HISTORY OF STATISTICS HIGHLIGHTS AND MILESTONES

Penny S. Reynolds (University of Florida)

Beginning with this issue of CHANCE magazine, we will include a section called "History of Statistics Highlights and Milestones." Aficionados of the ASA History of Statistics Special Interest Group (HoS-SIG) who follow us on Twitter (*@HOS_ASA*) will recognize these as entries from *On This Day in Statistics*—events and personalities featured in our daily Twitter mini-blogs.

Not a fan of Twitter? A complete text of Twitter entries for the year will be archived as *CHANCE* supplementary files linked to the History Chronicles column. These supplementary materials given below include milestones from the first half of 2023. A condensed list of highlights can be found in CHANCE, Vol. 36, No. 3, 2023, pp. 39 – 43.

History of Statistics January 2023

01 Jan Happy Fibonacci Day 1123. The series was named after Leonardo of Pisa (aka Fibonacci, Leonardo Bonacci, Leonardo of Pisa, or Leonardo Bigollo Pisano) (c. 1170 – c. 1240–50) His book *Liber Abaci* (1202) introduced the sequence, as well as the Hindu–Arabic numeral system, to Western Europe. He illustrated the computation of the series by the famous problem of rabbit population growth, starting from a pair of rabbits.

However, Stigler's Law applies: the pattern first showed up in the Chandaḥśāstra by Pingala in 200 BC.

Also, OTD 1924 Genichi Taguchi b (d 2 Jun 2012) Japan. Best known for his statistically-based Robust Design in quality engineering methodology, Taguchi loss function, "off-line quality control", and signal:noise ratio. Won Deming Prize 1960 for contributions to the global quality movement. He considered quality control was achieved by eliminating process variances before they occurred so experimental design was more important than process.

02 Jan Jerzy Neyman introduced 'confidence intervals' in a 1934 presentation attacking non-random "purposive" sampling. The term was a translation of the original Polish 'przedzial ufnosci'. However, in the vote of thanks Arthur Bowley wondered out loud if confidence intervals were just "a confidence trick", doubted their validity and even if they were any use.

Refs
1934 *J Royal Stat Soc*, 97, 558-606
1937 Phil Trans Royal Soc A 236 (767), 333–380.
1941 *Biometrika* 32(2), 128-150

03 Jan OTD 1917 David Finney b (d 12 Nov 2018) UK. CBE FRS. ASA Fellow 1951, President Royal Statistical Society 1973. Worked with Fisher at University College London and Frank Yates at Rothamsted. Brest known for pioneering statistical methods of pharmacovigilance, biological assay and probit analysis. He iInaugurated 3 statistics departments at Aberdeen, BARC and Edinburgh.

After he missed a whole term at Cambridge after a bout with typhoid he switched from maths to statistics. Under John Wishart he did applied consultation work and "learned a lot about handling data by machinery" namely the Brunsviga calculating machine. He claimed if it wasn't for typhoid he would have "ended as a rather poor school teacher of mathematics"

Ref https://www.jameslindlibrary.org/finney-dj-1965/

o4 Jan On 15 Mar 1986 *BMJ* led the charge on endorsing confidence intervals as the reporting standard "whenever an inference is to be made from the study results to the wider world". This was about 50 years after confidence intervals were introduced by Jerzy Neyman. In the same issue, Martin J Gardner and Doug Altman provided clear and user-friendly guidelines.

Refs
Langman ML 1986 Br Med J 292(6522):716;
Gardner MJ, Altman DG 1986 Br Med J 292(6522):746-50.

05 Jan OTD Aryness Joy Wickens b 1901 (d 2 Feb 1991) USA. ASA Fellow 1936 and second woman elected ASA President 1952 Acting Commissioner of Labor Statistics and one of the developers of US Consumer Price Index. In her ASA Presidential address pointed out deficiencies in statistical training

> *What is now most lacking in statistical training in colleges, I believe, is an adequate appreciation of the fact that statistics are tools to be applied to subject matter. A highly trained mathematical statistician without knowledge of the subject can be quite as dangerous as a subject-matter specialist trying to use statistical methods of which he has little knowledge*

She claimed to be "not much of a mathematician", saying "statistics are nothing to be afraid of as long as you know what question you want to answer and if you know arithmetic".

Fun fact: A feature story describing her career in the Prescott AZ Evening Courier newspaper was headlined "Highest paid career woman at $13000 resides on a farm" (p 10, 22 Nov 1954) under a recipe for sweet and sour spareribs. A LIFE magazine article was headed "Lady Civil Servants Dept. Of Labor"

Refs
Wickens (1953) Statistics and the Public Interest, Journal of the American Statistical Association, 48: 261, 1-14, doi: 10.1080/01621459.1953.10483451 Presidential Address 112th Annual Meeting of ASA

06 Jan Modern Medical Statistics: Who's on first? Before the BMJ Statistics Notes series by Altman and Bland made statistical methods more accessible to researchers, there was *The Lancet* and Austin Bradford Hill (1937)

In Jan-Apr 1937 Austin Bradford Hill published 17 articles in *The Lancet* on medical statistics. These later became the book *Principles of Medical Statistics* which ran to 12 editions over 55 years. He was paid 3 guineas (~£135 today)

07 Jan On 1 Aug 1936 Donald Mainland published (*Br Med J* 1936;2:1229) a simple method based on standard deviations for determining "limits for the true odds" in small samples. He acknowledged Fisher for his contributions to sample size calculations. He used published data from Tumarkin's study on Bell's palsy.

Tumarkin was more than narked: he complains about "the general principle that the clinician should claim no significance for his findings until they have passed the rigorous scrutiny of the statistician", what he interpreted as criticism of basic math errors (it was all true), and in any case he did not "wish to confuse the article with mathematics."

In his reply (*Br Med J* 1937;23 Jan:192) Mainland comments that Tumarkin's letter only "justifies rather than controverts the purpose of his article" about the need for medical statistics education, as it appeared that he "seems to have misunderstood the purpose principles, methods and results of the last three-quarters of it" The issue was not about "elementary mistakes in arithmetic but of comparing inappropriate data".

The first Hill heard of this was when Fisher wrote to him 9 Apr 1937 congratulating him on his book and then asked if he knew about Mainland "who has, I understand, a work on tests of significance in medicine now in the hands of the publishers." Apparently, Hill was alarmed and 'annoyed'.

o8 Jan OTD 1886 Lowell J Reed b (d 29 Apr 1966) USA. ASA Fellow 1927, ASA President 1951, Honorary Fellow Royal Statistical Society. Best known for Pearl-Reed Verhulst model of population growth, Reed-Merrell life-table, and the Frost-Reed epidemiology model, called "one of the most influential models in epidemiological reasoning"

Fun facts: Reed and Frost demonstrated the "workings of the model behind the mathematical formulae" with a simple mechanical analogue of coloured marbles rolling down a trough. 'Susceptibles' were green, 'infected' cases red, 'immune' blue and 'contact neutralizers' white. In 1951, Reed went on TV to show it to the public to increase lay understanding of epidemiology.

His New York Times obituary called him "a pioneer of biostatistics, the study of world population growth". This strange definition of biostatistics was probably because the reporter complained that "public health statistics was so new they couldn't find a definition in the Encyclopedia Britannica or Webster's "big unabridged dictionary."

Refs

New York Times 30 Apr 1966.

Cochran. RSS https://rss.onlinelibrary.wiley.com/doi/pdfdirect/10.1111/j.2397-2327.1967.tb02314.x

https://doi.org/10.1080/00031305.1966.10480412

Engelmann (2021) A box, a trough and marbles: How the Reed-Frost epidemic theory shaped epidemiological reasoning in the 20th

century. *HPLS* 43, 105 (2021). https://doi.org/10.1007/s40656-021-00445-z
"The Johns Hopkins Science Review": Youtube "Epidemic theory: What is it?"
https://www.youtube.com/watch?v=CmhL4rVLwn0

09 Jan OTD 1915 Mollie Orshansky b (d 18 Dec 2006) USA. ASA Fellow 1974. Best known for the Orshansky Poverty Thresholds, developed when she was at the Social Security Administration. These were the first US government-accepted measures of income adequacy.

Fun fact. Both Orshansky and the development of the Poverty Thresholds were features in an episode of the TV show *The West Wing* (The Indians in the Lobby, 21 Nov 2001). However, Orshansky and her work were grossly mischaracterised

10 Jan. DYK? Long before Ronald Fisher or Joseph Berkson, probably the earliest discussion of confounding as a type of experimental bias was by John Stuart Mill in his 1843 book *A System of Logic"* (Book III: Of Induction, Chap X: Of Plurality of Causes; and of the Intermixture of Effects)

> *We require also that none of the circumstances which we do know shall have effects susceptible of being confounded with those of the agents whose properties we wish to study. We take the utmost pains to exclude all causes capable of composition with the given cause; or if forced to let in any such causes, we take care to make them such that we can compute and allow for their influence, so that the effect of the given cause may, after the subduction of those other effects, be apparent as a residual phenomenon.*

11 Jan The percent symbol % began to take form in the Middle Ages. As early as 1339 in '*Rara Arithmetica*' the Italian term *per cento* is shortened to the letter p crossed by a strike over the tail. About 1435 somebody began abbreviating it further as 'pc' with a little 'o' on top. By 1684 it had evolved to a fraction o/o

Refs
Smith DE (1908) *Rara Arithmetica; A Catalogue of the Arithmetics Written before the Year MDCI, With Description of Those in the Library of George Arthur Plimpton, of New York.*
Smith DE 1923. *History of Mathematics*.

12 Jan OTD Three birthdays of prominent statisticians.

1. Archie Cochrane (1909-18 Jun 1988) Scotland. CBE The name behind Cochrane Library and the Cochrane Collaboration. Called the father of evidence–based medicine.
2 Stella Cunliffe (1917-20 Jan 2012) UK. MBE First woman President Royal Statistical Society 1975; Director of Statistics, UK Home Office. (Stay tuned - More to follow on 20 Jan post)
3. Yvonne Bishop (1925-26 May 2015) UK/USA. ASA Fellow 1975. Best known for THE book on multivariate statistics '*Discrete Multivariate Analysis: Theory and Practice'*

13 Jan OTD 1900 Gertrude Cox b (d 17 Oct 1978) USA. ASA Fellow 1947, ASA President 1956 (3rd woman), Elected National Academy of Sciences 1975, Best

known as a pioneer in applied experimental design (*Statistical methods* with G Snedecor is a classic), she racked up many First Woman achievements: founder and head of North Carolina State Experimental Statistics Department, founding member of International Biometric Society 1947, IBS president 1968, editor of *Biometrics* 1947 -1955; Honorary Fellow Royal Statistical Society 1959,

Her 3 principles for experimental design are still more than relevant: 1. A clear statement of objectives 2. Clearly described materials and methods (especially study design treatments, N, controls); 3. A statistical *analysis plan. Emphasizing randomization, replication, and experimental controls* All the data in the book were analyzed by her computing staff and personally checked by Cox and course students.

Refs
https://mathshistory.st-andrews.ac.uk/Biographies/Cox/
Anderson RL .1990
https://web.archive.org/web/20150326021456/http://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/cox-gertrude.pdf

14 Jan Introduced by Fisher (1926), the Latin square is a row-column experimental design for two blocking factors. However, it originated with Euler (1782) who called it a 'magic square' (De quadratis magicis). Each symbol occurs only once in each row or column. Traditionally Latin and/or Greek symbols were used to fill the grid, hence the names Latin and Graeco-Latin squares. Sudoku puzzles are Latin squares.

Fisher's centenary commemorative window showing a 7 x 7 Latin square was installed in Gonville and Caius College dining hall, Cambridge in 1989. It was removed in 2020 because of Fisher's eugenicist views.

Refs
Recherches sur une Nouvelle Espèces de Quarrés Magique," *Verh. uitgegeven door het Zeeuwsch Genootschap d. Wetensch. te Vlissingen*, 9, 85-232.

