

概率论与数理统计

ZZX-JLU

2022 年 11 月 8 日

目录

1	随机事件及其概率	1
1.1	随机试验	1
1.2	随机事件	2
1.2.1	随机事件的概念	2
1.2.2	随机事件的关系	2
1.2.3	随机事件的运算	3
1.3	随机事件的概率	5
1.3.1	频率	5
1.3.2	概率	6
1.3.3	古典概型	9
1.3.4	几何概型	10
1.4	条件概率	11
1.4.1	条件概率与乘法公式	11
1.4.2	全概率公式	12
1.4.3	贝叶斯公式	12
1.5	事件的独立性	13
1.6	伯努利概型	14
2	随机变量及其分布	16
2.1	随机变量的分布函数	16
2.1.1	随机变量	16
2.1.2	分布函数	16
2.2	离散型随机变量及其概率分布	17
2.3	连续型随机变量及其概率密度	18

2.4	常用的分布	20
2.4.1	(0-1) 分布	20
2.4.2	二项分布	20
2.4.3	泊松分布	21

1 随机事件及其概率

1.1 随机试验

在一定条件下必然出现的现象叫做**必然现象**. 在相同的条件下, 可能出现不同的结果, 而在试验或观测之前不能预知确切结果的现象叫做**随机现象**.

随机现象具有随机性和统计规律性.

- 随机性: 对随机现象进行观测时, 不能预先确定其结果.
- 统计规律性: 对随机现象进行大量重复观测后, 其结果往往会表现出某种规律性.

为了研究和揭示随机现象的统计规律性, 需要对随机现象进行大量重复的观察、测量或试验, 统称为试验.

如果试验具有以下特点:

1. 可重复性: 试验可以在相同条件下重复进行多次, 甚至进行无限次;
2. 可观测性: 每次试验的所有可能结果都是明确的、可以观测的, 并且试验的可能结果有两个或两个以上;
3. 随机性: 每次试验出现的结果是不确定的, 在试验之前无法预先确定究竟会出现哪一个结果,

则称之为**随机试验**, 简称为**试验**.

通常用字母 E 表示一个随机试验. 随机试验 E 的基本结果称为**样本点**, 用 ω 表示. 随机试验 E 的所有基本结果的集合称为**样本空间**, 用 $\Omega = \{\omega\}$ 表示.

1.2 随机事件

1.2.1 随机事件的概念

随机试验 E 的样本空间 $\Omega = \{\omega\}$ 的子集称为随机试验 E 的**随机事件** (random event), 简称为**事件** (event), 用大写字母 A, B, C 等表示.

设 $A \subseteq \Omega$, 如果试验结果 $\omega \in A$, 则称在这次试验中事件 A 发生; 如果 $\omega \notin A$, 则称事件 A 不发生.

由一个样本点 ω 组成的事件称为**基本事件**.

样本空间 Ω 本身也是 Ω 的子集, 它包含 Ω 的所有样本点, 在每次试验中 Ω 必然发生, 称为**必然事件**.

空集 \emptyset 也是 Ω 的子集, 它不包含任何样本点, 在每次试验中都不可能发生, 称为**不可能事件**.

在一个样本空间中, 如果只有有限个样本点, 则称它为**有限样本空间**; 如果有无限个样本点, 则称它为**无限样本空间**.

1.2.2 随机事件的关系

事件的包含

如果当事件 A 发生时事件 B 一定发生, 则称事件 B **包含** 事件 A , 记作 $A \subseteq B$.

对于任意事件 A , 有 $\emptyset \subseteq A \subseteq \Omega$.

如果 $A \subseteq B, B \subseteq C$, 则 $A \subseteq C$.

事件的相等

如果事件 A 和事件 B 相互包含, 即 $A \subseteq B$ 且 $B \subseteq A$, 则称事件 A 与事件 B **相等**, 记作 $A = B$.

事件的互不相容

如果事件 A 和事件 B 在同一次试验中不能同时发生, 则称事件 A 与事件 B 是**互不相容**的, 或称事件 A 与事件 B 是**互斥**的.

任意两个基本事件一定互斥.

事件的互逆

如果在每一次试验中事件 A 和事件 B 必有一个且仅有一个发生, 则称事件 A 与事件 B 是互逆的或对立的, 称其中的一个事件是另一个事件的逆事件, 记作 $\bar{A} = B$, 或 $\bar{B} = A$.

