College Mathematics Guidance Series
大学数学学习辅导丛书

概率论与数理统计
习题全解指南

浙大 • 第五版

□ 浙江大学 盛 骤 谢式千 潘承毅 编

高等教育出版社

College Mathematics Guidance Series
大学数学学习辅导丛书

概率论与数理统计
习题全解指南

浙大·第五版

□ 浙江大学 盛 骤 谢式千 潘承毅 编

高等教育出版社·北京

内容提要

本书是浙江大学盛骤等编的《概率论与数理统计》（第五版）的配套辅导书，全书按照主教材的要求和章节顺序进行编排，与主教材习题一致。本书对教材的300多道题目给出了解答，少数题目是一题多解，有些作了题目分析、解题思路分析和解答方法归纳，并指出易犯的错误，究其原因，澄清不正确的想法。通过本书的学习，可使读者提高分析问题和解题的能力，加深对基本内容的理解和掌握。

本书可作为理工科和其他非数学类专业的学生学习概率论与数理统计的参考书。

图书在版编目(CIP)数据

概率论与数理统计习题全解指南：浙大·五版 / 盛骤，谢式千，潘承毅编. --北京：高等教育出版社，2020.10

ISBN 978-7-04-051548-0

Ⅰ.①概… Ⅱ.①盛… ②谢… ③潘… Ⅲ.①概率论-高等学校-解题 ②数理统计-高等学校-解题 Ⅳ.①O21-44

中国版本图书馆CIP数据核字(2019)第040907号

Gailülun yu Shuli Tongji Xiti Quanjie Zhinan

策划编辑 李蕊　责任编辑 田玲　封面设计 王凌波　版式设计 马敬茹
插图绘制 于博　责任校对 刘娟娟　责任印制 存怡

出版发行 高等教育出版社
社　　址 北京市西城区德外大街4号
邮政编码 100120
印　　刷 大厂益利印刷有限公司
开　　本 787mm×960mm 1/16
印　　张 19.25
字　　数 350千字
购书热线 010-58581118
咨询电话 400-810-0598
网　　址 http://www.hep.edu.cn
　　　　 http://www.hep.com.cn
网上订购 http://www.hepmall.com.cn
　　　　 http://www.hepmall.com
　　　　 http://www.hepmall.cn
版　　次 2020年10月第1版
印　　次 2020年10月第1次印刷
定　　价 36.00元

物 料 号 51548-00

更多电子书，微信：taobao66taobao或淘宝店铺：菁华资料

前　言

本书是编者所编的《概率论与数理统计》(第五版)(高等教育出版社 2019 年出版)的习题解答。

本书按教材各章习题顺序编排，与教材的习题一致，少数题目有一题多解。在有些题解中，指出易犯的错误，究其原因，澄清不正确的想法。在有些题解中，指出在解题时应当注意的事项。

通过对本书的参考和学习，可使读者提高分析问题和解题的能力，加深对基本内容的理解和掌握，还会增强对学好本门课程的信心和兴趣。

我们希望读者先独立思考，自己解题，然后与题解进行对照。如果自己不动手去做题而照抄，那样学习，收效甚微。

本书可作为大学理科、工科学生学习概率论与数理统计课程的参考书，可供报考研究生的读者和工程技术人员参考。

本书承浙江大学范大茵教授详细审阅，她提出许多宝贵意见，对此我们表示衷心的感谢。

本书不足之处，诚恳地希望读者批评指正。

盛　骤　谢式千　潘承毅

2019 年 4 月

目　录

第一章 概率论的基本概念

1. 写出下列随机试验的样本空间 S：

(1) 记录一个班一次数学考试的平均分数(设以百分制记分).

(2) 生产产品直到有 10 件正品为止，记录生产产品的总件数.

(3) 对某工厂出厂的产品进行检查，合格的记上“正品”，不合格的记上“次品”，如连续查出了 2 件次品就停止检查，或检查了 4 件产品就停止检查，记录检查的结果.

(4) 在单位圆内任意取一点，记录它的坐标.

解 (1) 以 n 表示该班的学生数，总成绩的可能取值为 $0,1,2,3,\cdots,100n$，所以试验的样本空间为

$$S=\left\{\frac{i}{n}\,\middle|\,i=0,1,2,\cdots,100n\right\}.$$

(2) 设在生产第 10 件正品前共生产了 k 件不合格品，样本空间为 $S=\{10+k\,|\,k=0,1,2,\cdots\}$ 或写成 $S=\{10,11,12,\cdots\}$.

(3) 采用 0 表示检查到一件次品，以 1 表示检查到一件正品，例如 0110 表示第一次与第四次检查到次品，而第二次与第三次检查到的是正品，样本空间可表示为

$$S=\{00,100,0100,0101,0110,1100,1010,1011,0111,1101,1110,1111\}.$$

(4) 取一直角坐标系，则有 $S=\{(x,y)\,|\,x^2+y^2<1\}$，若取极坐标系，则有

$$S=\{(\rho,\theta)\,|\,0\leqslant\rho<1,0\leqslant\theta<2\pi\}.$$

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

(1) A 发生，B 与 C 不发生.

(2) A 与 B 都发生，而 C 不发生.

(3) A,B,C 中至少有一个发生.

(4) A,B,C 都发生.

(5) A,B,C 都不发生.

(6) A,B,C 中不多于一个发生.

(7) A,B,C 中不多于两个发生.

(8) A,B,C 中至少有两个发生.

解 以下分别用 $D_i(i=1,2,\cdots,8)$ 表示(1),(2),…,(8)中所给出的事件. 注意到一个事件不发生即为它的对立事件发生,例如事件 A 不发生即为 $\bar{A}$ 发生.

(1) A 发生,B 与 C 不发生,表示 $A,\bar{B},\bar{C}$ 同时发生,故 $D_1=A\bar{B}\bar{C}$ 或写成 $D_1=A-B-C$.

(2) A 与 B 都发生而 C 不发生,表示 $A,B,\bar{C}$ 同时发生,故 $D_2=AB\bar{C}$ 或写成 $D_2=AB-C$.

(3) 由和事件的含义知,事件 $A\cup B\cup C$ 即表示 A,B,C 中至少有一个发生,故 $D_3=A\cup B\cup C$.

也可以这样考虑:事件"A,B,C 至少有一个发生"是事件"A,B,C 都不发生"的对立事件,因此,$D_3=\overline{\bar{A}\bar{B}\bar{C}}$.

也可以这样考虑:事件"A,B,C 中至少有一个发生"表示三个事件中恰有一个发生或恰有两个发生或三个事件都发生,因此,D_3 又可写成

$$D_3=A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C\cup AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC\cup ABC.$$

(4) $D_4=ABC$.

(5) $D_5=\bar{A}\bar{B}\bar{C}$.

(6) "A,B,C 中不多于一个发生"表示 A,B,C 都不发生或 A,B,C 中恰有一个发生,因此,$D_6=\bar{A}\,\bar{B}\bar{C}\cup A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\,\bar{B}C$.

又"A,B,C 中不多于一个发生"表示"A,B,C 中至少有两个不发生",亦即 $\bar{A}\,\bar{B},\bar{B}\,\bar{C},\bar{A}\,\bar{C}$ 中至少有一个发生,因此又有 $D_6=\bar{A}\,\bar{B}\cup\bar{B}\,\bar{C}\cup\bar{C}\,\bar{A}$.

又"A,B,C 中不多于一个发生"是事件 $G=$"A,B,C 中至少有两个发生"的对立事件. 而事件 G 可写成 $G=AB\cup BC\cup CA$,因此又可将 D_6 写成

$$D_6=\overline{AB\cup BC\cup CA}=\overline{AB}\cap\overline{BC}\cap\overline{CA}.$$

(7) "A,B,C 中不多于两个发生"表示 A,B,C 都不发生或 A,B,C 中恰有一个发生或 A,B,C 中恰有两个发生. 因此,$D_7=\bar{A}\,\bar{B}\,\bar{C}\cup A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C\cup AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC$. 又"$A,B,C$ 中不多于两个发生"表示 A,B,C 中至少有一个不发生,亦即 $\bar{A},\bar{B},\bar{C}$ 中至少有一个发生,即有 $D_7=\bar{A}\cup\bar{B}\cup\bar{C}$.

又"A,B,C 中不多于两个发生"是事件"A,B,C 三个都发生"的对立事件,因此又有 $D_7=\overline{ABC}$.

(8) $D_8=AB\cup BC\cup CA$,也可写成 $D_8=ABC\cup\bar{A}BC\cup A\bar{B}C\cup AB\bar{C}$.

注:(i) 两事件的差可用对立事件来表示,例如 $A-B=A\bar{B}$,$A-BC=A\,\overline{BC}$.

(ii) 易犯的错误是,误将 $\overline{AB}$ 与 $\bar{A}\ \bar{B}$ 等同起来,事实上,$\overline{AB}=\bar{A}\cup\bar{B}\neq\bar{A}\bar{B}$,又如 $\overline{ABC}=\bar{A}\cup\bar{B}\cup\bar{C}\neq\bar{A}\ \bar{B}\ \bar{C}$.

(iii) 误以为 $S=A\cup B\cup C$，事实上，$S-A\cup B\cup C$ 可能不等于$\varnothing$，一般 $S\supset A\cup B\cup C$.

3. (1) 设 A,B,C 是三个事件，且 $P(A)=P(B)=P(C)=\frac{1}{4}$，$P(AB)=P(BC)=0$，$P(AC)=\frac{1}{8}$，求 A,B,C 至少有一个发生的概率.

(2) 已知 $P(A)=\frac{1}{2}$，$P(B)=\frac{1}{3}$，$P(C)=\frac{1}{5}$，$P(AB)=\frac{1}{10}$，$P(AC)=\frac{1}{15}$，$P(BC)=\frac{1}{20}$，$P(ABC)=\frac{1}{30}$，求 $A\cup B$，$\overline{A}\overline{B}$，$A\cup B\cup C$，$\overline{A}\overline{B}\overline{C}$，$\overline{A}\overline{B}C$，$\overline{A}\overline{B}\cup C$ 的概率.

(3) 已知 $P(A)=\frac{1}{2}$，(i) 若 A,B 互不相容，求 $P(A\overline{B})$，(ii) 若 $P(AB)=\frac{1}{8}$，求$P(A\overline{B})$.

解 (1)
$$\begin{aligned}&P(A\cup B\cup C)\\&=P(A)+P(B)+P(C)-P(AB)-P(BC)-P(AC)+P(ABC)\\&=\frac{5}{8}+P(ABC).\end{aligned}$$
由 $ABC\subset AB$，已知 $P(AB)=0$，故 $0\leqslant P(ABC)\leqslant P(AB)=0$，得$P(ABC)=0$. 所求概率为 $P(A\cup B\cup C)=\frac{5}{8}$.

(2) $P(A\cup B)=P(A)+P(B)-P(AB)=\frac{1}{2}+\frac{1}{3}-\frac{1}{10}=\frac{11}{15}$.

$P(\overline{A}\,\overline{B})=P(\overline{A\cup B})=1-P(A\cup B)=\frac{4}{15}$.

$$\begin{aligned}P(A\cup B\cup C)&=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)\\&=\frac{1}{2}+\frac{1}{3}+\frac{1}{5}-\frac{1}{10}-\frac{1}{15}-\frac{1}{20}+\frac{1}{30}=\frac{51}{60}=\frac{17}{20}.\end{aligned}$$

$P(\overline{A}\,\overline{B}\overline{C})=P(\overline{A\cup B\cup C})=1-P(A\cup B\cup C)=\frac{3}{20}$.

$$\begin{aligned}P(\overline{A}\,\overline{B}C)&=P(\overline{A}\,\overline{B}(S-\overline{C}))=P(\overline{A}\,\overline{B}-\overline{A}\,\overline{B}\overline{C})=P(\overline{A}\,\overline{B})-P(\overline{A}\,\overline{B}\overline{C})\\&=\frac{4}{15}-\frac{3}{20}=\frac{16-9}{60}=\frac{7}{60}.\end{aligned}$$

记 $p=P(\overline{A}\,\overline{B}\cup C)$，由加法公式

$$p=P(\overline{A}\,\overline{B})+P(C)-P(\overline{A}\,\overline{B}C)=\frac{4}{15}+\frac{1}{5}-\frac{7}{60}=\frac{7}{20}.$$

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

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

4. 设 A,B 是两个事件.

(1) 已知 $A\overline{B}=\overline{A}B$，验证 $A=B$.

(2) 验证事件 A 和事件 B 恰有一个发生的概率为 $P(A)+P(B)-2P(AB)$.

证 (1) 假设 $A\overline{B}=\overline{A}B$，故有 $(A\overline{B})\cup(AB)=(\overline{A}B)\cup(AB)$，从而 $A(\overline{B}\cup B)=(\overline{A}\cup A)B$，即 $AS=SB$，故有 $A=B$.

(2) A,B 恰好有一个发生的事件为 $A\overline{B}\cup\overline{A}B$，其概率为

$$\begin{aligned}P(A\overline{B}\cup\overline{A}B)&=P(A\overline{B})+P(\overline{A}B)=P(A(S-B))+P(B(S-A))\\&=P(A-AB)+P(B-AB)=P(A)+P(B)-2P(AB).\end{aligned}$$

5. 10 片药片中有 5 片是安慰剂.

(1) 从中任意抽取 5 片，求其中至少有 2 片是安慰剂的概率.

(2) 从中每次取一片，作不放回抽样，求前 3 次都取到安慰剂的概率.

解 (1) 所求概率为

$$\begin{aligned}p&=1-P(\text{取到的 5 片药片均不是安慰剂})\\&\quad-P(\text{取到的 5 片药片中只有 1 片是安慰剂})\\&=1-\binom{5}{0}\binom{10-5}{5}\Big/\binom{10}{5}-\binom{5}{1}\binom{10-5}{4}\Big/\binom{10}{5}=\frac{113}{126}.\end{aligned}$$

(2) 所求概率为 $p=\frac{5}{10}\times\frac{4}{9}\times\frac{3}{8}=\frac{1}{12}$.

6. 在房间里有 10 个人，分别佩戴从 1 号到 10 号的纪念章，任选 3 人记录其纪念章的号码.

(1) 求最小号码为 5 的概率.

(2) 求最大号码为 5 的概率.

解 E：在房间里任选 3 人，记录其佩戴的纪念章的号码. 10 人中任选 3 人共有 $\binom{10}{3}=120$ 种选法，此即为样本点的总数. 以 A 记事件"最小的号码为 5"，以 B 记事件"最大的号码为 5".

(1) 因选到的最小号码为 5，则其中一个号码为 5 且其余两个号码都大于 5，它们可从 6～10 这 5 个数中选取，故 $N(A)=\binom{5}{2}$，从而

$$P(A)=N(A)/N(S)=\binom{5}{2}\Big/\binom{10}{3}=\frac{1}{12}.$$

(2) 同理，$N(B)=\binom{4}{2}$，故

$$P(B)=N(B)/N(S)=\binom{4}{2}\Big/\binom{10}{3}=\frac{1}{20}.$$

7. 某油漆公司发出 17 桶油漆，其中白漆 10 桶、黑漆 4 桶、红漆 3 桶，在搬运中所有标签脱落，交货人随意将这些油漆发给顾客. 问一个订货为 4 桶白漆、3 桶黑漆和 2 桶红漆的顾客，能按所订颜色如数得到订货的概率是多少？

解　E：在 17 桶油漆中任取 9 桶给顾客. 以 A 表示事件"顾客取到 4 桶白漆、3 桶黑漆与 2 桶红漆"，则有 $N(S)=\binom{17}{9}$，$N(A)=\binom{10}{4}\binom{4}{3}\binom{3}{2}$，故

$$P(A)=N(A)/N(S)=\binom{10}{4}\binom{4}{3}\binom{3}{2}\Big/\binom{17}{9}=\frac{252}{2\ 431}.$$

8. 在 1 500 件产品中有 400 件次品、1 100 件正品. 任取 200 件.

(1) 求恰有 90 件次品的概率.

(2) 求至少有 2 件次品的概率.

解　E：从 1 500 件产品中任取 200 件产品. 以 A 表示事件"恰有 90 件次品"，以 B_i 表示事件"恰有 i 件次品"，$i=0,1$，以 C 表示事件"至少有 2 件次品".

(1)
$$N(S)=\binom{1\ 500}{200},$$

$$N(A)=\binom{400}{90}\binom{1\ 100}{200-90}=\binom{400}{90}\binom{1\ 100}{110},$$

故
$$P(A)=N(A)/N(S)=\binom{400}{90}\binom{1\ 100}{110}\Big/\binom{1\ 500}{200}.$$

(2) $C=S-B_0-B_1$，其中，B_0，B_1 互不相容，所以

$$\begin{aligned}P(C)&=P(S-B_0-B_1)=P(S-(B_0\cup B_1))\\&=1-P(B_0\cup B_1)=1-P(B_0)-P(B_1).\end{aligned}$$

因

$$N(B_0)=\binom{1\ 100}{200},\quad N(B_1)=\binom{400}{1}\binom{1\ 100}{199},$$

故

$$P(B_0)=\binom{1\ 100}{200}\Big/\binom{1\ 500}{200},\quad P(B_1)=\binom{400}{1}\binom{1\ 100}{199}\Big/\binom{1\ 500}{200},$$

因此有

$$P(C)=1-\binom{1\,100}{200}\Big/\binom{1\,500}{200}-\binom{400}{1}\binom{1\,100}{199}\Big/\binom{1\,500}{200}$$
$$=1-\left[\binom{1\,100}{200}+\binom{400}{1}\binom{1\,100}{199}\right]\Big/\binom{1\,500}{200}.$$

9. 从5双不同的鞋子中任取4只，问这4只鞋子中至少有两只配成一双的概率是多少？

解　E：从5双不同的鞋子中任取4只．以A表示事件“所取4只鞋子中至少有两只配成一双鞋子”，则$\overline{A}$表示事件“所取4只鞋子中无配对的”．先计算$P(\overline{A})$较为简便．以下按$N(\overline{A})$的不同求法，列出本题的3种解法，另外还给出一种直接求$P(A)$的解法．

解法(i)　考虑4只鞋子是有次序一只一只取出的．自5双(10只)鞋子中任取4只共有$10\times9\times8\times7$种取法，$N(S)=10\times9\times8\times7$．现在来求$N(\overline{A})$．第一只可以任意取，共有10种取法，第二只只能在剩下的9只中且除去与已取的第一只配对的8只鞋子中任取一只，共8种取法．同理第三只、第四只各有6种、4种取法，从而$N(\overline{A})=10\times8\times6\times4$．故

$$P(A)=1-P(\overline{A})=1-N(\overline{A})/N(S)$$
$$=1-\frac{10\times8\times6\times4}{10\times9\times8\times7}=\frac{13}{21}.$$

解法(ii)　从10只鞋子中任取4只，共有$\binom{10}{4}$种取法，即$N(S)=\binom{10}{4}$．为求$N(\overline{A})$，先从5双鞋子中任取4双共有$\binom{5}{4}$种取法，再自取出的每双鞋子中各取1只(在一双中取一只共有2种取法)，共有2^4种取法，即$N(\overline{A})=\binom{5}{4}2^4$．故

$$P(A)=1-P(\overline{A})=1-\frac{\binom{5}{4}2^4}{\binom{10}{4}}=\frac{13}{21}.$$

解法(iii)　现在来求$N(\overline{A})$．先从5只左脚鞋子中任取k只$(k=0,1,2,3,4)$，有$\binom{5}{k}$种取法，而剩下的$4-k$只鞋子只能从(不能与上述所取的配对的)$5-k$只右脚鞋子中选取，即对于每个固定的k，有$\binom{5}{k}\binom{5-k}{4-k}$种取法．故

$$N(\overline{A})=\sum_{k=0}^{4}\binom{5}{k}\binom{5-k}{4-k}=80.$$

故

$$P(A)=1-P(\overline{A})=1-N(\overline{A})/N(S)=1-\frac{80}{\binom{10}{4}}=\frac{13}{21}.$$

解法(iv) 以 A_i 表示事件"所取 4 只鞋子中恰能配成 i 双"($i=1,2$)，则 $A=A_1\cup A_2, A_1A_2=\varnothing$，故 $P(A)=P(A_1)+P(A_2)$. 因 A_2 为 4 只恰能配成 2 双，它可直接从 5 双鞋子中成双地取得，故 $N(A_2)=\binom{5}{2}$. $N(A_1)$的算法是：先从 5 双中取 1 双，共有$\binom{5}{1}$种取法，另外两只能从其他 8 只中取，共有$\binom{8}{2}$种取法，不过这种取法中将成双的也算在内了，应去掉. 从而

$$N(A_1)=\binom{5}{1}\left[\binom{8}{2}-\binom{4}{1}\right]=120.$$

$N(S)$仍为解法(ii)中的$\binom{10}{4}=210$ 种，故

$$P(A)=P(A_1)+P(A_2)=\frac{N(A_1)}{N(S)}+\frac{N(A_2)}{N(S)}$$
$$=\frac{120+10}{210}=\frac{13}{21}.$$

10. 在 11 张卡片上分别写上 probability 这 11 个字母，从中任意连抽 7 张，求其排列结果为 ability 的概率.

解法(i) E：自 11 个字母中随机地接连抽 7 个字母并依次排列. 将 11 个字母中的两个 b 看成是可分辨的，两个 i 也看成是可分辨的，$N(S)=A_{11}^{7}$. 以 A 记事件"排列结果为 ability"，则 $N(A)=4$(因 b 有两种取法，i 也有两种取法)，因而

$$P(A)=N(A)/N(S)=\frac{4}{A_{11}^{7}}=2.4\times10^{-6}.$$

解法(ii) 本题也可利用乘法定理来计算. 以 $A_1,B_2,I_3,L_4,I_5,T_6,Y_7$ 依次表示取得字母 a,b,i,l,i,t,y 的各事件，则所求概率为

$$\begin{aligned}P(A_1B_2I_3L_4I_5T_6Y_7)=&P(A_1)P(B_2\mid A_1)P(I_3\mid A_1B_2)\\&\times P(L_4\mid A_1B_2I_3)P(I_5\mid A_1B_2I_3L_4)\\&\times P(T_6\mid A_1B_2I_3L_4I_5)P(Y_7\mid A_1B_2I_3L_4I_5T_6)\end{aligned}$$

$$=\frac{1}{11}\times\frac{2}{10}\times\frac{2}{9}\times\frac{1}{8}\times\frac{1}{7}\times\frac{1}{6}\times\frac{1}{5}=\frac{4}{A_{11}^{7}}.$$

注：在解法(i)中仅当将两个 i 看成是可以区分的，两个 b 看成是可以区分的，才属于古典概型问题.

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

解 E：将 3 只球随机地放入 4 个杯子中去，易知共有 4^3 种放置法. 以 A_i 表示事件“杯子中球的最大个数为 i”，$i=1,2,3$.

A_3 只有当 3 只球放在同一杯子中时才能发生，有 4 个杯子可以任意选择，于是 $N(A_3)=\binom{4}{1}$，故

$$P(A_3)=N(A_3)/N(S)=\binom{4}{1}\Big/4^3=\frac{1}{16}.$$

A_1 只有当每个杯子最多放一只球时才能发生. 因而 $N(A_1)=4\times3\times2=A_4^3$，故

$$P(A_1)=N(A_1)/N(S)=A_4^3/4^3=\frac{3}{8}.$$

又 $A_1\cup A_2\cup A_3=S$，且 $A_iA_j=\varnothing\ (i\neq j)$，故 $P(A_1)+P(A_2)+P(A_3)=1$，从而

$$P(A_2)=1-\frac{1}{16}-\frac{3}{8}=\frac{9}{16}.$$

12. 50 只铆钉随机地取来用在 10 个部件上，其中有 3 只铆钉强度太弱. 每个部件用 3 只铆钉. 若将 3 只强度太弱的铆钉都装在一个部件上，则这个部件强度就太弱. 问发生一个部件强度太弱的概率是多少？

解 将部件自 1 到 10 编号. E：随机地取铆钉，使各部件都装 3 只铆钉. 以 A_i 表示事件“第 i 号部件强度太弱”. 由题设，仅当 3 只强度太弱的铆钉同时装在第 i 号部件上，A_i 才能发生. 由于从 50 只铆钉中任取 3 只装在第 i 号部件上共有 $\binom{50}{3}$ 种取法，强度太弱的铆钉仅有 3 只，它们都装在第 i 号部件上，只有 $\binom{3}{3}=1$ 种取法，故

$$P(A_i)=1\Big/\binom{50}{3}=\frac{1}{19\ 600},\quad i=1,2,\cdots,10,$$

且知 $A_1,A_2,\cdots,A_{10}$ 两两互不相容，因此，10 个部件中有一个强度太弱的概率为

$$p=P(A_1\cup A_2\cup\cdots\cup A_{10})=P(A_1)+P(A_2)+\cdots+P(A_{10})$$
$$=\frac{10}{19\ 600}=\frac{1}{1\ 960}.$$

13. 一俱乐部有 5 名一年级学生，2 名二年级学生，3 名三年级学生，2 名四年级学生.

(1) 在其中任选 4 名学生，求一、二、三、四年级的学生各一名的概率.

(2) 在其中任选 5 名学生，求一、二、三、四年级的学生均包含在内的概率.

解　(1) 共有 $5+2+3+2=12$ 名学生，在其中任选 4 名共有 $\binom{12}{4}=495$ 种选法，其中每年级各选 1 名的选法有 $\binom{5}{1}\binom{2}{1}\binom{3}{1}\binom{2}{1}=60$ 种选法，因此，所求概率为 $p=\frac{60}{495}=\frac{4}{33}$.

(2) 在 12 名学生中任选 5 名的选法共有 $\binom{12}{5}=792$ 种. 在每个年级中有一个年级取 2 名，而其他 3 个年级各取 1 名的取法共有

$$\binom{5}{2}\binom{2}{1}\binom{3}{1}\binom{2}{1}+\binom{5}{1}\binom{2}{2}\binom{3}{1}\binom{2}{1}+\binom{5}{1}\binom{2}{1}\binom{3}{2}\binom{2}{1}+\binom{5}{1}\binom{2}{1}\binom{3}{1}\binom{2}{2}=240(\text{种}).$$

于是所求的概率为

$$p=\frac{240}{792}=\frac{10}{33}.$$

14. (1) 已知 $P(\bar{A})=0.3$，$P(B)=0.4$，$P(A\bar{B})=0.5$，求条件概率 $P(B|A\cup\bar{B})$.

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

解　(1) $P(B|A\cup\bar{B})=\frac{P(B(A\cup\bar{B}))}{P(A\cup\bar{B})}=\frac{P(AB)}{P(A)+P(\bar{B})-P(A\bar{B})}$.

由题设得 $P(A)=1-P(\bar{A})=0.7$，$P(\bar{B})=1-P(B)=0.6$，$P(AB)=P(A(S-\bar{B}))=P(A)-P(A\bar{B})=0.2$，故

$$P(B|A\cup\bar{B})=\frac{0.2}{0.7+0.6-0.5}=0.25.$$

(2)
$$P(AB)=P(B|A)P(A)=\frac{1}{12},$$

$$P(B)=P(AB)/P(A|B)=\frac{1}{12}\Big/\frac{1}{2}=\frac{1}{6},$$

故

$$P(A\cup B)=P(A)+P(B)-P(AB)$$
$$=\frac{1}{4}+\frac{1}{6}-\frac{1}{12}=\frac{1}{3}.$$

15. 掷两颗骰子，已知两颗骰子点数之和为 7，求其中有一颗为 1 点的概率（用两种方法）.

解 E：掷两颗骰子，观察其出现之点数. 以 A 记事件"两颗骰子点数之和为 7"，以 B 记事件"两颗骰子中有一颗出现 1 点".

解法(i) 按条件概率的定义式：$P(B|A)=\dfrac{P(AB)}{P(A)}$来求条件概率. 设想两颗骰子是可分辨的，样本空间为

$$S=\{(1,1),(1,2),\cdots,(1,6),(2,1),(2,2),\cdots,(2,6),\cdots,(6,6)\},$$
$$A=\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\},$$

$AB=\{(1,6),(6,1)\}$. 现在 $N(S)=36,N(A)=6,N(AB)=2$，因此

$$P(B|A)=\frac{2/36}{6/36}=\frac{1}{3}.$$

解法(ii) 按条件概率的含义来求 $P(B|A)$. 样本空间原有 36 个样本点，现在知道了"A 已经发生"这一信息，根据这一信息，不在 A 中的样本点就不可能出现了，因而试验所有可能结果所成的集合就是 A，而 A 中共有 6 个可能结果，其中只有两个结果(1,6)和(6,1)有一颗骰子出现 1 点，因此

$$P(B|A)=\frac{2}{6}=\frac{1}{3}.$$

16. 据以往资料表明，某一 3 口之家，患某种传染病的概率有以下规律：

$$P\{孩子得病\}=0.6,\quad P\{母亲得病|孩子得病\}=0.5,$$
$$P\{父亲得病|母亲及孩子得病\}=0.4,$$

求母亲及孩子得病但父亲未得病的概率.

解 以 A 记事件"孩子得病"，以 B 记事件"母亲得病"，以 C 记事件"父亲得病"，按题意需要求 $P(AB\overline{C})$. 已知 $P(A)=0.6,P(B|A)=0.5,P(C|BA)=0.4$，由乘法定理得

$$P(AB\overline{C})=P(\overline{C}BA)=P(\overline{C}|BA)P(BA)$$
$$=P(\overline{C}|BA)P(B|A)P(A)$$
$$=[1-P(C|BA)]P(B|A)P(A)$$
$$=0.6\times0.5\times0.6=0.18.$$

17. 已知在 10 件产品中有 2 件次品，在其中取两次，每次任取一件，作不放

回抽样.求下列事件的概率：

(1) 两件都是正品.

(2) 两件都是次品.

(3) 一件是正品，一件是次品.

(4) 第二次取出的是次品.

解 E：在 10 件产品中(其中有 2 件次品)任取两次，每次取 1 件，作不放回抽样.以 $A_i(i=1,2)$ 表示事件"第 i 次抽出的是正品".因为是不放回抽样，所以

(1) $P(A_1A_2)=P(A_2|A_1)P(A_1)=\frac{7}{9}\times\frac{8}{10}=\frac{28}{45}.$

(2) $P(\overline{A}_1\overline{A}_2)=P(\overline{A}_2|\overline{A}_1)P(\overline{A}_1)=\frac{1}{9}\times\frac{2}{10}=\frac{1}{45}.$

(3)
$$\begin{aligned}&P(A_1\overline{A}_2\cup\overline{A}_1A_2)\\&=P(A_1\overline{A}_2)+P(\overline{A}_1A_2)\ (\text{因}(A_1\overline{A}_2)(\overline{A}_1A_2)=\varnothing)\\&=P(\overline{A}_2|A_1)P(A_1)+P(A_2|\overline{A}_1)P(\overline{A}_1)\\&=\frac{2}{9}\times\frac{8}{10}+\frac{8}{9}\times\frac{2}{10}=\frac{16}{45}.\end{aligned}$$

亦可利用(1)，(2)的结果.因为 $A_1A_2\cup\overline{A}_1\overline{A}_2\cup A_1\overline{A}_2\cup\overline{A}_1A_2=S$，且 A_1A_2，$\overline{A}_1\overline{A}_2$，$A_1\overline{A}_2$，$\overline{A}_1A_2$ 两两互不相容，故

$$P(A_1\overline{A}_2\cup\overline{A}_1A_2)=1-\frac{28}{45}-\frac{1}{45}=\frac{16}{45}.$$

(4)
$$\begin{aligned}P(\overline{A}_2)&=P((A_1\cup\overline{A}_1)\overline{A}_2)=P(A_1\overline{A}_2\cup\overline{A}_1\overline{A}_2)\\&=P(A_1\overline{A}_2)+P(\overline{A}_1\overline{A}_2)\\&=P(\overline{A}_2|A_1)P(A_1)+P(\overline{A}_2|\overline{A}_1)P(\overline{A}_1)\\&=\frac{2}{9}\times\frac{8}{10}+\frac{1}{9}\times\frac{2}{10}=\frac{1}{5}.\end{aligned}$$

18. 某人忘记了电话号码的最后一个数字，因而他随意地拨号.求他拨号不超过三次而接通所需电话的概率.若已知最后一个数字是奇数，则此概率是多少?

解法(i) 以 A_i 表示事件"第 i 次拨号接通电话"，$i=1,2,3$.以 A 表示事件"拨号不超过三次接通电话"，则有

$$A=A_1\cup\overline{A}_1A_2\cup\overline{A}_1\overline{A}_2A_3.$$

因 A_1，$\overline{A}_1A_2$，$\overline{A}_1\overline{A}_2A_3$ 两两互不相容，且 $P(A_1)=\frac{1}{10}$，

$$P(\overline{A}_1A_2)=P(A_2|\overline{A}_1)P(\overline{A}_1)=\frac{1}{9}\times\frac{9}{10}=\frac{1}{10},$$

$$P(\overline{A}_1\overline{A}_2A_3)=P(A_3\,|\,\overline{A}_1\overline{A}_2)P(\overline{A}_2\,|\,\overline{A}_1)P(\overline{A}_1)$$
$$=\frac{1}{8}\times\frac{8}{9}\times\frac{9}{10}=\frac{1}{10},$$

即有

$$P(A)=P(A_1)+P(\overline{A}_1A_2)+P(\overline{A}_1\overline{A}_2A_3)$$
$$=\frac{1}{10}+\frac{1}{10}+\frac{1}{10}=\frac{3}{10}.$$

当已知最后一位数是奇数时，所求概率为

$$p=\frac{1}{5}+\frac{1}{5}+\frac{1}{5}=\frac{3}{5}.$$

解法(ii) 沿用解法(i)的记号，知

$$P(A)=1-P\{拨号三次都接不通\}=1-P(\overline{A}_1\overline{A}_2\overline{A}_3)$$
$$=1-P(\overline{A}_3\,|\,\overline{A}_1\overline{A}_2)P(\overline{A}_2\,|\,\overline{A}_1)P(\overline{A}_1)$$
$$=1-\frac{7}{8}\times\frac{8}{9}\times\frac{9}{10}=\frac{3}{10}.$$

当已知最后一位是奇数时，所求概率为

$$p=1-\frac{2}{3}\times\frac{3}{4}\times\frac{4}{5}=\frac{3}{5}.$$

19. (1) 设甲袋中装有 n 只白球、m 只红球，乙袋中装有 N 只白球、M 只红球. 今从甲袋中任意取一只球放入乙袋中，再从乙袋中任意取一只球. 问取到白球的概率是多少?

(2) 第一只盒子装有 5 只红球、4 只白球，第二只盒子装有 4 只红球、5 只白球. 先从第一盒中任取 2 只球放入第二盒中去，然后从第二盒中任取一只球. 求取到白球的概率.

解法(i) (1) E:从甲袋任取一球放入乙袋(此为试验 E_1)，再从乙袋任取一球观察其颜色(此为试验 E_2). 试验 E 是由 E_1 和 E_2 合成的. 以 R 表示事件"从甲袋取得的是红球"，以 W 表示事件"从乙袋取得的是白球"，即有

$$W=SW=(R\cup\overline{R})W=RW\cup\overline{R}W,\ (RW)(\overline{R}W)=\varnothing.$$

于是
$$P(W)=P(RW)+P(\overline{R}W)$$
$$=P(W\,|\,R)P(R)+P(W\,|\,\overline{R})P(\overline{R}).$$

而
$$P(R)=\frac{m}{n+m},\quad P(\overline{R})=\frac{n}{n+m}.$$

在计算 $P(W\,|\,R)$，$P(W\,|\,\overline{R})$时，注意在试验 E_2 中，乙袋球数为 $N+M+1$ 只;在求 $P(W\,|\,R)$时，乙袋白球数为 N，但在求 $P(W\,|\,\overline{R})$时，乙袋白球数为 $N+1$，故

$$P(W)=\frac{N}{N+M+1}\frac{m}{n+m}+\frac{N+1}{N+M+1}\frac{n}{n+m}$$
$$=\frac{n+N(n+m)}{(n+m)(N+M+1)}.$$

(2) E:从第一盒中任取 2 只球放入第二盒(E_1)，再从第二盒任取一球观察其颜色(E_2). 以 $R_i(i=0,1,2)$表示事件“从第一盒中取得的球中有 i 只是红球”，以 W 表示事件“从第二盒取得一球是白球”. 由于 R_0,R_1,R_2 两两互不相容，且 $R_0\cup R_1\cup R_2=S$，故

$$W=SW=(R_0\cup R_1\cup R_2)W=R_0W\cup R_1W\cup R_2W.$$

从而

$$P(W)=P(R_0W)+P(R_1W)+P(R_2W)$$
$$=P(W|R_0)P(R_0)+P(W|R_1)P(R_1)+P(W|R_2)P(R_2).$$

而

$$P(R_0)=\binom{4}{2}\bigg/\binom{9}{2}=\frac{1}{6},\quad P(R_2)=\binom{5}{2}\bigg/\binom{9}{2}=\frac{5}{18},$$
$$P(R_1)=1-P(R_0)-P(R_2)=1-\frac{1}{6}-\frac{5}{18}=\frac{10}{18}.$$

并注意到，在试验 E_2 中第二盒球的个数为 11，故

$$P(W|R_0)=\frac{7}{11},\quad P(W|R_1)=\frac{6}{11},\quad P(W|R_2)=\frac{5}{11},$$

所以

$$P(W)=\frac{7}{11}\times\frac{1}{6}+\frac{6}{11}\times\frac{10}{18}+\frac{5}{11}\times\frac{5}{18}=\frac{53}{99}.$$

解法(ii) (1) 以 A 表示事件“最后取到的是白球”，以 B 表示事件“最后取到的是甲袋中的球”，因

$$A=SA=(B\cup\bar{B})A=BA\cup\bar{B}A,\quad (BA)(\bar{B}A)=\varnothing,$$

于是

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

而

$$P(B)=\frac{1}{N+M+1},\quad P(\bar{B})=\frac{N+M}{N+M+1}$$

(这是因为最后是从乙袋中取球的，此时乙袋中共有 $N+M+1$ 只球，其中只有一只是甲袋中的球).

又有
$$P(A|B)=\frac{n}{n+m},\quad P(A|\bar{B})=\frac{N}{M+N},$$
故
$$\begin{aligned}P(A)&=\frac{n}{n+m}\,\frac{1}{N+M+1}+\frac{N}{N+M}\,\frac{N+M}{N+M+1}\\&=\frac{n+N(n+m)}{(n+m)(N+M+1)}.\end{aligned}$$

(2) 以 A 表示事件“最后取到的是白球”，以 B 表示事件“最后取到的是第一盒中的球”，因
$$A=SA=(B\cup\bar{B})A=BA\cup\bar{B}A,\quad (BA)(\bar{B}A)=\varnothing$$
得
$$\begin{aligned}P(A)&=P(A|B)P(B)+P(A|\bar{B})P(\bar{B})\\&=\frac{4}{9}\times\frac{2}{11}+\frac{5}{9}\times\frac{9}{11}=\frac{53}{99}.\end{aligned}$$

20. 某种产品的商标为“MAXAM”，其中有 2 个字母脱落，有人捡起随意放回，求放回后仍为“MAXAM”的概率.

解 以 H_1,H_2,H_3,H_4,H_5 依次表示事件“脱落 M，M”“脱落 A，A”“脱落 M，A”“脱落 X，A”“脱落 X，M”，以 G 表示事件“放回后仍为 MAXAM”，所需求的是 $P(G)$. 可知 H_1,H_2,H_3,H_4,H_5 两两互不相容，且 $H_1\cup H_2\cup H_3\cup H_4\cup H_5=S$. 已知
$$P(H_1)=\binom{2}{2}\Big/\binom{5}{2}=\frac{1}{10},\quad P(H_2)=\binom{2}{2}\Big/\binom{5}{2}=\frac{1}{10},$$
$$P(H_3)=\binom{2}{1}\binom{2}{1}\Big/\binom{5}{2}=\frac{4}{10},\quad P(H_4)=\binom{1}{1}\binom{2}{1}\Big/\binom{5}{2}=\frac{2}{10},$$
$$P(H_5)=\binom{1}{1}\binom{2}{1}\Big/\binom{5}{2}=\frac{2}{10},$$
而
$$P(G|H_1)=P(G|H_2)=1,$$
$$P(G|H_3)=P(G|H_4)=P(G|H_5)=\frac{1}{2}.$$
由全概率公式得
$$P(G)=\sum_{i=1}^{5}P(G|H_i)P(H_i)=\frac{1}{10}+\frac{1}{10}+\frac{2}{10}+\frac{1}{10}+\frac{1}{10}=\frac{3}{5}.$$

21. 已知男子有 5%是色盲患者，女子有 0.25%是色盲患者. 今从男女人数相等的人群中随机地挑选一人，恰好是色盲患者，问此人是男性的概率是多少？

解　以 A 表示事件"选出的是男性"，则 $\bar{A}$ 表示事件"选出的是女性"，以 H 表示事件"选出的人患色盲"，则 $\bar{H}$ 表示"选出的人不患色盲". 由题设 $P(A)=P(\bar{A})=\frac{1}{2}$，$P(H|A)=0.05$，$P(H|\bar{A})=0.0025$，所需求的概率是 $P(A|H)$. 由贝叶斯公式得

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

$$=\frac{0.05\times\frac{1}{2}}{0.05\times\frac{1}{2}+0.0025\times\frac{1}{2}}=\frac{500}{525}=\frac{20}{21}.$$

22. 一学生接连参加同一课程的两次考试. 第一次及格的概率为 p，若第一次及格则第二次及格的概率也为 p；若第一次不及格则第二次及格的概率为 $\frac{p}{2}$.

(1) 若至少有一次及格则他能取得某种资格，求他取得该资格的概率.

(2) 若已知他第二次已经及格，求他第一次及格的概率.

解　E：一学生接连参加一门课程的两次考试. 以 A_i 表示事件"第 i 次考试及格"，$i=1,2$；以 A 表示"他能取得该种资格".

(1) 按题意 $A=A_1\cup\bar{A}_1A_2$. 因 $A_1\cap(\bar{A}_1A_2)=\varnothing$，且由已知条件：

$$P(A_1)=p,\quad P(\bar{A}_1)=1-p,$$

$$P(A_2|A_1)=p,\quad P(A_2|\bar{A}_1)=\frac{p}{2},$$

故

$$P(A)=P(A_1\cup\bar{A}_1A_2)=P(A_1)+P(\bar{A}_1A_2)$$

$$=p+P(A_2|\bar{A}_1)P(\bar{A}_1)=p+\frac{p}{2}(1-p)$$

$$=\frac{3}{2}p-\frac{1}{2}p^2.$$

(2) $P(A_1|A_2)=\frac{P(A_1A_2)}{P(A_2)}$

$$=\frac{P(A_2|A_1)P(A_1)}{P(A_2|A_1)P(A_1)+P(A_2|\bar{A}_1)P(\bar{A}_1)}$$

$$=\frac{p\cdot p}{p\cdot p+\frac{p}{2}(1-p)}=\frac{2p}{p+1}.$$

23. 将两信息分别编码为 A 和 B 传送出去，接收站收到时，A 被误收作 B 的概率为 0.02，而 B 被误收作 A 的概率为 0.01. 信息 A 与信息 B 传送的频繁

程度为 2∶1.若接收站收到的信息是 A,问原发信息是 A 的概率是多少?

解 以 D 表示事件"将信息 A 传递出去",则 $\bar{D}$ 表示事件"将信息 B 传递出去",以 R 表示事件"接收到信息 A",则 $\bar{R}$ 表示事件"接收到信息 B",按题意需求概率 $P(D|R)$. 已知 $P(\bar{R}|D)=0.02$,$P(R|\bar{D})=0.01$,且有 $P(D)/P(\bar{D})=2$,由于 $P(D)+P(\bar{D})=1$,得知 $P(D)=\frac{2}{3}$,$P(\bar{D})=\frac{1}{3}$. 由贝叶斯公式得到

$$P(D|R)=\frac{P(DR)}{P(R)}=\frac{P(R|D)P(D)}{P(R|D)P(D)+P(R|\bar{D})P(\bar{D})}$$

$$=\frac{(1-0.02)\times\frac{2}{3}}{(1-0.02)\times\frac{2}{3}+0.01\times\frac{1}{3}}=\frac{196}{197}.$$

24. 有两箱同种类的零件,第一箱装 50 只,其中 10 只一等品;第二箱装 30 只,其中 18 只一等品.今从两箱中任挑出一箱,然后从该箱中取零件两次,每次任取一只,作不放回抽样.求

(1) 第一次取到的零件是一等品的概率.

(2) 在第一次取到的零件是一等品的条件下,第二次取到的也是一等品的概率.

解 以 H 表示事件"从第一箱中取零件",则 $\bar{H}$ 表示事件"从第二箱中取零件".由已知条件 $P(H)=P(\bar{H})=\frac{1}{2}$. 又以 A_i 表示事件"第 i 次从箱中(不放回抽样)取得的是一等品",$i=1,2$.

(1) 由条件 $P(A_1|H)=\frac{1}{5}$,$P(A_1|\bar{H})=\frac{3}{5}$,故

$$P(A_1)=P(A_1|H)P(H)+P(A_1|\bar{H})P(\bar{H})$$

$$=\frac{1}{10}+\frac{3}{10}=\frac{2}{5}.$$

(2) 需要求的是 $P(A_2|A_1)$. 因 $P(A_2|A_1)=\frac{P(A_1A_2)}{P(A_1)}$,而

$$P(A_1A_2)=P(A_1A_2|H)P(H)+P(A_1A_2|\bar{H})P(\bar{H}),$$

由条件概率的含义,$P(A_1A_2|H)$表示在第一箱中取两次,每次取一只零件,作不放回抽样,且两次都取得一等品的概率.因第一箱共有 50 只零件,其中有 10 只一等品,故有 $P(A_1A_2|H)=\frac{10}{50}\times\frac{9}{49}$. 同理,$P(A_1A_2|\bar{H})=\frac{18}{30}\times\frac{17}{29}$,故有

$$P(A_2|A_1)=\frac{P(A_1A_2)}{P(A_1)}$$

$$=\frac{1}{P(A_1)}[P(A_1A_2\,|\,H)P(H)+P(A_1A_2\,|\,\overline{H})P(\overline{H})]$$

$$=\frac{5}{2}\left(\frac{10}{50}\times\frac{9}{49}\times\frac{1}{2}+\frac{18}{30}\times\frac{17}{29}\times\frac{1}{2}\right)$$

$$=\frac{1}{4}\left(\frac{9}{49}+\frac{51}{29}\right)=0.485\ 6.$$

25. 某人下午 5：00 下班，他所积累的资料表明：

到家时间	5：35～5：39	5：40～5：44	5：45～5：49	5：50～5：54	迟于 5：54
乘地铁的概率	0.10	0.25	0.45	0.15	0.05
乘汽车的概率	0.30	0.35	0.20	0.10	0.05

某日他抛一枚硬币决定乘地铁还是乘汽车，结果他是 5：47 到家的．试求他乘地铁回家的概率．

解　以 H 表示事件“乘地铁回家”，则 $\overline{H}$ 表示事件“乘汽车回家”．因到家时间为 5：47，它属于区间 5：45～5：49，以 T 记“到家时间在 5：45～5：49 之间”，则需要求的是概率 $P(H|T)$．已知 $P(T|H)=0.45$，$P(T|\overline{H})=0.20$，又因他是由掷硬币决定乘地铁还是乘汽车，因此，$P(H)=P(\overline{H})=0.5$．由贝叶斯公式得

$$P(H|T)=\frac{P(HT)}{P(T)}=\frac{P(T|H)P(H)}{P(T|H)P(H)+P(T|\overline{H})P(\overline{H})}$$

$$=\frac{0.45\times0.5}{0.45\times0.5+0.20\times0.5}=\frac{9}{13}.$$

26. 病树的主人外出，委托邻居浇水，设已知如果不浇水，树死去的概率为 0.8．若浇水则树死去的概率为 0.15．有 0.9 的把握确定邻居会记得浇水．

(1) 求主人回来时树还活着的概率．

(2) 若主人回来时树已死去，求邻居忘记浇水的概率．

解　(1) 记 A 为事件“树还活着”，记 W 为事件“邻居记得给树浇水”，即有

$$P(W)=0.9,\quad P(\overline{W})=0.1,\quad P(A|W)=0.85,\quad P(A|\overline{W})=0.2,$$

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

$$=0.85\times0.9+0.2\times0.1=0.785.$$

(2) $$P(\overline{W}|\overline{A})=\frac{P(\overline{A}|\overline{W})P(\overline{W})}{P(\overline{A})}=\frac{[1-P(A|\overline{W})]P(\overline{W})}{1-P(A)}$$

$$=\frac{0.8\times0.1}{0.215}=0.372.$$

27. 设本题涉及的事件均有意义. 设 A,B 都是事件.

(1) 已知 $P(A)>0$，证明 $P(AB|A)\geqslant P(AB|A\cup B)$.

(2) 若 $P(A|B)=1$，证明 $P(\bar{B}|\bar{A})=1$.

(3) 若设 C 也是事件，且有 $P(A|C)\geqslant P(B|C)$，$P(A|\bar{C})\geqslant P(B|\bar{C})$，证明 $P(A)\geqslant P(B)$.

证 (1) 若 $P(A)>0$，要证 $P(AB|A)\geqslant P(AB|A\cup B)$. 上式左边等于 $P(AB)/P(A)$，上式右边等于 $P(AB)/P(A\cup B)$. 因为 $A\cup B\supset A$，$P(A\cup B)\geqslant P(A)$，故有右$\leqslant$左.

(2) 由 $P(A|B)=1$ 得 $\dfrac{P(AB)}{P(B)}=1$，即

$$P(AB)=P(B). \qquad (*_1)$$

于是 $$P(\bar{B}|\bar{A})=P(\bar{A}\,\bar{B})/P(\bar{A})=P(\overline{A\cup B})/P(\bar{A})=\frac{1-P(A\cup B)}{1-P(A)}$$

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

由($*_1$)式得到

$$P(\bar{B}|\bar{A})=\frac{1-P(A)}{1-P(A)}=1.$$

(3) 由假设 $P(A|C)\geqslant P(B|C)$，即

$$P(AC)\geqslant P(BC). \qquad (*_2)$$

同样由 $P(A|\bar{C})\geqslant P(B|\bar{C})$ 就有

$$P(A\bar{C})\geqslant P(B\bar{C}). \qquad (*_3)$$

由($*_3$)式

$$P(A(S-C))\geqslant P(B(S-C)),$$

得 $$P(A)-P(AC)\geqslant P(B)-P(BC),$$

或 $$P(A)-P(B)\geqslant P(AC)-P(BC),$$

由($*_2$)式，得知

$$P(A)-P(B)\geqslant 0,\text{即 } P(A)\geqslant P(B).$$

28. 有两种花籽，发芽率分别为 0.8，0.9，从中各取一颗，设各花籽是否发芽相互独立. 求

(1) 这两颗花籽都能发芽的概率.

(2) 至少有一颗能发芽的概率.

(3) 恰有一颗能发芽的概率.

解 以 A,B 分别表示事件第一颗、第二颗花籽能发芽，即有 $P(A)=0.8$，$P(B)=0.9$.

(1) 由 A,B 相互独立，得两颗花籽都能发芽的概率为

$$P(AB)=P(A)P(B)=0.8\times0.9=0.72.$$

(2) 至少有一颗花籽能发芽的概率，即事件 $A\cup B$ 的概率为

$$P(A\cup B)=P(A)+P(B)-P(AB)=0.8+0.9-0.72=0.98.$$

(3) 恰有一颗花籽能发芽的概率，即为事件 $A\bar{B}\cup B\bar{A}$ 的概率，由第 4 题(2)得

$$P(A\bar{B}\cup\bar{A}B)=P(A)+P(B)-2P(AB)=0.8+0.9-2P(AB)$$
$$=0.8+0.9-2P(A)P(B)=1.7-2\times0.72=0.26.$$

29. 根据报道美国人血型的分布近似地为：A 型为 37%，O 型为 44%，B 型为 13%，AB 型为 6%. 夫妻拥有的血型是相互独立的.

(1) B 型的人只有输入 B，O 两种血型才安全. 若妻为 B 型，夫为何种血型未知，求夫是妻的安全输血者的概率.

(2) 随机地取一对夫妇，求妻为 B 型，夫为 A 型的概率.

(3) 随机地取一对夫妇，求其中一人为 A 型，另一人为 B 型的概率.

(4) 随机地取一对夫妇，求其中至少有一人是 O 型的概率.

解 (1) 由题意知夫血型应为 B，O 才为安全输血者. 因两种血型互不相容，故所求概率为

$$p_1=0.44+0.13=0.57.$$

(2) 因夫妻拥有的血型相互独立，于是所求概率为

$$p_2=0.13\times0.37=0.048\ 1.$$

(3) $p_3=2\times0.37\times0.13=0.096\ 2.$

(4) 有三种可能，即夫为 O，妻为非 O；妻为 O，夫为非 O；夫妻均为 O.

$$p_4=2\times0.44\times(1-0.44)+0.44\times0.44=0.686\ 4.$$

30. (1) 给出事件 A,B 的例子，使得

(i) $P(A|B)<P(A)$，(ii) $P(A|B)=P(A)$，(iii) $P(A|B)>P(A)$.

(2) 设事件 A,B,C 相互独立，证明(i)C 与 AB 相互独立，(ii)C 与 $A\cup B$ 相互独立.

(3) 设事件 A 的概率 $P(A)=0$，证明对于任意另一事件 B，有 A,B 相互独立.

(4) 证明事件 A,B 相互独立的充要条件是 $P(A|B)=P(A|\bar{B})$.

解 (1) 举例：

(i) 设试验为将骰子掷一次，事件 A 为“出现偶数点”，B 为“出现奇数点”，则 $P(A|B)=0$，$P(A)=\frac{1}{2}$，故 $P(A|B)<P(A)$.

(ii) 设试验为将骰子掷一次，A 同上，B 为“掷出点数$\geqslant1$”，则 $P(A|B)=$

$\frac{1}{2}$,而 $P(A)=\frac{1}{2}$,故 $P(A|B)=P(A)$.

(iii) 设试验为将骰子掷一次,A 同上,B 为"掷出点数$\geqslant 4$",则 $P(A|B)=\frac{2}{3}$,而 $P(A)=\frac{1}{2}$,故 $P(A|B)>P(A)$.

(2) 因 A,B,C 相互独立,故 $P(AB)=P(A)P(B)$,$P(BC)=P(B)P(C)$,$P(CA)=P(C)P(A)$,$P(ABC)=P(A)P(B)P(C)$,从而

(i) $P(C(AB))=P(CAB)=P(C)P(A)P(B)=P(C)P(AB)$,这表示 C 与 AB 相互独立.

(ii)
$$\begin{aligned}P(C(A\cup B))&=P(CA\cup CB)=P(CA)+P(CB)-P(CAB)\\&=P(C)P(A)+P(C)P(B)-P(C)P(A)P(B)\\&=P(C)[P(A)+P(B)-P(AB)]=P(C)P(A\cup B),\end{aligned}$$
这表示 C 与 $A\cup B$ 相互独立.

(3) 因 $AB\subset A$,故若 $P(A)=0$,则
$$0\leqslant P(AB)\leqslant P(A).$$
从而
$$P(AB)=0=P(B)\cdot 0=P(B)\cdot P(A),$$
按定义,A,B 相互独立.

(4) 必要性. 设 A,B 相互独立,则 $A,\bar{B}$ 也相互独立,从而知 $P(A|B)=P(A)$,$P(A|\bar{B})=P(A)$,故 $P(A|B)=P(A|\bar{B})$.

充分性. 设 $P(A|B)=P(A|\bar{B})$,按定义此式即表示
$$\frac{P(AB)}{P(B)}=\frac{P(A\bar{B})}{P(\bar{B})}.$$
由比例的性质得
$$\frac{P(AB)}{P(B)}=\frac{P(AB)+P(A\bar{B})}{P(B)+P(\bar{B})}=\frac{P(A(B\cup\bar{B}))}{1}=P(A).$$
即 $P(AB)=P(A)P(B)$,即 A,B 相互独立.

31. 设事件 A,B 的概率均大于零,说明以下的叙述(a) 必然对,(b) 必然错,(c) 可能对. 并说明理由.

(1) 若 A 与 B 互不相容,则它们相互独立.

(2) 若 A 与 B 相互独立,则它们互不相容.

(3) $P(A)=P(B)=0.6$,且 A,B 互不相容.

(4) $P(A)=P(B)=0.6$,且 A,B 相互独立.

解 (1) 必然错. 因若 A,B 互不相容,则 $0=P(AB)\neq P(A)P(B)$.

(2) 必然错. 因若 A,B 相互独立,则 $P(AB)=P(A)P(B)>0$.

(3) 必然错. 因若 A,B 互不相容,则 $P(A\cup B)=P(A)+P(B)=1.2$,这是

不对的.

(4) 可能对.

32. 有一种检验艾滋病毒的检验法，其结果有概率 0.005 误认为假阳性(即不带艾滋病毒者，经此检验法有 0.005 的概率被认为带艾滋病毒). 今有 140 名不带艾滋病毒的正常人全部接受此种检验，被报道至少有一人带艾滋病毒的概率为多少?

解　在本题中，这 140 人检查结果是相互独立的，这一假定是合理的，将人编号，第 i 号人检验结果以 A_i 表示正常，则 $\overline{A}_i$ 表示被报道为带艾滋病毒者，由题意知 $P(\overline{A}_i)=0.005$，从而 $P(A_i)=1-0.005=0.995$. 于是 140 人经检验至少有一人被报道呈阳性的概率为

$$p = P\{至少有一人呈阳性\} = 1 - P\{无人为阳性\}$$

$$= 1 - P\left(\bigcup_{i=1}^{140} A_i\right) = 1 - \prod_{i=1}^{140} P(A_i) = 1 - 0.995^{140}.$$

由 $140\lg 0.995=\lg 0.495\,7$，得

$$p=1-0.495\,7=0.504\,3.$$

这说明，即使无人带艾滋病毒，这样的检验法认为 140 人中至少有一人带艾滋病毒的概率大于 $\frac{1}{2}$.

33. 盒中有编号为 1,2,3,4 的 4 只球，随机地自盒中取一只球，事件 A 为"取得的是 1 号或 2 号球"，事件 B 为"取得的是 1 号或 3 号球"，事件 C 为"取得的是 1 号或 4 号球". 验证:

$$P(AB)=P(A)P(B),\quad P(AC)=P(A)P(C),\quad P(BC)=P(B)P(C),$$

但
$$P(ABC)\neq P(A)P(B)P(C),$$

即事件 A,B,C 两两独立，但 A,B,C 不是相互独立的.

证　以 $A_i(i=1,2,3,4)$ 表示取到第 i 号球，则

$$P(A_1)=P(A_2)=P(A_3)=P(A_4)=\frac{1}{4},$$

又 $A=A_1\cup A_2, B=A_1\cup A_3, C=A_1\cup A_4$，且 A_1,A_2,A_3,A_4 两两互不相容，故有

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

另外，$AB=A_1, AC=A_1, BC=A_1, ABC=A_1$，故

$$P(AB)=P(AC)=P(BC)=P(ABC)=P(A_1)=\frac{1}{4}.$$

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

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

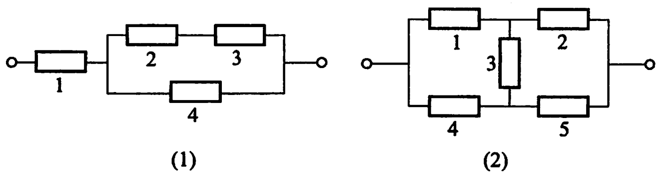
$$P(BC)=\frac{1}{4}=\frac{1}{2}\times\frac{1}{2}=P(B)P(C),$$

但 $P(ABC)=\frac{1}{4}\neq P(A)P(B)P(C)=\frac{1}{8}$.

34. 试分别求以下两个系统的可靠性：

(1) 设有 4 个独立工作的元件 1,2,3,4. 它们的可靠性分别为 p_1,p_2,p_3,p_4，将它们按题 1.34 图 1(1)的方式连接(称为并串联系统).

(2) 设有 5 个独立工作的元件 1,2,3,4,5. 它们的可靠性均为 p，将它们按题 1.34 图 1(2)的方式连接(称为桥式系统).



题 1.34 图 1

解 (1) 以 A_i 表示事件"第 i 只元件正常工作"，$i=1,2,3,4$，以 A 表示"系统正常工作"，已知各元件是否正常工作相互独立，且有 $P(A_i)=p_i\ (i=1,2,3,4)$. 由图知

$$A=A_1[(A_2A_3)\cup A_4]=A_1A_2A_3\cup A_1A_4.$$

由加法公式及各元件工作的独立性得到

$$\begin{aligned}P(A)&=P(A_1A_2A_3)+P(A_1A_4)-P[(A_1A_2A_3)\cap(A_1A_4)]\\&=P(A_1)P(A_2)P(A_3)+P(A_1)P(A_4)-P(A_1A_2A_3A_4)\\&=p_1p_2p_3+p_1p_4-P(A_1)P(A_2)P(A_3)P(A_4)\\&=p_1(p_4+p_2p_3-p_2p_3p_4).\end{aligned}$$

(2) 以 A_i 表示事件"第 i 只元件正常工作"，$i=1,2,3,4,5$，以 A 表示"系统正常工作"，已知各元件是否正常工作相互独立，且有 $P(A_i)=p\ (i=1,2,3,4,5)$.

解法(i)(路径穷举法) 根据系统逻辑框图(题 1.34 图 1(2))，将所有能使系统正常工作的路径一一列出，再利用概率的加法定理和乘法定理来计算系统的可靠性 $P(A)$. 由图知使桥式系统正常工作的路径有下列 4 条：12，45，135，432. 以 B_j 记事件第 $j\ (j=1,2,3,4)$ 条路径正常工作，即有

$$B_1=A_1A_2,\quad B_2=A_4A_5,\quad B_3=A_1A_3A_5,\quad B_4=A_4A_3A_2,$$

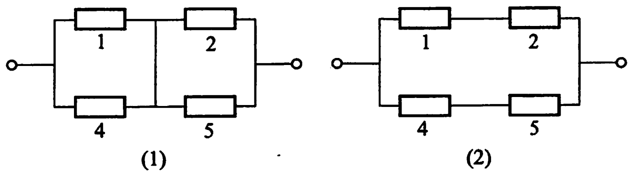
于是得系统的可靠性为

$$
\begin{aligned}
P(A)&=P(B_1\cup B_2\cup B_3\cup B_4)\\
&=\sum_{i=1}^{4}P(B_i)-\sum_{1\leqslant i<j\leqslant 4}P(B_iB_j)+\sum_{1\leqslant i<j<k\leqslant 4}P(B_iB_jB_k)-P(B_1B_2B_3B_4)\\
&=P(A_1A_2)+P(A_4A_5)+P(A_1A_3A_5)+P(A_4A_3A_2)-P(A_1A_2A_4A_5)\\
&\quad -P(A_1A_2A_3A_5)-P(A_1A_2A_3A_4)-P(A_1A_3A_4A_5)\\
&\quad -P(A_2A_3A_4A_5)-P(A_1A_2A_3A_4A_5)+P(A_1A_2A_3A_4A_5)\\
&\quad +P(A_1A_2A_3A_4A_5)+P(A_1A_2A_3A_4A_5)+P(A_1A_2A_3A_4A_5)\\
&\quad -P(A_1A_2A_3A_4A_5)\\
&=2p^2+2p^3-5p^4+2p^5.
\end{aligned}
$$

解法(ii)（全概率公式法）　按元件3处于正常工作与失效两种状态，将原系统简化为典型的并串联和串并联系统，再用全概率公式：

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

来计算原系统的可靠性.



题 1.34 图 2

当元件3正常工作时，系统简化成如题1.34图2(1)所示，当元件3失效时，系统简化成如题1.34图2(2)所示. 因此

$$P(A\mid A_3)=P[(A_1\cup A_4)(A_2\cup A_5)],$$

$$P(A\mid\overline{A}_3)=P(A_1A_2\cup A_4A_5).$$

故　$P(A)=P[(A_1\cup A_4)(A_2\cup A_5)]P(A_3)+P(A_1A_2\cup A_4A_5)P(\overline{A}_3)$.

注意到

$$P(A_1\cup A_4)=P(A_1)+P(A_4)-P(A_1A_4)=2p-p^2.$$

同样　$$P(A_2\cup A_5)=2p-p^2.$$

$$
\begin{aligned}
P(A_1A_2\cup A_4A_5)&=P(A_1)P(A_2)+P(A_4)P(A_5)-P(A_1)P(A_2)P(A_4)P(A_5)\\
&=2p^2-p^4.
\end{aligned}
$$

即得原系统的可靠性为

$$
\begin{aligned}
P(A)&=(2p-p^2)^2p+(2p^2-p^4)(1-p)\\
&=2p^2+2p^3-5p^4+2p^5.
\end{aligned}
$$

注：本题易犯的错误是在使用概率的加法公式时，没有注意两事件不是不相

容的，例如在(1)中应有 $P(A_1A_2A_3 \cup A_1A_4) = P(A_1A_2A_3) + P(A_1A_4) - P(A_1A_2A_3A_4)$，而错误地将最后一项 $P(A_1A_2A_3A_4)$漏了.

35. 如果一危险情况 C 发生时，一电路闭合并发出警报，我们可以借用两个或多个开关并联以改善可靠性. 在 C 发生时这些开关每一个都应闭合，且若至少一个开关闭合了，警报就发出. 如果两个这样的开关并联连接，它们每个具有 0.96 的可靠性(即在情况 C 发生时闭合的概率)，问这时系统的可靠性(即电路闭合的概率)是多少？如果需要有一个可靠性至少为 0.999 9 的系统，则至少需要用多少只开关并联？设各开关闭合与否是相互独立的.

解 以 A_i 表示事件"第 i 只开关闭合"，$i=1,2,\cdots,n$. 已知 $P(A_i)=0.96$，由此可得两只这样的开关并联而电路闭合的概率为(注意各开关闭合与否是相互独立的)

$$\begin{aligned} P(A_1 \cup A_2) &= P(A_1)+P(A_2)-P(A_1A_2) \\ &= P(A_1)+P(A_2)-P(A_1)P(A_2) \\ &= 2\times 0.96-0.96^2 = 0.998\ 4. \end{aligned}$$

设需要 n 只这样的开关并联，此时系统可靠性 $R=P\left(\bigcup\limits_{i=1}^{n} A_i\right)$，注意到

$$\overline{\bigcup_{i=1}^{n} A_i} = \overline{A_1 \cup A_2 \cup \cdots \cup A_n} = \overline{A}_1\overline{A}_2\cdots\overline{A}_n,$$

且由 $A_1,A_2,\cdots,A_n$ 的独立性推得 $\overline{A}_1,\overline{A}_2,\cdots,\overline{A}_n$ 也相互独立. 故

$$\begin{aligned} R &= P\left(\bigcup_{i=1}^{n} A_i\right) = 1-P\left(\overline{\bigcup_{i=1}^{n} A_i}\right) = 1-P(\overline{A}_1\overline{A}_2\cdots\overline{A}_n) \\ &= 1-P(\overline{A}_1)P(\overline{A}_2)\cdots P(\overline{A}_n) = 1-(1-0.96)^n. \end{aligned}$$

要使 $R\geqslant 0.999\ 9$，即要使 $1-0.04^n\geqslant 0.999\ 9$，亦即要使 $0.000\ 1\geqslant 0.04^n$. 故应有

$$n \geqslant \frac{\lg 0.000\ 1}{\lg 0.04} = \frac{4}{\lg 25} = \frac{4}{1.397\ 9} = 2.86.$$

因 n 为整数，故应有 $n\geqslant 3$，即至少要用 3 只开关并联.

36. 三人独立地去破译一份密码，已知各人能译出的概率分别为 $\frac{1}{5},\frac{1}{3},\frac{1}{4}$. 问三人中至少有一人能将此密码译出的概率是多少？

解 以 A_i 表示事件"第 i 人能译出密码"，$i=1,2,3$. 已知 $P(A_1)=\frac{1}{5}$，$P(A_2)=\frac{1}{3}$，$P(A_3)=\frac{1}{4}$，则至少有一人能译出密码的概率为

$$p=P(A_1\cup A_2\cup A_3)$$
$$=P(A_1)+P(A_2)+P(A_3)-P(A_1A_2)-P(A_1A_3)-P(A_2A_3)+P(A_1A_2A_3).$$

由独立性即得

$$p=\frac{1}{5}+\frac{1}{3}+\frac{1}{4}-\frac{1}{5}\times\frac{1}{3}-\frac{1}{5}\times\frac{1}{4}-\frac{1}{3}\times\frac{1}{4}+\frac{1}{5}\times\frac{1}{3}\times\frac{1}{4}=\frac{3}{5}.$$

也可以这样做，因 A_1,A_2,A_3 相互独立，知 $\overline{A}_1,\overline{A}_2,\overline{A}_3$ 也相互独立，即有

$$p=1-P(\overline{A}_1\overline{A}_2\overline{A}_3)=1-P(\overline{A}_1)P(\overline{A}_2)P(\overline{A}_3)$$
$$=1-\left(1-\frac{1}{5}\right)\left(1-\frac{1}{3}\right)\left(1-\frac{1}{4}\right)=\frac{3}{5}.$$

37. 设第一只盒子中装有 3 只蓝色球、2 只绿色球、2 只白色球，第二只盒子中装有 2 只蓝色球、3 只绿色球、4 只白色球. 独立地分别在两只盒子中各取一只球.

(1) 求至少有一只蓝色球的概率.

(2) 求有一只蓝色球、一只白色球的概率.

(3) 已知至少有一只蓝色球，求有一只蓝色球、一只白色球的概率.

解　以 B_i 记事件"从第 i 只盒子中取得一只蓝色球"，以 W_i 记事件"从第 i 只盒子中取得一只白色球"，$i=1,2$. 由题设在不同盒子中取球是相互独立的.

(1) 即需求 $P(B_1\cup B_2)$. 利用对立事件来求较方便，即有

$$P(B_1\cup B_2)=1-P(\overline{B_1\cup B_2})=1-P(\overline{B}_1\overline{B}_2)$$
$$=1-P(\overline{B}_1)P(\overline{B}_2)=1-\frac{4}{7}\times\frac{7}{9}=\frac{5}{9}.$$

(2) 即需求事件 $B_1W_2\cup B_2W_1$ 的概率，注意到 B_1,W_1 是互不相容的，即 $B_1W_1=\varnothing$，因而 $(B_1W_2)(B_2W_1)=\varnothing$，故有

$$P(B_1W_2\cup B_2W_1)=P(B_1W_2)+P(B_2W_1)$$
$$=P(B_1)P(W_2)+P(B_2)P(W_1)$$
$$=\frac{3}{7}\times\frac{4}{9}+\frac{2}{9}\times\frac{2}{7}=\frac{16}{63}.$$

(3) 即需求条件概率 $p=P(B_1W_2\cup B_2W_1\mid B_1\cup B_2)$. 因 $(B_1W_2\cup B_2W_1)\subset B_1\cup B_2$，故有

$$p=P[(B_1W_2\cup B_2W_1)(B_1\cup B_2)]/P(B_1\cup B_2)$$
$$=P(B_1W_2\cup B_2W_1)/P(B_1\cup B_2)=\frac{16}{35}.$$

38. 袋中装有 m 枚正品硬币、n 枚次品硬币(次品硬币的两面均印有国徽)，

在袋中任取一枚，将它投掷 r 次，已知每次都得到国徽. 问这枚硬币是正品的概率为多少？

解　以 T 记事件"将硬币投掷 r 次每次都出现国徽"，以 A 记事件"所取到的是正品"，由题设 $P(A)=\frac{m}{m+n}$，$P(\overline{A})=\frac{n}{m+n}$，$P(T|A)=\frac{1}{2^r}$，$P(T|\overline{A})=1$，需要求的是概率 $P(A|T)$. 由贝叶斯公式，所求概率为

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

$$=\left(\frac{1}{2^r}\times\frac{m}{m+n}\right)\bigg/\left(\frac{1}{2^r}\times\frac{m}{m+n}+\frac{n}{m+n}\right)=\frac{m}{m+2^r n}.$$

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

解　在被运输的物品中，随机取 3 件，相当于在物品中抽取 3 次，每次取一件，作不放回抽样. 又根据题中说明抽取一件后，不影响取后一件是否为好品的概率，已知当 A_1 发生时，一件产品是好品的概率为 $1-2\%=0.98$，从而随机取 3 件，它们都是好品的概率为 0.98^3，即

$$P(B|A_1)=0.98^3,$$

同样

$$P(B|A_2)=0.9^3,\quad P(B|A_3)=0.1^3.$$

又知

$$P(A_1)=0.8,\quad P(A_2)=0.15,\quad P(A_3)=0.05.$$

现在 $A_iA_j=\varnothing\,(i\neq j)$，$i,j=1,2,3$，且 $P(A_1\cup A_2\cup A_3)=P(A_1)+P(A_2)+P(A_3)=1$，由教材第一章 §5 的页下注知道此时全概率公式、贝叶斯公式都能够应用，由贝叶斯公式得到

$$P(A_1|B)=\frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1)+P(B|A_2)P(A_2)+P(B|A_3)P(A_3)}$$

$$=\frac{0.98^3\times0.8}{0.862\,4}=0.873\,1,$$

$$P(A_2|B)=\frac{0.9^3\times0.15}{0.862\,4}=0.126\,8,$$

$$P(A_3|B)=\frac{0.1^3\times 0.05}{0.862\ 4}=0.000\ 1.$$

40. 将 A,B,C 三个字母之一输入信道，输出为原字母的概率是 α，而输出为其他某一字母的概率都是 $\frac{1-\alpha}{2}$. 今将字母串 AAAA,BBBB,CCCC 之一输入信道，输入 AAAA,BBBB,CCCC 的概率分别为 $p_1,p_2,p_3(p_1+p_2+p_3=1)$，已知输出为 ABCA，问输入的是 AAAA 的概率是多少？（设信道传输各个字母的工作是相互独立的.）

解　以 A_1,B_1,C_1 分别表示事件“输入 AAAA”“输入 BBBB”“输入 CCCC”，以 D 表示事件“输出 ABCA”. 因事件 A_1,B_1,C_1 两两互不相容，且有 $P(A_1\cup B_1\cup C_1)=P(A_1)+P(B_1)+P(C_1)=p_1+p_2+p_3=1$，因此全概率公式和贝叶斯公式可以使用(参见教材第一章 §5 的页下注). 由贝叶斯公式有

$$P(A_1|D)=\frac{P(A_1D)}{P(D)}$$

$$=\frac{P(D|A_1)p_1}{P(D|A_1)p_1+P(D|B_1)p_2+P(D|C_1)p_3}.$$

在输入为 AAAA(即事件 A_1)输出为 ABCA(即事件 D)时，有两个字母为原字母，另两字母为其他字母，所以

$$P(D|A_1)=\alpha^2\left(\frac{1-\alpha}{2}\right)^2.$$

同理 $P(D|B_1)=P(D|C_1)=\alpha\left(\frac{1-\alpha}{2}\right)^3$. 代入上式并注意到 $p_1+p_2+p_3=1$，得到

$$P(A_1|D)=\frac{2\alpha p_1}{(3\alpha-1)p_1+1-\alpha}.$$

第二章　随机变量及其分布

1. 考虑为期一年的一张保险单，若投保人在投保后一年内因意外死亡，则公司赔付 20 万元；若投保人因其他原因死亡，则公司赔付 5 万元；若投保人在投保期末生存，则公司无须付给任何费用. 若投保人在一年内因意外死亡的概率为 0.000 2，因其他原因死亡的概率为 0.001 0，求公司赔付金额的分布律.

解　设赔付金额为 X（以万元计），由条件知 X 取值为 20，5，0，且已知 $P\{X=20\}=0.000\ 2$，$P\{X=5\}=0.001\ 0$，故 $P\{X=0\}=1-P\{X=20\}-P\{X=5\}=0.998\ 8$，即有分布律

X	20	5	0
p_k	0.000 2	0.001 0	0.998 8

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

(2) 将一颗骰子抛掷两次，以 X 表示两次中得到的小的点数，试求 X 的分布律.

解　(1) 从 1～5 五个正整数中随机取 3 个，以 X 表示 3 个数中的最大值. X 的可能值为 3，4，5. 在五个数中任取 3 个共有 $\binom{5}{3}=10$ 种取法.

$\{X=3\}$ 表示取出的 3 个数以 3 为最大值，其余两个数是 1，2，仅有这一种情况，故 $P\{X=3\}=1\Big/\binom{5}{3}=\frac{1}{10}$.

$\{X=4\}$ 表示取出的 3 个数以 4 为最大值，其余两个数可在 1，2，3 中任取 2 个，共有 $\binom{3}{2}$ 种取法，故 $P\{X=4\}=\binom{3}{2}\Big/\binom{5}{3}=\frac{3}{10}$.

$\{X=5\}$ 表示取出的 3 个数以 5 为最大值，其余两个数可在 1，2，3，4 中任取 2 个，共有 $\binom{4}{2}$ 种取法，故 $P\{X=5\}=\binom{4}{2}\Big/\binom{5}{3}=\frac{3}{5}$. $P\{X=5\}$ 也可由 $1-P\{X=3\}-P\{X=4\}$ 得到.

X 的分布律为

X	3	4	5
p_k	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{6}{10}$

(2) **解法(i)**　以 Y_1,Y_2 分别记第一次、第二次抛掷时骰子出现的点数，样本空间为

$$S=\{(y_1,y_2)\,|\,y_1=1,2,\cdots,6;y_2=1,2,\cdots,6\},$$

共有 $6\times6=36$ 个样本点.

$X=\min\{Y_1,Y_2\}$所有可能取的值为 1,2,3,4,5,6 这 6 个数，当且仅当以下三种情况之一发生时事件$\{X=k\}$ $(k=1,2,3,4,5,6)$发生：

(i) $Y_1=k$ 且 $Y_2=k+1,k+2,\cdots,6$ (共有 $6-k$ 个点)；

(ii) $Y_2=k$ 且 $Y_1=k+1,k+2,\cdots,6$ (共有 $6-k$ 个点)；

(iii) $Y_1=k$ 且 $Y_2=k$(仅有一个点).

因此事件"$X=k$"共包含$(6-k)+(6-k)+1=13-2k$ 个样本点，于是 X 的分布律为

$$P\{X=k\}=\frac{13-2k}{36},\quad k=1,2,3,4,5,6,$$

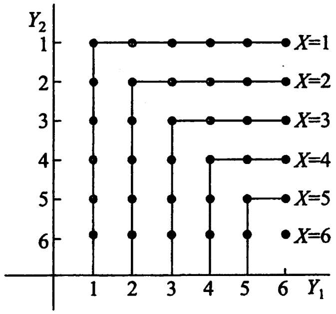
或写成表格形式

X	1	2	3	4	5	6
p_k	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$

解法(ii)　共有 $6\times6=36$ 个样本点，每个样本点发生的概率为 1/36，用题 2.2 图中的黑点表示各样本点. $X=\min\{Y_1,Y_2\}$所有可能取的值为 1,2,3,4,5,6. $P\{X=k\}$等于图中相应折线上各个黑点所对应的概率之和. 即有

X	1	2	3	4	5	6
p_k	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$

亦即　$P\{X=k\}=\dfrac{13-2k}{36}$, $k=1,2,3,4,5,6$.



题 2.2 图

3. 设在 15 只同类型的零件中有 2 只是次品，在其中取 3 次，每次任取 1 只，作不放回抽样. 以 X 表示取出的次品的只数.

（1）求 X 的分布律.

（2）画出分布律的图形.

解 （1）在 15 只零件(其中有 2 只次品)中抽样 3 次，每次任取 1 只作不放回抽样，以 X 表示所得的次品数，X 所有可能取的值为 0，1，2，且有

$$P\{X=0\}=\frac{13}{15}\times\frac{12}{14}\times\frac{11}{13}=\frac{22}{35},$$

$$P\{X=1\}=\frac{2}{15}\times\frac{13}{14}\times\frac{12}{13}+\frac{13}{15}\times\frac{2}{14}\times\frac{12}{13}+\frac{13}{15}\times\frac{12}{14}\times\frac{2}{13}$$

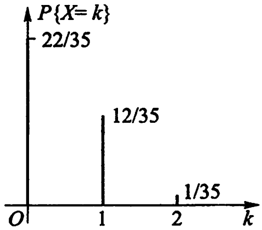
$$=3\left(\frac{2}{15}\times\frac{13}{14}\times\frac{12}{13}\right)=\frac{12}{35},$$

$$P\{X=2\}=1-P\{X=0\}-P\{X=1\}=\frac{1}{35}.$$

分布律为

X	0	1	2
p_k	$\frac{22}{35}$	$\frac{12}{35}$	$\frac{1}{35}$

（2）分布律的图形如题 2.3 图.



题 2.3 图

4. 进行重复独立试验，设每次试验成功的概率为 p，失败的概率为

$$q=1-p\ (0<p<1).$$

（1）将试验进行到出现一次成功为止，以 X 表示所需的试验次数，求 X 的分布律.（此时称 X 服从以 p 为参数的**几何分布**.）

（2）将试验进行到出现 r 次成功为止，以 Y 表示所需的试验次数，求 Y 的分布律.（此时称 Y 服从以 r,p 为参数的**帕斯卡分布或负二项分布**.）

（3）一篮球运动员的投篮命中率为 45%. 以 X 表示他首次投中时累计已投篮的次数，写出 X 的分布律，并计算 X 取偶数的概率.

解 （1）此试验至少做 1 次，此即 X 可能值的最小值. 若需做 k 次，则前 $k-1$次试验均失败，而最后一次成功，由于各次试验是相互独立的，故分布律为

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

（2）此试验至少做 r 次，若需做 k 次，则第 k 次必为成功，而前 $k-1$ 次中有 $r-1$ 次成功，由于各次试验是相互独立的，故分布律为

$$P\{Y=k\}=\binom{k-1}{r-1}p^rq^{k-r},\quad k=r,r+1,\cdots.$$

(3) 先写出 X 的分布律. 它是题(1)中 $p=0.45$ 的情形. 所求分布律为

$$P\{X=k\}=0.45\times 0.55^{k-1},\quad k=1,2,\cdots.$$

因 $\{X=j\}\cap\{X=k\}=\varnothing\ (j\neq k)$，故 X 取偶数的概率为

$$\begin{aligned}P\left\{\bigcup_{k=1}^{\infty}(X=2k)\right\}&=\sum_{k=1}^{\infty}P\{X=2k\}\\&=\sum_{k=1}^{\infty}0.45\times 0.55^{2k-1}=\frac{0.45\times 0.55}{1-0.55^2}=\frac{11}{31}.\end{aligned}$$

5. 一房间有 3 扇同样大小的窗子，其中只有一扇是打开的. 有一只鸟自开着的窗子飞入了房间，它只能从开着的窗子飞出去. 鸟在房间里飞来飞去，试图飞出房间. 假定鸟是没有记忆的，它飞向各扇窗子是随机的.

(1) 以 X 表示鸟为了飞出房间试飞的次数，求 X 的分布律.

(2) 户主声称，他养的一只鸟是有记忆的，它飞向任一窗子的尝试不多于一次. 以 Y 表示这只聪明的鸟为了飞出房间试飞的次数. 如户主所说是确实的，试求 Y 的分布律.

(3) 求试飞次数 X 小于 Y 的概率和试飞次数 Y 小于 X 的概率.

解 (1) 本题的试飞次数是指记录鸟儿飞向窗子的次数加上最后飞离房间的一次，其分布律为(参见第 4 题(1))

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

(2) 由题意 Y 的可能值为 1,2,3.

$\{Y=1\}$表明鸟儿从 3 扇窗子中选对了一扇，因对鸟儿而言，3 扇窗是等可能被选取的，故 $P\{Y=1\}=\frac{1}{3}$.

$\{Y=2\}$表明第一次试飞失败(选错了窗子)，失败方式有 2，故第一次失败的概率为$\frac{2}{3}$；第二次，鸟儿舍弃已飞过的那扇窗，而从余下的一开一关的两窗选一，成功机会为$\frac{1}{2}$，故 $P\{Y=2\}=\frac{2}{3}\times\frac{1}{2}=\frac{1}{3}$.

对有记忆鸟儿来说，$\sum\limits_{i=1}^{3}P\{Y=i\}=1$，故 $P\{Y=3\}=\frac{1}{3}$.

因此 Y 的分布律为

$$P\{Y=i\}=\frac{1}{3},\quad i=1,2,3.$$

(3) (i) $\{X<Y\}$可分解为下列 3 个两两互不相容的事件之和，即

$$\{X<Y\}=\{(X=1)\cap(Y=2)\}\cup\{(X=1)\cap(Y=3)\}$$

$$\cup\{(X=2)\cap(Y=3)\},$$

故

$$P\{X<Y\}=P\{(X=1)\cap(Y=2)\}+P\{(X=1)\cap(Y=3)\}+P\{(X=2)\cap(Y=3)\}.$$

因为两只鸟儿的行动是相互独立的，从而

$$P\{X<Y\}=P\{X=1\}P\{Y=2\}+P\{X=1\}P\{Y=3\}+P\{X=2\}P\{Y=3\}$$
$$=\frac{1}{3}\times\frac{1}{3}+\frac{1}{3}\times\frac{1}{3}+\frac{2}{9}\times\frac{1}{3}=\frac{8}{27}.$$

(ii) $P\{Y<X\}=1-P\{X<Y\}-P\{X=Y\}$

$$=1-\frac{8}{27}-\sum_{k=1}^{3}P\{(X=k)\cap(Y=k)\}$$
$$=1-\frac{8}{27}-\sum_{k=1}^{3}P\{X=k\}P\{Y=k\}$$
$$=1-\frac{8}{27}-\frac{1}{3}\times\frac{1}{3}-\frac{2}{9}\times\frac{1}{3}-\frac{4}{27}\times\frac{1}{3}=\frac{38}{81}.$$

6. 一大楼装有 5 台同类型的供水设备. 设每台设备是否被使用相互独立. 调查表明在任一时刻 t 每台设备被使用的概率为 0.1. 问在同一时刻，

(1) 恰有 2 台设备被使用的概率是多少?

(2) 至少有 3 台设备被使用的概率是多少?

(3) 至多有 3 台设备被使用的概率是多少?

(4) 至少有 1 台设备被使用的概率是多少?

解 以 X 表示同一时刻被使用的设备的个数，则

$$X\sim b(5,0.1).$$

(1) 所求的概率为

$$P\{X=2\}=\binom{5}{2}0.1^2(1-0.1)^3=0.0729.$$

(2) 所求的概率为

$$P\{X\geqslant 3\}=P\{X=3\}+P\{X=4\}+P\{X=5\}$$
$$=\binom{5}{3}0.1^3(1-0.1)^2+\binom{5}{4}0.1^4(1-0.1)+0.1^5$$
$$=0.0081+0.00045+0.00001=0.00856.$$

(3) 所求的概率为

$$P\{X\leqslant 3\}=1-P\{X=4\}-P\{X=5\}$$
$$=1-0.00045-0.00001=0.99954.$$

(4) 所求概率为

$$P\{X\geqslant 1\}=1-P\{X=0\}=1-(1-0.1)^5=0.40951.$$

7. 设事件 A 在每次试验中发生的概率为 0.3. 当 A 发生不少于 3 次时，指示灯发出信号.

(1) 进行了 5 次重复独立试验，求指示灯发出信号的概率.

(2) 进行了 7 次重复独立试验，求指示灯发出信号的概率.

解 (1) 以 X 表示在 5 次试验中事件 A 发生的次数，则 $X\sim b(5,0.3)$. 指示灯发出信号这一事件可表示为 $\{X\geqslant 3\}$，故所求的概率为

$$P\{X\geqslant 3\}=\binom{5}{3}0.3^3(1-0.3)^2+\binom{5}{4}0.3^4(1-0.3)+0.3^5=0.163.$$

(2) 以 Y 记在 7 次试验中事件 A 发生的次数，则 $Y\sim b(7,0.3)$. 故指示灯发出信号的概率为

$$\begin{aligned}P\{Y\geqslant 3\}&=1-P\{Y=0\}-P\{Y=1\}-P\{Y=2\}\\&=1-(1-0.3)^7-\binom{7}{1}(1-0.3)^6 0.3-\binom{7}{2}(1-0.3)^5 0.3^2\\&=0.353.\end{aligned}$$

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

(1) 两人投中次数相等的概率.

(2) 甲比乙投中次数多的概率.

解 以 X,Y 分别表示甲、乙投中的次数，则

$$X\sim b(3,0.6),\quad Y\sim b(3,0.7).$$

(1) 按题意需求事件 $\{X=Y\}$ 的概率，而事件 $\{X=Y\}$ 是下列 4 个两两互不相容的事件之和：

$$(X=0)\cap(Y=0),\quad (X=1)\cap(Y=1),$$
$$(X=2)\cap(Y=2),\quad (X=3)\cap(Y=3).$$

自然，甲、乙投中与否被认为是相互独立的，从而

$$\begin{aligned}P\{X=Y\}&=\sum_{i=0}^{3}P\{(X=i)\cap(Y=i)\}=\sum_{i=0}^{3}P\{X=i\}P\{Y=i\}\\&=(1-0.6)^3(1-0.7)^3+\binom{3}{1}0.6(1-0.6)^2\binom{3}{1}0.7(1-0.7)^2\\&\quad+\binom{3}{2}0.6^2(1-0.6)\binom{3}{2}0.7^2(1-0.7)+0.6^3\times 0.7^3\\&=0.001728+0.054432+0.190512+0.074088=0.321.\end{aligned}$$

(2) 按题意需求事件$\{X>Y\}$的概率，而事件$\{X>Y\}$可表示为下列两两互不相容的事件之和，即

$$\{X>Y\}=\{(X=1)\cap(Y=0)\}\cup\{(X=2)\cap(Y\leqslant 1)\}\cup\{(X=3)\cap(Y\leqslant 2)\}.$$

由于甲、乙投中与否相互独立，所以

$$\begin{aligned}P\{X>Y\}&=P\{(X=1)\cap(Y=0)\}\\&\quad+P\{(X=2)\cap(Y\leqslant 1)\}+P\{(X=3)\cap(Y\leqslant 2)\}\\&=P\{X=1\}P\{Y=0\}+P\{X=2\}P\{Y\leqslant 1\}+P\{X=3\}P\{Y\leqslant 2\}\\&=P\{X=1\}P\{Y=0\}+P\{X=2\}(P\{Y=0\}+P\{Y=1\})\\&\quad+P\{X=3\}(1-P\{Y=3\})\\&=\binom{3}{1}0.6(1-0.6)^2(1-0.7)^3+\binom{3}{2}0.6^2(1-0.6)\\&\quad\times\left[(1-0.7)^3+\binom{3}{1}0.7(1-0.7)^2\right]+0.6^3(1-0.7^3)\\&=0.007\,776+0.093\,312+0.141\,912=0.243.\end{aligned}$$

9. 有一大批产品，其验收方案如下，先作第一次检验：从中任取 10 件，经检验无次品时接受这批产品，次品数大于 2 时拒收；否则作第二次检验，其做法是从中再任取 5 件，仅当 5 件中无次品时接受这批产品. 若产品的次品率为 10%，求

(1) 这批产品经第一次检验就能接受的概率.

(2) 需作第二次检验的概率.

(3) 这批产品按第二次检验的标准被接受的概率.

(4) 这批产品在第一次检验未能作决定且第二次检验时被通过的概率.

(5) 这批产品被接受的概率.

解 由教材第二章 §2 例 2 的说明知，若以 X 表示所抽得的 10 件产品中所含的次品数，则

$$X\sim b(10,0.1),$$

又若以 Y 表示第二次抽检中出现的次品数，则

$$Y\sim b(5,0.1).$$

(1) 按题意所求概率为

$$P\{X=0\}=(1-0.1)^{10}=0.349.$$

(2) 需作第二次检验的概率为

$$P\{1\leqslant X\leqslant 2\}=\binom{10}{1}\times 0.1\times 0.9^9+\binom{10}{2}\times 0.1^2\times 0.9^8=0.581.$$

(3) 按第二次检验标准接受这批产品的概率为

$$P\{Y=0\}=0.9^5=0.590.$$

(4) 所求概率为

$$P\{(1\leqslant X\leqslant 2)\cap(Y=0)\}.$$

因为 X,Y 的取值被认为是放回抽样的结果，即都是独立试验的结果，因此，事件 $\{1\leqslant X\leqslant 2\}$ 与 $\{Y=0\}$ 是相互独立的，从而

$$\begin{aligned}P\{(1\leqslant X\leqslant 2)\cap(Y=0)\}&=P\{1\leqslant X\leqslant 2\}P\{Y=0\}\\&=0.581\times 0.590=0.343.\end{aligned}$$

(5) 这批产品被接受的概率为

$$\begin{aligned}P\{(X=0)\cup[(1\leqslant X\leqslant 2)\cap(Y=0)]\}&=P\{X=0\}+P\{(1\leqslant X\leqslant 2)\cap(Y=0)\}\\&=0.349+0.343=0.692.\end{aligned}$$

10. 有甲、乙两种味道和颜色都极为相似的名酒各 4 杯. 如果从中挑 4 杯，能将甲种酒全部挑出来，算是试验成功一次.

(1) 某人随机地去挑，问他试验成功一次的概率是多少?

(2) 某人声称他通过品尝能区分两种酒. 他连续试验 10 次，成功 3 次. 试推断他是猜对的，还是他确有区分的能力(设各次试验是相互独立的).

解 (1) 某人随机去挑，从 8 杯中挑取 4 杯共有 $\binom{8}{4}=70$ 种取法，其中只有一种是正确的. 故若某人随机去挑，试验成功一次的概率是 $p=\frac{1}{70}$.

(2) 为判断某人是否有区分能力，先假设“某人无区分能力”. 由(1)他挑对一次的概率为$\frac{1}{70}$，连续试验 10 次，则挑对次数 $X\sim b\left(10,\frac{1}{70}\right)$. 今

$$P\{X=3\}=\binom{10}{3}\left(\frac{1}{70}\right)^3\left(1-\frac{1}{70}\right)^7=3.16\times 10^{-4}.$$

不仅如此，

$$\begin{aligned}P\{X\geqslant 3\}&=1-\sum_{k=0}^{2}P\{X=k\}=1-\sum_{k=0}^{2}\binom{10}{k}\left(\frac{1}{70}\right)^k\left(1-\frac{1}{70}\right)^{10-k}\\&=3.24\times 10^{-4}.\end{aligned}$$

即试验 10 次，他挑对次数$\geqslant 3$ 的概率也仅为万分之三. 今事件 $\{X\geqslant 3\}$ 竟然发生了，按实际推断原理，应否定原假设“某人无区分能力”，而认为他确有区分能力.

11. 尽管在几何教科书中已经讲过仅用圆规和直尺三等分一个任意角是不可能的，但每一年总是有一些“发明者”撰写关于仅用圆规和直尺将角三等分的

文章.设某地区每年撰写此类文章的篇数 X 服从参数为 6 的泊松分布.求明年没有此类文章的概率.

解 由题设某地每年撰写此类文章的篇数 $X\sim\pi(6)$，因此，明年无此类文章的概率为

$$P\{X=0\}=\mathrm{e}^{-6}=2.5\times10^{-3}.$$

12. 一电话总机每分钟收到呼唤的次数服从参数为 4 的泊松分布.求

(1) 某一分钟恰有 8 次呼唤的概率.

(2) 某一分钟的呼唤次数大于 3 的概率.

解 以 X 记电话总机一分钟收到呼唤的次数，则有

$$X\sim\pi(4),\quad P\{X=k\}=\frac{4^k\mathrm{e}^{-4}}{k!},\ k=0,1,2,\cdots.$$

(1) 所求概率为

$$P\{X=8\}=\frac{4^8\mathrm{e}^{-4}}{8!}=0.029\ 8.$$

(2) 所求概率为

$$p=\sum_{k=4}^{\infty}P\{X=k\}=1-\sum_{k=0}^{3}P\{X=k\}$$
$$=1-\sum_{k=0}^{3}\frac{4^k\mathrm{e}^{-4}}{k!}=0.566\ 5.$$

13. 某一公安局在长度为 t 的时间间隔内收到的紧急呼救的次数 X 服从参数为$\frac{t}{2}$的泊松分布，而与时间间隔的起点无关(时间以 h 计).求

(1) 某一天中午 12 时至下午 3 时未收到紧急呼救的概率.

