

# 概率论与数理统计学习笔记

中国科学技术大学

彭煜峰

2024 年 1 月 14 日

# 前言

期末前夕，开始恶补概率论……主要参考 B 站孔祥仁老师的概率论与数理统计课程。

2024 年 1 月 14 日

中国科学技术大学

# 目录

<b>第一章 概率论的基本概念</b>	<b>1</b>
1.1 随机试验	1
1.1.1 名词	1
1.1.2 随机试验	1
1.2 样本空间与随机事件	1
1.2.1 样本空间	1
1.2.2 随机事件	2
1.3 事件间的关系及运算	2
1.3.1 事件关系	2
1.3.2 事件的运算律	3
1.4 频率与概率	4
1.4.1 频率	4
1.4.2 概率	4
1.5 古典概型	8
1.6 条件概率	8
1.7 乘法定理	10
1.8 全概率公式	12
1.8.1 完备事件群	12
1.8.2 全概率公式	12
1.9 贝叶斯公式	14
1.10 事件独立性	16
<b>第二章 随机变量及其分布</b>	<b>19</b>

2.1	随机变量	19
2.2	离散型随机变量及其分布	20
2.2.1	离散型随机变量	20
2.2.2	分布率及其性质	20
2.2.3	重要分布	21
2.3	随机变量的分布函数	25
2.4	连续型随机变量及其概率密度	26
2.4.1	连续型随机变量及其概率密度函数	26
2.4.2	重要的连续型随机变量分布	28
2.5	随机变量函数的分布	35
2.5.1	分布函数法	35
2.5.2	定理法	37
<b>第三章</b>	<b>多维随机变量及其分布</b>	<b>39</b>
3.1	二维随机变量	39
3.1.1	二维随机变量及其联合分布函数	39
3.1.2	离散型二维随机变量及联合分布律	40
3.1.3	连续型二维随机变量及联合概率密度	41
3.2	边缘分布	44
3.2.1	边缘分布函数	44
3.2.2	离散型随机变量的边缘分布率	44
3.2.3	连续型随机变量边缘概率密度	45
3.2.4	常用二维分布	46
3.3	条件分布	47
3.3.1	离散型二维随机变量的条件分布律	47
3.3.2	连续型二维随机变量的条件概率密度和条件分布函数	48
3.4	相互独立的随机变量	51
3.5	两个随机变量的函数的分布	55
3.5.1	两个随机变量的线性组合	55
3.5.2	非线性	59
3.5.3	取最值	60

<b>第四章</b>	<b>随机变量的数字特征</b>	<b>62</b>
4.1	数学期望 . . . . .	62
4.1.1	离散型随机变量的数学期望 . . . . .	62
4.1.2	连续型随机变量的数学期望 . . . . .	63
4.1.3	随机变量的函数的数学期望 . . . . .	64
4.1.4	数学期望的性质 . . . . .	66
4.1.5	服从特定分布的数学期望 . . . . .	67
4.2	方差 . . . . .	68
4.2.1	方差的性质 . . . . .	71
4.2.2	切比雪夫不等式 . . . . .	74
4.3	协方差与相关系数 . . . . .	75
4.4	矩以及协方差矩阵 . . . . .	77
4.4.1	矩 . . . . .	77
4.4.2	协方差矩阵 . . . . .	77
4.4.3	多维正态随机变量的性质 . . . . .	77
4.5	大数定理与中心极限定理 . . . . .	78
4.5.1	两个不等式 . . . . .	78
4.5.2	两个收敛 . . . . .	80
4.5.3	三个大数定律 . . . . .	81
4.5.4	三个中心极限定理 . . . . .	82
<b>第五章</b>	<b>统计学基本概念</b>	<b>85</b>
5.1	随机样本 . . . . .	85
5.2	抽样分布 . . . . .	85
5.2.1	统计量 . . . . .	85
5.2.2	统计量的分布 . . . . .	86
5.3	正态总体样本均值与样本方差的分布 . . . . .	87
5.3.1	四个等式 . . . . .	87
5.3.2	四个定理 . . . . .	88

第六章 参数点估计	90
6.1 概念 . . . . .	90
6.2 点估计 . . . . .	90

# 第一章 概率论的基本概念

## 1.1 随机试验

### 1.1.1 名词

定义 1.1.1. 确定性现象：结果呈现确定性的现象.

定义 1.1.2. 随机现象：在个别试验中呈现不确定性，但是在大量重复试验中，表现出统计规律性的现象.

### 1.1.2 随机试验

定义 1.1.3. 对随机现象的实现或对其观察称为随机试验，记为  $E$ .

例 1.1.1. 投币观察向上的面. (实现、观察)

例 1.1.2. 记录每个星期一的天气. (观察)

### 特点

1. 相同条件可重复.
2. 试验结果明确可知，且一般不止一个.
3. 试验前不能确定那个结果出现.

## 1.2 样本空间与随机事件

### 1.2.1 样本空间

定义 1.2.1. 将  $E$  的所有可能结果和组成的集合称为  $E$  的样本空间  $\sim \Omega$ .

**定义 1.2.2.** 样本空间  $\Omega$  中的元素即为样本点.

**例 1.2.1.** 写出下列试验的样本空间:

$E_1$  抛一枚硬币, 观察正面  $H$ , 反面  $T$  出现的情况.

答:  $\Omega = \{H, T\}$

$E_2$  将一枚硬币抛掷三次, 观察正面  $H$ , 反面  $T$  出现的情况.

答:  $\Omega = \{HHH, TTT, HHT, HTH, THH, HTT, THT, TTH\}$ .

$E_3$  记录某一地区一昼夜的最高气温和最低气温.

答: 记最高气温为  $x$ , 最低气温为  $y$ , 该地区历史最低气温 (不可能更低) 为  $T_1$ , 历史最高气温 (不可能更高) 为  $T_2$ , 则

$$\Omega = \{(x, y) | T_1 \leq y \leq x \leq T_2\}.$$

## 1.2.2 随机事件

**定义 1.2.3.** 称  $E$  的样本空间  $\Omega$  的子集为  $E$  的随机事件.

**定义 1.2.4.** 在一次试验中, 该子集的一个样本点出现, 称该事件发生.

**例 1.2.2.** 投一枚骰子, 将红色的点向上称作事件  $A$ . 在一次试验中 1 点向上, 请问事件  $A$  是否发生?

**解.** 样本空间  $\Omega = \{1, 2, 3, 4, 5, 6\}$ , 事件  $A = \{1, 4\}$ . 可知  $A \subseteq S$  且  $A$  中样本点  $\{1\}$  在该试验中出现, 因此事件  $A$  发生. □

**定义 1.2.5.** 由一个样本点组成的单点集叫做基本事件.

**定义 1.2.6.** 样本空间  $\Omega$  本身为一个必然事件.

**定义 1.2.7.** 事件集合中没有元素, 即为  $\emptyset$ , 称为不可能事件.

## 1.3 事件间的关系及运算

### 1.3.1 事件关系

包含关系

**定义 1.3.1.**  $A \subset B$  表示事件  $A$  包含于事件  $B$ , 如果事件  $A$  发生, 则事件  $B$  一定发生.



注. 若  $A \subset B$  且  $B \subset A$  则  $A = B$ . 这一性质常用于证明集合相等.

### 和事件 (并事件)

定义 1.3.2.  $A \cup B$  表示  $A$  事件与  $B$  事件至少发生一个, 也记作  $A + B$ .

### 积事件 (交事件)

定义 1.3.3.  $A \cap B$  表示事件  $A$  和事件  $B$  同时发生, 也记作  $AB$ .

### 差事件

定义 1.3.4. 事件  $A$  发生且事件  $B$  不发生, 记作  $A - B$ ; 事件  $B$  发生且事件  $A$  不发生, 记作  $B - A$ .

### 互斥事件 (互不相容)

定义 1.3.5. 事件  $A$  与事件  $B$ , 不能同时发生, 记作  $A \cap B = \emptyset$ .

### 逆事件 (对立事件)

定义 1.3.6. 事件  $A$  和  $B$  有且仅有一个发生, 记作  $A \cap B = \emptyset$  且  $A \cup B = \Omega$ .

注. 由此可知, 对立事件一定是互斥事件而互斥事件不一定是对立事件.

## 1.3.2 事件的运算律

定理 1.3.1. 交换律  $A \cup B = B \cup A$      $AB = BA$ .

定理 1.3.2. 结合律  $A \cup (B \cup C) = (A \cup B) \cup C$      $A(BC) = (AB)C$ .

定理 1.3.3. 分配律  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$      $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ .

定理 1.3.4. De Morgan's Law  $\overline{A \cup B} = \bar{A} \cap \bar{B}$      $\overline{A \cap B} = \bar{A} \cup \bar{B}$ .

## 1.4 频率与概率

### 1.4.1 频率

定义 1.4.1. 事件发生的频数 ( $n_A$ ) 与试验总次数 ( $n$ ) 之间的比值:

$$f_n(A) = \frac{n_A}{n}.$$

性质

1.  $0 \leq f_n(A) \leq 1$ .
2.  $f_n(\Omega) = 1$ .
3. 若  $A_1, A_2, A_3 \dots A_k$  为两两互不相容事件, 则

$$f_n(A_1 \cup A_2 \cup A_3 \cup \dots \cup A_k) = \sum_{i=1}^k f_n(A_i).$$

### 1.4.2 概率

用于衡量事件  $A$  发生的可能性的的大小, 一般用  $P$  来表示。

基本性质

1. 对于  $\forall A$ , 有  $P(A) \geq 0$ .
2.  $P(\Omega) = 1$ .
3. 可列可加性: 若  $A_1, A_2, A_3 \dots$  为两两互不相容事件, 则

$$P(A_1 \cup A_2 \cup A_3 \cup \dots) = \sum_{i=1}^{\infty} P(A_i).$$

重要性质

1.  $P(\emptyset) = 0$ .
2. 有限可加性: 若  $A_1, A_2, A_3 \dots A_n$  为两两互不相容事件, 则

$$P(A_1 \cup A_2 \cup A_3 \cup \dots \cup A_n) = \sum_{i=1}^n P(A_i).$$

3. 若  $A \subset B$ , 则  $P(B) \geq P(A)$ . 等号成立当且仅当  $A = B$ .

4. 对于  $\forall A$ , 必有  $P(A) \leq 1$ .
5. 对于  $\forall A$ , 有  $P(\bar{A}) = 1 - P(A)$ .
6. 对于  $\forall A, B$  有

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

7. 次可加性: 对于任意事件列  $A_1, A_2, \dots, A_n, \dots$ , 有  $P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n)$ .
8. 下连续性: 若事件列满足  $A_n \subset A_{n+1}, n = 1, 2, \dots$ , 则

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n).$$

9. 上连续性: 若事件列满足  $A_n \supset A_{n+1}, n = 1, 2, \dots$ , 则

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n).$$

可以将上述性质 6 推广至一般形式, 即为容斥原理:

**定理 1.4.1.** 容斥原理 对任意的事件  $A_1, A_2, \dots, A_n$ , 有

$$P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k) - \sum_{1 \leq i < j \leq n} P(A_i A_j) + \sum_{1 \leq i < j < k \leq n} P(A_i A_j A_k) - \dots + (-1)^{n-1} P(A_1 A_2 \dots A_n).$$

**证明.** 应用数学归纳法.  $n = 2$  时, 由于  $A_1 \cup A_2 = A_1 + A_2 - A_1 \cap A_2$ , 根据有限可加性, 有

$$P(A_1 \cup A_2) = P(A_1) + P(A_2 - A_1 \cap A_2) = P(A_1) + P(A_2) - P(A_1 A_2).$$

假设对  $n = k - 1$  成立, 当  $n = k$  时, 应用归纳假设前提有:

$$\begin{aligned} P\left(\bigcup_{i=1}^k A_i\right) &= P\left(\left(\bigcup_{i=1}^{k-1} A_i\right) \cup A_k\right) \\ &= P\left(\bigcup_{i=1}^{k-1} A_i\right) + P(A_k) - P\left(\left(\bigcup_{i=1}^{k-1} A_i\right) \cap A_k\right) \\ &= P\left(\bigcup_{i=1}^{k-1} A_i\right) + P(A_k) - P\left(\bigcup_{i=1}^{k-1} (A_i A_k)\right) \\ &= \sum_{i=1}^k P(A_i) - \sum_{1 \leq i < j \leq k} P(A_i A_j) + \sum_{1 \leq i < j < k \leq k} P(A_i A_j A_k) - \dots + (-1)^{k-1} P(A_1 A_2 \dots A_k). \end{aligned}$$

□



解. 由题意可知  $AB \subset C$ , 因此  $P(C) \geq P(AB)$ , 故 C 错误. 且  $C$  不一定等于  $A \cup B$ , D 错误. 又因为

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

可以推出

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

根据包含关系, 有:

$$\begin{aligned} P(C) &\geq P(AB) = P(A) + P(B) - P(A \cup B) \\ &\geq P(A) + P(B) - 1. \end{aligned}$$

故 B 正确. □

例 1.4.4. 随机事件  $A, B$  及其和事件  $A \cup B$  发生的概率分别为 0.4, 0.3, 0.6.  $\bar{B}$  表示  $B$  的对立事件, 那么积事件  $A\bar{B}$  发生的概率  $P(A\bar{B})$  为?

解. 根据概率的运算律, 有:

$$P(A\bar{B}) = P(A(\Omega - B)) = P(A\Omega - AB) = P(A - AB) = P(A) - P(AB).$$

上式中  $P(A - AB) = P(A) - P(AB)$  成立因为  $AB \subset A$ . 再根据容斥原理:

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

可以得出

$$P(AB) = P(A) + P(B) - P(A \cup B) = 0.1.$$

因此可以得出

$$P(A\bar{B}) = P(A) - P(AB) = 0.4 - 0.1 = 0.3.$$

□

## 1.5 古典概型

**定义 1.5.1.** 满足以下两个特性的概率模型为古典概型:

1. **有限性:**  $\Omega$  包含的样本点为有限个.
2. **等可能性:** 样本点 (基本事件) 发生的可能性相同.

## 1.6 条件概率

考虑事件  $A$  发生的情况下, 事件  $B$  发生的概率。

**例 1.6.1.** 将一枚硬币抛掷两次, 观察其出现正反面的情况. 设事件  $A$  为“至少有一次为  $H$ ”, 事件  $B$  为“两次掷出同一面”. 现在来求已知事件  $A$  发生的条件下事件  $B$  发生的概率.

**解.** 考虑样本空间:

$$\Omega = \{HH, HT, TH, TT\}$$

$$A = \{HH, HT, TH\}$$

$$B = \{HH, TT\}$$

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

□

**推论 1.6.1.** 设试验的样本空间的样本点总数为  $n$ ,  $A$  包含的样本点有  $m$  个 ( $m > 0$ ),  $AB$  包含的样本点有  $k$  个, 则

$$P(B|A) = \frac{k}{m} = \frac{k/n}{m/n} = \frac{P(AB)}{P(A)}.$$

**定义 1.6.1.** 条件概率 设  $A, B$  为两个事件, 且  $P(A) > 0$ , 则

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

**特点**

1. **非负性:** 对  $\forall B, P(B|A) \geq 0$ ;
2. **规范性:** 对于必然事件  $\Omega$ , 有  $P(\Omega|A) = 1$ ;

3. 可列可加性: 设  $B_1, B_2, B_3, \dots$  两两互不相容, 则有

$$P\left(\bigcup_{i=1}^{\infty} B_i|A\right) = \sum_{i=1}^{\infty} P(B_i|A).$$

4. 对  $\forall B_1, B_2$ , 有

$$P(B_1 \cup B_2|A) = P(B_1|A) + P(B_2|A) - P(B_1 B_2|A).$$

**例 1.6.2.** 一盒子装有 4 只产品, 3 只一等品, 1 只二等品, 从中取产品两次, 每次任取一只, 不放回抽样. 设事件  $A$  为“第一次取到的是一等品”, 事件  $B$  为“第二次取到的是一等品”. 试求条件概率  $P(B|A)$ .

解.

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{\frac{3}{4} \times \frac{2}{3}}{\frac{3}{4}} = \frac{2}{3}.$$

□

**例 1.6.3.** 设某动物活 20 年以上的概率是 0.8, 活 25 年以上的概率为 0.4. 现有只 20 岁的该动物, 问该动物能活到 25 岁以上的概率.

解. 依题意, 设事件  $A$  为动物活到 20 岁以上, 事件  $B$  为动物活到 25 岁以上. 现在要求  $P(B|A)$ . 根据条件概率公式, 有

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{0.4}{0.8} = 0.5$$

其中, 由于  $B \subset A$ , 因此  $P(AB) = P(B)$ .

□

**例 1.6.4.** 设  $A, B$  为随机事件, 且  $0 < P(A) < 1, P(B) > 0, P(B|A) = P(B|\bar{A})$ , 则必有?

- A.  $P(A|B) = P(\bar{A}|B)$ .                      B.  $P(A|B) \neq P(\bar{A}|B)$ .  
 C.  $P(AB) = P(A)P(B)$ .                      D.  $P(AB) \neq P(A)P(B)$ .

解.

$$\begin{aligned} P(B|A) &= P(B|\bar{A}) \\ \Rightarrow \frac{P(AB)}{P(A)} &= \frac{P(\bar{A}B)}{P(\bar{A})} = \frac{P(B) - P(AB)}{1 - P(A)} \\ \Rightarrow P(AB) - P(AB) \cdot P(A) &= P(A) \cdot P(B) - P(A) \cdot P(AB) \\ \Rightarrow P(AB) &= P(A) \cdot P(B). \end{aligned}$$

□

## 1.7 乘法定理

由条件概率:

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

可得乘法定理:

$$P(AB) = P(A) \cdot P(B|A).$$

**推论 1.7.1.** 选妃公式

$$P(ABC) = P(A) \cdot P(B|A) \cdot P(C|AB).$$

$$P(ABCD) = P(A) \cdot P(B|A) \cdot P(C|AB) \cdot P(D|ABC).$$

**例 1.7.1.** 设袋中有  $r$  只红球,  $t$  只白球. 每次自袋中任取一球, 观察其颜色然后放回, 并再放入  $a$  只与所取出的那只球同色的球. 若在袋中连续取球四次, 试求第一、二次取到红球且第三、四次取到白球的概率.

**解.** 设  $A_i$  为第  $i$  次取到红球,  $B_i$  为第  $i$  次取到白球, 题意要求  $P(A_1A_2B_3B_4)$ . 根据乘法定理:

$$\begin{aligned} P(A_1A_2B_3B_4) &= P(A_1) \cdot P(A_2|A_1) \cdot P(B_3|A_1A_2) \cdot P(B_4|A_1A_2B_3) \\ &= \frac{r}{r+t} \cdot \frac{r+a}{r+t+a} \cdot \frac{t}{r+t+2a} \cdot \frac{t+a}{r+t+3a}. \end{aligned}$$

□

**例 1.7.2.** 设某光学仪器厂制造的透镜, 第一次落下时打破的概率为  $\frac{1}{2}$ , 若第一次落下未打破, 第二次落下打破的概率为  $\frac{7}{10}$ , 第三次落下打破的概率为  $\frac{9}{10}$ . 求透镜落下三次而未打破的概率.

