Development of a Mathematical Formula for the Calculation of Fingerprint Probabilities Based on Individual Characteristics

JAMES W. OSTERBURG, T. PARTHASARATHY, T. E. S. RAGHAVAN, and STANLEY L. SCLOVE*

A method is developed for assigning a probability to a fingerprint, including a partial print, based on the number of individual (Galton) characteristics present. A multinomial model is used, the categories of which are the individual characteristics and combinations of them. The negative of the logarithm of the probability of any particular configuration is related to the entropy function of information theory. The parameters of the model are estimated from data (fingerprints). Confidence bounds are obtained for the negative log probability of any configuration.

KEY WORDS: Fingerprints; Identification; Criminalistics; Entropy; Information, Multinomial model.

1. INTRODUCTION AND SUMMARY

A fingerprint is left at the scene of a crime. A suspect is involved. His fingerprint matches the one at the scene of the crime. The suspect's lawyer argues that his client has an alibi. The suspect's print does match that at the scene of the crime, but it was a print of only one finger, and not a full print at that.

Assume for the sake of discussion that there is little evidence other than the fingerprint. Is the partial print so improbable that the suspect should be convicted on this evidence alone? In order to answer this question, a method for assigning a probability to a partial fingerprint is developed.

1.1 Background Information and Definitions

The bulb of each digit of the human hand contains friction ridges that form themselves into patterns thus providing a basis for classification. A fingerprint is an impression left by perspiration, grease, and oil which are present on the ridges. The potential impression is generally invisible; hence it is called a *latent* print, until it is developed by a special powder or by chemicals such as iodine vapor or silver nitrate solution. Latent prints are found at crime scenes. Often they are less than a full impression of all of the ridge lines; in this circumstance, they are called *partial* prints. For permanent recording, the bulbs of the ten fingers are inked, using printer's carbon ink, to preserve clearly all the ridge lines in detail. The inked fingers are carefully pressed onto a fingerprint card, a blank form specifically designed for classification and file purposes. Any latent impressions processed at a crime scene or at a forensic laboratory are compared against the recorded prints of suspects either present in the files or taken expressly for this purpose if the suspect has no criminal record. Ridge lines remain unchanged from birth until death.

Ridge lines fall into three major pattern types: loops (ca. 65 percent), whorls (ca. 30 percent), and arches (ca. 5 percent). Further subdivision within each pattern allows a classification scheme to be organized so that for the ten fingers over a thousand categories of fingerprint-pattern combinations result. Within each category there are many fingerprints from different individuals which, to the untrained eye, appear to be the same. This process of separation through classification results in relatively small sets of fingerprints which are of manageable proportions for the purpose of search and comparison.

The individuality of a fingerprint is not based on this classification scheme, but rather upon the ridge-line details, termed *Galton characteristics*, since Sir Francis Galton was among the first to study these individual characteristics systematically (Galton 1892). Whether in a loop, whorl, or arch, a ridge line may end abruptly, in which case the Galton detail is termed a ridge ending; approximately one-half of all individual characteristics are ridge endings. A ridge line may suddenly divide into two branches, much like a fork in a road; such a charac-

* James W. Osterburg is Professor and Head, Department of Criminal Justice, University of Illinois at Chicago Circle, Box 4348, Chicago IL 60680. At the time of this study T. Parthasarathy was Visiting Associate Professor, Department of Mathematics, University of Illinois at Chicago Circle; he is now associated with DePaul University, Chicago, IL 60614. T.E.S. Raghavan and Stanley L. Sclove are Associate Professors in the Department of Mathematics of the University of Illinois at Chicago Circle. Professor Osterburg provided the data, formulated the problems, and cooperated in writing the article. Professors Parthasarathy, Raghavan, and Sclove did the probability modeling and statistical analysis.

Portions of this article were presented in invited addresses by Sclove at the National Meeting of the Operations Research Society of America, Boston, April 22–24, 1974, and by Osterburg at the Annual Meeting of the International Association for Identification, Washington, D.C., July 28–August 1, 1974. This research was supported under a contract with the Center for Research in Criminal Justice, University of Illinois at Chicago Circle, Hans W. Mattick, Director. Computations were performed at the Computer Center of the University of Illinois at Chicago Circle.

