Great Lease, Arthur Guinness—Lovely Day for a Gosset!

By Stephen T. Ziliak

Stephen T. Ziliak is Professor of Economics, Roosevelt University, Chicago, Illinois. His e-mail address is <sziliak@roosevelt.edu>.

Prepared for *Journal of Wine Economics* (“Beeronomics” issue)

September 2009

Great Lease, Arthur Guinness—Lovely Day for a Gosset!

By Stephen T. Ziliak

Abstract: Small sample theory—the great innovation in statistical method in the period after Galton and Pearson—was ironically discovered by a brewer during routine work performed at a large brewery, Arthur Guinness, Son & Company, Ltd. For four decades William S. Gosset applied small sample experiments to the palpable end of improving, however gradually, the production and control of a consistent unpasteurized beer when packaged and sold at efficient economies of scale. Introducing, "Guinnessometrics." Annual output of stout at Guinness's Brewery may have topped 100 million gallons but Gosset's scientific knowledge was built one barleycorn at a time; in fact, the inventor of small sample theory worked closely with botanists and breeders. In the process, the brewer, William Sealy Gosset (1876-1937) aka "Student," an Oxford-trained chemist—though self-trained in statistics—solved a problem in the classical theory of errors which had eluded statisticians from Laplace to Pearson. In addition, though few have noticed, Gosset's exacting theory of errors, both random and real, marked a significant advance over ambiguous reports of plant life and fermentation asserted by chemists from Priestley and Lavoisier down to Pasteur and Johannsen, working at the Carlsberg Laboratory. Central to the Guinness brewer's success was his persistent economic interpretation of uncertainty, what Ziliak and McCloskey (2008) call the "size matters/how much" question of any series of experiments. An enlightened change in Guinness human resources policy gave an incentive structure that also seems to have nudged "Student," who rose in position to Head Brewer, to find a profit when the opportunity knocked. Beginning in 1893, Guinness vested "scientific brewers" such as Gosset with managerial authority. In fact Gosset was at times involved with price negotiations over hundreds of tons of barley and hops—perhaps hours or minutes before he ran (that is, calculated) a regression on related material. In brewing circles William Gosset is remembered less nowadays than he might be. He did not give two cents for arbitrary rules about statistical significance—at the 5% level or any level arbitrarily assumed. How the odds should be set depends on the importance of the issues at stake and the cost of getting new material, he said from 1904. Yet even in brewing journals, both academic and trade, and for the past 85 years, statistical significance at the 5% level continues to draw its arbitrary line segregating a meaningful from a non-meaningful result, a better barley from a worse.

Keywords: Guinness, "Student," small sample theory, economic approach to the logic of uncertainty, quantitative chemistry, Lavoisier, W. S. Gosset, scientific brewing, *t*-test.

Great Lease, Arthur Guinness—Lovely Day for a Gosset![1]

By Stephen T. Ziliak

Professor of Economics
Roosevelt University

Prelude: Accounting for "Arthur's Day"

In 2009 the Guinness Brewery (a Diageo company) celebrated its 250th birthday. Annual sales of stout may be flat of late but at "two billion pints" (Yenne 2007, p. 205) a year there is no need to blow out the candles. In fact in Dublin, Ireland, on September 24th, 2009 at precisely 17:59 local time, the mood was high as a year-long series of brewery-sponsored events, lectures, and celebrations came to a glorious head, on "Arthur's Day." This was a spectacular global phenomenon, with commerce and beer at its center. In fact, numerous economists joined brewers and other contestants, ladies and gentlemen, tossing back a few jars, increasing the sample size, and possibly cutting a rug out on the street to honor St. Guinness and one fateful if not liquid day in 1759.

The hopped-up salute was not to the 250 years since Adam Smith signed the *Theory of Moral Sentiments* (1759), however worthy in sober truth, but rather—and approximately—to "250 years since Arthur Guinness signed the 9,000 year lease on the St James's Gate Brewery in Dublin" (Reuters 2009).[*] Read that again: 250 years ago, Arthur Guinness signed a "9,000 year lease" for the "brewery, house property and land" (Lynch and Vaizey 1960, p. 70) at St. James's Gate at a rent a rational brewer was glad to pay: "£45 a year" (Lynch and Vaizey, p. 70; Yenne, p. 13)!

Arthur's impressive lease is on permanent display at The Guinness Storehouse Museum for mortgage bankers, lawyers, and other economic tourists to enjoy. It would seem regardless our eminent evangelical preachers of alcohol prohibition, such as the great Yale economist, Irving Fisher, are wrong. The Prohibitionist, such as Fisher (1926), is naturally troubled by the case of the sad and penniless drunk; but Fisher is nevertheless wrong about beer's role for individuals in the Good Society. Believe it or not, the other black export from Dublin, James Joyce's *Ulysses*, though frequently banned by U.S. churches and school boards, is right: beer, careful counting, and the bottom line *can* live happily ever after. As Joyce (1914, p. 79) put it in *Ulysses*: "Lord Ivaegh [aka "Edward Cecil Guinness," a Chief Officer of Guinness[2] who was in Joyce's time Mayor ("Ivaegh") of Dublin] once cashed a sevenfigure cheque for a million in the bank of Ireland. Shows you the money," Joyce said, "to be made out of porter."

---

[*] The historian and manager of Guinness Archives, Eibhlin Roche, explains that September 24th "is a date in the recent years, which the company has chosen to celebrate the life of Arthur Guinness . . . we have no evidence Arthur was brewing at St. James's Gate in September of that year" (Roche 2009).

Reuters stated in advance that Arthur's Day "will begin in the Irish capital and . . . carry on around the world in Lagos, Kuala Lumpur and New York"—as indeed it did.[3] But in the media, the full flavor of the day—the glamour, the pride, the virtue and excitement—seems understated. For instance, Sean Paul, a dancehall star from Jamaica, celebrated Arthur's Day by headlining a concert in Lagos, Nigeria—Nigeria being Guinness's third-largest market after the UK and Ireland (Yenne, p. 206)—while in Kuala Lumpur, Malaysia, another large market, The Black Eyed Peas were pounding out the jams along with the hogsheads of black stout.

Parts of the celebration sounded a more serious note. For example, the same family which bequeathed in 1908 Ivaegh Gardens to Dublin (Toibin 2009) and "Student's" *t* test to science (Student 1908a) gave a century later, in 2009, £6 million to establish the "Arthur Guinness Fund." According to the company website, the Fund will be used "to empower individuals with skills and opportunities to deliver a measured benefit to their communities."[4] "Measured benefit," the company said.

William's Day: Fermenting Small Sample Statistics

At Guinness clever uses of counting—"measured benefit"—did not end with Leopold Bloom or Arthur's lease. Far from it.

Arthur's Day, 2009 was quite properly focused on Arthur Guinness. But to a statistical scientist or scientific brewer, the golden moment—the thing you want to dance to —is what one might call "William's Day," by which is meant the 40-year long subsidy, invention and profitable application of "'Student's' test of significance" "(Fisher 1939, p. 1) and a half dozen other techniques of modern analysis and control, such as small sample "design of experiments" (Ziliak 2008, p. 212), control charts with "tolerance limits" (Shewhart 1939 [1986], p. 58) and "artificial randomization" for purposes of simulating the behavior of beer and beer inputs (Student 1908a, pp. 13-19; Pearson 1990, pp. 90-91).

The great innovation in statistics in the era after Galton and Pearson was made in the private sector of the economy, between 1904 and 1937, at Guinness's Laboratory, to the end of improving, however gradually, production of a consistent beer at efficient economies of scale.

Beginning in 1893 the Guinness brewery, inspired to act on a vision articulated first if not best by that same Lord Ivaegh of *Ulysses*, hired as a matter of policy exclusively first- class Oxford- and Cambridge-trained scientists to help put the brewery on a scientific basis (Dennison and McDonagh 1998, chp. 6; cf., in Denmark, The Carlsberg Laboratory, established in 1876 [Holter, H. and K. MaxMoller, eds. 1976.]). An unintended consequence of Guinness policy—of, in effect, Lord Iveagh's and others' push for scientific brewing—was a big revolution in scientific method.

In a firm pulsing with energetic brainiacs, William Sealy Gosset (1876-1937), an experimental brewer, was the battery-operated Energizer. From 1899 to 1937, inside the thick stone walls of St. James's Gate—but also out in the field, at his friend Edwin S. Beaven's malthouse and experimental barley farm, in Warminster, for instance; and at John H. Bennett's Irish malthouse and barley farms, scattered in and around Ballinacurra; and of course most famously speaking, at Karl Pearson's Biometric Laboratory, University College London—this little bunny—Gosset—an avid outdoor gardener and

Beethoven enthusiast—unknowingly caused a change in statistical and experimental theory, and thus scientific brewing.

Tragically, something happened in the history of statistics which caused neglect and finally disposal of the brewer's chief insights (Ziliak and McCloskey 2008). Ronald A. Fisher (1890-1962), the great mathematical statistician and eugenicist, and the brewer's younger friend (Gosset 1962), made invaluable contributions and additions to Gosset's small sample statistics (Fisher 1925a, 1925b, 1925c, 1935, etc.: see References). But he was a Prohibitionist of a different sort and in any case after Fisher's rise to fame the brewer's name faded fast and hard and with it the economic approach to the logic of uncertainty, small samples and large, has all but gone missing (Ziliak 2008).

In my book with Deirdre N. McCloskey, *The Cult of Statistical Significance* (2008), we show that 8 or 9 of every 10 articles published in the life and human sciences are misusing the techniques which Gosset pioneered at Guinness. The top-ranked journals of economics, management and biology are no exception. Not even today's leading scientific brewers seem to be aware of the history and problems with statistical significance, though originally discovered by William Gosset, a scientific brewer; one detects in today's brewery a rather too sharp line being drawn between science and business, between "technical excellence" (Bamforth 2002, p. 5) on the one hand and "the bottom line set of accounts" (p. 30; cf. Bamforth, p. 133) on the other.

