目 录

第一章　随机事件及其概率

在自然界和人类社会活动中，发生的现象是多种多样的．有一类现象在一定条件下必然会发生．例如，在标准大气压下，将水加热到100℃，水必然沸腾；向上抛一石子必然下落；两个同性的电荷一定互斥，等等．这类现象称为确定性现象．还有另一类现象，它在一定条件下可能发生，也可能不发生．例如，在相同条件下抛一枚均匀硬币，其结果可能是正面朝上，也可能是反面朝上，并且在每次抛掷之前无法肯定抛掷的结果是什么；新生婴儿，可能是男孩，也可能是女孩；向同一目标进行射击，可能命中目标，也可能不命中目标．我们称这一类现象为不确定性现象或随机现象．人们通过长期的实践及进一步研究后，发现这类现象在大量重复试验或观察下，其结果呈现出某种规律性．例如，一门火炮在一定条件下进行射击，个别炮弹的弹着点可能偏离目标而有随机性的误差，但大量炮弹的弹着点则表现出一定的规律性，如一定的命中率，一定的分布规律等．这种在个别试验中呈现不确定性，而在大量重复试验中又呈现某种统计规律性的现象叫随机现象．概率论和数理统计就是研究和揭示随机现象统计规律性的一门数学学科．它在自然科学与社会科学的各个领域中都有着广泛的应用．

§1.1　随机事件及运算

1.1.1　随机试验与随机事件

对随机现象进行研究时，人们通常要进行大量的观察、试验．在概率论中，把具有以下三个特征的试验称为随机试验．

(1) 可以在相同的条件下重复地进行；

(2) 每次试验的可能结果不止一个，并且事先明确试验的所有可能结果；

(3) 进行一次试验之前不能确定哪一个结果会出现．

随机试验简称为试验，是一个广泛的术语．它包括各种各样的科学试验，也包括对客观事物进行的“调查”、“观察”或“测量”等，随机试验通常用 E 来表示．下面举几个随机试验的例子：

E_1：抛一枚硬币，观察正面 H、反面 T 出现的情况；

E_2：抛两颗骰子，观察出现的点数之和；

E_3：记录某公共汽车站某日上午某时刻的等车人数；

E_4：在一批灯泡中任意抽取一只，测试它的寿命．

对每一随机试验，总要有一个观测目的，试验中可能观测到多种不同的结果．例如，从一批灯泡中任取一只进行通电试验．如果试验目的是检验产品是否合格，则试验结果为“合格品”或

“不合格品”；如果试验目的是测定其寿命，则试验结果为非负实数.

对于随机试验，尽管在每次试验之前不能预知其试验的结果，但试验的所有可能结果所组成的集合却是已知的. 我们称试验的所有可能结果组成的集合为样本空间，记为 S_i；样本空间的元素，即随机试验的每一个结果，称为样本点.

在前面所举的四个试验中，若以 S_i 表示试验 E_i 的样本空间，$i = 1,2,3,4$，则

$S_1 = \{H, T\}$；

$S_2 = \{2,3,\cdots,12\}$；

$S_3 = \{0,1,2,\cdots\}$；

$S_4 = \{t \mid t \geqslant 0\}$.

在试验中可能发生也可能不发生的事情称为随机事件，简称事件，以字母 $A, B, C, \cdots$ 表示. 有了样本空间的概念，我们可以定义样本空间 S 的任意一个子集为一个随机事件. 因此，随机事件就是试验的若干个结果组成的集合. 特别地，如果一个随机事件只含一个试验结果，则称此事件为基本事件.

在 E_1 中，$\{H\}$（表示“正面出现”）和 $\{T\}$（表示“反面出现”）为 E_1 的随机事件，它们都为基本事件.

在 E_2 中，$A = \{6\}$ 和 $B = \{k \mid k$ 为正整数且 $k > 6\}$ 为 E_2 的随机事件，E_2 的基本事件为 $A_k = \{k\}, k = 2,3,\cdots,12$.

在每次试验中，必然发生的事件称为必然事件，显然，样本空间作为一个事件为必然事件，记为 S. 例如，“上抛一石子必然下落” 就是必然事件. 空集 $\varnothing$ 不包含任何样本点，作为样本空间的子集，它在每次试验中必然不发生，称为不可能事件. 显然必然事件与不可能事件都是确定性的现象，但为了研究的方便，规定它们为随机事件.

1.1.2 事件的关系与运算

在实际问题中，往往要在同一个试验中同时研究几个事件以及它们之间的联系. 详细分析事件之间的关系，引进事件间的运算，不仅可以化复杂事件为简单事件，而且可以更好地解决相应的概率问题. 随机事件是一个集合，因此事件之间关系及运算可以按集合论中的知识处理，但应根据“事件发生与否” 给出它们在概率论中的提法.

设试验 E 的样本空间为 S，而 $A, B, C, A_k, B_k (k = 1,2,\cdots)$ 是 S 的子集.

1. 事件的包含与相等

若事件 A 中的每一个样本点都属于 B（见图 1－1），则称事件 B 包含事件 A，记作 $A \subset B$.

显然，这时事件 A 发生必然导致事件 B 发生. 故 B 包含 A，也常定义为：“若 A 发生必然导致 B 发生，则称 B 包含 A”.

对任一事件 A，有 $\varnothing \subset A \subset S$.

若 $A \subset B$ 且 $B \subset A$，则称事件 A 与事件 B 相等，记作 $A = B$.

2. 事件的和（或并）

至少属于 A 和 B 二者之一的所有样本点组成的集合（见图 1－2），称为事件 A 与事件 B 的和（或并），记作 $A \cup B$.

显然，事件 $A \cup B$ 发生，表示“事件 A 与事件 B 至少有一个发生”，即 $A \cup B = \{x \mid x \in A$ 或 $x \in B\}$. $A \cup B$ 也记为 $A + B$.

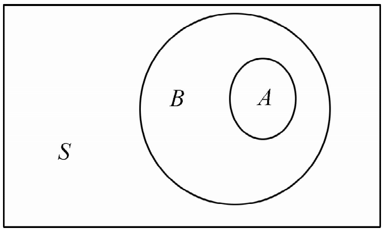

图 1－1

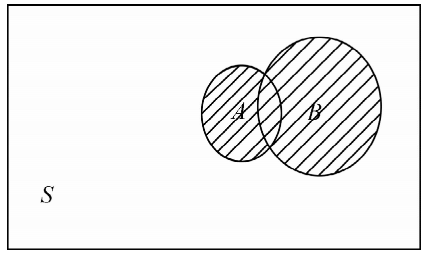

图 1－2

例如，在 E_2 中，若 A 表示"点数之和为偶数"，则 $A=\{2,4,6,8,10,12\}$，B 表示"点数之和大于 8"，$B=\{9,10,11,12\}$，则 $A\cup B=\{2,4,6,8,9,10,11,12\}$.

类似地，称 $\bigcup\limits_{i=1}^{n}A_i$ 为 n 个事件 $A_1,A_2,\cdots,A_n$ 的和事件；称 $\bigcup\limits_{i=1}^{\infty}A_i$ 为可列个事件 $A_1,A_2,\cdots$ 的和事件.

3. 事件的积(或交)

同时属于 A 与 B 的样本点的集合(见图 1－3) 称为 A 与 B 的积(或交)，记为 $A\cap B$ 或 AB.

显然，事件 AB 发生，表示"事件 A 与 B 同时发生"，即 $A\cap B=\{x\mid x\in A$ 且 $x\in B\}$.

例如，在 E_2 中，若设 $A=\{3,5,7,9\}$，$B=\{8,9,10\}$，则 $AB=\{9\}$.

$\bigcap\limits_{i=1}^{n}A_i$ 和 $\bigcap\limits_{i=1}^{\infty}A_i$ 的情况由读者自己完成.

4. 事件的差

包含在 A 中而不包含在 B 中的样本点的集合(见图 1－4)，称为 A 与 B 之差，记为 $A-B$.

显然，事件 $A-B$ 发生，表示"事件 A 发生，而事件 B 不发生". 即 $A-B=\{x\mid x\in A$ 且 $x\notin B\}$.

例如，在 E_2 中，若 $A=\{3,5,7,9\}$，$B=\{8,9,10\}$，则 $A-B=\{3,5,7\}$.

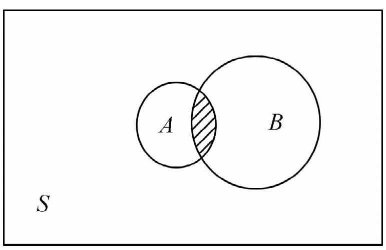

图 1－3

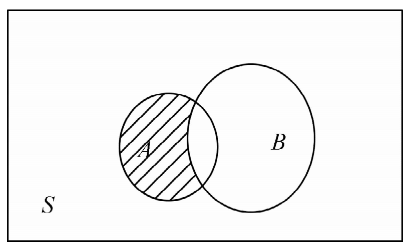

图 1－4

5. 事件互不相容(或互斥)

若 $A\cap B=\varnothing$，即事件 A 与事件 B 不能同时发生，或事件 A 与事件 B 没有共同的样本点，则称事件 A 与事件 B 是互不相容的，或互斥的，见图 1－5.

例如，在 E_2 中，若 $A=\{3,5,7,9,11\}$，$B=\{2,4,6,8,12\}$. 则 $AB=\varnothing$，因此 A 与 B 互斥.

在同一试验中，基本事件是两两互不相容的.

6. 对立(或逆) 事件

若 $A\cup B=S$ 且 $A\cap B=\varnothing$，则称事件 A 与事件 B 互为对立事件(或逆事件). 在一次试

验中，若事件 A 与 B 是对立事件，则其中必有一个发生，且仅有一个发生，A 的对立事件记为 $\overline{A}$，见图 1－6.

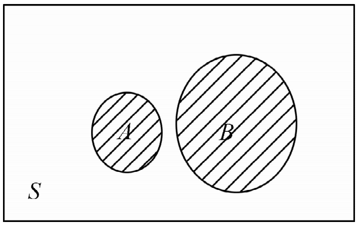

图 1－5

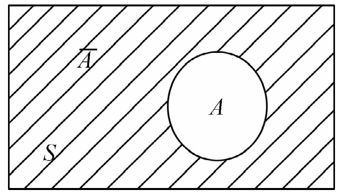

图 1－6

例如，在 E_2 中，若 $A=\{2,3,4\}$，$B=\{5,6,7,8,9,10,11,12\}$，则显然 $A\cup B=S$ 且 $A\cap B=\varnothing$，故 A 与 B 是对立事件，$B=\overline{A}$.

显然，$A\cup\overline{A}=S$，$A\cap\overline{A}=\varnothing$；

$A\cup S=S$，$A\cap S=A$；

$A\cup A=A$，$A\cap A=A$.

在进行事件的运算时，经常要用到如下规则：

交换律：$A\cup B=B\cup A$；$A\cap B=B\cap A$.

结合律：$(A\cup B)\cup C=A\cup(B\cup C)$；

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

分配律：$(A\cup B)\cap C=(A\cap C)\cup(B\cap C)$；

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

吸收律：$A\cup(A\cap B)=A$；

$A\cap(A\cup B)=A$.

对偶律：$\overline{A\cup B}=\overline{A}\cap\overline{B}$；$\overline{A\cap B}=\overline{A}\cup\overline{B}$.

