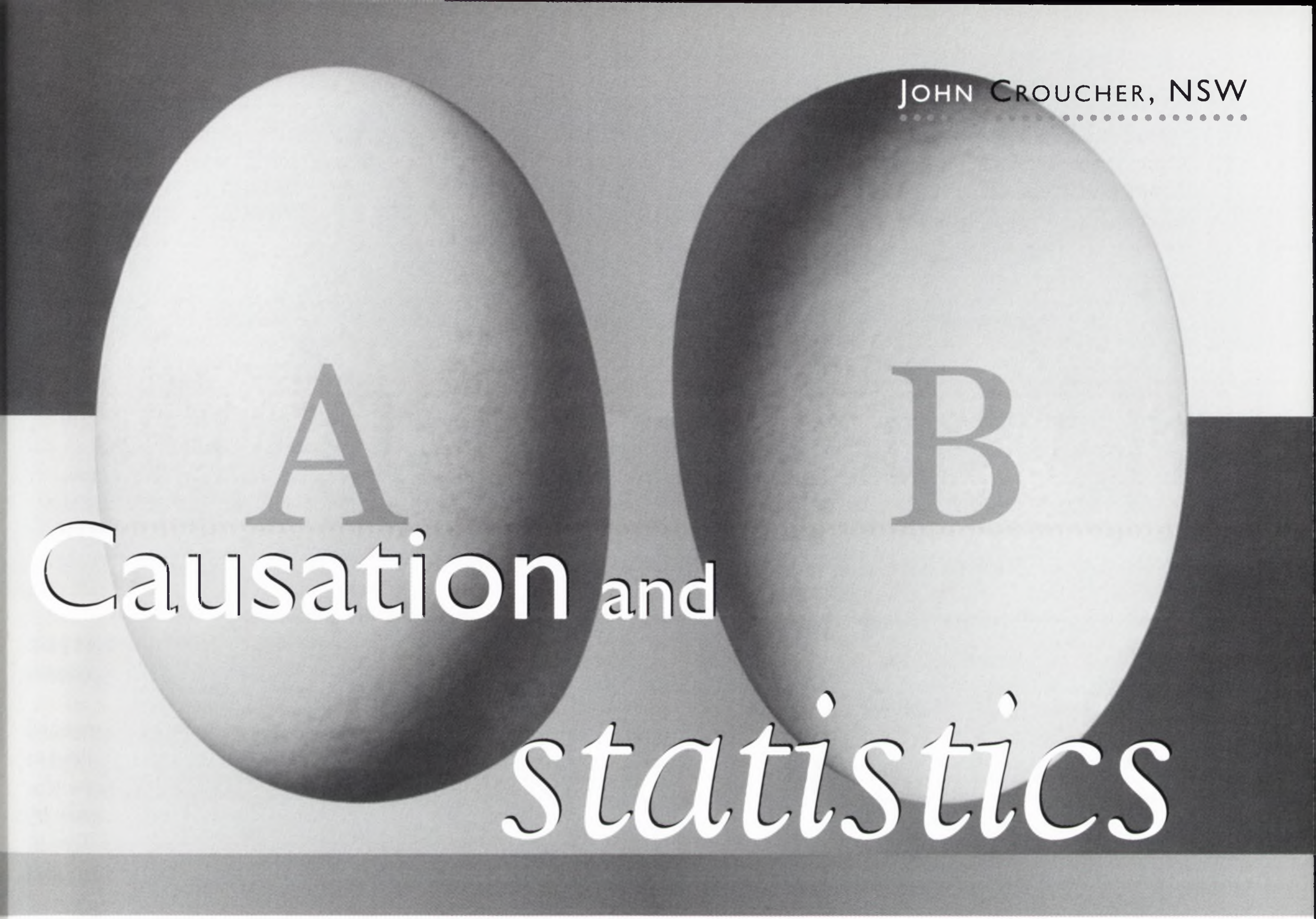
JOHN CROUCHER, NSW

Causation and statistics



The concepts of 'outcome' and 'causation' are fundamental in law and their proof is often dependent on proper statistical analysis.

INTRODUCTION

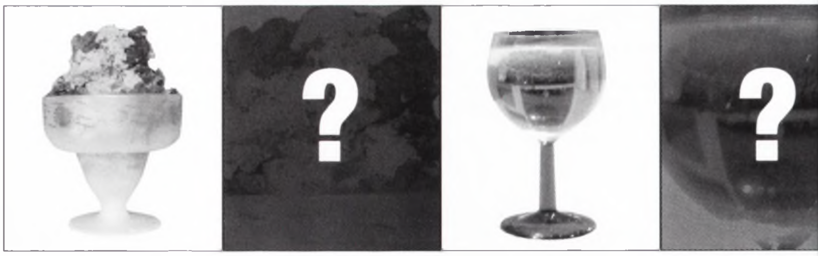
'Causality', or 'causation', is essentially the *relationship* between *causes* and *effects*, and may be expressed in terms that an *event* or *state of affairs*. *A* is said to be the *cause* of an event *B* if *A* is a *reason* that brings about the effect *B*. For instance, while it may be easy to conclude that pressing on the brake causes a car to stop, in general it can be a complex issue to determine that *A* is the *reason* that *B* occurs. It is important to clarify the relationship between causes and effects, as well as how, or even if, causes can bring about effects.

