



Statistics Born in Brewery

Lenka Viskotová, David Hampel

Department of Statistics and Operation Analysis, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic

* corresponding author: lenka.viskotova@mendelu.cz

Abstract

One of the most important statistical tools is the t-test and its associated Student distribution. These techniques resulted from the work of W. S. Gosset, who published under the pseudonym Student. This paper presents the biographical facts and context of Gosset's statistical research with his work at the Guinness Brewery. Gosset's collaboration with important personalities of that time and his work are described. His core article, which laid the foundations of the small sample theory, is discussed in more detail.

Key words: brewery Guinness, experiment evaluation, Gosset, small sample theory, student distribution, t-test

1 Introduction

A statistical t-test is a well-known technique, which belongs to the most important statistics tools. Its widespread use can be illustrated, inter alia, as follows: in Fermentation Industry there had been published by the date of April 18, 2018 at least 288 articles in which the application of the t-test appears. The t-test is commonly used for these purposes:

- A one-sample test for equality of a mean and a specified constant.
- A two-sample test for equality of means for independent samples.
- A "paired" test for equality of means for dependent samples.
- A test of whether the regression parameter is equal to zero or to some specified value.

This test was introduced at the beginning of the last century as a response to the problems of William Sealy Gosset, a brewer working for the Guinness brewery in Dublin, with the statistical check of beer production. Gosset, published under the pen name Student, had a great deal to the development and acceleration of the statistical thinking of his time.

In an interview of the journal *Statistical Science* (Laird, 1989) the distinguished statistician F. N. David talked about the history of statistics in the twenties and the thirties of the 20th century and in this context she mentioned among other important representatives of that period Gosset, Ronald Fisher, Egon S. Pearson and J. Neyman. David built herself as an observer who watches "all the protagonists from a worm's eye point of view". Her overall assessment is certainly interesting:

"I think he [Gosset] was really the big influence in statistics [...] he asked the questions and Pearson and Fisher put them into statistical language, and then Neyman came to work with the mathematics. But I think most of it came from Gosset."

David did not talk about all the statistic of that time; she focused on the development of sampling with relatively few observations. Even so, the above conclusion may seem surprising, because Gosset is known especially because of a single, though crucial, result: the Student t-test. Gosset's contribution is, however, much broader, and his link with the brewing industry, specifically with the position of the Guinness brewer, is really interesting.

The aim of this article is to summarize the basic facts and interesting features of Gosset's biography, including

his research in the field of statistics, which produced significant results. The structure of this paper is following: firstly, the important aspects of Gosset's life (drawn upon (Boland, 1984; Boland, 2011; Box, 1987)) are mentioned, as well as his links to the personages with whom he collaborated (Box, 1987; Lehmann 1999), the last part is dedicated to Gosset's work (Boland, 1984; Boland, 2011), including his most famous work (Student, 1908); more can be found in (Boland, 1984; Box, 1987; Lehmann, 1999).

2 Life of Gosset

William Sealy Gosset was born on June 13, 1876 in the English town of Canterbury as the oldest of five children in the family of a Colonel in the Royal Engineers. As a boy, he joined the Royal Military Academy in Woolwich to become a Royal Engineer, just like his father, but he had to leave the school due to impaired eyesight. Immediately afterwards Gosset gained a scholarship at Winchester College, and upon graduating he enrolled at New College in Oxford where he achieved degree in mathematics (1897) and chemistry (1899) with great school results. After studies, he started working as a brewer at the Guinness brewery in Dublin.

At that time, the Guinness Brewery introduced a new personnel policy and hired a number of talented young men with top-class chemistry studies at Oxford or Cambridge University and named them as brewers. Gosset joined the Guinness Brewery in Dublin on October 1, 1899 as the younger brewer. Other breweries employed chemists only occasionally in research laboratories, but Guinness also accepted them in top management positions originally held by Guinness family members only. When a new brewer had completed two years of practice in various parts of production under the supervision of experienced managers, he was in charge of particular department of brewery production and was solving research tasks. Given that Gosset also studied mathematics at Oxford, it was natural for him to pay attention to the use of mathematical methods in the brewing process.

