Winter 1964

The Use of Statistics in Criminalistics

C. R. Kingston

P. L. Kirk

THE USE OF STATISTICS IN CRIMINALISTICS

C. R. KINGSTON AND P. L. KIRK

C. R. Kingston is a graduate student in the School of Criminology in the University of California, Berkeley, where he is working toward a Doctorate of Criminology. He received a Master of Criminology degree from the same University in 1961 and has previously published a paper deaing with separation of the components of marijuana by gas chromatography. He presented this paper before the California Association of Criminalists.

P. L. Kirk, Professor of Criminology, University of California, Berkeley, is well known to our readers and workers in the field of police science. Professor Kirk has published extensively on research problems in this field and is the author of *Crime Investigation* (1953).—EDITOR.

This paper will present a relatively non-technical discussion of some aspects of the utilization of statistics in the field of criminalistics. A general concept of probability will be examined to see what relationship it might bear to the interpretative areas of criminalistics. Some specific conditions imposed by the nature of the statistical process will be outlined and then illustrated with some previously published data. It is hoped that this presentation will help to orient the reader toward a better understanding of the proper role of statistical methods in criminalistics.

JUSTIFICATION

There are many philosophically oriented fundamental ideas of probability. Perhaps the most practical basic approach to the subject lies in the concept of frequency, in which statements of probability express the relative frequencies of repeated events. Such generalizations are usually derived from prior observation of frequencies exhibited in a series of trials, plus the assumption that these same frequencies will recur in the future.

For practical use in criminalistics, it is of little interest what might happen in a long series of trials; the crime is committed only once. Of what use, then, is the above frequency concept? The answer to this lies in another way of looking at probability, which is to consider it as a degree of belief. For instance, in throwing a normal die, we would wish to decide just how much we believed that a given number, say six, will occur in the next toss. It is accepted today that the rational man will base the degree of belief upon the frequency concept. This means that the degree of belief that the number six will come up in the next toss would be about .17, which is about the proba-

bility that six will occur in an unbiased die. The exact interpretation of this can cause a bit of puzzlement. What the figure actually means is that out of many tosses of the die, six would be expected to occur in roughly 17% of the total number of tosses. But only the *next* toss is of interest, not a series of tosses, and six cannot occur in 17% of the next toss.

Under these circumstances the figure of .17 can only represent a guide to a decision. We wish to decide either that six will occur or will not, and the degree of belief in its occurrence, .17, will be a guide to this decision. In this case, if a decision were made solely on the basis of this figure, it would be that six will not occur.

Now consider a problem of evaluating the significance of the coincidence of several properties in two pieces of glass. Suppose that the probability of two fragments from different sources having this coincidence of properties is .005, and that the probability of such coincidence when they are from the same source is .999. What do these figures mean? They are simply guides for making a decision about the origin of the fragments, either in the form of an opinion by the expert or as a question of fact by the jury. Since the consequences of a decision will enter into the evaluation, no set figures can be given to establish the dividing point between the two possible decisions. The contrast between the requirement of "proof beyond reasonable doubt" that is demanded in criminal cases and the "preponderance of evidence" acceptable in civil cases illustrates the bearing of the consequences upon the making of a decision.

Now that some insight into the meaning of probability as applied to a single event has been gained, it is necessary to see whether or not such a proba-

bility is applicable in criminalistics, and whether it offers any advantages for making decisions. Only the more general aspects of this question will be considered here.

As shown above, a probability figure is related to past experience, and its accuracy will depend on the amount and applicability of this experience as well as upon the analytical processes applied to evaluate it. A decision is made by utilizing this probability, in conjunction with considerations as to the consequences of the decision, as a guide. If the term "estimate of the situation" is substituted for "probability" or "probability figure" in the above, the statements apply to decision-making processes used at present in criminalistics. Thus the difference between the generally-used approach and the statistical approach lies in the terms "estimate of the situation" and "probability figure".

In the majority of cases, the "estimate of the situation" is obtained through an intuitive evaluation based upon past experience—not only one's own experience, but often also that obtained indirectly from the study of others' combined experiences. This is unavoidable, since no one person could hope to have personal experience with the multitude of situations encountered in the practice of criminalistics. In the intuitive approach, therefore, there are two places where error can occur: one, in the communication of experience from different sources; and two, in the intuitive evaluation of these experiences, whether others' or one's own. The communication can suffer from semantic difficulties, while the evaluation is likely to be biased by the experimenter's own prejudices.