(2) 某一天中午 12 时至下午 5 时至少收到 1 次紧急呼救的概率.

解 已知

$$X\sim\pi\left(\frac{t}{2}\right).$$

(1) $t=3$，所求概率为

$$P\{X=0\}=\mathrm{e}^{-3/2}=0.223\ 1.$$

(2) $t=5$，所求概率为

$$P\{X\geqslant1\}=1-P\{X=0\}=1-\mathrm{e}^{-5/2}=0.917\ 9.$$

14. 某人家中在时间间隔 t(以 h 计)内接到电话的次数 X 服从参数为 $2t$ 的泊松分布.

(1) 若他外出计划用时 10 min，问其间电话铃响一次的概率是多少？

(2) 若他希望外出时没有电话的概率至少为 0.5，问他外出应控制的最长时

间是多少？

解 以 X 表示此人外出时电话铃响的次数，则 $X\sim\pi(2t)$，其中 t 表示外出的总时间，即 X 的分布律为 $P\{X=k\}=\dfrac{(2t)^k e^{-2t}}{k!}$，$k=0,1,2,\cdots$.

(1) $t=10/60=1/6$ 时，$X\sim\pi\left(2\times\dfrac{1}{6}\right)$，故所求概率为

$$P\{X=1\}=\frac{1}{3}e^{-1/3}=0.238\ 8.$$

(2) 设外出最长时间为 t(h)，因 $X\sim\pi(2t)$，无电话打进的概率为

$$P\{X=0\}=e^{-2t},$$

要使 $P\{X=0\}=e^{-2t}\geqslant 0.5$，即要使 $e^{2t}\leqslant 2$，由此得

$$t\leqslant\frac{1}{2}\ln 2=0.346\ 6(\mathrm{h}),$$

即外出时间应控制少于 20.79 min.

15. 保险公司在一天内承保了 5 000 张相同年龄、为期一年的寿险保单，每人一份. 在合同有效期内若投保人死亡，则公司需赔付 3 万元. 设在一年内，该年龄段的死亡率为 0.001 5，且各投保人是否死亡相互独立. 求该公司对于这批投保人的赔付总额不超过 30 万元的概率(利用泊松定理计算).

解 设这批投保人在一年内死亡的人数为 X，则 $X\sim b(5\ 000,0.001\ 5)$，因每死亡一人公司需赔付 3 万元，故公司赔付不超过 30 万元意味着在投保期内死亡人数不超过 30/3＝10(人)，从而所求概率为

$$P\{X\leqslant 10\}=\sum_{k=0}^{10}\binom{5\ 000}{k}0.001\ 5^k(1-0.001\ 5)^{5\ 000-k}.$$

若用泊松定理近似可以认为 $X\sim\pi(7.5)$，于是

$$P\{X\leqslant 10\}\approx\sum_{k=0}^{10}\frac{7.5^k e^{-7.5}}{k!}\overset{\text{查表}}{=\!=\!=}0.862\ 2.$$

16. 有一繁忙的汽车站，每天有大量汽车通过，设一辆汽车在一天的某段时间内出事故的概率为 0.000 1. 在某天的该时间段内有 1 000 辆汽车通过. 问出事故的车辆数不小于 2 的概率是多少？(利用泊松定理计算.)

解 以 X 表示汽车站某天该时间段内汽车出事故的辆数，由题设 $X\sim b(1\ 000,0.000\ 1)$，因 $n=1\ 000>100$，且 $np=0.1<10$，故可利用泊松定理计算 $P\{X\geqslant 2\}$，即令 $\lambda=np=0.1$，有

$$P\{X=k\}=\binom{n}{k}p^k(1-p)^{n-k}\approx\frac{\lambda^k e^{-\lambda}}{k!},$$

其中 $\lambda=np=0.1$，从而

$$P\{X\geqslant 2\}=1-P\{X=0\}-P\{X=1\}$$
$$\approx 1-e^{-0.1}-e^{-0.1}\times 0.1=0.0047.$$

17. (1) 设 X 服从(0－1)分布，其分布律为 $P\{X=k\}=p^k(1-p)^{1-k}, k=0, 1$，求 X 的分布函数，并作出其图形.

(2) 求第 2 题(1)中的随机变量的分布函数.

解 (1) X 服从(0－1)分布，分布律为

X	0	1
p_k	$1-p$	p

当 $x<0$ 时，$F(x)=P\{X\leqslant x\}=0$；

当 $0\leqslant x<1$ 时，

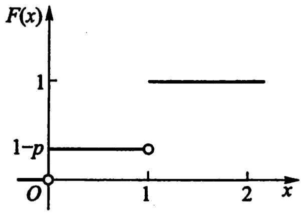
$$F(x)=P\{X\leqslant x\}=P\{X=0\}=1-p;$$

当 $x\geqslant 1$ 时，

$$\begin{aligned}F(x)&=P\{X\leqslant x\}\\&=P\{X=0\}+P\{X=1\}\\&=(1-p)+p=1,\end{aligned}$$

即有(如题 2.17 图)

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



题 2.17 图

(2) X 的分布律为

X	3	4	5
p_k	0.1	0.3	0.6

X 的分布函数为 $F(x)=P\{X\leqslant x\}$，即有

当 $x<3$ 时，

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

当 $3\leqslant x<4$ 时，

$$F(x)=P\{X\leqslant x\}=P\{X=3\}=0.1;$$

当 $4\leqslant x<5$ 时，

$$\begin{aligned}F(x)&=P\{X\leqslant x\}=P\{X=3\}+P\{X=4\}\\&=0.1+0.3=0.4;\end{aligned}$$

当 $x\geqslant 5$ 时，

$$F(x)=P\{X\leqslant x\}=P\{X=3\}+P\{X=4\}+P\{X=5\}=1,$$

故知

$$F(x)=\begin{cases}0, & x<3,\\ 0.1, & 3\leqslant x<4,\\ 0.4, & 4\leqslant x<5,\\ 1, & x\geqslant 5.\end{cases}$$

注意，$F(x)$的定义域总是$(-\infty,\infty)$，它是一个右连续函数，且有 $F(\infty)=1$. 读者可核对一下是否做对了.

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

解　X 的分布函数为

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

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

当 $0\leqslant x\leqslant a$ 时，按题意 $P\{0\leqslant X\leqslant x\}=kx$，$k$ 是某一常数，为了确定 k，取 $x=a$，得 $P\{0\leqslant X\leqslant a\}=ka$. 因我们只是在区间$[0,a]$上投掷质点，所以 $\{0\leqslant X\leqslant a\}$为必然事件，即有$1=P\{0\leqslant X\leqslant a\}=ka$，$k=\dfrac{1}{a}$，因此：

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

当 $x\geqslant a$ 时，按题意 $P\{X\leqslant x\}$为必然事件，即有

$$F(x)=P\{X\leqslant x\}=1,$$

故知

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

19. 以 X 表示某商店从早晨开始营业起直到第一个顾客到达的等待时间(以 min 计)，X 的分布函数是

$$F_X(x)=\begin{cases}1-\mathrm{e}^{-0.4x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

求下述概率：

(1) $P\{$至多 3 min$\}$.

(2) $P\{$至少 4 min$\}$.

(3) $P\{$3 min 至 4 min 之间$\}$.

(4) $P\{$至多 3 min 或至少 4 min$\}$.

(5) $P\{$恰好 2.5 min$\}$.

解　(1) $P\{$至多 3 min$\}=P\{X\leqslant 3\}=F_X(3)=1-e^{-1.2}$.

(2) $P\{$至少 4 min$\}=P\{X\geqslant 4\}=1-P\{X<4\}$

$$=1-P\{X\leqslant 4\}=1-F_X(4)=e^{-1.6}.$$

(因为 $F_X(x)$ 是指数分布随机变量 X 的分布函数，X 是连续型随机变量，故 $P\{X=4\}=0, P\{X<4\}=P\{X\leqslant 4\}$.)

(3) $P\{$3 min 至 4 min 之间$\}=P\{3\leqslant X\leqslant 4\}=P\{3<X\leqslant 4\}$

$$=F_X(4)-F_X(3)=e^{-1.2}-e^{-1.6}.$$

(4) $P\{$至多 3 min 或至少 4 min$\}=P\{(X\leqslant 3)\cup(X\geqslant 4)\}$

$$=P\{X\leqslant 3\}+P\{X\geqslant 4\}$$

$$=1-e^{-1.2}+e^{-1.6}.$$

(5) $P\{$恰好 2.5 min$\}=P\{X=2.5\}=0$.

20. 设随机变量 X 的分布函数为

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

(1) 求 $P\{X<2\}, P\{0<X\leqslant 3\}, P\left\{2<X<\frac{5}{2}\right\}$.

(2) 求概率密度 $f_X(x)$.

解　(1) 对应任意指定的实数 a，只要随机变量 X(不管什么类型)的分布函数在点 a 处连续，则 $P\{X=a\}=0$(参见教材第二章 §4 关于 $P\{X=a\}=0$ 的证明)，故有

$$P\{X<2\}=P\{X\leqslant 2\}=F_X(2)=\ln 2,$$

$$P\{0<X\leqslant 3\}=F_X(3)-F_X(0)=1-0=1.$$

$$P\left\{2<X<\frac{5}{2}\right\}=P\left\{2<X\leqslant\frac{5}{2}\right\}=F_X\left(\frac{5}{2}\right)-F_X(2)$$

$$=\ln\frac{5}{2}-\ln 2=\ln\frac{5}{4}.$$

(2) 由于在 $f_X(x)$ 的连续点处有 $\frac{d}{dx}F_X(x)=f_X(x)$，即有

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

因 $f_X(x)$在个别点的函数值可以是随意指定的有限值（参见教材第二章§4 页下注），这里我们指定 $f_X(1)=0, f_X(\mathrm{e})=0$.

21. 设随机变量 X 的概率密度为

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

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

求 X 的分布函数 $F(x)$，并画出(2)中的 $f(x)$及 $F(x)$的图形.

解　(1) 因概率密度 $f(x)$在 $x<1, x>2$ 处都等于零，即知

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

当 $x>2$ 时，$F(x)=\int_{-\infty}^{x} f(x)\mathrm{d}x=1-\int_{x}^{\infty} f(x)\mathrm{d}x=1-\int_{x}^{\infty} 0\mathrm{d}x=1$;

当 $1\leqslant x\leqslant 2$ 时，$F(x)=\int_{-\infty}^{x} f(x)\mathrm{d}x=\int_{-\infty}^{1} 0\mathrm{d}x+\int_{1}^{x} 2\left(1-\frac{1}{x^2}\right)\mathrm{d}x$

$$=2\left(x+\frac{1}{x}\right)\Big|_1^x=2\left(x+\frac{1}{x}-2\right).$$

故所求分布函数是

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

(2) 概率密度 $f(x)$在 $x<0, x\geqslant 2$ 处都等于零，即知

当 $x<0$ 时，$F(x)=0$；$x\geqslant 2$ 时，$F(x)=1$，又

当 $0\leqslant x<1$ 时，

$$F(x)=\int_{-\infty}^{x} f(x)\mathrm{d}x=\int_{-\infty}^{0} 0\mathrm{d}x+\int_{0}^{x} x\mathrm{d}x=\frac{x^2}{2}.$$

当 $1\leqslant x<2$ 时，

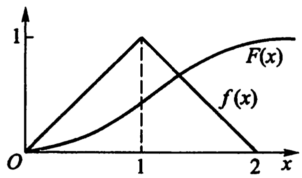
$$F(x)=\int_{-\infty}^{x} f(x)\mathrm{d}x=\int_{-\infty}^{0} 0\mathrm{d}x+\int_{0}^{1} x\mathrm{d}x+\int_{1}^{x}(2-x)\mathrm{d}x$$

$$=\frac{1}{2}+\left(2x-\frac{x^2}{2}\right)\Big|_1^x=2x-\frac{x^2}{2}-1.$$

故所求分布函数为

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

$f(x)$，$F(x)$的图形如题 2.21 图.



题 2.21 图

注：分段定义的连续型随机变量的分布函数 $F(x)$，由于 $F(x)$连续，它的定义域中各子区间的端点，只要求表达清楚，属于哪一个区间无关紧要，也没必要与 $f(x)$一致.

22. (1) 分子运动速度的绝对值 X 服从麦克斯韦(Maxwell)分布，其概率密度为

$$f(x)=\begin{cases}Ax^2\mathrm{e}^{-x^2/b}, & x>0,\\ 0, & \text{其他},\end{cases}$$

其中 $b=m/(2kT)$，k 为玻耳兹曼(Boltzmann)常数，T 为绝对温度，m 是分子的质量，试确定常数 A.

(2) 某人研究了英格兰在 1875—1951 年间，在矿山发生导致不少于 10 人死亡的事故的频繁程度，得知相继两次事故之间的时间 T(以日计)服从指数分布，其概率密度为

$$f_T(t)=\begin{cases}\dfrac{1}{241}\mathrm{e}^{-t/241}, & t>0,\\ 0, & \text{其他}.\end{cases}$$

求分布函数 $F_T(t)$，并求概率 $P\{50<T<100\}$.

解 (1) 因 X 的概率密度函数 $f(x)$具有性质

$$\int_{-\infty}^{\infty}f(x)\mathrm{d}x=1,$$

故由

$$1=\int_{-\infty}^{\infty}f(x)\mathrm{d}x=\int_{-\infty}^{0}0\mathrm{d}x+\int_{0}^{\infty}Ax^2\mathrm{e}^{-x^2/b}\mathrm{d}x$$

$$=-A\,\frac{b}{2}x\mathrm{e}^{-x^2/b}\Big|_0^{\infty}+\frac{Ab}{2}\int_0^{\infty}\mathrm{e}^{-x^2/b}\mathrm{d}x$$

$$\xlongequal{\text{令 } x/\sqrt{b}=u}\frac{Ab\sqrt{b}}{2}\int_0^{\infty}\mathrm{e}^{-u^2}\mathrm{d}u=\frac{Ab\sqrt{b}}{4}\sqrt{\pi}$$

$\left(\text{由}\int_{-\infty}^{\infty}\mathrm{e}^{-t^2/2}\mathrm{d}t=\sqrt{2\pi}\text{，知}\int_0^{\infty}\mathrm{e}^{-u^2}\mathrm{d}u=\frac{\sqrt{\pi}}{2}\right)$，得到

$$A=\frac{4}{b\sqrt{b\pi}}.$$

(2) 当 $t\leqslant 0$ 时，$F_T(t)=\int_{-\infty}^{t}f_T(t)\mathrm{d}t=\int_{-\infty}^{t}0\mathrm{d}t=0$；

当 $t>0$ 时，

$$F_T(t)=\int_{-\infty}^{t}f_T(t)\mathrm{d}t=\int_{-\infty}^{0}0\mathrm{d}t+\int_0^t\frac{1}{241}\mathrm{e}^{-t/241}\mathrm{d}t=1-\mathrm{e}^{-t/241},$$

故所求的分布函数为

$$F_T(t)=\begin{cases}1-\mathrm{e}^{-t/241}, & t>0,\\ 0, & \text{其他}.\end{cases}$$

而 $P\{50<T<100\}=F_T(100)-F_T(50)=\mathrm{e}^{-50/241}-\mathrm{e}^{-100/241}$.

23. 某种型号器件的寿命 X(以 h 计)具有概率密度

$$f(x)=\begin{cases}\dfrac{1\ 000}{x^2}, & x>1\ 000,\\ 0, & \text{其他}.\end{cases}$$

现有一大批此种器件(设各器件损坏与否相互独立)，任取 5 只，问其中至少有 2 只寿命大于 1 500 h 的概率是多少？

解 任取一只该种器件，其寿命大于 1 500 h 的概率为

$$p=\int_{1\,500}^{\infty}\frac{1\ 000}{x^2}\mathrm{d}x=-\frac{1\ 000}{x}\bigg|_{1\,500}^{\infty}=\frac{2}{3}.$$

任取 5 只这种器件，其中寿命大于 1 500 h 的只数记为 X，则 $X\sim b(5,2/3)$(理由参见教材第二章 §2 例 2). 故所求概率为

$$\begin{aligned}P\{X\geqslant 2\}&=1-P\{X=0\}-P\{X=1\}\\&=1-\left(1-\frac{2}{3}\right)^5-\binom{5}{1}\frac{2}{3}\left(1-\frac{2}{3}\right)^4=\frac{232}{243}.\end{aligned}$$

24. 设顾客在某银行的窗口等待服务的时间 X(以 min 计)服从指数分布，其概率密度为

$$f_X(x)=\begin{cases}\dfrac{1}{5}\mathrm{e}^{-x/5}, & x>0,\\ 0, & \text{其他}.\end{cases}$$

某顾客在窗口等待服务，若超过 10 min 他就离开. 他一个月要到银行 5 次. 以 Y 表示一个月内他未等到服务而离开窗口的次数. 写出 Y 的分布律，并求 $P\{Y\geqslant 1\}$.

解 顾客在窗口等待服务超过 10 min 的概率为

$$p = \int_{10}^{\infty} f_X(x)\mathrm{d}x = \int_{10}^{\infty} \frac{1}{5}\mathrm{e}^{-x/5}\mathrm{d}x = \mathrm{e}^{-2},$$

故顾客去银行一次因未等到服务而离开的概率为 e^{-2}. 从而 $Y \sim b(5, \mathrm{e}^{-2})$. Y 的分布律为

$$P\{Y = k\} = \binom{5}{k}(\mathrm{e}^{-2})^k(1-\mathrm{e}^{-2})^{5-k}, \quad k = 0,1,2,3,4,5.$$

$$P\{Y \geqslant 1\} = 1 - P\{Y = 0\} = 1 - (1-\mathrm{e}^{-2})^5 = 0.516\,7.$$

25. 设 K 在 $(0,5)$ 内服从均匀分布，求 x 的方程

$$4x^2 + 4Kx + K + 2 = 0$$

有实根的概率.

解 x 的二次方程 $4x^2+4Kx+K+2=0$ 有实根的充要条件是它的判别式

$$\Delta = (4K)^2 - 4 \times 4(K+2) \geqslant 0,$$

即
$$16(K+1)(K-2) \geqslant 0,$$

解得
$$K \geqslant 2, \text{或 } K \leqslant -1.$$

由假设 K 在区间 $(0,5)$ 上服从均匀分布，其概率密度为

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

故这个二次方程有实根的概率为

$$\begin{aligned} p &= P\{(K \geqslant 2) \cup (K \leqslant -1)\} = P\{K \geqslant 2\} + P\{K \leqslant -1\} \\ &= \int_2^{\infty} f_K(x)\mathrm{d}x + \int_{-\infty}^{-1} f_K(x)\mathrm{d}x = \int_2^5 \frac{1}{5}\mathrm{d}x + \int_{-\infty}^{-1} 0\mathrm{d}x = \frac{3}{5}. \end{aligned}$$

26. 设 $X \sim N(3, 2^2)$.

(1) 求 $P\{2<X\leqslant 5\}$, $P\{-4<X\leqslant 10\}$, $P\{|X|>2\}$, $P\{X>3\}$.

(2) 确定 c，使得 $P\{X>c\}=P\{X\leqslant c\}$.

(3) 设 d 满足 $P\{X>d\}\geqslant 0.9$，问 d 至多为多少?

解 (1) 因 $X \sim N(3, 2^2)$，故有

$$P\{a < X \leqslant b\} = P\left\{\frac{a-3}{2} < \frac{X-3}{2} \leqslant \frac{b-3}{2}\right\} = \Phi\left(\frac{b-3}{2}\right) - \Phi\left(\frac{a-3}{2}\right).$$

$$\begin{aligned} P\{2 < X \leqslant 5\} &= \Phi\left(\frac{5-3}{2}\right) - \Phi\left(\frac{2-3}{2}\right) = \Phi(1) - \Phi(-0.5) \\ &= \Phi(1) - [1 - \Phi(0.5)] = 0.841\,3 - 1 + 0.691\,5 \\ &= 0.532\,8. \end{aligned}$$

$$\begin{aligned}P\{-4<X\leqslant 10\}&=\Phi\left(\frac{10-3}{2}\right)-\Phi\left(\frac{-4-3}{2}\right)\\&=\Phi(3.5)-\Phi(-3.5)=2\Phi(3.5)-1\\&=2\times 0.9998-1=0.9996.\\P\{|X|>2\}&=1-P\{|X|\leqslant 2\}=1-P\{-2\leqslant X\leqslant 2\}\\&=1-\left[\Phi\left(\frac{2-3}{2}\right)-\Phi\left(\frac{-2-3}{2}\right)\right]\\&=1-\Phi(-0.5)+\Phi(-2.5)\\&=\Phi(0.5)+1-\Phi(2.5)\\&=0.6915+1-0.9938=0.6977.\\P\{X>3\}&=1-P(X\leqslant 3)=1-\Phi\left(\frac{3-3}{2}\right)\\&=1-\Phi(0)=1-0.5=0.5.\end{aligned}$$

(2) 由 $P\{X>c\}=P\{X\leqslant c\}$，得

$$1-P\{X\leqslant c\}=P\{X\leqslant c\},\quad P\{X\leqslant c\}=\frac{1}{2},$$

即有 $\Phi\left(\frac{c-3}{2}\right)=\frac{1}{2}=\Phi(0)$，于是$\frac{c-3}{2}=0$，$c=3$.

(3) $P\{X>d\}\geqslant 0.9$，即

$$1-\Phi\left(\frac{d-3}{2}\right)\geqslant 0.9,$$

$$\Phi\left(-\frac{d-3}{2}\right)\geqslant 0.9=\Phi(1.282),$$

因分布函数 $\Phi(x)$ 是一个不减函数，故有

$$-\frac{d-3}{2}\geqslant 1.282,$$

因此 $$d\leqslant 3+2\times(-1.282)=0.436.$$

27. 某地区 18 岁的女青年的血压(收缩压，以 mmHg 计，1 mmHg = 133.322 4 Pa)服从 $N(110,12^2)$分布. 在该地区任选一 18 岁的女青年，测量她的血压 X.

(1) 求 $P\{X\leqslant 105\}$，$P\{100<X\leqslant 120\}$.

(2) 确定最小的 x，使 $P\{X>x\}\leqslant 0.05$.

解 (1) 因为 $X\sim N(110,12^2)$，故有

$$\begin{aligned}P\{X\leqslant 105\}&=\Phi\left(\frac{105-110}{12}\right)=\Phi\left(\frac{-5}{12}\right)\\&=1-\Phi(0.417)=1-0.6617=0.3383.\end{aligned}$$

$$P\{100 < X \leqslant 120\} = \Phi\left(\frac{120-110}{12}\right) - \Phi\left(\frac{100-110}{12}\right)$$
$$= 2\Phi\left(\frac{10}{12}\right) - 1 = 2\Phi(0.833) - 1$$
$$= 2 \times 0.797\,6 - 1 = 0.595\,2.$$

(2) 要求 $P\{X>x\}\leqslant 0.05$. 因 $P\{X>x\}=1-P\{X\leqslant x\}=1-\Phi\left(\frac{x-110}{12}\right)$，即要求

$$1-\Phi\left(\frac{x-110}{12}\right)\leqslant 0.05,$$

即需
$$\Phi\left(\frac{x-110}{12}\right)\geqslant 0.95=\Phi(1.645).$$

由此得
$$\frac{x-110}{12}\geqslant 1.645, \quad x\geqslant 129.74.$$

故 x 的最小值为 129.74.

28. 由某机器生产的螺栓的长度(以 cm 计)服从参数 $\mu=10.05,\sigma=0.06$ 的正态分布. 规定长度在范围 10.05 ± 0.12 内为合格品，求一螺栓为不合格品的概率.

解 记螺栓的长度为 $X, X\sim N(10.05,0.06^2)$，螺栓不合格的概率为

$$1-P\{10.05-0.12 < X < 10.05+0.12\}$$
$$=1-\left[\Phi\left(\frac{10.05+0.12-10.05}{0.06}\right)-\Phi\left(\frac{10.05-0.12-10.05}{0.06}\right)\right]$$
$$=1-\Phi(2)+\Phi(-2)=0.045\,6.$$

29. 一工厂生产的某种元件的寿命 X(以 h 计)服从参数为 $\mu=160,\sigma(\sigma>0)$ 的正态分布. 若要求 $P\{120<X\leqslant 200\}\geqslant 0.80$，允许 σ 最大为多少？

解 $X\sim N(160,\sigma^2)$，今要求

$$P\{120 < X \leqslant 200\} = \Phi\left(\frac{40}{\sigma}\right)-\Phi\left(\frac{-40}{\sigma}\right)=2\Phi\left(\frac{40}{\sigma}\right)-1\geqslant 0.80,$$

即要求
$$\Phi\left(\frac{40}{\sigma}\right)\geqslant 0.9=\Phi(1.282),$$

应有
$$\frac{40}{\sigma}\geqslant 1.282, \quad \sigma\leqslant\frac{40}{1.282}=31.20,$$

即允许 σ 最大为 31.20.

30. 设在一电路中，电阻两端的电压(以 V 计)服从 $N(120,2^2)$ 分布，今独立测量了 5 次，试确定有 2 次测定值落在区间[118,122]之外的概率.

解 设第 i 次测定值为 $X_i, i=1,2,3,4,5$，则 $X_i\sim N(120,4)$.

$$P\{118\leqslant X_i\leqslant 122\}=\Phi\left(\frac{122-120}{2}\right)-\Phi\left(\frac{118-120}{2}\right)$$
$$=\Phi(1)-\Phi(-1)=2\Phi(1)-1=0.682\ 6.$$
$$P\{X_i\notin[118,122]\}=1-P\{118\leqslant X\leqslant 122\}=0.317\ 4,\quad i=1,2,3,4,5.$$

因诸 X_i 相互独立，故若以 Y 表示 5 次测量其测定值 X_i 落在[118,120]之外的个数，则 $Y\sim b(5,0.317\ 4)$，故所求概率为

$$P\{Y=2\}=\binom{5}{2}\times 0.317\ 4^2\times 0.682\ 6^3=0.320\ 4.$$

31. 某人上班，自家里去办公楼要经过一交通指示灯，这一指示灯有 80%时间亮红灯，此时他在指示灯旁等待直至绿灯亮. 等待时间在区间[0,30](以 s 计)上服从均匀分布. 以 X 表示他的等待时间，求 X 的分布函数 $F(x)$. 画出 $F(x)$ 的图形，并问 X 是否为连续型随机变量，是否为离散型的？(要说明理由.)

解　当他到达交通指示灯处时，若是亮绿灯，则等待时间 X 为零，亮红灯则等待时间 X 服从均匀分布. 记 A 为事件"指示灯亮绿灯"，对于固定的 $x\geqslant 0$，由全概率公式有

$$P\{X\leqslant x\}=P\{X\leqslant x|A\}P(A)+P\{X\leqslant x|\overline{A}\}P(\overline{A}),$$

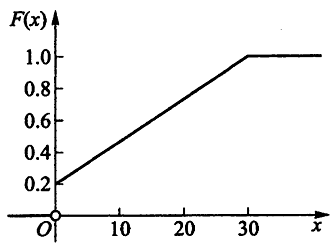
其中 $P\{X\leqslant x\mid A\}=1$，$P\{X\leqslant x\mid \overline{A}\}=\frac{x}{30}$ (当 $0\leqslant x\leqslant 30$)，$P\{X\leqslant x\mid \overline{A}\}=1$ (当 $x>30$). 由 $P(A)=0.2$ 得到

$$P\{X\leqslant x\}=1\times 0.2+\frac{x}{30}\times 0.8=0.2+\frac{0.8x}{30}\ (\text{当 } 0\leqslant x\leqslant 30);$$
$$P\{X\leqslant x\}=1\times 0.2+1\times 0.8=1\ (\text{当 } x>30).$$

于是得 X 的分布函数为

$$F(x)=P\{X\leqslant x\}=\begin{cases}0, & x<0,\\ 0.2+\dfrac{0.8x}{30}, & 0\leqslant x<30,\\ 1, & x\geqslant 30.\end{cases}$$

由题 2.31 图可看出，因 $F(x)$ 在 $x=0$ 处有不连续点，故随机变量 X 不是连续型的，又因不存在一个可列的点集，使得在这个点集上 X 取值的概率为 1，故随机变量 X 也不是离散型的，X 是混合型随机变量.



题 2.31 图

32. 设 $f(x)$，$g(x)$都是概率密度函数，求证

$$h(x)=\alpha f(x)+(1-\alpha)g(x),\quad 0\leqslant\alpha\leqslant 1$$

也是一个概率密度函数.

证 因 f,g 都是概率密度函数，故有

$$f(x)\geqslant 0,\quad g(x)\geqslant 0, \qquad (*_1)$$

且
$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1,\quad \int_{-\infty}^{\infty} g(x)\mathrm{d}x = 1. \qquad (*_2)$$

现在 $0\leqslant\alpha\leqslant 1$，故 $1-\alpha\geqslant 0$，由$(*_1)$有

$$\alpha f(x)\geqslant 0,\quad (1-\alpha)g(x)\geqslant 0,$$

于是
$$h(x)\geqslant 0.$$

又由$(*_2)$知

$$\int_{-\infty}^{\infty} h(x)\mathrm{d}x = \alpha\int_{-\infty}^{\infty} f(x)\mathrm{d}x + (1-\alpha)\int_{-\infty}^{\infty} g(x)\mathrm{d}x = \alpha + (1-\alpha) = 1,$$

所以 $h(x)$ 是一个概率密度函数.

33. 设随机变量 X 的分布律为

X	-2	-1	0	1	3
p_k	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{15}$	$\frac{11}{30}$

求 $Y=X^2$ 的分布律.

解 $Y=X^2$ 所有可能取值为 0,1,4,9.

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

$$P\{Y=1\}=P\{X^2=1\}=P\{(X=1)\cup(X=-1)\}$$
$$=P\{X=1\}+P\{X=-1\}=\frac{1}{15}+\frac{1}{6}=\frac{7}{30},$$

$$P\{Y=4\}=P\{X^2=4\}=P\{(X=2)\cup(X=-2)\}$$
$$=P\{X=2\}+P\{X=-2\}=0+\frac{1}{5}=\frac{1}{5},$$

$$P\{Y=9\}=P\{X^2=9\}=P\{(X=3)\cup(X=-3)\}$$
$$=P\{X=3\}+P\{X=-3\}=\frac{11}{30}+0=\frac{11}{30},$$

故 Y 的分布律为

Y	0	1	4	9
p_k	$\frac{1}{5}$	$\frac{7}{30}$	$\frac{1}{5}$	$\frac{11}{30}$

34. 设随机变量 X 在区间$(0,1)$内服从均匀分布.

(1) 求 $Y=e^X$ 的概率密度.

(2) 求 $Y=-2\ln X$ 的概率密度.

解　X 的概率密度为

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

分别记 X,Y 的分布函数为 $F_X(x),F_Y(y)$.

(1) 先来求 Y 的分布函数 $F_Y(y)$,因 $Y=e^X>0$,故当 $y\leqslant 0$ 时,$F_Y(y)=P\{Y\leqslant y\}=0$,从而 $f_Y(y)=0$. 当 $y>0$ 时,

$$\begin{aligned}F_Y(y)&=P\{Y\leqslant y\}=P\{e^X\leqslant y\}\\&=P\{X\leqslant \ln y\}=F_X(\ln y).\end{aligned}$$

将上式关于 y 求导,得

$$f_Y(y)=f_X(\ln y)\cdot\frac{1}{y}=\begin{cases}1\cdot\dfrac{1}{y}, & 0<\ln y<1,\\ 0, & \ln y<0\text{ 或 }\ln y>1\end{cases}$$

$$=\begin{cases}\dfrac{1}{y}, & 1<y<e,\\ 0, & 0<y<1\text{ 或 }y>e,\end{cases}$$

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

(2) 先来求 $F_Y(y)$. 当 X 在(0,1)内取值时 $Y>0$,故当 $y\leqslant 0$ 时,$F_Y(y)=0$,从而 $f_Y(y)=0$. 当 $y>0$ 时,

$$\begin{aligned}F_Y(y)&=P\{Y\leqslant y\}=P\{-2\ln X\leqslant y\}=P\{X\geqslant e^{-y/2}\}\\&=1-P\{X<e^{-y/2}\}=1-F_X(e^{-y/2}),\end{aligned}$$

于是

$$f_Y(y)=-f_X(e^{-y/2})\left(-\frac{1}{2}e^{-y/2}\right)=\frac{1}{2}e^{-y/2}.$$

故有

$$f_Y(y)=\begin{cases}\dfrac{1}{2}e^{-y/2}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

本题也可以利用教材第二章 §5 定理的结果直接解得:

(1) $Y=e^X$,即有 $y=g(x)=e^x$,在区间(0,1)上恒有 $g'(x)=e^x>0$,因此 $g(x)$严格单调增加,且 $g(x)$具有反函数 $x=h(y)=\ln y$,又 $h'(y)=\dfrac{1}{y}$,$g(0)=$

$1,g(1)=\mathrm{e}$，由教材第二章(5.2)式，得 $Y=\mathrm{e}^X$ 的概率密度为

$$f_Y(y)=\begin{cases} f_X(\ln y)\left|\dfrac{1}{y}\right|, & 1<y<\mathrm{e}, \\ 0, & \text{其他} \end{cases}=\begin{cases} \dfrac{1}{y}, & 1<y<\mathrm{e}, \\ 0, & \text{其他}. \end{cases}$$

(2) $Y=-2\ln X$，即有 $y=g(x)=-2\ln x$，在区间 $(0,1)$ 上恒有 $g'(x)=-\dfrac{2}{x}<0$，且 $g(x)$ 具有反函数 $x=h(y)=\mathrm{e}^{-y/2}$，又 $h'(y)=-\dfrac{1}{2}\mathrm{e}^{-y/2}$，$g(0)=\infty$，$g(1)=0$. 由教材第二章(5.2)式得 $Y=-2\ln X$ 的概率密度为

$$f_Y(y)=\begin{cases} f_X(\mathrm{e}^{-y/2})\left|-\dfrac{1}{2}\mathrm{e}^{-y/2}\right|, & 0<y<\infty, \\ 0, & \text{其他}, \end{cases}$$

即

$$f_Y(y)=\begin{cases} \dfrac{1}{2}\mathrm{e}^{-y/2}, & y>0, \\ 0, & \text{其他}. \end{cases}$$

注：利用教材第二章(5.2)式直接写出 $Y=g(X)$ 的概率密度时，应注意两点：

(i) 首先要检验函数 $y=g(x)$ 是否是严格单调的(一般可利用 $g'(x)$ 来检验)，如果不是严格单调的，那就不能采用这一公式.

(ii) 在公式中的 $h'(y)$ 要取绝对值，如忘记取绝对值，就可能犯严重的错误，例如在本题(2)中若忘记了将 $h'(y)$ 取绝对值，那么所得的概率密度就成为

$$\begin{cases} -\dfrac{1}{2}\mathrm{e}^{-y/2}, & y>0, \\ 0, & \text{其他}, \end{cases}$$

就有 $f_Y(y)\leqslant 0$ 了.

35. 设随机变量 $X\sim N(0,1)$.

(1) 求 $Y=\mathrm{e}^X$ 的概率密度.

(2) 求 $Y=2X^2+1$ 的概率密度.

(3) 求 $Y=|X|$ 的概率密度.

解 (1) 因为 $Y=\mathrm{e}^X$，故 Y 不取负值. 从而，若 $y<0$，则 $f_Y(y)=0$；若 $y>0$，注意到 $X\sim N(0,1)$，故 Y 的分布函数为

$$\begin{aligned} F_Y(y) &= P\{Y\leqslant y\}=P\{0<Y\leqslant y\}=P\{0<\mathrm{e}^X\leqslant y\} \\ &= P\{-\infty<X\leqslant \ln y\}=\Phi(\ln y). \end{aligned}$$

从而，$y>0$ 时，

$$f_Y(y)=\frac{\mathrm{d}}{\mathrm{d}y}F_Y(y)=\frac{\mathrm{d}}{\mathrm{d}x}\Phi(x)\Big|_{x=\ln y}\cdot\frac{1}{y}=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(\ln y)^2}\cdot\frac{1}{y}.$$

于是，$Y=\mathrm{e}^X$ 的概率密度为

$$f_Y(y)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\,y}e^{-(\ln y)^2/2}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

(2) 因 $Y=2X^2+1$，故 Y 在 $[1,\infty)$ 取值，从而 $y<1$ 时 $f_Y(y)=0$；若 $y\geqslant 1$，注意到 $X\sim N(0,1)$，故 Y 的分布函数为

$$\begin{aligned}F_Y(y)&=P\{Y\leqslant y\}=P\{2X^2+1\leqslant y\}\\&=P\left\{-\sqrt{\frac{y-1}{2}}\leqslant X\leqslant\sqrt{\frac{y-1}{2}}\right\}\\&=\Phi\left(\sqrt{\frac{y-1}{2}}\right)-\Phi\left(-\sqrt{\frac{y-1}{2}}\right)=2\Phi\left(\sqrt{\frac{y-1}{2}}\right)-1.\end{aligned}$$

故 $y>1$ 时，

$$\begin{aligned}f_Y(y)&=\frac{\mathrm{d}}{\mathrm{d}y}\left[2\Phi\left(\sqrt{\frac{y-1}{2}}\right)-1\right]\\&=2\cdot\frac{1}{\sqrt{2\pi}}e^{-(y-1)/4}\cdot\frac{1}{2}\cdot\frac{1}{\sqrt{2(y-1)}}\\&=\frac{1}{2\sqrt{\pi(y-1)}}e^{-(y-1)/4}.\end{aligned}$$

于是 $Y=2X^2+1$ 的概率密度为

$$f_Y(y)=\begin{cases}\dfrac{1}{2\sqrt{\pi(y-1)}}e^{-(y-1)/4}, & y>1,\\ 0, & \text{其他}.\end{cases}$$

(3) 对于 $Y=|X|$，显然，当 $y<0$ 时，$f_Y(y)=0$；当 $y\geqslant 0$ 时，注意到 $X\sim N(0,1)$，就有

$$\begin{aligned}F_Y(y)&=P\{0\leqslant Y\leqslant y\}=P\{|X|\leqslant y\}=P\{-y\leqslant X\leqslant y\}\\&=\Phi(y)-\Phi(-y)=2\Phi(y)-1.\end{aligned}$$

因此，$y>0$ 时，

$$f_Y(y)=\frac{\mathrm{d}}{\mathrm{d}y}F_Y(y)=\frac{\mathrm{d}}{\mathrm{d}y}[2\Phi(y)-1]=\frac{2}{\sqrt{2\pi}}e^{-y^2/2}.$$

故 $Y=|X|$ 的概率密度为

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

36. (1) 设随机变量 X 的概率密度为 $f(x)$，$-\infty<x<\infty$. 求 $Y=X^3$ 的概

率密度.

(2) 设随机变量 X 的概率密度为

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

求 $Y=X^2$ 的概率密度.

解 (1) $Y=X^3$，即有 $y=g(x)=x^3$，它严格单调增加，解得 $x=h(y)=y^{1/3}$，且有 $h'(y)=\frac{1}{3}y^{-2/3}$，由教材第二章(5.2)式得 $Y=X^3$ 的概率密度为

$$f_Y(y)=\frac{1}{3}y^{-2/3}f(y^{1/3}),\quad y\neq 0.$$

(2) $Y=X^2$，即有 $y=g(x)=x^2$，在 $x>0$ 时，$g(x)$ 严格单调增加，具有反函数 $x=h(y)=\sqrt{y}$，又有 $h'(y)=\frac{1}{2}y^{-1/2}$，由教材第二章(5.2)式得 $Y=X^2$ 的概率密度为

$$f_Y(y)=\begin{cases}\dfrac{1}{2\sqrt{y}}e^{-\sqrt{y}}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

37. 设随机变量 X 的概率密度为

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

求 $Y=\sin X$ 的概率密度.

解 X 在 $(0,\pi)$ 内取值时 $Y=\sin X$ 在 $(0,1)$ 内取值，故若 $y<0$ 或 $y>1$ 时 $f_Y(y)=0$. 若 $0\leqslant y\leqslant 1$，则 Y 的分布函数为

$$\begin{aligned}F_Y(y)&=P\{Y\leqslant y\}=P\{0\leqslant Y\leqslant y\}=P\{0\leqslant \sin X\leqslant y\}\\&=P\{(0\leqslant X\leqslant \arcsin y)\cup(\pi-\arcsin y\leqslant X\leqslant\pi)\}\\&=P\{0\leqslant X\leqslant\arcsin y\}+P\{\pi-\arcsin y\leqslant X\leqslant \pi\}\\&=\int_0^{\arcsin y}\frac{2x}{\pi^2}dx+\int_{\pi-\arcsin y}^{\pi}\frac{2x}{\pi^2}dx\\&=\frac{1}{\pi^2}(\arcsin y)^2+1-\frac{1}{\pi^2}(\pi-\arcsin y)^2=\frac{2}{\pi}\arcsin y.\end{aligned}$$

所以当 $0<y<1$ 时，

$$f_Y(y)=\frac{\mathrm{d}}{\mathrm{d}y}F_Y(y)=\frac{2}{\pi\sqrt{1-y^2}}.$$

因此，所求的概率密度为

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

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

解 电流 I 的概率密度为

$$f_I(i)=\begin{cases}\dfrac{1}{2}, & 9<i<11,\\ 0, & \text{其他}.\end{cases}$$

$W=2I^2$，即有 $w=g(i)=2i^2$，在 $i>0$ 时，$g(i)$严格单调增加，且有反函数 $i=h(w)=\left(\dfrac{w}{2}\right)^{\frac{1}{2}}$，而 $h'(w)=\dfrac{1}{2\sqrt{2}}w^{-\frac{1}{2}}$，$g(9)=162$，$g(11)=242$. 由教材第二章(5.2)式得 $W=2I^2$ 的概率密度为

$$f_W(w)=\begin{cases}\dfrac{1}{2}\left(\dfrac{1}{2\sqrt{2}}w^{-\frac{1}{2}}\right), & 162<w<242,\\ 0, & \text{其他},\end{cases}$$

即

$$f_W(w)=\begin{cases}\dfrac{1}{4}\dfrac{1}{\sqrt{2w}}, & 162<w<242,\\ 0, & \text{其他}.\end{cases}$$

39. 某物体的温度 T(以°F 计)是随机变量，且有 $T\sim N(98.6,2)$，已知 $\Theta=\dfrac{5}{9}(T-32)$，试求 Θ(以℃计) 的概率密度.

解 T 的概率密度为

$$f_T(t)=\frac{1}{2\sqrt{\pi}}\mathrm{e}^{-\frac{(t-98.6)^2}{4}}, \quad -\infty<t<\infty.$$

将 Θ 的分布函数记为 $F_\Theta(y)$，则有

$$F_\Theta(y)=P\{\Theta\leqslant y\}=P\left\{\frac{5}{9}(T-32)\leqslant y\right\}$$

$$=P\left\{T\leqslant\frac{9}{5}y+32\right\}=\int_{-\infty}^{\frac{9}{5}y+32}f_T(t)\,\mathrm{d}t.$$

将上式关于 y 求导得到 Θ 的概率密度为

$$f_\Theta(y)=f_T\left(\frac{9}{5}y+32\right)\times\frac{9}{5}=\frac{1}{2\sqrt{\pi}}\times\frac{9}{5}e^{-\frac{1}{4}\left(\frac{9}{5}y+32-98.6\right)^2},$$

即有

$$f_\Theta(y)=\frac{9}{10\sqrt{\pi}}e^{-\frac{81(y-37)^2}{100}}.$$

第三章　多维随机变量及其分布

1. 在一箱子中装有12只开关，其中2只是次品，在其中取两次，每次任取一只，考虑两种试验：(1) 放回抽样；(2) 不放回抽样. 我们定义随机变量 X,Y 如下：

$$X=\begin{cases}0, & \text{若第一次取出的是正品,}\\ 1, & \text{若第一次取出的是次品;}\end{cases}$$

$$Y=\begin{cases}0, & \text{若第二次取出的是正品,}\\ 1, & \text{若第二次取出的是次品.}\end{cases}$$

试分别就(1)，(2) 两种情况，写出 X 和 Y 的联合分布律.

解　(1) 放回抽样. 由教材第一章知第一次、第二次取到正品（或次品）的概率相同，且两次所得的结果相互独立，即有

$$P\{X=0\}=P\{Y=0\}=\frac{5}{6},$$

$$P\{X=1\}=P\{Y=1\}=\frac{1}{6},$$

且 $P\{X=i,Y=j\}=P\{X=i\}P\{Y=j\},i,j=0,1$，于是得 X 和 Y 的联合分布律为

$$P\{X=0,Y=0\}=P\{X=0\}P\{Y=0\}=\frac{25}{36},$$

$$P\{X=0,Y=1\}=P\{X=0\}P\{Y=1\}=\frac{5}{36},$$

$$P\{X=1,Y=0\}=P\{X=1\}P\{Y=0\}=\frac{5}{36},$$

$$P\{X=1,Y=1\}=P\{X=1\}P\{Y=1\}=\frac{1}{36}.$$

(2) 不放回抽样. 由乘法公式

$$P\{X=i,Y=j\}=P\{Y=j\mid X=i\}P\{X=i\},\quad i,j=0,1,$$

知 X 和 Y 的联合分布律为

$$P\{X=0,Y=0\}=\frac{9}{11}\times\frac{10}{12}=\frac{45}{66},$$

$$P\{X=0,Y=1\}=\frac{2}{11}\times\frac{10}{12}=\frac{10}{66},$$

$$P\{X=1,Y=0\}=\frac{10}{11}\times\frac{2}{12}=\frac{10}{66},$$

$$P\{X=1,Y=1\}=\frac{1}{11}\times\frac{2}{12}=\frac{1}{66}.$$

(1),(2) 两种情况下的 X 和 Y 的联合分布律的表格形式分别为

Y \ X	0	1
0	$\frac{25}{36}$	$\frac{5}{36}$
1	$\frac{5}{36}$	$\frac{1}{36}$

Y \ X	0	1
0	$\frac{45}{66}$	$\frac{10}{66}$
1	$\frac{10}{66}$	$\frac{1}{66}$

2. (1) 盒子里装有 3 只黑球、2 只红球、2 只白球，在其中任取4 只球. 以 X 表示取到黑球的只数，以 Y 表示取到红球的只数. 求 X 和 Y 的联合分布律.

(2) 在(1) 中求 $P\{X>Y\},P\{Y=2X\},P\{X+Y=3\},P\{X<3-Y\}$.

解 (1) 按古典概型计算. 自 7 只球中取 4 只，共有 $\binom{7}{4}=35$ 种取法. 在 4 只球中，黑球有 i 只，红球有 j 只(剩下 $4-i-j$ 只为白球) 的取法数为

$$N\{X=i,Y=j\}=\binom{3}{i}\binom{2}{j}\binom{2}{4-i-j},$$

$$i=0,1,2,3;j=0,1,2;i+j\leqslant 4.$$

于是

$$P\{X=0,Y=2\}=\binom{3}{0}\binom{2}{2}\binom{2}{2}\Big/35=\frac{1}{35}.$$

$$P\{X=1,Y=1\}=\binom{3}{1}\binom{2}{1}\binom{2}{2}\Big/35=\frac{6}{35}.$$

$$P\{X=1,Y=2\}=\binom{3}{1}\binom{2}{2}\binom{2}{1}\Big/35=\frac{6}{35}.$$

$$P\{X=2,Y=0\}=\binom{3}{2}\binom{2}{0}\binom{2}{2}\Big/35=\frac{3}{35}.$$

$$P\{X=2,Y=1\}=\binom{3}{2}\binom{2}{1}\binom{2}{1}\Big/35=\frac{12}{35}.$$

$$P\{X=2,Y=2\}=\binom{3}{2}\binom{2}{2}\binom{2}{0}\Big/35=\frac{3}{35}.$$

$$P\{X=3,Y=0\}=\binom{3}{3}\binom{2}{0}\binom{2}{1}\Big/35=\frac{2}{35}.$$

$$P\{X=3,Y=1\}=\binom{3}{3}\binom{2}{1}\binom{2}{0}\Big/35=\frac{2}{35}.$$

$$P\{X=0,Y=0\}=P\{X=0,Y=1\}=P\{X=1,Y=0\}$$
$$=P\{X=3,Y=2\}=0.$$

分布律为

Y \ X	0	1	2	3
0	0	0	$\frac{3}{35}$	$\frac{2}{35}$
1	0	$\frac{6}{35}$	$\frac{12}{35}$	$\frac{2}{35}$
2	$\frac{1}{35}$	$\frac{6}{35}$	$\frac{3}{35}$	0

(2) $P\{X>Y\}=P\{X=2,Y=0\}+P\{X=2,Y=1\}$
$+P\{X=3,Y=0\}+P\{X=3,Y=1\}$
$=\frac{3}{35}+\frac{12}{35}+\frac{2}{35}+\frac{2}{35}=\frac{19}{35}.$

$P\{Y=2X\}=P\{X=1,Y=2\}=\frac{6}{35}.$

$P\{X+Y=3\}=P\{X=1,Y=2\}+P\{X=2,Y=1\}+P\{X=3,Y=0\}$
$=\frac{6}{35}+\frac{12}{35}+\frac{2}{35}=\frac{20}{35}.$

$P\{X<3-Y\}=P\{X+Y<3\}$
$=P\{X=0,Y=2\}+P\{X=1,Y=1\}+P\{X=2,Y=0\}$
$=\frac{1}{35}+\frac{6}{35}+\frac{3}{35}=\frac{10}{35}.$

3. 设随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}k(6-x-y), & 0<x<2,2<y<4,\\ 0, & \text{其他}.\end{cases}$$

(1) 确定常数 k.

(2) 求 $P\{X<1,Y<3\}$.

(3) 求 $P\{X<1.5\}$.

(4) 求 $P\{X+Y\leqslant 4\}$.

解　(1) 由$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x,y)\mathrm{d}x\mathrm{d}y=1$,得

$$1=\int_2^4\mathrm{d}y\int_0^2 k(6-x-y)\mathrm{d}x=k\int_2^4\left[(6-y)x-\frac{1}{2}x^2\right]\Big|_{x=0}^{x=2}\mathrm{d}y$$

$$=k\int_2^4(12-2y-2)\mathrm{d}y=k(10y-y^2)\Big|_2^4=8k,$$

所以 $k=\frac{1}{8}$.

(2) $P\{X<1,Y<3\}=\int_2^3\mathrm{d}y\int_0^1\frac{1}{8}(6-x-y)\mathrm{d}x$

$$=\frac{1}{8}\int_2^3\left[(6-y)x-\frac{1}{2}x^2\right]\Big|_{x=0}^{x=1}\mathrm{d}y$$

$$=\frac{1}{8}\int_2^3\left(\frac{11}{2}-y\right)\mathrm{d}y=\frac{3}{8}.$$

(3) $P\{X<1.5\}=\int_2^4\mathrm{d}y\int_0^{1.5}\frac{1}{8}(6-x-y)\mathrm{d}x$

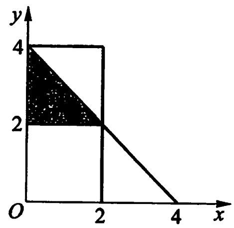
$$=\frac{1}{8}\int_2^4\left[(6-y)x-\frac{1}{2}x^2\right]\Big|_{x=0}^{x=1.5}\mathrm{d}y$$

$$=\frac{1}{8}\int_2^4\left(\frac{63}{8}-\frac{3}{2}y\right)\mathrm{d}y=\frac{27}{32}.$$

(4) 在 $f(x,y)\neq 0$ 的区域 $R:0\leqslant x\leqslant 2,2\leqslant y\leqslant 4$ 上作直线 $x+y=4$（如题 3.3 图），并记

$$G:\{(x,y)\mid 0\leqslant x\leqslant 2,\ 2\leqslant y\leqslant 4-x\},$$

则



题 3.3 图

$$P\{X+Y\leqslant 4\}=P\{(X,Y)\in G\}$$

$$=\iint_G f(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\int_2^4\mathrm{d}y\int_0^{4-y}\frac{1}{8}(6-x-y)\mathrm{d}x$$

$$=\frac{1}{8}\int_2^4\left[(6-y)x-\frac{1}{2}x^2\right]\Big|_{x=0}^{x=4-y}\mathrm{d}y$$

$$=\frac{1}{8}\int_2^4\left[(6-y)(4-y)-\frac{1}{2}(4-y)^2\right]\mathrm{d}y$$

$$=\frac{1}{8}\int_2^4\left[2(4-y)+\frac{1}{2}(4-y)^2\right]\mathrm{d}y$$

$$=\frac{1}{8}\left[-(4-y)^2-\frac{1}{6}(4-y)^3\right]\Big|_2^4=\frac{2}{3}.$$

4. 设 X,Y 都是非负的连续型随机变量，它们相互独立.

(1) 证明 $$P\{X<Y\}=\int_0^{\infty}F_X(x)f_Y(x)\mathrm{d}x,$$

其中 $F_X(x)$ 是 X 的分布函数，$f_Y(y)$ 是 Y 的概率密度.

(2) 设 X,Y 相互独立，其概率密度分别为

$$f_X(x)=\begin{cases}\lambda_1\mathrm{e}^{-\lambda_1 x}, & x>0,\\0, & \text{其他},\end{cases}\quad f_Y(y)=\begin{cases}\lambda_2\mathrm{e}^{-\lambda_2 y}, & y>0,\\0, & \text{其他},\end{cases}$$

求 $P\{X<Y\}$.

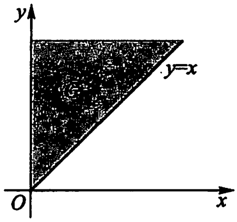
解 (1) 因 X,Y 为非负的相互独立的随机变量，故其概率密度为

$$f(x,y)=\begin{cases}f_X(x)f_Y(y), & x>0,y>0,\\0, & \text{其他}.\end{cases}$$

从而

$$P\{X<Y\}=\iint\limits_G f_X(x)f_Y(y)\mathrm{d}x\mathrm{d}y,$$

其中 G 为 $x\geqslant 0,y\geqslant x$ 界定的区域（如题 3.4 图），从而



题 3.4 图

$$\begin{aligned}P\{X<Y\}&=\int_0^{\infty}\int_0^{y}f_X(x)f_Y(y)\mathrm{d}x\mathrm{d}y\\&=\int_0^{\infty}f_Y(y)\left[\int_0^{y}f_X(x)\mathrm{d}x\right]\mathrm{d}y\\&=\int_0^{\infty}f_Y(y)F_X(y)\mathrm{d}y\\&=\int_0^{\infty}F_X(y)f_Y(y)\mathrm{d}y\\&=\int_0^{\infty}F_X(x)f_Y(x)\mathrm{d}x.\end{aligned}$$

(2) 由(1)

$$\begin{aligned}P\{X<Y\}&=\int_0^{\infty}(1-\mathrm{e}^{-\lambda_1 x})(\lambda_2\mathrm{e}^{-\lambda_2 x})\mathrm{d}x=\int_0^{\infty}[\lambda_2\mathrm{e}^{-\lambda_2 x}-\lambda_2\mathrm{e}^{-(\lambda_1+\lambda_2)x}]\mathrm{d}x\\&=\left[-\mathrm{e}^{-\lambda_2 x}+\frac{\lambda_2}{\lambda_1+\lambda_2}\mathrm{e}^{-(\lambda_1+\lambda_2)x}\right]\Bigg|_0^{\infty}=1-\frac{\lambda_2}{\lambda_1+\lambda_2}=\frac{\lambda_1}{\lambda_1+\lambda_2}.\end{aligned}$$

5. 设随机变量 (X,Y) 具有分布函数

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

求边缘分布函数.

解 $$F_X(x)=F(x,\infty)=\begin{cases}1-\mathrm{e}^{-x}, & x>0,\\0, & \text{其他}.\end{cases}$$

$$F_Y(y) = F(\infty, y) = \begin{cases} 1 - e^{-y}, & y > 0, \\ 0, & \text{其他}. \end{cases}$$

6. 将一枚硬币掷 3 次，以 X 表示前 2 次中出现 H 的次数，以 Y 表示 3 次中出现 H 的次数. 求 X,Y 的联合分布律以及 (X,Y) 的边缘分布律.

解法(i) 将试验的样本空间及 X,Y 取值的情况列表如下：

样本点	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
X 的值	2	2	1	1	1	1	0	0
Y 的值	3	2	2	2	1	1	1	0

X 所有可能取的值为 0,1,2；Y 所有可能取的值为 0,1,2,3，由于试验属等可能概型，容易得到 (X,Y) 取 (i,j)，$i=0,1,2$；$j=0,1,2,3$ 的概率. 例如

$$P\{X=1,Y=2\}=\frac{2}{8}=\frac{1}{4},\quad P\{X=2,Y=3\}=\frac{1}{8},$$

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

可得 X 和 Y 的联合分布律和 (X,Y) 的边缘分布律如下表所示.

$Y \backslash X$	0	1	2	$P\{Y=j\}$
0	$\frac{1}{8}$	0	0	$\frac{1}{8}$
1	$\frac{1}{8}$	$\frac{2}{8}$	0	$\frac{3}{8}$
2	0	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
3	0	0	$\frac{1}{8}$	$\frac{1}{8}$
$P\{X=i\}$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	1

解法(ii) $X \sim b\left(2,\frac{1}{2}\right)$，$Y$ 所有可能取的值为 0,1,2,3. 而当 $X=i$ $(i=0,1,2)$ 时，Y 取 i 的概率为 $\frac{1}{2}$，Y 取 $i+1$ 的概率也是 $\frac{1}{2}$，而取 $i,i+1$ 以外的值是不可能的（因第三次投掷不是出现 H 就是出现 T），知 $P\{X=i\}=\binom{2}{i}\frac{1}{4}$，$i=0,1,2$，故知

$$P\{X=0,Y=0\}=P\{Y=0\mid X=0\}P\{X=0\}$$
$$=\frac{1}{2}P\{X=0\}=\frac{1}{2}\times\frac{1}{4}=\frac{1}{8},$$
$$P\{X=0,Y=1\}=P\{Y=1\mid X=0\}P\{X=0\}$$
$$=\frac{1}{2}P\{X=0\}=\frac{1}{2}\times\frac{1}{4}=\frac{1}{8},$$
$$P\{X=1,Y=1\}=P\{Y=1\mid X=1\}P\{X=1\}$$
$$=\frac{1}{2}P\{X=1\}=\frac{1}{2}\times\frac{1}{2}=\frac{1}{4},$$
$$P\{X=1,Y=2\}=P\{Y=2\mid X=1\}P\{X=1\}$$
$$=\frac{1}{2}P\{X=1\}=\frac{1}{2}\times\frac{1}{2}=\frac{1}{4},$$
$$P\{X=2,Y=2\}=P\{Y=2\mid X=2\}P\{X=2\}$$
$$=\frac{1}{2}P\{X=2\}=\frac{1}{2}\times\frac{1}{4}=\frac{1}{8},$$
$$P\{X=2,Y=3\}=P\{Y=3\mid X=2\}P\{X=2\}$$
$$=\frac{1}{2}P\{X=2\}=\frac{1}{2}\times\frac{1}{4}=\frac{1}{8},$$
$$P\{X=0,Y=2\}=P\{X=0,\ Y=3\}$$
$$=P\{X=1,Y=0\}=P\{X=1,Y=3\}$$
$$=P\{X=2,Y=0\}=P\{X=2,Y=1\}=0.$$

所得 X 和 Y 的联合分布律与解法(i) 相同，即为上表所示.

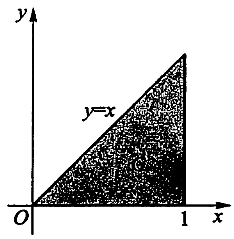
7. 设二维随机变量(X,Y) 的概率密度为

$$f(x,y)=\begin{cases}4.8y(2-x), & 0\leqslant x\leqslant 1,0\leqslant y\leqslant x,\\ 0, & \text{其他}.\end{cases}$$

求边缘概率密度.

解 (X,Y) 的概率密度 $f(x,y)$ 在区域 $G:\{(x,y)\mid 0\leqslant x\leqslant 1,0\leqslant y\leqslant x\}$ 外取零值. 如题 3.7 图有

$$f_X(x)=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}y$$
$$=\begin{cases}\int_0^x 4.8y(2-x)\mathrm{d}y, & 0\leqslant x\leqslant 1,\\ 0, & \text{其他}\end{cases}$$
$$=\begin{cases}2.4(2-x)x^2, & 0\leqslant x\leqslant 1,\\ 0, & \text{其他},\end{cases}$$



题 3.7 图

$$f_Y(y)=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}x$$
$$=\begin{cases}\int_y^1 4.8y(2-x)\mathrm{d}x, & 0\leqslant y\leqslant 1,\\ 0, & \text{其他}\end{cases}$$
$$=\begin{cases}2.4y(3-4y+y^2), & 0\leqslant y\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

注：在求边缘概率密度时，需画出(X,Y)的概率密度$f(x,y)\neq 0$的区域，这对于正确写出所需求的积分的上下限是很有帮助的.

8. 设二维随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}\mathrm{e}^{-y}, & 0<x<y,\\ 0, & \text{其他},\end{cases}$$

求边缘概率密度.

解 $$f_X(x)=\begin{cases}\int_x^{\infty}\mathrm{e}^{-y}\mathrm{d}y=-\mathrm{e}^{-y}\Big|_x^{\infty}=\mathrm{e}^{-x}, & x>0,\\ 0, & \text{其他},\end{cases}$$

$$f_Y(y)=\begin{cases}\int_0^{y}\mathrm{e}^{-y}\mathrm{d}x=y\mathrm{e}^{-y}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

9. 设二维随机变量(X,Y)的概率密度为

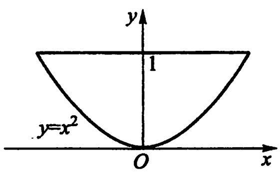
$$f(x,y)=\begin{cases}cx^2y, & x^2\leqslant y\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

(1) 确定常数c.

(2) 求边缘概率密度.

解 (1) 由于(如题3.9图)

$$1=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x,y)\mathrm{d}x\mathrm{d}y=\iint\limits_{x^2\leqslant y\leqslant 1}cx^2y\mathrm{d}x\mathrm{d}y$$
$$=c\int_{-1}^{1}x^2\mathrm{d}x\int_{x^2}^{1}y\mathrm{d}y=c\int_{-1}^{1}x^2\frac{y^2}{2}\Big|_{x^2}^{1}\mathrm{d}x$$
$$=c\int_0^1 x^2(1-x^4)\mathrm{d}x=\frac{4c}{21},$$



题3.9图

得 $c=\dfrac{21}{4}$.

(2) $$f_X(x)=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}y=\begin{cases}\int_{x^2}^{1}\frac{21}{4}x^2y\mathrm{d}y, & -1\leqslant x\leqslant 1,\\ 0, & \text{其他}\end{cases}$$

$$=\begin{cases}\dfrac{21}{8}x^2y^2\Big|_{x^2}^{1}=\dfrac{21}{8}x^2(1-x^4), & -1\leqslant x\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

$$f_Y(y)=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}x=\begin{cases}\displaystyle\int_{-\sqrt{y}}^{\sqrt{y}}\frac{21}{4}x^2y\mathrm{d}x, & 0\leqslant y\leqslant 1,\\ 0, & \text{其他},\end{cases}$$

$$=\begin{cases}\dfrac{7}{4}x^3y\Big|_{-\sqrt{y}}^{\sqrt{y}}=\dfrac{7}{2}y^{5/2}, & 0\leqslant y\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

10. 将某医药公司 8 月份和 9 月份收到的青霉素针剂的订货单数分别记为 X 和 Y. 据以往积累的资料知 X 和 Y 的联合分布律为

Y \ X	51	52	53	54	55
51	0.06	0.05	0.05	0.01	0.01
52	0.07	0.05	0.01	0.01	0.01
53	0.05	0.10	0.10	0.05	0.05
54	0.05	0.02	0.01	0.01	0.03
55	0.05	0.06	0.05	0.01	0.03

(1) 求边缘分布律.

(2) 求 8 月份的订单数为 51 时，9 月份订单数的条件分布律.

解 (1) (X,Y) 关于 X 的边缘分布律为

$$P\{X=i\}=\sum_{j=51}^{55}P\{X=i,Y=j\},\quad i=51,52,53,54,55.$$

将表中 $X=i$ 那一列的各数字相加，就得到概率 $P\{X=i\}$，例如

$$P\{X=52\}=0.05+0.05+0.10+0.02+0.06=0.28.$$

可得 (X,Y) 关于 X 的边缘分布律为

X	51	52	53	54	55
p_k	0.28	0.28	0.22	0.09	0.13

(X,Y) 关于 Y 的边缘分布律为

$$P\{Y=j\}=\sum_{i=51}^{55}P\{X=i,Y=j\},\quad j=51,52,53,54,55.$$

将表中 $Y=j$ 那一行的各数字相加，就得到概率 $P\{Y=j\}$，例如

$$P\{Y=53\}=0.05+0.10+0.10+0.05+0.05=0.35.$$

可得(X,Y)关于Y的边缘分布律为

Y	51	52	53	54	55
p_k	0.18	0.15	0.35	0.12	0.20

(2) 所需求的是条件分布律：

$$P\{Y=j\mid X=51\},\quad j=51,52,53,54,55.$$

由 $P\{Y=j\mid X=51\}=\dfrac{P\{X=51,Y=j\}}{P\{X=51\}}$ 知，只要将原表中第一列各数除以 $P\{X=51\}=0.28$，即得所求的条件分布律：

$Y=j$	51	52	53	54	55
$P\{Y=j\mid X=51\}$	$\frac{6}{28}$	$\frac{7}{28}$	$\frac{5}{28}$	$\frac{5}{28}$	$\frac{5}{28}$

11. 以X记某医院一天出生的婴儿的个数，Y记其中男婴的个数，设X和Y的联合分布律为

$$P\{X=n,Y=m\}=\frac{e^{-14}\times 7.14^m\times 6.86^{n-m}}{m!(n-m)!},$$
$$m=0,1,2,\cdots,n;\quad n=0,1,2,\cdots.$$

(1) 求边缘分布律.

(2) 求条件分布律.

(3) 特别，写出当 $X=20$ 时，Y 的条件分布律.

解 (1) $P\{X=n\}=\sum\limits_{m=0}^{n}P\{X=n,Y=m\}$

$$=\sum_{m=0}^{n}\frac{e^{-14}\times 7.14^m\times 6.86^{n-m}}{m!(n-m)!}$$

$$=\frac{e^{-14}}{n!}\sum_{m=0}^{n}\frac{n!}{m!(n-m)!}\times 7.14^m\times 6.86^{n-m}$$

$$=\frac{e^{-14}}{n!}(7.14+6.86)^n=\frac{14^n e^{-14}}{n!},\quad n=0,1,2,\cdots.$$

$$P\{Y=m\}=\sum_{n=m}^{\infty}P\{X=n,Y=m\}$$

$$= \sum_{n=m}^{\infty} \frac{e^{-14} \times 7.14^m \times 6.86^{n-m}}{m!(n-m)!}$$

$$= \frac{e^{-14}}{m!} 7.14^m \sum_{n=m}^{\infty} \frac{6.86^{n-m}}{(n-m)!} = \frac{e^{-14}}{m!} 7.14^m \sum_{k=0}^{\infty} \frac{6.86^k}{k!}$$

$$= \frac{e^{-14}}{m!} \times 7.14^m \times e^{6.86} = \frac{7.14^m e^{-7.14}}{m!}, \quad m = 0,1,2,\cdots.$$

亦即 $X \sim \pi(14), Y \sim \pi(7.14)$.

(2) 对于 $m = 0,1,2,\cdots$,

$$P\{X = n | Y = m\} = \frac{P\{X = n, Y = m\}}{P\{Y = m\}}$$

$$= \frac{e^{-14} \times 7.14^m \times 6.86^{n-m}}{m!(n-m)!} \Big/ \frac{7.14^m}{m!} e^{-7.14}$$

$$= \frac{6.86^{n-m}}{(n-m)!} e^{-6.86}, \quad n = m, m+1, \cdots.$$

对于 $n = 0,1,2,\cdots$,

$$P\{Y = m | X = n\} = \frac{P\{X = n, Y = m\}}{P\{X = n\}}$$

$$= \frac{e^{-14} \times 7.14^m \times 6.86^{n-m}}{m!(n-m)!} \Big/ \frac{14^n e^{-14}}{n!}$$

$$= \binom{n}{m} \left(\frac{7.14}{14}\right)^m \left(\frac{6.86}{14}\right)^{n-m}$$

$$= \binom{n}{m} \times 0.51^m \times 0.49^{n-m}, \quad m = 0,1,\cdots,n.$$

(3) 当 $X = 20$ 时,

$$P\left\{Y = m | X = 20\right\} = \binom{20}{m} \times 0.51^m \times 0.49^{20-m}, \quad m = 0,1,\cdots,20.$$

12. 求 §1 例 1 中的条件分布律：$P\{Y = k | X = i\}$.

解　在 §1 例 1 中，在 X 取为定值 i 之后，Y 是在 $1,2,\cdots,i$ 这 i 个数中等可能地取一个数，因此，条件分布律为

$$P\{Y = k | X = i\} = \frac{1}{i}, \quad k = 1,2,\cdots,i.$$

当 $i = 1$ 时，条件分布律为

$Y = k$	1
$P\{Y = k \mid X = 1\}$	1

当 $i=2$ 时，条件分布律为

$Y=k$	1	2
$P\{Y=k\mid X=2\}$	$\frac{1}{2}$	$\frac{1}{2}$

当 $i=3$ 时，条件分布律为

$Y=k$	1	2	3
$P\{Y=k\mid X=3\}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

当 $i=4$ 时，条件分布律为

$Y=k$	1	2	3	4
$P\{Y=k\mid X=4\}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

13. 在第 9 题中：

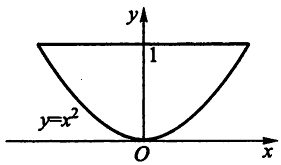
(1) 求条件概率密度 $f_{X|Y}(x\mid y)$，特别地，写出当 $Y=\frac{1}{2}$ 时 X 的条件概率密度.

(2) 求条件概率密度 $f_{Y|X}(y\mid x)$，特别地，分别写出当 $X=\frac{1}{3}$，$X=\frac{1}{2}$ 时 Y 的条件概率密度.

(3) 求条件概率

$$P\left\{Y\geqslant\frac{1}{4}\,\middle|\,X=\frac{1}{2}\right\},\quad P\left\{Y\geqslant\frac{3}{4}\,\middle|\,X=\frac{1}{2}\right\}.$$

解 在第 9 题中，有(如题 3.9 图)



题 3.9 图

$$f(x,y)=\begin{cases}\frac{21}{4}x^2y, & x^2\leqslant y\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

$$f_X(x)=\int_{x^2}^{1}\frac{21}{4}x^2y\,\mathrm{d}y=\frac{21}{8}x^2y^2\Big|_{y=x^2}^{y=1}$$

$$=\frac{21}{8}x^2(1-x^4),\quad -1\leqslant x\leqslant 1.$$

$$f_Y(y)=\int_{-\sqrt{y}}^{\sqrt{y}}\frac{21}{4}x^2y\,dx=\frac{7}{4}x^3y\Big|_{x=-\sqrt{y}}^{x=\sqrt{y}}=\frac{7}{2}y^{5/2},\quad 0\leqslant y\leqslant 1.$$

边缘概率密度为

$$f_X(x)=\begin{cases}\frac{21}{8}x^2(1-x^4), & -1\leqslant x\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

$$f_Y(y)=\begin{cases}\frac{7}{2}y^{5/2}, & 0\leqslant y\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

(1) 当 $0<y\leqslant 1$ 时，

$$f_{X|Y}(x|y)=\begin{cases}\dfrac{(21/4)x^2y}{(7/2)y^{5/2}}=\dfrac{3}{2}x^2y^{-3/2}, & -\sqrt{y}<x<\sqrt{y},\\ 0, & x\text{ 取其他值}.\end{cases}$$

当 $Y=\frac{1}{2}$ 时 X 的条件概率密度可自上式中令 $y=\frac{1}{2}$ 而得到

$$f_{X|Y}\left(x\,\Big|\,y=\frac{1}{2}\right)=\begin{cases}3\sqrt{2}x^2, & -\dfrac{\sqrt{2}}{2}<x<\dfrac{\sqrt{2}}{2},\\ 0, & x\text{ 取其他值}.\end{cases}$$

(2) 当 $-1<x<1$ 时，

$$f_{Y|X}(y|x)=\begin{cases}\dfrac{(21/4)x^2y}{(21/8)x^2(1-x^4)}=\dfrac{2y}{1-x^4}, & x^2<y<1,\\ 0, & y\text{ 取其他值}.\end{cases}$$

当 $X=\frac{1}{3}$ 时，Y 的条件概率密度为

$$f_{Y|X}\left(y\,\Big|\,x=\frac{1}{3}\right)=\begin{cases}\frac{81}{40}y, & \frac{1}{9}<y<1,\\ 0, & \text{其他}.\end{cases}$$

当 $X=\frac{1}{2}$ 时，Y 的条件概率密度为

$$f_{Y|X}\left(y\,\Big|\,x=\frac{1}{2}\right)=\begin{cases}\frac{32}{15}y, & \frac{1}{4}<y<1,\\ 0, & \text{其他}.\end{cases}$$

(3) $P\left\{Y\geqslant\frac{1}{4}\,\Big|\,X=\frac{1}{2}\right\}=\int_{\frac{1}{4}}^{1}f_{Y|X}\left(y\,\Big|\,x=\frac{1}{2}\right)dy=1.$

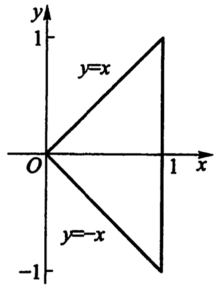
$$P\left\{Y\geqslant\frac{3}{4}\,\Big|\,X=\frac{1}{2}\right\}=\int_{\frac{3}{4}}^{1}f_{Y|X}\left(y\,\Big|\,x=\frac{1}{2}\right)dy=\frac{7}{15}.$$

14. 设随机变量(X,Y)的概率密度为

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

求条件概率密度 $f_{Y|X}(y|x)$，$f_{X|Y}(x|y)$.

解 如题 3.14 图，



题 3.14 图

$$f_X(x)=\begin{cases}\int_{-x}^{x}1\mathrm{d}y=2x, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$$

$$f_Y(y)=\begin{cases}\int_{y}^{1}1\mathrm{d}x=1-y, & 0<y<1,\\ \int_{-y}^{1}1\mathrm{d}x=1+y, & -1<y\leqslant 0,\\ 0, & \text{其他}.\end{cases}$$

当 $0<y<1$ 时，

$$f_{X|Y}(x|y)=\begin{cases}\dfrac{1}{1-y}, & y<x<1,\\ 0, & x\text{ 取其他值}.\end{cases}$$

当 $-1<y\leqslant 0$ 时，

$$f_{X|Y}(x|y)=\begin{cases}\dfrac{1}{1+y}, & -y<x<1,\\ 0, & x\text{ 取其他值}.\end{cases}$$

也可写成

$$f_Y(y)=\begin{cases}\int_{|y|}^{1}1\mathrm{d}x=1-|y|, & |y|<1,\\ 0, & \text{其他}.\end{cases}$$

因此，当 $|y|<1$ 时，

$$f_{X|Y}(x|y)=\begin{cases}\dfrac{1}{1-|y|}, & |y|<x<1,\\ 0, & x\text{ 取其他值}.\end{cases}$$

当 $0<x<1$ 时，

$$f_{Y|X}(y|x)=\begin{cases}\dfrac{1}{2x}, & |y|<x,\\ 0, & y\text{ 取其他值}.\end{cases}$$

15. 设随机变量 $X\sim U(0,1)$，当给定 $X=x$ 时，随机变量 Y 的条件概率密度为

$$f_{Y|X}(y \mid x)=\begin{cases}x, & 0<y<\dfrac{1}{x},\\ 0, & \text{其他}.\end{cases}$$

(1) 求 X 和 Y 的联合概率密度 $f(x,y)$.

(2) 求边缘概率密度 $f_Y(y)$，并画出它的图形.

(3) 求 $P\{X>Y\}$.

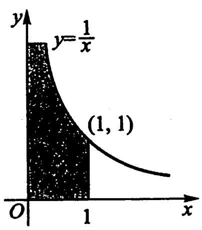
解　(1) 因 $f(x,y)=f_{Y|X}(y \mid x)f_X(x)$，今

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

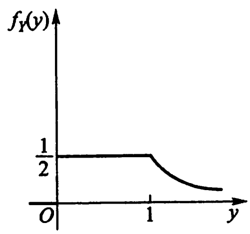
故

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

$f(x,y)$ 仅在区域 $D:\left\{(x,y)\,\middle|\,0<x<1,0<y<\dfrac{1}{x}\right\}$ 上不等于零，如题 3.15 图 1.



题 3.15 图 1



题 3.15 图 2

(2) 如题 3.15 图 2 有

$$\begin{aligned}f_Y(y)&=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}x\\ &=\begin{cases}\displaystyle\int_0^1 x\mathrm{d}x=\frac{1}{2}, & 0<y<1,\\ \displaystyle\int_0^{1/y}x\mathrm{d}x=\frac{1}{2y^2}, & 1\leqslant y<\infty,\\ 0, & \text{其他}.\end{cases}\end{aligned}$$

即

$$f_Y(y)=\begin{cases}\dfrac{1}{2}, & 0<y<1,\\ \dfrac{1}{2y^2}, & 1\leqslant y<\infty,\\ 0, & \text{其他}.\end{cases}$$

(3) $P\{X>Y\}=\iint\limits_{G}f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}y\int_y^1 x\,\mathrm{d}x=\int_0^1\frac{1}{2}(1-y^2)\mathrm{d}y=\frac{1}{3}.$

16. (1) 问第 1 题中的随机变量 X 和 Y 是否相互独立?

(2) 问第 14 题中的随机变量 X 和 Y 是否相互独立(需说明理由)?

解 (1) 在放回抽样时，X 和 Y 的联合分布律与边缘分布律如下表：

Y \ X	0	1	$P\{Y=j\}$
0	$\frac{25}{36}$	$\frac{5}{36}$	$\frac{5}{6}$
1	$\frac{5}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
$P\{X=i\}$	$\frac{5}{6}$	$\frac{1}{6}$	1

由于

$$P\{X=0,Y=0\}=\frac{25}{36}=P\{X=0\}P\{Y=0\},$$

$$P\{X=0,Y=1\}=\frac{5}{36}=P\{X=0\}P\{Y=1\},$$

$$P\{X=1,Y=0\}=\frac{5}{36}=P\{X=1\}P\{Y=0\},$$

$$P\{X=1,Y=1\}=\frac{1}{36}=P\{X=1\}P\{Y=1\},$$

故 X 与 Y 相互独立.

在不放回抽样时，X 和 Y 的联合分布律与边缘分布律如下表：

Y \ X	0	1	$P\{Y=j\}$
0	$\frac{45}{66}$	$\frac{10}{66}$	$\frac{5}{6}$
1	$\frac{10}{66}$	$\frac{1}{66}$	$\frac{1}{6}$
$P\{X=i\}$	$\frac{5}{6}$	$\frac{1}{6}$	1

由于 $P\{X=0,Y=0\}=\frac{15}{22}\neq P\{X=0\}P\{Y=0\}$，即知 X 和 Y 不是相互独立的.

(2) 在第 14 题中有(见第 14 题解答)：

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

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

而

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

在区域 $G:\left\{(x,y)\,\middle|\,|y|<x,0<x<1\right\}$ 上 $f(x,y)\neq f_X(x)f_Y(y)$，故 X 与 Y 不是相互独立的.

17. (1) 设随机变量(X,Y)具有分布函数

$$F(x,y)=\begin{cases}(1-e^{-\alpha x})y, & x\geqslant 0,0\leqslant y\leqslant 1,\\ 1-e^{-\alpha x}, & x\geqslant 0,y>1,\\ 0, & \text{其他}.\end{cases}\quad \alpha>0,$$

证明 X,Y 相互独立.

(2) 设随机变量(X,Y)具有分布律

$P\{X=x,Y=y\}=p^2(1-p)^{x+y-2}$，　$0<p<1$，x,y 均为正整数，

问 X,Y 是否相互独立？

解　(1) $F_X(x)=F(x,\infty)=\begin{cases}1-e^{-\alpha x}, & x\geqslant 0,\\ 0, & \text{其他}.\end{cases}$

$$F_Y(y)=F(\infty,y)=\begin{cases}y, & 0\leqslant y\leqslant 1,\\ 1, & y>1,\\ 0, & \text{其他}.\end{cases}$$

因为对于所有的 x,y 都有 $F(x,y)=F_X(x)F_Y(y)$，故 X,Y 相互独立.

$$(2)\ P\{X=x\}=\sum_{y=1}^{\infty}p^2(1-p)^{x+y-2}=p^2(1-p)^{x-1}\sum_{y=1}^{\infty}(1-p)^{y-1}$$
$$=p^2(1-p)^{x-1}\frac{1}{1-(1-p)}$$
$$=p(1-p)^{x-1},\quad x=1,2,\cdots,\text{其中 }0<p<1.$$

同理

$$P\{Y=y\}=p(1-p)^{y-1},\quad y=1,2,\cdots,\text{其中 }0<p<1.$$

因为对于所有正整数 x,y 都有

$$P\{X=x,Y=y\}=P\{X=x\}P\{Y=y\},$$

故 X,Y 相互独立.

18. 设 X 和 Y 是两个相互独立的随机变量，X 在区间 $(0,1)$ 上服从均匀分布，Y 的概率密度为

$$f_Y(y)=\begin{cases}\frac{1}{2}e^{-y/2}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

(1) 求 X 和 Y 的联合概率密度.

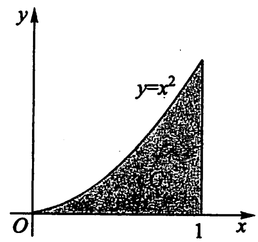
(2) 设含有 a 的二次方程为 $a^2+2Xa+Y=0$，试求 a 有实根的概率.

解 (1) 因 X 的概率密度为

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

且 X 和 Y 相互独立，故 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases}\frac{1}{2}e^{-y/2}, & 0<x<1, y>0,\\ 0, & 其他.\end{cases}$$



题 3.18 图

(2) a 的二次方程 $a^2+2Xa+Y=0$ 有实根的充要条件为判别式 $\Delta=4X^2-4Y\geqslant 0$，亦即

$$X^2\geqslant Y.$$

而

$$P\{X^2\geqslant Y\}=P\{(X,Y)\in G\},$$

其中 G 由曲线 $y=x^2, y=0, x=1$ 所围成(如题 3.18 图)，即有

$$\begin{aligned}P\{X^2\geqslant Y\}&=\iint_G f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^{x^2}\frac{1}{2}e^{-y/2}\mathrm{d}y\\&=\int_0^1[-e^{-y/2}]\Big|_0^{x^2}\mathrm{d}x=\int_0^1(1-e^{-x^2/2})\mathrm{d}x\\&=1-\int_0^1 e^{-x^2/2}\mathrm{d}x=1-\sqrt{2\pi}[\Phi(1)-\Phi(0)]\\&=1-\sqrt{2\pi}(0.8413-0.5)=0.1445.\end{aligned}$$

19. 进行打靶，设弹着点 $A(X,Y)$ 的坐标 X 和 Y 相互独立，且都服从 $N(0,1)$ 分布，规定

点 A 落在区域 $D_1=\{(x,y)\mid x^2+y^2\leqslant 1\}$ 得 2 分；

点 A 落在 $D_2=\{(x,y)\mid 1<x^2+y^2\leqslant 4\}$ 得 1 分；

点 A 落在 $D_3=\{(x,y)\mid x^2+y^2>4\}$ 得 0 分.

以 Z 记打靶的得分. 写出 X,Y 的联合概率密度，并求 Z 的分布律.

解 由题设知 X,Y 的概率密度为

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

$$f_Y(y)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-y^2/2},\quad -\infty<y<\infty,$$

且知 X 和 Y 相互独立，故 X 和 Y 的联合概率密度为

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

$$\begin{aligned}P\{(X,Y)\in D_1\}&=\iint\limits_{D_1}f(x,y)\mathrm{d}x\mathrm{d}y\\&\xlongequal{\text{用极坐标}}\int_0^{2\pi}\mathrm{d}\theta\int_0^1\frac{1}{2\pi}\mathrm{e}^{-r^2/2}r\mathrm{d}r\\&=2\pi\times\frac{1}{2\pi}\left[-\mathrm{e}^{-r^2/2}\right]\Big|_0^1=1-\mathrm{e}^{-1/2},\end{aligned}$$

$$\begin{aligned}P\{(X,Y)\in D_2\}&=\iint\limits_{D_2}f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^{2\pi}\mathrm{d}\theta\int_1^2\frac{1}{2\pi}\mathrm{e}^{-r^2/2}r\mathrm{d}r\\&=-\mathrm{e}^{-r^2/2}\Big|_1^2=\mathrm{e}^{-1/2}-\mathrm{e}^{-2},\end{aligned}$$

$$P\{(X,Y)\in D_3\}=1-(1-\mathrm{e}^{-1/2})-(\mathrm{e}^{-1/2}-\mathrm{e}^{-2})=\mathrm{e}^{-2},$$

故 Z 的分布律为

Z	0	1	2
p_k	e^{-2}	$\mathrm{e}^{-1/2}-\mathrm{e}^{-2}$	$1-\mathrm{e}^{-1/2}$

20. 设 X 和 Y 是相互独立的随机变量，其概率密度分别为

$$f_X(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x>0,\\0, & x\leqslant0,\end{cases}\qquad f_Y(y)=\begin{cases}\mu\mathrm{e}^{-\mu y}, & y>0,\\0, & y\leqslant0,\end{cases}$$

其中 $\lambda>0,\mu>0$ 是常数. 引入随机变量

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

(1) 求条件概率密度 $f_{X|Y}(x\mid y)$.

(2) 求 Z 的分布律和分布函数.

解　由于 X 和 Y 相互独立，(X,Y) 的概率密度 $f(x,y)=f_X(x)f_Y(y)$，即有

$$f(x,y)=\begin{cases}\lambda\mu\mathrm{e}^{-\lambda x-\mu y}, & x>0,y>0,\\0, & \text{其他}.\end{cases}$$

(1) 当 $y>0$ 时，有

$$f_{X|Y}(x|y)=f_X(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x>0,\\0, & \text{其他}.\end{cases}$$

(2) $P\{X \leqslant Y\} = \iint\limits_{G:x\leqslant y} f(x,y)\mathrm{d}x\mathrm{d}y = \int_0^{\infty}\mathrm{d}x\int_x^{\infty}\lambda\mu\mathrm{e}^{-\lambda x-\mu y}\mathrm{d}y$

$$= \int_0^{\infty}\left[-\lambda\mathrm{e}^{-\lambda x-\mu y}\right]\Big|_{y=x}^{y=\infty}\mathrm{d}x$$

$$= \int_0^{\infty}\lambda\mathrm{e}^{-(\lambda+\mu)x}\mathrm{d}x = \frac{\lambda}{\lambda+\mu},$$

而
$$P\{X > Y\} = 1 - P\{X \leqslant Y\} = \frac{\mu}{\lambda+\mu}.$$

故 Z 的分布律为

Z	0	1
p_k	$\frac{\mu}{\lambda+\mu}$	$\frac{\lambda}{\lambda+\mu}$

Z 的分布函数为

$$F_Z(z) = \begin{cases} 0, & z < 0, \\ \dfrac{\mu}{\lambda+\mu}, & 0 \leqslant z < 1, \\ 1, & z \geqslant 1. \end{cases}$$

21. 设随机变量(X,Y)的概率密度为

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

分别求(1) $Z = X+Y$,(2) $Z = XY$ 的概率密度.

解　记所需求的概率密度函数为 $f_Z(z)$.

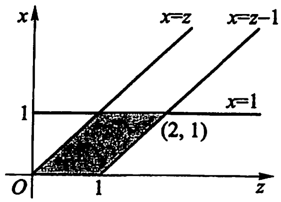
$$f(x,y) = \begin{cases} x+y, & 0 < x < 1, \quad 0 < y < 1, \\ 0, & \text{其他}. \end{cases}$$

(1) $Z = X+Y$.

$$f_Z(z) = \int_{-\infty}^{\infty} f(x,z-x)\mathrm{d}x. \tag{$*_1$}$$

仅当被积函数 $f(x,z-x) \neq 0$ 时，$f_Z(z) \neq 0$. 我们先找出使 $f(x,z-x) \neq 0$ 的 x,z 的变化范围. 从而可定出($*_1$)中积分(相对于不同 z 的值)的积分限，算出这一积分就可以了.

易知，仅当$\begin{cases} 0 < x < 1, \\ 0 < z-x < 1, \end{cases}$ 即$\begin{cases} 0 < x < 1, \\ z-1 < x < z \end{cases}$ 时，($*_1$)的被积函数不等于零，参考题 3.21 图 1，即得



题 3.21 图 1

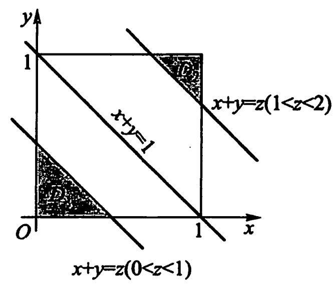
$$f_Z(z)=\int_{-\infty}^{\infty}f(x,z-x)\mathrm{d}x$$
$$=\begin{cases}\int_0^z[x+(z-x)]\mathrm{d}x, & 0<z<1,\\ \int_{z-1}^1[x+(z-x)]\mathrm{d}x, & 1\leqslant z<2,\\ 0, & \text{其他}.\end{cases}$$

即

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

本题也可利用分布函数来求 $f_Z(z)$，如下所示.

记 $Z=X+Y$ 的分布函数为 $F_Z(z)$，参考题 3.21 图 2 知



题 3.21 图 2

当 $z\leqslant 0$ 时，$F_Z(z)=0$.

当 $0<z<1$ 时，

$$F_Z(z)=P\{Z\leqslant z\}=P\{X+Y\leqslant z\}$$
$$=\iint\limits_{D_1}(x+y)\mathrm{d}x\mathrm{d}y=\int_0^z\mathrm{d}y\int_0^{z-y}(x+y)\mathrm{d}x=\frac{1}{3}z^3.$$

当 $1 \leqslant z < 2$ 时，因 $f(x,y)$ 只在矩形区域上 $\neq 0$，故

$$F_Z(z) = P\{Z \leqslant z\} = 1 - \iint_{D_2} f(x,y)\mathrm{d}x\mathrm{d}y$$

$$= 1 - \int_{z-1}^{1}\mathrm{d}y\int_{z-y}^{1}(x+y)\mathrm{d}x = -\frac{1}{3} + z^2 - \frac{1}{3}z^3.$$

当 $z \geqslant 2$ 时，$F_Z(z) = 1$.

故 $Z = X + Y$ 的分布函数为

$$F_Z(z) = \begin{cases} 0, & z \leqslant 0, \\ \dfrac{1}{3}z^3, & 0 < z < 1, \\ -\dfrac{1}{3} + z^2 - \dfrac{1}{3}z^3, & 1 \leqslant z < 2, \\ 1, & z \geqslant 2. \end{cases}$$

由此知 $Z = X + Y$ 的概率密度为

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

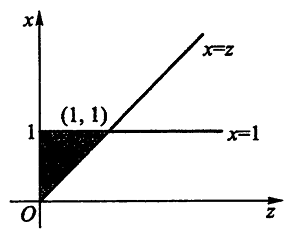
(2) $Z = XY$.

$$f_Z(z) = \int_{-\infty}^{\infty}\frac{1}{|x|}f\left(x,\frac{z}{x}\right)\mathrm{d}x.$$

易知仅当

$$\begin{cases} 0 < x < 1, \\ 0 < \dfrac{z}{x} < 1, \end{cases} \quad \text{即} \begin{cases} 0 < x < 1, \\ 0 < z < x \end{cases}$$

时，上述积分的被积函数不等于零，如题 3.21 图 3，即得



题 3.21 图 3

$$f_Z(z) = \int_{-\infty}^{\infty}\frac{1}{|x|}f\left(x,\frac{z}{x}\right)\mathrm{d}x$$

$$= \begin{cases} \displaystyle\int_z^1 \frac{1}{x}\left(x + \frac{z}{x}\right)\mathrm{d}x, & 0 < z < 1, \\ 0, & \text{其他}. \end{cases}$$

得

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

22. 设 X 和 Y 是两个相互独立的随机变量，其概率密度分别为

$$f_X(x)=\begin{cases}1, & 0\leqslant x\leqslant 1,\\ 0, & \text{其他},\end{cases}\qquad f_Y(y)=\begin{cases}e^{-y}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

求随机变量 $Z=X+Y$ 的概率密度.

解法(i)　利用公式

$$f_Z(z)=\int_{-\infty}^{\infty}f_X(z-y)f_Y(y)\mathrm{d}y,$$

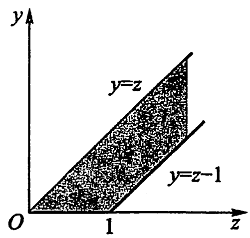
按函数 f_X, f_Y 的定义知,仅当

$$\begin{cases}0\leqslant z-y\leqslant 1,\\ y>0,\end{cases}$$

即

$$z-1\leqslant y\leqslant z,\quad y>0$$

时,上述积分的被积函数才不等于0,如题3.22图1知



题 3.22 图 1

$$f_Z(z)=\begin{cases}\displaystyle\int_0^z f_X(z-y)f_Y(y)\mathrm{d}y=\int_0^z 1\cdot e^{-y}\mathrm{d}y, & 0<z<1,\\ \displaystyle\int_{z-1}^z f_X(z-y)f_Y(y)\mathrm{d}y=\int_{z-1}^z 1\cdot e^{-y}\mathrm{d}y, & z\geqslant 1,\\ 0, & \text{其他}.\end{cases}$$

即有

$$f_Z(z)=\begin{cases}1-e^{-z}, & 0<z<1,\\ (e-1)e^{-z}, & z\geqslant 1,\\ 0, & \text{其他}.\end{cases}$$

若利用公式

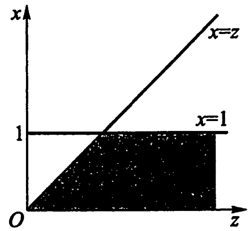
$$f_Z(z)=\int_{-\infty}^{\infty}f_X(x)f_Y(z-x)\mathrm{d}x$$

知仅当

$$\begin{cases}0\leqslant x\leqslant 1,\\ z-x>0,\end{cases}\quad \text{即}\quad \begin{cases}0\leqslant x\leqslant 1,\\ x<z\end{cases}$$

时，上述积分的被积函数才不会等于0，如题3.22图2知

$$f_Z(z)=\begin{cases}\int_0^z f_X(x)f_Y(z-x)\mathrm{d}x=\int_0^z \mathrm{e}^{-(z-x)}\mathrm{d}x, & 0<z<1,\\ \int_0^1 f_X(x)f_Y(z-x)\mathrm{d}x=\int_0^1 \mathrm{e}^{-(z-x)}\mathrm{d}x, & z\geqslant 1,\\ 0, & \text{其他}.\end{cases}$$



题3.22图2

即有

$$f_Z(z)=\begin{cases}1-\mathrm{e}^{-z}, & 0<z<1,\\ (\mathrm{e}-1)\mathrm{e}^{-z}, & z\geqslant 1,\\ 0, & \text{其他}.\end{cases}$$

解法(ii) 先求出 $Z=X+Y$ 的分布函数 $F_Z(z)$，然后将 $F_Z(z)$ 关于 z 求导从而得到 $f_Z(z)$.

