On the Origins of the .05 Level of Statistical Significance

MICHAEL COWLES *York University, Canada*
CAROLINE DAVIS *York University, Canada*

ABSTRACT: *Examination of the literature in statistics and probability that predates Fisher's Statistical Methods for Research Workers indicates that although Fisher is responsible for the first formal statement of the .05 criterion for statistical significance, the concept goes back much further. The move toward conventional levels for the rejection of the hypothesis of chance dates*

Cochran feels that Fisher was fairly casual about the choice, "as the words *convenient* and *prefers* have indicated" (p. 16). However, the statement quoted above leaves no doubt about Fisher's acceptance of the level as the critical cutoff point, once he had decided upon it.

that themselves were based on the notion of "chance" and the unlikelihood of an event occurring. In 1657 Huygens (1657/1970) published a tract *On Reasoning in Games of Dice*, that was based upon the exchanges between Pascal and Fermat

All persons conversant with statistics are aware that this supposition brings Variability within the grasp of the laws of Chance, with the result that the relative frequency of Deviations of different amounts admits of being calculated, when these amounts are measured in terms of any self-contained unit of variability—such as

The probable error is, quite clearly, not the most probable of all errors, and the use of the term *error* in describing the variation of human characteristics perhaps carries the analogy with measurement error distribution a shade too far.

[redacted]

Q is the symbol for the semi-interquartile range, defined as one half of the difference between the

[redacted]

the point of rejection. However, from an exami-

nation of the various examples of χ^2 calculations

presented, with their corresponding probability values, one can see the range within which what might be described as a mixture of intuitive and

What specifically determined the adoption of this convention is largely a matter of speculation. Perhaps it was a combination of the preferred use of the *PE* as a measure by early statisticians like Galton and the influence of Pearson and his state-

Although, strictly speaking, the conventional rejection level of 3*PE* is equivalent to two times the *SD* (in modern terminology, a *z* score of 2), which [redacted] at some point the events begin to contradict the expectations they have formed, they introduce *cause* and abandon the idea of chance. The point [redacted]

[redacted]

hazard a guess that Fisher simply rounded off this [redacted] the degree of discrepancy and how it is interpreted [redacted]

Fisher, R. A. The arrangement of field experiments. *Journal of the Ministry of Agriculture*, 1926, 33, 503–513.

everyday life. *British Journal of Psychology*, 1925, *15*, 318–355.

Galton, F. *Natural inheritance*. London: Macmillan, 1889.

Pearson, K. Contributions to the mathematical theory of evolution: I. On the dissection of asymmetrical frequency curves.

Pearson, K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen