DATA COLLECTION AND ANALYSIS IN FORENSIC SCIENCE

David Lucy
University of Edinburgh, United Kingdom
dlucy@maths.ed.ac.uk

The widespread international adoption of DNA technology in forensic science over the last twenty years or so has resulted in some standardised methods of data collection and data interpretation. The impetus generated by the systematic approach characteristic of forensic DNA has carried into other fields of forensic science, typically resulting in forensic scientists wondering whether the same approaches can be applied to their own specialisms. Workers in areas of forensic interest such as ballistics and trace evidence have for some time collected in a systematic manner data connected with those fields. However there are many more areas of forensic science which require large bodies of systematically collected data. Some of these areas are so rarely used in forensic science that the required data is not available, and for a few areas of evidence it is infeasible, if not impossible to collect suitable data.

INTRODUCTION

The one common theme running through forensic sciences is that of heterogeneity. That is the nature of the observations made, and the purpose to which they are put, are specific to a case, and of widely varying entities. For examples, in the last few years forensic scientists have described the systematic collection of decomposition odours (Vass *et al*., 2004), the frequency of MtDNA in the United Kingdom's dog population (Wetton *et al*., 2003) and the transport of human remains in the lower Mississippi (Bassett and Manhein, 2002). These examples are of the rarer types of observation made in connection with forensic sciences, there are many collections of more common types such as DNA data, footmark and foot-print data, and fingermark and fingerprint data.

The problem with forensic science is that in essence it deals with the reconstruction of unusual forms of human behaviour, and as such can require data on any facet of the observable or measurable world. Hence the data used in the forensic sciences is highly disparate and heterogeneous.

INTELLIGENCE AND EVIDENCE

A further complicating factor when considering data in forensic sciences is the use to which those data will be put. Fortunately there are really only two uses, intelligence and evidence.

Intelligence

In terms of Police work intelligence is information which points to other information. That is, intelligence may suggest that a specific person may have committed a crime, so might lead investigators to search that persons place of abode, or ask questions about that persons whereabouts at a specific time. Intelligence may suggest that a certain type of individual has committed a criminal offence, so investigators may focus upon gathering more information about individuals fitting a particular set of criteria. Within limits defined by common sense it does not matter where intelligence comes from, and how it is gathered, so long as it is of use to investigators. Collections of systematically gathered data are commonly used as part of an intelligence effort on the part of investigators. For example, in the United Kingdom the CATCHEM database of past cases of children who have been victims of homicide was employed in November 1999 to suggest possible locations for the remains of Arlene Fraser. Unfortunately Arlene Fraser has not been found to this day, however data from past homicide investigations has been systematically studied for value as a source of intelligence (Francis *et al*., 2004), and the study of these data has now become recommended practice by the Home Office to the United Kingdom Police Forces.

One might have thought that in jurisdictions such as those of the United Kingdom, with a long tradition of forensic science and intelligence led police work there would have also been a similarly long tradition of retaining, and systematically organising data about specific crimes.

Sadly this is not the case. Most of the datasets used for police intelligence are the work of individual officers, supported by their respective forces, rather than some centralised initiative by a government agency such as the Home Office.

The CATCHEM database for instance, despite having detailed data about many of the United Kingdom's child murders since 1960. It is one of the longest running crime databases in the United Kingdom, but was only started a couple of decades ago by Derbyshire Police, and has been maintained by Detective Inspector Chuck Burton. Scotland has a female murder database which has over 1,000 detailed descriptions of the circumstances, crime scene descriptions, and, where available, perpetrator descriptions. This database is currently being used to provide intelligence on possible linkages between cases which are current, and cases in the past for which no perpetrator was ever apprehended. Both these databases are the products of the efforts of particular police officers and police forces and not the Home Office. Until relatively recently this seems to have been the pattern for United Kingdom police and forensic science databases. An officer, or forensic scientist, would compile data peculiar to their specialism, often as part of a single investigation, that database would then prove useful in other investigations, and on an ad hoc and informal basis slowly become a national database.

