目　　录

习 题 一

1. 求例 1.2 及例 1.3 中的 $P(A)$,$P(B)$.

例 1.2 投掷两枚分币,则 A="两个都是正面朝上",B="两个都是正面朝下",……

解 投掷两枚分币共有四种可能情况:上上,下下,上下,下上. 在分币匀称的前提下,它们构成一等概基本事件组. 由古典概型(2.1)式,有

$$P(A)=P(\text{上上})=\frac{1}{4},P(B)=P(\text{下下})=\frac{1}{4}.$$

例 1.3 从 10 个同类产品(其中有 8 个正品,2 个次品)中,任意抽取 3 个. 那么

A="3 个都是正品",B="至少 1 个是次品". ……

解 由古典概型(2.1)式,我们有

$$P(A)=\frac{C_8^3}{C_{10}^3}=\frac{7}{15}.$$

为求 $P(B)$,按古典概型(2.1)式,分母还是 C_{10}^3;而分子应是 B,即"至少 1 个是次品"的情况数. 这个事件相对复杂些,一般有两种方法:

(1) 分解法 将"至少 1 个是次品"分解为"恰有 1 个是次品","恰有 2 个是次品"这两种情况,情况数分别为 $C_2^1 \cdot C_8^2$ 与 $C_2^2 \cdot C_8^1$. 于是由(2.1)式,有

$$P(B)=\frac{C_2^1C_8^2+C_2^2C_8^1}{C_{10}^3}=\frac{8}{15}.$$

(2) 逆向法 注意到在 C_{10}^3 个基本事件中,除了"至少 1 个是次品",就是"3 个都是正品",而后者的情况数是容易数出的,它是 C_8^3. 因此由(2.1)式有

$$P(B)=\frac{C_{10}^3-C_8^3}{C_{10}^3}=\frac{8}{15}.$$

评注 (1) 在用古典概型(2.1)式解题时,首先要正确找出一个等概基本事件组.在例1.2中,“上上”、“下下”、“一上一下”就不是一个等概基本事件组.

(2) 分解法与逆向法相比,本例中,后者简单得多.

2. 袋中有红、黄、白色球各一个,每次任取一个,有放回的抽三次,求下列事件的概率:A=“三个都是红的”=“全红”,B=“全黄”,C=“全白”,D=“颜色全同”,E=“全不同”,F=“不全同”,G=“无红”,H=“无黄”,I=“无白”,J=“无红且无黄”,K=“全红或全黄”.

解 有放回的抽取三次.共有 $3\times3\times3=27$ 种情况,它们构成了一个等概基本事件组.为求出那些事件的概率,按(2.1)式.现只需数清它们相应的情况数.显然,A,B,C 的情况数都是1,因此,

$$P(A)=P(B)=P(C)=\frac{1}{27},\text{而 } P(D)=\frac{3}{27}=\frac{1}{9}.$$

E 的情况数应是一个排列总数,是 $3!=6$,所以 $P(E)=\frac{6}{27}=\frac{2}{9}$.

F=“不全同”,采用逆向法,其情况数为 $27-3=24$,因此,$P(F)=\frac{24}{27}=\frac{8}{9}$.

G=“无红”,只有黄、白两种,情况数为 $2\times2\times2=8$,因此,

$$P(G)=P(H)=P(I)=\frac{8}{27}.$$

至于 J=“无红且无黄”,由字意即“全白”,所以

$$P(J)=P(C)=\frac{1}{27}.$$

显然有

$$P(K)=\frac{2}{27}.$$

3. 从一副扑克的52张牌中,任意抽取两张,问都是黑桃的概率有多大?

解 记 A 为所求事件.由于52张牌中有13张是黑桃,则

$$P(A)=\frac{C_{13}^{2}}{C_{52}^{2}}=\frac{1}{17}.$$

4. 在例 2.4 中，求至少有两件次品的概率.

例 2.4 100 件产品中，有 5 件次品，从中任取 50 件，……

解 用逆向法. 除了“至少有两件次品”，就是“至多有一件次品”. 再将后者分解为“没有次品”与“恰有一件次品”. 于是有

$$\begin{aligned}P(\text{“至少有两件次品”})&=\frac{C_{100}^{50}-(C_{95}^{50}+C_{95}^{49}C_{5}^{1})}{C_{100}^{50}}\\&=1-\frac{1\,739}{9\,603}\approx 0.82.\end{aligned}$$

5. 五人排队抓阄，决定谁取得一物(即五个阄中有四个是白阄，只有一个是有物之阄). (1) 问第三人抓到有物之阄的概率是多少？(2) 前三人之一抓到有物之阄的概率是多少？(3) 如果有两物(即五个阄中有两个是有物之阄)，问后两个人都抓不到有物之阄的概率是多少？

解 不妨先将五张阄编号如下：

×	×	×	×	√
①	②	③	④	⑤

五人排队抓阄，无非有如下结果：

$$\left.\begin{array}{ccccc}\text{第一人} & \text{第二人} & \text{第三人} & \text{第四人} & \text{第五人}\\ ① & ② & ③ & ④ & ⑤\\ ① & ② & ③ & ⑤ & ④\\ & & \cdots\cdots & & \\ ⑤ & ④ & ③ & ② & ①\end{array}\right\}\text{共 }5!=120$$

它们构成一等概基本事件组. “第三人抓到有物之阄”，即第三人抓到“√”，也即第三人抓到⑤，应由其中的①②⑤③④，①②⑤④③，…，④③⑤②①所构成，共 4！＝24 个基本事件. 按(2.1)式，

$$P(\text{第三人抓到有物之阄})=\frac{4!}{5!}=\frac{1}{5}.$$

同理,

$$P(\text{第一人抓到有物之阄})=\frac{1}{5},$$

$$P(\text{第二人抓到有物之阄})=\frac{1}{5},$$

所以

$$P(\text{前三人之一抓到有物之阄})=\frac{1+1+1}{5}=\frac{3}{5}.$$

为解(3),先将阄编号如下:

×	×	×	√	√
①	②	③	④	⑤

还是前面的 120 种情况构成一等概基本事件组.问题的关键是找出构成事件“后两个人都抓不到有物之阄”的基本事件数 m.

稍作具体分析便知,“后两个人都抓不到有物之阄”=“④、⑤在前三个位置”,而后者可分解为“④、⑤在第一、二个位置”,“④、⑤在第一、三个位置”,“④、⑤在第二、三个位置”;利用排列组合知识得它们所含的基本事件数都是 $2\times3!=12$. 于是按(2.1)式

$$P(\text{后两个人都抓不到有物之阄})=\frac{12+12+12}{120}=\frac{3}{10}.$$

习 题 二

1. 某产品 40 件，其中有 3 件次品. 现从中任取两件，求其中至少有一件次品的概率.

解法一 设A＝“至少有一件次品”.

B＝“没有次品”.

则 $A=\overline{B}$，于是 $P(A)=1-P(B)$. 而

$$P(B)=\frac{C_{37}^2}{C_{40}^2}=\frac{111}{130}.$$

故

$$P(A)=1-P(B)=\frac{19}{130}\approx 0.146.$$

解法二 A＝“至少有一件次品”＝“恰有一件次品”∪“恰有两件次品”，上式右边两事件是互不相容的，由加法公式，有

$$P(A)=\frac{C_3^1C_{37}^1}{C_{40}^2}+\frac{C_3^2C_{37}^0}{C_{40}^2}=\frac{19}{130}$$

2. 对立与互不相容有何异同？试举例说明.

答 事件 A 与 B 不能同时发生，或说 $AB=V$(不可能事件). 称它们为互不相容或互斥；再加上限制 $A\cup B=U$(必然事件)才称 A 与 B 对立. 这就是说，两事件对立一定互不相容，而互不相容不一定对立.

如上题中，记 A_1＝“恰有一件次品”，A_2＝“恰有两件次品”. 显然，A_1,A_2 互不相容，但它们并不对立，因为 $A_1\cup A_2\neq U$.

3. A,B,C 三事件互不相容与 $ABC=V$ 是否是一回事？为什么？

解 不是一回事. 如图 1，事件 A,B,C 分别表示打靶打中图形 A,B,C，显然有 $ABC=V$；但由于 A,B 是相容的. 因此 A,B,C 三事件不是互不相容的.

4. 从一副扑克牌的 13 张黑桃中，一张接一张地有放回的抽取 3 次，问没有同号的概率？

解 用古典概型求解. $13\times13\times13$ 个不同结果构成一等概基本事件组. 事件 A=“没有同号”，由其中的 $13\times12\times11$ 个(13 中取 3 的排列数)基本事件构成. 按(2.1)式，

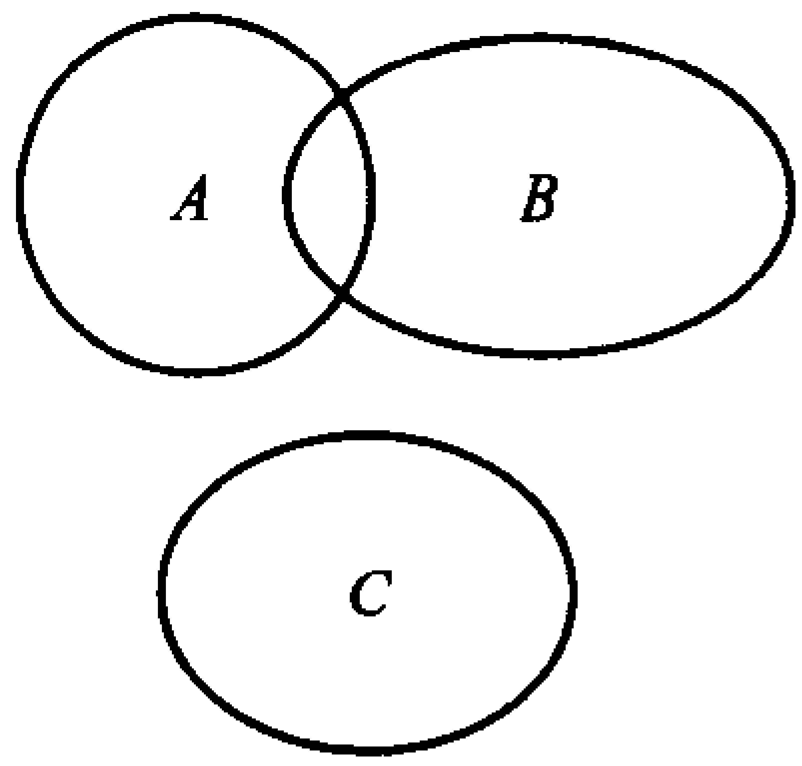

图 1

$$P(A)=\frac{13\times12\times11}{13^3}=\frac{132}{169}\approx0.781.$$

5. 同第 4 题，问抽到有同号的概率？

解 显然，B=“有同号”=$\overline{\text{“没有同号”}}=\overline{A}$，则

$$P(B)=1-P(A)=\frac{37}{169}\approx0.219.$$

6. 同第 4 题，问抽到的 3 张中最多只有两张同号的概率？

解 “最多只有两张同号”=$\overline{\text{“三张全同号”}}$，而

$$P(\text{三张全同号})=\frac{13}{13^3}=\frac{1}{169},$$

所以

$$P(\text{最多只有两张同号})=1-\frac{1}{169}\approx0.994.$$

7. 将习题一第 2 题中的条件改为：盒中有四个球，其中两个红球、一个黄球、一个白球，其他不变，求 $A,B,\cdots,K$ 的概率. 并求事件 L=“无红或无黄”的概率.

解 $P(A)=\frac{2^3}{4^3}=\frac{1}{8}$，　$P(B)=P(C)=\frac{1}{4^3}=\frac{1}{64}$，

$$P(D)=P(A\cup B\cup C)=P(A)+P(B)+P(C)=\frac{5}{32},$$

$$P(E)=\frac{2\cdot3!}{4^3}=\frac{3}{16},\quad P(F)=1-P(D)=\frac{27}{32},$$

$$P(G)=\frac{2^3}{4^3}=\frac{1}{8},\quad P(H)=\frac{3^3}{4^3}=\frac{27}{64}=P(I),$$

$$P(J)=P(C)=\frac{1}{64},$$

$$P(K)=P(A\cup B)=P(A)+P(B)=\frac{9}{64},$$

$$\begin{aligned}P(L)&=P(G\cup H)=P(G)+P(H)-P(GH)\\&=\frac{1}{8}+\frac{27}{64}-P(C)=\frac{17}{32}.\end{aligned}$$

8. 利用加法公式(2)导出三个事件的概率加法公式.

解 首先 $A\cup B\cup C=(A\cup B)\cup C$,对 $A\cup B,C$ 这两个事件,用加法公式(2)得

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

再对 AC,BC 这两个事件用加法公式(2)得

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

代入前式得

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

9*. 利用加法公式(2)和数学归纳法证明下列若尔当(Jordan)公式:设 $A_1,A_2,\cdots,A_n(n\geqslant 2)$ 是 n 个事件,则

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{k=1}^{n}(-1)^{k-1}\sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n}P\left(\bigcap_{l=1}^{k}A_{i_l}\right).\quad(0.1)$$

证 $n=2$ 时,(0.1)式即为加法公式(2).

现设(0.1)对某个特定的 n 成立,要证(0.1)对 $n+1$ 也成立,即要证

$$P\left(\bigcup_{i=1}^{n+1}A_i\right)=\sum_{k=1}^{n+1}(-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_k\leqslant n+1}P\left(\bigcap_{l=1}^{k}A_{i_l}\right).\quad(0.2)$$

(0.2)式左边

$$P\left(\bigcup_{i=1}^{n+1} A_i\right)=P\left(\left(\bigcup_{i=1}^{n} A_i\right)\cup A_{n+1}\right)$$

$$=P\left(\bigcup_{i=1}^{n} A_i\right)+P(A_{n+1})-P\left(\left(\bigcup_{i=1}^{n} A_i\right)\cap A_{n+1}\right)$$

$$=\sum_{k=1}^{n}(-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_k\leqslant n}P\left(\bigcap_{l=1}^{k} A_{i_l}\right)+P(A_{n+1})-$$

$$P\left(\bigcup_{i=1}^{n}(A_i\cap A_{n+1})\right)$$

$$=\sum_{k=1}^{n}(-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_k\leqslant n}P\left(\bigcap_{l=1}^{k} A_{i_l}\right)+P(A_{n+1})-$$

$$\sum_{k=1}^{n}(-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_k\leqslant n}P\left(\bigcap_{l=1}^{k}(A_{i_l}\cap A_{n+1})\right)$$

$$=\sum_{k=1}^{n}(-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_k\leqslant n}P\left(\bigcap_{l=1}^{k} A_{i_l}\right)+P(A_{n+1})-$$

$$\sum_{k=1}^{n}(-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_k\leqslant n}P\left(\left(\bigcap_{l=1}^{k} A_{i_l}\right)A_{n+1}\right)$$

(0.2)式右边

$$\sum_{k=1}^{n+1}(-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_k\leqslant n+1}P\left(\bigcap_{l=1}^{k} A_{i_l}\right)=\sum_{k=1}^{n+1}(-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_k=n+1}$$

$$P\left(\bigcap_{l=1}^{k} A_{i_l}\right)+\sum_{k=1}^{n+1}(-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_k\leqslant n}P\left(\bigcap_{l=1}^{k} A_{i_l}\right)\triangleq \mathrm{I}+\mathrm{II},$$

其中，$\mathrm{I}=\sum_{k=2}^{n+1}(-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_{k-1}\leqslant n}P\left(\bigcap_{l=1}^{k-1} A_{i_l}\cdot A_{n+1}\right)+P(A_{n+1})$

$$=-\sum_{k-1=1}^{n}(-1)^{k-1-1}\sum_{1\leqslant i_1<\cdots<i_{k-1}\leqslant n}P\left(\bigcap_{l=1}^{k-1} A_{i_l}\cdot A_{n+1}\right)+P(A_{n+1});$$

而对Ⅱ，注意到 k 取 $n+1$ 时，$1\leqslant i_1<\cdots<i_k\leqslant n$ 不可能成立，所以

$$\mathrm{II}=\sum_{k=1}^{n}(-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_k\leqslant n}P\left(\bigcap_{l=1}^{k} A_{i_l}\right)$$

因此，(0.2)式左边=(0.2)式右边．由数学归纳法．若尔当公式得证．

评注 本题的证明部分，非数学类读者可略过不看；不过，对$n=3$的情况，即第8题要掌握．若尔当公式，是若尔当在研究平面图形的面积时首先发现的．

习 题 三

1. 一个工人看管三台机床，在 1 h 内机床不需要工人照管的概率：第一台等于 0.9，第二台等于 0.8，第三台等于 0.7. 求在 1 h 内三台机床中最多有一台需要工人照管的概率（机床是否需要照管是相互独立的）.

解 记 B="三台机床中最多有一台需照管"，

A_i="第 i 台机床不需要照管"，$i=1,2,3$.

已知 $P(A_1)=0.9, P(A_2)=0.8, P(A_3)=0.7$. 为求 $P(B)$，我们注意到事件 B 与 A_1,A_2,A_3 有如下关系（为什么？请读者自行思考）：

$$B=A_1A_2A_3\cup\overline{A}_1A_2A_3\cup A_1\overline{A}_2A_3\cup A_1A_2\overline{A}_3.$$

上式右边的四个事件是互不相容的（为什么?）. 由加法公式有

$$P(B)=P(A_1A_2A_3)+P(\overline{A}_1A_2A_3)+P(A_1\overline{A}_2A_3)+P(A_1A_2\overline{A}_3)$$

由于 A_1,A_2,A_3 相互独立，则按乘法公式有

$$\begin{aligned}P(B)&=0.9\times0.8\times0.7+0.1\times0.8\times0.7+0.9\times0.2\times0.7+\\&\quad 0.9\times0.8\times0.3\\&=0.902.\end{aligned}$$

评注 找出 B 与 A_1,A_2,A_3 的关系是本题的关键.

2. 电路由电池 A 与两个并联的电池 B 及 C 串联而成. 设电池 A,B,C 损坏的概率分别是 0.3，0.2，0.2，求电路发生断电的概率（电池是否损坏是互不影响的）.

解 记 A="电池 A 损坏"，

B="电池 B 损坏"，

C="电池 C 损坏"，

D="电路断电".

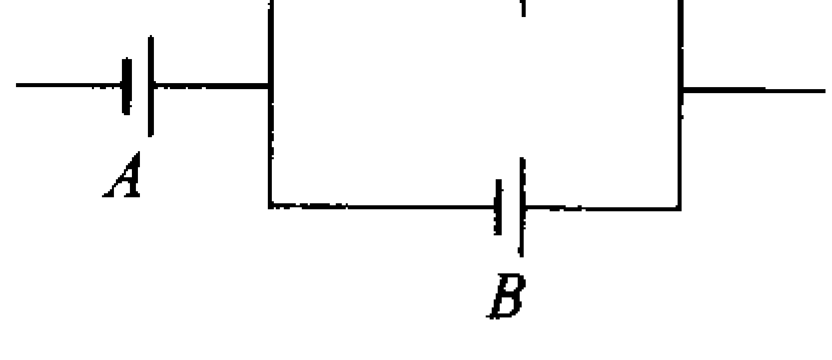

图 2

注意到 D 与 A,B,C 有如下关系（为什

么？请思考）：

$$D=A\cup BC,$$

因此 $P(D)=P(A)+P(BC)-P(ABC).$

再由 A,B,C 的相互独立性，得

$$\begin{aligned}P(D)&=P(A)+P(B)P(C)-P(A)P(B)P(C)\\&=0.3+0.2\times0.2-0.3\times0.2\times0.2=0.328.\end{aligned}$$

3. 某机械零件的加工由两道工序组成. 第一道工序的废品率为 0.015，第二道工序的废品率为 0.02. 假定两道工序出废品是彼此无关的，求产品的合格率.

解 首先注意，所谓一批产品的废品率是废品在该批产品中所占的比例，即废品数/产品总数；数值上它等于：从该批产品中任取一个，取到废品的概率. 这就是为什么该类问题可用概率方法来讨论解决的逻辑依据. 现设想从该批产品中任取一个. 记

A=“第一道工序产生的废品”，

B=“第二道工序产生的废品”，

C=“合格品”.

注意到 C 与 A,B 有如下关系：

$$C=\overline{A}\cap\overline{B},$$

由独立性， $P(C)=P(\overline{A})P(\overline{B})=(1-P(A))(1-P(B))$

$$=(1-0.015)(1-0.02)\approx0.965.$$

4. 在 $1,2,\cdots,100$ 中任取一数，问它既能被 2 整除又能被 5 整除的概率是多少？又它能被 2 整除或能被 5 整除的概率是多少？

解 记 A=“被 2 整除”，B=“被 5 整除”，则

$$P(\text{既能被 2 整除又能被 5 整除})=P(AB)$$

$$=P(\text{被 10 整除})=\frac{10}{100}=\frac{1}{10};$$

$$P(\text{被 2 整除或被 5 整除})=P(A\cup B)$$

$$=P(A)+P(B)-P(AB)=\frac{50}{100}+\frac{20}{100}-\frac{1}{10}=\frac{3}{5}.$$

5. 加工某一零件共需经过四道工序. 设第一，二，三，四道工

序的次品率分别是 2%,3%,5%,3%. 假定各道工序是互不影响的. 求加工出来的零件的次品率.

解 记 A_i="第 i 道工序产生的次品",$i=1,2,3,4$,

B="合格品",

则有

$$B=\overline{A}_1\overline{A}_2\overline{A}_3\overline{A}_4.$$

由独立性,

$$P(B)=P(\overline{A}_1)P(\overline{A}_2)P(\overline{A}_3)P(\overline{A}_4)$$
$$=(1-0.02)(1-0.03)(1-0.05)(1-0.03)=0.876.$$

所以,加工出来的零件的次品率为 0.124(或写作 12.4%).

6. 当掷五枚分币时,已知至少出现两个正面,问正面数刚好是三个的条件概率是什么?

解 记 A="至少出现两个正面",

B="恰好出现三个正面".

要求的是 $P(B|A)$.

由条件概率公式(5.1),

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

注意到有关系 $B\subset A$,所以 $AB=B$;因此上式右边的分子可改为 $P(B)$.

先求 B 的概率. 将五枚分币编号为 1,2,3,4,5,B 可分解为

$$\begin{cases}\text{正,正,正,反,反;(1,2,3 号分币为正面,4,5 号分币为反面. 下类推.)}\\ \text{正,正,反,正,反;}\\ \cdots\cdots\cdots\cdots\\ \text{反,反,正,正,正.}\end{cases}$$

共 C_5^3 个基本事件,而它们的概率,按独立性都是 $(1/2)^5$. 所以

$$P(B)=C_5^3(1/2)^5=\frac{10}{32}=\frac{5}{16}.$$

至于 $P(A)$,先对 $\overline{A}$="至多出现一个正面",进行分解:

$\overline{A}$=“恰出现0个正面”∪“恰出现一个正面”≡$C\cup D$. 类似上面求$P(B)$的方法可得：

$$P(C)=(1/2)^5=\frac{1}{32},$$

$$P(D)=\mathrm{C}_5^1(1/2)^5=\frac{5}{32},$$

因此

$$\begin{aligned}P(A)&=1-P(\overline{A})=1-(P(C)+P(D))\\&=1-\left(\frac{1}{32}+\frac{5}{32}\right)=\frac{13}{16},\end{aligned}$$

所以

$$P(B|A)=\frac{P(B)}{P(A)}=\frac{5}{13}.$$

习　题　四

1. 两台机床加工同样的零件，第一台出现废品的概率是 0.03，第二台出现废品的概率是 0.02. 加工出来的零件放在一起，并且已知第一台加工的零件比第二台加工的零件多一倍，求任意取出的零件是合格品的概率；又，如果任意取出的零件经检查是废品，求它是由第二台机床加工的概率.

解　任取一零件，记

A_i＝"它由第 i 台机床加工的"，$i=1.2$，

B＝"它是合格品".

由全概公式（A_1,A_2 构成一完备事件组），

$$P(\overline{B})=P(A_1)P(\overline{B}|A_1)+P(A_2)P(\overline{B}|A_2)$$
$$=\frac{2}{3}\times 0.03+\frac{1}{3}\times 0.02=\frac{0.08}{3}\approx 0.027,$$

因此 $P(B)=1-\dfrac{0.08}{3}\approx 0.973$.

本题第 2 问实际上求的是条件概率 $P(A_2|\overline{B})$. 由逆概公式，

$$P(A_2|\overline{B})=\frac{P(\overline{B}|A_2)P(A_2)}{P(\overline{B})}=\frac{0.02\times\frac{1}{3}}{\frac{1}{3}\times 0.08}=0.25.$$

2. 盒中放有 12 个乒乓球，其中有 9 个是新的. 第一次比赛时从中任取 3 个来用，比赛后仍放回盒中. 第二次比赛时再从盒中任取 3 个. 求第二次取出的球都是新球的概率. 又，已知第二次取出的球都是新球，求第一次取到都是新球的概率.

解　记A_i＝"第一次比赛时取到 i 个新球"，$i=0,1,2,3$；

B＝"第二次比赛时取到 3 个新球".

显然，A_0,A_1,A_2,A_3 构成一完备事件组. 由全概公式

$$P(B)=\sum_{i=0}^{3}P(A_i)P(B\mid A_i)$$

$$=\frac{C_9^0C_3^3}{C_{12}^3}\cdot\frac{C_9^3C_3^0}{C_{12}^3}+\frac{C_9^1C_3^2}{C_{12}^3}\cdot\frac{C_8^3C_4^0}{C_{12}^3}+\frac{C_9^2C_3^1}{C_{12}^3}\cdot\frac{C_7^3C_5^0}{C_{12}^3}+\frac{C_9^3C_3^0}{C_{12}^3}\cdot\frac{C_6^3C_6^0}{C_{12}^3}$$

$$=\frac{1}{(C_{12}^3)^2}\times(84+1\,512+3\,780+1\,680)=\frac{7\,056}{(220)^2}\approx 0.145\,8.$$

本题第 2 问即要求条件概率 $P(A_3|B)$. 由逆概公式

$$P(A_3|B)=\frac{P(B|A_3)P(A_3)}{P(B)}=\frac{1\,680}{7\,056}=\frac{5}{21}\approx 0.238\,1.$$

3. 有三只盒子，在甲盒中装有 2 枝红芯圆珠笔、4 枝蓝芯圆珠笔，乙盒中装有 4 枝红的、2 枝蓝的，丙盒中装有 3 枝红的、3 枝蓝的. 今从其中任取一枝. 设到三只盒子取物的机会相同，问：它是红芯圆珠笔的概率为多少？又若已知取得的是红的，它是从甲盒中取出的概率为多少？

解　记　A_1＝“取自甲盒”，

A_2＝“取自乙盒”，

A_3＝“取自丙盒”，

B＝“取到红的”.

显然，A_1,A_2,A_3 构成一完备事件组. 由全概公式

$$P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+P(A_3)P(B|A_3)$$

$$=\frac{1}{3}\times\frac{2}{6}+\frac{1}{3}\times\frac{4}{6}+\frac{1}{3}\times\frac{3}{6}=\frac{1}{2}.$$

本题第 2 问是求 $P(A_1|B)$，按逆概公式，

$$P(A_1|B)=\frac{P(B|A_1)P(A_1)}{P(B)}=\frac{\frac{2}{6}\times\frac{1}{3}}{\frac{1}{2}}=\frac{2}{9}.$$

习　题　五

1. 设某种型号的电阻的次品率为 0.01，现在从产品中抽取 4 个. 分别求出没有次品、有 1 个次品、有 2 个次品、有 3 个次品、全是次品的概率.

