BOOK REVIEW

"Student's" Collected Papers. Edited by E. S. PEARSON and JOHN WISHART. Biometrika Office, London. pp. xiv + 224. 1942. 15 shillings.

Since the question is still asked from time to time, it may be well to state that 'Student' was the pen name of William Sealy Gosset. After taking his degree in mathematics and natural sciences at Oxford, Student worked as a brewer for the well-known firm of Messrs. Guinness—in Ireland from 1899 till 1935 and thereafter in London as head of the new Guinness brewery established there. He died in 1937 at the age of 61. Between 1907 and 1937 he published twenty-one papers and a few notes; these are issued in the present volume as a tribute from a group of his relatives and friends.

Student's name is almost universally attached to a single discovery, the t-test, which requires for a complete proof more powerful mathematical methods than he devised. These facts may cause his papers to be regarded as museum pieces by people who have not read them. Actually, in many respects no better model than Student could be suggested for a young statistician today. Since we sometimes speak derisively of courses of lectures in statistics where the methods and ideas are "20 years out of date", it is interesting to find that many of the fundamental ideas considered most in need of emphasis by forward-looking teachers today were in fact emphasized by Student in his writings over 30 years ago.

For example, his classical paper on the t-test, published in 1908, opens as follows: "Any experiment may be regarded as forming an individual of a 'population' of experiments which might be performed under the same conditions. A series of experiments is a sample drawn from this population.

Now any series of experiments is only of value in so far as it enables us to form a judgment as to the statistical constants of the population to which the experiments belong".

The idea is elaborated in a second paper published in the same year. "Note that the indefinitely large population need not actually exist. In Mr. Hooker's case his sample was 21 years of farming under modern conditions in England, and included all the years about which information was obtainable. Probably it could not actually have been made much larger without loss of homogeneity, due to the mixing with farming under conditions not modern; but one can imagine the population indefinitely increased and the 21 years to be a sample from this." We note here a further observation, not always appreciated, that in some lines of research samples which are both large and *homogeneous* do not exist. A clear conception of the relation between sample and population is evident in all his work: one further quotation, the summary from his 1926 paper, *Mathematics and Agronomy*, will suffice. "To sum up, in planning agronomic experiments use plenty of replications and make quite sure your results are capable of being considered to be a random sample of the population about which you wish to draw conclusions."

Another noticeable feature of Student's work is his concern about the extent to which the mathematical model underlying his statistical techniques really represents the facts of the data. Further, he is always ready to give his opinion, from considered thought, as to whether the discrepancies between the model and the facts will affect the conclusions materially. In his first paper, *On the error of counting with a haemacytometer*, he shows that the Poisson distribution applies to the total number of cells or corpuscles which are found in a division of the haemacytometer field. He is careful, however, to point out three ways in which deviations from the Poisson law may occur in practice: (i) there may be interference between the particles, though this is considered unlikely in high dilutions (ii) there may be clumping and (iii) (probably the most important, in his view) the drops taken out for counting may not represent the bulk of the liquid, giving rise to an additional sampling error superimposed on the Poisson error. In a later paper *An explanation of deviations from Poisson's law in practice* (*1919*), he makes a critical examination of the assumptions necessary for Poisson's distribution, and shows the type of distribution to be expected from the invalidity of one or another of the assumptions.

Similarly, in the opening paragraphs of the 1908 paper on the t-test, he notes that the assumption of normality in the parent distribution may not hold in practice. Then follows his opinion on the importance of this assumption: "It appears probable that the deviation from normality must be very extreme to lead to serious error."

His earliest papers deal with attempts to find several exact small-sample distributions for which he and other workers in Britain felt a need. Lacking the mastery of probability theory necessary for obtaining a rigorous solution, he used a combination of mathematics, intelligent guesswork and inferences from repeated samplings which he drew. For example, the steps in his development of the t-distribution were as follows:

(i) The distribution of the mean square, s^2, being unknown to British statisticians at that time, had first to be found. Student calculated the first four moments by algebraic methods. Noticing that the condition for fitting a Pearson Type III curve ($2\beta_2 - 3\beta_1 - 6 = 0$) held exactly, he inferred correctly that s^2 followed a type III curve.

(ii) He was aware that the joint distribution of s^2 and the sample mean $\bar{x}$ was necessary for a solution of the problem and further that the form of the joint distribution depended on the correlation between $\bar{x}$ and s^2. Having shown by algebra that s^2 is uncorrelated with either $\bar{x}$ or $\bar{x}^2$, he assumed (without further comment) that the distributions of $\bar{x}$ and s^2 were independent.

(iii) The final step to the t-distribution was performed rigorously.

(iv) The result was then checked against empirical data. He used the heights of 3000 criminals, which provided 750 samples of 4. Another set of 750 samples was supplied by the lengths of the left middle fingers of the same criminals.