Causation is not only a key issue for a statistician, but plays an important role in law, especially the law of negligence. Indeed, statistics are often used, and misused, in law. One recent, celebrated example of the latter was the trial in the UK of Sally Clark, who was accused of murdering her two children.[1] In this instance, the question was whether the death of the children was caused by her or by some other event. The misuse of statistics at her trial (discussed later in this article) produced wildly distorted 'expert' evidence.

The question is the extent to which the law and statistics are relevant to each other. This article provides a window on how causation is considered from a statistical point of view. ▶

John Croucher is a Professor of Management at the Macquarie Graduate School of Management, Macquarie University, Sydney.
PHONE (02) 9850 8999 **EMAIL** jcrouche@gsm.mq.edu.au

RELATIONSHIPS AND FALLACIES



In a typical context, it is not uncommon for statisticians to first determine whether there is a degree of association between two events or variables.[2] This is usually measured by a *Pearson product moment correlation coefficient* (denoted by r) that will indicate whether the relationship, if any, is weak or strong and whether it is positive or negative. The formula is designed such that the value of r always lies between -1 and $+1$. Once it is calculated, the next step is to test it for *significance*. If it is found to be significant, then it may be possible to use the value of one of the variables to predict the value of the other. This type of calculation is referred to as *regression analysis* and can range from fitting a straight line (linear regression) to sophisticated curves.

Even if two variables are found to be significantly correlated, great care must be taken in interpreting this result. Every introductory statistics book will correctly emphasise that correlation does *not* imply causation. The establishing of cause and effect can be extremely difficult and lead to celebrated arguments, even among professionals. In everyday usage, we often take the expression '*A* causes *B*' to mean '*A* causes an increase in the *probability* of *B* occurring'.

For example, suppose it can be established that there is a signifcant positive correlation between the number of cigarettes smoked and the rate of lung cancer. This does not necessarily establish that smoking must be a *cause* of that increased cancer rate – it may well be that there exists a certain genetic defect that causes both cancer and a craving for nicotine.

Other instances of logical fallacies include:

- Teenagers eat a lot of chocolate. They also have skin problems. Does it necessarily follow that chocolate causes skin problems?
- The sale of ice cream is significantly positively correlated with crime rates. Is selling more ice cream going to lead to an increase in crime? Of course not. The explanation is that high temperatures increase crime rates (apparently by making people irritable) as well as ice cream sales. That is, they are both affected by a *third* variable (temperature) that makes it look like they have a relationship.
- The heights of schoolchildren and their reading comprehension may be positively correlated. However, this does not mean that tall children of the same age will read better. It simply means that both physical and mental attributes develop as a child gets older.
- Over the years, the funding for the arts in Australia has a very high correlation with the amount gambled on poker machines. Does this mean that increased funding is likely to find its way into the pokies?
- Causation in medical negligence claims is particularly problematical. For example, over time the amount spent on alcohol advertising and the amount spent by people purchasing alcohol may be highly correlated. But is it the advertising that causes people to begin drinking or drink more?
- After collecting considerable data, a researcher found a significant correlation in children between mathematical ability and shoe size. Does this mean that having big feet is somehow a cause of mathematical ability, or that mathematics skills cause a child's feet to grow?

In the above examples there is what is known as a 'spurious' or 'false' correlation. This is caused by the presence of a third variable, known as a *lurking variable*, which is actually causing the observed correlation. For instance, in the final example the lurking variable is *age*, since as children grow older they both learn more about mathematics and wear larger shoes. Two characteristics may appear to be strongly related, but this is due to the presence of the lurking variable that is not included in the study. Identifying such variables is not always easy. Take the efforts to show that there is not only a high correlation between heavy drinking and respiratory trouble, but that the former causes the latter. There may well be a lurking variable, for example, in the form of smoking. That is, most heavy drinkers could also be heavy smokers, and it is the smoking that causes the respiratory trouble and not the heavy drinking at all.[3] In some cases the statistician may calculate a *partial correlation*, this being the correlation between two variables when the effects of one or more related variables are removed.

> "Statistics are used, and misused, in law."

