第三部分

参考答案

第一章 概率论基础

1. (1) $\Omega = \{(i,j) : i,j = 1,\dots,6\}$,
   $A = \{(i,j) : i > j, i = 2,\dots,6, j = 1,\dots,5\}$,
   $B = \{(i,i) : i = 1,\dots,6\}$,
   $C = \{(i,j) : i + j = 10, i,j = 1,\dots,6\} = \{(4,6),(5,5),(6,4)\}$.
   (2) $\Omega = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$,
   $A = \{THH, THT, TTH, TTT\}$,
   $B = \{HHT, HTH, THH\}$,
   $C = \{TTT, HHH\}$.
   (3) $\Omega = \{(x,y) : x^2 + y^2 < 1\}$
   $A = \{(x,y) : x^2 + y^2 < 1/2\}$
   $B = \{(x,y) : 1/3 < x^2 + y^2 < 1/2\}$.

2. 略

3. (1) $A_1\bar{A}_2\bar{A}_3 + \bar{A}_1A_2\bar{A}_3 + \bar{A}_1\bar{A}_2A_3$,
   (2) $A_1 \cup A_2 \cup A_3$,
   (3) $A_1 \cap (A_2 \cup A_3)$,
   (4) $A_1\bar{A}_2\bar{A}_3 + \bar{A}_1A_2\bar{A}_3 + \bar{A}_1\bar{A}_2A_3 + \bar{A}_1\bar{A}_2\bar{A}_3$.

4. (1) $\emptyset$, (2)$\Omega = [0,2]$, (3) $\bar{A} = \{0 \le x \le 1/2\} \cup \{1 < x \le 2\}$, (4) $B = \{1/4 < x \le 3/2\}$

5. (1) $A \subset B$, 0.7 (2)$A \cup B = \Omega$, 0.5

6. 3/4

7. $4/11!$, $4/A_{11}^7$

8. 略

9. 41/63

10. $1 - \frac{\binom{365}{50}50!}{365^{50}}$

11. $\frac{a}{a+b}$

12. $\frac{\binom{3}{2}}{\binom{100}{2}} = 1/1650, \frac{\binom{97}{2}}{\binom{100}{2}} = 776/825$

13. $\frac{\binom{4}{m}3^{4-m}}{4^4}, m = 0, \ldots, 4.$

14. 3.1554

15. 1/4

16. 1/8

17. 0.504

18. $\frac{12!}{6^{12}2^6}$

19. 0.176

20. 0.0033

21. 0.7806

22. (1) 0.25, (2) 0.35, (3) 0.55, (4) 0.45.

23. 五局三胜更好

24. (1) $N \geq 2n-1, \frac{\binom{N-n+1}{n}}{\binom{N}{n}}$ (2) $N \geq 3n/2-1, n\, mod\, 2 = 0, \frac{\binom{N-n+1}{n/2}}{\binom{N}{n}}; n\, mod\, 2 = 1, 0.$
    (3) $N\ mod\ 2 = 0, \frac{\binom{N/2}{n}2^n}{\binom{N}{n}}; N\ mod\ 2 = 1, \frac{\binom{(N-1)/2}{n-1}2^{n-1}}{\binom{N}{n}}$

25. 19/36, 1/18

26. 5/11

27. 3/11

28. 5/12

29. $\frac{\binom{19}{8}11!}{19^8}$

30. (1) $\binom{11}{8}8!/11^8$ (2) $1/11^7$ (3) $\binom{8}{3}\binom{10}{5}5!/11^8$

31. $2/3$

32. $1/10, n=1$; $\frac{\binom{9}{n-1}+\binom{8}{n-1}-\binom{4}{n-1}}{\binom{10}{n}}, 2 \leq n \leq 5$; $\frac{\binom{9}{n-1}+\binom{8}{n-1}}{\binom{10}{n}}, 6 \leq n \leq 9$; $1, n=10$

33. $2(k-1)(n-k)/[(n(n-1)]$

34. $5/8$

35. $7/8$

36. $\frac{1}{9}\sum_{n=m}^{9}\frac{1}{n}$

37. $1/11$

38. (1) $3/10$ (2) $3/5$

39. (1) $1/2$ (2)$1/2$

40. $3/5$

41. $3/10$, $3/5$

42. (1) $1/2$ (2)$1-(5/6)^3$

43. (1) $11/36$ (2) $1/3$

44. $5/12$

45. $2/3$

46. $1/4$

47. 0.4, 0.4857

48. $6/7$

49. 0.893, 0.957

50. $\frac{rb(b+a)(r+a)}{(r+b)(r+b+a)(r+b+2a)(r+b+3a)}$

51. $[\binom{10}{3}+\binom{5}{3}]/\binom{15}{3}$

52. 0或1

53. 略

54. 略

55. 不独立

56. 略

57. 9

58. 0.182

59. (1) 0.26 (2) 0.96

60. (1)$1-(1-p)^m$ (2) $\binom{m}{10}p^{10}(1-p)^{m-10}$ (3) $1-\sum_{i=0}^{2}\binom{m}{i}p^i(1-p)^{m-i}$

61. 0.6928

62. 0.468

63. 略

64. 略

65. (1) $p_Ap_Bp_C$ (2)$1-(1-p_A)(1-p_B)(1-p_C)$ (3) $1-(1-p_A^2)(1-p_B^2)(1-p_C^2)$ (4) $p_D^2[1-(1-p_A)(1-p_B)(1-p_C)]$ (5) $2p_A(1-p_A)[1-(1-p_B)(1-p_Bp_C)]+p_A^2[1-(1-p_B)^2]$

66. $p^2(2-p^2)$, $p^2(2-p)^2$, 先并联在串联更稳定

67. $1-(1-p)^m, \binom{m}{10}p^{10}(1-p)^{m-10}, 1-\sum_{j=0}^{2}\binom{m}{j}p^j(1-p)^{m-j}$

68. $1-\prod_{i=1}^{n}p_i$ (2) $\prod_{i=1}^{n}(1-p_i)$ (3) $\sum_{i=1}^{n}[p_i\prod_{j=1,j\neq i}^{n}(1-p_j)]$

69. 来自第一个盒子的概率更大

70. 50%

71. $1/(m-1)$

72. 4/9

73. $[1-(1-p)^2][1-(1-p)^3]$

74. (1) 7/600 (2) 3/7

75. $4p/(3p+1)$

76. (1) 29/90 (2) 20/61

77. 0.9542

78. $(m-2)/(2m-2)$

79. 0.508

80. 4/9

81. (1) 91.3% (2) 99.9%

82. (1) 1/2 (2) 从第二个盒子取得的概率最大

83. 0.995%

84. (1)164/231 (2)20/41

85. (1)2/3 (2)2/3

86. 对乙更有利

87. 不会

第二章 随机变量及其分布

1. B.

2. (1) 若以$\{X=7\}$表示未获奖, 则 X 的分布律为

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \frac{1}{17721088} & \frac{15}{17721088} & \frac{162}{17721088} & \frac{7695}{17721088} & \frac{137475}{17721088} & \frac{1043640}{17721088} & \frac{16532100}{17721088} \end{pmatrix}.$$

(2) $A=\{X \le 6\}, \quad \mathbb{P}(A)=\frac{1188988}{17721088} \approx 6.71\%;$
$B=\{X=1 \text{ 或 } 2\}$ 或 $\{X=1\}\bigcup\{X=2\}, \quad \mathbb{P}(B)=0.00009\%.$

3. $p=0.1.$

4. $X \sim \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ \frac{1}{3} & \frac{2}{9} & \frac{4}{27} & \frac{8}{81} & \frac{16}{81} \end{pmatrix}.$

5. $X \sim \begin{pmatrix} 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} \end{pmatrix}.$

6. $X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{216} & \frac{7}{216} & \frac{19}{216} & \frac{37}{216} & \frac{61}{216} & \frac{91}{216} \end{pmatrix}.$

7. X 的分布律为 $X \sim \begin{pmatrix} 100 & 80 & 50 & -60 \\ 0.6 & 0.2 & 0.1 & 0.1 \end{pmatrix}.$

8. 分布律为 $\begin{pmatrix} 10 & 5 & 0 & -2 \\ 0.32768 & 0.4096 & 0.2048 & 0.05792 \end{pmatrix}.$

9. 随机变量 X的分布律为

$$\begin{aligned}
\mathbb{P}(X=100) &= \frac{2}{\binom{20}{10}} = \frac{1}{92378} \approx 1.08\times 10^{-5},\\
\mathbb{P}(X=20) &= \frac{2\binom{10}{9}\binom{10}{1}}{\binom{20}{10}} = \frac{100}{92378} \approx 1.08\times 10^{-3},\\
\mathbb{P}(X=5) &= \frac{2\binom{10}{8}\binom{10}{2}}{\binom{20}{10}} = \frac{2025}{92378} \approx 0.022\\
\mathbb{P}(X=0) &= 1-\frac{2126}{92378} \approx 0.977.
\end{aligned}$$

10. (1) 分布函数为

$$F(x)=\begin{cases} 0, & x<-1,\\ \frac{1}{4}, & -1\le x<1,\\ \frac{3}{4}, & 1\le x<2,\\ 1, & x\ge 2.\end{cases}$$

(2) $\frac{1}{4}$; $\frac{1}{2}$; $\frac{3}{4}$; $\frac{1}{4}$.

11. X 的分布律为 $X\sim\begin{pmatrix} 1 & 2 & 3\\ \frac{4}{5} & \frac{8}{45} & \frac{1}{45}\end{pmatrix}$; 分布函数为 $F(x)=\begin{cases} 0, & x<1,\\ \frac{4}{5}, & 1\le x<2,\\ \frac{44}{45}, & 2\le x<3,\\ 1, & x\ge 3.\end{cases}$

12. X 的分布律为 $X\sim\begin{pmatrix} -1 & 1 & 2\\ 0.4 & 0.4 & 0.2\end{pmatrix}$.

13. X 的分布律为 $\mathbb{P}(X=k)=p(1-p)^{k-1},\ k=1,2,\cdots$.

14. $\frac{3}{8}$; $\frac{9}{25}$.

15. $\frac{2}{3}$.

16. 0.936.

17. $\mathbb{P}(X\ge 1)=1-0.8^{20}$; X 的最有可能的取值为 4.

18. (1) 0.59526; (2) 0.4149.

19. 请 5 名董事代表比较好.

20. $p_4 = 0.6^4 \approx 0.13$; $\quad p_5 = \binom{4}{1} \times 0.4 \times 0.6^4 \approx 0.21$; $\quad p_6 = \binom{5}{2} \times 0.4^2 \times 0.6^4 \approx 0.21$; $p_7 = \binom{6}{3} \times 0.4^3 \times 0.6^4 \approx 0.17$. “三场两胜”制对乙队更有利.

21. (1) $\frac{3}{16}$; (2) $\frac{11}{16}$.

22. 所求的分布律为 $X \sim \begin{pmatrix} -1 & 1 & 2 & 3 \\ \frac{125}{216} & \frac{75}{216} & \frac{15}{216} & \frac{1}{216} \end{pmatrix}$.

23. $\frac{4}{3}\mathrm{e}^{-2} \approx 0.18$.

24. $\frac{27\mathrm{e}^2}{27\mathrm{e}^2+25} = 0.8886$.

25. 以 X 记一只昆虫产卵的个数, 则对任意整数$m, n \geq 0$, 由

$$\begin{aligned}
\mathbb{P}(Y=m, Z=n) &= \mathbb{P}(Y=m, Z=n | X=m+n)\mathbb{P}(X=m+n) \\
&= \binom{m+n}{m} p^m (1-p)^n \mathrm{e}^{-\lambda} \frac{\lambda^{m+n}}{(m+n)!} \\
&= \mathrm{e}^{-\lambda p} \frac{(\lambda p)^m}{m!} \cdot \mathrm{e}^{-\lambda(1-p)} \frac{[\lambda(1-p)]^n}{n!}
\end{aligned}$$

可知 Y 和 Z 分别服从参数为 λp 和 $\lambda(1-p)$ 的Poisson分布, 且相互独立.

26. e^{-1}.

27. 所求概率 $\approx 1 - 9\mathrm{e}^{-8} \approx 0.997$.

28. 在 10 秒内, 所求概率为 $1 - \mathrm{e}^{-0.512} \approx 0.401$; 而在 100 秒内, 则所求概率为 $1 - \mathrm{e}^{-5.12} \approx 0.994$.

29. 大约为 $(1+2.6)\mathrm{e}^{-(52\times 5\%)} \approx 0.27$.

30. (1) $1 - \mathrm{e}^{-0.01} \approx 0.0099$. (2) $10000 \cdot \ln 20 \approx 3$ 万.

31. $\{1, 2, \cdots, 99\}$ 上的均匀分布. 证明方法: 对总的投篮次数采用数学归纳法.

32. (1) $\frac{1}{4}, \frac{1}{12}, \frac{1}{6}$; (2) $\frac{1}{2}$.

33. $a = 1$, $\quad \mathbb{P}(X > \frac{\pi}{6}) = \frac{1}{2}$.

34. C.

35. C.

36. $a = \frac{5}{16}$, $b = \frac{7}{16}$.

37. 证明: 先证对任意有理数 $x \in (0,1)$, 我们有 $F(x) = x$. 事实上, 若 $x = \frac{m}{n}$, 其中 $m < n$ 为正整数, 则由题目条件可知 $F(0) = 0, F(1) = 1$, 且

$$F\Big(\frac{1}{n}\Big) - F(0) = F\Big(\frac{2}{n}\Big) - F\Big(\frac{1}{n}\Big) = \cdots = F(1) - F\Big(\frac{n-1}{n}\Big).$$

易知上式中各项之和为 1, 故每项均等于 $\frac{1}{n}$. 从而对任意 $m < n$, 有 $F(\frac{m}{n}) = \frac{m}{n}$. 再由连续函数的右连续性可知 $F(x) = x$ 对所有 $(0,1)$ 中的无理数也成立. 这就证明了 $X \sim U(0,1)$.

38. $a = \frac{1}{3},\ b = \frac{1}{2}$.

39. (1) $a = \frac{1}{\pi}$; (2) $F(x) = \frac{1}{2} + \frac{1}{\pi}\arctan x$; (3) $\frac{1}{2}$.

40. (1) $a = 1$; (2) $\mathbb{P}(|X| < \frac{1}{2}) = \frac{3}{4}$.

41. 密度函数为 $f(x) = \frac{3}{4}(2x - x^2),\ 0 < x < 2$; 分布函数为

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{3}{4}x^2 - \frac{1}{4}x^3, & 0 \le x < 2, \\ 1, & x \ge 2. \end{cases}$$

42. (1) $a = 1, b = -1$; (2) $f(x) = 2x\ln x - x,\ 1 < x < \mathrm{e}$.

43. (1) $a = \frac{1}{2}, b = \frac{1}{2}$; (2) $f(x) = \frac{1}{2}\mathrm{e}^{x}I(x < 0) + \frac{1}{2}\mathrm{e}^{-(x-1)}I(x > 1)$.

44. D.

45. $\frac{3}{5}$.

46. $\frac{1}{3}$.

47. $\frac{20}{27}$.

48. $1 - \mathrm{e}^{-1}$.

49. $\mathrm{e}^{-2} \approx 0.318; \mathrm{e}^{-2} \approx 0.318$.

50. $1 - (1 - \mathrm{e}^{-2})^5 \approx 0.517$.

51. (1) 证明: 先证必要性. 设 X 服从参数为 $\lambda > 0$ 的指数分布, 则由条件概率的定义, 有

$$\mathbb{P}(X \le t + x | X > t) = \frac{\mathbb{P}(t < X \le t + x)}{\mathbb{P}(X > t)} = \frac{F(t+x) - F(t)}{1 - F(t)},$$

其中 $F(x)$ 表示 X 的分布函数. 从而,

$$\mathbb{P}(X \le t+x|X>t) = \frac{\mathrm{e}^{-\lambda t} - \mathrm{e}^{-\lambda(t+x)}}{\mathrm{e}^{-\lambda t}} = 1 - \mathrm{e}^{-\lambda x} = F(x) = \mathbb{P}(X \le x).$$

下面来证充分性. 记函数 $g(x) = \ln(1-F(x))$, $x > 0$, 则将题目中的式子变形可得

$$g(t+x) = g(t) + g(x), \quad t, x > 0.$$

而上述函数方程具有唯一解: $g(x) = g(1)x$, 其中 $g(1)$可以为任意常数. 由题意进一步可知此时 $g(1) < 0$, 若记常数 $\lambda = -g(1)$, 则 $F(x) = 1 - \mathrm{e}^{-\lambda x}$. 这就证明了此时 X 服从参数为 λ 的指数分布. (2) 证明思路与 (1) 相同, 略.

52. (1) $\Phi(2) \approx 0.9773$, $2\Phi(2)-1 \approx 0.9546$; (2) a 满足 $\Phi(a) > 0.95$, 查表得 $a \ge 1.645$.

53. (1) $\mathbb{P}(0 \le X \le 4) = \Phi(1.5)+\Phi(0.5)-1 \approx 0.6247$, $\mathbb{P}(X > 2.4) = 1-\Phi(0.7) \approx 0.2420$, $\mathbb{P}(|X| > 2) = 2 - \Phi(1.5) - \Phi(0.5) \approx 0.3753$; (2) 由$\Phi(\frac{1-c}{2}) = \frac{2}{3}$, 查表可知 $c \approx 0.14$.

54. (1) 0.8; (2) $2\Phi(2) - 1 \approx 0.9546$.

55. (1) 两条路线满足要求的概率分别是 $\Phi(2)$ 和 $\Phi(2.5)$, 故选第二条路线; (2) 两条路线满足要求的概率分别是 $\Phi(1.5)$ 和 $\Phi(1.25)$, 故选第一条路线.

56. 所求概率为 $0.4[1-\Phi(0.8)] \approx 0.0848$.

57. 所求为$170 + 6u_{0.995} \approx 185.5$厘米.

58. A.

59. A.

60. A.

61. B.

62. (1) Y_1 的分布律为

$$Y_1 \sim \begin{pmatrix} -3 & -1 & 1 & 3 \\ 0.2 & 0.3 & 0.1 & 0.4 \end{pmatrix}.$$

(2) Y_2 的分布律为

$$Y_2 \sim \begin{pmatrix} 0 & 1 & 2 \\ 0.3 & 0.3 & 0.4 \end{pmatrix}.$$

(3) Y_3 的分布律为

$$Y_2 \sim \begin{pmatrix} 0 & 1 & 4 \\ 0.1 & 0.7 & 0.2 \end{pmatrix}.$$

63. $Y \sim \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$, $Z \sim \begin{pmatrix} 0 & \frac{\pi}{2} & \pi \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \end{pmatrix}$.

64. (1) $a = \frac{1}{2}$, $b = \frac{1}{\pi}$. (2) 不成立, 因为 $\int_{-\infty}^{\infty} |x| f(x) \mathrm{d}x = \infty$, 其中 $f(x)$ 是 X 的密度函数. (3) $p(y) = \frac{3(y-3)^2}{\pi[1+(y-3)^6]}$. (4) 验证随机变量 $Y = \frac{1}{X}$ 的密度函数为 $p(y) = \frac{1}{\pi(1+y^2)}$.

65. (1) Y_1 的密度函数为 $f_1(y) = y^{-1}$, $1 < y < \mathrm{e}$. (2) Y_2 的密度函数为 $f_2(y) = y^{-2}$, $y > 1$. (3) Y_3 的密度函数为 $f_3(y) = \lambda \mathrm{e}^{-\lambda y}$, $y > 0$, 即 Y_3 服从参数为 λ 的指数分布.

66. Y_1 的密度函数为 $f_1(y) = \frac{1}{\pi(1+y^2)}$; Y_2 的密度函数为 $f_2(y) = \frac{2}{\pi}\frac{1}{\sqrt{1-y^2}}$, $0 < y < 1$.

67. 证明: 首先注意到 Y 的取值范围为 (0,1). 而当 $0 < y < 1$ 时, 利用函数 $F(x)$ 的严格单调性, 则

$$\mathbb{P}(Y \le y) = \mathbb{P}(F(X) \le y) = \mathbb{P}(X \le F^{-1}(y)) = F(F^{-1}(y)) = y.$$

从而有 $Y \sim U(0,1)$.

68. Y_1 的密度函数为 $f_1(y) = \frac{1}{2\sqrt{y}}\mathrm{e}^{-\sqrt{y}}$, $y > 0$; Y_2 的密度函数为 $f_2(y) = 1$, $0 < y < 1$, 即 $Y_2 \sim U(0,1)$.

69. $g(x) = -2\ln(1-x)$.

70. Y 的分布律为

$$\mathbb{P}(Y = n) = \mathrm{e}^{-\lambda n}(1 - \mathrm{e}^{-\lambda}), \quad n = 0, 1, 2, \cdots,$$

即 Y 服从参数为 $1 - \mathrm{e}^{-\lambda}$ 的几何分布(取值范围为非负整值). 对任意 $n \ge 0$ 和 $0 < z < 1$, 由

$$\mathbb{P}(Z \le z | Y = n) = \frac{\int_n^{n+z} \lambda \mathrm{e}^{-\lambda x} \mathrm{d}x}{\mathrm{e}^{-\lambda n}(1 - \mathrm{e}^{-\lambda})} = \frac{1 - \mathrm{e}^{-\lambda z}}{1 - \mathrm{e}^{-\lambda}}$$

可知 Y 和 Z 相互独立, 且 Z 的密度函数为 $p(z) = \frac{\lambda \mathrm{e}^{-\lambda z}}{1-\mathrm{e}^{-\lambda}}$, $0 < z < 1$.

71. Y 的密度函数为

$$p(y)=\begin{cases}\lambda e^{-\lambda y}, & y>1,\\ \frac{\lambda}{2\sqrt{-y}}e^{-\lambda\sqrt{-y}}, & -1<y<0.\end{cases}$$

72. (1) Y_1 的密度函数为 $f_1(y)=\frac{1}{\sqrt{2\pi}y}e^{-\frac{(\ln y)^2}{2}}$, $y>0$. (2) Y_2 的密度函数为 $f_2(y)=\sqrt{\frac{2}{\pi}}e^{-\frac{y^2}{2}}$, $y>0$. (3) Y_3 的密度函数为 $f_3(y)=\frac{1}{2\sqrt{\pi(y-1)}}e^{-\frac{y-1}{4}}$, $y>1$.

73. (1) 先求出常数 a 的值. 由 $\int_{-\infty}^{\infty}f(x)\mathrm{d}x=1$ 可知 $a=9$. 当 $1\le y<2$ 时, 有

$$F_Y(y)=\mathbb{P}(Y=1)+\mathbb{P}(1<Y\le y)=\frac{2}{3}+\frac{y^3}{27}.$$

从而 Y 的分布函数为

$$F_Y(y)=\begin{cases}0, & y<1,\\ \frac{2}{3}+\frac{y^3}{27}, & 1\le y<2,\\ 1, & y\ge 2.\end{cases}$$

(2) 由全概率公式可知

$$\begin{aligned}\mathbb{P}(X\le Y) &= \mathbb{P}(X\le Y|Y=1)\mathbb{P}(Y=1)+\mathbb{P}(X\le Y|Y=2)\mathbb{P}(Y=2)\\ &\quad+\mathbb{P}(X\le Y|1<Y<2)\mathbb{P}(1<Y<2)\\ &= 0+\mathbb{P}(Y=2)+\mathbb{P}(1<X<2)\\ &= \frac{1}{27}+\frac{7}{27}=\frac{8}{27}.\end{aligned}$$

第三章　多维随机变量及其分布

1. C.

2. 13/48.

3. (1). (X, Y) 的分布为

$$\mathbb{P}(X=1, Y=-1)=\mathbb{P}(X=0, Y=0)=\mathbb{P}(X=1, Y=1)=1/3.$$

(2). Z 的分布为 $\mathbb{P}(Z=-1)=\mathbb{P}(Z=0)=\mathbb{P}(Z=1)=1/3$.

4. 随机变量 (X, Y) 的概率分布为:

$$\begin{aligned}&\mathbb{P}(X=0, Y=0)=\mathbb{P}(X=1, Y=0)=1/5, \quad \mathbb{P}(X=0, Y=1)=2/5,\\&\mathbb{P}(X=1, Y=1)=2/15, \qquad \mathbb{P}(X=0, Y=2)=1/15.\end{aligned}$$

5. (1) $\mathbb{P}(X=1|Z=0)=4/9$. (2) (X, Y) 的概率分布为

Y \ X	0	1	2
0	1/4	1/6	1/36
1	1/3	1/9	0
2	1/9	0	0

6. (X_1, X_2) 的联合分布为

$$\mathbb{P}(X_1=k, X_2=n-k)=C_k^n 2^{-n}, \quad k=0,1,2,\cdots,n.$$

X_1 和 X_2 的边缘分布相同, 且

$$\mathbb{P}(X_1=k)=\mathbb{P}(X_2=k)=C_k^n 2^{-n}, \quad k=0,1,2,\cdots,n.$$

7. (X,Y) 的联合分布为:

$$\mathbb{P}(X=0,Y=1)=\mathbb{P}(X=3,Y=3)=1/8,$$
$$\mathbb{P}(X=1,Y=1)=\mathbb{P}(X=2,Y=1)=3/8.$$

8. (X_1,X_2) 的联合分布为

$$\mathbb{P}(X_1=1,X_2=0)=0.8,\ \ \mathbb{P}(X_1=0,X_2=1)=0.1,$$
$$\mathbb{P}(X_1=0,X_2=0)=0.1,\ \ \mathbb{P}(X_1=1,X_2=1)=0.$$

9. (1) (X,Y) 的联合概率分布为

$$\mathbb{P}(X=i,Y=j)=\binom{13}{i}\binom{13}{j}\binom{26}{13-i-j}\Big/\binom{52}{13},\ \ i,j\geq 0,\ i+j\leq 13.$$

(2) 此时 Y 的条件概率分布为

$$\mathbb{P}(Y=j)=\binom{13}{j}\binom{26}{12-j}\Big/\binom{39}{12},\ \ 0\leq j\leq 12.$$

10. $a=0.3,b=0.2.$

11. (1) $\mathbb{P}(X=i,Y=j)=p^2(1-p)^{j-2},\ \ j=2,3,\cdots,i=1,2,\cdots,j-1.$ (2) $\mathbb{P}(X=i)=p(1-p)^{i-1},\ i=1,2,\cdots,$ $\quad\mathbb{P}(Y=j)=(j-1)p^2(1-p)^{j-2},\ j=2,3,\cdots.$

12. C.

13. $\mu(\mu^2+\sigma^2).$

14. 1/2.

15. 1/4.

16. 对离散随机变量, 举例如下:

$Y_1 \backslash X_1$	0	1	$p_{i\cdot}$
0	4/25	6/25	2/5
1	6/25	9/25	3/5
$p_{\cdot j}$	2/5	3/5	

$Y_2 \backslash X_2$	0	1	$p_{i\cdot}$
0	2/15	4/15	2/5
1	4/15	5/15	3/5
$p_{\cdot j}$	2/5	3/5	

则 (X_1, Y_1) 与 (X_2, Y_2) 的边缘分布相同但联合分布不同. 对连续随机变量, 例如 $N(0,0,1,1,0)$ 和 $N(0,0,1,1,1/2)$, 两分布的边缘分布相同但联合分布不同.

17. 假设 (X, Y) 的联合密度函数为

$$f(x,y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2}(1+\sin x \sin y), \quad x, y \in \mathbb{R}.$$

则 $X \sim N(0,1), Y \sim N(0,1)$, 但 (X, Y) 不是二元正态的.

18. (1) $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$;

(2) $f_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi\sigma_1^2(1-\rho^2)}} \exp\left\{ -\frac{(x-\mu_1-\rho\sigma_2^{-1}\sigma_1(y-\mu_2))^2}{2\sigma_1^2(1-\rho^2)} \right\}$, $x \in \mathbb{R}$,

$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi\sigma_2^2(1-\rho^2)}} \exp\left\{ -\frac{(y-\mu_2-\rho\sigma_1^{-1}\sigma_2(x-\mu_1))^2}{2\sigma_2^2(1-\rho^2)} \right\}$, $x \in \mathbb{R}$.

