

第一章 事件与概率

1.1 写出下列随机试验的样本空间及表示下列事件的样本点集合。

(1) 10 件产品中有 1 件是不合格品，从中任取 2 件得 1 件不合格品。

(2) 一个口袋中有 2 个白球、3 个黑球、4 个红球，从中任取一球，(i) 得白球，(ii) 得红球。

解 (1) 记 9 个合格品分别为 $\omega_1, \omega_2, \dots, \omega_9$ ，记不合格为次，则

$$\Omega = \{(\omega_1, \omega_2), (\omega_1, \omega_3), \dots, (\omega_1, \omega_9), (\omega_1, \text{次}), (\omega_2, \omega_3), (\omega_2, \omega_4), \dots, (\omega_2, \omega_9), (\omega_2, \text{次}),$$

$$(\omega_3, \omega_4), \dots, (\omega_3, \omega_9), (\omega_3, \text{次}), \dots, (\omega_9, \omega_{\text{次}}), (\omega_9, \text{次})\}$$

$$A = \{(\omega_1, \text{次}), (\omega_2, \text{次}), \dots, (\omega_9, \text{次})\}$$

(2) 记 2 个白球分别为 ω_1, ω_2 ，3 个黑球分别为 b_1, b_2, b_3 ，4 个红球分别为 r_1, r_2, r_3, r_4 。则 $\Omega = \{\omega_1, \omega_2, b_1, b_2, b_3, r_1, r_2, r_3, r_4\}$

$$(i) A = \{\omega_1, \omega_2\} \quad (ii) B = \{r_1, r_2, r_3, r_4\}$$

1.2 在数学系的学生中任选一名学生，令事件 A 表示被选学生是男生，事件 B 表示被选学生是三年级学生，事件 C 表示该生是运动员。

(1) 叙述 ABC 的意义。

(2) 在什么条件下 $ABC = C$ 成立？

(3) 什么时候关系式 $C \subset B$ 是正确的？

(4) 什么时候 $\bar{A} = B$ 成立？

解 (1) 事件 ABC 表示该是三年级男生，但不是运动员。

(2) $ABC = C$ 等价于 $C \subset AB$ ，表示全系运动员都有是三年级的男生。

(3) 当全系运动员都是三年级学生时。

(4) 当全系女生都在三年级并且三年级学生都是女生时。

1.3 一个工人生产了 n 个零件，以事件 A_i 表示他生产的第 i 个零件是合格品

($1 \leq i \leq n$)。用 A_i 表示下列事件：

(1) 没有一个零件是不合格品；

(2) 至少有一个零件是不合格品；

(3) 仅仅只有一个零件是不合格品；

(4) 至少有两个零件是不合格品。

$$\text{解 (1) } \bigcap_{i=1}^n A_i; \quad (2) \overline{\bigcap_{i=1}^n A_i} = \bigcup_{i=1}^n \bar{A}_i; \quad (3) \bigcup_{i=1}^n [\bar{A}_i (\bigcap_{\substack{j=1 \\ j \neq i}}^n A_j)];$$

(4) 原事件即“至少有两个零件是合格品”，可表示为 $\bigcup_{\substack{i,j=1 \\ i \neq j}}^n A_i A_j$ ；

1.4 证明下列各式：

(1) $A \cup B = B \cup A$;

(2) $A \cap B = B \cap A$

(3) $(A \cup B) \cup C = A \cup (B \cup C)$;

(4) $(A \cap B) \cap C = A \cap (B \cap C)$

(5) $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$

(6) $\overline{\bigcap_{i=1}^n A_i} = \bigcup_{i=1}^n \overline{A_i}$

证明 (1) — (4) 显然，(5) 和 (6) 的证法分别类似于课文第 10—12 页 (1.5) 式和 (1.6) 式的证法。

1.5 在分别写有 2、4、6、7、8、11、12、13 的八张卡片中任取两张，把卡片上的两个数字组成一个分数，求所得分数为既约分数的概率。

解 样本点总数为 $A_8^2 = 8 \times 7$ 。所得分数为既约分数必须分子分母或为 7、11、13 中的两个，或为 2、4、6、8、12 中的一个和 7、11、13 中的一个组合，所以事件 A “所得分数为既约分数” 包含 $A_3^2 + 2A_3^1 \times A_5^1 = 2 \times 3 \times 6$ 个样本点。于是

$$P(A) = \frac{2 \times 3 \times 6}{8 \times 7} = \frac{9}{14}.$$

1.6 有五条线段，长度分别为 1、3、5、7、9。从这五条线段中任取三条，求所取三条线段能构成一个三角形的概率。

解 样本点总数为 $\binom{5}{3} = 10$ 。所取三条线段能构成一个三角形，这三条线段必须是 3、5、7 或 3、7、9 或多或 5、7、9。所以事件 A “所取三条线段能构成一个三角形” 包含 3 个样本点，于是 $P(A) = \frac{3}{10}$ 。

1.7 一个小孩用 13 个字母 $A, A, A, C, E, H, I, I, M, M, N, T, T$ 作组字游戏。如果字母的各种排列是随机的（等可能的），问“恰好组成“MATHEMATICIAN”一词的概率为多大？

解 显然样本点总数为 $13!$ ，事件 A “恰好组成“MATHEMATICIAN”包含 $3!2!2!2!$ 个样本点。所以 $P(A) = \frac{3!2!2!2!}{13!} = \frac{48}{13!}$

1.8 在中国象棋的棋盘上任意地放上一只红“车”及一只黑“车”，求它们正好可以相互吃掉的概率。

解 任意固定红“车”的位置，黑“车”可处于 $9 \times 10 - 1 = 89$ 个不同位置，当

它处于和红“车”同行或同列的 $9+8=17$ 个位置之一时正好相互“吃掉”。故所求概率为

$$P(A) = \frac{17}{89}$$

1.9 一幢 10 层楼的楼房中的一架电梯，在底层登上 7 位乘客。电梯在每一层都停，乘客从第二层起离开电梯，假设每位乘客在哪一层离开电梯是等可能的，求没有两位及两位以上乘客在同一层离开的概率。

解 每位乘客可在除底层外的 9 层中任意一层离开电梯，现有 7 位乘客，所以样本点总数为 9^7 。事件 A “没有两位及两位以上乘客在同一层离开”相当于“从 9 层中任取 7 层，各有一位乘客离开电梯”。所以包含 A_9^7 个样本点，于是

$$P(A) = \frac{A_9^7}{9^7}.$$

1.10 某城市共有 10000 辆自行车，其牌照编号从 00001 到 10000。问事件“偶然遇到一辆自行车，其牌照号码中有数字 8”的概率为多大？

解 用 A 表示“牌照号码中有数字 8”，显然 $P(\bar{A}) = \frac{9^4}{10000} = \left(\frac{9}{10}\right)^4$ ，所以

$$P(A) = 1 - P(\bar{A}) = 1 - \frac{9^4}{10000} = 1 - \left(\frac{9}{10}\right)^4$$

1.11 任取一个正数，求下列事件的概率：

- (1) 该数的平方的末位数字是 1；
- (2) 该数的四次方的末位数字是 1；
- (3) 该数的立方的最后两位数字都是 1；

解 (1) 答案为 $\frac{1}{5}$ 。

(2) 当该数的末位数是 1、3、7、9 之一时，其四次方的末位数是 1，所以答案为 $\frac{4}{10} = \frac{2}{5}$

(3) 一个正整数的立方的最后两位数字决定于该数的最后两位数字，所以样本空间包含 10^2 个样本点。用事件 A 表示“该数的立方的最后两位数字都是 1”，则该数的最后一位数字必须是 1，设最后第二位数字为 a ，则该数的立方的最后两位数字为 1 和 $3a$ 的个位数，要使 $3a$ 的个位数是 1，必须 $a=7$ ，因此 A 所包含的样本点只有 71 这一点，于是

1.12 一个人把 6 根草掌握在手中，仅露出它们的头和尾。然后请另一个人把 6 个头两两相接，6 个尾也两两相接。求放开手以后 6 根草恰好连成一个环的概率。并把上述结果推广到 $2n$ 根草的情形。

解 (1) 6 根草的情形。取定一个头，它可以与其它的 5 个头之一相接，再取另一头，它又可以与其它未接过的 3 个之一相接，最后将剩下的两个头相接，故

对头而言有 $5 \cdot 3 \cdot 1$ 种接法，同样对尾也有 $5 \cdot 3 \cdot 1$ 种接法，所以样本点总数为 $(5 \cdot 3 \cdot 1)^2$ 。用 A 表示“6 根草恰好连成一个环”，这种连接，对头而言仍有 $5 \cdot 3 \cdot 1$ 种连接法，而对尾而言，任取一尾，它只能和未与它的头连接的另 4 根草的尾连接。再取另一尾，它只能和未与它的头连接的另 2 根草的尾连接，最后再将其余的尾连接成环，故尾的连接法为 $4 \cdot 2$ 。所以 A 包含的样本点数为 $(5 \cdot 3 \cdot 1)(4 \cdot 2)$ ，

$$\text{于是 } P(A) = \frac{(5 \cdot 3 \cdot 1)(4 \cdot 2)}{(5 \cdot 3 \cdot 1)^2} = \frac{8}{15}$$

(2) $2n$ 根草的情形和 (1) 类似得

1.13 把 n 个完全相同的球随机地放入 N 个盒子中（即球放入盒子后，只能区别盒子中球的个数，不能区别是哪个球进入某个盒子，这时也称球是不可辨的）。如果每一种放法都是等可能的，证明 (1) 某一个指定的盒子中恰好有 k 个球的

$$\text{的概率为 } \frac{\binom{N+n-k-2}{n-k}}{\binom{N+n-1}{n}}, \quad 0 \leq k \leq n$$

$$(2) \text{ 恰好有 } m \text{ 个盒的概率为 } \frac{\binom{N}{m} \binom{n-1}{N-m-1}}{\binom{N+n-1}{n}}, \quad N-n \leq m \leq N-1$$

$$(3) \text{ 指定的 } m \text{ 个盒中正好有 } j \text{ 个球的概率为 } \frac{\binom{m+j-1}{m-1} \binom{N-m+n-j-1}{n-j}}{\binom{N+n-1}{n}},$$

$$1 \leq m \leq N, 0 \leq j \leq N.$$

解 略。

1.14 某公共汽车站每隔 5 分钟有一辆汽车到达，乘客到达汽车站的时刻是任意的，求一个乘客候车时间不超过 3 分钟的概率。

$$\text{解 所求概率为 } P(A) = \frac{3}{5}$$

1.15 在 $\triangle ABC$ 中任取一点 P ，证明 $\triangle ABP$ 与 $\triangle ABC$ 的面积之比大于 $\frac{n-1}{n}$ 的概率为 $\frac{1}{n^2}$ 。

解 截取 $CD' = \frac{1}{n}CD$ ，当且仅当点 P 落入 $\triangle CA'B'$ 之内时 $\triangle ABP$ 与 $\triangle ABC$ 的面积之比大于 $\frac{n-1}{n}$ ，因此所求概率为

$$P(A) = \frac{\triangle A'B'C \text{ 有面积}}{\triangle ABC \text{ 的面积}} = \frac{\overline{CD'}^2}{\overline{CD}^2} = \frac{\frac{1}{n^2} \overline{CD}^2}{\overline{CD}^2} = \frac{1}{n^2}。$$

1.16 两艘轮船都要停靠同一个泊位，它们可能在一昼夜的任意时刻到达。设两船停靠泊位的时间分别为 1 小时与两小时，求有一艘船停靠泊位时必须等待一段时间的概率。

解 分别用 x, y 表示第一、二艘船到达泊位的时间。一艘船到达泊位时必须等待当且仅当 $0 \leq x - y \leq 2, 0 \leq y - x \leq 1$ 。因此所求概率为

$$P(A) = \frac{24^2 - \frac{1}{2} \times 23^2 - \frac{1}{2} \times 22^2}{24^2} \approx 0.121$$

1.17 在线段 AB 上任取三点 x_1, x_2, x_3 ，求：

(1) x_2 位于 x_1 与 x_3 之间的概率。

(2) Ax_1, Ax_2, Ax_3 能构成一个三角形的概率。

$$\text{解 (1) } P(A) = \frac{1}{3} \quad (2) \quad P(B) = \frac{1 - 3 \times \frac{1}{3} \times \frac{1}{2}}{1} = \frac{1}{2}$$

1.18 在平面上画有间隔为 d 的等距平行线，向平面任意地投掷一个三角形，该三角形的边长为 a, b, c （均小于 d ），求三角形与平行线相交的概率。

解 分别用 A_1, A_2, A_3 表示三角形的一个顶点与平行线相合，一条边与平行线相合，两条边与平行线相交，显然 $P(A_1) = P(A_2) = 0$ 。所求概率为 $P(A_3)$ 。分别用 $A_a, A_b, A_c, A_{ab}, A_{ac}, A_{bc}$ 表示边 a, b, c ，二边 ab, ac, bc 与平行线相交，则 $P(A_3) = P(A_{ab} \cup A_{ac} \cup A_{bc})$ 。显然 $P(A_a) = P(A_{ab}) + P(A_{ac})$ ， $P(A_b) = P(A_{ab}) + P(A_{bc})$ ， $P(A_c) = P(A_{ac}) + P(A_{bc})$ 。所以

$$P(A_3) = \frac{1}{2} [P(A_a) + P(A_b) + P(A_c)] = \frac{2}{2\pi d} (a + b + c) = \frac{1}{\pi d} (a + b + c)$$

（用例 1.12 的结果）

1.19 已知不可能事件的概率为零，现在问概率为零的事件是否一定为不可能事件？试举例说明之。

解 概率为零的事件不一定是不可可能事件。例如向长度为 1 的线段内随机投点。则事件 A “该点命中 AB 的中点” 的概率等于零，但 A 不是不可能事件。

1.20 甲、乙两人从装有 a 个白球与 b 个黑球的口袋中轮流摸取一球，甲先取，乙后取，每次取后都有不放回，直到两人中有一人取到白球时停止。试描述这一随机现象的概率空间，并求甲或乙先取到白球的概率。

解 ω_1 表示白, ω_2 表示黑白, ω_3 表示黑黑白, $\cdots \omega_{b+1}$ 表示 $\overbrace{\text{黑}\cdots\text{黑}}^{b\text{个}}\text{白}$,

则样本空间 $\Omega = \{\omega_1, \omega_2, \cdots, \omega_{b+1}\}$, 并且 $P(\{\omega_1\}) = \frac{a}{a+b}$,

$$P(\{\omega_2\}) = \frac{b}{a+b} \cdot \frac{a}{a+b-1}, \quad P(\{\omega_3\}) = \frac{b}{a+b} \cdot \frac{b-1}{a+b-1} \cdot \frac{a}{a+b-2}, \quad \cdots,$$

$$P(\{\omega_i\}) = \frac{b}{a+b} \cdot \frac{b-1}{a+b-1} \cdots \frac{b-(i-2)}{a+b-(i-2)} \cdot \frac{a}{a+b-(i-1)}$$

$$P(\{\omega_{b+1}\}) = \frac{b!a}{(a+b)(a+b-1)\cdots a}$$

甲取胜的概率为 $P(\{\omega_1\}) + P(\{\omega_3\}) + P(\{\omega_5\}) + \cdots$

乙取胜的概率为 $P(\{\omega_2\}) + P(\{\omega_4\}) + P(\{\omega_6\}) + \cdots$

1.21 设事件 A, B 及 $A \cup B$ 的概率分别为 p, q 及 r , 求 $P(AB)$, $P(\overline{AB})$, $P(\overline{A}B)$, $P(A\overline{B})$

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

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

$$P(\overline{AB}) = P(A - AB) = P(A) - P(AB) = r - q, \quad P(\overline{A}B) = r - p$$

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

1.22 设 A_1, A_2 为两个随机事件, 证明:

$$(1) \quad P(A_1 A_2) = 1 - P(\overline{A_1}) - P(\overline{A_2}) + P(\overline{A_1} \overline{A_2});$$

$$(2) \quad 1 - P(\overline{A_1}) - P(\overline{A_2}) \leq P(A_1 A_2) \leq P(A_1 \cup A_2) \leq P(A_1) + P(A_2).$$

证 明 (1)

$$P(A_1 A_2) = P(\overline{\overline{A_1} \cup \overline{A_2}}) = 1 - P(\overline{A_1} \cup \overline{A_2}) = 1 - P(\overline{A_1}) - P(\overline{A_2}) + P(\overline{A_1} \overline{A_2})$$

(2) 由(1)和 $P(\overline{A_1} \overline{A_2}) \geq 0$ 得第一个不等式, 由概率的单调性和半可加性分别得第二、三个不等式。

1.23 对于任意的随机事件 A, B, C , 证明:
 $P(AB) + P(AC) - P(BC) \leq P(A)$

$$\text{证明 } P(A) \geq P[A(B \cup C)] = P(AB) + P(AC) - P(ABC)$$

$$\geq P(AB) + P(AC) - P(BC)$$

1.24 在某城市中共发行三种报纸：甲、乙、丙。在这个城市的居民中，订甲报的有 45%，订乙报的有 35%，订丙报的有 30%，同时订甲、乙两报的有 10%，同时订甲、丙两报的有 8%，同时订乙、丙两报的有 5%，同时订三种报纸的有 3%，求下述百分比：

- (1) 只订甲报的；
- (2) 只订甲、乙两报的；
- (3) 只订一种报纸的；
- (4) 正好订两种报纸的；
- (5) 至少订一种报纸的；
- (6) 不订任何报纸的。

解 事件 A 表示订甲报，事件 B 表示订乙报，事件 C 表示订丙报。

$$(1) \quad P(\overline{ABC}) = P(A - (AB \cup AC)) = P(A) - P(AB \cup AC) = 30\%$$

$$(2) \quad P(\overline{AB}C) = P(AB - ABC) = 7\%$$

$$(3) \quad P(\overline{B}AC) = P(B) - [P(AB) + P(BC) - P(ABC)] = 23\%$$

$$P(\overline{C}AB) = P(C) - [P(AC) + P(BC) - P(ABC)] = 20\%$$

$$P(\overline{ABC} \cup \overline{BAC} + \overline{CAB}) = P(\overline{ABC}) + P(\overline{BAC}) + P(\overline{CAB}) = 73\%$$

$$(4) \quad P(\overline{ABC} + \overline{ACB} + \overline{BCA}) = P(\overline{ABC}) + P(\overline{ACB}) + P(\overline{BCA}) = 14\%$$

$$(5) \quad P(A + B + C) = 90\%$$

$$(6) \quad P(\overline{ABC}) = 1 - P(A + B + C) = 1 - 90\% = 10\%$$

1.26 某班有 n 个学生参加口试，考签共 N 张，每人抽到的考签用后即放回，在考试结束后，问至少有一张考没有被抽到的概率是多少？

解 用 A_i 表示“第 i 张考签没有被抽到”， $i = 1, 2, \dots, N$ 。要求 $P(\bigcup_{i=1}^N A_i)$ 。

$$P(A_i) = \left(\frac{N-1}{N}\right)^n, \quad P(A_i A_j) = \left(\frac{N-2}{N}\right)^n, \quad \dots, \quad P(A_1 \dots A_N) = \left(\frac{N-N}{N}\right)^n = 0$$

$$\sum_{i=1}^N P(A_i) = \binom{N}{1} \cdot \left(\frac{N-1}{N}\right)^n = (-1)^{1-1} \binom{N}{1} \left(\frac{N-1}{N}\right)^n$$

$$- \sum_{1 \leq i \leq j \leq N} P(A_i A_j) = - \binom{N}{2} \left(\frac{N-2}{N}\right)^n = (-1)^{2-1} \binom{N}{2} \left(\frac{N-2}{N}\right)^n, \quad \dots$$

$$\text{所以 } P(\bigcup_{i=1}^N A_i) = \sum_{i=1}^N (-1)^{i-1} \left(\frac{N-i}{N}\right)^n$$

1.27 从 n 阶行列式的一般展开式中任取一项，问这项包含主对角线元素的概率是多少？

解 n 阶行列式的展开式中，任一项略去符号不计都可表示为 $a_{1i_1}a_{2i_2}\cdots a_{ni_n}$ ，当且仅当 $1, 2, \dots, n$ 的排列 $(i_1 i_2 \cdots i_n)$ 中存在 k 使 $i_k = k$ 时这一项包含主对角线元素。用 A_k 表示事件“排列中 $i_k = k$ ”即第 k 个主对角线元素出现于展开式的某项中。则

$$P(A_i) = \frac{(n-1)!}{n!} \quad 1 \leq i \leq n \quad P(A_i A_j) = \frac{(n-2)!}{n!} \quad (1 \leq i < j \leq n), \dots\dots$$

$$\text{所以 } P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n (-1)^{i-1} \binom{n}{i} \frac{(n-i)!}{n!} = \sum_{i=1}^n (-1)^{i-1} \frac{1}{i!}$$

1.29 已知一个家庭中有三个小孩，且其中一个是女孩，求至少有一个男孩的概率（假设一个小孩是男孩或是女孩是等可能的）。

解 用 b, g 分别表示男孩和女孩。则样本空间为：

$$\Omega = \{(b, b, b), (b, b, g), (b, g, b), (g, b, b), (b, g, g), (g, b, g), (g, g, b), (g, g, g)\}$$

其中样本点依年龄大小的性别排列。 A 表示“有女孩”， B 表示“有男孩”，则

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{6/8}{7/8} = \frac{6}{7}$$

1.30 设 M 件产品中有 m 件是不合格品，从中任取两件，

(1) 在所取产品中有一件是不合格品的条件下，求另一件也是不合格品的概率。

(2) 在所取产品中有一件是合格品的条件下，求另一件也是不合格品的概率。

解 (1) 设 A 表示“所取产品中至少有一件是不合格品”， B 表示“所取产品都是不合格品”，则

$$P(A) = \frac{\binom{m}{2} + \binom{m}{1} \binom{M-m}{1}}{\binom{M}{2}} \quad P(B) = \frac{\binom{m}{2}}{\binom{M}{2}}$$

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{P(B)}{P(A)} = \frac{m-1}{2M-m-1}$$

(2) 设 C 表示“所取产品中至少有一件合格品”， D 表示“所取产品中有一件合格品，一件不合格品”。则

$$P(C) = \frac{\binom{m}{1} \binom{M-m}{1} + \binom{M-m}{2}}{\binom{M}{2}} \quad P(D) = \frac{\binom{m}{1} \binom{M-m}{1}}{\binom{M}{2}}$$

$$P(D|C) = \frac{P(CD)}{P(C)} = \frac{P(D)}{P(C)} = \frac{2m}{M+m-1}$$

1.31 n 个人用摸彩的方式决定谁得一张电影票，他们依次摸彩，求：

(1) 已知前 $k-1$ ($k \leq n$) 个人都没摸到，求第 k 个人摸到的概率；

(2) 第 k ($k \leq n$) 个人摸到的概率。

解 设 A_i 表示“第 i 个人摸到”， $i=1,2,\dots,n$ 。

$$(1) P(A_k | \overline{A_1} \cdots \overline{A_{k-1}}) = \frac{1}{n-(k-1)} = \frac{1}{n-k+1}$$

$$(2) P(A_k) = P(\overline{A_1} \cdots \overline{A_{k-1}} A_k) = \frac{n-1}{n} \cdot \frac{n-2}{n-1} \cdots \frac{1}{n-k+1} = \frac{1}{n}$$

1.32 已知一个母鸡生 k 个蛋的概率为 $\frac{\lambda^k}{k!} e^{-\lambda}$ ($\lambda > 0$)，而每一个蛋能孵化成小鸡的概率为 p ，证明：一个母鸡恰有 r 个下一代（即小鸡）的概率为 $\frac{(\lambda p)^r}{r!} e^{-\lambda p}$ 。

解 用 A_k 表示“母鸡生 k 个蛋”， B 表示“母鸡恰有 r 个下一代”，则

$$\begin{aligned} P(B) &= \sum_{k=r}^{\infty} P(A_k) P(B|A_k) = \sum_{k=r}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \cdot \binom{k}{r} \cdot p^r (1-p)^{k-r} \\ &= \frac{(\lambda p)^r}{r!} e^{-\lambda} \sum_{k=r}^{\infty} \frac{[\lambda(1-p)]^{k-r}}{(k-r)!} = \frac{(\lambda p)^r}{r!} e^{-\lambda} \cdot e^{\lambda(1-p)} \\ &= \frac{(\lambda p)^r}{r!} e^{-\lambda p} \end{aligned}$$

1.33 某射击小组共有 20 名射手，其中一级射手 4 人，二级射手 8 人，三级射手 7 人，四级射手一人，一、二、三、四级射手能通过选拔进入决赛的概率分别是 0.9、0.7、0.5、0.2，求在一组内任选一名射手，该射手能通过选拔进入决赛的概率。

解 用 A_k 表示“任选一名射手为 k 级”， $k=1,2,3,4$ ， B 表示“任选一名射手能进入决赛”，则

$$P(B) = \sum_{k=1}^4 P(A_k) P(B|A_k) = \frac{4}{20} \times 0.9 + \frac{8}{20} \times 0.7 + \frac{7}{20} \times 0.5 + \frac{1}{20} \times 0.2 = 0.645$$

1.34 在某工厂里有甲、乙、丙三台机器生产螺丝钉，它们的产量各占 25%，35%，40%，并在各自的产品里，不合格品各占有 5%，4%，2%。现在从产品中任

取一只恰是不合格品，问此不合格品是机器甲、乙、丙生产的概率分别等于多少？

解 用 A_1 表示“任取一只产品是甲台机器生产”

A_2 表示“任取一只产品是乙台机器生产”

A_3 表示“任取一只产品是丙台机器生产”

B 表示“任取一只产品恰是不合格品”。

则由贝叶斯公式：

$$P(A_1|B) = \frac{P(A_1)P(B|A_1)}{\sum_{k=1}^3 P(A_k)P(B|A_k)} = \frac{25}{69} \quad P(A_2|B) = \frac{P(A_2)P(B|A_2)}{\sum_{k=1}^3 P(A_k)P(B|A_k)} = \frac{28}{69}$$

$$P(A_3|B) = \frac{P(A_3)P(B|A_3)}{\sum_{k=1}^3 P(A_k)P(B|A_k)} = \frac{16}{69}$$

1.35 某工厂的车床、钻床、磨床、刨床的台数之比为 9:3:2:1，它们在一定时间内需要修理的概率之比为 1:2:3:1。当有一台机床需要修理时，问这台机床是车床的概率是多少？

$$\text{解 则 } P(A_1) = \frac{9}{15}, \quad P(A_2) = \frac{3}{15}, \quad P(A_3) = \frac{2}{15}, \quad P(A_4) = \frac{1}{15}$$

$$P(B|A_1) = \frac{1}{7}, \quad P(B|A_2) = \frac{2}{7}, \quad P(B|A_3) = \frac{3}{7}, \quad P(B|A_4) = \frac{1}{7}$$

$$\text{由贝叶斯公式得 } P(A_1|B) = \frac{P(A_1)P(B|A_1)}{\sum_{k=1}^4 P(A_k)P(B|A_k)} = \frac{9}{22}$$

1.36 有朋友自远方来访，他乘火车、轮船、汽车、飞机来的概率分别是 0.3、0.2、0.1、0.4。如果他乘火车、轮船、汽车来的话，迟到的概率分别是 $\frac{1}{4}$ 、 $\frac{1}{3}$ 、 $\frac{1}{12}$ ，而乘飞机不会迟到。结果他迟到了，试问他是乘火车来的概率是多少？

解 用 A_1 表示“朋友乘火车来”， A_2 表示“朋友乘轮船来”， A_3 表示“朋友乘汽车来”， A_4 表示“朋友乘飞机来”， B 表示“朋友迟到了”。

$$\text{则 } P(A_1|B) = \frac{P(A_1)P(B|A_1)}{\sum_{k=1}^4 P(A_k)P(B|A_k)} = \frac{1}{2}$$

1.37 证明：若三个事件 A 、 B 、 C 独立，则 $A \cup B$ 、 AB 及 $A - B$ 都与 C 独立。

$$\text{证明 (1) } P((A \cup B)C) = P(AC) + P(BC) - P(ABC)$$

$$= P(A \cup B)P(C)$$

$$(2) P(ABC) = P(A)P(B)P(C) = P(AB)P(C)$$

$$(3) P((A-B)C) = P((A-AB)C) = P(AC-ABC) = P(A-B)P(C)$$

1.38 试举例说明由 $P(ABC) = P(A)P(B)P(C)$ 不能推出 $P(AB) = P(A)P(B)$ 一定成立。

$$\text{解 设 } \Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}, \quad P(\{\omega_1\}) = \frac{1}{64}, \quad P(\{\omega_5\}) = \frac{18}{64},$$

$$P(\{\omega_2\}) = P(\{\omega_3\}) = P(\{\omega_4\}) = \frac{15}{64}, \quad A = \{\omega_1, \omega_2\}, \quad B = \{\omega_1, \omega_3\},$$

$$C = \{\omega_1, \omega_4\} \text{ 则 } P(A) = P(B) = P(C) = \frac{1}{64} + \frac{15}{64} = \frac{1}{4},$$

$$P(ABC) = P(\{\omega_1\}) = \frac{1}{64} = P(A)P(B)P(C)$$

$$\text{但是 } P(AB) = P(\{\omega_1\}) = \frac{1}{64} \neq P(A)P(B)$$

1.39 设 A_1, A_2, \dots, A_n 为 n 个相互独立的事件, 且 $P(A_k) = p_k (1 \leq k \leq n)$, 求下列事件的概率:

- (1) n 个事件全不发生;
- (2) n 个事件中至少发生一件;
- (3) n 个事件中恰好发生一件。

$$\text{解 (1) } P\left(\bigcap_{k=1}^n \bar{A}_k\right) = \prod_{k=1}^n P(\bar{A}_k) = \prod_{k=1}^n (1 - p_k)$$

$$(2) P\left(\bigcup_{k=1}^n A_k\right) = 1 - P\left(\bigcap_{k=1}^n \bar{A}_k\right) = 1 - \prod_{k=1}^n (1 - p_k)$$

$$(3) P\left(\bigcup_{k=1}^n (A_k \bigcap_{\substack{j=1 \\ j \neq k}}^n \bar{A}_j)\right) = \sum_{k=1}^n P\left(A_k \bigcap_{\substack{j=1 \\ j \neq k}}^n \bar{A}_j\right) = \sum_{k=1}^n [p_k \prod_{\substack{j=1 \\ j \neq k}}^n (1 - p_j)].$$

1.40 已知事件 A, B 相互独立且互不相容, 求 $\min(P(A), P(B))$ (注: $\min(x, y)$ 表示 x, y 中小的一个数)。

解 一方面 $P(A), P(B) \geq 0$, 另一方面 $P(A)P(B) = P(AB) = 0$, 即 $P(A), P(B)$ 中至少有一个等于 0, 所以 $\min(P(A), P(B)) = 0$.

