

• 成人教育数学辅导系列

概率论与数理统计**攻关**

上海交通大学数学系 编

上海交通大学出版社

内 容 提 要

本书共编选了概率论与数理统计习题 350 多题(均选自重点大学成人教育院校的教材)。其中“例题精解”180 多题,均有详解,有的题给出多种解法,对典型题或难题,还专门作分析、点评;其余“习题精选”170 余题,除标“*”号给出详解外都给出了答案或提示。

附录中收编了几家重点大学成人教育院校近年来本科生概率论与数理统计试卷五份,附答案与提示。

本书适合成人教育理、工、农、医科,经济管理和财经类各专业本、专科生阅读,也可作为教师的教学参考书。

前 言

众所周知,大学数学是理、工、农、医、管理等各科各专业共同的基础课,其重要性不言而喻。由于数学本身抽象深奥,使不少学生对学数学有畏惧感,特别是数学考试总有人过不了关,因而数学课历来被学生称为“霸王课”。

怎样才能学好数学呢?数学界的名师、学业有成的学子对此回答不尽相同:

“首先要喜欢数学,喜欢了,自然会下功夫去学好。”

“学数学要勤奋,要多动脑、多动手、多动口。”

“课前要预习,课上认真听,课后要复习,作业必须独立完成。”

上述回答有共同之处:学数学要花时间,要多做习题。我国著名数学家苏步青院士在求学期间就曾做过一万余道微积分题。多做习题自然要多花很多时间,这对成人教育学院的学生来讲难度较大:他们白天工作,夜晚上课,有的学生可能还要照顾家庭、子女,能用来做习题的时间实在不多。若遇难题,无处请教,宝贵的时间在苦思冥想中悄悄流失。向书本请教,也不失为一个好方法,但又难以觅到与教科书配套的辅导书。

呈现在眼前的这套成人教育数学辅导系列丛书,就是作为全国工科数学教学基地的上海交通大学数学系,专为接受成人高等教育的学生而组织编写的,作者都是教学第一线的老教师。他们根据数学教学大纲(本科非数学专业)的要求,精心编选了数百道习题,并对其中一半给予精解,即不但给出解题思路与方法(有的题给出多种解法),还对难点与易错之处进行分析、点评。通过对各种含不同知识点的典型例题的剖析,使读者加深对本课程基本概念、基本理论和基本方法的理解与掌握。每章的“习题精选”供读者练习之用,均给出了答案或提示。附录中收编了几家重点大学成人教育院校近年的试卷,供读者在复习迎

考时作练兵之用(附答案)。

由于非数学专业的每门数学课时都比较少,课堂上教师举例讲解的时间非常有限,所以这套成人教育数学辅导系列丛书,既是对课时不足的一种补偿,也是对学生的课外辅导。编者期望本丛书能使读者用不太多的时间比较扎实地掌握相关课程的基本知识,提高解题能力,攻下考试难关。

本系列丛书由《高等数学攻关》、《线性代数攻关》和《概率论与数理统计攻关》组成。其中《概率论与数理统计攻关》,前四章由冯卫国副教授编写,后两章由武爱文副教授编写,最后由冯卫国统稿。

随着社会和科学技术的发展,人们越来越重视对随机现象规律性的研究。例如,如果你投资股票、彩票、期货等,那么你必定要考虑投资的收益及风险问题;如果你是企业管理者,必定要考虑如何提高产品质量,使产品为顾客喜爱,使企业效益达到最大,等等。所以,掌握一些随机现象理论和研究方法,对你的日常工作和生活将带来很大益处。

概率论与数理统计是研究随机现象统计规律性的一门数学学科,学好这门课对你来说显然十分重要。由于本课程是随机数学,区别于其他传统数学,所以在学本课程时要特别注意它的特点。本书的概率统计习题大部分是从实际生活中提炼出来的,既真实、生动,又具有很强的实用性,因此,除了做好习题以外,还需要思考其本质问题,做到举一反三。

限于编者水平,难免有疏漏与不当之处,敬请同行与读者不吝赐教。

编 者
于上海交通大学
2005 年 8 月

目 录

第 1 章 概率论的基本概念	1
A 例题精解	3
B 习题精选	31
答案与提示	34
第 2 章 随机变量及其分布	38
A 例题精解	40
B 习题精选	64
答案与提示	69
第 3 章 多维随机变量及其分布	73
A 例题精解	76
B 习题精选	109
答案与提示	114
第 4 章 随机变量的数字特征	121
A 例题精解	123
B 习题精选	151
答案与提示	156
第 5 章 大数定律和中心极限定理	161
A 例题精解	164
B 习题精选	179
答案与提示	180

第 6 章 数理统计初步.....	184
A 例题精解	193
B 习题精选	213
答案与提示.....	218
附录 实战试卷及答案与提示.....	225
试卷(一).....	225
试卷(二).....	230
试卷(三).....	235
试卷(四).....	241
试卷(五).....	245

关键词

随机试验 随机事件 概率 古典概型 几何概型 随机事件的独立性 条件概率

教学要求

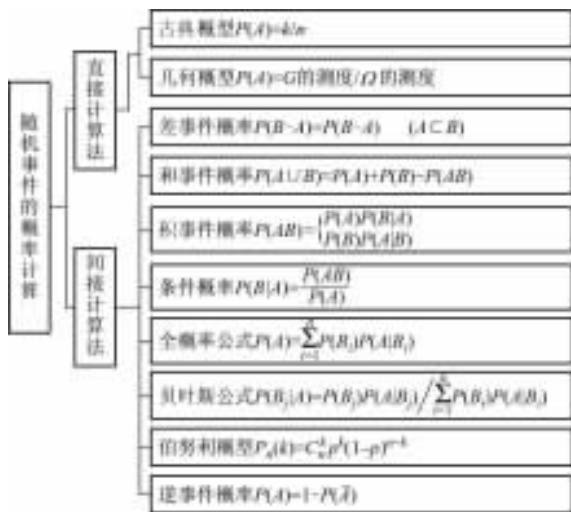
- (1) 理解随机试验、样本空间和随机事件的概念.
- (2) 熟练掌握随机事件之间的关系和运算.
- (3) 理解概率的定义,掌握概率的基本性质并能用来计算随机事件的概率.
- (4) 会计算古典概型和几何概型的概率.
- (5) 理解随机事件独立性概念,会利用事件独立性进行概率计算.
- (6) 理解条件概率的概念,会利用乘法公式、全概率公式和贝叶斯公式进行概率计算.

重点与难点

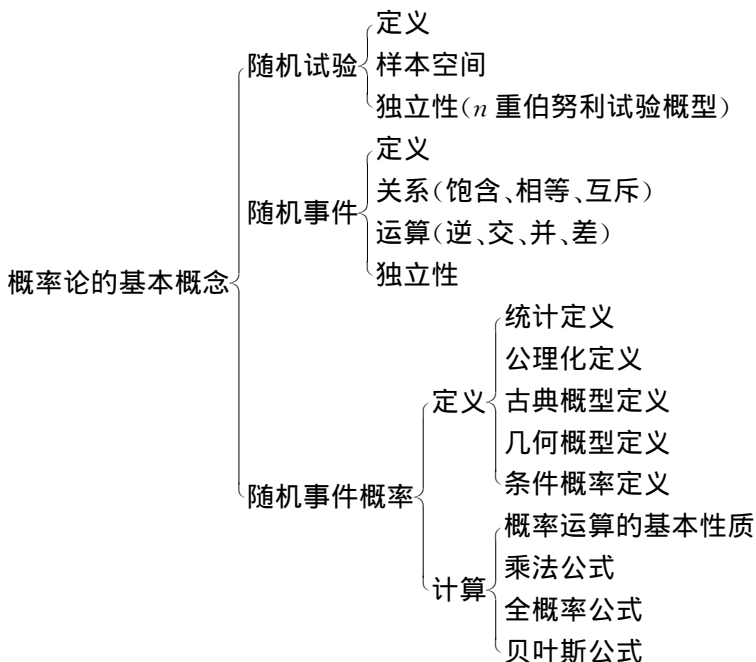
概率的计算是本章的重点与难点,关键在于会判别概率的各种类型,然后选择相应的公式进行计算.

方法提要

见下面框图:



本章知识结构



A 例题精解

【1-1】 一盒中装有 5 个同样的零件,其中编号为 1,2,3 的是正品零件,编号为 4,5 的是次品零件. 现从盒中先后取出 2 个零件,试写出:

- (1) 该试验的样本空间;
- (2) 取出的第一个零件是次品的事件 A ;
- (3) 取出的第一个零件是正品而第二个是次品的事件 B ;
- (4) 取出的 2 个零件均为次品的事件 C .

分析 随机试验的全体样本点的集合就是样本空间,而随机事件可表述为样本空间中的某个子集合. 本题的关键是要先找出全部样本点,而不能有一个遗漏.

解 设样本点 (i, j) 表示取出的第一个零件为 i 号,第二个零件为 j 号.

- (1) 该试验的样本空间

$$\Omega = \left\{ \begin{array}{cccc} (1,2) & (1,3) & (1,4) & (1,5) \\ (2,1) & (2,3) & (2,4) & (2,5) \\ (3,1) & (3,2) & (3,4) & (3,5) \\ (4,1) & (4,2) & (4,3) & (4,5) \\ (5,1) & (5,2) & (5,3) & (5,4) \end{array} \right\},$$

它所包含的样本点共有

$$C_5^1 C_4^1 = 20 \text{ 个.}$$

- (2)

$$A = \left\{ \begin{array}{cccc} (4,1) & (4,2) & (4,3) & (4,5) \\ (5,1) & (5,2) & (5,3) & (5,4) \end{array} \right\},$$

它所包含的样本点共有

$$C_2^1 C_4^1 = 8 \text{ 个.}$$

- (3)

$$B = \left\{ \begin{array}{ccc} (1,4) & (1,5) & (2,4) \\ (2,5) & (3,4) & (3,5) \end{array} \right\},$$

它所包含的样本点共有

$$C_3^1 C_2^1 = 6 \text{ 个.}$$

(4) $C = \{(4, 5), (5, 4)\}$, 仅含 $C_2^1 = 2$ 个样本点.

【1-2】 某学生做 3 道习题. A, B, C 分别表示一、二、三题做对. 试用 A, B, C 的运算关系表示下列事件:

- (1) 第一题做对, 第二、三题做错;
- (2) 第一、二题做对, 第三题做错;
- (3) 3 题都做对;
- (4) 3 题都做错;
- (5) 3 题中至少有 1 题做对;
- (6) 3 题中至少有 2 题做对;
- (7) 3 题中不多于 1 题做对;
- (8) 3 题中恰有 2 题做对.

分析 由于事件有各种运算关系, 所以答案形式不唯一.

解 (1) $\overline{A}\overline{B}\overline{C}$ 或 $A - B - C$;

(2) $AB\overline{C}$ 或 $AB - C$;

(3) ABC ;

(4) $\overline{A}\overline{B}\overline{C}$;

(5) $A \cup B \cup C$;

(6) $AB \cup AC \cup BC$;

(7) $\overline{A}\overline{B} \cup \overline{A}\overline{C} \cup \overline{B}\overline{C}$ (3 题中至少有 2 题做错)

或 $\overline{AB} \overline{AC} \overline{BC}$ (3 题中任 2 题都做错)

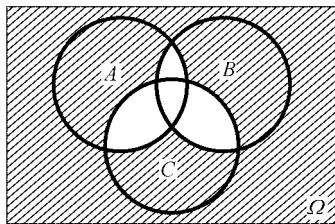
或 $\overline{A}\overline{B}\overline{C} \cup \overline{A}B\overline{C} \cup \overline{A}\overline{B}C \cup A\overline{B}\overline{C}$ (3 题中恰有 1 题做对或都做错);

(8) $AB\overline{C} \cup A\overline{B}C \cup \overline{A}BC$

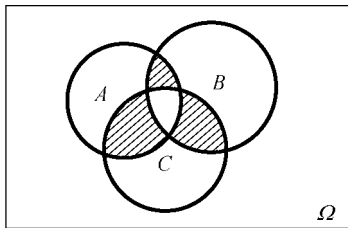
或 $AB \cup AC \cup BC - ABC$.

点评 题(4)和题(5)两个事件是对立事件, 故题(5)可表示成 $\overline{\overline{A}\overline{B}\overline{C}}$, 但不能表示成 $\Omega - \overline{A}\overline{B}\overline{C}$, 因为后一表示法不符合题目要求. 题(6)和题(7)也是对立事件.

若利用文氏图求解题(7), 题(8)更加直观方便(见图 1-1).



题(7)



题(8)

图 1-1

【1-3】 设 x 表示一个在数轴上某一段 $[-7, 8]$ 中作随机运动的质点的位置. 试说明下列各事件的关系:

$$A = \{x \mid -1 < x \leq 5\}, \quad B = \{x \mid 0 \leq x \leq 3\},$$

$$C = \{x \mid -7 \leq x \leq -1, 5 < x \leq 8\},$$

并写出下列事件:

$$(1) A(\overline{B \cup C}); \quad (2) \overline{(A - B)C}.$$

解 试验的样本空间

$$\Omega = \{x \mid -7 \leq x \leq 8\},$$

各事件的情况如图 1-2 所示.

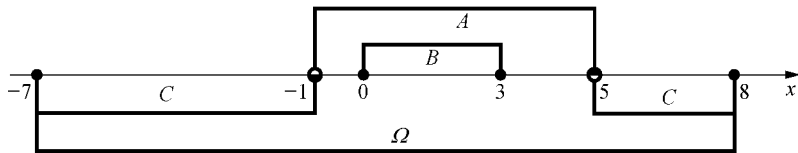


图 1-2

由图 1-2 可见, $B \subset A$; A 与 C , B 与 C 互不相容, 即 $AC = \emptyset$, $BC = \emptyset$; A 与 C 为对立事件, 即 $AC = \emptyset$, $A \cup C = \Omega$.

$$\begin{aligned} (1) A(\overline{B \cup C}) &= A\overline{B} \cup A\overline{C} = A\overline{B} = A - B \\ &= \{x \mid -1 < x < 0, 3 < x \leq 5\}; \end{aligned}$$

$$\begin{aligned} (2) \overline{(A - B)C} &= \overline{A\overline{B}C} = (\overline{A} \cup B)C = \overline{A}C \cup BC \\ &= C \cup \emptyset = C = \{x \mid -7 \leq x \leq -1, 5 < x \leq 8\}. \end{aligned}$$

【1-4】 在船舶与海洋工程系的学生中任选一名学生. 若事件 A 表示被选学生是男生, 事件 B 表示该学生是三年级学生, 事件 C 表示

该学生是运动员,则

- (1) 叙述事件 ABC 的意义.
- (2) 在什么条件下 $ABC=C$ 成立?
- (3) 什么时候关系式 $C\subset B$ 是正确的?
- (4) 什么时候 $\bar{A}=B$ 成立?

解 (1) 事件 ABC 表示挑选的是一名三年级男生,但不是运动员.

(2) 等式 $ABC=C$ 等价于 $C\subset AB$,即在“全系运动员都是三年级男生”这一条件下,等式 $ABC=C$ 成立.

(3) 在全系运动员都是三年级学生时, $C\subset B$ 是正确的.

(4) 当全系女生都在三年级,且三年级学生都是女生时 $\bar{A}=B$ 成立.

【1-5】 设 A, B 为随机事件,证明:

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

证 方法1 等价法

事件 $A - AB$ 等价于“ A 发生,而 A, B 不同时发生”等价于“ A 发生而 B 不发生”,从而事件 $(A - AB) \cup B$ 等价于“ A, B 中至少有一个发生”,即

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

方法2 推演法

$$\begin{aligned}(A - AB) \cup B &= A\bar{A}\bar{B} \cup B = A(\bar{A} \cup \bar{B}) \cup B = A\bar{A} \cup A\bar{B} \cup B \\&= A\bar{B} \cup B = A\bar{B} \cup B\bar{B} \cup AB \cup B \\&= (A \cup B)(\bar{B} \cup B) = A \cup B.\end{aligned}$$

点评 不能把数的运算律,如去括号、移项等用到事件的运算中去,否则会出错.例如:

① $(A - AB) \cup B \neq A - AB \cup B$, 因为运算顺序是先并后差;

② 不能由 $A\bar{B} \cup B = A\bar{B} \cup B\bar{B} \cup AB \cup B$ 移项 B , 推出

$$A\bar{B} = A\bar{B} \cup B\bar{B} \cup AB,$$

显然 $A\bar{B} \cup B\bar{B} \cup AB = A\bar{B} \cup AB = A \neq A\bar{B}$.

方法3 图示法

如图 1-3 所示.

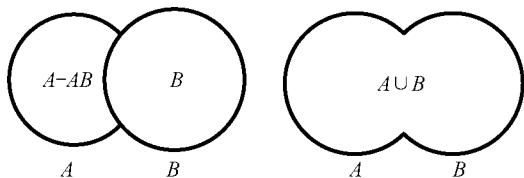


图 1-3

【1-6】 设事件 A, B 互不相容, 且 $P(A)=0.1, P(A \cup B)=0.7$. 求 $P(B), P(\overline{AB}), P(\overline{A}B)$ 和 $P(\overline{A}\overline{B})$.

解 由题设得 $P(AB)=P(\emptyset)=0$, 从而有

$$P(A \cup B) = P(A) + P(B) = 0.1 + P(B) = 0.7,$$

故

$$P(B) = 0.6.$$

$$\begin{aligned} P(\overline{AB}) &= P(A - B) = P(A - AB) = P(A) - P(AB) \\ &= P(A) = 0.1. \end{aligned}$$

同理

$$\begin{aligned} P(\overline{A}B) &= P(B - A) = P(B - AB) = P(B) - P(AB) \\ &= P(B) = 0.6. \end{aligned}$$

$$P(\overline{A}\overline{B}) = 1 - P(\overline{AB}) = 1 - P(A \cup B) = 1 - 0.7 = 0.3.$$

【1-7】 设事件 A, B 满足 $P(A)=0.7, P(A-B)=0.3$. 求 $P(\overline{AB})$.

解 方法 1

$$\begin{aligned} P(\overline{AB}) &= P(\overline{A} \cup \overline{B}) = P(\overline{A}) + P(\overline{B}) - P(\overline{A}\overline{B}) \\ &= 1 - P(A) + P(\overline{B} - \overline{A}\overline{B}) = 0.3 + P(\overline{B} - \overline{A}) \\ &= 0.3 + P(\overline{B}A) = 0.3 + P(A - B) \\ &= 0.3 + 0.3 = 0.6. \end{aligned}$$

方法 2 利用对立事件公式

$$\begin{aligned} P(A - B) &= P(A - AB) = P(A) - P(AB) \\ &= 0.7 - P(AB) = 0.3, \end{aligned}$$

$$P(AB) = 0.4,$$

$$P(\overline{AB}) = 1 - P(AB) = 0.6.$$

点评 题 1-6 和 1-7 都用到了概率的一个重要转换公式

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

它是由随机事件特殊的运算公式——差化积公式

$$A-B = A-AB = A\bar{B}$$

而得到.

【1-8】 事件 A, B, C 满足 $P(A) = P(B) = P(C) = \frac{1}{4}$, $P(AB) =$

$P(AC) = 0$, $P(BC) = \frac{1}{8}$, 试证: A, B, C 中至少有一个发生的概率为 $\frac{5}{8}$.

分析 本题的关键是需证明 $P(ABC) = 0$, 这样由广义加法公式便可得欲证结论.

证 由 $ABC \subset AB$ 得 $0 \leq P(ABC) \leq P(AB) = 0$, 从而得 $P(ABC) = 0$. 于是

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(AB) \\ &\quad - P(AC) - P(BC) + P(ABC) \\ &= 3 \times \frac{1}{4} - 0 - 0 - \frac{1}{8} + 0 = \frac{5}{8}. \end{aligned}$$

点评 初学者最容易发生如下误证:

因为 $P(AB) = 0$, 所以 $AB = \emptyset$, 于是 $ABC = \emptyset \cap C = \emptyset$, 从而 $P(ABC) = P(\emptyset) = 0$.

错误原因: 概率为零的事件未必是不可能事件.

【1-9】 设事件 A, B, C 同时发生必导致事件 D 发生, 则

$$P(A) + P(B) + P(C) \leq 2 + P(D).$$

分析 由于概率的性质很多, 所以证明方法也很多.

第一类方法是利用广义加法公式:

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

或

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

第二类方法是利用广义加法公式的推论

$$P(A \cup B \cup C) \leq P(A) + P(B) + P(C).$$

第三类方法是利用积化差公式

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

证 方法 1

由题设可知 $ABC \subset D$, 得 $P(ABC) \leq P(D)$,

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

$$\begin{aligned} P(AB \cup AC \cup BC) &= P(AB) + P(AC) + P(BC) \\ &\quad - 2P(ABC), \quad (2) \end{aligned}$$

式(1)+式(2)得

$$P(A) + P(B) + P(C) - P(ABC) \leq 2,$$

从而有

$$P(A) + P(B) + P(C) \leq 2 + P(ABC) \leq 2 + P(D).$$

方法 2 由 $P(A \cup B) \leq 1$ 可得 $P(A) + P(B) \leq 1 + P(AB)$, 两边加 $P(C)$ 并再一次应用广义加法公式, 得

$$\begin{aligned} P(A) + P(B) + P(C) &\leq 1 + P(AB) + P(C) \\ &\leq 1 + P(AB \cup C) + P(ABC) \\ &\leq 2 + P(ABC) \leq 2 + P(D). \end{aligned}$$

方法 3 由题设可知 $ABC \subset D$, 得 $P(\overline{ABC}) \geq P(\overline{D})$, 于是

$$\begin{aligned} P(A) + P(B) + P(C) &= 3 - P(\overline{A}) - P(\overline{B}) - P(\overline{C}) \\ &\leq 3 - P(\overline{A} \cup \overline{B} \cup \overline{C}) \\ &= 3 - P(\overline{ABC}) \leq 3 - P(\overline{D}) = 2 + P(D). \end{aligned}$$

方法 4 因为 $P(\overline{B}) \geq P(\overline{AB}) = P(A) - P(AB)$, 同理

$$P(\overline{C}) \geq P(\overline{ABC}) = P(AB) - P(ABC),$$

故

$$\begin{aligned} P(A) - P(\overline{B}) - P(\overline{C}) &\leq P(A) - [P(A) - P(AB)] \\ &\quad - [P(AB) - P(ABC)] \\ &= P(ABC) \leq P(D), \end{aligned}$$

于是有

$$P(A) + P(B) + P(C) \leq 2 + P(D).$$

【1-10】 大城市电话号码由 8 位数组成, 除首位数非 0 外, 其余位数可以是 0, 1, 2, ..., 9 中的任一数, 求电话号码是由完全不同的数字组成的概率.

解 设样本空间中基本事件总数 $n=10^8$.

有利场合是:首位数有 9 种取法,由于不重复,其余 7 位数有 P_9^7 种取法. 根据乘法原理,有利的基本事件数 $k=9P_9^7$. 故所求概率

$$p = \frac{k}{n} = \frac{9P_9^7}{10^8} = \frac{163\,296}{10^7} \approx 0.016\,3.$$

【1-11】 某人买了若干大小相同的新鲜鸭蛋,其中有 a 只青壳的, b 只白壳的. 他准备将青壳蛋加工成咸蛋,故将鸭蛋一只只从箱中摸出进行分类,求第 k 次摸出的是青壳蛋(设为事件 A)的概率 $P(A)$ ($1 \leq k \leq a+b$).

分析 本题是古典概型中非常典型的摸“球”概型. 由于对同一随机试验可作不同的设计,故求解方法也各不相同.

解 方法 1 (选排列)

将蛋从 1 至 $a+b$ 进行编号,基本事件总数为从 $a+b$ 只蛋中选出 k 只进行排列的排列数 P_{a+b}^k ,有利事件数为 $P_a^1 P_{a+b-1}^{k-1}$. 由古典概型即有

$$P(A) = \frac{P_a^1 P_{a+b-1}^{k-1}}{P_{a+b}^k} = \frac{a}{a+b}.$$

方法 2 (全排列)

仍将蛋编号,将所有的蛋摸出来排列在 $a+b$ 个位置上,基本事件总数为 $(a+b)!$,要求在第 k 个位置上放青壳蛋,有利事件数为 $a(a+b-1)!$,所以

$$P(A) = \frac{a(a+b-1)!}{(a+b)!} = \frac{a}{a+b}.$$

方法 3 (组合)

对同色蛋不加区别,构造的随机试验是将 $a+b$ 只蛋放在一直线的 $a+b$ 个位置上,且只需考虑青壳蛋的放法. 把 a 只青壳蛋放在 $a+b$ 个位置上的所有不同的放法总数为 C_{a+b}^a ,有利事件数为 C_{a+b-1}^{a-1} ,故

$$P(A) = \frac{C_{a+b-1}^{a-1}}{C_{a+b}^a} = \frac{a}{a+b}.$$

方法 4 构造的随机试验只考虑第 k 次摸出的是何种颜色的蛋,把所有的可能结果作为基本事件全体,显然总数为 $a+b$. 要求第 k 次摸出青壳蛋,有利事件数为 a ,所以

$$P(A) = \frac{a}{a+b}.$$

点评 前三个解法分别用了排列或组合的方法计算基本事件数,这是大家通常采用的方法. 由于随机试验设计的不同,所以样本空间的大小也不同,方法 4 的样本空间是最小的(因为再小就不能保证等可能性),因而也是最佳解法.

怎样能取到最小的样本空间而使计算大为简化呢? 关键在于抓住所求概率的事件的本质特点,而把无关的因素都丢掉不予考虑.

最后再强调两点:

① 本例虽然是求第 k 次摸出青壳蛋的概率,而结果却与摸蛋次序 k 无关,恒为一个常数.

② 若本题改为放回式摸蛋,则结果也是 $\frac{a}{a+b}$, 仍与 k 无关,读者不妨计算一下.

这一结果表明:在体育比赛或日常其他活动中常用抽签(无论是放回还是不放回)来作出某种选择或决定,对参与各方是公平合理的. 因为不管谁先抽还是后抽,各方中签的机会是均等的.

【1-12】 某县城有 1 万辆自行车,其牌照号码从 00 001 到 10 000. 求在该县城偶遇一辆自行车的牌照号码中有数字 8 的概率.

解 方法 1

基本事件总数 $n=10^4$. 有利场合共有 4 种情况:

个位取 8, 有 $C_{10}^1 C_{10}^1 C_{10}^1 \times 1$ 种不同取法;

十位取 8, 有 $C_{10}^1 C_{10}^1 \times 1 \times C_9^1$ 种不同取法;

百位取 8, 有 $C_{10}^1 \times 1 \times C_9^1 C_9^1$ 种不同取法;

千位取 8, 有 $1 \times C_9^1 C_9^1 C_9^1$ 种不同取法.

于是有利的基本事件总数

$$\begin{aligned} k &= C_{10}^1 C_{10}^1 C_{10}^1 + C_{10}^1 C_{10}^1 \times 1 \times C_9^1 + \\ &C_{10}^1 \times 1 \times C_9^1 C_9^1 + 1 \times C_9^1 C_9^1 C_9^1 = 3439, \end{aligned}$$

所求概率

$$P = \frac{k}{n} = 0.3439.$$

方法2 基本事件总数 $n=10^4$. 有利的基本事件数

$$k = C_4^1(C_9^1)^3 + C_4^2(C_9^1)^2 + C_4^3C_9^1 + C_4^4 = 3439,$$

$\downarrow$ 一个8 $\downarrow$ 二个8 $\downarrow$ 三个8 $\downarrow$ 四个8

所求概率

$$P = \frac{k}{n} = 0.3439.$$

方法3 设 A 表示牌号中有 8 的事件, A_i 为第 i 位上有 8 (从个位数起) 的事件 ($i=1, 2, 3, 4$). 则

$$A = \bigcup_{i=1}^4 A_i, \quad P(A_i) = 0.1 \quad (i = 1, 2, 3, 4).$$

显然 A_1, A_2, A_3, A_4 相互独立, 于是所求概率

$$\begin{aligned} P(A) &= P\left(\bigcup_{i=1}^4 A_i\right) = 1 - P\left(\bigcap_{i=1}^4 \overline{A_i}\right) = 1 - [P(\overline{A_i})]^4 \\ &= 1 - 0.9^4 = 0.3439. \end{aligned}$$

方法4 利用对立事件

$$P(\text{牌号中有 } 8) = 1 - P(\text{牌号中无 } 8) = 1 - \frac{9^4}{10^4} = 0.3439.$$

点评 直接利用对立事件是最佳解法.

【1-13】 一幢 12 层的大楼, 有 6 位乘客从底楼进入电梯, 电梯可停于 2 至 12 层的任一层. 若每位乘客在任一层离开电梯的可能性相同, 求下列事件的概率:

$A = \{\text{某指定的一层有 2 位乘客离开}\};$

$B = \{\text{至少有 2 位乘客在同一层离开}\};$

$C = \{\text{恰有 2 位乘客在同一层离开}\}.$

解 由于乘客均可能在 2 至 12 层的任一层离开, 所以基本事件总数 $n=11^6$.

某指定的一层有 2 位乘客离开, 6 位中任选 2 位有 C_6^2 种不同的组合方式, 其余 4 位可在另外 10 层中按任意方式离开, 则有 10^4 种不同的离开方式, 故有利于 A 发生的基本事件数 $k_A = C_6^2 \times 10^4$. 于是

$$P(A) = \frac{k_A}{n} \approx 0.0847.$$

考虑逆事件 $\bar{B} = \{\text{没有 2 位及 2 位以上乘客在同一层离开}\}$, 即 6

位乘客必须在 11 层中的任意 6 层每层任意离开 1 位,故有利于 $\bar{B}$ 发生的基本事件数为 $k_{\bar{B}} = C_{11}^6 \times 6!$ 于是

$$P(B) = 1 - P(\bar{B}) = 1 - \frac{k_{\bar{B}}}{n} \approx 0.8122.$$

恰有 2 位乘客在同一层离开,对楼层的选择方式有 C_{11}^1 种,乘客的组合方式有 C_6^2 种,而余下 4 位乘客在其余楼层不能再出现 2 位在同一层离开,他们离开的方式有如下三种情况:

- ① 一层离开 1 位,计有 $C_{10}^4 \times 4!$ 种不同方式;
- ② 4 位在同一层离开,计有 C_{10}^1 种不同方式;
- ③ 任意 1 位在其余 10 层中的任一层离开,其余 3 位在余下的 9 层中的任一层同时离开,计有 $C_4^1 C_{10}^1 C_9^1$ 种不同方式.

根据乘法原理与加法原理,可得有利于事件 C 发生的基本事件数合计为

$$k_C = C_{11}^1 C_6^2 (C_{10}^4 \times 4! + C_{10}^1 + C_4^1 C_{10}^1 C_9^1),$$

于是

$$P(C) = \frac{k_C}{n} \approx 0.5039.$$

【1-14】 3 封信随机地投向 1 至 4 号 4 个空邮筒. 求:

- (1) 第 2 号邮筒恰有一封信投入的概率;
- (2) 有 3 个邮筒各有一封信投入的概率.

分析 本题与上一题是同一类型的题目. 相当于 3 位乘客进入 5 层大楼底楼的电梯,求:

- (1) 某指定的一层有 1 位乘客离开;
- (2) 有 3 个楼层各有 1 位乘客离开.

解 基本事件总数 $n = 4^3 = 64$.

(1) 3 封信中任选一封投入 2 号邮筒,有 C_3^1 种选法,其余 2 封投入另外 3 个邮筒有 3^2 种投法,故有利的基本事件数 $k_1 = C_3^1 \times 3^2 = 27$. 于是

$$P(2 \text{ 号邮筒恰有一封信投入}) = \frac{k_1}{n} = \frac{27}{64}.$$

(2) 4 个邮筒任选 3 个有 C_4^3 种选法, 每个邮筒将被投入一封信有 $3!$ 种投法, 故有利的基本事件数 $k_2 = C_4^3 \times 3! = 24$, 于是

$$P(3 \text{ 个邮筒各有一封信投入}) = \frac{k_2}{n} = \frac{3}{8}.$$

点评 上述两题都属于球在盒中的分布问题, 这里“球”与“盒”的含义很广, 不少实际问题都可归结为这类问题.

【1-15】 一次投掷 2 颗骰子, 求出现的点数之和为奇数的概率.

解 方法 1 设随机事件 A 为“出现点数之和为奇数”. 以点数 (i, j) 表示基本事件 $(i, j=1 \sim 6)$, 则基本事件总数 $n=36$, 这 36 个基本事件组成等概样本空间(即每个基本事件发生的概率都相等的样本空间) 其中有利于 A 发生的基本事件数 $k=18$, 故

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

方法 2 每次试验可能出现的结果为(奇, 奇)、(奇, 偶)、(偶, 奇)、(偶, 偶), 这四个基本事件构成等概样本空间. 显然基本事件总数 $n=4$, A 包含的基本事件数 $k=2$, 故

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

方法 3 每次试验可能出现的结果为 $\omega_1 = \{\text{点数之和为奇数}\}$, $\omega_2 = \{\text{点数之和为偶数}\}$, ω_1, ω_2 构成等概样本空间, 显然基本事件总数 $n=2$, A 包含的基本事件数 $k=1$, 故

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

点评 本题说明同一问题可以取不同的样本空间, 但必须是等概样本空间, 否则会出错. 例如在方法 2 中若取(两个奇)、(一奇一偶)、(两个偶)作为基本事件组成样本空间, 则得出

$$P(A) = \frac{1}{3}.$$

错误的原因就是所构造的样本空间不是等概样本空间, 事实上

$$P(\text{两个奇}) = \frac{1}{4} \neq \frac{1}{2} = P(\text{一奇一偶}),$$

所以不能用古典概型公式计算.

【1-16】 从 6 双不同的手套中任取 4 只,求其中恰有一双配对的概率.

解 设恰有一双配对为事件 A ,基本事件总数 $n=C_{12}^4$.

方法 1 有利场合是:先从 6 双中任取一双,然后从剩下的 5 双中任取 2 双,再从每双中任取 1 只,则 A 包含的基本事件数 $k=C_6^1 C_5^2 C_2^1 C_2^1$, 于是

$$P(A) = \frac{k}{n} = \frac{16}{33}.$$

方法 2 有利场合是:先从 6 双中任取一双,然后按左、右手将剩下 10 只分成两摊,从两摊中任取一摊,再依次取出 2 只,则 A 包含的基本事件数 $k=C_6^1 C_2^1 C_5^1 C_4^1$, 于是

$$P(A) = \frac{k}{n} = \frac{16}{33}.$$

方法 3 有利场合是:先从 6 双中任取一双,然后从剩下 10 只中任取 2 只,再扣除其中能配对的数目,则 A 包含的基本事件数 $k=C_6^1 (C_{10}^2 - C_5^1)$, 于是

$$P(A) = \frac{k}{n} = \frac{16}{33}.$$

方法 4 利用逆事件

$$\begin{aligned} P(A) &= 1 - P(\bar{A}) = 1 - P(\text{无配对} \cup \text{全配对}) \\ &= 1 - P(\text{无配对}) - P(\text{全配对}) \\ &= 1 - \frac{C_6^4 \times 2^4}{C_{12}^4} - \frac{C_6^2}{C_{12}^4} = \frac{16}{33}. \end{aligned}$$

点评 与方法 1 对应的一个方法是先取不配对的两只,再取配对的一双,此时有利事件数 $k=C_6^2 C_2^1 C_2^1 C_4^1$. 注意:在这一解法中 6 双中任取 3 双的不同取法种数为 $C_6^1 C_5^2$ 或 $C_6^2 C_4^1$ 而不能为 $C_6^1 C_5^1 C_4^1$. 因为从 2 双中各取出 1 只属一类,另一双 2 只全取出属另一类,不能混淆.

【1-17】 两人约定晚上 7 点到 8 点在公园会面,试求一人要等另一人半小时以上的概率.

分析 这是几何概型的问题.

解 设两人到达公园的时刻分别为 x, y , 不妨将时间区间 $[19, 20]$ 平移至 $[0, 1]$, 则依题意有

$$0 \leq x \leq 1, \quad 0 \leq y \leq 1,$$

样本空间 Ω 是边长为 1 的正方形. 设事件 A 为“一人要等另一人半小时以上”, 则 A 发生的充要条件为 $|x - y| > 0.5$, 即当且仅当样本点 (x, y) 落在图 1-4 正方形 Ω 之内的区域 G 时, 事件 A 才会发生, 故

$$P(A) = \frac{G \text{ 的面积}}{\Omega \text{ 的面积}} = \frac{0.5 \times 0.5}{1} = 0.25.$$

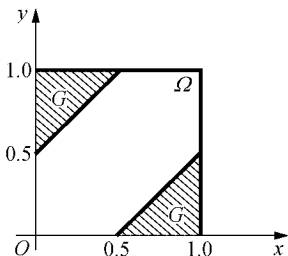


图 1-4

【1-18】 把长度为 a 的细铁杆任意截成 3 段, 试求它们可以构成一个三角形的概率.

解 设截得的 3 段铁杆的长度分别为 x, y 和 $a - x - y$, 则

$$0 < x < a, \quad 0 < y < a,$$

$$0 < a - x - y < a,$$

样本空间 Ω 是直角边长为 a 的等腰直角三角形. 三铁杆构成三角形的充要条件为

$$y + (a - x - y) > x,$$

$$x + (a - x - y) > y,$$

$$x + y > a - x - y,$$

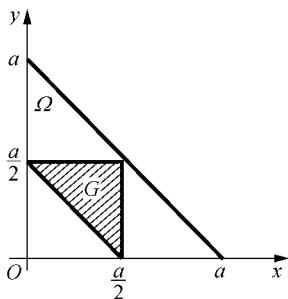


图 1-5

即

$$x < \frac{a}{2}, \quad y < \frac{a}{2}, \quad x + y > \frac{a}{2}.$$

当且仅当样本点 (x, y) 落在图 1-5 中的阴影区域 G 内时, 三铁杆才能构成三角形, 故所求的概率

$$P = \frac{G \text{ 的面积}}{\Omega \text{ 的面积}} = \frac{\frac{1}{2} \left(\frac{a}{2} \right)^2}{\frac{1}{2} a^2} = \frac{1}{4}.$$

点评 本题也可以在三维空间中求解, 即设截得的 3 段铁杆的长

度为 x, y, z , 它们能构成三角形的充要条件为

$$x + y > z, \quad x + z > y, \quad y + z > x, \quad x + y + z = a.$$

显然计算与作图的难度都增加了. 所以, 一般在低维空间能解决的问题不要放到高维空间中去解决.

【1-19】 从区间 $(0, 1)$ 中随机地取 2 个数, 求下列事件的概率

(1) 两数之和小于 $\frac{6}{5}$;

(2) 两数之积小于 $\frac{1}{4}$.

解 设从区间 $(0, 1)$ 内任取的 2 个数为 x 与 y , 则

$$0 < x < 1, \quad 0 < y < 1,$$

样本空间 Ω 是边长为 1 的正方形.

(1) 两数之和小于 $\frac{6}{5}$ 的充要条件是

$$x + y < \frac{6}{5}, \quad 0 < x < 1, \quad 0 < y < 1,$$

即当且仅当样本点 (x, y) 落在图 1-6 中的阴影区域 G 时, 事件“两数之和小于 $\frac{6}{5}$ ”才会发生, 故所求概率

$$\begin{aligned} P &= \frac{G \text{ 的面积}}{\Omega \text{ 的面积}} = \frac{1 - \frac{1}{2} \left(1 - \frac{1}{5}\right)^2}{1} \\ &= \frac{17}{25} = 0.68. \end{aligned}$$

(2) 两数之积小于 $\frac{1}{4}$ 的充要条件是

$$xy < \frac{1}{4}, \quad 0 < x < 1, \quad 0 < y < 1,$$

即当且仅当样本点 (x, y) 落在图 1-7 中的

阴影区域 G 时, 事件“两数之积小于 $\frac{1}{4}$ ”才会发生, 故所求概率

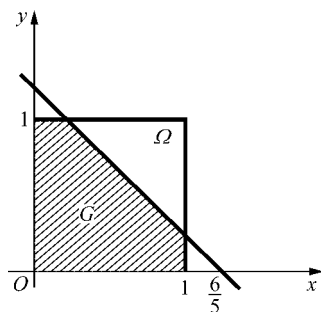


图 1-6

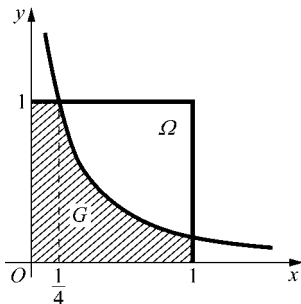


图 1-7

$$P = \frac{G \text{ 的面积}}{\Omega \text{ 的面积}} = \frac{1 \times \frac{1}{4} + \int_{\frac{1}{4}}^1 \frac{1}{4x} dx}{1} \\ = \frac{1}{4} + \frac{1}{2} \ln 2.$$

【1-20】 设事件 A, B, C 相互独立, 证明: 事件 $A \cup B, A - B, AB$ 都与 C 相互独立.

证 由于 A, B, C 相互独立, 故两两独立, 则有

$$\begin{aligned} \textcircled{1} P((A \cup B)C) &= P(AC \cup BC) = P(AC) + P(BC) - P(ABC) \\ &= P(A)P(C) + P(B)P(C) - P(A)P(B)P(C) \\ &= [P(A) + P(B) - P(AB)]P(C) \\ &= P(A \cup B)P(C), \end{aligned}$$

所以 $A \cup B$ 与 C 相互独立.

$$\begin{aligned} \textcircled{2} P((A - B)C) &= P(AC - BC) = P(AC - ABC) \\ &= P(AC) - P(ABC) \\ &= P(A)P(C) - P(A)P(B)P(C) \\ &= P(A - AB)P(C) = P(A - B)P(C), \end{aligned}$$

所以 $A - B$ 与 C 相互独立.

$$\textcircled{3} P((AB)C) = P(ABC) = P(A)P(B)P(C) = P(AB)P(C).$$

所以 AB 与 C 相互独立.

点评 一般都可利用事件独立的定义来判断事件是否独立和证明事件的独立性. 本题的条件与结论都比较简单, 即将三个相互独立的事件分成两个组, 经有限次运算后, 这两个组的事件仍相互独立. 下面将此题推广成如下命题:

设事件 $A_1, A_2, \dots, A_n$ 相互独立, 将它们任意分成 k 组, 因同一事件不能同时属于两个不同的组, 则对每组事件进行求和、差、积、逆等运算所得到的 k 个事件也相互独立.

例如事件 A, B, C, D, F, G 相互独立, 由上面命题可知, 事件 $A \cup \bar{C}, BF, D - G$ 也相互独立. 后面很多题, 例如题 1-23, 1-24 和 1-25 等都将用到本命题的结论.

【1-21】 袋中有 4 个乒乓球, 一个涂有红、黄、蓝 3 种颜色, 其余 3 个分别涂有红、黄、蓝的单一颜色. 现从袋中随机地取出 1 球, 设事件 A, B, C 分别为取出红、黄、蓝色球. 试证: A, B, C 两两独立, 但不相互独立.

证 由古典概型可得

$$P(AB) = \frac{1}{4} = \frac{1}{2} \times \frac{1}{2} = P(A)P(B),$$

$$P(AC) = \frac{1}{4} = \frac{1}{2} \times \frac{1}{2} = P(A)P(C),$$

$$P(BC) = \frac{1}{4} = \frac{1}{2} \times \frac{1}{2} = P(B)P(C),$$

$$P(ABC) = \frac{1}{4} \neq \frac{1}{8} = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = P(A)P(B)P(C),$$

所以事件 A, B, C 两两独立, 但不相互独立.

【1-22】 设 $P(A) > 0, P(B) > 0$.

(1) 如果 A 与 B 互不相容, 问 A 与 B 是否相互独立?

(2) 如果 A 与 B 相互独立, 问 A 与 B 是否互不相容?

解 (1) A 与 B 一定不相互独立. 用反证法证明如下: 若 A 与 B 相互独立, 由定义有 $P(AB) = P(A)P(B) > 0$, 而这与 $P(AB) = P(\emptyset) = 0$ 矛盾.

(2) A 与 B 不会互不相容. 用反证法证明如下: 若 A 与 B 互不相容, 即 $AB = \emptyset$, 则 $P(AB) = P(\emptyset) = 0$, 这与 A, B 相互独立, 从而有 $P(AB) = P(A)P(B) > 0$ 矛盾.

点评 事件相互独立与互不相容是两个不同的概念. “相互独立”是指事件 A, B 发生互不影响, “互不相容”是指事件 A, B 不能同时发生.

本题告诉我们, 在事件概率都不为零时, 有以下结论:

不相容 $\Rightarrow$ 不独立, 独立 $\Rightarrow$ 相容, 但其逆均不真.

【1-23】 某种彩票的中奖面为 10%, 某人一次买了 10 张, 求其中奖的概率.

解 方法 1 设 $A_i (i = 1 \sim 10)$ 表示第 i 张彩票中奖的事件, 则

$P(A_i)=0.1$, 显然, 每张彩票是否中奖是相互独立的. $A_1, A_2, \dots, A_{10}$ 中至少有一个事件发生, 某人便中奖了, 故

$$\begin{aligned} P(\text{某人中奖}) &= P\left(\bigcup_{i=1}^{10} A_i\right) = 1 - P\left(\overline{\bigcup_{i=1}^{10} A_i}\right) = 1 - P\left(\bigcap_{i=1}^{10} \overline{A_i}\right) \\ &\stackrel{\text{独立}}{=} 1 - \prod_{i=1}^{10} P(\overline{A_i}) = 1 - [1 - P(A_i)]^{10} \\ &= 1 - 0.9^{10} \approx 0.6513. \end{aligned}$$

方法 2 一次买 10 张彩票可视为进行 10 重伯努利(Bernoulli)试验, 每次试验成功的概率均为 0.1. 由逆事件公式可有

$$\begin{aligned} P(\text{某人中奖}) &= 1 - P(\text{某人不中奖}) = 1 - P_{10}(0) \\ &= 1 - C_{10}^0 \times 0.1^0 \times 0.9^{10} \approx 0.6513. \end{aligned}$$

点评 由方法 1 可得到求 n 个相互独立的事件的并事件的概率公式

$$P\left(\bigcup_{i=1}^n A_i\right) = 1 - \prod_{i=1}^n [1 - P(A_i)].$$

当 $P(A_i)=p$ ($i=1\sim n$) 时,

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

【1-24】 某企业招工需经过 4 项考核. 设能够通过第一、二、三、四项考核的概率分别为 0.9, 0.8, 0.7, 0.6, 且各项考核是独立的, 也未必按顺序进行, 每个应招者在四项考核中有一项不通过即被淘汰. 求:

- (1) 这次招工的淘汰率;
- (2) 虽通过第一、三项考核, 但仍被淘汰的概率.

解 设事件 $A_i = \{\text{第 } i \text{ 项考核通过}\}$ ($i=1, 2, 3, 4$), 则

$$P(A_1) = 0.9, P(A_2) = 0.8, P(A_3) = 0.7, P(A_4) = 0.6.$$

- (1) 这次招工的淘汰率

$$\begin{aligned} P(\overline{A_1 A_2 A_3 A_4}) &= 1 - P(A_1 A_2 A_3 A_4) \\ &= 1 - P(A_1)P(A_2)P(A_3)P(A_4) \\ &= 1 - 0.9 \times 0.8 \times 0.7 \times 0.6 = 0.6976, \end{aligned}$$

或者

$$\begin{aligned}
& P(\bar{A}_1 \cup A_1 \bar{A}_2 \cup A_1 A_2 \bar{A}_3 \cup A_1 A_2 A_3 \bar{A}_4) \\
&= P(\bar{A}_1) + P(A_1)P(\bar{A}_2) + P(A_1)P(A_2)P(\bar{A}_3) \\
&\quad + P(A_1)P(A_2)P(A_3)P(\bar{A}_4) \\
&= 0.1 + 0.18 + 0.216 + 0.2016 = 0.6976.
\end{aligned}$$

$$\begin{aligned}
(2) \quad P(A_1 A_3 \overline{A_2 A_4}) &= P(A_1)P(A_3)P(\overline{A_2 A_4}) \\
&= P(A_1)P(A_3)[1 - P(A_2)P(A_4)] \\
&= 0.9 \times 0.7 \times (1 - 0.8 \times 0.6) = 0.3276.
\end{aligned}$$

点评 淘汰率的计算用了两种不同的方法. 第一种是不考虑考核的顺序, 四项考核中至少有一项不通过的概率便是淘汰率; 第二种是假定考核按顺序进行, 应考者一旦某项考核未通过便遭淘汰, 不再参加后面项目的考核.

【1-25】 对同一目标进行 3 次独立射击, 第一、二、三次射击的命中率分别为 0.6, 0.8, 0.7. 求:

- (1) 恰有一次击中目标的概率 p_1 ;
- (2) 至少有一次击中目标的概率 p_2 .

解 设事件 $C_i = \{\text{第 } i \text{ 次击中目标}\} \quad (i=1, 2, 3)$.

- (1) 事件 $A = \{\text{恰有一次击中目标}\}$, 则

$$A = C_1 \bar{C}_2 \bar{C}_3 \cup \bar{C}_1 C_2 \bar{C}_3 \cup \bar{C}_1 \bar{C}_2 C_3.$$

上式右端三个事件互不相容, 由加法公式及事件的相互独立性可得

$$\begin{aligned}
p_1 &= P(A) = P(C_1 \bar{C}_2 \bar{C}_3) + P(\bar{C}_1 C_2 \bar{C}_3) + P(\bar{C}_1 \bar{C}_2 C_3) \\
&= 0.6 \times 0.2 \times 0.3 + 0.4 \times 0.8 \times 0.3 + 0.4 \times 0.2 \times 0.7 \\
&= 0.188.
\end{aligned}$$

- (2) 事件 $B = \{\text{至少有一次击中目标}\}$, 则 $\bar{B} = \bar{C}_1 \bar{C}_2 \bar{C}_3$, 于是

$$\begin{aligned}
p_2 &= P(B) = 1 - P(\bar{B}) = 1 - P(\bar{C}_1)P(\bar{C}_2)P(\bar{C}_3) \\
&= 1 - 0.4 \times 0.2 \times 0.3 = 0.976.
\end{aligned}$$

【1-26】 在 4 次重复独立的试验中, 事件 A 至少出现一次的概率为 $\frac{65}{81}$. 求在一次试验中 A 出现的概率.

解 本题为 4 重伯努利试验, 设 $P(A) = p$, 考虑对立事件的概率

$$P(4 \text{ 次试验中 } A \text{ 一次未出现}) = P_4(0) = C_4^0 p^0 (1-p)^4$$

$$\begin{aligned}
 &= (1-p)^4 = 1 - \frac{65}{81} = \frac{16}{81} \\
 &= \left(\frac{2}{3}\right)^4 \Rightarrow p = \frac{1}{3},
 \end{aligned}$$

所以 $P(A) = \frac{1}{3}$.

【1-27】 设事件 A, B 相互独立, 且 $P(A\bar{B}) = \frac{1}{4}$, $P(\bar{A}B) = \frac{1}{6}$. 求 $P(A), P(B)$.

分析 将 $P(A), P(B)$ 视为未知数 x, y , 则本题变为二元二次方程组的求解问题.

解 令 $P(A) = x, P(B) = y$, 由 A, B 的独立性可有

$$\begin{cases} P(A\bar{B}) = P(A)P(\bar{B}) = x(1-y) = \frac{1}{4}, \\ P(\bar{A}B) = P(\bar{A})P(B) = y(1-x) = \frac{1}{6}, \end{cases}$$

解此方程组得

$$\begin{cases} P(A) = x = \frac{1}{3}, \\ P(B) = y = \frac{1}{4}, \end{cases} \quad \text{或} \quad \begin{cases} P(A) = x = \frac{3}{4}, \\ P(B) = y = \frac{2}{3}. \end{cases}$$

点评 本题中事件 A, B 独立的条件在解题中不能不用, 否则会遇到解不出题的尴尬. 例如

$$\begin{aligned}
 \text{由 } P(A\bar{B}) &= P(\overline{A \cup B}) = 1 - P(A \cup B) \\
 &= 1 - [P(\bar{A}) + P(B) - P(\bar{A}B)]
 \end{aligned}$$

得

$$1 - P(A) + P(B) = \frac{11}{12}. \quad (1)$$

相仿可得

$$\begin{aligned}
 P(\bar{A}B) &= 1 - [P(A) + P(\bar{B}) - P(A\bar{B})] \\
 1 + P(A) - P(B) &= \frac{13}{12} \quad (2)
 \end{aligned}$$

式(1)+式(2)即得

$$2 = 2.$$

【1-28】 甲、乙两个乒乓球运动员进行比赛,已知每一局甲、乙获胜的概率各为 0.6 与 0.4,比赛可以采用 3 局 2 胜制或 5 局 3 胜制.问在哪一种赛制下甲获胜的可能性较大?

解 ① 若采用 3 局 2 胜制,甲在下列两种情况下获胜:

A_1 ——甲 2:0 胜, A_2 ——甲 2:1 胜.

则 $P(A_1) = P_2(2) = 0.6^2 = 0.36$,

$$P(A_2) = P_2(1) \times 0.6 = C_2^1 \times 0.6 \times 0.4 \times 0.6 = 0.288,$$

甲胜的概率

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) = 0.648.$$

② 若采用 5 局 3 胜制,甲在下列三种情况下获胜:

B_1 ——甲 3:0 胜, B_2 ——甲 3:1 胜, B_3 ——甲 3:2 胜.

则 $P(B_1) = P_3(3) = 0.6^3 = 0.216$,

$$P(B_2) = P_3(2) \times 0.6 = C_3^2 \times 0.6^2 \times 0.4 \times 0.6 = 0.2592,$$

$$P(B_3) = P_4(2) \times 0.6 = C_4^2 \times 0.6^2 \times 0.4^2 \times 0.6 = 0.20736,$$

甲胜的概率

$$P(B_1 \cup B_2 \cup B_3) = P(B_1) + P(B_2) + P(B_3) \approx 0.683.$$

比较①,②两种情况可知甲在 5 局 3 胜制下获胜可能性大.

【1-29】 设某种新农药研制成功的概率为 3.452% ,问科研人员至少进行多少次独立试验,才能使至少一次研制成功的概率不小于 90% ?

解 设至少要做 n 次试验, k 为 n 次试验中成功的次数,由题设可知

$$P(k \geq 1) = 1 - P(k < 1) = 1 - P_n(0) = 1 - (1 - 0.003452)^n \geq$$

$$0.9 \Rightarrow 0.996548^n \leq 0.1 \Rightarrow n \geq \frac{\ln 0.1}{\ln 0.996548} \approx 665.88.$$

取 $n=666$,即至少进行 666 次试验才能满足要求.

【1-30】 设事件 A, B 相互独立, $P(A \cup B) = 0.8$, $P(A|\bar{B}) = 0.2$,求 $P(B-A)$, $P(A-B)$.

解 方法1 设 $P(A)=x, P(B)=y$, 由

$$\begin{cases} P(A \cup B) = P(A) + P(B) - P(A)P(B) = 0.8, \\ P(A | \bar{B}) = \frac{P(A\bar{B})}{P(\bar{B})} = \frac{P(A-B)}{P(\bar{B})} = \frac{P(A) - P(A)P(B)}{P(\bar{B})} = 0.2, \end{cases}$$

可得方程组

$$\begin{cases} x + y - xy = 0.8, \\ x - xy = 0.2(1 - y), \end{cases} \Rightarrow \begin{cases} P(A) = x = 0.2, \\ P(B) = y = 0.75, \end{cases}$$

于是

$$P(B-A) = P(B-AB) = P(B) - P(A)P(B) = 0.6,$$

$$P(A-B) = P(A-AB) = P(A) - P(A)P(B) = 0.05.$$

方法2 A, B 独立 $\Rightarrow A, \bar{B}$ 独立 $\Rightarrow P(A|\bar{B}) = P(A) = 0.2$,

再由

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A)P(B) = 0.2 + 0.8P(B) \\ &= 0.8 \Rightarrow P(B) = 0.75, \end{aligned}$$

可得

$$P(B-A) = P(B\bar{A}) = P(B)P(\bar{A}) = 0.75 \times 0.8 = 0.6,$$

$$P(A-B) = P(A\bar{B}) = P(A)P(\bar{B}) = 0.2 \times 0.25 = 0.05.$$

点评 方法2较方法1计算简单, 得益于如下命题的应用:

四对事件 $A, B; A, \bar{B}; \bar{A}, B; \bar{A}, \bar{B}$ 中有一对相互独立, 则其余三对也相互独立.

【1-31】 设 $P(\bar{A})=0.3, P(B)=0.4, P(A\bar{B})=0.5$. 求 $P(B|A \cup \bar{B})$.

解 $P(A \cup \bar{B}) = P(A) + P(\bar{B}) - P(A\bar{B}) = 0.7 + 0.6 - 0.5 = 0.8$,

$$\begin{aligned} P(B(A \cup \bar{B})) &= P(AB) = P(A - \bar{B}) = P(A - A\bar{B}) \\ &= P(A) - P(A\bar{B}) = 0.7 - 0.5 = 0.2, \end{aligned}$$

由条件概率公式可得

$$P(B | A \cup \bar{B}) = \frac{P[B(A \cup \bar{B})]}{P(A \cup \bar{B})} = \frac{1}{4}.$$

点评 在已求得 $P(A \cup \bar{B}) = 0.8$ 后, 求 $P[B(A \cup \bar{B})]$ 的方法还有:

$$\textcircled{1} P(B(A \cup \bar{B})) = P(B) + P(A \cup \bar{B}) - P(B \cup A \cup \bar{B})$$

$$=0.4+0.8-1=0.2;$$

$$\begin{aligned}\textcircled{2} P(B(A\cup\bar{B})) &= P(A\cup\bar{B}) - P(\bar{B}(A\cup\bar{B})) \\ &= 0.8 - P(\bar{B}) = 0.2.\end{aligned}$$

【1-32】 设某家庭有 4 个孩子, 在已知至少有 1 个女孩的条件下, 求这个家庭至少有 1 个男孩的概率(假定生男生女的概率均为 $\frac{1}{2}$).

解 设事件 $A=\{\text{至少有 1 个男孩}\}$, $B=\{\text{至少有 1 个女孩}\}$. 由于生男生女等可能, 故

$$P(\bar{A}) = P(\bar{B}) = \left(\frac{1}{2}\right)^4 = \frac{1}{16}, \quad P(B) = \frac{15}{16}.$$

又 $\bar{A}\bar{B}=\emptyset$, 故

$$P(\bar{A}\bar{B}) = P(\bar{A} \cup \bar{B}) = P(\bar{A}) + P(\bar{B}) = \frac{1}{8}, \quad P(AB) = \frac{7}{8},$$

于是

$$P(A | B) = \frac{P(AB)}{P(B)} = \frac{14}{15} \approx 0.9333.$$

【1-33】 一批零件共 100 个, 次品率为 10%. 现不放回地逐个取零件, 求第一次取得次品且第二次取得正品的概率.

解 方法 1 设事件 A, B 分别表示第一次取得次品, 第二次取得正品. 由乘法公式可得

$$P(AB) = P(A)P(B | A) = \frac{1}{10} \times \frac{90}{99} = \frac{1}{11}.$$

方法 2 用超几何分布公式计算不放回地取一件次品一件正品的概率

$$\hat{p} = \frac{C_{10}^1 C_{90}^1}{C_{100}^2} = \frac{2}{11}.$$

于是第一次取得次品第二次取得正品的概率

$$p = \frac{\hat{p}}{2} = \frac{1}{11}.$$

方法 3 由古典概型公式可得所求概率

$$p = \frac{k}{n} = \frac{P_{10}^1 P_{90}^1}{P_{100}^2} = \frac{1}{11}.$$

点评 求解本题易犯如下错误:

$$\textcircled{1} P(AB) = \frac{C_{10}^1 C_{90}^1}{C_{100}^2} = \frac{2}{11},$$

要注意组合数表示只选不排的个数,是不讲排列次序的.

$$\textcircled{2} P(AB) = P(A)P(B) = \frac{10}{100} \times \frac{90}{99} = \frac{1}{11},$$

上式中有两个错误,一是 A 与 B 不独立,二是 $P(B) \neq \frac{90}{99}$,应为 $P(B) = \frac{9}{10}$.

【1-34】 10 件产品中有 3 件次品,从中任取 2 件. 在已知其中一件是次品的条件下,求另一件也是次品的概率.

解 设事件 $A = \{2 \text{ 件都是次品}\}$, $B = \{2 \text{ 件中至少有 1 件是次品}\}$. 由于 $A \subset B$, 故 $AB = A$. 由题设可知

$$P(A) = \frac{C_3^2}{C_{10}^2}, \quad P(B) = \frac{C_3^2 + C_3^1 C_7^1}{C_{10}^2},$$

于是

$$P(A | B) = \frac{P(AB)}{P(B)} = \frac{P(A)}{P(B)} = \frac{1}{8}.$$

点评 初学者求解本题往往会求什么设什么,比如:

设 $A = \{\text{另一件也是次品}\}$, $B = \{\text{所取 2 件中有 1 件次品}\}$, 则

$$P(A | B) = \frac{P(AB)}{P(B)} = \frac{\frac{C_3^2}{C_{10}^2}}{\frac{C_3^1 C_7^1}{C_{10}^2}} = \frac{1}{7}.$$

显然此结果是错误的,导致错误的原因是所设不合理.

【1-35】 小王忘了朋友家电话号码的最后一位数,他只能随意拨最后一个号,共拨了三次. 求第三次才拨通的概率.

解 设事件 $A_i = \{\text{第 } i \text{ 次拨通}\}$ ($i = 1, 2, 3$). 由乘法公式可得第三次才拨通的概率

$$P(\overline{A_1} \overline{A_2} A_3) = P(\overline{A_1})P(\overline{A_2} | \overline{A_1})P(A_3 | \overline{A_1} \overline{A_2})$$

$$= \frac{9}{10} \times \frac{8}{9} \times \frac{1}{8} = \frac{1}{10}.$$

点评 有些读者对上述解法有疑义,认为“第三次才拨通”,意味着前两次未拨通,故要用条件概率公式求解:

$$P(A_3 | \overline{A_1} \overline{A_2}) = \frac{1}{8}.$$

产生此误解的原因是未能仔细读题,未能分清条件概率与无条件概率的区别. 本题若改为:“……共拨了三次,已知前两次都未拨通. 求第三次拨通的概率.”此时,所求的才是条件概率.

【1-36】 盒中有 10 只羽毛球,其中有 6 只新球. 每次比赛时取出 2 只,用后放回. 求第二次比赛时取到的 2 只球都是新球的概率.

解 设事件 $A = \{\text{第二次比赛取得 2 只新球}\}$,

$B_i = \{\text{第一次比赛取到 } i \text{ 只新球}\} \quad (i=0,1,2)$, 则

$$P(B_i) = \frac{C_6^i C_4^{2-i}}{C_{10}^2}, \quad P(A | B_i) = \frac{C_{6-i}^2}{C_{10}^2}.$$

由全概率公式可得

$$\begin{aligned} P(A) &= \sum_{i=0}^2 P(B_i) P(A | B_i) \\ &= \frac{1}{(C_{10}^2)^2} (C_4^2 C_6^2 + C_6^1 C_4^1 C_5^2 + C_6^2 C_4^2) = \frac{28}{135} \approx 0.2074. \end{aligned}$$

【1-37】 已知某油田钻井队打的井出油的概率为 0.06,而出油的井恰位于有储油地质结构位置上的概率为 0.85,不出油的井位于有储油地质结构位置上的概率为 0.4. 求该钻井队:

(1) 打的井位于有储油地质结构位置上的概率;

(2) 在有储油结构位置上打的井出油的概率.

