The basics of probability – Part II

How to express probability distribution?

First let's define probability:

Let U denote the sample space for an experiment. A *probability distribution* on U is a function P which assigns a number, $P(\mathrm{A})$, between zero and one to any event $\mathrm{A} \subseteq \mathrm{U}$ such that:

$$P(\mathrm{A} \cup \mathrm{B}) = P(\mathrm{A}) + P(\mathrm{B}) \text{ if } \mathrm{A} \cap \mathrm{B} = \varnothing \qquad (1)$$

and so that $P(\mathrm{U}) = 1$.

Note:

A collection of events is *exhaustive* if, taken in totality, they account for all possible results or outcomes, i.e., $P(\mathrm{U}) = 1$.

Therefore, the probability distribution P describes how the total probability mass of 1 should be distributed on the sample space. That means the probability of an event cannot be less than 0 or greater than 1 (corresponding to 0% and 100%, respectively), and all probabilities (probability mass) add up to 1 (or 100%).

Formula (1) above, commonly known as the *addition rule* says that the probability of either event A or event B occurring is equal to the sum of the probabilities of the individual events if we assume that it is impossible to observe both A and B at the same time.

Example 2:

For the regular die throwing experiment, it is reasonable to think of all 6 possible outcomes as equally likely,

$$P(\{1\}) = P(\{2\}) = \ldots. = \mathrm{P}(\{6\})$$

and every throw is an independent event.

So, for an event A to have either pip number 1 or 6 to show up, i.e., $P(\mathrm{A}_1) = (\{1\})$ or $P(\mathrm{A}_2) = (\{6\})$, the probability distribution is given by

$$P(A) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

Joint and disjoint events – the basic probability rules

Let A and B be events from the sample space U. Then,

1. $P(\mathrm{A}) \leq P(\mathrm{B})$ if $\mathrm{A} \subseteq \mathrm{B}$ (2)

2. $P(\mathrm{A}^c) = 1 - P(\mathrm{A})$ (3)

3. $P(\mathrm{A} \cup \mathrm{B}) = P(\mathrm{A}) + \mathrm{P(B)} - P(\mathrm{A} \cap \mathrm{B})$ (for joint events) (4)

4. $P\,(\mathrm{A}_1 \cup \mathrm{A}_2 \cup \ldots \cup \mathrm{A}_k) = P\,(\mathrm{A}_1) + P\,(\mathrm{A}_2) + \ldots + P\,(\mathrm{A}_k)$
   if $\mathrm{A}_1, \mathrm{A}_2, \ldots, \mathrm{A}_k$ are pair-wise *disjoint* event,
   $\mathrm{A}_i \cap \mathrm{A}_j = \varnothing$ for all $i \neq j$. (5)

   Note:

   Any number of events are said to be disjoint or *mutually exclusive* if they have no overlap or commonality.

Example 3:

If we have two events: A = {1,3,5} and B = {4,5,6} and wish to calculate P(A ∪ B), we can do two ways:

a. We can calculate the probability by counting the elements. The union event A ∪ B = {1,3,4,5,6} has 5 elements and so the probability distribution yields the probability 5/6.

b. Alternatively, we can use the rule 3, which gives:

$$P(A \cup B) = \frac{1}{6} + \frac{1}{6} - \frac{1}{6} = \frac{5}{6}$$

since the intersection of A and B, A ∩ B = {5} has exactly one element. Had we only computed P(A) + P(B) = 1/6 + 1/6, we would have counted {5} twice, and rule 3 remedies this by subtracting P({5}) once.