15 Jan Before Tukey, there was Maurice Quenouille (9 Oct 1924-12 Dec 1973) In 1949 and 1956 he developed the jackknife, the precursor to bootstrapping. In 1958 Tukey gave it the cool name 'jack knife' and expanded the technique. The name was proposed because "it is a rough-and-ready tool that can improvise a solution for a variety of problems even though specific problems may be more efficiently solved with a purpose-designed tool.

Refs
Fernholz *Stat Sci* 2003; 18(3)336-40
Quenouille (1949). *J R Stat Soc B* 11: 68-84; (1956*) Biometrika* 43: 353-60, [?? 1, 2-1] Miller R 1974 *Biometrika* 61(1)1-15; Tukey 1958. *Ann Math Stat* 29: 614; Mosteller (1971) *Review Inter Stat Ins*t 39(3):363-368
Bio: Barnard GA (1977). *J R Stat Soc A* 140 (4): 568–569; Smith TMF (1978) *J R Stat Soc D* 27 (2): 65–86

16 Jan. From the Department of "Statistics Terms That (Fortunately) Never Caught On": In his classic 1979 paper on the bootstrap, Bradley Efron thanks friends for alternative name suggestions: Swiss Army Knife, Meat Axe, Swan

Dive, Jack Rabbit, Shotgun (Imagine describing these in papers: the Meat Axe resampling confidence interval is too perfect).

'Bootstrap' was derived from "*The Surprising Adventures of Baron Munchausen*"- the baron pulled himself out of a swamp by his own bootstraps.

Refs

Annals of Statistics 7(1): 1-26 (1979); *Significance* 16(1) 8-9 (2019)
"Aventures du baron de Münchhausen." Traduction nouvelle par Théophile Gautier fils; illustrées par Gustave Doré. Paris: Furne, Jouvet et cie, [ca. 1862]

17 Jan OTD 1911 Francis Galton d (b 16 Feb 1822) GB. KBE FRS Polymath – you name it he did it. Infamous hardcore eugenicist and misogynist. He coined the terms 'eugenics' in 1883 and nature vs nurture". But his statistical innovations include invention of questionnaires, variance and standard deviation, experimental derivation of normal distribution, correlation and regression, regression to the mean.
Also famous for exploration cartography, forensic fingerprinting, meteorology (first weather map), study of human intelligence, etc

In an effort to reach a wider audience, Galton worked on a novel entitled *Kantsaywhere* from May until December 1910. The novel described a utopia organized by a eugenic religion, designed to breed fitter and smarter people.

18 Jan OTD 2022 Sir David Roxbee Cox d (b 15 Jul 1924) UK. FRS FBA FRSE FRSC President Royal Statistical Society 1980. Called the most eminent statistician of the twentieth century." First recipient of International Prize in Statistics; Guy Gold and Silver, George Box, Copley medals. Best known for his "pioneering and hugely influential" statistical methods including logistic regression, Cox proportional hazards model. >18 books and >300 papers

Ref Reid N. 1994. *Statistical Science* 9(3): 439-455

19 Jan Fibonacci (1170-1250) may have been first to use the Latin 'additio'. It appeared in English as *addicion in* '*The crafte of nombrynge* (1300). The symbol + shows up about 1456 as a contraction of the Latin et (for 'and').

Robert Recorde popularised + and - in '*The Whetstone of Witte'* (1557): "There be other 2 signes in often use of which the first is made thus + and betokeneth more: the other is thus made – and betokeneth lesse."

Refs

Smith DE *History Of Mathematics*.
Recorde R
https://ia601604.us.archive.org/14/items/TheWhetstoneOfWitte/TheWhetstoneOfWitte_text.pdf

20 Jan OTD 2012 Stella Cunliffe d (b 12 Jan 1917) UK. MBE First woman President Royal Statistical Society 1975. As statistician for the Danish Bacon Company, she was in charge of WW2 London food rationing. Among first civilians to enter liberated Belsen. Pioneered statistically based quality control assessments at Guinness.

Later she became Home Office Director of Statistics, where she shaped key public policy on crime and punishment

She saw herself as a supremely practical statistician:

> "*Statistics is concerned much more with people than with vague ideas. If you do not understand human beings, you will not understand statistics either*."

Refs

Significance 2006 3(3)126-9;
JRSS A: Statistics in Society 175(4) 1057-9

21 Jan History of bar charts: William Playfair (1759-1824) is usually given the credit for developing the bar chart: In his book *The Commercial and Political Atlas; Representing, by Means of Stained Copper-Plate Charts, the Exports, Imports, and General Trade of England, at a Single View. To which are Added, Charts of the Revenue and Debts of Ireland, Done in the Same Manner by James Correy*. Playfair invented the line chart, times series plots, and the bar chart.

However, Nichole Oresme (c. 1320–1325 – 11 July 1382) in *De latitudinibus formarum*, 1486 shows a sketch of what has been described as a proto bar-chart. It was used for visualizing the velocity of an accelerating object against time.

22 Jan. In a paper read to the Royal Society on 17 Nov 1921 RA Fisher proposed a unifying theory of statistics, with focus on estimation and the 3 criteria of consistency, efficiency and sufficiency.
Fun fact: George Darmois (1888-1960) translated sufficient as 'exhaustif'.

Refs

Fisher RA 1922. On the mathematical foundations of theoretical statistics *Phil Trans Royal Soc Lond A* 222: 309-368 doi: 10.1098/rsta.1922.0009
Geisser, S. (1992). Introduction to Fisher (1922) On the Mathematical Foundations of Theoretical Statistics. In: Kotz, S., Johnson, N.L. (eds) *Breakthroughs in Statistics*. Springer Series in Statistics. Springer, New York, NY. https://doi.org/10.1007/978-1-4612-0919-5_1

23 Jan. Division as a concept was slow to catch on possibly because it lacked consistent notation. The horizontal bar — was first used in Arabic math, then introduced to Europe by Fibonacci in the 13th c. Michael Stiefel (1544) used both the bracket) and the letter D (1545). Rahn introduced the obelus ÷ in 1659, and De Morgan the forward slash / in 1845. The symbol $\overline{)}$ used in long division does not have a name.

24 Jan OTD 1895 Karl Pearson expanded on the definition of the histogram as a term for a common form of graphical representation, i.e., by columns marking as areas the frequency corresponding to the range of their base." He had previously introduced the term on 20 Nov 1891 in the lecture on 'Maps and Chartograms', the last of his Gresham lectures on the 'Geometry of Statistics'. Open to the public, the lectures showed how geometry could be used as 'a fundamental method of investigating and analyzing statistical material'.

Refs

Magnello ME (1996) *Br J Hist Sci*, 29(1): 43-63
Pearson K (1895) *Phil Trans Royal Soc A*, 186:343-414+16 plates

25 Jan Karl Pearson was appointed to the Gresham College Chair of Geometry in 1890. Between 17 Nov 1891 and 11 May 1894 he gave 38 Gresham lectures free to the public. His methods for teaching statistics were amazing: To teach probability he used dice, coins, roulette, lottery tickets, and once "scattered 10,000 pennies all over the lecture room floor and asked his students to pick them up and arrange them in heads or tails: 'the result was very nearly half heads and half tails, thus proving the law of averages and probability'.

After a lecture on experimental deductions which involved the use of 16,178 throws of the ball at the Monte Carlo Roulette Table, teetotums, and 2138 tickets from lotteries, one of his students remarked that the lecture was like 'an opera without the last act '. It is, perhaps, not surprising that the numbers of students 'increased five to ten-fold' in the first couple of years.

The Gresham College lectures gave Pearson the opportunity to develop the foundation of his statistical theory. The last ten resulted in his first two published statistical papers.

Refs

Pearson KP 1894 *Contributions to the mathematical theory of evolution'* (CMTE), *Philosophical Transactions* A, 185, 71-100;
Abstract in *Proceedings of the Royal Society* (1893), 54, 329-33;
'MTE. II. Skew variation in homogeneous material', *Philosophical Transactions A* (1895), 186, 343-414; Abstract in *Proceedings of the Royal Society* (1894), 57, 257-60

26 Jan During WWII, the Statistical Research Group (SRG) at Columbia University was formed in 1942 to support the war effort. The distinguished roster included WA Wallis, Hotelling, Wald, Wilks, Deming, and Milton Friedman. Work involved probability studies and development of statistical methods for predicting aircraft combat damage, bombing salvos, patterns for landmine clearance, and sequential analysis.

Refs

Rees 1980. *Am Math Monthly*, 87(8): 607-621; Wallis 1980. J *Am Stat Assoc* 75(379): 320-30
ly, 87(8): 607-621; Wallis 1980. *J Am Stat Assoc* 75(379): 320-30

27 Jan OTD 1832 Charles Dodgson b (d 14 Jan 1898) England, maths lecturer at Oxford. Worked on probability linear and matrix algebra, probability and symbolic logic. Developed Dodgson condensation for computing determinants of square matrices. However, he is best known as Lewis Carroll author of *Alice in Wonderland*. He always denied the story that Queen Victoria was so taken by AIW that she commanded that he dedicate his next book to her and he sent her *An Elementary Treatise on Determinants.*

28 Jan OTD 1892 Carlos Bonferroni b (d 18 Aug 1960) Trained as a musician then switched to maths. Best known for Bonferroni inequalities 1935-6. The Bonferroni adjustment for simultaneous inference is named for him but was actually developed by Olive Jean Dunn. Also worked on actuarial mathematics; probability and statistical mathematics; analysis, geometry and rational

mechanics. It is thought his work is not well known because his books were never typeset. He thought books too expensive for students (not wrong there) so he had his books printed from his handwritten teaching material.

Refs

Benedetti (1982). Carlo Emilio Bonferroni (1892-1960). *Metron*, **40**, N.3-4, 1-36.
https://mathshistory.st-andrews.ac.uk/Biographies/Bonferroni/

29 Jan The first blinded systematic controlled trial was performed in 1784 when the French Royal Commission on Animal Magnetism investigated medical claims of 'mesmerism', named after Franz Mesmer. The medical profession was understandably alarmed by the craze for its therapeutic use, especially since Mesmer was foreign, had no medical training and irritatingly was making lots of money.

The commission, which included Benjamin Franklin and Lavoisier, debunked claims w rigorous experimental methodology: method standardisation, controls, replication and even allocation concealment. Subjects were blindfolded and presented with sham or"mesmerised" items in unguessable order.

Refs

Gould SJ The Chain of Reason vs. The Chain of Thumbs. *Natural History* 89(7) (July 1989), *Bully for Brontosaurus,* 1991

30 Jan OTD 1928 Johannes Fibiger d (b 23 Apt 1863) Denmark. Investigating serum treatment of diphtheria, he conducted possibly the first clinical randomised controlled trial in 1898. It is called a milestone for use of random allocation "as a pivotal methodological principle".

However, his 1926 Nobel Prize in Medicine for his "discovery" of the Spiroptera carcinoma has been called "one of the biggest blunders made by the Karolinska Institute.

Ref

Hrobjartsson et al (1998). The controlled clinical trial turns 100 years: Fibiger's trial of serum treatment of diphtheria. *BMJ*, 317 (7167): 1243–1245.

31 Jan OTD 1919 Oscar Kempthorne b (d 15 Nov 2000) UK. ASA Fellow 1952 Best known for design of experiments and genetic statistics. Strongly promoted randomization and considered ANOVA to be "one of the most powerful statistical techniques". Part of the Rothamsted group 1941-6; founder of the "Iowa school" of design of experiments. Although he was often critical of both model-based inference and Bayesian design, paradoxically he acknowledged that "design of experiment is a Bayesian...process".

Refs

Folks JL (1995) *Statistical Science*, 10 (4): 321–336. doi:10.1214/ss/1177009867
Hinkelmann K (2001) *Statistical Science*. 16 (2): 169–183. doi:10.1214/ss/1009213289
AmStat News https://magazine.amstat.org/blog/2017/01/31/sih-kempthorne/

History of Statistics Twitter log February 2023

01 Feb On 17 Oct 1884 Charles Sanders Peirce (1839–1914) and Joseph Jastrow (1863-1944) USA presented results of possibly the first randomised AND blinded trial in experimental psychology. Order of stimulus presentation was formally randomised by shuffling ordinary playing cards and concealed from the test subject.