解　先分析随机试验“从产品中抽取 4 个”的含义，它等价于“从产品中每次抽取 1 个，无放回地抽 4 次”. 而后者，当产品的批量相当大时，可近似地按“从产品中每次抽取 1 个，有放回地抽 4 次”处理.

我们就按“有放回地抽 4 次”来解. 此时，可按独立试验序列概型来求，按公式

$p_k=P(k$ 个次品数$)=C_4^k(0.01)^k(0.99)^{4-k}$，$k=0,1,2,3,4$. 可算出：

$$p_0=0.960\,6,\ p_1=0.038\,8,\ p_2=0.000\,6,\ p_3\approx 0,\ p_4\approx 0.$$

2. 某类电灯泡使用时数在 1 000 h 以上的概率为 0.2，求三个灯泡在使用 1 000 h 以后最多只有一个坏的概率.

解　按独立试验序列概型，已知 P(一个灯泡在使用1 000 h以后坏)$=0.8$，$n=3$.

所以　P(三个灯泡在使用 1 000 h 以后恰有 0 个坏)$=$

$C_3^0(0.8)^0(0.2)^3=0.008$

P(三个灯泡在使用 1 000 h 以后恰有 1 个坏)$=$

$C_3^1 0.8(0.2)^2=0.096$

而“最多只有一个坏”=“恰有 0 个坏”∪“恰有 1 个坏”.

故所求概率为 $0.008+0.096=0.104$.

3. 有 6 个元件，它们断电的概率第一个为 $p_1=0.6$，第二个为 $p_2=0.2$，其余 4 个都为 $p_3=0.3$，求线路断电的概率，若

(1) 所有的元件串联；

(2) 元件按图 3 联接.

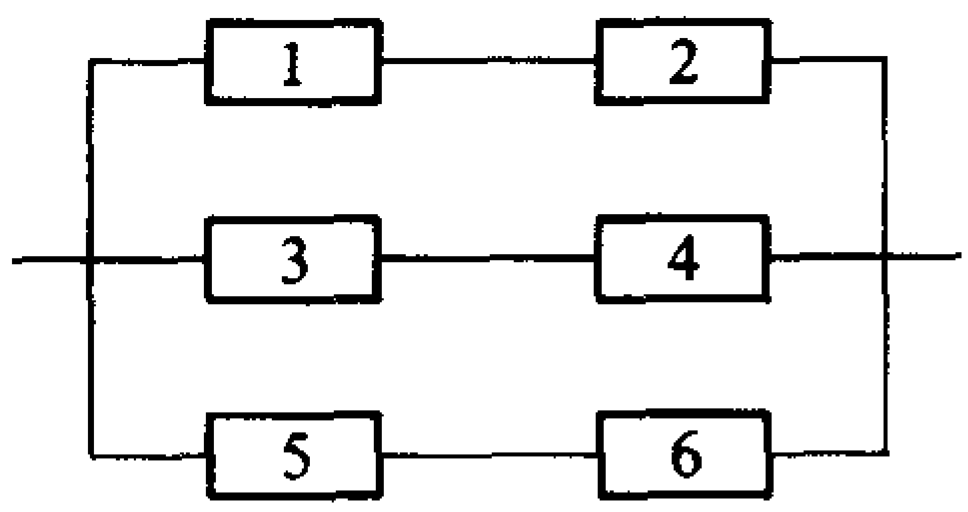

图 3

解 记 A_i =“第 i 个元件断电”, $i=1,2,3,4,5,6$.

(1) 串联

记 B_1 =“线路断电”. 为求 $P(B_1)$,先来找出 B_1 与 $A_1,\cdots,A_6$ 间的关系. 由于是串联,有

$$B_1=A_1\cup A_2\cup\cdots\cup A_6, \quad (\text{为什么? 请想清楚}).$$

因此,可利用 Jordan 公式来求出 $P(B_1)$;然而计算过于繁琐. 我们采用其对立事件,并利用对偶律(参见课文中(4.15)式),有

$$\overline{B}_1=\overline{A}_1\cap\overline{A}_2\cap\cdots\cap\overline{A}_6$$

再按 $A_1,\cdots,A_6$ 的独立性(这里作了 $A_1,\cdots,A_6$ 独立性的假定. 一般讲,是可以作这个假定的),得

$$P(\overline{B}_1)=P(\overline{A}_1)P(\overline{A}_2)\cdots P(\overline{A}_6).$$

由此算得

$$P(B_1)=0.923.$$

(2) 如图 3 联接

记 B_2 =“线路断电”,按图 3, B_2 与 $A_1,\cdots,A_6$ 间的关系如下

$B_2=(A_1\cup A_2)\cap(A_3\cup A_4)\cap(A_5\cup A_6)$. (为什么? 其原理跟习题二的第 2 题同)

于是有
$$\begin{aligned}P(B_2)&=P(A_1\cup A_2)P(A_3\cup A_4)P(A_5\cup A_6)\\&=(0.6+0.2-0.6\times0.2)(0.3+0.3-0.3\times0.3)^2\\&=0.177\end{aligned}$$

(这里,也假定 $A_1,\cdots,A_6$ 相互独立).

4. 设昆虫生产 k 个卵的概率 $p_k=\dfrac{\lambda^k}{k!}e^{-\lambda}(k=0,1,2,\cdots)$,又设一个虫卵能孵化为昆虫的概率等于 p. 若卵的孵化是互相独立的,问此昆虫的下一代有 l 条的概率是多少?

解 首先记 A_k =“此昆虫生产 k 个卵”, $k=0,1,2,\cdots$; B =“此

昆虫下一代有 l 条昆虫”.

这里要明确两个事实：

(1) $\{A_k, k=0,1,2,\cdots\}$ 构成一完备事件组；

(2) 当 $k<l$ 时，$P(B|A_k)=0$. 这是因为 k 个卵最多孵化出 k 条虫；

当 $k \geqslant l$ 时，$P(B|A_k)=C_k^l p^l(1-p)^{k-l}$. 这是按独立试验序列概型.

于是按全概公式

$$
\begin{aligned}
P(B) &= \sum_{k=0}^{\infty} P(A_k)P(B \mid A_k) \\
&= \sum_{k=l}^{\infty} \frac{\lambda^k}{k!}e^{-\lambda} \cdot C_k^l p^l(1-p)^{k-l} \\
&= e^{-\lambda} \sum_{k=l}^{\infty} \frac{\lambda^k}{k!} \frac{k!}{(k-l)!l!} p^l (1-p)^{k-l} \\
&= e^{-\lambda} \cdot \frac{\lambda^l}{l!} p^l \sum_{k-l=0}^{\infty} \frac{\lambda^{k-l}}{(k-l)!}(1-p)^{k-l} \\
&= \frac{e^{-\lambda}(\lambda p)^l}{l!} \sum_{m=0}^{\infty} \frac{[\lambda(1-p)]^m}{m!} = \frac{e^{-\lambda}(\lambda p)^l}{l!} e^{\lambda(1-p)} \\
&= \frac{(\lambda p)^l}{l!} e^{-\lambda p}\ .
\end{aligned}
$$

习　题　六

1. 求例 1.1 中“抽得的次品件数”X 的概率分布.

例 1.1　设有产品 100 件，其中有 5 件次品，95 件正品. 现从中随机抽取 20 件，……

解　$P(X=k)=C_5^k C_{95}^{20-k}/C_{100}^{20}, k=0,1,2,3,4,5.$

2. 求例 1.3 中“击中目标次数”X 的概率分布.

例 1.3　设某射手每次射击击中目标的概率是 0.8，现在连续射击 30 次，……

解　$P(X=k)=C_{30}^k(0.8)^k(0.2)^{30-k}, k=0,1,\cdots,30.$

3. 求例 1.4 中“所需射击次数”X 的概率分布.

例 1.4　某射手每次射击击中目标的概率是 0.8，现在连续向一目标射击，直到击中目标为止. ……

解　X 可能取的值是什么？不可能是零，但理论上可以取到一切正整数. 为找出它的概率分布，先来分析事件“$X=5$”的含义. 它意味着，前 4 次都未击中目标，而第 5 次击中了. 可以形象地表示为 “×，×，×，×，√”. 由独立性，

$$P(X=5)=P(\times\times\times\times\surd)=(0.2)^4 0.8.$$

所以，X 的概率分布

$$P(X=k)=(0.2)^{k-1}(0.8), k=1,2,\cdots.$$

4. 一批零件中有九个正品与三个废品，安装机器时，从这批零件中任取一个. 如果每次取出的废品不再放回，而再取一个零件，直到取得正品时为止. 求在取得正品以前已取出废品数的概率分布.

解　记 X 为取得正品以前已取出的废品数. 首先 X 可能取的值是 0,1,2,3. 逐个计算相应概率：

$$P(X=0)=P(\text{第一次就抽到正品})=\frac{9}{12}=\frac{3}{4}.$$

$$P(X=1)=P(\text{第一次抽到废品而第二次抽到正品})$$
$$=\frac{3}{12}\times\frac{9}{11}=\frac{9}{44}$$

类似有

$$P(X=2)=\frac{3}{12}\times\frac{2}{11}\times\frac{9}{10}=\frac{9}{220}$$

$$P(X=3)=\frac{3}{12}\times\frac{2}{11}\times\frac{1}{10}\times\frac{9}{9}=\frac{1}{220}$$

5. 抛掷一枚分币，直到出现“正面朝上”时为止，求抛掷次数的概率分布.

解 与第 3 题类似，可得抛掷次数 X 的概率分布为

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

6. 从一副扑克牌中发出 5 张，求其中黑桃张数的概率分布.

解 记 X=黑桃张数. 它的可能值是 0,1,2,3,4,5，其概率分布为

$$P(X=k)=\frac{C_{13}^{k}C_{39}^{5-k}}{C_{52}^{5}},\quad k=0,1,\cdots,5.$$

7. 设 X 服从泊松分布，且已知 $P(X=1)=P(X=2)$，求 $P(X=4)$.

解 （泊松分布有一个参数 λ. 不同的 λ 其相应的泊松分布不同；由条件“$P(X=1)=P(X=2)$”可将 λ 解出，从而求得 $P(X=4)$）. 具体过程如下：

已知 X 服从泊松分布，设参数为 λ，即

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

由 $P(X=1)=P(X=2)$ 得

$$\frac{\lambda}{1!}e^{-\lambda}=\frac{\lambda^2}{2!}e^{-\lambda},$$

解得 $\lambda=2$，从而

$$P(X=4)=\frac{2^4}{4!}e^{-2}=\frac{2}{3}e^{-2}=0.090\ 2.$$

8. 已知一电话交换台每分钟的呼唤次数服从参数为 4 的泊松分布. 求(1) 每分钟恰有 8 次呼唤的概率;(2) 每分钟呼唤次数大于 8 的概率.

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

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

(2) $P(X>8)=1-P(X\leqslant 8)=1-e^{-4}\left(1+4+\frac{4^2}{2!}+\cdots+\frac{4^8}{8!}\right)$

$$=1-0.978\ 6=0.021\ 4.$$

9. 设 X 服从泊松分布,分布律为

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

问 k 取何值时 $P(X=k)$ 最大?

解 我们先求出 $P(X=k+1)$ 与 $P(X=k)$ 的比值:

$$\frac{p_{k+1}}{p_k}=\frac{P(X=k+1)}{P(X=k)}=\frac{\lambda^{k+1}e^{-\lambda}/(k+1)!}{\lambda^k e^{-\lambda}/k!}=\frac{\lambda}{k+1}.$$

则 当 $k+1<\lambda$ 时 $p_k<p_{k+1}$,

当 $k+1=\lambda$ 时 $p_k=p_{k+1}$,

当 $k+1>\lambda$ 时 $p_k>p_{k+1}$

(请读者自己分析一下 p_k 随着 k 的增加的变化情况,并想像一下 p_k 的图形.)

因此不难看出:当 λ 为整数时,$k=\lambda-1$ 或 $k=\lambda$ 时,p_k 最大;

λ 不是整数时,$k=[\lambda]$时,p_k 最大.

综合以上,

$$k=\begin{cases}[\lambda], & \lambda \text{ 非整数},\\ \lambda-1 \text{ 或 } \lambda, & \lambda \text{ 整数}\end{cases}$$

时,$P(X=k)$最大.

10. 验证等式 $\sum\limits_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}=1\quad(\lambda>0)$.

(泊松分布的"总概率"为 1).

证 根据指数函数 e^x 的泰勒展式:

$$e^x=\sum_{k=0}^{\infty}\frac{x^k}{k!},$$

即得要验证的等式.

11. 利用恒等式

$$(1+x)^N=(1+x)^M(1+x)^{N-M}$$

两边 x^n 的系数相等,验证等式

$$\sum_{m=0}^{l}\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}=1$$

(超几何分布的"总概率"为 1),其中 N,M,n 是正整数,$N\geqslant M$,$N\geqslant n$,$l=\min(M,n)$.(当 $i>k$ 时规定 $C_k^i=0$).

证 将恒等式两边展开,得

$$\sum_{n=0}^{N}C_N^n x^n=\left(\sum_{m=0}^{M}C_M^m x^m\right)\left(\sum_{k=0}^{N-M}C_{N-M}^k x^k\right)$$

利用等式两边 x^n 的系数相等,有

$$C_N^n=\sum_{m=0}^{l}C_M^m C_{N-M}^{n-m}. \tag{0.3}$$

等式得以验证.

评注 等式(0.3),也可从 N 个不同元素中取 n 个的组合按某种方式分类计数而得到.还有,任一具有实际背景的离散型随机变量,其分布列之和必是 1.

习　题　七

1. 设随机变量 X 的概率密度为

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

求(1) 常数 C;(2) X 落在区间(0.3,0.7)内的概率.

解 (1) 既然 $p(x)$是概率密度,应有 $\int_{-\infty}^{+\infty}p(x)\mathrm{d}x=1$,因此,

$$\int_{-\infty}^{+\infty}p(x)\mathrm{d}x=\int_0^1 Cx\mathrm{d}x=C\frac{x^2}{2}\bigg|_0^1=\frac{C}{2}=1,$$

得 $C=2$.

(2) 所求概率

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

2. 随机变量 X 的概率密度为

$$p(x)=\begin{cases}\dfrac{C}{\sqrt{1-x^2}}, & |x|<1,\\ 0, & 其他.\end{cases}$$

求(1) 常数 C;(2) X 落在区间$\left(-\frac{1}{2},\frac{1}{2}\right)$内的概率.

解 (1) $$\int_{-\infty}^{+\infty}p(x)\mathrm{d}x=\int_{-1}^{1}\frac{C}{\sqrt{1-x^2}}\mathrm{d}x=C\arcsin x\bigg|_{-1}^{1}$$
$$=C\pi=1.$$

故 $C=\frac{1}{\pi}$.

(2) $$P\left(-\frac{1}{2}<X<\frac{1}{2}\right)=\int_{-\frac{1}{2}}^{\frac{1}{2}}\frac{\mathrm{d}x}{\pi\sqrt{1-x^2}}=\frac{1}{\pi}\arcsin x\bigg|_{-\frac{1}{2}}^{\frac{1}{2}}$$
$$=\frac{1}{3}.$$

3. 随机变量 X 的概率密度为

$$p(x)=Ce^{-|x|}, \quad -\infty<x<+\infty.$$

求(1) 常数 C;(2) X 落在区间(0,1)内的概率.

解 (1) $\int_{-\infty}^{+\infty} p(x)\mathrm{d}x = C\int_{-\infty}^{+\infty} e^{-|x|}\mathrm{d}x = C\cdot 2\int_0^{+\infty} e^{-x}\mathrm{d}x = 2C$

所以,$C=\frac{1}{2}$.

(2) $P(0<X<1)=\int_0^1 p(x)\mathrm{d}x = \frac{1}{2}\int_0^1 e^{-x}\mathrm{d}x = \frac{1}{2}(1-e^{-1})$.

4. 设 $X\sim N(1,(0.6)^2)$,求 $P(X>0)$和 $P(0.2<X<1.8)$

解 由 $X\sim N(1,0.6^2)$

$$P(X>0)=\int_0^{+\infty}\frac{1}{\sqrt{2\pi}\cdot 0.6}e^{-\frac{(x-1)^2}{2\cdot 0.6^2}}\mathrm{d}x$$

$$\xlongequal{\left(令\frac{x-1}{0.6}=t\right)}\int_{-1.667}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}\mathrm{d}t$$

$$\xlongequal{(对称性)}\int_{-\infty}^{1.667}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}\mathrm{d}t$$

$$=\Phi(1.667)=0.9522.$$

$$P(0.2<X<1.8)=\int_{0.2}^{1.8}\frac{1}{\sqrt{2\pi}0.6}e^{-\frac{(x-1)^2}{2\cdot 0.6^2}}\mathrm{d}x$$

$$\xlongequal{\left(令\frac{x-1}{0.6}=t\right)}\int_{-\frac{4}{3}}^{\frac{4}{3}}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\mathrm{d}t$$

$$=2\int_0^{\frac{4}{3}}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\mathrm{d}t = 2\left[\Phi\left(\frac{4}{3}\right)-\Phi(0)\right]$$

$$=2\times 0.4086=0.8172.$$

5. 设 $X\sim N(\mu,\sigma^2)$,对

$$P(\mu-k\sigma<X<\mu+k\sigma)=0.95,0.90,0.99,$$

分别找出相应的 k 值(查表). 又对于 k 的什么值有 $P(X>\mu-k\sigma)=0.95$?

解 采用解上题类似方法:定积分换元(积分变量替换)与利用对称性,可知

$$P(\mu-k\sigma<X<\mu+k\sigma)=2[\Phi(k)-0.5],$$

(注意,上式右边跟 μ,σ 无关.)则对应于 $P(\mu-k\sigma<X<\mu+k\sigma)=0.95,0.90,0.99$,有

$$\Phi(k)=\frac{0.95}{2}+0.5,\frac{0.90}{2}+0.5,\frac{0.99}{2}+0.5$$
$$=0.975,0.95,0.995.$$

查表得 $k=1.96,1.65,2.58$.

为解第 2 问,首先容易求得

$$P(X>\mu-k\sigma)=\Phi(k),$$

因此

$$k=1.65.$$

6. 乘以什么常数 C 将使 Ce^{-x^2+x} 变成概率密度函数?

解 由于

$$\int_{-\infty}^{+\infty} e^{-x^2+x}dx=\int_{-\infty}^{+\infty} e^{-(x-\frac{1}{2})^2+\frac{1}{4}}dx=e^{\frac{1}{4}}\int_{-\infty}^{+\infty} e^{-t^2}dt$$
$$=2e^{\frac{1}{4}}\int_0^{+\infty} e^{-t^2}dt=e^{\frac{1}{4}}\sqrt{\pi},$$

故要乘以常数 $\frac{1}{\sqrt{\pi}}e^{-\frac{1}{4}}$.

7. 设 X 的密度为

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

求 $Y=\ln X$ 的密度.

解 用分布函数法求 Y 的密度函数 $p_Y(y)$.

$$P(Y\leqslant y)=P(\ln X\leqslant y)=P(X\leqslant e^y)=F(e^y),$$

这里 $F(\cdot)$为 X 的分布函数.两边对 y 求导,得

$$p_Y(y)=F'(e^y)\cdot e^y=p_X(e^y)e^y$$
$$=\frac{2e^y}{\pi(e^{2y}+1)}.$$

8. 设 X 服从自由度为 k 的 χ^2 分布:

$$p_X(x)=\begin{cases}\dfrac{1}{2^{\frac{k}{2}}\Gamma\left(\dfrac{k}{2}\right)}x^{\frac{k}{2}-1}e^{-\frac{x}{2}}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

求 $Y=\sqrt{X/k}$ 的密度.

解 对于 $y>0$,有

$$P(Y\leqslant y)=P(\sqrt{X/k}\leqslant y)=P(X\leqslant ky^2)=F(ky^2).$$

两边求导,得

$$p_Y(y)=p_X(ky^2)\cdot 2ky=2\left(\frac{k}{2}\right)^{\frac{k}{2}}y^{k-1}e^{-\frac{ky^2}{2}}\Big/\Gamma\left(\frac{k}{2}\right).$$

对于 $y\leqslant 0$,显然有 $p_Y(y)=0$,故

$$p_Y(y)=\begin{cases}2\left(\dfrac{k}{2}\right)^{\frac{k}{2}}y^{k-1}e^{-\frac{ky^2}{2}}\Big/\Gamma\left(\dfrac{k}{2}\right), & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

9. 由统计物理学知道分子运动的速率 X 服从麦克斯韦(Maxwell)分布,即密度为

$$p_X(x)=\begin{cases}\dfrac{4x^2}{\alpha^3\sqrt{\pi}}e^{-\frac{x^2}{\alpha^2}}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

其中参数 $\alpha>0$. 求分子的动能 $Y=\frac{1}{2}mX^2$ 的密度.

解 对于 $y>0$,有

$$P(Y\leqslant y)=P\left(\frac{1}{2}mX^2\leqslant y\right).$$

因为 $p_X(x)$ 在 $x\leqslant 0$ 处为 0,即 X 只可能取正值,所以,对 $y>0$,

$$\text{事件“}\frac{1}{2}mX^2\leqslant y\text{”}=\text{事件“}X\leqslant\sqrt{\frac{2y}{m}}\text{”},$$

因此有

$$P(Y\leqslant y)=P\left(\frac{1}{2}mX^2\leqslant y\right)=P\left(X\leqslant\sqrt{\frac{2y}{m}}\right).$$

两边求导，得

$$p_Y(y)=p_X\left(\sqrt{\frac{2y}{m}}\right)\cdot\sqrt{\frac{2}{m}}\cdot\frac{1}{2}\cdot\frac{1}{\sqrt{y}}$$

$$=\frac{4\cdot\dfrac{2y}{m}}{\alpha^3\cdot\sqrt{\pi}}\mathrm{e}^{-\frac{2y}{\alpha^2 m}}\cdot\sqrt{\frac{1}{2m}}\frac{1}{\sqrt{y}}$$

$$=\frac{4\sqrt{2}}{\alpha^3\cdot\sqrt{\pi}m^{\frac{3}{2}}}y^{\frac{1}{2}}\mathrm{e}^{-\frac{2y}{m\alpha^2}}.$$

显然 $y\leqslant 0$ 时，$p_Y(y)=0$.（它是一个 Γ 分布.相应的参数为何？请读者自己思考.）

10. 设 $\ln X\sim N(1,2^2)$，求 $P\left(\frac{1}{2}<X<2\right)$.（$\ln 2=0.693$）.

解 $P\left(\frac{1}{2}<X<2\right)=P(-\ln 2<\ln X<\ln 2)$

$$=P\left(\frac{-\ln 2-1}{2}<\frac{\ln X-1}{2}<\frac{\ln 2-1}{2}\right)$$

$$=P\left(-0.8465<\frac{\ln X-1}{2}<-0.1535\right)$$

$$=\Phi(0.8465)-\Phi(0.1535)=0.2403.$$

11. 对球的直径作测量，设其均匀地分布在$[a,b]$内，求体积的密度函数.

解 设 X 是直径，X 是$[a,b]$内均匀分布的随机变量，因此其密度为

$$p_X(x)=\begin{cases}\dfrac{1}{b-a}, & x\in[a,b],\\ 0, & 其他.\end{cases}$$

球的体积 $Y=\frac{\pi}{6}X^3$. 它的分布函数是

$$P(Y\leqslant y)=P\left(\frac{\pi}{6}X^3\leqslant y\right)=P\left(X\leqslant\sqrt[3]{\frac{6y}{\pi}}\right).$$

求导，有

$$p_Y(y)=p_X\left(\sqrt[3]{\frac{6y}{\pi}}\right)\sqrt[3]{\frac{6}{\pi}}\frac{1}{3}y^{-\frac{2}{3}}$$

$$=\begin{cases}\dfrac{1}{b-a}\left(\dfrac{2}{9\pi}\right)^{\frac{1}{3}}y^{-\frac{2}{3}}, & 当\dfrac{\pi a^3}{6}\leqslant y\leqslant\dfrac{\pi b^3}{6},\\ 0, & 其他.\end{cases}$$

(上式右边的"当$\dfrac{\pi a^3}{6}\leqslant y\leqslant\dfrac{\pi b^3}{6}$",从直观看很显然;而从形式逻辑讲,它是由$\sqrt[3]{\dfrac{6y}{\pi}}\in[a,b]$而来.)

12. 点随机地落在中心在原点、半径为 R 的圆周上,并且对弧长是均匀地分布的.求落点的横坐标的概率密度.

解 设 Θ 为圆周上随机点 M 与原点连线和 Ox 轴的夹角,则 Θ 为随机变量;按题设,它在$[0,2\pi)$上均匀分布(弧长均匀分布与夹角均匀分布是等价的,它们之比是常数即 $S=R\Theta$),因此,其密度

$$p(\theta)=\begin{cases}\dfrac{1}{2\pi}, & 0\leqslant\theta<2\pi,\\ 0, & 其他.\end{cases}$$

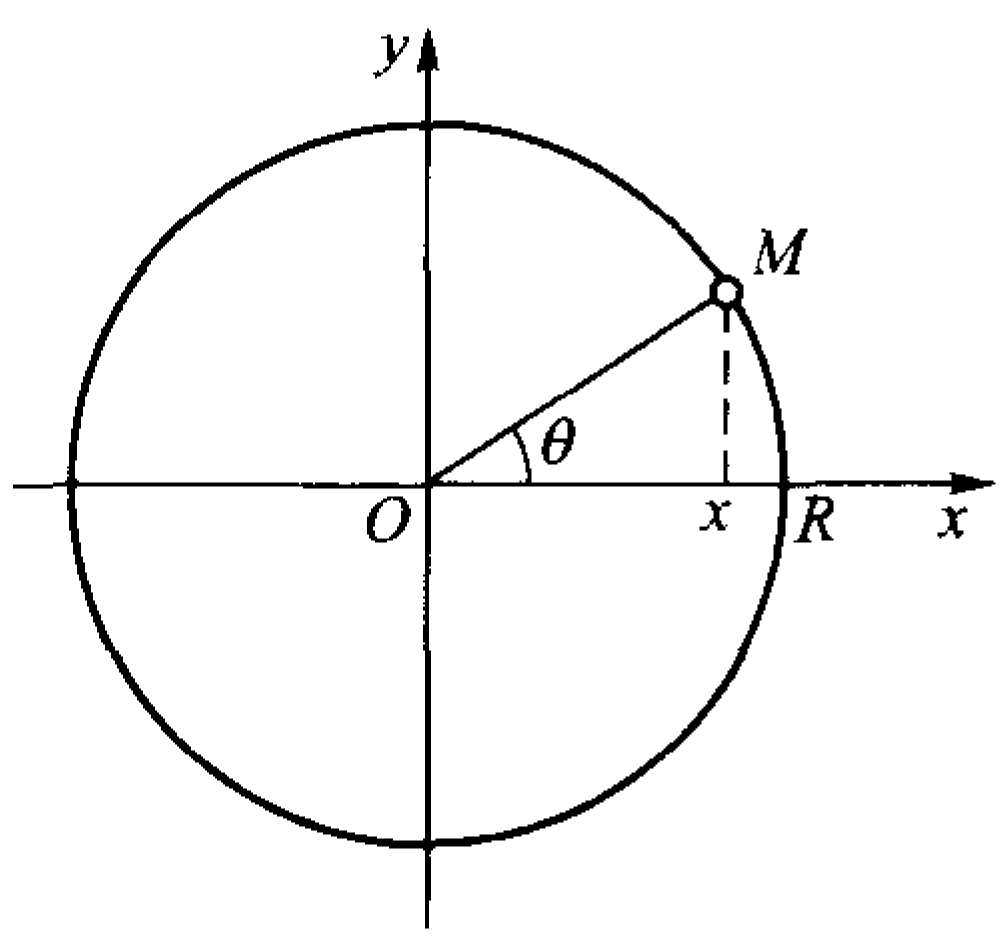

图 4

设 M 点的横坐标为 X,则有 $X=R\cos\Theta$(即 X 是随机变量 Θ 的函数,而 Θ 的密度已知,要求 X 的密度).还用分布函数法.对$|x|<$

R,我们有

$$F_X(x)=P(X\leqslant x)=P(R\cos\Theta\leqslant x)=P\left(\cos\Theta\leqslant\frac{x}{R}\right)$$

$$=P\left(\arccos\frac{x}{R}\leqslant\Theta\leqslant 2\pi-\arccos\frac{x}{R}\right)$$

$$\xlongequal{\text{由于}(\triangle)}\frac{1}{2\pi}\left[\left(2\pi-\arccos\frac{x}{R}\right)-\arccos\frac{x}{R}\right]$$

$$=1-\frac{1}{\pi}\arccos\frac{x}{R}.$$

两边求导,得

$$p_X(x)=-\frac{1}{\pi}\left(-\frac{1}{\sqrt{1-\left(\frac{x}{R}\right)^2}}\cdot\frac{1}{R}\right)=\frac{1}{\pi\sqrt{R^2-x^2}}.$$

当$|x|>R$时,显然$p_X(x)=0$,故

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

评注 在实际问题中,如果要找一个较为复杂的随机变量的密度,首先要寻找一个这样的随机变量:它的分布规律比较简单,又跟所关注的随机变量有函数关系,比如本题中的Θ.

