Probability Presentation

(from Wikipedia, the free encyclopedia)

In probability theory, the sample space, often denoted S, Ω or U (for "universe"), of an experiment or random trial is the set of all possible outcomes. For example, if the experiment is tossing a coin, the sample space is the set head, tail. For tossing a single die, the sample space is $\{1, 2, 3, 4, 5, 6\}$. Any subset of the sample space is usually called an event, while subsets of the sample space containing just a single element are called elementary events.

For some kinds of experiments, there may be two or more plausible sample spaces available. For example, when drawing a card from a standard deck of 52 playing cards, one possibility for the sample space could be the rank (Ace through King), while another could be the suit (clubs, diamonds, hearts, or spades). A complete description of outcomes, however, would specify both the denomination and the suit, and a sample space describing each individual card can be constructed as the Cartesian product of the two sample spaces noted above.

In probability theory, an event is a set of outcomes (a subset of the sample space) to which a probability is assigned. Typically, any subset of the sample space is an event (i.e. all elements of the power set of the sample space are events), but when defining a probability space it is possible to exclude certain subsets of the sample space from being events.

If we assemble a deck of 52 playing cards and two jokers, and draw a single card from the deck, then the sample space is a 54-element set, as each individual card is a possible outcome. An event, however, is any subset of the sample space, including any single-element set (an elementary event, of which there are 54, representing the 54 possible cards drawn from the deck), the empty set (which is defined to have probability zero) and the entire set of 54 cards, the sample space itself (which is defined to have probability one). Other events are sets are proper subsets of the sample space that contain multiple elements. So, for example, potential events include:

- "Red and black at the same time without being a joker" (0 elements),
- "The 5 of Hearts" (1 element),
- "A King" (4 elements),
- "A Face card" (12 elements),
- "A Spade" (13 elements),
- "A Face card or a red suit" (32 elements),
- "A card" (54 elements).

Since all events are sets, they are usually written as sets (e.g. $\{1, 2, 3\}$), and represented graphically using Venn diagrams. Venn diagrams are particularly useful for representing events because the probability of the event can be identified with the ratio of the area of the event and the area of the sample space. (Indeed, each of the axioms of probability, and the definition of conditional probability can be represented in this fashion.)

If two events A and B are complementary, then $\Pr(A) + \Pr(B) = 1$.

In probability theory, Boole's inequality (also known as the union bound) says that for any finite or countable set of events, the probability that at least one of the events happens is no greater than the sum of the probabilities of the individual events.

Formally, for a countable set of events $A1, A2, A3, \ldots$, we have $\Pr\left[\bigcup_i A_i\right] \leq \sum_i \Pr\left[A_i\right]$.

Conditional probability is the probability of some event A, assuming event B. Conditional probability is written P(A—B), and is read "the probability of A, given B". If A and B are events, and $P(B) > 0$, then $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$. Equivalently, we have $P(A \cap B) = P(A \mid B) \cdot P(B)$. If $P(A \cap B) = P(A)P(B)$, or equivalently, $P(A|B) = P(A)$, then we say that A and B are independent.

If $P(A \cap B) = 0$ and $P(B) \neq 0$, we say that A and B are mutually exclusive events. Then $P(A \mid B) = 0$ (i.e. the probability of A happening, given that B has happened, is nil since A cannot happen if B happens).

In mathematics, a random variable is discrete if its probability distribution is discrete; a discrete probability distribution is one that is fully characterized by a probability mass function. Thus X is a discrete random variable if $\sum_u \Pr(X = u) = 1$ as u runs through the set of all possible values of the random variable X. If X is a discrete random variable with values x1, x2, ... and corresponding probabilities p1, p2, ... which add up to 1, then E(X) can be computed as the sum or series $E(X) = \sum_i p_i x_i$.

If the probability distribution of X admits a probability density function f(x), then the expected value can be computed as $\mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x)\, dx$.

In combinatorial mathematics, the inclusion-exclusion principle states that if A1, ..., An are finite sets, then

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{i=1}^{n} |A_i| - \sum_{i,j\,:\,i<j} |A_i \cap A_j| + \sum_{i,j,k\,:\,i<j<k} |A_i \cap A_j \cap A_k| - \ \cdots\cdots\ \pm \left|A_1 \cap \cdots \cap A_n\right|.$$

where $|A|$ denotes the cardinality of the set A. For example, taking $n = 2$, we get a special case of double counting: in words, we can count the size of the union of sets A and B by adding $|A|$ and $|B|$ and then subtracting the size of their intersection. The name comes from the idea that the principle is based on over-generous inclusion, followed by compensating exclusion. When $n > 2$ the exclusion of the pairwise intersections is (possibly) too severe, and the correct formula is as shown with alternating signs.

