Contents

Chapter 2

Fundamental Concepts of Probability

2.1 Review: Set Theory

Set Definitions:

Definition 2.1 A set is a collection of objects(*abstract or concrete*), and a set is denoted by a capital letter.

Definition 2.2 The objects are called the "elements" of a set.

Example 2.1

A game of die casting:

Figure 2.1: The set representation of the game of a die casting game.

(1) $S = \{a_1, a_2, a_3, a_4, a_5, a_6\}$:sample space

(2) $a_i \in S$: element

(3) $A = \{x | x = a_1 \text{ or } a_2\}$:event, subset

(4) $B = \{a_3\}$:event, subset

Definition 2.3 Finite set: a set is "finite" if it has a finite number of elements.

Definition 2.4 Infinite set: an infinite set is a set which is NOT finite.

1. **Countable infinite set:** An (infinite) set is said to be countable iff its elements can be put into one-to-one correspondence with natural numbers $n = 1, 2, \ldots$.

   (e.g.) $A = \{m|m = 1, 3, 5, 7, 9, \ldots\}$, where $m = 2n - 1, \quad n = 1, 2, 3, \ldots$.

2. **Uncountable infinite set:** An (infinite) set is called uncountable, if it is NOT countable.

   (e.g.) $B = \{x|0 \leq x \leq 1\}$, where x is real number.

Definition 2.5 Subsets:

Figure 2.2: The subsets.

1. $C \subseteq A$ (improper subset) : if $a_i \in C$, then $a_i \in A$.
2. $B \subset A$ (proper subset) : $\exists$ at least one $a_i \ni$: $a_i \notin B$ and $a_i \in A$.
3. $C = A$ (equality) : if and only if (iff) $A \subseteq C$ and $A \supseteq C$

Definition 2.6 Complement of a set: $\overline{A}$ or A^c

$$\overline{A} = \{a|x \in S \text{ and } x \notin A\}$$

Figure 2.3: The complement of a set A.

Definition 2.7 Empty set: ϕ

ϕ is a set that has no element in it. (called "null set" as well.)

Definition 2.8 Disjoint sets:

Two sets A and B are called *disjoint* or *mutually exclusive* iff A and B have no common elements.

Definition 2.9 Union, Intersection, and Difference:

1. Union: $A \cup B = \{x | x \in A \text{ and/or } x \in B\}$
2. Intersection: $A \cap B = \{x | x \in A \text{ and } x \in B\}$
3. Difference: $A - B = \{x | x \in A \text{ but } x \notin B\}$

Venn Diagram: graphical representation of sets

Figure 2.4: A Venn diagram of sets.

Algebra of Sets:

1. Idempotent law:
   (a) $A \cup A = A$
   (b) $A \cap A = A$

2. Associative law:
   (a) $(A \cup B) \cup C = A \cup (B \cup C)$
   (b) $(A \cap B) \cap C = A \cap (B \cap C)$

3. Commutative law:
   (a) $A \cup B = B \cup A$
   (b) $A \cap B = B \cap A$

4. Distributive law:
   (a) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
   (b) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

5. Complement law:
   (a) $A \cup \overline{A} = S$
   (b) $A \cap \overline{A} = \phi$
   (c) $\overline{\overline{A}} = A$
   (d) $\overline{S} = \phi$ (or $\overline{\phi} = S$)

6. DeMorgan's law:
   (a) $\overline{A \cup B} = \overline{A} \cap \overline{B}$
   (b) $\overline{A \cap B} = \overline{A} \cup \overline{B}$

7. Identity law:

   (a) $A \cup \phi = A$
   (b) $A \cup S = S$
   (c) $A \cap S = A$
   (d) $A \cap \phi = \phi$

   Note: Check the above algebra using Venn diagram.

8. Duality principle:

   In a given set relation, if we replace;

$$\cup \longrightarrow \cap$$
$$\cap \longrightarrow \cup$$
$$S \longrightarrow \phi$$
$$\phi \longrightarrow S$$

   then, the equality still holds for the new set relation!!!

   remark: For the above 7 set relations, show the effectiveness of the "duality principle".

Analogy between probability concept and set theory;

Probability concept in chance experiments	**Set theory**
simple outcome	element of a set
sample sapce (a collection of all possible outcomes)	whole set $= S$
(compound) event	subset
mutually exclusive events	$A \cap B = \phi$
simultaneous occurrence of events	$A \cap B$
at least one event shows up	$A \cup B$
opposite event	$\overline{A}$
impossible event	ϕ
event A must occur	$A = S$
from the occurrence of event A follows the inevitable event B	$A \subset B$

NOTE:

1. Events are subsets of S!!!
2. The "head" and "tail" of a coin are good examples of opposite events.
3. Impossible event and a event with probability 0 are NOT the same in a rigorous sense!!!
4. In a chance experiment of casting a die, let $A = \{2 \text{ shows up}\}$, and $B = \{\text{ even numbers show up}\}$. Then, the occurrence of event A implies the occurrence of event B, i.e. $A \subset B$.

2.2 Probability Axiom and Probability Space

As we mentioned in Chapter 1, there are 3 basic components for the mathematical structure of probability concept:

1. Sample space S
2. Event A
3. Probability function $P(\cdot)$

$\Longrightarrow$ In order to apply the probability function $P(\cdot)$ to an event A(which is the variable of the function $P(\cdot)$), we therefore need *the domain* of event(i.e. variable):

Field ($\mathcal{F}$): [1]

Assuming S is a finite space, the *field* $\mathcal{F}$ is defined as a collection (set) of subsets of S satisfying the following conditions:

(1) If $A_1, A_2 \in \mathcal{F}$, then $A_1 \cup A_2 \in \mathcal{F}$.

(2) If $A \in \mathcal{F}$, then $\overline{A} \in \mathcal{F}$.

Remarks:

1. We can regard the "field" as a mathematical concept which is necessary to set up a function domain for the probability function $P(\cdot)$.

Figure 2.5: The domain and the range of $P(\cdot)$.