Ref Peirce CS, Jastrow J (1885) *Memoirs of the National Academy of Sciences*, 3, 73-83. http://psychclassics.yorku.ca/Peirce/small-diffs.html

02 Feb. Happy Groundhog Day! Unfortunately, groundhogs are mostly pretty rubbish at predicting spring. Recent analyses of 33 marmot prophets show almost no relationship between observed versus actual onset of spring (P = 0.28) although some are better than others. Oilsprings Ollie is #1 with >75% success (2023 prediction 6 more weeks of winter), Punxsutawney Phil (39-50%), and Wiarton Willie 27.3-36.4%

Refs

Ross et al (2021) *Weather, Climate and Society* 13:503-10, doi: 10.1175/WCAS-D-20-0171.1

https://fivethirtyeight.com/features/groundhogs-do-not-make-good-meteorologists/

https://uwimprint.ca/article/wiarton-willies-predictions-wildly-off-the-mark-much-of-the-time-uw-students-find/;

https://www.cbc.ca/news/canada/kitchener-waterloo/wiarton-willie-university-of-waterloo-students-1.5896555

03 Feb 3 Feb OTD 1877 Janet Lane-Claypon, Lady Forber b (d 17 Jul 1967) UK. One of the founders of modern epidemiology. Pioneered case-control and retrospective cohort study design. Modern re-analyses show breast cancer risk factors she identified are still consistent with current evidence.
Other firsts: Recognised the role of confounding in observational studies and analysed data; accordingly, the first use of t-tests in an epidemiological study; one of the first Dr-Dr (both MD and DSc); First woman awarded a British Medical Society research scholarship.

Refs

Winkelstein W Jr (2006). *Epidemiology* 17(6): 705; (2004) *Am J Epidemiol* 160(2):97-101. doi: 10.1093/aje/kwh185.

Press and Pharoah (2010) *Epidemiology* 21: 566-572

Mantel N, Haenszel W (1959) *J Natl Cancer Inst* 22:719–748.

04 Feb Karl Pearson conducted the first clinical meta-analysis in 1904 for a comparison of infection and mortality in soldiers either inoculated for typhoid or not. The last line of the tabulated results gives a pooled estimate of effect. His method for combining p-values was revisited by FN David 1934. In Jun 1938 his son Egon Pearson published one of the first formal methods for combining results from different trials. KP's method has been recently reinvented for genomic application.

Refs

Pearson K (1904). *BMJ* 3:1243-1246; David FN (1934) *Biometrika* 26 1–11;

Pearson E (1938) *Biometrika* 30: 134-148; Fisher RA 1948 *Amer Stat* 2:30-1;

Owen AB (2009) *Annals of Statistics* 37(6B) 3867-92

05 Feb OTD 2019 Edward Simpson d (b 10 Dec 1922) UK. CB Fellow Royal Statistical Society 1946. Best-known for Simpson's paradox (developed while a graduate student at Cambridge), and Simpson's diversity index. He was a Bletchley Park code breaker with Alan Turing 1942-5. In 2017, at the age of 95, he contributed book chapters in *The Bletchley Park Codebreakers,* on the crypt-analytic process Banburismus.

Fun fact. Simpson's paradox was featured in an episode of the TV cartoon show *The Simpsons*

Refs

Simpson EH (1951). The interpretation of interaction in contingency tables. J *Royal Stat Soc, Series B*. 13: 238–241; *Significance*, 7(2) June 2010 doi:10.1111/j.1740-9713.2010.00424.x

06 Feb The Greek letter sigma Σ was first used for 'summation' by Euler (1707-1783) in 1755. Lagrange (1736-1813) used it a bit, but its use didn't pick up steam (as it were) until 1829 when Fourier used it in *Theory of Heat*, and G.s. Klügel (1739-1812) in *Mathematiscehes Wörterbuch*

Ref Cajori F*. History of Mathematical Notations*

07 Feb John Napier (1 February 1550 – 4 April 1617) Scotland. 8th Laird of Merchiston invented logarithms (which he first called 'numeri artificiales') to be used as a shortcut to 'save time and limit slippery errors of calculations' (*Minifici Logarithmorum Canonis Descriptio,* 1614).

Logs got rave reviews from the beginning: Henry Briggs wrote to James Ussher on 10 March 1615, saying:

> "*Napper, Lord of Markinston, hath set my Head and Hands a Work, with his new and admirable Logarithms. I hope to see him this summer, if it please God, for I never saw a book which pleased me better or made me more wonder."*

In his English translation, *A Description of the Admirable Table of Logarithmes* (1616), Edward Wright said "*This new course of Logarithmes doth cleane take away all the difficultye that heretofore hath beene in mathematicall calculations*."

The word logarithmus may have been derived from 'logos' meaning 'reckoning' or 'calculation' but according to Briggs (1624), it was because logarithms are numbers that preserve the same ratio to each other. Jobst Bürgi (1620) called the logarithm Die Rothe Zahl for no better reason than the logarithms were printed in red and the antilogarithms in black in his book *Progress Tabulen*

Ref Smith DE *History of Mathematics.*

08 Feb OTD **1587** Mary Queen of Scots was executed at Fotheringhay on the orders of her cousin Elizabeth I. Also, OTD TODAY 2023 (436 years later) the results of the decryption of 50 of her cypher letters from her 19 years of captivity are now online. It required combination of modern computerized

cryptanalysis, manual codebreaking, and traditional linguistic and contextual analysis.

REF Lasry, Biermann, and Tomokiyo (2023) *Cryptologia*, doi 10.1080/01611194.2022.2160677

Also OTD **1700** Daniel Bernoulli b (d 17 Mar 1782). Best known for eponymous principle for conservation of energy. In statistics he did important work on probability and political economy, and made one of the first attempts (1766) to analyze censored data in a study of smallpox m("meanest man in maths") because they shared first place in a scientific contest and Johann apparently couldn't deal with the "shame of being compared Daniel's equal"

09 Feb George Snedecor (1881-1974) first coined the 'F' in 'F-statistic' and the F-distribution in his 1934 monograph in honour of Ronald Fisher. He also provided simple worked examples of analysis of variance and covariance, and tables of critical values corresponding to Fisher's 5% and 1% levels for different degrees of freedom. Unhappily Fisher refused to use F because unknown to Snedecor, Mahalanobis had already tabulated the variance ratio in 1932.

Refs

Snedecor 1934 *Calculation and Interpretation of Analysis of Variance and Covariance* Collegiate Press

Box JF. *R.A. Fisher: The Life of a Scientist*

10 Feb OTD 1916 Regina Loewenstein b (d 16 May 1999) USA. Public health statistician at Columbia, ASA Fellow 1976 and co-founder of ASA Caucus for Women In Statistics 1971. The Regina Loewenstein Prize "for Academic Excellence in Health Policy and Management" is given in her honour.

Refs. https://cwstat.org/; Golbeck A (2020) *Significance* Apr 42-4

11 Feb The 1959 paper on the Kaplan-Meier product-limit estimator method is now one of the most highly cited statistical papers of all time. But it was almost unnoticed for over 10 years. Originally developed for engineering applications, the method took off when Gehan 1969 used it to study cancer survival and Cox 1972 highlighted its importance for clinical survival analyses. Sadly, while Paul Meier became a statistics rock star, Ed Kaplan and his mathematical contributions have been overlooked.

Ref

Stalpers LJA Kaplan EL (2018) *BSHM Bulletin: J Brit Soc Hist Math*, 33:2, 109-135, doi: 10.1080/17498430.2018.1450055

12 Feb OTD 1864 Kate Claghorn b (d 22 May 1934) USA. First woman ASA Fellow 1918 (and the only one until 1937), First woman to publicly receive the PhD degree from Yale 1896. She is best known for prescient work on social statistics, especially on immigration, the juvenile legal system, and workplace sexual harassment. She was also a signatory on National Association for the Advancement of Colored People (NAACP) founding document in 1910.

In 1912, a journalist wrote:

Miss Kate Claghorn is holding down a man's job in the tenement house department because there was no man smart enough to fill it. Twice she

stood the test of an examination framed in Columbia University, which was designed, if anything, to eliminate women from the competition, but which in the end eliminated the men

13 Feb OTD 1766 Thomas Malthus b (d 29 Dec 1834) England. FRS. With Charles Babbage and Adolph Quetelet, he was cofounder in 1834 of the Statistical Society of London (now Royal Statistical Society). He is probably best known for his 1798 *Essay on the Principle of Population.* He argued against the contemporary view that societal progress was achievable by unchecked population growth.

The central premise of his book was that populations increase geometrically but production only arithmetically, with limits to growth imposed only by 'virtue', 'vice' and 'misery'. Paradoxically he disapproved of contraception. Darwin proposed Malthus' growth model as the mechanism for natural selection and speciation.

Refs

Magnello E 2009 *Significance* Jun 86-88

Mouat FJ *J Stat Soc Lond Jubilee Volume* (Jun. 22 - 24, 1885):14-71, 359-371

Hilts VL 1978. *Isis* 69(1):21-43

14 Feb Happy Valentine's Day! A (possibly) unexpected link with statistics is found in the life of Giacomo Casanova (2 Apr 1725 – 4 Jun 1798) world's best-known librarian and *checks notes * lover, sexual athlete, and cad. He was keen on applying 18th c probability theory to gambling strategies. His plan was to establish a national lottery system in Paris, with the sole goal of making as much money as possible in the shortest amount of time. It lasted a surprisingly long time, from 1757 to 1836.

Casanova claimed a mastery of probability that fooled advisors of King Louis XV, and even the mathematician d'Alembert. But he didn't really understand how probabilities worked. He bragged to his sponsors how his martingale system would guarantee a win every time. Every time he lost, the bet was redoubled until he won. Unfortunately win–loss outcomes of each bet are IID random variables, so the system only works if bets, time and $$ are nearly infinite. Casanova ended up in debtors' prison. He did realise that cheating is always an option

Refs

Casanova *Histoire de ma vie*

Stigler S 2022. *Casanova's Lottery: The History of a Revolutionary Game of Chance*, University of Chicago Press;

Haugen MW 2022. *Italian Studies*. doi.org/10.1080/00751634.2022.2069409

15 Feb OTD 1946 John Mauchly and J Presper Eckert unveiled the ENIAC: the world's first general-purpose electronic digital computer. Completely overlooked were 6 women, highly trained in mathematics, who did the programming: Fran Bilas, Jean Jennings Bartik, Ruth Lichterman, Kay McNulty, Betty Snyder, Marlyn Wescoff. They were not invited to the first celebration dinner or ENIAC's 50th Anniversary.

ENIAC had no keyboard, mouse, display, or programming languages. 'Programming' meant manual adjustment of complex wiring, switches, vacuum tubes, and cables; 'debugging a program' meant climbing inside to find bad connections. Setting up a single calculation could take days, and weeks for a full program.

Refs
Schatz K https://www.linkedin.com/pulse/how-six-girls-designed-first-all-electronic-computer-set-kate-schatz/
Jones B, Larsen L. 2019
https://www.digitaltrends.com/computing/remembering-eniac-and-the-women-who-programmed-it/
https://www.codecademy.com/resources/blog/eniac-six-women-programmed-computer/
Kleiman K. *The Secret History of the ENIAC Women* TEDx talks
https://www.youtube.com/watch?v=Zevt2blQyVs
Medium May 16 2017 https://timeline.com/women-pioneered-computer-programming-then-men-took-their-industry-over-c2959b822523

16 Feb OTD 1919 Nathan Mantel b (d 25 May 2002) USA. ASA Fellow 1962 Best known for the logrank test for survival data and the Mantel–Haenszel test and odds ratio for contingency table analysis. Saying that statistics has to be "math for something", he also made other major contributions to epidemiology including diagnostic tests, time-dependent covariate analysis, and logistic regression. During military service as "Clerk Non-typist' he worked on drug dosing allometry and probit analysis

Refs
1999 Special Issue: 30 December 1999 *Statistics in Medicine*, 18(24):3377–3513 Papers In Honour of Nathan Mantel's 80th birthday
Mantel N 1966. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports*. 50 (3): 163–70

17 Feb OTD 1890 Sir Ronald Aylmer Fisher b (d 29 Jul 1962) UK. FRS KB Guy Gold Medal 1946. A revolutionary genius, he was described by Bradley Efron as "the single most important figure in 20th century statistics". Made monumental and lasting contributions to both theoretical and applied statistics, especially design of experiments, randomization, and analysis of variance. He was also one of the three principal founders of population genetics. The International Biometric Society (he was first president) was founded in consequence of the "Fisherian revolution in research methodology."