对偶律在事件的运算中经常用到，它可以推广到更多个事件的情况，即

$$\overline{\bigcup_{i=1}^{n}A_i}=\bigcap_{i=1}^{n}\overline{A_i};$$

$$\overline{\bigcap_{i=1}^{n}A_i}=\bigcup_{i=1}^{n}\overline{A_i};$$

另外，还应注意到 $A-B=A\overline{B}$，$A=(AB)\cup(A\overline{B})$ 等.

例 1.1　设 A,B,C 表示三个随机事件，试将下列事件用 A,B,C 表示出来.

(1) A 出现，B,C 不出现；

(2) A,B 都出现，C 不出现；

(3) 三个事件至少有一个出现；

(4) 三个事件都不出现；

(5) 不多于一个事件出现；

(6) 三个事件至少有两个出现.

解　(1) A 出现，B,C 不出现：$A\overline{B}\,\overline{C}$；

(2) A,B 都出现，C 不出现：$AB\overline{C}$；

(3) 三个事件至少有一个出现：$A\cup B\cup C$；

(4) 三个事件都不出现：$\overline{A}\,\overline{B}\,\overline{C}$；

(5) 不多于一个事件出现，即 A,B,C 都不出现，或者 A 出现而 B,C 不出现，或者 B 出现而 A,C 不出现，或者 C 出现，A,B 不出现，故为：

$$\overline{A}\,\overline{B}\,\overline{C}+A\,\overline{B}\,\overline{C}+\overline{A}B\,\overline{C}+\overline{A}\,\overline{B}C;$$

(6) 三个事件至少有两个出现：

$$ABC+AB\,\overline{C}+A\,\overline{B}C+\overline{A}BC.$$

例 1.2　一工人生产了 n 个零件，设 A_i 表示“第 i 个零件是正品”$(i=1,2,\cdots,n)$. 试用文字叙述下列事件：(1) $\bigcap\limits_{i=1}^{n}A_i$，(2) $\overline{\bigcap\limits_{i=1}^{n}A_i}$，(3) $\bigcup\limits_{i=1}^{n}[\overline{A_i}\cap(\bigcap\limits_{\substack{k=1\\k\neq i}}^{n}A_k)]$.

解　(1) $\bigcap\limits_{i=1}^{n}A_i$：$n$ 个零件全是正品；

(2) $\overline{\bigcap\limits_{i=1}^{n}A_i}=\bigcup\limits_{i=1}^{n}\overline{A_i}$：至少有一个零件不是正品；

(3) $\bigcup\limits_{i=1}^{n}[\overline{A_i}\cap(\bigcap\limits_{\substack{k=1\\k\neq j}}^{n}A_k)]$：仅有一个零件不是正品.

§1.2　频率与概率

当做一个随机试验时，常常会发现有些事件出现的可能性大些，有些事件出现的可能性小些，有些事件出现的可能性彼此大致相同. 事件出现的可能性的大小，是客观存在的，它揭示了这些事件的内在统计规律. 在实际中常常希望知道某些事件在一次试验中发生的可能性究竟有多大. 例如，为了确定水坝的高度，就要知道河流在建造水坝地段每年洪水达到某一高度这一事件的可能性的大小. 为了合理地刻画事件在一次试验中发生的可能性大小，我们先引入频率的概念，进而引出表征事件在一次试验中发生的可能性大小的数字度量 —— 概率.

1.2.1　事件的频率

定义 1.1　设 A 是一个事件. 在相同条件下，进行 n 次试验，在这 n 次试验中，事件 A 发生的次数 n_A 称为事件 A 发生的频数. 比值 n_A/n 称为事件 A 发生的频率，并记成 $f_n(A)$.

由定义，不难发现频率具有如下性质：

(1) $0\leqslant f_n(A)\leqslant 1$；

(2) $f_n(S)=1$；

(3) 若 $A_1,A_2,\cdots,A_k$ 是两两互不相容的事件，则　$f_n(A_1\cup A_2\cup\cdots\cup A_k)=f_n(A_1)+f_n(A_2)+\cdots+f_n(A_k)$.

由于事件 A 在 n 次试验中发生的频率是它发生的次数与总试验次数之比，其大小表示 A 发生的频繁程度，频率越大，事件 A 发生的越频繁，这就意味着 A 在一次试验中发生的可能性越大. 因而，我们自然会想：能不能用频率来表示 A 在一次试验中发生的可能性的大小.

经验表明，当 n 较小时，频率 $f_n(A)$ 在 0 与 1 之间随机波动，其幅度较大，此时用频率来表示事件发生的可能性大小显然是不合适的. 而当 n 逐渐增大时，频率 $f_n(A)$ 逐渐稳定于某个常数. 例如，在抛硬币的试验中，观察出现正面的次数，请看下面抛硬币的实例，见表 1－1.

该试验表明，不管谁去抛硬币，当 n 逐渐增大时，频率 $f_n(H)$ 逐渐稳定于常数 0.5(表 1－1

中 n_H 表示 H 发生的频数，$f_n(H)$ 表示 H 发生的频率）. 这种在多次重复试验中出现的频率稳定性，通常称之为随机事件的统计规律性，它就是概率这一概念的经验基础. 我们称这个频率稳定值为随机事件 A 发生的概率，记为 $P(A)$.

表 1－1

试验者	n	n_H	$f_n(H)$
德摩根	2048	1061	0.5181
蒲丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

定义 1.2 在相同条件下进行大量重复试验，当试验次数充分大时，事件 A 的频率总在某个数值 p 附近摆动，则称 p 为事件 A 的概率，记为 $P(A)$，即 $P(A)=p$.

在后面章节中将会证明，当 n 很大时，用统计概率来度量事件发生的可能性大小是可行的.

1.2.2 事件的概率

在实际问题中，我们不可能，也没有必要对每个事件都做大量的试验，从中得到频率的稳定值，现在，我们从频率的稳定性和频率的性质出发，给出度量事件发生的可能性大小的量即概率的定义及性质.

定义 1.3 设 E 是随机试验，S 是它的样本空间. 对于 E 的每一事件 A，存在一实数，记作 $P(A)$，如果它满足条件

(1) 非负性：对任意事件 A，$P(A)\geqslant 0$；

(2) 规范性：$P(S)=1$；

(3) 可列可加性：对两两互不相容的事件 $A_k(k=1,2,\cdots)$ 有

$$P\left(\bigcup_{k=1}^{\infty} A_k\right)=\sum_{k=1}^{\infty} P(A_k) \tag{1.1}$$

则 $P(A)$ 为事件 A 的概率.

由定义 1.3 可以推得概率的一些重要性质.

性质 1.1 设 $\overline{A}$ 是 A 的对立事件，则

$$P(\overline{A})=1-P(A).$$

证明 因 $A\cup\overline{A}=S$，$A\cap\overline{A}=\varnothing$，$P(S)=1$，

故由概率的可加性，得 $1=P(S)=P(A\cup\overline{A})=P(A)+P(A)$

即得 $$P(\overline{A})=1-P(A)$$

性质 1.2 $P(\varnothing)=0$.

证明 由 $\overline{\varnothing}=S$，得 $P(\varnothing)=1-P(\overline{\varnothing})=1-P(S)=0$.

性质 1.3 若 $A_1,A_2,\cdots,A_n$ 是两两互不相容的事件，则有

$$P(A_1\cup A_2\cup\cdots\cup A_n)=P(A_1)+P(A_2)+\cdots+P(A_n).$$

证明 令 $A_{n+1}=A_{n+2}=\cdots=\varnothing$，则 $A_iA_j=\varnothing$，$i\neq j$，$i,j=1,2$，由概率的可列可加性得

$$P(A_1\cup A_2\cup\cdots\cup A_n)=P\left(\bigcup_{k=1}^{\infty} A_k\right)$$

$$= \sum_{k=1}^{n} P(A_k) + 0 = P(A_1) + P(A_2) + \cdots + P(A_n).$$

性质 1.4　设 A,B 为两个事件，且 $A \subset B$，则

$$P(A) \leqslant P(B), P(B-A) = P(B) - P(A).$$

证明　因为 $A \subset B$，所以 $B = A \cup (B-A)$（见图 1－1）

又$(B-A) \cap A = \varnothing$，得 $P(B) = P(A) + P(B-A)$.

于是　$P(B-A) = P(B) - P(A)$.

又因　$P(B-A) \geqslant 0$，故 $P(A) \leqslant P(B)$.

性质 1.5　对于任一事件 A，$P(A) \leqslant 1$.

证明　由 $A \subset S$，得 $P(A) \leqslant P(S) = 1$

故 $P(A) \leqslant 1$.

性质 1.6（加法公式）　对于任意两事件 A,B 有

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

证明　由图 1－2 可得 $A \cup B = A + (B - AB)$

又 $A \cap (B-AB) = \varnothing$，故 $P(A \cup B) = P(A) + P(B-AB)$.

由性质 1.4 得　$P(B-AB) = P(B) - P(AB)$

故

$$P(A \cup B) = P(A) + P(B) - P(AB) \tag{1.2}$$

由式(1.2)容易推广到三个事件，

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC) \tag{1.3}$$

对于 n 个事件的情形，可以用数学归纳法证得

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} p(A_i) - \sum_{1 \leqslant i < j \leqslant n} P(A_iA_j) + \sum_{1 \leqslant i < j < k \leqslant n} P(A_iA_jA_k) + \cdots + (-1)^{n-1} P(A_1A_2\cdots A_n) \tag{1.4}$$

例 1.3　设事件 A,B 的概率分别为$\frac{1}{3}$ 和$\frac{1}{2}$，求在下列三种情况下 $P(B\overline{A})$ 的值.

(1)A 与 B 互斥；(2)$A \subset B$；(3)$P(AB) = \frac{1}{8}$.

解　(1) 由图 1－5 得 $P(B\overline{A}) = P(B)$，

故　$P(B\overline{A}) = P(B) = \frac{1}{2}$.

(2) 由图 1－1 得　$P(B\overline{A}) = P(B) - P(A) = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$.

(3) 由图 1－4 得 $A \cup B = A \cup B\overline{A}$，且 $A \cap B\overline{A} = \varnothing$，

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

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

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

例 1.4　设 A,B 为两个事件，且 $P(A) = 0.6$，$P(B) = 0.5$，$P(A \cup B) = 0.7$，求 $P(A\overline{B})$，$P(\overline{AB})$，$P(\overline{A}\,\overline{B})$.

解 $P(AB)=P(A)+P(B)-P(A\cup B)=0.6+0.5-0.7=0.4$,

$P(A\overline{B})=P(A-B)=P(A-AB)=P(A)-P(AB)=0.2$,

$P(\overline{AB})=1-P(AB)=1-0.4=0.6$,

$P(\overline{A}\,\overline{B})=P(\overline{A\cup B})=1-P(A\cup B)=1-0.7=0.3.$

例 1.5 甲、乙两高射炮手,各自单独击中敌机的概率分别为 0.8 和 0.6,两人同时击中敌机的概率为 0.48,求敌机被击中的概率.

解 设 A 表示事件"甲击中敌机";B 表示事件"乙击中敌机";C 表示事件"敌机被击中". 由题意有 $C=A\cup B$,所以

$$\begin{aligned}P(C)&=P(A\cup B)=P(A)+P(B)-P(AB)\\&=0.8+0.6-0.48=0.92.\end{aligned}$$

§1.3 等可能概型(古典概型)

等可能概型是指在一次试验中,样本空间的每个样本点被取到的可能性相等的随机试验类型,这是一种最简单的概率类型.