The inclusion-exclusion principle can also be used in probability where it becomes:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P\left(A_i\right) - \sum_{i,j\,:\,i<j} P\left(A_i \cap A_j\right) + \sum_{i,j,k\,:\,i<j<k} P\left(A_i \cap A_j \cap A_k\right) - \cdots\cdots \pm P\left(\bigcap_{i=1}^{n} A_i\right)$$

The principle of inclusion-exclusion, combined with de Morgan's theorem, can be used to count the intersection of sets as well. Let A be some universal set such that $A_k \subseteq A$ for each k, and let $\overline{A_k}$ represent the complement of A_k with respect to A. Then we have

$$\left|\bigcap_{i=1}^{N} A_i\right| = \left|\overline{\bigcup_{i=1}^{N} \overline{A_i}}\right|$$

thereby turning the problem of finding an intersection into the problem of finding a union.

Examples

1. (Monty Hall problem) Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

2. (GRE math subject sample) Let x and y be uniformly distributed. independent random variables on $[0, 1]$. The probability that the distance between x and y is less than 1/2 is ()?
   (a) 1/4, (b) 1/3, (c) 1/2, (d) 2/3, (e) 3/4.

3. A drawer has red socks and black socks. You take two socks out at random and the probability that both are red is exactly 1/2. How many socks are in the drawer? Find the smallest possible solution. Can you find another solution?

4. A one-foot stick is broken at random in two places. What's the average length of the smallest piece? Middle piece? Largest piece?

5. Boston and Detroit are the only remaining contenders for the championship. Each has won 2/3 of its games. However, Boston has completed its schedule, while Detroit has three games left. If Detroit were to win two and lose one, it would finish in a tie with Boston, and have a 50% probability of winning the play-off and the championship. However, the team it is playing is slightly better than average, and instead of a 2/3 (66 2/3%) probability of winning, Detroit has only a 66% probability of winning each game. As a result, the probability of Detroit winning the championship is about

   (a) 45% (b) 47% (c) 49% (d) 51%

6. Player A has \$1 and B has \$2. They play a game where A has 2/3 chance of winning, and each time the winner gets \$1 from the loser. What's the probability that A goes bankrupt first?

7. (Putnam 2004-A1) Basketball star Shanille O'Keal's team statistician keeps track of the number, $S(N)$, of successful free throws she has made in her first N attempts of the season. Early in the season, $S(N)$ was less than 80% of N, but by the end of the season, $S(N)$ was more than 80% of N. Was there necessarily a moment in between when $S(N)$ was exactly 80% of N?

8. (MIT training) An unfair coin (probability p of showing heads) is tossed n times. What is the probability that the number of heads will be even?

9. Among integers not over 1000, how many of them cannot be divided by any of $3, 5, 7$?

10. (VT 2004) A computer is programmed to randomly generate a string of six symbols using only the letters A,B,C. What is the probability that the string will not contain three consecutive As?

11. (MIT training) Two persons agreed to meet in a definite place between noon and one o'clock. If either person arrives while the other is not present, he or she will wait for up to 15 minutes. Calculate the probability that the meeting will occur, assuming that the arrival times are independent and uniformly distributed between noon and one o'clock.

12. (Putnam 1961) Let α and β be given positive real numbers with $\alpha < \beta$. If two points are selected at random from a straight line segment of length β, what is the probability that the distance between them is at least α?

13. (Canadian Open Mathematics Challenge, 2002) Suppose that M is an integer with the property that if x is randomly chosen from the set $\{1, 2, 3, .., 999, 1000\}$, the probability that x is a divisor of M is $1/100$. If $M <= 1000$, determine the maximum possible value of M.

14. Two points are picked at random on the unit circle $x^2+y^2 = 1$. What is the probability that the chord joining the two points has length at least 1?

15. (Putnam 2001-A2) You have coins $C_1, C_2, \ldots, C_n$. For each k, C_k is biased so that, when tossed, it has probability $1/(2k + 1)$ of falling heads. If the n coins are tossed, what is the probability that the number of heads is odd? Express the answer as a rational function of n.

16. (Putnam 2002-B1) Shanille O'Keal shoots free throws on a basketball court. She hits the first and misses the second, and thereafter the probability that she hits the next shot is equal to the proportion of shots she has hit so far. What is the probability she hits exactly 50 of her first 100 shots?

Solution of Monty Hall problem: Once the host has opened a door, the car must be behind one of the two remaining doors. Since there is no way for the player to know which of these doors is the winning door, many people assume that each door has an equal probability and conclude that switching does not matter. However, as long as the host knows what is behind each door, always opens a door revealing a goat, and always makes the offer to switch, opening a losing door does not affect the probability of 1/3 that the car is behind the player's initially chosen door. As there remains only one other door, the probability that this door conceals the car must be 2/3 (1 - 1/3). The "equal probability" assumption, whilst being intuitively seductive, is actually incorrect; switching increases the chances of winning the car from 1/3 to 2/3. (see more from `http://en.wikipedia.org/wiki/Monty_Hall_problem`)