However, his undoubted racist and eugenicist views have triggered important debate about wider issues of diversity, representation, and academic values.

Refs
Efron B (1998) R. A. Fisher in the 21st century (1996 R. A. Fisher Lecture). *Statist. Sci*. 13 (2) 95 - 122, https://doi.org/10.1214/ss/1028905930
Box JF (1978) *R.A. Fisher: The life of a scientist*
Fienberg SE, Hinkley DV eds. (1980). *R.A. Fisher: An appreciation*.

Springer-Verlag.
Rao CR (1992). R. A. Fisher: The founder of modern statistics. *Statist. Sci* 7: 34–48. doi:10.1214/ss/1177011442.
Gonville and Caius College: *Fisher in the 21st Century (21-22 April 2022).* https://www.cai.cam.ac.uk/discover/fisher-21st-century

18 Feb OTD 1871 George Udny Yule b (d 26 June 1951) UK. Guy Gold Medal 1911, President Royal Statistical Society 1924-6. Frank Yates considered him one of the pioneers of modern statistics. Made important contributions to correlation, regression, time series analysis, and power law distributions. His use of autoregressive models to model Wolfer's sunspot time series is now a staple course exercise. His textbook '*Introduction to the Theory of Statistics'* went to 14 editions.

Originally an experimental physicist, he saw the light and switched to statistics when working as demonstrator for Karl Pearson at UCL. Pearson never took criticism well, so they fell out over obvious mathematical problems with Pearson's correlation and χ^2 tests. For many years he, Edgeworth, and Bowley were the only RSS members at all interested in mathematical statistics.

Fun facts: He liked fast driving, and after retirement he took up flying but because of his age he could not get insurance, so he bought his own plane. He began to study Latin to read Thomas à Kempis in the original, which led him to work on statistical methods as a tool for investigating disputed authorship 'The statistical study of literary vocabulary.'

Ref

Yates F (1952). *Obituary Notices of Fellows of the Royal Society*. 8 (21): 308–323. doi:10.1098/rsbm.1952.0020. S2CID 178300526.

19 Feb Statistical original sin. RA Fisher provides one of the earliest examples of pseudo-replication. Yes, he did. He analysed data quantifying the relationship between the number of fruit-fly eye facets and rearing temperature. There were 9 temperatures with one rearing bottle assigned to each, and 50-137 flies in each bottle. The significance test for slope used the correct experimental unit (bottle) with 7 degrees of freedom, but the test for linearity did not (814 df). Joseph Berkson, who did not care for significance tests, pointed out that the small P may have been an artefact of methodology without realising what the real issue was. Neither did Fisher who promptly attacked "It is not my purpose to make Dr. Berkson seem ridiculous, nor, of course, to prevent him from providing innocent amusement".

Refs

Berkson J. 1942. Tests of significance considered as evidence. *J Amer Stat Assoc* 37: 325–335.
Fisher RA 1943. *J Amer Stat Assoc* 38:103-4

20 Feb OTD 1762 Tobias Mayer d (b 17 Feb 1723) Best known for his lunar tables for finding longitude at sea, his 1750 work on lunar motion was an important forerunner of the least-squares method. He used data on the position of the Manilius crater to develop a set of 27 linear equations ('equations of condition') to solve for 3 unknown parameters.

Instead of choosing arbitrarily three equations, he devised a clever method using

all equations and obtained as good a solution as if he had used the method of least squares which was not yet known.

He also introduced the annoying ± x for the error.

Refs
Stigler SM (1986) *The history of statistics*. Belknap Press
Steele JM (2012). An Integrated Approach: Tobias Mayer. In: *Ancient Astronomical Observations and the Study of the Moon's Motion (1691-1757). Sources and Studies in the History of Mathematics and Physical Sciences.* Springer, Boston, MA. doi.org/10.1007/978-1-4614-2149-8_7

21 Feb DYK? Akaike's Information Criterion (AIC) for model selection traces its origins to a 1951 paper by Kullback and Leibler that quantified "information" in the context of RA Fisher's concept of sufficient statistics. The Kullback–Leibler divergence is a measure of similarity between probability distributions with major applications in information theory and cryptology. In turn it has roots in Boltzmann's 1877 concept of statistical entropy. Minimizing AIC in a statistical model system is essentially like maximizing entropy in a thermodynamic system.

Refs
Burnham KP, Anderson DR (2004) *Sociol Methods Res*, 33: 261–304, doi:10.1177/0049124104268644
Kullback S Leibler R (1951) *Annals Math Stat* 22:79-86

22 Feb OTD 1796 Adolphe Quetelet b (d 17 Feb 1874). Belgium. The man who gave us the Average Man ("l'homme moyen") and unhappily the Body Mass Index. Cofounder of the Royal Statistical Society. Among the first to apply the then-new science of probability and statistics to social science. His 1835 book *Essai De Physique Sociale* presented his theory of human variance around the mean with traits following a normal distribution.

He was statistical mentor to Florence Nightingale: they met in 1860 when he attended the International Statistical Congress in London and corresponded for years. Nightingale's copy of *Physique Sociale* is heavily annotated [Her first edition copy was later presented to Oxford]. She nagged Quetelet constantly by letter to produce a 2nd edition, which he finally got around to completing years later – he sent a copy to her in 1869.

Refs Diamond and Stone 1981. *J Royal Stat Soc Series A* 144(1): 66-79

23 Feb Period or comma? The notation separating integers from the fractional part of a decimal number has always been a bit confusing. USA uses the period, UK the raised period, and Europe the comma. The confusion goes back a long way; both John Napier 1617 and Edmund Wingate 1629 used both. Wingate said

> *'Notation conſiſts in the knowledge of two things, viz. the Order of places, and the Value of every place in any number. First, Diſtinguiſh by a Comma, or Point, every three places, beginning at the right hand, and proceeding towards the left, ſo will the aforeſaid number be diſtinguiſhed into parts, which may be called Periods.*

24 Feb The earliest fractions were reciprocals of integers, used by the Egyptians at least as early as 1000 BC. Adelard of Bath (1080?-1142-52?), is

credited with introducing the term ‘fraction’ (*fractiones*). Geoffrey Chaucer 1391 called it *fraccions* (disappointingly not in the *Canterbury Tales* but in *A Treatise On The Astrolabe*). Robert Recorde (1572) defined it as "*A Fraction in deede is a broken number*"

25 Feb OTD 2017 Dorothy P. Rice d (b 11 Jun 1922) USA. ASA Fellow 1977 Director of the National Center for Health Statistics 1976-1982 and development of the #National Death Index. Her pioneering research in aging andcost-of-illness led to the creation of Medicare.

26 Feb OTD 1985 Albert Turner Bharucha-Reid d (b 13 Nov 1927) USA Dean AandS Wayne State University, Distinguished Professor Atlanta. Best known for work on probability theory, Markov processes (essential for AI and ML) and stochastic epidemic models. The National Association of Mathematicians has named a lecture series in his honour.

27 Feb Before statistical physics, there was ….sociology? Yes. Both James Clark Maxwell and Ludwig Boltzmann independently proposed that Quetelet’s work on the error law and the mathematical treatment of variation could apply to the law of gases, and both used analogies with social statistics (1872-3).

28 Feb No birthday for Herman Hollerith this year as he was b 29 Feb 1860 (d 17 Nov 1929) Germany/USA. Inventor of the punched card tabulating machine (1884) and founder of a company that later became part of IBM. Best known for automating the tabulation work of the 1890 US census. In an 1888 test his system tabulated over 10,000 census records in 5.5 hours, compared with >45 h for the competition. This was the beginning of the data processing revolution, characterized by automation and mechanized binary code. However, punch cards were a fixture for decades; some of you out there doubtless remember carrying shoeboxes full of punch-cards to computer class (you young things don’t have to listen)

History of Statistics Twitter Log March 2023

01 Mar Although Euclid (300 BC) is credited with the first known use of exponents, the negative exponent was not mentioned until the 15th c by Nicolas Chuquet (1445?-1488). Credited with the first use of a negative number as a power, he also invented the radical symbol for roots and was first to represent powers with a raised number. His notation for $-12x^{-2}$ was $m.\ 12^{2m.}$

Not much is known about him. His book '*Le Triparty en la Science des Nombres*' was remarkably ahead of its time but was lost until rediscovered and published by Aristide Marre in 1880

Ref Cajori F. History of Mathematics.

02 Mar OTD (possibly) 1823 Carl Friedrich Gauss published his '*Theoria Combinationis Observationum Erroribus Minimis Obnoxiæ*' in which he implicitly developed the least-squares principle of best linear unbiased estimators (BLUE) as "error consistency". BLUE were only explicitly named in 1938 by FN David and Jerzy Neyman who called it "Markoff theorem on least squares". Fisher pointed out the precedence of Gauss to which Neyman replied that the principle "was developed by Gauss, but not in a very clear way. It was developed and put into practical form by…Markoff". So there Fisher.

Refs

Gauss CF (1823) Theoria Combinationis Observationum Erroribus Minimis Obnoxiæ: Pars Prior, Pars Posterior; Supplementum, Göttingen: Dieterische Universitäts-Druckerei

Chipman J 2014. Gauss-Markov Theorem *International Encyclopedia of Statistical Science:* 577-580

03 Mar For Women's History Month 2023, @HoS_ASA will occasionally feature some of the overlooked women in statistics. Cleo Youtz (1909–2005) was the long-time research assistant of Frederick Mosteller, and historian of Harvard Statistics Department She was listed as co-author on several important publications, but mostly served as 'computer', performing all the necessary, tedious, and intensive computations needed for Mosteller's mathematical statistics papers.

04 March @HOS_ASA is saddened to hear of the passing on 3 Mar 2023 of Barbara Everitt Bryant (b Apr 1926) ASA Fellow 1998, first woman to head the US Census Bureau 1989-1993. Under her direction the Bureau increased accuracy of the census, improved economic statistics. The modernizing of computing infrastructure and incorporation of total quality management concepts resulting in more streamlined and computer-intensive interview and census-taking processes.

She is best known for endorsing statistical methods for adjusting the 1990 Census figures for undercount of minorities. These were rejected for political reasons by then-Sec Commerce Robert Mosbacher. Bryant directed that a copy of the adjusted 1990 Census figures was to be deposited with ICPSR to be preserved as part of the historical record.

Refs

https://www.npr.org/2023/03/03/1072057378/first-woman-census-bureau-director-barbara-bryant

"Barbara Bryant, the first woman to head the U.S. census, has died at 96". *NPR. March 3, 2023*. Retrieved March 3, 2023.

04 Mar OTD 1897 Mary Eleanor Spear b (d 22 Jan 1986) USA. Pioneer data visualization specialist at the US Bureau of Labor Statistics. In particular, she pioneered the box plot, usually credited to John Tukey. She authored 2 books (1952, 1969) on effective graphic technique and presentation when all graphics and data visualization was done by hand.

For a really excellent overview of her contributions and methods of working, see Ben Jone's article in Medium https://t.co/vKaAEGM4KM

Also OTD 1906 Eleanor Josephine Macdonald b (d 26 Jul 2007) USA. First cancer epidemiologist. Established first cancer registry in US. Early proponent of concept of cancer as preventable disease and early adopter of computer technology to cancer research

05 March OTD 1574 William Oughtred b (d 30 Jun 1660) England. Best known for invention of slide rule based on logarithms (1622) and his popular book *Clavis Mathematicae* (*Key of Mathematics)* used as a textbook by Isaac Newton. He made the case for less words and more symbols in maths, and introduced x for multiplication, : for proportions, and abbreviations for trig functions "sin" and "cos". Instead of a period, he used a vertical line | to separate whole numbers from decimals.