**解.** 设  $A_i$  为第  $i$  次落下打破, 则题意要求  $P(\overline{A_1} \cap \overline{A_2} \cap \overline{A_3})$ . 根据乘法定理:

$$\begin{aligned} P(\overline{A_1} \cap \overline{A_2} \cap \overline{A_3}) &= P(\overline{A_1}) \cdot P(\overline{A_2}|\overline{A_1}) \cdot P(\overline{A_3}|\overline{A_1} \cap \overline{A_2}) \\ &= \frac{1}{2} \times \frac{3}{10} \times \frac{1}{10} \\ &= 1.5\%. \end{aligned}$$

□



**例 1.7.3.** 甲袋中 3 个白球 6 个黄球，乙袋中 5 个白球 4 个黄球；先从甲袋中任选一只放入乙袋，再从乙袋中任选一只放入甲；问：甲袋中白球的数目不发生变化的概率.

**解.** 题意要求两次选择选中同样颜色球的概率. 设  $A_i$  为第  $i$  次选中白球，则要求  $P((A_1 \cap A_2) \cup (\overline{A_1} \cap \overline{A_2})) = P(A_1 \cap A_2) + P(\overline{A_1} \cap \overline{A_2})$ .

$$P(A_1 \cap A_2) + P(\overline{A_1} \cap \overline{A_2}) = \frac{1}{3} \times \frac{3}{5} + \frac{2}{3} \times \frac{1}{2} = \frac{8}{15}.$$

其中  $P((A_1 \cap A_2) \cup (\overline{A_1} \cap \overline{A_2})) = P(A_1 \cap A_2) + P(\overline{A_1} \cap \overline{A_2})$  成立因为  $(A_1 \cap A_2)$  与  $(\overline{A_1} \cap \overline{A_2})$  互不相容. □

**例 1.7.4.** 100 件产品中有 10 件次品，用不放回方式每次抽取一件，连续抽 3 次，问第三次才抽到次品的概率.

**解.** 设  $A_i$  为第  $i$  次抽中次品，则题意要求  $P(\overline{A_1} \cdot \overline{A_2} \cdot A_3)$ . 根据乘法定理：

$$\begin{aligned} P(\overline{A_1} \cdot \overline{A_2} \cdot A_3) &= P(\overline{A_1}) \cdot P(\overline{A_2} | \overline{A_1}) \cdot P(A_3 | \overline{A_1} \cdot \overline{A_2}) \\ &= \frac{9}{10} \times \frac{89}{99} \times \frac{10}{98} \end{aligned}$$

□

## 1.8 全概率公式

### 1.8.1 完备事件群

**定义 1.8.1.** 试验  $E$  的样本空间为  $\Omega$ ,  $B_1, B_2, B_3, \dots, B_n$  为  $E$  中的一组事件, 若满足:

1.  $B_i \cap B_j = \emptyset$ , 其中  $i \neq j$  且  $i, j = 1, 2, 3, \dots, n$ ;
2.  $B_1 \cup B_2 \cup \dots \cup B_n = \Omega$ .

则称  $B_1, B_2, \dots, B_n$  为  $\Omega$  的一个划分 (完备事件群).

**注.** 注意到, 上述定义中的两个条件, 若只对两个事件成立, 即对  $B_1, B_2$  有  $B_1 \cap B_2 = \emptyset$  且  $B_1 \cup B_2 = \Omega$ . 则  $B_1, B_2$  为一组对立事件.

**例 1.8.1.** 设试验  $E$  为投掷一枚骰子观察点数, 其中有一下三个事件:

$$B_1 = \{1, 2, 3\};$$

$$B_2 = \{4, 5\};$$

$$B_3 = \{6\}.$$

根据定义,  $B_1, B_2, B_3$  显然是  $\Omega = \{1, 2, 3, 4, 5, 6\}$  的一个划分 (完备事件群).

### 1.8.2 全概率公式

**定理 1.8.1.** 全概率公式 设  $E$  的样本空间为  $\Omega$ ,  $A$  是  $E$  中的事件,  $B_1, B_2, \dots, B_n$  为  $\Omega$  的一组完备事件群, 且  $P(B_i) > 0$ , ( $i = 1, 2, 3, \dots, n$ ), 则

$$P(A) = \sum_{i=1}^n P(A|B_i) \cdot P(B_i).$$

**证明.** 事件  $A$  可以表示为

$$A \cap \Omega = A \cap \left( \bigcup_{i=1}^n B_i \right) = \bigcup_{i=1}^n (A \cap B_i).$$

又因为  $A \cap B_i \subset B_i$  且  $B_i, i = 1, 2, 3, \dots, n$  之两两互不相容, 所以对  $\forall i \neq j$  有:

$$(A \cap B_i) \cap (A \cap B_j) = \emptyset.$$

因此有:

$$\begin{aligned} P(A) &= P(A \cap \Omega) = P\left(A \cap \left(\bigcup_{i=1}^n B_i\right)\right) = P\left(\bigcup_{i=1}^n (A \cap B_i)\right) \\ &= \sum_{i=1}^n P(AB_i). \end{aligned}$$

根据乘法定理:

$$\begin{aligned} P(A) &= \sum_{i=1}^n P(AB_i) \\ &= \sum_{i=1}^n P(B_i) \cdot P(A|B_i). \end{aligned}$$

□

**例 1.8.2.** 一批产品共有 10 个正品和 2 个次品, 任意抽取两次, 不返回抽取. 问: 第二次抽到次品的概率.

**解.** 设事件  $A$  为第二次抽到次品; 事件  $B$  为第一次抽到正品. 显然事件  $B$  与事件  $\bar{B}$  构成样本空间  $\Omega$  的一个完备事件群. 根据全概率公式:

$$\begin{aligned} P(A) &= P(B) \cdot P(A|B) + P(\bar{B}) \cdot P(A|\bar{B}) \\ &= \frac{5}{6} \times \frac{2}{11} + \frac{1}{6} \times \frac{1}{11}. \end{aligned}$$

□

**注.** 上面的例题在构造完备事件群时只考虑了第一次抽取样品, 这是合理的. 因为第二次抽取样品要在第一次抽取之后, 由于抽样是不放回抽样, 第一次抽取的结果会影响第二次抽取时的概率, 所以第二次抽样的事件是第一次抽样事件的子集, 因此只需要考虑第一次抽样即可构造完备事件群. 这种方式在类似的情况中都成立.

**例 1.8.3.** 从 1, 2, 3, 4 中任取一个数, 记为  $x$ , 再从 1, 2, ...,  $x$  中任取一个数记为  $y$ . 求  $P\{y = 2\}$ .

**解.** 依题意  $x = i$  为第一次抽到  $i$ , 显然  $\{x = i | i = 1, 2, 3, 4\}$  为样本空间  $\Omega$  的一个完备事件群. 根据全概率公式:

$$\begin{aligned} P(y = 2) &= \sum_{i=1}^4 P(x = i) \cdot P(y = 2 | x = i) \\ &= \frac{13}{48}. \end{aligned}$$

□

**例 1.8.4.** 研究表明肺癌的患病概率为 0.1%, 在人群中 20% 是吸烟者, 他们患肺癌的概率为 0.4%, 求不吸烟者患肺癌的概率.

**解.** 患肺癌的概率受吸烟与否的影响, 因此设事件  $A$  为被抽样者吸烟, 则  $\bar{A}$  为被抽样者不吸烟. 事件  $A$  与  $\bar{A}$  构成样本空间  $\Omega$  的一个完备事件群. 设事件  $B$  为被抽样者患肺癌, 根据全概率公式:

$$\begin{aligned} P(B) &= P(A) \cdot P(B|A) + P(\bar{A}) \cdot P(B|\bar{A}) \\ &= 20\% \times 0.4\% + 80\% \times P(B|\bar{A}) \\ &= 0.1\%. \end{aligned}$$

则可得问题目标  $P(B|\bar{A})$ .

$$P(B|\bar{A}) = (0.001 - 0.0008) \div 0.8 = 0.025\%.$$

□

## 1.9 贝叶斯公式

**定理 1.9.1.** 设试验  $E$  的样本空间  $\Omega$ .  $A$  是  $E$  的一个事件.  $\{B_i | i = 1, 2, 3, \dots, n\}$  为  $\Omega$  的一个完备事件群. 且  $P(A) > 0, P(B_i) > 0$ . 那么

$$P(B_i|A) = \frac{P(B_i) \cdot P(A|B_i)}{\sum_{j=1}^n P(B_j) \cdot P(A|B_j)}. \quad (i = 1, 2, 3, \dots, n)$$

**证明.** 根据条件概率公式, 有:

$$P(B_i|A) = \frac{P(A \cdot B_i)}{P(A)}.$$

再对分母用全概率公式, 对分子用乘法原理得:

$$P(B_i|A) = \frac{P(A \cdot B_i)}{P(A)} = \frac{P(B_i) \cdot P(A|B_i)}{\sum_{j=1}^n P(B_j) \cdot P(A|B_j)}.$$

□

**例 1.9.1.** 某电子设备制造厂所用的元件是由三家元件制造厂提供的. 根据以往的记录有以下数据:

元件制造厂	次品率	提供元件的份额
1	0.02	0.15
2	0.01	0.80
3	0.03	0.05

设这三家工厂的产品再仓库中是均匀混合的, 且无区别的标志.

- (1) 在仓库中随机地取一只元件, 求它是次品的概率;
- (2) 在仓库中随机地取一只元件, 若已知取到的是次品, 求此次品出自三家工厂的概率.

**解.**

- (1) 在仓库中随机地取一只元件, 设  $A_i$  为此元件出自工厂  $i$ . 显然  $\{A_i\}$  为样本空间  $\Omega$  的一个完备事件群. 设事件  $B$  为该元件是次品, 根据全概率公式:

$$P(B) = \sum_{i=1}^3 P(A_i) \cdot P(B|A_i) = 0.15 \times 0.02 + 0.80 \times 0.01 + 0.05 \times 0.03 = 1.25\%.$$

- (2) 根据贝叶斯公式:

$$\begin{aligned} P(A_i|B) &= \frac{P(A_i) \cdot P(B|A_i)}{P(B)} \\ \Rightarrow P(A_1|B) &= \frac{P(A_1) \cdot P(B|A_1)}{P(B)} = \frac{0.15 \times 0.02}{0.0125} = 0.24 \\ \Rightarrow P(A_2|B) &= \frac{P(A_2) \cdot P(B|A_2)}{P(B)} = \frac{0.80 \times 0.01}{0.0125} = 0.64 \\ \Rightarrow P(A_3|B) &= \frac{P(A_3) \cdot P(B|A_3)}{P(B)} = \frac{0.05 \times 0.03}{0.0125} = 0.12. \end{aligned}$$

□

**例 1.9.2.** 对以往的数据分析结果表明, 当及其调整得良好时, 产品得合格率为 98%, 而当机器发生某种故障时, 其合格率为 55%. 每天早上机器开动时, 机器调整良好的概率为 95%. 现已知某日早上第一件产品是合格品, 求机器调整良好的概率.

**解.** 设事件  $A$  为机器调整良好, 事件  $B$  为生产的产品合格. 问题要求  $P(A|B)$ . 由于事件  $B$  的概率受到事件  $A$  的影响, 事件  $A$  与事件  $\bar{A}$  构成样本空间  $\Omega$  的一个完备事件群. 因此

模型满足贝叶斯公式的适用条件, 根据贝叶斯公式:

$$\begin{aligned} P(A|B) &= \frac{P(A) \cdot P(B|A)}{P(A) \cdot P(B|A) + P(\bar{A}) \cdot P(B|\bar{A})} \\ &= \frac{95\% \times 98\%}{95\% \times 98\% + 5\% \times 55\%} \\ &\approx 0.9713. \end{aligned}$$

□

注. 上例, 题干中说机器调整良好的概率为 95%, 这是在生产前根据以往的经验得出的, 我们称之为“先验概率”. 而在多加了一个条件“生产的第一件产品为合格品”后, 计算得出机器调整良好的概率约为 97%, 这是“后验概率”. 后验概率是在得到额外信息后对先验概率修正后的结果.

## 1.10 事件独立性

试验  $E$  的事件  $A, B$ . 若  $P(A) > 0$ , 就可以定义  $P(B|A)$ . 一般情况下  $P(B|A) \neq P(B) \implies$  事件  $A$  发生与否会对事件  $B$  发生的概率产生影响; 有的时候  $P(B|A) = P(B) \implies$  事件  $A$  发生与否不会对事件  $B$  发生的概率产生影响.

**例 1.10.1.** 随机试验  $E$ : 投两枚硬币 (甲和乙), 观察正反面出现的情况. 假定事件  $A$ : 甲币正面向上; 事件  $B$ : 乙币正面向上. 问:  $A$  事件发生与否是否会对  $B$  发生的概率产生影响.

解. 因为:

$$\begin{aligned} \Omega &= \{HH, HT, TH, TT\} \\ A &= \{HH, HT\} \quad P(A) = \frac{1}{2} \quad P(AB) = \frac{1}{4} \\ B &= \{HH, TH\} \quad P(B) = \frac{1}{2} \\ P(B|A) &= \frac{P(AB)}{P(A)} = \frac{1}{2} = P(B) \\ P(B|\bar{A}) &= \frac{P(\bar{A}B)}{P(\bar{A})} = \frac{1}{2} = P(B). \end{aligned}$$

所以没有影响. 同时, 运用同样的分析方法也可以得出  $B$  发生与否对  $A$  没有影响. □

注. 在这个例子中, 我们可以发现  $P(AB)$  恰好等于  $P(A) \cdot P(B)$ , 这是否是普遍性的结论呢?

定义 1.10.1. 假设  $A, B$  为两个事件, 如果满足:

$$P(AB) = P(A) \cdot P(B).$$

则事件  $A$  与事件  $B$  相互独立, 反之也成立.

定理 1.10.1. 设  $A, B, C$  三个事件, 有:

$$(1) P(AB) = P(A) \cdot P(B) \implies A \text{ 与 } B \text{ 相互独立};$$

$$(2) P(AC) = P(A) \cdot P(C) \implies A \text{ 与 } C \text{ 相互独立};$$

$$(3) P(CB) = P(C) \cdot P(B) \implies C \text{ 与 } B \text{ 相互独立};$$

$$(4) P(ABC) = P(A) \cdot P(B) \cdot P(C).$$

若  $A, B, C$  满足 (1), (2), (3) 式, 则  $A, B, C$  两两独立. 若满足 (1), (2), (3), (4) 式, 则  $A, B, C$  相互独立.

例 1.10.2. 盒中有编号为 1, 2, 3, 4 的 4 只球, 随机地从盒中取一只球, 事件  $A$  为“取得的是 1 号球或 2 号球”, 事件  $B$  为“取得的是 1 号或 3 号球”, 事件  $C$  为“取得的是 1 号或 4 号球”验证:

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

但  $P(ABC) \neq P(A) \cdot P(B) \cdot P(C)$ .

解. 根据题意:

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

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

$$\implies P(AB) = P(A) \cdot P(B)$$

$$P(AC) = P(A) \cdot P(C)$$

$$P(BC) = P(B) \cdot P(C)$$

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

□

**准则 1.10.2.** 若  $P(A) > 0, P(B) > 0$ , 则  $A$  与  $B$  相互独立与  $A$  与  $B$  不相容不能同时发生.

**证明.** 当  $P(A) > 0, P(B) > 0$  时, 若  $A$  与  $B$  不相容, 则有:

$$P(AB) = P(\emptyset) = 0 \neq P(A) \cdot P(B).$$

因此  $A$  与  $B$  不相互独立. □

**注.** 在上面的准则中, 若  $P(A) = 0$  或  $P(B) = 0$ , 且  $A$  与  $B$  不相容, 就有:

$$P(AB) = P(\emptyset) = 0 = P(A) \cdot P(B).$$

由此可见, 该准则的前提条件很重要.

**推论 1.10.3.** 若  $A, B$  相互独立, 则下列结论成立:

1.  $A$  与  $\bar{B}$  相互独立;
2.  $\bar{A}$  与  $B$  相互独立;
3.  $\bar{A}$  与  $\bar{B}$  相互独立.

**证明.** 这里只证明  $A$  与  $\bar{B}$  相互独立. 因为  $A, B$  相互独立:

$$\begin{aligned} A &= A\Omega = A(B \cup \bar{B}) = (A \cap B) \cup (A \cap \bar{B}) \\ \implies P(A) &= P(AB) + P(A\bar{B}) \\ \implies P(A) &= P(A) \cdot P(B) + P(A\bar{B}) \\ \implies P(A)(1 - P(B)) &= P(A\bar{B}) \\ \implies P(A) \cdot P(\bar{B}) &= P(A\bar{B}). \end{aligned}$$

因此  $A$  与  $\bar{B}$  相互独立. □

**推论 1.10.4.** 若事件群  $\{A_i | i = 1, 2, 3, \dots, n\}, (n \geq 2)$  相互独立, 则:

1. 任取其中  $k$  个事件也都相互独立;
2. 将  $\{A_i\}$  中任意  $j$  个事件换为其对立事件, 得到的新事件群也相互独立;



## 第二章 随机变量及其分布

### 2.1 随机变量

**例 2.1.1.** 设两枚硬币，观察正反面情况，用  $X$  记作得到正面的次数。分析  $X$  与该试验样本空间之间的关系。

**解.** 记样本空间为  $\Omega$ ，则

$$S = \{HH, HT, TH, TT\}$$

其中  $X$  与样本空间  $\Omega$  的对应关系有：

$$X = 2, HH$$

$$X = 1, HT$$

$$X = 1, TH$$

$$X = 0, TT$$

□

**注.** 上例中， $X$  就是变量，是应变量，将  $X$  看作定义在  $\Omega$  上的一个函数，假定  $\Omega$  内的样本点为  $e$ ，则：

$$X = X(e) = \begin{cases} 2 & e = HH \\ 1 & e = HT, TH \\ 0 & e = TT \end{cases}$$

**定义 2.1.1.** 随机试验  $E$  的样本空间  $\Omega = \{e\}$ ， $X = X(e)$  是定义在  $\Omega$  上的实单值函数，则称  $X = X(e)$  是随机变量。

注. 字母表示规则:

1. 随机变量: 用大写字母表示;
2. 实数: 用小写字母表示;
3. 某些试验: 结果本身就是一个数, 可以将试验结果本身作为随机变量.

## 2.2 离散型随机变量及其分布

### 2.2.1 离散型随机变量

定义 2.2.1. 取值是有限多个或者无限可列的随机变量称为离散型随机变量.

注. 需要掌握的目标:

1.  $X$  的所有取值;
2.  $X$  每个取值的概率.

上述两点的综合即为随机变量  $X$  的分布率.

### 2.2.2 分布率及其性质

例 2.2.1. 设一辆汽车在开往目的地的道路上需经过四组信号灯, 每组信号灯以  $p$  的概率禁止汽车通过. 以  $X$  表示汽车首次停下时, 它已通过的信号灯的组数 (各信号灯工作相互独立), 求随机变量  $X$  的分布率.

解. 第一步, 求随机变量  $X$  的所有可能取值:

$$X = 0, 1, 2, 3, 4$$

第二步, 计算每个取值对应的概率:

$$P\{X = 0\} = p$$

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

$$P\{X = 2\} = (1 - p)^2 p$$

$$P\{X = 3\} = (1 - p)^3 p$$

$$P\{X = 4\} = (1 - p)^4.$$

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

第三步, 画出分布图表. □

注. 由例可以发现, 分布率中的概率有两个性质:

1.  $P(X = i) \geq 0$ ;
2.  $\sum P(X = i) = 1$ .

### 2.2.3 重要分布

#### 0 - 1 分布 (两点分布)

定义 2.2.2. 设随机变量  $X$  只可能取到 0 和 1 两个值, 则它的分布率满足:

$$P\{X = k\} = p^k (1-p)^{1-k}, \quad k = 0, 1$$

称  $X$  服从以  $p$  为参数的 0 - 1 分布.

#### 伯努力试验及二项分布

定义 2.2.3. 如果试验  $E$  只有  $A, \bar{A}$  两个结果, 则试验称为伯努力试验. 将这样的试验重复独立地进行  $n$  次, 则称为  $n$  重伯努力试验.

注. 上述定义中:

1. 重复:  $P(A)$  保持不变;
2. 每次试验之间相互独立.

定义 2.2.4. 设  $X$  表示  $n$  重伯努力试验中  $A$  发生的次数 ( $X = 0, 1, 2, \dots, n$ ), 则它的分布率满足:

$$P\{X = k\} = C_n^k \cdot p^k (1-p)^{n-k}, \quad k = 0, 1, 2, 3, \dots, n$$

称  $X$  服从参数为  $n$  和  $p$  的二项分布, 记作  $X \sim B(n, p)$ .

例 2.2.2. 有 80 台同类型设备, 各台设备工作相互独立, 发生故障的概率均为 0.01 且一台设备的故障只能由一个人处理. 考虑两种配备维修工人的方法, 其一是由 4 人维护, 每人

负责 20 台；其二是由 3 人共同维护 80 台. 试比较这两种方法在设备发生故障时不能及时维修的概率.

**解.** 设事件  $A$  为设备发生故障时不能及时维修, 题目要求  $P(A)$ . 接下来分别考虑两种方案:

1. 对于方案一, 只要四个人中有一个人不能正常维修, 则事件  $A$  发生. 因此该方案下有:

$$P(A) = P(A_{\text{甲}} \cup A_{\text{乙}} \cup A_{\text{丙}} \cup A_{\text{丁}}).$$

对于  $P(A_i)$ ,  $i = \text{甲, 乙, 丙, 丁}$  可以通过二项分布来计算. 记  $X$  为同时发生故障的设备数, 则:

$$P(A_i) = 1 - \sum_{j=0}^1 P\{X = j\} = 0.0169.$$

因而  $P(A) = \sum_{\text{甲}}^{\text{丁}} P(A_i) = 0.0507$ .