19. (1) $a = 1/\pi^2, b = c = \pi/2$; (2) $\mathbb{P}(X > 0, Y > 0) = 1/4$; (3) X 和 Y 的边缘密度函数相同, 均为 $p(x) = 1/(\pi(1+x^2))$, $x \in \mathbb{R}$.

20. (X, Y) 的联合密度函数为

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

21. (1)

$$F(x,y) = \begin{cases} \sin x \sin y, & 0 < x < \pi/2,\ 0 < y < \pi/2, \\ \sin x, & 0 < x < \pi/2,\ y \geq \pi/2, \\ \sin y, & x \geq \pi/2,\ 0 < y < \pi/2, \\ 1, & x \geq \pi/2,\ y \geq \pi/2, \\ 0, & \text{其他}. \end{cases}$$

(2) $\mathbb{P}(0 < X < \pi/4, \pi/4 < Y < \pi/2) = (\sqrt{2}/2)(1-\sqrt{2}/2)$.

22. $A = 1/\pi$, 条件密度函数 $f_{Y|X}(y|x) = \frac{1}{\sqrt{\pi}} e^{-(x-y)^2}$, $y \in \mathbb{R}$.

23. $A = 6$, $\mathbb{P}(X \leq 0.25 | Y = 0.5) = 9/16$.

24. 随机变量 $Z = XY$ 的分布函数为

$$F_Z(z) = \begin{cases} 0.5\Phi(z), & z < 0, \\ 0.5 + 0.5\Phi(z), & z \geq 0. \end{cases}$$

25. $\mathbb{P}(Z = 0) = \mathbb{P}(Z = 1) = 1/2$.

26. (1) 条件密度函数 $f_{Y|X}(y|x) = \begin{cases} 1/x, & 0 < y < x, \\ 0, & \text{其他}. \end{cases}$ (2) $\mathbb{P}(X \le 1|Y \le 1) = \frac{e-2}{e-1}$.

27. (1) $c = 3/(\pi R^3)$, (2) $r^2(3R-2r)/R^3$.

28. (1) $A = 12$; (2) 独立; (3) $f_Z(z) = 12(e^{-3z} - e^{-4z})$, $(z > 0)$; (4) $f_{X|Z=1}(x) = \frac{e^x}{e-1}\ (0 < x < 1)$; $\mathbb{E}(X|Z=1) = \frac{1}{e-1}$.

29. (1) (X,Y) 的联合密度函数 $f(x,y) = \begin{cases} \frac{9y^2}{x}, & 0 < x < 1,\ 0 < y < x, \\ 0, & \text{其他}. \end{cases}$ (2) Y 的的边缘密度函数 $f_Y(y) = \begin{cases} -9y^2 \ln y, & 0 < y < 1, \\ 0, & \text{其他}. \end{cases}$

30. (1) (X,Y) 的联合分布为 $f(x,y) = \begin{cases} e^{-x}, & 0 < y < x, \\ 0, & \text{其他}. \end{cases}$, (2) Y 的密度函数为 $f_Y(y) = \begin{cases} e^{-y}, & y > 0, \\ 0, & y \le 0. \end{cases}$

31. $\mathbb{P}(Y = n) = 2^{-n-1}$, $n = 0, 1, 2, \cdots$.

32. 记 p 为Bernoulli试验中的成功概率, 则 $(X_1, X_2, \cdots, X_n)$ 的联合分布为

$$\mathbb{P}(X_1 = x_1, \cdots, X_n = x_n) = p^{\sum_i x_i}(1-p)^{n-\sum_i x_i}, \quad x_i = 0, 1;\ i = 1, 2, \cdots, n.$$

33. X 的边缘密度为

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

34. Z 的密度函数为 $f_Z(z) = \begin{cases} \frac{2-z}{z}, & 0 < z < 2 \\ 0, & \text{其他} \end{cases}$.

35. (1) 边缘密度 $f_X(x) = \begin{cases} x, & 0 < x \le 1, \\ 2-x, & 1 < x \le 2, \\ 0, & \text{其他}. \end{cases}$

(2) $f_{X|Y}(x|y) = \begin{cases} \frac{1}{2-2y}, & 0 < y < x < 2-y, \\ 0, & \text{其他}. \end{cases}$

36. (1) (X,Y) 的联合密度为 $f(x,y) = \begin{cases} 2, & 0 \le x \le y \le 1, \\ 0, & \text{其他}. \end{cases}$

(2) 边缘密度分别为

$$f_X(x)=\begin{cases} 2(1-x), & 0\le x\le 1,\\ 0, & 其他,\end{cases}\qquad f_Y(y)=\begin{cases} 2y, & 0\le y\le 1,\\ 0, & 其他.\end{cases}$$

(3) 对任意 $0\le y\le 1$, 条件密度为 $f_X(x|Y=y)=\begin{cases} 1/y, & 0\le x\le y,\\ 0, & 其他.\end{cases}$

(4) $\mathbb{E}(X|Y=y)=y/2$.

37. (1) X,Y 有相同的边缘密度 $f(x)=\begin{cases} x+1, & -1\le z<0,\\ 1-x, & 0\le x\le 1.\end{cases}$

(2) X 和 Y 不独立.

(3) $f_{Y|X}(y|x)=\frac{1}{2(1-x)}I(x-1\le y\le 1-x)$, 其中 $0<x<1$.

38. (U,V) 的联合概率分布为

$$\mathbb{P}(U=0,V=0)=\mathbb{P}(U=1,V=0)=1/4,$$
$$\mathbb{P}(U=1,V=1)=1/2,\ \ \mathbb{P}(U=0,V=1)=0.$$

39. (1) $(X_1,X_2,\cdots,X_n)$ 服从多项分布 $M(m;p_1,p_2,\cdots,p_n)$:

$$\mathbb{P}(X_1=m_1,\cdots,X_n=m_n)=\frac{m!}{m_1!m_2!\cdots m_n!}p_1^{m_1}p_2^{m_2}\cdots p_n^{m_n},$$

其中 $m_1,m_2,\cdots,m_n$ 是任一使得 $m_1+m_2+\cdots+m_n=m$ 的非负整数列.

(2) $X_k\sim B(m,p_k)$.

(3) $\mathbb{P}(X_1=m_1,X_2=m_2)=\frac{m!}{m_1!m_2!(m-m_1-m_2)!}p_1^{m_1}p_2^{m_2}(1-p_1-p_2)^{m-m_1-m_2}$, 其中 m_1,m_2 是任一使得 $m_1+m_2\le m$ 的非负整数.

(4) 条件分布为

$$\begin{aligned}&\mathbb{P}(X_2=m_2,\cdots,X_n=m_n|X_1=m_1)\\ =&\frac{(m-m_1)!}{m_2!\cdots m_n!}\Big(\frac{p_2}{1-p_1}\Big)^{m_2}\cdots\Big(\frac{p_n}{1-p_1}\Big)^{m_n},\end{aligned}$$

其中 $m_2,\cdots,m_n$ 是任一使得 $m_1+m_2+\cdots+m_n=m$ 的非负整数列.

40. X 的边缘密度为 $f_X(x)=\begin{cases} (3/4)(1-x^2), & |x|\le 1,\\ 0, & |x|>1.\end{cases}$

41. D.

42. A.

43. 1/5.

44. 3/4.

45. $\binom{n}{k}\left(\frac{\lambda}{\lambda+\mu}\right)^k\left(\frac{\mu}{\lambda+\mu}\right)^{n-k}$, $B(n,\lambda/(\lambda+\mu))$.

46. (1) (X,Y) 的联合分布为

$$\mathbb{P}(X=0,Y=-1)=\mathbb{P}(X=0,Y=1)=1/4,\ \mathbb{P}(X=1,Y=0)=1/2.$$

(2) X,Y 不独立.

47. (1) $\mathbb{P}(Y=m|X=n)=\binom{n}{m}p^m(1-p)^{n-m},\ \ m=0,1,2,\cdots,n;\ n=0,1,2,\cdots$.

(2) $\mathbb{P}(X=n,Y=m)=\binom{n}{m}p^m(1-p)^{n-m}\cdot\frac{\lambda^n}{n!}e^{-\lambda},\ \ m=0,1,2,\cdots,n;\ n=0,1,2,\cdots$.

48. (1) $\mathbb{P}(X=i,Y=j)=\frac{1}{6}\times\left(\frac{1}{3}\right)^{i-1}$, $i=1,2,\cdots,\ j=3,4,5,6$. $\mathbb{P}(X=i)=\frac{2}{3}\times\left(\frac{1}{3}\right)^{i-1}$, $i=1,2,\cdots,$ $\quad\mathbb{P}(Y=j)=\frac{1}{4}$, $j=3,4,5,6$.

(2) X,Y 独立.

49. (1) $\mathbb{P}(Z\le 1/2|X=0)=1/2$.

(2) Z 的密度函数为 $f_Z(z)=\begin{cases}1/3, & -1\le z<2,\\ 0, & \text{其他}.\end{cases}$

50. $(X,|Y|)$ 的联合密度函数为 $f(x,y)=\begin{cases}\sqrt{\frac{2}{\pi}}e^{-x-y^2/2}, & x\ge 0,y>0,\\ 0, & \text{其他}.\end{cases}$

51. (1) 条件密度为 $f_{Y|X=1/2}(y)=\frac{2+y}{4}$, $(|y|<1)$.

(2) 因为 $\mathbb{P}(X^2\le x,Y^2\le y)=\sqrt{x}\cdot\sqrt{y}=\mathbb{P}(X^2\le x)\mathbb{P}(Y^2\le y)$, 所以 X^2 和 Y^2 相互独立.

52. (1) X,Y 的分布相同, 均为 $(0,1)$ 上的均匀分布; (2) X 和 Y 相互独立.

53. (1) X,Y 的分布相同, 具有密度

$$f_X(z)=f_Y(z)=\begin{cases}\frac{2}{\pi}\sqrt{1-z^2}, & |z|\le 1,\\ 0, & \text{其他}.\end{cases}$$

(2) X 和 Y 不独立.

54. (1) 边缘密度分别为

$$f_X(x)=\begin{cases}3x^2, & 0<x<1,\\ 0, & \text{其他}.\end{cases}\qquad f_Y(y)=\begin{cases}\frac{3}{2}(1-y^2), & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

(2) 不独立. (3) $\mathbb{P}(X+Y\le 1)=3/8$.

55. (1) $F_X(x)=\begin{cases}1-(x+1)e^{-x}, & x>0\\ 0, & x\le 0\end{cases}$, $F_Y(y)=\begin{cases}\frac{y}{1+y}, & y>0\\ 0, & y\le 0.\end{cases}$.

(2) $f(x,y)=\begin{cases}\frac{xe^{-x}}{(1+y)^2}, & x>0,\ y>0\\ 0, & \text{其他}\end{cases}$,

$f_X(x)=\begin{cases}xe^{-x}, & x>0\\ 0, & x\le 0\end{cases}$,

$f_Y(y)=\begin{cases}\frac{1}{(1+y)^2}, & y>0\\ 0, & y\le 0\end{cases}$.

(3) 因为 $F(x,y)=F_X(x)F_Y(y)$ (或 $f(x,y)=f_X(x)f_Y(y)$), 所以 X,Y 独立.

56. (1) 联合密度为 $f(x,y)=\begin{cases}\frac{1}{2}e^{-y/2}, & 0<x<1,\ y>0\\ 0, & \text{其他}.\end{cases}$

(2) 二次方程 $a^2+2Xa+Y=0$ 有实根的概率为

$$\mathbb{P}(X^2\ge Y)=1-\sqrt{2\pi}(\Phi(1)-\Phi(0))=0.1446,$$

其中 $\Phi(x)$ 为标准正态的分布函数.

57. (X,Y), $(X,Z),(Y,Z)$ 的联合分布均为

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

而 X,Y,Z 的边缘分布相等且都等于

$$f(x)=\begin{cases}\frac{1}{2\pi}, & 0\le x\le 2\pi,\\ 0, & \text{其他}.\end{cases}$$

由此立得 X,Y,Z 两两独立. 但 $f_{X,Y,Z}(x,y,z)\ne f_X(x)f_Y(y)f_Z(z)$, 因此 X,Y,Z 不是相互独立的.

58. 因为

$$f_X(x) = f_Y(x) = \begin{cases} 1/2, & |x| < 1, \\ 0, & 其他, \end{cases}$$

所以 $f(x, y) \neq f_X(x) f_Y(y)$, 即 X, Y 不独立. 而

$$f_{X^2,Y^2}(x, y) = \begin{cases} 1/(4\sqrt{xy}), & 0 < x, y < 1, \\ 0, & 其他, \end{cases}$$

由此可得 X^2, Y^2 是相互独立的.

第四章 数字特征

1. (1) $\mathbb{E}X=\frac{1}{p}$; $\mathrm{Var}(X)=\frac{1-p}{p^2}$. (2) $\mathbb{E}X=\frac{r}{p}$; $\mathrm{Var}(X)=\frac{r(1-p)}{p^2}$.

2. 当 $p=0.5$ 时, 我们有 $\mathbb{E}X=4\times\frac{1}{8}+5\times\frac{1}{4}+6\times\frac{5}{16}+7\times\frac{5}{16}=\frac{93}{16}\approx 5.81$ 场. 对一般的 $0<p<1$, 记 $q=1-p$, 则

$$\begin{aligned}\mathbb{E}X &= 4\times(p^4+q^4)+5\times\binom{4}{1}\times(p^4q+pq^4)+6\times\binom{5}{2}\times(p^4q^2+p^2q^4)\\ &\quad +7\times\binom{6}{3}\times(p^4q^3+p^3q^4)\\ &= 4p^4(1+5q+15q^2+35q^3)+4q^4(1+5p+15p^2+35p^3).\end{aligned}$$

从而当 $p=0.6$ 时, 由上式可得 $\mathbb{E}X=\frac{17804}{3125}\approx 5.70$ 场.