In real life he was a clergyman. John Aubrey says he was "a *pittifulle preacher… because he never studied itt but bent all his thoughts on the Mathematiques*" and adds maliciously how two of his maths students "*studyed so muche*" they went mad.
Fun facts. The Oughtred Society was formed in 1991 for slide rule collectors. Because he was a staunch Royalist, he narrowly escaped execution by Oliver Cromwell. He dedicated the 2nd edition of Clavis 1647 to Sir Richard Onslow who helped save him.

Refs
Cajori F,1916. *William Oughtred, a Great Seventeenth Century Teacher of Mathematics,* Chicago.
https://www.maa.org/press/periodicals/convergence/mathematical-treasure-william-oughtred-s-the-key-of-the-mathematicks

06 Mar. The earliest use of 'big Pi' ∏ as the product symbol is a bit of a mystery. Cajori gives Gauss priority for mention in his 1812 '*Disquisitiones Generales circa Seriem Infinitam'* but this was really the pi function Π(k,z) not the product. ' Big Pi' was used by Riemann 1859 in *Über die Anzahl der Primzahlen unter einer gegebenen Größe but without commenting about where it came from.*

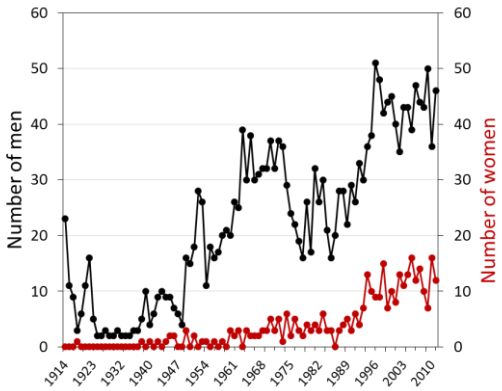
07 Mar OTD 1964 Samuel S. Wilks d (b 17 June 1906). USA. ASA Fellow 1940, ASA President 1950. Called the "Statesman of Statistics" he is best known for Wilks's lambda distribution, and with Walter Shewhart for editing the John Wiley 'Series in Statistics' that "helped change statistics from a discipline with a few isolated books in 1931 to a field with a large solid literature in 1964". He also developed Princeton as a major centre for quality control. The Samuel S. Wilks

Memorial Medal and Award was established in 1964

Fun fact: Fisher harshly criticised Wilks's 1962 book '*Mathematical Statistics*' because it was in paperback rather than hardcover

Refs Mosteller F (1964). *Amer Stat* 18 (2): 11–17.
Stephan FF et al *J Am Stat Assoc* 60(312):939-66.

08 Mar On International Womens' Day 2023 let's see how the number of women inducted as ASA Fellows has changed over the years:



Not too impressive ASA

09 Mar OTD 1908 Esther Seiden b (d 3 Jun 2014) Poland/Israel/USA. Elected to the International Statistical Institute and Fellow of the Institute of Mathematical Statistics. Best known for combinatorial theory, design of experiments (especially orthogonal arrays and problems of confounding in factorial designs) and finite geometry.

Refs
Samuel-Cahn, Ester (1992) *Stat Sci* 7 (3): 339–357, doi:10.1214/ss/1177011231
Ann. Statist., 1980, 8, 1284–1292.
Laycock PJ, Seiden E (1980) On a problem of repeated measurement design with treatment additivity. *Annals Stat 8*(6), 1284–1292.

10 Mar Multiplication is indicated by several very different symbols. William Oughtred used X (don't forget these were the days when symbolic representation of maths operations was fairly uncommon). Leibnitz hated X "as it is easily confounded with x; ... often I simply relate two quantities by an interposed dot and indicate multiplication by ZC · LM" (1698).
Possibly Thomas Harriot (1560-1621) used the dot earlier. The asterisk * was used by Johann Rahn in 1659 in '*Teutsche Algebra*'. and is used in computer coding for the same reasons given by Leibnitz.

Ref Cajori F. *A History of Mathematics*

11 Mar OTD 1941 Jessamine S. Whitney d (b 1880). USA. Public health statistician for US Census Bureau and Bureau of Labor Statistics; Pioneering studies of tuberculosis especially in relation to occupational health and "effects of tuberculosis migration". Appointed chief statistician of National Tuberculosis Association she was the first woman serving as national representative at the International conference on the Classification of Causes of Deaths in Paris 1929. In 1935 she was described as the "ranking woman vital statistician of this

country, and probably of the world."

Fun facts. When an English teacher in Puerto Rico, she was the first woman to drive a car 1904. She was also a hard-core baseball fan and possibly the earliest sabermetrician. She improved on Henry Chadwick's 1859 box score system, computing an index similar to the cost-of-living index, and claimed an 80% success rate in predicting winners of the World Series. She also wrote a syndicated baseball column for one season.

Refs

Whitney JS Study of Urban and Rural Tuberculosis Death Rates in New York State *Am J Public Health* 18(8), 978–984

NYT 9 Apr 1939

"Statistics came to her as easily as baseball". The Sheboygan Press. Wisconsin Friday, August 09, 1935

Atlanta Constitution 15 Feb 1914

12 Mar In March 1943 Abraham Wald (1902-1950) was presented with the problem of detecting anomalous production results for munitions. By Apr he developed the sequential probability ratio test SPRT. It was originally developed for military quality control studies. The US military thought it so important the paper was classified until the end of WWII. A similar test was developed independently by George Barnard (1946).

In 2003 SPRT was used by Prof Sir David Spiegelhalter and colleagues to show how it could have identified Harold Shipman as a serial killer much earlier than he was. At the Public Inquiry Prof Spiegelhalter testified that SPRT would have flagged up problematic deaths after ~40 victims by 1985. Shipman was arrested Sep 1998. He is thought to have killed >250 pts.

Medical applications of SPRT are complicated because they require case-mix risk adjustment so as not to unfairly penalise clinicians or hospitals for treating high-risk patients. However even simple graphics can show problems.

Refs

Wald A (1945) Sequential tests of statistical hypotheses. Ann Math Statist 16(2): 117–186. doi:10.1214/aoms/1177731118

Barnard GA (1946). Sequential tests in industrial statistics. J Royal Stat Soc, Suppl 8: 1–26.

Spiegelhalter D et al (2003). Int J Qual Health Care 15(1):7-13. doi: 10.1093/intqhc/15.1.7.

Tim Harford Catching a Killer Doctor | Cautionary Tales https://www.youtube.com/watch?v=Ad78IMHTfEw

Spiegelhalter D, The art of statistics. Pelican

Spiegelhalter D, Best N (2004) Significance 1(1) 10-12 https://dspiegel29.github.io/ArtofStatistics/10-3-4-Shipman-cumulative-frequency-and-SPRT/10-3-4-shipman-sprt-x.html

13 Mar DYK? Pierre-Simon Laplace (23 Mar 1749–5 Mar 1827) may have been the first to suggest that probability could be applied to medical data. In '*Essai Philosophique sur les Probabilités*' (1814) he says 'to discover the best treatment to use in curing a disease, it is sufficient to test each treatment on the same number of patients, while keeping all [other] circumstances perfectly similar. The superiority of the most beneficial treatment will become more and more

evident as this number is increased, and the calculus will yield the corresponding probability of its benefit and of the ratio by which it is greater than the others.'

Refs
Huth E (2008). *J Royal Soc Med*, 101: 205–212. DOI 10.1258/jrsm.2008.081008 2
www.jameslindlibrary.org

14 Mar. Laplace may have been the first to suggest application of 'probability calculus' to medicine, but so-called 'numeric methods' to compare efficacy of interventions were already in use. Pierre Charles-Alexandre Louis (1787–1872) assessed early versus delayed blood-letting on pneumonia-related mortality rates, demonstrating (1828) that early blood-letting (at least) was killing patients.

He argued that medical progress could not be made if each case was assessed individually. Instead, he promoted the importance of 'population' comparisons, and grouping of cases based on similarities in age, sex, diet, disease severity, and onset, forerunners to epidemiology and the modern clinical trial. He thought a sample size of 500 was enough for "certainty".

Refs.
Porter TM *The Rise of Statistical Thinking, 1820–1900*. Princeton
Morabia A (2006) *J Royal Soc Med* **99** (3): 158–160.

www.jameslindlibrary.org

15 Mar Louis-Dominique-Jules Gavarret (28 Jan 1809-30 Aug 1890) was the first (1840) to apply a forerunner of confidence intervals to clinical data, and described six principles for rigorous clinical trial conduct. Unfortunately, his book *Principes de Statistique Médicale* was almost forgotten by the end of the 19th century.

He first heard of 'le calcul des probabilités' in 1835 during a debate at the Académie des Sciences where Navier (of Navier-Stokes equation fame) outlined how it could be applied to therapeutic research (and incidentally demolished the debate opposition)

Using Louis' blood-letting data and the mathematics of "large numbers' he had learned as a student of Poisson, Gavarret calculated the range of possible errors of means ('limites d' oscillation') and argued against taking the point estimate at face value.

Louis reported 52 deaths out of 140 treated (0.37). Gavarret says "all that we have learned from the work of Louis, in reality, is that under the influence of the curative means used in his 140 observations the number of deaths must vary between 48 493 and 25 593 per 100 000 patients, or approximately between 49 and 26 per 100 patients.'

Gavarret's 6 principles for rigorous clinical trial conduct:

1. The patients have to be drawn exclusively in the same locality and from the same classes of the population.
2 The experienced illness has to have a precise diagnosis and perfect definition.

3. It has to be nosologically well delineated and separate from the illnesses resembling it most in this group.
4. The statistic within the makeup of the illness considered to be specific has to contain the precise indication of the number of cases within each of its varieties.
5. The medication tried has to be clearly formulated, as well as its main modifications for each of the varieties of the illness.
6. 'The medical statistician has to be competent.'

Refs
Huth E 2008. *J Royal Soc Med* 101:205-212
Beyneix A 2001. Le professeur Jules Gavarret (1809-1890) et l'application des méthodes mathématiques et physiques à la médecine. B*ull Académie nationale de médecine* 185(7): 1327-1335
Tröhler U 2020 *J Royal Soc Med* 113(9): 360–366 |

16 Mar OTD 1937 Amos Tversky b (d 2 Jun 1996) Israel. With long-time collaborator Danny Kahneman, they pioneered studies of decision-making, cognitive biases, risk assessment, and why people "struggle to think statistically". He revolutionized behavioural economics, which had in the past assumed decision-making was rational.

Their first joint paper "Belief in the Law of Small Numbers" described how people are not good "intuitive statisticians". People tend to assume small samples are highly representative of the population. Researcher sample size 'justifications' reflect "prevalent intuitive misconceptions of the extent of sampling variation". Gould (1992) says, "Tversky and Kahneman argue, correctly, I think, that our minds are not built (for whatever reason) to work by the rules of probability".

Because of later difficulties in replicating some key findings, Kahnman acknowledged that he himself "placed too much faith in underpowered studies".

Refs
Tversky A, Kahnemann D 1971 *Psychol Bull* 76(2): 105-110
Lewis M 2016. *The Undoing Project: A Friendship that Changed the World.* Penguin
Kahneman D 2011. *Thinking, Fast and Slow.*

17 Mar 68 yr ago this month (March 1955) David DB Duncan (16 Jun 1916-12 Jun 2006) published his multiple range test. It is one of the most cited statistical papers of all time. The test consists of pairwise comparisons between means using the studentized range distribution to determine critical values. It tends to protect against Type II error at the expense being too liberal with respect to Type I errors. It was developed as a modified version of his earlier 1951 multiple-F range. Duncan's tests ranked between the less conservative Fisher LSD and the more conservative Newman test.

Refs
Duncan DB (1955). Multiple range and multiple F tests. Biometrics, 11 (1): 1–42. doi:10.2307/3001478
Duncan DB (1977) Current Contents: Citation classics. Number 4, 24 Jan 1977, p 257.

