A method is developed for assigning a probability to a fingerprint, including a partial print, based on the number of individual (Gallon) characteristics present. A multimonial 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 (fingerprint). Confidence bounds are obtained for the negative log probability of any configuration.

KEY WORDS: Fingerprint; Identification; Criminalization; 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

*James W. Osterburg is Professor and Head, Department of Criminal Justice, University of Illinois at Chicago Circle, Box 4344, Chicago, IL 60604. At the time of this study, T. Parthasarathy was Visiting Associate Professor, Department of Mathematics, University of Illinois at Chicago Circle. J. L. Solove was employed by the Department of Mathematics of the University of Illinois at Chicago Circle. Professor Osterburg and Stanley L. Solove 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 Solove did the probability modeling and statistical analysis.

Portions of this article were presented in invited addresses by Solove 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 26-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. Maxick, 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.

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 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 an impression of all of the ridge lines; in this circumstance they are called partial prints. For permanent records the bulbs of the ten fingers are inked, using printless 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 file 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-
erm print Probabilities.

tic is termed a bifurcation (or fork). There is general
resent upon the ten types of ridge-line details that are
eful in the characterization of a fingerprint as
The nomenclature and shape of these individual
iexion are shown in Figure A.

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.

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 representa-
tiveness of the sample. The Galton characteristics are
“residential.” They are not genetic. With regard to these
characteristics, two siblings are no more alike than two
random persons. Therefore, with respect to these charac-
teristics, 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 our 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 rare, would also be considered as yielding an identification.

By taking the negative log probability of the multisomotial 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 2,356 occurrences or 0.265 per cell. Table 1 gives the distribution of the number of occurrences per cell, without regard to type.

1. Distribution of Number of Occurrences

<table>
<thead>
<tr>
<th>Number of occurrences</th>
<th>O</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of cells</td>
<td>6,584</td>
<td>1,594</td>
<td>320</td>
<td>72</td>
<td>19</td>
<td>2</td>
<td>8,591</td>
</tr>
</tbody>
</table>

3. PROBABILITIES OF VARIOUS CONFIGURATIONS

The model we employ in the present article can be summarized as follows.
The probability \( P \) of a configuration of \( k \) empty cells, \( k \) cells containing islands, \( k \) cells containing bridges, \( k \) cells containing deltas, and \( k \) cells containing two ending ridges, is estimated by \( P = P_0 P_1 \ldots P_n \). Let \( \mathcal{E} \) for entry \( \mathcal{E} \) (information) be defined as \( -\log_P P \) and \( \mathcal{E} = -\log_P S \).

We have

\[
\mathcal{E} = \sum_{i=1}^{11} i \bar{A}_i \log_P \bar{A}_i
\]

**Appendix B** gives confidence bounds for \( \mathcal{E} \).

### 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 say, 72 sq. mm. The estimated information (entropy, \( \log_P \) probability) is 10.9; a 90 percent confidence interval is (16.6, 20.3). Thus the probability \( P \) of such a configuration is approximately \( 10^{-9} \). Table 4 gives the estimated information for a few configurations.

### 4. Information for Several Configurations

<table>
<thead>
<tr>
<th>Area (sq. mm.)</th>
<th>Ending ridges</th>
<th>Trifurcations</th>
<th>One delta, 12 ending ridges</th>
<th>One delta, 3 trifurcations</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>17.3 (2)</td>
<td>15.1 (51)</td>
<td>157 (13)</td>
<td>17.0 (51)</td>
</tr>
<tr>
<td>100</td>
<td>31.1 (26)</td>
<td>20.4 (55)</td>
<td>25.5 (28)</td>
<td>23.9 (56)</td>
</tr>
</tbody>
</table>

**Note:** *Information* = \( -\log_P \) 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 in location. A delta is defined at the point where the axis moves away from the core of the fingerprint pattern, the ridge change from convex to concave. The pattern area is defined as the area between core and delta(s); the companion area, as the supplementary, outside area. The characteristic were lined up by the various pattern analysts. The ratio of the probability of occurrence of a characteristic in the companion area is 3.8 times that in the pattern area, i.e., characteristic are only 90 percent as prevalent in the companion area as in the pattern area. The model can easily be modified to take into account variation with respect to location by multiplying respective probability density function for pattern and companion area. When current print performance discrimination for wave, pattern and companion areas, such information could be used in the probability calculation. There may well be other sources of brevity, e.g., the core and delta areas may be determined in characterizing these other areas, and there may be an increase in density with increasing distance to the core. Methods for dealing with these sources of variation are discussed in [1975].
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 weight less than 20 would, according to accepted practice, yield an identification. The weights are given in Table 5.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Estimate of probability, p_i</th>
<th>Weight, -log_p_i</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delta</td>
<td>0.0198</td>
<td>3.70</td>
</tr>
<tr>
<td>Island</td>
<td>0.177</td>
<td>1.175</td>
</tr>
<tr>
<td>Bridge</td>
<td>0.122</td>
<td>1.175</td>
</tr>
<tr>
<td>Sper</td>
<td>0.0745</td>
<td>2.13</td>
</tr>
<tr>
<td>Dot</td>
<td>0.051</td>
<td>1.82</td>
</tr>
<tr>
<td>Ending ridge</td>
<td>0.0832</td>
<td>1.05</td>
</tr>
<tr>
<td>Fork</td>
<td>0.0382</td>
<td>1.42</td>
</tr>
<tr>
<td>Lake</td>
<td>0.00840</td>
<td>2.19</td>
</tr>
<tr>
<td>Trifurcation</td>
<td>0.0000822</td>
<td>3.24</td>
</tr>
<tr>
<td>Double bifurcation</td>
<td>0.00140</td>
<td>2.85</td>
</tr>
<tr>
<td>Broken ridge</td>
<td>0.0739</td>
<td>1.86</td>
</tr>
<tr>
<td>Other multiple</td>
<td>0.055</td>
<td>1.45</td>
</tr>
<tr>
<td>No characteristic</td>
<td>0.765</td>
<td>0.116</td>
</tr>
</tbody>
</table>

