MUTUALLY EXCLUSIVE AND INDEPENDENCE: UNRAVELLING BASIC MISCONCEPTIONS IN PROBABILITY THEORY

I.W. Kelly
University of Saskatchewan

F.W. Zwiers
Environment Canada

Two central concepts in probability theory are those of "independence" and of "mutually exclusive" events and their alternatives. In this article we provide for the instructor suggestions that can be used to equip students with an intuitive, comprehensive understanding of these basic concepts. Let us examine each of these concepts in turn along with common student misunderstandings.

1. Mutually Exclusive versus Non Mutually Exclusive Events

Events are mutually exclusive when the occurrence of one of the events rules out the possibility of the occurrence of the other events of concern. The outcomes, for example, on the toss of a single die are 1, 2, 3, 4, 5 or 6. The outcomes are all mutually exclusive because when a die is tossed and a number turns up, all the other numbers cannot occur. On the other hand, the events "eating rocky mountain oysters for breakfast" and "eating vegetable soup for breakfast" are not mutually exclusive since it is possible that one has both for breakfast, however improbable that might be.

The distinction between contradiction and contrariety should be introduced at this point to students. Two statements are contradictories when they cannot both be true, and cannot both be false. Two statements are contraries if they can both be false and a third statement, different from both, can be true. For example, "It is raining outside this building at this moment" and "It is not raining outside this building at this moment" are contradictories. However, the statements "All mathematicians are very intelligent" and "No mathematicians are very intelligent" are contraries since although both cannot be true, it is almost certainly the case that both are false. Some mathematicians are very intelligent and others are not.

There are mutually exclusive events of both types. In the die tossing example already described we could divide the sample space into two parts, say, "even" and "odd". These events are mutually exclusive and contradictory. By taking the sample space to be the outcomes 1, 2, 3, 4, 5 and 6 we have implicitly made an assumption about the operation of die tossing: namely, that it is impossible for the die to come to rest on a point or an edge. If the sample space were expanded to include these possibilities then the events "even" and "odd" are contraries. The reason for introducing these concepts is, of course, to ensure that the student is quite clear about the fact that a pair of mutually exclusive events are not necessarily complementary events.

The die example presented above has somewhat of a contrived air about it because most of us would accept the proposition that the probability of a die coming to rest on a point or an edge is zero. However, things are not always that clear cut in nature. Amongst humans there are males, females, and other borderline cases (e.g., Klinefelter's syndrome) which occur with nonzero probability. Thus, the events "male" and "female" are in fact contraries.

Another point that should be considered in class is the fact that very often in the real world it is <u>not always clear</u> whether or not two events are mutually exclusive. The examples we present to students when teaching the concept are, understandably, clearcut. For example, if we obtain a six on a single toss of a fair die, it rules out the possibility of obtaining a one, two, three, four, or five. In the real world we often do not have enough information to be sure. Not all students are aware that "Clark Kent" and "Superman" are <u>not</u> mutually exclusive. It would be useful to draw examples from nature where the distinction between contradictory, contrary, mutually exclusive and non-mutually exclusive are not always clear.

2. Independent versus Dependent Events

Events are independent when the occurrence (or nonoccurrence) of one of the events carries no information about the occurrence (or nonoccurrence) of the other event. Mathematically, two events A and B are considered to be independent if $P(A \cap B) = P(A) \cdot P(B)$. For example, if the probability that Obadiah has escargot for breakfast tomorrow is 0.4 and the probability that it will rain tomorrow is 0.3, then the probability that both events will occur tomorrow is $(0.4) \cdot (0.3)=0.12$. When we have more than two events, the situation becomes a bit more complicated – all possible combinations of component events must follow the multiplication rule. That is, each combination must also involve independent events.

Students have several difficulties with the distinction between independent and dependent events. The first parallels the problem that arises with mutually exclusive events, namely, determining when events in the real world are independent or dependent. Once again, we seldom help students to bridge the gap between the fuzzy distinctions apparent in nature and the very rigid distinctions made in mathematics. If we toss a pair of dice, the outcome that occurs on one due obviously (for most of us) does not influence what outcome will occur on the other die. In other cases one often needs expertise in a particular area to make a reasoned judgement whether particular events are independent or not. Many years of research were required to demonstrate that there is a dependent relationship between smoking and lung disease (see also, Ayton & Wright, 1985).

A second common misunderstanding involves interpreting a dependent relationship between events as a causal relationship. There are, of course, situations where this is plausible, for example, having tuberculosis is dependent on having tuberculosis bacilli in one's body. However, there are many examples of dependent relationships between events where no causal relationship is involved. For example, having a fire is dependent on the presence of oxygen, although the latter does not cause the former. This is something we can help students with simply by presenting examples.

