概率论与数理统计题目集

河北工程大学 信息与电气工程学院

计算机 2501 姜振坤

2025 年 12 月 3 日

使用说明[1]

目录

[1] 题目均来自民间数据, 答案仅供参考; 题目集已经同步开源到 GitHub:仓库地址

1 2024~2025 学年概率论计算机专业期末试卷 A 卷

1.1 选择题 (每小题 3 分，共 15 分)

题目 1

若事件 A 和 B 满足 $P(A\cup B)=P(A)+P(B)$，则 ( ).

(A) A 和 B 互不相容 (互斥)

(B) A 和 B 是对立事件

(C) AB 是一个小概率事件

(D) $P(A)=0$ 或 $P(B)=0$

解答 由概率的加法公式

$$P(A\cup B)=P(A)+P(B)-P(A\cap B),$$

故有

$$P(A)+P(B)-P(A\cap B)=P(A)+P(B) \implies P(A\cap B)=0.$$

因此 A 与 B 互不相容 (互斥).

答案:$\boxed{\text{A}}$.

题目 2

设随机变量 $X\sim N(\mu,9)$，$Y\sim N(\mu,4)$，记 $p_1=P(X\leq \mu-3)$，$p_2=P(X\geq \mu+4)$，则 ( ).

(A) $p_1=p_2$ (B) $p_1>p_2$ (C) $p_1<p_2$ (D) 以上均有可能

解答 对 $X\sim N(\mu,9)$(标准差为 3) 有

$$p_1=P(X\leq \mu-3)=P\left(\frac{X-\mu}{3}\leq -1\right)=\Phi(-1),$$

$$p_2=P(X\geq \mu+4)=P\left(\frac{X-\mu}{3}\geq \frac{4}{3}\right)=1-\Phi\left(\frac{4}{3}\right).$$

数值比较:$\Phi(-1)\approx 0.1587$，$1-\Phi(4/3)\approx 0.0918$，故 $p_1>p_2$.

答案:$\boxed{\text{B}}$.

题目 3

设随机变量 X 和 Y 独立同分布，且 X 的分布函数为 $F(x)$，则在 $Z=\max\{X,Y\}$ 分布函数为 ( ).

(A) $F^2(z)$

(B) $F(x)\cdot F(y)$

(C) $1-[1-F(z)]^2$

(D) $[1-F(x)][1-F(y)]$

解答 X,Y 独立同分布，分布函数为 $F(x)$.

$$F_Z(z)=P(\max\{X,Y\}\le z)=P(X\le z,\ Y\le z)=F(z)^2.$$

答案:$\boxed{\text{A}}$.

题目 4

设随机变量 X 和 Y 具有相同期望和方差，且相关系数 $\rho_{XY}=0$. 记 $Z_1=2X+Y$，$Z_2=2X-Y$，则 Z_1，Z_2 的相关系数为 ( ).

(A) 0

(B) $\frac{3}{5}$

(C) $\frac{3}{\sqrt{15}}$

(D) $\frac{3}{\sqrt{10}}$

解答 已知 $\mathrm{Var}(X)=\mathrm{Var}(Y)=\sigma^2$, $\mathrm{Cov}(X,Y)=0$.

$$Z_1=2X+Y,\qquad Z_2=2X-Y.$$

$$\mathrm{Cov}(Z_1,Z_2)=4\sigma^2-\sigma^2=3\sigma^2.$$

$$\mathrm{Var}(Z_1)=5\sigma^2,\qquad \mathrm{Var}(Z_2)=5\sigma^2.$$

相关系数:

$$\rho_{Z_1,Z_2}=\frac{3\sigma^2}{\sqrt{(5\sigma^2)(5\sigma^2)}}=\frac{3}{5}.$$

答案:$\boxed{\text{B}}$.

题目 5

设 $X_1,X_2,\cdots,X_n$ 是为独立同分布随机变量序列，且 $X_1\sim B(1,p)$，记 $\phi(x)$ 为标准正态分布函数，则 ( ).

A. $\lim_{n\to\infty} P\left\{\frac{\sum_{i=1}^{n} X_i - np}{n\sqrt{p}} \le x\right\} = \phi(x)$

B. $\lim_{n\to\infty} P\left\{\frac{\sum_{i=1}^{n} X_i - np}{(1-p)\sqrt{n}} \le x\right\} = \phi(x)$

C. $\lim_{n\to\infty} P\left\{\frac{\sum_{i=1}^{n} X_i - np}{\sqrt{np(1-p)}} \le x\right\} = \phi(x)$

D. $\lim_{n\to\infty} P\left\{\frac{\sum_{i=1}^{n} X_i - np}{n(1-p)p} \le x\right\} = \phi(x)$

解答 设 $X_i \sim B(1,p)$，令 $S_n = \sum_{i=1}^{n} X_i$.

$$E(S_n) = np, \qquad \mathrm{Var}(S_n) = np(1-p).$$

由中心极限定理，

$$\frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0,1),$$

从而

$$\lim_{n\to\infty} P\left\{\frac{S_n - np}{\sqrt{np(1-p)}} \le x\right\} = \phi(x).$$

答案:$\boxed{\text{C}}$.

1.2 填空题 (每小题 3 分，共 15 分)

题目 1

已知 $P(A\cup B) = 0.7$，$P(A) = 0.4$，当 A 与 B 不相容时，$P(B) =$ ________. 若 A 与 B 相互独立，则 $P(B) =$ ________

解答 已知 $P(A\cup B) = 0.7$, $P(A) = 0.4$.

1. 当 A 与 B 不相容 (互斥) 时，

$$P(A\cup B) = P(A) + P(B) \Rightarrow P(B) = 0.7 - 0.4 = 0.3.$$

2. 当 A 与 B 相互独立时，

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

代入数值得

$$0.7=0.4+P(B)-0.4P(B)=0.4+P(B)(1-0.4)=0.4+0.6P(B),$$

因此

$$0.6P(B)=0.3\Rightarrow P(B)=\frac{0.3}{0.6}=0.5.$$

题目 2

已知随机变量 X 的概率分布为

X	1	2	3
$P\{X=k\}$	$1/4$	$2\theta(1-\theta)$	$(1-\theta)^2$

则 $\theta=$ ________.

解答 已知分布

X	1	2	3
$P\{X=k\}$	$1/4$	$2\theta(1-\theta)$	$(1-\theta)^2$

概率之和为 1，于是

$$\frac{1}{4}+2\theta(1-\theta)+(1-\theta)^2=1.$$

展开并化简:

$$\frac{1}{4}+2\theta-2\theta^2+1-2\theta+\theta^2=1$$

$$\Rightarrow\frac{5}{4}-\theta^2=1\Rightarrow\theta^2=\frac{1}{4}.$$

由于 θ 为概率参数，取非负根:

$$\theta=\frac{1}{2}.$$

题目 3

已知随机变量 $X\sim N(2,\sigma^2)$，且 $P\{2<X<4\}=0.3$，则 $P\{X<0\}=$ ____________.

解答 已知 $X \sim N(2, \sigma^2)$，且

$$P\{2 < X < 4\} = 0.3.$$

令标准正态变量 $Z = \dfrac{X-2}{\sigma}$，则

$$P\{0 < Z < \frac{4-2}{\sigma}\} = 0.3.$$

记 $a = \dfrac{2}{\sigma}$，有

$$\Phi(a) - \Phi(0) = 0.3 \Rightarrow \Phi(a) - \frac{1}{2} = 0.3 \Rightarrow \Phi(a) = 0.8.$$

因此 $a = \Phi^{-1}(0.8)$(约为 0.8416)，故

$$\frac{2}{\sigma} = \Phi^{-1}(0.8) \quad (\text{用于中间计算}).$$

所求为

$$P\{X < 0\} = P\left\{Z < \frac{0-2}{\sigma}\right\} = P\{Z < -a\} = \Phi(-a).$$

利用标准正态对称性

$$\Phi(-a) = 1 - \Phi(a) = 1 - 0.8 = 0.2.$$

所以 $P\{X < 0\} = 0.2$.

题目 4

设两个相互独立的随机变量 X 和 Y 均服从二项分布 $B(1, \dfrac{1}{2})$，则 $P\{X \geqslant Y\} =$ ______________.

解答 设 X, Y 独立且服从 $B(1, \dfrac{1}{2})$(即各自取值为 0 或 1，且 $P(1) = \dfrac{1}{2}$). 列举四种等概率情况:

$$(0,0), (0,1), (1,0), (1,1),$$

各概率均为 $\dfrac{1}{4}$. 满足 $X \geq Y$ 的情形为 $(0,0), (1,0), (1,1)$，共三种，所以

$$P\{X \geq Y\} = \frac{3}{4}.$$

题目 5

设 $X_1, X_2, \cdots, X_n$ 是取自总体 X 的简单随机样本，已知 $E(X)=\mu$, $D(X)=\sigma^2$，则 $E(\bar{X})=$ ________, $D(\bar{X})=$ ________.

解答 若 $X_1, \dots, X_n$ 为来自总体 X 的简单随机样本，且 $E(X)=\mu$, $D(X)=\sigma^2$，样本均值 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n} X_i$. 由线性算子性质与独立性得

$$E(\bar{X})=\frac{1}{n}\sum_{i=1}^{n} E(X_i)=\frac{1}{n}\cdot n\mu=\mu,$$

$$D(\bar{X})=\frac{1}{n^2}\sum_{i=1}^{n} D(X_i)=\frac{1}{n^2}\cdot n\sigma^2=\frac{\sigma^2}{n}.$$

1.3 计算题 (共 70 分)

题目 1

学生连续参加同一课程的两次考试. 第一次及格的概率为 p. 若第一次及格，则第二次及格的概率也为 p. 若第一次不及格，则第二次及格的概率为 $\frac{p}{2}$.

(1) 若至少有一次及格他才能取得某种资格，求他取得该资格的概率.

(2) 若已知他第二次已经及格，求他第一次及格的概率.

解答 学生两次考试的关系: 第一次及格的概率为 p. 若第一次及格，则第二次及格的条件概率为 p; 若第一次不及格，则第二次及格的条件概率为 $\frac{p}{2}$.

设事件 A 表示“第一次及格”，B 表示“第二次及格”. 则

$$P(A)=p, \qquad P(B \mid A)=p, \qquad P(B \mid A^c)=\frac{p}{2}.$$

(1) 至少有一次及格的概率

$$P(A\cup B)=1-P(A^c\cap B^c)=1-P(A^c)P(B^c \mid A^c).$$

先计算 $P(B)$ 或直接用加法公式:

$$P(A\cup B)=P(A)+P(B)-P(A\cap B).$$

其中

$$P(A\cap B)=P(A)P(B\mid A)=p\cdot p=p^2,$$

且

$$P(B)=P(A)P(B\mid A)+P(A^c)P(B\mid A^c)=p\cdot p+(1-p)\cdot\frac{p}{2}=p^2+\frac{p(1-p)}{2}.$$

因此

$$\begin{aligned}P(A\cup B)&=P(A)+P(B)-P(A\cap B)\\&=p+\left(p^2+\frac{p(1-p)}{2}\right)-p^2\\&=p+\frac{p(1-p)}{2}\\&=p\left(1+\frac{1-p}{2}\right)=p\left(\frac{1+1-p}{2}\right)=p\left(\frac{1+p}{2}\right).\end{aligned}$$

化简得到

$$\boxed{P(\text{至少一次及格})=\frac{p(1+p)}{2}}.$$

(2) 已知第二次及格，求第一次及格的条件概率 $P(A\mid B)$. 由贝叶斯公式:

$$P(A\mid B)=\frac{P(A\cap B)}{P(B)}=\frac{p^2}{p^2+\dfrac{p(1-p)}{2}}.$$

对分母提取 p:

$$P(A\mid B)=\frac{p^2}{p\left(p+\dfrac{1-p}{2}\right)}=\frac{p}{p+\dfrac{1-p}{2}}=\frac{p}{\dfrac{p+1-p}{2}+\dfrac{p}{2}}.$$

更直接化简:

$$P(A\mid B)=\frac{p^2}{p^2+\dfrac{p(1-p)}{2}}=\frac{p}{p+\dfrac{1-p}{2}}=\frac{2p}{2p+1-p}=\frac{2p}{p+1}.$$

因此

$$\boxed{P(\text{第一次及格}\mid\text{第二次及格})=\frac{2p}{p+1}}.$$

题目 2

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

解答 袋中有编号 $1,2,3,4,5$ 的 5 个球，同时取 3 只. 令随机变量 X 为所取 3 只球中的最大号码.

所有取法等可能,样本点数为 $\binom{5}{3}=10$. 当最大号码为 k 时,另外两个球必须从 $\{1,\ldots,k-1\}$ 中取，且要取 2 个，不考虑顺序，所以

$$P(X=k)=\frac{\binom{k-1}{2}}{\binom{5}{3}},\qquad k=3,4,5.$$

具体数值:

$$\begin{aligned}P(X=3)&=\frac{\binom{2}{2}}{10}=\frac{1}{10},\\ P(X=4)&=\frac{\binom{3}{2}}{10}=\frac{3}{10},\\ P(X=5)&=\frac{\binom{4}{2}}{10}=\frac{6}{10}=\frac{3}{5}.\end{aligned}$$

写出分布律和分布函数 $D(x)$(即累积分布函数 $F_X(x)$):

概率质量函数 (分布律)

$$P(X=k)=\begin{cases}\dfrac{1}{10}, & k=3,\\ \dfrac{3}{10}, & k=4,\\ \dfrac{6}{10}, & k=5,\\ 0, & \text{其他}k.\end{cases}$$

累积分布函数 $D(x)=F_X(x)=P(X\le x)$:

$$D(x)=F_X(x)=\begin{cases}0, & x<3,\\ \dfrac{1}{10}, & 3\le x<4,\\ \dfrac{1+3}{10}=\dfrac{4}{10}, & 4\le x<5,\\ 1, & x\ge 5.\end{cases}$$

题目 3

设随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<1,\\ \ln x, & 1\le x<e,\\ 1, & x\ge e,\end{cases}$$

(1) 求 $P\{X<2\}$, $P\{0<X\le 3\}$, $P\{2<X<\frac{5}{2}\}$;

(2) 求概率密度 $f_X(x)$;

(3) 求 $Y=X^2$ 的概率密度.

解答 给出分布函数

$$F(x)=\begin{cases}0, & x<1,\\ \ln x, & 1\le x<e,\\ 1, & x\ge e,\end{cases}$$

其中 e 为自然常数.

(1) 计算概率:

$$P\{X<2\}=F(2^-)=\ln 2.$$

$$P\{0<X\le 3\}=F(3)-F(0)=1-0=1,\qquad (\text{因}3\ge e,\ F(3)=1)$$

$$P\left\{2<X<\frac{5}{2}\right\}=F\left(\frac{5}{2}\right)-F(2)=\ln\left(\frac{5}{2}\right)-\ln 2=\ln\left(\frac{5/2}{2}\right)=\ln\left(\frac{5}{4}\right).$$

(2) 概率密度 $f_X(x)$: 在 $1\le x<e$ 上 $F(x)=\ln x$ 可微，导数为 $1/x$; 其它处导数为 0(几乎处处). 因此

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

(在 $x=1$ 和 $x=e$ 点处可按几乎处处性质忽略点质量，因为 F 在这些点连续.)

