分布的方差为 0.05,问这一批滚珠的平均直径是否为 15.25 mm?($\alpha=0.05$)

**解** 用 $X$ 表示滚珠的直径,已知 $X\sim N(\mu,\sigma^2)$,其中 $\sigma^2=0.05$.这是一个已知方差,检验均值的问题.

首先提出零假设,写出基本假设 $H_0$ 的具体内容.这里我们要检验这批滚珠平均直径是否为 15.25,即 $H_0$: $\mu=15.25$;

然后选择一个统计量,即找一个(包括指定数值的)统计量,使得它在 $H_0$ 成立的条件下与一个(包括总体的待检验参数的)样本

函数有关，这里我们选前面所给出(包括指定数值 15.25)的 $U$ 统计量

$$U=\frac{\bar{x}-15.25}{\sigma_0/\sqrt{n}}.$$

在 $H_0$ 成立的条件下，$U$ 与(包括总体的待检参数 $\mu$ 的)样本函数

$$u=\frac{\bar{x}-\mu}{\sigma_0/\sqrt{n}}$$

都服从标准正态分布，即

$$U \xlongequal{\text{在 } H_0 \text{ 下}} u\sim N(0,1);$$

再由检验水平 $\alpha=0.05$ 选择区域

$$R_\alpha=\{(-\infty,\lambda_1)\cup(\lambda_2,+\infty)\},$$

使得

$$P\{u\in(-\infty,\lambda_1)\}=P\{u\in(\lambda_2,+\infty)\}=\frac{\alpha}{2},$$

即

$$P\{u\in R_\alpha\}=\alpha;$$

可见这里 $\{u\in R_\alpha\}$ 是一个小概率事件.

由于标准正态分布的对称性可知 $\lambda_2=-\lambda_1\xlongequal{\text{def}}\lambda$. 考虑到正态分布数值表的构造，令

$$\Phi(\lambda)=1-\frac{\alpha}{2}$$

可以找出临界值 $\lambda$；这里的 $\alpha=0.05$，根据 $\Phi(\lambda)=1-\dfrac{0.05}{2}=0.975$，查正态分布常值表(见附表 1)得到 $\lambda=1.96$，故否定域

$$R_\alpha=\{(-\infty,-1.96)\cup(1.96,+\infty)\}.$$

最后由样本计算统计量 $U$ 之值 $\hat{U}$，这里

$$\bar{x}=15.06,\quad \hat{U}=\frac{15.06-15.25}{\sqrt{0.05}/\sqrt{6}}=-2.08.$$

于是我们可以做出判断：若 $\hat{U}\in R_\alpha$，则否定 $H_0$，否则认为 $H_0$ 相

容. 例 1 中 $|\hat{U}|=2.08>1.96$, 即 $\hat{U}\in R_\alpha$.

这说明所给的样本值竟使"小概率事件"发生了，这是不合理的. 产生这个不合理现象的根源在于假定了 $H_0$ 是成立的，故应否定假设 $H_0$. 换句话说，这批滚珠平均直径不是 15.25 mm.

**例 2** 问例 1 中的这批滚珠的平均直径是否小于等于 15.25 mm? ($\alpha=0.05$)

我们仍用 $X$ 表示滚珠的直径，有 $X\sim N(\mu,\sigma^2)$，其中 $\sigma^2=0.05$. 这也是一个已知方差，检验均值的问题. 但是由于所提出零假设是 $H_0$: $\mu\leqslant 15.25$，从而使得所选取的统计量

$$U=\frac{\bar{x}-15.25}{\sigma_0/\sqrt{n}}$$

在 $H_0$: $\mu\leqslant 15.25$ 成立的条件下，有不等式

$$\frac{\bar{x}-15.25}{\sigma_0/\sqrt{n}}\leqslant\frac{\bar{x}-\mu}{\sigma_0/\sqrt{n}}.$$

因此 $U$ 与样本函数 $u=\dfrac{\bar{x}-\mu}{\sigma_0/\sqrt{n}}$ 有如下的关系：

$$U\leqslant u\sim N(0,1).$$

这样由检验水平 $\alpha$，选择 $R_\alpha=\{\lambda,+\infty\}$，使得

$$P\{u\in R_\alpha\}=\alpha.$$

可见这里 $\{u\in R_\alpha\}$ 是一个小概率事件. 由正态分布数值表，可令

$$\Phi(\lambda)=1-\alpha,$$

由此找 $\lambda$ 值. 这里的 $\alpha=0.05$，根据 $\Phi(\lambda)=1-\alpha=0.95$ 查正态分布数值表(见附表 1)得到 $\lambda=1.65$. 故否定域 $R_\alpha=\{1.65,+\infty\}$. 由于样本函数 $u$ 中含有总体未知参数 $\mu$，所以无法算出 $u$ 值. 由上面分析可见在 $H_0$ 成立的条件下有

$$U\leqslant u,$$

因而

$$\{U>\lambda\}\subset\{u>\lambda\},$$

故

$$P\{U > \lambda\} \leqslant P\{u > \lambda\} = \alpha.$$

这表明事件$\{U>\lambda\}$是概率较 $\alpha$ 更小的小概率事件. 由样本值 $x_1$, $x_2,\cdots,x_n$,算出 $\hat{U}=-2.08$. 于是我们可以做出这样的判断：若$\hat{U}\in R_\alpha$,则否定 $H_0$,否则认为 $H_0$ 相容. 例 2 中的$\hat{U}=-2.08<1.65$,即$\hat{U}\overline{\in}R_\alpha$.

**评注** 这说明小概率事件$\{U>\lambda\}$没有发生,即未发现什么不合理现象. 这时我们不能否定 $H_0$,故认为 $H_0$ 相容. 换句话说没有发现滚珠平均直径不小于等于 15.25 mm.

2. 未知方差,检验均值

关于一元正态总体当方差未知时,期望的检验程序我们不再写出,留给读者作为练习.

**例 3** 用某仪器间接测量温度,重复五次,所得数据是 1250℃,1265℃,1245℃,1260℃,1275℃,而用别的精确办法测得温度为 1277℃(可看作温度的真值),试问用此仪器间接测量温度有无系统偏差?($\alpha=0.05$)

**解** 用 $X$ 表示由这个仪器测得的数值,有 $X\sim N(\mu,\sigma^2)$,其中 $\sigma^2$ 未知,这是一个未知方差,检验均值问题.

提出零假设 $H_0$: $\mu=1277$;对于这类问题,我们选取一个包括指定数值 1277 的统计量

$$T=\frac{\bar{x}-1277}{S/\sqrt{n}},$$

其中

$$S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}.$$

在 $H_0$ 成立的条件下,$T$ 与样本函数

$$t=\frac{\bar{x}-\mu}{S/\sqrt{n}}$$

都服从 $t(n-1)$分布,即

$$T \xlongequal{\text{在 } H_0 \text{ 下}} t \sim t(n-1).$$

这里我们采取双边检验. 由检验水平 $\alpha$,选择

$$R_\alpha = \{(-\infty, \lambda_1) \cup (\lambda_2, +\infty)\}$$

且使得

$$P\{t \in (-\infty, \lambda_1)\} = P\{t \in (\lambda_2, +\infty)\} = \alpha/2,$$

即使得

$$P\{t \in R_\alpha\} = \alpha.$$

可见$\{t \in R_\alpha\}$是一个小概率事件. 由于 $t$ 分布的对称性,可知 $\lambda_2 = -\lambda_1 \xlongequal{\text{def}} \lambda$. 考虑到 $t$ 分布临界值表的构造(前面已介绍)可由 $t(n-1, \alpha)$查出 $\lambda_\alpha$ 之值.