TESTING CAUSATION AS A STATISTICS PROBLEM

In statistics, it is generally accepted that observational studies (like counting cancer cases among smokers) can give clues, but can never actually establish cause and effect. The standard procedure for testing and demonstrating causation here is the *randomised experiment*. For example, how would the relationship between cancer and smoking be tested in order to show the liklihood of causation? In this instance, a large number of people are randomly selected and divided into two groups. One group is forced to smoke while the other goup (the *control group*) would be prohibited from smoking (ideally in a *double-blind* setup). Whether one group developed a significantly higher lung cancer rate than the other would then be assessed. Obviously, for ethical reasons this particular experiment cannot be performed, but the method is widely applicable for similar types of experiments.

A *double-blind* technique describes an especially stringent way of conducting an experiment, usually on living, conscious, human subjects. In such an experiment, neither the individuals nor the researchers know who belongs to the control group. Only after all the data are recorded (and in some cases analysed) may researchers be permitted to learn which group individuals were in. Performing an experiment in such a double-blind manner is a way of reducing the influence of prejudices and unintentional physical cues on the results. In practice, every researcher who interacts with or treats a subject should be blinded, if an experiment is to be designated double-blind.

A particular instance of a double-blind scenario is when the treatment being tested is a drug. The appearance of the actual drug may be simulated with a colored pill or solution (that is, a *placebo*) that looks identical in all respects. However, surgical procedures present more difficult problems, since a surgeon inevitably knows whether it is the procedure or a sham that he or she is performing. The evaluation of such procedures can be 'approximately' double-blind if the researchers responsible for recording subjects' responses and analysing the data are blinded, but such tests are not strictly speaking double-blind. A *single-blind* experiment is designed so that individuals themselves do not know whether they are subjects or members of the experimental control group. However, the researchers knows to which group they belong.

Under certain assumptions, parts of the causal structure among several variables can be learned from full covariance or case data by the techniques of *path analysis* or *Bayesian networks*. In general, these inference algorithms search through the many possible causal structures among the variables, and remove those that are strongly incompatible with the observed correlations. This then leaves a set of possible causal relations, which should then be tested by designing appropriate experiments. If experimental data are already available, the algorithms can incorporate that as well.[4]

CAUSATION IN MEDICINE

There is a wide variety of situations in the medical arena in which a plaintiff attempts to show that a doctor's negligence has 'caused' the plaintiff to suffer some injury.[5] In this instance the problem lies in demonstrating the relationship between the doctor's mistake and the injury itself. It is usually necessary to rely on expert testimony to prove that if the doctor had not ▶

erred, the patient would have recovered or would have been injured to a lesser degree.

Sometimes causation is shown or discovered accidentally without the rigour of the double-blind experiment. Early evidence of this occurred when researchers attempted to find a 'cure' for certain medical ailments. In one such early example, Dr JGH Kramer in 1734 observed a significant correlation between scurvy and the eating of fruit and vegetables.[6] Quite simply, he found that soldiers who ate fruit and vegetables generally did not get scurvy while those who did not, generally speaking, got it. The difficulty was that the correlation was not perfect and there was no evidence as to why this should be so. It was just achieved by observation.

Another interesting example of this arose in the late 18th century when the cause of smallpox was unknown. All that was known was 'who' did not get smallpox. This included rural cattle and dairy workers who had previously had cowpox. This was noticed by the English physician, Edward Jenner, who speculated whether the unknown substance (cowpox pus with an unknown active ingredient), could be injected into people and thus prevent them getting smallpox.[7] On 14 May 1796, he injected a boy of about eight years of age with this mysterious substance and the experiment was successful. In this instance, the problem was solved by looking at who did *not* get smallpox without ever learning *why* the treatment worked. But the result was that smallpox epidemics were eliminated and the foundations of modern immunology as a science were established.

However, if affirmative proof is to be made, a more rigorous analysis is required. But even this can go awry if proper application is not made. One of the more recent famous cases involving the misuse of statistics was that of the UK solicitor, Sally Clark, who in 1999 was convicted of murdering her two infants.[8] Her first child had died in his sleep, aged 11 weeks, and the death was certified as natural causes, with evidence of respiratory infection. However, just 12 months later, when her second child died at the age of eight weeks, Sally was arrested and charged with murdering both children. The defence counsel claimed that both children had died of SIDS.

A paediatrician, Sir Roy Meadow, speaking as an expert witness for the prosecution, claimed that the chance of two children dying of cot death was 1 in 73 million. This would mean that such a double death would occur less often than once every 100 years in England. Meadow based his conclusion on his estimate that the chance of a randomly chosen baby in the socio-economic circumstances of that of Sally dying of SIDS was about 1 in 8,500. He therefore concluded that the chances of *two* such deaths could be obtained by squaring this value. This yields 1/8500 x 1/8500, or about the 1 in 73 million figure quoted, which has since been widely discredited.