(3) 设 $Y=X^2$. 由于 $X>0$ 且 $X\in[1,e)$，因此 $Y\in[1,e^2)$. 使用一变换公式: 若

$y \in (1, e^2)$，则 $x = \sqrt{y}$ 且

$$f_Y(y) = f_X(\sqrt{y}) \cdot \frac{1}{2\sqrt{y}} = \frac{1}{\sqrt{y}} \cdot \frac{1}{2\sqrt{y}} = \frac{1}{2y}.$$

在其它 y 上为 0. 所以

$$\boxed{f_Y(y) = \begin{cases} \frac{1}{2y}, & 1 < y < e^2, \\ 0, & \text{其他}y. \end{cases}}$$

题目 4

设随机变量 $X \sim U(0,1)$，当给定 $X = x$ 时，随机变量 Y 的条件概率密度为

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

(1) X 和 Y 的联合概率密度 $f(x,y)$;

(2) 边缘概率密度 $f_Y(y)$，并画出它的图形;

(3) $P\{X > Y\}$.

解答 $X \sim U(0,1)$(即 $f_X(x) = 1,\ 0 < x < 1$). 已知条件密度

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

(1) 联合密度

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

(2) 边缘密度 $f_Y(y)$: 固定 $y > 0$，对满足条件的 x 积分. 条件 $0 < y < 1/x$ 等价于 $x < 1/y$. 且还要求 $0 < x < 1$，因此积分区间为 $0 < x < \min(1,\ 1/y)$. 因此

$$f_Y(y) = \int_0^{\min(1,1/y)} x\, dx = \frac{1}{2}\big(\min(1,\ 1/y)\big)^2.$$

整理成分段形式:

$$\boxed{f_Y(y)=\begin{cases}\dfrac{1}{2}, & 0<y<1,\\ \dfrac{1}{2y^2}, & y\geq 1,\\ 0, & y\leq 0.\end{cases}}$$

图形说明: 在 $0<y<1$ 为常数 $\frac{1}{2}$; 在 $y\geq 1$ 单调衰减，按 $1/(2y^2)$ 下降，并在 $y=1$ 处连续 (两侧值均为 $1/2$).

(3) 计算 $P\{X>Y\}$:

$$P\{X>Y\}=\iint_{x>y} f_{X,Y}(x,y)\,dy\,dx.$$

更方便地，先对 x 取值:$x\in(0,1)$，对固定 x，y 的取值为 $0<y<\min(x,1/x)$. 但当 $0<x<1$ 时有 $1/x>1>x$，故 $\min(x,1/x)=x$. 于是

$$P\{X>Y\}=\int_0^1\int_0^x x\,dy\,dx=\int_0^1 x\cdot x\,dx=\int_0^1 x^2\,dx=\left[\frac{x^3}{3}\right]_0^1=\frac{1}{3}.$$

因此

$$\boxed{P\{X>Y\}=\frac{1}{3}.}$$

题目 5

设各零件的质量是都是随机变量，它们相互独立，且服从相同的分布，其数学期望为 0.5kg，均方差为 0.1kg，问 5000 个零件的总质量超过 2510kg 的概率是多少?

解答　设每个零件质量为独立同分布的随机变量，期望为 $\mu=0.5$ kg，题中给出“均方差为 0.1 kg”. 为明确起见，我们在此假定“均方差”指的是标准差 (即 $\sigma=0.1$ kg). 若题意为方差，请将下面的 σ 替换为 $\sqrt{0.1}$ 并重新计算.

设 X_i 为第 i 个零件质量，$E(X_i)=\mu=0.5$，$\mathrm{sd}(X_i)=\sigma=0.1$，且相互独立. 令总质量

$$S_n=\sum_{i=1}^n X_i,\quad n=5000.$$

则

$$E(S_n)=n\mu=5000\times 0.5=2500\ \text{kg},$$

方差

$$\mathrm{Var}(S_n) = n\sigma^2 = 5000 \times (0.1)^2 = 5000 \times 0.01 = 50,$$

标准差

$$\mathrm{sd}(S_n) = \sqrt{50} = \ 0.1\sqrt{5000} = 7.071067811 \text{ kg}.$$

利用中心极限定理，

$$P\{S_n > 2510\} \approx P\left\{Z > \frac{2510 - 2500}{\sqrt{50}}\right\} = 1 - \Phi\left(\frac{10}{\sqrt{50}}\right) = 1 - \Phi\left(\frac{10}{7.071067811}\right).$$

计算标准化值:

$$\frac{10}{7.071067811} \approx 1.41421356.$$

查标准正态表或用数值计算:

$$\Phi(1.41421356) \approx 0.92135 \quad \Rightarrow \quad P\{S_n > 2510\} \approx 1 - 0.92135 = 0.07865.$$

因此近似概率为

$$\boxed{P\{S_n > 2510\} \approx 0.0787\,}.$$

(若“均方差”原意为方差 $\mathrm{Var}(X_i) = 0.1$，则按相同步骤得到 $\mathrm{sd}(S_n) = \sqrt{5000 \times 0.1} \approx 223.607$，相应的 $z \approx 0.0447$，概率约 0.4822.)

题目 6

设总体 X 具有概率密度

$$f(x) = \begin{cases} \dfrac{1}{\theta^2} x e^{-x/\theta}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

其中 $\theta > 0$ 为未知参数，$X_1, X_2, \ldots, X_n$ 是来自 X 的样本，$x_1, x_2, \ldots, x_n$ 是相应的样本观测值.

(1) 求 θ 的最大似然估计量;

(2) 求 θ 的矩估计量和矩估计值 (已知 $E(X) = 2\theta$).

解答 总体密度

$$f(x) = \begin{cases} \dfrac{1}{\theta^2}\, x\, e^{-x/\theta}, & x > 0, \\ 0, & x \le 0, \end{cases} \qquad \theta > 0.$$

样本 $X_1,\dots,X_n$，观测值 $x_1,\dots,x_n$.

(1) 最大似然估计 (MLE):

似然函数 (关于参数 θ 的函数) 为

$$L(\theta)=\prod_{i=1}^{n}\frac{x_i}{\theta^2}e^{-x_i/\theta}=\left(\prod_{i=1}^{n}x_i\right)\theta^{-2n}\exp\Big(-\frac{\sum_{i=1}^{n}x_i}{\theta}\Big).$$

取对数似然

$$\ell(\theta)=\ln L(\theta)=\sum_{i=1}^{n}\ln x_i-2n\ln\theta-\frac{\sum_{i=1}^{n}x_i}{\theta}.$$

对 θ 求导并令导数为零:

$$\frac{d\ell}{d\theta}=-\frac{2n}{\theta}+\frac{\sum_{i=1}^{n}x_i}{\theta^2}=0.$$

解得

$$\sum_{i=1}^{n}x_i=2n\theta\quad\Rightarrow\quad\hat\theta_{\text{MLE}}=\frac{\sum_{i=1}^{n}x_i}{2n}=\frac{\overline{X}}{2}.$$

因此

$$\boxed{\hat\theta_{\text{MLE}}=\frac{\sum_{i=1}^{n}x_i}{2n}=\frac{\overline{X}}{2}\,.}$$

(2) 动差估计 (矩估计):

已知总体 $E(X)=2\theta$(题中已给出). 用第一阶矩匹配样本均值 $\overline{X}$:

$$\overline{X}\approx E(X)=2\theta\quad\Rightarrow\quad\hat\theta_{\text{MM}}=\frac{\overline{X}}{2}.$$

因此矩估计和矩估计值 (用观测数据替代 $\overline{X}$) 为

$$\boxed{\hat\theta_{\text{MM}}=\frac{\overline{X}}{2}\,.}$$

2 2024~2025 学年概率论公共课期末试卷 A 卷

2.1 选择题 (每小题 3 分，共 15 分)

题目 1

A,B,C 是相互独立的三个事件且 $P(A)=P(B)=P(C)=0.3$，则 $P(A\cup B\cup C)$ 的值为 (　　)

(A) 0.9　　(B) 0.3　　(C) 0.027　　(D) 0.657

解答　选 (D). 根据三个事件的概率加法公式和事件的独立性，有

$$\begin{aligned}P(A\cup B\cup C)&=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)\\&=P(A)+P(B)+P(C)-P(A)P(B)-P(A)P(C)-P(B)P(C)+P(A)P(B)P(C)\\&=0.3+0.3+0.3-0.09-0.09-0.09+0.027\\&=0.657.\end{aligned}$$

题目 2

设随机变量 $X\sim N(0,1)$，X 的分布函数为 $\Phi(x)$，则 $P\{|X|>2\}$ 的值为 (　　)

(A) $2[1-\Phi(2)]$　　(B) $2\Phi(2)-1$　　(C) $2-\Phi(2)$　　(D) $1-2\Phi(2)$

解答　选 (A). 根据正态分布函数的定义:

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

和正态分布的对称性

$$\Phi(-x)=1-\Phi(x),$$

得到

$$\begin{aligned}P\{|X|>2\}&=P\{X<-2\}+P\{X>2\}\\&=P\{X<-2\}+(1-P\{X\leqslant 2\})\\&=\Phi(-2)+[1-\Phi(2)]\\&=[1-\Phi(2)]+[1-\Phi(2)]\\&=2[1-\Phi(2)].\end{aligned}$$

题目 3

设随机变量 $X \sim \pi(2)$，则 $P\{X \leqslant 1\}$ 的值为 (　　)

(A) e^{-2}　　(B) $2e^{-2}$　　(C) $3e^{-2}$　　(D) $4e^{-2}$

解答　选 (C). 根据泊松分布的定义

$$P\{X=k\}=\frac{\lambda^k e^{-\lambda}}{k!}.$$

当 $\lambda=2$ 时，有

$$P\{X\leqslant 1\}=P\{X=0\}+P\{X=1\}=\frac{2^0e^{-2}}{0!}+\frac{2^1e^{-2}}{1!}=3e^{-2}.$$

题目 4

一盒中有 3 个红球，1 个白球，不放回取 2 个球，X 表示取到的红球数，$F(x)$ 是 X 的分布函数，则 $F(1.5)$ 的值为 (　　)

(A) 0　　(B) 0.25　　(C) 0.5　　(D) 0.75

解答　选 (C). 根据分布函数的定义:

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

有

$$F(1.5)=P\{X\leqslant 1.5\}=P\{X=0\}+P\{X=1\}=0+\frac{C_3^1C_1^1}{C_4^2}=\frac{1}{2}.$$

题目 5

已知 (X,Y) 的联合分布律为 $P\{X=1,Y=1\}=0.1$，$P\{X=1,Y=2\}=0.3$，$P\{X=2,Y=1\}=0.4$，$P\{X=2,Y=2\}=0.2$，$F(x,y)$ 是 (X,Y) 的分布函数，$F_X(x)$ 是 X 的边际分布函数，则以下结果正确的是 (　　)

(A) $F(1.5,2)=0.1$　　(B) $F_X(2.5)=1$　　(C) $F_X(1.5)=0$　　(D) $F(2,2)=0.2$

解答　选 (B). 二维随机变量的 (X,Y) 联合分布函数定义为

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

二维随机变量 (X,Y) 关于 X 和关于 Y 的边缘分布函数定义为

$$F_X(x)=P\{X\leqslant x\}=P\{X\leqslant x,Y<+\infty\}=F(x,+\infty).$$

$$F_Y(y)=P\{Y\leqslant y\}=P\{X\leqslant +\infty,Y<y\}=F(+\infty,y).$$

对于选项 (A)

$$\begin{aligned}F(1.5,2)&=P\{X\leqslant 1.5,Y\leqslant 2\}\\&=P\{X=1,Y=1\}+P\{X=1,Y=2\}\\&=0.1+0.3\\&=0.4.\end{aligned}$$

对于选项 (B)

$$\begin{aligned}F_X(2.5)&=P\{X\leqslant 2.5,Y\leqslant +\infty\}\\&=P\{X=1,Y=1\}+P\{X=1,Y=2\}+P\{X=2,Y=1\}+P\{X=2,Y=2\}\\&=0.1+0.3+0.4+0.2\\&=1.\end{aligned}$$

对于选项 (C)

$$F_X(1.5)=P\{X=1,Y=1\}+P\{X=1,Y=2\}=0.1+0.3=0.4.$$

对于选项 (D)

$$\begin{aligned}F(2,2)&=P\{X\leqslant 2,Y\leqslant 2\}\\&=P\{X=1,Y=1\}+P\{X=1,Y=2\}+P\{X=2,Y=1\}+P\{X=2,Y=2\}\\&=0.1+0.3+0.4+0.2\\&=1.\end{aligned}$$

2.2　填空题 (每小题 3 分，共 15 分)

题目 6

已知 $P(A\cup B)=0.7$，$P(A)=0.4$，当 A 与 B 不相容时，$P(B)=$ __________.

解答　根据概率的加法公式和事件 A, B 互不相容，有

$$\begin{cases} P(A\cup B)=P(A)+P(B)-P(AB), \\ P(AB)=P(\varnothing)=0. \end{cases} \implies P(B)=0.3.$$

题目 7

已知 (X,Y) 的联合分布律为

$X\setminus Y$	-1	0	1
-1	$1/18$	$2/18$	0
0	$3/18$	$4/18$	$5/18$
2	0	$2/18$	$1/18$

则 $P\{X\leqslant 0, |Y|<1\}=$ ________.

解答

$$\begin{aligned} P\{X\leqslant 0,|Y|<1\} &= P\{X\leqslant 0,-1<Y<1\} \\ &= P\{X=-1,Y=0\}+P\{X=0,Y=0\} \\ &= \frac{2}{18}+\frac{4}{18} \\ &= \frac{1}{3}. \end{aligned}$$

题目 8

设随机变量 X 服从参数 $\theta=2$ 的指数分布，则 $P\{X\geqslant 2\}=$ ________.

解答　随机变量 X 服从参数 $\theta=2$ 的指数分布

$$f(x)=\begin{cases} \frac{1}{2}\mathrm{e}^{-\frac{x}{2}}, & x>0, \\ 0, & \text{其他}. \end{cases}$$

在对应区间积分，得到

$$P\{X\geqslant 2\}=\int_2^{+\infty}\frac{1}{2}\mathrm{e}^{-\frac{x}{2}}\,\mathrm{d}x=-\int_2^{+\infty}\mathrm{e}^{-\frac{x}{2}}\mathrm{d}\left(-\frac{x}{2}\right)=-\left(\mathrm{e}^{-\frac{x}{2}}\Big|_2^{+\infty}\right)=\mathrm{e}^{-1}.$$

题目 9

设随机变量 $X \sim N(0,4)$，其概率密度函数为 $f_1(x)$，随机变量 $Y \sim U(-1,4)$，其概率密度函数为 $f_2(y)$. 若 $f(x)=\begin{cases} af_1(x), & x \leqslant 0, \\ bf_2(x), & x>0. \end{cases}$ 是概率密度函数，其中 $a>0, b>0$，则 a,b 应满足 __________.

解答　根据概率密度的性质

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

有

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x &= \int_{-\infty}^{0} af_1(x)\,\mathrm{d}x + \int_{0}^{+\infty} bf_2(x)\,\mathrm{d}x \\ 1 &= a\int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{8}}\,\mathrm{d}x + b\int_{0}^{4} \frac{1}{4-(-1)}\mathrm{d}x \\ 1 &= \frac{a}{2}+\frac{4b}{5}. \end{aligned}$$

综上，a,b 满足 $5a+8b=10$.

题目 10

随机变量 X 在区间 $(1,3)$ 上服从均匀分布，对 X 独立重复观察 3 次，则至少有 2 次观测值大于 1.5 的概率为 __________.

解答　随机变量 X 的概率密度为

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

观测值大于 1.5 的概率为

$$P\{X \geqslant 1.5\} = \int_{1.5}^{3} \frac{1}{2}\,\mathrm{d}x = \frac{3}{4}.$$

对 X 独立重复观察 3 次，至少有 2 次观测值大于 1.5 的概率为

$$P = C_3^2\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^1 + C_3^3\left(\frac{3}{4}\right)^3\left(\frac{1}{4}\right)^0 = \frac{27}{32}.$$

2.3 计算题 (共 70 分)

题目 11(10 分)

幼儿园举行家长开放日活动，小浩家参加活动的家长是父亲或者母亲. 父亲参加的概率是 0.9，若父亲参加，则母亲参加的概率为 0.1; 若父亲没有参加，则母亲参加的概率为 0.5. 求：

(1) 母亲参加的概率.