Recently the Home Office, with the formation of the Serious Crime Analysis Section and National Crime Faculty, has started to administer databases centrally for the whole of the United Kingdom with the instigation of such initiatives as the Homicide Index Database (Mayhew, 2003). The aim of this reform is to pull together many of the different databases which exist into what constitutes a single entity to provide data suitable for Police intelligence and to inform government policy. The question as to why this has not been done before might lie in the historical independence of police and government, and the independence of the various police forces which provide policing in the United Kingdom, however, one possible benefit is that detectives exercising a high degree of control over the databases can design them to be highly suited to intelligence work, rather than the needs of administrators and government.

Evidence

Evidence is a special subset of information about a case, and is that information which is used to prosecute, or exonerate, a defendant. Typically evidence will be information which links an individual to a crime, rather than a type of individual to a crime as might be the case with intelligence. Unlike intelligence, in most jurisdictions, evidence must be gathered within a strict set of regulations and laws. Contravention of these laws in the gathering of any piece of evidence would certainly mean that the evidence would not be accepted in court, and hence would not contribute in any meaningful way to a prosecution, or defence case. Contravention of admissibility rules on the part of the prosecution might also entail the court refusing to hear any element of the case against the defendant, in effect the case itself being ruled inadmissible.

In recent decades the statistical evaluation of evidence has been under intensive development, and with the rise of statistical evidence evaluation there is a concomitant need for reliable and representative data upon which to base any evaluation.

THE EVALUATION OF EVIDENCE

In the past where statistical science has been applied to the evaluation of evidence many different forms of analysis have been used. However, these days a consensus is emerging that the measure of evidential value, known as the likelihood ratio, is the most suitable summary measure of evidential worth. The measure has a number of desirable properties such as an explicit inferential structure, and forcing scientists to confine statements to the substance of their expertise, rather than allowing them to comment on statements concerning the ultimate issue which surround the notions of guilt or innocence.

A likelihood ratio, as its name implies, is simply the ratio of two likelihoods. One is the likelihood of the observations given a proposition which implies a defendant's guilt, the other is the likelihood of the observations given a proposition implying the defendant's innocence (Aitken and Taroni, 2004). A simple illustration might be a case where a crime has been committed, and it is known from an eyewitness that the perpetrator has a tattoo on their left hand. A suspect has been detained who also has a tattoo on their left hand. From of a survey of 100,000 people in the general population 100 individuals had indistinguishable tattoos on their left hands. The probability that a suspect would have a tattoo on the left hand were that suspect the perpetrator is 1 if we neglect the small probability that the tattoo could be made invisible by some removal process. From the population of 100,000 there are 99,999 individuals who were not the perpetrator, and 99 individuals who had tattoos on their left hands who were not the perpetrator. The probability of seeing a tattoo on the left hand of an individual from the population were they not the perpetrator is therefore 99/99,999 ≈ 0:001. A likelihood ratio would therefore be 1/0:001 = 1,000, which means that the observation of a tattoo on the left hand is 1,000 times more likely were the suspect the perpetrator than were any person from the population, other than the suspect, the perpetrator also has a tattoo on their left hand. From of a survey of 100,000 people in the general population 100 individuals had indistinguishable tattoos on their left hands. The probability that a suspect would have a tattoo on the left hand were that suspect the perpetrator is 1 if we neglect the small probability that the tattoo could be made invisible by some removal process. From the population of 100,000 there are 99,999 individuals who were not the perpetrator, and 99 individuals who had tattoos on their left hands who were not the perpetrator. The probability of seeing a tattoo on the left hand of an individual from the population were they not the perpetrator is therefore 99/99,999 ≈ 0:001. A likelihood ratio would therefore be 1/0:001 = 1,000, which means that the observation of a tattoo on the left hand is 1,000 times more likely were the suspect the perpetrator than were any person from the population, other than the suspect, the perpetrator.

From the simple example above, and real cases in the United Kingdom are sometimes as simple as that, it can be seen that the calculated likelihood ratio makes no statement about the probability of the guilt, or otherwise, of the suspect. It only refers to the observation of a tattoo on the left hand, and meets all the demands of the expert witnesses role in United Kingdom jurisdictions.

However there is a price for being able to give a clear, unambiguous and concise evidential value, that is one must know the frequency with which the observation of a tattoo on the left hand would be made in the general population. This places a necessary burden on the forensic science community to collect data from which to estimate the desired frequencies.