1.41 一个人的血型为 O, A, B, AB 型的概率分别为 0.46、0.40、0.11、0.03, 现在任意挑选五个人, 求下列事件的概率

- (1) 两个人为 O 型, 其它三个人分别为其它三种血型;
- (2) 三个人为 O 型, 两个人为 A 型;
- (3) 没有一人为 AB 。

解 (1) 从 5 个人任选 2 人为 O 型, 共有 $\binom{5}{2}$ 种可能, 在其余 3 人中任选一人
为 A 型, 共有三种可能, 在余下的 2 人中任选一人 B 型, 共有 2 种可能, 另一
人为 AB 型, 顺此所求概率为: $\binom{5}{2} \times 3 \times 2 \times 0.46^2 \times 0.40 \times 0.11 \times 0.13 \approx 0.0168$

$$(2) \binom{5}{3} \times 0.46^2 \times 0.40^2 \approx 0.1557$$

$$(3) (1 - 0.03)^5 \approx 0.8587$$

1.42 设有两门高射炮, 每一门击中目标的概率都是 0.6, 求同时发射一发炮弹而击中飞机的概率是多少? 又若有一架敌机入侵领空, 欲以 99% 以上的概率击中它, 问至少需要多少门高射炮。

解 用 A_k 表示 “第 k 门高射炮发射一发炮弹而击中飞机”, $k=1, 2, \dots$, B 表示 “击中飞机”。则 $P(A_k) = 0.6$, $k=1, 2, \dots$ 。

$$(1) P(A_1 \cup A_2) = 1 - P(\overline{A_1 A_2}) = 1 - 0.4^2 = 0.84$$

$$(2) P(A_1 \cup \dots \cup A_n) = 1 - P\left(\bigcap_{k=1}^n \overline{A_k}\right) = 1 - 0.4^n > 0.99, \quad n > \frac{\lg 0.01}{\lg 0.4} \approx 5.026$$

取 $n=6$ 。至少需要 6 门高射炮, 同时发射一发炮弹, 可保证 99% 的概率击中飞机。

1.43 做一系列独立的试验, 每次试验中成功的概率为 p , 求在成功 n 次之前已失败了 m 次的概率。

解 用 A 表示 “在成功 n 次之前已失败了 m 次”, B 表示 “在前 $n+m-1$ 次试验中失败了 m 次”, C 表示 “第 $n+m$ 次试验成功”

$$\begin{aligned} \text{则 } P(A) &= P(BC) = P(B)P(C) = \binom{n+m-1}{m} p^{n-1} (1-p)^m \cdot p \\ &= \binom{n+m-1}{m} p^n (1-p)^m \end{aligned}$$

1.45 某数学家有两盒火柴, 每盒都有 n 根火柴, 每次用火柴时他在两盒中任取一盒并从中抽出一根。求他用完一盒时另一盒中还有 r 根火柴 ($1 \leq r \leq n$) 的概率。

解 用 A_i 表示 “甲盒中尚余 i 根火柴”, 用 B_j 表示 “乙盒中尚余 j 根火柴”,
 C, D 分别表示 “第 $2n-r$ 次在甲盒取”, “第 $2n-r$ 次在乙盒取”, $A_0 B_r C$ 表示取

了 $2n-r$ 次火柴，且第 $2n-r$ 次是从甲盒中取的，即在前 $2n-r-1$ 在甲盒中取了

$$n-1, \text{ 其余在乙盒中取。所以 } P(A_0 B_r C) = \binom{2n-r-1}{n-1} \left(\frac{1}{2}\right)^{n-1} \cdot \left(\frac{1}{2}\right)^{n-r} \cdot \frac{1}{2}$$

由对称性知 $P(A_r B_0 C) = P(A_0 B_r D)$ ，所求概率为：

$$P(A_0 B_r C \cup A_r B_0 D) = 2P(A_0 B_r C) = \binom{2n-r-1}{n-1} \left(\frac{1}{2}\right)^{2n-r-1}$$

第二章 离散型随机变量

2.1 下列给出的是不是某个随机变量的分布列？

$$(1) \begin{pmatrix} 1 & 3 & 5 \\ 0.5 & 0.3 & 0.2 \end{pmatrix}$$

$$(2) \begin{pmatrix} 1 & 2 & 3 \\ 0.7 & 0.1 & 0.1 \end{pmatrix}$$

$$(3) \begin{pmatrix} 0 & 1 & 2 & \cdots & n & \cdots \\ \frac{1}{2} & \frac{1}{2}\left(\frac{1}{3}\right) & \frac{1}{2}\left(\frac{1}{3}\right)^2 & \cdots & \frac{1}{2}\left(\frac{1}{3}\right)^n & \cdots \end{pmatrix}$$

$$(4) \begin{pmatrix} 1 & 2 & \cdots & n & \cdots \\ \frac{1}{2} & \left(\frac{1}{2}\right)^2 & \cdots & \left(\frac{1}{2}\right)^n & \cdots \end{pmatrix}$$

解 (1) 是

(2) $0.7 + 0.1 + 0.1 \neq 1$ ，所以它不是随机变量的分布列。

(3) $\frac{1}{2} + \frac{1}{2}\left(\frac{1}{3}\right) + \frac{1}{2}\left(\frac{1}{3}\right)^2 + \cdots + \frac{1}{2}\left(\frac{1}{3}\right)^n + \cdots = \frac{3}{4}$ ，所以它不是随机变量的分布列。

(4) $\left(\frac{1}{2}\right)^n > 0$ ， n 为自然数，且 $\sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n = 1$ ，所以它是随机变量的分布列。

2.2 设随机变量 ξ 的分布列为： $P(\xi = k) = \frac{k}{15}, k=1,2,3,4,5$ ，求

(1) $P(\xi = 1 \text{ 或 } \xi = 2)$ ；

(2) $P\left(\frac{1}{2} < \xi < \frac{5}{2}\right)$ ； (3) $P(1 \leq \xi \leq 2)$ 。

解 (1) $P(\xi = 1 \text{ 或 } \xi = 2) = \frac{1}{15} + \frac{2}{15} = \frac{1}{5}$ ；

(2) $P\left(\frac{1}{2} < \xi < \frac{5}{2}\right) = P(\xi = 1) + P(\xi = 2) = \frac{1}{5}$ ；

(3) $P(1 \leq \xi \leq 2) = P(\xi = 1) + P(\xi = 2) = \frac{1}{5}$ 。

2.3 解 设随机变量 ξ 的分布列为 $P(\xi = i) = C \cdot \left(\frac{2}{3}\right)^i, i=1,2,3$ 。求 C 的值。

解 $C\left[\frac{2}{3}+\left(\frac{2}{3}\right)^2+\left(\frac{2}{3}\right)^3\right]=1$, 所以 $C=\frac{27}{38}$ 。

2.4 随机变量 ξ 只取正整数 N , 且 $P(\xi=N)$ 与 N^2 成反比, 求 ξ 的分布列。

解 根据题意知 $P(\xi=N)=\frac{C}{N^2}$, 其中常数 C 待定。由于 $\sum_{N=1}^{\infty} \frac{C}{N^2}=C \cdot \frac{\pi^2}{6}=1$, 所以 $C=\frac{6}{\pi^2}$, 即 ξ 的分布列为 $P(\xi=N)=\frac{6}{\pi^2 N^2}$, N 取正整数。

2.5 一个口袋中装有 m 个白球、 $n-m$ 个黑球, 不返回地连续从袋中取球, 直到取出黑球时停止。设此时取出了 ξ 个白球, 求 ξ 的分布列。

解 设 “ $\xi=k$ ” 表示前 k 次取出白球, 第 $k+1$ 次取出黑球, 则 ξ 的分布列为:

$$P(\xi=k)=\frac{m(m-1)\cdots(m-k+1)(n-m)}{n(n-1)\cdots(n-k)}, k=0,1,\cdots,m.$$

2.6 设某批电子管的合格品率为 $\frac{3}{4}$, 不合格品率为 $\frac{1}{4}$, 现在对该批电子管进行测试, 设第 ξ 次为首次测到合格品, 求 ξ 的分布列。

解 $P(\xi=k)=\left(\frac{1}{4}\right)^{k-1} \frac{3}{4}, k=1,2,\cdots$ 。

2.7 一个口袋中有 5 个同样大小的球, 编号为 1、2、3、4、5, 从中同时取出 3 只球, 以 ξ 表示取出球的取大号码, 求 ξ 的分布列。

$$\text{解 } P(\xi=k)=\frac{\binom{k-1}{2}}{\binom{5}{3}}, k=3,4,5.$$

2.8 抛掷一枚不均匀的硬币, 出现正面的概率为 p ($0 < p < 1$), 设 ξ 为一直掷到正、反面都出现时所需要的次数, 求 ξ 的分布列。

解 $P(\xi=k)=q^{k-1}p+p^{k-1}q, k=2,3,\cdots$, 其中 $q=1-p$ 。

2.9 两名篮球队员轮流投篮, 直到某人投中时为止, 如果第一名队员投中的概率为 0.4, 第二名队员投中的概率为 0.6, 求每名队员投篮次数的分布列。

解 设 ξ, η 表示第二名队员的投篮次数, 则

$$P(\xi=k)=0.6^{k-1}0.4^{k-1}0.4+0.6^k0.4^{k-1}0.6=0.76 \cdot 0.24^{k-1}, k=1,2,\cdots;$$

$$P(\eta = k) = 0.6^k 0.4^{k-1} 0.6 + 0.6^k 0.4^k 0.4 = 0.76 \cdot 0.6^k 0.4^{k-1}, k = 1, 2, \dots$$

2.10 设随机变量 ξ 服从普哇松分布, 且 $P(\xi = 1) = P(\xi = 2)$, 求 $P(\xi = 4)$ 。

$$\text{解 } P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda} (\lambda > 0) k = 0, 1, 2, \dots \quad \text{由 于 } \lambda e^{-\lambda} = \frac{\lambda^2}{2} e^{-\lambda}, \text{ 得}$$

$$\lambda_1 = 2, \lambda_2 = 0 \text{ (不合要求)。所以 } P(\xi = 4) = \frac{2^4}{4!} e^{-2} = \frac{2}{3} e^{-2}。$$

2.11 设某商店中每月销售某种商品的数量服从参数为 7 的普哇松分布, 问在月初进货时应进多少件此种商品, 才能保证当月不脱销的概率为 0.999。

解 设 ξ 为该种商品当月销售数, x 为该种商品每月进货数, 则

$$P(\xi \leq x) \geq 0.999. \text{ 查普哇松分布的数值表, 得 } x \geq 16。$$

2.12 如果在时间 t (分钟) 内, 通过某交叉路口的汽车数量服从参数与 t 成正比的普哇松分布。已知在一分钟内没有汽车通过的概率为 0.2, 求在 2 分钟内有多于一辆汽车通过的概率。

解 设 ξ 为时间 t 内通过交叉路口的汽车数, 则

$$P(\xi = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} (\lambda > 0), k = 0, 1, 2, \dots$$

$$t = 1 \text{ 时, } P(\xi = 0) = e^{-\lambda} = 0.2, \text{ 所以 } \lambda = \ln 5; \quad t = 2 \text{ 时, } \lambda t = 2 \ln 5, \text{ 因而}$$

$$P(\xi > 1) = 1 - P(\xi = 0) - P(\xi = 1) = (24 - \ln 25) / 25 \approx 0.83。$$

2.13 一本 500 页的书共有 500 个错误, 每个错误等可能地出现在每一页上 (每一页的印刷符号超过 500 个)。试求指定的一页上至少有三个错误的概率。

解 在指定的一页上出现某一个错误的概率 $p = \frac{1}{500}$, 因而, 至少出现三个错误的概率为

$$\sum_{k=3}^{500} \binom{500}{k} \left(\frac{1}{500}\right)^k \left(\frac{499}{500}\right)^{500-k} = 1 - \sum_{k=0}^2 \binom{500}{k} \left(\frac{1}{500}\right)^k \left(\frac{499}{500}\right)^{500-k}$$

利用普哇松定理求近似值, 取 $\lambda = np = 500 \times \frac{1}{500} = 1$, 于是上式右端等于

$$1 - \sum_{k=0}^2 \frac{1}{k!} e^{-1} = 1 - \frac{5}{2e} \approx 0.080301$$

2.14 某厂产品的不合格品率为 0.03, 现在要把产品装箱, 若要以不小于 0.9 的概率保证每箱中至少有 100 个合格品, 那么每箱至少应装多少个产品?

解 设每箱至少装 $100+x$ 个产品, 其中有 k 个次品, 则要求 x , 使

$$0.9 \leq \sum_{k=0}^x \binom{100+x}{k} 0.03^k 0.97^{100+x-k},$$

利用普哇松分布定理求近似值，取 $\lambda = (100+x) \times 0.03 \approx 3$ ，于是上式相当于

$$0.9 \leq \sum_{k=0}^x \frac{3^k}{k!} e^{-3}, \text{ 查普哇松分布数值表, 得 } x=5.$$

2.15 设二维随机变量 (ξ, η) 的联合分布列为：

$$P(\xi = n, \eta = m) = \frac{\lambda^n p^m (1-p)^{n-m}}{m!(n-m)!} e^{-\lambda} \quad (\lambda > 0, 0 < p < 1) \quad m = 0, 1, \dots, n \quad n = 0, 1, 2, \dots$$

求边际分布列。

$$\begin{aligned} \text{解 } P(\xi = n) &= \sum_{m=0}^n P(\xi = n, \eta = m) = \frac{\lambda^n e^{-\lambda}}{n!} \sum_{m=0}^n \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m} \\ &= \frac{\lambda^n e^{-\lambda}}{n!} \quad n = 0, 1, 2, \dots \end{aligned}$$

$$\begin{aligned} P(\eta = m) &= \sum_{n=0}^{\infty} P(\xi = n, \eta = m) = \frac{p^m e^{-\lambda}}{m!} \sum_{n=m}^{\infty} \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m} \\ &= \frac{(\lambda p)^m e^{-\lambda p}}{m!} \quad m = 0, 1, 2, \dots \end{aligned}$$

2.17 在一批产品中一等品占 50%，二等品占 30%，三等品占 20%。从中任取 4 件，设一、二、三等品的件数分别为 ξ 、 η 、 ζ ，求 (ξ, η, ζ) 的联合分布列与各自的边际分布列。

$$\text{解 } P(\xi = m, \eta = n, \zeta = k) = \frac{4!}{m!n!k!} 0.5^m 0.3^n 0.2^k, \quad m, n, k = 0, 1, 2, 3, 4 \quad m+n+k=4.$$

$$P(\xi = m) = \binom{4}{m} 0.5^m 0.5^{4-m}, \quad m = 0, 1, 2, 3, 4;$$

$$P(\eta = n) = \binom{4}{n} 0.3^n 0.7^{4-n}, \quad n = 0, 1, 2, 3, 4;$$

$$P(\zeta = k) = \binom{4}{k} 0.2^k 0.8^{4-k}, \quad k = 0, 1, 2, 3, 4.$$

2.18 抛掷三次均匀的硬币，以 ξ 表示出现正面的次数，以 η 表示正面出现次数与反面出现次数之差的绝对值，求 (ξ, η) 的联合分布列及边际分布列。

2.21 设随机变量 ξ 与 η 独立，且 $P(\xi = 1) = P(\eta = 1) = p > 0$ ，

又 $P(\xi = 0) = P(\eta = 0) = 1 - p > 0$, 定义 $\zeta = \begin{cases} 1 & \text{若 } \xi + \eta \text{ 为偶数} \\ 0 & \text{若 } \xi + \eta \text{ 为奇数} \end{cases}$, 问 p 取什么值

时 ξ 与 ζ 独立?

$$\text{解 } P(\zeta = 1) = P(\xi = 0)P(\eta = 0) + P(\xi = 1)P(\eta = 1) = (1 - p)^2 + p^2$$

$$P(\zeta = 0) = P(\xi = 0)P(\eta = 1) + P(\xi = 1)P(\eta = 0) = 2p(1 - p)$$

而 $P(\xi = 1, \zeta = 1) = P(\xi = 1, \eta = 1) = p^2$, 由 $P(\xi = 1, \zeta = 1) = P(\xi = 1)P(\zeta = 1)$ 得 $p = \frac{1}{2}$

2.22 设随机变量 ξ 与 η 独立, 且 $P(\xi = \pm 1) = P(\eta = \pm 1) = \frac{1}{2}$, 定义 $\zeta = \xi\eta$,

证明 ζ, ξ, η 两两独立, 但不相互独立。

$$\text{证明 } P(\zeta = 1) = P(\xi = 1)P(\eta = 1) + P(\xi = -1)P(\eta = -1) = \frac{1}{2}$$

$$P(\zeta = -1) = P(\xi = 1)P(\eta = -1) + P(\xi = -1)P(\eta = 1) = \frac{1}{2}$$

$$\text{因为 } P(\xi = 1, \zeta = 1) = P(\xi = 1, \eta = 1) = \frac{1}{4} = P(\xi = 1)P(\zeta = 1)$$

$$P(\xi = 1, \zeta = -1) = P(\xi = 1, \eta = -1) = \frac{1}{4} = P(\xi = 1)P(\zeta = -1)$$

$$P(\xi = -1, \zeta = 1) = P(\xi = -1, \eta = -1) = \frac{1}{4} = P(\xi = -1)P(\zeta = 1)$$

$$P(\xi = -1, \zeta = -1) = P(\xi = -1, \eta = 1) = \frac{1}{4} = P(\xi = -1)P(\zeta = -1)$$

所以 ζ, ξ 相互独立。同理 η 与 ζ 相互独立。

但是 $P(\xi = 1, \eta = 1, \zeta = 1) \neq P(\xi = 1)P(\eta = 1)P(\zeta = 1)$, 因而 ζ, ξ, η 不相互独立。

2.23 设随机变量 ξ 与 η 独立, 且只取值 1、2、3、4、5、6, 证明 $\xi + \eta$ 不服

从均匀分 (即不可能有 $P(\xi + \eta = k) = \frac{1}{11}, k = 2, 3, \dots, 12$ 。)

证明 设 $P(\xi = k) = p_k, P(\eta = k) = q_k, k = 1, 2, \dots, 6$ 。

若 $P(\xi + \eta = k) = \frac{1}{11}, k = 2, 3, \dots, 12$, 则

$$P(\xi + \eta = 2) = p_1 q_1 = \frac{1}{11} \quad (1)$$

$$P(\xi + \eta = 7) = p_1 q_6 + p_2 q_5 + \dots + p_6 q_1 = \frac{1}{11} \quad (2)$$

$$P(\xi + \eta = 12) = p_6 q_6 = \frac{1}{11} \quad (3)$$

将 (2) 式减去 (1) 式, 得: $(p_6 - p_1)q_1 < 0$, 于是 $p_6 < p_1$ 。同理 $q_6 < q_1$ 。

因此 $p_6 q_6 < p_1 q_1 = \frac{1}{11}$, 与 (3) 式矛盾。

2.24 已知随机变量 ξ 的分布列为 $\begin{pmatrix} 0 & \frac{\pi}{2} & \pi \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$, 求 $\eta = \frac{2}{3}\xi + 2$ 与 $\zeta = \cos\xi$ 的分布列。

解 η 分布列为 $P(\eta = 2) = \frac{1}{4}$, $P(\eta = 2 + \frac{\pi}{3}) = \frac{1}{2}$, $P(\eta = 2 + \frac{2\pi}{3}) = \frac{1}{4}$;

ζ 的分布列为 $P(\zeta = -1) = \frac{1}{4}$, $P(\zeta = 0) = \frac{1}{2}$, $P(\zeta = 1) = \frac{1}{4}$ 。

2.25 已知离散型随机变量 ξ 的分布列为 $\begin{pmatrix} -2 & -1 & 0 & 1 & 3 \\ \frac{1}{5} & \frac{1}{6} & \frac{1}{5} & \frac{1}{15} & \frac{3}{30} \end{pmatrix}$, 求 $\eta = \xi^2$ 的分布列。

解 $P(\eta = 0) = \frac{1}{5}$, $P(\eta = 1) = \frac{7}{30}$, $P(\eta = 4) = \frac{1}{5}$, $P(\eta = 9) = \frac{11}{30}$

2.26 设离散型随机变量 ξ 与 η 的分布列为 $\xi: \begin{pmatrix} 0 & 1 & 3 \\ \frac{1}{2} & \frac{3}{8} & \frac{1}{8} \end{pmatrix}$, $\eta: \begin{pmatrix} 0 & 1 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$,

且 ξ 与 η 相互独立, 求 $\zeta = \xi + \eta$ 的分布列。

解 $\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ \frac{1}{6} & \frac{11}{24} & \frac{1}{4} & \frac{3}{24} & \frac{1}{12} \end{pmatrix}$

2.27 设独立随机变量 ξ 与 η 分别服从二项分布: $b(k, n_1, p)$ 与 $b(k, n_2, p)$, 求 $\xi + \eta$ 的分布列。

解 设 ξ 为 n_1 重贝努里试验中事件 A 发生的次数 (在每次试验中 $P(A) = p$), η 为 n_2 重贝努里试验中事件 A 发生的次数 (在每次试验中 $P(A) = p$), 而 ξ 与 η 相互独立, 所以 $\xi + \eta$ 为 $n_1 + n_2$ 重贝努里试验中事件 A 发生的次数, 因而

$$P(\xi + \eta = k) = \binom{n_1 + n_2}{k} p^k q^{n_1 + n_2 - k}, \quad k = 0, 1, \dots, n_1 + n_2.$$

2.28 设 ξ 与 η 为独立同分布的离散型随机变量, 其分布列为

$$P(\xi = n) = P(\eta = n) = \frac{1}{2^n}, n = 1, 2, \dots$$

求 $\xi + \eta$ 的分布列。

$$\text{解 } P(\xi + \eta = n) = \sum_{k=1}^{n-1} P(\xi = k)P(\eta = n - k) = \sum_{k=1}^{n-1} \frac{1}{2^k} \cdot \frac{1}{2^{n-k}} = \frac{n-1}{2^n}$$

2.29 设随机变量 ξ 具有分布: $P(\xi = k) = \frac{1}{5}, k = 1, 2, 3, 4, 5$, 求 $E\xi$ 、 $E\xi^2$ 及 $E(\xi + 2)^2$ 。

$$\text{解, } E\xi = \frac{1}{5}(1 + 2 + 3 + 4 + 5) = 3, \quad E\xi^2 = \frac{1}{5}(1^2 + 2^2 + 3^2 + 4^2 + 5^2) = 11$$

$$E(\xi + 2)^2 = E\xi^2 + 4E\xi + 4 = 27$$

2.30 设随机变量 ξ 具有分布: $P(\xi = k) = \frac{1}{2^k}, k = 1, 2, \dots$, 求 $E\xi$ 及 $D\xi$ 。

$$\text{解 } E\xi = \sum_{k=1}^{\infty} \frac{k}{2^k} = \frac{1}{2} \sum_{k=1}^{\infty} k \left(\frac{1}{2}\right)^{k-1} = 2, \quad E\xi^2 = \sum_{k=1}^{\infty} \frac{k^2}{2^k} = \frac{1}{2} \sum_{k=1}^{\infty} k^2 \left(\frac{1}{2}\right)^{k-1} = 6$$

$$D\xi = E\xi^2 - (E\xi)^2 = 2$$

2.31 设离散型随机变量 ξ 的分布列为: $P[\xi = (-1)^k \frac{2^k}{k}] = \frac{1}{2^k}, k = 1, 2, \dots$, 问

ξ 是否有数学期望?

$$\text{解 } \sum_{k=1}^{\infty} |(-1)^k \frac{2^k}{k}| \cdot \frac{1}{2^k} = \sum_{k=1}^{\infty} \frac{1}{k}, \text{ 因为级数 } \sum_{k=1}^{\infty} \frac{1}{k} \text{ 发散, 所以 } \xi \text{ 没有数学期望。}$$

2.32 用天平称某种物品的重量 (砝码仅允许放在一个秤盘中), 物品的重量以相同的概率为 1 克、2 克、 \dots 、10 克, 现有三组砝码:

(甲组) 1, 2, 2, 5, 10 (克)

(乙组) 1, 2, 3, 4, 10 (克)

(丙组) 1, 1, 2, 5, 10 (克)

问哪一组砝码称重时所用的平均砝码数最少?

解 设 ξ_1 、 ξ_2 、 ξ_3 分别表示及甲组、乙组、丙组砝码称重时所用的砝码数, 则有

物品重量度	1	2	3	4	5	6	7	8	9	10
ξ_1	1	1	2	2	1	2	2	3	3	1
ξ_2	1	1	1	1	2	2	2	3	3	1

$$\xi_3 \quad 1 \quad 1 \quad 2 \quad 3 \quad 1 \quad 2 \quad 2 \quad 3 \quad 4 \quad 1$$

$$\text{于是 } E\xi_1 = \frac{1}{10}(1+1+2+2+1+2+2+3+3+1) = 1.8$$

$$E\xi_2 = \frac{1}{10}(1+1+1+1+2+2+2+3+3+1) = 1.7$$

$$E\xi_3 = \frac{1}{10}(1+1+2+3+1+2+2+3+4+1) = 2$$

所以，用乙组砝码称重时所用的平均砝码数最少。

2.33 某个边长为 500 米的正方形场地，用航空测量法测得边长的误差为：0 米的概率是 0.49， ± 10 米的概率各是 0.16， ± 20 米的概率各是 0.08， ± 30 米的概率各是 0.05，求场地面积的数学期望。

解 设场地面积为 S 米²，边长的误差为 ξ 米，则 $S = (\xi + 500)^2$ 且

$$E\xi = 0 \quad E\xi^2 = 2(10^2 \times 0.16 + 20^2 \times 0.08 + 30^2 \times 0.05) = 186$$

$$\text{所以 } ES = E(\xi + 500)^2 = E\xi^2 + 1000E\xi + 250000 = 250186(\text{米}^2)$$

2.34 对三架仪器进行检验，各仪器发生故障是独立的，且概率分别为 p_1 、 p_2 、 p_3 。试证发生故障的仪器数的数学期望 $p_1 + p_2 + p_3$ 。

$$\text{证 令 } \xi_i = \begin{cases} 1 & \text{第 } i \text{ 架仪器发生故障} \\ 0 & \text{第 } i \text{ 架仪器未发生故障} \end{cases} \quad i = 1, 2, 3$$

$$\xi \text{ 为发生故障的仪器数，则 } E\xi_i = P(\xi_i = 1) = p_i, i = 1, 2, 3,$$

$$\text{所以 } E\xi = E\xi_1 + E\xi_2 + E\xi_3 = p_1 + p_2 + p_3。$$

2.37 如果在 15000 件产品中有 1000 件不合格品，从中任意抽取 150 件进行检查，求查得不合格品数的数学期望。

解 设，

$$\text{则 } \eta_i \text{ 的分布列为 } \begin{pmatrix} 1 & 0 \\ \frac{1}{15} & \frac{14}{15} \end{pmatrix}, \text{ 因而 } E\eta_i = \frac{1}{15}。 \text{ 设 } \xi \text{ 为查得的不合格品数，}$$

则

$$\xi = \sum_{i=1}^{150} \eta_i, \text{ 所以 } E\xi = \sum_{i=1}^{150} E\eta_i = 10。$$

2.38 从数字 0, 1, ..., n 中任取两个不同的数字，求这两个数字之差的绝对值的数学期望。

$$\text{解 设 } \xi \text{ 为所选两个数字之差的绝对值，则 } P(\xi = k) = \frac{n-k+1}{\binom{n+1}{2}}, k = 1, 2, \dots, n,$$

$$\text{于是 } E\xi = \sum_{k=1}^n k \frac{n-k+1}{\binom{n+1}{2}} = \frac{2}{n(n+1)} \sum_{k=1}^n [(n+1)k - k^2] = \frac{n+2}{3}.$$

2.39 把数字 $1, 2, \dots, n$ 任意在排成一列, 如果数字 k 恰好出现在第 k 个位置上, 则称有一个匹配, 求匹配数的数学期望。

$$\text{解 设 } \xi_k = \begin{cases} 1 & \text{数字 } k \text{ 出现在第 } k \text{ 个位置上} \\ 0 & \text{数字 } k \text{ 不在第 } k \text{ 个位置上} \end{cases} \quad \text{则 } \xi_k \text{ 的分布列为: } \begin{pmatrix} 1 & 0 \\ \frac{1}{n} & 1 - \frac{1}{n} \end{pmatrix}$$

$$\text{于是 } E\xi_k = P(\xi_k = 1) = \frac{1}{n}, \text{ 设匹配数为 } \xi, \text{ 则 } \xi = \sum_{k=1}^n \xi_k, \text{ 因而 } E\xi = \sum_{k=1}^n E\xi_k = 1.$$

2.40 设 ξ 为取非负整数值的随机变量, 证明:

$$(1) \quad E\xi = \sum_{n=1}^{\infty} P(\xi \geq n);$$

$$(2) \quad D\xi = 2 \sum_{n=1}^{\infty} nP(\xi \geq n) - E\xi(E\xi + 1).$$

证明 (1) 由于 $E\xi = \sum_{n=0}^{\infty} nP(\xi = n)$ 存在, 所以该级数绝对收敛。从而

$$E\xi = \sum_{n=1}^{\infty} nP(\xi = n) = \sum_{n=1}^{\infty} \sum_{i=1}^n P(\xi = n) = \sum_{i=1}^{\infty} \sum_{n=i}^{\infty} P(\xi = n) = \sum_{i=1}^{\infty} P(\xi \geq i).$$

(2) $D\xi$ 存在, 所以级数 $E\xi^2 = \sum_{n=0}^{\infty} n^2 P(\xi = n)$ 也绝对收敛, 从而

$$D\xi = E\xi^2 + E\xi - E\xi(E\xi + 1) = \sum_{n=1}^{\infty} n(n+1)P(\xi = n) - E\xi(E\xi + 1)$$

$$= 2 \sum_{n=1}^{\infty} \sum_{i=1}^n iP(\xi = n) - E\xi(E\xi + 1) = 2 \sum_{i=1}^{\infty} \sum_{n=i}^{\infty} iP(\xi = n) - E\xi(E\xi + 1)$$

$$= 2 \sum_{n=1}^{\infty} nP(\xi \geq n) - E\xi(E\xi + 1).$$

2.41 在贝努里试验中, 每次试验成功的概率为 p , 试验进行到成功与失败均出现时停止, 求平均试验次数。

解 设成功与失败均出现时的试验次数为 ξ , 则

$$P(\xi \geq 1) = 1, \quad P(\xi \geq n) = p^{n-1} + q^{n-1}, n = 2, 3, \dots (q = 1 - p)$$

$$\text{利用上题的结论, } E\xi = P(\xi \geq 1) + \sum_{n=2}^{\infty} P(\xi \geq n) = 1 + \sum_{n=2}^{\infty} (p^{n-1} + q^{n-1})$$

$$= 1 + \frac{p}{1-p} + \frac{q}{1-q} = \frac{p^2 - p + 1}{p(1-p)}$$

2.42 从一个装有 m 个白球、 n 个黑球的袋中摸球，直至摸到白球时停止。如果(1)摸球是为返回的，(2)摸球是不返回的，试对这两种不同的摸球方式求：取出黑球数的数学期望。

解 略。

2.43 对一批产品进行检验，如果检查到第 n_0 件仍未发现不合格品就认为这批产品合格，如在尚未抽到第 n_0 件时已检查到不合格品即停止继续检查，且认为这批产品不合格。设产品数量很大，可以认为每次检查到不合格品的概率都是 p ，问平均每批要检查多少件？

解 略。

2.44 流水作业线上生产出的每个产品为不合格品的概率 p ，当生产出 k 个不合格品时即停工检修一次。求在两次检修之间产品总数的数学期望与方差。

解 设第 $i-1$ 个不合格出现后到第 i 个不合格品出现时的产品数为 ξ_i ，

$$i = 1, 2, \dots, k. \text{ 又在两次检修之间产品总数为 } \xi, \text{ 则 } \xi = \sum_{i=1}^k \xi_i.$$

因 ξ_i 独立同分布， $P(\xi_i = j) = q^{j-1}p, j = 1, 2, \dots (q = 1 - p)$ ，由此得：

$$E\xi_i = \sum_{j=1}^{\infty} jq^{j-1}p = \frac{1}{p}, \quad E\xi_i^2 = \sum_{j=1}^{\infty} j^2 q^{j-1}p = \frac{2-p}{p^2},$$

$$D\xi_i = E\xi_i^2 - (E\xi_i)^2 = \frac{1-p}{p^2}.$$

$$E\xi = \sum_{i=1}^k E\xi_i = \frac{k}{p}, \quad D\xi = \sum_{i=1}^k D\xi_i = \frac{k(1-p)}{p^2}.$$

2.46 设随机变量 ξ 与 η 独立，且方差存在，则有

$$D(\xi\eta) = D\xi \cdot D\eta + (E\xi)^2 \cdot D\eta + D\xi \cdot (E\eta)^2 \quad (\text{由此并可得 } D(\xi\eta) \geq D\xi \cdot D\eta)$$

$$\text{证明 } D(\xi\eta) = E\xi^2\eta^2 - (E\xi\eta)^2 = E\xi^2 E\eta^2 - (E\xi)^2 (E\eta)^2$$

$$\begin{aligned}
&= E\xi^2 E\eta^2 - E\xi^2 (E\eta)^2 + E\xi^2 (E\eta)^2 - (E\xi)^2 (E\eta)^2 \\
&= E\xi^2 D\eta - (E\eta)^2 D\xi = D\xi \cdot D\eta + (E\xi)^2 \cdot D\eta + D\xi \cdot (E\eta)^2
\end{aligned}$$

2.47 在整数 0 到 9 中先后按下列两种情况任取两个数, 记为 ξ 和 η : (1) 第一个数取后放回, 再取第二个数; (2) 第一个数取后不放回就取第二个数, 求在 $\eta = k (0 \leq k \leq 9)$ 的条件下 ξ 的分布列。

解 (1) $P(\xi = i | \eta = k) = \frac{1}{10} \quad i = 0, 1, \dots, 9.$

(2) $P(\xi = i | \eta = k) = \frac{1}{9} \quad (i = 0, 1, \dots, 9, i \neq k), \quad P(\xi = k | \eta = k) = 0$

2.49 在 n 次贝努里试验中, 事件 A 出现的概率为 p , 令

$$\xi_i = \begin{cases} 1 & \text{在第 } i \text{ 次试验中 } A \text{ 出现} \\ 0 & \text{在第 } i \text{ 次试验中 } A \text{ 不出现} \end{cases} \quad i = 1, 2, \dots, n$$

求在 $\xi_1 + \xi_2 + \dots + \xi_n = r (0 \leq r \leq n)$ 的条件下, $\xi_i (0 \leq i \leq n)$ 的分布列。

解 $P(\xi_i = 0 | \xi_1 + \xi_2 + \dots + \xi_n = r) = \frac{P(\xi_i = 0, \xi_1 + \dots + \xi_{i-1} + \xi_{i+1} + \dots + \xi_n = r)}{P(\xi_1 + \xi_2 + \dots + \xi_n = r)}$

$$= \frac{q \binom{n-1}{r} p^r q^{n-1-r}}{\binom{n}{r} p^r q^{n-r}} = \frac{n-r}{n}$$

$$P(\xi_i = 1 | \xi_1 + \xi_2 + \dots + \xi_n = r) = 1 - \frac{n-r}{n} = \frac{r}{n}.$$

2.50 设随机变量 ξ_1, ξ_2 相互独立, 分别服从参数为 λ_1 与 λ_2 的普哇松分布, 试证:

$$P(\xi_1 = k | \xi_1 + \xi_2 = n) = \binom{n}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^k \left(1 - \frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^{n-k}$$

