Kinship and Correlation

Francis Galton, FRS

Few intellectual pleasures are more keen than those enjoyed by a person who, while he is occupied in some special inquiry, suddenly perceives that it admits of a wide generalization, and that his results hold good in previously-unsuspected directions. The generalization of which I am about to speak arose in this way.

In a book of mine called "Natural Inheritance," published about a year ago, I showed that the problems of family likeness fell entirely within the scope of the higher laws of chance; that we were thereby rendered capable of defining the average amount of family likeness between kinsmen in each and every degree, and of expressing the frequency with which the family likeness will depart from its average amount to any specified extent. It followed, very unfortunately for the general reader, that the problems of family likeness do not admit of being properly expressed except in the technical language of the laws of chance, and that it is impossible to discuss them adequately except through the medium of mathematics.

After the proofs of my book had been finally revised and had passed out of my hands, it happened that there was a delay of a few months before its actual publication. In the interim I was busily at work upon a new inquiry that had been suggested to me by two concurrent circumstances. One was a renewed discussion among anthropologists as to the information that the length of a particular bone—say a solitary thigh-bone dug out of an ancient grave—might afford concerning the stature of the unknown man to whom it belonged. It seemed to me that the anthropologists had not discussed their facts in the best statistical manner, and that they ought to have adopted a different form of treatment to any they had hitherto tried. The other circumstance arose out of the interest excited by M. Alphonse Bertillon, who proved that it was feasible to identify old criminals by an anthropometric process. The man who was suspected of having been convicted before was variously measured, and his measures were compared with those of all the criminals who had previously passed through the same process. By a contrivance analogous in principle to that on which a dictionary is constructed, the search through a register containing many tens of thousands of measures was performed with unexpected ease and precision.

Then a question naturally arose as to the limits of refinement to which M. Bertillon's system could be carried advantageously. An additional *datum* was no doubt obtained through the measurement of each additional limb or other bodily dimension; but what was the corresponding increase of accuracy in the means of identification? The sizes of the various parts of the body of the same person are in some degree related together. A large glove or shoe suggests that the person to whom it belongs is a large man. But the knowledge that a man has a large glove *and* a large shoe does not give us very much more information than if our knowledge had been confined to only one of the two facts. It would be most incorrect to suppose that the accuracy of the anthropometric method of identification increases with the number of measures in anything like the same marvellous rapidity that the security afforded by the better description of locks increases with the number of wards. The depths of the wards are made to vary quite independently of each other; consequently the addition of each new ward *multiplies* the previous security. But the lengths of the various limbs and bodily dimensions of the same person do not vary independently; so that the addition of each new measure adds to the security of the identification in a constantly-lessening degree. It seemed important, as well as highly interesting, to investigate this subject.

These two problems—namely, that of estimating the stature of an unknown man from the length of one of his bones, and that of the relation between the various bodily dimensions of the same person—are clearly identical. I was able to attack them at once, from happening to possess a sufficient number of sets of measures of different persons, each of whom had been measured in various ways. My first step was to take a large sheet of paper, ruled crossways; to mark a scale appropriate to the stature across the top and another appropriate to the left cubit (that is, the length from the bent elbow to the extended finger-tips) down the side. Then I began to "plot" the pairs of observations of stature and cubit in the same persons. Suppose, for example, an entry had to be dealt with of stature 69 inches, cubit 19 inches; then I should put a pencil mark at the intersection of the lines that corresponded to those values. As I proceeded in this way, and as the number of marks upon the paper grew in number, the form of their general disposition became gradually more and more defined. Suddenly it struck me that their form was closely similar to that with which I had become very familiar when engaged in discussing kinships. There also I began with a sheet of paper, ruled crossways, with a scale across the top to refer to the statures of the sons, and another down the side for the statures of their fathers, and there also I had put a pencil mark at the spot appropriate to the stature of each son and to that of his father.