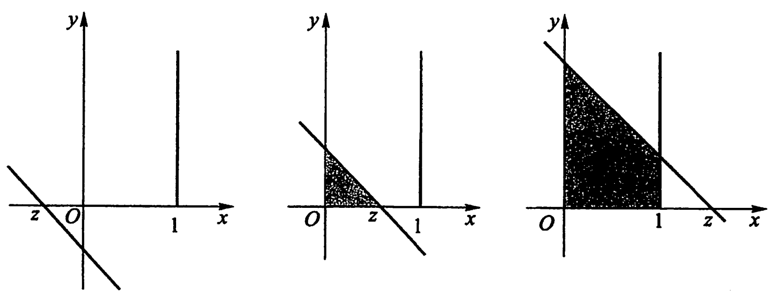
X 和 Y 的联合概率密度为

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

如果 $x+y\leqslant z$，那么 $y\leqslant z-x$，这就表明区域G:$\{(x,y)\mid x+y\leqslant z\}$ 位于直线 $x+y=z$ 的下方. 现就 z 的不同大小，画出区域 G:$\{(x,y)\mid x+y\leqslant z\}$ 与 $f(x,y)\neq 0$ 的区域 D:$\{(x,y)\mid 0\leqslant x\leqslant 1, y>0\}$ 的公共部分(有阴影的部分)如题3.22图3所示.

当 $z\leqslant 0$ 时，如题3.22图3(1)，有

$$F_Z(z)=P\{Z\leqslant z\}=P\{X+Y\leqslant z\}$$
$$=\iint\limits_{x+y\leqslant z} f(x,y)\mathrm{d}x\mathrm{d}y=\iint\limits_{x+y\leqslant z} 0\mathrm{d}x\mathrm{d}y=0,$$



(1) $x+y=z\leqslant 0$　(2) $x+y=z, 0<z<1$　(3) $x+y=z\geqslant 1$

题3.22图3

当 $0<z<1$ 时，如题 3.22 图 3(2)，有

$$F_Z(z)=P\{Z\leqslant z\}=P\{X+Y\leqslant z\}=\iint\limits_{x+y\leqslant z}f(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\int_0^z\mathrm{d}x\int_0^{z-x}\mathrm{e}^{-y}\mathrm{d}y=\int_0^z[1-\mathrm{e}^{-(z-x)}]\mathrm{d}x=z-1+\mathrm{e}^{-z},$$

当 $z\geqslant 1$ 时，如题 3.22 图 3(3)，有

$$F_Z(z)=P\{Z\leqslant z\}=P\{X+Y\leqslant z\}=\iint\limits_{x+y\leqslant z}f(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\int_0^1\mathrm{d}x\int_0^{z-x}\mathrm{e}^{-y}\mathrm{d}y=\int_0^1[1-\mathrm{e}^{-(z-x)}]\mathrm{d}x=1-\mathrm{e}^{-z+1}+\mathrm{e}^{-z},$$

得

$$F_Z(z)=\begin{cases}0, & z\leqslant 0,\\ z-1+\mathrm{e}^{-z}, & 0<z<1,\\ 1-\mathrm{e}^{-z+1}+\mathrm{e}^{-z}, & z\geqslant 1.\end{cases}$$

将 $F_Z(z)$ 关于 z 求导数，得到 Z 的概率密度为

$$f_Z(z)=\begin{cases}1-\mathrm{e}^{-z}, & 0<z<1,\\ (\mathrm{e}-1)\mathrm{e}^{-z}, & z\geqslant 1,\\ 0, & \text{其他}.\end{cases}$$

23. 某种商品一周的需求量是一个随机变量，其概率密度为

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

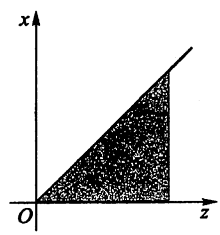
设各周的需求量是相互独立的. 求(1) 两周，(2) 三周的需求量的概率密度.

解 设某种商品在第 i 周的需求量为 $X_i(i=1,2,3)$，由题设 X_1,X_2,X_3 相互独立，并且有

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

(1) 记两周的需求量为 Z，即 $Z=X_1+X_2$，则 Z 的概率密度为

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



题 3.23 图

由 $f(t)$ 的定义，知仅当

$$\begin{cases}x>0,\\ z-x>0,\end{cases}\quad \text{亦即}\quad \begin{cases}x>0,\\ x<z\end{cases}$$

时上述积分的被积函数不等于零，于是(参见题 3.23 图)Z 的概率密度为

$$f_Z(z)=\begin{cases}\int_0^z f(x)f(z-x)\mathrm{d}x, & z>0,\\ 0, & \text{其他}\end{cases}$$

$$=\begin{cases}\int_0^z x\mathrm{e}^{-x}(z-x)\mathrm{e}^{-(z-x)}\mathrm{d}x, & z>0,\\ 0, & \text{其他}\end{cases}$$

$$=\begin{cases}\mathrm{e}^{-z}\int_0^z (xz-x^2)\mathrm{d}x=\dfrac{z^3\mathrm{e}^{-z}}{3!}, & z>0\\ 0, & \text{其他.}\end{cases}$$

(2) 记三周的需求量为 W，即 $W=Z+X_3$，因 X_1,X_2,X_3 相互独立，故 $Z=X_1+X_2$ 与 X_3 相互独立，从而 W 的概率密度为

$$f_W(u)=\int_{-\infty}^{\infty} f_Z(x)f_{X_3}(u-x)\mathrm{d}x.$$

由上述 $f_Z(z)$ 及 $f(t)$ 的定义，知仅当

$$\begin{cases}x>0,\\ u-x>0,\end{cases}\quad \text{亦即}\quad \begin{cases}x>0,\\ x<u\end{cases}$$

时，上述积分的被积函数不等于零，于是 W 的概率密度为

$$f_W(u)=\begin{cases}\int_0^u f_Z(x)f_{X_3}(u-x)\mathrm{d}x, & u>0,\\ 0, & \text{其他}\end{cases}$$

$$=\begin{cases}\int_0^u \dfrac{x^3\mathrm{e}^{-x}}{3!}(u-x)\mathrm{e}^{-(u-x)}\mathrm{d}x, & u>0,\\ 0, & \text{其他}\end{cases}$$

$$=\begin{cases}\dfrac{\mathrm{e}^{-u}}{3!}\int_0^u (x^3u-x^4)\mathrm{d}x=\dfrac{u^5\mathrm{e}^{-u}}{5!}, & u>0,\\ 0, & \text{其他.}\end{cases}$$

注：本题中我们假设第一周的需求量为 X_1，第二周的需求量为 X_2. 两周的需求量为 X_1+X_2，注意到，X_1,X_2 是相互独立的随机变量，虽然它们具有相同的分布，但它们的取值是相互独立的，因而两周的需求量不能写成 $2X$，而必须写成 X_1+X_2.

24. 设随机变量 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{2}(x+y)\mathrm{e}^{-(x+y)}, & x>0,y>0,\\ 0, & \text{其他.}\end{cases}$$

(1) 问 X 和 Y 是否相互独立？

(2) 求 $Z=X+Y$ 的概率密度.

解　(1) $f_X(x)=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}y=\int_0^{\infty}\frac{1}{2}(x+y)\mathrm{e}^{-(x+y)}\mathrm{d}y$

$$=\frac{1}{2}(x+y)\left[-\mathrm{e}^{-(x+y)}\right]\Big|_{y=0}^{y=\infty}+\frac{1}{2}\int_0^{\infty}\mathrm{e}^{-(x+y)}\mathrm{d}y$$

$$=\frac{1}{2}x\mathrm{e}^{-x}-\frac{1}{2}\mathrm{e}^{-(x+y)}\Big|_{y=0}^{y=\infty}=\frac{x+1}{2}\mathrm{e}^{-x},\quad x>0,$$

故 X 的概率密度为

$$f_X(x)=\begin{cases}\dfrac{x+1}{2}\mathrm{e}^{-x}, & x>0,\\ 0, & \text{其他}.\end{cases}$$

同理，Y 的概率密度为

$$f_Y(y)=\begin{cases}\dfrac{y+1}{2}\mathrm{e}^{-y}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

显然 $f_X(x)f_Y(y)\neq f(x,y)$，所以 X,Y 不相互独立.

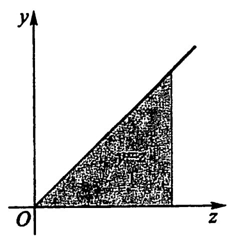
(2) 由教材第三章公式(5.1)可得 $Z=X+Y$ 的概率密度 $f_Z(z)$ 为

$$f_Z(z)=\int_{-\infty}^{\infty}f(z-y,y)\mathrm{d}y,$$

上述被积函数仅当

$$\begin{cases}z-y>0,\\ y>0,\end{cases}\quad 即\quad \begin{cases}y<z,\\ y>0\end{cases}$$

时才不会等于 0，由题 3.24 图得



题 3.24 图

$$f_Z(z)=\begin{cases}\displaystyle\int_0^z f(z-y,y)\mathrm{d}y, & z>0,\\ 0, & \text{其他}\end{cases}$$

$$=\begin{cases}\displaystyle\int_0^z\frac{1}{2}(z-y+y)\mathrm{e}^{-(z-y+y)}\mathrm{d}y, & z>0,\\ 0, & \text{其他}.\end{cases}$$

即有

$$f_Z(z)=\begin{cases}\dfrac{1}{2}\displaystyle\int_0^z z\mathrm{e}^{-z}\mathrm{d}y=\frac{1}{2}z^2\mathrm{e}^{-z}, & z>0,\\ 0, & \text{其他}.\end{cases}$$

25. 设随机变量 X,Y 相互独立，且具有相同的分布，它们的概率密度均为

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

求 $Z=X+Y$ 的概率密度.

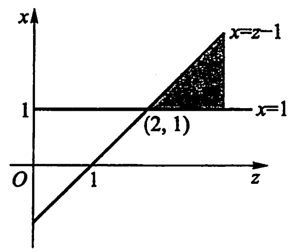
解 由卷积公式

$$f_Z(z)=\int_{-\infty}^{\infty}f_X(x)f_Y(z-x)\mathrm{d}x,$$

现在
$$f_X(x)=\begin{cases}e^{1-x}, & x>1,\\ 0, & \text{其他},\end{cases}$$

$$f_Y(y)=\begin{cases}e^{1-y}, & y>1,\\ 0, & \text{其他}.\end{cases}$$

仅当 $\begin{cases}x>1,\\ z-x>1,\end{cases}$ 即 $\begin{cases}x>1,\\ x<z-1\end{cases}$ 时，上述积分的被积函数不等于零，由题 3.25 图即得



题 3.25 图

$$f_Z(z)=\begin{cases}\displaystyle\int_1^{z-1}e^{1-x}e^{1-(z-x)}\mathrm{d}x=\int_1^{z-1}e^{2-z}\mathrm{d}x, & z>2,\\ 0, & \text{其他}.\end{cases}$$

得
$$f_Z(z)=\begin{cases}e^{2-z}(z-2), & z>2,\\ 0, & \text{其他}.\end{cases}$$

26. 设随机变量 X,Y 相互独立，它们的概率密度均为

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

求 $Z=\dfrac{Y}{X}$ 的概率密度.

解 $f_X(x)=\begin{cases}e^{-x}, & x>0,\\ 0, & \text{其他},\end{cases}$ $\quad f_Y(y)=\begin{cases}e^{-y}, & y>0,\\ 0, & \text{其他}.\end{cases}$

由公式
$$f_Z(z)=\int_{-\infty}^{\infty}|x|f_X(x)f_Y(xz)\mathrm{d}x,$$

仅当 $\begin{cases}x>0,\\ xz>0,\end{cases}$ 即 $\begin{cases}x>0,\\ z>0\end{cases}$ 时，上述积分的被积函数不等于零，于是当 $z>0$ 时有

$$f_Z(z)=\int_0^{\infty}xe^{-x}e^{-xz}\mathrm{d}x=\int_0^{\infty}xe^{-x(z+1)}\mathrm{d}x=\frac{1}{(z+1)^2}.$$

当 $z\leqslant 0$ 时 $f_Z(z)=0$，即

$$f_Z(z)=\begin{cases}\dfrac{1}{(z+1)^2}, & z>0,\\ 0, & z\leqslant 0.\end{cases}$$

27. 设随机变量 X,Y 相互独立. 它们都在区间 $(0,1)$ 上服从均匀分布. A 是以 X,Y 为边长的矩形的面积，求 A 的概率密度.

解　X,Y 的概率密度分别为

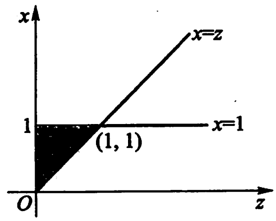
$$f_X(x)=\begin{cases}1, & 0<x<1,\\ 0, & \text{其他},\end{cases}\quad f_Y(y)=\begin{cases}1, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

面积 $A=XY$ 的概率密度为

$$f_Z(z)=\int_{-\infty}^{\infty}\frac{1}{|x|}f\left(x,\frac{z}{x}\right)\mathrm{d}x$$

$$=\int_{-\infty}^{\infty}\frac{1}{|x|}f_X(x)f_Y\left(\frac{z}{x}\right)\mathrm{d}x,$$

仅当$\begin{cases}0<x<1,\\ 0<\dfrac{z}{x}<1,\end{cases}$即$\begin{cases}0<x<1,\\ x>z>0\end{cases}$时上述积分的被积函数不等于零，由题 3.27 图得



题 3.27 图

$$f_Z(z)=\begin{cases}\displaystyle\int_z^1\frac{1}{x}\mathrm{d}x=-\ln z, & 0<z<1,\\ 0, & \text{其他}.\end{cases}$$

28. 设 X,Y 是相互独立的随机变量，它们都服从正态分布 $N(0,\sigma^2)$. 试验证随机变量 $Z=\sqrt{X^2+Y^2}$ 的概率密度为

$$f_Z(z)=\begin{cases}\dfrac{z}{\sigma^2}\mathrm{e}^{-z^2/(2\sigma^2)}, & z>0,\\ 0, & \text{其他}.\end{cases}$$

我们称 Z 服从参数为 σ $(\sigma>0)$ 的瑞利(Rayleigh) 分布.

证　先来求 Z 的分布函数 $F_Z(z)$，由于 $Z=\sqrt{X^2+Y^2}\geqslant 0$，知当 $z<0$ 时，$F_Z(z)=0$. 当 $z\geqslant 0$ 时有

$$F_Z(z)=P\left\{Z\leqslant z\right\}=P\left\{\sqrt{X^2+Y^2}\leqslant z\right\}$$

$$=P\left\{X^2+Y^2\leqslant z^2\right\}=\iint\limits_{x^2+y^2\leqslant z^2}f(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\iint\limits_{x^2+y^2\leqslant z^2}\frac{1}{2\pi\sigma^2}\mathrm{e}^{-(x^2+y^2)/(2\sigma^2)}\mathrm{d}x\mathrm{d}y$$

$$\xlongequal{\text{用极坐标}}\int_0^{2\pi}\mathrm{d}\theta\int_0^z\frac{1}{2\pi\sigma^2}\mathrm{e}^{-r^2/(2\sigma^2)}r\mathrm{d}r$$

$$=2\pi\cdot\frac{1}{2\pi}\left[-\mathrm{e}^{-r^2/(2\sigma^2)}\right]\Big|_0^z=1-\mathrm{e}^{-z^2/(2\sigma^2)}.$$

将 $F_Z(z)$ 关于 z 求导数，得 Z 的概率密度为

$$f_Z(z)=\begin{cases}\dfrac{z}{\sigma^2}\mathrm{e}^{-z^2/(2\sigma^2)}, & z>0,\\ 0, & \text{其他}.\end{cases}$$

29. 设随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}b\mathrm{e}^{-(x+y)}, & 0<x<1,0<y<\infty,\\ 0, & \text{其他}.\end{cases}$$

(1) 试确定常数 b.

(2) 求边缘概率密度 $f_X(x),f_Y(y)$.

(3) 求函数 $U=\max\{X,Y\}$ 的分布函数.

解　(1) 由

$$1=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^{\infty}\int_0^1 b\mathrm{e}^{-(x+y)}\mathrm{d}x\mathrm{d}y$$
$$=b\left(\int_0^{\infty}\mathrm{e}^{-y}\mathrm{d}y\right)\left(\int_0^1\mathrm{e}^{-x}\mathrm{d}x\right)=b(1-\mathrm{e}^{-1}),$$

得

$$b=\frac{1}{1-\mathrm{e}^{-1}}.$$

(2) $$f_X(x)=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}y=\begin{cases}\dfrac{1}{1-\mathrm{e}^{-1}}\displaystyle\int_0^{\infty}\mathrm{e}^{-x}\mathrm{e}^{-y}\mathrm{d}y=\dfrac{\mathrm{e}^{-x}}{1-\mathrm{e}^{-1}}, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$$

$$f_Y(y)=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}x=\begin{cases}\dfrac{1}{1-\mathrm{e}^{-1}}\displaystyle\int_0^1\mathrm{e}^{-x}\mathrm{e}^{-y}\mathrm{d}x=\mathrm{e}^{-y}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

(3) 由(2)知 $f(x,y)=f_X(x)f_Y(y)$，故 X,Y 相互独立. 分别记 $U=\max\{X,Y\}$，X 和 Y 的分布函数为 $F_U(u),F_X(x)$ 和 $F_Y(y)$，则有

$$F_U(u)=F_X(u)F_Y(u). \tag{$*_1$}$$

由(2)知

$$F_X(u)=\int_{-\infty}^{u}f_X(x)\mathrm{d}x=\begin{cases}0, & u<0,\\ \displaystyle\int_0^u\frac{\mathrm{e}^{-x}}{1-\mathrm{e}^{-1}}\mathrm{d}x, & 0\leqslant u<1,\\ 1, & u\geqslant 1\end{cases}$$

$$=\begin{cases}0, & u<0,\\ \dfrac{1-\mathrm{e}^{-u}}{1-\mathrm{e}^{-1}}, & 0\leqslant u<1,\\ 1, & u\geqslant 1.\end{cases}$$

$$F_Y(u)=\int_{-\infty}^{u}f_Y(y)\mathrm{d}y=\begin{cases}0, & u<0,\\ \int_0^u \mathrm{e}^{-y}\mathrm{d}y, & u\geqslant 0\end{cases}$$

$$=\begin{cases}0, & u<0,\\ 1-\mathrm{e}^{-u}, & u\geqslant 0.\end{cases}$$

将 $F_X(u),F_Y(u)$ 的表达式代入 $(*_1)$ 式，得到 $U=\max\{X,Y\}$ 的分布函数为

$$F_U(u)=\begin{cases}0, & u<0,\\ \dfrac{(1-\mathrm{e}^{-u})^2}{1-\mathrm{e}^{-1}}, & 0\leqslant u<1,\\ 1-\mathrm{e}^{-u}, & u\geqslant 1.\end{cases}$$

30. 设某种型号的电子元件的寿命（以 h 计）近似地服从正态分布 $N(160,20^2)$，随机地选取 4 只，求其中没有一只寿命小于 180 的概率.

解　以 $X_i(i=1,2,3,4)$ 记所选取的第 i 只元件的寿命，由题设一只元件寿命小于 180 h 的概率为

$$P\left\{X_i\leqslant 180\right\}=P\left\{\frac{X_i-160}{20}\leqslant\frac{180-160}{20}\right\}$$

$$=\Phi\left(\frac{180-160}{20}\right)=\Phi(1)=0.841\ 3.$$

可认为 X_1,X_2,X_3,X_4 相互独立，故选取的 4 只元件没有一只寿命小于 180 h 的概率为

$$\prod_{i=1}^{4}\left(1-P\left\{X_i\leqslant 180\right\}\right)=(1-0.841\ 3)^4=0.000\ 63.$$

31. 对某种电子装置的输出测量了 5 次，得到结果为 X_1,X_2,X_3,X_4,X_5. 设它们是相互独立的随机变量且都服从参数$\sigma=2$ 的瑞利分布.

(1) 求 $Z=\max\{X_1,X_2,X_3,X_4,X_5\}$ 的分布函数.

(2) 求 $P\{Z>4\}$.

解　参数 $\sigma=2$ 的瑞利分布，其概率密度为

$$f_X(x)=\begin{cases}\dfrac{x}{4}\mathrm{e}^{-x^2/8}, & x>0,\\ 0, & \text{其他}.\end{cases}$$

设其分布函数为 $F_X(x)$. 则当 $x<0$ 时，$F_X(x)=0$，当 $x\geqslant 0$ 时有

$$F_X(x)=\int_0^x\frac{x}{4}\mathrm{e}^{-x^2/8}\mathrm{d}x=-\mathrm{e}^{-x^2/8}\Big|_0^x=1-\mathrm{e}^{-x^2/8}.$$

即有

$$F_X(x)=\begin{cases}1-\mathrm{e}^{-x^2/8}, & x\geqslant 0,\\ 0, & x<0.\end{cases}$$

(1) 因 X_1,X_2,X_3,X_4,X_5 相互独立，且都服从参数 $\sigma=2$ 的瑞利分布，故 $Z=\max\{X_1,X_2,X_3,X_4,X_5\}$ 的分布函数为

$$F_Z(z)=[F_X(z)]^5=\begin{cases}(1-\mathrm{e}^{-z^2/8})^5, & z\geqslant 0,\\ 0, & z<0.\end{cases}$$

(2) $P\{Z>4\}=1-P\{Z\leqslant 4\}=1-F_Z(4)$

$$=1-(1-\mathrm{e}^{-2})^5=0.516\ 7.$$

32. 设随机变量 X,Y 相互独立，且服从同一分布，试证明

$$P\Big\{a<\min\{X,Y\}\leqslant b\Big\}=\Big(P\{X>a\}\Big)^2-\Big(P\{X>b\}\Big)^2\quad(a\leqslant b).$$

证 由题设 X 和 Y 相互独立，且服从同一分布，以 $F(x)$ 记它们的分布函数，又记 $N=\min\{X,Y\}$ 的分布函数为 $F_N(z)$，则

$$F_N(z)=1-[1-F(z)]^2,$$

于是

$$P\Big\{a<\min\{X,Y\}\leqslant b\Big\}=F_N(b)-F_N(a)=[1-F(a)]^2-[1-F(b)]^2.$$

因

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

从而

$$P\Big\{a<\min\{X,Y\}\leqslant b\Big\}=\Big(P\{X>a\}\Big)^2-\Big(P\{X>b\}\Big)^2.$$

33. 设 X,Y 是相互独立的随机变量，其分布律分别为

$$P\{X=k\}=p(k),\quad k=0,1,2,\cdots,$$
$$P\{Y=r\}=q(r),\quad r=0,1,2,\cdots.$$

证明随机变量 $Z=X+Y$ 的分布律为

$$P\{Z=i\}=\sum_{k=0}^{i}p(k)q(i-k),\quad i=0,1,2,\cdots.$$

证 随机变量 $Z=X+Y$ 的取值范围为 $0,1,2,\cdots$. 对于非负整数 i，$\{Z=i\}=\{X+Y=i\}$ 可按下列方式分解为若干个两两互不相容的事件之和：

$$\begin{aligned}\{Z=i\}&=\{X+Y=i\}\\&=\{X=0,Y=i\}\cup\{X=1,Y=i-1\}\cup\cdots\\&\quad\cup\{X=k,Y=i-k\}\cup\cdots\cup\{X=i,Y=0\}.\end{aligned}$$

又由 X,Y 的独立性知

$$P\{X=k,Y=i-k\}=P\{X=k\}P\{Y=i-k\}$$
$$=p(k)q(i-k),\quad k=0,1,2,\cdots,i.$$

因此

$$P\{Z=i\}=P\left\{\bigcup_{k=0}^{i}\{X=k,Y=i-k\}\right\}$$
$$=\sum_{k=0}^{i}P\{X=k,Y=i-k\}$$
$$=\sum_{k=0}^{i}p(k)q(i-k),\quad i=0,1,2,\cdots.$$

34. 设 X,Y 是相互独立的随机变量，$X\sim\pi(\lambda_1)$，$Y\sim\pi(\lambda_2)$. 证明

$$Z=X+Y\sim\pi(\lambda_1+\lambda_2).$$

证　因 $X\sim\pi(\lambda_1)$，$Y\sim\pi(\lambda_2)$，故

$$p(k)=P\{X=k\}=\frac{\lambda_1^k\mathrm{e}^{-\lambda_1}}{k!},\quad k=0,1,2,\cdots,$$

$$q(k)=P\{Y=k\}=\frac{\lambda_2^k\mathrm{e}^{-\lambda_2}}{k!},\quad k=0,1,2,\cdots.$$

而 $Z=X+Y$ 可能取的值为 $0,1,2,\cdots$，且 X,Y 相互独立. 由 33 题得

$$P\{Z=i\}=\sum_{k=0}^{i}\frac{\lambda_1^k\lambda_2^{i-k}}{k!(i-k)!}\mathrm{e}^{-(\lambda_1+\lambda_2)}$$
$$=\frac{\mathrm{e}^{-(\lambda_1+\lambda_2)}}{i!}\sum_{k=0}^{i}\frac{i!}{k!(i-k)!}\lambda_1^k\lambda_2^{i-k}$$
$$=\frac{\mathrm{e}^{-(\lambda_1+\lambda_2)}}{i!}\sum_{k=0}^{i}\binom{i}{k}\lambda_1^k\lambda_2^{i-k}=\frac{\mathrm{e}^{-(\lambda_1+\lambda_2)}}{i!}(\lambda_1+\lambda_2)^i$$
$$=\frac{(\lambda_1+\lambda_2)^i\mathrm{e}^{-(\lambda_1+\lambda_2)}}{i!},\quad i=0,1,2,\cdots,$$

即 $Z\sim\pi(\lambda_1+\lambda_2)$.

35. 设 X,Y 是相互独立的随机变量，$X\sim b(n_1,p)$，$Y\sim b(n_2,p)$. 证明

$$Z=X+Y\sim b(n_1+n_2,p).$$

证　因 $X\sim b(n_1,p)$，$Y\sim b(n_2,p)$，故

$$p(k)=P\{X=k\}=\binom{n_1}{k}p^k(1-p)^{n_1-k},\quad k=0,1,2,\cdots,n_1,$$

$$q(k)=P\{Y=k\}=\binom{n_2}{k}p^k(1-p)^{n_2-k},\quad k=0,1,2,\cdots,n_2.$$

而 $Z=X+Y$ 可能取的值为 $0,1,2,\cdots,n_1+n_2$，且 X,Y 相互独立，由 33 题得

$$P\{Z=i\}=\sum_{k=0}^{i}\binom{n_1}{k}p^k(1-p)^{n_1-k}\binom{n_2}{i-k}p^{i-k}(1-p)^{n_2-i+k}$$
$$=\left[\sum_{k=0}^{i}\binom{n_1}{k}\binom{n_2}{i-k}\right]p^i(1-p)^{n_1+n_2-i},$$
$$i=0,1,2,\cdots,n_1+n_2.$$

又

$$[p+(1-p)]^{n_1+n_2}=[p+(1-p)]^{n_1}[p+(1-p)]^{n_2}$$
$$=\left[\sum_{k=0}^{n_1}\binom{n_1}{k}p^k(1-p)^{n_1-k}\right]\left[\sum_{s=0}^{n_2}\binom{n_2}{s}p^s(1-p)^{n_2-s}\right].$$

比较上式两边展开式中 $p^i(1-p)^{n_1+n_2-i}$ 这一项的系数，知

$$\binom{n_1+n_2}{i}=\sum_{k=0}^{i}\binom{n_1}{k}\binom{n_2}{i-k},$$

从而

$$P\{Z=i\}=\binom{n_1+n_2}{i}p^i(1-p)^{n_1+n_2-i},\quad i=0,1,2,\cdots,n_1+n_2.$$

即
$$Z\sim b(n_1+n_2,p).$$

36. 设随机变量(X,Y)的分布律为

Y \ X	0	1	2	3	4	5
0	0.00	0.01	0.03	0.05	0.07	0.09
1	0.01	0.02	0.04	0.05	0.06	0.08
2	0.01	0.03	0.05	0.05	0.05	0.06
3	0.01	0.02	0.04	0.06	0.06	0.05

(1) 求 $P\{X=2\mid Y=2\}$，$P\{Y=3\mid X=0\}$.

(2) 求 $V=\max\{X,Y\}$ 的分布律.

(3) 求 $U=\min\{X,Y\}$ 的分布律.

(4) 求 $W=X+Y$ 的分布律.

解 (1) $P\{Y=2\}=\sum_{i=0}^{5}P\{X=i,Y=2\}$

$$=0.01+0.03+0.05+0.05+0.05+0.06=0.25,$$

$$P\{X=0\}=\sum_{j=0}^{3}P\{X=0,Y=j\}$$
$$=0.00+0.01+0.01+0.01=0.03,$$

故有

$$P\{X=2|Y=2\}=\frac{P\{X=2,Y=2\}}{P\{Y=2\}}=\frac{0.05}{0.25}=\frac{1}{5},$$

$$P\{Y=3|X=0\}=\frac{P\{X=0,Y=3\}}{P\{X=0\}}=\frac{0.01}{0.03}=\frac{1}{3}.$$

(2) $V=\max\{X,Y\}$ 所有可能的取值为 0,1,2,3,4,5.

$$\begin{aligned}\{V=i\}&=\{\max\{X,Y\}=i\}\\&=\{X=i,Y<i\}\cup\{X=i,Y=i\}\cup\{X<i,Y=i\}.\end{aligned}$$

上式右边三项两两互不相容,故有

$$\begin{aligned}P\{V=i\}&=P\{\max\{X,Y\}=i\}\\&=P\{X=i,Y<i\}+P\{X=i,Y=i\}+P\{X<i,Y=i\},\end{aligned}$$

例如

$$\begin{aligned}P\{V=2\}&=P\{X=2,Y=0\}+P\{X=2,Y=1\}\\&\quad+P\{X=2,Y=2\}+P\{X=0,Y=2\}\\&\quad+P\{X=1,Y=2\}\\&=0.03+0.04+0.05+0.01+0.03=0.16,\\P\{V=5\}&=P\{X=5,Y=0\}+P\{X=5,Y=1\}\\&\quad+P\{X=5,Y=2\}+P\{X=5,Y=3\}\\&=0.09+0.08+0.06+0.05=0.28.\end{aligned}$$

即有分布律

$V=\max\{X,Y\}$	0	1	2	3	4	5
p_k	0	0.04	0.16	0.28	0.24	0.28

(3) $U=\min\{X,Y\}$ 所有可能的取值为 0,1,2,3.

$$\begin{aligned}\{U=i\}&=\{\min\{X,Y\}=i\}\\&=\{X=i,Y>i\}\cup\{X=i,Y=i\}\cup\{X>i,Y=i\},\end{aligned}$$

$$P\{U=i\}=P\{X=i,Y>i\}+P\{X=i,Y=i\}+P\{X>i,Y=i\}.$$

例如

$$\begin{aligned}P\{U=2\}&=P\{X=2,Y=3\}+P\{X=2,Y=2\}\\&\quad+P\{X=3,Y=2\}+P\{X=4,Y=2\}\\&\quad+P\{X=5,Y=2\}\\&=0.04+0.05+0.05+0.05+0.06=0.25,\end{aligned}$$

即有

$U=\min\{X,Y\}$	0	1	2	3
p_k	0.28	0.30	0.25	0.17

(4) $W=X+Y$ 所有可能的取值为 $0,1,2,3,4,5,6,7,8$.

$$\{W=i\}=\{X+Y=i\}=\bigcup_{k=0}^{i}\{X=k,Y=i-k\},$$

$$P\{W=i\}=\sum_{k=0}^{i}P\{X=k,Y=i-k\}.$$

例如

$$\begin{aligned}P\{W=2\}&=P\{X=0,Y=2\}+P\{X=1,Y=1\}+P\{X=2,Y=0\}\\&=0.01+0.02+0.03=0.06,\\P\{W=5\}&=P\{X=0,Y=5\}+P\{X=1,Y=4\}\\&\quad+P\{X=2,Y=3\}+P\{X=3,Y=2\}\\&\quad+P\{X=4,Y=1\}+P\{X=5,Y=0\}\\&=0+0+0.04+0.05+0.06+0.09=0.24,\end{aligned}$$

即有分布律

$W=X+Y$	0	1	2	3	4	5	6	7	8
p_k	0	0.02	0.06	0.13	0.19	0.24	0.19	0.12	0.05

第四章　随机变量的数字特征

1. (1) 在下列句子中随机地取一个单词，以 X 表示取到的单词所包含的字母个数，写出 X 的分布律并求 $E(X)$.

"THE GIRL PUT ON HER BEAUTIFUL RED HAT".

(2) 在上述句子的 30 个字母中随机地取一个字母，以 Y 表示取到的字母所在单词所包含的字母数，写出 Y 的分布律并求 $E(Y)$.

(3) 一人掷骰子，如得 6 点则掷第 2 次，此时得分为 6＋第二次得到的点数；否则得分为他第一次掷得的点数，且不能再掷，求得分 X 的分布律及 $E(X)$.

解　(1) 随机试验属等可能概型. 所给句子共 8 个单词，其中含 2 个字母、含 4 个字母、含 9 个字母的各有一个单词，另有 5 个单词含 3 个字母，所以 X 的分布律为

X	2	3	4	9
p_k	$\frac{1}{8}$	$\frac{5}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

数学期望

$$E(X)=2\times\frac{1}{8}+3\times\frac{5}{8}+4\times\frac{1}{8}+9\times\frac{1}{8}=\frac{15}{4}.$$

(2) 随机试验属等可能概型，Y 的可能值也是 2,3,4,9. 样本空间 S 由各个字母组成，共有 30 个样本点，其中样本点属于$Y=2$ 的有 2 个，属于 $Y=3$ 的有 15 个，属于 $Y=4$ 的有 4 个，属于$Y=9$ 的有 9 个，所以 Y 的分布律为

Y	2	3	4	9
p_k	$\frac{2}{30}$	$\frac{15}{30}$	$\frac{4}{30}$	$\frac{9}{30}$

数学期望　$E(Y)=2\times\frac{2}{30}+3\times\frac{15}{30}+4\times\frac{4}{30}+9\times\frac{9}{30}=\frac{73}{15}.$

(3) 分布律为

X	1	2	3	4	5	7	8	9	10	11	12
p_k	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$

$$E(X)=1\times\frac{1}{6}+2\times\frac{1}{6}+3\times\frac{1}{6}+4\times\frac{1}{6}+5\times\frac{1}{6}+7\times\frac{1}{36}+8\times\frac{1}{36}+9\times\frac{1}{36}$$
$$+10\times\frac{1}{36}+11\times\frac{1}{36}+12\times\frac{1}{36}=\frac{49}{12}.$$

2. 某产品的次品率为0.1，检验员每天检验4次.每次随机地取10件产品进行检验，如发现其中的次品数多于1，就去调整设备.以 X 表示一天中调整设备的次数，试求 $E(X)$.（设诸产品是否为次品是相互独立的.）

解 先求检验一次，决定需要调整设备的概率.设抽检出次品件数为 Y，则 $Y\sim b(10,0.1)$.记需调整设备一次的概率为 p，则

$$p=P\{Y>1\}=1-P\{Y=0\}-P\{Y=1\}$$
$$=1-0.9^{10}-\binom{10}{1}\times0.9^{9}\times0.1=0.263\ 9.$$

又因各次检验结果相互独立，故

$$X\sim b(4,\ 0.263\ 9).$$

X 的分布律为

X	0	1	2	3	4
p_k	$(1-p)^4$	$4p(1-p)^3$	$6p^2(1-p)^2$	$4p^3(1-p)$	p^4

于是

$$E(X)=1\times4p(1-p)^3+2\times6p^2(1-p)^2+3\times4p^3(1-p)+4\times p^4$$
$$=4p=4\times0.263\ 9=1.055\ 6.$$

以后将会知道若 $X\sim b(n,p)$，则 $E(X)=np$.

3. 有3只球、4个盒子，盒子的编号为1,2,3,4.将球逐个独立地，随机地放入4个盒子中去.以 X 表示其中至少有一只球的盒子的最小号码（例如 $X=3$ 表示第1号、第2号盒子是空的，第3号盒子至少有一只球），试求 $E(X)$.

解法(i) 由于每只球都有4种放法，由乘法原理共有 $4^3=64$ 种放法.其中3只球都放在4号盒中的放置法仅有1种，从而

$$P\{X=4\}=\frac{1}{64}.$$

又 $\{X=3\}$ 表示事件“1,2号盒子都是空的，而3号盒子不空”.因1,2号盒子都空，球只能放置在3,4号两个盒子中，共有 2^3 种放置法，但其中有一种是3只球都放在4号盒子中，即3号盒子是空的，这不符合 $X=3$ 的要求，需除去，故有

$$P\{X=3\}=\frac{2^3-1}{64}=\frac{7}{64}.$$

同理可得

$$P\{X=2\}=\frac{3^3-2^3}{64}=\frac{19}{64},$$

$$P\{X=1\}=\frac{4^3-3^3}{64}=\frac{37}{64}.$$

因此

$$E(X)=\sum_{k=1}^{4}kP\{X=k\}=\frac{25}{16}.$$

注：$P\{X=1\}$ 也可由 $1-(P\{X=4\}+P\{X=3\}+P\{X=2\})$ 求得.

解法(ii) 以 $A_i(i=1,2,3,4)$ 记事件“第 i 号盒子是空盒”. $\{X=1\}$ 表示事件“第 1 号盒子中至少有一只球”，因此 $\{X=1\}=\bar{A}_1$，故

$$P\{X=1\}=P(\bar{A}_1)=1-P(A_1)=1-\left(\frac{3}{4}\right)^3=\frac{37}{64}.$$

（因第 1 号盒子为空盒，3 只球的每一只都只有 3 个盒子可以放，故 $P(A_1)=(3/4)^3$.）

$\{X=2\}$ 表示事件“第 1 号盒子为空盒且第 2 号盒子中至少有一只球”，因此 $\{X=2\}=A_1\bar{A}_2$. 故

$$\begin{aligned}P\{X=2\}&=P(A_1\bar{A}_2)=P(\bar{A}_2|A_1)P(A_1)\\&=[1-P(A_2|A_1)]P(A_1)\\&=\left[1-\left(\frac{2}{3}\right)^3\right]\left(\frac{3}{4}\right)^3=\frac{19}{64}.\end{aligned}$$

$\Big($因在第 1 号盒子是空盒的条件下，第 2 号盒子也是空盒，则 3 只球都只有 2 个盒子可以放，故 $P(A_2|A_1)=\left(\frac{2}{3}\right)^3$.$\Big)$

类似地，

$$\begin{aligned}P\{X=3\}&=P(A_1A_2\bar{A}_3)\\&=P(\bar{A}_3|A_1A_2)P(A_2|A_1)P(A_1)\\&=\left[1-\left(\frac{1}{2}\right)^3\right]\left(\frac{2}{3}\right)^3\left(\frac{3}{4}\right)^3=\frac{7}{64},\end{aligned}$$

$$P\{X=4\}=1-\frac{37}{64}-\frac{19}{64}-\frac{7}{64}=\frac{1}{64},$$

因此，$E(X)=\sum_{k=1}^{4}kP\{X=k\}=\frac{25}{16}$.

解法(iii) 将球编号. 以 X_1, X_2, X_3 分别记 1 号, 2 号, 3 号球所落入的盒子的号码数. 则 X_1, X_2, X_3 都是随机变量, 记 $X = \min\{X_1, X_2, X_3\}$, 按题意, 本题需要求的是

$$E(X) = E(\min\{X_1, X_2, X_3\}).$$

因 X_1, X_2, X_3 具有相同的分布律

X_j	1	2	3	4
p_k	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

因而 X_1, X_2, X_3 具有相同的分布函数

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

于是 $X = \min\{X_1, X_2, X_3\}$ 的分布函数为

$$F_{\min}(z) = 1 - [1 - F(z)]^3$$

$$= \begin{cases} 1 - (1 - 0)^3 = 0, & z < 1, \\ 1 - \left(1 - \frac{1}{4}\right)^3 = \frac{37}{64}, & 1 \leqslant z < 2, \\ 1 - \left(1 - \frac{2}{4}\right)^3 = \frac{56}{64}, & 2 \leqslant z < 3, \\ 1 - \left(1 - \frac{3}{4}\right)^3 = \frac{63}{64}, & 3 \leqslant z < 4, \\ 1 - (1 - 1)^3 = 1, & z \geqslant 4. \end{cases}$$

$X = \min\{X_1, X_2, X_3\}$ 的分布律为

X	1	2	3	4
p_k	$\frac{37}{64}$	$\frac{19}{64}$	$\frac{7}{64}$	$\frac{1}{64}$

得
$$E(X) = \frac{25}{16}.$$

4. (1) 设随机变量 X 的分布律为 $P\left\{X=(-1)^{j+1}\dfrac{3^j}{j}\right\}=\dfrac{2}{3^j},j=1,2,\cdots,$ 说明 X 的数学期望不存在.

(2) 一盒中装有一只黑球、一只白球，作摸球游戏，规则如下：一次从盒中随机摸一只球，若摸到白球，则游戏结束；若摸到黑球，放回再放入一只黑球，然后再从盒中随机地摸一只球. 试说明要游戏结束的摸球次数 X 的数学期望不存在.

解　(1) 因级数

$$\sum_{j=1}^{\infty}(-1)^{j+1}\frac{3^j}{j}P\left\{X=(-1)^{j+1}\frac{3^j}{j}\right\}$$
$$=\sum_{j=1}^{\infty}(-1)^{j+1}\frac{3^j}{j}\times\frac{2}{3^j}=2\sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j}$$

不绝对收敛，按定义 X 的数学期望不存在.

(2) 以 A_k 记事件"第 k 次摸球摸到黑球"，以 $\bar{A}_k$ 记事件"第 k 次摸球摸到白球"，以 C_k 表示事件"游戏在第 k 次摸球时结束"，$k=1,2,\cdots$. 按题意

$$C_k=A_1A_2\cdots A_{k-1}\bar{A}_k,$$

$$P(C_k)=P(\bar{A}_k\mid A_1A_2\cdots A_{k-1})P(A_{k-1}\mid A_1A_2\cdots A_{k-2})\cdots P(A_2\mid A_1)P(A_1).$$

$$P\{X=1\}=P(\bar{A}_1)=\frac{1}{2},$$

$$P\{X=2\}=P(A_1\bar{A}_2)=P(\bar{A}_2\mid A_1)P(A_1)=\frac{1}{3}\times\frac{1}{2},$$

$$P\{X=3\}=P(A_1A_2\bar{A}_3)=P(\bar{A}_3\mid A_1A_2)P(A_2\mid A_1)P(A_1)$$
$$=\frac{1}{4}\times\frac{2}{3}\times\frac{1}{2}=\frac{1}{4}\times\frac{1}{3},$$

$X=k$ 时，盒中共 $k+1$ 只球，其中只有一只是白球，故

$$P\{X=k\}=P(A_1\cdots A_{k-1}\bar{A}_k)$$
$$=P(\bar{A}_k\mid A_1A_2\cdots A_{k-1})P(A_{k-1}\mid A_1A_2\cdots A_{k-2})\cdots P(A_2\mid A_1)P(A_1)$$
$$=\frac{1}{k+1}\times\frac{k-1}{k}\times\frac{k-2}{k-1}\times\cdots\times\frac{2}{3}\times\frac{1}{2}=\frac{1}{k+1}\,\frac{1}{k}.$$

若 $E(X)$ 存在，则它应等于 $\sum\limits_{k=1}^{\infty}kP\{X=k\}$. 但

$$\sum_{k=1}^{\infty}kP\{X=k\}=\sum_{k=1}^{\infty}k\times\frac{1}{k+1}\times\frac{1}{k}=\sum_{k=1}^{\infty}\frac{1}{k+1}=\infty,$$

故 X 的数学期望不存在.

5. 设在某一规定的时间间隔里，某电气设备用于最大负荷的时间 X（以 min 计）是一个随机变量，其概率密度为

$$f(x)=\begin{cases}\dfrac{1}{1\,500^2}x, & 0\leqslant x\leqslant 1\,500,\\ \dfrac{-1}{1\,500^2}(x-3\,000), & 1\,500<x\leqslant 3\,000,\\ 0, & \text{其他}.\end{cases}$$

求 $E(X)$.

解 按连续型随机变量的数学期望的定义，有

$$\begin{aligned}E(X)&=\int_{-\infty}^{\infty}xf(x)\mathrm{d}x\\&=\int_{-\infty}^{0}xf(x)\mathrm{d}x+\int_{0}^{1\,500}xf(x)\mathrm{d}x\\&\quad+\int_{1\,500}^{3\,000}xf(x)\mathrm{d}x+\int_{3\,000}^{\infty}xf(x)\mathrm{d}x\\&=\int_{-\infty}^{0}x\cdot 0\mathrm{d}x+\int_{0}^{1\,500}x\cdot\frac{x}{1\,500^2}\mathrm{d}x\\&\quad+\int_{1\,500}^{3\,000}x\cdot\frac{-(x-3\,000)}{1\,500^2}\mathrm{d}x+\int_{3\,000}^{\infty}x\cdot 0\mathrm{d}x\\&=\frac{1}{1\,500^2}\frac{x^3}{3}\bigg|_0^{1\,500}+\frac{1}{1\,500^2}\left(3\,000\times\frac{x^2}{2}-\frac{x^3}{3}\right)\bigg|_{1\,500}^{3\,000}\\&=1\,500(\text{min}).\end{aligned}$$

6. (1) 设随机变量 X 的分布律为

X	-2	0	2
p_k	0.4	0.3	0.3

求 $E(X)$，$E(X^2)$，$E(3X^2+5)$.

(2) 设 $X\sim\pi(\lambda)$，求 $E\left(\dfrac{1}{X+1}\right)$.

解 (1) X 的分布律为

X	-2	0	2
p_k	0.4	0.3	0.3

$$E(X)=(-2)\times 0.4+0\times 0.3+2\times 0.3=-0.2.$$

由关于随机变量函数的数学期望的定理，知

$$E(X^2)=(-2)^2\times 0.4+0^2\times 0.3+2^2\times 0.3=2.8,$$

$$E(3X^2+5)=[3\times(-2)^2+5]\times 0.4+(3\times 0^2+5)\times 0.3+(3\times 2^2+5)\times 0.3$$
$$=13.4.$$

如利用数学期望的性质，则有

$$E(3X^2+5)=3E(X^2)+5=3\times 2.8+5=13.4.$$

(2) 因 $X\sim\pi(\lambda)$，故 $P\{X=k\}=\dfrac{\lambda^k e^{-\lambda}}{k!}$.

$$E\left(\frac{1}{X+1}\right)=\sum_{k=0}^{\infty}\frac{1}{k+1}P\{X=k\}=\sum_{k=0}^{\infty}\frac{1}{k+1}\frac{\lambda^k e^{-\lambda}}{k!}=\sum_{k=0}^{\infty}\frac{\lambda^k e^{-\lambda}}{(k+1)!}$$
$$=\frac{e^{-\lambda}}{\lambda}\sum_{k=0}^{\infty}\frac{\lambda^{k+1}}{(k+1)!}=\frac{e^{-\lambda}}{\lambda}\left(\sum_{j=1}^{\infty}\frac{\lambda^j}{j!}\right)=\frac{e^{-\lambda}}{\lambda}\left(\sum_{j=0}^{\infty}\frac{\lambda^j}{j!}-1\right)$$
$$=\frac{e^{-\lambda}}{\lambda}(e^{\lambda}-1)=\frac{1}{\lambda}(1-e^{-\lambda}).$$

7. (1) 设随机变量 X 的概率密度为

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

求(i)$Y=2X$,(ii)$Y=e^{-2X}$ 的数学期望.

(2) 设随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，且都服从(0,1)上的均匀分布. (i) 求$U=\max\{X_1,X_2,\cdots,X_n\}$的数学期望，(ii) 求 $V=\min\{X_1,X_2,\cdots,X_n\}$ 的数学期望.

解 (1) 由关于随机变量函数的数学期望的定理，知

(i) $E(Y)=E(2X)=\displaystyle\int_{-\infty}^{\infty}2xf(x)\mathrm{d}x$

$$=2\left(\int_{-\infty}^{0}x\cdot 0\mathrm{d}x+\int_{0}^{\infty}xe^{-x}\mathrm{d}x\right)$$
$$=2\left(-xe^{-x}\Big|_0^{\infty}+\int_0^{\infty}e^{-x}\mathrm{d}x\right)=-2e^{-x}\Big|_0^{\infty}=2;$$

(ii) $E(Y)=E(e^{-2X})=\displaystyle\int_0^{\infty}e^{-2x}\cdot e^{-x}\mathrm{d}x=\int_0^{\infty}e^{-3x}\mathrm{d}x$

$$=\frac{-1}{3}e^{-3x}\Big|_0^{\infty}=\frac{1}{3}.$$

(2) 因 $X_i\sim U(0,1)$，$i=1,2,\cdots,n$，X_i 的分布函数为

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

因 $X_1,X_2,\cdots,X_n$ 相互独立，故

(i) $U=\max\{X_1,X_2,\cdots,X_n\}$ 的分布函数为

$$F_U(u)=\begin{cases}0, & u<0,\\ u^n, & 0\leqslant u<1,\\ 1, & u\geqslant 1.\end{cases}$$

U 的概率密度为

$$f_U(u)=\begin{cases}nu^{n-1}, & 0<u<1,\\ 0, & \text{其他}.\end{cases}$$

$$E(U)=\int_{-\infty}^{\infty}uf_U(u)\mathrm{d}u=\int_0^1 u\cdot nu^{n-1}\mathrm{d}u=n\int_0^1 u^n\mathrm{d}u=\frac{n}{n+1}.$$

(ii) $V=\min\{X_1,X_2,\cdots,X_n\}$ 的分布函数为

$$F_V(v)=\begin{cases}0, & v<0,\\ 1-(1-v)^n, & 0\leqslant v<1,\\ 1, & v\geqslant 1.\end{cases}$$

V 的概率密度为

$$f_V(v)=\begin{cases}n(1-v)^{n-1}, & 0<v<1,\\ 0, & \text{其他}.\end{cases}$$

$$\begin{aligned}E(V)&=\int_{-\infty}^{\infty}vf_V(v)\mathrm{d}v=\int_0^1 vn(1-v)^{n-1}\mathrm{d}v\\&=-v(1-v)^n\Big|_0^1+\int_0^1(1-v)^n\mathrm{d}v\\&=-\frac{(1-v)^{n+1}}{n+1}\Big|_0^1=\frac{1}{n+1}.\end{aligned}$$

8. 设随机变量(X,Y)的分布律为

Y \ X	1	2	3
−1	0.2	0.1	0.0
0	0.1	0.0	0.3
1	0.1	0.1	0.1

(1) 求 $E(X),E(Y)$.

(2) 设 $Z=\dfrac{Y}{X}$，求 $E(Z)$.

(3) 设$Z=(X-Y)^2$，求 $E(Z)$.

解　由关于随机变量函数的数学期望 $E[g(X,Y)]$ 的定理，得

$$(1)\ E(X)=\sum_{i=1}^{3}\sum_{j=1}^{3}x_ip_{ij}$$
$$=1\times(0.2+0.1+0.1)+2\times(0.1+0+0.1)+3\times(0+0.3+0.1)$$
$$=2.$$

$$E(Y)=\sum_{j=1}^{3}\sum_{i=1}^{3}y_jp_{ij}$$
$$=(-1)\times(0.2+0.1+0)+0\times(0.1+0+0.3)+1\times(0.1+0.1+0.1)$$
$$=0.$$

$$(2)\ E(Z)=E\left(\frac{Y}{X}\right)$$
$$=\frac{-1}{1}P\{X=1,Y=-1\}+\frac{-1}{2}P\{X=2,Y=-1\}$$
$$+\frac{-1}{3}P\{X=3,Y=-1\}$$
$$+\frac{0}{1}P\{X=1,Y=0\}+\frac{0}{2}P\{X=2,Y=0\}$$
$$+\frac{0}{3}P\{X=3,Y=0\}+\frac{1}{1}P\{X=1,Y=1\}$$
$$+\frac{1}{2}P\{X=2,Y=1\}+\frac{1}{3}P\{X=3,Y=1\}$$
$$=-0.2-0.05+0.1+0.05+\frac{0.1}{3}=-\frac{1}{15}.$$

$$(3)\ E(Z)=E[(X-Y)^2]=\sum_{j=1}^{3}\sum_{i=1}^{3}(x_i-y_j)^2p_{ij}$$
$$=2^2\times0.2+3^2\times0.1+4^2\times0+1^2\times0.1+2^2\times0$$
$$+3^2\times0.3+0^2\times0.1+1^2\times0.1+2^2\times0.1$$
$$=5.$$

注：(i) 可先求出边缘分布律，然后求出 $E(X),E(Y)$.

(ii) 在(3)中可先算出 $Z=(X-Y)^2$ 的分布律

Z	0	1	4	9
p_k	0.1	0.2	0.3	0.4

然后求得 $E(Z)=\sum_{k=1}^{4}z_kp_k=5$.

9. (1) 设随机变量(X,Y)的概率密度为

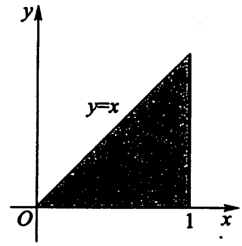
$$f(x,y)=\begin{cases}12y^2, & 0\leqslant y\leqslant x\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

求$E(X),E(Y),E(XY),E(X^2+Y^2)$.

(2) 设随机变量X,Y的联合概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{y}e^{-\left(y+\frac{x}{y}\right)}, & x>0,y>0,\\ 0, & \text{其他},\end{cases}$$

求$E(X),E(Y),E(XY)$.



题 4.9 图

解 (1) 各数学期望均可按照$E[g(X,Y)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y$计算. 因$f(x,y)$仅在有限区域$G:\{(x,y)\mid 0\leqslant y\leqslant x\leqslant 1\}$内不为零，故各数学期望均化为$G$（如题 4.9 图）上相应积分的计算.

$$E(X)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)\mathrm{d}x\mathrm{d}y=\iint_G x\cdot 12y^2\mathrm{d}x\mathrm{d}y$$

$$=\int_0^1\mathrm{d}x\int_0^x 12xy^2\mathrm{d}y=\frac{4}{5}.$$

$$E(Y)=\iint_G y\cdot 12y^2\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^x 12y^3\mathrm{d}y=\frac{3}{5}.$$

$$E(XY)=\iint_G xy\cdot 12y^2\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^x 12xy^3\mathrm{d}y=\frac{1}{2}.$$

$$E(X^2+Y^2)=\iint_G(x^2+y^2)12y^2\mathrm{d}x\mathrm{d}y$$

$$=\int_0^1\mathrm{d}x\int_0^x 12(x^2y^2+y^4)\mathrm{d}y=\frac{16}{15}.$$

(2) $$E(X)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)\mathrm{d}x\mathrm{d}y=\int_0^{\infty}\int_0^{\infty}\frac{x}{y}e^{-\left(y+\frac{x}{y}\right)}\mathrm{d}x\mathrm{d}y$$

$$=-\int_0^{\infty}e^{-y}\left[\int_0^{\infty}xe^{-x/y}\mathrm{d}\left(\frac{-x}{y}\right)\right]\mathrm{d}y$$

$$=-\int_0^{\infty}e^{-y}\left(xe^{-x/y}\Big|_0^{\infty}-\int_0^{\infty}e^{-x/y}\mathrm{d}x\right)\mathrm{d}y=\int_0^{\infty}e^{-y}y\mathrm{d}y=1.$$

$$E(Y)=\int_0^{\infty}\int_0^{\infty}e^{-(y+x/y)}\mathrm{d}x\mathrm{d}y=\int_0^{\infty}e^{-y}\int_0^{\infty}e^{-x/y}\mathrm{d}x\mathrm{d}y$$

$$=\int_0^{\infty}e^{-y}[-ye^{-x/y}]\Big|_0^{\infty}\mathrm{d}y=\int_0^{\infty}e^{-y}y\mathrm{d}y=1.$$

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)\mathrm{d}x\mathrm{d}y=\int_0^{\infty}\int_0^{\infty}xe^{-(y+x/y)}\mathrm{d}x\mathrm{d}y$$

$$=\int_0^{\infty}\mathrm{e}^{-y}\left(\int_0^{\infty}x\mathrm{e}^{-x/y}\mathrm{d}x\right)\mathrm{d}y.$$

而
$$\int_0^{\infty}x\mathrm{e}^{-x/y}\mathrm{d}x=-y\int_0^{\infty}x\mathrm{e}^{-x/y}\mathrm{d}\left(-\frac{x}{y}\right)=y^2,$$

故
$$E(XY)=\int_0^{\infty}y^2\mathrm{e}^{-y}\mathrm{d}y=\Gamma(3)^{①}=2.$$

10. (1) 设随机变量 $X\sim N(0,1)$，$Y\sim N(0,1)$ 且 X,Y 相互独立. 求 $E\left(\dfrac{X^2}{X^2+Y^2}\right)$.

(2) 一飞机进行空投物资作业，设目标点为原点 $O(0,0)$，物资着陆点为 (X,Y)，X,Y 相互独立，且设 $X\sim N(0,\sigma^2)$，$Y\sim N(0,\sigma^2)$，求原点到点 (X,Y) 间距离的数学期望.

解　(1) 由对称性知
$$E\left(\frac{X^2}{X^2+Y^2}\right)=E\left(\frac{Y^2}{X^2+Y^2}\right).$$

而
$$E\left(\frac{X^2}{X^2+Y^2}\right)+E\left(\frac{Y^2}{X^2+Y^2}\right)=E(1)=1,$$

故
$$E\left(\frac{X^2}{X^2+Y^2}\right)=\frac{1}{2}.$$

(2) 记原点到点 (X,Y) 的距离为 R，$R=\sqrt{X^2+Y^2}$，由题设 (X,Y) 的概率密度为
$$\begin{aligned}f(x,y)&=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-x^2/(2\sigma^2)}\cdot\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-y^2/(2\sigma^2)}\\&=\frac{1}{2\pi\sigma^2}\mathrm{e}^{-\frac{x^2+y^2}{2\sigma^2}},\quad -\infty<x<\infty,\quad -\infty<y<\infty.\end{aligned}$$
$$E(R)=E(\sqrt{X^2+Y^2})=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sqrt{x^2+y^2}\,\frac{1}{2\pi\sigma^2}\mathrm{e}^{-(x^2+y^2)/(2\sigma^2)}\mathrm{d}x\mathrm{d}y.$$

采用极坐标，
$$\begin{aligned}E(R)&=\int_0^{2\pi}\mathrm{d}\theta\int_0^{\infty}\frac{r}{2\pi\sigma^2}\mathrm{e}^{-r^2/(2\sigma^2)}r\mathrm{d}r\\&=2\pi\int_0^{\infty}\frac{1}{2\pi\sigma^2}r^2\mathrm{e}^{-r^2/(2\sigma^2)}\mathrm{d}r=\frac{1}{\sigma^2}\int_0^{\infty}r^2\mathrm{e}^{-r^2/(2\sigma^2)}\mathrm{d}r\end{aligned}$$

① Γ 函数：$\Gamma(\alpha)=\int_0^{\infty}x^{\alpha-1}\mathrm{e}^{-x}\mathrm{d}x,\alpha>0$，它具有性质：$\Gamma(\alpha+1)=\alpha\Gamma(\alpha),\alpha>0,\Gamma(1)=1,\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$，$\Gamma(n+1)=n\Gamma(n)=n!$ (n 为正整数).

$$=-\int_0^{\infty} r\mathrm{d}(\mathrm{e}^{-r^2/(2\sigma^2)}) = -r\mathrm{e}^{-r^2/(2\sigma^2)}\Big|_0^{\infty} + \int_0^{\infty} \mathrm{e}^{-r^2/(2\sigma^2)}\mathrm{d}r$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{e}^{-r^2/(2\sigma^2)}\mathrm{d}r = \frac{1}{2}\left[\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} \mathrm{e}^{-r^2/(2\sigma^2)}\mathrm{d}r\right]\sqrt{2\pi}\sigma$$

$$= \frac{1}{2}\times 1\times\sqrt{2\pi}\sigma = \sigma\sqrt{\frac{\pi}{2}}.$$

11. 一工厂生产的某种设备的寿命 X（以年计）服从指数分布，概率密度为

$$f(x) = \begin{cases} \frac{1}{4}\mathrm{e}^{-x/4}, & x > 0, \\ 0, & x \leqslant 0. \end{cases}$$

工厂规定，出售的设备若在售出一年之内损坏可予以调换. 若工厂售出一台设备赢利 100 元，调换一台设备厂方需花费 300 元. 试求厂方出售一台设备净赢利的数学期望.

解　一台设备在一年内调换的概率为

$$p = P\{X < 1\} = \int_0^1 \frac{1}{4}\mathrm{e}^{-x/4}\mathrm{d}x = -\mathrm{e}^{-x/4}\Big|_0^1 = 1 - \mathrm{e}^{-1/4}.$$

以 Y 记工厂售出一台设备的净赢利值，则 Y 具有分布律

Y	100	100 − 300
p_k	$\mathrm{e}^{-1/4}$	$1-\mathrm{e}^{-1/4}$

故有

$$E(Y) = 100\times\mathrm{e}^{-1/4} - 200(1-\mathrm{e}^{-1/4})$$
$$= 300\mathrm{e}^{-1/4} - 200 = 33.64(\text{元}).$$

12. 某车间生产的圆盘直径在区间 (a,b) 上服从均匀分布，试求圆盘面积的数学期望.

解　设圆盘直径为 X，按题设 X 具有概率密度

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

故圆盘面积 $A = \frac{1}{4}\pi X^2$ 的数学期望为

$$E\left(\frac{1}{4}\pi X^2\right) = \int_a^b \frac{1}{4}\pi x^2 \frac{1}{b-a}\mathrm{d}x = \frac{\pi}{12(b-a)}x^3\Big|_a^b$$

$$= \frac{\pi}{12}(b^2 + ab + a^2).$$

13. 设电压(以 V 计)$X \sim N(0,9)$. 将电压施加于一检波器，其输出电压为 $Y = 5X^2$，求输出电压 Y 的均值.

解法(i)　由 $X \sim N(0,9)$，即有 $E(X) = 0, D(X) = 9$.

$$E(Y) = E(5X^2) = 5E(X^2) = 5\{D(X) + [E(X)]^2\}$$
$$= 5 \times (9 + 0) = 45(\mathrm{V}).$$

解法(ii)　X 的概率密度为

$$f_X(x) = \frac{1}{3\sqrt{2\pi}} \mathrm{e}^{-x^2/18}, \quad -\infty < x < \infty.$$

$$E(Y) = E(5X^2) = 5E(X^2) = 5\int_{-\infty}^{\infty} \frac{x^2}{3\sqrt{2\pi}} \mathrm{e}^{-x^2/18} \mathrm{d}x$$
$$= \frac{5 \times 9}{3\sqrt{2\pi}} \left(-x\mathrm{e}^{-x^2/18} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \mathrm{e}^{-x^2/18} \mathrm{d}x \right)$$
$$= \frac{45}{3\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-x^2/18} \mathrm{d}x = 45 \int_{-\infty}^{\infty} f_X(x) \mathrm{d}x$$
$$= 45 \times 1 = 45(\mathrm{V}).$$

14. 设随机变量 X_1, X_2 的概率密度分别为

$$f_1(x) = \begin{cases} 2\mathrm{e}^{-2x}, & x > 0, \\ 0, & x \leqslant 0, \end{cases} \quad f_2(x) = \begin{cases} 4\mathrm{e}^{-4x}, & x > 0, \\ 0, & x \leqslant 0. \end{cases}$$

(1) 求 $E(X_1 + X_2), E(2X_1 - 3X_2^2)$.

(2) 又设 X_1, X_2 相互独立，求 $E(X_1 X_2)$.

解　若 X 服从以 θ 为参数的指数分布，其概率密度为

$$f(x) = \begin{cases} \dfrac{1}{\theta} \mathrm{e}^{-x/\theta}, & x > 0, \\ 0, & \text{其他}, \end{cases}$$

则 $E(X) = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x = \int_0^{\infty} x \frac{1}{\theta} \mathrm{e}^{-x/\theta} \mathrm{d}x$，令 $u = x/\theta$，得到

$$E(X) = \theta \int_0^{\infty} u\mathrm{e}^{-u} \mathrm{d}u = \theta\Gamma(2) = \theta\Gamma(1) = \theta,$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) \mathrm{d}x = \int_0^{\infty} x^2 \frac{1}{\theta} \mathrm{e}^{-x/\theta} \mathrm{d}x$$
$$= \theta^2 \int_0^{\infty} u^2 \mathrm{e}^{-u} \mathrm{d}u = \theta^2 \Gamma(3) \quad \left(\text{其中 } u = \frac{x}{\theta}\right)$$
$$= \theta^2 \cdot 2\Gamma(2) = \theta^2 \cdot 2\Gamma(1) = 2\theta^2,$$

故 $E(X_1) = \frac{1}{2}, E(X_2) = \frac{1}{4}, E(X_2^2) = 2 \times \left(\frac{1}{4}\right)^2 = \frac{1}{8}$，于是

(1) 由数学期望的性质，有

$$E(X_1+X_2)=E(X_1)+E(X_2)=\frac{3}{4},$$

$$E(2X_1-3X_2^2)=2E(X_1)-3E(X_2^2)=\frac{5}{8}.$$

(2) 因 X_1,X_2 相互独立，由数学期望的性质，有

$$E(X_1X_2)=E(X_1)E(X_2)=\frac{1}{2}\times\frac{1}{4}=\frac{1}{8}.$$

15. 将 n 只球（$1\sim n$ 号）随机地放进 n 个盒子（$1\sim n$ 号）中去，一个盒子装一只球. 若一只球装入与球同号的盒子中，则称为一个配对. 记 X 为总的配对数，求 $E(X)$.

解　引入随机变量

$$X_i=\begin{cases}1, & \text{若第 } i \text{ 号球装入第 } i \text{ 号盒子中},\\ 0, & \text{若第 } i \text{ 号球未装入第 } i \text{ 号盒子中},\end{cases}\quad i=1,2,\cdots,n,$$

则总的配对数 X 可表示成

$$X=X_1+X_2+\cdots+X_n.$$

显然
$$P\{X_i=1\}=\frac{1}{n},\quad i=1,2,\cdots,n.$$

X_i 的分布律为

X_i	0	1
p_k	$1-\frac{1}{n}$	$\frac{1}{n}$

即有 $E(X_i)=\frac{1}{n},i=1,2,\cdots,n$，于是

$$\begin{aligned}E(X)&=E(X_1+X_2+\cdots+X_n)\\&=E(X_1)+E(X_2)+\cdots+E(X_n)=1.\end{aligned}$$

16. 若有 n 把看上去样子相同的钥匙，其中只有一把能打开门上的锁，用它们去试开门上的锁. 设取到每只钥匙是等可能的. 若每把钥匙试开一次后除去，试用下面两种方法求试开次数 X 的数学期望.

(1) 写出 X 的分布律.

(2) 不写出 X 的分布律.

解　(1) 以 $A_k(k=1,2,\cdots,n)$ 表示事件"第 k 次试开是成功的". $\{X=k\}$ 表示前 $k-1$ 次所取的钥匙均未能打开门，而第 k 次所取的钥匙能将门打开. 即有

$$\begin{aligned}P\{X=k\}&=P(\bar{A}_1\bar{A}_2\cdots\bar{A}_{k-1}A_k)\\&=P(\bar{A}_1\bar{A}_2\cdots\bar{A}_{k-1})P(A_k\mid\bar{A}_1\bar{A}_2\cdots\bar{A}_{k-1})\\&=P(\bar{A}_1\bar{A}_2\cdots\bar{A}_{k-2})P(\bar{A}_{k-1}\mid\bar{A}_1\bar{A}_2\cdots\bar{A}_{k-2})P(A_k\mid\bar{A}_1\bar{A}_2\cdots\bar{A}_{k-1})\\&=\cdots\\&=P(\bar{A}_1)P(\bar{A}_2\mid\bar{A}_1)P(\bar{A}_3\mid\bar{A}_1\bar{A}_2)\cdots P(A_k\mid\bar{A}_1\bar{A}_2\cdots\bar{A}_{k-1})\\&=\frac{n-1}{n}\times\frac{n-2}{n-1}\times\cdots\times\frac{n-k+1}{n-k+2}\times\frac{1}{n-k+1}=\frac{1}{n},\end{aligned}$$

X 的分布律为

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

故

$$\begin{aligned}E(X)&=\sum_{k=1}^{n}kP\{X=k\}=\sum_{k=1}^{n}k\frac{1}{n}=\frac{1}{n}\sum_{k=1}^{n}k\\&=\frac{1}{n}\times\frac{n(n+1)}{2}=\frac{n+1}{2}.\end{aligned}$$

(2) 引入随机变量 X_k 如下：

$$X_1=1,$$

$$X_k=\begin{cases}1, & \text{前 } k-1 \text{ 次试开均未成功},\\0, & \text{前 } k-1 \text{ 次中有一次试开成功},\end{cases}\quad k=2,3,\cdots,n,$$

则

$$X=X_1+X_2+\cdots+X_n.$$

沿用(1)中的记号，则有

$$\begin{aligned}&E(X_1)=1,\\&E(X_k)=1\times P\{X_k=1\}=1\times P(\bar{A}_1\bar{A}_2\cdots\bar{A}_{k-1})\\&\quad=P(\bar{A}_1)P(\bar{A}_2\mid\bar{A}_1)\cdots P(\bar{A}_{k-1}\mid\bar{A}_1\bar{A}_2\cdots\bar{A}_{k-2})\\&\quad=\frac{n-1}{n}\times\frac{n-2}{n-1}\times\cdots\times\frac{n-(k-1)}{n-(k-2)}=\frac{n-k+1}{n},k=2,3,\cdots,n.\end{aligned}$$

故有

$$E(X)=1+\sum_{k=2}^{n}E(X_k)=1+\sum_{k=2}^{n}\frac{n-k+1}{n}=\frac{n+1}{2}.$$

17. 设 X 为随机变量，C 是常数，证明 $D(X)<E[(X-C)^2]$，对于 $C\neq E(X)$.（由于 $D(X)=E\{[X-E(X)]^2\}$，上式表明 $E[(X-C)^2]$ 当 $C=E(X)$ 时取到最小值.）

证　$E[(X-C)^2]=E(X^2-2CX+C^2)=E(X^2)-2CE(X)+C^2$

$$= E(X^2) - [E(X)]^2 + \{[E(X)]^2 - 2CE(X) + C^2\}$$
$$= D(X) + [E(X) - C]^2 \geqslant D(X).$$

等号仅当 $C = E(X)$ 时成立.

18. 设随机变量 X 服从瑞利分布，其概率密度为

$$f(x) = \begin{cases} \dfrac{x}{\sigma^2} e^{-x^2/(2\sigma^2)}, & x > 0, \\ 0, & x \leqslant 0, \end{cases}$$

其中 $\sigma > 0$ 是常数. 求 $E(X)$,$D(X)$.

解 $E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^{\infty} x \dfrac{x}{\sigma^2} e^{-x^2/(2\sigma^2)} dx.$

令 $u = x^2/(2\sigma^2)$，得到

$$E(X) = \sqrt{2}\sigma \int_0^{\infty} u^{1/2} e^{-u} du = \sqrt{2}\sigma \Gamma\left(\frac{3}{2}\right)①$$
$$= \sqrt{2}\sigma \frac{1}{2} \Gamma\left(\frac{1}{2}\right) = \sqrt{\frac{\pi}{2}}\sigma.$$
$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_0^{\infty} x^2 \frac{x}{\sigma^2} e^{-x^2/(2\sigma^2)} dx.$$

令 $u = x^2/(2\sigma^2)$，得到

$$E(X^2) = 2\sigma^2 \int_0^{\infty} u e^{-u} du = 2\sigma^2 \Gamma(2) = 2\sigma^2,$$

故

$$D(X) = E(X^2) - [E(X)]^2 = 2\sigma^2 - \frac{\pi}{2}\sigma^2 = \frac{4-\pi}{2}\sigma^2.$$

19. 设随机变量 X 服从 Γ 分布，其概率密度为

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

其中 $\alpha > 0, \beta > 0$ 是常数. 求 $E(X)$,$D(X)$.

解 $E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^{\infty} \dfrac{x}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} dx$

$$\xlongequal{令 u = x/\beta} \frac{\beta}{\Gamma(\alpha)} \int_0^{\infty} u^{\alpha} e^{-u} du = \frac{\beta}{\Gamma(\alpha)} \Gamma(\alpha+1)$$
$$= \frac{\beta}{\Gamma(\alpha)} \alpha \Gamma(\alpha) = \alpha\beta.$$

---

① 参见 101 页注.

$$E(X^2)=\int_{-\infty}^{\infty}x^2f(x)\mathrm{d}x=\int_0^{\infty}\frac{x^2}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-x/\beta}\mathrm{d}x$$

$$\xlongequal{令\ u=x/\beta}\frac{\beta^2}{\Gamma(\alpha)}\int_0^{\infty}u^{\alpha+1}\mathrm{e}^{-u}\mathrm{d}u=\frac{\beta^2}{\Gamma(\alpha)}\Gamma(\alpha+2)$$

$$=\frac{\beta^2}{\Gamma(\alpha)}(\alpha+1)\alpha\Gamma(\alpha)=\alpha(\alpha+1)\beta^2.$$

$$D(X)=\alpha(\alpha+1)\beta^2-(\alpha\beta)^2=\alpha\beta^2.$$

20. 设随机变量 X 服从几何分布，其分布律为

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

其中 $0<p<1$ 是常数．求 $E(X),D(X)$．

解 $E(X)=\sum_{n=1}^{\infty}nP\{X=n\}=\sum_{n=1}^{\infty}np(1-p)^{n-1}$

$$=p\sum_{n=1}^{\infty}n(1-p)^{n-1}=p\frac{1}{[1-(1-p)]^2}=\frac{1}{p}.$$

这是因为

$$\frac{1}{1-x}=1+x+x^2+\cdots+x^k+\cdots,\quad |x|<1,$$

两边对 x 求导，就有

$$\frac{1}{(1-x)^2}=1+2x+3x^2+\cdots+kx^{k-1}+\cdots,\quad |x|<1.\qquad (*_1)$$

又　$E[X(X+1)]=\sum_{n=1}^{\infty}n(n+1)P\{X=n\}=p\sum_{n=1}^{\infty}n(n+1)(1-p)^{n-1}.$

将上述$(*_1)$式两边关于 x 求导，就有

$$\frac{2}{(1-x)^3}=1\times2+2\times3x+\cdots+(k-1)\times kx^{k-2}+\cdots,\quad |x|<1,$$

由此知

$$E[X(X+1)]=p\frac{2}{[1-(1-p)]^3}=\frac{2}{p^2},$$

故

$$D(X)=E(X^2)-[E(X)]^2=E[X(X+1)-X]-[E(X)]^2$$

$$=E[X(X+1)]-E(X)-[E(X)]^2=\frac{2}{p^2}-\frac{1}{p}-\frac{1}{p^2}=\frac{1-p}{p^2}.$$

21. 设长方形的长（以 m 计）$X\sim U(0,2)$，已知长方形的周长（以 m 计）为 20．求长方形面积 A 的数学期望和方差．

解 长方形的长为 X，周长为 20，所以它的面积 A 为

$$A=X(10-X).$$

现在 $X \sim U(0,2)$，X 的概率密度为

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

所以

$$\begin{aligned}E(A)&=E[X(10-X)]=\int_0^2 x(10-x)\cdot\frac{1}{2}\mathrm{d}x\\&=\left(\frac{5}{2}x^2-\frac{1}{6}x^3\right)\Big|_0^2=\frac{26}{3}=8.67,\\E(A^2)&=E[X^2(10-X)^2]=\int_0^2 x^2(10-x)^2\cdot\frac{1}{2}\mathrm{d}x\\&=\frac{1}{2}\int_0^2(100x^2-20x^3+x^4)\mathrm{d}x=\frac{1\,448}{15}=96.53,\\D(A)&=E(A^2)-[E(A)]^2=\frac{1\,448}{15}-\left(\frac{26}{3}\right)^2=21.42.\end{aligned}$$

22. (1) 设随机变量 X_1,X_2,X_3,X_4 相互独立，且有 $E(X_i)=i,D(X_i)=5-i,i=1,2,3,4$. 设 $Y=2X_1-X_2+3X_3-\frac{1}{2}X_4$. 求 $E(Y),D(Y)$.

(2) 设随机变量 X,Y 相互独立，且 $X\sim N(720,30^2)$，$Y\sim N(640,25^2)$，求 $Z_1=2X+Y$，$Z_2=X-Y$ 的分布，并求概率 $P\{X>Y\}$，$P\{X+Y>1\,400\}$.

解 (1)
$$\begin{aligned}E(Y)&=E\left(2X_1-X_2+3X_3-\frac{1}{2}X_4\right)\\&=2E(X_1)-E(X_2)+3E(X_3)-\frac{1}{2}E(X_4)\\&=2\times1-2+3\times3-\frac{1}{2}\times4=7.\end{aligned}$$

因 X_1,X_2,X_3,X_4 相互独立，故有

$$\begin{aligned}D(Y)&=D\left(2X_1-X_2+3X_3-\frac{1}{2}X_4\right)\\&=4D(X_1)+D(X_2)+9D(X_3)+\frac{1}{4}D(X_4)\\&=4\times4+3+9\times2+\frac{1}{4}\times1=37.25.\end{aligned}$$

(2) 因 X,Y 相互独立，且 $X\sim N(720,30^2)$，$Y\sim N(640,25^2)$，故 $Z_1=2X+Y$，$Z_2=X-Y$ 均服从正态分布，且

$$\begin{aligned}E(Z_1)&=E(2X+Y)=2E(X)+E(Y)\\&=2\times720+640=2\,080,\end{aligned}$$

$$D(Z_1) = D(2X+Y) = 4D(X)+D(Y)$$
$$= 4\times 30^2 + 25^2 = 4\ 225,$$
$$E(Z_2) = E(X-Y) = E(X)-E(Y)$$
$$= 720 - 640 = 80,$$
$$D(Z_2) = D(X-Y) = D(X)+D(Y)$$
$$= 30^2 + 25^2 = 1\ 525,$$

故有
$$Z_1 \sim N(2\ 080, 4\ 225),\quad Z_2 \sim N(80, 1\ 525).$$
$$P\{X>Y\} = P\{X-Y>0\} = P\{Z_2>0\}$$
$$= 1 - P\{Z_2 \leqslant 0\} = 1 - \Phi\left(\frac{0-80}{\sqrt{1\ 525}}\right)$$
$$= \Phi(2.048\ 6) = 0.979\ 8.$$

又
$$X+Y \sim N(E(X)+E(Y), D(X)+D(Y)),$$
即
$$X+Y \sim N(1\ 360, 1\ 525).$$
故
$$P\{X+Y>1\ 400\} = 1 - P\{X+Y \leqslant 1\ 400\}$$
$$= 1 - \Phi\left(\frac{1\ 400 - 1\ 360}{\sqrt{1\ 525}}\right) = 1 - \Phi(1.02)$$
$$= 1 - 0.846\ 1 = 0.153\ 9.$$

23. 五家商店联营，它们每两周售出的某种农产品的数量（以 kg 计）分别为 X_1, X_2, X_3, X_4, X_5. 已知 $X_1 \sim N(200, 225)$, $X_2 \sim N(240, 240)$, $X_3 \sim N(180, 225)$, $X_4 \sim N(260, 265)$, $X_5 \sim N(320, 270)$, X_1, X_2, X_3, X_4, X_5 相互独立.

(1) 求五家商店两周的总销售量的均值和方差.

(2) 商店每隔两周进货一次，为了使新的供货到达前商店不会脱销的概率大于 0.99，问商店的仓库应至少储存多少千克该产品？

解　以 Y 记五家商店关于该种产品的总销售量，即 $Y = X_1 + X_2 + X_3 + X_4 + X_5$.

(1) 按题设 X_i $(i = 1,2,3,4,5)$ 相互独立且均服从正态分布，即有

$$E(Y) = \sum_{i=1}^{5} E(X_i) = 200 + 240 + 180 + 260 + 320 = 1\ 200,$$

$$D(Y) = \sum_{i=1}^{5} D(Y_i) = 225 + 240 + 225 + 265 + 270 = 1\ 225.$$

(2) 设仓库应至少储存 n 千克该产品，才能使该产品不脱销的概率大于 0.99，按题意，n 应满足条件

$$P\{Y \leqslant n\} > 0.99.$$

由于 $Y \sim N(1\ 200, 35^2)$，故有

$$P\{Y \leqslant n\} = P\left\{\frac{Y-1\ 200}{35} \leqslant \frac{n-1\ 200}{35}\right\} = \Phi\left(\frac{n-1\ 200}{35}\right),$$

因而上述不等式即为

$$\Phi\left(\frac{n-1\ 200}{35}\right) > 0.99 = \Phi(2.33),$$

从而 $\frac{n-1\ 200}{35} > 2.33$，故应有

$$n > 1\ 200 + 2.33 \times 35 = 1\ 281.55,$$

即需取 $n = 1\ 282$ kg.

24. 卡车装运水泥，设每袋水泥质量 X（以 kg 计）服从 $N(50, 2.5^2)$，问至多装多少袋水泥使总质量超过 2 000 的概率不大于 0.05？

解 设至多能装运 n 袋水泥，各袋水泥的质量分别为 $X_1, X_2, \cdots, X_n$，则

$$X_i \sim N(50, 2.5^2), \quad i = 1, 2, \cdots, n,$$

故卡车所装运水泥的总质量为

$$W = X_1 + X_2 + \cdots + X_n.$$

按题意 n 需满足

$$P\{W > 2\ 000\} \leqslant 0.05.$$

对于像这样的实际问题，认为 $X_1, X_2, \cdots, X_n$ 相互独立是适宜的，此时

$$E(W) = 50n, \quad D(W) = 2.5^2 n,$$

于是

$$W \sim N(50n, 2.5^2 n).$$

从而

$$P\{W > 2\ 000\} = 1 - \Phi\left(\frac{2\ 000 - 50n}{2.5\sqrt{n}}\right),$$

即 n 应满足

$$\Phi\left(\frac{2\ 000 - 50n}{2.5\sqrt{n}}\right) \geqslant 0.95 = \Phi(1.645).$$

故应有

$$\frac{2\ 000 - 50n}{2.5\sqrt{n}} \geqslant 1.645,$$

解得

$$\sqrt{n} \leqslant 6.283\ 6,$$

从而

$$n \leqslant 39.484.$$

故 n 至多取 39，即该卡车至多能装运 39 袋水泥，方能使超过 2 000 kg 的概率不大于 0.05.

（在这里我们指出，若设 $W = nX$，其中 $X \sim N(50, 2.5^2)$，而去求出 $n \approx 37$，

那就犯错误了，为什么？）

25. 设随机变量 X,Y 相互独立，且都服从(0,1)上的均匀分布.

(1) 求 $E(XY),E(X/Y),E[\ln(XY)],E(|Y-X|)$.

(2) 以 X,Y 为边长作一长方形，以 A,C 分别表示长方形的面积和周长，求 A 和 C 的相关系数.

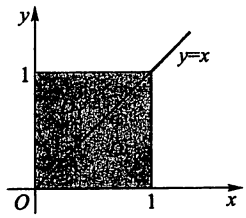
解 (1) X,Y 的概率密度都是

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

$$E(XY)=E(X)E(Y)=\frac{1}{2}\times\frac{1}{2}=\frac{1}{4}.$$

$E\left(\frac{X}{Y}\right)$不存在$\left(\text{因}\int_0^1\int_0^1\frac{x}{y}\mathrm{d}x\mathrm{d}y\ \text{发散}\right)$.

$$E[\ln(XY)]=\int_0^1\int_0^1(\ln x+\ln y)\mathrm{d}x\mathrm{d}y$$
$$=2\int_0^1\int_0^1(\ln x)\mathrm{d}x\mathrm{d}y=-2.$$



题 4.25 图

$E(|Y-X|)$

$$=\iint_D|y-x|\mathrm{d}x\mathrm{d}y\ (\text{如题 4.25 图 } D=D_1\cup D_2)$$

$$=2\iint_{D_1}(y-x)\mathrm{d}x\mathrm{d}y=2\int_0^1\int_x^1(y-x)\mathrm{d}y\mathrm{d}x=\frac{1}{3}.$$

(2) $A=XY,\quad C=2(X+Y),\quad \mathrm{Cov}(A,C)=E(AC)-E(A)E(C).$

$$AC=2X^2Y+2XY^2,$$

$$E(X^2)=E(Y^2)=D(X)+[E(X)]^2=\frac{1}{12}+\frac{1}{4}=\frac{1}{3}.$$

$$\begin{aligned}E(AC)&=2E(X^2Y)+2E(XY^2)\\&=2E(X^2)E(Y)+2E(X)E(Y^2)\\&=2\times\frac{1}{3}\times\frac{1}{2}+2\times\frac{1}{2}\times\frac{1}{3}=\frac{2}{3}.\end{aligned}$$

$$\begin{aligned}\mathrm{Cov}(A,C)&=E(AC)-E(A)E(C)\\&=\frac{2}{3}-\{E(X)E(Y)\times 2[E(X)+E(Y)]\}\\&=\frac{2}{3}-\left[\frac{1}{2}\times\frac{1}{2}\times 2\left(\frac{1}{2}+\frac{1}{2}\right)\right]=\frac{1}{6}.\end{aligned}$$

$$D(A)=E(X^2Y^2)-[E(X)E(Y)]^2=E(X^2)E(Y^2)-\left(\frac{1}{2}\times\frac{1}{2}\right)^2$$

$$=\left(\frac{1}{3}\right)^2-\left(\frac{1}{4}\right)^2=\frac{7}{144}.$$

$$D(C)=D(2X+2Y)=D(2X)+D(2Y)=4\times\frac{1}{12}+4\times\frac{1}{12}=\frac{2}{3}.$$

故 $$\rho_{AC}=\frac{\operatorname{Cov}(A,C)}{\sqrt{D(A)D(C)}}=\frac{1}{6}\Big/\sqrt{\frac{7}{144}\times\frac{2}{3}}=\sqrt{\frac{6}{7}}.$$

26. (1) 设随机变量 X_1,X_2,X_3 相互独立，且有 $X_1\sim b\left(4,\frac{1}{2}\right),X_2\sim b\left(6,\frac{1}{3}\right),X_3\sim b\left(6,\frac{1}{3}\right)$，求 $P\{X_1=2,X_2=2,X_3=5\},E(X_1X_2X_3),E(X_1-X_2),E(X_1-2X_2)$.

(2) 设 X,Y 是随机变量，且有 $E(X)=3,E(Y)=1,D(X)=4,D(Y)=9$，令 $Z=5X-Y+15$，分别在下列 3 种情况下求 $E(Z)$ 和 $D(Z)$.

(i) X,Y 相互独立，(ii) X,Y 不相关，(iii) X 与 Y 的相关系数为 0.25.

解 (1) $P\{X_1=2,X_2=2,X_3=5\}$

$$=P\{X_1=2\}P\{X_2=2\}P\{X_3=5\}.$$

因 $$P\{X_1=2\}=\binom{4}{2}\left(\frac{1}{2}\right)^2\left(1-\frac{1}{2}\right)^{4-2}=\binom{4}{2}\left(\frac{1}{2}\right)^4,$$

$$P\{X_2=2\}=\binom{6}{2}\left(\frac{1}{3}\right)^2\left(1-\frac{1}{3}\right)^{6-2}=\binom{6}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^4,$$

$$P\{X_3=5\}=\binom{6}{5}\left(\frac{1}{3}\right)^5\left(1-\frac{1}{3}\right)^{6-5}=\binom{6}{5}\left(\frac{1}{3}\right)^5\frac{2}{3},$$

故 $$P\{X_1=2,X_2=2,X_3=5\}=P\{X_1=2\}\cdot P\{X_2=2\}\cdot P\{X_3=5\}=0.002\,03,$$

$$E(X_1X_2X_3)=E(X_1)E(X_2)E(X_3)=\left(4\times\frac{1}{2}\right)\left(6\times\frac{1}{3}\right)\left(6\times\frac{1}{3}\right)=8.$$

$$E(X_1-X_2)=E(X_1)-E(X_2)=2-2=0.$$

$$E(X_1-2X_2)=E(X_1)-2E(X_2)=-2.$$

(2) 对于 $E(Z)$，在(i)，(ii)，(iii) 三种情况下都有

$$E(Z)=E(5X-Y+15)=5E(X)-E(Y)+15=15-1+15=29.$$

对于 $D(Z)$，(i) X,Y 相互独立，则

$$D(5X-Y+15)=D(5X-Y)=D(5X)+D(-Y)=25D(X)+D(Y)=25\times4+9=109.$$

(ii) X,Y 不相关，即 $\operatorname{Cov}(X,Y)=0$，

$$D(Z)=109.$$

(iii) $\rho_{XY} = 0.25$,则

$$\mathrm{Cov}(X,Y) = \sqrt{D(X)}\ \sqrt{D(Y)}\rho_{XY} = 2\times 3\times 0.25 = 1.5,$$

$$\begin{aligned} D(5X-Y+15) &= D(5X-Y) = 25D(X)+D(Y)-10\mathrm{Cov}(X,Y) \\ &= 100+9-10\times 1.5 = 94. \end{aligned}$$

27. 下列各对随机变量 X 和 Y,问哪几对是相互独立的?哪几对是不相关的?

(1) $X\sim U(0,1), Y = X^2$.

(2) $X\sim U(-1,1), Y = X^2$.

(3) $X = \cos V, Y = \sin V, V\sim U(0,2\pi)$.

若(X,Y)的概率密度为 $f(x,y)$,

(4) $f(x,y) = \begin{cases} x+y, & 0<x<1, 0<y<1, \\ 0, & \text{其他}. \end{cases}$

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

解　(1) $E(X) = \dfrac{1}{2}$,　$E(Y) = E(X^2) = \displaystyle\int_0^1 x^2\,\mathrm{d}x = \dfrac{1}{3}$,

$$E(XY) = E(X^3) = \int_0^1 x^3\,\mathrm{d}x = \frac{1}{4}.$$

$$\mathrm{Cov}(X,Y) = E(XY)-E(X)E(Y) = \frac{1}{4}-\frac{1}{2}\times\frac{1}{3}\neq 0.$$

故 X,Y 不相互独立,也不是不相关的.

(2) $E(X) = 0$,　$E(Y) = E(X^2) = \displaystyle\int_{-1}^1 \frac{1}{2}x^2\,\mathrm{d}x = \frac{1}{3}$,

$$E(XY) = E(X^3) = \int_{-1}^1 \frac{1}{2}x^3\,\mathrm{d}x = 0.$$

$$\mathrm{Cov}(X,Y) = E(XY)-E(X)E(Y) = 0-0 = 0.$$

故 X,Y 不相互独立,但不相关.

(3) $E(X) = \displaystyle\int_0^{2\pi}\frac{1}{2\pi}\cos v\,\mathrm{d}v = 0$,　$E(Y) = \displaystyle\int_0^{2\pi}\frac{1}{2\pi}\sin v\,\mathrm{d}v = 0$,

$$E(XY) = E(\sin V\cos V) = \frac{1}{2}E(\sin 2V) = \frac{1}{2}\int_0^{2\pi}\frac{1}{2\pi}\sin 2v\,\mathrm{d}v = 0,$$

$$\mathrm{Cov}(X,Y) = E(XY)-E(X)E(Y) = 0-0\times 0 = 0,$$

故 X,Y 不相互独立,但不相关.

(4) $f(x,y) = \begin{cases} x+y, & 0<x<1, 0<y<1, \\ 0, & \text{其他}. \end{cases}$

$$f_X(x)=\begin{cases}\int_0^1(x+y)\mathrm{d}y=x+\dfrac{1}{2}, & 0<x<1,\\ 0, & \text{其他},\end{cases}$$

$$f_Y(y)=\begin{cases}y+\dfrac{1}{2}, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

$f(x,y)$ 与 $f_X(x)f_Y(y)$ 在平面上不几乎处处相等，X,Y 不相互独立.

$$E(X)=\int_0^1 x\left(x+\frac{1}{2}\right)\mathrm{d}x=\frac{7}{12},\quad E(Y)=\frac{7}{12},$$

$$E(XY)=\int_0^1\int_0^1 xy(x+y)\mathrm{d}x\mathrm{d}y=\frac{1}{3}.$$

$$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)\neq 0.$$

故 X,Y 不是不相关的，因而一定也是不相互独立的.

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

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

$f(x,y)=f_X(x)f_Y(y)$ 对于任意 x,y 成立.

故 X,Y 相互独立，因此 X,Y 也是不相关的.

28. 设二维随机变量 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{\pi}, & x^2+y^2\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

试验证 X 和 Y 是不相关的，但 X 和 Y 不是相互独立的.

证 $E(X)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)\mathrm{d}x\mathrm{d}y=\iint\limits_{x^2+y^2\leqslant 1}\dfrac{x}{\pi}\mathrm{d}x\mathrm{d}y$

$$=\frac{1}{\pi}\int_{-1}^{1}\mathrm{d}y\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}x\mathrm{d}x=0.$$

同样 $E(Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)\mathrm{d}x\mathrm{d}y=\iint\limits_{x^2+y^2\leqslant 1}\dfrac{y}{\pi}\mathrm{d}x\mathrm{d}y=0,$

而 $E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)\mathrm{d}x\mathrm{d}y=\iint\limits_{x^2+y^2\leqslant 1}\dfrac{xy}{\pi}\mathrm{d}x\mathrm{d}y$

$$=\frac{1}{\pi}\int_{-1}^{1}y\mathrm{d}y\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}x\mathrm{d}x=0,$$

从而 $E(XY)=E(X)E(Y),$

这表明 X,Y 是不相关的.又

$$f_X(x)=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}y=\begin{cases}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\frac{1}{\pi}\mathrm{d}y=\frac{2}{\pi}\sqrt{1-x^2}, & -1<x<1,\\ 0, & \text{其他}.\end{cases}$$

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

显然 $f_X(x)f_Y(y)\neq f(x,y)$，故 X,Y 不是相互独立的.

29. 设随机变量(X,Y)的分布律为

Y \ X	-1	0	1
-1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
0	$\frac{1}{8}$	0	$\frac{1}{8}$
1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

验证 X 和 Y 是不相关的，但 X 和 Y 不是相互独立的.

证　先求出边缘分布律如下：

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

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

易见 $P\{X=0,Y=0\}=0\neq P\{X=0\}P\{Y=0\}$，故 X,Y 不是相互独立的.又知 X,Y 具有相同的分布律，且有

$$E(X)=E(Y)=(-1)\times\frac{3}{8}+1\times\frac{3}{8}=0.$$

又

$$E(XY)=\sum_{j=1}^{3}\sum_{i=1}^{3}x_iy_jp_{ij}$$

$$= (-1)\times(-1)\times\frac{1}{8}+(-1)\times 1\times\frac{1}{8}+1\times(-1)\times\frac{1}{8}+1\times 1\times\frac{1}{8}$$
$$= 0,$$

即有 $E(XY)=E(X)E(Y)$，故 X,Y 是不相关的.

30. 设 A 和 B 是试验 E 的两个事件，且 $P(A)>0,P(B)>0$，并定义随机变量 X,Y 如下：

$$X=\begin{cases}1, & 若 A 发生,\\ 0, & 若 A 不发生,\end{cases}\quad Y=\begin{cases}1, & 若 B 发生,\\ 0, & 若 B 不发生.\end{cases}$$

证明若 $\rho_{XY}=0$，则 X 和 Y 必定相互独立.