At the beginning of the 20th century, the Guinness Brewery established an experimental malt and brewery to test small samples of raw materials and changes in

beer technology. As an experimental brewer, Gosset was approached to assess the following problem: If we get the result based on the experimental series, what confidence can we have in that the result is correct and does not depend only on random changes in the process or observation?

In July 1905, during a vacation in England, Gosset was invited to the summer residence of Karl Pearson, already established statistics. They had a long discussion that Gosset considered very useful; later he stated that Pearson "was able in about half an hour to put me in the way of learning the practice of nearly all the methods then in use".



Figure 1 W. S. Gosset, 1899. Source: Box (1987)

Upon his return from the vacation, Gosset applied the scientific knowledge obtained from this meeting to solve problems arising from the experimental brewery. The management of the brewery was enthusiastic about Gosset's work, and Guinness arranged for Gosset to spend the academic year 1906–1907 at the Pearson's depart-

ment at University College London “to take up the study of the law of error¹, the working of which we have found of great service in the brewery”. Gosset, apart from other things, dealt with Pearson’s probability distributions and its correlation coefficients. He was aware that Pearson’s methods based on the large number of observations, that were widely used by biometricians at that time, would need to be modified for specific problems based on a small number of samples produced in the brewery.

In this period, the basis for Gosset’s most famous work, “The Probable Error of a Mean”, published in the journal *Biometrika* in 1908, was laid and will be discussed in more detail below. Guinness allowed Gosset to publish this article, but the management of the brewery decided that Gosset’s name should not appear in the publication. This regulation concerned not only Gosset, but also other brewers of Guinness. They were advised to use the pseudonym “Pupil” or “Student”. So Gosset decided to publish just under the name Student. Thanks to his mathematical abilities and experimental intuition he has created, with the help of R. A. Fisher², the method currently known as a Student’s t-test.

In 1907, Gosset was named the Brewer in charge of the Experimental Brewery at Guinness, and later he founded the Statistics Department. Gosset himself continued with statistical research only in his spare time, as statistical analyses did not belong to his workload.

Despite undisputable results, Gosset remained a modest man. Stella Cunliffe, his colleague, said in the context of Gosset (Cunliffe, 1976): “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 was married and had three children. He led a very active life and had a variety of hobbies (McMullen, 1970). He was a keen gardener with a special interest in growing fruit. In his own garden, he also bred two varieties of barley (known as Student I and Student II), using the acceleration of breeding by one generation grown in New Zealand. Gosset was also a skilled carpenter and built several ships. One of them, with a rudder at both ends, was specially designed for fly-fishers (the construction of this ship was described in 1936). Gosset was also a passionate fisherman, hunter, and golfer. As Fisher writes in his text in 1939 giving tribute to Student, his life

was full of fruitful scientific ideas, and his versatility far outweighed his research interests.

In 1935, Gosset left Dublin as he was transferred to a new Park Royal Brewery in London (the first Guinness brewery built outside Dublin). Here he became the Head Brewer, but only for a short time until his premature death on a heart attack on October 16, 1937.

3 Notable Collaborators of Gosset

It is undisputed that Gosset has influenced many of the then and future scientists and practitioners. The following part of paper will briefly discuss Gosset’s collaboration with scientists Fisher and Person, and barley breeder Edwin S. Beaven.

