Gosset At Guinness. J Hanley. SSC Annual Meeting, Winnipeg, June 13, 2017. james.hanley@mcgill.ca

Good day, and thank you to the organizers of this session. These are the topics I will cover. They are mostly different from those I covered in a 12 minute presentation at the Ottawa SSC meeting 9 years ago. 39 / 39

I gave a longer version of that at a special session in Dublin in 2008, on the 100th anniversary of Gosset's paper.

The outline then was as follows

has a link to my 'Student' webpage, our 2008 paper, as well as my presentations in 2008.

Today, I will follow this outline, based on material I have collected since then. But I will start again with a bit about Gosset and his paper.

So, who was this man who hid behind his non-de-plume or pen-name?

I will start at the back end, with the obituaries. The more personal one, well worth reading, was written by a Guinness colleague.

The second one was the 40-page appreciation by Egon Pearson, who organized the 2 appreciations in Biometrika.

R A Fisher had a 9-page appreciation in 'his own' journal, and this nice photo from around 1908.

This obituary in the London Times gives us the cause of his early death.

a few biographical details, and some insights into his life outside Guinness.

A few words about the spelling of his last name. Like Abraham de Moivre and Augustus De Morgan, he was of Huguenot origin, and the Gosset comes from the French, and has just 1 t – just as his statistical test is for just 1 sample.

Over the years, many writers got the spelling wrong. 9

You would have thought that at least Gosset's own employer would have it right. Here is their website in 2008, before IBC2008. Gosset was hired in 1899 as part of their move to Scientific Brewing, to stay ahead of their Danish competitor Carlsberg.

But they had had the spelling wrong – with 2 t's — on their website for several years. In 2008 I told them that if they claimed he was the father of modern statistics, they should at least get his name right.

They told me they were re-doing the main website. They did, and now leave his name off the main history. But in the new handout at the Storehouse, they do have it correct. Gosset with 1 t, just like the 1-sample. 41 / 390

Gosset spent a year's sabbatical with Karl Pearson, and the first paper from this involved counting yeast cells. It derives the Poisson distribution from scratch, without mentioning Poisson. It shows the first important application of that distribution. 37 / 427

Since our 2008 article and talks have already described the 1908 article, I am going to be brief about it. He starts out the same way we still do today,

but puts it very clearly and bluntly. What is a large n?

Instead of complaining, he has a plan.

Here are the 2 components, xbar and s-squared. Pearson insisted on n rather than n-1. But notice the ratio (he called it z) using the SD rather than the SE we use today.

His biggest problem was with the distribution of s-squared. His math-stat skills were not enough to get him past the 4th moment; it looked like s-squared followed a gamma distribution, but he couldn't be 100 percent sure. He though he had showed the independence of xbar and s, but in fact his proof was flawed. 55 / 564

Assuming he was right with both, then the rest is just a lot of calculus. He arrived at the pdf of z, then checked the smallest cases, and then made tables for n up to 10.

But here is the interesting and very modern section. He had done simulations even before he worked out the pdf, and used them again afterwards as a check on his results. Our 2008 paper describes how he simulated from what he was satisfied was a bivariate normal distribution. He could have used theoretical values, but instead, ever the applied type, he used the heights and finger lengths of criminals carefully measured by Scotland Yard police trainees.

Here are the grouped data on the 3000, with heights running from 4 feet 7 inches to 6 feet 5 inches. 21 / 697

Here are his steps, so clearly described that in 2007 we had no trouble replicating them (except we used R) 20 / 717

Our 2008 paper reports that his shuffling was just as good as our extensive shuffling via R. We had wondered whether his work was in the Guinness archives, but one of the speakers at the Dublin centenary was a very interesting Chicago economist named Stephen Ziliac. He told me that Gosset's notes, including the saved simulations, were in the Pearson archives at UCL. Apparently, when Gosset moved to London in 1935, Pearson was worried they might be lost, so he brought them to UCL. I had visited the Galton archives in 2005. I knew of the Pearson ones, and got to visit them the

year after the centenary, in 2009. Here are two notebooks belonging to Gosset.