已知 $\alpha = 0.05, n = 5$,由 $t(4, 0.05)$查 $t$ 分布临界值表,得到 $\lambda_\alpha = 2.776$. 故否定域为

$$R_\alpha = \{(-\infty, -2.776) \cup (2.776, +\infty)\}.$$

由样本值 $x_1, x_2, \cdots, x_n$ 计算

$$\bar{x} = \frac{1}{5}(1250 + \cdots + 1275) = 1259,$$

$$S^2 = \frac{1}{4}[(1250 - 1259)^2 + \cdots + (1275 - 1259)^2]$$

$$= 570 \times \frac{1}{4} = 142.5,$$

有

$$\hat{T} = \frac{1259 - 1277}{\sqrt{142.5/5}} = -3.37.$$

于是我们可以作出判断:若$\hat{T} \in R_\alpha$,则否定 $H_0$,否则认为 $H_0$ 相容. 本例中$\hat{T} = -3.37 < -2.776$,即$\hat{T} \in R_\alpha$,故结论为否定 $H_0$: $\mu = 1277$. 换句话说,该仪器间接测量温度有系统偏差.

**例 4** 在例 3 中,我们进一步问此仪器间接测量的温度是否偏低? ($\alpha = 0.05$)

**解** 用 $X$ 表示由这个仪器测得的数值，有 $X\sim N(\mu,\sigma^2)$，其中 $\sigma^2$ 未知，这是一个未知方差，检验均值问题.

提出零假设 $H_0$: $\mu\leqslant 1277$. 这里我们仍选取 $T$ 统计量

$$T=\frac{\bar{x}-1277}{S/\sqrt{n}}.$$

在 $H_0$: $\mu\leqslant 1277$ 成立的条件下，有不等式

$$\frac{\bar{x}-1277}{S/\sqrt{n}}\leqslant\frac{\bar{x}-\mu}{S/\sqrt{n}}.$$

因此 $T$ 与随机变量 $t=\dfrac{\bar{x}-\mu}{S/\sqrt{n}}$ 有如下的关系

$$T\leqslant t\sim t(n-1).$$

这里我们采取单边检验，由检验水平 $\alpha$，选择 $R_\alpha=\{\lambda,+\infty\}$，使得

$$P\{t>\lambda\}=\alpha.$$

根据 $t$ 分布数值表的构造，由 $t(n-1,2\alpha)$ 查得 $\lambda$，于是由

$$T\leqslant t$$

有

$$\{T>\lambda\}\subset\{t>\lambda\},$$

故

$$P\{T>\lambda\}\leqslant P\{t>\lambda\}=\alpha.$$

上式说明事件 $\left\{\dfrac{\bar{x}-1277}{S/\sqrt{n}}>\lambda\right\}$ 是概率比 $\alpha$ 更小的小概率事件.

由 $\alpha=0.05$，$n=5$，由 $t(4,0.10)$ 查 $t$ 分布临界值表，得到 $\lambda=2.132$，故否定域为

$$R_\alpha=\{2.132,+\infty\}.$$

由样本值算出

$$\bar{x}=1259,\quad S^2=142.5,\quad \hat{T}=-3.37.$$

由于 $\hat{T}=-3.37<2.132$，即 $\hat{T}\,\overline{\in}\,R_\alpha$，故结论为 $H_0$: $\mu\leqslant 1277$ 相容. 换句话说，没有发现此仪器间接测量温度偏高.

3. 方差的假设检验

关于一元正态总体当均值未知时，方差的检验程序：

(1) 提出零假设 $H_0$：$\sigma^2=\sigma_0^2$(或 $\sigma^2\leqslant\sigma_0^2$)；

(2) 由样本值 $x_1,x_2,\cdots,x_n$ 计算统计量

$$W=\frac{(n-1)S^2}{\sigma_0^2}$$

的数值$\hat{W}$；

(3) 对于检验水平 $\alpha$ 查 $\chi^2$ 分布临界值表 $\chi^2\left(n-1,1-\frac{\alpha}{2}\right)$ 和 $\chi^2\left(n-1,\frac{\alpha}{2}\right)$(或 $\chi^2(n-1,\alpha)$)得到临界值 $\lambda_1,\lambda_2$(或 $\lambda$)；

(4) 将$\hat{W}$与 $\lambda$ 进行比较，作出判断：当$\hat{W}<\lambda_1$ 或$\hat{W}>\lambda_2$(或$\hat{W}>\lambda$)时否定 $H_0$，否则认为 $H_0$ 相容.

**例 5** 在第七章例 7 中，问 5 至 6 岁的幼儿身高总体的方差是否为 49？($\alpha=0.05$)

**解** 用 $X$ 表示幼儿身高，有 $X\sim N(\mu,\sigma^2)$，其中 $\mu$ 未知. 这是一个未知均值，检验方差的问题.

这个问题的零假设是 $H_0$：$\sigma^2=49$. 我们选取统计量

$$W=\frac{(n-1)S^2}{49}.$$

由 $\alpha=0.05$，$n=9$，$\chi^2(8,0.975)$和 $\chi^2(8,0.025)$，查 $\chi^2$ 分布临界值表得到 $\lambda_1=2.18$，$\lambda_2=17.5$. 故否定域为

$$R_\alpha=\{(0,2.18)\cup(17.5,+\infty)\}.$$

再由样本值算出

$$\bar{x}=115,\quad S^2=55.25,\quad \hat{W}=9.02.$$

由于 $2.18<\hat{W}=9.02<17.5$，即$\hat{W}\in R_\alpha$，故结论为 $H_0$ 相容. 这就是说，没有发现身高的总体方差不等于 49.

**例 6** 问第七章例 7 中，5 至 6 岁幼儿身高的总体方差是否小于等于 49？($\alpha=0.05$)

**解** 用 $X$ 表示幼儿身高，有 $X \sim N(\mu,\sigma^2)$，其中 $\mu$ 未知，这是一个未知均值，检验方差的问题.

这个问题的零假设是 $H_0$：$\sigma^2 \leqslant 49$. 我们仍选取统计量

$$W = \frac{(n-1)S^2}{49}.$$

根据 $\alpha = 0.05$，$n = 9$，由 $\chi^2(8, 0.05)$ 查 $\chi^2$ 分布临界值表得 $\lambda = 15.5$，故否定域为

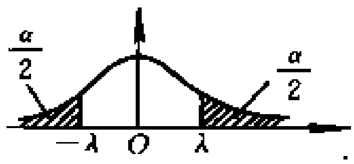
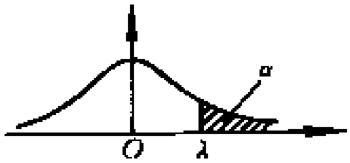
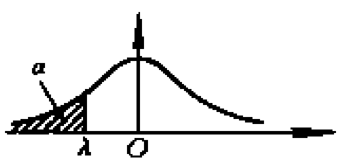
$$R_\alpha = \{15.5, +\infty\}.$$

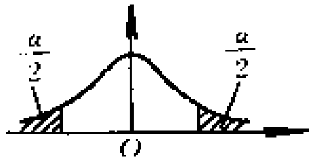
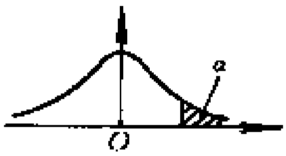
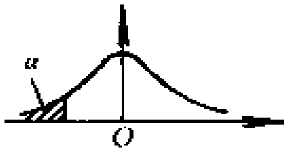
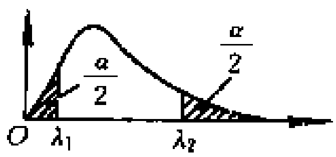
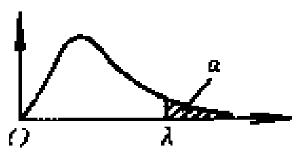
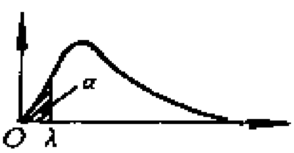
由样本值算出

$$\bar{x} = 115, \quad S^2 = 55.25, \quad \hat{W} = 9.02.$$

由于 $\hat{W} = 9.02 < 15.5$，即 $\hat{W} \in R_\alpha$，故结论为 $H_0$ 相容. 这就是说，没有发现身高的总体方差大于 49.

下面列表说明单个正态总体的均值和方差的假设检验.