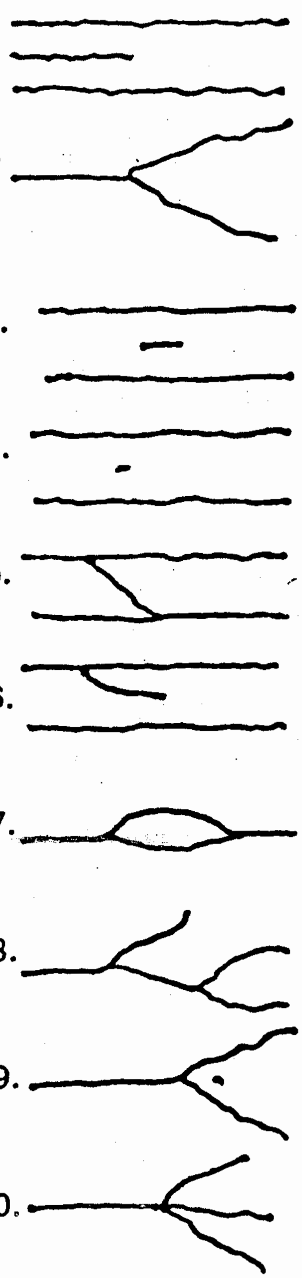
Thanks are extended to the referees, Associate Editor, and Editor for their helpful comments.

© Journal of the American Statistical Association
December 1977, Volume 72, Number 360
Applications Section

ɪtic is termed a bifurcation (or fork). There is general ement upon the ten types of ridge-line details that useful in the characterization of a fingerprint as ɪue. The nomenclature and shape of these individual lton) characteristics are shown in Figure A.

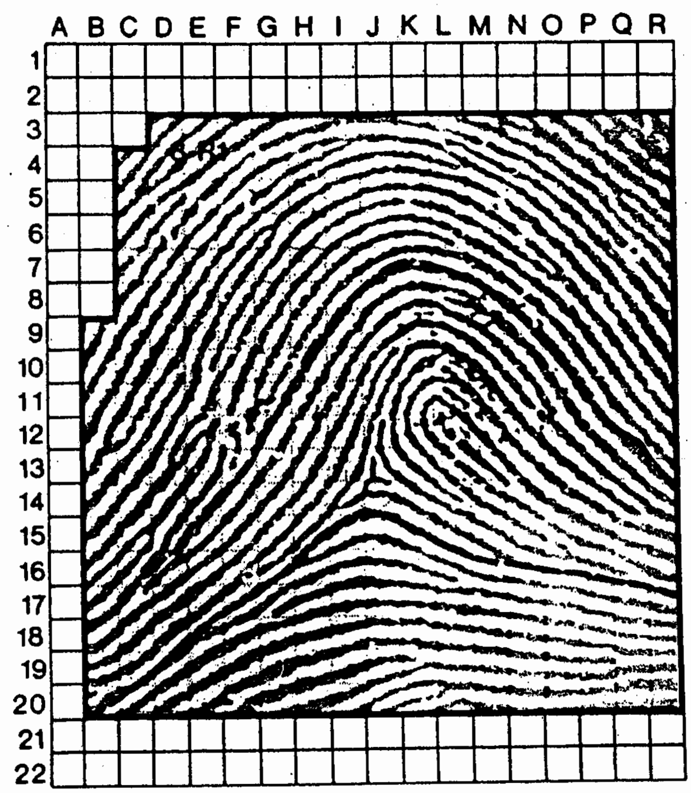
. Nomenclature and Shape of the Ten Individual ge Line (Galton) Characteristics.

Name	*Visual Appearance*
Ending ridge	1.
Fork (or bifurcation)	2.
Island ridge (or short ridge)	3.
Dot (or very short ridge)	4.
Bridge	5.
Spur (or hook)	6.
Eye (enclosure or island)	7.
Double bifurcation	8.
Delta	9.
Trifurcation	10.



The work here was done within the following framerk. A grid of one millimeter squares was placed over a ɜerprint. (See Figure B.) Each fingerprint is considered a *configuration* of the cells of the grid. For each cell the grid there are several possibilities: One or more of ten characteristics is there, or no characteristic is sent. Thus a configuration is a grid of cells. Each cell may be thought of as either being empty or else being occupied by one or more words. The words represent the characteristics present; e.g., if a cell contains the words "dot, dot, ending ridge," it means that the one square millimeter area corresponding to that cell contained two dots and one ending ridge.

