CHAPTER 6

Probability

6.1. Probability

6.1.1. Introduction. Assume that we perform an experiment such as tossing a coin or rolling a die. The set of possible outcomes is called the *sample space* of the experiment. An *event* is a subset of the sample space. For instance, if we toss a coin three times, the sample space is

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}\,.$$

The event "at least two heads in a row" would be the subset

$$E = \{HHH, HHT, THH\}\,.$$

If all possible outcomes of an experiment have the same likelihood of occurrence, then the probability of an event $A \subset \mathcal{S}$ is given by Laplace's rule:

$$P(E) = \frac{|E|}{|S|}\,.$$

For instance, the probability of getting at least two heads in a row in the above experiment is $3/8$.

6.1.2. Probability Function. In general the likelihood of different outcomes of an experiment may not be the same. In that case the probability of each possible outcome x is a function $P(x)$. This function verifies:

$$0 \leq P(x) \leq 1 \quad \text{for all } x \in S$$

and

$$\sum_{x \in S} P(x) = 1\,.$$

The probability of an event $E \subseteq S$ will be

$$P(E) = \sum_{x \in E} P(x)$$

Example: Assume that a die is loaded so that the probability of obtaining n point is proportional to n. Find the probability of getting an odd number when rolling that die.

Answer: First we must find the probability function $P(n)$ ($n = 1, 2, \dots, 6$). We are told that $P(n)$ is proportional to n, hence $P(n) = kn$. Since $P(S) = 1$ we have $P(1)+P(2)+\cdots P(6) = 1$, i.e., $k\cdot 1+k\cdot 2+\cdots + k \cdot 6 = 21k = 1$, so $k = 1/21$ and $P(n) = n/21$. Next we want to find the probability of $E = \{2, 4, 6\}$, i.e. $P(E) = P(2)+P(4)+P(6) = \frac{2}{21} + \frac{4}{21} + \frac{6}{21} = \boxed{\frac{12}{21}}$.

6.1.3. Properties of probability. Let P be a probability function on a sample space S. Then:

1. For every event $E \subseteq S$,
$$0 \leq P(E) \leq 1\,.$$
2. $P(\emptyset) = 0$, $P(S) = 1$.
3. For every event $E \subseteq S$, if $\overline{E} =$ is the complement of E ("not E") then
$$P(\overline{E}) = 1 - P(E)\,.$$
4. If $E_1, E_2 \subseteq S$ are two events, then
$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)\,.$$
In particular, if $E_1 \cap E_2 = \emptyset$ (E_1 and E_2 are *mutually exclusive*, i.e., they cannot happen at the same time) then
$$P(E_1 \cup E_2) = P(E_1) + P(E_2)\,.$$

Example: Find the probability of getting a sum different from 10 or 12 after rolling two dice. *Answer*: We can get 10 in 3 different ways: 4+6, 5+5, 6+4, so $P(10) = 3/36$. Similarly we get that $P(12) = 1/36$. Since they are mutually exclusive events, the probability of getting 10 or 12 is $P(10)+P(12) = 3/36+1/36 = 4/36 = 1/9$. So the probability of *not* getting 10 or 12 is $1 - 1/9 = 8/9$.

6.1.4. Conditional Probability. The *conditional probability* of an event E given F, represented $P(E \mid F)$, is the probability of E assuming that F has occurred. It is like restricting the sample space to F. Its value is
$$P(E \mid F) = \frac{P(E \cap F)}{P(F)}\,.$$

Example: Find the probability of obtaining a sum of 10 after rolling two fair dice. Find the probability of that event if we know that at least one of the dice shows 5 points.

Answer: We call $E =$ "obtaining sum 10" and $F =$ "at least one of the dice shows 5 points". The number of possible outcomes is $6 \times 6 = 36$. The event "obtaining a sum 10" is $E = \{(4,6),(5,5),(6,4)\}$, so $|E| = 3$. Hence the probability is $P(E) = |E|/|S| = 3/36 = 1/12$. Now, if we know that at least one of the dice shows 5 points then the sample space shrinks to

$$F = \{(1,5),(2,5),(3,5),(4,5),(5,5),(6,5),(5,1),(5,2),(5,3),(5,4),(5,6)\}\,,$$

so $|F| = 11$, and the ways to obtain a sum 10 are $E \cap F = \{(5,5)\}$, $|E \cap F| = 1$, so the probability is $P(E \mid F) = P(E \cap F)/P(F) = 1/11$.

6.1.5. Independent Events. Two events E and F are said to be *independent* if the probability of one of them does not depend on the other, e.g.:

$$P(E \mid F) = P(E)\,.$$

In this circumstances:

$$P(E \cap F) = P(E) \cdot P(F)\,.$$

Note that if E is independent of F then also F is independent of E, e.g., $P(F \mid E) = P(F)$.

Example: Assume that the probability that a shooter hits a target is $p = 0.7$, and that hitting the target in different shots are independent events. Find:

1. The probability that the shooter does not hit the target in one shot.
2. The probability that the shooter does not hit the target three times in a row.
3. The probability that the shooter hits the target at least once after shooting three times.

Answer:

1. $P(\text{not hitting the target in one shot}) = 1 - 0.7 = 0.3$.
2. $P(\text{not hitting the target three times in a row}) = 0.3^3 = 0.027$.
3. $P(\text{hitting the target at least once in three shots}) = 1-0.027 = 0.973$.

6.1.6. Bayes' Theorem. Suppose that a sample space S is partitioned into n classes $C_1, C_2, \dots, C_n$ which are pairwise mutually exclusive and whose union fills the whole sample space. Then for any event F we have

$$P(F) = \sum_{i=1}^{n} P(F \mid C_i)\, P(C_i)$$

and

$$P(C_j \mid F) = \frac{P(F \mid C_j)\, P(C_j)}{P(F)}\,.$$

Example: In a country with 100 million people 100 thousand of them have disease X. A test designed to detect the disease has a 99% probability of detecting it when administered to a person who has it, but it also has a 5% probability of giving a false positive when given to a person who does not have it. A person is given the test and it comes out positive. What is the probability that that person has the disease?

Answer: The classes are C_1 = "has the disease" and C_2 = "does not have the disease", and the event is F = "the test gives a positive". We have: $|S| = 100,000,000$, $|C_1| = 100,000$, $|C_2| = 99,900,000$, hence $P(C_1) = |C_1|/|S| = 0.001$, $P(C_2) = |C_2|/|S| = 0.999$. Also $P(F \mid C_1) = 0.99$, $P(F \mid C_2) = 0.05$. Hence:

$$\begin{aligned} P(F) &= P(F \mid C_1) \cdot P(C_1) + P(F \mid C_2) \cdot P(C_2) \\ &= 0.99 \cdot 0.001 + 0.05 \cdot 0.999 = 0.05094\,, \end{aligned}$$

and by Bayes' theorem:

$$\begin{aligned} P(C_1 \mid F) &= \frac{P(F \mid C_1) \cdot P(C_1)}{P(F)} = \frac{0.99 \cdot 0.001}{0.05094} \\ &= 0.019434628 \cdots \approx 2\%\,. \end{aligned}$$

(So the test is really of little use when given to a random person—however it might be useful in combination with other tests or other evidence that the person might have the disease.)