A probability figure can suffer from the same sources of error. That is, statistical information can be presented in such a way that the true meaning is obscured, and an experimenter can inject his own biases into the data being processed. The difference between the two approaches, then, lies in the fact that the means of minimizing the errors introduced are incorporated into a proper statistical approach and can be spelled out for the experimenter, whereas such definite controls are not easily applied to the intuitive approach. We believe that this is sufficient justification for advocating the use of statistics in criminalistics, for then the emphasis is upon a formal system rather than on a person.

CONDITIONS IMPOSED

In order to derive the maximum benefit from a statistical approach, the conditions and limitations imposed must be known and understood. Unfortunately, many opponents of the use of statistical methods in criminalistics have pronounced dogmatically that these conditions are so limiting as to invalidate completely any results. However, we know of no rational argument ever presented purporting to uphold this view. The difficulties inherent in the conditions should be judged only after they have been intelligently analysed, and not from general impressions which may stem from misunderstandings about the subject. Some of the limiting conditions will be discussed in this section.

Planning. A statistical analysis is not a remedy for poorly or inappropriately gathered data. If the data are to be applied to any particular problem of type of problem, it is essential to define exactly what is going to be demanded of the analysis. A method of analysis must be chosen with respect to the type of data which it is possible to gather. Then, and only then, are the data collected and the projected analysis applied. It is rare that data collected by one not versed in statistical techniques will allow the extraction of even fifty percent of the information that could be obtained from proper data with the same, or even less, effort and expense. When presented with inappropriate data, the statistician feels much the same as does the criminalist when the inexperienced investigator brings to the laboratory a paper bag into which he has stuffed the evidence from the scene of the crime along with the clothing of the suspect.

This does not imply that even inappropriate data cannot be analysed for mean values, variance, independence of variables, and so on, but these values will have meaning in the original problem only if the data have been properly collected with respect to the problem.

Randomness. The majority of cases in which the criminalist might employ statistical methods present two related aspects:

1. A problem arises which would require data on each member of a very large population for a 100%-accurate analysis.

2. The only practical way to approach the problem is to gather data on a limited selection of members from the population and then apply the results of this limited study to the problem.

In such a study, randomness enters into two different selection processes. One, unless a specific, partially non-randomized sampling scheme is planned, the sample taken from the population for study must reflect a random distribution of the

objects with respect to the properties of interest; and two, the particular item of evidence that is being examined must show a similar randomness with respect to these properties. The obvious way of achieving this in selecting a sample for study is to make a random selection from the population; that is, to select the items in such a way that any one item has the same probability of being chosen as any other item. In many cases it will not be possible to do this. For instance, in selecting a sample from headlight glass, it is not possible to go about smashing up people's headlights to obtain fragments thereof. Then the sampler must be content to take what is available if it can be reasonably inferred that this will reflect a random distribution with respect to the properties of interest. This was done in reference 2 which is discussed later.

The choice of the sample is more or less under the control of the experimenter, whereas the choice of evidence received is not. Therefore, it will inevitably be necessary to make a careful study of the type of evidence that is likely to appear in any one type of activity so that the population selected will be one from which the evidence will be equivalent to a random sample. Only under these conditions will the information obtained from a statistical analysis have meaning.

Independence. The idea of independence is often bandied about in discussions of statistical applications. It should be common knowleege to criminalists that properties must be statistically independent before the probability of a conjunction of these properties can be derived from the multiplication rule, that is by multiplying together their separate probabilities. Note that the term "statistically" is used here. Independence is defined by specific mathematical relationships and should always be tested by reference to these relationships.[1] To judge independence from our general concepts or information about the physical world can lead to error.[2]

It should be emphasized that independence is not necessary before a statistical analysis can be utilized. It is certainly a desirable property which greatly simplifies the work involved in most instances, but the problem is far from hopeless when

there is a lack of statistical independence among some of the properties.

Uncertainty. A statistical analysis is used when uncertainty must exist. If there were a way of arriving at a *certain* answer to a problem, statistical methods would not be used. But when uncertainty does exist, and a statistical approach is possible, then this approach is the best one available since it offers an index of the uncertainty based upon a precise and logical line of reasoning. It is almost incredible that this very reason for the development of statistical methods is seized upon by its opponents and used as an argument against the use of statistics. They will show that an occurrence which was deemed extremely improbable did actually happen and contend that therefore the methods by which its improbability was calculated cannot be trusted.