A first-hand study of Gosset at Guinness suggests that today's brewers, economists, and others would do well to stop thinking of observations and experiments as one thinks of pure mathematics—the legacy of Ronald Fisher. However large or small the scale of output if a beer lacking statistical control is predictably bland a statistic lacking economic control is certainly blind. This is the first assumption of "Guinnessometrics"—the original, pre-Fisherian strain of small-sample econometrics.

> My own war work [Gosset told Karl Pearson during the First World War] is obviously to brew Guinness stout in such a way as to waste as little labor and material as possible, and I am hoping to help to do something fairly creditable in that way."[5]

A Self-Trained Statistician

It seems he did. You don't have to tell a brewer that running a series of experiments on malt characteristics or of yield trials on barley can grow expensive, fast (Student 1911, 1923, 1926, 1938, 1942). Out in the field, for example, each experimental acre, farmer, and scientist employed has an opportunity cost measured by the value of alternative employments of scarce resources (such as being back home at the Brewery, in the Laboratory, analyzing data). Thus, beginning in 1904, a well-paid apprentice brewer deliberately pioneered small sample statistical analysis. Working with small samples was an economic, not a mathematical, decision. By 1907 (Ziliak 2008, p. 205) the brewer calculated and filled in a probability table, now called "Student's' *t*," with which he and any scientist could consult to find numerical limits to the precision of results found in random samples of barley or malt or what have you as small as N= 4 (Student 1908a, p. 19). "Student's" *t* was a major theoretical and practical breakthrough, however strange to the-then dominant Pearson school, which continued to labor on with large sample biometrics. Said Savage (1954, p. 116):

> Though observations are sometimes free, there is typically a cost associated with making them; information must typically be bought either from other people or, more often from nature, so to speak. The cost of information may be money, trouble, one's own life, that of another, or any of innumerable possibilities, but all can in principle be measured in terms of utility. The cost of an observation in utility may be negative as well as zero or positive; witness the cook that tastes the broth.

Or witness the chemist that tastes the beer, subject to a profit and loss function, the pioneer of Savage's and others' small sample-economic approach to the uncertainty of observations. William Sealy Gosset, an Oxford-trained chemist, is better known by his pen name, "Student," as in "'Student's' test of statistical significance" (E.S. Pearson 1937a; Fisher 1939, p. 1). Though self-taught in statistics, the ingenious brewer was able to solve a problem in the classical theory of errors which had eluded Laplace, Gauss, and others (save Helmert) in the history of statistics.

> "Student's" treatment of the problem of uncertainty of the mean of a set of observations derived from the normal law [wrote the great physicist Sir Harold Jeffreys] has the further merit of being accepted by all schools
>
> Jeffreys 1961, p. 378

> [I]t 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
>
> Fisher 1939, p. 9; italics in original.

"Student's" exacting approach to discovering the sources of errors, both random and real, in his and in others' experiments on fermentation and plant inputs, marked a significant advance over the approaches taken by previous students, such as for example Priestley, Pasteur, and even the great economist and chemist, Lavoisier (Schabas 2009).

Said the great experimental maltster and barley farmer, Edwin S. Beaven (1857-1941 [Beaven 1947]), in an October 19, 1937 *Times* obituary notice ("Mr. W.S. Gosset: The Interpretation of Statistics"):

> My friend of 30 years, William Sealy Gosset, who died suddenly from attack on Saturday, at the age of 61, was known to statisticians and economists all over the world by his pseudonym "Student," under which he was a frequent contributor to many journals. He was one of a new school of mathematicians who were founders of theories now generally accepted for the interpretation of industrial and other statistics.
>
> E.S. Beaven 1937, Guinness Archives, GDB/BR01/0964

Along with Beaven and Harold Jeffreys (1961), one of "Student's" closest "students" was Egon S. Pearson (1895-1980), a great statistician, historian, and industrial research worker who did far more for industry, science and statistics than is evident from mathematical papers he co-authored with Jerzy Neyman (Pearson 1990). When "Student" died Pearson told readers of *Nature*:

> Under Galton, Weldon and Pearson, the Biometric School had been mainly concerned with the handling of comparatively large samples from biological populations, but Gosset in his daily work was forced to attempt to draw conclusions, *leading to executive action*, from the analysis of relatively small numbers of observations.

E. S. Pearson 1937a, p. 838; italics added; Guinness Archives, GDB/BR01/0964

In the issue of the *Times* which ran Beaven's notice came a second of three separate notices authored by the grieving Pearson (Pearson's father, Karl, had died within the same year). Said Pearson again of the Guinness brewer:

> It is no exaggeration to say that the bulk of the pioneer work which has eventually brought these subjects within the range of statistical inquiry was Gosset's . . . The growing confidence which his firm placed in the use of statistical technique in research provided more valuable evidence than could be given by any amount of theoretical discussion on the point which it was his object to establish – namely, that useful conclusions, attainable in no other way, may be drawn from comparatively small series of observations , provided that care is taken in planning how the observations are collected and that appropriate methods of analysis are applied to the results.

E. S. Pearson 1937b, Guinness Archives, GDB/BR01/0964

The humble brewer's methods—much to his surprise—turned out to be general, of use to every statistical scientist. He (rather incorrectly) told Fisher in a letter of September 21, 1922:

> I am sending you a copy of Student's Tables as you are the only man that's ever likely to use them!"

W.S. Gosset to R.A. Fisher, in Gosset 1962, Letter no. 11, p. 21

Gosset, Fermentation, and the 18th Century Quantitative Revolution in Chemistry

Fermentation studies have been at the center of quantitative progress in science. Though Jeffreys, Fisher, Beaven, E.S. Pearson and others have not put it this way, "Student" finished a job begun in the 18th century by other quantitative chemists of fermentation, such as Priestley and Lavoisier (though the latter was assisted by Laplace) (Holmes 1985, p. 491, p. 160, *passim*).

Nearly 250 years ago—around the time that Arthur signed his lease—Joseph Priestley's discovery of carbon dioxide ("fixed air") occurred while Priestley was studying gases emitted from experimental beer he was brewing. A few years later, in 1788, Lavoisier's then earth-shattering discovery and most famous statement of "the principle of conservation of mass" came while Lavoisier was describing oenological experiments conducted by him, fermenting grapes into wine (Holmes 1985, p. 385).

Lavoisier—who was an economist, civil servant and social reformer as well as a philosopher and chemist (Schabas 2006, p. 51)—revolutionized the field of chemistry and put it on a firmer quantitative foundation. What's notable to an economist among Lavoisier's many contributions are his "balance sheet" (Holmes, p. 394-5) and "balanced equation," now familiar to students of the conservation principle. The Lavoisier balance sheets tell the mass of organic matter before and after transformation. "As the final step in making fermentation data ready for the *Elementary Treatise,*" Frederic Holmes, the eminent historian of science and leading Lavoisier student explains, "Lavoisier compiled a new set of tables showing the detailed balances of all the elements contained in each substance present before and after the operation, and 'recapitulation' in which he grouped the individual quantities by element rather than by substance." "These tables," remarks Holmes, "have been widely admired" (Holmes, p. 393).

In Lavoisier's own words:

> One can see that in order to reach a solution . . . it is necessary first to know well the analysis and the nature of the substances able to undergo fermentation; for nothing is created, either in the operations of art, or in those of nature, and one can state as a principle that in every operation there is an equal quantity of material before and after the operation; that the quality and the quantity of the principles are the same, and that there are nothing but changes, modifications.
>
> It is on this principle that the whole art of making experiments on chemistry is founded.
>
> Lavoisier 1788, "Elementary Treatise," quoted in Holmes 1985, p. 394, n19.

But full acceptance of Lavoisier's experiments on small-sample fermentation—his "tables"—was, as with his other experiments, limited by a lack of statistical control for degrees of uncertainty, random or real. Says Holmes in *Lavoisier and the Chemistry of Life* (1985, xviii): "assessments miss the real problems Lavoisier faced. . . The difficulties he encountered were those not of weighing accurately, *but of judging what was significant to weigh*; *of knowing how to estimate the quantities of substances he could not directly weigh;* of determining what adjustments he could legitimately make when his balance sheets did not work out perfectly, as they almost never did; *of deciding what degree of error was reasonable in a given type of experiment, when he lacked formal methods for computing expected errors*" (italics added).

"The most impressive aspect of his achievements," argues Holmes, "is not that he sometimes came out with figures so close to our own for the composition of this or that substance, but that he so consistently reached results that were meaningful. . . . [He worked not] *by routinely applying global methodological principles to every situation, but by analyzing each individual problem, and making complex judgments about the assumptions necessary to interpret each set of experimental measurements*" (Holmes, xviii-xix).

Lavoisier's "description of alcoholic fermentation signifies a major advance toward rigor in his handling of the processes involving plant and animal matter" (Holmes

1985, p. 396) but as Holmes notes, Lavoisier did not have a handle on the nature or degree of errors in his experiments; importantly, as Holmes also notes, neither did Lavoisier take a "global methodological" approach "to every situation." A prelude to Gosset.

Lavoisier obtained assistance on the design and interpretation of experiments from his friend, Laplace. But as Fisher (1939, p. 1) notes, even Laplace failed to see or deal with the problems "Student" perceived a century later, working as an experimental brewer at Guinness.