证明 $P(\xi_1 = k | \xi_1 + \xi_2 = n) = \frac{P(\xi_1 = k, \xi_1 + \xi_2 = n)}{P(\xi_1 + \xi_2 = n)}$

$$= \frac{P(\xi_1 = k) P(\xi_2 = n - k)}{P(\xi_1 + \xi_2 = n)}$$

由普哇松分布的可加性知 $\xi_1 + \xi_2$ 服从参数为 $\lambda_1 + \lambda_2$ 的普哇松分布, 所以

$$P(\xi_1 = k | \xi_1 + \xi_2 = n) = \frac{\frac{\lambda_1^k}{k!} e^{-\lambda_1} \cdot \frac{\lambda_2^{n-k}}{(n-k)!} e^{-\lambda_2}}{\frac{(\lambda_1 + \lambda_2)^n}{n!} e^{-(\lambda_1 + \lambda_2)}} = \binom{n}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^k \left(1 - \frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^{n-k}$$

2.51 设 $\xi_1, \xi_2, \dots, \xi_r$ 为 r 个相互独立随机变量, 且 $\xi_i (1 \leq i \leq r)$ 服从同一几何分布, 即有 $P(\xi_i = k) = qp^{k-1}, k=1, 2, \dots, (1 \leq i \leq r)$, 其中 $q=1-p$ 。试证明在 $\xi_1 + \xi_2 + \dots + \xi_r = n$ 的条件下, $(\xi_1, \xi_2, \dots, \xi_r)$ 的分布是均匀分布, 即

$$P(\xi_1 = n_1, \dots, \xi_r = n_r | \xi_1 + \xi_2 + \dots + \xi_r = n) = \frac{1}{\binom{n-1}{r-1}}, \text{ 其中 } n_1 + n_2 + \dots + n_r = n.$$

$$\begin{aligned} \text{证明} \quad P(\xi_1 = n_1, \dots, \xi_r = n_r | \xi_1 + \xi_2 + \dots + \xi_r = n) &= \frac{P(\xi_1 = n_1, \dots, \xi_r = n_r, \xi_1 + \dots + \xi_r = n)}{P(\xi_1 + \dots + \xi_r = n)} \\ &= \frac{P(\xi_1 = n_1, \dots, \xi_r = n_r)}{P(\xi_1 + \dots + \xi_r = n)} \end{aligned}$$

由于 $\xi_1, \xi_2, \dots, \xi_r$ 相互独立且服从同一几何分布, 所以

$$P(\xi_1 + \xi_2 + \dots + \xi_r = n) = \sum_{\substack{k_1 + \dots + k_r = n \\ k_i = 1, 2, \dots \\ i=1, \dots, r}} \left(\prod_{i=1}^r q \cdot p^{k_i-1} \right) = \binom{n-1}{r-1} q^r p^{n-r}.$$

$$\text{从而 } P(\xi_1 = n_1, \dots, \xi_r = n_r | \xi_1 + \xi_2 + \dots + \xi_r = n) = \frac{q^r p^{n-r}}{\binom{n-1}{r-1} q^r p^{n-r}} = \frac{1}{\binom{n-1}{r-1}}.$$

第三章

3.1 设随机变数 ξ 的分布函数为 $F(x)$, 试以 $F(x)$ 表示下列概率:

$$(1) P(\xi = a); (2) P(\xi \leq a); (3) P(\xi \geq a); (4) P(\xi > a)$$

解: (1) $P(\xi = a) = F(a+0) - F(a)$;

$$(2) P(\xi \leq a) = F(a+0);$$

$$(3) P(\xi \geq a) = 1 - F(a);$$

$$(4) P(\xi > a) = 1 - F(a+0).$$

3.2 函数 $F(x) = \frac{1}{1+x^2}$ 是否可以作为某一随机变量的分布函数, 如果

- (1) $-\infty < x < \infty$
- (2) $0 < x < \infty$, 在其它场合适当定义;
- (3) $-\infty < x < 0$, 在其它场合适当定义。

解: (1) $F(x)$ 在 $(-\infty, \infty)$ 内不单调, 因而不可能是随机变量的分布函数;

(2) $F(x)$ 在 $(0, \infty)$ 内单调下降, 因而也不可能是随机变量的分布函数;

(3) $F(x)$ 在 $(-\infty, 0)$ 内单调上升、连续且 $F(-\infty, 0)$, 若定义

$$\tilde{F}(x) = \begin{cases} F(x) & -\infty < x < 0 \\ 1 & x \geq 0 \end{cases}$$

则 $\tilde{F}(x)$ 可以是某一随机变量的分布函数。

3.3 函数 $\sin x$ 是不是某个随机变数 ξ 的分布密度? 如果 ξ 的取值范围为

- (1) $[0, \frac{\pi}{2}]$; (2) $[0, \pi]$; (3) $[0, \frac{3}{2}\pi]$ 。

解: (1) 当 $x \in [0, \frac{\pi}{2}]$ 时, $\sin x \geq 0$ 且 $\int_0^{\frac{\pi}{2}} \sin x dx = 1$, 所以 $\sin x$ 可以是某个随机变量的分布密度;

(2) 因为 $\int_0^{\pi} \sin x dx = 2 \neq 1$, 所以 $\sin x$ 不是随机变量的分布密度;

(3) 当 $x \in [\pi, \frac{3}{2}\pi]$ 时, $\sin x \leq 0$, 所以 $\sin x$ 不是随机变量的分布密度。

3.4 设随机变数 ξ 具有对称的分布密度函数 $p(x)$, 即 $p(x) = p(-x)$, 证明: 对任意的

$a > 0$, 有 (1) $F(-a) = 1 - F(a) = \frac{1}{2} - \int_0^a p(x) dx$;

(2) $P(|\xi| < a) = 2F(a) - 1$;

(3) $P(|\xi| > a) = 2[1 - F(a)]$ 。

$$\begin{aligned} \text{证: (1) } F(-a) &= \int_{-\infty}^{-a} p(x) dx = 1 - \int_{-a}^{\infty} p(x) dx \\ &= 1 + \int_a^{\infty} p(-x) dx = 1 - \int_{-\infty}^a p(x) dx \\ &= 1 - F(a) = 1 - \int_{-\infty}^0 p(x) dx \\ &\quad - \int_0^a p(x) dx = \frac{1}{2} - \int_0^a p(x) dx; \end{aligned}$$

(2) $P(|\xi| < a) = \int_{-a}^a p(x) dx = 2 \int_0^a p(x) dx$, 由 (1) 知

$$1 - F(a) = \frac{1}{2} - \int_0^a p(x) dx$$

故上式右端 $= 2F(a) - 1$;

(3) $P(|\xi| > a) = 1 - P(|\xi| < a) = 1 - [2F(a) - 1] = 2[1 - F(a)]$ 。

3.5 设 $F_1(x)$ 与 $F_2(x)$ 都是分布函数, 又 $a > 0, b > 0$ 是两个常数, 且 $a + b = 1$ 。证明

$$F(x) = aF_1(x) + bF_2(x)$$

也是一个分布函数, 并由此讨论, 分布函数是否只有离散型和连续型这两种类型?

证: 因为 $F_1(x)$ 与 $F_2(x)$ 都是分布函数, 当 $x_1 < x_2$ 时, $F_1(x_1) \leq F_1(x_2)$,

$F_2(x_1) \leq F_2(x_2)$, 于是

$$F(x_1) = aF_1(x_1) + bF_2(x_1) \leq aF_1(x_2) + bF_2(x_2) = F(x_2)$$

又

$$\lim_{x \rightarrow -\infty} F(x) = \lim_{x \rightarrow -\infty} [aF_1(x) + bF_2(x)] = 0$$

$$\lim_{x \rightarrow \infty} F(x) = \lim_{x \rightarrow \infty} [aF_1(x) + bF_2(x)] = a + b = 1$$

$$F(x-0) = aF_1(x-0) + bF_2(x-0) = aF_1(x) + bF_2(x) = F(x)$$

所以, $F(x)$ 也是分布函数。

取 $a = b = \frac{1}{2}$, 又令

$$F_1(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases} \quad F_2(x) = \begin{cases} 0 & x \leq 0 \\ x & 0 < x \leq 1 \\ 1 & x > 1 \end{cases}$$

这时

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1+x}{2} & 0 < x \leq 1 \\ 1 & x > 1 \end{cases}$$

显然, 与 $F(x)$ 对应的随机变量不是取有限个或可列个值, 故 $F(x)$ 不是离散型的, 而

$F(x)$ 不是连续函数, 所以它也不是连续型的。

3.6 设随机变数 ξ 的分布函数为

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

求相应的密度函数，并求 $P(\xi \leq 1)$ 。

解： $\frac{d}{dx}[1 - (1+x)e^{-x}] = xe^{-x}$ ，所以相应的密度函数为

$$p(x) = \begin{cases} xe^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$P(\xi \leq 1) = F(1) = 1 - \frac{2}{e}。$$

3.7 设随机变数 ξ 的分布函数为

$$F(x) = \begin{cases} 0 & x < 0 \\ Ax^2 & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

求常数 A 及密度函数。

解：因为 $F(1-0) = F(1)$ ，所以 $A=1$ ，密度函数为

$$p(x) = \begin{cases} 2x & 0 \leq x < 1 \\ 0 & \text{其它} \end{cases}$$

3.8 随机变数 ξ 的分布函数为 $F(x) = A + B \arctan x$ ，求常数 A 与 B 及相应的密度函数。

解：因为 $\lim_{x \rightarrow -\infty} F(x) = A + B(-\frac{\pi}{2}) = 0$

$$\lim_{x \rightarrow +\infty} F(x) = A + B\frac{\pi}{2} = 1$$

所以

$$A = \frac{1}{2}, B = \frac{1}{\pi}$$

因而

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan x, p(x) = F'(x) = \frac{1}{\pi(1+x^2)}。$$

3.9 已知随机变数 ξ 的分布函数为

$$p(x) = \begin{cases} x & 0 < x \leq 1 \\ 2-x & 1 < x \leq 2 \\ 0 & \text{其它} \end{cases}$$

(1) 求相应的分布函数 $F(x)$;

(2) 求 $P(\xi < 0.5), P(\xi > 1.3), P(0.2 < \xi < 1.2)$ 。

$$\text{解: } F(x) = \begin{cases} 0 & x \leq 0 \\ \int_0^x y dy = \frac{1}{2}x^2 & 0 < x \leq 1 \\ \int_0^1 y dy + \int_1^x (2-y) dy = 2x - \frac{1}{2}x^2 - 1 & 1 < x \leq 2 \\ 1 & x > 2 \end{cases}$$

$$P(\xi < 0.5) = F(0.5) = \frac{1}{8}$$

$$P(\xi > 1.3) = 1 - P(\xi \leq 1.3) = 1 - F(1.3) = 0.245$$

$$P(0.2 < \xi < 1.2) = F(1.2) - F(0.2) = 0.66$$

3.10 确定下列函数中的常数 A , 使该函数成为一元分布的密度函数。

(1) $p(x) = Ae^{-|x|}$;

(2) $p(x) = \begin{cases} A \cos x & -\frac{\pi}{2} \leq x \leq \frac{\pi}{2} \\ 0 & \text{其它} \end{cases}$

(3) $p(x) = \begin{cases} Ax^2 & 1 \leq x \leq 2 \\ Ax & 2 < x < 3 \\ 0 & \text{其它} \end{cases}$

解: (1) $\int_{-\infty}^{\infty} Ae^{-|x|} dx = 2A \int_0^{\infty} e^{-x} dx = 2A = 1$ 所以 $A = \frac{1}{2}$;

(2) $\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} A \cos x dx = 2A \int_0^{\frac{\pi}{2}} \cos x dx = 2A = 1$, 所以 $A = \frac{1}{2}$;

(3) $\int_1^2 Ax^2 dx + \int_2^3 Ax dx = \frac{29}{6}A = 1$, 所以 $A = \frac{6}{29}$ 。

3.12 在半径为 R , 球心为 O 的球内任取一点 P , 求 $\xi = OP$ 的分布函数。

解: 当 $0 \leq x \leq R$ 时

$$F(x) P(\xi < x) = \frac{\frac{4}{3}\pi x^3}{\frac{4}{3}\pi R^3} = \left(\frac{x}{R}\right)^3$$

所以

$$F(x) = \begin{cases} 0 & x < 0 \\ (\frac{x}{R})^3 & 0 \leq x \leq R \\ 1 & x > R \end{cases}$$

3.13 某城市每天用电量不超过一百万度，以 ξ 表示每天的耗电率（即用电量除以一百万度），它具有分布密度为

$$p(x) = \begin{cases} 12x(1-x)^2 & 0 < x < 1 \\ 0 & \text{其它} \end{cases}$$

若该城市每天的供电量仅有 80 万度，求供电量不够需要的概率是多少？如每天供电量 90 万度又是怎样呢？

$$\text{解： } P(\xi > 0.8) = \int_{0.8}^1 12x(1-x)^2 dx = 0.0272$$

$$P(\xi > 0.9) = \int_{0.9}^1 12x(1-x)^2 dx = 0.0037$$

因此，若该城市每天的供电量为 80 万度，供电量不够需要的概率为 0.0272，若每天的供电量为 90 万度，则供电量不够需要的概率为 0.0037。

3.14 设随机变数 ξ 服从 $(0, 5)$ 上的均匀分布，求方程

$$4x^2 + 4\xi x + \xi + 2 = 0$$

有实根的概率。

解：当且仅当

$$(4\xi)^2 - 16(\xi + 2) \geq 0 \quad (1)$$

成立时，方程 $4x^2 + 4\xi x + \xi + 2 = 0$ 有实根。不等式 (1) 的解为： $\xi \geq 2$ 或 $\xi \leq -1$ 。

因此，该方程有实根的概率

$$p = P(\xi \geq 2) + P(\xi \leq -1) = P(\xi \geq 2) = \int_2^5 \frac{1}{5} dx = \frac{3}{5}.$$

3.17 某种电池的寿命 ξ 服从正态 $N(a, \sigma^2)$ 分布，其中 $a = 300$ （小时）， $\sigma = 35$ （小时）

(1) 求电池寿命在 250 小时以上的概率；

(2) 求 x ，使寿命在 $a-x$ 与 $a+x$ 之间的概率不小于 0.9。

$$\text{解：(1) } P(\xi > 250) = P\left(\frac{\xi - 300}{35} > -1.43\right)$$

$$= P\left(\frac{\xi - 300}{35} < 1.43\right) = \Phi(1.43) \approx 0.9236;$$

$$(2) P(a-x < \xi < a+x) = P\left(-\frac{x}{35} < \frac{\xi - 300}{35} < \frac{x}{35}\right)$$

$$= \Phi\left(\frac{x}{35}\right) - \Phi\left(-\frac{x}{35}\right) = 2\Phi\left(\frac{x}{35}\right) - 1 \geq 0.9$$

即

$$\Phi\left(\frac{x}{35}\right) \geq 0.95$$

所以

$$\frac{x}{35} \geq 1.65$$

即

$$x \geq 57.75$$

3.18 设 $\Phi(x)$ 为 $N(0,1)$ 分布的分布函数, 证明当 $x > 0$ 时, 有

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \cdot \frac{1}{x} > 1 - \Phi(x) > \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left(\frac{1}{x} - \frac{1}{x^3} \right)$$

$$\text{证: } 1 - \Phi(x) = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{y^2}{2}} dy$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \cdot \frac{1}{x} - \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \frac{1}{y^2} e^{-\frac{y^2}{2}} dy$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left(\frac{1}{x} - \frac{1}{x^3} \right) + \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \frac{3}{y^4} e^{-\frac{y^2}{2}} dy$$

所以

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \cdot \frac{1}{x} > 1 - \Phi(x) > \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left(\frac{1}{x} - \frac{1}{x^3} \right).$$

3.21 证明: 二元函数

$$F(x, y) = \begin{cases} 1 & x + y > 0 \\ 0 & x + y \leq 0 \end{cases}$$

对每个变元单调非降, 左连续, 且 $F(-\infty, y) = F(x, -\infty) = 0$, $F(-\infty, +\infty) = 0$, 但是

$F(x, y)$ 并不是一个分布函数。

证: (1) 设 $\Delta x > 0$,

若 $x + y > 0$, 由于 $x + \Delta x + y > 0$, 所以 $F(x, y) = F(x + \Delta x, y) = 1$,

若 $x + y \leq 0$, 则 $F(x, y) = 0$ 。当 $x + \Delta x + y \leq 0$ 时, $F(x + \Delta x, y) = 0$;

当 $x + \Delta x + y > 0$ 时, $F(x + \Delta x, y) = 1$ 。所以 $F(x, y) \leq F(x + \Delta x, y)$ 。

可见, $F(x, y)$ 对 x 非降。同理, $F(x, y)$ 对 y 非降。

(2) $x + y \leq 0$ 时

$$\lim_{\Delta x \downarrow 0} F(x - \Delta x, y) = \lim_{\Delta y \downarrow 0} F(x, y - \Delta y) = 0 = F(x, y),$$

$x + y > 0$ 时,

$$\lim_{\Delta x \downarrow 0} F(x - \Delta x, y) = \lim_{\Delta y \downarrow 0} F(x, y - \Delta y) = 1 = F(x, y),$$

所以 $F(x, y)$ 对 x 、 y 左连续。

$$(3) \quad F(-\infty, y) = F(x, -\infty) = 0, \quad F(+\infty, +\infty) = 0.$$

$$(4) \quad P(0 \leq \xi < 2, 0 \leq \eta < 2) = F(2, 2) - F(2, 0) - F(0, 2) + F(0, 0) = -1,$$

所以 $F(x, y)$ 不是一个分布函数。

3.23 设二维随机变数 (ξ, η) 的密度

$$p(x, y) = \begin{cases} \frac{1}{2} \sin(x + y) & 0 \leq x \leq \frac{\pi}{2}, 0 \leq y \leq \frac{\pi}{2} \\ 0 & \text{其它} \end{cases}$$

求 (ξ, η) 的分布函数。

解：当 $0 \leq x \leq \frac{\pi}{2}$, $0 \leq y \leq \frac{\pi}{2}$ 时,

$$F(x, y) = P(\xi < x, \eta < y)$$

$$\begin{aligned} &= \int_0^x \int_0^y \frac{1}{2} \sin(t + s) ds dt \\ &= \frac{1}{2} \int_0^x [\cot - \cos(t + y)] dt \\ &= \frac{1}{2} [\sin x + \sin y - \sin(x + y)], \text{ 所以} \end{aligned}$$

$$F(x, y) = \begin{cases} 0 & (x < 0) \cup (y < 0) \\ \frac{1}{2} [\sin x + \sin y - \sin(x + y)] & 0 \leq x \leq \frac{\pi}{2}, 0 \leq y \leq \frac{\pi}{2} \\ \frac{1}{2} (\sin x + 1 - \cos x) & 0 \leq x \leq \frac{\pi}{2}, y > \frac{\pi}{2} \\ \frac{1}{2} (1 + \sin y - \cos y) & x > \frac{\pi}{2}, 0 \leq y \leq \frac{\pi}{2} \\ 1 & x > \frac{\pi}{2}, y > \frac{\pi}{2} \end{cases}$$

3.24 设二维随机变数 (ξ, η) 的联合密度为

$$p(x, y) = \begin{cases} ke^{-3x-4y} & x > 0, y > 0 \\ 0 & \text{其它} \end{cases}$$

- (1) 求常数 k ;
 (2) 求相应的分布函数;
 (3) 求 $P(0 < \xi < 1, 0 < \eta < 2)$ 。

解: (1) $\int_0^\infty \int_0^\infty ke^{-3x-4y} dx dy = \frac{k}{4} \int_0^\infty e^{-3x} dx = \frac{k}{12},$

所以 $k = 12$;

(2) $x > 0, y > 0$ 时,

$$\begin{aligned} F(x, y) &= \int_0^x \int_y^\infty 12e^{-3t-4s} dt ds = 12 \left(\int_0^x e^{-3t} dt \right) \left(\int_y^\infty e^{-4s} ds \right) \\ &= (1 - e^{-3x})(1 - e^{-4y}), \text{ 所以} \end{aligned}$$

$$F(x, y) = \begin{cases} (1 - e^{-3x})(1 - e^{-4y}) & x > 0, y > 0 \\ 0 & \text{其它} \end{cases}$$

(3) $P(0 < \xi < 1, 0 < \eta < 2)$

$$= F(1, 2) - F(0, 2) - F(1, 0) + F(0, 0)$$

$$= 1 - e^{-3} - e^{-8} + e^{-11}。$$

3. 25 设二维随机变数 (ξ, η) 有密度函数

$$p(x, y) = \frac{A}{\pi^2(16 + x^2)(25 + y^2)}$$

求常数 A 及 (ξ, η) 的密度函数。

解:

$$\begin{aligned} & \int_{-\infty}^\infty \int_{-\infty}^\infty p(x, y) dx dy \\ &= \int_{-\infty}^\infty \int_{-\infty}^\infty \frac{A}{\pi^2(16 + x^2)(25 + y^2)} dx dy \\ &= \frac{4A}{\pi^2} \int_0^\infty \frac{dx}{16 + x^2} \int_0^\infty \frac{dy}{25 + y^2} = \frac{A}{20} = 1 \end{aligned}$$

所以, $A = 20$;

$$\begin{aligned}
 F(x, y) &= \int_{-\infty}^x \int_{-\infty}^y p(t, s) dt ds \\
 &= \frac{20}{\pi^2} \int_{-\infty}^x \int_{-\infty}^y \frac{dt ds}{(16+t^2)(25+s^2)} \\
 &= \frac{20}{\pi^2} \left(\int_{-\infty}^x \frac{dt}{16+t^2} \right) \left(\int_{-\infty}^y \frac{ds}{25+s^2} \right) \\
 &= \frac{1}{\pi^2} \left(\arctg \frac{x}{4} + \frac{\pi}{2} \right) \left(\arctg \frac{y}{5} + \frac{\pi}{2} \right)
 \end{aligned}$$

3.26 设二维随机变数 (ξ, η) 的密度函数为

$$p(x, y) = \begin{cases} 4xy & 0 < x < 1, 0 < y < 1 \\ 0 & \text{其它} \end{cases}$$

求 (1) $P(0 < \xi < \frac{1}{2}, \frac{1}{4} < \eta < 1)$; (2) $P(\xi = \eta)$; (3) $P(\xi < \eta)$; (4) $P(\xi \leq \eta)$ 。

解:

$$(1) \quad P(0 < \xi < \frac{1}{2}, \frac{1}{4} < \eta < 1) = \int_0^{\frac{1}{2}} \int_{\frac{1}{4}}^1 4xy dx dy = 4 \int_0^{\frac{1}{2}} x dx \int_{\frac{1}{4}}^1 y dy = \frac{15}{64};$$

$$(2) \quad P(\xi = \eta) = \int_{x=y} 4xy dx dy = 0;$$

$$(3) \quad P(\xi < \eta) = \iint_{x < y} 4xy dx dy = \int_0^1 \int_x^1 4xy dy dx = \int_0^1 2(x - x^2) dx = \frac{1}{2};$$

$$(4) \quad P(\xi \leq \eta) = \frac{1}{2}$$

3.28 设 (ξ, η) 的密度函数为

$$p(x, y) = \begin{cases} \frac{1}{2} & 0 \leq x \leq 1, 0 \leq y \leq 2 \\ 0 & \text{其它} \end{cases}$$

求 ξ 与 η 中至少有一个小于 $\frac{1}{2}$ 的概率。

解:

$$\begin{aligned}
 P[(\xi < \frac{1}{2}) \cup (\eta < \frac{1}{2})] &= 1 - P(\xi \geq \frac{1}{2}, \eta \geq \frac{1}{2}) \\
 &= 1 - \int_{\frac{1}{2}}^1 \int_{\frac{1}{2}}^2 p(x, y) dx dy = 1 - \int_{\frac{1}{2}}^1 \int_{\frac{1}{2}}^2 \frac{1}{2} dx dy = \frac{5}{8}
 \end{aligned}$$

3.30 一个电子器件包含两个主要组件，分别以 ξ 和 η 表示这两个组件的寿命（以小时计），

设 (ξ, η) 的分布函数为

$$F(x, y) = \begin{cases} 1 - e^{-0.01x} - e^{-0.01y} + e^{-0.01(x+y)} & x \geq 0, y \geq 0 \\ 0 & \text{其它} \end{cases}$$

求两个组件的寿命都超过 120 的概率。

解：

$$\begin{aligned} P(\xi > 120, \eta > 120) &= 1 - P[(\xi \leq 120) \cup (\eta \leq 120)] \\ &= 1 - P(\xi \leq 120) - P(\eta \leq 120) + P(\xi \leq 120, \eta \leq 120) \\ &= 1 - F(120 + 0, \infty) - F(\infty, 120 + 0) + F(120 + 0, 120 + 0) \\ &= 1 - (1 - e^{-1.2}) - (1 - e^{-1.2}) + (1 - 2e^{-1.2} + e^{-2.4}) \\ &= e^{-2.4} \approx 0.09 \end{aligned}$$

3.31 设 $p_1(x), p_2(x)$ 都是一维分布的密度函数，为使

$$p(x, y) = p_1(x)p_2(y) + h(x, y)$$

成为一个二维分布的密度函数，问其中的 $h(x, y)$ 必需且只需满足什么条件？

解：若 $p(x, y)$ 为二维分布的密度函数，则

$$p(x, y) \geq 0, \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) dx dy = 1$$

所以条件 (1) $h(x, y) \leq p_1(x)p_2(y)$; (2) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) dx dy = 0$ 得到满足。

反之，若条件 (1), (2) 满足，则

$$p(x, y) \geq 0, \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) dx dy = 1$$

$p(x, y)$ 为二维分布的密度函数。

因此，为使 $p(x, y)$ 成为二维分布的密度函数， $h(x, y)$ 必需且只需满足条件 (1) 和 (2)。

3.32 设二维随机变数 (ξ, η) 具有下列密度函数，求边际分布。

$$(1) \quad p(x, y) = \begin{cases} \frac{2e^{-y+1}}{x^3} & x > 1, y > 1 \\ 0 & \text{其它} \end{cases}$$

$$(2) \quad p(x, y) = \begin{cases} \frac{1}{\pi} e^{-\frac{1}{2}(x^2+y^2)} & x > 0, y \leq 0 \text{ 或 } x \leq 0, y > 0 \\ 0 & \text{其它} \end{cases}$$

$$(3) \quad p(x, y) = \begin{cases} \frac{1}{\Gamma(k_1)\Gamma(k_2)} x^{k_1-1} (y-x)^{k_2-1} e^{-y} & 0 < x < y \\ 0 & \text{其它} \end{cases}$$

解：(1) $p_{\xi}(x) = \int_1^{\infty} \frac{2e^{-y+1}}{x^3} dy = \frac{2}{x^3}, (x > 1) \quad p_{\xi}(x) = 0, (x \leq 1)$

$$p_{\xi}(x) = \int_1^{\infty} \frac{2e^{-y+1}}{x^3} dx = e^{-y+1}, (y > 1) \quad p_{\xi}(x) = 0, (y \leq 1)$$

(2) $x > 0$ 时,

$$p_{\xi}(x) = \int_{-\infty}^0 \frac{1}{\pi} e^{-\frac{1}{2}(x^2+y^2)} dy = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$x \leq 0$ 时,

$$p_{\xi}(x) = \int_0^{\infty} \frac{1}{\pi} e^{-\frac{1}{2}(x^2+y^2)} dy = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

所以, $p_{\xi}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ 。同理, $p_{\xi}(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$ 。

$$(3) \quad p_{\xi}(x) = \frac{x^{k_1-1}}{\Gamma(k_1)\Gamma(k_2)} \int_x^{\infty} (y-x)^{k_2-1} e^{-y} dy = \frac{1}{\Gamma(k_1)} x^{k_2-1} e^{-x}, (x > 0)$$

$$p_{\xi}(x) = 0, (x \leq 0)$$

$$p_{\eta}(y) = \frac{e^{-y}}{\Gamma(k_1)\Gamma(k_2)} \int_0^y x^{k_1-1} (y-x)^{k_2-1} dx = \frac{1}{\Gamma(k_1+k_2)} y^{k_1+k_2-1}, (y > 0)$$

$$p_{\eta}(y) = 0, (y \leq 0)$$

3.34 证明：若随机变数 ξ 只取一个值 a ，则 ξ 与任意的随机变数 η 独立。

证： ξ 的分布函数为

$$F_{\xi}(x) = \begin{cases} 0 & x \leq a \\ 1 & x > a \end{cases}$$

设 η 的分布函数、 (ξ, η) 的联合分布函数分别为 $F_{\eta}(y), F(x, y)$ 。

当 $x \leq a$ 时, $F(x, y) = P(\xi < x, \eta < y) = 0 = F_{\xi}(x)F_{\eta}(y)$ 。当 $x > a$ 时,

$F(x, y) = P(\xi < x, \eta < y) = P(\eta < y) = F_{\xi}(x)F_{\eta}(y)$ 。所以, 对任意实数 x, y , 都有

$F(x, y) = F_{\xi}(x)F_{\eta}(y)$, 故 ξ 与 η 相互独立。

3.35 证明：若随机变数 ξ 与自己独立, 则必有常数 c , 使 $P(\xi = c) = 1$ 。

证：由于 $P(\xi < x) = P(\xi < x, \xi < x) = P(\xi < x)P(\xi < x)$, 所以 $F(x) = [F(x)]^2$,

$F(x) = 0$ 或 1 。由于 $F(-\infty) = 0, F(+\infty) = 1$, $F(x)$ 非降、左连续, 所以必有常数 c , 使得

$$F(x) = \begin{cases} 0 & x \leq c \\ 1 & x > c \end{cases}$$

故 $P(\xi = c) = 1$ 。

3.36 设二维随机变量 (ξ, η) 的密度函数为

$$p(x, y) = \begin{cases} \frac{1}{\pi} & x^2 + y^2 \leq 1 \\ 0 & \text{其它} \end{cases}$$

问 ξ 与 η 是否独立? 是否不相关?

$$\text{解: } p_{\xi}(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{dy}{\pi} = \frac{2\sqrt{1-x^2}}{\pi}, (|x| \leq 1); p_{\xi}(x) = 0, (|x| > 1)。$$

$$\text{同理, } p_{\eta}(y) = \frac{2\sqrt{1-y^2}}{\pi}, (|y| \leq 1); p_{\eta}(y) = 0, (|y| > 1)。$$

由于 $p(x, y) \neq p_{\xi}(x)p_{\eta}(y)$, 所以 ξ 与 η 不相互独立。

又因 $p(x, y), p_{\xi}(x), p_{\eta}(y)$ 关于 x 或关于 y 都是偶函数, 因而 $E\xi = E\eta = E(\xi\eta) = 0$,

故 $\text{cov}(\xi, \eta) = 0$, ξ 与 η 不相关。

3.41 设某类电子管的寿命(以小时计)具有如下分布密度:

$$p(x) = \begin{cases} \frac{100}{x^2} & x > 100 \\ 0 & x \leq 100 \end{cases}$$

一台电子管收音机在开初使用的 150 小时中, 三个这类管子没有一个要替换的概率是多少? 三个这类管子全部要替换的概率又是多少? (假设这三个管子的寿命分布是相互独立的)

解: 设这类电子管的寿命为 ξ , 则

$$P(\xi > 150) = \int_{150}^{\infty} \frac{100}{x^2} dx = \frac{2}{3}$$

所以三个这类管子没有一个要替换的概率为 $(\frac{2}{3})^3 = \frac{8}{27}$; 三个这类管子全部要替换的概率是 $(1 - \frac{2}{3})^3 = \frac{1}{27}$ 。

3.44 对球的直径作近似测量, 设其值均匀分布在区间 $[a, b]$ 内, 求球体积的密度函数。

解：设球的直径为 ξ ，则其体积为 $\eta = \frac{1}{6}\pi\xi^3$ 。 $y = \frac{1}{6}\pi x^3$ 的反函数 $x = \sqrt[3]{6y/\pi}$, $dx = 2/\sqrt[3]{36\pi y^2} dy$ 。由 ξ 的密度函数 $p_\xi(x) = 1/(b-a)$, $a \leq x \leq b$, 得 η 的密度函数为

$$p_\eta(y) = \begin{cases} \frac{2}{(b-a) \cdot \sqrt[3]{36\pi y^2}} & \frac{\pi}{6} a^3 \leq y \leq \frac{\pi}{6} b^3, \\ 0 & \text{其它。} \end{cases}$$

3.45 设随机变数 ξ 服从 $N(0,1)$ 分布, 求 $|\xi|$ 的分布密度。

解：在 $x \geq 0$ 时,

$$P(|\xi| < x) = P(-x < \xi < x) = \int_{-x}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt。$$

所以 $|\xi|$ 的分布密度

$$p_{|\xi|}(x) = \sqrt{2/\pi} \cdot e^{-x^2/2}, (x \geq 0); p_{|\xi|}(x) = 0, (x < 0)。$$