Although the uncertainty involved in a situation to which a statistical analysis is applicable constitutes a limitation on the appropriateness of *any* decisions made, the fact of uncertainty validates the contention that the statistical approach offers the *best possible* solution to the problem. This is not meant to condone the misuse of statistics just because such an approach is indicated. It is undoubtedly true that serious errors have been made in applying incorrect statistical methods to the evaluation of physical evidence, but such misuse does not support the generalization that statistics cannot be properly used in criminalistics at all.

AN ILLUSTRATION

Some of the questions raised in the above discussion will be illustrated by the problem of comparing glass fragments. The problem will be limited to the determination of whether or not the fragments were at one time part of a particular larger glass object. Only two properties, density or specific gravity, and refractive index, will be considered. Two publications have been chosen for reference since they present some of the actual pertinent data. The publications will be called reference 1 (ref. 1)[3] and reference 2 (ref. 2),[4] respectively. Ref. 1 presents the results of determinations of the specific gravity and refractive index of 100 glass samples collected from miscellaneous glass objects. Ref. 2 presents the results of determinations of the density (20°C, g/ml), refractive index (20°C), hard-

---

[1] The simplest relationship is that the probabilities do combine by the multiplication rule; e.g. that $P(AB) = P(A) \times P(B)$.

[2] For an example of this see: FELLER, WILLIAM, AN INTRODUCTION TO PROBABILITY THEORY AND ITS APPLICATIONS, Vol. 1; John Wiley and Sons, Inc., N.Y.; 1st ed., p. 86, example (d).

[3] GAMBLE, LUCY, BURD, DAVID Q., KIRK, PAUL L., J. CRIM. LAW & CRIMINOL., *33:* 416 (1943).

[4] NELSON, D. F.; THE ANALYST, *84:* 388 (1959).

ness and color of 50 samples from different Lucas 700 headlamps obtained from motor-repair depots in one city.

Planning. In ref. 1 there was no initial planning, relative to a subsequent statistical analysis, before collecting the samples, although planning is shown relative to the overall aim of the study, which was to examine the significance of the measured properties in glass origin determinations. Since no statistical handling was attempted, the lack of prior considerations in this respect is understandable. Several considerations would have to be taken into account if useful and reasonably accurate information were to be obtained. These include the type of crime and its location, and the bearing of these two factors upon the choice of population or populations to be studied. The sampling should then be planned so as to reflect any significant effects of these variables.

In ref. 2 it is stated that the data therein are sufficient to describe a particular population. Here a definite statistical problem is stated: describing a particular population. Two factors led the experimenter to choose a limited population consisting only of Lucas 700 headlamp glass, the factors being, (1) the evidence received was from a Lucas 700 headlamp, and (2) the majority of British cars in New Zealand were equipped with this same type of headlamp. The variables involved here are quite limited; the main concern would be the effects of different batches of glass on the distribution of the properties. The bimodal distribution of the densities might be attributable to some chance difference in different batches of glass or, as suggested in the article, to some change in manufacturing processes. This would be a difficult thing to foresee, and would be difficult to prepare for in a planning stage. However, such occurrences should be considered and incorporated into the overall planning wherever possible. Complete records, including detailed information about the samples, might make it possible to explain any unexpected behavior of the data.

It is suggested that a general discussion of the variables considered and taken into account, as well as the bearing of the choice of population upon the type of evidence being considered, be made a part of every report involving statistical evaluations. Such inclusions would serve the double purpose of helping the criminalist to assess the applicability of the data to his particular problem and of awakening a general awareness of the problems to be considered in using statistical methods in criminalistics.

Randomness. In ref. 1 it is stated that the glass samples were "collected at random". The persons undertaking the study collected any stray pieces of glass that might be observed in the course of their daily routines. (This explains the "unknown" category listed in the publication.) This manner of collection may have indeed resulted in a random sample, but experience has shown that true randomness is difficult to achieve on an intuitive basis. The human mind seems to have channels that work against randomness, and that tend to introduce an unconscious bias into any selection process. This is one reason for the considerable expenditure of time in the construction of tables of random numbers, which can be used truly to randomize a selection process. Another mental process that works against randomness is the tendency to wish to cover a given area evenly. For instance, if different colored beads are scattered about a checker board, there is a tendency in sampling the beads to take one bead from each square and call this a random sample. Obviously it is not, since the beads remaining on a square after a selection there have a zero probability (no chance) of future selection. For random sampling, each bead must always have the same probability of selection as any other bead left on the board. However, the above sampling method on the checker board might be appropriate for a stratified sampling scheme; but then the data would be analysed differently.

