128

CONTRIBUTION TO A DISCUSSION OF J. NEYMAN'S PAPER ON STATISTICAL PROBLEMS IN AGRICULTURAL EXPERIMENTATION*

(incomplete extract)

* Neyman, J. (1935) Statistical problems in agricultural experimentation. *Journal of the Royal Statistical Society, Supplement*, 2: 107-154.

Professor Fisher said: . . . .

In the parts of the paper to which the greater space was devoted, Dr. Neyman had arrived, or thought he had arrived, at somewhat novel conclusions respecting the Latin Square scheme of lay-out; he had concluded that both it and the arrangement in Randomized Blocks were biased, and that in the case of the Latin Square the test of significance was vitiated. . . . .

If Prof. Fisher remembers aright, it was only about a year since another academic mathematician from abroad had been as much excited about having proved that the Latin Square was mathematically exact, as Dr. Neyman seemed to be at having proved it inaccurate.

The conflict of opinion suggested that there was something faulty about the purely mathematical approach. What was more surprising was the apparent inability to grasp the very simple argument by which the unbiased character of the test of significance might be demonstrated. In order to put this argument at its simplest, he would take a particular class of Latin Squares, namely the 5×5 squares from which Graeco-Latin Squares could be constructed. The number of these was

$$6 \times 24 \times 120$$

But the factor 120 would be ignored, since different squares arrived at by permuting the treatments among themselves, leaving the same groups of plots to be treated alike, would on the hypothesis to be tested, namely, that the treatments have no effects on the yields, lead to the same analysis of variance.

The fact that made this class of squares particularly simple was that corresponding to any one set of 120, there were three others related to it in such a way that any group of plots that received the same treatment in one set of arrangements received all five different treatments in the other sets. He would put on the board such a quadruplet of arrangements. We might represent by p, q, r, and s the sums of squares ascribable to treatments for these four arrangements; when any set of yields whatever were ascribed to the different plots.

	ABCDE BCDEA CDEAB DEABC EABCD	ABCDE DEABC BCDEA EABCD CDEAB	ABCDE CDEAB EABCD BCDEA DEABC	ABCDE EABCD DEABC CDEAB BCDEA
Sum of squares due to treatment ...	p	q	r	s
Sum of squares due to error	$q+r+s$	$p+r+s$	$p+q+s$	$p+q+r$

Then since on the hypotheses to be tested the treatments were wholly without effect, it followed that when the amount p was ascribed to treatments, the amount ascribable to error would be

$$q + r + s,$$

for these represented mutually orthogonal groups of four degrees of freedom each, being also orthogonal equally to treatments, rows and columns. Similarly, when q was ascribed to treatments, the amount ascribed to error would be

$$p + r + s,$$

and so in the other cases.

Randomization was designed to ensure that if any one of these arrangements might be chosen, the other three would be chosen equally frequently. Since there were in each case four degrees of freedom for treatment, and twelve for error, it was apparent that the mean square had in each case the average value

$$\tfrac{1}{16}(p + q + r + s).$$

The argument flowed directly from a mere understanding of the arithmetical processes used in the analysis, and showed that the elaborate specification of the hypothetical means, interactions, and correlations was quite irrelevant to a recognition of the unbiased nature of the test. . . . .