3. 证明: (1) 我们只需要证明第一个等式. 由期望的定义可知,

$$\begin{aligned}\mathbb{E}X &= \sum_{k=1}^{\infty}k\mathbb{P}(X=k)=\sum_{k=1}^{\infty}\sum_{n=1}^{k}\mathbb{P}(X=k)\\ &= \sum_{n=1}^{\infty}\sum_{k=n}^{\infty}\mathbb{P}(X=k)=\sum_{n=1}^{\infty}\mathbb{P}(X\geq n).\end{aligned}$$

(2) 证明思路与 (1) 类似. 设 $f(x)$ 为其密度函数, 则

$$\begin{aligned}\mathbb{E}X &= \int_0^{\infty}tf(t)\mathrm{d}t=\int_0^{\infty}f(t)\mathrm{d}t\int_0^t\mathrm{d}x\\ &= \int_0^{\infty}\mathrm{d}x\int_x^{\infty}f(t)\mathrm{d}t=\int_0^{\infty}(1-F(x))\mathrm{d}x.\end{aligned}$$

(3) 以 $I(A)$ 表示事件 A的示性函数, 则由

$$\begin{aligned}\mathbb{E}X &= \mathbb{E}\Big[\int_0^X\mathrm{d}x\Big]=\mathbb{E}\Big[\int_0^{\infty}I(X>x)\mathrm{d}x\Big]\\ &= \int_0^{\infty}\mathbb{E}[I(X>x)]\mathrm{d}x=\int_0^{\infty}\mathbb{P}(X>x)\mathrm{d}x\end{aligned}$$

可知结论成立.

4. 以 X 表示满足要求的随机数的个数, 由 $\mathbb{P}(X>n)=\frac{1}{n!}, n=1,2,\cdots$, 可知所求即为 $\mathbb{E}X=\mathrm{e}$.

5. $n=8, p=0.3$.

6. 2.

7. $c=\frac{1}{\sqrt{\pi}}\mathrm{e}^{-\frac{1}{4}}$, $\mathbb{E}X=\operatorname{Var}(X)=\frac{1}{2}$.

8. $\mathbb{E}X=\operatorname{Var}(X)=n+1$.

9. (1) $\mathbb{E}X=\sqrt{\frac{\pi}{2}}\sigma$, $\operatorname{Var}(X)=\frac{4-\pi}{2}\sigma^2$.

   (2) $\mathbb{E}X=\frac{\alpha}{\alpha+\beta}$, $\operatorname{Var}(X)=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.

   (3) $\mathbb{E}X=\lambda\Gamma(1+\frac{1}{k})$, $\operatorname{Var}(X)=\lambda^2[\Gamma(1+\frac{2}{k})-(\Gamma(1+\frac{1}{k}))^2]$.

10. $a=12, b=-12, c=3$.

11. e^{-1}.

12. 1.

13. 7.5.

14. $[-1,1]$ 中的任一点都是 X 的中位数.

15. X 的 α 分位数为 $\Phi^{-1}(\alpha)$, 其中 $\Phi(x)$ 为标准正态分布函数而 $\Phi^{-1}(x)$ 为 $\Phi(x)$ 的逆函数.

16. EX.

17. $m(X)$ (X 的中位数).

18. $n(1-\frac{1}{n})^n$; e^{-1}.

19. (1) $\mathbb{E}[X_n]=n\sum\limits_{k=1}^{n}\frac{1}{k}=12\sum\limits_{k=1}^{12}\frac{1}{k}\approx 37.24$; (2) 1.

20. $\frac{na}{a+b}$.

21. 2.

22. 3.

23. 证明: 不妨设

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

其中 $p_i > 0(i = 1, 2, \cdots, k), \sum_{i=1}^{k} p_i = 1$, 且 $x_1 > x_2 > \cdots > x_k > 0$. 由于

$$\mathbb{E}[X^n] = \sum_{i=1}^{k} x_i^n p_i,$$

我们有

$$\lim_{n\to\infty} \frac{\mathbb{E}[X^n]}{x_1^n} = p_i.$$

由此,

$$\lim_{n\to\infty} \frac{\mathbb{E}[X^{n+1}]}{\mathbb{E}[X^n]} = \lim_{n\to\infty} \frac{\mathbb{E}[X^{n+1}]}{x_1^{n+1}} \frac{x_1^n \cdot x_1}{\mathbb{E}[X^n]} = x_1 = \max_{1\le i\le k} x_i.$$

24. $(p\mathrm{e}^t + 1 - p)^n; \exp\{\lambda(\mathrm{e}^t - 1)\}; \frac{\lambda}{\lambda - t}; \exp\{\mu t + \frac{\sigma^2 t^2}{2}\}$.

25. (1) 验证 $Y = cX$ 的密度函数为 $p(y) = \frac{\lambda}{c}\mathrm{e}^{-\lambda y/c}$, $y > 0$. (2) $n!/\lambda^n$.

26. $\mathbb{E}Y = 2\mathrm{e}; \mathbb{E}Z = 2(1 - \ln 2)$.

27. X 的密度函数为

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\Big\{ -\frac{(\ln x - \mu)^2}{2\sigma^2} \Big\}, \quad x > 0;$$

期望和方差分别为

$$\mathbb{E}X = \mathrm{e}^{\mu + \frac{\sigma^2}{2}}, \quad \mathrm{Var}(X) = (\mathrm{e}^{\sigma^2} - 1)\,\mathrm{e}^{2\mu + \sigma^2}.$$

28. $0, \frac{2}{\pi}, 0$.

29. $\frac{an + bm}{a + b}$.

30. (1) $\frac{8}{9}$; (2) $\sqrt{\frac{2}{\pi}}$.

31. $\frac{\ln 2}{\pi} + \frac{1}{2}$.

32. (1) Y 的分布函数为

$$F(y)=\begin{cases} 0, & y<0,\\ \frac{3y}{4}, & 0\le y<1,\\ \frac{y}{4}+\frac{1}{2}, & 1\le y<2,\\ 1, & y\ge 2.\end{cases}$$

(2) $\mathbb{E}Y=\frac{3}{4}$.

33. (1) $f_Z(z)=\frac{2}{\sqrt{\pi}}e^{-z^2}$, $z>0$; $\mathbb{E}[Z]=\frac{1}{\sqrt{\pi}}$; (2) $\mathbb{E}[U]=1+\frac{1}{2\sqrt{\pi}}$, $\mathbb{E}[V]=1-\frac{1}{2\sqrt{\pi}}$

34. (1)$\mathbb{E}[X_2|X_1=k]=\frac{(m-k)p_2}{1-p_1}$, $\quad \mathrm{Var}(X_2|X_1=k)=\frac{(m-k)p_2(1-p_1-p_2)}{(1-p_1)^2}$.
(2)$\mathbb{E}[X_1+X_2]=m(p_1+p_2)$,
$\mathrm{Var}(X_1+\cdots+X_k)=m\sum_{i=1}^k p_i(1-p_i)-2m\sum_{i=1}^{k-1}\sum_{j=i+1}^k p_ip_j$,$k=1,\ldots,n$.

35. $\frac{2(21-\sqrt{3})}{73}$

36. (1) $\mathbb{E}((X+Y-3\mu)_+)=\frac{2\sigma}{\sqrt{2\pi}}$ (2)$\mathrm{Var}((X+Y-3\mu)_+)=4\left(1-\frac{1}{2\pi}\right)\sigma^2$

37. (1)$f_Y(y)=\frac{1}{3}\sum_{k=0}^2\frac{1}{\sqrt{2\pi}}e^{-\frac{(y-k)^2}{2}}$, $\quad \mathbb{E}(Y)=1$.
(2)$F_{X+Y}(z)=\frac{1}{3}\sum_{k=0}^2\phi(z-2k)$, 其中$\phi$为标准正态分布的分布函数.
(3)$\frac{2}{3}$

38. 1.

39. 期望为$\frac{1}{\lambda}\sum_{k=1}^n\frac{1}{k}$.

40. 12.5.

41. $\mathbb{E}(\min\{X_1,X_2\})=\frac{1}{4}$, $\quad \mathbb{E}(\max\{X_1,X_2\})=\frac{3}{4}$.

42. 略.

43. 答案：D.

44. $\mathbb{E}[X_1+X_2+X_3]=0$, $\quad \mathrm{Var}(X_1+X_2+X_3)=\frac{1}{3}$.

45. 略.

46. 期望为1000，方差为200.

47. 均值为4.2，方差为0.072.

48. (1)$\log n$,　　(2)证明略.

49. $\frac{121}{12}$

50. 2

51. $\frac{2}{3}$

52. (1)$\mathbb{E}[y]=1$, $\mathbb{E}[Z]=3$ (2)$\mathrm{Var}(Y)=1$, $\mathrm{Var}(Z)=5$

53. $\frac{k}{2}(n+1)$, $\frac{k}{12}(n^2-1)$

54. $12p(1-p)+3$

55. $\frac{1}{3}$

56. B

57. -1

58. $-\frac{1}{4}$

59. C

60. 0

61. 0.275, -0.005

62. (1)3.5, $\frac{35}{12}$, $\frac{161}{36}$, 1.97 (2)$\frac{154}{9}$, $\frac{35}{24}$

63. (1)$(\alpha^2-\beta^2)\sigma^2$; (2) $\alpha=\beta$或$\alpha=-\beta$

64. $\alpha=\beta$或$\alpha=-\beta$

65. B

66. 0, 不独立

67. 2000, $\frac{35000000}{9}$

68. 证明略

69. (1) 证明略; (2) 0.5

70. 证明略

71. 证明略

72. $\frac{n}{p}, \frac{np}{(1-p)^2}, \frac{npq}{(1-p)^2} + \frac{nq(1-q)}{p^2}$

73. $0, x^2$

74. $(1)\lambda b$ $(2)\lambda^2 b + \lambda a$

75. $\frac{a}{1-p_0}$, $\frac{\sigma^2(1-p_0)-a^2 p_0}{(1-p_0)^2}$

76. 31.5

77. 1, $\frac{5}{3}$

78. 5.625

79. $2p$

80. 33/20

81. $(1)\binom{m}{k}(\frac{\lambda}{\lambda+\mu})^k(\frac{\mu}{\lambda+\mu})^{m-k}$, $\frac{m\lambda}{\lambda+\mu}$ $(2)\frac{\binom{n}{k}\binom{n}{m-k}}{\Sigma_{i=1}^{m}\binom{n}{i}\binom{n}{m-i}}$, $\frac{m}{2}$

82. $\frac{1}{2}z$

83. 证明略

第五章 极限理论

1. 180000
2. $1/\epsilon$
3. 4
4. 5
5. $b + a^2$
6. $0.8, 0.97$
7. $\frac{\sum_{i=1}^{n} X_i^2 - n\alpha_2}{\sqrt{n(\alpha_4 - \alpha_2^2)}} \xrightarrow{\mathrm{d}} N(0,1).$
8. 0.79
9. 0.305
10. 0.08
11. 0.608
12. 0.876, 0.0002
13. (1)0.95, (2)9
14. 1167
15. (1)0.18, (2)443
16. (1)0.82, (2)10
17. 0.967

第六章 数理统计的基本概念

1. 在试验之前样本观测值是未知的,所以可以看成是随机变量; 而当试验完成之后样本又是一组确定的值, 故可以看成是一组确定的值. 视样本为随机变量时候方便研究其性质.

2. 总体即为该人射击的所有可能环数, 令为$\Omega = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$; 记他射击5次的环数为$X_1, \ldots, X_5$, 即为样本, 样本值分别为8,9,7,10和6.

3. 总体为全班同学, 样本为选择的5名同学. 抽样结束后, 则样本值为特定的5名同学.

4. 记总体为X, 则在不放回方式下$\mathbb{P}(X = k) = \binom{2}{k}\binom{4}{2-k}/\binom{6}{2}, k = 0, 1, 2.$; 在有放回方式下$\mathbb{P}(X = k) = \binom{2}{k}(2/6)^k(4/6)^{2-k}, k = 0, 1, 2.$

5. 记总体为X, 则$X \sim B(1, p)$. 记样本为$X_1, \ldots, X_{10}$, 则样本分布为$\mathbb{P}(X_1 = x_1, \ldots, X_{10} = x_{10}) = p^{\sum_{i=1}^{10} x_i}(1-p)^{10-\sum_{i=1}^{10} x_i}$.

6. 记物体真实长度为a, 则测量值依赖于测量误差, 因此可记总体为$X = a + \epsilon$, 由于测量误差常常由多种不同类型误差综合组成, 故由中心极限定理, 可以假设$\epsilon \sim N(0, \sigma^2)$, 因此总体$X \sim N(a, \sigma^2)$, 这里$\sigma^2$反映了测量精度.

7. 可以取总体为指数分布或者威布尔分布. 由于寿命为非负值, 且失效率往往会随时间增加(为简单记, 也常假设为常数), 因此指数分布或威布尔分布是合适的.

8. $\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = 1/\binom{M}{m}\binom{N-M}{n-m}$, 其中$(x_1, \ldots, x_m) \subset \{a_1, \ldots, a_M\}$, $(x_{m+1}, \ldots, x_n) \subset \{a_{M+1}, \ldots, a_N\}$.

9. (1) 样本空间为$\Omega = \{(x_1, \ldots, x_5) : x_i = 0, 1, 1 \le i \le 5\}$, 抽样分布为$\mathbb{P}(X_1 = x_1, \ldots, X_5 = x_5) = p^{\sum_{i=1}^{5} x_i}(1-p)^{5-\sum_{i=1}^{5} x_i}$. (2) $X_1 + X_2$和$\min_{1 \le i \le 5} X_i$为统计量, 其余因为依赖于未知参数p, 因此不是统计量.

10.

$$F_n(x) = \begin{cases} 0, & x < -0.3 \\ 1/10, & -0.3 \leq x < -0.1 \\ 3/10, & -0.1 \leq x < 0.3 \\ 4/10, & 0.3 \leq x < 0.4 \\ 5/10, & 0.4 \leq x < 0.5 \\ 6/10, & 0.5 \leq x < 0.6 \\ 7/10, & 0.6 \leq x < 0.9 \\ 8/10, & 0.9 \leq x < 1.7 \\ 9/10, & 1.7 \leq x < 2.6 \\ 1, & 2.6 \leq x. \end{cases}$$

11. 样本均值73.98929, 样本标准差0.03590596.

