A Biographical Glimpse of William Sealy Gosset

PHILIP J. BOLAND*

A brief historical survey is given of the man who became famous as the "Student" of Student's test of significance. Excerpts from various articles and letters are used to highlight the contact Gosset had with both Karl Pearson and Ronald A. Fisher.

KEY WORDS: Pearson curves; Student's-*t* distribution; Correlation coefficients; Small-sample statistics.

William Sealy Gosset, alias "Student," was an immensely talented scientist of diverse interests, but he will be remembered primarily for his contributions to the development of modern statistics. Born in Canterbury in 1876, he was educated at Winchester and New College, Oxford, where he studied chemistry and mathematics.

At the turn of the 19th century, Arthur Guinness, Son & Co. became interested in hiring scientists to analyze data concerned with various aspects of its brewing process. Gosset was to be one of the first of these scientists, and so it was that in 1899 he moved to Dublin to take up a job as a brewer at St. James' Gate. In 1935 he left Dublin to become head brewer at the new Guinness Park Royal brewery in London, but he died soon thereafter at the young age of 61 in 1937.

After initially finding his feet at the brewery in Dublin, Gosset wrote a report for Guinness in 1904 called "The Application of the Law of Error to Work of the Brewery." The report emphasized the importance of probability theory in setting an exact value on the results of brewery experiments, many of which were probable but not certain. Most of the report was the classic theory of errors (Airy and Merriman) being applied to brewery analysis, but it also showed signs of a curious mind at work exploring new statistical horizons. The report concluded that a mathematician should be consulted about special problems with small samples in the brewery. This led to Gosset's first meeting with Karl Pearson in 1905.

Karl Pearson (1857–1936) headed, at University College London, an industrious biometric laboratory that was much concerned with large-sample statistical analysis. Pearson had developed an extensive family of distribution curves, written an important paper introducing the χ^2 goodness of fit criterion, and initiated the journal *Biometrika*. (See Fig. 1.) Gosset was introduced by Pearson to correlation coefficients and large-sample theory. Soon afterward correlation coefficients of various types were being used extensively in new work at the brewery. Gosset soon realized, however, that modifications of Pearson's large-sample theory were needed for the special small-sample problems that were encountered in the brewery.

In 1906, Gosset received a year's leave from Guinness for specialized study. Most of the year he spent in close contact with Pearson's biometric laboratory in London, where he worked on small-sample problems and obtained good foundations in Pearson's theory of curves, correlation coefficients, and so on. It was during this time that he laid the basis for his most famous work, "The Probable Error of a Mean," which was published in *Biometrika* in 1908.

Student published 22 papers, the first of which was entitled "On the Error of Counting With a Haemacytometer" (*Biometrika*, 1907). In it, Student illustrated the practical use of the Poisson distribution in counting the number of yeast cells on a square of a haemacytometer. Up until just before World War II, Guinness would not allow its employees to publish under their own names, and hence Gosset chose to write under the pseudonym of "Student."

Most statistical analysis at that time dealt with large-sample theory. When investigating the mean of, say, a normal population, it was standard procedure to (a) calculate the sample mean $\bar{x}$, (b) assume the "population" standard deviation $= (\sum_1^n (x_i - \bar{x})^2/n)^{1/2}$, and (c) use the "normal" probability tables to make statements about the mean. This was, of course, reasonable for large n, but what about when dealing with small sample sizes? It was on this account that Student wrote in his classic paper "The Probable Error of a Mean":

> The usual method of determining the probability that the mean of the population lies within a given distance of the mean of the sample is to assume a normal distribution about the mean of the sample with a standard deviation equal to $s/\sqrt{n}$, where s is the standard deviation of the sample and to use the tables of the probability integral . . .
>
> There are experiments, however, which cannot be easily repeated very often; in such cases it is sometimes necessary to judge of the certainty of the results from a very small sample, which itself affords the only indication of the variability. Some chemical, many biological, and most agricultural and large scale experiments belong to this class, which has hitherto been almost outside the range of statistical enquiry.
>
> Again, although it is well known that the method of using the normal curve is only trustworthy when the sample is "large", no one has yet told us very clearly where the limit between large and small samples is to be drawn.
>
> The aim of the present paper is to determine the point at which we may use the tables of the probability integral in judging of the significance of the mean of a series of experiments, and to furnish alternative tables for use when the number of experiments is too few.

