What is a Probability?

Probabilities are numbers between 0 and 1 that indicate the likelihood of an event. Generally, the statement that the probability of hitting a target- that is being fired at- is 0.9 (or 90%) indicates that we are more likely to hit the target than to miss it. More precisely, the statement means that on average, 9 out of every 10 shots are expected to hit the target. Therefore, probabilities close to 1indicate events that are more likely, while numbers close to zero indicate events that are less likely. A probability of 1 indicates certainty, while a probability of 0 indicates an impossible event. It is important to immediately make the following observations:

1. Probabilities have meaning only before events occur. It makes no sense to say that the probability of rain yesterday is 0.5. It either rained or it did not. This also follows from the fact that probabilities are used to predict uncertain events. Prediction can only be associated with future events. Probabilities have meaning only for repeated events
2. Probabilities have meaning only for repeated events. In the example above, if we fire at the target only once, we will either hit it or miss it with that single shot. Only when we repeat the firings can we talk about the average number of hitting. Similarly, if we say that the probability of surviving a surgery is 90%, this means that on average, 90 out of every 100 patients would survive this surgery. This information is very useful to the hospital administering the surgery. However, for a patient who would undergo the surgery only once, this information has little value. Regardless of the probability value, the patient is either going to survive the surgery or fail it.
3. The accuracy of a probability estimate increases as the number of repetitions increases. When we say that a 0.9 probability of hitting a target indicates that on average, every 9 out of 10 shout will hit, we really mean if we took 10 shots and recorded the hits, then repeated this process many times, we expect the average of all recorded hits to be close to 9. Therefore, with a 0.9 probability, hitting 900 out of 1000 is more likely than hitting 9 out of 10.

Assigning Probabilities

The fact that a 90% probability indicates an event that would occur on average in 9 out of 10 trials leads to a direct way of assigning probabilities. We would simply calculate the probability of an event as the ratio between the observed number of occurrences of that event (frequency of the event) and the total number of trials. In other words, we may say:

$$P[A] = \frac{Number\ of\ times\ A\ occured\ (Frequency\ of\ A)}{Total\ number\ of\ trials} = \frac{f_A}{N} \qquad [1]$$

Where A denotes the event of interest. For example, if we want to find the probability of obtaining a defective item from a production line, we may collect a sample from the line and count the number of defectives in the sample. In this case, the number of trials is the sample size and the number of times the event occurred is the number of defectives in the sample. Equation 1 may then be used to assign the probability. As in point 3 above, we expect our accuracy in estimating the probability to increase as *N* increases.

The only problem with this approach, is the time and cost associated with observing the occurrence of the events. In some cases, however, the frequency of the event and the total number of possible trials (or outcomes) may be theoretically counted, eliminating the need for observation. In this case, we can modify equation [1] to:

$$P[A] = \frac{Number\ of\ times\ A\ may\ occure}{Total\ number\ of\ ways\ to\ proceed} \qquad [2]$$

For example, if we are interested in the event of obtaining a defective item from a box that is known to have 3 defective items and 7 good ones. The probability is simply 0.3, calculated as: $P[A] = \frac{Number\ of\ ways\ to\ obatin\ a\ defective\ item}{Total\ number\ of\ ways\ to\ obtain\ an\ item} = \frac{3}{10} = 0.3$.

Equation 2 is sometimes called the *classical method* of assigning probabilities.

To proceed further in this topic, it is helpful to introduce some counting techniques.

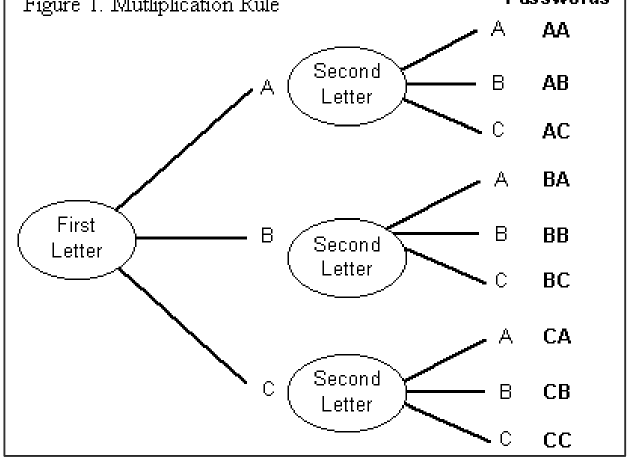
Counting Techniques

In general, if we can divide the trials under consideration to stages (say *r* stages) where each stage may be conducted in n_i ways ($i = 1, 2, \ldots r$), then the total number of trials (N) may be calculated as: $N = n_1 n_2 \ldots n_r$. This is known as the *multiplication rule*.

