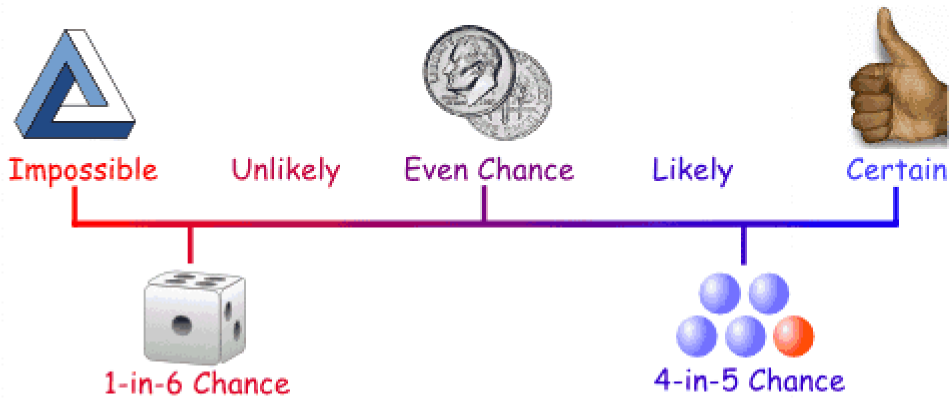
Critical Reasoning 10 - Probability

with guest editor Katherine Eyal



Probability is the extent to which an event is likely to occur, while **probability theory** is that branch of mathematics that deals with quantities having random distributions. (OED) The point of this study unit is not to master a mathematical field but to become literate in probability in the same way that we are with numeracy. This is an especially important (and often neglected) skill set for philosophers because, with the exception of logical proofs, philosophical theories are only more or less probable. Those that can be decided empirically have usually been incorporated into sciences such as Physics or Psychology, whereas until only very recently there has grown up a new field of **Experimental Philosophy**, where philosophical hypotheses are tested directly. However these still yield a statistical probability rather than a categorical yes or no. It is important therefore to be able to interpret such probabilities and to know how they should be contextualised, compared or combined.

Probability theory has its historical origins in the study of games of chance such as dice or cards. Obviously there was a profit motive in beating or simply understand the odds, therefore such games were studied especially closely. Unlike meteorologists who express the probability of precipitation as a percentage, mathematicians use a dimensionless scale from 0 to 1, with 0 denoting impossibility, and 1 certainty, and any value in between being less or more likely respectively. Note that probabilities can never be negative or exceed 1. Furthermore the sum of probabilities over all possible outcomes will always be 1.

Suppose we begin with the toss of a coin. Since there are two possible outcomes, heads or tails, which are equally likely, the probability of *either* heads or *either* tails is ½ or 0.5 while the probability of turning up either heads *or* tails is 1 or certainty. Suppose however we want to calculate the probability of drawing an ace from a standard pack of cards (with the jokers removed.) There are 52 such cards with 4 suits (clubs, diamonds, hearts and spades) with 13 cards in each. Either way, we may reason that, because there are 4 aces in every pack of 52 cards, the probability of drawing an ace should be 4/52 ≈ 0,077. Alternatively we may argue that because there is only one ace in every suite of 13 cards and we are certain to draw *a* card from 1 suit, the probability of drawing an ace should be 1/13 ≈ 0,077. In both cases we are calculating the probability of an event $P(E)$ by dividing the number of **favourable events** (*e.g.* a coin landing heads or a card being drawn an ace) by the number of **possible outcomes** (*i.e.* the total number of possible results.) Thus,

$$P(E) = \frac{Number\ of\ favorable\ events}{Number\ of\ possible\ outcomes} \qquad (...1)$$

And the opposite or **complement** of an event, say of not drawing an ace or not throwing a six, is given by

$$P(not\ E) = 1 - P(E) \qquad (...2)$$

Modern textbooks give a more detailed and nuanced treatment of probability. However the "classical model of probability," from which the above formula for $P(E)$ is derived, is adequate so long as we are dealing with **discrete probability distributions**. These deal with events that occur in countable **sample spaces** (*i.e.* we can count the set of all possible outcomes.) Coin tosses, throwing a die or picking a single card are examples of such events. Furthermore, they are also said to be **independent** because the occurrence of one event has no influence on the next (throwing and obtaining heads on one coin toss has no relationship with whether we obtain heads or tails on the next throw). Discrete probability distributions are also said to be **mutually exclusive** if the occurrence of one event precludes that of another, such as heads precluding tails and *vice-versa*.