*Philip J. Boland is statutory lecturer in Mathematics at University College, Dublin, Belfield, Dublin 4, Ireland. An earlier version of this article appeared in the No. 8 (1983) issue of the *Irish Mathematical Society Newsletter*.

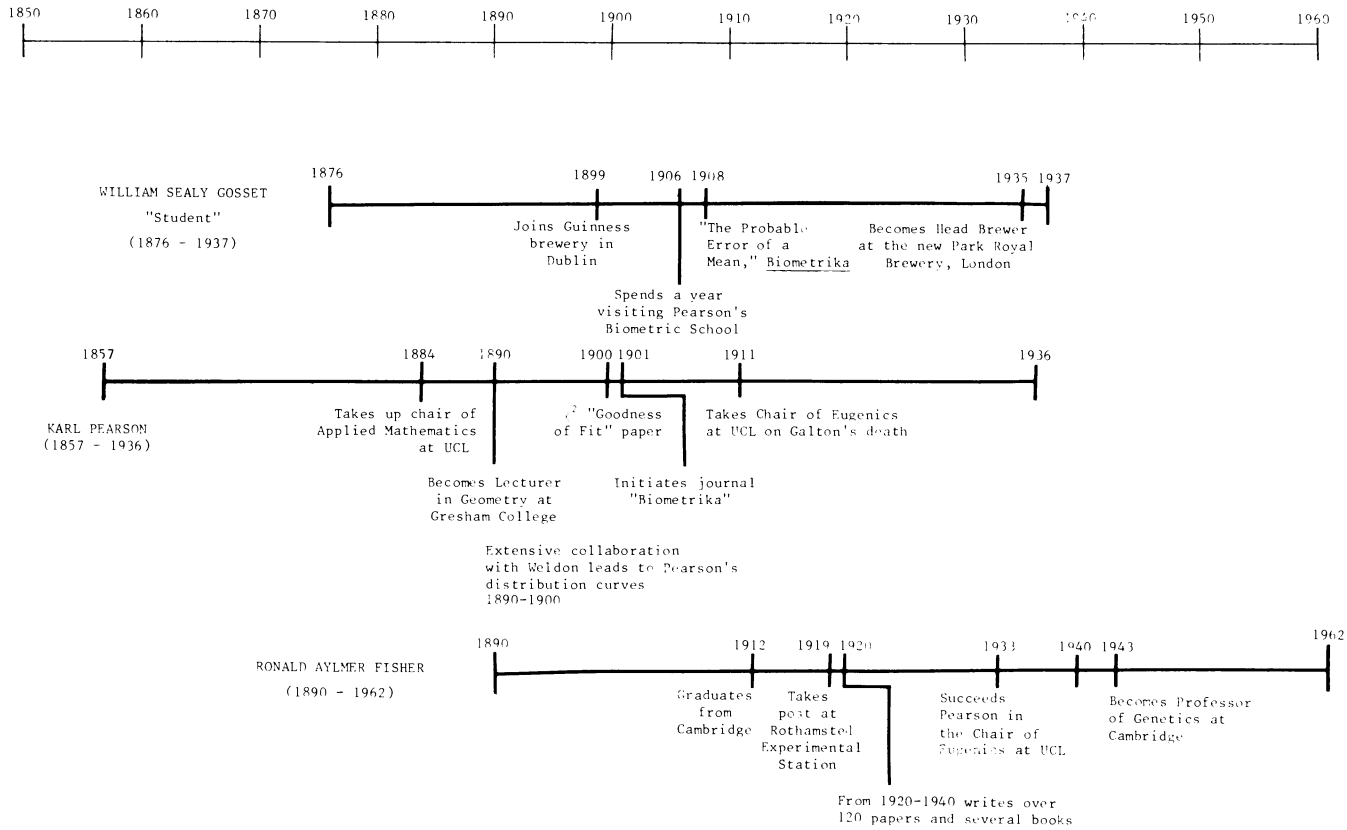

Figure 1. Chronology of Careers.