3.1 Gosset and R. A. Fisher

In 1912, a 22-year-old student at Cambridge, R. A. Fisher, was introduced to Gosset by his doctoral supervisor, astronomer Frederick Stratton. This acquaintance led to mutual co-operation, thanks to which Gosset later received from Fisher the evidence of the distribution of Z-statistics. Gosset wanted Fisher to derive distributions of other statistics, such as correlation, F-test of equality of variances, regression parameters; the real basis of these problems surely captured Fisher. Gosset also inspired Fisher to break the assumption of normality for individual observations. Of course, Fisher wrote a number of original papers for the small sample theory, which in no way build on Gosset. These include, for example, variance-stabilizing transformation, permutation tests, the design of experiments (including randomization), and the concepts of sufficiency and of likelihood. But still, we can assert that Gosset influenced Fisher’s scientific orientation to a large extent.

3.2 Gosset and E. S. Pearson³

Pearson’s contributions to the small sample theory are of a dual nature. On the first hand, it consists of simulation studies of robustness published in 1931 in “The analysis of variance in cases of non-normal variation”, the second branch is his work with Neyman, when he developed the theory formulated into the so-called Neyman-Pearson theorem. Key ideas for both were brought by Gosset. As Pearson, in the commentary of Gosset’s letter of 1926, himself acknowledges (Pearson, 1990): “His letter left me with two fundamental ideas: (a) The rational human

¹ The notion „normal distribution“ is currently used for this term.

² Gosset and Fisher’s cooperation will be discussed later.

³ Egon S. Pearson was the son of Karel Pearson, founder of mathematical statistics.

mind did not discard a hypothesis unless it could conceive at least one plausible alternative hypothesis. (b) It was desirable to explore the sensitivity of his z-test to departures from normality in the population, i.e. the question which was later to be termed by G. E. P. Box that of robustness." Here, F. N. David's assessment that the main areas of Pearson's research actually stemmed from Gosset's thoughts, is confirmed.

3.3 Gosset and E. S. Beaven

Beaven was a Warminster maltster whose passion was to breed barley. In 1904, Beaven started cooperation with Guinness, who began to make intense interest in the production of malt including the influence of barley varieties on it. At that time, Beaven had already selected lines of several varieties, and with his friends from Cambridge University, Biffin and Wood, began to hybridize them to create better varieties of barley. They became enthusiastic genetics and breeders, among other things thanks to Mendel's hybridization articles. Beaven also started to organize experiments to test certain hybrids with each other and against standard varieties, and at this moment, in 1907, Gosset entered into his research. Gosset analysed experiments with respect to the differences between pairs of neighbouring.

4 Work of Gosset

In 1904, Gosset wrote for Guinness an internal report entitled "The Application of the 'Law of Error' to Work of the Brewery", in which he highlighted the importance of using probability theory in evaluating the results of the brewery experiments. In 1905, when correlation methods were frequently used in the brewery, he published the internal report "The Pearson Co-Efficiency of Correlation". The Guinness Board endorsed this paper as follows: "It is to be noted that for the present the Karl Pearson method of correlation is to be adopted in all reports which admit of it being done, and this endorsement is to be regarded as a formal instruction in the matter." The next Gosset report from 1906, "Our present knowledge of the relationship between life and Ready Formed Sugars" mentions that "of Malts made from the same barley those with more R.F.S. (Ready Formed Sugars) would also produce the less stable beer." The Board Endorsement on this finding said "This report calls for no formal endorsement, but is especially interesting as illustrating the great utility of the new statistical method introduced in the Brewery by Mr. Gosset."

William Sealy Gosset was a scientist whose scientific work was in direct connection with the practice. He published 22 scientific papers and 355 reports for Guinness

Brewery, see (Boland, 2011). Gosset's research and developed methods reflected his efforts to find solutions to problems associated with the production of stout, in response to variations in barley, hops, malt, etc. The individual reports then described the results of his experiments, as an example, "The value of barley analysis in predicting the extract in Experimental Brewery" (1906), "The connection between moisture in malt and life" (1909), "Arsenic in malt" (1921), "Effect of Temperature on Sales of Guinness's Stout" (1929), "Correlation between Sales and Unemployment" (1934), etc.

