Comment

Kenneth Lange

1. GENERAL REMARKS

Let me congratulate Berry on a provocative and entertaining paper. The Bayesian perspective provides a rational framework for overcoming some of the inconsistencies of previous analyses. A particularly novel feature of Berry's approach is how the normal kernel estimates of the prior combine with the normal error distributions of the log band weights to give the simple formula (9) for the likelihood ratio R. I like the term "identity index" coined by Devlin, Risch and Roeder (1991) for the amended ratio (10). Having such straightforward formulas as (9) and (10) should promote the Bayesian cause in this forensic context. I agree with Berry that as a safeguard the kernel estimate should include the suspect. With this safeguard, a low value of his constant b seems reasonable. In my opinion, taking b as high as 5 distorts the fitted histograms too much. I also agree with Berry that band shifting should be correctable by regression against monomorphic probes of known band weights. This last issue is discussed at length in a recent paper by McNally, Baird, McElfresh, Eisenberg and Balazs (1990).

Berry does not dwell on possible departures from assumptions. There are two key independence assumptions. One, Hardy–Weinberg equilibrium, requires that the maternal and paternal bands of a person be independently and identically distributed at each genetic locus analyzed. (An allele is one of the finite number of possible qualitative variants at a genetic locus.) Because of band overlap, neither the number nor the population frequencies of the underlying alleles for the DNA fingerprinting loci are known. Despite these uncertainties, the concerns expressed by Lander (1989a) about violations of Hardy–Weinberg equilibrium have largely been laid to rest by Devlin, Risch and Roeder (1990). There is little a priori reason to expect significant departures since equilibrium is reached in a single generation in a well-mixed population.

The issue of linkage equilibrium is more vexing. This has to do with the independence of alleles, and consequently bands, between loci. If linkage disequilibrium holds, then knowing, say, a person's maternal allele at one locus tells us something about his maternal allele at a second locus. Even with perfect mixing of two ancestral populations, the departure from equilibrium is at most halved each generation. This optimal convergence rate holds for loci on different chromosomes. For closely spaced loci on the same chromosome, the convergence rate can be much slower. I am not inclined to go so far as Cohen (1990) in questioning the whole enterprise of DNA fingerprinting because of the potential lack of linkage equilibrium. In well-established racial groups, it should not be an issue. However, there is a clear need for further research to estimate the extent of linkage disequilibrium in typical American populations and to correct for it in the forensic calculations.

Berry tends to gloss over a few other complications. For instance, deciding the right reference population for the blood on Castro's watch more than just complicates the algebra. Presumably, one needs to define some kind of prior for weighting the various racial groups who might contribute such blood. Was Castro a gang member or habitual criminal? If so, then the evidence presented in the trial might sway the jury towards one prior rather than another. The same question arises in paternity calculations. Berry would probably agree that visual examination of the child might fail to establish with certainty the race of its father. Although these ambiguities can be captured in an appropriate Bayesian framework (Devlin, Risch and Roeder, 1991), I share with Geisser (1990) anxiety about the ability of judges and juries to adjust posteriors to priors. Nevertheless, these complications should not be allowed to obscure the strength of the Bayesian approach.

2. COMPARISONS WITH LESS POLYMORPHIC MARKERS

On a more fundamental biological level, I question why forensic laboratories are so enamored with the highly polymorphic VNTR loci. The large number of alleles present at such loci must be balanced against the uncertainties introduced by not being able to determine genotype qualitatively. For loci with fewer alleles, genotype is unambiguous and inferences are less subject to the doubts engendered by complicated modeling assumptions. For example, it is far easier to check Hardy–Weinberg and

Kenneth Lange is Professor and Chair, Department of Biomathematics, School of Medicine, University of California, Los Angeles, Los Angeles, California 90024.

linkage equilibrium for co-dominant loci with just a few alleles. (Co-dominance means that one allele never masks the presence of the second allele at the same locus. Hence, both alleles can be uniquely determined.) The information lost by going to simpler loci can be easily recovered by including more loci. If these loci are well chosen, then the number of tests might at most double or triple. Because of the importance of most criminal and civil litigation using DNA fingerprinting, the objection of greater cost is not compelling. The fragility and paucity of the crime sample furnishes a better argument, but with the use of the polymerase chain reaction (PCR) it is possible to amplify incredibly small amounts of DNA. It is even possible to test a single human sperm for its DNA content (Li et al., 1988). Since each sperm cell possesses only a single allele at a given locus, sperm typing incidently provides a method for avoiding the coalescence of bands seen in heterozygotes with closely spaced band weights.