"The Probable Error of a Mean" was one of the first papers in which a clear distinction was made between population parameters and sample estimates of them (for example, (μ, σ^2) vs. $(\bar{x}, s^2)$). Assuming a random sample $(x_1, \ldots, x_n)$ from a normal population with mean μ and variance σ^2, Gosset went on to calculate the first 4 moments of $s^2 = \sum_1^n (x_i - \bar{x})^2/n$, and he then *inferred* that the distribution of s^2 is given by a Pearson type III curve. In modern language, Gosset inferred that $\sum_1^n (x_i - \bar{x})^2/\sigma^2$ is a χ^2 random variable with $n - 1$ degrees of freedom. This result had actually been discovered earlier but was unknown to English biometricians in 1908. Next Gosset showed that $\bar{x}$ and s^2 were uncorrelated, and inferred (again correctly but without sufficient reason) their independence. He then derived the distribution of $z = (\bar{x} - \mu)/s$ to be of the form $y(z) = C(1 + z^2)^{-n/2}$. This was, of course, the main result of the paper. In about 1922, the use of z was replaced by R.A. Fisher and Gosset with $t = \sqrt{n-1}\, z$, and hence it is now known as Student's t. (See Eisenhart 1979.)

Gosset checked the adequacy of his theoretical distributions for s^2 and z (equivalently t) on the basis of a data set of the height and middle finger length of 3,000 criminals. This was one of the first examples of a "simulated" experiment to be used in statistical research. Tables for his z statistic were also given, and Gosset concluded his paper with four practical illustrations of his method (i.e., the "t-test").

Because of the general lack of interest in "small-sample" statistics at the time, Gosset's method was not extensively used outside of the brewery for many years. Mainly as a result of the promotional efforts of R.A. Fisher (1890–1962), it eventually became widely used. R.A. Fisher was a student at Cambridge in 1912 when he first wrote to Gosset with a rigorous proof of the frequency distribution of the z statistic. In a subsequent letter, Fisher showed that one should be dividing by $n - 1$ instead of n in the formula for the sample standard deviation. Not fully understanding Fisher's mathematics, he wrote to Pearson (see E.S. Pearson 1970b):

12th September 1912 — Woodlands, Monkstown, Co. Dublin.

Dear Pearson,

I am enclosing a letter which gives a proof of my formulae for the frequency distribution of $z(=x/s)$, where x is the distance of the mean of n observations from the general mean and s is the S.D. of the n observations. Would you mind looking at it for me; I don't feel at home in more than three dimensions even if I could understand it otherwise.

The question arose because this man's tutor is a Caius man whom I have met when I visit my agricultural friends at Cambridge and as he is an astronomer he has applied what you may call Airy to their statistics and I have fallen upon him for being out of date. Well, this chap Fisher produced a paper giving 'A new criterion of probability' or something of the sort. A neat but as far as I could understand it, quite unpractical and unserviceable way of looking at things. (I understood it when I read it but it's gone out of my head and as you shall hear, I have lost it.) By means of this he thought he proved that the proper formula for the S.D. is

$$\frac{\Sigma(x-m)^2}{n} \quad \text{vice} \quad \frac{\Sigma(x-m)^2}{n-1}.$$

This, Stratton, the tutor, made him send me and with some exertion I mastered it, spotted the fallacy (as I believe) and wrote him a letter showing, I hope, an intelligent interest in the matter and incidentally making a blunder. To this he replied with two foolscap pages covered with mathematics of the deepest dye in which he proved, by using n dimensions that the formula was, after all

$$\frac{\Sigma(x-m)^2}{n-1}$$

and of course exposed my mistake. I couldn't understand his stuff and wrote and said I was going to study it when I had time. I actually took it up to the Lakes with me—and lost it!