For example, consider forming a password consisting of two letters out of the first three letters in the alphabet (a, b, and c). We may divide this problem to two stages, each stage corresponding to choosing a letter. In the first stage we may choose any of the three letters (i.e., $n_1 = 3$). Assuming that we may use repeated letters, the second stage may also proceed in 3 ways (i.e., $n_2 = 3$). In this case, $N = n_1 n_2 = 3*3 = 9$.

The tree diagram to the right may be used to explain the counting rule above. Choosing the first letter may result in A, B, or C. For each of these results, the second choice could again be A, B, or C. The possible passwords are therefore (AA, AB, AC, BA, BB, BC, CA, CB, CC). We may be interested in the probability of an event (E) of having a password that consists of two identical letters (AA, BB, CC). Three passwords (out of 9) fulfill this condition or represent this event. Thus, $P[E] = \frac{3}{9} = 0.33$

Figure 1. Mutliplication Rule

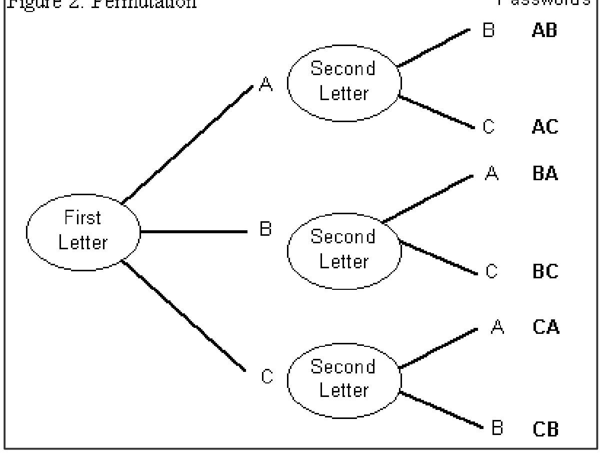


The *multiplication rule* is the general rule for counting. However, there is a special case of the rule when the each n_i is smaller by one than the one before it. In other words, if we assume that $n_1 = n$, then $n_2 = n\text{-}1$, $n_3 = n\text{-}2$, … , *etc.* In this case, $N = n(n\text{-}1)(n\text{-}2) \ldots (n\text{-}r\text{+}1)$. By multiplying and dividing by $(n\text{-}r)(n\text{-}r\text{-}1) \ldots (1)$, we get: $N = \frac{n!}{(n-r)!}$

This special case is called the *permutation* rule and results when we are arranging r objects out of n without replacement. It is sometimes written as ${}^{n}P_{r} = \frac{n!}{(n-r)!}$, where P stands for *permutation*. For example, consider the situation above when a letter may not be repeated. In this case, we may choose any of the three letters in the first stage (i.e., $n_1 = 3$). However, once a letter is chosen, it may not be chosen in the second stage (i.e., $n_2 = 2$). In this case, $N = n_1 n_2 = 3*2 = 6$. Or, ${}^{3}P_{2} = \frac{3!}{(3-2)!} = \frac{6}{1} = 6$.

The tree diagram to the right represent this new case. The 6 possible outcomes are (AB, AC, BA, BC, CA, CB). In this case, P[E] = 0. i.e., it is impossible to have a password consisting of two identical letters. If we define a new event (F) as a password consisting of two adjacent letters in the alphabet, we can easily see that $P[F] = \frac{4}{9} = 0.44$ in the first case (with identical letters), while $P[F] = \frac{4}{6} = 0.67$ in the second case.