This was no doubt powerful evidence for a jury. In the judge's summing up, he told the jury that 'although we do not convict people in these courts on statistics…the statistics in this case do seem compelling'. He added: 'This may be part of the evidence to which you attach some significance.' Sally Clark was convicted and sentenced to life imprisonment. In January 2003 the conviction was quashed on a second appeal, the particular grounds being that crucial medical evidence that would have assisted her case was unavailable

"A famous case involving the misuse of statistics was that of UK solicitor, Sally Clark, who in 1999 was convicted of murdering her two infants."

> "Problems arise when statistics and causation issues become scrambled."

at the time of her original trial or first appeal.[9] Furthermore, Lord Justice Kay, the senior judge hearing the appeal, commented upon the power of the erroneous 1 in 73 million 'statistical evidence' as 'dramatic evidence … that one could confidently expect to have a dramatic impact on the jury'.

Statisticians commenting on the Clark case had been very disturbed from the outset that Meadow had made a serious statistical error that had no doubt influenced the jury, especially given the comments of the trial judge in his summing up. In particular, it prompted a letter from Professor Peter Green, President of the Royal Statistical Society, to the Lord Chancellor that outlined the statistical flaws made at the trial and imploring him 'to ensure that statistical evidence is presented only by qualified statistical experts, as would be the case for any other form of expert'.[10] There has been much written on the statistical errors made in the Sally Clark trial and they serve as an excellent example of just how things can go terribly wrong if they are accepted as fact. They can also have a compelling influence on the weighing up of evidence.[11]

REMARKS

In a legal context, the notion of causation relates to the issue of whether the defendant(s) can be said to have caused the plaintiff's injuries such that they are liable to compensate the plaintiff's losses. Causation provides the link between a finding of fault on the part of the defendant, and the obligation to pay damages to the plaintiff. There are a number of types of causation, including *factual causation* (cause and effect), *legal causation* (relating to the closeness of the connection between the defendant's conduct and the plaintiff's injuries) or the relative responsibility of more than one wrongdoer to compensate the plaintiff's injuries.

There is no end to the type and scope of problems in which the researcher is attempting to show that some kind of causation exists. In a topical example, there are many people who believe that electro-magnetic fields (EMF) are the cause of much human illness.[12] These ailments include childhood leukemia, other cancers, high blood pressures, the aggravation of other diseases, and electrical sensitivity syndrome. Indeed, EMF has become a legal issue as people attempt to seek compensation for alleged injuries and try to have power lines relocated away from their homes. The statistical approach to uncovering any causation effect would be by rigorous testing of the available data, but even this would be a long, painstaking task. Similarly, the link between brain cancer or tumours and the prolonged use of mobile phones is also a very topical issue.

Statistics certainly has a place in law and, while it has proven ways of showing the probability of causation, it also has limitations. Problems arise when statistics and causation issues become scrambled in law. This article has offered some examples of how statisticians consider causation and there have been a number of research papers on these and similar concerns. There have also been a number of texts devoted entirely to the topic.[13] In each case the questions to be dealt with usually involve complex scientific matters in which it is essential that a sound legal decision be made. This is what makes the task so difficult and time-consuming. PL

Endnotes: **1** JS Croucher, 'Assessing the statistical reliability of witness evidence', (2003), *Australian Bar Review*, Vol. 23, 173–83. **2** JS Croucher, *Statistics: Making Business Decisions*, (2002), McGraw-Hill Australia, Sydney, Chapter S8. **3** See RA Johnson and GK Bhattacharyya, *Statistical Concepts and Methods*, 3rd edition, (1996) John Wiley and Sons Inc, New York. **4** See *Wikipedia*, URL: http://en.wikipedia.org/wiki/Causality **5** CSR Limited. Submission to the review of the law of negligence. **6** *The medical research: an overview*, see URL: http://medicolegal.tripod.com/statisticsrhowdocsknow.htm. **7** See RB Fisher, *Edward Jenner: 1749–1823*, (1991), Andre Deutsch Ltd, London. **8** Case No. 1999/07495/Y3. There have been two appeals [2000] EWCA Crim 54; [2003] EWCA Crim 216. On the second appeal the convictions were quashed. **9** [2003] EWCA Crim 216. **10** See URL http://www.rss.org.uk/archive/reports/sc_letter.html **11** See URL http://plus.maths.org/issue21/features/clark/ **12** See EW Campion, 'Power Lines, Cancer and Risk', (1997), *New England Journal of Medicine*, 337, 44–6. **13** HLA Hart and T Honore, *Causation and the Law*, (1985), 2nd ed., Clarendon Press, Oxford. Also R Goldberg, *Causation and Risk in the Law of Torts: Scientific Evidence and Medicinal Product Liability*, (1999), Hart Publishing, Oregon.