The first Gosset's scientific work had a title "On the Error of Counting with a Haemacytometer" (published in *Biometrika* in 1907). In this work he demonstrated the practical use of the Poisson distribution for yeast cells counting with hemocytometer.

Guinness, as a large barley consumer, had a great interest in conducting agricultural experiments. Gosset was involved in the planning and evaluation of these experiments, many of them were carried out under the supervision of the Irish Department of Agriculture. Gosset's knowledge and recommendations on the issue had considerable respect among the professional public, and he also maintained extensive correspondence with other experimentalists. It could be said that Gosset was a pioneer in the field of analysis of variance and experimental design. And he made Fisher begin to tackle those areas.

As can be seen in (Boland, 2011), Guinness decided to start advertising in 1929. This decision was preceded by extensive debates, as relations describing the effect of advertising on sales are very complicated. Gosset created an analysis that pointed to the fact that in the first part of 1929 "there had been a fairly close correspondence between expenditure on advertising and variations in sales, but that more recently sales have continued to rise in spite of the decrease of money spent. The various methods of market research in general use today were of course hardly known at this time, so we can say that Gosset build up basic concepts of statistical evaluation of advertising" (Dennison and MacDonagh, 1998).

In 1930, Gosset wrote the report "On the Employment of Labor Staff" comparing the age distribution of the workers in the brewery with the situation in 1914. He set up interesting forecasts and recommendations for hiring of labourers. In another report, he presents to the brewery an actuarial analysis (Dennison and MacDonagh, 1998) needed for the revision of the pension and life insurance system. Gosset investigated the mortality rate of the brewery workers in the period 1878–1902 and showed that their mortality was lower than that of skilled workers registered in Dublin trade unions, but

significantly higher than amongst the same workers in England and higher than the mortality rate used at that time in the insurance of healthy men⁴.

5 Main paper of Gosset

Gosset gained his knowledge of statistics from the books “G. B. Airy: On the Algebraic and Numerical Theory of Errors of Observations and the Combination of Observations” from 1879 and “M. Merriman: Method of Least Squares” (1884). Both publications emphasize that errors are usually sums of a large number of independent small components and therefore are approximately normally distributed. The normal distribution is called the “Law of Probability of Errors”. Merriman in his book states: “Whatever may be thought of the theoretical deductions of the law of probability of error, there can be no doubt that its practical demonstration by experience is entirely satisfactory.” Airy is more cautious in his claims: “It must always be borne in mind that the law of frequency of errors does not exactly hold except the number of errors is infinitely great. With a limited number of errors the law will be imperfectly followed; and the deductions, made on the supposition that the law is strictly followed, will be or may be inaccurate or inconsistent.” It can be said, therefore, that Gosset’s assumption of normality was based on the established habits of that time.

In 1903, Gosset was able to calculate a standard error. In 1904, he wrote a report on this topic for the brewery, and one year later he was given the opportunity to consult unresolved problems with Karl Pearson, an already well-known biometrician.

One of the main difficulties in the evaluation of experimental data was the fact that the samples were always small. For example, barley experiments took place on four farms, where each variety was tested on one plot. Gosset realized that estimating the mean based on only four samples could not be accurate and the error of estimating the standard deviation cannot be insignificant. The key question for Gosset was how much the error limits should be wider if we work with the mean and the standard deviations instead of the theoretical parameters.

Gosset’s research has given rise to a new methodology that deals with small-sample problems. According to the paper (Lehmann, 1999), the basis of this methodology was laid down in three stages. In the first stage, in cooperation with Fisher, distributions of the statistics

were developed to test means, variances and correlation and regression coefficients assuming normality. In the second stage, with the help of E. Pearson, the robustness of these distributions was verified for the situation when the assumption of normality was broken. In the final stage, thanks to Neyman and E. Pearson, the foundations for a suitable choice of statistical tests were established not only for the case of normality but also for a general distribution.