2. We cannot use the sample space S itself as a domain of $P(\cdot)$, since $P(\cdot)$ is a function of events(which are usually the combination of elements in S), NOT a function of elements.
3. Therefore, the domain of $P(\cdot)$, i.e. $\mathcal{F}$ must be closed by basic set operation $\ni$: union($\cup$), intersection($\cap$), and complement(^{c}).

[1]The field is sometimes called as "Algebra" as well.

Question:
The above dfinition of field implies that *since* $A_1 \cup A_2$ *and* $\overline{A}$ *could be the events that we want to compute the probability, they must also be within the domain* $\mathcal{F}$ *so that we can apply the set function* $P(\cdot)$.
Then what about another set algebra, i.e. the *intersection*?

Answer: Above definition of field *implicitely* contains the following statement as well:

$$\text{If } A_1, A_2 \in \mathcal{F}, \text{ then } A_1 \cap A_2 \in \mathcal{F}$$

proof:

If $A_1, A_2 \in \mathcal{F}$, then $\overline{A_1}, \overline{A_2} \in \mathcal{F}$. (from (2))

If $\overline{A_1}, \overline{A_2} \in \mathcal{F}$, then $\overline{A_1} \cup \overline{A_2} \in \mathcal{F}$. (from (1))

If $\overline{A_1} \cup \overline{A_2} \in \mathcal{F}$, then $\overline{\overline{A_1} \cup \overline{A_2}} = A_1 \cap A_2 \in \mathcal{F}$. (from (2)) Q.E.D.

Example 2.2

Given a sample sapce S as: $S = \{1, 2, 3, 4, 5, 6\}$, determine whether the following class of subsets can be a field.

$$\{\phi, S, A_1 = \{1, 3, 5\}, A_2 = \{2, 4, 6\}\}$$

Solution: You can easily check the conditions (1) and (2) of a field are met, and the answer is YES.

(cf) Other possibilities of field:

(1) $\{\phi, S, \{1, 2, 3\}, \{4, 5, 6\}\}$

(2) $\{\phi, S, \{1\}, \{2, 3, 4, 5, 6\}\}$

(3) $\{\phi, S\}$

$\vdots$

$\vdots$

Remarks:

1. Notice that he *subsets* A, B, C in the sample space S are the *elements* of the field $\mathcal{F}$. [2]

Figure 2.6: Comparison b/w the sample space and the field as sets.

2. If the sample space S has an infinite number of elements, then $\mathcal{F}$ is said to be the σ-*field*, and the necessary condtions for the sigma-field are:

   (1) If $A_i \in \mathcal{F} \quad i = 1, 2, 3, \ldots$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

   (2) If $A \in \mathcal{F}$, then $\overline{A} \in \mathcal{F}$.

Set function:

A set function operates on a set and assigns a *real* number to each set.

Figure 2.7: A set function $G(\cdot)$.

Note: Strictly speaking, a set function should be defined on a field, not the samples space.

[2] Recall that the field $\mathcal{F}$ is a set of subsets.

Probability Axiom:

A set function $P(\cdot)$ is called a *probability* on the subsets of the sample space S iff $P(A)$ satisfies the following 3 axioms:

1. $P(A) \geq 0, \quad \forall A \in \mathcal{F}$ of S
2. $P(S) = 1$
3. If $A_i \cap A_j = \phi, \quad \forall i \neq j$ (i.e. disjoint), then

$$P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$$

NOTE: $P(A)$ can be arbitrarily assigned provided that it satisfies the above 3 axioms.

(e.g.) A die need not necessarily be fair!!! For example, we could assign the probabilities as: $P(a_1) = P(a_2) = P(a_3) = \frac{1}{4}$ and $P(a_4) = P(a_5) = P(a_6) = \frac{1}{12}$, and still satisfies all of the probability axioms.

Example 2.3

We repeat the example **??**, which is an experiment of rolling a fair die, and check the validity of the probability axioms.

Solution: The sample space S is as follows:

$$S = \{1, 2, 3, 4, 5, 6\}$$

(a) $P(\{a_i\}) = \frac{1}{6} \geq 0, \ \forall \, i.$

(b) $P(\{a_2\} \cup \{a_3\}) = P(\{a_2\}) + P(\{a_3\}) = \frac{1}{3}.$

(c) $P(S) = P(\{a_1\} \cup \{a_2\} \cup \cdots \cup \{a_6\}) = \sum_{i=1}^{6} P(\{a_i\}) = 6 \cdot \frac{1}{6} = 1.$

Example 2.4

Consider a game of spinning a wheel, where real numbers from 0 to 100 are marked along the perimeter, so the sample space is: [3]

$$S = \{x | 0 < x \leq 100\}, \quad \text{real numbers}$$

Figure 2.8: A game of spinning a wheel.

Assuming the wheel is fair, it seems to be reasonable to assign probabilities as follows:

$$P(A) = \frac{x_2 - x_1}{100}, \quad \text{where } A = \{x_1 < x \leq x_2\}, \;\; x_2 \geq x_1$$

Check the vailidity of the above assigned $P(\cdot)$.

Solution:

Axiom 1: $P(A) = \frac{x_2 - x_1}{100} \geq 0$ since $x_2 \geq x_1$

Axiom 2: $P(S) = \frac{100-0}{100} = 1$

Axiom 3: Define an event A_n as: $A_n = \{x_{n-1} < x \leq x_n\}$ where $x_n = \frac{100}{N} \cdot n$, $n = 1, 2, 3, \ldots, N$ and $x_0 = 0$, then,

$$\begin{aligned} P(A_n) &= \frac{x_n - x_{n-1}}{100} \\ &= \frac{1}{100} \cdot (n - n + 1)\frac{100}{N} \\ &= \frac{1}{N} \end{aligned}$$

Since $\{A_n\}_{n=1}^{N}$ are disjoint, we have:

$$1 = P\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} P(A_n) = \frac{1}{N} \cdot N = 1 = P(S)$$

[3]This is similar to a very famous TV game show called "wheel of fortune".