Likewise Pasteur's *Studies on Fermentation* (1879) made seminal contributions to experimental science. "As early as the 1870s, Pasteur was awarded American patents for his methods of manufacturing and preserving beer and wine" (Geison 1995, p. 266). But how much did chance, mis-measurement, inconsistent procedures, large standard errors and the like affect Pasteur's conclusions? For a sobering answer, see Geison 1995, pp. 237-47 and Pasteur 1879, pp. 31-32.

Until the early 1900s no chemist or student of fermentation—certainly no member of "the cult of Pasteur" (Geison, p. 264)—had conquered error analysis as Gosset was about to. For instance, although several of Horace T. Brown's contributions to quantitative chemistry and scientific brewing were original and commercially useful (Stewart 2009), the great brewer from England, like the great brewer from France, eschewed the classical theory of errors. The venerated Horace Brown established the first chemistry laboratory at Guinness, as described by Brown in two thick volumes, *Transactions of the Guinness Research Laboratory* (1903, Arthur Guinness, Son, & Co. Ltd., Guinness Archives). In *Transactions* one finds hundreds of pages thick with quantitative chemical description—instructions on what you'd expect such as, for example, how to determine nitrogen content in barley. As Stewart explains, Brown's work on modification and flocculation may have been seminal but Brown did know how to use the classical theory of errors to adjust, if necessary, his favored hypotheses. For example in 1903, in small sample experiments on the "correlation of 'mealiness' and density in barley" (Brown, p. 99) the man who was hired to advise Guinness on standards did not take any actions to account for random or real errors in procedures (Brown 1903, pp. 99-100). Until "Student," few chemists did.

"Mistakes are the portals of discovery," a famous Dubliner allegedly said. Let's admit in advance that "Student" made a lot of mistakes, starting with a (small) miscalculation of the probable error of a mean, in 1908 (Box, 1978, pp. 70-1). Yet "Student" acknowledged his mistakes; he thanked others for pointing them out, when they did, and normally he tried to help his interlocutor in return for what "Student" considered a kind gesture, fixing errors. That was his business. He always thought of mistakes, his errors and those of others', as portals of discovery.

"Student": A Porter Brewer's Tale

William Gosset was born June 13, 1876 in Canterbury, England. He was cradled in the gentry and educated at Winchester School and Oxford University, New College.[6] The great unknown of statistical science worked his entire adult life—1899 to 1937—as an experimental brewer for one employer, Arthur Guinness, Son & Company, Ltd., Dublin, St. James's Gate. Gosset was a master brewer and rose in fact to the top of the top of the brewing industry: Head Brewer of Guinness. By all accounts it was a happy and comfortable life for the Canterbury lad who was widely admired and loved. A friend from childhood recalled that Gosset possessed "an immovable foundation of niceness" (L. McMullen 1939, p. 208). Friends and colleagues such as Karl Pearson, Ronald Fisher, Egon S. Pearson, Florence David, Udny Yule, and others whole-heartedly agreed. Said Yule, "He is a very pleasant chap. Not at all the autocrat of the *t* table" (quoted in M. Kendall, 1952, p. 159).

Gosset was in 1899 a energetic—if slightly loony—23 year-old gentleman scientist.[7] He possessed a wickedly fertile imagination and more energy and focus than a St. Bernard in a snowstorm. An obsessive observer, counter, cyclist, and cricket nut, the self-styled brewer had a sizzle for invention, experiment, and the great outdoors. Guinness was giving its brewery a radical make-over, and the nice chemist was one of the men it brought in to help.

The look-touch-and-sniff approach of 18th century craft brewers had satisfied the Guinness bottom line since its founding in 1759. No longer. The extent of Guinness's market was in a sense already large—in fact, largest in the world. By 1914, annual production at Gosset's brewery would surpass 2 million hogsheads or 4.35 billion pints.[8] But until the 1890s, the extent of the market was limited by the guildsmen's division of labor (Dennison and MacDonagh, p. 23, 38; O'Grada, pp. 304-5).

The Guinness future was in "scientific brewing"—large-scale, industrially controlled brewing—wherein all factors of production, from barley breeding to taste testing, are controlled, improved, and confirmed by experimental science. A degree from either Oxford or Cambridge in a natural science was a minimum requirement for a Guinness brewer in the new era (Dennison and MacDonagh, 1998, ch. 6), whereas family ties had secured employment in the past. Danish and German breweries were transforming similarly—perhaps even a little ahead of Guinness's imposing pace. For discussion of the early history of Guinness, and for industry background, see Lynch and Vaizey (1960), Mathias (1959), Holter and MaxMoller, eds. (1976), and Bamforth (2002).

But Guinness took brewing science to another level – the economic level. Edward Cecil Guinness (1847-1927) was the extraordinary capitalist here. Himself a master brewer and one-time sheriff of Dublin, the first Lord Iveagh went on to become a legendary entrepreneur, philanthropist, chief executive officer, and long-time Chairman of the Board of Guinness (Dennison and MacDonagh, pp. 5-15). His great brewery was unique in a number of ways, and especially—for the future development of statistics—by vesting selectively chosen brewer-scientists, such as Gosset, with economic authority. With scientific brewers in managerial position, the theory was, experiments could shine a light on the bottom line and the bottom line a light on the experiments.[9]

Guinness recruited other scientific minds besides Horace T. Brown, who was after all a short-term consultant on chemical analysis, not a salaried brewer, the position of distinction. Over the years, scientific brewers and the statisticians serving them, such as A. McMullen, F. Watson, T. B. Case, G. F. Story (aka "Sophister"), E. S. Somerfield (aka "Mathetes"), the statistician Stella Cunliff—who rose to president of the Royal Statistical Society—and a Head Brewer, Geoffrey S. Phillpotts—Gosset's brother-in-law—did a lot to improve statistical Guinness (Dennison and MacDonagh, pp. 80-82).

But no one simulated the golden harp as much—or with as much salience, applicability, and profit—as did Gosset. On August 4, 1899, C. D. La Touche, then managing director, recorded in a note that "Mr Gosset" graduated Winchester as "Scholar of New College [Oxford University], [earning a] 1st Class in Mathematical Moderations, Trinity Term 1897, and 1st Class in Chemistry, July 1899. He is short-sighted and wears spectacles," added La Touche. "Seems generally speaking suitable."[10]

From 1899 to 1906, Gosset was Apprentice Brewer, mostly in the "Experimental Brewery," a miniature brewery close to the Main.[11] In 1904 he began to tackle the problem of making an inference from small samples of malt and hops, two of the crucial inputs to the beer. With Board endorsement, he spent in 1906-1907 a sabbatical year at Karl Pearson's Biometrics Laboratory, University College London, where his general logic of small samples—and his groundbreaking article on "The Probable Error of a Mean"—fermented (Student, 1908a; E. S. Pearson 1990, pp. 47-48). Gosset married in 1906 a national field hockey player and coach, Marjory Surtees Phillpotts, with whom he sired three children.

He signed his published articles and notes with a pseudonym, which was standard business practice. Search his published works for references to beer and odds are you'll fail. Why he chose to be known as "Student" is not known for sure but here is a clue: besides being a humble man there is stamped on the cover of his early notebook a manufacturer's imprint, "The *Student's* Science Notebook, Eason and Son, Ltd., Dublin and Belfast." [12]

In 1907 Gosset returned to Dublin as Head Experimental Brewer, a position he held through calm and turbulent times until 1935. (He volunteered to fight in the First World War but, like Fisher, his application was denied by reason of short-sightedness.) In the early 1920s Gosset became Head of the Statistics Department he established at Guinness. His first task: to estimate the effect of their first-ever advertising campaign on beer sales in Scotland. In 1935 Gosset was promoted to Head Brewer in charge of the new plant at London, Park Royal (now closed), and in September, 1937, he was appointed Head Brewer of all Guinness.[13] Pounding out up to 100 million gallons of Guinness annually, Gosset introduced the quantitative side of scientific brewing, and with it, a storehouse of statistical and experimental theory and tools.

On Guinness's clock, the anonymous "Student" invented or inspired solution concepts which today represent about a dozen different research programs in statistics, econometrics, agronomy, decision science, brewing, and industrial quality control.[14]

Gosset died of a heart attack, his third in close succession, in Beaconsfield, England, October 16, 1937.

The Economic Origins of "Student's" *t*

Porter is a name given to a dark and bitter beer with a good head. In the nineteenth and early twentieth century, porter was alternatively called "stout" or "stout porter." Stout is the name used now to describe a beer such as Guinness that is bitter on the up-take (bitterness being a function of both the quantity and quality of hops added per barrel of malt) and yet smooth, mellow, and slightly smoky on the finish. Black-ruby tint arises like the smoky finish from roasted barley or "malt"—the distinguishing ingredient of stout.

Brewing "experimentally" introduced challenges and trade-offs. For example, in Gosset's time, Guinness stout was a completely natural and unpasteurized beer. In keg, cask, or bottle, the life of a natural beer is numbered in days.[15] Yet Guinness's beer was shipped worldwide, on an increasingly large scale. Hop is a natural and effective preservative, but it is bitter, and brings bacteria and pests. A heavily hopped beer, such as "Foreign Extra Stout," would last longer than the light hopped, but it would continue to condition on the ocean voyage, becoming increasingly bitter. This was a trade-off to be estimated. In 1911, Guinness carried out an international taste test by shipping bottles of Foreign Extra Stout to agents located in Rio de Janiero, Auckland, San Francisco and other cities. The object was to "ascertain whether any distinctions could be drawn as regards behaviour, . . . acidity in bottle, or flavour" (Case 1911, p. 284).

Another challenge was that Guinness, a wholesale dealer only, pursued an unusual pricing strategy. Nominal product price the company tried to hold constant and many years it did (measured in Dublin prices).[16] At the same time, between 1887 and 1914, output more than *doubled*. Plant size expanded, too, and with it the capital/output ratio. So the crucial question facing scientific brewers was: How can experimental science advance economies of scale in brewing? How can inferential statistics help to control and improve produce and at the same time reduce average total costs for the firm?