One way of checking the randomness of the sampling would be to investigate the possible presence of bias for or against any portions of the population. One such bias might favor objects that are found in the street, though these are not necessarily typical of the population in general. An obvious example of such a class of objects is bottle fragments. Indeed, the original data sheets for ref. 1 show that 40% of the samples consisted of bottle glass. This percentage may have resulted from bias, or it may actually reflect the distribution of the population chosen.

The appropriateness of the population to be sampled would of course be determined from a study of the evidence likely to be received. Since the concern in ref. 1 was with criminal acts in general, an extensive population is clearly indicated. Still, if certain types of crime were known to produce the bulk of evidence where glass was

involved, and therefore limited the majority of evidential glass to a few types of objects, to be valid a general sampling would have to reflect this limitation. For instance, if burglary and automobile accidents comprised a large percentage of activities where glass appeared as evidence, there might be a significantly larger percentage of building window glass, automobile window glass, and automobile headlight glass as evidence than glass from all other sources combined. The chosen population should then reflect this distribution of sources.

In ref. 2, the evidence is indicated to be glass from Lucas 700 headlamps that is in evidence from automobile accidents. The applicability of the entire population of Lucas 700 headlamps would intuitively seem reasonable. If, however, one particular make of automobile happened to be designed so that it was more disposed to headlamp breakage than other makes, and if the manufacturer had purchased a large order of Lucas 700 headlamps, which just happened to have been molded from one deviant batch of glass, the result might be a significant bias in the population resulting from accidents.[5]

Some other considerations in connection with the sampling from various motor-repair depots might be made. The sources of the headlamps sampled in the depots should be questioned; whether they are from lamps broken in accidents or from burnout replacements might be important. The makes of automobiles that the depots handle could be considered; the samples from depots that handle only one make of automobile could be significantly different from samples obtained in depots that handle a large variety of makes.[6]

[5] There is no implication here that the more remote possibilities mentioned might actually obtain; the important factor is that they are recognized as possibilities and then rejected, rather than rejected by virtue of not being considered. Such possibilities might rapidly be ruled out by a criminalist who has had practical contact with the physical situation, but could be considered as serious possibilities by those who have not had such contact.

[6] Dr. Nelson was kind enough to examine this paper and offer some comments. It seems appropriate to mention some of these which demonstrate the criminalist's point of view.

". . . it (manufacturer purchasing a large batch of headlamps) is *unlikely* because the purchase of headlamps by car manufacturers is *probably* a continuing process rather than a series of "batch processes" repeated once or twice per year."

"If cars are in accidents severe enough to break headlamps, one would not *expect* great differences in

It should be obvious that most of the points mentioned in the above discussion of randomness must be considered in the planning stage, and settled before the data are obtained.

Independence. Suppose that previous experiments had been conducted with glass fragments, and that any one experiment considered only density or refractive index, but not both, and that these experiments showed that the probability of the occurrence of density X was P(X) and of refractive index Y was P(Y). Would the probability of both X and Y occurring simultaneously in one fragment of glass, say P(X,Y), be equal to the product of the separate probabilities; i.e., would $P(X,Y) = P(X) \times P(Y)$? This would be true only if the two properties were independent. Now consider Figure 1. The scatter diagram is reproduced from the original data for ref. 1. There is a very definite visual indication of a correlation between the two properties under consideration. If this is in fact the case, these two properties are not independent, and the multiplication rule would not be applicable. If it were concluded that the data gave a reasonable fit to a bivariate normal distribution, it would be possible to utilize existing tables to find the probability P(X,Y).