DATA COLLECTION FOR EVIDENTIAL EVALUATION

From the tattoo example above there is an obvious requirement for systematically collected, high quality data on objects of forensic interest. Objects of forensic interest vary, both in the ease of data collection and in the type of data being collected.

Univariate Categorical Data

Much evidential evaluation for the criminal courts in the United Kingdom involves the observation of some trait connected with the perpetrator, and a trait which indistinguishable from that of the perpetrator connected with a defendant. These may well include tattoos on the left hand, but there are many other examples, these include: types of footwear where shoe marks have been found at the scene of a crime, types of vehicle where there may be limited registration information, types of clothing where a fragment of fabric has been found at a crime scene, and types of cable ties that have been used as ligatures are all examples where observations of common types of object have been assessed for courts in the United Kingdom. Fortunately most of these objects are manufactured objects, and data on the numbers of any specific object is usually available from the makers or importers of those commercial objects. It is more difficult to find the size of the population from which the specific object has been taken. For instance an importer may say that 12,000 pairs of a particular shoe type were sold, but how many pairs of shoes should be considered when estimating the frequency of the particular shoes? Fortunately data is usually available from manufacturers' and importers' organisations to make some estimate of the population size. However, there are instances, such as cable-ties improvised for use as ligatures, where we can know the numbers of a particular cable tie type, and the numbers of all cable ties, but the relevant population is all objects which could be used a ligatures. This population might include old bits of rope, electrical flex and objects of clothing. As we can never know how many old bits of rope suitable for use as ligatures exist, then some deliberately conservative estimate of the population is usually made.

Multivariate Categorical Data

Some forensically relevant observations of everyday objects can become problematic for the investigator. This is especially true where categorical data may be multivariate. A well known example is a case from the United States of America. In *People v. Collins* (Fairley and Mosteller, 1977, p. 355-379) the case was one of street robbery where an elderly lady was pushed to the ground in a Los Angeles alleyway. She managed to see a young man running from the scene. She then noticed her purse missing. Another man at the end of the alleyway noticed a woman who he described as Caucasian, with her hair in a blond ponytail, run out of the alleyway and into a yellow car parked on the other side of the road. The car was being driven by a black man with a moustache and beard.

A couple who fitted this description were arrested and brought to trial. An expert witness, described as an instructor in mathematics at a state college instructed the court to think of all the characteristics as being independent.

The probabilities assumed during the trial for each of these traits were: yellow car 1/10, man with moustache 1/4, girl with ponytail 1/10, girl with Blonde hair 1/3, black man with beard 1/10, mixed race couple in car 1/1000.

The probabilities given above are not taken from survey data, but instead were provided by the instructor who also testified to the third law of probability for independent events. The jury were instructed to insert their own estimates of the frequencies if they thought the frequency to be wrong. The calculated probability of finding at random suspects matching all the criteria above was about 1 in 12 million. This may seem a fairly convincing analysis, but as we know, it is necessary to examine the interrelationships between observations which may be dependent on each other, and the true probability for observing these traits may be somewhat higher than 1 in 12 million. The problem is where does one find data on the frequency for all the characteristics in question? It is extremely unlikely that any pre-existing survey could supply an estimate of the frequency, so a survey may have to be constructed especially for the specific case. This would entail much expenditure of public money, an expenditure which, unlike the databases of more common forensic materials, would be unlikely to be used again.

Continuous Data

Trace evidence from objects such as glass and individual fibers are specialisms in the forensic sciences in their own right. As these types of materials are a feature of many criminal cases in which forensic science is used so forensic scientists have large existing datasets, usually organised at a regional level, available to them.

Usually the observations on these items are multivariate and continuous, although it is still traditional to record refractive index as the sole measurement from some forms of glass. It is only recently that likelihood ratio methods have been applied to these types of measurements (Curran, 2000; Aitken *et al*., 2006). The methods for evidence evaluation for continuous data were first suggested by Lindley (1977). The calculation of the probability of any observation given the proposition that two objects come from the same source, require that the forensic scientist estimate the within object variability. This means that replicate measurements have to be made for all objects which appear in the dataset. For example, for a database of elemental measurements from glass one would have to include window glass, glass from cars and glass from bottles jars and other containers. Then each of the specimens selected to be represented in this database would have to have a number of sub-specimens measured, that is the bottle or window pane would have to be broken up, some fragments selected, and the elemental compositions of these fragments measured.