Take the choice of hops, for example. In 1898, Guinness used 4.72 million imperial pounds of the fruitful yellow cone. The traditional method of choosing hops based on looks or fragrance wasn't efficient or reliable on this large scale. But was it any more reliable to take small samples out of a larger quantity of hops, test them for certain key characteristics, and then draw an inference about the general quality of the whole lot?

In 1898, Thomas B. Case, Guinness's first scientific brewer, led a team to address this question. Case and his team felt that the key characteristic was the degree of soft resins in the hops. Thus, Case analyzed the average percentage of soft and hard resins found in samples of 50 grams taken from several seasons of American and Kent hops (Case, 1898, p. 47). He compared his samples with those of a cooperating scientist named Briant. For example, Case examined 11 samples of hops (n= 11) from Kent in 1897, finding on average 8.1 percent soft resins content. Briant examined 14 samples drawn from the same lot, finding 8.4 percent soft resins—a difference of 0.3 percentage points. The mean difference between their two samples of "American, 1895" was even higher, at 0.7 percent (soft resins) and 1.0 percent (hard).

How should these results be interpreted? Case wrote: "We could not . . . support the conclusion that there are no differences between pockets of the same lot." He worried about "defects" in his procedure, especially the "difficulty of sampling." The more fundamental difficulty, however, was a lack of knowledge about inferential statistics, period. Lacking a theory of how to draw inferences from small samples, Case had no basis for evaluating whether observed differences represented random error from the samples or actual differences in the population.

Other experiments were conducted on barley yield, malt extract, kiln-drying. In 1899 Gosset was recruited by Case, and by 1904 the brewer was determined to solve—he did not realize at the time—a classical problem in the theory of errors. Economic considerations tended to leave the brewers with small samples, Gosset noticed, time and again. In barley yield experiments, for example, it was common to test new varieties and treatments on plots which were 2-4 acres in size. Experimental plots could be further subdivided, a practice which Gosset encouraged and labeled "the principle of maximum contiguity" (Student , 1923 [1942], p. 95). The constraint on samples available for analysis is any case understandable given the opportunity cost of the commercial farmers who where commissioned by Guinness to run the experiments, not to mention the cost of the Guinness scientists' own time.[17] Mixing new beers and malts took time, too, placing a limit on the size of those samples. As Gosset (1962, Letter no. 1, Sept. 15, 1915, emphasis added) wrote to Ronald Fisher in 1915: "Experiments naturally required a solution of the mean/S.D. problem and the Experimental Brewery which concerns such things as the connection between the analysis of malt or hops, and the behavior of the beer, *and which takes a day to each unit of the experiment, thus limiting the numbers.*"

In early November, 1904, Gosset discussed his first break-through in an internal report entitled "The Application of the 'Law of Error' to the Work of the Brewery" (Gosset, 1904 *Lab. R.,* Nov. 3, 1904, p. 3). Gosset (p. 3) wrote:

> Results are only valuable when the amount by which they probably differ from the truth is so small as to be insignificant for the purposes of the experiment. What the odds should be depends—
> 1. On the degree of accuracy which the nature of the experiment allows, and
> 2. On the importance of the issues at stake.

Two features of Gosset's report are especially worth highlighting here. First, he suggested that judgments about "significant" differences were not a purely probabilistic exercise: they depend on the "importance of the issues at stake." Second, Gosset underscored a positive correlation in the normal distribution curve between "the square root of the number of observations" and the level of statistical significance. Other things equal, he wrote, "the greater the number of observations of which means are taken [the larger the sample size], the smaller the [probable or standard] error" (p. 5). "And the curve which represents their frequency of error," he illustrated, "becomes taller and narrower" (p. 7).

Since its discovery in the early nineteenth century, tables of the normal probability curve had been created for large samples; Stigler (1986, 1999) offers a useful early history of the normal distribution. The relation between sample size and "significance" was rarely explored. For example, while looking at biometric samples with up to thousands of observations, Karl Pearson declared that a result departing by more than three standard deviations is "definitely significant."[18] Yet Gosset, a self-trained statistician, found that at such large samples, nearly everything is *statistically* "significant"—though not, in Gosset's terms, economically or scientifically "important." Regardless, Gosset didn't have the luxury of large samples. One of his earliest experiments employed a sample size of 2 (Gosset 1904, p. 7) and in fact in "The Probable Error of a Mean" he calculated a *t* statistic for n=2 (1908a, p. 23).

Gosset's analysis focused on malt extract, which was measured in "degrees saccharine" per barrel of 168 lbs. malt.[19] At the time, an extract in the neighborhood of 133° gave the targeted level of alcohol content for Guinness's beer. A higher extract affected the life of the beer, and also the alcohol content—which in turn affected the excise tax paid on alcoholic beverages. In Gosset's view, ± .5 was a difference or error in malt extract level which Guinness and its customers could swallow. "It might be maintained," he said, "that malt extract "should be [estimated] within .5 of the true result with a probability of 10 to 1" (p. 7). Using mean differences of extract values with samples drawn from the Main and Experimental breweries, he then calculated the odds of observing the stipulated accuracy for small and large numbers of observations (p. 7). He found at:

	Odds in favour of smaller error than .5 [are]
2 observations	4:1
3 "	7:1
4 "	12:1
5 "	19:1
82 "	practically infinite

Thus, Gosset (p. 8) concluded, "In order to get the accuracy we require [that is, 10 to 1 odds with .5 accuracy], we must, therefore, take the mean of [at least] four determinations."

The report, showing how to achieve quality control of output through small sample estimation, was instantly hailed by the Board. (See, for a modern treatment of a similar problem, Zellner 2005, pp. 18, 19, and the very illuminating Table 1 on p. 23). But Gosset himself wasn't convinced: "We have been met with the difficulty," he cautioned, "that none of our books mentions the odds, which are conveniently accepted as being sufficient to establish any conclusion." He said (p. 12), "It might be of assistance to us to consult some mathematical physicist on the matter." Perhaps, Gosset thought, there is a conventional rule about how to set the level of significance, and a mathematical physicist might be able to tell him. Board Endorsement No. 62, signed March 9, 1905, explains: "Mr. Case will make arrangements for Mr. Gosset to have an interview with Prof. Karl Pearson."[20]

Professor Pearson, who was the leading figure in the era before Fisher, was willing to meet Mr. Gosset at Pearson's summer home, in July 1905. His only request was that Gosset write him a letter in advance, detailing his question (for a discussion of Pearson's life, Porter, 2005, is a fine contemporary starting point). In the letter, Gosset actually answers the very question he poses. Gosset's letter of 1905 to Karl Pearson (quoted in E. S. Pearson 1939, pp. 215-216; first italics in original) invents an economic approach to the logic of uncertainty:

> *My original question and its modified form.* When I first reported on the subject [in the 1904 internal memo], I thought that perhaps there might be some degree of probability which is conventionally treated as sufficient in such work as ours and I advised that some outside authority in mathematics should be consulted as to what certainty is required to aim at in large scale work. However it would appear that in such work as ours *the degree of certainty to be aimed at* must depend on the *pecuniary advantage to be gained by following the result of the experiment, compared with the*

> *increased cost of the new method, if any, and the cost of each experiment.*
> This is one of the points on which I should like advice.

Notice that in Gosset's view, setting the "degree of probability" to be "treated as sufficient" is not to be made "conventionally" or by "some outside authority in mathematics." The 5% rule—the convention in every science--implies that no rational decision can be made if the odds are not 19 to 1 (.95/.05) or better that the observed result, however important or not, is there by chance. Gosset, a profit center at Guinness, though still an Apprentice, thought better. Instead of a fixed rule, the young Bernoulli said, the "degree of certainty to be aimed at" depends on the opportunity cost of following a result as if true, added to the opportunity cost of conducting the experiment itself. So in Gosset's view odds of 5 or 6 to 1 (.85/.15) might suffice for claiming a commercially or biometrically speaking important "finding": it depends on the size of the prize and the cost of pursuing it, every time. Gosset never deviated from this central position.[21]

The July 1905 meeting took place. Pearson apparently did not answer the "modified" question, but Gosset spent the next year working on his logic of small samples while on sabbatical at Pearson's laboratory.

As early as 1904 Gosset began to use regression analysis, a technique he learned by reading Merriman's *Method of Least Squares* (1884) and G. B. Airy's *The Theory of Errors of Observation* (1861).[22] In 1908 he used regression to revisit Case's 1898 "hops input" versus "life of beer" question. In a fantastic analytical leap, Gosset (1908, p. 145)—assisted only by slide rule and a mechanical calculator—estimated parabolas of the form

$$L = A + BH^2,$$

where L = life [of beer] in days
A = life in days of no-hopped beer depending on conditions
H = lbs. of hops
B = a constant [slope parameter] depending on the hops and on the conditions.

"No-hop brewings" (A) could survive between "12.2 and 16.7 days," he found after numerous repetitions of the experiment under same and different conditions, whereas "hopped" (B) could live for a month or beyond.

To Gosset's persistently economic approach, the effect sizes he estimated with "life" regressions were only the beginning. A few months later he noted the "advantage to replace a vague character like increase in life [the dependent variable] . . . . by a single definite value which can be directly converted into £ *s. d.*"—pounds, shillings, and pence (Gosset, 1909 *Lab. R.*, p. 211). From partial correlations which he had calculated on the percentage of soft resins and "[brewing] value," he argued, "We can find an equation giving the probable value for any given percentage of soft resins. The equation is V = 2.82 + 10.78S, where V is the per cent Value compared with the 'standard' hop, and S is the soft resins measure from 9 per cent … . [E]each one per cent of soft resins makes a difference of 10.78 per cent. in the value of the hops. This," he said, "at the average [1909] price of hops, represents about 8*s*. per [hundredweight]." At a 1909 input of 6.79 million pounds of hops, Gosset discovered a big economic difference.