12. C

13. D

14. $F_{2,2}$

15. $a = 1/20, b = 1/100$.

16. t_2

17. $F_{10,5}$

18. (a) $N(a, \sigma^2/n)$; (b) 利用 $n\bar{X} \sim \mathbb{P}(n\lambda)$ 易得; (c) 利用 $2n\lambda\bar{X} \sim \chi^2_{2n}$ 易得.

19. $(n-1)p(1-p)/n$

20. t_{n-1}

21. t_{n+m-2}

22. (1) 299 (2) $(n-1)/(n+1)$

第七章 参数估计

1. 0.38

2. 利用EX, EX^2得到$\hat{p}_1 = n_1/n$, $\hat{p}_2 = n_2/n$.

3. (1) $2\bar{X}+1$ (2) $\hat{\theta} = 1 - S^2/\bar{X}$ (3) $2/\bar{X}$ (4) $1-\bar{X}/a_2$, 其中a_2为二阶原点矩. (5) $\bar{X}$

4. (1) $3\bar{X}$ (2) $(2\bar{X}-1)/(1-\bar{X})$ (3) $(\bar{X})^2/(1-\bar{X})^2$ (4) $\bar{X}/(\bar{X}-c)$ (5) $2\bar{X}$ (6) $\bar{X}$

5. (1) $\sqrt{\pi}/2\bar{X}$ (2) $Var(\hat{\theta}) = \frac{3\pi-8}{8n}\theta^2$

6. (1) $2\bar{X}-1/2$ (2) 不是无偏估计

7. (1) $e^{-\bar{X}}$ (2) $e^{-1.12}$

8. 2

9. $\hat{\theta} = 0.3866$.

10. (1) $max_i X_i + 1$ (2) $\hat{\theta} = \bar{X}$ (3) $\hat{\theta} = 2/\bar{X}$ (4) 无显式解, $\hat{\theta} = \arg\max \theta^{\sum x_i}/(-ln(1-\theta))^n$ (5) $\hat{\theta} = \bar{X}$

11. (1) 方程$\sum_i \frac{1}{\theta - x_i} = \frac{2n}{\theta}$的根 (2) $\hat{\theta} = -\frac{n}{\sum_i log(x_i)} - 1$ (3) $\hat{\theta} = [-n/\sum_i log(x_i)]^2$ (4) $\hat{\theta} = n/(\sum_i log(x_i) - nlogc)$ (5) 方程$\sum_i \frac{1}{x_i(\theta - x_i)} = \frac{3}{\theta^2}$的根 (6) $\hat{\theta} = \frac{2}{\frac{1}{n}\sum_i 1/x_i}$

12. (1) 利用密度变换公式和指数分布的无记忆性. (2) $\hat{\lambda} = T/(2r)$

13. $\hat{\theta} = \frac{2n_1+n_2}{2n}$

14. $\hat{\theta}_1 = \min_i X_i, \hat{\theta}_2 = \max_i X_i$.

15. $1-\Phi(\frac{2-\bar{X}}{\hat{\sigma}}), \hat{\sigma} = \frac{1}{n}\sum_i (X_i - \bar{X})^2$.

16. $\hat{\theta}=\frac{n_0}{n_0+n_1}, n_0=\sum_i I(0<X_i<1), n_1=\sum_i I(1\le X_i<2)$

17. (1) $\sqrt{\pi\theta}2, \theta$ (2) $\frac{1}{n}\sum_i X_i^2$ (3) $a=\theta$

18. 略

19. (1) $g(t)=9t^8/\theta^9 I(0<t<\theta$ (2) $10/9$

20. (1) $1/(2(n-1))$ (2) $1/n$

21. $a=n_1/(n_1+n_2)$

22. $a_i=\frac{1/\sigma_i^2}{\sum_i 1/\sigma_i^2}$

23. (1) $\hat{\theta}=\max_i X_i/c$ (2) $\tilde{\theta}=2\bar{X}/(c+1)$, 为无偏估计

24. $\bar{X}, S^2+\bar{X}$

25. $\hat{p}_3$的方差最小

26. $a_1=0, a_2=a_3=1/n, Var(T)=\theta(1-\theta)/n$

27. 矩估计为$\hat{\theta}=N(2-\bar{X})=N(2-(n_1+2n_2)/n)$; 最大似然估计为$\tilde{\theta}=N*4(n-n_2)/(3n)$

28. (1) $\hat{\theta}_1=\bar{X}-\sigma, \hat{\theta}_2=\min_i X_i$ (2) $\tilde{\theta}_1=\hat{\theta}_1, \tilde{\theta}_2=\hat{\theta}_2-\sigma/n$ (3) $\tilde{\theta}_2$更优

29. (1) $\hat{\sigma}=\sqrt{\pi/2}\frac{1}{n}\sum_i|X_i|$ (2) $\hat{\sigma}=S$, 其中$S^2=\frac{1}{n}\sum_i X_i^2$.

30. $\frac{1}{n}\sum_i I(X_i>1)$

31. 利用EY, EY^2得到矩估计 $\hat{\mu}=-\frac{ln(\sum_i Y_i^2)}{2}+2ln(\sum_i Y_i)-\frac{3}{2}ln(n), \hat{\sigma}^2=ln(\sum_i Y_i^2)-2ln(\sum_i Y_i)+ln(n)$. (2) 最大似然估计为$\tilde{\mu}=\sum_i ln(Y_i)/n, \tilde{\sigma}^2=\frac{1}{n}\sum_i(ln(Y_i)-\tilde{\mu})^2$.

32. 矩估计$\hat{a}=\bar{X}, \hat{\sigma}=S^2/2$; 最大似然估计$\tilde{a}=m, \tilde{\sigma}=\frac{1}{n}\sum_i|X_i-m|$，其中$m$为样本中位数.

33. (1) 矩估计 $\hat{\lambda}=[\Gamma(1/\alpha+1)/\bar{X}]^\alpha$ (2) 最大似然估计为$\tilde{\lambda}=1/\sum_i X_i^\alpha$.

34. $[\max_i X_i-1/2, \min_i X_i+1/2]$

35. $\max_i X_i/2$, 可修正为$\frac{n+1}{2n+1}\max_i X_i$

36. $\hat{\mu}_1=\bar{X}, \hat{\mu}_2=\bar{Y}, \hat{\sigma}^2=\frac{1}{n+m}[\sum_i(X_i-\bar{X})^2+\sum_j(Y_j-\bar{Y})^2]$.

37. $\hat{\mu} = \frac{2m\bar{X}+n\bar{Y}}{2m+n}, \hat{\sigma}^2 = \frac{1}{n+m}[\sum_i (X_i - \hat{\mu})^2 + \frac{1}{2}\sum_j (Y_j - \hat{\mu})^2]$.

38. 15000

39. k/n

40. (1) $\hat{\mu}* = \min_i X_i$, 可以修正为$\hat{\mu}^{**} = \hat{\mu} * -1/n$ (2) $\hat{\mu} = \bar{X}$ (3) $\hat{\mu}^{**}$更有效

41. 略

42. 略

43. 不是

44. $\hat{\theta}_1 = \min_i X_i, \hat{\theta}_2 = \max_i X_i$

45. 后者

46. 前者更有效

47. 1)证明略, 2) $a = \frac{n_1}{n_1+n_2+n_3}, b = \frac{n_2}{n_1+n_2+n_3}, c = \frac{n_3}{n_1+n_2+n_3}$.

48. 1) $a_n = 2$, $b_n = n+1$, $c_n = \frac{n+1}{n}$, 2) $\hat{\theta}_3$更有效.

49. 1) $a_n = 1$, $b_n = n$, 2) $\hat{\theta}_1$ 更有效.

50. 1) $a_n = -1$, $b_n = -\frac{1}{n}$, 2) $\hat{\theta}_2$ 更有效.

51. 1) $\hat{\theta}_M = (3-\bar{X})/5 = 2/5$, $\hat{\theta}_L = \frac{n-n_3}{3n} = \frac{4}{15}$, 2) $\hat{\theta}_M$,$\hat{\theta}_L$都为θ的无偏估计, 3)$\hat{\theta}_L$更有效.

52. 证明略

53. 证明略

54. 证明略

55. $c_n = \alpha^{-1/n}$.

56. 1)证明略, 2) $1-\alpha$ 置信区间 $\left[\frac{\chi^2_{2n}(1-\alpha/2)}{\sum_{i=1}^n 2X_i}, \frac{\chi^2_{2n}(\alpha/2)}{\sum_{i=1}^n 2X_i}\right]$.

57. $n = 25$.

58. $n = 166$.

59. C.

60. A.

61. $[46.76, 49.24]$.

62. $[7712.17, 8287.83]$.

63. $[2.120, 2.145]$.

64. 1) [119.80, 124.54], 2) [118.69, 127.53], 3) [118.43, 124.01].

65. 1) $e^{\mu+1/2}$, 2)[-0.922,1.038], 3) [0.657,4.654].

66. 1)[214.112, 225.888], 2) [154.44,165.55].

67. 1)[71.03, 110.97], 2) [72.22,114.96], 3)[-10.06, 4.88].

68. $[0.307, 7.060]$.

69. 1)[43.04, 52.96], 2) [29.20, 234.89].

70. 1)[3.189, 3.211], 2) [0.024,0.043].

71. 1)[0.141, 0.893], 2) [0.149, 1.050].

72. 1)[60.92, 384.28], [68.16, 316.67], 2) [65.46, 461.14], [73.60,374.50]

73. $[8.86, 18.57], [78.58, 344.93]$.

74. 1)$[7400.41, 8904.23]$, 2)$[5816.02, 7077.04]$.

75. 1)[431.37, 2721.25], 2) [494.01, 3116.42], 3)[60.36, 380.75].

76. $[0.015, 0.045]$.

77. $[0.00215, 0.01170]$.

78. 1) $\left[\frac{t}{s} \pm \sqrt{\frac{t/s(1-t/s)}{s}}u_{\alpha/2}\right]$, 2) $\left[\frac{r}{\frac{t}{s}+\sqrt{\frac{t/s(1-t/s)}{s}}u_{\alpha/2}}, \frac{r}{\frac{t}{s}-\sqrt{\frac{t/s(1-t/s)}{s}}u_{\alpha/2}}\right]$.

79. $[341, 478]$.

80. $1-\alpha$置信上界 $\alpha^{1/n}\theta$, $1-\alpha$ 置信下界 $(1-\alpha)^{1/n}\theta$.

81. 至少6块.

82. 792.

83. 1)0.0368, 2)0.0340.

84. 1)120.21, 2)119.54, 3)118.97.

85. 41147.53

86. 1) 1593.426, 2) 464.82

第八章 假设检验

1. (1) 拒绝域 $\{\bar{X} > 0.6645\}$, (2) 第二类错误概率0.5576.
2. α应取大些，因为“减少次品混入正品的可能性”为减少第二类错误的概率.
3. (1) 0.1, 2) V_3.
4. 63/64, 1/16, 7/8
5. (1)0.287, 0.046, 0.304, 0.046 (2) 0.046, 0.713, 0.916.
6. (1) $2.5^n/\theta^n$, $(2.5/3)^n$ (2) 17.
7. 拒绝H_0.
8. 接受H_0.
9. 接受H_0.
10. 25.
11. 原假设：废品率为0.06, 备择假设：废品率小于0.06, 2) 0.05, 3) 不支持备择假设.
12. 合格.
13. (1)接受H_0, (2)拒绝H_0 , (3)拒绝H_0.
14. 备择假设:平均起薪大于7700, p值0.0207.
15. 备择假设都是平均身高大于120, (1)接受备择假设,(2)拒绝备择假设, (3)拒绝备择假设.
16. (1)备择假设：包月客户一个月平均通话时间大于190分钟, 接受备择假设. 2)备择假设：按流量收费的客户一个月平均通话时间小于190分钟, 接受备择假设.

17. 备择假设：更换经营策略后平均销量大于更换经营策略前平均销量, 拒绝备择假设.

18. 接受原假设.

19. 不能认为该班的数学平均成绩为 87 分.

20. 接受原假设.

21. 仍然显著高于全市平均成绩.

22. 显著小于全国人口出生率, p值5.110357e-72.

23. 备择假设：合格, 接受备择假设.

24. 比旧工艺有所提高.

25. 在 0.05的显著性水平下检验机器工作良好, 在0.01的显著性水平下不能认为检验机器工作良好.

26. 接受H_0.

27. (1)接受H_1, (2) 拒绝H_1.

28. 接受H_1.

29. 接受H_0.

30. (1)接受H_1, (2) 接受H_1.

31. C

32. C

33. B

34. 由题意，令$Z_i = X_i - Y_i$为各男孩 A, B 材料耐磨性的差异，若 A, B 材料无差异，则在n个试验单元中Z_i取" $+$ "和" $-$ "的概率都为$\frac{1}{2}$,因此假设检验问题转化为$n_+ \sim b(n,p)$，$0 \le p \le 1$, 检验问题为

$$H_0:\ p = \tfrac{1}{2},\ \ H_1:\ p \neq \tfrac{1}{2}.$$

拒绝域$D = \{n_+ \ge c, n_+ \le d\}$ 在本例中$n = 10$,$\alpha = 0.05$, 查二分分布表知

$$\sum_{k=0}^{1}\binom{10}{k}(\tfrac{1}{2})^{10}=0.011$$
$$\sum_{k=0}^{2}\binom{10}{k}(\tfrac{1}{2})^{10}=0.055$$

故水平$\alpha=0.05$的符号检验的拒绝域为$D=\{n_+\geq 1, n_+\leq 9\}$,样本中的差值取正数的个数为$n_+=2$,所以我们认为原假设成立，即两种材料的耐磨性无显著性差异.