Now he sends this to me. It seemed to me that if it's all right perhaps you might like to put the proof in a note. It's so nice and mathematical that it might appeal to some people. In any case I should be glad of your opinion of it.

[*The rest of the letter is concerned with tuberculosis death rates, a matter about which Gosset was already in correspondence with Pearson.*]

Yours very sincerely,

W.S. Gosset

An indication of the importance Karl Pearson put on Gosset's small sample theory at that time can be seen from his reply to this letter on 17 September 1912. Commenting on the subject of the standard deviation, Pearson (1970a) remarked that it made little difference whether the sum of squares was divided by n or by $n-1$, "because only naughty brewers take n so small that the difference is not of the order of the probable error!" R.A. Fisher, on the other hand, felt that Gosset never fully realized the importance of his discovery. In a tribute to Student in 1939, Fisher wrote:

How did it come about that a man of Student's interests and training should have made an advance of fundamental mathematical importance, the possibility of which had been overlooked by the very brilliant mathematicians who have studied the Theory of Errors?...

One immense advantage which Student possessed was his concern with, and responsibility for, the practical interpretation of experimental data. If more mathematicians shared this advantage there can be no doubt that mathematical research would be more fruitfully directed than it often is.

Gosset was also keenly interested in the problem of determining the distribution of the sample correlation coefficient r of two normal random variables (which were, say, uncorrelated). This was the subject of another of his papers, entitled "The Probable Error of a Correlation Coefficient." Using the same criminal data as in "The Probable Error of a Mean," he simulated an experiment in which he observed 750 sample values of r (sample size 4) from a bivariate normal population with no correlation. A graphical representation of the results suggested to him that for sample size 4, the distribution of r was rectangular (i.e., uniform on $[-1, 1]$). With his knowledge of Pearson's family of distribution curves, he felt a Pearson type II curve was the only one suitable for the distribution of r based on samples of size n. He wrote: "...working from $y = y_0(1-x^2)^0$ for samples of size 4, I guessed the formula $y = y_0(1-x^2)^{(n-4)/2}$." He then showed by simulation that the formula worked well for samples of size 8. He made some comments about the distribution of r when sampling from a population with (nonzero) correlation, but this more general problem was out of his grasp. Student (1908b) concludes the paper as follows:

It has been shown that when there is no correlation between two normally distributed random variables, $y = y_0(1-x^2)^{(n-4)/2}$ gives fairly closely the distribution of r found from samples of size n.

Next the general problem has been stated and three distributions of r have been given which show the sort of variation which must occur. I hope they may serve as illustrations for the successful solver of the problem.

The successful solver was, of course, R.A. Fisher, who published the results (which justified Gosset's "guesswork") in *Biometrika* in 1915.

Guinness, a large consumer of barley, was very interested in agricultural experimentation. Gosset eventually became involved in the planning and interpretation of such experiments, many of them carried out under the supervision of the Irish Department of Agriculture. His knowledge and advice on such matters was held in high regard and he corresponded extensively with many other experimentalists. It can be said that Gosset did a considerable amount of pioneering work in the areas of analysis of variance and experimental design, and he had a lot to do with Fisher's getting so interested in those areas. He did, however, come into open controversy with R.A. Fisher and his Rothamsted school on the subject of balance (as opposed to randomness) in experimental design. In a letter to Fisher on 18 April 1928, he wrote:

The fact is that there are two principles involved in the Latin square of which I attach the greater importance to the balancing of the error and you to the randomisation. It is my opinion that in the great majority of cases the randomisation is supplied to any properly balanced experiment by the soil itself though of course where the ground has been used for experimenting before or for any other reason has met with a "straight edge" lack of uniformity in recent years it is better to supply it artificially.

