Again, it can be observed that Table 4 is an alternative form of Table 6.

The readers may try to prove the *duals* of the above two examples, namely,

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

and

$$(A \cap B)' = A' \cup B',$$

which are obtained by interchanging $\cup$ and $\cap$.

****

When G.M. Hardy visited S. Ramanujan he told him that the number of the taxi in which he came, 1729, looked rather unattractive. Ramanujan immediately denied this, saying that it was the least number which could be expressed as the sum of two cubes in two different ways ; that is, $1729 = 12^3 + 1^3 = 10^3 + 9^3$.

oooooooooooooooooooooooooooooooo

EQUAL LIKELIHOOD AND INDEPENDENCE

Louis H. Y. Chen
Department of Mathematics
University of Singapore

1. *Introduction.* The notion of equal likelihood is, in some sense, closely tied with that of independence in the Theory of Probability. Often this fact is either overlooked or insufficiently emphasized in the teaching of elementary courses. This is perhaps due to the difficulty in making the relation between the two notions precise at the level concerned. However, a good understanding of the relation is necessary in order to explain the equivalence of two methods which are often employed in the solutions of a large number of elementary problems.

To illustrate the last point, consider the following simple example: A fair coin is tossed, a fair die is rolled and a ball is drawn at random from an urn containing 2 black and 8 white balls. What is the probability of the event that the coin falls heads, an even number appears on the die and a white ball is drawn? One method of solution is as follows: Since the sample space consists of $2\times6\times10 = 120$ equally likely outcomes (each of which can be represented by a triple such as (Head,5,black ball)) and the event consists of $1\times3\times8 = 24$ outcomes, it follows that the probability is $24/120 = 1/5$. Alternatively, one could first calculate the probabilities of the coin falling heads, of an even number appearing on the die and of a white ball being drawn, which are 1/2, 1/2 and 4/5 respectively. Then one could use independence to conclude that the probability of the event is $(1/2)\times(1/2)\times(4/5) = 1/5$. The two solutions just discussed employ different concepts. It is not immediately clear that they should be equivalent. Moreover, what is the underlying sample space in the second solution?

The following is an attempt to answer these questions in elementary terms. Some basic knowledge of Probability will be assumed. The relation between equal likelihood and independence will be stated in two forms which will be proved as two theorems. In order to show that the usefulness of the theorems is not confined to the present discussion, one of them will be applied to obtain a simpler proof of an interesting lemma in [1], pp. 192 - 193.

2. *Equal likelihood and independence.* Very often, the experiment of interest consists of a number of other experiments. Such an experiment is a compound experiment. In the above example, the compound experiment consists of three experiments, namely tossing a coin, rolling a die and drawing a ball from an urn. Let $\Omega_1,\ldots,\Omega_n$ be the sample spaces of n experiments, say $\mathcal{E}_1,\ldots,\mathcal{E}_n$. The natural sample space of the compound experiment consisting of these experiments (denoted by $\mathcal{E}_1\times\ldots\times\mathcal{E}_n$) is the Cartesian product $\Omega_1\times\ldots\times\Omega_n = \{(\omega_1,\ldots,\omega_n) = \omega_i \varepsilon\ \Omega_i,\ i=1,\ldots,n\}$. An event in the compound experiment $\mathcal{E}_1\times\ldots\times\mathcal{E}_n$, which is of the form $\Omega_1\times\ldots\times A_i\times\ldots\times\Omega_n$, where $A_i \subseteq \Omega_i$, is called an event referring to the i_{th} experiment $\mathcal{E}_i$, $i=1,\ldots,n$. It is said to occur if and only if A_i occurs in the i^{th} experiment $\mathcal{E}_i$. In the present discussion, all sample spaces are assumed to be finite, i.e. all experiments are assumed to have a finite number of outcomes.

Let P_i be the probability associated with the experiment $\mathcal{E}_i$, $i=1,\ldots,n$. The question arises as to whether there exists a probability P associated with the compound experiment $\mathcal{E}_1\times\ldots\times\mathcal{E}_n$, which has the following properties:

$$P(\widetilde{A}_i) = P_i(A_i),\quad A_i \subseteq \Omega_i,\quad i = 1,\ldots,n, \tag{2.1}$$

and

$$P\Big(\bigcap_{i=1}^{n} \widetilde{A}_i\Big) = \prod_{i=1}^{n} P(\widetilde{A}_i),\quad A_i \subseteq \Omega_i,\ i=1,\ldots,n, \tag{2.2}$$

where $\widetilde{A}_i = \Omega_1\times\ldots\times A_i\times\ldots\times\Omega_n$, $i = 1,\ldots,n$.

