Essential Probability Theory

Matthew Tien

No matter where you end up in life, a basic understanding of probability will serve you extremely well - it is one field of mathematics applicable to day-to-day life. Its use is also central to many fields of physics, computer science, mathematics, and genetics. This skill module will introduce fundamental concepts of probability and ask you to solve problems along the way.

Counting If you plan on taking any probability class it is essential that you understand how to properly account for any possibility and combination. This is why it essential we start with counting. Let us start with words. An English word is made up of at least one letter from the English Alphabet (A-Z). For every letter there are 26 possible letters to create the word. Let us try to count all possible combinations of words of length four.

$$\_\ \_\ \_\ \_ = \underline{\text{(A-Z)}} \text{ X } \underline{\text{(A-Z)}} \text{ X } \underline{\text{(A-Z)}} \text{ X } \underline{\text{(A-Z)}}$$

For every letter in the word, there are 26 letters to pick from A-Z. So we can write the following:

$$\underline{26} \text{ X } \underline{26} \text{ X } \underline{26} \text{ X } \underline{26} = (26)^4 = 456{,}976$$

This means we have 456,976 potential combinations of all four letter words. Now let us put down the parameter that you are not allowed to repeat a letter after it has been used.

$$\_\ \_\ \_\ \_ = \underline{\text{(A-Z)}} \text{ X } \underline{\text{(A-Z)-1}} \text{ X } \underline{\text{(A-Z)-2}} \text{ X } \underline{\text{(A-Z)-3}}$$

So if we calculated the possibilities.

$$\underline{26} \text{ X } \underline{25} \text{ X } \underline{24} \text{ X } \underline{23} = 358{,}800$$

Now say we are going to use the entire alphabet to make a word that does not repeat letters (scrambling the alphabet out of order). What are the combinations here?

$$\underline{26} \text{ X } \underline{25} \text{ X } \ldots \text{ X } \underline{2} \text{ X } \underline{1} = 26! = 4.03291461 \times 10^{26}$$

Here we are introducing the factorial (!) operation. The factorial operation multiplies all the positive integers from the number to 0. It is also important to note that 0! = 1. We can relook at the four letter example where you are not allowed to repeat letters to make words

$$\underline{26} \text{ X } \underline{25} \text{ X } \underline{24} \text{ X } \underline{23} = \frac{26!}{22!} = 358{,}800$$

Problem 1: How many different combinations of License Plates are there? Requirements are three letters followed by three digits with repeats. Tip: draw spaces and write the numerical possibility per space.

Events When mathematicians talk about the probability of some particular thing happening, they say use the word event to denote some type of specific outcome. A six-sided die turning up a five is an event, a coin landing on tails is an event, and you successfully passing this skill module are an event (albeit an unlikely one). Mathematically, we would write `the probability that event A occurs' as P(A).

Though probability is a theoretical science, we can still talk about conducting experiments and the results of those experiments. Rolling a six-sided die constitutes an experiment with possible outcomes of the roll turning up a one, two, three, four, five, or six. The collection of all possible outcomes of an experiment is called the sample space of that experiment, sometimes denoted by *S*. The sample space of the roll of a die is {1 ; 2 ; 3 ; 4 ; 5 ; 6}.

Problem 2: In the case of a die, what is the probability that you will roll a one or a two or a three or a four or a five or a six? What does this tell you about the sum of the individual event probabilities of an experiment's sample space?

Independent Events Probability starts to get interesting when you consider multiple events and sequences of events. For instance, you might be interested in calculating the probability that a DNA sequence contains 20 Guanine bases in a row (20 different events), or the probability that you get an even number when rolling a die (3 separate events; a 2 or a 4 or a 6). But first there is an extremely important distinction to make; two events can be independent of or dependent on one another. When you flip two coins, each outcome is completely independent of one another; one coin turning up heads will not affect the second flip in the least. In this special case, you can calculate the probability of two independent events, A and B, occurring together as:

$$P(AB) = P(A) * P(B)$$

In many realistic situations, however, events are often linked. Learning to recognize the interdependency of events is an extremely important skill.

Problem 3: You learn from the weather channel that out of the 90 days of summer in Austin, it rained on 20 of them. If you picked one day at random from those days, what is the probability that it would be a rainy day? Can you reliably calculate the probability that two consecutive days will be rainy? Is this different than calculating the probability that two days chosen at random will be rainy? Why or why not?

Problem 4: A single human chromosome contains approximately 100 million base pairs. Assume that each of the four DNA bases (guanine, adenine, cytosine, thymine) is equally likely to appear at any given position. What is the probability that there is a sequence of a single base repeating exactly 20 times in this chromosome? Genome analyses have revealed that sequences longer than 20 pairs are quite common. What does this tell you about the assumption of random placement of bases? What might the functional explanation for this phenomenon be?

Problem 5: Two brothers apply for a job as a pair and each has his own car. On any given morning, brother A's car will function with probability .8 and brother B's car will function with probability .9 (these two events are also independent). The boss of the company denies them the job because he needs them to be able to be at work with greater than 95% probability. Has the boss wrongly refused them the job, and if so, why?

Conditional Probability Another interesting question that probability theory can help answer is: Given that event A has occurred, what is the probability that event B will occur" Conditional Probability has innumerable applications, and is of fundamental importance to artificial intelligence learning algorithms. We write 'the probability of event B occurring given that event A has occurred' as P(B|A) and can calculate it using the following formula:

$$P(B|A) = \frac{P(AB)}{P(A)}$$

Problem 6: Suppose the probability that Dr. Laude gives your first inquiry a passing grade is $\frac{1}{3}$*. If you turn in a passing first inquiry, the probability that your second inquiry will be given a passing grade increases to* $\frac{2}{3}$ *and if your first inquiry sucks, the probability that your second inquiry will pass decreases to* $\frac{1}{6}$*. Dr. Laude loses your first inquiry but not your second inquiry. Given that your second inquiry passes, what is the probability that your first passed? Hint: Calculating P(second inquiry passing) requires that you consider that your first inquiry could have passed or failed.*

Bonus Problems! Getting either of these correct will grant you a passing grade on this skill module.

Bonus 1: 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, chickens. The car and the chickens were placed randomly behind the doors before the show. The rules of the game show are as follows: After you have chosen a door, the door remains closed for the time being. The game show host, Pat, who knows what is behind the doors, now has to open one of the two remaining doors, and the door he opens must have a chicken behind it. If both remaining doors have chickens behind them, he chooses one randomly. After Pat opens a door with a chicken, he will ask you to decide whether you want to stay with your first choice or to switch to the last remaining door. Imagine that you chose Door 1 and the host opens Door 3, which has a chicken. He then asks you "Do you want to switch to Door Number 2?" Is it to your advantage to change your choice? Explain thoroughly.

Bonus 2: If you choose three points at random from an infinite plane, what is the probability that the triangle that the points form will be obtuse? If you cannot answer this question, identify the assumption(s) that this question makes.