NOTE:

Let $N \to \infty$, then $x_n - x_{n-1} \to 0$ and $A_n \to x_n$(a specific number), and thus:

$\lim_{N\to\infty} P(A_n) = P(x_n) = \lim_{N\to\infty} \frac{1}{N} = 0$

$\Longrightarrow$ Event $\{x = x_n\}$ is an event with probability zero.

$\Longrightarrow$ Impossible event $\neq$ event w/ probability 0

Remarks:

1. Event w/ zero probability is NOT an impossible event, rather it means that the event may occur once, but *never again*!!!
2. In a similar way, event $\{x \neq x_n\}$ is an event w/ probability 1, but that does NOT mean that the event MUST occur(i.e. NOT a whole set).

Example 2.5

Consider a game of casting two dice. In this case, there are 36 possible outcomes, and if dice are fair wetend to assign probabilities as : [4]

$$P(a_i) = \frac{1}{36}, \quad \forall\, i = 1, 2, 3, \ldots, 36$$

Figure 2.9: Sample space of rolling two dice.

What is the probability of getting sum of 6?

[4]Strictly speaking the expreesion should be $P(\{a_i\}) = \frac{1}{36}$, since $P(\cdot)$ is a set function. However, for notational convenience, let us use them interchangeably.

Solution:

The event that we are trying to compute its probability is as follows:

$$\begin{aligned} A = \{(i,j) | i+j=6\} &= \{(1,5),(5,1),(2,4),(4,2),(3,3)\} \\ &= \bigcup_{i+j=6,\ 1\le i,j\le 6} (i,j) \\ &: \text{union of disjoint sets} \end{aligned}$$

Therefore, applying the probability axiom3, we have:

$$P(A) = \sum_{i+j=6,\ 1\le i,j\le 6} P((i,j)) = \frac{5}{36}$$

Terminology:

1. Sample space S
2. Field $\mathcal{F}$
3. Measurable space $(S, \mathcal{F})$: [5]

   A sample space S and a field $\mathcal{F}$ of subsets of S with property that S is the union of all members of $\mathcal{F}$
4. Probability space $(S, \mathcal{F}, P)$

NOTE:

The probability space $(S, \mathcal{F}, P)$ is the basic mathematical model for a chance experiment. So, when we say "probability", there always IS an associated probability space, and we start from this probability space to calculate the probability of various complex events.

[5] Remember that a vector space is defined as a set of vectors, which is closed by (1) vector addition and (2) scalar product.

2.3 Joint Probability

Given $(S, \mathcal{F}, P)$, the joint probability $P(A \cap B)$ of two events A and B of a probability space satisfies the following equation:

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

Figure 2.10: The Venn diagram for the joint probability $P(A \cap B)$.

Proof: [6]

We can check from above diagram that:

$$\begin{cases} A \cup B &= A \cup (B - (A \cap B)) \quad : \text{union of disjoint sets} \\ B &= (A \cap B) \cup (B - (A \cap B)) \quad : \text{union of disjoint sets} \end{cases}$$

Therefore, we have, by the axiom #3 of probability, the following probability relations:

$$P(A \cup B) = P(A) + P\left(B - (A \cap B)\right) \tag{2.1}$$

$$P(B) = P(A \cap B) + P\left(B - (A \cap B)\right) \tag{2.2}$$

Substracting (2.2) from (2.1), we get

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

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

Q.E.D [7]

Note:

1. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
2. $P(A \cup B) = P(A) + P(B)$ if and only if (iff) $A \cap B = \phi$

[6]All we know about probability so far is the 3 axioms on probability, and we must rely only on these axioms to prove prbability relations!!!

[7]Q.E.D. is the abbreviation of *Quad Erat Demonstrandum* in Latin, which means the end of demonstration.

Assignment:

Show that

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

.

Figure 2.11: The Venn diagram for the complement probability $P(\overline{A})$.

proof:

¿From the complement law of set relation, the sample space S can be expressed as the union of two *disjoint* sets as:

$$S = A \cup \overline{A}$$

By the axiom #2 and #3 of probability, we have:

$$1 = P(S) = P(A) + P(\overline{A})$$

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

2.4 Conditional Probability, Total Probability Law, and Baye's Theorem

2.4.1 Conditional Probability

Given $(S, \mathcal{F}, P)$, suppose that $P(B)$ of an event B is non-zero, i.e. $P(B) > 0$. Then, we define the conditional probability of an event A given B by the following expression:

$$P(B|A) \triangleq \frac{P(A \cap B)}{P(B)}$$

Remarks:

1. Note that if $P(A) \neq 0$ as well, we have:

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

2. In the case of the conditional probability, we can regard the definition as confining the sample sapce (or the whole set) S to the event B. In other words, the ordinary probability of an event A is the *conditional probability* of event A given S:

   proof:

$$A = A \cap S \quad \text{(by identity law)}$$

$$P(S) = 1 \qquad \text{(by axiom \#2)}$$

$$\longrightarrow \quad P(A) = P(A \cap S) = \frac{P(A \cap S)}{P(S)} = P(A|S)$$

Question: Is $P(A|B)$ defined above a valid probability? In other words, does $P(A|B)$ satisfy the three axioms of probability?

Answer: YES.

check:

1) Since the event $A \cap B$ which is a subset of S is an element of the field $\mathcal{F}$, we can apply the probability function $P(\cdot)$, and by the assumption $P(B) \neq 0$. Therefore, we have:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \geq 0 \quad : \text{by axiom \#1}$$

2)

$$P(S|B) = \frac{P(S \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

3) Let A_1 and A_2 be two disjoint sets, then [8]

$$\begin{aligned} P(A_1 \cup A_2|B) &= \frac{P((A_1 \cup A_2) \cap B)}{P(B)} \\ &= \frac{P((A_1 \cap B) \cup (A_2) \cap B))}{P(B)} \\ &= \frac{P(A_1 \cap B) + P(A_2) \cap B)}{P(B)} \\ &= P(A_1|B) + P(A_2|B) \end{aligned}$$

Example 2.6

A box contains 100 resistors, which are composed of 22Ω, 47Ω, and 100Ω with 5% and 10% tolerance as follows:

Ω	5%	10%	total
22	10	14	24
47	28	16	44
100	24	8	32
	62	38	100

Soppose we draw a 22Ω resistor, then what is the probability that its tolerance is 10%?