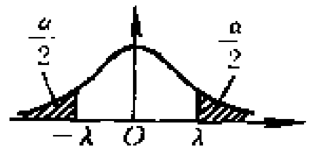
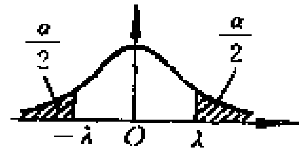
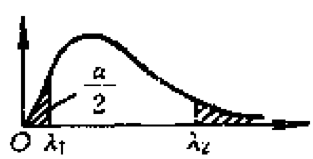
| 条件 | 假设 $H_0$ | 统计量 | 应查分布表 | 拒绝域 |
|---|---|---|---|---|
| 已知 $\sigma^2 = \sigma_0^2$ | $\mu = \mu_0$ | $U = \dfrac{\overline{X} - \mu_0}{\dfrac{\sigma_0}{\sqrt{n}}}$ | $\Phi(\lambda) = 1 - \dfrac{\alpha}{2}$ | $\lvert u \rvert > \lambda$   |
| | $\mu \leqslant \mu_0$ | | $\Phi(\lambda) = 1 - \alpha$ | $u > \lambda$   |
| | $\mu \geqslant \mu_0$ | | | $u < -\lambda$   |

| 条件 | 假设 $H_0$ | 统计量 | 应查分布表 | 拒绝域 |
|---|---|---|---|---|
| $\sigma^2$ 未知 | $\mu=\mu_0$ | $T=\dfrac{\bar{X}-\mu_0}{\dfrac{S}{\sqrt{n}}}$ 其中 $S^2$ 为样本方差 | $t(n-1,\alpha)$ | $\|t\|>\lambda$   |
| | $\mu\leqslant\mu_0$ | | $t(n-1,2\alpha)$ | $t>\lambda$   |
| | $\mu\geqslant\mu_0$ | | | $t<-\lambda$   |
| $\mu$ 未知 | $\sigma^2=\sigma_0^2$ | $W=\dfrac{(n-1)S^2}{\sigma_0^2}$ | $\chi^2\left(n-1,1-\dfrac{\alpha}{2}\right)\Rightarrow\lambda_1$ $\chi^2\left(n-1,\dfrac{\alpha}{2}\right)\Rightarrow\lambda_2$ | $0<w<\lambda_1$ 或 $w>\lambda_2$   |
| | $\sigma^2\leqslant\sigma_0^2$ | | $\chi^2(n-1,\alpha)$ | $w>\lambda$   |
| | $\sigma^2\geqslant\sigma_0^2$ | | $\chi^2(n-1,1-\alpha)$ | $0<w<\lambda$   |

其中 $u,t,w$ 分别是统计量 $U$、$T$、$W$ 所对应且具有确定分布的样本函数.

**方法 8.2**(双正态总体假设检验基本步骤)

设 $X_1,X_2,\cdots,X_{n_1}$ 为总体 $N(\mu_1,\sigma_1^2)$ 的一个样本,$Y_1,Y_2,\cdots,Y_{n_2}$ 为总体 $N(\mu_2,\sigma_2^2)$ 的一个样本,且两者相互独立,它们的均值分别为 $\bar X,\bar Y$;方差分别为 $S_1^2,S_2^2$,则关于总体均值和方差的假设检验如下表所示.

| 条件 | 假设 $H_0$ | 统计量 | 应查分布表 | 拒绝域 |
|---|---|---|---|---|
| 已知 $\sigma_1^2$ $\sigma_2^2$ | $\mu_1=\mu_2$ | $U=\dfrac{\bar X-\bar Y}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$ | $\Phi(\lambda)=1-\dfrac{\alpha}{2}$ | $\lvert\tilde U\rvert>\lambda$  |
| $\sigma_1^2=\sigma_2^2$ 但其值未知 | $\mu_1=\mu_2$ | $T=\dfrac{\bar X-\bar Y}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ 其中 $S_w^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$ | $t(n_1+n_2-2,\alpha)$ | $\lvert\tilde T\rvert>\lambda$  |
| $\mu_1,\mu_2$ 未知 | $\sigma_1^2=\sigma_2^2$ | $F=\dfrac{S_1^2}{S_2^2}$ | $f\left(n_1-1,n_2-1,1-\dfrac{\alpha}{2}\right)\Rightarrow\lambda_1$ $f\left(n_1-1,n_2-1,\dfrac{\alpha}{2}\right)\Rightarrow\lambda_2$ | $0<\tilde F<\lambda_1$ 或 $\tilde F>\lambda_2$  |

表中 $\tilde U,\tilde T,\tilde F$ 分别是统计量 $U$、$T$、$F$ 所对应且具有确定分布的样本函数.

**例 7** 为比较甲、乙两种安眠药的疗效,将 20 名患者分成两组,每组 10 人,如服药后延长的睡眠时间分别服从正态分布,其数据为(单位:小时):甲:5.5,4.6,4.4,3.4,1.9,1.6,1.1,0.8,0.1,-0.1;乙:3.7,3.4,2.0,2.0,0.8,0.7,0,-0.1,

$-0.2$，$-1.6$. 问在显著性水平 $\alpha=0.05$ 下两种药的疗效有无显著差别.

**解** 设甲药服后延长的睡眠时间 $X\sim N(\mu_1,\sigma_1^2)$，乙药服后延长的睡眠时间 $Y\sim N(\mu_2,\sigma_2^2)$，其中 $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$ 均为未知，先在 $\mu_1$，$\mu_2$ 未知的条件下检验假设 $H_0$：$\sigma_1^2=\sigma_2^2$. 所用统计量为 $F=S_1^2/S_2^2$，由题给数据得：$n_1=10,n_2=10,\overline{X}=2.33,\overline{Y}=0.75,S_1^2=4.01,S_2^2=3.2$，于是

$$F=\frac{S_1^2}{S_2^2}=1.25.$$

查 $f(9,9,0.025)$，得 $\lambda_2=4.03$，从而 $f(9,9,0.975)=\frac{1}{4.03}$，由于 $\frac{1}{4.03}<1.25<4.03$，即样本 $(X_1,\cdots,X_{10},Y_1,\cdots,Y_{10})$ 不在假设 $H_0$：$\sigma_1^2=\sigma_2^2$ 的拒绝域内，因此在显著性水平 $\alpha=0.05$ 下可认为 $\sigma_1^2=\sigma_2^2$

其次，在 $\sigma_1^2=\sigma_2^2$ 但其值未知的条件下，检验假设 $H_0'$：$\mu_1=\mu_2$，所用统计量为

$$T=\frac{\overline{X}-\overline{Y}}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}},$$

其中

$$S_w=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}=\sqrt{\frac{9\times4.01+9\times3.2}{18}}=1.899.$$

计算后得

$$T=\frac{2.33-0.75}{1.899\times\sqrt{\frac{1}{10}+\frac{1}{10}}}=1.86.$$

查 $t(18,0.05)$ 得到 $\lambda=2.101$，由于

$$|1.86|<2.101,$$

即样本$(X_1,\cdots,X_{10},Y_1,\cdots,Y_{10})$不在假设$H_0'$: $\mu_1=\mu_2$的拒绝域内，因此在显著性水平$\alpha=0.05$下可认为$\mu_1=\mu_2$

综合上述所讨论的结果，可以认为两种安眠药疗效无显著差异.

## 三、典型例题分析

### (一) 填空题

1. 设$X_1,\cdots,X_n$是来自正态总体$N(\mu,\sigma^2)$的简单随机样本，其中参数$\mu$和$\sigma$未知，记$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，$Q^2=\sum_{i=1}^{n}(X_i-\overline{X})^2$，则假设$H_0$: $\mu=0$的$t$检验使用统计量________.

答案是：$T=\frac{\overline{X}}{Q}\sqrt{n(n-1)}$.

**分析**　由题设$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{1}{n-1}Q^2$，代入公式得

$$T=\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\xlongequal{\mu_0=0}\frac{\overline{X}}{S/\sqrt{n}}=\frac{\overline{X}}{Q}\sqrt{n(n-1)}.$$

### (二) 解答题

1. 设某厂生产的一种钢索，其断裂强度$X$(kg/cm$^2$)服从正态分布$N(\mu,40^2)$. 从中选取一个容量为9的样本，得$\overline{X}=780$ kg/cm$^2$. 能否据此认为这批钢索的断裂强度为800 kg/cm$^2$($\alpha=0.05$)?