Gosset was not completely happy with Fisher's randomized blocks because he felt that often a greater accuracy could be obtained with a balanced arrangement within the blocks. He was furthermore unwilling to use a plot arrangement determined by the toss of a coin, if the arrangement so obtained was biased in relation to already known fertility knowledge of the field. In his final paper, "Comparison Between Balanced and Random Arrangements of Field Plots," published posthumously, Gosset wrote:

It is, of course, perfectly true that in the long run, taking all possible arrangements, exactly as many misleading conclusions will be drawn as are allowed for in the tables, and anyone prepared to spend a blameless life in repeating an experiment would doubtless confirm this; nevertheless, it would be pedantic to continue with an arrangement of plots known beforehand to be likely to lead to a misleading conclusion.

Although this was a subject on which they never agreed, their friendship and mutual respect for one another did not suffer. In his tribute to Student, Fisher (1939) wrote

Certainly though he practised it, he did not consistently appreciate the necessity of randomization, or the theoretical impossibility of obtaining systematic designs of which both the real and the esti-

mated error shall be less than those given by the same plots when randomized; this special failure was perhaps only a sign of his loyalty to colleagues whose work was in this respect open to criticism.

The theory of evolution is another area of interest that intrigued Gosset. Darwin's Theory of Natural Selection suffered a setback at the turn of the century with the discovery of Mendel's work. Many geneticists came to believe that hereditary traits are conditioned by a limited number of genes and that selection involving a limited number of genes cannot lead to important evolutionary effects. They felt that evolutionary progress and development must occur as the result of new mutations.

Gosset (and Fisher), however, did not believe in this mutation theory. About 1932 Gosset became interested in some experimental work of H.L. Winter involving oil in maize. He felt that here there was evidence of a large number of genes at work, and that by (natural) selection important evolutionary changes could take place. Using Winter's work, Gosset wrote two papers in support of the Theory of Natural Selection ("Evolution by Selection—The Implications of Winter's Selection Experiment," *Eugenics Review,* 1933, and "A Calculation of the Minimum Number of Genes in Winter's Selection Experiment," *Annals of Eugenics,* 1934). He sought the support of Fisher (and his mathematical ability) in this cause, and persuaded him to write up a note in this direction for *Nature.* The following letter, written to Fisher in early 1933 on this matter, is a good illustration of Gosset's lively sense of humor.

Holly House, Blackrock,
Co. Dublin.
16.1.33

Dear Fisher,

When I persuaded you to write up the mathematics of myriad gene selection in Nature I was so pleased with the idea of having got it done properly, that I overlooked the fact that I have put you in the position of appearing to "butt in".

That being so I am taking the liberty of suggesting an opening sentence, not of course that I would wish you to use exactly my words, but something of the sort should give you a locus standi.

"In the January number of the Eugenics Review, Student has drawn attention to an experiment in selection carried out by Winter and has shown that the remarkable result of that experiment is consistent with the theory that the genes which affect the percentage of oil in maize are to be numbered by hundreds; further that given such large numbers of genes continued selective breeding will necessarily result in the production of individuals and ultimately of sub-races completely outside the original range of variation.

The argument has, of course, a mathematical basis and Student has invited me to examine it in a more general way than he has considered appropriate to an article in the Eugenics Review."

Of course if you want to be truthful you can substitute "can himself" (which has the merit of brevity) for "considered appropriate . . . Review" but people who don't know Student might consider it rude, which would be a pity.

And here I think I hear you murmur "Damn the man, why doesn't he refrain from teaching his granny. He's as fussy about his little bit of stuff as a hen with one chick".

To which I reply, "I am, curse you, for the very good reason that I'll never have the chance to incubate an egg which interests me so much".

Cluck. Cluck. Cluck. Cluck.

Fact is that until just recently I was so much taken up with the first part of the thing, "myriad genes", that I overlooked the fact that the second is really an essential cog in the mechanism of Darwinian selection. For at least twenty-five years I've been reading that the continued accumulation of infinitesimal variations can do nothing and all the time I've felt in my bones that Darwin was right.