Figure 2. Permutation

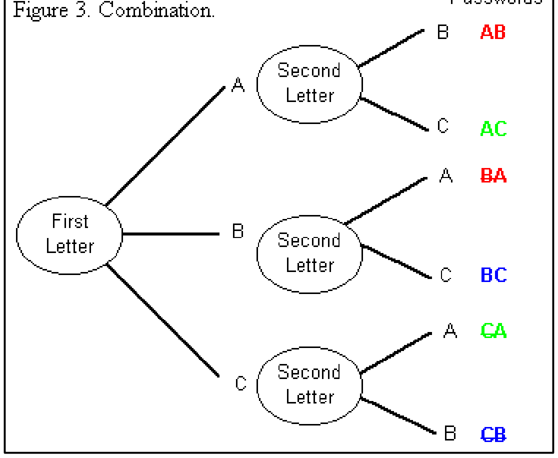


A special case of the permutation rule results when the order of objects is not important and is known as *Combination*. In this case, we are combining (or choosing) r object out of a total of n without replacement. Since there are $r!$ possible repetitions in this process (${}^{r}P_{r}$, arranging r object out of r), it is easy to see that the total number of possible combinations is ${}^{n}C_{r} = \binom{n}{r} = \frac{n!}{r!(n-r)!}$, where ${}^{n}C_{r}$ stands for combining (or choosing r objects out of n). In the password example, if the order of letters is not important, Figure 3 Shows that half of the outcomes would be eliminated resulting in only 3 outcomes. This may be calculated as:

$$N = {}^{3}C_{2} = \binom{3}{2} = \frac{3!}{2!(3-2)!} = \frac{6}{2*1} = 3$$

Notice that: ${}^{n}C_{n-r} = {}^{n}C_{r}$ (e.g., ${}^{3}C_{2} = {}^{3}C_{1} = 3$).

Figure 3. Combination.



To summarize, the *multiplication rule* states that the number of possible outcomes is the product of the outcomes of all stages. *Permutations* are a special case when each stage has one less outcome than the stage before it. It has its application in arranging objects where order is important. *Combinations* are yet a special case from *Permutations* when the order is not important. It has its application in selecting or choosing objects.

Example:

In how many different ways can a true-false test consisting of 10 questions be answered?

Solution:

If we take the questions one at a time, we have a series of 10 stages each with two possible outcomes. Applying the multiplication rule, $N = (2)(2) \ldots (2) = 2^{10} = 1024$.

Example:

In how many ways can 5 starting positions on a basketball team be filled with 8 men who can play any of the positions.

Solution:

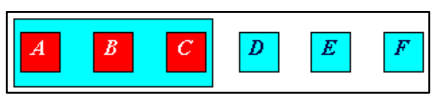
Since the 5 positions are different, the order of selection is important resulting in a permutation of 5 out of 8. $\therefore N = {}^{8}P_{5} = \dfrac{8!}{(8-5)!} = \dfrac{40320}{6} = 6720$. L. K. Gaafar

Example:

a. In how many ways can 6 people be lined up to get on a bus?
b. What is the probability that certain 3 persons will always follow each other?
c. If a certain 2 persons refuse to follow each other, how many ways are possible? What is the probability of this event?
d. In how many ways can 3 people be chosen to get on the next bus?

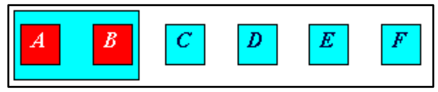
Solution:

a. Since order is important, we will be arranging the 6 people, i.e. $N = {}^{6}P_{6} = \dfrac{6!}{(6-6)!} = 720$
b. As the adjacent figure shows, we will assume that the 3 persons following each other form one block within which the can be arranged in



3! = 6 ways. This block with the other 3 persons may be arranged in 4! = 24 ways. Therefore, using the multiplication rule, N = (6)(24) = 144. The probability of this event is $\frac{144}{720} = 0.20$.

c. Let us consider the complement, i.e., two people following each other. Similar to part "b", this may happen in 2!5! = 240 ways.



Therefore, N = 720 – 240 = 480. The probability of this event is $\frac{480}{720} = 0.667$.

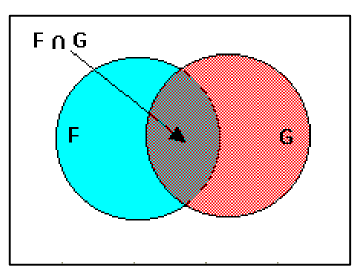
d. Now order is not important as all three will get on the bus. Therefore, $N = {}^6C_3 = \binom{6}{3} = \frac{6!}{3!(6-3)!} = \frac{720}{6*6} = 20$.

Combining probabilities

The presentation above may be simplified using *Set* representations. For that, we define a set that contains all possible outcomes and call it the *Sample Space (S)*. Therefore, the *Sample Space* is the set of all possible outcomes. An event, then, is any subset of the sample space. Assuming that all elements are equally likely, probabilities may be defined as the ratio between the number of elements in the event set to the number of elements in the sample space. An Empty Set (*f*) is used to denote impossible events (*f* = {})

In the password example above, the sample space that corresponds to Figure 1 is: S = {AA, AB, AC, BA, BB, BC, CA, CB, CC}, and the set that corresponds to event E is E = {AA, BB, CC}. Notice that E is a subset of S ($E \subset S$). Counting the number of elements in each set, $P[E] = \frac{3}{9} = 0.33$. The set representing event F (two adjacent letters) is F = {AB, BA, BC, CB}, $P[F] = \frac{4}{9} = 0.44$. We may also be interested in a new event G, in which a password starts with B. Again from Figure 1, G = {BA, BB, BC}, $P[G] = \frac{3}{9} = 0.33$.

Now, what if we are interested in the event of having a password starting with a B *and* consisting of two similar letters (H)? This new event satisfies both E and G at the same time and may be defined as the *intersection* of the two sets, or $H = F \cap G$. H is the set that includes the common elements between F and G, H = {BA, BC}, $P[H] = \frac{2}{9} = 0.22$. Venn diagrams are used to simplify the concepts of probability or event combining. For example, in the Venn diagram to the right both events F and G are represented as circles, while the outer box represents the sample space. $F \cap G$ is the common area or the

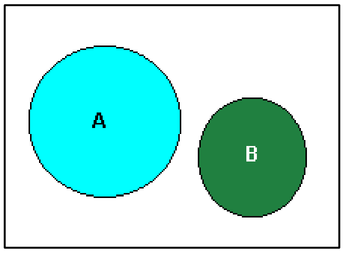


intersection of the two circles.

If we are interested in the event (I) of having a password that either starts with B or consists of two adjacent letters. Obviously, both F and G satisfy I, and we say that I is the union of F and G, $I = F \cup G$. The set I then consists of all elements in both F and G (without repetitions), i.e., I = {AB, BA, BC, CB, BB}, $P[I] = \frac{5}{9} = 0.56$. On the Venn diagram, $F \cup G$ is the area covered by both circles. If event areas on the Venn diagrams were scaled according to the probability of the event, it is easy to see that $P(F \cup G) = P(F) + P(G) - P(F \cap G)$. Subtracting $F \cap G$ is done because on the Venn diagram, $F \cap G$ is part of both circles and is added twice when we add $F+G$, and hence, it must be subtracted once. $P(F \cup G) = P(F) + P(G) - P(F \cap G) = \frac{4}{9} + \frac{3}{9} - \frac{2}{9} = \frac{5}{9} = 0.56$

Associated with every event is the complement of the event, which is everything in the sample space that is not in the event. Therefore, the complement of F may simply be defined as "not F" as is denoted by F'. In the example above, F' = {AA, AC, BB, CA, CC}. On the Venn diagram, F' is the area within the box falling outside the blue circle. Clearly, $S' = \phi$.

Events that do not intersect are called mutually exclusive, because they cannot occur simultaneously. In the Venn diagram to the right, if the outcome is in B, it cannot be in A at the same time. Therefore, events A and B are mutually exclusive. Clearly, $A \cap B = \phi$, $P(A \cap B) = 0$.



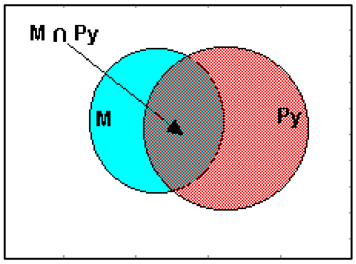
Example:

In a group of 80 students, 48 studied mathematics, 64 studied physics and 40 studied both topics. If one of these students is selected at random, find the probability that:

a. The student took mathematics or physics
b. The student did not take mathematics
c. The student did not take any of the subjects
d. The student took mathematics but not physics

Solution:

We will use M for mathematics and Ph for physics
$P(M) = 48/80 = 0.60$, $P(Py) = 64/80 = 0.80$,
$P(M \cap Py) = 40/80 = 0.50$.



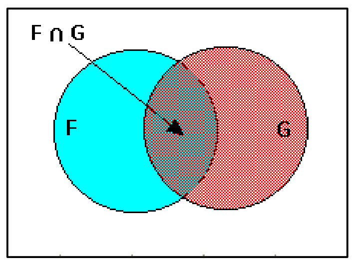
a. $P(M \cup Py) = P(M) + P(Py) - P(M \cap Py) = 0.6 + 0.8 - 0.50 = 0.90$.
b. $P(M') = 1 - P(M) = 1 - 0.60 = 0.40$
c. This is the complement of the 'a,.' i.e., $P((M \cup Py)') = 1 - P(M \cup Py) = 1 - 0.90 = 0.10$.

d. To exclude *Py* and consider *M* only, we take the intersection of *M* with the complement of *Py*, i.e., $M \cap Py'$. $P(M \cap Py') = P(M) - P(M \cap Py) = 0.60 - 0.50 = 0.10$.

<u>Conditional Probability</u>

In the discussion above, we calculated probabilities using the sample space as our reference. In other words, we knew that the outcome we are looking for is in the sample space, but were interested in it being in a given part of the sample space (an event). Now, we consider the case where we know more about the outcome. Specifically, we may be interested in the probability of an event, knowing that another event occurred. This is called conditional probability as we are calculating the probability of an event on the condition of the occurrence of another. For example, in Figure 1, we may be interested in the probability that the password starts with B, given that it consists of two adjacent letters, that is we want the probability of *G* given that *F* occurred. In this case, our outcomes are limited to elements of *F* = {AB, BA, BC, CB}, and *G* can only occur through its common elements with *F*, i.e., {BA, BC}. Therefore, once *F* occurs, *G* has a chance of 2 out of 4 of occurring. In other words, $P(G \backslash F) = \frac{2}{4} = 0.50$, where $P(G \backslash F)$ is read as probability *G* given *F*, and means that we are interested in the probability of *G* occurring given that *F* already occurred.

Conditional probabilities may also be explained using a Venn diagram. On the diagram to the right if we know that *F* occurred, our outcome possibilities will be limited to the blue circle. It is obvious that under this condition, *G* can only occur through its common area (intersection) with *F*. The chances of that happening is the ratio of this common area to the area representing all possible outcomes *F*, i.e., $P(G \backslash F) = \frac{P(G \cap F)}{P(F)} = \frac{2}{9} \div \frac{4}{9} = 0.50$.



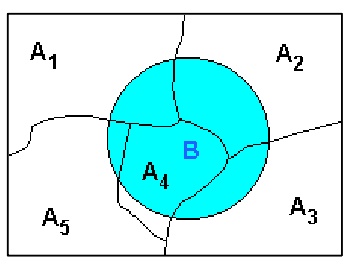
In General, $P(A \backslash B) = \frac{P(A \cap B)}{P(B)}$ [3]

Equation 3 may also be expressed as $P(B \backslash A) = \frac{P(A \cap B)}{P(A)}$ [4]

From 4, $P(A \cap B) = P(B \backslash A)P(A)$, by substituting in 3, we get:

$$P(A \backslash B) = \frac{P(B \backslash A)P(A)}{P(B)} \quad [5]$$

Equation 5 is known as Bayes' rule and is useful in calculating conditional probabilities without going through the intersection probabilities.