**解**　$H_0$: $\mu=800$，$H_1$: $\mu\neq800$，这里$\sigma^2=40^2$为已知，选取检验统计量$U=\frac{\overline{X}-800}{\sigma/\sqrt{n}}\sim N(0,1)$，$\alpha=0.05$，$u_{\alpha/2}=1.96$，$|U|=\left|\frac{780-800}{40/\sqrt{9}}\right|=1.5<1.96$. 故可接受$H_0$，即可认为这批钢索的断裂强度为800 kg/cm$^2$.

2. 设某次考试的考生成绩服从正态分布，从中随机地抽取36位考生的成绩，算得平均成绩为66.5分，标准差为15分，问在显

著性水平 0.05 下，是否可以认为这次考试全体考生的平均成绩为 70 分？并给出检验过程.

附表：$t$ 分布表

$$P\{t(n) \leqslant t_p(n)\} = p$$

| $t_p(n)$ \ $p$ / $n$ | 0.95 | 0.975 |
|---|---|---|
| 35 | 1.6896 | 2.0301 |
| 36 | 1.6883 | 2.0281 |

**解** 设该次考试的考生成绩为 $X$，则 $X \sim N(\mu, \sigma^2)$. 把从 $X$ 中抽取的容量为 $n$ 的样本均值记为 $\overline{X}$，样本标准差记为 $S$. 本题是在显著性水平 $\alpha = 0.05$ 下检验假设

$$H_0: \mu = 70; \quad H_1: \mu \neq 70,$$

拒绝域为

$$|t| = \frac{|\bar{x} - 70|}{S}\sqrt{n} \geqslant t_{1-\frac{\alpha}{2}}(n-1).$$

由 $n = 36, \bar{x} = 66.5, S = 15, t_{0.975}(36-1) = 2.0301$，算得

$$|t| = \frac{|66.5 - 70|\sqrt{36}}{15} = 1.4 < 2.0301.$$

所以接受假设 $H_0: \mu = 70$，即在显著性水平 0.05 下，可以认为这次考试全体考生的平均成绩为 70 分.

3. 一细纱车间纺出某种细纱支数标准差为 1.2. 从某日纺出的一批细纱中，随机地抽 16 缕进行支数测量，算得样本标准差 $S = 2.1$，问纱的均匀度有无显著变化($\alpha = 0.05$)？假定总体分布是正态的.

**解** 该日纺出纱的支数构成一个正态总体. 按题意要检验 $H_0: \sigma^2 = 1.2^2, H_1: \sigma^2 \neq 1.2^2$.

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{15 \times 2.1^2}{1.2^2} = 45.94,$$

由 $\alpha = 0.05$，自由度 $n-1 = 15$，查 $\chi^2$ 分布表得 $\chi^2_{0.975}(15) = 6.262$，

$\chi^2_{0.025}(15)=27.488$，现在 $\chi^2=45.94>27.488$，拒绝 $H_0$，因而这天细纱均匀度有显著变化.

4. 罐头的细菌含量按规定标准必须小于 62.0，现从一批罐头中抽取 9 个，检验其细菌含量，经计算得 $\bar{x}=62.5$，$S=0.3$. 问这批罐头的质量是否完全符合标准($\alpha=0.05$)?

**解** 设罐头的细菌含量 $X$ 服从正态分布，依题意检验 $H_0$: $\mu=62.0$，$H_1$: $\mu>62.0$.

$$t=\frac{\bar{x}-\mu_0}{S/\sqrt{n}}=\frac{62.5-62.0}{0.3/\sqrt{9}}=5,$$

$\alpha=0.05$，自由度 $n-1=8$，$t_\alpha(8)=1.8595$，$W=[1.8595,+\infty)$，现 $t=5\in W$，拒绝 $H_0$，即认为这批罐头的质量不符合标准.

5. 设某产品的某项质量指标服从正态分布，已知它的标准差 $\sigma=150$. 现从一批产品中随机地抽取了 26 个，测得该项指标的平均值为 1637. 问能否认为这批产品的该项指标值为 1600($\alpha=0.05$)?

**解** (1) 提出零假设：$H_0$：$\mu=1600$；

(2) 选择统计量：$U=\dfrac{\bar{x}-1600}{150/\sqrt{26}}$；

(3) 查正态分布数值表 $\Phi(x)=0.975$，得到 $\lambda=1.96$；

(4) 计算统计量的值：

$$\hat{U}=\frac{1637-1600}{150/\sqrt{26}}\approx1.258;$$

(5) 结论：$|\hat{U}|<1.96$，未落入否定域，因此不能否定这批产品该项指标为 1600.

6. 食品厂用自动装罐机装罐头食品，每罐标准重量为 500 克，每隔一定时间需要检验机器的工作情况. 现抽得 10 罐，测得其重量(单位：g)：

495，510，505，498，503，492，502，512，497，506.

假定重量 $\xi$ 服从正态分布 $N(\mu,\sigma^2)$，试问机器工作是否正常($\alpha=$

0.10)?

**解** (1) 提出零假设：$H_0: \mu=500$;

(2) 选择统计量：$T=\dfrac{\bar{x}-500}{S/\sqrt{10}}$;

(3) 查 $t(9,0.10)$，得到 $\lambda=1.833$;

(4) 计算统计量的值：

$$\hat{T}=\frac{502-500}{6.5/\sqrt{10}}\approx 0.973.$$

(5) 结论：$|\hat{T}|<1.833$，故在显著性水平 $\alpha=0.10$ 下，可以认为自动装罐机的工作正常.

7. 已知罐头番茄汁中维生素C(Vc)的含量服从正态分布.按照规定Vc的平均含量不得少于21 mg.现从一批罐头中取了17罐，算得Vc含量的平均值 $\bar{x}=23$，$S^2=3.98^2$，问该批罐头Vc的含量是否合格?

**解** (1) 提出零假设：$H_0: \mu\leqslant 21$;

(2) 选择统计量：$T=\dfrac{\bar{x}-21}{S/\sqrt{17}}$;

(3) 查 $t(16,0.10)$ 得到 $\lambda=1.746$;

(4) 计算统计量的值：

$$\hat{T}=\frac{23-21}{3.98/\sqrt{17}}\approx 2.07;$$

(5) 结论：因为 $2.07>1.746$，所以 $\hat{T}$ 落入否定域，因此该批罐头Vc的含量合格.

8. 设某市犯罪青少年的年龄构成服从正态分布，今随机地抽取9名罪犯，其年龄如下：

$$22,\ 17,\ 19,\ 25,\ 25,\ 18,\ 16,\ 23,\ 24,$$

试以95%的概率判断犯罪青少年的年龄是否为18岁.

**解** (1) 提出零假设：$H_0: \mu=18$;

(2) 选择统计量：$T=\dfrac{\bar{x}-18}{S/\sqrt{9}}$;

(3) 查 $t(8,0.05)$得到 $\lambda=2.306$;

(4) 计算统计量的值：$\hat{T}=\dfrac{21-18}{\sqrt{12.5/3}}\approx 2.55$;

(5) 结论：因为$|\hat{T}|>2.306$，即$\hat{T}$落入否定域，因此能以95%的把握推断该市犯罪少年的平均年龄不是18岁.

9. 测得两批电子器件的样品的电阻(欧)为

| | | | | | | |
|---|---|---|---|---|---|---|
| $A$ 批($x$) | 0.140 | 0.138 | 0.143 | 0.142 | 0.144 | 0.137 |
| $B$ 批($y$) | 0.135 | 0.140 | 0.142 | 0.136 | 0.138 | 0.140 |

设这两批器材的电阻值总体分别服从分布 $N(\mu_1,\sigma_1^2)$，$N(\mu_2,\sigma_2^2)$，且这两样本独立.

