From: "A Critical Analysis of Quantitative Fingerprint Individuality Models"

David A. Stoney and John I. Thornton

Journal of Forensic Sciences, JFSCA, Vol. 31, No. 4, Oct. 1986, pp. 1187-1216.

Galton Model, Epoch 1892

Description of the Galton Model

Galton [1] made the first attempt to quantify fingerprint individuality. His basic approach was to divide a fingerprint into small regions, such that the ridge detail within each region could be treated as an independent variable.

Galton worked with photographic enlargements of fingerprints. The enlargements were placed on the floor, and paper squares of various sizes were allowed to fall haphazardly on the enlarged fingerprint. Galton then attempted to reconstruct the ridge detail which was masked by the paper squares, given the surrounding ridge flow. He sought the size of square region where he could successfully predict the actual ridge detail with a frequency of one half. Galton found that for a square region "six ridge intervals" on a side he was able to predict the hidden detail correctly with a frequency of one third, and consequently concluded that a square region with five ridge intervals on a side was very nearly the size he was seeking.

To ensure that any errors would *overestimate* the chance of fingerprint duplication, Galton used a six ridge interval square region, and then assumed a probability of one half for finding the existing minutia configuration, given the surrounding ridges. The total area of a complete fingerprint was estimated to consist of twenty four such square regions. Assuming independence among these regions, Galton calculated the probability of a specific fingerprint configuration, given the surrounding ridges, $P(C/R)$, using Eq 1.

$$P(C/R) = (1/2)^{24} = 5.96 \times 10^{-8} \qquad \text{(Eq 1)}$$

Galton next estimated the chance that a particular configuration of surrounding ridges would occur. Two factors were considered: (1) the occurrence of general fingerprint pattern type, and (2) the occurrence of the correct number of ridges entering and exiting each of the 24 regions. Galton estimated the probability for coincidence of pattern type, (termed factor b by Galton), as 1/16, and the probability that the correct number of ridges would enter and exit each region, (termed factor c by Galton), as 1/256. The latter estimate was largely arbitrary, and both were presented by Galton as grossly overestimating the "true" probabilities.

Combining the frequencies of finding the necessary ridge pattern outside the six ridge interval regions with the frequencies of finding all necessary ridge detail within the regions, Galton then predicted the probability of finding any given fingerprint, $P(FP)$, using Eq 2.

$$P(FP) = (1/16)(1/256)(1/2)^{24} = 1.45 \times 10^{-11} \qquad \text{(Eq 2)}$$

Assuming a world population of approximately 16 billion human fingers, Galton concluded that, given any particular finger, the odds of finding another finger which showed the same ridge detail would be approximately one in four.

Discussion of the Galton Model

Galton's model has been criticized by Roxburgh [*16*], by Pearson [*17*], and by Kingston [*20,21*]. Most of this criticism has focused on Galton's basic assumption that, given the surrounding ridges, there is probability of one half for the occurrence of any particular ridge configuration in one of his six ridge interval regions.

Pearson considered this assumption "drastic" and suggested an alternative approach for determining the probability of a particular configuration. Assuming that the position of a minutia may be resolved to within 1 square ridge interval, there would be 36 possible minutia locations within 1 of Galton's regions. Assuming 1 minutia in each of 24 independent regions, Pearson calculated the probability of any given configuration using Eq 3.

$$P(C/R) = (1/16)(1/256)(1/16)^{24} = 1.09 \times 10^{-41} \quad \text{(Eq 3)}$$

Pearson noted that the actual probability would be smaller for two reasons, first, because minutiae are not uniformly restricted to a single minutia in each Galton region, and secondly because of variability in minutia type.

Roxburgh's criticism of Galton's model [*4*] is more fundamental. He noted that Galton investigated only variation within single fingerprints, whereas his conclusions concerned variation among different fingerprints. This is a basic confusion by Galton of "within-group" and "between-group" variation. Roxburgh presented a series of illustrations showing that these two levels of variation need have no relationship with one another. Roxburgh conceded that Galton calculated the probability that Galton could reconstruct any particular print wholly in square regions, six ridge intervals on a side. Roxburgh argued, however, that the probability of one half for a correct guess is influenced by the size of the region relative to the ridge characteristics rather than by the variation or distribution of the characteristics themselves. If a one ridge interval square region were used an observer could always guess correctly. Under these circumstances it would be possible to reconstruct any particular print, given the ridges surrounding the squares, and yet not be able to say anything about variation between fingerprints.

Roxburgh pointed out that Galton's analysis proceeds as if he had surveyed a number of fingerprints, comparing square regions in corresponding positions within the prints. Had Galton done this, Roxburgh would have agreed with the analysis. The actual experiments, however, were quite different, and as a result Roxburgh rejected Galton's model.

Kingston [*20,21*] made somewhat the same point, noting that Galton's ability to guess the content of a square region is not an indication of the variation in actual fingerprint patterns. If Galton had shown that his region could contain only two configurations, given the surrounding patterns, Kingston would have accepted the basis for Galton's calculations. Seeing no evidence to support this contention, however, Kingston also rejected Galton's model.

In the view of the present authors, the criticisms of Kingston and Roxburgh are only partially valid. Galton intended his factors of b and c to summarize much of the variation among fingerprints. His factor b accounted for variation in general pattern type, and his factor c accounted for variation in the number of ridges entering and leaving each square region. Clearly the values of c would change radically if the size of the region were to vary. In particular, for the limiting cases where the ability to guess the content of the region approached certainty, the factor c would become very small. Unfortunately, Galton did not consider these factors in any detail, and instead chose arbitrary and excessively large estimates for both factors.

If we accept the concept of Galton's factors b and c, the question becomes whether or not Galton's experiments reasonably approximate a survey of corresponding regions in different fingerprints. It is clear that Galton had this in mind when he wrote [*23*]:

> When the reconstructed squares were wrong, they had none the less a natural appearance. . . . Being so familiar with the run of these ridges in fingerprints, I can speak with confidence on this.

> My assumption is that any one of these reconstructions represents lineations that might have occurred in Nature, in association with the conditions outside the square, just as well as the lineations of the actual prints.

Galton continued, making a further assumption:

> . . . when the surrounding conditions alone are taken into account, the ridges within their limits may either run in the observed way or in a different way, the chance of these two contrasted events being taken (for safety's sake) as approximately equal.

The weakness of Galton's model lies in the magnitude of the above approximation and in the arbitrary value chosen for c. We may justly criticize his final figure as a gross underestimate of fingerprint variability. Pearson's calculations of the variability in one of Galton's regions may be closer to reality, *but both his hypothesis and Galton's remain untested*.