1.3.1 古典概型

最先涉及的求概率问题都满足"各可能结果具有等可能性" 这一假设. 例如,在游戏中使用的骰子是一个匀质的正方体,使得掷出 1 ~ 6 各个点数的可能性相同,以保证游戏的公平.

定义 1.4 一般地,若随机试验 E 具有如下特点:

(1) 每次试验只有有限个可能的试验结果. 即试验的样本空间只包含有限个样本点,

$$S=\{e_1,e_2,\cdots,e_n\};$$

(2) 试验中每个基本事件发生的可能性相同,则对试验 E 中的任意事件 A,其概率的计算公式为

$$P(A)=\frac{m}{n}=\frac{A\text{ 包含的基本事件数}}{S\text{ 中基本事件的总数}}\tag{1.5}$$

这样定义的概率称作古典概率,具有这两个特点的试验称为等可能概型,它在概率论发展初期曾是主要研究对象,所以也称为古典概型. 用公式(1.5) 计算古典概率时,首先必须判断相应的试验一定是古典概型,在此基础上需知道事件 A 所包含的基本事件个数 m 及基本事件总数 n. 这种计算常常用到排列与组合的知识,现举例如下.

例 1.6 抛掷一颗匀质骰子,观察出现的点数,求"出现的点数是不小于 3 的偶数" 的概率.

解 设 A 表示"出现的点数是不小于 3 的偶数",则基本事件总数 $n=6$,A 包含的基本事件是"出现 4 点" 和"出现 6 点",即 $m=2$,故

$$P(A)=\frac{2}{6}=\frac{1}{3}.$$

从式(1.5) 不难得到古典概率的性质:

(1) 设 A 为任一事件,则 $0\leqslant P(A)\leqslant 1$;

(2) 必然事件的概率等于 1,即 $P(S)=1$;

(3) 设事件 $A_1,A_2,\cdots,A_m(m\leqslant n)$ 两两互不相容,则 $P(\bigcup\limits_{i=1}^{m}A_i)=\sum\limits_{i=1}^{m}P(A_i)$.

例 1.7　某城市的电话号码是7位数，某人忘记了他朋友的电话号码的后四位，于是随便拨号，求他拨一次号就拨通电话的概率.

解　由于此人记得前三位数，所以只考虑后四位数，每位数有 10 种取法，故样本点总数为 10^4，设 A 表示“拨到正确号码”，则 A 中只有一个样本点，于是

$$P(A)=\frac{1}{10^4}=0.0001.$$

这个事件的概率很小，我们称之为“小概率事件”. 人们在实践中总结出一条原理. 概率很小的事件在一次试验中几乎不可能发生. 这称为小概率实际推断原理.

例 1.8　一只箱中共装有 100 件某产品，其中有 8 件次品，余下为正品，今从中任取出 5 件，求(1) 至多 1 件次品，(2) 至少 2 件次品的概率.

解　设 A 表示“至多 1 件次品”，B 表示“至少 2 件次品”. 则

$$P(A)=\frac{C_8^0C_{92}^5+C_8^1C_{92}^4}{C_{100}^5}=0.95,$$

又显然 $B=\overline{A}$

故　$P(B)=1-P(A)=0.05.$

例 1.9　将 10 本书任意放在书架上，求其中指定的 3 本书靠在一起的概率.

解　将 10 本书的每一种排列看作基本事件，则基本事件的总数为 10!.

设 A 表示事件“指定的 3 本书靠在一起”，如果将 3 本书看作一本书与剩下的 7 本书进行排列，则有 8!种，而 3 本书靠在一起的排法有 3!种，故 A 中包含的基本事件个数为 $8!\cdot 3!$. 所以

$$P(A)=\frac{8!3!}{10!}=\frac{1}{15}=0.067.$$

例 1.10　一口袋装有 6 只球，其中，4 只白球，2 只红球，从袋中取球两次，每次随机地取 1 只，考虑两种取球方式：(a) 第 1 次取 1 只球，观察其颜色后放回袋中，搅匀后再取 1 球，这种取球方式叫做放回抽样；(b) 第 1 次取 1 球不放回袋中，第 2 次从剩余的球中再取 1 球，这种取球方式叫做不放回抽样. 试分别就上面两种情况求

(1) 取到的 2 只球均为白球的概率；

(2) 取到的 2 只球颜色相同的概率；

(3) 取到的 2 只球中至少有 1 只白球的概率.

解　设 A 为事件“取到的两只球均为白球”，B 为事件“取到的 2 只球均为红球”，C 为事件“取到的 2 只球至少有 1 只白球”，则 $AB=\varnothing$，$C=\overline{B}$.

分析：在袋中依次取 2 只球，每一种取法为一个基本事件，显然样本空间中仅包含有限个元素，由已知，每个基本事件发生的可能性均等，所以可选用古典概型，利用式(1.5) 来计算事件的概率.

(a) 放回抽样的情况

第 1 次从袋中取球 6 只球可供抽取，第 2 次也有 6 只球可供抽取. 由组合法的乘法原理共有 6×6 种取法，即样本空间中基本事件总数为 6×6. 对于事件 A 而言，由于第 1 次有 4 只白球可供抽取，第 2 次也有 4 只白球可供抽取，由乘法原理共有 4×4 种取法，即 A 中包含 4×4 个元素. 同理，B 中包含 2×2 个元素，于是

$$P(A)=\frac{C_4^1\cdot C_4^1}{C_6^1\cdot C_6^1}=\frac{4\times 4}{6\times 6}=\frac{4}{9},$$

$$P(B)=\frac{C_2^1\cdot C_2^1}{C_6^1\cdot C_6^1}=\frac{2\times 2}{6\times 6}=\frac{1}{9},$$

由于 $AB=\varnothing$，得

$$P(A\cup B)=P(A)+P(B)=\frac{5}{9},$$

$$P(C)=P(\overline{B})=1-P(B)=\frac{8}{9}.$$

(b) 不放回抽样的情况

此时，样本空间中基本事件总数为 6×5，事件 A 所包含的基本事件数为 4×3，事件 B 所包含的基本事件数为 2×1，以下可由读者自己完成.

例 1.11(公平的抽签) 盒中有彩券 $m+n$ 张，其中 m 张有奖，n 张无奖. 现随机地一张一张取出，试求第 k 张为有奖彩券的概率($1\leqslant k\leqslant m+n$).

解 从 $m+n$ 张彩券中不放回地一张一张地任取 k 张，有 P_{m+n}^k 种取法. 设 A 为事件"取出的 k 张彩券中，第 k 张是有奖彩券"，完成事件 A 分两个步骤：① 从 m 张有奖彩券中任取一张排在第 k 个位置上，有 m 种取法；② 从剩下的 $m+n-1$ 张中任取 $k-1$ 张，放在前 $k-1$ 个位置上，有 P_{m+n-1}^{k-1} 种. 因而所求的概率为

$$P(A)=\frac{P_m^1P_{m+n-1}^{k-1}}{P_{m+n}^k}=\frac{m}{m+n}.$$

这个结果与 k 值无关，这表明是先抽还是后抽，每个人中签的概率都是一样的，均为 $\frac{m}{m+n}$，与抽签的次序无关. 这从理论上说明了平常人们采用的"抓阄儿"的办法是公平合理的.

1.3.2 几何概型

当联系于某一随机现象的样本空间，可用欧氏空间的某一区域 S 表示，其样本点具有类似于古典概型中的等可能性，而将古典概型中的有限性推广到无限性，就得到几何概型. 一般地说，具有下列特点的概率问题称为几何概型：

(1) 每次试验的结果(即基本事件)具有无限多个，且全体结果可用一个有度量的几何区域来表示；

(2) 在每次试验中，各个基本事件发生的可能性都相同.

这里所说几何区域的度量，是指直线段或曲线段的长度，曲面片的面积，空间立体的体积.

对于几何概型中事件的概率可由下述定义所确定.

定义 1.5 设样本空间 S 所对应的区域仍记为 S，事件 A 所对应的区域仍记为 A，$A\subset S$，则定义事件 A 的概率 $P(A)$ 为

$$P(A)=\frac{A\text{ 的度量}}{S\text{ 的度量}} \tag{1.6}$$

由式(1.6)所确定的概率称为几何概率.

例 1.12 设电台每到整点均报时，一人早上醒来打开收音机，求他等待时间不超过 10 分钟就能听到电台报时的概率.

解 显然样本空间 $S=[0,60]$(单位：分钟).

设 A 表示"等待时间不超过 10 分钟"，则 $A=[50,60]$.

从而

$$P(A)=\frac{A\text{ 的长度}}{S\text{ 的长度}}=\frac{10}{60}=\frac{1}{6}.$$

例 1.13(约会问题)　甲、乙二人约定于 9 时到 10 时之间在某地会面,先到的等候 20 分钟,过时就离去.假定每个人在指定的 1 小时内任一时刻到达是等可能的,求这两人能会面的概率.

解　设 x,y 分别表示甲、乙两人约定地点的时刻,从 9 时算起,单位取分钟,则 $\begin{cases}0\leqslant x\leqslant 60,\\0\leqslant y\leqslant 60.\end{cases}$ $A=\{$两人能够会面$\}$ 的充要条件是

$$|x-y|\leqslant 20.$$

在平面上建立坐标系(图 1－7),则 (x,y) 的所有可能结果对应边长为 60 的正方形 G,而可能会面时间对应图 1－7 中阴影部分 g. 这是一个几何概率问题,由式(1.6)有

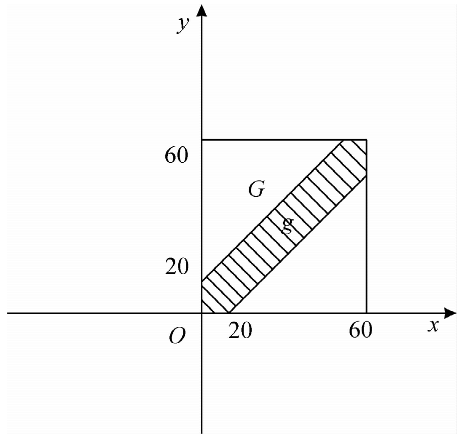

图 1－7

$$P(A)=\frac{g\text{ 的面积}}{G\text{ 的面积}}=\frac{60^2-40^2}{60^2}=\frac{5}{9}.$$

§1.4　条件概率

在自然界和人类社会活动中,有许多事物往往是相互联系、互相影响的. 因此,在实际问题中,除了要考虑事件 A 的概率 $P(A)$ 之外,还要考虑事件 A 在"某事件 B 已经发生"这一附加条件下的概率. 这样的概率,人们称之为条件概率,记作 $P(A\mid B)$.

例如　将一枚硬币抛掷两次,观察其出现正反两面的情况,设事件 B 为"至少有一次为正面",事件 A 为"两次掷出同一面". 现在来求已知事件 B 发生的条件下事件 A 发生的概率.

分析　设 H 为正面,T 为反面

$S=\{HH,HT,TH,TT\}$. $B=\{HH,HT,TH\}$, $A=\{HH,TT\}$, $P(A)=\frac{2}{4}=\frac{1}{2}$. 事件 B 已知发生的条件下事件 A 发生的概率 $P(A\mid B)=\frac{1}{3}=\frac{1/4}{3/4}=\frac{P(AB)}{P(B)}\neq P(A)$.