显然, $\overline{\bar{A}} = A$.

1.2.3 随机事件的运算

事件的并

如果事件 A 和事件 B 至少有一个发生, 则这样的—个事件称为事件 A 与事件 B 的并事件或和事件, 记作 $A \cup B$.

$$A \cup B = \{\omega \mid \omega \in A \text{ 或 } \omega \in B\}$$

事件 A 和事件 B 作为样本空间 Ω 的子集, 并事件 $A \cup B$ 就是子集 A 与 B 的并集. 对于任何事件 A 与 B , 有

$$A \cup A = A$$

$$A \cup \emptyset = A$$

$$A \cup B = B \cup A$$

$$A \cup \bar{A} = \Omega$$

$$A \subseteq A \cup B$$

$$B \subseteq A \cup B$$

如果 $A \subseteq B$, 则有 $A \cup B = B$.

事件的并可以推广到多个事件的情形:

$$\bigcup_{i=1}^n A_i = \{\text{事件 } A_1, A_2, \dots, A_n \text{ 中至少有一个发生}\}$$

$$\bigcup_{i=1}^{\infty} A_i = \{\text{事件 } A_1, A_2, \dots, A_n, \dots \text{ 中至少有一个发生}\}$$

事件的交

如果事件 A 和事件 B 同时发生, 则这样的—个事件称为事件 A 与事件 B 的交事件或积事件, 记作 $A \cap B$ 或 AB .

$$A \cap B = \{\omega \mid \omega \in A \text{ 且 } \omega \in B\}$$

事件 A 和事件 B 作为样本空间 Ω 的子集, 交事件 $A \cap B$ 就是子集 A 与 B 的交集. 对于任何事件 A 与 B , 有

$$A \cap A = A$$

$$A \cap \emptyset = \emptyset$$

$$A \cap B = B \cap A$$

$$A \cap \bar{A} = \emptyset$$

$$A \cap B \subseteq A$$

$$A \cap B \subseteq B$$

如果 $A \subseteq B$, 则有 $A \cap B = A$.

如果 A 与 B 互不相容, 则有 $A \cap B = \emptyset$.

事件的交可以推广到多个事件的情形:

$$\bigcap_{i=1}^n A_i = \{\text{事件 } A_1, A_2, \dots, A_n \text{ 同时发生}\}$$

$$\bigcap_{i=1}^{\infty} A_i = \{\text{事件 } A_1, A_2, \dots, A_n, \dots \text{ 同时发生}\}$$

事件的差

如果事件 A 发生而事件 B 不发生, 则这样的—个事件称为事件 A 与事件 B 的差事件, 记作 $A - B$.

$$A - B = \{\omega \mid \omega \in A \text{ 且 } \omega \notin B\}$$

对于任何事件 A 与 B , 有

$$A - A = \emptyset$$

$$A - \emptyset = A$$

$$A - B = A - AB = A\bar{B}$$

$$\Omega - A = \bar{A}$$

$$A - \Omega = \emptyset$$

$$(A - B) \cup B = (B - A) \cup A = A \cup B$$

$$A \cup B = A \cup (B - AB) = B \cup (A - AB)$$

$A - B, AB, B - A$ 两两互斥, 且 $A \cup B = (A - B) \cup AB \cup (B - A)$, $A = (A - B) \cup AB$, $B = (B - A) \cup AB$.

随机事件的运算规律

1. 交换律: $A \cup B = B \cup A$, $AB = BA$.
2. 结合律: $(A \cup B) \cup C = A \cup (B \cup C)$, $(AB)C = A(BC)$.
3. 分配律: $A(B \cup C) = (AB) \cup (AC)$, $A \cup (BC) = (A \cup B)(A \cup C)$.
4. 对偶律: $\overline{A \cup B} = \overline{A} \overline{B}$, $\overline{AB} = \overline{A} \cup \overline{B}$.

对于多个随机事件, 以上的运算规律也成立.

1.3 随机事件的概率

1.3.1 频率

定义 1.1 设在相同的条件下进行的 n 次试验中, 事件 A 发生了 n_A 次, 则称 n_A 为事件 A 发生的频数, 称比值 $\frac{n_A}{n}$ 为事件 A 发生的频率, 记作 $f_n(A)$, 即

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

事件 A 发生的频率反映了事件 A 在 n 次试验中发生的频繁程度. 频率越大, 表明事件 A 的发生越频繁, 从而可知事件 A 在一次试验中发生的可能性越大.