18 Mar Before Doug Altman there was Marc Daniels (11 Jan 1907-3 Mar 1953). UK. An overlooked pioneer of clinical trial design, conduct, and notably,

reporting. His reporting guidelines were first published in 1951 for trials of anti-tuberculosis chemotherapy. However, it was not until 1995 that CONSORT was established as standard.

Doug Altman quoted Daniels at Evidence Live 2010*:* "Description of the procedures of the trial is indispensable; failure on this point leaves one in considerable doubt concerning the validity of some published work". Yet clinical trial reporting is still poor.

Refs

Daniels M (1951) Clinical evaluation of chemotherapy in tuberculosis. *British Medical Bulletin* 7:320-326.

https://www.jameslindlibrary.org/daniels-m-1951/

Crofton J (2005). Marc Daniels (1907-1953), a pioneer in establishing standards for clinical trial methods and reporting. *JLL Bulletin: Commentaries on the history of treatment evaluation* (https://www.jameslindlibrary.org/articles/marc-daniels-1907-1953-a-pioneer-in-establishing-standards-for-clinical-trial-methods-and-reporting/)

Obituary *British Medical Journal* 07 Mar 1953

19 Mar OTD 1936 Edgar Sydenstricker d (b 15 Jul 1881) USA. ASA Fellow 1922. An overlooked pioneer of health economics especially with respect to income-related health inequalities. Appointed first statistician for the US Public Health Service in 1915. He was one of the developers of the US National Health Survey. Documented impact of poverty and income disparity on health especially for marginalised groups: immigrants, garment workers, the rural poor. Notable reports included those on the 1914 influenza epidemic and pellagra.

King said he "came as close as a human being can to be being the perfect scientist." However, he is almost unknown to modern health economists. Clarke and Erreygers think his sudden death and relatively few collaborators prevented his contributions from being better known.

Fun fact: Born in China to missionary parents, his sister was the winner of the Nobel Prize in Literature (and civil rights advocate) Pearl S Buck

Refs

Clarke P, Erreygers G (2022) *Eur J Hist Econ Thought* 29(6): 1066-1088, doi: 10.1080/09672567.2022.21367312022.

Collier I. 2020 https://www.irwincollier.com/chicago-1907-08-economist-turned-epidemiologist-edgar-sydenstricker/

King WI 1936 [Obituary] *J Amer Stat Assoc*, 31(194): 411-414

20 Mar The first recorded statistical consult was probably 1654. Antoine Gombaud (1607–1684), self-styled Chevalier de Méré, wrote to Blaise Pascal with a probability problem. Except probability hadn't been invented yet. It is unknown if he thought the solution would "only take a minute".

Like other clients since then, de Méré actually wanted answers to several questions. Common games of chance in 18th c France were betting on the event of getting at least one 1 (ace) in four rolls of a dice or getting at least one double ace with 24 throws of a pair of dice. He thought they should be equiprobable, but he was losing money. He also posed the classic 'problem of

points': what is the most fair way of dividing the stake between two players when the game was interrupted before the end?

Pascal wrote to Pierre de Fermat and in a series of letters over the next year they developed the foundational concepts for the modern theory of probability. Fermat solved the problem by tabulating all possible continuations. Pascal improved on it of using expected values instead of odds, and manipulations of a triangular array of the binomial coefficients (Pascal's triangle). This was revolutionary as up to that time mathematicians thought it was impossible to predict the probability of an event.
de Méré later went swanning about claiming he had discovered probability theory himself.

Refs
Devlin K. 2010. *The Unfinished Game: Pascal, Fermat, and the Seventeenth-Century Letter that Made the World Modern: A Tale of How Mathematics is Really Done*
Letters: Fermat and Pascal on probability.
https://www.york.ac.uk/depts/maths/histstat/pascal.pdf

21 Mar OTD 1768 Jean-Baptiste Joseph Fourier b (d 16 May 1830) France. Pioneering physicist and mathematician. Best known to statisticians for the Fourier series, first presented in a paper to the Paris Institute 21 Dec 1807, 'Memoire sur la propagation de la chaleur'. He described an analytical theory of heat, modelled as a set of partial differential equations and solved by a trigonometric series (linear combinations of sines and cosines). Lagrange and Laplace were among the judges who didn't like it at all, complaining it lacked generality and rigour. As a result, Fourier did not publish his results until 1822. Even so he was still scooped by Gauss, who had extended work on trigonometric interpolation to periodic functions as early as 1805. This work was subsequently rediscovered by Colley and Tukey in 1965 as the fast Fourier transform, revolutionising digital computing

Refs
Fourier 1807 Memoire sur la propagation de la chaleur (On the Propagation of Heat in Solid Bodies) 1807.
Heideman M, Johnson D, Burrus C.1984. Gauss and the history of the fast Fourier transform" (PDF). IEEE ASSP Magazine. 1 (4): 14–21.
Colley JW 1988. Mikrochim. Acta [Wien], III, 33--45
doi:10.1109/MASSP.1984.1162257
Stigler SM. Gauss and the Invention of Least Squares," *Annals of Statist.,* 9(3), 1981, pp. 465–474.

22 Mar OTD 1973 Hilda Geiringer d (b 28 Sep 1893) Austria IMS Fellow. Recognised as "a first-rate scholar of great experience and accomplishment" in mathematical statistics, she is also a 'missed genius'. She was denied a permanent faculty position at US research universities because of her gender.

Her PhD thesis 1917 developed a generalised theory for a Fourier series in two variables. She was first woman lecturer 1927 in applied mathematics in Germany (at Humboldt University of Berlin), where she developed the Geiringer equations and slip-line theory for analysis of metals deformation. She fled Nazi Germany for Istanbul 1934 then went to the US 1938 when it was no longer safe.

However, US immigration quotas had been met. Non-quota visas reserved for scientists and scholars often excluded women because - Catch-22 - women were unable to get professor positions that were usually given to men. She had worked closely with Richard von Mises, who had already obtained a position at Harvard (and therefore his visa). He worked frantically to find her a position at a women's college in the USA so she could get out of Istanbul (she faced almost certain death if deported back to Germany).

She eventually obtained an unpaid position at Bryn Mawr. Later she was lecturer in a number of women's colleges and managed to do some research at Wheaton. She obtained a part-time job at Harvard, completing and editing unfinished works of von Mises whom she married in 1943.

A typical response to her job applications: "I am sure that our President would not approve of a woman. We have some women on our staff, so it is not merely prejudice against women, yet it is partly that, for we do not want to bring in more if we can get men." In 1942, she gave an advanced summer course on the geometrical foundations of mechanics at Brown, who profited from the sale of her lecture notes, although never offering her employment (or an apology apparently)

Refs
Mcneill L The woman who reshaped maths.
https://bbc.com/future/article/20191031-hilda-geiringer-mathematician-who-fled-the-nazis
Leff L https://northeastern.edu/refugeescholars/about

23 Mar OTD Pierre-Simon Laplace b (d 05 Mar 1927) FR. Called 'one of the greatest scientists of all time', his prodigious mathematical talent alienated many of his contemporaries, mostly because he told them he was best mathematician in France, which in fairness he was.

His primary interests in celestial mechanics led him to develop the Bayesian interpretation of probability, and a precursor of least-squares (he minimised the maximum error not the sum of squared errors)

His 'Théorie analytique des probabilités' (1812) discusses probability distributions, compound probabilities, and Buffon's Needle Problem. In 'Essai philosophique sur les probabilités' (1814) he sets out the principles of Bayesian probability

Refs
Stigler S (1986) History of Statistics. Belknap
https://www.maa.org/press/periodicals/convergence/mathematical-treasures-pierre-simon-laplace-on-probability

24 Mar In 1936 RA Fisher started using the symbol μ instead of *m* to indicate the mean of the normal distribution in *Statistical Methods for Research Workers* 6th edition. Surprisingly μ was not in standard use until then. However, he had adopted Pearson's use of $\bar{x}$ for the sample mean back in 1912. $\bar{x}$ is apparently left over from the time when physicists (such as Clark Maxwell and Thomson and Tait) indicated averages with a bar. Karl Pearson was trained in physics, so he used the bar for both sample means and expected values (later replaced by E).

Refs
Maxwell 1867. On the Dynamical Theory of Gases *Philosophical Transactions of the Royal Society*, 157: 64
Thomson W Tait PG 1879. *Treatise on Natural Philosophy*
Fisher 1912 On an Absolute Criterion for Fitting Frequency Curves. *Statistical Science* 12(1) 39-41 (1997)

25 Mar Origins of 'degrees of freedom' can be a bit confusing to trace b/c nobody explicitly called them that until RA Fisher in 1922. The concept first showed up in 1821 in Gauss (of course) "Theoria combinationis observationum erroribus minimis obnoxiae".

In letters to Karl Pearson (1893), Francis Edgeworth explained it in terms of geometric dimensions and explicitly corrected for restrictions on the sum of counts. They were discussing the problem of goodness of fit tests for multinomial distributions.

In 1922 Fisher introduced the term in connection with Pearson's landmark χ2 test. Pearson had erroneously claimed no correction was needed when parameters were estimated under the null hypothesis. Based on his new conception of parametric families Fisher was able to correct it and properly define df, and a feud was born

Refs
Fisher RA. 1922. On the interpretation of χ2 from contingency tables, and the calculation of P. *J Roy Stat Soc*. 85: 87–94
Stigler S. 1986. *History of Statistics*; Stigler S. 2008 *Stat Sci* 23(2): 261–27

26 Mar J J Sylvester (3 Sep 1814 – 15 Mar 1897) UK. FRS. He once said that he could "*without immodesty lay claim to the appellation of Mathematical Adam, as I believe that I have given more names (passed into general circulation) of the creatures of mathematical reason than all the other mathematicians of the age combined"*.

He was certainly not wrong. In the course of important work on matrix theory, he coined the terms 'matrix' 'discriminant' 'canonical form', "hessian' 'latent roots', "Jacobian' 'linear independence', sequence'. It was perhaps inevitable that some didn't stick to the wall – these included catalecticant, meiocatalecticizant.

He was also a pioneer of graph theory. In an 1878 note in *Nature* he used the word 'graph' (in the sense of graph theory) for the first time, and coined the terms X-, Y-, and Z-coordinates.

Refs
Sylvester JJ. The collected mathematical papers 1814-1897. 1888 *Nature* 37, 152. (1851) Phil Mag, series 2, (1851), 391-410.
Wilson R, Watkins JJ, Parks DJ. 2023 Graph Theory in America: The First Hundred Years
James IM. 1997. Notes and Records of the Royal Society of London 51(2): 247-261
https://ia804505.us.archive.org/29/items/collectedmathema04sylvuoft/collectedmathema04sylvuoft.pdf

Biggs, Lloyd and Wilson *Graph Theory* 1736-1936

27 Mar OTD 2015 Janet L. Norwood d (b 11 Dec 1923) USA. ASA Fellow 1974. Honorary Fellow Royal Statistical Society, ASA President 1989 (6th woman), Vice-president ISI. First woman Commissioner of the US Bureau of Labor Statistics , served under two US presidents ,and was confirmed by the US Senate several times. Awarded Presidential Rank as Distinguished Executive in the US Senior Executive Service. She advocated the benefits of social and behavioural science on economic indicators, creating a 'cognitive laboratory' within BLS to increase its efficiency and the quality of its output. She had a reputation for being "objective, methodical, unflappable under sometimes hostile congressional questioning" and "helped bring recognition to female presence and leadership" for numerous government agencies and professional organizations
The Janet L. Norwood Award for Outstanding Achievement by a Woman in the Statistical Sciences is offered by the University of Alabama and ASA.

Refs

https://magazine.amstat.org/blog/2016/12/08/sih-norwood/
https://www.nytimes.com/2015/04/01/business/janet-norwood-dies-at-91-led-labor-statistics-bureau.html
https://www.bls.gov/bls/history/commissioners/norwood.htm
Fienberg SE. A conversation with Janet L Norwood, https://www.stat.cmu.edu/tr/tr603/tr603.html

28 Mar OTD 1809 Carl Friedrich Gauss (1777–1855) finishes '*Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium*'. Of interest to statisticians is how he showed derivations of measurement error distributions led to the method of least squares.

