Forensic Science International 214 (2012) 189–194

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

The use of the likelihood ratio for evaluative and investigative purposes in comparative forensic handwriting examination

F. Taroni [a,*], R. Marquis [a], M. Schmittbuhl [b], A. Biedermann [a], A. Thiéry [a], S. Bozza [c]

[a] University of Lausanne, School of Criminal Justice, Lausanne, Switzerland
[b] University of Strasbourg, Rsearch Team UMR 977 INSERM, Faculty of Dentistry, Strasbourg, France
[c] University Ca' Foscari, Department of Economics, Venice, Italy

ARTICLE INFO

Article history:
Received 21 December 2010
Received in revised form 27 April 2011
Accepted 5 August 2011
Available online 9 September 2011

Keywords:
Evidence evaluation
Bayes' factor
Likelihood ratio
Handwriting evidence
Sex determination
Investigation

ABSTRACT

This paper extends previous research and discussion on the use of multivariate continuous data, which are about to become more prevalent in forensic science. As an illustrative example, attention is drawn here on the area of comparative handwriting examinations. Multivariate continuous data can be obtained in this field by analysing the contour shape of loop characters through Fourier analysis. This methodology, based on existing research in this area, allows one describe in detail the morphology of character contours throughout a set of variables. This paper uses data collected from female and male writers to conduct a comparative analysis of likelihood ratio based evidence assessment procedures in both, evaluative and investigative proceedings. While the use of likelihood ratios in the former situation is now rather well established (typically, in order to discriminate between propositions of authorship of a given individual versus another, unknown individual), focus on the investigative setting still remains rather beyond considerations in practice. This paper seeks to highlight that investigative settings, too, can represent an area of application for which the likelihood ratio can offer a logical support. As an example, the inference of gender of the writer of an incriminated handwritten text is forwarded, analysed and discussed in this paper. The more general viewpoint according to which likelihood ratio analyses can be helpful for investigative proceedings is supported here through various simulations. These offer a characterisation of the robustness of the proposed likelihood ratio methodology.

1. Introduction

The forensic examination of handwriting examination involves the description of handwriting features along with the study of their range of variation. The characterization of writing habits relies on an essentially 'subjective'[1] approach, largely dependent on examiners, while Courts seem less and less at ease with such evaluations [15,16,18,8].

Various studies have been undertaken with the aim of reducing the subjective part of the handwriting analysis process [17,21,3,20]. Marquis et al. [12,13], for example, have proposed a procedure based on Fourier analysis that allows one to describe the contour shape of loops of characters and to study their intra- and inter-variability through a set of variables (so called *Fourier descriptors*). Recent developments (see, for example, [11]) focused on improving the understanding of the degree of intra- and inter-individual variability of selected handwriting features that are used in comparative processes. This objective is met through a research project that relies on the development of a formal procedure to assist forensic examiners in the analysis of handwriting, notably in selecting relevant handwriting features that describe the shape of characters as well as their standardized and automated extraction on samples collected from documents under investigation.

The main objective of this paper is to conceptualize and develop frameworks for assessing, articulating and communicating the evidential value of handwriting features. This objective is approached through an inferentially rigorous methodology – Bayesian, in essence – for evidence evaluation in forensic science.

The paper is organised as follows. Section 2 starts by presenting the *Bayes factor*, the principal statistic discussed throughout this paper and known in forensic science contexts more often as the *likelihood ratio*. Material and data are presented in Section 3. Section 4 introduces the proposed inferential model for the assessment of the value of handwriting evidence in an evaluative framework. Section 5 develops the inferential approach for the investigative usage of handwriting evidence, along with a

* Corresponding author.
E-mail address: franco.taroni@unil.ch (F. Taroni).

[1] The term 'subjective' is not understood here in the sense of 'arbitrary', but in the sense of 'personalistic', that is intimately related to a given individual (e.g., a scientist) who is concerned with a particular inference task.

0379-0738/$ – see front matter © 2011 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.forsciint.2011.08.007

measures of its robustness for this purpose. A discussion and conclusions are presented in Section 6.

2. The use of the *Bayes factor* for the evaluation of scientific evidence

In forensic science, statistical concepts are now widely used for assessing the probative value of various kinds of scientific evidence. In the context, evidence usually consists of particular observations and/or measurements made on, for example, a series of handwritten characters on a questioned document and those made on a undisputed text written by a suspect. During the evaluative process, the assessment of such evidence is typically conducted by deriving a statistic, such as the *likelihood ratio* (hereafter written *LR*, for short). This metric represents a concept that provides a balanced measure of the degree to which evidence at hand is capable of discriminating between competing propositions that are forwarded, for example, by opposing parties at trial [10]. Examples for such propositions could be 'the suspect is the author of the questioned text (H_1)' or 'an unknown person is the author (H_2)'. The use of the likelihood ratio in forensic science is now well established, presented and discussed in both theory and practice [2].