13. 设随机变量X的分布函数为

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

(1) 求$P(X\leqslant 2)$,$P(X>3)$;

(2) 求X的密度函数$p(x)$.

解 (1) $P(X\leqslant 2)=F(2)=1-e^{-2}=0.864\,7.$

$P(X>3)=1-P(X\leqslant 3)=1-(1-e^{-3})=e^{-3}=0.049\,8.$

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

14. 设随机变量X的分布密度为

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

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

求 X 的分布函数 $F(x)$，并作出(2)中 $p(x)$ 与 $F(x)$ 的图形.

解 (1) 因 $F(x)=\int_{-\infty}^{x}p(t)\mathrm{d}t$，

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

当 $-1<x<1$ 时，

$$F(x)=\frac{2}{\pi}\int_{-1}^{x}\sqrt{1-t^2}\mathrm{d}t=\frac{2}{\pi}\cdot\frac{1}{2}\left(t\sqrt{1-t^2}+\arcsin t\right)\Big|_{-1}^{x}=\frac{1}{\pi}x\sqrt{1-x^2}+\frac{1}{\pi}\arcsin x+\frac{1}{2};$$

当 $x\geqslant 1$ 时， $F(x)=\dfrac{2}{\pi}\int_{-1}^{1}\sqrt{1-t^2}\mathrm{d}t=1$，

所以

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

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

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

当 $0\leqslant x<1$ 时，$F(x)=\int_{0}^{x}t\mathrm{d}t=\dfrac{x^2}{2}$；

当 $1\leqslant x<2$ 时，$F(x)=\int_{0}^{1}t\mathrm{d}t+\int_{1}^{x}(2-t)\mathrm{d}t=\dfrac{1}{2}+2t\Big|_{1}^{x}-\dfrac{t^2}{2}\Big|_{1}^{x}$

$$=2x-\frac{x^2}{2}-1;$$

当 $x\geqslant 2$ 时，$F(x)=1$.

综合得

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

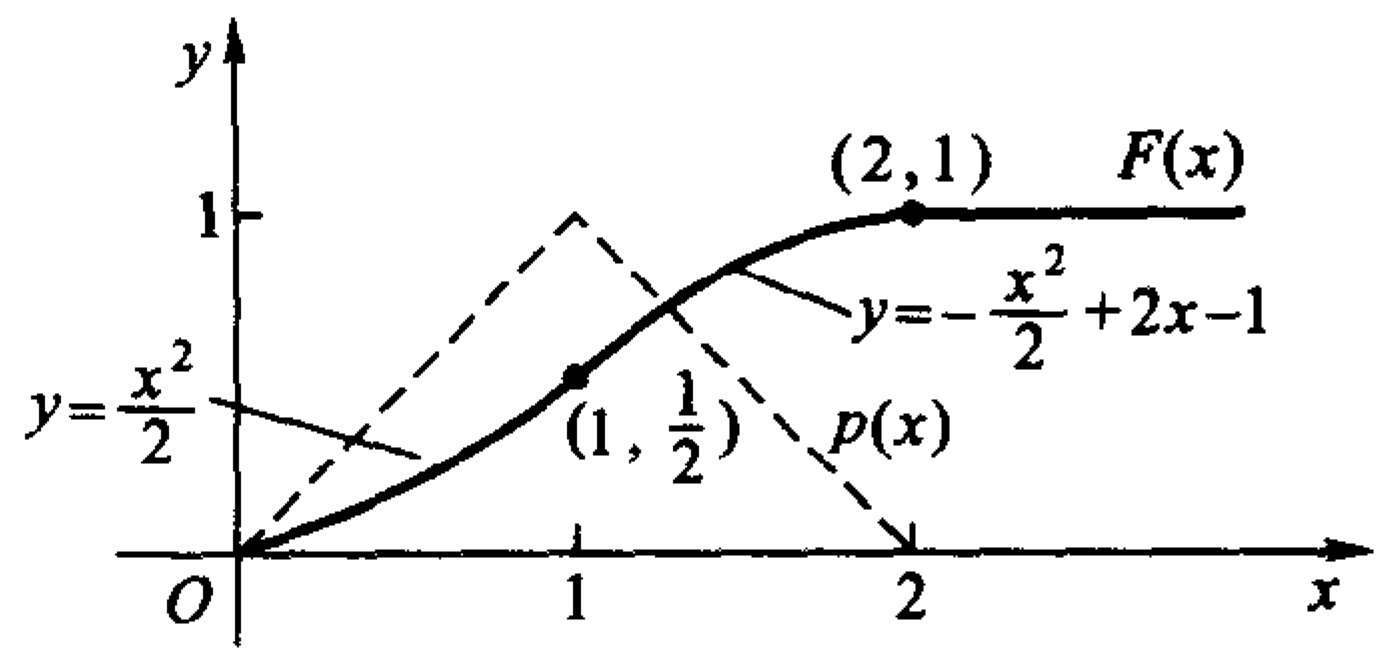

图 5

评注 从 $p(x)$ 来求 $F(x)$，说来简单，做起来有点麻烦. 这是因为 $p(x)$ 一般都是分段表示的函数. 至于区间端点是否带等号，倒不必在意，因为有密度的所谓连续型随机变量，其分布函数至少是连续的.

15. 某产品的质量指标 $X\sim N(160,\sigma^2)$，若要求

$$P(120<X<200)\geqslant 0.80,$$

问：允许 σ 最多为多少？

解 令 $Y=\dfrac{X-160}{\sigma}$，则 $Y\sim N(0,1)$.

$$P(120<X<200)=P\left(\frac{120-160}{\sigma}<\frac{X-160}{\sigma}<\frac{200-160}{\sigma}\right)$$

$$=P\left(-\frac{40}{\sigma}<Y<\frac{40}{\sigma}\right)$$

$$=2\left[\Phi\left(\frac{40}{\sigma}\right)-0.5\right]\geqslant 0.80,$$

得
$$\Phi\left(\frac{40}{\sigma}\right)\geqslant 0.9.$$

查附表 1,有$\frac{40}{\sigma}\geqslant 1.28$,即$\sigma\leqslant 31.25$. σ最多为 31.25.

16. 如果 X 的分布函数 $F(x)$具有连续的导函数 $F'(x)$,试证:$F'(x)$是 X 的分布密度.

证 对于任意常数 $a,b(a<b)$,按分布函数定义有

$$P(a<X\leqslant b)=F(b)-F(a).$$

对上式右边,由微积分基本公式,有

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

因此

$$P(a<X\leqslant b)=\int_a^b F'(x)\mathrm{d}x.$$

这表明,$F'(x)$为 X 的密度函数.

17*. 设 X 的分布函数 $F(x)$满足下列条件:

① $F(x)$连续;

② 存在 $x_1<x_2<\cdots<x_n\ (n\geqslant 1)$. 在区间$(-\infty,x_1)$,$(x_1,x_2)$,…,$(x_{n-1},x_n)$,$(x_n,+\infty)$上 $F'(x)$存在且连续.

令

$$f(x)=\begin{cases}F'(x), & \text{当 } F'(x)\text{存在时},\\ 0, & \text{当 } F'(x)\text{不存在时}.\end{cases}$$

试证:$f(x)$是 X 的分布密度

证 对于任意 $a<b$,若(a,b)内不含 $x_1,x_2,\cdots,x_n$ 中任一点,即(a,b)被②中$(n+1)$个区间的某个区间包含,则由微积分基本公式,有

$$P(a<X\leqslant b)=\int_a^b F'(x)\mathrm{d}x$$

因此

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

若(a,b)内含 $x_1,x_2,\cdots,x_n$ 中的点，不妨记 x_i 为其中的最小值，x_j 为其中的最大值，则有

$$\begin{aligned}
P(a<X\leqslant b) &= F(b)-F(a) \\
&= F(x_i)-F(a)+F(x_{i+1})-F(x_i)+\cdots+F(b)-F(x_j) \\
&= \int_a^{x_i} F'(x)\mathrm{d}x+\int_{x_i}^{x_{i+1}} F'(x)\mathrm{d}x+\cdots+\int_{x_j}^{b} F'(x)\mathrm{d}x \\
&= \int_a^{x_i} f(x)\mathrm{d}x+\int_{x_i}^{x_{i+1}} f(x)\mathrm{d}x+\cdots+\int_{x_j}^{b} f(x)\mathrm{d}x \\
&= \int_a^b f(x)\mathrm{d}x,
\end{aligned}$$

即对任 $a<b$，都有

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

这表明 $f(x)$是 X 的分布密度.

18*. 设随机变量 X 的分布函数不是连续函数，试证明：随机变量 $Y=F(X)$一定不服从$(0,1)$上的均匀分布.

证 事实上，由概率分布函数 $F(x)$的性质(2)(即“不减”)知，$F(x)$在任意点 x_0 处，其左右极限都存在，即

$$F(x_0-0)=\lim_{x\to x_0^-}F(x),$$

$$F(x_0+0)=\lim_{x\to x_0^+}F(x).$$

已知 $F(x)$不是连续函数，这表明至少存在一个 x_0 使 $F(x_0-0)\neq F(x_0+0)$，即 $F(x_0-0)<F(x_0+0)$. 因此，$Y=F(X)$这个随机变量取不到$(F(x_0-0),F(x_0+0))$中的值，所以它不服从$(0,1)$上的均匀分布.

习 题 八

1. 已知随机变量 X 的概率分布为

$$P(X=k)=\frac{1}{10},\quad k=2,4,\cdots,18,20,$$

求 $E(X)$.

解 由定义(本章(1.2)式)有

$$E(X)=\frac{2}{10}+\frac{4}{10}+\cdots+\frac{20}{10}=\frac{2}{10}(1+2+\cdots+10)=11.$$

2. 两台生产同一零件的车床,一天生产的产品中次品数的概率分布分别是

甲	0	1	2	3
p	0.4	0.3	0.2	0.1

乙	0	1	2	3
p	0.3	0.5	0.2	0

如果两台机床的产量相同,问哪台机床好?

解 记X=甲机床一天内生产的次品数,

Y=乙机床一天内生产的次品数,

则

$$E(X)=0\times0.4+1\times0.3+2\times0.2+3\times0.1=1,$$

$$E(Y)=0\times0.3+1\times0.5+2\times0.2+3\times0=0.9,$$

$E(Y)<E(X)$,故乙机床好.

3. 某射手每次射击击中目标的概率都是 0.8,现连续向目标射击,直到第一次击中为止.求“射击次数”X 的期望.

解 首先 X 的概率分布为

$$P(X=k)=(0.2)^{k-1}0.8,k=1,2,\cdots;$$

$$E(X)=\sum_{k=1}^{\infty}kP(X=k)=\sum_{k=1}^{\infty}0.8k(0.2)^{k-1}$$

$$=0.8\sum_{k=1}^{\infty}k(0.2)^{k-1}.$$

由级数逐项微商性质，有

$$\sum_{k=1}^{\infty} kx^{k-1} = \left(\sum_{k=1}^{\infty} x^k\right)' = \left(\sum_{k=0}^{\infty} x^k\right)' = \left(\frac{1}{1-x}\right)' = \frac{1}{(1-x)^2}.$$

得 $$E(X) = 0.8\sum_{k=1}^{\infty} k(0.2)^{k-1} = 0.8 \cdot \frac{1}{(1-0.2)^2} = \frac{1}{0.8} = 1.25.$$

4. 推导超几何分布的期望计算公式.

解 X 服从参数为 n,N,M 的超几何分布，即

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

$$E(X) = \sum_{k=0}^{l} k\frac{C_M^k C_{N-M}^{n-k}}{C_N^n} = \sum_{k=1}^{l}\frac{M \cdot C_{M-1}^{k-1} C_{(N-1)-(M-1)}^{(n-1)-(k-1)}}{C_{N-1}^{n-1} \cdot \dfrac{N}{n}}$$

$$= n\frac{M}{N}\sum_{k-1=0}^{l-1}\frac{C_{M-1}^{k-1} C_{(N-1)-(M-1)}^{(n-1)-(k-1)}}{C_{N-1}^{n-1}} = n\frac{M}{N}.$$

这是因为任一概率分布的总概率等于 1，参数为 $n-1,N-1,M-1$的超几何分布也不例外.

5. 盒中有五个球，三白二黑. 从中随机抽取两个球，求“抽得的白球数”X 的期望.

解 X 服从参数为 $n=2,N=5,M=3$ 的超几何分布，按上题计算公式，有

$$E(X)=n\frac{M}{N}=\frac{6}{5}=1.2.$$

6. 射击比赛，每人射四次(每次一发)，约定全部不中得 0 分，只中一弹得 15 分，中二弹得 30 分，中三弹得 55 分，中四弹得 100 分. 甲每次射击命中率为 3/5，问他期望能得多少分?

解 令 $X=$“甲的得分”. 按独立试验序列模型，X 的概率分布：

$$P(X=0)=C_4^0\left(\frac{3}{5}\right)^0\left(\frac{2}{5}\right)^4=\frac{16}{625},$$

$$P(X=15)=C_4^1\left(\frac{3}{5}\right)^1\left(\frac{2}{5}\right)^3=\frac{4\times 24}{625},$$

$$P(X=30)=C_4^2\left(\frac{3}{5}\right)^2\left(\frac{2}{5}\right)^2=\frac{6\times 36}{625},$$

$$P(X=55)=C_4^3\left(\frac{3}{5}\right)^3\left(\frac{2}{5}\right)^1=\frac{8\times 27}{625},$$

$$P(X=100)=C_4^4\left(\frac{3}{5}\right)^4\left(\frac{2}{5}\right)^0=\frac{81}{625}.$$

由期望公式得

$$E(X)=44.64.$$

7*. 某射手每次射中目标的概率是 p，现携有 10 发子弹准备对一目标连续射击(每次打一发)，一旦射中或子弹打完了就立刻转移到别的地方. 问：他在转移前平均射击几次？

解 令 X=转移前射击次数，它的可能值为 $1,2,\cdots,10$. 先来找出它的分布：

$$P(X=k)=(1-p)^{k-1}p,k=1,2,\cdots,9,$$
$$P(X=10)=(1-p)^9.$$

(上面两个等式的第一个，跟本章习题 3 的分析相同，而第二个等式是因为"$X=10$"="前 9 次都不中".)平均射击次数

$$E(X)=\sum_{k=1}^{9}k(1-p)^{k-1}\cdot p+10(1-p)^9.$$

但

$$\sum_{k=1}^{9}kx^{k-1}=\left[\sum_{k=1}^{9}x^k\right]'=\left[\sum_{k=0}^{9}x^k\right]'=\left(\frac{1-x^{10}}{1-x}\right)'$$
$$=\frac{1-10x^9+9x^{10}}{(1-x)^2},$$

因此

$$E(X)=p\,\frac{1-10(1-p)^9+9(1-p)^{10}}{p^2}+10(1-p)^9$$
$$=\frac{1}{p}[1-(1-p)^{10}].$$

习 题 九

1. 设随机变量 X 的概率密度为

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

求 $E(X)$.

解 $E(X)=\int_{-\infty}^{+\infty}xp(x)\mathrm{d}x=\int_0^1 x\cdot 2x\mathrm{d}x=\frac{2}{3}x^3\Big|_0^1=\frac{2}{3}$.

2. 设随机变量 X 的密度为

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

求 $E(X)$.

解 $E(X)=\int_{-\infty}^{+\infty}xp(x)\mathrm{d}x=\int_{-1}^{1}\frac{x}{\pi\sqrt{1-x^2}}\mathrm{d}x$.

由对称性(被积函数为奇函数,积分区间原点对称),积分值为零,即

$$E(X)=0.$$

3. 设随机变量 X 的概率密度为

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

求 $E(X)$.

解 因该密度为偶函数,故 $E(X)=0$.

4. 设 $X\sim N(0,\sigma^2)$,求 $E(X^n)$.

解 $E(X^n)=\int_{-\infty}^{+\infty}x^n\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\mathrm{d}x$.

当 n 为奇数时,被积函数为奇函数,则 $E(X^n)=0$;当 $n=2k$ 时,

$$E(X^n)=2\int_0^{+\infty}x^{2k}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}dx \quad \left(令\frac{x}{\sigma}=t\right)$$

$$=\frac{2\sigma^{2k}}{\sqrt{2\pi}}\int_0^{+\infty}t^{2k}e^{-\frac{t^2}{2}}dt \quad \left(令\frac{t^2}{2}=u\right)$$

$$=\frac{\sigma^{2k}2^k}{\sqrt{\pi}}\int_0^{+\infty}u^{k-\frac{1}{2}}e^{-u}du$$

$$=\frac{\sigma^{2k}2^k}{\sqrt{\pi}}\cdot\Gamma\left(k+\frac{1}{2}\right).$$

由 Γ 函数的递推公式与 $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$，有

$$\Gamma\left(k+\frac{1}{2}\right)=\left(k-\frac{1}{2}\right)\left(k-\frac{3}{2}\right)\cdots\frac{1}{2}\Gamma\left(\frac{1}{2}\right)$$

$$=\frac{(2k-1)!!}{2^k}\sqrt{\pi}.$$

得

$$E(X^{2k})=\sigma^{2k}(2k-1)!!$$

5. 对球的直径作近似测量，设其值均匀地分布在区间$[a,b]$内，求球体积的均值.

解 直径 X 是一随机变量，其密度为

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

体积 $Y=\frac{1}{6}\pi X^3$ 的均值

$$E(Y)=\int_a^b\frac{1}{6}\pi x^3\frac{1}{b-a}dx=\frac{\pi}{24}(b+a)(b^2+a^2).$$

评注 本题利用随机变量函数的期望公式进行计算. 非常方便，它不必找出 Y 的密度函数.

6. 点随机地落在中心在原点、半径为 R 的圆周上，并对弧长是均匀分布的. 求落点横坐标的均值.

解 参见习题七第 12 题的解. 我们知道落点横坐标 X 与落点极角 Θ 有如下关系：

$$X=R\cos\Theta.$$

Θ 在$[0,2\pi)$上均匀分布,因此

$$E(X)=\int_0^{2\pi}R\cos\theta\cdot\frac{1}{2\pi}\mathrm{d}\theta=0.$$

7. 设 X 的密度函数 $p(x)$满足

$$p(c+x)=p(c-x),\quad (x>0)$$

其中 c 为一常数,又 $\int_{-\infty}^{+\infty}|x|p(x)\mathrm{d}x$ 收敛,求证:

$$E(X)=c.$$

解 $E(X)-c=\int_{-\infty}^{+\infty}(x-c)p(x)\mathrm{d}x\quad$ (令 $y=x-c$)

$$=\int_{-\infty}^{+\infty}yp(y+c)\mathrm{d}y.$$

记 $g(y)\equiv p(y+c)$,则由题设知

$$g(-y)=p(-y+c)=p(c+y)=g(y).$$

这表明 $p(y+c)$是偶函数,所以有

$$\int_{-\infty}^{+\infty}yp(y+c)\mathrm{d}y=0,$$

因此

$$E(X)=c.$$

习 题 十

1. 对于习题八的习题 1、3 中的随机变量，分别求出它们的方差.

解 对习题八第 1 题，$P(X=k)=\frac{1}{10}$，$k=2,4,\cdots,20$，因此，

$$E(X^2)=\sum x_i^2 p_i=2^2\times\frac{1}{10}+4^2\times\frac{1}{10}+\cdots+20^2\times\frac{1}{10}$$

$$=\frac{2^2}{10}(1^2+2^2+\cdots+10^2)=\frac{4}{10}\times\frac{10\times 11\times 21}{6}=154.$$

前已求出 $E(X)=11$，所以

$$D(X)=E(X^2)-[E(X)]^2=33.$$

对习题八第 3 题，$P(X=k)=(0.2)^{k-1}\cdot 0.8$，$k=1,2,\cdots$，因此 $E(X^2)=0.8\times\sum_{k=1}^{\infty}k^2(0.2)^{k-1}$ $\left(\text{前已求出 } E(X)=\frac{1}{0.8}=1.25\right)$.

现求 $\sum_{k=1}^{\infty}k^2x^{k-1}$ 的一般公式.

$$\sum_{k=1}^{\infty}k^2x^{k-1}=\sum_{k=1}^{\infty}(k+1)kx^{k-1}-\sum_{k=1}^{\infty}kx^{k-1}$$

$$=\left(\sum_{k=1}^{\infty}x^{k+1}\right)''-\left(\sum_{k=1}^{\infty}x^k\right)'=\left(\frac{1}{1-x}\right)''-\left(\frac{1}{1-x}\right)'$$

$$=\frac{2}{(1-x)^3}-\frac{1}{(1-x)^2}=\frac{1+x}{(1-x)^3}.$$

将 $x=0.2$ 代入得

$$E(X^2)=0.8\times\frac{1.2}{0.8^3}=\frac{1.2}{0.8^2},$$

$$D(X)=E(X^2)-E^2(X)=\frac{1.2}{0.8^2}-\frac{1}{0.8^2}=0.312\,5.$$

2. 对于习题九的1,2,3,6题中的随机变量,分别求出它们的方差.

解 (1) 习题九第1题

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

已求出 $E(X)=\frac{2}{3}$,现求 $E(X^2)$.

$$E(X^2)=\int_{-\infty}^{+\infty}x^2p(x)\mathrm{d}x=\int_0^1 x^2\cdot 2x\mathrm{d}x=\frac{2}{4}x^4\Big|_0^1=\frac{1}{2},$$

故

$$D(X)=\frac{1}{2}-\left(\frac{2}{3}\right)^2=\frac{1}{18}.$$

(2) 习题九第2题

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

已求出 $E(X)=0$,因此

$$D(X)=E(X^2)=\int_{-1}^1 x^2\frac{1}{\pi\sqrt{1-x^2}}\mathrm{d}x=\frac{2}{\pi}\int_0^1\frac{x^2}{\sqrt{1-x^2}}\mathrm{d}x$$

$$(x=\sin t)=\frac{2}{\pi}\int_0^{\frac{\pi}{2}}\frac{\sin^2 t}{\cos t}\cos t\mathrm{d}t$$

$$=\frac{2}{\pi}\cdot\int_0^{\frac{\pi}{2}}\frac{1-\cos 2t}{2}\mathrm{d}t=\frac{1}{2}.$$

(3) 习题九第3题

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

已求出 $E(X)=0$,因此

$$D(X)=E(X^2)=\frac{1}{2}\int_{-\infty}^{+\infty}x^2\mathrm{e}^{-|x|}\mathrm{d}x=\int_0^{+\infty}x^2\mathrm{e}^{-x}\mathrm{d}x=\Gamma(3)=2.$$

(4) 习题九第6题

$$X=R\cos\Theta,\quad \Theta\text{ 在}[0,2\pi)\text{上均匀分布}.$$

已求出 $E(X)=0$，因此

$$D(X)=E(X^2)=\int_0^{2\pi}R^2\cos^2\theta\,\frac{1}{2\pi}\mathrm{d}\theta=\frac{R^2}{2\pi}\int_0^{2\pi}\frac{1+\cos 2\theta}{2}\mathrm{d}\theta$$

$$=\frac{R^2}{2\pi}\cdot\frac{2\pi}{2}=\frac{R^2}{2}.$$

3. 设 X 服从参数为 $N,M,n(n\leqslant N-M)$ 的超几何分布，即

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

试证明

$$D(X)=\frac{nM(N-n)(N-M)}{N^2(N-1)}.$$

解 由习题八第 4 题知 $E(X)=\dfrac{nM}{N}$，用类似方法求 $E(X^2)$.

$$E(X^2)=\sum_{k=1}^{l}k^2\,\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}=\sum_{k=1}^{l}[k(k-1)+k]\,\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$$

$$=\sum_{k=2}^{l}k(k-1)\,\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}+E(X)$$

$$=\left[\sum_{k=2}^{l}\frac{C_{M-2}^{k-2}C_{(N-2)-(M-2)}^{(n-2)-(k-2)}}{C_{N-2}^{n-2}}\right]\frac{n(n-1)\cdot M(M-1)}{N(N-1)}+\frac{nM}{N}.$$

利用上式中方括号内等于 1（将求和指标 k 换为 $k'=k-2$，而记 $l'=l-2$，于是就成为参数是 $N-2,M-2,n-2$ 的超几何分布的总概率和），再由 $D(X)=E(X^2)-(E(X))^2$ 整理即得.

4. 已知 $X\sim N(\mu,\sigma^2)$，$Y=\mathrm{e}^X$，求 Y 的分布密度，并计算 $E(Y)$ 和 $D(Y)$.