证 X,Y 的分布律分别为

X	0	1
p_k	$P(\bar{A})$	$P(A)$

Y	0	1
p_k	$P(\bar{B})$	$P(B)$

由 X,Y 的定义，XY 只能取 0,1 两个值，且

$$P\{XY=1\}=P\{X=1,Y=1\}=P(AB),$$

于是得 XY 的分布律为

XY	0	1
p_k	$1-P(AB)$	$P(AB)$

即得 $\qquad E(X)=P(A),\quad E(Y)=P(B),\quad E(XY)=P(AB).$

由假设 $\rho_{XY}=0$，得 $E(XY)=E(X)E(Y)$，即

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

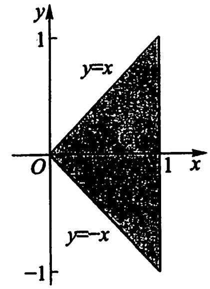
故知 A 与 B 相互独立. 从而知 A 与 $\bar{B}$、$\bar{A}$ 与 B、$\bar{A}$ 与 $\bar{B}$ 也相互独立，于是

$$P\{X=1,Y=1\}=P(AB)=P(A)P(B)=P\{X=1\}P\{Y=1\},$$
$$P\{X=1,Y=0\}=P(A\bar{B})=P(A)P(\bar{B})=P\{X=1\}P\{Y=0\},$$
$$P\{X=0,Y=1\}=P(\bar{A}B)=P(\bar{A})P(B)=P\{X=0\}P\{Y=1\},$$
$$P\{X=0,Y=0\}=P(\bar{A}\bar{B})=P(\bar{A})P(\bar{B})=P\{X=0\}P\{Y=0\},$$

故 X,Y 相互独立.

31. 设随机变量 (X,Y) 具有概率密度

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



题 4.31 图

求 $E(X),E(Y),\mathrm{Cov}(X,Y)$.

解　注意到 $f(x,y)$ 只在区域 G: $\{(x,y)\,|\,|y|<x,0<x<1\}$(题 4.31 图) 上不等于零,故有

$$E(X)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)\mathrm{d}x\mathrm{d}y=\iint\limits_{G}x\mathrm{d}x\mathrm{d}y$$

$$=\int_0^1\mathrm{d}x\int_{-x}^{x}x\mathrm{d}y=\int_0^1 2x^2\mathrm{d}x=\frac{2}{3},$$

$$E(Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)\mathrm{d}x\mathrm{d}y=\iint\limits_{G}y\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_{-x}^{x}y\mathrm{d}y=0,$$

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)\mathrm{d}x\mathrm{d}y=\iint\limits_{G}xy\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_{-x}^{x}xy\mathrm{d}y=0,$$

$$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=0.$$

32. 设随机变量 (X,Y) 具有概率密度

$$f(x,y)=\begin{cases}\dfrac{1}{8}(x+y), & 0\leqslant x\leqslant 2,0\leqslant y\leqslant 2,\\0, & \text{其他}.\end{cases}$$

求 $E(X),E(Y),\mathrm{Cov}(X,Y),\rho_{XY},D(X+Y)$.

解　注意到 $f(x,y)$ 在区域 G:$\{(x,y)\,|\,0<x<2,0<y<2\}$ 上不等于零,故有

$$E(X)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)\mathrm{d}x\mathrm{d}y=\int_0^2\mathrm{d}x\int_0^2\frac{x}{8}(x+y)\mathrm{d}y$$

$$=\int_0^2\frac{x}{8}\left(xy+\frac{1}{2}y^2\right)\Big|_0^2\mathrm{d}x=\int_0^2\frac{x}{4}(x+1)\mathrm{d}x=\frac{7}{6},$$

$$E(X^2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x^2f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^2\mathrm{d}x\int_0^2\frac{x^2}{8}(x+y)\mathrm{d}y$$

$$=\frac{1}{8}\int_0^2 x^2\left(xy+\frac{1}{2}y^2\right)\Big|_0^2\mathrm{d}x=\frac{1}{4}\int_0^2(x^3+x^2)\mathrm{d}x=\frac{5}{3},$$

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)\mathrm{d}x\mathrm{d}y=\int_0^2\mathrm{d}x\int_0^2\frac{xy}{8}(x+y)\mathrm{d}y$$

$$=\frac{1}{4}\int_0^2\left(x^2+\frac{4x}{3}\right)\mathrm{d}x=\frac{4}{3}.$$

由 x,y 在 $f(x,y)$ 的表达式中的对称性(即在表达式 $f(x,y)$ 中将 x 和 y 互换,表达式不变),得知

$$E(Y)=E(X)=\frac{7}{6},\quad E(Y^2)=E(X^2)=\frac{5}{3},$$

且有 $$D(Y)=D(X)=E(X^2)-[E(X)]^2=\frac{5}{3}-\left(\frac{7}{6}\right)^2=\frac{11}{36}.$$

而 $$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=\frac{4}{3}-\frac{49}{36}=-\frac{1}{36},$$

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=-\frac{1}{11},$$

$$D(X+Y)=D(X)+D(Y)+2\mathrm{Cov}(X,Y)=\frac{5}{9}.$$

33. 设随机变量 $X\sim N(\mu,\sigma^2)$，$Y\sim N(\mu,\sigma^2)$，且设 X,Y 相互独立，试求 $Z_1=\alpha X+\beta Y$ 和 $Z_2=\alpha X-\beta Y$ 的相关系数(其中 α,β 是不为零的常数).

解法(i) $$\begin{aligned}\mathrm{Cov}(Z_1,Z_2)&=\mathrm{Cov}(\alpha X+\beta Y,\alpha X-\beta Y)\\&=\alpha^2\mathrm{Cov}(X,X)-\alpha\beta\mathrm{Cov}(X,Y)+\alpha\beta\mathrm{Cov}(Y,X)\\&\quad-\beta^2\mathrm{Cov}(Y,Y)\\&=\alpha^2D(X)-\beta^2D(Y)=(\alpha^2-\beta^2)\sigma^2,\end{aligned}$$

而 $$\begin{aligned}D(Z_1)&=D(\alpha X+\beta Y)\\&=\alpha^2D(X)+\beta^2D(Y)+2\mathrm{Cov}(\alpha X,\beta Y)=(\alpha^2+\beta^2)\sigma^2,\\D(Z_2)&=D(\alpha X-\beta Y)\\&=\alpha^2D(X)+\beta^2D(Y)-2\mathrm{Cov}(\alpha X,\beta Y)=(\alpha^2+\beta^2)\sigma^2,\end{aligned}$$

故 $$\rho_{Z_1Z_2}=\frac{(\alpha^2-\beta^2)\sigma^2}{\sqrt{D(Z_1)D(Z_2)}}=\frac{\alpha^2-\beta^2}{\alpha^2+\beta^2}.$$

解法(ii)

$$\begin{aligned}\mathrm{Cov}(Z_1,Z_2)&=E(Z_1Z_2)-E(Z_1)E(Z_2)\\&=E(\alpha^2X^2-\beta^2Y^2)-[\alpha E(X)+\beta E(Y)][\alpha E(X)-\beta E(Y)]\\&=\alpha^2E(X^2)-\beta^2E(Y^2)-\{\alpha^2[E(X)]^2-\beta^2[E(Y)]^2\}\\&=\alpha^2\{E(X^2)-[E(X)]^2\}-\beta^2\{E(Y^2)-[E(Y)]^2\}\\&=\alpha^2D(X)-\beta^2D(Y)=(\alpha^2-\beta^2)\sigma^2.\end{aligned}$$

$$D(Z_1)=D(\alpha X+\beta Y)=\alpha^2D(X)+\beta^2D(Y)=(\alpha^2+\beta^2)\sigma^2,$$

$$D(Z_2)=D(\alpha X-\beta Y)=\alpha^2D(X)+\beta^2D(Y)=(\alpha^2+\beta^2)\sigma^2,$$

故 $$\rho_{Z_1Z_2}=\frac{(\alpha^2-\beta^2)\sigma^2}{\sqrt{D(Z_1)D(Z_2)}}=\frac{\alpha^2-\beta^2}{\alpha^2+\beta^2}.$$

34. (1) 设随机变量 $W=(aX+3Y)^2$，$E(X)=E(Y)=0$，$D(X)=4$，$D(Y)=16$，$\rho_{XY}=-0.5$. 求常数 a 使 $E(W)$ 为最小，并求 $E(W)$ 的最小值.

(2) 设随机变量 (X,Y) 服从二维正态分布，且有 $D(X)=\sigma_X^2$，$D(Y)=\sigma_Y^2$. 证

明当 $a^2=\sigma_X^2/\sigma_Y^2$ 时，随机变量 $W=X-aY$ 与 $V=X+aY$ 相互独立.

解　(1) $E(W)=E[(aX+3Y)^2]=a^2E(X^2)+6aE(XY)+9E(Y^2)$,

$$E(X^2)=D(X)+[E(X)]^2=4,$$

$$E(Y^2)=D(Y)+[E(Y)]^2=16,$$

$$E(XY)=\mathrm{Cov}(X,Y)+E(X)E(Y)=\rho_{XY}\sqrt{D(X)D(Y)}=-4,$$

故

$$E(W)=4a^2-24a+144=4(a-3)^2+108,$$

故当 $a=3$ 时 $E(W)$ 取最小值，$\min\{E(W)\}=108$.

(2) 因为 (X,Y) 是二维正态变量，而 W 与 V 分别是 X,Y 的线性组合，故由 n 维正态随机变量的性质 3° 知 (W,V) 也是二维正态变量. 现在 $a^2=\sigma_X^2/\sigma_Y^2$，故有

$$\begin{aligned}\mathrm{Cov}(W,V)&=\mathrm{Cov}(X-aY,X+aY)\\&=\mathrm{Cov}(X,X)-a^2\mathrm{Cov}(Y,Y)=\sigma_X^2-a^2\sigma_Y^2=0,\end{aligned}$$

即知 W 与 V 不相关. 又因 (W,V) 是二维正态变量，故知 W 与 V 是相互独立的.

35. 设随机变量 (X,Y) 服从二维正态分布，且 $X\sim N(0,3)$，$Y\sim N(0,4)$，相关系数 $\rho_{XY}=-\dfrac{1}{4}$，试写出 X 和 Y 的联合概率密度.

解　因 $\mu_1=\mu_2=0$，$\sigma_1=\sqrt{3}$，$\sigma_2=2$，$\rho_{XY}=-\dfrac{1}{4}$，故 X 和 Y 的联合概率密度为

$$\begin{aligned}f(x,y)&=\frac{1}{4\sqrt{3}\pi\sqrt{1-1/16}}\exp\left\{\frac{-1}{2(1-1/16)}\left(\frac{x^2}{3}+\frac{xy}{4\sqrt{3}}+\frac{y^2}{4}\right)\right\}\\&=\frac{1}{3\sqrt{5}\pi}\exp\left\{\frac{-8}{15}\left(\frac{x^2}{3}+\frac{xy}{4\sqrt{3}}+\frac{y^2}{4}\right)\right\}.\end{aligned}$$

36. 已知正常男性成人血液中，每一毫升所含白细胞数的均值是 7 300，均方差是 700. 利用切比雪夫不等式估计每毫升含白细胞数在 5 200 ～ 9 400 的概率 p.

解　以 X 表示每毫升含白细胞数，由题设

$$E(X)=\mu=7\,300,\qquad \sqrt{D(X)}=\sigma=700,$$

而概率

$$\begin{aligned}p&=P\{5\,200<X<9\,400\}\\&=P\{-2\,100<X-7\,300<2\,100\}\\&=P\{|X-7\,300|<2\,100\}.\end{aligned}$$

在切比雪夫不等式

$$P\{|X-\mu|<\varepsilon\}\geqslant 1-\frac{\sigma^2}{\varepsilon^2}$$

中，取 $\varepsilon=2\ 100$，此时 $1-\frac{\sigma^2}{\varepsilon^2}=1-\frac{700^2}{2\ 100^2}=\frac{8}{9}$，即知

$$p=P\{|X-7\ 300|<2\ 100\}\geqslant\frac{8}{9}.$$

37. 对于两个随机变量 V,W，若 $E(V^2),E(W^2)$ 存在，证明

$$[E(VW)]^2\leqslant E(V^2)E(W^2). \tag{$*_1$}$$

这一不等式称为柯西-施瓦茨(Cauchy-Schwarz) 不等式.

证 若 $E(V^2)=0$，则 $P\{V=0\}=1$ (因 $E(V^2)=D(V)+[E(V)]^2=0$，得 $D(V)=0$ 且 $E(V)=0$，由方差性质 4° 即得 $P\{V=0\}=1$). 由此 $P\{VW=0\}=1$，因此，$E(VW)=0$，此时不等式($*_1$)得证. 同样对于 $E(W^2)=0$ 时，不等式($*_1$)也成立. 以下设 $E(V^2)>0,E(W^2)>0$. 考虑实变量 t 的函数：

$$q(t)=E[(V+tW)^2]=E(V^2)+2tE(VW)+t^2E(W^2).$$

因为对于任意 t，$E[(V+tW)^2]\geqslant 0$，$E(W^2)>0$，故二次三项式 $q(t)$ 的判别式：

$$\Delta=4[E(VW)]^2-4E(V^2)E(W^2)\leqslant 0,$$

即有

$$[E(VW)]^2\leqslant E(V^2)E(W^2).$$

38. 分位数(分位点).

定义 设连续型随机变量 X 的分布函数为 $F(x)$，概率密度函数为 $f(x)$.

1° 对于任意正数 $\alpha(0<\alpha<1)$，称满足条件

$$P\{X\leqslant x_{\underline{\alpha}}\}=F(x_{\underline{\alpha}})=\int_{-\infty}^{x_{\underline{\alpha}}}f(x)\mathrm{d}x=\alpha$$

的数 $x_{\underline{\alpha}}$ 为此分布的 α **分位数或下** α **分位数**.

2° 对于任意正数 $\alpha(0<\alpha<1)$，称满足条件

$$P\{X>x_\alpha\}=1-F(x_\alpha)=\int_{x_\alpha}^{\infty}f(x)\mathrm{d}x=\alpha$$

的数 x_α 为此分布的**上** α **分位数**.

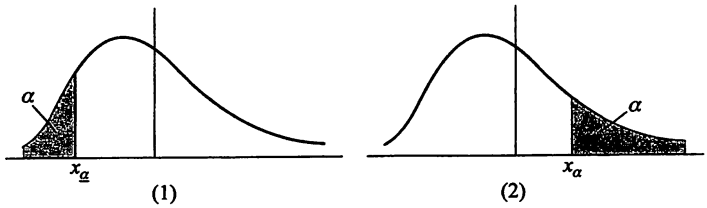
特别，当 $\alpha=0.5$ 时，

$$F(x_{0.5})=F(x_{\underline{0.5}})=\int_{0.5}^{\infty}f(x)\mathrm{d}x=0.5,$$

$x_{0.5}$ 称为此分布的**中位数**.

下 α 分位数 $x_{\underline{\alpha}}$ 将概率密度曲线下的面积分为两部分，左侧的面积恰为 α(见题

4.38 图(1)). 上 α 分位数 x_α 也将概率密度曲线下的面积分为两部分，右侧的面积恰为 α(见题 4.38 图(2)).



题 4.38 图

下 α 分位数与上 α 分位数有以下的关系：

$$x_\alpha = x_{\underline{1-\alpha}}, \quad x_{\underline{\alpha}} = x_{1-\alpha}.$$

类似地，可定义离散型随机变量 X 的分位数.

定义　对于任意正数 $\alpha(0<\alpha<1)$，称满足条件

$$P\{X < x_{\underline{\alpha}}\} \leqslant \alpha \quad 且 \quad P\{X \leqslant x_{\underline{\alpha}}\} \geqslant \alpha$$

的数 $x_{\underline{\alpha}}$ 为此分布的 α **分位数**或**下 α 分位数**.

(1) 设 X 的概率密度为

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

试求 X 的中位数 M.

(2) 设 X 服从柯西分布，其概率密度为

$$f(x) = \frac{b}{\pi[(x-a)^2+b^2]}, \quad b > 0.$$

试求 X 的中位数 M.

解　设 $F(x)$ 为分布函数.

(1) M 应满足 $F(M) = \frac{1}{2}$. 即

$$\frac{1}{2} = F(M) = P\{X \leqslant M\} = \int_0^M 2e^{-2x}\,dx = -e^{-2x}\Big|_0^M = 1 - e^{-2M},$$

故
$$e^{-2M} = \frac{1}{2}, \quad e^{2M} = 2,$$

得
$$M = \frac{1}{2}\ln 2.$$

此即为所求的中位数.

(2) 由

更多电子书，微信：taobao66taobao或淘宝店铺：菁华资料

$$\frac{1}{2}=F(M)=P\{X\leqslant M\}=\int_{-\infty}^{M}\frac{b}{\pi[(x-a)^2+b^2]}\mathrm{d}x$$

$$=\frac{1}{\pi}\arctan\frac{x-a}{b}\Big|_{-\infty}^{M}=\frac{1}{\pi}\arctan\frac{M-a}{b}+\frac{1}{2},$$

得 $M-a=0$，即知中位数 $M=a$.

另外，易知 X 的概率密度 $f(x)$ 的图形关于直线 $x=a$ 是对称的. 即知

$$P\{X\leqslant a\}=\int_{-\infty}^{a}f(x)\mathrm{d}x=\frac{1}{2}.$$

故中位数为 $M=a$.

第五章 大数定律及中心极限定理

1. 据以往经验，某种电器元件的寿命服从均值为 100 h 的指数分布，现随机地取 16 只，设它们的寿命是相互独立的. 求这 16 只元件的寿命的总和大于 1 920 h 的概率.

解 以 X_i $(i=1,2,\cdots,16)$ 记第 i 只元件的寿命，以 T 记16 只元件寿命的总和：$T=\sum_{i=1}^{16}X_i$，按题设 $E(X_i)=100, D(X_i)=100^2$，由中心极限定理知 $\dfrac{T-16\times 100}{\sqrt{16}\sqrt{100^2}}$ 近似地服从 $N(0,1)$ 分布，故所求概率为

$$
\begin{aligned}
P\{T>1\ 920\} &= 1-P\{T\leqslant 1\ 920\} \\
&= 1-P\left\{\frac{T-16\times 100}{\sqrt{16}\sqrt{100^2}}\leqslant \frac{1\ 920-16\times 100}{\sqrt{16}\sqrt{100^2}}\right\} \\
&\approx 1-\Phi\left(\frac{1\ 920-1\ 600}{400}\right)=1-\Phi(0.8) \\
&= 1-0.788\ 1=0.211\ 9.
\end{aligned}
$$

2. (1) 一保险公司有 10 000 个汽车投保人，每个投保人索赔金额的数学期望为 280 美元，标准差为 800 美元，求索赔总金额超过 2 700 000 美元的概率.

(2) 一公司有 50 张签约保险单，各张保险单的索赔金额 $X_i, i=1,2,\cdots,50$(以千美元计)服从韦布尔(Weibull)分布，均值 $E(X_i)=5$，方差 $D(X_i)=6$，求 50 张保险单索赔的合计金额大于 300 的概率(设各保险单索赔金额是相互独立的).

解 (1) 记第 i 人的索赔金额为 X_i，则由已知条件

$$E(X_i)=280,\quad D(X_i)=800^2.$$

要计算

$$p_1=P\left\{\sum_{i=1}^{10\ 000}X_i>2\ 700\ 000\right\},$$

因各投保人索赔金额是独立的，$n=10\ 000$ 很大. 故由中心极限定理，近似地有

$$\bar{X}=\frac{1}{10\ 000}\sum_{i=1}^{10\ 000}X_i\sim N\left(280,\frac{800^2}{100^2}\right),$$

故 $$p_1=P\{\bar{X}>270\}\approx 1-\Phi\left(\frac{270-280}{8}\right)=1-\Phi\left(-\frac{5}{4}\right)$$

$$= \Phi\left(\frac{5}{4}\right) = \Phi(1.25) = 0.894\ 4.$$

(2) $E(X_i) = 5, D(X_i) = 6, n = 50$. 故

$$p = P\left\{\sum_{i=1}^{50} X_i > 300\right\} \approx 1 - \Phi\left(\frac{300 - 50 \times 5}{\sqrt{50 \times 6}}\right)$$

$$= 1 - \Phi\left(\frac{50}{\sqrt{300}}\right) = 1 - \Phi(2.89) = 0.001\ 9.$$

这与情况(1)相反.(1)的概率为0.894 4,表明可能性很大.而(2)表明可能性太小了,大约500次索赔中出现 >300 的只有一次.

3. 计算器在进行加法时,将每个加数舍入最靠近它的整数,设所有舍入误差相互独立且在 $(-0.5, 0.5)$ 上服从均匀分布.

(1) 将1 500个数相加,问误差总和的绝对值超过15的概率是多少?

(2) 最多可有几个数相加使得误差总和的绝对值小于10的概率不小于0.90?

解 设第 k 个加数的舍入误差为 $X_k(k = 1, 2, \cdots, 1\ 500)$,已知 X_k 在 $(-0.5, 0.5)$ 上服从均匀分布,故知 $E(X_k) = 0, D(X_k) = \frac{1}{12}$.

(1) 记 $X = \sum_{k=1}^{1\ 500} X_k$,由中心极限定理,当 n 充分大时有近似公式

$$P\left\{\frac{\sum_{k=1}^{1\ 500} X_k - 1\ 500 \times 0}{\sqrt{1\ 500}\sqrt{1/12}} \leqslant x\right\} \approx \Phi(x).$$

于是

$$P\{|X| > 15\} = 1 - P\{|X| \leqslant 15\} = 1 - P\{-15 \leqslant X \leqslant 15\}$$

$$= 1 - P\left\{\frac{-15 - 0}{\sqrt{1\ 500}\sqrt{1/12}} \leqslant \frac{X - 0}{\sqrt{1\ 500}\sqrt{1/12}} \leqslant \frac{15 - 0}{\sqrt{1\ 500}\sqrt{1/12}}\right\}$$

$$\approx 1 - \left[\Phi\left(\frac{15}{\sqrt{1\ 500}\sqrt{1/12}}\right) - \Phi\left(\frac{-15}{\sqrt{1\ 500}\sqrt{1/12}}\right)\right]$$

$$= 1 - \left[2\Phi\left(\frac{15}{\sqrt{1\ 500/12}}\right) - 1\right] = 1 - [2\Phi(1.342) - 1]$$

$$= 2(1 - 0.909\ 9) = 0.180\ 2.$$

即误差总和的绝对值超过15的概率近似地为0.180 2.

(2) 设最多有 n 个数相加,使误差总和 $Y = \sum_{k=1}^{n} X_k$ 符合要求,即要确定 n,使

$$P\{|Y|<10\}\geqslant 0.90.$$

由中心极限定理，当 n 充分大时有近似公式

$$P\left\{\frac{Y-0}{\sqrt{n}\ \sqrt{1/12}}\leqslant x\right\}\approx \Phi(x).$$

于是 $P\{|Y|<10\}=P\{-10<Y<10\}$

$$=P\left\{\frac{-10}{\sqrt{n}\ \sqrt{1/12}}<\frac{Y}{\sqrt{n}\ \sqrt{1/12}}<\frac{10}{\sqrt{n}\ \sqrt{1/12}}\right\}$$

$$\approx \Phi\left(\frac{10}{\sqrt{n/12}}\right)-\Phi\left(\frac{-10}{\sqrt{n/12}}\right)=2\Phi\left(\frac{10}{\sqrt{n/12}}\right)-1.$$

因而 n 需满足

$$2\Phi\left(\frac{10}{\sqrt{n/12}}\right)-1\geqslant 0.90,$$

亦即 n 需满足

$$\Phi\left(\frac{10}{\sqrt{n/12}}\right)\geqslant 0.95=\Phi(1.645),$$

即 n 应满足

$$\frac{10}{\sqrt{n/12}}\geqslant 1.645,$$

由此得 $$n\leqslant 443.46.$$

因 n 为正整数，因而所求的 n 为 443. 故最多只能有 443 个数加在一起，才能使得误差总和的绝对值小于 10 的概率不小于 0.90.

4. 设各零件的质量都是随机变量，它们相互独立，且服从相同的分布，其数学期望为 0.5 kg，均方差为 0.1 kg，问 5 000 个零件的总质量超过 2 510 kg 的概率是多少？

解　以 $X_i(i=1,2,\cdots,5\ 000)$ 记第 i 个零件的质量，以 W 记 5 000 个零件的总质量：$W=\sum_{i=1}^{5\ 000}X_i$. 按题设 $E(X_i)=0.5$，$D(X_i)=0.1^2$，由中心极限定理，可知 $\dfrac{W-5\ 000\times 0.5}{\sqrt{5\ 000}\times 0.1}$ 近似地服从 $N(0,1)$ 分布，故所求概率为

$$P\{W>2\ 510\}=1-P\{W\leqslant 2\ 510\}$$

$$=1-P\left\{\frac{W-5\ 000\times 0.5}{\sqrt{5\ 000}\times 0.1}\leqslant\frac{2\ 510-5\ 000\times 0.5}{\sqrt{5\ 000}\times 0.1}\right\}$$

$$\approx 1-\Phi\left(\frac{2\ 510-5\ 000\times 0.5}{\sqrt{5\ 000}\times 0.1}\right)=1-\Phi(\sqrt{2})$$

$$= 1 - 0.9213 = 0.0787.$$

5. 有一批建筑房屋用的木柱，其中80%的长度不小于3 m，现从这批木柱中随机地取100根，求其中至少有30根短于3 m的概率.

解 按题意，可认为100根木柱是从为数甚多的木柱中抽取得到的，因而可当作放回抽样来看待.将检查一根木柱看它是否短于3 m看成是一次试验，检查100根木柱相当于做100重伯努利试验.以X记被抽取的100根木柱中长度短于3 m的根数，则$X \sim b(100,0.2)$. 于是由教材第五章 §2 定理3得

$$\begin{aligned}
P\{X \geqslant 30\} &= P\{30 \leqslant X < \infty\} \\
&= P\left\{\frac{30-100\times0.2}{\sqrt{100\times0.2\times0.8}} \leqslant \frac{X-100\times0.2}{\sqrt{100\times0.2\times0.8}} < \frac{\infty-100\times0.2}{\sqrt{100\times0.2\times0.8}}\right\} \\
&\approx \Phi(\infty) - \Phi\left(\frac{30-20}{\sqrt{16}}\right) \\
&= 1-\Phi(2.5) = 1-0.9938 = 0.0062.
\end{aligned}$$

本题也可以这样做，引入随机变量：

$$X_k = \begin{cases} 1, & \text{若第 } k \text{ 根木柱短于 } 3\text{ m}, \\ 0, & \text{若第 } k \text{ 根木柱不短于 } 3\text{ m}, \end{cases} \quad k = 1,2,\cdots,100.$$

于是$E(X_k) = 0.2, D(X_k) = 0.2 \times 0.8$. 以$X$表示100根木柱中短于3 m的根数，则$X = \sum_{k=1}^{100} X_k$. 由中心极限定理有

$$\begin{aligned}
P\{X \geqslant 30\} &= P\{30 \leqslant X < \infty\} \\
&= P\left\{\frac{30-100\times0.2}{\sqrt{100}\sqrt{0.2\times0.8}} \leqslant \frac{\sum_{k=1}^{100} X_k - 100\times0.2}{\sqrt{100}\sqrt{0.2\times0.8}} < \frac{\infty-100\times0.2}{\sqrt{100}\sqrt{0.2\times0.8}}\right\} \\
&\approx \Phi(\infty) - \Phi\left(\frac{30-20}{\sqrt{16}}\right) = 1-\Phi(2.5) = 0.0062.
\end{aligned}$$

6. 一工人修理一台机器需两个阶段，第一阶段所需时间(以h计)服从均值为0.2的指数分布，第二阶段所需时间服从均值为0.3的指数分布，且与第一阶段独立.现有20台机器需要修理，求他在8 h内完成的概率.

解 设修理第i $(i = 1,2,\cdots,20)$台机器，第一阶段耗时为X_i，第二阶段为Y_i，则共耗时$Z_i = X_i + Y_i$，今已知$E(X_i) = 0.2, E(Y_i) = 0.3$，故$E(Z_i) = 0.5$. $D(Z_i) = D(X_i) + D(Y_i) = 0.2^2 + 0.3^2 = 0.13$. 20台机器需要修理的时间可认为近似服从正态分布，即有

$$\sum_{i=1}^{20} Z_i \sim N(20\times 0.5, 20\times 0.13) = N(10, 2.6).$$

所求概率为

$$p = P\left\{\sum_{i=1}^{20} Z_i \leqslant 8\right\} \approx \Phi\left(\frac{8-20\times 0.5}{\sqrt{20\times 0.13}}\right)$$
$$= \Phi\left(-\frac{2}{1.612\,5}\right) = \Phi(-1.24) = 0.107\,5,$$

即不大可能在 8 h 内完成全部工作.

7. 一食品店有三种蛋糕出售，由于售出哪一种蛋糕是随机的，因而售出一只蛋糕的价格是一个随机变量，它取 1 元、1.2 元、1.5 元各个值的概率分别为 0.3，0.2，0.5. 若售出 300 只蛋糕.

(1) 求收入至少 400 元的概率.

(2) 求售出价格为1.2 元的蛋糕多于 60 只的概率.

解　设第 i 只蛋糕的价格为 $X_i, i=1,2,\cdots,300$，则 X_i 的分布律为

X_i	1	1.2	1.5
p_k	0.3	0.2	0.5

由此得

$$E(X_i) = 1\times 0.3 + 1.2\times 0.2 + 1.5\times 0.5 = 1.29,$$
$$E(X_i^2) = 1^2\times 0.3 + 1.2^2\times 0.2 + 1.5^2\times 0.5 = 1.713,$$

故　$$D(X_i) = E(X_i^2) - [E(X_i)]^2 = 0.048\,9.$$

(1) 以 X 表示这天的总收入，则 $X=\sum_{i=1}^{300} X_i$，由中心极限定理得

$$P\{X \geqslant 400\} = P\{400 \leqslant X < \infty\}$$
$$= P\left\{\frac{400-300\times 1.29}{\sqrt{300}\sqrt{0.048\,9}} \leqslant \frac{\sum_{i=1}^{300} X_i - 300\times 1.29}{\sqrt{300}\sqrt{0.048\,9}} < \frac{\infty - 300\times 1.29}{\sqrt{300}\sqrt{0.048\,9}}\right\}$$
$$\approx 1 - \Phi(3.39) = 1 - 0.999\,7 = 0.000\,3.$$

(2) 以 Y 记 300 只蛋糕中售价为 1.2 元的蛋糕的只数，于是 $Y \sim b(300, 0.2)$. $E(Y) = 300\times 0.2$，$D(Y) = 300\times 0.2\times 0.8$，由棣莫弗-拉普拉斯定理得

$$P\{Y > 60\} = 1 - P\{Y \leqslant 60\}$$

$$= 1 - P\left\{\frac{Y - 300 \times 0.2}{\sqrt{300 \times 0.2 \times 0.8}} \leqslant \frac{60 - 300 \times 0.2}{\sqrt{300 \times 0.2 \times 0.8}}\right\}$$

$$\approx 1 - \Phi\left(\frac{60 - 300 \times 0.2}{\sqrt{300 \times 0.2 \times 0.8}}\right) = 1 - \Phi(0) = 0.5.$$

8. 一复杂的系统由100个相互独立起作用的部件所组成，在整个运行期间每个部件损坏的概率为0.1. 为了使整个系统起作用，至少必须有85个部件正常工作，求整个系统起作用的概率.

解 将观察一个部件是否正常工作看成是一次试验，由于各部件是否正常工作是相互独立的，因而观察100个部件是否正常工作是做100重伯努利试验，以X表示100个部件中正常工作的部件数，则$X \sim b(100, 0.9)$，按题意需求概率$P\{X \geqslant 85\}$，由棣莫弗-拉普拉斯定理知$\dfrac{X - 100 \times 0.9}{\sqrt{100 \times 0.9 \times 0.1}}$近似地服从标准正态分布$N(0,1)$，故所求概率为

$$P\{X \geqslant 85\} = P\{85 \leqslant X < \infty\}$$

$$= P\left\{\frac{85 - 100 \times 0.9}{\sqrt{100 \times 0.9 \times 0.1}} \leqslant \frac{X - 100 \times 0.9}{\sqrt{100 \times 0.9 \times 0.1}} < \frac{\infty - 100 \times 0.9}{\sqrt{100 \times 0.9 \times 0.1}}\right\}$$

$$\approx 1 - \Phi\left(-\frac{5}{3}\right) = 0.952\,5.$$

9. 已知在某十字路口，一周事故发生数的数学期望为2.2，标准差为1.4.

(1) 以$\overline{X}$表示一年(以52周计)此十字路口事故发生数的算术平均，试用中心极限定理求$\overline{X}$的近似分布，并求$P\{\overline{X} < 2\}$.

(2) 求一年事故发生数小于100的概率.

解 (1) $E(\overline{X}) = E(X) = 2.2, \quad D(\overline{X}) = \dfrac{D(X)}{52} = \dfrac{1.4^2}{52}$,

由中心极限定理，可认为$\overline{X} \sim N(2.2, 1.4^2/52)$.

$$P\{\overline{X} < 2\} \approx \Phi\left(\frac{2 - 2.2}{1.4/\sqrt{52}}\right) = \Phi\left(\frac{-0.2 \times \sqrt{52}}{1.4}\right) = \Phi(-1.030)$$

$$= 1 - \Phi(1.030) = 1 - 0.848\,5 = 0.151\,5.$$

(2) 一年52周，设各周事故发生数为$X_1, X_2, \cdots, X_{52}$. 则需计算$p = P\left\{\sum_{i=1}^{52} X_i < 100\right\}$，即$P\{52\overline{X} < 100\}$. 用中心极限定理可知所求概率为

$$p = P\left\{52\overline{X} < 100\right\} = P\left\{\overline{X} < \frac{100}{52}\right\} \approx \Phi\left[\frac{\left(\frac{100}{52} - 2.2\right)\sqrt{52}}{1.4}\right]$$

$$= \Phi(-1.426) = 1 - 0.9230 = 0.0770.$$

10. 某种小汽车氧化氮的排放量的数学期望为 0.9 g/km，标准差为 1.9 g/km，某汽车公司有这种小汽车 100 辆，以 $\overline{X}$ 表示这些车辆氧化氮排放量的算术平均，问当 L 为何值时 $\overline{X} > L$ 的概率不超过 0.01？

解　设以 $X_i(i = 1,2,\cdots,100)$ 表示第 i 辆小汽车氧化氮的排放量，则

$$\overline{X} = \frac{1}{100}\sum_{i=1}^{100} X_i.$$

由已知条件 $E(X_i) = 0.9, D(X_i) = 1.9^2$ 得

$$E(\overline{X}) = 0.9,\quad D(\overline{X}) = \frac{1.9^2}{100}.$$

各辆汽车氧化氮的排放量相互独立，故可认为近似地有

$$\overline{X} \sim N\left(0.9, \frac{1.9^2}{100}\right).$$

需要计算的是满足

$$P\left\{\overline{X} > L\right\} \leqslant 0.01$$

的最小值 L. 由中心极限定理

$$P\left\{\overline{X} > L\right\} = P\left\{\frac{\overline{X} - 0.9}{0.19} > \frac{L - 0.9}{0.19}\right\} \leqslant 0.01.$$

L 应为满足

$$1 - \Phi\left(\frac{L - 0.9}{0.19}\right) \leqslant 0.01$$

的最小值，即

$$\Phi\left(\frac{L - 0.9}{0.19}\right) \geqslant 0.99 = \Phi(2.33),$$

即
$$\frac{L - 0.9}{0.19} \geqslant 2.33,$$

故
$$L \geqslant 0.9 + 0.19 \times 2.33 = 1.3427,$$

应取 $L = 1.3427$ g/km.

11. 随机地选取两组学生，每组 80 人，分别在两个实验室里测量某种化合物的 pH. 各人测量的结果是随机变量，它们相互独立，服从同一分布，数学期望为 5，方差为 0.3，以 $\overline{X}, \overline{Y}$ 分别表示第一组和第二组所得结果的算术平均.

(1) 求$P\{4.9<\bar{X}<5.1\}$.

(2) 求 $P\{-0.1<\bar{X}-\bar{Y}<0.1\}$.

解 由题设 $E(\bar{X})=5, D(\bar{X})=D(\bar{Y})=0.3/80$.

(1) 由中心极限定理知 $\bar{X}$ 近似服从 $N(5,0.3/80)$，故

$$\begin{aligned}P\{4.9<\bar{X}<5.1\}&=P\left\{\frac{4.9-5}{\sqrt{0.3/80}}<\frac{\bar{X}-5}{\sqrt{0.3/80}}<\frac{5.1-5}{\sqrt{0.3/80}}\right\}\\&\approx\Phi\left(\frac{5.1-5}{\sqrt{0.3/80}}\right)-\Phi\left(\frac{4.9-5}{\sqrt{0.3/80}}\right)\\&=2\Phi(1.63)-1=2\times0.948\,4-1=0.896\,8.\end{aligned}$$

(2) 因 $E(\bar{X}-\bar{Y})=E(\bar{X})-E(\bar{Y})=0, D(\bar{X}-\bar{Y})=D(\bar{X})+D(\bar{Y})=0.3/40$，由中心极限定理

$$\begin{aligned}P\{-0.1<\bar{X}-\bar{Y}<0.1\}&=P\left\{\frac{-0.1-0}{\sqrt{0.3/40}}<\frac{(\bar{X}-\bar{Y})-0}{\sqrt{0.3/40}}<\frac{0.1-0}{\sqrt{0.3/40}}\right\}\\&\approx\Phi\left(\frac{0.1-0}{\sqrt{0.3/40}}\right)-\Phi\left(\frac{-0.1-0}{\sqrt{0.3/40}}\right)\\&=2\Phi(1.15)-1=2\times0.874\,9-1=0.749\,8.\end{aligned}$$

12. 一公寓有 200 户住户，一户住户拥有汽车辆数 X 的分布律为

X	0	1	2
p_k	0.1	0.6	0.3

问需要多少车位，才能使每辆汽车都具有一个车位的概率至少为 0.95?

解 设需要车位数为 n，且设第 i $(i=1,2,\cdots,200)$ 户有车辆数为 X_i，则由 X_i 的分布律知

$$E(X_i)=0\times0.1+1\times0.6+2\times0.3=1.2,$$

$$E(X_i^2)=0^2\times0.1+1^2\times0.6+2^2\times0.3=1.8,$$

故 $$D(X_i)=E(X_i^2)-[E(X_i)]^2=1.8-1.2^2=0.36.$$

因共有 200 户，各户占有车位数相互独立. 从而近似地有

$$\sum_{i=1}^{200}X_i\sim N(200\times1.2,\ 200\times0.36).$$

今要求车位数 n 满足

$$0.95\leqslant P\left\{\sum_{i=1}^{200}X_i\leqslant n\right\},$$

由正态近似知，上式中 n 应满足

$$0.95 \leqslant \Phi\left(\frac{n-200\times 1.2}{\sqrt{200\times 0.36}}\right)=\Phi\left(\frac{n-240}{\sqrt{72}}\right),$$

因 $0.95=\Phi(1.645)$，从而由 $\Phi(x)$ 的单调性知

$$\frac{n-240}{\sqrt{72}} \geqslant 1.645,$$

故 $n \geqslant 240+1.645\times\sqrt{72}=253.96$. 由此知至少需 254 个车位.

13. 某种电子器件的寿命(以 h 计)具有数学期望 μ(未知)，方差 $\sigma^2=400$. 为了估计 μ，随机地取 n 只这种器件，在时刻 $t=0$ 投入测试(测试是相互独立的)直到失效，测得其寿命为 $X_1,X_2,\cdots,X_n$，以 $\bar{X}=\frac{1}{n}\sum\limits_{i=1}^{n}X_i$ 作为 μ 的估计，为使 $P\{|\bar{X}-\mu|<1\}\geqslant 0.95$，问 n 至少为多少?

解　由教材第五章 §2 定理 1 可知，当 n 充分大时，

$$\frac{\sum\limits_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}=\frac{\frac{1}{n}\sum\limits_{i=1}^{n}X_i-\mu}{\sigma/\sqrt{n}}\overset{\text{近似地}}{\sim}N(0,1),$$

即

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\overset{\text{近似地}}{\sim}N(0,1).$$

由题设 $D(X_i)=400\ (i=1,2,\cdots,n)$，即有 $\sigma=\sqrt{400}$，于是 $\dfrac{\bar{X}-\mu}{\sqrt{400}/\sqrt{n}}=\dfrac{\bar{X}-\mu}{20/\sqrt{n}}$ 近似地服从 $N(0,1)$ 分布，即有

$$\begin{aligned}P\{|\bar{X}-\mu|<1\}&=P\{-1<\bar{X}-\mu<1\}\\&=P\left\{\frac{-1}{20/\sqrt{n}}<\frac{\bar{X}-\mu}{20/\sqrt{n}}<\frac{1}{20/\sqrt{n}}\right\}\\&\approx\Phi\left(\frac{1}{20/\sqrt{n}}\right)-\Phi\left(\frac{-1}{20/\sqrt{n}}\right)=2\Phi\left(\frac{1}{20/\sqrt{n}}\right)-1.\end{aligned}$$

现在要求 $P\{|\bar{X}-\mu|<1\}\geqslant 0.95$，即要求

$$2\Phi\left(\frac{1}{20/\sqrt{n}}\right)-1\geqslant 0.95,$$

亦即要求

$$\Phi\left(\frac{1}{20/\sqrt{n}}\right) \geqslant 0.975 = \Phi(1.96),$$

故需要

$$\frac{1}{20/\sqrt{n}} \geqslant 1.96,$$

即

$$n \geqslant (20 \times 1.96)^2 = 1\ 536.64.$$

因 n 为正整数，故 n 至少为 1 537.

14. 某药厂断言，该厂生产的某种药品对于医治一种疑难血液病的治愈率为 0.8，医院任意抽查 100 个服用此药品的患者，若其中多于 75 人治愈，就接受此断言，否则就拒绝此断言.

(1) 若实际上此药品对这种疾病的治愈率是 0.8，问接受这一断言的概率是多少？

(2) 若实际上此药品对这种疾病的治愈率为 0.7，问接受这一断言的概率是多少？

解 由药厂断言来看 100 人中治愈人数 $X \sim b(100,0.8)$.

(1) 在治愈率与实际情况相符合条件下，接受药厂断言的概率即为 $P\{X > 75\}$. 由中心极限定理知近似地有 $X \sim N(100 \times 0.8,\ 100 \times 0.8 \times 0.2) = N(80,4^2)$，于是所求概率

$$\begin{aligned} p_1 &= P\{X > 75\} \approx 1 - \Phi\left(\frac{75-80}{4}\right) = 1 - \Phi\left(\frac{-5}{4}\right) \\ &= \Phi(1.25) = 0.894\ 4. \end{aligned}$$

(2) 若实际上治疗率为 0.7，即 $X \sim b(100,0.7)$，则治愈人数 X 近似地服从正态分布，即有

$$X \sim N(100 \times 0.7, 100 \times 0.7 \times 0.3).$$

所求概率

$$\begin{aligned} p_2 &= P\{X > 75\} \approx 1 - \Phi\left(\frac{75 - 100 \times 0.7}{\sqrt{100 \times 0.7 \times 0.3}}\right) \\ &= 1 - \Phi\left(\frac{5}{\sqrt{21}}\right) = 1 - \Phi(1.09) \\ &= 1 - 0.862\ 1 = 0.137\ 9. \end{aligned}$$

第六章　样本及抽样分布

1. 在总体 $N(52,6.3^2)$ 中随机抽取一容量为 36 的样本，求样本均值 $\bar{X}$ 落在 50.8 到 53.8 之间的概率.

解　$n=36,\bar{X}=\frac{1}{36}\sum_{i=1}^{36}X_i$，因总体 $X\sim N(52,6.3^2)$，故 $E(\bar{X})=52,D(\bar{X})=6.3^2/36=1.05^2,\bar{X}\sim N(52,1.05^2)$. 从而

$$\begin{aligned}P\{50.8<\bar{X}<53.8\}&=P\left\{\frac{50.8-52}{1.05}<\frac{\bar{X}-52}{1.05}<\frac{53.8-52}{1.05}\right\}\\&=\Phi\left(\frac{53.8-52}{1.05}\right)-\Phi\left(\frac{50.8-52}{1.05}\right)\\&=\Phi(1.71)-\Phi(-1.14)\\&=\Phi(1.71)+\Phi(1.14)-1=0.8293.\end{aligned}$$

2. 在总体 $N(12,4)$ 中随机抽一容量为 5 的样本 X_1,X_2,X_3,X_4,X_5.

(1) 求样本均值与总体均值之差的绝对值大于 1 的概率.

(2) 求概率 $P\{\max\{X_1,X_2,X_3,X_4,X_5\}>15\},P\{\min\{X_1,X_2,X_3,X_4,X_5\}<10\}$.

解　(1) $\bar{X}=\frac{1}{5}\sum_{i=1}^{5}X_i$，因总体 $X\sim N(12,4)$，故 $\bar{X}\sim N\left(12,\frac{4}{5}\right)$，从而

$$\begin{aligned}P\{|\bar{X}-12|>1\}&=1-P\{|\bar{X}-12|\leqslant 1\}\\&=1-P\{-1\leqslant\bar{X}-12\leqslant 1\}\\&=1-P\left\{-\frac{1}{\sqrt{4/5}}\leqslant\frac{\bar{X}-12}{\sqrt{4/5}}\leqslant\frac{1}{\sqrt{4/5}}\right\}\\&=1-\left[\Phi\left(\frac{1}{\sqrt{4/5}}\right)-\Phi\left(\frac{-1}{\sqrt{4/5}}\right)\right]=2-2\Phi(1.12)\\&=2(1-0.8686)=0.2628.\end{aligned}$$

(2) 因 X_i 的分布函数为 $\Phi\left(\frac{x-12}{2}\right)$，故 $M=\max\{X_1,X_2,X_3,X_4,X_5\}$ 的分布函数为

$$F_M(x)=\left[\Phi\left(\frac{x-12}{2}\right)\right]^5,$$

因而

$$\begin{aligned}&P\{\max\{X_1,X_2,X_3,X_4,X_5\}>15\}\\&\quad=P\{M>15\}=1-P\{M\leqslant 15\}=1-F_M(15)\\&\quad=1-\left[\Phi\left(\frac{15-12}{2}\right)\right]^5=1-0.933\ 2^5=0.292\ 3.\end{aligned}$$

记 $N=\min\{X_1,X_2,X_3,X_4,X_5\}$，则 N 的分布函数为

$$F_N(x)=1-\left[1-\Phi\left(\frac{x-12}{2}\right)\right]^5,$$

故

$$\begin{aligned}&P\{\min\{X_1,X_2,X_3,X_4,X_5\}<10\}\\&\quad=P\{N<10\}=1-\left[1-\Phi\left(\frac{10-12}{2}\right)\right]^5=1-[1-\Phi(-1)]^5\\&\quad=1-[\Phi(1)]^5=1-0.841\ 3^5=0.578\ 5.\end{aligned}$$

3. 求总体 $N(20,3)$ 的容量分别为 10，15 的两独立样本均值差的绝对值大于 0.3 的概率.

解 将总体 $N(20,3)$ 的容量分别为 10，15 的两独立样本的均值分别记作 $\bar{X},\bar{Y}$，则 $\bar{X}\sim N(20,3/10)$，$\bar{Y}\sim N(20,3/15)$，从而

$$\bar{X}-\bar{Y}\sim N(20-20,3/10+3/15),$$

即

$$\bar{X}-\bar{Y}\sim N(0,1/2),$$

故所求概率为

$$\begin{aligned}p&=P\{|\bar{X}-\bar{Y}|>0.3\}=1-P\{|\bar{X}-\bar{Y}|\leqslant 0.3\}\\&=1-P\left\{\frac{-0.3}{\sqrt{1/2}}\leqslant\frac{\bar{X}-\bar{Y}}{\sqrt{1/2}}\leqslant\frac{0.3}{\sqrt{1/2}}\right\}\\&=2-2\Phi(0.42)=2(1-0.662\ 8)=0.674\ 4.\end{aligned}$$

注意 $D(\bar{X}-\bar{Y})=D(\bar{X})+D(\bar{Y})$.

4. (1) 设样本 $X_1,X_2,\cdots,X_6$ 来自总体 $N(0,1)$，$Y=(X_1+X_2+X_3)^2+(X_4+X_5+X_6)^2$，试确定常数 C 使 CY 服从χ^2分布.

(2) 设样本 $X_1,X_2,\cdots,X_5$ 来自总体 $N(0,1)$，$Y=\dfrac{C(X_1+X_2)}{(X_3^2+X_4^2+X_5^2)^{1/2}}$，试确定常数 C 使 Y 服从 t 分布.

(3) 已知总体 $X\sim t(n)$，求证 $X^2\sim F(1,n)$.

解 (1) 因 $X_1,X_2,\cdots,X_6$ 是总体 $N(0,1)$ 的样本，故

$$X_1+X_2+X_3\sim N(0,3),\quad X_4+X_5+X_6\sim N(0,3),$$

且两者相互独立. 因此

$$\frac{X_1+X_2+X_3}{\sqrt{3}}\sim N(0,1),\quad \frac{X_4+X_5+X_6}{\sqrt{3}}\sim N(0,1),$$

且两者相互独立. 按χ^2分布的定义

$$\frac{(X_1+X_2+X_3)^2}{3}+\frac{(X_4+X_5+X_6)^2}{3}\sim\chi^2(2),$$

即$\frac{1}{3}Y\sim\chi^2(2)$,即知 $C=\frac{1}{3}$.

(2) 因 $X_1,X_2,\cdots,X_5$ 是总体 $N(0,1)$的样本,故 $X_1+X_2\sim N(0,2)$,即有

$$\frac{X_1+X_2}{\sqrt{2}}\sim N(0,1).$$

而

$$X_3^2+X_4^2+X_5^2\sim\chi^2(3).$$

且$\frac{X_1+X_2}{\sqrt{2}}$与 $X_3^2+X_4^2+X_5^2$ 相互独立,于是

$$\frac{(X_1+X_2)/\sqrt{2}}{\sqrt{(X_3^2+X_4^2+X_5^2)/3}}=\sqrt{\frac{3}{2}}\frac{X_1+X_2}{(X_3^2+X_4^2+X_5^2)^{1/2}}\sim t(3),$$

因此所求的常数 $C=\sqrt{\frac{3}{2}}$.

(3) 按定义总体 $X\sim t(n)$,故 X 可表示成

$$X=\frac{Z}{\sqrt{Y/n}},$$

其中,$Y\sim\chi^2(n)$,$Z\sim N(0,1)$且 Z 与 Y 相互独立,从而

$$X^2=\frac{Z^2}{Y/n}.$$

由于 $Z\sim N(0,1)$,$Z^2\sim\chi^2(1)$,上式右端分子 $Z^2\sim\chi^2(1)$,分母中 $Y\sim\chi^2(n)$,又由 Z 与 Y 相互独立,知 Z^2 与 Y 相互独立. 按 F 分布的定义得

$$X^2\sim F(1,n).$$

5. (1) 已知某种能力测试的得分服从正态分布 $N(\mu,\sigma^2)$,随机取 10 个人参与这一测试. 求他们得分的联合概率密度,并求这 10 个人得分的平均值小于 μ 的概率.

(2) 在(1)中设 $\mu=62$,$\sigma^2=25$,若得分超过 70 就能得奖,求至少有一人得奖的概率.

解　(1) 10 个人的得分分别记为 $X_1,X_2,\cdots,X_{10}$. 它们的联合概率密度为

$$f(x_1,x_2,\cdots,x_{10})=\prod_{i=1}^{10}\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-(x_i-\mu)^2}{2\sigma^2}},$$

$$\bar{X}=\frac{1}{10}\sum_{i=1}^{10}X_i\sim N\left(\mu,\frac{\sigma^2}{10}\right),$$

$$P\{\bar{X}<\mu\}=\Phi\left(\frac{\mu-\mu}{\sigma/\sqrt{10}}\right)=\Phi(0)=\frac{1}{2}.$$

(2) 若一人得奖的概率为 p，则得奖人数 $Y\sim b(10,p)$. 此处 p 是随机选取一人，其考分 X 在 70 分以上的概率. 因 $X\sim N(62,25)$，故

$$p=P\{X>70\}=1-P\{X\leqslant 70\}=1-\Phi\left(\frac{70-62}{\sqrt{25}}\right)$$
$$=1-\Phi(1.6)=1-0.945\ 2=0.054\ 8.$$

至少一人得奖的概率为

$$P\{Y\geqslant 1\}=1-0.945\ 2^{10}=0.431.$$

6. 设总体 $X\sim b(1,p)$，$X_1,X_2,\cdots,X_n$ 是来自 X 的样本.

(1) 求 $(X_1,X_2,\cdots,X_n)$ 的分布律.

(2) 求 $\sum\limits_{i=1}^{n}X_i$ 的分布律.

(3) 求 $E(\bar{X})$，$D(\bar{X})$，$E(S^2)$.

解 (1) 因 $X_1,X_2,\cdots,X_n$ 相互独立，且有 $X_i\sim b(1,p)$，$i=1,2,\cdots,n$，即 X_i 具有分布律 $P\{X_i=x_i\}=p^{x_i}(1-p)^{1-x_i}$，$x_i=0,1$，因此 $(X_1,X_2,\cdots,X_n)$ 的分布律为

$$P\{X_1=x_1,X_2=x_2,\cdots,X_n=x_n\}$$
$$=\prod_{i=1}^{n}P\{X_i=x_i\}=\prod_{i=1}^{n}[p^{x_i}(1-p)^{1-x_i}]=p^{\sum\limits_{i=1}^{n}x_i}(1-p)^{n-\sum\limits_{i=1}^{n}x_i}.$$

(2) 因 $X_1,X_2,\cdots,X_n$ 相互独立，且有 $X_i\sim b(1,p)$，$i=1,2,\cdots,n$，故 $\sum\limits_{i=1}^{n}X_i\sim b(n,p)$，其分布律为

$$P\left\{\sum_{i=1}^{n}X_i=k\right\}=\binom{n}{k}p^k(1-p)^{n-k},\quad k=0,1,2,\cdots,n.$$

(3) 由于总体 $X\sim b(1,p)$，$E(X)=p$，$D(X)=p(1-p)$，故有

$$E(\bar{X})=p,\quad D(\bar{X})=\frac{p(1-p)}{n},\quad E(S^2)=D(X)=p(1-p).$$

7. 设总体 $X\sim\chi^2(n)$，$X_1,X_2,\cdots,X_{10}$ 是来自 X 的样本，求 $E(\bar{X})$，$D(\bar{X})$，$E(S^2)$.

解　因 $E(X)=n, D(X)=2n$，故有

$$E(\bar{X})=n, \quad D(\bar{X})=\frac{2n}{10}=\frac{n}{5},$$

而

$$E(S^2)=D(X)=2n.$$

8. 设总体 $X\sim N(\mu,\sigma^2)$，$X_1,X_2,\cdots,X_{10}$是来自 X 的样本.

(1) 写出 $X_1,X_2,\cdots,X_{10}$的联合概率密度.

(2) 写出 $\bar{X}$ 的概率密度.

解　(1) 由假设 $X_i(i=1,2,\cdots,10)$的概率密度为

$$f_{X_i}(x_i)=\frac{1}{\sqrt{2\pi}\sigma}e^{-(x_i-\mu)^2/(2\sigma^2)},$$

故 $X_1,X_2,\cdots,X_{10}$的联合概率密度为

$$\prod_{i=1}^{10} f_{X_i}(x_i) = \prod_{i=1}^{10}\frac{1}{\sqrt{2\pi}\sigma}e^{-(x_i-\mu)^2/(2\sigma^2)}$$

$$= \frac{1}{(2\pi\sigma^2)^5}e^{-\sum\limits_{i=1}^{10}(x_i-\mu)^2/(2\sigma^2)}.$$

(2) $\bar{X}\sim N\left(\mu,\dfrac{\sigma^2}{10}\right)$，故 $\bar{X}$ 的概率密度为

$$f_{\bar{X}}(x)=\frac{\sqrt{10}}{\sqrt{2\pi}\sigma}e^{-5(x-\mu)^2/\sigma^2}=\frac{\sqrt{5}}{\sqrt{\pi}\sigma}e^{-5(x-\mu)^2/\sigma^2}.$$

9. 设在总体 $N(\mu,\sigma^2)$中抽得一容量为 16 的样本，这里 μ,σ^2 均未知.

(1) 求 $P\{S^2/\sigma^2\leqslant 2.041\}$，其中 S^2 为样本方差.

(2) 求 $D(S^2)$.

解　(1) 因为$\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$，现在 $n=16$，即有$\dfrac{15S^2}{\sigma^2}\sim\chi^2(15)$，故有

$$p= P\{S^2/\sigma^2 \leqslant 2.041\} = P\{15S^2/\sigma^2 \leqslant 15\times 2.041\}$$
$$= P\{15S^2/\sigma^2 \leqslant 30.615\} = 1 - P\{15S^2/\sigma^2 > 30.615\}.$$

查χ^2分布表得$\chi^2_{0.01}(15)=30.577$，从而知 $p=1-0.01=0.99$.

(2) 由 $15S^2/\sigma^2\sim\chi^2(15)$，得

$$D(15S^2/\sigma^2) = 2\times 15 = 30,$$

即

$$\frac{15^2}{\sigma^4}D(S^2) = 30, \quad D(S^2) = \frac{2\sigma^4}{15}.$$

10. 下面列出了 30 个美国 NBA 球员的体重(以磅计，1 磅＝0.454 kg)数据. 这些数据是从美国 NBA 球队 1990—1991 年赛季的花名册中抽样得到的.

225	232	232	245	235	245	270	225	240	240
217	195	225	185	200	220	200	210	271	240
220	230	215	252	225	220	206	185	227	236

(1) 画出这些数据的频率直方图(提示：最大和最小观察值分别为 271 和 185，区间[184.5,271.5]包含所有数据，将整个区间分为 5 等份，为计算方便，将区间调整为(179.5,279.5).

(2) 作出这些数据的箱线图.

解 (1) 最大和最小观察值分别为 271 和 185，考虑到这些数据是将实测数据经四舍五入后得到的，取区间 $I=[184.5,271.5]$使得所有实测数据都落在 I 上. 将区间 I 等分为若干小区间，小区间的个数与数据个数 n 有关，取为$\sqrt{n}$左右为佳. 现在取小区间的个数为 5，于是小区间的长度为$(271.5-184.5)/5=17.4$. 这一长度使用起来不方便. 为此，将区间 I 的下限延伸至 179.5，上限延伸至 279.5. 这样小区间的长度调整为

$$\Delta=(279.5-179.5)/5=20.$$

数出落在每个小区间内的数据的个数 $f_i, i=1,2,3,4,5$，算出数据落在各个小区间的频率 $f_i/n(n=30, i=1,2,3,4,5)$，所得结果列表如下：

组限	频数 f_i	频率 f_i/n	累积频率
179.5～199.5	3	0.10	0.10
199.5～219.5	6	0.20	0.30
219.5～239.5	13	0.43	0.73
239.5～259.5	6	0.20	0.93
259.5～279.5	2	0.07	1.00

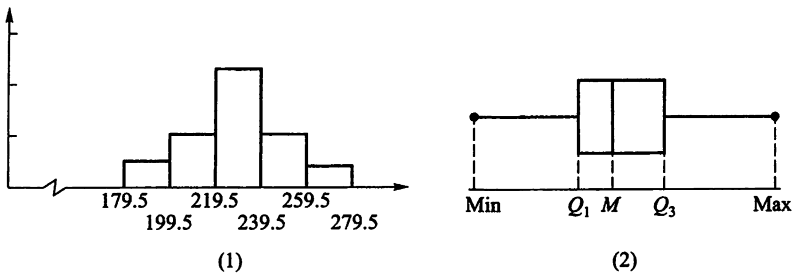
在每个小区间上作以对应的频率除以 Δ 为高(即以$(f_i/n)/\Delta$ 为高)、以小区间为底的小长方形. 小长方形的面积就是$[(f_i/n)/\Delta]\times\Delta=f_i/n$. 画出图形，这就是所求的频率直方图(如题 6.10 图(1)).

(2) 将 $n=30$ 个数据按自小到大的次序排序得到

185	185	195	200	200	206	210	215	217	220
220	220	225	225	225	225	227	230	232	232
235	236	240	240	240	245	245	252	270	271

下面来求第一四分位数 Q_1，中位数 M，第三四分位数 Q_3.

因 $np=30\times0.25=7.5$，故 Q_1 位于左起第$[7.5]+1=8$ 处，即有 $Q_1=215$.



题 6.10 图

因 $np=30\times0.5=15$，故 $M=Q_2$ 是这 30 个数最中间两个数的平均值，即有 $Q_2=M=\frac{1}{2}(225+225)=225$.

因 $np=30\times0.75=22.5$，故 Q_3 位于左起第$[22.5]+1=23$ 处，即有 $Q_3=240$. 又 Min=185，Max=271.

根据 Min，Q_1，M，Q_3，Max 这 5 点作出箱线图如题 6.10 图(2)所示. 从上述两个图能看出数据的分布关于中心线比较对称.

11. 截尾均值　设数据集包含 n 个数据，将这些数据自小到大排序为

$$x_{(1)}\leqslant x_{(2)}\leqslant\cdots\leqslant x_{(n)},$$

删去 $100\alpha\%$个数值小的数，同时删去 $100\alpha\%$个数值大的数，将留下的数据取算术平均，记为 $\bar{x}_\alpha$，即

$$\bar{x}_\alpha=\frac{x_{([n\alpha]+1)}+\cdots+x_{(n-[n\alpha])}}{n-2[n\alpha]}$$

其中$[n\alpha]$是小于或等于 $n\alpha$ 的最大整数(一般取 α 为 0.1～0.2). $\bar{x}_\alpha$ 称为 $100\alpha\%$截尾均值. 例如对于第 10 题中的 30 个数据，取 $\alpha=0.1$，则有$[n\alpha]=[30\times0.1]=3$，得 $100\times0.1\%$截尾均值为

$$\bar{x}_\alpha=\frac{200+200+\cdots+245+245}{30-6}=225.4167.$$

若数据来自某一总体的样本，则 $\bar{x}_\alpha$ 是一个统计量. $\bar{x}_\alpha$ 不受样本的极端值的影响. 截尾均值在实际应用问题中是常会用到的.

试求第 10 题的 30 个数据的 $\alpha=0.2$ 的截尾均值.

解　$\alpha=0.2$，$[n\alpha]=[30\times0.2]=6$，$100\times0.2\%$截尾均值为

$$\bar{x}_\alpha=\frac{210+215+\cdots+240+240}{30-12}=226.3333.$$

第七章 参数估计

1. 随机地取 8 只活塞环，测得它们的直径为(以 mm 计)

$$\begin{array}{cccc} 74.001 & 74.005 & 74.003 & 74.001 \\ 74.000 & 73.998 & 74.006 & 74.002 \end{array}$$

试求总体均值 μ 及方差 σ^2 的矩估计值，并求样本方差 s^2.

解 总体均值 μ 的矩估计值、总体方差 σ^2 的矩估计值分别为

$$\hat{\mu}=\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i,\quad \hat{\sigma^2}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2.$$

由给出的观察值得

$$\begin{aligned}\hat{\mu}&=74+\frac{1}{8}[0.001+0.005+0.003+0.001+0+(-0.002)+0.006+0.002]\\&=74.002,\end{aligned}$$

$$\begin{aligned}\hat{\sigma^2}&=\frac{1}{8}\sum_{i=1}^{8}(x_i-\bar{x})^2\\&=\frac{1}{8}[(-0.001)^2+0.003^2+0.001^2\\&\quad+(-0.001)^2+(-0.002)^2+(-0.004)^2+0.004^2+0^2]\\&=6\times10^{-6},\end{aligned}$$

$$s^2=\frac{1}{7}\sum_{i=1}^{8}(x_i-\bar{x})^2=\frac{8}{7}\hat{\sigma^2}=\frac{8}{7}\times6\times10^{-6}=6.86\times10^{-6}.$$

事实上，只需将观察值输入具有统计功能的计算器，就能直接读出 $\bar{x}$，$\sqrt{\hat{\sigma^2}}$ 和 s. 读者应掌握计算器的统计功能的用法.

2. 设 $X_1,X_2,\cdots,X_n$ 为总体 X 的一个样本，$x_1,x_2,\cdots,x_n$ 为一相应的样本值. 求下列各总体的概率密度或分布律中的未知参数的矩估计量和矩估计值.

(1) $f(x)=\begin{cases}\theta c^{\theta}x^{-(\theta+1)}, & x>c,\\ 0, & 其他,\end{cases}$

其中 $c>0$ 为已知，$\theta>1$，θ 为未知参数.

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

其中 $\theta>0$，θ 为未知参数.

(3) $P\{X=x\}=\binom{m}{x}p^x(1-p)^{m-x}, x=0,1,2,\cdots,m,$

其中 $0<p<1$，p 为未知参数.

解　(1) $\mu_1=\int_{-\infty}^{\infty}xf(x)\mathrm{d}x=\int_c^{\infty}x\theta c^\theta x^{-(\theta+1)}\mathrm{d}x$

$$=\theta c^\theta\int_c^{\infty}x^{-\theta}\mathrm{d}x=\frac{\theta c^\theta x^{-\theta+1}}{-\theta+1}\Bigg|_c^{\infty}=\frac{c\theta}{\theta-1},$$

由此得
$$\theta=\frac{\mu_1}{\mu_1-c}.$$

在上式中以 $\bar{X}$ 代替 μ_1，得到 θ 的矩估计量和矩估计值分别为

$$\hat{\theta}=\frac{\bar{X}}{\bar{X}-c},\quad \hat{\theta}=\frac{\bar{x}}{\bar{x}-c}.$$

(2) $\mu_1=\int_0^1 x\sqrt{\theta}x^{\sqrt{\theta}-1}\mathrm{d}x=\int_0^1\sqrt{\theta}x^{\sqrt{\theta}}\mathrm{d}x=\frac{\sqrt{\theta}}{\sqrt{\theta}+1}.$

由此得
$$\theta=\left(\frac{\mu_1}{1-\mu_1}\right)^2.$$

在上式中以 $\bar{X}$ 代替 μ_1，得 θ 的矩估计量和矩估计值分别为

$$\hat{\theta}=\left[\frac{\bar{X}}{1-\bar{X}}\right]^2,\quad \hat{\theta}=\left(\frac{\bar{x}}{1-\bar{x}}\right)^2.$$

(3) 因 $\mu_1=E(X)=mp$，得 $p=\frac{\mu_1}{m}$，以 $\bar{X}$ 代替 μ_1，得到 p 的矩估计量和矩估计值分别为

$$\hat{p}=\frac{\bar{X}}{m},\quad \hat{p}=\frac{\bar{x}}{m}.$$

3. 求上题中各未知参数的最大似然估计值和估计量.

解　(1) 设 $x_1,x_2,\cdots,x_n$ 是一个样本值. 似然函数为

$$L=L(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^n\theta c^\theta x_i^{-(\theta+1)}$$
$$=(\theta c^\theta)^n\prod_{i=1}^n x_i^{-(\theta+1)}=(\theta c^\theta)^n\left(\prod_{i=1}^n x_i\right)^{-(\theta+1)}.$$
$$\ln L=n(\ln\theta+\theta\ln c)-(\theta+1)\ln\prod_{i=1}^n x_i.$$

令
$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L = n\left(\frac{1}{\theta}+\ln c\right)-\sum_{i=1}^{n}\ln x_i = 0,$$
得 θ 的最大似然估计值为
$$\hat{\theta} = 1\Bigg/\left(\frac{1}{n}\sum_{i=1}^{n}\ln x_i - \ln c\right),$$
θ 的最大似然估计量为
$$\hat{\theta} = 1\Bigg/\left(\frac{1}{n}\sum_{i=1}^{n}\ln X_i - \ln c\right).$$

(2) 设 $x_1, x_2, \cdots, x_n$ 是一个样本值. 似然函数为
$$L = \prod_{i=1}^{n}\sqrt{\theta}x_i^{\sqrt{\theta}-1} = \theta^{n/2}\left(\prod_{i=1}^{n}x_i\right)^{\sqrt{\theta}-1},$$
$$\ln L = \frac{n}{2}\ln\theta + (\sqrt{\theta}-1)\sum_{i=1}^{n}\ln x_i.$$
令
$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L = \frac{n}{2\theta}+\frac{1}{2\sqrt{\theta}}\sum_{i=1}^{n}\ln x_i = 0,$$
得 θ 的最大似然估计值为
$$\hat{\theta} = \frac{n^2}{\left(\sum_{i=1}^{n}\ln x_i\right)^2},$$
θ 的最大似然估计量为
$$\hat{\theta} = \frac{n^2}{\left(\sum_{i=1}^{n}\ln X_i\right)^2}.$$

(3) 设 $x_1, x_2, \cdots, x_n$ 是一个样本值. 似然函数为
$$L = \prod_{i=1}^{n}P\{X_i = x_i\} = \prod_{i=1}^{n}\binom{m}{x_i}p^{x_i}(1-p)^{m-x_i}$$
$$= \left[\prod_{i=1}^{n}\binom{m}{x_i}\right]p^{\sum\limits_{i=1}^{n}x_i}(1-p)^{nm-\sum\limits_{i=1}^{n}x_i},$$
$$\ln L = \ln\prod_{i=1}^{n}\binom{m}{x_i}+\left(\sum_{i=1}^{n}x_i\right)\ln p+\left(nm-\sum_{i=1}^{n}x_i\right)\ln(1-p),$$
令
$$\frac{\mathrm{d}}{\mathrm{d}p}\ln L = \left(\sum_{i=1}^{n}x_i\right)\frac{1}{p}-\left(nm-\sum_{i=1}^{n}x_i\right)\frac{1}{1-p} = 0,$$
得 p 的最大似然估计值为

$$\hat{p}=\frac{\sum_{i=1}^{n}x_i}{nm}=\frac{\bar{x}}{m},\quad 其中\ \bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i,$$

p 的最大似然估计量为

$$\hat{p}=\frac{\bar{X}}{m}.$$

4. (1) 设总体 X 具有分布律

X	1	2	3
p_k	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

其中 $\theta(0<\theta<1)$为未知参数. 已知取得了样本值 $x_1=1,x_2=2,x_3=1$. 试求 θ 的矩估计值和最大似然估计值.

(2) 设 $X_1,X_2,\cdots,X_n$ 是来自参数为 λ 的泊松分布总体的一个样本，试求 λ 的最大似然估计量及矩估计量.

(3) 设随机变量 X 服从以 r,p 为参数的负二项分布，其分布律为

$$P\{X=x_k\}=\binom{x_k-1}{r-1}p^r(1-p)^{x_k-r},\quad x_k=r,r+1,\cdots,$$

其中 r 已知，p 未知. 设有样本值 $x_1,x_2,\cdots,x_n$，试求 p 的最大似然估计值.

解 (1) (i) $\mu_1=\theta^2+2\times 2\theta(1-\theta)+3\times(1-\theta)^2=3-2\theta.$

解得 $\theta=\frac{1}{2}(3-\mu_1)$，故得 θ 的矩估计值为

$$\hat{\theta}=\frac{1}{2}(3-\bar{x}).$$

令 $\bar{x}=\frac{1}{3}(x_1+x_2+x_3)=\frac{1}{3}(1+2+1)=\frac{4}{3}$，故 θ 的矩估计值为 $\hat{\theta}=\frac{5}{6}$.

(ii) 由给定的样本值，得似然函数为

$$\begin{aligned}L&=\prod_{i=1}^{3}P\{X_i=x_i\}=P\{X_1=1\}P\{X_2=2\}P\{X_3=1\}\\&=\theta^2\cdot 2\theta(1-\theta)\cdot\theta^2=2\theta^5(1-\theta),\end{aligned}$$

$$\ln L=\ln 2+5\ln\theta+\ln(1-\theta).$$

令

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L=\frac{5}{\theta}-\frac{1}{1-\theta}=0,$$

得 θ 的最大似然估计值为 $\hat{\theta}=\frac{5}{6}$.

(2) (i) 设 $x_1,x_2,\cdots,x_n$ 是相应于样本 $X_1,X_2,\cdots,X_n$ 的样本值，则似然函数为

$$L=\prod_{i=1}^{n}\frac{\lambda^{x_i}\mathrm{e}^{-\lambda}}{x_i!}=\mathrm{e}^{-n\lambda}\lambda^{\sum\limits_{i=1}^{n}x_i}\Big/\prod_{i=1}^{n}x_i!,$$

$$\ln L=-n\lambda+\left(\sum_{i=1}^{n}x_i\right)\ln\lambda-\ln\prod_{i=1}^{n}x_i!.$$

令
$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\ln L=-n+\left(\sum_{i=1}^{n}x_i\right)\Big/\lambda=0,$$

得 λ 的最大似然估计值为 $\hat{\lambda}=\frac{1}{n}\sum\limits_{i=1}^{n}x_i=\bar{x}$，最大似然估计量为

$$\hat{\lambda}=\bar{X}.$$

(ii) 因 $\mu_1=E(X)=\lambda$，故 λ 的矩估计量也是 $\hat{\lambda}=\bar{X}$.

(3) 似然函数为

$$L(p)=\prod_{k=1}^{n}P\{X=x_k\}=\prod_{k=1}^{n}\binom{x_k-1}{r-1}p^r(1-p)^{x_k-r}$$
$$=\left[\prod_{k=1}^{n}\binom{x_k-1}{r-1}\right]p^{nr}(1-p)^{\sum\limits_{i=1}^{n}x_i-nr},$$

$$\ln L=\ln C+nr\ln p+\left(\sum_{k=1}^{n}x_k-nr\right)\ln(1-p),\quad C\text{ 为常数}.$$

令
$$\frac{\mathrm{d}}{\mathrm{d}p}(\ln L)=\frac{nr}{p}-\frac{1}{1-p}\left(\sum_{k=1}^{n}x_k-nr\right)=0,$$

得 p 的最大似然估计值为

$$\hat{p}=\frac{r}{\bar{x}}.$$

5. 设某种电子器件的寿命(以 h 计)T 服从双参数的指数分布，其概率密度为

$$f(t)=\begin{cases}\dfrac{1}{\theta}\mathrm{e}^{-(t-c)/\theta}, & t\geqslant c,\\ 0, & \text{其他},\end{cases}$$

其中 $c,\theta\ (c,\theta>0)$ 为未知参数. 自一批这种器件中随机地取 n 件进行寿命试验. 设它们的失效时间依次为 $x_1\leqslant x_2\leqslant\cdots\leqslant x_n$.

(1) 求 c 与 θ 的最大似然估计值.

(2) 求 c 与 θ 的矩估计量.

解 (1) 似然函数为

$$L(\theta,c)=L(x_1,x_2,\cdots,x_n;\theta,c)$$

$$=\begin{cases}\prod_{i=1}^{n}\frac{1}{\theta}e^{-(x_i-c)/\theta}, & x_1,x_2,\cdots,x_n\geqslant c,\\ 0, & \text{其他}.\end{cases}$$

由题设 $x_1\leqslant x_2\leqslant\cdots\leqslant x_n$，故 $x_1,x_2,\cdots,x_n\geqslant c$ 相当于 $x_1\geqslant c$，因而上式相当于

$$L(\theta,c)=\begin{cases}\frac{1}{\theta^n}e^{-\left(\sum\limits_{i=1}^{n}x_i-nc\right)/\theta}, & c\leqslant x_1,\\ 0, & c>x_1.\end{cases}$$

可知当 $c\leqslant x_1$ 时，$L(\theta,c)$随 c 的增加而递增，而当 $c>x_1$ 时 $L(\theta,c)=0$，因而对于固定的 θ，$L(\theta,c)$在 $c=x_1$ 取到最大值，从而知应取 $\hat{c}=x_1$.

另外，当 $c\leqslant x_1$ 时，将 $L(\theta,c)$取自然对数得

$$\ln L(\theta,c)=-n\ln\theta-\frac{1}{\theta}\left(\sum_{i=1}^{n}x_i-nc\right).$$

令 $\frac{\partial}{\partial\theta}\ln L(\theta,c)=0$，得

$$\frac{\partial}{\partial\theta}\ln L(\theta,c)=-\frac{n}{\theta}+\frac{\sum\limits_{i=1}^{n}x_i-nc}{\theta^2}=0,$$

于是

$$\theta=\bar{x}-c.$$

由此可知 c,θ 的最大似然估计值为

$$\begin{cases}\hat{c}=x_1,\\ \hat{\theta}=\bar{x}-x_1.\end{cases}$$

(2) $\mu_1=\int_{-\infty}^{\infty}tf(t)dt=\int_{c}^{\infty}\frac{t}{\theta}e^{-(t-c)/\theta}dt$,

令 $u=\frac{t-c}{\theta}$，得

$$\mu_1=\int_0^{\infty}(\theta u+c)e^{-u}du=c+\theta\Gamma(2)=c+\theta,$$

$$\mu_2=\int_{-\infty}^{\infty}t^2f(t)dt=\int_c^{\infty}\frac{t^2}{\theta}e^{-(t-c)/\theta}dt,$$

令 $u=\frac{t-c}{\theta}$，得

$$\mu_2=\int_0^{\infty}(\theta u+c)^2e^{-u}du=\theta^2\Gamma(3)+2c\theta\Gamma(2)+c^2=2\theta^2+2c\theta+c^2$$
$$=(c+\theta)^2+\theta^2,$$

由此得

$$\theta=\sqrt{\mu_2-\mu_1^2},\quad c=\mu_1-\sqrt{\mu_2-\mu_1^2}.$$

将上两式中的 μ_1,μ_2 分别换成 $A_1=\bar{X},A_2=\frac{1}{n}\sum_{i=1}^{n}X_i^2$，并注意到 $A_2-A_1^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$，就得到 θ 及 c 的矩估计量为

$$\hat{\theta}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2},\quad \hat{c}=\bar{X}-\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2}.$$

6. 一地质学家为研究密歇根湖湖滩地区的岩石成分，随机地自该地区取 100 个样品，每个样品有 10 块石子，记录了每个样品中属石灰石的石子数. 假设这 100 次观察相互独立，并且由过去经验知，它们都服从参数为 $m=10,p$ 的二项分布，p 是这地区一块石子是石灰石的概率. 求 p 的最大似然估计值. 该地质学家所得的数据如下：

样品中属石灰石的石子数 i	0	1	2	3	4	5	6	7	8	9	10
观察到 i 块石灰石的样品个数	0	1	6	7	23	26	21	12	3	1	0

解 设 X 为一个样品中属石灰石的石子数，则 $X\sim b(10,p)$，由本章习题第 3 题知 p 的最大似然估计值为

$$\hat{p}=\frac{\bar{x}}{10}.$$

由给出的数据得

$$\begin{aligned}\bar{x}&=\frac{1}{100}(0\times0+1\times1+6\times2+7\times3+\cdots+3\times8+1\times9+0\times10)\\&=4.99,\end{aligned}$$

于是
$$\hat{p}=\frac{4.99}{10}=0.499.$$

7. (1) 设 $X_1,X_2,\cdots,X_n$ 是来自总体 X 的一个样本，且 $X\sim\pi(\lambda)$，求 $P\{X=0\}$ 的最大似然估计值.

(2) 某铁路局证实一个扳道员在五年内所引起的严重事故的次数服从泊松分布. 求一个扳道员在五年内未引起严重事故的概率 p 的最大似然估计值. 使用下面 122 个观察值. 下表中，r 表示一扳道员五年中引起严重事故的次数，s 表示观察到的扳道员人数.

r	0	1	2	3	4	5
s	44	42	21	9	4	2

解 (1) 设 $x_1,x_2,\cdots,x_n$ 是相应于样本 $X_1,X_2,\cdots,X_n$ 的样本值. 本题需求

$$P\{X=0\}=\frac{\lambda^0 e^{-\lambda}}{0!}=e^{-\lambda}$$

的最大似然估计.

由第 4 题知 λ 的最大似然估计值 $\hat{\lambda}=\bar{x}$，又由于函数 $u=e^{-\lambda}$ 具有单值反函数：$\lambda=-\ln u$，由最大似然估计的不变性知 $P\{X=0\}=e^{-\lambda}$ 的最大似然估计值为

$$\hat{P}\{X=0\}=e^{-\bar{x}}.$$

(2) 由所给数据，得

$$\begin{aligned}\bar{x}&=\frac{1}{122}\sum_{i=1}^{122}x_i\\&=\frac{1}{122}(44\times 0+42\times 1+21\times 2+9\times 3+4\times 4+2\times 5)=\frac{137}{122}.\end{aligned}$$

由(1)知，扳道员五年内未引起严重事故的概率 $p=P\{X=0\}=e^{-\lambda}$ 的最大似然估计值为

$$\hat{p}=\hat{P}\{X=0\}=e^{-\bar{x}}=e^{-137/122}=0.325\ 3.$$

8. (1) 设 $X_1,X_2,\cdots,X_n$ 是来自概率密度为

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

的总体的样本，θ 未知，求 $U=e^{-1/\theta}$ 的最大似然估计值.

(2) 设 $X_1,X_2,\cdots,X_n$ 是来自正态总体 $N(\mu,1)$ 的样本，μ 未知，求 $\theta=P\{X>2\}$ 的最大似然估计值.

(3) 设 $x_1,x_2,\cdots,x_n$ 是来自总体 $b(m,\theta)$ 的样本值，又 $\theta=\frac{1}{3}(1+\beta)$，求 β 的最大似然估计值.

解 (1) 先求 θ 的最大似然估计. 似然函数为

$$L(\theta)=\prod_{i=1}^{n}\theta x_i^{\theta-1}=\theta^n\left(\prod_{i=1}^{n}x_i\right)^{\theta-1},$$

$$\ln L(\theta)=n\ln\theta+(\theta-1)\ln\left(\prod_{i=1}^{n}x_i\right).$$

令

$$\frac{d\ln L(\theta)}{d\theta}=\frac{n}{\theta}+\sum_{i=1}^{n}\ln x_i=0,$$

得 θ 的最大似然估计值为

$$\hat{\theta}=\frac{-n}{\sum\limits_{i=1}^{n}\ln x_i}. \qquad (*_1)$$

$U=\mathrm{e}^{-1/\theta}$具有单调反函数，故由最大似然估计的不变性知U的最大似然估计值为

$$\hat{U}=\mathrm{e}^{-1/\hat{\theta}},\quad \text{其中 } \hat{\theta} \text{ 由}(*_1)\text{所确定.}$$

(2) 已知μ的最大似然估计为$\hat{\mu}=\bar{x}$. 而$\theta=P\{X>2\}=1-P\{X\leqslant 2\}=1-\Phi(2-\mu)$具有单调反函数. 由最大似然估计的不变性得$\theta=P\{X>2\}$的最大似然估计值为

$$\hat{\theta}=1-\Phi(2-\hat{\mu})=1-\Phi(2-\bar{x}).$$

(3) 由本章习题第 3 题知二项分布$X\sim b(m,\theta)$的参数θ的最大似然估计为$\hat{\theta}=\bar{x}/m$. 由最大似然估计的不变性得$\beta=3\theta-1$的最大似然估计值为

$$\hat{\beta}=3\hat{\theta}-1=\frac{3\bar{x}}{m}-1.$$

9. (1) 验证教材第六章 § 3 定理 5 中的统计量

$$S_W^2=\frac{n_1-1}{n_1+n_2-2}S_1^2+\frac{n_2-1}{n_1+n_2-2}S_2^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$

是两总体公共方差σ^2的无偏估计量(S_W^2称为σ^2的合并估计).

(2) 设总体X的数学期望为μ, $X_1,X_2,\cdots,X_n$是来自X的样本, $a_1,a_2,\cdots,a_n$是任意常数, 验证$\left(\sum\limits_{i=1}^{n}a_iX_i\right)\Big/\sum\limits_{i=1}^{n}a_i\ \left(\text{其中}\sum\limits_{i=1}^{n}a_i\neq 0\right)$是$\mu$的无偏估计量.

证 (1) 注意到$E(S_1^2)=E(S_2^2)=\sigma^2$, 即有

$$\begin{aligned}E(S_W^2)&=E\left[\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}\right]\\&=\frac{1}{n_1+n_2-2}[(n_1-1)E(S_1^2)+(n_2-1)E(S_2^2)]\\&=\sigma^2,\end{aligned}$$

故S_W^2是σ^2的无偏估计量.

(2) $$\begin{aligned}E\left[\left(\sum_{i=1}^{n}a_iX_i\right)\Big/\sum_{i=1}^{n}a_i\right]&=\sum_{i=1}^{n}E(a_iX_i)\Big/\sum_{i=1}^{n}a_i\\&=\sum_{i=1}^{n}a_iE(X_i)\Big/\sum_{i=1}^{n}a_i=\mu\sum_{i=1}^{n}a_i\Big/\sum_{i=1}^{n}a_i=\mu,\end{aligned}$$

这就证明了$\left(\sum\limits_{i=1}^{n}a_iX_i\right)\Big/\sum\limits_{i=1}^{n}a_i$是$\mu$的无偏估计量.

10. 设$X_1,X_2,\cdots,X_n$是来自总体X的一个样本, 设$E(X)=\mu, D(X)=\sigma^2$.

(1) 确定常数c, 使$c\sum\limits_{i=1}^{n-1}(X_{i+1}-X_i)^2$为$\sigma^2$的无偏估计.

(2) 确定常数 c，使 $\bar{X}^2-cS^2$ 是 μ^2 的无偏估计($\bar{X},S^2$ 是样本均值和样本方差).

解 (1)
$$E\left[c\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2\right]=c\sum_{i=1}^{n-1}E[(X_{i+1}-X_i)^2]$$
$$=c\sum_{i=1}^{n-1}\{D(X_{i+1}-X_i)+[E(X_{i+1}-X_i)]^2\}$$
$$=c\sum_{i=1}^{n-1}[D(X_{i+1})+D(X_i)]=2\sigma^2(n-1)c$$

(因 X_{i+1},X_i 相互独立且 $E(X_{i+1}-X_i)=0$). 要使
$$E\left[c\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2\right]=2\sigma^2(n-1)c=\sigma^2,$$
应取 $c=\dfrac{1}{2(n-1)}$.

(2) 要使
$$E\left(\bar{X}^2-cS^2\right)=E(\bar{X}^2)-cE(S^2)=\left(\frac{\sigma^2}{n}+\mu^2\right)-c\sigma^2=\mu^2,$$
应取 $c=\dfrac{1}{n}$.

11. 设总体 X 的概率密度为
$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta}x^{(1-\theta)/\theta}, & 0<x<1,\\ 0, & \text{其他}.\end{cases}\quad 0<\theta<\infty,$$
$X_1,X_2,\cdots,X_n$ 是来自总体 X 的样本.

(1) 验证 θ 的最大似然估计量是 $\hat{\theta}=-\dfrac{1}{n}\sum\limits_{i=1}^{n}\ln X_i$.

(2) 证明 $\hat{\theta}$ 是 θ 的无偏估计量.

证 (1) 似然函数为
$$L(\theta)=\frac{1}{\theta^n}\prod_{i=1}^{n}x_i^{(1-\theta)/\theta},$$
$$\ln L(\theta)=-n\ln\theta+\frac{1-\theta}{\theta}\ln\prod_{i=1}^{n}x_i.$$
令
$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta)=-\frac{n}{\theta}+\left(\sum_{i=1}^{n}\ln x_i\right)\left(\frac{-1}{\theta^2}\right)=0,$$
得
$$-n\theta=\sum_{i=1}^{n}\ln x_i.$$

于是得到 θ 的最大似然估计量为

$$\hat{\theta}=\frac{-1}{n}\sum_{i=1}^{n}\ln X_i.$$

（2）因
$$E(-\ln X)=\int_0^1(-\ln x)\cdot\frac{1}{\theta}x^{\frac{1}{\theta}-1}\mathrm{d}x$$
$$=-x^{\frac{1}{\theta}}\ln x\Big|_0^1+\int_0^1\frac{1}{x}x^{\frac{1}{\theta}}\mathrm{d}x=\theta.$$

从而知

$$E(\hat{\theta})=\frac{1}{n}\sum_{i=1}^{n}E(-\ln X_i)=\frac{1}{n}\cdot n\theta=\theta,$$

故 $\hat{\theta}$ 为 θ 的无偏估计量.

12. 设 X_1,X_2,X_3,X_4 是来自均值为 θ 的指数分布总体的样本，其中 θ 未知. 设有估计量

$$T_1=\frac{1}{6}(X_1+X_2)+\frac{1}{3}(X_3+X_4),$$

$$T_2=\frac{1}{5}(X_1+2X_2+3X_3+4X_4),$$

$$T_3=\frac{1}{4}(X_1+X_2+X_3+X_4).$$

（1）指出 T_1,T_2,T_3 中哪几个是 θ 的无偏估计量.

（2）在上述 θ 的无偏估计中指出哪一个较为有效.

解 （1）已知对于均值为 θ 的指数分布总体 X，有 $E(X)=\theta,D(X)=\theta^2$，于是 $E(X_i)=\theta,D(X_i)=\theta^2,i=1,2,3,4$. 所以

$$E(T_1)=\frac{1}{6}[E(X_1)+E(X_2)]+\frac{1}{3}[E(X_3)+E(X_4)]$$
$$=\frac{\theta}{3}+\frac{2\theta}{3}=\theta,$$

$$E(T_2)=\frac{1}{5}[E(X_1)+2E(X_2)+3E(X_3)+4E(X_4)]$$
$$=\frac{1}{5}(\theta+2\theta+3\theta+4\theta)=2\theta,$$

$$E(T_3)=\frac{1}{4}[E(X_1)+E(X_2)+E(X_3)+E(X_4)]$$
$$=\frac{1}{4}(\theta+\theta+\theta+\theta)=\theta.$$

以上结果表明 T_1,T_3 都是 θ 的无偏估计量，但 T_2 不是 θ 的无偏估计量.

(2) 又
$$D(T_1)=D\left[\frac{1}{6}(X_1+X_2)+\frac{1}{3}(X_3+X_4)\right]$$
$$=\frac{1}{36}D(X_1+X_2)+\frac{1}{9}D(X_3+X_4)$$
$$=\frac{1}{36}[D(X_1)+D(X_2)]+\frac{1}{9}[D(X_3)+D(X_4)]$$
$$=\frac{2\theta^2}{36}+\frac{2\theta^2}{9}=\frac{5}{18}\theta^2,$$

而
$$D(T_3)=D\left[\frac{1}{4}(X_1+X_2+X_3+X_4)\right]=\frac{1}{16}\sum_{i=1}^{4}D(X_i)$$
$$=\frac{4}{16}\theta^2=\frac{1}{4}\theta^2<D(T_1),$$

故统计量 T_3 较 T_1 有效.

13. (1) 设$\hat{\theta}$是参数 θ 的无偏估计，且有 $D(\hat{\theta})>0$，试证$\hat{\theta}^2=(\hat{\theta})^2$ 不是 θ^2 的无偏估计量.

(2) 试证明均匀分布
$$f(x)=\begin{cases}\dfrac{1}{\theta}, & 0<x\leqslant\theta,\\ 0, & \text{其他}\end{cases}$$
中未知参数 θ 的最大似然估计量不是无偏的.

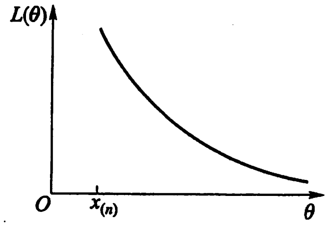
证 (1) 由 $D(\hat{\theta})>0$ 及 $E(\hat{\theta})=\theta$，得知$\hat{\theta}^2=(\hat{\theta})^2$ 的数学期望为
$$E(\hat{\theta}^2)=E[(\hat{\theta})^2]=D(\hat{\theta})+[E(\hat{\theta})]^2=D(\hat{\theta})+\theta^2>\theta^2,$$
故$\hat{\theta}^2=(\hat{\theta})^2$ 不是 θ^2 的无偏估计量.

(2) 似然函数为
$$L(\theta)=\begin{cases}\dfrac{1}{\theta^n}, & 0<x_1,x_2,\cdots,x_n\leqslant\theta,\\ 0, & \text{其他.}\end{cases}$$
记 $x_{(n)}=\max\{x_1,x_2,\cdots,x_n\}$，由于 $x_1,x_2,\cdots,x_n\leqslant\theta$ 相当于 $x_{(n)}\leqslant\theta$，因而上式相当于
$$L(\theta)=\begin{cases}\dfrac{1}{\theta^n}, & \theta\geqslant x_{(n)},\\ 0, & \theta<x_{(n)}.\end{cases}$$
可知当 $\theta<x_{(n)}$ 时，$L(\theta)=0$；而当 $\theta\geqslant x_{(n)}$ 时，$L(\theta)$ 随 θ 的增加而减少；故 $L(\theta)$ 在 $x_{(n)}$ 处取到最大值(题 7.13 图)，即得 θ 的最大似然估计值为

$$\hat{\theta}=x_{(n)}=\max\{x_1,x_2,\cdots,x_n\},$$

于是 θ 的最大似然估计量为

$$\hat{\theta}=\max\{X_1,X_2,\cdots,X_n\}.$$



题 7.13 图

本题总体 X 的分布函数为

$$F(x)=\begin{cases}0, & x<0,\\ \dfrac{x}{\theta}, & 0\leqslant x\leqslant\theta,\\ 1, & x>\theta.\end{cases}$$

得 $\hat{\theta}=\max\{X_1,X_2,\cdots,X_n\}$ 的分布函数为

$$F_{\hat{\theta}}(z)=F_{\max}(z)=[F(z)]^n=\begin{cases}0, & z<0,\\ \left(\dfrac{z}{\theta}\right)^n, & 0\leqslant z\leqslant\theta,\\ 1, & z>\theta.\end{cases}$$

由此推得 $\hat{\theta}$ 的概率密度为

$$f_{\hat{\theta}}(z)=f_{\max}(z)=\begin{cases}\dfrac{n}{\theta}\left(\dfrac{z}{\theta}\right)^{n-1}, & 0<z<\theta,\\ 0, & \text{其他},\end{cases}$$

于是

$$E(\hat{\theta})=\int_{-\infty}^{\infty}zf_{\hat{\theta}}(z)\mathrm{d}z=\int_0^{\theta}n\left(\frac{z}{\theta}\right)^n\mathrm{d}z$$

$$=\frac{n\theta}{n+1}\left(\frac{z}{\theta}\right)^{n+1}\Bigg|_0^{\theta}=\frac{n\theta}{n+1}\neq\theta,$$

故 $\hat{\theta}$ 不是 θ 的无偏估计量.

14. 设从均值为 μ，方差为 $\sigma^2>0$ 的总体中分别抽取容量为 n_1,n_2 的两独立样本. $\bar{X}_1$ 和 $\bar{X}_2$ 分别是两样本的均值. 试证：对于任意常数 a,b $(a+b=1)$，$Y=a\bar{X}_1+b\bar{X}_2$ 都是 μ 的无偏估计，并确定常数 a,b 使 $D(Y)$ 达到最小.

解 由 $E(\bar{X}_1)=\mu=E(\bar{X}_2)$ 以及 $a+b=1$，得知

$$E(Y)=E(a\bar{X}_1+b\bar{X}_2)=aE(\bar{X}_1)+bE(\bar{X}_2)$$

$$=a\mu+b\mu=(a+b)\mu=\mu,$$

即对于任意 a,b，只要 $a+b=1$，则 Y 都是 μ 的无偏估计量. 又

$$D(\bar{X}_1)=\frac{\sigma^2}{n_1},\quad D(\bar{X}_2)=\frac{\sigma^2}{n_2},$$

且 $\bar{X}_1,\bar{X}_2$ 相互独立，由此得

$$D(Y)=D(a\overline{X}_1+b\overline{X}_2)=D(a\overline{X}_1)+D(b\overline{X}_2)$$
$$=a^2D(\overline{X}_1)+b^2D(\overline{X}_2)=\left(\frac{a^2}{n_1}+\frac{b^2}{n_2}\right)\sigma^2.$$

将 $b=1-a$ 代入上式，得到

$$D(Y)=\left[\frac{a^2}{n_1}+\frac{(1-a)^2}{n_2}\right]\sigma^2.$$

令
$$\frac{\mathrm{d}}{\mathrm{d}a}D(Y)=\left[\frac{2a}{n_1}-\frac{2(1-a)}{n_2}\right]\sigma^2=0,$$

得
$$a=\frac{n_1}{n_1+n_2},$$

从而
$$b=1-a=\frac{n_2}{n_1+n_2},$$

又由于
$$\frac{\mathrm{d}^2}{\mathrm{d}a^2}D(Y)=\left(\frac{2}{n_1}+\frac{2}{n_2}\right)\sigma^2>0,$$

故知当 $a=\dfrac{n_1}{n_1+n_2}$，$b=\dfrac{n_2}{n_1+n_2}$ 时，$D(Y)$ 达到最小.