B. Cell Grid Superimposed on Fingerprint. Middle Section Between Dark Lines is the Pattern Area. The Core is in Cell 12L; the Delta is in Cell 13J.



A *match* between a suspect's full print and a partial print exists when there is a section of the full print that is the same as the partial print. Since we are working in terms of a grid of cells, for our purposes this match exists when a grid can be laid on the full print in such a way that the resulting configuration contains a section which is the same as the configuration corresponding to the partial print.

The fingerprints which were studied were enlarged to ten times actual size, making a full rolled print about 8″ by 10″. The cells of the grid were one centimeter square after enlargement. Members of the project staff coded the ten Galton characteristics, cell by cell. (See Appendix A for precise working definitions of the characteristics.) Thirty-nine prints were coded. (Osterburg had earlier examined 40 prints, from 40 different individuals, but one was misplaced or borrowed and not returned leaving 39 for reexamination.) There is no problem with representativeness of the sample. The Galton characteristics are "accidental." They are not genetic. With regard to these characteristics, two siblings are no more alike than two random persons. Therefore, with respect to these characteristics, each and every person is representative. The sampling model, then, is one of fixed effects, different for each individual. This is inherent in the nature of an identification problem.

1.2 Partial Fingerprints

Accepted practice among fingerprint experts in America indicates that the limiting or weakest, yet unequivocal, identification is that which is based on the presence of 12 ridge endings alone. We show that the negative log probability of 12 ridge endings in a partial print of about 50 to 100 square millimeters is about 20, i.e., the probability is about 10^{-20}. The occurrence of only three trifurcations (the rarest characteristic) in a partial print of this size has an estimated probability that is even smaller. Hence such a configuration, being rarer, would also be considered as yielding an identification.

By taking the negative log probability of the multinomial distribution of the characteristics (the result being the conventional entropy measure employed in information theory), a weight was obtained for each of the ten characteristics, such that the negative log probability of any mix of characteristics is equal to the sum of the weights of the characteristics in the mix. Any configuration resulting in a sum greater than about 20 would, according to accepted practice, then yield an identification.

This article represents an attempt to quantify what is meant by "reasonable doubt" in one field of physical evidence, the analysis of fingerprints. The probabilities in question are small. We shall be concerned not with whether a probability is on the order of one in a thousand but whether it is one in a million or one in a million millions. Such probabilities take on meaning when they are considered in terms of the corresponding expected numbers in large populations such as the adult population of the United States.

Some consideration is given toward the end of the article to a conditional probability of relevance to inference in criminalistics. This is the *probability of identity*, the probability that a person whose print matches a latent print is the person who made that print.

2. DATA DESCRIPTION

By an *occurrence* we mean the occurrence of any one of the ten Galton characteristics. The 39 fingerprints used yielded a total of 8,591 cells which could be coded. In all there were 2,536 occurrences or 0.295 per cell. Table 1 gives the distribution of the number of occurrences per cell, without regard to type.

1. Distribution of Number of Occurrences

Number of occurrences	0	1	2	3	4	5	Total
Number of cells	6,584	1,594	320	72	19	2	8,591
Proportion of cells	.766	.185	.0372	.00838	.0022	.00023	1.00

The abbreviations used for the characteristics are as follows.

- B: bridge
- D: dot
- E: ending ridge
- F: fork (bifurcation)
- I: Island
- L: lake (eye)
- O: delta
- S: spur
- T: trifurcation
- Z: double bifurcation

The symbol DE denotes the occurrence of one dot and one ending ridge in a cell; BEE would denote the occurrence of a bridge and two ending ridges in a cell; etc. Altogether 54 combinations occurred including one DEEEE and one DDDDE. Table 2 gives the distribution of these cell configurations. Note in particular that 77 percent of the cells were empty; i.e., the probability that a cell is occupied is .23.