[8] Note that $A_1 \cap B$ and $A_2 \cap B$ are also disjoint to each other.

Solution:

Figure 2.12: The samples space composed of 100 points.

Let's define the following events:

$$A = \text{Draw a } 22\Omega \text{ resistor}$$

$$B = \text{Draw a } 10\% \text{ tolerance resistor}$$

Then, the probability we want to compute is the following conditional probability:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{14}{100} / \frac{24}{100} = \frac{14}{24} \approx 0.58$$

(cf) Note that:

(1) $P(B) = \frac{38}{100}$, and thus $P(B) < P(B|A)$.

(2) By the condition of event A, the new sample sapce S' is now shrinked and composed of 24 resistors(i.e. 22Ω resistors), rather than total of 100 resistors.

2.4.2 Total probability Law

Suppose we are given N disjoint events $\{B_n\}_{n=1}^{N}$ whose union equals to the sample space S. Then, the probability of *any* event A defined on S can be expressed as:

$$\begin{aligned} P(A) &= \sum_{i=1}^{N} P(A|B_i) \cdot P(B_i) \\ &= \sum_{i=1}^{N} P(A \cap B_i) \\ &\quad : \text{total probability law} \end{aligned}$$

Proof:

Figure 2.13: The Venn diagram of a sample sapce as a union of N disjoint events.

We have:

$$S = \bigcup_{i=1}^{N} B_i$$

where B_i's are nutually exclusive events. Therefore, from the identity law of sets,

$$\begin{aligned} A = A \cap S &= A \cap \left\{ \bigcup_{i=1}^{N} B_i \right\} \\ &= \bigcup_{i=1}^{N} \{A \cap B_i\} \quad \text{: distributive law} \end{aligned}$$

which is a union of disjoint events as well.

Now, the probability of event A is:

$$\begin{aligned} P(A) &= P\left[\bigcup_{i=1}^{N} \{A \cap B_i\} \right] \\ &= \sum_{i=1}^{N} P\left(A \cap B_i\right) \qquad \text{:by axiom \#3} \\ &= \sum_{i=1}^{N} P(A|B_i) \cdot P(B_i) \end{aligned}$$

q.e.d.

2.4.3 Baye's Theorem

Soppose the probabilities of two events A and B are not zero, i.e. $P(A) \neq 0$, $P(B) \neq 0$. Then, following probability relation holds between the two events:

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

or, equivalently

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

Proof: (easy!)

¿From the definition of the conditional probability and from the fact that $P(A) \neq 0$ and $P(B) \neq 0$, we have;

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

By equating the last two expressions in above equation, we get

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

q.e.d.

Another form of total probability law [9]

Figure 2.14: The Venn diagram of a sample sapce as a union of N disjoint events.(revisited)

Note that:

(a) $\{B_n\}_{n=1}^{N}$ are disjoint, i.e. $B_i \cap B_j = \phi, \ \forall\, i \neq j$.

(b) $S = \bigcup_{n=1}^{N} B_n$.

[9] This comes from combining the Baye's theorem and the total probability law.

We know from above that for any $B_n \subset S$, following holds:

$$P(B_n|A) = \frac{P(A|B_n)P(B_n)}{P(A)}$$

By applying the total probability law to the denominator of above equation, we derive:

$$P(B_n|A) = \frac{P(A|B_n)P(B_n)}{\sum_{i=1}^{N} P(A|B_i)P(B_i)}$$

: total probability law

where $P(B_i)$ is called the "*a priori* probability"(i.e. before experiment) whereas $P(B_i|A)$ is called "*a posteriori* probability"(i.e. after some experiment) of event B_i.

Example 2.7

Radar detection problem:

Figure 2.15: Radar detection environment.

Define the following events which could happen in running a radar system.

R: Radar reports a target

T: Target exists

N: No target exists

Figure 2.16: The sample space of radar detection environment.

Then, we usually have the following informations on our hand about the radar system: [10]

(i) $P(T) = 0.2$: the probability that a target shows up.

(ii) $P(N) = 0.8$: the probability that no target shows up.

(iii) $P(R|T) = 0.9$: detection probability.

(iv) $P(R|N) = 0.01$: false alarm probability

Question: Suppose the radar reports that an enemy fighter plane is approaching. Then, what are the probabilities that there **actually** is and is not an enemy airplane?

Solution:

These probabilities correspond to the *a posteriori* probabilities, and are as follows:

$$
\begin{aligned}
P(T|R) = \frac{P(R|T)P(T)}{P(R)} &= \frac{P(R|T)P(T)}{P(R|T)P(T) + P(R|N)P(N)} \\
&= \frac{0.9 \times 0.2}{0.9 \times 0.2 + 0.01 \times 0.8} \\
&= \frac{0.18}{0.188} \approx 0.96
\end{aligned}
$$

$$
\begin{aligned}
P(N|R) = \frac{P(R|N)P(N)}{P(R)} &= \frac{P(R|N)P(N)}{P(R|T)P(T) + P(R|N)P(N)} \\
&= \frac{0.01 \times 0.8}{0.9 \times 0.2 + 0.01 \times 0.8} \\
&= \frac{0.008}{0.188} \approx 0.04
\end{aligned}
$$

Note: Comparison between a priori & a posteriori probabilities:

(1) $P(T|R) >> P(T)$

(2) $P(N|R) << P(N)$

What do these results mean?

[10] Among these, $P(T)$ and $P(N)$ correspond to the *a priori* probabilities, whereas $P(R|T)$ and $P(R|N)$ are the design factors of a radar system representing the performance of a radar.

Example 2.8

Digital(binary) communication problem:

Figure 2.17: Digital communication environment.

Define the following events which could happen in digital communication system.

$$
\begin{aligned}
B_0 &= \{\text{signal 0 is sent}\} \\
B_1 &= \{\text{signal 1 is sent}\} \\
A_0 &= \{\text{signal 0 is received}\} \\
A_1 &= \{\text{signal 1 is received}\}
\end{aligned}
$$

Figure 2.18: The sample space of digital communication system.