If a group of mutually exclusive events A_1, A_2, … , A_k completely partition the sample space, any event may be constructed by summing its intersections with the A_i's. For example, in the Venn diagram to the right, $B = B \cap A_1 + B \cap A_2 + \dots + B \cap A_5$, and $P(B) = P(B \cap A_1) + P(B \cap A_2) + \dots + P(B \cap A_5)$, or in general,

$$P(B) = \sum_{i=1}^{k} P(B \cap A_i), \qquad [6]$$

Since $P(B \cap A_i) = P(B \setminus A_i)P(A_i)$, Equation 6 may be expressed as:

$$P(B) = \sum_{i=1}^{k} P(B \setminus A_i)P(A_i) \qquad [7]$$

Equation 7 is very useful in calculating indirect probabilities. We will call it the *probability assembly rule* in future references.

Finally, if P(A\B) = P(A), then A is not affected by B, or A is *independent* of B as its probability is not affected by the occurrence of B. Using Equation 4, $P(A) = P(A \setminus B) = \frac{P(A \cap B)}{P(B)}$, which leads to:

$$P(A \cap B) = P(A)P(B) \qquad [8]$$

Equation 8 is very useful in calculating the intersection (AND) probability of two *independent* events.

Example:

A system consists of two components. The probability that the second component functions in a satisfactory manner during its design life is 0.9, the probability that at least one of the two components does so is 0.96, and the probability that both components do so is 0.75. Given that the first component functions in a satisfactory manner throughout its design life, what is the probability that the second one does also? [0.926]

We will use the following notation
S: Second component functions in a satisfactory manner
F: First component functions in a satisfactory manner

Solution:

Given information:
P(S) = 0.9,
$P(S \cup F) = 0.96$, this is a direct translation of the statement “the probability that at least one of the two components does so is 0.96.” At least one means S or F, i.e. the union.
$P(S \cap F) = 0.75$, this is a direct translation of the statement “the probability that both components do so is 0.75.” Both means S and F, i.e. the intersection.

Required
P(S\F) = ?

$$P(S \setminus F) = \frac{P(S \cap F)}{P(F)},$$

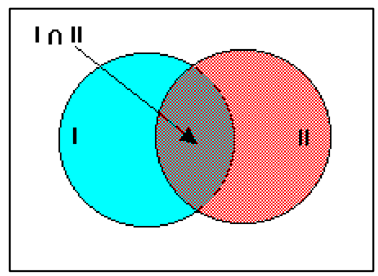
P(S ∪ F) = P(S) + P(F) - P(S ∩ F) ⇒ P(F) = P(S ∪ F) + P(S ∩ F) - P(S)
∴ P(F) = 0.96 + 0.75 – 0.9 = 0.81

$$P(S \setminus F) = \frac{P(S \cap F)}{P(F)} = \frac{0.75}{0.81} = 0.926$$

Example:

An exam has two questions; I and II. The probability that a student will solve **only** question I correctly is 0.4. The analogous probability for question II is 0.45. The probability that a student will solve either question I or II correctly is 0.9. Find the probability that the student will correctly solve:
a) both questions. *[0.05]*
b) exactly one question. *[0.85]*
c) none of the question. *[0.1]*
d) question II given that he solved question I correctly. *[0.11]*

Solution:



Given Information
P(I ∩ II') = 0.40
P(II ∩ I') = 0.45
P(I ∪ II) = 0.90

Required
a. P(I ∩ II) = ?
b. P((I ∩ II') ∪ (II ∩ I')) = ? (I only OR II only)
c. P(I ∪ II)' = ?
d. P(II\I) = ?

a. P(I ∩ II) = P(I ∪ II) - P(I ∩ II') - P(II ∩ I') = 0.90 – 0.40 – 0.45 = 0.05
b. P((I ∩ II') ∪ (II ∩ I')) = 0.40 + 0.45 = 0.85
c. P(I ∪ II)' = 1 - P(I ∪ II) = 1- 0.9 = 0.10
d. $P(II \setminus I) = \frac{P(II \cap I)}{P(I)}$, P(I) = P(I ∩ II') + P(I ∩ II) = 0.40 + 0.05 = 0.45

$$\therefore \; P(II \setminus I) = \frac{P(II \cap I)}{P(I)} = \frac{0.05}{0.45} = 0.11$$