(1) 检验假设($\alpha=0.05$)

$$H_0:\ \sigma_1^2=\sigma_2^2,\quad H_1:\ \sigma_1^2\neq\sigma_2^2.$$

(2) 在(1)的基础上检验假设($\alpha=0.05$)

$$H_0:\ \mu_1=\mu_2,\quad H_1:\ \mu_1\neq\mu_2.$$

**解** (1) $H_0:\ \sigma_1^2=\sigma_2^2$, $H_1:\ \sigma_1^2\neq\sigma_2^2$.

$$n_1=n_2=6,\quad \alpha=0.05,\quad F_{0.025}(5,5)=7.15,$$

$$F_{0.975}(5,5)=\frac{1}{F_{0.025}(5,5)}=0.14,$$

$$\sum x_i=0.844,\quad \sum y_i=0.831,$$

$$\sum x_i^2=0.118762,\quad \sum y_i^2=0.115129,$$

$$\sum(x_i-\bar{x})^2=0.00003933,$$

$$\sum(y_i-\bar{y})^2=0.0000355,$$

$$S_1^2=\frac{1}{5}\sum(x_i-\bar{x})^2=0.000007866,$$

$$S_2^2=\frac{1}{5}\sum(y_i-\bar{y})^2=0.0000071.$$

$$F=\frac{S_1^2}{S_2^2}=\frac{0.000007866}{0.0000071}=1.1079.$$

易见 $0.14<F<7.15$,故接受 $H_0$,即认为这两个总体的方差相等.

(2) $H_0$: $\mu_1=\mu_2$, $H_1$: $\mu_1\neq\mu_2$.

$\alpha=0.05$, 自由度 $n_1+n_2-2=10$, $t_{\alpha/2}(n_1+n_2-2)=t_{0.025}(10)=2.2281$.

$$t=\frac{\bar{x}-\bar{y}}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

$$=\frac{\frac{0.844}{6}-\frac{0.831}{6}}{\sqrt{\frac{0.00003933+0.0000355}{10}}\sqrt{\frac{1}{6}+\frac{1}{6}}}$$

$$=1.3927<2.2281,$$

接受 $H_0$,即认为这两批电阻的均值无显著差异.

## 四、练习题

1. 从已知标准差 $\sigma=5.2$ 的正态总体中,抽取容量为 16 的样本,算得样本均值 $\bar{x}=27.56$. 试在显著水平 $\alpha=0.05$ 之下,检验假设 $H_0$: $\mu=26$.

2. 测定某种溶液中的水分,由它的 10 个测定值算出: $\bar{x}=0.452\%$, $S=0.037\%$. 设测定值总体服从正态分布. 试在 $\alpha=0.05$ 之下,分别检验假设:

(1) $H_0$: $\mu=0.5\%$; (2) $H_0$: $\sigma=0.04\%$.

3. 已知维尼纶纤度在正常条件下服从正态分布,且标准差 $\sigma=0.048$. 从某天产品中抽取 5 根纤维,测得其纤度为1.32,1.55,1.36,1.40,1.44. 问这一天纤度的总体标准差是否正常($\alpha=0.05$)?

4. 由经验知道某种零件重量 $X\sim N(\mu,\sigma^2)$, $\mu=15$, $\sigma^2=0.05$.

技术革新后，抽了 6 个样品，测得重量(以 g 为单位)为

$$14.7,\ 15.1,\ 14.8,\ 15.0,\ 15.2,\ 14.6.$$

已知方差不变，问平均重量是否为 15？($\alpha=0.05$)

5. 问第 4 题中零件的平均重量是否小于等于 15？($\alpha=0.01$)

6. 根据长期资料的分析，知道某种钢筋的强度服从正态分布. 今随机抽取六根钢筋进行强度试验，测得强度(以 MPa 为单位)为

$$48.5,\ 49,\ 53.5,\ 49.5,\ 56.0,\ 52.5.$$

问：能否认为该种钢筋的平均强度为 52.0？

7. 某车间生产铜丝，生产一向比较稳定，今从产品中随机抽出 10 根检查折断力，得数据如下(以 N 为单位)

$$578,\ 572,\ 570,\ 568,\ 572,\ 570,\ 570,\ 572,\ 596,\ 584.$$

问：是否可相信该车间生产的铜丝其折断力的方差为 64？

8. 在两个工厂生产的蓄电池中，分别取 10 个测量蓄电池的电容量值(单位：安培小时)，数据如下：

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $A$ 厂 | 146 | 141 | 135 | 142 | 140 | 143 | 138 | 137 | 142 | 137 |
| $B$ 厂 | 141 | 143 | 139 | 139 | 140 | 141 | 138 | 140 | 142 | 138 |

(1) 试检验两厂蓄电池的性能有无显著差异($\alpha=0.05$)？

(2) 说明检验要作哪些假定.

## 习题答案与提示

**1.** $H_0$：$\mu=26$ 相容.

**2.** 拒绝 $H_0$：$\mu=0.5\%$，$H_0$：$\sigma=0.04\%$相容.

**3.** 不正常.　**4.** $H_0$：$\mu=15$ 相容.　**5.** $H_0$：$\mu\leqslant15$ 相容.

**6.** $H_0$：$\mu=52.0$ 相容.　**7.** $H_0$：$\sigma^2=64$ 相容.

**8.** (1) 无显著差异；(2) 双正态总体，且方差相等.

## 附表 1 正态分布数值表

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

| $x$ | $\Phi(x)$ | $x$ | $\Phi(x)$ | $x$ | $\Phi(x)$ | $x$ | $\Phi(x)$ |
|---|---|---|---|---|---|---|---|
| 0.00 | 0.5000 | 0.80 | 0.7881 | 1.60 | 0.9452 | 2.35 | 0.9906 |
| 0.05 | 0.5199 | 0.85 | 0.8023 | 1.65 | 0.9505 | 2.40 | 0.9918 |
| 0.10 | 0.5398 | 0.90 | 0.8159 | 1.70 | 0.9554 | 2.45 | 0.9929 |
| 0.15 | 0.5596 | 0.95 | 0.8289 | 1.75 | 0.9599 | 2.50 | 0.9938 |
| 0.20 | 0.5793 | 1.00 | 0.8413 | 1.80 | 0.9641 | 2.55 | 0.9946 |
| 0.25 | 0.5987 | 1.05 | 0.8531 | 1.85 | 0.9678 | 2.58 | 0.9951 |
| 0.30 | 0.6179 | 1.10 | 0.8643 | 1.90 | 0.9713 | 2.60 | 0.9953 |
| 0.35 | 0.6368 | 1.15 | 0.8749 | 1.95 | 0.9744 | 2.65 | 0.9960 |
| 0.40 | 0.6554 | 1.20 | 0.8849 | 1.96 | 0.9750 | 2.70 | 0.9965 |
| 0.45 | 0.6736 | 1.25 | 0.8944 | 2.00 | 0.9772 | 2.75 | 0.9970 |
| 0.50 | 0.6915 | 1.30 | 0.9032 | 2.05 | 0.9798 | 2.80 | 0.9974 |
| 0.55 | 0.7088 | 1.35 | 0.9115 | 2.10 | 0.9821 | 2.85 | 0.9978 |
| 0.60 | 0.7257 | 1.40 | 0.9192 | 2.15 | 0.9842 | 2.90 | 0.9981 |
| 0.65 | 0.7422 | 1.45 | 0.9265 | 2.20 | 0.9861 | 2.95 | 0.9984 |
| 0.70 | 0.7580 | 1.50 | 0.9332 | 2.25 | 0.9878 | 3.00 | 0.9987 |
| 0.75 | 0.7734 | 1.55 | 0.9394 | 2.30 | 0.9893 | 4.00 | 1.0000 |

