MEMORANDUM

No 30/2003

A Finer Point in Forensic Identification.

Halvor Mehlum

ISSN: 0801-1117

Department of Economics
University of Oslo

This series is published by the **University of Oslo** **Department of Economics** P. O.Box 1095 Blindern N-0317 OSLO Norway Telephone: + 47 22855127 Fax: + 47 22855035 Internet: http://www.oekonomi.uio.no/ e-mail: econdep@econ.uio.no	In co-operation with **The Frisch Centre for Economic Research** Gaustadalleén 21 N-0371 OSLO Norway Telephone: +47 22 95 88 20 Fax: +47 22 95 88 25 Internet: http://www.frisch.uio.no/ e-mail: frisch@frisch.uio.no

List of the last 10 Memoranda:

No 29	Svenn-Erik Mamelund Effects of the Spanish Influenza Pandemic of 1918-19 on Later Life Mortality of Norwegian Cohorts Born About 1900. 30 pp.
No 28	Fedor Iskhakov Quasi-dynamic forward-looking model for joint household retirement decision under AFP scheme. 59 pp.
No 27	Henrik Wiig The Productivity of Social Capital. 22 pp.
No 26	Tao Zhang Identifying treatment effects of active labour market programmes for Norwegian adults. 45 pp.
No 25	Tao Zhang A Monte Carlo study on non-parametric estimation of duration models with unobserved heterogeneity. 89 pp.
No 24	Karine Nyborg and Kjetil Telle The Role of Warnings in Regulation: Keeping Control with Less Punishment. 31 pp.
No 23	Maria Kalvarskaia Savings behaviour when households havean access to occupational pensions. 50 pp.
No 22	Rolf Golombek and Michael Hoel Climate Policy under Technology Spillovers. 37 pp.
No 21	Jo Thori Lind Fractionalization and the size of government. 50 pp.
No 20	Michael Hoel, Tor Iversen, Tore Nilssen and Jon Vislie Genetic testing and repulsion from chance. 20pp.

A complete list of this memo-series is available in a PDF® format at:
http://www.oekonomi.uio.no/memo/

A Finer Point in Forensic Identification*

Halvor Mehlum†

Abstract

In this note I bring a new aspect into the so called Island Problem. Given that only cases where there is a suspect reaches the court, what is the consequence for the probability of guilt? I find that it indeed matters for the results that court cases are selected in this way. The analysis illustrates the general point that the exact protocol by which data are generated is an essential part of the information that should be used when analyzing data.

1 Introduction

In California in 1968 a couple with certain characteristics committed a robbery. The Collins couple, who had identical characteristics, was arrested and put on trial. In court the prosecutor claimed that the characteristics would appear in a randomly chosen couple with the probability 1/12,000,000 and that the couple thus had to be guilty. The suspected couple was convicted but, based on an argument that was closer to the mark (but not enterily correct), the California Supreme Court later overturned the conviction.

*I am grateful to Tore Schweder, to partcipants at the Joseph Bell Workshop in Edinburgh and to seminar participants at the University of Oslo.

†Department of Economics, University of Oslo P.O. Box 1095, Blindern N-0317 Oslo, Norway. E-mail: halvor.mehlum@econ.uio.no.

The Collins case and a stylized version of it, the "Island problem", has a central place in the forensic statistics literature. In its stylized form it can be formulated as a beads in an urn problem: An urn contains a known number of beads with different colours. Then a crime happens: A bead is drawn, its colour, which happens to be red, is observed. The bead then escapes back to the urn. Now, a search for a red bead is carried out and a read bead is indeed found. What is the probability that the first and the second red beads are identical? That is, what is the probability that the suspected bead is really the guilty bead? In the literature, one central theme is how to properly extract information from the circumstances of the case and the search protocol used when finding the suspect. Central contributions starts with Yellin (1979) and Eggleston (1983) and continue with Dawid (1994), Balding and Donnelly (1995), and Dawid and Mortera (1996). The authors show how changes in the assumption regarding the search procedure change the results. For example, it matters for the analysis whether the search is a full search or not. These points in addition to several other refinements, like unreliable evidence, are taken well care of in these papers.