频率的基本性质:

性质 1(非负性) 对于任意事件 A , 有 $f_n(A) \geq 0$.

性质 2(规范性) 对于必然事件 Ω , 有 $f_n(\Omega) = 1$.

性质 3(有限可加性) 对于两两互不相容的事件 A_1, A_2, \dots, A_m (即当 $i \neq j$ 时, 有 $A_i A_j = \emptyset$, $i, j = 1, 2, \dots, m$), 有

$$f_n\left(\bigcup_{i=1}^m A_i\right) = \sum_{i=1}^m f_n(A_i)$$

在相同的条件下重复进行 n 次试验, 当 n 增大时, 事件 A 发生的频率 $f_n(A)$ 呈现出稳定性, 逐渐稳定于某一常数 p . 用这一常数表示事件 A 发生的可能性大小, 称为事件 A 的概率, 记为 $P(A)$, 即 $P(A) = p$.

当 n 很大时, 可以用频率 $f_n(A)$ 作为概率 $P(A)$ 的近似值.

1.3.2 概率

定义 1.2 设随机试验 E 的样本空间为 Ω , 如果对于 E 的每一个事件 A , 有唯一的实数 $P(A)$ 和它对应, 并且这个事件的函数 $P(A)$ 满足以下条件:

1. 非负性: 对于任意事件 A , 有 $P(A) \geq 0$;
2. 规范性: 对于必然事件 Ω , 有 $P(\Omega) = 1$;
3. 可列可加性: 对于两两互不相容的事件 A_1, A_2, \dots , 有

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

则称 $P(A)$ 为事件 A 的概率 (probability) .

性质 1 对于不可能事件 \emptyset , 有 $P(\emptyset) = 0$.

证明: 因为 $\emptyset = \emptyset \cup \emptyset \cup \dots$, 根据概率的可列可加性, 有

$$P(\emptyset) = P(\emptyset) + P(\emptyset) + \dots$$

由概率的非负性知 $P(\emptyset) \geq 0$, 因此 $P(\emptyset) = 0$. □

性质 2(有限可加性) 对于两两互不相容的事件 A_1, A_2, \dots, A_n , 有

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

证明: 令 $A_i = \emptyset$ ($i = n+1, n+2, \dots$), 根据概率的可列可加性及性质 1, 有

$$P\left(\bigcup_{i=1}^n A_i\right) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) = \sum_{i=1}^n P(A_i)$$

□

性质 3 对于任一事件 A , 有 $P(\bar{A}) = 1 - P(A)$.

证明: 因为 $A \cup \bar{A} = \Omega$, 且 $A\bar{A} = \emptyset$, 由性质 2 及概率的规范性, 得

$$P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A}) = 1$$

即

$$P(\bar{A}) = 1 - P(A)$$

□

性质 4 如果 $A \subseteq B$, 则有 $P(B - A) = P(B) - P(A)$, $P(A) \leq P(B)$.

证明: 因为 $A \subseteq B$, 从而有 $B = A \cup (B - A)$, 且 $A(B - A) = \emptyset$, 由性质 2 可得

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

所以

$$P(B - A) = P(B) - P(A)$$

由于 $P(B - A) \geq 0$, 因此 $P(A) \leq P(B)$.

□

性质 5 对于任一事件 A , 有 $P(A) \leq 1$.

证明: 因为 $A \subseteq \Omega$, 由性质 4 及概率的规范性, 可得

$$P(A) \leq P(\Omega) = 1$$

□

性质 6(概率的减法公式) 对于任意两个事件 A 与 B , 有

$$P(B - A) = P(B) - P(AB) \quad (1-1)$$

证明: 由于 $B - A = B - AB$, 而 $AB \subseteq B$, 根据性质 4 可得

$$P(B - A) = P(B - AB) = P(B) - P(AB)$$

□

性质 7 对于任意两个事件 A 与 B , 有

$$P(A \cup B) = P(A) + P(B) - P(AB) \quad (1-2)$$

$$P(A \cup B) \leq P(A) + P(B)$$

证明: 因为 $A \cup B = A \cup (B - AB)$, 且 $A(B - AB) = \emptyset$, $AB \subseteq B$, 由性质 2 及性质 4 可得

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

由于 $P(AB) \geq 0$, 因此 $P(A \cup B) \leq P(A) + P(B)$.