## 附表 2　$t$ 分布临界值表

| $n$ \ $\lambda$ \ $\alpha$ | 0.10 | 0.05 | 0.01 | $n$ \ $\lambda$ \ $\alpha$ | 0.10 | 0.05 | 0.01 |
|---|---|---|---|---|---|---|---|
| 1 | 6.314 | 12.706 | 63.657 | 18 | 1.734 | 2.101 | 2.878 |
| 2 | 2.920 | 4.303 | 9.925 | 19 | 1.729 | 2.093 | 2.861 |
| 3 | 2.353 | 3.182 | 5.841 | 20 | 1.725 | 2.086 | 2.845 |
| 4 | 2.132 | 2.776 | 4.604 | 21 | 1.721 | 2.080 | 2.831 |
| 5 | 2.015 | 2.571 | 4.032 | 22 | 1.717 | 2.074 | 2.819 |
| 6 | 1.943 | 2.447 | 3.707 | 23 | 1.714 | 2.069 | 2.807 |
| 7 | 1.895 | 2.365 | 3.499 | 24 | 1.711 | 2.064 | 2.797 |
| 8 | 1.860 | 2.306 | 3.355 | 25 | 1.708 | 2.060 | 2.787 |
| 9 | 1.833 | 2.262 | 3.250 | 26 | 1.706 | 2.056 | 2.779 |
| 10 | 1.812 | 2.228 | 3.169 | 27 | 1.703 | 2.052 | 2.771 |
| 11 | 1.796 | 2.201 | 3.106 | 28 | 1.701 | 2.048 | 2.763 |
| 12 | 1.782 | 2.179 | 3.055 | 29 | 1.699 | 2.045 | 2.756 |
| 13 | 1.771 | 2.160 | 3.012 | 30 | 1.697 | 2.042 | 2.750 |
| 14 | 1.761 | 2.145 | 2.977 | 40 | 1.684 | 2.021 | 2.704 |
| 15 | 1.753 | 2.131 | 2.947 | 60 | 1.671 | 2.000 | 2.660 |
| 16 | 1.746 | 2.120 | 2.921 | 120 | 1.658 | 1.980 | 2.617 |
| 17 | 1.740 | 2.110 | 2.898 | $\infty$ | 1.645 | 1.960 | 2.576 |

［注］ $n$：自由度，$\lambda$：临界值，$P\{|t|>\lambda\}=\alpha$.

## 附表 3　$\chi^2$ 分布临界值表

| $\lambda$ $\alpha$ / $n$ | 0.975 | 0.05 | 0.025 | 0.01 |
|---|---|---|---|---|
| 1 | 0.00098 | 3.84 | 5.05 | 6.63 |
| 2 | 0.0506 | 5.99 | 7.38 | 9.21 |
| 3 | 0.216 | 7.81 | 9.35 | 11.3 |
| 4 | 0.484 | 9.49 | 11.1 | 13.3 |
| 5 | 0.831 | 11.07 | 12.8 | 15.1 |
| 6 | 1.24 | 12.6 | 14.4 | 16.8 |
| 7 | 1.69 | 14.1 | 16.0 | 18.5 |
| 8 | 2.18 | 15.5 | 17.5 | 20.1 |
| 9 | 2.70 | 16.9 | 19.0 | 21.7 |
| 10 | 3.25 | 18.3 | 20.5 | 23.2 |
| 11 | 3.82 | 19.7 | 21.9 | 24.7 |
| 12 | 4.40 | 21.0 | 23.3 | 26.2 |
| 13 | 5.01 | 22.4 | 24.7 | 27.7 |
| 14 | 5.63 | 23.7 | 26.1 | 29.1 |
| 15 | 6.26 | 25.0 | 27.5 | 30.6 |
| 16 | 6.91 | 26.3 | 28.8 | 32.0 |
| 17 | 7.56 | 27.6 | 30.2 | 33.4 |
| 18 | 8.23 | 28.9 | 31.5 | 34.8 |
| 19 | 8.91 | 30.1 | 32.9 | 36.2 |
| 20 | 9.59 | 31.4 | 34.2 | 37.6 |
| 21 | 10.3 | 32.7 | 35.5 | 38.9 |
| 22 | 11.0 | 33.9 | 36.8 | 40.3 |
| 23 | 11.7 | 35.2 | 38.1 | 41.6 |
| 24 | 12.4 | 36.4 | 39.4 | 43.0 |
| 25 | 13.1 | 37.7 | 40.6 | 44.3 |
| 26 | 13.8 | 38.9 | 41.9 | 45.6 |
| 27 | 14.6 | 40.1 | 43.2 | 47.0 |
| 28 | 15.3 | 41.3 | 44.5 | 48.3 |
| 29 | 16.0 | 42.6 | 45.7 | 49.6 |
| 30 | 16.8 | 43.8 | 47.0 | 50.9 |

[注]　$n$：自由度，$\lambda$：临界值，$P\{\chi^2>\lambda\}=\alpha$.

## 附表4　$F$分布临界值表($\alpha=0.05$)

| $n_2$ \ $n_1$ ($\lambda$) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 12 | 24 | ∞ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 161.4 | 199.5 | 215.7 | 224.6 | 230.2 | 234.0 | 236.8 | 238.9 | 243.9 | 249.1 | 254.3 |
| 2 | 18.5 | 19.0 | 19.2 | 19.2 | 19.3 | 19.3 | 19.4 | 19.4 | 19.4 | 19.5 | 19.5 |
| 3 | 10.1 | 9.55 | 9.28 | 9.12 | 9.01 | 8.94 | 8.89 | 8.85 | 8.74 | 8.64 | 8.53 |
| 4 | 7.71 | 6.94 | 6.59 | 6.39 | 6.26 | 6.16 | 6.09 | 6.04 | 5.91 | 5.77 | 5.63 |
| 5 | 6.61 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.88 | 4.82 | 4.68 | 4.53 | 4.36 |
| 6 | 5.99 | 5.14 | 4.76 | 4.53 | 4.39 | 4.28 | 4.21 | 4.15 | 4.00 | 3.84 | 3.67 |
| 7 | 5.59 | 4.74 | 4.35 | 4.12 | 3.97 | 3.87 | 3.79 | 3.73 | 3.57 | 3.41 | 3.23 |
| 8 | 5.32 | 4.46 | 4.07 | 3.84 | 3.69 | 3.58 | 3.50 | 3.44 | 3.28 | 3.12 | 2.93 |
| 9 | 5.12 | 4.26 | 3.86 | 3.63 | 3.48 | 3.37 | 3.29 | 3.23 | 3.07 | 2.90 | 2.71 |
| 10 | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.14 | 3.07 | 2.91 | 2.74 | 2.54 |
| 11 | 4.84 | 3.98 | 3.59 | 3.36 | 3.20 | 3.09 | 3.01 | 2.95 | 2.79 | 2.61 | 2.40 |
| 12 | 4.75 | 3.89 | 3.49 | 3.26 | 3.11 | 3.00 | 2.91 | 2.85 | 2.69 | 2.51 | 2.30 |
| 13 | 4.67 | 3.81 | 3.41 | 3.18 | 3.03 | 2.92 | 2.83 | 2.77 | 2.60 | 2.42 | 2.21 |
| 14 | 4.60 | 3.74 | 3.34 | 3.11 | 2.96 | 2.85 | 2.76 | 2.70 | 2.53 | 2.35 | 2.13 |
| 15 | 4.54 | 3.68 | 3.29 | 3.06 | 2.90 | 2.79 | 2.71 | 2.64 | 2.48 | 2.29 | 2.07 |
| 16 | 4.49 | 3.63 | 3.24 | 3.01 | 2.85 | 2.74 | 2.66 | 2.59 | 2.42 | 2.24 | 2.01 |
| 17 | 4.45 | 3.59 | 3.20 | 2.96 | 2.81 | 2.70 | 2.61 | 2.55 | 2.38 | 2.19 | 1.96 |
| 18 | 4.41 | 3.55 | 3.16 | 2.93 | 2.77 | 2.66 | 2.58 | 2.51 | 2.34 | 2.15 | 1.92 |
| 19 | 4.38 | 3.52 | 3.13 | 2.90 | 2.74 | 2.63 | 2.54 | 2.48 | 2.31 | 2.11 | 1.88 |
| 20 | 4.35 | 3.49 | 3.10 | 2.87 | 2.71 | 2.60 | 2.51 | 2.45 | 2.28 | 2.08 | 1.84 |
| 21 | 4.32 | 3.47 | 3.07 | 2.84 | 2.68 | 2.57 | 2.49 | 2.42 | 2.25 | 2.05 | 1.81 |
| 22 | 4.30 | 3.44 | 3.05 | 2.82 | 2.66 | 2.55 | 2.46 | 2.40 | 2.23 | 2.03 | 1.78 |
| 23 | 4.28 | 3.42 | 3.03 | 2.80 | 2.64 | 2.53 | 2.44 | 2.37 | 2.20 | 2.01 | 1.76 |
| 24 | 4.26 | 3.40 | 3.01 | 2.78 | 2.62 | 2.51 | 2.42 | 2.36 | 2.18 | 1.98 | 1.73 |
| 25 | 4.24 | 3.39 | 2.99 | 2.76 | 2.60 | 2.49 | 2.40 | 2.34 | 2.16 | 1.96 | 1.71 |
| 26 | 4.23 | 3.37 | 2.98 | 2.74 | 2.59 | 2.47 | 2.39 | 2.32 | 2.15 | 1.95 | 1.69 |
| 27 | 4.21 | 3.35 | 2.96 | 2.73 | 2.57 | 2.46 | 2.37 | 2.31 | 2.13 | 1.93 | 1.67 |
| 28 | 4.20 | 3.34 | 2.95 | 2.71 | 2.56 | 2.45 | 2.36 | 2.29 | 2.12 | 1.91 | 1.65 |
| 29 | 4.18 | 3.33 | 2.93 | 2.70 | 2.55 | 2.43 | 2.35 | 2.28 | 2.10 | 1.90 | 1.64 |
| 30 | 4.17 | 3.32 | 2.92 | 2.69 | 2.53 | 2.42 | 2.33 | 2.27 | 2.09 | 1.89 | 1.62 |
| 40 | 4.08 | 3.23 | 2.84 | 2.61 | 2.45 | 2.34 | 2.25 | 2.18 | 2.00 | 1.79 | 1.51 |
| 60 | 4.00 | 3.15 | 2.76 | 2.53 | 2.37 | 2.25 | 2.17 | 2.10 | 1.92 | 1.70 | 1.39 |
| 120 | 3.92 | 3.07 | 2.68 | 2.45 | 2.29 | 2.17 | 2.09 | 2.02 | 1.83 | 1.61 | 1.25 |
| ∞ | 3.84 | 3.00 | 2.60 | 2.37 | 2.21 | 2.10 | 2.01 | 1.94 | 1.75 | 1.52 | 1.00 |