2. 三人共同维护时, 当同时四台设备发生故障则不能及时维护, 因此根据二项分布:

$$P(A) = 1 - \sum_{i=0}^3 P\{X = i\} = 0.0087.$$

3. 经过比较, 显然第二种方案比较好.

□

## 泊松分布

**定义 2.2.5.** 设随机变量  $X = 0, 1, 2, 3, \dots$ , 每个取值的概率满足:

$$P\{X = k\} = \frac{\lambda^k}{k!} \cdot e^{-\lambda}. \quad k = 0, 1, 2, 3, \dots, \lambda > 0$$

称  $X$  服从参数为  $\lambda$  的泊松分布, 记为  $X \sim P(\lambda)$ . (也记作  $X \sim \pi(\lambda)$ ).

**例 2.2.3.** 已知  $X \sim \pi(4)$ , 求  $P\{X \leq 3\}$ .

**解.**

$$P\{X \leq 3\} = \sum_{i=0}^3 P\{X = i\} = \sum_{i=0}^3 \frac{4^i}{i!} \cdot e^{-4} = \frac{71}{3} e^{-4}.$$

□

注. 泊松分布的应用场景:

1. 一本书, 一页中发生的印刷错误的数量;
2. 某医院一天内急诊病人的人数;
3. 某提取, 一个时间间隔内, 发生的交通事故的次数.

**定理 2.2.1.** 泊松定理 设  $\lambda$  是一个正常数,  $n$  为任意正整数, 且  $n \cdot p = \lambda$ . 则对任意固定的非负整数  $k$  有:

$$\lim_{n \rightarrow \infty} C_n^k \cdot p^k (1-p)^{n-k} = \frac{\lambda^k}{k!} \cdot e^{-\lambda}.$$

注. 上述定理右侧的形式与泊松分布的分布律形式相同; 左边是对二项分布的分布律求极限. 该等式表明, 对一个伯努利试验, 重复独立地做非常多的次数时, 二项分布的分布律趋近于泊松分布的分布律.

接下来给出泊松定理的证明.

**证明.**

$$\begin{aligned} & \lim_{n \rightarrow \infty} C_n^k \cdot p^k (1-p)^{n-k} \\ &= \lim_{n \rightarrow \infty} \left[ \frac{n \cdot (n-1) \cdot (n-2) \cdots (n-k+1)}{k!} \cdot \left(\frac{\lambda}{n}\right)^k \cdot \left(1 - \frac{\lambda}{n}\right)^{n-k} \right] \\ &= \lim_{n \rightarrow \infty} \left[ \frac{\lambda^k}{k!} \left( \frac{n \cdot (n-1) \cdot (n-2) \cdots (n-k+1)}{n^k} \right) \cdot \left(1 - \frac{\lambda}{n}\right)^n \cdot \left(1 - \frac{\lambda}{n}\right)^{-k} \right] \\ &= \lim_{n \rightarrow \infty} \frac{\lambda^k}{k!} \left[ 1 \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \right] \left(1 - \frac{\lambda}{n}\right)^n \cdot \left(1 - \frac{\lambda}{n}\right)^{-k} \\ &= \frac{\lambda^k}{k!} \cdot e^{-\lambda}. \end{aligned}$$

□

**例 2.2.4.** 计算机硬件公司制造某种特殊型号的微型芯片, 次品率达到 0.1%, 各芯片成为次品相互独立. 求在 1000 只产品中至少有 2 只次品的概率. 以  $X$  记产品中的次品数,  $X \sim B(1000, 0.001)$ .

**解.** 由于  $X \sim B(1000, 0.001)$ , 其中参数  $n$  非常大, 因此考虑用泊松分布近似. 由于

$n \cdot p = 1$ , 因此考虑  $X \sim \pi(1)$ :

$$\begin{aligned} P\{X \geq 2\} &= 1 - P\{X < 2\} = 1 - \sum_{i=0}^1 P\{X = i\} \\ &\sim 1 - e^{-1} \cdot \left( \sum_{i=0}^1 \frac{1^i}{i!} \right) = \frac{e-2}{e}. \end{aligned}$$

□

**例 2.2.5.** 随机变量  $X$  服从泊松分布, 并且已知  $P\{X = 1\} = P\{X = 2\}$ , 求  $P\{X = 4\}$ .

**解.** 根据题目所给条件, 设  $X \sim \pi(\lambda)$  则有:

$$P\{X = 1\} = \lambda \cdot e^{-\lambda} = P\{X = 2\} = \frac{\lambda^2}{2} \cdot e^{-\lambda} \implies \lambda = 2.$$

因此  $X \sim \pi(2)$ , 得出:

$$P\{X = 4\} = \frac{2^4}{24} \cdot e^{-2} = \frac{4}{3 \cdot e^2}.$$

□

## 几何分布

**定义 2.2.6.** 伯努利试验“成功”的概率为  $p$ . 试验第  $k$  次才成功的概率.

$$P\{X = k\} = (1-p)^{k-1} \cdot p, \quad k = 1, 2, 3, \dots$$

称  $X$  的分布律服从参数为  $p$  的几何分布, 记为  $X \sim Ge(p)$ .

**例 2.2.6.** 一个盒子里有 100 颗球, 98 颗白球; 两颗红球. 每次抽 1 个观察颜色, 有放回抽烟个, 抽到红球就停止. 假设  $X$  代表停止时抽取的次数, 求:

- (1)  $X$  的分布率;
- (2)  $P\{X = 6\}$ .

**解.**

- (1) 设“成功”表示抽到红球, 则  $X = k$  表示第  $k$  次才成功. 根据集合分布:

$$P\{X = k\} = \left( \frac{98}{100} \right)^{k-1} \times \frac{2}{100}.$$

- (2)

$$P\{X = 6\} = 0.98^5 \times 0.02 = 0.018$$

□

**例 2.2.7.** 一批产品不合格的概率为 0.05, 对该批产品进行逐个抽查, 求第一次抽到不合格产品时, 检查的次数  $X$  的分布率.

**解.** 定义“成功”: 抽到不合格产品, 则:

$$P\{X = k\} = 0.95^{k-1} \times 0.05$$

□

### 超几何分布

从有限  $N$  个物品 (其中含有  $D$  个特定物品), 从  $N$  中抽出  $n$  个物品, 考虑其中包含  $k$  个特定物品的概率.

**例 2.2.8.** 设有  $N$  件产品, 其中有  $D$  件次品, 今从中任取  $n$  件, 求其中恰有  $k(k \leq D)$  件次品的概率.

**解.** 设随机变量  $X$  表示取中次品的数量, 则题目要求  $P\{X = k\}$ , 则:

$$P\{X = k\} = \frac{C_D^k \cdot C_{N-D}^{n-k}}{C_N^n}$$

□

**定义 2.2.7.** 随机变量  $X$  的分布率为如下形式:

$$P\{X = k\} = \frac{C_D^k \cdot C_{N-D}^{n-k}}{C_N^n}$$

称  $X$  满足参数为  $n, D, N$  的超几何分布, 记作  $X \sim H(n, D, N)$ .

## 2.3 随机变量的分布函数

**定义 2.3.1.** 设随机变量  $X$ ,  $x$  为任意实数, 则称

$$F(x) = P\{X \leq x\}, \quad -\infty < x < +\infty.$$

为  $X$  的分布函数.

## 2.4 连续型随机变量及其概率密度

### 2.4.1 连续型随机变量及其概率密度函数

**定义 2.4.1.** 随机变量  $X$  的分布函数为  $F(x)$ , 若  $F(x)$  可由一个非负的可积函数  $f(x)$  积分得到:

$$F(x) = \int_{-\infty}^x f(t) dt.$$

则称  $X$  是连续型随机变量, 其概率密度函数是  $f(x)$ .

**注.** 连续型随机变量  $X$  的分布函数是连续函数.

#### 概率密度的性质

1.  $f(x) \geq 0$ ;
2.  $\int_{-\infty}^{+\infty} f(x) dx = 1$ ;
3. 对于  $\forall x_1, x_2$  且  $x_1 \leq x_2$ , 则:

$$P\{x_1 < X \leq x_2\} = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x) dx.$$

4. 如果  $f(x)$  在  $x = x_0$  处连续, 则  $F'(x_0) = f(x_0)$ ;
5.  $P\{X = a\} = 0$ ;
6. 对于连续型随机变量, 有

$$\begin{aligned} P\{x_1 \leq X \leq x_2\} &= P\{X \leq x_2\} - P\{X \leq x_1\} + P\{X = x_1\} \\ &= P\{X \leq x_2\} - P\{X \leq x_1\} \end{aligned}$$

#### 概率分布

随机变量的概率分布  $\implies$  分布函数 +  $\begin{cases} \text{分布律} & \text{离散型随机变量} \\ \text{概率密度函数} & \text{连续型随机变量} \end{cases}$

**例 2.4.1.** 一个靶子半径为 2 m 的圆盘, 设击中靶上任一同心圆盘上的点的概率与该圆盘的面积成正比, 并设射击都能中靶, 以  $X$  表示弹着点与圆心的距离. 试求随机变量  $X$  的分布律.



解. 根据题意, 以圆盘的圆心为原点在圆盘上建立极坐标系  $(x, \theta)$ , 则可以得出:

$$F(x) = P\{X \leq x\} = \frac{\int_0^x 2\pi x dx}{\int_0^R 2\pi x dx} = \frac{x^2}{R^2} = \frac{x^2}{4}.$$

结合  $x$  在实数域上的所有取值, 可以得到:

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x^2}{4} & 0 \leq x \leq 2 \\ 1 & \text{others.} \end{cases}$$

□

例 2.4.2. 设随机变量  $X$  具有概率密度函数

$$f(x) = \begin{cases} kx & 0 \leq x < 3 \\ 2 - \frac{x}{2} & 3 \leq x \leq 4 \\ 0 & \text{others.} \end{cases}$$

- (1) 确定常数  $k$ ;
- (2) 求  $X$  的分布函数  $F(x)$ ;
- (3) 求  $P\{1 < x < \frac{7}{2}\}$ .

解.

- (1)  $f(x)$  为概率密度函数需满足  $\int_{-\infty}^{\infty} f(x) dx = 1$ , 因此有

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x) dx &= \int_0^3 kx dx + \int_3^4 2 - \frac{x}{2} dx = 1 \\ &\Rightarrow \frac{9k}{2} = \frac{3}{4} \\ &\Rightarrow k = \frac{1}{6} \end{aligned}$$

- (2) 根据分布函数和概率密度函数之间的关系, 可以得出

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x^2}{12} & 0 \leq x < 3 \\ 2x - 3 - \frac{x^2}{4} & 3 \leq x < 4 \\ 1 & 4 \leq x. \end{cases}$$

(3) 根据概率分布函数的定义有

$$P\left\{1 < x < \frac{7}{2}\right\} = F\left(\frac{7}{2}\right) - F(1) = \frac{15}{16} - \frac{1}{12} = \frac{41}{48}.$$

□

## 2.4.2 重要的连续型随机变量分布

### 均匀分布

定义 2.4.2.  $X$  的概率密度为:

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{others.} \end{cases}$$

称  $X$  在区间  $(a, b)$  服从均匀分布, 记作  $X \sim U(a, b)$ .

注. 根据分布函数和概率密度函数之间的关系, 可以得出均匀分布的概率分布函数

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

例 2.4.3. 设电阻值  $R$  为随机变量, 均匀分布在  $900 \Omega \sim 1100 \Omega$ . 求  $R$  的概率密度函数及落在  $950 \Omega \sim 1050 \Omega$  的概率.

解. 由均匀分布可以得出  $R$  的概率密度函数为

$$f(r) = \begin{cases} \frac{1}{200} & 900 < r < 1100 \\ 0 & \text{others.} \end{cases}$$

所求区间内概率为

$$P\{950 \leq R \leq 1050\} = \int_{950}^{1050} \frac{1}{200} dr = \frac{1}{2}.$$

□

例 2.4.4. 若随机变量  $Y$  在  $(1, 6)$  上服从均匀分布, 求方程  $x^2 + Yx + 1 = 0$  有实根的概率.

解. 随机变量有实根需满足

$$Y^2 - 4 \geq 0 \implies Y^2 \geq 4 \implies Y \geq 2$$

根据均匀分布可以得出随机变量  $Y$  的概率密度函数

$$f(y) = \begin{cases} \frac{1}{5} & 1 < x < 6 \\ 0 & \text{others.} \end{cases}$$

由此得出  $Y \geq 2$  的概率为

$$P\{Y \geq 2\} = 1 - P\{Y \leq 2\} = 1 - \int_1^2 \frac{1}{5} dy = \frac{4}{5}.$$

□

**例 2.4.5.** 设随机变量  $X$  在区间  $(2, 5)$  上服从均匀分布, 现对  $X$  进行三次独立观测, 则至少有两次观测值大于 3 的概率.

解. 根据均匀分布, 可以得出随机变量  $X$  的概率密度函数

$$f(x) = \begin{cases} \frac{1}{3} & 2 < x < 5 \\ 0 & \text{others.} \end{cases}$$

由此得出每次观测, 值大于 3 的概率为

$$p = \int_3^5 \frac{1}{3} dx = \frac{2}{3}.$$

设随机变量  $Y$  为三次观测中值大于 3 的次数. 则  $Y \sim B(3, \frac{2}{3})$ , 可以得出

$$P\{Y \geq 2\} = P\{Y = 2\} + P\{Y = 3\} = C_3^2 \cdot p^2 (1-p) + p^3 = \frac{20}{27}.$$

□

## 指数分布

**定义 2.4.3.** 随机变量  $X$  的概率密度函数为

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

其中参数  $\lambda > 0$ . 称  $X$  服从指数分布, 记作  $X \sim E(\lambda)$ .

注. 指数分布的概率分布函数并不能简单通过积分求得, 可以记忆一下:

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

**定理 2.4.1.** 指数分布的无记忆性 对于任意  $s > 0, t > 0$ , 有

$$P\{X > t\} = P\{X > s + t | X > s\}.$$

注. 对于无记忆性, 可以这样理解: 对一个没有明显老化的产品, 设随即变量  $X$  表示产品的寿命. 则上述定理等式左边表示使用时间大于  $t$  的概率; 右边表示在产品已经用了  $s$  时间的基础上, 还能用的时间大于  $t$  的概率.

**证明.** 指数分布的无记忆性 对于等式的左边

$$P\{X > t\} = 1 - P\{X \leq t\} = 1 - F(t) = 1 - (1 - e^{-\lambda t}) = e^{-\lambda t}.$$

对于等式的右边有

$$\begin{aligned} P\{X > s + t | X > s\} &= \frac{P\{(X > s + t) \cap (X > s)\}}{P\{X > s\}} = \frac{P\{X > s + t\}}{P\{X > s\}} = \frac{1 - F(s + t)}{1 - F(s)} \\ &= \frac{1 - (1 - e^{-\lambda(s+t)})}{1 - (1 - e^{-\lambda s})} = e^{-\lambda t}. \end{aligned}$$

所以左右两式相等. □

**例 2.4.6.** 设打一次电话所用时间  $X$  (min) 服从参数  $\lambda = 0.1$  的指数分布, 如果某人刚好在你前面走进电话亭, 求你等待的时间:

- (1) 超过 10 分钟的概率;
- (2) 在 10 分钟到 20 分钟之间的概率.

**解.** 由随机变量  $X \sim E(\lambda = 0.1)$  可以得出

$$f(x) = \begin{cases} 0.1e^{-0.1x} & x > 0 \\ 0 & \text{others.} \end{cases}$$

因此可以计算概率

(1)

$$P\{X > 10\} = 1 - P\{X \leq 10\} = 1 - F(10) = 1 - (1 - e^{-1}) = \frac{1}{e}.$$

(2)

$$P\{10 < X \leq 20\} = F(20) - F(10) = (1 - e^{-2}) - (1 - e^{-1}) = e^{-1} - e^{-2}.$$

□

### 正态分布

定义 2.4.4. 若连续型随机变量  $X$  的概率密度为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad -\infty < x < +\infty.$$

其中参数  $\mu$  和  $\sigma$  都是大于 0 的常数. 称这样的  $X$  服从参数为  $\mu$  和  $\sigma$  的正态分布或高斯 (Gauss) 分布, 记作  $X \sim N(\mu, \sigma^2)$ .

注. 若正态分布的参数满足

1.  $\mu = 0$ ;
2.  $\sigma = 1$ .

则称分布为标准正态分布. 其概率密度函数为

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\}.$$

正态分布的概率分布函数为

$$F(x) = P\{X \leq x\} = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(t-\mu)^2}{2\sigma^2}\right\} dt.$$

$$\text{当 } \begin{cases} \mu = 0 \\ \sigma = 1 \end{cases} \text{ 时有 } \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{t^2}{2}\right\} dt.$$

注. 正态分布的性质:

1. 概率密度曲线关于  $X = \mu$  对称;
2.  $f(x)_{max} = f(\mu) = \frac{1}{\sqrt{2\pi}\sigma}$ ;
3.  $P\{\mu - h < X \leq \mu\} = P\{\mu < X \leq \mu + h\}$ ;

4. 对于标准正态分布, 有  $\Phi(x) + \Phi(-x) = 1$ .

引理 2.4.2. 设随机变量  $X \sim N(\mu, \sigma^2)$ ,  $Z = \frac{X-\mu}{\sigma}$ , 则  $Z \sim N(0, 1)$ .

证明. 设  $F_Z(x)$  为随机变量  $Z$  的概率分布函数, 则

$$\begin{aligned} F_Z(x) &= P\{Z \leq x\} = P\left\{\frac{X-\mu}{\sigma} \leq x\right\} = P\{X \leq x\sigma + \mu\} \\ &= \int_{-\infty}^{\mu+x\sigma} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(t-\mu)^2}{2\sigma^2}\right\} dt \quad \text{令 } u = \frac{t-\mu}{\sigma} \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x \exp\left\{-\frac{u^2}{2}\right\} du \\ &= \Phi(x). \end{aligned}$$

□

例 2.4.7. 设随机变量  $X$  服从正态分布  $N(\mu, \sigma^2)$ , 则随  $\sigma$  增大, 概率  $P\{|X-\mu| < \sigma\}$  如何变化?

解. 由于

$$\begin{aligned} P\{|X-\mu| < \sigma\} &= P\{-\sigma < X-\mu < \sigma\} = P\left\{-\frac{\sigma}{\sigma} < \frac{X-\mu}{\sigma} < \frac{\sigma}{\sigma}\right\} \\ &= P\left\{-1 < \frac{X-\mu}{\sigma} < 1\right\} = \Phi(1) - \Phi(-1) = \text{const.} \end{aligned}$$

所以保持不变.

□

例 2.4.8. 假设测量的随机误差  $X \sim N(0, 10^2)$ , 求在 100 次独立重复测量中, 至少有三次测量的误差的绝对值大于 19.6 的概率  $\alpha$ , 并利用泊松分布求出  $\alpha$  的近似值.

解. 每一次测量的随机误差  $X$  服从正态分布, 设  $Y = \frac{X}{10}$  服从标准正态分布, 则

$$\begin{aligned} P\{|X| > 19.6\} &= P\{X < -19.6, X > 19.6\} = P\{Y < -1.96, Y > 1.96\} \\ &= \Phi(-1.96) + (1 - \Phi(1.96)) = 1 - 0.9750 + 1 - 0.9750 = 0.05. \end{aligned}$$

由此得出每次测量误差绝对值大于 19.6 的概率  $p = 0.05$ .

