

第 1 章 随机事件与概率

概率论是研究随机现象的统计规律的学科,是统计学的理论基础.事件和概率是概率论中最基本的两个概念.本章重点介绍概率论的两个最基本概念:随机事件及其概率.主要内容包括:随机试验、样本空间、随机事件的频率与概率、条件概率以及事件的独立性等概率论中的最基本概念.介绍古典概型和几何概型、条件概率、乘法公式、全概率公式与贝叶斯公式及二项概率公式等计算概率的基本公式.这些内容是进一步学习概率论的基础.

1.1 随 机 事 件

1.1.1 随机现象

在自然界和人类社会生活中普遍存在着两类现象,一类是在一定条件下必然发生的现象,称为**确定性现象**.例如,在没有外力作用下,作匀速直线运动的物体必然继续作匀速直线运动;人从地面向上抛起的石块经过一段时间必然落到地面;在标准大气压下,水加热到 100°C 必然会沸腾;等等.

另一类现象是在一定条件下可能发生也可能不发生的现象,称为**随机现象**.例如,投掷一枚硬币,我们不能事先预知将出现正面还是反面;投掷一枚骰子出现的点数也不能预知;从一大批产品中任取一个产品,这个产品可能是正品,也可能是废品,其结果带有偶然性.人们通过长期实践并深入研究之后,发现这类现象在大量重复试验或观察下,它的结果却呈现出某种规律性.概率论与数理统计就是研究和揭示随机现象这种规律性的一门数学学科.其理论和方法被广泛地应用于自然科学、社会科学、工程技术和经济管理等诸多领域.

1.1.2 随机试验与随机事件

为了研究和揭示随机现象的统计规律性,我们要对随机现象进行大量重复观察.我们把观察的过程称为试验,满足下列条件的试验称为**随机试验**,本书以下简称**试验**.一般地,一个随机试验要求满足下列条件:

- (1) 可重复性:试验可以在相同条件下重复进行多次,甚至进行无限多次;
- (2) 可观测性:每次试验的结果具有多种可能性,并能事先明确试验的所有可能结果;
- (3) 随机性:试验之前不能准确预言该次试验将出现哪一种结果.

我们用 E 表示一个试验, 用 ω 表示随机试验 E 的最基本的结果, 称为**样本点**, 用 $\Omega = \{\omega\}$ 表示随机试验 E 的最基本结果的集合, 称为**样本空间**或**基本空间**.

例 1.1.1 在投掷一枚硬币观察其出现正面还是出现反面的试验中, 有两个样本点: 正面、反面. 样本空间为 $\Omega = \{\text{正面}, \text{反面}\}$. 记 $\omega_1 = \text{“正面”}$, $\omega_2 = \text{“反面”}$, 则样本空间可表示为

$$\Omega = \{\omega_1, \omega_2\}.$$

例 1.1.2 投掷一枚骰子, 观察出现的点数, 则基本结果是“出现 i 点”, 分别记为 $\omega_i (i = 1, 2, 3, 4, 5, 6)$, 则试验的样本空间为

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}.$$

也可简记为 $\Omega = \{1, 2, 3, 4, 5, 6\}$.

例 1.1.3 将一枚硬币抛掷两次, 观察正面出现的次数, 记为 $\omega_i (i = 0, 1, 2)$, 则样本空间为

$$\Omega = \{\omega_0, \omega_1, \omega_2\}.$$

也可简记为 $\Omega = \{0, 1, 2\}$.

例 1.1.4 记录某电话台在一分钟内接到的呼叫次数, 则样本空间为

$$\Omega = \{0, 1, 2, \dots\}.$$

例 1.1.5 测量某一零件的长度, 考察其观测结果与真正长度的误差, 样本空间可取作 $[-M, M]$, 其中 M 为最大误差. 如果无法确定这一最大值, 将 Ω 取作 $(-\infty, +\infty)$ 也可.

对于随机现象, 我们关心的通常是在随机试验中某一结果是否会出现, 或会出现什么结果, 这些结果称为**随机事件**. 习惯上, 用大写字母 A, B, C 等表示.

例如在例 1.1.2 中, 样本空间 $\Omega = \{1, 2, 3, 4, 5, 6\}$, 如果以 A 表示“得到的为偶数点”, 则显然 $A = \{2, 4, 6\}$, B 表示“得到的点数大于或等于 3”, 则 $B = \{3, 4, 5, 6\}$. 这里 A, B 均为随机事件, 它们都由 Ω 中的若干个样本点所构成.