In regard to assumption (3), a study of inter-cell dependence (Selove 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 noniliated occupied cells and replacing p_i 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, n milliseconds in width by t milliseconds in length. Take an average full-finger print to be approximately a rectangle of about 10 by 20 mm. Then the number of possible positions for the latent print, inside of the full print, is (15 - e + 1)(20 - t + 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 X 10 possible positions. Multiplying by 10 fingers gives 10 X 10 X 10 X 10 = 10^4 total possible positions.

The occurrence of the configuration in one of these positions does not exclude the occurrence in one or more other positions, but since P(A) or B or ...) is at most P(A) + P(B) + ... an upper bound is obtained by multiplying by the total number of possible positions. Consequently, for this size of latent print, the negative log 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 \( x \) latent print, this decrease should be approximately \( \log_{10}(10^{10} - e + 1)(120 - t + 1) \). This analysis could be refined to take into account non-rectangular 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 probability to allow for the possibility that the full configuration could occur on any one of ten fingers. For a given (rectangular) area \( x = w \times l \approx A sq. mm. \), the function \( 10^{10}(10 - e + 1)(120 - t + 1) \) is maximized when \( w = 4(4l/21) \). The maximum value is \( 4(4/21 - \sqrt{A}) \). The corresponding maximum decrease in negative log probability is approximately \( \log_{10}(10^{4l/21} - \sqrt{A}) = 1 + 2 \log_{10}(4l/21 - \sqrt{A}) = 1 + 2 \log_{10}(18.8 - \sqrt{A}) \).

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
A Poisson distribution is used by Kingston to explain the total number of characteristics. As such, he uses a normal distribution to calculate the total number of occurrences, which follows the Citing (marginal) distributions of the characteristics. We did not find Poisson distributions to be in the description per cell of the number of characteristics, though it might be adequate for small areas. Our method does not impose distributions for the different characteristics and is simpler to apply. Kingston’s method is rather difficult to apply; he suggests estimating the probability of only one configuration (92), a configuration of four empty ridges, one fork, one recurve, in an area of about 43 sq. mm. (pp. 91). He obtains a value of $9.1 \times 10^{-4}$, or about $10^{-4}$, the probability of this assortment. The following are used to arrive at this value.

1. Through a fairly complicated procedure developed in Chapter V (pp. 20-91), a probability of $1.944 \times 10^{-10}$ for obtaining occurrences in a particular geometrical configuration is derived.
2. By reference to his calibration of expected counts of occurrence vs. area, one obtains an expected counts of 4.5 for the size area. One then finds from a table of the Poisson distribution that the probability of six exceeds occurring, using the mean value of 4.5, is about .1230. Combining this with the result of Step (1) gives $1.944 \times 1.322 \times 10^{-10} = 2.596 \times 10^{-10}$.
3. One then uses the relative frequencies of .659 for an ending ridge, .341 for a fork, and .280 for a recurve to obtain a conditional probability of $0.659 \times 0.341 \times 0.280 = 3.744 \times 10^{-10}$ for the assortment. (A “recurve” or “recurring ridge” is not one of the ten standard characteristics; the relative frequency of .28 was “a somewhat arbitrary value” (p. 92.).
4. The unconditional probability is then $3.356 \times 3.744 \times 10^{-10} = 9.1 \times 10^{-4}$.

In other words, it would appear to be difficult to obtain 1.322 occurrences to accompany point estimates obtained by Kingston’s procedure.

Compute the probability of this same configuration, allowing, using the weights given in Table 5. For this estimate a weight for the “recurve,” by interpolation from Table 6.

### Table 5

<table>
<thead>
<tr>
<th>Occurrence</th>
<th>Kingston’s relative frequency</th>
<th>Our weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fork</td>
<td>.341</td>
<td>1.42</td>
</tr>
<tr>
<td>Recurve</td>
<td>.250</td>
<td>1.32</td>
</tr>
<tr>
<td>Dot</td>
<td>.083</td>
<td>1.32</td>
</tr>
</tbody>
</table>

\[1.42 = \left( \frac{.341}{.250} \right) \times \left( \frac{.250}{.083} \right) = 1.32 \times 1.32 = 1.32^2 \]

\[
\begin{align*}
\text{Total} & = 11.7 \\
\end{align*}
\]

Thus we estimate the probability to be $10^{-4}$, compared to Kingston’s $9.1 \times 10^{-4}$, which is about $10^{-4}$—very close agreement in terms of order of magnitude.