解 设事件 $A = \{\text{打的井位于有储油地质结构位置上}\}$, $B = \{\text{打的井出油}\}$, 则由题设可知

$$P(B) = 0.06, \quad P(A | B) = 0.85, \quad P(A | \bar{B}) = 0.4.$$

(1) 由全概率公式可得

$$\begin{aligned} P(A) &= P(B) P(A | B) + P(\bar{B}) P(A | \bar{B}) \\ &= 0.06 \times 0.85 + 0.94 \times 0.4 = 0.427. \end{aligned}$$

(2) 由贝叶斯公式可得

$$P(B|A) = \frac{P(B)P(A|B)}{P(A)} = \frac{0.06 \times 0.85}{0.427} = \frac{51}{427} \approx 0.1194.$$

【1-38】 某厂卡车运送防“非典”用品下乡,顶层装 10 个纸箱,其中 5 箱为民用口罩、2 箱为医用口罩、3 箱为消毒棉花.到目的地时发现丢失了一箱,但不知丢失的是哪一箱.现从剩下的 9 箱中任意打开 2 箱,结果都是民用口罩,求丢失的一箱也是民用口罩的概率.

解 设事件 $A = \{\text{任意打开 2 箱都是民用口罩}\}$,

$B_k = \{\text{丢失的一箱为 } k\}, k=1, 2, 3$ 分别表示民用口罩、医用口罩、消毒棉花.则 B_1, B_2, B_3 构成完备事件组,且

$$P(B_1) = \frac{1}{2}, \quad P(B_2) = \frac{3}{10}, \quad P(B_3) = \frac{1}{5},$$

$$P(A|B_1) = \frac{C_4^2}{C_9^2}, \quad P(A|B_2) = P(A|B_3) = \frac{C_5^2}{C_9^2}.$$

由全概率公式可得

$$P(A) = \sum_{k=1}^3 P(B_k)P(A|B_k) = \frac{8}{36},$$

由贝叶斯公式即得所求概率

$$P(B_1|A) = \frac{P(B_1)P(A|B_1)}{P(A)} = \frac{3}{36} \times \frac{36}{8} = \frac{3}{8}.$$

点评 通过求解以上三题,似乎可以得出一个结论:用全概率公式或贝叶斯公式求解的题目,题中给出的已知数据都比较多,一般至少有两对,即 $P(B_1), P(B_2); P(A|B_1), P(A|B_2)$. 此时若求的是无条件概率,则用全概率公式;求的是条件概率,则用贝叶斯公式.反之,如果题中有两对或两对以上这样的数据,是否一定要用全概率公式或贝叶斯公式来求解呢?这倒未必.比如题 1-38 就有非常简单的求解方法——缩减样本空间法.事实上,去掉打开的 2 箱民用口罩,基本事件总数 $n=8$,有利事件发生的基本事件数 $k=3$,故 $P(B_1|A) = \frac{3}{8}$. 这个例子告诉我们:要善于抓住问题的本质,灵活运用各种方法,不要拘泥于某些固定的解题套路.

全概率公式是一种“化整为零”、“化难为易”的方法,它不仅使计算简单规则,而且使分析问题的思路变得非常清晰,从而降低了解题的难度.下面一题便是很好的说明.

【1-39】 甲、乙两人比赛射击,每回射击胜者得 1 分. 每回射击甲、乙获胜的概率分别为 α 与 β ($\alpha + \beta = 1$). 比赛进行到有一人比对方多 2 分为止. 多 2 分者为胜者,求甲最终获胜的概率.

解 设事件 $A = \{\text{甲最终获胜}\}$, $B_1 = \{\text{第一、二回射击中甲均获胜}\}$, $B_2 = \{\text{第一、二回射击中乙均获胜}\}$, $B_3 = \{\text{第一、二回射击中甲、乙各胜一回}\}$. 显然 B_1, B_2, B_3 构成完备事件组,根据独立性有

$$P(B_1) = \alpha^2, \quad P(B_2) = \beta^2, \quad P(B_3) = 2\alpha\beta,$$

根据题设比赛规则可得

$$P(A | B_1) = 1, \quad P(A | B_2) = 0, \quad P(A | B_3) = P(A).$$

由全概率公式可知

$$P(A) = \sum_{i=1}^3 P(B_i)P(A | B_i) = \alpha^2 + 2\alpha\beta P(A),$$

于是所求概率

$$P(A) = \frac{\alpha^2}{1 - 2\alpha\beta}.$$

点评 通过对题意的分析,找到一个完备事件组是用全概率公式解题的关键之处.

【1-40】 一种传染病在某市的发病率为 4%. 为查出这种传染病,医院采用一种检验法,该方法能使 98% 的患有此病的人被检出阳性,但也会有 3% 未患此病的人被检出阳性. 现某人被此法检出阳性,求此人确定患这种传染病的概率.

解 设事件 $A = \{\text{被查者结果为阳性}\}$, $B = \{\text{被查者患此传染病}\}$, 则

$$P(B) = 0.04, \quad P(\bar{B}) = 0.96,$$

$$P(A | B) = 0.98, \quad P(A | \bar{B}) = 0.03.$$

由贝叶斯公式可得

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

$$\begin{aligned}
&= \frac{0.04 \times 0.98}{0.04 \times 0.98 + 0.96 \times 0.03} \\
&= \frac{392}{680} \approx 0.576.
\end{aligned}$$

【1-41】 袋中有 6 张相同的卡片,上面分别标有数 0,1,2,3,4,5. 现从袋中任意摸出两张卡片,已知其上数字之和大于 6,试判断先摸出的一张卡片上的数字最可能是几?

解 设事件 $A = \{\text{两数之和大于 } 6\}$,

$B_i = \{\text{先摸出的数字为 } i\} \quad (i=0,1,2,3,4,5)$,

依题意可知

$$P(B_i) = \frac{1}{6} \quad (i=0,1,2,3,4,5), P(A | B_0) = P(A | B_1) = 0,$$

$$P(A | B_2) = \frac{1}{5}, \quad P(A | B_3) = P(A | B_4) = \frac{2}{5}, \quad P(A | B_5) = \frac{3}{5}.$$

由全概率公式即得

$$P(A) = \sum_{i=0}^5 P(B_i)P(A | B_i) = \frac{1}{6} \times \frac{1+2+2+3}{5} = \frac{4}{15};$$

又由贝叶斯公式可求得

$$P(B_5 | A) = \frac{P(B_5)P(A | B_5)}{P(A)} = \frac{3}{8},$$

$$P(B_4 | A) = P(B_3 | A) = \frac{2}{8},$$

$$P(B_2 | A) = \frac{1}{8},$$

$$P(B_1 | A) = P(B_0 | A) = 0.$$

因为

$$P(B_5 | A) > P(B_k | A) \quad (k=0,1,2,3,4),$$

所以先摸出的卡片上的数字最可能的是 5.

点评 本题表明,使后验概率 $P(B_i | A)$ 达到最大值的事件 B_i 是最有可能发生的,故贝叶斯公式提供了一种统计推断的方法.

B 习题精选

【1-42】 甲、乙两人同时向一目标射击, 设事件 $A = \{\text{甲未命中目标, 乙命中目标}\}$, 则其对立事件 $\bar{A} =$ _____.

- (A) $\{\text{甲命中目标, 乙未命中目标}\}$;
- (B) $\{\text{甲命中目标或乙未命中目标}\}$;
- (C) $\{\text{甲命中目标或乙命中目标}\}$;
- (D) $\{\text{甲、乙都命中目标}\}$.

【1-43】 设随机事件 A, B 互不相容, 则必有 _____.

- (A) $P(\bar{A} \cup \bar{B}) = 1$;
- (B) $P(\bar{A} \cap \bar{B}) = 1$;
- (C) $P(A) = 1 - P(B)$;
- (D) $P(AB) = P(A)P(B)$.

【1-44】 某厂甲、乙两个车间生产同一种产品, 今从即将出厂的该种产品中任取一件. 设事件 $A = \{\text{取得的是甲车间生产的产品}\}$, $B = \{\text{取得的是乙车间生产的产品}\}$, $C = \{\text{取得的是次品}\}$, 则取得的产品是甲车间生产的正品的概率是 _____.

- (A) $P(\bar{C}|A)$;
- (B) $P(\bar{C}|\bar{B})$;
- (C) $P(\bar{B} \cup \bar{C})$;
- (D) $P(\bar{A} \cup \bar{C})$.

【1-45】 下列各命题中正确的是 _____.

- (A) 若事件 A 与 B 相互独立, 则 $P(A|B) = P(B)$;
- (B) 若事件 A 与 B 相互独立, 则 $P(A) = P(B)$;
- (C) 若事件 A 与 B 相互独立, 则 $P(A \cup B) = P(A) + P(B)$;
- (D) 若事件 A 与 B 相互独立, 则 $P(\bar{A}\bar{B}) = P(\bar{A})P(\bar{B})$.

【1-46】 设 $0 < P(A) < 1$, $0 < P(B) < 1$, $P(A|B) + P(\bar{A}|\bar{B}) = 1$, 则 _____.

- (A) 事件 A, B 互不相容;
- (B) 事件 A, B 相互对立;
- (C) 事件 A, B 相互独立;
- (D) 事件 A, B 未必独立.

【1-47】 某城市发行日报和晚报. 经调查, 在这两种报纸的订户中, 订阅日报的有 40%, 订阅晚报的有 55%, 同时订阅两种报纸的有 10%. 求:

- (1) 订户只订日报的概率;
- (2) 订户只订一种报纸的概率.

【1-48】 10 件产品中有正品 7 件,从中任取 3 件,求下列事件的概率:

- (1) 3 件全是正品;
- (2) 3 件中至少有 1 件次品;
- (3) 3 件中有 1 件正品.

【1-49】 12 个球队参与足球联赛. 第一阶段为分组赛, 每组 6 个队, 求最强的 2 个队分在不同组的概率.

【1-50】 从 1, 2, 3, 4, 5, 6, 7 七个数码中, 任取三个不同数码排成三位数, 求:

- (1) 所得三位数为奇数的概率;
- (2) 所得三位数大于 300 的概率.

【1-51】 设每人生于一年中任意一个月都是等可能的, 求下列事件的概率:

- (1) 6 个人的生日在六个不同的月份;
- (2) 6 个人的生日恰在两个月中;
- (3) 6 个人中至少有 2 人的生日在同一个月.

【1-52】 袋中有红、黄、蓝色球各 2 个, 每次任取 1 个, 求:

- (1) 有放回地取两次, 取到的球中无红球或无黄球的概率;
- (2) 无放回地取两次, 取到的球中无红球或无黄球的概率.

【1-53】 一盒子中有 6 个乒乓球, 其中有 2 个新球, 每次任取 1 个, 连取两次, 按放回和不放回两种形式, 求至少有一次取到新球的概率.

【1-54】 一工人看管 4 台同类型的机床, 在 1 小时内, 每台机床需要照顾的概率均为 0.7, 且需要照顾是相互独立的. 求:

- (1) 1 小时内至少有 1 台机床不需要照顾的概率;
- (2) 1 小时内至多有 1 台机床需要照顾的概率.

【1-55】 某股份公司董事会由 7 名董事组成. 现董事会欲对某事可行与否进行表决, 并按多数人意见作出决策. 若每个董事贡献正确意

见的百分比是 0.8,求董事会作出正确决策的概率(设每个董事都是独立提供意见的).

【1-56】 设 $0 < P(A) < 1, P(B|A) = P(B|\bar{A})$, 证明: A 与 B 相互独立.

【1-57】 设事件 A, B 相互独立, 两个事件仅发生 A 的概率或仅发生 B 的概率都是 0.25, 求 $P(A)$ 与 $P(B)$.

【1-58】 2 架轰炸机各自袭击一个目标, 它们的命中率分别为 0.7 与 0.8, 求至少命中一个目标的概率.

【1-59】 设有 2 门高射炮, 每门击中飞机的概率都是 0.6, 求同时发射一发炮弹而击中飞机的概率. 又若有一架敌机侵犯, 问至少需配备多少门这样的高射炮, 才能以 99% 的概率击中它?

【1-60】 冰箱厂甲、乙两条生产线分别生产了 520 台和 480 台冰箱, 它们的次品率分别为 0.005 和 0.007. 现从这 1000 台冰箱中任抽一台, 求抽得的是甲生产线生产的且为正品的概率.

【1-61】 人的血型为 O, A, B, AB 型的概率分别为 0.46, 0.40, 0.11, 0.03. 任意挑选 4 人, 求下列事件的概率:

- (1) 1 人为 O 型, 其他 3 人分别为其他三种血型;
- (2) 3 人为 O 型, 1 人为 B 型;
- (3) 没有 1 人为 AB 型.

【1-62】 某大卖场供应的微波炉中, 甲、乙、丙三厂的产品各占 50%、40%、10%, 而三厂产品的合格率分别为 95%、85%、80%, 求:

- (1) 买到的一台微波炉是合格品的概率;
- (2) 已知买到的微波炉是合格品, 则它是甲厂生产的概率为多少?

【1-63】 已知 3% 的男人和 0.5% 的女人是色盲者. 现随机挑选一人, 此人恰为色盲患者, 求此人是男人的概率(假设男女各占总人数的一半).

【1-64】 某一城市有 25% 的汽车废气排放量超过规定, 一辆废气排放量超标的汽车不能通过检验站检验的概率是 0.99, 但一辆废气排放量未超标的汽车也有 0.05 的概率不能通过检验. 求一辆未通过检验的汽车, 它的废气排放量超标的概率.

【1-65】* 某企业因技术和设备落后以及市场竞争激烈而不断亏损. 为改变这种状况, 该企业决定分别独立开发两种新产品, 其开发成功的希望分别为 80% 和 85%. 又根据各方面的情况分析认为:

(1) 若两项新产品皆研制成功, 该企业有 90% 的希望改变亏损状况;

(2) 若仅有一项研制成功, 该企业有 50% 的希望改变亏损状况;

(3) 若两项都未研制成功, 仅有 10% 的希望改变亏损状况.

试问该企业采取这种决策能改变亏损状况的把握有多大?

【1-66】* 多项选择题是各类考试中常采用的一种题型, 它要求考生在列出的 4 个备选答案中至少选出 2 个作为答案. 考生只有全部选对才算对, 错选、多选、漏选都算错. 设考生中确实理解题意而答对某多项选择题的比例为 30%, 而其余的考生只能乱猜. 已知某考生该道多项选择题答对了, 求他确实知道正确答案的概率.

答案与提示

【1-42】 选 B. 设 B, C 分别为甲, 乙命中目标, 则 $A = \overline{B}C, \overline{A} = B \cup \overline{C}$.

【1-43】 选 A. $AB = \emptyset \Rightarrow \overline{AB} = \overline{A} \cup \overline{B} = \Omega$.

【1-44】 选 D. $\overline{A \cup C} = \overline{A} \cap \overline{C}$.

【1-45】 选 D. 若 A, B 独立, 则 $\overline{A}, \overline{B}$ 也独立.

【1-46】 选 C. $P(A|B) + P(\overline{A}|\overline{B}) = 1 \Rightarrow P(AB) = P(A)P(B)$.

【1-47】 (1) $P(A\overline{B}) = 0.3$; (2) $P(A\overline{B} \cup \overline{A}B) = 0.75$. 其中 A ——订日报, B ——订晚报.

【1-48】 (1) $\frac{7}{24} \approx 0.292$; (2) $\frac{17}{24} \approx 0.708$; (3) $\frac{21}{120} = 0.175$.

均用超几何分布计算, 题(2)是题(1)的对立事件.

【1-49】 $\frac{21}{77} \approx 0.272$. 应用超几何分布, 相当于从 12 支球队中选出 6 支球队.

【1-50】 (1) $\frac{C_4^1 P_6^2}{P_7^3} = \frac{4}{7}$; (2) $\frac{C_5^1 P_6^2}{P_7^3} = \frac{5}{7}$. 本题属古典概型.

【1-51】 (1) $\frac{C_{12}^6 \cdot 6!}{12^6} \approx 0.2228$; (2) $\frac{C_{12}^2 (2^6 - 2)}{12^6} \approx 0.00137$;

(3) 0.7772. 题(3)是题(1)的对立事件.

【1-52】 (1) $P(A \cup B) = \frac{7}{9}$, A ——无红球, B ——无黄球. 本题还可用伯努利概型求解.

(2) $P(A \cup B) = \frac{11}{15} \approx 0.733$. 本题还可用超几何分布求解.

【1-53】 (1) $P(A_1 \cup A_2) = 1 - P(\overline{A_1})P(\overline{A_2}) = \frac{5}{9}$, A_i ——第 i 次取到新球 ($i=1, 2$). 本题还可用伯努利概型求解.

(2) $P(A_1 \cup A_2) = 1 - P(\overline{A_1})P(\overline{A_2} | \overline{A_1}) = \frac{3}{5}$. 本题还可用超几何分布求解.

【1-54】 (1) $P\left(\bigcup_{i=1}^4 A_i\right) = 1 - \prod_{i=1}^4 P(\overline{A_i}) = 0.9919$, A_i ——第 i 台机床需照顾 ($i=1 \sim 4$).

(2) $P(\overline{A_1}\overline{A_2}\overline{A_3}\overline{A_4} \cup A_1\overline{A_2}\overline{A_3}\overline{A_4} \cup \overline{A_1}A_2\overline{A_3}\overline{A_4} \cup \overline{A_1}\overline{A_2}A_3\overline{A_4} \cup \overline{A_1}\overline{A_2}\overline{A_3}A_4) = 0.0837$, 或 $P(X \leq 1) = P_4(0) + P_4(1) = 0.0837$, 其中 $X \sim B(4, 0.7)$.

【1-55】 $P(X \geq 4) = 1 - P(X < 4) = 1 - \sum_{k=0}^3 P_7(k) \approx 0.9667$, 其中 X ——作出正确意见的董事人数, $X \sim B(7, 0.8)$.

【1-56】 只需证明 $P(AB) = P(A)P(B)$.

【1-57】 $P(A) = P(B) = 0.5$. 设 $P(A) = x, P(B) = y$, 建立关于 x, y 的二元方程组并解之.

【1-58】 $p = 0.94$. 利用广义加法公式及独立性.

【1-59】 $p = 0.84, n = 6$ 门. 利用相互独立的 n 个事件的并的计算公式:

$$P\left(\bigcup_{i=1}^n A_i\right) = 1 - (1-p)^n, \quad p = P(A_i) \quad (i = 1 \sim n).$$

【1-60】 $P(AB) = 0.5174$. A ——甲生产线生产, B ——正品, 用乘法公式.

【1-61】 (1) $p \approx 0.0146$; (2) $p \approx 0.0428$; (3) $p = 0.8853$. 利用独立性.

【1-62】 (1) $P(A) = 0.895$. 全概率公式;

$$(2) P(B_1|A) = \frac{95}{179} \approx 0.5307, \text{ 贝叶斯公式.}$$

【1-63】 $P(B|A) = \frac{6}{7}$, A ——选到的是色盲者, B ——选到的是男人, 贝叶斯公式.

【1-64】 $P(B|A) = \frac{33}{38} \approx 0.8684$, A ——汽车未通过检验, B ——废气排放量超标, 贝叶斯公式.

【1-65】* 设事件 $A = \{\text{采取此决策改变亏损状况}\}$,

$$B_i = \{\text{第 } i \text{ 项新产品研制成功}\} \quad (i=1, 2),$$

则 B_1, B_2 相互独立, 且 $B_1 B_2, B_1 \overline{B_2}, \overline{B_1} B_2, \overline{B_1} \overline{B_2}$ 构成完备事件组, 由全概率公式可得

$$\begin{aligned} P(A) &= P(B_1 B_2)P(A|B_1 B_2) + P(B_1 \overline{B_2})P(A|B_1 \overline{B_2}) \\ &\quad + P(\overline{B_1} B_2)P(A|\overline{B_1} B_2) + P(\overline{B_1} \overline{B_2})P(A|\overline{B_1} \overline{B_2}) \\ &= 0.8 \times 0.85 \times 0.9 + 0.8 \times 0.15 \times 0.5 + 0.2 \times 0.85 \times 0.5 \\ &\quad + 0.2 \times 0.15 \times 0.1 \\ &= 0.76. \end{aligned}$$

【1-66】* 设事件 $A = \{\text{答对该道题}\}$, $B = \{\text{确实理解题意而答对题}\}$, 显然 $P(A|B) = 1$, $P(A|\overline{B}) = \frac{1}{C_4^2 + C_4^3 + C_4^4} = \frac{1}{11}$. 由全概率公式可得

$$\begin{aligned} P(A) &= P(B)P(A|B) + P(\overline{B})P(A|\overline{B}) \\ &= \frac{3}{10} \times 1 + \frac{7}{10} \times \frac{1}{11} = \frac{40}{110}, \end{aligned}$$

由贝叶斯公式即得所求概率

$$P(B | A) = \frac{P(B)P(A | B)}{P(A)} = \frac{3}{10} \times \frac{110}{40} = \frac{33}{40} = 0.825.$$

关键词

随机变量 随机变量的分布函数 离散型随机变量的分布律 连续型随机变量的密度函数 0-1 分布 二项分布 泊松分布 均匀分布 指数分布 正态分布 随机变量函数的分布

教学要求

- (1) 理解随机变量的概念, 会计算由随机变量表达的随机事件的概率.
- (2) 理解随机变量的分布函数的概念, 了解分布函数的性质, 会用分布函数计算概率.
- (3) 理解离散型随机变量及其分布律的概念, 了解分布律的性质, 掌握常见离散分布的应用.
- (4) 理解连续型随机变量及其密度函数的概念, 了解密度函数的性质, 掌握常见连续分布的应用.
- (5) 已知分布函数, 会求分布律或密度函数; 已知分布律或密度函数, 会求分布函数.
- (6) 会求简单的随机变量的函数的分布.

重点

随机变量的概念; 分布函数、分布律和密度函数的概念及性质; 分布函数、分布律和密度函数的计算.

难点

用随机变量描述随机事件; 求随机变量的分布函数、分布律和密度

函数;求随机变量的函数的分布函数、分布律和密度函数.

方法提要

(1) 当分布函数 $F(x)$ 中含有待定常数时,常利用

$$\lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow +\infty} F(x) = 1, \quad F(x_0 + 0) = F(x_0)$$

来确定该常数;

当分布律或密度函数含有待定常数时,常利用

$$\sum_{k=1}^{\infty} P(X = x_k) = 1 \quad \text{或} \quad \int_{-\infty}^{+\infty} f(x) dx = 1$$

来确定该常数.

(2) 求离散型随机变量 X 的分布律时,首先要确定 X 所有可能取的值,其次求出对应于各取值的事件的概率,最后注意验证 $\sum_{k=1}^{\infty} P(X = x_k) = 1$, 否则不正确.

(3) 求离散型随机变量 X 的分布函数 $F(x)$ 时,若 X 的可能取值为 n 个,则 $F(x)$ 为 $n+1$ 段分段函数,其图形是右连续的阶梯曲线.

(4) 由密度函数 $f(x)$ 求分布函数 $F(x)$,要在相应的区间段把 $F(x)$ 写成 $f(x)$ 的变上限积分.

(5) 已知 X 的分布,求 $Y=g(X)$ 的分布:

① 若 $X \sim P(X=x_i)=p_i \quad (i=1,2,\cdots)$, 则

$$P(Y=y_j) = \sum_{g(x_i)=y_j} P(X=x_i) \quad (j=1,2,\cdots);$$

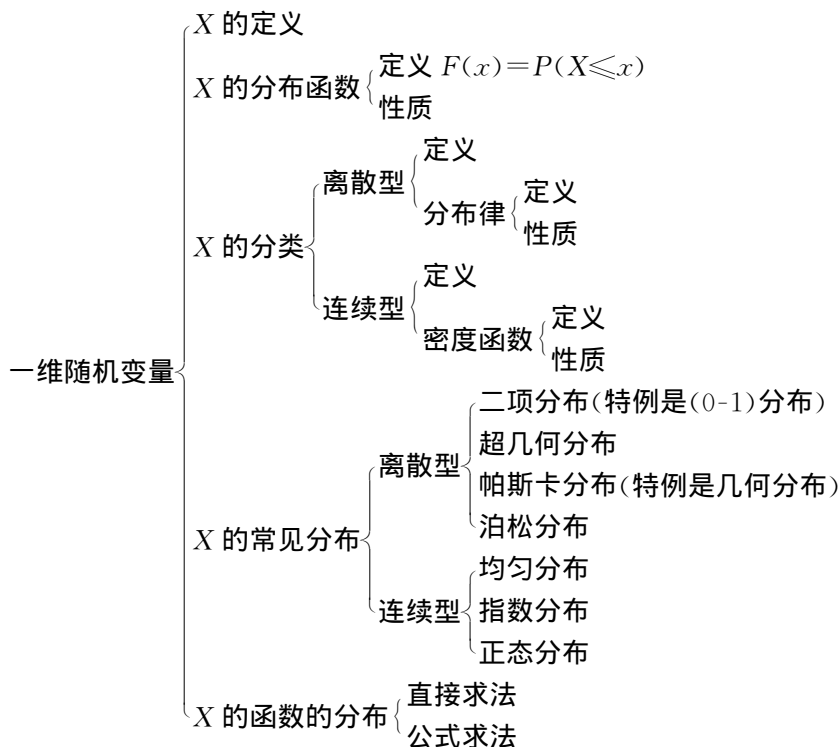
② 若 $X \sim f(x)$, 则

$$f_Y(y) = f(x) |x'|,$$

当 $y=g(x)$ 严格单调,反函数 $x=g^{-1}(y)$ 有连续导数;或

$$f_Y(y) = \sum_{i=1}^{\infty} f(x_i) |x'_i|,$$

其中 $y=g(x)$ 在不相重叠的区间 $I_1, I_2, \cdots$ 上逐段严格单调,其反函数 $x_i=g_i^{-1}(y)$ 均有连续导数;或先求 $F_Y(y)=P(g(X) \leq y)$, 再得 $f_Y(y) = F'_Y(y)$.



A 例题精解

【2-1】 一批产品分一、二、三等, 其中一等品是二等品的 3 倍, 三等品是二等品的 $\frac{1}{6}$. 从这批产品中随机地抽取一个检验质量, 用随机变量描述检验的可能结果, 并写出其分布律和分布函数.

解 随机变量 $X=i$ 表示抽检到的是 i 等品, $i=1, 2, 3$. 这三个事件的概率可由一、二、三等品的比例 18:6:1 求得:

$$P(X=1) = \frac{18}{25} = 0.72,$$

$$P(X=2) = \frac{6}{25} = 0.24,$$

$$P(X=3) = \frac{1}{25} = 0.04.$$

X 的分布律为

X	1	2	3
P	0.72	0.24	0.04

X 的分布函数

$$F(x) = \begin{cases} 0 & (x < 1), \\ 0.72 & (1 \leq x < 2), \\ 0.96 & (2 \leq x < 3), \\ 1 & (x \geq 3). \end{cases}$$

【2-2】 一批零件中有 9 个正品和 3 个次品. 安装机器时从这批零件中任取一个, 若取出次品不再放回, 求在取得正品以前已取得次品数的概率分布.

解 设在取得正品以前已取得的次品数为 X , 由题设可知, X 的可能取值为 0, 1, 2, 3.

事件 $(X=k)$ 表示前 $k-1$ 次取到次品, 第 k 次首次取到正品. 因 $0 \leq k-1 \leq 3$, 故 $1 \leq x \leq 4$, 于是记 $A_i = \{\text{第 } i \text{ 次取得正品}\} (i=1, 2, 3, 4)$, 则

$$P(X=0) = P(A_1) = \frac{3}{4},$$

$$P(X=1) = P(\overline{A_1}A_2) = P(\overline{A_1})P(A_2 | \overline{A_1}) = \frac{3}{12} \times \frac{9}{11} = \frac{9}{44},$$

$$\begin{aligned} P(X=2) &= P(\overline{A_1}\overline{A_2}A_3) = P(\overline{A_1})P(\overline{A_2} | \overline{A_1})P(A_3 | \overline{A_1}\overline{A_2}) \\ &= \frac{3}{12} \times \frac{2}{11} \times \frac{9}{10} = \frac{9}{220}, \end{aligned}$$

$$P(X=3) = P(\overline{A_1}\overline{A_2}\overline{A_3}A_4) = \frac{3}{12} \times \frac{2}{11} \times \frac{1}{10} \times \frac{9}{9} = \frac{1}{220}.$$

故 X 的概率分布为

X	0	1	2	3
P	$\frac{3}{4}$	$\frac{9}{44}$	$\frac{9}{220}$	$\frac{1}{220}$

点评 把事件 $(X=k)$ 表达成 $\bar{A}_1 \cdots \bar{A}_{k-1} A_k$, 然后利用乘法公式是计算本题的关键. 这是一种常用的方法. 当然, 对较简单的题目可不必转化而直接计算, 如题 2-1 和 2-3.

【2-3】 从 5 个数 1, 2, 3, 4, 5 中任取 3 个数, 记为 a, b, c . 求:

(1) $X = \min\{a, b, c\}$ 的分布律及 $P(X > 3)$;

(2) $Y = \max\{a, b, c\}$ 的分布律及 $P(Y \leq 4)$.

解 (1) X 的可能取值为 1, 2, 3, 取这些值的概率可通过古典概型求出. 样本空间中基本事件总数 $n = C_5^3 = 10$, 有利于事件 $(X=1)$, $(X=2)$, $(X=3)$ 发生的基本事件数分别为

$$k_1 = C_4^2 = 6, \quad k_2 = C_3^2 = 3, \quad k_3 = C_2^2 = 1,$$

由此得

$$P(X=1) = 0.6, \quad P(X=2) = 0.3, \quad P(X=3) = 0.1,$$

即

$$X \sim \begin{pmatrix} 1 & 2 & 3 \\ 0.6 & 0.3 & 0.1 \end{pmatrix}.$$

显然 $P(X > 3) = 0$.

(2) 仿题(1)可求得

$$Y \sim \begin{pmatrix} 3 & 4 & 5 \\ 0.1 & 0.3 & 0.6 \end{pmatrix} \quad P(Y \leq 4) = 0.4.$$

【2-4】 设某种独立试验每次试验成功的概率为 0.75, 以 X 表示首次成功所需试验次数. 试写出 X 的分布律, 并计算 X 取偶数的概率.

解 方法 1 设事件 $A_k = \{\text{第 } k \text{ 次试验成功}\} (k=1, 2, 3, \cdots)$, 则 $P(A_k) = 0.75$, 事件 $(X=k) = \bar{A}_1 \bar{A}_2 \cdots \bar{A}_{k-1} A_k$, 由题设可知各事件相互独立, 从而得 X 的分布律

$$P(X=k) = P(\bar{A}_1)P(\bar{A}_2) \cdots P(\bar{A}_{k-1})P(A_k)$$

$$= 0.75 \times 0.25^{k-1} \quad (k = 1, 2, 3, \dots),$$

$$\begin{aligned} P(X \text{ 取偶数}) &= \sum_{n=1}^{\infty} P(X = 2n) = \sum_{n=1}^{\infty} 0.75 \times 0.25^{2n-1} \\ &= \frac{0.75 \times 0.25}{1 - 0.25^2} = 0.2. \end{aligned}$$

方法 2 由题设可知 X 服从参数 $p=0.75$ 的几何分布:

$$P(X = k) = p(1-p)^{k-1} \quad (k = 1, 2, 3, \dots).$$

所以 X 的分布律为

$$P(X = k) = 0.75(1 - 0.75)^{k-1} \quad (k = 1, 2, 3, \dots).$$

由上述分布律可知 $P(X \text{ 取奇数}) = 4P(X \text{ 取偶数})$,

再由分布律的归一性可知 $P(X \text{ 取奇数}) + P(X \text{ 取偶数}) = 1$,

于是得

$$P(X \text{ 取偶数}) = 0.2.$$

点评 无论是求分布律还是计算概率,方法 2 都优于方法 1. 概率论中最常见的分布也就是十几个,不仅要记住这些分布的公式,更要清楚每个常见分布的应用场合. 比如几何分布应用于:独立试验中的试验首次成功(或失败)的场合.

本题要求的是取值无穷的分布律,一般不用表格形式表达,而用如下解析形式表达:

$$P(X = k) = p_k \quad (k = 1, 2, 3, \dots).$$

初学者往往只记得写出上式的前面一半,而忽略后面一半,这是犯“捡了芝麻,丢了西瓜”的错误. 因为说白了,分布律就是一群概率,前面一半只是一个概率,后面才是一群.

【2-5】 某人有 10 把外形相似的钥匙,其中只有 1 把能打开房门,但不知是哪一把,只好逐把试开. 求此人直至将门打开所需的试开次数的分布律.

解 设随机变量 X 为试开次数,则 X 可能取值为 $1, 2, \dots, 10$. 设事件 $A_k = \{\text{第 } k \text{ 次试开将门打开}\}, k = 1, 2, \dots, 10$. 由乘法公式得 X 的分布律

$$P(X = k) = P(\overline{A_1} \overline{A_2} \cdots \overline{A_{k-1}} A_k)$$

$$\begin{aligned}
 &= P(\overline{A_1})P(\overline{A_2} | \overline{A_1}) \cdots P(\overline{A_{k-1}} | \overline{A_1} \overline{A_2} \cdots \overline{A_{k-2}}) \cdot \\
 &\quad P(A_k | \overline{A_1} \overline{A_2} \cdots \overline{A_{k-1}}) \\
 &= \frac{9}{10} \times \frac{8}{9} \times \cdots \times \frac{10-k+1}{10-k+2} \times \frac{1}{10-k+1} \\
 &= \frac{1}{10} \quad (k=1, 2, \cdots, 10).
 \end{aligned}$$

【2-6】 $P(X=k)=\theta\lambda^k$ ($k=1, 2, \cdots$) 为 X 的分布律的充要条件为_____.

(A) $\theta>0$ 且 $0<\lambda<1$;

(B) $\theta=1-\lambda$ 且 $0<\lambda<1$;

(C) $\theta=\lambda^{-1}-1$ 且 $\lambda<1$;

(D) $\lambda=(1+\theta)^{-1}$ 且 $\theta>0$.

解 由归一性可知

$$\sum_{k=1}^{\infty} \theta\lambda^k = \theta\lambda(1+\lambda+\lambda^2+\cdots) = \frac{\theta\lambda}{1-\lambda} = 1,$$

解得 $\lambda=(1+\theta)^{-1}$. 由非负性可知 $\theta\lambda^k \geq 0$, 得 $\theta>0$, 所以应选 D.

【2-7】 设随机变量 X 的分布律

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

其中 $\lambda>0$ 是已知常数, 试确定常数 a .

解 根据常用函数的泰勒(Taylor)级数公式

$$e^x = 1 + x + \frac{x^2}{2} + \cdots + \frac{x^k}{k!} + \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad (-\infty < x < \infty)$$

以及分布律的归一性, 可得

$$\sum_{k=0}^{\infty} a \frac{\lambda^k}{k!} = a \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = ae^\lambda = 1,$$

所以

$$a = e^{-\lambda}.$$

【2-8】 抛掷硬币 10 次, 求:

(1) 国徽向上次数 X 的分布律;

(2) 国徽向上次数不小于 2 的概率.

解 (1) 每次抛掷硬币国徽向上的概率均为 0.5, 抛掷 10 次可视为进行 10 次独立试验, 故 $X \sim B(10, 0.5)$, 其分布律

$$P(X=k) = C_{10}^k \cdot 0.5^k (1-0.5)^{10-k}$$

$$= C_{10}^k \cdot 0.5^{10} \quad (k=0,1,\cdots,10).$$

$$(2) P(X \geq 2) = 1 - P(X < 2) = 1 - P(X=0) - P(X=1)$$

$$= 1 - 0.5^{10} - C_{10}^1 \cdot 0.5^{10} \approx 0.9893.$$

【2-9】 同时掷 2 颗骰子,求:

(1) 2 颗骰子中出现的较大点数 X 的分布律;

(2) 2 颗骰子中出现的较小点数 Y 的分布律.

解 (1) 基本事件总数 $n=6^2=36$.

事件 $(X=1)$ 仅包含 1 个基本事件;

事件 $(X=k) \quad (k=2,3,4,5,6)$ 分两种情况:一是 2 颗骰子都出现 k 点,含 1 个基本事件;二是一颗骰子出现 k 点,另一颗骰子的点数小于 k ,含 $C_2^1 C_{k-1}^1$ 个基本事件.故有利于事件 $(X=k)$ 的基本事件数合计为 $1 + C_2^1 C_{k-1}^1 = 2k-1$.

由古典概率的定义可得

$$P(X=k) = \frac{2k-1}{36} \quad (k=1,2,3,4,5,6),$$

即 X 的分布律

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{36} & \frac{1}{12} & \frac{5}{36} & \frac{7}{36} & \frac{1}{4} & \frac{11}{36} \end{pmatrix}.$$

(2) 相仿可得

$$P(Y=m) = \frac{13-2m}{36} \quad (m=1,2,3,4,5,6),$$

即

$$Y \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{11}{36} & \frac{1}{4} & \frac{7}{36} & \frac{5}{36} & \frac{1}{12} & \frac{1}{36} \end{pmatrix}.$$

【2-10】 一批产品包括 50 件正品、3 件次品. 现有放回地随机抽取,每次一件,直到取得正品为止. 求抽取次数 X 的分布律.

解 由于是放回抽样,所以 X 的可能取值为无限个. 事件 $(X=k)$ 表示第 k 次抽到正品,前 $k-1$ 次抽到次品,且各事件相互独立,从而得

X 的分布律

$$P(X=k) = \frac{50}{53} \left(\frac{3}{53} \right)^{k-1} \quad (k=1, 2, \dots).$$

点评 本题是放回抽样,故属于几何分布;若改成不放回抽样,则是取值有限的分布.

【2-11】 自动生产线在调整之后出现废品的概率为 p ,当生产过程中出现废品时立即重新进行调整.求在两次调整之间生产的合格品数 X 的概率分布.

解 方法 1 事件 $(X=0)$ 表示调整后生产的第一个产品是废品,则

$$P(X=0) = p;$$

事件 $(X=1)$ 表示调整后生产的第一个产品是合格品,而第二个产品是废品,则

$$P(X=1) = p(1-p);$$

事件 $(X=2)$ 表示调整后生产的前两个产品是合格品,而第三个产品是废品,则

$$P(X=2) = p(1-p)^2.$$

依次类推,可得两次调整之间生产的合格品数 X 的分布律

$$P(X=k) = p(1-p)^k \quad (k=0, 1, 2, \dots).$$

方法 2 由题设可知,事件 $(X=n)$ 表示首次出现废品之前已生产了 n 个合格品,即为“首次成功”的概率模型,故只要在几何分布 $(P(X=k)=p(1-p)^{k-1}, k=1, 2, \dots)$ 中,将 $k-1$ 换成 k ,于是得 X 的分布律

$$P(X=k) = p(1-p)^k \quad (k=0, 1, 2, \dots).$$

【2-12】 已知胃癌患者占 2% ,试求:

(1) 为发现一例胃癌患者至少要检查 50 人的概率;

(2) 为使发现一例患胃癌者的概率不小于 80% ,至少要对多少人作胃镜检查?

解 设 X 表示首次发现一例胃癌患者所需检查的人数,这是“首次成功”的概率模型,故 X 服从参数 $p=0.002$ 的几何分布: $X \sim$

$G(0.002)$.

$$\begin{aligned}(1) P(X \geq 50) &= \sum_{k=50}^{\infty} P(X=k) = \sum_{k=50}^{\infty} p(1-p)^{k-1} \\&= p(1-p)^{49} [1 + (1-p) + (1-p)^2 + \cdots] \\&= p(1-p)^{49} \times \frac{1}{1-(1-p)} \\&= (1-p)^{49} = 0.998^{49} \approx 0.9066.\end{aligned}$$

(2) 设至少要对 n 个人作胃镜检查,依题意应有

$$P(X \leq n) \geq 0.8.$$

而

$$\begin{aligned}P(X \leq n) &= 1 - P(X > n) = 1 - \sum_{k=n+1}^{\infty} P(X=k) \\&= 1 - \sum_{k=n+1}^{\infty} p(1-p)^{k-1} = 1 - (1-p)^n \\&= 1 - 0.998^n,\end{aligned}$$

故 $1 - 0.998^n \geq 0.8$, 解得 $n \geq \frac{\ln 0.2}{\ln 0.998} \approx 804.7$, 即至少要对 805 人作胃镜检查才能使发现一例胃癌患者的概率不小于 80%.

【2-13】 一电话交换台每分钟的呼唤次数 $X \sim \pi(4)$. 求:

- (1) 每分钟恰有 6 次呼唤的概率;
- (2) 每分钟的呼唤次数不超过 10 次的概率.

解 (1) $P(X=6) = \frac{4^6}{6!} e^{-4} \approx 0.104$;

$$(2) P(X \leq 10) = \sum_{k=0}^{10} P(X=k) = 1 - \sum_{k=11}^{\infty} \frac{4^k}{k!} e^{-4},$$

查泊松分布表, 当 $\lambda=4$ 时, $\sum_{k=11}^{\infty} \frac{4^k}{k!} e^{-4} = 0.0028$, 所以

$$P(X \leq 10) \approx 1 - 0.003 = 0.997.$$

【2-14】 一总机共有 200 台分机, 拥有外线 14 条. 若每台分机向总机要外线的概率为 2.5%, 求:

- (1) 每台分机向总机要外线时, 能及时得到满足的概率;

(2) 同时向总机要外线的分机的最可能台数.

解 设向总机要外线的分机台数为 X , 每台分机要外线的概率 $p=0.025$, 故可将 200 台分机要外线视为 200 重伯努利试验, 即 $X \sim B(200, 0.025)$.

(1) 泊松分布是二项分布(即伯努利分布)的极限分布, 故当 $n=200$ 较大, $p=0.025$ 较小时, $X \stackrel{\text{近似}}{\sim} \pi(\lambda)$, $\lambda=np=5$. 所求概率的事件等价于事件 $(X \leq 14)$, 于是所求概率

$$P(X \leq 14) = 1 - P(X > 14) \approx 1 - \sum_{k=15}^{\infty} \frac{5^k}{k!} e^{-5} \approx 0.9998.$$

$$(2) k = [(n+1)p] = [201 \times 0.025] = [5.025] = 5,$$

所以同时向总机要外线的分机的最可能台数为 5.

【2-15】 某急救中心在长度为 t 的时间间隔内收到紧急呼救的次数 $X \sim \pi\left(\frac{t}{2}\right)$, 而与时间间隔的起点无关(时间以小时计). 试求:

(1) 某天 12:00~15:00 没有收到紧急呼救的概率;

(2) 某天 12:00~17:00 至少收到 1 次紧急呼救的概率.

解 (1) 当 $t=15-12=3$ 时, $X \sim \pi\left(\frac{3}{2}\right)$, 从而所求概率

$$P(X=0) = \frac{\left(\frac{3}{2}\right)^0}{0!} e^{-\frac{3}{2}} = e^{-\frac{3}{2}}.$$

(2) 当 $t=17-12=5$ 时, $X \sim \pi\left(\frac{5}{2}\right)$, 从而所求概率

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(X=0) = 1 - e^{-\frac{5}{2}}.$$

【2-16】 设连续型随机变量 X 的分布函数

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

求: (1) 系数 A ;

$$(2) P(X < 2), P(0 < X \leq 3), P\left(2 < X < \frac{5}{2}\right);$$

(3) X 的概率密度函数 $f(x)$.

解 (1) 方法 1 (归一性)

$$\text{由 } \int_{-\infty}^{+\infty} f(x) dx = \int_1^e f(x) dx = F(x) \Big|_1^e = A \ln x \Big|_1^e = A(1-0) = 1 \Rightarrow A = 1.$$

方法 2 (连续性)

$$\text{由 } F(e-0) = F(e) \Rightarrow A \ln e = 1 \Rightarrow A = 1 \text{ (其中 } F(e-0) = \lim_{x \rightarrow e^-} F(x) \text{)}.$$

方法 3 (有关命题: 连续随机变量取任何定值的概率为零)

$$\text{由 } P(X=e) = F(e) - F(e-0) = 1 - A \ln e = 0 \Rightarrow A = 1.$$

$$(2) P(X < 2) = P(X \leq 2) = F(2) = \ln 2;$$

$$P(0 < X \leq 3) = F(3) - F(0) = 1 - 0 = 1;$$

$$\begin{aligned} P\left(2 < X < \frac{5}{2}\right) &= P\left(2 < X \leq \frac{5}{2}\right) = F\left(\frac{5}{2}\right) - F(2) \\ &= \ln \frac{5}{2} - \ln 2 = \ln \frac{5}{4}. \end{aligned}$$

$$(3) f(x) = F'(x) = \begin{cases} \frac{1}{x} & (1 < x < e) \\ 0 & (\text{其他}). \end{cases}$$

点评 本题求系数 A 牵涉到分布函数 $F(x)$ 与密度函数 $f(x)$ 的性质及两者间的关系等, 概念性很强, 稍有疏忽便会出错. 例如下面的等式与推导都是不成立的:

$$\textcircled{1} F(x) = \int_1^e f(x) dx = A - 0 = 1;$$

$$\textcircled{2} F(+\infty) = 1 \Rightarrow A \ln e = 1;$$

$$\textcircled{3} \int_{-\infty}^{+\infty} A \ln x dx = A(x \ln x - x) \Big|_1^e = A = 1;$$

$$\textcircled{4} F(e+0) = F(e) \Rightarrow A \ln e = 1;$$

$$\textcircled{5} P(X=e) = A = 1.$$

关于题(3)中的求导问题, 详见题 2-19 中第(2)小题及后面的点评.

【2-17】 设连续型随机变量 X 的密度函数

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

且 $P(X < 1) = \frac{1}{2}$, 求:

(1) 系数 A, B ;

(2) X 的分布函数 $F(x)$.

解 (1) 由 $P(X < 1) = \int_0^1 Ax \, dx = \left. \frac{A}{2} x^2 \right|_0^1 = \frac{A}{2} = \frac{1}{2} \Rightarrow A = 1$,

再由 $\int_{-\infty}^{+\infty} f(x) \, dx = \int_0^1 x \, dx + \int_1^2 (B-x) \, dx = \frac{1}{2} + B - \frac{3}{2} = 1 \Rightarrow B = 2$.

(2) 当 $x < 0$ 时, $F(x) = \int_{-\infty}^x f(t) \, dt = \int_{-\infty}^x 0 \, dt = 0$;

当 $0 \leq x < 1$ 时, $F(x) = \int_{-\infty}^x f(t) \, dt = \int_{-\infty}^0 0 \, dt + \int_0^x t \, dt = \frac{x^2}{2}$;

当 $1 \leq x < 2$ 时, $F(x) = \int_{-\infty}^0 0 \, dt + \int_0^1 t \, dt + \int_1^x (2-t) \, dt = -\frac{x^2}{2} + 2x - 1$;

当 $x \geq 2$ 时, $F(x) = \int_{-\infty}^0 0 \, dt + \int_0^1 t \, dt + \int_1^2 (2-t) \, dt + \int_2^x 0 \, dt = 1$.

综合得

$$F(x) = \begin{cases} 0 & (x < 0), \\ \frac{x^2}{2} & (0 \leq x < 1), \\ -\frac{x^2}{2} + 2x - 1 & (1 \leq x < 2), \\ 1 & (x \geq 2). \end{cases}$$

点评 分布函数 $F(x)$ 通常定义在 $(-\infty, +\infty)$ 上; 当密度函数 $f(x)$ 是 n 段分段函数时, 则 $F(x)$ 为 $n+1$ 段分段函数.

本题最易产生错误之外:

当 $1 \leq x < 2$ 时, $F(x) = \int_1^x (2-t) \, dt = -\frac{x^2}{2} + 2x - \frac{3}{2}$;

当 $x \geq 2$ 时, $F(x) = \int_2^x 0 \, dt = 0$.

上述错误起因于对分布函数的概念理解不清. 例如, 当 $1 \leq x < 2$ 时, $F(x) = P(X \leq x)$ 是表示在区间 $(-\infty, x]$ 上对 $f(x)$ 的积分, 而不是在区间 $[1, x]$ 上对 $f(x)$ 的积分.

【2-18】 某型号电子管的寿命 X (以小时计) 的密度函数

$$f(x) = \begin{cases} \frac{1000}{x^2} & (x > 1000), \\ 0 & (x \leq 1000). \end{cases}$$

某一电子设备内装有 3 个这样的电子管, 试求电子管使用 1500 小时内没有一个损坏的概率和只有一个损坏的概率.

分析 本题属伯努利概型, 损坏的电子管数 $Y \sim B(3, p)$, 参数 $p = P(X < 1500)$ 为电子管在 1500 小时内损坏的概率, 它可以通过密度函数的积分求出, 然后计算 $P_3(0)$ 与 $P_3(1)$.

$$\begin{aligned} \text{解 } P(X < 1500) &= \int_{-\infty}^{1500} f(x) \, dx = \int_{1000}^{1500} \frac{1000}{x^2} \, dx = -\frac{1000}{x} \Big|_{1000}^{1500} \\ &= -\frac{2}{3} + 1 = \frac{1}{3}, \end{aligned}$$

所求概率

$$P(\text{无损坏}) = P_3(0) = C_3^0 \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^3 = \frac{8}{27},$$

$$P(\text{只有一个损坏}) = P_3(1) = C_3^1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^2 = \frac{4}{9}.$$

【2-19】 设连续型随机变量 X 的分布函数

$$F(x) = \begin{cases} 0, & x < -a; \\ A + B \arcsin \frac{x}{a}, & -a \leq x < a; \\ 1, & x \geq a \end{cases} \quad (a > 0).$$

求: (1) 系数 A, B ; (2) 密度函数 $f(x)$.

解 (1) 由于 X 是连续型随机变量, 故分布函数 $F(x)$ 连续, 特别在 $x = -a$ 和 $x = a$ 处连续. 因此

$$F(-a+0) = F(-a-0), \quad F(a-0) = F(a),$$

$$\text{即} \quad A - \frac{\pi}{2}B = 0, \quad A + \frac{\pi}{2}B = 1,$$

$$\text{解得} \quad A = \frac{1}{2}, \quad B = \frac{1}{\pi}.$$

$$(2) \quad f(x) = F'(x) = \begin{cases} \frac{1}{\pi \sqrt{a^2 - x^2}} & (-a < x < a), \\ 0 & (\text{其他}). \end{cases}$$

点评 对分段函数的求导,在其各个子开区间内可直接按初等函数的求导公式和法则求导,而在其分段点处应按导数的定义求导.故应有

$$f(x) = F'(x) = \begin{cases} 0 & (x < -a), \\ \frac{1}{\pi \sqrt{a^2 - x^2}} & (-a < x < a), \\ 0 & (x > a). \end{cases}$$

当 $x = -a$ 时,其左右导数分别为

$$\begin{aligned} F'_-(-a) &= \lim_{x \rightarrow -a^-} \frac{F(x) - F(-a)}{x - (-a)} = \lim_{x \rightarrow -a^-} \frac{0 - \left[\frac{1}{2} + \frac{1}{\pi} \left(-\frac{\pi}{2} \right) \right]}{x + a} = 0, \\ F'_+(-a) &= \lim_{x \rightarrow -a^+} \frac{F(x) - F(-a)}{x - (-a)} \\ &= \lim_{x \rightarrow -a^+} \frac{\frac{1}{2} + \frac{1}{\pi} \arcsin \frac{x}{a} - \left(\frac{1}{2} - \frac{1}{\pi} \times \frac{\pi}{2} \right)}{x + a} \\ &= \lim_{x \rightarrow -a^+} \frac{\frac{1}{2} + \frac{1}{\pi} \arcsin \frac{x}{a}}{x + a} \stackrel{\left(\frac{0}{0} \right)}{=} \lim_{x \rightarrow -a^+} \frac{1}{\pi \sqrt{a^2 - x^2}} = \infty, \end{aligned}$$

可见 $F(x)$ 在点 $x = -a$ 处不可导. 同理, $F(x)$ 在点 $x = a$ 处也不可导. 在这两点处可补充定义: 当 $x = -a$ 和 $x = a$ 时, $f(x) = 0$ (补充定义其他非零值也可以, 因为这两个点的函数值并不影响 $f(x)$ 在一个区间上的积分值). 此时有 $F(x) = \int_{-\infty}^x f(t) dt$ 成立, 因此 X 的密度函数

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

所以,由 $F(x)$ 求 $f(x)$ 可直接对 $F(x)$ 按求导公式和法则求导,而在分段点处一律令 $f(x) = F'(x) = 0$. 前面题 2-16 就是按此方法处理的.

【2-20】 服从拉普拉斯(Laplace)分布的随机变量 X 的密度函数

$$f(x) = Ae^{-|x|}, \quad x \in (-\infty, +\infty).$$

求系数 A 及分布函数 $F(x)$.

解 由 $\int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^{+\infty} Ae^{-|x|} dx = 2A \int_0^{+\infty} e^{-x} dx = 2A = 1 \Rightarrow$
 $A = \frac{1}{2}.$

当 $x < 0$ 时

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^x \frac{1}{2} e^t dt = \frac{1}{2} e^t \Big|_{-\infty}^x = \frac{1}{2} e^x,$$

当 $x \geq 0$ 时

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(t) dt = \int_{-\infty}^0 \frac{1}{2} e^t dt + \int_0^x \frac{1}{2} e^{-t} dt \\ &= \frac{1}{2} e^t \Big|_{-\infty}^0 - \frac{1}{2} e^{-t} \Big|_0^x = \frac{1}{2} - \frac{1}{2} e^{-x} + \frac{1}{2} \\ &= 1 - \frac{1}{2} e^{-x}, \end{aligned}$$

从而得

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

点评 当密度函数 $f(x)$ 的自变量带绝对值符号时,求分布函数 $F(x)$ 计算变限积分前,需将定义域分为 $x < 0$ 与 $x \geq 0$ 两部分.

【2-21】 公共汽车站每隔 10 分钟有一辆汽车通过,乘客在 10 分钟内任一时刻到达汽车站是等可能的. 求乘客候车时间不超过 7 分钟

的概率. 下面三种解法哪个正确?

解法 1 本题属古典概型. 由题设知 $n=10, k=7$, 故

$$P(\text{候车不超过 } 7 \text{ 分钟}) = \frac{k}{n} = \frac{7}{10} = 0.7.$$

解法 2 设乘客到达车站时刻为 X , 由题设知 $X \sim U(0, 10)$,

$$P(\text{候车不超过 } 7 \text{ 分钟}) = P(0 < X < 7) = \int_0^7 \frac{1}{10} dx = 0.7.$$

解法 3 设乘客到达车站时刻为 X , 由题设知 $X \sim U(0, 10)$,

$$P(\text{候车不超过 } 7 \text{ 分钟}) = P(3 < X < 10) = \int_3^{10} \frac{1}{10} dx = 0.7.$$

答 本题样本空间中基本事件总数无限, 故不属于古典概型, 解法 1 错误; 解法 2 也错, 因为乘客在两班汽车间的 $0 \sim 7$ 分钟内到达, 并不意味着候车时间不超过 7 分钟; 所以只有解法 3 是正确的.

【2-22】 秒表的最小刻度值为 $1\% \text{ s}$ (秒). 若计时精度取最近的刻度值, 求使用该秒表计时产生的随机误差 X 的概率密度函数, 并计算误差的绝对值不超过 $4\%_0 \text{ s}$ 的概率.

分析 若秒针正好停在两个最小刻度值中间, 不管取前一个刻度值还是后一个刻度值, 都将产生最大的误差 -0.005 s 与 0.005 s . 由于秒针停在何处是十分随机的, 换言之, 停在两个最小刻度值之间的任何一个点处都是等可能的, 所以可认为随机误差 X 服从区间 $[-0.005, 0.005]$ 上的均匀分布.

解 由题设知 $X \sim U[-0.005, 0.005]$, 则其密度函数

$$f(x) = \begin{cases} 100 & (-0.005 \leq x \leq 0.005), \\ 0 & (\text{其他}), \end{cases}$$

所求概率

$$P(|X| < 0.004) = \int_{-0.004}^{0.004} 100 dx = 0.8.$$

【2-23】 设 $Y \sim U(a, 5)$, 且方程 $x^2 + Yx + 0.75Y + 1 = 0$ 没有实根的概率为 0.25, 求常数 a .

解 题设一元二次方程无实根, 则其判别式

$$\Delta = Y^2 - 4(0.75Y + 1) < 0,$$

解得 $-1 < Y < 4$. 因 $Y \sim U(a, 5)$, 其密度函数

$$f(y) = \begin{cases} \frac{1}{5-a} & (a < y < 5), \\ 0 & (\text{其他}), \end{cases}$$

所以由 $P(-1 < Y < 4) = 0.25$ 可得

$$\int_{-1}^4 f(y) dy = \int_{-1}^4 \frac{1}{5-a} dy = \frac{5}{5-a} = 0.25,$$

解得 $a = -15$.

点评 上述解法都是粗心人的解法, 它只求出了一个解, 还漏了另外一个解. 事实上由于 a 是个未确定的常数, 在求得 $-1 < Y < 4$ 后, 需要在三个区间 $(-\infty, -1)$, $[-1, 4)$, $[4, 5)$ 上分别对 a 进行考察:

① 当 $a < -1$ 时, 则由

$$P(\text{方程无实根}) = P(-1 < Y < 4) = \int_{-1}^4 \frac{1}{5-a} dy = \frac{5}{5-a} = 0.25$$

可得 $a = -15$;

② 当 $-1 \leq a < 4$ 时, 则由

$$\begin{aligned} P(\text{方程无实根}) &= P(-1 < Y < 4) = \int_{-1}^a 0 dy + \int_a^4 \frac{1}{5-a} dy \\ &= \frac{4-a}{5-a} = 0.25 \end{aligned}$$

可得 $a = \frac{11}{3}$;

③ 当 $4 \leq a < 5$ 时, 则与题设矛盾, 因为

$$P(\text{方程无实根}) = P(-1 < Y < 4) = 0 \neq 0.25.$$

综上所述, $a = -15$ 或 $a = \frac{11}{3}$.

【2-24】 某电脑显示器的使用寿命(单位: kh, 千小时) X 服从参数为 $\lambda = \frac{1}{50}$ 的指数分布. 生产厂家承诺: 购买者使用一年内显示器损坏将免费予以更换.

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

- (2) 求显示器至少可以使用 10 000 h 的概率;
 (3) 已知某台显示器已经使用了 10 000 h, 求其至少还能再用 10 000 小时的概率.

解 $X \sim E\left(\frac{1}{50}\right)$, 其分布函数

$$F(x) = \begin{cases} 1 - e^{-\frac{x}{50}} & (x > 0), \\ 0 & (x \leq 0). \end{cases}$$

$$(1) P(\text{免费更换}) = P(X < 2) = P(X \leq 2) = F(2)$$

$$= 1 - e^{-0.04} \approx 0.0392.$$

$$(2) P(X \geq 10) = 1 - P(X \leq 10) = 1 - F(10)$$

$$= e^{-0.2} \approx 0.8187.$$

(3) 由指数分布的“无记忆法”

$$P(X \geq 20 | X \geq 10) = P(X \geq 10) \approx 0.8187.$$

点评 本题的单位是 kh, 计算时必须将 2 000 h, 10 000 h 化作 2 与 10(kh). 密度函数与分布函数都可用来计算概率, 利用前者需计算积分, 利用后者仅需计算函数值, 故常利用后者即分布函数来计算概率.

【2-25】 设打一次电话所用的时间 X (单位: min, 分) 服从参数为 0.1 的指数分布. 若有一人刚好在你前面走进公共电话间, 试求你等待时间在 5 分钟到一刻钟之间的概率.

解 $X \sim E(0.1)$, 其分布函数

$$F(x) = \begin{cases} 1 - e^{-0.1x} & (x > 0), \\ 0 & (x \leq 0). \end{cases}$$

所求概率

$$P(5 \leq X \leq 15) = F(15) - F(5) = 1 - e^{-1.5} - (1 - e^{-0.5})$$

$$= e^{-0.5} - e^{-1.5} \approx 0.3834.$$

【2-26】 设 $X \sim N(5, 4)$, 确定 c 使 $P(X < c) = P(X > c)$.

解 方法 1 由 $P(X < c) = P(X > c) \Rightarrow F(c) = 1 - F(c) \Rightarrow F(c) =$

$$\Phi\left(\frac{c-5}{2}\right) = 0.5 \Rightarrow \frac{c-5}{2} = 0 \Rightarrow c = 5.$$

方法 2 利用正态分布的对称性. 因为 $X \sim N(5, 4)$, $P(X < c) =$

$P(X > c)$, 所以 $c = \mu = 5$.

【2-27】 某人需乘车到机场, 现有两条路线可供选择. 第一条路线较短, 但交通比较拥挤, 到达机场所需时间 X (单位为 min, 分) 服从正态分布 $N(50, 100)$. 第二条路线较长, 但出现意外阻塞较少, 所需时间 $X \sim N(60, 16)$.

(1) 若有 70 min 可用, 应走哪一条路线, 使及时赶到机场的概率较大?

(2) 若有 65 min 可用, 又应选哪条路线?

解 (1) 走第一条路线, 及时赶到机场的概率

$$\begin{aligned} P_1(0 < X \leq 70) &= \Phi\left(\frac{70-50}{10}\right) - \Phi\left(\frac{0-50}{10}\right) \\ &= \Phi(2) - \Phi(-5) \approx \Phi(2). \end{aligned}$$

走第二条路线, 及时赶到机场的概率

$$\begin{aligned} P_2(0 < X \leq 70) &= \Phi\left(\frac{70-60}{4}\right) - \Phi\left(\frac{0-60}{4}\right) \\ &= \Phi(2.5) - \Phi(-15) \approx \Phi(2.5). \end{aligned}$$

由于 $\Phi(2.5) > \Phi(2)$, 所以若有 70 min 可用, 应选择第二条路线.

(2) 相仿得

$$P_1(0 < X \leq 65) \approx \Phi(1.5) > \Phi(1.25) \approx P_2(0 < X \leq 65).$$

若有 65 min 可用, 应选择第一条路线.

【2-28】 设某机器生产的螺栓的长度 X (单位: cm) 服从正态分布 $N(10.05, 0.06^2)$. 规定长度在范围 10.05 ± 0.12 内为合格, 求螺栓不合格的概率.

解 螺栓合格的概率

$$\begin{aligned} P(10.05 - 0.12 < X < 10.05 + 0.12) &= \Phi(2) - \Phi(-2) \\ &= 2\Phi(2) - 1 = 0.9544, \end{aligned}$$

故螺栓不合格的概率

$$p = 1 - 0.9544 = 0.0456.$$

点评 上述两题都是正态分布的应用, 可以说正态分布是自然界中最常见的一种分布, 因而是概率论中最重要的一种分布. 理论上已证

明:如果某个数量指标呈现随机性是由很多随机因素影响的结果,而每个随机因素的影响又都不太大,则此数量指标服从正态分布.另外,正态分布具有良好的性质,理论上研究得比较透彻,许多其他分布常可用正态分布作为其近似分布,这就使得概率论基本上是围绕正态分布发展起来的.

【2-29】 设随机变量 $X \sim N(\mu, \sigma^2)$, 则随 σ 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.

解 因 $X \sim N(\mu, \sigma^2)$, 故

$$\begin{aligned} P(|X - \mu| < \sigma) &= P(\mu - \sigma < X < \mu + \sigma) \\ &= \Phi\left(\frac{\mu + \sigma - \mu}{\sigma}\right) - \Phi\left(\frac{\mu - \sigma - \mu}{\sigma}\right) \\ &= \Phi(1) - \Phi(-1) = 2\Phi(1) - 1, \end{aligned}$$

这个概率是个定值, 与 μ, σ 无关, 所以选 C.

【2-30】 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 3 次测量误差的绝对值大于 21.7 的概率.

解 因 $X \sim N(0, 10^2)$, 故每次测量误差的绝对值大于 21.7 的概率

$$\begin{aligned} p &= P(|X| > 21.7) = 1 - P(|X| \leq 21.7) \\ &= 1 - P(-21.7 \leq X \leq 21.7) \\ &= 1 - \left[\Phi\left(\frac{21.7}{10}\right) - \Phi\left(\frac{-21.7}{10}\right) \right] \\ &= 1 - [\Phi(2.17) + \Phi(2.17) - 1] \\ &= 2[1 - \Phi(2.17)] = 2(1 - 0.985) = 0.03. \end{aligned}$$

设在 100 次独立重复测量中, “测量误差的绝对值大于 21.7” 出现的次数为 Y , 则 $Y \sim B(100, 0.03)$, 由于 $n=100$ 较大, $p=0.03$ 较小, 故可用泊松分布近似, 取 $\lambda = np = 3$, 则 $Y \overset{\text{近似}}{\sim} \pi(3)$, 于是所求概率

$$P(Y \geq 3) = 1 - P(Y < 3) = 1 - \sum_{k=0}^2 P_{100}(k)$$

$$\begin{aligned} &\approx 1 - \sum_{k=0}^2 \frac{3^k}{k!} e^{-3} = 1 - e^{-3}(1 + 3 + 4.5) \\ &= 1 - 8.5e^{-3} \approx 0.5768. \end{aligned}$$

点评 本题是牵涉到两个随机变量、三个分布的综合题,其中二项分布的一个未知参数 p 需通过已知的正态分布求出.