Assume the following probabilities are known: [11]

(i) $P(B_0) = 0.4, \quad P(A_0|B_0) = 0.9 \quad P(A_1|B_0) = 0.1$

(ii) $P(B_1) = 0.6, \quad P(A_1|B_1) = 0.9 \quad P(A_0|B_1) = 0.1$

Question: Suppose the receiver gets a signal 1. [12] Then, what is the probability that the transmitter **actually** sent the signal 1?(i.e. errorless reception.)

[11] Among these, $P(B_0)$ and $P(B_1)$ correspond to the *a priori* probabilities, whereas $P(A_0|B_0)$, $P(A_1|B_1)$ (errorless transmission) and $P(A_1|B_0)$, $P(A_0|B_1)$ (transmission error) represent the performance of the transmission line(or channel).

[12] This is an estimation problem and we will briefly deal with this issue at later part of this class.

Solution:

These probabilities as well correspond to the *a posteriori* probabilities, and are as follows:

$$
\begin{aligned}
P(B_1|A_1) &= \frac{P(B_1 \cap A_1)}{P(A_1)} \\
&= \frac{P(A_1|B_1)P(B_1)}{P(A_1|B_1)P(B_1) + P(A_1|B_0)P(B_0)} \\
&= \frac{0.9 \times 0.6}{0.9 \times 0.6 + 0.1 \times 0.4} \\
&= \frac{0.54}{0.58} \approx 0.93
\end{aligned}
$$

(cf) What is the probability that the transmitter **actually** sent a signal 1, if the receiver gets the signal 0?

$$
\begin{aligned}
P(B_1|A_0) &= \frac{P(A_0|B_1)P(B_1)}{P(A_1|B_1)P(B_1) + P(A_1|B_0)P(B_0)} = \frac{0.1 \times 0.6}{0.1 \times 0.6 + 0.9 \times 0.4} \\
&= \frac{1}{7} \approx 0.143
\end{aligned}
$$

Note: Comparison between a priori & a posteriori probabilities:

(1) $P(B_1|A_1) >> P(B_1)$

(2) $P(B_1|A_0) << P(B_1)$

What do these results mean?

2.5 Independent Events

Given a probability space $(S, \mathcal{F}, P)$, two events A and B such that $P(A) \neq 0$ and $P(B) \neq 0$ are said to be *statistically independent* if and only if: [13]

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

Fact: Equivalent conditions for $P(A|B) = P(A)$ are:

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

(ii) $P(A \cap B) = P(A) \cdot P(B)$

proof: From the given condition of independence;

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

$$\longrightarrow \quad \frac{P(A \cap B)}{P(B)} = P(A) \quad \Longrightarrow \quad P(A \cap B) = P(A) \cdot P(B)$$

$$\longrightarrow \quad \frac{P(A \cap B)}{P(A)} = P(B) \quad \Longrightarrow \quad P(B|A) = P(B)$$

(e.g.) Define the following three events:

$$\begin{aligned} A &= \{\text{rainy}\} \\ B &= \{\text{attend the probability \& random process class}\} \\ C &= \{\text{bring umbrella to the class}\} \end{aligned}$$

Then obviously(or hopefully) it must be $P(B|A) = P(B)$ whereas $P(C|A) >> P(C)$, which means the events A and B are independent while events A and C are NOT independent(i.e. dependent).

[13] The a priori probability $P(A)$ is equal to the a posteriori probability $P(A|B)$, which means the event B has no effect whatsoever on the event A.

NOTE:
Two events A and B CANNOT be both mutually exclusive and statistically independent at the same time.

why?

1. Mutually exclusive:

$$A \cap B = \phi \quad \rightarrow \quad P(A \cap B) = 0$$

2. Statistically independent:

$$P(A \cap B) = P(A)P(B) \neq 0 \quad \rightarrow \quad A \cap B \neq \phi$$

   This is due to the fact that $P(A) \neq 0$ and $P(B) \neq 0$, and it means there MUST $\exists$ intersection for independent events.

   In conclusion, events A and B can be independent to each other only under the assuption that they can occur simultaneously!!!

Three events(A, B, C):

Fact: If events A, B, C are piecewise independent, that does *not necessarily* mean that they are independent as a triple, and vice versa, i.e.;

$$\left.\begin{array}{l} P(A \cap B) = P(A) \cdot P(B) \\ P(B \cap C) = P(B) \cdot P(C) \\ P(A \cap C) = P(A) \cdot P(C) \end{array}\right\} \overset{(\times)}{\Longleftrightarrow} P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$$

Example 2.9

Suppose we have the four sided dice which is assumed to be fair, i.e.

$$P\left(\{\omega_i\}\right) = \frac{1}{4}, \quad \forall i = 1, 2, 3, 4$$

Then, consider the following three events:

$$A = \{\omega_1, \omega_2\}$$
$$B = \{\omega_1, \omega_3\}$$
$$C = \{\omega_1, \omega_4\}$$

Are these 3 events statistically independent to each other?

Figure 2.19: The sample space for the experiment of throwing 4 sided dice.

Solution:

Since each element of the sample space S is mutually exclusive to each other, we have the following probabilities for the above defines 3 events:

$$P\left(A\right) = P\left(\{\omega_1\}\right) + P\left(\{\omega_2\}\right) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$
$$P\left(B\right) = P\left(\{\omega_1\}\right) + P\left(\{\omega_3\}\right) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$
$$P\left(C\right) = P\left(\{\omega_1\}\right) + P\left(\{\omega_4\}\right) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

Then,

$$P(A \cap B) = P\left(\{\omega_1\}\right) = \frac{1}{4} \equiv P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$
$$P(B \cap C) = P\left(\{\omega_1\}\right) = \frac{1}{4} \equiv P(B) \cdot P(C) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$
$$P(A \cap C) = P\left(\{\omega_1\}\right) = \frac{1}{4} \equiv P(A) \cdot P(C) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

which means that the 3 events are *pairwise* independent.