35. 由题意，设 $X_1,\cdots,X_m\ iid\sim N(\mu_1,\sigma_1^2)$, $Y_1,\cdots,Y_m\ iid\sim N(\mu_2,\sigma_2^2)$,此处$m=9$,故此问题属于Behrens-Fisher问题，用基于t分布的检验方法. 只能在方差未知的一般情形下，检验假设

$$H_0:\ \mu_1-8\geq\mu_2,\ \ H_1:\ \mu_1-8<\mu_2.$$

检验的拒绝域为

$$D=\{(X,Y):T=\tfrac{\overline{X}-\overline{Y}-8}{\sqrt{S_1^2/m+S_2^2/m}}<-t_r(\alpha)\}.$$

首先计算t分布的自由度为

$$r=\tfrac{(S_1^2/m+S_2^2/m)^2}{S_1^4/m^2(m-1)+S_2^4/m^2(m-1)}=15.99\approx 16,$$

查表得$t_{16}(0.05)=1.7459$,由数据得

$$T=\tfrac{\overline{X}-\overline{Y}-8}{\sqrt{S_1^2/m+S_2^2/m}}=0.019>-1.7459$$

即认为原假设成立，俱乐部的宣传可信.

36. 由题意，在$\sigma_1^2=\sigma_2^2=\sigma^2$未知时，检验假设

$$H_0:\ \mu_2-\mu_1\leq -8,\ \ H_1:\ \mu_2-\mu_1>-8.$$

检验的拒绝域为

$$T_w=\tfrac{\overline{Y}-\overline{X}+8}{S_w}\sqrt{\tfrac{mn}{m+n}}$$
$$S_w^2=\tfrac{(m-1)S_1^2+(n-1)S_2^2}{m+n-2}$$
$$D=\{(X,Y):T_w>t_{m+n-2}(\alpha)\}$$

经计算得$T_w=-0.027<1.7459$,所以我们认为原假设成立，俱乐部的宣传可信.

37. 由题意，在$\sigma_1^2=\sigma_2^2=\sigma^2$未知时，检验假设

$$H_0:\ \mu_2-\mu_1=0,\ \ H_1:\ \mu_2-\mu_1\neq 0.$$

检验的拒绝域为

$$T_w=\frac{\overline{Y}-\overline{X}}{S_w}\sqrt{\frac{mn}{m+n}}$$
$$S_w^2=\frac{(m-1)S_1^2+(n-1)S_2^2}{m+n-2}$$
$$D=\{(X,Y):|T_w|>t_{m+n-2}(\tfrac{\alpha}{2})\}$$

经计算得$|T_w|=2.4627>2.0739$,所以我们认为原假设不成立，甲乙两种方法有显著不同.

38. 略

39. 略

40. (1) 由题意，在正态总体均值未知的情况下，检验假设

$$H_0:\ \sigma_1^2=\sigma_2^2,\ \ H_1:\ \sigma_1^2\neq\sigma_2^2.$$

检验的拒绝域为

$$F=\frac{S_2^2}{S_1^2}$$
$$F<F_{n-1,m-1}(1-\alpha/2)\text{或}$$
$$F>F_{n-1,m-1}(\alpha/2)$$

经计算得，$S_1^2=129447.6$,$S_2^2=181026.8$,$F_{7,6}(0.025)=5.12$,$F_{7,6}(1-0.025)=0.175$. 显然不在拒绝域内，所以我们认为两者的方差相等 (2) 由题意，在正态总体均值方差未知的情况下，检验假设

$$H_0:\ \mu_2-\mu_1\geq 0,\ \ H_1:\ \mu_2-\mu_1<0.$$

检验的拒绝域为

$$D=\{(X,Y):T=\frac{\overline{Y}-\overline{X}}{\sqrt{S_1^2/m+S_2^2/n}}<-t_r(\alpha)\}.$$

首先计算t分布的自由度为

$$r=\frac{(S_1^2/m+S_2^2/n)^2}{S_1^4/m^2(m-1)+S_2^4/n^2(n-1)}=12.99\approx 13,$$

查表得$t_{13}(0.05)=1.7709$,由数据得

$$T=\frac{\overline{Y}-\overline{X}}{\sqrt{S_1^2/m+S_2^2/n}}=1.848>-1.7709$$

即认为原假设成立，甲企业的平均工资低于乙企业.

41. 略

42. 由题意，在正态总体均值未知的情况下，检验假设

$$H_0:\ \sigma_1^2\geq\sigma_2^2,\ \ H_1:\ \sigma_1^2<\sigma_2^2.$$

检验的拒绝域为

$$F=\frac{S_2^2}{S_1^2}$$
$$F>F_{n-1,m-1}(\alpha)$$

经计算得，$S_1^2=0.0015$,$S_2^2=0.0086$,$F_{8,8}(0.05)=3.44$. 显然$F=\frac{0.0086}{0.0015}>3.44$,所以原假设不成立，即$\sigma_1^2<\sigma_2^2$.

43. 假定病人服用A种药一个固定时间后身体细胞内药的浓度服从$N(\mu_A,\sigma_A^2)$,服用B种药一个固定时间后身体细胞内药的浓度服从$N(\mu_B,\sigma_B^2)$.现将两个总体中分别抽取了容量为$m=8,n=6$的两个样本.欲检验$H_0:\sigma_A^2\geqslant\sigma_B^2\ v.s.\ H_1:\sigma_A^2<\sigma_B^2$.检验统计量为:$F=S_A^2/S_B^2$,其中$S_A^2,S_B^2$分别是两个样本的方差.显著性水平为$\alpha$时的检验规则为:
当$F\leqslant F_{1-\alpha}(m-1,n-1)$时拒绝$H_0$.
由样本数据计算得:$S_A^2=0.0192,S_B^2=0.0335,F=0.5731$,而$\alpha=0.1$ 时,
$F_{0.9}(7,5)=\frac{1}{F_{0.1}(5,7)}=\frac{1}{2.88}=0.3472$, 因此无法拒绝$H_0$.

44. 由题意，在$\sigma_1^2=\sigma_2^2=\sigma^2$未知时，检验假设

$$H_0:\ \mu_2-\mu_1=0,\ \ H_1:\ \mu_2-\mu_1\neq 0.$$

检验的拒绝域为

$$T_w=\frac{\overline{Y}-\overline{X}}{S_w}\sqrt{\frac{mn}{m+n}}$$
$$S_w^2=\frac{(m-1)S_1^2+(n-1)S_2^2}{m+n-2}$$
$$D=\{(X,Y):|T_w|>t_{m+n-2}(\tfrac{\alpha}{2})\}$$

经计算得$|T_w| = 2.245 < 2.3646$,所以我们认为原假设成立，甲乙两处煤矿含灰率无显著差异.

45. 由题意，在$\sigma_1^2 = 0.04$,$\sigma_2^2 = 0.09$的情况下，检验假设

$$H_0:\ \mu_2 - \mu_1 = 0,\ \ H_1:\ \mu_2 - \mu_1 \neq 0.$$

检验的拒绝域为

$$U = \frac{\overline{Y} - \overline{X}}{\sqrt{\sigma_1^2/m + \sigma_2^2/n}}$$
$$|U| > u_{\alpha/2}$$

经计算得$|U| = 0.1463 < 1.96$,所以我们认为原假设成立，μ_1,μ_2 无显著差异.

46. 由题意，在正态总体均值方差未知的情况下，检验假设

$$H_0:\ \mu_2 - \mu_1 = 0,\ \ H_1:\ \mu_2 - \mu_1 \neq 0.$$

检验的拒绝域为

$$D = \{(X,Y) : |T| = |\frac{\overline{Y} - \overline{X}}{\sqrt{S_1^2/m + S_2^2/n}}| > t_r(\alpha/2)\}.$$

首先计算t分布的自由度为

$$r = \frac{(S_1^2/m + S_2^2/n)^2}{S_1^4/m^2(m-1) + S_2^4/n^2(n-1)} = 9.97 \approx 10,$$

查表得$t_10(0.025) = 2.228$,由数据得

$$|T| = |\frac{\overline{Y} - \overline{X}}{\sqrt{S_1^2/m + S_2^2/n}}| = 1.37 < 2.228$$

即认为原假设成立，两批电子原件电阻无显著差异.

47. 略

48. 略

49. 略

50. 略

51. 由题意，在正态总体均值方差未知的情况下，检验假设

$$H_0:\ \mu_2-\mu_1=0,\ \ H_1:\ \mu_2-\mu_1\neq 0.$$

检验的拒绝域为

$$D=\{(X,Y):|T|=|\tfrac{\overline{Y}-\overline{X}}{\sqrt{S_1^2/m+S_2^2/n}}|>t_r(\alpha/2)\}.$$

首先计算t分布的自由度为

$$r=\tfrac{(S_1^2/m+S_2^2/n)^2}{S_1^4/m^2(m-1)+S_2^4/n^2(n-1)}=15.82\approx 16,$$

查表得$t_16(0.025)=2.120$,由数据得

$$|T|=|\tfrac{\overline{Y}-\overline{X}}{\sqrt{S_1^2/m+S_2^2/n}}|=0.108<2.120$$

即认为原假设成立，两种工艺对产品该性能指标的影响无显著性差异.

52. 由题意,(1)假定旧肥料产量服从$N(\mu_A,\sigma_A^2)$,新肥料产量服从$N(\mu_B,\sigma_B^2)$. 现将两个总体中分别抽取了容量为$m=6,n=6$ 的两个样本.欲检验$H_0:\sigma_A^2=\sigma_B^2\ v.s.\ H_1:\sigma_A^2\neq\sigma_B^2$.检验统计量为:$F=S_A^2/S_B^2$,其中$S_A^2,S_B^2$分别是两个样本的方差.显著性水平为$\alpha$时的检验规则为:
当$F\geqslant F_{\alpha/2}(m-1,n-1)$或$F\leqslant F_{1-\alpha/2}(m-1,n-1)$时拒绝$H_0$.
由样本数据计算得:$S_A^2=5.6,S_B^2=4,F=1.4$, 而$\alpha=0.05$ 时,$F_{0.025}(5,5)=7.146$, $F_{0.975}(5,5)=\frac{1}{F_{0.025}(5,5)}=\frac{1}{7.146}=0.1399$, 因此无法拒绝$H_0$.
(2)假定旧肥料产量服从$N(\mu_A,\sigma_A^2)$,新肥料产量服从$N(\mu_B,\sigma_B^2)$. 现将两个总体中分别抽取了容量为$m=6,n=6$的两个样本.欲检验$H_0:\mu_A\geqslant\mu_B\ v.s.\ H_1:\mu_A<\mu_B$.因为由第一问,我们可假定$\sigma_A^2=\sigma_B^2$,所以,检验统计量为:$t=\frac{\overline{X}-\overline{Y}}{S_w\sqrt{\frac{1}{m}+\frac{1}{n}}}$,$S_w^2=\frac{(m-1)S_A^2+(n-1)S_B^2}{(m+n-2)}$, 其中$\overline{X},\overline{Y}$ 分别是两个样本的均值,S_A^2,S_B^2 分别是两个样本的方差.显著性水平为α时的检验规则为:
当$t\leqslant -t_\alpha(m+n-2)$时拒绝$H_0$.
由样本数据计算得:$\overline{X}=16,\overline{Y}=19,S_A^2=5.6,S_B^2=4,t=-2.3717$, 而$\alpha=0.05$时,$-t_{0.05}(10)=-1.812$, 因此拒绝$H_0$.

53. 记$X=1$表示该人年龄在65岁以上，则检验问题可表示为

$$H_0:\ \mathbb{P}(X=1)=0.1355,\mathbb{P}(X\neq 1)=0.8645.$$

令$\nu_1=57$,$\nu_2=343$则

$$k_0 = \frac{(57-400*0.1355)^2}{400*0.1355} + \frac{(343-400*0.8645)^2}{400*0.8645} = 0.1673,$$

拟合优度

$$\mathbb{P}(\chi_1^2(0.05)) = 3.841$$

故我们认为原假设成立，即该市现在老年人所占比例较2000年普查时无变化.

54. 检验问题为

$$H_0 \text{:吸烟与患慢性气管炎无关.}$$

在列联表检验中,$n = 1000$,$r = s = 2$,查表$\chi^2_{(r-1)(s-1)}(0.05) = 3.841$ 检验统计量为

$$K = \frac{339(43\times121-13\times162)^2}{205\times134\times56\times283} = 7.469 > 3.841$$

故我们拒绝原假设. 这表明吸烟中患慢性气管炎比例较高.

55. 检验问题为

$$H_0 \text{:施肥无效果.}$$

在列联表检验中,$n = 1000$,$r = s = 2$,查表$\chi^2_{(r-1)(s-1)}(0.01) = 6.64$ 检验统计量为

$$K = \frac{1000(53\times117-47\times783)^2}{100\times900\times164\times836} = 75.884 > 6.64$$

故我们拒绝原假设. 这表明施肥的效果显著.

56. 略

57. 略

58. 检验问题为

$$H_0 \text{:红球个数为5}$$

此处 $n = 112$,

$$np_0 = 112 \times \frac{\binom{3}{3}}{\binom{8}{3}} = 2$$
$$np_1 = 112 \times \frac{\binom{5}{1}\binom{3}{2}}{\binom{8}{3}} = 30$$
$$np_2 = 112 \times \frac{\binom{5}{2}\binom{3}{1}}{\binom{3}{1}} = 60$$
$$np_3 = 112 \times \frac{\binom{5}{3}}{\binom{3}{1}} = 20$$

拟合优度

$$s = 2\ \mathbb{P}(\chi^2_{4-2-1}(0.05)) = 3.841$$

$$K = \frac{(1-2)^2}{2} + \frac{(31-30)^2}{30} + \frac{(55-60)^2}{60} + \frac{(25-20)^2}{20} = 2.2 < 3.841$$

即接受 H_0,红球的个数为5.

59. 检验问题为

$$H_0:\text{各工厂的质量一致}$$

此例中$r = 3, s = 3$,而$(r-1)(s-1) = 4$,查表得$\chi^2_4(0.05) = 9.488$,计算检验统计量

$$\hat{K}^\star = n(\sum_{i=1}^r \sum_{j=1}^s \frac{n_{ij}^2}{n_{i.}n_{.j}} - 1) = 15.41 > 9.488$$

因此否定 H_0,即认为各工厂质量不一样，通过比较各工厂各品质产品百分比，我们可以认为甲工厂质量最高，丙工厂质量最低.