15. 设有 k 台仪器，已知用第 i 台仪器测量时，测定值总体的标准差为 $\sigma_i(i=1,2,\cdots,k)$. 用这些仪器独立地对某一物理量 θ 各观察一次，分别得到 $X_1,X_2,\cdots,X_k$. 设仪器都没有系统误差，即 $E(X_i)=\theta\ (i=1,2,\cdots,k)$. 问 $a_1,a_2,\cdots,a_k$ 取何值，方能使使用 $\hat{\theta}=\sum\limits_{i=1}^{k}a_iX_i$ 估计 θ 时，$\hat{\theta}$ 是无偏的，并且 $D(\hat{\theta})$ 最小？

解 要使 $\hat{\theta}$ 是 θ 的无偏估计，则必须

$$\theta=E(\hat{\theta})=E(a_1X_1+a_2X_2+\cdots+a_kX_k)=(a_1+a_2+\cdots+a_k)\theta.$$

即必须
$$a_1+a_2+\cdots+a_k=1. \qquad (*_1)$$

又由题设知 $D(X_i)=\sigma_i^2$，$i=1,2,\cdots,k$，且 $X_1,X_2,\cdots,X_k$ 相互独立，故有

$$D(\hat{\theta})=D\left(\sum_{i=1}^{k}a_iX_i\right)=\sum_{i=1}^{k}D(a_iX_i)=\sum_{i=1}^{k}a_i^2D(X_i)$$
$$=a_1^2\sigma_1^2+a_2^2\sigma_2^2+\cdots+a_k^2\sigma_k^2.$$

为求 $D(\hat{\theta})$ 在条件 $(*_1)$ 下的最小值，用拉格朗日乘数法，作函数

$$g(a_1,a_2,\cdots,a_k,\lambda)=a_1^2\sigma_1^2+a_2^2\sigma_2^2+\cdots+a_k^2\sigma_k^2+\lambda(a_1+a_2+\cdots+a_k-1).$$

令

$$\frac{\partial g}{\partial a_1}=2a_1\sigma_1^2+\lambda=0,\quad \frac{\partial g}{\partial a_2}=2a_2\sigma_2^2+\lambda=0,\cdots,$$

$$\frac{\partial g}{\partial a_k}=2a_k\sigma_k^2+\lambda=0,\quad \frac{\partial g}{\partial \lambda}=\sum_{i=1}^{k}a_i-1=0.$$

得

$$a_1=-\frac{\lambda}{2\sigma_1^2},\quad a_2=-\frac{\lambda}{2\sigma_2^2},\quad \cdots,\quad a_k=-\frac{\lambda}{2\sigma_k^2},\quad \sum_{i=1}^{k}a_i=1. \qquad (*_2)$$

将前 k 个式子代入末式，得

$$-\frac{\lambda}{2}\sum_{i=1}^{k}\frac{1}{\sigma_i^2}=1.$$

在上式中记 $\sum\limits_{i=1}^{k}\frac{1}{\sigma_i^2}=\frac{1}{\sigma_0^2}$，即有 $-\frac{\lambda}{2}=\sigma_0^2$，代入$(*_2)$式，得

$$a_1=\frac{\sigma_0^2}{\sigma_1^2},\quad a_2=\frac{\sigma_0^2}{\sigma_2^2},\quad \cdots,\quad a_k=\frac{\sigma_0^2}{\sigma_k^2},$$

即当 $a_i=\frac{\sigma_0^2}{\sigma_i^2}$，$i=1,2,\cdots,k$，用 $\hat{\theta}=\sum\limits_{i=1}^{k}a_iX_i$ 估计 θ 时，$\hat{\theta}$ 是无偏的且其方差最小(相对于其他无偏估计来说)．

16． 设某种清漆的 9 个样品，其干燥时间(以 h 计)分别为

6.0　5.7　5.8　6.5　7.0　6.3　5.6　6.1　5.0

设干燥时间总体服从正态分布 $N(\mu,\sigma^2)$．在下述情形下，求 μ 的置信水平为 0.95 的置信区间．

(1) 若由以往经验知 $\sigma=0.6$(h)．

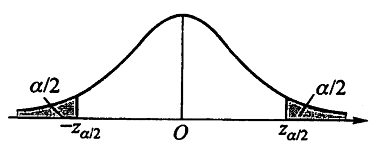
(2) 若 σ 为未知．

解　(1) 总体 $X\sim N(\mu,\sigma^2)$，μ 未知，σ^2 已知，$X_1,X_2,\cdots,X_n$ 是来自 X 的样本，需要求 μ 的置信水平为 $1-\alpha$ 的置信区间，即需确定随机区间$(\underline{\theta},\bar{\theta})$，使得

$$P\{\underline{\theta}<\mu<\bar{\theta}\}=1-\alpha.$$

考虑　$P\left\{?<\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}<?\right\}=1-\alpha$，

因 $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$，如题 7.16 图有



题 7.16 图

$$P\left\{-z_{\alpha/2}<\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}<z_{\alpha/2}\right\}=1-\alpha.$$

在{　}内的不等式中解出 μ，得

$$P\left\{\bar{X}-\frac{\sigma}{\sqrt{n}}z_{\alpha/2}<\mu<\bar{X}+\frac{\sigma}{\sqrt{n}}z_{\alpha/2}\right\}=1-\alpha,$$

即得 μ 的一个置信水平为 $1-\alpha$ 的置信区间为

$$\left(\bar{X}-\frac{\sigma}{\sqrt{n}}z_{\alpha/2},\quad \bar{X}+\frac{\sigma}{\sqrt{n}}z_{\alpha/2}\right).$$

今 $n=9,\sigma=0.6,1-\alpha=0.95,\alpha/2=0.025,z_{0.025}=1.96$，并算得 $\bar{x}=6$，得到 μ 的一个置信水平为 0.95 的置信区间为

$$\left(6\pm\frac{0.6}{3}z_{0.025}\right)=(6\pm0.2\times1.96)=(5.608,\ 6.392).$$

(2) σ 为未知，由于

$$\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1),$$

有
$$P\left\{-t_{\alpha/2}(n-1)<\frac{\bar{X}-\mu}{S/\sqrt{n}}<t_{\alpha/2}(n-1)\right\}=1-\alpha.$$

在{ }内的不等式中解出 μ，得

$$P\left\{\bar{X}-\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)<\mu<\bar{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right\}=1-\alpha,$$

即得 μ 的一个置信水平为 $1-\alpha$ 的置信区间为

$$\left(\bar{X}-\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1),\quad \bar{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right).$$

今 $n=9,1-\alpha=0.95,\alpha/2=0.025,t_{0.025}(8)=2.306$，并算得 $\bar{x}=6,s^2=0.33$，得到 μ 的一个置信水平为 0.95 的置信区间为

$$\left(6\pm\frac{\sqrt{0.33}}{3}\times2.306\right)=(6\pm0.442)=(5.558,\ 6.442).$$

17. 分别使用金球和铂球测定引力常数(以 $10^{-11}\mathrm{m}^3\cdot\mathrm{kg}^{-1}\cdot\mathrm{s}^{-2}$ 计).

(1) 用金球测定观察值为

6.683　6.681　6.676　6.678　6.679　6.672

(2) 用铂球测定观察值为

6.661　6.661　6.667　6.667　6.664

设测定值总体为 $N(\mu,\sigma^2)$，μ,σ^2 均为未知. 试就(1),(2)两种情况分别求 μ 的置信水平为 0.9 的置信区间，并求 σ^2 的置信水平为 0.9 的置信区间.

解 (1) $n=6,\bar{x}=6.678,s^2=0.003\ 87^2,1-\alpha=0.9,\alpha/2=0.05,t_{\alpha/2}(n-1)=t_{0.05}(5)=2.015\ 0$，得 μ 的一个置信水平为 0.9 的置信区间为

$$\left(\bar{x}\pm\frac{s}{\sqrt{n}}t_{\alpha/2}(n-1)\right)=\left(6.678\pm\frac{0.003\ 87}{\sqrt{6}}\times2.015\ 0\right)$$
$$=(6.678\pm0.003)=(6.675,\ 6.681).$$

由于

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1),$$

有
$$P\left\{\chi^2_{1-\alpha/2}(n-1)<\frac{(n-1)S^2}{\sigma^2}<\chi^2_{\alpha/2}(n-1)\right\}=1-\alpha.$$

在{　}内的不等式中解出 σ^2，得

$$P\left\{\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}<\sigma^2<\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right\}=1-\alpha,$$

即得 σ^2 的一个置信水平为 $1-\alpha$ 的置信区间为

$$\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)},\quad \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right).$$

查表知 $\chi^2_{\alpha/2}(n-1)=\chi^2_{0.05}(5)=11.070$，$\chi^2_{1-\alpha/2}(5)=\chi^2_{0.95}(5)=1.145$，得到 σ^2 的一个置信水平为 0.9 的置信区间为

$$\left(\frac{(6-1)\times 0.003\ 87^2}{11.070},\quad \frac{(6-1)\times 0.003\ 87^2}{1.145}\right)=(6.8\times 10^{-6},\ 6.5\times 10^{-5}).$$

(2) $n=5$，$\bar{x}=6.664$，$s^2=0.003^2$，$t_{0.05}(4)=2.131\ 8$，$\chi^2_{0.05}(4)=9.488$，$\chi^2_{0.95}(4)=0.711$，得到 μ 的一个置信水平为0.9的置信区间为

$$\left(6.664\pm\frac{0.003}{\sqrt{5}}\times 2.131\ 8\right)=(6.664\pm 0.003)=(6.661,\ 6.667),$$

σ^2 的一个置信水平为 0.9 的置信区间为

$$\left(\frac{(5-1)\times 0.003^2}{9.488},\ \frac{(5-1)\times 0.003^2}{0.711}\right)=(3.8\times 10^{-6},\ 5.06\times 10^{-5}).$$

18. 随机地取某种炮弹 9 发做试验，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 的置信水平为 0.95 的置信区间.

解　令 $n=9$，$s=11$，$1-\alpha=0.95$，$\alpha/2=0.025$，$\chi^2_{\alpha/2}(n-1)=\chi^2_{0.025}(8)=17.534$，$\chi^2_{1-\alpha/2}(n-1)=\chi^2_{0.975}(8)=2.180$，得到标准差 σ 的一个置信水平为 0.95 的置信区间为

$$\left(\frac{\sqrt{n-1}S}{\sqrt{\chi^2_{\alpha/2}(n-1)}},\ \frac{\sqrt{n-1}S}{\sqrt{\chi^2_{1-\alpha/2}(n-1)}}\right)=\left(\frac{\sqrt{8}\times 11}{\sqrt{17.534}},\ \frac{\sqrt{8}\times 11}{\sqrt{2.180}}\right)$$
$$=(7.4,\ 21.1).$$

19. 设 $X_1,X_2,\cdots,X_n$ 是来自分布 $N(\mu,\sigma^2)$ 的样本，μ 已知，σ 未知.

(1) 验证 $\sum\limits_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}\sim\chi^2(n)$. 利用这一结果构造 σ^2 的置信水平为$1-\alpha$

的置信区间.

（2）设 $\mu=6.5$，且有样本值 7.5，2.0，12.1，8.8，9.4，7.3，1.9，2.8，7.0，7.3. 试求 σ 的置信水平为 0.95 的置信区间.

解　(1) 因 $X_i \sim N(\mu,\sigma^2)$，故 $\dfrac{X_i-\mu}{\sigma} \sim N(0,1)$，$i=1,2,\cdots,n$.

由 $\dfrac{X_1-\mu}{\sigma},\dfrac{X_2-\mu}{\sigma},\cdots,\dfrac{X_n-\mu}{\sigma}$ 相互独立，得

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n).$$

于是有
$$P\left\{\chi^2_{1-\alpha/2}(n) < \sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2} < \chi^2_{\alpha/2}(n)\right\} = 1-\alpha,$$

即有
$$P\left\{\frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)} < \sigma^2 < \frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}\right\} = 1-\alpha.$$

得 σ^2 的置信水平为 $1-\alpha$ 的置信区间为

$$\left(\frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)},\quad \frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}\right).$$

（2）现在 $n=10$，$\mu=6.5$，$1-\alpha=0.95$，$\alpha=0.05$，由样本值经计算得 $\sum\limits_{i=1}^{10}(X_i-\mu)^2=102.69$，查表知，$\chi^2_{0.025}(10)=20.483$，$\chi^2_{0.975}(10)=3.247$，于是 σ^2 的置信水平为 0.95 的置信区间为(5.013，31.626). σ 的置信水平为 0.95 的置信区间为(2.239，5.624).

20. 在第 17 题中，设用金球和用铂球测定时测定值总体的方差相等. 求两个测定值总体均值差的置信水平为 0.90 的置信区间.

解　由于 $\sigma_1^2=\sigma_2^2=\sigma^2$ 未知，因而

$$\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_W\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \sim t(n_1+n_2-2),$$

其中 $S_W=\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$，即有

$$P\left\{-t_{\alpha/2}(n_1+n_2-2) \leqslant \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_W\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \leqslant t_{\alpha/2}(n_1+n_2-2)\right\}=1-\alpha.$$

在{ }内的不等式中解出 $\mu_1-\mu_2$，得

$$P\left\{\bar{X}_1-\bar{X}_2-t_{\alpha/2}(n_1+n_2-2)S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right.$$

$$\left.\leqslant\mu_1-\mu_2\leqslant\bar{X}_1-\bar{X}_2+t_{\alpha/2}(n_1+n_2-2)S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right\}$$

$$=1-\alpha,$$

即得 $\mu_1-\mu_2$ 的一个置信水平为 $1-\alpha$ 的置信区间为

$$\left(\bar{X}_1-\bar{X}_2\pm t_{\alpha/2}(n_1+n_2-2)S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right).$$

今 $n_1+n_2-2=9$，$t_{\alpha/2}(n_1+n_2-2)=t_{0.05}(9)=1.8331$，$\bar{x}_1-\bar{x}_2=0.014$，得到两个测定值总体均值差 $\mu_{金}-\mu_{铂}$ 的一个置信水平为 0.90 的置信区间为

$$\left(0.014\pm1.8331\sqrt{\frac{5\times0.00387^2+4\times0.003^2}{9}}\sqrt{\frac{1}{6}+\frac{1}{5}}\right)$$

$$=(0.014\pm0.004)=(0.010,\ 0.018).$$

21. 随机地从 A 批导线中抽 4 根，又从 B 批导线中抽 5 根，测得电阻（以 Ω 计）为

A 批导线：　0.143　0.142　0.143　0.137

B 批导线：　0.140　0.142　0.136　0.138　0.140

设测定数据分别来自分布 $N(\mu_1,\sigma^2)$，$N(\mu_2,\sigma^2)$，且两样本相互独立. 又 μ_1,μ_2,σ^2 均为未知. 试求 $\mu_1-\mu_2$ 的置信水平为 0.95 的置信区间.

解　将 A 批导线测定数据的均值、方差分别记为 $\bar{x}_1,s_1^2$；B 批的均值、方差分别记为 $\bar{x}_2,s_2^2$. 则有

$$n_1=4,\quad \bar{x}_1=0.14125,\quad 3s_1^2=0.00002475;$$

$$n_2=5,\quad \bar{x}_2=0.1392,\quad 4s_2^2=0.0000208,$$

$$s_w^2=\frac{3s_1^2+4s_2^2}{7}=0.00255^2,\quad 1-\alpha=0.95,\quad \alpha/2=0.025,$$

$$t_{\alpha/2}(n_1+n_2-2)=t_{0.025}(7)=2.3646.$$

得 $\mu_1-\mu_2$ 的一个置信水平为 0.95 的置信区间

$$\left(\bar{x}_1-\bar{x}_2\pm t_{\alpha/2}(n_1+n_2-2)s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$$

$$=\left(0.00205\pm2.3646\times0.00255\times\sqrt{\frac{1}{4}+\frac{1}{5}}\right)$$

$$=(0.002\pm0.004)=(-0.002,\ 0.006).$$

22. 研究两种固体燃料火箭推进器的燃烧率. 设两者都服从正态分布，并且

已知燃烧率的标准差均近似地为 0.05 cm/s，取样本容量为 $n_1=n_2=20$. 得燃烧率的样本均值分别为 $\bar{x}_1=18$ cm/s，$\bar{x}_2=24$ cm/s，设两样本独立. 求两燃烧率总体均值差$\mu_1-\mu_2$的置信水平为 0.99 的置信区间.

解　分别记两种固体燃料火箭推进器的燃烧率总体为 X_1 和 X_2，按题意$X_1\sim N(\mu_1,\sigma_1^2)$，$X_2\sim N(\mu_2,\sigma_2^2)$，其中 μ_1,μ_2 未知，σ_1^2,σ_2^2 已知，两样本独立，此时有

$$\bar{X}_1-\bar{X}_2\sim N\left(\mu_1-\mu_2,\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right),$$

$$\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\sim N(0,1),$$

因而有

$$P\left\{-z_{\alpha/2}<\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}<z_{\alpha/2}\right\}=1-\alpha.$$

在{　}内的不等式中解出 $\mu_1-\mu_2$，得

$$P\left\{(\bar{X}_1-\bar{X}_2)-z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2<(\bar{X}_1-\bar{X}_2)+z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right\}$$
$$=1-\alpha,$$

即得 $\mu_1-\mu_2$ 的一个置信水平为 $1-\alpha$ 的置信区间为

$$\left(\bar{X}_1-\bar{X}_2\pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right).$$

今 $\sigma_1=\sigma_2=0.05$，$n_1=n_2=20$，$\bar{x}_1=18$，$\bar{x}_2=24$，$1-\alpha=0.99$，$\alpha/2=0.005$，$z_{0.005}=2.57$，得到 $\mu_1-\mu_2$ 的一个置信水平为 0.99 的置信区间为

$$\left(18-24\pm2.57\sqrt{\frac{0.05^2}{20}+\frac{0.05^2}{20}}\right)=(-6\pm0.04)=(-6.04,\ -5.96).$$

23. 设两位化验员 A,B 独立地对某种聚合物含氯量用相同的方法各做 10 次测定，其测定值的样本方差依次为 $s_A^2=0.541\,9$，$s_B^2=0.606\,5$. 设 σ_A^2,σ_B^2 分别为 A,B 所测定的测定值总体的方差. 设总体均为正态的，且两样本独立. 求方差比 σ_A^2/σ_B^2 的置信水平为 0.95 的置信区间.

解　本题是两个总体均值均未知时，求两总体方差比的置信区间的问题. 已知 $n_1=n_2=10$，$s_A^2=0.541\,9$，$s_B^2=0.606\,5$，$1-\alpha=0.95$，$\frac{\alpha}{2}=0.025$，$F_{0.025}(9,9)=4.03$，$F_{0.975}(9,9)=0.248\,1$，得 σ_A^2/σ_B^2 的一个置信水平为 0.95 的置信区间为

$$\left(\frac{s_A^2}{s_B^2}\frac{1}{F_{0.025}(9,9)},\quad\frac{s_A^2}{s_B^2}\frac{1}{F_{0.975}(9,9)}\right)=(0.222,\ 3.601).$$

24. 在一批货物的容量为 100 的样本中，经检验发现有 16 只次品，试求这批货物次品率的置信水平为 0.95 的置信区间.

解 本题是(0−1)分布总体 X 的参数的区间估计问题，现在样本容量 $n=100$ 是一个大样本，X 的分布律为

$$f(x,p)=p^x(1-p)^{1-x},\quad x=0,1,$$

其中 p 为待估参数，已知 $E(X)=p,D(X)=p(1-p)$. 设 $X_1,X_2,\cdots,X_n$ 是一个样本，由中心极限定理知，近似地有

$$\frac{\sum_{i=1}^{n}X_i-np}{\sqrt{np(1-p)}}=\frac{n\bar{X}-np}{\sqrt{np(1-p)}}\overset{\text{近似地}}{\sim}N(0,1),$$

于是有

$$P\left\{-z_{\alpha/2}<\frac{n\bar{X}-np}{\sqrt{np(1-p)}}<z_{\alpha/2}\right\}\approx 1-\alpha.$$

在{ }的不等式中解出 p，得

$$P\{p_1<p<p_2\}\approx 1-\alpha,$$

其中

$$p_1=\frac{1}{2a}(-b-\sqrt{b^2-4ac}),\quad p_2=\frac{1}{2a}(-b+\sqrt{b^2-4ac}),$$

$$a=n+z_{\alpha/2}^2,\quad b=-\left(2n\bar{X}+z_{\alpha/2}^2\right),\quad c=n\bar{X}^2,$$

即得 p 的一个置信水平为 $1-\alpha$ 的近似置信区间为

$$(p_1,\ p_2).$$

今 $n=100,\bar{x}=0.16,1-\alpha=0.95,\alpha/2=0.025,z_{0.025}=1.96$，经计算得 $a=103.84,b=-35.84,c=2.56,p_1=0.101,p_2=0.244$，得这批货物的次品率的一个置信水平为 0.95 的近似置信区间为

$$(0.101,\ 0.244).$$

25. (1) 求第 16 题中 μ 的置信水平为 0.95 的单侧置信上限.

(2) 求第 21 题中 $\mu_1-\mu_2$ 的置信水平为 0.95 的单侧置信下限.

(3) 求第 23 题中方差比 σ_A^2/σ_B^2 的置信水平为 0.95 的单侧置信上限.

解 (1) 总体 $X\sim N(\mu,\sigma^2)$，μ 未知，先设 $\sigma=0.6$，$X_1,X_2,\cdots,X_n$ 是来自 X 的样本，需要求 μ 的置信水平为 $1-\alpha=0.95$ 的单侧置信上限. 即需确定统计量 $\bar{\mu}$ 使得

$$P\{\mu<\bar{\mu}\}=1-\alpha.$$

考虑

$$P\left\{?<\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right\}=1-\alpha.$$

(因在这里 μ 前有"－"号，故在{　}内写上不等式"? $<\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$".)

由于 $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$，知

$$P\left\{-z_\alpha<\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right\}=1-\alpha,$$

在{　}内的不等式中解出 μ，得

$$P\left\{\mu<\overline{X}+\frac{\sigma}{\sqrt{n}}z_\alpha\right\}=1-\alpha,$$

于是得 μ 的置信水平为 $1-\alpha$ 的单侧置信上限为

$$\overline{\mu}=\overline{X}+\frac{\sigma}{\sqrt{n}}z_\alpha.$$

在第 16 题中，$n=9,\bar{x}=6,\sigma=0.6,\alpha=0.05,z_\alpha=1.645$，得所求的单侧置信上限为

$$\overline{\mu}=6+\frac{0.6}{\sqrt{9}}\times1.645=6.329.$$

类似地，当 σ 未知时，可得 μ 的置信水平为 $1-\alpha$ 的单侧置信上限为

$$\overline{\mu}=\overline{X}+\frac{S}{\sqrt{n}}t_\alpha(n-1).$$

在第 16 题中，$n=9,\bar{x}=6,s=0.574\ 5,t_\alpha(n-1)=t_{0.05}(8)=1.859\ 5$. 得所求的单侧置信上限为

$$\overline{\mu}=6+\frac{0.574\ 5}{\sqrt{9}}\times1.859\ 5=6.356.$$

(2) 在第 21 题中，$\bar{x}_1-\bar{x}_2=0.002\ 05,n_1=4,n_2=5,1-\alpha=0.95,\alpha=0.05$，$\mu_1-\mu_2$ 的置信水平为 0.95 的单侧置信下限为

$$\begin{aligned}\underline{\mu_1-\mu_2}&=\bar{x}_1-\bar{x}_2-t_\alpha(n_1+n_2-2)s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\\&=0.002\ 05-1.894\ 6\times0.002\ 55\sqrt{\frac{1}{4}+\frac{1}{5}}\\&=-0.001\ 2.\end{aligned}$$

(3) 在第 23 题中，$s_A^2/s_B^2=0.541\ 9/0.606\ 5,n_1=n_2=10,1-\alpha=0.95,\alpha=0.05$. 本题需要求的是"?"满足

$$P\left\{\frac{\sigma_A^2}{\sigma_B^2}<?\right\}=0.95,$$

注意到
$$\frac{S_A^2/S_B^2}{\sigma_A^2/\sigma_B^2} \sim F(9,9),$$

即有
$$P\left\{\frac{S_A^2/S_B^2}{\sigma_A^2/\sigma_B^2} > F_{0.95}(9,9)\right\} = 0.95,$$

由{ }中的不等式解出$\dfrac{\sigma_A^2}{\sigma_B^2}<\dfrac{S_A^2/S_B^2}{F_{0.95}(9,9)}$，于是有
$$P\left\{\frac{\sigma_A^2}{\sigma_B^2}<\frac{S_A^2}{S_B^2}\frac{1}{F_{0.95}(9,9)}\right\}=0.95.$$

故得$\dfrac{\sigma_A^2}{\sigma_B^2}$的置信水平为 0.95 的单侧置信上限为
$$\overline{\left(\frac{\sigma_A^2}{\sigma_B^2}\right)} = \frac{s_A^2}{s_B^2}\frac{1}{F_{0.95}(9,9)} = \frac{0.541\ 9}{0.606\ 5}F_{0.05}(9,9)$$
$$= \frac{0.541\ 9}{0.606\ 5} \times 3.18 = 2.84.$$

26. 为研究某种汽车轮胎的磨损特性，随机地选择 16 只轮胎，每只轮胎行驶到磨坏为止，记录所行驶的路程(以 km 计)如下：

41 250	40 187	43 175	41 010	39 265	41 872	42 654	41 287
38 970	40 200	42 550	41 095	40 680	43 500	39 775	40 400

假设这些数据来自正态总体 $N(\mu,\sigma^2)$，其中 μ,σ^2 未知，试求 μ 的置信水平为 0.95 的单侧置信下限.

解 由$\dfrac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$，有
$$P\left\{\frac{\bar{X}-\mu}{S/\sqrt{n}}<t_\alpha(n-1)\right\}=1-\alpha,$$

即
$$P\left\{\mu>\bar{X}-\frac{S}{\sqrt{n}}t_\alpha(n-1)\right\}=1-\alpha.$$

得 μ 的置信水平为 $1-\alpha$ 的单侧置信下限为
$$\underline{\mu}=\bar{X}-\frac{S}{\sqrt{n}}t_\alpha(n-1).$$

今 $n=16,\bar{x}=41\ 116.875,s=1\ 346.842,1-\alpha=0.95,\alpha=0.05,t_\alpha(n-1)=t_{0.05}(15)=1.753\ 1$，故得 μ 的置信水平为 0.95 的单侧置信下限为
$$\underline{\mu}=41\ 116.875-\frac{1\ 346.842}{4}\times1.753\ 1=40\ 527.$$

27. 科学上的重大发现往往是由年轻人做出的. 下面列出了自 16 世纪初期

至 20 世纪早期的十二项重大发现的发现者和他们发现时的年龄：

发现内容	发现者	发现时间	年龄
1. 地球绕太阳运转	哥白尼(Copernicus)	1513	40
2. 望远镜、天文学的基本定律	伽利略(Galileo)	1600	36
3. 运动原理、重力、微积分	牛顿(Newton)	1665	22
4. 电的本质	富兰克林(Franklin)	1746	40
5. 燃烧是与氧气联系着的	拉瓦锡 (Lavoisier)	1774	31
6. 地球是在渐进过程中演化成的	莱尔(Lyell)	1830	33
7. 自然选择控制演化的证据	达尔文 (Darwin)	1858	49
8. 光的场方程	麦克斯韦(Maxwell)	1864	33
9. 放射性	居里夫人(Marie Curie)	1898	31
10. 量子论	普朗克(Planck)	1901	43
11. 狭义相对论，$E=mc^2$	爱因斯坦(Einstein)	1905	26
12. 量子论的数学基础	薛定谔(Schrödinger)	1926	39

设样本来自正态总体，试求发现者的平均年龄 μ 的置信水平为 0.95 的单侧置信上限.

解　μ 的置信水平为 $1-\alpha$ 的单侧置信上限为

$$\bar{\mu}=\bar{X}+\frac{S}{\sqrt{n}}t_{\alpha}(n-1).$$

现在 $n=12, \bar{x}=35.25, s=7.48, 1-\alpha=0.95, \alpha=0.05, t_{\alpha}(n-1)=t_{0.05}(11)=1.7959$，故得

$$\bar{\mu}=35.25+\frac{7.48}{\sqrt{12}}\times 1.7959=39.13,$$

约为 39 岁零 1 个月.

第八章 假设检验

1. 某批矿砂的 5 个样品中的镍含量(以%计)，经测定为

$$3.25 \quad 3.27 \quad 3.24 \quad 3.26 \quad 3.24$$

设测定值总体服从正态分布，但参数均未知，问在 $\alpha=0.01$ 下能否接受假设：这批矿砂的镍含量的均值为 3.25？

解 按题意总体 $X\sim N(\mu,\sigma^2)$，μ,σ^2 均未知，要求在显著性水平 $\alpha=0.01$ 下检验假设

$$H_0:\mu=3.25,\quad H_1:\mu\neq3.25.$$

因 σ^2 未知，故采用 t 检验，取检验统计量为 $t=\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}}$. 今 $n=5$，$\bar{x}=3.252$，$s=0.013$，$\alpha=0.01$，$t_{\alpha/2}(n-1)=t_{0.005}(4)=4.6041$，拒绝域为

$$|t|=\left|\frac{\bar{x}-\mu_0}{s/\sqrt{n}}\right|\geqslant t_{\alpha/2}(n-1)=4.6041.$$

因 $|t|$ 的观察值 $|t|=\left|\dfrac{3.252-3.25}{0.013/\sqrt{5}}\right|=0.344<4.6041$，不落在拒绝域之内，故在显著性水平 $\alpha=0.01$ 下接受原假设 H_0，即认为这批矿砂镍含量的均值为 3.25.

2. 如果一个矩形的宽度 w 与长度 l 的比 $\dfrac{w}{l}=\dfrac{1}{2}(\sqrt{5}-1)\approx0.618$，这样的矩形称为黄金矩形. 这种尺寸的矩形使人们看上去有良好的感觉. 现代的建筑构件(如窗架)、工艺品(如图片镜框)，甚至司机的执照、商业的信用卡等常常都是采用黄金矩形. 下面列出某工艺品工厂随机取的 20 个矩形的宽度与长度的比值：

0.693	0.749	0.654	0.670	0.662	0.672	0.615	0.606	0.690	0.628
0.668	0.611	0.606	0.609	0.601	0.553	0.570	0.844	0.576	0.933

设这一工厂生产的矩形的宽度与长度的比值总体服从正态分布，其均值为 μ，方差为 σ^2，μ,σ^2 均未知. 试检验假设(取 $\alpha=0.05$)

$$H_0:\mu=0.618,\quad H_1:\mu\neq0.618.$$

解 本题要求在显著性水平 $\alpha=0.05$ 下，检验正态总体均值的假设

$$H_0:\mu=0.618,\quad H_1:\mu\neq0.618.$$

因 σ 未知，故采用 t 检验. 因 $n=20$，$\bar{x}=0.6605$，$s=0.0925$，$\alpha=0.05$，$t_{\alpha/2}(n-1)=t_{0.025}(19)=2.093$，拒绝域为

$$|t|=\left|\frac{\bar{x}-\mu_0}{s/\sqrt{n}}\right|\geqslant t_{0.025}(19)=2.093.$$

今观察值

$$|t|=\left|\frac{0.6605-0.618}{0.0925/\sqrt{20}}\right|=2.055<2.093,$$

不落在拒绝域之内，故在显著性水平 $\alpha=0.05$ 下接受原假设 $H_0:\mu=0.618$.

3. 要求一种元件平均使用寿命不得低于 1 000 h，生产者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为 $\sigma=100$ h 的正态分布. 试在显著性水平 $\alpha=0.05$ 下判断这批元件是否合格？设总体均值为 μ，μ 未知. 即需检验假设 $H_0:\mu\geqslant 1\,000$，$H_1:\mu<1\,000$.

解 本题要求在显著性水平 $\alpha=0.05$ 下，检验正态总体均值的假设

$$H_0:\mu\geqslant 1\,000,\quad H_1:\mu<1\,000.$$

因 σ^2 已知，故采用 Z 检验，取检验统计量为 $Z=\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$，今 $n=25$，$\bar{x}=950$，$\sigma=100$，$\alpha=0.05$，$z_{0.05}=1.645$，拒绝域为

$$z=\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\leqslant -z_\alpha=-1.645.$$

因 Z 的观察值为 $z=\dfrac{950-1\,000}{100/\sqrt{25}}=-2.5<-1.645$，落在拒绝域内，故在显著性水平 $\alpha=0.05$ 下拒绝原假设 H_0，认为这批元件不合格.

4. 下面列出的是某工厂随机选取的 20 只部件的装配时间（以 min 计）：

9.8	10.4	10.6	9.6	9.7	9.9	10.9	11.1	9.6	10.2
10.3	9.6	9.9	11.2	10.6	9.8	10.5	10.1	10.5	9.7

设装配时间的总体服从正态分布 $N(\mu,\sigma^2)$，μ,σ^2 均未知. 是否可以认为装配时间的均值 μ 显著大于 10（取 $\alpha=0.05$）？

解 我们取 H_0 为维持原状，即取 H_0 为 $\mu\leqslant 10$，取 H_1 为 $\mu>10$（事实上我们关心的是 μ 有没有大于 10），即需在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\mu\leqslant 10,\quad H_1:\mu>10.$$

因为 σ 未知，故采用 t 检验，现在 $n=20$，$\bar{x}=10.2$，$s=0.5099$，$t_{0.05}(n-1)=t_{0.05}(19)=1.7291$，拒绝域为

$$t=\frac{\bar{x}-\mu_0}{s/\sqrt{n}}\geqslant t_\alpha(n-1)=1.7291.$$

因 t 的观察值

$$t=\frac{10.2-10}{0.5099/\sqrt{20}}=1.754>1.7291,$$

落在拒绝域内，故在显著性水平 $\alpha=0.05$ 下拒绝原假设 H_0，认为装配时间的均值显著大于 10.

5. 按规定，100 g 罐头番茄汁中的平均维生素 C 含量不得少于 21 mg/g. 现从工厂的产品中抽取 17 个罐头，其 100 g 番茄汁中，测得维生素 C 含量（以 mg/g 计）记录如下：

16 25 21 20 23 21 19 15 13 23 17 20 29 18 22 16 22

设维生素 C 含量服从正态分布 $N(\mu,\sigma^2)$，μ,σ^2 均未知，问这批罐头是否符合要求（取显著性水平 $\alpha=0.05$）？

解 本题需检验假设（$\alpha=0.05$）

$$H_0:\mu\geqslant 21,\quad H_1:\mu<21.$$

今 $n=17,\bar{x}=20,s=3.984,t_{0.05}(16)=1.7459$,

$$t=\frac{20-21}{3.984/\sqrt{17}}=-1.035>-1.7459.$$

故接受 H_0，认为这批罐头是符合规定的.

6. 下表分别给出两位文学家马克·吐温（Mark Twain）的 8 篇小品文以及斯诺特格拉斯（Snodgrass）的 10 篇小品文中由 3 个字母组成的单词的比例：

马克·吐温	0.225	0.262	0.217	0.240	0.230	0.229	0.235	0.217		
斯诺特格拉斯	0.209	0.205	0.196	0.210	0.202	0.207	0.224	0.223	0.220	0.201

设两组数据分别来自正态总体，且两总体方差相等，但参数均未知. 两样本相互独立. 问两位作家所写的小品文中包含由 3 个字母组成的单词的比例是否有显著的差异（取 $\alpha=0.05$）？

解 按题意总体 $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，两总体的方差相等，均等于 σ^2，σ^2 未知，两样本相互独立，本题需在显著性水平$\alpha=0.05$下检验假设

$$H_0:\mu_1=\mu_2,\quad H_1:\mu_1\neq\mu_2.$$

采用 t 检验，取检验统计量为 $t=\dfrac{\bar{X}-\bar{Y}-\delta}{S_W\sqrt{1/n_1+1/n_2}}$，今 $n_1=8$，$n_2=10$，$\bar{x}=0.2319$，$\bar{y}=0.2097$，$s_1^2=0.0146^2$，$s_2^2=0.0097^2$，$s_w^2=\dfrac{(8-1)s_1^2+(10-1)s_2^2}{8+10-2}=0.012^2$，$\delta=0$，$t_{0.025}(16)=2.1199$. 拒绝域为

$$|t|=\left|\frac{\bar{x}-\bar{y}}{s_w\sqrt{1/n_1+1/n_2}}\right|\geqslant t_{\alpha/2}(n_1+n_2-2)=2.1199.$$

因观察值

$$|t|=\left|\frac{0.2319-0.2097}{0.012\sqrt{1/8+1/10}}\right|=3.900>2.1199,$$

落在拒绝域之内，故拒绝 H_0，认为两位作家所写的小品文中包含由 3 个字母组成的单词的比例有显著的差异.

7. 在 20 世纪 70 年代后期人们发现，在酿造啤酒时，麦芽在干燥过程中形成致癌物质 N－亚硝基二甲胺(NDMA). 到了 20 世纪 80 年代初期开发了一种新的麦芽干燥过程. 下面分别给出新老两种过程中形成的 NDMA 含量(以 10 亿份中的份数计)：

老过程	6	4	5	5	6	5	5	6	4	6	7	4
新过程	2	1	2	2	1	0	3	2	1	0	1	3

设两样本分别来自正态总体，且两总体的方差相等，但参数均未知. 两样本独立. 分别以 μ_1,μ_2 记对应于老、新过程的总体的均值，试检验假设($\alpha=0.05$)

$$H_0:\mu_1-\mu_2\leqslant 2,\quad H_1:\mu_1-\mu_2>2.$$

解 本题要求在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\mu_1-\mu_2\leqslant 2,\quad H_1:\mu_1-\mu_2>2.$$

采用 t 检验. 今 $n_1=n_2=12,\bar{x}_1=5.25,\bar{x}_2=1.5,(n_1-1)s_1^2=10.25,(n_2-1)s_2^2=11,s_w^2=\dfrac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}=\dfrac{11s_1^2+11s_2^2}{22}=0.9828^2,t_\alpha(n_1+n_2-2)=t_{0.05}(22)=1.7171,\delta=2$，拒绝域为

$$t=\frac{\bar{x}_1-\bar{x}_2-2}{s_w\sqrt{1/n_1+1/n_2}}\geqslant t_\alpha(n_1+n_2-2)=1.7171.$$

因 t 的观察值 $t=\dfrac{5.25-1.5-2}{0.9828\sqrt{1/12+1/12}}=4.3616>1.7171$，落在拒绝域内，故在显著性水平 $\alpha=0.05$ 下拒绝 H_0，认为老新过程的均值差大于 2.

8. 随机地选了 8 个人，分别测量了他们在早晨起床时和晚上就寝时的身高(以 cm 计)，得到以下的数据：

序号	1	2	3	4	5	6	7	8
早上(x_i)	172	168	180	181	160	163	165	177
晚上(y_i)	172	167	177	179	159	161	166	175

设各对数据的差 $D_i=X_i-Y_i(i=1,2,\cdots,8)$ 是来自正态总体 $N(\mu_D,\sigma_D^2)$ 的样本，μ_D,σ_D^2 均未知.问是否可以认为早晨的身高比晚上的身高要高(取$\alpha=0.05$)?

解 题中的数据属成对数据，且可认为成对数据之差来自正态总体 $N(\mu_D,\sigma_D^2)$，本题要求在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\mu_D\leqslant 0,\quad H_1:\mu_D>0.$$

采用 t 检验，取检验统计量 $t=\dfrac{\overline{D}-0}{S_D/\sqrt{n}}$，今 $n=8$，各对数据之差 $d_i=x_i-y_i$ $(i=1,2,\cdots,8)$ 依次为 $0,1,3,2,1,2,-1,2$，由此得 $\overline{d}=\dfrac{1}{8}\sum\limits_{i=1}^{8}d_i=1.25$，$s_d=1.281\ 7$，$\alpha=0.05$，$t_\alpha(7)=t_{0.05}(7)=1.894\ 6$，拒绝域为

$$t=\frac{\overline{d}-0}{s_d/\sqrt{n}}\geqslant t_\alpha(n-1)=1.894\ 6.$$

因 t 的观察值 $t=\dfrac{1.25-0}{1.281\ 7/\sqrt{8}}=2.758>1.894\ 6$，落在拒绝域之内，故在显著性水平 $\alpha=0.05$ 下拒绝 H_0，认为早晨的身高比晚上的身高要高.

9. 为了比较用来做鞋子后跟的两种材料的质量，选取了 15 名男子(他们的生活条件各不相同)，每人穿一双新鞋，其中一只是以材料 A 做后跟，另一只以材料 B 做后跟，其厚度均为 10 mm. 过了一个月再测量厚度，得到数据(以 mm 计)如下：

男子	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
材料 $A(x_i)$	6.6	7.0	8.3	8.2	5.2	9.3	7.9	8.5	7.8	7.5	6.1	8.9	6.1	9.4	9.1
材料 $B(y_i)$	7.4	5.4	8.8	8.0	6.8	9.1	6.3	7.5	7.0	6.5	4.4	7.7	4.2	9.4	9.1

设 $D_i=X_i-Y_i$ $(i=1,2,\cdots,15)$ 是来自正态总体 $N(\mu_D,\sigma_D^2)$ 的样本，μ_D,σ_D^2 均未知.问是否可以认为以材料 A 制成的后跟比材料 B 的耐穿(取 $\alpha=0.05$)?

解 题中的数据属成对数据.本题要求在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\mu_D\leqslant 0,\quad H_1:\mu_D>0.$$

采用 t 检验，取检验统计量同第 8 题.今 $n=15$，$d_i=x_i-y_i(i=1,2,\cdots,15)$ 依次为 $-0.8,1.6,-0.5,0.2,-1.6,0.2,1.6,1.0,0.8,1.0,1.7,1.2,1.9,0,0$，$\overline{d}=0.553\ 3$，$s_d=1.022\ 5$，$\alpha=0.05$，$t_\alpha(n-1)=t_{0.05}(14)=1.761\ 3$，拒绝域为

$$t=\frac{\overline{d}-0}{s_d/\sqrt{n}}\geqslant t_\alpha(n-1)=1.761\ 3.$$

因 t 的观察值 $t=\dfrac{0.553\ 3-0}{1.022\ 5/\sqrt{15}}=2.095\ 8>1.761\ 3$，所以在 $\alpha=0.05$下拒绝 H_0，

认为材料 A 的制品较材料 B 的耐穿.

10. 为了试验两种不同的某谷物的种子的优劣，选取了 10 块土质不同的土地，并将每块土地分为面积相同的两部分，分别种植这两种种子. 设在每块土地的两部分人工管理等条件完全一样. 下面给出各块土地上的单位面积产量：

土地编号 i	1	2	3	4	5	6	7	8	9	10
种子 $A(x_i)$	23	35	29	42	39	29	37	34	35	28
种子 $B(y_i)$	26	39	35	40	38	24	36	27	41	27

设 $D_i=X_i-Y_i\ (i=1,2,\cdots,10)$ 是来自正态总体 $N(\mu_D,\sigma_D^2)$ 的样本，μ_D,σ_D^2 均未知. 问以这两种种子种植的谷物的产量是否有显著的差异(取 $\alpha=0.05$)？

解 题中的数据属成对数据，本题要求在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\mu_D=0,\quad H_1:\mu_D\neq 0.$$

采用 t 检验，取检验统计量同第 8 题. 今 $n=10$，$d_i=x_i-y_i\ (i=1,2,\cdots,10)$ 依次为 $-3,-4,-6,2,1,5,1,7,-6,1$，$\bar{d}=-0.2$，$s_d=4.44$，$\alpha=0.05$，$t_{\alpha/2}(n-1)=t_{0.025}(9)=2.2622$. 拒绝域为

$$|t|\geqslant t_{\alpha/2}(n-1)=2.2622.$$

因 $|t|$ 的观察值 $|t|=\left|\dfrac{-0.2-0}{4.44/\sqrt{10}}\right|=0.142<2.2622$，故在 $\alpha=0.05$ 下接受 H_0，认为这两种种子的谷物的产量无显著差异.

11. 一种混杂的小麦品种，株高的标准差为 $\sigma_0=14$ cm，经提纯后随机抽取 10 株，它们的株高(以 cm 计)为

90　105　101　95　100　100　101　105　93　97

考察提纯后的群体是否比原群体整齐？取显著性水平 $\alpha=0.01$，并设小麦株高服从 $N(\mu,\sigma^2)$.

解 需检验假设($\alpha=0.01$)

$$H_0:\sigma\geqslant\sigma_0,\quad H_1:\sigma<\sigma_0\quad(\sigma_0=14).$$

采用 χ^2 检验. 拒绝域为

$$\chi^2=\frac{(n-1)s^2}{\sigma_0^2}\leqslant\chi^2_{1-\alpha}(9).$$

现在 $n=10$，$\chi^2_{1-0.01}(9)=2.088$，$s^2=24.233$.

$$\frac{(n-1)s^2}{\sigma_0^2}=\frac{218.1}{14^2}=1.11<2.088.$$

故拒绝 H_0，认为提纯后的群体比原群体整齐.

12. 某种导线，要求其电阻的标准差不得超过 0.005 Ω，今在生产的一批导线中取样品 9 根，测得 $s=0.007\ \Omega$，设总体为正态分布，参数均未知. 问在显著性水平 $\alpha=0.05$ 下能否认为这批导线的标准差显著地偏大？

解 本题要求在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\sigma\leqslant 0.005,\quad H_1:\sigma>0.005\quad (\sigma_0=0.005).$$

采用χ^2检验，取检验统计量为$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$. 今 $n=9$，$s^2=0.007^2$，$\alpha=0.05$，$\chi_\alpha^2(n-1)=\chi_{0.05}^2(8)=15.507$，拒绝域为

$$\chi^2=\frac{(n-1)s^2}{\sigma_0^2}\geqslant\chi_\alpha^2(n-1)=15.507.$$

因χ^2的观察值$\chi^2=\dfrac{8\times 0.007^2}{0.005^2}=15.68>15.507$，落在拒绝域内，故在显著性水平 $\alpha=0.05$ 下拒绝 H_0，即认为这批导线的标准差显著地偏大.

13. 在第 2 题中记总体的标准差为 σ，试检验假设（取 $\alpha=0.05$）

$$H_0:\sigma^2=0.11^2,\quad H_1:\sigma^2\neq 0.11^2.$$

解 本题为第 2 题的方差检验问题，即要求在显著性水平 $\alpha=0.05$ 下，检验假设

$$H_0:\sigma^2=0.11^2,\quad H_1:\sigma^2\neq 0.11^2\quad (\sigma_0=0.11^2).$$

采用χ^2检验. 今 $n=20$，$s^2=0.092\,5^2$，$\alpha=0.05$，$\chi_{\alpha/2}^2(n-1)=\chi_{0.025}^2(19)=32.852$，$\chi_{1-\alpha/2}^2(n-1)=\chi_{0.975}^2(19)=8.906$，拒绝域为

$$\chi^2=\frac{(n-1)s^2}{\sigma_0^2}\geqslant\chi_{\alpha/2}^2(n-1)=32.852,$$

或

$$\chi^2=\frac{(n-1)s^2}{\sigma_0^2}\leqslant\chi_{1-\alpha/2}^2(n-1)=8.906.$$

因χ^2的观察值为$\chi^2=\dfrac{(20-1)\times 0.092\,5^2}{0.11^2}=13.44$，又因为

$$8.906<13.44<32.852,$$

即χ^2的观察值不落在拒绝域之内，故在显著性水平$\alpha=0.05$下接受 H_0，即认为 $\sigma^2=0.11^2$.

14. 测定某种溶液中的水分，它的 10 个测定值给出 $s=0.037\%$，设测定值总体为正态分布，σ^2 为总体方差，σ^2 未知. 试在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\sigma\geqslant 0.04\%,\quad H_1:\sigma<0.04\%.$$

解 本题要求在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\sigma\geqslant 0.04\%,\quad H_1:\sigma<0.04\%\quad(\sigma_0=0.04\%).$$

采用χ^2检验. 今 $n=10, s=0.037\%, \alpha=0.05, \chi^2_{1-\alpha}(n-1)=\chi^2_{0.95}(9)=3.325$，拒绝域为

$$\chi^2=\frac{(n-1)s^2}{\sigma_0^2}\leqslant\chi^2_{1-\alpha}(n-1)=3.325.$$

因χ^2的观察值为$\chi^2=\dfrac{(10-1)\times 0.037^2}{0.04^2}=7.701>3.325$，不落在拒绝域内，故在显著性水平 $\alpha=0.05$ 下接受 $H_0:\sigma\geqslant 0.04\%$.

15. 在第 6 题中分别记两个总体的方差为 σ_1^2 和 σ_2^2. 试检验假设(取 $\alpha=0.05$)

$$H_0:\sigma_1^2=\sigma_2^2,\quad H_1:\sigma_1^2\neq\sigma_2^2,$$

以说明在第 6 题中我们假设 $\sigma_1^2=\sigma_2^2$ 是合理的.

解 本题为第 6 题的两总体方差的检验问题，即要求在显著性水平 $\alpha=0.05$ 下，检验假设

$$H_0:\sigma_1^2=\sigma_2^2,\quad H_1:\sigma_1^2\neq\sigma_2^2.$$

采用 F 检验，取检验统计量为 $F=S_1^2/S_2^2$. 今 $n_1=8, n_2=10, s_1^2=0.014\ 6^2$, $s_2^2=0.009\ 7^2, \alpha=0.05, F_{\alpha/2}(n_1-1,n_2-1)=F_{0.025}(7,9)=4.20, F_{1-\alpha/2}(n_1-1, n_2-1)=1/F_{\alpha/2}(n_2-1,n_1-1)=1/F_{0.025}(9,7)=1/4.82=0.207$. 拒绝域为

$$F=s_1^2/s_2^2\geqslant F_{\alpha/2}(n_1-1,n_2-1)=4.20,$$

或

$$F=s_1^2/s_2^2\leqslant F_{1-\alpha/2}(n_1-1,n_2-1)=0.207.$$

今 F 的观察值为 $F=0.014\ 6^2/0.009\ 7^2=2.27$，因 $0.207<2.27<4.20$，不落在拒绝域中，故在显著性水平 $\alpha=0.05$ 下接受 $H_0:\sigma_1^2=\sigma_2^2$，认为两总体方差相等.

注：在采用 F 检验法检验双边假设 $H_0:\sigma_1^2=\sigma_2^2, H_1:\sigma_1^2\neq\sigma_2^2$ 时，接受域为

$$F_{1-\alpha/2}(n_1-1,n_2-1)<s_1^2/s_2^2<F_{\alpha/2}(n_1-1,n_2-1).$$

易犯的错误是，将接受域误写成 $s_1^2/s_2^2<F_{\alpha/2}(n_1-1,n_2-1)$，或误写成 $s_1^2/s_2^2>F_{1-\alpha/2}(n_1-1,n_2-1)$.

16. 在第 7 题中分别记两个总体的方差为 σ_1^2 和 σ_2^2. 试检验假设(取 $\alpha=0.05$)

$$H_0:\sigma_1^2=\sigma_2^2,\quad H_1:\sigma_1^2\neq\sigma_2^2,$$

以说明在第 7 题中我们假设 $\sigma_1^2=\sigma_2^2$ 是合理的.

解 本题为第 7 题的两总体方差的检验问题，即要求在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\sigma_1^2=\sigma_2^2,\quad H_1:\sigma_1^2\neq\sigma_2^2.$$

采用 F 检验. 今 $n_1=n_2=12, s_1^2=0.9318, s_2^2=1, \alpha=0.05, F_{\alpha/2}(n_1-1,n_2-1)=F_{0.025}(11,11)=3.48, F_{1-\alpha/2}(n_1-1,n_2-1)=1/F_{\alpha/2}(n_2-1,n_1-1)=1/F_{0.025}(11,11)=1/3.48=0.287$,拒绝域为

$$F=s_1^2/s_2^2 \geqslant F_{\alpha/2}(n_1-1,n_2-1)=3.48,$$

或

$$F=s_1^2/s_2^2 \leqslant F_{1-\alpha/2}(n_1-1,n_2-1)=0.287.$$

今 F 的观察值为 $F=0.9318$,因 $0.287<0.9318<3.48$,不落在拒绝域之内,故在显著性水平 $\alpha=0.05$ 下接受 $H_0:\sigma_1^2=\sigma_2^2$,认为两总体方差相等.

17. 两种小麦品种从播种到抽穗所需的天数如下:

x	101	100	99	99	98	100	98	99	99	99
y	100	98	100	99	98	99	98	98	99	100

设两样本依次来自正态总体 $N(\mu_1,\sigma_1^2), N(\mu_2,\sigma_2^2), \mu_i, \sigma_i (i=1,2)$ 均未知,两样本相互独立.

(1) 试检验假设 $H_0:\sigma_1^2=\sigma_2^2, H_1:\sigma_1^2\neq\sigma_2^2$ (取 $\alpha=0.05$).

(2) 若能接受 H_0,接着检验假设 $H_0':\mu_1=\mu_2, H_1':\mu_1\neq\mu_2$ (取 $\alpha=0.05$).

解 本题需检验

(1) $H_0:\sigma_1^2=\sigma_2^2,\quad H_1:\sigma_1^2\neq\sigma_2^2\quad(\alpha=0.05)$.

(2) $H_0':\mu_1=\mu_2,\quad H_1':\mu_1\neq\mu_2\quad(\alpha=0.05)$.

今 $n_1=10, n_2=10, \bar{x}_1=99.2, s_1^2=0.84, \bar{x}_2=98.9, s_2^2=0.77$.

(1) $s_1^2/s_2^2=1.09$,而 $F_{0.025}(9,9)=4.03, F_{0.975}(9,9)=\dfrac{1}{4.03}$,

$$\frac{1}{4.03}<1.09<4.03.$$

故接受 H_0,认为两者方差相等.

(2) $s_w^2=\dfrac{9\times0.84+9\times0.77}{18}=0.805$.

$$|t|=\frac{99.2-98.9}{\sqrt{0.805}\times\sqrt{1/10+1/10}}=0.748<t_{0.025}(18)=2.1009.$$

故接受 H_0',认为所需天数相同.

18. 用一种叫"混乱指标"的尺度去衡量工程师的英语文章的可理解性,对混乱指标的打分越低表示可理解性越高. 分别随机选取 13 篇刊载在工程杂志上的论文,以及 10 篇未出版的学术报告,对它们的打分列于下表:

工程杂志上的论文(数据Ⅰ)	未出版的学术报告(数据Ⅱ)
1.79 1.75 1.67 1.65	2.39 2.51 2.86
1.87 1.74 1.94	2.56 2.29 2.49
1.62 2.06 1.33	2.36 2.58
1.96 1.69 1.70	2.62 2.41

设数据Ⅰ,Ⅱ分别来自正态总体 $N(\mu_1,\sigma_1^2)$,$N(\mu_2,\sigma_2^2)$,$\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$ 均未知,两样本相互独立.

(1) 试检验假设 $H_0:\sigma_1^2=\sigma_2^2$,$H_1:\sigma_1^2\neq\sigma_2^2$ (取 $\alpha=0.1$).

(2) 若能接受 H_0,接着检验假设 $H_0':\mu_1=\mu_2$,$H_1':\mu_1\neq\mu_2$ (取 $\alpha=0.1$).

解 (1) $n_1=13,n_2=10,s_1^2=0.034,s_2^2=0.0264,\alpha=0.1,F_{0.05}(12,9)=3.07,F_{1-0.05}(12,9)=\dfrac{1}{F_{0.05}(9,12)}=\dfrac{1}{2.80}=0.357,s_1^2/s_2^2=1.288$. 由于

$$0.357<s_1^2/s_2^2<3.07,$$

故接受 H_0,认为两总体方差相等.

(2) 由(1)可认为 $\sigma_1^2=\sigma_2^2$,接着检验 $H_0':\mu_1=\mu_2$,$H_1':\mu_1\neq\mu_2$. 经计算 $\bar{x}_1=1.752,\bar{x}_2=2.507$.

$$s_w^2=\frac{12\times0.034+9\times0.0264}{13+10-2}=0.0307,$$

$$|t|=\left|\frac{1.752-2.507}{\sqrt{0.0307}\times\sqrt{1/13+1/10}}\right|=10.244.$$

而 $t_{0.05}(13+10-2)=t_{0.05}(21)=1.7207$,故拒绝 H_0',认为杂志上刊载的论文与未出版的学术报告的可理解性有显著差异.

注:在采用 t 检验法检验有关两个正态总体均值差的假设时,先要检查一下两总体的方差是否相等.若在题目中未指明两总体方差相等时,需先用 F 检验法来检验方差齐性,只有当经 F 检验认为两总体方差相等时,才能用 t 检验法来检验有关均值差的假设,如上题所示.

19. 有两台机器生产金属部件.分别在两台机器所生产的部件中各取一容量 $n_1=60$,$n_2=40$ 的样本,测得部件质量(以 kg 计)的样本方差分别为 $s_1^2=15.46$,$s_2^2=9.66$.设两样本相互独立.两总体分别服从 $N(\mu_1,\sigma_1^2)$,$N(\mu_2,\sigma_2^2)$ 分布,$\mu_i,\sigma_i^2(i=1,2)$均未知.试在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\sigma_1^2\leqslant\sigma_2^2,\quad H_1:\sigma_1^2>\sigma_2^2.$$

解 本题要求在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\sigma_1^2\leqslant\sigma_2^2,\quad H_1:\sigma_1^2>\sigma_2^2.$$

采用 F 检验. 今 $n_1=60, n_2=40, s_1^2=15.46, s_2^2=9.66, \alpha=0.05, F_\alpha(n_1-1,n_2-1)=F_{0.05}(59,39)=1.64$. 拒绝域为

$$F=s_1^2/s_2^2 \geqslant F_\alpha(n_1-1,n_2-1)=1.64.$$

因 F 的观察值 $F=15.46/9.66=1.60<1.64$，不落在拒绝域之内，故在显著性水平 $\alpha=0.05$ 下接受 H_0，认为 σ_1^2 不大于 σ_2^2.

20. 设需要对某一正态总体的均值进行假设检验

$$H_0:\mu\geqslant 15, \quad H_1:\mu<15.$$

已知 $\sigma^2=2.5$. 取 $\alpha=0.05$. 若要求当 H_1 中的 $\mu\leqslant 13$ 时犯第Ⅱ类错误的概率不超过 $\beta=0.05$，求所需的样本容量.

解 在显著性水平 $\alpha=0.05$ 下检验 $H_0:\mu\geqslant 15, H_1:\mu<15$，且已知 $\sigma^2=2.5$. 因要求当 $\mu\in H_1$ 且 $\mu\leqslant 13=15-\delta$ 时，犯第Ⅱ类错误的概率不超过 $\beta=0.05$，因此，$\delta=2$，所要求的样本容量 n 满足的关系式为

$$\sqrt{n} \geqslant (z_\alpha+z_\beta)\sigma/\delta,$$

其中 $z_\alpha=z_{0.05}=1.645, z_\beta=z_{0.05}=1.645, \sigma=\sqrt{2.5}$，故

$$n \geqslant 1.645^2 \times 2.5 = 6.765.$$

因此，取 $n\geqslant 7$ 就能满足题中的要求.

21. 电池在货架上滞留的时间不能太长. 下面给出某商店随机选取的 8 只电池的货架滞留时间(以天计)：

108　124　124　106　138　163　159　134

设数据来自正态总体 $N(\mu,\sigma^2)$，μ,σ^2 未知.

(1) 试检验假设 $H_0:\mu\leqslant 125, H_1:\mu>125$，取 $\alpha=0.05$.

(2) 若要求在上述 H_1 中 $(\mu-125)/\sigma\geqslant 1.4$ 时，犯第Ⅱ类错误的概率不超过 $\beta=0.1$，求所需的样本容量.

解 (1) 本题要求在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\mu\leqslant 125, \quad H_1:\mu>125.$$

采用 t 检验，检验统计量为 $t=\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}}$. 今 $n=8, \bar{x}=132, s^2=21.08^2, \alpha=0.05$, $t_\alpha(n-1)=t_{0.05}(7)=1.8946$. 拒绝域为

$$t\geqslant t_\alpha(n-1)=1.8946.$$

因 t 的观察值 $t=\dfrac{132-125}{21.08/\sqrt{8}}=0.939<1.8946$，不落在拒绝域之内，故在显著性水平 $\alpha=0.05$ 下接受 $H_0:\mu\leqslant 125$.

(2) 查教材附表 7，得 $n\geqslant 7$.

22. 一药厂生产一种新的止痛片，厂方希望验证服用新药片后至开始起作用的时间间隔较原有止痛片至少缩短一半，因此厂方提出需检验假设

$$H_0:\mu_1\leqslant 2\mu_2,\quad H_1:\mu_1>2\mu_2.$$

此处 μ_1,μ_2 分别是服用原有止痛片和服用新止痛片后至起作用的时间间隔的总体的均值. 设两总体均为正态且方差分别为已知值 σ_1^2,σ_2^2. 现分别在两总体中取一样本 $X_1,X_2,\cdots,X_{n_1}$ 和 $Y_1,Y_2,\cdots,Y_{n_2}$，设两个样本独立. 试给出上述假设 H_0 的拒绝域，取显著性水平为 α.

解　本题是在显著性水平 α 下，检验假设

$$H_0:\mu_1\leqslant 2\mu_2,\quad H_1:\mu_1>2\mu_2.$$

已知 $X_i\sim N(\mu_1,\sigma_1^2),i=1,2,\cdots,n_1;Y_i\sim N(\mu_2,\sigma_2^2),i=1,2,\cdots,n_2$，且两样本 $X_1,X_2,\cdots,X_{n_1}$ 和 $Y_1,Y_2,\cdots,Y_{n_2}$ 独立. 记

$$\bar{X}=\frac{1}{n_1}\sum_{i=1}^{n_1}X_i,\quad \bar{Y}=\frac{1}{n_2}\sum_{i=1}^{n_2}Y_i.$$

拒绝域的形式为

$$\bar{x}-2\bar{y}\geqslant k.$$

下面来确定 k.

$$\begin{aligned}P\{H_0\text{ 为真拒绝 }H_0\}&=P_{H_0}\left\{\bar{X}-2\bar{Y}\geqslant k\right\}\\&=P_{\mu_1-2\mu_2\leqslant 0}\left\{\frac{\bar{X}-2\bar{Y}}{\sqrt{\sigma_1^2/n_1+4\sigma_2^2/n_2}}\geqslant\frac{k}{\sqrt{\sigma_1^2/n_1+4\sigma_2^2/n_2}}\right\}\\&\leqslant P_{\mu_1-2\mu_2\leqslant 0}\left\{\frac{(\bar{X}-2\bar{Y})-(\mu_1-2\mu_2)}{\sqrt{\sigma_1^2/n_1+4\sigma_2^2/n_2}}\geqslant\frac{k}{\sqrt{\sigma_1^2/n_1+4\sigma_2^2/n_2}}\right\}\\&\quad(\text{因 }\mu_1-2\mu_2\leqslant 0),\end{aligned}$$

要控制 $P\{H_0\text{ 为真拒绝 }H_0\}\leqslant\alpha$，只需令上式右边 $=\alpha$. 由于

$$\frac{(\bar{X}-2\bar{Y})-(\mu_1-2\mu_2)}{\sqrt{\sigma_1^2/n_1+4\sigma_2^2/n_2}}\sim N(0,1),$$

即得 $\dfrac{k}{\sqrt{\sigma_1^2/n_1+4\sigma_2^2/n_2}}=z_\alpha$，因而 $k=z_\alpha\sqrt{\sigma_1^2/n_1+4\sigma_2^2/n_2}$，故得在给定的显著性水平 α 下，检验的拒绝域为

$$\bar{x}-2\bar{y}\geqslant z_\alpha\sqrt{\sigma_1^2/n_1+4\sigma_2^2/n_2},$$

即

$$\frac{\bar{x}-2\bar{y}}{\sqrt{\sigma_1^2/n_1+4\sigma_2^2/n_2}}\geqslant z_\alpha.$$

23. 检查了一本书的100页，记录各页中印刷错误的个数，其结果为

错误个数 f_i	0	1	2	3	4	5	6	≥7
含 f_i 个错误的页数	36	40	19	2	0	2	1	0

问能否认为一页的印刷错误的个数服从泊松分布(取 $\alpha=0.05$)？

解 记一页的印刷错误数为 X，按题意需在显著性水平 $\alpha=0.05$ 下检验假设

H_0：X 的分布律为

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

因参数 λ 未知，应先根据观察值，用最大似然估计法来求 λ 的估计(在 H_0 下)。可知 λ 的最大似然估计值为 $\hat{\lambda}=\bar{x}=1$。在 X 服从泊松分布的假设下，X 的所有可能取得的值为 $\Omega=\{0,1,2,\cdots\}$，将 Ω 分成如表8.1左起第一栏所示的两两不相交的子集 $A_0,A_1,\cdots,A_7$，接着根据估计式

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

来计算有关概率的估计，得到

$$\hat{p}_0=\hat{P}\{X=0\}=0.3679,$$

$$\hat{p}_1=\hat{P}\{X=1\}=0.3679,$$

$$\hat{p}_2=\hat{P}\{X=2\}=0.1839,$$

$$\hat{p}_3=\hat{P}\{X=3\}=0.0613,$$

$$\hat{p}_4=\hat{P}\{X=4\}=0.0153,$$

$$\hat{p}_5=\hat{P}\{X=5\}=0.0031,$$

$$\hat{p}_6=\hat{P}\{X=6\}=0.0005,$$

$$\hat{p}_7=\hat{P}\{X>6\}=1-\sum_{i=0}^{6}\hat{p}_i=0.0001.$$

计算结果列表如下(表8.1中 $n=100$)：

表 8.1

A_i	f_i	$\hat{p}_i$	$n\hat{p}_i$	$f_i^2/(n\hat{p}_i)$
$A_0:\{X=0\}$	36	0.367 9	36.79	35.227 0
$A_1:\{X=1\}$	40	0.367 9	36.79	43.490 1
$A_2:\{X=2\}$	19	0.183 9	18.39	19.630 2
$A_3:\{X=3\}$	2 } 5	0.061 3	6.13 } 8.03	3.113 3
$A_4:\{X=4\}$	0	0.015 3	1.53	
$A_5:\{X=5\}$	2	0.003 1	0.31	
$A_6:\{X=6\}$	1	0.000 5	0.05	
$A_7:\{X\geqslant 7\}$	0	0.000 1	0.01	

$\sum=101.460\ 6$

今$\chi^2=101.460\ 6-100=1.460\ 6$，因估计了一个参数 λ，故 $r=1$，又经并组后，只有 4 组，故 $k=4$，$\alpha=0.05$，$\chi_\alpha^2(k-r-1)=\chi_{0.05}^2(2)=5.992>1.460\ 6=\chi^2$，故在显著性水平 $\alpha=0.05$ 下接受假设 H_0，即认为样本来自泊松分布的总体.

注：本题中的原假设 H_0 所给出的分布律为

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

其中参数 $\lambda>0$ 未给出，这里要求检验的是印刷错误数 X 的分布的形式是否属于泊松分布类，而不是某一特定的泊松分布. 当接受 H_0 时，我们就认为 X 的分布的形式属于泊松分布类；当拒绝 H_0 时就认为 X 的分布的形式不属于泊松分布类.

24. 在一批灯泡中抽取 300 只做寿命试验，其结果如下：

寿命 t/h	$0\leqslant t\leqslant 100$	$100<t\leqslant 200$	$200<t\leqslant 300$	$t>300$
灯泡数	121	78	43	58

取 $\alpha=0.05$，试检验假设

H_0：灯泡寿命服从指数分布

$$f(t)=\begin{cases}0.005e^{-0.005t}, & t\geqslant 0,\\ 0, & t<0.\end{cases}$$

解 本题是在显著性水平 $\alpha=0.05$ 下，检验假设 H_0：灯泡寿命 X 服从指数分布，其概率密度为

$$f(t)=\begin{cases}0.005e^{-0.005t}, & t\geqslant 0,\\ 0, & t<0.\end{cases}$$

$n=300$. 在 H_0 为真的假设下，X 可能取值的范围为 $\Omega=[0,\infty)$. 将 Ω 分成互不相交的 4 个部分 A_1,A_2,A_3,A_4 如表 8.2. 以 A_i 记事件 $\{X\in A_i\}$，$i=1,2,3,4$. 若 H_0 为真，X 的分布函数为

$$F(t)=\begin{cases}1-e^{-0.005t}, & t\geqslant 0,\\ 0, & t<0.\end{cases}$$

得知

$$\begin{aligned}p_i&=P(A_i)=P\{a_i<X\leqslant a_{i+1}\}\\ &=F(a_{i+1})-F(a_i),\quad i=1,2,3,4.\end{aligned}$$

于是

$$p_1=P(A_1)=F(100)-F(0)=1-e^{-0.5}=0.393\ 5,$$
$$p_2=P(A_2)=F(200)-F(100)=e^{-0.5}-e^{-1}=0.238\ 7,$$
$$p_3=P(A_3)=F(300)-F(200)=e^{-1}-e^{-1.5}=0.144\ 7,$$
$$p_4=1-\sum_{i=1}^{3}p_i=0.223\ 1.$$

计算结果列表如下(表 8.2 中 $n=300$)：

表 8.2

A_i	f_i	p_i	np_i	$f_i^2/(np_i)$
$A_1:0\leqslant t\leqslant 100$	121	0.393 5	118.05	124.023 7
$A_2:100<t\leqslant 200$	78	0.238 7	71.61	84.960 2
$A_3:200<t\leqslant 300$	43	0.144 7	43.41	42.593 9
$A_4:t>300$	58	0.223 1	66.93	50.261 5

$\sum=301.839\ 3$

今 $\chi^2=301.839\ 3-300=1.839\ 3$. 由 $\alpha=0.05,k=4,r=0$ 知 $\chi_\alpha^2(k-r-1)=\chi_{0.05}^2(3)=7.815>1.839\ 3=\chi^2$. 故在显著性水平 $\alpha=0.05$ 下，接受假设 H_0，认为这批灯泡寿命服从指数分布，其概率密度为

$$f(t)=\begin{cases}0.005e^{-0.005t}, & t\geqslant 0,\\ 0, & \text{其他}.\end{cases}$$

25. 下面给出了随机选取的某大学一年级学生(200 名)一次数学考试的成绩.

(1) 画出数据的直方图.

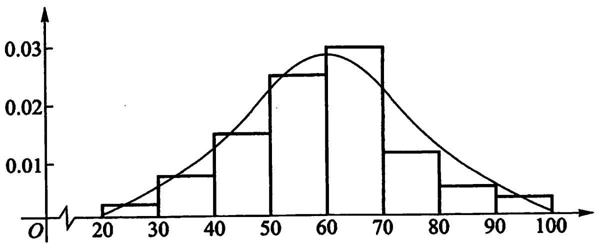
(2) 试取 $\alpha=0.1$ 检验数据来自正态总体 $N(60,15^2)$.

分数 x	$20\leqslant x\leqslant 30$	$30< x\leqslant 40$	$40< x\leqslant 50$	$50< x\leqslant 60$	$60< x\leqslant 70$	$70< x\leqslant 80$	$80< x\leqslant 90$	$90< x\leqslant 100$
学生数	5	15	30	51	60	23	10	6

解　(1) 将区间[20,100]等分为 8 个小区间，每个小区间的长度为 $\Delta=10$. 写出落在每个小区间内的数据频数 f_i，算出频率 f_i/n $(n=200, i=1,2,\cdots,8)$，列表如表 8.3. 在第 i 个区间上以 Δ 为底，以 $(f_i/n)/\Delta$ 为高作小长方形，得直方图如题 8.25 图.

表 8.3

组限	频数 f_i	频率 f_i/n	累积频率
$20\leqslant x\leqslant 30$	5	0.025	0.025
$30< x\leqslant 40$	15	0.075	0.100
$40< x\leqslant 50$	30	0.150	0.250
$50< x\leqslant 60$	51	0.255	0.505
$60< x\leqslant 70$	60	0.300	0.805
$70< x\leqslant 80$	23	0.115	0.920
$80< x\leqslant 90$	10	0.050	0.970
$90< x\leqslant 100$	6	0.030	1



题 8.25 图

(2) 本题要求在显著性水平 $\alpha=0.1$ 下检验假设

$$H_0: \quad \text{数据 } X \text{ 来自正态总体 } X\sim N(60,15^2).$$

即需检验 X 的概率密度为

$$f(x)=\frac{1}{15\sqrt{2\pi}}\mathrm{e}^{-\frac{(x-60)^2}{2\times 15^2}}, \quad -\infty<x<\infty.$$

将在 H_0 下 X 可能取值的区间 $(-\infty,\infty)$ 分为 8 个两两不相交的小区间 $A_1,A_2,\cdots,A_8$. 用 A_i 记事件"X 的观察值落在 A_i 内". 以 $f_i(i=1,2,\cdots,8)$ 记样

本观察值 $x_1,x_2,\cdots,x_{200}$ 中落在 A_i 的个数. 现取 A_i 如表 8.4 中第一列所示，其中 $A_1=(-\infty,30]$，$A_8=(90,\infty)$，其他与表 8.3 中所列的相同.

此处 $n=200$，若 H_0 为真，则有

$$p_1=P\{-\infty<X\leqslant 30\}=\Phi\left(\frac{30-60}{15}\right)$$
$$=\Phi(-2)=1-\Phi(2)=1-0.977\,2=0.022\,8,$$

$$p_2=P\{30<X\leqslant 40\}=\Phi\left(\frac{40-60}{15}\right)-\Phi\left(\frac{30-60}{15}\right)$$
$$=\Phi(-1.33)-\Phi(-2)=\Phi(2)-\Phi(1.33)$$
$$=0.977\,2-0.908\,2=0.069\,0,$$

$$p_3=P\{40<X\leqslant 50\}=\Phi\left(\frac{50-60}{15}\right)-\Phi\left(\frac{40-60}{15}\right)$$
$$=\Phi(-0.67)-\Phi(-1.33)=\Phi(1.33)-\Phi(0.67)$$
$$=0.908\,2-0.748\,6=0.159\,6,$$

$$p_4=P\{50<X\leqslant 60\}=\Phi(0)-\Phi\left(\frac{50-60}{15}\right)$$
$$=0.5-\Phi(-0.67)=\Phi(0.67)-0.5=0.748\,6-0.5=0.248\,6,$$

$$p_5=P\{60<X\leqslant 70\}=\Phi\left(\frac{70-60}{15}\right)-\Phi(0)$$
$$=\Phi(0.67)-0.5=0.748\,6-0.5=0.248\,6,$$

$$p_6=P\{70<X\leqslant 80\}=\Phi\left(\frac{80-60}{15}\right)-\Phi\left(\frac{70-60}{15}\right)$$
$$=\Phi(1.33)-\Phi(0.67)=0.908\,2-0.748\,6=0.159\,6,$$

$$p_7=P\{80<X\leqslant 90\}=\Phi\left(\frac{90-60}{15}\right)-\Phi\left(\frac{80-60}{15}\right)$$
$$=\Phi(2)-\Phi(1.33)=0.977\,2-0.908\,2=0.069\,0,$$

$$p_8=P\{90<X<\infty\}=1-\sum_{i=1}^{7}p_i=1-0.977\,2=0.022\,8.$$

计算结果列于表 8.4：

表 8.4

A_i	f_i	p_i	np_i	$f_i^2/(np_i)$
$A_1:(-\infty,30]$ $A_2:(30,40]$	$\left.\begin{matrix}5\\15\end{matrix}\right\}20$	0.022 8 0.069 0	$\left.\begin{matrix}4.56\\13.80\end{matrix}\right\}18.36$	21.786 5
$A_3:(40,50]$	30	0.159 6	31.92	28.195 5

续表

A_i	f_i	p_i	np_i	$f_i^2/(np_i)$
$A_4:(50,60]$	51	0.248 6	49.72	52.313 0
$A_5:(60,70]$	60	0.248 6	49.72	72.405 5
$A_6:(70,80]$	23	0.159 6	31.92	16.572 7
$A_7:(80,90]$	10 } 16	0.069 0	13.80 } 18.36	13.943 4
$A_8:(90,\infty)$	6	0.022 8	4.56	

$\sum=205.216\ 6$

因此$\chi^2=205.216\ 6-200=5.216\ 6$. 因 $\alpha=0.1,k=6,r=0$,有

$$\chi_\alpha^2(k-r-1)=\chi_{0.1}^2(5)=9.236>5.216\ 6,$$

故在显著性水平 $\alpha=0.1$ 下接受 H_0,即认为考试成绩的数据来自正态总体 $N(60,15^2)$.

26. 袋中装有 8 只球,其中红球数未知. 在其中任取 3 只,记录红球的只数 X,然后放回,再任取 3 只,记录红球的只数,然后放回. 如此重复进行了 112 次,其结果如下:

X	0	1	2	3
次数	1	31	55	25

试取 $\alpha=0.05$ 检验假设

H_0:X 服从超几何分布

$$P\{X=k\}=\binom{5}{k}\binom{3}{3-k}\Big/\binom{8}{3},\quad k=0,1,2,3.$$

即检验假设 H_0:红球的只数为 5.

解 本题要求在显著性水平 $\alpha=0.05$ 下检验假设

H_0:X 服从超几何分布

$$P\{X=k\}=\binom{5}{k}\binom{3}{3-k}\Big/\binom{8}{3},\quad k=0,1,2,3.$$

此处 $n=112$. 若 H_0 为真,则可按所给分布律计算,得如下的概率

$$p_1=P\{X=0\}=\binom{5}{0}\binom{3}{3}\Big/\binom{8}{3}=\frac{1}{56},$$

$$p_2=P\{X=1\}=\binom{5}{1}\binom{3}{2}\Big/\binom{8}{3}=\frac{15}{56},$$

$$p_3=P\{X=2\}=\binom{5}{2}\binom{3}{1}\Big/\binom{8}{3}=\frac{30}{56},$$

$$p_4=P\{X=3\}=\binom{5}{3}\binom{3}{0}\Big/\binom{8}{3}=\frac{10}{56}.$$

计算结果列于表 8.5：

表 8.5

A_i	f_i	p_i	np_i	$f_i^2/(np_i)$
$A_1:\{X=0\}$	1 } 32	$\frac{1}{56}$	2 } 32	32
$A_2:\{X=1\}$	31	$\frac{15}{56}$	30	
$A_3:\{X=2\}$	55	$\frac{30}{56}$	60	50.416 7
$A_4:\{X=3\}$	25	$\frac{10}{56}$	20	31.25

$\sum=113.666\ 7$

因此$\chi^2=113.666\ 7-112=1.666\ 7$. 因 $\alpha=0.05$, $k=3$, $r=0$, 有$\chi_\alpha^2(k-r-1)=\chi_{0.05}^2(2)=5.992>1.666\ 7$，故在显著性水平$\alpha=0.05$下接受 H_0，即认为 X 服从超几何分布

$$P\{X=k\}=\binom{5}{k}\binom{3}{3-k}\Big/\binom{8}{3},\quad k=0,1,2,3.$$

27. 一农场 10 年前在一鱼塘中按比例 20 : 15 : 40 : 25 投放了四种鱼：鲑鱼、鲈鱼、竹夹鱼和鲇鱼的鱼苗，现在在鱼塘里获得一样本如下：

序号	1	2	3	4	
种类	鲑鱼	鲈鱼	竹夹鱼	鲇鱼	
数量(条)	132	100	200	168	$\sum=600$

试取 $\alpha=0.05$，检验各类鱼数量的比例较 10 年前是否有显著的改变.

解 以 X 记鱼种类的序号，按题意需检验假设

H_0: X 的分布律为

X	1	2	3	4
p_k	0.20	0.15	0.40	0.25

现在获得一样本如下：

X	1	2	3	4	
种类	鲑鱼	鲈鱼	竹夹鱼	鲇鱼	
数量(条)	132	100	200	168	$\sum=600$
A_i	A_1	A_2	A_3	A_4	

将在 H_0 下，X 可能取值的全体分成 4 个两两不相交的子集 A_1,A_2,A_3,A_4（以一鱼种作为一个子集）. 所需计算列表如下（表 8.6 中 $n=600$）：

表 8.6

A_i	f_i	p_i	np_i	$f_i^2/(np_i)$
A_1	132	0.20	120	145.20
A_2	100	0.15	90	111.11
A_3	200	0.40	240	166.67
A_4	168	0.25	150	188.16

$\sum=611.14$

现在 $\chi^2=611.14-600=11.14$，$k=4$，$r=0$，$\chi_{0.05}^2(k-r-1)=\chi_{0.05}^2(3)=7.815<11.14$. 故拒绝 H_0，认为各鱼类数量之比相对于 10 年前有显著的改变.

28. 某种鸟在起飞前，双足齐跳的次数 X 服从几何分布，其分布律为

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

今获得一样本如下：

x	1	2	3	4	5	6	7	8	9	10	11	12	$\geqslant 13$
观察到 x 的次数	48	31	20	9	6	5	4	2	1	1	2	1	0

(1) 求 p 的最大似然估计值.

(2) 取 $\alpha=0.05$，检验假设 H_0：数据来自总体

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

解 (1) 设 $x_1,x_2,\cdots,x_n$ 是一个样本值，似然函数为

$$L=L(x_1,x_2,\cdots,x_n;p)=\prod_{i=1}^{n}[p^{x_i-1}(1-p)]=(1-p)^n p^{\sum_{i=1}^{n}x_i-n},$$

$$\ln L=n\ln(1-p)+\left(\sum_{i=1}^{n}x_i-n\right)\ln p,$$

令
$$\frac{\mathrm{d}}{\mathrm{d}p}\ln L=\frac{-n}{1-p}+\frac{\sum_{i=1}^{n}x_i-n}{p}=0,$$
得 p 的最大似然估计值为 $\hat{p}=\left(\sum_{i=1}^{n}x_i-n\right)\Big/\sum_{i=1}^{n}x_i=1-\frac{1}{\bar{x}}$.

本题中 $n=48+31+20+9+6+5+4+2+1+1+2+1=130$，

$$\sum_{i=1}^{n}x_i=1\times48+2\times31+3\times20+4\times9+5\times6+6\times5+7\times4$$
$$+8\times2+9\times1+10\times1+11\times2+12\times1=363.$$
$$\bar{x}=\frac{363}{130}.$$

得 p 的最大似然估计值为

$$\hat{p}=1-\frac{1}{\bar{x}}=1-\frac{130}{363}=0.641\ 9.$$

(2) 检验假设($\alpha=0.05$)

H_0:数据来自总体

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

在 X 服从几何分布的假设下，X 的所有可能取的值为 $\Omega=\{1,2,\cdots\}$，将 Ω 分成两两不相交的 7 个部分 $A_1=\{1\},A_2=\{2\},\cdots,A_6=\{6\},A_7=\{7,8,\cdots\}$，利用估计式

$$\hat{p}_i=\hat{P}\{X=i\}=\hat{p}^{i-1}(1-\hat{p}),\quad i=1,2,\cdots,6$$

来计算有关概率的估计，得到

$\hat{p}_1=0.358\ 1$，$\hat{p}_2=0.229\ 9$，$\hat{p}_3=0.147\ 5$，$\hat{p}_4=0.094\ 7$，

$\hat{p}_5=0.060\ 8$，$\hat{p}_6=0.039\ 0$，$\hat{p}_7\doteq1-\sum_{i=1}^{6}\hat{p}_i=0.070\ 0$.

所需计算列表如下(表 8.7 中 $n=130$)：

表 8.7

A_i	f_i	$\hat{p}_i$	$n\hat{p}_i$	$f_i^2/(n\hat{p}_i)$
A_1	48	0.358 1	46.553	49.492
A_2	31	0.229 9	29.887	32.154
A_3	20	0.147 5	19.175	20.860
A_4	9	0.094 7	12.311	6.579

续表

A_i	f_i	$\hat{p}_i$	$n\hat{p}_i$	$f_i^2/(n\hat{p}_i)$
A_5	6	0.060 8	7.904	4.555
A_6	5	0.039 0	5.070	4.931
A_7	11	0.070 0	9.100	13.297

$\sum=131.868$

于是$\chi^2=131.868-130=1.868$，因为$\chi^2_{0.05}(k-r-1)=\chi^2_{0.05}(7-1-1)=\chi^2_{0.05}(5)$，查表$\chi^2_{0.05}(5)=11.070>1.868$，从而接受 H_0.

29. (1) 设总体服从 $N(\mu,100)$，μ 未知，现有样本 $n=16$，$\bar{x}=13.5$，试检验假设 $H_0:\mu\leqslant 10$，$H_1:\mu>10$，(i) 取 $\alpha=0.05$，(ii) 取 $\alpha=0.10$，(iii) 求 H_0 可被拒绝的最小显著性水平.

(2) 考察生长在老鼠身上的肿块的大小. 以 X 表示在老鼠身上生长了 15 天的肿块的直径(以 mm 计)，设 $X\sim N(\mu,\sigma^2)$，μ,σ^2 均未知. 今随机地取 9 只老鼠(在它们身上的肿块都长了 15 天)，测得 $\bar{x}=4.3$，$s=1.2$，试取 $\alpha=0.05$，用 p 值法检验假设 $H_0:\mu=4.0$，$H_1:\mu\neq 4.0$，求出 p 值.

(3) 用 p 值法检验 §2 例 4 的检验问题.

(4) 用 p 值法检验第 27 题中的检验问题.

解 (1) 用 Z 检验法. 检验问题的拒绝域为

$$Z=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\geqslant z_\alpha.$$

现在 $n=16$，$\mu_0=10$，$\bar{x}=13.5$，$\sigma=10$，得拒绝域为

$$z=\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}=1.4\geqslant z_\alpha.$$

(i) 当 $\alpha=0.05$ 时，$z_\alpha=1.645>1.4$，故接受 H_0.

(ii) 当 $\alpha=0.10$ 时，$z_\alpha=1.282<1.4$，故拒绝 H_0.

(iii) p 值$=P\{Z\geqslant 1.4\}=1-\Phi(1.4)=1-0.919\ 2=0.080\ 8$.
因而拒绝 H_0 的最小显著性水平为 0.080 8.

(2) 用 t 检验法，检验统计量为

$$t=\frac{\overline{X}-\mu_0}{S/\sqrt{n}}.$$

当 H_0 为真时，$t\sim t(n-1)$，现在 $n=9$，$\mu_0=4.0$，$\bar{x}=4.3$，$s=1.2$，得检验统计量的观察值为

$$t_0=\frac{4.3-4.0}{1.2/\sqrt{9}}=0.75.$$

此为双边检验，由计算机得

$$p\text{ 值}=2\times P\{t>0.75\}=0.474\ 7>\alpha=0.05,$$

故接受 H_0.

(3) 用 t 检验法，检验统计量为

$$t=\frac{\overline{D}-0}{S_D/\sqrt{n}}.$$

当 H_0 为真时，$t\sim t(n-1)$，现在 $n=8,\overline{d}=-0.062\ 5,s_d=0.076\ 5$，得统计量的观察值为

$$t_0=-2.311.$$

此为单边检验，由计算机得

$$p\text{ 值}=P\{t<-2.311\}=P\{t>2.311\}=0.027\ 1<\alpha=0.05.$$

故拒绝 H_0.

(4) 用χ^2检验法，检验统计量为

$$\chi^2=\sum_{i=1}^{k}\frac{f_i^2}{np_i}-n.$$

当 H_0 为真时，$\chi^2\sim\chi^2(3)$. 以数据代入得检验统计量的观察值为

$$\chi_0^2=11.14.$$

此为右边检验，由计算机得

$$p\text{ 值}=P\{\chi^2\geqslant 11.14\}=0.011\ 0<\alpha=0.05.$$

故拒绝 H_0.

第九章 方差分析及回归分析

以下约定各个习题均符合涉及的方差分析模型或回归分析模型所要求的条件.

1. 今有某种型号的电池三批，它们分别是 A,B,C 三个工厂所生产的. 为评比其质量，各随机抽取 5 只电池为样品，经试验得其寿命(以 h 计)如下：

A	B	C
40　42	26　28	39　50
48　45	34　32	40　50
38	30	43

试在显著性水平 0.05 下检验电池的平均寿命有无显著的差异. 若差异是显著的，试求均值差 $\mu_A-\mu_B$，$\mu_A-\mu_C$ 和 $\mu_B-\mu_C$ 的置信水平为 95% 的置信区间.

解 以 μ_A,μ_B,μ_C 依次表示由工厂 A,B,C 生产的电池的平均寿命.

本题要求在 $\alpha=0.05$ 下检验假设

$$H_0:\ \mu_A=\mu_B=\mu_C,$$

$$H_1:\ \mu_A,\mu_B,\mu_C \text{ 不全相等}.$$

将自 A,B,C 三工厂抽得的样本依次记为 $x_{11},x_{21},\cdots,x_{n_1 1};x_{12},x_{22},\cdots,x_{n_2 2};x_{13},x_{23},\cdots,x_{n_3 3}$. 今 $s=3,n_1=n_2=n_3=5,n=n_1+n_2+n_3=15$，

$$T_{\cdot 1}=\sum_{i=1}^{5}x_{i1}=213,\quad T_{\cdot 2}=\sum_{i=1}^{5}x_{i2}=150,$$

$$T_{\cdot 3}=\sum_{i=1}^{5}x_{i3}=222,\quad T_{\cdot\cdot}=\sum_{j=1}^{3}\sum_{i=1}^{5}x_{ij}=585.$$

$$S_T=\sum_{j=1}^{3}\sum_{i=1}^{n_j}x_{ij}^2-T_{\cdot\cdot}^2/n=23\ 647-22\ 815=832,$$

$$S_A=\sum_{j=1}^{3}T_{\cdot j}^2/n_j-T_{\cdot\cdot}^2/n=23\ 430.6-22\ 815=615.6,$$

$$S_E=S_T-S_A=216.4,$$

S_T,S_A,S_E 的自由度分别为 $n-1=15-1=14,s-1=2,n-s=15-3=12$，从而得方差分析表如下：

方差来源	平方和	自由度	均方	F 比（$\alpha=0.05$）
因素 A	615.6	2	$\bar{S}_A=307.8$	$\bar{S}_A/\bar{S}_E=17.07$
误差 E	216.4	12	$\bar{S}_E=18.03$	
总和 T	832	14		

因 $F_{0.05}(2,12)=3.89$，$F=17.07>3.89$，故在显著性水平 0.05 下拒绝 H_0，认为平均寿命的差异是显著的.

下面来求 $\mu_A-\mu_B$，$\mu_A-\mu_C$ 和 $\mu_B-\mu_C$ 的置信水平为 $1-\alpha=0.95$ 的置信区间.

现在

$$t_{\alpha/2}(n-s)=t_{0.025}(15-3)=2.1788,$$

$$t_{0.025}(12)\sqrt{\bar{S}_E\left(\frac{1}{n_i}+\frac{1}{n_k}\right)}=5.85 \quad (i,k=1,2,3),$$

$$\hat{\mu}_A=\bar{x}_{\cdot 1}=T_{\cdot 1}/5=42.6,$$

$$\hat{\mu}_B=\bar{x}_{\cdot 2}=T_{\cdot 2}/5=30,$$

$$\hat{\mu}_C=\bar{x}_{\cdot 3}=T_{\cdot 3}/5=44.4,$$

从而分别得 $\mu_A-\mu_B$，$\mu_A-\mu_C$，$\mu_B-\mu_C$ 的一个置信水平为 95%的置信区间为

$$(\hat{\mu}_A-\hat{\mu}_B\pm 5.85)=(6.75,\ 18.45),$$

$$(\hat{\mu}_A-\hat{\mu}_C\pm 5.85)=(-7.65,\ 4.05),$$

$$(\hat{\mu}_B-\hat{\mu}_C\pm 5.85)=(-20.25,\ -8.55).$$

2. 为了寻找飞机控制板上仪器表的最佳布置，试验了三个方案，观察领航员在紧急情况的反应时间（以 $\frac{1}{10}$ s 计），随机地选择 28 名领航员，得到他们对于不同的布置方案的反应时间如下：

方案Ⅰ	14	13	9	15	11	13	14	11				
方案Ⅱ	10	12	7	11	8	12	9	10	13	9	10	9
方案Ⅲ	11	5	9	10	6	8	8	7				

试在显著性水平 0.05 下检验各个方案的反应时间有无显著差异. 若有差异，试求 $\mu_1-\mu_2$，$\mu_1-\mu_3$，$\mu_2-\mu_3$ 的置信水平为 0.95 的置信区间.

解 记第 i 种方案为 A_i，$i=1,2,3$. 各方案构成的总体平均反应时间为 μ_1，

μ_2,μ_3.则所述问题为:在显著性水平 $\alpha=0.05$ 下,检验假设

$$H_0:\mu_1=\mu_2=\mu_3,$$
$$H_1:\mu_1,\mu_2,\mu_3\text{ 不全相等.}$$

本题中,$n_1=8,n_2=12,n_3=8,n=28,s=3,T_{\cdot 1}=100,T_{\cdot 2}=120,T_{\cdot 3}=64$,$T_{\cdot\cdot}=284$.

$$S_T=\sum_{j=1}^{3}\sum_{i=1}^{n_j}x_{ij}^2-T_{\cdot\cdot}^2/n=3\ 052-284^2/28=171.43,$$
$$S_A=\sum_{j=1}^{3}T_{\cdot j}^2/n_j-T_{\cdot\cdot}^2/n=2\ 962-2\ 880.57=81.43,$$
$$S_E=S_T-S_A=90.$$

又 S_T,S_A,S_E 的自由度分别为 $n-1=28-1=27,s-1=3-1=2,n-s=28-3=25$,从而得方差分析表如下:

方差来源	平方和	自由度	均方	F 比 ($\alpha=0.05$)
因素 A	81.43	2	$\bar{S}_A=40.715$	$\bar{S}_A/\bar{S}_E=11.3$
误差 E	90	25	$\bar{S}_E=3.6$	
总和 T	171.43	27		

因 $F_{0.05}(2,25)=3.39,F=11.3>3.39$.故在显著性水平 $\alpha=0.05$ 下拒绝 H_0,认为差异是显著的.

以下来求置信水平为 $1-\alpha=0.95$ 的置信区间,今 $n_1=n_3=8,n_2=12$,$t_{0.025}(25)=2.059\ 5$,

$$t_{\alpha/2}(25)\sqrt{\bar{S}_E\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}=t_{\alpha/2}(25)\sqrt{\bar{S}_E\left(\frac{1}{n_2}+\frac{1}{n_3}\right)}$$
$$=2.059\ 5\times\sqrt{3.6\times\left(\frac{1}{8}+\frac{1}{12}\right)}=1.78,$$
$$t_{\alpha/2}(25)\sqrt{\bar{S}_E\left(\frac{1}{n_1}+\frac{1}{n_3}\right)}=2.059\ 5\times\sqrt{3.6\times\left(\frac{1}{8}+\frac{1}{8}\right)}=1.95,$$
$$\hat{\mu}_1=\bar{x}_{\cdot 1}=T_{\cdot 1}/8=12.5,$$
$$\hat{\mu}_2=\bar{x}_{\cdot 2}=T_{\cdot 2}/12=10,$$
$$\hat{\mu}_3=\bar{x}_{\cdot 3}=T_{\cdot 3}/8=8,$$

从而分别得 $\mu_1-\mu_2,\mu_1-\mu_3,\mu_2-\mu_3$ 的一个置信水平为 0.95 的置信区间为

$$(\hat{\mu}_1-\hat{\mu}_2\pm 1.78)=(0.72,\ 4.28),$$
$$(\hat{\mu}_1-\hat{\mu}_3\pm 1.95)=(2.55,\ 6.45),$$

$$(\hat{\mu}_2-\hat{\mu}_3\pm1.78)=(0.22,\ 3.78).$$

由此可见，若仅从得到的样本作出决策，则以方案Ⅲ为佳.

3. 某防治站对 4 个林场的松毛虫密度进行调查，每个林场调查 5 块地得资料如下表：

地点	松毛虫密度(头/标准地)				
A_1	192	189	176	185	190
A_2	190	201	187	196	200
A_3	188	179	191	183	194
A_4	187	180	188	175	182

判断 4 个林场松毛虫密度有无显著差异，取显著性水平 $\alpha=0.05$.