解 由 $F_Y(y)=P(Y\leqslant y)=P(\mathrm{e}^X\leqslant y)=P(X\leqslant\ln y)$，$y>0$ 得

$$p_Y(y)=F_Y'(y)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(\ln y-\mu)^2}{2\sigma^2}}\cdot\frac{1}{y},\quad y>0.$$

$$E(Y)=E(\mathrm{e}^X)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}\mathrm{e}^x\cdot\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x$$

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

注意到上式中的方括号内为

$$(x-\mu)^2-2\sigma^2 x=[x-(\mu+\sigma^2)]^2-2\mu\sigma^2-\sigma^4.$$

得

$$E(Y)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{1}{2\sigma^2}[x-(\mu+\sigma^2)]^2}\cdot\mathrm{e}^{\mu+\frac{\sigma^2}{2}}\mathrm{d}x$$

$$=\mathrm{e}^{\mu+\frac{\sigma^2}{2}}.$$

类似可得

$$E(Y^2)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{1}{2\sigma^2}[x-(\mu+2\sigma^2)]^2}\mathrm{e}^{2\mu+2\sigma^2}\mathrm{d}x$$

$$=\mathrm{e}^{2\mu+2\sigma^2}.$$

所以

$$D(Y)=E(Y^2)-E^2(Y)=\mathrm{e}^{2\mu+\sigma^2}(\mathrm{e}^{\sigma^2}-1).$$

5．设轮船横向摇摆的随机振幅 X 的概率密度为

$$p(x)=Ax\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\quad(x>0).$$

求(1) A;(2) 遇到大于其振幅均值的概率是多少？(3) X 的方差.

解 (1) $\int_{-\infty}^{+\infty}p(x)\mathrm{d}x=\int_0^{+\infty}Ax\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\mathrm{d}x=A\int_0^{+\infty}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\cdot\sigma^2\mathrm{d}\left(\frac{x^2}{2\sigma^2}\right)$

$$=A\sigma^2\int_0^{+\infty}\mathrm{e}^{-t}\mathrm{d}t=A\sigma^2.$$

于是

$$A=\frac{1}{\sigma^2}.$$

(2) 先求均值 $E(X)$.

$$E(X)=\int_0^{+\infty}\frac{1}{\sigma^2}x^2\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\mathrm{d}x\quad\text{（利用正态密度的方差）}$$

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

$$=\sqrt{\frac{\pi}{2}}\sigma.$$

$$P\left(X>\sqrt{\frac{\pi}{2}}\sigma\right)=\int_{\sqrt{\frac{\pi}{2}}\sigma}^{+\infty}\frac{1}{\sigma^2}x\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\mathrm{d}x=-\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\Big|_{\sqrt{\frac{\pi}{2}}\sigma}^{+\infty}$$

$$=\mathrm{e}^{-\frac{\pi}{4}}.$$

(3) $$E(X^2)=\int_0^{+\infty}\frac{x^3}{\sigma^2}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\mathrm{d}x=2\sigma^2\int_0^{+\infty}\frac{x^2}{2\sigma^2}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\mathrm{d}\left(\frac{x^2}{2\sigma^2}\right)$$

$$=2\sigma^2\int_0^{+\infty}t\mathrm{e}^{-t}\mathrm{d}t=2\sigma^2,$$

所以

$$D(X)=E(X^2)-E^2(X)=2\sigma^2-\frac{\pi}{2}\sigma^2=\sigma^2\left(2-\frac{\pi}{2}\right).$$

6. 设 X 的密度为

$$p(x)=\frac{x^m}{m!}\mathrm{e}^{-x},\quad x\geqslant 0.$$

试证

$$P(0<X<2(m+1))\geqslant\frac{m}{m+1}.$$

证 $$E(X)=\int_0^{+\infty}x\frac{x^m}{m!}\mathrm{e}^{-x}\mathrm{d}x=\frac{1}{m!}\int_0^{+\infty}x^{(m+2)-1}\mathrm{e}^{-x}\mathrm{d}x$$

$$=\frac{1}{m!}\Gamma(m+2)=\frac{(m+1)!}{m!}=m+1.$$

$$E(X^2)=\int_0^{+\infty}x^2\frac{x^m}{m!}\mathrm{e}^{-x}\mathrm{d}x=\frac{1}{m!}\Gamma(m+3)=(m+2)(m+1).$$

因此

$$D(X)=(m+2)(m+1)-(m+1)^2=m+1.$$

由切比雪夫不等式

$$P(|X-E(X)|\geqslant\varepsilon)\leqslant\frac{D(X)}{\varepsilon^2}\quad(\varepsilon>0)$$

得

$$P(|X-E(X)|<\varepsilon)\geqslant 1-\frac{D(X)}{\varepsilon^2}.$$

取 $\varepsilon=m+1$，对该随机变量 X 有

$$P(|X-(m+1)|<m+1)\geqslant 1-\frac{m+1}{(m+1)^2},$$

即

$$P(0<X<2(m+1))\geqslant \frac{m}{m+1}.$$

7. 设 $X\sim$ 贝塔分布，即它的密度为

$$p(x)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1},\quad 0<x<1.\quad (\alpha>0,\beta>0)$$

求 $E(X)$，$D(X)$.

解 我们知道，对 $B(\alpha,\beta)\equiv\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\mathrm{d}x$，有 $B(\alpha,\beta)=\frac{\Gamma(\alpha)\cdot\Gamma(\beta)}{\Gamma(\alpha+\beta)}$. 另外，对 $\Gamma(\alpha)$ 有递推公式 $\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$.
因此有，

$$\begin{aligned}E(X)&=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 x^{\alpha}(1-x)^{\beta-1}\mathrm{d}x\\&=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot\frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+1+\beta)}\\&=\frac{\alpha}{\alpha+\beta}.\end{aligned}$$

类似有，

$$E(X^2)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 x^{\alpha+1}(1-x)^{\beta-1}\mathrm{d}x=\frac{(\alpha+1)\alpha}{(\alpha+\beta+1)(\alpha+\beta)}.$$

因此

$$D(X)=E(X^2)-E^2(X)=\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}.$$

8. 对一目标进行射击，直到击中 r 次为止. 如果每次射击的命中率为 p，求需射击次数的均值与方差.

解 用 X 表示需射击次数，则可能值为 $r,r+1,\cdots$. 它的分布是(为什么？可先对 $r=3,k=6$ 来具体分析射击记录).

$$P(X=k)=C_{k-1}^{r-1}p^r q^{k-r},\quad k=r,r+1,\cdots.$$

因为 X 具有实际背景，所以有

$$\sum_{k=r}^{\infty} C_{k-1}^{r-1} p^r q^{k-r} = 1. \quad (0.4)$$

(上式也可利用 $(1-x)^{-r}$ 的泰勒展开式来证：

$$(1-x)^{-r} = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k = \sum_{k=0}^{\infty} \frac{(k+r-1)!}{(r-1)!k!} x^k$$

$$= \sum_{k=0}^{\infty} C_{k+r-1}^{r-1} x^k .)$$

现来求

$$E(X) = \sum_{k=r}^{\infty} k \cdot C_{k-1}^{r-1} p^r q^{k-r} = \sum_{k=r}^{\infty} k \frac{(k-1)!}{(r-1)!(k-r)!} p^r q^{k-r}$$

$$= r \cdot \sum_{k=r}^{\infty} \frac{k!}{r!(k-r)!} p^{r+1} \cdot p^{-1} q^{k-r}$$

$$= \frac{r}{p} \cdot \sum_{k+1=r+1}^{\infty} C_{(k+1)-1}^{(r+1)-1} p^{r+1} q^{(k+1)-(r+1)}$$

$$\xlongequal{\text{由(0.4)式}} \frac{r}{p}.$$

类似可求得 $E(X^2)=\frac{r(q+r)}{p^2}$，因此得

$$D(X)=E(X^2)-(E(X))^2=\frac{r(q+r)}{p^2}-\left(\frac{r}{p}\right)^2=\frac{rq}{p^2}.$$

习题十一

1. 离散型随机向量(X,Y)有如下的概率分布，求边缘分布. 又，X,Y是否独立？

$X \backslash Y$	0	1	2	3	4	5	6	
0	0.202	0.174	0.113	0.062	0.049	0.023	0.004	**0.627**
1	0	0.099	0.064	0.040	0.031	0.020	0.006	**0.260**
2	0	0	0.031	0.025	0.018	0.013	0.008	**0.095**
3	0	0	0	0.001	0.002	0.004	0.011	**0.018**
	0.202	**0.273**	**0.208**	**0.128**	**0.100**	**0.060**	**0.029**	

解 由边缘分布定义知，X的分布(相对向量(X,Y)而言，称边缘分布)：

$$P(X=0)=0.202+0.174+\cdots+0.004=0.627.$$

即上表中$X=0$的那一行元素相加；类似可得

$$P(X=1)=0.260, P(X=2)=0.095, P(X=3)=0.018.$$

(将它们填在表的右“边缘”). 按列相加，得Y的分布

$$P(Y=0)=0.202,\cdots,P(Y=6)=0.029.$$

至于X,Y是否独立，因

$P(X=1,Y=0)=0$，而$P(X=1)\cdot P(Y=0)=0.260\times 0.202\neq 0$，所以$X,Y$不独立.

2. 随机向量(X,Y)在矩形区域$D=\{(x,y): a<x<b, c<y<d\}$内服从均匀分布，求联合密度与边缘密度. 又，随机变量X,Y是否独立？

解 由题设知，联合密度$p(x,y)$在区域D上取常数，即

$$p(x,y)=\begin{cases}A, & (x,y)\in D,\\ 0, & \text{其他}.\end{cases}$$

因 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x\mathrm{d}y=1$,代入得

$$\iint_D A\mathrm{d}x\mathrm{d}y=A\iint_D\mathrm{d}x\mathrm{d}y=1,$$

$$A=\frac{1}{\iint_D\mathrm{d}x\mathrm{d}y}=\frac{1}{(b-a)(d-c)}.$$

由边缘密度公式,X 的密度(相对于向量(X,Y);称边缘密度)为

$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y\xlongequal{(a<x<b)}\int_c^d\frac{1}{(b-a)(d-c)}\mathrm{d}y=\frac{1}{b-a}.$$

而当 $x\overline{\in}(a,b)$时,$\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y=0$,所以

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

类似可得

$$p_Y(y)=\begin{cases}\dfrac{1}{d-c}, & c<y<d,\\ 0, & \text{其他}.\end{cases}$$

因为 $p(x,y)=p_X(x)\cdot p_Y(y)$,故 X,Y 相互独立.

3. 随机向量(X,Y)的联合密度

$$p(x,y)=\begin{cases}c(R-\sqrt{x^2+y^2}), & x^2+y^2<R^2,\\ 0, & x^2+y^2\geqslant R^2.\end{cases}$$

求(1) 系数 c;(2) 随机向量落在圆 $x^2+y^2\leqslant r^2(r<R)$内的概率.

解 (1) 由于 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x\mathrm{d}y=1$,同时有

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x\mathrm{d}y=\iint_{x^2+y^2<R^2}c(R-\sqrt{x^2+y^2})\mathrm{d}x\mathrm{d}y$$

$$=cR\cdot\pi R^2-c\int_0^R\int_0^{2\pi}\rho\cdot\rho\mathrm{d}\theta\mathrm{d}\rho$$

$$=c\pi R^3-c2\pi\frac{R^3}{3}=\frac{c\pi R^3}{3},$$

所以
$$c=\frac{3}{\pi R^3}.$$

(2) 记 $G=\{(x,y):x^2+y^2\leqslant r^2\}$,则按(1.6)式有

$$P((X,Y)\in G)=\iint_G p(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\iint_{x^2+y^2\leqslant r^2}\frac{3}{\pi R^3}(R-\sqrt{x^2+y^2})\mathrm{d}x\mathrm{d}y$$

$$=\frac{3}{\pi R^2}\pi r^2-\frac{3}{\pi R^3}\int_0^r\int_0^{2\pi}\rho\cdot\rho\mathrm{d}\theta\mathrm{d}\rho$$

$$=\frac{3r^2}{R^2}\left(1-\frac{2r}{3R}\right).$$

4. 设(X,Y)的联合密度是

$$p(x,y)=\frac{c}{(1+x^2)(1+y^2)}.$$

求:(1) 系数 c;(2) (X,Y)落在以(0,0),(0,1),(1,0),(1,1)为顶点的正方形内的概率;(3) X,Y 是否独立?

解 (1) $\displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x\mathrm{d}y=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{c}{(1+x^2)(1+y^2)}\mathrm{d}x\mathrm{d}y$

$$=c\int_{-\infty}^{+\infty}\frac{1}{1+x^2}\mathrm{d}x\cdot\int_{-\infty}^{+\infty}\frac{1}{1+y^2}\mathrm{d}y$$

$$=c\arctan x\Big|_{-\infty}^{+\infty}\cdot\arctan y\Big|_{-\infty}^{+\infty}$$

$$=c\pi^2,$$

所以 $c=\dfrac{1}{\pi^2}$.

(2) 记 $G=$“以(0,0),(0,1),(1,0),(1,1)为顶点的正方形”

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

按(1.6)式有

$$P((X,Y)\in G)=\iint_G p(x,y)\mathrm{d}x\mathrm{d}y$$
$$=\iint_G \frac{1}{\pi^2}\frac{1}{(1+x^2)(1+y^2)}\mathrm{d}x\mathrm{d}y$$
$$=\frac{1}{\pi^2}\int_0^1\frac{\mathrm{d}x}{1+x^2}\cdot\int_0^1\frac{\mathrm{d}y}{1+y^2}=\frac{1}{\pi^2}\left(\frac{\pi}{4}\right)^2=\frac{1}{16}.$$

(3) 不难算出边缘密度：

$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y=\int_{-\infty}^{+\infty}\frac{1}{\pi^2}\frac{\mathrm{d}y}{(1+x^2)(1+y^2)}$$
$$=\frac{1}{\pi^2}\cdot\frac{1}{1+x^2}\int_{-\infty}^{+\infty}\frac{\mathrm{d}y}{1+y^2}$$
$$=\frac{1}{\pi}\cdot\frac{1}{1+x^2};$$

类似有

$$p_Y(y)=\frac{1}{\pi}\cdot\frac{1}{1+y^2}.$$

最后有

$$p(x,y)=p_X(x)\cdot p_Y(y),$$

这表明 X,Y 相互独立.

5. 设(X,Y)的联合密度是

$$p(x,y)=\begin{cases}A\sin(x+y), & 0<x<\frac{\pi}{2},0<y<\frac{\pi}{2};\\ 0, & \text{其他}.\end{cases}$$

求：(1) 系数 A；(2) 边缘密度.

解 (1) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x\mathrm{d}y=\int_0^{\frac{\pi}{2}}\mathrm{d}x\int_0^{\frac{\pi}{2}}A\sin(x+y)\mathrm{d}y$

$$=-A\int_0^{\frac{\pi}{2}}[\cos(x+y)]_0^{\frac{\pi}{2}}\mathrm{d}x$$
$$=A\int_0^{\frac{\pi}{2}}(\sin x+\cos x)\mathrm{d}x=2A.$$

所以 $A=\frac{1}{2}$.

(2) $p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y\xlongequal{\left(0<x<\frac{\pi}{2}\right)}\int_0^{\frac{\pi}{2}}\frac{1}{2}\sin(x+y)\mathrm{d}y$

$$=-\frac{1}{2}\cos(x+y)\Big|_0^{\frac{\pi}{2}}=\frac{1}{2}(\sin x+\cos x),$$

而当 $x\overline{\in}\left(0,\frac{\pi}{2}\right)$时，$p_X(x)$为 0，故

$$p_X(x)=\begin{cases}\frac{1}{2}(\sin x+\cos x), & 0<x<\frac{\pi}{2},\\ 0, & \text{其他}.\end{cases}$$

类似(或按 $p(x,y)$的对称性)有

$$p_Y(y)=\begin{cases}\frac{1}{2}(\sin y+\cos y), & 0<y<\frac{\pi}{2},\\ 0, & \text{其他}.\end{cases}$$

6. 一机器制造直径为 X 的圆轴，另一机器制造内径为 Y 的轴衬，设(X,Y)的联合密度为

$$p(x,y)=\begin{cases}2\,500, & 0.49<x<0.51,0.51<y<0.53;\\ 0, & \text{其他}.\end{cases}$$

若轴衬的内径与轴的直径之差大于 0.004 且小于 0.036，则两者可以相适衬. 求任一轴与任一轴衬相适衬的概率.

解 所求概率为

$$P(0.004<Y-X<0.036).$$

记 $D=\{(x,y):0.49<x<0.51,0.51<y<0.53\}$，

$G=\{(x,y):(x,y)\in D,\text{且 }0.004<y-x<0.036\}$.

如图 6 所示，D 是一正方形，((X,Y)在 D 上均匀分布)，G 的形状如何？按 G 的定义，它由两条直线切割 D 所得，一条记为 $L_1:y-x=0.036$，另一条记为 $L_2:y-x=0.004$，

则 $P(0.004<Y-X<0.036)=\iint\limits_{0.004<y-x<0.036}p(x,y)\mathrm{d}x\mathrm{d}y=2\,500\iint\limits_G\mathrm{d}x\mathrm{d}y$，而 G 的面积$=D$ 的面积$-$(两小三角形面积之和)，

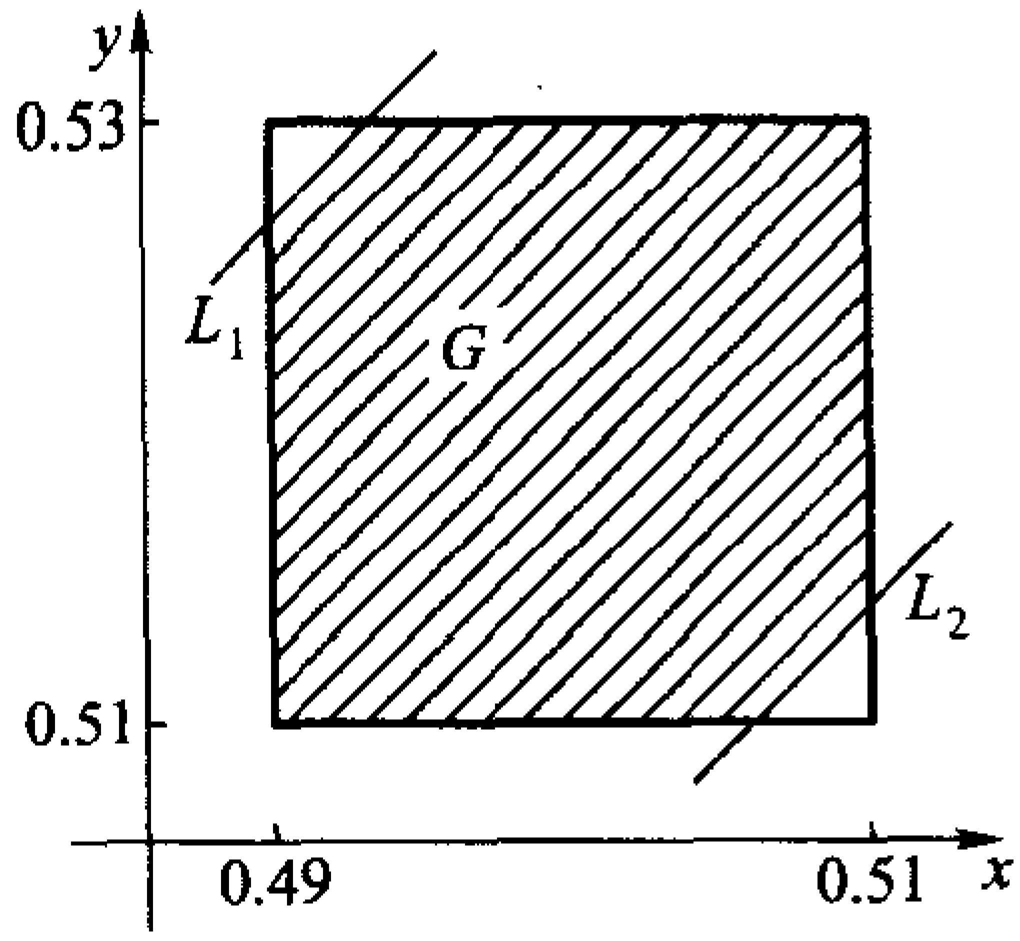

图 6

由直线 L_1, L_2 的方程知，两小三角形都是等腰直角三角形，直角边长为 0.004，因此

$$G\text{ 的面积}=(0.02)^2-(0.004)^2,$$

由此，

$$P(0.004<Y-X<0.036)=2\,500[(0.02)^2-(0.004)^2]$$
$$=0.96.$$

评注 本题的关键是弄清楚事件“相适衬”.

① 首先，由题意可知“相适衬”=“$0.004<Y-X<0.036$”.

② 上式右边，无非限制了(X,Y)的取值范围，若记

$$G^*=\{(x,y):0.004<y-x<0.036\},$$

则“$0.004<Y-X<0.036$”=“(X,Y)取值于 G^*”

=“$(X,Y)\in G^*$”.

(G^* 在平面上是一由 L_1, L_2 界成的带状区域；而 G 是 G^* 与 D 的交集.)

7. 设(X,Y)服从

$$D=\left\{(x,y):\frac{(x+y)^2}{2a^2}+\frac{(x-y)^2}{2b^2}\leqslant 1\right\}$$

上的均匀分布，求(X,Y)的联合密度 $p(x,y)$.

解 由题设知，

$$p(x,y)=\begin{cases}\dfrac{1}{S}, & (x,y)\in D,\\ 0, & \text{其他},\end{cases}$$

其中 S 为 D 的面积.

区域 D 实际上是由椭圆围成的.作旋转坐标变换

$$\begin{cases}x'=\dfrac{1}{\sqrt{2}}(x+y),\\ y'=\dfrac{1}{\sqrt{2}}(x-y),\end{cases}$$

区域 D 在新坐标系下,$D=\left\{(x',y'):\dfrac{x'^2}{a^2}+\dfrac{y'^2}{b^2}\leqslant 1\right\}$. 我们知道,椭圆面积为 πab,此即为所求的 S.

8. 对于下列三组参数.写出二维正态随机向量的联合密度与边缘密度(表略).

解 只需分别将三组参数代入本章(1.10)式即得,无非想要体会一下二元正态密度的复杂性,它主要由 $\rho(|\rho|<1)$ 引起的:$\rho\neq 0$ 时,e 的上角有交叉乘积项 $(x-\mu_1)(y-\mu_2)$;当 $\rho=0$ 时比较简单,此时二元密度可分解为两个一元正态密度的乘积(因此,X, Y 相互独立,且分别服从相应的一元正态分布).

习题十二

1. 设 X,Y 相互独立，其密度分别为

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

求 $X+Y$ 的密度.

解 由本章(2.2)式知 $Z=X+Y$ 的密度为

$$p_Z(z)=\int_{-\infty}^{+\infty}p(x,z-x)\mathrm{d}x,$$

其中 $p(x,y)$ 是 X,Y 的联合密度. 再由题设 X,Y 独立，上式可进一步表述为

$$p_Z(z)=\int_{-\infty}^{+\infty}p_X(x)p_Y(z-x)\mathrm{d}x.$$

(这时，应将这里具体的 $p_X(x)$，$p_Y(y)$ 代入上式右边. 不过，由于它们都是分段函数，情况比较复杂.)

先将 $p_X(x)$ 具体化(请注意积分上下限)，得

$$p_Z(z)=\int_0^1 1\cdot p_Y(z-x)\mathrm{d}x=\int_0^1 p_Y(z-x)\mathrm{d}x.$$

再将 $p_Y(z-x)$ 具体化(而最终求出 $p_Z(z)$)，不过需对 z 的取值分情况讨论：

$z\leqslant 0$ 时，由于 $p_Y(z-x)$ 中的 $z-x$，在积分变量 x 限制在 $(0,1)$ 时，它总是小于 0，因此按这里的 $p_Y(y)$，有 $p_Y(z-x)=0$，故

$$p_Z(z)=\int_0^1 p_Y(z-x)\mathrm{d}x=0.$$

$0<z<1$ 时，由于 x 在 $(0,1)$ 中取不同值时，$z-x$ 的取值并不单纯，因此对 $p_Y(z-x)$ 需分段表示.

$$p_Y(z-x)=\begin{cases}\mathrm{e}^{-(z-x)}, & z-x>0(\text{即 } x<z),\\ 0, & z-x\leqslant 0(\text{即 } x\geqslant z),\end{cases}$$

所以

$$p_Z(z)=\int_0^1 p_Y(z-x)\mathrm{d}x=\int_0^z \mathrm{e}^{-(z-x)}\mathrm{d}x=\mathrm{e}^{-z}(\mathrm{e}^z-1)=1-\mathrm{e}^{-z}.$$

$z\geqslant 1$ 时，对 $z-x$，在积分变量 x 限制在$(0,1)$中时，它总是大于 0，因此 $p_Y(z-x)=\mathrm{e}^{-(z-x)}$. 所以

$$p_Z(z)=\int_0^1 p_Y(z-x)\mathrm{d}x=\int_0^1 \mathrm{e}^{-(z-x)}\mathrm{d}x=\mathrm{e}^{-z}(\mathrm{e}-1)$$

综合以上得

$$p_Z(z)=\begin{cases}0, & z\leqslant 0,\\ 1-\mathrm{e}^{-z}, & 0<z<1,\\ \mathrm{e}^{-z}(\mathrm{e}-1), & z\geqslant 1.\end{cases}$$

2. 设 X,Y 相互独立，分别服从自由度为 k_1,k_2 的 χ^2 分布，即

$$p_X(x)=\begin{cases}\dfrac{1}{2^{\frac{k_1}{2}}\Gamma\left(\dfrac{k_1}{2}\right)}x^{\frac{k_1}{2}-1}\mathrm{e}^{-\frac{x}{2}}, & x>0,\\ 0, & x\leqslant 0;\end{cases}$$

$$p_Y(y)=\begin{cases}\dfrac{1}{2^{\frac{k_2}{2}}\Gamma\left(\dfrac{k_2}{2}\right)}y^{\frac{k_2}{2}-1}\mathrm{e}^{-\frac{y}{2}}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

证明 $X+Y$ 也服从χ^2分布，自由度为 k_1+k_2.

证 由(2.2)式及独立性，$Z=X+Y$ 的密度

$$\begin{aligned}p_Z(z)&=\int_{-\infty}^{+\infty}p_X(x)p_Y(z-x)\mathrm{d}x\\ &=\int_0^{+\infty}\frac{1}{2^{\frac{k_1}{2}}\Gamma\left(\frac{k_1}{2}\right)}x^{\frac{k_1}{2}-1}\mathrm{e}^{-\frac{x}{2}}\cdot p_Y(z-x)\mathrm{d}x\\ &\xlongequal{(z>0)}\frac{1}{2^{\frac{k_1+k_2}{2}}\Gamma\left(\frac{k_1}{2}\right)\Gamma\left(\frac{k_2}{2}\right)}\int_0^z x^{\frac{k_1}{2}-1}\mathrm{e}^{-\frac{x}{2}}\cdot\\ &\qquad(z-x)^{\frac{k_2}{2}-1}\mathrm{e}^{-\frac{1}{2}(z-x)}\mathrm{d}x\\ &=A\cdot\mathrm{e}^{-\frac{1}{2}z}\int_0^z x^{\frac{k_1}{2}-1}(z-x)^{\frac{k_2}{2}-1}\mathrm{d}x\end{aligned}$$

$$\xlongequal{\left(令\frac{x}{z}=t\right)} A\cdot e^{-\frac{z}{2}}\cdot z^{\frac{k_1+k_2}{2}-1}\cdot\int_0^1 t^{\frac{k_1}{2}-1}(1-t)^{\frac{k_2}{2}-1}\mathrm{d}t$$

利用 $\int_0^1 t^{\frac{k_1}{2}-1}(1-t)^{\frac{k_2}{2}-1}\mathrm{d}t=B\left(\frac{k_1}{2},\frac{k_2}{2}\right)=\dfrac{\Gamma\left(\frac{k_1}{2}\right)\Gamma\left(\frac{k_2}{2}\right)}{\Gamma\left(\frac{k_1}{2}+\frac{k_2}{2}\right)}$，得

$$p_Z(z)=\begin{cases}\dfrac{1}{2^{\frac{k_1+k_2}{2}}\cdot\Gamma\left(\dfrac{k_1+k_2}{2}\right)}z^{\frac{k_1+k_2}{2}-1}e^{-\frac{z}{2}}, & z>0,\\ 0, & z\leqslant 0.\end{cases}$$

3. 设 $X_1\sim N(\mu_1,\sigma_1^2)$，$X_2\sim N(\mu_2,\sigma_2^2)$，且 X_1,X_2 相互独立，求证：

(1) $X_1+X_2\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$.

