4.C. Definition of Probability

Probability is a measure of the likelihood that something may happen.

In physics, it is associated with random experiments in which repeated observations under identical settings do not produce identical results. The classic example is the repeated throwing of a dice.

Note: for the sake of clarity, errors due to imprecision of the measuring equipments are assumed negligible in the following discussion. However, in the theory of error analysis, such errors are treated as results of "random experiments".

In the deterministic Newtonian physics, randomness is caused by "hidden" variables that the experimenter either neglected or could not control. In quantum physics, randomness is built-in.

There are many ways to define probability. Two of the best known first attempts are:

1. Classical (or a priori) approach.

   Consider a random experiment that can result in n different but equally likely outcomes.
   If m of these outcomes correspond to an event A, then the probability of A is

   $$P(A) = \frac{m}{n}$$

2. Frequency (or a posteriori) approach.

   If, in N measurements of a random experiment, event A was observed to occur M times, then the probability of A is defined as

   $$P(A) = \frac{M}{N} \qquad \text{if} \qquad N > N_0$$

   where N_0 is some threshold number.

Though still used in most applications, both approaches are deemed to lack conceptual vigor. The 1st because of the vagueness of "equally likely". The 2nd because of that of N_0 .

The remedy is the axiomatic approach based on the theories of set and measure.

To begin, the set S of all possible outcomes of a random experiment is called the sample space. Each outcome is called a sample point. S is the universal set of our problem.

An event A is any subset of S. If the sample point of an observed outcome is an element of A, then the event A has occurred. The set, or more generally class, C of all A is therefore the power set of S.

Two events A & B are called mutually exclusive if

$$A \cap B = \Phi$$

where Φ is the empty set.

To each event A is associated a real number $P(A)$, where P is called a probability function that satisfies the following axioms.

Axiom 1. $P(A) \geq 0 \qquad \forall\ A$

Axiom 2. $P(S) = 1$

Axiom 3. If events A & B are mutually exclusive, then

$$P(A \cup B) = P(A) + P(B) \qquad (4.2)$$

which is just the rule of addition.

Axioms 1 & 3 imply

$$A \subset B \quad \Leftrightarrow \quad P(A) < P(B)$$

so that $P(A)$ is proportional to the "size" of A.

Let $B \subset A$, then

$$A = \left(A \backslash B \right) \cup B \qquad (4.2a)$$

where $A \backslash B$ is the complement of B with respect to A, i.e., $A \backslash B$ contains all elements of A not in B. Putting (4.2a) into (4.2) gives

$$P(A) = P\left(A \backslash B \right) + P(B)$$

$$\rightarrow \quad P\left(A \backslash B \right) = P(A) - P(B) \qquad (4.2b)$$

Combining axioms 2 & 3, we have

$$P(S \cup \Phi) = P(S) + P(\Phi)$$

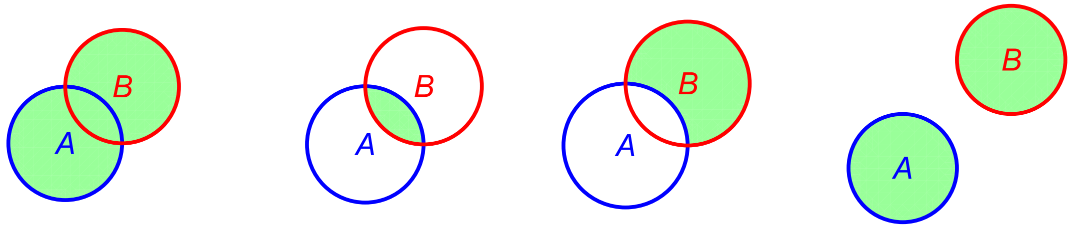
$$= P(S) \quad \rightarrow \quad P(\Phi) = 0 \qquad (4.2c)$$

If A & B are not mutually exclusive, then Fig.4.1a&c show that

$$A \cup B = A \cup \left[B \backslash (A \cap B) \right]$$

$$\rightarrow \quad P(A \cup B) = P(A) + P\left[B \backslash (A \cap B) \right] \qquad \text{[(4.2) used.]}$$

$$= P(A) + P(B) - P(A \cap B) \qquad \text{[(4.2a) used.]} \qquad (4.1)$$



(a) (b) (c) (d)

Fig.4.1. Green areas show (a) $A \cup B$, (b) $A \cap B$, (c) $B \backslash (A \cap B)$ and (d) $A \cup B$.
In (d), A & B are mutually exclusive so that $A \cap B = \Phi$.