However, note that:

$$P(A \cap B \cap C) = P\left(\{\omega_1\}\right) = \frac{1}{4} \not\equiv P(A) \cdot P(B) \cdot P(C) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

which means that even if they are independent pairwisely, they may NOT be independent as a triple.

Example 2.10

Suppose we have the 12 sided dice which is assumed to be fair, i.e.

$$S = \{i \mid 1 \leq i \leq 12 : \text{ integer}\} \quad \text{and}$$

$$P(\{\omega_i\}) = \frac{1}{12}, \quad \forall i = 1, 2, 3, \ldots, 12$$

Let three events be defined as follows:

$$A = \{\omega_1, \omega_2, \omega_4, \omega_{12}\}$$

$$B = \{\omega_1, \omega_2, \omega_5, \omega_6, \omega_7, \omega_8\}$$

$$C = \{\omega_1, \omega_5, \omega_6, \omega_7, \omega_9, \omega_{12}\}$$

Are these 3 events statistically independent to each other?

Figure 2.20: The sample space for the experiment of throwing 12 sided dice.

Solution:

Since each element of the sample space S is mutually exclusive to each other, we have the following probabilities for the above defines 3 events:

$$P(A) = \frac{1}{12} \cdot 4 = \frac{1}{3}$$

$$P(B) = \frac{1}{12} \cdot 6 = \frac{1}{2}$$

$$P(C) = \frac{1}{12} \cdot 6 = \frac{1}{2}$$

Then,

$$P(A \cap B \cap C) = P(\{\omega_1\}) = \frac{1}{12} \equiv P(A) \cdot P(B) \cdot P(C) = \frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{12}$$

$$P(A \cap B) = P(\{\omega_1, \omega_2\}) = \frac{2}{12} \equiv P(A) \cdot P(B) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$$

$$P(A \cap C) = P(\{\omega_1, \omega_{12}\}) = \frac{2}{12} \equiv P(A) \cdot P(C) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$$

However, note that:

$$P(B \cap C) = P(\{\omega_1, \omega_5, \omega_6, \omega_7\}) = \frac{4}{12} \not\equiv P(B) \cdot P(C) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

Conclusion:
In the case of 3 events, they are said to be statistically independent iff they are pairwise independent **and** independent as a triple as well.

In general, if events $A_1, A_2, \dots, A_N$ are statistically independent, they MUST satisfy all of the following conditions:

$$P(A_i \cap A_j) = P(A_i) \cdot P(A_j) \quad \forall\, i \neq j$$
$$P(A_i \cap A_j \cap A_k) = P(A_i) \cdot P(A_j) \cdot P(A_k) \quad \forall\, i \neq j \neq k$$
$$\vdots$$
$$\vdots$$
$$P(A_1 \cap A_2 \cap \cdots \cap A_N) = P(A_1) \cdot P(A_2) \cdots P(A_N)$$

Question: How many conditions are there for N events $A_1, A_2, \dots, A_N$ to be statistically independent?

Answer: The total number M of conditions is as follows:

$$\begin{aligned} M = \binom{N}{2} + \binom{N}{3} + \cdots + \binom{N}{N} &= \sum_{k=0}^{N} \binom{N}{k} - \binom{N}{0} - \binom{N}{1} \\ &= 2^N - 1 - N \end{aligned}$$

(cf) Binomial expansion:

$$(1+x)^N = \sum_{k=0}^{N} \binom{N}{k} 1^k \cdot x^{N-k}$$

Let $x = 1$ in the above equation, then we have;

$$2^N = \sum_{k=0}^{N} \binom{N}{k}$$

e.g.

(i) In the case of two events, $N = 2$ and $M = 1$.

(ii) In the case of three events, $N = 3$ and $M = 4$.

2.6 Theory of Counting(Combinatorial Analysis)

1. Fundamental principle of counting:

If event A can occur in n ways, and event B can occur in m ways, then

1. Event $A \times B$ [14] can occur in $n \cdot m$ ways.
2. If event A and event B are *mutually exclusive* ot *disjoint*, then event A or B, (i.e. $A \cap B$) can occur only in $n + m$ ways.

2. Permutations:

1. An arrangement of n objects in a given *order* is called the permutation of n objects:

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 3 \cdot 2 \cdot 1$$

2. An arrangement of any $r(\leq n)$ objects out of n objects in a given *order* is called the permutation of n objects taken r at a time (or r objects out of n): [15]

$$P_r^n = \frac{n!}{(n-r)!}$$

3. Combinations:

The number of ways of selecting r objects from a lot of n objects where *order* does not count: [16] [17]

$$C_r^n = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

[14]This is an event in the combined experiments, which will be discussed in the next section.

[15]Note that $P_r^n = n(n-1)(n-2)\cdots(n-r+1)$, and thus $n! = n \cdot (n-1) \cdot (n-2) \cdots 3 \cdot 2 \cdot 1 = P_r^n \cdot (n-r)(n-r-1)\cdots 321$. Therefore, $P_r^n = \frac{n!}{(n-r)!}$.

[16]The number of ways for electing r objects from n objects and arranging them in *order* is $C_r^n \cdot r! \equiv P_r^n$. Therefore, $C_r^n = \frac{P_r^n}{r!}$.

[17]Since the order does not count, and there are $r!$ different ways of arranging r objects in order, C_r^n must be P_r^n divided by $r!$.

4. Sampling:

Figure 2.21: A box containing n objects.

1. The number of ways N for ampling of n objects taken r at a time (or r objects out of n) *with replacement*:

$$N = \underbrace{n \times n \times \cdots \times n}_{r} = n^r$$

2. The number of ways $N^{'}$ for sampling *without replacement*:

$$N^{'} = n \times (n-1) \times \cdots\cdots \times (n-r+1) = P_r^n$$

Example 2.11

Consider an experiment of throwing a fair coin 3 times, where the sample space S of each experiment is as follows:

$$S = \{H, T\}$$

Then, we have:

(i) Total number of ways(or events) that can occur $= 2^3 = 8$.