However, the French mathematician Adrien-Marie Legendre (1752–183) claimed precedence. An appendix to his 1805 '*Nouvelles Méthodes pour la Détermination des Orbites des Comètes' "Sur la Méthode des moindres quarrés*" was the first clear statement of the method. He was incensed when Gauss first referred to it as * OUR * method and then went about saying he had been using it since 1795 when he was 18.

In 1812 Laplace wrote about it in his 'Théorie Analytique des Probabilités' and showed how the least squares estimate has the smallest expected error.

Refs

Stigler S. 1986. The History of Statistics: The Measurement of Uncertainty before 1900.
Ruger HA Walker H (Transl) Legendre on least squares, https://york.ac.uk/depts/maths/histstat/legendre.pdf

29 Mar OTD 1983 Maurice Kendall d (b 6 Sep 1907) UK. FBA KBE Guy Silver 1945 and Gold 1968 Medals, ASA Fellow 1950, 1968 Honorary President ISI; President Royal Statistical Society 1961. Known for his tau and rank correlation coefficient, and contributions to k-statistics, cluster analysis, factor analysis, and time series analysis. In 1939 Kendall and Babington-Smith built one of the first early mechanical devices to produce random digits. It was a rotating disc divided into 10 sections for the digits, which was randomly illuminated by an electric spark in a dark room. They also formulated a series of tests for statistical randomness.

Quote: *If they would only do as he [Thomas Bayes] did and publish posthumously we should all be saved a lot of trouble.*

Refs

Stuart A 1984 obituary Sir Maurice Kendall, 1907-1983 J Royal Stat Soc. Series A 147(1) : 120-122

Herbert A David and Wayne A Fuller (2007) Sir Maurice Kendall (1907–1983), Amer Stat 61(1): 41-46, DOI: 10.1198/000313007X169055

Keith Ord (1984) In Memoriam: Maurice George Kendall, Amer Stat 38(1): 36-37, DOI: 10.1080/00031305.1984.10482870

Bartholomew DJ The Statistician 32 (1983) 445-446

30 Mar Before Rothamsted there was Jethro Tull. Yes. Baptised OTD 1674 (d 21 Feb 1741, earliest proponent of a scientific empirical approach to agricultural research and helped bring about 18th c British Agricultural Revolution. He was heavily criticised in his lifetime for his revolutionary rejection of the poetry of Virgil as scientific authority in favour of actual experiments.

Fun fact. The British rock band Jethro Tull was given their name by a booking agent who clearly knew his English agricultural history. According to Ian Anderson, they not particularly good when they first started and so had trouble getting club bookings. As a way around that they pretended to be somebody different every week. The time they got invited back to play a return engagement meant they had to stay with the name they already had that week, which was Jethro Tull.

Refs

Porter R, ed. 2003. *The Cambridge History of Science, 4: The eighteenth century*

Warkentin BP. 2000. *Can. J. Soil Sci*. 80: 391–393.

Sayre L. 2010. *Physics and Chemistry of the Earth* 35(15):851-9

31 Mar OTD 1971 William J Youden d (b 12 Apr 1900) Australia/USA. ASA Fellow 1951; Honorary Fellow Royal Statistical Society 1965, ASA Wilks Award 1969. Best known his J statistic for summarising the performance of diagnostic tests, he developed innovative experimental designs (e.g. Youden square), restricted randomisation, and new methods of analysis. He was an effective communicator of DOE and statistical analysis to non-statistician researchers.

Each year ASA awards the *W. J. Youden Award in Interlaboratory Testing* to authors "of publications that make outstanding contributions to the design and/or analysis of inter-laboratory tests or describe ingenious approaches to the planning and evaluation of data from such tests."

Fun Facts: Youden communicated his new designs to RA Fisher in 1936. Fisher called the new experiment designs 'Youden squares' in his work on *Statistical Tables* published jointly with Frank Yates in 1938. Youden squares played an important role in World War II being used for experimental trials in engineering and other scientific projects.

Refs

J Quality Technology, 1972 491) 3-6

Eisenhart C, Rosenblatt JR 1972 Annals Math Stat 43 (34) :1035-40

Preece DA 1990 Fifty years of Youden squares: A review. *Bull. Inst. Math. Appl.* **26** (4):65-75.

The HOS-ASA Twitter Diaries: APRIL 2023

01 April DYK? A brilliant 13th c Benedictine monk Udo of Aachen developed early basics of probability theory, an empirical geometric probability method for estimating pi (500 years earlier than de Buffon's similar needle problem, 1777) and even a forerunner of the Mandelbrot set, However…

JUST KIDDING, this is an APRIL FOOL'S JOKE. Both Udo and the research of Prof Schipke of Harvard originated in the fertile brain of writer Ray Girvan in 1999. If nothing else, it shows why we should be sceptical of ChatGPT: the Schipke and Eberhardt reference doesn't exist either.

Refs

Girven R 1999 The Mandelbrot Monk
https://users.math.yale.edu/public_html/People/frame/Fractals/MandelSet/MandelMonk/MandelMonk.html

Schipke RJ, Eberhardt A. 1999. The forgotten genius of Udo von Aachen. *Harvard J Historical Mathematics* 32:34-77

02 April Marie-Sophie Germain (01 Apr 1776- 27 Jun 1831) France. First woman to win the Paris Académie des Sciences Grand Prize 1816 for work on elasticity theory. Under the male pseudonym Antoine-August LeBlanc she corresponded with Legendre on number theory (some of her work is included in the *Supplement to 'Théorie des Nombres')* and with Gauss on Fermat's Last Theorem. First woman allowed to attend lectures at the Institut de France.

The Prix Sophie Germain, awarded annually by the Foundation Sophie Germain, is conferred by the Académie des Sciences to honour a French mathematician for mathematics research Her parents originally opposed her math studies to the extent of taking away candles, fire, and her clothes to prevent her studying. Gauss recommended that she be awarded an honorary degree, but she died of breast cancer before it could happen.

Refs

Singh, Simon (1997). Math's Hidden Woman
https://www.pbs.org/wgbh/nova/physics/sophie-germain.html
Cipra BA 2008 A Woman Who Counted. *Science* 319 (5865): 899

03 April In April 1901 the first use of 'bimodal' was used by RW Strong in a study of morphometric variation in shrikes. The study was a test of the hypothesis that geographic differentiation in measured traits was a mechanism for speciation. He used frequency graphs, correlations, and CB Davenport's "Precise Criterion", the hot statistical methods at the time.

In 1895 Karl Pearson had coined the term 'mode' and introduced both correlation and frequency distributions as practical methods of analysis. Davenport introduced Pearson's methods to the USA in his 1899 text '*Statistical methods with special reference to biological variation'*. There were a lot of mistakes in the first edition.

04 April From Richard Coles: "It is the centenary of the death of the cleric and mathematician John Venn, inventor of the Venn Diagram, who resigned his Orders in 1883 and built a cricket ball bowling machine which bowled out the Australian first XI in 1909. He is remembered thus at his Cambridge college."

@HoS_ASA can't top that but here is another less well-known statistician: OTD 1951 Charles P Winsor d (b 19 Jun 1895) USA. ASA Fellow 1949. First VP of ENAR. Best known for developing 'winsorization', the 'clipping' of outliers by setting them to pre-specified data percentiles. He also did important work on small samples, Gompertz growth curves, and linear regression with error in both variables.

Refs

Cochran WG 1951. Human Biology 23(2):73.
Biometrics 7 (2): 221. June 1951.

05 April In April 1959 Nathan Mantel and William Haenzel (USA) published a method for effective use of heterogeneous retrospective study data. Now one of the top-cited statistics papers of all time, it has been described as "one of the major contributions to both statistics and epidemiology" and was one of the earliest examples of statistically-based analytical, rather than descriptive, epidemiology. The Mantel-Haenszel test is a staple for analysing stratified or matched categorical data.

Refs

Mantel N, Haenszel W. 1959. Statistical aspects of the analysis of data from retrospective studies of disease. *J Nat Cancer Inst* 22:719-48
This Week's Citation Classic 26, 29 Jun 1981
Hankey B 1997*. Stat Sci* 12(2):108–112

06 April. The development of complex numbers was ….complex. Imaginary numbers were introduced as a method for solving cubic equations as early as 1530 by Nicolo Tartaglia, then ~~stolen~~ appropriated by Girolamo Cardano in *Ars Magna* 1539. Rafael Bombelli in *L'Algebra* represented $\sqrt{-1}$ as pdm, short for 'piu di meno' ("plus of minus"). His terminology quickly got a bit convoluted; $i^2 = -1$ was called "plus of minus times minus of minus makes minus". Rene Descartes 1637 is credited with first introducing the term 'imaginary': "...tant les vrayes racines que les fausses ne sont pas tousiours réelles; mais quelquefois seulement imaginaires." A 1668 book review (yes, really) of James Gregorie's *Geometriæ pars Universalis* was the first to use the term in English.

Refs

Bagni GT. 2009. Bombelli's Algebra (1572) and a new mathematical object. *For the Learning of Mathematics* 29 (2): 29-31
1668 An Account of Fome Books, *Philosophical Transactions* 3(35)

07 April OTD 1761 Thomas Bayes d (b c. 1701) FRS Best known for his eponymous theorem, he in fact never published it; his notes were edited posthumously by his friend Richard Price 1761. A solution to a problem of inverse probability was presented in *An Essay towards solving a Problem in the Doctrine of Chances*, published 01 Jan 1763. Much of his early work was related to infinite series and numerical analysis. He never published those either, and similar results are usually attributed to Lagrange 1772, 1792.

Bayes probably studied mathematics with James Gregory (or Gregorie; spelling, even of people's own names, tended to be erratic) at Edinburgh 1720-1. He may have become interested in probability either directly from de Moivre who was in London by then, or after reading Simpson 1755. On 8 Apr 1742 Bayes was

nominated to the Royal Society based on his "known merit, well skilled in Geometry and all parts of Mathematical and Philosophical Learning".

Fun fact: Birth dates for Bayes and his siblings are unknown. This was because Bayes was Presbyterian, and therefore a religious nonconformist, or Dissenter. Birth and baptismal registers for Dissenter churches were not considered legal documents and were either hidden or not kept at all as a result of active religious discrimination.

Refs

Bayes 1763 An Essay towards solving a Problem in the Doctrine of Chances. *Philosophical Transactions*. 53: 370–418.

Bellhouse DR 2004. Statistical Science 19(1):3–43

Stigler SM 1986 *The History of Statistics: The Measurement of Uncertainty Before 1900*. Belknap

08 Apr OTD 1964 George Box and David R Cox read their paper "An Analysis of Transformations" at a research methods meeting of the Royal Statistical Society. The paper began with their names, then found the topic.

In the UK the phrase 'Box and Cox' refers to "people who always miss each other andthus are never together". The original Box and Cox was a one-act comic play by JM Morton (premiered 1847) about two men tricked by their landlady into occupying the same room without being aware of each other's existence. (There is also 1867 musical version by Burnand and Sullivan). Box said it was obvious that any paper he and Cox would write would have to be on transformations.

Refs

Box GEP, Cox DR. 1964 *J Royal Stat Soc Series B* 26: 211-243

Sakia RM 1992 *J Royal Stat Soc Series D* 41: 169-178

deGroot M 1987. *Statist Sci* 2(3): 239-258

Phrase origin https://wordhistories.net/2017/07/10/box-and-cox-origin/

09 Apr OTD 1810 Pierre-Simon de Laplace (23 Mar 1749-5 Mar 1827) reads his paper on the Central (aka "Fundamental") Limit Theorem to the Académie de Sciences. It took him about 40 yrs to develop. It was devised as a tool to solve other mathematical problems, such as "the comet problem", method of least squares, and risk in games of chance. Laplace did not explicitly state a theorem - the Central Limit Theorem was precisely formulated in modern times in the 1920s by von Mises, Pólya, Lindeberg among others.