60. H_0:A和B两种蜗牛的分布在3种珊瑚礁中都是一样的,H_1:A 和B两种蜗牛的分布在3种珊瑚礁中是有差异的,假设显著性水平$\alpha = 0.05$,计算卡方值公式为$\chi^2 = 61 \times (\frac{6^2}{28\times13} + \frac{8^2}{28\times29} + \frac{14^2}{28\times19} + \frac{7^2}{33\times13} + \frac{21^2}{33\times29} + \frac{5^2}{33\times19} - 1) = 9.8238$,
而$\chi^2_{0.05}(2) = 5.99$,所以,说明A和B两种蜗牛的分布在3 种珊瑚礁中是有差异的.

61. H_0:产品合格率与班次无关,H_1:产品合格率与班次有关,假设显著性水平$\alpha = 0.05$,计算卡方值公式为$\chi^2 = \frac{323\times(|232\times18-19\times54|-323/2)^2}{(232+19)\times(54+18)\times(232+18)\times(19+54)} = 15.0846$,
而$\chi^2_{0.05}(1) = 3.841$,所以,产品合格率与班次有关.

62. H_0:男性和女性对这三种类型的啤酒的偏好无显著差异,H_1:男性和女性对这三种类型的啤酒的偏好有显著差异,假设显著性水平$\alpha = 0.05$, 计算卡方值公式为$\chi^2 = 300\times(\frac{49^2}{180\times100} + \frac{31^2}{180\times51} + \frac{100^2}{180\times149} + \frac{51^2}{120\times100} + \frac{20^2}{120\times51} + \frac{49^2}{120\times149} - 1) = 8.1968$,
而$\chi^2_{0.05}(2) = 5.991$,所以,男性和女性对这三种类型的啤酒的偏好有显著差异.

63. 检验问题为

$$H_0 : \text{每页上印刷错误的个数服从泊松分布}$$

将数据重新分组，使得每组的错页个数不少于5，见下表 Possion分布中参数λ的MLE为

错误的个数f_i	0	1	2	≥ 3
含f_i个错误的页数	86	40	19	5

$$\hat{\lambda} = \overline{X} = \frac{100}{150} = \frac{2}{3}$$

检验问题可视为

$$H_0 : r.v.X \sim \text{Poisson 分布 } \mathbb{P}(\tfrac{2}{3})$$

此处$n = 150$,

$$\begin{aligned} np_0 &= 150 \times \frac{(\frac{2}{3})^0 e^{-\frac{2}{3}}}{0!} = 77.01 \\ np_1 &= 150 \times \frac{(\frac{2}{3})^1 e^{-\frac{2}{3}}}{1!} = 51.34 \\ np_2 &= 150 \times \frac{(\frac{2}{3})^2 e^{-\frac{2}{3}}}{2!} = 17.11 \\ np_3 &= 4.53 \end{aligned}$$

从而有

$$K_n^\star = \frac{(86-77.01)^2}{77.01} + \frac{(40-51.34)^2}{51.34} + \frac{(19-17.11)^2}{17.11} + \frac{(5-4.53)^2}{4.53} = 3.812$$

查表得$\chi_2^2(0.05) = 5.991$,即我们认为原假设成立，符合泊松分布.

64. 由题意，检验问题为

$$H_0 : r.v.X \sim N(60, 15^2)$$

我们先计算$r.v.X$在每个区间中的概率. 区间的个数$r = 8$,$\hat{p}_i = \mathbb{P}(a_{i-1} \leq X \leq a_i) = \Phi(u_i) - \Phi(u_{i-1}), i = 1, \cdots, 8,$ 其中$u_i = (a_i - 60)/15, a_0 = -\infty, a_8 = +\infty$,故有

$$\hat{p_1} = \mathbb{P}(-\infty < X < 30) = \Phi(-2) = 0.023$$
$$\hat{p_2} = \mathbb{P}(30 < X < 40) = \Phi(-\tfrac{4}{3}) - \Phi(-2) = 0.068$$

以此类推得$\hat{p_3} = 0.161, \hat{p_4} = 0.248, \hat{p_5} = 0.248, \hat{p_6} = 0.161, \hat{p_7} = 0.068, \hat{p_8} = 0.023$, 此时我们计算$K_n^\star = 6.557$, 查表得$\chi^2_{8-2-1}(0.05) = 11.071$,故我们认为原假设成立，符合正态分布.

65. 假定甲厂电视机的寿命服从$N(\mu_A, \sigma_A^2)$,乙厂电视机的寿命服从$N(\mu_B, \sigma_B^2)$.现将两个总体中分别抽取了容量为$m = 10, n = 10$的两个样本.欲检验$H_0 : \sigma_A^2 = \sigma_B^2$ $v.s.$ $H_1 : \sigma_A^2 \neq \sigma_B^2$.检验统计量为:$F = S_A^2/S_B^2$,其中$S_A^2, S_B^2$分别是两个样本的方差.显著性水平为$\alpha$时的检验规则为:
当$F \geqslant F_{\alpha/2}(m-1, n-1)$或$F \leqslant F_{1-\alpha/2}(m-1, n-1)$时拒绝$H_0$.
由样本数据计算得:$S_A^2 = 3.8333, S_B^2 = 4.9889, F = 0.7684$, 而$\alpha = 0.05$ 时,$F_{0.025}(9,9) = 4.03, F_{0.975}(9,9) = \frac{1}{F_{0.025}(9,9)} = \frac{1}{4.03} = 0.2481$, 因此无法拒绝$H_0$.
接下来,欲检验$H_0 : \mu_A = \mu_B$ $v.s.$ $H_1 : \mu_A \neq \mu_B$.因为由上述检验,我们可假定$\sigma_A^2 = \sigma_B^2$,所以,检验统计量为:$t = \frac{\overline{X} - \overline{Y}}{S_w\sqrt{\frac{1}{m} + \frac{1}{n}}}, S_w^2 = \frac{(m-1)S_A^2 + (n-1)S_B^2}{(m+n-2)}$, 其中$\overline{X}, \overline{Y}$分别是两个样本的均值,$S_A^2, S_B^2$ 分别是两个样本的方差.显著性水平为α时的检验规则为:
当$t \geqslant t_{\alpha/2}(m+n-2)$或$t \leqslant -t_{\alpha/2}(m+n-2)$时拒绝$H_0$.
由样本数据计算得:$\overline{X} = 8.5, \overline{Y} = 7.1, S_A^2 = 3.8333, S_B^2 = 4.9889, t = 1.4905$, 而$\alpha = 0.05$ 时,$t_{0.025}(18) = 2.101, -t_{0.025}(18) = -2.101$, 因此不能拒绝$H_0$.

66. H_0:星座与是否当美国总统无关,H_1:星座与是否当美国总统有关,假设显著性水平$\alpha = 0.05$, 若假设星座与是否当美国总统无关,可计算出每个星座的理想频数为3.7,则计算卡方值公式为$\chi^2 = \frac{(5-3.7)^2\times2+(3-3.7)^2\times2+(2-3.7)^2\times2+(4-3.7)^2\times6}{3.7} = 2.8865$,
而$\chi^2_{0.05}(11) = 24.725$,所以,星座与是否当美国总统无关.

67. 略

68. H_0:两家工厂的产品质量等级相同,H_1:两家工厂的产品质量等级不同,计算卡方值公式为$\chi^2 = 210 \times (\frac{58^2}{110\times98} + \frac{40^2}{110\times80} + \frac{12^2}{110\times32} + \frac{40^2}{100\times98} + \frac{40^2}{100\times80} + \frac{20^2}{100\times32} - 1) = 4.8409$,
假设显著性水平$\alpha = 0.1$,而$\chi^2_{0.1}(2) = 4.61$, 所以,两家工厂的产品质量等级不同.
假设显著性水平$\alpha = 0.05$,而$\chi^2_{0.05}(2) = 5.991$, 所以,两家工厂的产品质量等级相同.

69. 略

70. 略

71. 由题意, 总体 $X \sim N(\mu_1, \sigma^2)$, $Y \sim N(\mu_2, \sigma^2)$, σ^2 未知, 且两样本相互独立. 检验问题

$$H_0:\ \mu_1 = \mu_2,\ \ H_1:\ \mu_1 \neq \mu_2.$$

检验的拒绝域为

$$|T| = \left|\frac{\bar{X} - \bar{Y}}{S_w\sqrt{1/m + 1/n}}\right| > t_{m+n-2}(\alpha/2),$$

其中, $m = 8$, $n = 10$, $\bar{X} = 0.2319$, $\bar{Y} = 0.2097$, $S_1^2 = 0.0146^2$, $S_2^2 = 0.0097^2$, $S_w^2 = \frac{(8-1)S_1^2 + (10-1)S_2^2}{8+10-2} = 0.012^2$, $t_{16}(0.025) = 2.1199$. 由于

$$|T| = \left|\frac{0.2319 - 0.2097}{0.012\sqrt{1/8 + 1/10}}\right| = 3.900 > 2.1199,$$

故否定 H_0, 即认为两位作家所写的小品文中包含有 3 个字母组成的单字的比例有显著的差异.

72. (1) 由题意 $m = 13$, $n = 10$, $\alpha = 0.1$, $F_{12,9}(0.05) = 3.07$, $F_{12,9}(0.95) = 1/F_{9,12}(0.05) = 1/2.80 = 0.357$, $S_1^2 = 0.034$, $S_2^2 = 0.0264$. 由于

$$F_{12,9}(0.95) < \frac{S_1^2}{S_2^2} = 1.288 < F_{12,9}(0.05),$$

故接受 H_0, 即认为两总体方差相等.

(2) 由 (1) 可认为 $\sigma_1^2 = \sigma_2^2$, 接着检验 $H_0':\ \mu_1 = \mu_2$, $H_1':\ \mu_1 \neq \mu_2$. 计算可得 $\bar{X} = 1.752$, $\bar{Y} = 2.507$, $S_w^2 = \frac{12\times 0.034 + 9\times 0.0264}{13+10-2} = 0.03075$. 由于

$$|T| = \left|\frac{1.752 - 2.507}{\sqrt{0.03075}\sqrt{1/12 + 1/9}}\right| = 9.764 > t_{13+10-2}(0.1/2) = t_{21}(0.05) = 1.7207,$$

故拒绝 H_0', 认为杂志上刊载的论文与未出版的学术报告的可理解性有显著差异.

73. (1) 将区间 $[20, 100]$ 等分为 8个小区间, 每个小区间的长度为 $\Delta = 10$. 记落在每个小区间内的数据频数, 在第 i 个区间上以 Δ 为底, 以 f_i 为高作小长方形, 得直方图如图 9.1:

(2) 在显著水平 $\alpha = 0.1$ 下检验假设

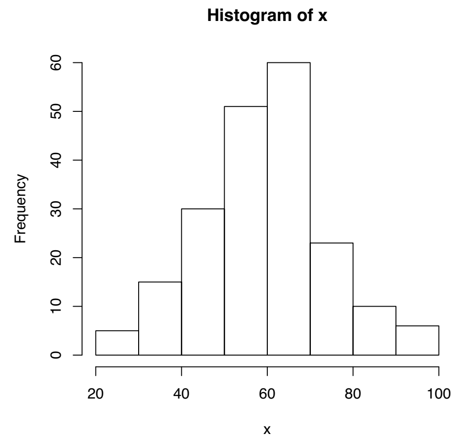

图 9.1: “某大学一年级学生(200名)一次数学考试的成绩”

$$H_0: \text{数据 } X \text{ 来自正态总体 } X \sim N(60, 15^2),$$

若 H_0为真, 则有

$$p_1 = \mathbb{P}(-\infty < X \le 30) = \Phi(\frac{30-60}{15}) = 1 - \Phi(2) = 0.0228,$$

同理可得, $p_2 = \mathbb{P}(30 < X \le 40) = 0.0690$, $p_3 = \mathbb{P}(40 < X \le 50) = 0.1596$, $p_4 = \mathbb{P}(50 < X \le 60) = 0.2486$, $p_5 = \mathbb{P}(60 < X \le 70) = 0.2486$, $p_6 = \mathbb{P}(70 < X \le 80) = 0.1596$, $p_7 = \mathbb{P}(80 < X \le 90) = 0.0690$, $p_8 = \mathbb{P}(90 < X < \infty) = 0.0228$.

因此, $\chi^2 = \sum_{i=1}^{8} f_i^2/(np_i) - 200 = 6.4145$. 因 $\alpha = 0.1$, $k = 8$, $r = 0$, 有 $\chi^2_{k-r-1}(0.1) = \chi^2_7(0.1) = 12.017 > 6.4145$, 故在显著性水平 $\alpha = 0.1$ 下接受 H_0, 即认为考试成绩的数据来自正态总体 $N(60, 15^2)$.

方法二: 若考虑把 $\{X \le 30\}$ 设为一组, $\{X > 80\}$ 设为一组, 那么经计算可得 $\chi^2 = \sum_{i=1}^{6} \tilde{f}_i^2/(n\tilde{p}_i) - 200 = 5.2164 < \chi^2_{6-0-1}(0.1) = 9.236$, 因此可得到上述相同结论.

74. 在显著性水平 $\alpha = 0.05$ 下检验假设

$$H_0: X \text{ 服从超几何分布 } P\{X = k\} = \binom{5}{k}\binom{3}{3-k}\Big/\binom{8}{3}, \ \ k = 0, 1, 2, 3.$$

此处 $n = 112$. 若 H_0 为真, 则可按所给分布律计算, 得如下的概率

$$p_1 = P\{X = 0\} = \binom{5}{0}\binom{3}{3}\Big/\binom{8}{3} = 1/56, \quad p_2 = P\{X = 1\} = \binom{5}{1}\binom{3}{2}\Big/\binom{8}{3} = 15/56,$$
$$p_3 = P\{X = 2\} = \binom{5}{2}\binom{3}{1}\Big/\binom{8}{3} = 30/56, \quad p_4 = P\{X = 3\} = \binom{5}{3}\binom{3}{0}\Big/\binom{8}{3} = 10/56.$$

因此 $\chi^2 = \sum_{i=1}^4 f_i^2/(np_i) - 112 = 114.2 - 112 = 2.2 < \chi_3^2(0.05) = 7.815$, 故在显著性水平 $\alpha = 0.05$ 下接受 H_0, 即认为 X 服从超几何分布 $P\{X=k\} = \binom{5}{k}\binom{3}{3-k}/\binom{8}{3}$, $k = 0,1,2,3$.