Suppose that we wanted to demonstrate that a particular coin was fair. We would do a little **statistical experiment** and flip it as many times as we had the patience to do so. We would record how many times heads (or tails) were obtained, and how many times we threw the coin. Now the **relative frequency** or proportion of times that an event occurred can be calculated by

$$P(E) = \frac{number\ of\ obserations\ \ of\ an\ event}{number\ of\ time\ the\ experiment\ was\ repeated} \qquad (...3)$$

Now according to the **law of large numbers**, if the outcomes of an experiment are independent and it is performed repeatedly, then the proportion of favourable outcomes will approach its theoretical probability. Therefore if a coin is fair, the proportion of heads (or tails) should approach its theoretical probability of 1 in 2, or 0,5 as we flip the coin repeatedly. Of course we may have a run of heads (or tails) but as the number of flips increases their relative contribution to the overall proportion of heads (or tails) will be diminished. As discussed in Critical Reasoning 06 concerning the **gambler's fallacy**, there is a tendency to believe that tails is somehow "due" after a run of a certain number of heads (and *vice-versa,*) while we know however that the probability of either remains the same 0,5 because the events (tosses) are independent of one another. (Had we used a coin made of butter or some other such material things might have been different.)

Often however, we are interested in some combination of probabilities such as "A and B" or "A or B" or even" A given that B." Statisticians tend to prefer the language of set theory, rather than logic in representing such combinations. Thus "A and B" is denoted "A ∩ B" (the intersection of sets A and B) such that for two independent events,

$$P(A \cap B) = P(A)P(B) \qquad (...4)$$

The following examples are drawn from "Wikipedia: Probability" *E.g.* if two different coins are tossed at the same time or even if one coin is tossed on two different occasions, the joint probability of both turning up heads is ½ x ½ = ¼ or 0,25.

However if two events are mutually exclusive then probability of *either* of them occurring "A or B" is denoted by "A ∪ B" (the union of sets A and B) such that,

$$P(A \cup B) = P(A) + P(B) \qquad (...5)$$

E.g. Because the chances of rolling one number or another in single throw of a die are mutually exclusive, the chance of rolling a 1 *or* a 2, say, is ⅙ + ⅙ = ⅓ or ≈ 0.33

In the case that events are *not* mutually exclusive then the probability of either of them occurring is equal to the sum of their individual probabilities minus their joint probability, thus:

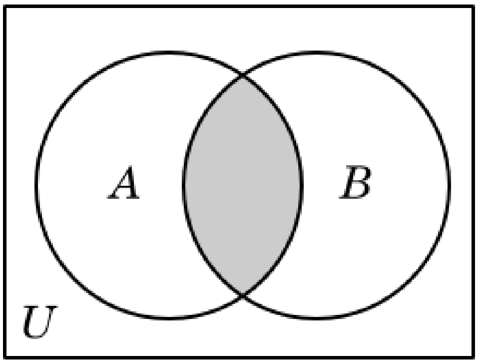
$$P(A\ or\ B) = P(A) + P(B) - P(A\ and\ B) \qquad (\ldots 6)$$

Alternatively,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \qquad (\ldots 7)$$

E.g. What is the chance that a single card drawn at random from standard deck will be either a heart *or* a face card? Here we have three possibilities to consider: the possibility of any card being a heart (13 in 52 because there are 13 hearts in every pack); the possibility of any card being a face card (12 in 52 because there are 12 face cards in every pack); and the possibility of any card being both a heart and a face card which we don't want to count twice (3 in 52 because there are only 3 face cards that are also hearts in every pack.) Now we simply substitute these values into eqn. 7 above so that we get,

$$P(A \cup B) = \frac{13}{52} + \frac{12}{52} - \frac{3}{52} = \frac{22}{52} \text{ or } \frac{11}{26} \approx 0{,}42$$



Sets provide a helpful way of visualising this problem. Suppose that "A" represents the set of all hearts in a pack and "B" the set of all face cards in a pack. There will be 13 elements in set A {A♥; 2♥; 3♥; …; K♥} as well as 12 elements in set B {J♥;J♠; J♦; J♣; Q♥; Q♠; Q♦; Q♣; K♥; K♠; K♦; K♣ }. The shaded area, the intersection, contains only 3 cards that are common to both sets A ∩ B {J♥; Q♥; K♥}. While we want to count these cards, we don't want to count them twice when we tally the cards in A and then tally them in B. The way that eqn. 7 gets round this is to count them twice anyway: "$P(A) + P(B)$"on the left, and then subtract them once: "$-P(A \cap B)$" on the right. Of course it is not actual cards we are counting but probabilities of cards being drawn at random that we are counting. However, so long as thinking about a problem informally at first helps us think about it more rigorously later, we are justified in doing so.