□

式 (1-2) 称为概率的**加法公式**.

加法公式可以推广到任意有限个事件的情形: 设 A_1, A_2, \dots, A_n 是 n 个随机事件, 则有

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i A_j) + \sum_{1 \leq i < j < k \leq n} P(A_i A_j A_k) + \dots + (-1)^{n-1} P(A_1 A_2 \dots A_n) \quad (1-3)$$

式 (1-3) 称为概率的**一般加法公式**.

【例 1.1】 设 A 和 B 是同一试验 E 的两个随机事件, 证明

$$1 - P(\bar{A}) - P(\bar{B}) \leq P(AB) \leq P(A \cup B)$$

证明: 因为 $AB \subseteq A \subseteq (A \cup B)$, 所以

$$P(AB) \leq P(A \cup B)$$

由概率的性质 7、性质 3 及事件的对偶律, 可得

$$P(\bar{A}) + P(\bar{B}) \geq P(\bar{A} \cup \bar{B}) = P(\overline{AB}) = 1 - P(AB)$$

因此

$$1 - P(\bar{A}) - P(\bar{B}) \leq P(AB)$$

□

1.3.3 古典概型

如果随机试验具有以下两个特点:

1. 试验的样本空间只包含有限个样本点;
2. 在试验中每个基本事件发生的可能性相同,

则称这种试验为**等可能概型**或**古典概型** (classic probability model) .

设试验 E 是古典概型, 样本空间为 $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$, 则基本事件 $\{\omega_1\}, \{\omega_2\}, \dots, \{\omega_n\}$ 两两互不相容, 且

$$\Omega = \{\omega_1\} \cup \{\omega_2\} \cup \dots \cup \{\omega_n\}$$

由于 $P(\Omega) = 1$ 及 $P(\{\omega_1\}) = P(\{\omega_2\}) = \cdots = P(\{\omega_n\})$, 因此

$$P(\{\omega_1\}) = P(\{\omega_2\}) = \cdots = P(\{\omega_n\}) = \frac{1}{n}$$

如果事件 A 包含 k 个基本事件, $A = \{\omega_{i_1}\} \cup \{\omega_{i_2}\} \cup \cdots \cup \{\omega_{i_k}\}$, 其中 i_1, i_2, \cdots, i_k 是 $1, 2, \cdots, n$ 中某 k 个不同的数, 则有

$$P(A) = P(\{\omega_{i_1}\}) + P(\{\omega_{i_2}\}) + \cdots + P(\{\omega_{i_k}\}) = \frac{k}{n}$$

即

$$P(A) = \frac{A \text{ 包含的基本事件个数}}{\Omega \text{ 包含的基本事件总数}}$$

1.3.4 几何概型

如果随机试验是将一个点随机地投到某一区域 Ω 内, 而这个随机点落在 Ω 中任意两个度量相等的子区域内的可能性是一样的, 则称这样的试验属于**几何概型** (geometric probability model) .

注: Ω 可以是直线上的某一区间, 也可以是平面或空间内的某一区域. 区域的度量是指直线上区间的长度, 或者平面内区域的面积, 或者空间内区域的体积.

对于任何有度量的子区域 $A \subseteq \Omega$, 定义事件 $A = \text{“随机点落在区域 } A \text{ 内”}$ 的概率为

$$P(A) = \frac{A \text{ 的度量}}{\Omega \text{ 的度量}}$$

【例 1.2】 在平面上画有等距离的平行线, 平行线间的距离为 $2a$ ($a > 0$). 向该平面任意投掷一枚长为 $2l$ ($l < a$) 的圆柱形的针, 求此针与任一平行线相交的概率.

解: 针投在该平面上, 设 x 为针的中点 M 到最近的一条平行线的距离, φ 为针与此直线的交角, 如图 1.1 所示, 则有 $0 \leq x \leq a$, $0 \leq \varphi \leq \pi$.

针与最近的一条平行线相交的充分必要条件是 $x \leq l \sin \varphi$.