定义 1.6　在事件 B 已发生的条件下,事件 A 发生的概率,称为在事件 B 发生的条件下,事件 A 发生的条件概率,记为 $P(A\mid B)$.

若 $P(B)>0$,规定 $P(A\mid B)=\dfrac{P(AB)}{P(B)}$;　(1.7)

若 $P(A)>0$,规定 $P(B\mid A)=\dfrac{P(AB)}{P(A)}$.　(1.8)

由定义 1.6 不难验证下列定理成立.

定理 1.1　设 A 为任一给定的事件且 $P(A)>0$,则条件概率 $P(B\mid A)$ 符合概率定义三条件,在缩减的样本空间,即事件 A 发生的样本点组成的集合中,

(1) 非负性:对于任一事件 B, $0\leqslant P(B\mid A)\leqslant 1$;

(2) 规范性：$P(S \mid A) = 1$；

(3) 可列可加性：设 $B_1, B_2, \cdots, B_n, \cdots$ 是两两互斥的事件，则

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

证明 (1) $0 \leqslant P(B \mid A) = \dfrac{P(AB)}{P(A)} \leqslant 1$；

(2) $P(S \mid A) = \dfrac{P(SA)}{P(A)} = \dfrac{P(A)}{P(A)} = 1$；

(3) $$P(\bigcup_{i=1}^{\infty} B_i \mid A) = \frac{P((\bigcup_{i=1}^{\infty} B_i)A)}{P(A)} = \frac{P(\bigcup_{i=1}^{\infty} B_iA)}{P(A)}$$

$$= \frac{\sum_{i=1}^{\infty} P(B_iA)}{P(A)} = \sum_{i=1}^{\infty} \frac{P(B_iA)}{P(A)} = \sum_{i=1}^{\infty} P(B_i \mid A)$$

既然条件概率符合上述三个条件，故概率的性质也适用于条件概率. 例如，对任意事件 A，有 $P(A \mid \overline{B}) = 1 - P(\overline{A} \mid \overline{B})$.

例 1.14 有外观相同的三极管 6 只，按电流放大系数分类，4 只属甲类，2 只属乙类. 不放回地抽取三极管两次，每次只抽一只. 求在第一次抽到是甲类三极管的条件下，第二次又抽到甲类三极管的概率.

解 设 A_i 表示事件"第 i 次抽到的是甲类三极管"，$i = 1, 2$，则

A_1A_2 表示"两次抽到的都是甲类三极管". 依题意可知 $P(A_1A_2) = \dfrac{4 \times 3}{6 \times 5} = \dfrac{12}{30} = \dfrac{2}{5}$，$P(A_1) = \dfrac{4}{6} = \dfrac{2}{3}$，可得

$$P(A_2 \mid A_1) = \frac{P(A_1A_2)}{P(A_1)} = \frac{2/5}{2/3} = \frac{3}{5}.$$

1.42 乘法公式

若 $P(B) > 0$，由条件概率公式定义，可得

$$P(AB) = P(B)P(A \mid B). \tag{1.9}$$

若 $P(A) > 0$，则 $P(AB) = P(A)P(B \mid A)$. (1.10)

通常称(1.9)式和(1.10)式为概率的乘法公式.

乘法公式也可以推广到多个事件的情况.

例如，设 A, B, C 为 3 个事件且 $P(AB) > 0$，则有

$$P(ABC) = P(C \mid AB)P(B \mid A)P(A) \tag{1.11}$$

一般地，设 $A_1, A_2, \cdots, A_n$ 为 n 个事件，$n \geqslant 2$ 且 $P(A_1A_2\cdots A_{n-1}) > 0$，则有

$$P(A_1A_2\cdots A_n) = P(A_n \mid A_1A_2\cdots A_{n-1}) \times P(A_{n-1} \mid A_1\cdots A_{n-2}) \times \cdots \times P(A_2 \mid A_1)P(A_1) \tag{1.12}$$

例 1.15 已知 $P(A) = \dfrac{1}{4}$，$P(B \mid A) = \dfrac{1}{3}$，$P(A \mid B) = \dfrac{1}{2}$，求 $P(A \cup B)$.

解 由于 $P(A \cup B) = P(A) + P(B) - P(AB)$，而

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

即
$$P(AB)=\frac{1}{4}\times\frac{1}{3}=P(B)\times\frac{1}{2}.$$
于是
$$P(AB)=\frac{1}{12},P(B)=\frac{1}{6}.$$
故
$$P(A\cup B)=\frac{1}{4}+\frac{1}{6}-\frac{1}{12}=\frac{1}{3}.$$

例 1.16　设袋中装有 r 只红球，t 只白球. 每次自袋中任取一只球，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 $A_i(i=1,2,3,4)$ 为事件"第 i 次取到红球"，则 $\overline{A}_3,\overline{A}_4$ 为事件第三、四次取到白球因此所求概率为
$$\begin{aligned}P(A_1A_2\overline{A}_3\overline{A}_4)&=P(\overline{A}_4\mid A_1A_2\overline{A}_3)P(\overline{A}_3\mid A_1A_2)P(A_2\mid A_1)P(A_1)\\&=\frac{t+a}{r+t+3a}\cdot\frac{t}{r+t+2a}\cdot\frac{r+a}{r+t+a}\cdot\frac{r}{r+t}.\end{aligned}$$

例 1.17　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$\frac{7}{10}$；若前两次落下未打破，第三次落下打破的概率为$\frac{9}{10}$. 试求透镜落下三次而未打破的概率.

解　设 $A_i(i=1,2,3)$ 表示事件"透镜第 i 次落下打破"，B 表示事件"透镜落下三次而未打破".

因为 $B=\overline{A}_1\overline{A}_2\overline{A}_3$，

所以
$$\begin{aligned}P(B)&=P(\overline{A}_1\overline{A}_2\overline{A}_3)=P(\overline{A}_3\mid\overline{A}_1\overline{A}_2)P(\overline{A}_2\mid\overline{A}_1)P(\overline{A}_1)\\&=(1-\frac{9}{10})(1-\frac{7}{10})(1-\frac{1}{2})=\frac{3}{200}.\end{aligned}$$

1.4.3　全概率与贝叶斯公式

在解决较复杂的概率问题时，人们希望能把所涉及的复杂事件分解为简单事件之和. 下面建立两个用来计算概率的重要公式. 先介绍样本空间的划分的定义.

定义 1.7　设 S 为试验 E 的样本空间，$B_1,B_2,\cdots,B_n$ 为 E 的一组事件. 若

(1)$B_1,B_2,\cdots,B_n$ 两两互斥，即 $B_iB_j=\varnothing(i\neq j,i,j=1,2,\cdots,n)$；

(2)$B_1\cup B_2\cup\cdots\cup B_n=S$，即这 n 个事件的和为必然事件，则称 $B_1,B_2,\cdots,B_n$ 为样本空间 S 的一个划分，或称它为一个完备事件组.

若 $B_1,B_2,\cdots,B_n$ 是样本空间的一个完备事件组，那么，对每次试验，事件 $B_1,B_2,\cdots,B_n$ 中必有一个且仅有一个发生.

例如，A 与$\overline{A}$ 就构成一个完备事件组. 又如试验 E"掷一颗骰子观察其点数"，它的样本空间 $S=\{1,2,3,4,5,6\}$. E 的一组事件 $B_1=\{1,2\}$，$B_2=\{3,4,5\}$，$B_3=\{6\}$ 是 S 的一个划分，而事件组 $C_1=\{1,2,3\}$，$C_2=\{3,4\}$，$C_3=\{5,6\}$ 不是 S 的划分.

定理 1.2(全概率公式)　如果事件组 $B_1,B_2,\cdots,B_n$ 为样本空间 S 的一个划分，$P(B_i)>0(i=1,2,\cdots,n)$，则对任一事件 A，有
$$P(A)=\sum_{i=1}^{n}P(B_i)P(A\mid B_i)\tag{1.13}$$

证明　因为 $A=AS=A(B_1\cup B_2\cup\cdots\cup B_n)=AB_1\cup AB_2\cup\cdots\cup AB_n$，又 $B_1,B_2,\cdots,B_n$ 两两互斥，所以 $AB_1,AB_2,\cdots,AB_n$ 两两互斥. 由加法公式，得

$$P(A)=P(AB_1)+P(AB_2)+\cdots+P(AB_n).$$

又由乘法公式，得

$$P(AB_i)=P(B_i)P(A\mid B_i),i=1,2,\cdots,n,P(B_i)>0,$$

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

在很多实际问题中 $P(A)$ 不容易直接求得，但却容易找到 S 的一个完备事件组 $B_1,B_2,\cdots,B_n$，且 $P(B_i)$ 和 $P(A\mid B_i)$ 或为已知，或容易求得，那么就可以根据(1.13)式求出 $P(A)$. 应用公式时，关键是要找到合适的完备事件组，选取的完备事件组 $B_1,B_2\cdots,B_n$ 应使 $B(B_i)$ 及 $P(A\mid B_i)$ 容易求得.

另一个重要公式是下述的贝叶斯公式.

定理 1.3　设 $B_1,B_2,\cdots,B_n$ 为 S 的一个划分且 $P(B_i)>0$，则对任一事件 $A(P(A)>0)$，有

$$P(B_i\mid A)=\frac{P(A\mid B_i)P(B_i)}{\sum\limits_{j=1}^{n}P(A\mid B_j)P(B_j)},i=1,2,\cdots,n \tag{1.14}$$

证　由全概率公式及条件概率定义得

$$P(B_i\mid A)=\frac{P(B_iA)}{P(A)}=\frac{P(A\mid B_i)P(B_i)}{\sum\limits_{j=1}^{n}P(A\mid B_j)P(B_j)},i=1,2,\cdots,n.$$

(1.14)式称为贝叶斯公式，它是概率论中的一个著名公式，1763年由英国哲学家贝叶斯首先提出的. 从形式上看，贝叶斯公式把一个简单的条件概率 $P(B_i\mid A)$ 表示成了很复杂的形式，但在许多问题中，公式(1.14)右端的量 $P(B_j)$ 和 $P(A\mid B_j)$ 或为已知，或容易求得. 因此，这个公式提供了计算事件条件概率的一个有效途径.

例 1.18　有一批同型号的产品，已知其中由一厂生产的占30%，二厂生产的占50%，三厂生产的占20%，又知这三个厂的产品次品率分别为2%，1%，1%. 问从这批产品中任取一件是次品的概率是多少？

解　设事件 A 为“任取一件为次品”，事件 B_i 为“任取一件为 i 厂的产品”，$i=1,2,3$.