Very often the probability of one event is contingent upon another event, in which case we are dealing with **conditional probability**. The expression "$P(A|B)$" is used to denote conditional probability and is usually read as "the probability of *A*, given *B*." Suppose, for example, that an opaque bag contains 4 balls: 2 red and 2 blue. The chance of drawing out a red ball is clearly ½ (2 reds out of a possible 4.) However, the chance of drawing *another* red ball depends very much upon which ball was taken out the first time round: If a red ball was taken out first, the probability of drawing out another is ⅓ (1 remaining red out of a possible 3) however, had a blue ball been taken out first, the probability of drawing out another red would be ⅔ (2 remaining reds out of a possible 3.) (Wikipedia: Probability)

Mathematically, conditional probability is expressed as the ratio of the joint probability A and B and that of B alone thus,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B\,|A)P(A)}{P(B)} \qquad (...8)$$

with the restriction that all terms are positive and that $P(B)$ cannot be zero.

E.g. What is the probability of drawing two aces from the same pack? Call event *A*: "drawing an ace first" and event *B*: "drawing an ace second." What we are looking for is $P(A\ and\ B)$ or $P(A \cap B)$. For event A, $P(A) = \frac{4}{52}$ because there are 4 aces in a pack of 52 cards. For event B, $P(B|A) = \frac{3}{51}$ because there are only 3 aces left in our pack that now has 51 cards. So by making $P(A \cap B)$ the subject of the formula for $P(B|A)$ in this case, we can determine:

$$P(A \cap B) = P(A)P(B|A) = \frac{4}{52} \text{ x } \frac{3}{51} = 0{,}0045 \text{ or just under half a percent.}$$

Combinations and Permutations

In Critical Reasoning 05 we alluded to combinations and permutations because we required knowing how many ways there are of arranging just so many possible T's and F's when constructing truth tables. There we noted that **combination** refers to the number of ways of selecting members from a group such that the order does not matter, whereas for **permutation** the order *is* taken into account. The website mathsisfun.com, from which we have borrowed the following explanation and diagram, suggests an easy way for remembering the difference between the two: "My fruit salad is a combination of apples, grapes and bananas," is indeed a combination because we don't care about the order of the ingredients. It's still the same fruit salad. On the other hand, "The combination to the safe is 472," is actually a permutation because we must take the order of the numbers into account: 2-7-4 or 7-2-4 will not open the safe. So strictly speaking the combination lock at right is actually a **permutation** lock.

Often, when calculating probabilities we are required to know precisely how many ways there are to select a certain number of objects (in order or irrespective of order) from a given group. Firstly, we will have to introduce the notion of a factorial. *n!* (say: "n factorial") is simply the product of all the whole numbers from 1 to *n*. Thus,

5*!* = 1 x 2 x 3 x 4 x 5

7*!* = 1 x 2 x 3 x 4 x 5 x 6 x 7

1*!* = 1 but note that 0*!* = 1 by definition.

Factorials have some special properties which we should observe: They are natural numbers.

n! ∈ {ℕ}

Nor can they be negative numbers or fractions because we have defined them that way.

n! = *n*(*n* -1)*!*

Any factorial is the same as the product of *n* and the factorial of the number before, hence the term (n -1)*!*.

E.g. 5*!* = 1 x 2 x 3 x 4 x 5 = 120 = 5 x (1 x 2 x 3 x 4)

Also observe that:

$(n \pm m)! \neq n! \pm m!$

$(n \times m)! \neq n! \times m!$

The factorial sign does not distribute across addition, subtraction or multiplication, however division is discussed below.

The following examples are drawn from the website betterexplained.com by way of introduction:

E.g. How many ways can 3 medals (gold, silver and bronze) be awarded to 8 contestants?

Obviously, if you qualify for a medal, the order in which they are awarded matters very much to you; therefore we are dealing with a permutation. When we award the gold medal we have 8 choices available to us. When next we award the silver medal we have 7 choices remaining because we have already handed out one medal. Finally, when we come to award the bronze medal we have 6 remaining choices out of the original 8 because, by this time, we have already handed out two medals. After this we had no more medals to distribute. So the total number of ways of distributing the medals we had was:

8 x 7 x 6 = 336 ways to order 3 out of 8

But this looks just like the first three terms of 8*!* only written in descending order, which doesn't change the magnitude because *a* x *b* x *c* = *c* x *b* x *a* (commutation.) However, we can't simply use 8*!* because it has too many terms *i.e.* 5 x 4 x 3 x 2 x 1. So somehow we have to find a way to "stop" factorial 8 at 5. If we divide the factorial 8 by 5 factorial we do just that because the terms 5 to 1 cancel top and bottom, thus:

$$\frac{8!}{5!} = \frac{8.7.6.5.4.3.2.1}{5.4.3.2.1} = 8.7.6 = 336$$

The usual way that mathematicians represent this is:

$$\frac{8!}{(8-3)!}$$

or in general,

$$P_{(n,k)} = \frac{n!}{(n-k)!} \qquad \text{(say: "n pick k")} \qquad (...9)$$

where *n* is the number of items in total and *k* is the number we want to pick in a certain order.