Recent papers by Litt and Luty (1989) and Weber and May (1989) introduce classes of DNA polymorphisms based on TG or CA repeats. These loci are not as polymorphic as the current loci, but they do present large numbers of alleles that the above authors claim can be scored qualitatively. My contention about remedying the loss in information can be illustrated by computing the exclusion probability for a locus with n alleles. This is the probability that a random criminal and a random, unrelated suspect do not share matching genotypes; it measures a locus' power to exonerate an innocent person. Suppose the ith allele at the locus has population frequency p_i. If all the alleles are co-dominant, then it follows that the exclusion probability is

$$\begin{aligned} e &= \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} 2p_ip_j(1-2p_ip_j) + \sum_{i=1}^{n} p_i^2(1-p_i^2) \\ &= 1 - 2\sum_{i=1}^{n}\sum_{j=1}^{n} p_i^2p_j^2 + \sum_{i=1}^{n} p_i^4 \\ &= 1 - 2\left(\sum_{i=1}^{n} p_i^2\right)^2 + \sum_{i=1}^{n} p_i^4. \end{aligned}$$

In the first formula above, the double sum corresponds to the criminal being heterozygous while the single sum corresponds to the criminal being homozygous. It can be shown that e is maximized when all $p_i = 1/n$. In this case $e = 1 - 2/n^2 + 1/n^3$. The nonexclusion probability $e^* = 1 - e = 2/n^2 - 1/n^3$ can be compared to the nonexclusion probability $e_1^*e_2^*$ afforded by two independent loci with n_1 and n_2 alleles, respectively. By the argument already given, e_i^* has maximum value $e_i^* = 2/n_i^2 - 1/n_i^3$. If we choose $n_1 = n_2 = \sqrt{n}$, then

$$e_1^*e_2^* = \left(\frac{2}{n} - \frac{1}{n^{3/2}}\right)^2 = \frac{4}{n^2} - \frac{4}{n^{5/2}} + \frac{1}{n^3},$$

and the maximum nonexlusion afforded by two independent loci is less than twice that of the more polymorphic single locus. For example with $n = 16$, $e^* \simeq .008$ and $e_1^*e_2^* \simeq .012$.

The same conclusion can be reached by computing the Kullback–Liebler information K under the assumption that the suspect is guilty (Chernoff, 1979). This expected loglikelihood is defined for a locus with n co-dominant alleles by

$$\begin{aligned} K &= \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} 2p_ip_j \log\left[\frac{2p_ip_j}{(2p_ip_j)^2}\right] \\ &\quad + \sum_{i=1}^{n} p_i^2 \log\left[\frac{p_i^2}{(p_i^2)^2}\right] \\ &= -\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} 2p_ip_j \log(p_ip_j) \\ &\quad - \log(2)\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} 2p_ip_j - \sum_{i=1}^{n} p_i^2\log(p_i^2) \\ &= -\sum_{i=1}^{n}\sum_{j=1}^{n} p_ip_j(\log(p_i)+\log(p_j)) \\ &\quad - \log(2)\left(1 - \sum_{i=1}^{n} p_i^2\right) \\ &= -2\sum_{i=1}^{n} p_i\log(p_i) - \log(2)\left(1 - \sum_{i=1}^{n} p_i^2\right). \end{aligned}$$