由此可见, 准确地讲, **随机事件**是由若干个样本点组成的集合, 或者说是样本空间的子集. 称某个事件 A 发生, 当且仅当该事件所包含的某个样本点出现.

由一个样本点组成的事件, 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**.

例 1.1.6 在抛掷一枚骰子的实验中, 分别记“点数是 1”为 A , “点数是偶数”为 B , “点数大于或等于 2”为 C , “点数大于 6”为 D , “点数小于或等于 6”为 F , 则 $A = \{1\}, B = \{2, 4, 6\}, C = \{2, 3, 4, 5, 6\}, D = \emptyset, F = \Omega = \{1, 2, 3, 4, 5, 6\}$.

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**; 如果有无限多个样本点, 则称它为**无限样本空间**. 例 1.1.1 ~ 例 1.1.3 中的样本空间都是有限样本空间, 在例 1.1.4 和例 1.1.5 中的样本空间都是无限样本空间.

1.1.3 随机事件间的关系和事件的运算

在一个随机试验中, 一般有很多随机事件. 为了通过对简单事件的研究来掌握比较复杂的事件的规律, 需要研究事件间的关系及事件的运算. 由于事件是样本空间的子集, 因此事件间的关系及运算与集合间的关系及运算是相互对应的. 关键是要理解事件间的关系及运算的概率含义.

在以下的讨论中, 假定 Ω 是试验 E 的样本空间, 所论及的事件都是同一试验 E 的事件.

1. 事件的包含

如果事件 A 发生必然导致事件 B 发生, 则称事件 B **包含** 事件 A , 或称事件 A **包含于** 事件 B , 或称事件 A 是事件 B 的子事件, 记作

$$B \supset A \quad \text{或} \quad A \subset B.$$

易知, 对任意事件 A , 有

$$\emptyset \subset A \subset \Omega.$$

2. 事件的相等

如果事件 A 包含事件 B , 事件 B 也包含事件 A , 则称事件 A 与事件 B **相等**, 记作 $A = B$. 事件 A 与事件 B 相等, 表明 A 和 B 是样本空间的同一子集.

3. 事件的并 (或和)

如果事件 A 和事件 B 至少有一个发生, 则称这样的一个事件为事件 A 与事件 B 的**并或和**, 记作 $A \cup B$, 即

$$A \cup B = \{A \text{ 发生或 } B \text{ 发生}\} = \{\omega | \omega \in A \text{ 或 } \omega \in B\}.$$

在例 1.1.6 中, $A \cup B = \{1, 2, 4, 6\}$.

事件 A 和事件 B 作为样本空间 Ω 的子集, 事件 $A \cup B$ 就是子集 A 与 B 的**并集**.

4. 事件的交 (或积)

如果事件 A 和事件 B 同时发生, 则称这样的一个事件为事件 A 与事件 B 的**交或积**, 记作 $A \cap B$ 或 AB , 即

$$A \cap B = \{A \text{ 发生且 } B \text{ 发生}\} = \{\omega | \omega \in A \text{ 且 } \omega \in B\}.$$

在例 1.1.6 中, $B \cap C = \{2, 4, 6\}$.

事件 A 和事件 B 作为样本空间 Ω 的子集, 事件 $A \cap B$ 就是子集 A 与 B 的交集.

5. 事件的差

如果事件 A 发生而事件 B 不发生, 则称这样的一个事件为事件 A 与 B 的差事件, 记作 $A - B$, 即

$$A - B = \{A \text{ 发生但 } B \text{ 不发生}\} = \{\omega | \omega \in A \text{ 但 } \omega \notin B\}.$$

在例 1.1.6 中, $C - B = \{3, 5\}$.

6. 互不相容事件

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 互斥.

在例 1.1.6 中, 事件 A 和事件 B 是互不相容的.

7. 对立事件 (或互逆事件)

如果在每一次试验中事件 A 和事件 B 都有一个且仅有一个发生, 则称事件 A 与事件 B 是对立的或互逆的, 其中的一个事件是另一个事件的对立事件, 记作 $\bar{A} = B$, 或 $\bar{B} = A$. 显然 $\bar{\bar{A}} = A$.

在例 1.1.6 中, 事件 A 与事件 C 是对立的.

8. 有限个或可数无穷多个事件的并与交

设有 n 个事件 $A_1, A_2, \dots, A_n$, 则称“ $A_1, A_2, \dots, A_n$ 至少有一个发生”这一事件为事件 $A_1, A_2, \dots, A_n$ 的并, 记作 $A_1 \cup A_2 \cup \dots \cup A_n$ 或 $\bigcup_{i=1}^n A_i$. 称“ $A_1, A_2, \dots, A_n$ 同时发生”这一事件为事件 $A_1, A_2, \dots, A_n$ 的交, 记作 $A_1 A_2 \dots A_n$ 或 $\bigcap_{i=1}^n A_i$.