Cluck. Cluck. Cluck.

And now I have been vouchsafed a vision, . . . and am filled with insufferable conceit . . . for the nonce I too am among the prophets, a mere Obadiah, but still among the prophets. And if any one were to offer to make me a Doctor of Divinity on the strength of it, I'd accept with conscious pride and flaunt a scarlet gown through the scandalized streets of Oxford.

Bear with me, Fisher, laugh with me tonight; tomorrow . . . when I'm sane again . . . when I know that my little bit was discovered in 1896 and put into better words than mine often since then when I have been shown that my essential cog will hardly ever fit into the machine and when it does is a clog . . . then I'll laugh with you—at myself.

Cluck. Cluck.

Yrs. v. sincerely,
W.S. Gosset

That Gosset was a talented, but modest, man there can be no doubt. Stella Cunliffe (1976) relates the following about Gosset:

Once, in 1937, a young statistician who went to consult him, said pompously:

"On behalf of fellow statisticians, I would like to thank you for all that you have done for the advancement of statistics."

to which Gosset replied:

"Oh, that's nothing—Fisher would have discovered it all anyway."

Gosset led a very full life and pursued a variety of hobbies (see McMullen 1970). He was a keen gardener with a particular interest in fruit. In the 1920's, he developed some logan-raspberry hybrids, two of which went by the names of "jamberry" and "Paddyberry." He made two barley crosses (known as Student I and Student II) in his own garden and accelerated their development by having one generation grown in New Zealand. Gosset was also a good carpenter and built several boats. One of these, built with a rudder at each end, was for the particular benefit of fly fishermen. The design of this boat was described in *Field* in March 1936. Gosset also took an avid interest in fishing, hunting, and golfing. However, as Fisher writes in his tribute to Student in 1939,

His life was one full of fruitful scientific ideas and his versatility extended beyond his interests in research. In spite of his many activities it is the student of Student's test of significance who has won, and deserved to win, a unique place in the history of scientific method.

[*Received September 1983. Revised December 1983.*]

REFERENCES

BOX, JOAN FISHER (1981), "Gosset, Fisher, and the *t*-Distribution," *The American Statistician,* 35, 61–67.

CUNLIFFE, STELLA (1976), "Interaction," *Journal of the Royal Statistical Society,* Ser. A, 139, Part 1, 1–19.

EISENHART, CHURCHILL (1979), "On the Transition From Student's *z* to Student's *t*," *The American Statistician,* 33, 6–10.

FISHER, R.A. (1939), "Student," *Annals of Eugenics,* 1–9.

GOSSET, W.S. (1970), *Letters from W.S. Gosset to R.A. Fisher, 1915–1936* (with summaries by R.A. Fisher and a foreword by L. McMullen), Dublin: Arthur Guinness, Son & Co. (Dublin) Ltd. Issued for private circulation.

McMULLEN, L. (1970), "Student as a Man," *Studies in the History of Statistics and Probability,* Volume I, Charles Griffin and Co., 355–360.

PEARSON, E.S. (1970a), "Student as Statistician," *Studies in the History of Statistics and Probability,* Volume I, Charles Griffin and Co., 360–403.

——— (1970b), "Some Early Correspondence between W.S. Gosset, R.A. Fisher, and Karl Pearson, With Notes and Comments," *Studies in the History of Statistics and Probability,* Volume I, Charles Griffin and Co., 405–417.

Student (1908a), "The Probable Error of a Mean," *Biometrika*, 6, 1–25.

——— (1908b), "The Probable Error of a Correlation Coefficient," *Biometrika,* 7, 302–310.

——— (1938), "Comparison Between Balanced and Random Arrangements of Field Plots," *Biometrika,* 29, 363–379.

——— (1942), *Student's Collected Papers,* eds. E.S. Pearson and J. Wishart, with foreword by Launce McMullen, Cambridge: Cambridge University Press. Published for the *Biometrika* Trustees.