"The probable error of the prediction is large," Gosset however cautioned, "being about 6.6 per cent." But the noisy resins variable did not stop Gosset from making a judgment about resins' *economic* importance to brewing value. "Of this [probable error] some 3.2 per cent. is due to errors of analysis and sampling," he said, "leaving [a residual experimental error] due to brewing errors and other factors not included in the analysis" (p. 212). But with the new if imperfect and noisy method of making inferences from small samples Gosset was able to reject about one-third of the "standard" hops that unscientific methods had previously commended (pp. 212-13). Again the Board cheered.

What Gosset did later for the bottom line in barley is hopped up by several orders of magnitude. In variety trials designed and co-conducted with E.S. Beaven (1857-1941), Gosset biometrically proved the commercial value of three varieties that would eventually be grown on "well over five million acres of [English] land" (Beaven, 1947, p. xiv). The Gosset-enhanced barleys were used to brew Guinness but also to make breakfast cereals and feed livestock.

Precision matters, as in soft resins and malt extract. But high statistical significance ranked low in "Student's" ordering. "Student" was looking for practical guidelines on how to improve and maintain beer quality and production without raising the costs. He would not have dreamed of stopping an experiment on grounds that a result reached or failed to reach an arbitrary level of statistical significance.

Additional examples of Gosset's disregard for the 5 percent rule are easy to cite. A few more shall suffice. In 1927, when Fisher was actively promoting the 5 percent rule, Gosset considered "p=.13" a "fairly good fit" (Student, 1927 [1942], p. 147).

On May 18, 1929, Gosset wrote a letter to Egon S. Pearson (1895-1980), repeating to the son the same message that had been delivered to the father (as reprinted in Pearson 1939, p. 244; italics added).

> I fancy you give me credit for being a more systematic cove than I really am ["Student" said] in the matter of limits of significance. What would actually happen would be that I should make out P_t (normal) and say to myself "that would be about 50:1; pretty good but as it may not be normal we'd best not be too certain" . . . and whether one would be content with that or would require further work would *depend on the importance of the conclusion and the difficulty of obtaining suitable experience.*

In 1937 Gosset again wrote to Egon (who was by then the editor of *Biometrika* and also the chair of University College of London's brilliant Statistics Department), this time using no uncertain terms (as quoted in Pearson 1939, p. 244; emphasis in original):

> Obviously the important thing . . . is to have a low real error, not to have a [statistically] "significant" result at a particular station. The latter seems to me to be nearly valueless in itself. . . . Experiments at a single station [that is, tests of statistical significance on a single set of data] *are* almost valueless. . . . You want to be able to say not only "We have significant evidence that if farmers in general do this they will make money by it", but also "we have found it so in nineteen cases out of twenty and we are finding out why it

> doesn't work in the twentieth." To do that you have to be as sure as possible which is the 20th—your real error must be small.

In short, "Student" saw statistical significance at any level as being "nearly valueless" in itself.[23]

Fisher's 5 Percent Rule

"Student" was Ronald A. Fisher's older friend and behind-the-scenes mentor. Fisher was in 1912 a student at Cambridge when he inquired in a letter about a mistake in Gosset's "The Probable Error of a Mean" (Box, 1978, pp. 70-1). He proved that Gosset's test statistic should be divided by what he would later call "degrees of freedom" (n – 1) not by total sample size (n) (Student, 1925, pp. 105-6). Gosset was grateful, and he asked Karl Pearson to publish the young man's correction in *Biometrika*. Although Gosset and Fisher would not meet in person until 1922, a friendship developed in prolific correspondence, with over 150 scientific letters surviving, mostly from Gosset to Fisher (Gosset, 1962).

Fisher developed his own philosophy of "Student's" methods while claiming to teach the original. For example, in *Statistical Methods* Fisher (1925a [1941], p. 42, italics supplied) wrote: "The value for which P=.05, or 1 in 20, is 1.96 or nearly 2; *it is convenient to take this point as a limit in judging* whether a deviation is to be considered significant or not. Deviations exceeding twice the standard deviation [said Fisher] *are thus formally regarded as significant."*

Fisher was a rhetorical magician: notice how his 5 percent rule evolved in consecutive sentences from a rule of convenience to a formal regard. Look again at Fisher's "formal" .05-rule and recall Gosset's life regression, which had a probable error of .066.[24] Following Fisher's rule at Guinness would have brought reduced profit and quality to the firm and customers.

Fisher took a hard line on the 5 percent rule. For example, he (1926b, p. 504, italics supplied) wrote:

> It is convenient to draw the line at about the level at which we can say: "*Either there is something in the treatment, or a coincidence has occurred* such as does not occur more than once in twenty trials." . . . Personally, the writer prefers to set a low standard of significance at the 5 per cent point, and *ignore entirely all results which fail to reach this level.*

In 1935 Fisher (1935a [1960], p. 13, italics supplied) declared in his other classic book of methods, *The Design of Experiments*:

> *It is usual and convenient for experimenters to take 5 per cent. as a standard* level of significance, in the sense that they are prepared to *ignore all results which fail to reach this standard*, and, by this means, *to eliminate from further discussion* the greater part of the fluctuations which chance causes have introduced into their experimental results.

But as Tim Harford, "The Undercover Economist," wrote in *Financial Times* (2009):

> Fisher proposed ignoring any finding that failed to reach the 95 per cent confidence level. In other words, until the odds against a pattern having emerged by chance are 19 to 1 against, disregard the pattern completely. . . That might seem a reasonable precaution – and it is certainly standard practice today – but a sharp line for statistical significance makes no sense, and it has a cost.

Fisher took from the brewer only what he needed. He (1935a, in Savage 1971a, pp. 471-2, italics supplied) readily acknowledged that testing a null hypothesis did not necessarily provide much information about underlying true values. Said he:

> *A null hypothesis may, indeed, contain arbitrary elements,* and in more complicated cases often does so: as, for example, if it should assert that the death-rates of two groups of animals are equal, without specifying what these death-rates actually are. *In such cases it is evidently the equality rather than any particular values of the death-rates that the experiment is designed to test, and possibly to disprove.*

"Student" did not admire Fisher's null hypothesis test procedure. "Student" focused on estimation, prior probability, alternative hypotheses, power, questions of "how much" (for example, Student 1926 [1942], p. 126). "Student" himself was a philosophical Bayesian, Jeffreys (1961, pp. 379-81) noticed, and took in his laboratory work a Bayesian and decision-theoretic approach (for example, Gosset 1909 *Lab. R.*, pp. 205-6).

Fisher tended to take the opposite approach (Zabell, 1989). He sought retreat from the consequences of being incorrect in judgment, and indeed from measuring or evaluating practical consequences in any currency at all. Said Fisher (1955a, p. 75, italics supplied):

> *Finally, in inductive inference we introduce no cost functions for faulty judgments . . . In fact, scientific research is not geared to maximize the profits* of any particular organization, but is rather an attempt to improve public knowledge *undertaken as an act of faith . . . We make no attempt to evaluate these consequences, and do not assume that they are capable of evaluation in any currency.*

Contrast again "Student," working at the then-largest for-profit brewery in the world: "the degree of certainty to be aimed at must depend on the *pecuniary advantage to be gained by following the result of the experiment, compared with the increased cost of the new method, if any, and the cost of each experiment.*"

Reasons for widespread acceptance of Fisher's 5 percent philosophy are too complex to disentangle here. Ziliak and McCloskey (2008, chp. 20-23) make an attempt to do so.[25] But the fact that Fisher's way has been widely accepted is obvious in tables reporting "significant" results in economics and other sciences.

Statistics, Brewing, and the Economic Approach Today

By the 1950s Fisher was campaigning heavily to discourage statisticians from taking economic perspectives, those "evaluations," and he won more converts than even he realized. "In the U. S. also the great importance of organized technology has I think made it easy to confuse the process appropriate for drawing correct conclusions [by 'correct' he means conforming with his 5% philosophy], with those aimed rather at, let us say, speeding production, or saving money" (Fisher 1955, p. 70). In life, the Guinness employee who died in '37 was by inclination and duty unable to go into full rhetorical battle against his friend, the imperious academic mathematician (Gosset 1936; cf. Student 1938). But as Bruno de Finetti wrote in an important if neglected symposium,"The economic approach seems (if not rejected owing to aristocratic or puritanic taboos) the only device apt to distinguish neatly what is or is not contradictory in the logic of uncertainty" (de Finetti 1971, pp. 486-7). Gosset, a profit-seeker, agreed. Fisher's "faith" in "Natural Knowledge" is yeasty, even as metaphysics; his null hypothesis testing procedure, minus power and a loss function, and his uniform "5% rule of statistical significance" without a Bayesian prior, may have made sense to some in the early part of the 20th century when professional associations for many disciplines were finding their quantitative orientations (Ziliak and McCloskey 2008, chp. 24). But Fisherian statistics stripped of economics—today's convention—is small beer compared to alternatives such as Guinnessometrics (see also: Zellner 2005, 1984; Press 2005; Jeffreys 1961).