解 记 A_i 林场的平均松毛虫密度为 $\mu_i, i=1,2,3,4$，则所述问题为在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0: \mu_1=\mu_2=\mu_3=\mu_4,$$

$$H_1: \mu_1,\mu_2,\mu_3,\mu_4 \text{ 不全相等}.$$

今 $s=4, n_1=n_2=n_3=n_4=5, n=20$. $T_{\cdot 1}=932, T_{\cdot 2}=974, T_{\cdot 3}=935, T_{\cdot 4}=912$. $T_{\cdot\cdot}=3\ 753$.

$$S_T=\sum_{j=1}^{4}\sum_{i=1}^{5}x_{ij}^2-T_{\cdot\cdot}^2/n=705\ 225-3\ 753^2/20=974.55.$$

$$S_A=\sum_{j=1}^{4}\frac{T_{\cdot j}^2}{5}-T_{\cdot\cdot}^2/n=704\ 653.8-704\ 250.45=403.35.$$

$$S_E=S_T-S_A=571.2.$$

S_T, S_A, S_E 的自由度分别为 $n-1=19, s-1=3, n-s=20-4=16$，从而得方差分析表如下：

方差来源	平方和	自由度	均方	F 比
因素 A	403.35	3	134.45	3.766
误差 E	571.2	16	35.7	
总和 T	974.55	19		

因 $F_{0.05}(3,16)=3.24, F=3.766>3.24$，故在显著性水平 0.05 下拒绝 H_0，认为差异是显著的.

4. 一试验用来比较 4 种不同药品解除外科手术后疼痛的延续时间(以 h 计)，结果如下表：

药品	时间长度(h)				
A	8	6	4	2	
B	6	6	4	4	
C	8	10	10	10	12
D	4	4	2		

试在显著性水平 $\alpha=0.05$ 下检验各种药品对解除疼痛的延续时间有无显著差异.

解　将题中表略作改变以便于计算，如下：

药品号	时间长度(h)					$T_{\cdot j}$
1	8	6	4	2		20
2	6	6	4	4		20
3	8	10	10	10	12	50
4	4	4	2			10
					$T_{\cdot\cdot}$	100

并用 μ_i 表示第 i 号药品的平均缓解疼痛的延续时间，$i=1,2,3,4$，则所述问题为在显著性水平 $\alpha=0.05$ 下，检验假设

$$H_0:\mu_1=\mu_2=\mu_3=\mu_4,$$

$$H_1:\mu_1,\mu_2,\mu_3,\mu_4\text{ 不全相等}.$$

本题中 $n_1=n_2=4,n_3=5,n_4=3,n=16,s=4,T_{\cdot j}$ 及 $T_{\cdot\cdot}$ 见上表，于是

$$S_T=\sum_{j=1}^{4}\sum_{i=1}^{n_j}x_{ij}^2-T_{\cdot\cdot}^2/n=768-625=143.$$

$$S_A=\sum_{j=1}^{4}T_{\cdot j}^2/n_j-T_{\cdot\cdot}^2/n=733.33-625=108.33.$$

$$S_E=S_T-S_A=34.67.$$

S_T,S_A,S_E 的自由度分别为 $n-1=15,s-1=3,n-s=12$，从而得方差分析表如下：

方差来源	平方和	自由度	均方	F 比
因素 A	108.33	3	36.11	12.49
误差 E	34.67	12	2.89	
总和 T	143	15		

因 $F_{0.05}(3,12)=3.49$，$F=12.49>3.49$，故在显著性水平 0.05 下，认为药品对解除疼痛的延续时间的差异是显著的.

5. 将抗生素注入人体会产生抗生素与血浆蛋白质结合的现象，以致减少了药效. 下表列出 5 种常用的抗生素注入牛的体内时，抗生素与血浆蛋白质结合的百分比(以%计).

青霉素	四环素	链霉素	红霉素	氯霉素
29.6	27.3	5.8	21.6	29.2
24.3	32.6	6.2	17.4	32.8
28.5	30.8	11.0	18.3	25.0
32.0	34.8	8.3	19.0	24.2

试在显著性水平 $\alpha=0.05$ 下检验这些百分比的均值有无显著的差异.

解 以 $\mu_1,\mu_2,\mu_3,\mu_4,\mu_5$ 依次表示青霉素、四环素、链霉素、红霉素、氯霉素与血浆蛋白质结合的百分比的均值. 本题需检验假设

$$H_0: \quad \mu_1=\mu_2=\mu_3=\mu_4=\mu_5,$$
$$H_1: \quad \mu_1,\mu_2,\mu_3,\mu_4,\mu_5 \text{ 不全相等}.$$

现在 $n_1=n_2=n_3=n_4=n_5=4, n=\sum_{i=1}^{5} n_i=20, T_{\cdot 1}=114.4, T_{\cdot 2}=125.5, T_{\cdot 3}=31.3, T_{\cdot 4}=76.3, T_{\cdot 5}=111.2, T_{\cdot\cdot}=458.7$.

$$\begin{aligned} S_T &= \sum_{j=1}^{5}\sum_{i=1}^{4} x_{ij}^2 - T_{\cdot\cdot}^2/n \\ &= 12\,136.93 - 10\,520.284\,5 = 1\,616.645\,5, \\ S_A &= \sum_{j=1}^{5} T_{\cdot j}^2/n_j - T_{\cdot\cdot}^2/n \\ &= 12\,001.107\,5 - 10\,520.284\,5 = 1\,480.823, \\ S_E &= S_T - S_A \\ &= 1\,616.645\,5 - 1\,480.823 = 135.822\,5. \end{aligned}$$

S_T, S_A, S_E的自由度分别为 19,4,15，从而得方差分析表如下：

方差来源	平方和	自由度	均方	F 比
因素 A	1 480.823	4	370.205 8	40.89
误差 E	135.822 5	15	9.054 8	
总和 T	1 616.645 5	19		

因 $F_{0.05}(4,15)=3.06, F=40.89>3.06$，故在显著性水平 0.05 下拒绝 H_0，认为百分比的均值有显著的差异.

6. 下表给出某种化工过程在三种浓度、四种温度水平下得率(以%计)的数据：

浓度(因素 A)	温度(因素 B)							
	10℃		24℃		38℃		52℃	
2%	14	10	11	11	13	9	10	12
4%	9	7	10	8	7	11	6	10
6%	5	11	13	14	12	13	14	10

试在显著性水平 $\alpha=0.05$ 下检验：在不同浓度下得率的均值是否有显著差异，在不同温度下得率的均值是否有显著差异，交互作用的效应是否显著.

解　将浓度 A 的效应记为 $\alpha_1,\alpha_2,\alpha_3$，将温度 B 的效应记为 $\beta_1,\beta_2,\beta_3,\beta_4$，交互作用 $A\times B$ 的效应记为 γ_{ij}，$i=1,2,3$，$j=1,2,3,4$，按题意需检验假设

$$\begin{cases}H_{01}:\ \alpha_1=\alpha_2=\alpha_3=0,\\ H_{11}:\ \alpha_1,\alpha_2,\alpha_3\text{ 不全为 }0.\end{cases}$$

$$\begin{cases}H_{02}:\ \beta_1=\beta_2=\beta_3=\beta_4=0,\\ H_{12}:\ \beta_1,\beta_2,\beta_3,\beta_4\text{ 不全为 }0.\end{cases}$$

$$\begin{cases}H_{03}:\ \gamma_{11}=\gamma_{12}=\cdots=\gamma_{34}=0,\\ H_{13}:\ \gamma_{11},\gamma_{12},\cdots,\gamma_{34}\text{ 不全为 }0.\end{cases}$$

$T_{ij\cdot},T_{i\cdot\cdot},T_{\cdot j\cdot},T_{\cdot\cdot\cdot}$ 的计算如下表：

	$T_{ij\cdot}$				$T_{i\cdot\cdot}$
	24	22	22	22	90
	16	18	18	16	68
	16	27	25	24	92
$T_{\cdot j\cdot}$	56	67	65	62	$T_{\cdot\cdot\cdot}=250$

今 $r=3,s=4,t=2$，故有

$$S_T=\sum_{i=1}^{3}\sum_{j=1}^{4}\sum_{k=1}^{2}x_{ijk}^2-\frac{T_{\cdot\cdot\cdot}^2}{rst}$$

$$=(14^2+10^2+\cdots+10^2)-\frac{250^2}{24}$$

$$=2\,752-2\,604.166\,7=147.833\,3.$$

$$S_A=\frac{1}{st}\sum_{i=1}^{3}T_{i\cdot\cdot}^2-\frac{T_{\cdot\cdot\cdot}^2}{rst}$$

$$=\frac{1}{4\times 2}\times(90^2+68^2+92^2)-2\,604.166\,7$$

$$=2\,648.5-2\,604.166\,7=44.333\,3,$$

$$S_B=\frac{1}{rt}\sum_{j=1}^{4}T_{\cdot j\cdot}^2-\frac{T_{\cdots}^2}{rst}$$

$$=\frac{1}{3\times 2}\times(56^2+67^2+65^2+62^2)-2\,604.166\,7$$

$$=2\,615.666\,7-2\,604.166\,7=11.5,$$

$$S_{A\times B}=\frac{1}{t}\sum_{i=1}^{3}\sum_{j=1}^{4}T_{ij\cdot}^2-\frac{T_{\cdots}^2}{rst}-S_A-S_B$$

$$=\frac{1}{2}\times(24^2+16^2+\cdots+24^2)-2\,604.166\,7-S_A-S_B$$

$$=2\,687-2\,604.166\,7-44.333\,3-11.5=27,$$

$$S_E=S_T-S_A-S_B-S_{A\times B}=65.$$

$S_T,S_A,S_B,S_{A\times B},S_E$ 的自由度分别为 $rst-1=23,r-1=2,s-1=3,(r-1)\cdot(s-1)=6,rs(t-1)=12$. 从而得方差分析表如下：

方差来源	平方和	自由度	均方	F 比
因素 A	44.333 3	2	$\bar{S}_A=22.166\,7$	$F_A=\bar{S}_A/\bar{S}_E=4.09$
因素 B	11.5	3	$\bar{S}_B=3.833\,3$	$F_B=\bar{S}_B/\bar{S}_E=0.71$
交互作用 $A\times B$	27	6	$\bar{S}_{A\times B}=4.5$	$F_{A\times B}=\bar{S}_{A\times B}/\bar{S}_E=0.83$
误差 E	65	12	$\bar{S}_E=5.416\,7$	
总和 T	147.833 3	23		

因 $F_{0.05}(2,12)=3.89<4.09$，故拒绝 H_{01}；$F_{0.05}(3,12)=3.49>0.71$，故接受 H_{02}；$F_{0.05}(6,12)=3.00>0.83$，故接受 H_{03}，即认为在不同的浓度下得率均值的差异显著，而在不同的温度下得率均值的差异以及交互作用的效应均不显著.

因交互作用的效应不显著，我们将 $A\times B$ 一栏的平方和与自由度分别加到误差 E 这一栏中去，作为新的误差项，重新作方差分析，以提高分析的精度，现作出方差分析表如下：

方差来源	平方和	自由度	均方	F 比
因素 A	44.333 3	2	$\bar{S}_A=22.166\,7$	$F_A=\bar{S}_A/\bar{S}_E=4.34$
因素 B	11.5	3	$\bar{S}_B=3.833\,3$	$F_B=\bar{S}_B/\bar{S}_E=0.75$
误差 E	92	18	$\bar{S}_E=5.111\,1$	
总和 T	147.833 3	23		

因 $F_{0.05}(2,18)=3.55<4.34$，故拒绝 H_{01}；$F_{0.05}(3,18)=3.16>0.75$，故接受 H_{02}，即认为在不同浓度下得率均值的差异显著，而在不同温度下得率均值的差异不显著，这一结论与刚才的结论一样.

7. 为了研究某种金属管防腐蚀的功能，考虑了 4 种不同的涂料涂层. 将金属管埋设在 3 种不同性质的土壤中，经历了一定时间，测得金属管腐蚀的最大深度如下表所示(以 mm 计)：

涂层(因素 A)	土壤类型(因素 B)		
	1	2	3
1	1.63	1.35	1.27
2	1.34	1.30	1.22
3	1.19	1.14	1.27
4	1.30	1.09	1.32

试取显著性水平 $\alpha=0.05$ 检验在不同涂层下腐蚀的最大深度的平均值有无显著差异，在不同土壤下腐蚀的最大深度的平均值有无显著差异. 设两因素间没有交互作用效应.

解 本题是双因素无重复试验的方差分析问题. 需要在显著性水平 $\alpha=0.05$ 下，检验假设

$$\begin{cases} H_{01}: & \alpha_1=\alpha_2=\alpha_3=\alpha_4=0, \\ H_{11}: & \alpha_1,\alpha_2,\alpha_3,\alpha_4 \text{ 不全为零}. \end{cases}$$

$$\begin{cases} H_{02}: & \beta_1=\beta_2=\beta_3=0, \\ H_{12}: & \beta_1,\beta_2,\beta_3 \text{ 不全为零}. \end{cases}$$

今 $r=4, s=3, n=3\times4=12$，

$$T_{1\cdot}=4.25,\quad T_{2\cdot}=3.86,\quad T_{3\cdot}=3.6,\quad T_{4\cdot}=3.71,$$

$$T_{\cdot1}=5.46,\quad T_{\cdot2}=4.88,\quad T_{\cdot3}=5.08,\quad T_{\cdot\cdot}=15.42.$$

$$S_T=\sum_{i=1}^{4}\sum_{j=1}^{3}x_{ij}^2-\frac{T_{\cdot\cdot}^2}{rs}$$

$$=20.0154-19.8147=0.2007,$$

$$S_A=\sum_{i=1}^{4}\frac{T_{i\cdot}^2}{s}-\frac{T_{\cdot\cdot}^2}{rs}$$

$$=19.8954-19.8147=0.0807,$$

$$S_B=\sum_{j=1}^{3}\frac{T_{\cdot j}^2}{r}-\frac{T_{\cdot\cdot}^2}{rs}$$

$$=19.8581-19.8147=0.0434,$$

$$S_E = S_T - S_A - S_B$$
$$= 0.2007 - 0.0807 - 0.0434 = 0.0766.$$

S_T, S_A, S_B, S_E 的自由度分别是 $rs-1=11, r-1=3, s-1=2, (r-1)\times(s-1)=6$，从而得方差分析表如下：

方差来源	平方和	自由度	均　方	F 比
因素 A	0.080 7	3	$\bar{S}_A=0.0269$	$\bar{S}_A/\bar{S}_E=2.106$
因素 B	0.043 4	2	$\bar{S}_B=0.0217$	$\bar{S}_B/\bar{S}_E=1.70$
误差 E	0.076 6	6	$\bar{S}_E=0.01277$	
总　和	0.200 7	11		

由于 $F_{0.05}(3,6)=4.76>2.106, F_{0.05}(2,6)=5.14>1.70$，故接受 H_{01} 及 H_{02}，即认为不同涂层和不同土壤下腐蚀的最大深度的平均值均无显著差异.

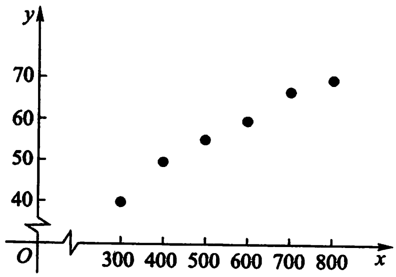
8. 下表数据是退火温度 x(以℃计)对黄铜延性 Y 效应的试验结果，Y 是以延长度计算的.

x(℃)	300	400	500	600	700	800
y(%)	40	50	55	60	67	70

画出散点图并求 Y 对于 x 的线性回归方程.

解　散点图如题 9.8 图. 从图上看，取回归函数为 x 的线性函数 $a+bx$ 是合适的. 现在 $n=6$，为求线性回归方程，所需计算列表如下：

	x	y	x^2	xy
	300	40	90 000	12 000
	400	50	160 000	20 000
	500	55	250 000	27 500
	600	60	360 000	36 000
	700	67	490 000	46 900
	800	70	640 000	56 000
Σ	3 300	342	1 990 000	198 400



题 9.8 图

$$S_{xx}=1\,990\,000-\frac{1}{6}\times 3\,300^2=175\,000,$$

$$S_{xy}=198\,400-\frac{1}{6}\times 3\,300\times 342=10\,300,$$

从而

$$\hat{b}=\frac{S_{xy}}{S_{xx}}=0.058\ 857,$$

$$\hat{a}=\frac{1}{6}\times 342-\frac{1}{6}\times 3\ 300\times 0.058\ 857=24.628\ 65.$$

线性回归方程为

$$\hat{y}=24.628\ 7+0.058\ 86x.$$

9. 在钢线碳含量对于电阻的效应的研究中，得到以下的数据：

碳含量 $x(\%)$	0.10	0.30	0.40	0.55	0.70	0.80	0.95
20℃时电阻 $y(\mu\Omega)$	15	18	19	21	22.6	23.8	26

(1) 画出散点图.

(2) 求线性回归方程 $\hat{y}=\hat{a}+\hat{b}x$.

(3) 求 ε 的方差 σ^2 的无偏估计.

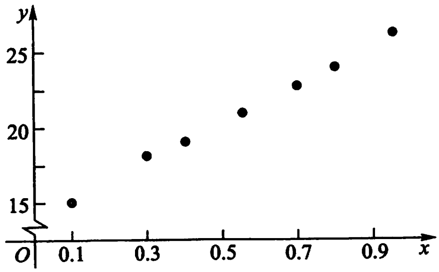
(4) 检验假设 $H_0:b=0,H_1:b\neq 0$.

(5) 若回归效果显著，求 b 的置信水平为 0.95 的置信区间.

(6) 求 $x=0.50$ 处 $\mu(x)$ 的置信水平为 0.95 的置信区间.

(7) 求 $x=0.50$ 处观察值 Y 的置信水平为 0.95 的预测区间.

解　(1) 散点图如题 9.9 图. 从图上看取回归函数为 x 的线性函数 $a+bx$ 是合适的.



题 9.9 图

(2) 由给定的数据经计算，得到 $n=7,\sum x_i=3.8,\sum y_i=145.4,\sum x_i^2=2.595,\sum y_i^2=3\ 104.2,\sum x_iy_i=85.61$.

$$S_{xx}=\sum x_i^2-\frac{1}{n}(\sum x_i)^2=0.532\ 142\ 857,$$

$$S_{yy}=\sum y_i^2-\frac{1}{n}(\sum y_i)^2=84.034\ 285\ 71,$$

$$S_{xy}=\sum x_iy_i-\frac{1}{n}(\sum x_i)(\sum y_i)=6.678\ 571\ 429.$$

$$\hat{b}=\frac{S_{xy}}{S_{xx}}=12.550\ 335\ 57,$$

$$\hat{a}=\frac{1}{n}\sum y_i-\frac{\hat{b}}{n}\sum x_i=13.958\ 389\ 26,$$

所以线性回归方程为

$$\hat{y}=13.958\ 4+12.550\ 3x.$$

(3) $\hat{\sigma}^2=Q_e/(n-2)=\dfrac{S_{yy}-\hat{b}S_{xy}}{5}=0.043\ 194\ 63.$

(4) $H_0:b=0,H_1:b\neq0.$

检验统计量为

$$t=\frac{\hat{b}}{\sqrt{\hat{\sigma}^2}}\sqrt{S_{xx}},$$

拒绝域为

$$|t|\geqslant t_{\alpha/2}(n-2)=t_{\alpha/2}(5).$$

因 t 的观察值为 $t=44.050\ 9\gg t_{0.005}(5)=4.032\ 2$(表载 t 分布的临界值 α 最小的是 0.005)，由此知回归效果是极其显著的.

(5) 给定的置信水平为 $1-\alpha=0.95$，故 $\frac{\alpha}{2}=0.025$，$t_{\alpha/2}(n-2)=t_{0.025}(5)=2.570\ 6$. 即知

$$t_{\alpha/2}(n-2)\frac{\hat{\sigma}}{\sqrt{S_{xx}}}=t_{0.025}(5)\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$$
$$=2.570\ 6\times0.284\ 91=0.732\ 4.$$

得 b 的置信水平为 0.95 的置信区间为

$$\left(\hat{b}\pm t_{\alpha/2}(n-2)\sqrt{\hat{\sigma}^2/S_{xx}}\right)=(12.550\ 3\pm0.732\ 4)$$
$$=(11.817\ 9,\ 13.282\ 7).$$

(6) $x=x_0=0.50$ 处对应的 Y 的估计值为

$$\hat{y}_0=13.958\ 4+12.550\ 3\times0.50=20.233\ 55,$$

置信水平为 $1-\alpha=0.95$，$\alpha/2=0.025$，$t_{\alpha/2}(n-2)=t_{0.025}(5)=2.570\ 6$. 故

$$t_{\alpha/2}(n-2)\hat{\sigma}\sqrt{\frac{1}{n}+(x_0-\bar{x})^2/S_{xx}}=0.204\ 4,$$

从而得 $\mu(x_0)$ 的一个置信水平为 0.95 的置信区间为

$$(\hat{y}_0 \pm 0.204\,4) = (20.03,\ 20.44).$$

(7) 以上已求得 $\hat{y}_0 = \hat{a} + \hat{b}x_0 = 20.233\,55$，$t_{\alpha/2}(n-2)$同上，可得

$$t_{\alpha/2}(n-2)\hat{\sigma}\sqrt{1+\frac{1}{n}+(x_0-\bar{x})^2/S_{xx}} = 0.572\,0\ .$$

于是得 $x=0.50$ 处，观察值 Y 的一个置信水平为 0.95 的预测区间为

$$(\hat{y}_0 \pm 0.572\,0) = (19.66,\ 20.81).$$

10. 下表列出了 18 名 5～8 岁儿童的体重(这是容易测得的)和体积(这是难以测量的)：

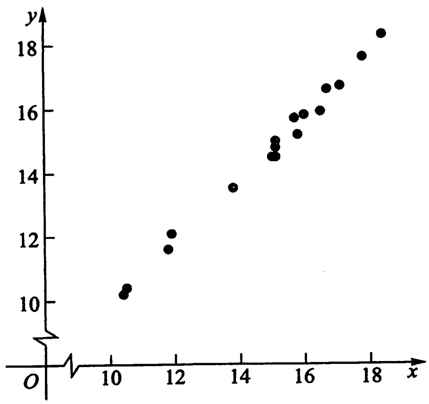
体重 x(kg)	17.1	10.5	13.8	15.7	11.9	10.4	15.0	16.0	17.8
体积 y(dm^3)	16.7	10.4	13.5	15.7	11.6	10.2	14.5	15.8	17.6
体重 x(kg)	15.8	15.1	12.1	18.4	17.1	16.7	16.5	15.1	15.1
体积 y(dm^3)	15.2	14.8	11.9	18.3	16.7	16.6	15.9	15.1	14.5

(1) 画出散点图.

(2) 求 Y 关于 x 的线性回归方程 $\hat{y} = \hat{a} + \hat{b}x$.

(3) 求 $x=14.0$ 时 Y 的置信水平为 0.95 的预测区间.

解 (1) 散点图如题 9.10 图.



题 9.10 图

(2) $n=18$，$\sum x_i = 270.1$，$\sum x_i^2 = 4\,149.39$，

$\bar{x} = 15.005\,555\,56$， $\sum y_i = 265$， $\sum y_i^2 = 3\,996.14$，

$\bar{y} = 14.722\,222\,22$， $\sum x_i y_i = 4\,071.71$，

$$S_{xx}=\sum x_i^2-\frac{1}{n}(\sum x_i)^2=96.389\ 444\ 44,$$

$$S_{yy}=\sum y_i^2-\frac{1}{n}(\sum y_i)^2=94.751\ 111\ 11,$$

$$S_{xy}=\sum x_i y_i-\frac{1}{n}(\sum x_i)(\sum y_i)=95.237\ 777\ 78,$$

$$\hat{b}=\frac{S_{xy}}{S_{xx}}=0.988\ 051\ 942\approx 0.988,$$

$$\hat{a}=\frac{1}{n}\sum y_i-\frac{\hat{b}}{n}\sum x_i=-0.104\ 046\ 085\approx -0.104.$$

故线性回归方程为

$$\hat{y}=-0.104+0.988x.$$

(3) $1-\alpha=0.95, \alpha=0.05$,

$t_{0.025}(16)=2.119\ 9$,

$Q_e=S_{yy}-\hat{b}S_{xy}=0.651\ 239\ 832$,

$\hat{\sigma}^2=\frac{Q_e}{16}=0.201\ 748\ 579^2$,

$t_{\alpha/2}(16)\hat{\sigma}\sqrt{1+\frac{1}{18}+(14.0-\bar{x})^2/S_{xx}}=0.442.$

在 $x=14.0$ 处 Y 的观察值的预测值为

$$\hat{y}|_{x=14.0}=-0.104+0.988x|_{x=14.0}=13.728,$$

得 $x=14.0$ 处观察值 Y 的一个置信水平为 0.95 的预测区间为

$$(13.728\pm 0.442)=(13.29, 14.17).$$

11. 蟋蟀用一个翅膀在另一翅膀上快速地滑动，从而发出吱吱喳喳的叫声. 生物学家知道叫声的频率 x 与气温 Y 具有线性关系. 下表列出了 15 对频率与气温间的对应关系的观察结果：

频率 x_i(叫声数/s)	20.0	16.0	19.8	18.4	17.1	15.5	14.7	17.1
气温 y_i(℃)	31.4	22.0	34.1	29.1	27.0	24.0	20.9	27.8
频率 x_i(叫声数/s)	15.4	16.2	15.0	17.2	16.0	17.0	14.4	
气温 y_i(℃)	20.8	28.5	26.4	28.1	27.0	28.6	24.6	

试求 Y 关于 x 的线性回归方程.

解 本题需求出 Y 关于 x 的线性回归函数 $a+bx$. 为此，先将需要的计算列表如下：

	x	y	x^2	xy
	20.0	31.4	400	628
	16.0	22.0	256	352
	19.8	34.1	392.04	675.18
	18.4	29.1	338.56	535.44
	17.1	27.0	292.41	461.7
	15.5	24.0	240.25	372
	14.7	20.9	216.09	307.23
	17.1	27.8	292.41	475.38
	15.4	20.8	237.16	320.32
	16.2	28.5	262.44	461.7
	15.0	26.4	225	396
	17.2	28.1	295.84	483.32
	16.0	27.0	256	432
	17.0	28.6	289	486.2
	14.4	24.6	207.36	354.24
$\sum$	249.8	400.3	4 200.56	6 740.71

$$S_{xx}=\sum x_i^2-\frac{1}{n}(\sum x_i)^2=4\ 200.56-\frac{1}{15}\times 249.8^2=40.557\ 333,$$

$$S_{xy}=\sum x_i y_i-\frac{1}{n}\sum x_i\sum y_i=6\ 740.71-\frac{1}{15}\times 249.8\times 400.3=74.380\ 667,$$

$$\hat{b}=\frac{S_{xy}}{S_{xx}}=1.833\ 963,$$

$$\hat{a}=\frac{1}{n}\sum y_i-\frac{\hat{b}}{n}\sum x_i=\frac{1}{15}\times 400.3-\frac{1}{15}\times 1.833\ 963\times 249.8=-3.854\ 930,$$

故线性回归方程为

$$\hat{y}=-3.854\ 93+1.833\ 96x.$$

12. 下面列出了自 1952—2004 年各届奥林匹克运动会男子 10 000 m 赛跑的冠军的成绩(时间以 min 计).

年份(x)	1952	1956	1960	1964	1968	1972	1976
成绩(y)	29.3	28.8	28.5	28.4	29.4	27.6	27.7
年份(x)	1980	1984	1988	1992	1996	2000	2004
成绩(y)	27.7	27.8	27.4	27.8	27.1	27.3	27.1

(1) 求 Y 关于 x 的线性回归方程 $\hat{y}=\hat{a}+\hat{b}x$.

(2) 检验假设 $H_0:b=0, H_1:b\neq0$(显著性水平 $\alpha=0.05$).

(3) 求 2008 年冠军成绩的预测值.

解 1. 成绩关于奥运届次的回归方程

对数据作变换.(i) 时间 x 原取值改为 1,2,3,…(即自 1952 年算作奥运男子万米赛跑的第一次记录，其后为第二次、第三次等).(ii) 把万米记录均减去 20 min 来算(这样在使用经验回归方程时，得到的时间加上 20 就是实际所要求的时间). 得经整理的数据及计算如下表：

x	y	x^2	y^2	xy
1	9.3	1	86.49	9.3
2	8.8	4	77.44	17.6
3	8.5	9	72.25	25.5
4	8.4	16	70.56	33.6
5	9.4	25	88.36	47
6	7.6	36	57.76	45.6
7	7.7	49	59.29	53.9
8	7.7	64	59.29	61.6
9	7.8	81	60.84	70.2
10	7.4	100	54.76	74
11	7.8	121	60.84	85.8
12	7.1	144	50.41	85.2
13	7.3	169	53.29	94.9
14	7.1	196	50.41	99.4
Σ 105	111.9	1 015	901.99	803.6

$$S_{xx}=\sum x_i^2-\frac{1}{n}(\sum x_i)^2=1\ 015-\frac{1}{14}\times105^2=227.5,$$

$$S_{xy}=\sum x_iy_i-\frac{1}{n}\sum x_i\sum y_i=803.6-\frac{1}{14}\times105\times111.9=-35.65,$$

$$S_{yy}=\sum y_i^2-\frac{1}{n}(\sum y_i)^2=901.99-\frac{1}{14}\times111.9^2=7.589\ 285\ 7.$$

(1) 设所要求的回归函数为 $a+bx$,则

$$\hat{b}=\frac{S_{xy}}{S_{xx}}=-0.156\ 7,$$

$$\hat{a}=\frac{1}{n}\sum y_i-\frac{\hat{b}}{n}\sum x_i=9.168.$$

故经验回归方程为

$$\hat{y}=9.168-0.156\,7x. \qquad (*_1)$$

(2) 需在显著性水平 0.05 下检验假设

$$H_0: b=0, \quad H_1: b\neq 0.$$

为此先计算

$$Q_e=S_{yy}-\hat{b}S_{xy}=7.589\,285\,7-(-0.156\,7)\times(-35.65)=2.002\,93,$$

$$\hat{\sigma}^2=\frac{Q_e}{n-2}=0.166\,911.$$

查表得知 $t_{0.025}(12)=2.178\,8$，今观察值

$$|t|=\frac{|\hat{b}|}{\hat{\sigma}}\sqrt{S_{xx}}=5.785\,2>2.178\,8.$$

故在显著性水平 $\alpha=0.05$ 下拒绝 H_0，认为回归效果是显著的.

(3) 预测值. 2008 年相当于第 15 次，即在($*_1$)式中令 $x=15$ 得

$$\hat{y}\Big|_{x=15}=9.168-0.156\,7\times 15=6.82(\text{min}).$$

注意，我们在以上作计算时，将历届记录的成绩都减去 20 min，因此，2008 年男子万米赛跑冠军成绩的预测值是 6.82+20=26.82(min).

2. 成绩关于年份的回归方程

此时，$S_{xx}=3\,640$，$S_{xy}=-142.6$. 回归方程为

$$\hat{y}=105.482\,6-0.039\,18x.$$

13. 以 x 与 Y 分别表示人的脚长与手长(均以英寸计，1 英寸=2.54 厘米)，下面列出了 15 名女子的脚的长度 x 与手的长度 Y 的样本值：

x	9.00	8.50	9.25	9.75	9.00	10.00	9.50	9.00
y	6.50	6.25	7.25	7.00	6.75	7.00	6.50	7.00
x	9.25	9.50	9.25	10.00	10.00	9.75	9.50	
y	7.00	7.00	7.00	7.50	7.25	7.25	7.25	

(1) 试求 Y 关于 x 的线性回归方程 $\hat{y}=\hat{a}+\hat{b}x$.

(2) 求 b 的置信水平为 0.95 的置信区间.

解　先作必要的计算见下表：

x	y	x^2	y^2	xy
9.00	6.50	81	42.25	58.5
8.50	6.25	72.25	39.062 5	53.125
9.25	7.25	85.562 5	52.562 5	67.062 5
9.75	7.00	95.062 5	49	68.25
9.00	6.75	81	45.562 5	60.75
10.00	7.00	100	49	70
9.50	6.50	90.25	42.25	61.75
9.00	7.00	81	49	63
9.25	7.00	85.562 5	49	64.75
9.50	7.00	90.25	49	66.5
9.25	7.00	85.562 5	49	64.75
10.00	7.50	100	56.25	75
10.00	7.25	100	52.562 5	72.5
9.75	7.25	95.062 5	52.562 5	70.687 5
9.50	7.25	90.25	52.562 5	68.875
$\sum$ 141.25	104.5	1 332.812 5	729.625	985.5

$$S_{xx}=\sum x_i^2-\frac{1}{n}(\sum x_i)^2=1\ 332.812\ 5-\frac{1}{15}\times 141.25^2=2.708\ 333,$$

$$S_{xy}=\sum x_i y_i-\frac{1}{n}\sum x_i\sum y_i=985.5-\frac{1}{15}\times 141.25\times 104.5=1.458\ 333,$$

$$S_{yy}=\sum y_i^2-\frac{1}{n}(\sum y_i)^2=729.625-\frac{1}{15}\times 104.5^2=1.608\ 333.$$

从而

(1)
$$\hat{b}=\frac{S_{xy}}{S_{xx}}=0.538\ 46,$$

$$\hat{a}=\frac{1}{n}\sum y_i-\frac{\hat{b}}{n}\sum x_i=\frac{1}{15}\times 104.5-\frac{0.538\ 46}{15}\times 141.25=1.896.$$

所求的线性回归方程为

$$\hat{y}=1.896+0.538\ 46x.$$

(2) 先计算

$$Q_e=S_{yy}-\hat{b}S_{xy}=1.608\ 333-0.538\ 46\times 1.458\ 333=0.823\ 079.$$

因 $n=15$，故

$$\hat{\sigma}^2=Q_e/(n-2)=0.063\ 31.$$

因 $t_{0.025}(13)=2.160\ 4$，所以 b 的置信水平为 0.95 的置信区间为

$$\left(\hat{b}\pm t_{\alpha/2}(n-2)\frac{\hat{\sigma}}{\sqrt{S_{xx}}}\right)=\left(0.538\ 46\pm 2.160\ 4\times\frac{\sqrt{0.063\ 31}}{\sqrt{2.708\ 333}}\right)$$
$$=(0.538\ 46\pm 0.330\ 31)=(0.208,\ 0.869).$$

14. 槲寄生是一种寄生在大树上部树枝上的寄生植物. 它喜欢寄生在年轻的大树上. 下面给出在一定条件下完成的试验中采集的数据：

大树的年龄 x(年)	3	4	9	15	40
每株大树上槲寄生的株数 y	28	10	15	6	1
	33	36	22	14	1
	22	24	10	9	

(1) 作出 (x_i, y_i) 的散点图.

(2) 令 $z_i=\ln y_i$，作出 (x_i, z_i) 的散点图.

(3) 以模型 $Y=ae^{bx}\varepsilon$，$\ln\varepsilon\sim N(0,\sigma^2)$ 拟合数据，其中 a,b,σ^2 与 x 无关. 试求曲线回归方程 $\hat{y}=\hat{a}\exp\{\hat{b}x\}$.

解　(1) 散点图如题 9.14 图(1).

(2) 令 $z_i=\ln y_i$，得数据如下表：

x_i	3	4	9	15	40
z_i	3.33	2.30	2.71	1.79	0
	3.50	3.58	3.09	2.64	0
	3.09	3.18	2.30	2.20	

由此作 (x_i, z_i) 的散点图如题 9.14 图(2). 上表仅供作散点图之用，作数值计算时，可直接由计算器(机)求得精度更高的数据.

(3) 将 $Y=ae^{bx}\varepsilon$ 取对数，得

$$\ln Y=\ln a+bx+\ln\varepsilon.$$

令 $Z=\ln Y$，则回归模型为

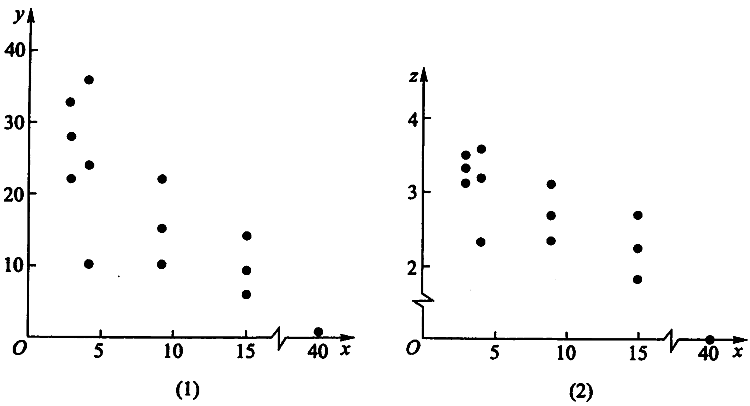
$$Z=\ln a+bx+\ln\varepsilon,$$

其中 $\ln\varepsilon\sim N(0,\sigma^2)$.

令 $n=14$，$\sum x_i=173$，$\sum x_i^2=4\ 193$ $\left(\text{其中}\sum\text{表示}\sum\limits_{i=1}^{n}\right)$，

$\sum z_i=33.713\ 631\ 51$，　$S_{xx}=2\ 055.214\ 286$，

$\sum x_i z_i=238.351\ 615\ 4$，　$S_{xz}=-178.252\ 545\ 4$，　$\hat{b}=-0.086\ 731\ 854$，



题 9.14 图

$\widehat{\ln a}=3.47987445$，　$e^{\widehat{\ln a}}=32.45564701$，

得曲线回归方程为

$$\hat{y}=32.4556e^{-0.0867319x}.$$

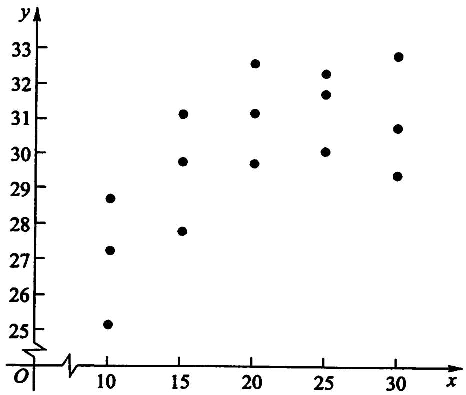
15. 一种合金在某种添加剂的不同浓度之下，各做三次试验，得数据如下：

浓度 x	10.0	15.0	20.0	25.0	30.0
抗压强度 y	25.2	29.8	31.2	31.7	29.4
	27.3	31.1	32.6	30.1	30.8
	28.7	27.8	29.7	32.3	32.8

(1) 作散点图.

(2) 以模型 $Y=b_0+b_1x+b_2x^2+\varepsilon,\varepsilon\sim N(0,\sigma^2)$ 拟合数据，其中 b_0,b_1,b_2,σ^2 与 x 无关. 求回归方程 $\hat{y}=\hat{b}_0+\hat{b}_1x+\hat{b}_2x^2$.

解 (1) 散点图如题 9.15 图.



题 9.15 图

(2) 令 $x_1=x, x_2=x^2$，则题中假设的模型可写成

$$Y=b_0+b_1x_1+b_2x_2+\varepsilon,\quad \varepsilon\sim N(0,\sigma^2).$$

本题要求利用给定的数据来估计系数 b_0, b_1, b_2.

引入矩阵

$$\boldsymbol{X}=\begin{bmatrix}1&10&100\\1&10&100\\1&10&100\\1&15&225\\1&15&225\\1&15&225\\1&20&400\\1&20&400\\1&20&400\\1&25&625\\1&25&625\\1&25&625\\1&30&900\\1&30&900\\1&30&900\end{bmatrix},\quad \boldsymbol{Y}=\begin{bmatrix}25.2\\27.3\\28.7\\29.8\\31.1\\27.8\\31.2\\32.6\\29.7\\31.7\\30.1\\32.3\\29.4\\30.8\\32.8\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}b_0\\b_1\\b_2\end{bmatrix}.$$

经计算得

$$\boldsymbol{X}^{\mathrm T}\boldsymbol{X}=\begin{bmatrix}15&300&6\ 750\\300&6\ 750&165\ 000\\6\ 750&165\ 000&4\ 263\ 750\end{bmatrix}=15\begin{bmatrix}1&20&450\\20&450&11\ 000\\450&11\ 000&284\ 250\end{bmatrix},$$

$$\boldsymbol{X}^{\mathrm T}\boldsymbol{Y}=\begin{bmatrix}450.5\\9\ 155\\207\ 990\end{bmatrix},$$

$$(\boldsymbol{X}^{\mathrm T}\boldsymbol{X})^{-1}=\frac{1}{26\ 250}\begin{bmatrix}138\ 250&-14\ 700&350\\-14\ 700&1\ 635&-40\\350&-40&1\end{bmatrix}.$$

得正规方程组的解为

$$\hat{\boldsymbol{B}}=\begin{bmatrix}\hat b_0\\\hat b_1\\\hat b_2\end{bmatrix}=(\boldsymbol{X}^{\mathrm T}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm T}\boldsymbol{Y}=\begin{bmatrix}19.033\ 33\\1.008\ 57\\-0.020\ 38\end{bmatrix}.$$

故回归方程为

$$\hat{y}=19.033\ 33+1.008\ 57x_1-0.020\ 38x_2,$$

即

$$\hat{y}=19.033\ 33+1.008\ 57x-0.020\ 38x^2.$$

16. 某种化工产品的得率 Y 与反应温度 x_1、反应时间 x_2 及某反应物浓度 x_3 有关. 今得试验结果如下表所示，其中 x_1,x_2,x_3 均为二水平且均以编码形式表达.

x_1	-1	-1	-1	-1	1	1	1	1
x_2	-1	-1	1	1	-1	-1	1	1
x_3	-1	1	-1	1	-1	1	-1	1
得率	7.6	10.3	9.2	10.2	8.4	11.1	9.8	12.6

（1）设 $\mu(x_1,x_2,x_3)=b_0+b_1x_1+b_2x_2+b_3x_3$，求 Y 的多元线性回归方程.

（2）若认为反应时间不影响得率，即认为

$$\mu(x_1,x_2,x_3)=\beta_0+\beta_1x_1+\beta_3x_3,$$

求 Y 的多元线性回归方程.

解 （1）引入矩阵

$$\boldsymbol{X}=\begin{bmatrix}1&-1&-1&-1\\1&-1&-1&1\\1&-1&1&-1\\1&-1&1&1\\1&1&-1&-1\\1&1&-1&1\\1&1&1&-1\\1&1&1&1\end{bmatrix},\quad \boldsymbol{Y}=\begin{bmatrix}7.6\\10.3\\9.2\\10.2\\8.4\\11.1\\9.8\\12.6\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}b_0\\b_1\\b_2\\b_3\end{bmatrix},$$

则所要求的线性回归模型为

$$Y=b_0+b_1x_1+b_2x_2+b_3x_3+\varepsilon,\quad \varepsilon\sim N(0,\sigma^2).$$

其正规方程为

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}.$$

易得

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}=\begin{bmatrix}8&0&0&0\\0&8&0&0\\0&0&8&0\\0&0&0&8\end{bmatrix},\quad \boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}=\begin{bmatrix}79.2\\4.6\\4.4\\9.2\end{bmatrix},$$

故

$$(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}=\operatorname{diag}\left(\frac{1}{8},\frac{1}{8},\frac{1}{8},\frac{1}{8}\right),$$

所以

$$\hat{\boldsymbol{B}}=\begin{bmatrix}\hat{b}_0\\ \hat{b}_1\\ \hat{b}_2\\ \hat{b}_3\end{bmatrix}=(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}=\begin{bmatrix}9.9\\ 0.575\\ 0.55\\ 1.15\end{bmatrix}.$$

所以多元线性回归方程为

$$\hat{y}=9.9+0.575x_1+0.55x_2+1.15x_3.$$

(2) 若认为 $\mu(x_1,x_2,x_3)=\beta_0+\beta_1x_1+\beta_3x_3$，则引入的8×3矩阵就是上述 $\boldsymbol{X}$ 中删去第 3 列后所得的矩阵，即

$$\boldsymbol{M}=\begin{bmatrix}1&-1&-1\\1&-1&1\\1&-1&-1\\1&-1&1\\1&1&-1\\1&1&1\\1&1&-1\\1&1&1\end{bmatrix},\quad \boldsymbol{\beta}=\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_3\end{bmatrix},$$

模型

$$Y=\beta_0+\beta_1x_1+\beta_3x_3+\varepsilon,\quad \varepsilon\sim N(0,\sigma^2)$$

的正规方程为

$$\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{\beta}=\boldsymbol{M}^{\mathrm{T}}\boldsymbol{Y}$$

($\boldsymbol{Y}$ 见(1))，则有

$$\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M}=\begin{bmatrix}8&0&0\\0&8&0\\0&0&8\end{bmatrix},\quad \boldsymbol{M}^{\mathrm{T}}\boldsymbol{Y}=\begin{bmatrix}79.2\\ 4.6\\ 9.2\end{bmatrix},$$

故

$$(\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M})^{-1}=\operatorname{diag}\left(\frac{1}{8},\frac{1}{8},\frac{1}{8}\right),$$

$$\hat{\boldsymbol{\beta}}=\begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \hat{\beta}_3\end{bmatrix}=\begin{bmatrix}9.9\\ 0.575\\ 1.15\end{bmatrix},$$

得多元线性回归方程为

$$\hat{y}=9.9+0.575x_1+1.15x_3.$$

第十二章 随机过程

1. 利用抛掷一枚硬币的试验定义一随机过程

$$X(t)=\begin{cases}\cos \pi t, & \text{出现 } H, \\ 2t, & \text{出现 } T,\end{cases} \quad t\in(-\infty,\infty).$$

假设 $P(H)=P(T)=\frac{1}{2}$，试确定 $X(t)$ 的

(1) 一维分布函数 $F\left(x;\frac{1}{2}\right)$，$F(x;1)$.

(2) 二维分布函数 $F\left(x_1,x_2;\frac{1}{2},1\right)$.

解 (1) 由 $X(t)$ 的定义

$$X\left(\frac{1}{2}\right)=\begin{cases}0, & \text{出现 } H, \\ 1, & \text{出现 } T.\end{cases}$$

这一离散型随机变量的分布律为

$X\left(\frac{1}{2}\right)$	0	1
p_k	$\frac{1}{2}$	$\frac{1}{2}$

其分布函数为

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

同理
$$X(1)=\begin{cases}-1, & \text{出现 } H, \\ 2, & \text{出现 } T,\end{cases}$$

其分布律为

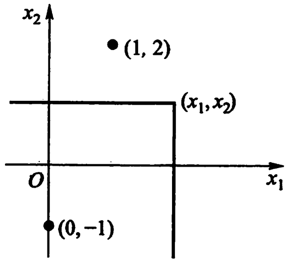
$X(1)$	-1	2
p_k	$\frac{1}{2}$	$\frac{1}{2}$

分布函数为

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

(2) 当 $t_1=\dfrac{1}{2},t_2=1$ 时，$\left(X\left(\dfrac{1}{2}\right),X(1)\right)$是一个二维离散型随机变量，且当硬币出现 H 时，它的取值为(0,−1)；当硬币出现 T 时，它的取值为(1,2). 由于硬币出现 H、出现 T 的概率均为$\dfrac{1}{2}$，因此 $X\left(\dfrac{1}{2}\right)$与 $X(1)$的联合分布律为

$X(1)$ \ $X\left(\frac{1}{2}\right)$	0	1
−1	$\frac{1}{2}$	0
2	0	$\frac{1}{2}$



题 12.1 图

$\left(X\left(\dfrac{1}{2}\right),X(1)\right)$的分布函数为

$$F\left(x_1,x_2;\frac{1}{2},1\right)=P\left\{X\left(\frac{1}{2}\right)\leqslant x_1,X(1)\leqslant x_2\right\}.$$

由题 12.1 图知，当 $x_1<0,-\infty<x_2<\infty$时，

$$F\left(x_1,x_2;\frac{1}{2},1\right)=0;$$

当 $x_1\geqslant 0,x_2<-1$ 时，

$$F\left(x_1,x_2;\frac{1}{2},1\right)=0;$$

当 $0\leqslant x_1<1,x_2\geqslant -1$ 时，

$$F\left(x_1,x_2;\frac{1}{2},1\right)=P\left\{\left(X\left(\frac{1}{2}\right),X(1)\right)=(0,-1)\right\}=\frac{1}{2};$$

当 $x_1\geqslant 1,-1\leqslant x_2<2$ 时，

$$F\left(x_1,x_2;\frac{1}{2},1\right)=P\left\{\left(X\left(\frac{1}{2}\right),X(1)\right)=(0,-1)\right\}=\frac{1}{2};$$

当 $x_1\geqslant 1,x_2\geqslant 2$ 时，

$$F\left(x_1,x_2;\frac{1}{2},1\right)=P\left\{\left(X\left(\frac{1}{2}\right),X(1)\right)=(0,-1)\right\}$$
$$+P\left\{\left(X\left(\frac{1}{2}\right),X(1)\right)=(1,2)\right\}$$
$$=\frac{1}{2}+\frac{1}{2}=1,$$

所以分布函数为

$$F\left(x_1,x_2;\frac{1}{2},1\right)=\begin{cases}0, & x_1<0,-\infty<x_2<\infty,\\ 0, & x_1\geqslant 0,x_2<-1,\\ \frac{1}{2}, & 0\leqslant x_1<1,x_2\geqslant -1,\\ \frac{1}{2}, & x_1\geqslant 1,-1\leqslant x_2<2,\\ 1, & x_1\geqslant 1,x_2\geqslant 2.\end{cases}$$

2. 给定随机过程$\{X(t),t\in T\}$，x 是任一实数，定义另一个随机过程

$$Y(t)=\begin{cases}1, & X(t)\leqslant x,\\ 0, & X(t)>x,\end{cases}\quad t\in T.$$

试将 $Y(t)$的均值函数和自相关函数用随机过程 $X(t)$的一维和二维分布函数来表示.

解 设随机过程$\{X(t),t\in T\}$的一维分布函数为 $F_X(x;t)$，二维分布函数为 $F_X(x_1,x_2;t_1,t_2)$，固定 t 时，$Y(t)$是服从(0－1)分布的随机变量，其分布律为

$Y(t)$	0	1
p_k	$P\{X(t)>x\}$	$P\{X(t)\leqslant x\}$

于是 $Y(t)$的均值为

$$\mu_Y(t)=E[Y(t)]=0\times P\{X(t)>x\}+1\times P\{X(t)\leqslant x\}$$
$$=P\{X(t)\leqslant x\}=F_X(x;t),$$

又随机变量 $Y(t_1)$和 $Y(t_2)$的联合分布律为

$Y(t_2)$ \ $Y(t_1)$	0	1
0	$P\{X(t_1)>x,X(t_2)>x\}$	$P\{X(t_1)\leqslant x,X(t_2)>x\}$
1	$P\{X(t_1)>x,X(t_2)\leqslant x\}$	$P\{X(t_1)\leqslant x,X(t_2)\leqslant x\}$

由教材第四章 §1 公式(1.6)及二维分布函数的定义，有

$$R_Y(t_1,t_2)=E[Y(t_1)Y(t_2)]$$

$$\begin{aligned}
&=0\times 0\times P\{X(t_1)>x,X(t_2)>x\}\\
&\quad+0\times 1\times P\{X(t_1)>x,X(t_2)\leqslant x\}\\
&\quad+1\times 0\times P\{X(t_1)\leqslant x,X(t_2)>x\}\\
&\quad+1\times 1\times P\{X(t_1)\leqslant x,X(t_2)\leqslant x\}\\
&=P\{X(t_1)\leqslant x,X(t_2)\leqslant x\}=F_X(x,x;t_1,t_2).
\end{aligned}$$

3. 设随机过程 $X(t)=e^{-At},t>0$，其中 A 是在区间 $(0,a)$ 上服从均匀分布的随机变量，试求 $X(t)$ 的均值函数和自相关函数.

解 由关于随机变量函数的数学期望的定理知道 $X(t)$ 的均值函数为

$$\begin{aligned}
\mu_X(t)&=E[X(t)]=E(e^{-At})=\int_0^a e^{-ut}\times\frac{1}{a}du\\
&=\frac{1}{at}(1-e^{-at}),\quad t>0.
\end{aligned}$$

自相关函数为

$$\begin{aligned}
R_X(t_1,t_2)&=E[X(t_1)X(t_2)]=E(e^{-At_1}\cdot e^{-At_2})\\
&=E[e^{-(t_1+t_2)A}]=\int_0^a e^{-(t_1+t_2)u}\times\frac{1}{a}du\\
&=\frac{1}{a(t_1+t_2)}[1-e^{-a(t_1+t_2)}],\quad t_1,t_2>0.
\end{aligned}$$

4. 设随机过程 $X(t)\equiv X$，其中 X 是一随机变量，$E(X)=a,D(X)=\sigma^2(\sigma>0)$，试求 $X(t)$ 的均值函数和协方差函数.

解

$$\mu_X(t)=E[X(t)]=E(X)=a.$$

$$\begin{aligned}
C_X(t_1,t_2)&=E\{[X(t_1)-\mu_X(t_1)][X(t_2)-\mu_X(t_2)]\}\\
&=E[(X-a)^2]=D(X)=\sigma^2.
\end{aligned}$$

5. 已知随机过程 $\{X(t),t\in T\}$ 的均值函数 $\mu_X(t)$ 和协方差函数 $C_X(t_1,t_2)$，$\varphi(t)$ 是普通的函数，试求随机过程 $Y(t)=X(t)+\varphi(t)$ 的均值函数和协方差函数.

解

$$\begin{aligned}
\mu_Y(t)&=E[Y(t)]=E[X(t)+\varphi(t)]\\
&=E[X(t)]+E[\varphi(t)]=\mu_X(t)+\varphi(t).
\end{aligned}$$

$$\begin{aligned}
C_Y(t_1,t_2)&=E\{[Y(t_1)-\mu_Y(t_1)][Y(t_2)-\mu_Y(t_2)]\}\\
&=E\{[X(t_1)+\varphi(t_1)-\mu_X(t_1)-\varphi(t_1)]\\
&\quad\times[X(t_2)+\varphi(t_2)-\mu_X(t_2)-\varphi(t_2)]\}\\
&=E\{[X(t_1)-\mu_X(t_1)][X(t_2)-\mu_X(t_2)]\}\\
&=C_X(t_1,t_2).
\end{aligned}$$

6. 给定一随机过程 $\{X(t),t\in T\}$ 和常数 a，试以 $X(t)$ 的自相关函数表示出

随机过程 $Y(t)=X(t+a)-X(t),t\in T$ 的自相关函数.

解　设 $X(t)$ 的自相关函数为 $R_X(t_1,t_2),t_1,t_2\in T$，按定义，$Y(t)$ 的自相关函数为

$$\begin{aligned}R_Y(t_1,t_2)&=E[Y(t_1)Y(t_2)]\\&=E\{[X(t_1+a)-X(t_1)][X(t_2+a)-X(t_2)]\}\\&=E[X(t_1+a)X(t_2+a)]-E[X(t_1)X(t_2+a)]\\&\quad-E[X(t_1+a)X(t_2)]+E[X(t_1)X(t_2)]\\&=R_X(t_1+a,t_2+a)-R_X(t_1,t_2+a)-R_X(t_1+a,t_2)+R_X(t_1,t_2).\end{aligned}$$

7. 设 $Z(t)=X+Yt,t\in(-\infty,\infty)$，若已知二维随机变量 (X,Y) 的协方差矩阵为

$$\begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix},$$

试求随机过程 $Z(t)$ 的协方差函数.

解　根据第四章 §4 协方差矩阵的定义及题设知，

$$E[(X-\mu_X)^2]=\sigma_1^2,\quad E[(X-\mu_X)(Y-\mu_Y)]=\rho\sigma_1\sigma_2,\quad E[(Y-\mu_Y)^2]=\sigma_2^2,$$

又　$\mu_Z(t)=E(X+Yt)=E(X)+tE(Y)=\mu_X+t\mu_Y$，

$$Z(t)-\mu_Z(t)=(X+tY)-(\mu_X+t\mu_Y)=(X-\mu_X)+t(Y-\mu_Y),$$

所以

$$\begin{aligned}C_Z(t_1,t_2)&=E\{[Z(t_1)-\mu_Z(t_1)][Z(t_2)-\mu_Z(t_2)]\}\\&=E\{[(X-\mu_X)+t_1(Y-\mu_Y)][(X-\mu_X)+t_2(Y-\mu_Y)]\}\\&=\sigma_1^2+(t_1+t_2)\rho\sigma_1\sigma_2+t_1t_2\sigma_2^2.\end{aligned}$$

8. 设 $X(t)=At+B,t\in(-\infty,\infty)$，式中 A,B 是相互独立，且都服从正态分布 $N(0,\sigma^2)$ 的随机变量，试证明 $X(t)$ 是一正态过程，并求出它的相关函数（协方差函数）.

解　由题设，A,B 是相互独立的正态变量，所以 (A,B) 是二维正态变量，对于任意一组实数 $t_1,t_2,\cdots,t_n\in T$，

$$X(t_i)=At_i+B,\quad i=1,2,\cdots,n$$

都是 A,B 的线性组合，于是根据教材第四章 §4 中 n 维正态变量的性质 3°知 $(X(t_1),X(t_2),\cdots,X(t_n))$ 是 n 维正态变量，再由 n,t_i 的任意性，得知 $X(t)$ 是正态过程. 而

$$\begin{aligned}R_X(t_1,t_2)&=E[X(t_1)X(t_2)]=E[(At_1+B)(At_2+B)]\\&=t_1t_2E(A^2)+(t_1+t_2)E(AB)+E(B^2),\quad t_1,t_2\in T.\end{aligned}$$

因 $A\sim N(0,\sigma^2),B\sim N(0,\sigma^2)$，且 A,B 相互独立，即有 $E(A)=E(B)=0$，

$E(A^2)=E(B^2)=\sigma^2, E(AB)=E(A)E(B)=0$，故

$$R_X(t_1,t_2)=(t_1t_2+1)\sigma^2,$$

又因 $\mu_X(t)=E(At+B)=tE(A)+E(B)=0$，故

$$C_X(t_1,t_2)=R_X(t_1,t_2).$$

9. 设随机过程 $X(t)$ 与 $Y(t)$，$t\in T$ 不相关，试用它们的均值函数与协方差函数来表示随机过程

$$Z(t)=a(t)X(t)+b(t)Y(t)+c(t),\quad t\in T$$

的均值函数和自协方差函数，其中 $a(t),b(t),c(t)$ 是普通的函数.

解
$$\begin{aligned}\mu_Z(t)&=E[Z(t)]\\&=a(t)E[X(t)]+b(t)E[Y(t)]+c(t)\\&=a(t)\mu_X(t)+b(t)\mu_Y(t)+c(t),\quad t\in T.\end{aligned}$$

从而

$$Z(t)-\mu_Z(t)=a(t)[X(t)-\mu_X(t)]+b(t)[Y(t)-\mu_Y(t)].$$

注意到 $X(t)$ 与 $Y(t)$ 不相关，于是它们的互协方差函数为零，即

$$\begin{aligned}C_{XY}(t_1,t_2)&=\mathrm{Cov}[X(t_1),Y(t_2)]\\&=E\{[X(t_1)-\mu_X(t_1)][Y(t_2)-\mu_Y(t_2)]\}=0,\quad t_1,t_2\in T,\end{aligned}$$

所以 $Z(t)$ 的自协方差函数

$$\begin{aligned}C_Z(t_1,t_2)&=\mathrm{Cov}[Z(t_1),Z(t_2)]\\&=E\{[Z(t_1)-\mu_Z(t_1)][Z(t_2)-\mu_Z(t_2)]\}\\&=a(t_1)a(t_2)E\{[X(t_1)-\mu_X(t_1)][X(t_2)-\mu_X(t_2)]\}\\&\quad+b(t_1)b(t_2)E\{[Y(t_1)-\mu_Y(t_1)][Y(t_2)-\mu_Y(t_2)]\}\\&=a(t_1)a(t_2)C_X(t_1,t_2)+b(t_1)b(t_2)C_Y(t_1,t_2),\quad t_1,t_2\in T.\end{aligned}$$

10. 设 $X(t)$ 和 $Y(t)$，$t>0$ 是两个相互独立的，分别具有强度 λ 和 μ 的泊松过程，试证

$$S(t)=X(t)+Y(t)$$

是具有强度 $\lambda+\mu$ 的泊松过程.

证 本题用另一种形式的定义来证明较为方便，需要证(i) $S(t)$ 是独立增量过程，(ii)对于任意的 $t>t_0\geqslant 0$，$S(t)-S(t_0)\sim\pi((\lambda+\mu)(t-t_0))$，(iii) $S(0)=0$. 下面按(iii)，(i)，(ii)的次序论证. 已知条件是

(a) $X(t),Y(t)$ $(t\geqslant 0)$ 分别是强度为 λ,μ 的泊松过程；

(b) 过程 $X(t)$ $(t\geqslant 0)$ 与 $Y(t)$ $(t\geqslant 0)$ 相互独立.

(iii) 因 $X(t),Y(t)$ $(t\geqslant 0)$ 都是泊松过程，所以 $X(0)=0,Y(0)=0$，今

$S(t)=X(t)+Y(t)$，从而得

$$S(0)=X(0)+Y(0)=0.$$

(i) 在 $t\geqslant 0$ 上任取点 $0\leqslant t_0<t_1<t_2<\cdots<t_n$，作 $S(t)$ 的增量

$$S(t_i)-S(t_{i-1})=X(t_i)+Y(t_i)-[X(t_{i-1})+Y(t_{i-1})]$$

$$\xlongequal{\text{记成}}U_i+V_i,$$

此处，$U_i=X(t_i)-X(t_{i-1})$，$V_i=Y(t_i)-Y(t_{i-1})$，$i=1,2,\cdots,n$.

对于任意实数 $x_1,x_2,\cdots,x_n,y_1,y_2,\cdots,y_n$ 有

$$P\{U_1\leqslant x_1,U_2\leqslant x_2,\cdots,U_n\leqslant x_n,V_1\leqslant y_1,V_2\leqslant y_2,\cdots,V_n\leqslant y_n\}$$

$$=P\{U_1\leqslant x_1,U_2\leqslant x_2,\cdots,U_n\leqslant x_n\}P\{V_1\leqslant y_1,V_2\leqslant y_2,\cdots,V_n\leqslant y_n\}$$

(由 $X(t)$ 与 $Y(t)$ 相互独立)

$$=P\{U_1\leqslant x_1\}P\{U_2\leqslant x_2\}\cdots P\{U_n\leqslant x_n\}P\{V_1\leqslant y_1\}P\{V_2\leqslant y_2\}\cdots P\{V_n\leqslant y_n\}$$

(由 $X(t),Y(t)$ 是独立增量过程)，

即知 $U_1,U_2,\cdots,U_n,V_1,V_2,\cdots,V_n$ 相互独立，从而 $U_1+V_1,U_2+V_2,\cdots,U_n+V_n$ 相互独立. 因此 $S(t)$ 是独立增量过程.

(ii) 由条件(a)知，对于任意 $t>t_0\geqslant 0$，

$$X(t)-X(t_0)\sim\pi(\lambda(t-t_0)),\quad Y(t)-Y(t_0)\sim\pi(\mu(t-t_0)).$$

由 $X(t),Y(t)$ 的独立性知 $X(t)-X(t_0)$ 与 $Y(t)-Y(t_0)$ 相互独立，从而由第三章习题 34 知

$$\begin{aligned}S(t)-S(t_0)&=X(t)+Y(t)-[X(t_0)+Y(t_0)]\\&=X(t)-X(t_0)+Y(t)-Y(t_0)\sim\pi((\lambda+\mu)(t-t_0)).\end{aligned}$$

11. 设 $\{W(t),t\geqslant 0\}$ 是以 σ^2 为参数的维纳过程，求下列过程的协方差函数：

(1) $W(t)+At$，A 为常数.

(2) $W(t)+Xt$，X 为与 $\{W(t),t\geqslant 0\}$ 相互独立的标准正态随机变量.

(3) $aW\left(\dfrac{t}{a^2}\right)$，$a$ 为正常数.

解　因 $W(t)$ 是维纳过程，故有

$\mu_W(t)=E[W(t)]=0$，$C_W(s,t)=E[W(s)W(t)]=\sigma^2\min\{s,t\}$，　$s,t\geqslant 0$.

(1) 记 $Z_1(t)=W(t)+At$，则有

$$\mu_{Z_1}(t)=E[W(t)]+E(At)=At,$$

故

$$\begin{aligned}C_{Z_1}(s,t)&=E\{[Z_1(s)-\mu_{Z_1}(s)][Z_1(t)-\mu_{Z_1}(t)]\}\\&=E[W(s)W(t)]=\sigma^2\min\{s,t\},\quad s,t\geqslant 0.\end{aligned}$$

(2) 记 $Z_2(t)=W(t)+Xt$，由题设 $W(t)$ 与 X 独立，$X\sim N(0,1)$，知

$$E(X)=0,\quad E(X^2)=1,\quad E[W(t)X]=E[W(t)]E(X)=0,$$
$$\mu_{Z_2}(t)=E[W(t)]+tE(X)=0,$$

故

$$\begin{aligned}C_{Z_2}(s,t)&=E\{[W(s)+Xs][W(t)+Xt]\}\\&=E[W(s)W(t)]+E(stX^2)\\&=\sigma^2\min\{s,t\}+st,\quad s,t\geqslant 0.\end{aligned}$$

(3) 记 $Z_3(t)=aW\left(\frac{t}{a^2}\right)$，则有 $\mu_{Z_3}(t)=aE\left[W\left(\frac{t}{a^2}\right)\right]=0$，故

$$\begin{aligned}C_{Z_3}(s,t)&=a^2E\left[W\left(\frac{s}{a^2}\right)W\left(\frac{t}{a^2}\right)\right]\\&=a^2\sigma^2\min\left\{\frac{s}{a^2},\frac{t}{a^2}\right\}=\sigma^2\min\{s,t\},\quad s,t\geqslant 0.\end{aligned}$$

第十三章　马尔可夫链

1. 从数 $1,2,\cdots,N$ 中任取一数，记为 X_1；再从 $1,2,\cdots,X_1$ 中任取一数，记为 X_2；如此继续，从 $1,2,\cdots,X_{n-1}$ 中任取一数，记为 X_n. 说明$\{X_n,n\geqslant 1\}$构成一时齐马尔可夫链，并写出它的状态空间和一步转移概率矩阵.

解　随机序列$\{X_n,n\geqslant 1\}$的状态空间 $I=\{1,2,\cdots,N\}$. X_n 在 $1,2,\cdots,X_{n-1}$ 中均匀取值. 对于任意整数 $1\leqslant a_{n-1}\leqslant\cdots\leqslant a_2\leqslant a_1\leqslant N$，有

$$P\left\{X_n=a_n\,\middle|\,X_{n-1}=a_{n-1},X_{n-2}=a_{n-2},\cdots,X_2=a_2,X_1=a_1\right\}$$

$$=\begin{cases}\dfrac{1}{a_{n-1}}, & a_n=1,2,\cdots,a_{n-1},\\ 0, & a_n\text{ 为其他值}\end{cases}$$

$$=P\left\{X_n=a_n\,\middle|\,X_{n-1}=a_{n-1}\right\},$$

故$\{X_n,n\geqslant 1\}$具有无后效性，即它是一个马尔可夫链.

按题意一步转移概率

$$p_{ij}=P\{X_{m+1}=j\,|\,X_m=i\}=\begin{cases}\dfrac{1}{i}, & 1\leqslant j\leqslant i,\\ 0, & j>i,\end{cases}\quad i=1,2,\cdots,N.$$

它们都只与 i,j 有关而与起始时刻 m 无关，因此$\{X_n,n\geqslant 1\}$是时齐马尔可夫链，且它的一步转移概率矩阵为

$$\boldsymbol{P}=\begin{bmatrix}1 & & & & & \\ \frac{1}{2} & \frac{1}{2} & & & \mathbf{0} & \\ \vdots & \vdots & \ddots & & & \\ \frac{1}{i} & \frac{1}{i} & \cdots & \frac{1}{i} & & \\ \vdots & \vdots & & \vdots & \ddots & \\ \frac{1}{N} & \frac{1}{N} & \cdots & \frac{1}{N} & \cdots & \frac{1}{N}\end{bmatrix}.$$

注：n 步转移是由相继的 n 次一步转移而完成的，所以只要一步转移概率与起始时刻 m 无关(那么，n 步转移概率也一定与 m 无关)，马尔可夫链就必定是时齐的.

2. 设 $X_0=1,X_1,X_2,\cdots,X_n,\cdots$是相互独立且都以概率$p$ $(0<p<1)$取值 1，以概率 $q=1-p$ 取值 0 的随机变量序列，令 $S_n=\sum\limits_{k=0}^{n}X_k$，证明$\{S_n,n\geqslant 0\}$构成一马尔可夫链，并写出它的状态空间和一步转移概率矩阵.

解　$S_n=\sum\limits_{k=0}^{n}X_k$，于是 $S_n=S_{n-1}+X_n$，由此可知，在 S_{n-1}给定的条件下，S_n的取值仅依赖于 X_n，此时 S_n 的取值与$(S_1,S_2,\cdots,S_{n-2})$的取值相互独立，即对于任意整数$1\leqslant i_1\leqslant i_2\leqslant\cdots\leqslant i_n$，有

$$P\{S_n=i_n|S_{n-1}=i_{n-1},S_{n-2}=i_{n-2},\cdots,S_1=i_1\}$$
$$=P\{S_{n-1}+X_n=i_n|S_{n-1}=i_{n-1},S_{n-2}=i_{n-2},\cdots,S_1=i_1\}$$
$$=P\{X_n=i_n-i_{n-1}|S_{n-1}=i_{n-1},S_{n-2}=i_{n-2},\cdots,S_1=i_1\}$$
$$=P\{X_n=i_n-i_{n-1}|S_{n-1}=i_{n-1}\}$$

（由于 X_n 与 $X_{n-2},X_{n-3},\cdots,X_1$ 相互独立，故 X_n 与$(S_{n-2},S_{n-3},\cdots,S_1)$相互独立）

$$=P\{X_n+S_{n-1}=(i_n-i_{n-1})+i_{n-1}|S_{n-1}=i_{n-1}\}$$
$$=P\{S_n=i_n|S_{n-1}=i_{n-1}\},$$

所以$\{S_n,n\geqslant 0\}$ 是一个马尔可夫链. 状态空间为 $I=\{1,2,\cdots\}$，又由 $X_1,X_2,\cdots$的独立性知一步转移概率为

$$\begin{aligned}p_{ij}&=P\{S_{m+1}=j|S_m=i\}=P\{S_m+X_{m+1}=j|S_m=i\}\\&=P\{X_{m+1}=j-i|S_m=i\}=P\{X_{m+1}=j-i\}\\&=\begin{cases}p, & j=i+1,\\ q, & j=i, \qquad i=1,2,\cdots.\\ 0, & \text{其他},\end{cases}\end{aligned}$$

因 p_{ij}只与i,j有关而与起始时刻m 无关，因此，它是时齐马尔可夫链. 一步转移概率矩阵为

$$\boldsymbol{P}=\begin{bmatrix}q & p & 0 & \cdots & & \\ 0 & q & p & 0 & \cdots & \\ 0 & 0 & q & p & 0 & \cdots\\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{bmatrix}.$$

3.（传染模型）考虑某种传染病在 N 个人中传染，假设

（1）在每个单位时间内此 N 个人中恰有两人互相接触，且一切成对的接触是等可能的.

（2）当健康人与患者接触时，被传染上的概率是 α.

（3）患者康复的概率为 0，健康人如不与患者接触，得病的概率也是 0.

以 X_n 表示第 n 个单位时间内的患者人数.试说明这种传染过程,即 $\{X_n,n\geqslant 0\}$ 是一马尔可夫链,并写出它的状态空间和一步转移概率矩阵.

解　$\{X_n,n\geqslant 0\}$ 的状态空间 $I=\{0,1,2,\cdots,N\}$,X_n 的取值仅与 X_{n-1} 的取值以及第 n 个单位时间内的人群的成对接触情况有关,所以 $\{X_n,n\geqslant 0\}$ 是一个马尔可夫链.

由假设(3),一旦患者人数(状态)为 0 或 N,则其人数不会再改变,用相应的转移概率可表示为

$$\begin{aligned}
p_{00}&=P\{X_{m+1}=0\mid X_m=0\}=1,\\
p_{0j}&=P\{X_{m+1}=j\mid X_m=0\}=0,\quad j\neq 0,\\
p_{NN}&=P\{X_{m+1}=N\mid X_m=N\}=1,\\
p_{Nj}&=P\{X_{m+1}=j\mid X_m=N\}=0,\quad j\neq N.
\end{aligned}$$

对其他任意一个状态 $i=1,2,\cdots,N-1$,转移概率 $P\{X_{m+1}=i+1\mid X_m=i\}$ 表示在时刻 m 恰有 i 个人患病的条件下,在第 $m+1$ 个单位时间,随机接触的两个人$\left(共\binom{N}{2}种方式\right)$恰为一人是健康人,一人是患者$\left(共\binom{i}{1}\binom{N-i}{1}种方式\right)$,且健康人被传染而患病的概率为

$$\begin{aligned}
P\{X_{m+1}=i+1\mid X_m=i\}&=\frac{\binom{i}{1}\binom{N-i}{1}}{\binom{N}{2}}\cdot\alpha\\
&=\frac{2\alpha i(N-i)}{N(N-1)},
\end{aligned}$$

将上式右边记为 α_i,于是,转移概率

$$\begin{aligned}
p_{i,i+1}&=P\{X_{m+1}=i+1\mid X_m=i\}=\alpha_i,\\
p_{ii}&=P\{X_{m+1}=i\mid X_m=i\}=1-\alpha_i,\\
p_{ij}&=P\{X_{m+1}=j\mid X_m=i\}=0,\quad j\neq i,i+1.
\end{aligned}$$

由于上述所有转移概率都仅与 i,j 有关而与起始时刻 m 无关,所以传染模型 $\{X_n,n\geqslant 0\}$ 是一个时齐马尔可夫链,且它的一步转移概率矩阵为

$$\boldsymbol{P}=\begin{bmatrix}
1 & 0 & 0 & 0 & \cdots & 0 & 0\\
0 & 1-\alpha_1 & \alpha_1 & 0 & \cdots & 0 & 0\\
0 & 0 & 1-\alpha_2 & \alpha_2 & \cdots & 0 & 0\\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\
0 & 0 & 0 & 0 & \cdots & 1-\alpha_{N-1} & \alpha_{N-1}\\
0 & 0 & 0 & 0 & \cdots & 0 & 1
\end{bmatrix}.$$

4. 设马尔可夫链$\{X_n,n\geqslant 0\}$的状态空间为$I=\{1,2,3\}$，初始分布为$p_1(0)=\frac{1}{4},p_2(0)=\frac{1}{2},p_3(0)=\frac{1}{4}$，一步转移概率矩阵为

$$\boldsymbol{P}=\begin{bmatrix}\frac{1}{4} & \frac{3}{4} & 0\\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}\\ 0 & \frac{1}{4} & \frac{3}{4}\end{bmatrix}.$$

(1) 计算$P\{X_0=1,X_1=2,X_2=2\}$.

(2) 证明$P\{X_1=2,X_2=2\mid X_0=1\}=p_{12}p_{22}$.

(3) 计算$p_{12}(2)=P\{X_2=2\mid X_0=1\}$.

(4) 计算$p_2(2)=P\{X_2=2\}$.

解 先计算二步转移概率矩阵

$$\boldsymbol{P}(2)=\boldsymbol{P}^2=\begin{bmatrix}\frac{5}{16} & \frac{7}{16} & \frac{4}{16}\\ \frac{7}{36} & \frac{16}{36} & \frac{13}{36}\\ \frac{4}{48} & \frac{13}{48} & \frac{31}{48}\end{bmatrix}.$$

(1) 因$p_1(0)=P\{X_0=1\}$，即有

$$\begin{aligned}&P\{X_0=1,X_1=2,X_2=2\}\\&=P\{X_0=1\}P\{X_1=2\mid X_0=1\}P\{X_2=2\mid X_0=1,X_1=2\}\\&=P\{X_0=1\}P\{X_1=2\mid X_0=1\}P\{X_2=2\mid X_1=2\}\\&=p_1(0)p_{12}p_{22}=\frac{1}{4}\times\frac{3}{4}\times\frac{1}{3}=\frac{1}{16}.\end{aligned}$$

(2) $$\begin{aligned}&P\{X_1=2,X_2=2\mid X_0=1\}\\&=P\{X_0=1,X_1=2,X_2=2\}/P\{X_0=1\}\\&\xlongequal{\text{由(1)}}p_{12}p_{22}.\end{aligned}$$

(3) 由 K－C 方程，

$$\begin{aligned}p_{12}(2)&=p_{11}p_{12}+p_{12}p_{22}+p_{13}p_{32}\\&=\frac{1}{4}\times\frac{3}{4}+\frac{3}{4}\times\frac{1}{3}+0\times\frac{1}{4}=\frac{7}{16}.\end{aligned}$$

$p_{12}(2)$也可以直接从二步转移概率矩阵获得.

(4) $p_2(2)=P\{X_2=2\}$

$$
\begin{aligned}
&=P\{X_2=2\mid X_0=1\}P\{X_0=1\}\\
&\quad+P\{X_2=2\mid X_0=2\}P\{X_0=2\}\\
&\quad+P\{X_2=2\mid X_0=3\}P\{X_0=3\}\\
&=p_{12}(2)p_1(0)+p_{22}(2)p_2(0)+p_{32}(2)p_3(0)\\
&=\frac{7}{16}\times\frac{1}{4}+\frac{16}{36}\times\frac{1}{2}+\frac{13}{48}\times\frac{1}{4}=\frac{115}{288}=0.399\ 3.
\end{aligned}
$$

5. 说明如何得到§1例4中的转移概率 p_{23} 和 p_{34}.

解　转移概率 p_{23} 和 p_{34} 都代表事件 A:系统中有顾客接受服务但还没有离开和 B:有一个新的顾客进入系统同时发生. 由于 A 的概率为 $1-p$,B 的概率为 q,可见

$$p_{23}=p_{34}=q(1-p).$$

6. 用§1例4中转移概率 $p_{ij}(i\neq j)$ 导出转移概率 p_{ii}.

解　如果系统当前处于状态 i,则下一步只有两种可能:仍然处于状态 i,其概率为 p_{ii} 或转移到其他状态 $j\neq i$,其概率为 $\sum\limits_{i\neq j}p_{ij}$. 所以 $p_{ii}+\sum\limits_{i\neq j}p_{ij}=1$. 这样我们就得到

$$p_{ii}=1-\sum_{i\neq j}p_{ij}.$$

7. 证明公式(3.1).

证法(i)　为求矩阵 $\boldsymbol{P}(n)=\boldsymbol{P}^n$,先作矩阵 $\boldsymbol{P}$ 的相似变换,令

$$\boldsymbol{P}=\begin{bmatrix}1-a & a\\ b & 1-b\end{bmatrix}\quad(0<a,b<1),$$

令 $\det(\lambda\boldsymbol{I}-\boldsymbol{P})=0$,得

$$[\lambda-(1-a)][\lambda-(1-b)]-ab=0,$$

解得特征根 $\lambda=1$ 或 $1-a-b$.

当 $\lambda=1$ 时,由 $(\boldsymbol{I}-\boldsymbol{P})\begin{bmatrix}x_1\\x_2\end{bmatrix}=\boldsymbol{0}$,即 $\begin{bmatrix}a & -a\\ -b & b\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\boldsymbol{0}$,得一特征向量 $\boldsymbol{\varepsilon}_1=\begin{bmatrix}1\\1\end{bmatrix}$.

当 $\lambda=1-a-b$ 时,由 $[(1-a-b)\boldsymbol{I}-\boldsymbol{P}]\begin{bmatrix}x_1\\x_2\end{bmatrix}=\boldsymbol{0}$,即 $\begin{bmatrix}-b & -a\\ -b & -a\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\boldsymbol{0}$,得一特征向量 $\boldsymbol{\varepsilon}_2=\begin{bmatrix}-a\\b\end{bmatrix}$,于是

$$\boldsymbol{H}=\begin{bmatrix}1 & -a\\ 1 & b\end{bmatrix},\quad \boldsymbol{H}^{-1}=\frac{1}{a+b}\begin{bmatrix}b & a\\ -1 & 1\end{bmatrix},$$

因而

$$\boldsymbol{P}=\boldsymbol{H}\begin{bmatrix}1 & 0\\ 0 & 1-a-b\end{bmatrix}\boldsymbol{H}^{-1},$$

从而

$$\begin{aligned}
\boldsymbol{P}^n &= \boldsymbol{H}\begin{bmatrix}1 & 0\\ 0 & 1-a-b\end{bmatrix}^n \boldsymbol{H}^{-1}=\boldsymbol{H}\begin{bmatrix}1 & 0\\ 0 & (1-a-b)^n\end{bmatrix}\boldsymbol{H}^{-1}\\
&=\boldsymbol{H}\left(\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}+\begin{bmatrix}0 & 0\\ 0 & (1-a-b)^n\end{bmatrix}\right)\boldsymbol{H}^{-1}\\
&=\begin{bmatrix}1 & -a\\ 1 & b\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\cdot\frac{1}{a+b}\begin{bmatrix}b & a\\ -1 & 1\end{bmatrix}\\
&\quad+\begin{bmatrix}1 & -a\\ 1 & b\end{bmatrix}\begin{bmatrix}0 & 0\\ 0 & (1-a-b)^n\end{bmatrix}\frac{1}{a+b}\begin{bmatrix}b & a\\ -1 & 1\end{bmatrix}\\
&=\frac{1}{a+b}\begin{bmatrix}b & a\\ b & a\end{bmatrix}+\frac{(1-a-b)^n}{a+b}\begin{bmatrix}a & -a\\ -b & b\end{bmatrix}.
\end{aligned}$$

证法(ii) 因已有结果，故可用数学归纳法证明之，$n=1$ 时，

$$\begin{aligned}
\boldsymbol{P}(1) &= \frac{1}{a+b}\begin{bmatrix}b & a\\ b & a\end{bmatrix}+\frac{1-a-b}{a+b}\begin{bmatrix}a & -a\\ -b & b\end{bmatrix}\\
&=\begin{bmatrix}1-a & a\\ b & 1-b\end{bmatrix}=\boldsymbol{P},
\end{aligned}$$

即当 $n=1$ 时，教材第十三章§3 公式(3.1)成立. 今假设 $n-1$ 时公式(3.1)成立，现计算

$$\begin{aligned}
\boldsymbol{P}(n) &= \boldsymbol{P}(n-1)\boldsymbol{P}\\
&=\left(\frac{1}{a+b}\begin{bmatrix}b & a\\ b & a\end{bmatrix}+\frac{(1-a-b)^{n-1}}{a+b}\begin{bmatrix}a & -a\\ -b & b\end{bmatrix}\right)\begin{bmatrix}1-a & a\\ b & 1-b\end{bmatrix}\\
&=\frac{1}{a+b}\begin{bmatrix}b & a\\ b & a\end{bmatrix}\begin{bmatrix}1-a & a\\ b & 1-b\end{bmatrix}\\
&\quad+\frac{(1-a-b)^{n-1}}{a+b}\begin{bmatrix}a & -a\\ -b & b\end{bmatrix}\begin{bmatrix}1-a & a\\ b & 1-b\end{bmatrix}\\
&=\frac{1}{a+b}\begin{bmatrix}b & a\\ b & a\end{bmatrix}+\frac{(1-a-b)^n}{a+b}\begin{bmatrix}a & -a\\ -b & b\end{bmatrix},
\end{aligned}$$

即公式(3.1)当 n 时也成立. 由数学归纳法，公式(3.1)对一切正整数都成立.

8. 设任意相继的两天中，雨天转晴天的概率为 $\dfrac{1}{3}$，晴天转雨天的概率为 $\dfrac{1}{2}$，任一天晴或雨互为逆事件. 以 0 表示晴天状态，以 1 表示雨天状态，X_n 表示第 n 天的状态(0 或 1). 试写出马尔可夫链 $\{X_n, n\geqslant 1\}$ 的一步转移概率矩阵. 又若已知 5 月 1 日为晴天，问 5 月 3 日为晴天，5 月 5 日为雨天的概率各等于多少?

解　以 0 表示晴天,1 表示雨天,已知 $p_{01}=\frac{1}{2}$,$p_{10}=\frac{1}{3}$,且已知任一天晴或雨是互逆事件,故 $p_{00}=\frac{1}{2}$,$p_{11}=\frac{2}{3}$,于是,若以 X_n 表示第 n 天的状态,则$\{X_n,n\geqslant 1\}$的一步转移概率矩阵为

$$\boldsymbol{P}=\begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{3} & \frac{2}{3}\end{bmatrix},$$

又已知 5 月 1 日为晴天,则 5 月 3 日为晴天的概率为 $p_{00}(2)$,而 5 月 5 日为雨天的概率为 $p_{01}(4)$,今求得

$$\boldsymbol{P}^2=\begin{bmatrix}\frac{5}{12} & \frac{7}{12}\\ \frac{7}{18} & \frac{11}{18}\end{bmatrix},\qquad \boldsymbol{P}^4=\begin{bmatrix}* & \frac{5}{12}\cdot\frac{7}{12}+\frac{7}{12}\cdot\frac{11}{18}\\ * & *\end{bmatrix},$$

故
$$p_{00}(2)=\frac{5}{12}=0.416\,7,$$

$$p_{01}(4)=\frac{5}{12}\times\frac{7}{12}+\frac{7}{12}\times\frac{11}{18}=\frac{259}{432}=0.599\,5.$$

9. 在一计算系统中,每一循环具有误差的概率取决于先前一个循环是否有误差.以 0 表示误差状态,以 1 表示无误差状态.设状态的一步转移概率矩阵为

$$\boldsymbol{P}=\begin{bmatrix}0.75 & 0.25\\ 0.5 & 0.5\end{bmatrix}.$$

试证明相应的时齐马尔可夫链具有遍历性,并求其极限分布.

(1) 用定义解.

(2) 利用遍历性定理解.

解　已知时齐马尔可夫链的一步转移概率矩阵为

$$\boldsymbol{P}=\begin{bmatrix}0.75 & 0.25\\ 0.5 & 0.5\end{bmatrix}.$$

(1) 在此应用教材第十三章 §3 公式(3.1),现在 $a=0.25$,$b=0.5$,即有

$$\boldsymbol{P}(n)=\boldsymbol{P}^n$$
$$=\frac{1}{0.75}\begin{bmatrix}0.5 & 0.25\\ 0.5 & 0.25\end{bmatrix}+\frac{(1-0.25-0.5)^n}{0.75}\begin{bmatrix}0.25 & -0.25\\ -0.5 & 0.5\end{bmatrix},\quad n=1,2,\cdots.$$

在上式中令 $n\to\infty$,得到

$$\lim_{n\to\infty}\boldsymbol{P}(n)=\frac{1}{0.75}\begin{bmatrix}0.5 & 0.25\\0.5 & 0.25\end{bmatrix}=\begin{bmatrix}\frac{2}{3} & \frac{1}{3}\\ \frac{2}{3} & \frac{1}{3}\end{bmatrix}.$$

由定义知此时齐马尔可夫链具有遍历性,其极限分布为 $\boldsymbol{\pi}=\left(\frac{2}{3},\frac{1}{3}\right)$.

(2) 因存在 $m=1$,使对于所有 $i,j,p_{ij}(1)=p_{ij}>0$ (即 $\boldsymbol{P}(1)$无零元),按定理知相应的时齐马尔可夫链是遍历的,其极限分布 $\boldsymbol{\pi}=(\pi_1,\pi_2)$满足方程组

$$\boldsymbol{\pi}=\boldsymbol{\pi P},$$

即

$$(\pi_1,\pi_2)=(\pi_1,\pi_2)\begin{bmatrix}0.75 & 0.25\\0.5 & 0.5\end{bmatrix}$$

$$=(0.75\pi_1+0.5\pi_2,0.25\pi_1+0.5\pi_2),$$

且有 $\pi_1+\pi_2=1$,由此得方程组

$$\pi_1=0.75\pi_1+0.5\pi_2,\quad \pi_2=0.25\pi_1+0.5\pi_2,\quad \pi_1+\pi_2=1.$$

解此方程组得 $\pi_1=\frac{2}{3},\pi_2=\frac{1}{3}$,即得极限分布为

$$\boldsymbol{\pi}=\left(\frac{2}{3},\frac{1}{3}\right).$$

10. 设时齐马尔可夫链的一步转移概率矩阵为

$$\boldsymbol{P}=\begin{bmatrix}q & p & 0\\q & 0 & p\\0 & q & p\end{bmatrix},\quad q=1-p,p\in(0,1).$$

试证明此链具有遍历性,并求其极限分布.

解 已知时齐马尔可夫链的一步转移概率矩阵为

$$\boldsymbol{P}=\begin{bmatrix}q & p & 0\\q & 0 & p\\0 & q & p\end{bmatrix},\quad q=1-p,0<p<1,$$

即有

$$\boldsymbol{P}^2=\begin{bmatrix}q & pq & p^2\\q^2 & 2pq & p^2\\q^2 & pq & p\end{bmatrix},$$

$\boldsymbol{P}(2)=\boldsymbol{P}^2$(即 $m=2$ 时)无零元,按教材第十三章§3 定理知此链具有遍历性,且其极限分布$\boldsymbol{\pi}=(\pi_1,\pi_2,\pi_3)$由方程组

$$\boldsymbol{\pi}=\boldsymbol{\pi P},\quad \pi_1+\pi_2+\pi_3=1$$

确定，亦即满足方程组

$$\begin{cases}\pi_1=q\pi_1+q\pi_2,\\ \pi_2=p\pi_1+q\pi_3,\\ \pi_3=p\pi_2+p\pi_3,\\ \pi_1+\pi_2+\pi_3=1.\end{cases}$$

方程组可改写成

$$p\pi_1=q\pi_2,\quad q\pi_3=p\pi_2,\quad \pi_1+\pi_2+\pi_3=1.$$

由此得极限分布

$$\boldsymbol{\pi}=\left(\frac{q^2}{1-pq},\frac{pq}{1-pq},\frac{p^2}{1-pq}\right).$$

若经适当恒等变形，π_j 也可写成

$$\pi_j=\frac{1-\dfrac{p}{q}}{1-\left(\dfrac{p}{q}\right)^3}\cdot\left(\frac{p}{q}\right)^{j-1},\quad j=1,2,3.$$

11. 设时齐马尔可夫链的一步转移概率矩阵为

$$\boldsymbol{P}=\begin{bmatrix}\frac{1}{2} & \frac{1}{2} & 0\\ \frac{1}{2} & \frac{1}{2} & 0\\ 0 & 0 & 1\end{bmatrix},$$

试证明此链不具有遍历性.

证　这一马氏链的一步转移概率矩阵可写成对角块矩阵：

$$\boldsymbol{P}=\begin{bmatrix}\frac{1}{2} & \frac{1}{2} & 0\\ \frac{1}{2} & \frac{1}{2} & 0\\ 0 & 0 & 1\end{bmatrix}=\begin{bmatrix}\boldsymbol{M} & \\ & 1\end{bmatrix},$$

其中 $\boldsymbol{M}=\begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2}\end{bmatrix}$. 故对于任意正整数 n，有

$$\boldsymbol{P}^n=\begin{bmatrix}\boldsymbol{M}^n & \\ & 1\end{bmatrix},$$

得到 $p_{13}(n)=p_{23}(n)=0$，$p_{33}(n)=1$，因此对于固定的 $j=3$，有：$\lim\limits_{n\to\infty}p_{1j}(n)=\lim\limits_{n\to\infty}p_{2j}(n)=0$，而$\lim\limits_{n\to\infty}p_{3j}(n)=1$，这表示 $p_{ij}(n)$的极限与 i 有关，故此链不是遍历的.

或者这样做，经计算知 $\boldsymbol{P}(2)=\boldsymbol{P}^2=\boldsymbol{P}$，即可知 $\boldsymbol{P}(n)=\boldsymbol{P}$. 由此，$\lim\limits_{n\to\infty}p_{13}(n)=\lim\limits_{n\to\infty}p_{23}(n)=0$，$\lim\limits_{n\to\infty}p_{33}(n)=1$，即对于固定的 $j=3$，极限$\lim\limits_{n\to\infty}p_{ij}(n)$与 i 有关，根据定义知此链不是遍历的.

第十四章 平稳随机过程

1. 设有随机过程 $X(t)=A\cos(\omega t+\Theta),t\in(-\infty,\infty)$，其中 A 是服从瑞利分布的随机变量，其概率密度为

$$f(a)=\begin{cases}\dfrac{a}{\sigma^2}e^{-a^2/(2\sigma^2)}, & a>0,\\ 0, & a\leqslant 0,\end{cases}$$

Θ 是在 $(0,2\pi)$ 上服从均匀分布且与 A 相互独立的随机变量，ω 是一常数，问 $X(t)$ 是不是平稳过程？

解 $X(t)X(t+\tau)=A^2\cos(\omega t+\Theta)\cos[\omega(t+\tau)+\Theta]$.

$$E(A^2)=\int_0^{\infty}\frac{x^3}{\sigma^2}e^{-x^2/(2\sigma^2)}dx \xlongequal{\text{令 } x^2/(2\sigma^2)=u} \int_0^{\infty}2\sigma^2 ue^{-u}du$$
$$=2\sigma^2\Gamma(2)=2\sigma^2.$$

$$E[\cos(\omega t+\Theta)]=\frac{1}{2\pi}\int_0^{2\pi}\cos(\omega t+\theta)d\theta=0.$$

$$E\{\cos(\omega t+\Theta)\cos[\omega(t+\tau)+\Theta]\}$$
$$=\frac{1}{2\pi}\int_0^{2\pi}\cos(\omega t+\theta)\cos[\omega(t+\tau)+\theta]d\theta$$
$$=\frac{1}{2\pi}\int_0^{2\pi}\frac{1}{2}[\cos\omega\tau+\cos(2\omega t+\omega\tau+2\theta)]d\theta=\frac{1}{2}\cos\omega\tau.$$

由题设 A,Θ 相互独立，所以 A 的函数和 Θ 的函数也相互独立（教材第三章 §4 定理），根据教材第四章 §1 数学期望的性质 4°，

$$\mu_X(t)=E[X(t)]=E(A)E[\cos(\omega t+\Theta)]=0,$$
$$R_X(t,t+\tau)=E[X(t)X(t+\tau)]$$
$$=E(A^2)E\{\cos(\omega t+\Theta)\cos[\omega(t+\tau)+\Theta]\}$$
$$=2\sigma^2\times\frac{1}{2}\cos\omega\tau=\sigma^2\cos\omega\tau.$$

这说明 $X(t)=A\cos(\omega t+\Theta)$ 的均值为常数，而自相关函数只与时间差 τ 有关，按定义 $X(t)$ 是平稳过程.

2. 设 $X(t)$ 和 $Y(t)$ 是相互独立的平稳过程. 试证以下随机过程也是平稳过程：

(1) $Z_1(t)=X(t)Y(t)$.

更多电子书，微信：taobao66taobao或淘宝店铺：菁华资料

(2) $Z_2(t)=X(t)+Y(t)$.

证 因 $X(t),Y(t)$都是平稳过程，所以对于任意 $t,t+\tau\in T$，有

$$E[X(t)]=\mu_X(\text{常数}),\quad E[Y(t)]=\mu_Y(\text{常数}),$$

$$E[X(t)X(t+\tau)]=R_X(\tau),\quad E[Y(t)Y(t+\tau)]=R_Y(\tau),$$

又因 $X(t),Y(t)$相互独立，从而

(1) $E[Z_1(t)]=E[X(t)]E[Y(t)]=\mu_X\mu_Y$(常数)，

$$\begin{aligned}R_{Z_1}(t,t+\tau)&=E[Z_1(t)Z_1(t+\tau)]\\&=E[X(t)Y(t)X(t+\tau)Y(t+\tau)]\\&=E[X(t)X(t+\tau)]E[Y(t)Y(t+\tau)]\\&=R_X(\tau)R_Y(\tau)\quad(\tau\text{ 的函数}).\end{aligned}$$

这说明 $Z_1(t)$的均值是常数，而自相关函数只与时间差 τ 有关，按定义 $Z_1(t)$是平稳过程.