(ii) The number of ways (or events) that exactly two heads occur $= C_2^3 = 3$. [18]

(iii) Thus, the probability of having exactly two heads $= P(\text{ 2 heads}) = \frac{3}{8}$.

[18] Notice that the order does not count in this case, and thus corresponding 3 events are respectively $\{H, H, T\}$, $\{T, H, H\}$, $\{H, T, H\}$.

Example 2.12

Three balls, each of which is colored red, white, and blue, are in a box.

Figure 2.22: A box containing a red, a white, and a blue balls.

$$S = \{R, W, B\}$$

Then, what is the total number of different ways for the following experiments?

(i) Take two balls successively one at a time with replacement: $N = 3^2 = 9.$

(ii) Take two balls*in order* without replacement: $N = P_2^3 = \frac{3!}{1!} = 6.$

(iii) Take two balls without replacement: $N = C_2^3 = 3.$

2.7 Product Space (Order Space)

Definition 2.10 The product space S formed by spaces S_1 and S_2 is defined as a set of pairs from each space, and denoted by:

$$S = S_1 \times S_2 \triangleq \{(x, y) \mid x \in S_1 \text{ and } y \in S_2\}$$

Example 2.13

xy-plane is the product spcae of x-axis and the y-axis:

Figure 2.23: The xy-plane as a product space.

product space:

$(X - \text{space}) \times (Y - \text{space}) = (XY - \text{space})$, i.e.;

$(x - \text{axis}) \times (y - \text{axis}) = (xy - \text{plane})$

Note:
It is often called a "order space" since the order is important $\ni$:

$$S_1 \times S - 2 \neq S_2 \times S_1$$

(e.g.) Let two spaces S_1 and S_2 as follows

$$S_1 = \{Y_1 \mid a \leq Y_1 \leq b\}; \quad \text{set of real numbers } [a, b]$$

$$S_2 = \{Y_2 \mid c \leq Y_2 \leq d\}; \quad \text{set of real numbers } [c, d]$$

Figure 2.24: The product spaces $S_1 \times S_2$ and $S_2 \times S_1$.

Example 2.14

Consider the twp samples spaces given below:

$$S_1 = \text{sample space } \#1 \triangleq \{\omega_1, \omega_2, \omega_3\}$$

$$S_2 = \text{sample space } \#2 \triangleq \{y_1, y_2\}$$

Then the product space formed by S_1 and S_2 is :

$$\begin{aligned} S = S_1 \times S_2 &= \{(\omega_i, y_j)\}_{i=1,2,3\ j=1,2,} \\ &= \{(\omega_1, y_1), (\omega_1, y_2), (\omega_2, y_1), (\omega_2, y_2), (\omega_3, y_1), (\omega_3, y_2)\} \end{aligned}$$

Application to probability space:

The concept of product space is now being applied to the probability space of combined experiments, where we do more that one experiments sequentially. As we mentioned in chapter 1, we need three basic components for this combined experiment as well in order to form a probability space $(S, \mathcal{F}, P)$, which are described one by one below:

(1) Sample space (S):

Definition 2.11 Let S_1 be the sample space of subexperiment #1, and let S_2 be the sample space of subexperiment #2. Then, the sample space of the *combined experiment* S is defined as follows:

$$S = S_1 \times S_2 \triangleq \{(\omega_1, \omega_2) \mid \omega_1 \in S_1 \text{ and } \omega_2 \in S_2\}$$

Example 2.15

Consider the two sample spaces given below:

$$S_1 = \text{sample space for tossing a coin}$$

$$S_2 = \text{sample space for rolling a die}$$

i.e.

$$S_1 = \{H, T\}$$

$$S_2 = \{1, 2, 3, 4, 5, 6\}$$

Then, the samples space S of the combined experiment (i.e tosiing a coin, and THEN rolling a die) is given as follows:

$$S = S_1 \times S_2 = \{(H, 1), (H, 2), \ldots, (H, 6), (T, 1), (T, 2), \ldots, (T, 6)\}$$

Product events defined on a product space:

Consider two events A and B each of which is defined on the sample space S_1 and S_2 respectively as in the figure below:

Figure 2.25: An event defined on the product space.

Then, an event C on S, formed by the event A on S_1 and the event B on S_2, is defined as follows:

$$\begin{aligned} C = A \times B &\triangleq \{(\omega_1, \omega_2) \mid \omega_1 \in A \text{ and } \omega_2 \in B\} \\ &= (A \times S_2) \cap (S_1 \times B) \end{aligned}$$

Remark: Notice that:

(i) Event A on $S_1 \equiv$ event $A \times S_2$ on S.

(ii) Event B on $S_2 \equiv$ event $S_1 \times B$ on S.

(2) Field ($\mathcal{F}$):

Definition 2.12 The field $\mathcal{F}$ of the product space $S = S_1 \times S_2$ is the set (or collection) of subsets in S such that:

(i) If $A \times B \in \mathcal{F}$, then $\overline{A \times B} \in \mathcal{F}$.

(ii) If $A_1 \times B_1 \in \mathcal{F}$ and $A_2 \times B_2 \in \mathcal{F}$, then $(A_1 \times B_1) \cap (A_2 \times B_2) \in \mathcal{F}$

and we denote it by:

$$\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2$$

where $\mathcal{F}_1$ is the field of S_1, and $\mathcal{F}_2$ is the field of S_2 respectively.

(3) Probability (P):

Let:

1) A set function P be defined over $\mathcal{F}$, i.e. $P(A)$ where $A \in \mathcal{F}$.

2) A set function P_1 be defined over $\mathcal{F}_\infty$, i.e. $P(A_1)$ where $A_1 \in \mathcal{F}_1$.

3) A set function P_2 be defined over $\mathcal{F}_\in$, i.e. $P(A_2)$ where $A_2 \in \mathcal{F}_2$.