设随机变量  $Z$  表示 100 次试验中, 测量误差绝对值大于 19.6 的次数, 则  $Z \sim B(100, 0.05)$

$$\begin{aligned} P\{Z \geq 3\} &= 1 - P\{Z < 3\} = 1 - \sum_{i=0}^2 P\{Z = i\} \\ &= 1 - \left[(1-p)^{100} + C_{100}^1 \cdot p(1-p)^{99} + C_{100}^2 \cdot p^2(1-p)^{98}\right] \\ &= 0.882. \end{aligned}$$

根据泊松定理, 随机变量  $Z$  可以用泊松分布近似, 由于  $n \cdot p = 5$ , 因此  $Z \sim \pi(5)$ , 由此可得

$$\begin{aligned}\alpha = P\{Z \geq 3\} &= 1 - \sum_{i=0}^2 P\{Z = i\} \\ &= 1 - e^{-5} \left(1 + 5 + \frac{5^2}{2}\right) = 0.875.\end{aligned}$$

□

注.  $3\sigma$  法则:  $X \sim N(\mu, \sigma^2) \Rightarrow F(x) = P\{X \leq x\} = P\left\{\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right\} = \Phi\left(\frac{x-\mu}{\sigma}\right)$ .

$$\begin{aligned}P\{x_1 < X \leq x_2\} &= P\left\{\frac{x_1-\mu}{\sigma} < \frac{X-\mu}{\sigma} \leq \frac{x_2-\mu}{\sigma}\right\} \\ &= P\left\{\frac{X-\mu}{\sigma} \leq \frac{x_2-\mu}{\sigma}\right\} - P\left\{\frac{X-\mu}{\sigma} \leq \frac{x_1-\mu}{\sigma}\right\} \\ &= \Phi\left(\frac{x_2-\mu}{\sigma}\right) - \Phi\left(\frac{x_1-\mu}{\sigma}\right).\end{aligned}$$

⇒

$$\begin{aligned}P\{\mu - \sigma < X \leq \mu + \sigma\} &= \Phi\left(\frac{\mu + \sigma - \mu}{\sigma}\right) - \Phi\left(\frac{\mu - \sigma - \mu}{\sigma}\right) \\ &= \Phi(1) - \Phi(-1) = \Phi(1) - (1 - \Phi(1)) = 2\Phi(1) - 1 \\ &= 68.26\%.\end{aligned}$$

$$P\{\mu - 2\sigma < X \leq \mu + 2\sigma\} = 2\Phi(2) - 1 = 95.44\%.$$

$$P\{\mu - 3\sigma < X \leq \mu + 3\sigma\} = 2\Phi(3) - 1 = 99.74\%.$$

定义 2.4.5. 设随机变量  $X \sim N(0, 1)$ , 若  $Z_\alpha$  满足

$$P\{X > Z_\alpha\} = \alpha, \quad 0 < \alpha < 1.$$

则  $Z_\alpha$  是  $X$  的上  $\alpha$  分位点.

注. 上  $\alpha$  分位点指的是: 在标准正态分布图像中, 满足  $X = Z_\alpha$  右侧的面积等于  $\alpha$  的  $X$  的值. 且  $-Z_\alpha = Z_{1-\alpha}$ .

例 2.4.9. 将一温度调节器放置在贮存着某种液体的容器内, 调节器整定在  $d^\circ\text{C}$ , 液体的温度  $X$  是一个随机变量, 且  $X \sim N(d, 0.5^2)$ .

(1) 若  $d = 90^\circ\text{C}$ , 求  $X$  小于  $89^\circ\text{C}$  的概率.

(2) 若要求保持液体的温度至少为  $80^\circ\text{C}$  的概率不低于 0.99, 求  $d$  至少为多少.

解.

(1) 若  $d = 90^\circ\text{C}$ , 则

$$P\{X \leq 89\} = P\left\{\frac{X-d}{0.5} \leq \frac{89-90}{0.5}\right\} = \Phi(-2) = 0.0228.$$

(2) 保持液体温度至少为  $80^\circ\text{C}$ , 意味  $X \geq 80$ , 要求  $P\{X \geq 80\} \geq 0.99$ , 则

$$\begin{aligned} P\{X \geq 80\} &= 1 - P\{X \leq 80\} + P\{X = 80\} = 1 - P\{X \leq 80\} \\ &= 1 - P\left\{\frac{X-d}{0.5} \leq \frac{80-d}{0.5}\right\} = 1 - \Phi\left(\frac{80-d}{0.5}\right) \geq 0.99 \\ &\Rightarrow \Phi\left(\frac{80-d}{0.5}\right) \leq 0.01 \Rightarrow \Phi\left(\frac{d-80}{0.5}\right) \geq 0.99 \\ &\Rightarrow \frac{d-80}{0.5} \geq 2.327 \Rightarrow 81.1625^\circ\text{C}. \end{aligned}$$

□

**例 2.4.10.** 某单位招聘 155 人, 按考试成绩录用, 共有 526 人报名, 假设报名者的考试成绩  $X \sim N(\mu, \sigma^2)$ . 已知 90 分以上的 12 人, 60 分以下的 83 人, 若从高到低依次录取, 某人成绩为 78 分, 问此人能否被录取.

解. 根据信息, 录取率  $p = \frac{155}{526} = 29.47\%$ , 先计算正态分布的参数. 已知 90 分以上的有 12 人, 因此

$$\begin{aligned} P\{X \leq 90\} &= \Phi\left(\frac{90-\mu}{\sigma}\right) = 1 - P\{X \geq 90\} = 1 - \frac{12}{526} = 0.9772 \\ &\Rightarrow \frac{90-\mu}{\sigma} = 2. \end{aligned}$$

又因为 60 分以下的有 83 人, 所以

$$\begin{aligned} P\{X \leq 60\} &= \Phi\left\{\frac{60-\mu}{\sigma}\right\} = \frac{83}{526} = 15.78\% = 1 - \Phi\left(\frac{\mu-60}{\sigma}\right) \\ &\Rightarrow \frac{\mu-60}{\sigma} = 1. \end{aligned}$$

综合上述, 可以解得

$$\begin{cases} \mu = 70 \\ \sigma = 10. \end{cases}$$



因此  $X \sim N(70, 100)$ .

根据录取规则, 若该应聘者被录取, 则它的成绩应在前 29.47%. 即  $P\{X > 78\} \leq 29.47\%$ , 于是根据

$$P\{X > 78\} = 1 - \Phi\left\{\frac{8}{10}\right\} = 21.19\% \leq 29.47\%.$$

因此可以被录取. □

## 2.5 随机变量函数的分布

### 2.5.1 分布函数法

例 2.5.1. 设随机变量  $X$  具有以下分布律, 试求  $Y = (X - 1)^2$  的分布律.

$X$	-1	0	1	2
$p_k$	0.2	0.3	0.1	0.4

解. 根据随机变量  $X$  与  $Y$  的关系, 可以解得  $Y$  的取值为  $\{0, 1, 4\}$ .

$$P\{Y = 0\} = P\{(X - 1)^2 = 0\} = P\{X = 1\} = 0.1$$

$$P\{Y = 1\} = P\{(X - 1)^2 = 1\} = P\{X = 0 \cup X = 2\} = P\{X = 0\} + P\{X = 2\} = 0.7$$

$$P\{Y = 4\} = P\{(X - 1)^2 = 4\} = P\{X = -1\} = 0.2.$$

因此  $Y$  得分布律为

$Y$	0	1	4
$p_k$	0.1	0.7	0.2

□

例 2.5.2. 随机变量  $X$  具有得概率密度如下

$$f_X(x) = \begin{cases} \frac{x}{8} & 0 < x < 4 \\ 0 & \text{others.} \end{cases}$$

求随机变量  $Y = 2X + 8$  得概率密度函数.

解. 先求随机变量  $Y$  的概率分布函数.

$$\begin{aligned} F_Y(y) &= P\{Y \leq y\} = P\{2X + 8 \leq y\} = P\left\{X \leq \frac{y-8}{2}\right\} \\ &= \int_0^{\frac{y-8}{2}} \frac{x}{8} dx = \frac{1}{64}y^2 - \frac{1}{4}y + 1, \quad 8 < y < 16. \end{aligned}$$

由此得出

$$F_Y(y) = \begin{cases} 0 & y < 8 \\ \frac{1}{64}y^2 - \frac{1}{4}y + 1 & 8 \leq y < 16 \\ 1 & 16 \leq y. \end{cases}$$

通过概率分布函数和概率密度函数的关系, 可求得随机变量  $Y$  的概率密度函数.

$$f_Y(y) = \begin{cases} \frac{1}{32}y - \frac{1}{4} & 8 < y < 16 \\ 0 & \text{others.} \end{cases}$$

□

**例 2.5.3.** 设随机变量  $X$  具有概率密度函数  $f_x(x)$ , 求  $Y = X^2$  的概率密度函数.

解. 先求  $Y$  的概率分布函数.

$$F_Y(y) = P\{Y \leq y\} = P\{X^2 \leq y\} = P\{-\sqrt{y} \leq X \leq \sqrt{y}\} = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

再根据概率分布函数和概率密度函数之间的关系, 求得  $Y$  得概率密度函数.

$$\begin{aligned} f_Y(y) &= \frac{dF_Y(y)}{dy} = \frac{d(F_X(\sqrt{y}) - F_X(-\sqrt{y}))}{dy} \\ &= \frac{dF_X(\sqrt{y})}{dy} - \frac{dF_X(-\sqrt{y})}{dy} \\ &= \frac{1}{2\sqrt{y}} \cdot (f_x(\sqrt{y}) + f_x(-\sqrt{y})). \end{aligned}$$

⇒

$$f_y(y) = \begin{cases} 0 & y \leq 0 \\ \frac{1}{2\sqrt{y}} \cdot (f_x(\sqrt{y}) + f_x(-\sqrt{y})) & 0 < y. \end{cases}$$

□

例 2.5.4. 设随机变量  $X$  有如下概率密度函数

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

求随机变量  $Y = e^X$  的概率密度函数  $f_Y(y)$ .

解. 先求出随机变量  $Y$  的取值范围. 由  $X \geq 0$  得出  $Y \geq 1$ . 再有

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

又因为  $f_Y(y) = F'_Y(y)$

$$f_Y(y) = \frac{1}{y} f_X(\ln y) = \begin{cases} \frac{1}{y^2} & 1 \leq y \\ 0 & y < 1. \end{cases}$$

□

## 2.5.2 定理法

定理 2.5.1. 设随机变量  $X$  具有概率密度  $f_X(x)$ ,  $-\infty < x < +\infty$ , 又设函数  $g(x)$  处处可导且有  $g(x)$  单调, 则  $Y = g(X)$  是连续型随机变量, 其概率密度为

$$f_Y(y) = \begin{cases} f_X[h(y)] \cdot |h'(y)| & \alpha < y < \beta \\ 0 & \text{others.} \end{cases}$$

其中  $\alpha = \min\{g(-\infty), g(+\infty)\}$ ,  $\beta = \max\{g(-\infty), g(+\infty)\}$ ,  $h(y)$  是  $g(x)$  的反函数.

证明. 简单证明 先计算  $Y$  的概率密度函数.

$$F_Y(y) = P\{Y \leq y\} = P\{g(x) \leq y\} = P\{X \leq h(y)\} = F_X(h(y)).$$

再由概率分布函数和概率密度函数的关系, 可以得出

$$f_Y(y) = F'_Y(y) = f_X(h(y)) \cdot h'(y).$$

绝对值只是因为反函数求导可能正可能负, 这与  $g(x)$  的单调性有关, 该定理只是分布函数法的理论概括. □

例 2.5.5. 设随机变量  $X$  有如下概率密度函数

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

求随机变量  $Y = e^X$  的概率密度函数  $f_Y(y)$ .

解. 运用定理, 先找定理中的元素.  $g(x) = e^x$  单调且处处可导,  $Y = g(X) = e^X$ , 得出  $X = \ln Y$ . 因此  $g(x)$  的反函数为  $h(y) = \ln y$ .  $\alpha = g(0) = 1$ ,  $\beta = g(+\infty) = +\infty$ .

$$f_Y(y) = \begin{cases} f_X(h(y)) \cdot |h'(y)| = \frac{1}{y^2} & 1 < y < +\infty \\ 0 & \text{others.} \end{cases}$$

□

# 第三章 多维随机变量及其分布

## 3.1 二维随机变量

### 3.1.1 二维随机变量及其联合分布函数

定义 3.1.1. 试验  $E$  的样本空间为  $\Omega = \{e\}$ ,  $e$  由  $(X, Y)$  构成.  $X, Y$  是定义在  $\Omega$  里的随机变量. 称  $(X, Y)$  为二维随机变量.

定义 3.1.2. 称

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

为随机变量  $X$  和  $Y$  的联合分布函数.

注. 通常记  $Z = (X, Y)$ , 此时随机变量  $Z$  就是一个二维随机变量.

性质

1.  $F(x, y)$  是一个不减函数;
2.  $0 \leq F(x, y) \leq 1$ ; 特别的有:  
 $F(-\infty, y) = F(x, -\infty) = F(-\infty, -\infty) = 0$ ;  
 $F(+\infty, +\infty) = 1$ ;
3.  $F(x, y)$  是关于  $x$  或  $y$  的右连续函数.

例 3.1.1. 用分布函数表示如下概率:

$$P\{x_1 < X \leq x_2, y_1 < Y \leq y_2\}.$$

解. 根据分布函数的定义:

$$\begin{aligned} & P\{x_1 < X \leq x_2, y_1 < Y \leq y_2\} \\ &= P\{X \leq x_2, Y \leq y_2\} - P\{X \leq x_1, Y \leq y_2\} \\ &\quad - P\{X \leq x_2, Y \leq y_1\} + P\{X \leq x_1, Y \leq y_1\} \\ &= F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1) \end{aligned}$$

□

注. 上例可以从几何的角度思考. 概率  $P\{x_1 < X \leq x_2, y_1 < Y \leq y_2\}$  表示的概率在二维实数平面上为一个由  $x_1, x_2, y_1, y_2$  围成的矩形, 只需要利用满足定律的方式将对应的举行表示出来即可.

### 3.1.2 离散型二维随机变量及联合分布律

定义 3.1.3. 随机变量  $Z = (X, Y)$  所有的取值为有限对或无限可列对, 称为离散型二维随机变量.

定义 3.1.4. 离散型二维随机变量的联合分布率:

$$P\{X = x_i, Y = y_j\}, \quad i, j = 1, 2, 3, \dots$$

性质

1.  $P_{i,j} \geq 0$ ;
2.  $\sum P_{i,j} = 1$ .

例 3.1.2. 随机变量  $X$  在 1, 2, 3, 4 中等可能地取一个值, 另一个随机变量  $Y$  在  $1 \sim X$  中等可能地取一个整数值. 求  $(X, Y)$  的联合分布率.

解. 随机变量  $X$  其取值可以为 1, 2, 3, 4, 只有当  $X$  的取值确定时,  $Y$  的取值范围才能确定. 因此根据定义有:

$$P\{X = i, Y = j\} = \frac{1}{4} \times \frac{1}{i} = \frac{1}{4 \cdot i} \quad i = 1, 2, 3, 4; j = 1, \dots, i$$

由此可以画出表格.

□

## 3.1.3 连续型二维随机变量及联合概率密度

**定义 3.1.5.** 若二维随机变量  $(X, Y)$  的联合分布函数为  $F(x, y)$ , 存在  $f(x, y)$  非负可积, 使得:

$$F(x, y) = \int_{-\infty}^y \int_{-\infty}^x f(u, v) du, dv$$

称  $X, Y$  为连续型二维随机变量. 函数  $f(x, y)$  为  $(X, Y)$  的联合概率密度.

性质

1.  $f(x, y) \geq 0$ ;
2.  $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) = 1$ ;
3. 设  $G$  是  $\{x, O, y\}$  平面上的某个区域, 则  $(X, Y)$  落在  $G$  区域的概率

$$P\{(X, Y) \in G\} = \iint_G f(x, y) dx dy.$$

4.  $f(x, y)$  在  $(x, y)$  上连续:

$$\frac{\partial^2 F(x, y)}{\partial x \partial y} = \frac{\partial^2 F(x, y)}{\partial y \partial x} = f(x, y).$$

**例 3.1.3.** 设二维随机变量  $(X, Y)$  具有概率密度

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

- (1) 求分布函数  $F(x, y)$ ;
- (2) 求概率  $P\{Y \leq X\}$ .

解.

- (1) 根据联合分布函数的定义:

$$\begin{aligned} F(x, y) &= P\{X \leq x, Y \leq y\} \\ &= \int_{-\infty}^y \int_{-\infty}^x f(u, v) du dv \\ &= \int_0^y \int_0^x 2e^{-(2u+v)} du dv \\ &= \int_0^y e^{-v} (1 - e^{-2x}) dv \\ &= (1 - e^{-2x})(1 - e^{-y}) \end{aligned}$$

由此得出联合分布函数:

$$F(x, y) = \begin{cases} (1 - e^{-2x})(1 - e^{-y}) & x > 0, y > 0 \\ 0 & \text{others.} \end{cases}$$

(2) 根据定义, 可以得出:

$$\begin{aligned} P\{Y \leq X\} &= \int_{-\infty}^{+\infty} \int_y^{+\infty} f(x, y) dx dy \\ &= \int_0^{+\infty} \int_y^{+\infty} 2e^{-(2x+y)} dx dy \\ &= \int_0^{+\infty} e^{-3y} dy \\ &= \frac{1}{3} \end{aligned}$$

□

例 3.1.4. 设二维随机变量  $(X, Y)$  的概率密度为

$$f(x, y) = \begin{cases} 6x & 0 \leq x \leq y \leq 1 \\ 0 & \text{others.} \end{cases}$$

求  $P\{X + Y \leq 1\}$ .

解. 根据概率密度函数, 可以确定概率非 0 的区域为直线  $y = x$  与  $y = 1$  和  $x = 0$  围成的区域. 要求的概率为:

$$\begin{aligned} P\{X + Y \leq 1\} &= P\{Y \leq 1 - X\} \\ &= \int_0^{\frac{1}{2}} \int_x^{1-x} 6x dy dx \\ &= \int_0^{\frac{1}{2}} 6x - 12x^2 dx \\ &= \frac{1}{4} \end{aligned}$$

□

注. 上例的积分区间需根据概率密度函数的定义域得到, 利用作图确定最后的积分区域.



例 3.1.5. 已知随机变量  $X$  和  $Y$  的联合概率密度为

$$\varphi(x, y) = \begin{cases} 4xy & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{others.} \end{cases}$$

求  $X$  和  $Y$  的联合分布函数  $F(x, y)$ .

解. 概率密度的非零区域是一个矩形, 根据分布函数的定义:

$$\begin{aligned} F(x, y) &= P\{X \leq x, Y \leq y\} \\ &= \int_0^y \int_0^x 4xy \, dx \, dy \\ &= \int_0^y 2x^2y \, dy \\ &= x^2y^2 \quad (0 \leq x \leq 1, 0 \leq y \leq 1) \end{aligned}$$

$$\begin{aligned} F(x, y) &= P\{X \leq x, Y \leq y\} \\ &= \int_0^1 \int_0^x 4xy \, dx \, dy \\ &= \int_0^1 2x^2y \, dy \\ &= x^2 \quad (0 \leq x \leq 1, y > 1) \end{aligned}$$

$$\begin{aligned} F(x, y) &= P\{X \leq x, Y \leq y\} \\ &= \int_0^1 \int_0^x 4xy \, dx \, dy \\ &= \int_0^1 2x^2y \, dy \\ &= y^2 \quad (0 \leq y \leq 1, x > 1) \end{aligned}$$

由此得出分布函数为

$$F(x, y) = \begin{cases} x^2y^2 & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 1 & x > 1, y > 1 \\ x^2 & y > 1, 0 \leq x \leq 1 \\ y^2 & x > 1, 0 \leq y \leq 1 \\ 0 & x < 0, y < 0 \end{cases}$$

□

## 3.2 边缘分布

### 3.2.1 边缘分布函数

**定义 3.2.1.** 设二维随机变量  $(X, Y)$  的联合分布函数为  $F(x, y)$ . 若有:

1. 随机变量  $X$  的概率分布函数为  $F_X(x)$ ;
2. 随机变量  $Y$  的概率分布函数为  $F_Y(y)$ .

则称  $F_X$  和  $F_Y$  为二维随机变量  $(X, Y)$  的边缘分布函数.

**注.** 边缘分布函数可以由联合分布函数确定:

$$F_X(x) = P\{X \leq x\} = P\{(X \leq x) \cap \Omega\} = P\{(X \leq x) \cap (Y \leq +\infty)\} = F(x, +\infty).$$

同理可以得到:

$$F_Y(y) = F(+\infty, y).$$

### 3.2.2 离散型随机变量的边缘分布率

**例 3.2.1.** 一整数  $N$  等可能地在  $1, 2, 3, \dots, 10$  十个值中取一个值. 设  $D = D(N)$  是能整除  $N$  的正整数的个数,  $F = F(N)$  是能整除  $N$  的素数的个数. 试写出  $D$  和  $F$  的联合分布律. 并求边缘分布律.