方法二: 若把 $\{X=0\}$ 和 $\{X=1\}$ 设为一组, 那么可以得到$\chi^2 = \sum_{i=1}^3 \tilde{f}_i^2/(n\tilde{p}_i) - 112 = 1.6667 < \chi_2^2(0.05) = 5.991$, 因此可得到上述相同结论.

75. 以 X 记鱼种类的序号, 按题意需检验假设

$$H_0 : \mathbb{P}(X=1) = 0.20, \mathbb{P}(X=2) = 0.15, \mathbb{P}(X=3) = 0.40, \mathbb{P}(X=4) = 0.25.$$

由题意 $n = 600$, $f_1 = 132$, $f_2 = 100$, $f_3 = 200$, $f_4 = 168$, 因此, $\chi^2 = \sum_{i=1}^4 f_i^2/(np_i) - 600 = 11.1378 > \chi_3^2(0.05) = 7.815$. 故拒绝 H_0, 认为各鱼类数量之比相对于 10 年期有显著的改变.

第九章 回归分析

1. (1) 散点图如图9.2:

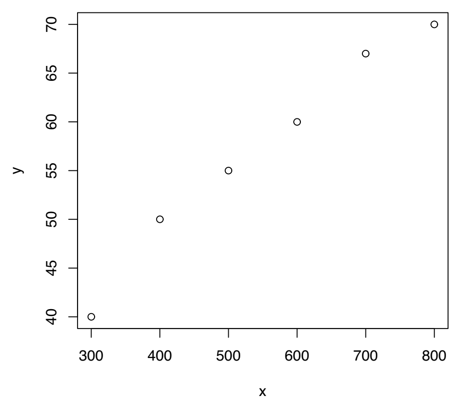

图 9.2: "退火温度x(° C)对黄铜延性Y的效应"

(2) 由上图可以看出取回归函数为 x 的线性函数 $a+bx$ 是合适的. 经计算得到

$$S_{xx}=\sum x_i^2-\frac{1}{n}(\sum x_i)^2=1990000-\frac{1}{6}\times 3300^2=175000,$$

$$S_{xy}=\sum x_iy_i-\frac{1}{n}(\sum x_i)(\sum y_i)=198400-\frac{1}{6}\times 3300\times 342=10300,$$

从而

$$\hat{b}=\frac{S_{xy}}{S_{xx}}=\frac{10300}{175000}=0.058857,$$

$$\hat{a}=\bar{y}-\hat{b}\bar{x}=\frac{1}{6}\times 342-0.058857\times\frac{1}{6}\times 3300=24.62857.$$

则回归方程为

$$\hat{y}=24.62857+0.058857x.$$

2. 经计算得到

$$S_{xx}=\sum x_i^2-\frac{1}{n}(\sum x_i)^2=4200.56-\frac{1}{15}\times 249.8^2=40.55733,$$

$$S_{xy}=\sum x_iy_i-\frac{1}{n}(\sum x_i)(\sum y_i)=6740.71-\frac{1}{15}\times 249.8\times 400.3=74.38067,$$

因此

$$\hat{b}=\frac{S_{xy}}{S_{xx}}=1.833963,$$

$$\hat{a}=\bar{y}-\hat{b}\bar{x}=\frac{1}{15}\times 400.3-1.833963\times\frac{1}{15}\times 249.8=-3.854938,$$

故回归方程为

$$\hat{y}=-3.854938+1.833963x.$$

3. 经计算得到

$$S_{xx}=\sum x_i^2-\frac{1}{n}(\sum x_i)^2=1332.8125-\frac{1}{15}\times 141.25^2=2.708333,$$

$$S_{yy}=\sum y_i^2-\frac{1}{n}(\sum y_i)^2=729.625-\frac{1}{15}\times 104.5^2=1.608333,$$

$$S_{xy}=\sum x_iy_i-\frac{1}{n}(\sum x_i)(\sum y_i)=985.5-\frac{1}{15}\times 141.25\times 104.5=1.458333,$$

(1)

$$\hat{b}=\frac{S_{xy}}{S_{xx}}=0.53846,$$

$$\hat{a}=\bar{y}-\hat{b}\bar{x}=\frac{1}{15}\times 104.5-0.53846\times\frac{1}{15}\times 141.25=1.896,$$

故回归方程为

$$\hat{y}=1.896+0.53846x.$$

(2) 先计算 $Q_e=S_{yy}-\hat{b}S_{xy}=1.608333-0.53846\times 1.458333=0.823079$, 因 $n=15$, 故 $\hat{\sigma^2}=Q_e/(n-2)=0.06331$. 查表可得 $t_{13}(0.025)=2.1604$. 所以 b 的置信水平为 0.95 的置信区间为

$$(\hat{b}\pm t_{n-2}(\alpha/2)\frac{\hat{\sigma}}{\sqrt{S_{xx}}})=(0.53846\pm 2.1604\times\frac{\sqrt{0.06331}}{\sqrt{2.708333}})=(0.208,0.869).$$

4. 略

5. 略

6. 略

7. 略

8. 略

9. 略

10. 略

11. 略

12. 略

13. (1) 散点图如图9.3:

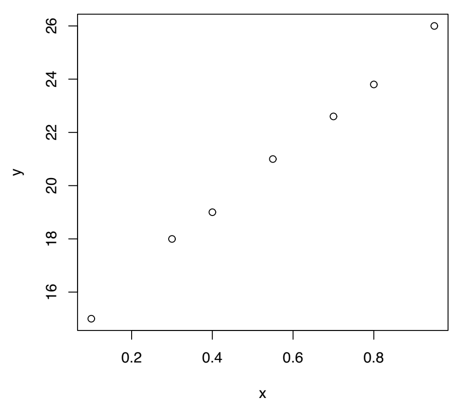

图 9.3: "钢线碳 x 含量对于电阻 y 的效应"

(2) 由题意 $n=7$, 经计算得 $\sum x_i=3.8$, $\sum y_i=145.4$, $\sum x_i^2=2.595$, $y_i^2=3104.2$, $\sum x_iy_i=85.61$,

$$S_{xx}=\sum x_i^2-\frac{1}{n}(\sum x_i)^2=0.5321429,$$

$$S_{yy}=\sum y_i^2-\frac{1}{n}(\sum y_i)^2=84.03429,$$

$$S_{xy}=\sum x_iy_i-\frac{1}{n}(\sum x_i)(\sum y_i)=6.678571.$$

因此

$$\hat{b}=\frac{S_{xy}}{S_{xx}}=12.55034,$$

$$\hat{a}=\bar{y}-\hat{b}\bar{x}=13.95839,$$

故回归方程为

$$\hat{y} = 13.95839 + 12.55034x.$$

(3) $\hat{\sigma^2} = Q_e/(n-2) = (S_{yy} - \hat{b}S_{xy})/5 = 0.04319463.$

(4) 检验统计量 T 使得

$$|T| = |\frac{\hat{b}}{\sqrt{\hat{\sigma^2}}}\sqrt{S_{xx}}| = 44.0509 \gg t_{n-2}(\alpha/2) = t_5(0.005) = 4.0322,$$

由此可知回归效果是极其显著的.

(5) b 的置信水平为 0.95 (即 $\alpha = 0.05$)的置信区间:

$$\left(\hat{b} \pm t_{n-2}(\alpha/2)\frac{\hat{\sigma}}{\sqrt{S_{xx}}}\right) = \left(12.5503 \pm 2.5706 \times \frac{\sqrt{0.0432}}{\sqrt{0.5321}}\right) = (11.81785, 13.28275).$$

(6) $x_0 = 0.50$ 处对应的 Y 对应的估计值为

$$\hat{y_0} = 13.9584 + 12.5503 \times 0.50 = 20.23355.$$

因此$x = 0.50$处$\mu(x)$ 的置信水平为0.95的置信区间为

$$\left(\hat{y_0} \pm t_{n-2}(\alpha/2)\hat{\sigma}\sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}\right) = (20.23355 \pm 0.2044) = (20.03, 20.44).$$

(7) $x = 0.50$处观察值Y的置信水平为0.95的置信区间为

$$\left(\hat{y_0} \pm t_{n-2}(\alpha/2)\hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}\right) = (20.23355 \pm 0.5720) = (19.66, 20.81).$$

14. (1) 在图 9.4 表示.

(2) 在图 9.5 表示.

(3) 将 $Y = a\exp(bx)\varepsilon$ 取对数, 得

$$\ln Y = \ln a + bx + \ln\varepsilon.$$

令 $Z = \ln Y$, 则回归模型为

$$Z = \ln a + bx + \ln\varepsilon,$$

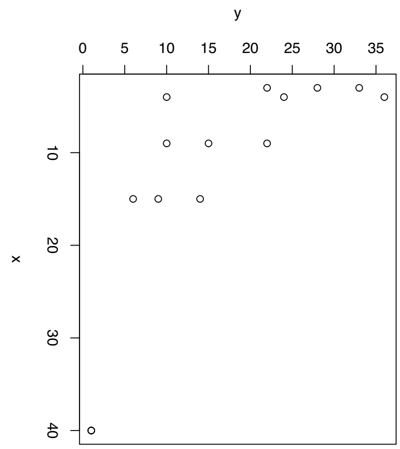

图 9.4: “大树的年龄 x 和槲寄生的株数 y 的关系”

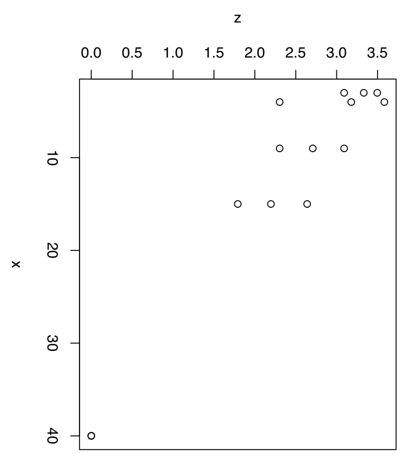

图 9.5: “大树的年龄 x 和槲寄生的株数 y 的函数 $\ln y$ 的关系”

其中 $\varepsilon \sim N(0,\sigma^2)$.

由题意 $n=14$, 经计算 $\sum x_i = 173$, $\sum x_i^2 = 4193$, $\sum z_i = 33.71363$, $\sum x_i z_i = 238.3516$, $S_{xx} = 2055.214$, $S_{xz} = -178.2525$, 于是得到 $\hat{b} = -0.08673185$, $\widehat{\ln a} = 3.479874$, $\exp(\widehat{\ln a}) = 32.45565$, 则曲线回归方程为

$$\hat{y} = 32.45565 e^{-0.08673x}.$$

15. (1) 散点图如图 9.6

(2) 令 $x_1 = x$, $x_2 = x^2$, 则题中假设的模型可写成

$$Y = b_0 + b_1 x_1 + b_2 x_2 + \varepsilon, \;\; \varepsilon \sim N(0,\sigma^2).$$

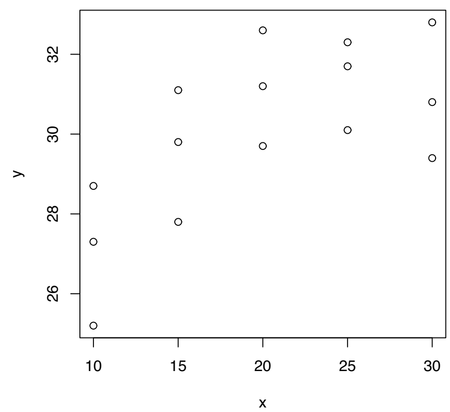

图 9.6: “一种合金在不同浓度的某种添加剂下的抗压强度”

记

$$X=\begin{bmatrix}1&10&100\\1&10&100\\1&10&100\\1&15&225\\1&15&225\\1&15&225\\1&20&400\\1&20&400\\1&20&400\\1&25&625\\1&25&625\\1&25&625\\1&30&900\\1&30&900\\1&30&900\end{bmatrix},\quad Y=\begin{bmatrix}25.2\\27.3\\28.7\\29.8\\31.1\\27.8\\31.2\\32.6\\29.7\\31.7\\30.1\\32.3\\29.4\\30.8\\32.8\end{bmatrix},\quad B=\begin{bmatrix}b_0\\b_1\\b_2\end{bmatrix}.$$

可以得到系数的估计为

$$\hat{B}=\begin{bmatrix}\hat{b}_0\\\hat{b}_1\\\hat{b}_2\end{bmatrix}=(X^\top X)^{-1}X^\top Y=\begin{bmatrix}19.03333\\1.00857\\-0.02038\end{bmatrix}.$$

故回归方程为

$$\hat{y} = 19.03333 + 1.00857x_1 - 0.02038x_2,$$

即

$$\hat{y} = 19.03333 + 1.00857x - 0.02038x^2.$$

16. (1) 记

$$X = \begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} 7.6 \\ 10.3 \\ 9.2 \\ 10.2 \\ 8.4 \\ 11.1 \\ 9.8 \\ 12.6 \end{bmatrix}, \quad B = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix},$$

所要求的线性回归模型为

$$Y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \varepsilon, \ \ \varepsilon \sim N(0, \sigma^2),$$

于是得到系数的估计

$$\hat{B} = \begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_3 \end{bmatrix} = (X^\top X)^{-1}X^\top Y = \begin{bmatrix} 9.9 \\ 0.575 \\ 0.55 \\ 1.15 \end{bmatrix}.$$

故 Y 的多元回归方程为

$$\hat{y} = 9.9 + 0.575x_1 + 0.55x_2 + 1.15x_3.$$

(2) 根据题意, 新模型

$$Y = b_0 + b_1x_1 + b_3x_3 + \varepsilon, \ \ \varepsilon \sim N(0, \sigma^2),$$

记 M 为 X 删去第三列后所得的矩阵, 新的系数向量为 $\beta = [\beta_0, \beta_1, \beta_3]^\top$, 于是得到系数的估计

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} 9.9 \\ 0.575 \\ 1.15 \end{bmatrix},$$

故多元回归方程为

$$\hat{y} = 9.9 + 0.575x_1 + 1.15x_3.$$