【2-31】 设 X 是 $[0, 1]$ 上的连续型随机变量, 且 $P(X \leq 0.29) = 0.75$. 若 $Y = 1 - X$, 试确定常数 c , 使 $P(Y \leq c) = 0.25$.

解 由 $P(X \leq 0.29) = 0.75 \Rightarrow P(1 - Y \leq 0.29) = 0.75 \Rightarrow P(Y \geq 0.71) = 0.75 \Rightarrow 1 - P(Y \leq 0.71) = 0.75 \Rightarrow P(Y \leq 0.71) = 0.25$, 所以 $c = 0.71$, 使 $P(Y \leq c) = 0.25$.

【2-32】 对球的直径进行测量, 测量值在区间 $[a, b]$ 上服从均匀分布, 求球体积的概率密度函数.

分析 球的直径是个随机变量, 且其分布已知. 球的体积是个随机因变量, 它是直径的函数, 其分布可以通过已知分布求出.

解 设球直径为 X , 则 $X \sim U[a, b]$, 其密度函数

$$f_X(x) = \begin{cases} \frac{1}{b-a} & (a \leq x \leq b), \\ 0 & (\text{其他}). \end{cases}$$

设球体积为 V , 则 $V = \frac{1}{6} \pi X^3$, 即有严格单调增函数 $v = \frac{1}{6} \pi x^3$. 其

反函数 $x = \sqrt[3]{\frac{6v}{\pi}}$, $\frac{1}{6} \pi a^3 \leq v \leq \frac{1}{6} \pi b^3$. $x' = \sqrt[3]{\frac{2}{9\pi}} v^{-\frac{2}{3}}$, 由随机变量的函数的密度函数的公式得 V 的密度函数

$$f_V(v) = f_X(x) |x'| = \begin{cases} \frac{1}{b-a} \sqrt[3]{\frac{2}{9\pi}} v^{-\frac{2}{3}} & \left(\frac{1}{6} \pi a^3 \leq v \leq \frac{1}{6} \pi b^3 \right), \\ 0 & (\text{其他}). \end{cases}$$

点评 必须注意应用密度函数公式的条件是: 函数在定义域内严格单调, 其反函数有连续导数. 在使反函数无意义的点 y 处, 补充定义 $f_Y(y) = 0$. 若函数不严格单调, 则先将定义域区间划分为若干个子区间, 使函数在每个子区间上严格单调, 然后应用公式, 最后叠加各子区

间上求得的 $f_X(x_i) |x'_i| (i=1, 2, \dots, n)$, n 为子区间个数 (详见后面三例).

【2-33】 由统计物理学知, 分子运动速度的绝对值 X 服从麦克斯韦(Maxwell)分布, 其密度函数

$$f_X(x) = \begin{cases} \frac{2\sqrt{2}x^2}{\sigma^3\sqrt{\pi}} e^{-\frac{x^2}{\sigma^2}} & (x > 0), \\ 0 & (x \leq 0). \end{cases}$$

其中 $\sigma > 0$ 为常数. 求分子动能 $Y = \frac{1}{2}mX^2$ (m 为分子质量) 的概率密度函数.

解 $y = \frac{1}{2}mx^2$ 在 $x > 0$ 时严格单调增, 其反函数 $x = \sqrt{\frac{2y}{m}}$ ($y > 0$), $x' = \frac{1}{\sqrt{2m}}y^{-\frac{1}{2}}$, 于是动能 Y 的密度函数

$$f_Y(y) = f_X(x) |x'| = \begin{cases} \frac{4\sqrt{y}}{m^{\frac{3}{2}}\sigma^3\sqrt{\pi}} e^{-\frac{2y}{m\sigma^2}} & (y > 0), \\ 0 & (y \leq 0). \end{cases}$$

【2-34】 设 $X \sim N(0, 1)$, 求下列各随机变量的函数的密度函数:

(1) $Y = e^X$; (2) $Y = |X|$; (3) $Y = 2X^2 + 1$.

解 由题设可知 X 的密度函数

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

(1) $y = e^x$ 在定义域上严格单调增, 其反函数 $x = \ln y$ ($y > 0$), $x' =$

$\frac{1}{y}$, 则 Y 的密度函数

$$f_Y(y) = f_X(x) |x'| = \begin{cases} \frac{1}{y\sqrt{2\pi}} e^{-\frac{(\ln y)^2}{2}} & (y > 0), \\ 0 & (y \leq 0). \end{cases}$$

(2) 方法 1 $y = |x|$ 在 $(-\infty, 0]$ 上严格单调减, 其反函数 $x_1 = -y$, $x'_1 = -1$; $y = |x|$ 在 $(0, +\infty)$ 上严格单调增, 其反函数 $x_2 = y$,

$x'_2=1$. 于是 Y 的密度函数

$$f_Y(y) = f_X(x_1) |x'_1| + f_X(x_2) |x'_2| = \begin{cases} \sqrt{\frac{2}{\pi}} e^{-\frac{y^2}{2}} & (y \geq 0) \\ 0 & (y < 0). \end{cases}$$

方法 2 $y=|x| \geq 0$, 当 $y < 0$ 时, Y 的分布函数

$$F_Y(y) = P(Y \leq y) = P(|X| \leq y) = P(\emptyset) = 0 \\ \Rightarrow f_Y(y) = F'_Y(y) = 0;$$

当 $y \geq 0$, 记 X 的分布函数为 $F_X(x)$,

$$F_Y(y) = P(|X| \leq y) = P(-y \leq X \leq y) = F_X(y) - F_X(-y).$$

于是

$$f_Y(y) = F'_Y(y) = f_X(y) + f_X(-y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{(-y)^2}{2}} \\ = \sqrt{\frac{2}{\pi}} e^{-\frac{y^2}{2}},$$

故 Y 的密度函数

$$f_Y(y) = \begin{cases} \sqrt{\frac{2}{\pi}} e^{-\frac{y^2}{2}} & (y \geq 0), \\ 0 & (y < 0). \end{cases}$$

$$(3) \text{ 由 } y=2x^2+1 \text{ 得 } x_{1,2} = \pm \sqrt{\frac{y-1}{2}}, \quad x'_{1,2} = \pm \frac{1}{2\sqrt{2(y-1)}}$$

($y > 1$) 于是 Y 的密度函数

$$f_Y(y) = f_X(x_1) |x'_1| + f_X(x_2) |x'_2| \\ = \begin{cases} \frac{1}{2\sqrt{\pi(y-1)}} e^{-\frac{y-1}{4}} & (y > 1), \\ 0 & (y \leq 1). \end{cases}$$

点评 求连续随机变量函数的概率密度函数一般有两种方法:

- ① 公式法 如题(2)的方法 1;
- ② 分布函数法 如题(2)的方法 2.

【2-35】 设 $X \sim U\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, 求 $Y = \cos X$ 的密度函数.

解 由题设知 X 的密度函数

$$f_X(x) = \begin{cases} \frac{1}{\pi} & \left(-\frac{\pi}{2} \leq x \leq \frac{\pi}{2}\right), \\ 0 & (\text{其他}). \end{cases}$$

方法 1 (公式法)

$y = \cos x$ 在区间 $\left[-\frac{\pi}{2}, 0\right]$ 与 $\left(0, \frac{\pi}{2}\right]$ 上分别为严格单调增与严格单调减, 其反函数分别为 $x_1 = -\arccos y$ 与 $x_2 = \arccos y$, 对其求导得

$$x'_1 = \frac{1}{\sqrt{1-y^2}}, \quad x'_2 = -\frac{1}{\sqrt{1-y^2}},$$

于是 Y 的密度函数

$$\begin{aligned} f_Y(y) &= f_X(x_1) |x'_1| + f_X(x_2) |x'_2| \\ &= \frac{f_X(-\arccos y)}{\sqrt{1-y^2}} + \frac{f_X(\arccos y)}{\sqrt{1-y^2}} \\ &= \begin{cases} \frac{2}{\pi \sqrt{1-y^2}} & (0 < y < 1), \\ 0 & (\text{其他}). \end{cases} \end{aligned}$$

方法 2 (分布函数法)

当 $y \leq 0$ 或 $y \geq 1$ 时, $F_Y(y) = 0$, 从而 $f_Y(y) = 0$;

当 $0 < y < 1$ 时, 满足 $\cos X \leq y$ 的 X 在区间 $\left[-\frac{\pi}{2}, -\arccos y\right]$ 和 $\left[\arccos y, \frac{\pi}{2}\right]$ 上, 故

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(\cos X \leq y) \\ &= P\left(-\frac{\pi}{2} \leq X \leq -\arccos y\right) + P\left(\arccos y \leq X \leq \frac{\pi}{2}\right) \\ &= \int_{-\frac{\pi}{2}}^{-\arccos y} \frac{1}{\pi} dt + \int_{\arccos y}^{\frac{\pi}{2}} \frac{1}{\pi} dt, \end{aligned}$$

求导得 Y 的密度函数

$$f_Y(y) = F'_Y(y) = \begin{cases} \frac{1}{\pi} \left(\frac{1}{\sqrt{1-y^2}} + \frac{1}{\sqrt{1-y^2}} \right) & (0 < y < 1) \\ 0 & (\text{其他}) \end{cases}$$

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

【2-36】 设随机变量 X 的分布函数为 $F(x)$, 求 $Y = -\frac{X}{3} + 2$ 的分布函数. 下面解法是否正确? 为什么?

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P\left(-\frac{X}{3} + 2 \leq y\right) \\ &= P(X \geq 6 - 3y) = 1 - P(X < 6 - 3y) \\ &= 1 - F(6 - 3y). \end{aligned}$$

解 若 X 是连续型随机变量, 则结果正确. 但题中并未给出这一条件, 所以解法不正确. 错误在于最后一步等式不成立, 即一般

$$P(X < a) \neq F(a).$$

事实上, $P(X < a) = P(X \leq a) - P(X = a)$

$$= F(a) - [F(a) - F(a-0)] = F(a-0),$$

其中 $F(a-0) = \lim_{x \rightarrow a^-} F(x)$

本题的正确答案为

$$F_Y(y) = 1 - F(6 - 3y - 0).$$

【2-37】 设随机变量 X 的分布函数 $F(x)$ 是连续函数, 求 $Y = F(X)$ 的分布函数. 下面的解法是否正确? 为什么?

(1) 当 $y < 0$ 时, $F_Y(y) = P(Y \leq y) = P(F(X) \leq y) = P(\emptyset) = 0$;

(2) 当 $y \geq 1$ 时, $F_Y(y) = P(Y \leq y) = P(F(X) \leq y) = P(\Omega) = 1$;

(3) 当 $0 \leq y < 1$ 时, $F_Y(y) = P(Y \leq y) = P(F(X) \leq y)$
 $= P(X \leq F^{-1}(y)) = F(F^{-1}(y)) = y,$

故 $Y = F(X)$ 的分布函数

$$F_Y(y) = \begin{cases} 0 & (y < 0), \\ y & (0 \leq y < 1), \\ 1 & (y \geq 1). \end{cases}$$

解 不正确. 错在第三步的第三个等号处, 因为事件 $(F(X) \leq y)$ 与 $(X \leq F^{-1}(y))$ 未必等价. 尽管题中已设 $F(x)$ 是连续函数, 但它未必

严格单调,因而其反函数也未必存在.

正确的求解方法是利用分布函数单调不减这一基本性质.

当 $0 \leq y < 1$ 时

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(F(X) \leq F(x)) = P(X \leq x) \\ &= F(x) = y. \end{aligned}$$

B 习题精选

【2-38】 求随机变量 X 的分布律中的未知参数值:

(1) $P(X=k) = a\left(\frac{2}{3}\right)^k \quad (k=0,1,2)$, 则 $a = \underline{\hspace{2cm}}$.

(2) $P(X=k) = b\left(\frac{1}{2}\right)^k \quad (k=1,2,\dots)$, 则 $b = \underline{\hspace{2cm}}$.

【2-39】 设随机变量 $X \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0.2+a & 0.1 & 0.3+b & c \end{pmatrix}$, 则分布律中未知的三个参数 a, b, c 应满足 .

【2-40】 设 $G(x)$ 与 $H(x)$ 分别为随机变量 Y 与 Z 的分布函数, 为使 $F(x) = aG(x) - bH(x)$ 也能成为某一随机变量的分布函数, 在下列各组数值中应取 .

(A) $a = \frac{1}{2}, b = -\frac{3}{2}$;

(B) $a = -\frac{1}{2}, b = \frac{3}{2}$;

(C) $a = \frac{3}{5}, b = -\frac{7}{5}$;

(D) $a = \frac{3}{5}, b = -\frac{2}{5}$.

【2-41】 设 a, b 为任意实数, $a < b$. 已知 $P(X \leq b) \geq 1 - \beta, P(X > a) \geq 1 - \alpha$, 则必有 $P(a < X \leq b)$.

(A) $\geq 1 - (\alpha + \beta)$;

(B) $\geq \alpha + \beta$;

(C) $\leq 1 - (\alpha + \beta)$;

(D) $\leq \alpha + \beta$.

【2-42】 设随机变量 X 的分布函数

$$F(x) = \begin{cases} 0 & (x < 0), \\ x + 0.3 & (0 \leq x < 0.6), \\ 1 & (x \geq 0.6), \end{cases}$$

则 $P(X=0.6)=$ _____.

- (A) 0; (B) 0.1; (C) 0.9; (D) 1.

【2-43】 设随机变量 X 的绝对值不大于 1, 且 $P(X=-1)=\frac{1}{3}$,

$P(X=1)=\frac{1}{6}$, 则 $P(-1<X<1)=$ _____.

【2-44】 游船上有供水龙头 20 个, 每一龙头被打开的可能性为 0.1, 记 X 为同时被打开的水龙头个数, 则 $P(X\geq 2)=$ _____.

【2-45】 设随机变量 $X\sim B(3, p)$, $Y\sim B(4, p)$. 若 $P(X\geq 1)=\frac{7}{8}$, 则 $P(Y\geq 1)=$ _____.

【2-46】 尽管在几何教科书中已指出: 用圆规和直尺三等分一个任意角是不可能的, 但每年总有一些“发明者”撰写关于用圆规和直尺将角三等分. 某地区每年撰写此类文章的篇数 X 服从参数为 2 的泊松分布, 求明年仍有此类文章的概率.

【2-47】 某保险公司有 2 500 名同一年龄段的人参加人寿保险, 统计资料表明该年龄段的人一年中死亡的概率为 2‰. 每个参保人在年初须交保险费 12 元, 而死亡时家属可获得 2 000 元赔偿金, 求:

- (1) 保险公司亏本的概率;
- (2) 保险公司在此项目上年获利不少于 20 000 元的概率;
- (3) 保险公司一年中最可能付出的赔偿金额.

【2-48】 某宾馆有 250 间客房, 每间按有一部电话分机, 分机在任意时刻要外线的概率为 2%, 问服务总台至少需安装多少条外线, 才能以不低于 90% 的概率保证分机能及时接通外线?

【2-49】* 设在一段时间内进入商场的人数服从参数为 λ 的泊松分布, 每个顾客购买某家用电器的概率为 p , 而每个顾客是否购买该电器相互独立. 证明: 进入商场的顾客购买这种电器的人数服从参数为 λp 的泊松分布.

【2-50】 设连续型随机变量 X 的密度函数与分布函数分别为 $f(x)$ 与 $F(x)$, 则有_____.

$$(A) F(x) = \int_0^x f(t) dt;$$

$$(B) F(x) = \int_{-\infty}^{+\infty} f(x) dx;$$

$$(C) F(x) = 1 - \int_x^{+\infty} f(t) dt;$$

$$(D) F(x) = 1 + \int_x^{+\infty} f(t) dt.$$

【2-51】 若随机变量 X 的可能值充满区间_____, 而在此区间外取值的概率为零, 则 $f(x) = \cos x$ 可成为 X 的密度函数.

$$(A) [0, \pi];$$

$$(B) \left[\frac{\pi}{2}, \pi\right];$$

$$(C) \left[-\frac{\pi}{2}, \frac{\pi}{2}\right];$$

$$(D) \left[-\frac{\pi}{2}, 0\right].$$

【2-52】 设 X 的密度函数

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

则 $P(X \leq 0.3) = \underline{\hspace{2cm}}$, $P(0.3 < X < 0.6) = \underline{\hspace{2cm}}$, $P(X = 0.6) = \underline{\hspace{2cm}}$, $P(X < 0.6 | X > 0.3) = \underline{\hspace{2cm}}$.

【2-53】 设连续型随机变量 X 的分布函数

$$F(x) = \begin{cases} A + Be^{-\lambda x} & (x > 0), \\ 0 & (x \leq 0), \end{cases}$$

则 $A = \underline{\hspace{2cm}}$, $B = \underline{\hspace{2cm}}$, 其密度函数 $f(x) = \underline{\hspace{2cm}}$.

【2-54】 设连续型随机变量 X 的密度函数

$$f(x) = \begin{cases} A \sqrt{1-x^2} & (-1 \leq x \leq 1), \\ 0 & (\text{其他}), \end{cases}$$

则 $A = \underline{\hspace{2cm}}$.

$$(A) \frac{1}{\pi};$$

$$(B) \frac{2}{\pi};$$

$$(C) \frac{1}{2\pi};$$

$$(D) \frac{1}{4\pi}.$$

【2-55】 设随机变量 $Y \sim U[0, 8]$, 则方程 $x^2 + Yx + 1 = 0$ 有实根的概率为_____.

$$(A) \frac{1}{8};$$

$$(B) \frac{1}{4};$$

$$(C) \frac{1}{2};$$

$$(D) \frac{3}{4}.$$

【2-56】 设随机变量 $X \sim N(\mu, 9)$, $Y \sim N(\mu, 16)$, 令

$$p_1 = P(X \leq \mu - 3), \quad p_2 = P(Y \geq \mu + 4),$$

则_____成立.

- (A) 对任何实数 μ , 都有 $p_1 = p_2$;
- (B) 对任何实数 μ , 都有 $p_1 \neq p_2$;
- (C) 仅对部分 μ 值, 才有 $p_1 = p_2$;
- (D) 不能确定 p_1, p_2 大小.

【2-57】 设随机变量 X 服从参数 $\lambda = \frac{1}{8}$ 的指数分布, 则 $P(2 < X < 8) =$ _____.

- (A) $\frac{2}{8} \int_2^8 e^{-\frac{x}{8}} dx$;
- (B) $\int_2^8 e^{-\frac{x}{8}} dx$;
- (C) $\frac{1}{8} (e^{-\frac{1}{4}} - e^{-1})$;
- (D) $e^{-\frac{1}{4}} - e^{-1}$.

【2-58】 设 $X \sim N(1, 2)$, $f(x), F(x)$ 分别为 X 的密度函数和分布函数, 则下面结论中错误的是_____.

- (A) $f(x)$ 的图形关于直线 $x=1$ 对称;
- (B) $f(x)$ 的最大值为 $\frac{1}{2\sqrt{\pi}}$;
- (C) $F(0) = \frac{1}{2}$;
- (D) $F(x)$ 表示 $f(x)$ 在 $(-\infty, x)$ 上的积分.

【2-59】 设 $X \sim N(\mu, 1)$, 分布函数为 $F(x)$, 则对任意实数 x , 有_____.

- (A) $F(\mu+x) + F(\mu-x) = 1$;
- (B) $F(\mu+x) = F(\mu-x)$;
- (C) $F(x+\mu) + F(x-\mu) = 1$;
- (D) $F(x+\mu) = F(x-\mu)$.

【2-60】 假设一电路装有 2 个同类型的电子元件, 其工作状态相互独立, 且无故障工作时间都服从参数为 5 的指数分布. 当 2 个元件都无故障时, 电路正常工作, 否则不能正常工作, 求电路正常工作时间 T 的概率分布函数.

【2-61】 设某河流每年的最高洪水水位 X (单位: m) 具有密度

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

今要修建河堤能防御百年一遇的洪水,河堤至少要修多高(计算河堤的起点与水位的起点相同)?

【2-62】 某厂生产的电子管寿命 $X \sim N(16, \sigma^2)$ (单位: kh). 若要求 $P(12 \leq X \leq 20) \geq 0.95$, 问允许 σ 最大为多少?

【2-63】 抽样表明: 某市新生儿体重 $X \sim N(3.4, \sigma^2)$ (单位: kg). 已知该市新生儿体重不足 2 kg 的占 3.1%, 试求该市新生儿体重超过 4 kg 的百分比.

【2-64】 设随机变量 X 的密度函数

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

则 X^2 服从_____.

- (A) 参数为 1 的指数分布; (B) 参数为 2 的指数分布;
(C) 参数为 0,1 的均匀分布; (D) 参数为 0,2 的均匀分布.

【2-65】 设随机变量 $X \sim E(2)$, 则 $Y = 1 - e^{-2X}$ _____.

- (A) 服从 $(0, 1)$ 上的均匀分布;
(B) 仍服从指数分布;
(C) 服从正态分布;
(D) 服从参数为 2 的泊松分布.

【2-66】 设随机变量 $X \sim E(4)$, 则 $Y = \frac{1}{3}X^3 + 2$ 的密度函数 $f_Y(y) =$ _____.

【2-67】 设 $X \sim N(-1, 1)$, $Y = e^X$, 则 $P(Y > 1) =$ _____.

【2-68】 设 X 的分布律为 $\begin{bmatrix} 0 & \frac{\pi}{2} & \pi \\ 0.1 & 0.6 & 0.3 \end{bmatrix}$, 求:

- (1) $Y = \frac{1}{2}X + 1$ 的分布律; (2) $Y = \cos^2 X$ 的分布律.

【2-69】 对一圆片直径进行测量,测量值 X 服从均匀分布 $U(5, 6)$. 求圆面积 Y 的密度函数.

【2-70】 设随机变量 X 服从柯西(Cauchy)分布,其分布函数

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

求随机变量 $Y=1-\sqrt[3]{X}$ 的密度函数.

【2-71】 设星球 A 至最近的星球 B 的距离 X 的分布函数

$$F(x) = \begin{cases} 1 - e^{-\frac{4}{3}\pi k x^3} & (x \geq 0), \\ 0 & (x < 0). \end{cases}$$

求 B 对 A 的引力 $Y=\frac{k}{X^2}$ ($k>0$ 为常数) 的密度函数.

答案与提示

【2-38】 (1) $a=\frac{9}{19}$; (2) $b=1$.

【2-39】 $a+b+c=0.4$ 且 $a>-0.2, b>-0.4, c>0$.

【2-40】 选 D. 从 $F(+\infty)=1$ 着手.

【2-41】 选 A. 由 $P(a < X \leq b) = P(X \leq b) - P(X \leq a)$ 推导出结果.

【2-42】 选 B. $P(X=0.6) = F(0.6) - F(0.6-0)$.

【2-43】 $\frac{1}{2}$. $P(-1 < X < 1) = P(-1 \leq X \leq 1) - P(X = -1) - P(X = 1)$.

【2-44】 0.6082. $X \sim B(20, 0.1)$.

【2-45】 $\frac{15}{16}$. 先通过 $P(X \geq 1) = \frac{7}{8}$ 求出 $p = \frac{1}{2}$.

【2-46】 0.8709. $X \sim \pi(2), P(X \geq 1) = 1 - P(X < 1)$.

【2-47】 (1) 0.000069; (2) 0.616.

$X \sim B(2500, 0.002) \Rightarrow X \overset{\text{近似}}{\sim} \pi(5)$

(3) 10000 元. $k = \lceil (n+1)p \rceil = 5$.

【2-48】 至少安装 8 条外线

$$X \sim B(250, 0.02) \Rightarrow X \overset{\text{近似}}{\sim} \pi(5)$$

【2-49】* 设 X ——进入商场人数, 则 $X \sim \pi(\lambda)$, 即

$$P(X = m) = \frac{\lambda^m}{m!} e^{-\lambda} \quad (m = 0, 1, 2, \dots).$$

设 Y ——购买电器人数, 由题设知 Y 的条件分布为 $B(m, p)$, 即

$$P(Y = k | X = m) = C_m^k p^k (1-p)^{m-k} \quad (k = 0, 1, 2, \dots, m).$$

由全概率公式, 得

$$\begin{aligned} P(Y = k) &= \sum_{m=0}^{\infty} P(X = m) P(Y = k | X = m) \\ &= \sum_{m=0}^{\infty} \frac{\lambda^m e^{-\lambda}}{m!} C_m^k p^k (1-p)^{m-k} \\ &= \sum_{m=0}^{\infty} \frac{m!}{k! (m-k)!} \frac{1}{m!} \lambda^m e^{-\lambda} p^k (1-p)^{m-k} \\ &= \frac{(\lambda p)^k}{k!} e^{-\lambda} \sum_{m=k}^{\infty} \frac{\lambda^{m-k} (1-p)^{m-k}}{(m-k)!} \quad (\text{令 } n = m - k) \\ &= \frac{(\lambda p)^k}{k!} e^{-\lambda} \sum_{n=0}^{\infty} \frac{[\lambda(1-p)]^n}{n!} = \frac{(\lambda p)^k}{k!} e^{-\lambda} e^{\lambda(1-p)} \\ &= \frac{(\lambda p)^k}{k!} e^{-\lambda p} \quad (k = 0, 1, 2, \dots). \end{aligned}$$

所以 $Y \sim \pi(\lambda p)$.

$$\text{【2-50】 选 C. } F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^{+\infty} f(t) dt + \int_{+\infty}^x f(t) dt.$$

$$\text{【2-51】 选 D. } \cos x \text{ 在 } \left[-\frac{\pi}{2}, 0\right] \text{ 上非负, 且积分为 1.}$$

$$\text{【2-52】 } 0.027, 0.189, 0, \frac{189}{973} \approx 0.194.$$

$$\text{【2-53】 } 1, -1, f(x) = \begin{cases} \lambda e^{-\lambda x} & (x > 0), \\ 0 & (x \leq 0). \end{cases}$$

$$\text{【2-54】 选 B. } \int_{-1}^1 A \sqrt{1-x^2} dx = 1 \text{ 作三角代换 } x = \sin t.$$

【2-55】 选 D. 由判别式得 $Y < -2$ (舍去), $Y > 2$.

【2-56】 选 A. $p_1 = \Phi(-1)$, $p_2 = 1 - \Phi(1)$.

【2-57】 选 D. $P(2 < X < 8) = \int_2^8 \frac{1}{8} e^{-\frac{x}{8}} dx$.

【2-58】 选 C. $F(0) = \Phi\left(\frac{0-1}{\sqrt{2}}\right) \neq \Phi(0) = \frac{1}{2}$.

【2-59】 选 A. $F(\mu+x) + F(\mu-x) = \Phi(x) + \Phi(-x)$.

【2-60】 $G(t) = \begin{cases} 1 - e^{-10t} & (t > 0), \\ 0 & (t \leq 0). \end{cases}$

以 $X_i (i=1, 2)$ 表示第 i 个电子元件无故障工作的时间, 则 $X_i \sim E(5)$, 其分布函数为 $F(x)$. 设 T 的分布函数为 $G(t)$, 可推得 $G(t) = 1 - [1 - F(t)]^2$.

【2-61】 河堤高度 $h > 10(\text{m})$. 由 $P(X > h) < 0.01$ 推出.

【2-62】 2.041. 利用标准正态分布表.

【2-63】 21.19%. 先由 $P(X < 2) = \Phi\left(\frac{2-3.4}{\sigma}\right) = 0.031$ 求出 σ .

【2-64】 选 C. 用公式法求出 X^2 的密度函数.

【2-65】 选 A. 用公式法求出 Y 的密度函数.

【2-66】 $f_Y(y) = \begin{cases} \frac{4e^{-4(3y-6)}^{\frac{1}{3}}}{(3y-6)^{\frac{2}{3}}} & (y > 2), \\ 0 & (y \leq 2). \end{cases}$

【2-67】 0.1587. $P(Y > 1) = \int_1^{+\infty} f_Y(y) dy = \cdots = 1 - \Phi(1)$.

【2-68】 (1) $\begin{pmatrix} 1 & \frac{\pi}{4} + 1 & \frac{\pi}{2} + 1 \\ 0.1 & 0.6 & 0.3 \end{pmatrix}$; (2) $\begin{pmatrix} 0 & 1 \\ 0.6 & 0.4 \end{pmatrix}$.

【2-69】 $f_Y(y) = \begin{cases} \frac{1}{\sqrt{\pi y}} & \left(\frac{25}{4}\pi < y < 9\pi\right), \\ 0 & (\text{其他}). \end{cases}$

【2-70】 $f_Y(y) = \frac{3(1-y)^2}{\pi[1+(1-y)^6]} \quad (-\infty < y < +\infty)$.

$$\textbf{【2-71】} \quad f_Y(y) = \begin{cases} 2\pi\lambda k^{\frac{3}{2}} y^{-\frac{5}{2}} e^{-\frac{4}{3}\pi\lambda(\frac{k}{y})^{\frac{3}{2}}} & (y > 0), \\ 0 & (y \leq 0). \end{cases}$$

关键词

二维随机变量 (X,Y) (X,Y) 的联合分布函数 (X,Y) 的联合分布律 (X,Y) 的联合密度函数 (X,Y) 的边缘分布 条件分布 二维均匀分布 二维正态分布 二个随机变量函数的分布

教学要求

- (1) 了解二维(多维)随机变量的概念.
- (2) 了解二维随机变量的分布函数的概念和性质.
- (3) 了解二维离散型随机变量的分布律及其性质;了解二维连续型随机变量的密度函数及其性质.
- (4) 掌握二维随机变量的联合分布、边缘分布和条件分布的计算.
- (5) 理解随机变量的独立性概念,会利用独立性进行概率计算.
- (6) 了解二维均匀分布和二维正态分布,并会作有关的概率计算.
- (7) 会求两个独立的随机变量的简单函数的分布.

重点

二维随机变量的概念;联合分布函数、联合分布律和联合密度函数的概念及性质;边缘分布;随机变量独立性;两个随机变量的函数的分布,特别是 $Z=X+Y$, $M=\max(X,Y)$, $N=\min(X,Y)$ 的分布.

难点

已知联合分布律或联合密度函数,求联合分布函数(一般不作此要求);

已知 $(X,Y) \sim f(x,y)$, 计算 $P((X,Y) \in D)$;

已知 (X,Y) 的分布,求 $Z=X+Y$ 的分布.

方法提要

(1) 当 $F(x,y)$ 含有若干待定常数时,常利用

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

来确定这些待定常数;

当联合分布律或联合密度函数含有待定常数时,常利用

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P(X=x_i, Y=y_j) = 1 \text{ 或 } \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$$

来确定该常数.

(2) 求离散型随机变量 (X,Y) 的分布律,若 X,Y 分别有 n 个和 m 个取值,则 (X,Y) 共有 $n \times m$ 个取值,其对应于各取值的事件中的联合概率 p_{ij} 一般可利用乘法公式求出,有时也可由古典概型直接求出,最后

不忘验证 $\sum_{i=1}^n \sum_{j=1}^m p_{ij} = 1$.

$$(3) \text{ 若 } (X,Y) \sim f(x,y) = \begin{cases} g(x,y) & ((x,y) \in G), \\ 0 & ((x,y) \notin G), \end{cases}$$

$$\text{则 } P((X,Y) \in D) = \iint_D f(x,y) dx dy = \iint_{D \cap G} g(x,y) dx dy,$$

$D \cap G$ 表示平面上区域 D 与区域 G 的公共部分.

(4) 若 $(X,Y) \sim F(x,y)$,则

$$X \text{ 与 } Y \text{ 相互独立} \Leftrightarrow F(x,y) = F_X(x)F_Y(y).$$

离散情形: X 与 Y 相互独立 $\Leftrightarrow p_{ij} = p_{i \cdot} p_{\cdot j} \quad (\forall i, j)$

$\Leftrightarrow$ 联合概率矩阵秩为1.

连续情形: X 与 Y 相互独立 $\Leftrightarrow f(x,y) = f_X(x)f_Y(y)$.

(5) 若 $(X,Y) \sim f(x,y)$,则 $Z=aX+bY(a \neq 0, b \neq 0)$ 的密度函数

$$f_Z(z) = \frac{1}{|b|} \int_{-\infty}^{+\infty} f\left(x, \frac{z-ax}{b}\right) dx = \frac{1}{|a|} \int_{-\infty}^{+\infty} f\left(\frac{z-by}{a}, y\right) dy;$$

若 X 与 Y 相互独立, $X \sim f_X(x), Y \sim f_Y(y)$,则 $Z=aX+bY(a \neq 0, b \neq 0)$ 的密度函数

$$f_Z(z) = \frac{1}{|b|} \int_{-\infty}^{+\infty} f_X(x) f_Y\left(\frac{z-ax}{b}\right) dx = \frac{1}{|a|} \int_{-\infty}^{+\infty} f_X\left(\frac{z-by}{a}\right) f_Y(y) dy.$$

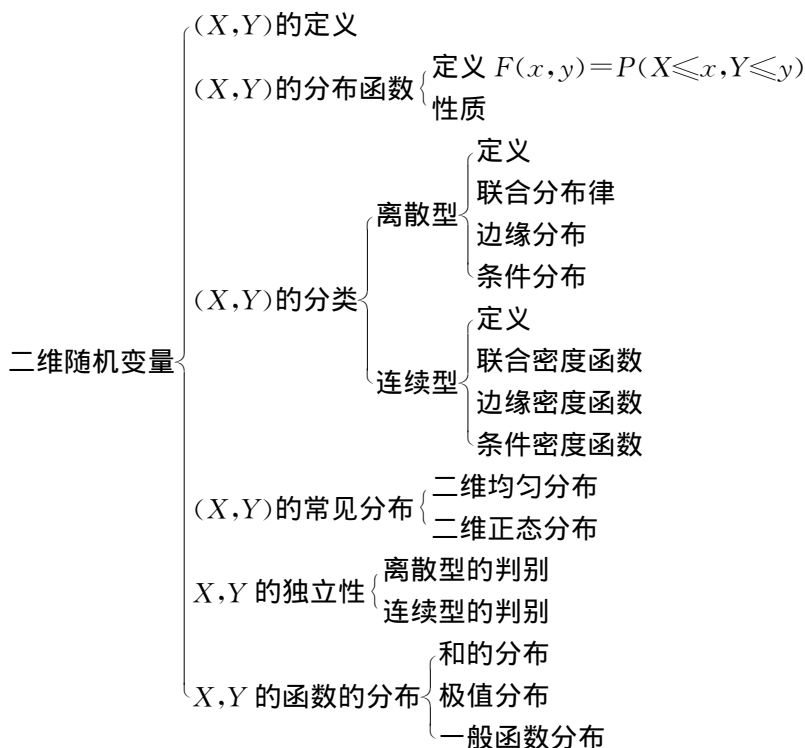
计算中会遇到两个难点:怎样找准 z 轴上的分界点;积分限如何演变. 采用不等式组定限法或作图法均可化解这两个难点.

一般,若 $(X,Y) \sim f(x,y)$, $Z=g(X,Y)$ 的分布函数和密度函数

$$F_Z(z) = P(Z \leq z) = P(g(X,Y) \leq z) = \iint_{g(x,y) \leq z} f(x,y) dx dy,$$

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \frac{d}{dz} \iint_{g(x,y) \leq z} f(x,y) dx dy.$$

本章知识结构



A 例题精解

【3-1】 某电子显示屏第一排有 2 个灯管,第二排有 3 个灯管. 令 X, Y 分别表示在某一时段内第一排与第二排烧坏的灯管数, (X, Y) 的联合分布律如下:

$X \backslash Y$	0	1	2	3
0	0.10	0.13	0.09	0.04
1	0.14	0.11	0.07	0.03
2	0.12	0.10	0.05	0.02

计算在该时段内下列事件的概率:

- (1) 第一排没有灯管烧坏;
- (2) 第一排与第二排烧坏的灯管数相等;
- (3) 第一排烧坏的灯管数不超过第二排烧坏的灯管数.

解 (1) 由联合分布律得 X 的边缘分布律

$$\begin{pmatrix} 0 & 1 & 2 \\ 0.36 & 0.35 & 0.29 \end{pmatrix},$$

于是第一排没有灯管烧坏的概率

$$P(X=0) = 0.36.$$

$$\begin{aligned} (2) \quad P(\text{两排烧坏的灯管数相等}) &= \sum_{i=0}^2 P(X=i, Y=i) = \sum_{i=0}^2 p_{ii} \\ &= P_{00} + P_{11} + P_{22} = 0.10 + 0.11 + 0.05 = 0.26. \end{aligned}$$

(3) 用逆事件计算, 所求概率

$$p = 1 - p_{10} - p_{20} - p_{21} = 1 - 0.14 - 0.12 - 0.10 = 0.64.$$

【3-2】 某校新选出的学生会 6 名女委员中文、理、工科各占 $\frac{1}{6}$,

$\frac{1}{3}, \frac{1}{2}$. 现从中随机指定 2 人为学生会主席候选人. 令 X, Y 分别为候选

人来自文、理科的人数. 求:

(1) (X, Y) 的联合分布律;

(2) (X, Y) 关于 X, Y 的边缘分布律.

解 X 与 Y 的可能取值分别为 $0, 1$ 与 $0, 1, 2$.

(1) 由古典概型可求得联合概率

$$p_{00} = P(X = 0, Y = 0) = \frac{C_3^2}{C_6^2} = \frac{3}{15},$$

$$p_{01} = P(X = 0, Y = 1) = \frac{C_2^1 C_3^1}{C_6^2} = \frac{6}{15},$$

$$p_{02} = P(X = 0, Y = 2) = \frac{C_2^2}{C_6^2} = \frac{1}{15},$$

$$p_{10} = P(X = 1, Y = 0) = \frac{C_1^1 C_3^1}{C_6^2} = \frac{3}{15},$$

$$p_{11} = P(X = 1, Y = 1) = \frac{C_1^1 C_2^1}{C_6^2} = \frac{2}{15},$$

$$p_{12} = P(X = 1, Y = 2) = 0.$$

于是得 (X, Y) 的联合分布律:

$\begin{matrix} Y \\ X \end{matrix}$	0	1	2
0	$\frac{3}{15}$	$\frac{6}{15}$	$\frac{1}{15}$
1	$\frac{3}{15}$	$\frac{2}{15}$	0

(2) 由联合分布律得边缘分布律

$$X \sim \begin{pmatrix} 0 & 1 \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}, \quad Y \sim \begin{pmatrix} 0 & 1 & 2 \\ \frac{6}{15} & \frac{8}{15} & \frac{1}{15} \end{pmatrix}.$$

点评 求联合概率一般有两种方法:

① 古典概型(仅适用于特殊场合);

② 乘法公式(适用于一般场合)

$$p_{ij} = P(X = i, Y = j) = P(X = i)P(Y = j | X = i).$$

本题用方法①较为简便. 若用②计算就稍嫌麻烦:

$$\begin{aligned} p_{00} &= P(X = 0, Y = 0) = P(X = 0)P(Y = 0 | X = 0) \\ &= \frac{C_5^2 C_3^2}{C_8^2 C_5^2} = \frac{3}{5}. \end{aligned}$$

【3-3】 盒里 10 个产品中有 2 个次品. 每次从盒里任取一个产品, 共取 2 次. 定义随机变量 X, Y 如下:

$$X = \begin{cases} 1 & (\text{第一次取出次品}), \\ 0 & (\text{第一次取出正品}), \end{cases} \quad Y = \begin{cases} 1 & (\text{第二次取出次品}), \\ 0 & (\text{第二次取出正品}). \end{cases}$$

分别就以下两种情况求出 (X, Y) 的联合分布律: (1) 放回抽样; (2) 不放回抽样.

分析 本题放回抽样与不放回抽样的区别在于: 放回抽样时, X 与 Y 是相互独立的, 而不放回抽样时, X 与 Y 是不独立的.

解 (1) 放回抽样:

$$P(X = 0, Y = 0) = P(X = 0)P(Y = 0) = \frac{8}{10} \times \frac{8}{10} = 0.64,$$

$$P(X = 1, Y = 0) = P(X = 1)P(Y = 0) = \frac{2}{10} \times \frac{8}{10} = 0.16,$$

$$P(X = 0, Y = 1) = P(X = 0)P(Y = 1) = \frac{8}{10} \times \frac{2}{10} = 0.16,$$

$$P(X = 1, Y = 1) = P(X = 1)P(Y = 1) = \frac{2}{10} \times \frac{2}{10} = 0.04,$$

故 (X, Y) 的联合分布律

X \ Y		
	0	1
0	0.64	0.16
1	0.16	0.04

(2) 不放回抽样:

$$P(X = 0, Y = 0) = \frac{C_8^2}{C_{10}^2} = \frac{28}{45},$$

$$P(X=0, Y=1) = \frac{C_8^1 C_2^1}{C_{10}^2} = \frac{8}{45},$$

$$P(X=1, Y=0) = \frac{C_2^1 C_8^1}{C_{10}^2} = \frac{8}{45},$$

$$P(X=1, Y=1) = \frac{C_2^2}{C_{10}^2} = \frac{1}{45},$$

故 (X, Y) 的联合分布律为

X \ Y	0	1
0	$\frac{28}{45}$	$\frac{8}{45}$
1	$\frac{8}{45}$	$\frac{1}{45}$

点评 由本题看到, 抽样方式的不同会引起联合分布律的不同, 但未必引起边缘分布律的不同. 不管抽样方式如何, 边缘分布律都为 $B(1, 0.2)$, 即

$$X \sim \begin{pmatrix} 0 & 1 \\ 0.8 & 0.2 \end{pmatrix}, \quad Y \sim \begin{pmatrix} 0 & 1 \\ 0.8 & 0.2 \end{pmatrix}.$$

【3-4】 考虑伯努利独立试验序列. 设每次试验成功的概率为 $p(0 < p < 1)$. 令 X 表示第一次成功之前的失败次数, Y 表示前两次成功之间的失败次数. 求 (X, Y) 的联合分布律及边缘分布律.

解 事件 $(X=i, Y=j)$ 表示试验共进行了 $i+j+2$ 次, 其中第 $i+1$ 次和第 $i+j+2$ 次是成功的, 其余 $i+j$ 次试验都是失败的. 由试验的独立性可得 (X, Y) 的联合分布律

$$\begin{aligned} P(X=i, Y=j) &= (1-p)^i p (1-p)^j p \\ &= p^2 (1-p)^{i+j} \quad (i, j = 0, 1, 2, \dots). \end{aligned}$$

X 的边缘分布律

$$P(X=i) = p(1-p)^i \quad (i = 0, 1, 2, \dots).$$

Y 的边缘分布律

$$P(Y=j) = p(1-p)^j \quad (j = 0, 1, 2, \dots).$$

点评 分布律的表示法一般有三种:

- ① 表格形式(适用于随机变量取有限个值的情形);
- ② 解析形式(经常采用的形式);
- ③ 图像形式(最少采用的形式).

【3-5】 X 表示随机地在 $1\sim 4$ 的 4 个整数中取出的一个整数, Y 表示在 $1\sim X$ 中随机地取出的一个整数. 求 (X, Y) 的联合分布.

解 方法 1 (表格形式)

显然 $P(X=k, Y>k)=0$,

$$P(X=1, Y=1)=P(X=1)P(Y=1|X=1)=\frac{1}{4}\times 1=\frac{1}{4},$$

$$P(X=2, Y=j)=P(X=2)P(Y=j|X=2)=\frac{1}{4}\times \frac{1}{2}=\frac{1}{8}$$

$$(j=1, 2),$$

$$P(X=3, Y=l)=P(X=3)P(Y=l|X=3)=\frac{1}{4}\times \frac{1}{3}=\frac{1}{12}$$

$$(l=1, 2, 3),$$

$$P(X=4, Y=m)=P(X=4)P(Y=m|X=4)=\frac{1}{4}\times \frac{1}{4}=\frac{1}{16}$$

$$(m=1, 2, 3, 4),$$

于是 (X, Y) 的联合分布律为

$X \backslash Y$	1	2	3	4
1	$\frac{1}{4}$	0	0	0
2	$\frac{1}{8}$	$\frac{1}{8}$	0	0
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	0
4	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$

方法 2 (解析形式)

$$P_{ij} = P(X = i, Y = j) = \begin{cases} \frac{1}{4i} & (j \leq i), \\ 0 & (j > i) \end{cases} \quad (i = 1, 2, 3, 4).$$

点评 两种方法各有优点. 表格形式, 直观明了; 解析形式, 结构严谨, 便于推理应用.

【3-6】 设函数 $g(x)$ 满足 $g(x) \geq 0$, $\int_0^{+\infty} g(x) dx = 1$. 问二元函数

$$f(x, y) = \begin{cases} \frac{g(\sqrt{x^2 + y^2})}{\pi \sqrt{x^2 + y^2}} & (x > 0, y > 0), \\ 0 & (\text{其他}), \end{cases}$$

能否成为某连续型随机变量 (X, Y) 的联合密度函数?

解 当 $x > 0, y > 0$ 时, $\sqrt{x^2 + y^2} > 0, g(\sqrt{x^2 + y^2}) \geq 0$, 从而 $f(x, y) \geq 0$, 即 $f(x, y)$ 满足非负性. 再看 $f(x, y)$ 是否满足归一性:

$$\begin{aligned} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy &= \int_0^{+\infty} \int_0^{+\infty} \frac{g(\sqrt{x^2 + y^2})}{\pi \sqrt{x^2 + y^2}} dx dy \\ &\stackrel{\substack{x = r \cos \theta \\ y = r \sin \theta}}{=} \int_0^{\frac{\pi}{2}} d\theta \int_0^{+\infty} \frac{g(r)}{\pi r} r dr \\ &= \frac{1}{2} \int_0^{+\infty} g(r) dr = \frac{1}{2} \neq 1, \end{aligned}$$

归一性不满足, 所以 $f(x, y)$ 不能成为 (X, Y) 的联合密度函数.

点评 显然二元函数

$$f(x, y) = \begin{cases} \frac{2g(\sqrt{x^2 + y^2})}{\pi \sqrt{x^2 + y^2}} & (x > 0, y > 0), \\ 0 & (\text{其他}) \end{cases}$$

可作为某连续型随机变量 (X, Y) 的联合密度函数, 其中 $g(x) \geq 0$ 且

$$\int_0^{+\infty} g(x) dx = 1.$$

【3-7】 设二维随机变量 (X, Y) 的联合密度函数

$$f(x, y) = \begin{cases} k(6 - x - y) & (0 < x < 2, 2 < y < 4), \\ 0 & (\text{其他}). \end{cases}$$

计算: (1) 常数 k ; (2) $P(X \leq 1, Y \leq 3)$; (3) $P(X \leq 1.5)$;
(4) $P(X+Y \leq 4)$.

解 (1) 由归一性可得

$$\begin{aligned}\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy &= \int_0^2 \int_2^4 k(6-x-y) dy dx = k \int_0^2 (6-2x) dx \\ &= 8k = 1,\end{aligned}$$

所以 $k = \frac{1}{8}$.

$$\begin{aligned}(2) P(X \leq 1, Y \leq 3) &= \iint_{\substack{x \leq 1 \\ y \leq 3}} f(x, y) dx dy = \int_0^1 \int_2^3 \frac{1}{8} (6-x-y) dy dx \\ &= \frac{1}{8} \int_0^1 \left(\frac{7}{2} - x \right) dx = \frac{3}{8}.\end{aligned}$$

(3) 方法 1

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \begin{cases} \int_2^4 \frac{1}{8} (6-x-y) dy = \frac{3}{4} - \frac{1}{4}x & (0 < x < 2), \\ 0 & (\text{其他}), \end{cases}$$

$$P(X \leq 1.5) = \int_{-\infty}^{1.5} f_X(x) dx = \int_0^{1.5} \left(\frac{3}{4} - \frac{1}{4}x \right) dx = \frac{27}{32}.$$

方法 2

$$\begin{aligned}P(X \leq 1.5) &= \iint_{x \leq 1.5} f(x, y) dx dy = \int_{-\infty}^{1.5} \int_{-\infty}^{+\infty} f(x, y) dy dx \\ &= \int_0^{1.5} \int_2^4 \frac{1}{8} (6-x-y) dy dx \\ &= \int_0^{1.5} \left(\frac{3}{4} - \frac{1}{4}x \right) dx = \frac{27}{32}.\end{aligned}$$

$$\begin{aligned}(4) P(X+Y \leq 4) &= \iint_{x+y \leq 4} f(x, y) dx dy \\ &= \int_0^2 \int_2^{4-x} \frac{1}{8} (6-x-y) dy dx \\ &= \int_0^2 \left(6-4x + \frac{1}{2}x^2 \right) dx = \frac{2}{3}.\end{aligned}$$

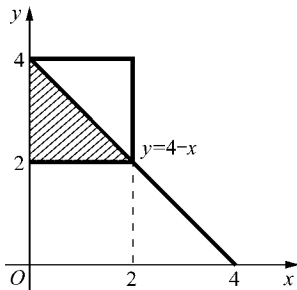


图 3-1

点评 本题的概率计算都是二重积分的计算,而化二重积分为累次积分的关键之处是找对每个积分的上、下限.下面将问题与解题方法作一般归纳.

设 $(X, Y) \sim f(x, y) \begin{cases} \neq 0 & ((x, y) \in G), \\ = 0 & ((x, y) \notin G). \end{cases}$ 则

$$\begin{aligned} P((X, Y) \in D) &= \iint_{(x, y) \in D} f(x, y) dx dy \\ &= \int_{x_1}^{x_2} dx \int_{y_1(x)}^{y_2(x)} f(x, y) dy. \end{aligned}$$

累次积分的积分区域是区域 D 与 G 的交 (见图 3-2). 化二重积分为累次积分还有个积分次序的问题,即先积 x 还是先积 y , 这要视具体题目而定.

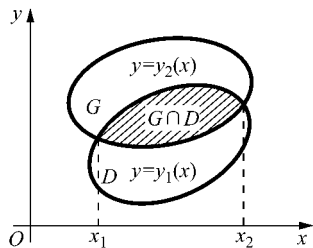


图 3-2

【3-8】 设二维随机变量 (X, Y) 的联合密度函数

$$f(x, y) = \begin{cases} 8xy & (0 \leq x < y, 0 \leq y \leq 1), \\ 0 & (\text{其他}). \end{cases}$$

(1) 计算 $P(X+Y \geq 1), P(X < 0.5)$;

(2) 求边缘密度函数 $f_X(x), f_Y(y)$.

$$\begin{aligned} \text{解} \quad (1) \quad P(X+Y \geq 1) &= \iint_{x+y \geq 1} f(x, y) dx dy \\ &= \int_{0.5}^1 dy \int_{1-y}^y 8xy dx \\ &= \frac{5}{6} (\text{见图 3-3}), \end{aligned}$$

$$\begin{aligned} P(X < 0.5) &= \iint_{x < 0.5} f(x, y) dx dy \\ &= \int_0^{0.5} dx \int_x^1 8xy dy \\ &= \frac{7}{16} (\text{见图 3-4}). \end{aligned}$$

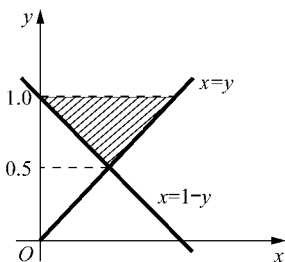


图 3-3

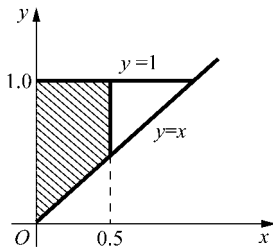


图 3-4

$$(2) f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \begin{cases} \int_x^1 8xy dy = 4x - 4x^3 & (0 \leq x < 1), \\ 0 & (\text{其他}). \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \begin{cases} \int_0^y 8xy dx = 4y^3 & (0 \leq y < 1), \\ 0 & (\text{其他}). \end{cases}$$

点评 若将题(1)中的两个二重积分的积分次序对调,则会增加计算量,因为此时

$$P(X+Y \geq 1) = \int_0^{0.5} dx \int_{1-x}^1 8xy dy + \int_{0.5}^1 dx \int_x^1 8xy dy,$$

$$P(X < 0.5) = \int_0^{0.5} dy \int_0^y 8xy dx + \int_{0.5}^1 dy \int_0^{0.5} 8xy dx.$$

【3-9】 设二维随机变量 (X, Y) 的联合密度函数

$$f(x, y) = \begin{cases} k(2 - \sqrt{x^2 + y^2}) & (x^2 + y^2 \leq 4), \\ 0 & (\text{其他}). \end{cases}$$

求:(1) 常数 k ;

(2) (X, Y) 在以原点为圆心、以1为半径的圆域内取值的概率.

解 (1) 由归一性可知

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = \int_0^{2\pi} d\theta \int_0^2 k(2-r)r dr = \frac{8}{3} \pi k = 1,$$

$$\text{解得 } k = \frac{3}{8\pi}.$$

$$\begin{aligned}
 (2) P(X^2 + Y^2 < 1) &= \iint_{x^2 + y^2 < 1} f(x, y) dx dy \\
 &= \int_0^{2\pi} d\theta \int_0^1 \frac{3}{8\pi} (2-r) dr \\
 &= \frac{1}{2} \text{ (见图 3-5).}
 \end{aligned}$$

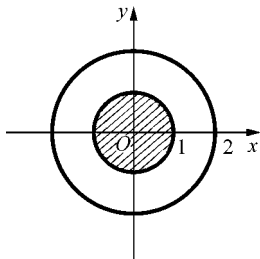


图 3-5

【3-10】 设二维随机变量 (X, Y) 的联合密度函数

$$f(x, y) = \begin{cases} x^2 + \frac{1}{3}xy & (0 \leq x \leq 1, 0 \leq y \leq 2), \\ 0 & \text{(其他).} \end{cases}$$

计算 $P(X+Y < 2)$.

解 先求出累次积分的积分区域:

$$\begin{aligned}
 D \cap G &= \{(x, y) \mid x + y < 2\} \cap \{(x, y) \mid 0 < x < 1, 0 < y < 2\} \\
 &= \{(x, y) \mid 0 < x < 1, 0 < y < 2 - x\}.
 \end{aligned}$$

于是所求概率

$$\begin{aligned}
 P(X+Y < 2) &= \iint_{x+y < 2} f(x, y) dx dy = \int_0^1 dx \int_0^{2-x} \left(x^2 + \frac{1}{3}xy\right) dy \\
 &= \int_0^1 \left(\frac{2}{3}x + \frac{4}{3}x^2 - \frac{5}{6}x^3\right) dx \\
 &= \frac{41}{72} \approx 0.5694.
 \end{aligned}$$

【3-11】 设 (X, Y) 在曲线 $y = x^2, y = 4$ 所围成的区域 G 内 (见图 3-6) 服从均匀分布, 求边缘密度函数 $f_X(x), f_Y(y)$.

解 区域 G 的面积

$$\begin{aligned}
 S &= \iint_G dx dy = 2 \int_0^2 dx \int_{x^2}^4 dy \\
 &= 2 \int_0^2 (4 - x^2) dx = \frac{32}{3},
 \end{aligned}$$

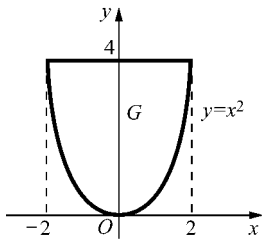


图 3-6

所以, (X, Y) 的联合密度函数

$$f(x, y) = \begin{cases} \frac{3}{32} & (-2 < x < 2, x^2 < y < 4), \\ 0 & (\text{其他}). \end{cases}$$

当 $|x| \geq 2$ 时, $f(x, y) = 0$, 故 $f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = 0$;

当 $|x| < 2$ 时,

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \int_{x^2}^4 \frac{3}{32} dy = \frac{3}{32} (4 - x^2).$$

从而

$$f_X(x) = \begin{cases} \frac{3}{32} (4 - x^2) & (|x| < 2), \\ 0 & (|x| \geq 2). \end{cases}$$

当 $y \leq 0$ 或 $y \geq 4$ 时, $f(x, y) = 0$, 故 $f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = 0$;

当 $0 < y < 4$ 时,

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{3}{32} dx = \frac{3}{16} \sqrt{y}.$$

从而

$$f_Y(y) = \begin{cases} \frac{3}{16} \sqrt{y} & (0 < y < 4), \\ 0 & (\text{其他}). \end{cases}$$

【3-12】 设随机变量 (X, Y) 在矩形域 $a \leq x \leq b, c \leq y \leq d$ 内服从均匀分布. 求:

(1) 密度函数 $f(x, y)$ 和分布函数 $F(x, y)$;

(2) 边缘密度函数 $f_X(x), f_Y(y)$.

解 (1) 矩形面积 $S = (b-a)(d-c)$, 故 (X, Y) 的联合密度函数

$$f(x, y) = \begin{cases} \frac{1}{(b-a)(d-c)} & (a \leq x \leq b, c \leq y \leq d), \\ 0 & (\text{其他}). \end{cases}$$

由于 $f(x, y)$ 仅在矩形域 $a \leq x \leq b, c \leq y \leq d$ 内取非零值, 过矩形各顶点分别作与 x 轴和 y 轴平行的直线, 这 4 条直线将整个平面分成 9 个区域, 使 $F(x, y)$ 取不同值的有 5 个区间, 如图 3-7 所示.

① 当 $x < a$ 或 $y < c$ 时, $f(x, y) = 0$,
从而 $F(x, y) = 0$;

② 当 $a \leq x \leq b, c \leq y \leq d$ 时,

$$\begin{aligned} F(x, y) &= \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv \\ &= \int_a^x \int_c^y \frac{1}{(b-a)(d-c)} dv du \\ &= \frac{(x-a)(y-c)}{(b-a)(d-c)}; \end{aligned}$$

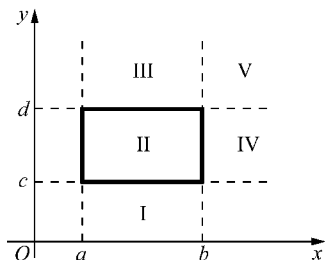


图 3-7

③ 当 $a \leq x \leq b, y > d$ 时,

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv = \int_a^x \int_c^d \frac{1}{(b-a)(d-c)} dv du = \frac{x-a}{b-a};$$

④ 当 $x > b, c \leq y \leq d$ 时,

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv = \int_a^b \int_c^y \frac{1}{(b-a)(d-c)} dv du = \frac{y-c}{d-c};$$

⑤ 当 $x > b, y > d$ 时,

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv = \int_a^b \int_c^d \frac{1}{(b-a)(d-c)} dv du = 1.$$

综上所述, (X, Y) 的联合分布函数

$$F(x, y) = \begin{cases} 0 & (x < a \text{ 或 } y < c), \\ \frac{(x-a)(y-c)}{(b-a)(d-c)} & (a \leq x \leq b, c \leq y \leq d), \\ \frac{x-a}{b-a} & (a \leq x \leq b, y > d), \\ \frac{y-c}{d-c} & (x > b, c \leq y \leq d), \\ 1 & (x > b, y > d). \end{cases}$$

(2) 方法 1

$$F_X(x) = F(x, +\infty) = \begin{cases} 0 & (x < a) \\ \frac{x-a}{b-a} & (a \leq x \leq b), \\ 1 & (x > b); \end{cases}$$

$$f_X(x) = F'_X(x) = \begin{cases} \frac{1}{b-a} & (a \leq x \leq b), \\ 0 & (\text{其他}). \end{cases}$$

$$F_Y(y) = F(+\infty, y) = \begin{cases} 0 & (y < c) \\ \frac{y-c}{d-c} & (c \leq y \leq d), \\ 1 & (y > d); \end{cases}$$

$$f_Y(y) = F'_Y(y) = \begin{cases} \frac{1}{d-c} & (c \leq y \leq d), \\ 0 & (\text{其他}). \end{cases}$$

方法 2

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \begin{cases} \frac{1}{b-a} & (a \leq x \leq b), \\ 0 & (\text{其他}). \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \begin{cases} \frac{1}{d-c} & (c \leq y \leq d), \\ 0 & (\text{其他}). \end{cases}$$

点评 已知分布函数求密度函数比较容易,反之不太容易.当密度函数 $f(x, y)$ 为分段函数,即其取值因区域不同而不同时,必将导致分布函数 $F(x, y)$ 也为分段函数,且分段数多于前者.

$F(x, y)$ 的定义域为整个二维平面,如何分区域呢?首先在 xOy 平面上画出 $f(x, y)$ 取非零值的区域 G .若 G 是矩形域,再过矩形域的各顶点分别作平行于 x 轴和 y 轴的直线,这些直线将平面分成若干子区域,在各个子区域内分别找出 $f(x, y)$ 的非零值区域.这一步相当于确定了重积分化为累次积分时积分的上、下限.计算重积分便得 $F(x, y)$ 在各区域的表达式,最后加以综合即得联合分布函数 $F(x, y)$.

如果 $f(x, y)$ 取非零值的区域 G 不是矩形域,则解题过程更复杂一些,但基本思路与方法不变.

值得强调的是,使 $F(x, y)$ 取不同值的子区域不管有几个(比如本题的 5 个),这些子区域必须充满全平面而没有遗漏.

【3-13】 设随机变量 (X, Y) 的联合密度函数

$$f(x, y) = \begin{cases} 3x & (0 \leq x \leq 1, 0 \leq y \leq x), \\ 0 & (\text{其他}). \end{cases}$$

(1) 求边缘密度函数;

(2) X 与 Y 是否相互独立?

解 (1) 当 $0 \leq x \leq 1$ 时

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \int_0^x 3x dy = 3x^2,$$

所以 (X, Y) 关于 X 的边缘密度函数

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

当 $0 \leq y \leq 1$ 时

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_y^1 3x dx = \frac{3}{2}(1 - y^2),$$

所以 (X, Y) 关于 Y 的边缘密度函数

$$f_Y(y) = \begin{cases} \frac{3}{2}(1 - y^2) & (0 \leq y \leq 1), \\ 0 & (\text{其他}). \end{cases}$$

(2) 因为 $f(1, 0) = 3 \neq 3 \times \frac{3}{2} = f_X(1)f_Y(0)$,

所以 X 与 Y 不相互独立.

【3-14】 设随机变量 (X, Y) 的联合密度函数

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

(1) 求边缘密度函数;

(2) X 与 Y 是否相互独立?

解 (1) 由边缘密度函数公式得

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{+\infty} f(x, y) dy = \int_{-\infty}^{+\infty} \frac{1}{\pi^2(x^2 + y^2 + x^2y^2 + 1)} dy \\ &= \frac{1}{\pi(1 + x^2)} \int_{-\infty}^{+\infty} \frac{1}{\pi(1 + y^2)} dy = \frac{1}{\pi(1 + x^2)} \cdot 1 \\ &= \frac{1}{\pi(1 + x^2)} \quad (x \in (-\infty, +\infty)). \end{aligned}$$

同理可得

$$f_Y(y) = \frac{1}{\pi(1+y^2)} \quad (y \in (-\infty, +\infty)).$$

(2) 显然, 任给 (x, y) 恒有 $f(x, y) = f_X(x)f_Y(y)$, 所以 X 与 Y 相互独立.

【3-15】 设随机变量 X 与 Y 相互独立, 它们的密度函数分别为

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

$$f_Y(y) = \begin{cases} \lambda e^{-\lambda y} & (y > 0), \\ 0 & (y \leq 0), \end{cases}$$

其中 $\lambda > 0$ 为常数. 求 (X, Y) 的联合密度函数及 $P(X \geq Y)$.