### 4. PROBABILITY OF IDENTITY

A measure $P$ has been obtained for a latent fingerprint, what should we then do with this number? What inference are we to draw?

Let $a$ be the person who committed the crime and $b$ be a suspect with the same characteristics. The logical identity $a = b$ means that $a$ and $b$ are the same. Let $C$ be the random variable giving the number of persons with the same characteristics as $a$. Then the conditional probability $P(b = a|C \geq 1)$ has been termed by Cullison (1966) the “probability of identity.” We denote it by $P(Id)$. This probability has been considered by others as well as Cullison including Kingston (1964, 1965a); see also Kingston (1965b). Note that

\[
P(Id) = P(b = a|C \geq 1)
\]

\[
= \sum_{i=1}^{\infty} P(C = k \text{ and } b = a)/P(C \geq 1).
\]

Now, $P(C = k \text{ and } b = a) = P(b = a|C = k) \times P(C = k)$, and if we take $P(b = a|C = k) = 1/k$, this is $(1/k)P(C = k)$, and

\[
P(Id) = \sum_{k=1}^{\infty} 1/k \times P(C = k)/P(C \geq 1)
\]

\[
= \sum_{k=1}^{\infty} 1/k \times P(C = k)/P(C \geq 1) 
\]

It is clear that the limit of $P(Id)$ as $C$ becomes statistically 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 parameters. Hence calculations can be performed using a Poisson distribution with parameter $\lambda$. It is small values of $\lambda$ that are of interest in criminalistics. Some values of $P(Id)$ and $1 - P(Id)$, the probability of “identifying” a wrong person, are given in Table 7. Some of these are given by Kingston (1964, page 22).

### Table 6

<table>
<thead>
<tr>
<th>Occurrence</th>
<th>Kingston’s relative frequency</th>
<th>Our weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fork</td>
<td>.341</td>
<td>1.42</td>
</tr>
<tr>
<td>Recurve</td>
<td>.250</td>
<td>1.32</td>
</tr>
<tr>
<td>Dot</td>
<td>.083</td>
<td>1.32</td>
</tr>
</tbody>
</table>

\[1.42 = \left( \frac{.341}{.250} \right) \times \left( \frac{.250}{.083} \right) = 1.32 \times 1.32 = 1.32^2 \]

\[
\begin{align*}
\text{Total} & = 11.7 \\
\end{align*}
\]

### Table 7

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$P(Id)$</th>
<th>$1 - P(Id)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>.5607</td>
<td>.4393</td>
</tr>
<tr>
<td>1</td>
<td>.7670</td>
<td>.2330</td>
</tr>
<tr>
<td>0.5</td>
<td>.9506</td>
<td>.0494</td>
</tr>
<tr>
<td>0.1</td>
<td>.9749</td>
<td>.0251</td>
</tr>
<tr>
<td>0.01</td>
<td>.9975</td>
<td>.0025</td>
</tr>
<tr>
<td>0.001</td>
<td>.99975</td>
<td>.00025</td>
</tr>
</tbody>
</table>

APPENDIX A: THE GALTON CHARACTERISTICS

Definitions of some of the Galton characteristics were related 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. Cooks 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 INTERVALS FOR THE ENTROPY

The negative log probabilities considered are given by the expression

\[ E = \log_2 P = -\sum_{i=1}^{N} k_i \log_2 p_i, \]

where \( p_i = 1 - p_1 - p_2 - \ldots - p_{i-1} \) and \( k_i = e - c_i - \ldots - k_{i-1} \), with \( i \) being the total number of cells in the print. We have \( \log_2 E = (\log_2 e) (\log_2 P) = (\log_2 e) H \), where \( H = \log_2 P \); thus \( \text{var}(E) = (\log_2 e)^2 \text{var}(H) \). The asymptotic variance of \( H \) is given by Bowman et al. (1971); this gives

\[ \text{var}(E) \approx \frac{1}{\sum_{i=1}^{N} k_i / p_i} \]

This variance is estimated by substituting the estimates \( \hat{p}_i \) based on \( n = 8,595 \) 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 \( \hat{p}_i = 60/k_i = 12 \), and the other \( k_i \) are 0. From Table 3 we have \( p_5 = 0.766 \) and \( p_6 = 0.052 \). Thus

\[ E = -12 \log_2 (0.0523) - 60 \log_2 (0.766) = 19.9 \]

and

\[ \text{var}(E) = 0.010 (0.0423)(0.766) + (0.0523)^2 = 0.0734 \]

The corresponding standard deviation is 0.2731 = 0.16, thus a 95 percent confidence interval is obtained from the point estimate by adding and subtracting 1.96(0.2731) = 0.3.

[Received June 1975. Revised July 1977.]

REFERENCES