设有可数个事件 $A_1, A_2, \dots, A_n, \dots$, 则称“ $A_1, A_2, \dots, A_n, \dots$ 至少有一个发生”这一事件为事件 $A_1, A_2, \dots, A_n, \dots$ 的并, 记作 $\bigcup_{i=1}^{\infty} A_i$. 称“ $A_1, A_2, \dots, A_n, \dots$ 同时发生”这一事件为事件 $A_1, A_2, \dots, A_n, \dots$ 的交, 记作 $\bigcap_{i=1}^{\infty} A_i$.

9. 完备事件组

设 $A_1, A_2, \dots, A_n, \dots$ 是有限个或可数无穷多个事件, 如果满足:

(1) $A_i A_j = \emptyset, i \neq j, i, j = 1, 2, \dots$;

(2) $\bigcup_i A_i = \Omega$.

则称 $A_1, A_2, \dots, A_n, \dots$ 为样本空间 Ω 的一个完备事件组或 (有限) 分割.

10. 事件间的关系与运算的文氏图

上述关于事件间的各种关系与运算可直观地用图形 (文氏图) 来表示 (见图 1.1).

类似集合的运算, 事件的运算有以下运算法则:

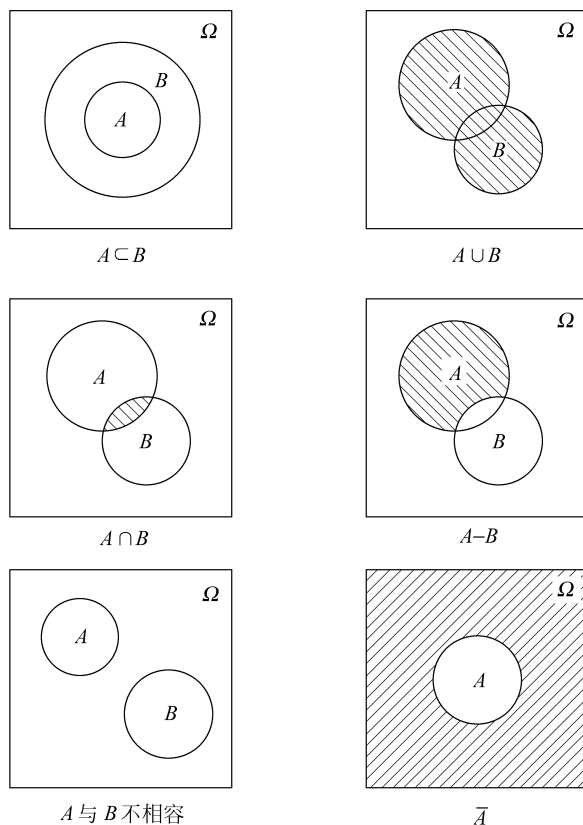


图 1.1 事件间的关系与运算的文氏图

- (1) 交换律: $A \cup B = B \cup A, A \cap B = B \cap A$;
- (2) 结合律: $(A \cup B) \cup C = A \cup (B \cup C), (AB)C = A(BC)$;
- (3) 分配律: $(A \cup B) \cap C = (A \cap C) \cup (B \cap C),$
 $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$;
- (4) 德摩根 (De Morgan) 定律 (对偶律):

$$\overline{\bigcup_{i=1}^n A_i} = \bigcap_{i=1}^n \bar{A}_i, \quad \overline{\bigcap_{i=1}^n A_i} = \bigcup_{i=1}^n \bar{A}_i.$$

对于多个随机事件, 以上的运算法则也是成立的.

例 1.1.7 某工人加工了三个零件, 设 A_i 为事件“加工的第 i 个零件是合格品” ($i = 1, 2, 3$), 试用 A_1, A_2, A_3 表示下列事件:

- (1) 只有第一个零件是合格品;
- (2) 只有一个零件是合格品;
- (3) 至少有一个零件是合格品;

(4) 最多有一个零件是合格品.