Gosset’s work “The Probable Error of a Mean” published in *Biometrika* in 1908 was the first one where population parameters were clearly distinguished from their sample estimates. Gosset introduced his article as follows (Student, 1908):

“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. In a great number of cases the question finally turns on the value of a mean, either directly, or as the mean difference between the two quantities. If the number of experiments be very large, we may have precise information as to the value of the mean, but if our sample be small, we have two sources of uncertainty:

1. owing to the ‘error of random sampling’ the mean of our series of experiments deviates more or less widely from the mean of the population, and
2. the sample is not sufficiently large to determine what is the law of distribution of individuals.”

The essence of Gosset’s finding is to derive the distribution of the random variable

$$Z = (m_x - \mu)/S,$$

where m_x denotes the average and $S^2 = \Sigma(X_i - m_x)^2/n$, while X_1, \dots, X_n are independent and identically distributed random variables with normal distribution $N(\mu, \sigma^2)$. Note that the current definition of Z and S^2 differs from the Gosset notation. For Gosset, who was more a chemist than a mathematician, the derivation of the formula of this distribution was a difficult task. Although he managed to get the right answer, he was unable to provide an accurate proof. In order to obtain the distribution of Z , he calculated the correlation m_x and S and identified them as

⁴ Gosset considered the spread of tuberculosis as one of the major causes of increased mortality in both Dublin and the Guinness Brewery.

independent after finding that they are uncorrelated. The distribution of Z was derived, and this is what we perceive as the main result of his research. In about 1922, Z was replaced by Fisher a Gosset with the new distribution $t = (n-1)^{1/2}Z$, which we now know under the name Student (Eisenhart, 1979).

However, although Gosset had invested a great deal of effort in this result, this was not why the paper's influence remained so crucial. The main benefit was that he created a new point of view on the issue. He pointed to the need to work with methods suitable for small samples where the assumption of normality is not reasonable. He also brought the key idea that accurate results can be obtained if we know the distribution of observed values.

"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."

Calculating of these tables required knowledge of the normal curve of the appropriate distribution of the observations. However, Gosset states that for small samples "it is usual, however, to assume a normal distribution [...], since some law of distribution must be assumed it is better to work with a curve whose area and coordinates are tabled, and whose properties are well known. This assumption is accordingly made in the present paper, so that its conclusions are not strictly applicable to populations known not to be normally distributed; yet it appears probable that the deviation from normality must be very extreme to lead to serious error". This idea is illustrated in Figure 2.

Gosset illustrates the use of his distribution on three examples, including a paired comparison experiment, which was converted into one-sample using differences. The paper was supplemented by a Z-distribution table for sample sizes 4 to 10. Later, the table was extended for observation numbers from 2 to 30 (Student, 1917).

The importance of the whole article can be read from the original Gosset's conclusions (cited from the original paper (Student, 1908)):

1. A curve has been found representing the frequency distribution of standard deviations of samples drawn from a normal population.
2. A curve has been found representing the frequency distribution of values of the means of such samples,

when these values are measured from the mean of the population in terms of the standard deviation of the sample.

3. It has been shown that this curve represents the facts fairly well even when the distribution of the population is not strictly normal.
4. Tables are given by which it can be judged whether a series of experiments, however short, have given a result which conforms to any required standard of accuracy or whether it is necessary to continue the investigation.

Gosset did not write any more work on a small-samples distribution. This is usually explained by the fact that his main job was the brewer's work. Although he himself denies this reason when he explains to Fisher (August 14, 1924): "By the way it is not time but ability which has prevented me following up my work by more on your line." But (Fisher, 1939) can see other causes: "Probably he felt that, had the problem really been so important as it had once seemed, the leading authorities in English statistics (i.e. Karl Pearson) would at least have given the encouragement of recommending the use of his method; and better still, would have sought again similar advantages in more complex problems. Five years, however, passed without the writers in *Biometrika*, the journal in which he had published, showing any significance of his work. This weighty apathy must greatly have chilled his enthusiasm."