The next example from the same website, explains how combinations are derived. According to the author, "If we want to figure out how many combinations we have, we just **create all the**

permutations and divide by all the redundancies." A redundancy in this sense would be the same items picked over again, only in a different order (which for combinations doesn't matter.)

> *E.g.* Suppose that this time round instead we only had 3 identical cans of cool drink that we wanted to hand out to our 8 contestants. How many ways can they be distributed?

At first, we would have 3 choices for the first person, then 2 choices for the second and finally only 1 choice for the last person *i.e.* 3 x 2 x 1 = 6 ways to arrange 3 people. In general there are *n!* ways of arranging *n* number of objects. So for this example, the number of combinations is equal to the number of permutations (336, see above) divided by the number of redundancies (6) for each permutation or,

$\frac{336}{6}$ = 56 combinations.

In general:

$$C_{(n,k)} = \frac{P_{(n,k)}}{k!} \qquad (\ldots 10)$$

which, if we substitute the formula for $P_{(n,k)}$ above, we get,

$$C_{(n,k)} = \frac{n!}{(n-k)!k!} = \binom{n}{k} = \binom{n}{n-k} \qquad (\ldots 11)$$

This expression is so important in mathematics that it assigned its own notation (see the big brackets to the right above.) It is also known as the **binomial coefficient** (of which more below.) Once we know how many combinations or permutations are available to us we can simply use (…11) above to calculate the probability of an event.

> *E.g.* In the game of pool there are 15 coloured balls which are numbered accordingly, plus the white cue ball which makes 16. If a London bookie offers you £ 600 in a bet if you draw out any 2 balls of your naming (without replacement) from out of an opaque bag but says he will keep your £100 bet if you don't, what are the odds of you winning the bet?

We can express the possible outcomes as the combination of 16 balls taken 2 at a time, or $C_{(16,2)}$ because the order in which we draw out the balls is not important. Now,

$$C_{(16,2)} = \frac{16!}{(16-2)!2!} = \frac{16!}{14!.2!} = \frac{16.15.14!}{14!.2.1} = 120 \quad \text{(Note how the 14!'s cancel.)}$$

And since there is only one combination of two balls you are allowed to pick per bet, your chances of doing so are a slim $\frac{1}{120}$ or about 0.83%. Clearly this is not a bet you should wager.

> *E.g.* Suppose, furthermore that the bookie grows impatient with you refusing his bet and offers you a much larger prise of £1200 if you can draw out 3 balls (without replacement) in any specific order of your choosing, but ups the cost of your bet to £150, which he will keep if, either you pull out the wrong balls or the right balls but in the wrong order. Should you reconsider his wager?

This time the order does matter, therefore we must express the possible outcomes as a permutation of 16 ball from which we want to pick 3 in the correct sequence, or

$$P_{(16,3)} = \frac{16!}{(16-3)!} = \frac{16!}{13!} = \frac{16.15.14.13!}{13!} = 16 \times 15 \times 14 = 3360 \text{ possible ordered outcomes.}$$

And since there is only one order of balls that bookie would be satisfied with as a win, the odds are an even slimmer $\frac{1}{3360}$ or about 0.03%. Only a fool would bet on such odds, which is why they call gambling, in general and the lottery in particular, **the stupid tax**!

The Binomial Distribution

Although the **binomial theorem** is so important in Mathematics that is required to be proven by every Mathematics student at first year level, the inclusion of binomial distributions in Research Methodology texts in the Humanities is somewhat reluctant. On the one hand, you will be required to know about them but on the other, you won't be examined concerning them, except perhaps a maximum of one question in a roundabout way.

The binomial distribution is usually introduced as a probability model for experiments involving coin tosses; however it is applicable to any situation involving a random variable that has two discrete outcomes, be they: Yes or no; HIV+ or HIV-; pass or fail; male or female; immunised or not immunised; employed or unemployed… and so on.