**解.** 根据题意:

$N$	1	2	3	4	5	6	7	8	9	10
$D$	1	2	2	3	2	4	2	4	3	4
$F$	0	1	1	1	1	2	1	1	1	2

由此得到  $(D, F)$  的联合分布律和边缘分布率

	$D$	1	2	3	4	$F$
$F$	0	$\frac{1}{10}$	0	0	0	$\frac{1}{10}$
	1	0	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{1}{10}$	$\frac{3}{10}$
	2	0	0	0	$\frac{1}{5}$	$\frac{1}{5}$
	$D$	$\frac{1}{10}$	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{3}{10}$	

□

## 3.2.3 连续型随机变量边缘概率密度

根据边缘概率分布函数:

$$F_X(x) = F(x, +\infty).$$

可以得到:

$$F_X(x) = \int_{-\infty}^x \int_{-\infty}^{+\infty} f(t, y) dy dt.$$

其中  $f(x, y)$  是二维随机变量  $(X, Y)$  的联合概率密度.

接下来求随机变量  $X$  的边缘概率密度:

$$f_X(x) = \frac{dF_X(x)}{dx} = \int_{-\infty}^{+\infty} f(x, y) dy.$$

同理可得:

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \int_{-\infty}^{+\infty} f(x, y) dx.$$

例 3.2.2. 设随机变量  $X$  和  $Y$  具有联合概率密度:

$$f(x, y) = \begin{cases} 6 & x^2 \leq y \leq x \\ 0 & \text{others.} \end{cases}$$

求边缘概率密度  $f_X(x)$  和  $f_Y(y)$ .

解. 根据概率密度函数, 可以确定概率密度非零的区域为直线  $y = x$  与曲线  $y = x^2$  围成的区域, 因此:

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{+\infty} f(x, y) dy \\ &= \int_{x^2}^x 6 dy \\ &= 6x - 6x^2, 0 \leq x \leq 1. \end{aligned}$$

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{+\infty} f(x, y) dx \\ &= \int_y^{\sqrt{y}} 6 dx \\ &= 6\sqrt{y} - 6y, 0 \leq y \leq 1. \end{aligned}$$

□

例 3.2.3. 设二维随机变量  $(X, Y)$  的概率密度为:

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

求  $(X, Y)$  的边缘概率密度  $f_X(x)$  和  $f_Y(y)$ .

解. 根据联合概率密度函数, 可以确定概率密度非零的区域为  $y = 0$ ,  $x = 1$  和  $y = 2x$  围城的三角形区域, 因此:

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{+\infty} f(x, y) dy \\ &= \int_0^{2x} 1 dy \\ &= 2x, 0 < x < 1 \end{aligned}$$

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{+\infty} f(x, y) dx \\ &= \int_{\frac{y}{2}}^1 1 dx \\ &= \frac{2-y}{2}, 0 < y < 2 \end{aligned}$$

□

### 3.2.4 常用二维分布

#### 二维均匀分布

回顾一维随机变量的均匀分布, 我们说当一维随机变量  $X$  服从均匀分布, 有  $X = x$  在实数轴  $x$  上任意一点的  $f(x)$  都相等, 且有  $\int f(x) dx = 1$ . 即  $y = f(x)$  与  $x$  围城的面积为 1. 因此若  $X$  的取值范围为  $x_1 = a \leq x_2 = b$ , 则随机变量  $X$  的概率密度  $f(x) = \frac{1}{b-a}$ , 从而使  $\int f(x) = (b-a) f(x) = 1$  成立.

类比一维随机变量, 二维随机变量  $(X, Y)$  服从均匀分布, 意味着以  $X, Y$  为坐标构成的点落在二维平面内的概率为常数. 若点  $(X, Y)$  只会落在区域  $G$  中, 那么  $f(x, y) = \frac{1}{|G|}$ , 其中  $|G|$  表示区域  $G$  的面积. 由此得出  $\iint f(x, y) = |G| \cdot f(x, y) = 1$ . 即以区域  $G$  为底面,  $f(x, y)$  为高的物体体积为 1.

二维均匀分布的概率密度函数表示为:

$$f(x, y) = \begin{cases} \frac{1}{|G|} & (x, y) \in G \\ 0 & \text{others.} \end{cases}$$

### 二维正态分布

注. 等待补充

注. 联合概率密度可以确定边缘概率密度, 但是边缘概率密度无法确定联合概率密度.

## 3.3 条件分布

### 3.3.1 离散型二维随机变量的条件分布律

例 3.3.1. 在一汽车工厂中, 一辆汽车有两道工序是由机器人完成的. 其一是紧固 3 只螺栓, 其二是焊接 2 处焊点. 以  $X$  表示由机器人紧固的螺栓紧固的不良的数目;  $Y$  表示由机器人焊接的不良焊点的数目. 据积累的资料知  $(X, Y)$  具有以下分布律:

$Y \backslash X$	0	1	2	3	$P\{Y = j\}$
0	0.840	0.030	0.020	0.010	0.900
1	0.060	0.010	0.008	0.002	0.080
2	0.010	0.005	0.004	0.001	0.020
$P\{X = i\}$	0.910	0.045	0.032	0.013	1.000

(1) 求在  $X = 1$  的条件下,  $Y$  的条件分布律;

(2) 求在  $Y = 0$  的条件下,  $X$  的条件分布律.

解. 根据条件概率计算公式:

$$P\{Y = i | X = j\} = \frac{P\{Y = i, X = j\}}{P\{X = j\}}.$$

同理:

$$P\{X = i | Y = j\} = \frac{P\{X = i, Y = j\}}{P\{Y = j\}}.$$

因此可以做如下计算:

(1)

$$\begin{aligned}
 P\{Y=0|X=1\} &= \frac{P\{Y=0, X=1\}}{P\{X=1\}} = \frac{0.003}{0.045} = \frac{2}{3} \\
 P\{Y=1|X=1\} &= \frac{P\{Y=1, X=1\}}{P\{X=1\}} = \frac{0.010}{0.045} = \frac{2}{9} \\
 P\{Y=2|X=1\} &= \frac{P\{Y=2, X=1\}}{P\{X=1\}} = \frac{0.005}{0.045} = \frac{1}{9}.
 \end{aligned}$$

(2)

$$\begin{aligned}
 P\{X=0|Y=0\} &= \frac{P\{X=0, Y=0\}}{P\{Y=0\}} = \frac{0.840}{0.900} = \frac{4}{5} \\
 P\{X=1|Y=0\} &= \frac{P\{X=1, Y=0\}}{P\{Y=0\}} = \frac{0.030}{0.900} = \frac{1}{30} \\
 P\{X=2|Y=0\} &= \frac{P\{X=2, Y=0\}}{P\{Y=0\}} = \frac{0.020}{0.900} = \frac{2}{90} \\
 P\{X=3|Y=0\} &= \frac{P\{X=3, Y=0\}}{P\{Y=0\}} = \frac{0.010}{0.900} = \frac{1}{90}.
 \end{aligned}$$

□

**定义 3.3.1.** 设二维离散型随机变量  $(X, Y)$ . 称

$$P\{X=x|Y=y\} = \frac{P\{X=x, Y=y\}}{P\{Y=y\}}.$$

为在随机变量  $Y=y$  的条件下, 随机变量  $X$  的条件分布率.

### 3.3.2 连续型二维随机变量的条件概率密度和条件分布函数

**定义 3.3.2.** 设二维连续型随机变量  $(X, Y)$  具有概率密度  $f(x, y)$ . 而  $(X, Y)$  关于  $Y$  的边缘概率密度为  $f_Y(y)$ . 称

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}.$$

为在  $Y=y$  的条件下,  $X$  的条件概率密度.

得到条件概率密度后, 根据概率密度与分布函数的关系, 可以得到:

$$F_{X|Y}(x|y) = \int_{-\infty}^x \frac{f(x, y)}{f_Y(y)} dx.$$

即在  $Y=y$  的条件下, 随机变量  $X$  的条件分布函数.

**例 3.3.2.** 设  $G$  是平面上的有界区域, 其面积为  $A$ . 若二维随机变量  $(X, Y)$  具有概率密度:

$$f(x, y) = \begin{cases} \frac{1}{A} & (x, y) \in G \\ 0 & \text{others.} \end{cases}$$

则称  $(X, Y)$  在  $G$  上服从均匀分布. 现设二维随机变量  $(X, Y)$  在圆域  $x^2 + y^2 \leq 1$  上服从均匀分布, 求条件概率密度  $f_{X|Y}(x|y)$ .

**解.** 二维随机变量  $(X, Y)$  在  $x^2 + y^2 \leq 1$  上服从均匀分布, 可以得出其概率密度函数为:

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

根据二维连续型随即变量的条件概率密计算公式, 可以得出:

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f(x, y)}{f_Y(y)} \\ &= \frac{\frac{1}{\pi}}{\int_{-\infty}^{+\infty} f(x, y) dx} \\ &= \frac{\frac{1}{\pi}}{\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{1}{\pi} dx} \\ &= \frac{1}{2\sqrt{1-y^2}}, \quad -\sqrt{1-y^2} \leq x \leq \sqrt{1-y^2}. \end{aligned}$$

□

**注.** 在上例的基础上, 可以求概率  $P\{X > \frac{1}{2} | Y = 0\}$ .

$$P\left\{X > \frac{1}{2} | Y = 0\right\} = \int_{\frac{1}{2}}^1 \frac{1}{2} dx = \frac{1}{4}.$$

**例 3.3.3.** 随机变量  $X$  在  $(0, 1)$  上随机等可能地取值, 当观察到  $X = x$  ( $0 < x < 1$ ) 时, 数  $Y$  在区间  $(x, 1)$  上随机等可能地取值. 求  $Y$  的概率密度  $f_Y(y)$ .

**解.**  $Y$  的概率密度  $f_Y(y)$ , 可以通过二维随机变量  $(X, Y)$  的联合概率密度得到. 设随机变量  $(X, Y)$  的联合概率密度为  $f(x, y)$ , 则根据条件概率密度与联合概率密度的关系有:

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} \implies f(x, y) = f_{Y|X}(y|x) \cdot f_X(x).$$

题干指出  $X$  在  $(0 < x < 1)$  上服从均匀分布,  $Y$  在  $(x, 1)$  上服从均匀分布, 因此可以写出  $X$  的边缘概率密度和  $Y$  的条件概率密度:

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

因此二维随机变量  $(X, Y)$  的联合概率密度为:

$$f(x, y) = \begin{cases} \frac{1}{1-x} & 0 < x < y < 1 \\ 0 & \text{others.} \end{cases}$$

再根据联合概率密度和边缘概率密度的关系有:

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{+\infty} f(x, y) dx \\ &= \int_0^y \frac{1}{1-x} dx \\ &= -\ln(1-y), \quad 0 < y < 1 \end{aligned}$$

□

注. 在上例的基础上, 可以尝试计算概率  $P\{Y + X > 1\}$ .

$$\begin{aligned} P\{X + Y > 1\} &= P\{Y > 1 - X\} = \int_{\frac{1}{2}}^1 \int_{1-y}^y f(x, y) dx dy \\ &= \int_{\frac{1}{2}}^1 \int_{1-y}^y \frac{1}{1-x} dx dy \\ &= \int_{\frac{1}{2}}^1 \ln \frac{y}{1-y} dy \\ &= 0.693 \end{aligned}$$



### 3.4 相互独立的随机变量

定义 3.4.1. 设随机变量  $X, Y$ , 如果有

$$P\{X \leq x, Y \leq y\} = P\{X \leq x\} \cdot P\{Y \leq y\}.$$

则称随机变量  $X, Y$  相互独立.

注. 若  $X, Y$  具有联合分布函数  $F(x, y)$  则,  $X, Y$  相互独立等价于:

$$F(x, y) = F_X(x) \cdot F_Y(y).$$

对于连续型相互独立的随机变量  $X, Y$ . 有:

$$f(x, y) = f_X(x) \cdot f_Y(y).$$

对于离散型相互独立的随机变量有:

$$P_{i,j} = P_i \cdot P_j.$$

例 3.4.1. 若  $X, Y$  具有联合分布率如下:

Y \ X	0	1	$P\{Y = j\}$
1	1/6	2/6	1/2
2	1/6	2/6	1/2
$P\{X = i\}$	1/3	2/3	1

判断  $X, Y$  是否相互独立.

解. 根据离散型随机变量的定义, 若  $X, Y$  相互独立, 需满足  $P_{i,j} = P_i \cdot P_j$ :

$$P_{0,1} = \frac{1}{6} = P\{X = 0\} \cdot P\{Y = 1\} = \frac{1}{3} \cdot \frac{1}{2}$$

$$P_{0,2} = \frac{1}{6} = P\{X = 0\} \cdot P\{Y = 2\} = \frac{1}{3} \cdot \frac{1}{2}$$

$$P_{1,1} = \frac{2}{6} = P\{X = 1\} \cdot P\{Y = 1\} = \frac{2}{3} \cdot \frac{1}{2}$$

$$P_{1,2} = \frac{2}{6} = P\{X = 1\} \cdot P\{Y = 2\} = \frac{2}{3} \cdot \frac{1}{2}.$$

因此  $X, Y$  相互独立. □

注. 二维正态随机变量  $X, Y$  相互独立, 则其参数  $\rho = 0$ .

例 3.4.2. 一负责人到达办公室的时间均匀分布在  $8 \sim 12$  时, 他的秘书到达办公室的时间均匀分布在  $7 \sim 9$  时, 设他们两人到达的时间相互独立, 求他们到达办公室的时间相差不超过五分钟的概率.

解. 设随机变量  $X$  为负责人到达办公室的时间,  $Y$  为秘书到达办公室的时间. 要求  $P\{|X - Y| \leq 1/12\}$ . 根据题干, 随机变量  $X, Y$  分别服从连续型的均匀分布, 因此可以得出其各自的概率密度函数:

$$f_X(x) = \begin{cases} \frac{1}{4} & 8 \leq x \leq 12 \\ 0 & \text{others.} \end{cases} \quad f_Y(y) = \begin{cases} \frac{1}{2} & 7 \leq y \leq 9 \\ 0 & \text{others.} \end{cases}$$

又因为  $X, Y$  相互独立, 因此可以得出  $(X, Y)$  的联合概率密度函数:

$$f(x, y) = f_X(x) \cdot f_Y(y) = \begin{cases} \frac{1}{8} & 8 \leq x \leq 12, 7 \leq y \leq 9 \\ 0 & \text{others.} \end{cases}$$

接下来求目标概率:

$$\begin{aligned} P\left\{|X - Y| \leq \frac{1}{12}\right\} &= P\left\{-\frac{1}{12} \leq X - Y \leq \frac{1}{12}\right\} \\ &= \iint_{|X - Y| \leq \frac{1}{12}} \frac{1}{8} dy dx \\ &= \frac{1}{48}. \end{aligned}$$

□

例 3.4.3. 设随机变量  $(X, Y)$  的概率密度为:

$$f(x, y) = \begin{cases} 6xy^2 & 0 < x < 1, 0 < y < 1 \\ 0 & \text{others.} \end{cases}$$

证明  $X$  与  $Y$  相互独立.

解. 若  $X, Y$  相互独立, 应该满足:  $f(x, y) = f_X(x) \cdot f_Y(y)$ . 下面求  $f_X(x)$ :

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{+\infty} f(x, y) dy \\ &= \int_0^1 6xy^2 dy \\ &= 2x, 0 < x < 1. \end{aligned}$$

再求  $f_Y(y)$ :

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{+\infty} f(x, y) dx \\ &= \int_0^1 6xy^2 dx \\ &= 3y^2, 0 < y < 1. \end{aligned}$$

验证:

$$f_X(x) \cdot f_Y(y) = 2x \cdot 3y^2 = 6xy^2 = f(x, y).$$

满足要求, 因此  $X, Y$  相互独立. □

**例 3.4.4.** 一个电子仪器由两个部件构成, 以  $X, Y$  分别表示两个部件的寿命 (千小时). 已知  $X, Y$  的联合分布函数为:

$$F(x, y) = \begin{cases} 1 - e^{-0.5x} - e^{-0.5y} + e^{-0.5(x+y)} & x \geq 0, y \geq 0 \\ 0 & \text{others.} \end{cases}$$

- (1) 问  $X, Y$  是否独立.
- (2) 求两个部件的寿命都超过 100 小时的概率  $\alpha$ .

**解.** 已知联合概率分布函数, 可以得到边缘分布函数:

$$F_X(x) = F(x, +\infty) = 1 - e^{-0.5x}. \quad F_Y(y) = F(+\infty, y) = 1 - e^{-0.5y}.$$

验证独立性:

$$F_X(x) \cdot F_Y(y) = 1 - e^{-0.5x} - e^{-0.5y} + e^{-0.5(x+y)} = F(x, y).$$

成立, 因此  $X, Y$  相互独立.

接下来求  $\alpha$ :

$$\begin{aligned} \alpha &= P\{X > 0.1, Y > 0.1\} \\ &= 1 - F(0.1, +\infty) - F(+\infty, 0.1) + F(0.1, 0.1) \\ &= 1 - 2(1 - e^{-0.05}) + (1 - e^{-0.05} - e^{-0.05} + e^{-0.1}) \\ &= e^{-0.1}. \end{aligned}$$

□

注. 注意到上例中最后  $\alpha$  的计算还可以根据独立性得到:

$$\alpha = P\{X > 0.1, Y > 0.1\} = P\{X > 0.1\} \cdot P\{Y > 0.1\} = (1 - F_X(0.1))(1 - F_Y(0.1)).$$

例 3.4.5. 设随机变量  $X, Y$  相互独立, 他们的密度函数分别为

$$f_X(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & x \leq 0. \end{cases} \quad f_Y(y) = \begin{cases} e^{-y} & y > 0 \\ 0 & y \leq 0. \end{cases}$$

(1)  $(X, Y)$  的联合概率密度.

(2) 求  $P\{X \leq 1 | Y > 0\}$ .

解. 由于随机变量  $X$  和  $Y$  相互独立:

$$f(x, y) = f_X(x) \cdot f_Y(y) \implies f(x, y) = \begin{cases} e^{-(x+y)} & x > 0, y > 0 \\ 0 & \text{others.} \end{cases}$$

接下来求  $P\{X \leq 1 | Y > 0\}$ :

$$\begin{aligned} P\{X \leq 1 | Y > 0\} &= \frac{P\{X \leq 1, Y > 0\}}{P\{Y > 0\}} \\ &= \frac{P\{X \leq 1\} \cdot P\{Y > 0\}}{P\{Y > 0\}} \\ &= P\{X \leq 1\} \\ &= \int_0^1 e^{-x} dx \\ &= 1 - \frac{1}{e} \end{aligned}$$

□

### 3.5 两个随机变量的函数的分布

#### 3.5.1 两个随机变量的线性组合

形如  $Z = X + Y$  的随机变量。若随机变量是离散型：

**例 3.5.1.** 已知随机变量  $(X, Y)$  的联合分布率为：试求  $Z_1 = X + Y$  的分布律。

		$X$		
		1	2	3
$Y$	1	1/5	0	1/5
	2	1/5	1/5	1/5

**解.** 随机变量  $Z_1$  可能的取值为 2, 3, 4, 5, 因此根据分布律的定义：

$$P\{Z_1 = 2\} = P\{X = 1, Y = 1\} = \frac{1}{5}$$

$$P\{Z_1 = 3\} = P\{(X = 1, Y = 2) + (X = 2, Y = 1)\} = \frac{1}{5} + 0 = \frac{1}{5}$$

同理可求得所有的概率，然后得到分布律。 □

若随机变量是连续型。考虑  $Z$  的概率分布函数，根据定义有：

$$\begin{aligned} F_Z(z) &= P\{Z \leq z\} = P\{X + Y \leq z\} \\ &= \iint_G f(x, y) dx dy \end{aligned}$$

其中积分区域  $G$  表示满足  $X + Y \leq z$  的区域，概率密度  $f(x, y)$  是二维随机变量  $(X, Y)$  的联合概率密度。

继续往下推导，可以得到随机变量  $Z$  的概率密度：

$$\begin{aligned} F_Z(z) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{z-y} f(x, y) dx, dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^z f(u-y, y) du dy \quad u = x + y \\ &= \int_{-\infty}^z \int_{-\infty}^{+\infty} f(u-y, y) dy, du \end{aligned}$$

接再来对  $F_Z(z)$  求导：

$$f_Z(z) = \frac{dF_Z}{dz} = \int_{-\infty}^{+\infty} f(z-y, y) dy$$

这就是随机变量  $Z$  的概率密度. 同理可以证明:

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, z-x) dx$$

根据随机变量相互独立的性质, 可以推出当  $X, Y$  相互独立时:

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) \cdot f_Y(y) dy = \int_{-\infty}^{+\infty} f_X(x) \cdot f_Y(z-x) dx$$

这个公式也叫做卷积公式. 记作:  $f_X * f_Y$ .