This zooms in on the one for the cell-count data. More interesting than the title, is that the booksellers, Easons, called it 'The Student's Science Notebook.'

This is the one I was interested in. 8 / 868

Here is the inside. 4 / 872

Here is the first page, in blocks of 4. To simplify calculations, he refers to the values by the row and column they were in; the first number in each pair the finger, the second the height. 37 / 909

and this last one is page 13, so it checks out.. 12 x (10x6)

+ (5x6) = 750. 18 / 931

One interesting consequence of sampling from binned data is that in a few instances, all 4 persons were from the same 1 inch bin, so s=0 and z is infinite. What would you do today in such cases? He replaced the infinity by the largest finite z he encountered. 49 / 980

I don't have time to cover these other aspects, but you can read about them in our 2008 paper and slides, and on my website.

After the centenary session, the Irish Statistical Association gave each of the 4 speakers a plaque containing Gosset's return in the Census of Ireland in 1911. Library and Archives Canada had already digitized and put online some of the older Canadian censuses, and then helped the Irish Government to do the same for the Irish Censuses of 1901 and 1911. 60 / 1065

When we zoom in, you can see the name Gosset, his wife, 3 children aged 4, 2 and 1 month, his mother, and 3 Irish domestic servants (a maid. a cook, and a nurse). Gosset lists himself as a brewer.

Here is the return made by my grandfather, who lists himself, my grandmother, the uncle I am called after, aged 5, an uncle who died young, my father John, aged 2, my aunt aged 5 months, and my great-grandfather. Others remarked later that Gosset was not a great filler-in-of-forms, and it does look like my grandfather was more careful to read all the instructions first. 65 / 1170

So, how quickly did his distribution spread? Not very quickly, as Fisher's daughter Joan (who was married to George Box) tells us. It is not all that surprizing, given that it was just for 1 sample situations. It was Fisher who in 1925 showed that it was a special case of a much wider family where the numerator involves a Gaussian random variable, and the denominator the square root of an independent chi-squared random variable. 75 / 1245

The first use outside of Guinness was by this epidemiologist, 10 / 1255

who is also considered to have carried out the first 20th century case control study – on the risk factors for breast cancer.

In Ireland the word from doctors was that Guinness was good for nursing mothers, and I also knew of instances where some mothers added a little Guinness directly to the baby's bottle. One of the other ads was "My goodness, my Guinness." 42 / 1320

Lane-Claypon's 1912 report to the Government was on the value of boiled milk for infants.

She got the data from an infant consultation service in Berlin that saw about 100 babies a day (from working class families), and that visited the mothers at home. Mothers who could not breastfeed were supplied with milk from dairies operated by the city.

She formed 2 non-overlapping series, one fed on the breast, and one on boiled cows' milk [milk supplied from dairies operated by city] 23 / 1402

The 300 in the control series began being seen at various times up to 18 weeks, and were followed for various lengths of time. (The mothers got a small 'nursing-bonus each week) 32 / 1434

There were 204 in the boiled milk series series. They had to come for consultation every 7-10 days to renew their milk-card. Weighings were at the consultations.

Here are the average weighs for each series for each 8-day periods. 12 / 1473

and here they are on a graph. 7 / 1480

Now to the analysis, and to the 'question suggested by the curves'. Notice the causal phrase 'due to.'

This is an interesting way to approach things: even before considering other causal factors, try to rule out chance/sampling error.

She carried out a separate 2-independent-samples z-test for each of the 3 periods with at least 160 observations in each series. She did 2 things differently from today: she grouped the weights into 200 gram bins to save arithmetic, and (like Student) she used 'Probable' errors rather than 'Standard' errors of the means. The PE is about 2/3 of the SE. So it takes 3 PE's rather than 2 for 'statistical significance. All 3 ratios are well past that threshold. 80 / 1598

But, she is still not ready to jump to conclusions until she had ruled out the other factors: she devotes considerable space to these. 24 / 1622

I will skip to the beginning of the time-curves, where 10

/ 1632

the ratios are not so extreme.

and she is quite aware of this limitation.

