Chapter 1 Finite Trials

A finite trial is about a test that has only a finite number of different outcomes. To explain the rudimentary concepts of probability theory, the main object of this chapter is restricted to finite trial. In studying finite trials, it is sufficient to learn finite sets and finite union formula. Moreover, if one could grasp the idea of probability theory by finite trials, then with the methods of analysis, it is easy to proceed to infinite trial, which is the object of modern probability theory.

§ 1.1 Probability Space

In the trial of examining the point of a rolled dice, there are six possibilities of 1, 2, 3, 4, 5, and 6. In this case, 1, 2, . . ., 6 are known as the **sample points**, and the set of sample points $\{1, 2, 3, 4, 5, 6\}$ is called the **sample space**. One could regard the sample point and sample space by the notion of random trial. In accordance to whether the sample space of a certain trial is a finite set or an infinite set, the trial is said to be a **finite trial** or an **infinite trial**. This chapter considers only the finite trial, or simply the trial.

Let Ω be the sample space of a trial $\boldsymbol{T}$. When A is a random subset of Ω, "sample points belonging to A appear" as a result of trial $\boldsymbol{T}$ is simply said as A **occurs**. In this sense, A as the subset of Ω is also known as **event** A.

When the compliment set A^c of A occurs is the same as saying that when A does not occur. In this aspect, A^c is said to be the **compliment event** of A. Since the union set of A and B $A \cup B$ is the same as saying that of the two events A and B at least one of them occurs, $A \cup B$ is called the **union event**. As the intersection set $A \cap B$ of A and B is the same as saying that both A and B occur, $A \cap B$ is called the **intersection event**. When the difference $A \backslash B$ occurs, it the same as saying that A occurs and B does not occur; $A \backslash B$ is called the **difference event**.

What "A ⊂ B" means is that "if event A occurs, then event B must also occur." In this instance, $B \backslash A$ is known as the **eigen difference**, and denoted by $B - A$. Henceforth, when $B - A$ is written, $A \subset B$ is implicitly assumed. $A \cap B = \emptyset$ i.e. A, B are mutually disjoint is the same as saying that events A, B do not occur at the same time (**exclusive events**). In this case, $A \cup B$ is called the **direct union**, and denoted by $A + B$. Analogous to the implicit assumption of eigen difference, when $A + B$ is written, it is implicitly assumed that A, B are mutually disjoint (i.e. A, B are mutually exclusive events). The same applies to a direct sum of more than two sets:

$$A_1 + A_2 + \cdots + A_n \qquad \left(\text{or expressed as } \sum_{i=1}^{n} A_i\right)$$

As mentioned above, the sample space Ω of a rolled dice is given by

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

and by the same token, the probability $P(A)$ of event A occurring is given by

$$P(A) = \frac{\sharp A}{6} \qquad (\sharp A = \text{the number of points of } A)\,.$$

Let $\boldsymbol{T}$ be a generic trial, and Ω its sample space. With respect to an arbitrary $A \subset \Omega$, the probability of A occurring is written as $P(A)$. Now $P(A)$ is a function of set A. This set function is said to be the **probability law** of trial T. From the intuitive meaning of probability, of course, $P(A)$ comes with the following properties.

(P.1) $\qquad P(A) \geq 0\,,$

(P.2) (**additive**) $\qquad P(A + B) = P(A) + P(B)\,,$

(P.3) $\qquad P(\Omega) = 1\,.$

In general, the set function P on finite set Ω possessing these properties is called the **probability measure**, and Ω that comes with P is called the **probability space** (Ω, P).

From the above observation, trial $\boldsymbol{T}$'s probability law P is the probability measure on $\boldsymbol{T}$'s sample space Ω. To the sample space of trial $\boldsymbol{T}$, the probability law P of $\boldsymbol{T}$ is given, and (Ω, P) is said to be the probability space of $\boldsymbol{T}$. As the probability space of $\boldsymbol{T}$ exhaustingly exhibits all mathematical aspects of $\boldsymbol{T}$, to research on $\boldsymbol{T}$ is nothing but to examine (Ω, P).

Theorem 1.1 Let P be the probability measure on Ω.

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

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

(iii) $P(A^c) = 1 - P(A)$,

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

(v) $P(A) = \sum_{\omega \in A} P\{\omega\}$. (Hence, for all $\omega \in \Omega$, if $P\{\omega\}$ is given, then P is completely determined)

Proof (i) Consider (P.2) and use the method of induction.

(ii) As $B \supset A$ is implicitly assumed,

$$B = (B - A) + A\,.$$

Apply (P.2) to this expression.

(iii) Let $B = \Omega$ in (ii).

(iv) Let $C = A \cap B$, $A_1 = A - C$, $B_1 = B - C$.

$$A = A_1 + C\,, \qquad B = B_1 + C\,, \qquad A \cup B = A_1 + B_1 + C$$

Then apply (i).

(v) To $A = \sum_{a \in A} \{a\}$, apply (i). ∎

Though an event is described by a set, sometimes it is expressed as the condition $\alpha(\omega)$ concerning the sample point ω. Since the sample points that fulfill condition $\alpha(\omega)$ are the result of a trial, it is said that α **occurs**. In this sense, condition $\alpha(\omega)$ is also called event.

Let $A = \{\omega | \alpha(\omega)\}$. α occurs is the same as A occurs. It follows that the probability of α occurring is equal to $P\{\omega | \alpha(\omega)\}$.

The negation condition of α is denoted by $\alpha^{\neg}$, and $\alpha^{\neg}$ is the compliment event of α. Condition 'α or β' is denoted by $\alpha \vee \beta$. Since

$$\{\omega | \alpha(\omega) \vee \beta(\omega)\} = \{\omega | \alpha(\omega)\} \cup \{\omega | \beta(\omega)\}\,,$$

$\alpha \vee \beta$ is the disjoint union event of α and β. Similarly, condition 'α and β' is denoted by $\alpha \wedge \beta$, and it is the disjoint intersection event of α and β.

Exercise 1.1 With respect to probability measure P, prove the following **inclusion-exclusion formula**:

(i) $$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{i_i < i_2 < \cdots < i_k} P\left(\bigcap_{\kappa=1}^{k} A_{i_\kappa}\right)$$

(ii) the formula resulting from the exchange of $\bigcup$ with $\bigcap$ in the above formula.

[Hint] Consider (iv) of Theorem 1.1 and use the induction method with respect to n.

Christopher Ting
Translation Version 1.0
Date: 10 May, 2017