2. Distribution of Cell Configurations

Cell configuration	Frequency: Number of cells	Frequency: Percent of cells
Empty	6,584	76.6%
E	715	8.32
F	328	3.82
I	152	1.77
D	130	1.51
EE	119	1.39
B	105	1.22
S	64	0.745
L	55	0.640
EL	32	0.372
DE	32	0.372
EEE	21	0.244
EI	21	0.244
O	17	0.198
DD	15	0.175
BE	13	0.151
Z	12	0.140
DI	11	0.128
EEEE	10	0.116
ES	10	0.116
DDI	10	0.116
II	9	0.105
FI	9	0.105
BF	7	0.0815
DEE	7	0.0815
FF	5	0.0582
T	5	0.0582
EEF	4	0.0466
BEE	4	0.0466
EII	4	0.0466
FL	3	0.0349
BB	3	0.0349
FS	2	0.0233
BD	2	0.0233
DDE	2	0.0233
LL	2	0.0233
Other (19 other multiple occurrences)	67	0.780
Total	8,591 cells	100.0%

3. PROBABILITIES OF VARIOUS CONFIGURATIONS

The model we employ in the present article can be summarized as follows.

1. A fingerprint is characterized as a configuration of the cells of a grid.
2. For any cell there are 13 possibilities; either the cell is empty, or one of the following twelve possibilities has occurred: B, D, E, F, Z, L, O, S, T; Z, EE (broken ridge), or other multple occurrence. (By multiple occurrence we mean more than one occurrence in a cell.)
3. The cells are statistically independent.

he probability of a given configuration is, under this iel, given by the point multinomial probability disution, $p_0^{k_0}p_1^{k_1}\ldots p_{12}^{k_{12}}$, where the k_i's, $i = 0, 1, \ldots,$ are nonnegative integers summing to t, the total iber of cells in the print, and the p_i's are the probities of the 13 possibilities and hence sum to one.

n regard to assumption (2), in the selection of the tinomial categories we gave priority to the ten idard Galton characteristics used in practice, occur; as singletons. The number of possible combinations iltiple occurrences) of these individual characteristics normous. Among the combinations, we selected the ible ridge ending because it was the most frequent;), it includes a broken ridge, which is different from a ge coming to an end as shown in Figure A. A conuence of lumping rare multiple occurrences together the category, other multiple occurrence, is to give benefit of the doubt to the suspect, in the sense of ing a conservative, i.e., large, probability estimate for given configuration.

The consequences of assumption (3) have been studied clove 1976); in the following section, we mention how s study qualifies the results of the present article.

Estimates of Probabilities

The parameters p_i of the model were estimated from e data. See Table 3. The standard deviation of the timate of any one of the p_i's is $[p_i(1-p_i)/n]^{\frac{1}{2}}$, where $= 8{,}591$ cells. Estimates of the standard deviations are so given in Table 3.

3. Estimates of Probability Parameters

arameter	Cell configuration	Frequency	Estimate of probability parameter	Estimated standard deviation of estimate
p_0	Empty	6,584	.766	.0045
p_1	Island (I)	152	.0177	.0014
p_2	Bridge (B)	105	.0122	.0012
p_3	Spur (S)	64	.00745	.00093
p_4	Dot (D)	130	.0151	.0013
p_5	Ending ridge (E)	715	.0832	.0030
p_6	Fork (F)	328	.0382	.0021
p_7	Lake (L)	55	.00640	.00086
p_8	Trifurcation (T)	5	.000582	.00024
p_9	Double bifurcation (Z)	12	.00140	.00040
p_{10}	Delta (O)	17	.00198	.00048
p_{11}	Broken ridge (or EE)	119	.0139	.0013
p_{12}	Other multiple occurrence	305	.0355	.0020
otal		8,591	1.0	

ource: Table 2.

The probability P of a configuration of k_0 empty cells, k_1 cells containing islands, k_2 cells containing bridges, ..., k_{10} cells containing deltas, k_{11} cells containing two ending ridges, and k_{12} cells containing other multiple occurrences, is estimated by $\hat{P} = \hat{p}_0^{k_0}\hat{p}_1^{k_1}\ldots\hat{p}_{12}^{k_{12}}$. Let E, for entropy (information), be defined as $-\log_{10} P$ and $\hat{E} = -\log_{10} \hat{P}$. We have

$$E = -\sum_{i=0}^{12} k_i \log_{10} p_i \ .$$

Appendix B gives confidence bounds for E.[1]