The property (2.1) is essential in order that P is meaningful, whereas the property (2.2) stipulates the independence of $\tilde{A}_1, \ldots, \tilde{A}_n$ for any $A_i \subseteq \Omega_i$, i=1,...,n . In a loose sense, (2.2) says that the experiments $\mathcal{E}_1, \ldots, \mathcal{E}_n$ are independent. Noting that $\bigcap_{i=1}^{n} \tilde{A}_i = A_1 \times \ldots \times A_n$ and that $\{(\omega_1, \ldots, \omega_n)\} = \{\omega_1\} \times \ldots \times \{\omega_n\}$, one immediately sees that such a probability P does exist. Indeed, it is given by

$$(2.3) \quad P(\{(\omega_1, \ldots \omega_n)\}) = \prod_{i=1}^{n} P_i(\{\omega_i\}), \ (\omega_1, \ldots, \omega_n) \in \Omega_1 \times \ldots \times \Omega_n .$$

Furthermore, it is unique. The reader is advised to verify that the probability P given by (2.3) is the only probability associated with $\mathcal{E}_1 \times \ldots \times \mathcal{E}_n$, which satisfies (2.1) and (2.2). The probability P given by (2.3) is called the product probability of $P_1, \ldots, P_n$ and is denoted by $P_1 \times \ldots \times P_n$.

A probability is said to be uniform if it attributes equal probabilities to all the outcomes in the associated sample space, i.e. all outcomes are equally likely. Similarly, a random variable or a random vector is said to have a uniform distribution, if all its values have equal probabilities. The theorems are now stated and proved as follows:

Theorem 2.1. Let $\Omega_1, \ldots, \Omega_n$ be sample spaces. If the probability P_i associated with Ω_i is uniform, i=1,...,n, then the product probability $P_1 \times \ldots \times P_n$ associated with $\Omega_1 \times \ldots \times \Omega_n$ is uniform. Conversely, if a probability P associated with $\Omega_1 \times \ldots \times \Omega_n$ is uniform, then there exist unique probabilities P_i' associated with Ω_i, i=1,...,n, such that $P = P_1' \times \ldots \times P_n'$. Moreover each P_i' must necessarily be uniform.

Proof. Let Ω_i consist of k_i outcomes, i=1,...,n. The uniformity of P_i implies that

$$P_i(\{\omega_i\}) = 1/k_i, \quad \omega_i \in \Omega_i, \quad i = 1, \ldots, n$$

Thus for every $(\omega_1, \ldots, \omega_n) \in \Omega_1 \times \ldots \times \Omega_n$, we have

$$\begin{aligned} &P_1 \times \ldots \times P_n (\{(\omega_1, \ldots, \omega_n)\}) \\ &= P_1 \times \ldots \times P_n (\{\omega_1\} \times \ldots \times \{\omega_n\}) \\ &= \prod_{i=1}^{n} P_i(\{\omega_i\}) = 1/k_1 \ldots k_n . \end{aligned}$$

This proves the uniformity of $P_1 \times \ldots \times P_n$.

Conversely, suppose P is uniform. Then for every $\omega_i \varepsilon \Omega_i$, we have

$$P(\Omega_1 \times \ldots \times \{\omega_i\} \times \ldots \times \Omega_n)$$
$$= \Sigma P(\{(\omega_1, \ldots, \omega_n)\})$$
$$= \Sigma 1/k_1 \ldots k_n = 1/k_i ,$$

where the summation is taken over the set $\Omega_1 \times \ldots \times \{\omega_1\} \times \ldots \times \Omega_n$ and $i=1,\ldots,n$. Define P_i' by

$$P_i'(\{\omega_i\}) = P(\Omega_1 \times \ldots \times \{\omega_i\} \times \ldots \times \Omega_n), \omega_i \varepsilon \Omega_i, \ i=1,\ldots,n.$$