3.46 设随机变数 ξ 服从 $N(a, \sigma^2)$ 分布, 求 e^ξ 的分布密度。

解： $y = e^x$ 的反函数 $x = \ln y$, $dx = 1/y \cdot dy$ 。由 ξ 服从 $N(a, \sigma^2)$ 分布, 推得 $\eta = e^\xi$ 的分布密度为

$$p_\eta(y) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma y} \cdot \exp\left\{-\frac{1}{2\sigma^2}(\ln y - a)^2\right\} & y > 0, \\ 0 & y \leq 0. \end{cases}$$

3.47 随机变数 ξ 在任一有限区间 $[a, b]$ 上的概率均大于 0 (例如正态分布等), 其分布函数为 $F_\xi(x)$, 又 η 服从 $[0,1]$ 上的均匀分布。证明 $\zeta = F_\xi^{-1}(\eta)$ 的分布函数与 ξ 的分布函数相同。

解：因为 ξ 在任一有限区间 $[a, b]$ 上的概率均大于 0, 所以 $F_\xi(x)$ 是严格上升函数。由于 $[0,1]$ 上的均匀分布, 所以 ζ 的分布函数 $F_\zeta(x) = P(\xi < x) = P(F_\xi^{-1}(\eta) < x) = P(\eta < F_\xi(x)) = F_\xi(x)$, 对任意的 x 都成立。所以 ζ 与 ξ 的分布函数相同。

3.48 设随机变量 ξ 与 η 独立, 求 $\xi + \eta$ 的分布密度。若 (1) ξ 与 η 分布服从 (a, b) 及 (α, β) 上的均匀分布, 且 $a < \alpha < b < \beta$; (2) ξ 与 η 分别服从 $(-a, 0)$ 及 $(0, a)$ 上的均匀分

布, $a > 0$ 。

解 (1) $p_{\xi}(x) = 1/(b-a), a < x < b; p_{\xi}(x) = 0$, 其它。

$$p_{\eta}(x) = 1/(\beta - \alpha), \alpha < x < \beta; p_{\eta}(y) = 0, \text{ 其它。}$$

$$\begin{aligned} p_{\xi+\eta}(x) &= \int_{-\infty}^{\infty} p_{\xi}(x-y) \cdot p_{\eta}(y) dy \\ &= \int_{\max(x-b, \alpha)}^{\min(x-a, \beta)} \frac{1}{(b-a)(\beta-\alpha)} dy \end{aligned}$$

$= [\min(x-a, \beta) - \max(x-b, \alpha)] / [(b-a)(\beta-\alpha)], a+\alpha < x < b+\beta; p_{\xi+\eta}(x) = 0$, 其它。

(2) $p_{\xi}(x) = 1/a, -a < x < 0; p_{\xi}(x) = 0$, 其它,

$$p_{\eta}(x) = 1/a, 0 < x < a, p_{\eta}(x) = 0, \text{ 其它。}$$

$$\begin{aligned} p_{\xi+\eta}(x) &= \int_{-\infty}^{\infty} p_{\xi}(x-y) \cdot p_{\eta}(y) dy = \int_{\max(x, 0)}^{\min(x+a, a)} 1/a^2 dy \\ &= [\min(x+a, a) - \max(x, 0)] / a^2 \\ &= \frac{a-|x|}{a^2}, -a < x < a, p_{\xi+\eta}(x) = 0, \text{ 其它} \end{aligned}$$

3.49 设随机变量 ξ 与 η 独立, 服从相同的拉普拉斯分布, 其密度函数为

$$p(x) = \frac{1}{2a} \cdot e^{-|x|/a}, \quad (a > 0)$$

求 $\xi + \eta$ 的密度函数。

解:
$$p_{\xi}(x) = p_{\eta}(x) = \frac{1}{2a} \cdot e^{-|x|/a},$$

$$p_{\xi+\eta}(x) = \int_{-\infty}^{\infty} p_{\xi}(x-y) \cdot p_{\eta}(y) dy,$$

当 $x \geq 0$ 时,

$$\begin{aligned} p_{\xi+\eta}(x) &= \int_{-\infty}^{\infty} \frac{1}{4a^2} \exp\left\{-\frac{|x-y|+|y|}{a}\right\} dy \\ &= \frac{1}{4a^2} \left[\int_{-\infty}^0 e^{-\frac{x-y-y}{a}} dy + \int_0^x e^{-\frac{x-y+y}{a}} dy + \int_x^{\infty} e^{-\frac{y-x+y}{a}} dy \right] \\ &= \frac{1}{4a} \left(1 + \frac{x}{a}\right) e^{-x/a} \end{aligned}$$

当 $x < 0$ 时,

$$p_{\xi+\eta}(x) = \frac{1}{4a^2} \left[\int_{-\infty}^x e^{-\frac{x-y-y}{a}} dy + \int_x^0 e^{-\frac{y-x-y}{a}} dy + \int_0^{\infty} e^{-\frac{y-x+y}{a}} dy \right] = \frac{1}{4a} \left(1 - \frac{x}{a}\right) e^{-x/a}$$

所以

$$p_{\xi+\eta}(x) = \frac{1}{4a^2} (a + |x|) e^{-|x|/a}$$

3.50 设随机变量 ξ 与 η 独立, 服从相同的柯西分布, 其密度函数为

$$p(x) = \frac{1}{\pi(1+x^2)}$$

证明: $\varsigma = \frac{1}{2}(\xi + \eta)$ 也服从同一分布。

证:

$$\begin{aligned} p_{\xi+\eta}(y) &= \int_{-\infty}^{\infty} \frac{1}{\pi^2} \frac{1}{1+x^2} \frac{1}{1+(y-x)^2} dx \\ &= \frac{1}{\pi^2 y(y^2+4)} \int_{-\infty}^{\infty} \left[\frac{2x+y}{x^2+1} - \frac{2(x-y)-y}{(x-y)^2+1} \right] dx \\ &= \frac{1}{\pi^2 y(y^2+4)} [\ln(x^2+1) + y \arctg x - \ln((x-y)^2+1) + y \arctg(x-y)] \Big|_{-\infty}^{\infty} \\ &= \frac{2}{\pi(y^2+4)} \end{aligned}$$

所以

$$p_{\frac{1}{2}(\xi+\eta)}(z) = \frac{2}{\pi[(2z)^2+4]} = \frac{1}{\pi(1+z^2)}$$

即 $\varsigma = \frac{1}{2}(\xi + \eta)$ 也服从相同的柯西分布。

3.51 设随机变量 ξ 与 η 独立, 分别具有密度函数

$$p_{\xi}(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

$$p_{\eta}(x) = \begin{cases} \mu e^{-\mu x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

(其中 $\lambda > 0, \mu > 0$), 求 $\xi + \eta$ 的分布密度。

解: $x > 0$ 时,

$$\begin{aligned}
 p_{\xi+\eta}(x) &= \int_0^x \mu e^{-\mu(x-y)} \lambda e^{-\lambda y} dy \\
 &= \mu \lambda e^{-\mu x} \int_0^x e^{-(\lambda-\mu)y} dy \\
 &= \begin{cases} \frac{\mu \lambda}{(\lambda-\mu)} [e^{-\mu x} e^{-\lambda x}], & \lambda \neq \mu \\ \lambda^2 x e^{-\lambda x}, & \lambda = \mu \end{cases}
 \end{aligned}$$

$x \leq 0$ 时,

$$p_{\xi+\eta}(x) = 0$$

3.53 设随机变量 ξ 与 η 独立, 都服从(0,1)上的均匀分布, 求 $|\xi - \eta|$ 的分布。

解: $-\eta$ 服从(-1,0)上的均匀分布, 据 3.48(2)知,

$$p_{\xi-\eta}(x) = [\min(x+1, 1) - \max(x, 0)] = \begin{cases} x+1 & -1 < x \leq 0 \\ 1-x & 0 < x < 1 \end{cases}$$

在 $0 < x < 1$ 时, $|\xi - \eta|$ 的分布函数

$$\begin{aligned}
 F(x) &= P(|\xi - \eta| < x) = P(-x < \xi - \eta < x) \\
 &= \int_{-x}^0 (t+1) dt + \int_0^x (1-t) dt = 2x - x^2
 \end{aligned}$$

所以 $|\xi - \eta|$ 的分布密度为

$$p_{|\xi-\eta|}(x) = \begin{cases} 2(1-x) & 0 \leq x < 1 \\ 0 & \text{其它} \end{cases}$$

3.54 设随机变量 ξ 与 η 独立, 分别服从参数为 λ 与 μ 的指数分布, 求 $\xi - \eta$ 的分布密度。

解: 由 $p_{\eta}(x) = \mu e^{-\mu x}, x > 0$ 得 $p_{-\eta}(x) = \mu e^{\mu x}, x < 0$, 所以

$$p_{\xi-\eta}(x) = \int_{-\infty}^{\infty} p_{\xi}(y) p_{-\eta}(x-y) dy$$

在 $x \leq 0$ 时,

$$p_{\xi-\eta}(x) = \int_0^{\infty} \lambda e^{-\lambda y} \mu e^{\mu(x-y)} dy = \frac{\lambda \mu e^{\mu x}}{(\lambda + \mu)}$$

在 $x > 0$ 时,

$$p_{\xi-\eta}(x) = \int_x^{\infty} \lambda e^{-\lambda y} \mu e^{\mu(x-y)} dy = \frac{\lambda \mu e^{-\lambda x}}{(\lambda + \mu)}$$

所以

$$p_{\xi+\eta}(x) = \begin{cases} \frac{\lambda\mu e^{\mu x}}{(\lambda+\mu)} & x \leq 0 \\ \frac{\lambda\mu e^{-\lambda x}}{(\lambda+\mu)} & x > 0 \end{cases}$$

3.56 设随机变量 ξ 与 η 独立, 且分别具有密度函数为

$$p_{\xi}(x) = \begin{cases} \frac{1}{\pi\sqrt{1-x^2}} & |x| < 1 \\ 0 & |x| \geq 1 \end{cases}$$

$$p_{\eta}(y) = \begin{cases} xe^{-x^2/2} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

证明 $\xi\eta$ 服从 $N(0,1)$ 分布。

证: 由 $p_{\eta}(x) = xe^{-x^2/2}, x > 0$ 得 $p_{\xi/\eta}(x) = x^{-3}e^{-1/2x^2}, x > 0$ 。故

$$p_{\xi\eta}(y) = p_{\xi/\eta}(y) = \int_{-\infty}^{\infty} |x| p_{\xi}(yx) p_{\eta}(x) dx$$

令 $1/2x^2 = u + y^2/2$, 则

$$p_{\xi\eta}(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \int_0^{\infty} u^{-1/2} e^{-u} du = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

所以 $\xi\eta$ 服从 $N(0,1)$ 分布。

3.58 设随机变量 ξ 与 η 独立, 都服从 $(0, a)$ 上的均匀分布, 求 ξ/η 的密度函数。

$$\text{解: } p_{\xi/\eta}(x) = \int_{-\infty}^{\infty} p_{\xi}(xz) p_{\eta}(z) |z| dz = \frac{1}{a} \int_0^{\infty} zp_{\xi}(xz) dz$$

当 $0 < x \leq 1$ 时,

$$p_{\xi/\eta}(x) = \frac{1}{a^2} \int_0^a z dz = \frac{1}{2}$$

当 $x > 1$ 时

$$p_{\xi/\eta}(x) = \frac{1}{a^2} \int_0^{1/x} z dz = \frac{1}{2x^2}$$

所以 ξ/η 的密度函数为

$$p_{\xi/\eta}(x) = \begin{cases} 0 & x \leq 0 \\ 1/2 & 0 < x \leq 1 \\ 1/2x^2 & x > 1 \end{cases}$$

3.59 设随机变量 ξ 与 η 独立，都服从参数为 λ 的指数分布，求 ξ/η 的密度函数。

解：在 $x \geq 0$ 时，

$$\begin{aligned} p_{\xi/\eta}(x) &= \int_{-\infty}^{\infty} p_{\xi}(xy) p_{\eta}(y) |y| dy \\ &= \int_0^{\infty} \lambda^2 e^{-\lambda xy} e^{-\lambda y} y dy = \frac{1}{(x+1)^2} \end{aligned}$$

在 $x < 0$ 时， $p_{\xi/\eta}(x) = 0$ 。

3.60 设二维随机变量 (ξ, η) 的联合分布密度为

$$p(x, y) = \begin{cases} \frac{1+xy}{4} & |x| < 1, |y| < 1 \\ 0 & \text{其它} \end{cases}$$

证明： ξ 与 η 不独立，但 ξ^2 与 η^2 独立。

证：由于 $p(x, y) \neq p_{\xi}(x) p_{\eta}(y)$ ，所以 ξ 与 η 不独立。由于

$$\begin{aligned} P(\xi^2 < x) &= \begin{cases} 1 & x > 1 \\ \int_{-\sqrt{x}}^{\sqrt{x}} \left(\int_{-1}^1 \frac{1+ty}{4} dy \right) dt = \sqrt{x} & 0 < x \leq 1 \\ 0 & x \leq 0 \end{cases} \\ P(\eta^2 < y) &= \begin{cases} 1 & y > 1 \\ \int_{-\sqrt{y}}^{\sqrt{y}} \left(\int_{-1}^1 \frac{1+tx}{4} dx \right) dt = \sqrt{y} & 0 < y \leq 1 \\ 0 & y \leq 0 \end{cases} \\ P(\xi^2 < x, \eta^2 < y) &= \begin{cases} 1 & x, y > 1 \\ \sqrt{x} & 0 < x \leq 1, y > 1 \\ \sqrt{y} & x > 1, 0 < y \leq 1 \\ \sqrt{xy} & 0 < x, y \leq 1 \\ 0 & \text{其它} \end{cases} \end{aligned}$$

所以对一切的 x, y ，都有 $P(\xi^2 < x, \eta^2 < y) = P(\xi^2 < x) P(\eta^2 < y)$ ，故 ξ^2 与 η^2 相互独立。

3.61 设随机变量 ξ 具有密度函数

$$p(x) = \begin{cases} \frac{2}{\pi} \cos^2 x & -\frac{\pi}{2} \leq x \leq \frac{\pi}{2} \\ 0 & \text{其它} \end{cases}$$

求 $E\xi, D\xi$ 。

解: $E\xi = \int_{-\pi/2}^{\pi/2} x \frac{2}{\pi} \cos^2 x dx = 0$

$$D\xi = E\xi^2 = \int_{-\pi/2}^{\pi/2} x^2 \frac{2}{\pi} \cos^2 x dx = \frac{\pi^2}{12} - \frac{1}{2}$$

3.62 设随机变量 ξ 具有密度函数

$$p(x) = \begin{cases} x & 0 < x \leq 1 \\ 2-x & 1 < x < 2 \\ 0 & \text{其它} \end{cases}$$

求 $E\xi$ 及 $D\xi$ 。

解 $E\xi = \int_0^1 x^2 dx + \int_1^2 x(2-x) dx = 1,$

$$E\xi^2 = \int_0^1 x^3 dx + \int_1^2 x^2(2-x) dx = 7/6,$$

$$D\xi = E\xi^2 - (E\xi)^2 = 1/6.$$

3.63 设随机变量 ξ 的分布函数为

$$F(x) = \begin{cases} 0 & x < -1 \\ a + b \arcsin x & -1 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

试确定常数 (a, b) , 并求 $E\xi$ 与 $D\xi$ 。

解: 由分布函数的左连续性,

$$\begin{cases} a + b \cdot \arcsin 1 = 1, \\ a + b \cdot \arcsin 0 = 0, \end{cases}$$

故 $a = 1/2, b = 1/\pi$ 。

$$E\xi = \int_{-1}^1 x \cdot d\left(\frac{1}{2} + \frac{1}{\pi} \arcsin x\right)$$

$$= \int_{-1}^1 \frac{x}{\pi \sqrt{1-x^2}} dx = 0,$$

$$D\xi = E\xi = \int_{-1}^1 \frac{x}{\pi\sqrt{1-x^2}} dx = \frac{2}{\pi} \int_0^1 \frac{x^2 dx}{\sqrt{1-x^2}} = \frac{2}{\pi} \int_0^{\pi/2} \sin^2 t dt = 1/2。$$

3.64 随机变量 ξ 具有密度函数

$$p(x) = \begin{cases} A \cdot x^\alpha \cdot e^{-x/\beta}, & x < 0 \\ 0, & x \leq 0 \end{cases}$$

其中 $\alpha > 1, \beta > 0$, 求常数 $A, E\xi$ 及 $D\xi$ 。

$$\begin{aligned} \text{解: } 1 &= \int_0^\infty A \cdot x^\alpha \cdot e^{-x/\beta} dx = A \cdot \int_0^\beta \beta^{\alpha+1} y^\alpha e^{-y} dy \\ &= A\beta^{\alpha+1} T(\alpha+1), \end{aligned}$$

故

$$A = \frac{1}{\beta^{\alpha+1} \cdot T(\alpha+1)}。$$

$$E\xi = \int_0^\infty A \cdot x^{\alpha+1} \cdot e^{-x/\beta} dx = A \cdot \beta^{\alpha+2} \cdot T(\alpha+2) = (\alpha+1)\beta,$$

$$E\xi^2 = \int_0^\infty A \cdot x^{\alpha+2} \cdot e^{-x/\beta} dx = A \cdot \beta^{\alpha+3} \cdot T(\alpha+3)$$

$$= (\alpha+1)(\alpha+2)\beta^2$$

$$D\xi = E\xi^2 - (E\xi)^2 = (\alpha+1)\beta^2$$

3.66 设随机变量 ξ 服从 $(-\frac{1}{2}, \frac{1}{2})$ 上的均匀分布, 求 $\eta = \sin \pi \xi$ 的数学期望与方差。

$$\text{解: } E\eta = \int_{-\frac{1}{2}}^{\frac{1}{2}} \sin \pi x dx = 0,$$

$$D\eta = E\eta^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} \sin^2 \pi x dx = 1/2。$$

3.67 地下铁道列车的运行间隔时间为五分钟, 一个旅客在任意时刻进入月台, 求候车时间的数学期望与方差。

解: 设旅客候车时间为 ξ (秒), 则 ξ 服从 $[0, 300]$ 上的均匀分布, 则

$$E\xi = \int_0^{300} \frac{1}{300} \cdot x \cdot dx = 150(\text{秒}),$$

$$E\xi^2 = \int_0^{300} \frac{1}{300} \cdot x^2 \cdot dx = 30000(\text{秒}^2),$$

$$D\xi = 30000 - 150^2 = 7500(\text{秒}^2)。$$

3.71 设 $\xi_1, \xi_2, \dots, \xi_n$ 为正的且独立同分布的随机变量 (分布为连续型或离散型), 证明: 对任意的 $k (1 \leq k \leq n)$, 有

$$E\left(\frac{\xi_1 + \dots + \xi_k}{\xi_1 + \dots + \xi_n}\right) = \frac{k}{n}.$$

证: $\xi_j / \sum_{i=1}^n \xi_i$ 同分布 ($j=1, \dots, n$), 又 $\left|\xi_j / \sum_{i=1}^n \xi_i\right| \leq 1$, 所以 $E\left[\xi_j / \sum_{i=1}^n \xi_i\right]$ 都存在且相等

($j=1, \dots, n$)。由于 $1 = E\left[\sum_{i=1}^n \xi_i / \sum_{i=1}^n \xi_i\right] = n \cdot E\left[\xi_1 / \sum_{i=1}^n \xi_i\right]$, 所以

$$E\left(\frac{\xi_1 + \dots + \xi_k}{\xi_1 + \dots + \xi_n}\right) = k \cdot E\left[\xi_1 / \sum_{i=1}^n \xi_i\right] = \frac{k}{n}.$$

3.72 设 ξ 是非负连续型随机变量, 证明: 对 $x > 0$, 有

$$P(\xi < x) \geq 1 - \frac{E\xi}{x}.$$

证: $P(\xi < x) = \int_0^x p_\xi(t) dt = 1 - \int_x^\infty p_\xi(t) dt$
 $\geq 1 - \int_x^\infty \frac{t}{x} \cdot p_\xi(t) dt \geq 1 - \frac{1}{x} \int_0^\infty t \cdot p_\xi(t) dt$
 $= 1 - \frac{E\xi}{x}.$

3.73 若对连续型随机变量 ξ , 有 $E|\xi|^r < \infty (r > 0)$, 证明有 $P(|\xi| > \varepsilon) \leq \frac{E|\xi|^r}{\varepsilon^r}$.

证: $P(|\xi| > \varepsilon) = \int_{|x| > \varepsilon} p_\xi(x) dx \leq \int_{|x| > \varepsilon} \frac{|x|^r}{\varepsilon^r} \cdot p_\xi(x) dx$
 $\leq \frac{1}{\varepsilon^r} \int_{-\infty}^\infty |x|^r \cdot p_\xi(x) dx = E|\xi|^r / \varepsilon^r.$

3.75 已知随机变量 ξ 与 η 的相关系数为 ρ , 求 $\xi_1 = a\xi + b$ 与 $\eta_1 = c\eta + d$ 的相关系数, 其中 a, b, c, d 均为常数, a, c 皆不为零。

解: $\rho_{\xi_1 \eta_1} = \frac{E[(\xi_1 - E\xi_1) \cdot (\eta_1 - E\eta_1)]}{\sqrt{E(\xi_1 - E\xi_1)^2} \cdot \sqrt{E(\eta_1 - E\eta_1)^2}}$
 $= \frac{ac \cdot \text{cov}(\xi, \eta)}{|a| \cdot \sqrt{D\xi} \cdot |c| \cdot \sqrt{D\eta}}$

$$= \frac{ac}{|ac|} \cdot \rho = \begin{cases} \rho & ac > 0 \\ -\rho & ac < 0 \end{cases}$$

3.81 设随机变量 $\xi_1, \xi_2, \dots, \xi_n$ 中任意两个的相关系数都是 ρ ，试证： $\rho \geq -\frac{1}{n-1}$ 。

$$\begin{aligned} \text{证: } 0 &\leq E\left[\sum_{i=1}^n (\xi_i - E\xi_i)\right]^2 \\ &= \sum_{i=1}^n D\xi_i + 2\rho \sum_{1 \leq i < j \leq n} \sqrt{D\xi_i} \cdot \sqrt{D\xi_j} \\ &\leq \sum_{i=1}^n D\xi_i + \rho \cdot \sum_{1 \leq i < j \leq n} (D\xi_i + D\xi_j) \\ &= \sum_{i=1}^n D\xi_i [1 + \rho(n-1)], \end{aligned}$$

$$\text{故 } 1 + \rho(n-1) \geq 0, \rho \geq -\frac{1}{n-1}.$$

3.84 证明下述不等式（设 ξ, η 都是连续型或离散型随机变量）：

(1) 若 ξ 与 η 都有 $p \geq 1$ 阶矩，则有

$$\begin{aligned} [E|\xi + \eta|^p]^{1/p} &\leq [E|\xi|^p]^{1/p} + [E|\eta|^p]^{1/p} \\ E|\xi + \eta|^p &\leq 2^{p-1} (E|\xi|^p + E|\eta|^p) \end{aligned}$$

(2) 若 ξ 与 η 都具有 $p > 0$ 阶矩，则

$$E|\xi + \eta|^p \leq 2^p (E|\xi|^p + E|\eta|^p)$$

证：(1) $p \geq 1$ 时， $[E|\xi + \eta|^p]^{1/p} \leq [E|\xi|^p]^{1/p} + [E|\eta|^p]^{1/p}$ 即所谓的明可夫斯基不等式，证明略。

在 $p \geq 1$ 时， $|x|^p$ 是 x 的下凸函数，故

$$\left| \frac{x+y}{2} \right|^p \leq \frac{|x|^p + |y|^p}{2}$$

即

$$|x+y|^p \leq 2^{p-1} (|x|^p + |y|^p)$$

故

$$E|\xi + \eta|^p \leq 2^{p-1} (E|\xi|^p + E|\eta|^p)$$

(2) 在 $p > 0$ 时， $|x+y|^p \leq (|x| + |y|)^p \leq 2|x|^p + 2|y|^p = 2^p (|x|^p + |y|^p)$ ，故

$$E|\xi + \eta|^p \leq 2^p (E|\xi|^p + E|\eta|^p)$$

3.88 设二维随机变量 (ξ, η) 的联合分布密度为

$$p(x, y) = \begin{cases} \frac{(n-1)(n-2)}{(1+x+y)^n} & x > 0, y > 0 \\ 0 & \text{其它} \end{cases}$$

其中 $n > 2$ 。求 $\xi = 1$ 条件下 η 的条件分布密度。

$$\text{解: } p_{\xi}(x) = \int_0^{\infty} \frac{(n-1)(n-2)}{(1+x+y)^n} dy = \frac{n-2}{(1+x)^{n-1}}, x > 0. \text{ 故}$$

$$p_{\eta|\xi}(y|1) = \begin{cases} 2^{n-1} (n-1) / (2+y)^n & y > 0 \\ 0 & \text{其它} \end{cases}$$

3.89 设随机变量 ξ 服从 $N(m, \tau^2)$ 分布, 随机变量 η 在 $\xi = x$ 时的条件分布为 $N(x, \sigma^2)$,

求 η 的分布及 ξ 关于 η 的条件分布。

$$\text{解: } p(x, y) = p_{\xi}(x) \cdot p_{\eta|\xi}(y|x) = \frac{1}{2\pi\sigma\tau} \exp \left\{ -\frac{(x-m)^2}{2\tau^2} - \frac{(y-x)^2}{2\sigma^2} \right\}$$

$$\begin{aligned} p_{\eta}(y) &= \int_{-\infty}^{\infty} p(x, y) dx = \frac{1}{2\pi\tau\sigma} \cdot \exp \left\{ -\frac{(y-m)^2}{2(\sigma^2 + \tau^2)} \right\} \cdot \int_{-\infty}^{\infty} \exp \left\{ -\frac{\sigma^2 + \tau^2}{2\tau^2\sigma^2} \cdot \left[x - \frac{m\sigma^2 + y\tau^2}{\sigma^2 + \tau^2} \right]^2 \right\} dx \\ &= \frac{1}{\sqrt{2\pi(\tau^2 + \sigma^2)}} \cdot \exp \left\{ -\frac{(y-m)^2}{2(\sigma^2 + \tau^2)} \right\}, \end{aligned}$$

$$\text{故 } \eta \sim N(m, \sigma^2 + \tau^2) \cdot p_{\xi|\eta}(x|y)$$

$$= p(x, y) / p_{\eta}(y) = \sqrt{\tau^2 + \sigma^2} / (\sqrt{2\pi} \tau \sigma) \cdot \exp \left\{ -\frac{(\sigma^2 + \tau^2)}{2\tau^2\sigma^2} \cdot \left[x - \frac{\sigma^2 m + \tau^2 y}{\sigma^2 + \tau^2} \right]^2 \right\},$$

$$\text{故在 } \eta = y \text{ 时, } \xi \text{ 的条件分布为 } N\left(\frac{\sigma^2 m + \tau^2 y}{\sigma^2 + \tau^2}, \frac{\sqrt{\sigma^2 + \tau^2}}{\sigma\tau}\right).$$

3.90 设 $\xi_1, \xi_2, \dots, \xi_n, \dots$ 为具有数学期望的独立随机变量序列, 随机变量 η 只取正整数值,

且与 $\{\xi_n, n \geq 1\}$ 独立, 证明:

$$E \sum_{k=1}^{\eta} \xi_k = \sum_{k=1}^{\infty} E \xi_k \cdot P(\eta \geq k)$$

$$\begin{aligned}
\text{证: } E \sum_{k=1}^{\eta} \xi_k &= E \left[E \left(\sum_{k=1}^{\eta} \xi_k / \eta \right) \right] \\
&= \sum_{s=1}^{\infty} E \left(\sum_{k=1}^s \xi_k \right) \cdot P(\eta = s) \\
&= \sum_{s=1}^{\infty} \left(\sum_{k=1}^s E \xi_k \right) \cdot P(\eta = s) \\
&= \sum_{k=1}^{\infty} E \xi_k \cdot \left(\sum_{s=k}^{\infty} P(\eta = s) \right) \\
&= \sum_{k=1}^{\infty} E \xi_k \cdot P(\eta \geq k)
\end{aligned}$$

3.91 求下列连续型分布的特征函数:

(1) $(-a, a)$ 上的均匀分布 ($a > 0$),

(2) 柯西分布, 其密度函数为

$$p(x) = \frac{a}{\pi} \cdot \frac{1}{(x-b)^2 + a^2}, \quad (a > 0)$$

(3) T -分布, 其密度函数为

$$p(x) = \begin{cases} \frac{\beta^\alpha}{T(\alpha)} \cdot x^{\alpha-1} \cdot e^{-\beta x} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (\alpha > 0, \beta > 0)$$

$$\text{解: (1) } \varphi(t) = \int_{-a}^a e^{itx} \cdot \frac{1}{2a} \cdot dx = \frac{\sin at}{at}$$

(2)

$$\varphi(t) = \int_{-a}^a e^{itx} \cdot \frac{a}{\pi} \cdot \frac{1}{(x-b)^2 + a^2} dx = \frac{a}{\pi} \cdot e^{itb} \cdot \int_{-\infty}^{\infty} \frac{e^{itu}}{u^2 + a^2} du = \frac{2a}{\pi} \cdot e^{itb} \int_0^{\infty} \frac{\cos tu}{u^2 + a^2} du$$

由拉普拉斯积分

$$\int_0^{\infty} \frac{\cos \beta x}{\alpha^2 + x^2} dx = \frac{\pi}{2\alpha} e^{-\alpha\beta}, \quad (\alpha, \beta > 0), \text{ 得 } \varphi(t) = e^{ibt-a|t|}$$

(3)

$$\varphi(t) = \int_0^{\infty} e^{itx} \cdot \beta^\alpha / \Gamma(\alpha) \cdot x^{\alpha-1} \cdot e^{-\beta x} dx = \beta^\alpha / \Gamma(\alpha) \cdot \int_0^{\infty} e^{(it-\beta)x} \cdot x^{\alpha-1} dx = \beta^\alpha / \Gamma(\alpha) \cdot \Gamma(\alpha) / (\beta - it)^\alpha = (1 - \frac{it}{\beta})^{-\alpha}$$

$$= (1 - \frac{it}{\beta})^{-\alpha}$$

3.93 若 $\varphi(t)$ 是特征函数, 证明下列函数也是特征函数: (1)

$$\varphi(-t); (2) |\varphi(t)|^2; (3) [\varphi(t)]^2 \quad (n \text{ 为正整数})$$

证：(1) 若 $\varphi(t)$ 是随机变量 ξ 的特征函数，则 $\varphi(-t)$ 是随机变量 $\eta = -\xi$ 的特征函数；

(2) 若 ξ 与 η 独立同分布，其特征函数为 $\varphi(t)$ 。则 $|\varphi(t)|^2 = \varphi(t) \cdot \varphi(-t)$ 是随机变量 $\zeta = \xi - \eta$ 的特征函数；

(3) 若 ξ_1, \dots, ξ_n 独立分布，其特征函数为 $\varphi(t)$ 。则 $[\varphi(t)]^n$ 是随机变量 $\eta = \sum_{i=1}^n \xi_i$ 的特征函数。

3.94 证明下列函数是特征函数，并找出相应的分布函数：

$$(1) \cos t; (2) \cos^2 t; (3) \frac{1}{1+it}; (4) \left(\frac{\sin t}{t}\right)^2; (5) \frac{1}{2e^{-it}-1}.$$

证：(1) $\cos t = \frac{1}{2} \cdot e^{it} + \frac{1}{2} \cdot e^{-it}$ ，所以 $\cos t$ 是两点分布

ξ	-1	1
P	1/2	1/2

的特征函数。

(2) $\cos^2 t = \frac{1}{2} + \frac{1}{4} \cdot e^{2it} + \frac{1}{4} \cdot e^{-2it}$ ，所以 $\cos^2 t$ 是三点分布

ξ	-2	0	2
P	1/4	1/2	1/4

的特征函数。

(3) 密度函数为 $p(x) = e^{-x}, x \geq 0; p(x) = 0, x < 0$ 的指数分布的特征函数为 $\frac{1}{1-it}$ ，所以 $\frac{1}{1+it}$ 是密度函数为 $p(x) = e^x, x \leq 0; p(x) = 0, x > 0$ 的分布的特征函数。

(4) $[-1,1]$ 上均匀分布的特征函数为 $\frac{\sin t}{t}$ ，所以互相独立且同为 $[-1,1]$ 上均匀分布的两个随机变量和的特征函数为 $(\frac{\sin t}{t})^2$ ，即 $(\frac{\sin t}{t})^2$ 是密度函数为

$$p(x) = \begin{cases} (2+x)/4 & -2 \leq x < 0 \\ (2-x)/4 & 0 \leq x \leq 2 \\ 0 & \text{其它} \end{cases}$$

的分布的特征函数。

(5) $\frac{1}{2e^{-it}-1} = \sum_{k=1}^{\infty} \frac{1}{2^k} e^{ikt}$, 所以 $\frac{1}{2e^{-it}-1}$ 是几何分布

$$P(\xi = k) = \frac{1}{2^k}, k=1,2,3,\dots$$

的特征函数。

3.95 试举一个满足 (1) $\varphi(-t) = \overline{\varphi(t)}$, (2) $|\varphi(t)| \leq \varphi(0) = 1$, 但是 $\varphi(t)$ 不是特征函数的例子。