解 用 A, B, C, D 分别表示 (1)、(2)、(3)、(4) 所述事件, 则有

(1) 事件 A 发生, 即“只有第一个零件是合格品”这一事件发生, 即事件 A_1 发生并且 A_2, A_3 都不发生, 故

$$A = A_1 \overline{A_2} \overline{A_3}.$$

(2) 事件 B 发生, 意味着在三个零件中有一个是合格品, 并且另外两个是不合格品, 因此

$$B = A_1 \overline{A_2} \overline{A_3} \cup \overline{A_1} A_2 \overline{A_3} \cup \overline{A_1} \overline{A_2} A_3.$$

(3) 事件 C 发生, 即在三个零件中至少有一个是合格品, 故

$$C = A_1 \cup A_2 \cup A_3.$$

(4) 事件 D 发生, 也就是事件“在三个零件中任意两个都不能同时为合格品”发生, 从而

$$D = \overline{A_1} \overline{A_2} \cup \overline{A_1} \overline{A_3} \cup \overline{A_2} \overline{A_3}.$$

也可以表示成

$$D = \overline{A_1} \overline{A_2} \overline{A_3} \cup A_1 \overline{A_2} \overline{A_3} \cup \overline{A_1} A_2 \overline{A_3} \cup \overline{A_1} \overline{A_2} A_3.$$

例 1.1.8 一名射手连续向某个目标射击三次, 事件 A_i 表示该射手第 i 次射击时击中目标 ($i = 1, 2, 3$). 试用文字叙述下列事件:

$$A_1 \cup A_2; \quad A_1 \cup A_2 \cup A_3; \quad A_1 A_2 A_3; \quad A_3 - A_2; \quad A_3 \overline{A_2}; \quad \overline{A_1 \cup A_2};$$

$$\overline{A_1} \overline{A_2}; \quad \overline{A_2} \cup \overline{A_3}; \quad \overline{A_2 A_3}; \quad A_1 A_2 \cup A_1 A_3 \cup A_2 A_3.$$

解 $A_1 \cup A_2$: 前两次中至少有一次击中目标;

$A_1 \cup A_2 \cup A_3$: 三次射击中至少有一次击中目标;

$A_1 A_2 A_3$: 三次射击都击中了目标;

$A_3 - A_2 = A_3 \overline{A_2}$: 第三次击中但第二次未击中;

$\overline{A_1 \cup A_2} = \overline{A_1} \overline{A_2}$: 前两次均未击中目标;

$\overline{A_2} \cup \overline{A_3} = \overline{A_2 A_3}$: 后两次中至少有一次未击中目标;

$A_1 A_2 \cup A_1 A_3 \cup A_2 A_3$: 三次射击中至少有两次击中目标.

例 1.1.9 化简事件 $\overline{(\overline{A} \overline{B} \cup C)} \overline{AC}$.

解 原式 $= \overline{(\overline{A} \overline{B}) \cup C} \cup (AC) = (\overline{\overline{A} \overline{B} \overline{C}}) \cup (AC) = (A \cup B) \overline{C} \cup (AC)$
 $= (A \overline{C}) \cup (B \overline{C}) \cup (AC) = A(\overline{C} \cup C) \cup (B \overline{C}) = A \cup (B \overline{C}).$

1.2 随机事件的频率与概率

1.2.1 频率

概率是概率论中最基本的概念. 在引入此概念之前, 需要介绍频率的概念.

定义 1.2.1 设在相同条件下进行的 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 $\frac{n_A}{n}$ 为事件 A 发生的**频率**, 记作 $f_n(A)$, 即

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

由频率的定义易见其具有下列基本性质:

- (1) $0 \leq f_n(A) \leq 1$;
- (2) $f_n(\Omega) = 1$;
- (3) 若 $A_1, A_2, \dots, A_m$ 是两两互不相容的事件, 则

$$f_n(A_1 \cup A_2 \cup \dots \cup A_m) = f_n(A_1) + f_n(A_2) + \dots + f_n(A_m).$$

事件 A 的频率反映了事件 A 发生的频繁程度. 频率越大, 事件 A 发生越频繁, 这意味着 A 在一次试验中发生的可能性越大.

通过实践人们发现, 当试验次数 n 很大时, 事件 A 发生的频率总会在某个确定的数值附近摆动, 随机事件的频率的这一特性称为频率的稳定性.

历史上有不少人通过投掷硬币的试验研究过频率的稳定性, 表 1.1 列出了他们的试验记录.

表 1.1

试验者	抛掷次数	正面朝上的次数	正面朝上的频率
德摩根	2048	1061	0.5181
蒲 丰	4040	2048	0.5069
费歇尔	10000	4979	0.4979
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

从表 1.1 可以看出频率有随机波动性. 当抛掷次数 n 较小时, 频率随机波动的幅度较大. 但是, 随着抛掷次数 n 的增大, 频率呈现出稳定性, 大致在 0.5 附近摆动, 且逐渐稳定于常数 0.5. 而且研究任何随机事件都有这样一个客观存在的常数与之对应. 这种“频率稳定性”即通常所说的统计规律性, 它是概率这一概念的经验基础. 而所谓事件发生的可能性大小, 就是这个“频率的稳定值”.