Scientific brewers today might wish to swallow a little more Gosset than they do now. For example, Charles Bamforth's excellent *Standards of Brewing* (2002) is a trifle quiet on topics concerning errors of experiments and the difference between statistical and brewing significance. His chapter three, "Statistics and Process Control," is designed to introduce brewers and managers to small sample significance testing and sound interpretation. But Bamforth doesn't make the economic points that Gosset articulated to brewers so long ago. Instead, he mocks the money makers, such as CEOs, and glorifies the stance of the non-pecuniary scientist—the softer side of Fisher (pp. 5, 137, 138). Unlike Gosset or Lavoisier, Bamforth sounds frustrated by money words: "I was from the world where the prime currency was scientific intellect," he writes, "the *raison d'etre* was technical excellence, the beauty was in the consistency of the beer and not the bottom line of a set of accounts" (Bamforth, p. 5). To Gosset, aesthetic and economic choices go together and must be faced head on. For example, Gosset was unimpressed with "validity" as it is conventionally defined in Fisher's purely mathematical system. Said Gosset (Student 1938, p. 206):

> I personally choose the method which is most likely to be profitable when designing the experiment rather than use Prof. Fisher's system of a posteriori choice* which has always seemed to me to savour rather too much of "heads I win, tails you lose".[26]

(Incidentally, the key parameters measured and reported in Bamforth's statistical tables of chemical analyses of barley (chp. 5, p. 62), malt (p. 82), hops (pp. 96-7), and beer ("QC Checks on Beer," p. 144), were, though "state of the art" in Bamforth 2002, originated and statistically updated annually or semi-annually by Gosset and others at the Guinness brewery, aided by an early version of Student's *t*-test (e.g., Gosset 1908, 1909).)

But the main problem in Bamforth's book is that by sneering at economics and profit he fails to warn readers about the "significance" mistake, Fisher's non-interpretative-interpretation: for example, Bamforth misses a chance on page 133 in his discussion of taste testing -"difference testing" (p. 133), as he misleadingly calls it. As a result, Bamforth fails to make the economic points about "real error" and "profit" and "power" and "loss" which Gosset, again, had made to Fisher, the two Pearsons and others statisticians, come to think of it not so long ago. (To better see that Gosset's world is not so far away from our own—that Gosset's approach versus Fisher's is not "just" "history"—consider that Gosset's friend and adopted student, Egon S. Pearson, was the Chair of the UCL Statistics Department immediately before the job was taken over by Dennis Lindley. Professor Lindley is, like Pearson and Gosset before him, an eminent Neo-Bayesian statistician and in addition a recent contributor to this *Journal*.)

In Graham G. Stewart's 2009 "Horace Brown Medal Lecture: Forty Years of Brewing Research," published in the *Journal of the Institute of Brewing*, the distinguished scientist and brewer of 40 years, Dr. Stewart, reports statistics on "beer foam collapse" (Stewart 2009, p. 21) and other uncertain organic phenomena. Unlike Bamforth, whose book is in fact thoroughly quantitative and mostly up-to-date, Stewart doesn't report on variance or dispersion—at all! His article is filled with authoritative sounding reports: but the tables and figures and offers of interpretation are presented without any indication of the form of error, sample variance, design of experiment, or real measurement error, such as mis-counts. His article, to celebrate his and Horace T. Brown's distinguished service to the profession is in fact innocent of the classical theory of errors, circa 1898, at Guinness, the Case level, pre-Gosset. For example, Stewart's "Table IX" reports on "Vitality of brewing yeast strains after a 96 h fermentation of synthetic media"* in six separate "strains" of experimental beer (Stewart, p. 21):

[Strain]	Glucose	Maltose
Ale 1	0.8	1.3
Ale 2	0.9	1.3
Ale 3	1.1	1.4
Lager 1	0.7	0.9
Lager 2	0.8	1.2
Lager 3	0.9	1.0

*Peptone – yeast extract – 4% sugar medium. Acidification power test" (Stewart 2009, Table IX, p. 21).

But in Stewart's "Horace Brown Medal Lecture" (2009), the reader can't judge from the figures and surrounding text which if any of Stewart's pair-wise differences of maltose and glucose content are substantively or even statistically significantly different (N =6). For the six different beers, three ales and three lagers, the eminent brewer has not provided sufficient information to prove the existence of either "Gosset"- or "Fisher"-significance, neither brewing- nor statistical significance. He didn't establish a null hypothesis or minimum difference that would be pay the brewer or challenge the chemist. He didn't calculate a standard deviation. Yet here is how Stewart interprets his Table:

"For all six strains studied, cells cultured in maltose consistently had higher viabilities and enhanced vitalities compared to their glucose cultured counterparts. Reasons for these differences are not immediately apparent" (Graham G. Stewart 2009, pp. 21-22). Perhaps a small dose of "Student" would help to make things apparent. Similar failure to control for and report the brewing and economic meaning of differences (for example, p. 66, p. 758, p. 763) and of the non-random sources of error obstructs full acceptance of Dr. Stewart's otherwise brilliant *Handbook of Brewing* (2006, with F. G. Priest).

Likewise in the "new" histories of experimental chemistry, debates about statistical inference have not risen to the level of Gosset versus Fisher, not even close. In Steven Shapin's and Simon Schaffer's *Leviathan and the Air Pump*, for example, the authors claim to ground their discussion of Robert Boyle's experimentalism with N= 2 experiments by Boyle, yet giving no indication of the size and properties of real or random errors. Of Shapin's and Schaffer's conclusions one historian remarked: "That [n = 2 with no discussion of error terms] is a slender foundation on which to raise the superstructure that they erect to depict the strategies Boyle used to "compel assent" to the "matters of fact" that his experiments yielded" (Holmes 2004, p. 150). Clearly even historians find a need for economic application of the theory of errors.

In the final analysis, the lesson of the Battle of Guinness's Lab versus Fisher's is probably this: Don't export your bottom line judgments to the stream of consciousness poet, lunatic, or vague mathematician. You would have to be a penniless drunk—or a noble person temporarily seduced or controlled by one or more false gods—to stumble so deliberately, yes? To discover the price of hops that will beat or cover hop cost at the margin, science, statistics and economics are three of the firm's necessary inputs. Odds are The Story of Guinnessometrics is not known as much as it might be as some of the scientific personnel employed at today's leading breweries are still unsure about it.

Acknowledgments

A version of this paper was presented May 29, 2009 at the conference, "Beeronomics: The Economics of Beer and Brewing," LICOS/Katholieke Universiteit-Leuven, Belgium. I would like to thank members of the Beeronomics Organizing Committee and other conference participants for their many interesting remarks, beers, and t-statistics. (By the way, to those who couldn't travel: the brewery tours, at n =3, were perfectly selected!) For critical and/or encouraging comments related to the history of this paper I gratefully acknowledge Julian Alston, Tim Harford, Oliver Hawkins (at BBC Radio 4, "More or Less"), Deirdre McCloskey, Jacques Kibambe Ngoie, John Nye, Damiaan Persyn, Jo Swinnen, Bert Van Landeghem, Arnold Zellner, and Jim Ziliak. At the Guinness Archives (Diageo, St. James's Gate) it is my pleasure to thank archivists' Eibhlin Roche and Deirdre Flood for their impeccable good taste, efficiency, and kind and knowledgeable assistance. At the Special Collections Library, University College London, I owe a debt of gratitude to Gillian Furlong, Mandy Wise, and Steven Wright for patient kindnesses extended by each of them since 2004. I thank Guinness (Diageo) and the Special Collections Library at UCL for permission to cite selections of copyrighted material kept in the Archives. Finally, I owe "a pint of plain" to Prof. John Mullahy, for pointing my attention to a colorful flying toucan. Any errors are my own.

References

Primary Sources
(detailed citations in text)

Gosset 1962: *Letters of William Sealy Gosset to R. A. Fisher. Vols. 1-5,* Eckhart Library, University of Chicago. Private circulation.

Gosset Collection, Guinness Archives, Dublin (Diageo).

Guinness Archives, Dublin (Diageo)

Especially the fifteen volumes of *Laboratory Reports,* 1898-1912 (abbreviated as *Lab. R.*), and:

Brown, Horace T., ed. 1903. *Transactions of the Guinness Research Laboratory, Vol. I, Part I.* Dublin: Arthur Guinness, Son and Co., Ltd.

Case, Thomas B. 1898. "Hops, Season 1897," *Laboratory Report,* Vol. 1, Aug. 31, Arthur Guinness & Son, Ltd., pp. 47–57.

Case, Thomas B. 1911. "Comparison of Intermediate and Ordinary F. E. S. [Foreign Extra Stout] Under Trade Conditions." *Laboratory Report*,Vol. 12, Dec. 13th, Arthur Guinness Son, Ltd., Board Endorsement No. 2868, pp. 284–307.

Gosset, William S. [Student] 1904. "The Application of the 'Law of Error' to the Work of the Brewery." *Laboratory Report*, Vol. 8, Nov. 3, Arthur Guinness & Son, Ltd., pp. 3–16 and unnumbered appendix.

Gosset, William S. 1908. "The Present Position of Our Knowledge of the Connection between Life and Hops in the Experimental Brewery." *Laboratory Report,* Vol. 10, Oct. 13, Arthur Guinness & Son, Ltd., pp. 137–150.

Gosset, William S. 1909. "The Brewing of the Experimental Hop Farm Hops, 1907 Crop (PartII), Together with a Note on the Present Method of Hop Analysis." Laboratory Report, Vol. 10, Jan. 19, Arthur Guinness & Son, Ltd., pp. 202–220.

Pearson Papers (Karl Pearson, Egon Pearson, Gosset, Fisher, Neyman files), University College London (UCL), Special Collections.

Secondary Sources

Beaven, Edwin S. 1947. *Barley: Fifty Years of Observation and Experiment* (London: Duckworth). With a statistical appendix by Geoffrey S. Phillpotts, who was Head Brewer of Guinness, 1937-1947, and a close friend and brother-in-law to W. S. Gosset.

Berger, James O. 2003. "Could Fisher, Jeffreys and Neyman Have Agreed on Testing?" *Statistical Science* 18(1): 1-32.