A set $\{A_j,\ j = 1,\ ..,\ m\}$ of mutually exclusive events that satisfies

$$A_1 \cup ... \cup A_m = S$$

is called exhaustive. (4.2) then gives

$$P(A_1) + ... + P(A_m) = 1 \qquad (4.3)$$

Two events A & B are called independent if and only if

$$P(A \cap B) = P(A)\, P(B) \neq 0 \qquad (4.4)$$

The conditional probability is defined as

$P(B \mid A) \equiv$ probability of A provided B has already occurred. (4.5a)

Applying the rule of multiplication, the probability that both events have occurred is

$$P(A \cap B) = P(B)\, P(B \mid A) = P(A)\, P(A \mid B) \qquad (4.6)$$

$$\rightarrow \quad P(B \mid A) = \frac{P(A \cap B)}{P(B)} \qquad P(A \mid B) = \frac{P(A \cap B)}{P(A)} \tag{4.5}$$

In case A & B are independent, (4.5) reduces to

$$P(B \mid A) = P(A)\,P(B) = P(A \mid B) \tag{4.7}$$

Ex.4.2.

Consider a sample space of two events A & B. Given

$$P(A) = \frac{3}{5} \qquad P(B) = \frac{2}{3} \qquad P(A \cup B) = 1 \tag{1}$$

compute $P(A \cap B)$, $P(B \mid A)$ & $P(A \mid B)$.

Answer

(4.1) gives

$$\begin{aligned} P(A \cap B) &= P(A) + P(B) - P(A \cup B) \\ &= \frac{3}{5} + \frac{2}{3} - 1 = \frac{4}{15} \end{aligned}$$

which means A & B are not mutually exclusive.

Since

$$P(A)\,P(B) = \frac{3}{5} \times \frac{2}{3} = \frac{2}{5} \neq P(A \cap B)$$

A & B are also not independent.

Using (4.5), we have

$$P(B \mid A) = \frac{P(A \cap B)}{P(B)} = \frac{4}{15} \Big/ \frac{2}{3} = \frac{2}{5}$$

$$P(A \mid B) = \frac{P(A \cap B)}{P(A)} = \frac{4}{15} \Big/ \frac{3}{5} = \frac{4}{9}$$

Thus, $\frac{4}{9}$ of the points in A lies in $A \cap B$, and hence also belong to B. Similarly, $\frac{2}{5}$ of the points in B lies in $A \cap B$, and hence also belong to A.

Code

```
In[∘]:= disk[{x0_, y0_}, r_] := (x - x0)^2 + (y - y0)^2 ≤ r^2

In[∘]:= grc = Graphics[{Thick,
        Blue, Circle[-.5 {1, 1}, 1], Text["A", -.5 {1, 1}],
        Red, Circle[.5 {1, 1}, 1], Text["B", .5 {1, 1}]
       }];

In[∘]:= L = 1.1 √2 ;
```

In[∘]:=
```
(* A ⋃ B *)
lu = RegionPlot[{ disk[-.5 {1, 1}, 1] || disk[.5 {1, 1}, 1]}, {x, -L, L}, {y, -L, L},
   PlotStyle → {Opacity[.4], Green}, BoundaryStyle → None,
   Frame → None
  ];
Show[{lu, grc}]
```

In[∘]:=
```
(* A ⋂ B *)
li = RegionPlot[{ disk[-.5 {1, 1}, 1] && disk[.5 {1, 1}, 1]}, {x, -L, L},
   {y, -L, L}, PlotStyle → {Opacity[.4], Green}, BoundaryStyle → None,
   Frame → None
  ];
Show[{li, grc}]
```

In[∘]:=
```
(* B\A *)
lm = RegionPlot[{ ! disk[-.5 {1, 1}, 1] && disk[.5 {1, 1}, 1]}, {x, -L, L}, {y, -L, L},
   PlotStyle → {Opacity[.4], Green}, BoundaryStyle → None,
   Frame → None
  ];
Show[{lm, grc}]
```

In[∘]:=
```
grcX = Graphics[{Thick,
    Blue, Circle[-1 {1, 1}, 1], Text["A", -{1, 1}],
    Red, Circle[{1, 1}, 1], Text["B", {1, 1}]
   }];
```

In[∘]:=
```
LX = 1.5 √2 ;
luX = RegionPlot[{ disk[-1 {1, 1}, 1] || disk[{1, 1}, 1]}, {x, -LX, LX}, {y, -LX, LX},
   PlotStyle → {Opacity[.4], Green}, BoundaryStyle → None,
   Frame → None
  ];
Show[{luX, grcX}]
```