(2) 在已知母亲参加的条件下，父亲参加的概率.

解答 设事件 A 为父亲参加开放日，事件 B 为母亲参加开放日.

$$\begin{cases} P(A) = 0.9, & P(\overline{A}) = 0.1, \\ P(B|A) = 0.9, & P(\overline{B}|A) = 0.1, \\ P(B|\overline{A}) = 0.5, & P(\overline{B}|\overline{A}) = 0.5. \end{cases}$$

(1) 根据全概率公式，母亲参加的概率为

$$P(B) = P(B|A)P(A) + P(B|\overline{A})P(\overline{A}) = 0.1 \times 0.9 + 0.5 \times 0.1 = 0.14.$$

(2) 根据贝叶斯公式，在已知母亲参加的条件下，父亲参加的概率为

$$\begin{aligned} P(A|B) &= \frac{P(B|A_i)P(A_i)}{\sum\limits_{j=1}^{2} P(B|A_j)P(A_j)} \\ &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\overline{A})P(\overline{A})} \\ &= \frac{0.1 \times 0.9}{0.1 \times 0.9 + 0.5 \times 0.1} \\ &= \frac{9}{14}. \end{aligned}$$

题目 12(10 分)

在一次打靶训练中，某运动员给自己定的目标是要一直打到命中 10 环为止. 已知他每次命中 10 环的概率为 p，$0 < p < 1$. 设各次是否命中相互独立，求：

(1) 他打靶次数的分布律.

(2) 如果他打了 n 次还没有命中，还需要进行 m 次的概率.

(3) 如果他打了 5 分钟还没有命中，还需要进行 m 次的概率.

解答　(1) 用随机变量 X 表示打靶的次数，其分布律为

$$P\{X=k\}=(1-p)^{k-1}p.$$

(2) 用随机变量 X 表示已经射击的次数，Y 表示继续射击的次数，结合条件概率公式，有

$$\begin{aligned}P\{Y=m|X=n\}&=\frac{P\{Y=m,X=n\}}{P\{X=n\}}\\&=\frac{(1-p)^{n+m-1}p}{(1-p)^n}\\&=(1-p)^{m-1}p.\end{aligned}$$

(3) 根据第 (2) 小题的结果，继续射击的次数与之前的射击结果没有关系

$$P=(1-p)^{m-1}p.$$

题目 13(12 分)

设 A,B 为随机事件，且 $P(A)=0.3,P(B)=0.12,P(AB)=0.06$，令

$$X=\begin{cases}1, & A\text{发生},\\0, & A\text{不发生},\end{cases}\quad Y=\begin{cases}1, & B\text{发生},\\0, & B\text{不发生}.\end{cases}$$

求 (X,Y) 的联合概率密度及边缘分布律.

解答　根据事件运算的分配律

$$P(A\overline{B})=P(A(S-B))=P(A)-P(AB)=0.24.$$

$$P(\overline{A}B)=P((S-A)B)=P(B)-P(AB)=0.06.$$

根据事件运算的德摩根律

$$P(\overline{A}\overline{B})=P(\overline{A\cup B})=1-[P(A)+P(B)-P(AB)]=0.64.$$

得到 (X, Y) 的联合概率密度及边缘分布律

$X \setminus Y$	0	1	$P\{X=i\}$
0	0.64	0.24	0.88
1	0.06	0.06	0.12
$P\{Y=j\}$	0.70	0.30	1

题目 14(12 分)

某化合物的酒精含量百分比 X 是随机变量，其概率密度函数为:

$$f(x)=\begin{cases} c(1-x), & 0.3<x<0.7, \\ 0, & \text{其他}. \end{cases}$$

此化合物的成本为每升 10 元，售价为每升 40 元或 60 元. 若 $0.4<X<0.6$ ，则售价以 0.9 的概率为 60 元，否则以概率 0.2 为 60 元. 以 Y 表示每升的利润，求:

(1) 常数 c.

(2) X 的分布函数 $F(x)$.

(3) Y 的分布律和分布函数.

解答 (1) 根据概率密度的性质

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x=\int_{0.3}^{0.7} c(1-x)\,\mathrm{d}x=c\left(x-\frac{1}{2}x^2\right)\bigg|_{0.3}^{0.7}=1,$$

解得

$$c=5.$$

(2) 分布函数 $F(x)$ 定义为

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

情形一

当 $x \leqslant 0.3$ 时

$$F(x)=\int_{-\infty}^{x} 0\,\mathrm{d}x=0.$$

情形二

当 $0.3 < x < 0.7$ 时

$$F(x)=\int_{-\infty}^{0.3}0\,\mathrm{d}x+\int_{0.3}^{x}5(1-x)\,\mathrm{d}x=\left(-\frac{5}{2}x^2+5x\right)\Big|_{0.3}^{x}=-\frac{5}{2}x^2+5x-\frac{51}{40}.$$

情形三

当 $x \geqslant 0.7$ 时

$$F(x)=\int_{-\infty}^{0.3}0\,\mathrm{d}x+\int_{0.3}^{0.7}5(1-x)\,\mathrm{d}x+\int_{0.7}^{x}0\,\mathrm{d}x=1.$$

综上所述，分布函数 $F(x)$ 为

$$F(x)=\int_{-\infty}^{x}f(x)\,\mathrm{d}x=\begin{cases}0, & x\leqslant\dfrac{3}{10},\\ -\dfrac{5}{2}x^2+5x-\dfrac{51}{40}, & \dfrac{3}{10}<x<\dfrac{7}{10},\\ 1, & x\geqslant\dfrac{7}{10}.\end{cases}$$

(3) 根据全概率公式，有

$$\begin{aligned}P\{Y=50,0.4<X<0.6\}&=P\{Y=30|0.4<X<0.6\}P\{0.4<X<0.6\}\\&=0.9[F(0.6)-F(0.4)]\\&=0.9\times0.5\\&=0.45.\end{aligned}$$

$$\begin{aligned}P\{Y=30,0.4<X<0.6\}&=P\{Y=30|0.4<X<0.6\}P\{0.4<X<0.6\}\\&=0.1[F(0.6)-F(0.4)]\\&=0.1\times0.5\\&=0.05.\end{aligned}$$

$$\begin{aligned}P\{Y=50,X\notin(0.4,0.6)\}&=P\{Y=50|X\notin(0.4,0.6)\}P\{X\notin(0.4,0.6)\}\\&=0.2[F(0.4)-F(0.3)+F(0.7)-F(0.6)]\\&=0.2\times0.5\\&=0.1.\end{aligned}$$

$$\begin{aligned}P\{Y=30, X \notin (0.4,0.6)\} &= P\{Y=30|X \notin (0.4,0.6)\}P\{X \notin (0.4,0.6)\}\\ &= 0.8[F(0.4)-F(0.3)+F(0.7)-F(0.6)]\\ &= 0.8\times 0.5\\ &= 0.4.\end{aligned}$$

$$P\{Y=50\} = P\{Y=50, 0.4<X<0.6\} + P\{Y=50, X \notin (0.4,0.6)\} = 0.55.$$

$$P\{Y=30\} = P\{Y=30, 0.4<X<0.6\} + P\{Y=30, X \notin (0.4,0.6)\} = 0.45.$$

综上，Y 的分布律为

Y	30	50
p_k	0.45	0.55

分布函数 $F(y)$ 为

$$F(y) = \begin{cases} 0, & y<30,\\ \dfrac{9}{20}, & 30 \leqslant y < 50,\\ 1, & y \geqslant 50.\end{cases}$$

题目 15(12 分)

设 X 的概率密度

$$f(x) = \begin{cases} \dfrac{1}{4}, & -2<x<0,\\ \dfrac{1}{6}, & 0 \leqslant x < 3,\\ 0, & \text{其他}.\end{cases}$$

令 $Y=X^4$，求：

(1) Y 的概率密度 $f_Y(y)$.

(2) $P\left\{X \leqslant -\dfrac{1}{2}, Y \leqslant 1\right\}$.

解答 (1) 根据分布函数的定义

$$F_Y(y) = P\{Y \leqslant y\} = P\left\{X^4 \leqslant y\right\} = P\left\{-y^{\frac{1}{4}} \leqslant X \leqslant y^{\frac{1}{4}}\right\} = F_X\left(y^{\frac{1}{4}}\right) - F_X\left(-y^{\frac{1}{4}}\right),$$

将 $F_Y(y)$ 对 y 求导得到 $f_Y(y)$

$$\begin{aligned} f_Y(y) = F_Y'(y) &= F_X'\left(y^{\frac{1}{4}}\right) - F_X'\left(-y^{\frac{1}{4}}\right) \\ &= f_X\left(y^{\frac{1}{4}}\right)\left(\frac{1}{4}y^{-\frac{3}{4}}\right) - f_X\left(-y^{\frac{1}{4}}\right)\left(-\frac{1}{4}y^{-\frac{3}{4}}\right) \\ &= \frac{1}{4}y^{-\frac{3}{4}}\left[f_X\left(y^{\frac{1}{4}}\right) + f_X\left(-y^{\frac{1}{4}}\right)\right], \end{aligned}$$

综上，Y 的概率密度函数为

$$f_Y(y) = \begin{cases} \frac{5}{48}y^{-\frac{3}{4}}, & 0 \leqslant y < 16, \\ \frac{1}{24}y^{-\frac{3}{4}}, & 16 \leqslant y < 81, \\ 0, & \text{其他}. \end{cases}$$

(2)

$$\begin{aligned} P\left\{X \leqslant -\frac{1}{2}, Y \leqslant 1\right\} &= P\left\{X \leqslant -\frac{1}{2}, X^4 \leqslant 1\right\} \\ &= P\left\{X \leqslant -\frac{1}{2}, -1 \leqslant X \leqslant 1\right\} \\ &= P\left\{-1 \leqslant X \leqslant -\frac{1}{2}\right\} \\ &= \int_{-1}^{-\frac{1}{2}} \frac{1}{4}\,\mathrm{d}x \\ &= \frac{1}{8}. \end{aligned}$$

题目 16(14 分)

设 (X,Y) 的概率密度为

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

求：

(1) 常数 k.

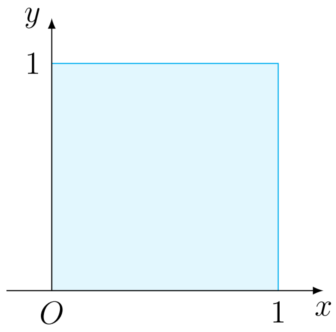
(2) $f_X(x), f_Y(y)$.

(3) $P\{X+Y>1\}$.

解答 (1) 根据概率密度的性质

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

结合如图所示的积分区域



得到

$$\int_0^1\int_0^1 kx^2y\,\mathrm{d}x\mathrm{d}y = \int_0^1 kx^2\mathrm{d}x\int_0^1 y\,\mathrm{d}y = \int_0^1 kx^2\left(\frac{y^2}{2}\bigg|_0^1\right)\mathrm{d}x = \frac{k}{6}x^3\bigg|_0^1 = 1,$$

解得

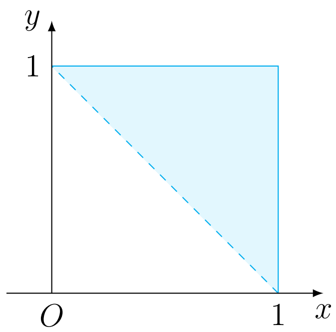
$$k = 6.$$

(2) 根据边缘概率密度的定义

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}y = \begin{cases}\displaystyle\int_0^1 6x^2y\,\mathrm{d}y, & 0 \leqslant x \leqslant 1,\\ 0, & \text{其他}.\end{cases} = \begin{cases}3x^2, & 0 \leqslant x \leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}y = \begin{cases}\displaystyle\int_0^1 6x^2y\,\mathrm{d}x, & 0 \leqslant y \leqslant 1,\\ 0, & \text{其他}.\end{cases} = \begin{cases}2y, & 0 \leqslant y \leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

(3) 新的积分区域如图所示



根据概率密度的性质，得到

$$P\{X+Y>1\}=\int_0^1 \mathrm{d}x\int_{1-x}^1 6x^2y\,\mathrm{d}y=\int_0^1 3x^2\left[1-(1-x)^2\right]\,\mathrm{d}x=\frac{9}{10}.$$

3 概率论的基本概念

题目 2

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

(1) A 发生，B 与 C 不发生.

(2) A 与 B 都发生，而 C 不发生.

(3) A, B, C 中至少有一个发生.

(4) A, B, C 都发生.

(5) A, B, C 都不发生.

(6) A, B, C 中不多于一个发生.

(7) A, B, C 中不多于两个发生.

(8) A, B, C 中至少有两个发生.

解答

(1) $A\overline{BC}$

(2) $AB\overline{C}$

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

(4) ABC

(5) $\overline{ABC}$

(6) $\overline{AB} \cup \overline{AC} \cup \overline{BC}$

(7) $\overline{A} \cup \overline{B} \cup \overline{C}$

(8) $AB \cup AC \cup BC$

题目 3

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

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

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

解答 (1) 根据多个事件的加法公式

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

(2) (i) 根据概率的加法公式

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

(ii) 根据事件运算的德摩根律

$$P\left(\overline{A}\overline{B}\right)=P\left(\overline{A\cup B}\right)=1-P(A\cup B)=1-\frac{11}{15}=\frac{4}{15}.$$

(iii) 根据多个事件的加法公式

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

(iv) 根据事件运算的德摩根律

$$P\left(\overline{A}\overline{B}\overline{C}\right)=P\left(\overline{A\cup B\cup C}\right)=1-P(A\cup B\cup C)=1-\frac{17}{20}=\frac{3}{20}.$$

(v) 设样本空间为 S，利用 $\overline{C}=S-C$ 转化

$$P(\overline{A}\overline{B}C)=P\left(\overline{A}\overline{B}(S-\overline{C})\right)=P\left(\overline{A}\overline{B}\right)-P\left(\overline{A}\overline{B}\overline{C}\right)=\frac{4}{15}-\frac{3}{20}=\frac{7}{60}.$$

(vi) 将 $\overline{A}\overline{B}$ 视作一个事件，根据加法公式

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

(3) 设样本空间为 S

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

(i) 因为 A, B 互不相容，所以 $P(AB) = 0$

$$P(A\overline{B}) = P(A) - P(AB) = \frac{1}{2} - 0 = \frac{1}{2}.$$

(ii) 若 $P(AB) = \dfrac{1}{8}$

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

题目 4

设 A, B 是两个事件.

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

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

解答 (1) 设样本空间为 S

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

(2) 设样本空间为 S

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

题目 6

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

(1) 求最小号码为 5 的概率.

(2) 求最大号码为 5 的概率.

解答 (1) 事件 A_1 包含的基本事件数为 C_5^2，S 中基本事件的总数为 C_{10}^3

$$P(A_1)=\frac{C_5^2}{C_{10}^3}=\frac{1}{12}.$$

(2) 事件 A_2 包含的基本事件数为 C_4^2，S 中基本事件的总数为 C_{10}^3

$$P(A_2)=\frac{C_4^2}{C_{10}^3}=\frac{1}{20}.$$

题目 7

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

解答 事件 A 包含的基本事件数为 $C_{10}^4C_4^3C_3^2$，S 中基本事件的总数为 C_{17}^9.

$$P(A)=\frac{C_{10}^4C_4^3C_3^2}{C_{17}^9}=\frac{252}{2431}.$$

补充习题

某厂家称一批数量为 1000 件的产品的次品率为 5%. 现从该批产品中有放回地抽取了 30 件，经检验发现有次品 5 件，问该厂家是否谎报了次品率？

解答 若次品率为厂家声称的 5%，那么当有放回地抽检 30 件时，出现次品的数量应为

$$30\times 5\%=1.5.$$

而实际抽检出 5 个次品，这说明厂家谎报了次品率. **方法二**假设商家如实地上报次品率，那么有放回地抽取 30 件，出现 5 件次品的概率为

$$P=C_{30}^5(5\%)^5(1-5\%)^{25}\approx 0.01235.$$

这是一个很低概率的事件，说明厂家谎报了次品率.