Box, Joan Fisher. 1978. *R. A. Fisher: The Life of a Scientist.* New York: Wiley & Sons.

De Finetti, Bruno. 1971 [1976]. "Comments on Savage's "On Rereading R. A. Fisher." *Annals of Statistics* 4(3): 486-7.

Dennison, S. R., and Oliver MacDonagh. 1998. *Guinness 1886-1939*. Cork: Cork University Press.

Efron, Bradley. 1998. "R. A. Fisher in the 21st Century." *Statistical Science* 13(2): 95-122.

Fisher, Irving. 1926. *Prohibition at its Worst* (New York: Macmillan).

Fisher, R. A. 1922a. "On the Mathematical Foundations of Theoretical Statistics." *Philos. Trans. Roy. Soc. London Ser. A* 222: 309-368.

Fisher, R. A. 1922b. "On the Interpretation of χ^2 from Contingency Tables, and the Calculation of P," *Journal of the Royal Statistical Society* 85 (Jan.): 87-94.

Fisher, R. A. 1925a [1941]. *Statistical Methods for Research Workers*. New York: G. E. Stechart and Co. Originally published Edinburgh: Oliver and Boyd. Eighth of thirteen editions, in at least seven languages.

Fisher, R. A. 1925b. "Applications of 'Student's' distribution." *Metron* V(3, Dec.): 90-104.

Fisher, R. A. 1925c. "Expansion of 'Student's' Integral in Powers of n^{-1}." *Metron* V(3, Dec.): 110-120.

Fisher, R. A. 1926. "Arrangement of Field Experiments." *Journal of Ministry of Agriculture* XXXIII: 503-13.

Fisher, R. A. 1935. *The Design of Experiments.* Edinburgh: Oliver & Boyd. Reprinted in eight editions and at least four languages.

Fisher, R. A. 1939. "'Student'." *Annals of Eugenics* 9: 1-9.

Fisher, R. A. 1950. *Contributions to Mathematical Statistics*. New York: Wiley. Ed. Walter A. Shewhart.

Fisher, R. A. 1955. "Statistical Methods and Scientific Induction." *Journal of the Royal Statistical Society, Series B (Methodological)*, Vol. 17, No. 1, pp. 69-78.

Fisher, R. A. 1956 [1959]. *Statistical Methods and Scientific Inference.* New York: Hafner. Second edition.

Fisher, R. A., and Frank Yates. 1938 [1963]. *Statistical Tables for Biological, Agricultural and Medical Research.* Edinburgh: Oliver and Boyd. Sixth edition.

Geison, Gerald L. 1995. *The Private Science of Louis Pasteur* (Princeton: Princeton University Press).

Geweke, John. 2005. *Contemporary Bayesian Econometrics and Statistics.* New York: Wiley.

Gosset, William S. [see Student]. 1936. "Co-Operation in Large-Scale Experiments." *Supplement to the Journal of the Royal Statistical Society* 3(2): 115-36.

Hanley, James A., Marilyse Julien, and Erica E.M. Moodie. 2008. "Student's *z*, *t*, and *s*: What if Gosset had *R*? *The American Statistician* 62 (1, Feb.): 1-6.

Harford, Tim. 2009. "Enlightened Research, Fuelled by the Dark Stuff," *Financial Times*, Feb. 7, 2009. http://timharford.com/2009/02/enlightened-research-fuelled-by-the-dark-stuff/

Holmes, Frederic L. 1985. *Lavoisier and the Chemistry of Life: An Exploration of Scientific Community* (Madison: University of Wisconsin Press).

Holmes, Frederic L. 2004. *Investigative Pathways: Patterns and Stages in the Careers of Experimental Scientists* (London: Yale University Press).

Holter, H. and K. MaxMoller, eds. 1976. *The Carlsberg Laboratory, 1876-1976.* Copenhagen: Rhodos and the Carlsberg Foundation.

Hotelling, Harold. 1930. "British Statistics and Statisticians Today," *Journal of the American Statistical Association* 25 (170, June): 186-90.

Hotelling, Harold. 1931. "The Generalization of Student's Ratio." *Annals of Mathematical Statistics* 2: 360-78.

Hotelling, Harold. 1951. "The Impact of R. A. Fisher on Statistics," *Journal of the American Statistical Association* 46 (243, Mar.): 35-46.

Jeffreys, Harold. 1939 [1961]. *Theory of Probability*. Third revised edition. London: Oxford University Press.
Kendall, Maurice G. 1952. "George Udny Yule, 1871-1951." *J. Royal Stat. Soc.* 115A: 156-61.
Kendall, Maurice G. 1978. "The History of Statistical Method," pp. 1093-1101, in W. Kruskal and J. Tanur, eds., *International Encyclopedia of Statistics, Vol. 2.*
Kruskal, William H., and Judith M. Tanur. 1978. *International Encyclopedia of Statistics, Vols. 1-2.* New York: The Free Press.
Lynch, Patrick, and John Vaizey. 1960. *Guinness's Brewery in the Irish Economy, 1758-1876* (London: Cambridge University Press).
Macdonnell, William R. 1901. "On Criminal Anthropometry and the Identification of Criminals." *Biometrika*, Vol. 1, No. 2. (Jan., 1902): 177-227.
Matthias, Peter. 1959. *The Brewing Industry in England, 1700-1830*. London: Cambridge University Press.
McCloskey, Deirdre N., and Stephen T. Ziliak. 1996. "The Standard Error of Regressions." *Journal of Economic Literature* 34 (March 1996): 97-114.
McMullen, Launce. 1939. "'Student' as a Man." *Biometrika* 30(3/4, Jan.): 205-10.
Neyman, Jerzy, and E. S. Pearson. 1938. "Note on Some Points on 'Student's' Paper on 'Comparison Between Balanced and Random Arrangements of Field Plots.'" *Biometrika* 29(3/4, Feb.): 379-88.
Neyman, Jerzy. 1961. "Silver Jubilee of my Dispute with Fisher," *Journal of Operations Research* (Society of Japan) 3: 145-54.
O'Grada, Cormac. 1995. *Ireland: A New Economic History* (Oxford: Oxford University Press).
Pasteur. Louis. 1879. *Studies of Fermentation: The Diseases of Beer—Their Causes and the Means of Preventing Them.* London: Macmillan, trans. Frank Faulkner and D. Constable Robb.
Pearson, Egon S. 1937a. "Mr. W. S. Gosset: 'Student.' *Nature* 140 (Nov. 13): 838; Guinness Archives.
Pearson, Egon S. 1937b. [Obituary page] *London Times*, Oct. 19, 1937;Guinness Archives.
Pearson, Egon S. 1939. "'Student' as Statistician." *Biometrika* 30 (3/4, Jan.): 210-50.
Pearson, Egon S. 1966. "The Neyman-Pearson Story: 1926-34." Pp. 1-24 in F. N. David, ed., *Research Papers in Statistics: Festschrift for J. Neyman* (1966).
Pearson, Egon S. 1968. "Studies in the History of Probability and Statistics. XX: Some Early Correspondence Between W. S. Gosset, R. A. Fisher, and Karl Pearson, with Notes and Comments." *Biometrika 55* (3, Nov.): 445-57.
Pearson, Egon S. 1990 [posthumous]. *'Student': A Statistical Biography of William Sealy Gosset.* Oxford: Clarendon Press. Edited and augmented by R. L. Plackett, with the assistance of G. A. Barnard.
Pearson, Karl. 1914. *Tables for Statisticians and Biometricians, Part I* (London: Biometrics Laboratory, University College London).
Persyn, Damiaan. 2009. Personal communication to the author, email, Dec. 4, 2008, Roosevelt University, Chicago.
Porter, Theodore M. 2004. *Karl Pearson: The Scientific Life in a Statistical Age.* Princeton: Princeton University Press.

Press, S. James. 2003. *Subjective and Objective Bayesian Statistics.* New York: Wiley.

Priest, Fergus G. and Graham G. Stewart. 2006. *Handbook of Brewing*. London: CRC Press, Taylor & Francis.

Roche, Eibhlin. 2009. Personal communication, Guinness Archives (Diageo), Dublin, September 17, 2009.

Savage, Leonard J. 1971 [1976 posthumous]. "On Re-Reading R. A. Fisher." *Annals of Statistics* 4(3): 441-500.

Schabas, Margaret. 2006. *The Natural Origins of Economics*. Chicago: University of Chicago Press.

Shewhart, Walter A. 1939 [1986]. *Statistical Method from the Viewpoint of Quality Control* (New York: Dover). Edited with a new foreword by W. Edwards Deming.

Stewart, Graham G. 2009. "The Horace Brown Medal Lecture: Forty Years of Brewing Research," *Journal of the Institute of Brewing*115(1), 3–29, 2009.

Stigler, Stephen. 1986. *The History of Statistics.* Cambridge: Harvard University Press.

Stigler, Stephen. 1999. *Statistics on the Table.* Cambridge: Harvard University Press.

Student [W. S. Gosset]. 1907. "On the Error of Counting with a Haemacytometer." *Biometrika* 5(3, Feb.): 351-60.

Student. 1908a. "The Probable Error of a Mean." *Biometrika* VI (1, March): pp. 1-24.

Student. 1908b. "The Probable Error of a Correlation Coefficient." *Biometrika* (2/3, Oct.): 300-310.

Student. 1911. "Appendix to Mercer and Hall's paper on 'The Experimental Error of Field Trials,'" *Journal of Agricultural Science* IV: 128-131.

Student 1914. (See Pearson, Karl, 1914.)

Student. 1917. "Tables for Estimating the Probability that the Mean of a Unique Sample of Observations Lies Between - ∞ and Any Given Distance of the Mean of the Population from which the Sample is Drawn." *Biometrika* 11(4, May): 414-17.