$B_1\cup B_2\cup B_3=S,B_iB_j=\varnothing,i,j=1,2,3.$

由全概率公式得

$$P(A)=P(A\mid B_1)P(B_1)+P(A\mid B_2)P(B_2)+P(A\mid B_3)P(B_3).$$

依题意知　$P(B_1)=0.3,P(B_2)=0.5,P(B_3)=0.2,$

$P(A\mid B_1)=0.02,P(A\mid B_2)=0.01,P(A\mid B_3)=0.01,$

故
$$\begin{aligned}P(A)&=P(A\mid B_2)P(B_1)+P(A\mid B_2)P(B_2)+P(A\mid B_3)P(B_3)\\&=0.02\times0.3+0.01\times0.5+0.01\times0.2=0.013.\end{aligned}$$

例 1.19　对以往数据分析结果表明，当机器调整得良好时，产品的合格率为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%，试求已知某日早上第一件产品是合格品时，机器调整得良好的概率是多少？

解　设 A 为事件“产品合格”，B 为事件“机器调整良好”. 已知 $P(A\mid B)=0.98,P(A\mid\overline{B})=$

$0.55, P(B)=0.95, P(\overline{B})=0.05$,所需求的概率为 $P(B\mid A)$. 由贝叶斯公式得

$$P(B\mid A)=\frac{P(A\mid B)P(B)}{P(A\mid B)P(B)+P(A\mid \overline{B})P(\overline{B})}$$

$$=\frac{0.98\times 0.95}{0.98\times 0.95+0.55\times 0.05}=0.97$$

这就是说,当生产出第一件产品是合格品时,此时机器调整良好的概率为 0.97. 这里,概率 0.95,是由以往的数据分析得到的,叫做先验概率. 而在得到信息(即生产出的第一件产品是合作品)之后再重新加以修正的概率(0.97)叫做后验概率. 有了后验概率我们就能对机器的情况有进一步的了解.

例 1.20　设甲袋中装有 2 个白球,1 个黑球;乙袋中装有 1 个白球,2 个黑球;现从甲袋中任取 1 个球放入乙袋中,再从乙袋中任取一个球:(1) 求从乙袋中取出的球为白球的概率;(2) 已知从乙袋中取出的球为白球,求从甲袋中取到白球的概率.

解　设 B_1, B_2 分别表示"从甲袋中取出的球为白球,黑球放入乙袋"的事件,A 表示"从乙袋中取出的球为白球"的事件,按题意

$$P(B_1)=\frac{2}{3}, P(B_2)=\frac{1}{3}, P(A\mid B_1)=\frac{1}{2}, P(A\mid B_2)=\frac{1}{4},$$

B_1, B_2 为一完备事件组.

(1) 由全概率公式

$$P(A)=P(B_1)P(A\mid B_1)+P(B_2)P(A\mid B_2)=\frac{2}{3}\cdot\frac{1}{2}+\frac{1}{3}\cdot\frac{1}{4}=\frac{5}{12};$$

(2) 由贝叶斯公式

$$P(B_1\mid A)=\frac{P(B_1)P(A\mid B_1)}{P(B_1)P(A\mid B_1)+P(B_2)P(A\mid B_2)}=\frac{\frac{2}{3}\times\frac{1}{2}}{\frac{5}{12}}=\frac{4}{5}.$$

§1.5　事件的独立性

设 A 和 B 是两个事件,若 $P(B)>0$,则可定义条件概率 $P(A\mid B)$,它表示在事件 B 发生的条件下,事件 A 发生的概率;而 $P(A)$ 表示不管事件 B 是否发生,事件 A 发生的概率. 若 $P(A\mid B)=P(A)$,则表明事件 B 的发生并不影响事件 A 发生的概率,这时称事件 A 与 B 相互独立,并且乘法公式变为

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

我们可以用这个公式来刻画事件的独立性.

定义 1.8　设 A, B 是两个事件,若

$$P(AB)=P(A)P(B) \tag{1.15}$$

成立,则称事件 A 与事件 B 是相互独立的.

在实际应用中,两个事件是否相互独立,我们常常不是根据式(1.15)来判断的,而是根据这两个事件的发生是否相互影响来判断的. 例如,甲、乙两人向同一目标射击,彼此互不相干,则甲、乙各自是否击中目标这类事件是相互独立的;又如,对某一物体进行多次测量时,不同次测

量的测量误差都可认为是相互独立的.

容易知道,若 $P(A)>0, P(B)>0$,则 A,B 相互独立与 A,B 互不相容不能同时成立.

定理 1.4 设 A,B 是两事件,且 $P(A)>0$.若 A,B 相互独立,则 $P(B \mid A)=P(B)$.反之亦然.

定理的正确性是显然的.

定理 1.5 若 A,B 相互独立,则下列各对事件,$\overline{A}$ 与 B,A 与 $\overline{B}$,$\overline{A}$ 与 $\overline{B}$ 也相互独立.

证明 先证 A 与 $\overline{B}$ 相互独立.因为

$$A=AB \cup A\overline{B} \text{ 且}(AB)(A\overline{B})=\varnothing,$$

所以

$$P(A)=P(AB)+P(A\overline{B}),$$

即

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

又因为 A、B 相互独立,所以有

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

因而

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

从而 A 与 $\overline{B}$ 相互独立.由此可立即推出 $\overline{A}$ 与 $\overline{B}$ 相互独立.再由 $\overline{\overline{B}}=B$,又推出 $\overline{A}$ 与 B 相互独立.

下面我们将独立性的概念推广到三个事件的情况.

定义 1.9 设 A,B,C 是三个事件,如果有

$$\left.\begin{aligned} P(AB) &= P(A)P(B), \\ P(BC) &= P(B)P(C), \\ P(AC) &= P(A)P(C), \end{aligned}\right\} \tag{1.16}$$

则称事件 A,B,C 两两相互独立.

若不仅式(1.16)成立,而且

$$P(ABC)=P(A)P(B)P(C) \tag{1.17}$$

也成立,则称 A,B,C 是相互独立的.

由定义可知,若三个事件相互独立,则它们一定是两两相互独立;但两两相互独立不一定是相互独立.

一般地,n 个事件相互独立的定义如下:

定义 1.10 设有 n 个事件 $A_1,A_2,\cdots,A_n$.若对于其中任意两个事件 $A_i,A_j(i\neq j, i,j=1,2,\cdots,n)$ 都有 $P(A_iA_j)=P(A_i)P(A_j)$,则称事件 $A_1,A_2,\cdots,A_n$ 两两独立;若对于其中任意 k 个事件($k=2,3,\cdots,n$),都有 $P(A_{i1}A_{i2}\cdots A_{ik})=P(A_{i1})\cdot P(A_{i2})\cdots P(A_{ik})$,则称事件 $A_1,A_2,\cdots,A_n$ 相互独立.

从定义可知,若 n 个事件相互独立,则它们中任何 $k(2\leqslant k\leqslant n)$ 个事件也是相互独立的.

不难证明,若 $A_1,A_2,\cdots,A_n$ 相互独立,将其中任意多个事件换成它们的对立事件而得到的 n 个事件仍相互独立.

例 1.21 一个均匀正四面体,它的 3 个面分别涂上了红色、黄色、蓝色,第 4 个面同时涂上了红、黄、蓝 3 色,随机抛掷这个正四面体,观察与桌面接触的那个面所涂的颜色.假设以 A,B,C 分别表示接触面涂有红、黄、蓝色,试求 $P(A),P(B),P(C),P(AB),P(AC),P(BC)$ 和 $P(ABC)$,并问 A,B,C 是否两两独立?是否相互独立?

解　易知

$P(A)=P(B)=P(C)\neq\dfrac{1}{2}$,

$P(AB)=P(BC)=P(AC)=\dfrac{1}{4}$,

$P(ABC)=\dfrac{1}{4}$.

由于　$P(AB)=P(A)P(B)$,$P(AC)=P(A)P(C)$,$P(BC)=P(B)P(C)$,所以 A,B,C 两两独立,但

$$P(ABC)\neq P(A)P(B)P(C)$$

故 A,B,C 不相互独立.

例 1.22　设每一名机枪射击手击落飞机的概率都是 0.2,若 10 名机枪射击手同时向一架飞机射击,问击落飞机的概率是多少?

解　设事件 A_i 为"第 i 个射手击落飞机",$i=1,2,\cdots,10$,事件 B 为"击落飞机",则

$$\begin{aligned}B&=A_1\cup A_2\cup\cdots\cup A_{10},\\P(B)&=P(A_1\cup A_2\cup\cdots\cup A_{10})\\&=1-P(\overline{A_1\cup A_1\cup\cdots\cup A_{10}})\\&=1-P(\overline{A}_1\,\overline{A}_2\cdots\overline{A}_{10})\\&=1-P(\overline{A}_1)P(\overline{A}_2)\cdots P(\overline{A}_{10})=1-(0.8)^{10}=0.893\end{aligned}$$

例 1.23　同时抛掷一对骰子,共抛两次,求两次所得点数分别为 7 与 11 的概率.

解　设事件 A_i 为"第 i 次得 7 点",$i=1,2$. 设事件 B_i 为"i 次得 11 点",$i=1,2$. 事件 A 为两次所得点数分别为 7 与 11,则有

$$\begin{aligned}P(A)&=P(A_1B_2\cup B_1A_2)=P(A_1B_2)+P(B_1A_2)\\&=P(A_1)P(B_2)+P(B_1)P(A_2)\\&=\frac{6}{36}\times\frac{2}{36}+\frac{2}{36}\times\frac{6}{36}=\frac{1}{54}.\end{aligned}$$

例 1.24　在可靠性理论中,每个元件(或系统)能正常工作的概率,称为元件(或系统)的可靠性. 如图所示,设有 4 个独立工作的元件 1,2,3,4 按先串联再并联的方式联结(称为串并联系统),设第 i 个元件的可靠性为 $p_i(i=1,2,3,4)$,试求系统的可靠性.

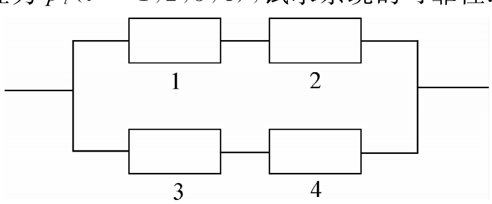

图 1－8

解　设 $A_i(i=1,2,3,4)$ 表示事件第 i 个元件正常工作,以 A 表示系统正常工作.

则有　$$A=A_1A_2\cup A_3A_4$$

由事件的独立性,得系统的可靠性:

$$\begin{aligned}P(A)&=P(A_1A_2)+P(A_3A_4)-P(A_1A_2A_3A_4)\\&=P(A_1)P(A_2)+P(A_3)P(A_4)-P(A_1)P(A_2)P(A_3)P(A_4)\end{aligned}$$

$= p_1p_2 + p_3p_4 - p_1p_2p_3p_4$.

例 1.25 设 $0 < P(A) < 1, 0 < P(B) < 1, P(A \mid B) + P(\overline{A} \mid \overline{B}) = 1$,试证事件 A 与 B 相互独立.

证明 因为 $0 < P(A) < 1, 0 < P(B) < 1$,所以

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

$$P(\overline{A} \mid \overline{B}) = \frac{P(\overline{A}\,\overline{B})}{P(\overline{B})} = \frac{1 - P(A + B)}{1 - P(B)} = \frac{1 - P(A) - P(B) + P(AB)}{1 - P(B)}$$

又因为
$$P(A \mid B) + P(\overline{A} \mid \overline{B}) = 1$$

所以
$$\frac{1 - P(A) - P(B) + P(AB)}{1 - P(B)} = \frac{P(B) - P(AB)}{P(B)}$$

化简,得:$P(AB) = P(A)P(B)$

所以 事件 A、B 相互独立.