In this note I focus on a new issue. In short, my question is as follows: Given that the search procedure rarely produces a suspect, and assuming that for each case with a successful search there are several cases with unsuccessful searches, what is then the appropriate analysis of the problem? As in the other contributions, I consider a highly stylized and abstract version of the case. In contrast to the other authors I include in the analysis the fact that only a selected subset of criminal cases reaches the court. (We are never asked to analyze the probability of guilt faced with a suspect *that does not* fit the characteristics.) I will explain this argument using the helpful 'beads in an urn' analogy. I will first present two central solutions from the literature; based on full search, and on random search. I will then show how the solutions changes when taking account of the fact that we are faced with a selected case.

2 The "beads in an urn"

Consider an urn containing N beads. The N beads in the urn may be a sample from a underlying population with M different colours. Now, a robbery happens: One bead, the robber-bead, is drawn at random from the urn. All beads are equally likely to be drawn. The colour of the robber-bead is observed and then put back in - it happens to be red. Let this part of the evidence be denoted $F_1 =$ 'bead drawn at random is red'.

The question now is: How should F_1 affect our belief regarding the number of red beads in the urn? When the bead we draw at random is seen to be red, we will adjust our belief about the likely number of red beads in the urn. The only knowledge we have before picking a red bead is that the number of red beads X is distributed $\mathrm{bin}\,(N, p)$ with N and p known. Given the evidence F_1 the distribution of X may be updated using Bayes' formula:

$$P\left(X = n|F_1\right) = P\left(X = n\right) \frac{P\left(F_1|X = n\right)}{P\left(F_1\right)} = \binom{N}{n} p^n \left(1-p\right)^{(N-n)} \frac{n/N}{p} \tag{1}$$

By re-arranging,

$$P\left(X = n|F_1\right) = \binom{N-1}{n-1} p^{n-1} \left(1-p\right)^{N-n-(n-1)} \tag{2}$$

which is the unconditional probability of there being $n-1$ red beads among the $N-1$ non-robber beads.

Now a search for a red bead is conducted. I will first discuss two possible search procedures: Search until success, which is the implicit assumption in Yellin (1979), and random search, which was introduced by Dawid (1994).[1]

2.1 Search until success

A search through the urn is carried out until a red bead, the suspected bead, is found. If no record is made of the time used or the number of beads screened, no additional evidence is gained in the process of finding the suspected bead. This search protocol always produces a suspect and the question of guilt G is the question of whether the suspected bead is identical to the robber-bead. The larger the number of red beads, the lower is the probability of guilt. For a given number of red beads, $X = n$, the probability of guilt is $1/n$. Using 2, the probability of guilt is

$$P(G|F_1) = \sum_{n=1}^{N} \frac{1}{n} \binom{N-1}{n-1} p^{n-1} (1-p)^{(N-n)} = \frac{1-(1-p)^N}{Np} = \frac{1}{E(X|X \geq 1)}$$

Hence, the probability of guilt $P(G|F_1)$ has a simple solution which happens to be identical to the inverse of $E(X|X \geq 1)$.

This solution is different from the solution following from the Supreme Court's erroneous argument. The Supreme Court's interpretation of the evidence can be formulated as F_0 ='there is at least one read bead'. In that case the probability of guilt is $P(G|F_0) = E(X^{-1}|X \geq 1)$. Given this particular relationship between the expressions for $P(G|F_1)$ and $P(G|F_0)$ it follows from Jensen's inequality that $P(G|F_1) < P(G|F_0)$. Hence compared to Yellin the Supreme Court overstated the probability of guilt.