(2) $E[Z_2(t)]=E[X(t)]+E[Y(t)]=\mu_X+\mu_Y$(常数)，

$$\begin{aligned}R_{Z_2}(t,t+\tau)&=E\{[X(t)+Y(t)][X(t+\tau)+Y(t+\tau)]\}\\&=E[X(t)X(t+\tau)]+E[Y(t)Y(t+\tau)]\\&\quad+E[X(t)Y(t+\tau)]+E[X(t+\tau)Y(t)]\\&=R_X(\tau)+R_Y(\tau)+2\mu_X\mu_Y\quad(\tau\text{ 的函数}),\end{aligned}$$

故 $Z_2(t)$是平稳过程.

3. 设$\{X(t),t\in(-\infty,\infty)\}$是平稳过程，$R_X(\tau)$是其自相关函数，$a$ 是常数，试问随机过程

$$Y(t)=X(t+a)-X(t)$$

是不是平稳过程？为什么？

解 由题设 $X(t)$的均值为 μ_X(常数)，自相关函数为 $R_X(\tau)$，由此，

$$\begin{aligned}E[Y(t)]&=E[X(t+a)-X(t)]\\&=E[X(t+a)]-E[X(t)]=\mu_X-\mu_X=0,\end{aligned}$$

$$\begin{aligned}R_Y(t,t+\tau)&=E[Y(t)Y(t+\tau)]\\&=E\{[X(t+a)-X(t)][X(t+\tau+a)-X(t+\tau)]\}\\&=E[X(t+a)X(t+\tau+a)]-E[X(t+a)X(t+\tau)]\\&\quad-E[X(t)X(t+\tau+a)]+E[X(t)X(t+\tau)]\\&=2R_X(\tau)-R_X(\tau-a)-R_X(\tau+a).\end{aligned}$$

$Y(t)$的均值为常数，又因 a 是常数，自相关函数 $R_Y(t,t+\tau)$仅与时间差 τ 有关，故 $Y(t)$是平稳过程.

4. 设$\{N(t),t\geqslant0\}$是强度为 λ 的泊松过程，定义随机过程 $Y(t)=N(t+L)-$

$N(t)$，其中常数 $L>0$. 试求 $Y(t)$ 的均值函数和自相关函数，并问 $Y(t)$ 是否是平稳过程?

解　因 $N(t)$ 是强度为 λ 的泊松过程，所以 $N(t+L)-N(t)\sim\pi(\lambda L)$，从而

$$E[Y(t)]=E[N(t+L)-N(t)]=\lambda L,$$

$$\begin{aligned}R_Y(s,t)&=E\{[N(s+L)-N(s)][N(t+L)-N(t)]\}\\&=E[N(s+L)N(t+L)]+E[N(s)N(t)]\\&\quad-E[N(s+L)N(t)]-E[N(s)N(t+L)].\end{aligned}$$

因 $E[N(s)N(t)]=\lambda^2 st+\lambda\min\{s,t\},s,t\geqslant 0$，故

$$\begin{aligned}R_Y(s,t)=\lambda^2L^2+\lambda[&\min\{s+L,t+L\}\\&+\min\{s,t\}-\min\{s+L,t\}-\min\{s,t+L\}].\end{aligned}\qquad (*_1)$$

记 $t-s=\tau$，由 $R_Y(s,t)=R_Y(t,s)$，即知 $R_Y(s,t)$ 关于 s,t 是对称的，故不妨暂时认为 $s\leqslant t$. 于是上式右边等于

$$\begin{aligned}&\lambda^2L^2+\lambda(s+L)+\lambda s-\lambda\min\{s+L,t\}-\lambda s\\&=\lambda^2L^2+\lambda(s+L)-\lambda\min\{s+L,t\}\\&=\begin{cases}\lambda^2L^2+\lambda(L-t+s), & 若\ t\leqslant s+L,\\ \lambda^2L^2, & 若\ t>s+L\end{cases}\\&=\begin{cases}\lambda^2L^2+\lambda(L-\tau), & 若\ 0\leqslant\tau\leqslant L,\\ \lambda^2L^2, & 若\ \tau>L,\end{cases}\end{aligned}$$

由 s,t 的对称性，即知

$$R_Y(s,t)=\begin{cases}\lambda^2L^2+\lambda(L-|\tau|), & 若\ |\tau|\leqslant L,\\ \lambda^2L^2, & 若\ |\tau|>L,\end{cases}$$

且知 $Y(t)$ 是平稳过程.

上述 $(*_1)$ 式也可以由先求出 $Y(t)$ 的协方差函数而获得，事实上

$$\begin{aligned}C_Y(t,t+\tau)&=\mathrm{Cov}[Y(t),Y(t+\tau)]\\&=\mathrm{Cov}[N(t+L)-N(t),N(t+\tau+L)-N(t+\tau)]\\&=\mathrm{Cov}[N(t+L),N(t+\tau+L)]-\mathrm{Cov}[N(t),N(t+\tau+L)]\\&\quad-\mathrm{Cov}[N(t+L),N(t+\tau)]+\mathrm{Cov}[N(t),N(t+\tau)]\\&=\lambda\min\{t+L,t+\tau+L\}-\lambda\min\{t,t+\tau+L\}\\&\quad-\lambda\min\{t+L,t+\tau\}+\lambda\min\{t,\ t+\tau\}.\end{aligned}$$

于是得到自相关函数

$$R_Y(t,t+\tau)=C_Y(t,t+\tau)+\mu_Y^2=C_Y(t,t+\tau)+\lambda^2L^2,$$

此即为 $(*_1)$ 式.

5. 设平稳过程 $\{X(t),t\in(-\infty,\infty)\}$ 的自相关函数为 $R_X(\tau)=\mathrm{e}^{-a|\tau|}(1+$

$a|\tau|)$，其中常数 $a>0$，而 $E[X(t)]=0$. 试问 $X(t)$ 的均值是否具有各态历经性？为什么？

解法(i) 记 $J=\int_0^{2T}\left(1-\frac{\tau}{2T}\right)[R_X(\tau)-\mu_X^2]\mathrm{d}\tau$，由题设，得

$$\begin{aligned}J&=\int_0^{2T}\left(1-\frac{\tau}{2T}\right)[\mathrm{e}^{-a|\tau|}(1+a|\tau|)]\mathrm{d}\tau\\&=\int_0^{2T}\left(1-\frac{\tau}{2T}\right)[\mathrm{e}^{-a\tau}(1+a\tau)]\mathrm{d}\tau\\&\xlongequal{\text{令 } a\tau=t}\int_0^{2aT}\left(1-\frac{t}{2aT}\right)(1+t)\mathrm{e}^{-t}\,\frac{1}{a}\mathrm{d}t\\&=\frac{1}{a}\int_0^{2aT}\left[1+\left(1-\frac{1}{2aT}\right)t-\frac{t^2}{2aT}\right]\mathrm{e}^{-t}\mathrm{d}t\\&=\frac{\mathrm{e}^{-t}}{a}\left[\frac{t^2}{2aT}-\left(1-\frac{3}{2aT}\right)t-2+\frac{3}{2aT}\right]\Bigg|_0^{2aT}\\&=\frac{1}{a}\left[\left(1+\frac{3}{2aT}\right)\mathrm{e}^{-2aT}+2-\frac{3}{2aT}\right].\end{aligned}$$

故

$$\lim_{T\to\infty}\frac{1}{T}J=\lim_{T\to\infty}\frac{1}{a}\left[\left(\frac{1}{T}+\frac{3}{2aT^2}\right)\mathrm{e}^{-2aT}+\frac{2}{T}-\frac{3}{2aT^2}\right]=0.$$

根据教材第十四章 §2 定理 1 知随机过程 $X(t)$ 的均值具有各态历经性.

解法(ii) 由于自相关函数当 $\tau\to\infty$ 时的极限

$$\lim_{\tau\to\infty}R_X(\tau)=\lim_{\tau\to\infty}\mathrm{e}^{-a|\tau|}(1+a|\tau|)=0=\mu_X^2,$$

根据教材第十四章 §2 定理 1 的推论知，随机过程 $X(t)$ 的均值具有各态历经性.

6. 第 1 题中的随机过程 $X(t)=A\cos(\omega t+\Theta)$ 是否是各态历经过程？为什么？

解 由第 1 题知随机过程 $X(t)=A\cos(\omega t+\Theta)$ 的均值 $\mu_X=0$，自相关函数 $R_X(\tau)=\sigma^2\cos\omega\tau$，现在来计算 $X(t)$ 的相应的时间平均：

$$\begin{aligned}\langle X(t)\rangle&=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}A\cos(\omega t+\Theta)\mathrm{d}t\\&=\lim_{T\to\infty}\frac{A[\sin(\omega T+\Theta)-\sin(-\omega T+\Theta)]}{2\omega T}=0,\end{aligned}$$

$$\begin{aligned}\langle X(t)X(t+\tau)\rangle&=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}A^2\cos(\omega t+\Theta)\cos(\omega(t+\tau)+\Theta)\mathrm{d}t\\&=\lim_{T\to\infty}\frac{A^2}{4T}\int_{-T}^{T}[\cos\omega\tau+\cos(2(\omega t+\Theta)+\omega\tau)]\mathrm{d}t\\&=\frac{A^2}{2}\cos\omega\tau.\end{aligned}$$

由$\langle X(t)\rangle=\mu_X$，知 $X(t)$的均值具有各态历经性，但 $P\{A^2=2\sigma^2\}\neq 1$，因而$\langle X(t)X(t+\tau)\rangle$不可能以概率1等于$R_X(\tau)$，即自相关函数不具有各态历经性，所以$X(t)$不是各态历经过程.

7. (1) 设 $C_X(\tau)$是平稳过程 $X(t)$的协方差函数. 试证若 $C_X(\tau)$绝对可积，即

$$\int_{-\infty}^{\infty}|C_X(\tau)|\mathrm{d}\tau<\infty,$$

则 $X(t)$的均值具有各态历经性.

(2) 证明本章§1例1中的随机相位周期过程 $X(t)=s(t+\Theta)$是各态历经过程.

证 (1) 由于 $R_X(\tau)-\mu_X^2=C_X(\tau)$，因此

$$\begin{aligned}0&\leqslant\left|\frac{1}{T}\int_0^{2T}\left(1-\frac{\tau}{2T}\right)[R_X(\tau)-\mu_X^2]\mathrm{d}\tau\right|\\&=\left|\frac{1}{T}\int_0^{2T}\left(1-\frac{\tau}{2T}\right)C_X(\tau)\mathrm{d}\tau\right|\\&\leqslant\frac{1}{T}\int_0^{2T}|C_X(\tau)|\mathrm{d}\tau\leqslant\frac{1}{T}\int_{-\infty}^{\infty}|C_X(\tau)|\mathrm{d}\tau.\end{aligned}$$

按题设 $C_X(\tau)$的积分绝对收敛，故当 $T\to\infty$时上式右方趋于零，从而

$$\lim_{T\to\infty}\frac{1}{T}\int_0^{2T}\left(1-\frac{\tau}{2T}\right)[R_X(\tau)-\mu_X^2]\mathrm{d}\tau=0.$$

所以，由教材第十四章§2定理1，$X(t)$的均值是各态历经的.

本题给出了均值具有各态历经性的一个充分条件，即若

$$\int_{-\infty}^{\infty}|C_X(\tau)|\mathrm{d}\tau<\infty,$$

则 $X(t)$的均值是各态历经的.

(2)
$$\langle X(t)\rangle=\lim_{G\to\infty}\frac{1}{2G}\int_{-G}^{G}X(t)\mathrm{d}t=\lim_{G\to\infty}\frac{1}{2G}\int_{-G}^{G}s(t+\Theta)\mathrm{d}t.$$

因对于任意正数 G，总存在正整数 n，使 $nT\leqslant G<(n+1)T$，故

$$\frac{1}{2G}\int_{-G}^{G}s(t+\Theta)\mathrm{d}t=\frac{1}{2nT}\frac{nT}{G}\left[\int_{-nT}^{nT}s(t+\Theta)\mathrm{d}t+\int_{-G}^{-nT}s(t+\Theta)\mathrm{d}t+\int_{nT}^{G}s(t+\Theta)\mathrm{d}t\right].$$

注意到

$$\lim_{n\to\infty}\frac{nT}{G}=1,\quad\lim_{n\to\infty}\frac{1}{2nT}\int_{-G}^{-nT}s(t+\Theta)\mathrm{d}t=0,$$

$$\lim_{n\to\infty}\frac{1}{2nT}\int_{nT}^{G}s(t+\Theta)\mathrm{d}t=0,$$

得到

$$\langle X(t)\rangle=\lim_{n\to\infty}\frac{1}{2nT}\int_{-nT}^{nT}s(t+\Theta)\mathrm{d}t=\lim_{n\to\infty}\frac{1}{2nT}\sum_{i=-n}^{n-1}\int_{iT}^{(i+1)T}s(t+\Theta)\mathrm{d}t$$

$$=\frac{1}{T}\int_0^T s(t+\Theta)\mathrm{d}t=\frac{1}{T}\int_{\Theta}^{\Theta+T}s(u)\mathrm{d}u=\frac{1}{T}\int_0^T s(u)\mathrm{d}u\ (\text{由周期性}).$$

在第十四章§1例2中已知 $E[X(t)]=\frac{1}{T}\int_0^T s(u)\mathrm{d}u$，故有

$$\langle X(t)\rangle=E[X(t)].$$

同样可证

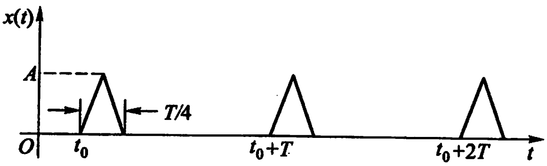
$$\langle X(t)X(t+\tau)\rangle=E[X(t)X(t+\tau)]=R_X(\tau),$$

因此 $X(t)$的均值和自相关函数都具有各态历经性，从而得 $X(t)$是各态历经过程.

8. 设 $X(t)$是随机相位周期过程，题14.8图表示它的一个样本函数 $x(t)$，其中周期 T 和波幅 A 都是常数，而相位 t_0 是在$(0,T)$上服从均匀分布的随机变量.

(1) 求 μ_X，Ψ_X^2.

(2) 求$\langle X(t)\rangle$和$\langle X^2(t)\rangle$.



题 14.8 图

解 依照随机相位周期过程的定义，在此

$$X(t)=s(t+t_0),$$

其中

$$s(t)=\begin{cases}\frac{8A}{T}, & 0\leqslant t\leqslant\frac{T}{8},\\ -\frac{8A}{T}\left(t-\frac{T}{4}\right), & \frac{T}{8}<t\leqslant\frac{T}{4},\\ 0, & \frac{T}{4}<t\leqslant\tau\end{cases}$$

且 $s(t+T)=s(t)$.

(1) 依照教材第十四章§1例2的结果

$$\mu_X=\frac{1}{T}\int_0^T s(t)\mathrm{d}t=\frac{A}{8},$$

$$\Psi_X^2=\frac{1}{T}\int_0^T s^2(t)\mathrm{d}t=\frac{A^2}{12}.$$

(2) 参见本章习题第 7 题(2)，利用各态历经性有

$$\langle X(t)\rangle = E[X(t)] = \mu_X = \frac{A}{8},$$

$$\langle X^2(t)\rangle = E[X^2(t)] = \Psi_X^2 = \frac{A^2}{12}.$$

9. 设平稳过程 $X(t)$ 的自相关函数为 $R_X(\tau)$，证明

$$P\{|X(t+\tau)-X(t)|\geqslant a\}\leqslant 2[R_X(0)-R_X(\tau)]/a^2,\quad a>0.$$

证　命题似与切比雪夫不等式相关联. 因 $X(t)$ 是平稳过程，故有

$$E[X(t)] = \mu_X \text{（常数）},$$

$$E[X(t)X(t+\tau)] = R_X(\tau),$$

$$E[X^2(t)] = E[X^2(t+\tau)] = R_X(0).$$

记 $Y(t)=X(t+\tau)-X(t)$，即有

$$E[Y(t)] = E[X(t+\tau)]-E[X(t)] = \mu_X-\mu_X = 0,$$

$$\begin{aligned} D[Y(t)] &= E[Y^2(t)] = E\{[X(t+\tau)-X(t)]^2\} \\ &= E[X^2(t+\tau)]-2E[X(t)X(t+\tau)]+E[X^2(t)] \\ &= 2[R_X(0)-R_X(\tau)]. \end{aligned}$$

对 $Y(t)=X(t+\tau)-X(t)$ 应用切比雪夫不等式，即有

$$P\{|X(t+\tau)-X(t)|\geqslant a\}\leqslant 2[R_X(0)-R_X(\tau)]/a^2,\quad a>0.$$

10. 设 $X(t)$ 为平稳过程，其自相关函数 $R_X(\tau)$ 是以 T_0 为周期的函数. 证明 $X(t)$ 是周期为 T_0 的平稳过程.

证　记 $Y(t)=X(t+T_0)-X(t)$，由于 $X(t)$ 是平稳过程，故 $Y(t)$ 也是平稳过程(见本章第 3 题)，且

$$E[Y(t)] = E[X(t+T_0)]-E[X(t)] = \mu_X-\mu_X = 0,$$

$$D[Y(t)] = E[Y^2(t)] = 2[R_X(0)-R_X(T_0)]$$

(运算过程见上题)，又，按题设 $R_X(\tau)$ 具有周期 T_0，故 $R_X(0)=R_X(T_0)$，即有 $D[Y(t)]=0$，由教材第四章§2 方差性质 4°知，对于任意 t，有 $P\{Y(t)=0\}=1$，或即

$$P\{X(t+T_0)=X(t)\}=1,$$

即在概率 1 的意义下，$X(t)$ 是以 T_0 为周期的平稳过程.

11. 设 $X(t)$ 是雷达的发射信号，遇目标后返回接收机的微弱信号是 $aX(t-\tau_1)$，$a\ll 1$，τ_1 是信号返回时间，由于接收到的信号总是伴有噪声的，记噪声为 $N(t)$，于是接收到的全信号为

$$Y(t)=aX(t-\tau_1)+N(t).$$

(1) 若 $X(t)$ 和 $N(t)$ 是平稳相关的，证明 $X(t)$ 和 $Y(t)$ 也平稳相关.

(2) 在(1)的条件下，假设 $N(t)$ 的均值为零且与 $X(t)$ 是相互独立的，求 $R_{XY}(\tau)$（这是利用互相关函数从全信号中检测小信号的相关接收法）.

解 (1)
$$\begin{aligned}R_{XY}(t,t+\tau)&=E[X(t)Y(t+\tau)]\\&=E\{X(t)[aX(t+\tau-\tau_1)+N(t+\tau)]\}\\&=aR_X(\tau-\tau_1)+R_{XN}(\tau),\end{aligned}$$
因而知 $R_{XY}(t,t+\tau)$ 与 t 无关，故 $X(t)$ 与 $Y(t)$ 平稳相关.

(2) 由题设 $\mu_N=0$，且 $X(t)$ 与 $N(t)$ 相互独立，即有
$$R_{XN}(\tau)=E[X(t)N(t+\tau)]=E[X(t)]E[N(t+\tau)]=\mu_X\mu_N=0.$$
由(1)得
$$R_{XY}(\tau)=aR_X(\tau-\tau_1).$$
因 $R_X(0)\geqslant R_X(\tau)$，即 $R_X(\tau)$ 在 $\tau=0$ 时取到最大值，故当调整 τ 使 $R_{XY}(\tau)$ 的测量值为最大时，就有 $\tau=\tau_1$，由此可推算得目标的距离.

12. 平稳过程 $\{X(t),t\in(-\infty,\infty)\}$ 的自相关函数为
$$R_X(\tau)=4\mathrm{e}^{-|\tau|}\cos\pi\tau+\cos 3\pi\tau,$$
求：(1) $X(t)$ 的均方值. (2) $X(t)$ 的谱密度.

解 (1) 由题设 $R_X(\tau)=4\mathrm{e}^{-|\tau|}\cos\pi\tau+\cos 3\pi\tau$，即得 $X(t)$ 的均方值为
$$\Psi_X^2=R_X(0)=4+1=5.$$

(2) 由教材第十四章 § 4 维纳-辛钦公式(4.15)，$X(t)$ 的谱密度可用傅里叶变换表示为
$$\begin{aligned}S_X(\omega)&=\mathscr{F}[R_X(\tau)]=\mathscr{F}[4\mathrm{e}^{-|\tau|}\cos\pi\tau+\cos 3\pi\tau]\\&=4\mathscr{F}[\mathrm{e}^{-|\tau|}\cos\pi\tau]+\mathscr{F}[\cos 3\pi\tau].\end{aligned}$$
由教材第十四章 § 4 例 1 及表 14—1 第 3、7 栏得到
$$\mathscr{F}[\mathrm{e}^{-|\tau|}\cos\pi\tau]=\frac{1}{1+(\omega-\pi)^2}+\frac{1}{1+(\omega+\pi)^2},$$
$$\mathscr{F}[\cos 3\pi\tau]=\pi[\delta(\omega+3\pi)+\delta(\omega-3\pi)].$$
所以
$$S_X(\omega)=4\left[\frac{1}{1+(\omega-\pi)^2}+\frac{1}{1+(\omega+\pi)^2}\right]+\pi[\delta(\omega+3\pi)+\delta(\omega-3\pi)].$$

13. 已知平稳过程 $X(t)$ 的谱密度为
$$S_X(\omega)=\frac{\omega^2}{\omega^4+3\omega^2+2},$$
求 $X(t)$ 的均方值.

解法(i) 将有理谱密度分解成两个部分分式之和
$$S_X(\omega)=\frac{\omega^2}{\omega^4+3\omega^2+2}=\frac{2}{\omega^2+2}-\frac{1}{\omega^2+1},$$

代入教材第十四章维纳-辛钦公式(4.16)，并令 $\tau=0$，得到

$$\Psi_X^2=R_X(0)=\frac{1}{2\pi}\int_{-\infty}^{\infty}S_X(\omega)\mathrm{d}\omega=\frac{1}{\pi}\int_0^{\infty}\left(\frac{2}{\omega^2+2}-\frac{1}{\omega^2+1}\right)\mathrm{d}\omega$$

$$=\frac{1}{\pi}\left(\sqrt{2}\arctan\frac{\omega}{\sqrt{2}}\Big|_0^{\infty}-\arctan\omega\Big|_0^{\infty}\right)=\frac{1}{2}(\sqrt{2}-1).$$

解法(ii)　主要是借此题展示求与有理谱密度相对应的自相关函数的方法，注意到教材第十四章公式(4.16)就是对 $S_X(\omega)$ 求逆傅里叶变换，即有

$$R_X(\tau)=\mathscr{F}^{-1}[S_X(\omega)]=\mathscr{F}^{-1}\left[\frac{2}{\omega^2+2}-\frac{1}{\omega^2+1}\right]$$

$$=\mathscr{F}^{-1}\left[\frac{2}{\omega^2+2}\right]-\mathscr{F}^{-1}\left[\frac{1}{\omega^2+1}\right].$$

查教材表 14—1 第 1 栏得

$$\mathscr{F}^{-1}\left[\frac{2}{\omega^2+2}\right]=\frac{1}{\sqrt{2}}\mathscr{F}^{-1}\left[\frac{2\sqrt{2}}{\omega^2+(\sqrt{2})^2}\right]=\frac{1}{\sqrt{2}}\mathrm{e}^{-\sqrt{2}|\tau|},$$

$$\mathscr{F}^{-1}\left[\frac{1}{\omega^2+1}\right]=\frac{1}{2}\mathscr{F}^{-1}\left[\frac{2}{\omega^2+1}\right]=\frac{1}{2}\mathrm{e}^{-|\tau|},$$

所以 $X(t)$ 的自相关函数为

$$R_X(\tau)=\frac{1}{\sqrt{2}}\mathrm{e}^{-\sqrt{2}|\tau|}-\frac{1}{2}\mathrm{e}^{-|\tau|}.$$

由此，$X(t)$ 的均方值为

$$\Psi_X^2=R_X(0)=\frac{1}{\sqrt{2}}-\frac{1}{2}=\frac{1}{2}(\sqrt{2}-1).$$

14. 已知平稳过程 $X(t)$ 的自相关函数为

$$R_X(\tau)=\begin{cases}1-\dfrac{|\tau|}{T}, & |\tau|\leqslant T,\\ 0, & |\tau|>T,\end{cases}$$

求谱密度 $S_X(\omega)$.

解　引用教材第十四章公式(4.18)，得

$$S_X(\omega)=2\int_0^{\infty}R_X(\tau)\cos\omega\tau\mathrm{d}\tau=2\int_0^{T}\left(1-\frac{\tau}{T}\right)\cos\omega\tau\mathrm{d}\tau$$

$$=2\left(1-\frac{\tau}{T}\right)\frac{\sin\omega\tau}{\omega}\Big|_0^{T}+\frac{2}{\omega T}\int_0^{T}\sin\omega\tau\mathrm{d}\tau$$

$$=\frac{2}{\omega^2T}(1-\cos\omega T)=\frac{4}{\omega^2T}\sin^2\frac{\omega T}{2}.$$

15. 已知平稳过程 $X(t)$ 的谱密度为

$$S_X(\omega)=\begin{cases}8\delta(\omega)+20\left(1-\dfrac{|\omega|}{10}\right), & |\omega|<10,\\ 0, & |\omega|\geqslant 10,\end{cases}$$

求 $X(t)$的自相关函数.

解 引用公式(4.16)(教材第十四章 §4)得

$$R_X(\tau)=\frac{1}{2\pi}\int_{-10}^{10}\left[8\delta(\omega)+20\left(1-\frac{|\omega|}{10}\right)\right]e^{i\omega\tau}d\omega$$

$$=\frac{4}{\pi}\int_{-10}^{10}\delta(\omega)e^{i\omega\tau}d\omega+\frac{10}{\pi}\int_{-10}^{10}\left(1-\frac{|\omega|}{10}\right)e^{i\omega\tau}d\omega.$$

因当 $\omega\neq 0$ 时,$\delta(\omega)=0$,故

$$\frac{4}{\pi}\int_{-10}^{10}\delta(\omega)e^{i\omega\tau}d\omega=\frac{4}{\pi}\int_{-\infty}^{\infty}\delta(\omega)e^{i\omega\tau}d\omega=\frac{4}{\pi},$$

$$\frac{10}{\pi}\int_{-10}^{10}\left(1-\frac{|\omega|}{10}\right)e^{i\omega\tau}d\omega=\frac{10}{\pi}\int_{-10}^{10}\left(1-\frac{|\omega|}{10}\right)(\cos\omega\tau+i\sin\omega\tau)d\omega$$

$$=\frac{20}{\pi}\int_{0}^{10}\left(1-\frac{\omega}{10}\right)\cos\omega\tau d\omega$$

$$=\frac{20}{\pi}\left[\left(1-\frac{\omega}{10}\right)\frac{\sin\omega\tau}{\tau}\Bigg|_0^{10}+\frac{1}{10\tau}\int_0^{10}\sin\omega\tau d\omega\right]$$

$$=\frac{20}{\pi}\cdot\frac{1}{10\tau^2}(1-\cos 10\tau)=\frac{4}{\pi}\frac{\sin^2 5\tau}{\tau^2},$$

于是

$$R_X(\tau)=\frac{4}{\pi}\left(1+\frac{\sin^2 5\tau}{\tau^2}\right).$$

16. 记随机过程

$$Y(t)=X(t)\cos(\omega_0 t+\Theta),\quad t\in(-\infty,\infty),$$

其中 $X(t)$是平稳过程,Θ 为在区间$(0,2\pi)$上均匀分布的随机变量,ω_0 为常数,且 $X(t)$与 Θ 相互独立. 记 $X(t)$的自相关函数为 $R_X(\tau)$,功率谱密度为 $S_X(\omega)$. 试证:

(1) $Y(t)$是平稳过程,且它的自相关函数

$$R_Y(\tau)=\frac{1}{2}R_X(\tau)\cos\omega_0\tau.$$

(2) $Y(t)$的功率谱密度为

$$S_Y(\omega)=\frac{1}{4}[S_X(\omega-\omega_0)+S_X(\omega+\omega_0)].$$

证 (1) $\mu_Y=E[Y(t)]=E[X(t)\cos(\omega_0 t+\Theta)]$

$$=E[X(t)]E[\cos(\omega_0 t+\Theta)]$$

$$= \mu_X \cdot \frac{1}{2\pi}\int_0^{2\pi}\cos(\omega_0 t+\theta)\mathrm{d}\theta = \mu_X \cdot 0 = 0,$$

$$\begin{aligned} R_Y(t,t+\tau) &= E[Y(t)Y(t+\tau)] \\ &= E[X(t)\cos(\omega_0 t+\Theta)\cdot X(t+\tau)\cos(\omega_0(t+\tau)+\Theta)] \\ &= E[X(t)X(t+\tau)]E[\cos(\omega_0 t+\Theta)\cos(\omega_0 t+\omega_0\tau+\Theta)] \\ &= R_X(\tau)\cdot\frac{1}{2\pi}\int_0^{2\pi}\frac{1}{2}[\cos\omega_0\tau+\cos(2\omega_0 t+\omega_0\tau+2\theta)]\mathrm{d}\theta \\ &= R_X(\tau)\cdot\frac{1}{2}\cos\omega_0\tau\ (\text{是 } \tau \text{ 的函数}), \end{aligned}$$

故 $Y(t)$是平稳过程，并且它的自相关函数为

$$R_Y(\tau)=\frac{1}{2}R_X(\tau)\cos\omega_0\tau.$$

(2)
$$\begin{aligned} S_Y(\omega) &= \int_{-\infty}^{\infty}R_Y(\tau)\mathrm{e}^{-\mathrm{i}\omega\tau}\mathrm{d}\tau \\ &= \int_{-\infty}^{\infty}\frac{1}{2}R_X(\tau)\cos\omega_0\tau\cdot\mathrm{e}^{-\mathrm{i}\omega\tau}\mathrm{d}\tau \\ &= \frac{1}{2}\int_{-\infty}^{\infty}R_X(\tau)\frac{\mathrm{e}^{\mathrm{i}\omega_0\tau}+\mathrm{e}^{-\mathrm{i}\omega_0\tau}}{2}\mathrm{e}^{-\mathrm{i}\omega\tau}\mathrm{d}\tau \\ &= \frac{1}{4}\left[\int_{-\infty}^{\infty}R_X(\tau)\mathrm{e}^{-\mathrm{i}(\omega-\omega_0)\tau}\mathrm{d}\tau+\int_{-\infty}^{\infty}R_X(\tau)\mathrm{e}^{-\mathrm{i}(\omega+\omega_0)\tau}\mathrm{d}\tau\right] \\ &= \frac{1}{4}[S_X(\omega-\omega_0)+S_X(\omega+\omega_0)]. \end{aligned}$$

17. 设平稳过程 $X(t)$的谱密度为 $S_X(\omega)$，证明 $Y(t)=X(t)+X(t-T)$的谱密度是

$$S_Y(\omega)=2S_X(\omega)(1+\cos\omega T).$$

证
$$\begin{aligned} R_Y(t,t+\tau) &= E\{[X(t)+X(t-T)][X(t+\tau)+X(t+\tau-T)]\} \\ &= E[X(t)X(t+\tau)]+E[X(t-T)X(t-T+\tau)] \\ &\quad +E[X(t-T)X(t+\tau)]+E[X(t)X(t+\tau-T)] \\ &= 2R_X(\tau)+R_X(\tau+T)+R_X(\tau-T). \end{aligned}$$

即有
$$R_Y(\tau)=2R_X(\tau)+R_X(\tau+T)+R_X(\tau-T).$$

由题设 $\int_{-\infty}^{\infty}R_X(\tau)\mathrm{e}^{-\mathrm{i}\omega\tau}\mathrm{d}\tau=S_X(\omega)$，即有

$$\begin{aligned} S_Y(\omega) &= \int_{-\infty}^{\infty}R_Y(\tau)\mathrm{e}^{-\mathrm{i}\omega\tau}\mathrm{d}\tau \\ &= \int_{-\infty}^{\infty}[2R_X(\tau)+R_X(\tau+T)+R_X(\tau-T)]\mathrm{e}^{-\mathrm{i}\omega\tau}\mathrm{d}\tau \end{aligned}$$

$$=2S_X(\omega)+\mathrm{e}^{\mathrm{i}\omega T}\int_{-\infty}^{\infty}R_X(\tau+T)\mathrm{e}^{-\mathrm{i}\omega(\tau+T)}\mathrm{d}\tau+\mathrm{e}^{-\mathrm{i}\omega T}\int_{-\infty}^{\infty}R_X(\tau-T)\mathrm{e}^{-\mathrm{i}\omega(\tau-T)}\mathrm{d}\tau$$

$$=2S_X(\omega)+\mathrm{e}^{\mathrm{i}\omega T}S_X(\omega)+\mathrm{e}^{-\mathrm{i}\omega T}S_X(\omega)$$

$$=2S_X(\omega)\left(1+\frac{\mathrm{e}^{\mathrm{i}\omega T}+\mathrm{e}^{-\mathrm{i}\omega T}}{2}\right)$$

$$=2S_X(\omega)(1+\cos\omega T).$$

第十五章 时间序列分析

1. 用延迟算子表示下列模型：

(1) $X_t-0.5X_{t-1}=\varepsilon_t$.

(2) $X_t=\varepsilon_t-0.7\varepsilon_{t-1}-0.24\varepsilon_{t-2}$.

(3) $X_t-0.5X_{t-1}=\varepsilon_t-0.7\varepsilon_{t-1}-0.24\varepsilon_{t-2}$.

(4) $X_t-1.5X_{t-1}+0.5X_{t-2}=\varepsilon_t$.

(5) $X_t-X_{t-1}=\varepsilon_t-0.5\varepsilon_{t-1}$.

解 以上模型可用延迟算子 B 表示为

(1) $(1-0.5B)X_t=\varepsilon_t$.

(2) $X_t=(1-0.7B-0.24B^2)\varepsilon_t$.

(3) $(1-0.5B)X_t=(1-0.7B-0.24B^2)\varepsilon_t$.

(4) $(1-1.5B+0.5B^2)X_t=\varepsilon_t$.

(5) $(1-B)X_t=(1-0.5B)\varepsilon_t$.

2. 将上题中的模型(1)—(5)按 ARMA(p,q)分类.

解 上题中的模型分类如下：

(1) ARMA(1,0).

(2) ARMA(0,2).

(3) ARMA(1,2).

(4) ARMA(2,0).

(5) ARMA(1,1).

3. 证明自相关函数的性质：(1)$\rho_k=\rho_{-k}$和(2)$|\rho_k|\leqslant 1$.

证 (1) 由 X_t 的平稳性，

$$\gamma_k=E[(X_t-\mu)(X_{t+k}-\mu)]$$

与 t 无关. 在上式中令 $t=s-k$ 得到

$$\gamma_k=E[(X_{s-k}-\mu)(X_s-\mu)]=\gamma_{-k}.$$

将上式两端除以 γ_0 即得 $\rho_k=\rho_{-k}$.

(2) 若 $\gamma_k=0$ 则 $\rho_k=0$，显然满足不等式. 所以只需考虑 $\gamma_k\neq 0$. 由 X_t 的平稳性可知

$$E[(X_t-\mu)^2]=E[(X_{t+k}-\mu)^2]=\gamma_0.$$

于是对任意 $s\in\mathbf{R}$ 我们有

$$
\begin{aligned}
0 &\leqslant E\{[X_t-\mu-s(X_{t+k}-\mu)]^2\} \\
&=E[(X_t-\mu)^2]-2sE[(X_t-\mu)(X_{t+k}-\mu)]+s^2E[(X_{t+k}-\mu)^2] \\
&=\gamma_0-2s\gamma_k+s^2\gamma_0.
\end{aligned}
$$

令 $s=\gamma_0/\gamma_k=1/\rho_k$ 得到

$$0\leqslant-\gamma_0+\gamma_0/\rho_k^2.$$

可见 $\rho_k^2\leqslant 1$，也就是 $|\rho_k|\leqslant 1$.

4. 求 $X_t=\varepsilon_t-0.5\varepsilon_{t-1}-0.24\varepsilon_{t-2}$ 的自相关函数.

解 这是一个系数 $\theta_1=0.5,\theta_2=0.24$ 的 MA(2)模型. 由教材第十五章式(2.7)可知

$$
\begin{aligned}
&\rho_0=1, \\
&\rho_1=\frac{-0.5+0.5\times0.24}{1+0.5^2+0.24^2}=\frac{-0.38}{1+0.5^2+0.24^2}, \\
&\rho_k=0,\quad k>2.
\end{aligned}
$$

5. 将§2例中运用的 R 程序用于下列模型：

(1) $X_t-0.5X_{t-1}=\varepsilon_t$.

(2) $X_t=\varepsilon_t-0.7\varepsilon_{t-1}-0.24\varepsilon_{t-2}$.

解 下面的两组指令可分别用来模拟生成 $n=200$ 项对应于模型(1)和(2)的时间序列并作出它们的 ACF 和 PACF 的图示.

指令组(1)：

```
>sim.ar<-arima.sim(list(ar=c(0.5)),n=200)
>par(mfrow=c(2,1))
>acf(sim.ar)
>pacf(sim.ar)
```

指令组(2)：

```
>sim.ma<-arima.sim(list(ma=c(-0.7,-0.24)),n=200)
>par(mfrow=c(2,1))
>acf(sim.ma)
>pacf(sim.ma)
```

每次 R 模拟生成的序列是不一样的. 因此作出的 ACF 和 PACF 的图示也会有所不同.

6. 反复运行§3例2中的 arima.sim()和 eacf()至少 100 次. 函数正确识

别模型的频率有多大？与同学的结果作综合比较.这个频率稳定吗？

解　先用下面的指令调用 R 中的库 fBasics 和 TSA：

```
>library(fBasics)
>library(TSA)
```

再反复运行下面的指令：

```
>x<-arima.sim(list(order=c(2,0,1),ar=c(0.3,-0.7),ma=c(0.5)),n=200)
>eacf(x,ar.max=6,ma.max=6)
```

并对结果进行统计.

7. 讨论 MA(1)和 MA(6)模型对§3 例 3 中 10 年期国债利率的一阶差分时间序列的适用性.

解　直接使用§3 例 3 的方法及 R 指令.唯一需要改变的是在调用 arima()时分别用 order=c(0,0,1)和 order=c(0,0,6)代替 order=c(1,0,0).细节如下:先从数字课程网站下载数据文件 shuju.csv,存入子目录 C:/R-example(或任何其他子目录),并用指令

```
>setwd("C:/R-example")
```

将上述子目录设置为 R 的工作目录.然后用下面的指令读入数据并计算差分序列.

```
>bond10=read.csv(file="shuju.csv",head=TRUE,sep=",")
>l=length(bond10$Yield) % 计算序列长度
>x<-diff(bond10$Yield) % 计算利率差分序列
```

先讨论 MA(1)模型对此数据的适用性.用 arima()函数来估计参数如下.

```
>ma1=arima(x,order=c(0,0,1))
>ma1
Call:
arima(x=x,order=c(0,0,1))
Coefficients:
          ma1      intercept
          0.2507    -0.0174
s.e.      0.0967     0.0242
sigma^2 estimated as 0.04468:log likelihood=16.05,aic= -26.1
```

再用 Box-Ljung 方法来考核模型.

```
>Box.test(ma1$residuals,lag=12,type='Ljung',fitdf=1)
```

```
        Box - Ljung test
data:ma1 $ residuals
X - squared = 11.042,df = 11,p - value = 0.4398
```

调用这个函数时一般用 lag=$n/10$ 来给出自由度，并用 fitdf=$p+q$ 给出约束的数目．于是对于此例自由度就是 df=lag$-p-q=11$．$\chi^2(11)=11.042$，由χ^2分布表（附表 5）可以查出，该值小于以 0.05 的显著性水平拒绝假设 $H_0: \hat{\rho}_1=\hat{\rho}_2=\cdots=\hat{\rho}_{11}=0$ 的临界点 19.675．另外 p 值为 0.439 8，远大于拒绝假设 H_0 的显著性水平 0.05．因而按 0.05 的显著性水平，这个模型可以通过考核．

类似地讨论 MA(6)模型对此数据的适用性．用 arima()函数来估计参数如下．

```
>ma6 = arima(x,order = c(0,0,6))
>ma6
Call:
arima(x = x,order = c(0,0,6))
Coefficients:
         ma1      ma2      ma3      ma4      ma5      ma6    intercept
      0.1662  -0.0924   0.0515   0.0399  -0.0866  -0.1736    -0.0169
s.e.  0.0970   0.0971   0.1029   0.1097   0.0891   0.0974     0.0174
sigma^2 estimated as 0.04292:log likelihood = 18.35,aic = -20.69
```

再用 Box - Ljung 方法来考核模型．

```
>Box.test(ma6 $ residuals,lag = 12,type = 'Ljung',fitdf = 6)
  Box - Ljung test
data:ma6 $ residuals
X - squared = 7.5551,df = 6,p - value = 0.2725
```

p 值为 0.272 5，大于拒绝假设 H_0 的显著性水平 0.05．按 0.05 的显著性水平，MA(6)模型通过考核．

§3 例 3 和这里的讨论显示同一随机序列很可能有多个可用的模型．

第十六章　选做习题

1. 一打靶场备有 5 支某种型号的枪，其中 3 支已经校正，2 支未经校正. 某人使用已校正的枪击中目标的概率为 p_1，使用未经校正的枪击中目标的概率为 p_2. 他随机地取一支枪进行射击，已知他射击了 5 次，都未击中，求他使用的是已校正的枪的概率(设各次射击的结果相互独立).

解　以 M 表示事件“射击 5 次均未击中”，以 C 表示事件“取得的枪是已经校正的”，则 $P(C)=\frac{3}{5}$，$P(\overline{C})=\frac{2}{5}$，又，按题设 $P(M|C)=(1-p_1)^5$，$P(M|\overline{C})=(1-p_2)^5$，由贝叶斯公式

$$P(C|M)=\frac{P(MC)}{P(M)}=\frac{P(M|C)P(C)}{P(M|C)P(C)+P(M|\overline{C})P(\overline{C})}$$

$$=\frac{(1-p_1)^5\times\frac{3}{5}}{(1-p_1)^5\times\frac{3}{5}+(1-p_2)^5\times\frac{2}{5}}$$

$$=\frac{3(1-p_1)^5}{3(1-p_1)^5+2(1-p_2)^5}.$$

2. 某人共买了 11 个水果，其中有 3 个是二级品，8 个是一级品. 随机地将水果分给 A,B,C 三人，各人分别得到 4 个、6 个、1 个.

(1) 求 C 未拿到二级品的概率.

(2) 已知 C 未拿到二级品，求 A,B 均拿到二级品的概率.

(3) 求 A,B 均拿到二级品而 C 未拿到二级品的概率.

解　以 A,B,C 分别表示事件“A,B,C 取到二级品”，则 $\overline{A},\overline{B},\overline{C}$ 分别表示事件“A,B,C 未取到二级品”.

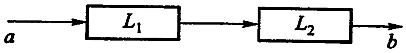
(1) $P(\overline{C})=\frac{8}{11}$.

(2) 就是需要求 $P(AB|\overline{C})$. 已知 C 未取到二级品，这时 A,B 将 7 个一级品和 3 个二级品全部分掉. 而 A,B 均取到二级品，只需 A 取到 1 个至 2 个二级品，其他的为一级品. 于是

$$P(AB\mid\overline{C})=\frac{\binom{3}{1}\binom{7}{3}}{\binom{10}{4}}+\frac{\binom{3}{2}\binom{7}{2}}{\binom{10}{4}}=\frac{4}{5}.$$

(3) $P(AB\overline{C})=P(AB\mid\overline{C})P(\overline{C})=\dfrac{32}{55}.$

3. 一系统 L 由两个只能传输字符 0 和 1 的独立工作的子系统 L_1 与 L_2 串联而成(如题 16.3 图)，每个子系统输入为 0 输出为 0 的概率为 p $(0<p<1)$；而输入为 1 输出为 1 的概率也是 p. 今在图中 a 端输入字符 1，求系统 L 的 b 端输出字符 0 的概率.



题 16.3 图

解　"系统 L 的输入为 1 输出为 0"这一事件(记为 $L(1\to 0)$)是两个互不相容事件之和，即

$$L(1\to 0)=L_1(1\to 1)L_2(1\to 0)\cup L_1(1\to 0)L_2(0\to 0).$$

这里的记号"$L_1(1\to 1)$"表示事件"子系统 L_1 的输入为 1 输出为 1". 其余 3 个记号的含义类似. 于是由子系统工作的独立性得

$$\begin{aligned}P\{L(1\to 0)\}&=P\{L_1(1\to 1)L_2(1\to 0)\}+P\{L_1(1\to 0)L_2(0\to 0)\}\\&=P\{L_1(1\to 1)\}P\{L_2(1\to 0)\}+P\{L_1(1\to 0)\}P\{L_2(0\to 0)\}\\&=p(1-p)+(1-p)p=2p(1-p).\end{aligned}$$

4. 甲、乙两人轮流掷一颗骰子，每轮掷一次，谁先掷得 6 点谁得胜，从甲开始掷，问甲、乙得胜的概率各为多少?

解　以 A_i 表示事件"第 i 次投掷时投掷者才得 6 点". 事件 A_i 发生，表示在前 $i-1$ 次甲或乙均未得 6 点，而在第 i 次投掷时甲或乙得 6 点. 因各次投掷相互独立，故有

$$P(A_i)=\left(\frac{5}{6}\right)^{i-1}\frac{1}{6}.$$

因甲为首掷，故甲掷奇数轮次，从而甲胜的概率为

$$\begin{aligned}P\{\text{甲胜}\}&=P\{A_1\cup A_3\cup A_5\cup\cdots\}\\&=P(A_1)+P(A_3)+P(A_5)+\cdots\\&\quad(\text{因 }A_1,A_2,\cdots\text{ 两两互不相容})\\&=\frac{1}{6}\left[1+\left(\frac{5}{6}\right)^2+\left(\frac{5}{6}\right)^4+\cdots\right]\end{aligned}$$

$$=\frac{1}{6}\frac{1}{1-\left(\frac{5}{6}\right)^2}=\frac{6}{11}.$$

同样，乙胜的概率为

$$\begin{aligned}P\{\text{乙胜}\}&=P\{A_2\cup A_4\cup A_6\cup\cdots\}\\&=P(A_2)+P(A_4)+P(A_6)+\cdots\\&=\frac{1}{6}\left[\frac{5}{6}+\left(\frac{5}{6}\right)^3+\left(\frac{5}{6}\right)^5+\cdots\right]=\frac{5}{11}.\end{aligned}$$

5. 将一颗骰子掷两次，考虑事件 A＝"第一次掷得点数 2 或 5"，B＝"两次点数之和至少为 7"，求 $P(A)$，$P(B)$，并问事件 A,B 是否相互独立？

解　将骰子掷一次共有 6 种等可能结果，故 $P(A)=\frac{2}{6}=\frac{1}{3}$. 设以 X_i 表示第 i 次掷出的骰子的点数，则

$$P(B)=P\{X_1+X_2\geqslant 7\}=1-P\{X_1+X_2\leqslant 6\}.$$

因将骰子掷两次共有 36 个样本点，其中 $X_1+X_2\leqslant 6$ 有 $X_1+X_2=2,3,4,5,6$ 共 5 种情况，这 5 种情况分别含有 1,2,3,4,5 个样本点，故

$$P(B)=1-\frac{1+2+3+4+5}{36}=1-\frac{5}{12}=\frac{7}{12}.$$

以 (X_1,X_2) 记两次投掷的结果，则 AB 共有(2,5)，(2,6)，(5,2)，(5,3)，(5,4)，(5,5)，(5,6)这 7 个样本点. 故 $P(AB)=\frac{7}{36}$.

今有

$$P(A)P(B)=\frac{1}{3}\times\frac{7}{12}=\frac{7}{36}=P(AB).$$

按定义 A,B 相互独立.

6. A,B 两人轮流射击，每次每人射击一枪，射击的次序为 $A,B,A,B,A,\cdots$，射击直至击中两枪为止. 设每人击中的概率均为 p，且各次击中与否相互独立. 求击中的两枪是由同一人射击的概率.

解　A 总是在奇数轮射击，B 在偶数轮射击. 先考虑 A 击中两枪的情况. 以 A_{2n+1} 表示事件"A 在第 $2n+1$ 轮($n=1,2,\cdots$)射击时又一次击中，射击在此时结束". A_{2n+1} 发生表示"前 $2n$ 轮中 A 共射击 n 枪而其中击中一枪，且 A 在第 $2n+1$ 轮时击中第二枪"(这一事件记为 C)，同时"B 在前 $2n$ 轮中共射击 n 枪但一枪未中"(这一事件记为 D)，因此

$$\begin{aligned}P(A_{2n+1})&=P(CD)=P(C)P(D)\\&=\left[\binom{n}{1}p(1-p)^{n-1}p\right](1-p)^n=np^2(1-p)^{2n-1}.\end{aligned}$$

注意到 $A_3, A_5, A_7, \cdots$ 两两互不相容，故由 A 击中了两枪而结束射击（这一事件仍记为 A）的概率为

$$P(A)=P\left(\bigcup_{n=1}^{\infty} A_{2n+1}\right)=\sum_{n=1}^{\infty}P(A_{2n+1})=\sum_{n=1}^{\infty}np^2(1-p)^{2n-1}$$

$$=p^2(1-p)\sum_{n=1}^{\infty}n[(1-p)^2]^{n-1}$$

$$=p^2(1-p)\frac{1}{[1-(1-p)^2]^2}=\frac{1-p}{(2-p)^2}.$$

$\left(\text{此处级数求和用到公式}\frac{1}{(1-x)^2}=\sum_{n=1}^{\infty}nx^{n-1},|x|<1.\right.$ 这一公式可由等比级数 $\left.\frac{1}{1-x}=\sum_{n=0}^{\infty}x^n,|x|<1 \text{ 两边求导而得到.}\right)$

若两枪均由 B 击中，以 $B_{2(n+1)}$ 表示事件“B 在第 $2(n+1)$ 轮（$n=1,2,\cdots$）射击时又一次击中，射击在此时结束”. $B_{2(n+1)}$ 发生表示在前 $2n+1$ 轮中 B 射击 n 枪其中击中一枪，且 B 在第 $2(n+1)$ 轮时击中第二枪，同时 A 在前 $2n+1$ 轮中共射击 $n+1$ 枪，但一枪未中. 注意到 $A_4, A_6, A_8, \cdots$ 两两互不相容，故 B 击中了两枪而结束射击（这一事件仍记为 B）的概率为

$$P(B)=P\left(\bigcup_{n=1}^{\infty} B_{2(n+1)}\right)=\sum_{n=1}^{\infty}\binom{n}{1}p(1-p)^{n-1}p(1-p)^{n+1}$$

$$=\sum_{n=1}^{\infty}np^2(1-p)^{2n}=p^2(1-p)^2\sum_{n=1}^{\infty}n[(1-p)^2]^{n-1}$$

$$=p^2(1-p)^2\frac{1}{[1-(1-p)^2]^2}=\frac{(1-p)^2}{(2-p)^2}.$$

因此由一人击中两枪的概率为

$$P(A\cup B)=P(A)+P(B)=\frac{1-p}{(2-p)^2}+\frac{(1-p)^2}{(2-p)^2}=\frac{1-p}{2-p}.$$

7. 有 3 个独立工作的元件 1、元件 2、元件 3，它们的可靠性分别为 p_1, p_2, p_3. 设由它们组成一个“3 个元件取 2 个元件的表决系统”，记为 $2/3[G]$. 这一系统的运行方式是当且仅当 3 个元件中至少有 2 个正常工作时这一系统正常工作. 求这一 $2/3[G]$ 系统的可靠性.

解 以 A_i 表示事件“第 i 个元件正常工作”，以 G 表示事件“$2/3[G]$ 系统正常工作”，则 G 可表示为下述两两互不相容的事件之和：

$$G=A_1A_2\overline{A}_3\cup A_1\overline{A}_2A_3\cup\overline{A}_1A_2A_3\cup A_1A_2A_3.$$

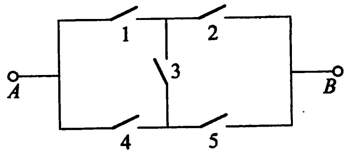
因 A_1, A_2, A_3 相互独立，故有

$$P(G)=P(A_1A_2\overline{A}_3)+P(A_1\overline{A}_2A_3)+P(\overline{A}_1A_2A_3)+P(A_1A_2A_3)$$
$$=P(A_1)P(A_2)P(\overline{A}_3)+P(A_1)P(\overline{A}_2)P(A_3)$$
$$+P(\overline{A}_1)P(A_2)P(A_3)+P(A_1)P(A_2)P(A_3)$$
$$=p_1p_2(1-p_3)+p_1(1-p_2)p_3+(1-p_1)p_2p_3+p_1p_2p_3.$$

8. 在如题 16.8 图所示的桥式结构的电路中，第 i 个继电器触点闭合的概率为 $p_i,i=1,2,3,4,5$. 各继电器工作相互独立.

(1) 以继电器触点 1 是否闭合为条件，求 A 到 B 之间为通路的概率.

(2) 已知 A 到 B 为通路的条件下，求继电器触点 3 闭合的概率.



题 16.8 图

解 以 F 表示事件"A 点到 B 点为通路"，以 C_i 表示事件"继电器触点 i 闭合"，$i=1,2,3,4,5$，各继电器工作相互独立.

(1) 得

$$P(F)=P(F|C_1)P(C_1)+P(F|\overline{C}_1)P(\overline{C}_1).$$

而 $$P(F|C_1)=P(C_2\cup C_3C_5\cup C_4C_5)$$
$$=P(C_2)+P(C_3C_5)+P(C_4C_5)-P(C_2C_3C_5)-P(C_2C_4C_5)$$
$$-P(C_3C_4C_5)+P(C_2C_3C_4C_5)$$
$$=p_2+p_3p_5+p_4p_5-p_2p_3p_5-p_2p_4p_5-p_3p_4p_5+p_2p_3p_4p_5,$$
$$P(F|\overline{C}_1)=P(C_4C_5\cup C_2C_3C_4)=p_4p_5+p_2p_3p_4-p_2p_3p_4p_5,$$

故 $P(F)=P(F|C_1)p_1+P(F|\overline{C}_1)(1-p_1)$，其中

$$P(F\mid C_1)=p_2+p_3p_5+p_4p_5-p_2p_3p_5-p_2p_4p_5-p_3p_4p_5+p_2p_3p_4p_5,$$
$$P(F\mid \overline{C}_1)=p_4p_5+p_2p_3p_4-p_2p_3p_4p_5.$$

(2) 令 $q_i=1-p_i$，则

$$P(C_3|F)=\frac{P(F|C_3)P(C_3)}{P(F)}=\frac{[1-P(\overline{C}_1\overline{C}_4\cup\overline{C}_2\overline{C}_5)]P(C_3)}{P(F)}$$
$$=\frac{(1-q_1q_4-q_2q_5+q_1q_2q_4q_5)p_3}{P(F)},$$

$P(F)$的表达式由(1)决定.

9. 进行非学历考试，规定考甲、乙两门课程，每门课程考试第一次未通过都只允许考第二次. 考生仅在课程甲通过后才能考课程乙，如两门课程都通过可获得一张资格证书. 设某考生通过课程甲的各次考试的概率为 p_1，通过课程乙的各次考试的概率为 p_2，设各次考试的结果相互独立. 又设考生参加考试直至获得资格证书或者不准予再考为止. 以 X 表示考生总共需考试的次数. 求 X 的分布律.

解 按题意知考生总共至少需考 2 次而最多只能考 4 次. 以 A_i 表示事件"课程甲在考第 i 次时通过"，$i=1,2$；以 B_i 表示事件"课程乙在考第 i 次时通过"，$i=1,2$.

事件$\{X=2\}$表示考生总共考 2 次，这一事件只在下列两种互不相容的情况下发生，一种是课程甲、乙都在第一次考试时通过，亦即 A_1B_1 发生(此时他得到证书)；另一种是课程甲在第一次、第二次考试均未通过，亦即 $\overline{A}_1\overline{A}_2$ 发生(此时他不准再考). 故

$$\{X=2\}=A_1B_1\cup\overline{A}_1\overline{A}_2,$$

同样

$$\{X=3\}=A_1\overline{B}_1B_2\cup\overline{A}_1A_2B_1\cup A_1\overline{B}_1\overline{B}_2,$$

$$\{X=4\}=\overline{A}_1A_2\overline{B}_1B_2\cup\overline{A}_1A_2\overline{B}_1\overline{B}_2.$$

得 X 的分布律为

$$\begin{aligned}P\{X=2\}&=P(A_1B_1\cup\overline{A}_1\overline{A}_2)=P(A_1B_1)+P(\overline{A}_1\overline{A}_2)\\&=P(A_1)P(B_1)+P(\overline{A}_1)P(\overline{A}_2)\\&=p_1p_2+(1-p_1)(1-p_1),\\P\{X=3\}&=P(A_1\overline{B}_1B_2\cup\overline{A}_1A_2B_1\cup A_1\overline{B}_1\overline{B}_2)\\&=P(A_1\overline{B}_1\cup\overline{A}_1A_2B_1)\\&=p_1(1-p_2)+(1-p_1)p_1p_2,\\P\{X=4\}&=P(\overline{A}_1A_2\overline{B}_1B_2\cup\overline{A}_1A_2\overline{B}_1\overline{B}_2)\\&=P(\overline{A}_1A_2\overline{B}_1)=(1-p_1)p_1(1-p_2).\end{aligned}$$

则考生总共需考试的次数 X 的分布律为

X	2	3	4
p_k	$p_1p_2+(1-p_1)^2$	$p_1(1-p_2)+(1-p_1)p_1p_2$	$(1-p_1)p_1(1-p_2)$

例如，若 $p_1=\frac{3}{4}$，$p_2=\frac{1}{2}$，则有分布律

X	2	3	4
p_k	$\frac{14}{32}$	$\frac{15}{32}$	$\frac{3}{32}$

10. (1) 5 只电池中有 2 只是次品，每次取一只测试，直到将 2 只次品都找到. 设第 2 只次品在第 X ($X=2,3,4,5$)次找到，求 X 的分布律(注：实际上第 5 次检测可无须进行).

(2) 5 只电池中 2 只是次品，每次取一只，直到找出 2 只次品或 3 只正品为

止.写出需要测试的次数的分布律.

解　(1) X 可能取的值为 2,3,4,5.

$P\{X=2\}=P\{$第 1 次、第 2 次都取到一只次品$\}=\dfrac{2}{5}\times\dfrac{1}{4}=\dfrac{1}{10}$.

$P\{X=3\}=P\{($前两次取到一只次品$)\cap($第 3 次取到一只次品$)\}$

$=P\{$第 3 次取到一只次品$\mid$前两次取到一只次品$\}$

$\times P\{$前两次取到一只次品$\}$

$$=\frac{1}{3}\times\left(\frac{2}{5}\times\frac{3}{4}+\frac{3}{5}\times\frac{2}{4}\right)=\frac{2}{10}.$$

$P\{X=4\}=P\{($前 3 次取到一只次品$)\cap($第 4 次取到一只次品$)\}$

$=P\{$第 4 次取到一只次品$\mid$前 3 次取到一只次品$\}$

$\times P\{$前 3 次取到一只次品$\}$

$$=\frac{1}{2}\times\left(\frac{2}{5}\times\frac{3}{4}\times\frac{2}{3}+\frac{3}{5}\times\frac{2}{4}\times\frac{2}{3}+\frac{3}{5}\times\frac{2}{4}\times\frac{2}{3}\right)=\frac{3}{10}.$$

$$P\{X=5\}=1-P\{X=2\}-P\{X=3\}-P\{X=4\}=\frac{4}{10}.$$

得分布律为

X	2	3	4	5
p_k	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{4}{10}$

(2) 以 Y 表示所需测试的次数，则 Y 的可能取值为 2,3,4.

$$P\{Y=2\}=P\{X=2\}=\frac{1}{10}.$$

$\{Y=3\}$表示“前 3 次取到都是正品”或“第二只次品在第 3 次取到”，故

$$P\{Y=3\}=P\{\text{前 3 次取到的都是正品}\}+P\{X=3\}$$

$$=\frac{3}{5}\times\frac{2}{4}\times\frac{1}{3}+\frac{2}{10}=\frac{3}{10}.$$

$$P\{Y=4\}=1-P\{Y=2\}-P\{Y=3\}=\frac{6}{10}.$$

Y 的分布律为

Y	2	3	4
p_k	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{6}{10}$

11. 向某一目标发射炮弹，设炮弹弹着点离目标的距离为 R（以 10 m 计），R 服从瑞利分布，其概率密度为

$$f_R(r)=\begin{cases}\dfrac{2r}{25}\mathrm{e}^{-r^2/25}, & r>0,\\ 0, & r\leqslant 0.\end{cases}$$

若弹着点离目标不超过 5 个单位时，目标被摧毁.

(1) 求发射一枚炮弹能摧毁目标的概率.

(2) 为使至少有一枚炮弹能摧毁目标的概率不小于 0.94，问最少需要独立发射多少枚炮弹？

解 (1) 所求概率为

$$\begin{aligned}P\{R\leqslant 5\}&=\int_{-\infty}^{5}f_R(r)\mathrm{d}r=\int_0^5\frac{2r}{25}\mathrm{e}^{-r^2/25}\mathrm{d}r\\&=-\mathrm{e}^{-r^2/25}\Big|_0^5=1-\mathrm{e}^{-1}=0.632.\end{aligned}$$

(2) 设发射 n 枚炮弹，则这 n 枚炮弹都不能摧毁目标的概率为$(1-0.632)^n$，故至少有一枚炮弹能摧毁目标的概率为

$$1-(1-0.632)^n.$$

按题意需求最小的 n 使得

$$1-(1-0.632)^n\geqslant 0.94.$$

即

$$0.368^n\leqslant 0.06,$$

$$n\geqslant\frac{\ln 0.06}{\ln 0.368}=2.81.$$

故最少需要独立发射 3 枚炮弹.

12. 设一枚深水炸弹击沉一潜艇的概率为$\dfrac{1}{3}$，击伤的概率为$\dfrac{1}{2}$，击不中的概率为$\dfrac{1}{6}$. 并设击伤两次也会导致潜艇下沉. 求施放 4 枚深水炸弹能击沉潜艇的概率.（提示：先求击不沉的概率.）

解 "击沉"的逆事件为事件"击不沉"，击不沉潜艇仅出现于下述两种互不相容的情况：(1)4 枚深水炸弹全击不中潜艇（这一事件记为 A），(2)一枚击伤潜艇而另三枚击不中潜艇（这一事件记为 B）. 各枚炸弹袭击效果被认为是相互独立的. 故有

$$P(A)=\left(\frac{1}{6}\right)^4,\quad P(B)=\binom{4}{1}\times\frac{1}{2}\times\left(\frac{1}{6}\right)^3$$

(因击伤潜艇的炸弹可以是 4 枚的任一枚)，又 A,B 是互不相容的，于是，击不沉潜艇的概率为

$$P(A\cup B)=P(A)+P(B)=\frac{13}{6^4}.$$

因此，击沉潜艇的概率为

$$p=1-P(A\cup B)=1-\frac{13}{6^4}=0.989\ 97.$$

13. 一盒中装有 4 只白球、8 只黑球，从中取 3 只球，每次一只，作不放回抽样.

(1) 求第 1 次和第 3 次都取到白球的概率.(提示：考虑第 2 次的抽取.)

(2) 求在第 1 次取到白球的条件下，前 3 次都取到白球的概率.

解 以 A_1,A_2,A_3 分别表示第 1,2,3 次取到白球.

(1) $$\begin{aligned}P(A_1A_3)&=P[A_1A_3(A_2\cup\overline{A}_2)]=P(A_1A_2A_3)+P(A_1\overline{A}_2A_3)\\&=P(A_3|A_1A_2)P(A_2|A_1)P(A_1)+P(A_3|A_1\overline{A}_2)P(\overline{A}_2|A_1)P(A_1)\\&=\frac{2}{10}\times\frac{3}{11}\times\frac{4}{12}+\frac{3}{10}\times\frac{8}{11}\times\frac{4}{12}=\frac{1}{11}.\end{aligned}$$

(2) $$P(A_1A_2A_3|A_1)=\frac{P(A_1A_2A_3)}{P(A_1)}=\frac{\frac{4}{12}\times\frac{3}{11}\times\frac{2}{10}}{\frac{4}{12}}=\frac{6}{110}=\frac{3}{55}.$$

14. 设元件的寿命 T (以 h 计)服从指数分布，分布函数为

$$F(t)=\begin{cases}1-\mathrm{e}^{-0.03t}, & t>0,\\0, & t\leqslant 0.\end{cases}$$

(1) 已知元件至少工作了 30 h，求它能再至少工作 20 h 的概率.

(2) 由 3 个独立工作的此种元件组成一个 2/3[G]系统(参见第 7 题). 求这一系统的寿命 $X>20$ 的概率.

解 (1) 由指数分布的无记忆性(教材第二章 §4)知所求概率为

$$\begin{aligned}p&=P\{T>50|T>30\}=P\{T>20\}\\&=1-F(20)=\mathrm{e}^{-0.6}=0.548\ 8.\end{aligned}$$

(2) 由第 7 题知 2/3[G]系统的寿命 $X>20$ 的概率为

$$P\{X>20\}=3p^2(1-p)+p^3=p^2(3-2p)=0.573\ 0.$$

15. (1) 已知随机变量 X 的概率密度为 $f_X(x)=\frac{1}{2}\mathrm{e}^{-|x|}$，$-\infty<x<\infty$，求 X 的分布函数.

(2) 已知随机变量 X 的分布函数为 $F_X(x)$，另有随机变量

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

试求 Y 的分布律和分布函数.

解 (1) 由于

$$f_X(x)=\begin{cases}\dfrac{1}{2}e^{x}, & -\infty<x<0,\\ \dfrac{1}{2}e^{-x}, & 0\leqslant x<\infty.\end{cases}$$

当 $x<0$ 时,分布函数

$$F_X(x)=\int_{-\infty}^{x}f_X(x)dx=\int_{-\infty}^{x}\frac{1}{2}e^{x}dx=\frac{1}{2}e^{x}\Big|_{-\infty}^{x}=\frac{1}{2}e^{x},$$

当 $x\geqslant 0$ 时,分布函数

$$F_X(x)=\int_{-\infty}^{x}f_X(x)dx=\int_{-\infty}^{0}\frac{1}{2}e^{x}dx+\int_{0}^{x}\frac{1}{2}e^{-x}dx$$

$$=\frac{1}{2}+\frac{1}{2}-\frac{1}{2}e^{-x}=1-\frac{1}{2}e^{-x}.$$

故所求分布函数为

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

(2) $P\{Y=-1\}=P\{X\leqslant 0\}=F_X(0)=\dfrac{1}{2}$,

$P\{Y=1\}=1-P\{Y=-1\}=1-\dfrac{1}{2}=\dfrac{1}{2}$.

分布律为

Y	-1	1
p_k	$\frac{1}{2}$	$\frac{1}{2}$

分布函数为

$$F_Y(y)=\begin{cases}0, & y<-1,\\ \dfrac{1}{2}, & -1\leqslant y<1,\\ 1, & y\geqslant 1.\end{cases}$$

16. (1) 设随机变量 X 服从泊松分布，其分布律为

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

问当 k 取何值时 $P\{X=k\}$ 为最大?

(2) 设随机变量 X 服从二项分布，其分布律为

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

问当 k 取何值时 $P\{X=k\}$ 为最大?

解 (1) 由

$$\frac{P\{X=k\}}{P\{X=k-1\}}=\frac{\lambda^k e^{-\lambda}}{k!}\frac{(k-1)!}{\lambda^{k-1}e^{-\lambda}}$$

$$=\frac{\lambda}{k}\begin{cases}>1, & \text{当 } k<\lambda,\\ =1, & \text{当 } k=\lambda,\quad k=1,2,\cdots\\ <1, & \text{当 } k>\lambda,\end{cases}$$

知道，当 $k<\lambda$ 时，$P\{X=k\}$ 随 k 增大而递增；当 $k>\lambda$ 时，$P\{X=k\}$ 随 k 增大而递减. 从而，若 λ 为正整数，则当 $k=\lambda$ 时，$P\{X=\lambda\}=P\{X=\lambda-1\}$ 为概率的最大值，即当 $k=\lambda$ 或 $k=\lambda-1$ 时概率都取到最大值. 若 λ 不是正整数，令 $k_0=[\lambda]$（即 k_0 是 λ 的整数部分），则 $k_0<\lambda<k_0+1$，此时有

$$P\{X=k_0-1\}<P\{X=k_0\},\quad P\{X=k_0\}>P\{X=k_0+1\},$$

因此，可推得 $P\{X=k_0\}=P\{X=[\lambda]\}$ 为概率的最大值.

(2) 由

$$\frac{P\{X=k\}}{P\{X=k-1\}}=\frac{(n-k+1)p}{k(1-p)}$$

$$=1+\frac{(n+1)p-k}{k(1-p)}\begin{cases}>1, & \text{当 } k<(n+1)p,\\ =1, & \text{当 } k=(n+1)p,\quad k=1,2,\cdots,n\\ <1, & \text{当 } k>(n+1)p,\end{cases}$$

知道，当 $k<(n+1)p$ 时，$P\{X=k\}$ 随 k 增大而递增；当 $k>(n+1)p$ 时，$P\{X=k\}$ 随 k 增大而递减. 从而，若 $(n+1)p$ 为正整数，则当 $k=(n+1)p$ 时，$P\{X=(n+1)p\}=P\{X=(n+1)p-1\}$ 为概率的最大值，即当 $k=(n+1)p$ 或 $k=(n+1)p-1$ 时概率都取到最大值. 若 $(n+1)p$ 不是正整数，令 $k_0=[(n+1)p]$，则 $k_0<(n+1)p<k_0+1$，此时有

$$P\{X=k_0-1\}<P\{X=k_0\},$$
$$P\{X=k_0\}>P\{X=k_0+1\},$$

可推得 $P\{X=k_0\}=P\{X=[(n+1)p]\}$ 为概率的最大值.

17. 若离散型随机变量 X 具有分布律

X	1	2	…	n
p_k	$\frac{1}{n}$	$\frac{1}{n}$	…	$\frac{1}{n}$

则称 X 服从取值为 $1,2,\cdots,n$ 的离散型均匀分布. 对于任意非负实数 x，记$[x]$为不超过 x 的最大整数. 设随机变量 $U\sim U(0,1)$，证明 $X=[nU]+1$ 服从取值为 $1,2,\cdots,n$ 的离散型均匀分布.

证 对于 $i=1,2,\cdots,n$，

$$P\{X=i\}=P\{[nU]+1=i\}=P\{[nU]=i-1\}$$
$$=P\{i-1\leqslant nU<i\}=P\left\{\frac{i-1}{n}\leqslant U<\frac{i}{n}\right\}=\frac{1}{n}.$$

18. 设随机变量 $X\sim U(-1,2)$，求 $Y=|X|$ 的概率密度.

解 X 的概率密度为

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

记 X 的分布函数为 $F_X(x)$. 先来求 Y 的分布函数 $F_Y(y)$.

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

当 $y>0$ 时，$F_Y(y)=P\{|X|\leqslant y\}=P\{-y\leqslant X\leqslant y\}=F_X(y)-F_X(-y)$.

将 $F_Y(y)$关于 y 求导数可得 Y 的概率密度 $f_Y(y)$如下：

$$f_Y(y)=\begin{cases}f_X(y)+f_X(-y), & y>0,\\ 0, & \text{其他}.\end{cases}$$

当 $0<y<1$ 时，$-1<-y<0$. 因而 $f_X(y)=\frac{1}{3}$，$f_X(-y)=\frac{1}{3}$，此时

$$f_Y(y)=\frac{1}{3}+\frac{1}{3}=\frac{2}{3}.$$

当 $1\leqslant y<2$ 时，$-2<-y\leqslant -1$，因而 $f_X(y)=\frac{1}{3}$，$f_X(-y)=0$，此时

$$f_Y(y)=\frac{1}{3}.$$

当 $y>2$ 时，$f_X(y)=0$，$f_X(-y)=0$，因而 $f_Y(y)=0$. 故

$$f_Y(y)=\begin{cases}\frac{2}{3}, & 0<y<1,\\ \frac{1}{3}, & 1\leqslant y<2,\\ 0, & \text{其他}.\end{cases}$$

19. 设随机变量 X 的概率密度为

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

求 $Y=\dfrac{1}{X}$的概率密度.

解　因函数 $y=g(x)=\dfrac{1}{x}$严格单调减少，它的反函数 $h(y)=\dfrac{1}{y}$. 当 $0<x<\infty$时，$0<y<\infty$. 由教材第二章公式(5.2)得 Y 的概率密度为

$$f_Y(y)=\begin{cases}f_X[h(y)]\cdot|h'(y)|, & 0<y<\infty,\\ 0, & y\leqslant 0\end{cases}$$

$$=\begin{cases}f_X\left(\dfrac{1}{y}\right)\dfrac{1}{y^2}, & 0<y<\infty,\\ 0, & y\leqslant 0.\end{cases}$$

因而

$$f_Y(y)=\begin{cases}0, & y\leqslant 0,\\ \dfrac{1}{2}\,\dfrac{1}{y^2}, & 0<\dfrac{1}{y}<1,\\ \dfrac{1}{2}\,\dfrac{1}{\left(\dfrac{1}{y}\right)^2}\,\dfrac{1}{y^2}, & 1\leqslant\dfrac{1}{y}<\infty.\end{cases}$$

即

$$f_Y(y)=\begin{cases}0, & y\leqslant 0,\\ \dfrac{1}{2}, & 0<y\leqslant 1,\\ \dfrac{1}{2y^2}, & 1<y<\infty.\end{cases}$$

本题 X 和$\dfrac{1}{X}$的概率密度相同(除去原点外).

20. 设随机变量 X 服从参数为$\dfrac{1}{\lambda}$的指数分布. 验证随机变量 $Y=[X]$服从参数为 $1-\mathrm{e}^{-\lambda}$的几何分布. 这一事实表明连续型随机变量的函数可以是离散型随机变量.

证　X 的概率密度为 $f_X(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x>0,\\ 0, & \text{其他}.\end{cases}$ X 的值域为$(0,\infty)$，故 $Y=[X]$的值域是$\{0,1,2,\cdots\}$，Y 是离散型随机变量. 对于任意非负整数 y 有

$$
\begin{aligned}
P\{Y=y\}&=P\{[X]=y\}=P\{y\leqslant X<y+1\}\\
&=\int_y^{y+1}\lambda e^{-\lambda x}dx=e^{-\lambda y}-e^{-\lambda(y+1)}\\
&=(1-e^{-\lambda})(e^{-\lambda})^y\\
&=(1-e^{-\lambda})[1-(1-e^{-\lambda})]^y,\quad y=0,1,2,\cdots.
\end{aligned}
$$

这就是说 Y 服从参数为 $1-e^{-\lambda}$ 的几何分布①. 这表示一个连续型随机变量经变换变成了离散型随机变量.

21. 投掷一枚硬币直至正面出现为止，引入随机变量

$$X=\text{投掷总次数}.$$

$$Y=\begin{cases}1, & \text{若首次投掷得到正面},\\ 0, & \text{若首次投掷得到反面}.\end{cases}$$

(1) 求 X 和 Y 的联合分布律及边缘分布律.

(2) 求条件概率 $P\{X=1|Y=1\}, P\{Y=2|X=1\}$.

解 (1) Y 的可能值是 $0,1$，X 的可能值是 $1,2,3,\cdots$.

$$
\begin{aligned}
P\{X=1,Y=1\}&=P\{Y=1|X=1\}P\{X=1\}\\
&=1\times\frac{1}{2}=\frac{1}{2}.
\end{aligned}
$$

(因 $X=1$ 必定首次得正面，故 $P\{Y=1|X=1\}=1$.)

若 $k>1$，

$$
\begin{aligned}
P\{X=k,Y=1\}&=P\{Y=1|X=k\}P\{X=k\}\\
&=0\times\frac{1}{2^k}=0.
\end{aligned}
$$

(因 $X=k>1$，首次得正面是不可能的，故 $P\{Y=1|X=k\}=0, k=2,3,\cdots$.)

$$
\begin{aligned}
P\{X=1,Y=0\}&=P\{Y=0|X=1\}P\{X=1\}\\
&=0\times\frac{1}{2}=0.
\end{aligned}
$$

(因 $X=1$ 必须首次得正面，故 $P\{Y=0|X=1\}=0$.)

当 $k>1$，

$$
\begin{aligned}
P\{X=k,Y=0\}&=P\{Y=0|X=k\}P\{X=k\}\\
&=1\times\frac{1}{2^k},\quad k=2,3,\cdots.
\end{aligned}
$$

(因 $X=k>1$，必定首次得反面，故 $P\{Y=0|X=k\}=1$.)

---

① 几何分布 X 的分布律为 $P\{X=x\}=p(1-p)^{x-1}, x=1,2,\cdots$. 令 $Y=X-1$，分布律又可写成 $P\{Y=y\}=p(1-p)^y, y=0,1,2,\cdots$.

综上，得(X,Y)的分布律及边缘分布律如下：

$Y \backslash X$	1	2	3	4	…	$P\{Y=j\}$
0	0	$\frac{1}{2^2}$	$\frac{1}{2^3}$	$\frac{1}{2^4}$	…	$\frac{1}{2}$
1	$\frac{1}{2}$	0	0	0	…	$\frac{1}{2}$
$P\{X=i\}$	$\frac{1}{2}$	$\frac{1}{2^2}$	$\frac{1}{2^3}$	$\frac{1}{2^4}$	…	1

(2) $P\{X=1|Y=1\}=\dfrac{P\{X=1,Y=1\}}{P\{Y=1\}}=\dfrac{1/2}{1/2}=1,$

$$P\{Y=2|X=1\}=\frac{P\{X=1,Y=2\}}{P\{X=1\}}=0.$$

22. 设随机变量 $X\sim\pi(\lambda)$，随机变量 $Y=\max\{X,2\}$. 试求X和Y的联合分布律及边缘分布律.

解 X 的分布律为

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

X 的可能值是 $0,1,2,\cdots$;Y 的可能值为 $2,3,4,\cdots$.

$$\begin{aligned}P\{X=0,Y=2\}&=P\{Y=2|X=0\}P\{X=0\}\\&=1\times P\{X=0\}=e^{-\lambda}.\\P\{X=1,Y=2\}&=P\{Y=2|X=1\}P\{X=1\}\\&=1\times P\{X=1\}=\lambda e^{-\lambda}.\end{aligned}$$

$i\geqslant 2$ 时，

$$\begin{aligned}P\{X=i,Y=j\}&=P\{Y=j|X=i\}P\{X=i\}\\&=\begin{cases}1\times P\{X=i\}, & j=i,\\0\times P\{X=i\}, & j\neq i\end{cases}\\&=\begin{cases}\dfrac{\lambda^i e^{-\lambda}}{i!}, & j=i,\\0, & j\neq i,\end{cases}\quad j=2,3,4,\cdots.\end{aligned}$$

即得 X,Y 的联合分布律及边缘分布律为

$Y \backslash X$	0	1	2	3	4	5	…	$P\{Y=j\}$
2	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2 e^{-\lambda}}{2!}$	0	0	0	…	$\sum_{i=0}^{2}\frac{\lambda^i e^{-\lambda}}{i!}$

续表

Y \ X	0	1	2	3	4	5	…	$P\{Y=j\}$
3	0	0	0	$\frac{\lambda^3 e^{-\lambda}}{3!}$	0	0	…	$\frac{\lambda^3 e^{-\lambda}}{3!}$
4	0	0	0	0	$\frac{\lambda^4 e^{-\lambda}}{4!}$	0	…	$\frac{\lambda^4 e^{-\lambda}}{4!}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮		⋮
$P\{X=i\}$	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2 e^{-\lambda}}{2!}$	$\frac{\lambda^3 e^{-\lambda}}{3!}$	$\frac{\lambda^4 e^{-\lambda}}{4!}$	…	…	1

23. 设 X,Y 是相互独立的泊松随机变量,参数分别为 λ_1,λ_2,求给定 $X+Y=n$ 的条件下 X 的条件分布.

解 $P\{X=k|X+Y=n\}$

$$=\frac{P\{X=k,X+Y=n\}}{P\{X+Y=n\}}$$

$$=\frac{P\{X=k,Y=n-k\}}{P\{X+Y=n\}}\xlongequal{\text{由独立性}}\frac{P\{X=k\}P\{Y=n-k\}}{P\{X+Y=n\}}$$

$$=\frac{\lambda_1^k e^{-\lambda_1}}{k!}\cdot\frac{\lambda_2^{n-k}e^{-\lambda_2}}{(n-k)!}\left[\frac{(\lambda_1+\lambda_2)^n e^{-(\lambda_1+\lambda_2)}}{n!}\right]^{-1}=\frac{n!}{(n-k)!k!}\frac{\lambda_1^k\lambda_2^{n-k}}{(\lambda_1+\lambda_2)^n}$$

$$=\binom{n}{k}\frac{\lambda_1^k\lambda_2^{n-k}}{(\lambda_1+\lambda_2)^n}=\binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k}.$$

这就是说给定条件 $X+Y=n$ 下 X 的条件分布为以 $n,\lambda_1/(\lambda_1+\lambda_2)$ 为参数的二项分布.

24. 一教授将两篇论文分别交给两个打字员打印. 以 X,Y 分别表示第一篇和第二篇论文的打印错误. 设 $X\sim\pi(\lambda)$,$Y\sim\pi(\mu)$,X,Y 相互独立.

(1) 求 X,Y 的联合分布律.

(2) 求两篇论文总共至多有 1 个打印错误的概率.

解 (1) X,Y 的联合分布律为

$$P\{X=x,Y=y\}=\frac{\lambda^x e^{-\lambda}}{x!}\cdot\frac{\mu^y e^{-\mu}}{y!}=\frac{\lambda^x\mu^y e^{-(\lambda+\mu)}}{x!y!},\quad x,y=0,1,2,\cdots.$$

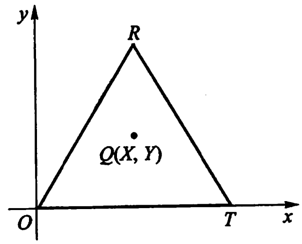
(2) 两篇论文总共至多有 1 个打印错误的概率为

$$\begin{aligned}P\{X+Y\leqslant 1\}&=P(\{X+Y=0\}\cup\{X+Y=1\})\\&=P\{X=0,Y=0\}+P\{X=1,Y=0\}+P\{X=0,Y=1\}\\&=e^{-(\lambda+\mu)}+\lambda e^{-(\lambda+\mu)}+\mu e^{-(\lambda+\mu)}=e^{-(\lambda+\mu)}(1+\lambda+\mu).\end{aligned}$$

25. 一等边三角形 ROT（如题 16.25 图）的边长为 1，在三角形内随机地取点 $Q(X,Y)$（意指随机点 (X,Y) 在三角形 ROT 内均匀分布）.

(1) 写出随机变量 (X,Y) 的概率密度.

(2) 求点 Q 到底边 OT 的距离的分布函数.



题 16.25 图

解 (1) 因三角形 ROT 的面积为 $\sqrt{3}/4$，故 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases}\dfrac{4}{\sqrt{3}}, & 0\leqslant y\leqslant\sqrt{3}x, 0\leqslant y\leqslant-\sqrt{3}(x-1),\\ 0, & \text{其他}.\end{cases}$$

(2) 点 $Q(X,Y)$ 到底边 OT 的距离就是 Y，因而求 Q 到 OT 的距离的分布函数，就是求 (X,Y) 关于 Y 的边缘分布函数，现在

$$f_Y(y)=\int_{y/\sqrt{3}}^{1-y/\sqrt{3}}f(x,y)\mathrm{d}x=\frac{4}{\sqrt{3}}\left(1-\frac{2y}{\sqrt{3}}\right),\quad 0<y<\frac{\sqrt{3}}{2},$$

从而

$$f_Y(y)=\begin{cases}\dfrac{4}{\sqrt{3}}\left(1-\dfrac{2y}{\sqrt{3}}\right), & 0<y<\dfrac{\sqrt{3}}{2},\\ 0, & \text{其他}.\end{cases}$$

Y 的分布函数为

$$F_Y(y)=\begin{cases}0, & y<0,\\ \dfrac{4}{\sqrt{3}}y-\dfrac{4}{3}y^2, & 0\leqslant y<\dfrac{\sqrt{3}}{2},\\ 1, & y\geqslant\dfrac{\sqrt{3}}{2}.\end{cases}$$

26. 设随机变量 (X,Y) 具有概率密度

$$f(x,y)=\begin{cases}xe^{-x(y+1)}, & x>0,y>0,\\ 0, & \text{其他}.\end{cases}$$

(1) 求边缘概率密度 $f_X(x)$，$f_Y(y)$.

(2) 求条件概率密度 $f_{X|Y}(x|y)$，$f_{Y|X}(y|x)$.

解 (1) 当 $x>0$ 时，

$$f_X(x)=\int_0^{\infty}xe^{-x(y+1)}\mathrm{d}y=e^{-x}\left[-e^{-xy}\right]\Big|_{y=0}^{y=\infty}=e^{-x},$$

当 $y>0$ 时，

$$f_Y(y)=\int_0^{\infty}xe^{-x(y+1)}\mathrm{d}x=\frac{-xe^{-x(y+1)}}{y+1}\Bigg|_{x=0}^{x=\infty}+\frac{1}{y+1}\int_0^{\infty}e^{-x(y+1)}\mathrm{d}x$$

$$=\frac{-\mathrm{e}^{-x(y+1)}}{(y+1)^2}\Bigg|_{x=0}^{x=\infty}=\frac{1}{(y+1)^2}.$$

故边缘概率密度分别是

$$f_X(x)=\begin{cases}\mathrm{e}^{-x}, & x>0,\\ 0, & \text{其他}.\end{cases}\quad f_Y(y)=\begin{cases}\dfrac{1}{(y+1)^2}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

(2) 条件概率密度：

当 $x>0$ 时，

$$f_{Y|X}(y|x)=\begin{cases}\dfrac{x\mathrm{e}^{-x(y+1)}}{\mathrm{e}^{-x}}, & y>0,\\ 0, & y\text{ 取其他值}\end{cases}$$

$$=\begin{cases}x\mathrm{e}^{-xy}, & y>0,\\ 0, & y\text{ 取其他值}.\end{cases}$$

当 $y>0$ 时，

$$f_{X|Y}(x|y)=\begin{cases}\dfrac{x\mathrm{e}^{-x(y+1)}}{1/(y+1)^2}, & x>0,\\ 0, & x\text{ 取其他值}\end{cases}$$

$$=\begin{cases}x(y+1)^2\mathrm{e}^{-x(y+1)}, & x>0,\\ 0, & x\text{ 取其他值}.\end{cases}$$

27. 设有随机变量 U 和 V，它们都仅取 $1,-1$ 两个值. 已知

$$P\{U=1\}=\frac{1}{2},$$

$$P\{V=1|U=1\}=\frac{1}{3}=P\{V=-1|U=-1\}.$$

(1) 求 U 和 V 的联合分布律.

(2) 求 x 的方程 $x^2+Ux+V=0$ 至少有一个实根的概率.

(3) 求 x 的方程 $x^2+(U+V)x+U+V=0$ 至少有一个实根的概率.

解 (1) $P\{U=1,V=1\}=P\{V=1|U=1\}P\{U=1\}$

$$=\frac{1}{3}\times\frac{1}{2}=\frac{1}{6}.$$

$$P\{U=-1,V=-1\}=P\{V=-1|U=-1\}P\{U=-1\}$$

$$=\frac{1}{3}\times(1-P\{U=1\})$$

$$=\frac{1}{3}\times\frac{1}{2}=\frac{1}{6}.$$

$$P\{U=1,V=-1\}=P\{V=-1|U=1\}P\{U=1\}$$

$$=(1-P\{V=1|U=1\})P\{U=1\}$$
$$=\frac{2}{3}\times\frac{1}{2}=\frac{2}{6}.$$
$$P\{U=-1,V=1\}=P\{V=1|U=-1\}P\{U=-1\}$$
$$=(1-P\{V=-1|U=-1\})P\{U=-1\}$$
$$=\frac{2}{3}\times\frac{1}{2}=\frac{2}{6}.$$

U,V 的联合分布律为

V \ U	-1	1
-1	$\frac{1}{6}$	$\frac{2}{6}$
1	$\frac{2}{6}$	$\frac{1}{6}$

(2) 方程 $x^2+Ux+V=0$ 当且仅当在 $\Delta=U^2-4V\geqslant 0$ 时至少有一实根，因而所求的概率为

$$P\{\Delta\geqslant 0\}=P\{U^2-4V\geqslant 0\}=P\{V=-1\}=\frac{1}{2}.$$

(3) 方程 $x^2+(U+V)x+U+V=0$ 当且仅当在 $\Delta=(U+V)^2-4(U+V)\geqslant 0$ 时至少有一实根，因而所求的概率为

$$P\{\Delta\geqslant 0\}=P\{U=-1,V=-1\}+P\{U=-1,V=1\}+P\{U=1,V=-1\}$$
$$=\frac{5}{6}.$$

28. 某图书馆一天的读者人数 $X\sim\pi(\lambda)$，任一读者借书的概率为 p，各读者借书与否相互独立. 记一天读者借书的人数为 Y，求 X 和 Y 的联合分布律.

解 读者借书人数的可能值为 $Y=0,1,2,\cdots,Y\leqslant X$，

$$P\{X=k,Y=i\}=P\{Y=i|X=k\}P\{X=k\}$$
$$=\binom{k}{i}p^i(1-p)^{k-i}\frac{\lambda^k e^{-\lambda}}{k!},\quad k=0,1,2,\cdots;\ i=0,1,2,\cdots,k.$$

29. 设随机变量 X,Y 相互独立，且都服从均匀分布 $U(0,1)$，求两变量之一至少为另一变量之值之两倍的概率.

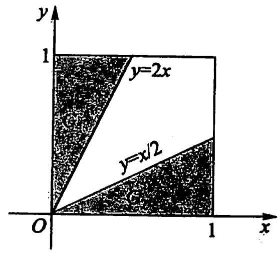
解 按题意知，(X,Y) 在区域：$G=\{(x,y)|0<x<1,0<y<1\}$ 服从均匀分布，其概率密度为

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

所求概率为

$$\begin{aligned} p &= P\{Y > 2X\} + P\{X > 2Y\} \\ &= \iint_{G_1} f(x,y)\mathrm{d}x\mathrm{d}y + \iint_{G_2} f(x,y)\mathrm{d}x\mathrm{d}y \\ &= G_1 \text{ 的面积} + G_2 \text{ 的面积} = \frac{1}{2}, \end{aligned}$$

G_1，G_2 见题 16.29 图.



题 16.29 图

30. 一家公司有一份保单招标，两家保险公司竞标. 规定标书的保险费必须在 20 万元至 22 万元之间. 若两份标书保险费相差 2 000 元或 2 000 元以上，招标公司将选择报价低者，否则就重新招标. 设两家保险公司的报价是相互独立的，且都在 20 万元至 22 万元之间均匀分布. 试求招标公司需重新招标的概率.

解　设以 X，Y 分别表示两家保险公司提出的保费.

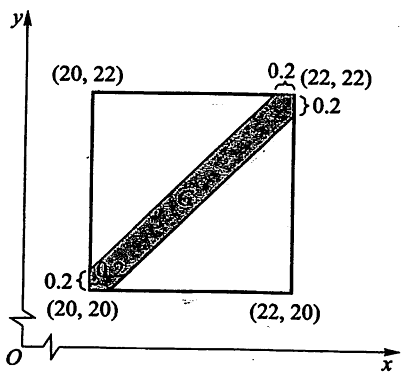
由假设 X 和 Y 的概率密度均为

$$f(u)=\begin{cases}\dfrac{1}{2}, & 20<u<22,\\ 0, & \text{其他}.\end{cases}$$

因 X，Y 相互独立，故 (X,Y) 的概率密度为

$$f(x,y)=f_X(x)f_Y(y)=\begin{cases}\dfrac{1}{4}, & 20<x<22, 20<y<22,\\ 0, & \text{其他}.\end{cases}$$

按题意需求概率 $P\{|X-Y|<0.2\}$. 画出区域 $\{(x,y) \mid |x-y|<0.2\}$，以及正方形 $\{(x,y) \mid 20<x<22, 20<y<22\}$，如题 16.30 图，它们公共部分的面积 G 为



题 16.30 图

$$G=\text{正方形面积}-2\times\text{三角形面积}=4-1.8\times1.8=0.76.$$

所求概率$=\dfrac{0.76}{2\times2}=0.19.$

31. 设随机变量 $X\sim N(0,\sigma_1^2)$，$Y\sim N(0,\sigma_2^2)$，且 X,Y 相互独立，求概率

$$P\{0<\sigma_2X-\sigma_1Y<2\sigma_1\sigma_2\}.$$

解 因 X,Y 相互独立，其线性组合 $\sigma_2X-\sigma_1Y$ 仍为正态变量，而

$$E(\sigma_2X-\sigma_1Y)=\sigma_2E(X)-\sigma_1E(Y)=0,$$

$$D(\sigma_2X-\sigma_1Y)=\sigma_2^2D(X)+\sigma_1^2D(Y)=2\sigma_1^2\sigma_2^2,$$

故 $\sigma_2X-\sigma_1Y\sim N(0,2\sigma_1^2\sigma_2^2)$. 因而

$$\begin{aligned}P\{0<\sigma_2X-\sigma_1Y<2\sigma_1\sigma_2\}&=P\left\{0<\frac{\sigma_2X-\sigma_1Y-0}{\sqrt{2\sigma_1^2\sigma_2^2}}\leqslant\frac{2\sigma_1\sigma_2-0}{\sqrt{2\sigma_1^2\sigma_2^2}}\right\}\\&=\Phi\left(\frac{2\sigma_1\sigma_2}{\sqrt{2\sigma_1^2\sigma_2^2}}\right)-\Phi(0)=\Phi(\sqrt{2})-0.5\\&=0.920\ 7-0.5=0.420\ 7.\end{aligned}$$

32. NBA 篮球赛中有这样的规律，两支实力相当的球队比赛时，每节主队得分与客队得分之差为正态随机变量，均值为 1.5，方差为 6，并且假设四节的比分差是相互独立的. 问：

(1) 主队胜的概率有多大？

(2) 在前半场主队落后 5 分的情况下，主队得胜的概率有多大？

(3) 在第一节主队赢 5 分的情况下，主队得胜的概率有多大？

解 以 $X_i(i=1,2,3,4)$ 记主队在第 i 节的得分与客队在第 i 节的得分之差，则有 $X_i\sim N(1.5,6)$，$\sum_{i=1}^{4}X_i\sim N(4\times1.5,4\times6)$. 记 Z 为标准正态随机变量.