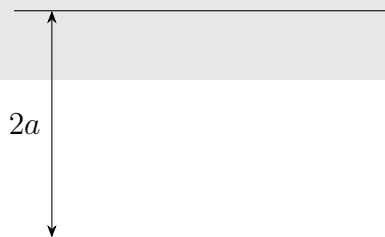


图 1.1

1.4 条件概率

1.4.1 条件概率与乘法公式

定义 1.3 设 A 和 B 是试验 E 的两个事件, 且 $P(A) > 0$, 称 $\frac{P(AB)}{P(A)}$ 为在事件 A 已经发生的条件下, 事件 B 发生的**条件概率** (conditional probability), 记为 $P(B | A)$, 即

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

对于任意两个事件 A 和 B , 如果 $P(A) > 0$, 则有

$$P(AB) = P(A) P(B | A) \quad (1-4)$$

式 (1-4) 称为概率的**乘法公式**.

同样可以在 $P(B) > 0$ 时, 定义在事件 B 已经发生的条件下, 事件 A 发生的条件概率为

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

在 $P(A) > 0, P(B) > 0$ 的条件下, 有

$$P(AB) = P(A) P(B | A) = P(B) P(A | B)$$

条件概率具有如下性质:

性质 1(非负性) 对任意事件 B , 有 $P(B | A) \geq 0$.

性质 2(规范性) 对于必然事件 Ω , 有 $P(\Omega | A) = 1$.

性质 3(可列可加性) 对于两两互不相容的事件 B_1, B_2, \dots , 有

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

可由条件概率的三个基本性质推导出其他性质, 例如

$$P(\emptyset | A) = 0$$

$$P(\bar{B} | A) = 1 - P(B | A)$$

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

可以把乘法公式推广到有限个事件的交的情况: 设 A_1, A_2, \dots, A_n 是同一试验的事件, 且 $P(A_1 A_2 \cdots A_{n-1}) > 0$, 则有

$$P(A_1 A_2 \cdots A_n) = P(A_1) P(A_2 | A_1) P(A_3 | A_1 A_2) \cdots P(A_n | A_1 A_2 \cdots A_{n-1}) \quad (1-5)$$

1.4.2 全概率公式

设试验 E 的样本空间为 Ω , 事件 A_1, A_2, \dots, A_n 两两互不相容, 且 $\bigcup_{i=1}^n A_i = \Omega$, 则称 A_1, A_2, \dots, A_n 为样本空间 Ω 的一个**分割**或**完全事件组**.

如果 $P(A_i) > 0$ ($i = 1, 2, \dots, n$), 则对任意事件 B , 有

$$B = B\Omega = B \left(\bigcup_{i=1}^n A_i \right) = \bigcup_{i=1}^n (A_i B)$$

这里 $(A_i B) \cap (A_j B) = \emptyset$ ($i \neq j, i, j = 1, 2, \dots, n$), 由概率的有限可加性得

$$P(B) = P\left(\bigcup_{i=1}^n (A_i B)\right) = \sum_{i=1}^n P(A_i B)$$

由乘法公式得

$$P(B) = \sum_{i=1}^n P(A_i) P(B | A_i) \quad (1-6)$$

式 (1-6) 称为**全概率公式** (total probability formula) .

1.4.3 贝叶斯公式

设试验 E 的样本空间为 Ω , 事件 A_1, A_2, \dots, A_n 是 Ω 的一个分割, 且 $P(A_i) > 0$ ($i = 1, 2, \dots, n$). 对于任一事件 B , 如果 $P(B) > 0$, 由乘法公式可得

$$P(A_i B) = P(B) P(A_i | B) = P(A_i) P(B | A_i)$$

由此得

$$P(A_i | B) = \frac{P(A_i) P(B | A_i)}{P(B)}$$

利用全概率公式, 得

$$P(A_i | B) = \frac{P(A_i) P(B | A_i)}{\sum_{j=1}^n P(A_j) P(B | A_j)} \quad (1-7)$$

式 (1-7) 称为贝叶斯公式 (Bayes formula) .

1.5 事件的独立性

定义 1.4 设 A 与 B 是同一试验 E 的两个事件, 如果 $P(AB) = P(A)P(B)$, 则称事件 A 与事件 B 是相互独立的.

对于同一试验 E 的两个事件 A 与 B , 如果 $P(A) > 0$, 则 A 与 B 相互独立的充分必要条件是 $P(B | A) = P(B)$; 如果 $P(B) > 0$, 则 A 与 B 相互独立的充分必要条件是 $P(A | B) = P(A)$.

结论: 如果事件 A 与事件 B 相互独立, 则事件 A 与事件 \bar{B} 相互独立.