[注]　表中$n_1$是第一自由度(分子的自由度);

$n_2$是第二自由度(分母的自由度);

$\lambda$是临界值,$P\{F>\lambda\}=\alpha=0.05$.

## 附表5 $F$分布临界值表($\alpha=0.025$)

| $n_2$ \ $n_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 12 | 24 | ∞ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 648.8 | 799.5 | 864.2 | 899.6 | 921.8 | 937.1 | 948.2 | 956.7 | 976.7 | 997.2 | 1018.3 |
| 2 | 38.51 | 39.00 | 39.17 | 39.25 | 39.30 | 39.33 | 39.36 | 39.37 | 39.41 | 39.46 | 39.5 |
| 3 | 17.44 | 16.04 | 15.44 | 15.10 | 14.88 | 14.73 | 14.62 | 14.54 | 14.34 | 14.12 | 13.9 |
| 4 | 12.22 | 10.65 | 9.98 | 9.60 | 9.36 | 9.20 | 9.07 | 8.98 | 8.75 | 8.51 | 8.26 |
| 5 | 10.01 | 8.43 | 7.76 | 7.39 | 7.15 | 6.98 | 6.85 | 6.76 | 6.52 | 6.28 | 6.02 |
| 6 | 8.81 | 7.26 | 6.60 | 6.23 | 5.99 | 5.82 | 5.70 | 5.60 | 5.37 | 5.12 | 4.85 |
| 7 | 8.07 | 6.54 | 5.89 | 5.52 | 5.29 | 5.12 | 4.99 | 4.90 | 4.67 | 4.42 | 4.14 |
| 8 | 7.57 | 6.06 | 5.42 | 5.05 | 4.82 | 4.65 | 4.53 | 4.43 | 4.20 | 3.95 | 3.67 |
| 9 | 7.21 | 5.71 | 5.08 | 4.72 | 4.48 | 4.32 | 4.20 | 4.10 | 3.87 | 3.61 | 3.33 |
| 10 | 6.94 | 5.46 | 4.83 | 4.47 | 4.24 | 4.07 | 3.95 | 3.85 | 3.62 | 3.37 | 3.08 |
| 11 | 6.72 | 5.26 | 4.63 | 4.28 | 4.04 | 3.88 | 3.76 | 3.66 | 3.43 | 3.17 | 2.88 |
| 12 | 6.55 | 5.10 | 4.47 | 4.12 | 3.89 | 3.73 | 3.61 | 3.51 | 3.28 | 3.02 | 2.73 |
| 13 | 6.41 | 4.97 | 4.35 | 4.00 | 3.77 | 3.60 | 3.48 | 3.39 | 3.15 | 2.89 | 2.60 |
| 14 | 6.30 | 4.86 | 4.24 | 3.89 | 3.66 | 3.50 | 3.38 | 3.29 | 3.05 | 2.79 | 2.49 |
| 15 | 6.20 | 4.77 | 4.15 | 3.80 | 3.58 | 3.41 | 3.29 | 3.20 | 2.96 | 2.70 | 2.40 |
| 16 | 6.12 | 4.69 | 4.08 | 3.73 | 3.50 | 3.34 | 3.22 | 3.12 | 2.89 | 2.63 | 2.32 |
| 17 | 6.04 | 4.62 | 4.01 | 3.66 | 3.44 | 3.28 | 3.16 | 3.06 | 2.82 | 2.56 | 2.25 |
| 18 | 5.98 | 4.56 | 3.95 | 3.61 | 3.38 | 3.22 | 3.10 | 3.01 | 2.77 | 2.50 | 2.19 |
| 19 | 5.92 | 4.51 | 3.90 | 3.56 | 3.33 | 3.17 | 3.05 | 2.96 | 2.72 | 2.45 | 2.13 |
| 20 | 5.87 | 4.46 | 3.86 | 3.51 | 3.29 | 3.13 | 3.01 | 2.91 | 2.68 | 2.41 | 2.09 |
| 21 | 5.83 | 4.42 | 3.82 | 3.48 | 3.25 | 3.09 | 2.97 | 2.87 | 2.64 | 2.37 | 2.04 |
| 22 | 5.79 | 4.38 | 3.78 | 3.44 | 3.22 | 3.05 | 2.93 | 2.84 | 2.60 | 2.33 | 2.00 |
| 23 | 5.75 | 4.35 | 3.75 | 3.41 | 3.18 | 3.02 | 2.90 | 2.81 | 2.57 | 2.30 | 1.97 |
| 24 | 5.72 | 4.32 | 3.72 | 3.38 | 3.15 | 2.99 | 2.87 | 2.78 | 2.54 | 2.27 | 1.94 |
| 25 | 5.69 | 4.29 | 3.69 | 3.35 | 3.13 | 2.97 | 2.85 | 2.75 | 2.51 | 2.24 | 1.91 |
| 26 | 5.66 | 4.27 | 3.67 | 3.33 | 3.10 | 2.94 | 2.82 | 2.73 | 2.49 | 2.22 | 1.88 |
| 27 | 5.63 | 4.24 | 3.65 | 3.31 | 3.08 | 2.92 | 2.80 | 2.71 | 2.47 | 2.19 | 1.85 |
| 28 | 5.61 | 4.22 | 3.63 | 3.29 | 3.06 | 2.90 | 2.78 | 2.69 | 2.45 | 2.17 | 1.83 |
| 29 | 5.59 | 4.20 | 3.61 | 3.27 | 3.04 | 2.88 | 2.76 | 2.67 | 2.43 | 2.15 | 1.81 |
| 30 | 5.57 | 4.18 | 3.59 | 3.25 | 3.03 | 2.87 | 2.75 | 2.65 | 2.41 | 2.14 | 1.79 |
| 40 | 5.42 | 4.05 | 3.46 | 3.13 | 2.90 | 2.74 | 2.62 | 2.53 | 2.29 | 2.01 | 1.64 |
| 60 | 5.29 | 3.93 | 3.34 | 3.01 | 2.79 | 2.63 | 2.51 | 2.41 | 2.17 | 1.88 | 1.48 |
| 120 | 5.15 | 3.80 | 3.23 | 2.89 | 2.67 | 2.52 | 2.39 | 2.30 | 2.05 | 1.76 | 1.31 |
| ∞ | 5.02 | 3.69 | 3.12 | 2.79 | 2.57 | 2.41 | 2.29 | 2.19 | 1.94 | 1.64 | 1.00 |

[注] 表中$n_1$是第一自由度(分子的自由度);

$n_2$是第二自由度(分母的自由度);

$\lambda$是临界值,$P\{F>\lambda\}=\alpha=0.025$.