解: 令

$$\varphi(t) = \begin{cases} 1 & t=0 \\ 0 & t \neq 0 \end{cases}$$

则 $\varphi(t)$ 满足 (1), (2), 但 $\varphi(t)$ 在 $t=0$ 点不连续, 故 $\varphi(t)$ 不是特征函数。

3.96 证明函数

$$\varphi(t) = \begin{cases} 1 - \frac{|t|}{a} & |t| \leq a \\ 0 & |t| > a \end{cases} \quad (a > 0)$$

是特征函数, 并求出它的分布函数。

解: 由于

$$\int_{-\infty}^{\infty} |\varphi(t)| dt = \int_{-a}^a \left(1 - \frac{|t|}{a}\right) dt = a < \infty$$

故欲证 $\varphi(t)$ 是特征函数, 仅须验证

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \cdot \varphi(t) dt = \frac{1}{2\pi} \int_{-a}^a e^{-itx} \cdot \left(1 - \frac{|t|}{a}\right) dt = \frac{1}{\pi} \int_0^a \left(1 - \frac{t}{a}\right) \cos tx dt = \frac{1}{\pi} \cdot \frac{1 - \cos ax}{ax^2}$$

是密度函数由于 $p(x) \geq 0$,

$$\int_{-\infty}^{\infty} p(x) dx = \frac{a}{\pi} \int_0^{\infty} \sin^2 \frac{ax}{2} \left/ \left(\frac{ax}{2}\right)^2 \right. dx = \frac{2}{\pi} \int_0^{\infty} \frac{\sin^2 y}{y^2} dy = 1,$$

所以 $\varphi(t)$ 为特征函数, 其分布函数为

$$F(x) = \int_{-\infty}^x \frac{1}{\pi} \cdot \frac{1 - \cos at}{at^2} dt.$$

3.97 设 $\varphi(t)$ 是一个特征函数。 $h > 0$, 证明:

$$\varphi_h(t) = \varphi(t) \cdot \frac{\sin th}{th}$$

也是特征函数。

证：设 ξ 与 η 相互独立， ξ 的特征函数为 $\varphi(t)$ ， η 服从 $[-h, h]$ 上的均匀分布， η 的特征函数为 $\frac{\sin th}{th}$ ，则 $\frac{\sin th}{th}$ 是 $\xi + \eta$ 的特征函数。

3.98 设 $\xi_1, \xi_2, \dots, \xi_n$ 为 n 个独立同柯西分布的随机变量，证明 $\frac{1}{n} \sum_{i=1}^n \xi_i$ 与 ξ_1 有相同的分布。

证：柯西分布 $p(x) = \frac{a}{\pi} \cdot \frac{1}{(x-b)^2 + a^2}$ 的特征函数 $\varphi(t) = e^{ibt-a|t|}$ 。故 $\frac{1}{n} \cdot \sum_{i=1}^n \xi_i$ 的特征函数为

$$\left[\varphi\left(\frac{t}{n}\right) \right]^n = e^{ibt-a|t|}。所以 \frac{1}{n} \cdot \sum_{i=1}^n \xi_i 与同分布。$$

3.99 设 $\xi_1, \xi_2, \dots, \xi_n$ 为独立同 T -分布的随机变量，求 $\sum_{i=1}^n \xi_i$ 的分布。

解： T -分布 $p(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$ ， $x > 0$ ； $p(x) = 0$ ， $x \leq 0$ 的特征函数

$\varphi(t) = \left(1 - \frac{it}{\beta}\right)^{-\alpha}$ 。故 $\sum_{i=1}^n \xi_i$ 的特征函数为

$$[\varphi(t)]^n = \left(1 - \frac{it}{\beta}\right)^{-n\alpha},$$

所以 $\sum_{i=1}^n \xi_i$ 也是 T -分布，其密度函数为 $p(x) = \frac{\beta^{n\alpha}}{\Gamma(n\alpha)} \cdot x^{n\alpha-1} \cdot e^{-\beta x}$ ， $x > 0$ ； $p(x) = 0$ ， $x \leq 0$ 。

3.100 设二维随机变量 (ξ, η) 具有联合密度函数为

$$p(x, y) = \begin{cases} \frac{1}{4} [1 + xy(x^2 + y^2)] & |x| < 1, |y| < 1 \\ 0 & \text{其它} \end{cases}$$

证明： $\xi + \eta$ 的特征函数等于 ξ ， η 的特征函数的乘积，但是 ξ 与 η 并不相互独立。

证： $p_{\xi+\eta}(z) = \int_{-\infty}^{\infty} p(x, z-x) dx$

$$= \begin{cases} (2+x)/4 & -2 < x < 0 \\ (2-x)/4 & 0 \leq x < 2 \\ 0 & \text{其它}。 \end{cases}$$

$\xi + \eta$ 的特征函数为 $\left(\frac{\sin t}{t}\right)^2$ 。

$$p_{\xi}(x) = 1/2, -1 < x < 1; p_{\xi}(x) = 0, |x| \geq 1. p_{\eta}(y) = 1/2, -1 < y < 1; p_{\eta}(y) = 0, |y| \geq 1.$$

故 ξ 与 η 的特征函数皆为 $\frac{\sin t}{t}$ ，所以 $\xi + \eta$ 的特征函数等于 ξ 、 η 的特征函数的乘积。由

$p(x, y) = p_{\xi}(x) \cdot p_{\eta}(y)$ ，故 ξ 与 η 不互相独立。

3.101 设随机变量 ξ 服从柯西分布，其特征函数为 $e^{-|t|}$ ，又令 $\eta = a\xi (a > 0)$ ，证明 $\xi + \eta$ 的特征函数等于 ξ 、 η 的特征函数的乘积，但 ξ 与 η 不独立。

证：由 ξ 的特征函数 $\varphi_{\xi}(t) = e^{-|t|}$ 推得， $\eta = a\xi$ 与 $\xi + \eta$ 的特征函数分别为 $\varphi_{\eta}(t) = e^{-a|t|}$ 与 $\varphi_{\xi+\eta}(t) = e^{-(a+1)|t|}$ ，故 $\varphi_{\xi+\eta}(t) = \varphi_{\xi}(t) \cdot \varphi_{\eta}(t)$ 。

倘若 ξ 与 η 相互独立，令 ξ 的分布函数为 $F(x)$ ，则

$$F(x) = P(\xi < x, \eta < ax) = P(\xi < x) \cdot P(\eta < ax) = P(\xi < x) \cdot P(\xi < x) = [F(x)]^2,$$

故 $F(x) = 0$ 或 1 ，此与 ξ 服从柯西分布相矛盾，故 ξ 与 η 互不独立。

3.102 判别下列函数是否为特征函数（说明理由）：

$$(1) \sin t; (2) \frac{1-t}{1+t^2}; (3) \ln(e+|t|); (4) \frac{1}{1-|t|}; (5) \frac{1}{(1+t^2)^2}.$$

解：(1) 不是，因为 $\sin 0 \neq 1$ 。

(2) 不是，因为当 $-1 < t < 0$ 时， $\frac{1-t}{1+t^2} > 1$ 。

(3) 不是，因为 $\ln(e+|t|) \leq 1$ 不成立

(4) 不是，因为 $\varphi(t) = \frac{1}{1-|t|} \neq \overline{\varphi(-t)}$ 。

(5) 是的，拉普拉斯分布 $p(x) = \frac{1}{2} \cdot e^{-|x|}$ 的特征函数为 $\frac{1}{1+t^2}$ ，所以 $\frac{1}{(1+t^2)^2}$ 也是特征

函数。

第四章 大数定律与中心极限定理

4.1 设 $D(x)$ 为退化分布：

$$D(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$$

讨论下列分布函数列的极限是否仍是分布函数？

(1) $\{D(x+n)\}$; (2) $\{D(x+\frac{1}{n})\}$; (3) $\{D(x-\frac{1}{n})\}$, 其中 $n=1, 2, \dots$

解：(1) (2) 不是；(3) 是。

4.2 设分布函数 $F_n(x)$ 如下定义：

$$F_n(x) = \begin{cases} 0 & x \leq -n \\ \frac{x+n}{2n} & -n < x \leq n \\ 1 & x > n \end{cases}$$

问 $F(x) = \lim_{n \rightarrow \infty} F_n(x)$ 是分布函数吗？

解：不是。

4.3 设分布函数列 $\{F_n(x)\}$ 弱收敛于分布函数 $F(x)$ ，且 $F(x)$ 为连续函数，则

$\{F_n(x)\}$ 在 $(-\infty, \infty)$ 上一致收敛于 $F(x)$ 。

证：对任意的 $\varepsilon > 0$ ，取 M 充分大，使有

$$1 - F(x) < \varepsilon, \forall x \geq M; F(x) < \varepsilon, \forall x \leq -M$$

对上述取定的 M ，因为 $F(x)$ 在 $[-M, M]$ 上一致连续，故可取它的 k 分点：

$x_1 = -M < x_2 < \dots < x_{k-1} < x_k = M$ ，使有 $F(x_{i+1}) - F(x_i) < \varepsilon, 1 \leq i < k$ ，再令

$x_0 = -\infty, x_{k+1} = \infty$ ，则有

$$F(x_{i+1}) - F(x_i) < \varepsilon, 0 \leq i < k+1$$

(1)

这时存在 N ，使得当 $n > N$ 时有

$$|F_n(x_i) - F(x_i)| < \varepsilon, 0 \leq i \leq k+1$$

(2)

成立，对任意的 $x \in (-\infty, \infty)$ ，必存在某个 $i (0 \leq i \leq k)$ ，使得 $x \in (x_i, x_{i+1})$ ，由 (2)

知当 $n > N$ 时有

$$F_n(x) \leq F_n(x_{i+1}) < F(x_{i+1}) + \varepsilon$$

(3)

$$F_n(x) \geq F_n(x_i) > F(x_i) - \varepsilon$$

(4)

由 (1), (3), (4) 可得

$$F_n(x) - F(x) < F(x_{i+1}) - F(x) + \varepsilon \leq F(x_{i+1}) - F(x_i) + \varepsilon < 2\varepsilon,$$

$$F_n(x) - F(x) > F(x_i) - F(x) - \varepsilon \geq F(x_i) - F(x_{i+1}) - \varepsilon > -2\varepsilon,$$

即有 $|F_n(x) - F(x)| < 2\varepsilon$ 成立, 结论得证。

4.5 设随机变量序列 $\{\xi_n\}$ 同时依概率收敛于随机变量 ξ 与 η , 证明这时必有

$$P(\xi = \eta) = 1.$$

证: 对任意的 $\varepsilon > 0$ 有 $(|\xi - \eta| \geq \varepsilon) \subset \left[\left(|\xi_n - \eta| \geq \frac{\varepsilon}{2} \right) \right]$, 故

$$0 \leq P(|\xi - \eta| \geq \varepsilon) \leq P\left(|\xi - \xi_n| \geq \frac{\varepsilon}{2}\right) + P\left(|\xi_n - \eta| \geq \frac{\varepsilon}{2}\right) \rightarrow 0, n \rightarrow \infty$$

即对任意的 $\varepsilon > 0$ 有 $P(|\xi - \eta| \geq \varepsilon) = 0$ 成立, 于是有

$$P(\xi \neq \eta) = P\left[\bigcup_{k=1}^{\infty} \left(|\xi - \eta| \geq \frac{1}{k}\right)\right] \leq \sum_{k=1}^{\infty} P\left(|\xi - \eta| \geq \frac{1}{k}\right) = 0$$

从而 $P(\xi = \eta) = 1$ 成立, 结论得证。

4.6 设随机变量序列 $\{\xi_n\}$, $\{\eta_n\}$ 分别依概率收敛于随机变量 ξ 与 η , 证明:

$$(1) \xi_n + \eta_n \xrightarrow{P} \xi + \eta; (2) \xi_n \times \eta_n \xrightarrow{P} \xi \times \eta.$$

证: (1) 因为 $(|\xi + \eta - \xi_n - \eta_n| \geq \varepsilon) \subset \left[\left(|\xi - \xi_n| \geq \frac{\varepsilon}{2} \right) \cup \left(|\eta - \eta_n| \geq \frac{\varepsilon}{2} \right) \right]$ 故

$$0 \leq P(|\xi + \eta - \xi_n - \eta_n| \geq \varepsilon) \leq P\left(|\xi - \xi_n| \geq \frac{\varepsilon}{2}\right) + P\left(|\eta - \eta_n| \geq \frac{\varepsilon}{2}\right) \rightarrow 0, n \rightarrow \infty$$

即 $\xi_n + \eta_n \xrightarrow{P} \xi + \eta$ 成立。

(2) 先证明这时必有 $\xi_n^2 \xrightarrow{P} \xi^2$ 。对任给的 $\varepsilon > 0, \delta > 0$ 取 M 足够大

$\left(\frac{\varepsilon}{M} \leq 1\right)$, 使有 $P\left(|\xi| > \frac{M-1}{2}\right) < \delta$ 成立, 对取定的 M , 存在 N , 当 $n > N$ 时有

$$P(|\xi_n - \xi| \geq 1) \leq P\left(|\xi_n - \xi| \geq \frac{\varepsilon}{M}\right) < \delta \text{ 成立这时}$$

$$\begin{aligned}
P(|\xi_n + \xi| > M) &\leq P(|\xi_n - \xi| + |2\xi| > M) \\
&= P\{(|\xi_n - \xi| + |2\xi| > M) \cap (|\xi_n - \xi| < 1)\} \\
&\quad + P\{(|\xi_n - \xi| + |2\xi| > M) \cap (|\xi_n - \xi| \geq 1)\} \\
&\leq P(|2\xi| > M - 1) + P(|\xi_n - \xi| \geq 1) < 2\delta
\end{aligned}$$

从而有

$$\begin{aligned}
P(|\xi_n^2 - \xi^2| \geq \varepsilon) &= P(|\xi_n - \xi| |\xi_n + \xi| \geq \varepsilon) \\
&= P\{(|\xi_n - \xi| |\xi_n + \xi| \geq \varepsilon) \cap (|\xi_n + \xi| \leq M)\} \\
&\quad + P\{(|\xi_n - \xi| |\xi_n + \xi| \geq \varepsilon) \cap (|\xi_n + \xi| > M)\} \\
&\leq P(|\xi_n - \xi| \geq \frac{\varepsilon}{M}) + P(|\xi_n + \xi| > M) < 3\delta
\end{aligned}$$

由 ε, δ 的任意性知 $\xi_n^2 \xrightarrow{P} \xi^2$, 同理可证 $\eta_n^2 \xrightarrow{P} \eta^2$, 由前述 (1) 有

$$2\xi_n\eta_n = (\xi_n + \eta_n)^2 - \xi_n^2 - \eta_n^2 \xrightarrow{P} (\xi + \eta)^2 - \xi^2 - \eta^2 = 2\xi\eta$$

故 $\xi_n \times \eta_n \xrightarrow{P} \xi \times \eta$, 结论成立。

4.7 设随机变量序列 $\xi_n \xrightarrow{P} a$, $a \neq 0$ 是一个常数, 且 $\xi_n \neq 0$, 证明 $\frac{1}{\xi_n} \xrightarrow{P} \frac{1}{a}$ 。

证：不妨设 $a > 0$ 对任意的 $0 < \varepsilon < a$, 当 $|\xi_n - a| < \varepsilon$ 时有

$$|\xi_n a| = |a^2 + a(\xi_n - a)| \geq a^2 - a\varepsilon,$$

因而 $\left(\left|\frac{\xi_n - a}{\xi_n a}\right| \geq \varepsilon\right) \subset \left(\left|\frac{\xi_n - a}{a^2 - a\varepsilon}\right| \geq \varepsilon\right)$ 。于是有

$$\begin{aligned}
0 &\leq P\left(\left|\frac{1}{\xi_n} - \frac{1}{a}\right| \geq \varepsilon\right) \\
&= P\left\{\left(\left|\frac{\xi_n - a}{\xi_n a}\right| \geq \varepsilon\right) \cap (|\xi_n - a| < \varepsilon)\right\} + P\left\{\left(\left|\frac{\xi_n - a}{\xi_n a}\right| \leq \varepsilon\right) \cap (|\xi_n - a| \geq \varepsilon)\right\} \\
&\leq P\left(\left|\frac{\xi_n - a}{a^2 - a\varepsilon}\right| \geq \varepsilon\right) + P(|\xi_n - a| \geq \varepsilon) \rightarrow 0, n \rightarrow \infty.
\end{aligned}$$

结论成立。

4.9 证明随机变量序列 $\{\xi_n\}$ 依概率收敛于随机变量 ξ 的充要条件为：

$$E \frac{|\xi_n - \xi|}{1 + |\xi_n - \xi|} \rightarrow 0, n \rightarrow \infty$$

证：充分性，令 $f(x) = \frac{x}{1+x}$ ， $x > 0$ ，则 $f'(x) = \frac{1}{(1+x)^2} > 0, x > 0$ ，故 $f(x)$ 是

$x(x > 0)$ 的单调上升函数，因而 $(|\xi_n - \xi| > \varepsilon) \subset \left(\frac{|\xi_n - \xi|}{1 + |\xi_n - \xi|} > \frac{\varepsilon}{1 + \varepsilon} \right)$ ，于是有

$$\begin{aligned} P(|\xi_n - \xi| > \varepsilon) &\leq P\left(\frac{|\xi_n - \xi|}{1 + |\xi_n - \xi|} > \frac{\varepsilon}{1 + \varepsilon}\right) \\ &\leq \frac{1 + \varepsilon}{\varepsilon} E \frac{|\xi_n - \xi|}{1 + |\xi_n - \xi|} \rightarrow 0, n \rightarrow \infty \end{aligned}$$

对任意的 $\varepsilon > 0$ 成立，充分性得证。

必要性，对任给的 $\varepsilon > 0$ ，令 $A_\varepsilon = \{\omega : |\xi_n - \xi| > \varepsilon\}$ ，因为 $\xi_n \xrightarrow{P} \xi$ ，故存在

充分大的 N 使得当 $n \geq N$ 时有 $P(A_\varepsilon) < \varepsilon$ ，于是有

$$\begin{aligned} E \frac{|\xi_n - \xi|}{1 + |\xi_n - \xi|} &= E\left(\frac{|\xi_n - \xi|}{1 + |\xi_n - \xi|} I_{A_\varepsilon}\right) + E\left(\frac{|\xi_n - \xi|}{1 + |\xi_n - \xi|} I_{A_\varepsilon^c}\right) \\ &\leq P(A_\varepsilon) + \varepsilon < 2\varepsilon, \end{aligned}$$

由 ε 的任意性知 $E \frac{|\xi_n - \xi|}{1 + |\xi_n - \xi|} \rightarrow 0, n \rightarrow \infty$ ，结论为真。

4.10 设随机变量 ξ_n 按分布收敛于随机变量 ξ ，又数列 $a_n \rightarrow a$ ， $b_n \rightarrow b$ ，证明 $a_n \xi_n + b_n$ 也按分布收敛于 $a\xi + b$ 。

证：先证明 $a\xi_n$ 按分布收敛于 $a\xi$ 。 $a = 0$ 时为显然，不妨设 $a > 0$ （ $a < 0$ 时的修改为显然），若 $a\xi$ ， ξ ， $a\xi_n$ ， ξ_n 的分布函数分别记作 $F_{a\xi}(\cdot)$ ， $F_\xi(\cdot)$ ，

$F_{a\xi_n}(\cdot)$ 与 $F_n(\cdot)$ ，则 $F_{a\xi}(x) = F_\xi\left(\frac{x}{a}\right)$ ，当 x 是 $F_{a\xi}(\cdot)$ 的连续点时， $\frac{x}{a}$ 是 $F_\xi(\cdot)$ 的连续点，

于是有

$$\lim_{n \rightarrow \infty} F_{a\xi_n}(x) = \lim_{n \rightarrow \infty} F_n\left(\frac{x}{a}\right) = \lim_{n \rightarrow \infty} F_\xi\left(\frac{x}{a}\right) = F_{a\xi}(x)$$

成立，结论为真。由 4.12 知 $\xi_n(a_n - a) \xrightarrow{P} 0$ ，再由 4.6(1) 知 $\xi_n(a_n - a) + b_n \xrightarrow{P} b$ ，于是由前述结论及 4.11 知 $\xi_n a_n + b_n = a \xi_n + (a_n - a) \xi_n + b_n$ 按分布收敛于 $a \xi + b$ ，结论得证。

4.11 设随机变量序列 $\{\xi_n\}$ 按分布收敛于随机变量 ξ ，随机变量序列 $\{\eta_n\}$ 依概率收敛于常数 a ，证明 $\xi_n + \eta_n$ 按分布收敛于 $\xi + a$ 。

证：记 ξ, ξ_n 的分布函数分别为 $F(x), F_n(x)$ ，则 $\xi + a$ 的分布函数为 $F(x - a)$ ，设 x 是 $F(x - a)$ 的连续点，则对任给的 $\varepsilon > 0$ ，存在 $\delta > 0$ ，使当 $0 < \varepsilon' < \delta$ 时有

$$|F(x - a \pm \varepsilon') - F(x - a)| < \delta \quad (1)$$

现任取 $0 < \varepsilon_1 < \varepsilon_2 < \delta$ ，使得 $x - a + \varepsilon_1, x - a - \varepsilon_2$ 都是 $F(\cdot)$ 的连续点，这时存在 N_1 ，当 $n \geq N_1$ 时有

$$|F(x - a + \varepsilon_1) - F_n(x - a + \varepsilon_1)| < \varepsilon \quad (2)$$

$$|F(x - a - \varepsilon_2) - F_n(x - a - \varepsilon_2)| < \varepsilon \quad (3)$$

对取定的 ε_1 ，存在 N_2 ，当 $n \geq N_2$ 时有

$$P(|\eta_n - a| \geq \varepsilon_1) < \varepsilon \quad (4)$$

于是当 $n \geq \max(N_1, N_2)$ 时，由 (1)，(2)，(4) 式有

$$\begin{aligned} & P(\xi_n + \eta_n - a < x - a) \\ &= P\{(\xi_n + \eta_n - a < x - a) \cap (|\eta_n - a| < \varepsilon_1)\} + P\{(\xi_n + \eta_n - a < x - a) \cap (|\eta_n - a| \geq \varepsilon_1)\} \\ &\leq P(\xi_n < x - a + \varepsilon_1) + P(|\eta_n - a| \geq \varepsilon_1) < F(x - a) + 3\varepsilon \end{aligned} \quad (5)$$

又因为

$$\begin{aligned} P(\xi_n < x - a - \varepsilon_2) &= P\{[\xi_n + \eta_n - (\eta_n - a) < x - \varepsilon_2] \cap (|\eta_n - a| < \varepsilon_2)\} \\ &\quad + P\{(\xi_n < x - a - \varepsilon_2) \cap (|\eta_n - a| \geq \varepsilon_2)\} \end{aligned}$$

于是由 (1)，(3)，(4) 式有

$$\begin{aligned} P(\xi_n + \eta_n - a < x - a) &\geq P\{[\xi_n + \eta_n - (\eta_n - a) < x - \varepsilon_2] \cap (|\eta_n - a| < \varepsilon_2)\} \\ &\geq P(\xi_n < x - a - \varepsilon_2) - P(|\eta_n - a| \geq \varepsilon_2) \geq F(x - a) - 3\varepsilon \end{aligned}$$

(6)

由 (5), (6) 两式可得

$$|P(\xi_n + \eta_n - a < x - a) - F(x - a)| < 3\varepsilon$$

由 ε 的任意性即知 $\xi_n + \eta_n$ 按分布收敛于 $\xi + a$, 结论得证。

4.12 设随机变量序列 $\{\xi_n\}$ 按分布收敛于 ξ , 随机变量序列 $\{\eta_n\}$ 依概率收敛于 0, 证明

$$\xi_n \eta_n \xrightarrow{P} 0.$$

证: 记 ξ, ξ_n 的分布函数分别为 $F(x), F_n(x)$, 对任给的 $\varepsilon > 0$, 取 $a > 0, b > 0$ 足够大, 使 $-a, b$ 是 $F(x)$ 的连续点且

$$1 - F(b) < \varepsilon, F(-a) < \varepsilon$$

因为 $F_n(x) \xrightarrow{w} F(x)$, 故存在 N_1 , 当 $n \geq N_1$ 时有

$$1 - F_n(b) < 2\varepsilon, F_n(-a) < 2\varepsilon$$

令 $M = \max(a, b)$, 因为 $\eta_n \xrightarrow{P} 0$, 故存在 N_2 , 当 $n \geq N_2$ 时有

$$P(|\eta_n| > \frac{\varepsilon}{M}) < \varepsilon$$

而

$$\begin{aligned} P(|\xi_n \eta_n| > \varepsilon) &= P\{(|\xi_n \eta_n| > \varepsilon) \cap [(-a \leq \xi_n < b) \cap (|\eta_n| \leq \frac{\varepsilon}{M})]\} \\ &+ P\{(|\xi_n \eta_n| > \varepsilon) \cap \overline{(-a \leq \xi_n < b)} \cup (|\eta_n| > \frac{\varepsilon}{M})\} = I_1 + I_2 \end{aligned}$$

其中 $I_1 = 0$, 当 $n \geq \max(N_1, N_2)$ 时有

$$\begin{aligned} P\{(|\xi_n \eta_n| > \varepsilon) \cap \overline{(-a \leq \xi_n < b)}\} &\leq P\{\overline{(-a \leq \xi_n < b)}\} \\ &= P\{(\xi_n < -a) \cup (\xi_n \geq b)\} = F_n(-a) + [1 - F_n(b)] < 4\varepsilon \end{aligned}$$

因而 $P(|\xi_n \eta_n| > \varepsilon) = I_2 < 5\varepsilon$, 由 ε 的任意性知 $\xi_n \eta_n \xrightarrow{P} 0$, 结论为真。

4.13 设随机变量 ξ_n 服从柯西分布, 其密度函数为

$$p_n(x) = \frac{n}{\pi(1 + n^2 x^2)}$$

证明 $\xi_n \xrightarrow{P} 0, n \rightarrow \infty$ 。

证：对任意的 $\varepsilon > 0$ ，有

$$P(|\xi_n| \leq \varepsilon) = \int_{-\varepsilon}^{\varepsilon} \frac{n}{\pi(1+n^2x^2)} dx = \int_{-n\varepsilon}^{n\varepsilon} \frac{1}{\pi(1+t^2)} dt \rightarrow 1, n \rightarrow \infty$$

故 $\xi_n \xrightarrow{P} 0, n \rightarrow \infty$ 。

4.14 设 $\{\xi_n\}$ 为一列独立同分布随机变量，其密度函数为

$$p(x) = \begin{cases} 1/\beta & 0 < x < \beta \\ 0 & \text{其它} \end{cases}$$

其中 $\beta > 0$ 为常数，令 $\eta_n = \max(\xi_1, \xi_2, \dots, \xi_n)$ ，证明 $\eta_n \xrightarrow{P} \beta$ 。

证：对任意的 n ， $0 < \eta_n < \beta$ 为显然，这时有

$$P(\eta_n < x) = \prod_{i=1}^n P(\xi_i < x) = \prod_{i=1}^n \int_0^x \frac{1}{\beta} dx = \left(\frac{x}{\beta}\right)^n, 0 < x < \beta$$

$$P(\eta_n < x) = 0, x \leq 0; P(\eta_n < x) = 1, x \geq \beta$$

对任意的 $\varepsilon > 0 (\varepsilon < \beta)$ ，有

$$P(|\eta_n - \beta| > \varepsilon) = P(\eta_n < \beta - \varepsilon) = \left(\frac{\beta - \varepsilon}{\beta}\right)^n \rightarrow 0, n \rightarrow \infty$$

故 $\eta_n \xrightarrow{P} \beta$ 成立，结论得证。

4.15 设 $\{\xi_n\}$ 为一列独立同分布随机变量，其密度函数为

$$p(x) = \begin{cases} e^{-(x-a)} & x \geq a \\ 0 & x < a \end{cases}$$

令 $\eta_n = \min(\xi_1, \xi_2, \dots, \xi_n)$ ，证明 $\eta_n \xrightarrow{P} a$ 。

证：设 ξ_i 的分布函数为 $F(x)$ ，有

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

这时有

$$P(\eta_n \geq x) = \prod_{i=1}^n P(\xi_i \geq x) = [1 - F(x)]^n = e^{-n(x-a)}, x > a$$

对任意的 $\varepsilon > 0$, 有

$$P(|\eta_n - a| \geq \varepsilon) = P(\eta_n - a \geq \varepsilon) = e^{-n\varepsilon} \rightarrow 0, n \rightarrow \infty$$

故 $\eta_n \xrightarrow{P} a$ 成立, 结论得证。

4.17 设 $\{\xi_n\}$ 为一列独立同分布随机变量, 都服从 (0,1) 上的均匀分布, 若

$\eta_n = (\prod_{k=1}^n \xi_k)^{\frac{1}{n}}$, 证明 $\eta_n \xrightarrow{P} c$ (c 为常数), 并求出 c 。

证: 这时 $\{\ln \xi_n\}$ 也是独立同分布随机变量序列, 且

$$E\xi_n = \int_0^1 \ln x dx = -1$$

由辛钦大数定律知 $\{\ln \xi_n\}$ 服从大数定理, 即有 $\frac{1}{n} \sum_{i=1}^n \ln \xi_i \xrightarrow{P} -1$, 令 $f(x) = e^x$, 则

$f(x)$ 是直线上的连续函数, 由 4.8 题知

$$\left(\prod_{i=1}^n \xi_i\right)^{\frac{1}{n}} = e^{\frac{1}{n} \sum_{i=1}^n \ln \xi_i} \xrightarrow{P} e^{-1} = c$$

结论成立。

4.18 设 $\{\xi_n\}$ 为一列独立同分布随机变量, 每个随机变量的期望为 a , 且方差存在,

证明 $\frac{2}{n(n+1)} \sum_{k=1}^n k \xi_k \xrightarrow{P} a$ 。

证: 已知 $E\xi_n = a$, 记 $D\xi_n = \sigma^2$, 令 $\eta_n = \frac{2}{n(n+1)} \sum_{k=1}^n k \xi_k$, 则

$$E\eta_n = \frac{2}{n(n+1)} \sum_{k=1}^n k a = a$$

$$D\eta_n = \frac{4}{n^2(n+1)^2} \sum_{k=1}^n k^2 \sigma^2 \leq \frac{4\sigma^2}{n+1}$$

对任给的 $\varepsilon > 0$, 由契贝晓夫不等式有

$$P(|\eta_n - a| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} D\eta_n \leq \frac{1}{\varepsilon^2} \frac{4\sigma^2}{n+1} \rightarrow 0, n \rightarrow \infty$$

故 $\eta_n \xrightarrow{P} a$, 结论得证。

4.19 设 $\{\xi_n\}$ 为一列独立同分布随机变量, 且 $D\xi_n = \sigma^2$ 存在, 数学期望为零, 证

明 $\frac{1}{n} \sum_{k=1}^n \xi_k^2 \xrightarrow{P} \sigma^2$ 。

证：这时 $\{\xi_n^2\}$ 仍独立同分布，且 $E\xi_n^2 = D\xi_n = \sigma^2 < \infty$ ，由辛钦大数定律知结论成立。

4.21 设随机变量序列 $\{\xi_n\}$ 按分布收敛于随机变量 ξ ，又随机变量序列 $\{\eta_n\}$ 依概率收敛于常数 $a(a \neq 0), \eta_n \neq 0$ ，则 $\{\xi_n/\eta_n\}$ 按分布收敛于 ξ/a 。

证：由 4.7 题知 $\frac{1}{\eta_n} - \frac{1}{a} \xrightarrow{P} 0$ ，于是由 4.12 题有 $\xi_n(\frac{1}{\eta_n} - \frac{1}{a}) \xrightarrow{P} 0$ ，而 $\frac{\xi_n}{a}$ 按分布收敛于 $\frac{\xi}{a}$ （见 4.10 题的证明），因而由 4.11 题知

$$\frac{\xi_n}{\eta_n} = \xi_n \left(\frac{1}{\eta_n} - \frac{1}{a} \right) + \frac{\xi_n}{a}$$

按分布收敛于 $\frac{\xi}{a}$ ，结论成立。

4.22 设 $\{\xi_n^2\}$ 为独立同 $N(0,1)$ 分布的随机变量序列，证明 $n\xi_{n+1}/\sum_{k=1}^n \xi_k^2$ 的分布函数弱收敛于 $N(0,1)$ 分布。

证：这时 $\{\xi_n^2\}$ 也为独立同分布随机变量序列，且 $E\xi_n^2 = 1$ ，由辛钦大数定律知

$\frac{1}{n} \sum_{i=1}^n \xi_i^2 \xrightarrow{P} 1$ ，又 ξ_{n+1} 服从 $N(0,1)$ 分布，当然弱收敛于 $N(0,1)$ 分布，由 4.21 题

即知 η_n 按分布收敛于 $N(0,1)$ 分布，结论得证。

4.23 如果随机变量序列 $\{\xi_n\}$ ，当 $n \rightarrow \infty$ 时有 $\frac{1}{n^2} D\left(\sum_{k=1}^n \xi_k\right) \rightarrow 0$ ，证明 $\{\xi_n\}$ 服从大