题目 8

在 1500 件产品中有 400 件次品、1100 件正品. 任取 200 件.

(1) 求恰有 90 件次品的概率.

(2) 求至少有 2 件次品的概率.

解答 (1) 设事件 A_1 为恰有 90 件次品

$$P(A_1) = \frac{C_{400}^{90}C_{1100}^{110}}{C_{1500}^{200}}.$$

(2) 设事件 A_2 为至少有 2 件次品，考虑其对立事件 $\overline{A_2} = \{\{\text{没有次品}\}, \{\text{只有 1 件次品}\}\}$

$$P(A_2) = 1 - P(\overline{A_2}) = 1 - \frac{C_{400}^{0}C_{1100}^{200}}{C_{1500}^{200}} - \frac{C_{400}^{1}C_{1100}^{199}}{C_{1500}^{200}}.$$

题目 11

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

解答 设杯中最多有 i 个球为事件 A_i

$$P(A_1) = \frac{C_4^1C_3^1C_2^1}{C_4^1C_4^1C_4^1} = \frac{6}{16}.$$

$$P(A_2) = \frac{C_3^2C_4^1C_3^1}{C_4^1C_4^1C_4^1} = \frac{9}{16}.$$

$$P(A_3) = \frac{C_4^1}{C_4^1C_4^1C_4^1} = \frac{1}{16}.$$

题目 14

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

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

解答 (1) 根据题设条件，有

$$\begin{cases} P(\overline{A}) = 1 - P(A) = 0.3, \\ P(B) = 1 - P(\overline{B}) = 0.4, \\ P(A\overline{B}) = P(A) - P(AB) = 0.5. \end{cases} \implies \begin{cases} P(A) = 0.7, \\ P(\overline{B}) = 0.6, \\ P(AB) = 0.2. \end{cases}$$

再根据条件概率公式和事件运算的分配律，有

$$P(B|A \cup \overline{B}) = \frac{P(B(A \cup \overline{B}))}{P(A \cup \overline{B})} = \frac{P(BA \cup B\overline{B})}{P(A \cup \overline{B})} = \frac{P(AB)}{P(A) + P(\overline{B}) - P(A\overline{B})} = \frac{1}{4}.$$

(2) 根据题设条件和乘法定理

$$\begin{cases} P(AB) = P(B|A)P(A), \\ P(AB) = P(A|B)P(B). \end{cases} \Longrightarrow \begin{cases} P(B) = \dfrac{1}{6}, \\ P(AB) = \dfrac{1}{12}. \end{cases}$$

进一步有

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

题目 16

据以往资料表明，某 3 口之家，患某种传染病的概率有以下规律：

$$P\{孩子得病\} = 0.6,$$

$$P\{母亲得病|孩子得病\} = 0.5,$$

$$P\{父亲得病|母亲及孩子得病\} = 0.4,$$

求母亲及孩子得病但父亲未得病的概率.

解答 设样本空间为 S，事件 A 为孩子得病，事件 B 为母亲得病，事件 C 为父亲得病

$$\begin{cases} P(A) = 0.6, \\ P(B|A) = \dfrac{P(AB)}{P(A)} = 0.5, \\ P(C|BA) = \dfrac{P(ABC)}{P(AB)} = 0.4. \end{cases} \Longrightarrow \begin{cases} P(AB) = 0.3, \\ P(ABC) = 0.12. \end{cases}$$

进一步有

$$P(AB\overline{C}) = P(AB(S-C)) = P(AB) - P(ABC) = 0.18.$$

题目 37

设第一只盒子中装有 3 只蓝球，2 只绿球，2 只白球；第二只盒子中装有 2 只蓝球，3 只绿球，4 只白球. 独立地分别在两只盒子中各取一只球.

(1) 求至少有一只蓝球的概率.

(2) 求有一只蓝球、一只白球的概率.

(3) 已知至少有一只蓝球，求有一只蓝球一只白球的概率.

解答 (1) 设事件 A 为至少有一只蓝球，考虑其对立事件 $\overline{A}$

$$P(A) = 1 - P(\overline{A}) = 1 - \frac{C_4^1}{C_7^1}\frac{C_7^1}{C_9^1} = \frac{5}{9}.$$

(2) 设事件 B 为取出一只蓝球和一只白球

$$P(B) = \frac{C_3^1C_4^1}{C_7^1C_9^1} + \frac{C_2^1C_2^1}{C_7^1C_9^1} = \frac{16}{63}.$$

(3) 根据 (1)(2) 题的结论和条件概率公式

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{16/63}{5/9} = \frac{16}{35}.$$

4 随机变量及其分布

补充习题

设某便利店一天中的顾客人数服从参数为 10 的泊松分布，并且这些到店顾客的购物行为彼此独立，其购物概率为 0.6，求每天到便利店且购物的顾客人数的分布.

解答 我们约定用随机变量 X 表示到店的人数，Y 表示到店且购物的人数. 设当天有 $r(r=0,1,2,\cdots)$ 位顾客到店，其中 $k(k\leqslant r)$ 位顾客在到店后选择购物，那么

$$P\{Y=k|X=r\}=C_r^k(0.6)^k(1-0.6)^{r-k}.$$

再根据全概率公式

$$\begin{aligned}P\{Y=k\}&=\sum_{r=k}^{\infty}P\{Y=k|X=r\}P\{X=r\}\\&=\sum_{r=k}^{\infty}C_r^k(0.6)^k(1-0.6)^{r-k}\frac{10^r\mathrm{e}^{-10}}{r!}\\&=\sum_{r=k}^{\infty}\frac{r!}{k!(r-k)!}(0.6)^k(0.4)^{r-k}\frac{10^r\mathrm{e}^{-10}}{r!}.\end{aligned}$$

约去 $r!$ 后，进一步令 $r-k=i$

$$\begin{aligned}P\{Y=k\}&=\sum_{i=0}^{\infty}\frac{1}{k!i!}(0.6)^k(0.4)^i10^{i+k}\mathrm{e}^{-10}\\&=\frac{0.6^k10^k\mathrm{e}^{-10}}{k!}\sum_{i=0}^{\infty}\frac{0.4^i10^i}{i!}\\&=\frac{6^k\mathrm{e}^{-10}}{k!}\sum_{i=0}^{\infty}\frac{4^i}{i!}\\&=\frac{6^k\mathrm{e}^{-6}}{k!}.\end{aligned}$$

题目 2

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

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

解答 (1) 随机变量 X 的取值范围为 $3,4,5$

$$
\begin{aligned}
P\{X=3\} &= \frac{C_2^2}{C_5^3} = \frac{1}{10},\\
P\{X=4\} &= \frac{C_3^2}{C_5^3} = \frac{3}{10},\\
P\{X=5\} &= \frac{C_4^2}{C_5^3} = \frac{6}{10}.
\end{aligned}
$$

其分布律为

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

(2) 随机变量 X 的取值范围为 $1,2,3,4,5,6$

$$
\begin{aligned}
P\{X=1\} &= A_2^2\frac{C_5^1}{C_6^1C_6^1} + \frac{1}{C_6^1C_6^1} = \frac{11}{36},\\
P\{X=2\} &= A_2^2\frac{C_4^1}{C_6^1C_6^1} + \frac{1}{C_6^1C_6^1} = \frac{9}{36},\\
P\{X=3\} &= A_2^2\frac{C_3^1}{C_6^1C_6^1} + \frac{1}{C_6^1C_6^1} = \frac{7}{36},\\
P\{X=4\} &= A_2^2\frac{C_2^1}{C_6^1C_6^1} + \frac{1}{C_6^1C_6^1} = \frac{5}{36},\\
P\{X=5\} &= A_2^2\frac{C_1^1}{C_6^1C_6^1} + \frac{1}{C_6^1C_6^1} = \frac{3}{36},\\
P\{X=6\} &= \frac{1}{C_6^1C_6^1} = \frac{1}{36}.
\end{aligned}
$$

其分布律为

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

题目 6

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

(1) 恰有 2 台设备被使用的概率是多少？

(2) 至少有 3 台设备被使用的概率是多少？

(3) 至多有 3 台设备被使用的概率是多少？

(4) 至少有 1 台设备被使用的概率是多少？

解答 我们约定随机变量 X 为正在被使用的设备数量，其分布列为

$$P\{X=k\}=C_5^k(0.1)^k(1-0.1)^{5-k},(k=0,1,2,3,4,5)$$

X	0	1	2	3	4	5
p_k	0.59049	0.32805	0.0729	0.0081	0.00045	0.00001

(1) 恰有 2 台设备被使用的概率

$$P\{X=2\}=C_5^2(0.1)^2(0.9)^3=0.0729.$$

(2) 至少有 3 台设备被使用的概率

$$P\{X\geqslant 3\}=\sum_{k=3}^{5}p_k=0.00856.$$

(3) 至多有 3 台设备被使用的概率

$$P\{X\leqslant 3\}=\sum_{k=0}^{3}p_k=0.99954.$$

(4) 至少有 1 台设备被使用的概率

$$P\{X\geqslant 1\}=\sum_{k=1}^{5}p_k=0.40951.$$

题目 14

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

(1) 若他外出计划用时 10min ，问其间有电话铃响一次的概率是多少？

(2) 若他希望外出时没有电话的概率至少为 0.5，问他外出应控制最长时间是多少？

解答 (1) 当他外出 $t=\dfrac{1}{6}$ h 时，其间电话铃响 $k=1$ 次的概率

$$P\{X=1\}=\frac{(2t)^1\mathrm{e}^{-2t}}{1!}=\frac{1}{3}\mathrm{e}^{-\frac{1}{3}}.$$

(2) 他外出时没有电话的概率至少为 0.5

$$P\{X=0\}=\frac{(2t)^0\mathrm{e}^{-2t}}{0!}\geqslant 0.5.$$

解得最长的外出时间为 $t\leqslant\dfrac{\ln 2}{2}$ h.

题目 16

有一繁忙的汽车站，每天有大量汽车通过，设一辆汽车在一天的某段时间内出事故的概率为 0.0001. 在某天的该时间段内有 1000 辆汽车通过. 问出事故的车辆数不小于 2 的概率是多少？（利用泊松定理计算）

解答 用随机变量 X 表示出事故的车辆数，且 $X\sim\pi(\lambda)$，其中 $\lambda=np=0.1$

$$\begin{aligned}P\{X\geqslant 2\}&=1-P\{X=0\}-P\{X=1\}\\&=1-\frac{0.1^0\mathrm{e}^{-0.1}}{0!}-\frac{0.1^1\mathrm{e}^{-0.1}}{1!}\\&\approx 0.00468.\end{aligned}$$

题目 17

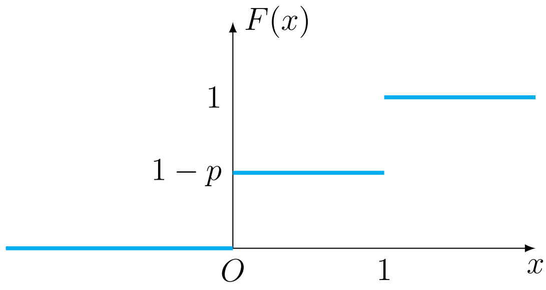
(1) 设 X 服从 $(0-1)$ 分布，其分布律为 $P\{X=k\}=p^k(1-p)^{1-k},(k=0,1)$. 求 X 的分布函数，并作出其图形.

(2) 求第 2 题 (1) 中的随机变量的分布函数.

解答 (1) 根据 $P\{X=0\}=1-p$，$P\{X=1\}=p$ ，得到分布函数 $F(x)$

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

作出分布函数图形



(2) 分布函数为

$$F(x)=\begin{cases}0, & x<3,\\ \dfrac{1}{10}, & 3\leqslant x<4,\\ \dfrac{4}{10}, & 4\leqslant x<5,\\ 1, & x\geqslant 5.\end{cases}$$

题目 20

设随机变量 X 的分布函数为

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

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

(2) 求概率密度 $f_X(x)$.

解答 (1) 根据分布函数的定义

$$P\{X<2\}=F(2)=\ln 2.$$

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

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

(2) 概率密度为

$$f_X(x)=F_X'(x)=\begin{cases}\dfrac{1}{x}, & 1\leqslant x<\mathrm{e},\\ 0, & \text{其他}.\end{cases}$$

题目 24

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

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

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

解答 随机变量 X 的分布函数为

$$F_X(x)=\int_{-\infty}^{x} f(x)\,\mathrm{d}x=\int_0^x \frac{1}{5}\mathrm{e}^{-\frac{x}{5}}\mathrm{d}x=\begin{cases}1-\mathrm{e}^{-\frac{x}{5}}, & x>0,\\ 0, & \text{其他}.\end{cases}$$

顾客在银行等待时间大于 10min 的概率为

$$P\{X>10\}=1-P\{X\leqslant 10\}=1-\left(1-\mathrm{e}^{-2}\right)=\mathrm{e}^{-2}.$$

设一个月内顾客有 k 次因为等待超时而离开窗口，那么 Y 的分布律为

$$P\{Y=k\}=C_5^k\left(\mathrm{e}^{-2}\right)^k\left(1-\mathrm{e}^{-2}\right)^{5-k}\,(k=0,1,2,3,4,5).$$

进一步得到

$$P\{Y\geqslant 1\}=1-P\{Y=0\}=1-\left(1-\mathrm{e}^{-2}\right)^5\approx 0.5167.$$

题目 27

某地区 18 岁的女青年的血压（收缩压以 mmHg 计）服从 $N(110,12^2)$ 分布. 在该地区任选一 18 岁的女青年，测量她的血压 X. 求

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

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

解答 (1) 把正态分布线性变换成标准正态分布，得到

$$\begin{aligned}P\{X \leqslant 105\} &= P\left\{\frac{X-110}{12} \leqslant \frac{105-110}{12}\right\} \\ &= \Phi\left(-\frac{5}{12}\right) \\ &= 1-\Phi\left(\frac{5}{12}\right) \\ &\approx 0.3385.\end{aligned}$$

$$\begin{aligned}P\{100 < X \leqslant 120\} &= P\left\{\frac{100-110}{12} < \frac{X-110}{12} \leqslant \frac{120-110}{12}\right\} \\ &= \Phi\left(\frac{5}{6}\right) - \Phi\left(-\frac{5}{6}\right) \\ &= 2\Phi\left(\frac{5}{6}\right) - 1 \\ &\approx 0.5953.\end{aligned}$$

(2) 题意等价为 $1-P\{X \leqslant x\} \leqslant 0.05$ ，再把正态分布线性变换成标准正态分布，得到

$$\begin{aligned}P\{X \leqslant x\} &\geqslant 0.95 \\ P\left\{\frac{X-110}{12} \leqslant \frac{x-110}{12}\right\} &\geqslant 0.95 \\ \Phi\left(\frac{x-110}{12}\right) &\geqslant 0.95\end{aligned}$$

查表得到 $\left(\dfrac{x-110}{12}\right)_{\min} = 1.65$，也即 $x_{\min} = 129.8$.

题目 33

设随机变量 X 的分布律为

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

求 $Y = X^2$ 的分布律.