3.2 Configurations Yielding Identification

Lower limits on the number of characteristics sufficient to constitute *legal identification* have been set from time to time. Though these limits have varied from country to country, bureau to bureau, and expert to expert, a number frequently mentioned is 12. Since ridge endings are the most common of the ten characteristics, then the limiting, or weakest, identification is that based on 12 ridge endings alone. Consequently, as an example consider a configuration of 12 ridge endings in a print of area of, say, 72 sq. mm. The estimated information (negative $\log_{10}$ probability) is 19.9; a 95 percent confidence interval is (19.6, 20.2). Thus the probability P of such a configuration is approximately 10^{-20}. Table 4 gives the estimated information for a few configurations.

4. Information for Several Configurations[a]

Area (sq. mm.)	Ending ridges	Trifurcations	One delta, 12 ending ridges	One delta, three trifurcations
50	17.3 (.12)	15.1 (.51)	19.7 (.13)	17.5 (.51)
100	23.1 (.26)	20.9 (.55)	25.5 (.26)	23.3 (.55)

[a] Numbers in parentheses are estimated standard deviations.
NOTE: "Information" is negative $\log_{10}$ probability.

The occurrence of only three trifurcations (the rarest characteristic) in a print has an even smaller probability than that of 12 ridge endings. Hence such a configuration, being rarer than 12 ridge endings, would also be considered as yielding an identification.

[1] *Variation with respect to location.* A *delta* is defined as the point where, as one moves out from the core of the fingerprint pattern, the ridges change from concave to convex. The *pattern* area is defined as the area between core and delta(s); the nonpattern area, as the complementary, outside area. The characteristics were coded by location (pattern vs. nonpattern area). The ratio of the probability of occurrence of a characteristic in the nonpattern area is .598 times that in the pattern area; i.e., characteristics are only 60 percent as prevalent in the nonpattern area as in the pattern area. The model can easily be modified to take into account variation with respect to location by multiplying separate probability density functions for pattern and nonpattern areas. When a latent print permits discrimination between pattern and nonpattern areas, such information could be used in the probability calculation. There may well be other sources of heterogeneity; e.g., the core and delta areas may be denser in characteristics than other areas, and there may be an increase in density with decreasing distance to the core. Methods for dealing with these sources of variation are discussed in Sclove (1976).

By taking the negative log of the multinomial probability (thus obtaining the conventional entropy measure employed in information theory), we obtain weights for the possibilities, such that the negative log probability of any mix of characteristics is equal to the sum of the weights of the characteristics in the mix. Any configuration resulting in a sum greater than 20 would, according to accepted practice, yield an identification. The weights are given in Table 5.

5. Weights for the Characteristics

Cell configuration	*Estimate of probability, p_i*	*Weight, $-\log_{10} p_i$*
Delta	.00198	2.70
Island	.0177	1.75
Bridge	.0122	1.91
Spur	.00745	2.13
Dot	.0151	1.82
Ending ridge	.0832	1.08
Fork	.0382	1.42
Lake	.00640	2.19
Trifurcation	.000582	3.24
Double bifurcation	.00140	2.85
Broken ridge	.0139	1.86
Other multiple occurrence	.0355	1.45
No characteristic (empty cell)	.766	0.116

In regard to assumption (3), a study of inter-cell dependence (Sclove 1976) indicates that the probability a cell is occupied increases if adjacent cells are occupied. In fact, the probability a cell is occupied is about .40 if six of the eight adjacent cells are occupied, compared to an overall probability of occupancy of .24. A reasonably conservative upper bound can be obtained by doubling the probabilities for nonisolated occupied cells and replacing p_0 by one. Generally, we should expect a change of one or two units in the log probability. Such a change is unimportant for our purposes since we are concerned with deciding whether a probability is on the order of one in a million or one in a million millions, not with deciding whether a probability is .1 or .01. The approximations of the present article should give results even more accurate than that except for configurations where most of the occurrences are clustered together in neighboring cells.

Another demonstration of robustness with respect to assumption (3) was provided in the course of the present study when, using independent Poisson within-cell distributions for the characteristics, we repeated the computations with various assumed cell sizes. The estimates of negative log probability did not vary much with respect to cell size. One can then argue that if there is some cell size for which independence is a good approximation, then the estimates based on the particular cell size of one sq. mm. are accurate since the estimates do not vary much with respect to cell size.