## 附表6　$F$ 分布临界值表($\alpha=0.01$)

| $\lambda$ $n_1$ / $n_2$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 12 | 24 | $\infty$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 4052 | 4999 | 5403 | 5625 | 5764 | 5859 | 5928 | 5982 | 6106 | 6234 | 6366 |
| 2 | 98.5 | 99.0 | 99.2 | 99.2 | 99.3 | 99.3 | 99.4 | 99.4 | 99.4 | 99.5 | 99.5 |
| 3 | 34.1 | 30.8 | 29.5 | 28.7 | 28.2 | 27.9 | 27.7 | 27.5 | 27.1 | 26.6 | 26.1 |
| 4 | 21.2 | 18.0 | 16.7 | 16.0 | 15.5 | 15.2 | 15.0 | 14.8 | 14.4 | 13.9 | 13.5 |
| 5 | 16.3 | 13.3 | 12.1 | 11.4 | 11.0 | 10.7 | 10.5 | 10.3 | 9.89 | 9.47 | 9.02 |
| 6 | 13.7 | 10.9 | 9.78 | 9.15 | 8.75 | 8.47 | 8.26 | 8.10 | 7.72 | 7.31 | 6.88 |
| 7 | 12.2 | 9.55 | 8.45 | 7.85 | 7.46 | 7.19 | 6.99 | 6.84 | 6.47 | 6.07 | 5.65 |
| 8 | 11.3 | 8.65 | 7.59 | 7.01 | 6.63 | 6.37 | 6.18 | 6.03 | 5.67 | 5.28 | 4.86 |
| 9 | 10.6 | 8.02 | 6.99 | 6.42 | 4.06 | 5.80 | 5.61 | 5.47 | 5.11 | 4.73 | 4.31 |
| 10 | 10.0 | 7.56 | 6.55 | 5.99 | 5.64 | 5.39 | 5.20 | 5.06 | 4.71 | 4.33 | 3.91 |
| 11 | 9.65 | 7.21 | 6.22 | 5.67 | 5.32 | 5.07 | 4.89 | 4.74 | 4.40 | 4.02 | 3.60 |
| 12 | 9.33 | 6.93 | 5.95 | 5.41 | 5.06 | 4.82 | 4.64 | 4.50 | 4.16 | 3.78 | 3.36 |
| 13 | 9.07 | 6.70 | 5.74 | 5.21 | 4.86 | 4.62 | 4.44 | 4.30 | 3.96 | 3.59 | 3.17 |
| 14 | 8.86 | 6.51 | 5.56 | 5.04 | 4.69 | 4.46 | 4.28 | 4.14 | 3.80 | 3.43 | 3.00 |
| 15 | 8.68 | 6.36 | 5.42 | 4.89 | 4.56 | 4.32 | 4.14 | 4.00 | 3.67 | 3.20 | 2.87 |
| 16 | 8.53 | 6.23 | 5.29 | 4.77 | 4.44 | 4.20 | 4.03 | 3.89 | 3.55 | 3.18 | 2.75 |
| 17 | 8.40 | 6.11 | 5.18 | 4.67 | 4.34 | 4.10 | 3.93 | 3.79 | 3.46 | 3.08 | 2.65 |
| 18 | 8.29 | 6.01 | 5.09 | 4.58 | 4.25 | 4.01 | 3.84 | 3.71 | 3.37 | 3.00 | 2.57 |
| 19 | 8.18 | 5.93 | 5.01 | 4.50 | 4.17 | 3.94 | 3.77 | 3.63 | 3.30 | 2.92 | 2.49 |
| 20 | 8.10 | 5.85 | 4.94 | 4.43 | 4.10 | 3.87 | 3.70 | 3.56 | 3.23 | 2.86 | 2.42 |
| 21 | 8.02 | 5.78 | 4.87 | 4.37 | 4.04 | 3.81 | 3.64 | 3.51 | 3.17 | 2.80 | 2.36 |
| 22 | 7.95 | 5.72 | 4.82 | 4.31 | 3.99 | 3.76 | 3.59 | 3.45 | 3.12 | 2.75 | 2.31 |
| 23 | 7.88 | 5.66 | 4.76 | 4.26 | 3.94 | 3.71 | 3.54 | 3.41 | 3.07 | 2.70 | 2.26 |
| 24 | 7.82 | 5.61 | 4.72 | 4.22 | 3.90 | 3.67 | 3.50 | 3.36 | 3.03 | 2.66 | 2.21 |
| 25 | 7.77 | 5.57 | 4.68 | 4.18 | 3.85 | 3.63 | 3.46 | 3.32 | 2.99 | 2.62 | 2.17 |
| 26 | 7.72 | 5.53 | 4.64 | 4.14 | 3.82 | 3.59 | 3.42 | 3.29 | 2.96 | 2.58 | 2.13 |
| 27 | 7.68 | 5.49 | 4.60 | 4.11 | 3.78 | 3.56 | 3.39 | 3.26 | 2.93 | 2.55 | 2.10 |
| 28 | 7.64 | 5.45 | 4.57 | 4.07 | 3.75 | 3.53 | 3.36 | 3.23 | 2.90 | 2.52 | 2.06 |
| 29 | 7.60 | 5.42 | 4.54 | 4.04 | 3.73 | 3.50 | 3.33 | 3.20 | 2.87 | 2.49 | 2.03 |
| 30 | 7.56 | 5.39 | 4.51 | 4.02 | 3.70 | 3.47 | 3.30 | 3.17 | 2.84 | 2.47 | 2.01 |
| 40 | 7.31 | 5.18 | 4.31 | 3.83 | 3.51 | 3.29 | 3.12 | 2.99 | 2.66 | 2.29 | 1.80 |
| 60 | 7.08 | 4.98 | 4.13 | 3.65 | 3.34 | 3.12 | 2.95 | 2.82 | 2.50 | 2.12 | 1.60 |
| 120 | 6.85 | 4.79 | 3.95 | 3.48 | 3.17 | 2.96 | 2.79 | 2.66 | 2.34 | 1.95 | 1.38 |
| $\infty$ | 6.63 | 4.61 | 3.78 | 3.32 | 3.02 | 2.80 | 2.64 | 2.51 | 2.18 | 1.79 | 1.00 |

［注］ 表中$n_1$是第一自由度(分子的自由度)；

$n_2$是第二自由度(分母的自由度)；

$\lambda$是临界值，$P\{F>\lambda\}=\alpha=0.01$.

责任编辑 刘 勇 封面设计 林胜利

2001年研究生入学考试应试指导丛书

2001

2001

Entrance Exams for MD

## 公共课系列

1. 政治理论应试指导
2. 政治理论模拟试卷
3. 数学应试指导（工学类）
4. 数学模拟试卷（工学类）
5. 数学应试指导（经济学类）
6. 微积分
7. 线性代数
8. 概率论与数理统计
9. 数学模拟试卷（经济学类）
10. 英语应试指导与模拟试题
11. 英语完形填空
12. 英语语法与词汇
13. 英语阅读翻译
14. 英语写作
15. 时事政治应试指导

## 法律硕士联考系列

1. 刑法应试指导
2. 刑法模拟试卷
3. 民法应试指导
4. 民法模拟试卷
5. 综合考试应试指导
6. 综合考试模拟试卷

## MBA 联考系列

1. 管理应试指导与模拟试题
2. 英语应试指导与模拟试题
3. 数学应试指导与模拟试题
4. 语文应试指导与模拟试题
5. 逻辑应试指导与模拟试题

ISBN 7-301-04480-1

9 787301 044803 >

ISBN 7-301-04480-1/G·564 定价：14.50元