解答 根据

$$P\{Y=0\} = P\{X=0\} = \frac{6}{30}.$$

$$P\{Y=1\}=P\{X=-1\}+P\{X=1\}=\frac{7}{30}.$$

$$P\{Y=4\}=P\{X=-2\}=\frac{6}{30}.$$

$$P\{Y=9\}=P\{X=3\}=\frac{11}{30}.$$

得到 Y 的分布律

X	0	1	4	9
p_k	$\frac{6}{30}$	$\frac{7}{30}$	$\frac{6}{30}$	$\frac{11}{30}$

题目 34

设随机变量 X 在区间 $(0,1)$ 服从均匀分布.

(1) 求 $Y=\mathrm{e}^X$ 的概率密度.

(2) 求 $Y=-2\ln X$ 的概率密度.

解答 随机变量 X 在区间 $(0,1)$ 服从均匀分布，概率密度为

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

(1) 根据 $Y=\mathrm{e}^X$，得到 Y 的分布函数

$$F_Y(y)=P\{Y\leqslant y\}=P\left\{\mathrm{e}^X\leqslant y\right\}=P\left\{X\leqslant \ln y\right\}=F_X(\ln y).$$

将 $F_Y(y)$ 对 y 求导，得到 $Y=\mathrm{e}^X$ 的概率密度

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

(2) 根据 $Y=-2\ln X$，得到 Y 的分布函数

$$F_Y(y)=P\{Y\leqslant y\}=P\left\{X\geqslant \mathrm{e}^{-\frac{y}{2}}\right\}=1-P\left\{X<\mathrm{e}^{-\frac{y}{2}}\right\}=1-F_X\left(\mathrm{e}^{-\frac{y}{2}}\right).$$

将 $F_Y(y)$ 对 y 求导，得到 $Y=-2\ln X$ 的概率密度

$$f_Y(y)=\left[1-F_X\left(\mathrm{e}^{-\frac{y}{2}}\right)\right]'=-f_X\left(\mathrm{e}^{-\frac{y}{2}}\right)\left(\mathrm{e}^{-\frac{y}{2}}\right)'=\begin{cases}\dfrac{1}{2}\mathrm{e}^{-\frac{y}{2}}, & 0<\mathrm{e}^{-\frac{y}{2}}<1,\\ 0, & \text{其他}.\end{cases}=\begin{cases}\dfrac{1}{2}\mathrm{e}^{-\frac{y}{2}}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

题目 35

设 $X\sim N(0,1)$.

(1) 求 $Y=\mathrm{e}^X$ 的概率密度.

(2) 求 $Y=2X^2+1$ 的概率密度.

(3) 求 $Y=|X|$ 的概率密度.

解答 随机变量 $X\sim N(0,1)$ ，其概率密度为

$$f(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}.$$

(1) 根据 $Y=\mathrm{e}^X$ 得到 Y 的分布函数

$$F_Y(y)=P\{Y\leqslant y\}=P\{\mathrm{e}^X\leqslant y\}=P\{X\leqslant\ln y\}=F_X(\ln y).$$

将 $F_Y(y)$ 对 y 求导，得到 $Y=\mathrm{e}^X$ 的概率密度

$$f_Y(y)=F_X'(\ln y)=f_X(\ln y)\cdot\frac{1}{y}=\begin{cases}\dfrac{1}{\sqrt{2\pi}y}\mathrm{e}^{-\frac{\ln^2 y}{2}}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

(2) 根据 $Y=2X^2+1$ 得到 Y 的分布函数

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

将 $F_Y(y)$ 对 y 求导，得到 $Y=2X^2+1$ 的概率密度

$$
\begin{aligned}
f_Y(y) &= F_X'\left(\sqrt{\frac{y-1}{2}}\right) - F_X'\left(\sqrt{\frac{y-1}{2}}\right)\\
&= f_X\left(\sqrt{\frac{y-1}{2}}\right)\left(\sqrt{\frac{y-1}{2}}\right)' - f_X\left(-\sqrt{\frac{y-1}{2}}\right)\left(-\sqrt{\frac{y-1}{2}}\right)'\\
&= \begin{cases} \dfrac{1}{2\sqrt{\pi(y-1)}}\mathrm{e}^{-\frac{y-1}{4}}, & y \geqslant 1,\\ 0, & \text{其他}. \end{cases}
\end{aligned}
$$

(3) 根据 $Y=|X|$ 得到 Y 的分布函数

$$
F_Y(y) = P\{Y \leqslant y\} = P\{|X| \leqslant y\} = P\{-y \leqslant X \leqslant y\} = F_X(y) - F_X(-y).
$$

将 $F_Y(y)$ 对 y 求导，得到 $Y=|X|$ 的概率密度

$$
f_Y(y) = F_X'(y) - F_X'(-y) = f_X(y) - [-f_X(-y)] = \begin{cases} \sqrt{\dfrac{2}{\pi}}\mathrm{e}^{-\frac{y^2}{2}}, & y \geqslant 0,\\ 0, & \text{其他}. \end{cases}
$$

5 多维随机变量及其分布

题目 1

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

$$X=\begin{cases}0, & \text{若第一次取出的是正品,}\\ 1, & \text{若第一次取出的是次品,}\end{cases}\qquad Y=\begin{cases}0, & \text{若第二次取出的是正品,}\\ 1, & \text{若第二次取出的是次品.}\end{cases}$$

试分别就 (1)(2) 两种情况，写出 X 和 Y 的联合分布律.

解答 (1) 进行放回抽样时

$$P\{X=0,Y=0\}=\frac{C_{10}^1C_{10}^1}{C_{12}^1C_{12}^1}=\frac{25}{36}.$$

$$P\{X=0,Y=1\}=\frac{C_{10}^1C_{2}^1}{C_{12}^1C_{12}^1}=\frac{5}{36}.$$

$$P\{X=1,Y=0\}=\frac{C_{2}^1C_{10}^1}{C_{12}^1C_{12}^1}=\frac{5}{36}.$$

$$P\{X=1,Y=1\}=\frac{C_{2}^1C_{2}^1}{C_{12}^1C_{12}^1}=\frac{1}{36}.$$

进一步，得到 X 和 Y 的联合分布律：

$Y\setminus X$	0	1	$P\{Y=j\}$
0	25/36	5/36	30/36
1	5/36	1/36	6/36
$P\{X=i\}$	30/36	6/36	1

(2) 进行不放回抽样时

$$P\{X=0,Y=0\}=\frac{C_{10}^1C_{9}^1}{C_{12}^1C_{11}^1}=\frac{45}{66},$$

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

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

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

进一步，得到 X 和 Y 的联合分布律：

$Y \setminus X$	0	1	$P\{Y=j\}$
0	45/66	10/66	55/66
1	10/66	1/66	11/66
$P\{X=i\}$	55/66	11/66	1

题目 2

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

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

解答 (1) 设取出黑球 k（$k=0,1,2,3$）个，取出红球 r（$r=0,1,2$）个，取出白球 $4-k-r$ 个.

$$P\{X=k, Y=r\} = \frac{C_4^k C_2^r C_2^{4-k-r}}{C_7^4}(k+r \leqslant 4),$$

进一步，得到 X 和 Y 的联合分布律：

$Y \setminus X$	0	1	2	3	$P\{Y=j\}$
0	0	0	3/35	2/35	5/35
1	0	6/35	12/35	2/35	20/35
2	1/35	6/35	3/35	0	10/35
$P\{X=i\}$	1/35	12/35	18/35	4/35	1

(2) 根据 (1) 题的结论

$$P\{X>Y\} = \frac{19}{35},$$

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

$$P\{X+Y=3\} = \frac{20}{35},$$

$$P\{X<3-Y\} = \frac{10}{35}.$$

题目 5

设随机变量 (X,Y) 的概率密度为

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

求边缘分布函数.

解答 根据边缘分布函数的定义，X 的边缘分布函数为

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

同理，Y 的边缘分布函数为

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

题目 9

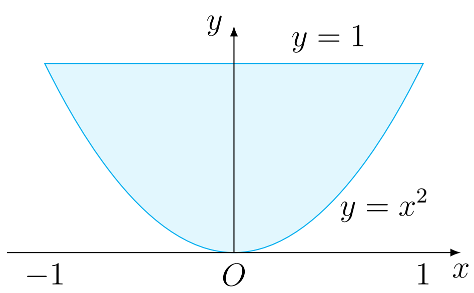
设二维随机变量 (X,Y) 的概率密度为

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

(1) 确定常数 c.

(2) 求边缘概率密度.

解答 (1) 概率密度不为零的区域如图所示



根据概率密度 $f(x,y)$ 的性质

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}x\mathrm{d}y = \int_{-1}^{1}\int_{x^2}^{1} cx^2y\,\mathrm{d}x\mathrm{d}y = c\int_{-1}^{1} x^2\left(\frac{1}{2}-\frac{1}{2}x^4\right)\mathrm{d}x = \frac{4}{21}c = 1.$$

解得

$$c = \frac{21}{4}.$$

(2) 根据边缘概率密度的定义，X 的边缘概率密度为

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}y = \int_{x^2}^{1}\frac{21}{4}x^2y\,\mathrm{d}y = \begin{cases}\dfrac{21}{8}x^2(1-x^4), & -1\leqslant x\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

同理，Y 的边缘概率密度为

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}x = \int_{-\sqrt{y}}^{\sqrt{y}}\frac{21}{4}x^2y\,\mathrm{d}x = \begin{cases}\dfrac{7}{2}y^{\frac{5}{2}}, & 0\leqslant y\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

题目 11

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

$$P\{X=n,Y=m\} = \frac{\mathrm{e}^{-14}(7.14)^m(6.86)^{n-m}}{m!(n-m)!}$$

$$(n=0,1,2,\cdots,\quad m=0,1,2,\cdots n)$$

(1) 求边缘分布律.

(2) 求条件分布律.

(3) 特别，写出当 $X=20$ 时，Y 的条件分布律.

解答 (1) 根据边缘分布函数的定义，X 的边缘分布函数为

$$\begin{aligned}
P\{X=n\} &= \sum_{m=0}^{n} \frac{\mathrm{e}^{-14}(7.14)^m(6.86)^{n-m}}{m!(n-m)!} \\
&= \sum_{m=0}^{n} \frac{\mathrm{e}^{-14}}{n!} \frac{n!}{m!(n-m)!}(7.14)^m(6.86)^{n-m} \\
&= \frac{\mathrm{e}^{-14}}{n!} \sum_{m=0}^{n} C_n^m (7.14)^m(6.86)^{n-m} \\
&= \frac{\mathrm{e}^{-14}}{n!}(7.14+6.86)^n \\
&= \frac{14^n\mathrm{e}^{-14}}{n!}(n=0,1,2,\cdots).
\end{aligned}$$

同理可得 Y 的边缘分布函数

$$\begin{aligned}
P\{Y=m\} &= \sum_{n=m}^{\infty} \frac{\mathrm{e}^{-14}(7.14)^m(6.86)^{n-m}}{m!(n-m)!} \\
&= \frac{\mathrm{e}^{-14}(7.14)^m}{m!} \sum_{n=m}^{\infty} \frac{(6.86)^{n-m}}{(n-m)!} \\
&= \frac{\mathrm{e}^{-14}(7.14)^m}{m!}\mathrm{e}^{6.86} \\
&= \frac{(7.14)^m\mathrm{e}^{-7.14}}{m!}(m=0,1,2,\cdots).
\end{aligned}$$

(2) 根据条件概率公式，得到 X 的条件分布律

$$\begin{aligned}
P\{X=n|Y=m\} &= \frac{P\{X=n,Y=m\}}{P\{Y=m\}} \\
&= \frac{\mathrm{e}^{-14}(7.14)^m(6.86)^{n-m}}{m!(n-m)!} \bigg/ \frac{(7.14)^m\mathrm{e}^{-7.14}}{m!} \\
&= \frac{(6.86)^{n-m}\mathrm{e}^{-6.86}}{(n-m)!}(n-m=0,1,2,\cdots).
\end{aligned}$$

同理，得到 Y 的条件分布律

$$\begin{aligned}
P\{Y=m|X=n\} &= \frac{P\{X=n,Y=m\}}{P\{X=n\}} \\
&= \frac{\mathrm{e}^{-14}(7.14)^m(6.86)^{n-m}}{m!(n-m)!} \bigg/ \frac{14^n\mathrm{e}^{-14}}{n!} \\
&= \frac{n!}{m!(n-m)!}\left(\frac{7.14}{14}\right)^m\left(\frac{6.86}{14}\right)^{n-m} \\
&= C_n^m(0.51)^m(0.49)^{n-m}(m=0,1,\cdots,n).
\end{aligned}$$

(3) 特别地，当 $X=30$ 时

$$P\{Y=m|X=20\}=C_{20}^{m}(0.51)^{m}(0.49)^{20-m}(m=0,1,2,\cdots,n).$$

题目 18

设 X 和 Y 是两个相互独立的随机变量，X 在区间 $(0,1)$ 上服从均匀分布，Y 的概率密度为

$$f_Y(y)=\begin{cases}\frac{1}{2}\mathrm{e}^{-\frac{y}{2}}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

(1) 求 X 和 Y 的联合概率密度.

(2) 设有 a 的二次方程为 $a^2+2Xa+Y=0$，试求 a 有实根的概率.

解答 (1) 随机变量 X 在区间 $(0,1)$ 上服从均匀分布，其概率密度为

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

由于 X 和 Y 相互独立，那么

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

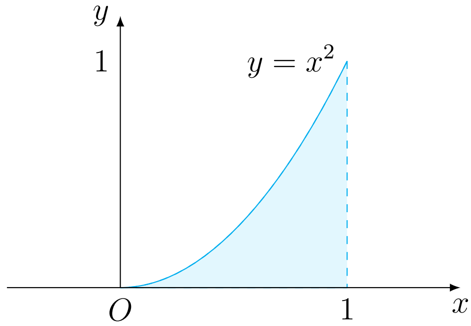
(2) 关于 a 的二次方程有实根，需满足判别式

$$\Delta=4X^2-4Y\geqslant 0.$$

即

$$X^2\geqslant Y$$

概率密度不为零的区域如图所示



与之对应的概率为

$$P\{X^2 \geqslant Y\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}x\mathrm{d}y = \int_0^1 \mathrm{d}x \int_0^{x^2} \frac{1}{2}\mathrm{e}^{-\frac{y}{2}}\mathrm{d}y = \int_0^1 1-\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x.$$

凑出含有标准正态分布的因式

$$P\{X^2 \geqslant Y\} = 1-\sqrt{2\pi}\int_0^1 \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x = 1-\sqrt{2\pi}\,[\Phi(1)-\Phi(0)] \approx 0.1445.$$

或者直接用计算器求定积分

$$P\{X^2 \geqslant Y\} = 1-\int_0^1 \mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x \approx 0.1444.$$

题目 24

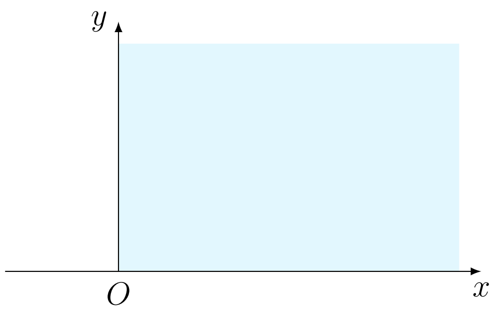
设随机变量 (X,Y) 的概率密度为

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

(1) 问 X 和 Y 是否相互独立?

(2) 求 $Z = X+Y$ 的概率密度.