Note the close relationship of K to Shannon information. The maximum of K occurs when all $p_i = 1/n$. At this point

$$K = 2\log(n) - \log(2)\left(1 - \frac{1}{n}\right).$$

Since Kullback–Liebler numbers add for independent experiments, the combined Kullback–Liebler information for two independent loci with n_1 and n_2 alleles, respectively, has maximum

$$\begin{aligned} K_1 + K_2 &= 2\log(n_1) - \log(2)\left(1 - \frac{1}{n_1}\right) \\ &\quad + 2\log(n_2) - \log(2)\left(1 - \frac{1}{n_2}\right). \end{aligned}$$

Again when $n_1 = n_2 = \sqrt{n}$,

$$K_1 + K_2 = 2\log(n) - 2\log(2)\left(1 - \frac{1}{\sqrt{n}}\right).$$

For $n = 16$, the maxima are $K = 4.90$ and $K_1 + K_2 = 4.51$. Thus, there is only a slight loss in information in using two independent loci.

To summarize verbally the above mathematics, a locus with n alleles is maximally informative when all alleles are equally frequent in the underlying population. The condition of equally frequent alleles is admittedly extreme, but one might approximate it by the appropriate choice of test loci. Under the assumption of maximum information, two loci with $\sqrt{n}$ alleles each are jointly about as informative as a single locus with n alleles. If the more informative single locus suffers from band overlap, its information content is diminished. One can take the approach of Berry and try to extract the most information by dealing with the quantitative measures directly. By comparison, the FBI's method of preset bins loses some information in discretizing the problem. According to the above mathematical argument, these bins should be equally probable rather than equally spaced.

In closing, let me stress that the rapid rate of innovation in molecular genetics is apt to overcome the technical problems such as band overlap associated with the current DNA fingerprint loci. These loci are typed by a procedure called Southern blotting. The alternative PCR techniques advocated by Weber and May (1989) and Budowle, Chakraborty, Giusti, Eisenberg and Allen (1991) are exquisitely sensitive to minute amounts of DNA and can avoid the problem of band overlap. However, PCR sensitivity can be so extreme that contamination by exogeneous DNA is troublesome. It is not now clear which technology will prevail. I prefer the PCR technology since it permits more loci to be typed from a small crime sample. As I have attempted to argue, most of the controversies over Hardy- Weinberg equilibrium and, particularly, linkage equilibrium will dissipate with better defined loci. In any event, we should welcome the inevitable improvements in technology even if the statistical issues become less interesting. Justice will be better served by greater genetic clarity.

ACKNOWLEDGMENTS

Research support in part by the University of California, Los Angeles; Harvard University; and USPHS Grant CA 16042.

Comment

Herman Chernoff

1. INTRODUCTION

Berry sets two objectives in his abstract. One is to introduce the Bayesian approach to the forensic use of DNA evidence, and the other is to compare that approach with that of "match/binning." The latter is criticized as giving results that are too extreme and for failing to distinguish, in principle, between results that barely fail to fall in the appropriate bin and those that are way out. As he points out, two potential observations that are very close to one another could lead to drastically different conclusions. As I understand it, he seems to suggest that the users of this approach may have recognized this problem in the Castro case and reacted by an ex post facto widening of the bin when the observation barely failed to fall in a bin suggesting guilt.

In my opinion the Bayesian approach is well suited for this subject and deserves to be developed as a useful tool. This approach has several difficulties, some of which are addressed by Berry. One of these is that of educating the members of the legal system and the potential jurors. Another is the use of density estimation to determine the frequency distribution of band weights.

Several issues will be discussed here. The match/binning approach, as described, doesn't make much inferential sense, and if the Bayesian approach should be compared with something, it should be with a more or less classical significance or frequentist Neyman–Pearson (NP) competitor. While binning does replace a continuous analysis by a discrete analysis that leads to aggravating discontinuities, one ought to evaluate the resulting cost in loss of efficiency before outlawing the practice of binning. To do so, we shall review briefly

Herman Chernoff is Professor of Statistics, Harvard University, Science Center, 1 Oxford Street, Cambridge, Massachusetts 02138.