证明: $A = A(B \cup \bar{B}) = (AB) \cup (A\bar{B})$, 而 $(AB)(A\bar{B}) = \emptyset$, 所以

$$P(A) = P((AB) \cup (A\bar{B})) = P(AB) + P(A\bar{B})$$

如果 A 与 B 相互独立, 则 $P(AB) = P(A)P(B)$, 代入上式可得

$$P(A) = P(A)P(B) + P(A\bar{B})$$

由此得

$$\begin{aligned} P(A\bar{B}) &= P(A) - P(A)P(B) \\ &= P(A)[1 - P(B)] \\ &= P(A)P(\bar{B}) \end{aligned}$$

因此事件 A 与事件 \bar{B} 是相互独立的. □

同理, 如果事件 A 与事件 B 相互独立, 则事件 \bar{A} 与事件 B 相互独立, 事件 \bar{A} 与事件 \bar{B} 相互独立.

定义 1.5 对于同一试验 E 的三个事件 A, B, C , 如果满足

$$P(AB) = P(A)P(B)$$

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

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

则称三个事件 A, B, C 是**两两相互独立**的.

定义 1.6 如果三个事件 A, B, C 是两两相互独立的, 并且有 $P(ABC) = P(A)P(B)P(C)$, 则称三个事件 A, B, C 是**相互独立**的.

定义 1.7 设 A_1, A_2, \dots, A_n 是同一试验 E 的 n 个事件, 如果对于任意正整数 k 及这 n 个事件中的任意 k ($2 \leq k \leq n$) 个事件 $A_{i_1}, A_{i_2}, \dots, A_{i_k}$, 都有等式

$$P(A_{i_1}A_{i_2}\cdots A_{i_k}) = P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k})$$

则称这 n 个事件 A_1, A_2, \dots, A_n 是**相互独立**的.

1.6 伯努利概型

如果将试验 E 重复执行 n 次, 在每一次试验中, 事件 A 或者发生, 或者不发生. 假设每次试验的结果互不影响, 即在每次试验中事件 A 发生的概率保持不变, 不受其他各次试验结果的影响, 则称这 n 次试验相互独立.

如果试验 E 只有两个可能的对立结果 A 和 \bar{A} , 并且 $P(A) = p$, $P(\bar{A}) = 1 - p$, 其中 $0 < p < 1$. 将试验 E 独立地重复进行 n 次所构成的一个试验叫做 n 重伯努利试验, 简称为伯努利试验 (Bernoulli experiment) 或伯努利概型 (Bernoulli probability model).

n 重伯努利试验的基本事件可记为 $\omega = \omega_1\omega_2\cdots\omega_n$, 其中 ω_i ($1 \leq i \leq n$) 为 A 或者为 \bar{A} , 即 ω 是从 A 及 \bar{A} 中每次取 1 个, 独立地重复取 n 次的一种排列, 共有 2^n 个基本事件.

如果 ω 中有 k 个 A , 则必有 $n-k$ 个 \bar{A} , 由独立性可得这一基本事件的概率为 $p^k(1-p)^{n-k}$.

由于在 2^n 个基本事件中, 共有 C_n^k 个含 k 个 A 及 $n-k$ 个 \bar{A} , 因此在 n 次独立重复试验中, 事件 A 恰好发生 k 次的概率 $P_n(k)$ 为

$$P_n(k) = C_n^k p^k (1-p)^{n-k} \quad k = 0, 1, 2, \dots, n \quad (1-8)$$

由二项式定理可得

$$\sum_{k=0}^n P_n(k) = \sum_{k=0}^n C_n^k p^k (1-p)^{n-k} = [p + (1-p)]^n = 1$$

由此可见, $C_n^k p^k (1-p)^{n-k}$ 是二项展开式中的一项, 因此式 (1-8) 又称为二项概率公式.

2 随机变量及其分布

2.1 随机变量的分布函数

2.1.1 随机变量

定义 2.1 设随机试验 E 的样本空间为 $\Omega = \{\omega\}$. 如果对于每一个 $\omega \in \Omega$, 都有一个实数 $X(\omega)$ 与之对应, 则称 $X = X(\omega)$ 为**随机变量** (random variable) .

随机变量常用大写字母 X, Y, Z 等表示.

2.1.2 分布函数

定义 2.2 设 X 是一个随机变量, 对于任意实数 x , 令 $F(x) = P\{X \leq x\}$, 称 $F(x)$ 为随机变量 X 的**分布函数** (cumulative distribution function) .