According to Kruger & Janeke (2012 p. 41) the binomial distribution applies to variables with the following properties:

- They must consist of a fixed number of trials. Although probabilities can be calculated for trials of various lengths, the lengths of each trial must be kept constant for the probability distribution to be determined.
- They must have two, "mutually exclusive and collectively exhaustive events, typically labelled as 'success' and 'failure'." If someone asks, "Are you pregnant?" you cannot reply, "Only slightly" or "Mostly not."
- The probability of being a success *p* or not a success 1 - *p* must remain constant throughout a trial. You cannot, for example, begin with a fair coin and progress to one that comes up heads with a probability of 0.6 towards the end of the trial.
- The outcome of one trial must be independent of any other trial.

Using the binomial distribution allows one to calculate probabilities without having to calculate all the possible outcomes and then apply the multiplicative rule above. The distribution itself may be symmetrical where p = 0.5, regardless of the number *n*. On the one hand, the distribution may be positively skewed (to the right) if $p < 0.5$ or negatively skewed (to the left) if $p > 0.5$. On the other hand, the distribution will become more symmetrical as *p* approaches 0.5 and/or as *n* becomes larger. (Kruger & Janeke *l.c.*)

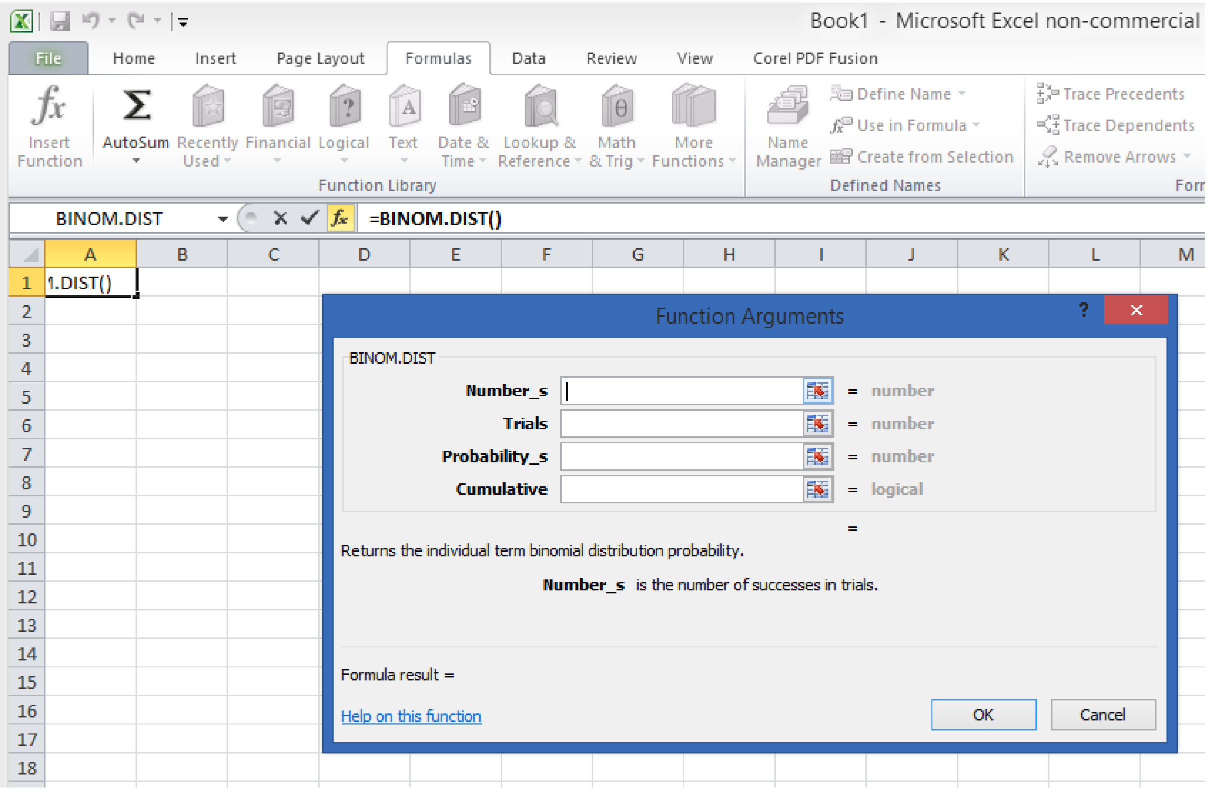
The probability mass function $P(x)$ for binomial probabilities involves 3 variables, where x is the probability of "success" for n number of trials and for a probability p for any single successful outcome. Thus,

$$P(x,n,p) = \frac{n!}{x!\,(n-x)!}p^x(1-p)^{n-x}$$

But since we recognise the left-hand side of this equation as a binomial coefficient of the type $\binom{n}{k}$ that we learned above, we can simplify the notation, somewhat to:

$$P(x,n,p) = \binom{n}{x}p^x(1-p)^{n-x}$$

However this won't reduce the number of steps required in actually manually calculating this probability. The easiest way to handle this formula is to let a spreadsheet program like Excel calculate the probabilities you require using the "BINOM.DIST" function under "more functions" / "statistical". When you select this option you will be presented with the following window:



Simply fill in the fields as follows:

Number_s: Enter the number of successes to be included in the calculation. (Most introductory examples will just involve 1.)

Trials: Enter the number of trials.

Probability_s: Enter the probability of any single successful outcome. (0 to 1)

Cumulative: Enter "FALSE" (unless you really do want cumulative probabilities.)

And click "OK". The good news is that you will not be required to memorise or calculate the probability mass function for binomial probabilities in an undergraduate exam. Just be sure you know what it is about and when it would be appropriate to use it, if you had to.

Bayesian Analysis

We have already encountered conditional probabilities above expressed as "*A*, given *B*" such that:

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

So named after Thomas Bayes (1701 - 1761) the English statistician, philosopher and clergyman who formulated a special case, **Bayes' theorem** is used in statistical inference as well as a wide number of calculations involving probability. In public life, for example, the correct use of Bayes' theorem in several of high profile court cases has been used to exonerate falsely accused persons. However because Bayesian inference or interpretation has a number of important epistemic implications it has become a valuable tool for contemporary philosophers, especially within the Philosophy of Science.

Fittingly, only probably (not) a portrait of the Rev. Thomas Bayes used in the book History of Life Insurance *(1936). No earlier images survive.*

To continue where we left off drawing aces from a pack, consider the following unlikely example, not for its philosophical weight but for its intuitive ease of grasp. (Adapted from Wikipedia: Bayes' theorem.)

Suppose a man tells you that he had an interesting conversation with someone on a train. Without knowing who this person was, you reckon that the probability that the man had a conversation with a woman to be 50% or 0,5, since both men and women use trains in equal numbers, you assume. Now suppose that the man mentions that the person he conversed with had long hair. Like most people, based on the new information, you would probably be more inclined to believe that the person was a women rather than a man. But how certain can one be, given that there are a good many men who do wear their hair long?

In order to proceed we need to know (or estimate) some probabilities: Before we knew about the long hair, we supposed the probability of the conversationalist being a woman at, let us call this, $P(W)$ = 0,5. If we also knew that 75% of women wear their hair long, then we can represent this conditional probability as $P(L|W)$ = 0,75 which reads "the probability of *L* given *W* is 0,75." Similarly if we had found out that 15% of men wear their hair long, we can represent that conditional probability as $P(L|M)$ = 0,15. Since we are assuming that every human is either a male or female, we can treat Mas the complement of $W$and *vice versa*. Given that a proportion of both men and women wear their hair long, we also need to calculate $P(L)$, the probability that any random person might have long hair. In this example $P(L)$ is the sum of the individual probabilities for L given W x the probability of W and the probability of L given M x the probability of M. In symbols:

$$P(L) = P(L|W)P(W) + P(L|M)P(M)$$

Since we already know these individual probabilities, we can substitute them into the above equation, thus:

$$P(L) = (0{,}75)(0{,}5) + (0{,}15)(0{,}5) = 0{,}45$$

Now we can express the probability of the conversationalist being a woman, given that the person in question was sporting long hair, as an instance of Bayes' theorem and, while we are at it, substitute in the probabilities that we know:

$$P(W|L) = \frac{P(L|W)P(W)}{P(L)} = \frac{(0{,}75)(0{,}5)}{0{,}45} = 0{,}83$$

This shows that that we were correct in assuming that the conversationalist was probably a woman, however now we have actually quantified that probability based on a handful of assumptions.

This little example belies the central role that probability theorists consign to Bayes' theorem; indeed Sir Harold Jeffreys declared that Bayes' theorem "is to the theory of probability what Pythagoras's theorem is to geometry." (1973, p. 31)

Bayesian confirmation theory is the most successful attempt yet, by the logical positivist movement, in treating inductive reasoning in accordance with principles of probability. According to Bogen (2005) on whose entry the remainder of this section is based, it is an attempt "to provide a uniform, general account of scientific knowledge." (p. 81) Suppose that H is an hypothesis, and E is some piece of evidence, then $P(H|E)$ is the probability or degree of belief that one might assign to H on discovering E to be true. Accordingly, $P(H|E)$ varies with:

1. $P(H)$, the prior probability, which is the degree of confidence in H before observing E.
2. $P(E|H)$, which is the probability of expecting to observe E given that H is true (not to be confused with $P(H|E)$ where H is a function of E, rater than the other way round.)
3. $P(E)$, sometimes termed the marginal likelihood or "model evidence" which is the prior degree of confidence given to E regardless of whether or not H is true.