解 由于 X 与 Y 相互独立, 故 (X, Y) 的联合密度函数

$$f(x, y) = f_X(x)f_Y(y) = \begin{cases} \frac{\lambda e^{-\lambda y}}{b-a} & (a < x < b, y > 0), \\ 0 & (\text{其他}). \end{cases}$$

(1) 当 $b \leq 0$ 时, 使 $f(x, y)$ 取非零值的区域 $G = \{(x, y) | a < x < b, y > 0\}$ 与区域 $D = \{(x, y) | x \geq y\}$ 的交集是空集 (见图 3-8), 所以

$$P(X \geq Y) = \iint_{x \geq y} f(x, y) dx dy = 0.$$

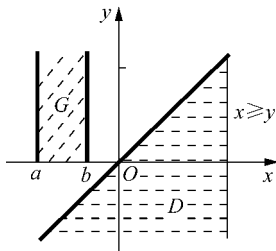


图 3-8

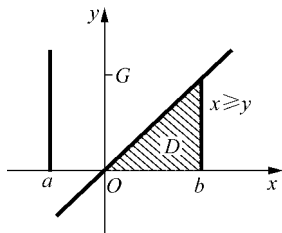


图 3-9

(2) 当 $a < 0 < b$ 时 (见图 3-9)

$$G \cap D = \{(x, y) | 0 < x < b, 0 < y \leq x\},$$

从而

$$\begin{aligned}
 P(X \geq Y) &= \iint_{x \geq y} f(x, y) dx dy = \int_0^b dx \int_0^x \frac{\lambda e^{-\lambda y}}{b-a} dy \\
 &= \frac{1}{b-a} \int_0^b (1 - e^{-\lambda x}) dx = \frac{\lambda b + e^{-\lambda b} - 1}{(b-a)\lambda} \\
 &= 1 + \frac{\lambda a + e^{-\lambda b} - 1}{(b-a)\lambda}.
 \end{aligned}$$

(3) 当 $0 \leq a < b$ 时 (见图 3-10)

$$G \cap D = \{(x, y) \mid a < x < b, 0 < y \leq x\},$$

$$\begin{aligned}
 P(X \geq Y) &= \iint_{x \geq y} f(x, y) dx dy \\
 &= \int_a^b dx \int_0^x \frac{\lambda e^{-\lambda y}}{b-a} dy \\
 &= \frac{1}{b-a} \int_a^b (1 - e^{-\lambda x}) dx \\
 &= 1 + \frac{e^{-\lambda b} - e^{-\lambda a}}{(b-a)\lambda}.
 \end{aligned}$$

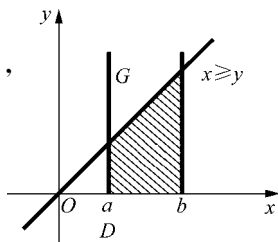


图 3-10

综上所述,可得

$$P(X \geq Y) = \begin{cases} 0 & (b \leq 0) \\ 1 + \frac{a\lambda + e^{-\lambda b} - 1}{(b-a)\lambda} & (a < 0 < b), \\ 1 + \frac{e^{-\lambda b} - e^{-\lambda a}}{(b-a)\lambda} & (0 \leq a < b). \end{cases}$$

点评 本题最常见的一种误解是未对 a, b 的正负性进行讨论,而直接求得 $P(X \geq Y) = \int_a^b dx \int_0^x \frac{\lambda e^{-\lambda y}}{b-a} dy = 1 + \frac{e^{-\lambda b} - e^{-\lambda a}}{(b-a)\lambda}$. 事实上这样求解心中已默认 $0 \leq a < b$, 但是题中的 a, b 并未指定是正数. 即使都是负的,其定义的 $f_X(x)$ 以及要求的 $f(x, y)$ 都存在且有意义. 所以正确求解必须按题解中的三种情况讨论.

【3-16】 设二维随机变量 (X, Y) 的密度函数

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

(1) 求条件密度函数 $f_{X|Y}(x|y), f_{Y|X}(y|x)$;

(2) X 与 Y 是否相互独立?

解 (1) 先求边缘密度函数, 再根据公式求条件密度函数.

当 $0 < x < 1$ 时, $f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \int_{-x}^x dy = 2x$, 从而

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

当 $|y| < 1$ 时, $f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_{|y|}^1 dx = 1 - |y|$, 从而

$$f_Y(y) = \begin{cases} 1 - |y| & (|y| < 1), \\ 0 & (|y| \geq 1). \end{cases}$$

于是当 $|y| < 1$ 时

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \begin{cases} \frac{1}{1 - |y|} & (|y| < x < 1), \\ 0 & (\text{其他}). \end{cases}$$

当 $0 < x < 1$ 时

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} = \begin{cases} \frac{1}{2x} & (|y| < x), \\ 0 & (|y| \geq x). \end{cases}$$

(2) 因为 $f(0.5, 0.2) = 1 \neq 1 \times 0.8 = f_X(0.5)f_Y(0.2)$, 所以 X 与 Y 不相互独立.

【3-17】 设随机变量 X 与 Y 相互独立, 且都服从参数为 1 的指数分布, 求 (X, Y) 的联合密度函数及条件密度函数 $f_{X|Y}(x|y)$.

解 由题设知

$$f_X(x) = \begin{cases} e^{-x} & (x \geq 0), \\ 0 & (x < 0), \end{cases}$$

$$f_Y(y) = \begin{cases} e^{-y} & (y \geq 0), \\ 0 & (y < 0). \end{cases}$$

又因 X 与 Y 相互独立, 可得 (X, Y) 的联合密度函数

$$f(x, y) = f_X(x)f_Y(y) = \begin{cases} e^{-(x+y)} & (x \geq 0, y \geq 0), \\ 0 & (\text{其他}). \end{cases}$$

X 的条件密度函数(当 $y \geq 0$ 时)

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \begin{cases} e^{-x} & (x \geq 0), \\ 0 & (x < 0). \end{cases}$$

点评 ① 由于 X 与 Y 相互独立, 条件密度函数就是无条件密度函数, 故当 $y \geq 0$ 时

$$f_{X|Y}(x|y) = f_X(x) = \begin{cases} e^{-x} & (x \geq 0), \\ 0 & (x < 0). \end{cases}$$

② X 的条件密度函数仍是一元函数, 故不能写成

$$f_{X|Y}(x|y) = \begin{cases} e^{-x} & (x \geq 0, y \geq 0), \\ 0 & (\text{其他}). \end{cases}$$

【3-18】 设随机变量 (X, Y) 的联合密度函数

$$f(x, y) = \begin{cases} \frac{21}{4}x^2y & (x^2 \leq y \leq 1), \\ 0 & (\text{其他}). \end{cases}$$

求: (1) 条件密度函数 $f_{X|Y}(x|y)$, 并写出 $Y = \frac{1}{2}$ 时 X 的条件密度函数;

(2) 条件密度函数 $f_{Y|X}(y|x)$, 并写出 $X = \frac{1}{3}$ 时 Y 的条件密度函数;

(3) $P\left(Y \geq \frac{3}{4} \mid X = \frac{1}{2}\right)$.

解 当 $-1 \leq x \leq 1$ 时

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \int_{x^2}^1 \frac{21}{4}x^2y dy = \frac{21}{8}x^2(1 - x^4);$$

当 $0 \leq y \leq 1$ 时

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{21}{4}x^2y dx = \frac{7}{2}y^{\frac{5}{2}}.$$

(1) 当 $0 < y \leq 1$ 时

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \begin{cases} \frac{3}{2}x^2y^{-\frac{3}{2}} & (-\sqrt{y} < x < \sqrt{y}), \\ 0 & (\text{其他}); \end{cases}$$

$$f_{X|Y}\left(x \mid y = \frac{1}{2}\right) = \begin{cases} 3\sqrt{2}x^2 & \left(-\frac{1}{\sqrt{2}} < x < \frac{1}{\sqrt{2}}\right), \\ 0 & (\text{其他}). \end{cases}$$

(2) 当 $-1 < x < 1$ 时

$$f_{Y|X}(y \mid x) = \frac{f(x, y)}{f_X(x)} = \begin{cases} \frac{2y}{1-x^4} & (x^2 < y < 1), \\ 0 & (\text{其他}); \end{cases}$$

$$f_{Y|X}\left(y \mid x = \frac{1}{3}\right) = \begin{cases} \frac{81}{40}y & \left(\frac{1}{9} < y < 1\right), \\ 0 & (\text{其他}). \end{cases}$$

$$(3) f_{Y|X}\left(y \mid x = \frac{1}{2}\right) = \begin{cases} \frac{32}{15}y & \left(\frac{1}{4} < y < 1\right), \\ 0 & (\text{其他}), \end{cases}$$

则

$$P\left(Y \geq \frac{3}{4} \mid X = \frac{1}{2}\right) = \int_{y \geq \frac{3}{4}} f_{Y|X}\left(y \mid x = \frac{1}{2}\right) dy = \int_{\frac{3}{4}}^1 \frac{32}{15}y dy = \frac{7}{15}.$$

【3-19】 设随机变量 X 与 Y 相互独立, X 在区间 $(0, 1)$ 上服从均匀分布, Y 服从参数为 $\frac{1}{2}$ 的指数分布. 求:

(1) (X, Y) 的联合密度函数;

(2) 方程 $x^2 + 2xX + Y = 0$ 关于 x 有实根的概率.

解 (1) 由题设知

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

$$f_Y(y) = \begin{cases} \frac{1}{2}e^{-\frac{1}{2}y} & (y > 0), \\ 0 & (y \leq 0). \end{cases}$$

又因 X 与 Y 相互独立, 可得 (X, Y) 的联合密度

$$f(x, y) = \begin{cases} \frac{1}{2}e^{-\frac{1}{2}y} & (0 < x < 1, y > 0), \\ 0 & (\text{其他}). \end{cases}$$

(2) 方程有实根,故 $\Delta=4X^2-4Y\geqslant 0$,即 $X^2\geqslant Y$,

$$\begin{aligned} P(\text{方程有实根}) &= P(X^2 \geqslant Y) = \iint_{x^2 \geqslant y} f(x, y) dx dy = \int_0^1 dx \int_0^{x^2} \frac{1}{2} e^{-\frac{1}{2}y} dy \\ &= \int_0^1 (1 - e^{-\frac{x^2}{2}}) dx = 1 - \sqrt{2\pi} \int_0^1 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= 1 - \sqrt{2\pi} [\Phi(1) - \Phi(0)] \\ &\approx 0.1445. \end{aligned}$$

【3-20】 设随机变量 (X, Y) 的分布律如下表所示:

$\begin{matrix} X \\ Y \end{matrix}$	1	2	3
1	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{18}$
2	$\frac{1}{3}$	α	β

问表中 α, β 取什么值时 X 与 Y 相互独立?

解 方法 1 要使 X, Y 独立, 必有 $P(X=2, Y=1) = P(X=2) \cdot P(Y=1)$, 即 $\frac{1}{9} = \left(\frac{1}{9} + \alpha\right) \left(\frac{1}{6} + \frac{1}{9} + \frac{1}{18}\right)$, 解得 $\alpha = \frac{2}{9}$. 再由分布律的归一性得 $\alpha + \beta = 1 - \frac{1}{3} - \frac{1}{6} - \frac{1}{9} - \frac{1}{18} = \frac{1}{3}$, 解得 $\beta = \frac{1}{9}$.

所以当 $\alpha = \frac{2}{9}, \beta = \frac{1}{9}$ 时, X 与 Y 相互独立.

方法 2 X 与 Y 相互独立 $\Leftrightarrow$ 联合概率矩阵秩为 1.

要使联合概率矩阵 $\begin{pmatrix} \frac{1}{6} & \frac{1}{9} & \frac{1}{18} \\ \frac{1}{3} & \alpha & \beta \end{pmatrix}$ 秩为 1, 只需任两行(或列)对应

元素成比例. 由第一列可见, 第一行与第二行之比为 $1:2$, 所以 $\alpha = \frac{2}{9}$, $\beta = \frac{1}{9}$ 时, 矩阵秩为 1, 从而保证 X 与 Y 相互独立.

点评 方法 1 是一般教科书上介绍的,方法 2 由于是新提出的,故在目前的教科书上未曾见过. 它用线性代数的知识解决概率论中的问题,其特点是运算简便. 如对本题,几乎是看一眼,就能说出答案.

方法 2 的提出者是上海交通大学数学系张跃辉教授,其论文发表在《海南大学学报》上(1999 年 12 月第 4 期).

【3-21】 设随机变量 X 和 Y 相互独立,且 $X \sim B(n, p), Y \sim B(n, p)$. 证明: $Z = X + Y \sim B(2n, p)$.

证 由题设知 X 与 Y 的分布律

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

$$P(Y = j) = C_n^j p^j (1-p)^{n-j} \quad (j = 0, 1, 2, \dots, n).$$

由 X 与 Y 的独立性可得

$$\begin{aligned} P(Z = k) &= P(X + Y = k) = \sum_{m=0}^k P(X = m, Y = k - m) \\ &= \sum_{m=0}^k P(X = m) P(Y = k - m) \\ &= \sum_{m=0}^k C_n^m p^m (1-p)^{n-m} C_n^{k-m} p^{k-m} (1-p)^{n-k+m} \\ &= p^k (1-p)^{2n-k} \sum_{m=0}^k C_n^m C_n^{k-m} \\ &= C_{2n}^k p^k (1-p)^{2n-k} \quad (k = 0, 1, 2, \dots, 2n). \end{aligned}$$

此即表示 $Z = X + Y \sim B(2n, p)$.

点评 本题证明中有两个关键之处:

(1) $P(X + Y = k) = \sum_{m=0}^k P(X = m, Y = k - m)$. 事实上,事件 $(X + Y = k)$ 与 $\bigcup_{m=0}^k P(X = m, Y = k - m)$ 等价,后者是 $k+1$ 个两两互不相容事件之并.

(2) $\sum_{m=0}^k C_n^m C_n^{k-m} = C_{2n}^k$. 这是组合恒等式.

本题的结论告诉我们,二项分布在独立的条件下具有可加性[设两

个独立的随机变量服从同一分布(参数可以不同),若它们的和仍服从同一分布,则称此分布具有可加性].

相仿可以证明更一般的结论:

若随机变量 X, Y 相互独立,其分布律分别如下:

$$P(X=i)=p(i) \quad (i=0,1,2,\cdots),$$

$$P(Y=j)=q(j) \quad (j=0,1,2,\cdots),$$

则 $Z=X+Y$ 的分布律

$$P(Z=k)=\sum_{m=0}^k p(m)q(k-m) \quad (k=0,1,2,\cdots).$$

上式称为离散型卷积公式,专门用于求独立的离散型随机变量和的分布.

【3-22】 设随机变量 X, Y 相互独立,且 $X \sim \pi(\lambda), Y \sim \pi(\mu)$, 求 $Z=X+Y$ 的分布律.

解 由题设知

$$P(X=i)=\frac{\lambda^i}{i!}e^{-\lambda}=p(i) \quad (i=0,1,2,\cdots).$$

$$P(Y=j)=\frac{\mu^j}{j!}e^{-\mu}=q(j) \quad (j=0,1,2,\cdots).$$

由上述卷积公式可得 $Z=X+Y$ 的分布律

$$\begin{aligned} P(Z=k) &= \sum_{m=0}^k p(m)q(k-m) = \sum_{m=0}^k \frac{\lambda^m}{m!}e^{-\lambda} \frac{\mu^{k-m}}{(k-m)!}e^{-\mu} \\ &= \frac{e^{-(\lambda+\mu)}}{k!} \sum_{m=0}^k \frac{k!}{m!(k-m)!} \lambda^m \mu^{k-m} \\ &= \frac{(\lambda+\mu)^k}{k!} e^{-(\lambda+\mu)} \quad (k=0,1,2,\cdots). \end{aligned}$$

点评 由本题结论得如下命题:

若 $X \sim \pi(\lambda), Y \sim \pi(\mu)$, 且 X 与 Y 相互独立,则 $X+Y \sim \pi(\lambda+\mu)$.
即泊松分布具有可加性.

【3-23】 设 $X \sim E(1), Y \sim E(1)$, 且 X, Y 相互独立. 求随机变量 $Z=X+Y$ 的密度函数.

解 由题设可知 X, Y 的密度函数

$$f_X(x) = \begin{cases} e^{-x} & (x > 0), \\ 0 & (x \leq 0); \end{cases}$$

$$f_Y(y) = \begin{cases} e^{-y} & (y > 0), \\ 0 & (y \leq 0). \end{cases}$$

由于 X, Y 相互独立, 故可用卷积公式计算, 即有

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) dx.$$

上式被积函数不为零的区域为 $\begin{cases} x > 0, \\ z-x > 0, \end{cases}$ 即 $0 < x < z$.

当 $z \leq 0$ 时, $f_Z(z) = 0$,

当 $z > 0$ 时, $f_Z(z) = \int_0^z e^{-x} e^{-(z-x)} dx = \int_0^z e^{-z} dx = ze^{-z}$,

所以 Z 的密度函数

$$f_Z(z) = \begin{cases} ze^{-z} & (z > 0), \\ 0 & (z \leq 0). \end{cases}$$

点评 显然 Z 不再服从指数分布, 所以指数分布不具有可加性.

【3-24】 设 $X \sim U(0, 1), Y \sim E(1)$, 且 X, Y 相互独立. 求随机变量 $Z = X + Y$ 的密度函数.

解 由题设可知 X, Y 的密度函数

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

$$f_Y(y) = \begin{cases} e^{-y} & (y > 0), \\ 0 & (y \leq 0). \end{cases}$$

由于 X, Y 相互独立, 故可用卷积公式计算, 即有

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) dx. \quad (*)$$

上式被积函数不为零的区域为 $\begin{cases} 0 < x < 1, \\ z-x > 0, \end{cases}$ 即 $\begin{cases} 0 < x < 1, \\ x < z. \end{cases}$

由上面不等式组边边相等得 z 轴上的分界点 $z=0$ 与 $z=1$, 它们将 z 轴分成三段:

① 当 $z < 0$ 时, 不等式组无解, $f_Z(z) = 0$.

② 当 $0 \leq z \leq 1$ 时, 不等式组解为 $0 < x < z$, 此即为 (*) 式积分具体计算时积分的上、下限.

$$f_Z(z) = \int_0^z 1 \cdot e^{-(z-x)} dx = e^{-z} \int_0^z e^x dx = 1 - e^{-z}.$$

③ 当 $z > 1$ 时, 不等式组的解为 $0 < x < 1$, 于是

$$f_Z(z) = \int_0^1 e^{-(z-x)} dx = e^{-z} \int_0^1 e^x dx = (e-1)e^{-z}.$$

所以 z 的密度函数

$$f_Z(z) = \begin{cases} 0 & (z < 0), \\ 1 - e^{-z} & (0 \leq z \leq 1), \\ (e-1)e^{-z} & (z > 1). \end{cases}$$

点评 求 $Z=X+Y$ 的分布密度既是本章中的要点之一, 也是本章中的难点. 所谓难点有两个: 一是找准 z 轴上的分界点; 二是化重积分为累次积分时积分限如何确定, 或是按卷积公式计算时积分限如何演变. 在题 3-23 和 3-24 中都采用了不等式组定限法, 化解了这两个难点. 当然采用作图法也可解决这两个难点, 由于一般教科书都有此介绍, 这里不再赘述.

【3-25】 设随机变量 (X, Y) 的联合密度函数

$$f(x, y) = \begin{cases} 2e^{-(x+2y)} & (x > 0, y > 0), \\ 0 & (\text{其他}). \end{cases}$$

求随机变量 $Z=X-2Y$ 的密度函数.

分析 本题中的联合密度函数为可分离变量, 即当 $x > 0, y > 0$ 时,

$$f(x, y) = e^{-x} \cdot 2e^{-2y},$$

故 X, Y 相互独立, 满足使用卷积公式的条件之一, 但 Y 的系数是 -2 不是 1 , 所以不能直接用卷积公式. 下面将卷积公式予以推广:

① 若 $(X, Y) \sim f(x, y)$, 则 $Z=aX+bY(a \neq 0, b \neq 0)$ 的密度函数

$$f_Z(z) = \frac{1}{|b|} \int_{-\infty}^{+\infty} f\left(x, \frac{z-ax}{b}\right) dx = \frac{1}{|a|} \int_{-\infty}^{+\infty} f\left(\frac{z-by}{a}, y\right) dy.$$

② 若 $X \sim f_X(x), Y \sim f_Y(y)$, 且 X, Y 相互独立, 则 $Z=aX+bY(a \neq 0, b \neq 0)$ 的密度函数

$$\begin{aligned} f_Z(z) &= \frac{1}{|b|} \int_{-\infty}^{+\infty} f_X(x) f_Y\left(\frac{z-ax}{b}\right) dx \\ &= \frac{1}{|a|} \int_{-\infty}^{+\infty} f_X\left(\frac{z-by}{a}\right) f_Y(y) dy. \end{aligned}$$

解 由联合密度函数易得边缘密度函数

$$\begin{aligned} f_X(x) &= \begin{cases} e^{-x} & (x > 0), \\ 0 & (x \leq 0), \end{cases} \\ f_Y(y) &= \begin{cases} 2e^{-2y} & (y > 0), \\ 0 & (y \leq 0). \end{cases} \end{aligned}$$

显然 X, Y 相互独立, 由推广的卷积公式可得

$$f_Z(z) = \frac{1}{2} \int_{-\infty}^{+\infty} f_X(x) f_Y\left(\frac{x-z}{2}\right) dx.$$

上式被积函数不为零的区域 $\begin{cases} x > 0, \\ \frac{x-z}{2} > 0, \end{cases}$ 即 $\begin{cases} x > 0, \\ x > z. \end{cases}$ 由此得 z 轴上一个分

界点 $z=0$.

① 当 $z < 0$ 时, 不等式组的解为 $x > 0$, 于是

$$f_Z(z) = \frac{1}{2} \int_0^{+\infty} e^{-x} \cdot 2e^{-(x-z)} dx = \frac{1}{2} e^z \int_0^{+\infty} 2e^{-2x} dx = \frac{1}{2} e^z.$$

② 当 $z \geq 0$ 时, 不等式组的解为 $x > z$, 于是

$$f_Z(z) = \frac{1}{2} \int_z^{+\infty} e^{-x} \cdot 2e^{-(x-z)} dx = \frac{1}{2} e^z \int_z^{+\infty} 2e^{-2x} dx = \frac{1}{2} e^{-z}.$$

所以 Z 的密度函数

$$f_Z(z) = \begin{cases} \frac{1}{2} e^z & (z < 0), \\ \frac{1}{2} e^{-z} & (z \geq 0). \end{cases}$$

点评 本题也可利用分布函数法求解, 即先求出 $F_Z(z)$, 再求导得 $f_Z(z)$, 但此法计算过程会冗长一些.

【3-26】 设 X 与 Y 是两个相互独立的随机变量, 它们服从同一分布, 其密度函数

$$f(t) = \begin{cases} \frac{1\,000}{t^2} & (t > 1\,000), \\ 0 & (t \leq 1\,000). \end{cases}$$

求随机变量 $Z = \frac{X}{Y}$ 的密度函数.

解 方法 1 (分布函数法)

由题设可知, (X, Y) 的联合密度函数

$$f(x, y) = f_X(x)f_Y(y) = \begin{cases} \left(\frac{1\,000}{xy}\right)^2 & (x > 1\,000, y > 1\,000), \\ 0 & (\text{其他}). \end{cases}$$

当 $z \leq 0$ 时, 显然有 $F_Z(z) = 0$;

当 $0 < z < 1$ 时,

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P\left(\frac{X}{Y} \leq z\right) = \iint_{\frac{x}{z} \leq y} f(x, y) dx dy \\ &= \int_{1\,000}^{+\infty} dx \int_{\frac{x}{z}}^{+\infty} \left(\frac{1\,000}{xy}\right)^2 dy \\ &= \int_{1\,000}^{+\infty} \frac{1\,000^2 z}{x^3} dx = \frac{1}{2} z; \end{aligned}$$

当 $z \geq 1$ 时,

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P\left(\frac{X}{Y} \leq z\right) = \iint_{x \leq yz} f(x, y) dx dy \\ &= \int_{1\,000}^{+\infty} dy \int_{1\,000}^{yz} \left(\frac{1\,000}{xy}\right)^2 dx \\ &= \int_{1\,000}^{+\infty} \frac{1\,000}{y^2} \left(1 - \frac{1\,000}{yz}\right) dy = 1 - \frac{1}{2z}. \end{aligned}$$

于是

$$F_Z(z) = \begin{cases} 0 & (z \leq 0), \\ \frac{1}{2} z & (0 < z < 1), \\ 1 - \frac{1}{2z} & (z \geq 1). \end{cases}$$

所以 $Z = \frac{X}{Y}$ 的密度函数

$$f_Z(z) = F'_Z(z) = \begin{cases} 0 & (z \leq 0), \\ \frac{1}{2} & (0 < z < 1), \\ \frac{1}{2z^2} & (z \geq 1). \end{cases}$$

方法 2 (公式法)

$$f_Z(z) = \int_{-\infty}^{+\infty} |y| f(yz, y) dy$$

$$\stackrel{\text{独立}}{=} \int_{-\infty}^{+\infty} |y| f_X(yz) f_Y(y) dy \quad (\text{推导见后面}).$$

上式被积函数不为零的区域为 $\begin{cases} yz > 1000, \\ y > 1000, \end{cases}$ 即 $\begin{cases} y > \frac{1000}{z}, \\ y > 1000. \end{cases}$ 显然 $z =$

0 为 z 轴的一个分界点, 令 $\frac{1000}{z} = 1000$ 得 z 轴上另一个分界点 $z = 1$.

当 $z \leq 0$ 时, 不等式组无解, $f_Z(z) = 0$;

当 $0 < z < 1$ 时, 不等式组的解为 $y > \frac{1000}{z}$, 于是

$$f_Z(z) = \int_{\frac{1000}{z}}^{+\infty} y \frac{1000}{y^2 z^2} \frac{1000}{y^2} dy = \frac{1000^2}{z^2} \int_{\frac{1000}{z}}^{+\infty} \frac{1}{y^3} dy = \frac{1}{2};$$

当 $z \geq 1$ 时, 不等式组的解为 $y > 1000$, 于是

$$f_Z(z) = \int_{1000}^{+\infty} \frac{1000^2}{z^2 y^3} dy = \frac{1}{2z^2}.$$

所以 $Z = \frac{X}{Y}$ 的密度函数

$$f_Z(z) = \begin{cases} 0 & (z \leq 0), \\ \frac{1}{2} & (0 < z < 1), \\ \frac{1}{2z^2} & (z \geq 1). \end{cases}$$

关于两个随机变量商的密度函数公式的推导:

设二维随机变量 (X, Y) 的密度函数为 $f(x, y)$, 引入新的二维随机变量 (Z, V) , 其联合密度函数记为 $g(z, v)$. 新老随机变量的关系是 $Z = \frac{X}{Y}, V = Y$, 即有

$$\begin{cases} z = \frac{x}{y}, \\ v = y, \end{cases} \Rightarrow \begin{cases} x = zv, \\ y = v. \end{cases}$$

其雅可比(Jacobi)行列式

$$J = \begin{vmatrix} \frac{\partial(zv)}{\partial z} & \frac{\partial(zv)}{\partial v} \\ \frac{\partial(v)}{\partial z} & \frac{\partial(v)}{\partial v} \end{vmatrix} = \begin{vmatrix} v & z \\ 0 & 1 \end{vmatrix} = v,$$

于是

$$g(z, v) = f(x, y) |J| = f(zv, v) |v|$$

则 $Z = \frac{X}{Y}$ 的密度函数

$$f_Z(z) = \int_{-\infty}^{+\infty} g(z, v) dv = \int_{-\infty}^{+\infty} f(zv, v) |v| dv.$$

注: 由 Z 的分布函数出发亦可推导出上述公式, 只是推导过程稍微复杂些.

【3-27】 有一射手向某靶射击, 设靶心坐标为原点, 而弹着点坐标为 (X, Y) , 且服从正态分布, 密度函数

$$f(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}.$$

求弹着点与靶心距离 $Z = \sqrt{X^2 + Y^2}$ 的分布函数和密度函数.

解 由于 z 是非负的, 故当 $z < 0$ 时, 分布函数 $F_Z(z) = 0$.

当 $z > 0$ 时, 分布函数

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P(\sqrt{X^2 + Y^2} \leq z) \\ &= \iint_{\sqrt{x^2+y^2} \leq z} f(x, y) dx dy = \iint_{\sqrt{x^2+y^2} \leq z} \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} dx dy, \end{aligned}$$

采用极坐标计算二重积分, 求 $x = r \cos \theta, y = r \sin \theta$, 则雅可比行列式

$$J = \begin{vmatrix} \cos \theta & \sin \theta \\ -r \sin \theta & r \cos \theta \end{vmatrix} = r, \quad dx dy = r dr d\theta. \text{ 于是}$$

$$\begin{aligned} F_Z(z) &= \int_0^{2\pi} d\theta \int_0^z \frac{1}{2\pi\sigma^2} e^{-\frac{r^2}{2\sigma^2}} r dr = \frac{1}{\sigma^2} \int_0^z e^{-\frac{r^2}{2\sigma^2}} r dr \\ &= 1 - e^{-\frac{z^2}{2\sigma^2}}. \end{aligned}$$

所求分布函数

$$F_Z(z) = \begin{cases} 1 - e^{-\frac{z^2}{2\sigma^2}} & (z \geq 0), \\ 0 & (z < 0); \end{cases}$$

密度函数

$$f_Z(z) = F'_Z(z) = \begin{cases} \frac{z}{\sigma^2} e^{-\frac{z^2}{2\sigma^2}} & (z \geq 0), \\ 0 & (z < 0). \end{cases}$$

注: 称 Z 服从参数为 $\sigma(>0)$ 的瑞利(Rayleigh)分布.

【3-28】 设随机变量 X, Y 相互独立, 且都服从 $[0, 1]$ 区间上的均匀分布. 求:

- (1) $Z = \max(X, Y)$ 的密度函数 $f_Z(z)$;
- (2) $W = \min(X, Y)$ 的密度函数 $f_W(w)$.

解 X, Y 的分布函数和密度函数均为

$$\begin{aligned} F(x) &= \begin{cases} 0 & (x < 0), \\ x & (0 \leq x \leq 1), \\ 1 & (x > 1), \end{cases} \\ f(x) &= \begin{cases} 1 & (0 \leq x \leq 1), \\ 0 & (\text{其他}). \end{cases} \end{aligned}$$

(1) $Z = \max(X, Y)$ 的分布函数 $F_Z(z) = [F(z)]^2$, 于是 Z 的密度函数

$$f_Z(z) = F'_Z(z) = 2F(z)f(z) = \begin{cases} 2z & (0 \leq z \leq 1), \\ 0 & (\text{其他}). \end{cases}$$

(2) $W = \min(X, Y)$ 的分布函数 $F_W(w) = 1 - [1 - F(w)]^2$, 于是 W 的密度函数

$$f_W(w) = F'_W(w) = 2[1 - F(w)]f(w) = \begin{cases} 2(1-w) & (0 \leq w \leq 1), \\ 0 & (\text{其他}). \end{cases}$$

【3-29】 设随机变量 (X, Y) 的联合分布律如下表所示:

$\begin{matrix} Y \\ X \end{matrix}$	-1	0	1
1	0.1	0.1	0.2
2	0.3	0	0.3

试求: (1) $P(Y=0|X=1)$;

(2) 在 $Y=-1$ 的条件下, X 的条件分布律;

(3) $Z=2X-Y$ 的分布律;

(4) $U=\max(X, Y)$ 和 $V=\min(X, Y)$ 的分布律.

解 (1) $P(Y=0|X=1) = \frac{P(X=1, Y=0)}{P(X=1)} = \frac{0.1}{0.1+0.1+0.2} = 0.25.$

(2) $P(X=1|Y=-1) = \frac{P(X=1, Y=-1)}{P(Y=-1)} = \frac{0.1}{0.1+0.3} = 0.25,$

$P(X=2|Y=-1) = \frac{P(X=2, Y=-1)}{P(Y=-1)} = \frac{0.3}{0.1+0.3} = 0.75,$

故所求条件分布律如下表所示:

$X Y=-1$	1	2
p_k	0.25	0.75

(3)

(X, Y)	(1, -1)	(1, 0)	(1, 1)	(2, -1)	(2, 0)	(2, 1)
$Z=2X-Y$	3	2	1	5	4	3
p_{ij}	0.1	0.1	0.2	0.3	0	0.3

所求分布律

$$Z \sim \begin{pmatrix} 1 & 2 & 3 & 5 \\ 0.2 & 0.1 & 0.4 & 0.3 \end{pmatrix}.$$

(4) $P(U=1) = P(X=1, Y=-1) + P(X=1, Y=0) + P(X=1, Y=1)$

$$=1)=0.1+0.1+0.2=0.4,$$

$$\begin{aligned} P(U=2) &= P(X=2, Y=-1) + P(X=2, Y=0) + P(X=2, Y \\ &=1)=0.3+0+0.3=0.6, \end{aligned}$$

所以 $U=\max(X, Y)$ 的分布律

$$U \sim \begin{pmatrix} 1 & 2 \\ 0.4 & 0.6 \end{pmatrix}.$$

$$\begin{aligned} P(V=-1) &= P(X=1, Y=-1) + P(X=2, Y=-1) \\ &= 0.1+0.3=0.4, \end{aligned}$$

$$\begin{aligned} P(V=0) &= P(X=1, Y=0) + P(X=2, Y=0) \\ &= 0.1+0=0.1, \end{aligned}$$

$$\begin{aligned} P(V=1) &= P(X=1, Y=1) + P(X=2, Y=1) \\ &= 0.2+0.3=0.5, \end{aligned}$$

所以 $V=\min(X, Y)$ 的分布律

$$V \sim \begin{pmatrix} -1 & 0 & 1 \\ 0.4 & 0.1 & 0.5 \end{pmatrix}.$$

【3-30】 设 $X_1, X_2, \dots, X_n$ 相互独立, 且都服从(0-1)分布,

$$P(X_i = k) = p^k q^{1-k} \quad (k = 0, 1, 0 < p = 1 - q < 1; i = 1, 2, \dots, n),$$

证明: $Z_n = \sum_{i=1}^n X_i$ 服从二项分布 $B(n, p)$.

分析 由于 n 是自然数, 故可用数学归纳法证明.

证 当 $m=1$ 时, (0-1)分布即 $B(1, p)$ 分布.

假设当 $m=n-1$ 时, 结论成立, 即

$$Z_{n-1} = \sum_{i=1}^{n-1} X_i \sim B(n-1, p) \text{ 为真.}$$

下面证明 $m=n$ 时, 结论也成立.

$$Z_n = \sum_{i=1}^{n-1} X_i + X_n = Z_{n-1} + X_n,$$

其中 $Z_{n-1} \sim B(n-1, p), X_n \sim B(1, p)$, 且 Z_{n-1} 与 X_n 相互独立.

当 $j=0$ 时

$$P(Z_n = 0) = P(Z_{n-1} = 0, X_n = 0) = P(Z_{n-1} = 0)P(X_n = 0)$$

$$= q^{r-1}q = q^n.$$

当 $j=1, 2, \dots, n$ 时

$$\begin{aligned} P(Z_n = j) &= P((Z_{n-1} = j, X_n = 0) \cup (Z_{n-1} = j-1, X_n = 1)) \\ &= P(Z_{n-1} = j)P(X_n = 0) + P(Z_{n-1} = j-1)P(X_n = 1) \\ &= C_{n-1}^j p^j q^{n-1-j} + C_{n-1}^{j-1} p^{j-1} q^{n-1-(j-1)} p \\ &= (C_{n-1}^j + C_{n-1}^{j-1}) p^j q^{n-j} \\ &= C_n^j p^j q^{n-j}. \end{aligned}$$

故当 $m=n$ 时, $Z_n = \sum_{i=1}^n X_i \sim B(n, p)$.

【3-31】 设随机变量 X 与 Y 相互独立, (X, Y) 的联合分布律如下表所示:

X \ Y	Y	
	y_1	y_2
x_1	p_{11}	p_{12}
x_2	p_{21}	p_{22}

证明: $p_{11} p_{22} = p_{12} p_{21}$.

证 方法 1 (利用定义)

因为 X 与 Y 相互独立, 所以 $p_{ij} = p_{i\cdot} p_{\cdot j}$ ($i, j=1, 2$). 由 (X, Y) 的联合分布律得 X 和 Y 的边缘分布律

$$X \sim \begin{pmatrix} x_1 & x_2 \\ p_{11} + p_{12} & p_{21} + p_{22} \end{pmatrix}, \quad Y \sim \begin{pmatrix} y_1 & y_2 \\ p_{11} + p_{21} & p_{12} + p_{22} \end{pmatrix},$$

于是

$$\begin{aligned} p_{11} p_{12} &= [(p_{11} + p_{12})(p_{11} + p_{21})][(p_{21} + p_{22})(p_{12} + p_{22})] \\ &= [(p_{11} + p_{12})(p_{12} + p_{22})][(p_{21} + p_{22})(p_{11} + p_{22})] = p_{12} p_{21}. \end{aligned}$$

方法 2 (利用命题)

X 与 Y 相互独立 $\Leftrightarrow$ 联合分布矩阵秩为 1

$\Rightarrow$ 联合分布矩阵两列向量对应分量成比例

$$\Rightarrow \frac{p_{11}}{p_{21}} = \frac{p_{12}}{p_{22}} \Rightarrow p_{11} p_{22} = p_{12} p_{21}.$$

点评 方法 2 较方法 1 简便. 有人认为方法 2 有问题, 理由是若 $p_{21}=0$, 则 p_{21} 不能作分母, 上述等式便无意义. 事实上在 X 与 Y 相互独立的条件下, $p_{ij} (i, j=1, 2)$ 都不为零. 若有一个 p_{ij} 为零, 则可推出其余三个都为零, 而这是不可能的.

【3-32】 设随机变量 X 与 Y 相互独立, 且都服从 $[0, 1]$ 上的均匀分布, 求 $Z=|X-Y|$ 的分布函数和密度函数.

解 由题设可知 X 和 Y 的密度函数

$$f_X(x) = \begin{cases} 1 & (0 \leq x \leq 1), \\ 0 & (\text{其他}); \end{cases}$$

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

因 X 与 Y 相互独立, 故 (X, Y) 的联合密度

$$f(x, y) = f_X(x)f_Y(y) = \begin{cases} 1 & (0 \leq x \leq 1, 0 \leq y \leq 1), \\ 0 & (\text{其他}). \end{cases}$$

当 $z < 0$ 时

$$F_Z(z) = P(Z \leq z) = P(|X-Y| \leq z) = P(\emptyset) = 0.$$

当 $0 \leq z < 1$ 时

$$\begin{aligned} F_Z(z) &= P(|X-Y| \leq z) = \iint_{|x-y| \leq z} f(x, y) dx dy \\ &= \iint_{G \cap D} dx dy = 1 - (1-z)^2. \end{aligned}$$

当 $z \geq 1$ 时;

$$F_Z(z) = P(|X-Y| \leq z) = P(\Omega) = 1.$$

于是 $Z=|X-Y|$ 的分布函数

$$F_Z(z) = \begin{cases} 0 & (z < 0), \\ 1 - (1-z)^2 & (0 \leq z < 1), \\ 1 & (z \geq 1). \end{cases}$$

Z 的密度函数

$$f_Z(z) = F'_Z(z) = \begin{cases} 2(1-z) & (0 \leq z < 1), \\ 0 & (\text{其他}). \end{cases}$$

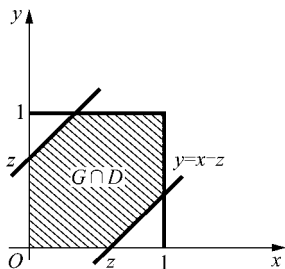


图 3-11

B 习题精选

【3-33】 设随机变量 (X, Y) 的联合分布函数为 $F(x, y)$, 则 $P(X > a, Y > b) = \underline{\hspace{2cm}}$.

- (A) $1 - F(a, b)$;
(B) $F(a, +\infty) + F(+\infty, b)$;
(C) $F(a, b) + 1 - F(a, +\infty) - F(+\infty, b)$;
(D) $F(a, b) - 1 + F(a, +\infty) - F(+\infty, b)$.

【3-34】 用随机变量 (X, Y) 的分布函数 $F(x, y)$ 表示如下概率:

- (1) $P(X > 1) = \underline{\hspace{2cm}}$;
(2) $P(X = 1, Y \leq 3) = \underline{\hspace{2cm}}$;
(3) $P(1 \leq X \leq 2, Y \leq 3) = \underline{\hspace{2cm}}$.

【3-35】 设 (X, Y) 的联合分布函数为 $F(x, y)$, 则其边缘分布函数 $F_X(x) = \underline{\hspace{2cm}}$.

- (A) $\lim_{x \rightarrow +\infty} F(x, y)$; (B) $\lim_{y \rightarrow +\infty} F(x, y)$;
(C) $F(0, y)$; (D) $F(x, 0)$.

【3-36】 设随机变量 (X, Y) 只取下列数组中的值:

$$(-1, 0), (0, 1), (2, 0), (2, 1),$$

且取这些值的相应概率依次为 $\frac{1}{c}, \frac{1}{2c}, \frac{1}{4c}, \frac{5}{4c}$, 则 $c = \underline{\hspace{2cm}}$.

- (A) 2; (B) 3; (C) 4; (D) 5.

【3-37】 设二元函数

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

问: $F(x, y)$ 能否成为某二维随机变量 (X, Y) 的分布函数? 为什么?

【3-38】 设随机变量 (X, Y) 在区域 $G = \{(x, y) | 0 \leq x \leq 1, 0 \leq y \leq 2\}$ 上服从均匀分布, 则 $P(Y > X^2) = \underline{\hspace{2cm}}$.

- (A) $\frac{1}{6}$; (B) $\frac{1}{3}$; (C) $\frac{2}{3}$; (D) $\frac{5}{6}$.

【3-39】 设随机变量 (X, Y) 的联合密度函数

$$f(x, y) = \begin{cases} A(x^2 + y^2) & (x^2 + y^2 \leq 1), \\ 0 & (\text{其他}), \end{cases}$$

则 $A = \underline{\hspace{2cm}}, P\left(X^2 + Y^2 < \frac{1}{9}\right) = \underline{\hspace{2cm}}$.

【3-40】 将 A, B 两枚硬币各投掷一次, 以 X 表示 A 币得到的正面数, 以 Y 表示 A, B 两枚硬币得到的正面总数, 求 (X, Y) 的联合分布律及边缘分布律.

【3-41】 设 $X \sim B(1, 0.6)$. 当 $X=0$ 时, $Y \sim \begin{pmatrix} 0 & 1 & 2 \\ 0.3 & 0.2 & 0.5 \end{pmatrix}$; 当

$X=1$ 时, $Y \sim \begin{pmatrix} 0 & 1 & 2 \\ 0.4 & 0.2 & 0.4 \end{pmatrix}$. 求:

- (1) (X, Y) 的联合分布律及边缘分布律;
- (2) $P(X > 0, Y > 1), P(X - Y \geq 0), P(X + Y \leq 2)$;
- (3) 当 $Y \neq 2$ 时关于 X 的条件分布.

【3-42】 设 (X, Y) 的联合密度函数

$$f(x, y) = \begin{cases} Ae^{-2x-3y} & (x > 0, y > 0), \\ 0 & (\text{其他}). \end{cases}$$

求:

- (1) 常数 A ;
- (2) $P(X < 1, Y < 2)$;
- (3) $f_X(x), f_Y(y)$;
- (4) $P(X + Y < 3)$;
- (5) $f_{X|Y}(x|y)$;
- (6) $P(X < 1|Y < 2)$.

【3-43】 设 (X, Y) 服从区域 $G = \{(x, y) | 0 \leq y \leq 1 - x^2\}$ 上的均匀分布, 设区域 $D = \{(x, y) | y \geq x^2\}$. 求:

- (1) (X, Y) 的联合密度函数;
- (2) X 和 Y 的边缘密度函数;
- (3) $P((X, Y) \in D)$;
- (4) $X = \frac{1}{2}$ 时 Y 的条件密度函数.

【3-44】 设随机变量 X 与 Y 相互独立, 且 $X \sim \begin{pmatrix} 0 & 1 \\ 0.2 & 0.8 \end{pmatrix}, Y \sim$

$\begin{pmatrix} 0 & 1 \\ 0.2 & 0.8 \end{pmatrix}$, 则必有_____.

- (A) $X=Y$; (B) $P(X=Y)=0$;
(C) $P(X=Y)=0.68$; (D) $P(X=Y)=1$.

【3-45】 设两个相互独立的随机变量 X 与 Y 分别服从正态分布 $N(1,2)$ 与 $N(-1,2)$, 则_____.

- (A) $P(X-Y \leq 0) = \frac{1}{2}$; (B) $P(X-Y \leq 1) = \frac{1}{2}$;
(C) $P(X+Y \leq 0) = \frac{1}{2}$; (D) $P(X+Y \leq 1) = \frac{1}{2}$.

【3-46】 设随机变量 X 与 Y 相互独立, 且都服从 $(0,1)$ 区间上的均匀分布, 则服从相应区间或区域上的均匀分布的随机变量是_____.

- (A) X^2 ; (B) $X-Y$;
(C) $X+Y$; (D) (X,Y) .

【3-47】 随机变量关于 X 和 Y 的边缘密度可以由它们的联合密度确定, 联合分布_____由边缘分布确定.

- (A) 不能; (B) 为正态分布时可以;
(C) 也可; (D) 在 X 与 Y 相互独立时可以.

【3-48】 设离散随机变量 (X,Y) 的联合分布律如下表所示:

X \ Y	1	2	3
1	0.08	a	0.12
2	b	c	0.18

且 X 与 Y 相互独立, 则 $a=$ _____, $b=$ _____, $c=$ _____.

【3-49】 设随机变量 X 与 Y 相互独立且同分布, X 的密度函数

$$f(x) = \begin{cases} \frac{3}{8}x^2 & (0 \leq x \leq 2), \\ 0 & (\text{其他}). \end{cases}$$

若 $P((X>a) \cup (Y>a)) = \frac{3}{4}$, 则 $a=$ _____.

【3-50】 设随机变量 $X \sim \begin{pmatrix} 0 & 1 \\ 0.5 & 0.5 \end{pmatrix}$, $Y \sim \begin{pmatrix} -1 & 0 & 1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix}$,

且 $P(XY=0)=1$.

(1) 求 X 和 Y 的联合分布律;

(2) 问 X 和 Y 是否相互独立? 为什么?

【3-51】 一电脑专卖店每周五进货, 以备周六、日 2 天销售, 根据以往统计, 这 2 天销售台数 X 、 Y 相互独立, 且

$$X \sim \begin{pmatrix} 7 & 8 & 9 \\ 0.2 & 0.6 & 0.2 \end{pmatrix}, \quad Y \sim \begin{pmatrix} 10 & 11 & 12 \\ 0.3 & 0.6 & 0.1 \end{pmatrix}.$$

求进货 20 台不够卖的概率以及进货 18 台够卖的概率.

【3-52】 设二维随机变量 (X, Y) 在单位圆域 $G = \{(x, y) | x^2 + y^2 < r^2\}$ 内服从均匀分布.

(1) 求 (X, Y) 的联合密度;

(2) (X, Y) 的边缘分布还是均匀分布吗?

(3) (X, Y) 的条件分布是否均匀分布?

(4) X 与 Y 是否相互独立?

【3-53】 设二维随机变量 (X, Y) 在区域 $G = \{(x, y) | x > 0, 0 < y < 2 - 2x\}$ 上服从均匀分布. 试验证:

(1) (X, Y) 的边缘分布不是均匀分布;

(2) (X, Y) 的条件分布仍是均匀分布;

(3) X 与 Y 不相互独立.

【3-54】 设随机变量 (X, Y) 的分布函数

$$F(x, y) = \begin{cases} 0 & (x < 0 \text{ 或 } y < 0), \\ xy & (0 \leq x < 1, 0 \leq y < 1), \\ x & (0 \leq x < 1, y \geq 1), \\ y & (x \geq 1, 0 \leq y < 1), \\ 1 & (x \geq 1, y \geq 1). \end{cases}$$

(1) 求 (X, Y) 的联合密度函数 $f(x, y)$ 和边缘分布函数 $F_X(x), F_Y(y)$;

(2) X 与 Y 是否相互独立?

【3-55】 设 (X, Y) 的联合分布函数

$$F(x, y) = G(x)[H(y) - H(-\infty)]$$

且 $G(+\infty), H(+\infty), H(-\infty)$ 都存在, 试证: X 与 Y 相互独立.

【3-56】 设随机变量 X, Y 相互独立, 且 $X \sim B(1, p), Y \sim \pi(\lambda)$, 则 $X+Y$ _____.

- (A) 服从两点分布; (B) 服从泊松分布;
(C) 为二维随机变量; (D) 仍为一维随机变量.

【3-57】 设随机变量 $X \sim N(0, 1), Y \sim N(0, 1)$, 则 $X+Y$ _____.

- (A) 不一定服从正态分布;
(B) 服从正态分布 $N(0, 2)$;
(C) 服从正态分布 $N(0, \sqrt{2})$;
(D) 服从正态分布 $N(0, 1)$.

【3-58】 设随机变量 X, Y 满足 $P(X \leq 1, Y \leq 1) = \frac{3}{7}, P(X \leq 1) = P(Y \leq 1) = \frac{4}{7}$, 则 $P(\min(X, Y) \leq 1) =$ _____.

- (A) $\frac{5}{7}$; (B) $\frac{3}{7}$; (C) $\frac{2}{7}$; (D) $\frac{1}{7}$.

【3-59】 设随机变量 (X, Y) 的联合分布律如下表所示:

X \ Y	Y		
	-1	0	2
1	0.25	0.1	0.3
2	0.15	0.05	0.15

则 $Z = 2X - Y$ 的分布律为 _____; $W = XY$ 的分布律为 _____.

【3-60】 设随机变量 X, Y 相互独立, 且都服从参数为 $p = 0.5$ 的 $(0-1)$ 分布, 则 $Z = \max(X, Y)$ 的分布律为 _____.

【3-61】 设随机变量 X, Y 相互独立, 且都服从参数 $\lambda = 1$ 的指数分布, 则 $Z = \max(X, Y)$ 的密度函数 $f_Z(z) =$ _____.

【3-62】 设二维随机变量 (X, Y) 的密度函数

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

求 $Z=2X-Y$ 的密度函数 $f_Z(z)$.

【3-63】 设随机变量 X, Y 相互独立, 其密度函数分别为

$$f_X(x) = \begin{cases} \frac{1}{3} & (-1 \leq x \leq 2), \\ 0 & (\text{其他}); \end{cases}$$
$$f_Y(y) = \begin{cases} 2y & (0 \leq y \leq 1), \\ 0 & (\text{其他}). \end{cases}$$

求 $Z=X+Y$ 的密度函数 $f_Z(z)$.

【3-64】 设随机变量 (X, Y) 的联合密度函数

$$f(x, y) = \begin{cases} \frac{1}{2} & (0 \leq x \leq 2, 0 \leq y \leq 1), \\ 0 & (\text{其他}). \end{cases}$$

求边长为 X 和 Y 的矩形面积 S 的分布函数 $F(s)$.

【3-65】 在区间 $(0, 1)$ 上任意掷两个点 X 与 Y , 求两点间距离 $Z=|X-Y|$ 的分布函数 $F(z)$.

答案与提示

【3-33】 选 C. 根据二元分布函数的几何意义.

【3-34】 (1) $1-F(1, +\infty)$; (2) $F(1, 3)-F(1-0, 3)$;

(3) $F(2, 3)-F(1-0, 3)$, 其中 $F(1-0, 3)=\lim_{x \rightarrow 1^-} F(x, 3)$.

【3-35】 选 B. $F_X(x)=P(X \leq x)=P(X \leq x, Y \leq +\infty)=F(x, +\infty)$.

【3-36】 选 B. 利用概率的归一性.

【3-37】 否. 因为 $F(+\infty, +\infty)=1$ 不满足.

【3-38】 选 D. $P(Y > X^2) = \iint_{y > x^2} f(x, y) dx dy = \int_0^1 dx \int_{x^2}^2 \frac{1}{2} dy$.

【3-39】 $\frac{2}{\pi}, \frac{1}{81}$.

【3-40】

X \ Y	0	1	2
	0	1	2
0	$\frac{1}{4}$	$\frac{1}{4}$	0
1	0	$\frac{1}{4}$	$\frac{1}{4}$

$$X \sim \begin{pmatrix} 0 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad Y \sim \begin{pmatrix} 0 & 1 & 2 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}.$$

【3-41】 (1)

X \ Y	0	1	2
	0	1	2
0	0.12	0.08	0.2
1	0.24	0.12	0.24

$$X \sim \begin{pmatrix} 0 & 1 \\ 0.4 & 0.6 \end{pmatrix}, \quad Y \sim \begin{pmatrix} 0 & 1 & 2 \\ 0.36 & 0.2 & 0.44 \end{pmatrix}.$$

提示:利用乘法公式求联合概率

$$p_{ij} = P(X=i, Y=j) = P(X=i)P(Y=j | X=i).$$

(2) 0.24, 0.48, 0.76;

(3)

X Y≠2	0	1
p	$\frac{5}{14}$	$\frac{9}{14}$

提示:利用条件概率公式

$$P(X=i | Y \neq 2) = \frac{P(X=i, Y \neq 2)}{P(Y \neq 2)} \quad (i=0,1).$$

【3-42】 (1) $A=6$;

(2) $(1-e^{-2})(1-e^{-6})$;

$$(3) f_X(x) = \begin{cases} 2e^{-2x} & (x > 0), \\ 0 & (x \leq 0), \end{cases} \quad f_Y(y) = \begin{cases} 3e^{-3y} & (y > 0), \\ 0 & (y \leq 0); \end{cases}$$

$$(4) 1 - 3e^{-6} + 2e^{-9};$$

$$(5) \text{ 当 } y > 0 \text{ 时, } f_{X|Y}(x|y) = \begin{cases} 2e^{-2x} & (x > 0), \\ 0 & (x \leq 0); \end{cases}$$

$$(6) 1 - e^{-2}.$$

提示:题(5)、(6)两题可利用 X, Y 的独立性直接求得,即

$$f_{X|Y}(x|y) = f_X(x), \quad P(X < 1 | Y < 2) = P(X < 1).$$

$$\textbf{【3-43】} \quad (1) f(x, y) = \begin{cases} \frac{3}{4} & (0 \leq y \leq 1 - x^2), \\ 0 & (\text{其他}); \end{cases}$$

$$(2) f_X(x) = \begin{cases} \frac{3}{4}(1 - x^2) & (|x| < 1), \\ 0 & (|x| \geq 1), \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{3}{2}\sqrt{1 - y^2} & (0 < y < 1), \\ 0 & (\text{其他}); \end{cases}$$

$$(3) P((X, Y) \in D) = \iint_{D \cap G} \frac{3}{4} dx dy = \int_{-\frac{1}{\sqrt{2}}}^{\frac{1}{\sqrt{2}}} dx \int_{x^2}^{1-x^2} \frac{3}{4} dy = \frac{\sqrt{2}}{2};$$

$$(4) f_{Y|X}\left(y \middle| \frac{1}{2}\right) = \frac{f\left(\frac{1}{2}, y\right)}{f_X\left(\frac{1}{2}\right)} = \begin{cases} \frac{4}{3} & (0 < y < \frac{3}{4}), \\ 0 & (\text{其他}). \end{cases}$$

【3-44】 选 C. 提示:先利用独立性求出 (X, Y) 的联合分布律,再计算 $P(X=Y) = P(X=0, Y=0) + P(X=1, Y=1)$.

【3-45】 选 C. 提示:利用正态分布的对称性,并注意到

$$X + Y \sim N(0, 4).$$

【3-46】 选 D. 因为当 X, Y 独立时,二维随机变量 (X, Y) 的密度函数 $f(x, y) = f_X(x)f_Y(y) = \begin{cases} 1 & (0 < x < 1, 0 < y < 1), \\ 0 & (\text{其他}), \end{cases}$ 即

$$(X, Y) \sim U[0, 1; 0, 1]$$

【3-47】 选 D. 当 X, Y 独立时, 联合分布必可由边缘分布确定.

【3-48】 $a=0.2, b=0.12, c=0.3$.

提示: 参阅题 3-20 方法 2.

【3-49】 $\sqrt[3]{4}$. 提示: 设事件 $(X > a) = A$, 则有 $1 - [P(\bar{A})]^2 = \frac{3}{4}$,

其中

$$P(A) = \int_0^a \frac{3}{8} x^2 dx = \frac{a^3}{8}.$$

【3-50】 (1)

$X \backslash Y$	-1	0	1
0	0.25	0	0.25
1	0	0.5	0

提示: 由 $P(XY=0)=1$ 可得 $P(X=1, Y=-1)=P(X=1, Y=1)=0$.

(2) X 与 Y 不独立. 因为 $P(X=0, Y=0)=0 \neq 0.25=P(X=0) \cdot P(Y=0)$.

【3-51】 $P(X+Y > 20) = P(X=9, Y=12) = P(X=9)P(Y=12) = 0.02$;

$P(X+Y \leq 18) = P(X=7, Y=10) + P(X=7, Y=11) + P(X=8, Y=10) = 0.54$.

【3-52】 (1) $f(x, y) = \begin{cases} \frac{1}{\pi r^2} & (x^2 + y^2 < r^2), \\ 0 & (x^2 + y^2 \geq r^2). \end{cases}$

(2) 边缘分布不再是均匀分布. 事实上

$$f_X(x) = \begin{cases} \frac{2}{\pi r^2} \sqrt{r^2 - x^2} & (|x| < r), \\ 0 & (|x| \geq r); \end{cases}$$
$$f_Y(y) = \begin{cases} \frac{2}{\pi r^2} \sqrt{r^2 - y^2} & (|y| < r), \\ 0 & (|y| \geq r). \end{cases}$$

(3) 条件分布是均匀分布. 事实上

$$\text{当 } |x| < r \text{ 时, } f_{Y|X}(y|x) = \begin{cases} \frac{1}{2\sqrt{r^2-x^2}} & (|y| < \sqrt{r^2-x^2}), \\ 0 & (|y| \geq \sqrt{r^2-x^2}); \end{cases}$$

$$\text{当 } |y| < r \text{ 时, } f_{X|Y}(x|y) = \begin{cases} \frac{1}{2\sqrt{r^2-y^2}} & (|x| < \sqrt{r^2-y^2}), \\ 0 & (|x| \geq \sqrt{r^2-y^2}). \end{cases}$$

(4) X 与 Y 不相互独立. 事实上

$$f(0,0) = \frac{1}{\pi r^2} \neq \frac{4}{\pi r^2} = f_X(0)f_Y(0).$$

【3-53】 (1) $f_X(x) = \begin{cases} 2-2x & (0 < x < 1), \\ 0 & (\text{其他}), \end{cases}$

$$f_Y(y) = \begin{cases} 1-\frac{y}{2} & (0 < y < 2), \\ 0 & (\text{其他}), \end{cases}$$

故均不是均匀分布.

$$(2) \text{ 当 } 0 < x < 1 \text{ 时, } f_{Y|X}(y|x) = \begin{cases} \frac{1}{2-2x} & (0 < y < 2-2x), \\ 0 & (\text{其他}), \end{cases}$$

$$\text{当 } 0 < y < 2 \text{ 时, } f_{X|Y}(x|y) = \begin{cases} \frac{2}{2-y} & (0 < x < 1-\frac{y}{2}), \\ 0 & (\text{其他}), \end{cases}$$

故条件分布仍是均匀分布.

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

$$f\left(\frac{1}{2}, \frac{1}{2}\right) = 1 \neq \frac{3}{4} = f_X\left(\frac{1}{2}\right)f_Y\left(\frac{1}{2}\right), X \text{ 与 } Y \text{ 不相互独立.}$$

【3-54】 (1) $f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = \begin{cases} 1 & (0 < x < 1, 0 < y < 1), \\ 0 & (\text{其他}); \end{cases}$

$$F_X(x) = F(x, +\infty) = \begin{cases} 0 & (x < 0), \\ x & (0 \leq x < 1), \\ 1 & (x \geq 1), \end{cases}$$

$$F_Y(y) = F(+\infty, y) = \begin{cases} 0 & (y < 0), \\ y & (0 \leq y < 1), \\ 1 & (y \geq 1); \end{cases}$$

(2) $F(x, y) = F_X(x)F_Y(y)$, X 与 Y 相互独立.

【3-55】 提示: 由 $F(+\infty, +\infty) = 1$ 可得 $G(+\infty)[H(+\infty) - H(-\infty)] = 1$, 然后推出 $F_X(x)F_Y(y) = F(x, y)$.

【3-56】 选 D. n 维随机变量的函数 $Z = f(X_1, \dots, X_n)$ 自然是一维随机变量.

【3-57】 选 A. 题中无独立条件, 故不一定服从正态分布.

【3-58】 选 A. 借助图示及概率的归一性.

【3-59】 $Z = 2X - Y \sim \begin{pmatrix} 0 & 2 & 3 & 4 & 5 \\ 0.3 & 0.25 & 0.25 & 0.05 & 0.15 \end{pmatrix},$

$$W = XY \sim \begin{pmatrix} -2 & -1 & 0 & 2 & 4 \\ 0.15 & 0.25 & 0.15 & 0.3 & 0.15 \end{pmatrix}.$$

【3-60】 $Z = \max(X, Y) \sim \begin{pmatrix} 0 & 1 \\ 0.25 & 0.75 \end{pmatrix}.$

【3-61】 $f_Z(z) = \begin{cases} 2e^{-z}(1-e^{-z}) & (z > 0), \\ 0 & (z \leq 0). \end{cases}$

【3-62】 $f_Z(z) = \begin{cases} 1 - \frac{z}{2} & (0 < z < 2), \\ 0 & (\text{其他}). \end{cases}$

【3-63】 $f_Z(z) = \begin{cases} \frac{2}{3}z + \frac{1}{3}z^2 + \frac{1}{3} & (-1 \leq z < 0), \\ \frac{1}{3} & (0 \leq z < 2), \\ \frac{4}{3}z - \frac{1}{3}z^2 - 1 & (2 \leq z < 3), \\ 0 & (\text{其他}). \end{cases}$

【3-64】 $F(s) = \begin{cases} 0 & (s < 0), \\ \frac{1}{2}(1 + \ln 2 - \ln s)s & (0 \leq s < 2), \\ 1 & (s \geq 2). \end{cases}$

$$\textbf{【3-65】} \quad F_Z(z) = \begin{cases} 0 & (z < 0), \\ 2z - z^2 & (0 \leq z < 1), \\ 1 & (z \geq 1). \end{cases}$$

关键词

数学期望 方差 标准差 协方差 相关系数 矩相关性 标准化随机变量

教学要求

(1) 理解随机变量的数字特征——数学期望、方差、协方差、相关系数的概念.

(2) 熟记常用分布的期望和方差.

(3) 会运用数字特征的基本性质计算具体分布的数字特征.

(4) 会计算随机变量函数的期望 $E(g(X))$, $E(g(x, y))$, 并掌握其实际应用.

(5) 了解切比雪夫不等式的条件和结论, 并会用来估计概率.

(6) 理解随机变量的相关性与独立性之间的关系.

重点

期望和方差的概念、性质及计算; 利用性质计算随机变量函数的期望.

难点

计算随机变量函数的期望(实际上, 随机变量的方差、协方差、 k 阶原点矩、中心矩等的计算都可归结为随机变量函数的期望的计算).