Due to the general lack of interest in statistics applied to small-sample observations, Gosset's approach was not used for many years except for the research in brewery. Ultimately, it was expanded only thanks to Fisher's considerable promotional effort.

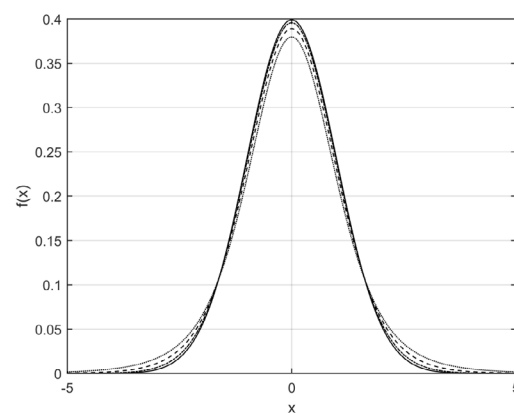


Figure 2 Comparison of probability density graphs. By solid line is depicted density of standardised normal distribution, dotted line Student distribution with 5 degrees of freedom, dashed line Student distribution with 10 degrees of freedom, dash-dotted line Student distribution with 30 degrees of freedom.

6 Conclusions

W. S. Gosset, who published under the pseudonym “Student”, was an important, though at the time not properly appreciated personality of the first third of the 20th century. He laid out the basics of statistical analysis of experiments with a relatively small number of observations. His work had a bearing on disciplines such as statistical production management, planning and evaluation of field and laboratory experiments, social and economic statistics. His thoughts fundamentally influenced the work of other personages of his time, especially Fisher. It is obvious that the problems addressed by Gosset would have been later solved by somebody else; but it would have happened with a considerable time span, which would probably have affected the level of contemporary knowledge.

7 References

- Boland, P. J., 1984: A Biographical Glimpse of William Sealy Gosset. *The American Statistician*, 38(3): 179–183.
- Boland, P. J., 2011: William Sealy Gosset – An Inspiring ‘Student’. Proc. 58th World Statistical Congress (Session STS028), Int. Statistical Inst., Dublin.
- Box, J. F., 1987: Guinness, Gosset, Fisher, and Small Samples. *Statistical Science*, 29(1): 45–52. <https://doi.org/10.1214/ss/1177013437>
- Cunliffe, S., 1976: Interaction. *Journal of the Royal Statistical Society, Ser. A*, 139(1): 1–19. <https://doi.org/10.2307/2344381>
- Dennison, S. R., MacDonagh, O., 1998: Guinness 1886–1939: From Incorporation to the Second World War. Cork University Press.
- Eisenhart, C., 1979: On the Transition from Student’s z to Student’s t. *The American Statistician*, 33: 6–10. <https://doi.org/10.2307/2683058>
- Fisher, R.A., 1939: ‘Student’. *Annals of Eugenics*, 9, 1–9. <https://doi.org/10.1111/j.1469-1809.1939.tb02192.x>
- Laird, N. M., 1989: A conversation with F. N. David. *Statistical Science*, 4(4): 235–246. <https://doi.org/10.1214/ss/1177012487>
- Lehmann, E. L., 1999: “Student” and Small-Sample Theory. *Statistical Science*. 14(4): 418–426. <https://doi.org/10.1214/ss/1009212520>
- McMullen, L., 1970: Student as a Man. *Studies in the History of Statistics and Probability, Volume I*, Charles Griffin and Co., p. 355–360.
- Pearson, E. S., 1990: ‘Student’ – A Statistical Biography of William Sealy Gosset (R. L. Plackett and G. A. Barnard, Eds.). Clarendon Press, Oxford.
- Student, 1908: The Probable Error of a Mean. *Biometrika*, 6(1): 1–25.
- Student, 1917: Tables for estimating the probability that the mean of a unique sample of observations lies between $-\infty$ and any given distance of the mean of the population from which the sample is drawn. *Biometrika*, 11: 414–417.