Then, since we have;

$$\text{event } A \text{ defined on } S_1 \equiv \text{event } (A \times S_2) \text{ defined on } S$$

$$\text{event } B \text{ defined on } S_2 \equiv \text{event } (S_1 \times B) \text{ defined on } S$$

$$\text{event } A \times B \text{ defined on } S \equiv \text{event } (A \times S_2) \cap (S_1 \times B) \text{ defined on } S$$

we have the following probabilities for the events:

$$P_1(A) = P(A \times S_2) \tag{2.3}$$

$$P_2(B) = P(S_1 \times B) \tag{2.4}$$

$$P(A \times B) = P\left[(A \times S_2) \cap (S_1 \times B)\right] \tag{2.5}$$

NOTE:
For a special case that subexperiment #1 is independent of subexperiment #2, we can re-write (2.5) using (2.3) and (2.4) as follows:

$$P\left(A \times B\right) = P_1\left(A\right) \cdot P_2\left(B\right)$$

Remarks: Importance of the above result is:

In principle, if we want to compute the probabilities for combined experiments, we first have to form a measureable space $(S, \mathcal{F})$ of that combined experiment, i.e. $S_1 \times S_2$ and $\mathcal{F}_1 \times calF_2$. We then must define a probability function $P(\cdot)$, which can be applied to that measurable space, and the probablity of a specific event $A \times B$ is calculated via the function $P(\cdot)$.
However.....
if the subexperiments are *independent*, we do not need to form the probability space $S, \mathcal{F}, P)$ for the combined experiment. All we have to do is to compute the probability of the event A on the probability space $(S_1, \mathcal{F}_1, P_1)$, i.e. $P_1(A)$, and the probability of the event B on $(S_2, \mathcal{F}_2, P_2)$, i.e. $P_2(B)$, and multiply them to get $P(A \times B) = P_1(A)P_2(B)$.

This provides us with much easier way of computing probabilities on a product space!!!

(e.g.)

In a combined experiment of tossing a coin and rolling a die sequentially, where

$$S_1 = \{H, T\}$$

$$S_2 = \{1, 2, 3, 4, 5, 6\}$$

The probability of getting a "head" and a number "1" can directly be computed as:

$$P\left[(H, 1)\right] = P_1(H) \cdot P_2(1) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}$$

This is done without forming the probability space of the combined experiment, and that is possible since two experiments "tossing coin" and "rolling a die" have nothing to do with each other, that is they are *independent*.

Extension:

In general, the product space formed by spaces $S_1, S_2, \ldots, S_N$ is denoted and defined as follows:

$$\begin{aligned} S &= S_! \times S_2 \times \cdots \times S_N \\ &\triangleq \{(\omega_1, \omega_2, \ldots, \omega_N) \mid \omega_i \in S_i \quad i = 1, 2, 3, \ldots, N\} \end{aligned}$$

where S is the sample space for the combined experiment, and S_i's are the sample space of the subexperiment $\#i,\ i = 1, 2, \ldots, N$.

Bernoulli Trials: [19]

Problem statement:
For an experiment for which $\exists$ only two possible outcomes $\{A, \overline{A}\}$ on any trial, we try N successive experiments, and determine the probability of the event $\ni$: outcome A is observed exactly k times out of N trials.

$\equiv$ **combination of N identical subexperiments** (e.g. coin tossing etc.)

$$S = S_1 \times S_2 \times \cdots \times S_N \triangleq \times_{i=1}^{N} S_i$$

Derivation:

Suppose events $\{A$ and/or $\overline{A}\}$ are statistically independent for every trial, i.e. we assume *independent experiments* [20]
Let the probability of each outcome is the same for every trial as:

$$P(A) = p, \quad \text{and} \quad P(\overline{A}) = 1 - p$$

As a specific event of getting the outcome A k times out of N trials, consider the event:

$$\underbrace{AA\cdots A}_{k}\underbrace{\overline{A}\overline{A}\cdots\overline{A}}_{N-k}$$

The the probability of above specific event is as follows:

$$\begin{aligned} P\left(\underbrace{AA\cdots A}_{k}\underbrace{\overline{A}\overline{A}\cdots\overline{A}}_{N-k}\right) &= \underbrace{p \cdot p \cdot \cdots \cdot p}_{k}\underbrace{(1-p)\cdot(1-p)\cdots(1-p)}_{N-k} \\ &= p^k \cdot (1-p)^{N-k} \end{aligned}$$

Since we have C_k^N different ways of getting such an event as given above, and since they are all mutually exclusive, we can add up each probability C_k^N times to get:

$$P\left(A\text{occus exactly } k \text{ times}\right) = C_k^N \cdot p^k \cdot (1-p)^{N-k}$$

[19] This is a typical example of a product space formed by independent sub-spaces.
[20] Be careful that this does not mean that A and $\overline{A}$ are independent, which cannot be true.

Example 2.16

A manufacturing process produces parts which are 10% defective. 10 of the parts are selected at random.

Then, assuming that the defectiveness of each part is independent of each other,

(a) What is the probability that there $\exists$ 2 or less defective parts? (event A)

(b) What is the probability that there $\exists$ 9 defective parts out of 10? (event B)

Solution:

Let D denote the being a "defective" part, and G denote the being a "good" part. Then, we have:

Figure 2.26: The sample space of testing quality of a manufactured item.

$$P(D) = \frac{1}{10} = p$$

$$P(G) = \frac{9}{10} = 1 - p$$

(a) $P(A)$:

$$\begin{aligned} P(A) &= P_0 + P_1 + P_2 \\ &= \sum_{k=0}^{2} C_k^{10} p^k (1-p)^{10-k} \\ &= C_0^{10} 0.1^0 0.9^{10} + C_1^{10} 0.1^1 0.9^9 + C_2^{10} 0.1^2 0.9^8 \\ &= 0.9^{10} + 0.9^9 + \frac{45}{100} \cdot 0.9^8 \\ &\approx 0.93 \end{aligned}$$

where P_i denotes the probability that there $\exists$ exactly i defective parts out of 10.

(b) $P(B)$:

$$\begin{aligned} P(B) &= C_9^{10} 0.1^9 0.9^1 \\ &= 10 \times 10^{-9} \times 0.9 \\ &= 0.9 \times 10^{-9} \\ &\approx 0 \end{aligned}$$