[1]Balding and Donnelly (1995) and Dawid and Mortera (1996) discus several varieties of these search strategies.

2.2 Random search

By design, the above search always produces a suspect and provides no additional information about the distribution of X. Another alternative is the random search where only one bead is picked at random. If this bead is not red it is definitely not guilty. If it is red, however, the evidence is F_2 = 'second bead drawn at random is red', and the distribution of X should be updated accordingly. Using Bayes' formula we get:

$$\begin{aligned} P(X=n|F_1 \cap F_2) &= P(X=n|F_1)\,\frac{P(F_2|X=n \cap F_1)}{P(F_2|F_1)} \\ &= P(X=n|F_1)\,\frac{n}{1+(N-1)\,p} \end{aligned} \tag{3}$$

The probability of guilt is now

$$P(G|F_1 \cap F_2) = \sum_{n=1}^{N} \frac{1}{n}\frac{n}{1+(N-1)\,p} P(X=n|F_1) = \frac{1}{1+(N-1)\,p} \tag{4}$$

$$P(G|F_1 \cap F_2) = \frac{1}{E(X|F_1)} \tag{5}$$

As $P(G|F_1) = E(X^{-1}|F_1)$ again it follows from Jensen's inequality that $P(G|F_1 \cap F_2) < P(G|F_1)$. The probability of guilt is thus even lower in the random search than in full search.

This updating assumes that we are faced with *one* experiment, where two beads drawn with replacement by chance happen to be red. From (5) it follows that the probability that the second bead is red, given F_1, can be written as

$$E(X/N|F_1) = \frac{1}{N P(G|F_1 \cap F_2)}$$

This probability is less than one, and may be very low in realistic cases. With the random search protocol and if only cases with successful searches are taken to court, there will be an interesting selection of cases. I will in the following show how the analysis changes when taking account of the possibility that we are faced with a selected case.

2.3 Random search until case with suspect.

As before consider an urn containing N beads drawn from an underlying population of M colours. Now, one of a never ending series of robberies happens, one bead is drawn and put back in. Then a second bead is drawn. If the beads are not the same colour, the case is closed and the next potential case occurs. A potential case is characterized by the the joint frequency of beads by colour X_i, $\sum_{i=1}^{M} X_i = N$, and the draw of two beads with replacement. The iteration of cases lasts until there is a case with a suspected bead of the same colour as the first.

Assume that after an unknown number of iterations (i.e. potential cases) there is a case where the first and second beads are of the same colour. Assume further that the two beads are red. When this case arrives the question is: Are the beads identical? Let the evidence be denoted $F_3 =$ 'First case with two similar beads involves two red beads'. Let X_i, $i \in [1, M]$, denote the number of beads of colour i, where $i = 1$ is the colour red, then Baye's formula gives

$$\begin{aligned} P\left(X_1 = n | F_3\right) &= P\left(X = n\right) \frac{P\left(F_3 | X_1 = n\right)}{P\left(F_3\right)} \iff \\ P\left(X_1 = n | F_3\right) &= P\left(X = n\right) \frac{E\left(n^2 / \sum_{i=1}^{M} X_i{}^2 | X_1 = n\right)}{E\left(X_1{}^2 / \sum_{i=1}^{M} X_i{}^2\right)}, \end{aligned} \tag{6}$$

This expression can generally not be simplified further as the simultaneous distribution of all the M colours is involved. In order to get an idea on how updating based on F_3 compares to updating based on $F_1 \cap F_2$ I will look at some numerical illustrations.