**例 3.5.2.** (独立同分布) 设  $X$  和  $Y$  是两个相互独立的随机变量. 它们都服从  $N(0, 1)$  分布. 求  $Z = X + Y$  的概率密度.

解. 利用卷积公式:

$$\begin{aligned} f_Z(z) &= f_X(x) * f_Y(y) \\ &= \int_{-\infty}^{+\infty} f_X(x) \cdot f_Y(z-x) dx \end{aligned}$$

带入标准正态分布的概率密度函数:

$$\begin{aligned} f_X(x) \cdot f_Y(z-x) &= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} \cdot \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(z-x)^2}{2}\right\} \\ &= \frac{1}{2\pi} \exp\left\{-\frac{x^2 + (z-x)^2}{2}\right\} \\ &= \frac{1}{2\pi} \exp\left\{-\frac{z^2}{4}\right\} \cdot \exp\left\{-\left(x - \frac{z}{2}\right)^2\right\} \end{aligned}$$

带入卷积公式:

$$\begin{aligned} f_Z(z) &= \frac{1}{2\pi} \exp\left\{-\frac{z^2}{4}\right\} \int_{-\infty}^{+\infty} \exp\left\{-\left(x - \frac{z}{2}\right)^2\right\} dx \\ &= \frac{1}{2\pi} \exp\left\{-\frac{z^2}{4}\right\} \int_{-\infty}^{+\infty} e^{-t^2} dt \quad t = x - \frac{z}{2} \\ &= \frac{1}{2\sqrt{\pi}} \exp\left\{-\frac{z^2}{4}\right\} \end{aligned}$$

□

注. 将上例的结果同正态分布的定义对比:

$$f_Z(z) = \frac{1}{\sqrt{2}\sqrt{2\pi}} \exp\left\{-\frac{(z-0)^2}{2(\sqrt{2})^2}\right\}$$

可以得出  $\sigma_Z = \sqrt{2}$ ,  $\mu_Z = 0$ . 即  $Z \sim N(0, 2) \implies Z \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$ . 跟一般地, 有限个相互独立的正态随机变量的线性组合也服从正态分布.

**例 3.5.3.** 在一简单电路中, 两电阻  $R_1$  和  $R_2$  串联, 设  $R_1 + R_2$  相互独立, 它们的概率密度均为:

$$f(x) = \begin{cases} \frac{10-x}{50} & 0 \leq x \leq 10 \\ 0 & \text{others.} \end{cases}$$

求总电阻  $R = R_1 + R_2$  的概率密度.

**解.** 根据卷积公式;

$$f_R(r) = f_{R_1}(r_1) * f_{R_2}(r_2)$$

接下来讨论  $r_1$  的取值范围. 根据题干可以得到:

$$\begin{cases} 0 \leq r_1 \leq 10 \\ 0 \leq r_2 = r - r_1 \leq 10 \end{cases} \quad \begin{cases} 0 \leq r_1 \leq 10 \\ r - 10 \leq r_1 \leq r \end{cases}$$

因此确定积分区域为由  $r_1 = 0$ ,  $r_1 = 10$ ,  $r_1 = r$ ,  $r_1 = r - 10$  围城的楔形区域. 在这个区域上做卷积得到:

$$f_{R_1} * f_{R_2} = \begin{cases} \frac{1}{15000} (600r - 60r^2 + r^3) & 0 \leq r < 10 \\ \frac{1}{15000} (20 - r)^3 & 10 \leq r < 20 \\ 0 & \text{others.} \end{cases}$$

□

**例 3.5.4.** 设二维随机变量  $(X, Y)$  的概率密度为:

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

求  $Z = 2X - Y$  的概率密度  $f_Z(z)$ .

解. (分布函数法) 要求随机变量  $Z$  的概率密度, 可以先求  $Z$  的分布函数.

$$F_Z(z) = \iint_G f(x, y) dx dy$$

其中积分区域  $G$  是满足  $0 < x < 1, 0 < y < 2x, 2x - y \leq z$  的区域. 由前面的不等式, 可以推出  $0 < z < 2$ .

$$\begin{aligned} F_Z(z) &= \int_0^{\frac{z}{2}} \int_0^{2x} f(x, y) dy dx + \int_{\frac{z}{2}}^1 \int_{2x-z}^{2x} f(x, y) dy dx \\ &= \frac{z^2}{4} + z - \frac{z^2}{2} \\ &= z - \frac{z^2}{4} \end{aligned}$$

由此得  $Z$  的概率密度:

$$F_Z(z) = \begin{cases} 0 & z \leq 0 \\ z - \frac{z^2}{4} & 0 < z < 2 \\ 1 & z \geq 2 \end{cases}$$

接下来对  $z$  求导得出概率密度:

$$f_Z(z) = \begin{cases} 1 - \frac{z}{2} & 0 < z < 2 \\ 0 & \text{others.} \end{cases}$$

□

解. (公式法) 由之前提到的公式, 可以写出:

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, 2x - z) dx$$

根据题干所给信息, 解出  $z$  的取值范围:  $0 < z < 2x$ . 由此得出概率密度  $f(x, 2x - z)$  的非零区域为  $z = 0, z = 2x$  与  $x = 1$  围城的三角形区域. 在这个区域上对  $x$  进行积分:

$$\begin{aligned} f_Z(x) &= \int_{\frac{z}{2}}^1 1 dx \\ &= 1 - \frac{z}{2} \end{aligned}$$

□



## 3.5.2 非线性

**例 3.5.5.** 设二维随机变量  $(X, Y)$  在矩形  $G = \{(x, y) | 0 \leq x \leq 2, 0 \leq y \leq 1\}$  上服从均匀分布, 试求边长为  $X$  和  $Y$  的矩形的面积  $S$  的概率密度.

**解.** 显然要求的随机变量  $S = XY$ . 先求  $S$  的分布函数.

$$F_S(s) = P\{XY \leq s\} = \iint_D f(x, y).$$

其中区域  $D$  表示的是由曲线  $y = \frac{s}{x}$ ,  $x = 2$ ,  $y = 0$  和  $y = 1$  围城的区域. 由于二维随机变量  $(X, Y)$  在  $G$  上服从均匀分布, 可以写出  $X$  和  $Y$  的联合概率密度为:

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

带入积分:

$$\begin{aligned} F_S(s) &= \int_0^s \int_0^1 \frac{1}{2} dy dx + \int_s^2 \int_0^{\frac{s}{x}} \frac{1}{2} dy dx \\ &= \frac{s}{2} \left(1 + \ln \frac{2}{s}\right) \end{aligned}$$

由此得到概率分布函数:

$$F_S(s) = \begin{cases} 0 & s \leq 0 \\ \frac{s}{2} \left(1 + \ln \frac{2}{s}\right) & 0 < s \leq 2 \\ 1 & s > 2 \end{cases}$$

对概率分布函数求导, 得到概率密度函数:

$$f_S(s) = \begin{cases} \frac{1}{2} \ln \frac{s}{2} & 0 < s \leq 2 \\ 0 & \text{others.} \end{cases}$$

□

**例 3.5.6.** 设随机变量  $X$  与  $Y$  相互独立, 分别服从参数为  $\lambda_1$  与  $\lambda_2$  的指数分布, 求  $Z = \frac{X}{Y}$  的密度函数.

**解.** 先求  $Z$  的概率分布函数:

$$F_Z(z) = P\left\{\frac{X}{Y} \leq z\right\} = \iint_G f(x, y) dx dy$$

因为随机变量  $X$  与  $Y$  相互独立, 可以得出联合概率密度:

$$f(x, y) = f_X(x) \cdot f_Y(y) = \lambda_1 e^{-\lambda_1 x} \cdot \lambda_2 e^{-\lambda_2 y} = \lambda_1 \lambda_2 e^{-(\lambda_1 x + \lambda_2 y)}$$

根据定义, 计算概率分布函数:

$$\begin{aligned} F_Z(z) &= \int_0^{+\infty} \int_0^{zy} \lambda_1 \lambda_2 e^{-(\lambda_1 x + \lambda_2 y)} dx dy \\ &= 1 - \frac{\lambda_2}{\lambda_2 + \lambda_1 z} \quad z > 0 \end{aligned}$$

求导得到概率密度函数:

$$f_Z(z) = \begin{cases} \frac{\lambda_1 \lambda_2}{(\lambda_2 + \lambda_1 z)^2} & z > 0 \\ 0 & \text{others.} \end{cases}$$

□

### 3.5.3 取最值

形如  $Z = \max\{X, Y\}$  的随机变量. 这里要求随机变量  $X$  与  $Y$  相互独立. 可以得出  $Z$  的分布函数:

$$F_Z(z) = P\{Z \leq z\} = P\{\max\{X, Y\} \leq z\} = P\{X \leq z, Y \leq z\} = F_X(z) \cdot F_Y(z).$$

更一般地有:  $Z = \max\{X_1, X_2, \dots, X_n\}$  的概率分布函数为:

$$F_Z(z) = \prod_{i=1}^n F_{X_i}(z)$$

其中  $X_i$  之间相互独立, 若这些  $X_i$  全部服从相同的分布函数, 有  $F_Z(z) = (F_X(z))^n$ .

若  $Z = \min\{X, Y\}$ , 则有:

$$\begin{aligned} F_Z(z) &= P\{Z \leq z\} = P\{\min\{X, Y\} \leq z\} = 1 - P\{\min\{X, Y\} > z\} \\ &= 1 - P\{X > z, Y > z\} \\ &= 1 - P\{X > z\} \cdot P\{Y > z\} \\ &= 1 - (1 - F_X(z))(1 - F_Y(z)) \end{aligned}$$

其中  $X$  与  $Y$  相互独立.

同样的对于更一般的情况, 对于  $Z = \min \{X_1, X_2, \dots, X_n\}$ , 其中  $X_i$  相互独立. 则  $Z$  的分布函数为:

$$F_Z(z) = 1 - \prod_{i=1}^n (1 - F_{X_i}(z))$$

若这些  $X_i$  独立同分布, 还可以推出  $F_Z(z) = 1 - (1 - F_X(z))^n$ .

# 第四章 随机变量的数字特征

## 4.1 数学期望

### 4.1.1 离散型随机变量的数学期望

若随机变量  $X$  具有取值  $x_i$  和对应的概率  $p_i$  ( $i = 1, 2, 3, \dots$ ). 则它的数学期望为:

$$E(X) = \sum_{i=1}^{+\infty} x_i p_i$$

这里要求求和  $\sum x_i p_i$  必须绝对收敛, 否则随机变量  $X$  无期望.

**例 4.1.1.** 按规定, 某车站每天 8:00 ~ 9:00, 9:00 ~ 10:00 都恰有一辆客车到站, 但到站的时刻是随机的, 且两者到站的时间相互独立. 其规律为:

到站时刻	8:00	8:30	8:50
	9:10	9:30	9:50
概率	$\frac{1}{6}$	$\frac{3}{6}$	$\frac{2}{6}$

一旅客 8:20 到车站, 求他候车时间的数学期望.

**解.** 设随机变量  $X$  为旅客候车的时间.  $X$  可能的取值有  $X = 10, 30, 50, 70, 90$ . 对应的概率为:

$$\begin{aligned} P\{X = 10\} &= \frac{3}{6} & P\{X = 30\} &= \frac{2}{6} \\ P\{X = 50\} &= \frac{1}{36} & P\{X = 70\} &= \frac{3}{36} \\ P\{X = 90\} &= \frac{2}{36} \end{aligned}$$

因此根据公式算的  $E(X) = 27.22 \text{ min}$ . □

## 4.1.2 连续型随机变量的数学期望

随机变量  $X$  是连续型随机变量，可以得出：

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

只要上述积分满足绝对收敛，则称其为数学变量  $X$  的数学期望。

例 4.1.2. 设随机变量的分布函数为：

$$F(x) = \begin{cases} 1 - \frac{4}{x^2} & x \geq 2 \\ 0 & x < 2 \end{cases}$$

求  $X$  的数学期望。

解. 先求随机变量  $X$  的概率密度：

$$f(x) = \frac{8}{x^3} \quad x \geq 2$$

接着根据定义：

$$E(X) = \int_2^{+\infty} x \frac{8}{x^3} dx = 4$$

□

例 4.1.3. 两个相互独立工作的电子装置，它们的寿命用  $X_k$  来表示，服从同一个指数分布，其概率密度为

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad \theta > 0.$$

若将这两个电子装置串联组成整机，求整机寿命  $N$  的数学期望。

解. 根据题干，可以推断出整机的寿命应该小于  $X_k$  中最小的哪个，由此得出  $N = \min\{X_1, X_2\}$ . 接下来计算随机变量  $N$  的概率密度函数。

由于  $X_k$  独立同分布，可以得出：

$$F_N(n) = 1 - (1 - F_X(n))^2$$

根据题式概率密度, 可以求得:

$$F_X(n) = \int_0^n \frac{1}{\theta} e^{-\frac{x}{\theta}} dx = 1 - e^{-\frac{n}{\theta}}$$

由此得出:

$$F_N(n) = 1 - e^{-\frac{2n}{\theta}}$$

对  $F_N(n)$  求导得到  $N$  的概率密度:

$$f_N(n) = \frac{2}{\theta} e^{-\frac{2n}{\theta}}$$

可见随机变量  $N$  服从参数为  $\frac{2}{\theta}$  的指数分布. 根据期望的定义, 可以得出:

$$E(N) = \int_0^{+\infty} \frac{2n}{\theta} e^{-\frac{2n}{\theta}} dn = \frac{\theta}{2}$$

□

### 4.1.3 随机变量的函数的数学期望

设随机变量  $Y$  是随机变量  $X$  的函数:  $Y = g(X)$  且  $g$  连续.

1. 若随机变量  $X$  为离散型随机变量. 则  $E(Y) = \sum_{k=1}^{+\infty} g(x_k) p_k$ ;
2. 若随机变量  $X$  为连续型随机变量. 则  $E(Y) = \int_{-\infty}^{+\infty} g(x) f(x) dx$ , 其中  $f(x)$  是  $X$  的概率密度.

若  $Z$  是  $(X, Y)$  的函数,  $Z = g(X, Y)$  且  $g$  连续.

1.  $(X, Y)$  为二维离散型随机变量:  $P\{X = x_i, Y = y_j\} = p_{ij}$ . 则

$$E(Z) = \sum_{j=1}^{+\infty} \sum_{i=1}^{+\infty} g(x_i, y_j) \cdot p_{ij}$$

2. 若  $X, Y$  为二维连续型随机变量:  $f(x, y)$ . 有

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy$$

例 4.1.4. 设风速  $V$  在  $(0, a)$  上服从均匀分布, 即具有概率密度;

$$f(v) = \begin{cases} \frac{1}{a} & 0 < v < a \\ 0 & \text{others.} \end{cases}$$

又设飞机机翼受到的正压力  $W$  是  $V$  的函数:  $W = kV^2 (k > 0)$ . 求  $W$  的数学期望.

解. 带入公式:

$$E(W) = \int_0^a kv^2 \frac{1}{a} dv = \frac{ka^2}{3}$$

□

注. 上面的例子也可以利用分布函数求得, 先求  $W$  的概率分布函数, 然后求得  $W$  的概率密度函数, 最后带入公式:

$$E(W) = \int_{-\infty}^{+\infty} wf(w) dw$$

结果是一致的.

例 4.1.5. 设随机变量  $(X, Y)$  的概率密度:

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

求数学期望  $E(Y)$ ,  $E(\frac{1}{XY})$ .

解. 首先计算  $E(Y)$ . 有联合概率密度求  $Y$  的边缘概率密度:

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_y^{+\infty} \frac{3}{2x^3y^2} dx = \frac{3}{4y^4} \quad 1 < y$$

$$f_Y(y) = \int_{\frac{1}{y}}^{+\infty} \frac{3}{2x^3y^2} dx = \frac{3}{4} \quad 0 < y < 1$$

根据定义:

$$E(Y) = \int_0^1 y \cdot \frac{3}{4} dy + \int_1^{+\infty} y \cdot \frac{3}{4y^4} dy = \frac{3}{8} + \frac{3}{8} = \frac{3}{4}$$

上述积分不收敛, 因此根据公式算的

接下来计算  $E(\frac{1}{XY})$ , 令  $Z = \frac{1}{XY}$ . 带入公式:

$$E(Z) = \int_1^{+\infty} \int_{\frac{1}{x}}^x \frac{1}{xy} \cdot \frac{3}{2x^3y^2} dy dx = \frac{3}{5}$$

□

注. 对于  $E(Y)$  的计算, 还可以将  $Y$  视为  $(X, Y)$  的函数, 然后直接带入公式:

$$E(Y) = \int_1^{+\infty} \int_{\frac{1}{x}}^x y \cdot \frac{3}{2x^3y^2} dy dx = \frac{3}{4}$$

从原理上来说, 两种方法是相通的, 用分布函数法可以推导出上面的公式法.

## 4.1.4 数学期望的性质

1. 设  $C$  为常数,  $E(C) = C$ ;
2.  $E(CX) = CE(X)$ ;
3.  $E(X + Y) = E(X) + E(Y)$ ;
4.  $X$  与  $Y$  相互独立, 则  $E(XY) = E(X)E(Y)$ .

**例 4.1.6.** 一民航送客车载有 20 位乘客自机场开出, 旅客有 10 个站可以下车 (每位乘客在每个车站下车是等可能的并且是否下车相互独立), 如果到达一个车站没有旅客下车就不停车. 以  $X$  表示停车的次数, 求  $E(X)$ .

**解.** 引入随机变量  $Y_i$  ( $i = 1, 2, 3, \dots, 10$ ),  $Y_i = 0$  表示在第  $i$  个车站没有人下车, 若等于 1 则表示有人下车. 随机变量  $X$  与  $Y$  之间的关系为  $X = \sum_{i=1}^{10} Y_i$ . 根据数学期望的性质:

$$E(X) = E\left(\sum_{i=1}^{10} Y_i\right) = \sum_{i=1}^{10} E(Y_i)$$

由于每位乘客在每个车站下车等可能且相互独立, 因此所有的  $Y_i$  是等价的, 记为  $Y$ :

$$P\{Y = k\} = \begin{cases} \left(\frac{9}{10}\right)^{20} & \text{所有人都不下车即 } k = 0 \\ 1 - \left(\frac{9}{10}\right)^{20} & \text{至少有一有人下车即 } k = 1 \end{cases}$$

因此可以计算出:

$$E(X) = 10E(Y) = 10 \times 10 \times \left[1 - \left(\frac{9}{10}\right)^{20}\right]$$

□

**例 4.1.7.** 设一电路中电流  $I$  与电阻  $R$  是两个相互独立的随机变量, 其概率密度分别为:

$$g(i) = \begin{cases} 2i & 0 \leq i \leq 1 \\ 0 & \text{others.} \end{cases} \quad h(r) = \begin{cases} \frac{r^2}{9} & 0 \leq r \leq 3 \\ 0 & \text{others.} \end{cases}$$

试求电压  $V = IR$  的平均值.

**解.** 电压的平均值即随机变量  $V$  的数学期望. 由于  $I, R$  相互独立, 根据数学期望的性质:

$$E(V) = E(IR) = E(I) \cdot E(R) = \int_0^1 i \cdot 2i \, di \int_0^3 r \cdot \frac{r^2}{9} \, dr = \frac{3}{2}$$

□



## 4.1.5 服从特定分布的数学期望

例 4.1.8. 设随机变量  $X \sim \pi(\lambda)$ , 求  $E(X)$ .

解. 根据泊松分布的定义:

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

由期望的定义:

$$E(X) = \sum_{k=1}^{+\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \lambda$$

□

解. 设随机变量  $X \sim U(a, b)$ , 求  $E(X)$ .

□

解. 根据均匀分布的定义:

$$f(x) = \frac{1}{b-a}$$

由数学期望的定义:

$$E(X) = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}$$

□

下表是对之前的所有常见分布的性质的总结:

分布类型	符号表示	分布率或概率密度	期望	方差
0-1 分布	$B(1, p)$	$P\{X = k\} = p^k (1-p)^{1-k}$	$p$	$p(1-p)$
二项分布	$B(n, p)$	$P\{X = k\} = \binom{n}{k} p^k (1-p)^{n-k}$	$np$	$np(1-p)$
泊松分布	$\pi(\lambda)$	$P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}$	$\lambda$	$\lambda$
几何分布	$G(p)$	$P\{X = k\} = p(1-p)^{k-1}$	$\frac{1}{p}$	
均匀分布	$U(a, b)$	$f(x) = \frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布	$E(\lambda)$	$f(x) = \lambda e^{-\lambda x}$	$\frac{1}{\lambda}$	$\theta^2$
正态分布	$N(\mu, \sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$	$\mu$	$\sigma^2$

## 4.2 方差

方差是用来衡量随机变量与自身均值偏移程度的数字特征。

$$E\{|X - E(X)|\} \implies E\{(X - E(X))^2\} = D(X)$$

**定义 4.2.1.** (方差) 设随机变量  $X$ , 若  $(X - E(X))^2$  的期望存在, 称它的期望为  $X$  的方差, 记作:

$$D(X) = E\{(X - E(X))^2\}$$

注. 从上面的定义来看, 方差的本质其实就是随机变量函数  $g(X) = (X - E(X))^2$  的期望. 而这个函数刻画的就是随机变量与自身均值的偏移水平.