方法提要

(1) 求 X 的数学期望:

$$E(X) = \begin{cases} \sum_{k=1}^{\infty} x_k p_k & (\text{离散型}), \\ \int_{-\infty}^{+\infty} x f(x) dx & (\text{连续型}). \end{cases}$$

当离散型随机变量 X 的分布律不易求时, 可尝试随机变量分解法, 即 $X = \sum_{i=1}^n X_i$, 其中 X_i 的分布律易求, 于是

$$E(X) = \sum_{i=1}^n E(X_i).$$

(2) 求 X 的方差:

$$D(X) = E(X - E(X))^2 = \begin{cases} \sum_{k=1}^{\infty} [x_k - E(X)]^2 p_k & (\text{离散型}), \\ \int_{-\infty}^{+\infty} [x - E(x)]^2 f(x) dx & (\text{连续性}). \end{cases}$$

不管 X 离散还是连续, 常用的方差计算公式为

$$D(X) = E(X^2) - [E(X)]^2.$$

(3) 求随机变量函数的数学期望:

$$E(g(X)) = \begin{cases} \sum_{k=1}^{\infty} g(x_k) p_k & (\text{离散型}), \\ \int_{-\infty}^{+\infty} g(x) f(x) dx & (\text{连续型}); \end{cases}$$

$$E(g(X, Y)) = \begin{cases} \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} g(x_i, y_j) p_{ij} & (\text{离散型}), \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy & (\text{连续型}). \end{cases}$$

(4) 求协方差和相关系数:

$$\text{cov}(X, Y) = E([X - E(X)][Y - E(Y)]) = E(XY) - E(X)E(Y)$$

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sqrt{D(X)} \sqrt{D(Y)}}$$

计算中常用到方差与协方差或方差与相关系数的关系式

$$D(X \pm Y) = \begin{cases} D(X) + D(Y) \pm 2\text{cov}(X, Y), \\ D(X) + D(Y) \pm 2\rho_{XY} \sqrt{D(X)} \sqrt{D(Y)}. \end{cases}$$

(5) 判断有关概念时常用的五条条件等价命题:

X 与 Y 不相关 $\Leftrightarrow \rho_{XY} = 0$

$\Leftrightarrow \text{cov}(X, Y) = 0$

$\Leftrightarrow E(XY) = E(X)E(Y)$

$\Leftrightarrow D(X \pm Y) = D(X) + D(Y).$

本章知识结构



A 例题精解

【4-1】 某商店有 8 台打印机, 其中 2 台为次品, 现从中随机取 3 台, 设 X 为抽取的次品台数, 求 $E(X)$.

解 方法 1 X 的可能取值为 0, 1, 2, 相应的概率

$$P(X=0) = \frac{C_6^3}{C_8^3} = \frac{5}{14}, P(X=1) = \frac{C_2^1 C_6^2}{C_8^3} = \frac{15}{28},$$

$$P(X=2) = \frac{C_2^2 C_6^1}{C_8^3} = \frac{3}{28},$$

$$E(X) = 0 \times \frac{5}{14} + 1 \times \frac{15}{28} + 2 \times \frac{3}{28} = \frac{3}{4}.$$

方法 2 由题意可知 $X \sim B\left(3, \frac{1}{4}\right)$, 所以 $E(X) = 3 \times \frac{1}{4} = \frac{3}{4}$.

点评 求离散型随机变量的数学期望, 一般的方法是先求出随机变量的分布律, 然后按定义 $E(X) = \sum_k x_k p_k$ 计算. 若随机变量服从常见的分布, 则可直接利用该分布的期望公式求之. 方法 2 就是一例: 若 $X \sim B(n, p)$ 则 $E(X) = np$.

【4-2】某种按新配方试制的中成药在 500 名病人中进行临床试验, 有一半人服用, 另一半人未服. 一周后, 有 280 人痊愈, 其中 240 人服了新药. 试用概率统计方法说明新药的疗效.

解 设随机变量 X 表示服过新药的病人的痊愈情况:

$$X = \begin{cases} 1 & (\text{病人痊愈}), \\ 0 & (\text{病人未痊愈}); \end{cases}$$

Y 表示未服过新药的病人的痊愈情况:

$$Y = \begin{cases} 1 & (\text{病人痊愈}), \\ 0 & (\text{病人未痊愈}). \end{cases}$$

$$\text{由题设得 } X \sim \begin{bmatrix} 0 & 1 \\ \frac{1}{25} & \frac{24}{25} \end{bmatrix}, \quad Y \sim \begin{bmatrix} 0 & 1 \\ \frac{21}{25} & \frac{4}{25} \end{bmatrix},$$

$$E(X) = 0 \times \frac{1}{25} + 1 \times \frac{24}{25} = \frac{24}{25},$$

$$E(Y) = 0 \times \frac{21}{25} + 1 \times \frac{4}{25} = \frac{4}{25},$$

因为 $E(X) > E(Y)$, 所以新药疗效显著.

点评 有人以服药与未服药后的痊愈率作比较, 即

$$P(X=1) = \frac{24}{25} > \frac{4}{25} = P(Y=1)$$

来说明新药疗效显著,似乎也是行得通的,但不如用数学期望进行比较来得全面、科学.

【4-3】 一盒零件中有 9 件合格品,3 件次品. 安装设备时,从这盒零件中任取一件. 若取出的是次品不再放回. 求在取得合格品前已取出的次品数的数学期望.

解 设 X 为取得合格品前已取出的次品数,则 X 可能取值为 0, 1, 2, 3, 取这些值的概率分别为

$$P(X=0) = \frac{9}{12} = \frac{3}{4},$$

$$P(X=1) = \frac{3}{12} \times \frac{9}{11} = \frac{9}{44},$$

$$P(X=2) = \frac{3}{12} \times \frac{2}{11} \times \frac{9}{10} = \frac{9}{220},$$

$$P(X=3) = \frac{3}{12} \times \frac{2}{11} \times \frac{1}{10} \times \frac{9}{9} = \frac{1}{220},$$

所以

$$E(X) = 0 \times \frac{3}{4} + 1 \times \frac{9}{44} + 2 \times \frac{9}{220} + 3 \times \frac{1}{220} = 0.3.$$

点评 本题求概率用的是乘法公式,例如

$$\begin{aligned} P(X=2) &= P(\overline{A_1} \overline{A_2} A_3) = P(\overline{A_1})P(\overline{A_2} | \overline{A_1})P(A_3 | \overline{A_1} \overline{A_2}) \\ &= \frac{3}{12} \times \frac{2}{11} \times \frac{9}{10} = \frac{9}{220}. \end{aligned}$$

其中事件 A_i 表示第 i 次取得合格品 ($i=1, 2, 3$).

【4-4】 已知离散型随机变量 X 的可能取值为 $-1, 0, 1$, $E(X) = 0.1$, $E(X^2) = 0.9$. 求 X 的分布律.

解 设 $X \sim \begin{pmatrix} -1 & 0 & 1 \\ p_1 & 1-p_1-p_2 & p_2 \end{pmatrix}$, 由题设可得关于 p_1, p_2 的方程组

$$\begin{cases} E(X) = p_2 - p_1 = 0.1, \\ E(X^2) = p_1 + p_2 = 0.9, \end{cases} \Rightarrow \begin{cases} p_1 = 0.4, \\ p_2 = 0.5, \end{cases}$$

所以 X 的分布律为

$$\begin{pmatrix} -1 & 0 & 1 \\ 0.4 & 0.1 & 0.5 \end{pmatrix}.$$

【4-5】 4个人在一楼进入电梯,该楼共有11层.若每个乘客在任
何一层楼走出电梯的概率相同.求直到电梯中的乘客全走空时,电梯需
停次数的数学期望.

解 方法1 设电梯需停次数为随机变量 X , 则 X 可能取值为 1,
2, 3, 4, 由古典概型得

$$P(X=1) = \frac{C_4^1 C_{10}^1}{10^4} = \frac{1}{1000},$$

$$P(X=2) = \frac{(C_4^1 C_3^3 + C_4^2 C_2^2 + C_4^3 C_1^1) C_{10}^2}{10^4} = \frac{63}{1000},$$

$$P(X=3) = \frac{(C_4^1 C_3^1 C_2^2 + C_4^1 C_3^2 C_1^1 + C_4^2 C_2^1 C_1^1) C_{10}^3}{10^4} = \frac{432}{1000},$$

$$P(X=4) = \frac{C_4^1 C_3^1 C_2^1 C_1^1 C_{10}^4}{10^4} = \frac{504}{1000},$$

$$E(X) = 1 \times \frac{1}{1000} + 2 \times \frac{63}{1000} + 3 \times \frac{432}{1000} + 4 \times \frac{504}{1000} = 3.439.$$

方法2 设 $X_i = \begin{cases} 1 & (\text{电梯在第 } i \text{ 层停}) \\ 0 & (\text{电梯在第 } i \text{ 层不停}) \end{cases} \quad (i=1, 2, \dots, 10),$

X 表示电梯需停次数, 则 $X = \sum_{i=1}^{10} X_i$. 易得 X_i 的分布律

X_i	1	0
p	$1 - \left(1 - \frac{1}{10}\right)^4$	$\left(1 - \frac{1}{10}\right)^4$

$$E(X_i) = 1 - \left(1 - \frac{1}{10}\right)^4 = 0.3439,$$

$$E(X) = \sum_{i=1}^{10} E(X_i) = 10 \times 0.3439 = 3.439.$$

点评 方法1按部就班地计算,费时费力.试想本题若改成10个
人在一楼进入电梯,再用方法1求解就勉为其难了(因为计算量极大).
而用方法2几乎不增加计算量就可得出

$$E(X) = \sum_{i=1}^{10} E(X_i) = 10 \left[1 - \left(1 - \frac{1}{10} \right)^{10} \right] \approx 6.513.$$

本题的难点是求出 X 的分布律,方法 2 引进(0-1)随机变量以及利用数学期望的性质,巧妙地绕过了“求分布律”这道坎,并使得计算极其简单.

可称方法 2 为随机变量分解法.

【4-6】 已知离散型随机变量 X 的分布函数

$$F(x) = \begin{cases} 0 & (x < -2), \\ 0.4 & (-2 \leq x < 0), \\ 0.6 & (0 \leq x < 1), \\ 0.9 & (1 \leq x < 3), \\ 1 & (x \geq 3), \end{cases}$$

求 $E(1-2X)$.

分析 不能直接通过分布函数求数学期望,只能通过公式

$$P(X = x_k) = F(x_k) - F(x_{k-1}) \quad (x_{k-1} < x_k, k = 1, 2, \dots)$$

先求出 X 的分布律,再求 $E(X)$.

解 由题设知随机变量 X 可能取值为 $-2, 0, 1, 3$, 取这些值的概率

$$P(X = -2) = F(-2) - F(-\infty) = 0.4 - 0 = 0.4,$$

$$P(X = 0) = F(0) - F(-2) = 0.6 - 0.4 = 0.2,$$

$$P(X = 1) = F(1) - F(0) = 0.9 - 0.6 = 0.3,$$

$$P(X = 3) = F(3) - F(1) = 1 - 0.9 = 0.1,$$

$$E(X) = -2 \times 0.4 + 0 \times 0.2 + 1 \times 0.3 + 3 \times 0.1 = -0.2.$$

由数学期望的性质可得

$$E(1-2X) = 1 - 2E(X) = 1.4.$$

【4-7】 对球的直径作近似测量,其值均匀分布在 $[a, b]$ 上. 求球的体积的数学期望.

解 设球的直径为随机变量 X , 球的体积为随机因变量 V , 则 $V =$

$\frac{\pi}{6} X^3$, X 的密度函数

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

于是

$$\begin{aligned} E(V) &= \frac{\pi}{6} E(X^3) = \frac{\pi}{6} \int_a^b x^3 \frac{1}{b-a} dx = \frac{\pi}{24(b-a)} x^4 \Big|_a^b \\ &= \frac{\pi}{24} (a+b)(a^2+b^2). \end{aligned}$$

点评 上述解法直接利用了求连续随机变量函数的期望的公式, 从而绕开了求体积 V 的密度函数这一步. 若不绕开的话, 计算将很繁琐.

【4-8】 设随机变量服从以下各分布, 分别求它们的期望.

(1) 拉普拉斯分布

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

(2) 麦克斯韦分布

$$f(x) = \begin{cases} \frac{4x^2}{a^3\sqrt{\pi}} e^{-\frac{x^2}{a^2}} & (x > 0), \\ 0 & (x \leq 0), \end{cases} \quad a > 0;$$

(3) 瑞利分布

$$f(x) = \begin{cases} \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} & (x \geq 0), \\ 0 & (x < 0), \end{cases} \quad \sigma > 0.$$

解 (1) 被积函数 $xf(x) = \frac{x}{2} e^{-|x|}$ 是奇函数, 且积分区间 $(-\infty, +\infty)$ 关于原点对称, 故

$$E(X) = \int_{-\infty}^{+\infty} \frac{x}{2} e^{-|x|} dx = 0;$$

$$(2) E(X) = \int_{-\infty}^{+\infty} xf(x) dx = \frac{4}{a^3\sqrt{\pi}} \int_0^{+\infty} x^3 e^{-\frac{x^2}{a^2}} dx$$

$$\stackrel{t=\frac{x^2}{a^2}}{=} \frac{2a}{\sqrt{\pi}} \int_0^{+\infty} te^{-t} dt = \frac{2a}{\sqrt{\pi}} (-te^{-t} - e^{-t}) \Big|_0^{+\infty} = \frac{2a}{\sqrt{\pi}};$$

$$(3) E(X) = \int_0^{+\infty} \frac{x^2}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} dx \stackrel{t=\frac{x^2}{2\sigma^2}}{=} \sqrt{2}\sigma \int_0^{+\infty} t^{\frac{1}{2}} e^{-t} dt$$

$$= \sqrt{2}\sigma \Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{2}}{2}\sigma \Gamma\left(\frac{1}{2}\right) = \sqrt{\frac{\pi}{2}}\sigma.$$

点评 求连续型随机变量的期望时,常会遇到计算积分的困难,题(2)和题(3)中的积分都是同一类型的积分,前者能用分部积分法,后者分部积分法失效,只能借助一类特殊函数—— Γ 函数及其递推公式求解.

伽马函数 $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt \quad (x > 0).$

递推公式 $\Gamma(x+1) = x\Gamma(x).$

特殊值 $\Gamma(1)=1, \quad \Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}, \quad \Gamma(n+1)=n!$

一般地,遇到形如 $\int_0^{+\infty} x^\alpha e^{-\frac{x^n}{\beta}} dx \quad (\beta > 0)$ 的积分可利用换元法 $t = \frac{x^n}{\beta}$ 化为 Γ 函数进行计算.

【4-9】 设随机变量 X, Y 的密度函数分别如下:

$$f_X(x) = \begin{cases} 2e^{-2x} & (x > 0), \\ 0 & (x \leq 0), \end{cases}$$

$$f_Y(y) = \begin{cases} 4e^{-4y} & (y > 0), \\ 0 & (y \leq 0). \end{cases}$$

求 $E(X+Y), E(2X-3Y^2).$

解 $E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^{+\infty} 2xe^{-2x}dx = \int_0^{+\infty} -x d(e^{-2x})$

$$= -xe^{-2x} \Big|_0^{+\infty} + \int_0^{+\infty} e^{-2x} dx = -\frac{1}{2}e^{-2x} \Big|_0^{+\infty} = \frac{1}{2},$$

相仿可得 $E(Y) = \int_{-\infty}^{+\infty} yf(y)dy = \frac{1}{4}$, 从而

$$E(X+Y) = E(X) + E(Y) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4};$$

因为

$$\begin{aligned} E(Y^2) &= \int_0^{+\infty} 4y^2 e^{-4y} dy = -y^2 e^{-4y} \Big|_0^{+\infty} + \int_0^{+\infty} 2ye^{-4y} dy \\ &= \frac{1}{2} \int_0^{+\infty} 4ye^{-4y} dy = \frac{1}{8}, \end{aligned}$$

所以 $E(2X-3Y^2) = 2E(X) - 3E(Y^2) = 1 - \frac{3}{8} = \frac{5}{8}.$

点评 利用常见分布的期望公式,可避免复杂的积分计算. 例如若 $X \sim E(\lambda)$, 则 $E(X) = \frac{1}{\lambda}.$

由题设可知 $X \sim E(2), Y \sim E(4)$, 则 $E(X) = \frac{1}{2}, E(Y) = \frac{1}{4}$, 于是 $E(X+Y) = E(X) + E(Y) = \frac{3}{4}.$

【4-10】 某人有 n 把钥匙, 其中只有一把能打开门, 从中任取一把试开, 试过的不再重复, 直至把门打开为止. 求试开次数的数学期望和方差.

解 设事件 $A_i = \{\text{第 } i \text{ 次试开打开门}\}, i=1, 2, \dots, n.$ 设试开次数为 X , 则 X 的分布律

$$\begin{aligned} p_i &= P(X=i) = P(\bar{A}_1 \bar{A}_2 \cdots \bar{A}_{i-1} A_i) = P(\bar{A}_1) P(\bar{A}_2 | \bar{A}_1) \cdots \\ &\quad P(\bar{A}_{i-1} | \bar{A}_1 \bar{A}_2 \cdots \bar{A}_{i-2}) P(A_i | \bar{A}_1 \bar{A}_2 \cdots \bar{A}_{i-1}) \\ &= \frac{n-1}{n} \cdot \frac{n-2}{n-1} \cdot \cdots \cdot \frac{n-(i-1)}{n-(i-2)} \cdot \frac{1}{n-(i-1)} \\ &= \frac{1}{n} \quad (i=1, 2, \dots, n), \end{aligned}$$

X 的数学期望

$$E(X) = \sum_{i=1}^n i p_i = \frac{1}{n} \sum_{i=1}^n i = \frac{n+1}{2};$$

$$E(X^2) = \sum_{i=1}^n i^2 p_i = \frac{1}{n} \sum_{i=1}^n i^2 = \frac{1}{6} (n+1)(2n+1),$$

X 的方差

$$D(X) = E(X^2) - [E(X)]^2 = \frac{1}{12} (n^2 - 1).$$

【4-11】 抛掷 6 颗骰子,求出现的点数之和的数学期望和方差.

分析 本题若设 6 颗骰子的点数之和为 X ,则 X 可以取 31 个数: 6, 7, $\dots$, 36, 求 X 的分布律的计算量显然很大. 避免直接求 X 的分布律的好方法是将 X 分解成 6 个服从 6 点分布的随机变量之和.

解 设 $X_i (i=1, 2, \dots, 6)$ 表示第 i 个骰子出现的点数, 则 6 颗骰子出现的点数之和 $X = \sum_{i=1}^6 X_i$. 由

$$X_i \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$$
$$X_i^2 \sim \begin{pmatrix} 1 & 4 & 9 & 16 & 25 & 36 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$$

可得

$$E(X_i) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} \\ = \frac{7}{2},$$

$$E(X_i^2) = 1 \times \frac{1}{6} + 4 \times \frac{1}{6} + 9 \times \frac{1}{6} + 16 \times \frac{1}{6} + 25 \times \frac{1}{6} + 36 \times \frac{1}{6} \\ = \frac{91}{6},$$

$$D(X_i) = E(X_i^2) - [E(X_i)]^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12},$$

所以所求数学期望

$$E(X) = E\left(\sum_{i=1}^6 X_i\right) = \sum_{i=1}^6 E(X_i) = 6 \times \frac{7}{2} = 21.$$

再由 $X_1, X_2, \dots, X_6$ 的相互独立, 得 X 的方差

$$D(X) = D\left(\sum_{i=1}^6 X_i\right) = \sum_{i=1}^6 D(X_i) = 6 \times \frac{35}{12} = \frac{35}{2}.$$

【4-12】 随机变量 X 与 Y 相互独立, 它们的密度函数分别是

$$f_X(x) = \frac{1}{2\sqrt{\pi}} e^{\frac{-x^2+2x-1}{4}} \quad (-\infty < x < +\infty),$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-(0.5y^2+2y+4)} \quad (-\infty < y < +\infty).$$

设随机变量 $Z=2X-Y+8$, 求 Z 的期望与方差.

分析 本题若直接用定义求解, 则需先求出 Z 的密度函数, 然后再计算积分, 整个过程的计算量将会很大. 较简便的方法是先弄清楚 X, Y 服从什么分布及分布的参数, 然后利用期望和方差的性质求之.

解 改写题中给出的密度函数:

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

$$f_Y(y) = \frac{1}{\sqrt{2\pi} \cdot 1} e^{\frac{-(y+2)^2}{2 \times 1^2}} \quad (-\infty < y < +\infty),$$

可知 $X \sim N(1, 2)$, $Y \sim N(-2, 1)$, 即有 $E(X)=1, D(X)=2, E(Y)=-2, D(Y)=1$. 又 X, Y 相互独立, 据期望和方差的性质有

$$E(Z) = E(2X - Y + 8) = 2E(X) - E(Y) + 8 = 12,$$

$$D(Z) = D(2X - Y + 8) = 4D(X) + D(Y) = 9.$$

【4-13】 设随机变量 X 的密度函数

$$f(x) = \begin{cases} ax & (0 < x < 2), \\ bx + c & (2 \leq x \leq 4), \\ 0 & (\text{其他}). \end{cases}$$

已知 $E(X)=2, P(1 < X < 3) = \frac{3}{4}$. 求:

(1) a, b, c 的值;

(2) 随机变量 $Y=e^X$ 的数学期望与方差.

解 (1) 由归一性及题设可有

$$\int_{-\infty}^{+\infty} f(x) dx = \int_0^2 ax dx + \int_2^4 (bx + c) dx = 2a + 6b + 2c = 1,$$

$$E(X) = \int_0^2 ax^2 dx + \int_2^4 (bx^2 + cx) dx = \frac{8}{3}a + \frac{56}{3}b + 6c = 2,$$

$$P(1 < X < 3) = \int_1^2 ax \, dx + \int_2^3 (bx + c) \, dx = \frac{3}{2}a + \frac{5}{2}b + c = \frac{3}{4},$$

联立上面三式可得

$$a = \frac{1}{4}, \quad b = -\frac{1}{4}, \quad c = 1.$$

$$\begin{aligned} (2) \quad E(Y) &= E(e^X) = \int_{-\infty}^{+\infty} e^x f(x) \, dx \\ &= \int_0^2 \frac{x}{4} e^x \, dx + \int_2^4 \left(1 - \frac{x}{4}\right) e^x \, dx \\ &= \frac{1}{4} (e^2 - 1)^2, \end{aligned}$$

$$\begin{aligned} E(Y^2) &= E(e^{2X}) = \int_{-\infty}^{+\infty} e^{2x} f(x) \, dx \\ &= \frac{1}{16} (e^4 - 1)^2, \end{aligned}$$

$$D(Y) = E(Y^2) - [E(Y)]^2 = \frac{1}{4} e^2 (e^2 - 1)^2.$$

点评 常数 a, b, c 还可通过分布函数 $F(x) = \int_{-\infty}^x f(t) \, dt$ 及其性质求得. 事实上,

$$F(X) = \begin{cases} 0 & (x < 0), \\ \frac{a}{2} x^2 & (0 \leq x < 2), \\ \frac{b}{2} x^2 + cx + 2(a - b - c) & (2 \leq x < 4), \\ 1 & (x \geq 4). \end{cases}$$

由 $F(4-0) = F(4)$ 得 $2a + 6b + 2c = 1$, 由 $P(1 < X < 3) = F(3) - F(1) = \frac{3}{4}$ 得 $\frac{3}{2}a + \frac{5}{2}b + c = \frac{3}{4}$.

【4-14】 设 (X, Y) 的联合分布律如下:

Y \ X			
	1	2	3
-1	0.2	0.1	0
0	0.1	0	0.3
1	0.1	0.1	0.1

(1) 设 $Z = \frac{Y}{X}$, 求 $E(Z)$;

(2) 设 $W = X + Y$, 求 $D(W)$.

解

(X,Y)	(1,-1)	(1,0)	(1,1)	(2,-1)	(2,0)	(2,1)	(3,-1)	(3,0)	(3,1)
Z	-1	0	1	$-\frac{1}{2}$	0	$\frac{1}{2}$	$-\frac{1}{3}$	0	$\frac{1}{3}$
W	0	1	2	1	2	3	2	3	4
p	0.2	0.1	0.1	0.1	0	0.1	0	0.3	0.1

取同值的概率相加得分布律

$$Z \sim \begin{pmatrix} -1 & -\frac{1}{2} & -\frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{2} & 1 \\ 0.2 & 0.1 & 0 & 0.4 & 0.1 & 0.1 & 0.1 \end{pmatrix},$$

$$W \sim \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 0.2 & 0.2 & 0.1 & 0.4 & 0.1 \end{pmatrix}.$$

$$\begin{aligned} (1) E(Z) &= -1 \times 0.2 - \frac{1}{2} \times 0.1 - \frac{1}{3} \times 0 + 0 \times 0.4 + \frac{1}{3} \times 0.1 \\ &\quad + \frac{1}{2} \times 0.1 + 1 \times 0.1 = -\frac{1}{15}; \end{aligned}$$

$$\begin{aligned} (2) E(W) &= 0 \times 0.2 + 1 \times 0.2 + 2 \times 0.1 + 3 \times 0.4 + 4 \times 0.1 = 2, \\ E(W^2) &= 0 \times 0.2 + 1 \times 0.2 + 4 \times 0.1 + 9 \times 0.4 + 16 \times 0.1 \\ &= 5.8, \end{aligned}$$

$$D(W) = E(W^2) - [E(W)]^2 = 1.8.$$

【4-15】 将 n 个球 ($1 \sim n$ 号) 随机地放进 n 只盒子 ($1 \sim n$) 中, 每盒

装一个球. 若球装入与球同号的盒子中称为一个配对, 记 X 为配对个数. 求 $E(X), D(X)$.

分析 为避免直接求 X 的分布律(因为它很不容易求), 需先将 X 分解为 n 个服从(0-1)分布的随机变量之和, 然后利用期望性质和方差公式计算之.

解 设 $X_i = \begin{cases} 1 & (i \text{ 号球入 } i \text{ 号盒}), \\ 0 & (\text{其他}), \end{cases} \quad i = 1, 2, \dots, n$, 则 $X =$

$\sum_{i=1}^n X_i$, 显然 $X_i \sim \begin{pmatrix} 0 & 1 \\ 1 - \frac{1}{n} & \frac{1}{n} \end{pmatrix}$, 于是

$$E(X) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) = \sum_{i=1}^n \frac{1}{n} = 1.$$

又

$$X_i^2 \sim \begin{pmatrix} 0 & 1 \\ 1 - \frac{1}{n} & \frac{1}{n} \end{pmatrix}, \quad X_i X_j \sim \begin{pmatrix} 0 & 1 \\ 1 - \frac{1}{n(n-1)} & \frac{1}{n(n-1)} \end{pmatrix} \\ (i, j = 1, 2, \dots, n),$$

可得

$$E(X_i^2) = \frac{1}{n}, \quad E(X_i X_j) = \frac{1}{n(n-1)},$$

$$\begin{aligned} E(X^2) &= E\left(\sum_{i=1}^n X_i\right)^2 = E\left(\sum_{i=1}^n X_i^2 + 2 \sum_{1 \leq i < j \leq n} X_i X_j\right) \\ &= \sum_{i=1}^n E(X_i^2) + 2 \sum_{1 \leq i < j \leq n} E(X_i X_j) \\ &= \sum_{i=1}^n \frac{1}{n} + 2 \sum_{1 \leq i < j \leq n} \frac{1}{n(n-1)} \\ &= 1 + 2C_n^2 \cdot \frac{1}{n(n-1)} = 2, \end{aligned}$$

所以

$$D(X) = E(X^2) - [E(X)]^2 = 1.$$

点评 必须注意, 引进的(0-1)分布的随机变量 $X_1, X_2, \dots, X_n$ 不是相互独立的. 若不然极易犯如下两个错误:

$$\textcircled{1} X_i \sim B\left(1, \frac{1}{n}\right) \xRightarrow{(*)} X = \sum_{i=1}^n X_i \sim B\left(n, \frac{1}{n}\right) \Rightarrow E(X) = n \times \frac{1}{n} = 1, D(X) = n \times \frac{1}{n} \left(1 - \frac{1}{n}\right) = 1 - \frac{1}{n}.$$

$$\textcircled{2} D(X) = D\left(\sum_{i=1}^n X_i\right) \xRightarrow{(*)} \sum_{i=1}^n D(X_i) = \sum_{i=1}^n \{E(X_i^2) - [E(X_i)]^2\} = \sum_{i=1}^n \left(\frac{1}{n} - \frac{1}{n^2}\right) = 1 - \frac{1}{n}. \text{ 其中标 } * \text{ 的为错误之处.}$$

【4-16】 证明:在一次试验中,事件 A 发生的次数 X 的方差满足 $D(X) \leq \frac{1}{4}$.

证 设 $P(A) = p$, 则 $X \sim B(1, p)$, 于是 $D(X) = p(1-p)$

方法 1 (公式法) $ab \leq \left(\frac{a+b}{2}\right)^2, a, b \geq 0$.

取 $a = p, b = 1-p$, 得

$$D(X) = p(1-p) \leq \left(\frac{p+1-p}{2}\right)^2 = \frac{1}{4}.$$

方法 2 (配方法)

$$D(X) = p(1-p) = \frac{1}{4} - \left(\frac{1}{2} - p\right)^2 \leq \frac{1}{4}.$$

方法 3 (微分法)

$$\text{令 } f(p) = p(1-p), f'(p) = 1-2p = 0 \Rightarrow p = \frac{1}{2},$$

$f''(p) = -2 < 0 \Rightarrow f\left(\frac{1}{2}\right)$ 为 $f(p)$ 的最大值, 所以

$$D(X) = f(p) \leq f\left(\frac{1}{2}\right) = \frac{1}{4}.$$

方法 4 (方差性质) $D(X) \leq E(X-c)^2, c$ 为任意常数.

$$\text{取 } c = \frac{1}{2}, E(X-c)^2 = E\left(X - \frac{1}{2}\right)^2 = \left(0 - \frac{1}{2}\right)^2 (1-p) +$$

$$\left(1 - \frac{1}{2}\right)^2 p = \frac{1}{4}, \text{ 所以}$$

$$D(X) \leq E(X-c)^2 = \frac{1}{4}.$$

点评 有人先设 $P(A) = \frac{1}{2}$, 即 $X \sim B\left(1, \frac{1}{2}\right)$, 然后证明 $D(X) \leq \frac{1}{4}$. 这种证法是不可取的, 因为它失去了一般性.

【4-17】 汽车站分别于每小时的 10 分, 30 分和 55 分发车. 若乘客不知发车时间, 在每小时内的任一时刻随机到达车站, 求乘客等候时间的数学期望(精确到秒).

分析 乘客到达车站的时间 T 是个随机变量, 而等候的时间也是随机变量, 且可以视为 T 的函数 $g(T)$. T 服从均匀分布, 通过积分可以求得 $E(g(T))$.

解 设随机变量 T 为乘客到达车站的时间, 则 $T \sim U[0, 60]$ 其密度函数

$$f(t) = \begin{cases} \frac{1}{60} & (0 \leq t \leq 60), \\ 0 & (\text{其他}). \end{cases}$$

乘客需要等候的时间为随机变量 Y , 则

$$Y = g(T) = \begin{cases} 10 - T & (0 \leq T < 10), \\ 30 - T & (10 \leq T < 30), \\ 55 - T & (30 \leq T < 55), \\ 70 - T & (55 \leq T \leq 60), \end{cases}$$

$$\begin{aligned} E(Y) &= \int_{-\infty}^{+\infty} g(t) f(t) dt = \frac{1}{60} \left[\int_0^{10} (10-t) dt + \int_{10}^{30} (30-t) dt \right. \\ &\quad \left. + \int_{30}^{55} (55-t) dt + \int_{55}^{60} (70-t) dt \right] \\ &= \frac{625}{60} = 10 \text{ 分 } 25 \text{ 秒}. \end{aligned}$$

点评 这是一般的大多数人采用的解法. 下面介绍一个特殊的解法.

解 设乘客需等候的时间为 T , 则其密度函数

$$\varphi(t) = \begin{cases} 3a & (0 \leq t < 15), \\ 2a & (15 \leq t < 20), \\ a & (20 \leq t \leq 25), \\ 0 & (\text{其他}), \end{cases} \quad a \text{ 为待定常数.}$$

由 $\int_{-\infty}^{+\infty} \varphi(t) dt = 60a = 1 \Rightarrow a = \frac{1}{60}$, 于是等候时间的数学期望

$$\begin{aligned} E(T) &= \int_{-\infty}^{+\infty} t\varphi(t) dt = \int_0^{15} \frac{t}{20} dt + \int_{15}^{20} \frac{t}{30} dt + \int_{20}^{25} \frac{t}{60} dt \\ &= \frac{625}{60} = 10 \text{ 分 } 25 \text{ 秒.} \end{aligned}$$

【4-18】 某意外事故 A 发生的概率为 p . 若 A 发生, 保险公司要赔偿给投保者 M 元. 为使公司的期望收益达到 $0.05M$ 元, 公司将要求投保者交纳多少保费?

分析 这是数学期望的应用题. 公司在每一投保者身上获取的收益为随机变量, 投保者需交纳的保费则是一个常量.

解 设投保者需交纳保费 x 元, 公司在每一投保者身上获取的收益为随机变量 X , 由题设可知 X 服从两点分布, 即

$$X \sim \begin{pmatrix} x & x-M \\ 1-p & p \end{pmatrix}.$$

由 $E(X) = x(1-p) + p(x-M) = 0.05M$ 得 $x = (0.05 + p)M$.

所以投保者需交纳 $(0.05 + p)M$ 元保费.

点评 本题的关键之处是正确求得 X 的分布律. 部分学生算得的结果是: 需交纳 $\frac{(0.05+p)M}{1-p}$ 元保费, 其错误原因是 X 的分布律

$$\text{为} \begin{pmatrix} x & -M \\ 1-p & p \end{pmatrix}.$$

【4-19】 公共汽车车门的高度是按男子与车门顶碰头的机会在 0.5% 以下来设计的. 设男子身高服从正态分布 $N(176, 64)$ (单位: cm). 问车门高度应如何确定?

解 设车门高度为 h (cm), 男子身高为随机变量 X , 则 $X \sim N(176, 64)$, 依题意可知

$$P(X > h) < 0.005 \Rightarrow P(X \leq h) > 0.995 \text{ 即 } \Phi\left(\frac{h-176}{8}\right) > 0.995.$$

查标准正态分布表,得

$$\frac{h-176}{8} \geq 2.58 \Rightarrow h \geq 196.64.$$

即车门高度可确定为 196.64 cm.

【4-20】 用卡车装运水泥. 设每袋水泥的重量 X (单位: kg) 服从 $N(50, 2.5^2)$. 问最多装多少袋水泥使总重量超过 2 000 的概率不大于 5%. 下面给出两种解法, 判断哪个正确, 并说明理由.

解法 1 设最多装 m 袋水泥, 则总重量为 mX , 据题设应有

$$\begin{aligned} P(mX > 2000) &= 1 - P(mX \leq 2000) = 1 - P\left(X \leq \frac{2000}{m}\right) \\ &= 1 - \Phi\left[\frac{\frac{2000}{m} - 50}{2.5}\right] \leq 0.05, \end{aligned}$$

$$\Phi\left[\frac{\frac{2000}{m} - 50}{2.5}\right] \geq 0.95.$$

查标准正态分布表 $\frac{\frac{2000}{m} - 50}{2.5} \geq 1.645, m \leq 36.96 \approx 37$, 所以最多装 37 袋水泥.

解法 2 设最多装 m 袋水泥, 再设 X_i 表示第 i 袋水泥重量, 则总重量 $Y = \sum_{i=1}^m X_i$. 由于 $X_1, X_2, \dots, X_m$ 相互独立, 据期望和方差的性质可知

$$E(Y) = \sum_{i=1}^m E(X_i) = 50m, \quad D(Y) = \sum_{i=1}^m D(X_i) = 2.5^2 m,$$

由正态分布的可加性可得

$$Y \sim N(50m, 2.5^2 m).$$

由题设有

$$P(Y > 2000) = 1 - \Phi\left(\frac{2000 - 50m}{2.5 \sqrt{m}}\right) \leq 0.05.$$

查标准正态分布表 $\frac{2000-50m}{2.5\sqrt{m}} \geq 1.645, m \leq 39.43$, 所以最多装 39 袋水泥.

解 解法 2 正确.

解法 1 中以 mX 表示总量是欠妥的, 它以一袋的随机重量代表其他每一袋的重量, 这种做法不妨称为是“半随机”的, 而解法 2 中的总量 $\sum_{i=1}^m X_i$ 则是完全彻底的随机.

两个解法若不放在一起, 初学者单看方法 1 是很难发现其错误的. 这里介绍如何发现错误的方法: 对所得结果进行检验, 看是否符合实际. 例如: 解法 1 的结果是 37,

$$50 \times 37 = 1850 < 1900.$$

由于本题方差较小, 数据波动(即每袋重量波动)不大, 故装 37 袋总重量不超过 1900, 这与题中要求超过 2000 差距太大, 由此断定 37 是错误结果.

【4-21】 设随机变量 X 与 Y 相互独立, 其密度函数分别为

$$f_X(x) = \begin{cases} 2x & (0 \leq x \leq 1), \\ 0 & (\text{其他}); \end{cases}$$

$$f_Y(y) = \begin{cases} e^{-(y-5)} & (y > 5), \\ 0 & (y \leq 5). \end{cases}$$

求: $E(XY), D(X-Y)$.

解 方法 1 直接用二维随机变量函数的期望公式计算.

$$\begin{aligned} E(XY) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f(x, y) dx dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f_X(x) f_Y(y) dx dy \\ &= \int_0^1 2x^2 dx \int_5^{+\infty} ye^{-(y-5)} dy = \frac{2}{3} \int_5^{+\infty} -y d(e^{-y+5}) \\ &= \frac{2}{3} (-ye^{-y+5} - e^{-y+5}) \Big|_5^{+\infty} = \frac{2}{3} (5 + 1) = 4. \end{aligned}$$

$$E(X-Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x-y) f(x, y) dx dy$$

$$\begin{aligned}
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x-y) f_X(x) f_Y(y) dx dy \\
&= \int_0^1 2x dx \int_5^{+\infty} (x-y) e^{-y+5} dy = \int_0^1 2x(x-6) dx \\
&= \left(\frac{2}{3} x^3 - 6x^2 \right) \Big|_0^1 = -\frac{16}{3},
\end{aligned}$$

$$\begin{aligned}
E[(X-Y)^2] &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x-y)^2 f(x,y) dx dy \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x-y)^2 f_X(x) f_Y(y) dx dy \\
&= \int_0^1 2x dx \int_5^{+\infty} (x^2 - 2xy + y^2) e^{-y+5} dy \\
&= \int_0^1 2x(x^2 - 12x + 37) dx \\
&= \left(\frac{x^4}{2} - 8x^3 + 37x^2 \right) \Big|_0^1 = \frac{59}{2},
\end{aligned}$$

从而所求方差

$$D(X-Y) = E((X-Y)^2) - [E(X-Y)]^2 = \frac{59}{2} - \left(-\frac{16}{3}\right)^2 = \frac{19}{18}.$$

方法 2 利用期望与方差的性质.

$$E(X) = \int_{-\infty}^{+\infty} x f_X(x) dx = \int_0^1 2x^2 dx = \frac{2}{3},$$

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f_X(x) dx = \int_0^1 2x^3 dx = \frac{1}{2},$$

$$E(Y) = \int_{-\infty}^{+\infty} y f_Y(y) dy = \int_5^{+\infty} y e^{-y+5} dy = 6,$$

$$E(Y^2) = \int_{-\infty}^{+\infty} y^2 f_Y(y) dy = \int_5^{+\infty} y^2 e^{-y+5} dy = 37,$$

$$D(X) = E(X^2) - [E(X)]^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18},$$

$$D(Y) = E(Y^2) - [E(Y)]^2 = 37 - 6^2 = 1.$$

由于 X 与 Y 相互独立, 故有

$$E(XY) = E(X)E(Y) = \frac{2}{3} \times 6 = 4,$$

$$D(X-Y) = D(X) + D(Y) = \frac{1}{18} + 1 = \frac{19}{18}.$$

点评 显然方法 2 较方法 1 计算简单.

【4-22】 设随机变量 X, Y 相互独立, 且都服从区间 $(0, 1)$ 上的均匀分布. 令 $M = \max(X, Y), N = \min(X, Y)$, 求 M 与 N 的数学期望和方差.

分析 M, N 是两个随机变量的函数, 故可用二重积分计算期望与方差; 由于 M, N 仍是一维随机变量, 故在求得 M, N 的密度函数后也可用一维积分计算期望和方差.

解 方法 1 按二维随机变量处理.

由题设可知 (X, Y) 的联合密度

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

$$\text{又} \quad \max(x, y) = \begin{cases} x & (x \geq y), \\ y & (x < y), \end{cases} \quad \min(x, y) = \begin{cases} y & (x \geq y), \\ x & (x < y). \end{cases}$$

于是

$$\begin{aligned} E(M) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \max(x, y) f(x, y) dx dy \\ &= \iint_{y \leq x} x f(x, y) dx dy + \iint_{y > x} y f(x, y) dx dy \\ &= \int_0^1 dx \int_0^x x dy + \int_0^1 dx \int_x^1 y dy \\ &= \frac{1}{3} + \frac{1}{3} = \frac{2}{3}. \end{aligned}$$

同理

$$\begin{aligned} E(N) &= \int_0^1 dx \int_0^x y dy + \int_0^1 dx \int_x^1 x dy \\ &= \frac{1}{6} + \frac{1}{6} = \frac{1}{3}. \end{aligned}$$

$$E(M^2) = \int_0^1 dx \int_0^x x^2 dy + \int_0^1 dx \int_x^1 y^2 dy$$

$$= \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

$$\begin{aligned} E(N^2) &= \int_0^1 dx \int_0^x y^2 dy + \int_0^1 dx \int_x^1 x^2 dy \\ &= \frac{1}{12} + \frac{1}{12} = \frac{1}{6}. \end{aligned}$$

所以

$$D(M) = E(M^2) - [E(M)]^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18},$$

$$D(N) = E(N^2) - [E(N)]^2 = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{18}.$$

方法 2 按一维随机变量处理.

X, Y 均服从 $(0, 1)$ 上的均匀分布, 故其分布函数和密度函数为

$$F(x) = \begin{cases} 0 & (x < 0), \\ x & (0 \leq x < 1), \\ 1 & (x \geq 1); \end{cases} \quad f(x) = \begin{cases} 1 & (0 < x < 1), \\ 0 & (\text{其他}). \end{cases}$$

于是 M, N 的分布函数

$$F_M(z) = [F(z)]^2, \quad F_N(z) = 1 - [1 - F(z)]^2,$$

M, N 的密度函数

$$f_M(z) = F'_M(z) = 2F(z)f(z) = \begin{cases} 2z & (0 < z < 1), \\ 0 & (\text{其他}); \end{cases}$$

$$f_N(z) = F'_N(z) = 2[1 - F(z)]f(z) = \begin{cases} 2(1 - z) & (0 < z < 1), \\ 0 & (\text{其他}). \end{cases}$$

故所求的数学期望

$$E(M) = \int_{-\infty}^{+\infty} z f_M(z) dz = \int_0^1 2z^2 dz = \frac{2}{3},$$

$$E(N) = \int_{-\infty}^{+\infty} z f_N(z) dz = \int_0^1 2z(1 - z) dz = \frac{1}{3}.$$

又

$$E(M^2) = \int_{-\infty}^{+\infty} z^2 f_M(z) dz = \int_0^1 2z^3 dz = \frac{1}{2},$$

$$E(N^2) = \int_{-\infty}^{+\infty} z^2 f_N(z) dz = \int_0^1 2z^2(1-z) dz = \frac{1}{6}.$$

故所求的方差

$$D(M) = E(M^2) - [E(M)]^2 = \frac{1}{18},$$

$$D(N) = E(N^2) - [E(N)]^2 = \frac{1}{18}.$$

点评 对本题而言,两种方法各有千秋,且计算量也相当.

【4-23】 设随机变量 X, Y 相互独立,且都服从正态分布 $N(0, 0.5)$,

求 $D(|X-Y|)$.

解 方法 1 按二维随机变量处理.

由题设可知 (X, Y) 的联合密度函数

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

$$\begin{aligned} E(|X-Y|) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |x-y| \frac{1}{\pi} e^{-(x^2+y^2)} dx dy \\ &= \frac{1}{\pi} \int_0^{2\pi} d\theta \int_0^{+\infty} r |\cos\theta - \sin\theta| e^{-r^2} r dr \\ &= \frac{4}{\pi} \int_0^{\frac{\pi}{2}} |\cos\theta - \sin\theta| d\theta \int_0^{+\infty} r^2 e^{-r^2} dr \\ &= \frac{4}{\pi} \left(\int_0^{\frac{\pi}{4}} (\cos\theta - \sin\theta) d\theta + \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} (\sin\theta - \cos\theta) d\theta \right) \\ &\quad \left(-\frac{1}{2} r e^{-r^2} \Big|_0^{+\infty} + \frac{1}{2} \int_0^{+\infty} e^{-r^2} dr \right) \\ &= \frac{4}{\pi} \times \sqrt{2} \times \frac{\sqrt{\pi}}{4} = \sqrt{\frac{2}{\pi}}, \end{aligned}$$

$$\begin{aligned} E(|X-Y|^2) &= E(X^2 - 2XY + Y^2) = E(X^2) - 2E(X)E(Y) + E(Y^2) \\ &= D(X) + D(Y) = 0.5 + 0.5 = 1, \end{aligned}$$

所以

$$D(|X-Y|) = E(|X-Y|^2) - (E|X-Y|)^2 = 1 - \frac{2}{\pi}.$$

方法 2 按一维随机变量处理.

令 $Z=X-Y$, 由于相互独立的正态变量的线性组合仍是正态变量, 故 $Z \sim N(0, 1)$. 由 $E(Z)=0$ 得 $E(Z^2)=D(Z)=1$.

$$\begin{aligned} E(|Z|) &= \int_{-\infty}^{+\infty} |z| \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} z e^{-\frac{z^2}{2}} dz \\ &= \sqrt{\frac{2}{\pi}} \left(-e^{-\frac{z^2}{2}} \right) \Big|_0^{+\infty} = \sqrt{\frac{2}{\pi}}, \end{aligned}$$

于是

$$D(|X-Y|) = D(|Z|) = E(Z^2) - [E(|Z|)]^2 = 1 - \frac{2}{\pi}.$$

点评 与上一题不同, 这里用一维随机变量方法处理二维随机变量函数的数字特征问题, 比用二维随机变量方法来处理要好得多.

本题出现了绝对值, 不少学生按过去的思路会这样求解:

① 当 $X-Y \geq 0$ 时, 由 X, Y 的独立性及方差的性质可得

$$D(|X-Y|) = D(X-Y) = D(X) + D(Y) = 0.5 + 0.5 = 1;$$

② 当 $X-Y < 0$ 时,

$$D(|X-Y|) = D(Y-X) = D(X) + D(Y) = 0.5 + 0.5 = 1,$$

综合①, ②得

$$D(|X-Y|) = 1.$$

这一解法粗看似乎无可挑剔, 仔细一琢磨, 发现确有问题. 问题出在对随机变量的概念未真正理解, 将 X, Y 视为一般非随机变量, 故而将问题分成两部分分别求解. 事实上, 由于 X, Y 的随机性, 使之不能保持恒有 $X-Y \geq 0$ 或 $X-Y < 0$, 因此分两段讨论这一方法本身就是错误的.

【4-24】 设连续型随机变量 X 的一切可能值在区间 $[a, b]$ 内, 其密度函数为 $f(x)$. 证明:

$$(1) a \leq E(X) \leq b;$$

$$(2) D(X) \leq \frac{(b-a)^2}{4}.$$

证 (1) 由 $f(x)$ 的归一性及期望的定义可有

$$a = a \cdot 1 = a \int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^{+\infty} a f(x) dx \leq \int_{-\infty}^{+\infty} x f(x) dx = E(X).$$

相仿可证 $b \geq E(X)$, 所以

$$a \leq E(X) \leq b.$$

(2) 利用方差性质可有 $D(X) \leq E(X-c)^2$, c 为任意常数.

取 $c = \frac{a+b}{2}$, 则有

$$D(X) \leq E(X-c)^2 = E\left(X - \frac{a+b}{2}\right)^2 \leq E\left(b - \frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{4}.$$

点评 证明中常见如下错误:

① 由题设知 $X \sim U[a, b]$.

题中仅设连续型随机变量在 $[a, b]$ 内取值, 这并不意味着 X 服从均匀分布, 故接下去的证明失去一般性.

$$\textcircled{2} D(X) = E(X^2) - [E(X)]^2 \leq b^2 - a^2 \leq \frac{(b-a)^2}{4}.$$

当 $a=1, b=2$ 时, 得 $3 \leq \frac{1}{4}$, 矛盾.

$$\textcircled{3} D(X) \leq E(X-c)^2 \leq E(b-c)^2.$$

在 c 未确定时, 未必有 $E(X-c)^2 \leq E(b-c)^2$. 例如 $X=1, b=2, c=3$, 得 $4 \leq 1$, 矛盾.

【4-25】 某地区种植某种农作物, 根据以往数据统计平均亩产量 (单位: kg) 是 412, 产量的均方差是 16. 试估计今年的亩产量与 412 的偏差不小于 47 的概率.

分析 尽管不知道农作物的亩产量服从什么分布, 但只要知道亩产量的期望和方差, 便可利用切比雪夫不等式对亩产量的有关问题进行概率估算.

解 设随机变量 X 为某农作物的亩产量, 则 $E(X)=412, D(X)=16^2$, 由切比雪夫不等式可有

$$\begin{aligned} P(|X - E(X)| \geq 47) &= P(|X - 412| \geq 47) \leq \frac{D(X)}{47^2} \\ &= \frac{16^2}{47^2} \approx 0.1159. \end{aligned}$$

【4-26】 设二维随机变量 (X, Y) 的联合分布律如下:

X \ Y	0	1
0	0.1	0.3
1	0.2	0.4

求: $\text{cov}(X, Y), \rho_{XY}$.

解 由联合分布得边缘分布 $X \sim B(1, 0.7), Y \sim B(1, 0.6)$, 故 $E(X) = 0.7, E(Y) = 0.6, D(X) = 0.21, D(Y) = 0.24$.

又 $XY \sim B(1, 0.4), E(XY) = 0.4$, 从而

$$\text{cov}(X, Y) = E(XY) - E(X)E(Y) = 0.4 - 0.7 \times 0.6 = -0.02,$$

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sqrt{D(X)} \sqrt{D(Y)}} = \frac{-0.02}{\sqrt{0.21} \sqrt{0.24}} = -\frac{\sqrt{14}}{42} \approx -0.09.$$

【4-27】 设二维随机变量 (X, Y) 的联合密度函数

$$f(x, y) = \begin{cases} \frac{1}{8}(x+y) & (0 \leq x \leq 2, 0 \leq y \leq 2), \\ 0 & (\text{其他}). \end{cases}$$

求: $\text{cov}(X, Y), \rho_{XY}$.

解 由对称性可知 $E(X) = E(Y), D(X) = D(Y)$, 而

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xf(x, y) dx dy = \int_0^2 x dx \int_0^2 \frac{1}{8}(x+y) dy \\ &= \int_0^2 \left(\frac{x^2}{4} + \frac{x}{4} \right) dx = \frac{7}{6}, \end{aligned}$$

$$\begin{aligned} E(XY) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xyf(x, y) dx dy = \int_0^2 x dx \int_0^2 \frac{1}{8}(xy + y^2) dy \\ &= \int_0^2 \left(\frac{x^2}{4} + \frac{x}{3} \right) dx = \frac{4}{3}, \end{aligned}$$

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^2 f(x, y) dx dy = \int_0^2 x^2 dx \int_0^2 \frac{1}{8}(x+y) dy \\ &= \int_0^2 \left(\frac{x^3}{4} + \frac{x^2}{4} \right) dx = \frac{5}{3}, \end{aligned}$$

$$D(X) = E(X^2) - [E(X)]^2 = \frac{5}{3} - \left(\frac{7}{6} \right)^2 = \frac{11}{36}.$$

于是所求协方差和相关系数:

$$\text{cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{4}{3} - \left(\frac{7}{6}\right)^2 = -\frac{1}{36},$$

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sqrt{D(X)} \sqrt{D(Y)}} = \frac{-\frac{1}{36}}{\frac{11}{36}} = -\frac{1}{11}.$$

点评 注意到联合密度函数关于 x 与 y 的对称性, 便可减少近一半的计算量.

【4-28】 设二维随机变量 (X, Y) 的联合分布律如下:

$\begin{array}{c} X \\ Y \end{array}$	-1	0	1
-1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
0	$\frac{1}{8}$	0	$\frac{1}{8}$
1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

试验证: X 与 Y 是不相关的, 但 X 和 Y 不相互独立.

证 由联合分布律可得

$$X \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{3}{8} & \frac{1}{4} & \frac{3}{8} \end{pmatrix}, \quad Y \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{3}{8} & \frac{1}{4} & \frac{3}{8} \end{pmatrix},$$

$$XY \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix},$$

从而

$$E(X) = E(Y) = E(XY) = 0$$

$$\Rightarrow \text{cov}(X, Y) = E(XY) - E(X)E(Y) = 0,$$

即 $\rho_{XY} = 0$, 故 X 与 Y 是不相关的. 但

$$P(X = 0, Y = 0) = 0 \neq \frac{1}{16} = P(X = 0)P(Y = 0),$$

所以 X 和 Y 不相互独立.

【4-29】 设随机变量 $Y=2X+3$, 且 $D(X)=4$, 求 $\text{cov}(X, Y)$ 和 ρ_{XY} .

分析 本题是利用协方差的各性质求协方差.

$$\begin{aligned}\text{解 } \text{cov}(X, Y) &= \text{cov}(X, 2X+3) = \text{cov}(X, 2X) + \text{cov}(X, 3) \\ &= 2\text{cov}(X, X) + 0 = 2D(X) = 8;\end{aligned}$$

$$D(Y) = D(2X+3) = D(2X) + 0 = 4D(X) = 16,$$

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sqrt{D(X)} \sqrt{D(Y)}} = \frac{8}{\sqrt{4} \sqrt{16}} = 1.$$

点评 若知道相关系数取 ± 1 的充要条件, 则不用计算, 由题设 $Y=2X+3$ 就可推得 $\rho_{XY}=1$.

【4-30】 已知二维随机变量服从正态分布, 且 $E(X)=E(Y)=0$, $D(X)=4, D(Y)=9, \text{cov}(X, Y)=3$, 求 (X, Y) 的联合密度函数 $f(x, y)$.

$$\text{解 } f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2 \sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)} \left(\frac{x^2}{\sigma_1^2} - 2\rho\frac{xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2} \right)},$$

其中

$$\sigma_1 = \sqrt{D(X)} = 2, \sigma_2 = \sqrt{D(Y)} = 3,$$

$$\rho = \frac{\text{cov}(X, Y)}{\sqrt{D(X)} \sqrt{D(Y)}} = \frac{3}{2 \times 3} = \frac{1}{2},$$

代入上式可得

$$f(x, y) = \frac{\sqrt{3}}{18\pi} e^{-\frac{2}{3} \left(\frac{x^2}{4} - \frac{xy}{6} + \frac{y^2}{9} \right)}.$$

【4-31】 设随机变量 X 和 Y 的方差分别为 4 和 9, 相关系数为 0.5, 求 $D(X+Y)$ 和 $D(X-Y)$

$$\text{解 } D(X \pm Y) = D(X) + D(Y) \pm 2\text{cov}(X, Y),$$

将 $D(X)=4, D(Y)=9, \text{cov}(X, Y)=\rho_{XY} \sqrt{D(X)} \sqrt{D(Y)}=0.5 \times 2 \times 3=3$ 代入上式, 得

$$D(X+Y) = 4 + 9 + 2 \times 3 = 19,$$

$$D(X-Y) = 4 + 9 - 2 \times 3 = 7.$$

【4-32】 (1) 求题 4-26 中 (X, Y) 的协方差矩阵 C ;

(2) 求题 4-27 中 (X, Y) 的协方差矩阵 C .

解 $C = (c_{ij})_{2 \times 2}$, 其中 $c_{ij} = c_{ji} = \text{cov}(X_i, X_j) \ (i, j = 1, 2)$.

$$(1) \ c_{11} = \text{cov}(X, X) = D(X) = 0.21,$$

$$c_{22} = \text{cov}(Y, Y) = D(Y) = 0.24,$$

$$c_{12} = \text{cov}(X, Y) = -0.02 = c_{21},$$

所以

$$C = \begin{pmatrix} 0.21 & -0.02 \\ -0.02 & 0.24 \end{pmatrix};$$

(2) 相仿

$$C = \begin{pmatrix} \frac{11}{36} & -\frac{1}{36} \\ -\frac{1}{36} & \frac{11}{36} \end{pmatrix} = \frac{1}{36} \begin{pmatrix} 11 & -1 \\ -1 & 11 \end{pmatrix}.$$

【4-33】 设 A 和 B 是试验 E 的两个随机事件, 且 $P(A) > 0$, $P(B) > 0$. 定义随机变量 X 和 Y 如下:

$$X = \begin{cases} 1 & (A \text{ 发生}), \\ 0 & (\bar{A} \text{ 发生}), \end{cases} \quad Y = \begin{cases} 1 & (B \text{ 发生}), \\ 0 & (\bar{B} \text{ 发生}). \end{cases}$$

证明: 若相关系数 $\rho_{XY} = 0$, 则 X 和 Y 必定相互独立.

分析 要证明离散型随机变量 X, Y 相互独立, 即需证明 (X, Y) 的联合分布律等于边缘分布律的乘积, 即

$$P(X = i, Y = j) = P(X = i)P(Y = j) \ (i, j = 0, 1).$$

证 由题设可知

$$X \sim \begin{pmatrix} 0 & 1 \\ 1 - P(A) & P(A) \end{pmatrix},$$

$$Y \sim \begin{pmatrix} 0 & 1 \\ 1 - P(B) & P(B) \end{pmatrix},$$

$$XY \sim \begin{pmatrix} 0 & 1 \\ 1 - P(AB) & P(AB) \end{pmatrix},$$

故 $E(X) = P(A)$, $E(Y) = P(B)$, $E(XY) = P(AB)$.

由于 $\rho_{XY} = 0$, 即 $\text{cov}(X, Y) = E(XY) - E(X)E(Y) = 0$, 亦即 $E(XY) = E(X)E(Y)$, 故 $P(AB) = P(A)P(B)$, 从而有

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

由 A, B 相互独立可得 $\bar{A}$ 和 $\bar{B}, A$ 和 $\bar{B}, \bar{A}$ 和 B 也相互独立, 故又有

$$P(X = 0, Y = 0) = P(X = 0)P(Y = 0),$$

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

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

所以 X 和 Y 必定相互独立.

点评 按如下方法证明是不正确的:

$$\begin{aligned} \text{由 } \rho_{XY} = 0 &\Rightarrow E(XY) = E(X)E(Y) \Rightarrow P(AB) = P(A)P(B) \\ &\Rightarrow A \text{ 和 } B \text{ 相互独立} \Rightarrow X \text{ 和 } Y \text{ 相互独立.} \end{aligned}$$

错误在于仅证明了一个等式 $P(X=1, Y=1) = P(X=1)P(Y=1)$ 成立, 没有证明所有等式成立.

B 习题精选

【4-34】 设随机变量 $X \sim \begin{pmatrix} a & b \\ 0.6 & p \end{pmatrix} (a < b)$, 又 $E(X) = 1.4$,

$D(X) = 0.24$, 则 a, b 的值为_____.

(A) $a=1, b=2$; (B) $a=-1, b=2$;

(C) $a=1, b=-2$; (D) $a=0, b=1$.

【4-35】 对两个仪器进行独立试验. 设这两个仪器发生故障的概率分别为 p_1 与 p_2 , 则发生故障的仪器数的数学期望为_____.

(A) $p_1 p_2$; (B) $p_1 + p_2$;

(C) $p_1 + (1 - p_2)$; (D) $p_1(1 - p_2) + p_2(1 - p_1)$.

【4-36】 设 X 为随机变量, $E(X) = \mu, D(X) = \sigma^2 (\sigma > 0)$, 则对任意常数 c , 必有_____.

(A) $E(X - c)^2 = E(X^2) - c^2$;

(B) $E(X - c)^2 = E(X - \mu)^2$;

(C) $E(X - c)^2 < E(X - \mu)^2$;

(D) $E(X-c)^2 \geq E(X-\mu)^2$.

【4-37】 人的体重 $X \sim N(100, 100)$, 记 Y 为 10 个人的平均体重, 则_____.

(A) $E(Y)=100, D(Y)=100$;

(B) $E(Y)=100, D(Y)=10$;

(C) $E(Y)=10, D(Y)=100$;

(D) $E(Y)=10, D(Y)=10$.

【4-38】 设 $X \sim B(n, p)$, 且 $E(X)=2.4, D(X)=1.44$, 则 n, p 的值为_____.

(A) $n=12, p=0.2$;

(B) $n=8, p=0.3$;

(C) $n=6, p=0.4$;

(D) $n=4, p=0.6$.

【4-39】 设 X 服从指数分布 $E(\lambda)$, 则 $P(X > E(X)) =$ _____.

【4-40】 设 X 服从泊松分布 $\pi(2)$, 则 $E(3X-2) =$ _____.

【4-41】 设 X 服从泊松分布 $\pi(\lambda)$, 且 $E[(X-1)(X-2)] = 1$, 则 $\lambda =$ _____.

【4-42】 设 $X \sim B(n, p), Y = 3^X - 2$, 则 $E(Y) =$ _____.

【4-43】 设 X, Y 相互独立, 均服从 $N(\mu, \sigma^2)$, 令 $Z = 3X - 2Y$, 则 $E(Z) =$ _____, $D(Z) =$ _____.

【4-44】 一批产品中有一、二、三等品及废品 4 种, 相应的概率分别为 0.8, 0.15, 0.04 及 0.01, 若其产值分别为 20 元、18 元、15 元及 0 元, 求产品的平均产值.

【4-45】 某车间完成生产线改造的天数 X 是一随机变量, 其分布律 $X \sim \begin{pmatrix} 26 & 27 & 28 & 29 & 30 \\ 0.1 & 0.2 & 0.4 & 0.2 & 0.1 \end{pmatrix}$, 所得利润(单位: 万元)为 $Y = 5(29 - X)$. 求:

(1) $E(X)$; (2) $E(Y)$.

【4-46】 (有奖销售) 某商场举办购物有奖活动, 每购 1 000 份物品中有一等奖 1 名, 奖金 500 元; 二等奖 3 名, 奖金 100 元; 三等奖 16 名, 奖金 50 元; 四等奖 100 名, 可得价值 5 元的奖品一份. 商场把每份价值为 7.5 元的物品以 10 元出售, 求每个顾客买一份商品平均付多

少钱？

【4-47】 设离散随机变量 X 的分布律

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

求 $Y=\sin\left(\frac{\pi}{2}X\right)$ 的数学期望和方差.

【4-48】 设连续随机变量 X 的密度函数

$$f(x) = \begin{cases} \frac{18}{(x+3)^3} & (x > 0), \\ 0 & (x \leq 0). \end{cases}$$

求 $Y=X+3$ 的数学期望.

【4-49】 设在某一规定的时段内,其电气设备用于最大负荷的时间 X (以分计)是一个随机变量,其密度函数

$$f(x) = \begin{cases} \frac{x}{(1500)^2} & (0 \leq x < 1500), \\ \frac{3000-x}{(1500)^2} & (1500 \leq x \leq 3000), \\ 0 & (\text{其他}). \end{cases}$$

求 X 的数学期望.

【4-50】 设随机变量 X 的密度函数

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

求 $E(X)$ 和 $D(X)$.

【4-51】 设二维随机变量 (X,Y) 的联合分布律如下:

$\begin{matrix} X \\ Y \end{matrix}$	-1	0	2
-1	0.05	0.15	0.25
2	0.2	0.3	0.05

求: (1) $E(X), E(Y), D(X), D(Y)$;

(2) $E(X-Y), D(X-Y)$;

(3) $E(XY), D(XY)$;

(4) $E(\min(X, Y)), D(\min(X, Y))$.

【4-52】 设二维随机变量 (X, Y) 的联合密度

$$f(x, y) = \begin{cases} \frac{1}{\pi R^2} & (x^2 + y^2 < R^2), \\ 0 & (x^2 + y^2 \geq R^2). \end{cases}$$

令 $Z = \sqrt{X^2 + Y^2}$, 求 $E(Z)$ 和 $D(Z)$.

【4-53】 设 X, Y 为任意随机变量, 且 $\text{cov}(X, Y) = 0$, 则必有_____.