There is a tendency in references to the evidential value of various properties to pick some probability figure as representative of each property and, with proper warning about the independence requirement, to show that the properties considered have a probability of occurrence equal to the product of the separate probabilities. This is an extremely dangerous and misleading practice. First of all, it fosters a general concept that this is the proper way to analyse statistical data, and thus leads to the incorrect application of the multiplication rule when the independence of the properties is not easily checked. Next, the use of a single probability is not correct unless the properties all show a uniform distribution, which is extremely rare in practice. Mollification is sometimes offered with such statements to the effect that the estimate is conservative or that it sets some sort of maximum bound to the probability figure. If the calculation does indeed show the maximum bound on the probability for any property combination within the distribution, there can be no quarrel as to its technical correctness, but it still tends to

disposition to headlamp breakage. The headlamps studied were broken, not burnt-out."
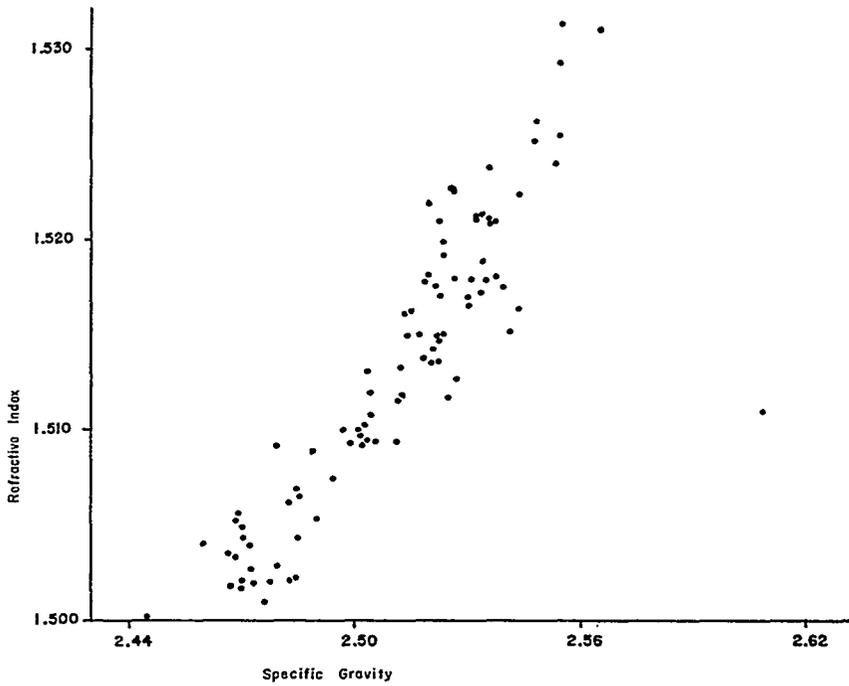
The italics are the authors'.

FIGURE 1

mislead persons not sufficiently versed in statistical methods. Further, if this conservative estimate shows too large a probability, the properties involved may be deemed inappropriate for evidential use. But combinations that define an area remote from the central portion of the overall distribution may show a sufficiently low probability to be of value. The general and uncritical use of the multiplication rule has caused more criticism about the application of statistics in criminal investigation than any other factor.

Figure 2 is a scatter diagram derived from the tabulated data in ref. 2, and shows an interesting distribution of points. It is stated that there does
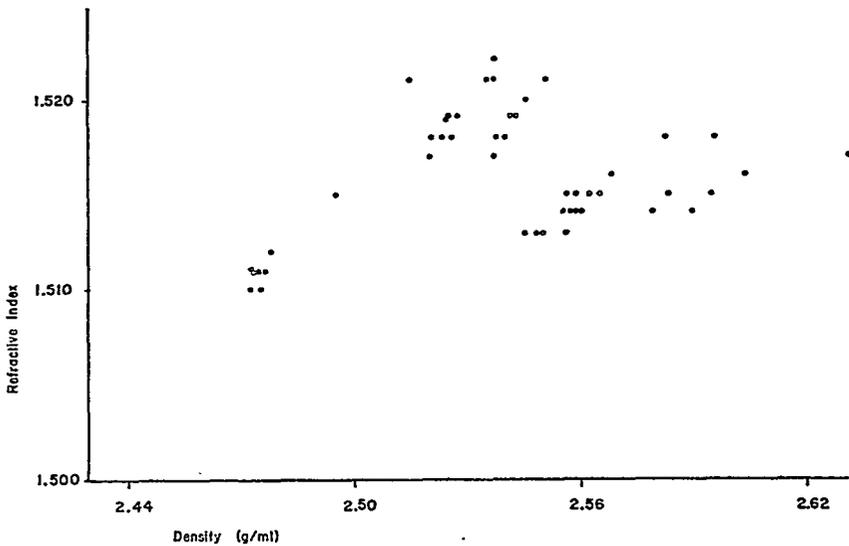