数定律（马尔柯夫大数定律）

证：由契贝晓夫不等式即得。

4.26 在贝努里试验中，事件 A 出现的概率为 p ，令

$$\xi_n = \begin{cases} 1, & \text{若在第 } n \text{ 次及第 } n+1 \text{ 次实验中 } A \text{ 出现} \\ 0, & \text{其它} \end{cases}$$

证明 $\{\xi_n\}$ 服从大数定律。

证： $\{\xi_n\}$ 为同分布随机变量序列，且 $E\xi_n = E\xi_n^2 = p^2$ ，因而 $D\xi_n = p^2(1-p^2) \leq 1$ ，又当 $|i-j| \geq 2$ 时， ξ_i 与 ξ_j 独立，由 4.24 知 $\{\xi_n\}$ 服从大数定律，结论得证。

4.28 设 $\{\xi_n\}$ 为一列独立同分布随机变量，方差存在，又 $\sum_{n=1}^{\infty} a_n$ 为绝对收敛级数，

令 $\eta_n = n \sum_{i=1}^n \xi_i$ ，则 $\{a_n \eta_n\}$ 服从大数定律。

证：不妨设 $E\xi_n = 0$ 。否则令 $\xi'_n = \xi_n - E\xi_n$ ，并讨论 $\{\xi'_n\}$ 即可。记 $E\xi_n^2 = \sigma^2$ ，又

$c = \sum_{n=1}^{\infty} |a_n| < \infty$ 。因为 $\sum_{i=1}^n a_i \eta_i = \sum_{i=1}^n a_i (\sum_{k=1}^i \xi_k) = \sum_{k=1}^n \xi_k (\sum_{i=k}^n a_i)$ ，故有

$$\frac{1}{n^2} D(\sum_{i=1}^n a_i \eta_i) = \frac{1}{n^2} E[\sum_{k=1}^n \xi_k (\sum_{i=k}^n a_i)]^2 = \frac{\sigma^2}{n^2} \sum_{k=1}^n (\sum_{i=k}^n a_i)^2 \leq \frac{c^2 \sigma^2}{n} \rightarrow 0, n \rightarrow \infty$$

由 4.23 知 $\{a_n \eta_n\}$ 服从大数定律，结论得证。

4.30 设 $\{\xi_n\}$ 为一列独立同分布随机变量，共同分布为

$$P(\xi_n = \frac{2^k}{k^2}) = \frac{1}{2^k}, k=1,2,\dots$$

试问 $\{\xi_n\}$ 是否服从大数定律？

答：因为 $E\xi_n$ 存在，由辛钦大数定律知 $\{\xi_n\}$ 服从大数定律。

4.31 设 $\{\xi_n\}$ 为一列独立同分布随机变量，共同分布为

$$P(\xi_n = k) = \frac{c}{k^2 \log^2 k}, k=2,3,\dots$$

其中 $c = (\sum_{k=2}^{\infty} \frac{1}{k^2 \log^2 k})^{-1}$ ，问 $\{\xi_n\}$ 是否服从大数定律？

答：因为 $E\xi_n$ 存在，由辛钦大数定律知 $\{\xi_n\}$ 服从大数定律。

4.32 如果要估计抛掷一枚图钉时尖头朝上的概率，为了有 95% 以上的把握保证所观察到的频率与概率 p 的差小于 $p/10$ ，问至少应该做多少次试验？

解：令

$$\xi_n = \begin{cases} 1 & \text{第}n\text{次试验时图钉的尖头朝上} \\ 0 & \text{其它} \end{cases}$$

据题意选取试验次数 n 应满足 $P(|\frac{\sum_{i=1}^n \xi_i}{n} - p| < \frac{p}{10}) \geq 0.95$, 因为 n 比较大, 由中心极限定理有

$$\begin{aligned} P(|\frac{\sum_{i=1}^n \xi_i}{n} - p| < \frac{p}{10}) &= P(|\frac{\sum_{i=1}^n (\xi_i - p)}{\sqrt{npq}}| < \frac{1}{10} \sqrt{\frac{np}{q}}) \\ &\approx \int_{\frac{1}{10} \sqrt{\frac{np}{q}}}^{\frac{1}{10} \sqrt{\frac{np}{q}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \geq 0.95 \end{aligned}$$

故应取 $\frac{1}{10} \sqrt{\frac{np}{q}} = 2$, 即 $n = 400 \frac{q}{p}$, 但图钉底部重, 尖头轻, 由直观判断有

$$p \geq \frac{1}{2}, \text{ 因而}$$

$$\frac{q}{p} \leq 1, \text{ 故可取 } n = 400.$$

4.33 一本书共有一百万个印刷符号, 排版时每个符号被排错的概率为 0.0001, 校对时每个排版错误被改正的概率为 0.9, 求在校对后错误不多于 15 个的概率。
解: 令

$$\xi_i = \begin{cases} 1 & \text{第}i\text{个印刷符号被排错且校对后仍错误} \\ 0 & \text{其它} \end{cases}$$

因为排版与校对是两个独立的工序, 因而

$$p = P(\xi_i = 1) = 0.0001 \times 0.1 = 10^{-5}, P(\xi_i = 0) = q = 1 - p$$

$\{\xi_i\}$ 是独立同分布随机变量序列, $E\xi_i = p$, 令 $\eta_n = \sum_{i=1}^n \xi_i$, 其中 $n = 10^6$, 由中心

极限定理有

$$P(\eta_n \leq 15) = P(\frac{\eta_n - np}{\sqrt{npq}} \leq \frac{15 - np}{\sqrt{npq}} = b) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^b e^{-\frac{x^2}{2}} dx$$

其中 $b \approx \frac{5}{\sqrt{10}} \approx 1.58$, 查 $N(0,1)$ 分布表即可得 $P(\eta_n \leq 15) \approx 0.94$, 即在校对后错误

不多于 15 个的概率。

4.34 在一家保险公司里有 10000 个人参加保险, 每人每年付 12 元保险费, 在一

年里一个人死亡的概率为 0.006, 死亡时家属可向保险公司领得 1000 元, 问:

(1) 保险公司亏本的概率多大?

(2) 保险公司一年的利润不少于 40000 元, 60000 元, 80000 元的概率各为多大?

解: 保险公司一年的总收入为 120000 元, 这时

(1) 若一年中死亡人数 > 120 , 则公司亏本;

(2) 若一年中死亡人数 ≤ 80 , 则利润中死亡人数 ≥ 40000 元;

若一年中死亡人数 ≤ 60 , 则利润中死亡人数 ≥ 60000 元;

若一年中死亡人数 ≤ 40 , 则利润中死亡人数 ≥ 80000 元;

令

$$\xi_i = \begin{cases} 1 & \text{第 } i \text{ 个人在一年内死亡} \\ 0 & \text{第 } i \text{ 个人在一年内活着} \end{cases}$$

则 $P(\xi_i = 1) = 0.006 = p$, 记 $\eta_n = \sum_{i=1}^n \xi_i$, $n=10000$ 已足够大, 于是由中心极限定理

可得欲求事件的概率为

(1)

$$P(\eta_n > 120) = 1 - P\left(\frac{\eta_n - np}{\sqrt{npq}} \leq \frac{120 - np}{\sqrt{npq}} = b\right) \approx 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^b e^{-\frac{x^2}{2}} dx \approx 0 \quad (\text{其中 } b \approx \frac{60}{7.723})$$

同理可求得

(2)

$$P(\eta_n \leq 80) \approx 0.995 \quad (\text{对应的 } b \approx 2.59)$$

$$P(\eta_n \leq 60) \approx 0.5 \quad (\text{对应的 } b = 0)$$

$$P(\eta_n \leq 40) \approx 0.005 \quad (\text{对应的 } b \approx -2.59)$$

4.35 有一批种子, 其中良种占 $\frac{1}{6}$, 从中任取 6000 粒, 问能以 0.99 的概率保证

其中良种的比例与 $\frac{1}{6}$ 相差多少?

解: 令

$$\xi_i = \begin{cases} 1 & \text{第 } i \text{ 粒为良种} \\ 0 & \text{第 } i \text{ 粒不是良种} \end{cases}$$

则 $P(\xi_i = 1) = \frac{1}{6}$, 记 $p = \frac{1}{6}$, $\eta_n = \sum_{i=1}^n \xi_i$, 其中 $n = 6000$, 据题意即要求 α 使满足

$$P\left(\left|\frac{\eta_n}{n} - \frac{1}{6}\right| \leq \alpha\right) \geq 0.99. \quad \text{令 } q = 1 - p, b = \frac{n\alpha}{\sqrt{npq}}, \text{ 因为 } n \text{ 很大, 由中心极限定理有}$$

$$P(|\frac{\eta_n}{n} - \frac{1}{6}| \leq \alpha) = P(-b \leq \frac{\eta_n - np}{\sqrt{npq}} \leq b) \approx \frac{1}{\sqrt{2\pi}} \int_{-b}^b e^{-\frac{x^2}{2}} dx \geq 0.99$$

由 $N(0,1)$ 分布表知当 $b=2.60$ 时即能满足上述不等式，于是知

$\alpha = \frac{b}{n} \sqrt{npq} \approx 1.25 \times 10^{-4}$ ，即能以 0.99 的概率保证其中良种的比例与 $\frac{1}{6}$ 相差不超

过 1.25×10^{-4} 。

4.36 若某产品的不合格率为 0.005，任取 10000 件，问不合格品不多于 70 件的概率等于多少？

解：令

$$\xi_i = \begin{cases} 1 & \text{第 } i \text{ 件为不合格品} \\ 0 & \text{第 } i \text{ 件为合格品} \end{cases}$$

则 $p = P(\xi_i = 1) = 0.005$ ，记 $q = 1 - p$ ， $\eta_n = \sum_{i=1}^n \xi_i$ ，其中 $n = 10000$ ，记 $b = \frac{70 - np}{\sqrt{npq}}$ ，

由中心极限定理有

$$P(\eta_n \leq 70) = P(\frac{\eta_n - np}{\sqrt{npq}} \leq b) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^b e^{-\frac{x^2}{2}} dx \approx 0.998$$

即不合格品不多于 70 件的概率约等于 0.998。

4.37 某螺丝钉厂的不合格品率为 0.01，问一盒中应装多少只螺丝钉才能使其中含有 100 只合格品的概率不小于 0.95？

解：令

$$\xi_i = \begin{cases} 1 & \text{第 } i \text{ 只是合格品} \\ 0 & \text{第 } i \text{ 只是不合格品} \end{cases}$$

则 $p = P(\xi_i = 1) = 0.99$ ，记 $q = 1 - p$ ， $b = \frac{100 - np}{\sqrt{npq}}$ ， $\eta_n = \sum_{i=1}^n \xi_i$ ，其中 n 尚待确定，它

应满足 $P(\eta_n < 100) \leq 0.05$ ，由中心极限定理有

$$P(\eta_n < 100) = P(\frac{\eta_n - np}{\sqrt{npq}} < b) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^b e^{-\frac{x^2}{2}} dx \leq 0.05$$

查 $N(0,1)$ 分布表可取 $b = -1.65$ ，由此求得 $n = 103$ ，即在一盒中应装 103 只螺丝钉才能使其中含有 100 只合格品的概率不小于 0.95。

4.39 用特征函数的方法证明“二项分布收敛于普哇松分布”的普哇松定理。

证：设 $\{\xi_i^n\}_{1 \leq i \leq n}$ 独立同二项分布，即

$$P(\xi_i^n = 1) = p_n, P(\xi_i^n = 0) = q_n = 1 - p_n, 1 \leq i \leq n$$

ξ_i^n 的特征函数为 $(q_n + p_n e^{it})$, 记 $\eta_n = \sum_{i=1}^n \xi_i^n$, η_n 的特征函数记作 $\varphi_n(t)$, 因为

$np_n \rightarrow \lambda$, 故 $p_n = \frac{\lambda}{n} + o(\frac{1}{n})$, $q_n = 1 - \frac{\lambda}{n} + o(\frac{1}{n})$, 于是有

$$\begin{aligned}\varphi_n(t) &= (q_n + p_n e^{it})^n = \left(1 - \frac{\lambda}{n} + \frac{\lambda}{n} e^{it} + o(1)\right)^n \\ &= \left[1 + \frac{1}{n} \lambda (e^{it} - 1) + o\left(\frac{1}{n}\right)\right]^{\frac{n}{\lambda(e^{it}-1)} \lambda(e^{it}-1)} \\ &\rightarrow e^{\lambda(e^{it}-1)}, n \rightarrow \infty\end{aligned}$$

而 $e^{\lambda(e^{it}-1)}$ 是参数为 λ 的普哇松分布的特征函数, 由特征函数的逆极限定理即知定理成立, 证毕。

4.40 设随机变量 ξ_α 服从 Γ ---分布, 其分布密度为

$$p_\alpha(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (\alpha > 0, \beta > 0)$$

证: 当 $\alpha \rightarrow \infty$ 时, $\frac{\beta \xi_\alpha - \alpha}{\sqrt{\alpha}}$ 的分布函数弱收敛于 $N(0,1)$ 分布。

证: ξ_α 的特征函数为 $\varphi_\alpha(t) = (1 - \frac{it}{\beta})^{-\alpha}$, 易知 $\frac{\beta \xi_\alpha - \alpha}{\sqrt{\alpha}}$ 的特征函数为

$$g_\alpha(t) = e^{-i\sqrt{\alpha}t} (1 - \frac{it}{\sqrt{\alpha}})^{-\alpha} = e^{-i\sqrt{\alpha}t - \alpha \ln(1 - \frac{it}{\sqrt{\alpha}})}$$

而

$$\ln(1 - \frac{it}{\sqrt{\alpha}}) = -\frac{it}{\sqrt{\alpha}} + \frac{1}{2} \frac{t^2}{\alpha} - \frac{1}{3} \frac{it^3}{\alpha\sqrt{\alpha}} + o(\frac{t^3}{\alpha\sqrt{\alpha}})$$

因而有

$$-i\sqrt{\alpha}t - \alpha \ln(1 - \frac{it}{\sqrt{\alpha}}) = -\frac{t^2}{2} + \frac{1}{3} \frac{it^3}{\sqrt{\alpha}} + \alpha o(\frac{t^3}{\alpha\sqrt{\alpha}}) \rightarrow -\frac{t^2}{2}, \alpha \rightarrow \infty$$

故 $\lim_{\alpha \rightarrow \infty} g_\alpha(t) = e^{-\frac{t^2}{2}}$, 所以相应的分布函数弱收敛于 $N(0,1)$ 分布, 命题得证。

4.41 设 $\{\xi_n\}$ 为一列独立同分布随机变量, 且 ξ_n 服从 $(-n, n)$ 上的均匀分布, 证明

对 $\{\xi_n\}$ 成立中心极限定理。

证：易知 $E\xi_n = 0, D\xi_n = E\xi_n^2 = \int_{-n}^n \frac{x^2}{2n} dx = \frac{n^2}{3}$ ，于是

$$B_n^2 = \sum_{k=1}^n D\xi_k = \sum_{k=1}^n \frac{k^2}{3} = \frac{1}{18} n(n+1)(2n+1)$$

故 $B_n > \frac{n^{3/2}}{3}$ ，对任意的 $\tau > 0$ ，存在 N ，使当 $n \geq N$ 时有 $\frac{\sqrt{n}}{3}\tau > 1$ ，因而 $B_n\tau > n$ ，

从而当 $n \geq N$ ， $\int_{|x| > B_n\tau} x^2 dF_k(x) = 0$ ，若 $k \leq n$ ，由此知

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^2} \sum_{k=1}^n \int_{|x| > B_n\tau} x^2 dF_k(x) = 0$$

即林德贝尔格条件满足，所以对 $\{\xi_n\}$ 成立中心极限定理，结论得证。

4.42 设 $\{\xi_n\}, \{\eta_n\}$ 皆为独立同分布随机变量序列，且 $\{\xi_n\}$ 与 $\{\eta_n\}$ 独立，其中

$E\xi_n = 0, D\xi_n = 1; P(\eta_n = \pm 1) = \frac{1}{2}, n = 1, 2, \dots$ ，证明： $s_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \xi_i \eta_i$ 的分布函数弱

收敛于正态分布 $N(0,1)$ 。

证：这时 $\{\xi_n \eta_n\}$ 仍是独立同分布随机变量序列，易知有

$$E(\xi_n \eta_n) = 0, D(\xi_n \eta_n) = E(\xi_n \eta_n)^2 = E\xi_n^2 = 1$$

由林德贝尔格——勒维中心极限定理知： $s_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \xi_i \eta_i$ 的分布函数弱收敛于正态

分布 $N(0,1)$ ，结论得证。

4.45 利用中心极限定理证明：

$$\left(\sum_{k=0}^n \frac{n^k}{k!} \right) e^{-n} \rightarrow \frac{1}{2}, n \rightarrow \infty$$

证：设 $\{\xi_n\}$ 是独立同分布随机变量序列，共同分布为 $\lambda = 1$ 的 Poisson 分布，故

$E\xi_n = D\xi_n = 1, B_n^2 = \sum_{k=1}^n D\xi_k = n$ ，由林德贝尔格——勒维中心极限定理知

$$P\left(\sum_{k=1}^n \xi_k < n\right) = P\left(\frac{\sum_{k=1}^n (\xi_k - E\xi_k)}{B_n} < 0\right) \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-\frac{t^2}{2}} dt = \frac{1}{2}, n \rightarrow \infty$$

由 Poisson 分布的可加性知 $\sum_{k=1}^n \xi_k$ 服从参数为 n 的 Poisson 分布, 因而

$$P\left(\sum_{k=1}^n \xi_k < n\right) = \sum_{k=0}^{n-1} \frac{n^k}{k!} e^{-n}, \text{ 但 } \frac{n^n}{n!} e^{-n} \rightarrow 0 \quad (n \rightarrow \infty), \text{ 所以}$$

$$\left(\sum_{k=0}^n \frac{n^k}{k!}\right) e^{-n} = P\left(\sum_{k=1}^n \xi_k < n\right) + \frac{n^n}{n!} e^{-n} \rightarrow \frac{1}{2}, n \rightarrow \infty$$

成立, 结论得证。

第五章习题

1. 设 X_1, X_2, \dots, X_6 是来自服从参数为 λ 的泊松分布 $P(\lambda)$ 的样本, 试写出样本的联合分布律。

2. 设 X_1, X_2, \dots, X_6 是来自 $(0, \theta)$ 上的均匀分布的样本, $\theta > 0$ 未知

(1) 写出样本的联合密度函数;

(2) 指出下列样本函数中哪些是统计量, 哪些不是? 为什么?

$$T_1 = \frac{X_1 + X_2 + \dots + X_6}{6}, T_2 = X_6 - \theta, T_3 = X_6 - E(X_1), T_4 = \max(X_1, X_2, \dots, X_6)$$

(3) 设样本的一组观察是: 0.5, 1, 0.7, 0.6, 1, 1, 写出样本均值、样本方差和标准差。

3. 查表求 $\chi_{0.99}^2(12)$, $\chi_{0.01}^2(12)$, $t_{0.99}(12)$, $t_{0.01}(12)$ 。

4. 设 $T \sim t(10)$, 求常数 c , 使 $P(T > c) = 0.95$ 。

5. 设 X_1, X_2, \dots, X_n 是来自正态总体 $N(0, \sigma^2)$ 的样本, 试证:

$$(1) \frac{1}{\sigma^2} \sum_{i=1}^n X_i^2 \sim \chi^2(n);$$

$$(2) \frac{1}{n\sigma^2} \left(\sum_{i=1}^n X_i\right)^2 \sim \chi^2(1).$$

6. 设 X_1, X_2, \dots, X_5 是独立且服从相同分布的随机变量, 且每一个 $X_i (i=1, 2, \dots, 5)$ 都服从 $N(0, 1)$ 。

(1) 试给出常数 c , 使得 $c(X_1^2 + X_2^2)$ 服从 χ^2 分布, 并指出它的自由度;

(2) 试给出常数 d , 使得 $d \frac{X_1 + X_2}{\sqrt{X_3^2 + X_4^2 + X_5^2}}$ 服从 t 分布, 并指出它的自由度。

7. 设 (X_1, X_2, \dots, X_n) 是取自总体 X 的一个样本, 在下列三种情况下, 分别求 $E(\bar{X}), D(\bar{X}), E(S^2)$: (1) $X \sim B(1, p)$; (2) $X \sim E(\lambda)$; (3) $X \sim R(0, 2\theta)$, 其中 $\theta > 0$ 。

8. 某市有 100000 个年满 18 岁的居民, 他们中 10% 年收入超过 1 万, 20% 受过高等教育。今从中抽取 1600 人的随机样本, 求:

- (1) 样本中不少于 11% 的人年收入超过 1 万的概率;
- (2) 样本中 19% 和 21% 之间的人受过高等教育的概率。

9. 设总体 $X \sim N(40, 5^2)$, (1) 抽取容量为 36 的样本, 求 $P\{38 \leq \bar{X} \leq 43\}$;
 (2) 抽取容量为 64 的样本, 求 $P\{|\bar{X} - 40| < 1\}$;
 (3) 取样本容量 n 多大时, 才能使 $P\{|\bar{X} - 40| < 1\} = 0.95$ 。

10. 设总体 $X \sim N(\mu, \sigma^2)$, μ, σ^2 皆未知, 已知样本容量 $n=16$, 样本均值 $\bar{x}=12.5$, 修正样本方差 $s^{*2}=5.333$, 求 $P\{|\bar{X} - \mu| < 0.4\}$ 。

11. 设是 $X_1, X_2, \dots, X_n, X_{n+1}, \dots, X_{n+m}$ 来自正态总体 $N(0, \sigma^2)$, 容量为 $n+m$ 的样本, 求下列统计量的抽样分布:

$$(1) Y = \frac{1}{\sigma^2} \sum_{i=1}^{n+m} X_i^2;$$

$$(2) Z = \frac{\sqrt{m} \sum_{i=1}^n X_i}{\sqrt{n} \sqrt{\sum_{i=n+1}^{n+m} X_i^2}};$$

$$(3) F = \frac{m \sum_{i=1}^n X_i^2}{n \sum_{i=n+1}^{n+m} X_i^2}.$$

12. 若 $T \sim t(n)$, 则 T^2 服从什么分布?

13. 设 X_1, X_2, \dots, X_n 是来自泊松分布 $P(\lambda)$ 的一个样本, \bar{X} 与 S^2 分别为样本均值与样本方差, 试求 $E(\bar{X}), Var(\bar{X}), E(S^2)$ 。

14. 某区有 25000 户家庭, 10% 的家庭没有汽车, 今有 1600 户家庭的随机样本, 试求: 9%~11% 之间的样本家庭没有汽车的概率。

习题解答

1. 解

$$\begin{aligned}
 f(x_1, x_2, \dots, x_6) &= e^{-\lambda} \frac{\lambda^{x_1}}{x_1!} \cdot e^{-\lambda} \frac{\lambda^{x_2}}{x_2!} \cdots e^{-\lambda} \frac{\lambda^{x_6}}{x_6!} \\
 &= e^{-6\lambda} \frac{\lambda^{\sum_{i=1}^6 x_i}}{\prod_{i=1}^6 x_i!} \quad x_1, x_2, \dots, x_6 = 0, 1, 2, \dots
 \end{aligned}$$

2. 解

$$(1) \quad f(x_1, x_2, \dots, x_6) = \begin{cases} \theta^{-6} & 0 < x_1, x_2, \dots, x_6 < \theta \\ 0 & \text{其他} \end{cases}$$

(2) T_1 和 T_4 是, T_2 和 T_3 不是。因为 T_1 和 T_4 中不含总体中的唯一未知参数 θ , 而 T_2 和 T_3 中含有未知参数 θ 。

$$(3) \quad \text{样本均值} \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{6} \sum_{i=1}^6 X_i = \frac{1}{6} (0.5 + 1 + 0.7 + 0.6 + 1 + 1) = 0.8$$

$$\begin{aligned}
 \text{样本方差} \quad S^2 &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{6} \sum_{i=1}^6 (X_i - \bar{X})^2 \\
 &= \frac{1}{6} [(-0.3)^2 + (0.2)^2 + (-0.1)^2 + (-0.2)^2 + (0.2)^2 + (0.2)^2] = 0.0433
 \end{aligned}$$

$$\text{样本标准差} \quad S = \sqrt{S^2} = \sqrt{0.0433} = 0.2082。$$

$$3. \quad \text{解} \quad \chi_{0.99}^2(12) = 26.217, \quad \chi_{0.01}^2(12) = 3.571, \quad t_{0.99}(12) = 2.6810, \\
 t_{0.01}(12) = -2.6810。$$

4. 解 由 t 分布关于纵轴对称, 所以 $P(T > c) = 0.95$ 即为 $P(T > -c) = 0.05$ 。
由附表 5.6 可查得 $-c = 1.81$, 所以 $c = -1.81$ 。

5. 证明:

(1) $\frac{X_i}{\sigma}$ 独立同分布于 $N(0,1)$, 由 χ^2 分布的定义, $\sum_{i=1}^n \left(\frac{X_i}{\sigma}\right)^2 \sim \chi^2(n)$, 即 $\frac{1}{\sigma^2} \sum_{i=1}^n X_i^2 \sim \chi^2(n)$ 。

(2) 易见, $\sum_{i=1}^n X_i \sim N(0, n\sigma^2)$, 即 $\frac{\sum_{i=1}^n X_i}{\sqrt{n\sigma^2}} \sim N(0,1)$, 由 χ^2 分布的定义,

$$\left(\frac{\sum_{i=1}^n X_i}{\sqrt{n\sigma^2}} \right)^2 \sim \chi^2(1), \quad \text{即} \quad \frac{1}{n\sigma^2} \left(\sum_{i=1}^n X_i \right)^2 \sim \chi^2(1)。$$

6. 解

(1) 易见, $X_1^2 + X_2^2$ 即为二个独立的服从 $N(0,1)$ 的随机变量平方和, 服从 $\chi^2(2)$ 分布, 即 $c=1$; 自由度为 2。

(2) 由于 $X_1 + X_2 \sim N(0,2)$, 则 $\frac{X_1 + X_2}{\sqrt{2}} \sim N(0,1)$ 。

又 $X_3^2 + X_4^2 + X_5^2 \sim \chi^2(3)$, $\frac{X_1 + X_2}{\sqrt{2}}$ 与 $X_3^2 + X_4^2 + X_5^2$ 相互独立, 则

$$\frac{(X_1 + X_2)/\sqrt{2}}{\sqrt{(X_3^2 + X_4^2 + X_5^2)/3}} \sim t(3)$$

$$\frac{\sqrt{6}}{2} \frac{X_1 + X_2}{\sqrt{X_3^2 + X_4^2 + X_5^2}} \sim t(3)$$

即

即 $d = \frac{\sqrt{6}}{2}$, 自由度为 3。

7. 解

(1) $X \sim B(1, p)$

$$E(X) = p, E(X^2) = p, D(X) = p(1-p)$$

$$E(\bar{X}) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = p$$

$$D(\bar{X}) = D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n D(X_i) = \frac{p(1-p)}{n}$$

$$\begin{aligned} E(S^2) &= E\left(\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2\right) = \frac{1}{n} E\left(\sum_{i=1}^n X_i^2 - n\bar{X}^2\right) = \frac{1}{n} \left(\sum_{i=1}^n E(X_i^2) - nE(\bar{X}^2)\right) \\ &= \frac{1}{n} \left[\sum_{i=1}^n \left(D(X_i) + (E(X_i))^2\right) - n\left(D(\bar{X}) + (E(\bar{X}))^2\right)\right] \\ &= \frac{1}{n} \left[np - n\left(\frac{p(1-p)}{n} + p^2\right) \right] \\ &= \left(1 - \frac{1}{n}\right)p(1-p) \end{aligned}$$

(2) $X \sim E(\lambda)$

$$E(X) = \frac{1}{\lambda}, D(X) = \frac{1}{\lambda^2},$$

$$E(\bar{X}) = \frac{1}{\lambda}$$

$$D(\bar{X}) = \frac{1}{n\lambda^2}$$

$$E(S^2) = \frac{1}{n} \left[\sum_{i=1}^n \{D(X_i) + (E(X_i))^2\} - n\{D(\bar{X}) + (E(\bar{X}))^2\} \right] = \left(1 - \frac{1}{n}\right) \frac{1}{\lambda^2}$$

(3) $X \sim R(0, 2\theta)$, 其中 $\theta > 0$

$$E(X) = \theta$$

$$D(X) = \frac{\theta^2}{3}$$

$$E(\bar{X}) = \theta$$

$$D(\bar{X}) = \frac{\theta^2}{3n}$$

$$E(S^2) = \frac{1}{n} \left[\sum_{i=1}^n \{D(X_i) + (E(X_i))^2\} - n\{D(\bar{X}) + (E(\bar{X}))^2\} \right] = \left(1 - \frac{1}{n}\right) \frac{\theta^2}{3}$$

8. 解

(1) 引入新变量:

$$\begin{cases} X_i = & 1, \text{第 } i \text{ 个样本居民年收入超过 1 万} \\ & 0, \text{第 } i \text{ 个样本居民年收入没超过 1 万} \end{cases}$$

其中 $i = 1, 2, \dots, n, n = 1600$

易见: $p = P(X_i = 1) = 0.1$

又因 $n = 1600 \ll N = 100000$, 故可以近似看成有放回抽样, X_1, X_2, \dots, X_n 相互独立。

$$\mu = E(X_i) = 0.1, \sigma = \sqrt{D(X_i)} = \sqrt{0.1 \times 0.9} = 0.3$$

样本中年收入超过 1 万的比例即为 \bar{X} , 由于 $n = 1600$ 较大, 可以使用渐近分布求

解, 即 $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, 所求概率即为

$$\begin{aligned} P(\bar{X} \geq 11\%) &= 1 - P(\bar{X} \leq 0.11) = 1 - P\left(\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \leq \frac{40(0.11 - 0.1)}{0.3}\right) \\ &= 1 - \Phi\left(\frac{4}{3}\right) = 1 - 0.9082 = 0.0918 \end{aligned}$$

(2) 同 (1) 解法

引入新变量:

$$\begin{cases} X_i = & 1, \text{第 } i \text{ 个样本居民受过高等教育} \\ & 0, \text{第 } i \text{ 个样本居民未受过高等教育} \end{cases}$$

其中 $i = 1, 2, \dots, n, n = 1600$

$$p = P(X_i = 1) = 0.2$$

$$\mu = 0.2, \sigma = \sqrt{0.2 \times 0.8} = 0.4$$

$$P(19\% \leq \bar{X} \leq 21\%) = P\left(\frac{40(0.19 - 0.2)}{0.4} \leq \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \leq \frac{40(0.21 - 0.2)}{0.4}\right) \\ = \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 = 2 \times 0.8413 - 1 = 0.6826$$

答：(1) 样本中不少于 11% 的人年收入超过 1 万的概率为 0.0918；

(2) 样本中 19% 和 21% 之间的人受过高等教育的概率为 0.6826。

9. 0.9916, 0.8904, 96。

10. 0.5。

11. (1) $\chi^2(n+m)$; (2) $t(m)$; (3) $F(n, m)$ 。

12. $F(1, n)$ 。

13. $\lambda, \frac{\lambda}{n}, \left(1 - \frac{1}{n}\right)\lambda$ 。

14. 0.8164。

第六章习题

1. 设 X_1, X_2, \dots, X_n 是取自总体 X 的一个样本，在下列情形下，试求总体参数的矩估计与最大似然估计：

(1) $X \sim B(n, p)$ ，其中 p 未知， $0 < p < 1$ ；

(2) $X \sim E(\lambda)$ ，其中 λ 未知， $\lambda > 0$ 。

2. 设 X_1, X_2, \dots, X_n 是取自总体 X 的一个样本，其中 X 服从参数为 λ 的泊松分布，其中 λ 未知， $\lambda > 0$ ，求 λ 的矩估计与最大似然估计，如得到一组样本观测值

X	0	1	2	3	4
频数	17	20	10	2	1

求 λ 的矩估计值与最大似然估计值。

3. 设 X_1, X_2, \dots, X_n 是取自总体 X 的一个样本，其中 X 服从区间 $(0, \theta)$ 的均匀分布，其中 $\theta > 0$ 未知，求 θ 的矩估计。

4. 设 X_1, X_2, \dots, X_n 是取自总体 X 的一个样本， X 的密度函数为

$$f(x) = \begin{cases} \frac{2x}{\theta^2} & 0 < x < \theta \\ 0 & \text{其他} \end{cases}$$

其中 $\theta > 0$ 未知，求 θ 的矩估计。

5. 设 X_1, X_2, \dots, X_n 是取自总体 X 的一个样本， X 的密度函数为

$$f(x) = \begin{cases} (\theta+1)x^\theta & 0 < x < 1 \\ 0 & \text{其他} \end{cases}$$

其中 $\theta > 0$ 未知，求 θ 的矩估计和最大似然估计。

6. 设 X_1, X_2, \dots, X_n 是取自总体 X 的一个样本, 总体 X 服从参数为 p 的几何分布, 即 $P(X=x) = p(1-p)^{x-1}, (x=1, 2, 3, \dots)$, 其中 p 未知, $0 < p < 1$, 求 p 的最大似然估计。

7. 已知某路口车辆经过的时间间隔服从指数分布 $E(\lambda)$, 其中 $\lambda > 0$ 未知, 现在观测到六个时间间隔数据 (单位: s): 1.8, 3.2, 4.8, 4.5, 2.5, 试求该路口车辆经过的平均时间间隔的矩估计值与最大似然估计值。