(A) X, Y 相互独立; (B) $E(XY) = E(X)E(Y)$;

(C) $D(XY) = D(X)D(Y)$; (D) 以上都不对.

【4-54】 设 X, Y 为随机变量, 若 $D(X+Y) = D(X) + D(Y)$, 则必有_____.

(A) X, Y 相互独立; (B) X, Y 不独立;

(C) X 与 Y 相关; (D) X 与 Y 不相关.

【4-55】 随机变量 X 和 Y 满足 $D(X-Y) = D(X) + D(Y)$ 是 X 和 Y _____.

(A) 独立的必要条件, 但不是充分条件;

(B) 独立的充分条件, 但不是必要条件;

(C) 不相关的必要条件, 但不是充分条件;

(D) 不相关的充分条件, 但不是必要条件.

【4-56】 设随机变量 X 和 Y 独立同分布, 记 $U = X+Y, V = X-Y$, 则 U 和 V 必_____.

(A) 相互独立; (B) 不独立;

(C) 相关系数为零; (D) 相关系数不为零.

【4-57】 设存在常数 $a, b (a \neq 0)$ 使 $P(Y = aX + b) = 1$, 则必有_____.

(A) $\rho_{XY} = a$; (B) $\rho_{XY} = \frac{a}{|a|}$;

(C) $\rho_{XY} = 1$; (D) $\rho_{XY} = -1$.

【4-58】 设 $(X,Y) \sim N(1,2^2;0,3^2;0.5)$, 则 $\text{cov}(X-Y,Y) = \underline{\hspace{2cm}}$.

【4-59】 设 $(X,Y) \sim N(\mu_1, \sigma^2; \mu_2, \sigma^2; \rho)$ ($\sigma > 0$), 则当 $\rho = \underline{\hspace{2cm}}$

时, 随机变量 $\frac{1}{2}(X+Y)$ 与 $\frac{1}{3}(X-2Y)$ 不相关.

【4-60】 设 $(X,Y) \sim N(0,1;0,1;0.5)$, 则 $D(3X-2Y) = \underline{\hspace{2cm}}$.

【4-61】 设 $X \sim N(4,4)$, 由切比雪夫不等式估计 $P(|X-4| \geq 6) \underline{\hspace{2cm}}$.

【4-62】 投掷一枚均匀硬币 100 次, 正面向上出现的频率在 0.4 与 0.6 之间的概率不少于 $\underline{\hspace{2cm}}$.

【4-63】 设 $(X,Y) \sim N(0,2;1,4;-0.5)$, 则 (X,Y) 的协方差矩阵 $C = \underline{\hspace{2cm}}$.

【4-64】 某箱子装有 100 件产品, 其中一、二、三等品分别为 80, 10, 10 件. 现从箱中随机取一件, 记

$$X_i = \begin{cases} 1 & (\text{若取到 } i \text{ 等品}), \\ 0 & (\text{其他}), \end{cases} \quad i = 1, 2, 3.$$

求: (1) (X_1, X_2) 的联合分布律;

(2) X_1 和 X_2 的相关系数 $\rho_{X_1 X_2}$.

【4-65】 设二维随机变量 (X,Y) 的联合密度函数

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

求: (1) $\text{cov}(X,Y)$;

(2) (X,Y) 的协方差矩阵.

【4-66】 设随机变量 $X \sim N(1,16)$, $Y \sim N(1,9)$, $\rho_{XY} = 0.5$. 令 $Z = \frac{X}{2} + \frac{Y}{3}$,

求: (1) $E(Z), D(Z)$;

(2) ρ_{YZ} ;

(3) Y 与 Z 是否相互独立, 为什么?

【4-67】 设 X, Y 都是标准化随机变量, $\rho_{XY} = \frac{1}{2}$. 令 $U = aX, V = bX + cY$, 试确定 a, b, c 的值, 使 $D(U) = D(V) = 1$, 且 U 与 V 不相关.

【4-68】 若随机变量 X 和 Y 相互独立,证明:

$$D(XY) = D(X)D(Y) + [E(X)]^2 D(Y) + [E(Y)]^2 D(X).$$

【4-69】* 若 X 和 Y 都是只能取两个值的随机变量,证明:如果它们不相关,则一定相互独立.

【4-70】* 设 X, Y 均为标准化的随机变量,证明

$$E[\max(X^2, Y^2)] \leq 1 + \sqrt{1 - \rho_{XY}^2}.$$

答案与提示

【4-34】 选 A. 只要解关于 a, b 的方程组

$$\begin{cases} 0.6a + 0.4b = 1.4, \\ 0.6a^2 + 0.4b^2 = 0.24 + 1.4^2. \end{cases}$$

【4-35】 选 B. 发生故障的仪器数 $X = X_1 + X_2, X_i \sim \begin{pmatrix} 0 & 1 \\ 1-p_i & p_i \end{pmatrix}, i=1, 2$. 再利用期望的性质.

【4-36】 选 D. 由方差性质 $D(X) \leq E(X-c)^2$ (c 为任意常数) 即得.

【4-37】 选 B. C, D 显然错, $D(Y) = D\left(\frac{1}{10} \sum_{i=1}^{10} X_i\right) = \frac{1}{100} \sum_{i=1}^{10} D(X_i) = \frac{10 \times 100}{100} = 10$.

【4-38】 选 C. 由 $np=2.4, np(1-p)=1.44$ 解得.

【4-39】 e^{-1} . $P(X > E(X)) = P\left(X > \frac{1}{\lambda}\right) = 1 - F\left(\frac{1}{\lambda}\right)$, $F(x)$ 为 X 的分布函数.

【4-40】 4.

【4-41】 $\lambda=1$. $E((X-1)(X-2)) = \lambda^2 - 2\lambda + 1$.

【4-42】 $(2p+1)^n - 2$. 利用二项展开式 $(a+b)^n = \sum_{k=0}^n C_n^k a^k b^{n-k}$ 及随机变量函数的期望公式可得

$$E(3^X) = \sum_{k=0}^n 3^k C_n^k p^k (1-p)^{n-k} = (2p+1)^n.$$

【4-43】 $E(Z)=\mu, D(Z)=13\sigma^2.$

【4-44】 19.3 元.

【4-45】 (1) $E(X)=28(\text{天})$; (2) $E(Y)=5(\text{万元})$.

【4-46】 7.9 元. 提示: 设 X 为买一份商品的付钱数, 则

$$X \sim \begin{pmatrix} -490 & -90 & -40 & 5 & 10 \\ 0.001 & 0.003 & 0.016 & 0.1 & 0.88 \end{pmatrix}.$$

【4-47】 $E(Y)=\frac{2}{5}, D(Y)=\frac{38}{75}.$ 提示: 利用等比数列求和可得

$$Y \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{2}{15} & \frac{1}{3} & \frac{8}{15} \end{pmatrix}.$$

【4-48】 6.

【4-49】 1500(分). 通过分段积分求得.

【4-50】 $E(X)=0, D(X)=\frac{1}{6}.$

【4-51】 (1) $E(X)=0.35, E(Y)=0.65, D(X)=1.3275,$
 $D(Y)=2.2275;$

(2) $E(X-Y)=-0.3, D(X-Y)=E((X-Y)^2)-[E(X-Y)]^2=$
 $5.31;$

(3) $E(XY)=-0.65, D(XY)=2.2275;$

(4) $E(\min(X, Y))=-0.55, D(\min(X, Y))=0.5475.$

【4-52】 $E(Z)=\frac{2}{3}R, D(Z)=\frac{1}{18}R^2.$ 利用极坐标计算二重

积分.

【4-53】 选 B. $\text{cov}(X, Y)=0 \Leftrightarrow E(XY)=E(X)E(Y).$

【4-54】 选 D. $D(X+Y)=D(X)+D(Y) \Leftrightarrow X$ 与 Y 不相关.

【4-55】 选 A. X 和 Y 独立 $\Rightarrow D(X-Y)=D(X)+D(Y).$

【4-56】 选 C. $\text{cov}(U, V)=D(X)-D(Y)=0 \Leftrightarrow \rho_{UV}=0.$

【4-57】 选 B. 若 $a>0$ 则 $\rho_{XY}=1$; 若 $a<0$, 则 $\rho_{XY}=-1.$

【4-58】 $-6. \quad \text{cov}(X-Y, Y) = \rho_{XY} \sqrt{D(X)} \sqrt{D(Y)} - D(Y).$

【4-59】 $-1. \quad \text{cov}\left(\frac{X+Y}{2}, \frac{X-2Y}{3}\right) = \frac{\sigma^2}{6}(-1-\rho) = 0.$

【4-60】 7.

$$D(3X-2Y) = 9D(X) + 4D(Y) - 12\rho_{XY} \sqrt{D(X)} \sqrt{D(Y)}.$$

【4-61】 $\leq \frac{1}{9}.$

【4-62】 $\frac{3}{4}.$ 利用切比雪夫不等式.

【4-63】 $\begin{pmatrix} 2 & -\sqrt{2} \\ -\sqrt{2} & 4 \end{pmatrix}.$

【4-64】 (1)

$X_1 \backslash X_2$	0	1
0	0.1	0.8
1	0.1	0

(2) $\rho_{X_1 X_2} = -\frac{2}{3}.$

【4-65】 (1) $-\frac{1}{36};$ (2) $\frac{1}{36} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$

【4-66】 (1) $E(Z) = \frac{5}{6}, \quad D(Z) = 7;$ (2) $\rho_{YZ} = \frac{2\sqrt{7}}{7};$

(3) 不独立, 因为 $\rho_{YZ} \neq 0.$

【4-67】 $a = \pm 1, b = \frac{1}{\sqrt{3}}, c = -\frac{2}{\sqrt{3}};$

或 $a = \pm 1, b = -\frac{1}{\sqrt{3}}, c = \frac{2}{\sqrt{3}}.$

【4-68】 提示: ① $D(XY) = E(X^2 Y^2) - [E(XY)]^2$, ② X, Y 独立 $\Rightarrow X^2, Y^2$ 独立, ③ $E(X^2) = D(X) + [E(X)]^2.$

【4-69】* 设 (X, Y) 的联合分布律和边缘分布律如下:

Y \ X	x_1	x_2	$p_{\cdot j}$
	x_1	x_2	$p_{\cdot j}$
y_1	a	b	q
y_2	c	d	$1-q$
$p_{i \cdot}$	p	$1-p$	

则 XY 的分布律为

XY	$x_1 y_1$	$x_1 y_2$	$x_2 y_1$	$x_2 y_2$
p	a	c	b	d

于是

$$E(X) = px_1 + (1-p)x_2, \quad E(Y) = qy_1 + (1-q)y_2,$$

$$E(XY) = ax_1y_1 + cx_1y_2 + bx_2y_1 + dx_2y_2.$$

由于 X 与 Y 不相关, 即 $E(XY) = E(X)E(Y)$, 亦即

$$ax_1y_1 + cx_1y_2 + bx_2y_1 + dx_2y_2$$

$$= pqx_1y_1 + p(1-q)x_1y_2 + (1-p)qx_2y_1 + (1-p)(1-q)x_2y_2.$$

比较上面等式两边系数可得

$$a = pq, \quad c = p(1-q), \quad b = (1-p)q, \quad d = (1-p)(1-q),$$

即

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j) \quad (i, j = 1, 2).$$

所以 X 和 Y 相互独立.

$$\text{【4-70】}^* \quad \text{由 } \max(X^2, Y^2) = \begin{cases} X^2 & (X^2 \geq Y^2), \\ Y^2 & (X^2 < Y^2) \end{cases}$$

可得

$$\max(X^2, Y^2) = \frac{1}{2}(X^2 + Y^2) + \frac{1}{2}|X^2 - Y^2|,$$

于是据柯西-许瓦兹不等式并注意到 $E(X^2) = E(Y^2) = 1, E(XY) = \rho_{XY}$, 有

$$E(\max(X^2, Y^2)) = \frac{1}{2}[E(X^2) + E(Y^2)] + \frac{1}{2}E(|X+Y| |X-Y|)$$

$$\begin{aligned}
&\leq 1 + \frac{1}{2} \sqrt{E|X+Y|^2 \cdot E|X-Y|^2} \\
&= 1 + \frac{1}{2} \sqrt{E(X^2 + 2XY + Y^2) \cdot E(X^2 - 2XY + Y^2)} \\
&= 1 + \frac{1}{2} \sqrt{2[1 + E(XY)] \cdot 2[1 - E(XY)]} \\
&= 1 + \sqrt{1 - [E(XY)]^2} = 1 + \sqrt{1 - \rho_{XY}^2}.
\end{aligned}$$

关键词

依概率收敛 伯努利大数定律 切比雪夫大数定律 独立同分布的中心极限定理 德莫佛-拉普拉斯中心极限定理

教学要求

- (1) 了解伯努利大数定律与切比雪夫大数定律.
- (2) 了解独立同分布的中心极限定理与德莫佛-拉普拉斯中心极限定理.

重点

各定理的内容;正态分布在近似计算中的应用.

难点

本章内容是数理统计的理论基础,难度较大.但由于这部分内容教学大纲要求较弱,所以了解各定理的内容和思想、会利用中心极限定理作简单的近似计算即可.

主要内容与方法提要

(1) 大数定律

1) 依概率收敛

若有常数 a , 对随机变量序列 $\{X_n\}$ 及对任意 $\epsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P(|X_n - a| < \epsilon) = 1,$$

则称 $\{X_n\}$ 依概率收敛于 a , 简化为 $X_n \xrightarrow{P} a$.

依概率收敛与高等数学中的收敛是两个不同的概念. 在高等数学中, 序列 $\{x_n\}$ 是普通的变量, $\{x_n\}$ 收敛于常数 a ($\lim_{n \rightarrow \infty} x_n = a$) 指对任意给定的 $\varepsilon > 0, N > 0$, 使得对所有的 $n > N$, 恒有 $|x_n - a| < \varepsilon$, 并且不会有例外.

在概率论中, 序列 $\{X_n\}$ 是随机变量, $\{X_n\}$ 依概率收敛于 a 是指对任意给定的 $\varepsilon > 0$, 当 n 充分大时, 事件 $(|X_n - a| < \varepsilon)$ 发生的概率很大, 接近于 1 (即 $\lim_{n \rightarrow \infty} P(|X_n - a| < \varepsilon) = 1$), 但事件 $(|X_n - a| \geq \varepsilon)$ 还有可能发生, 只是发生的可能性很小.

2) 切比雪夫大数定律

设随机变量序列 $\{X_k\}$ 相互独立, 且具有相同的方差和期望, $E(X_k) = \mu, D(X_k) = \sigma^2 (k=1, 2, \dots)$. 又设前 n 个随机变量的算术平均 $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$, 则对任意的正数 ε , 有

$$\lim_{n \rightarrow \infty} P(|\bar{X} - \mu| < \varepsilon) = \lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{k=1}^n X_k - \mu\right| < \varepsilon\right) = 1.$$

3) 伯努利大数定律

设 n_A 是 n 次独立重复试验中事件 A 发生的次数, p 是事件 A 在每次试验中发生的概率, 则对任意正数 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{n_A}{n} - p\right| < \varepsilon\right) = 1.$$

大数定律刻画了随机现象在大量重复下发生频率具一定稳定性的定律, 也就是揭示了大量随机现象的平均结果一般都具有某种稳定性的规律.

(2) 中心极限定理

1) 独立同分布的中心极限定理

设 $\{X_k\}$ 为相互独立且服从同一分布的随机变量序列, 对每一随机变量都有有限的 $E(X_k) = \mu$ 及有限的方差 $D(X_k) = \sigma^2 \neq 0 (k=1, 2, \dots)$, 则对前 n 个随机变量和 $\sum_{k=1}^n X_k$ 进行标准化变换后

$$Y_n = \frac{\sum_{k=1}^n X_k - E\left(\sum_{k=1}^n X_k\right)}{\sqrt{D\left(\sum_{k=1}^n X_k\right)}} = \frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n}\sigma}.$$

此时对任意的 x 都有

$$\lim_{n \rightarrow \infty} P(Y_n \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x).$$

本定理主要适用以下两个方面:

① 求随机变量落在某区间内的概率. 对于此类情况, 先构造独立同分布且期望和方差均已知的随机变量序列 $\{X_k\}$, 然后将所求事件的概率转化为该随机变量前 n 项之和 $X = \sum_{k=1}^n X_k$, 最后当 n 很大时, 利用本中心极限定理, 可求出所需的概率值.

② 已知随机变量前 n 项和 $X = \sum_{k=1}^n X_k$ 取值的概率, 求满足条件的最小 n . 对于此类问题, 解法与①相反: 首先利用本中心极限定理将已知概率转化成一个与 n 有关的标准正态分布函数 $\Phi(f(n))$, 查表得 $f(n)=c$, 再解出 $f(n)=c$ 中的未知数 n .

中心极限定理阐明了即使原来并不服从正态分布的那些独立的随机变量, 但它们的总和的分布渐近地服从正态分布.

2) 德莫佛-拉普拉斯(De Moivre-Laplace)中心极限定理

设随机变量 $\eta_n (n=1, 2, \dots)$ 服从参数为 $n, p (0 < p < 1)$ 的二项分布, 则对于任意 x , 有

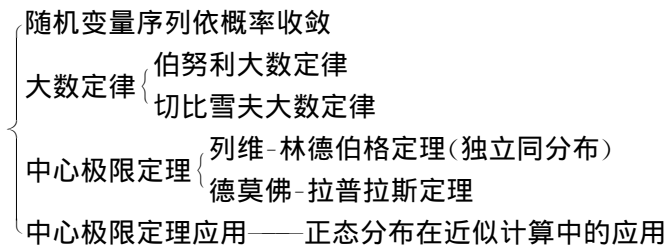
$$\lim_{n \rightarrow \infty} P\left(\frac{\eta_n - np}{\sqrt{np(1-p)}} \leq x\right) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x).$$

随机变量 η_n 可看成 n 个相互独立的服从(0-1)分布的随机变量之和, 即 $\eta_n = \sum_{k=1}^n X_k$, 则德莫佛-拉普拉斯中心极限定理就是独立同分布中心极限定理的一个特例.

当 n 很大时, 除了可利用本定理来作为二项分布的近似计算以外,

由泊松定理可知,也可利用泊松分布来作为二项分布的近似计算.当 p 很小, n 很大时,利用泊松分布来作为二项分布的近似计算时,精度较高.但 p 较大,且 n 很大,则利用本中心极限定理来作为二项分布的近似计算,精度稍差.

本章知识结构



A 例题精解

【5-1】 某计算机系统有 100 个终端,每个终端有 2% 的时间在使用,若各个终端使用与否相互独立,试分别用二项分布、泊松分布和中心极限定理计算至少一个终端使用的概率.

解 设

$$X_i = \begin{cases} 1 & (\text{第 } i \text{ 个终端在使用}), \\ 0 & (\text{第 } i \text{ 个终端未在使用}), \end{cases} \quad i = 1, 2, \dots, n.$$

由题意可知, $P(X_i=1)=0.02$, $P(X_i=0)=0.98$. 若 X 为使用中终端

的总数,则有 $X = \sum_{i=1}^n X_i$.

(1) 用二项分布求解

由于 $X_1, X_2, \dots, X_n$ 是相互独立的,并且均服从(0-1)分布,因此

$X = \sum_{i=1}^n X_i$ 服从二项分布,即 $X_n \sim B(n, p) = P(X=k) = C_n^k p^k (1-p)^{n-k}$, $k=0, 1, \dots, n$.

当 $n=100$, $p=0.02$ 时,解得

$$P(X \geq 1) = 1 - P(X = 0) = 1 - C_{100}^0 \times 0.02^0 \times 0.98^{100} \approx 0.8674.$$

(2) 用泊松分布求解

根据泊松定理, 当 $n(n \geq 20)$ 很大, $p(p \leq 0.05)$ 很小时, 有以下近似公式:

$$P(X = k) = C_n^k p^k (1-p)^{n-k} \approx P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \text{ 其中 } \lambda = np.$$

由于 $\lambda = np = 100 \times 0.02 = 2$, 解得

$$P(X \geq 1) = 1 - P\{X = 0\} \approx 1 - \frac{2^0}{0!} e^{-2} \approx 0.8647.$$

(3) 用中心极限定理求解

已知 $E(X_i) = 0.02, D(X_i) = 0.0196$, 由 $X = \sum_{i=1}^n X_i$ 及 $n = 100$, 算得

$$E(X) = \sum_{i=1}^{100} E(X_i) = 100 \times 0.02 = 2,$$

$$D(X) = \sum_{i=1}^{100} D(X_i) = 100 \times 0.0196 = 1.96.$$

最后解得

$$\begin{aligned} P(X \geq 1) &= P\left(\frac{X - E(X)}{\sqrt{D(X)}} \geq \frac{1 - E(X)}{\sqrt{D(X)}}\right) \\ &= P\left(\frac{X_n - 2}{\sqrt{1.96}} \geq \frac{1 - 2}{\sqrt{1.96}}\right) \\ &\approx 1 - \Phi\left(\frac{1 - 2}{\sqrt{1.96}}\right) = 1 - [1 - \Phi(0.71)] \\ &= \Phi(0.71) \approx 0.7611. \end{aligned}$$

点评 从本题中所用的三种算法可看到精确计算用二项分布, 泊松分布及中心极限定理用于近似计算. 若 $n \geq 20, p \leq 0.05$ 时, 利用泊松分布来作为二项分布的近似计算则近似程度较好.

【5-2】 独立地多次测量一个物理量, 每次测量产生的随机误差都服从 $(-1, 1)$ 内的均匀分布. 若把 75 次测量的算术平均 $\bar{X}$ 作为测量结果, 分别用切比雪夫不等式和中心极限定理估计 $\bar{X}$ 与真值之差的绝

对值小于 $\frac{1}{9}$ 的概率.

分析 切比雪夫不等式的一个重要应用,就是估计随机变量落在区间 (a,b) 内的概率 $P(a < X < b)$. 关键是将待估概率 $P(a < X < b)$ 化为 $P(|X - E(X)| < \epsilon)$ [或 $P(|X - E(X)| \geq \epsilon)$] 的形式,方法是将不等式 $a < X < b$ 的各端同减去 $E(X)$. 用切比雪夫不等式估计随机变量落入区间内概率,所得的结果虽然比较粗糙,但其优点是无需知道 X 的分布,只要知道其期望 $E(X)$ 和方差 $D(X)$ 就行了.

独立同分布中心极限定理对分布无限制,只要求 $\{X_n\}$ 相互独立同分布,期望和方差有限,不论 X_n 是离散型或是连续型随机变量均可使用.

在独立同分布随机变量序列的中心极限定理的应用中,当涉及求独立同分布随机变量和的有关概率时,有以下几个步骤:

① 正确选取满足定理条件的独立同分布的随机变量序列 $\{X_n\}$;

② 将 $\sum_{i=1}^n X_i$ 标准化: $\frac{\sum_{i=1}^n X_i - nE(X_i)}{\sqrt{nD(X_i)}}$;

③ 由公式

$$\lim_{n \rightarrow \infty} P \left\{ \frac{\sum_{i=1}^n X_i - nE(X_i)}{\sqrt{nD(X_i)}} \leq x \right\} = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x)$$

计算有关概率;

④ 一般使用如下公式:

$$P(a \leq \sum_{i=1}^n X_i \leq b) \approx \Phi\left(\frac{b - nE(X_i)}{\sqrt{nD(X_i)}}\right) - \Phi\left(\frac{a - nE(X_i)}{\sqrt{nD(X_i)}}\right).$$

解 设第 i 次测量的结果为 $X_i = a + Y_i (i=1, 2, \dots, n)$, 其中 a 为真值, Y_i 为测量误差. 由题意可知, Y_i 服从 $(-1, 1)$ 内的均匀分布. 利用均匀分布的相应结果,可得 $E(Y_i) = 0$ 且 $D(Y_i) = \frac{1}{3}$. 又算术平均

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{n} \sum_{i=1}^n (a + Y_i) = a + \frac{1}{n} \sum_{i=1}^n Y_i,$$

且求得

$$E(\bar{X}) = E\left(a + \frac{1}{n} \sum_{i=1}^n Y_i\right) = a + \frac{1}{n} \sum_{i=1}^n E(Y_i) = a,$$

$$D(\bar{X}) = D\left(a + \frac{1}{n} \sum_{i=1}^n Y_i\right) = \frac{1}{n^2} \sum_{i=1}^n D(Y_i) = \frac{1}{3n}.$$

(1) 利用切比雪夫不等式求解

已知 $n=75, D(\bar{X}) = \frac{1}{225}$, 有

$$\begin{aligned} P\left(|\bar{X} - a| < \frac{1}{9}\right) &= P\left(|\bar{X} - E(\bar{X})| < \frac{1}{9}\right) \\ &\geq 1 - \frac{D(\bar{X})}{\left(\frac{1}{9}\right)^2} = 1 - \frac{81}{225} \\ &= 0.64. \end{aligned}$$

(2) 利用独立同分布中心极限定理求解

$$\begin{aligned} P\left(|\bar{X} - a| < \frac{1}{9}\right) &= P\left[\frac{|\bar{X} - E(\bar{X})|}{\sqrt{D(\bar{X})}} < \frac{1}{9\sqrt{D(\bar{X})}}\right] \\ &\approx 2\Phi\left(\frac{1}{9\sqrt{D(\bar{X})}}\right) - 1 \\ &= 2\Phi\left(\frac{5}{3}\right) - 1 \approx 2\Phi(1.67) - 1 \\ &\approx 0.905. \end{aligned}$$

【5-3】 为了确定事件 A 的概率, 进行了一系列试验. 在 100 次试验中, 事件 A 发生了 36 次. 若取频率 0.36 作为事件 A 的概率 p 的近似值, 求误差小于 0.05 的概率.

分析 本题要利用德莫佛·拉普拉斯中心极限定理求解. 这个定理是独立同分布中心极限定理的特例, 实际上引入随机变量

$$X_i = \begin{cases} 1 & (\text{第 } i \text{ 次试验中事件 } A \text{ 发生}), \\ 0 & (\text{第 } i \text{ 次试验中事件 } A \text{ 未发生}), \end{cases} \quad i = 1, 2, \dots, n,$$

则在 n 重贝努利试验中, 事件 A 发生的总次数 X 为 $X = \sum_{i=1}^n X_i$.

又因 $X_1, X_2, \dots, X_n$ 独立同分布, 且 $E(X_i) = p, D(X_i) = p(1-p)$, 则由独立同分布中心极限定理, 有

$$\begin{aligned}\lim_{n \rightarrow \infty} P\left(\frac{X - np}{\sqrt{np(1-p)}} \leq x\right) &= \lim_{n \rightarrow \infty} P\left[\frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}} \leq x\right] \\ &= \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x).\end{aligned}$$

这个定理表明了, 服从二项分布的随机变量, 经标准化后, 其极限分布为标准正态分布. 这就为二项分布提供了一种十分简便的近似计算概率的方法, 即若 $X \sim B(n, p)$, 当 n 充分大时, 有

$$\begin{aligned}P(a \leq X \leq b) &= \sum_{k=a}^b C_n^k p^k (1-p)^{n-k} \\ &\approx \Phi\left(\frac{b - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - np}{\sqrt{np(1-p)}}\right).\end{aligned}$$

解 设事件 A 在一次试验中发生的概率为 $P(A) = p$, 并且令

$$X_i = \begin{cases} 1 & (\text{第 } i \text{ 次试验中 } A \text{ 发生}), \\ 0 & (\text{第 } i \text{ 次试验中 } A \text{ 未发生}), \end{cases} \quad i = 1, 2, \dots, n.$$

显然有 $P(X_i = 1) = p, P(X_i = 0) = 1 - p$ 以及 $E(X_i) = p, D(X_i) = p(1-p)$.

又设 X_n 为 n 次试验中事件 A 发生的次数, 有 $X = \sum_{i=1}^n X_i$. 则频率

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \text{ 且算得 } E(\bar{X}) = p, D(\bar{X}) = \frac{p(1-p)}{n}.$$

已知 $p = 0.36, n = 100$, 根据德莫佛·拉普拉斯中心极限定理求得

$$\begin{aligned}P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - p\right| < 0.05\right) &= P\left[\frac{\left|\frac{1}{n} \sum_{i=1}^n X_i - p\right|}{\sqrt{\frac{p(1-p)}{n}}} < \frac{0.05}{\sqrt{\frac{p(1-p)}{n}}}\right] \\ &\approx 2\Phi\left[\frac{0.05\sqrt{n}}{\sqrt{p(1-p)}}\right] - 1 \approx 2\Phi(1.042) - 1\end{aligned}$$

≈ 0.706 .

【5-4】 一通信系统拥有 50 台相互独立起作用的交换机,在系统运行期间,每台交换机能清晰接受信号的概率为 0.9. 系统正常工作时,要求能清晰接受信号的交换机至少 45 台. 求该通信系统能正常工作的概率.

解 设

$$X_i = \begin{cases} 1 & (\text{第 } i \text{ 台交换机正常工作}), \\ 0 & (\text{第 } i \text{ 台交换机不正常工作}), \end{cases} \quad i = 1, 2, \dots, 50.$$

显然有 $P(X_i=1)=0.9, P(X_i=0)=0.1$ 且 $E(X_i)=0.9, D(X_i)=0.09$.

又设 X 为正常工作交换机的总台数,有 $X = \sum_{i=1}^{50} X_i$. 由于 $X_1, X_2, \dots, X_{50}$ 是相互独立的,因此 X 服从二项分作 $B(50, 0.9)$ 以及 $E(X) = np = 50 \times 0.9 = 45, D(X) = np(1-p) = 50 \times 0.9 \times 0.1 = 4.5$.

利用德莫佛-拉普拉斯中心极限定理,可求得

$$\begin{aligned} P(X \geq 45) &= P\left(\sum_{i=1}^{50} X_i \geq 45\right) \\ &= P\left[\frac{\sum_{i=1}^{50} X_i - E(X)}{\sqrt{D(X)}} \geq \frac{45 - E(X)}{\sqrt{D(X)}}\right] \\ &\approx 1 - \Phi\left(\frac{45 - E(X)}{\sqrt{D(X)}}\right) \\ &= 1 - \Phi(0) = 0.5. \end{aligned}$$

【5-5】 某大学共有 4900 个学生,已知每天晚上每个学生到阅览室去学习的概率为 0.1. 问阅览室要准备多少个座位,才能以 99% 的概率保证每个去阅览室的学生都有座位?

分析 假设每个学生是否去阅览室是相互独立的,本问题则是已

知 $X_n = \sum_{i=1}^n X_i$ 取值的概率 $P(0 \leq X \leq k) = 0.99$, 然后求出 k . 方法是先利用德莫佛-拉普拉斯中心极限定理将所求的概率转换为标准正态分

布 $\Phi(x)$ 与 k 有关的函数值 $\Phi[a(k)]$, 通过查表求出 $a(k)$, 然后再解出 k .

解 设

$$X_i = \begin{cases} 1 & (\text{第 } i \text{ 个学生去阅览室}), \\ 0 & (\text{第 } i \text{ 个学生不去阅览室}), \end{cases} \quad i = 1, 2, \dots, n.$$

所以有 $P(X_i=1)=0.1, P\{X_i=0\}=0.9$ 且 $E(X_i)=0.1, D(X_i)=0.09$.

又设去阅览室学生人数总和 $X = \sum_{i=1}^n X_i$, 则 X 服从二项分布 $B(n, p)$. 当 $n=4900, p=0.1$ 时, 有 $E(X)=np=490, D(X)=np(1-p)=441$.

利用德莫佛-拉普拉斯中心极限定理, 可解得

$$\begin{aligned} P(0 \leq X \leq k) &= P\left(\frac{-E(X)}{\sqrt{D(X)}} \leq \frac{X-E(X)}{\sqrt{D(X)}} \leq \frac{k-E(X)}{\sqrt{D(X)}}\right) \\ &= P\left(\frac{-490}{21} \leq \frac{X_n-490}{21} \leq \frac{k-490}{21}\right) \\ &\approx \Phi\left(\frac{k-490}{21}\right) - \left(1 - \Phi\left(\frac{490}{21}\right)\right) \\ &= \Phi\left(\frac{k-490}{21}\right) + \Phi\left(\frac{490}{21}\right) - 1. \end{aligned}$$

根据题意, 有 $0.99 = P(0 \leq X \leq k) \approx \Phi\left(\frac{k-490}{21}\right) + \Phi(23.3) - 1$

$$\Phi\left(\frac{k-490}{21}\right) \approx 0.99,$$

查表得

$$\frac{k-490}{21} = 2.325,$$

最后解得

$$k = 21 \times 2.325 + 490 = 538.85,$$

即至少要准备 539 个座位, 才能以 99% 的概率保证每个去阅览室的學生都有座位.

【5-6】 一批种子中良种占 $1/6$, 从中任取 6 000 粒. 问: 能以 0.99

的概率保证其中良种的比例与 $1/6$ 相差多少？此时相应的良种数落在哪个范围？

解 设

$$X_i = \begin{cases} 1 & (\text{第 } i \text{ 粒种子是良种}), \\ 0 & (\text{第 } i \text{ 粒种子不是良种}), \end{cases} \quad i = 1, 2, \dots, n.$$

$$\text{有 } P(X_i=1)=\frac{1}{6}, P(X_i=0)=\frac{5}{6}, \text{ 且 } E(X_i)=\frac{1}{6}, D(X_i)=\frac{5}{36}.$$

再设 6 000 粒种子中的良种总数 $X = \sum_{i=1}^n X_i$, 则 X 服从二项分布

$$B(n, p), \text{ 其中 } n=6\,000, p=\frac{1}{6}. \text{ 算得 } E(X)=np=6\,000 \times \frac{1}{6}=1\,000,$$

$$D(X)=np(1-p)=6\,000 \times \frac{1}{6} \times \frac{5}{6} = \frac{2\,500}{3} \approx 833.3.$$

由题意以及利用中心极限定理可得

$$\begin{aligned} P\left(\left|\frac{X}{6\,000} - \frac{1}{6}\right| \leq k\right) &= P\left[\frac{\left|\frac{X}{6\,000} - E\left(\frac{X}{6\,000}\right)\right|}{\sqrt{D\left(\frac{X}{6\,000}\right)}} \leq \frac{k}{\sqrt{D\left(\frac{X}{6\,000}\right)}}\right] \\ &\approx 2\Phi\left[\frac{k}{\sqrt{D\left(\frac{X}{6\,000}\right)}}\right] - 1 \\ &= 2\Phi\left(\frac{6\,000k}{28.87}\right) - 1. \end{aligned}$$

由

$$2\Phi\left(\frac{6\,000k}{28.87}\right) - 1 = 0.99$$

解得

$$\Phi\left(\frac{6\,000k}{28.87}\right) = 0.995,$$

通过查表并求得

$$\frac{6\,000k}{28.87} = 2.57,$$

$$k = \frac{2.57 \times 28.87}{6\,000} \approx 0.012.$$

另外由

$$\left| \frac{X}{6\,000} - \frac{1}{6} \right| \leqslant 0.012$$

解得

$$928 \leqslant X \leqslant 1\,072,$$

即能以 0.99 的概率保证其中良种的比例与 $\frac{1}{6}$ 相差 0.012, 此时相应良种数落在 $[928, 1\,072]$ 之间.

【5-7】 煤气公司供应某地区 1 000 户居民用气, 各户用气情况相互独立. 已知每户每日用气量(单位: 度)在 $[0, 20]$ 之间上均匀分布. 求:

(1) 这 1 000 户居民每日用气量超过 10 100 度的概率;

(2) 要求以 99% 的概率保证该地区居民能正常用上煤气, 煤气公司每日至少需向该地区供应多少度煤气?

解 设 X_i 为第 i 户居民的每日用气量 ($i=1, 2, \dots, n$). 根据题意可知, X_i 服从 $[0, 20]$ 上的均匀分布且 $E(X_i)=10, D(X_i)=\frac{100}{3}$.

再设 1 000 户居民的用气总量为 $X = \sum_{i=1}^n X_i$, 有 $E(X) = nE(X_i) = 1\,000 \times 10 = 10\,000, D(X) = nD(X_i) = \frac{100\,000}{3}$, 其中 $n = 1\,000$.

(1) 利用独立同分布中心极限定理, 可得

$$\begin{aligned} P(X \geqslant 10\,100) &= 1 - P(X < 10\,100) \\ &= 1 - P\left[\frac{X - 10\,000}{\sqrt{\frac{100\,000}{3}}} < \frac{10\,100 - 10\,000}{\sqrt{\frac{100\,000}{3}}} \right] \\ &\approx 1 - \Phi(0.5478) \approx 1 - 0.7088 = 0.2912. \end{aligned}$$

(2) 根据题意, 有

$$P\{X \leqslant n\} = P\left[\frac{X - 10\,000}{\sqrt{\frac{100\,000}{3}}} \leqslant \frac{n - 10\,000}{\sqrt{\frac{100\,000}{3}}} \right]$$

$$\approx \Phi\left(\frac{n-10\,000}{182.57}\right) \geqslant 0.99,$$

查表并解得

$$\frac{n-10\,000}{182.57} \geqslant 2.33,$$

$$n \geqslant 2.33 \times 182.57 + 10\,000 = 10\,425.4,$$

即 1 000 户居民每日用气量超过 10 100 度的概率为 0.291 2, 若要求以 99% 的概率保证该地区居民能正常用上煤气, 煤气公司每日至少需向该地区供气 10 425.4 度.

【5-8】 一射手打靶, 他得 0 分、2 分、3 分、4 分、5 分的概率分别为 0.1, 0.2, 0.2, 0.4. 此射手独立打靶 200 次, 求其得分在 650 至 750 分之间的概率.

解 设 X_i 为第 i 次射击所得的分数 ($i=1, 2, \dots, 200$), X_i 的分布律为

X_i	0	2	3	4	5
P	0.1	0.1	0.2	0.2	0.4

算得 $E(X_i)=3.6$, $E(X_i^2)=15.4$, $D(X_i)=E(X_i^2)-E^2(X_i)=15.4-3.6^2=2.44$.

另设 X 为 200 次射击所得的总分, 即有 $X = \sum_{i=1}^{200} X_i$, 此时 $E(X) = \sum_{i=1}^{200} E(X_i) = 200 \times 3.6 = 720$ 和 $D(X) = \sum_{i=1}^{200} D(X_i) = 200 \times 2.44 = 488$.

利用独立同分布中心极限定理可得所求的概率

$$\begin{aligned} P(650 \leqslant X \leqslant 750) &= P\left(\frac{650-E(X)}{\sqrt{D(X)}} \leqslant \frac{X-E(X)}{\sqrt{D(X)}} \leqslant \frac{750-E(X)}{\sqrt{D(X)}}\right) \\ &= P\left(\frac{650-720}{\sqrt{488}} \leqslant \frac{X-720}{\sqrt{488}} \leqslant \frac{750-720}{\sqrt{488}}\right) \\ &\approx \Phi(1.36) + \Phi(3.17) - 1 \\ &= 0.9131 + 0.9992 - 1 = 0.9123. \end{aligned}$$

【5-9】 售报员在“东方书报亭”内售报,每个过路人在此买报的概率为 $1/5$. 若有 600 人路过此书报亭,求售出报纸数目不超过 180 份的概率(假设路过报亭的买报人每人仅买 1 份报纸).

解 设

$$X_i = \begin{cases} 1 & (\text{第 } i \text{ 个过路人买报纸}), \\ 0 & (\text{第 } i \text{ 个过路人未买报纸}), \end{cases} \quad i = 1, 2, \dots, 600.$$

有 $P(X_i=1)=\frac{1}{5}, P(X_i=0)=\frac{4}{5}$ 和 $E(X_i)=\frac{1}{5}, D(X_i)=\frac{4}{25}$.

另设 X 为买报纸人的总数,即有 $X = \sum_{i=1}^{600} X_i$. 由于 $X_1, X_2, \dots, X_{600}$ 是相互独立的且服从同一分布,便有 X 服从二项分布 $B(n, p)$, 其中 $n=600, p=\frac{1}{5}$ 以及 $E(X) = np = 200, D(X) = np(1-p) = 96$.

利用德莫佛-拉普拉斯中心极限定理,可求得概率

$$\begin{aligned} P(0 \leq X \leq 180) &= P\left(\frac{-E(X)}{\sqrt{D(X)}} \leq \frac{X-E(X)}{\sqrt{D(X)}} \leq \frac{180-E(X)}{\sqrt{D(X)}}\right) \\ &= P\left(\frac{-200}{\sqrt{96}} \leq \frac{X-200}{\sqrt{96}} \leq \frac{180-200}{\sqrt{96}}\right) \\ &\approx \Phi\left(\frac{-20}{\sqrt{96}}\right) - \Phi\left(\frac{-200}{\sqrt{96}}\right) = \Phi(-2.04) - \Phi(-20.4) \\ &= \Phi(20.4) - \Phi(2.04) \approx 1 - 0.9793 \\ &= 0.0207. \end{aligned}$$

【5-10】 某运输公司有 500 辆汽车参加保险,在一年里汽车出事故的概率为 0.006,参保的汽车每年交 800 元的保费.若出事故保险公司最多可赔偿 50 000 元,试分别计算保险公司亏本的概率及一年中盈利不小于 200 000 元的概率.

分析 每辆汽车一年的保险费为 800 元,5 000 辆汽车一年的保险费为 $500 \times 800 = 40$ 万元,若保险公司不亏本即每年赔偿费用不能大于 40 万元,出事故的车辆数不能多于 $400\,000 \div 50\,000 = 8$ 辆,于是本题要求一年中出事故的车辆数大于 8 辆的概率.若公司要盈利 20 万元,则求出事故的车辆数不能大于 4 辆的概率.

解 设

$$X_i = \begin{cases} 1 & (\text{第 } i \text{ 辆汽车出事故}), \\ 0 & (\text{第 } i \text{ 辆汽车不出事故}), \end{cases} \quad i = 1, 2, \dots, 500.$$

由题意可知, $P(X_i = 1) = 0.006$, $P(X_i = 0) = 0.994$ 且有 $E(X_i) = 0.006$, $D(X_i) = 5.964 \times 10^{-3}$. 又设 X 为每年出事故汽车总数, 有 $X =$

$$\sum_{i=1}^{500} X_i.$$

由于 $X_1, X_2, \dots, X_{500}$ 是相互独立且均服从 0-1 分布, 所以 X 服从二项分布 $B(500, 0.006)$. 由二项分布的性质可知, $E(X) = np = 500 \times 0.006 = 3$, $D(X) = np(1-p) = 500 \times 5.964 \times 10^{-3} = 2.982$.

利用德莫佛-拉普拉斯中心极限定理可求得以下结果:

① 保险公司亏本的概率

$$\begin{aligned} P(X \geq 8) &= P\left(\frac{X - E(X)}{\sqrt{D(X)}} \geq \frac{8 - E(X)}{\sqrt{D(X)}}\right) \\ &= P\left(\frac{X - E(X)}{\sqrt{D(X)}} \geq \frac{8 - 3}{\sqrt{2.982}}\right) \\ &\approx 1 - \Phi(2.9) = 1 - 0.9981 = 0.0019. \end{aligned}$$

② 一年中盈利不小于 20 万元的概率

$$\begin{aligned} P(X \leq 4) &= P\left(\frac{X - E(X)}{\sqrt{D(X)}} \leq \frac{4 - E(X)}{\sqrt{D(X)}}\right) \\ &= P\left(\frac{X - 3}{\sqrt{2.982}} \leq \frac{4 - 3}{\sqrt{2.982}}\right) \\ &\approx \Phi(0.579) = 0.7190. \end{aligned}$$

【5-11】 假设在一本书中, 每页印刷错误的个数服从参数为 2 的泊松分布. 试用独立同分布中心极限定理估计一本 300 页的书中, 印刷错误超过 580 个的概率.

解 设第 i 页的印刷错误的个数为 $X_i (i = 1, 2, \dots, 300)$, 则 X_i 服从参数为 2 的泊松分布, 且有 $E(X_i) = 2$, $D(X_i) = 2$. 另设 X 为 300 页中印刷错误的总数, 显然有 $X = \sum_{i=1}^{300} X_i$. 另有 $E(X) = \sum_{i=1}^{300} E(X_i) =$

$300 \times 2 = 600, D(X) = \sum_{i=1}^{300} D(X_i) = 300 \times 2 = 600$. 根据题意以及独立同分布中心极限定理, 可求得

$$\begin{aligned} P(X \geq 580) &= P\left(\frac{X - E(X)}{\sqrt{D(X)}} \geq \frac{580 - E(X)}{\sqrt{D(X)}}\right) \\ &= P\left(\frac{X - E(X)}{\sqrt{D(X)}} \geq \frac{580 - 600}{\sqrt{600}}\right) \\ &\approx 1 - \Phi(-0.82) = \Phi(0.82) = 0.7939. \end{aligned}$$

【5-12】 银行为支付某日将到期的债券需准备一笔现金. 已知这批债券共发放了 500 张, 每张需付本息 1 000 元. 设持券人(1 人 1 券)到期日到银行领取本息的概率为 0.4. 问: 银行于该日应准备多少现金才能以 99.9% 的把握满足客户的兑换.

分析 若该日银行准备的现金金额为 n 元, X 为去银行兑换的人数, 则由题意可知 $P(1\,000X \leq n) = 0.999$, 然后利用中心极限定理求出 n .

解 设

$$X_i = \begin{cases} 1 & (\text{第 } i \text{ 个人到期去银行兑换}), \\ 0 & (\text{第 } i \text{ 个人到期未去银行兑换}), \end{cases} \quad i = 1, 2, \dots, 500.$$

由题意可知, $P(X_i = 1) = 0.4, P(X_i = 0) = 0.6$ 且 $E(X_i) = 0.4, D(X_i) = 0.24$. 另设该日到银行兑换人的总数为 X , 则有 $X = \sum_{i=1}^{500} X_i$. 由于 $X_1, X_2, \dots, X_{500}$ 是相互独立的且它们均服从 0-1 分布, 所以 X 服从二项分布 $B(500, 0.4)$. 由二项分布的性质可求得 $E(X) = np = 500 \times 0.4 = 200, D(X) = np(1-p) = 500 \times 0.24 = 120$.

利用德莫佛-拉普拉斯中心极限定理可求得

$$\begin{aligned} P(1\,000X \leq n) &= P\left(X \leq \frac{n}{1\,000}\right) \\ &= P\left[\frac{X - E(X)}{\sqrt{D(X)}} \leq \frac{\frac{n}{1\,000} - E(X)}{\sqrt{D(X)}}\right] \end{aligned}$$

$$= P\left[\frac{X-200}{\sqrt{120}} \leq \frac{\frac{n}{1000}-200}{\sqrt{120}}\right]$$

$$\approx \Phi\left[\frac{\frac{n}{1000}-200}{\sqrt{120}}\right],$$

再由 $\Phi\left[\frac{\frac{n}{1000}-200}{\sqrt{120}}\right]=0.999$ 查表得

$$\frac{\frac{n}{1000}-200}{\sqrt{120}} = 3.1$$

$$n = 233958.8,$$

即该日银行只需准备约 234 000 元就能以 99.9% 的把握满足客户的兑换。

【5-13】 从某厂生产的一批同型号的电子元件中抽取 395 件检验合格率。次品率 p 未知,要通过次品的相对频率来估计,估计的可靠度大于 95%。(1) 求绝对误差界 ϵ ; (2) 如果样本中有 $\frac{1}{10}$ 是次品,应对 p 作怎样的估计。

分析 设在 n 重贝努利试验中,事件 A 发生的次数为 X_n ,在一次试验中事件 A 发生的概率为 $P(A)=p$,则 $X_n \sim B(n, p)$ 。由德莫佛·拉普拉斯中心极限定理可知,当 n 很大时,有

$$X_n \sim N(np, np(1-p)),$$

于是用频率估计概率误差的可靠度

$$\begin{aligned}\beta &= P\left(\left|\frac{X_n}{n}-p\right| \leq \epsilon\right) = P\left(p-\epsilon \leq \frac{X_n}{n} \leq p+\epsilon\right) \\ &= P(np-n\epsilon \leq X_n \leq np+n\epsilon) \\ &= \Phi\left(\frac{n\epsilon}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{-n\epsilon}{\sqrt{np(1-p)}}\right) \\ &= 2\Phi\left(\epsilon \sqrt{\frac{n}{p(1-p)}}\right) - 1,\end{aligned}$$

得近似公式

$$\beta = P\left(\left|\frac{X_n}{n} - p\right| \leq \varepsilon\right) = 2\Phi\left(\varepsilon \sqrt{\frac{n}{p(1-p)}}\right) - 1. \quad (*)$$

其中有四个互相联系的指标: 试验总次数 n , 事件 A 一次试验中发生的概率 p , 误差界 ε 以及用频率估计概率所产生的误差的可靠度 β . 只要知道其中任意三个量, 就可求出另一个.

解 (1) 已知 $n=395$, 由公式 (*) 可得

$$\begin{aligned}\beta &= \Phi\left(\varepsilon \sqrt{\frac{n}{p(1-p)}}\right) - 1 \\ &= 2\Phi\left(\varepsilon \sqrt{\frac{395}{p(1-p)}}\right) - 1 = 0.95,\end{aligned}$$

故
$$\Phi\left(\varepsilon \sqrt{\frac{395}{p(1-p)}}\right) = 0.975,$$

查表得 $\varepsilon \sqrt{\frac{395}{p(1-p)}} = 1.96$, 从而 $\varepsilon = 1.96 \sqrt{\frac{p(1-p)}{395}}.$

由于 $p(1-p) = p - p^2 = \frac{1}{4} - \left(p - \frac{1}{2}\right)^2 \leq \frac{1}{4}$, 于是

$$\varepsilon = 1.96 \sqrt{\frac{p(1-p)}{395}} \leq 1.96 \sqrt{\frac{1}{4 \times 395}} = 0.05,$$

即绝对误差界 $\varepsilon \leq 0.05$.

(2) 由题(1)已得

$$P\left(\left|\frac{X_n}{n} - p\right| \leq 0.05\right) \geq 0.95,$$

此时, p 满足

$$\left|\frac{X_n}{n} - p\right| \leq 0.05,$$

而由题设可知 $\frac{X_n}{n} = \frac{1}{10} = 0.1$, 故 $0.05 \leq p \leq 0.15$, 即估计这批元件的次品率 p 在 $0.05 \sim 0.15$ 之间. p 在这个范围内的可靠度 $\beta = 0.95$.

B 习题精选

【5-14】 设 $X_1, X_2, \dots, X_n, \dots$ 相互独立, 且都服从参数为 $\lambda (\lambda > 0)$ 的指数分布, 则下列选项正确的是_____.

$$(A) \lim_{n \rightarrow \infty} P \left[\frac{\sum_{i=1}^n X_i - \frac{n}{\lambda}}{\frac{n}{\lambda^2}} \leq x \right] = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x);$$

$$(B) \lim_{n \rightarrow \infty} P \left[\frac{\lambda \sum_{i=1}^n X_i - n}{\sqrt{n}} \leq x \right] = \Phi(x);$$

$$(C) \lim_{n \rightarrow \infty} P \left[\frac{\sum_{i=1}^n X_i - \frac{1}{\lambda}}{\frac{1}{\lambda^2}} \leq x \right] = \Phi(x);$$

$$(D) \lim_{n \rightarrow \infty} P \left[\frac{\lambda \sum_{i=1}^n X_i - n}{n} \leq x \right] = \Phi(x).$$

【5-15】 设随机变量 $X_1, X_2, \dots, X_n, \dots$ 相互独立, $X = \sum_{i=1}^n X_i$, 由独立同分布中心极限定理可知, 当 n 充分大时, X 近似服从正态分布, 只要 $X_1, X_2, \dots, X_n, \dots$ _____.

- (A) 有相同的数学期望; (B) 有相同的方差;
(C) 服从同一指数分布; (D) 服从同一离散分布.

【5-16】 一随机变量 X 的 $E(X)=12, D(X)=9$. 用切比雪夫不等式估计:

- (A) $P(6 < X < 8) \geq$ _____;
(B) $P(3 < X < 21) \geq$ _____.

【5-17】 随机变量 X 的 $E(X)=\mu, D(X)=\sigma^2$, 用切比雪夫不等式估计: $P(|X-\mu| < k\sigma) \geq$ _____.

【5-18】 设 $X_i (i=1, 2, \dots, 50)$ 是相互独立的随机变量, 且都服从泊松分布, 参数 $\lambda=0.03$. 令 $X = \sum_{i=1}^{50} X_i$, 则用中心极限定理计算: $P(X>3) \approx$ _____ (精确到 2 位小数).

【5-19】 用切比雪夫不等式确定当掷一均匀硬币时, 需掷多少次, 才能保证使得出现正面的频率在 0.4 至 0.6 之间的概率不小于 90%, 并用中心极限定理计算同一问题.

【5-20】 某药厂宣称: 该厂生产的某种药品对于医治一种疑难的血液病的治愈率为 80%. 医院检验员任意抽查了 100 个服用此药品的病人, 如果其中多于 75 人治愈, 就接受这一断言, 认为该厂没有作虚假宣传, 否则就认为该厂宣传不符合实际. 问:

(1) 若实际上此药品对该病治愈率确实为 80%, 接受厂方宣传的概率是多少?

(2) 实际上此药品对该病治愈率是 70%, 接受厂方宣传的概率是多少?

【5-21】 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) 利用德莫佛·拉普拉斯中心极限定理, 求被盗索赔户不少于 14 户, 且不多于 30 户的概率.

【5-22】 某大型商场每天接待顾客 10 000 人, 设每位顾客的消费额(元)服从 $[10, 1000]$ 上的均匀分布, 且顾客的消费额是相互独立的, 试求该商场的消费额(元)在平均消费额上下浮动不超过 20 000 元的概率.

【5-23】 某厂生产的灯泡, 合格率为 0.96. 某天生产了 1 000 只灯泡, 试问能以 95% 的把握保证在这 1 000 只灯泡中合格灯泡的个数在哪个范围内?

答案与提示

【5-14】 B.

【5-15】 A 不知期望、方差是否有限, B 及 D 同 A. C 满足独立同分布及期望, 方差有限的要求, 故 C 入选.

【5-16】 A 填 $\frac{3}{4}$; B 填 $\frac{8}{9}$.

【5-17】 取 $\varepsilon = k\sigma$, $P(|X - \mu| < k\sigma) \geq 1 - \frac{\sigma^2}{k^2\sigma^2} = 1 - \frac{1}{k^2}$, 故填 $1 - \frac{1}{k^2}$.

【5-18】 由 $E(X_i) = 0.03$, $D(X_i) = 0.03$, 可得

$$\begin{aligned} P(X > 3) &= 1 - P(X \leq 3) \\ &\approx 1 - \Phi\left(\frac{3 - 0.03 \times 50}{\sqrt{0.03 \times 50}}\right) \\ &= 1 - \Phi(1.225) = 0.11. \end{aligned}$$

故填 0.11.

【5-19】 设需掷 n 次, 用 X_n 表示正面出现的次数, 则 $X_n \sim B\left(n, \frac{1}{2}\right)$,

由切比雪夫不等式得

$$\begin{aligned} P\left(0.4 < \frac{X_n}{n} < 0.6\right) &= P\left(\left|\frac{X_n}{n} - 0.5\right| < 0.1\right) \\ &\geq 1 - \frac{100}{4n} \geq 0.90, \end{aligned}$$

所以 $n \geq \frac{1000}{4} = 250$.

利用德莫佛·拉普拉斯中心极限定理可得

$$\begin{aligned} P\left(0.4 < \frac{X_n}{n} < 0.6\right) &= P\left[\frac{\left|X_n - \frac{n}{2}\right|}{\sqrt{n \times \frac{1}{2} \times \frac{1}{2}}} < \frac{0.1\sqrt{n}}{\sqrt{\frac{1}{2} \times \frac{1}{2}}}\right] \\ &= 2\Phi(0.2\sqrt{n}) - 1 \geq 0.90, \\ \Phi(0.2\sqrt{n}) &\geq 0.95, \quad 0.2\sqrt{n} \geq 1.645, \end{aligned}$$

所以 $n \geq 68$.

【5-20】 设 X 为服用此药品的病人中被治愈的人数, 则 $X \sim$

$B(100, p)$. 本题实际上利用德莫佛·拉普拉斯中心极限定理来近似地求 $P(X > 75)$.

$n=100$, 则 $E(X)=100p, D(X)=100p(1-p)$. 有

$$\begin{aligned} P(X > 75) &= P\left(\frac{X-100p}{\sqrt{100p(1-p)}} > \frac{75-100p}{\sqrt{100p(1-p)}}\right) \\ &\approx 1 - \Phi\left(\frac{75-100p}{\sqrt{100p(1-p)}}\right). \end{aligned}$$

对题(1) $P(X > 75) \approx 1 - \Phi\left(-\frac{5}{4}\right) = \Phi(1.25) = 0.8944$.

对题(2) $P(X > 75) \approx 1 - \Phi\left(\frac{5}{4.58}\right) = \Phi(1.09) = 0.1379$.

可见, 当厂方宣传符合实际时, 接受这一宣传的概率为 0.8944; 而当厂方宣传言过其实时, 接受这一宣传的概率为 0.1379.

【5-21】 (1) 由于 $X \sim B(100, 0.2)$, 故 X 的分布律为

$$P(X = k) = C_{100}^k (0.2)^k (0.8)^{100-k} \quad (k = 0, 1, 2, \dots, 100);$$

(2) 由于 $E(X) = np = 100 \times 0.2 = 20, D(X) = np(1-p) = 100 \times 0.2 \times 0.8 = 16$, 故

$$\begin{aligned} P(14 \leq X \leq 30) &\approx \Phi\left(\frac{30-20}{4}\right) - \Phi\left(\frac{14-20}{4}\right) \\ &= \Phi(2.5) - \Phi(-1.5) = \Phi(2.5) + \Phi(1.5) - 1 \\ &= 0.927. \end{aligned}$$

【5-22】 设第 i 位顾客的消费额为 $X_i (i=1, 2, \dots, 10\,000)$, 则该

商店的消费总额 $X = \sum_{i=1}^{10\,000} X_i$. 又 $E(X_i) = 450, D(X_i) = \frac{900^2}{12}$, 故

$E(X) = 45 \times 10^5, D(X) = 10^4 \times \frac{900^2}{12}$. 于是所求概率

$$\begin{aligned} &P(|X - 45 \times 10^5| \leq 20\,000) \\ &\approx \Phi\left[\frac{2 \times 10^4}{\sqrt{10^4 \times \frac{900^2}{12}}}\right] - \Phi\left[\frac{-2 \times 10^4}{\sqrt{10^4 \times \frac{900^2}{12}}}\right] \\ &= 2\Phi\left(\frac{4\sqrt{3}}{9}\right) - 1 \end{aligned}$$

$$\approx 2\Phi(0.77) - 1 = 0.56.$$

【5-23】 已知 $n=1000, p=0.96, \beta=0.95$, 要求 X_n 所在的范围. 先求 ϵ . 由公式(*)有

$$\beta = P\left(\left|\frac{X_n}{n} - p\right| \leq \epsilon\right) = 2\Phi\left(\epsilon \sqrt{\frac{n}{p(1-p)}}\right) - 1 = 0.95,$$

即

$$\Phi\left(\epsilon \sqrt{\frac{n}{p(1-p)}}\right) = 0.975,$$

查表得 $\epsilon \sqrt{\frac{n}{p(1-p)}} = 1.96$, 故

$$\epsilon = 1.96 \sqrt{\frac{p(1-p)}{n}}.$$

此时, X_n 满足

$$\left|\frac{X_n}{n} - p\right| \leq \epsilon,$$

将 $n=1000, p=0.96$ 代入上式得

$$948 < X_n < 972,$$

即合格的个数在 948 个到 972 个之间.

关键词

总体 样本 统计量 常用统计量 抽样分布 χ^2 分布 t 分布
 F 分布 参数估计 点估计 点估计的评价标准 区间估计 假设
检验 统计假设 两类错误 拒绝域

教学要求

(1) 理解总体、简单随机样本、统计量的概念,掌握样本均值、样本方差及样本矩的计算.

(2) 了解 χ^2 分布、 t 分布和 F 分布的定义及性质,了解分位数的概念并会查表计算.

(3) 了解正态总体常用统计量的分布.

(4) 理解参数的点估计概念,掌握用矩估计和极大似然估计法计算参数的估计量.

(5) 了解估计量的评价标准,并会验证.

(6) 理解区间估计的概念,掌握单个正态总体的均值和方差和置信区间的求法.

(7) 理解显著性检验的基本思想,掌握假设检验的基本步骤,了解假设检验可能产生的两类错误.

(8) 掌握单个正态总体的均值和方差的假设检验.

重点

总体、样本、统计量的概念; χ^2 分布、 t 分布、 F 分布的定义;正态总体常用统计量的分布;参数的矩估计和极大似然估计;估计量的无偏性、有效性;单个正态总体均值和方差的置信区间;正态总体参数的假

设检验.

难点

参数的极大似然估计法;假设检验所用统计量的选择及单、双侧假设检验的区分.

主要内容与方法提要

(1) 基本概念

1) 总体与样本

我们把研究对象的全体称为总体. 为研究方便起见,通常把一个带有确定分布的随机变量 X 称为总体. 组成总体的每个元素称为个体.

从总体 X 中随机抽取一部分个体 $X_1, X_2, \dots, X_n$, 称为来自总体 X 的一个容量为 n 的一个样本.

若 $X_1, X_2, \dots, X_n$ 相互独立并且每一 X_i 和 X 服从同一分布, 则称 $(X_1, X_2, \dots, X_n)$ 为简单随机样本.

2) 统计量

设 $(X_1, X_2, \dots, X_n)$ 为来自总体 X 的一个简单随机样本, $g(X_1, X_2, \dots, X_n)$ 为一个 n 元连续函数并且 g 中不含任何未知参数, 则称 g 为一个统计量.

3) 常用统计量

① 样本均值

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i;$$

② 样本方差

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2;$$

③ 样本 k 阶原点矩

$$A_k = \frac{1}{n} \sum_{i=1}^n X_i^k (k = 1, 2, \dots);$$

④ 样本 k 阶中心矩

$$B_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k (k = 1, 2, \dots).$$

4) 抽样分布

统计量的分布称为抽样分布.

① χ^2 分布

设 $X_1, X_2, \dots, X_n$ 相互独立, 且都服从标准正态分布 $N(0, 1)$, 则

称随机变量 $\chi^2 = \sum_{i=1}^n X_i^2$ 服从自由度为 n 的 χ^2 分布, 记作 $\chi^2 \sim \chi^2(n)$.

其密度函数

$$f(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} & (x > 0), \\ 0, & (x \leq 0). \end{cases}$$

性质 1 χ^2 分布具有可加性, 即若 $\chi_1^2 \sim \chi^2(n_1)$, $\chi_2^2 \sim \chi^2(n_2)$, 且 χ_1^2 与 χ_2^2 相互独立, 则 $\chi_1^2 + \chi_2^2 \sim \chi^2(n_1 + n_2)$.

性质 2 若 $\chi^2 \sim \chi^2(n)$, 则 $E(\chi^2) = n$, $D(\chi^2) = 2n$.

称满足 $P(\chi^2 \geq \chi_\alpha^2(n)) = \int_{\chi_\alpha^2(n)}^{+\infty} f(x) dx = \alpha$ ($0 < \alpha < 1$) 的点

$\chi_\alpha^2(n)$ 为 χ^2 分布的 α 分位点.

当 $n > 45$ 时, 近似地有 $\chi_\alpha^2(n) \approx \frac{1}{2}(z_\alpha + \sqrt{2n+1})^2$.

② T 分布

设随机变量 $X \sim N(0, 1)$, $Y \sim \chi^2(n)$, 且 X 与 Y 相互独立, 则随机

变量 $T = \frac{X}{\sqrt{\frac{Y}{n}}}$ 服从自由度为 n 的 T 分布, 记作 $T \sim t(n)$. T 分布密度函

数

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2}) \sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} \quad (-\infty < x < +\infty).$$

性质 1 T 分布的密度函数 $f(x)$ 为偶函数.

性质 2 当 n 充分大时($n > 45$), T 分布近似标准正态分布.