例 1.26 验收 100 件产品的方案如下:从中任取 3 件进行独立地测试,如果至少有一件被断定为次品,则拒绝接收这批产品.设一件次品经测试后被断定为次品的概率为 0.95,一件正品经测试后被断定为正品的概率为 0.99,并已知这 100 件产品中恰有 4 件次品,求此产品能被接收的概率.

解 设 A 表示事件"此批产品被接收",B_i 表示事件"取出的产品中恰有 i 件是次品",$i = 0,1,2,3$,则

$$P(B_0) = \frac{C_{96}^3}{C_{100}^3}, P(B_1) = \frac{C_4^1C_{96}^2}{C_{100}^3},$$

$$P(B_2) = \frac{C_4^2C_{96}^1}{C_{100}^3}, P(B_3) = \frac{C_4^3}{C_{100}^3}.$$

由假设,三次测试是相互独立的,于是

$P(A \mid B_0) = 0.99^3, P(A \mid B_1) = 0.99^2(1 - .095)$,

$P(A \mid B_2) = 0.99(1 - 0.95)^2, P(A \mid B_3) = (1 - 0.95)^3$.

由全概率公式,得

$$P(A) = \sum_{i=0}^{3} P(A \mid B_i)P(B_i) \approx 0.8629$$

所以,此批产品能被接收的概率为 0.8629.

习题一

1.写出下列随机试验的样本空间:

(1) 记录一个小班一次数学考试的平均分数(设以百分制记分).

(2) 生产产品直到有 10 件正品为止,记录生产产品的总件数.

(3) 在单位圆内任意取一点,记录它的坐标.

(4) 一尺之棰折成三段,观察各段的长度.

2.一位工作生产四个零件,以事件 A_i 表示他生产的第i 个零件是次品,$i = 1,2,3,4$,用它们的运算关系表示如下事件:

(1) 全是合格品;

(2) 至少有一件次品;

(3) 仅有一件是次品.

3. 设 A,B,C 为 3 事件,试用事件的运算关系表示下列事件:

(1)A,B,C 都不发生;

(2)A 与 B 都发生,但 C 不发生;

(3)A,B,C 中不多于两个发生;

(4)A,B,C 中 A,B 至少有一个发生,C 不发生.

4. 在某城市中发行三种报纸:甲、乙、丙,用 A、B、C 分别表示"订阅甲报"、"订阅乙报"、"订阅丙报",试求下列各事件.(1) 只订甲报;(2) 只订甲、乙两报;(3) 只订一种报纸;(4) 正好订两种报纸;(5) 至少订一种报纸;(6) 不订任何报纸.

5. 设一批零件,有正品也有次品,从这批零件中任意抽取 7 件,设 A 表示事件"抽到的次品数不多于3",B 表示事件"抽到的次品数为奇数". 试问:事件 $A\cup B$、AB、$A-B$、$\overline{B}$ 各表示什么意思?

6. 设样本空间 $S=\{x\mid 0\leqslant x\leqslant 2\}$,事件 $A=\{x\mid 0.5\leqslant x<1\}$,$B=\{x\mid 0.8<x\leqslant 1.6\}$,具体写出下列各事件:(1)$AB$;(2)$A-B$;(3) $\overline{A-B}$;(4) $\overline{A\cup B}$.

7. 设 A,B 是两事件且 $P(A)=0.6$,$P(B)=0.7$. 问:

(1) 在什么条件下,$P(AB)$ 取到最大值,最大值是多少?

(2) 在什么条件下,$P(AB)$ 取到最小值,最小值是多少?

8. 已知事件 A、B 的概率分别为 $P(A)=0.7$,$P(B)=0.6$,且 $P(AB)=0.4$,求 $P(A\cup B)$,$P(A-B)$,$P(\overline{A}\cup\overline{B})$,$P(\overline{A}\cup B)$,$P(\overline{A}B)$,$P(A\cup\overline{A}B)$.

9. 某城市共发行 A,B,C 三种报纸,调查表明居民家庭中订购 C 报的占 30%,同时订购 A,B 两报的点 10%,同时订购 A,C 及 B,C 两报的各占 8%,5%,三报都订的占 3%,今在该城中任找一户,问该户(1) 只订 A、B 两报;(2) 只订 C 报的概率各为多少?

10. 一个 5 位数字的号码锁,每位上有 0,1,…,9 共计 10 个数码,若不知道该锁号码,问开一次锁就把该锁打开的概率是多大?

11. 设有 N 件产品,其中有 D 件次品,今从中任取 n 件,问其中恰有 $k(k\leqslant D)$ 件次品的概率是多少?

12. 设 A,B,C 是 3 个事件且 $P(A)=P(B)=P(C)=\frac{1}{4}$,$P(AB)=P(BC)=P(AC)=\frac{1}{8}$,求 A,B,C 至少有一个发生的概率.

13. 一批产品共 200 件,共中恰有 6 件废品,求:(1) 这批产品的废品率;(2) 任取 3 件,至多有 1 件废品的概率.

14. 在房间里有 10 个人,分别佩戴从 1 号到 10 号纪念章,任选 3 人记录其纪念章的号码.(1) 求最小号码为 5 的概率.(2) 求最大号码为 5 的概率.

15. 一袋中有 7 个白球和 5 个红球,从中摸取二次,每次一球. 设 A 表示"两次都取到红球",B 表示"至少一次取到红球" 请在(1) 有放回抽样;(2) 不放回抽样条件下求 $P(A)$,$P(B)$.

16. 10 名学生入围知识竞赛决赛,共有 20 道竞赛题,其中数学、英语、历史学科分别为 4,6,10 道题,每人随机抽取 2 道题,求:

(1)3 号学生抽到英语题的概率;

(2)3 号学生抽到 2 道英语题,5 号学生抽到 1 道英语题与 1 道历史题的概率.

17.从 5 双不同的鞋子中任取 4 只,问这 4 只鞋子中至少有两只配成一双的概率是多少?

18.在 $1 \sim 2000$ 的整数中随机地取一个数,问取到的整数既不能被 6 整除,又不能被 8 整除的概率是多少?

19.甲、乙两艘油轮驶向一个不能同时停泊两艘油轮的码头,它们都将在某日8时至20时抵达码头,甲轮卸完油要 1 小时,乙轮要 2 小时,假设每艘油轮在8时至20时的每一时刻抵达码头的可能性相同.

(1) 求甲、乙两轮都不需要等候空出码头的概率;

(2) 设 A 表示甲、乙同一时刻抵达码头,问 A 是否为不可能事件,并求 $P(A)$.

20.将一根长为 l 的棍子任意地折成 3 段,求此 3 段能构成一个三角形的概率.

21.由长期统计资料得知,某一地区在 4 月份下雨(记作事件 A) 的概率为$\dfrac{4}{15}$,刮风(用 B 表示) 的概率为$\dfrac{7}{15}$,既刮风又下雨的概率为$\dfrac{1}{10}$,求 $P(A \mid B)$、$P(B \mid A)$、$P(A \cup B)$.

22.某种动物由出生算起活到20岁以上的概率为 0.8,活到25岁以上的概率为 0.4,如果现在有一个 20 岁的这种动物,问它能活到 25 岁以上的概率是多少?

23.包装了的玻璃器皿第一次扔下被打破的概率为 0.4,若未破,第二次扔下被打破的概率为 0.6,若又未破,第三次扔下被打破的概率为 0.9,今将这种包装了的器皿连续扔 3 次,求打破的概率.

24.甲、乙两选手进行乒乓球单打比赛,甲选手发球成功后,乙选手回球失误的概率为 0.3,若乙选手回球成功,甲选手回球失误的概率为 0.4,若甲选手回球成功,乙选手再次回球失误的概率为 0.5,试计算这几个回合中,乙选手输掉 1 分的概率.

25.在对空演习中,某高射炮的目标是正在飞行中的一架战斗机,已知该炮能击中发动机、机舱及其他部位的概率分别为 0.10,0.08,0.39.又若击中上述各部位而使飞机坠毁的概率分别是 0.95,0.89,0.51.试求该炮任意发射一炮弹使飞机坠毁的概率.

26.5 张卡片上分别标有数字 1,2,3,4,5,每次从中任取一张,连取两次.

(1) 若第1次取出的卡片不放回,求第2次取出的卡片上的数字大于第1次取出的卡片上的数字的概率;

(2) 若第1次取出的卡片放回,求第2次取出的卡片上的数字大于第1次取出的卡片上的数字的概率.

27.设盒中装有 5 只晶体管,其中 3 只是正品,2 只是次品,现从盒中随机地摸出 2 只并换进 2 只正品,之后,再从盒中摸出 2 只,求第二次摸出的 2 只都是正品的概率.

28.某试卷中一道选择题有 6 个答案,其中,只有一个正确,考生不知道正确答案的概率为$\dfrac{1}{4}$,不知道正确答案的考生可以猜,设猜对的概率为$\dfrac{1}{6}$,现已知某考生答对了,问他猜对此题的概率有多大?

29.将两个信息分别编码为 A 和 B 传递出去,接收站收到时,A 被误收作 B 的概率为 0.02,而 B 被误收作 A 的概率为 0.01,信息 A 与信息 B 传递的频繁程度为 $2:1$.若接收站收到的信息是 A,问原发信息是 A 的概率是多少?

30.设有来自三个地区的各 10 名、15 名和 25 名考生的报名表,其中女生的报名表分别为 3 份、7 份和 5 份.随机取一个地区的报名表,从中先后抽取两份:

(1) 求先抽取的一份是女生表的概率；

(2) 已知后抽到的一份是男生表，求先抽到的一份是女生表的概率.

31. 证明：若 $P(A)=1$，则 A 与任意事件 B 相互独立.

32. 加工一个产品要经过三道工序，第一、二、三道工序不出废品的概率分别为 0.9，0.95，0.8，若假定各工序是否出废品为独立的，求经过三道工序而不出废品的概率.

33. 三个人独立地去破译一个密码，他们能破译的概率分别为 $\frac{1}{5}$，$\frac{1}{3}$，$\frac{1}{4}$，若让他们共同破译的概率是多少？

34. 设 $0<P(A)<1$，$0<P(B)<1$，$P(A\mid B)+P(\overline{A}\mid\overline{B})=1$，试证事件 A 与 B 相互独立.

35. 玻璃杯成箱出售，每箱 20 只，各箱含有 0，1，2 只次品的概率分别是 0.7，0.2 和 0.1，顾客购买时，任取一箱打开随机查看 4 只，若无次品，就买下该箱玻璃杯，否则退回，试求：

(1) 顾客买下该箱的概率；

(2) 顾客买下的一箱中，仍有次品的概率.

第二章　随机变量及其分布

在实际问题中,随机试验的结果可以用数量来表示,为了方便地研究随机试验的各种结果及各种结果发生的概率,我们引入了随机变量的概念.本章通过引入随机变量及其分布的概念,借助微积分这一工具来全面地研究随机试验的结果,揭示随机现象的统计规律性.

§2.1　随机变量

有些随机试验的结果本身与数值有关(本身就是一个数).例如,掷一颗骰子面上出现的点数,记录电话的呼唤次数,抽样检查产品时废品的个数等,所得的可能结果都是数量,而在有些随机试验中,随机试验的结果看来与数值无关,但我们可以引进一个变量来表示它的各种结果.也就是说,把试验结果数值化.正如裁判员在运动场上不叫运动员的名字而叫号码一样,二者建立了一种对应关系.