Ignoring background assumptions, Bayes' theorem says the probability of H, given E should vary directly with (1) and (2) and inversely with (3), thus:

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)} \qquad (\ldots 12)$$

Incorporating background beliefs (B) on which such expectations are contingent yields:

$$P(H|E) = \frac{P(E|H \cap B)P(H|B)}{P(E|B)} \qquad (\ldots 13)$$

According to Bogen, the Bayesian approach is attractive because: "it avoids technical difficulties which beset its rivals; it treats epistemic support quantitatively; it seems to shed light on disagreements (emphasised by Kuhn) among scientists over the epistemic bearing of evidence [and it] applies to reasoning from uncertain evidence." (p. 82)

On the other hand Bogen mentions two challenges to the Bayesian approach, including that

- "… applications to real world cases are clouded by the apparent arbitrariness of its assignment of numbers to prior degrees of confidence." Arguably however, this is down to a psychological limitation on our part rather than logical fault of the theory: we are simply not used to stipulating degrees of prior belief in dimensionless increments from 0 to 1.

- "It has trouble explaining how a theory can be tested against old evidence already accepted with certainty." In such a case $P(H)$ and $P(E)$ are effectively equal to 1 and therefore $P(H|E)$ can be no different from its prior probability.

Lastly, Bayesian epistemology is saddled with the same presumption as traditional epistemology in that the laws of logic, including Bayesian confirmation theory itself, are exempt from revision on the basis of empirical evidence.

Notwithstanding, Bayesian analysis is useful in delivering quantifiable probabilities in vastly more contexts than such peculiarities as mentioned by Bogen. Indeed its attraction for philosophers is that it does so even from uncertain evidence.

Tasks

The following statements contain fallacious reasoning based on misinterpretations of probability. Identify the error(s) in each case and amend them where possible. (The last two are adapted from Wikipedia: Base rate fallacy)

1. Jake is on a winning streak of a lifetime at Roulette. Given that the odds of winning diminish over time, he is even more likely to loose on the next round.
2. If something can go wrong, it will go wrong.
3. Due to the impossibility of predicting the precise trajectory of all the balls from one draw to the next and the fact that no two outcomes can ever be the same, it is impossible to calculate the exact probability of winning the lottery from one game to the next.
4. Fiona (which is a heathen name) is always seen in gothic attire, listening to death metal. On the preponderance of probabilities alone, she is more likely to be a Satanist than a Christian.
5. Police use breathalysers which always indicate when a driver is drunk, however they also display a false positive in 5% of cases. Given that 1 in 1000 drives is drunk at the wheel at any given time, a driver's chances of being drunk as indicated by a breathalyser is 95%.

Feedback

1. It is true that because the odds are always in favour of the house, otherwise they would be out of business; the longer one plays the more chances a casino will have of reliving you of your money. However, because the outcome of one game of Roulette is independent of that of another, the chances of winning or losing, from one game to the next, will be the same, irrespective of whether one were on a winning or losing "streak."

2. This is a statement of Murphy's Law. As it stands it is an invalid, deductive argument because one cannot validly argue from probable premises to a certain conclusion. When we come to the subject of modal logic it will be clear that "It is possible that *p*" does not imply that "It is necessary that *p*." Of course, when one is repeatedly frustrated, Murphy's Law can have a *subjective* ring of truth.

3. It is true that it is physically impossible to precisely predict the protracted motion of any three or more bodies (known as the three body problem,) just as it is true to say that the motion of even one cast die cannot be predicted remotely with any certainty. However just because a physical outcome may be impossible to predict does not imply that the distribution of statistical outcomes cannot be known. In the same way that one can never know the fall of a die, one can however know the probability of such an outcome and that if cast repeatedly it will approach its theoretical probability of 1 in 6.