(2) $\dfrac{X_1+X_2}{2}\sim N\left(\dfrac{\mu_1+\mu_2}{2},\dfrac{\sigma_1^2+\sigma_2^2}{4}\right)$.

证 (1) 为简化证明过程，令

$$Y_i=X_i-\mu_i,\quad i=1,2.$$

则由题设知 $Y_i\sim N(0,\sigma_i^2)$，$i=1,2$，且 Y_1,Y_2 独立. 现证：

$$Y_1+Y_2\sim N(0,\sigma_1^2+\sigma_2^2).$$

（数学上，为简化叙述，将上述过程只用“不妨假定 $\mu_1=0,\mu_2=0$”带过）.

记 $Z=Y_1+Y_2$，则 Z 的密度

$$\begin{aligned}p_Z(z)&=\int_{-\infty}^{+\infty}p_{Y_1}(y)p_{Y_2}(z-y)\mathrm{d}y\\&=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{y^2}{2\sigma_1^2}}\cdot\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(z-y)^2}{2\sigma_2^2}}\mathrm{d}y\\&=\frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{+\infty}e^{-\frac{1}{2}\left[\frac{y^2}{\sigma_1^2}+\frac{(z-y)^2}{\sigma_2^2}\right]}\mathrm{d}y\\&=\frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{+\infty}e^{-\frac{1}{2}\left[\frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\cdot\sigma_2^2}\left(y-\frac{\sigma_1^2 z}{\sigma_1^2+\sigma_2^2}\right)^2+\frac{1}{\sigma_1^2+\sigma_2^2}z^2\right]}\mathrm{d}y\end{aligned}$$

（将 e 的上角方括号内的关于 y 的二次三项式配方）

$$=\frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}}\mathrm{e}^{\frac{1}{2}\cdot\frac{z^2}{\sigma_1^2+\sigma_2^2}}.$$

这表明 $Z\sim N(0,\sigma_1^2+\sigma_2^2)$，因而 $X_1+X_2\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$.

(2) 由 $\frac{X_1+X_2}{2}=\frac{1}{2}(X_1+X_2)$，利用(1)立得.

4. 设 X,Y 独立同分布，密度为 $p(\cdot)$，分布函数为 $F(\cdot)$，求 $\min(X,Y)$ 的密度.

解 记 $Z=\min(X,Y)$. 则有

$$\begin{aligned}F_Z(z)&=P(Z\leqslant z)=1-P(Z>z)=1-P(X>z,Y>z)\\&=1-P(X>z)P(Y>z)=1-[1-F(z)]^2.\end{aligned}$$

因此，

$$\begin{aligned}p_Z(z)&=\{1-[1-F(z)]^2\}'\\&=2[1-F(z)]p(z).\end{aligned}$$

5. 设系统 L 由两个相互独立的子系统 L_1,L_2 联接而成. 联接的方式分别为(1) 串联，(2) 并联，(3) 备用(当系统 L_1 损坏时，系统 L_2 开始工作). 如图 7 所示，已知 L_1,L_2 的寿命分别为 X,Y，

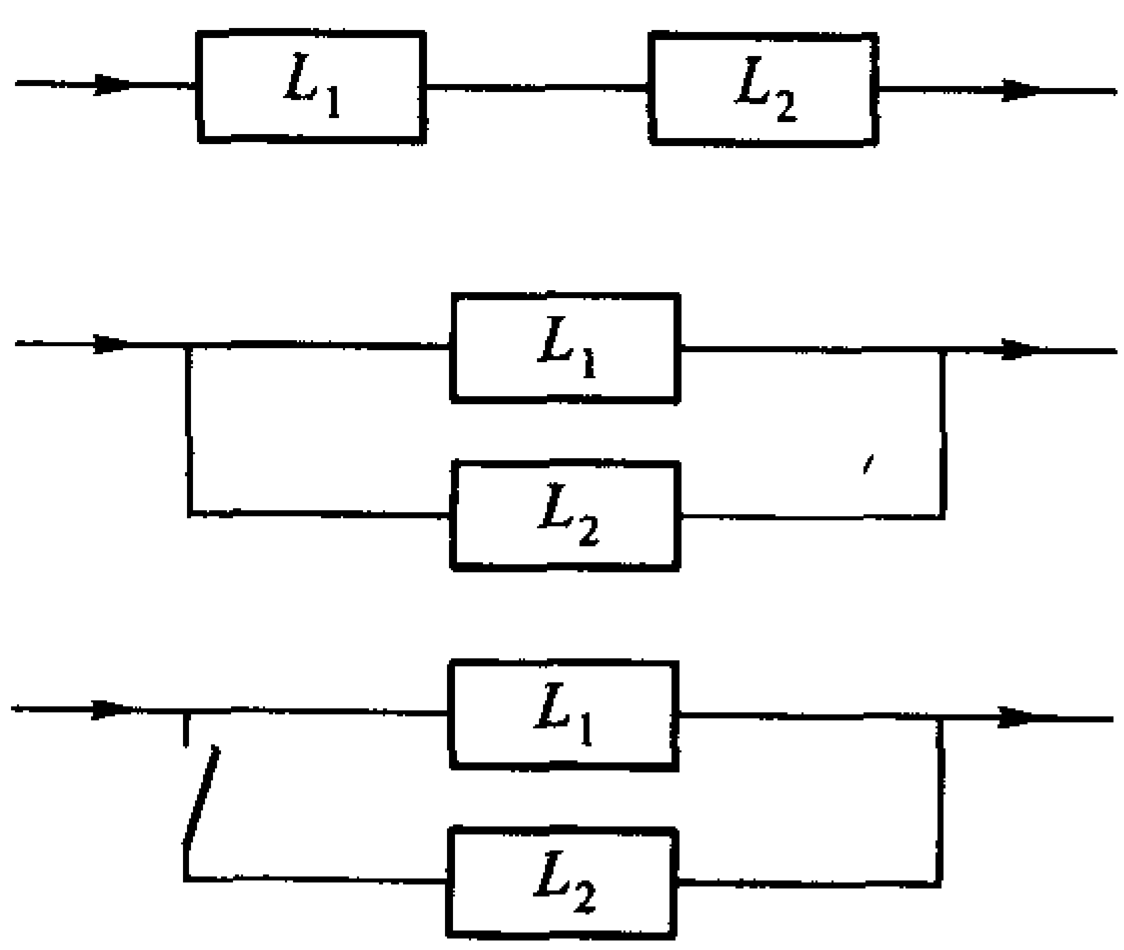

图 7

概率密度分别为

$$p_X(x)=\begin{cases}\alpha\mathrm{e}^{-\alpha x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}\quad p_Y(y)=\begin{cases}\beta\mathrm{e}^{-\beta y}, & y>0,\\ 0, & y\leqslant 0,\end{cases}$$

其中 $\alpha>0,\beta>0$，且 $\alpha\neq\beta$. 试分别就这三种联接方式写出系统 L 的

寿命 Z 的概率密度.

解 (1) 串联

系统 L 的寿命 Z 与两个子系统 L_1,L_2 的寿命 X,Y 应有关系

$$Z=\min(X,Y).$$

于是
$$\begin{aligned}F_Z(z)&=P(Z\leqslant z)=1-P(Z>z)\\&=1-P(X>z)P(Y>z)\\&=1-(1-F_X(z))(1-F_Y(z)).\end{aligned}$$

而 $F_X(z)=\int_{-\infty}^{z}p_X(x)\mathrm{d}x\xlongequal{(z>0)}\int_0^z\alpha\mathrm{e}^{-\alpha x}\mathrm{d}x=-\mathrm{e}^{-\alpha x}\Big|_0^z=1-\mathrm{e}^{-\alpha z}$.

类似有

$$F_Y(z)=\begin{cases}1-\mathrm{e}^{-\beta z}, & z>0,\\0, & z\leqslant 0.\end{cases}$$

代入上式得

$$F_Z(z)=1-\mathrm{e}^{-(\alpha+\beta)z},z>0.$$

求导得

$$p_Z(z)=\begin{cases}(\alpha+\beta)\mathrm{e}^{-(\alpha+\beta)z}, & z>0,\\0, & z\leqslant 0.\end{cases}$$

(2) 并联

此时,有关系

$$Z=\max(X,Y),$$

于是

$$\begin{aligned}F_Z(z)&=P(Z\leqslant z)=P(X\leqslant z,Y\leqslant z)=F_X(z)F_Y(z)\\&=(1-\mathrm{e}^{-\alpha z})(1-\mathrm{e}^{-\beta z}),\quad z>0.\end{aligned}$$

求导得

$$p_Z(z)=\begin{cases}\alpha\mathrm{e}^{-\alpha z}+\beta\mathrm{e}^{-\beta z}-(\alpha+\beta)\mathrm{e}^{-(\alpha+\beta)z}, & z>0,\\0, & z\leqslant 0.\end{cases}$$

(3) 备用

此时,有关系

$Z=X+Y$. 由本章(2.2)式

$$p_Z(z)=\int_{-\infty}^{+\infty}p_X(x)p_Y(z-x)\mathrm{d}x$$

$$= \int_0^{+\infty} \alpha e^{-\alpha x} \cdot p_Y(z-x)\,dx \quad （先具体化\ p_X(x)）$$

$$\xlongequal{(z>0)} \int_0^z \alpha e^{-\alpha x} \beta e^{-\beta(z-x)}\,dx$$

$$= \alpha\beta e^{-\beta z} \int_0^z e^{(\beta-\alpha)x}\,dx \quad （z \leqslant 0\ 时，p_Y(z-x)=0，所以积$$

分为 0）

$$= \frac{\alpha\beta}{\beta-\alpha}(e^{-\alpha z} - e^{-\beta z}).$$

评注 本题是一道很好的习题.弄清 Z 与 X,Y 的关系，是本题的基本点.另外，$\beta=\alpha$ 时，又怎样呢？

6. 设某种商品一周的需要量是一随机变量，其密度为

$$p(x) = \begin{cases} xe^{-x}, & x>0, \\ 0, & x \leqslant 0. \end{cases}$$

如果各周的需要量是相互独立的，试求(1)两周，(2)三周的需要量的概率密度.

(1) 两周需要量 Z

若记 X_1 = 第一周的需要量，

X_2 = 第二周的需要量，

则 $Z = X_1 + X_2$.

$$p_Z(z) = \int_{-\infty}^{+\infty} p_1(x) p_2(z-x)\,dx = \int_0^{+\infty} xe^{-x} p_2(z-x)\,dx$$

$$\xlongequal{(z>0)} \int_0^z xe^{-x}(z-x)e^{-(z-x)}\,dx$$

$$= e^{-z}\int_0^z x(z-x)\,dx$$

$$= \frac{z^3}{6}e^{-z},$$

所以

$$p_Z(z) = \begin{cases} \dfrac{z^3}{6}e^{-z}, & z>0, \\ 0, & z \leqslant 0. \end{cases}$$

(2) 三周需要量 Y

$Y=X_1+X_2+X_3$,(X_3 是第三周的需要量.)

则 $Y=Z+X_3$.

$$p_Y(y)=\int_{-\infty}^{+\infty}p_Z(x)p_3(y-x)\mathrm{d}x \xlongequal{(y>0)} \int_0^y \frac{x^3}{6}\mathrm{e}^{-x}\cdot (y-x)\mathrm{e}^{-(y-x)}\mathrm{d}x$$

$$=\mathrm{e}^{-y}\int_0^y \frac{x^3}{6}(y-x)\mathrm{d}x=\frac{y^5}{120}\mathrm{e}^{-y},$$

所以

$$p_Y(y)=\begin{cases}\dfrac{y^5}{120}\mathrm{e}^{-y}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

评注 解本题时,最易犯的错误是认为 $Z=2X$. 这显然不合理,因为第二周的实际需要量一般与第一周的实际需要量不同. 另外,如果要求 n 周的需要量呢? 我们可用下列引理:

引理 设 $X_1\sim\Gamma(\alpha_1,\beta)$, $X_2\sim\Gamma(\alpha_2,\beta)$, X_1,X_2 独立,则 $X_1+X_2\sim\Gamma(\alpha_1+\alpha_2,\beta)$.(引理的证明可类似于对本章习题 2 的证明,这里从略. 另外由于 χ^2 分布是 Γ 分布族中的一类,习题 2 也是本引理的一个推论).

不难看出本题一周的需要量服从 $\Gamma(2,1)$. 因此,按该引理,n 周的需要量应服从 $\Gamma(2+2+\cdots+2,1)$,即 $\Gamma(2n,1)$分布.

习 题 十 三

1. 设(X,Y)的联合密度

$$p(x,y)=\begin{cases}4xy\mathrm{e}^{-(x^2+y^2)}, & x>0,y>0,\\ 0, & \text{其他}.\end{cases}$$

求 $Z=\sqrt{X^2+Y^2}$的均值(用两种方法).

解法一 用(3.1)式

$$\begin{aligned}E(Z) &= E(\sqrt{X^2+Y^2})=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\sqrt{x^2+y^2}p(x,y)\mathrm{d}x\mathrm{d}y\\ &=\int_0^{+\infty}\int_0^{+\infty}\sqrt{x^2+y^2}4xy\mathrm{e}^{-(x^2+y^2)}\mathrm{d}x\mathrm{d}y\\ &=\int_0^{\frac{\pi}{2}}\mathrm{d}\theta\int_0^{+\infty}\rho\cdot 4\rho^2\sin\theta\cos\theta\mathrm{e}^{-\rho^2}\rho\mathrm{d}\rho\\ &=\int_0^{+\infty}2\rho^4\mathrm{e}^{-\rho^2}\mathrm{d}\rho \xlongequal{(t=\rho^2)}\int_0^{+\infty}t^{\frac{3}{2}}\mathrm{e}^{-t}\mathrm{d}t=\Gamma\left(\frac{5}{2}\right)=\frac{3}{4}\sqrt{\pi}.\end{aligned}$$

解法二 先找 Z 的密度 $p_Z(z)$.

$$\begin{aligned}F_Z(z) &= P(Z\leqslant z)=P(\sqrt{X^2+Y^2}\leqslant z)\\ &=\iint\limits_{\substack{\sqrt{x^2+y^2}\leqslant z\\ x>0,y>0}}4xy\mathrm{e}^{-(x^2+y^2)}\mathrm{d}x\mathrm{d}y\\ &\xlongequal{(z>0)}\int_0^{\frac{\pi}{2}}\mathrm{d}\theta\int_0^z 4\rho^2\sin\theta\cos\theta\mathrm{e}^{-\rho^2}\rho\mathrm{d}x\mathrm{d}y\\ &=\int_0^z 2\rho^3\mathrm{e}^{-\rho^2}\mathrm{d}\rho.\end{aligned}$$

所以

$$p_Z(z)=\begin{cases}2z^3\mathrm{e}^{-z^2}, & z>0,\\ 0, & z\leqslant 0.\end{cases}$$

于是 $E(Z)=\int_0^{\infty} z\cdot 2z^3\mathrm{e}^{-z^2}\,\mathrm{d}z=\frac{3}{4}\sqrt{\pi}.$

2. 证明：如果 X 与 Y 独立，则

$D(X\cdot Y)=D(X)D(Y)+(E(X))^2D(Y)+(E(Y))^2D(X).$

证 我们有 $D(X\cdot Y)=E(X^2\cdot Y^2)-(E(X\cdot Y))^2$. 由 X,Y 相互独立性，按定义 1.5 容易直接证明 X^2,Y^2 也相互独立（实际上，可以证明若 X,Y 相互独立，$f(\cdot),g(\cdot)$ 是所谓"贝尔函数"——相当广泛的一类函数，则 $f(X),g(Y)$ 也相互独立）. 于是有

$$E(X^2\cdot Y^2)=E(X^2)E(Y^2).$$

连同等式 $E(X\cdot Y)=E(X)E(Y)$，经简单的算术运算即可得要证的等式.

3. 设 (X,Y) 的联合密度

$$p(x,y)=\frac{A}{(x^2+y^2+1)^2}\quad (A \text{ 是常数}),$$

试求出 A 的数值，并问 σ_{XX} 与 σ_{YY} 是否存在？

解 $$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)\,\mathrm{d}x\mathrm{d}y=\int_0^{2\pi}\mathrm{d}\theta\int_0^{+\infty}\frac{A}{(r^2+1)^2}r\,\mathrm{d}r$$

$$=A\pi\int_0^{+\infty}\frac{\mathrm{d}(r^2)}{(r^2+1)^2}=A\pi,$$

因此 $$A=\frac{1}{\pi}.$$

$$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}x\frac{1}{\pi(x^2+y^2+1)^2}\mathrm{d}x\mathrm{d}y$$

$$=\int_{-\infty}^{+\infty}\mathrm{d}y\int_{-\infty}^{+\infty}\frac{x}{\pi(x^2+y^2+1)^2}\mathrm{d}x$$

（由对称性知）$=0$.

类似有

$$E(Y)=0.$$

下面计算σ_{XX}，σ_{YY}.

$$\sigma_{XX}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-E(X))^2\frac{\mathrm{d}x\mathrm{d}y}{\pi(x^2+y^2+1)^2}$$
$$=\int_0^{2\pi}\cos^2\theta\mathrm{d}\theta\int_0^{+\infty}\frac{r^2\cdot r}{\pi(r^2+1)^2}\mathrm{d}r=\infty.$$

类似有$\sigma_{YY}=\infty$. 这表明σ_{XX}，σ_{YY}都不存在.

4. 设(X,Y)服从区域$D=\{(x,y):0<x<1,0<y<x\}$上的均匀分布，求相关系数ρ.

解 区域D如图8所示，其面积等于$\frac{1}{2}$，则(X,Y)的联合密度

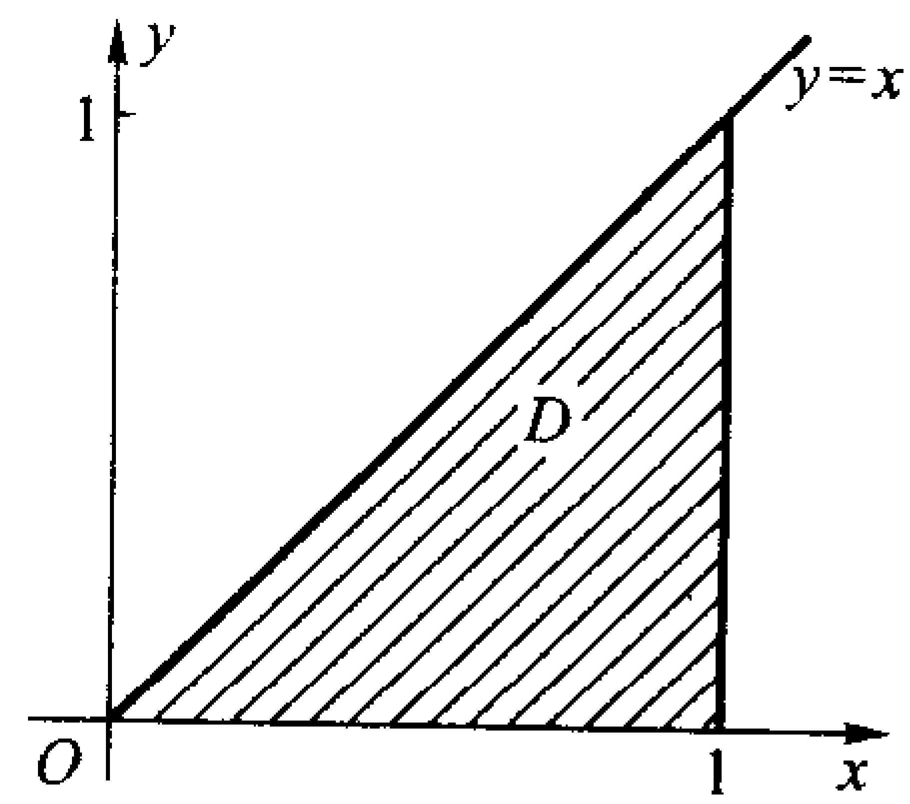

图 8

$$p(x,y)=\begin{cases}2, & (x,y)\in D,\\ 0, & \text{其他}.\end{cases}$$

$$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xp(x,y)\mathrm{d}x\mathrm{d}y$$
$$=\iint_D 2x\mathrm{d}x\mathrm{d}y=\int_0^1 2x\mathrm{d}x\int_0^x\mathrm{d}y$$
$$=\int_0^1 2x^2\mathrm{d}x=\frac{2}{3}.$$

$$E(Y)=\iint_D 2y\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^x 2y\mathrm{d}y=\int_0^1 x^2\mathrm{d}x=\frac{1}{3}.$$

$$\sigma_{XX}=\iint_D\left(x-\frac{2}{3}\right)^2\cdot 2\mathrm{d}x\mathrm{d}y=\int_0^1 2\left(x-\frac{2}{3}\right)^2\mathrm{d}x\int_0^x\mathrm{d}y$$
$$=\int_0^1 2x\left(x-\frac{2}{3}\right)^2\mathrm{d}x=\frac{1}{18}.$$

$$\sigma_{YY}=\iint_D\left(y-\frac{1}{3}\right)^2\cdot 2\mathrm{d}x\mathrm{d}y=\int_0^1\left(y-\frac{1}{3}\right)^2\cdot 2\mathrm{d}y\int_y^1\mathrm{d}x$$
$$=\int_0^1 2\left(y-\frac{1}{3}\right)^2(1-y)\mathrm{d}y=\frac{1}{18}.$$

$$\sigma_{XY}=\iint_D\left(x-\frac{2}{3}\right)\left(y-\frac{1}{3}\right)2\mathrm{d}x\mathrm{d}y=\int_0^1 2\left(x-\frac{2}{3}\right)\mathrm{d}x\int_0^x\left(y-\frac{1}{3}\right)\mathrm{d}y$$

$$= \frac{1}{36}.$$

因此

$$\rho = \frac{\sigma_{XY}}{\sqrt{\sigma_{XX}}\sqrt{\sigma_{YY}}} = \frac{1}{2}.$$

5. 设 X_1, X_2 独立，概率密度分别为

$$p_1(x) = \begin{cases} 2x, & 0 \leqslant x \leqslant 1, \\ 0, & \text{其他}. \end{cases} \qquad p_2(x) = \begin{cases} e^{-(x-5)}, & x > 5, \\ 0, & \text{其他}. \end{cases}$$

求 $E(X_1 \cdot X_2)$.

解 由独立性 $E(X_1 \cdot X_2) = E(X_1)E(X_2)$

$$= \int_0^1 x(2x)\mathrm{d}x \cdot \int_5^{+\infty} x e^{-(x-5)}\mathrm{d}x$$

$$= \frac{2}{3} \cdot \int_0^{+\infty} (t+5)e^{-t}\mathrm{d}t = \frac{2}{3}(1+5)$$

$$= 4.$$

6. 已知 $D(X)=25, D(Y)=36, \rho=0.4$，求 $D(X+Y)$ 及 $D(X-Y)$.

解 $D(X+Y) = E[(X+Y)-E(X+Y)]^2$

$$= E[(X-E(X))+(Y-E(Y))]^2$$

$$= D(X)+D(Y)+2\sigma_{XY}.$$

而 $\sigma_{XY} = \rho\sqrt{D(X)D(Y)}$，故得

$$D(X+Y) = 25+36+2\left(0.4\sqrt{25\times 36}\right) = 85.$$

类似可得 $D(X-Y) = 25+36-2\left(0.4 \cdot \sqrt{25\times 36}\right) = 37$.

7. 设 (X,Y) 服从二维正态分布，$E(X)=E(Y)=0, D(X)=a^2$，$D(Y)=b^2, \rho=0$. 求 (X,Y) 落在区域 $D=\{(x,y): \frac{x^2}{a^2}+\frac{y^2}{b^2} \leqslant k^2\}$ 中的概率.

解 (X,Y) 的密度

$$p(x,y) = \frac{1}{2\pi ab} e^{-\frac{1}{2}\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right)}.$$

所求概率

$$P((X,Y)\in D)=\iint\limits_D \frac{1}{2\pi ab}e^{-\frac{1}{2}\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right)}\mathrm{d}x\mathrm{d}y.$$

对此二重积分,作广义极坐标变换$\begin{cases}x=ar\cos\theta\\ y=br\sin\theta\end{cases}\begin{pmatrix}r\geqslant 0,\\ 0<\theta\leqslant 2\pi\end{pmatrix}$

$\left(\text{该变换的雅可比行列式}\frac{\partial(x,y)}{\partial(r,\theta)}=abr\right)$,得

$$P((X,Y)\in D)=\int_0^{2\pi}\mathrm{d}\theta\int_0^k\frac{1}{2\pi ab}e^{-\frac{1}{2}r^2}abr\mathrm{d}r=\int_0^k re^{-\frac{r^2}{2}}\mathrm{d}r$$

$$=1-e^{-\frac{k^2}{2}}.$$

评注 求解一个二重积分,无非通过累次积分来做.重要的是,先看看能否找出一个变换,使其比较简单;所谓简单,是指被积函数与积分区域都简单.比如被积函数可分解,积分区域为矩形区域.当然,作变换时不要丢了雅可比行列式.

8. 直接验证:若 $Y=a+bX$,则 $\rho=\begin{cases}1, & \text{当 } b>0,\\ -1, & \text{当 } b<0.\end{cases}$

证 $\sigma_{YY}=E(Y-E(Y))^2=E(a+bX-(a+bE(X)))^2=b^2D(X)$,

$$\sigma_{XY}=E(X-E(X))(Y-E(Y))=bD(X),$$

于是

$$\rho=\frac{\sigma_{XY}}{\sqrt{\sigma_{XX}\sigma_{YY}}}=\frac{bD(X)}{|b|D(X)}=\begin{cases}1, & b>0,\\ -1, & b<0.\end{cases}$$

9. 设 $X\sim N(0,1)$,而 $Y=X^n$(n 是正整数),求 ρ_{XY}.

解 因 $E(X)=0$,所以

$$\sigma_{XY}=E(X-E(X))(Y-E(Y))=E(X\cdot Y)=E(X^{n+1}).$$

而 $\sigma_{YY}=E(Y^2)-(E(Y))^2=E(X^{2n})-(E(X^n))^2$.又知 $\sigma_{XX}=1$.因此,为求 $\rho_{XY}=\frac{\sigma_{XY}}{\sqrt{\sigma_{XX}\cdot\sigma_{YY}}}$,只需找出 $E(X^m)$的一般公式即可.