称满足 $P(T \geq t_\alpha(n)) = \int_{t_\alpha(n)}^{+\infty} f(x) dx = \alpha$ ($0 < \alpha < 1$) 的点 $t_\alpha(n)$ 为 T 分布的 α 分位点. T 分布具有对称性, 故 $t_{1-\alpha}(n) = -t_\alpha(n)$. 另外, 当 $n > 45$ 时, $t_\alpha(n) \approx z_\alpha$.

③ F 分布

设随机变量 X 与 Y 相互独立, 且分别服从自由度为 n, m 的 χ^2 分布, 即 $X \sim \chi^2(n), Y \sim \chi^2(m)$, 则称随机变量

$$F = \frac{\frac{X}{n}}{\frac{Y}{m}}$$

服从自由度为 (n, m) 的 F 分布, 记为 $F \sim F(n, m)$, 其中 n, m 分别称为第一、二自由度. F 分布的密度函数

$$f(x) = \begin{cases} \frac{\Gamma\left(\frac{n+m}{2}\right)\left(\frac{n}{m}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)\left(1+\frac{n}{m}x\right)^{\frac{n+m}{2}}} x^{\frac{n}{2}-1} & (x > 0), \\ 0 & (x \leq 0). \end{cases}$$

称满足 $P(F \geq F_\alpha(n, m)) = \int_{F_\alpha(n, m)}^{+\infty} f(x) dx = \alpha$ ($0 < \alpha < 1$) 的点 $F_\alpha(n, m)$ 为 F 分布的 α 分位点, 且有

$$F_{1-\alpha}(n, m) = \frac{1}{F_\alpha(m, n)}.$$

5) 正态总体的抽样分布

设 $X_1, X_2, \dots, X_n$ 为来自正态总体 $N(\mu, \sigma^2)$ 的样本, $\bar{X}$ 是样本均值, S^2 是样本方差, 则

$$\textcircled{1} \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right),$$

$$\textcircled{2} \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \sim \chi^2(n),$$

$$\textcircled{3} \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1),$$

$$\textcircled{4} T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1).$$

(2) 参数估计

利用统计量去估计总体的未知参数称为参数估计. 设 $(X_1, X_2, \dots, X_n)$ 为来自总体 X 的一个样本, θ 是总体的未知参数, 若用一个统计量 $\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$ 来估计 θ , 则称 $\hat{\theta}$ 为未知参数 θ 的估计量; 若一次抽取得到样本的观察值为 $(x_1, x_2, \dots, x_n)$, 则称 $\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$ 为 θ 的估计值.

1) 点估计

设 θ 是总体 X 的一个未知参数, 用统计量 $\hat{\theta}(X_1, X_2, \dots, X_n)$ 去估计 θ , 称为对 θ 作点估计. 常用点估计法有以下两种:

① 矩估计法

用样本矩估计总体矩, 从而得到未知参数的估计量称为参数 θ 的矩估计.

② 极大似然估计法

使得似然函数 $L(x_1, x_2, \dots, x_n; \theta)$ 取极大值的估计量称为未知参数 θ 的极大似然估计.

2) 点估计的评价标准

① 无偏性

设 $\hat{\theta}$ 为参数 θ 的估计量, 若 $E(\hat{\theta}) = \theta$, 则称 $\hat{\theta}$ 为 θ 的无偏估计量.

② 一致性

设 $\hat{\theta}$ 为参数 θ 的估计量, 若对任意给定的 $\epsilon > 0$, 都有

$$\lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| < \epsilon) = 1,$$

即 $\hat{\theta}$ 依概率收敛于参数 θ , 则称 $\hat{\theta}$ 为 θ 的一致估计量.

③ 有效性

设 $\hat{\theta}_1, \hat{\theta}_2$ 为 θ 的两个无偏估计量, 若 $D(\hat{\theta}_1) < D(\hat{\theta}_2)$, 则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效.

注: 样本均值 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ 是总体均值 $E(X)$ 的无偏、一致估计量, 样本方差 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ 是总体方差 $D(X)$ 的无偏、一致估计量.

3) 参数的区间估计

设 θ 是总体 X 的一个待估参数, $\hat{\theta}_1(X_1, X_2, \dots, X_n)$ 和 $\hat{\theta}_2(X_1, X_2, \dots, X_n)$ 是两个统计量 ($\hat{\theta}_1 < \hat{\theta}_2$), 把待估参数 θ 的取值范围估计在区间 $(\hat{\theta}_1, \hat{\theta}_2)$ 内, 称为参数的区间估计.

① 置信度与置信区间

设 θ 为总体 X 的未知参数, $(X_1, X_2, \dots, X_n)$ 为来自总体 X 的一个样本, 构造两个统计量 $\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \dots, X_n)$ 和 $\hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \dots, X_n)$, 使对任意给定的 $\alpha (0 < \alpha < 1)$, 有

$$P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = 1 - \alpha$$

成立, 则称 $(\hat{\theta}_1, \hat{\theta}_2)$ 为 θ 的一个置信度为 $1 - \alpha$ 的置信区间, α 为显著性水平, $\hat{\theta}_1$ 为置信下限, $\hat{\theta}_2$ 为置信上限.

注: 由于 $(\hat{\theta}_1, \hat{\theta}_2)$ 是随机区间, 而 θ 是一个客观存在的未知常数, 所以确切的解释应该是随机区间 $(\hat{\theta}_1, \hat{\theta}_2)$ 包含 θ 的可能性为 $1 - \alpha$.

② 正态总体参数的区间估计

设总体 $X \sim N(\mu, \sigma^2)$, $(X_1, X_2, \dots, X_n)$ 为来自总体 X 的一个样本, 置信度为 $1 - \alpha$.

(i) 方差 σ^2 已知, 求期望 μ 的区间估计

此时用于估计的统计量

$$U = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1),$$

则置信度为 $1-\alpha$ 的置信区间为

$$\left(\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right).$$

(ii) 方差 σ^2 未知, 求期望 μ 的区间估计
用于估计的统计量

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1),$$

则置信度为 $1-\alpha$ 的置信区间为

$$\left(\bar{X} - t_{\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{n}}, \bar{X} + t_{\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{n}} \right).$$

(iii) 均值 μ 已知, 求方差 σ^2 的区间估计
用于估计的统计量

$$\chi^2 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \sim \chi^2(n),$$

则置信度为 $1-\alpha$ 的置信区间为

$$\left[\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{\frac{\alpha}{2}}^2(n)}, \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{1-\frac{\alpha}{2}}^2(n)} \right].$$

(iv) 均值 μ 未知, 求方差 σ^2 的区间估计
用于估计的统计量

$$\chi^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1).$$

以置信度为 $1-\alpha$ 的置信区间为

$$\left(\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}^2(n-1)}, \frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)} \right).$$

(3) 假设检验

1) 统计假设

假设检验是统计推断的重要内容之一. 设总体 X 的分布已确定, 而分布中含有未知参数, 现对总体中的未知参数的值作出假设, 这种假设称为原假设 (记为 H_0). 然后从总体中抽取容量为 n 的样本, 并构造一种检验法则, 再利用样本按检验法则对假设 H_0 进行检验, 得出接受或拒绝假设 H_0 的结论, 称为参数的假设检验.

2) 两类错误

拒绝 H_0 还是接受 H_0 都要承担一定的风险, 前者有可能将正确的假设误认为是错误的, 后者可能将错误的假设误认为是正确的. 也就是说, 在假设检验中, 可能会犯下面的两类错误:

① 第 I 类错误 在 H_0 为真时, 作出拒绝 H_0 的判断称为犯第 I 类错误. 犯第 I 类错误的概率

$$P(\text{拒绝 } H_0 \mid \text{当 } H_0 \text{ 为真}) = \alpha.$$

② 第 II 类错误 在 H_0 为不真时, 作出接受 H_0 的判断称为犯第 II 类错误. 犯第 II 类错误的概率

$$P(\text{接受 } H_0 \mid \text{当 } H_0 \text{ 为不真}) = \beta.$$

当样本容量 n 一定时, α 和 β 不能同时减少, 减少其中一个, 另一个会增大. 要同时减小犯两类错误的概率, 只有增加样本容量 n . 在实际问题中总是采用控制犯第 I 类错误的概率 α 的方法.

3) 假设检验的基本步骤

① 据题意提出原假设 H_0 和备择假设 H_1 .

② 当 H_0 为真时, 构造检验统计量 $G(X_1, X_2, \dots, X_n)$, 以及确认 G 有确定的分布.

③ 由给定的显著性水平 α , 根据统计量 G 所服从的分布表, 定出临界值 λ , 使得

$$P(|G| > \lambda) = \alpha,$$

从而求出 H_0 的拒绝域

$$W = (-\infty, \lambda] \cup [\lambda, +\infty).$$

④ 由样本观察值计算 G 的统计值 G_0 .

⑤ 作出判断, 若 $G_0 \in W$, 则拒绝 H_0 . 若 $G_0 \notin W$, 则接受 H_0 .

4) 正态总体的假设检验

设总体 $X \sim N(\mu, \sigma^2)$, $\bar{x}$ 和 s^2 分别为样本观察值 $(x_1, x_2, \dots, x_n)$ 的均值和方差, 显著性水平为 $\alpha (0 < \alpha < 1)$.

① 方差 σ^2 已知, 检验期望 μ

若 $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$,

则检验统计值 $U_0 = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$,

拒绝域 $W = (-\infty, -z_{\frac{\alpha}{2}}] \cup [z_{\frac{\alpha}{2}}, +\infty)$.

若 $H_0: \mu = \mu_0, H_1: \mu > \mu_0$ ($H_1: \mu < \mu_0$),

则拒绝域 $W = [z_\alpha, +\infty)$ ($W = (-\infty, -z_\alpha]$).

② 方差 σ^2 未知, 检验期望 μ

若 $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$,

则检验统计值 $T_0 = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$,

拒绝域 $W = (-\infty, -t_{\frac{\alpha}{2}}(n-1)] \cup [t_{\frac{\alpha}{2}}(n-1), +\infty)$.

若 $H_0: \mu = \mu_0, H_1: \mu > \mu_0$ ($H_1: \mu < \mu_0$),

则拒绝域 $W = [t_\alpha(n-1), +\infty)$ ($W = (-\infty, -t_\alpha(n-1)]$).

③ 期望 μ 已知, 检验方差 σ^2

若 $H_0: \sigma^2 = \sigma_0^2, H_1: \sigma^2 \neq \sigma_0^2$,

则检验统计值 $\chi_0^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{\sigma_0^2}$

拒绝域 $W = [0, \chi_{1-\frac{\alpha}{2}}^2(n)] \cup [\chi_{\frac{\alpha}{2}}^2(n), +\infty)$.

若 $H_0: \sigma^2 = \sigma_0^2, H_1: \sigma^2 > \sigma_0^2$ ($H_1: \sigma^2 < \sigma_0^2$),

则拒绝域 $W = [\chi_\alpha^2(n), +\infty)$ ($W = [0, \chi_{1-\alpha}^2(n)]$).

④ 期望 μ 未知, 检验方差 σ^2

若 $H_0: \sigma^2 = \sigma_0^2, H_1: \sigma^2 \neq \sigma_0^2$,

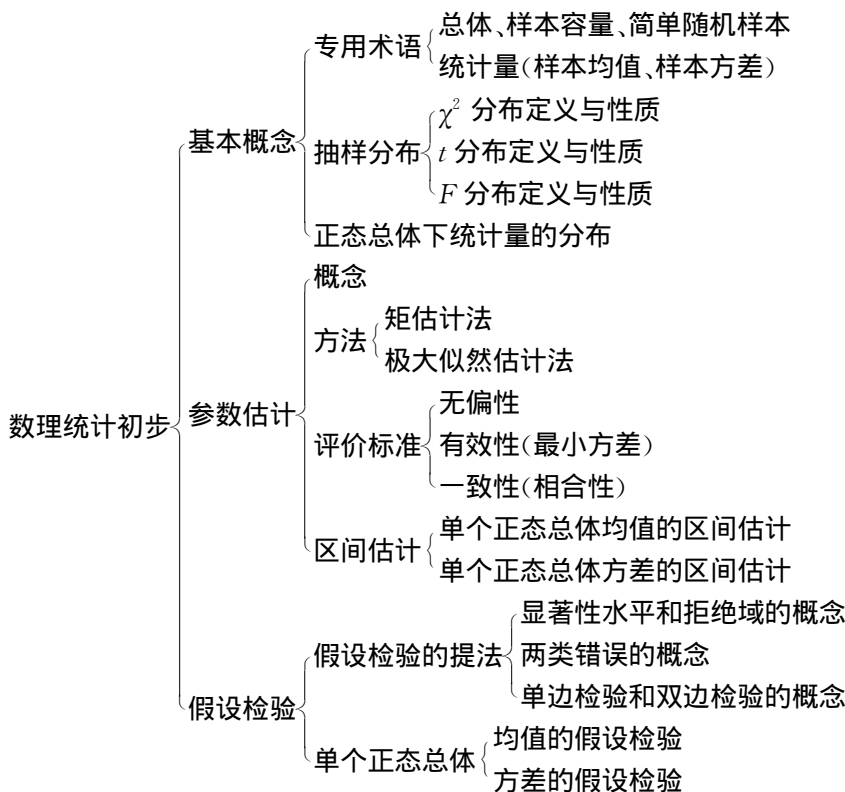
则检验统计值 $\chi_0^2 = \frac{(n-1)S^2}{\sigma^2}$,

拒绝域 $W = [0, \chi_{1-\frac{\alpha}{2}}^2(n-1)] \cup [\chi_{\frac{\alpha}{2}}^2(n-1), +\infty)$.

若 $H_0: \sigma^2 = \sigma_0^2, H_1: \sigma^2 > \sigma_0^2 (H_1: \sigma^2 < \sigma_0^2)$,

则拒绝域 $W = [\chi_{\alpha}^2(n-1), +\infty) (W = [0, \chi_{1-\alpha}^2(n-1)])$.

本章知识结构



A 例题精解

【6-1】 设 $\bar{X}$ 是来自总体 $X \sim N(80, 12^2)$ 的一个容量为 8 的样本均

值,计算 $P(68 < \bar{X} < 92)$.

分析 设 $X_1, X_2, \dots, X_n$ 是从正态总体 $N(\mu, \sigma^2)$ 中抽取的一个简单随机样本, 则样本均值 $\bar{X}$ 服从正态分布 $N\left(\mu, \frac{\sigma^2}{n}\right)$. 而经过标准化变换后, 有 $\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$.

解 已知 $E(X) = \mu = 80, D(X) = \sigma^2 = 12^2$, 则 $E(\bar{X}) = 80, D(\bar{X}) = \frac{\sigma^2}{n} = \frac{12^2}{8} = 18$. 解得

$$\begin{aligned} P(68 < \bar{X} < 92) &= P\left(\frac{68-80}{\sqrt{18}} < \frac{\bar{X}-80}{\sqrt{18}} < \frac{92-80}{\sqrt{18}}\right) \\ &= \Phi(2.83) - \Phi(-2.83) \\ &= 2\Phi(2.83) - 1 = 2 \times 0.9977 - 1 \\ &= 0.9954. \end{aligned}$$

【6-2】 总体 $X \sim N(30, 9^2)$, 从总体 X 抽取一个容量为 36 的样本. 求: 样本均值与总体均值的差的绝对值小于 2 的概率.

解 已知 $E(X) = \mu = 30, D(X) = \sigma^2 = 9^2$, 有 $E(\bar{X}) = \mu = 30, D(\bar{X}) = \frac{\sigma^2}{n} = \frac{9^2}{36} = \frac{9}{4}$. 由题意可解得

$$\begin{aligned} P(|\bar{X} - \mu| < 2) &= P\left(\frac{|\bar{X} - E(\bar{X})|}{\sqrt{D(\bar{X})}} < \frac{2}{\sqrt{D(\bar{X})}}\right) \\ &= 2\Phi\left(\frac{2}{\sqrt{D(\bar{X})}}\right) - 1 = 2\Phi\left(\frac{4}{3}\right) - 1 \\ &= 2 \times 0.9082 - 1 = 0.8164. \end{aligned}$$

【6-3】 设总体 $X \sim N(\mu, 4)$, 若要以 95% 的概率保证样本均值 $\bar{X}$ 与总体期望 μ 的偏差小于 0.1, 问样本容量 n 应取多大?

解 由于 $X \sim N(\mu, 4)$, 故 $\bar{X} \sim N\left(\mu, \frac{4}{n}\right)$, 所以

$$P(|\bar{X} - \mu| < 0.1) = P\left[\frac{|\bar{X} - \mu|}{\frac{2}{\sqrt{n}}} < \frac{0.1}{\frac{2}{\sqrt{n}}}\right]$$

$$= 2\Phi(0.05\sqrt{n}) - 1 = 0.95.$$

由 $\Phi(0.05\sqrt{n}) = 0.975$, 查表得 $0.05\sqrt{n} = 1.96$, 故 $n = 1536.64$, 可取 $n = 1537$.

【6-4】 设某厂生产的电子元件的使用寿命 $X \sim N(1000, \sigma^2)$ (单位: 小时). 现抽取 9 个元件测试, 得到 $\bar{x} = 940, s^2 = 100^2$, 问 $P(\bar{X} < 940)$ 是多少?

分析 由于本题中 σ^2 未知, 故不能用 $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ 来解题, 可利用

$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$ 服从自由度为 $n-1$ 的 t 分布来求解.

解 由于

$$\begin{aligned} P(\bar{X} < 940) &= P\left[\frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} < \frac{940 - 1000}{\frac{100}{\sqrt{9}}}\right] \\ &= P\left[\frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} < -1.8\right] = \alpha, \end{aligned}$$

并且 $T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$, 利用 t 分布的对称性可有

$$P(T < -t_\alpha(n-1)) = P(T > t_\alpha(n-1)) = \alpha.$$

令 $t_\alpha(8) = 1.8 (n=9)$, 由于 t 分布表上查不到相应 α 的值, 只能利用相邻的两个值来采用线性插值法计算 α 的值. 由 $t_{0.10}(8) = 1.3968, t_{0.05}(8) = 1.8595$, 得 $\alpha = 0.1081t - 0.1510$. 当 $t = 1.8$ 时, 解得 $\alpha \approx 0.044$.

【6-5】 设 $X_1, X_2, \dots, X_{10}$ 为总体 $X \sim N(\mu, \sigma^2)$ 的一个样本, 试求:

$$(1) P\left(0.26\sigma^2 \leq \frac{1}{10} \sum_{i=1}^{10} (X_i - \bar{X})^2 \leq 2.3\sigma^2\right);$$

$$(2) P\left(0.26\sigma^2 \leq \frac{1}{10} \sum_{i=1}^{10} (X_i - \mu)^2 \leq 2.3\sigma^2\right).$$

解 (1) 求解本题要利用以下结果:

$$\frac{\sum_{i=1}^{10} (X_i - \bar{X})^2}{\sigma^2} \sim \chi^2(9).$$

可解得

$$\begin{aligned} & P\left(0.26\sigma^2 \leq \frac{1}{10} \sum_{i=1}^{10} (X_i - \bar{X})^2 \leq 2.3\sigma^2\right) \\ &= P\left[2.6 \leq \frac{\sum_{i=1}^{10} (X_i - \bar{X})^2}{\sigma^2} \leq 23\right] \\ &= P(2.6 \leq \chi^2(9) \leq 23) \\ &= P(\chi^2(9) > 2.6) - P(\chi^2(9) > 23) \\ &= 0.975 - 0.005 = 0.97. \end{aligned}$$

(2) 本题则要利用以下结果:

$$\frac{\sum_{i=1}^{10} (X_i - \mu)^2}{\sigma^2} \sim \chi^2(10).$$

可解得

$$\begin{aligned} & P\left(0.26\sigma^2 \leq \frac{1}{10} \sum_{i=1}^{10} (X_i - \mu)^2 \leq 2.3\sigma^2\right) \\ &= P\left(2.6 \leq \sum_{i=1}^{10} (X_i - \mu)^2 \leq 23\right) \\ &= P(2.6 \leq \chi^2(10) \leq 23) \\ &= P(\chi^2(10) > 2.6) - P(\chi^2(10) > 23) \\ &= 0.99 - 0.01 = 0.98. \end{aligned}$$

【6-6】 设 $X_1, X_2, \dots, X_{10}$ 为来自总体 $X \sim N(0, 0.3^2)$ 的样本, 求

概率 $P\left\{\sum_{i=1}^{10} X_i^2 > 1.44\right\}$.

分析 因为 $X_1, X_2, \dots, X_{10}$ 是一个简单随机样本, 因此 $X_1, X_2, \dots, X_{10}$ 相互独立且与总体 X 有相同分布; 然后对 X_i 进行标准化变换, 再利用 χ^2 分布的定义和性质即可求得相应的结果.

解 已知 $X \sim N(0, 0.3^2)$, 则 $X_i \sim N(0, 0.3^2)$. 标准化变换后, 得 $\frac{X_i}{0.3} \sim N(0, 1)$. 由 χ^2 分布的定义得 $\left(\frac{X_i}{0.3}\right)^2 \sim \chi^2(1)$. 再利用 χ^2 分布的可加性, 有 $\sum_{i=1}^{10} \left(\frac{X_i}{0.3}\right)^2 \sim \chi^2(10)$. 最后解得

$$\begin{aligned} P\left(\sum_{i=1}^{10} X_i^2 > 1.44\right) \\ &= P\left(\sum_{i=1}^{10} \left(\frac{X_i}{0.3}\right)^2 > \frac{1.44}{0.3^2}\right) \\ &= P(\chi^2(10) > 16) \\ &= 0.1. \end{aligned}$$

【6-7】 设 X_1, X_2, X_3, X_4 是来自正态总体 $N(0, 1)$ 的样本. 设

$$X = a(X_1 - 2X_2) + b(3X_3 - 4X_4)^2,$$

则当 a, b 为何值时, X 服从 χ^2 分布? 并求出其自由度.

解 由于 $X_1 - 2X_2 \sim N(0, 5)$, $3X_3 - 4X_4 \sim N(0, 25)$, 因此有 $\frac{1}{\sqrt{5}}(X_1 - 2X_2) \sim N(0, 1)$ 及 $\frac{1}{5}(3X_3 - 4X_4)^2 \sim \chi^2(1)$. 再由 χ^2 分布的可加性可得 $\frac{1}{5}(X_1 - 2X_2)^2 + \frac{1}{25}(3X_3 - 4X_4)^2 \sim \chi^2(2)$. 应取 $a = \frac{1}{5}, b = \frac{1}{25}$, 此时 X 服从自由度为 2 的 χ^2 分布.

点评 本题的解题方法实际上是利用 χ^2 分布的定义和性质, 应将 X 配成 n 个服从标准正态分布的独立随机变量的平方和.

【6-8】 设总体 X, Y 相互独立, 且都服从正态分布 $N(0, 3^2)$, 而 $X_1, X_2, \dots, X_9$ 和 $Y_1, Y_2, \dots, Y_9$ 分别来自总体 X 和 Y 的简单随机样本. 问: 统计量

$$T = \frac{X_1 + X_2 + \dots + X_9}{\sqrt{Y_1^2 + Y_2^2 + \dots + Y_9^2}}$$

服从什么分布?

解 由于 $X_i \sim N(0, 3^2)$, 故 $\sum_{i=1}^9 X_i \sim N(0, 9 \times 3^2)$. 经标准化变换

后,得 $\frac{1}{9} \sum_{i=1}^9 X_i \sim N(0,1)$.

又 $Y_i \sim N(0,3^2)$, 故 $\frac{Y_i}{3} \sim N(0,1)$, 从而 $\left(\frac{Y_i}{3}\right)^2 \sim \chi^2(1)$. 再根据 χ^2 分布的可加性, 可知 $\sum_{i=1}^9 \left(\frac{Y_i}{3}\right)^2 = \frac{1}{9} \sum_{i=1}^9 Y_i^2 \sim \chi^2(9)$.

最后由 T 分布的定义可得

$$T = \frac{X_1 + X_2 + \cdots + X_9}{\sqrt{Y_1^2 + Y_2^2 + \cdots + Y_9^2}} = \frac{\frac{1}{9} \sum_{i=1}^9 X_i}{\sqrt{\frac{1}{9} \sum_{i=1}^9 Y_i^2}} \sim t(9),$$

即统计量 T 服从自由度为 9 的 t 分布.

【6-9】 设总体 X 的分布律为

X	0	1	2
P	$\frac{1}{2}$	p	$\frac{1}{2} - p$

来自 X 的样本为 $X_1, X_2, \cdots, X_n$, 求 p 的矩估计.

分析 对于分布中含有单个未知参数情形, 若能求出 $E(X) = g(\theta)$, (其中 θ 为未知), 则可用 $\bar{X}$ 来取代 $E(X)$, 以 $\hat{\theta}$ 取代 θ , 得 $\bar{X} = g(\hat{\theta})$. 然后解方程, 解得 $\hat{\theta} = g^{-1}(\bar{X})$, 即为未知参数 θ 的矩估计量.

解 先求出总体 X 的期望

$$E(X) = 0 \times \frac{1}{2} + 1 \times p + 2 \times \left(\frac{1}{2} - p\right) = p + 1 - 2p = 1 - p,$$

然后用样本的均值 $\bar{X}$ 取代 $E(X)$, $\hat{p}$ 取代 p , 解得

$$\bar{X} = 1 - \hat{p},$$

$$\hat{p} = 1 - \bar{X},$$

即为未知参数 p 的矩估计量.

【6-10】 设总体 X 的密度函数如下:

$$f(x) = \begin{cases} \theta x^{\theta-1}, & x \in (0, 1); \\ 0, & x \notin (0, 1). \end{cases}$$

$X_1, X_2, \dots, X_n$ 为来自 X 的样本, 试求: 未知参数 θ 的 (1) 矩估计; (2) 极大似然估计.

解 (1) 首先求出总体 X 的期望

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^1 x\theta x^{\theta-1}dx = \frac{\theta}{\theta+1},$$

然后用样本的均值 $\bar{X}$ 和 $\hat{\theta}$ 分别取代上式中的 $E(X)$ 和 θ , 可解得未知参数 θ 的矩估计量

$$\hat{\theta} = \frac{\bar{X}}{1 - \bar{X}}.$$

(2) 用极大似然估计法求解未知参数 θ 时, 关键是先要求出似然函数 L 及对数似然函数 $\ln L$.

设在一次试验中, 样本的观察值为 $x_1, x_2, \dots, x_n$, 即事件 $(X_1 \leq x_1), (X_2 \leq x_2), \dots, (X_n \leq x_n)$ 同时发生, 且 $X_1, X_2, \dots, X_n$ 相互独立, 此时样本 $X_1, X_2, \dots, X_n$ 的联合密度

$$\begin{aligned} f_n(x_1, x_2, \dots, x_n; \theta) &= f(x_1, \theta) \cdots f(x_n, \theta) \\ &= \theta x_1^{\theta-1} \cdots \theta x_n^{\theta-1} \\ &= \theta^n x_1^{\theta-1} \cdots x_n^{\theta-1}, \end{aligned}$$

得似然函数

$$L(\theta) = \theta^n x_1^{\theta-1} x_2^{\theta-1} \cdots x_n^{\theta-1},$$

上式两边取对数, 得

$$\ln L(\theta) = n \ln \theta + (\theta - 1) \sum_{i=1}^n \ln x_i,$$

对未知参数 θ 求导, 得

$$\frac{d(\ln L)}{d\theta} = \frac{n}{\theta} + \sum_{i=1}^n \ln x_i.$$

解对数似然方程:

$$\frac{d(\ln L)}{d\theta} = \frac{n}{\theta} + \sum_{i=1}^n \ln x_i = 0,$$

$$\theta = -\frac{n}{\sum_{i=1}^n \ln x_i}.$$

可容易看到上述解为唯一解. 因此未知参数 θ 的极大似然估计量

$$\hat{\theta} = -\frac{n}{\sum_{i=1}^n \ln X_i}.$$

点评 在一般情况下, 由对数似然方程组求得的是唯一解的话, 那么此解即为未知参数的极大似然估计.

【6-11】 设总体 $X \sim E(\lambda)$, $X_1, X_2, \dots, X_n$ 为来自 X 的样本, 试求: 未知参数 λ 的 (1) 矩估计; (2) 极大似然估计.

解 已知总体 X 服从参数为 λ 的指数分布,

$$X \sim f(x) = \begin{cases} \lambda e^{-\lambda x} & (x \geq 0), \\ 0 & (x < 0), \end{cases}$$

并且有 $E(X) = \frac{1}{\lambda}$.

(1) 用样本的均值 $\bar{X}$ 以及 $\hat{\lambda}$ 分别取代等式 $E(X) = \frac{1}{\lambda}$ 中的 $E(X)$ 和 λ , 可解得未知参数 λ 的矩估计量

$$\hat{\lambda} = \frac{1}{\bar{X}}.$$

(2) 设 $x_1, x_2, \dots, x_n$ 为总体 X 的一个样本观察值, 事件 $\{X_1 \leq x_1\}, \{X_2 \leq x_2\}, \dots, \{X_n \leq x_n\}$ 同时发生且相互独立, 得到样本 $X_1, X_2, \dots, X_n$ 的联合密度

$$\begin{aligned} f_n(x_1, \dots, x_n; \lambda) &= \lambda e^{-\lambda x_1} \lambda e^{-\lambda x_2} \cdots \lambda e^{-\lambda x_n} \\ &= \lambda^n e^{-\lambda \sum_{i=1}^n x_i} \quad (x_i \geq 0), \end{aligned}$$

则似然函数

$$L(\lambda) = \lambda^n e^{-\lambda \sum_{i=1}^n x_i},$$

两边取对数, 得

$$\ln L(\lambda) = n \ln \lambda - \lambda \sum_{i=1}^n x_i,$$

然后对 λ 求导得到对数似然方程, 再求解该方程, 解得

$$\frac{d(\ln L)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^n x_i = 0,$$

$$\lambda = \frac{1}{\frac{1}{n} \sum_{i=1}^n x_i}.$$

由于求得的是唯一解,故未知参数 λ 的极大似然估计量

$$\hat{\lambda} = \frac{1}{\bar{X}}.$$

【6-12】 设总体 $X \sim U[a, b]$, $X_1, X_2, \dots, X_n$ 为来自 X 的样本, 试求: 未知参数 a, b 的极大似然估计.

分析 已知总体 X 的密度函数

$$f(x; a, b) = \begin{cases} \frac{1}{b-a}, & x \in [a, b], \\ 0, & x \notin [a, b], \end{cases}$$

则对于样本 $X_1, X_2, \dots, X_n$ 的一次观察值 $x_1, x_2, \dots, x_n$ 的似然函数

$$L(a, b) = \begin{cases} \frac{1}{(b-a)^n} & (a \leq x_1, x_2, \dots, x_n \leq b), \\ 0, & (\text{其他}), \end{cases}$$

显然 L 不是 a, b 的可微函数, 因此不能通过解似然方程组求 a 和 b 的极大似然估计. 事实上, 由对数似然方程组

$$\frac{\partial \ln L}{\partial a} = \frac{n}{(b-a)^{n+1}} = 0,$$

$$\frac{\partial \ln L}{\partial b} = -\frac{n}{(b-a)^{n+1}} = 0,$$

解不出 a 和 b . 但并不意味着用极大似然估计的方法求解失败, 我们可通过极大似然估计的定义来直接求解.

解 我们直接求 L 的最大值点, 由 L 的表达式可以看出, a 与 b 越靠近, L 就越大, 但同时又必须满足 $a \leq x_1, x_2, \dots, x_n \leq b$, 所以取 $a = \min_{1 \leq i \leq n} \{x_i\}$, $b = \max_{1 \leq i \leq n} \{x_i\}$ 时, L 达到最大, 即 a, b 的极大似然估计值

$$\hat{a} = \min\{x_1, x_2, \dots, x_n\} = x_{(1)},$$

$$\hat{b} = \max\{x_1, x_2, \dots, x_n\} = x_{(n)};$$

a, b 的极大似然估计量

$$\hat{a} = \min\{X_1, X_2, \dots, X_n\} = X_{(1)},$$

$$\hat{b} = \max\{X_1, X_2, \dots, X_n\} = X_{(n)}.$$

【6-13】 设 $X_1, X_2, \dots, X_9$ 为来自总体 X 的样本, 总体 X 的期望

$E(X) = \mu$ 已知, 但方差未知. 求: 常数 c 使得 $S_9^2 = c \sum_{i=1}^9 (X_i - \mu)^2$ 为 σ^2 的无偏估计.

分析 由无偏性定义, 即 $E(S_9^2) = E(c \sum_{i=1}^9 (X_i - \mu)^2) = \sigma^2$ 来求出 c .

解 因为 $E(S_9^2) = c \sum_{i=1}^9 E(X_i - \mu)^2 = c \sum_{i=1}^9 E(X_i^2 - 2\mu X_i + \mu^2) = c \sum_{i=1}^9 [E(X_i^2) - \mu^2] = c9\sigma^2$, 故应有 $E(S_9^2) = c9\sigma^2 = \sigma^2$, 解得 $c = \frac{1}{9}$.

【6-14】 设总体 X 服从二项分布, 即 X 的分布律为

$$P(X = k) = C_n^k p^k (1-p)^{n-k}, k = 0, 1, \dots, n \text{ (其中 } p \text{ 未知)}.$$

又 $X_1, X_2, \dots, X_N$ 为来自总体 X 的样本, 求 p^2 的无偏估计量.

分析 判断估计量的无偏性归结为求估计量的数学期望.

方法 1 已知 $E(X) = np$, $D(X) = np(1-p) = np - np^2$, 又 $D(X) = E(X^2) - E^2(X)$, 故

$$\begin{aligned} E(X^2) &= D(X) + E^2(X) = np - np^2 + n^2 p^2 \\ &= E(X) + n(n-1)p^2, \end{aligned}$$

于是

$$E\left(\frac{X(X-1)}{n(n-1)}\right) = \frac{1}{n(n-1)} E(X(X-1)) = p^2.$$

若样本的值 $\frac{1}{N} \sum_{i=1}^N X_i(X_i - 1)$ 取代 $E(X(X-1))$, 并令 $\hat{p}^2 =$

$\frac{1}{Nn(n-1)} \sum_{i=1}^N X_i(X_i - 1)$, 则 $\hat{p}^2$ 是 p^2 的无偏估计量.

方法 2 由于 $E(X_i) = E(X) = np$, $D(X_i) = D(X) = np(1-p) = np - np^2$, 又 $D(X_i) = E(X_i^2) - E^2(X_i)$, 故

$$\begin{aligned} E(X_i^2) &= D(X_i) + E^2(X_i) = np - np^2 + n^2 p^2 \\ &= E(X_i) + n(n-1)p^2, \end{aligned}$$

于是

$$E\left(\frac{X_i(X_i-1)}{n(n-1)}\right) = p^2.$$

若 $\hat{p}^2 = \frac{X_i(X_i-1)}{n(n-1)} (i=1, 2, \dots, n)$, 则 $\hat{p}^2$ 是 p^2 的无偏估计量.

方法3 已知 $E(X) = np, D(X) = np - n^2 p = E(X) - n^2 p^2$, 得

$$\frac{1}{n}[E(X) - D(X)] = p^2.$$

令 $\hat{p}^2 = \frac{1}{n}(\bar{X} - S^2)$, 其中 $\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i, S^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2$, 则 $\hat{p}^2$ 是 p^2 的无偏估计量.

点评 由本题可见, 一个未知参数可以有不同的无偏估计量.

【6-15】 某钢缆厂质量管理部门定期抽出一批钢缆作抗拉强度试验. 经验表明, 某种型号钢缆的抗拉强度服从正态分布, 标准差为 100 kPa. 今从一批该种型号的钢缆中抽出 16 根组成一个标本, 经试验后求得平均抗拉强度为 3 100 kPa, 以 95% 的可靠性估计这批钢缆的平均抗拉强度的范围.

分析 对正态总体 $X \sim N(\mu, \sigma^2)$ 的期望和方差的区间估计, 一般步骤: 首先是根据题意确定是要求哪一个未知参数的置信区间, 置信度是多少? 其次是利用抽样分布定理, 构造含有未知参数即待估参数的有确定分布的随机变量; 然后根据随机变量的分布, 对给定的置信度 $1-\alpha$, 查表定出临界值; 最后将不等式变形, 求出参数即待估参数的置信区间.

对于本题中若设 X 为钢缆抗拉强度, 则 $X \sim N(\mu, \sigma^2)$, 其中 μ 未知, $\sigma^2 = 100^2$, 则 σ^2 为已知. $X_1, X_2, \dots, X_n$ 为来自总体 X 的一个样本, 此时求得未知参数 μ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right).$$

解 已知 $1-\alpha = 0.95$, 即 $\alpha = 0.05, n = 16, \bar{x} = 3\,100$, 并且查表求得 $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$. 因此解得 μ 的置信度为 0.95 的置信区间为

$$\left(\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$$

$$= \left(3100 - 1.96 \frac{100}{\sqrt{16}}, \quad 3100 + 1.96 \frac{100}{\sqrt{16}} \right) \\ = (3051, \quad 3149).$$

【6-16】 从某校男生中随机抽取 36 名, 测量并计算得平均身高为 170 cm, 标准差为 5 cm. 假定身高是一个正态变量, 求该校男生平均身高的 95% 的置信区间.

分析 设 X 为身高, 则 $X \sim N(\mu, \sigma^2)$, 其中 μ, σ^2 均未知. 由于 σ^2 未知, 因此未知参数 μ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\bar{X} - t_{\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{n}}, \quad \bar{X} + t_{\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{n}} \right).$$

解 已知 $\alpha=0.05, n=36, t_{\frac{\alpha}{2}}(n-1)=t_{0.025}(35)=2.0301, \bar{x}=170, s=5$. 由此求得该校男生平均身高 μ 的置信度为 0.95 的置信区间为

$$\left(\bar{x} - t_{\frac{\alpha}{2}}(n-1) \frac{s}{\sqrt{n}}, \quad \bar{x} + t_{\frac{\alpha}{2}}(n-1) \frac{s}{\sqrt{n}} \right) \\ = \left(170 - 2.0301 \frac{5}{\sqrt{36}}, \quad 170 + 2.0301 \frac{5}{\sqrt{36}} \right) \\ = (168.31, \quad 171.69).$$

【6-17】 调查 100 人中每个人的吸烟量, 得知平均值为 12 支, 标准差 5 支, 求平均吸烟量 95% 的置信区间.

解 设 X 为吸烟量, $X \sim N(\mu, \sigma^2)$, 其中 μ, σ^2 均为未知. 又 $X_1, X_2, \dots, X_n$ 为总体 X 的一个样本, 由于 σ^2 未知, 故 μ 的置信度为 $1-\alpha$ 的置信区间为 $\left(\bar{X} - t_{\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{n}}, \bar{X} + t_{\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{n}} \right)$.

已知 $\alpha=0.05, n=100, t_{\frac{\alpha}{2}}(n-1)=t_{0.025}(99) \approx z_{0.025}=1.96, \bar{x}=12, s=5$. 由此算得平均吸烟量 μ 的置信度为 0.95 的置信区间为

$$\left(\bar{x} - t_{\frac{\alpha}{2}}(n-1) \frac{s}{\sqrt{n}}, \quad \bar{x} + t_{\frac{\alpha}{2}}(n-1) \frac{s}{\sqrt{n}} \right) \\ = \left(12 - 1.96 \frac{5}{\sqrt{100}}, \quad 12 + 1.96 \frac{5}{\sqrt{100}} \right)$$

$$= (11.02, 12.98).$$

【6-18】 出于质量管理的需要,某厂技术员想知道用自动车床加工的一批零件长度的标准差.为此抽取 5 个零件,测量长度(单位:mm)为 22.3, 21.6, 22.0, 21.8, 21.4. 假设零件长度服从正态分布,分别求总体方差和标准差的 95% 的置信区间.

分析 若设零件的长度为 X , 则 $X \sim N(\mu, \sigma^2)$, 其中 μ 与 σ^2 均为未知. 由于 μ 未知, 有总体方差 σ^2 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}^2(n-1)}, \frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)} \right),$$

均方差 σ 的置信度为 $1-\alpha$ 的置信区间是

$$\left(\sqrt{\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}^2(n-1)}}, \sqrt{\frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)}} \right).$$

解 由样本的观察值算得 $\bar{x} = \frac{1}{5}(22.3 + 21.6 + 22.0 + 21.8 +$

$$21.4) = 21.82, s^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right) = 0.122. \text{ 又 } \alpha = 0.05, n = 5,$$

$\chi_{\frac{\alpha}{2}}^2(n-1) = \chi_{0.025}^2(4) = 11.143, \chi_{1-\frac{\alpha}{2}}^2(n-1) = \chi_{0.975}^2(4) = 0.484$, 求得方差 σ^2 的置信度为 0.95 的置信区间为

$$\begin{aligned} & \left(\frac{(n-1)s^2}{\chi_{\frac{\alpha}{2}}^2(n-1)}, \frac{(n-1)s^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)} \right) \\ &= \left(\frac{4 \times 0.122}{11.143}, \frac{4 \times 0.122}{0.484} \right) \\ &= (0.0438, 1.0083), \end{aligned}$$

以及标准差 σ 的置信度为 0.95 的置信区间

$$\begin{aligned} & \left(\sqrt{\frac{(n-1)s^2}{\chi_{\frac{\alpha}{2}}^2(n-1)}}, \sqrt{\frac{(n-1)s^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)}} \right) \\ &= (\sqrt{0.0438}, \sqrt{1.0083}) \\ &= (0.2093, 1.0041). \end{aligned}$$

【6-19】 化肥厂用自动包装机包装化肥,每包的重量 $X \sim N(100, 1.2^2)$ 单位为 kg. 某日开工后,为了确定这天包装机工作是否正常,随

机抽取 9 袋化肥,称得重量如下:

99.3, 98.7, 100.5, 101.2, 98.3, 99.7, 99.5, 102.1, 100.5

设方差稳定不变,问:这一天包装机包装化肥的平均重量是否为 100 kg (取 $\alpha=0.1$)?

分析 假设检验的基本思想为依据实际推断原理,即假设小概率事件在一次试验中几乎是不会发生的,来进行推断.在原假设 H_0 成立的条件下,样本值落入否定域的概率为 α ,这是小概率事件,不应发生.因此,若试验结果落入否定域,则拒绝 H_0 而接受备择假设 H_1 ,即认为 H_0 不成立而 H_1 成立;否则接受 H_0 .

假设检验的基本步骤:

- ① 根据题意提出原假设 H_0 和备择假设 H_1 ;
- ② 假设 H_0 成立,并选择合适的检验统计量以及确定其分布;
- ③ 给定显著性水平 α ,确定检验的拒绝域 W ;
- ④ 作判断:由样本的观察值计算统计量的值.若统计值落入拒绝域 W 中,则拒绝 H_0 ;否则接受 H_0 .

解 根据题意提出假设

$$H_0: \mu = 100; \quad H_1: \mu \neq 100.$$

这是一个双侧假设检验问题.由于方差 σ^2 已知用 U 检验法.

设 H_0 成立,则检验统计量

$$U = \frac{\bar{X} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} \sim N(0, 1).$$

已知 $\alpha=0.1, z_{\frac{\alpha}{2}} = z_{0.05} = 1.645$,故拒绝域

$$W = (-\infty, -1.645] \cup [1.645, +\infty).$$

又因 $n=9, \mu_0=100, \sigma_0=1.2, \bar{x} = \frac{1}{9} \sum_{i=1}^9 x_i = 99.98$,可算得检验统计值

$$U_0 = \frac{\bar{x} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} = \frac{99.98 - 100}{\frac{1.2}{\sqrt{9}}} = -0.05 \notin W.$$

由于 $U_0 \notin W$, 因此接受 H_0 , 即包装机包装化肥的平均重量是 100 kg.

【6-20】 已知某制砖厂生产的砖的抗断强度 X 服从正态分布, 今从该厂所产的一批砖中, 随机抽取 6 块, 测得抗断强度 (单位: MPa) 如下:

$$3.26, 2.97, 3.16, 3.00, 3.19, 3.10$$

试问: 这批砖的平均抗断强度是否为 3.25 (取 $\alpha=0.05$) ?

解 由题意提出假设

$$H_0: \mu = 3.25; \quad H_1: \mu \neq 3.25.$$

这是一个双侧假设检验问题. 由于方差 σ^2 未知, 用 T 检验法.

设 H_0 成立, 则检验统计量

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1).$$

已知 $\alpha=0.05, n=6, t_{\frac{\alpha}{2}}(n-1) = t_{0.025}(5) = 2.5706$, 故拒绝域

$$\begin{aligned} W &= (-\infty, -t_{\frac{\alpha}{2}}(n-1)] \cup [t_{\frac{\alpha}{2}}(n-1), +\infty) \\ &= (-\infty, -2.5706] \cup [2.5706, +\infty). \end{aligned}$$

又算得 $\bar{x} = 3.11, s^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right) = 0.1937^2, \mu_0 = 3.25$, 以及检验统计值

$$T_0 = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{3.11 - 3.25}{\frac{0.1937}{\sqrt{6}}} = -1.7704 \notin W.$$

由于 $T_0 \notin W$, 接受 H_0 , 即可认为这批砖的平均抗断强度为 3.25.

【6-21】 由经验知某零件重量 $X \sim N(15, 0.05^2)$, 技术革新后, 抽出 6 个零件, 测得重量 (单位: g) 如下:

$$14.7, 15.1, 14.8, 15.0, 15.2, 14.6$$

已知方差不变, 问: 平均重量是否不低于 15g (取 $\alpha=0.05$) ?

分析 按题意可知本题是一个单侧假设检验问题, 若提出假设 $H_0: \mu=15; H_1: \mu>15$, 由于算得 $\bar{x}=14.9 < \mu_0=15$, 可知这是错的. 因

为要检验均值 μ , 应该利用样本观察值的均值 $\bar{x}$ 的表现趋势来决定 H_1 . 正确的是: $H_0: \mu = 15; H_1: \mu < 15$.

解 由题意可知本题是一个单侧假设检验问题. 又算得 $\bar{x} = 14.9 < \mu_0 = 15$, 因此可先提出 H_1 然后再提出 H_0 . 故提出左侧假设

$$H_0: \mu = 15; \quad H_1: \mu < 15.$$

由于方差已知, 用 U 检验法. 设 H_0 成立, 则检验统计量

$$U = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1).$$

由 $\alpha = 0.05, z_\alpha = z_{0.05} = 1.645$, 得拒绝域

$$W = (-\infty, -z_\alpha] = (-\infty, -1.645].$$

又 $\bar{x} = 14.9, \mu_0 = 15, \sigma_0 = 0.05$, 算得检验统计值

$$U_0 = \frac{\bar{x} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} = \frac{14.9 - 15}{\frac{0.05}{\sqrt{6}}} = -4.899 \in W$$

由于 U_0 落在 W 内, 因此拒绝 H_0 , 接受 H_1 , 即平均重量低于 15 g.

【6-22】 某市 1 月份对零售商店的调查结果表明, 市郊食品店的月平均销售额为 25 000 元. 在下一月份中, 取出 16 个市郊食品店的一个样本, 其平均月销售额为 26 600 元, 销售额的标准差为 3 200 元. 试问: 能否得出结论说, 从上次调查以来, 平均月销售额提高了 (假定市郊食品店的月销售额服从正态分布, 取 $\alpha = 0.05$) ?

分析 设月销售额为 X 元, 则 $X \sim N(\mu, \sigma^2)$, 其中方差 σ^2 未知, 要检验均值 μ . 根据题意可知本题是一个单侧假设检验问题; 另算得样本观察值的均值 $\bar{x} = 26 600 > \mu_0 = 25 000$, 因此本题是右侧假设检验问题.

解 提出假设

$$H_0: \mu = \mu_0 = 25 000; \quad H_1: \mu > \mu_0 = 25 000.$$

由于方差 σ^2 未知, 用 T 检验法. 设 H_0 成立, 则检验统计量

$$T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} \sim t(n-1).$$

已知 $\alpha = 0.05, n = 16, t_\alpha(n-1) = t_{0.05}(15) = 1.7531$, 得拒绝域

$$W = [1.7531, +\infty).$$

又 $\bar{x}=26\,600, s=3\,200, \mu_0=25\,000$, 算得检验统计值

$$T_0 = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{26\,600 - 25\,000}{\frac{3\,200}{\sqrt{16}}} = 2 \in W.$$

由于 T_0 落在 W 内, 则拒绝 H_0 , 即平均月销售额提高了.

【6-23】 为考察一鱼塘中某种鱼的含汞量 (mg/kg), 随机地取 10 条鱼测得各条鱼的含汞量如下:

$$0.8, 1.6, 0.9, 0.8, 1.2, 0.4, 0.7, 1.0, 1.2, 1.1$$

设鱼的含汞量 X 服从正态分布 $N(\mu, \sigma^2)$, σ^2 未知, 试检验假设 $H_0: \mu=1.2$ (取 $\alpha=0.10$).

分析 对于本题中提出检验假设 $H_0: \mu=1.2$, 可用两种方法来求解: 一是用双侧假设检验法, 即 $H_0: \mu=1.2, H_1: \mu \neq 1.2$; 二是由算得样本观察值的均值 $\bar{x}=0.97 < \mu_0=1.2$. 因此可用左侧假设检验法来解.

解 方法 1 用双侧假设检验法

首先提假设

$$H_0: \mu = 1.2; \quad H_1: \mu \neq 1.2.$$

当 H_0 为真时, 由于 σ^2 未知, 用 T 检验法. 此时检验统计量

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1).$$

已知 $\alpha=0.10, n=10, t_{\frac{\alpha}{2}}(n-1)=t_{0.05}(9)=1.8331$, 得拒绝域

$$W = (-\infty, -1.8331] \cup [1.8331, +\infty).$$

另外有 $\bar{x}=0.97, s^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right) = 0.33^2, \mu_0 = 1.2$, 可算得

$$T_0 = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{0.97 - 1.2}{\frac{0.33}{\sqrt{10}}} = -2.2 \in W.$$

由于 $T_0 \in W$, 拒绝 H_0 , 即鱼的平均含汞量 $\mu \neq 1.2$.

以上可看出用双侧检验法得出鱼的平均含汞量 $\mu \neq 1.2$ 有差异, 但无法判断 μ 的变化趋势究竟是变大还是变小, 于是可利用以下的左侧

假设检验问题来求解.

方法 2 用左侧假设检验法

首先提出假设

$$H_0: \mu = 1.2; \quad H_1: \mu < 1.2.$$

当 H_0 为真时,有

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1).$$

已知 $\alpha=0.10, n=10, t_{\alpha}(n-1)=t_{0.10}(9)=1.3830$, 则拒绝域

$$W = (-\infty, -1.3830],$$

可算得

$$T_0 = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{0.97 - 1.2}{\frac{0.33}{\sqrt{10}}} = -2.2 \in W,$$

拒绝 H_0 , 即接受 H_1 , 可见鱼的平均含汞量比预期的 $\mu=1.2$ 偏小.

【6-24】 某车间生产铜丝, 生产一向较稳定, 可认为其折断力 X 服从正态分布. 今从产品中随机抽取 10 根检查折断力, 得数据如下(单位: N):

$$578, 572, 570, 568, 572, 570, 572, 570, 596, 584$$

是否可相信该车间的铜丝折断力的方差为 64(取 $\alpha=0.05$)?

解 本题是期望 μ 未知, 要检验方差 σ^2 的双侧假设检验问题. 提出假设

$$H_0: \sigma^2 = 64; \quad H_1: \sigma^2 \neq 64.$$

若 H_0 为真, 用 χ^2 检验法求解, 此时检验统计量

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1).$$

已知 $\alpha=0.05, n=10, \chi_{\frac{\alpha}{2}}^2(n-1)=\chi_{0.025}^2(9)=19.023, \chi_{1-\frac{\alpha}{2}}^2(n-1)=\chi_{0.975}^2(9)=2.700$, 则拒绝域

$$W = [0, 2.700] \cup [19.023, +\infty).$$

另有 $\bar{x}=575.2, s^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right) = 8.70^2 = 75.73, \sigma_0^2 =$

64, 可求得

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{9 \times 75.73}{64} = 10.64 \notin W.$$

由于 $\chi_0^2 \notin W$, 接受 H_0 , 即可认为该车间生产的铜丝折断力的方差为 64.

【6-25】 已知维尼纶纤维的纤度在正常条件下服从正态分布, 且标准差为 0.048. 从某天生产的产品中抽取 5 根纤维, 测得其纤度如下:

$$1.32, 1.55, 1.36, 1.40, 1.44$$

问: 这一天纤度的总体标准差是否正常 (取 $\alpha=0.05$) ?

分析 本题可用双侧或单侧假设检验法求解. 由于样本观察值算得的标准差 $s=0.0882 > \sigma_0=0.048$, 故用右侧假设检验法求解.

解 设维尼纶纤度为 X , 则 $X \sim N(\mu, \sigma^2)$, 其中 μ, σ^2 均未知.

方法 1 用双侧假设检验法

首先提出假设

$$H_0: \sigma = 0.048; \quad H_1: \sigma \neq 0.048.$$

设 H_0 成立, 由于 μ 未知, 用 χ^2 检验法, 有

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1).$$

已知 $\alpha=0.05, n=5, \chi_{\frac{\alpha}{2}}^2(n-1) = \chi_{0.025}^2(4) = 11.143, \chi_{1-\frac{\alpha}{2}}^2(n-1) = \chi_{0.975}^2(4) = 0.484$, 求得拒绝域

$$W = [0, 0.484] \cup [11.143, +\infty).$$

又因 $\bar{x} = 1.414, s^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right) = 0.0882^2, \sigma_0^2 = 0.048^2$, 可得

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{4 \times 0.0882^2}{0.048^2} = 13.51 \in W.$$

由于 $\chi_0^2 \in W$, 则拒绝 H_0 , 可认为纤度的总体标准差不正常.

方法 2 用右侧假设检验法

提出假设

$$H_0: \sigma = 0.048; \quad H_1: \sigma > 0.048.$$

设 H_0 成立时,有

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1).$$

已知 $\alpha=0.05, n=5, \chi_{\alpha}^2(n-1)=\chi_{0.05}^2(4)=9.488$, 则拒绝域

$$W = [9.488, +\infty).$$

另外算得 $\bar{x}=1.414, s^2 = \frac{1}{n-1}(\sum_{i=1}^n x_i^2 - n\bar{x}^2) = 0.0882^2, \sigma_0^2 = 0.048^2$, 以及

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{4 \times 0.0882^2}{0.048^2} = 13.51 \in W.$$

由于 $\chi_0^2 \in W$, 拒绝 H_0 , 接受 H_1 , 则认为该天生产的维尼纶纤维的纤度的标准差不正常而且偏大.

【6-26】 为研究某种合金的铜含量, 取 6 个试块测得铜含量(%)如下:

$$8.03, 9.99, 9.92, 7.75, 11.65, 14.64$$

设铜含量服从正态分布, 试问:

(1) 在显著性水平 $\alpha=0.05$ 下能否认为铜含量的均值显著大于 9.5%?

(2) 在显著性水平 $\alpha=0.01$ 下能否认为铜含量的方差不超过 5.5%?

分析 本题是检验均值 μ 和方差 σ^2 的单侧假设检验问题. 由于样本观察值的均值 $\bar{x}=10.33 > \mu_0=9.5$ 和方差 $s^2=6.51 > \sigma_0^2=5.5$, 对 μ 和 σ^2 来说均为右侧假设检验问题.

解 设 X 为某种合金的铜含量, 即 $X \sim N(\mu, \sigma^2)$, 其中 μ, σ^2 均未知.

(1) 提出假设

$$H_0: \mu = 9.5; \quad H_1: \mu > 9.5.$$

设 H_0 成立, 由于 σ^2 未知, 可用 T 检验法, 算得检验统计量

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1).$$

已知 $\alpha=0.05, n=6, t_{\alpha}(n-1)=t_{0.05}(5)=2.0150$, 则拒绝域

$$W = [2.0150, +\infty).$$

另算得 $\bar{x}=10.33, s^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right) = 2.55^2, \mu_0 = 9.5$, 以及

$$T_0 = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{10.33 - 9.5}{\frac{2.55}{\sqrt{6}}} = 0.797 \notin W.$$

由于 $T_0 \notin W$, 接受 H_0 , 可认为铜含量的均值不显著大于 9.5% .

(2) 提出假设

$$H_0: \sigma^2 = 5.5; \quad H_1: \sigma^2 > 5.5.$$

设 H_0 为真, 用 χ^2 检验法, 算得检验统计量

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1).$$

已知 $\alpha=0.01, n=6, \chi_{\alpha}^2(n-1)=\chi_{0.01}^2(5)=15.086$, 求得拒绝域

$$W = [15.086, +\infty).$$

又因 $\bar{x}=10.33, s^2=6.51=2.55^2, \sigma_0^2=5.5$, 可算得检验统计值

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{5 \times 6.51}{5.5} = 5.92 \notin W.$$

由于 $\chi_0^2 \notin W$, 接受 H_0 , 可认为铜含量的方差不超过 5.5% .

B 习题精选

【6-27】 设 X 服从正态分布, $E(X) = -1, E(X^2) = 4, \bar{X} =$

$\frac{1}{n} \sum_{i=1}^n X_i$, 则 $\bar{X}$ 服从的分布是_____.

(A) $N\left(-1, \frac{3}{n}\right)$;

(B) $N(-1, 1)$;

(C) $N\left(-\frac{1}{n}, 4\right)$;

(D) $N\left(-\frac{1}{n}, \frac{1}{n}\right)$.

【6-28】 设总体 X 服从正态分布 $N(\mu, \sigma^2)$, 其中 μ, σ^2 是未知参数. $X_1, X_2, \dots, X_n$ 是来自 X 的样本, 则下列结论正确的是_____.

$$(A) S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi^2(n-1);$$

$$(B) S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi^2(n-1);$$

$$(C) \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi^2(n-1);$$

$$(D) \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi^2(n).$$

【6-29】 设总体 $X \sim B(1, p)$, 其中 p 为未知参数. 设 $X_1, X_2, \dots, X_n$ 是来自总体 X 的一个样本, $\bar{X}$ 是样本的均值, 则 $P\left(\bar{X} = \frac{k}{n}\right) = \underline{\hspace{2cm}}$.

$$(A) p;$$

$$(B) 1-p;$$

$$(C) C_n^k p^k (1-p)^{n-k};$$

$$(D) C_n^k (1-p)^k p^{n-k}.$$

【6-30】 设 $X_1, X_2, \dots, X_{10}$ 为 $N(0, 0.3^2)$ 的一个样本, 则 $P\left(\sum_{i=1}^{10} X_i^2 > 0.4379\right) \underline{\hspace{2cm}}$.

$$(A) 0.9;$$

$$(B) 0.1;$$

$$(C) 0.2;$$

$$(D) 0.3.$$

【6-31】 设 $X_1, X_2, \dots, X_n$ 是来自总体 $X \sim N(\mu, \sigma^2)$ 的一个样本, $\bar{X}$ 是样本的均值, 记

$$S_1^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2, \quad S_2^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2,$$

$$S_3^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \mu)^2, \quad S_4^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2,$$

则服从自由度为 $n-1$ 的 t 分布的随机变量 $T = \underline{\hspace{2cm}}$.

$$(A) \frac{\bar{X} - \mu}{\frac{S_1}{\sqrt{n-1}}};$$

$$(B) \frac{\bar{X} - \mu}{\frac{S_2}{\sqrt{n-1}}};$$

$$(C) \frac{\bar{X} - \mu}{\frac{S_3}{\sqrt{n}}};$$

$$(D) \frac{\bar{X} - \mu}{\frac{S_4}{\sqrt{n}}}.$$

【6-32】 设 $X_1, X_2, \dots, X_n$ 是来自总体 X 的一个样本, 总体 $X \sim U[a, b]$, 则 $E(\bar{X}) = \underline{\hspace{2cm}}$, $D(\bar{X}) = \underline{\hspace{2cm}}$, $E(S^2) = \underline{\hspace{2cm}}$.

【6-33】 设 $X_1, X_2, \dots, X_n$ 为来自总体 $X \sim \chi^2(10)$ 的样本, 则统计量 $Y = \sum_{i=1}^n X_i \sim \underline{\hspace{2cm}}$.

【6-34】 若一个样本的观察值为 $0, 0, 1, 1, 0, 1$, 则总体均值的矩估计值为 $\underline{\hspace{2cm}}$, 总体方差的矩估计值为 $\underline{\hspace{2cm}}$.

【6-35】 设 $1, 0, 0, 1, 1$ 是来自两点分布总体 $B(1, p)$ 的样本观察值, 则参数 $q = 1 - p$ 的矩估计值为 $\underline{\hspace{2cm}}$.

【6-36】 若由总体 $F(x; \theta)$ (θ 为未知参数) 的样本观察值所求得 $P(35.5 < \theta < 35.9) = 0.95$, 则称 $\underline{\hspace{2cm}}$ 是 θ 的置信度为 $\underline{\hspace{2cm}}$ 的置信区间.

【6-37】 设总体 X 服从正态分布 $N(\mu, \sigma^2)$, 其中 μ, σ^2 皆未知, 则 μ 的置信度为 95% 的置信区间是 $\underline{\hspace{2cm}}$.

【6-38】 假设检验所依据的原理是 $\underline{\hspace{2cm}}$.

【6-39】 用某仪器间接测量温度, 重复 5 次, 测得结果为 1250°C , 1265°C , 1245°C , 1260°C , 1275°C . 设测量值 X 服从正态分布 $N(\mu, \sigma^2)$, 检验水平 $\alpha = 0.05$, 那么应取原假设 $H_0: \underline{\hspace{2cm}}$, 备择假设 $H_1: \underline{\hspace{2cm}}$, 又检验统计量为 $\underline{\hspace{2cm}}$.

【6-40】 设总体 $X \sim N(\mu, \sigma^2)$, μ 未知. 欲在显著水平 α 下检验假设 $H_0: \sigma^2 = \sigma_0^2$, $H_1: \sigma^2 < \sigma_0^2$, 应由样本的观察值 $x_1, x_2, \dots, x_n$ 计算统计量 $\underline{\hspace{2cm}}$ 的值. 如果 $\underline{\hspace{2cm}}$ 时, 则拒绝 H_0 ; 如果 $\underline{\hspace{2cm}}$, 则接受 H_0 .

【6-41】 在总体 $N(52, 6.3^2)$ 中随机抽一容量为 36 样本, 求样本均值 $\bar{X}$ 落在 50.8 到 53.8 之间的概率.

【6-42】 设总体 $X \sim N(\mu, \sigma^2)$. 已知样本容量 $n = 24$, 样本方差 $s^2 = 12.5227$, 求总体标准差大于 3 的概率.

【6-43】 设总体 $X \sim \chi^2(n)$, $X_1, X_2, \dots, X_{10}$ 为来自总体 X 的一个样本, 求 $E(\bar{X}), D(\bar{X}), E(S^2)$.

【6-44】 设总体 $X \sim N(\mu, \sigma^2)$, 其中 μ 和 σ^2 均为未知. 已知 $n = 16$, 样本均值 $\bar{x} = 12.5$, 样本方差 $s^2 = 5.333$, 求: $P(|\bar{X} - \mu| < 0.4)$.

【6-45】 设 X_1, X_2, X_3, X_4 是来自正态总体 $N(0, 1)$ 的样本, 又设 $Y = (X_1 + X_2)^2 + (X_3 + X_4)^2$, 求常数 c 使得 cY 服从 χ^2 分布.

【6-46】 设 $X_1, X_2, \dots, X_n$ 为来自总体 X 的一个样本, X 服从均匀分布 $U[1, \theta]$, 其中 θ 未知.

(1) 求 θ 的矩估计量;

(2) 当样本的观察值为 1.3, 1.8, 1.27, 1.35, 1.62, 1.55 时, 求 θ 的矩估计值.

【6-47】 设某市的新生儿体重 X (单位: g) 服从正态分布 $N(\mu, \sigma^2)$, 现测量 10 名新生儿的体重如下:

3 100, 3 480, 2 520, 3 700, 2 520,

3 200, 2 800, 3 800, 3 020, 3 260.

(1) 求参数 μ, σ^2 的矩估计;

(2) 求参数 σ^2 的一个无偏估计.