Before proceeding further, some addition notes on terminology are introduced here. When comparing alternative propositions, it is often useful to consider the odds form of the Bayes' theorem, which provides a concise description of how evidence, say $\mathbf{y}$, alters the odds in favor of a given proposition:

$$\frac{Pr(H_1|\mathbf{y})}{Pr(H_2|\mathbf{y})} = \frac{Pr(\mathbf{y}|H_1)}{Pr(\mathbf{y}|H_2)} \times \frac{Pr(H_1)}{Pr(H_2)}.$$

This is an expression of the way in which a prior opinion is updated to a posterior opinion through consideration of evidence, that is a transformation obtained by multiplying the prior odds by a quantity known as the *Bayes factor* (BF for short):

$$BF = \frac{Pr(\mathbf{y}|H_1)}{Pr(\mathbf{y}|H_2)}.$$

The Bayes factor measures the change produced by the evidence in the odds when going from the prior to the posterior distribution in favor of one scientific theory – as referred to, for example, in the more general context of inference in science – as opposed to another [7]. In many cases, the competing hypotheses have single distributions (i.e., simple versus simple hypotheses), and it can readily be shown that the Bayes factor is just the likelihood ratio of H_1 to H_2, and depends only upon the sample data [22]. There may however be other cases where composite hypotheses are compared and unknown parameters are involved. Here, the Bayes factor still has the form of a likelihood ratio, but it is the ratio of two marginal likelihoods obtained by integrating over the parameter space:

$$Pr(\mathbf{y}|H_k) = \int Pr(\mathbf{y}|\theta_k, H_k)\pi(\theta_k|H_k)d\theta_k \quad k = 1, 2, \tag{1}$$

where θ_k is the unknown parameter under hypothesis H_k, $\pi(\theta_k|H_k)$ is its prior density, and $Pr(\mathbf{y}|\theta_k, H_k)$ is the likelihood function. The Bayes factor is the ratio of weighted likelihoods under the competing hypotheses, and it appears that it no longer depends only upon the sample data.

Generally, it can be said that statistics such as the Bayes factor can be used for two main purposes in forensic science:

1. A first purpose consists of assigning a value for a given item of evidence. This refers to the *evaluative* level at which forensic scientists operate. Evaluating a piece of evidence means that the scientist provides an expression of the value of the evidence in support – which may be positive, negative or neutral – of a hypothesis of interest. This represents the task briefly introduced in the introductory lines of this Section. A main aspect of this level of operation is that the scientist does *not* express an opinion about a proposition itself. This is a main difference with respect to the second purpose outlined hereafter.
2. A second purpose is that of providing information to the police. Here, the scientist acts at an *investigative* level. At this stage, the scientist tries to answer questions such as 'what happened?'. The forensic scientist is said to be 'crime focused' and observes evidence which forms the basis to generate hypotheses and suggestions for explanations, in order to give guidance to police investigators.

In accordance with current forensic literature, the term *likelihood ratio* will be used in what follows instead of *Bayes factor*. It will be understood as a ratio of marginal likelihoods, as noted in Eq. (1), in presence of composite hypotheses. The forthcoming sections illustrate examples of applications for the two above mentioned *operational levels* and offer a theoretical presentation of the implemented statistical models. Section 3 will start by introducing the data collected.

3. Materials and data

This paper focuses on the evaluative and investigative value of a single letter. Loops of the letter d were extracted on documents written by 80 writers (21 males and 59 females) who produced a total of 7290 loops of letter d. Fourier coefficients [9,19] have been extracted from 5826 of these characters. The remaining letters could not be analysed with this method because they did not consist of a closed loop, or, in some instances, could not be considered as simple curves. Here, a simple closed curve is considered as one that is characterised by the fact that each radius starting from the barycentre of the contour crosses this contour only once. Note that *Fourier coefficients* rather than Fourier descriptors as used in [12,13] have been used in the research presented here. Technical challenges, such as the bimodality described in [13], favour the use of Fourier coefficients. Both methodologies are mathematically related and offer a comparable level of discrimination between writers. Validation tests ensure the reliability and efficiency of Fourier coefficients in discriminative examination [23]. Preliminary analyses showed that pairs of Fourier coefficients of order higher than 4 can be neglected, so that the shape of each analysed character can be described through $p = 8$ variables representing the first four pairs of coefficients, say (a_i, b_i), for $i = 1, \ldots, 4$.

4. Assessing handwriting evidence at evaluative level

In the kind of settings considered here, experts must handle items of evidence, that is handwritten characters on a questioned document, that can be quantitatively described by several variables. In other words, they must handle multivariate data. One of the skepticisms forwarded against the use of multivariate statistical techniques in forensic science to deal with multivariate continuous data is the lack of background information from which one could estimate the parameters of the adopted statistical models. In fact, there are cases in which it was incorrectly assumed that variables are independent so as to lead to a reduction in the number of parameters to be estimated.

A statistical model for the evaluation of evidence in terms of a likelihood ratio for multivariate data, for use within an evaluative perspective, has been proposed in [1]. This development was made in the context of elemental composition of glass fragments. That model assumes two sources of variation (that between measurements within the same group, and that between groups) and a constant variation within sources. More recently, further Bayesian methodology has been developed which is able to handle the non-constant variation within sources that characterize, in particular, handwriting scenarios. This is worth considering because each writer presents a peculiar individual variability which cannot