首先,当 m 为奇数时,显然有 $E(X^m)=0$.

当 m 为偶数时，

$$E(X^m)=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}x^m e^{-\frac{x^2}{2}}\mathrm{d}x=\frac{\sqrt{2}}{\sqrt{\pi}}\int_0^{+\infty}x^m e^{-\frac{x^2}{2}}\mathrm{d}x$$

$$\left(\text{作变换 } t=\frac{x^2}{2}\right)=\frac{\sqrt{2}}{\sqrt{\pi}}2^{\frac{m-1}{2}}\int_0^{+\infty}t^{\frac{m-1}{2}}e^{-t}\mathrm{d}t$$

$$=\frac{1}{\sqrt{\pi}}2^{\frac{m}{2}}\Gamma\left(\frac{m+1}{2}\right),$$

而 $\Gamma\left(\frac{m+1}{2}\right)=\Gamma\left(\frac{m}{2}+\frac{1}{2}\right)=\left(\frac{m}{2}-\frac{1}{2}\right)\left(\frac{m}{2}-1-\frac{1}{2}\right)\cdots\frac{1}{2}\Gamma\left(\frac{1}{2}\right)$

$$=\frac{(m-1)!!}{2^{\frac{m}{2}}}\sqrt{\pi},$$

所以，此时有

$$E(X^m)=(m-1)!!.$$

因此，当 n 为偶数时，$\sigma_{XY}=0$，得 $\rho_{XY}=0$；而当 n 为奇数时，

$$\sigma_{XY}=E(X^{n+1})=(n+1-1)!!=n!!$$

$$\sigma_{YY}=E(X^{2n})=(2n-1)!!$$

得

$$\rho_{XY}=\frac{n!!}{\sqrt{(2n-1)!!}}.$$

习 题 十 四

1. 已知 X,Y,Z 的联合密度为

$$p(x,y,z)=\begin{cases}e^{-(x+y+z)}, & x>0,y>0,z>0,\\ 0, & \text{其他},\end{cases}$$

分别求出 X,Y,Z 的单个密度. X,Y,Z 相互独立吗?

解 由本章(4.3)式,X 的密度

$$p_X(x)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y,z)\mathrm{d}y\mathrm{d}z$$

$$=\begin{cases}\int_0^{+\infty}\int_0^{+\infty}e^{-(x+y+z)}\mathrm{d}y\mathrm{d}z=e^{-x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

类似有 $p_Y(y)=\begin{cases}e^{-y}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$ $p_Z(z)=\begin{cases}e^{-z}, & z>0,\\ 0, & z\leqslant 0.\end{cases}$

于是有

$$p(x,y,z)=p_X(x)\cdot p_Y(y)\cdot p_Z(z).$$

所以 X,Y,Z 相互独立.

2. 设 $X_1,X_2,\cdots,X_n$ 独立同分布,都服从 $N(\mu,\sigma^2)$,求 $(X_1,X_2,\cdots,X_n)$ 的联合密度

解 已知 X_i 的密度

$$p_{X_i}(x_i)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x_i-\mu)^2}{2\sigma^2}},\quad i=1,2,\cdots,n;$$

又知 $X_1,X_2,\cdots,X_n$ 相互独立,则$(X_1,X_2,\cdots,X_n)$的联合密度

$$p(x_1,x_2,\cdots,x_n)=p_{X_1}(x_1)p_{X_2}(x_2)\cdots p_{X_n}(x_n)$$

$$=\frac{1}{(2\pi)^{\frac{n}{2}}\sigma^n}e^{-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2}$$

3. 设 X,Y,Z 独立同分布,服从 $N(0,1)$,求 $\sqrt{X^2+Y^2+Z^2}$ 的

概率密度.

解 (X,Y,Z)的联合密度

$$p(x,y,z)=\frac{1}{(2\pi)^{\frac{3}{2}}}e^{-\frac{1}{2}(x^2+y^2+z^2)},$$

则$U=\sqrt{X^2+Y^2+Z^2}$的分布函数为(对 $u>0$)

$$\begin{aligned}F_U(u)&=P\left(\sqrt{X^2+Y^2+Z^2}\leqslant u\right)\\&=\iiint\limits_{\sqrt{x^2+y^2+z^2}\leqslant u}\frac{1}{(2\pi)^{\frac{3}{2}}}e^{-\frac{1}{2}(x^2+y^2+z^2)}dxdydz\end{aligned}$$

$$\left(\text{作球坐标变换}\quad\begin{cases}x=r\sin\varphi\cos\theta\\y=r\sin\varphi\sin\theta\\z=r\cos\varphi\end{cases}\right)$$

$$\begin{aligned}&=\int_0^{\pi}d\varphi\int_0^{2\pi}d\theta\int_0^u\frac{1}{(2\pi)^{\frac{3}{2}}}e^{-\frac{1}{2}r^2}\cdot r^2\sin\varphi dr\\&=\frac{2\cdot2\pi}{(2\pi)^{\frac{3}{2}}}\int_0^u r^2e^{-\frac{1}{2}r^2}dr=\sqrt{\frac{2}{\pi}}\int_0^u r^2e^{-\frac{1}{2}r^2}dr.\end{aligned}$$

所以

$$p_U(u)=\begin{cases}\sqrt{\dfrac{2}{\pi}}u^2e^{-\frac{1}{2}u^2}, & u>0,\\0, & u\leqslant0.\end{cases}$$

4. 设 $X_1,X_2,\cdots,X_n$ 独立同分布,分布的均值是μ,方差是σ^2,而$Y=\frac{1}{n}(X_1+X_2+\cdots+X_n)$,求 $E(Y),D(Y)$.

解
$$\begin{aligned}E(Y)&=E\left(\frac{1}{n}(X_1+X_2+\cdots+X_n)\right)\\&=\frac{1}{n}(E(X_1)+\cdots+E(X_n))=\mu.\end{aligned}$$

$$\begin{aligned}D(Y)&=D\left(\frac{1}{n}(X_1+X_2+\cdots+X_n)\right)\\&=\frac{1}{n^2}(D(X_1)+\cdots+D(X_n))=\frac{\sigma^2}{n}.\end{aligned}$$

5. 设 $X_1,X_2,\cdots,X_n$ 独立，都服从参数为 $m,\eta(m>0,\eta>0)$ 的韦布尔分布 $\left(p(x)=\frac{m}{\eta^m}x^{m-1}\mathrm{e}^{-\left(\frac{x}{\eta}\right)^m},x>0\right)$，试证明：$\min(X_1,X_2,\cdots,X_n)$ 仍服从韦布尔分布.

证 首先，由参数为 m,η 的韦布尔分布的密度可知，其相应的分布函数

$$F(x)=\begin{cases}1-\mathrm{e}^{-\left(\frac{x}{\eta}\right)^m}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

记 $Z=\min(X_1,X_2,\cdots,X_n)$，则类似于习题十二的第 4 题，我们有

$F_Z(z)=1-(1-F(z))^n$，其中 $F(z)$ 是 X_1 的分布函数，

因此有

$$F_Z(z)=\begin{cases}1-\left(\mathrm{e}^{-\left(\frac{z}{\eta}\right)^m}\right)^n\\ 0\end{cases}=\begin{cases}1-\mathrm{e}^{-n\left(\frac{z}{\eta}\right)^m}, & z>0,\\ 0, & z\leqslant 0.\end{cases}$$

这表明 Z 服从参数为 $m,\eta/n^{\frac{1}{m}}$ 的韦布尔分布.

6. 将 n 只球放入 M 只盒子中，设每只球落入各个盒子是等可能的，求有球的盒子数 X 的均值. (提示：引进随机变量

$$X_i=\begin{cases}1, & \text{第 } i \text{ 只盒子中有球}\\ 0, & \text{第 } i \text{ 只盒子中无球}\end{cases}\quad i=1,2,\cdots,M,$$

显然有 $X=\sum\limits_{i=1}^{M}X_i$)

解 对提示中引进的随机变量 $X_i(i=1,2,\cdots,M)$，有

$P(X_i=0)=P(\text{第 } i \text{ 只盒中无球})$

$=P($第 1 个球没有落入第 i 只盒子，

第 2 个球没有落入第 i 只盒子，

…………

第 n 个球没有落入第 i 只盒子$)$

$$=\left(1-\frac{1}{M}\right)^n,$$

因此

$$P(X_i=1)=1-\left(1-\frac{1}{M}\right)^n,$$

所以

$$E(X_i)=1-\left(1-\frac{1}{M}\right)^n.$$

由期望的可加性，

$$E(X)=E\Big(\sum_{i=1}^{M}X_i\Big)=\sum_{i=1}^{M}E(X_i)=M\left[1-\left(1-\frac{1}{M}\right)^n\right].$$

评注 求随机变量的期望，并不一定要先找出它的分布．本题中将它分解为一些比较简单的随机变量之和，值得品味．

7．求事件在 n 次独立试验中发生次数的均值与方差．如果该事件在第 i 次试验中发生的概率等于 $p_i(i=1,2,\cdots,n)$

解 引进随机变量 $X_i(i=1,2,\cdots,n)$：

$$X_i=\begin{cases}1, & \text{第 } i \text{ 次试验中该事件发生,}\\ 0, & \text{第 } i \text{ 次试验中该事件不发生.}\end{cases}$$

则在 n 次试验中该事件发生次数 X 有分解式

$$X=\sum_{i=1}^{n}X_i$$

由于 X_i 服从两点分布，则有

$$E(X_i)=p_i,\quad D(X_i)=p_i(1-p_i).\quad i=1,2,\cdots,n.$$

因此

$$E(X)=\sum_{i=1}^{n}E(X_i)=\sum_{i=1}^{n}p_i.$$

再由独立性

$$D(X)=\sum_{i=1}^{n}D(X_i)=\sum_{i=1}^{n}p_i(1-p_i).$$

8．对于随机变量 X,Y,Z，已知

$$E(X)=E(Y)=1,\quad E(Z)=-1,\quad D(X)=D(Y)=D(Z)=1,$$

$$\rho_{XY}=0,\rho_{XZ}=\frac{1}{2},\rho_{YZ}=-\frac{1}{2}.$$

求 $E(X+Y+Z)$，$D(X+Y+Z)$

解 $E(X+Y+Z)=E(X)+E(Y)+E(Z)=1+1-1=1.$

$$\begin{aligned}D(X+Y+Z)&=E(X+Y+Z-E(X+Y+Z))^2\\&=E(X-E(X)+Y-E(Y)+Z-E(Z))^2\\&=D(X)+D(Y)+D(Z)+2\rho_{XY}\sqrt{D(X)D(Y)}+\\&\quad 2\rho_{XZ}\sqrt{D(X)D(Z)}+2\rho_{YZ}\sqrt{D(Y)D(Z)}\\&=3+2\cdot\frac{1}{2}+2\cdot\left(-\frac{1}{2}\right)=3.\end{aligned}$$

9. 设 $X_1,X_2,\cdots,X_n$ 相互独立，且 $X_i\sim N(\mu,\sigma^2)$. $(i=1,2,\cdots,n)$. 试证明：

$$\frac{X_1+X_2+\cdots+X_n}{n}\sim N\left(\mu,\frac{\sigma^2}{n}\right)$$

证 由习题十二第3题知：

$$X_1+X_2\sim N(2\mu,2\sigma^2),X_1+X_2+X_3\sim N(3\mu,3\sigma^2),\cdots$$

$$X_1+X_2+\cdots+X_n\sim N(n\mu,n\sigma^2).$$

因此 $$\overline{X}\equiv\frac{X_1+X_2+\cdots+X_n}{n}\sim N\left(\mu,\frac{\sigma^2}{n}\right).$$

习 题 十 五

1. 设 X 与 Y 相互独立，X 服从泊松分布，$E(X)=\lambda_1$，Y 也服从泊松分布，$E(Y)=\lambda_2$，试在 $X+Y=n$ 的条件下求出 X 的条件分布.

解 $X+Y=n$ 的条件下，X 的可能值为 $0,1,\cdots,n$，其条件分布

$$P(X=k\mid X+Y=n)=\frac{P(X=k,X+Y=n)}{P(X+Y=n)}=\frac{P(X=k)P(Y=n-k)}{P(X+Y=n)},$$

而

$$P(X=k)=\frac{\lambda_1^k}{k!}e^{-\lambda_1},P(Y=n-k)=\frac{\lambda_2^{n-k}}{(n-k)!}e^{-\lambda_2}.$$

对于分母，先将事件$\{X+Y=n\}$分解为$\bigcup\limits_{l=0}^{n}\{X=l,Y=n-l\}$，再由概率加法公式与 X,Y 独立性，有

$$\begin{aligned}P(X+Y=n)&=P\Big(\bigcup_{l=0}^{n}(X=l,Y=n-l)\Big)\\&=\sum_{l=0}^{n}P(X=l)P(Y=n-l)\\&=\sum_{l=0}^{n}\frac{\lambda_1^l}{l!}e^{-\lambda_1}\cdot\frac{\lambda_2^{n-l}}{(n-l)!}e^{-\lambda_2}\\&=\frac{e^{-(\lambda_1+\lambda_2)}}{n!}\sum_{l=0}^{n}C_n^l\lambda_1^l\lambda_2^{n-l}\\&=\frac{(\lambda_1+\lambda_2)^n}{n!}e^{-(\lambda_1+\lambda_2)}\end{aligned}$$

（这表明 $X+Y$ 服从参数为$\lambda_1+\lambda_2$ 的泊松分布）.

因此得

$$p(X=k\mid X+Y=n)=\frac{\dfrac{\lambda_1^k}{k!}e^{-\lambda_1}\cdot\dfrac{\lambda_2^{n-k}}{(n-k)!}e^{-\lambda_2}}{\dfrac{(\lambda_1+\lambda_2)^n}{n!}e^{-(\lambda_1+\lambda_2)}}$$

$$=C_n^k\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k}.$$

这表明条件分布为 $B\left(n,\dfrac{\lambda_1}{\lambda_1+\lambda_2}\right)$$\left(\text{参数为 } n,\dfrac{\lambda_1}{\lambda_1+\lambda_2}\text{的二项分布}\right)$.

2. 一只小猫不幸陷进一个有三扇门的大山洞中，第一扇门通到一条通道，沿此通道走 2h 后它可到达地面. 第二扇门通到另一个通道，沿它走 3h 后又回到原处，第三扇门通到第三个通道，沿它走 5h 后也会使小猫回到原处. 假定这只小猫总是等可能地在三扇门中任意选择一个. 试计算这只小猫到达地面的时间的期望.

解 设小猫到达地面所需的时间为 X，而用 Y 表示开始时小猫所选择的门编号，即

$$Y=\begin{cases}1, & \text{选第一扇门,}\\ 2, & \text{选第二扇门,}\\ 3, & \text{选第三扇门,}\end{cases}$$

则 $P(Y=1)=P(Y=2)=P(Y=3)=\dfrac{1}{3}$. 于是，

$$E(X)=\sum_{i=1}^{3}P(Y=i)E(X\mid Y=i)=\frac{1}{3}\sum_{i=1}^{3}E(X\mid Y=i).$$

显然，$Y=1$ 时 $X=2$，故 $E(X|Y=1)=2$；

$Y=2$ 时，3 小时后回到原处，故 $E(X|Y=2)=3+E(X)$；

$Y=3$ 时，5 小时后回到原处，故 $E(X|Y=3)=5+E(X)$.

代入上式得

$$E(X)=\frac{1}{3}(2+3+E(X)+5+E(X)).$$

因此，

$$E(X)=10\text{h}.$$

（从数学的严密性，还需论证 $E(X)$ 的存在且有限. 为此，令

τ 表示小猫到达地面之前在第 2 或第 3 通道经过的次数，它的分布：$P(\tau=k)=\left(\frac{2}{3}\right)^k \frac{1}{3}$， $(k=0,1,\cdots)$ 它的期望是存在且有限的，事实上，$E(\tau)=\sum_{k=0}^{\infty} k\left(\frac{2}{3}\right)^k \cdot \frac{1}{3}=2$.

显然有关系式

$$X \leqslant 5\tau+2,$$

故 $E(X)$ 存在且有限.）

评注 细心而有耐心的读者，也可用初等的传统方法，在分析综合后得出

$$E(X)=\sum_{K=1}^{\infty} \frac{1}{3^K}\left(2^{K-1} \cdot 2+(K-1) 2^{K-2} \cdot 8\right),$$

因而得到同样的结果. 可要是情况再复杂些呢？(5.6)式实在是一有力工具.

3. 设 X, Y 都是离散型随机变量，$E(Y^2)$ 存在，

$$\varphi(x)=\begin{cases}E(Y \mid X=x), & \text{当 } P(X=x)>0, \\ 0, & \text{其他}.\end{cases}$$

试证明：对任何非负函数 $\psi(x)$，只要 $E(\psi(X))^2$ 存在，必成立

$$E(\varphi(X)-Y)^2 \leqslant E(\psi(X)-Y)^2. \quad (0.5)$$

证 事实上，
$$\begin{aligned}(\psi(X)-Y)^2 &=(\varphi(X)-Y+\psi(X)-\varphi(X))^2 \\ &=(\varphi(X)-Y)^2+(\psi(X)-\varphi(X))^2+ \\ &\quad 2(\varphi(X)-Y)(\psi(X)-\varphi(X)).\end{aligned}$$

因此，为证(0.5)，只需证

$$E(\varphi(X)-Y)(\psi(X)-\varphi(X))=0. \quad (0.6)$$

设 X 的可能值为 $x_1, x_2, \cdots$，于是

$$E[Y(\psi(X)-\varphi(X))]=\sum_k P(X=x_k) E(Y(\psi(X)-\varphi(X)) \mid X=x_k)$$

$$= \sum_k P(X = x_k)[\psi(x_k) - \varphi(x_k)]E(Y \mid X = x_k)$$
$$= \sum P(X=x_k)(\psi(x_k)-\varphi(x_k))\varphi(x_k)$$
$$= E[\varphi(X)(\psi(X)-\varphi(X))].$$

这表明(0.5)式成立,因而(0.6)式成立.

4. 设一天走进某百货商店的顾客数是均值为 1 200 的随机变量,又设这些顾客所花的钱数是相互独立的,均值为 50 元的随机变量.又设任一顾客所花的钱数和进入该商店的总人数相互独立.试问该商店一天的平均营业额是多少?

解 设顾客数为 N,第 i 个顾客所花的钱为 X_i,($i=1,2,\cdots,N$).则商店一天的营业额为 $\sum_{i=1}^{N} X_i$.

$$\text{平均营业额} = E\left(\sum_{i=1}^{N} X_i\right)$$
$$= \sum_{k=1}^{\infty} P(N = k)E\left(\sum_{i=1}^{N} X_i \mid N = k\right)$$
$$= \sum P(N = k)E\left(\sum_{i=1}^{k} X_i \mid N = k\right)$$
$$= \sum P(N = k)(E(X_1) + E(X_2) + \cdots + E(X_k))$$
$$= 50\sum P(N = k)\cdot k = 50E(N) = 60\,000(\text{元})$$

习 题 十 六

1. 某食品厂为加强质量管理，对某天生产的罐头抽查了 100 个(数据如下表). 试画直方图；它是否近似服从正态分布？

100 个罐头样品的净重数据(单位：g)：

342	340	348	346	343	342	346	341	344	348
346	346	340	344	342	344	345	340	344	344
343	344	342	343	345	339	350	337	345	349
336	348	344	345	332	342	342	340	350	343
347	340	344	353	340	340	356	346	345	346
340	339	342	352	342	350	348	344	350	335
340	338	345	345	349	336	342	338	343	343
341	347	341	347	344	339	347	348	343	347
346	344	345	350	341	338	343	339	343	346
342	339	343	350	341	346	341	345	344	342

解 分 4 步解.

① 找最大、最小值，分别为 $M=356$，$m=332$.

② 取起点 $a=331.5$，终点 $b=357.5$；取 $k=12$，共分 $k+1=13$ 组，组距为 2.

③ 分组及频数：

分组 i	频数 ν_i	分组 i	频数 ν_i
331.5～333.5	1	345.5～347.5	14
333.5～335.5	1	347.5～349.5	7
335.5～337.5	3	349.5～351.5	6
337.5～339.5	8	351.5～353.5	2
339.5～341.5	15	353.5～355.5	0
341.5～343.5	21	355.5～357.5	1
343.5～345.5	21		

④ 作直方图

记 $f_i=\frac{\nu_i}{n}=\frac{\nu_i}{100}$,$(i=0,1,\cdots,12)$

$$y_i=\frac{f_i}{t_{i+1}-t_i}=\frac{f_i}{2}=\frac{\nu_i}{200}.\ (i=0,1,\cdots,12)$$

得 $y_0=0.005$,$y_1=0.005$,$y_2=0.015$,$y_3=0.04$,$y_4=0.075$,$y_5=0.105$,$y_6=0.105$,$y_7=0.07$,$y_8=0.035$,$y_9=0.03$,$y_{10}=0.01$,$y_{11}=0$,$y_{12}=0.005$.

以 13 个 y_i 为高,相应组的两端点连线为底,在 Oxy 平面上画出 13 个并立的长方形,即得直方图(图略).它近似服从正态分布.

2. 设 $x_1,x_2,\cdots,x_n$ 是来自正态分布 $N(\mu,\sigma^2)$ 的样本值.μ 已知,求 σ^2 的最大似然估计量.

解 似然函数

$$L(\mu,\sigma^2;x_1,x_2,\cdots,x_n)=\frac{1}{(2\pi)^{\frac{n}{2}}\sigma^n}\mathrm{e}^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{n}(x_i-\mu)^2}$$

取对数,

$$\ln L=-\frac{n}{2}\ln(2\pi)-\frac{n}{2}\ln\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2.$$

将上式对 σ^2 求导(注意,这里是对 σ^2,而不是 σ 求导),令导数等于 0,有

$$\frac{\partial\ln L}{\partial(\sigma^2)}=-\frac{n}{2}\cdot\frac{1}{\sigma^2}+\frac{1}{2\sigma^4}\sum(x_i-\mu)^2=0,$$

所以
$$\hat{\sigma}^2=\frac{1}{n}\sum_{1}^{n}(x_i-\mu)^2.$$

3. 设 $x_1,x_2,\cdots,x_n$ 是来自正态分布 $N(\mu,1)$ 的样本值,求 μ 的最大似然估计量.

解 似然函数

$$L(\mu;x_1,x_2,\cdots,x_n)=\frac{1}{(2\pi)^{\frac{n}{2}}}\mathrm{e}^{-\frac{1}{2}\sum(x_i-\mu)^2},$$

取对数,

$$\ln L = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum(x_i-\mu)^2.$$

将上式对 μ 求导，令导数等于 0，有：

$$\frac{\partial \ln L}{\partial \mu} = -\frac{1}{2}\sum 2(x_i-\mu)(-1)=0,$$

所以
$$\hat{\mu}=\frac{1}{n}\sum x_i=\bar{x}.$$

4. 设 X 服从区间$[0,\lambda]$($\lambda>0$)上的均匀分布，λ 是未知参数. 设 $x_1,x_2,\cdots,x_n$ 是 X 的样本值，试求出 λ 的最大似然估计量和矩估计量.

解 似然函数

$$L(\lambda;x_1,x_2,\cdots,x_n)=\begin{cases}\dfrac{1}{\lambda^n}, & 0\leqslant x_1,x_2,\cdots,x_n\leqslant\lambda,\\ 0, & \text{其他}.\end{cases}$$

我们看出，参数 λ 越小，似然函数值越大；另一方面，λ 又不能太小，它必须"$\geqslant x_1,x_2,\cdots,x_n$"（注意：$x_1,x_2,\cdots,x_n$ 是取定的样本值），否则，L 值为 0. 因此，最大似然估计量$\hat{\lambda}$应是在限制条件"$\geqslant x_1,\cdots,x_n$"下的最小值，所以

$$\hat{\lambda}=\max\{x_1,x_2,\cdots,x_n\}.$$

现在求 λ 的矩估计量. 由于 $E(X)=\dfrac{\lambda}{2}$，得 $\lambda=2E(X)$，所以 λ 的矩估计量

$$\tilde{\lambda}=2\,\bar{x}.$$

5. 对 § 5 的例 5.1，分别对置信度 0.99，0.90 找出均值的置信区间.

解 在 § 5 例 5.1 的讨论中，已求出$\bar{x}=15.06$，并对置信度为 0.95 时，找出了置信区间

$$\left(\bar{x}-1.96\sqrt{\frac{D(X)}{n}},\bar{x}+1.96\sqrt{\frac{D(X)}{n}}\right)$$，其中 $n=6$，$D(X)=0.05$.

注意到这里的1.96，实际上满足 $\Phi(1.96)=\dfrac{1+0.95}{2}=0.975$，即它是标准正态随机变量的0.975分位数.（可由附表1查到）. 因此，为解本题只需分别找出 $\dfrac{1+0.99}{2}$（即0.995）与 $\dfrac{1+0.90}{2}$（即0.95）的（标准正态随机变量）分位数即可. 它们分别为2.58，1.65. 所以相应的置信区间是

$$\left(\bar{x}-2.58\sqrt{\frac{D(X)}{n}},\bar{x}+2.58\sqrt{\frac{D(X)}{n}}\right)=(15.06-0.24,15.06+0.24)$$

与

$$\left(\bar{x}-1.65\sqrt{\frac{D(X)}{n}},\bar{x}+1.65\sqrt{\frac{D(X)}{n}}\right)=(15.06-0.15,15.06+0.15).$$

6. 对§5的例5.2，分别对置信度0.99与0.90，找出均值的置信区间.

解 在§5例5.2的讨论中，已求出 $\bar{x}=1\,259$，$\sqrt{\dfrac{S^2}{n}}=5.339$，并对置信度为0.95的情况，找出了置信区间

$$\left(\bar{x}-2.776\sqrt{\frac{S^2}{n}},\bar{x}+2.776\sqrt{\frac{S^2}{n}}\right)$$

注意：这里的2.776是从附表2（对自由度为5－1，$\alpha=0.05$）直接查得（它实质上是自由度为4的 t 分布的0.975分位数，但由于附表2跟附表1的结构上的差异，查表方式不一样）. 因此，为解本题，只需直接查附表2. 得临界值

4.604 （自由度 $5-1=4$，$\alpha=0.01$），

2.132 （自由度 $5-1=4$，$\alpha=0.10$），

所以相应的置信区间是

$$\left(\bar{x}-4.604\sqrt{\frac{S^2}{n}},\bar{x}+4.604\sqrt{\frac{S^2}{n}}\right)=(1\,259-24.6,1\,259+24.6),$$

$$\left(\bar{x}-2.132\sqrt{\frac{S^2}{n}},\bar{x}+2.132\sqrt{\frac{S^2}{n}}\right)=(1\,259-11.4,1\,259+11.4).$$

7. 已知样本 3.3，−0.3，−0.6，−0.9，求具有 $\sigma=3$ 的正态分布的均值的置信区间（置信度为 0.95）. 如果 σ 未知，问均值的置信区间为何？

解 $\bar{x}=\frac{1}{4}(3.3-0.3-0.6-0.9)=0.375$，

$$\sqrt{\frac{\sigma^2}{n}}=\sqrt{\frac{9}{4}}=1.5.$$

由于 $\Phi(1.96)=0.975$，因此均值的置信度为 0.95 的置信区间

$$(0.375-1.96\times1.5,0.375+1.96\times1.5)$$
$$=(0.375-2.94,0.375+2.94).$$

如果 σ 未知，计算得 $S^2=(1.965)^2$，$\sqrt{\frac{S^2}{n}}=0.982\,5$. 对自由度 $=4-1=3$，$\alpha=0.05$，查附表 2 得 $\lambda=3.182$，所以均值的置信度为 0.95 的置信区间

$$(0.375-3.182\times0.982\,5,0.375+3.182\times0.982\,5)$$
$$=(0.375-3.126,0.375+3.126).$$

8. 对某一距离进行 5 次独立测量，得（单位：m）：

2 781，2 836，2 807，2 763，2 858，

已知测量无系统误差，求该距离的置信度为 0.95 的置信区间（测量值可认为服从正态分布）.