随机变量 X 的分布函数 $F(x)$ 是定义在 $(-\infty, +\infty)$ 上的函数, 是随机事件 $\{X \leq x\}$ 发生的概率. 分布函数值 $F(a)$ 表示 X 落在区间 $(-\infty, a]$ 上的概率.

分布函数的基本性质如下.

性质 1 对于任意实数 x , 有 $0 \leq F(x) \leq 1$.

性质 2 对于任意实数 $x_1, x_2 (x_1 < x_2)$, 有 $P\{x_1 < X \leq x_2\} = F(x_2) - F(x_1)$.

证明：对于任意实数 $x_1, x_2 (x_1 < x_2)$ ，由于

$$\{x_1 < X \leq x_2\} = \{X \leq x_2\} - \{X \leq x_1\}$$

所以有

$$\begin{aligned} P\{x_1 < X \leq x_2\} &= P\{X \leq x_2\} - P\{X \leq x_1\} \\ &= F(x_2) - F(x_1) \end{aligned}$$

□

性质 3

$$\begin{aligned} F(-\infty) &= \lim_{x \rightarrow -\infty} F(x) = 0 \\ F(+\infty) &= \lim_{x \rightarrow +\infty} F(x) = 1 \end{aligned}$$

性质 4 $F(x)$ 处处右连续，即 $F(x^+) = F(x)$.

性质 5 $F(x)$ 是一个单调不减函数.

证明：对于任意实数 $x_1, x_2 (x_1 < x_2)$ ，有

$$F(x_2) - F(x_1) = P\{x_1 < X \leq x_2\} \geq 0$$

因此 $F(x)$ 是单调不减函数.

□

2.2 离散型随机变量及其概率分布

定义 2.3 如果一个随机变量 X 所有可能取到的不相同的值是有限个或可列无限多个，并且以确定的概率取这些不同的值，则称 X 为离散型随机变量 (discrete random variable).

定义 2.4 设离散型随机变量 X 所有可能取的值为 $x_k (k = 1, 2, \dots)$, X 取各个可能值的概率, 即事件 $\{X = x_k\}$ 的概率为

$$P\{X = x_k\} = p_k \quad k = 1, 2, \dots \quad (2-1)$$

并且 p_k 满足以下两个条件:

1. $p_k \geq 0$;
2. $\sum_{k=1}^{\infty} p_k = 1$,

则称式 (2-1) 为离散型随机变量 X 的**概率分布** (probability distribution) 或**分布律**.

概率分布也可以用如下的表格来表示:

X	x_1	x_2	\cdots	x_k	\cdots
P	p_1	p_2	\cdots	p_k	\cdots

概率分布反映了离散型随机变量的统计规律性.

对于任意实数 x , 随机事件 $\{X \leq x\}$ 可以表示成 $\bigcup_{x_k \leq x} \{X = x_k\}$. 由于 $x_k (k = 1, 2, \dots)$ 互不相同, 根据概率的可加性, 可得离散型随机变量 X 的分布函数为

$$F(x) = P\{X \leq x\} = \sum_{x_k \leq x} P\{X = x_k\} = \sum_{x_k \leq x} p_k$$

2.3 连续型随机变量及其概率密度

定义 2.5 对于随机变量 X 的分布函数 $F(x)$, 如果存在非负函数 $f(x)$, 使得对任意的 x , 都有 $F(x) = \int_{-\infty}^x f(t) dt$, 则称随机变量 X 是**连续型随机变量** (continuous random variable), 其中函数 $f(x)$ 叫做 X 的**概率密度函数** (probability density function), 简称为**概率密度** (probability density), 记作 $X \sim f(x)$.

由定义 2.5 可知, 连续型随机变量的分布函数处处连续.

概率密度 $f(x)$ 的性质如下.

性质 1 $f(x) \geq 0$

性质 2 $\int_{-\infty}^{+\infty} f(x) dx = 1$

满足性质 1 与性质 2 的函数 $f(x)$ 必然是某随机变量的概率密度.

性质 3 对于任意实数 $a, b (a < b)$, 有

$$P\{a < X \leq b\} = F(b) - F(a) = \int_a^b f(x) dx$$

由以上性质可知, 概率密度曲线总是位于 x 轴上方, 并且介于它和 x 轴之间的面积等于 1; 随机变量落在区间 $(a, b]$ 的概率 $P\{a < X \leq b\}$ 等于区间 $(a, b]$ 上曲线 $y = f(x)$ 之下的曲边梯形的面积.