解答 (1) 概率密度不为零的区域如图所示



根据边缘概率密度的定义，X 的边缘概率密度为

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}y = \int_0^{+\infty} \frac{1}{2}(x+y)\mathrm{e}^{-(x+y)}\,\mathrm{d}y \\
&= -\frac{1}{2}(x+y)\mathrm{e}^{-(x+y)}\Big|_{y=0}^{y=+\infty} - \int_0^{+\infty} \frac{1}{2}\mathrm{e}^{-(x+y)}\mathrm{d}[-(x+y)] \\
&= \frac{x}{2}\mathrm{e}^{-x} - \left[\frac{1}{2}\mathrm{e}^{-(x+y)}\right]\Big|_{y=0}^{y=+\infty} \\
&= \frac{x}{2}\mathrm{e}^{-x} + \frac{1}{2}\mathrm{e}^{-x}.
\end{aligned}
$$

即

$$
f_X(x) = \begin{cases} \dfrac{\mathrm{e}^{-x}}{2}(x+1), & x > 0, \\ 0, & 其他. \end{cases}
$$

同理，Y 的边缘概率密度为

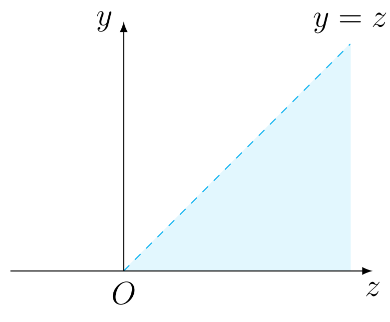
$$
f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}x = \begin{cases} \dfrac{\mathrm{e}^{-y}}{2}(y+1), & y > 0, \\ 0, & 其他. \end{cases}
$$

由此可知

$$
f_X(x)f_Y(y) \neq f(x,y).
$$

即 X 和 Y 不独立.

(2) 概率密度不为零的区域如图所示



根据两个随机变量的函数分布，$Z = X + Y$ 的概率密度可表示为

$$
f_Z(z) = \int_{-\infty}^{+\infty} f(z-y,y)\,\mathrm{d}y = \int_0^z \frac{1}{2}(z-y+y)\mathrm{e}^{-(z-y+y)}\mathrm{d}y = \begin{cases} \dfrac{z^2}{2}\mathrm{e}^{-z}, & z > 0, z > y \\ 0, & 其他. \end{cases}
$$

6 随机变量的数字特征

题目 2

某产品的次品率为 0.1，检验员每天检验 4 次. 每次随机地取 10 件产品进行检验，如发现其中的次品数多于 1，就去调整设备. 以 X 表示一天中调整设备的次数，试求 $E(X)$.（设诸产品是否为次品是相互独立的.）

解答 抽检时次品数多于 1 ，即需要调整设备的概率为

$$\begin{aligned} p &= 1 - C_{10}^{0}(0.1)^{0}(1-0.1)^{10} - C_{10}^{1}(0.1)^{1}(1-0.1)^{9} \\ &= 1 - (0.9)^{10} - (0.9)^{9} \\ &\approx 0.2639. \end{aligned}$$

调整设备的次数 X 所服从的分布律为

$$p_k = C_4^k p^k (1-p)^{4-k}.$$

X	0	1	2	3	4
p_k	$(1-p)^4$	$4p(1-p)^3$	$6p^2(1-p)^2$	$4p^3(1-p)$	p^4

其数学期望为

$$E(X) = \sum_{k=0}^{4} p_k X_k = 0 + 4p(1-p)^3 + 12p^2(1-p)^2 + 12p^3(1-p) + 4p^4 \approx 1.0556.$$

题目 4

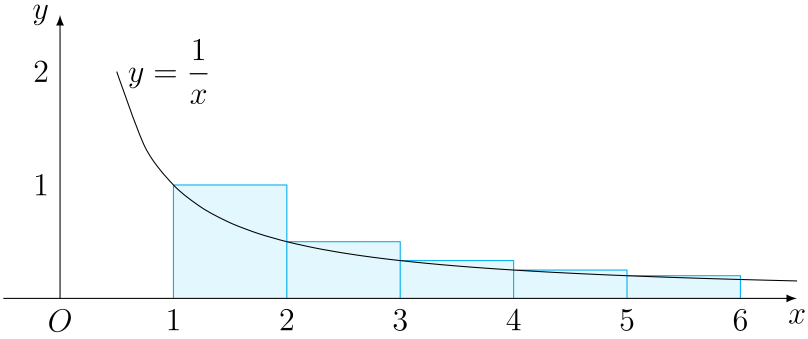
(1) 设随机变量 X 的分布律为 $P\left\{X = (-1)^{j+1}\frac{3^j}{j}\right\} = \frac{2}{3^j}, j = 1, 2, \cdots$，说明 X 的数学期望不存在.

(2) 一盒中装有一只黑球，一只白球，作摸球游戏，规则如下：一次从盒中随机摸一只球，若摸到白球，则游戏结束，摸到黑球放回再放入一只黑球，然后再从盒中随机地摸一只球. 试说明要游戏结束的摸球次数 X 的数学期望不存在.

解答 (1) 根据数学期望的定义

$$E(X) = \sum_{k=1}^{\infty} x_k p_k = \sum_{k=1}^{\infty} (-1)^{k+1}\frac{3^k}{k}\cdot\frac{2}{3^k} = 2\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k} < 2\sum_{k=1}^{\infty}\left|\frac{1}{k}\right|.$$

根据判断调和级数敛散性的积分放缩法



$$\sum_{k=1}^{\infty}\frac{1}{k}>\int_{1}^{+\infty}\frac{1}{x}\,\mathrm{d}x\to\infty.$$

这说明调和级数 $\sum\limits_{k=1}^{\infty}\frac{1}{k}$ 不绝对收敛，级数 $2\sum\limits_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}$ 也不绝对收敛，所以 X 的数学期望不存在.

题目 14

设随机变量 X_1,X_2 的概率密度分别为

$$f_1(x)=\begin{cases}2\mathrm{e}^{-2x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}\qquad f_2(x)=\begin{cases}4\mathrm{e}^{-4x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

(1) 求 $E(X_1+X_2),E\left(2X_1-3X_2^2\right)$.

(2) 又设 X_1,X_2 相互独立，求 $E(X_1X_2)$.

解答 如果随机变量 X 服从参数为 θ 的指数分布，那么其数学期望为

$$\begin{aligned}E(X)&=\int_{-\infty}^{+\infty}xf(x)\,\mathrm{d}x=\int_{0}^{+\infty}\frac{x}{\theta}\mathrm{e}^{-\frac{x}{\theta}}\,\mathrm{d}x\\&=-x\mathrm{e}^{-\frac{x}{\theta}}\Big|_0^{+\infty}-\int_{0}^{+\infty}\mathrm{e}^{-\frac{x}{\theta}}\,\mathrm{d}(-x)\\&=0-\theta\mathrm{e}^{-\frac{x}{\theta}}\Big|_0^{+\infty}\\&=\theta.\end{aligned}\tag{6.1}$$

类似地，X^2 的数学期望为

$$\begin{aligned}
E(X^2) &= \int_{-\infty}^{+\infty} x^2 f(x)\,\mathrm{d}x = \int_0^{+\infty} \frac{x^2}{\theta}\mathrm{e}^{-\frac{x}{\theta}}\mathrm{d}x \\
&= -x^2\mathrm{e}^{-\frac{x}{\theta}}\Big|_0^{\infty} - \int_0^{+\infty} \mathrm{e}^{-\frac{x}{\theta}}\,\mathrm{d}(-x^2) \\
&= 0 + \int_0^{+\infty} 2x\mathrm{e}^{-\frac{x}{\theta}}\,\mathrm{d}x \\
&= 2\theta \int_0^{+\infty} \frac{x}{\theta}\mathrm{e}^{-\frac{x}{\theta}}\,\mathrm{d}x,
\end{aligned} \tag{6.2}$$

根据式 (6.1) 中已经证明的结论

$$E(X^2) = 2\theta \cdot \theta = 2\theta^2. \tag{6.3}$$

根据方差的性质

$$D(X) = E(X^2) - E^2(X) = 2\theta^2 - \theta^2 = \theta^2. \tag{6.4}$$

(1) 考虑到 $X_1 \sim E\left(\frac{1}{2}\right), X_2 \sim E\left(\frac{1}{4}\right)$

$$E(X_1 + X_2) = E(X_1) + E(X_2) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}.$$

$$E\left(2X_1 - 3X_2^2\right) = 2E(X_1) - 3E(X_2^2) = 2 \cdot \frac{1}{2} - 3 \cdot 2\left(\frac{1}{4}\right)^2 = \frac{5}{8}.$$

(2) 如果 X_1, X_2 相互独立，那么

$$E(X_1X_2) = E(X_1)E(X_2) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}.$$

题目 22

设随机变量 X_1, X_2, X_3, X_4 相互独立，且有 $E(X_i) = i$，$D(X) = 5 - i$，$i = 1, 2, 3, 4$. 设 $Y = 2X_1 - X_2 + 3X_3 - \frac{1}{2}X_4$，求 $E(Y), D(Y)$.

解答 根据数学期望的性质

$$\begin{aligned}E(Y) &= E\left(2X_1 - X_2 + 3X_3 - \frac{1}{2}X_4\right)\\ &= 2E(X_1) - E(X_2) + 3E(X_3) - \frac{1}{2}E(X_4)\\ &= 2 - 2 + 3\times 3 - \frac{1}{2}\times 4\\ &= 7.\end{aligned}$$

根据方差的性质

$$\begin{aligned}D(Y) &= D\left(2X_1 - X_2 + 3X_3 - \frac{1}{2}X_4\right)\\ &= 4D(X_1) + D(X_2) + 9D(X_3) + \frac{1}{4}D(X_4)\\ &= 4\times 4 + 3 + 9\times 2 + \frac{1}{4}\times 1\\ &= \frac{149}{4}.\end{aligned}$$

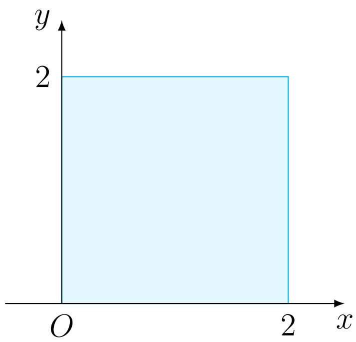
题目 32

设随机变量 (X,Y) 具有概率密度

$$f(x,y) = \begin{cases}\frac{1}{8}(x+y), & 0\leqslant x\leqslant 2, 0\leqslant y\leqslant 2,\\ 0, & \text{其他}.\end{cases}$$

求 $E(X), E(Y), \mathrm{Cov}(X,Y), \rho_{XY}, D(X+Y)$.

解答 概率密度不为零的区域如图所示



根据数学期望的定义，X 的数学期望为

$$\begin{aligned}E(X) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf(x,y)\,\mathrm{d}x\mathrm{d}y = \int_0^2 \mathrm{d}x \int_0^2 \frac{x}{8}(x+y)\,\mathrm{d}y \\ &= \int_0^2 \left(\frac{x^2}{8}y + \frac{x}{8}\frac{y^2}{2}\right)\Big|_{y=0}^{y=2} \mathrm{d}x = \int_0^2 \frac{x^2}{4} + \frac{x}{4}\,\mathrm{d}x \\ &= \left(\frac{x^3}{12} + \frac{x^2}{8}\right)\Big|_{x=0}^{x=2} = \frac{7}{6}.\end{aligned}$$

X^2 的数学期望为

$$\begin{aligned}E(X^2) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^2f(x,y)\,\mathrm{d}x\mathrm{d}y = \int_0^2 \mathrm{d}x \int_0^2 \frac{x^2}{8}(x+y)\,\mathrm{d}y \\ &= \int_0^2 \left(\frac{x^3}{8}y + \frac{x^2}{8}\frac{y^2}{2}\right)\Big|_{y=0}^{y=2} \mathrm{d}x = \int_0^2 \frac{x^3}{4} + \frac{x^2}{4}\,\mathrm{d}x \\ &= \left(\frac{x^4}{16} + \frac{x^3}{12}\right)\Big|_{x=0}^{x=2} = \frac{5}{3}.\end{aligned}$$

XY 的数学期望为

$$\begin{aligned}E(XY) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x,y)\,\mathrm{d}x\mathrm{d}y = \int_0^2 \mathrm{d}x \int_0^2 \frac{xy}{8}(x+y)\,\mathrm{d}y \\ &= \int_0^2 \left(\frac{x^2}{8}\frac{y^2}{2} + \frac{x}{8}\frac{y^3}{3}\right)\Big|_{y=0}^{y=2} \mathrm{d}x = \int_0^2 \frac{x^2}{4} + \frac{x}{3}\,\mathrm{d}x \\ &= \left(\frac{x^3}{12} + \frac{x^2}{6}\right)\Big|_{x=0}^{x=2} = \frac{4}{3}.\end{aligned}$$

同理

$$E(Y) = \frac{7}{6}, \quad E(Y^2) = \frac{5}{3}.$$

根据数学期望的性质

$$D(X) = E\left\{[X - E(X)]^2\right\} = E(X^2) - E^2(X) = \frac{5}{3} - \left(\frac{7}{6}\right)^2 = \frac{11}{36}.$$

$$D(Y) = E\left\{[Y - E(Y)]^2\right\} = E(Y^2) - E^2(Y) = \frac{5}{3} - \left(\frac{7}{6}\right)^2 = \frac{11}{36}.$$

根据协方差的定义

$$\mathrm{Cov}(X,Y) = E\{[X - E(X)][Y - E(Y)]\} = E(XY) - E(X)E(Y) = \frac{4}{3} - \frac{7}{6} \times \frac{7}{6} = -\frac{1}{36}.$$

根据相关系数的定义和方差的性质

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = \frac{-\frac{1}{36}}{\sqrt{\frac{11}{36}}\sqrt{\frac{11}{36}}} = -\frac{1}{11}.$$

根据方差的性质

$$\begin{aligned}
D(X+Y) &= E\left\{[(X+Y)-E(X+Y)]^2\right\} \\
&= E\left\{[(X-E(X))+(Y-E(Y))]^2\right\} \\
&= E\left\{[X-E(X)]^2\right\} + E\left\{[Y-E(Y)]^2\right\} + 2E\left\{[X-E(X)][Y-E(Y)]\right\} \\
&= D(X)+D(Y)+2\,\mathrm{Cov}(X,Y) \\
&= \frac{11}{36}+\frac{11}{36}+2\times\left(-\frac{1}{36}\right) \\
&= \frac{5}{9}.
\end{aligned}$$

题目 33

设随机变量 $X \sim N(\mu,\sigma^2)$，$Y \sim N(\mu,\sigma^2)$，且设 X,Y 相互独立，试求 $Z_1 = \alpha X + \beta Y$ 和 $Z_2 = \alpha X - \beta Y$ 的相关系数（其中 α,β 是不为零的常数）.

解答 根据协方差的定义和数学期望的性质，结合 X,Y 的独立性

$$\begin{aligned}
\mathrm{Cov}(Z_1,Z_2) &= E(Z_1Z_2)-E(Z_1)E(Z_2) \\
&= E(\alpha^2X^2-\beta^2Y^2)-E(\alpha X+\beta Y)E(\alpha X-\beta Y) \\
&= \alpha^2E(X^2)-\beta^2E(Y^2)-[\alpha E(X)+\beta E(Y)][\alpha E(X)-\beta E(Y)] \\
&= \alpha^2[E(X^2)-E^2(X)]-\beta^2[E(Y^2)-E^2(Y)] \\
&= (\alpha^2-\beta^2)\sigma^2.
\end{aligned}$$

根据方差的性质

$$D(Z_1) = D(\alpha X+\beta Y) = \alpha^2D(X)+\beta^2D(Y) = (\alpha^2+\beta^2)\sigma^2.$$

$$D(Z_2) = D(\alpha X-\beta Y) = \alpha^2D(X)+\beta^2D(Y) = (\alpha^2+\beta^2)\sigma^2.$$

相关系数为

$$\rho_{Z_1Z_2} = \frac{\mathrm{Cov}(Z_1,Z_2)}{\sqrt{D(Z_1)}\sqrt{D(Z_2)}} = \frac{(\alpha^2-\beta^2)\sigma^2}{\sqrt{(\alpha^2+\beta^2)\sigma^2}\sqrt{(\alpha^2+\beta^2)\sigma^2}} = \frac{\alpha^2-\beta^2}{\alpha^2+\beta^2}.$$

7 样本及抽样分布

题目 1

在总体 $N(52, 6.3^2)$ 中随机抽取一容量为 36 的样本，求样本均值落在 50.8 到 53.8 之间的概率.