**定义 4.2.2.** (标准差) 随机变量的标准差为  $\sigma(X)$ .

$$\sigma(X) = \sqrt{D(X)}$$

根据上节期望的计算, 可以总结出方差的计算方法:

1. 若随机变量  $X$  为离散型随机变量:

$$D(X) = \sum_{k=1}^{\infty} g(x_k) \cdot p_k = \sum_{k=1}^{\infty} (x_k - E(X))^2 \cdot p_k$$

2. 若随机变量  $X$  为连续型随机变量:

$$D(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 \cdot f(x) dx$$

**定理 4.2.1.** 对任意随机变量  $X$  有:

$$D(X) = E(X^2) - E^2(X)$$

**证明.** 根据方差的定义及期望的性质:

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

□

例 4.2.1. 设随机变量  $X$  具有数学期望  $E(X) = \mu$ , 方差  $D(X) = \sigma^2$ . 记:

$$X^* = \frac{X - \mu}{\sigma}$$

求期望和方差.

解. 首先求期望, 根据期望的性质:

$$E(X^*) = \frac{1}{\sigma} (E(X) - E(\mu)) = \frac{1}{\sigma} E(X) - \frac{1}{\sigma} \mu = 0$$

接下来求方差:

$$\begin{aligned} D(X^*) &= E(X^{*2}) - E^2(X^*) = \frac{1}{\sigma^2} E(X^2 - 2\mu X + \mu^2) = \frac{1}{\sigma^2} (E(X^2) - 2\mu E(X) + \mu^2) \\ &= \frac{1}{\sigma^2} (\sigma^2 + \mu^2 - 2\mu^2 + \mu^2) = 1 \end{aligned}$$

□

注. 上例可以作为一个性质, 形如:

$$X^* = \frac{X - \mu}{\sigma}$$

的变换成为标准化变换. 在正态分布中常用.

例 4.2.2. 设随机变量  $X$  具有  $(0-1)$  分布, 其分布律为:

$$P\{X=0\} = 1-p, \quad P\{X=1\} = p.$$

求数学期望.

解. 首先考虑求解随机变量  $X$  的数学期望:

$$E(X) = p$$

接下来求解方差;

$$D(X) = E(X^2) - E^2(X) = E(X^2) - p^2$$

对于期望  $E(X^2)$ , 可以采用随机变量函数的期望的求解方法:

$$E(X^2) = \sum_{k=0}^1 x_k^2 \cdot p_k = p$$

综上所述:

$$D(X) = p - p^2 = p(1-p).$$

□

例 4.2.3. 设随机变量  $X \sim \pi(\lambda)$ , 求  $D(X)$ .

解. 泊松分布下, 随机变量  $X$  的数学期望为  $\lambda$ . 接下来求  $E(X^2)$ .

$$E(X^2) = \sum_{k=0}^{+\infty} k^2 \cdot \frac{\lambda^k}{k!} e^{-\lambda}$$

这里发现求解比较困难, 于是采用下面这种求解方法.

首先对  $E(X^2)$  进行变换:

$$E(X^2) = E(X(X-1) + X)$$

然后根据期望的性质:

$$\begin{aligned} E(X^2) &= E(X(X-1)) + E(X) \\ &= \sum_{k=0}^{+\infty} k(k-1) \frac{\lambda^k}{k!} e^{-\lambda} + \lambda \\ &= e^{-\lambda} \sum_{k=2}^{+\infty} \frac{\lambda^k}{(k-2)!} + \lambda \\ &= e^{-\lambda} \lambda^2 \sum_{k=2}^{+\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda \\ &= e^{-\lambda} \lambda^2 \sum_{t=0}^{+\infty} \frac{\lambda^t}{t!} + \lambda \\ &= \lambda^2 + \lambda \end{aligned}$$

因此:

$$D(X) = E(X^2) - E^2(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$$

□

例 4.2.4. 设随机变量  $X \sim U(a, b)$ , 求  $D(X)$ .

解. 对于均匀分布的随机变量  $X$ , 其期望为:

$$E(X) = \frac{a+b}{2}$$

下面求  $E(X^2)$ . 先求  $Y = X^2$  的概率密度:

$$f_Y(y) = \frac{1}{\sqrt{y}(b-a)}$$

接下来求  $Y$  的期望:

$$E(Y) = \int_{a^2}^{b^2} \frac{1}{\sqrt{y}(b-a)} = \frac{2(b^2 + ab + a^2)}{3}$$

因此, 随机变量  $X$  的方差为:

$$D(X) = \frac{2(b^2 + ab + a^2)}{3} - \frac{a^2 + 2ab + b^2}{4} = \frac{(a-b)^2}{12}$$

□

**例 4.2.5.** 设随即变量  $X$  指数分布, 其概率密度为:

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad \theta > 0$$

求  $E(X)$  和  $D(X)$ .

**解.** 根据期望的定义:

$$E(X) = \int_0^{+\infty} \frac{x}{\theta} e^{-\frac{x}{\theta}} dx = \theta$$

接下来求  $E(X^2)$ .

设  $Y = X^2 \in (0, +\infty)$ , 有:

$$f_Y(y) =$$

最后求得:

$$D(X) = 2\theta^2 - \theta^2 = \theta^2$$

□

### 4.2.1 方差的性质

设  $C$  为常数, 随机变量  $X, Y$ . 有:

1.  $D(C) = 0$ ;
2.  $D(CX) = C^2 D(X)$ ;
3.  $D(C + X) = D(X)$ ;

4.  $D(X \pm Y) = D(X) + D(Y) + 2E\{(X - E(X))(Y - E(Y))\}$ . 特别地如果  $X$  和  $Y$  相互独立, 有

$$D(X \pm Y) = D(X) + D(Y)$$

5.  $D(X) = 0$  等价于随机变量  $X$  以概率为 1 取到常数  $E(X)$ .

**例 4.2.6.** 设随机变量  $X \sim B(n, p)$ , 求  $E(X)$ ,  $D(X)$ .

**解.** 设随机变量  $Y_i$ :

$$Y_i = \begin{cases} 1 & \text{第 } k \text{ 次试验成功} \\ 0 & \text{第 } k \text{ 次试验失败} \end{cases}$$

由题意得:

$$E(X) = E\left(\sum Y_i\right) = np$$

由于每次伯努利试验之间相互独立, 因此:

$$D(X) = \sum_{i=1}^n D(Y_i) = np(1-p)$$

□

**例 4.2.7.** 设随机变量  $X \sim N(\mu, \sigma^2)$ , 求  $E(X)$ ,  $D(X)$ .

**解.** 设随机变量  $Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$ . 且有:

$$f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

可以求得  $E(X) = \mu$ . 下面求  $D(Z)$ :

$$D(Z) = E(Z^2) - E^2(Z) = 1$$

有因为  $X = \sigma Z + \mu$ , 根据方差的性质:

$$D(X) = D(\sigma Z + \mu) = \sigma^2 D(Z) = \sigma^2$$

□

**例 4.2.8.** 设活塞的直径  $X \sim N(22.40, 0.03^2)$ , 气缸直径  $Y \sim N(22.50, 0.04^2)$ ,  $X$  与  $Y$  相互独立. 任取一只活塞, 一只气缸, 求活塞能装入气缸的概率.

解. 要求活塞能够装入气缸的概率, 即求:

$$P\{X < Y\}$$

设随机变量  $Z = Y - X$ , 由于  $X, Y$  相互独立, 根据期望和方差的性质, 可以推出:  $Z \sim N(\mu_Y - \mu_X, \sigma_Y^2 + \sigma_X^2)$ .

$$P\{Z > 0\} = 1 - P(Z \leq 0) = 1 - P\left\{\frac{Z - \mu_Z}{\sigma_Z} \leq -\frac{\mu_Z}{\sigma_Z}\right\} = 1 - \Phi(-2) = 0.9772$$

□

注. 有上例可以总结出正态分布的性质: 若  $X, Y$  相互独立. 且  $X \sim N(\mu_X, \sigma_X^2)$ ,  $Y \sim N(\mu_Y, \sigma_Y^2)$ . 设随机变量  $Z = CX + DY$ , 其中  $C$  与  $D$  是常数. 则根据方差和期望的性质, 有:

$$\mu_Z = E(Z) = E(CX + DY) = CE(X) + DE(Y) = C\mu_X + D\mu_Y$$

$$\sigma_Z^2 = D(Z) = D(CX + DY) = D(CX) + D(DY) = \sigma_X^2 + \sigma_Y^2$$

因此  $Z \sim N(C\mu_X + D\mu_Y, \sigma_X^2 + \sigma_Y^2)$ .

## 4.2.2 切比雪夫不等式

**定理 4.2.2.** (切比雪夫不等式) 设随机变量  $X$  满足:  $E(X) = \mu$ ,  $D(X) = \sigma^2$ . 对于任何一个大于 0 的常数  $\varepsilon$ , 都有以下不等式成立:

$$P\{|X - \mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}$$

称之为“切比雪夫不等式”.

**证明.**

$$\begin{aligned} P\{|X - \mu| \geq \varepsilon\} &= P\{(X - \mu \geq \varepsilon) \cup (X - \mu \leq -\varepsilon)\} \\ &= P\{X - \mu \geq \varepsilon\} + P\{X - \mu \leq -\varepsilon\} \\ &= \int_{\mu+\varepsilon}^{+\infty} f(x) dx + \int_{-\infty}^{\mu-\varepsilon} f(x) dx \\ &\leq \frac{|x - \mu|^2}{\varepsilon^2} \left( \int_{\mu+\varepsilon}^{+\infty} f(x) dx + \int_{-\infty}^{\mu-\varepsilon} f(x) dx \right) \\ &= \frac{1}{\varepsilon^2} \left( \int_{\mu+\varepsilon}^{+\infty} (x - \mu)^2 \cdot f(x) dx + \int_{-\infty}^{\mu-\varepsilon} (x - \mu)^2 \cdot f(x) dx \right) \\ &\leq \frac{1}{\varepsilon} \int_{-\infty}^{+\infty} (x - \mu)^2 \cdot f(x) dx \\ &= \frac{\sigma^2}{\varepsilon^2} \end{aligned}$$

□

**注.** 对于切比雪夫不等式, 还有一种表达:

$$P\{|X - \mu| < \varepsilon\} \geq 1 - \frac{\sigma^2}{\varepsilon^2}$$

这个很好理解. 切比雪夫不等式是一个用来估计概率的一个工具.

**例 4.2.9.** 对于随机变量  $X$ , 有:

$$D(X) = 0 \implies P\{X = E(X)\} = 1$$

**证明.** 根据切比雪夫不等式, 有:

$$P\{|X - E(X)| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2} = \frac{D(X)}{\varepsilon^2} = 0$$



又因为对于  $\forall X$ , 有  $P\{X\} \geq 0$ , 因此

$$P\{|X - E(X)| \geq \varepsilon\} = 0 \implies P\{|X - E(X)| < \varepsilon\} = 1$$

其中  $\varepsilon$  为任意大于 0 的正数. 因此:

$$\lim_{\varepsilon \rightarrow 0} P\{|X - E(X)| < \varepsilon\} = 1 \implies P\{X = E(X)\} = 1$$

□

注. 上例是方差的第 4 条性质的必要性的证明. 要注意, 若随机变量  $X$  取到  $E(X)$  是必然事件, 则有  $D(X) = 0$ , 反之不一定成立. 这里体现出了“以概率为 1 取到期望”和“必然事件”之间的差异.

### 4.3 协方差与相关系数

设随机变量  $X, Y$ , 有:

$$D(X + Y) = D(X) + D(Y) + 2E\{(X - E(X))(Y - E(Y))\}$$

当  $X$  与  $Y$  相互独立, 则:

$$E\{(X - E(X))(Y - E(Y))\} = 0$$

反之, 当:

$$E\{(X - E(X))(Y - E(Y))\} \neq 0$$

说明  $X$  与  $Y$  不相互独立.

当随机变量  $X$  与  $Y$  不相互独立, 说明两者之间有关系, 可能关系如下:

1. 线性关系 (线性相关/相关):  $Y = aX + b$ ;
2. 非线性关系:  $Y = \sum_{i=0}^n a_i \cdot X^i$ .

**定义 4.3.1.** (协方差与相关系数) 设随机变量  $X$  与  $Y$ , 定义:

$$E\{(X - E(X))(Y - E(Y))\}$$

为  $X$  与  $Y$  的协方差. 记作  $Cov(X, Y)$ .

定义 4.3.2. 设随机变量  $X, Y$ . 定义:

$$\rho_{XY} = \frac{Cov(X, Y)}{\sqrt{D(X)} \cdot \sqrt{D(Y)}}$$

为  $X$  与  $Y$  的 (线性) 相关系数. 反应  $X$  与  $Y$  的 (线性) 相关程度.

协方差有如下性质:

1.  $Cov(X, Y) = Cov(Y, X)$ ;
2.  $Cov(X, X) = D(X)$ ;
3.  $D(X, Y) = D(X) + D(Y) + 2Cov(X, Y)$ ;
- 4.

$$\begin{aligned} Cov(X, Y) &= E\{(X - E(X))(Y - E(Y))\} \\ &= E\{XY - XE(Y) - YE(E(X)) + E(X)E(Y)\} \\ &= E(XY) - E(Y)E(X) - E(X)E(Y) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

5.  $Cov(aX, bY) = abCov(X, Y)$ ;
6.  $Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y)$ .

相关系数有如下性质:

1.  $|\rho_{XY}| \leq 1$ ;
2.  $|\rho_{XY}| = 1 \implies X$  与  $Y$  是线性关系;
3.  $|\rho_{XY}|$  越靠近 1,  $X$  与  $Y$  之间的 (线性) 相关性越大;
4.  $\rho_{XY} = 0$ , 随机变量  $X$  与  $Y$  不 (线性) 相关.

注. 若随机变量  $X$  与  $Y$  相互独立  $\iff$  无任何关系  $\implies$  无线性关系  $\iff$  不相关. 特别地, 对于服从二维正态分布的随机变量  $(X, Y)$ , 如果  $\rho = 0$ , 则  $X$  与  $Y$  独立.

## 4.4 矩以及协方差矩阵

### 4.4.1 矩

**定义 4.4.1.** 设随机变量  $X$ . 若  $E(|X|^k) < +\infty$ , 则  $E((X-c)^k)$  称为  $X$  关于  $c$  的  $k$  阶矩. 称  $\alpha_k = E(X^k)$  为随机变量  $X$  的  $k$  阶原点矩, 称  $E((X-E(X))^k)$  为  $X$  的  $k$  阶中心矩.

**定义 4.4.2.** 设随机变量  $X, Y$ . 若  $E(X^k Y^l)$  存在, 则称之为  $X$  和  $Y$  的  $k+l$  阶混合原点矩. 若  $E\{(X-E(X))^k (Y-E(Y))^l\}$  存在, 称其为  $X$  和  $Y$  的  $k+l$  阶混合中心矩.

**注.** 由上面的定义可知: 随机变量  $X$  的期望  $E(X)$  为其一阶原点矩, 方差  $D(X) = E\{(X-E(X))^2\}$  为其二阶中心矩, 随机变量  $X, Y$  的协方差

$$\text{Cov}\{X, Y\} = E\{(X-E(X))(Y-E(Y))\}$$

为  $X, Y$  的二阶混合中心矩.

### 4.4.2 协方差矩阵

设随机变量  $(X_1, X_2)$ , 如果它们所有的矩都存在, 则存在 4 个中心矩:

$$\begin{aligned} C_{11} &= E\{(X_1 - E(X_1))^2\} & C_{12} &= E\{(X_1 - E(X_1))(X_2 - E(X_2))\} \\ C_{21} &= E\{(X_2 - E(X_2))(X_1 - E(X_1))\} & C_{22} &= E\{(X_2 - E(X_2))^2\} \end{aligned}$$

将上述中心矩写成矩阵:

$$\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} = \begin{pmatrix} D(X_1) & \text{Cov}(X_1, X_2) \\ \text{Cov}(X_2, X_1) & D(X_2) \end{pmatrix}$$

这便是协方差矩阵.

### 4.4.3 多维正态随机变量的性质

设  $(X_1, X_2, X_3, \dots, X_n)$  服从  $n$  维正态分布:

1.  $X_i$  服从正态分布; 若  $X_i$  之间相互独立且均服从正态分布, 则  $X_1, X_2, \dots, X_n$  服从  $n$  维正态分布;

2.  $X_i$  的任意线性组合服从一维正态分布;
3. 设  $Y_j$  是  $X_i$  的线性组合, 则  $(Y_1, Y_2, \dots, Y_n)$  服从  $n$  维正态分布.
4. 如果  $X_i$  之间都相互独立  $\Leftrightarrow X_i$  与  $X_j$  之间两两不相关.

**例 4.4.1.** 设随机变量  $X$  在  $[a, b]$  上服从均匀分布, 求  $X$  的  $k$  阶原点矩和三阶中心矩.

**解.** 首先求  $k$  阶原点矩:

$$E(|X|^k) = \int_a^b x^k \cdot \frac{1}{b-a} dx = \frac{1}{k+1} \cdot \frac{b^{k+1} - a^{k+1}}{b-a}$$

再求三阶中心矩:

$$E((X - E(X))^3) = \int_a^b \left(x - \frac{a+b}{2}\right)^3 \cdot \frac{1}{b-a} dx = 0$$

□

**例 4.4.2.** 设二维随机变量  $(X, Y)$  的协方差矩阵为:

$$C = \begin{pmatrix} 1 & -1 \\ -1 & 9 \end{pmatrix}$$

求  $\rho_{XY}$ .

**解.** 根据协方差矩阵可以得到:  $D(X) = 1, D(Y) = 9, Cov(X, Y) = -1$ . 根据相关系数的定义:

$$\rho_{XY} = \frac{Cov(X, Y)}{\sqrt{D(X)} \cdot \sqrt{D(Y)}} = -\frac{1}{3}$$

□

## 4.5 大数定理与中心极限定理

### 4.5.1 两个不等式

马尔可夫不等式

**定理 4.5.1.** (马尔可夫不等式) 若  $Y$  是只取非负数的随机变量.  $E(Y)$  存在, 那么对于  $\forall \varepsilon > 0$ , 则有:

$$P\{Y \geq \varepsilon\} \leq \frac{E(Y)}{\varepsilon}$$

证明. 这里只证明连续型随机变量的情况. 根据期望的定义:

$$\begin{aligned}
 E(Y) &= \int_0^{+\infty} y \cdot f(y) dy \\
 &\geq \int_{\varepsilon}^{+\infty} y \cdot f(y) dy \\
 &\geq \int_{\varepsilon}^{+\infty} \varepsilon \cdot f(y) dy \\
 &= \varepsilon \int_{\varepsilon}^{+\infty} f(y) dy \\
 &= \varepsilon \cdot P\{Y \geq \varepsilon\} \\
 \Rightarrow P\{Y \geq \varepsilon\} &\leq \frac{E(Y)}{\varepsilon}
 \end{aligned}$$

□

### 切比雪夫不等式

**定理 4.5.2.** (切比雪夫不等式) 设随机变量  $X$  满足:  $E(X) = \mu$ ,  $D(X) = \sigma^2$ . 对于任何一个大于 0 的常数  $\varepsilon$ , 都有以下不等式成立:

$$P\{|X - \mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}$$

称之为“切比雪夫不等式”.

**证明.** (利用马尔可夫不等式) 令  $|X - \mu| = Y \geq 0$ , 显然  $Y$  符合马尔可夫不等式要求. 有:

$$E(Y^2) = E\{(X - E(X))^2\} = D(X)$$

带入马尔可夫不等式:

$$P\{|X - \mu| \geq \varepsilon\} = P\{Y \geq \varepsilon\} = P\{Y^2 \geq \varepsilon^2\} \leq \frac{E(Y^2)}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2}$$

□

注. 上述两个不等式的作用都是对概率进行粗略的估计.