2.4 Numerical illustrations

Assume that the underlying distribution amounts to a multinomial distribution where p_i is the probability of colour i. I will consider two cases

1. The Handy case, where $N = 5$ and $P\,(\text{red bead}) = p_1 = 1/6$
2. The Island problem, where $N = 100$ and $P\,(\text{red bead}) = p_1 = 0.004$

for various values of M and p. The parameters of Island problem is taken from Eggleston (1983). For each of these cases I look at three different sets of assumptions regarding the parameters of the multinomial distribution.

i) *Red and green*, where $M = 2$ and $p_2 = 1 - p_1$

ii) *Red and palette*, where M is enormous and $p_2 = \cdots = p_M = (1 - p_1)\,/(M - 1)$

iii) *All equal*, where $p_1 = p_2 = \cdots = p_M$ and $M = 1/p_1$

The calculation of the results for i) and ii) are straight forward. The calculations for iii) are more demanding. Note however that as all colours in that case have the same probability there is no information in what exact colour is involved in the case. By construction the parameters of the Handy case makes it identical to the game of poker dice[2] as analyzed in Epstein (1967 p154) and the essential probabilities can be taken from there. For the Island problem some CPU time is needed.[3] The results for the probability

[2] Poker dice is a variant of yatzee with one throw and seven hands: no two alike, one pair, two pairs, three alike, full house, four alike, and five alike.

[3] The complete calculations involve integrating over all 190 mill. partitions of the number 100. I approximate by integrating over the 600,000 partitions with the highest probability. These account for $1 - 10^{-6}$ of the probability mass.

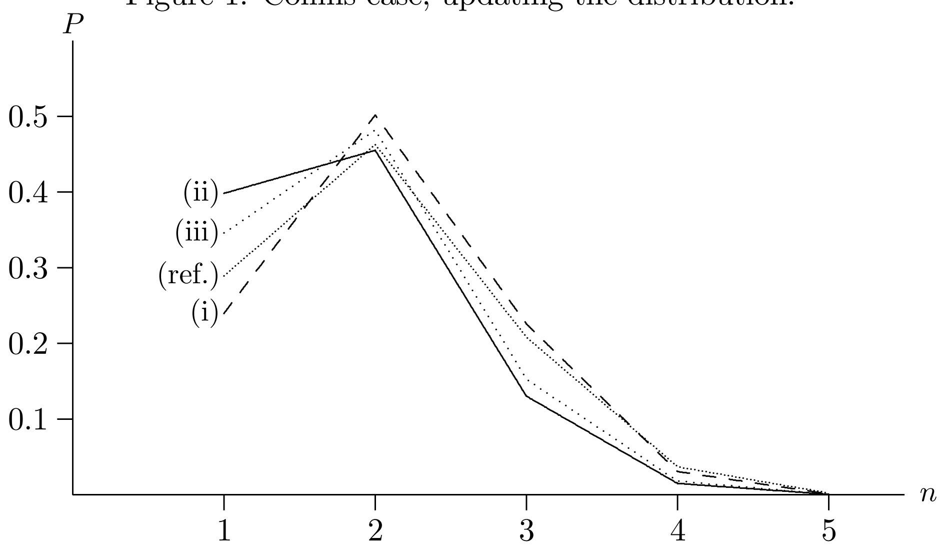
of guilt are summarized in Table 1. The table also includes the result of random search $P(G|F_1 \cap F_2)$ as a reference (ref.).

Table 1: Probability of guilt, $P(G|F_3)$ in the iterated case and in the reference.

		Handy case	Island problem
(i)	Red and green	.574	0.712
(ii)	Red and palette	.674	0.722
(iii)	All equal	.643	0.719
(ref.)	Random search, $P(G\|F_1 \cap F_2)$	.600	0.716

The results show that the probability of guilt $P(G|F_3)$ indeed varies with the distributional assumption regarding all the colours. The probability of guilt may be above or below the reference case of random search $P(G|F_1 \cap F_2)$. Some additional insight can be gained by looking at the distribution for the updated distribution of X_1. For the Handy case the updated distribution $P(X_1 = n|F_1)$ is given in Figure 1 . Consider $n = 1$, here

Figure 1: Collins case, updating the distribution.