The preceding discussion has dealt with the assignment of a probability P to the occurrence of a given configuration in a given set of cells. For inferential purposes it is necessary to estimate the probability that a person has this configuration *somewhere on his fingers*. To obtain a conservative (i.e., large) estimate of this probability, one assumes that the latent print in question could have been made by any part of any finger of a person. Suppose the latent print is approximately rectangular, w millimeters in width by ℓ millimeters in length. Take an average full-finger print to be approximately a rectangle of about 15 mm. by 20 mm. Then the number of possible positions for the latent print, inside of the full print, is $(15 - w + 1)(20 - \ell + 1)$; e.g., if the width is 6 mm. and the length is 11 mm. (corresponding to a latent print with an area of about 66 sq. mm.), this is about 10×10 possible positions. Multiplying by 10 fingers gives $10 \times 100 = 1{,}000$ or 10^3 total possible positions.

The occurrence of the configuration in one of these positions does not exclude its occurrence in one or more other positions, but since $P(A \text{ or } B \text{ or } \ldots)$ is at most $P(A) + P(B) + \ldots$, an upper bound is obtained by multiplying by the total number of possible positions. Consequently, for this size of latent print, the negative $\log_{10}$ probability of a given configuration in a given set of 50 to 100 cells ought to be decreased by about three in instances where it is assumed that the configuration could occur anywhere on any one of a person's ten fingers. More generally, for a $w \times \ell$ latent print, this decrease should be approximately $\log_{10}[10(15 - w + 1)(20 - \ell + 1)]$. This analysis could be refined to take into account nonrectangular shapes, but the adjustment suggested here would appear to be of an appropriate order of magnitude.

For a full print, the number one is subtracted from the negative $\log_{10}$ probability to allow for the possibility that the full configuration could occur on any one of ten fingers.

For a given (rectangular) area $= w \times \ell = A$ sq. mm., the function $(15 - w + 1)(20 - \ell + 1)$ is maximized when $w = 4(A/21)^{\frac{1}{2}}$. The maximum value is $(4\sqrt{21} - \sqrt{A})^2$. The corresponding maximum decrease in negative log probability is approximately

$$\log [10(4\sqrt{21} - \sqrt{A})^2 = 1 + 2 \log |4\sqrt{21} - \sqrt{A}|$$
$$= 1 + 2 \log |18.3 - \sqrt{A}| \ .$$

For an area $A = 100$ sq. mm., this is $1 + 2 \log 8.3 = 2.84$. For an area $A = 50$ sq. mm., it is $1 + 2 \log 11.2 = 3.09$. So for a print between 50 and 100 sq. mm. in area, this maximum adjustment is, then, about three. This changes the negative log probability of 20 for 12 ridge endings to about 17.

3.3 Kingston's Method

Kingston (1964) made an interesting study of the problem of modeling the occurrence of individual characteristics and computing fingerprint probabilities. His approach to computing fingerprint probabilities is based on the relative frequencies of the various types of characteristics. The relative frequency of a given type of characteristic is the proportion of that type, among all acteristics. A Poisson distribution is used by Kingston he distribution of the total number of characteristics. gston thus uses a conditional multinomial distribution n the total number of occurrences; it follows that the lting (marginal) distributions of the characteristics Poisson. We did not find Poisson distributions to be quate in describing the distribution per cell of the l number of characteristics, though it might be adete for small areas. Our method does not impose sson distributions for the different characteristics and ns simpler to apply.

Kingston's method is rather difficult to apply; he aclly estimates the probability of only one configuration 92), a configuration of four ending ridges, one fork, one recurve, in an area of about 43 sq. mm. (pp. 91). He obtains a value of 9.1×10^{-14}, or about 10^{-13}, the probability of this assortment. The following s are used to arrive at this.

