ELEMENTARY PROBABILITY

Summary measures of a statistical experiment. Let us toss the die 30 times and let us record the value assumed by the random variable at each toss:

$$1, 2, 5, 3, \ldots, 4, 6, 2, 1.$$

To summarise this information, we may construct a *frequency table:*

x	f	r
1	8	8/30
2	7	7/30
3	5	5/30
4	5	5/30
6	3	3/30
6	2	2/30
	30	1

Here,

$$f_i = \text{frequency},$$

$$n = \sum f_i = \text{sample size},$$

$$r_i = \frac{f_i}{n} = \text{relative frequency}.$$

In this case, the order in which the numbers occur is of no interest; and, therefore, there is no loss of information in creating this summary.

We can describe the outcome of the experiment more economically by calculating various summary statistics.

First, there is the *mean* of the sample

$$\bar{x} = \frac{\sum x_i f_i}{n} = \sum x_i r_i.$$

The *variance* is a measure of the dispersion of the sample relative to the mean, and it is the average of the squared deviations. It is defined by

$$\begin{aligned} s^2 &= \sum \frac{(x_i - \bar{x})^2 f_i}{n} = \sum (x_i - \bar{x})^2 r_i \\ &= \sum (x_i^2 - x_i\bar{x} - \bar{x}x_i + \{\bar{x}\}^2) r_i \\ &= \sum x_i^2 r_i - \{\bar{x}\}^2, \end{aligned}$$

which follows since $\bar{x}$ is a constant that is not amenable to the averaging operation.

THE CONCEPT OF PROBABILITY

It is tempting to define the probabilities of the various outcomes as the limits of the corresponding empirical relative frequencies as the sample size n tends to infinity.

We may say that a sequence of numbers $\{r_i; i = 1, 2, 3, \ldots\}$ tends to a limit p if, for every number ϵ, be it ever so small, there exists a number $n = n(\epsilon)$, which is a function of ϵ, such that

$$|r_i - p| < \epsilon \qquad \text{for all} \quad i > n.$$

We may denote this by writing $\lim(n \to \infty) r_n = p$

However, this cannot serve in an definition of probability, since there is always a chance that, by a run of aberrant outcomes, the value of r_i will break the bounds of the neighbourhood $(p - \epsilon, p - \epsilon)$.

All that we can say is that probability of this happening becomes vanishingly small as $n \to \infty$. But, in making this statement, we have invoked a notion of probability, which is the very thing that we are trying to define. Clearly, such a limiting process cannot serve to define probabilities.

The consensus of probability theorists that dates from the 1930s is that, in general, there is no alternative but to take an axiomatic approach to probability, as we shall do in the following lectures. For the present, we shall consider probabilities that are based upon the classical notions of equally likely outcomes, such as the outcomes that characterise games of chance.

THE ELEMENTARY RULES OF PROBABILITY

The union of mutually exclusive events.

We might ask ourselves, what is the probability of getting either a 2 or a 4 or a 6 in a single toss of the die? We can denote the event in question by $A_2 \cup A_4 \cup A_6$. The corresponding probability is

$$P(A_2 \cup A_4 \cup A_6) = P(A_2) + P(A_4) + P(A_6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}.$$

The result depends on the following principle:

> *The probability of the union of mutually exclusive events is the sum of their separate probabilities.*

Observe that $\cup$ (union) means *OR* in the inclusive sense. That is to say, $A_1 \cup A_2$ means A_1 *OR* A_2 *OR BOTH*. In this case, *BOTH* is not possible, since the events are mutually exclusive.

By extending the argument, we can see that

$$P(A_1 \cup A_2 \cup \cdots \cup A_6) = 1,$$

which is to say that the probability of a certainty is unity.

Tossing a pair of die: various outcomes

For the red die, the outcomes are $\{A_i; i = 1, 2, \ldots, 6\}$ and the corresponding numbers $\{x_i; i = 1, \ldots, 6\} = \{1, 2, \ldots, 6\}$.

For the blue die, the outcomes are $\{B_j; i = 1, 2, \ldots, 6\}$ and the corresponding numbers $\{y_j; i = 1, \ldots, 6\} = \{1, 2, \ldots, 6\}$.

Also, we may define a further set of events $\{C_k; k = 1, 2, \ldots, 6\}$ and the corresponding numbers $\{z_{k=i+j} = x_i + y_j\} = \{2, 3, \ldots, 12\}$.

Table 1. The outcomes from tossing a red die and a blue die, highlighting the outcomes for which the joint store is 5.

	1	*2*	*3*	*4*	*5*	*6*
1	2	3	4	(**5**)	6	7
2	3	4	(**5**)	6	7	8
3	4	(**5**)	6	7	8	9
4	(**5**)	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

The intersection of events, or their joint or sequential occurrence.