Clearly P_i' is uniform. Now

$$P(A_1 \times \ldots \times A_n) = \Sigma \ P(\{(\omega_1, \ldots, \omega_n)\})$$
$$= \Sigma 1/k_1 \ldots k_n = \prod_{i=1}^{n} N(A_i)/k_i$$
$$= \prod_{i=1}^{n} P_i'(A_i),$$

where the summation is taken over the set $A_1 \ldots A_n$ and $N(A_i)$ denotes the number of elements in the set $A_i, i=1,\ldots,n$. Thus $P=P_1 \times \ldots \times P_n$. Let P_i'' be another probability associated with Ω_i, $i = 1,\ldots,n$, such that $P = P_1'' \times \ldots \times P_n''$. Then for every $A_i \subseteq \Omega_i$,

$$P_i'(A_i) = P(\Omega_1 \times \ldots \times A_i \times \ldots \times \Omega_n)$$
$$= P_1''(\Omega_1) \ldots P_i''(A_i) \ldots P_n''(\Omega_n)$$
$$= P_i''(A_i).$$

This proves the uniqueness of P_i'. Hence the theorem.

In a loose sense, Theorem 2.1 says that the outcomes of a compound experiment $\mathcal{E}_1 \times \ldots \times \mathcal{E}_n$ are equally likely if and only if the experiments $\mathcal{E}_1, \ldots, \mathcal{E}_n$ are independent and the outcomes of each $\mathcal{E}_i$ are equally likely. This property carries over to random variables. In fact, it will be more vividly exhibited in terms of random variables and will not depend on the structure of the underlying sample space. The next theorem illustrates this point.

Theorem 2.2.: Let $X_1, \dots, X_n$ be discrete random variables defined on a sample space Ω. The random vector $(X_1, \dots, X_n)$ has a uniform distribution if and only if $X_1, \dots, X_n$ are independent and each X_i has a uniform distribution.

The proof of this theorem is similar to and even simpler than that of the preceding theorem. It is therefore left to the reader.

3. *Applications.* In order to explain the equivalence of the two methods of solution in the above example in the proper context, it is necessary that the same sample space, namely $\Omega = \Omega_1 \times \dots \times \Omega_n$ (in this example, n=3) must be used in both cases. In the first method, equal likelihood of the outcomes in Ω is assumed. This is the same as assuming the probability associated with Ω, say P, to be uniform. Let P_i be the probability associated with Ω_i, $i=1,\dots,n$. In the second method, the probability associated with Ω is actually the product probability $P_1 \times \dots \times P_n$ with each P_i assumed to be uniform. By Theorem 2.1, $P = P_1 \times \dots \times P_n$. This proves the equivalence of the two methods.

A simpler proof of a lemma in [1], pp.192-193, will now be discussed. Let Ω be the sample space of all $n!$ distinct permutations $(a_1, \dots, a_n)$ of the integers $(1,\dots,n)$, where each permutation has probability $1/n!$. For each i, $i=1,\dots,n$, and each $\omega = (a_1, \dots, a_n) \in \Omega$, let X_i be the number of "inversions" caused by i in ω, i.e. $X_i(\omega) = m$ if and only if i precedes exactly m of the integers $1,\dots,i-1$ in the permutation ω. The lemma states that the random variables $X_1, \dots, X_n$ are independent and each X_i has a uniform distribution, i.e.

$$P(X_i = m) = 1/i, \quad 0 \leqslant m \leqslant i-1, \; i = 1,\dots,n.$$

This result is far from being obvious and is difficult to prove directly. But, in view of Theorem 2.2, one only needs to show that the random vector $(X_1, \dots, X_n)$ has a uniform distribution. Indeed, it is not difficult to see that the mapping defined by $(a_1, \dots, a_n) \longmapsto (X_1(\omega), \dots, X_n(\omega))$ for every $\omega = (a_1, \dots, a_n) \in \Omega$ is one-to-one and onto from Ω to $N_0 \times N_1 \times \dots \times N_{n-1}$, where $N_i = \{0, 1, \dots, i\}$, $i = 1,\dots,n-1$. Thus for every value $(c_1, \dots c_n)$ of $(X_1, \dots X_n)$, we have $P(X_1 = c_1, \dots, X_n = c_n) = 1/n!$.

This proves the lemma.

Reference

[1] Chung, Kai Lai, *A Course in Probability Theory*
Harcourt, Brace & World, Inc., New York, 1968.