   To win the South African lottery one must correctly draw six numbers, in any sequence, corresponding to 6 out of 49 balls. Without recourse to any formula one should be able to intuitively "see" that there are 49 possibilities in drawing the first ball, 48 in drawing the second, 47 in drawing the third… and 44 in drawing the last. Therefore there are:

   49 x 48 x 47 x 46 x 45 x 44 = 10 068 347 520

   possibilities. However, each possible group of six numbers can be drawn in different ways depending on the order of the numbers draw: 6 for the first, 5 for the second, 4 for the third… and 1 for the last, or

   6 x 5 x 4 x 3 x 2 x 1 = 720

   possible ways or "picks." And because each ball is as likely to be picked as the next, and that the likelihood *a* number been drawn each time is 1, there are:

   $\frac{10\ 068\ 347\ 520}{720}$ = 13 983 816

   different groups of six balls that can be drawn at any one time. Therefore the probability of picking the right numbers in any draw is:

   $\frac{1}{13983816}$ = 0,000 000 071 516

   which is vanishingly small. To put this number into perspective, it is roughly the same probability as obtaining 24 consecutive heads in the flipping of coin. Hence the name "the stupid tax"- Anyone would be "stupid" to bet against such odds. And yet, if one does not play, one's chances of winning are precisely zero.

   This explanation was adapted from the website http://ms.mcmaster.ca/fred/Lotto/ by Fred M. Hoppe, although many of you will recognise it as simply a combination "n choose k" where n = 49 and k = 6 so that, in a single line, we can calculate:

   $C_{(49,6)} = \frac{49!}{6!} = \frac{10\ 068\ 347\ 520}{720}$ = 13 983 816

   As to the claim that no two lottery outcomes can ever be the same, this is both factually and statistically false. In 2009 for example, the Bulgarian lottery turned up the same number sequence twice within five days. The probability of two such a sequences being drawn twice are vanishingly small. However, it is no more remote than the probability of any one sequence of numbers being drawn once, because the outcome of any two draws are independent of each other, just as the chance of rolling a six again, after just having rolled a six, is still ⅙.

4. This is a truly awful example of the base rate fallacy, where the person making the statement has simply ignored the base rate (frequency) of being a Christian (there are an estimated 2 billion globally,) which is several orders of magnitude higher than that of being a Satanist (estimated to be in the thousands.) (Robinson, 2006) For good measure there is also a red herring thrown in about having a "heathen" name, as caricatured in Charles Dickens's 1850 novel, *David Copperfield*. (See David's aunt, Betsey Trotwood's remarks about Peggotty.)

5. This is another instance of the base rate fallacy, however to calculate the actual probability that a random driver is drunk as indicated by a breathalyser requires recourse to Bayes' Theorem. The conditional probability we are interested in can be represented as:

$$p(D|B)$$

where "D" means that the driver is actually drunk and "B" that the breathalyser indicates that a driver is drunk. According to Bayes' Theorem:

$$p(D|B) = \frac{p(B|D)p(D)}{p(B)}$$

From the information given, we known that:

$p(D)$ = 0,001 the 1 in 1000 drivers that are driving drunk at any time.

$p(S)$ = 0,999 where "S" means that the driver is sober.

$p(B|D)$ = 1 100% of breathalysed drunks are correctly indicated as being drunk.

$p(B|S)$ =0,05 5% of cases falsely indicated as drunk when they are actually sober.

The formula above also requires that we know the value of $p(B)$ which we can calculate from the existing information:

$$p(B) = p(B|D)p(D) + p(B|S)p(S) \ = 0{,}05095$$

As you can see, this is the sum of probabilities of being someone indicated as being drunk, given that that someone is in fact driving drunk plus the probability of somebody being falsely indicated as being drunk when in fact that person is sober.

Now substituting these values into Bayes' Theorem above, we get

$$p(D|B) = \frac{1(0.001)}{0.05095} = 0{,}019627$$

which is about fifty times smaller than most people would guess. Of course, such considerations can never exonerate ever actually driving drunk.

The next Critical Reasoning unit will introduce the very important Predicate Calculus by which statements or propositions can be expressed more naturally in terms of their subjects and predicates, rather than just one letter for an entire proposition.

References

"Better explained" website:

http://betterexplained.com/articles/easy-permutations-and-combinations/

BOGEN, J. (2005) Bayesian confirmation theory **In:** HONDERICH, T. Ed. (2005)*The Oxford Companion to Philosophy,* (2nd ed.) Oxford University Press : New York p.81 - 82

KRUGER, P. & JANEKE, H. C. (2012) *Psychological Research - Study Guide for PYC3704*. UNISA Department of Psychology

JEFFREYS, H. (1973) *Scientific Inference* (3rd ed.) Cambridge University Press

ROBINSON, B.A. (2006) *Religious Satanism, 16th century Satanism, Satanic Dabbling, etc*. Ontario Consultants on Religious Tolerance: http://www.religioustolerance.org/satanism.htm