But she is also aware of Student's Table that corrects for this.

Student's Table is only for 1-sample problems, so in her first calculation she takes the mean in the control group as 'mu zero' and uses the 9 d.f 1 sample test, obtaining p=0.02. But then she moves 'mu zero' up by 2 PE's and now the p-value is 0.14. 49 / 1707

So she concludes that the difference might be just sampling error. Of course, today we see a lot more complexities, both with the comparability of the series, and the structure of the data. It would be great if someone were to track down the raw data, and see if the conclusions would change. 53 / 1760

Except for the sleep example in his 1908 paper, Gosset's papers dealt with the statistics of agricultural experimentation. As Ziliak tells us, he got into quite an argument (albeit a civilized one) with Fisher about split plot designs and randomization. Ziliak and Stigler are continuing it today. But there is this one Biometrika paper of Gosset's dealing with human experimentation that every teacher and student should know about. It concerns a massive cluster-randomized and individually-randomized trial carried out in 67 primary schools in Scotland. 1/2 the students got 3/4

of a pint of milk each day, and the other 1/2 (like the fourth

little piggie in the child's story) got none. 111 / 1871

He strarts out, as good referees should, with high praise

but then here comes the 'but...' 6 / 1887

The most serious of his 4 criticisms concerns the allocation. The teachers put the poorly-nourished children in the milk groups so the control group had a big advantage to start with.

You can see it here, in every age-group. 8 / 1926

In fact this is how Gosset drew it (there were no figures in the 20 page report).

It was the same with the girls. 7 / 1950

Gosset had several suggestions for avoiding bias and reducing error in any future comparisons of milk versus control, and he suggested identical twins for comparing raw and pasteurized milk. Fisher weighted in on the raw vs. pasteurized question, but it was Karl Pearson and his colleague who said let's make the most of the already collected data. 57 / 2007

and that's what Ethel Elderton did. Her graphics are quite telling. And the story is a very modern one. Again, it would be great if someone could find these cards - it sounds like they were returned to the Department of Health for Scotland. 44 / 2051

In 2008, during the International Biometrics Conference in Dublin, the Irish Statistical Association unveiled a plaque in the Guinness Storehouse, Ireland's most popular tourist spot. 25 / 2076

Better still, Gosset's grandchildren came over from England to attend. His grandson (himself a chemist) told us he must be the only one at IBC who had met Gosset: he was born a few months before Gosset died in 1937, and his parents brought him into the hospital to visit his grandfather, who had had a number of heart attacks. The granddaughter was born after Gosset died.

The 2 of them unveiled the plaque.

On my website link to Peter Macdonald's photos, you can you can see all the other statisticians who lined up to have their photo taken. 25 / 2175

I already mentioned the special session, with 3 talks and one discussant. The 3rd talk was by this economist, Stephen Ziliak, who is quite a fan of Gosset. He doesn't like p-values, or Fisher's glorification of randomization. My Gosset website has a link to his site, and his books and other writings. He calls his site Ziliak's 'Library of Oomph & Precision.' 62 / 2237

Here the grand-daughter listening to the Discussion by 8

this man. He knew Fisher, and Egon Pearson Between them, Carl, Egon, and himself, were the editors of Biometrika for 90 years. 22 / 2274

Here is a close up of the plaque. Should the single quotes should be around Student or around the t?

First, only in school do data come from the back of the book. Real-data matter, so stay very close to them.

Gosset's rucksack model was vey heavy, and he was only allowed use it when Guinness was not doing annual reports etc. Fisher had a dedicated, but desktop model.

You guys have it easy. But you need to think harder and more imaginatively about simulations.

Writing , and communication more generally, are key. If asked, Yogi would have said statistics is 40Lastly, become students of the history of statistics. And, like me, have fun in your job. 16 / 2404