FIGURE 2

not seem to be any correlation between the two properties except for some samples in the lower density group. In this case, with the exception of values in the lower density group, we would have $P(X,Y) = P(X) \times P(Y)$. Note that variables are used in the expression to indicate that the probabilities are not fixed, but assume a value dependent on the value of the property. The multiplication rule is used only as a simple way of arriving at the combined probability, and should not exclude the consideration of the distribution relative to density and refractive index as a bivariate distribution over the plane.

But now look at the diagram from a slightly different viewpoint. There is one group of points that lies roughly along a 45° line from the lower left corner, and then another group that lies on a line at about 30° and which is to the right of the first group. The first group is close to what might be expected from an inspection of the grouping in Figure 1, whereas the second group would be a bit surprising if the data in ref. 1 were considered as typical of glass in general. If instead of looking at the distribution as a single bivariate one, it is viewed as two bivariate distributions, as represented by the two indicated groupings, then the visual indication is that a correlation within each group exists which approximates that seen in Figure 1.

If now, as an exercise, the second group in Figure 2 is shifted to the left, it will fill up a "hole" that appears in the first group, thus bringing the data more into line with that in ref. 1. At times, such grouped deviation of values may be indicative of a consistent bias occurring in the measuring process. This would be especially suspect if the group of objects showing this shift were examined at a separate time or even consecutively. Another cause of such a shift could lie in a bias that occurred in the sampling. Suspicion would be enhanced here if the shifted group of glass came from one particular repair depot. The difference in distribution between the two sets of data in these two studies is interesting, and poses a problem to be considered in future research.[7]

CONCLUSION

In conclusion we will discuss some of the more immediate implications of the foregoing discussion

[7] We consider the possibility of an actual bias in this particular example to be remote, and are only suggesting that such questions be considered when planning, executing or reading statistically oriented studies.

in the field of criminalistics. First of all, there is no need for all criminalists to become statisticians in order to utilize statistics in their profession, but there is a need for them to become aware of some of the more general aspects of statistics.

Next, the determination of a suitable population for study should preferably be based upon the collected experiences of many criminalists. Perhaps the organization of a national evidence report might be desirable. Summary information could be submitted by interested laboratories at specified intervals. This data could then be used to choose appropriate populations for study, as well as to compile other useful statistics about the occurrence and examination of physical evidence. An analysis of this information could then be disseminated for general use, followed later by the results of a study on a sample from the chosen population.

The use of statistical evaluations of physical evidence cannot be rushed. It will take time to determine appropriate areas of applicability, to decide upon the proper methods of analysis, to gather the necessary data, and to establish a confidence in the results both among the criminalists and in the courts. At first, the responsibility will fall upon a few with the necessary time and interest to study the problems involved and to work out the fundamental concepts and methods. But even here, the aggregate considerations, both favorable and unfavorable, of criminalists in general will be of significant value.

One way to develop some further insight into the problems involved and some methods of approaching these problems from a general statistical point of view is for the criminalist to reflect for a while on the logic used in forming interpretative opinions. As an example, consider the opinion that a particular suspect left the partial fingerprint in evidence, the basis being a coincidence of 12 elements with no observable differences in the patterns of the evidential print and an area of the print taken from the suspect. What logic leads from the basis to the opinion? Even though the opinion is considered by the courts as an acceptable one from the given basis, there is some disagreement on the connecting logic. Some will say that the connection depends alone on experience, others will argue that the connection is essentially a statistical one. A satisfying discussion of this would be too long to include in this paper; perhaps the following two points will help to indicate that

the problem is a bit more complex than might be evident at first.

1) The bulk of verifiable experience relates either to an entire set of prints as taken on a fingerprint card or a complete fingerprint, not to partial prints.

2) No statistical analysis appropriate to partial prints has, to our knowledge, ever been published.

It is hoped that this paper has stimulated the reader to more serious thought about the role of statistics in criminalistics as well as pointing the way to a clearer view of the subject. In discussing the illustrative investigations it has not been the purpose here to pass judgment upon the methods used by the authors of the articles, or to make precise analyses of the data, but only to indicate the kind of considerations which must be taken into account.