解答 根据正态总体的样本均值与样本方差的分布

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) = N(52, 1.05^2).$$

从而有

$$\begin{aligned} P\left\{50.8 < \overline{X} < 53.8\right\} &= P\left\{\frac{50.8-52}{1.05} < \frac{\overline{X}-52}{1.05} < \frac{53.8-52}{1.05}\right\} \\ &= P\left\{-\frac{8}{7} < \frac{\overline{X}-52}{1.05} < \frac{12}{7}\right\} \\ &= \Phi\left(\frac{12}{7}\right) - \Phi\left(-\frac{8}{7}\right) \\ &= 0.8302. \end{aligned}$$

题目 2

在总体 $N(12, 4)$ 中随机抽一容量为 5 的样本 X_1, X_2, X_3, X_4, X_5.

(1) 求样本均值与总体均值之差的绝对值大于 1 的概率.

(2) 求概率 $P\{\max\{X_1, X_2, X_3, X_4, X_5\} > 15\}$，$P\{\min\{X_1, X_2, X_3, X_4, X_5\} < 10\}$.

解答 (1) 由于样本 X_1, X_2, X_3, X_4, X_5 来自正态总体，所以

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) = N\left(12, 0.8\right).$$

样本均值与总体均值之差的绝对值大于 1 的概率

$$\begin{aligned} P\{|\overline{X}-\mu| > 1\} &= 1 - P\{\overline{X}-\mu \leqslant 1\} \\ &= 1 - P\{-1 \leqslant \overline{X} - 12 \leqslant 1\} \\ &= 1 - P\left\{-\frac{1}{\sqrt{0.8}} \leqslant \frac{\overline{X}-12}{\sqrt{0.8}} \leqslant \frac{1}{\sqrt{0.8}}\right\} \\ &= 1 - \left[\Phi\left(\frac{1}{\sqrt{0.8}}\right) - \Phi\left(-\frac{1}{\sqrt{0.8}}\right)\right] \\ &= 0.2636. \end{aligned}$$

(2) 因 X_i 的分布函数为 $\Phi\left(\frac{x-12}{2}\right)$，故 $M=\max\{X_1,X_2,X_3,X_4,X_5\}$ 的分布函数为

$$F_M(x)=\left[\Phi\left(\frac{x-12}{2}\right)\right]^5,$$

因此

$$\begin{aligned}P\{\max\{X_1,X_2,X_3,X_4,X_5\}>15\}&=P\{M>15\}\\&=1-P\{M>15\}\\&=1-F_M(15)\\&=1-\left[\Phi\left(\frac{15-12}{2}\right)\right]^5\\&=0.2923.\end{aligned}$$

记 $N=\min\{X_1,X_2,X_3,X_4,X_5\}$，则 N 的分布函数为

$$F_N(x)=1-\left[1-\Phi\left(\frac{x-12}{2}\right)\right]^5,$$

因此

$$\begin{aligned}P\{\min\{X_1,X_2,X_3,X_4,X_5\}<10\}&=P\{N<10\}\\&=1-\left[1-\Phi\left(\frac{10-12}{2}\right)\right]^5\\&=1-[1-\Phi(-1)]^5\\&=1-\Phi(1)^5\\&=0.5785.\end{aligned}$$

题目 4

(1) 设样本 $X_1,X_2,\cdots,X_6$ 来自总体 $N(0,1)$，$Y=(X_1+X_2+X_3)^2+(X_4+X_5+X_6)^2$，试确定常数 C 使 CY 服从 χ^2 分布.

(2) 设样本 $X_1,X_2,\cdots,X_5$ 来自总体 $N(0,1)$，$Y=\dfrac{C(X_1+X_2)}{\sqrt{X_3^2+X_4^2+X_5^2}}$，试确定常数 C 使 Y 服从 t 分布.

解答 (1) 因为样本来自正态总体 $N(0,1)$，所以

$$X_1+X_2+X_3\sim N(0,3),\quad X_4+X_5+X_6\sim N(0,3),$$

将二者都化为标准正态分布

$$\frac{X_1+X_2+X_3}{\sqrt{3}} \sim N(0,1), \quad \frac{X_4+X_5+X_6}{\sqrt{3}} \sim N(0,1),$$

由于样本间彼此独立，所以根据 χ^2 分布的定义，有

$$\left(\frac{X_1+X_2+X_3}{\sqrt{3}}\right)^2 + \left(\frac{X_4+X_5+X_6}{\sqrt{3}}\right)^2 \sim \chi^2(2),$$

与 Y 对比系数，得到

$$C = \frac{1}{3}.$$

(2) 考虑到 $X_1, X_2, \cdots, X_n$ 是总体 $N(0,1)$ 的样本，所以

$$X_1+X_2 \sim N(0,2), \quad \frac{X_1+X_2}{\sqrt{2}} \sim N(0,1),$$

另一方面

$$X_3^2+X_4^2+X_5^2 \sim \chi^2(3),$$

结合 $\dfrac{X_1+X_2}{\sqrt{2}}$ 与 $X_3^2+X_4^2+X_5^2$ 的独立性

$$\frac{\dfrac{X_1+X_2}{\sqrt{2}}}{\sqrt{(X_3^2+X_4^2+X_5^2)/3}} = \sqrt{\frac{3}{2}}\frac{X_1+X_2}{\sqrt{X_3^2+X_4^2+X_5^2}} \sim t(3),$$

综上

$$C = \sqrt{\frac{3}{2}}.$$

题目 6

设总体 $X \sim b(1,p)$，$X_1, X_2, \cdots X_n$ 是来自 X 的样本.

(1) 求 $(X_1, X_2, \cdots, X_n)$ 的分布律.

(2) 求 $\sum\limits_{i=1}^{n} X_i$ 的分布律.

(3) 求 $E(\overline{X}), D(\overline{X}), E(S^2)$.

解答 (1) 由于样本 $X_i \sim b(1,p)$，所以

$$P\{X_i = x_i\} = p^{x_i}(1-p)^{1-x_i}(x_i = 0, 1).$$

又因为样本间彼此独立，所以

$$\begin{aligned}P\{X_1,X_2,\cdots,X_n\} &= P\{X_1=x_1\}\cdot P\{X_2=x_2\}\cdots P\{X_n=x_n\}\\ &= p^{x_1}(1-p)^{1-x_1}\cdot p^{x_2}(1-p)^{1-x_2}\cdots p^{x_n}(1-p)^{1-x_n}\\ &= \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}\\ &= p^{\sum_{i=1}^{n} x_i}(1-p)^{\sum_{i=1}^{n} 1-x_i}.\end{aligned}$$

(2) 由于样本 $X_i \sim b(1,p)$ 且彼此相互独立，这显然满足二项分布的定义，所以

$$\sum_{i=1}^{n} X_i \sim b(n,p)$$

其分布律为

$$P\left\{\sum_{i=1}^{n} X_i = k\right\} = C_n^k p^k (1-p)^{n-k}$$

(3) 由于总体 $X \sim b(1,p)$，$E(X)=p$，$D(X)=p(1-p)$，所以

$$E(\overline{X}) = \mu = p,$$

$$D(\overline{X}) = \frac{\sigma^2}{n} = \frac{p(1-p)}{n},$$

$$E(S^2) = \sigma^2 = p(1-p).$$

8 参数估计

题目 4

(1) 设总体 X 具有分布律

X	1	2	3
p_k	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

其中 $\theta(0<\theta<1)$ 为未知参数. 已知取得了样本值 $x_1=1,x_2=2,x_3=1$. 试求 θ 的矩估计值和最大似然估计值.

解答 (1) 对于矩估计值

$$\mu_1=E(X)=\sum_{k=1}^{\infty}x_kp_k=\theta^2+4\theta(1-\theta)+3(1-\theta)^2,$$

解得

$$\theta=\frac{1}{2}(3-\mu_1),$$

所以 θ 的估计量为

$$\hat{\theta}=\frac{1}{2}(3-\overline{x})=\frac{5}{6}.$$

对于最大似然估计值

$$L(\theta)=\prod_{i=1}^{3}P\{X_i=x_i\}=\theta^2\cdot 2\theta(1-\theta)\cdot\theta^2=2\theta^5(1-\theta),$$

两边取对数

$$\ln L(\theta)=\ln 2+5\ln\theta+\ln(1-\theta),$$

两边求导，令导函数为零

$$\ln L'(\theta)=\frac{5}{\theta}-\frac{1}{1-\theta}=0,$$

所以 θ 的最大似然估计值为

$$\hat{\theta}=\frac{5}{6}.$$

题目 5(2)

设某种电子器件的寿命（以 h 计）T 服从双参数的指数分布，其概率密度为

$$f(t)=\begin{cases}\dfrac{1}{\theta}\mathrm{e}^{-\frac{t-c}{\theta}}, & t\geqslant c,\\ 0, & \text{其他}.\end{cases}$$

其中 $c,\theta(c,\theta>0)$ 为未知参数. 自一批这种器件中随机地取 n 件进行寿命试验. 设它们的失效时间依次为 $x_1\leqslant x_2\leqslant\cdots\leqslant x_n$.

(1) 求 θ 与 c 的最大似然估计值.

(2) 求 θ 与 c 的矩估计值.

解答 (2) 对于矩估计值

$$\mu_l=\int_{-\infty}^{+\infty}t^l f(t)\,\mathrm{d}t=\int_c^{+\infty}\frac{t^l}{\theta}\mathrm{e}^{-\frac{t-c}{\theta}}\,\mathrm{d}t,$$

令 $x=\dfrac{t-c}{\theta}$ 有

$$\mu_1=\int_0^{+\infty}\frac{\theta x+c}{\theta}\mathrm{e}^{-x}\,\mathrm{d}(\theta x+c)=\int_0^{+\infty}(\theta x+c)\mathrm{e}^{-x}\,\mathrm{d}x=\theta+c,$$

$$\mu_2=\int_0^{+\infty}\frac{(\theta x+c)^2}{\theta}\mathrm{e}^{-x}\,\mathrm{d}(\theta x+c)=\int_0^{+\infty}\left(\theta^2x^2+2\theta cx+c^2\right)\mathrm{e}^{-x}\,\mathrm{d}x=2\theta^2+2\theta c+c^2.$$

反解得到

$$\begin{cases}\theta=\sqrt{\mu_2-\mu_1^2}\\ c=\mu_1-\sqrt{\mu_2-\mu_1^2}\end{cases}$$

带入样本数据，有

$$\hat{\theta}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2}$$

$$\hat{c}=\overline{X}-\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2}$$

题目 8

(1) 设 $X_1, X_2, \cdots, X_n$ 是来自概率密度为

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

的总体的样本，θ 未知，求 $U=\mathrm{e}^{-\frac{1}{\theta}}$ 的最大似然估计值.

(2) 设 $X_1, X_2, \cdots, X_n$ 是来自正态总体 $N(\mu, 1)$ 的样本. μ 未知，求 $\theta=P\{X\geqslant 2\}$ 的最大似然估计值.

解答 (1) 考虑到样本间的独立性

$$P\{X_1, X_2, \cdots, X_n\}=P\{X_1=x_1\}\cdot P\{X_2=x_2\}\cdots P\{X_n=x_n\},$$

取似然函数为

$$L(\theta)=\prod_{i=1}^{n}\theta x_i^{\theta-1}=\theta^n\prod_{i=1}^{n}x_i^{\theta-1}=\theta^n\left(\prod_{i=1}^{n}x_i\right)^{\theta-1},$$

两边取对数

$$\ln L(\theta)=n\ln\theta+(\theta-1)\ln\left(\prod_{i=1}^{n}x_i\right)=n\ln\theta+(\theta-1)\sum_{i=1}^{n}\ln x_i,$$

两边求导并令导函数为零

$$\ln L'(\theta)=\frac{n}{\theta}+\sum_{i=1}^{n}\ln x_i=0,$$

所以 θ 的最大似然估计值

$$\hat{\theta}=-\frac{n}{\sum\limits_{i=1}^{n}\ln x_i},$$

考虑到 $U=\mathrm{e}^{-\frac{1}{\theta}}$ 具有单调反函数，所以 U 的最大似然估计值为

$$\widehat{U}=\mathrm{e}^{-\frac{1}{\hat{\theta}}}.$$

题目 11

设总体 X 的概率密度为

$$f(x;\theta)=\begin{cases}\frac{1}{\theta}x^{\frac{1-\theta}{\theta}}, & 0<x<1,\\ 0, & \text{其他}.\end{cases}\quad 0<\theta<+\infty.$$

$X_1,X_2,\cdots,X_n$ 是来自总体 X 的样本.

(1) 验证 θ 的最大似然估计量是 $\hat{\theta}=-\frac{1}{n}\sum\limits_{i=1}^{n}\ln X_i$.

(2) 证明 $\hat{\theta}$ 是 θ 的无偏估计量.

解答 (1) 似然函数为

$$L(\theta)=\prod_{i=1}^{n}\frac{1}{\theta}x_i^{\frac{1-\theta}{\theta}}=\frac{1}{\theta^n}\prod_{i=1}^{n}x_i^{\frac{1-\theta}{\theta}}=\frac{1}{\theta^n}\left(\prod_{i=1}^{n}x_i\right)^{\frac{1-\theta}{\theta}},$$

两边取对数

$$\ln L(\theta)=-n\ln\theta+\frac{1-\theta}{\theta}\ln\prod_{i=1}x_i,$$

两边求导并令导函数为零

$$\ln L'(\theta)=-\frac{n}{\theta}-\frac{1}{\theta^2}\sum_{i=1}^{n}\ln x_i=0,$$

找到 θ 的最大似然估计量

$$\hat{\theta}=-\frac{1}{n}\sum_{i=1}^{n}\ln x_i.$$

(2) 估计量 $\hat{\theta}$ 的数学期望为

$$E(\hat{\theta})=E\left(-\frac{1}{n}\sum_{i=1}^{n}\ln x_i\right)=-\frac{1}{n}E\left(\sum_{i=1}^{n}\ln x_i\right)=-\frac{1}{n}\sum_{i=1}^{n}E(\ln x_i).$$

考虑到

$$\begin{aligned}E(\ln x)&=\int_0^1\ln x\cdot\frac{1}{\theta}x^{\frac{1-\theta}{\theta}}\,\mathrm{d}x=\int_0^1\ln x\,\mathrm{d}\left(x^{\frac{1}{\theta}}\right)\\&=x^{\frac{1}{\theta}}\ln x\Big|_0^1-\int_0^1 x^{\frac{1-\theta}{\theta}}\,\mathrm{d}(\ln x)\\&=0-\theta x^{\frac{1}{\theta}}\Big|_0^1=-\theta.\end{aligned}$$

于是

$$E(\hat{\theta}) = -\frac{1}{n}E\left(\sum_{i=1}^{n}\ln x_i\right) = -\frac{1}{n}\sum_{i=1}^{n}E(\ln x_i) = -\frac{1}{n}\cdot(-n\theta) = \theta.$$

所以 $\hat{\theta}$ 是 θ 的无偏估计量.

题目 12

设 X_1, X_2, X_3, X_4 是来自均值为 θ 的指数分布总体的样本，其中 θ 未知. 设有估计量

$$T_1 = \frac{1}{6}(X_1 + X_2) + \frac{1}{3}(X_3 + X_4),$$

$$T_2 = \frac{1}{5}(X_1 + 2X_2 + 3X_4 + 4X_4),$$

$$T_3 = \frac{1}{4}(X_1 + X_2 + X_3 + X_4).$$

(1) 指出 T_1, T_2, T_3 中哪几个是 θ 的无偏估计量.