An issue that sometimes arises with students who have done reading in physics or philosophy concerns whether events can ever really be independent. For example, some writers (e.g., Capra, 1975) take a holistic approach to the universe in which the universe is considered to be a web of relationships where all things communicate intimately with one another and all being is shared. Statisticians sometimes suggest similar things, adding to student confusion. Hays (1981 p. 293), for example, tells us that "There is surely nothing on earth that is completely independent of anything else". A student coming across such a statement would be understandably confused. The student might reason that if nothing is completely independent of anything else, then how can we apply probability formulas that assume independence of events? The answer is that application of such formulas doesn't need to assume that the events are completely independent of each other, only that any relationship is negligible. For example, every body in the universe has some interactive gravitational attraction with every other body, but the gravitational attraction between a human being and a star light years away is so minute that it can be considered non-existent. Underwood (1957, p. 6) put it well:

> *The length of an astronomer's toenails isn't related to phases of the moon; the color of the secretary's hair isn't related to the height to which the corn grows in an Iowa field, and a pygmy tribe in New Guinea has little influence on the alcoholic consumption of a truck driver in Brooklyn.*

As a final point, we should point out that independence is a difficult idea that is clearly understood in relation to specific examples – we talk about conducting an experiment in class and then talk about whether two events (describing sets of possible outcomes of the experiment) are independent. We use phrases such as "events A and B are independent if knowledge about whether A has occurred provides us with no knowledge about whether B has occurred". In a very subtle way an element of time is hinted at in such a statement and it often confuses students. It is therefore important to emphasize to students that <u>whatever</u> the temporal relationship between two (or more) independent events, a knowledge of the occurrence (or nonoccurrence) of any of the events provides one with no knowledge of the future <u>or</u> past outcomes of any of the other events.

In order to understand how this problem with time can arise we will consider the simple example of a coin tossing experiment in which a fair coin is tossed six times. Associated with the experiment is a sample space, the list of all possible outcomes of that experiment, and a collection of events. The latter represent more general descriptions of outcomes of the experiment and can usually occur in more than one way. For example, the event "three heads come up in the six tosses of a coin, can occur in 20 ways. Now suppose that you are blindfolded before the experiment is conducted and that you have the opportunity to bet that the last two tosses of the coin will come up heads. Quick mental arithmetic tells you that this can happen one time in four. Now the coin is tossed six times and, still blindfolded, you are told that the outcome was three heads and given an opportunity to revise your bet. You realize that there are only four ways in which the outcome of six tosses of a coin can result in a total of three heads with a head on each of the last two tosses. Therefore, the relative chances of being a winner once you've been given some knowledge of the outcome of the experiment has been reduced from one-in-four to one-in-five. The events "three heads in six tosses of a coin" and "the last two tosses in a sequence of six tosses are heads" are not independent.

The element of time in the above example relates not to the way in which the experiment was conducted, but to the way in which we think. We were given some knowledge about the outcome of the experiment, namely that three heads had occurred. This in turn gives us some knowledge about the likelihood of other events that may have occurred. We now think that it is less likely that there was a head on each of the last two tosses than before the experiment was conducted. This is essentially what we mean when we say that two events are dependent – knowledge that one event has occurred conveys information about whether or not the second event also occurred. The events "heads on the first toss" and "heads on the third toss" are independent, because when we re-evaluate the likelihood that second event occurred in light of the occurrence of the first event, we see that we have no more knowledge about the outcome of the second event.

The element of time which we mentioned is associated with this process of re-evaluation of probabilities after some information about the outcome of the experiment is available. These *a posteriori*, or conditional probabilities can only be applied to decision making, such as whether to continue a bet or raise the ante in a poker game, after some information about the outcome of the experiment is available.

These probabilities are evaluated by conceptually repeating the experiment with a restricted sample space. We feel that students have some inkling of what goes on, but that we don't explain these concepts to them carefully enough. We suspect that they do feel that something happens as time goes on but don't really understand its mechanics. We should point out, using simple examples such as coin tossing or card games, that we construct new probability models for the experiment as it progresses (or after the fact) which are conditional upon information which we have received about the outcome of the experiment to that point. With these ideas about conditional probability in place it can be shown that two events, say A and B, are independent if the *a priori* probability of A is equal to the *a posteriori* probability of A given B. We think that such an approach is much easier for the student to understand than the standard approach which is found in most text books. The latter consists of stating the usual definition of independence, i.e., that A and B are independent if

$$P(A \cap B) = P(A) \cdot P(B)$$

and then presenting a few examples to illustrate the truth of the definition. We often confuse the student through our use of language, as discussed above, and the failure to link the ideas of independence and conditional probability.

3. Confusion Between Independent Events and Mutually Exclusive Events

Students not only have difficulty with the notions of mutually exclusive and independent events; they very often confuse the two. Most of the confusion arises because we, as instructors, do not take the time to relate the two concepts.

We frequently answer the question "If A and B are mutually exclusive, does it follow that they are not independent?" With the reply, "well, no. For example, . . ." and then trot out the pathological example of a pair of events of probability zero without further explanation. Except in certain pathological cases, the concept of mutually exclusive events is the exact antithesis of the concept of independent events. A nice illustration is provided by Hays (1981, p. 43-44). Suppose that all men are either "bald" or have a "full head of hair". These are mutually exclusive. Let us say the probability of selecting a bald man from the population is 0.60; the probability of selecting a hairy man would be 0.40. If these two events were independent, then the probability of selecting a man who is both bald with a head full of hair would be equal to (0.60).(0.40). But the probability of such a joint event is, of course, zero. Mutually exclusive events are (almost) never independent. Ad hoc explanations tend to deal with specific aberrations that may lead to further confusions and moreover avoid dealing with the fundamental principles upon which the student's question is based.

By planning for instruction of these two fundamental concepts we can insure that the student's understanding is built up systematically. This provides a firmer foundation upon which the student can acquire a grasp of probability theory.

References

Ayton P., & Wright, G. (1985). Thinking with probabilities. Teaching statistics, 7(2), 37-40.

Capra, F. (1975). The tao of physics. Berkeley: Shambala Books.

Hays, W. L. (1981). Statistics. New York, N.Y.: Holt, Rinehart & Winston.

Underwood, B.J. (1957). Psychological research. New York, N.Y.: Appleton-Century-Crofts.