定义 1.2.2 在相同条件下重复进行 n 次试验, 如果当 n 增大时, 事件 A 的频率 $\frac{n_A}{n}$ 稳定地在某一常数 p 附近摆动, 则称常数 p 为事件 A 的**概率**, 记为 $P(A) = p$.

概率的这个定义, 称为概率的统计定义. 根据这一定义, 可以把大量重复试验所得到的事件的频率作为事件概率的近似值.

1.2.2 概率

任何一个数学概念都是对现实世界的抽象, 这种抽象使得其具有广泛的适应性, 并成为进一步数学推理的基础. 1933 年, 苏联数学家柯尔莫戈洛夫综合已有的大量成果, 提出了概率的公理化结构, 明确定义了概率等基本概念, 使得概率论成为严谨的数学分支, 推动了概率论的发展.

定义 1.2.3 设随机试验 E 的样本空间为 Ω , 如果对于 Ω 中的每一个事件 A , 有惟一的实数 $P(A)$ 和它对应, 并且这一事件的函数 $P(A)$ 满足以下条件:

- (1) 非负性: 对于任一事件 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 $A_1, A_2, \dots$ (即当 $i \neq j$ 时, 有 $A_i A_j = \emptyset, i, j = 1, 2, \dots$), 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**.

由概率的公理化定义, 可以推导出概率的一些性质.

性质 1.2.1 对于不可能事件 $\emptyset$, 有 $P(\emptyset) = 0$.

证明 取 $A_i = \emptyset (i = 1, 2, \dots)$, 显然这是一列两两不相容的事件, 且 $\bigcup_{i=1}^{\infty} A_i = \emptyset$, 由定义 1.2.3 的 (3), 有

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

因此

$$P(\emptyset) = 0.$$

□

性质 1.2.2 对于两两互不相容的事件 $A_1, A_2, \dots, A_n$, 有

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

这一性质称为概率的**有限可加性**.

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 1.2.1, 有

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

性质 1.2.3 对任一事件 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 1.2.2 得 $P(A \cup \bar{A}) = P(A) + P(\bar{A}) = P(\Omega)$, 即

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

故

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

性质 1.2.4 如果 $A \subset B$, 则有

$$P(B - A) = P(B) - P(A), \quad P(A) \leq P(B).$$

证明 因为 $A \subset B$, 从而有 $B = A \cup (B - A)$, 且 $A(B - A) = \emptyset$, 由性质 1.2.2 得

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

所以

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

由于 $P(B - A) \geq 0$, 因此

$$P(A) \leq P(B). \quad \square$$

性质 1.2.5 对任一事件 A , 有 $P(A) \leq 1$.

证明 因为 $A \subset \Omega$, 由性质 1.2.4 及概率的规范性, 可得

$$P(A) \leq P(\Omega) = 1. \quad \square$$

性质 1.2.6 对于任意两个事件 A, B , 有

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

证明 由于 $B - A = B - AB$, 而 $AB \subset B$, 根据性质 1.2.4, 可得

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

□

性质 1.2.6 称为概率的**减法公式**.

性质 1.2.7 对于任意两个事件 A, B , 有

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

$$P(A \cup B) \leqslant P(A) + P(B).$$

证明 因为 $A \cup B = A \cup (B - AB)$, 且 $A(B - AB) = \emptyset, AB \subset B$, 由性质 1.2.2 及性质 1.2.4, 可得

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

由于 $P(AB) \geqslant 0$, 因此

$$P(A \cup B) \leqslant P(A) + P(B).$$

□

性质 1.2.7 中的第一个公式称为概率的**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

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

这个公式称为概率的**一般加法公式**.

例 1.2.1 已知 $P(\bar{A}) = 0.5$, $P(\bar{A}B) = 0.2$, $P(B) = 0.4$, 求: (1) $P(AB)$; (2) $P(A - B)$; (3) $P(A \cup B)$; (4) $P(\bar{A} \bar{B})$.

解 (1) 因为 $AB \cup \bar{A}B = B$, 且 AB 与 $\bar{A}B$ 互不相容, 故有

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

于是

$$P(AB) = P(B) - P(\bar{A}B) = 0.4 - 0.2 = 0.2;$$

$$(2) \quad P(A) = 1 - P(\bar{A}) = 1 - 0.5 = 0.5,$$

$$P(A - B) = P(A) - P(AB) = 0.5 - 0.2 = 0.3;$$