**例 4.5.1.** 设随机变量  $X$ .  $E(X) = \mu$ ,  $D(X) = \sigma^2 > 0$ . 估计  $P\{|X - \mu| \geq 6\sigma\}$ .

解. 利用切比雪夫不等式:

$$P\{|X - \mu| \geq 6\sigma\} \leq \frac{\sigma^2}{(6\sigma)^2} = \frac{1}{36}$$

□

例 4.5.2. 设随机变量  $X \sim N(\mu, \sigma^2)$ . 求  $P\{|X - \mu| \geq 6\sigma\}$ .

解.

$$\begin{aligned} P\{|X - \mu| \geq 6\sigma\} &= 1 - P\{-6\sigma < X - \mu < 6\sigma\} \\ &= 1 - P\left\{-6 < \frac{X - \mu}{\sigma} < 6\right\} \\ &= 1 - (P\{Z < 6\} - P\{Z \leq -6\}) \\ &= 1 - (2\Phi(6) - 1) \\ &= 2 - 2\Phi(6) = 0.0000068 \end{aligned}$$

□

注. 从上面的两个例子可以看出, 切比雪夫不等式 (马尔可夫不等式) 对于概率的估计是很粗略的.

## 4.5.2 两个收敛

### 数列收敛和函数收敛

定义 4.5.1. (数列收敛) 设数列  $\{X_n\}$ , 存在  $a$ . 若对于  $\forall \varepsilon$ , 总存在  $N \in \mathbb{Z}$ , 使得当  $n > N$  时, 有:

$$|X_n - a| < \varepsilon \quad (4.1)$$

称数列  $\{X_n\}$  收敛于  $a$ . 这是一个必然事件. 记作:  $\lim_{n \rightarrow +\infty} X_n = a$ .

定义 4.5.2. (函数收敛)

$$|f(x) - A| < \varepsilon \implies \lim_{x \rightarrow +\infty} f(x) = A$$

### 依概率收敛

定义 4.5.3. 随机变量  $X$  与随机变量得序列  $\{X_n\}$ . 若对于  $\forall \varepsilon > 0$  或常数  $a$ . 有:

$$\lim_{n \rightarrow \infty} P\{|X_n - X| \geq \varepsilon\} = 0 \quad \text{or} \quad P\{|X_n - a| < \varepsilon\} = 1$$

成立. 则称  $\{X_n\}$  依概率收敛于  $X$ . 记作:  $X_n \xrightarrow{P} X$  或  $\lim_{n \rightarrow +\infty} X_n = a(\text{in } P)$ .

注. 考虑抛硬币的例子, 记  $X$  为抛硬币正面向上的频率.  $X_n$  为第  $n$  次试验时正面向上的频率. 当  $n$  无限大时, 常识指出频率趋近于 0.5. 但是在所有的实验中, 总有一段时间的试验不是服从逐渐趋近 0.5 的, 但是总体来看服从尝试规律. 这样的收敛我们称为依概率收敛. 在这个例子中, 可以说抛硬币的频率依概率收敛至 0.5.

例 4.5.3. 设随机变量序列  $\{X_n\}$ .  $X_n$  的概率密度为:

$$f_n(x) = \frac{n}{\pi(1+n^2x^2)} \quad -\infty < x < +\infty$$

试证明  $X_n \xrightarrow{P} 0$ .

证明. 根据依概率收敛的定义:

$$\begin{aligned} P\{|X_n - 0| < \varepsilon\} &= \int_{-\varepsilon}^{+\varepsilon} f_n(x) dx \\ &= \int_{-\varepsilon}^{+\varepsilon} \frac{n}{\pi(1+n^2x^2)} dx \\ &= \frac{2}{\pi} \arctan n\varepsilon \end{aligned}$$

求极限得:  $\lim_{n \rightarrow \infty} P = 1$ . □

### 4.5.3 三个大数定律

#### 伯努利大数定律

定理 4.5.3. (伯努利大数定理) 在  $n$  次重复独立试验 ( $n$  重伯努利试验) 中, 事件  $A$  发生次数为  $f_A$ ,  $p$  为事件  $A$  发生的概率. 如果对于  $\forall \varepsilon > 0$  有:

$$P\left\{\left|\frac{f_A}{n} - p\right| < \varepsilon\right\} = 1$$

则说明实验次数足够多时, 事件  $A$  发生的频率依概率收敛于概率.

证明. 利用切比雪夫不等式可以证明. □

注. 上述定理说明, 可以用大量试验的频率来估计概率.

## 切比雪夫大数定律

**定理 4.5.4.** (切比雪夫大数定律) 设随机变量序列  $\{X_n\}$  相互独立. 且  $D(X_n)$  存在一致的上界,  $E(X_n)$  存在. 如果对  $\forall \varepsilon > 0$  有:

$$\lim_{n \rightarrow +\infty} P \left\{ \left| \frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n E(X_i) \right| < \varepsilon \right\} = 1$$

则说明随机变量序列的均值依概率收敛于期望的均值.

**证明.** (利用切比雪夫不等式) 首先计算随机变量序列的均值的期望:

$$E \left( \frac{1}{n} \sum_{i=1}^n X_i \right) = \frac{1}{n} E \left( \sum_{i=1}^n X_i \right) = \frac{1}{n} \sum_{i=1}^n E(X_i)$$

满足切比雪夫不等式的要求, 带入得:

$$P \left\{ \left| \frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n E(X_i) \right| < \varepsilon \right\} \geq 1 - \frac{D(Z)}{\varepsilon^2}$$

其中  $Z = \frac{1}{n} \sum_{i=1}^n X_i$ . 因此  $D(Z) = \frac{1}{n^2} D(\sum_{i=1}^n X_i) \leq \frac{M}{n^2}$ . 推出:

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n E(X_i) \right| < \varepsilon \right\} = 1 - \frac{M}{n^2 \varepsilon} = 1$$

其中  $M$  为方差的上界. □

## 辛钦大数定律

**定理 4.5.5.** (辛钦大数定律) 随机变量序列  $\{X_n\}$  独立同分布, 且  $E(X_i) = \mu$ . 如果对  $\forall \varepsilon > 0$  有:

$$\lim_{n \rightarrow +\infty} P \left\{ \left| \frac{1}{n} \sum_{i=1}^n X_i - \mu \right| < \varepsilon \right\} = 1$$

说明均值依概率收敛于期望.

## 4.5.4 三个中心极限定理

## 林德伯格-莱维中心极限定理

**定理 4.5.6.** (林德伯格-莱维中心极限定理) 设随机变量序列  $\{X_n\}$  独立同分布,  $E(X_i) = \mu$ ,  $D(X_i) = \sigma^2 < +\infty$ . 如果对  $\forall x \in \mathbb{R}$ , 都有:

$$\lim_{n \rightarrow +\infty} P \left\{ \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n} \cdot \sigma} \leq x \right\} = \Phi(x)$$

说明可以通过上述变换将随机变量序列转化成标准正态分布.



## 李雅普诺夫中心极限定理

**定理 4.5.7.** 设随机变量序列  $\{X_n\}$  相互独立. 有  $E(X_n) = \mu_n$ ,  $D(X_n) = \sigma_n^2 < +\infty$ . 假设每个  $X_n$  对  $\sum_{n=1}^m X_n$  影响不大, 记  $S_m = (\sum_{n=1}^m \sigma_n^2)^{\frac{1}{2}}$ , 则:

$$\lim_{m \rightarrow +\infty} P \left\{ \frac{1}{S_m} \cdot \sum_{n=1}^m (X_n - \mu_n) \leq x \right\} = \Phi(x)$$

**例 4.5.4.** 一加法器同时收到 20 个噪声电压  $V_k$  ( $k = 1, 2, 3, \dots, 20$ ). 设它们是相互独立的随机变量, 且都在区间  $(0, 10)$  上服从均匀分布. 记  $V = \sum_{k=1}^{20} V_k$ . 求  $P\{V > 105\}$  的近似值.

**解.** 因为随机变量序列  $\{V_n\}$  独立同分布, 因此考虑林德伯格-莱维中心极限定理.

$$P\{V > 105\} = P \left\{ \frac{V - n\mu}{\sqrt{n}\sigma} > \frac{105 - 20\mu}{\sqrt{20}\sigma} \right\} \rightarrow 1 - \Phi \left( \frac{105 - 20\mu}{\sqrt{20}\sigma} \right)$$

其中  $\mu = 5$ ,  $\sigma^2 = \frac{100}{12} = 8.3$ . 带入得:

$$P\{V > 105\} \rightarrow 1 - \Phi(0.387) = 0.348$$

□

## 棣莫弗-拉普拉斯中心极限定理

**定理 4.5.8.** 设随机变量  $Y_n \sim B(n, p)$ , 对于任意实数  $x$  有:

$$\lim_{n \rightarrow \infty} P \left\{ \frac{Y_n - np}{\sqrt{np(1-p)}} \leq x \right\} = \Phi(x)$$

还可以定义随机变量序列  $\{X_n\}$ . 其中每个  $X_i \sim B(1, p)$ , 则  $Y_n = \sum_{i=1}^n X_i$ .

**注.** 上述定理说明当试验次数足够大时, 可以用正态分布来近似二项分布.

**例 4.5.5.** 一船舶在某海区航行, 已知每遭受一次波浪的冲击, 纵摇大于  $3^\circ$  的概率为  $p = \frac{1}{3}$ . 若船舶遭受了 90000 次破浪冲击, 问其中有 29500 ~ 30500 次纵摇角度大于  $3^\circ$  的概率.

**解.** 设随机变量  $X \sim B(90000, \frac{1}{3})$  表示纵摇次数大于  $3^\circ$  的次数. 考虑棣莫弗-拉普拉斯中心极限定理.

$$\begin{aligned} P\{29500 \leq X \leq 30500\} &= P \left\{ \frac{29500 - 30000}{141.4} \leq \frac{X - 30000}{\sqrt{30000(1 - \frac{1}{3})}} \leq \frac{30500 - 30000}{141.4} \right\} \\ &= 2\Phi(3.53) - 1 = 0.9995 \end{aligned}$$

□

**例 4.5.6.** 对于一个学生而言, 来参加家长会的家长人数是一个随机变量, 设一个学生无家长、1 名家长、2 名家长来参会的概率分别为 0.05、0.8、0.15. 若学校共有 400 名学生, 设各学生参加会议的家长人数相互独立, 且服从同一分布.

- (1) 求参加会议的家长人数  $X$  超过 450 的概率;
- (2) 求有 1 名家长来参加会议的学生人数不多于 340 的概率.

**解.** 首先求  $P\{X > 450\}$ . 对于每个  $X_i$ , 可以求出其共同的期望和方差:  $\mu = 1.1, \sigma^2 = 0.19$ . 而随机变量  $X = \sum_{i=1}^n X_i$ . 根据林德伯格-莱维中心极限定理:

$$P\{X > 450\} = P\left\{\frac{X - n\mu}{\sqrt{n}\sigma} > \frac{450 - 440}{20 \times \sqrt{0.19}}\right\} \rightarrow 1 - \Phi(1.147) = 0.1251$$

设随机变量  $Y$  表示有 1 名家长来参会的学生的人数.  $Y \sim B(400, 0.8)$ , 考虑棣莫弗-拉普拉斯中心极限定理.

$$P\{Y \leq 340\} = P\left\{\frac{Y - np}{\sqrt{np(1-p)}} \leq \frac{340 - 320}{\sqrt{320 \times 0.2}}\right\} \rightarrow \Phi(2.5) = 0.9938$$

□

# 第五章 统计学基本概念

## 5.1 随机样本

定义 5.1.1. 试验中全部可能被观测到的值被称为总体. 每一个可观测到的值为个体.

定义 5.1.2. 从总体中抽取部分个体来判断总体.

定义 5.1.3. 对总体  $X$  进行一次随机观察记录结果, 即从总体中抽取了一个个体. 然后在相同条件下, 进行  $n$  次独立重复实验, 产生了  $X_n = X_1, X_2, \dots, X_n$  等多个随机变量. 随机变量序列  $\{X_n\}$  叫做样本, 与  $X$  服从相同分布. 当  $X_i$  的具体取值  $x_i$  叫做样本值.

## 5.2 抽样分布

### 5.2.1 统计量

定义 5.2.1. 设  $X_1, X_2, \dots, X_n$  是来自总体的一个样本.  $g(X_1, X_2, \dots, X_n)$  是  $X_i$  的函数. 若  $g$  中不包含未知参数. 则称  $g(x_1, x_2, \dots, x_n)$  为一个统计量.  $g(x_1, x_2, \dots, x_n)$  为统计量的观测值.

#### 典型统计量

定义 5.2.2. (样本均值)

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

定义 5.2.3. (样本方差、标准差)

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \qquad S = \sqrt{S^2}$$

定义 5.2.4. 样本的  $k$  阶原点矩、中心矩.

$$a_k = \frac{1}{n} X_i^k \qquad m_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k$$

## 5.2.2 统计量的分布

### 卡方分布

定义 5.2.5. 样本  $X_1, X_2, \dots, X_n$  来自标准正态总体:  $X \sim N(0, 1)$ . 称统计量:

$$\chi^2 = \sum_{i=1}^n X_n^2$$

服从自由度为  $n$  的  $\chi^2$  分布. 记为  $X \sim \chi_n^2$ .

概率密度:

$$f(x) = \frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}} e^{-\frac{x}{2}} x^{\frac{n-2}{2}}$$

性质:

1.  $E(\chi^2) = n$ ;
2.  $D(\chi^2) = 2n$ .
3. 若  $X \sim \chi_m^2, Y \sim \chi_n^2$ ,  $X$  与  $Y$  相互独立, 则  $Z = X + Y \sim \chi_{m+n}^2$ ;
4. 若  $X \sim \chi_n^2$ , 记  $P\{X > c\} = \alpha \implies c = \chi_n^2(\alpha)$  (上  $\alpha$  分位点).

### t 分布

定义 5.2.6. 总体  $X \sim N(0, 1), Y \sim \chi_n^2$ .  $X$  与  $Y$  相互独立.

$$t = \frac{X}{\sqrt{Y/n}}$$

服从自由度为  $n$  的  $t$  分布. 记作:  $t \sim t_n$ .

性质:

1.  $E(t_n) = 0 \quad n \geq 2$ ;
2.  $D(t_n) = \frac{n}{n-2} \quad n \geq 3$ .

## F 分布

定义 5.2.7. 若  $X \sim \chi_n^2$ ,  $Y \sim \chi_m^2$ .  $X$  与  $Y$  相互独立, 则:

$$F = \frac{X/m}{Y/n}$$

服从自由度为  $n, m$  的  $F$  分布. 记作  $F \sim F_{m,n}$ .

性质:

$$1. E(F_{m,n}) = \frac{n}{n-2}, D(F_{m,n}) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$$

## 5.3 正态总体样本均值与样本方差的分布

## 5.3.1 四个等式

假设有总体  $X$ , 有均值  $\mu$ , 方差  $\sigma^2$ . 随机变量序列  $X_1, X_2, \dots, X_n$  为来自  $X$  的一个样本. 样本均值  $\bar{X}$ , 样本方差  $S^2$ . 有:

$$E(\bar{X}) = \mu \quad (5.1)$$

$$D(\bar{X}) = \frac{\sigma^2}{n} \quad (5.2)$$

$$E(S^2) = \sigma^2 \quad (5.3)$$

$$D(S^2) = \frac{2\sigma^4}{n-1} \quad (5.4)$$

例 5.3.1. 设总体  $X$  服从正态分布  $N(\mu, \sigma^2)$ . 从该总体中抽取简单随机样本  $X_i, i = 1, \dots, 2n$ . 其样本均值为  $\bar{X} = \frac{1}{2n} \sum_{i=1}^{2n} X_i$ , 试求统计量:

$$Y = \sum_{i=1}^n (X_i + X_{i+n} - 2\bar{X})^2 \quad (5.5)$$

的数学期望  $E(Y)$ .

解. 考虑将总体  $X$  进行拆分:

$$T_i = X_i + X_{i+n} \sim N(2\mu, 2\sigma)$$

形成新的总体  $T \sim N(2\mu, 2\sigma^2)$ . 可求得样本  $\{T_i\}$  的均值  $\bar{T}$ .

$$\bar{T} = \frac{1}{n} \sum_{i=1}^n T_i = \frac{1}{n} \sum_{i=1}^{2n} X_i = 2\bar{X}$$

样本方差:

$$S_T = \frac{1}{n-1} \sum_{i=1}^n (T_i - \bar{T})^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i + X_{i+n} - 2\bar{X})^2 = \frac{Y}{n-1}$$

因此:

$$E(Y) = (n-1)E(S_T) = (n-1)\sigma_T^2 = 2(n-1)\sigma^2$$

□

### 5.3.2 四个定理

设样本  $\{X_i\} (i = 1, \dots, m)$  来自总体  $X \sim N(\mu_1, \sigma_1^2)$ . 样本  $\{Y_i\} (i = 1, \dots, n)$  来自总体  $Y \sim N(\mu_2, \sigma_2^2)$ .

**定理 5.3.1.** 独立的正态随机变量线性组合服从正态分布, 即

$$T = \sum_{k=1}^m c_k X_k \sim N \left( \mu_1 \sum_{k=1}^m c_k, \sigma_1^2 \sum_{k=1}^m c_k^2 \right)$$

特别地, 当  $c_1 = c_2 = \dots = c_k$  时, 得到的就是样本均值的分布:

$$\bar{X} \sim N \left( \mu_1, \frac{\sigma_1^2}{b} \right)$$

**定理 5.3.2.** 样本方差  $S_X^2 = \frac{1}{n-1} \sum_{i=1}^m (X_i - \bar{X})^2$ . 则

$$T = \frac{(m-1)S_X^2}{\sigma_1^2} \sim \chi_{m-1}^2$$

且  $\bar{X}$  与  $S^2$  相互独立. 进而:

$$W = \frac{\sqrt{m}(\bar{X} - \mu_1)}{S_X} \sim t_{m-1}$$

**定理 5.3.3.**  $X$  与  $Y$  相互独立, 则:

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_T} \sqrt{\frac{mn}{n+m}} \sim t_{m+n-2}$$

其中  $(n+m-2)S_T^2 = (m-1)S_X^2 + (n-1)S_Y^2$ , 且:

$$S_X^2 = \frac{1}{m-1} \sum_{i=1}^m (X_i - \bar{X})^2 \qquad S_Y^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2$$

定理 5.3.4.  $X$  与  $Y$  相互独立, 则:

$$F = \frac{S_X^2}{S_T^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{m-1, n-1}$$

例 5.3.2. 设总体  $X \sim N(\mu_1, \sigma_1^2)$ ,  $Y \sim N(\mu_2, \sigma_2^2)$ . 从两个总体中分别抽样得:  $n_1 = 8$  求概率  $P\{\sigma_1^2 > \sigma_2^2\}$ .

解. 首先对原式做变换:

$$P\{\sigma_1^2 > \sigma_2^2\} = P\left\{\frac{\sigma_1^2}{\sigma_2^2} > 1\right\} = P\left\{\frac{S_Y}{S_X} \cdot \frac{\sigma_1^2}{\sigma_2^2} > \frac{S_Y}{S_X}\right\} = \alpha$$

$\Rightarrow$

$$\alpha = 0.95$$

□

# 第六章 参数点估计

## 6.1 概念

定义 6.1.1. 用样本的统计量来估计总体的参数.

估计方法

$$\text{点估计} = \begin{cases} a: \text{矩估计法} \\ b: \text{最大似然估计法} \end{cases} \quad \text{区间估计} = \begin{cases} a: \text{估计} \\ b: \text{概率} \end{cases}$$

## 6.2 点估计

定义 6.2.1. 总体  $X$  的分布函数的形式已知:  $F(x; \theta)$ , 用样本来估计总体参数的值. 称这种估计方法为点估计.

例 6.2.1. 在某炸药厂, 一天中发生着火现象的次数  $X$  是一个随机变量. 假设它服从以  $\lambda > 0$  为参数的泊松分布, 参数  $\lambda$  未知. 现有以下的样本值. 试估计参数  $\lambda$ .

着火次数 $k$	0	1	2	3	4	5	6	$\geq 7$	
发生 $k$ 次着火的天数 $n_k$	75	90	54	22	6	2	1	0	$\sum = 250$

解. 已知泊松分布的参数  $\lambda$  为其总体的期望, 因此用样本均值来估计:

$$\bar{X} = \frac{1}{250} \sum_{k=0}^{\infty} kn_k = 1.22$$

因此  $\lambda$  的估计值为 1.22. □

定义 6.2.2.