解 记 X = 测量值. 由题设 $X\sim N(\mu,\sigma^2)$，而 μ 即为该距离，σ^2 未知，对 $n=5$ 的样本值，算得

$$\bar{x}=2\,809,\ \sqrt{\frac{S^2}{n}}=17.37.$$

查附表 2（$\alpha=0.05$，自由度 $=5-1=4$）得 $\lambda=2.776$，故 $\lambda\sqrt{\frac{S^2}{n}}=48.2$. 所以，所求的置信区间为 $(2\,809-48.2,2\,809+48.2)$.

9. 为了估计灯泡使用时数的均值 μ 及标准差 σ，测试 10 个灯泡，得 $\bar{x}=1\,500$ h，$S=20$ h. 如果已知灯泡使用时数是服从正态分布的，求 μ 及 σ 的置信区间（置信度为 0.95）.

解 先求均值 μ 的置信区间.

$$\bar{x}=1\,500,\sqrt{\frac{S^2}{n}}=\frac{20}{\sqrt{10}}=6.325,$$

查附表 2(自由度$=10-1=9$,$\alpha=0.05$)得 $\lambda=2.262$,于是 $\lambda\sqrt{\frac{S^2}{n}}=14.3$,故 μ 的置信度为 0.95 的置信区间是

$$(1\,500-14.3,1\,500+14.3).$$

再求标准差 σ 的置信区间. 由本章 §6 中(6.6)式知 σ 的置信区间为

$$\left(\sqrt{\frac{\sum(x_i-\bar{x})^2}{\lambda_2}},\sqrt{\frac{\sum(x_i-\bar{x})^2}{\lambda_1}}\right),$$

其中 λ_1,λ_2 是 $\chi^2(n-1)$的分位数,可由附表 3 查出. 本题中

$$\sum(x_i-\bar{x})^2=9\cdot S^2=3\,600.$$

对于自由度$=10-1=9$,查附表 3,得

$$P(\chi^2>2.70)=0.975,P(\chi^2>19.0)=0.025.$$

于是 $$\lambda_1=2.70,\lambda_2=19.0.$$

因此

$$\left(\sqrt{\frac{\sum(x_i-\bar{x})^2}{\lambda_2}},\sqrt{\frac{\sum(x_i-\bar{x})^2}{\lambda_1}}\right)=\left(\sqrt{\frac{3\,600}{19.0}},\sqrt{\frac{3\,600}{2.70}}\right)=(13.8,36.5).$$

10. 测量铝的相对密度 16 次,测得$\bar{x}=2.705$,$S=0.029$. 试求铝的相对密度的置信区间(设测量值服从正态分布,置信度为 0.95).

解 $\bar{x}=2.705,\sqrt{\frac{S^2}{n}}=\frac{S}{4}=0.007\,25.$

查附表 2(自由度$=16-1=15$,$\alpha=0.05$)得 $\lambda=2.131$,于是

$$\lambda\sqrt{\frac{S^2}{n}}=0.015.$$

故铝的相对密度的置信度为 0.95 的置信区间为

$$(2.705-0.015,2.705+0.015).$$

11. 设 $X\sim N(\mu,\sigma^2)$，$x_1,x_2,\cdots,x_n$ 是其样本值. 如果 σ^2 已知，问：n 取多大值时，方能保证 μ 的置信度为 0.95 的置信区间的长度不大于给定的 L？

解 由于 μ 的置信区间长度（置信度为 0.95）为

$$2\times 1.96\sqrt{\frac{\sigma^2}{n}},$$

由题意，它不大于 L，所以

$$n\geqslant\frac{(2\times 1.96)^2\cdot\sigma^2}{L^2}=15.37\sigma^2/L^2.$$

12. 随机地从甲批导线中抽取 4 根，从乙批导线中抽取 5 根，测得其电阻为（单位：Ω）

甲批导线：0.143，0.142，0.143，0.137.

乙批导线：0.140，0.142，0.136，0.138，0.140.

设甲、乙两批导线的电阻分别服从 $N(\mu_1,\sigma^2)$，$N(\mu_2,\sigma^2)$（并且它们相互独立），σ^2 已知为 $0.002\,5^2$，但 μ_1,μ_2 均未知，试求 $\mu_1-\mu_2$ 的置信度为 0.95 的置信区间.

解 记 X 为甲批导线的电阻，Y 为乙批导线的电阻.

先计算样本均值差

$$\begin{aligned}\bar{x}-\bar{y}=&\frac{1}{4}(0.143+0.142+0.143+0.137)-\\&\frac{1}{5}(0.140+0.142+0.136+0.138+0.140)\\=&0.002.\end{aligned}$$

我们知道（参见习题十二第 3 题），$\overline{X}\sim N\left(\mu_1,\frac{\sigma^2}{4}\right)$，$\overline{Y}\sim N\left(\mu_2,\frac{\sigma^2}{5}\right)$，

还有

$$\overline{X}-\overline{Y}\sim N\left(\mu_1-\mu_2,\frac{\sigma^2}{4}+\frac{\sigma^2}{5}\right),$$

所以

$$P\left\{\left|\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\left(\frac{1}{4}+\frac{1}{5}\right)\sigma^2}}\right|<1.96\right\}=0.95.$$

由此得 $\mu_1-\mu_2$ 的置信度为 0.95 的置信区间：

$$\left(\bar{x}-\bar{y}-1.96\sqrt{\frac{1}{4}+\frac{1}{5}}\sigma,\bar{x}-\bar{y}+1.96\sqrt{\frac{1}{4}+\frac{1}{5}}\sigma\right),$$

而 $1.96\sqrt{\frac{1}{4}+\frac{1}{5}}\sigma=0.003$，所以所求的置信区间是

$$(0.002-0.003,0.002+0.003).$$

习 题 十 七

1. 由经验知某零件质量 $X \sim N(\mu, \sigma^2)$，$\mu = 15$，$\sigma^2 = 0.05$，技术革新后，抽了六个样品，测得质量为(单位：g)：

$$14.7, 15.1, 14.8, 15.0, 15.2, 14.6.$$

已知方差不变，问平均质量是否仍为 15 g($\alpha = 0.05$)？

解 提出零假设 $H_0: \mu = 15$. 计算

$$|u| = \left| \frac{\bar{x} - 15}{\sqrt{\frac{\sigma^2}{n}}} \right| = \left| \frac{14.9 - 15}{\sqrt{\frac{0.05}{6}}} \right| = 1.095;$$

查正态分布表. 对 $\alpha = 0.05$，得临界值 $\lambda = 1.96$，即 $P(|u| > 1.96) = 0.05$；由于 $1.095 < 1.96$，故 H_0 相容，即不能否定 H_0，亦即还不能认为技术革新后零件平均质量起了变化.

2. 糖厂用自动打包机打包. 每包标准质量为 100kg. 每天开工后需要检验一次打包机工作是否正常，即检查打包机是否有系统偏差. 某日开工后测得 n 包质量(单位：kg)如下：

$$99.3, 98.7, 100.5, 101.2, 98.3, 99.7, 99.5, 102.1, 100.5.$$

问：该日打包机工作是否正常？($\alpha = 0.05$；已知包质量服从正态分布.)

解 该日包质量 $X \sim N(\mu, \sigma^2)$. σ^2 未知，$n = 9$.

零假设 $H_0: \mu = 100$；

算得 $\bar{x} = 99.98$，$S^2 = (1.21)^2$. 因此统计量

$$T = \frac{\bar{x} - 100}{\sqrt{\frac{S^2}{n}}} = \frac{99.98 - 100}{\sqrt{\frac{(1.21)^2}{9}}} = -0.05;$$

查附表 2($\alpha = 0.05$；自由度 $= 9 - 1 = 8$)得 $\lambda = 2.306$；由于 $0.05 < 2.306$，所以，不能否定 H_0，即没有发现打包机有系统偏差.

3. 正常人的脉搏平均为 72 次/分. 某医生测得 10 例慢性四

乙基铅中毒患者的脉搏(次/分)如下:

$$54,67,68,78,70,66,67,70,65,69.$$

问:四乙基铅中毒者和正常人的脉搏有无显著性差异?(已知四乙基铅中毒者的脉搏服从正态分布).

解 铅中毒患者脉搏 $X \sim N(\mu,\sigma^2)$,σ^2 未知,$n=10$.

$$H_0:\mu=72;$$

算得$\bar{x}=67.4$,$S^2=5.93^2$.统计量

$$T=\frac{\bar{x}-72}{\sqrt{\frac{S^2}{n}}}=\frac{67.4-72}{\sqrt{\frac{5.93^2}{10}}}=\frac{-4.6}{1.88}=-2.45;$$

查附表 2($\alpha=0.05$,自由度$=10-1=9$)得 $\lambda=2.262$;由于 $|T|=2.45>2.262$,所以否定 H_0,即铅中毒患者与正常人的脉搏有显著差异.

4. 用热敏电阻测温仪间接测量地热勘探井底温度,重复测量 7 次,测得温度(℃):

$$112.0,113.4,111.2,112.0,114.5,112.9,113.6.$$

而用某精确办法测得温度为 112.6℃(可看作温度真值),试问用热敏电阻测温仪间接测温有无系统偏差($\alpha=0.05$)?

解 可以认为热敏电阻测温仪的测量值 $X \sim N(\mu,\sigma^2)$.

$$H_0:\mu=112.6;$$

算得$\bar{x}=112.8$,$S^2=(1.136)^2$,$\sqrt{\frac{S^2}{n}}=0.4293$.统计量

$$T=\frac{\bar{x}-112.6}{\sqrt{\frac{S^2}{n}}}=0.4659;$$

查附表 2($\alpha=0.05$,自由度$=7-1=6$),得 $\lambda=2.447$;由于 $|T|<\lambda$,不能否定 H_0,即没有发现测温仪有系统偏差.

5. 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 $S=0.007$ Ω. 设总体为正态分布. 问在水平 $\alpha=0.05$ 下能认为这批导线的标准差显著地

偏大吗？

解 这批导线的电阻 $X\sim N(\mu,\sigma^2)$，μ 未知，$n=9$.

$$H_0:\sigma^2\leqslant 0.005^2;$$

算得统计量

$$W=\frac{\sum(x_i-\bar{x})^2}{\sigma_0^2}=\frac{(9-1)0.007^2}{0.005^2}=15.68;$$

查附表 3($\alpha=0.05$，自由度$=9-1=8$)得 $\lambda=15.5$；由于 $W>\lambda$，所以否定 H_0，即可以认为这批导线电阻的标准差偏大.

6. 机床厂某日从两台机器所加工的同一种零件中，分别抽若干个样品测量零件尺寸(cm)，得：

第一台机器：6.2，5.7，6.5，6.0，6.3，5.8，5.7，6.0，6.0，5.8，6.0；

第二台机器：5.6，5.9，5.6，5.7，5.8，6.0，5.5，5.7，5.5.

问：这两台机器的加工精度是否有显著性差异($\alpha=0.05$)？

解 第一台机器所加工的零件尺寸 $X\sim N(\mu_1,\sigma_1^2)$，

第二台机器所加工的零件尺寸 $Y\sim N(\mu_2,\sigma_2^2)$.

$$H_0:\sigma_1=\sigma_2;$$

现 $n_1=11$，$n_2=9$，并算得 $S_1^2=0.064$，$S_2^2=0.030$.

统计量 $F=\dfrac{S_1^2}{S_2^2}=2.13$；

查附表 5($\alpha=0.025$，第一自由度$=11-1=10$，第二自由度$=9-1=8$)，得 $\lambda_2=4.30$. 再查附表 5($\alpha=0.025$，第一自由度$=8$，第二自由度$=10$)得 $\lambda_1=\dfrac{1}{3.85}$；而

$$\frac{1}{3.85}<2.13<4.30,$$

故 H_0 相容，即还没有发现这两台机器加工精度有显著性差异.

7. 检查了 26 匹马，测得每 100 mL 的血清中，所含的无机磷平均为 3.29 mg，标准差为 0.34 mg；又检查了 18 头羊，100 mL 的血清中含无机磷平均为 3.96 mg，标准差为 0.40 mg. 试以 0.05 的

检验水平. 检验马与羊的血清中含无机磷的量是否有显著性差异?

解 马(每 100 mL 血清中)含无机磷 $X\sim N(\mu_1,\sigma^2)$, $n_1=26$;

羊(每 100 mL 血清中)含无机磷 $Y\sim N(\mu_2,\sigma^2)$, $n_2=18$.

(这里马、羊的方差相等是由经验知道的. 自然,也可以利用 § 4 中所介绍的方法,先直接对这批数据进行方差相等的检验.)

$$H_0:\mu_1=\mu_2;$$

现已算得 $\bar{x}=3.29$, $\bar{y}=3.96$, $S_1=0.34$, $S_2=0.40$. 统计量

$$T=\frac{\bar{x}-\bar{y}}{\sqrt{\sum(x_i-\bar{x})^2+\sum(y_i-\bar{y})^2}}\cdot\sqrt{\frac{n_1\cdot n_2(n_1+n_2-2)}{n_1+n_2}}.$$

将 $\sum(x_i-\bar{x})^2=(n_1-1)S_1^2=25\cdot(0.34)^2$,

$\sum(y_i-\bar{y})^2=(n_2-1)S_2^2=17\cdot(0.40)^2$

代入,得到

$$T=\frac{3.29\quad 3.96}{\sqrt{25(0.34)^2+17(0.40)^2}}\sqrt{\frac{26\times18\times42}{44}}=-5.98;$$

查附表 2($\alpha=0.05$, 自由度 $=26+18-2=42$), 得 $\lambda=2.021$; 现

$$|T|=5.98>\lambda,$$

否定 H_0, 即认为它们有显著性差异.

8. 十个失眠患者,服用甲、乙两种安眠药,延长睡眠的时间如下表所示

	a	b	c	d	e	f	g	h	i	j
甲	1.9	0.8	1.1	0.1	−0.1	4.4	5.5	1.6	4.6	3.4
乙	0.7	−1.6	−0.2	−1.2	−0.1	3.4	3.7	0.8	0	2.0

问这两种安眠药的疗效有无显著性差异(可以认为服用两种安眠药后增加的睡眠时间之差近似服从正态分布)($\alpha=0.05$)?

解 X = 服用甲种安眠药后增加的睡眠时间,

Y = 服用乙种安眠药后增加的睡眠时间,

$Z=X-Y\sim N(\mu,\sigma^2)$.

对 Z 有 $n=10$ 个样本值：

$$1.2,2.4,1.3,1.3,0,1.0,1.8,0.8,4.6,1.4.$$

得

$$\bar{z}=1.58, S_z^2=1.23^2.$$

待验假设 $H_0:\mu=0$；

统计量

$$T=\frac{\bar{z}-0}{\sqrt{\frac{S^2}{n}}}=\frac{1.58}{1.23}\sqrt{10}=4.06;$$

查附表 2($\alpha=0.05$，自由度$=10-1=9$)，得 $\lambda=2.262$；现 $4.06>2.262$. 所以否定 H_0，即甲、乙两种安眠药的疗效有显著性差异.

9. 比较甲、乙两种安眠药的疗效，将 20 个患者分成两组，每组 10 人；甲组病人服用甲种安眠药，乙组病人服用乙种安眠药. 如服药后延长的睡眠时间分别近似服从正态分布，其数据仍如上题(自然，数据不是两两成对了). 问这两种安眠药的疗效有无显著性差异($\alpha=0.05$)?

解 X,Y 同上题. 我们假定它们的方差相同(事实上，可先用 § 4 中所介绍的方法. 进行方差相同的检验). 因此，$X\sim N(\mu_1,\sigma^2)$，$Y\sim N(\mu_2,\sigma^2)$，$n_1=n_2=10=n$.

待验假设 $H_0:\mu_1=\mu_2$；

使用统计量

$$T=\frac{\bar{x}-\bar{y}}{\sqrt{\frac{S_1^2+S_2^2}{n}}}.$$

现算得

$$\bar{x}=2.33, \bar{y}=0.75, S_1^2=4, S_2^2=1.789^2;$$

代入得

$$T=\frac{2.33-0.75}{\sqrt{\frac{4+(1.789)^2}{10}}}=1.86;$$

查附表 2($\alpha=0.05$，自由度$=20-2=18$)得 $\lambda=2.101$；由于

$$1.86<2.101,$$

故 H_0 相容，不能否定 H_0，即没有发现甲、乙两种安眠药的疗效有显著性差异.

10. 在一正 20 面体的 20 个面上，分别标以数字 0，1，…，9，每个数字在两个面上标出. 为检验其匀称性，共作 800 次投掷试验，数字 0，1，…，9 朝正上方的次数如下：

数字	0	1	2	3	4	5	6	7	8	9
频数	74	92	83	79	80	73	77	75	76	91

问：该正 20 面体是否匀称？

解 对随机变量 X = 朝正上方的数字，提出待验假设 H_0：$p_i=P(X=i)=\frac{1}{10}$，$i=0,1,\cdots,9$.（即 H_0：该正 20 面体是匀称的.）

为计算统计量 $\chi^2=\sum\limits_{i=0}^{9}\frac{(\nu_i-np_i)^2}{np_i}$ 的值，列表如下 $\left(np_i=\frac{800}{10}=80\right)$：

数字	0	1	2	3	4	5	6	7	8	9
频数 ν_i	74	92	83	79	80	73	77	75	76	91
ν_i-np_i	-6	12	3	-1	0	-7	-3	-5	-4	11
$(\nu_i-np_i)^2$	36	144	9	1	0	49	9	25	16	121

得 $\chi^2=5.125$；查附表 3（$\alpha=0.05$，自由度 $=10-1=9$），得 $\lambda=16.9$，而

$$5.125<16.9,$$

故 H_0 相容，即没有发现该正 20 面体不匀称.

11. 某工厂采用新法处理废水，对处理后的水测量所含某种有毒物质的浓度，得到 10 个数据（单位：μg/L）：

$$22,14,17,13,21,16,15,16,19,18.$$

而以往用老法处理废水后，该种有毒物质的平均浓度为 $19\mu g/L$. 问：新法是否比老法效果好？（检验水平 $\alpha=0.05$.）

解 有毒物质浓度的测量值 $X\sim N(\mu,\sigma^2)$.

$$H_0:\mu>19;$$

由于 σ^2 未知，我们采用统计量 T.

$$T=\frac{\bar{x}-19}{\sqrt{\frac{S^2}{n}}}.$$

算得 $\bar{x}=17.1, S^2=2.92^2$. 代入有

$$T=\frac{17.1-19}{2.92}\sqrt{10}=-2.06;$$

查附表 2（注意到这是单边 $\mu>\mu_0$，检验水平 0.05；自由度 $=10-1=9$），得 $\lambda=-1.833$；现

$$T<\lambda,$$

否定 H_0，即认为新方法比老方法显著地好.

习 题 十 八

1. 炼铝厂测得所产铸模用的铝的硬度 x 与抗张强度 y 数据如下：

x	68	53	70	84	60	72	51	83	70	64
y	288	293	349	343	290	354	283	324	340	286

求 y 对 x 的回归直线.

解 先求几个基本量：

$$\bar{x}=67.5,\bar{y}=315.$$

$$l_{xx}=\sum(x_i-\bar{x})^2=1\,096.5,l_{yy}=\sum(y_i-\bar{y})^2=7\,870.$$

$$l_{xy}=\sum(x_i-\bar{x})(y_i-\bar{y})=2\,047.$$

从而得

$$\hat{b}=\frac{l_{xy}}{l_{xx}}=\frac{2\,047}{1\,096.5}=1.867,$$

$$\hat{a}=\bar{y}-\hat{b}\bar{x}=188.99,$$

回归直线方程：$\hat{y}=188.99+1.867x$.

2. 检验第 1 题所得回归直线的显著性.

解 所谓对回归直线的显著性进行检验是指，在假定（即大前提）

$$Y_i=a+bx_i+\varepsilon_i,\quad i=1,2,\cdots,n,$$

且 $\varepsilon_i(i=1,2,\cdots,n)$ 独立同分布 $N(0,\sigma^2)$ 下，对零假设 $H_0:b=0$ 进行检验. 具体过程如下：

由上题的几个基本量，可算得

$$U=\hat{b}\cdot l_{xy}=3\,821.75;$$

再由平方和分解公式，可算得

$$Q=l_{yy}-U=4\,048.25,$$

则统计量 $F=\dfrac{U}{Q/(n-2)}=\dfrac{8U}{Q}=7.55$；

查附表 5($\alpha=0.05$，第一自由度 1，第二自由度 $=n-2=8$)得 $\lambda=5.32$. 现 $F>\lambda$，H_0 被否定，即回归直线显著.

3. 对于第 1 题所讨论的问题，试预报铝的硬度 $x=65$ 时的抗张强度 y.

解 首先 $x_0=65$ 时，代入回归直线方程，得

$$\hat{y}_0=188.99+1.867\times 65\approx 310,$$

则预报区间为

$$(310-d,310+d),$$

其中

$$d=\lambda S\sqrt{1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{l_{xx}}}.$$

查自由度为 $n-2=8$ 的 t 分布临界值表得 $\lambda=2.306(\alpha=0.05)$. 又知

$$S=\sqrt{\frac{Q}{n-2}}=\sqrt{\frac{4\,048.25}{8}}=22.5,$$

$$\sqrt{1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{l_{xx}}}=\sqrt{1+\frac{1}{10}+\frac{(65-67.5)^2}{1\,096.5}}=1.05,$$

代入得 $d=54$，故预报区间($\alpha=0.05$，即置信度 0.95)为

$$(310-54,310+54).$$

习题十九

（本章的两个习题，实际上是对三个实例施行正交设计并对试验结果进行计算与分析. 为节省篇幅，这里省略了题目，并综合一起给出题解.）

解　(1)（对第1个实例）

因素 / 列号 / 试验号	上升温度	保温时间	出炉温度	硬度合格率(%)
	1	2	3	
1	1(800℃)	1(6 h)	1(400℃)	100
2	2(820℃)	1	2(500℃)	45
3	1	2(8 h)	2	85
4	2	2	1	70
Ⅰ＝位级1的两次合格率之和	185	145	170	总和＝300
Ⅱ＝位级2的两次合格率之和	115	155	130	
R＝Ⅰ，Ⅱ中“大－小”	70	10	40	

可能好配合为：上升温度 800℃，保温时间 8 h，出炉温度 400℃.

(2)（对第2个实例）

可能好配合为：品种——窄叶青，密度——30万棵/亩，施肥量——10斤/亩.

密度、氮肥量的趋势图见图 9.

试验号 \ 列号 \ 因素	品种 1	密度（万棵/亩） 2	施肥量（斤/亩） 3	亩产（斤）
1	1(窄叶青)	1(30)	3(15)	839
2	2(南二矮)	1	1(10)	761
3	3(珍珠矮)	1	2(5)	688
4	1	2(25)	2	734
5	2	2	3	774
6	3	2	1	754
7	1	3(20)	1	843
8	2	3	2	676
9	3	3	3	726
Ⅰ＝位级 1 的三次产量之和	2 416	2 288	2 358	总和＝6 795
Ⅱ＝位级 2 的三次产量之和	2 211	2 262	2 098	
Ⅲ＝位级 3 的三次产量之和	2 168	2 245	2 339	
R＝Ⅰ、Ⅱ、Ⅲ中“大－小”	248	43	260	

从图 9 看出，产量随密度的增大而提高，因此再做试验应考虑到再加密的位级；而氮肥用量为 10 斤/亩或 15 斤/亩时产量相差无几，应在用量为 10 左右再做试验.

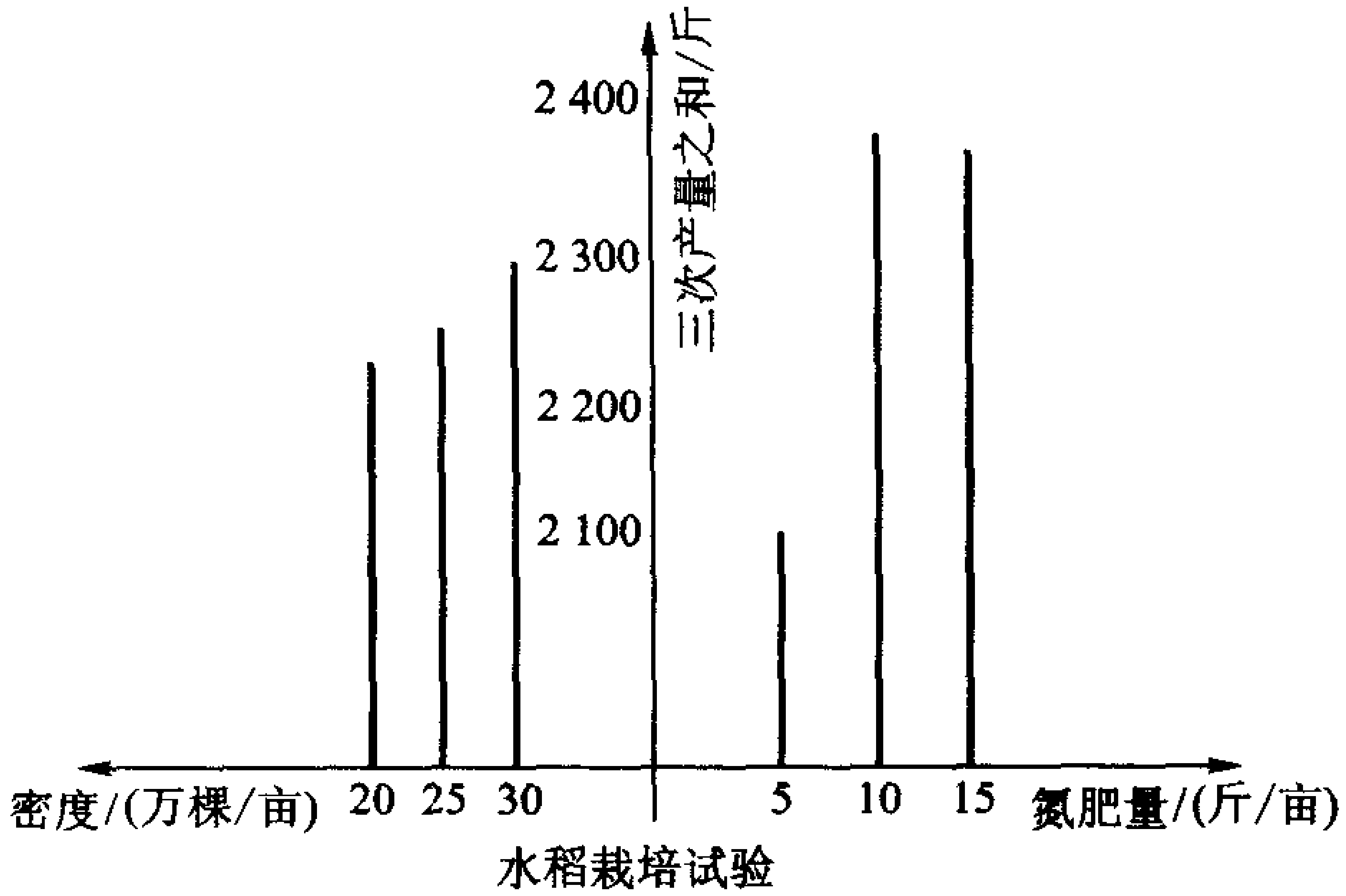

图 9

趋势图直观形象，它为进一步试验提供指导性建议.