Student. 1923. "On Testing Varieties of Cereals." *Biometrika* 15 (3/4, Dec.): 271-293.

Student. 1925. "New Tables for Testing the Significance of Observations." *Metron* V(3, Dec.): 105-108.

Student. 1926 [1942]. "Mathematics and Agronomy," *Journal of the American Society of Agronomy* 18; reprinted: Pp. 121-34 in E. S. Pearson and J. Wishart, eds., *Student's Collected Papers* (1942).

Student. 1927. "Errors of Routine Analysis." *Biometrika* 19(1/2, July): 151-64.

Student. 1931a. "The Lanarkshire Milk Experiment." *Biometrika* 23(3/4, Dec.): 398-406.

Student. 1931b. "On the '*z'* Test." *Biometrika* 23(3/4, Dec.): 407-8

Student. 1931c. "Yield Trials," *Bailliere's Encyclopedia of Scientific Agriculture,* pp. 1342-1360, reprinted in *"Student's" Collected Papers* (1942, pp. 150-168).

Student. 1938 [posthumous]. "Comparison between Balanced and Random Arrangements of Field Plots." *Biometrika* 29 (3/4, Feb.): 363-78.

Student. 1942 [posthumous]. *"Student's" Collected Papers* (London: Biometrika Office). Eds. E. S. Pearson and John Wishart.

Toibin, Colin. 2009. "Old, Heavy, Bitter, and Divine: An Irish Novelist on his World-Wide Quest for Perfect Pint of Guinness," *Wall Street Journal*, Sept. 12, 2009.

Yates, Frank. 1951. "The Influence of *Statistical Methods for Research Workers* on the Development of the Science of Statistics," *Journal of the American Statistical Association* 46**:** 19-34.

Yates, Frank, and Kenneth Mather. 1963. "Ronald Aylmer Fisher," *Biograph. Mem. Fell. R. Soc. Lond.* 9: 91–129.

Yenne, Bill. 2007. *Guinness: The 250-Year Quest for the Perfect Pint.* Hoboken: John Wiley & Sons

Zabell, Sandy. 1989. "R. A. Fisher on the History of Inverse Probability." *Statistical Science* 4(3): 247-63.

Zellner, Arnold. 1984. *Basic Issues in Econometrics*. Chicago: University of Chicago Press.

Zellner, Arnold. 2005. *Statistics, Econometrics, and Forecasting* (Cambridge: Cambridge University Press). The Stone Lectures in Economics.

Ziliak, Stephen T. 2008. "Guinnessometrics: The Economic Foundation of 'Student's' t," *Journal of Economic Perspectives* 22 (4, Fall): 199-216

Ziliak, Stephen T. 2008. "R. A. Fisher and the Copyright History of 'Student's' *t*," Roosevelt University, Department of Economics, http://faculty.roosevelt.edu/Ziliak

Ziliak, Stephen T., and Deirdre N. McCloskey. 2008. *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives* (Ann Arbor: University of Michigan Press).

---

[1] For my friends the oenologists and distillers, "Lovely Day for a Guinness!" is a popular Guinness advertisement and jingle which is known and sung by beer drinkers the world over.

[2] A time-line of Guinness executive staff members, 1886-1939, is found in Dennison and MacDonagh 1998, p. 271.

[3] Reuters, Sept. 8, 2009, "Guinness Announces Further Details for Arthur's Day Global Celebration," http://www.reuters.com/article/pressRelease/idUS186320+08-Sep-2009+PRN20090908

[4] Guinness.com website, Sept. 13, 2009: http://www2.guinness.com/en-gb/Pages/250-fund.aspx

[5] Letter of W. S. Gosset to Karl Pearson, Sept. 1 1915, quoted in E.S. Pearson 1990 (posthumous), p. 19.

[6] Guinness Archives, GDB/C004.06/0001.04 (File: "William Sealy Gosset, Memoranda and correspondence regarding Gosset's recruitment to Junior Brewer").

[7] "Loony" is a trait "Student" never outgrew: see, for example, how Student 1927 [1942], p. 145n, uses two kangaroos and a platypus to explain the meaning of kurtosis.

[8] Guinness Archives, GDB C004.06/0016 (File: "Comparative Statement and Summary of Financial Operations"); author calculations.

[9] Perhaps the industry should consider adopting some or all of the old Guinness approach, which is collecting money making and science into the same job: as Bamforth (p. 12) says, "The traditional approach is to have two separate bodies (bringing with them, regrettably, a degree of mutual distrust and distaste). The production guys make the stuff, and the quality control folk pull it apart with the aid of countless measurements." Of his own experience in a large-scale brewery in England, Bamforth laments: "When I was a quality assurance manager, I felt very much a second-class citizen to the head brewer" (p. 10).

[10] C. D. La Touche, Aug. 4, 1899, in GDB/C004.06/0001.04 (File: "William Sealy Gosset, Memoranda and correspondence regarding Gosset's recruitment to Junior Brewer").

[11] Guinness Archives, GDB/C004.06/0001.04 (File: "William Sealy Gosset, Recruitment papers").

[12] Folder 282, "Student's Haemocytometer Paper on Yeast-cell Counts," 1905-1907, Pearson Papers, UCL. Dennison and MacDonagh (1998, p. 90n9) suggest the "Student" name originates with the managing director La Touche but the Guinness historians show no evidence for it and apparently they were not aware of "Student's" eponymous notebook, located in London.

[13] Guinness Archives, GDB/C004.09/0004.14 (File: "Research Papers from Second Volume of the Brewery History—Brewers and Directors); Letter of W.S. Gosset to E. Somerfield, Oct. 1, 1937, GDB/BR01/0964 (File: "Correspondence").

[14] Harold Hotelling (1930, p. 189), a vice president of the American Statistical Association and a teacher of many leading economists and econometricians, wrote: "I have heard guesses in this country, identifying 'Student' with Egon S. Pearson and the Prince of Wales."

[15] Damiaan Persyn, a Ph.D. student in economics and a research-econometrician at VIVES and LICOS, Katholieke Universiteit-Leuven, Belgium, tells the story a little differently. At a famous Lambic brewery and even more famous Trappist brewery, both which Persyn and I toured together—I won't say which ones—brewers claim their unpasteurized beer is good for two decades or more (as Persyn pointed out, many claim the maximum limit allowable under EU policy). "However, I've read its taste described as "a complex aroma of leather, horse blanket, spice, and many other earthy components," Persyn said, "it seems to me that Gosset made the commercially correct judgement when calculating the optimal trade-off regarding the hop contents of Guinness: when buying Kriek at Mr. Hanssens place, you have to specifically state that it is beer you want to buy

and not vegetables or eggs, as his main income is from farming" (Persyn, personal communication, Dec. 4, 2008).

[16] From 1887 to the First World War, the average price of Guinness stayed between 2.6 and 2.8 British pounds per hogshead (one hogshead= 52 U.S. gallons): "Comparative Statement and Summary of Financial Operations," GDB C004.06/0016, Guinness Archives; author calculations.

[17] Student, 1931, p. 1342 [reprinted in Student 1942, p. 150]; Beaven, 1947, p. 164.

[18] *K. P. Lectures Volume I* [Gosset's Classroom Notebook], p. 13, 1906-7, Pearson Papers, Gosset file, UCL.

[19] The formula is: Malt extract = ([Specific gravity of the wort] – 1000) x 4.67. Page 2 of *Lab. R., Vol. VII*, No. 5, Oct. 25, 1906, "The Relationship Between Laboratory and Brewery Extracts, Introduction and Part I," by Alan Jackson (with the assistance of W. S. Gosset).

[20] Also see: "The Pearson Co-efficient of Correlation," *Lab. R.*, No. 2, Vol. III, 30th August, 1905, by W. S. Gosset, Guinness Archives.

[21] See, for example, Student 1923, p. 271, paragraph one, and Student 1931c, p. 1342, paragraph one, reprinted in Student 1942, p. 90 and p. 150.

[22] Gosset 1904, *Lab. R.,* op cit., p. 4n.

[23] Few realize that Neyman-Pearson "power" can also be traced to "Student," though Egon Pearson himself credited him. "Student" intuited the idea of power in two letters of May 1926 to Egon Pearson (as told in Pearson 1966, pp. 4-11; Pearson 1939, p. 242; Letter no. 1, p. 1, Green Box, Pearson Papers, Egon Collection, UCL).

[24] The probable error is .6745 times the standard error. Thus by today's measure the odds of observing Gosset's hops coefficients were less still—far below the 19 to 1 odds required by the rigid 5 percent rule. The odds were good enough for Guinness to justify a gamble on a new hops strategy.

[25] For example, consider the transfer in copyright holding on *t* itself, 1908 to 1938. To Fisher's 1922 request to 'quote' "Student's" revised table, "Student" replied: "As to 'quoting' the table in *Biometrika* it depends just what you mean by quoting. . . I don't think, if I were Editor, that I would allow much more than a reference!" Fisher finally did "quote" "Student's" complete tables, many times over. But ever since Fisher and Yates (1938), which became the major reference book, most researchers are not made aware of "Student's" earlier copyrights on the tables (Student 1908a, 1914, 1917, 1925). Fisher and Yates (1938, pp. 41-43) did not mention "Student's" copyrights and neither have volumes of textbooks ever since (Ziliak 2008; Ziliak and McCloskey 2008, chp. 22).

---

[26] Gosset's reference (*) is to Fisher's "*Statistical Methods for Research Workers* § 24.1 (5th ed.), p. 125" (Student 1942, p. 206, footnote). By "heads I win, tails you lose " Gosset was referring to Fisher's 5% rule of statistical significance. See Ziliak and McCloskey 2008 for the full argument.