(1) $$\begin{aligned}P\left\{\sum_{i=1}^{4}X_i>0\right\}&=P\left\{\frac{\sum_{i=1}^{4}X_i-4\times1.5}{\sqrt{4\times6}}>\frac{-6}{\sqrt{4\times6}}\right\}\\&=P\{Z>-1.224\ 7\}=0.889\ 7.\end{aligned}$$

(2) 由独立性

$$\begin{aligned}P\left\{\sum_{i=1}^{4}X_i>0\,\middle|\,\sum_{i=1}^{2}X_i=-5\right\}&=P\{X_3+X_4>5\}\\&=P\left\{\frac{X_3+X_4-3}{\sqrt{2\times6}}>\frac{5-3}{\sqrt{12}}\right\}=P\left\{Z>\frac{\sqrt{3}}{3}\right\}\\&=P\{Z>0.577\ 4\}=0.281\ 8.\end{aligned}$$

(3) $P\left\{\sum_{i=1}^{4}X_i>0\,\middle|\,X_1=5\right\}=P\{5+X_2+X_3+X_4>0\}$

$$
\begin{aligned}
&=P\{X_2+X_3+X_4>-5\}\\
&=P\left\{\frac{X_2+X_3+X_4-4.5}{\sqrt{3\times 6}}>\frac{-5-4.5}{\sqrt{18}}\right\}\\
&=P\left\{Z>\frac{-9.5}{\sqrt{18}}\right\}=P\{Z>-2.239\,2\}\\
&=0.987\,4.
\end{aligned}
$$

33. 产品的某种性能指标的测量值 X 是随机变量，设 X 的概率密度为

$$
f_X(x)=\begin{cases}x\mathrm{e}^{-\frac{1}{2}x^2}, & x>0,\\ 0, & 其他.\end{cases}
$$

测量误差 $Y\sim U(-\varepsilon,\varepsilon)$，$X,Y$ 相互独立. 求 $Z=X+Y$ 的概率密度 $f_Z(z)$，并验证

$$
P\{Z>\varepsilon\}=\frac{1}{2\varepsilon}\int_0^{2\varepsilon}\mathrm{e}^{-\frac{u^2}{2}}\mathrm{d}u.
$$

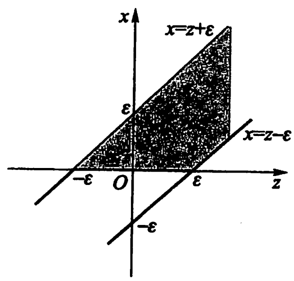
解 (1) Y 的概率密度为

$$
f_Y(y)=\begin{cases}\dfrac{1}{2\varepsilon}, & -\varepsilon<y<\varepsilon,\\ 0, & 其他.\end{cases}
$$

故 $Z=X+Y$ 的概率密度为

$$
f_Z(z)=f_X*f_Y=\int_{-\infty}^{\infty}f_X(x)f_Y(z-x)\mathrm{d}x,
$$

仅当
$$
\begin{cases}x>0,\\ -\varepsilon<z-x<\varepsilon,\end{cases}\quad 即\begin{cases}x>0,\\ z-\varepsilon<x<z+\varepsilon\end{cases}
$$
时上述积分的被积函数不等于零，参考题 16.33 图，得



题 16.33 图

$$
f_Z(z)=\begin{cases}\dfrac{1}{2\varepsilon}\displaystyle\int_0^{z+\varepsilon}x\mathrm{e}^{-\frac{1}{2}x^2}\mathrm{d}x, & -\varepsilon<z<\varepsilon,\\ \dfrac{1}{2\varepsilon}\displaystyle\int_{z-\varepsilon}^{z+\varepsilon}x\mathrm{e}^{-\frac{1}{2}x^2}\mathrm{d}x, & z\geqslant\varepsilon,\\ 0, & 其他\end{cases}
$$

$$
=\begin{cases}\dfrac{1}{2\varepsilon}\left[1-\mathrm{e}^{-\frac{1}{2}(z+\varepsilon)^2}\right], & -\varepsilon<z<\varepsilon,\\ \dfrac{1}{2\varepsilon}\left[\mathrm{e}^{-\frac{1}{2}(z-\varepsilon)^2}-\mathrm{e}^{-\frac{1}{2}(z+\varepsilon)^2}\right], & z\geqslant\varepsilon,\\ 0, & 其他.\end{cases}
$$

(2) $P\{Z>\varepsilon\}=\int_{\varepsilon}^{\infty}f_Z(z)\mathrm{d}z$

$$=\frac{1}{2\varepsilon}\left[\int_{\varepsilon}^{\infty}\mathrm{e}^{-\frac{1}{2}(z-\varepsilon)^2}\mathrm{d}z-\int_{\varepsilon}^{\infty}\mathrm{e}^{-\frac{1}{2}(z+\varepsilon)^2}\mathrm{d}z\right]$$

$$\xlongequal{\text{记成}}\frac{1}{2\varepsilon}(\mathrm{I}+\mathrm{II}),$$

其中

$$\mathrm{I}=\int_{\varepsilon}^{\infty}\mathrm{e}^{-\frac{1}{2}(z-\varepsilon)^2}\mathrm{d}z\xlongequal{\text{令 } z-\varepsilon=u}\int_{0}^{\infty}\mathrm{e}^{-\frac{1}{2}u^2}\mathrm{d}u,$$

$$\mathrm{II}=-\int_{\varepsilon}^{\infty}\mathrm{e}^{-\frac{1}{2}(z+\varepsilon)^2}\mathrm{d}z\xlongequal{\text{令 } z+\varepsilon=u}-\int_{2\varepsilon}^{\infty}\mathrm{e}^{-\frac{1}{2}u^2}\mathrm{d}u.$$

于是

$$P\{Z>\varepsilon\}=\frac{1}{2\varepsilon}(\mathrm{I}+\mathrm{II})=\frac{1}{2\varepsilon}\int_{0}^{2\varepsilon}\mathrm{e}^{-\frac{1}{2}u^2}\mathrm{d}u.$$

34. 在一化学过程中，产品中有份额 X 为杂质，而在杂质中有份额 Y 是有害的，而其余部分不影响产品的质量. 设 $X\sim U(0,0.1)$，$Y\sim U(0,0.5)$，且 X 和 Y 相互独立，求产品中有害杂质份额 Z 的概率密度.

解　因 $Z=XY$，$X\sim U(0,0.1)$，$Y\sim U(0,0.5)$且 X,Y 相互独立. 于是 Z 的概率密度为

$$f_Z(z)=\int_{-\infty}^{\infty}\frac{1}{|x|}f_1(x)f_2\left(\frac{z}{x}\right)\mathrm{d}x,\qquad (*_1)$$

其中，

$$f_1(x)=\begin{cases}10, & 0<x<0.1,\\ 0, & \text{其他},\end{cases}$$

$$f_2(x)=\begin{cases}2, & 0<x<0.5,\\ 0, & \text{其他}.\end{cases}$$

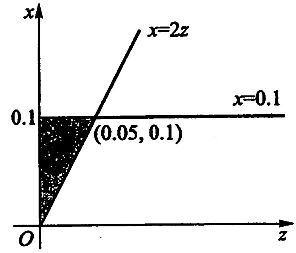
易知仅当

$$\begin{cases}0<x<0.1,\\ 0<z/x<0.5,\end{cases}\quad \text{即}\begin{cases}0<x<0.1,\\ 0<2z<x\end{cases}$$

时($*_1$)中的被积函数不等于零，参考题 16.34 图即得

$$f_Z(z)=\begin{cases}\int_{2z}^{0.1}10\cdot 2\cdot\frac{1}{x}\mathrm{d}x, & 0<z<0.05,\\ 0, & \text{其他}\end{cases}$$

$$=\begin{cases}20\ln x\Big|_{2z}^{0.1}, & 0<z<0.05,\\ 0, & \text{其他}\end{cases}$$



题 16.34 图

$$=\begin{cases}-20\ln(20z), & 0<z<0.05,\\ 0, & 其他.\end{cases}$$

35. 设随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}e^{-y}, & 0<x<y,\\ 0, & 其他.\end{cases}$$

(1) 求(X,Y)的边缘概率密度.

(2) 问 X,Y 是否相互独立?

(3) 求 $X+Y$ 的概率密度 $f_{X+Y}(z)$.

(4) 求条件概率密度 $f_{X|Y}(x|y)$.

(5) 求条件概率 $P\{X>3|Y<5\}$.

(6) 求条件概率 $P\{X>3|Y=5\}$.

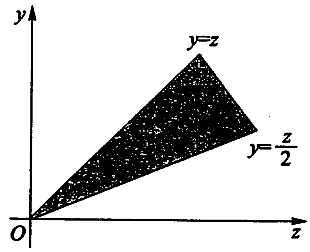
解 (1) $f_X(x)=\begin{cases}\int_x^{\infty}e^{-y}dy=e^{-x}, & x>0,\\ 0, & 其他.\end{cases}$

$$f_Y(y)=\begin{cases}\int_0^{y}e^{-y}dx=ye^{-y}, & y>0,\\ 0, & 其他.\end{cases}$$

(2) X,Y 不是相互独立的.

(3) $f_{X+Y}(z)=\int_{-\infty}^{\infty}f(z-y,y)dy.$

仅当 $0<z-y<y$,即$\begin{cases}2y>z,\\ y>0,\\ y<z\end{cases}$时被积函数不为零.



题 16.35 图 1

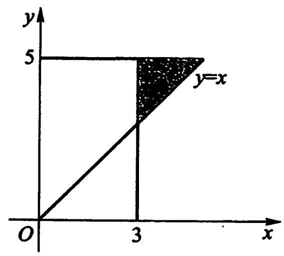
如题 16.35 图 1 得

$$f_{X+Y}(z)=\begin{cases}\int_{z/2}^{z}e^{-y}dy=e^{-z/2}-e^{-z}, & z>0,\\ 0, & 其他.\end{cases}$$

(4) 对于 $y>0$,

$$f_{X|Y}(x|y)=\begin{cases}\dfrac{e^{-y}}{ye^{-y}}=\dfrac{1}{y}, & 0<x<y,\\ 0, & 其他.\end{cases}$$

即对于固定的 $y\ (y>0)$,X 的条件分布是区间$(0,y)$上的均匀分布.



题 16.35 图 2

(5) 如题 16.35 图 2,条件概率为

$$P\{X>3|Y<5\}=\frac{P\{X>3,Y<5\}}{P\{Y<5\}}$$

$$=\frac{\iint\limits_D e^{-y}\mathrm{d}x\mathrm{d}y}{\int_0^5 f_Y(y)\mathrm{d}y},$$

$$分子=\int_3^5\int_x^5 e^{-y}\mathrm{d}y\mathrm{d}x=\int_3^5[-e^{-y}]\Big|_x^5\mathrm{d}x$$

$$=\int_3^5(-e^{-5}+e^{-x})\mathrm{d}x=-3e^{-5}+e^{-3},$$

$$分母=\int_0^5 f_Y(y)\mathrm{d}y=\int_0^5 ye^{-y}\mathrm{d}y$$

$$=-ye^{-y}\Big|_0^5+\int_0^5 e^{-y}\mathrm{d}y=-6e^{-5}+1,$$

故
$$P\{X>3|Y<5\}=0.030\ 82.$$

(6) $$f_{X|Y}(x|5)=\begin{cases}\dfrac{1}{5}, & 0<x<5,\\ 0, & 其他.\end{cases}$$

$$P\{X>3|Y=5\}=\int_3^5\frac{1}{5}\mathrm{d}x=\frac{2}{5}.$$

36. 设某图书馆的读者借阅甲种图书的概率为 p，借阅乙种图书的概率为 α，设每人借阅甲、乙图书的行动相互独立，读者之间的行动也相互独立.

(1) 某天恰有 n 个读者，求借阅甲种图书的人数的数学期望.

(2) 某天恰有 n 个读者，求甲、乙两种图书中至少借阅一种的人数的数学期望.

解 (1) 以 X 表示某天读者中借阅甲种图书的人数，因各人借阅甲种图书的概率均为 p，且由题设各人是否借阅相互独立，故 $X\sim b(n,p)$，因此

$$E(X)=np.$$

(2) 以 A 表示事件“读者借阅甲种图书”，以 B 表示事件“读者借阅乙种图书”，则就该读者而言，有

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

借阅两种图书的行动相互独立，故 $P(AB)=P(A)P(B)$，按题设

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

以 Y 表示至少借阅一种图书的人数，由题设各人是否借阅相互独立，知 $Y\sim b(n,p+\alpha-p\alpha)$，故

$$E(Y)=n(p+\alpha-p\alpha).$$

也可这样做. 引入随机变量：

$$Z_i=\begin{cases}1, & 若第\ i\ 个读者至少借阅甲、乙两种图书的一种,\\ 0, & 若第\ i\ 个读者不借阅甲、乙两种图书的任一种,\end{cases}\quad i=1,2,\cdots,n.$$

则

$$Y=\sum_{i=1}^{n}Z_i,$$

$$E(Y)=E\left(\sum_{i=1}^{n}Z_i\right)=\sum_{i=1}^{n}E(Z_i)=n(p+\alpha-p\alpha).$$

这里不需假设读者之间的行动相互独立.（对第(1)小题也类似.）

37. 某种鸟在某时间区间$(0,t_0]$下蛋数为 1～5 只，下 r 只蛋的概率与 r 成正比. 一个拾蛋人在时刻 t_0 去收集鸟蛋，但他仅当鸟窝中多于 3 只蛋时才从中取走一只蛋. 在某处有这种鸟的鸟窝 6 个（每个鸟窝保存完好，各鸟窝中蛋的只数相互独立).

(1) 写出一个鸟窝中鸟蛋只数 X 的分布律.

(2) 对于指定的一个鸟窝，求拾蛋人在该鸟窝中拾到一只蛋的概率.

(3) 求拾蛋人在 6 个鸟窝中拾到蛋的总数 Y 的分布律及数学期望.

(4) 求 $P\{Y<4\}$，$P\{Y>4\}$.

(5) 当一个拾蛋人在这 6 个鸟窝中拾过蛋后，紧接着又有一个拾蛋人到这些鸟窝中拾蛋，也仅当鸟窝中多于 3 只蛋时，拾取一只蛋，求第二个拾蛋人拾得蛋数 Z 的数学期望.

解 (1) 设该种鸟在$(0,t_0]$内下蛋数为 X，按题意 $P\{X=r\}=Cr$，$r=1,2,3,4,5$，其中，C 为待定常数. 因 $\sum\limits_{r=1}^{5}P\{X=r\}=1$，即有 $\sum\limits_{r=1}^{5}Cr=15C=1$，所以 $C=\frac{1}{15}$，因此，X 的分布律为

$$P\{X=r\}=\frac{1}{15}r,\quad r=1,2,3,4,5.$$

(2) 因当且仅当窝中蛋数多于 3 时，拾蛋人从中取走一只蛋，故拾蛋人在该窝中拾取一只蛋的概率为

$$P\{X>3\}=P\{X=4\}+P\{X=5\}=\frac{4}{15}+\frac{5}{15}=\frac{3}{5}.$$

(3) 记拾蛋人在 6 个鸟窝中拾到蛋的总数为 Y，则 $Y\sim b\left(6,\frac{3}{5}\right)$，故

$$E(Y)=6\times\frac{3}{5}=\frac{18}{5}.$$

(4) $P\{Y<4\}=1-P\{Y=4\}-P\{Y=5\}-P\{Y=6\}$

$$=1-\binom{6}{4}\left(\frac{3}{5}\right)^4\left(\frac{2}{5}\right)^2-\binom{6}{5}\left(\frac{3}{5}\right)^5\frac{2}{5}-\left(\frac{3}{5}\right)^6=0.456,$$

$P\{Y>4\}=P\{Y=5\}+P\{Y=6\}=0.233.$

(5) 第二个拾蛋人仅当鸟窝中最初有 5 只蛋时，他才能从该窝中拾到一只蛋，故他在一个鸟窝中拾到一只蛋的概率为 $p=P\{X=5\}=\frac{1}{3}$. 以 Z 记第二个拾蛋人拾到蛋的总数，则 $Z\sim b\left(6,\frac{1}{3}\right)$，故有 $E(Z)=6\times\frac{1}{3}=2$.

38. 设袋中有 r 只白球、$N-r$ 只黑球. 在袋中取球 $n(n\leqslant r)$ 次，每次任取一只作不放回抽样，以 Y 表示取到白球的个数，求 $E(Y)$.

解 引入随机变量 X_i：

$$X_i=\begin{cases}1, & \text{若第 } i \text{ 次取到白球},\\ 0, & \text{若第 } i \text{ 次取到黑球},\end{cases} \quad i=1,2,\cdots,n,$$

则 n 次取球得到的白球数

$$Y=X_1+X_2+\cdots+X_n.$$

而

$$P\{X_i=1\}=P\{\text{第 } i \text{ 次取球得到白球}\}=\frac{r}{N},$$

X_i 的分布律为

X_i	0	1
p_k	$1-\frac{r}{N}$	$\frac{r}{N}$

$i=1,2,\cdots,n,$

即知 X_i 的数学期望为 $E(X_i)=\frac{r}{N}$. 于是得 Y 的数学期望为

$$E(Y)=E\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}E(X_i)=n\times\frac{r}{N}=\frac{nr}{N}.$$

本题也可以按以下的方式写出 Y 的表达式，从而求得 $E(Y)$. 将球编号，引入随机变量 X_i：

$$X_i=\begin{cases}1, & \text{若第 } i \text{ 号白球被取到},\\ 0, & \text{若第 } i \text{ 号白球未被取到},\end{cases} \quad i=1,2,\cdots,r,$$

则

$$Y=X_1+X_2+\cdots+X_r.$$

事件 $\{X_i=1\}$ 发生，表示在袋中取球 n 次，每次任取一只作不放回抽样时，第 i 号白球被取到. 因为事件 $\{X_i=1\}$ 可以在第 1 次、第 2 次、…、第 n 次取球，这 n 种两两互不相容的情况发生，且每次取到第 i 号白球的概率都是 $\frac{1}{N}$. 因此，

$$P\{X_i=1\}=\frac{1}{N}+\frac{1}{N}+\cdots+\frac{1}{N}=\frac{n}{N}, \quad i=1,2,\cdots,r.$$

这样
$$E(X_i)=\frac{n}{N},$$
从而
$$E(Y)=\sum_{i=1}^{r}E(X_i)=\frac{nr}{N}.$$

39. 抛一颗骰子直到所有点数全部出现为止，求所需抛掷次数 Y 的数学期望.

解 引入随机变量 $X_i, i=1,2,3,4,5,6$，如下：

$X_1=1$，

$X_2=$第一个点数得到后，等待第二个不同点数所需的等待次数，

$X_3=$第一、第二两点数得到后，等待第三个不同点数所需的等待次数，X_4,X_5,X_6 的意义类似. 则所需抛掷的总次数为
$$Y=X_1+X_2+\cdots+X_6.$$
因第一个点数得到后，掷一次得第二个不相同点数的概率为$\frac{5}{6}$，因此 X_2 的分布律为
$$P\{X_2=k\}=\frac{5}{6}\left(\frac{1}{6}\right)^{k-1},\quad k=1,2,\cdots.$$
即 X_2 服从参数 $p=\frac{5}{6}$的几何分布. 又因得到两个不相同的点数后，掷一次得第三个不相同点数的概率为$\frac{4}{6}$，故 X_3 服从参数$p=\frac{4}{6}$的几何分布，其分布律为
$$P\{X_3=k\}=\frac{4}{6}\left(\frac{2}{6}\right)^{k-1},\quad k=1,2,\cdots.$$
同样，X_4,X_5,X_6 的分布律分别为
$$P\{X_4=k\}=\frac{3}{6}\left(\frac{3}{6}\right)^{k-1},\quad k=1,2,\cdots,$$
$$P\{X_5=k\}=\frac{2}{6}\left(\frac{4}{6}\right)^{k-1},\quad k=1,2,\cdots,$$
$$P\{X_6=k\}=\frac{1}{6}\left(\frac{5}{6}\right)^{k-1},\quad k=1,2,\cdots.$$
因几何分布 $P\{X=k\}=p(1-p)^{k-1},k=1,2,\cdots$的数学期望为（参见本书第四章习题第 20 题）
$$E(X)=\frac{1}{p}.$$
所以
$$E(Y)=E\left(\sum_{i=1}^{6}X_i\right)=E(X_1)+\sum_{i=2}^{6}E(X_i)$$
$$=1+\left(\frac{6}{5}+\frac{6}{4}+\frac{6}{3}+\frac{6}{2}+\frac{6}{1}\right)=14.7.$$

40. 设随机变量 X,Y 相互独立，且 X,Y 分别服从以 $\frac{1}{\alpha},\frac{1}{\beta}$ 为均值的指数分布. 求 $E(X^2+Ye^{-X})$.

解
$$\begin{aligned}E(X^2+Ye^{-X})&=E(X^2)+E(Y)E(e^{-X})\\&=D(X)+[E(X)]^2+E(Y)\cdot\int_0^\infty e^{-t}\cdot\alpha e^{-\alpha t}dt\\&=\frac{1}{\alpha^2}+\frac{1}{\alpha^2}+\frac{1}{\beta}\int_0^\infty \alpha e^{-(\alpha+1)t}dt\\&=\frac{2}{\alpha^2}+\frac{\alpha}{\beta(\alpha+1)}.\end{aligned}$$

41. 一酒吧间柜台前有 6 张凳子，服务员预测，若两个陌生人进来就座的话，他们之间至少相隔两张凳子.

(1) 若真有两个陌生人入内，他们随机地就座，问服务员预言为真的概率是多少?

(2) 设两位顾客是随机就座的，求顾客之间凳子数的数学期望.

解 (1) 将凳子按自左至右编号，设服务员预言为真. (A) 若第一位顾客就座于 1 号，则另一位顾客可坐 4 或 5 或 6 号共三种坐法，(B) 若第一位顾客就座于 2 号，则另一位顾客可坐在 5 或 6 号共两种坐法，(C) 若第一位顾客就座于 3 号，则另一位顾客只可就座于 6 号，只有一种坐法. 综合(A)、(B)、(C)三种情况共计 6 种坐法. 同样，若第一位顾客分别就座于 6 号、5 号、4 号，则另一位顾客也有 6 种坐法，因此两人共有 $6\times2=12$ 种坐法，若两人随机就座共有 $A_6^2=30$ 种坐法，故服务员预言为真的概率是

$$p=\frac{12}{30}=\frac{2}{5}.$$

(2) 若两位顾客是随机就座的，以 Y 记两顾客间的凳子数，则 Y 可能取的值为 0,1,2,3,4. 可知 Y 的分布律为

Y	0	1	2	3	4
p_k	$\frac{5}{15}$	$\frac{4}{15}$	$\frac{3}{15}$	$\frac{2}{15}$	$\frac{1}{15}$

于是 $$E(Y)=0\times\frac{5}{15}+1\times\frac{4}{15}+2\times\frac{3}{15}+3\times\frac{2}{15}+4\times\frac{1}{15}=\frac{4}{3}.$$

42. 设随机变量 $X_1,X_2,\cdots,X_{100}$ 相互独立，且都服从 $U(0,1)$，又设 $Y=X_1\cdot X_2\cdot\cdots\cdot X_{100}$，求概率 $P\{Y<10^{-40}\}$ 的近似值.

解 所求概率为

$$p=P\{Y<10^{-40}\}=P\{\ln Y<-40\ln 10\}$$
$$=P\left\{\ln\prod_{i=1}^{100}X_i<-40\ln 10\right\}=P\left\{\sum_{i=1}^{100}\ln X_i<-92.1\right\}.$$

因 $X_1,X_2,\cdots,X_{100}$ 相互独立且都服从 $U(0,1)$ 分布，知 $\ln X_1,\ln X_2,\cdots,\ln X_{100}$ 也相互独立，且服从同一分布，又 $X_i\sim U(0,1)$，其概率密度为

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

故有 $E(\ln X_i)=\int_0^1\ln x\mathrm{d}x=-1$，$E(\ln^2 X_i)=\int_0^1\ln^2 x\mathrm{d}x=2$，$D(\ln X_i)=2-1=1$. 由中心极限定理得

$$p=P\left\{\sum_{i=1}^{100}\ln X_i<-92.1\right\}$$
$$=P\left\{\frac{\sum_{i=1}^{100}\ln X_i-100\times(-1)}{\sqrt{100}\sqrt{1}}<\frac{-92.1+100}{\sqrt{100}\sqrt{1}}\right\}$$
$$\approx\Phi\left(\frac{-92.1+100}{\sqrt{100}}\right)=\Phi(0.79)=0.7852.$$

43. 来自某个城市的长途电话呼唤的持续时间 X(以 min 计)是一个随机变量，它的分布函数是

$$F(x)=\begin{cases}1-\dfrac{1}{2}\mathrm{e}^{-\frac{x}{3}}-\dfrac{1}{2}\mathrm{e}^{-\left[\frac{x}{3}\right]}, & x\geqslant 0,\\ 0, & x<0\end{cases}$$

(其中 $\left[\dfrac{x}{3}\right]$ 是不大于 $\dfrac{x}{3}$ 的最大整数).

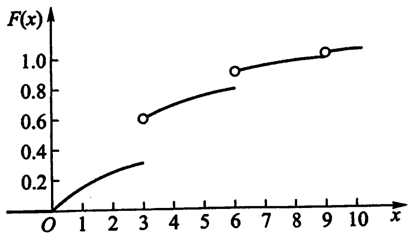
(1) 画出 $F(x)$ 的图形.

(2) 说明 X 是什么类型的随机变量.

(3) 求 $P\{X=4\}$，$P\{X=3\}$，$P\{X<4\}$，$P\{X>6\}$.（提示：$P\{X=a\}=F(a)-F(a-0)$.）

解 (1) $F(x)$ 的图形如题 16.43 图所示.

(2) $F(x)$ 的所有不连续点为 $3k(k=1,2,\cdots)$，X 取这些值的概率的总和为



题 16.43 图

$$\sum_{k=1}^{\infty}P\{X=3k\}=\sum_{k=1}^{\infty}[F(3k)-F(3k-0)]$$

$$=\sum_{k=1}^{\infty}\left\{\left(1-\frac{1}{2}e^{-\frac{3k}{3}}-\frac{1}{2}e^{-\frac{3k}{3}}\right)-\left[1-\frac{1}{2}e^{-\frac{3k}{3}}-\frac{1}{2}e^{-\left(\frac{3k}{3}-1\right)}\right]\right\}$$

$$=\sum_{k=1}^{\infty}\frac{1}{2}(e^{1-k}-e^{-k})=\frac{1}{2}(e-1)\sum_{k=1}^{\infty}e^{-k}=\frac{1}{2}.$$

注意到，在 $F(x)$ 的任一连续点 a 处有 $P\{X=a\}=0$；又由于 $\sum_{k=1}^{\infty}P\{X=3k\}=\frac{1}{2}$，因此，不可能取到可列多个值 $x_1,x_2,\cdots$，使得 $\sum_{k=1}^{\infty}P\{X=x_k\}=1$，故 X 不是离散型随机变量. 又由于 $F(x)$ 不是连续函数，故 X 也不是连续型随机变量.

(3) $P\{X=4\}=0$.

$$P\{X=3\}=F(3)-F(3-0)$$

$$=\left(1-\frac{1}{2}e^{-1}-\frac{1}{2}e^{-1}\right)-\left[1-\frac{1}{2}e^{-1}-\frac{1}{2}e^{-(1-1)}\right]$$

$$=\frac{1}{2}(1-e^{-1})=0.316.$$

$$P\{X<4\}=F(4)-P\{X=4\}=1-\frac{1}{2}e^{-4/3}-\frac{1}{2}e^{-1}-0=0.684.$$

$$P\{X>6\}=1-F(6)=1-\left(1-\frac{1}{2}e^{-2}-\frac{1}{2}e^{-2}\right)=e^{-2}=0.135.$$

44. 一汽车保险公司分析一组(250 人)签约的客户中的赔付情况. 据历史数据分析，在未来的一周中一组客户中至少提出一项索赔的客户数 X 占 10%. 写出 X 的分布，并求 $X>250\times0.12$ (即 $X>30$)的概率. 设各客户是否提出索赔相互独立.

解 按题意知 $X\sim b(250,0.10)$. 现在需要求

$$P\{X>30\}=\sum_{x=31}^{250}\binom{250}{x}\times0.10^{x}\times0.90^{250-x},$$

即需求
$$P\{X>30\}=1-\sum_{x=0}^{30}\binom{250}{x}\times 0.10^{x}\times 0.90^{250-x}.$$
由拉普拉斯定理得
$$P\{X>30\}\approx 1-\Phi\left(\frac{30-250\times 0.10}{\sqrt{250\times 0.10\times 0.90}}\right)$$
$$=1-\Phi(1.054)=1-0.853\ 1=0.146\ 9.$$

45. 在区间$(0,1)$随机地取一点 X. 定义 $Y=\min\{X,0.75\}$.

(1) 求随机变量 Y 的值域.

(2) 求 Y 的分布函数,并画出它的图形.

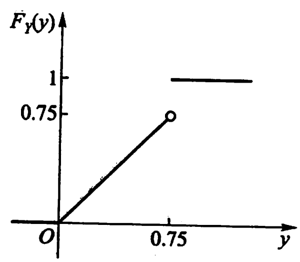
(3) 说明 Y 不是连续型随机变量,Y 也不是离散型随机变量.

解 (1) 因 $Y=\min\{X,0.75\}$,故 $Y\leqslant X$ 且 $Y\leqslant 0.75$. 又由于 X 的值域是$(0,1)$,知 Y 的值域为$(0,0.75]$.

(2) 由(1)知当 $y<0$ 时 $F_Y(y)=P\{Y\leqslant y\}=0$,当 $y\geqslant 0.75$ 时,$F_Y(y)=P\{Y\leqslant y\}=1$.

当 $0\leqslant y<0.75$ 时,事件$\{Y\leqslant y\}$表示 X 是在$(0,y]$随机取的一点. 故有
$$F_Y(y)=\begin{cases}0, & y<0,\\ y, & 0\leqslant y<0.75,\\ 1, & y\geqslant 0.75.\end{cases}$$
$F_Y(y)$的图形如题 16.45 图所示.

(3) 从题 16.45 图看出,$F_Y(y)$在点 $y=0.75$ 处不连续,故它不是连续型随机变量. $F_Y(y)$只有一个不连续点 $y=0.75$. 注意到在 $F_Y(y)$的任一连续点 a 处,有 $P\{Y=a\}=0$,而在不连续点 $y=0.75$处,$P\{Y=0.75\}=F_Y(0.75)-F_Y(0.75-0)=0.25$. 故不可能取到可列多个值 $y_1,y_2,\cdots$,使得 $\sum\limits_{k=1}^{\infty}P\{Y=y_k\}=1$,故 Y 不是离散型随机变量.



题 16.45 图

46. 设 X_1,X_2 是数学期望为 θ 的指数分布总体 X 的容量为 2 的样本,设 $Y=\sqrt{X_1X_2}$,试证明 $E\left(\frac{4Y}{\pi}\right)=\theta$.

证 X 的概率密度为
$$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-x/\theta}, & x>0,\\ 0, & \text{其他},\end{cases}\qquad \theta>0.$$

$$E(\sqrt{X})=\int_{-\infty}^{\infty}\sqrt{x}f(x)\mathrm{d}x=\int_{0}^{\infty}\frac{\sqrt{x}}{\theta}\mathrm{e}^{-x/\theta}\mathrm{d}x$$

$$\xlongequal{令 x/\theta=u}\int_{0}^{\infty}\sqrt{\theta}u^{\frac{1}{2}}\mathrm{e}^{-u}\mathrm{d}u=\sqrt{\theta}\Gamma\left(\frac{3}{2}\right)$$

$$=\sqrt{\theta}\,\frac{1}{2}\Gamma\left(\frac{1}{2}\right)=\frac{1}{2}\sqrt{\theta}\sqrt{\pi}=\frac{1}{2}\sqrt{\theta\pi}.$$

因 X_1,X_2 相互独立，从而 $\sqrt{X_1}$，$\sqrt{X_2}$ 相互独立，且

$$E(\sqrt{X_1})=E(\sqrt{X_2})=E(\sqrt{X}).$$

故由数学期望的性质，得

$$E\left(\frac{4Y}{\pi}\right)=\frac{4}{\pi}E(\sqrt{X_1}\sqrt{X_2})=\frac{4}{\pi}E(\sqrt{X_1})E(\sqrt{X_2})$$

$$=\frac{4}{\pi}\left(\frac{1}{2}\sqrt{\theta\pi}\right)^2=\theta.$$

47. 设总体 $X\sim N(\mu,\sigma^2)$，$X_1,X_2,\cdots,X_n$ 是一个样本. $\bar{X},S^2$ 分别为样本均值和样本方差，试证

$$E\left[(\bar{X}S^2)^2\right]=\left(\frac{\sigma^2}{n}+\mu^2\right)\left(\frac{2\sigma^4}{n-1}+\sigma^4\right).$$

证 因 $\bar{X},S^2$ 分别是正态总体 $N(\mu,\sigma^2)$ 的容量为 n 的样本均值和样本方差，于是 $\bar{X}$ 与 S^2 相互独立，故 $\bar{X}^2$ 与 $(S^2)^2$ 也相互独立. 从而

$$E\left[(\bar{X}S^2)^2\right]=E(\bar{X}^2)E[(S^2)^2]$$

$$=\left\{D(\bar{X})+[E(\bar{X})]^2\right\}\{D(S^2)+[E(S^2)]^2\},\qquad (*_1)$$

$$E(\bar{X})=\mu,\quad D(\bar{X})=\frac{\sigma^2}{n},$$

$$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1).$$

又由 χ^2 分布的性质知，

$$E\left[\frac{(n-1)S^2}{\sigma^2}\right]=n-1,\quad D\left[\frac{(n-1)S^2}{\sigma^2}\right]=2(n-1),$$

得

$$E(S^2)=\sigma^2,\quad D(S^2)=\frac{2\sigma^4}{n-1}.$$

将这些结果代入 $(*_1)$ 式，得

$$E\left[(\bar{X}S^2)^2\right]=\left(\frac{\sigma^2}{n}+\mu^2\right)\left(\frac{2\sigma^4}{n-1}+\sigma^4\right).$$

48. 设总体 X 具有概率密度

$$f(x)=\begin{cases}\dfrac{1}{\theta^2}xe^{-x/\theta}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

其中 $\theta>0$ 为未知参数，$X_1,X_2,\cdots,X_n$ 是来自 X 的样本，$x_1,x_2,\cdots,x_n$ 是相应的样本观察值.

(1) 求 θ 的最大似然估计量.

(2) 求 θ 的矩估计量.

(3) 问求得的估计量是否是无偏估计量?

解 (1) 由 X 的样本观察值 $x_1,x_2,\cdots,x_n$ 以及 X 的概率密度的形式，得似然函数为

$$L=L(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n}\frac{1}{\theta^2}x_i e^{-x_i/\theta}$$

$$=\frac{1}{\theta^{2n}}\left(\prod_{i=1}^{n}x_i\right)e^{-\sum\limits_{i=1}^{n}x_i/\theta},$$

故
$$\ln L=-2n\ln\theta+\ln\left(\prod_{i=1}^{n}x_i\right)-\frac{1}{\theta}\sum_{i=1}^{n}x_i,$$

令
$$\frac{d}{d\theta}\ln L=\frac{-2n}{\theta}+\frac{1}{\theta^2}\sum_{i=1}^{n}x_i=0,$$

得 θ 的最大似然估计值为

$$\hat{\theta}=\frac{1}{2n}\sum_{i=1}^{n}x_i=\frac{1}{2}\bar{x}.$$

θ 的最大似然估计量为

$$\hat{\theta}=\frac{1}{2}\bar{X}.$$

(2) $$\mu_1=E(X)=\int_{-\infty}^{\infty}xf(x)dx=\int_0^{\infty}\frac{1}{\theta^2}x^2e^{-x/\theta}dx$$

$$\xlongequal{\text{令 } x/\theta=u}\theta\int_0^{\infty}u^2e^{-u}du=\theta\Gamma(3)$$

$$=\theta\cdot 2\Gamma(2)=\theta\cdot 2\cdot 1\cdot\Gamma(1)=2\theta.$$

得
$$\theta=\frac{1}{2}\mu_1.$$

以 A_1 代替上式中的 μ_1，得 θ 的矩估计量为

$$\hat{\theta}=\frac{1}{2}A_1=\frac{1}{2}\bar{X},$$

它与最大似然估计量相一致.

(3) 因 $E(\hat{\theta})=\frac{1}{2}E(\bar{X})=\frac{1}{2}E(X)=\frac{1}{2}\times 2\theta=\theta$. 故 $\hat{\theta}=\frac{1}{2}\bar{X}$ 是 θ 的无偏估计量.

49. 设 $X_1,X_2,\cdots,X_{n_1}$ 以及 $Y_1,Y_2,\cdots,Y_{n_2}$ 为分别来自总体 $N(\mu_1,\sigma^2)$ 与 $N(\mu_2,\sigma^2)$ 的样本，且它们相互独立. μ_1,μ_2,σ^2 均未知，试求 μ_1,μ_2,σ^2 的最大似然估计量.

解　设给定的两独立样本的相应样本值分别为 $x_1,x_2,\cdots,x_{n_1};y_1,y_2,\cdots,y_{n_2}$，将它们代入相应的概率密度，然后相乘，得似然函数为

$$L(x_1,x_2,\cdots,x_{n_1},y_1,y_2,\cdots,y_{n_2};\mu_1,\mu_2,\sigma^2)$$

$$=\left[\prod_{i=1}^{n_1}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-(x_i-\mu_1)^2/(2\sigma^2)}\right]\left[\prod_{j=1}^{n_2}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-(y_j-\mu_2)^2/(2\sigma^2)}\right],$$

$$\ln L=(n_1+n_2)\ln\frac{1}{\sqrt{2\pi}}-\frac{n_1+n_2}{2}\ln\sigma^2-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n_1}(x_i-\mu_1)^2+\sum_{j=1}^{n_2}(y_j-\mu_2)^2\right],$$

令

$$\begin{cases}\dfrac{\partial\ln L}{\partial\mu_1}=\dfrac{1}{\sigma^2}\displaystyle\sum_{i=1}^{n_1}(x_i-\mu_1)=0,\\ \dfrac{\partial\ln L}{\partial\mu_2}=\dfrac{1}{\sigma^2}\displaystyle\sum_{j=1}^{n_2}(y_j-\mu_2)=0,\\ \dfrac{\partial\ln L}{\partial\sigma^2}=-\dfrac{n_1+n_2}{2}\dfrac{1}{\sigma^2}+\dfrac{1}{2(\sigma^2)^2}\left[\displaystyle\sum_{i=1}^{n_1}(x_i-\mu_1)^2+\sum_{j=1}^{n_2}(y_j-\mu_2)^2\right]=0,\end{cases}$$

得 μ_1,μ_2,σ^2 的最大似然估计值为

$$\hat{\mu}_1=\frac{1}{n_1}\sum_{i=1}^{n_1}x_i=\bar{x},\quad \hat{\mu}_2=\frac{1}{n_2}\sum_{j=1}^{n_2}y_j=\bar{y},$$

$$\hat{\sigma}^2=\frac{\sum_{i=1}^{n_1}(x_i-\bar{x})^2+\sum_{j=1}^{n_2}(y_j-\bar{y})^2}{n_1+n_2}.$$

μ_1,μ_2,σ^2 的最大似然估计量为

$$\hat{\mu}_1=\bar{X},\quad \hat{\mu}_2=\bar{Y},\quad \hat{\sigma}^2=\frac{\sum_{i=1}^{n_1}(X_i-\bar{X})^2+\sum_{j=1}^{n_2}(Y_j-\bar{Y})^2}{n_1+n_2}.$$

50. 为了研究一批贮存着的产品的可靠性，在产品投入贮存时，即在时刻 $t_0=0$ 时，随机地选定 n 件产品，然后在预先规定的时刻 $t_1,t_2,\cdots,t_k$ 取出来进行检测（检测时确定已失效的去掉，将未失效的继续投入贮存），今得到以下的寿命试验数据：

检测时刻(月)	t_1	t_2	…	t_k		
区间$(t_{i-1},t_i]$	$(0,t_1]$	$(t_1,t_2]$	…	$(t_{k-1},t_k]$	(t_k,∞)	
在$(t_{i-1},t_i]$的失效数	d_1	d_2	…	d_k	s	$\sum_{i=1}^{k}d_i+s=n$

这种数据称为区间数据. 设产品寿命 T 服从指数分布，其概率密度为

$$f(t)=\begin{cases}\lambda e^{-\lambda t}, & t>0,\\ 0, & \text{其他},\end{cases}\quad \lambda>0 \text{ 未知}.$$

(1) 试基于上述数据写出 λ 的对数似然方程.

(2) 设 $d_1<n, s<n$. 我们可以用数值解法求得 λ 的最大似然估计值，在计算机上计算是容易的. 特别，取检测时间是等间隔的，即取 $t_i=it_1, i=1,2,\cdots,k$. 验证，此时可得 λ 的最大似然估计为 $\hat{\lambda}=\dfrac{1}{t_1}\ln\left[1+\dfrac{n-s}{\sum\limits_{i=2}^{k}(i-1)d_i+sk}\right]$.

解　(1) 由假设产品寿命 T 服从指数分布，其分布函数为

$$F(t)=\begin{cases}1-e^{-\lambda t}, & t>0,\\ 0, & \text{其他}.\end{cases}$$

则产品在区间$(t_{i-1},t_i]$失效的概率为

$$P\{t_{i-1}<T\leqslant t_i\}=F(t_i)-F(t_{i-1})=e^{-\lambda t_{i-1}}-e^{-\lambda t_i},\quad i=1,2,\cdots,k,$$

这里 $t_0=0$. 产品直至 t_k 未失效的概率为

$$P\{T>t_k\}=1-F(t_k)=e^{-\lambda t_k}.$$

因而事件"n 件产品分别在区间$(0,t_1],(t_1,t_2],\cdots,(t_{k-1},t_k]$失效 $d_1,d_2,\cdots,d_k$ 件，而直至 t_k 还有 s 件未失效"的概率为

$$\begin{aligned}L(\lambda)&=\left\{\prod_{i=1}^{k}[F(t_i)-F(t_{i-1})]^{d_i}\right\}[1-F(t_k)]^s\\&=\left[\prod_{i=1}^{k}(e^{-\lambda t_{i-1}}-e^{-\lambda t_i})^{d_i}\right](e^{-\lambda t_k})^s,\end{aligned}$$

这就是样本的似然函数. 取对数得

$$\ln L(\lambda)=\sum_{i=1}^{k}d_i\ln(e^{-\lambda t_{i-1}}-e^{-\lambda t_i})-st_k\lambda.$$

对数似然方程为$\dfrac{d}{d\lambda}\ln L(\lambda)=0$，即为

$$\frac{d}{d\lambda}\ln L(\lambda)=\sum_{i=1}^{k}\frac{d_i[e^{-\lambda t_{i-1}}(-t_{i-1})-e^{-\lambda t_i}(-t_i)]}{e^{-\lambda t_{i-1}}-e^{-\lambda t_i}}-st_k=0.$$

上式中间第一项即为

$$\sum_{i=1}^{k}\frac{d_i[t_i-t_{i-1}e^{\lambda(t_i-t_{i-1})}]}{e^{\lambda(t_i-t_{i-1})}-1}=\sum_{i=1}^{k}\frac{d_i(t_i-t_{i-1})-d_it_{i-1}[e^{\lambda(t_i-t_{i-1})}-1]}{e^{\lambda(t_i-t_{i-1})}-1}$$

$$=\sum_{i=1}^{k}\frac{d_i(t_i-t_{i-1})}{e^{\lambda(t_i-t_{i-1})}-1}-\sum_{i=2}^{k}d_it_{i-1},$$

于是得对数似然方程为

$$\sum_{i=1}^{k}\frac{d_i(t_i-t_{i-1})}{e^{\lambda(t_i-t_{i-1})}-1}-\sum_{i=2}^{k}d_it_{i-1}-st_k=0.$$

(2) 若 $t_i=it_1,i=1,2,\cdots,k$，则上述方程化成

$$\sum_{i=1}^{k}\frac{d_i}{e^{\lambda t_1}-1}-\sum_{i=2}^{k}(i-1)d_i-sk=0.$$

即
$$\frac{1}{e^{\lambda t_1}-1}\sum_{i=1}^{k}d_i-\sum_{i=2}^{k}(i-1)d_i-sk=0.$$

即
$$e^{\lambda t_1}=1+\frac{\sum_{i=1}^{k}d_i}{\sum_{i=2}^{k}(i-1)d_i+sk},$$

即
$$e^{\lambda t_1}=1+\frac{n-s}{\sum_{i=2}^{k}(i-1)d_i+sk}.$$

从而得 λ 的最大似然估计为

$$\hat{\lambda}=\frac{1}{t_1}\ln\left[1+\frac{n-s}{\sum_{i=2}^{k}(i-1)d_i+sk}\right].$$

51. 设某种电子器件的寿命 T(以 h 计)服从指数分布，概率密度为

$$f(t)=\begin{cases}\lambda e^{-\lambda t}, & t>0,\\ 0, & \text{其他},\end{cases}$$

其中 $\lambda>0$ 未知. 从这批器件中任取 n 只在时刻 $t=0$ 时投入独立寿命试验. 试验进行到预定时间 T_0 结束. 此时，有 k $(0<k<n)$ 只器件失效，试求 λ 的最大似然估计.

解　考虑事件 $A=$“试验直至时间 T_0 为止，有 k 只器件失效，而有 $n-k$ 只未失效”的概率. 记 T 的分布函数为 $F(t)$，

$$F(t)=\begin{cases}1-e^{-\lambda t}, & t>0,\\ 0, & \text{其他}.\end{cases}$$

一只器件在 $t=0$ 时投入试验，在时间 T_0 以前失效的概率为 $P\{T\leqslant T_0\}=F(T_0)=1-e^{-\lambda T_0}$；而在时间 T_0 未失效的概率为 $P\{T>T_0\}=1-F(T_0)=e^{-\lambda T_0}$. 由于各

只器件的试验是相互独立的，因此，事件 A 的概率为

$$L(\lambda)=\binom{n}{k}(1-\mathrm{e}^{-\lambda T_0})^k(\mathrm{e}^{-\lambda T_0})^{n-k},$$

这就是所求的似然函数. 取对数得

$$\ln L(\lambda)=\ln\binom{n}{k}+k\ln(1-\mathrm{e}^{-\lambda T_0})+(n-k)(-\lambda T_0),$$

令
$$\frac{\mathrm{d}\ln L(\lambda)}{\mathrm{d}\lambda}=\frac{kT_0\mathrm{e}^{-\lambda T_0}}{1-\mathrm{e}^{-\lambda T_0}}-(n-k)T_0=0,$$

得
$$n\mathrm{e}^{-\lambda T_0}=n-k.$$

解得 λ 的最大似然估计为

$$\hat{\lambda}=\frac{1}{T_0}\ln\frac{n}{n-k}.$$

52. 设系统由两个独立工作的成败型元件串联而成(成败型元件只有两种状态:正常工作或失效). 元件 1、元件 2 的可靠性分别为 p_1,p_2，它们均未知. 随机地取 N 个系统投入试验，当系统中至少有一个元件失效时系统失效，现得到以下的试验数据:n_1 表示仅元件 1 失效的系统数，n_2 表示仅元件 2 失效的系统数，n_{12} 表示元件 1、元件 2 至少有一个失效的系统数，s 表示未失效的系统数. $n_1+n_2+n_{12}+s=N$. 这里 n_{12} 为隐蔽数据，也就是只知系统失效，但不能知道是由元件 1 或元件 2 单独失效引起的，还是由元件 1,2 均失效引起的，设隐蔽与系统失效的真正原因独立.

(1) 试写出 p_1,p_2 的似然函数.

(2) 设有系统寿命试验数据 $N=20,n_1=5,n_2=3,n_{12}=1,s=11$. 试求 p_1,p_2 的最大似然估计值.

解 (1) 为了写出似然函数，现在来求取到现有样本的概率. 因共有 N 个系统，因而似然函数是 N 个因子的乘积，其中

对应于 n_1 个仅元件 1 失效的系统有 n_1 个因子:$[(1-p_1)p_2]^{n_1}$;

对应于 n_2 个仅元件 2 失效的系统有 n_2 个因子:$[(1-p_2)p_1]^{n_2}$;

对应于 n_{12} 个元件 1,2 至少有一个失效的系统有 n_{12} 个因子:$(1-p_1p_2)^{n_{12}}$;

对应于 s 个未失效的系统有 s 个因子:$(p_1p_2)^s$.

故得似然函数为

$$\begin{aligned}L(p_1,p_2)&=[(1-p_1)p_2]^{n_1}[(1-p_2)p_1]^{n_2}(1-p_1p_2)^{n_{12}}(p_1p_2)^s\\&=(1-p_1)^{n_1}(1-p_2)^{n_2}(1-p_1p_2)^{n_{12}}p_1{}^{n_2+s}p_2{}^{n_1+s}.\end{aligned}$$

(2) 以 $N=20,n_1=5,n_2=3,n_{12}=1,s=11$ 代入上式，得似然函数为

$$L(p_1,p_2)=(1-p_1)^5(1-p_2)^3(1-p_1p_2)p_1^{14}p_2^{16}.$$

$$\ln L(p_1,p_2)=5\ln(1-p_1)+3\ln(1-p_2)+\ln(1-p_1p_2)+14\ln p_1+16\ln p_2.$$

令

$$\begin{cases} \dfrac{\partial \ln L}{\partial p_1} = \dfrac{-5}{1-p_1} + \dfrac{14}{p_1} - \dfrac{p_2}{1-p_1p_2} = 0, & (*_1) \\ \dfrac{\partial \ln L}{\partial p_2} = \dfrac{-3}{1-p_2} + \dfrac{16}{p_2} - \dfrac{p_1}{1-p_1p_2} = 0, & (*_2) \end{cases}$$

$(*_1)\times p_1-(*_2)\times p_2$，得

$$\frac{14-19p_1}{1-p_1}=\frac{16-19p_2}{1-p_2},$$

即有

$$-5p_2=-3p_1-2,$$

$$p_2=\frac{3p_1+2}{5}.$$

将 p_2 代入$(*_1)$，经化简得

$$12p_1^3-p_1^2-25p_1+14=0,$$

或即

$$(p_1-1)(12p_1^2+11p_1-14)=0.$$

解得 $p_1=1$(不合理，舍去)，$p_1=0.715\ 0$，于是

$$p_2=\frac{1}{5}(2+3p_1)=0.829\ 0.$$

即得 p_1,p_2 的最大似然估计值为

$$\hat{p}_1=0.715\ 0,\quad \hat{p}_2=0.829\ 0.$$

53. (1) 设总体 X 具有分布律

X	1	2	3
p_k	θ	θ	$1-2\theta$

$\theta>0$ 未知，今有样本

1 1 1 3 2 1 3 2 2 1 2 2 3 1 1 2

试求 θ 的最大似然估计值和矩估计值.

(2) 设总体 X 服从 Γ 分布，其概率密度为

$$f(x)=\begin{cases} \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}, & x>0, \\ 0, & \text{其他}. \end{cases}$$

其形状参数 $\alpha>0$ 为已知，尺度参数 $\beta>0$ 未知. 今有样本值 $x_1,x_2,\cdots,x_n$，求 β 的最大似然估计值.

解 (1) $n=16$，样本值 1 出现 7 次，2 出现 6 次，3 出现 3 次，故似然函数为

$$L=\theta^7\cdot\theta^6\cdot(1-2\theta)^3=\theta^{13}(1-2\theta)^3.$$

$$\ln L=13\ln\theta+3\ln(1-2\theta).$$

令
$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L=\frac{13}{\theta}-\frac{6}{1-2\theta}=0,$$

得 θ 的最大似然估计值为

$$\hat{\theta}=\frac{13}{32}.$$

下面来求矩估计值，令

$$\mu_1=E(X)=1\cdot\theta+2\cdot\theta+3\cdot(1-2\theta)=3-3\theta,$$

解得
$$\theta=1-\frac{1}{3}\mu_1.$$

于是 θ 的矩估计值为

$$\hat{\theta}=1-\frac{1}{3}\bar{x}=1-\frac{1}{3}\times\frac{7}{4}=\frac{5}{12}.$$

(2) 对于样本值 $x_1,x_2,\cdots,x_n$，似然函数为

$$L=L(x_1,x_2,\cdots,x_n;\beta)=\prod_{i=1}^{n}\frac{1}{\beta^{\alpha}\Gamma(\alpha)}x_i^{\alpha-1}\mathrm{e}^{-x_i/\beta}$$
$$=\frac{1}{\beta^{n\alpha}[\Gamma(\alpha)]^n}\left(\prod_{i=1}^{n}x_i\right)^{\alpha-1}\mathrm{e}^{-\frac{1}{\beta}\sum\limits_{i=1}^{n}x_i}.$$

$$\ln L=-n\alpha\ln\beta-n\ln\Gamma(\alpha)+(\alpha-1)\ln\prod_{i=1}^{n}x_i-\frac{1}{\beta}\sum_{i=1}^{n}x_i.$$

令
$$\frac{\mathrm{d}\ln L}{\mathrm{d}\beta}=\frac{-n\alpha}{\beta}+\frac{1}{\beta^2}\sum_{i=1}^{n}x_i=0,$$

得 β 的最大似然估计值为

$$\hat{\beta}=\frac{1}{\alpha n}\sum_{i=1}^{n}x_i=\frac{\bar{x}}{\alpha}.$$

54. (1) 设随机变量 $Z=\ln X\sim N(\mu,\sigma^2)$，即 X 服从对数正态分布，验证

$$E(X)=\mathrm{e}^{\mu+\frac{1}{2}\sigma^2}.$$

(2) 设自(1)中总体 X 中取一容量为 n 的样本 $x_1,x_2,\cdots,x_n$，求 $E(X)$ 的最大似然估计. 此处设 μ,σ^2 均为未知.

(3) 已知在文学家萧伯纳的 *An Intelligent Woman's Guide to Socialism* 一书中，一个句子的单词数近似地服从对数正态分布，设 μ 及 σ^2 为未知. 今自该书中随机地取 20 个句子，这些句子中的单词数分别为

52	24	15	67	15	22	63	26	16	32
7	33	28	14	7	29	10	6	59	30

问这本书中，一个句子单词数均值的最大似然估计值等于多少？

解　(1) 由 $Z=\ln X$ 得 $X=e^Z$，而 $Z\sim N(\mu,\sigma^2)$，故

$$E(X)=E(e^Z)=\int_{-\infty}^{\infty}e^z\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(z-\mu)^2}{2\sigma^2}}dz$$

$$=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}e^{-\frac{(z-\mu)^2}{2\sigma^2}+z}dz.$$

由于

$$\frac{-(z-\mu)^2}{2\sigma^2}+z=\frac{-1}{2\sigma^2}(z^2-2\mu z+\mu^2-2\sigma^2 z)$$

$$=\frac{-1}{2\sigma^2}[z^2-2(\mu+\sigma^2)z+\mu^2]$$

$$=\frac{-1}{2\sigma^2}[(z-\mu-\sigma^2)^2+\mu^2-(\mu+\sigma^2)^2]$$

$$=\frac{-1}{2\sigma^2}(z-\mu-\sigma^2)^2+\mu+\frac{\sigma^2}{2},$$

故有

$$E(X)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}e^{-(z-\mu-\sigma^2)^2/(2\sigma^2)}\cdot e^{\mu+\sigma^2/2}dz$$

$$=e^{\mu+\sigma^2/2}\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}e^{-(z-\mu-\sigma^2)^2/(2\sigma^2)}dz=e^{\mu+\sigma^2/2}.$$

(2) 为求 $E(X)=e^{\mu+\sigma^2/2}$ 的最大似然估计，需先求 μ,σ^2 的最大似然估计. 为此，先来求 X 的概率密度.

$X=e^Z$，而 $Z\sim N(\mu,\sigma^2)$. 记 Z 的概率密度为 $f_Z(z)$，由于函数 $x=e^z$ 严格单调增加，其反函数为 $z=\ln x$，故由教材第二章公式(5.2)知，X 的概率密度为

$$f_X(x)=\begin{cases}f_Z(\ln x)\cdot|(\ln x)'|, & x>0,\\ 0, & \text{其他}\end{cases}$$

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

接着来求 μ 和 σ^2 的最大似然估计. 对于样本值 $x_1,x_2,\cdots,x_n$，似然函数为

$$L=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n\left(\prod_{i=1}^n\frac{1}{x_i}\right)\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n(\ln x_i-\mu)^2\right\}.$$

$$\ln L=n\ln\frac{1}{\sqrt{2\pi}}-\frac{n}{2}\ln\sigma^2+\ln\left(\prod_{i=1}^n\frac{1}{x_i}\right)-\frac{1}{2\sigma^2}\sum_{i=1}^n(\ln x_i-\mu)^2.$$

令

$$\begin{cases} \dfrac{\partial}{\partial \mu} \ln L = \dfrac{1}{\sigma^2}\sum_{i=1}^{n}(\ln x_i - \mu) = 0, \\ \dfrac{\partial}{\partial \sigma^2} \ln L = -\dfrac{n}{2\sigma^2} + \dfrac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(\ln x_i - \mu)^2 = 0, \end{cases}$$

得 μ,σ^2 的最大似然估计值分别为

$$\begin{cases} \hat{\mu} = \dfrac{1}{n}\sum_{i=1}^{n}\ln x_i, \\ \hat{\sigma^2} = \dfrac{1}{n}\sum_{i=1}^{n}(\ln x_i - \hat{\mu})^2. \end{cases}$$

注意到 $E(X)=e^{\mu+\sigma^2/2}$，由最大似然估计量的不变性知[①] $E(X)$的最大似然估计为

$$\begin{aligned} \widehat{E(X)} &= \exp\{\hat{\mu}+\hat{\sigma^2}/2\} \\ &= \exp\left\{\frac{1}{n}\sum_{i=1}^{n}\ln x_i + \frac{1}{2n}\sum_{i=1}^{n}\left(\ln x_i - \frac{1}{n}\sum_{j=1}^{n}\ln x_j\right)^2\right\}. \end{aligned}$$

(3) 将所给的 20 个数取对数，经计算得到 $\hat{\mu}$ 和 $\hat{\sigma^2}$ 分别为

$$\hat{\mu} = 3.089\,033, \quad \hat{\sigma^2} = 0.508\,131.$$

故
$$\widehat{E(X)} = \exp\left\{3.089\,033 + \frac{1}{2}\times 0.508\,131\right\} = 28.306\,7.$$

55. 考虑进行定数截尾寿命试验，假设将随机抽取的 n 件产品在时间 $t=0$ 时同时投入试验. 试验进行到 m 件($m<n$)产品失效时停止，m 件失效产品的失效时间分别为

---

① 详细地说有如下的结果. 定理：设函数 $\boldsymbol{u}=\boldsymbol{h}(\boldsymbol{\theta})$，$\boldsymbol{\theta}\in\Theta$ 有唯一的反函数，其中 $\boldsymbol{u}=\begin{bmatrix}u_1\\ \vdots\\ u_r\end{bmatrix}$，$\boldsymbol{h}(\boldsymbol{\theta})=\begin{bmatrix}h_1(\theta)\\ \vdots\\ h_r(\theta)\end{bmatrix}$，$\boldsymbol{\theta}=\begin{bmatrix}\theta_1\\ \vdots\\ \theta_r\end{bmatrix}$，则若 $\hat{\boldsymbol{\theta}}$ 是 $\boldsymbol{\theta}$ 的最大似然估计，那么 $\hat{\boldsymbol{u}}=\boldsymbol{h}(\hat{\boldsymbol{\theta}})$ 是 $\boldsymbol{h}(\boldsymbol{\theta})$ 的最大似然估计. 现在 $\boldsymbol{u}=\begin{bmatrix}a\\ b\end{bmatrix}=\begin{bmatrix}e^{\mu+\sigma^2/2}\\ \mu\end{bmatrix}$ 有唯一的反函数：$\begin{bmatrix}\mu\\ \sigma\end{bmatrix}=\begin{bmatrix}b\\ 2(\ln a-b)\end{bmatrix}$，故 $\boldsymbol{u}=\begin{bmatrix}a\\ b\end{bmatrix}$ 的最大似然估计为 $\begin{bmatrix}\hat{a}\\ \hat{b}\end{bmatrix}=\begin{bmatrix}\exp\{\hat{\mu}+\hat{\sigma^2}/2\}\\ \hat{\mu}\end{bmatrix}$，即 a 有最大似然估计 $\hat{a}=\exp\{\hat{\mu}+\hat{\sigma^2}/2\}$.

$$0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_m.$$

t_m 是第 m 件产品的失效时间. 设产品的寿命分布为韦布尔分布,其概率密度为

$$f(t)=\begin{cases}\dfrac{\beta}{\eta^\beta}t^{\beta-1}\mathrm{e}^{-\left(\frac{t}{\eta}\right)^\beta}, & t>0,\\ 0, & \text{其他}.\end{cases}$$

其中参数 $\beta>0$ 已知. 求参数 η 的最大似然估计.

解　$f(t)=\begin{cases}\dfrac{\beta}{\eta^\beta}t^{\beta-1}\mathrm{e}^{-\left(\frac{t}{\eta}\right)^\beta}, & t>0,\\ 0, & \text{其他},\end{cases}$,$\beta>0$ 已知,使用定数截尾数据,截尾数为 $m\ (<n)$. 则似然方程为

$$\begin{aligned}L &= \left[\binom{n}{m}\prod_{i=1}^{m}\frac{\beta}{\eta^\beta}t_i^{\beta-1}\mathrm{e}^{-\left(\frac{t_i}{\eta}\right)^\beta}\right]\mathrm{e}^{-\left(\frac{t_m}{\eta}\right)^{\beta}(n-m)}\\ &=\binom{n}{m}\left(\frac{\beta}{\eta^\beta}\right)^m\left(\prod_{i=1}^{m}t_i^{\beta-1}\right)\exp\left\{-\sum_{i=1}^{m}\left(\frac{t_i}{\eta}\right)^\beta-\left(\frac{t_m}{\eta}\right)^\beta(n-m)\right\}.\end{aligned}$$

$$\begin{aligned}\ln L &= \ln\left[\binom{n}{m}\beta^m\prod_{i=1}^{m}t_i^{\beta-1}\right]-m\beta\ln\eta-\frac{1}{\eta^\beta}\left[\sum_{i=1}^{m}t_i^\beta+(n-m)t_m^\beta\right]\\ &=\ln\left[\binom{n}{m}\beta^m\prod_{i=1}^{m}t_i^{\beta-1}\right]-m\beta\ln\eta-\frac{T_m}{\eta^\beta},\end{aligned}$$

其中 $T_m=\sum\limits_{i=1}^{m}t_i^\beta+(n-m)t_m^\beta$. 令

$$\frac{\mathrm{d}\ln L}{\mathrm{d}\eta}=-\frac{m\beta}{\eta}+\frac{\beta T_m}{\eta^{\beta+1}}=0.$$

得

$$\eta^\beta=\frac{T_m}{m},$$

于是得 η 的最大似然估计为

$$\hat{\eta}=\left(\frac{T_m}{m}\right)^{1/\beta},$$

其中 $T_m=\sum\limits_{i=1}^{m}t_i^\beta+(n-m)t_m^\beta$.

56. 设某大城市郊区的一条林荫道两旁开设了许多小商店,这些商店的开设延续时间(以月计)是一个随机变量,现随机地取 30 家商店,将它们的延续时间按自小到大排序,选其中前 8 家商店,它们的延续时间分别是

3.2　3.9　5.9　6.5　16.5　20.3　40.4　50.9

假设商店开设延续时间的长度是韦布尔随机变量.其概率密度为

$$f(x)=\begin{cases}\dfrac{\beta}{\eta^{\beta}}x^{\beta-1}e^{-\left(\frac{x}{\eta}\right)^{\beta}}, & x>0,\\ 0, & \text{其他},\end{cases}$$

其中 $\beta=0.8$.

(1) 试用上题结果,写出 η 的最大似然估计.

(2) 按(1)的结果求商店开设延续时间至少为 2 年的概率的估计.

解 (1) 现在 $n=30, m=8, \beta=0.8, t_m=50.9$,

$$\begin{aligned}T_m &= \sum_{i=1}^{8} t_i^{0.8}+22\times 50.9^{0.8}\\ &=2.536+2.971+4.137+4.470+9.419+11.117\\ &\quad +19.280+23.194+22\times 50.9^{0.8}\\ &=77.124+22\times 23.194=587.392.\end{aligned}$$

$$\hat{\eta}=\left(\frac{T_m}{m}\right)^{1/0.8}=\left(\frac{587.392}{8}\right)^{1/0.8}=214.930.$$

(2) 2 年＝24 月.韦布尔分布的分布函数为

$$F(x)=\begin{cases}1-e^{-\left(\frac{x}{\eta}\right)^{\beta}}, & x>0,\\ 0, & \text{其他}.\end{cases}$$

于是商店开设延续时间至少为 24 个月的概率为

$$p=P\{X>24\}=1-F(24)=e^{-\left(\frac{x}{\eta}\right)^{\beta}}\Big|_{x=24}=e^{-(24/214.930)^{0.8}}=0.841.$$

57. 设分别自总体 $N(\mu_1,\sigma^2)$ 和 $N(\mu_2,\sigma^2)$ 中抽取容量为 n_1,n_2 的两独立样本.其样本方差分别为 S_1^2,S_2^2.试证,对于任意常数 a,b $(a+b=1)$, $Z=aS_1^2+bS_2^2$ 都是 σ^2 的无偏估计,并确定常数 a,b,使 $D(Z)$ 达到最小.

解 因 $E(S_1^2)=\sigma^2, E(S_2^2)=\sigma^2$,且 $a+b=1$,于是

$$\begin{aligned}E(aS_1^2+bS_2^2)&=aE(S_1^2)+bE(S_2^2)\\ &=a\sigma^2+b\sigma^2=\sigma^2(a+b)=\sigma^2.\end{aligned}$$

故对于任意常数 a,b,只要 $a+b=1$, $Z=aS_1^2+bS_2^2$ 都是 σ^2 的无偏估计.

因 S_1^2 与 S_2^2 相互独立,故

$$\begin{aligned}D(Z)&=D(aS_1^2+bS_2^2)=D(aS_1^2)+D(bS_2^2)\\ &=D\left[\frac{a\sigma^2}{n_1-1}\frac{(n_1-1)S_1^2}{\sigma^2}\right]+D\left[\frac{b\sigma^2}{n_2-1}\frac{(n_2-1)S_2^2}{\sigma^2}\right]\\ &=\left(\frac{a\sigma^2}{n_1-1}\right)^2 D\left[\frac{(n_1-1)S_1^2}{\sigma^2}\right]+\left(\frac{b\sigma^2}{n_2-1}\right)^2 D\left[\frac{(n_2-1)S_2^2}{\sigma^2}\right]\end{aligned}$$

$$=\frac{a^2\sigma^4}{(n_1-1)^2}\cdot 2(n_1-1)+\frac{b^2\sigma^4}{(n_2-1)^2}\cdot 2(n_2-1)$$

$$=2\sigma^4\left(\frac{a^2}{n_1-1}+\frac{b^2}{n_2-1}\right).$$

$\Bigg($因为正态总体 $N(\mu,\sigma^2)$ 的样本方差 $S^2=\frac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\bar{X})^2$ 有 $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$，而$\chi^2(n-1)$分布的方差为$2(n-1)$. $\Bigg)$

记 $F(a,b)=\frac{a^2}{n_1-1}+\frac{b^2}{n_2-1}$，由 $a+b=1$，$F(a,b)$可化成

$$F(a,1-a)=\frac{a^2}{n_1-1}+\frac{(1-a)^2}{n_2-1}.$$

令
$$\frac{\mathrm{d}}{\mathrm{d}a}F(a,1-a)=\frac{2a}{n_1-1}-\frac{2(1-a)}{n_2-1}=0,$$

得
$$a=\frac{n_1-1}{n_1+n_2-2}\ \left(\text{此时 } b=1-a=\frac{n_2-1}{n_1+n_2-2}\right).$$

而
$$\frac{\mathrm{d}^2}{\mathrm{d}a^2}F(a,1-a)=\frac{2}{n_1-1}+\frac{2}{n_2-1}>0,$$

即知当 $a=\frac{n_1-1}{n_1+n_2-2}$时 $F(a,1-a)$取最小值. 这就是说，当 $a=\frac{n_1-1}{n_1+n_2-2}$，$b=\frac{n_2-1}{n_1+n_2-2}$时 $D(Z)$取到最小值.

58. 设总体 $X\sim N(\mu,\sigma^2)$，$X_1,X_2,\cdots,X_n$ 是来自 X 的样本. 已知样本方差 $S^2=\frac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\bar{X})^2$ 是 σ^2 的无偏估计. 验证样本标准差 S 不是标准差 σ 的无偏估计.

证 因 $Y=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$，而 $S=\frac{\sigma}{\sqrt{n-1}}\sqrt{Y}$是 Y 的函数，故

$$E(S)=E\left(\frac{\sigma}{\sqrt{n-1}}Y^{\frac{1}{2}}\right)=\int_0^\infty\frac{\sigma}{\sqrt{n-1}}y^{\frac{1}{2}}f_{\chi^2(n-1)}(y)\mathrm{d}y,$$

其中 $f_{\chi^2(n-1)}(y)$是$\chi^2(n-1)$分布的概率密度. 将 $f_{\chi^2(n-1)}(y)$的表达式代入上式右边，得到

$$E(S)=\frac{\sigma}{\sqrt{n-1}}\int_0^\infty y^{\frac{1}{2}}\frac{1}{2^{(n-1)/2}\Gamma\left(\frac{n-1}{2}\right)}y^{(n-1)/2-1}\mathrm{e}^{-y/2}\mathrm{d}y$$

$$=\frac{\sigma}{\sqrt{n-1}\,\Gamma\left(\frac{n-1}{2}\right)2^{(n-1)/2}}\int_0^\infty y^{n/2-1}\mathrm{e}^{-y/2}\mathrm{d}y$$

$$\xlongequal{令 y/2=t} \frac{\sqrt{2}\sigma}{\sqrt{n-1}\Gamma\left(\frac{n-1}{2}\right)}\int_0^\infty t^{n/2-1}e^{-t}dt$$

$$=\sqrt{\frac{2}{n-1}}\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\sigma \quad \left(因\int_0^\infty t^{n/2-1}e^{-t}dt=\Gamma\left(\frac{n}{2}\right)\right).$$

因此，$E(S)\neq\sigma$，即 S 不是 σ 的无偏估计.

59. 设总体 X 服从指数分布，其概率密度为

$$f(x)=\begin{cases}\frac{1}{\theta}e^{-x/\theta}, & x>0,\\ 0, & 其他.\end{cases}$$

$\theta>0$ 未知. 从总体中抽取一容量为 n 的样本 $X_1,X_2,\cdots,X_n$.

(1) 证明$\frac{2n\bar{X}}{\theta}\sim\chi^2(2n)$.

(2) 求 θ 的置信水平为 $1-\alpha$ 的单侧置信下限.

(3) 某种元件的寿命(以 h 计)服从上述指数分布，现从中抽得一容量$n=16$的样本，测得样本均值为 5 010 h，试求元件的平均寿命的置信水平为 0.90 的单侧置信下限.

解 (1) 令 $Z=\frac{2X}{\theta}$，因 $z=\frac{2x}{\theta}$为严格单调函数，其反函数为 $x=\frac{\theta}{2}z$，故由教材第二章(5.2)式，知 Z 的概率密度为

$$f_Z(z)=\begin{cases}f\left(\frac{\theta}{2}z\right)\left|\left(\frac{\theta}{2}z\right)'\right|=\frac{1}{2}e^{-z/2}, & z>0,\\ 0, & 其他.\end{cases}$$

它是$\chi^2(2)$分布的概率密度. 也就是说

$$\frac{2X}{\theta}\sim\chi^2(2).$$

$X_1,X_2,\cdots,X_n$ 是来自 X 的样本，因此 $X_1,X_2,\cdots,X_n$ 相互独立，且都与 X 有相同的分布. 这样就有

$$\frac{2X_i}{\theta}\sim\chi^2(2),\quad i=1,2,\cdots,n.$$

再由χ^2分布的可加性，得

$$\frac{2n\bar{X}}{\theta}=\sum_{i=1}^n\frac{2X_i}{\theta}\sim\chi^2(2n).$$

(2) 因
$$P\left\{\frac{2n\overline{X}}{\theta}<\chi_{\alpha}^{2}(2n)\right\}=1-\alpha,$$
即有
$$P\left\{\frac{2n\overline{X}}{\chi_{\alpha}^{2}(2n)}<\theta\right\}=1-\alpha,$$
故 θ 的置信水平为 $1-\alpha$ 的单侧置信下限是 $\underline{\theta}=\dfrac{2n\overline{X}}{\chi_{\alpha}^{2}(2n)}$.

(3) 今 $n=16,\bar{x}=5\ 010,1-\alpha=0.90,\alpha=0.10,\chi_{\alpha}^{2}(2n)=\chi_{0.10}^{2}(32)=42.585$.

故
$$\underline{\theta}=2\times16\times\frac{5\ 010}{42.585}=3\ 764.7.$$

60. 设总体 $X\sim U(0,\theta)$，$X_1,X_2,\cdots,X_n$ 是来自 X 的样本.

(1) 验证 $Y=\max\{X_1,X_2,\cdots,X_n\}$的分布函数为
$$F_Y(y)=\begin{cases}0, & y<0,\\ \dfrac{y^n}{\theta^n}, & 0\leqslant y<\theta,\\ 1, & y\geqslant\theta.\end{cases}$$

(2) 验证 $U=\dfrac{Y}{\theta}$的概率密度为
$$f_U(u)=\begin{cases}nu^{n-1}, & 0<u<1,\\ 0, & \text{其他}.\end{cases}$$

(3) 给定正数 α，$0<\alpha<1$，求 U 的分布的上$\dfrac{\alpha}{2}$分位数 $h_{\alpha/2}$ 以及上 $1-\dfrac{\alpha}{2}$分位数 $h_{1-\alpha/2}$.

(4) 利用(2)，(3)求参数 θ 的置信水平为 $1-\alpha$ 的置信区间.

(5) 设某人上班的等车时间 $X\sim U(0,\theta)$，θ 未知. 现在有样本 $x_1=4.2,x_2=3.5,x_3=1.7,x_4=1.2,x_5=2.4$，求 θ 的置信水平为 0.95 的置信区间.

解 (1) $X\sim U(0,\theta)$，它的分布函数为
$$F(x)=\begin{cases}0, & x<0,\\ \dfrac{x}{\theta}, & 0\leqslant x<\theta,\\ 1, & x\geqslant\theta.\end{cases}$$
因 $X_1,X_2,\cdots,X_n$ 是来自 X 的样本，故它们相互独立，且都与 X 有相同的分布. 从而 $Y=\max\{X_1,X_2,\cdots,X_n\}$的分布函数为

$$F_Y(y)=[F(y)]^n=\begin{cases}0, & y<0,\\ \dfrac{y^n}{\theta^n}, & 0\leqslant y<\theta,\\ 1, & y\geqslant\theta.\end{cases}$$

(2) $U=\dfrac{Y}{\theta}$的分布函数为

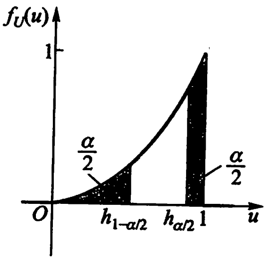
$$F_U(u)=P\{U\leqslant u\}=P\{Y\leqslant\theta u\}=F_Y(\theta u)$$
$$=\begin{cases}0, & u<0,\\ u^n, & 0\leqslant u<1,\\ 1, & u\geqslant 1.\end{cases}$$

将 $F_U(u)$关于 u 求导得 U 的概率密度为

$$f_U(u)=\begin{cases}nu^{n-1}, & 0<u<1,\\ 0, & \text{其他}.\end{cases}$$

(3) 如题 16.60 图，上$\dfrac{\alpha}{2}$分位数 $h_{\alpha/2}$应满足

$$P\{U>h_{\alpha/2}\}=\frac{\alpha}{2}.$$



题 16.60 图

即
$$1-F_U(h_{\alpha/2})=\frac{\alpha}{2},$$

或即
$$1-(h_{\alpha/2})^n=\frac{\alpha}{2},$$

因而
$$h_{\alpha/2}=\left(1-\frac{\alpha}{2}\right)^{\frac{1}{n}}.$$

同样，$h_{1-\alpha/2}$应满足

$$P\{U\leqslant h_{1-\alpha/2}\}=\frac{\alpha}{2},$$

即
$$F_U(h_{1-\alpha/2})=\frac{\alpha}{2},$$

或即
$$(h_{1-\alpha/2})^n=\frac{\alpha}{2},$$

因而
$$h_{1-\alpha/2}=\left(\frac{\alpha}{2}\right)^{\frac{1}{n}}.$$

(4) 考虑到

$$P\{h_{1-\alpha/2}<U<h_{\alpha/2}\}=1-\alpha,$$

即
$$P\left\{h_{1-\alpha/2}<\frac{Y}{\theta}<h_{\alpha/2}\right\}=1-\alpha,$$

因此
$$P\left\{\frac{Y}{h_{\alpha/2}}<\theta<\frac{Y}{h_{1-\alpha/2}}\right\}=1-\alpha.$$

即

$$P\left\{\frac{\max\{X_1,X_2,\cdots,X_n\}}{\left(1-\frac{\alpha}{2}\right)^{\frac{1}{n}}}<\theta<\frac{\max\{X_1,X_2,\cdots,X_n\}}{\left(\frac{\alpha}{2}\right)^{\frac{1}{n}}}\right\}=1-\alpha.$$

故得总体 $X\sim U(0,\theta)$的未知参数 θ 的置信水平为 $1-\alpha$ 的一个置信区间为

$$\left(\left(1-\frac{\alpha}{2}\right)^{-\frac{1}{n}}\max\{X_1,X_2,\cdots,X_n\},\quad \left(\frac{\alpha}{2}\right)^{-\frac{1}{n}}\max\{X_1,X_2,\cdots,X_n\}\right).$$

(5) 现在 $n=5$，$\max\{x_1,x_2,x_3,x_4,x_5\}=4.2$，$1-\frac{\alpha}{2}=0.975$，$\frac{\alpha}{2}=0.025$，所求 θ 的置信水平为 0.95 的置信区间为

$$(0.975^{-\frac{1}{5}}\times 4.2,\quad 0.025^{-\frac{1}{5}}\times 4.2)=(4.22,\ 8.78).$$

61. 设总体 X 服从指数分布，概率密度为

$$f(x)=\begin{cases}\frac{1}{\theta}e^{-\frac{x}{\theta}}, & x>0,\\ 0, & \text{其他},\end{cases}\quad \theta>0.$$

设 $X_1,X_2,\cdots,X_n$ 是来自 X 的一个样本. 试取第 59 题中当 $\theta=\theta_0$ 时的统计量 $\chi^2=\frac{2n\overline{X}}{\theta_0}$作为检验统计量，检验假设 $H_0:\theta=\theta_0$，$H_1:\theta\neq\theta_0$. 取显著性水平为 α (注意：$E(\overline{X})=\theta$).

设某种电子元件的寿命(以 h 计)服从均值为 θ 的指数分布，随机取 12 只元件测得它们的寿命分别为

340 430 560 920 1 380 1 520 1 660 1 770 2 100 2 320 2 350 2 650

试取显著性水平 $\alpha=0.05$，检验假设 $H_0:\theta=1\,450$，$H_1:\theta\neq 1\,450$.

解 按题设总体 X 服从指数分布，其概率密度为

$$f(x)=\begin{cases}\frac{1}{\theta}e^{-\frac{x}{\theta}}, & x>0,\\ 0, & \text{其他},\end{cases}\quad \theta>0.$$

$X_1,X_2,\cdots,X_n$ 是来自 X 的样本，现在来检验假设(取显著性水平为 α)

$$H_0:\theta=\theta_0,\quad H_1:\theta\neq\theta_0.$$

取检验统计量

$$\chi^2=\frac{2n\overline{X}}{\theta_0},$$

当 H_0 为真时，由于 $E(\overline{X})=E(X)=\theta_0$，$\chi^2=\frac{2n\overline{X}}{\theta_0}$与 $2n$ 接近，而当 H_1 为真时，

$\chi^2=\dfrac{2n\overline{X}}{\theta_0}$倾向于偏离 $2n$，因此拒绝域具有以下的形式：

$$\frac{2n\overline{X}}{\theta_0}\leqslant k_1 \quad 或 \quad \frac{2n\overline{X}}{\theta_0}\geqslant k_2,$$

此处 k_1,k_2 的值由下式确定：

$$P\{当\ H_0\ 为真时拒绝\ H_0\}=P_{\theta_0}\left\{\left(\frac{2n\overline{X}}{\theta_0}\leqslant k_1\right)\cup\left(\frac{2n\overline{X}}{\theta_0}\geqslant k_2\right)\right\}=\alpha,$$

取

$$P_{\theta_0}\left\{\frac{2n\overline{X}}{\theta_0}\leqslant k_1\right\}=\frac{\alpha}{2},\quad P_{\theta_0}\left\{\frac{2n\overline{X}}{\theta_0}\geqslant k_2\right\}=\frac{\alpha}{2},$$

由第 59 题当 H_0 为真时，$\dfrac{2n\overline{X}}{\theta_0}\sim\chi^2(2n)$. 故得 $k_1=\chi^2_{1-\alpha/2}(2n)$，$k_2=\chi^2_{\alpha/2}(2n)$. 于是得拒绝域为

$$\chi^2=\frac{2n\overline{X}}{\theta_0}\leqslant\chi^2_{1-\alpha/2}(2n) \quad 或 \quad \chi^2=\frac{2n\overline{X}}{\theta_0}\geqslant\chi^2_{\alpha/2}(2n).$$

现在要检验假设

$$H_0:\theta=\theta_0=1\ 450,\quad H_1:\theta\neq 1\ 450.$$

今 $\alpha=0.05$，$n=12$，$\bar{x}=1\ 500$，$\chi^2_{\alpha/2}(2n)=\chi^2_{0.025}(24)=39.364$，$\chi^2_{1-\alpha/2}(24)=\chi^2_{0.975}(24)=12.401$，$\dfrac{2n\bar{x}}{\theta_0}=2\times12\times\dfrac{1\ 500}{1\ 450}=24.828$. 即有 $12.401<24.828<39.364$，故接受 H_0. 认为这批电子元件寿命的均值 $\theta=1\ 450$ h.

62. 经过十一年的试验，达尔文于 1876 年得到 15 对玉米样品的数据如下表，每对作物除授粉方式不同外，其他条件都是相同的. 试用逐对比较法检验不同授粉方式对玉米高度是否有显著的影响($\alpha=0.05$). 问应增设什么条件才能用逐对比较法进行检验？

授粉方式	1	2	3	4	5	6	7
异株授粉的作物高度(x_i)	$23\frac{1}{8}$	12	$20\frac{3}{8}$	22	$19\frac{1}{8}$	$21\frac{4}{8}$	$22\frac{1}{8}$
同株授粉的作物高度(y_i)	$27\frac{3}{8}$	21	20	20	$19\frac{3}{8}$	$18\frac{5}{8}$	$18\frac{5}{8}$

授粉方式	8	9	10	11	12	13	14	15
异株授粉的作物高度(x_i)	$20\frac{3}{8}$	$18\frac{2}{8}$	$21\frac{5}{8}$	$23\frac{2}{8}$	21	$22\frac{1}{8}$	23	12
同株授粉的作物高度(y_i)	$15\frac{2}{8}$	$16\frac{4}{8}$	18	$16\frac{2}{8}$	18	$12\frac{6}{8}$	$15\frac{4}{8}$	18

解 本题是历史上第一个对比试验的结果. 我们用逐对比较法来检验. 计算 x_i 与 y_i 的差 $d_i=x_i-y_i$，得到

$$\begin{array}{cccccccc} -4.25 & -9 & 0.375 & 2 & -0.25 & 2.875 & 3.5 & \\ 5.125 & 1.75 & 3.625 & 7 & 3 & 9.375 & 7.5 & -6 \end{array}$$

今要求在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\mu_D=0,\quad H_1:\mu_D\neq 0.$$

现在 $n=15,\alpha=0.05,t_{\alpha/2}(n-1)=t_{0.025}(14)=2.1448$，由所给的数据得 $\bar{d}=1.775,s_d=5.051$.

$$|t|=\left|\frac{\bar{d}}{s_d/\sqrt{n}}\right|=1.3610<2.1448.$$

故接受 H_0，即认为两种授粉方式对玉米高度无显著影响.

用逐对比较法作检验时，一般应假定各对数据之差 $D_1,D_2,\cdots,D_n$ 构成正态总体的一个样本. 不过这种假定，通常体现于做对比试验的要求上.

63. 一内科医生声称，如果患者每天傍晚聆听一种特殊的轻音乐会降低血压(舒张压，以 mmHg 计，1 mmHg=133.322 4 Pa). 今选取了 10 个患者在试验之前和试验之后分别测量了血压，得到以下的数据：

患者	1	2	3	4	5	6	7	8	9	10
试验之前(x_i)	86	92	95	84	80	78	98	95	94	96
试验之后(y_i)	84	83	81	78	82	74	86	85	80	82

设 $D_i=X_i-Y_i\ (i=1,2,\cdots,10)$ 为来自正态总体 $N(\mu_D,\sigma_D^2)$ 的样本，μ_D,σ_D^2 均未知. 试检验是否可以认为医生的意见是对的(取 $\alpha=0.05$).

解 本题亦宜采用逐对比较法，即在显著性水平 $\alpha=0.05$ 下检验假设 $H_0:\mu_D\leqslant 0,H_1:\mu_D>0$. 检验统计量为 $t=\dfrac{\overline{D}}{S_D/\sqrt{n}}$，当 t 的观察值 $t\geqslant t_\alpha(n-1)$ 时拒绝 H_0. 今 $n=10,D_i=X_i-Y_i$ 的观察值为

$$2\quad 9\quad 14\quad 6\quad -2\quad 4\quad 12\quad 10\quad 14\quad 14$$

得 $\bar{d}=8.3,s_d=5.6184$，查表 $t_\alpha(n-1)=t_{0.05}(9)=1.8331$，检验统计量 t 的观察值

$$t=\frac{8.3}{5.6184/\sqrt{10}}=4.67>1.8331.$$

从而在显著性水平 0.05 下拒绝 H_0，认为医生的意见是对的.

64. 以下是各种颜色的汽车的销售情况：

颜色	红	黄	蓝	绿	棕
车辆数	40	64	46	36	14

试检验顾客对这些颜色是否有偏爱，即检验销售情况是否是均匀的(取 $\alpha=0.05$).

解 以 $P(\times)$表示一顾客买一辆×色的汽车的概率. 本题要求根据销售记录，在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0: P(红)=P(黄)=P(蓝)=P(绿)=P(棕)=0.2.$$

现在 $n=200$，所需计算列表如下：

车辆颜色	f_i	p_i	np_i	$f_i^2/(np_i)$
红	40	0.2	40	40
黄	64	0.2	40	102.4
蓝	46	0.2	40	52.9
绿	36	0.2	40	32.4
棕	14	0.2	40	4.9

$\sum=232.6$

$\chi_\alpha^2(k-r-1)=\chi_{0.05}^2(5-0-1)=\chi_{0.05}^2(4)=9.488$. 而观察值 $\chi^2=232.6-200=32.6>9.488$，故在 $\alpha=0.05$ 下拒绝 H_0，认为顾客对颜色是有偏爱的.

65. 某种闪光灯，每盏灯含 4 个电池. 随机地取 150 盏灯，经检测得到以下的数据：

一盏灯损坏的电池数 x	0	1	2	3	4
灯的盏数	26	51	47	16	10

试取 $\alpha=0.05$ 检验一盏灯损坏的电池数 $X\sim b(4,\theta)$（θ 未知）.

解 本题要求在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0: X\sim b(4,\theta),$$

此处 θ 为未知. 故需在 H_0 下用最大似然估计法估计 θ. 由 $X\sim b(4,\theta)$知 θ 的最大似然估计值为 $\hat{\theta}=\bar{x}/4$（见第七章题 3），即有

$$\hat{\theta}=\frac{1}{4}\times\frac{0\times26+1\times51+2\times47+3\times16+4\times10}{150}=\frac{233}{600}.$$

其余所需的计算列表如下：

A_i	f_i	$\hat{p}_i$	$n\hat{p}_i$	$f_i^2/(n\hat{p}_i)$
$X=0$	26	0.139 978	20.996 7	32.195 5
$X=1$	51	0.355 475	53.321 25	48.779 8
$X=2$	47	0.338 525	50.778 8	43.502 4
$X=3$	16 }	0.143 281	21.492 15 }	27.145 0
$X=4$	10 }	0.022 741	3.411 15 }	

$\sum=151.623$

$\chi_\alpha^2(k-r-1)=\chi_{0.05}^2(4-1-1)=\chi_{0.05}^2(2)=5.992$，而观察值$\chi^2=151.623-150=1.623<5.992$. 故接受 H_0，认为 $X\sim b(4,\theta)$.

66. 临界闪烁频率(cff)是人眼对于闪烁光源能够分辨出它在闪烁的最高频率(以 Hz 计). 超过 cff 的频率，即使光源实际是在闪烁的，人看起来也是连续的(不闪烁的). 一项研究旨在判定 cff 的均值是否与人眼的虹膜颜色有关，所得数据如下：

虹膜颜色	棕色	绿色	蓝色
临界闪烁频率(cff)	26.8　26.3 27.9　24.8 23.7　25.7 25.0　24.5	26.4　29.1 24.2 28.0 26.9	25.7　29.4 27.2　28.3 29.9 28.5

试在显著性水平 0.05 下，检验各种虹膜颜色相应的 cff 的均值有无显著的差异. 设各个总体服从正态分布，且方差相等，不同颜色下的样本之间相互独立.

解　将棕色、绿色、蓝色虹膜的 cff 均值分别记作 μ_1,μ_2,μ_3. 本题要求在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\mu_1=\mu_2=\mu_3,\quad H_1:\mu_1,\mu_2,\mu_3\text{ 不全相等.}$$

现在，

$$n_1=8,n_2=5,n_3=6,n=n_1+n_2+n_3=19,s=3,$$
$$T_{\cdot1}=204.7,T_{\cdot2}=134.6,T_{\cdot3}=169,T_{\cdot\cdot}=508.3.$$
$$S_T=\sum x_{ij}^2-T_{\cdot\cdot}^2/n=13\,659.67-13\,598.362\,63=61.307\,37,$$
$$S_A=\sum T_{\cdot j}^2/n_j-T_{\cdot\cdot}^2/n=13\,621.359\,92-13\,598.362\,63=22.997\,29,$$
$$S_E=S_T-S_A=38.310\,08.$$

S_T,S_A,S_E 的自由度分别为 $n-1=18,s-1=2,n-s=16$. 经计算得方差分析表如下：

方差来源	平方和	自由度	均方	F 比($\alpha=0.05$)
因素 A	22.997 29	2	$\bar{S}_A=11.498\,645$	$\bar{S}_A/\bar{S}_E=4.802$
误差 E	38.310 08	16	$\bar{S}_E=2.394\,38$	
总和 T	61.307 37	18		

而 $F_{0.05}(2,16)=3.63<4.802$，故拒绝 H_0，认为有显著差异.

67. 下面列出了挪威人自 1938—1947 年间年人均脂肪消耗量与患动脉粥样硬化症而死亡的死亡率之间相关的一组数据：

年份	1938	1939	1940	1941	1942	1943	1944	1945	1946	1947
脂肪消耗量 x (kg/人年)	14.4	16.0	11.6	11.0	10.0	9.6	9.2	10.4	11.4	12.5
死亡率 y ($1/(10^5$ 人年))	29.1	29.7	29.2	26.0	24.0	23.1	23.0	23.1	25.2	26.1

设对于给定的 x，Y 为正态变量，且方差与 x 无关.

(1) 求回归直线方程 $\hat{y}=\hat{a}+\hat{b}x$.

(2) 在显著性水平 $\alpha=0.05$ 下检验假设 $H_0:b=0,H_1:b\neq0$.

(3) 求 $\hat{y}|_{x=13}$.

(4) 求 $x=13$ 处 $\mu(x)$的置信水平为0.95的置信区间.

(5) 求 $x=13$ 处 Y 的新观察值 Y_0 的置信水平为0.95的预测区间.

解 (1) 现在 $n=10$. $\sum x_i=116.1$，$\sum y_i=258.5$，$\sum x_iy_i=3\ 046.09$，$\sum x_i^2=1\ 390.09$，$\sum y_i^2=6\ 746.01$，$S_{xx}=\sum x_i^2-\frac{1}{n}(\sum x_i)^2=42.169$，$S_{yy}=\sum y_i^2-\frac{1}{n}(\sum y_i)^2=63.785$，$S_{xy}=\sum x_iy_i-\frac{1}{n}(\sum x_i)(\sum y_i)=44.905$.(此处$\sum$表示 $\sum\limits_{i=1}^{n}$.)

$$\hat{b}=\frac{S_{xy}}{S_{xx}}=1.064\ 881\ 8,\quad \hat{a}=\bar{y}-\bar{x}\hat{b}=13.486\ 722\ 3.$$

故回归直线方程为

$$\hat{y}=13.487+1.065x.$$

(2) 要检验假设 $H_0:b=0,H_1:b\neq0$. 现在 $Q_e=S_{yy}-\hat{b}S_{xy}=15.966\ 483$，$\hat{\sigma}^2=\frac{Q_e}{n-2}=1.995\ 810$，$\hat{\sigma}=1.412\ 732$. 查表 $t_{0.05/2}(n-2)=t_{0.025}(8)=2.306\ 0$，$H_0$ 的拒绝域为 $|t|=\frac{|\hat{b}|}{\hat{\sigma}}\times\sqrt{S_{xx}}\geqslant2.306\ 0$. $|t|$的观察值为 $|t|=4.894\ 8>2.306\ 0$，故在显著性水平 $\alpha=0.05$ 下拒绝 H_0，认为回归效果是显著的.

(3) $x=13$ 处 y 的回归函数值为

$$\hat{y}\big|_{x=13}=13.487+1.065\times13=27.332.$$

(4) $1-\alpha=0.95$，$\frac{\alpha}{2}=0.025$，$t_{\alpha/2}(n-2)=t_{0.025}(8)=2.306\ 0$，从而 $t_{\alpha/2}(n-2)\hat{\sigma}\sqrt{1/n+(x_0-\bar{x})^2/S_{xx}}=1.244$. 故在 $x=13$ 处 $\mu(x)$的一个置信水平

为 0.95 的置信区间是

$$(27.332 \pm 1.244).$$

(5) $t_{\alpha/2}(n-2)\hat{\sigma}\sqrt{1+1/n+(x_0-\bar{x})^2/S_{xx}}=3.487$，故在 $x=13$ 处 Y 的新观察值 Y_0 的一个置信水平为 0.95 的预测区间为

$$(27.332 \pm 3.487).$$

68. 下面给出 1924—1992 年奥林匹克运动会女子 100 m 仰泳的最佳成绩(以 s 计)(其中 1940 年及 1944 年未举行奥运会)：

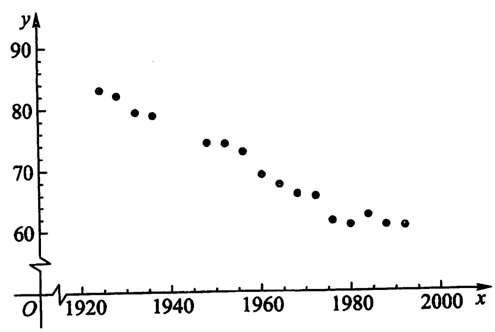
年份	1924	1928	1932	1936	1948	1952	1956	1960
成绩	83.2	82.2	79.4	78.9	74.4	74.3	72.9	69.3
年份	1964	1968	1972	1976	1980	1984	1988	1992
成绩	67.7	66.2	65.8	61.8	60.9	62.6	60.9	60.7

(1) 画出散点图.

(2) 求成绩关于年份的线性回归方程.

(3) 检验回归效果是否显著(取 $\alpha=0.05$).

解 (1) 散点图如题 16.68 图所示.



题 16.68 图

(2) $\sum x_i=31\ 360$，$\sum y_i=1\ 121.2$，$n=16$，

$S_{xx}=7\ 168$，$S_{xy}=-2\ 574.4$，$S_{yy}=948.79$.

回归方程为

$$\hat{y}=774.012\ 5-0.359\ 15x.$$

(3) $|t|=23.13>t_{0.025}(14)=2.144\ 8$，拒绝 H_0，回归效果是非常显著的.

□ 概率论与数理统计　第五版
浙江大学 盛 骤 谢式千 潘承毅 编

□ 概率论与数理统计附册
学习辅导与习题选解　浙大·第五版
浙江大学 盛 骤 谢式千 潘承毅 编

■ 概率论与数理统计
习题全解指南　浙大·第五版
浙江大学 盛 骤 谢式千 潘承毅 编

ISBN 978-7-04-051548-0

定价 36.00元