1. Through a fairly complicated procedure developed in Chapter V (pp. 25–81), a probability of 1.914×10^{-10} for obtaining occurrences in a particular geometrical configuration is derived.
2. By reference to his calibration of expected number of occurrences vs. area, one obtains an expected count of 4.5 for this size area. One then finds from a table of the Poisson distribution that the probability of six counts occurring, using the mean value of 4.5, is about .1252. Combining this with the result of Step (1.) gives $1.914 \times 1.252 \times 10^{-11} = 2.396 \times 10^{-11}$.
3. One then uses the relative frequencies of .459 for an ending ridge, .341 for a fork, and .250 for a recurve to obtain a conditional probability of $.459^4 \times .341 \times .250 = 3.784 \times 10^{-3}$ for this assortment. [A "recurve" or "recurving ridge" is not one of the ten standard characteristics; the relative frequency of .25 was "a somewhat arbitrary value" (p. 93).]
4. The unconditional probability is then $2.396 \times 3.784 \times 10^{-14} = 9.1 \times 10^{-14}$.

Needless to say, it would appear to be difficult to obtain fidence intervals to accompany point estimates obed by Kingston's procedure.

We compute the probability of this same configuration, ollows, using the weights given in Table 5. For this estimate a weight for the "recurve," by interpolation n Table 6.

6. Estimation of Weight for "Recurve"

Occurrence	*Kingston's relative frequency*	*Our weight*
Fork	.341	1.42
Recurve	.250	x[a]
Dot	.083	1.82

[a] $1.42 + [(.250 - .083)/(.341 - .083)](1.82 - 1.42) = 1.68$.

We then compute as follows:

4 ridge endings:	4×1.08	=	4.3
1 fork:	1×1.42	=	1.4
1 recurve:	1×1.68	=	1.7
37 empty cells:	37×0.116	=	4.3
Total			11.7

Thus we estimate the probability to be $10^{-11.7}$, compared to Kingston's 9.1×10^{-14}, which is about 10^{-13}—very close agreement in terms of order of magnitude.

4. PROBABILITY OF IDENTITY

Once a probability P has been obtained for a latent fingerprint, what should we then do with this number? What inferences are to be drawn?

Let a be the person who committed the crime and b be a suspect with the same characteristics. The logical identity $b \equiv a$ denotes the event that these persons are one and the same. Let C be the random variable giving the number of persons with the same characteristics as a. Then the conditional probability $P(b \equiv a | C \geq 1)$ has been termed by Cullison (1969) the "probability of identity." We denote it by $P(\mathrm{Id})$. This probability has been considered by others as well as Cullison including Kingston (1964, 1965a); see also Kingston (1965b). Note that

$$P(\mathrm{Id}) = P(b \equiv a | C \geq 1) = \sum_{k=1}^{\infty} P(C = k \text{ and } b \equiv a)/P(C \geq 1) .$$

Now, $P(C = k \text{ and } b \equiv a) = P(b \equiv a | C = k) \cdot P(C = k)$. If we take $P(b \equiv a | C = k) = 1/k$, this is $(1/k)P(C = k)$, and

$$P(\mathrm{Id}) = \sum_{k=1}^{\infty} k^{-1} P(C = k)/P(C \geq 1) = \sum_{k=1}^{\infty} k^{-1} P(C = k | C \geq 1) = E(C^{-1} | C \geq 1) .$$

It is clear that the limit of $P(\mathrm{Id})$ as C becomes stochastically small should be one. It is reasonable to take C to be distributed according to a hypergeometric distribution or a binomial distribution, with small probability parameter. Hence calculations can be performed using a Poisson distribution with parameter λ. It is small values of λ that are of interest in criminalistics. Some values of $P(\mathrm{Id})$ and $1 - P(\mathrm{Id})$, the probability of "identifying" a wrong person, are given in Table 7. Some of these are given by Kingston (1964, page 22).

7. Values of the Probability of Identity

λ	*P(Id)*	*1 − P(Id)*
2	.5607	.4393
1	.7670	.2330
0.2	.9506	.0494
0.1	.9749	.0251
0.01	.9975	.0025
0.001	.99975	.00025

APPENDIX A: THE GALTON CHARACTERISTICS

Definitions of some of the Galton characteristics were refined by us by means of precise working definitions, which were necessary to accomplish the coding.

A *bridge* was defined as less than two centimeters in length in the enlarged photograph (i.e., two millimeters in actuality); otherwise, it would be coded as a fork.

A *dot* was defined as being large enough to encompass one pore. Smaller "dots" were not counted; larger "dots" were coded as short ridges.

Distinct breaks in ridges were coded as two separate *ending ridges* to distinguish such breaks from ridges simply coming to an end.