(3)（对第 3 个实例）

可能好配合：pH9～10，不加凝聚剂，用 NaOH 作沉淀剂，不加 $CaCl_2$（废水浓度的极差很小，意味着浓、稀对结果影响不大，即废水浓度稀浓都可以）. 从 pH 的趋势图（图 10）看出，pH 从 9～11 都可以.

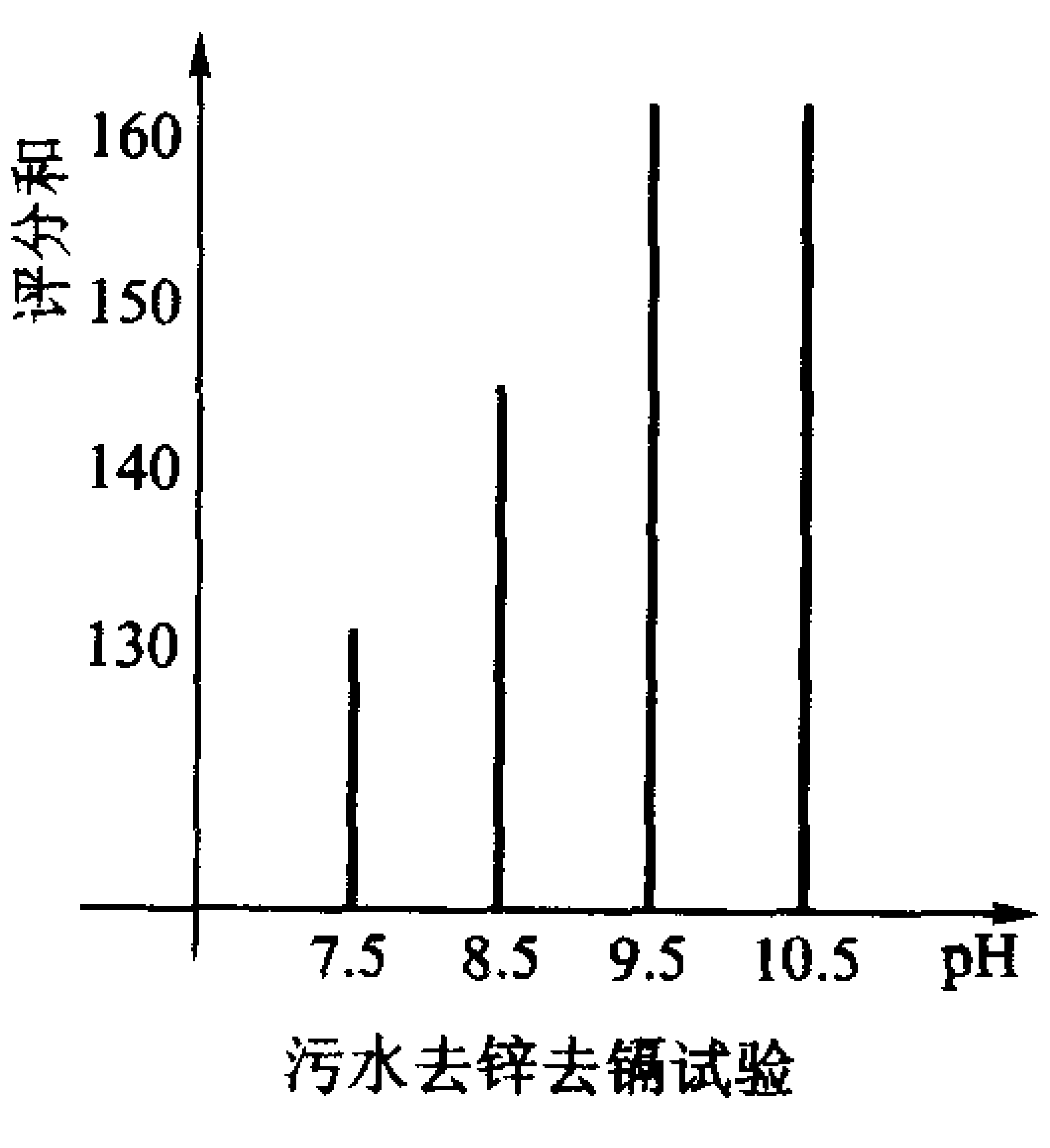

图 10

因素 列号 试验号	pH 1	凝聚剂 2	沉淀剂 3	$CaCl_2$ 4	废水浓度 5	综合评分
1	1(7～8)	1(加)	2(Na_2CO_3)	2(加)	1(稀)	45
2	3(9～10)	2(不加)	2	1(不加)	1	70
3	2(8～9)	2	2	2	2(浓)	55
4	4(10～11)	1	2	1	2	65
5	1	2	1(NaOH)	1	2	85
6	3	1	1	2	2	95
7	2	1	1	1	1	90
8	4	2	1	2	1	100
Ⅰ=位级1评分之和	130	295	370	310	305	总和=605
Ⅱ=位级2评分之和	145	310	235	295	300	
Ⅲ=位级3评分之和	165					
Ⅳ=位级4评分之和	165					
极差R=Ⅰ,Ⅱ,Ⅲ,Ⅳ中“最大－最小”	35	15	135	15	5	

习 题 二 十

1. 设 X 是离散型或连续型随机变量，$E(X^2)$存在. 试证明：为了使 $E(X-a)^2$（a 是实数）达到最小值，必须且只须 $a=E(X)$.

证 注意到

$$\begin{aligned}E(X-a)^2 &= E(X-E(X)+E(X)-a)^2\\ &= E(X-E(X))^2+(E(X)-a)^2 \text{（交叉项等于 0）}.\end{aligned}$$

上式右端第 1 项为 X 的方差，与 a 无关；而第 2 项当且仅当 $a=E(X)$时达到最小（$=0$）. 命题得证.

2. 设 X 是离散型或连续型随机变量，$E(X)$存在. 试证明：为了使 $E|X-a|$（a 是实数）达到最小值，必须且只须 a 是 X 的中位数.

证 首先按中位数 $x_{\frac{1}{2}}$ 的定义（见第三章§5 的 3）.

$$P(X<x_{\frac{1}{2}})\leqslant\frac{1}{2}\leqslant P(X\leqslant x_{\frac{1}{2}}). \qquad (0.7)$$

对连续型随机变量 X，其中位数 $x_{\frac{1}{2}}$ 有 $F(x_{\frac{1}{2}})=\dfrac{1}{2}$. 对任实数 a，有

$$\begin{aligned}E|X-a| &= \int_{-\infty}^{+\infty}|x-a|\,p(x)\mathrm{d}x\\ &= \int_{-\infty}^{a}(a-x)p(x)\mathrm{d}x+\int_{a}^{+\infty}(x-a)p(x)\mathrm{d}x\end{aligned}$$

（选定一个中位数 $x_{\frac{1}{2}}$）$= \displaystyle\int_{-\infty}^{x_{\frac{1}{2}}}(a-x)p(x)\mathrm{d}x+\int_{x_{\frac{1}{2}}}^{a}(a-x)p(x)\mathrm{d}x+$

$$\int_{x_{\frac{1}{2}}}^{+\infty}(x-a)p(x)\mathrm{d}x-\int_{x_{\frac{1}{2}}}^{a}(x-a)p(x)\mathrm{d}x.$$

$$=aF(x_{\frac{1}{2}})-\int_{-\infty}^{x_{\frac{1}{2}}}xp(x)\mathrm{d}x-a(1-F(x_{\frac{1}{2}}))+$$

$$\int_{x_{\frac{1}{2}}}^{+\infty} xp(x)\mathrm{d}x + 2\int_{x_{\frac{1}{2}}}^{a} (a-x)p(x)\mathrm{d}x$$

$$= \int_{x_{\frac{1}{2}}}^{+\infty} xp(x)\mathrm{d}x - \int_{-\infty}^{x_{\frac{1}{2}}} xp(x)\mathrm{d}x +$$

$$2\int_{x_{\frac{1}{2}}}^{a} (a-x)p(x)\mathrm{d}x.$$

注意到上式右端第 1,2 项与 a 无关，而第 3 项对任意 a，有

$$\int_{x_{\frac{1}{2}}}^{a} (a-x)p(x)\mathrm{d}x \geqslant 0.$$

而当且仅当 a 为 X 的中位数时，取到等号，即取到最小值.

当 X 为离散型随机变量时，由于 $F(x_{\frac{1}{2}})$ 不一定为 $\frac{1}{2}$，而给证明带来不小的困难. 这里仅给感兴趣的读者予以参考.

记 $p_i = P(X = x_i)$，$i = 1,2,\cdots$. 对 $a \geqslant x_{\frac{1}{2}}$，

$$E\mid X-a\mid = \sum_i \mid x_i - a\mid p_i = \sum_{x_i\leqslant a}(a-x_i)p_i + \sum_{a<x_i}(x_i-a)p_i$$

$$= \sum_{x_i\leqslant x_{\frac{1}{2}}}(a-x_i)p_i + \sum_{x_{\frac{1}{2}}<x_i\leqslant a}(a-x_i)p_i +$$

$$\sum_{x_{\frac{1}{2}}<x_i}(x_i-a)p_i - \sum_{x_{\frac{1}{2}}<x_i\leqslant a}(x_i-a)p_i$$

$$= a\sum_{x_i\leqslant x_{\frac{1}{2}}} p_i - \sum_{x_i\leqslant x_{\frac{1}{2}}} x_ip_i - a\sum_{x_{\frac{1}{2}}<x_i} p_i + \sum_{x_{\frac{1}{2}}<x_i} x_ip_i +$$

$$2\sum_{x_{\frac{1}{2}}<x_i\leqslant a}(a-x_i)p_i$$

$$= a\Big(2\sum_{x_i\leqslant x_{\frac{1}{2}}} p_i - 1\Big) + \Big(\sum_{x_{\frac{1}{2}}<x_i} x_ip_i - \sum_{x_i\leqslant x_{\frac{1}{2}}} x_ip_i\Big) +$$

$$2\sum_{x_{\frac{1}{2}}<x_i\leqslant a}(a-x_i)p_i.$$

上式右端第 1 项即为 $a(2F(x_{\frac{1}{2}})-1)$. 按(0.7)，有 $F(x_{\frac{1}{2}}) \geqslant \frac{1}{2}$，

因此对 $a \geqslant x_{\frac{1}{2}}$，有

$$a(2F(x_{\frac{1}{2}})-1) \geqslant x_{\frac{1}{2}}(2F(x_{\frac{1}{2}})-1), \tag{0.8}$$

且当 $a=x_{\frac{1}{2}}$ 取到等号；第 2 项与 a 无关；至于第 3 项，有

$$2\sum_{x_{\frac{1}{2}}<x_i \leqslant a}(a-x_i)p_i \geqslant 0, \tag{0.9}$$

且当 $a=x_{\frac{1}{2}}$ 时取到等号.（再将(0.8)，(0.9)联立. 若它们同时成立，a 必是中位数.）

对 $a \leqslant x_{\frac{1}{2}}$，类似有（细心的读者可能会看到，其实有很大的不同.）

$$E \mid X-a \mid = a(2\sum_{x_i<x_{\frac{1}{2}}} p_i - 1) + (\sum_{x_{\frac{1}{2}} \leqslant x_i} x_i p_i - \sum_{x_i<x_{\frac{1}{2}}} x_i p_i) + 2\sum_{a \leqslant x_i < x_{\frac{1}{2}}}(x_i-a)p_i$$

上式右端第 1 项，按(0.7)有 $\sum_{x_i<x_{\frac{1}{2}}} p_i = P(X<x_{\frac{1}{2}}) \leqslant \frac{1}{2}$，所以对

$a \leqslant x_{\frac{1}{2}}$. 有

$$a(2\sum_{x_i<x_{\frac{1}{2}}} p_i - 1) \geqslant x_{\frac{1}{2}}(2\sum_{x_i<x_{\frac{1}{2}}} p_i - 1).$$

综合以上，命题得证.

3. 设 $X_1, X_2, \cdots, X_n$ 是来自泊松分布的样本，

$$P(X_i=k)=\frac{\lambda^k}{k!}e^{-\lambda}, (\lambda>0, k=0,1,\cdots)$$

设 λ 的先验分布是参数为 α, β 的 Γ 分布，损失函数是 $(\lambda-a)^2$，试求 λ 的贝叶斯估计.

解 由对本章例 2.2 的讨论知，在平方损失下，λ 的后验分布的期望，就是 λ 的贝叶斯估计. 因此我们先来找出 λ 的后验分布密度 $\xi(\lambda \mid x_1, x_2, \cdots, x_n)$.

按本章(2.3)式

$$\xi(\lambda \mid x_1, x_2, \cdots, x_n) = \frac{\left(\prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda}\right) \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}}{\int_{\lambda>0} \left[\prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}\right] d\lambda}$$

$$= \frac{\frac{\beta^\alpha}{\prod_{i=1}^{n}(x_i!) \cdot \Gamma(\alpha)} \lambda^{\sum x_i+\alpha-1} e^{-(n+\beta)\lambda}}{\frac{\beta^\alpha}{\prod_{i=1}^{n}(x_i!) \cdot \Gamma(\alpha)} \int_0^{+\infty} \lambda^{\sum x_i+\alpha-1} e^{-(n+\beta)\lambda} d\lambda}$$

$$= \frac{(n+\beta)^{\sum x_i+\alpha}}{\Gamma(\sum x_i+\alpha)} \cdot \lambda^{\sum x_i+\alpha-1} e^{-(n+\beta)\lambda}, \lambda>0.$$

这表明 $\xi(\lambda | x_1, x_2, \cdots, x_n)$为参数$(\sum x_i+\alpha, n+\beta)$的 Γ 分布.

由于 $\Gamma(\alpha,\beta)$的期望为 α/β(参见第三章末的附表),所以后验分布的期望,亦即 λ 的贝叶斯估计

$$\hat{\lambda} = \frac{\sum x_i+\alpha}{n+\beta}.$$

4. 设 $X_1, X_2, \cdots, X_n$ 是来自伯努利分布的样本($n\geqslant 3$),

$$P(X_i=1)=\theta=1-P(X_i=0), \theta\in(0,1).$$

设 θ 的先验分布是(0,1)上的均匀分布,损失函数是

$$L(\theta,a) = \left(\frac{\theta-a}{\theta(1-\theta)}\right)^2.$$

试求 θ 的贝叶斯估计.

解 由本章例 2.2 知,θ 的后验密度 $\xi(\theta | x_1, x_2, \cdots, x_n)$在题设情况下是参数为$(\sum x_i+1, n-\sum x_i+1)$的贝塔分布. 由所给的损失函数 $L(\theta,a)$知,平均损失

$$\int_0^1 L(\theta,a)\xi(\theta \mid x_1, x_2, \cdots, x_n) d\theta$$

$$= C\int_0^1 \left(\frac{\theta-a}{\theta(1-\theta)}\right)^2 \cdot \theta^{\sum x_i+1-1}(1-\theta)^{n-\sum x_i+1-1} d\theta$$

$$= C\int_0^1 (\theta-a)^2 \cdot \theta^{\sum x_i-2}(1-\theta)^{n-\sum x_i-2}\mathrm{d}\theta.$$

a 取何值使它最小呢？由上式知，它已转化为"平方损失"的问题. 从例 2.2 中对(2.6)式的讨论知，a 应取以 $\theta^{\sum x_i-2}(1-\theta)^{n-\sum x_i-2}$ 为"核"的密度的期望（所谓密度的"核"，就是密度函数去掉与 θ 无关的系数后剩下的只跟 θ 有关的东西），显然该密度为贝塔分布，参数为($\sum x_i-1, n-\sum x_i-1$).

由于贝塔分布 $\boldsymbol{B}(\alpha,\beta)$ 的期望为 $\alpha/(\alpha+\beta)$（参见第三章末的附表），所以 θ 的贝叶斯估计

$$\hat{\theta}=\frac{\sum x_i-1}{n-2}$$

5. 设 $X_1,X_2,\cdots,X_n$ 是来自正态分布 $N(\theta,1)$ 的样本，θ 是未知参数. 给定检验问题：

$$H_1:\theta\leqslant\theta_0 \longleftrightarrow H_2:\theta>\theta_0$$

用 a_i 表示"接受假设 H_i"$(i=1,2)$. 设损失函数如下：

$$L(\theta,a_1)=\begin{cases}0 & \theta\leqslant\theta_0\\ k(\theta-\theta_0) & \theta>\theta_0\end{cases}$$

$$L(\theta,a_2)=\begin{cases}k|\theta-\theta_0| & \theta\leqslant\theta_0\\ 0 & \theta>\theta_0\end{cases}$$

其中 k 是正常数. 若 θ 的先验分布是 $N(\mu_0,\sigma_0^2)$，试求出贝叶斯检验.

解 首先，由本章定理 3.4. θ 的后验密度 $\xi(\theta|x_1,x_2,\cdots,x_n)$ 为正态分布 $N(\mu^*,\sigma^{*2})$，(参见本章(3.1)，(3.2)式)

$$\mu^*=\frac{\mu_0\sigma^2+n\sigma_0^2\cdot\bar{x}}{\sigma^2+n\sigma_0^2}=\frac{\mu_0+n\sigma_0^2\cdot\bar{x}}{1+n\sigma_0^2},$$

$$(\sigma^*)^2=\frac{\sigma^2\sigma_0^2}{\sigma^2+n\sigma_0^2}=\frac{\sigma_0^2}{1+n\sigma_0^2}. \tag{0.10}$$

再由题设的损失函数找出行动 $a_i(i=1,2)$ 的平均损失.

$$\rho(a_i)=\int_\Theta L(\theta,a_i)\xi(\theta\mid x_1,x_2,\cdots,x_n)\mathrm{d}\theta.\ (i=1,2)$$

先求

$$\rho(a_1) = \int_{\theta_0}^{+\infty} k(\theta - \theta_0) \frac{1}{\sqrt{2\pi}\sigma^*} e^{-\frac{1}{2}\left(\frac{\theta-\mu^*}{\sigma^*}\right)^2} d\theta$$

$$\xlongequal{\left(\frac{\theta-\mu^*}{\sigma^*}=t\right)} k\int_{t_0}^{+\infty} (\sigma^* t + \mu^* - \theta_0) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt.$$

其中 $t_0 = \dfrac{\theta_0 - \mu^*}{\sigma^*}$.

可算得

$$\rho(a_1) = k\sigma^* [t_0\Phi(t_0) - t_0 + \varphi(t_0)]$$

其中 $\Phi(\cdot)$,$\varphi(\cdot)$分别为标准正态分布函数与密度函数.

类似可算得

$$\rho(a_2) = \int_{-\infty}^{\theta_0} k(\theta_0 - \theta) \frac{1}{\sqrt{2\pi}\ \sigma^*} e^{-\frac{1}{2}\left(\frac{\theta-\mu^*}{\sigma^*}\right)^2} d\theta$$

$$= k\sigma^* [t_0\Phi(t_0) + \varphi(t_0)].$$

注意到

$$\rho(a_1) = \rho(a_2) - k\sigma^* t_0 = \rho(a_2) - k(\theta_0 - \mu^*),$$

因此得贝叶斯检验:

当 $k(\theta_0 - \mu^*) > 0$,即 $\mu^* < \theta_0$ 时,取 a_1,否则取 a_2.将(0.10)代入,得:当$\bar{x} < \theta_0 + \dfrac{\theta_0 - \mu_0}{n\sigma_0^2}$,接受假设 H_1;否则接受 H_2.

6. 地质学家要根据某地区的地层结构来判断该地是否蕴藏石油.地层结构总是 0,1 两种状态之一;用 θ_0 表示该地无油,θ_1 表示该地有油.已知有下列概率分布规律(其中 x 表示地层结构的状态,θ 表示石油的状态):

θ \ x	0	1
θ_0(无油)	0.6	0.4
θ_1(有油)	0.3	0.7

它表示如果该地区无油，那么地层结构呈现状态 0 的概率为 0.6，呈现状态 1 的概率为 0.4；如果该地蕴藏石油，那么地层结构呈现状态 0 的概率是 0.3，呈现状态 1 的概率为 0.7. 土地所有者希望根据地质学家对地层结构的分析决定自己投资钻探石油，还是出卖土地所有权或者在该地区开辟旅游点，分别记这些行动为 a_1，a_2，a_3. 行动空间 $A=\{a_1,a_2,a_3\}$，土地所有者权衡利弊之后取损失函数 $L(\theta,a)$为

$L(\theta,a)$ \ a / θ	a_1（自己投资钻探）	a_2（出卖所有权）	a_3（开辟旅游点）
θ_0（无油）	12	1	6
θ_1（有油）	0	7	5

试写出可供土地所有者选择的全部决策（函数）及其风险. 求出 minimax 决策. 若 θ 的先验分布是：$\xi(\theta_0)=0.2$，$\xi(\theta_1)=0.8$，试求出贝叶斯决策.

解 本题与例 1.1 完全类似（样本空间$\{0,1\}$，行动空间$\{a_1,a_2,a_3\}$，参数空间$\{\theta_0,\theta_1\}$）.

共有 9 个决策函数，$\delta_1,\delta_2,\cdots,\delta_9$ 如下表

x \ δ	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9
0	a_1	a_1	a_1	a_2	a_2	a_2	a_3	a_3	a_3
1	a_1	a_2	a_3	a_1	a_2	a_3	a_1	a_2	a_3

它们的风险，按定义

$$
\begin{aligned}
R(\theta,\delta_i)&=E[L(\theta,\delta_i(X))]\\
&=L(\theta,a_1)P(\delta_i(X)=a_1)+L(\theta,a_2)P(\delta_i(X)=a_2)\\
&\quad+L(\theta,a_3)P(\delta_i(X)=a_3).
\end{aligned}
$$

这是因为，对给定的 θ 与 δ_i，X 的取值并不确定. 为具体找出，要利用题目中的两张表.

$L(\theta,a)$ a / θ	a_1 (自己投资钻探)	a_2 (出卖所有权)	a_3 (开辟旅游点)
θ_0(无油)	12	1	6
θ_1(有油)	0	7	5

x / θ	0	1
θ_0	0.6	0.4
θ_1	0.3	0.7

$$\begin{aligned} R(\theta_0,\delta_i) &= L(\theta_0,a_1)P(\delta_i(X)=a_1)+L(\theta_0,a_2)P(\delta_i(X)=a_2)+ \\ &\quad L(\theta_0,a_3)P(\delta_i(X)=a_3) \\ &= 12\times P(\delta_i(X)=a_1)+1\times P(\delta_i(X)=a_2)+ \\ &\quad 6\times P(\delta_i(X)=a_3). \end{aligned}$$

注意:这里的 $P(\delta_i(X)=a_j)$ 确切地讲是条件概率,即 $P(\delta_i(X)=a_j|\theta_0)$. 它们的值,由上面的决策函数表及右边小表容易找出. 从而求出 $R(\theta_0,\delta_i)$(见下表).

δ	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9
$R(\theta_0,\delta)$	12	7.6	9.6	5.4	1	3	8.4	4	6
$R(\theta_1,\delta)$	0	4.9	3.5	2.1	7	5.6	1.5	6.4	5

上表中的 $R(\theta_1,\delta)$ 是由下式求得:

$$\begin{aligned} R(\theta_1,\delta_i) &= 0\times P(\delta_i(X)=a_1|\theta_1)+7\times P(\delta_i(X)=a_2|\theta_1) \\ &\quad +5\times P(\delta_i(X)=a_3|\theta_1). \end{aligned}$$

从上表看出 δ_4 为 minimax 决策,即

$X=0$ 时取行动 a_2(出卖所有权);

$X=1$ 时取行动 a_1(自己投资钻探).

下面来求先验分布 $\xi(\theta_0)=0.2(\xi(\theta_1)=0.8)$ 时的贝叶斯决策.

先找后验分布

$$\begin{aligned} \xi(\theta_0|X=0) &= \frac{P(X=0|\theta_0)\xi(\theta_0)}{P(X=0)} \\ &= \frac{P(X=0|\theta_0)\xi(\theta_0)}{P(X=0|\theta_0)\xi(\theta_0)+P(X=0|\theta_1)\xi(\theta_1)} \end{aligned}$$

$$=\frac{0.6\times0.2}{0.6\times0.2+0.3\times0.8}=\frac{1}{3},$$

$$\xi(\theta_1\mid X=0)=\frac{2}{3};$$

类似，$\xi(\theta_0\mid X=1)=\dfrac{0.4\times0.2}{0.4\times0.2+0.7\times0.8}=\dfrac{1}{8}$,

$$\xi(\theta_1\mid X=1)=\frac{7}{8}.$$

为找对给定的 X(0 或 1)，取何行动(a_1,a_2,a_3)使平均损失最小.我们分别计算(对 $i=0,1,j=1,2,3$)

$$\int_\Theta L(\theta,a_j)\xi(\theta\mid X=i)\mathrm{d}\theta.$$

对离散型即计算

$$L(\theta_0,a_j)\cdot\xi(\theta_0\mid X=i)+L(\theta_1,a_j)\xi(\theta_1\mid X=i).$$

利用损失函数表与上面算出的后验分布 $\xi(\theta\mid X)$. 可得下表(平均损失表)：

	a_1	a_2	a_3
$X=0$	$\frac{1}{3}\times12+\frac{2}{3}\times0=4$	$\frac{1}{3}\times1+\frac{2}{3}\times7=5$	$\frac{1}{3}\times6+\frac{2}{3}\times5=5.33$
$X=1$	$\frac{1}{8}\times12+\frac{7}{8}\times0=1.5$	$\frac{1}{8}\times1+\frac{7}{8}\times7=6.25$	$\frac{1}{8}\times6+\frac{7}{8}\times5=5.13$

由上表知，贝叶斯决策为：

$X=0$ 时取行动 a_1，

$X=1$ 时取行动 a_1.

郑重声明

策划编辑　徐　可
责任编辑　田　军
封面设计　王凌波
责任绘图　朱　静
版式设计　王　莹
责任校对　杨雪莲
责任印制　宋克学

［Ｇｅｎｅｒａｌ Ｉｎｆｏｒｍａｔｉｏｎ］
书名＝概率统计讲义习题解答
作者＝汪仁官　　刘婉如　　冯泰编
页数＝１０６
ＳＳ号＝１１４００９１４
出版日期＝２００４年０８月第１版