性质 4 如果 $f(x)$ 在点 x 处连续, 则有 $F'(x) = f(x)$.

由性质 4 可知, 在 $f(x)$ 的连续点有

$$\begin{aligned} f(x) &= \lim_{\Delta x \rightarrow 0^+} \frac{F(x + \Delta x) - F(x)}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0^+} \frac{P\{x < X \leq x + \Delta x\}}{\Delta x} \end{aligned}$$

可见概率密度反映了随机变量在点 x 处概率分布的密集程度. $f(x)$ 的大小能反映出随机变量 X 在点 x 附近取值的可能性大小, 即概率的大小. 因此, 用概率密度描述连续型随机变量的分布比用分布函数更直观. 当不考虑高阶无穷小时, 有

$$P\{x < X \leq x + \Delta x\} \approx f(x)\Delta x$$

对于 X 的任意一个可取的值 x , 设 $\Delta x > 0$, 由于事件 $\{X = x\} \subseteq \{x - \Delta x < X \leq x\}$, 因此有

$$0 \leq P\{X = x\} \leq P\{x - \Delta x < X \leq x\} = F(x) - F(x - \Delta x)$$

令 $\Delta x \rightarrow 0$, 可得

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

因此, 连续型随机变量取任意指定实数的概率均为零. 据此, 在计算连续型随机变量在某一区间取值的概率时, 可以不区分该区间是开区间或闭区间或半开半闭区间, 即有

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

2.4 常用的分布

2.4.1 (0-1) 分布

定义 2.6 如果离散型随机变量 X 只取 0 与 1 两个值, 其概率分布为

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

或写成

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

则称随机变量 X 服从参数为 p 的 **(0-1) 分布**或**两点分布** (two-point distribution) .

服从两点分布的随机变量 X 的概率分布也可以写成

X	0	1
P	$1 - p$	p

2.4.2 二项分布

在 n 重伯努利试验中, 如果以 X 表示事件 A 出现的次数, 则 X 是一个离散型随机变量, 它的所有可能取值是 $0, 1, 2, \dots, n$. 设 $P(A) = p (0 < p < 1)$, 则由二项概率公式 (式 (1-8)) 可得

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

定义 2.7 如果随机变量 X 的概率分布为

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

则称随机变量 X 服从参数为 n, p 的**二项分布** (binomial distribution), 记作 $X \sim B(n, p)$.

由定义 2.7 可得

$$\begin{aligned} P\{X = k\} &\geq 0 \\ \sum_{k=0}^n P\{X = k\} &= \sum_{k=0}^n C_n^k p^k (1-p)^{n-k} = [p + (1-p)]^n = 1 \end{aligned}$$

特别地, 当 $n = 1$ 时, 二项分布 $B(1, p)$ 的概率分布为

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

这就是 (0-1) 分布. 因此, (0-1) 分布是二项分布的特例.

2.4.3 泊松分布

定义 2.8 如果离散型随机变量 X 的所有可能取值为 $0, 1, 2, \dots$, 并且

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

其中 $\lambda > 0$ 是常数, 则称随机变量 X 服从参数为 λ 的**泊松分布** (poisson distribution), 记作 $X \sim P(\lambda)$ 或 $X \sim \pi(\lambda)$.

易知

$$\begin{aligned} \frac{\lambda^k e^{-\lambda}}{k!} &> 0 \\ \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1 \end{aligned}$$

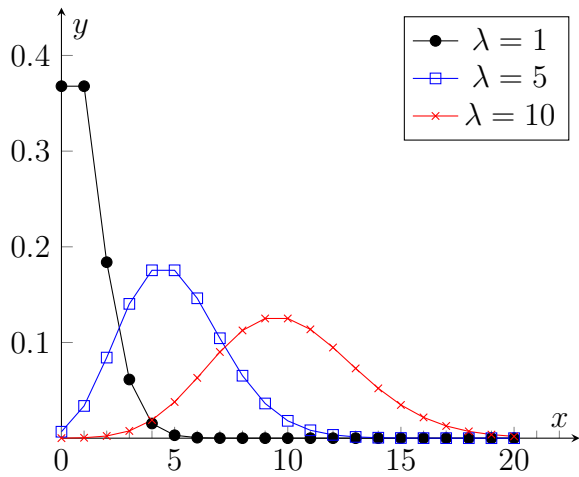


图 2.1 泊松分布