We may ask what is the probability of the event C_5, that the combined score of the dice is 5. The event is defined by

$$C_5 = \{(A_1 \cap B_4) \cup (A_2 \cap B_3) \cup (A_3 \cap B_2) \cup (A_4 \cap B_1)\}.$$

This is a union of mutually exclusive events, so the first rule of probability gives

$$P(C_5) = P(A_1 \cap B_4) + P(A_2 \cap B_3) + P(A_3 \cap B_2) + P(A_4 \cap B_1)\},$$

But A_i and B_j are independent events for all i, j. Therefore,

$$P(A_i \cap B_j) = P(A_i) \times P(B_j) = \frac{1}{36} \quad \text{for all} \quad i, j;$$

and it follows that

$$P(C_5) = P(A_1)P(B_4) + P(A_2)P(B_3) + P(A_3)P(B_2) + P(A_4)P(B_1) = \frac{4}{36}.$$

The principle that has been invoked in solving this problem, which has provided the probabilities of the events $A_i \cap B_j$, is the following:

> *The probability of the joint occurrence of two statistically independent events is the product of their individual or marginal probabilities. Thus, if A and B are statistically independent, then $P(A \cap B) = P(A) \times P(B)$.*

Conditional probabilities.

Consider a set of events $\{D_k; k = 2, 3, \ldots, 12\}$ corresponding to $x+y = z \geq k \in \{2, 3, \ldots, 12\}$. For example, there is the event

$$D_8 = C_8 \cup C_9 \cup \cdots \cup C_{12},$$

whereby the joint score equals of exceeds eight, which has the probability

$$P(D_8) = P(C_8) + P(C_9) + \cdots + P(C_{12}) = \frac{15}{36}.$$

Table 2. The outcomes from tossing a red die and a blue die, highlighting the outcomes for which the score on the red die is 4, by the numbers enclosed by brackets [,] as well as the outcomes for which the joint score exceeds 7, by the numbers enclosed by parentheses (,).

	1	*2*	*3*	*4*	*5*	*6*
1	2	3	4	[**5**]	6	7
2	3	4	5	[**6**]	7	(**8**)
3	4	5	6	[**7**]	(**8**)	(**9**)
4	5	6	7	[(**8**)]	(**9**)	(**10**)
5	6	7	(**8**)	[(**9**)]	(**10**)	(**11**)
6	7	(**8**)	(**9**)	[(**10**)]	(**11**)	(**12**)

Conditional probabilities: continued

What is the value of the probability $P(D_8|A_4)$ that the event D_8 will occur when the event A_4 is already know to have occurred, which is probability that $x + y \geq 8$ given that $x = 4$?

The question concerns the event $D_8 \cap A_4$ for which the unconditional probability is

$$P(D_8 \cap A_4) = 3/36.$$

This event is to be considered not within the entire sample space S, but only within the context of A_4.

Since the occurrence of A_4 is now a certainty, and the probabilities of its constituent events, which are mutually exclusive and exhaustive, must sum to unity.

This can be achieved by re-scaling the probabilities in question, and the appropriate scaling factor is $1/P(A_4)$. Thus, the conditional probability is

$$P(D_8|A_4) = \frac{P(D_8 \cap A_4)}{P(A_4)} = \frac{3/36}{1/6} = \frac{1}{2}.$$

By this form of reasoning, we can arrive at the following law of probability:

> *The conditional probability of the occurrence of the event A given that the event B has occurred is given by the probability of their joint occurrence divided by the probability of B. Thus $P(A|B) = P(A \cap B)/P(B)$.*

The joint occurrence of non-exclusive events.

Consider the probability of the event $A_4 \cup D_8$, which is the probability of getting $x = 4$ for the score on the red die or of getting $x + y \geq 8$ for the joint score, or of getting both of these outcomes at the same time.

Since $A_4 \cap D_8 \neq \emptyset$, the law of probability concerning mutually exclusive outcomes, cannot be invoked directly. It would lead to the double counting of those events that are indicated in Table 2 by the cells bearing numbers that are surrounded both by brackets and by parentheses. Thus $P(A_4 \cup D_8) \neq P(A_4) + P(D_8)$.

The avoidance of double counting leads to the formula

$$\begin{aligned} P(A_4 \cup D_8) &= P(A_4) + P(D_8) - P(A_4 \cap D_8) \\ &= \frac{6}{36} + \frac{15}{36} - \frac{3}{36} = \frac{1}{2}. \end{aligned}$$

By this form of reasoning, we can arrive at the following law of probability:

> *The probability that either of events A and B will occur, or that both of them will occur, is equal to the sum of their separate probabilities less the probability or their joint occurrence. Thus*

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