Scientists have for some time been aware of the need to estimate within object variability as part of systematic data collection. The Federal Bureau of Investigation in the United States has systematically collected data on glass which includes replicate measurements (Koons and Buscaglia, 2002), as has the Institute of Forensic Research in Poland. These efforts are worth making for collections of data which are going to be useful to forensic scientists over a long period of time.

CONCLUSIONS

The fact that police investigators require organised and structured data for intelligence and evidential purposes, and that forensic scientists have a similar data requirement so that they may evaluate their evidence is beyond dispute. However, beyond this obvious need no generalisations hold. The types of data, and the observations to which those data relate take on no particular universal form as, except in some special cases where object types are a repeated feature of forensic science, the data needs are very heterogeneous, and it cannot be predicted in advance of any specific investigation which may be required.

Of those data types which are a constant feature of police intelligence and forensic investigation data are available, and in the United Kingdom at least, are now beginning to be integrated into larger data entities which will aid both government, and investigators. However, there will always be observations connected with any specific investigation which will require the ad hoc assembly of data. For some observations, such as highly multivariate observations of everyday characteristics seen in People v. Collins it is improbable that any data will be available which will fully meet the requirements of statistical evidence evaluation.

Finally a mention should be made of the numerous DNA databases which are being constructed on a regional basis throughout the world. The observations of DNA profiles are structured in such a way that they are multivariate and discrete. There is no within person variability, and all the loci represented are selected in such a way that they are independent. These databases present, at least in principle, few problems from the point of view of data collection and evaluation. They are used for intelligence, and evidence, and represent some form of epistemological ideal in terms of structure, ease of use, and interpretive power.

REFERENCES

Aitken, C. G. G., Lucy, D., Zadora, G. and Curran, J. (2006). Evaluation of trace evidence for three-level multivariate data with use of graphical models. *Computational Statistics and Data Analysis*, 50(10), 2571-2588.

Aitken, C. G. G. and Taroni, F. (2004). *Statistics and the Evaluation of Evidence for Forensic Scientists* (2nd edition). Chichseter: John Wiley and Sons.

Balding, D. J. (2002). The DNA database search controversy. *Biometrics*, 58(1), 241-244.

Bassett, H. E. and Manhein, M. H. (2002). Fluvial transport of human remains in the lower Mississippi River. *Journal of Forensic Sciences*, 47(4), 719-724.

Curran, J. M., Hicks, T. N., and Buckleton, J. S. (2000). *Forensic Interpretation of Glass Evidence*. Boca Raton: CRC Press.

Fairley, W. and Mosteller, W. B. (1977). *Statistics and Public Policy*. Reading, MA: Addison-Wesley.

Francis, B., Barry, J., Bowater, R., Miller, N., Soothill, K., and Ackerley, E. (2004). Using homicide data to assist murder investigations. London: Home Office Online Report 26/04.

Koons, R. D. and Buscaglia, J. (2002). Interpretation of glass composition measurements: the effects of match criteria on discrimination capability. *Journal of Forensic Sciences*, 47, 505-512.

Lindley, D. V. (1977). A problem in forensic science. *Biometrika*, 64, 207-213.

Mayhew, P. (2003). Review of information on homicide: Discussion document. *National Statistics Quality Review Series,* 25. London: The Home Office.

Vass, A. A., Smith, R. R., Thompson, C. V., Burnett, M. N., Wolf, D. A., Synstelien, J. A., Dulgerian, N. and Eckenrode, B. A. (2004). Decompositional odor analysis database. *Journal of Forensic Science*, 49(4), 760-769.

Wetton, J. H., Higgs, J. E., Spriggs, A. C., Roney, C. A., Tsang, C. S. F. and Foster, A. P. (2003). Mitochondial profiling of dog hairs. *Forensic Science International*, 133(3), 235-241.