'Red and palette' have the highest probability while 'Red and green' have the lowest. The reason for 'Red and palette' to be at the top is that the likelihood of F_3 = 'First case with two similar beads involves two red beads' is quite high in the case of one red bead and four others of different colours. The reason for 'red and green' to be at the bottom is that the likelihood of F_3 is quite low in the case of one red and four green beads. For

$n \geq 2$, however, this order is turned on its head. For all n the 'All equal' is in-between the two extremes. The random search updating (ref.) follows its own distinct path.

Comparing the Handy case with the Island problem it seems clear that the difference between the cases declines with N. This feature can be confirmed when looking at (6). To put it loosely: when the number of beads is large and when there is a large number of colours, each with small probabilities, then $\sum {X_i}^2$ can be treated as a constant independent of X_1 for small X_1. Then, it follows that (6), for small n, can be simplified as follows

$$P\left(X_1 = n | F_3\right) \approx P\left(X_1 = n\right) \frac{n^2}{E\left(X_1^2\right)} = P\left(X_1 = n | F_1 \cap F_2\right) \tag{7}$$

the last equality follows as

$$P\left(X_1 = n | F_1 \cap F_2\right) = P\left(X_1 = n\right) \frac{P\left(\left(F_1 \cap F_2\right) | X_1 = n\right)}{P\left(F_1 \cap F_2\right)} = P\left(X_1 = n\right) \frac{n^2/N^2}{E\left(X_1^2/N^2\right)}$$

The approximation 7 is accurate when 'red' is quite rare and when there is a large number of other features in the population. That the feature 'red' is rare is an implicit condition for the problem to be relevant in the context of evidence in a court case. Hence, empirically the consequence of taking account of the selection is likely to be modest. This may be comforting as there would be gigantic challenges involved if one where to determine the simultaneous distribution of all possible characteristics in the population of possible culprits.

3 Discussion

The analysis illustrates the general point that the interpretation of statistical data must take into account exactly how the data was collected. What is the protocol and where does it end? In the solution to the Island problem Yellin (1979) only considered the part of the evidence relating to the fact that the guilty happened to have certain characteristics. Later Dawid(1994) brought into the picture the fact that search had produced a suspect with the same characteristics. Both Yellin and Dawid start from the premise that the case is given. My own argument is that the case may itself be selected in a stochastic process. Only cases where there is a suspect with identical characteristics as the guilty is brought before the court.

The protocol I discuss is one of several possibilities. Another palatable assumption could be: Only cases where there is a suspect with identical characteristics as the guilty *and* where the characteristics are rare is brought before the court. If this was part of the protocol the solution would be affected. If, for example, it also happened to be the case that red was the only colour considered to be 'rare', then a court case with a red suspect would amount to the information 'there is at least one read bead'. This information happens to be identical to what the Supreme Court, though erroneous and via a different argument, extracted from the circumstances of the Collins case.

References

Balding; David J. and Peter Donnelly (1995) "Inference in Forensic Identification" *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 158 (1), 21-53.

Dawid, A. P. (1994) "The Island Problem: Coherent Use of Identification Evidence" *Aspects of Uncertainty. A Tribute to D. V. Lindley*, Wiley, New York, 159–170.

Dawid, A. P. and J. Mortera (1996) "Coherent Analysis of Forensic Identification Evidence" *Journal of the Royal Statistical Society. Series B (Methodological)*, 58 (2), 425-443.

Eggleston, R (1983) *Evidence, proof and probability 2nd ed.* , London: Weidenfeld and Nicolson.

Epstein, R. A. (1967) *The Theory of Gambling and Statistical Logic*, Academic Press, New York.

Turner, T. R. (2002) "The Collins Case: A 'Beads in Urns' Model" *The Mathematical Scientist*, 27, 80-84.

Yellin, J. (1979) "Review of ′Evidence, Proof and Probability (by R. Eggleston)" *Journal of Economic Literature,* 17, 583-584.