例 2.1　考察抛硬币试验,结果是:“出现正面”或“出现反面”.虽然其结果并不表示为数量,但可以把试验结果数量化,如令

$$X = X(e) = \begin{cases} 0, e = T, \\ 1, e = H, \end{cases} \tag{2.1}$$

则这个有两个可能值的变量 X 代表了抛一枚硬币这一试验的结果.

上面这个例子中的试验结果能用一个数 X 来表示,这个数 X 随着试验结果不同而变化,因而它是样本点的函数,即 X 是定义在样本空间上的函数,这种函数就是我们本章研究的随机变量.下面我们给出随机变量的定义.

定义 2.1　设随机试验 E 的样本空间为 S,如果对于每一个 $e \in S$,有一个实数 $X(e)$ 与之对应,则将单值实值函数 $X = X(e)$ 叫做样本空间 S 上的随机变量.通常用大写字母 $X,Y,Z,\cdots$ 表示随机变量,用小写字母 $x,y,z,\cdots$ 表示它们可能取的值.

引入随机变量以后,随机事件就可以用随机变量来描述了.例如,设 X 表示电话变换台在一段时间内接到的呼叫次数,则“$0 \leqslant X \leqslant 3$”表示“呼叫次数不超过三次”的事件;“$X > 5$”表示“呼叫次数大于5”的事件.若 X 是式(2.1)所规定的随机变量,则“$X = 1$”表示“正面”,而“$X = 0$”表示“反面”.于是,对事件的研究就可转化为对随机变量的研究了.

随机变量通常分为两类:如果随机变量 X 的所有取值可以逐个列举出来(即有限个或可列无限个).则 X 称为离散型随机变量;如果随机变量 X 的所有可能取值不可以逐个列举出来,则 X 称为非离散型随机变最.非离散型随机变量范围很广,其中最重要,实际工作也常遇到的是连续型随机变量.

§2.2　离散型随机变量

2.2.1　离散型随机变量的分布律

我们对随机试验的研究，不仅关心试验会出现什么样的结果，而且更重要的是要研究这些结果将以多大的概率出现，也就是对随机变量而言，我们不但要知道它取什么样的数值，而且要知道它取这些值的概率.

对离散型随机变量，描述它的概率特征的，主要是分布律.

定义 2.2　设离散型随机变量 X 所有可能取值为 $x_k(k=1,2,3,\cdots)$

$$P\{X=x_k\}=p_k, k=1,2,3,\cdots, \tag{2.2}$$

则式(2.2)称为离散型随机变量 X 的概率分布或分布律.

X 的分布律常用表格表示：

X	x_1	x_2	$\cdots$	x_n	$\cdots$
p_k	p_1	p_2	$\cdots$	p_n	$\cdots$

有时亦写成

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_n & \cdots \\ p_1 & p_2 & \cdots & p_n & \cdots \end{pmatrix}$$

由概率的定义可知：

(1) $p_k \geqslant 0(k=1,2,\cdots)$；

(2) $\sum\limits_{k=1}^{\infty} p_k = 1$.

例 2.2　设一汽车在开往目的地的道路上需路经四个交叉路口，遇到红灯的概率为 p. 以 X 表示汽车首次停下时，它已通过的信号灯的灯数（设通过四个信号灯以后停下且各组信号灯的工作是相互独立的），求 X 的分布律.

解　由题意，随机变量 X 的所有可能取值为 0，1，2，3，4. $\{X=0\}$ 表示已通过的信号灯数为 0，即第一盏信号灯是红灯，故 $P\{X=0\}=P$.

$\{X=1\}$ 表示已通过的信号灯数为 1（即第一盏信号灯是绿灯，第二盏信号灯为红灯），故 $P\{x=1\}=(1-p)p$，同理可得

$P\{X=2\}=(1-p)^2p, P\{X=3\}=(1-p)^3p, P\{X=4\}=(1-p)^5$.

则 X 的分布律为

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

例 2.3　设随机变量的概率分布为：

$$P\{x=k\}=a\frac{\lambda}{k!}, k=0,1,2,\cdots, \lambda>0$$

试确定常数 a.

解　依据概率分布性质：

$$\begin{cases} P\{X=k\} \geqslant 0, \\ \sum\limits_{k} P\{X=k\} = 1 \end{cases} \text{则有} \begin{cases} a \geqslant 0 \\ \sum\limits_{k=1}^{\infty} a\dfrac{\lambda^k}{k!} = a\mathrm{e}^{\lambda} = 1 \end{cases}$$

从中解得：$a = \mathrm{e}^{-\lambda}$.

2.2.2　常见的离散型分布

下面介绍一些常见的离散型随机变量的分布律.

1. 两点分布

设随机变量 X 只可能取 0 与 1 两个值，它的分布律为

X	0	1
p_k	$1-p$	p

则称 X 服从(0－1)分布或两点分布.

在例 2.1 中，随机变量 X 服从(0－1)分布.

其分布律为

X	0	1
p_k	$\frac{1}{2}$	$\frac{1}{2}$

例 2.4　200 件产品中，有 190 件合格品，10 件不合格品，现在从中随机抽取一件，那么，若规定

$$X = \begin{cases} 1, \text{取得不合格品}, \\ 0, \text{取得合格品}. \end{cases}$$

X	0	1
p_k	$\frac{190}{200}$	$\frac{10}{200}$

则随机变量 X 服从(0－1)分布.

两点分布是最简单的一种分布，任何一个只有两种可能结果的随机现象，比如新生婴儿是男还是女，明天是否下雨，种子是否发芽等，都属于两点分布.

2. 伯努利试验、二项分布

设试验 E 只有两个可能结果 A 及 $\overline{A}$，则称 E 为伯努利试验. 将试验 E 在相同条件下重复进行 n 次，如果各次试验的结果是相互独立的，则称这 n 次重复的独立试验为 n 重伯努利试验，相应的数学模型叫伯努利概型.

在一次试验中，事件 A 发生的概率为 $p(0<p<1)$，用 X 表示 n 次试验中事件 A 发生的次数，则 X 是一个随机变量，它的可能取值是 $0,1,2,\cdots,n$，分布律为

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

记 $q=1-p$，即有 $P\{X=k\} = C_n^k p^k q^{n-k}$，$k=0,1,2,\cdots,n$. 我们称随机变量 X 服从参数为

n,p 的二项分布，记为 $X \sim B(n,p)$.

容易验证，它满足概率分布的两个性质：

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

(2)$\sum\limits_{k=0}^{n} P\{X=k\}=\sum\limits_{k=0}^{n} C_n^k p^k(1-p)^{n-k}=(p+q)^n=1$.

注意到 $C_n^k p^k q^{n-k}$ 恰好是二项式 $(p+q)^n$ 的展开式中的一般项，故称满足 $P\{X=k\}=C_n^k p^k q^{n-k}$ 式的随机变量 X 服从二项分布，当 $n=1$ 时，二项分布就是两点分布：$P\{X=k\}=p^k q^{n-k}(k=0,1)$

例 2.5　设有 N 件产品，其中有 M 件次品. 现进行 n 次有放回的抽样，每次抽取一件. 求这 n 次中共抽到的次品数 X 的概率分布.

解　由于抽样是有放回的，因此这是 n 重伯努利试验，若以 A 表示一次抽样中抽到次品这一事件，则

$$p=P(A)=\frac{M}{N}$$

故 $X \sim B(n,\frac{M}{N})$，即

$$P\{X=k\}=C_n^k\left(\frac{M}{N}\right)^k\left(1-\frac{M}{N}\right)^{n-k}, k=0,1,2,\cdots,n$$

二项分布 $B(n,p)$ 和两点分布 $B(1,p)$ 还有一层密切关系. 仍设一个试验只有两个结果：A 和 $\overline{A}$，且 $P(A)=p$，现将试验独立进行 n 次，记 X 为 n 次试验中结果 A 出现的次数，则 $X \sim B(n,p)$，若记 X_i 为第 i 次试验中结果 A 出现的次数，即

$$X_i=\begin{cases}1,\text{第 } i \text{ 次试验中结果 } A \text{ 出现，}\\ 0,\text{第 } i \text{ 次试验中结果 } A \text{ 不出现.}\end{cases} \quad i=1,2,\cdots,n,$$

则 $X_i \sim B(1,p)$，并且 $X_i,X_2,\cdots,X_n$ 相互独立(相互独立的概念见第三章). 根据 X 和 $X_1,X_2,\cdots,X_n$ 的定义，自然有

$$X=X_1+X_2+\cdots+X_n$$

这个式子表明，一个服从二项分布的随机变量可以表示成 n 个相互独立的，且服从两点分布的随机变量之和，这个很重要的事实在以后的讨论中要多次用到.

例 2.6　按规定，某种型号电子元件的使用寿命超过 1500 小时的为一级品，已知一大批产品的一级品率为 0.2，现在从中随机地抽查 20 只. 问 20 只元件中恰有 k 只($k=0,1,\cdots,20$) 为一级品的概率是多少？

解　这是不放回抽样. 但由于这批元件的总数很大，且抽查的元件的数量相对于元件的总数来说又很小，因而可以当作放回抽样来处理，这样做会有一些误差，但误差不大. 我们将检查 1 只元件看它是否为一级品看成一次试验，检查 20 只元件相当于做 20 重伯努利试验. 设 X 为 20 只元件中一级品的只数，那么，X 是一个随机变量，且有 $X \sim B(20,0.2)$. 故所求概率为

$$P\{X=k\}=C_{20}^k(0.2)^k(0.8)^{20-k}, k=0,1,\cdots,20.$$

将计算的结果列表如下：

X	0	1	2	3	4	5
p_k	0.012	0.058	0.137	0.205	0.218	0.175

X	6	7	8	9	10	$k \geqslant 11$
p_k	0.109	0.055	0.022	0.007	0.002	< 0.001

为了对本题的结果有一个直观了解，我们作出上表的图形，如图

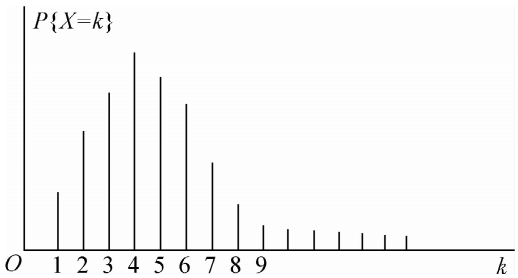

图 2－1

从图 2－1 中看到，当 k 增加时，概率 $P\{X = k\}$ 先是随之增加，直至达到最大值(本例中当 $k = 4$ 时取到最大值)，随后单调减少. 一般地，对于固定的 n 及 p，二项分布 $B(n,p)$ 都具有这一性质.

例 2.7 设每台自动机床在运行过程中需要维修的概率均为 $p = 0.01$，并且各机床是否维修各自独立. 如果

(1) 每名维修工人负责看管 20 台机床；

(2)3 名维修工人共同看管 80 台机床，

求不能及时维修的概率.