A *spur* was defined as being less than two centimeters in length in the enlargement (i.e., two millimeters in actuality); otherwise, it was ~~coded as a fork~~. A spur was counted only once: The end of a spur was not counted as a ridge ending.

The sizes in these rules of thumb are of an order of magnitude suggested by T. Dickerson Cooke[2] of the Institute of Applied Science, Chicago, Illinois and are consistent with recommendations of the Committee on Standardization of the International Association for Identification.

APPENDIX B: DERIVATION OF CONFIDENCE BOUNDS FOR THE ENTROPY

The negative log probabilities considered are given by the expression

$$E = \log_{10} P = -\sum_{i=0}^{12} k_i \log_{10} p_i \ ,$$

where $p_0 = 1 - p_1 - p_2 - \ldots - p_{12}$ and $k_0 = t - k_1 - k_2 - \ldots - k_{12}$, with t being the total number of cells in the print. We have $\hat{E} = \log_{10} \hat{P} = (\log_{10} e)(\log_e \hat{P}) = (\log_{10} e)H$, where $H = \log_e \hat{P}$; thus $\text{var}(\hat{E}) = (\log_{10} e)^2 \text{var}(H)$. The asymptotic variance of H is given by Bowman et al. (1971); this gives

$$\text{var}(\hat{E}) \cong n^{-1}(\log_{10} e)^2 [\sum_{i=0}^{12} k_i^2/p_i - t^2] \ .$$

This variance is estimated by substituting the estimates $\hat{p}_i$ based on $n = 8{,}591$ cells. As an example, for 12 ending ridges and no other characteristics in a print of area $t = 72$ cells, we have 60 empty cells. Hence $k_0 = 60$, $k_5 = 12$, and the other k_i's are 0. From Table 3 we have $\hat{p}_0 = .766$ and $\hat{p}_5 = .0832$. Thus

$$\hat{E} = -12(\log_{10} .0832) - 60(\log_{10} .766) = 19.9 \ ,$$

and

$$\text{var}(\hat{E}) = 8{,}591^{-1}(0.434)^2[(60^2/.766) + (12^2/.0832) - 72^2] = 0.0273 \ .$$

The corresponding standard deviation is $0.0273^{\frac{1}{2}} = 0.165$. Thus a 95 percent confidence interval is obtained from the point estimate by adding and subtracting $1.96(0.165) = 0.3$.

[*Received June 1975. Revised July 1977.*]

REFERENCES

Bowman, K.O., Hutcheson, K., Odum, E.P., and Shenton, L.R. (1971), "Comments on the Distribution of Indices of Diversity," in *Statistical Ecology, Volume 3: Many Species Populations, Ecosystems, and Systems Analysis*, eds. G.P. Patil, E.C. Pielou, and W.E. Waters, University Park, Pa: Pennsylvania State University Press, 315–366.

Cullison, Alan D. (1969), "Identification by Probabilities and Trial by Arithmetic (A Lesson for Beginners in How to Be Wrong with Greater Precision)," *Houston Law Review*, 6, 471–518.

Galton, Francis (1892), *Finger Prints*, London: Macmillan; republication (1965), New York: DaCapo Press.

Kingston, C.R. (1964), "Probabilistic Analysis of Partial Fingerprint Patterns," Ph.D. dissertation, Dept. of Criminology, University of California, Berkeley.

——— (1965a), "Applications of Probability Theory in Criminalistics—I," *Journal of the American Statistical Association*, 60, 70–80.

——— (1965b), "Applications of Probability Theory in Criminalistics—II," *Journal of the American Statistical Association*, 60, 1028–1034.

Osterburg, James W. (1964), "An Inquiry into the Nature of Proof: the Identity of Fingerprints," *Journal of Forensic Sciences*, 9, 413–427.

——— (1969), "The Evaluation of Physical Evidence in Criminalistics: Subjective or Objective Process?" *Journal of Criminal Law, Criminology and Police Science*, 60, 97–101.

Sclove, Stanley L. (1976), "The Occurrence of Fingerprint Characteristics as a Two-Dimensional Process," Center for Research in Criminal Justice and Department of Mathematics, University of Illinois at Chicago Circle.

[2] Cooke, T. Dickerson, personal communication (1974).