【6-48】 某铁路局证实, 一个扳道员在 5 年内所引起的严重事故次数服从参数为 λ 的泊松分布. 设 r 表示一扳道员在 5 年内引起的严重事故次数, t 表示观察到的扳道员人数, 有

r	0	1	2	3	4	5
t	44	42	21	9	4	2

求未知参数 λ 的极大似然估计值以及求出一个扳道员在 5 年内未引起严重事故的概率.

【6-49】 设总体 X 的概率密度

$$f(x; \alpha) = \begin{cases} (\alpha + 1)x^\alpha, & x \in (0, 1), \\ 0, & x \notin (0, 1), \end{cases} \quad \text{其中 } \alpha \text{ 未知且 } \alpha > -1.$$

求: (1) α 的矩估计量; (2) α 的极大似然估计量.

【6-50】 设 $X_1, X_2, \dots, X_n$ 是来自参数为 λ 的泊松分布的一个样本, 求 λ^2 的无偏估计.

【6-51】 设 $X_1, X_2, \dots, X_n$ 是取自正态总体 $N(\mu, \sigma^2)$ 的样本, 试适当选择 C , 使 $S^2 = C \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2$ 为 σ^2 的无偏估计.

【6-52】 设某种清漆的 9 个样品,其干燥时间(单位:h)如下:

6.0, 5.7, 5.8, 6.5, 7.0, 6.3, 5.6, 6.1, 5.0

若干燥时间 X 服从正态分布 $N(\mu, \sigma^2)$, 求 μ 的置信度为 0.95 的置信区间: (1) 设由以往经验知 $\sigma=0.6$ h, (2) 设 σ 为未知.

【6-53】 随机地取某种炮弹 9 发做实验, 得炮口速度的样本标准差 $s=11$ (m/s), 设炮口速度服从正态分布, 求这种炮弹的炮口速度的标准差 σ 的置信度为 0.95 的置信区间.

【6-54】 设某地旅游者日消费额服从正态 $N(\mu, \sigma^2)$ 分布, 且标准差 $\sigma=12$ 元. 今对该地旅游者的日平均消费额进行估计, 为了能以 95% 的置信度相信这种估计误差的绝对值会小于 2 元, 问至少要调查多少人?

【6-55】 为确定某种溶液中的甲醛浓度, 取得 4 个独立测量值的样本, 并算得样本均值 $\bar{x}=8.34\%$, 样本标准差为 $s=0.03\%$. 设被测总体近似地服从正态分布, $\alpha=0.05$, 试分别求出 μ, σ^2 的置信区间.

【6-56】 设某次考试的考生的成绩服从正态分布, 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. 在显著性水平 $\alpha=0.05$ 下, 是否可以认为这次考试全体考生的平均成绩为 70 分?

【6-57】 要求一种元件的使用寿命不得低于 1000 h. 今从一批这种元件中随机抽取 25 件, 测得其寿命的平均值为 950 h. 已知该种元件寿命服从标准差 $\sigma=100$ h 的正态分布, 试在显著性水平 $\alpha=0.05$ 下确定这批元件是否合格(设总体均值为 μ)?

【6-58】 原有一台仪器, 测量电阻时误差相应的方差是 0.0006. 现有一台新仪器, 对一个电阻器测量了 10 次, 测得的值(单位: Ω) 如下:

1.101, 1.103, 1.105, 1.098, 1.099

1.101, 1.104, 1.095, 1.100, 1.100

取 $\alpha=0.10$, 问: 假定测量所得电阻值服从正态分布, 新仪器的精度是否比原有的仪器好?

【6-59】 用自动包装机包装食盐, 设每袋盐的净重服从正态分

布. 规定每袋盐的标准重量为 500 g, 标准差不能超过 10 g. 某天开工以后, 为了检查机器是否工作正常, 从已包装好的食盐中随机地抽取 9 袋, 测得其净重(g)如下:

497, 507, 510, 475, 484, 488, 524, 491, 515

问: 这天自动包装机工作是否正常($\alpha=0.05$)?

答案与提示

【6-27】 选 A. 由 $D(X) = E(X^2) - E^2(X) = 4 - 1 = 3$, 又 $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$, 故

$$\bar{X} \sim N(-1, \frac{3}{n}).$$

【6-28】 选 C.

【6-29】 选 C. X 服从(0-1)分布, 即 $\begin{array}{c|cc} X & 0 & 1 \\ \hline P & 1-p & p \end{array}$, 则 $\sum_{i=1}^n X_i \sim B(n, p)$. 故 $P(\sum_{i=1}^n X_i = k) = C_n^k p^k (1-p)^{n-k}$, 得 $P(\bar{X} = \frac{k}{n}) = P(\sum_{i=1}^n X_i = k) = C_n^k p^k (1-p)^{n-k}$. 因此选 C.

【6-30】 选 A. 由 $X_i \sim N(0, 0.3^2)$, 故 $\frac{X_i}{0.3} \sim N(0, 1), i=1, 2, \dots, 10$.

$$\chi^2 = \sum_{i=1}^{10} \left(\frac{X_i}{0.3} \right)^2 = \sum_{i=1}^{10} \frac{X_i^2}{0.3^2} \sim \chi^2(10),$$

从而

$$\begin{aligned} P\left(\sum_{i=1}^{10} X_i^2 > 0.4379\right) &= P\left(\sum_{i=1}^{10} \frac{X_i^2}{0.3^2} > 4.865\right) \\ &= P(\chi^2 > 4.865) = 0.9. \end{aligned}$$

【6-31】 选 B. 由 T 分布的定义可知, 当 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

时,有 $T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$. 因为 $T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n-1}}} = \frac{\bar{X} - \mu}{\sqrt{\frac{n-1}{n}} \frac{S}{\sqrt{n-1}}} =$

$\frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$, 故选 B.

【6-32】 $E(\bar{X}) = \frac{1}{2}(a+b)$, $D(\bar{X}) = \frac{1}{12n}(b-a)^2$, $E(S^2) = \frac{1}{12}(b-a)^2$.

【6-33】 $Y = \sum_{i=1}^n X_i \sim \chi^2(10n)$.

【6-34】 $\frac{1}{2}, \frac{1}{4}$.

【6-35】 $\hat{q} = \frac{2}{5}$.

【6-36】 (35.5, 35.9), 0.95.

【6-37】 $\left(\bar{X} - t_{0.025}(n-1) \frac{S}{\sqrt{n}}, \bar{X} + t_{0.025}(n-1) \frac{S}{\sqrt{n}} \right)$.

【6-38】 实际推断原理即小概率事件在一次试验中几乎是不可发生的.

【6-39】 提出假设 $H_0: \mu = 1277$ ($H_0: \mu \geq 1277$), $H_1: \mu < 1277$, 检验统计量 $T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$.

【6-40】 $\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2}$, $\chi_0^2 \in W = [0, \chi_{1-\alpha}^2(n-1)]$, $\chi_0^2 \notin W = [0, \chi_{1-\alpha}^2(n-1)]$.

【6-41】 因为 $\bar{X} \sim N\left(52, \frac{6.3^2}{36}\right)$, 所以

$$P(50.8 < \bar{X} < 53.8) = P\left(\frac{50.8 - 52}{\frac{6.3}{6}} < \frac{\bar{X} - 52}{\frac{6.3}{6}} < \frac{53.8 - 52}{\frac{6.3}{6}}\right)$$

$$\begin{aligned}
&= P\left[-\frac{8}{7} < \frac{\bar{X}-52}{\frac{6.3}{6}} < \frac{12}{7}\right] \\
&= \Phi\left(\frac{12}{7}\right) - \Phi\left(-\frac{8}{7}\right) = \Phi\left(\frac{12}{7}\right) + \Phi\left(\frac{8}{7}\right) - 1 \\
&\approx 0.9564 + 0.8729 - 1 = 0.8293.
\end{aligned}$$

【6-42】 由 $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$, 已知 $n=24$, 故 $\chi^2 = \frac{23 \times s^2}{\sigma^2} \sim \chi^2(23)$,

$$\begin{aligned}
P(\sigma > 3) &= P\left(\frac{1}{\sigma^2} < \frac{1}{9}\right) \\
&= P\left(\frac{23 \times s^2}{\sigma^2} < \frac{23 \times 12.5227}{9}\right) \\
&= P(\chi^2 < 32) = 1 - P(\chi^2 \geq 32).
\end{aligned}$$

查表得

$$P(\sigma > 3) = 1 - 0.1 = 0.9.$$

【6-43】 设 $\bar{X} = \frac{1}{10} \sum_{i=1}^{10} X_i$, 且由 χ^2 分布的性质知 $E(X) = n$,

$$D(X) = 2n. \text{ 有 } E(\bar{X}) = E(X) = 10, D(\bar{X}) = \frac{D(X)}{10} = \frac{2 \times 10}{10} = 2,$$

$$E(S^2) = D(X) = 2 \times 10 = 20.$$

【6-44】 由于 σ 未知, 要利用 $T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$. 已知 $n=16$,

$$\bar{x}=12.5, s^2=5.333=2.309^2, \text{ 有 } T = \frac{\bar{X} - \mu}{0.5773} \sim t(15). \text{ 算得}$$

$$\begin{aligned}
P(|\bar{X} - \mu| < 0.4) &= P\left(\left|\frac{\bar{X} - \mu}{0.5773}\right| < 0.692\right) \\
&= P(|T| < 0.692) \\
&= 1 - 2P(T \geq 0.692) \\
&= 1 - 2 \times 0.25 = 0.5.
\end{aligned}$$

【6-45】 因 $X_1 + X_2 \sim N(0, 2)$, $X_3 + X_4 \sim N(0, 2)$, 故

$$\left(\frac{X_1 + X_2}{\sqrt{2}}\right)^2 + \left(\frac{X_3 + X_4}{\sqrt{2}}\right)^2 \sim \chi^2(2),$$

即应取 $c = \frac{1}{2}$.

【6-46】 (1) 总体 X 的概率密度

$$f(x; \theta) = \begin{cases} \frac{1}{\theta-1}, & x \in [1, \theta], \\ 0, & x \notin [1, \theta], \end{cases}$$

则有 $E(X) = \frac{1+\theta}{2}$, 用 $\bar{X}$ 取代 $E(X)$ 和 $\hat{\theta}$ 取代 θ , 并解出 $\hat{\theta}$, 得 $\hat{\theta} = 2\bar{X} - 1$.

(2) 算得样本的观察值 $\bar{x} = \frac{1}{6} \sum_{i=1}^6 x_i = \frac{1}{6} (1.3 + 1.8 + 1.27 + 1.35 + 1.62 + 1.55) = 1.4817$, 所以有 $\hat{\theta} = 2\bar{x} - 1 = 1.9633$.

【6-47】 (1) μ, σ^2 的矩估计分别是样本的一阶原点矩和二阶中心矩, 即 $\hat{\mu} = \bar{x} = \frac{1}{10} \sum_{i=1}^{10} x_i = 3140, \hat{\sigma}^2 = \frac{1}{10} \sum_{i=1}^{10} (x_i - \bar{x})^2 = 178320$.

(2) 由于样本方差 S^2 是总体方差 σ^2 的无偏估计, 因此 σ^2 的无偏估计, $\hat{\sigma}^2 = s^2 = \frac{1}{9} \sum_{i=1}^{10} (x_i - \bar{x})^2 = 198133$

【6-48】 已知 $X \sim P(X=k; \lambda) = \frac{\lambda^k}{k!} e^{-\lambda}$, 其中 $k=0, 1, 2, \dots$ 且 $\lambda > 0$, 设 $x_1, x_2, \dots, x_n$ 为来自 X 的一个样本观察值, 则 $X_1, X_2, \dots, X_n$ 的联合分布律

$$\begin{aligned} P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n; \lambda) &= \frac{\lambda^{x_1}}{x_1!} e^{-\lambda} \frac{\lambda^{x_2}}{x_2!} e^{-\lambda} \dots \frac{\lambda^{x_n}}{x_n!} e^{-\lambda} \\ &= e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n (x_i!)} \end{aligned}$$

对数似然函数 $\ln L(\lambda) = -n\lambda + \sum_{i=1}^n x_i \ln \lambda - \ln \left(\prod_{i=1}^n x_i! \right)$,

对数似然方程为

$$\frac{d \ln L}{d \lambda} = -n + \frac{\sum_{i=1}^n x_i}{\lambda} = 0,$$

解得

$$\lambda = \frac{1}{n} \sum_{i=1}^n x_i.$$

因仅有一个解,所以参数 λ 的极大似然估计量

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}.$$

由提供的数据表可求得 λ 的极大似然估计值

$$\begin{aligned}\hat{\lambda} = \bar{x} &= \frac{1}{122} (0 \times 44 + 1 \times 42 + 2 \times 21 + 3 \times 9 + 4 \times 4 + 5 \times 2) \\ &= \frac{137}{122} = 1.1230.\end{aligned}$$

又由题意可知,即求一个扳道员在 5 年内未引起严重事故的概率

$$p = P(X = 0) = e^{-1.1230} = 0.3253.$$

【6-49】 (1) α 的矩估计量 $\hat{\alpha} = \frac{2\bar{X}-1}{2-\bar{X}}.$

(2) α 的极大似然估计量

$$\hat{\alpha} = -\frac{n}{\sum_{i=1}^n \ln X_i} - 1.$$

【6-50】 因 $E(X_i) = E(X) = \lambda, D(X_i) = D(X) = \lambda$, 故 $E(X_i^2) = D(X_i) + E^2(X_i) = E(X_i) + \lambda^2$. 于是 $E(X_i^2 - X_i) = \lambda^2$, 从而取 $\hat{\lambda}^2 = X_i^2 - X_i (i=1, 2, \dots, n)$, 均从 λ^2 的无偏估计. 另外取 $\hat{\lambda}^2 = \frac{1}{n} \sum_{i=1}^n X_i(X_i - 1)$ 也是 λ^2 的无偏估计.

【6-51】 因为 $E(S^2) = E\left(C \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2\right)$

$$\begin{aligned}&= CE\left(\sum_{i=1}^{n-1} (X_{i+1}^2 - 2X_i X_{i+1} + X_i^2)\right) \\ &= C \sum_{i=1}^{n-1} [E(X_{i+1}^2) - 2E(X_i X_{i+1}) + E(X_i^2)].\end{aligned}$$

又 $E(X^2) = D(X) + [E(X)]^2, D(X) = \sigma^2, E(X) = \mu$ 以及 X_{i+1} 与 X_i 相互独立, 故 $E(X_i X_{i+1}) = E(X_i)E(X_{i+1})$, 可得

$$\begin{aligned} E(S^2) &= C \sum_{i=1}^{n-1} [(\sigma^2 + \mu^2) - 2\mu^2 + (\sigma^2 + \mu^2)] \\ &= C \sum_{i=1}^{n-1} 2\sigma^2 = C \cdot 2(n-1)\sigma^2. \end{aligned}$$

若 $E(S^2) = \sigma^2$, 则 S^2 是 σ^2 的无偏估计, 应该选 $c = \frac{1}{2(n-1)}$.

【6-52】 (1) 当 σ^2 已知时, μ 的置信度为 0.95 的置信区间是

$$\left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}, \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}} \right).$$

取 $\alpha = 0.05$, $z_{\frac{\alpha}{2}} = 1.96$, $n = 9$, $\sigma = 0.6$, $\bar{x} = 6$, 算得置信区间为 (5.608, 6.392).

(2) 当 σ^2 未知时, μ 的置信度为 0.95 的置信区间是

$$\left(\bar{X} - \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1), \bar{X} + \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1) \right).$$

取 $\alpha = 0.05$, $n-1 = 8$, $t_{\frac{\alpha}{2}}(n-1) = 2.3060$, $\bar{x} = 6$, $s^2 = 0.33$, 代入上式算得置信区间为 (5.558, 6.442).

【6-53】 由于 μ 未知, σ 的置信度为 0.95 的置信区间为

$$\left[\frac{\sqrt{n-1}S}{\sqrt{\chi_{\frac{\alpha}{2}}^2(n-1)}}, \frac{\sqrt{n-1}S}{\sqrt{\chi_{1-\frac{\alpha}{2}}^2(n-1)}} \right].$$

取 $s = 11$, $n-1 = 8$, $\alpha = 0.05$, $\chi_{\frac{\alpha}{2}}^2(8) = 17.535$, $\chi_{1-\frac{\alpha}{2}}^2(8) = 2.180$, 代入上式算得置信区间为 (7.43, 21.07).

【6-54】 已知 $\alpha = 0.05$, $\sigma = 12$, $|\bar{X} - \mu| < 2$, $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$.

由于

$$P\left(|\bar{X} - \mu| < z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha,$$

所以 $z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 1.96 \times \frac{12}{\sqrt{n}} = 2$, 解得 $n = \left(\frac{1.96 \times 12}{2}\right)^2 \approx 138.3$, 即至少需要调查 139 人.

【6-55】 μ 的置信区间为 (8.292%, 8.388%), σ^2 的置信区间为 $(0.00029 \times 10^{-4}, 0.0125 \times 10^{-4})$.

【6-56】 (1) 设 $H_0: \mu = 70, H_1: \mu \neq 70$, 求得 $T_0 = 1.4 \notin W = (-\infty, -2.0301] \cup [2.0301, +\infty)$, 则接受 H_0 , 即可认为考生的平均分为 70 分.

(2) 设 $H_0: \mu = 70, H_1: \mu < 70$, 算得 $T_0 = -1.4 \notin W = (-\infty, -1.6896)$, 则接受 H_0 .

【6-57】 设 $H_0: \mu = 1000, H_1: \mu < 1000$, 算得 $z_0 = -2.5 \in W = (-\infty, -1.65] \cup [1.65, +\infty)$, 故拒绝 H_0 , 可认为元件的使用寿命偏低.

【6-58】 设 $H_0: \sigma^2 = 0.0006, H_1: \sigma^2 < 0.0006$, 算得

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{9 \times 0.0000784}{0.0006} = 1.176 \in W = [0, 4.168],$$

故拒绝 H_0 , 可认为新仪器的精度比原有的仪器好.

【6-59】 (1) 先检验假设 $H_0: \mu = 500, H_1: \mu \neq 500$, 算得 $T_0 = -0.187 \notin W = (-\infty, -2.306] \cup [2.306, +\infty)$, 故接受 H_0 , 可认为机器包装食盐的均值是 500 g, 没有产生系统误差.

(2) 再检验假设 $H_0: \sigma^2 = 100, H_1: \sigma^2 > 100$, 算得 $\chi_0^2 = 20.56 \in W = [15.5, +\infty)$, 故拒绝 H_0 , 接受 H_1 , 即可认为其标准差超过了 10 g.

由上可知, 这天机器自动包装食盐, 虽然没有产生系统误差, 但生产不够稳定(方差偏大), 从而认为这天自动包装机工作不正常.

附录 实战试卷及答案与提示

试 卷 (一)

(时间 90 分钟)

一、选择题(每题 3 分,共 18 分)

1. 口袋中有 5 只红球,3 只白球,不放回任取 3 只,其中至少有 2 只是红球的概率为_____.

- (A) $\frac{3}{8}$; (B) $\frac{15}{28}$; (C) $\frac{5}{7}$; (D) $\frac{5}{28}$.

2. 设 X 的分布函数 $F(x) = \begin{cases} 0 & (x < 0), \\ x^3 & (0 \leq x \leq 1), \\ 1 & (x > 1), \end{cases}$ 则 $E(X) = \underline{\hspace{2cm}}$.

- (A) $\int_0^{+\infty} x^4 dx$; (B) $\int_0^1 3x^3 dx$;
(C) $\int_0^1 3x^2 dx$; (D) $\int_0^1 x^3 dx + \int_1^{+\infty} x dx$.

3. 设 (X, Y) 的联合分布函数为 $F(x, y)$, 则其边缘分布函数 $F_X(x)$ 为_____.

- (A) $F(0, y)$; (B) $F(x, 0)$;
(C) $\lim_{x \rightarrow +\infty} F(x, y)$; (D) $\lim_{y \rightarrow +\infty} F(x, y)$.

4. 已知 $X \sim B(n, p)$, 则 $\frac{D(X)}{E(X)} = \underline{\hspace{2cm}}$.

- (A) n ; (B) $1-p$; (C) p ; (D) $\frac{1}{1-p}$.

5. 设 X, Y 为两个任意的随机变量, 已知 $\text{cov}(X, Y) = 0$, 则必

有_____.

- (A) X, Y 相互独立; (B) X, Y 不相互独立;
(C) $E(XY) = E(X)E(Y)$; (D) $D(X+Y) = D(X)D(Y)$.

6. 设 $X \sim N(\mu, \sigma^2)$, 且 σ^2 未知, 则 μ 的置信度为 0.95 的置信区间为_____.

- (A) $\left(\bar{X} \pm \frac{S}{\sqrt{n}} t_{0.025}\right)$; (B) $\left(\bar{X} \pm \frac{\sigma}{\sqrt{n}} t_{0.025}\right)$;
(C) $\left(\bar{X} \pm \frac{S}{\sqrt{n}} z_{0.025}\right)$; (D) $\left(\bar{X} \pm \frac{\sigma}{\sqrt{n}} z_{0.025}\right)$.

二、填空题(每题 3 分,共 24 分)

1. 已知 $P(B) = 0.5, P(A-B) = 0.3$, 则 $P(\overline{A}\overline{B}) =$ _____.

2. 已知随机事件 A 的概率 $P(A) = 0.5$, 随机事件 B 的概率 $P(B) = 0.6$, 条件概率 $P(B|A) = 0.8$, 则事件 $A \cup B$ 的概率 $P(A \cup B) =$ _____.

3. 已知随机变量 X 的密度函数 $f(x) = \begin{cases} ax^2 & (0 < x < 2) \\ 0 & (\text{其他}) \end{cases}$, 则 $a =$ _____.

4. 已知随机变量 $X \sim \frac{X}{P} \begin{array}{c|cccc} & -1 & 0 & 1 & 2 \\ \hline & 0.1 & a & 0.3 & b \end{array}$, 且 $P(X \leq 1.2) = 0.7$, 则 $a =$ _____, $b =$ _____.

5. 已知随机变量 X, Y 有 $D(X) = 4, D(Y) = 1, \rho_{XY} = 0.5$, 则 $D(2X - 3Y) =$ _____.

6. 已知 (X, Y) 的分布律为

$X \backslash Y$	1	2	3	4
1.5	0.1	0	0.05	0.15
2.5	0.05	0.15	0.05	0
3.5	0.1	0.2	0.05	0.1

则 $P(Y < 3) =$ _____.

7. 已知 $X \sim P\{X=k\} = \frac{2^k e^{-2}}{k!} \quad (k=0, 1, 2, \dots)$, $Y = 3X - 2$,

则 $E(Y) =$ _____.

8. 已知 $X_i (i=1, 2, 3, 4, 5)$ 是取自 $X \sim N(1, 4)$ 的样本, $Y_j (j=1, 2, 3, 4)$ 是取自 $Y \sim N(0, 2)$ 的样本, 且 X, Y 相互独立, 则统计量 $\bar{X} - \bar{Y} \sim$ _____.

三、计算题(每题 10 分, 共 50 分)

1. 有三只箱子, 甲箱有球 4 黑 2 白, 乙箱有球 3 黑 3 白, 丙箱有球 3 黑 6 白. 现随机取一只箱子, 再从箱子中任取一球, 试求:

(1) 取到的是白球的概率;

(2) 取到的是甲箱的白球的概率.

2. 设随机变量 (X, Y) 的联合分布律为

$X \backslash Y$	-1	1	2
-1	5/20	2/20	6/20
2	3/20	3/20	1/20

(1) 问: X 和 Y 是否相互独立? 为什么?

(2) 求: $X+Y, X-Y$ 的分布律.

3. 随机变量 (X, Y) 的密度函数

$$f(x, y) = \begin{cases} A(x+y) & (0 \leq x \leq 1, 0 \leq y \leq 2), \\ 0 & (\text{其他}). \end{cases}$$

试求: (1) A ;

(2) 边缘分布密度函数 $f_X(x), f_Y(y)$;

(3) 概率 $P(X+Y \leq 1)$.

4. 设某种晶体管的寿命 X (小时) 的密度函数

$$f(x) = \begin{cases} \frac{100}{x^2} & (x > 100), \\ 0 & (x \leq 100). \end{cases}$$

(1) 求一个晶体管在使用 150 小时后仍未损坏的概率 $P(X > 150)$;

(2) 设在一个仪器中装有 3 个独立工作的这种晶体管, 求当仪器工作 150 小时后恰有两个晶体管损坏的概率.

5. 已知 $X_1, X_2, \dots, X_n$ 是取自于总体 X 的样本, 且总体 X 的密度函数

$$f(x; \theta) = \begin{cases} \theta x^{\theta-1} & (0 < x < 1, \theta > 0), \\ 0 & (\text{其他}), \end{cases}$$

试求: 参数 θ 的矩估计量 $\hat{\theta}$, 极大似然估计量 $\hat{\theta}_L$.

四、证明题(共 8 分)

设随机变量 X 和 Y 相互独立, 且方差 $D(X), D(Y), D(XY)$ 存在, 证明:

$$D(XY) \geq D(X)D(Y).$$

试卷(一)答案与提示

一、选择题

1. 选 C. $p = \frac{C_{55}^2 C_{33}^1}{C_{88}^3} + \frac{C_{55}^3}{C_{88}^3}$. 2. 选 B. $E(X) = \int_0^1 x F'(x) dx$.
3. 选 D. 4. 选 B. 5. 选 C. $\text{cov}(X, Y) = 0 \Leftrightarrow E(XY) = E(X)E(Y)$. 6. 选 A.

二、填空题

1. 0.15. 2. 0.7. 3. $a = \frac{3}{8}$. 4. $a = 0.3, b = 0.3$, 5. 13.
6. 0.6. 7. 4. 8. $N(1, 1.3)$.

三、计算题

1. (1) $P(A) = \sum_{i=1}^3 P(B_i)P(A | B_i) = 0.5$; (2) $P(AB_1) = \frac{1}{9}$.

2. (1) 不独立, 因为

$$P(X=-1, Y=-1) = 0.25 \neq 0.26 = 0.65 \times 0.4 \\ = P(X=-1)P(Y=-1).$$

$$(2) X+Y \sim \begin{pmatrix} -2 & 0 & 1 & 3 & 4 \\ 0.25 & 0.1 & 0.45 & 0.15 & 0.05 \end{pmatrix},$$

$$X-Y \sim \begin{pmatrix} -3 & -2 & 0 & 1 & 3 \\ 0.3 & 0.1 & 0.3 & 0.15 & 0.15 \end{pmatrix}.$$

3. (1) $A=1/3$;

$$(2) f_X(x) = \begin{cases} \frac{2x}{3} + \frac{2}{3} & (0 \leq x \leq 1), \\ 0 & (\text{其他}), \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{y}{3} + \frac{1}{6} & (0 \leq y \leq 2), \\ 0 & (\text{其他}); \end{cases}$$

$$(3) P(X+Y \leq 1) = \frac{1}{3} \int_0^1 dx \int_0^{1-x} (x+y) dy = \frac{1}{9}.$$

$$4. (1) P(X > 150) = \frac{2}{3}; \quad (2) P_3(2) = \frac{2}{9}.$$

$$5. \hat{\theta}_{\text{矩}} = \frac{\bar{X}}{1 - \bar{X}}; \quad \hat{\theta}_L = \frac{-n}{\sum_{i=1}^n \ln(X_i)}.$$

四、证明题

证明: 由题设条件知, X, Y 独立 $\Rightarrow X^2, Y^2$ 独立,

$$\begin{aligned} D(XY) &= E(X^2Y^2) - (E(XY))^2 = E(X^2)E(Y^2) - E^2(X)E^2(Y) \\ &= (D(X) + E^2(X))(D(Y) + E^2(Y)) - E^2(X)E^2(Y) \\ &= D(X)D(Y) + E^2(X)D(Y) + D(X)E^2(Y) \geq D(X)D(Y). \end{aligned}$$

试 卷 (二)

(时间 90 分钟)

一、选择题(每题 3 分,共 24 分)

1. 设事件 A 表示“甲种产品畅销,乙种产品滞销”,其对立事件为_____.

- (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”.

2. 设 $B \subset A$, 则下面正确的等式是_____.

- (A) $P(\overline{AB}) = 1 - P(A)$;
- (B) $P(\overline{B} - \overline{A}) = P(\overline{B}) - P(\overline{A})$;
- (C) $P(B|A) = P(B)$;
- (D) $P(A|\overline{B}) = P(A)$.

3. 离散随机变量 X 的分布函数为 $F(x)$, 且 $x_{k-1} < x_k < x_{k+1}$, 则 $P(X=x_k) =$ _____.

- (A) $P(x_{k-1} \leq X \leq x_k)$;
- (B) $F(x_{k+1}) - F(x_{k-1})$;
- (C) $P(x_{k-1} < X < x_{k+1})$;
- (D) $F(x_k) - F(x_{k-1})$.

4. 设随机变量 X 的分布律为 $P(X=k) = \frac{k}{15} (k=1, 2, 3, 4, 5)$, 则

$P(0.5 < X < 2.5)$ 的值是_____.

- (A) 0.6; (B) 0.2;
- (C) 0.4; (D) 0.8.

5. 设随机变量 X, Y 相互独立, $X \sim N(0, 1), Y \sim N(1, 1)$, 则_____.

$$(A) P(X+Y \leq 0) = \frac{1}{2};$$

$$(B) P(X+Y \leq 1) = \frac{1}{2};$$

$$(C) P(X-Y \leq 0) = \frac{1}{2};$$

$$(D) P(X-Y \leq 1) = \frac{1}{2}.$$

6. 设 10 个电子管的寿命 $X_i (i=1 \sim 10)$ 独立同分布, 且 $D(X_i) = A (i=1 \sim 10)$, 则 10 个电子管的平均寿命 Y 的方差 $D(Y) = \underline{\hspace{2cm}}$.

$$(A) A; \quad (B) 0.1A;$$

$$(C) 0.2A; \quad (D) 10A.$$

7. 设 $(X_1, X_2, \dots, X_n)$ 为总体 $X \sim N(0, 1)$ 的一个样本, $\bar{X}$ 为样本均值, S^2 为样本方差, 则有 $\underline{\hspace{2cm}}$.

$$(A) \bar{X} \sim N(0, 1);$$

$$(B) n\bar{X} \sim N(0, 1);$$

$$(C) \bar{X}/S \sim t(n-1);$$

$$(D) (n-1)X_1^2 / \sum_{i=2}^n X_i^2 \sim F(1, n-1).$$

8. 设 $X_1, X_2, \dots, X_n$ 是取自总体 X 的样本, 则 $D(X) = \sigma^2$ 的无偏估计量为 $\underline{\hspace{2cm}}$.

$$(A) \frac{1}{n} \sum_{i=1}^{n-1} (X_i - \bar{X})^2; \quad (B) \frac{1}{n-1} \sum_{i=1}^{n-1} (X_i - \bar{X})^2;$$

$$(C) \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2; \quad (D) \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

二、填空题(每题 5 分, 共 30 分)

1. 用随机事件 A, B, C 表示事件 $D = \{A, B, C \text{ 中恰有两个发生}\} = \underline{\hspace{2cm}}$.

2. 设 $P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}$, 且三事件 A_1, A_2, A_3 相互独

立,则三事件中至少发生一个的概率为_____,三事件中恰好发生一个的概率为_____.

3. 设 $X \sim U(0, 2)$, 则随机变量 $Y = X^2$ 在 $(0, 4)$ 内的概率密度函数为_____.

4. 若随机变量 X 满足 $E(X^2) = 200, D(X) = 100$, 则 $E(X) =$ _____.

5. 设样本观察值为 5, 10, 15, 20, 15, 则样本均值、方差 $\bar{x} =$ _____, $s^2 =$ _____.

6. 原假设 H_0 不真时, 作出接受 H_0 的决策, 称为犯第_____类错误; 原假设 H_0 为真时, 作出拒绝 H_0 的决策, 称为犯第_____类错误.

三、计算题(每题 10 分, 共 40 分)

1. 编号为 1, 2, 3 的三台仪器正在工作的概率分别为 0.9, 0.8 和 0.4, 从中任选一台.

(1) 求此台仪器正在工作的概率;

(2) 已知选到的仪器正在工作, 求它编号为 2 的概率.

2. 随机变量 X 的密度函数 $f(x) = \begin{cases} Ax^2 & (0 \leq x \leq 2), \\ 0 & (\text{其他}). \end{cases}$

试求:

(1) 系数 A ;

(2) 分布函数 $F(x)$;

(3) 概率 $P(1 \leq X \leq 2)$.

3. 设随机变量 $X \sim U(0, 1)$ (均匀分布), $Y \sim E(1)$ (指数分布), 且它们相互独立, 计算 $P(X > Y)$.

4. (1) 设某种产品的一项质量指标 $X \sim N(1600, 150^2)$, 现从一批产品中随机地抽取 26 件, 测得该指标的均值 $\bar{x} = 1637$, 试以 $\alpha = 0.05$ 检验这批产品的质量指标是否合格.

(2) 设某种电阻器的电阻值 $X \sim N(\mu, 60)$, μ 未知. 某天抽取 10 只这种电阻器, 测得电阻值的方差 $s^2 = 87.682$, 问方差有无显著变化 (取 $\alpha = 10\%$) ?

$$[z_{0.025} = 1.96, z_{0.05} = 1.65, t_{0.05}(9) = 1.8331, \chi^2_{0.05}(9) = 16.919, \chi^2_{0.95}(9) = 3.325]$$

四、证明题(6分)

设随机事件 A 和 B 相互独立, 证明: $\bar{A}$ 和 $\bar{B}$ 也相互独立.

试卷(二)答案与提示

一、选择题

1. 选 D. $\overline{AB} = \bar{A} \cup B$, 其中 A 为甲种产品畅销, B 为乙种产品畅销.
2. 选 B. 利用事件差的概率性质.
3. 选 D. 因为 $P(X=a) = F(a) - F(a-0)$.
4. 选 B. $P(X=1) + P(X=2)$.
5. 选 B. 利用正态分布的对称性 $P(X \leq \mu) = 1/2$.
6. 选 B. 由 $D(Y) = D\left(\frac{1}{10} \sum_{i=1}^{10} X_i\right) = \frac{1}{100} \sum_{i=1}^{10} D(X_i)$ 解得.
7. 选 D. 根据 F 分布定义.
8. 选 D. 因为样本方差 S^2 是 $D(X) = \sigma^2$ 的无偏估计量.

二、填空题

1. $D = AB\bar{C} \cup A\bar{B}C \cup \bar{A}BC$.
2. $\frac{19}{27}, \frac{4}{9}$.
3. $f(y) = \begin{cases} \frac{0.25}{\sqrt{y}} & (0 < y < 4), \\ 0 & (\text{其他}). \end{cases}$
4. 10.
5. 13, 32.5.

6. 二,一.

三、计算题

$$1. (1) P(A) = \sum_{i=1}^3 P(B_i)P(A | B_i) = 0.7;$$

$$(2) P(B_2 | A) = \frac{8}{21} \approx 0.38.$$

$$2. (1) A = \frac{3}{8}; \quad (2) F(x) = \begin{cases} 0 & (x < 0), \\ \frac{x^3}{8} & (0 \leq x < 2), \\ 1 & (x \geq 2); \end{cases}$$

$$(3) P(1 \leq X \leq 2) = \frac{7}{8}.$$

$$3. P(X > Y) = \iint_{x > y} f(x, y) dx dy = \int_0^1 dx \int_0^x e^{-y} dy = e^{-1}.$$

$$4. (1) H_0: \mu = 1600, H_1: \mu \neq 1600; \text{拒绝域 } D: |U| = \left| \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{26}}} \right| >$$

$z_{0.025} = 1.96$. 代入数据得 $|U| = 1.2578 < 1.96$, 落在 D 外, 故接受 H_0 , 即质量指标合格.

$$(2) H_0: \sigma^2 = 60, H_1: \sigma^2 \neq 60. \text{拒绝域: } D: \chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \geq \chi_{0.05}^2(9) =$$

$$16.919 \text{ 或 } \chi^2 \leq \chi_{0.95}^2(9) = 3.325. \text{代入数据得 } \chi^2 = \frac{9 \times 87.682}{60} = 13.152 \notin$$

D , 故接受 H_0 即方差无显著变化.

四、证明题

$$\begin{aligned} P(\bar{A} \bar{B}) &= 1 - P(A \cup B) = 1 - P(A) - P(B) + P(A)P(B) \\ &= P(\bar{A}) - P(B)[1 - P(A)] = P(\bar{A}) - P(B)P(\bar{A}) \\ &= P(\bar{A})[1 - P(B)] = P(\bar{A})P(\bar{B}). \end{aligned}$$

试 卷 (三)

(时间 90 分钟)

一、选择题(每题 3 分,共 24 分)

1. 设事件 A 和 B 互斥, $P(A) > 0, P(B) > 0$, 则下列结论中一定成立的有_____.

- (A) $\bar{A}$ 和 $\bar{B}$ 互不相容; (B) A, B 为对立事件;
(C) A 和 B 相互独立; (D) A 和 B 不独立.

2. 一盒零件有 5 个正品, 2 个次品, 不放回任取 3 个, 其中至少有 2 个正品的概率为_____.

- (A) $\frac{2}{7}$; (B) $\frac{4}{7}$; (C) $\frac{5}{7}$; (D) $\frac{6}{7}$.

3. 某人射击中靶的概率为 0.75. 若射击直到中靶为止, 则射击次数为 3 的概率为_____.

- (A) $(0.75)^3$; (B) $0.75(0.25)^2$;
(C) $0.25(0.75)^2$; (D) $(0.25)^3$.

4. 下列各函数中可以作为某个随机变量 X 的分布函数的是_____.

(A) $F(x) = \sin x$;

(B) $F(x) = \frac{1}{1+x^2}$;

(C) $F(x) = \begin{cases} \frac{1}{1+x^2} & (x \leq 0), \\ 1 & (x > 0); \end{cases}$

(D) $F(x) = \begin{cases} 0 & (x < 0), \\ 1.1 & (0 \leq x \leq 1), \\ 1 & (x > 1). \end{cases}$

5. 设随机变量 $X \sim B(1, p), Y \sim \pi(\lambda)$, 且 X, Y 相互独立, 则 $X + Y$ _____.

(A) 是一维随机变量; (B) 是二维随机变量;

(C) 服从两点分布; (D) 服从泊松分布.

6. 设随机变量 X 的分布律为 $P(X=a)=0.6, P(X=b)=p$ ($a < b$). 又 $E(X)=1.4, D(X)=0.24$, 则 a, b 的值为_____.

(A) $a=1, b=2$; (B) $a=-1, b=2$;

(C) $a=1, b=-2$; (D) $a=0, b=1$.

7. 设 $X_1, X_2, \dots, X_n$ 是来自正态总体 $N(\mu, 1)$ 的一个简单随机样本, $\bar{X}, S^2$ 分别为样本均值和样本方差, 则_____.

(A) $\bar{X} \sim N(0, 1)$;

(B) $\sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi^2(n-1)$;

(C) $\sum_{i=1}^n (X_i - \mu)^2 \sim \chi^2(n)$;

(D) $\frac{\bar{X}}{S/\sqrt{n-1}} \sim t(n-1)$.

8. 在 H_0 为原假设, H_1 为备择假设的假设检验中, 若显著性水平为 α , 则_____.

(A) $P(\text{接受 } H_0 | H_0 \text{ 成立}) = \alpha$;

(B) $P(\text{接受 } H_1 | H_1 \text{ 成立}) = \alpha$;

(C) $P(\text{接受 } H_1 | H_0 \text{ 成立}) = \alpha$;

(D) $P(\text{接受 } H_0 | H_1 \text{ 成立}) = \alpha$.

二、填空题(每题 5 分, 共 30 分)

1. 设 $P(A)=3P(B)=\frac{2}{3}$, A 和 B 都不发生的概率是 A 和 B 同时发生的概率的 2 倍, 则 $P(A-B)=$ _____.

2. 设 A, B 为两随机事件, 已知 $P(A)=0.7=0.3+P(B), P(A \cup B)=0.8$, 则 $P(A|\bar{A} \cup B)=$ _____.

$$3. \text{ 设随机变量 } X \text{ 的密度函数 } f(x) = \begin{cases} \frac{1}{3} & (0 \leq x \leq 1), \\ \frac{2}{9} & (3 \leq x \leq 6), \\ 0 & (\text{其他}), \end{cases} \text{ 若 } k \text{ 满足}$$

$P(X \geq k) = \frac{2}{3}$, 则 k 的取值范围是_____.

4. 设随机变量 $X \sim N(1.04, 1)$, 已知 $P(X \leq 3) = 0.975$, 则 $P(X \leq -0.92) =$ _____.

5. 设随机变量 X, Y 满足 $D(X) = 4, D(Y) = 1, D(3X - 2Y) = 28$, 则 $\rho_{XY} =$ _____.

6. 设总体 $X \sim U(0, \theta)$, $X_1, X_2, \dots, X_n$ 为总体的一个样本, 则未知参数 θ 的矩估计量为_____; 极大似然估计量为_____.

三、计算题(每题 10 分, 共 40 分)

1. 某电脑公司组装的电脑所用的显示屏是由 3 家工厂提供的(数据见下表), 现从待出厂的电脑中任抽一台检验发现是次品(设为事件 A), 原因是显示屏有问题.

(1) 求 $P(A)$;

(2) 有问题的显示屏由哪家厂提供的可能性最大?

显示屏制造厂	提供份额	次品率
1	0.15	0.03
2	0.60	0.01
3	0.25	0.02

2. 设随机变量 (X, Y) 的联合密度函数

$$f(x, y) = \begin{cases} 3x & (0 \leq x \leq 1, 0 \leq y \leq x), \\ 0 & (\text{其他}). \end{cases}$$

(1) 求边缘密度函数 $f_X(x)$ 与 $f_Y(y)$;

(2) X 和 Y 是否相互独立? 为什么?

(3) 计算 $P(X+Y>1)$.

3. 某意外事故 A 发生的概率为 p . 若 A 发生, 保险公司要赔偿给投保者 M 元. 为使公司的期望收益达到 $0.05M$ 元, 公司将要求投保者交纳多少保费?

4. 机器自动包装食盐, 设每袋盐的净重服从正态分布, 规定每袋盐的标准重量为 500 g , 标准差不能超过 10 g . 某天开工后, 为了检验机器是否正常工作, 从已经包装好的食盐中随机抽取 9 袋, 测得 $\bar{X}=499$, $S^2=16.03^2$. 问这天自动包装机工作是否正常 ($\alpha=0.05$)? 即检验: (1) $H_0: \mu=500, H_1: \mu \neq 500$; (2) $H_0: \sigma^2 \leq 10^2, H_1: \sigma^2 > 10^2$.
[$t_{0.025}(8)=2.306, t_{0.025}(9)=2.262, \chi_{0.025}^2(8)=17.535, \chi_{0.025}^2(9)=19.023,$
 $t_{0.05}(8)=1.8595, t_{0.05}(9)=1.8331, \chi_{0.05}^2(8)=15.507, \chi_{0.05}^2(9)=16.919]$

四、证明题(6分)

设事件 A, B, C 同时发生必导致事件 D 发生, 证明: $P(A) + P(B) + P(C) \leq 2 + P(D)$.

试卷(三)答案与提示

一、选择题

1. 选 D. 可以证明: 当 $P(A)>0, P(B)>0$ 时, 事件 A 和 B 互斥, 与 A 和 B 独立不能同时成立, 故只有 A 和 B 不独立.

2. 选 D. 利用超几何分布计算

$P(\text{至少有 } 2 \text{ 个正品}) = P(\text{恰有 } 2 \text{ 个正品}) + P(3 \text{ 个都是正品}).$

3. 选 B. 射击次数服从几何分布.

4. 选 C. 只有 C 中的 $F(x)$ 满足分布函数的三个性质.

5. 选 A. $X+Y=Z, Z$ 显然是一维随机变量.

6. 选 A. 由 $\begin{cases} 0.6a+0.4b=1.4, \\ 0.6a^2+0.4b^2=0.24+1.4^2 \end{cases}$ 解得.

7. 选 C. 因为 $X_i - \mu \sim N(0, 1)$, 再由 χ^2 分布定义.

8. 选 D. 只有 D 中表达的是犯第一类弃真错误的概率.

二、填空题

1. $\frac{5}{9}$. 2. 0.55 3. $k \in [1, 3]$. 4. 0.025. 5. 0.5. 6. $\hat{\theta} =$

$$2\bar{X}; \hat{\theta}_L = \max\{X_1, X_2, \dots, X_n\}.$$

三、计算题

1. (1) $P(A) = \sum_{i=1}^3 P(B_i)P(A | B_i) = 0.0155;$

(2) $P(B_1 | A) = 0.290, P(B_2 | A) = 0.387, P(B_3 | A) = 0.323.$

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

$$f_Y(y) = \begin{cases} \frac{3(1-y^2)}{2} & (0 \leq y \leq 1), \\ 0 & (\text{其他}). \end{cases}$$

(2) 不独立.

(3) $P(X+Y > 1) = \int_{0.5}^1 dx \int_{1-x}^x 3x dy = \frac{5}{8}.$

3. 公司收益 $X \sim \begin{array}{c|cc} & a & a-M \\ \hline P & 1-p & p \end{array}$, 投保者交纳的保费为 a .

$$E(X) = a(1-p) + p(a-M) = 0.05M \Rightarrow a = (0.05 + p)M.$$

4. (1) $H_0: \mu=500, H_1: \mu \neq 500$; 拒绝域 $D: \left| \frac{\bar{X}-500}{\frac{S}{3}} \right| > t_{0.025}(8) =$

2.306. 代入数据得 T 的观察值 $T_0 = -\frac{3}{16.03} = -0.187$, 因 $T_0 \notin D$, 故接受 H_0 .

(2) $H_0: \sigma^2 \leq 10^2, H_1: \sigma^2 > 10^2$. 拒绝域 $D: \chi^2 = \frac{(n-1)S^2}{\sigma_0^2} > \chi_{0.05}^2(8) =$

15.507. 代入数据得 $\frac{8 \times 16.03^2}{100} = 20.56 \in D$, 故应拒绝 H_0 .

四、证明题

证明: 由题设条件知 $ABC \subset D \Rightarrow P(ABC) \leq P(D)$,
 $P(A) + P(B) - P(AB) \leq 1 \Rightarrow P(A) + P(B) \leq 1 + P(AB)$
 $\Rightarrow P(A) + P(B) + P(C) \leq 1 + P(AB) + P(C)$
 $= 1 + P(AB \cup C) + P(ABC)$
 $\leq 2 + P(ABC) \leq 2 + P(D).$

试 卷 (四)

(时间 120 分钟)

一、填空题(每题 3 分,共 30 分)

1. 如果 $P(A)=0.4, P(B)=0.3, P(A \cup B)=0.5$, 则 $P(A\bar{B})=$ _____.
2. 设 A, B, C 是三个事件, 且 $P(A)=P(B)=P(C)=1/4, P(AB)=P(BC)=0, P(AC)=1/8$, 则 A, B, C 至少发生一个的概率为 _____.
3. 设随机变量 X 的分布函数

$$F(x) = \begin{cases} 0 & (\text{当 } x < -1 \text{ 时}), \\ 0.3 & (\text{当 } -1 \leq x < 1 \text{ 时}), \\ 0.7 & (\text{当 } 1 \leq x < 3 \text{ 时}), \\ 1 & (\text{当 } 3 \leq x \text{ 时}), \end{cases}$$

则 X 的概率分布律为 _____.

4. 已知 $D(X)=4, D(Y)=9, D(X-Y)=12$, 则 X 与 Y 间的相关系数 $\rho=$ _____.

5. 设随机变量 X 和 Y 相互独立, 且服从同一分布, X 的分布律 $P(X=0)=P(X=1)=\frac{1}{2}$, 则 $Z=\max(X, Y)$ 的分布律为 _____.

6. 设 X, Y 相互独立, 且都服从标准正态分布, 则 $Z=\frac{X}{\sqrt{Y^2}}$ 服从 _____ 分布(同时写出分布的参数).

7. 设 $X_1, X_2, \dots, X_{20}$ 是来自总体 $N(\mu, \sigma^2)$ 的样本, 则 $\frac{1}{\sigma^2} \sum_{i=1}^{20} (X_i - \mu)^2$ 服从 _____ 分布(同时写出分布的参数).

8. 设 X_1, X_2, X_3 为从总体 X 中抽取的容量为 3 的样本, 总体均值为 θ , 总体方差为 σ^2 . 记 $\hat{\theta}_1 = \frac{1}{6} X_1 + \frac{1}{3} X_2 + \frac{1}{2} X_3, \hat{\theta}_2 = \frac{1}{3} X_1 + \frac{1}{3} X_2 +$

$\frac{1}{3}X_3, \hat{\theta}_3 = \frac{1}{4}X_1 + \frac{1}{2}X_2 + \frac{1}{4}X_3$ 分别为未知参数 θ 的估计, 则_____为 θ 的无偏估计, 且此三个估计中_____最有效.

9. 如果一个假设检验问题的显著性水平为 0.05, 那么犯第一类错误的概率是_____.

二、(10 分) 在一道答案有 4 种选择的单项选择题测验中, 若一个学生不知道题目的正确答案, 他就从 4 个答案中任选 1 个. 已知有 80% 的学生知道正确答案, 现在某个学生答对了此题, 问他确实知道正确答案的概率为多少?

三、(20 分) 设随机变量 X 和 Y 的联合密度函数

$$f(x, y) = \begin{cases} ce^{-y} & (0 \leq x \leq y), \\ 0 & (\text{其他}). \end{cases}$$

- (1) 求常数 c ;
- (2) 求 X 和 Y 各自的边缘密度函数;
- (3) X 和 Y 是否独立? 为什么?
- (4) $P(X+2Y < 1)$.

四、(10 分) 将一枚均匀硬币掷 400 次, 计算正面出现的次数大于 220 的概率.

五、(10 分) 设 $X_1, X_2, \dots, X_n$ 是来自参数为 λ 的泊松分布总体 X 的一个样本, 试求:

- (1) λ 的极大似然估计;
- (2) $P(X=0)$ 的极大似然估计.

六、(10 分) 设总体 X 以概率 $\frac{1}{\theta}$ 取值 $1, 2, \dots, \theta$, 求未知参数 θ 的矩估计.

七、(10 分)某电子产品的一个指标服从正态分布,从某天生产的产品中抽取 15 个产品,测得该指标的样本均值为 2.80,样本标准差为 0.215.

(1) 取显著性水平 $\alpha=0.05$,是否可以认为该指标的平均值显著地大于 2?

(2) 求该指标的方差的 0.95 的置信区间.

以下分位数全部为下侧 α 分位数:

$\Phi(1)=0.8413; \Phi(1.645)=0.95; \Phi(1.95)=0.974; \Phi(1.96)=0.975;$

$\Phi(2)=0.9773; \Phi(2.05)=0.9798; \Phi(3)=0.9985;$

$\chi^2_{0.95}(15)=24.996; \chi^2_{0.975}(15)=27.488; \chi^2_{0.95}(14)=23.685; \chi^2_{0.975}(14)=$

$26.119; \chi^2_{0.05}(15)=7.261; \chi^2_{0.025}(15)=6.262; \chi^2_{0.05}(14)=6.571; \chi^2_{0.025}(14)=$

$5.629; t_{0.95}(15)=1.7531; t_{0.975}(15)=2.1315; t_{0.95}(14)=1.7613; t_{0.975}(14)=$

2.1448.

试卷(四)答案与提示

一、填空题

1. 0.2. $P(A\bar{B})=P(A-AB)=P(A)-P(AB).$

2. $\frac{5}{8}$. 由 $ABC \subset AB \Rightarrow 0 \leq P(ABC) \leq P(AB)=0.$

3. $X \sim \begin{pmatrix} -1 & 1 & 3 \\ 0.3 & 0.4 & 0.3 \end{pmatrix}$. 由 $P(X=a)=F(a)-F(a-0)$ 求得.

4. $\rho=\frac{1}{12}$. $D(X-Y)=D(X)+D(Y)-2\rho\sqrt{D(X)}\sqrt{D(Y)}.$

5. $Z=\max(X,Y) \sim \begin{pmatrix} 0 & 1 \\ 0.25 & 0.75 \end{pmatrix}.$

6. $t(1)$ 分布. 由 t 分布定义.

7. $\chi^2(20)$. 根据 χ^2 分布定义.

8. $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3; \hat{\theta}_2$ 最有效.

9. $\alpha=0.05$.

二、计算题

$P(B|A)=0.9412$. A 表事件“某学生答对此题”, B 为“某学生确实知道正确答案”, 利用贝叶斯公式求解.

三、(1) $c=1$;

$$(2) f_X(x) = \begin{cases} e^{-x} & (x > 0), \\ 0 & (x \leq 0), \end{cases}$$

$$f_Y(y) = \begin{cases} ye^{-y} & (y > 0), \\ 0 & (y \leq 0); \end{cases}$$

(3) X 和 Y 不独立, 因为 $f(x, y) \neq f_X(x)f_Y(y)$;

$$(4) 1 - 3e^{-\frac{1}{3}} + 2e^{-\frac{1}{2}}. P(X + 2Y < 1) = \iint_{x+2y < 1} f(x, y) dx dy =$$

$$\int_0^{\frac{1}{3}} dx \int_x^{\frac{1}{2}-\frac{x}{2}} e^{-y} dy.$$

四、0.0227 提示: 利用中心极限定理.

$$\text{五、(1) } \hat{\lambda}_L = \bar{X}; \text{提示: 似然函数为 } L(\lambda) = \frac{e^{-n\lambda} \lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}.$$

(2) $\hat{P}(X=0)_L = e^{-\bar{X}}$. 提示: 利用极大似然估计不变性.

六、 $\hat{\theta}_{\text{矩}} = 2\bar{X} - 1$. 提示: 令 $E(X) = \frac{1}{\theta}(1+2+\dots+\theta) = \bar{X}$.

七、(1) $H_0: \mu=2, H_1: \mu>2$; 拒绝域 $D: T = \frac{\bar{X}-2}{\frac{0.215}{\sqrt{15}}} > t_{0.95}(14) =$

1.7613, 代入数据得 $T=14.4111 > 1.7613$, 落在 D 内, 故拒绝 H_0 , 即指标均值显著大于 2.

(2) (0.025, 0.115).

试 卷 (五)

(时间 120 分钟)

一、判断题(每题 2 分,共 16 分)

1. 设样本空间 $\Omega = \{e_1, e_2, e_3, e_4, e_5\}$, $A = (e_1, e_3, e_5)$, 则 $P(A) = 0.6$. ()
2. 如果 $P(A) = P(B) = 0.5$, 则 $P(AB) = P(\overline{A}\overline{B})$. ()
3. 设 n 次独立重复试验中, 事件 A 出现的次数为 X , 则 $4n$ 次独立重复试验中, A 出现的次数为 $4X$. ()
4. 如果随机变量 $X \sim N(\mu, \sigma^2)$, 则 $\frac{(\mu - X)}{\sigma} \sim N(0, 1)$. ()
5. 二维均匀分布的边缘分布仍是均匀分布. ()
6. 若随机变量 X 的数学期望存在, 则 X 的方差也存在. ()
7. 样本二阶中心矩不是总体方差的无偏估计. ()
8. 假设检验中, 样本容量确定时, 犯弃真错误和取伪错误的概率不能同时减小. ()

二、选择题(每题 4 分,共 24 分)

1. 如果 $P(A) + P(B) > 1$, 则 A 与 B 必定_____
(A) 独立; (B) 不独立;
(C) 相容; (D) 不相容.
2. 设随机变量 X 的密度函数为 $f(x)$, 如果_____, 则恒有 $0 \leq f(x) \leq 1$.
(A) $X \sim N(0, 1)$; (B) $X \sim N(0, \sigma^2)$;
(C) $X \sim N(-1, \sigma^2)$; (D) $X \sim N(\mu, \sigma^2)$.
3. 人的体重为随机变量 X , $E(X) = a$, $D(X) = b$, 10 个人的平均体重记为 Y , 则_____.
(A) $E(Y) = a$; (B) $E(Y) = 0.1a$;

(C) $D(Y)=0.01b$;

(D) $D(Y)=b$.

4. 设 (X, Y) 的联合概率密度

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

则 X 和 Y 为_____的随机变量.

(A) 独立同分布;

(B) 独立不同分布;

(C) 不独立同分布;

(D) 不独立不同分布.

5. 设 X 为随机变量, 若 $E(X^2)=1.1$, $D(X)=0.1$, 则一定有_____.

(A) $P(-1 < X < 1) \geq 0.9$; (B) $P(0 < X < 2) \geq 0.9$;

(C) $P(X+1 \geq 1) < 0.9$; (D) $P(|X| \geq 1) \leq 0.1$.

6. 设 $X \sim N(0, 1)$, $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$, 服从自由度为 $(n-1)$ 的 χ^2 分布的随机变量是_____.

(A) $\sum_{i=1}^n X_i^2$;

(B) S^2 ;

(C) $(n-1)\bar{X}^2$;

(D) $(n-1)S^2$.

三、计算题(共 60 分).

1. (10 分) 某工厂有四种机床: 车床、钻床、磨床和刨床, 其台数之比为 $9:3:2:1$, 而在一定时间内需要修理的台数之比为 $1:2:3:1$:

1. 当有一台机床需要修理时, 问这台机床是车床的概率是多少?

2. (15 分) 设二维随机变量 (X, Y) 的联合密度函数

$$f(x, y) = \begin{cases} cxy & (0 \leq x \leq 1, 0 \leq y \leq 1), \\ 0 & (\text{其他}). \end{cases}$$

试求: (1) 系数 c ;

(2) X 和 Y 各自的边缘密度函数;

(3) $P(X < Y)$;

(4) X 和 Y 相互独立吗? 为什么?

3. (10 分)某工厂有 100 台机器,各台机器独立工作,每台机器的开工率为 0.8,工作时各需要 1 kW 电力.问供电局至少要供应多少电力,才能以 97.5%的把握保证正常生产?

4. (10 分)设总体 X 的密度函数

$$f(x) = \begin{cases} \theta x^{\theta-1} & (0 \leq x \leq 1), \\ 0 & (\text{其他}), \end{cases}$$

其中 θ 未知.求 θ 的矩法估计.

5. (15 分)正常人的脉搏平均为 72 次/分.某医生测得 10 例慢性铅中毒患者的脉搏均值为 67.4 次/分,标准差为 5.929.设人的脉搏次数/分近似服从正态分布.

(1) 取 $\alpha=0.05$,是否可以认为铅中毒患者的脉搏均值显著地小于 72 次/分.

(2) 求铅中毒患者脉搏均值的 0.95 的置信区间.

以下分位数全部为下侧 α 分位数:

$\Phi(0.50)=0.6915$; $\Phi(0.55)=0.7088$; $\Phi(1.00)=0.8413$; $\Phi(1.05)=0.8531$; $\Phi(1.65)=0.95$;

$\Phi(1.95)=0.9744$; $\Phi(1.96)=0.9750$; $\Phi(2.00)=0.9772$; $\Phi(2.05)=0.9798$; $\Phi(3)=0.9987$;

$\chi^2_{0.95}(9)=16.919$; $\chi^2_{0.975}(9)=19.023$; $\chi^2_{0.95}(10)=18.307$; $\chi^2_{0.975}(10)=20.483$;

$\chi^2_{0.05}(9)=3.325$; $\chi^2_{0.025}(9)=2.7$; $\chi^2_{0.05}(10)=3.94$; $\chi^2_{0.025}(10)=3.247$;
 $t_{0.95}(9)=1.8331$; $t_{0.975}(9)=2.2622$; $t_{0.95}(10)=1.8125$; $t_{0.975}(10)=2.2281$.

试卷(五)答案与提示

一、判断题

1. 非. 因为样本空间中每一样本点 e_i 未必等可能发生.

2. 是. 由 $P(AB)=1-P(\bar{A}\cup\bar{B})=1-P(\bar{A})-P(\bar{B})+P(\bar{A}\bar{B})$

可得.

3. 非. $4n$ 次独立重复试验中, A 出现的次数为另一随机变量 Y .
4. 是. 因为正态变量的线性函数仍为正态变量, 且

$$E\left(\frac{\mu - X}{\sigma}\right) = 0, \quad D\left(\frac{\mu - X}{\sigma}\right) = 1.$$

5. 非. 因为圆域上的均匀分布的边缘分布不再是均匀分布.

6. 非. 因为 X 的一阶原点矩的存在并不能保证 X 的二阶中心矩的存在.

7. 是. 由于 $E(S_n^2) = \left(\frac{n-1}{n}S^2\right) = \frac{n-1}{n}\sigma^2 \neq \sigma^2$, 其中 $\sigma^2 = D(X)$.

8. 是. 在 U 检验法的单边(右边)假设检验中可得到 $z_\alpha + z_\beta =$

$\frac{\delta\sqrt{n}}{\sigma_0}$, 当 n 固定时, 等式右边是常数, 从而有

$$\alpha \uparrow \Rightarrow z_\alpha \downarrow \Rightarrow z_\beta \uparrow \Rightarrow \beta \downarrow \quad \text{或} \quad \beta \uparrow \Rightarrow z_\beta \downarrow \Rightarrow z_\alpha \uparrow \Rightarrow \alpha \downarrow,$$

即两类错误的概率形成跷跷板现象.

二、选择题

1. 选 C. 若 A 和 B 不相容, 即 $AB = \Phi \Rightarrow P(AB) = 0$, 从而得如下矛盾:

$$P(A \cup B) = P(A) + P(B) > 1.$$

2. 选 A. 当 $X \sim N(0, 1)$ 时, 有 $0 \leq f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \leq \frac{1}{\sqrt{2\pi}} \leq 1$.

3. 选 A. $E(Y) = E\left(\frac{1}{10} \sum_{i=1}^{10} X_i\right) = \frac{1}{10} \sum_{i=1}^{10} E(X_i) = a$.

4. 选 C. 由于 $f(x, y)$ 在圆域 $x^2 + y^2 < 1$ 内取非零值, 而圆域 $x^2 + y^2 < 1$ 无法分离变量, 故 X 和 Y 不独立; 又由于 $f(x, y)$ 的对称性, $f_X(x)$ 与 $f_Y(y)$ 相同.

5. 选 B. 由切比雪夫不等式 $P(0 < X < 2) = P(|X - 1| < 1) \geq 1 -$

$$\frac{0.1}{1^2} = 0.9.$$

6. 选 D. 由定理 $\frac{(n-1)S^2}{\sigma^2} = (n-1)S^2 \sim \chi^2(n-1)$.

三、计算题

1. $P(B_1|A) = \frac{9}{22} \approx 0.409$. A 表事件“有一台机床需要修理”, B_i ($i=1, 2, 3, 4$) 分别表示“车床、钻床、磨床和刨床”.

$$P(A) = \sum_{i=1}^4 P(B_i)P(A|B_i) = \frac{22}{105}.$$

2. (1) $c=4$; (2) $f_X(x) = \begin{cases} 2x & (0 < x < 1), \\ 0 & (\text{其他}), \end{cases}$

$f_Y(y) = \begin{cases} 2y & (0 < y < 1), \\ 0 & (\text{其他}); \end{cases}$ (3) 0.5 . $P(X < Y) = \iint_{x < y} f(x, y) dx dy =$

$\int_0^1 4x dx \int_x^1 y dy = 0.5$, 或者利用 $f(x, y)$ 的对称性直接得 $P(X < Y) = 0.5$; (4) X 和 Y 相互独立, 因为 $f(x, y) = f_X(x)f_Y(y)$.

3. 87.84(kW). 提示: $X \sim B(100, 0.8)$, 利用中心极限定理.

4. $\hat{\theta}_{\text{矩}} = \frac{\bar{X}}{1-\bar{X}}$. 提示: 令 $E(X) = \int_0^1 \theta x^\theta dx = \bar{X}$.

5. (1) $H_0: \mu=72, H_1: \mu < 72$; 拒绝域 $D: T = \frac{\bar{X}-72}{\frac{5.929}{\sqrt{10}}} < -t_{0.95}(9) =$

-1.8331 , 代入数据得 $T = -2.4534 < -1.8331$, 落在 D 内, 故拒绝 H_0 , 即铅中毒患者的脉搏均值显著地小于 72 次/分.

(2) (67.159, 71.641).