解 (1) 这是 $n = 20$ 重伯努利试验，参数 $p = 0.01$，需维修的机床数 $X \sim B(20,0.01)$，故不能及时维修的概率为

$$\begin{aligned} P\{X > 1\} &= 1 - \sum_{k=0}^{1} P\{X = k\} \\ &= 1 - \sum_{k=0}^{1} C_{20}^{k} \cdot 0.01^{k} \cdot 0.99^{20-k} \\ &= 1 - 0.99^{20} - 20 \times 0.01 \times 0.99^{19} \\ &= 0.169 \end{aligned}$$

(2) 这是 $n = 80$ 重伯努利试验，参数 $p = 0.01$，需要维修的机床数 $X \sim B(80,0.01)$，3 名维修工人共同看管 80 台机床时不能及时维修的概率为

$$P\{X > 3\} = 1 - \sum_{k=0}^{3} C_{80}^{k} \cdot 0.01^{k} \cdot 0.99^{80-k} = 0.0087.$$

我们发现，在后一种情况尽管任务重了，但工作效率不仅没有降低，反而提高了.

直接计算上式是很麻烦的，而且 n 愈大有关二项分布的计算量愈大. 法国数学家泊松在研究二发布的近似计算时证明了当 n 很大而 p 很小时，

$$C_n^k p^k (1-p)^{n-k} \approx \frac{\lambda k}{k\,!} e^{-\lambda}, \lambda = np.$$

并提出了泊松分布.

3. 泊松分布

如果随机变量 X 的概率分布为

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

其中,$\lambda > 0$ 是常数,则称 X 服从参数为 λ 的泊松分布,记为 $X \sim \pi(\lambda)$.

显然,$P\{X = k\} \geqslant 0 (k = 0, 1, \cdots)$;

$$\sum_{k=0}^{\infty} P\{X = k\} = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k\,!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k\,!} = e^{-\lambda} e^{\lambda} = 1.$$

在实际问题中,有许多随机变量服从泊松分布.例如,一段时间内电话交换台收到的呼叫次数,某路口通过的车辆数,粮食种子中含的杂草种子数,一年内发生洪水的次数,等等,都可用泊松分布描述.

例 2.8　有一繁忙的汽车站,每天有大量汽车通过,设每辆汽车,在一天的某段时间内出事故的概率为 0.0001,在每天的该段时间内有 1000 辆汽车通过,问出事故的次数不小于 2 的概率是多少?

解　设 1000 辆车通过,出事故的次数为 X,则 $X \sim B(1000, 0.0001)$,由 $np = 1000 \times 0.0001 = 0.1$ 知近似地有 $X \sim \pi(0.1)$,则由泊松分布的定义

$$\begin{aligned} P\{X \geqslant 2\} &= 1 - P\{X = 0\} - P\{X = 1\} \\ &= 1 - \frac{e^{-0.1}}{0!} - \frac{0.1 \cdot e^{0.1}}{1!} = 0.0047. \end{aligned}$$

4. 几何分布

若随机变量 X 的分布律为

X	1	2	$\cdots$	k	$\cdots$
p_k	p	qp	$\cdots$	$q^{k-1}p$	$\cdots$

$, p + q = 1,$

则称 X 服从几何分布.

例 2.9　设某批产品的次品率为 p,对该批产品做有放回的抽样检查,直到第一次抽到一只次品为止(在此之前抽到的全是正品),那么所抽到的产品数 X 是一个随机变量,求 X 的分布律.

解　X 所取的可能值是 $1, 2, 3, \cdots$.设 A_i 表示"抽到的第 i 个产品是正品",

$$\begin{aligned} P\{X = k\} &= P\{A_1 A_2 \cdots A_{k-1} \overline{A}_k\} \\ &= P(A_1) \cdot P(A_2) \cdots P(A_{k-1}) \cdot P(\overline{A}_k) \\ &= \underbrace{(1-p)(1-p)\cdots(1-p)}_{(k-1)} \cdot p = q^{k-1} p. \qquad (k = 1, 2, \cdots) \end{aligned}$$

所以 X 服从几何分布.

几何分布可作为描述某个试验"首次成功"的概率模型.

§2.3　随机变量的分布函数

在处理实际问题中，人们常常关心的是一个随机变量 X 落入某个区间 $(a,b]$ 内的概率. 例如，参军青年关心的是它的身高是否达到标准，而不关心其身高是否刚好等于某个数字. 注意到概率 $P\{a < X \leqslant b\} = P\{X \leqslant b\} - P\{X \leqslant a\}$，为此我们引入分布函数的概念.

定义 2.3　设 X 是一个随机变量，x 是任意实数，函数

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

称为 X 的分布函数：

分布函数 $F(x)$ 主要研究随机变量在某一区间内取值的概率情况，它是 x 的一个普通实函数. 它的定义域是 $(-\infty, +\infty)$，故确定分布函数时，要把 x 放在整个数轴上进行讨论.

由定义，对于任意实数 $x_1, x_2 (x_1 < x_2)$，有

$$\begin{aligned} P\{x_1 < X \leqslant x_2\} &= P\{X \leqslant x_2\} - P\{X \leqslant x_1\} \\ &= F(x_2) - F(x_1) \end{aligned} \tag{2.3}$$

(2.3) 式表明，随机变量 X 落在区间 $(a,b]$ 上的概率可以通过 X 的分布函数来计算，而分布函数是一个普通的函数，正因为这个缘故，才能用微积分的工具来研究随机变量.

例 2.10　设随机变量 X 的分布律为

X	-1	2	3
p_k	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$

求 X 的分布函数.

解　由分布律可知：

当 $x < -1$ 时，$P\{X \leqslant x\} = P(\varnothing) = 0$；

当 $-1 \leqslant x < 2$ 时，$P\{X \leqslant x\} = P\{X = -1\} = \frac{1}{2}$；

当 $2 \leqslant x < 3$ 时，$P\{X \leqslant x\} = P\{X = -1\} + P\{X = 2\} = \frac{5}{6}$；

当 $x \geqslant 3$ 时，$P\{X \leqslant x\} = P(S) = 1$.

于是 X 的分布函数为

$$F(x) = \begin{cases} 0, x < -1 \\ \frac{1}{2}, -1 \leqslant x < 2 \\ \frac{5}{6}, 2 \leqslant x < 3, \\ 1, x \geqslant 3 \end{cases}$$

$F(x)$ 的图形如图 2－2 所示，它是一条阶梯形曲线，在 $x = -1, 2, 3$ 处，分别有跳跃值 $\frac{1}{2}, \frac{5}{6}$, 1.

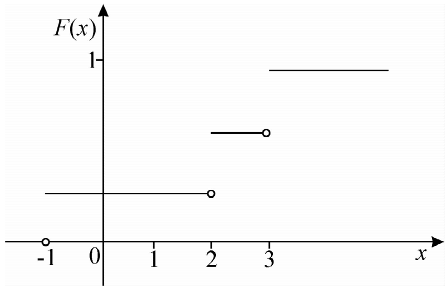

图 2－2

一般地,设离散型随机变量 X 的分布律为

$$P\{X=x_k\}=p_k,k=1,2,\cdots,$$

则其分布函数为

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

其中,和式是对所有满足 $x_k\leqslant x$ 的指标进行的.

分布函数具有如下一些性质:

(1)$0\leqslant F(x)\leqslant 1$ 且 $\lim\limits_{x\to-\infty}F(x)=F(-\infty)=0$, $\lim\limits_{x\to+\infty}F(x)=F(+\infty)=1$.

这是由于 $F(x)=P\{X\leqslant x\}$,而 $0\leqslant P\{X\leqslant x\}\leqslant 1$ 且 $\{X\leqslant -\infty\}=\varnothing$,$\{X\leqslant +\infty\}=S$.

(2)$F(x)$ 是不减函数,即当 $x_1<x_2$ 时,$F(x_1)\leqslant F(x_2)$.

事实上,由式(2.3),对于任意实数 $x_1,x_2(x_1<x_2)$,有

$F(x_2)-F(x_1)=P\{x_1<X\leqslant x_2\}\geqslant 0$,即 $F(x_1)\leqslant F(x_2)$.

(3)$F(x)$ 是右连续的,即 $F(x+0)=F(x)$(证略).

(4) 对每个 x_0,$P\{X=x_0\}=F(x_0)-F(x_0-0)$.

证明由读者自己思考.

例 2.11　设随机变量 X 的分布律为

X	0	1	2
P_k	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$

求:(1)X 的分布函数;(2)$P\{X\leqslant \frac{1}{2}\}$;(3)$P\{\frac{1}{2}<X\leqslant \frac{3}{2}\}$;(4)$P\{1\leqslant X\leqslant 2\}$.

解　(1) 由概率的有限可加性,得所求分布函数为

$$F(x)=\begin{cases}0, & x<0,\\ \dfrac{1}{3}, & 0\leqslant x<1,\\ \dfrac{1}{3}+\dfrac{1}{6}, & 1\leqslant x<2,\\ \dfrac{1}{3}+\dfrac{1}{6}+\dfrac{1}{2}, & x\geqslant 2.\end{cases}$$

即

$$F(x)=\begin{cases}0, & x<0,\\ \dfrac{1}{3}, & 0\leqslant x<1,\\ \dfrac{1}{2}, & 1\leqslant x<2\\ 1, & x\geqslant 2.\end{cases}$$

(2)$P\{X\leqslant \frac{1}{2}\}=F(\frac{1}{2})=\frac{1}{3}$

(3)$P\{\frac{1}{2}<X\leqslant \frac{3}{2}\}=F(\frac{3}{2})-F(\frac{1}{2})=\frac{1}{2}-\frac{1}{3}=\frac{1}{6}$;

(4)$P\{1\leqslant X\leqslant 2\}=F(2)-F(1)+P\{X=1\}=1-\frac{1}{2}+\frac{1}{6}=\frac{2}{3}$.

§2.4　连续型随机变量

2.4.1　连续型随机变量的概率密度

在 §2.2 中，我们讨论了一类重要的随机变量 —— 离散型随机变量.在这一节中我们来讨论另一类重要的随机变量 —— 连续型随机变量，我们先引入概率密度函数的概率.

定义 2.4　如果对于随机变量 X 的分布函数 $F(x)$，存在非负函数 $f(x)$，使对于任意实数 x，有

$$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t, \tag{2.4}$$

则称 X 为连续型随机变量，其中 $f(x)$ 称为 X 的概率密度函数，简称概率密度或密度.

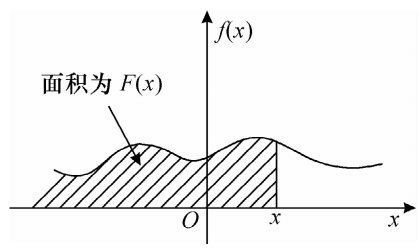

图 2－3

$F(x)$ 与其密度 $f(x)$ 关系的几何解释见图 2－3.

由积分上限函数的性质易知，尽管密度函数可以不连续，但对连续型随机变量 X 而言，其分布函数一定连续且唯一.在实际应用中遇到的基本上是离散型和连续型随机变量，在本书中只讨论这两种随机变量.

由定义 2.4 知道，概率密度 $f(x)$ 具有以下性质：

(1)$f(x)\geqslant 0$；

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

(3) 对于任意实数 $x_1,x_2(x_1\leqslant x_2)$，

$P\{x_1<X\leqslant x_2\}=F(x_2)-F(x_1)=\int_{x_1}^{x_2}f(x)\mathrm{d}x$；

(4) 对 $f(x)$ 的连续点，有

$$F'(x)=f(x) \tag{2.5}$$

式(2.4)与式(2.5)表示了分布函数与概率密度间的两个关系，利用这些关系，可以根据分布函数和概率密度中的一个推出另一个.