(2) 在上述 θ 的无偏估计中指出哪一个较为有效.

解答 (1) 根据无偏估计的定义，上述估计量的数学期望为

$$E(T_1) = \frac{1}{6}E(X_1 + X_2) + \frac{1}{3}E(X_3 + X_4) = \frac{\theta}{3} + \frac{2\theta}{3} = \theta,$$

$$E(T_2) = \frac{1}{5}[E(X_1) + 2E(X_2) + 3E(X_3) + 4E(X_4)] = \frac{1}{5}(\theta + 2\theta + 3\theta + 4\theta) = 2\theta,$$

$$E(T_3) = \frac{1}{4}[E(X_1) + E(X_2) + E(X_3) + E(X_4)] = \frac{1}{4}(\theta + \theta + \theta + \theta) = \theta.$$

所以只有 T_1 和 T_3 是 θ 的无偏估计量.

(2) 比较两个无偏量估计有效性的依据是方差大小，其中

$$\begin{aligned} D(T_1) &= D\left[\frac{1}{6}(X_1 + X_2) + \frac{1}{3}(X_3 + X_4)\right] \\ &= \frac{1}{36}[D(X_1) + D(X_2)] + \frac{1}{9}[D(X_3) + D(X_4)] \\ &= \frac{1}{36}(\theta^2 + \theta^2) + \frac{4}{36}(\theta^2 + \theta^2) \\ &= \frac{10}{36}\theta^2, \end{aligned}$$

$$
\begin{aligned}
D(T_3) &= D\left[\frac{1}{4}(X_1+X_2+X_3+X_4)\right] \\
&= \frac{1}{16}[D(X_1)+D(X_2)+D(X_3)+D(X_4)] \\
&= \frac{1}{16}[\theta^2+\theta^2+\theta^2+\theta^2] \\
&= \frac{9}{36}\theta^2.
\end{aligned}
$$

由于

$$D(T_1) > D(T_3).$$

这说明估计量 T_3 较 T_1 有效.

9 假设检验

例题 1（双边检验）

设一车床生产的纽扣直径服从正态分布. 根据以往的经验，当车床工作正常时，生产的纽扣的平均直径 $\mu_0 = 26\,\mathrm{mm}$，方差 $\sigma^2 = 5.2\,\mathrm{mm}^2$. 某天开工一段时间后，为检验车床生产是否正常，从刚生产的纽扣中随机抽检了 100 颗，测得其观测值 $(x_1, x_2, \cdots, x_{100})$ 的样本均值 $\bar{x} = 26.56\,\mathrm{mm}$. 假定所生产的纽扣的精度保持不变，试分别在显著性水平 $\alpha_1 = 0.05, \alpha_2 = 0.01$ 下检验这天改车床的生产是否正常.

解答　由于方差 σ^2 已知，考虑采用 z 检验法. **第一步**作统计假设

$$H_0 : \mu = \mu_0 = 26, \quad H_1 : \mu \neq 26.$$

第二步选取检验统计量

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}.$$

第三步选取拒绝域

$$C = \{|z| \geqslant z_{\frac{\alpha}{2}}\} = \left\{\left|\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right| \geqslant z_{\frac{\alpha}{2}}\right\}.$$

当 $\alpha_1 = 0.05$ 时

$$C_1 = \{|z| \geqslant z_{0.025}\} = \{|z| \geqslant 1.96\}.$$

当 $\alpha_2 = 0.01$ 时

$$C_1 = \{|z| \geqslant z_{0.005}\} = \{|z| \geqslant 2.58\}.$$

第四步代入实测值

$$z = \frac{26.56 - 26}{\sqrt{5.2}/\sqrt{100}} = 2.4558.$$

$$z \in C_1, \quad z \notin C_2.$$

故 $\alpha_1 = 0.05$ 的情况下拒绝 H_0，认为该车床不正常，$\alpha_2 = 0.01$ 的情况下接受 H_0，认为该车床生产正常.

例题 2（左边检验）

有一批枪弹，出厂时的初速度（单位：m/s）服从正态分布 $N(950, 10^2)$. 经过较长时间储存后，现取出 9 发枪弹试射，测得其初速度如下：

$$914 \quad 920 \quad 910 \quad 934 \quad 953 \quad 945 \quad 912 \quad 924 \quad 940$$

假定 $\sigma_0^2 = 10^2$ 不变，试在显著性水平 $\alpha = 0.05$ 下检验这批枪弹的初速度是否变小.

解答　由于方差 σ^2 已知，考虑采用 z 检验法. **第一步**作统计假设

$$H_0 : \mu = \mu_0 = 950, \quad H_1 : \mu < \mu_0 = 950.$$

第二步选取检验统计量

$$z = \frac{\bar{x} - \mu_0}{\sigma_0/\sqrt{n}}.$$

第三步选取拒绝域

$$C = \{z \leqslant -z_\alpha\} = \left\{\left|\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right| \leqslant -z_\alpha\right\} = \{z \leqslant -z_{0.05}\} = \{z \leqslant -1.645\}.$$

第四步代入实测值

$$z = \frac{928 - 985}{10/\sqrt{9}} = -6.6 \in C.$$

故 $\alpha = 0.05$ 的情况下拒绝 H_0，认为枪弹的初速度变小.

例题 3（左边检验）

要求一种元件平均使用寿命不得低于 1000h，生产者从一批这种元件中随机抽取 25 件，测得其寿命的平均值为 950h. 已知该种元件寿命服从标准差为 $\sigma = 100\,\text{h}$ 的正态分布. 试在显著性水平 $\alpha = 0.05$ 下判断这批元件是否合格.

解答　由于方差 σ^2 已知，考虑采用 z 检验法. **第一步**作统计假设

$$H_0 : \mu = \mu_0 = 1000, \quad H_1 : \mu < \mu_0 = 1000.$$

第二步选取检验统计量

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}.$$

第三步选取拒绝域

$$C = \{z \leqslant -z_\alpha\} = \left\{\left|\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right| \leqslant -z_\alpha\right\} = \{z \leqslant -z_{0.05}\} = \{z \leqslant -1.645\}.$$

第四步代入实测值

$$z = \frac{950 - 1000}{100/\sqrt{25}} = -2.5 \in C.$$

故 $\alpha = 0.05$ 的情况下拒绝 H_0，认为这批元件不合格.

例题 4（右边检验）

公司从生产商购买牛奶. 公司怀疑生产商在牛奶中掺水以牟利. 通过测定牛奶的冰点，可以检验出牛奶是否掺水. 天然牛奶的冰点温度近似服从正态分布，均值 $\mu_0 = -0.545°C$，标注差 $\sigma = 0.008°C$. 牛奶掺水可使冰点温度升高而接近于水的冰点温度（0°C）. 测得生产商提交的 5 批牛奶的冰点温度，其均值为 $\bar{x} = -0.535°C$，问是否可以认为生产商在牛奶中掺了水？取 $\alpha = 0.05$.

解答　由于方差 σ^2 已知，考虑采用 z 检验法. **第一步**作统计假设

$$H_0 : \mu = \mu_0 = -0.545, \quad H_1 : \mu > \mu_0 = 0.545.$$

第二步选取检验统计量

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}.$$

第三步选取拒绝域

$$C = \{z \geqslant z_\alpha\} = \left\{\left|\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right| \geqslant z_\alpha\right\} = \{z \geqslant z_{0.05}\} = \{z \geqslant -1.645\}.$$

第四步代入实测值

$$z = \frac{-0.535 + 0.545}{0.008/\sqrt{5}} = -2.7951 \in C.$$

故 $\alpha = 0.05$ 的情况下拒绝 H_0，认为牛奶掺水.

例题 1（双边检验）

某批矿砂的 5 个样品中的镍含量，经测定为（%）

$$3.25 \quad 3.27 \quad 3.24 \quad 3.26 \quad 3.24$$

设测定值总体服从正态分布，但参数均未知. 问在 $\alpha = 0.01$ 下能否接受假设：这批矿砂的镍含量的均值为 3.25.

解答　由于方差 σ^2 未知，考虑采用 t 检验法. **第一步**作统计假设

$$H_0 : \mu = \mu_0 = 3.25, \quad H_1 : \mu \neq \mu_0 = 3.25.$$

第二步选取检验统计量

$$t = \frac{\bar{x} - \mu_0}{S/\sqrt{n}}.$$

第三步选取拒绝域

$$C=\{|t|\geqslant t_{\frac{\alpha}{2}}(n-1)\}=\left\{\left|\frac{\bar{x}-\mu_0}{S/\sqrt{n}}\right|\geqslant t_{\frac{\alpha}{2}}(n-1)\right\}=\{t\geqslant t_{0.005}(4)\}=\{t\geqslant 4.60413\}.$$

第四步代入实测值 $\bar{x}=3.252$，$S=0.013$

$$t=\frac{-0.535+0.545}{0.013/\sqrt{5}}=0.344\notin C.$$

故 $\alpha=0.01$ 的情况下接受 H_0，认为这批砂矿的镍含量的均值为 3.25.

例题 2（左边检验）

按规定，100g 罐头番茄汁中的平均维生素 C 含量不得少于 21mg/g. 现从工厂的产品中抽取 17 个罐头，其 100g 番茄汁中，测得维生素 C 含量（mg/g）记录如下：

16 25 21 20 23 21 19 15 13 23 17 20 29 18 22 16 22

设维生素含量服从正态分布 $N(\mu,\sigma^2),\mu,\sigma^2$ 均未知，问这批罐头是否符合要求（取显著性水平 $\alpha=0.05$）.

解答　由于方差 σ^2 未知，考虑采用 t 检验法. **第一步**作统计假设

$$H_0:\mu\geqslant\mu_0=21,\quad H_1:\mu<\mu_0=21.$$

第二步选取检验统计量

$$t=\frac{\bar{x}-\mu_0}{S/\sqrt{n}}.$$

第三步选取拒绝域

$$C=\{t<-t_\alpha(n-1)\}=\left\{\left|\frac{\bar{x}-\mu_0}{S/\sqrt{n}}\right|<-t_\alpha(n-1)\right\}=\{t<-t_{0.05}(16)\}=\{t<-1.7459\}.$$

第四步代入实测值 $\bar{x}=20$，$S=3.984$

$$t=\frac{20-21}{3.984/\sqrt{17}}=-1.035\notin C.$$

故 $\alpha=0.05$ 的情况下接受 H_0，认为这批罐头符合要求.

例题 3（右边检验）

下面列出的是某工厂随机选取的 20 只部件的装配时间（min）

9.8	10.4	10.6	9.6	9.7	9.9	10.9	11.1	9.6	10.2
10.3	9.6	9.9	11.2	10.6	9.8	10.5	10.1	10.5	9.7

设装配时间的总体服从正态分布 $N(\mu,\sigma^2),\mu,\sigma^2$ 均未知. 是否可以认为装配时间的均值显著大于 10（取 $\alpha=0.05$）?

解答　由于方差 σ^2 未知，考虑采用 t 检验法. **第一步**作统计假设

$$H_0:\mu\leqslant\mu_0=10,\quad H_1:\mu>\mu_0=10.$$

第二步选取检验统计量

$$t=\frac{\bar{x}-\mu_0}{S/\sqrt{n}}.$$

第三步选取拒绝域

$$C=\{t\geqslant t_\alpha(n-1)\}=\left\{\left|\frac{\bar{x}-\mu_0}{S/\sqrt{n}}\right|\geqslant t_\alpha(n-1)\right\}=\{t\geqslant t_{0.05}(19)\}=\{t\geqslant 1.7291\}.$$

第四步代入实测值 $\bar{x}=10.2$，$S=0.5099$

$$t=\frac{10.2-10}{0.5099/\sqrt{20}}=1.754\in C.$$

故 $\alpha=0.05$ 的情况下拒绝 H_0，认为装配时间的均值显著大于 10.

例题 1（双边检验）

某厂生产的某种型号的电池，其寿命（以 h 计）长期以来服从方差 $\sigma^2=5000$ 的正态分布，现有一批这种电池，从它的生产情况来看，寿命的波动性有所改变. 现随机取 26 只电池，测出其寿命的样本方差 $S^2=9200$. 问根据这一数据能否推断这批电池的寿命的波动性较以往的有显著的变化（取 $\alpha=0.02$）?

解答　由于均值 μ 未知，考虑采用 χ^2 检验法. **第一步**作统计假设

$$H_0:\sigma^2=\sigma_0^2=5000,\quad H_1:\sigma^2\neq\sigma_0^2=5000.$$

第二步选取检验统计量

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}.$$

第三步选取拒绝域

$$\begin{aligned} C &= \left\{\chi^2 \leqslant \chi^2_{1-\frac{\alpha}{2}}(n-1)\right\} \cup \left\{\chi^2 \geqslant \chi^2_{\frac{\alpha}{2}}(n-1)\right\} \\ &= \{\chi^2 \leqslant \chi^2_{0.99}(25)\} \cup \{\chi^2 \geqslant \chi^2_{0.01}(25)\} \\ &= \{\chi^2 \leqslant 11.524\} \cup \{\chi^2 \geqslant 44.314\} \end{aligned}$$

第四步代入实测值

$$\chi^2 = \frac{(26-1)\times 9200}{5000} = 46 \in C.$$

故 $\alpha = 0.05$ 的情况下拒绝 H_0，认为这批电池的寿命的波动性较以往的有显著的变化.

例题 2（左边检验）

一种混杂的小麦品种，株高的标准差为 $\sigma_0 = 14\,\text{cm}$，经提纯后随机抽取 10 株，它们的株高（以 cm 计）为

90　105　101　95　100　100　101　105　93　97

考察提纯后群体是否比原群体更整齐？取显著性水平 $\alpha = 0.01$，并设小麦株高服从 $N(\mu, \sigma^2)$.

解答　由于均值 μ 未知，考虑采用 χ^2 检验法. **第一步**作统计假设

$$H_0 : \sigma^2 \geqslant \sigma_0^2 = 14^2, \quad H_1 : \sigma^2 < \sigma_0^2 = 14^2.$$

第二步选取检验统计量

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}.$$

第三步选取拒绝域

$$C = \{\chi^2 \leqslant \chi^2_{1-\alpha}(n-1)\} = \{\chi^2 \leqslant \chi^2_{0.99}(9)\} = \{\chi^2 \leqslant 2.088\}$$

第四步代入实测值 $S^2 = 24.2333$

$$\chi^2 = \frac{(10-1)\times 24.23333}{14^2} = 1.1127 \in C.$$

故 $\alpha = 0.05$ 的情况下拒绝 H_0，认为提纯后群体比原群体更整齐.

例题 3（右边检验）

某种导线，要求其电阻的标准差不得超过 0.005Ω，今在生产的一批线中取样品 9 根，测得 $S = 0.007\Omega$，设总体为正态分布，参数均未知. 问在显性水平 $\alpha = 0.05$ 下能否认为这批导线的标准差显著地偏大？

解答　由于均值 μ 未知，考虑采用 χ^2 检验法. **第一步**作统计假设

$$H_0 : \sigma^2 \leqslant \sigma_0^2 = 0.005^2, \quad H_1 : \sigma^2 > \sigma_0^2 = 14^2.$$

第二步选取检验统计量

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}.$$

第三步选取拒绝域

$$C = \left\{\chi^2 \geqslant \chi_\alpha^2(n-1)\right\} = \left\{\chi^2 \geqslant \chi_{0.05}^2(8)\right\} = \left\{\chi^2 \geqslant 15.507\right\}$$

第四步代入实测值 $S = 0.007$

$$\chi^2 = \frac{(9-1) \times 0.007^2}{0.005^2} = 15.68 \in C.$$

故 $\alpha = 0.05$ 的情况下拒绝 H_0，认为这批导线的标准差显著地偏大.