In 1908 Student also published the distribution of the sample correlation coefficient r, in a bivariate normal population in which ρ is zero. His interest in the problem arose from a meeting of the Royal Statistical Society, in which an inquiry was raised about the significance of correlation coefficients derived from small numbers of cases. The authorities present, Yule, Hooker and Edgeworth, ventured their opinion on a specific instance, but had no general answer. Accordingly, Student set out to find the general distribution of r, without assuming ρ zero. His mathematics did not take him far; providing the solution only for samples of 2, where the frequency is zero except at -1 and $+1$. In order to learn what empirical results could teach about the form of the distribution, he then drew four sets of 750 samples, the sets having $\rho = 0, n = 4$; $\rho = 0, n = 8$; $\rho = .66, n = 4$; and $\rho = .66, n = 8$, respectively.

He considered first the set having $\rho = 0, n = 4$. Now, for samples of 4, the exact distribution, $C(1 - r^2)^{\frac{1}{2}(n-4)}$, reduces to a rectangular distribution. For his purpose, it would seem that Student had been unlucky in his choice of $n = 4$, since the rectangular distribution appears to furnish little or no clue as to the general form of distribution when n differs from 4. Actually, things could not have turned out better. Referring, as in the case of the distribution of s^2, to Pearson's curves, he selected a type II on account of the limited range of the distribution of r. The fitted curve was found to be $C\left(1 - \frac{r^2}{1.076}\right)^{0.272}$. This suggested to Student that the true form ought to be $C(1-r^2)^0$, giving a rectangular distribution; the general form was guessed as $C(1 - r^2)^{\frac{1}{2}(n-4)}$. The samples of 8 were then used merely as corroborative material. One wonders how Student would have fared, both in this case and in the case of s^2, if the Pearson system of curves had not been familiar to him. The method did not of course provide the much more complex distribution when ρ is not zero; however, Student was able to write down several important properties of the distribution for the guidance of mathematical statisticians who might be interested in the problem.

By similar methods he attempted (1921) to find the frequency distribution and the efficiency of Spearman's rank correlation coefficient. Much later (1936), he remarked in connection with this coefficient that, though it was not fully efficient, "it was so simple that when playing with *other* people's figures, for instance on a railway journey, it was the obvious one to use."

His paper, *The elimination of spurious correlation due to position in time and space*, though brief and incomplete in its treatment of the problem, established him as one of the pioneers in the use of the variate-difference method. His suggestion was to correlate the nth differences of x with those of y for $n = 1, 2, \cdots$ until a point was reached where the correlation ceased to change. This limiting correlation would be free from spurious effects due to trend factors. Such methods, he remarks, help to show "whether there really *is* a close connexion between the female cancer death rate and the quantity of imported apples consumed per head."

His association with field experiments conducted in Ireland for the purpose of developing high yielding varieties of barley suitable for brewing led to an interest in the design of experiments, to which he devoted seven papers. His philosophy on this subject, though it appeared to be increasingly out of harmony with the philosophy underlying Fisher's methods, was a natural outgrowth of Student's general attitude towards the applications of statistical techniques. He believed that, from general experience of soil fertility patterns plus local knowledge of the history and topography of the particular field in question, one could construct a systematic design which would be more accurate than the most appropriate randomized design. Secondly, although systematic dèsigns rely on nature, as it were, to supply the random element necessary for the application of probability theory, he believed, again from knowledge of field conditions, that the discrepancies between the mathematical model and the true fertility pattern would not seriously vitiate the estimate of error. How far he would have carried the advocacy of systematic arrangements is not clear. Nearly all of his discussions refer to arrangements for comparing only two varieties. Moreover, he stressed repeatedly the necessity for carrying out such trials over several seasons and at a considerable number of places, in order to sample weather and soil variations. In such cases the statistical significance of a result at a single place did not interest Student greatly.

A few quotations of his opinions on more detailed questions illustrate his understanding of the attitude of the farmer and agronomist and his direct mode of expression. Speaking of experimentation with large plots, he says "it has the advantage that the farmer, who always has a healthy contempt for gardening, may pay some attention to the results." On adjustments to yield data to allow for variations in soil fertility, he writes: "There is a great disadvantage in correcting any figures for position, inasmuch as it savors of cooking, and besides the corrected figures do not represent anything real." Long experience had made him more and more convinced that a serial correlation is inevitable in routine chemical analyses; accordingly, he advises: "The chemist who wishes to impress his clients will therefore arrange to do repetition analyses as nearly as possible at the same time, but if he wishes to diminish his real error he will separate them by as wide an interval of time as possible."

Re-reading of Student's papers has been a keen pleasure. It is hoped that the volume will enjoy a wide circulation amongst statisticians.

W. G. Cochran