8. 设总体 X 的密度函数为 $f(x, \sigma) = \frac{1}{2\sigma} e^{-\frac{|x|}{\sigma}}, (-\infty < x < +\infty)$, 其中 $\sigma > 0$ 未知, 设 X_1, X_2, \dots, X_n 是取自这个总体的一个样本, 试求 σ 的最大似然估计。

9. 在第 3 题中 θ 的矩估计是否是 θ 的无偏估计?

$$\text{解 } E(\hat{\theta}) = E(2\bar{X}) = 2E(\bar{X}) = 2E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{2}{n} \sum_{i=1}^n E(X_i) = \frac{2}{n} \sum_{i=1}^n \frac{\theta}{2} = \theta$$

故 θ 的矩估计量 $2\bar{X}$ 是 θ 的无偏估计。

10. 试证第 8 题中 σ 的最大似然估计是 σ 的无偏估计。

11. 设 X_1, X_2, X_3 为总体 $X \sim N(\mu, \sigma^2)$ 的样本, 证明

$$\hat{\mu}_1 = \frac{1}{6} X_1 + \frac{1}{3} X_2 + \frac{1}{2} X_3$$

$$\hat{\mu}_2 = \frac{2}{5} X_1 + \frac{1}{5} X_2 + \frac{2}{5} X_3$$

都是总体均值 μ 的无偏估计, 并进一步判断哪一个估计有效。

12. 设 X_1, X_2, \dots, X_n 是取自总体 $X \sim N(0, \sigma^2)$ 的一个样本, 其中 $\sigma^2 > 0$ 未知, 令 $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n X_i^2$, 试证 $\hat{\sigma}^2$ 是 σ^2 的相合估计。

13. 某车间生产滚珠, 从长期实践中知道, 滚珠直径 X 服从正态分布 $N(\mu, 0.2^2)$, 从某天生产的产品中随机抽取 6 个, 量得直径如下 (单位: mm): 14.7, 15.0, 14.9, 14.8, 15.2, 15.1, 求 μ 的 0.9 双侧置信区间和 0.99 双侧置信区间。

14. 假定某商店中一种商品的月销售量服从正态分布 $N(\mu, \sigma^2)$, σ 未知。为了合理的确定对该商品的进货量, 需对 μ 和 σ 作估计, 为此随机抽取七个月, 其销售量分别为: 64, 57, 49, 81, 76, 70, 59, 试求 μ 的双侧 0.95 置信区间和方差 σ^2 的双侧 0.9 置信区间。

15. 随机地取某种子弹 9 发作试验, 测得子弹速度的 $\bar{s} = 11$, 设子弹速度服从正态分布 $N(\mu, \sigma^2)$, 求这种子弹速度的标准差 σ 和方差 σ^2 的双侧 0.95 置信区间。

16. 已知某炼铁厂的铁水含碳量 (1%) 正常情况下服从正态分布 $N(\mu, \sigma^2)$, 且标准差 $\sigma = 0.108$ 。现测量五炉铁水, 其含碳量分别是: 4.28, 4.4, 4.42, 4.35, 4.37 (1%), 试求未知参数 μ 的单侧置信水平为 0.95 的置信下限和置信上限。

17. 某单位职工每天的医疗费服从正态分布 $N(\mu, \sigma^2)$, 现抽查了 25 天, 得

$\bar{x} = 170$ 元, $s^* = 30$ 元, 求职工每天医疗费均值 μ 的双侧 0.95 置信区间。

18. 某食品加工厂有甲乙两条加工猪肉罐头的生产线。设罐头质量服从正态分布并假设甲生产线与乙生产线互不影响。从甲生产线并假设抽取 10 只管头测得其平均质量 $\bar{x} = 501g$, 已知其总体标准差 $\sigma_1 = 5g$; 从乙生产线抽取 20 只罐头测得其平均质量 $\bar{y} = 498g$, 已知其总体标准差 $\sigma_2 = 4g$, 求甲乙两条猪肉罐头生产线生产罐头质量的均值差 $\mu_1 - \mu_2$ 的双侧 0.99 置信区间。

19. 为了比较甲、乙两种显像管的使用寿命 X 和 Y , 随机的抽取甲、乙两种显像管各 10 只, 得数据 x_1, \dots, x_{10} 和 y_1, \dots, y_{10} (单位: $10^4 h$), 且由此算得

$\bar{x} = 2.33, \bar{y} = 0.75, \sum_{i=1}^{10} (x_i - \bar{x})^2 = 27.5, \sum_{i=1}^{10} (y_i - \bar{y})^2 = 19.2$, 假定两种显像管的使用寿命均服从正态分布, 且由生产过程知道它们的方差相等。试求两个总体均值之差 $\mu_1 - \mu_2$ 的双侧 0.95 置信区间。

20. 在 3091 个男生, 3581 个女生组成的总体中, 随机不放回地抽取 100 人, 观察其中男生的成数, 要求计算样本中男生成数的 SE。

21. 抽取 1000 人的随机样本估计一个大的人口总体中拥有私人汽车的人的百分数, 样本中有 543 人拥有私人汽车, (1) 求样本中拥有私人汽车的人的百分数的 SE; (2) 求总体中拥有私人汽车的人的百分数的 95% 的置信区间。

习题解答

1. 解 (1) $E(X) = p$, 故 p 的矩估计量有 $\hat{p} = \bar{X}$ 。

另, X 的分布律为 $P(X = x) = p^x (1-p)^{1-x}, x = 0, 1$, 故似然函数为

$$L(p) = p^{\sum_{i=1}^n X_i} (1-p)^{n - \sum_{i=1}^n X_i}$$

对数似然函数为:

$$\ln L(p) = \left(\sum_{i=1}^n X_i \right) \ln p + \left(n - \sum_{i=1}^n X_i \right) \ln(1-p)$$

$$\frac{d \ln L(p)}{dp} = \frac{\sum_{i=1}^n X_i}{p} - \frac{n - \sum_{i=1}^n X_i}{1-p} = 0$$

令

$$\hat{p} = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$$

解得 p 的最大似然估计量

可以看出 p 的矩估计量与最大似然估计量是相同的。

(2) $E(X) = \frac{1}{\lambda}$, 令 $\frac{1}{\lambda} = \bar{X}$, 故 λ 的矩估计量 $\hat{\lambda} = \frac{1}{\bar{X}}$ 。

另, X 的密度函数为

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

故似然函数为

$$L(\lambda) = \begin{cases} \lambda^n e^{-\lambda \sum_{i=1}^n X_i} & X_i > 0, i = 1, 2, \dots, n \\ 0 & \text{其他} \end{cases}$$

对数似然函数为

$$\begin{aligned} \ln L(\lambda) &= n \ln \lambda - \lambda \sum_{i=1}^n X_i \\ \frac{d \ln L(\lambda)}{d\lambda} &= \frac{n}{\lambda} - \sum_{i=1}^n X_i = 0 \\ \hat{\lambda} &= \frac{n}{\sum_{i=1}^n X_i} = \frac{1}{\bar{X}} \end{aligned}$$

解得 λ 的最大似然估计量

可以看出 λ 的矩估计量与最大似然估计量是相同的。

2. 解 $E(X) = \lambda$, 故 λ 的矩估计量 $\hat{\lambda} = \bar{X}$ 。

由样本观测值可算得

$$\bar{X} = \frac{0 \times 17 + 1 \times 20 + 2 \times 10 + 3 \times 2 + 4 \times 1}{50} = 1$$

另, X 的分布律为

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

故似然函数为

$$L(\lambda) = e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^n X_i}}{X_1! \cdots X_n!}, X_i = 0, 1, 2, \dots, i = 1, 2, \dots, n$$

对数似然函数为

$$\begin{aligned} \ln L(\lambda) &= -n\lambda + \left(\sum_{i=1}^n X_i \right) \ln \lambda - \sum_{i=1}^n \ln(X_i!) \\ \frac{d \ln L(\lambda)}{d\lambda} &= -n + \frac{\sum_{i=1}^n X_i}{\lambda} = 0 \\ \hat{\lambda} &= \frac{\sum_{i=1}^n X_i}{n} = \bar{X} \end{aligned}$$

解得 λ 的最大似然估计量

故 λ 的最大似然估计值 $\hat{\lambda} = 1$ 。

3. 解 $E(X) = \frac{\theta}{2}$, 令 $\frac{\theta}{2} = \bar{X}$, 故 θ 的矩估计量 $\hat{\theta} = 2\bar{X}$ 。

4. 解 $E(X) = \int_0^{\theta} x \cdot \frac{2x}{\theta^2} dx = \frac{2}{3} \theta$, 令 $\frac{2}{3} \theta = \bar{X}$, 故 θ 的矩估计量为 $\hat{\theta} = \frac{3}{2} \bar{X}$ 。

5. 解 $E(X) = \int_0^1 x \cdot (\theta+1)x^{\theta} dx = \frac{\theta+1}{\theta+2}$, 令 $\frac{\theta+1}{\theta+2} = \bar{X}$, 故 θ 的矩估计量为

$\hat{\theta} = \frac{1-2\bar{X}}{\bar{X}-1}$, 另, 似然函数

$$L(\theta) = \begin{cases} (\theta+1)^n \prod_{i=1}^n X_i^\theta & 0 < X_i < 1 \\ 0 & \text{其他} \end{cases}$$

对数似然函数为

$$\ln L(\theta) = n \ln(\theta+1) + \theta \sum_{i=1}^n \ln X_i$$

$$\frac{d \ln L(\theta)}{d\theta} = \frac{n}{\theta+1} + \sum_{i=1}^n \ln X_i = 0$$

$$\hat{\theta} = -1 - \frac{n}{\sum_{i=1}^n X_i} = -1 - \frac{1}{\bar{X}}$$

解得 θ 的最大似然估计量为

6. 解 似然函数 $L(p) = p^n (1-p)^{\sum_{i=1}^n X_i - n}$
对数似然函数

$$\ln L(p) = n \ln p + \left(\sum_{i=1}^n X_i - n \right) \ln(1-p)$$

$$\frac{d \ln L(p)}{dp} = \frac{n}{p} - \frac{\sum_{i=1}^n X_i - n}{1-p} = 0$$

解得 p 的最大似然估计量为 $\hat{p} = \frac{1}{\bar{X}}$ 。

7. 解 根据习题 1 的结果, λ 的矩估计和最大似然估计量都为 $\frac{1}{\bar{X}}$, 故平均时间

间隔的矩估计和最大似然估计都为 $\frac{1}{\hat{\lambda}}$, 即为 \bar{X} 。

由样本观测值可算得 $\bar{X} = \frac{1}{6}(1.8+3.2+4+8+4.5+2.5) = 4$ 。

$$L(\sigma) = \frac{1}{(2\sigma)^n} e^{-\frac{\sum_{i=1}^n |X_i|}{\sigma}}$$

8. 解 似然函数
对数似然函数为

$$\ln L(\sigma) = -n \ln(2\sigma) - \frac{1}{\sigma} \sum_{i=1}^n |X_i|$$

$$\frac{d \ln L(\sigma)}{d\sigma} = -\frac{n}{\sigma} + \frac{\sum_{i=1}^n |X_i|}{\sigma^2} = 0$$

得 σ 的最大似然估计量为 $\hat{\sigma} = \frac{1}{n} \sum_{i=1}^n |X_i|$ 。

9. 解 $E(\hat{\theta}) = E(2\bar{X}) = 2E(\bar{X}) = 2E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{2}{n} \sum_{i=1}^n E(X_i) = \frac{2}{n} \sum_{i=1}^n \frac{\theta}{2} = \theta$

故 θ 的矩估计量 $2\bar{X}$ 是 θ 的无偏估计。

$$10. \text{ 证明: } E(\hat{\sigma}) = E\left(\frac{1}{n} \sum_{i=1}^n |X_i|\right) = \frac{1}{n} \sum_{i=1}^n E(|X_i|)$$

$$= \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{+\infty} |x| \cdot \frac{1}{2\sigma} e^{-\frac{|x|}{\sigma}} dx = \frac{1}{n} \sum_{i=1}^n 2 \int_0^{+\infty} x \cdot \frac{1}{2\sigma} e^{-\frac{x}{\sigma}} dx = \sigma$$

故 σ 的最大似然估计 $\hat{\sigma} = \frac{1}{n} \sum_{i=1}^n |X_i|$ 是 σ 的无偏估计。

$$11. \text{ 证明 } E(\hat{\mu}_1) = E\left(\frac{1}{6}X_1 + \frac{1}{3}X_2 + \frac{1}{2}X_3\right)$$

$$= \frac{1}{6}E(X_1) + \frac{1}{3}E(X_2) + \frac{1}{2}E(X_3)$$

$$= \left(\frac{1}{6} + \frac{1}{3} + \frac{1}{2}\right)E(X) = E(X) = \mu$$

$$E(\hat{\mu}_2) = E\left(\frac{2}{5}X_1 + \frac{1}{5}X_2 + \frac{2}{5}X_3\right)$$

$$= \frac{2}{5}E(X_1) + \frac{1}{5}E(X_2) + \frac{2}{5}E(X_3)$$

$$= \left(\frac{2}{5} + \frac{1}{5} + \frac{2}{5}\right)E(X) = E(X) = \mu$$

所以 $\hat{\mu}_1, \hat{\mu}_2$ 都是总体均值 μ 的无偏估计。

$$\text{又 } D(\hat{\mu}_1) = D\left(\frac{X_1}{6} + \frac{X_2}{3} + \frac{X_3}{2}\right)$$

$$= \frac{1}{36}D(X_1) + \frac{1}{9}D(X_2) + \frac{1}{4}D(X_3)$$

$$= \left(\frac{1}{36} + \frac{1}{9} + \frac{1}{4}\right)D(X) = \frac{7}{18}D(X) = \frac{7}{18}\sigma^2$$

$$D(\hat{\mu}_2) = D\left(\frac{2}{5}X_1 + \frac{1}{5}X_2 + \frac{2}{5}X_3\right)$$

$$= \frac{4}{25}D(X_1) + \frac{1}{25}D(X_2) + \frac{4}{25}D(X_3)$$

$$= \frac{9}{25}D(X) = \frac{9}{25}\sigma^2$$

可见 $D(\hat{\mu}_2) < D(\hat{\mu}_1)$, 所以二个估计量中 $\hat{\mu}_2$ 更有效。

$$12. \text{ 证明 } E(\hat{\sigma}^2) = E\left(\frac{1}{n} \sum_{i=1}^n X_i^2\right) = \frac{1}{n} \sum_{i=1}^n E(X_i^2) = \sigma^2$$

$$\text{又 } \frac{1}{\sigma^2} \sum_{i=1}^n X_i^2 \sim \chi^2(n),$$

$$\text{由公式 (9), } D\left(\frac{1}{\sigma^2} \sum_{i=1}^n X_i^2\right) = 2n,$$

$$\text{故 } D(\hat{\sigma}^2) = D\left(\frac{1}{\sigma^2} \sum_{i=1}^n X_i^2\right) \cdot \frac{\sigma^4}{n^2} = \frac{2\sigma^4}{n}.$$

由切比雪夫不等式, 当 $n \rightarrow \infty$, 对任给 $\varepsilon > 0$,

$$P(|\hat{\sigma}^2 - \sigma^2| > \varepsilon) \leq \frac{D(\hat{\sigma}^2)}{\varepsilon^2} = \frac{2\sigma^4}{n\varepsilon^2} \rightarrow 0$$

即 $\hat{\sigma}^2$ 是 σ^2 的相合估计。

13. 解 由于 $\sigma^2 = 0.2^2$ 已知, 所以选用 μ 的 $1-\alpha$ 置信区间 $\left[\bar{X} - u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$ 。

当 $1-\alpha=0.9$, 查表得 $u_{1-\frac{\alpha}{2}} = u_{0.95} = 1.64$, 当 $1-\alpha=0.99$, 查表得 $u_{1-\frac{\alpha}{2}} = u_{0.995} = 2.576$ 。 $\bar{x} = 14.95, n = 6$,

代入数据得 μ 的双侧 0.9 置信区间观测值为 $\left[14.95 - 1.64 \cdot \frac{0.2}{\sqrt{6}}, 14.95 + 1.64 \cdot \frac{0.2}{\sqrt{6}} \right]$, 即为 $[14.82, 15.08]$ 。

μ 的双侧 0.99 置信区间观测值为 $\left[14.95 - 2.576 \cdot \frac{0.2}{\sqrt{6}}, 14.95 + 2.576 \cdot \frac{0.2}{\sqrt{6}} \right]$, 即为 $[14.74, 15.16]$ 。

14. 解 由于 μ 和 σ 都未知, 故 μ 的 $1-\alpha$ 双侧置信区间为

$$\left[\bar{X} - t_{1-\frac{\alpha}{2}}(n-1) \frac{S^*}{\sqrt{n}}, \bar{X} + t_{1-\frac{\alpha}{2}}(n-1) \frac{S^*}{\sqrt{n}} \right],$$

σ^2 的 $1-\alpha$ 双侧置信区间为

$$\left[\frac{nS^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)}, \frac{nS^2}{\chi_{\frac{\alpha}{2}}^2(n-1)} \right],$$

代入数据得

$$\bar{x} = 65.14, s^2 = 108.41, s^* = 11.25, t_{0.975}(6) = 2.45, n = 7, \chi_{0.95}^2(6) \chi_{0.05}^2(6) = 1.635,$$

μ 的 0.95 双侧置信区间观测值为 $\left[65.14 - 2.45 \times \frac{11.25}{\sqrt{7}}, 65.14 + 2.45 \times \frac{11.25}{\sqrt{7}} \right]$, 即为 $[54.74, 75.54]$ 。

σ^2 的 0.9 双侧置信区间观测值为 $\left[\frac{7 \times 108.41}{12.592}, \frac{7 \times 108.41}{1.635} \right]$, 即为 $[60.3, 464.14]$ 。

15. 解 由于 μ 未知, 故 σ^2 的双侧置信区间为 $\left[\frac{(n-1)S^{*2}}{\chi_{1-\frac{\alpha}{2}}^2(n-1)}, \frac{(n-1)S^{*2}}{\chi_{\frac{\alpha}{2}}^2(n-1)} \right]$, 代入数据得 $n = 9, S^{*2} = 121, \chi_{0.975}^2(8) = 17.535, \chi_{0.025}^2(8) = 2.18$,

σ^2 的 0.95 双侧置信区间观测值为 $\left[\frac{8 \times 121}{17.535}, \frac{8 \times 121}{2.18} \right]$, 即为 $[55.204, 444.037]$ 。故 σ 的 0.95 双侧置信区间观测值为 $[\sqrt{55.204}, \sqrt{444.037}]$, 即为 $[7.43, 21.07]$ 。

16. 解 由于 $\sigma = 0.108$ 已知, 故 μ 的 $1-\alpha$ 单侧置信下限为 $\bar{X} - u_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$, μ 的

$1-\alpha$ 单侧置信上限为 $\bar{X} + u_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$, 代入数据得 $\bar{x} = 4.364(\%), u_{0.95} = 1.645, n = 5$, 故

μ 的 0.95 单侧置信下限观测值为 $4.364 - 1.645 \cdot \frac{0.108}{\sqrt{5}} = 4.285$, μ 的 0.95 单侧置信

上限观测值为 $4.364 + 1.645 \cdot \frac{0.108}{\sqrt{5}} = 4.443$ 。

17. 解 由于 σ^2 未知, 故 μ 的 $1-\alpha$ 双侧置信区间为 $\left[\bar{X} - t_{1-\frac{\alpha}{2}} \frac{S^*}{\sqrt{n}}, \bar{X} + t_{1-\frac{\alpha}{2}} \frac{S^*}{\sqrt{n}} \right]$, 代入数据得 $\bar{x} = 170, s^* = 30, n = 25, t_{0.975}(24) = 2.0639$, 故 μ 的 0.95 双侧置信区间观测值为 $\left[170 - 2.0639 \frac{30}{\sqrt{24}}, 170 + 2.0639 \frac{30}{\sqrt{24}} \right]$, 即为 $[157.4, 182.6]$ 。

18. 解 由于 $\sigma_1 = 5g, \sigma_2 = 4g$ 已知, 故 $\mu_1 - \mu_2$ 的 $1-\alpha$ 的双侧置信区间为

$$\left[\bar{X} - \bar{Y} - u_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}, \bar{X} - \bar{Y} + u_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}} \right]$$

代入数据得 $\bar{x} = 501, \bar{y} = 498, m = 10, n = 20, \sigma_1^2 = 25, \sigma_2^2 = 16, u_{0.995} = 2.576$, 故 $\mu_1 - \mu_2$ 的 0.99 双侧置信区间观测值为 $\left[501 - 498 - 2.576 \sqrt{\frac{25}{10} + \frac{16}{20}}, 501 - 498 + 2.576 \sqrt{\frac{25}{10} + \frac{16}{20}} \right]$, 即为 $[-1.68, 7.68]$ 。

19. 解 由于 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 未知, 故 $\mu_1 - \mu_2$ 的 $1-\alpha$ 双侧置信区间为

$$\left[\bar{X} - \bar{Y} - t_{1-\frac{\alpha}{2}}(m+n-2)S_w \sqrt{\frac{1}{m} + \frac{1}{n}}, \bar{X} - \bar{Y} + t_{1-\frac{\alpha}{2}}(m+n-2)S_w \sqrt{\frac{1}{m} + \frac{1}{n}} \right]$$

其中 $S_w^2 = \frac{1}{m+n-2} \left[\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{i=1}^n (Y_i - \bar{Y})^2 \right]$,

代入数据得 $\bar{x} = 2.33, \bar{y} = 0.75, m = 10 = n, s_w = 1.611, t_{0.975}(18) = 2.1009$, 故 $\mu_1 - \mu_2$ 的 0.95 双侧置信区间观测值为

$$\left[2.33 - 0.75 - 2.1009 \times 1.611 \sqrt{\frac{1}{10} + \frac{1}{10}}, 2.33 - 0.75 + 2.1009 \times 1.611 \sqrt{\frac{1}{10} + \frac{1}{10}} \right],$$

即为 $[0.066, 3.094]$ 。

20. 解 由于样本大小 $n = 100$ 相对于总体容量 $N = 6672$ 来说很小, 因此可使用有放回抽样的公式。

样本成数 $\bar{x} = 100 \times \frac{3091}{6672} \approx 46$, 估计 $\hat{\sigma} = \sqrt{46 \times 54} \approx 50$, 标准差 SE 的估计为 $SE = \frac{50}{\sqrt{100}} = 5$ 。

21. 解 $\bar{x} = \frac{543}{1000} \times 100 = 54.3(\%), \hat{\sigma} = \sqrt{54.3 \times 45.7} \approx 49.8$,

$$\hat{SE} = \frac{498}{\sqrt{1000}} \approx 1.575, u_{1-\frac{\alpha}{2}} \cdot \hat{SE} = u_{0.975} \times 1.575 = 3.087,$$

故

所以总体中拥有私人汽车的人的百分数的 95% 的置信区间观测值为 (51.213, 57.387)。

第七章 假设检验

7.1 设总体 $\xi \sim N(\mu, \sigma^2)$, 其中参数 μ, σ^2 为未知, 试指出下面统计假设中哪些是简单假设, 哪些是复合假设:

$$(1) H_0: \mu = 0, \sigma = 1; \quad (2) H_0: \mu = 0, \sigma > 1;$$

$$(3) H_0: \mu < 3, \sigma = 1; \quad (4) H_0: 0 < \mu < 3;$$

$$(5) H_0: \mu = 0.$$

解: (1) 是简单假设, 其余为复合假设

7.2 设 $\xi_1, \xi_2, \dots, \xi_{25}$ 取自正态总体 $N(\mu, 9)$, 其中参数 μ 未知, \bar{x} 是子样均值, 如

对检验问题 $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$ 取检验的拒绝域:

$c = \{(x_1, x_2, \dots, x_{25}) : |\bar{x} - \mu_0| \geq c\}$, 试决定常数 c , 使检验的显著性水平为 0.05

解: 因为 $\xi \sim N(\mu, 9)$, 故 $\bar{\xi} \sim N(\mu, \frac{9}{25})$

在 H_0 成立的条件下,

$$\begin{aligned} P_0(|\bar{\xi} - \mu_0| \geq c) &= P(|\bar{\xi} - \mu_0| / \frac{3}{5} \geq \frac{5c}{3}) \\ &= 2 \left[1 - \Phi\left(\frac{5c}{3}\right) \right] = 0.05 \end{aligned}$$

$\Phi(\frac{5c}{3}) = 0.975, \frac{5c}{3} = 1.96$, 所以 $c = 1.176$ 。

7.3 设子样 $\xi_1, \xi_2, \dots, \xi_{25}$ 取自正态总体 $N(\mu, \sigma_0^2)$, σ_0^2 已知, 对假设检验

$H_0: \mu = \mu_0, H_1: \mu > \mu_0$, 取临界域 $c = \{(x_1, x_2, \dots, x_n) : \bar{\xi} > c_0\}$,

(1) 求此检验犯第一类错误概率为 α 时, 犯第二类错误的概率 β , 并讨论它们之间的关系;

(2) 设 $\mu_0=0.05$, $\sigma_0^2=0.004$, $\alpha=0.05$, $n=9$, 求 $\mu=0.65$ 时不犯第二类错误的概率。

解: (1) 在 H_0 成立的条件下, $\bar{\xi} \sim N(\mu_0, \frac{\sigma_0^2}{n})$, 此时

$$\alpha = P_0(\bar{\xi} \geq c_0) = P_0\left(\frac{\bar{\xi} - \mu_0}{\sigma_0} \sqrt{n} \geq \frac{c_0 - \mu_0}{\sigma_0} \sqrt{n}\right)$$

所以, $\frac{c_0 - \mu_0}{\sigma_0} \sqrt{n} = \mu_{1-\alpha}$, 由此式解出 $c_0 = \frac{\sigma_0}{\sqrt{n}} \mu_{1-\alpha} + \mu_0$

在 H_1 成立的条件下, $\bar{\xi} \sim N(\mu, \frac{\sigma_0^2}{n})$, 此时

$$\begin{aligned} \beta &= P_1(\bar{\xi} < c_0) = P_1\left(\frac{\bar{\xi} - \mu}{\sigma_0} \sqrt{n} < \frac{c_0 - \mu}{\sigma_0} \sqrt{n}\right) \\ &= \Phi\left(\frac{c_0 - \mu}{\sigma_0} \sqrt{n}\right) = \Phi\left(\frac{\frac{\sigma_0}{\sqrt{n}} \mu_{1-\alpha} + \mu_0 - \mu}{\sigma_0} \sqrt{n}\right) \\ &= \Phi\left(\mu_{1-\alpha} - \frac{\mu - \mu_0}{\sigma_0} \sqrt{n}\right) \end{aligned}$$

由此可知, 当 α 增加时, $\mu_{1-\alpha}$ 减小, 从而 β 减小; 反之当 α 减少时, 则 β 增加。

(2) 不犯第二类错误的概率为

$$\begin{aligned} 1 - \beta &= 1 - \Phi\left(\mu_{1-\alpha} - \frac{\mu - \mu_0}{\sigma_0} \sqrt{n}\right) \\ &= 1 - \Phi\left(\mu_{0.95} - \frac{0.65 - 0.05}{0.2} 3\right) \\ &= 1 - \Phi(-0.605) = \Phi(0.605) = 0.7274 \end{aligned}$$

7.4 设一个单一观测的 ξ 子样取自分布密度函数为 $f(x)$ 的母体, 对 $f(x)$ 考虑统计假设:

$$H_0: f_0(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{其他} \end{cases} \quad H_1: f_1(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{其他} \end{cases}$$

试求一个检验函数使犯第一, 二类错误的概率满足 $\alpha + 2\beta = \min$, 并求其最小值。

解 设检验函数为

$$\phi(x) = \begin{cases} 1 & x \in c \\ 0 & \text{其他} \end{cases} \quad (c \text{ 为检验的拒绝域})$$

$$\begin{aligned} \alpha + 2\beta &= P_0(x \in c) + 2P_1(x \in \bar{c}) \\ &= P_0(x \in c) + 2[1 - P_1(x \in c)] \\ &= E_0\phi(x) + 2[1 - E_1\phi(x)] \\ &= \int_0^1 \phi(x) dx + 2(1 - \int_0^1 2x\phi(x) dx) \\ &= 2 + \int_0^1 (1 - 4x)\phi(x) dx \end{aligned}$$

要使 $\alpha + 2\beta = \min$, 当 $1 - 4x \geq 0$ 时, $\phi(x) = 0$

当 $1 - 4x < 0$ 时, $\phi(x) = 1$

所以检验函数应取 $\phi(x) = \begin{cases} 1 & x \leq \frac{1}{4} \\ 0 & x > \frac{1}{4} \end{cases}$, 此时, $\alpha + 2\beta = 2 + \int_0^1 (1 - 4x) dx = \frac{7}{8}$ 。

7.5 设某产品指标服从正态分布, 它的根方差 σ 已知为 150 小时。今由一批产品中随机抽取了 26 个, 测得指标的平均值为 1637 小时, 问在 5% 的显著性水平下, 能否认为该批产品指标为 1600 小时?

解 总体 $\xi \sim N(\mu, 150^2)$, 对假设, $H_0: \mu = 1600$, 采用 U 检验法, 在 H_0 为真时, 检验统计量

$$u = \frac{\bar{x} - \mu_0}{\sigma_0} \sqrt{n} = 1.2578$$

临界值 $u_{1-\alpha/2} = u_{0.975} = 1.96$

$|u| < u_{1-\alpha/2}$, 故接受 H_0 。

7.6 某电器零件的平均电阻一直保持在 2.64Ω , 根方差保持在 0.06Ω , 改变加工工艺后, 测得 100 个零件, 其平均电阻为 2.62Ω , 根方差不变, 问新工艺对此零件的电阻有无显著差异? 去显著性水平 $\alpha = 0.01$ 。

解 设改变工艺后电器的电阻为随机变量 ξ , 则 $E\xi = \mu$ 未知, $D\xi = (0.06)^2$,

假设为 $H_0: \mu = 2.64$ ，统计量

$$u = \frac{\bar{\xi} - \mu_0}{\sigma} \sqrt{n} = -3.33$$

由于 $u_{1-\alpha/2} = u_{0.995} = 2.10 < |u|$ ，故拒绝原假设。即新工艺对电阻有显著差异。

7.7 有甲乙两个检验员，对同样的试样进行分析，各人实验分析的结果如下：

实验号	1	2	3	4	5	6	7	8
甲	4.3	3.2	8	3.5	3.5	4.8	3.3	3.9
乙	3.7	4.1	3.8	3.8	4.6	3.9	2.8	4.4

试问甲乙两人的实验分析之间有无显著差异？

解 此问题可以归结为判断 $\xi = x_1 - x_2$ 是否服从正态分布 $N(0, \sigma^2)$ ，其中 σ^2 未知，即要检验假设 $H_0: \mu = 0$ 。

由 t 检验的统计量 $t = \frac{\bar{\xi} - \mu_0}{s_n^*} \sqrt{n} = \frac{0.1 - 0}{0.727} \sqrt{8} = -0.389$

取 $\alpha = 0.10$ ，又由于， $t_{0.95}(7) = 1.8946 > |t|$ ，故接受 H_0

7.8 某纺织厂在正常工作条件下，平均每台布机每小时经纱断头率为 0.973 根，每台布机的平均断头率的根方差为 0.162 根，该厂作轻浆试验，将轻纱上浆率减低 20%，在 200 台布机上进行实验，结果平均每台每小时轻纱断头次数为 0.994 根，根方差为 0.16，问新的上浆率能否推广？取显著性水平 0.05。

解 设减低上浆率后的每台布机断头率为随机变量 η ，有子样试验可得其均值和方差的无偏估计为 0.994 及 $s_n^{*2} = (0.16)^2$ ，问新上浆率能否推广就要分析每台布机的平均断头率是否增大，即要检验

$$H_0: E\eta = 0.973 \leftrightarrow H_1: E\eta > 0.973$$

由于 $D\eta$ 未知，且 n 较大，用 t 检验，统计量为

$$t = \frac{\bar{\eta} - \mu_0}{s_n^*} \sqrt{n} = \frac{0.994 - 0.973}{0.16} \sqrt{200} = 1.856$$

查表知 $t_{0.95}(199) = 1.645$ ，故拒绝原假设，不能推广。

7.9 在十块土地上试种甲乙两种作物，所得产量分别为 $(x_1, x_2, \dots, x_{10})$ ， $(y_1, y_2, \dots, y_{10})$ ，假设作物产量服从正态分布，并计算得 $\bar{x} = 30.97$ ， $\bar{y} = 21.79$ ， $s_x^* = 26.7$ ， $s_y^* = 12.1$ 取显著性水平 0.01，问是否可认为两个品种的产量没有显著性差别？

解 甲作物产量 $\xi \sim N(\mu_1, \sigma_1^2)$ ，乙作物产量 $\eta \sim N(\mu_2, \sigma_2^2)$ ，即要检验

$$H_0: \mu_1 = \mu_2$$

由于 σ_1^2 ， σ_2^2 未知，要用两子样 t 检验来检验假设 $H'_0: \sigma_1^2 = \sigma_2^2$ ，由 F 检验，统计量为

$$F = s_1^{*2} / s_2^{*2} = 26.7^2 / 12.1^2 = 4.869 < F_{0.995}(9, 9) = 6.54 \quad (\text{取显著性水平 } 0.01)$$

故接受假设 $H'_0: \sigma_1^2 = \sigma_2^2$ ，于是对于要检验的假设 $H_0: \mu_1 \neq \mu_2$ 取统计量

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{(n_1 - 1)s_1^{*2} + (n_2 - 1)s_2^{*2}}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} = 0.99$$

又 $\alpha = 0.01$ 时， $t_{0.995}(18) = 2.878 > |t|$ ，所以接受原假设，即两品种的产量没有显著性差别。

7.10 有甲、乙两台机床，加工同样产品，从这两台机床加工的产品中随机地抽取若干产品，测得产品直径为（单位：mm）：

甲 20.5，19.8，19.7，20.4，20.1，20.0，19.6，19.9

乙 19.7，20.8，20.5，19.8，19.4，20.6，19.2。

试比较甲乙两台机床加工的精度有无显著差异？显著性水平为 $\alpha = 0.05$ 。

解：假定甲产品直径服从 $N(\mu_1, \sigma_1^2)$ ，由子样观察值计算得 $\bar{x} = 20.00$ ，

$$s_{n_1}^{*2} = (0.3207)^2 = 0.1029。$$

乙产品直径服从 $N(\mu_2, \sigma_2^2)$ ，由子样观察值计算得 $\bar{y} = 20.00$ ， $s_{n_2}^{*2} = 0.3967$ 。

要比较两台机床加工的精度，既要检验

$$H_0: \sigma_1^2 = \sigma_2^2$$

由 F-检验

$$F = \frac{S_{n_1}^{*2}}{S_{n_2}^{*2}} = \frac{0.1029}{0.3967} = 0.2594$$

$\alpha = 0.05$ 时查表得: $F_{0.975}(7.6) = 5.70$,

$$F_{0.025}(7.6) = \frac{1}{F_{0.975}(6.7)} = \frac{1}{5.12} = 0.1953$$

由于 $F_{0.025}(7.6) < F < F_{0.975}(7.6)$, 所以接受 H_0 , 即不能认为两台机床的加工精度有显著差异。

7.11 随机从一批钉子中抽取 16 枚, 测得其长度为 (cm)

2.14 2.10 2.13 2.15 2.13 2.12 2.13 2.10
2.15 2.12 2.14 2.10 2.13 2.11 2.14 2.11

设钉长服从正态分布, 分别对下面两个情况求出总体均值 μ 的 90% 的置信区间

(1) $\sigma = 0.01cm$;

(2) σ 未知

解 (1) 由子样函数 $U = \frac{\bar{\xi} - \mu}{\sigma} \sqrt{n} \sim N(0,1)$, $p(|U| < u_{0.95}) = 0.90$, 可求 μ 的置信

区间

$$\text{置信下限 } \bar{\xi} - \frac{u_{0.95}\sigma}{\sqrt{n}} = 2.121$$

$$\text{置信上限 } \bar{\xi} + \frac{u_{0.95}\sigma}{\sqrt{n}} = 2.129$$

(2) 在 σ 未知时, 由子样函数 $t = \frac{\bar{\xi} - \mu}{S_n^*} \sqrt{n} \sim t(n-1)$, $p(|t| < t_{0.95}(n-1)) = 0.90$ 可

求得 μ 置信区间为

$$\text{置信下限 } \bar{\xi} - \frac{t_{0.95}(15)S_n^*}{\sqrt{n}} = 2.1175$$

$$\text{置信上限 } \bar{\xi} + \frac{t_{0.95}(15)S_n^*}{\sqrt{n}} = 2.1325$$

7.12 包糖机某日开工包糖，抽取 12 包糖，称得重量为

9.9 10.1 10.3 10.4 10.5 10.2 9.7 9.8 10.1 10.0 9.8 10.3

假定重量服从正态分布，试由此数据对该机器所包糖的平均重量 求置信水平为 95%的区间估计。

解 由于 σ 未知，用统计量 $t = \frac{\bar{\xi} - \mu}{s_n^*} \sqrt{n} \sim t(n-1)$ ，计算各数据值后可以得到均值的

的置信区间，置信上限为 $\bar{\xi} + \frac{t_{0.975}(11)s_n^*}{\sqrt{n}} = 10.2556$ ，下限为

$$\bar{\xi} - \frac{t_{0.975}(11)s_n^*}{\sqrt{n}} = 9.9284$$

7.13 随机取 9 发炮弹做实验，得炮口速度的方差的无偏估计 $s_n^{*2} = 11$ (米/秒)²，

设炮口速度服从正态分布，分别求出炮口速度的标准差 σ 和方差 σ^2 的置信水平为 90%的置信区间。

解 选取统计量 $\frac{(n-1)s_n^{*2}}{\sigma^2} \sim \chi^2(n-1)$ ，可得 σ^2 的置信区间为：

$$\left(\frac{(n-1)s_n^{*2}}{\chi_{1-\alpha/2}^2(n-1)}, \frac{(n-1)s_n^{*2}}{\chi_{\alpha/2}^2(n-1)} \right) = (5.6749, 32.199)$$

因为

$$P\left(\frac{(n-1)s_n^{*2}}{\chi_{1-\alpha/2}^2(n-1)} < \sigma^2 < \frac{(n-1)s_n^{*2}}{\chi_{\alpha/2}^2(n-1)}\right) = P\left(\frac{\sqrt{(n-1)s_n^{*2}}}{\sqrt{\chi_{1-\alpha/2}^2(n-1)}} < \sigma < \frac{\sqrt{(n-1)s_n^{*2}}}{\sqrt{\chi_{\alpha/2}^2(n-1)}}\right) = 1 - \alpha$$

故，标准差的置信区间取方差的根方即可。

7.14 假设六个整数 1, 2, 3, 4, 5, 6 被随机地选择，重复 60 次独立实验中出现 1, 2, 3, 4, 5, 6 的次数分别为 13, 19, 11, 8, 5, 4。问在 5%的显著性水平下是否可以认为下列假设成立：

$$H_0: p(\xi = 1) = p(\xi = 2) = \cdots = p(\xi = 6) = \frac{1}{6}。$$

解：用 χ^2 - 拟合优度检验，如果 H_0 成立

$$\chi^2 = \sum_{i=1}^6 \frac{(n_i - np_i)^2}{np_i} \sim \chi^2(5)$$

列表计算 χ^2 的观察值：

组数 i	频数 n_i	np_i	$n_i - np_i$	$(n_i - np_i)^2 / np_i$
1	13	10	3	0.9
2	19	10	9	8.1
3	11	10	1	0.1
4	8	10	-2	0.4
5	5	10	-5	2.5
6	4	10	-6	3.6

$$\chi^2 = 15.6, \quad \chi^2_{0.95}(5) = 11.07$$

由于 $\chi^2 > \chi^2_{0.95}(5)$ ，所以拒绝 H_0 。即等概率的假设不成立。

7.15 对某型号电缆进行耐压测试实验，记录 43 根电缆的最低击穿电压，数据列表如下：

测试电压	3.8	3.9	4.0	4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8
击穿频数	1	1	1	2	7	8	8	4	6	4	1

试对电缆耐压数据作分析检验（用概率图纸法和 χ^2 - 拟合优度检验）。

解：用正态概率纸检验出数据基本上服从正态分布，下面 χ^2 - 拟合优度检验假设

$$H_0: \xi \sim N(\hat{\mu}, \hat{\sigma}^2)$$

其中 $\hat{\mu}, \hat{\sigma}^2$ 为 μ 和 σ^2 的极大似然估计，其观察值

$$\hat{\mu} = \bar{x} = 4.3744 \qquad \hat{\sigma}^2 = s_n^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = 0.04842$$

所以要检验的假设

$$H_0: \xi \sim N(4.3744, 0.04842)$$

分组列表计算 χ^2 - 统计量的观察值。

组 距 $x_{i-1} \quad x_i$	频数 n_i	标准化区间 $y_{i-1} \quad y_i$	$p_i = \Phi(y_i) - \Phi(y_{i-1})$	np_i	$(n_i - np_i)^2 / np_i$
$-\infty \quad 4.1$	5	$-\infty \quad -1.25$	0.1056	4.5408	0.0464
$4.1 \quad 4.2$	7	$-1.25 \quad -0.79$	0.1087	4.6741	1.1574
$4.2 \quad 4.3$	8	$-0.79 \quad -0.34$	0.1526	6.5618	0.2152
$4.3 \quad 4.5$	12	$-0.34 \quad 0.57$	0.3488	14.9984	0.5994
$4.5 \quad 4.6$	6	$0.57 \quad 1.03$	0.1328	5.7104	0.0147
$4.6 \quad \infty$	5	$0.31 \quad \infty$	0.1515	6.5145	0.3521

$$\chi^2 = \sum_{i=1}^n \frac{(np_i - n_i)^2}{np_i} = 2.4852$$

用 $\alpha = 0.1$ 查表 $\chi^2_{0.9}(6-2-1) = \chi^2_{0.9}(3) = 6.251$ 由于 $\chi^2 < \chi^2_{0.9}(3)$ ，所以不能否定正态分布的假设。

7.16 用手枪对 100 个靶各打 10 发，只记录命中或不命中，射击结果列表如下

命中数 x_i : 0 1 2 3 4 5 6 7 8 9 10

频 数 f_i : 0 2 4 10 22 26 18 12 4 2 0

在显著水平 $\alpha = 0.05$ 下用 χ^2 拟合优度检验法检验射击结果所服从的分布。

解 对每一靶打一发，只记录命中或不命中可用二点分布描述，而对一个靶打十发，其射击结果可用二项分布 $b(K; 10, p)$ 来描述，其中 p 未知，可求其极大似然估计为

$$\hat{p} = \bar{x} = \frac{1}{100} \sum_{i=0}^{10} f_i x_i = 0.5$$

设 ξ 是十发射击中射中靶的个数，建立假设

$$H_0: p(\xi = k) = \binom{10}{K} (0.5)^K (0.5)^{10-K}, K = 0, 1, \dots, 10$$

用 χ^2 拟合优度检验法列表如下：

i	n_i	p_i	np_i	$(n_i - np_i)^2 / np_i$
0	0	0.000977	0.098	0.098
1	2	0.009765	0.976	1.074
2	4	0.043945	4.395	0.036
3	10	0.117188	11.719	0.252
4	22	0.205212	20.521	0.107
5	26	0.246094	24.609	0.079
6	18	0.205212	20.521	0.310
7	12	0.117188	11.719	0.007
8	4	0.043945	4.395	0.036
9	2	0.009765	0.976	1.074
10	0	0.000977	0.098	0.098

$$\chi^2 = \sum_{i=0}^{10} \frac{(np_i - n_i)^2}{np_i} = 3.171$$

取 $\alpha = 0.05$ ， $\chi_{0.95}^2(11-1-1) = \chi_{0.95}^2(9) = 16.919$

由于 $\chi^2 < \chi_{0.95}^2(9)$ ，所以接受 H_0 。

7.17 在某细纱机上进行断头率测定，试验锭子总数为 440，测得断头总次数为

292 次只锭子的断头次数纪律于下表。问每只锭子的纺纱条件是否相同？

每锭断头数	0	1	2	3	4	5	6	7	9
锭数（实测）	263	112	38	19	3	1	1	0	3

解：如果各个锭子的纺纱条件无差异，则所有锭子断头次数服从同一个普哇松分布，所以问题是要检验每只锭子的断头数 $\xi \sim p(K;\lambda)$ 。其中 λ 未知，求其极大似然估计为 $\lambda = \bar{x} = \frac{292}{440} = 0.66$ ，建立假设 $H_0: \xi \sim p(K;0.66)$ ，由 χ^2 拟合优度检验。

列表

温度		60	65	70	75	80
i	断头数 K	n_i	p_i	np_i	$(n_i - np_i)^2 / np_i$	
1	0	268	0.5169	227.41	5.568	
2	1	112	0.3411	150.09	9.668	
3	2	38	0.1126	49.53	2.684	
4	3	19	0.0247	10.897	6.026	
5	4-8	8	0.0047	2.068	17.016	

$$\chi^2 = \sum_{i=0}^5 \frac{(np_i - n_i)^2}{np_i} = 40.962$$

取 $\alpha = 0.05$ ， $\chi^2_{0.95}(5-1-1) = \chi^2_{0.95}(3) = 7.815$ ，

取 $\alpha = 0.01$ ， $\chi^2_{0.99}(5-1-1) = \chi^2_{0.99}(3) = 11.345$

由于 $\chi^2 > \chi^2_{0.99}(3)$ ，所以拒绝 H_0 。即认为每只锭子纺纱条件不相同。

第八章 方差分析和回归分析

8.1 考察温度对某一化工产品得率的影响，选了五种不同的温度，在同一温度下做了三次实验，测得其得率如下，试分析温度对得率有无显著影响。

	90	91	96	84	84
得率	92	93	96	83	89
	88	92	93	83	82

解 把原始数据均减去 90 后可列出如下计算表和方差分析表， r 表示因子水平数， t 为重复实验次数。

$$r=5, t=3, n=rt=15$$

计算表						
温度	60	65	70	75	80	
y_{ij}	0	1	6	-6	-6	
	5	3	6	-7	-4	
	-2	2	3	-2	-8	
$y_{i\cdot}$	0	6	15	-15	-18	$\sum_i y_{i\cdot} = -12$

$$\sum_i \sum_j y_{ij}^2 = 308, \sum_i y_{i\cdot}^2 = 810, \frac{\left(\sum_i \sum_j y_{ij}\right)^2}{n} = 9.6$$

$$S_A = \frac{1}{3} \times 810 - 9.6 = 260.4$$

$$S_T = 308 - 9.6 = 298.4$$

$$S_e = S_T - S_A = 38$$

方差分析表				
来源	平方和	自由度	均方和	F 比
温度	260.4	4	65.1	17.1
e	38	10	3.8	
总和	298.4	17	$F_{0.99}(4,10)=6$	

由于 $F=17.1>6$ ，所以在 $\alpha=0.01$ 水平上认为温度对得率有显著影响。

8.2 下面记录了三位操作工分别在四台不同机器上操作三天的日产量：

机 器	操 作 工								
	甲			乙			丙		
A_1	15	15	17	17	19	16	16	18	21
A_2	17	17	17	15	15	15	19	22	22
A_3	15	17	16	18	17	16	18	18	18
A_4	18	20	22	15	16	17	17	17	17

试在显著性水平 $\alpha=0.05$ 下检验：

- (1) 操作工之间有无显著性差异？
- (2) 机器之间的差异是否显著？
- (3) 操作工与机器的交互作用是否显著？

解 用 r 表示机器的水平数， s 表示操作工的水平数， t 表示重复实验次数，列出计算表和方差分析表：

$$r=4, s=3, t=3, n=rst=36$$

y_{ij}	甲	乙	丙	$y_{i.}$
A_1	47	54	55	156
A_2	51	45	63	159
A_3	48	51	54	153
A_4	60	48	51	159
$y_{.j}$	206	198	223	627

$$\sum_i \sum_j \sum_k y_{ijk}^2 = 11065, \quad \sum_i \sum_j y_{ik.}^2 = 33071$$

$$\sum_i y_{i..}^2 = 98307, \quad \sum_j y_{.j.}^2 = 131369, \quad \frac{(\sum_i \sum_j \sum_k y_{ijk})^2}{n} = 10920.25$$

$$S_A = \frac{1}{9} \times 98307 - 10920.25 = 2.75$$

$$S_B = \frac{1}{12} \times 131369 - 10920.25 = 27.17$$

$$S_{A \times B} = \frac{1}{3} \times 33071 - 10920.25 - 2.75 - 27.17 = 73.50$$

$$S_T = 11065 - 10920.25 = 144.75$$

$$S_e = 144.75 - 2.75 - 27.17 - 77.50 = 41.33$$

方差分析表

来源	平方和	自由度	均方和	F 比
机器 A	2.75	3	0.92	<1
操作工 B	27.17	2	13.59	7.90
交互作用 $A \times B$	73.50	6	12.25	7.12
e	41.33	24	1.72	
总和	144.75	35	$F_{0.95}(2, 24) = 3.40$ $F_{0.95}(6, 24) = 2.51$	

由于 $F_B = 7.90 > 3.40$, $F_{A \times B} = 7.12 > 2.51$, 所以在 $\alpha = 0.05$ 水平上, 操作工有显著差异, 机器之间无显著差异, 交互作用有显著差异。

8.3 通过原点的一元线性回归模型时怎样的? 通过原点的二元线性回归模型是怎样的? 分别写出结构矩阵 X , 正规方程组的系数矩阵 $X'X$, 常数项矩阵 $X'Y$, 并写出回归系数的最小二乘法估计公式。

解 通过原点的一元线性回归模型:

$$\begin{cases} y_\alpha = \beta x_\alpha + \varepsilon_\alpha & \alpha = 1, 2, \dots, N \\ \text{各 } \varepsilon_\alpha \text{ 独立同分布, } \varepsilon_\alpha \sim N(0, \sigma^2) \end{cases}$$

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}, \quad X'X = \sum_{\alpha=1}^N x_\alpha^2, \quad X'Y = (x_1, x_2, \dots, x_N) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} = \sum_{\alpha=1}^N x_\alpha y_\alpha$$

β 的最小二乘估计为

$$\hat{\beta} = (X'X)^{-1} X'Y = \sum_{\alpha=1}^N x_\alpha y_\alpha / \sum_{\alpha=1}^N x_\alpha^2$$

通过原点的二元线性回归模型:

$$\begin{cases} y_{\alpha} = \beta_1 x_{\alpha 1} + \beta_2 x_{\alpha 2} + \varepsilon_{\alpha} & \alpha = 1, 2, \dots, N \\ \text{各 } \varepsilon_{\alpha} \text{ 独立同分布, } \varepsilon_{\alpha} \sim N(0, \sigma^2) \end{cases}$$

$$X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ \vdots & \vdots \\ x_{N1} & x_{N2} \end{pmatrix}, \quad X'X = \begin{pmatrix} \sum_{\alpha} x_{\alpha 1}^2 & \sum_{\alpha} x_{\alpha 1} x_{\alpha 2} \\ \sum_{\alpha} x_{\alpha 1} x_{\alpha 2} & \sum_{\alpha} x_{\alpha 2}^2 \end{pmatrix},$$

$$X'Y = \begin{pmatrix} \sum_{\alpha} x_{\alpha 1} y_{\alpha} \\ \sum_{\alpha} x_{\alpha 2} y_{\alpha} \end{pmatrix}$$

β_1, β_2 的最小二乘估计为:

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = (X'X)^{-1} X'Y$$

8.4 对不同的元麦堆测得如下数据:

堆 号	1	2	3	4	5	6
重量 p	2813	2705	11103	2590	2131	5181
跨度 l	3.25	3.20	5.07	3.14	2.90	4.02

试求重量对跨度的回归方程, 并求出根方差 σ 的估计值。

解 设所求回归方程为 $\hat{p} = \hat{\beta}_0 + \hat{\beta}_1 l$, 由数据可以求出:

$$\sum_{\alpha} p_{\alpha} = 26523, \sum_{\alpha} p_{\alpha} l_{\alpha} = 109230.58, \sum_{\alpha} p_{\alpha}^2 = 176598625$$

$$\sum_{\alpha} l_{\alpha} = 21.58, \sum_{\alpha} l_{\alpha}^2 = 80.9374, N = 6$$

由最小二乘法估计公式可知

$$\hat{\beta}_1 = \frac{\sum_{\alpha} p_{\alpha} l_{\alpha} - \frac{1}{N} p_{\alpha} \sum_{\alpha} l_{\alpha}}{\sum_{\alpha} l_{\alpha}^2 - \frac{1}{N} \left(\sum_{\alpha} l_{\alpha} \right)^2} = 4165.85$$

$$\hat{\beta}_0 = \frac{1}{N} \sum_{\alpha} p_{\alpha} - \hat{\beta}_1 \frac{1}{N} \sum_{\alpha} l_{\alpha} = -10562$$

故可得回归方程: $\hat{p} = -10562 + 4165.85l$

σ^2 的估计是

$$\hat{\sigma}^2 = \frac{1}{N-2} \left\{ \left[\sum_{\alpha} p_{\alpha}^2 - \frac{1}{N} \left(\sum_{\alpha} p_{\alpha} \right)^2 \right] - \hat{\beta}_1 \left[\sum_{\alpha} p_{\alpha} l_{\alpha} - \frac{1}{N} \left(\sum_{\alpha} p_{\alpha} \right) \left(\sum_{\alpha} l_{\alpha} \right) \right] \right\}$$

$$= 428538$$

则 σ 的估计为 655

8.5 设

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 (3x_i^2 - 2) + \varepsilon_i, i=1,2,3$$

$$x_1 = -1, x_2 = 0, x_3 = 1$$

$\varepsilon_1, \varepsilon_2, \varepsilon_3$ 相互独立同服从于 $N(0, \sigma^2)$ 。

(1) 写出矩阵 X

(2) 求 $\beta_0, \beta_1, \beta_2$ 的最小二乘估计

(3) 证明当 $\beta_2 = 0$ 时, β_0, β_1 的最小二乘估计不变

解 (1) $X = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix}$

(2) $X'X = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix}, X'Y = \begin{pmatrix} y_1 + y_2 + y_3 \\ -y_1 + y_3 \\ y_1 - 2y_2 + y_3 \end{pmatrix}$, 则, $\beta_0, \beta_1, \beta_2$ 的最小二乘估计是

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = (X'X)^{-1} X'Y = \begin{pmatrix} \frac{1}{3}(y_1 + y_2 + y_3) \\ \frac{1}{2}(-y_1 + y_3) \\ \frac{1}{6}(y_1 - 2y_2 + y_3) \end{pmatrix}$$

(3) 若 $\beta_2 = 0$, 此时模型成为:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, i=1,2,3, \text{ 则对应的}$$

$$X = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}, X'X = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}, X'Y = \begin{pmatrix} y_1 + y_2 + y_3 \\ -y_1 + y_3 \end{pmatrix}, \beta_0, \beta_1 \text{ 的最小二乘估计是}$$

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = (X'X)^{-1} X'Y = \begin{pmatrix} \frac{1}{3}(y_1 + y_2 + y_3) \\ \frac{1}{2}(-y_1 + y_3) \end{pmatrix}$$

8.6 若 y 与 x 有下述关系:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_p x^p + \varepsilon$$

其中 $\varepsilon \sim N(0, \sigma^2)$ 从中获得了 n 组独立观测值 (x_α, y_α) , 能否求出 $\beta_0, \beta_1, \dots, \beta_p$ 的最小二乘估计, 试写出最小二乘估计的公式, 能否检验假设

$$H_0: \beta_i = 0$$

试写出检验的拒绝域。

解 若记 $X_{\alpha i} = x_\alpha^i, \bar{X}_i = \frac{1}{n} \sum_{\alpha=1}^n x_\alpha^i \quad \alpha=1, \dots, n, i=1, \dots, p$

$$l_{ij} = \sum_{\alpha=1}^n (X_{\alpha i} - \bar{X}_i)(X_{\alpha j} - \bar{X}_j) \quad i, j=1, \dots, p$$

$$l_{i0} = \sum_{\alpha=1}^n (X_{\alpha i} - \bar{X}_i)(y_\alpha - \bar{y}) \quad i=1, \dots, p$$

则 β_1, \dots, β_p 的最小二乘估计为下述方程组的解:

$$\begin{cases} l_{11}\hat{\beta}_1 + l_{12}\hat{\beta}_2 + \cdots + l_{1p}\hat{\beta}_p = l_{10} \\ l_{21}\hat{\beta}_1 + l_{22}\hat{\beta}_2 + \cdots + l_{2p}\hat{\beta}_p = l_{20} \\ \quad \quad \quad \dots \quad \quad \quad \dots \\ l_{p1}\hat{\beta}_1 + l_{p2}\hat{\beta}_2 + \cdots + l_{pp}\hat{\beta}_p = l_{p0} \end{cases} \quad (*)$$

β_0 的最小二乘估计为:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{X}_1 - \cdots - \hat{\beta}_p \bar{X}_p$$

若把方程组 (*) 的系数矩阵记为 L , 则 $L = (l_{ij})$, 又记 $L^{-1} = (l^{ij})$, 则在显著性水平 α 上检验 $H_0: \beta_i = 0$ 的拒绝域是:

$$F_i = \frac{\hat{\beta}_i^2}{l^{ii} \hat{\sigma}^2} > F_{1-\alpha}(1, n-p-1)$$

其中, $\hat{\sigma}^2 = \frac{1}{n-p-1} \{ \sum_{\alpha} (y_{\alpha} - \bar{y})^2 - \hat{\beta}_1 l_{10} - \cdots - \hat{\beta}_p l_{p0} \}$

8.7 某医院用光色比色计检验尿贡时, 得尿贡含量与肖光系数读数的结果如下:

尿贡含量 x	2	4	6	8	10
肖光系数 y	64	138	205	285	360

已知它们之间有下列关系式:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad i=1,2,3,4,5$$

各 ε_i 相互独立, 均服从 $N(0, \sigma^2)$ 分布, 试求 β_0, β_1 的最小二乘估计, 并给出检验假设

$$H_0: \beta_1 = 0$$

的拒绝域。

解 由数据可以求得, $n=5$

$$\sum_{\alpha} x_{\alpha} = 30, \quad \bar{x} = 6$$

$$\sum_{\alpha} y_{\alpha} = 1052, \quad \bar{y} = 210.4$$

$$\sum_{\alpha} x_{\alpha}^2 = 220, \quad \sum_{\alpha} x_{\alpha} y_{\alpha} = 7790, \quad \sum_{\alpha} y_{\alpha}^2 = 275990$$

$$l_{xx} = 40, l_{xy} = 1478, l_{yy} = 54649.2$$

则, 最小二乘估计为:

$$\hat{\beta}_0 = -11.3, \quad \hat{\beta}_1 = 36.95$$

检验假设 $H_0: \beta_1 = 0$ 可用统计量

$$F = \frac{\hat{\beta}_1 l_{xy}}{(l_{yy} - \hat{\beta}_1 l_{xy}) / (n-2)} = 4416 > F_{1-\alpha}(1, 3) = 34.1, \alpha = 0.01$$

因此, 拒绝原假设。

8.8 研究同一地区土壤中所含植物可给态磷的情况, 得到 18 组数据如下, 其中,

x_1 ——土壤内所含无机磷浓度

x_2 ——土壤内溶于 K_2CO_3 溶液并受溴化物水解的有机磷浓度

x_3 ——土壤内溶于 K_2CO_3 溶液但不溶于溴化物的有机磷浓度

y ——载在 $20^\circ C$ 土壤内的玉米中可给态磷的浓度

已知 y 与 x_1, x_2, x_3 之间有下列关系：

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i \quad i=1, 2, \dots, 18$$

各 ε_i 相互独立，均服从 $N(0, \sigma^2)$ 分布，试求出回归方程，并对方程及各因子的显著性进行检验。

土壤样本	x_1	x_2	x_3	y
1	0.4	53	158	64
2	0.4	23	163	60
3	3.1	19	37	71
4	0.6	34	157	61
5	4.7	24	59	54
6	1.7	65	123	77
7	9.4	44	46	81
8	10.1	31	117	93
9	11.6	29	173	93
10	12.6	58	112	51
11	10.9	37	111	76
12	23.1	46	114	96
13	23.1	50	134	77
14	21.6	44	73	93
15	23.1	56	168	95
16	1.9	36	143	54
17	26.8	28	202	168
18	29.9	51	124	99

由上述数据可以求得下面的结果：

$$p=3, n=18$$

$$\bar{x}_1 = 11.94, \bar{x}_2 = 42.11, \bar{x}_3 = 123, \bar{y} = 81.28$$

$$L = \begin{pmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{23} \\ l_{31} & l_{32} & l_{33} \end{pmatrix} = \begin{pmatrix} 1752.9644 & 1085.6111 & 1200.0000 \\ 1085.6111 & 3155.7778 & 3364.0000 \\ 1200.0000 & 3364.0000 & 35572.0000 \end{pmatrix}$$

$$l = \begin{pmatrix} l_{10} \\ l_{20} \\ l_{30} \end{pmatrix} = \begin{pmatrix} 3231.4778 \\ 2216.4445 \\ 7953.0000 \end{pmatrix}$$

$$L^{-1} = \begin{pmatrix} l^{11} & l^{12} & l^{13} \\ l^{21} & l^{22} & l^{23} \\ l^{31} & l^{32} & l^{33} \end{pmatrix} = \begin{pmatrix} 0.000725 & -0.000248 & -0.000001 \\ -0.000248 & 0.000437 & -0.000033 \\ -0.000001 & -0.000033 & 0.000031 \end{pmatrix}$$

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix} = L^{-1}l = \begin{pmatrix} 1.784780 \\ -0.083397 \\ 0.161133 \end{pmatrix}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}_1 - \hat{\beta}_2 \bar{x}_2 - \hat{\beta}_3 \bar{x}_3 = 43.652198$$

所求得的回归方程为

$$\hat{y} = 43.65 + 1.78x_1 - 0.08x_2 + 0.16x_3$$

记

$$S_T = \sum_{\alpha} (y_{\alpha} - \bar{y})^2 = 12389.6111$$

$$S_R = \sum_{j=1}^3 \hat{\beta}_j l_{j0} = 6806.1115$$

$$S_e = S_T - S_R = 5583.4997$$

对方乘作检验的 F 统计量为:

$$F = \frac{S_R / p}{S_e / (n - p - 1)} = 5.6885 > F_{1-0.05}(3, 14) = 3.34$$

故在 $\alpha = 0.05$ 的水平上方程是显著的。

对各因子作 F 检验的统计量分别为

$$F_1 = \frac{\hat{\beta}_1^2}{l^{11} S_e / (n - p - 1)} = 11.02 > F_{0.95}(1, 14) = 4.60$$

$$F_2 = \frac{\hat{\beta}_2^2}{\hat{l}^2 S_e / (n-p-1)} = 0.0399 < F_{0.95}(1,14) = 4.60$$

$$F_3 = \frac{\hat{\beta}_3^2}{\hat{l}^3 S_e / (n-p-1)} = 2.0822 < F_{0.95}(1,14) = 4.60$$

故在 $\alpha = 0.05$ 的水平上, x_1 是显著的, x_2 与 x_3 是不显著的。

8.8 某种膨胀合金含有两种主要成分, 做了一批试验如表所示, 从中发现这两种成分含量和 x 与合金的膨胀数 y 之间有一定关系。

- (1) 试确定 x 与 y 之间的关系表达式
- (2) 求出其中系数的最小二乘估计
- (3) 对回归方程及各项作显著性检验

试验号	金属成分和 x	膨胀系数 y
1	37.0	3.40
2	37.0	3.00
3	38.0	3.00
4	38.5	3.27
5	39.0	2.10
6	39.5	1.83
7	40.0	1.53
8	40.5	1.70
9	41.0	1.80
10	41.5	1.90
11	42.0	2.35
12	42.5	2.54
13	43.0	3.90

解 (1) 由散点图可知 y 与 x 的关系为:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$$

并可假设 $\varepsilon \sim N(0, \sigma^2)$ 。

- (2) 由以上数据可求得:

$$p=2, \quad n=13$$

$$\bar{x} = 40, \quad \bar{x}_2 = \overline{x^2} = 1603.5, \quad \bar{y} = 2.409231$$

$$L = \begin{pmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{pmatrix} = \begin{pmatrix} 45.50 & 3640.00 \\ 3640.00 & 291325 \end{pmatrix}$$

$$l = \begin{pmatrix} l_{10} \\ l_{20} \end{pmatrix} = \begin{pmatrix} -6.37 \\ -490.08 \end{pmatrix}$$

$$L^{-1} = \begin{pmatrix} l^{11} & l^{12} \\ l^{21} & l^{22} \end{pmatrix} = \begin{pmatrix} 51.170829 & -0.639361 \\ -0.369361 & 0.007992 \end{pmatrix}$$

据最小二乘估计为：

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = L^{-1}l = \begin{pmatrix} -12.620320 \\ 0.156004 \end{pmatrix}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} - \hat{\beta}_2 \bar{x}_2 = 257.069610$$

则回归方程为：

$$\hat{y} = 257.070 - 12.620x + 0.156x^2$$

(3) 对方程作检验：

$$S_T = \sum_{\alpha} (y_{\alpha} - \bar{y})^2 = 5.019692$$

$$S_R = \hat{\beta}_1 l_{10} + \hat{\beta}_2 l_{20} = 3.9370$$

$$S_e = S_T - S_R = 1.0827$$

$$F = \frac{S_R / p}{S_e / (n - p - 1)} = 18.18 > F_{1-0.05}(2, 10) = 4.10$$

故在 $\alpha = 0.05$ 的水平上方程是显著的。

对 x 及 x^2 项作检验：

$$F_1 = \frac{\hat{\beta}_1^2}{l^{11} S_e / (n - p - 1)} = 28.7483 > F_{0.95}(1, 10) = 4.96$$

$$F_2 = \frac{\hat{\beta}_2^2}{l^{22} S_e / (n - p - 1)} = 28.1261 > F_{0.95}(1, 10) = 4.96$$

故方程中两项均为显著。