Refs

Laplace 1810. Mémoire sur les approximations des formules qui sont fonctions des très grands nombreset sur leur application aux probabilités"

Fischer H 2010 *A History Of The Central Limit Theorem: From Classical To Modern Probability Theory*. Springer

10 Apr OTD 1755 Thomas Simpson (20 Aug 1710-14 May 1761, England) read his paper on error distributions to the Royal Society: "On the Advantage of taking the Mean of a Number of Obfervations in Practical Aftronomy". Considered a milestone in statistical inference, it focused on observation error rather than the observations themselves and was the earliest demonstration of the law of large numbers by attempting to show that the arithmetic mean is nearer the 'truth' than any single observation.

Stigler says it was the "critical step… to an applicable quantification of uncertainty". However, Karl Pearson, no fan, said it was "nothing but a boiling down of De Moivre's *Doctrine of Chances"* and sniffed that he was "an unpleasant and truculent writer of cheap textbooks, not a great mathematician like De Moivre". Simpson was elected a Fellow of both the Royal Society 1745 and the Royal Swedish Academy of Sciences 1758, so not everyone shared Pearson's opinion.

He also developed the eponymous Simpson's rule for approximation of definite integrals. The attribution, as often in mathematics, can be debated: this rule had been 'found' 100 years earlier by Johannes Kepler - in German it is called Keplersche Fassregel.

Simpson was a bit of a scandal in his time. At the age of 19, he married his landlady, a 50 yr old widow with 2 children. A report that he had frightened a girl into fits by 'raising the devil' (apparently, he or an assistant had frightened her by dressing up as a devil during an astrology session) meant that he and his family had to skip town in a hurry.

Under a variety of pseudonyms such as Marmaduke Hodgson, Hurlothrumbo, Kubernetes, Patrick O'Cavannah, and Anthony Shallow (I know right?), he answered mathematical problems posed in the *Ladies Diary*, the magazine he edited. It was his obvious mathematical skills demonstrated in these solutions which first brought him to the attention of other mathematicians of the day. Other periodicals to which he contributed were the *Gentleman's Magazine, Miscellanea Curiosa Mathematica*, and the *Gentleman's Diary*.

Refs
Simpson *Philosophical Transactions, Royal Society of London* 49(49) 1755
Stigler SM 1986 *The History of Statistics: The Measurement of Uncertainty Before 1900*. Belknap
Shoesmith E 1985 Thomas Simpson and the arithmetic mean, *Historia Mathematica,* 12(4): 352-355, doi.org/10.1016/0315-0860(85)90044-8.

11 Apr OTD 1907 Henry Scheffé b (d 5 Jun 1977) USA. ASA Fellow 1951, ASA Vice-President 1954-6, President Institute of Mathematical Statistics. Best known for Scheffé's correction for multiple comparisons, and optimal tests for equality of two normal variances. He also did important work on quality control, mixture designs, and nonparametric statistics.

Fun fact. He switched to statistics when he found out that some of the results in his PhD dissertation he thought were original had actually been proved by Gauss. He first studied electrical engineering at the Polytechnic Institute of Brooklyn. His only bad grade was D in mechanical engineering which was later changed to an E.

Refs
Daniel C, Lehmann EL 1979 *Annals Stat,* 7(6) 1149-61

12 Apr OTD 1878 Kirstine Smith b (d 11 Nov 1939). Denmark. Considered by Karl Pearson to be his most brilliant student. As a mathematical statistician, she invented the field of optimal designs. Her PhD dissertation described G-optimal designs for polynomial regression (1918).

It was her 1916 paper on method of moments for minimising the $\chi 2$ statistic as a measure of goodness of fit that sparked what evolved in a major row between Pearson and Fisher, which lasted in the published literature for decades. Fisher promoted the "Gaussian method" (what he would later call maximum likelihood) over method of moments. Pearson defended her against Fisher to the end of his life (1936). Smith apparently didn't care; she returned to Denmark to teach high school.

Refs

Smith K 1916. On the "best" values of the constants in frequency distributions. *Biometrika* 11 262–276

Smith K 1918. On the standard deviations of adjusted and interpolated values of an observed polynomial function and its constants, and the guidance they give towards a proper choice of the distribution of observations. *Biometrika* 12, 1-85

Pearson K 1936. Method of Moments and Method of Maximum Likelihood *Biometrika* 28(1/2) 34-*59*

Stigler S 2005 Fisher in 1921. *Stat Sci* 20(1): 32-49

Guttorp P, Lindgren G. 2009. Karl Pearson and the Scandinavian school of statistics. *Intern Stat Review*, 77: 64.

13 Apr 1912 Herman Otto (HO) Hartley b (d 30 Dec 1980) Germany/USA. ASA Fellow 1953, ASA President 1979. Founder Texas AandM University Department of Statistics. A long-time collaborator with Egon Pearson, they were the first to use the terms "Type I and Type II errors" in 1933. Before that, they used "the first and second source of error" (*Biometrika* 20A, 1928) and "Errors of first and second kind" (*Phil Trans Royal Soc Lond* A, 1933), but apparently didn't like them much.

Fun fact: Hartley was short, and Pearson was very tall so Hartley's classic introduction was "Never were there two more appropriate statisticians to work on the concept of range statistics."

Refs

Hartley HO 1980. Statistics as a science and a profession. *J Amer Stat Assoc* 75, 1-6.

David HA 1982. *Intern Stat Review* 50(3), 327-30

Hartley HO, Pearson ES 1933. *Proc Cambr Philos Soc* 24:492-510

@AmStatNews https://magazine.amstat.org/wp-content/uploads/2017/08/hartleyho.pdf

14 Apr 1629 Christiaan Huygens b (d 8 Jul 1695) Called the "greatest scientist in Europe" his 14-page pamphlet *De Ratiociniis in Ludo Aleae* (published 19 Apr 1657) was the most important work on probability for the rest of the century. It develops the beginning of game theory, describes the problem of points, and develops expected values as generalization of the weighted average.

> *That any one Chance or Expectation to win any thing is worth just such a Sum, as wou'd procure in the same Chance and Expectation at a fair Lay. ... If I expect a or b, and have an equal chance of gaining them, my Expectation is worth (a+b)/2.*

Blaise Pascal's *Traité du triangle arithmétique* (1665) was more comprehensive. However, because Huygens' work was written in Latin, it was more popular among European mathematicians. Franz van Schooten had translated it from the original Dutch and then published it in own book *Exercitationum Mathematicarum* (1657) giving it even more traction. It received yet more attention with further translations (notably the translation into English in 1692 by John Arbuthnot) and commentaries.

Huygens became interested in probability and games of chance after reading Fermat and Pascal. But he didn't stop there: he founded the study of mathematical physics and was a prolific inventor. Among many ground-breaking inventions he developed the pendulum clock (the most accurate timekeeper until the 1930s) and the balance- spring pocket watch (although Hooke may have had precedence)

Refs

Transl. *de Ratiociniis In Ludo Aleae*
https://math.dartmouth.edu/~doyle/docs/huygens/huygens.pdf

Stigler SM 2007. Chance Is 350 Years Old. *Chance*, 20 (4): 26–30. doi:10.1080/09332480.2007.10722870

Christiann Huygens
http://www.proevenvanvroeger.nl/eindopdrachten/huygens/huygensfamily.pdf

Heyde CC et al (eds) 2001 *Statisticians of the Centuries* Springer

15 Apr OTD 1707 Leonhard Euler b (d 18 Sep 1783). Laplace called him the "Master of us all". Immensely prolific (>850 publications, plus 13 children), he created the whole form of modern mathematical thinking by using and inventing consistent terminology and notation such as the familiar summation Σ, Euler's number e = 2.71828, i for $\sqrt{-1}$, the concept of functions f(x), beta and gamma functions, use of exponential functions and logarithms in analytic proofs, power series solutions of differential equations,…..

He also pioneered diverse fields such as graph theory and topology and did important work in astronomy, fluid mechanics, music theory, analytic geometry, trigonometry, calculus, and number theory.

Fun fact. On the day of his death, he gave a mathematics lesson to one of his grandchildren, calculated the motion of balloons, then chatted with friends about the recently discovered planet Uranus.

Also, OTD 1821 Carl Friedrich Gauss (1777-1855) reads part I of what was later called the Gauss-Markov theorem on complex numbers. Most 17th and 18th century writers spoke of a + bi as an imaginary quantity. Carl Friedrich Gauss (1777-1855) saw the desirability of having different names for ai and a + bi, so he gave to the latter the Latin expression *numeros integros complexos*.

16 Apr OTD 1894 Jerzy Neyman b (d 5 Aug 1981). Poland/USA. ASA Fellow 1942; FRS, Guy Gold Medal 1966; ASA Wilks Award 1968; US National Medal of Science 1969. Best known for developing modern concepts of statistical hypothesis testing, estimation, sampling, and design of experiments. Introduced the confidence interval in 1937. With Egon Pearson developed the 'hypothesis

testing' paradigm (1933) contrasting to Fisher's 'significance testing' of the 'null' (much to Fisher's often-expressed ire).

In 1925 he joined Karl Pearson's team on a fellowship but was disappointed to find the work done was 'old-fashioned 'and Pearson himself was 'surprisingly ignorant of modern mathematics'. He went to Paris but returned to London in 1934 after KP retired. Much of his joint work with Egon Pearson throughout this time (1928-33) had been carried out by letter.

Fisher first welcomed Neyman as an ally in the effort "to free statistics from unwarranted Bayesian assumptions". However, to say relations deteriorated was an understatement. Neyman later said that Fisher's 'unique style involving torrents of derogatory remarks', insults, and polemics, represents an unusual phenomenon, probably unparalleled in the history of science". He spent the rest of his life in Berkeley where he established a major statistical center.

Refs

Neyman J 1937. *Phil Trans Royal Soc London. Series A*, 236 (767): 333–380

Kendall DG et al. 1982. Biographical Memoirs of Fellows of the Royal Society. 28: 379–412. doi:10.1098/rsbm.1982.0015

Lehmann EL 1994 Dictionary of Scientific Biography

Neyman 1961 Silver Jubilee of My Dispute with Fisher. *J Operation Research* Soc Japan, 145-54)

17 Apr In April 1912, as a third-year undergraduate student at Cambridge, RA Fisher published his first mathematical paper 'On an Absolute Criterion for Fitting Frequency Curves' This was the first iteration of what would later become the method of maximum likelihood. He finally settled on the name "maximum likelihood" in 1922. Derived from the theory of errors taught by his tutor the astronomer FJM Stratton, it features a critique of the method of moments favoured by Pearson.

Refs

Fisher RA. 1912. *Messenger of Mathematics* 41, 155-160

Aldrich J. 1997. *Statistical Science*, 12(3) 162-176

18 Apr OTD 1991 Sir Austin Bradford Hill d (b 8 Jul 1897) UK. CBE FRS President Royal Statistical Society 1950, Guy Gold Medal 1953. Peter Armitage called him the "world's leading medical statistician". Possibly best known for establishing the causal link between smoking and lung cancer, he pioneered both the modern randomised controlled trial (pertussis, streptomycin for TB), and the "Bradford Hill" criteria for determining causal associations in observational epidemiological data.

His 1937 series of papers 'Principles of Medical Statistics ' in *The Lancet* was the forerunner of the highly popular *BMJ Statistics Notes* by Martin Bland and Doug Altman almost 50 years later. They were written for clinicians in plain language with sparing use of equations. They proved so popular they were printed as a book, running to 12 editions.

Refs

Doll R 2003. Fisher and Bradford Hill: Their personal impact. *Intern J Epidem,* 32 (6): 929–931, discussion 931–8. doi:10.1093/ije/dyg287; Discussion pp 945-8.

Armitage P 1991 *J Royal Stat Soc A* 154(3): 482–484,
Farewell V, Johnson A 2012. *J Royal Soc Med* 105(11): 483-489

19 Apr OTD 1922 Ronald Fisher published The One Ring to Rule Them All. He first read his paper "On the mathematical foundations of theoretical statistics" to the Royal Statistical Society on 17 Nov 1921. This paper marked a seismic shift in statistics, which up until then lacked a unifying structure. In this ground-breaking paper he explicitly distinguishes between a parameter and a statistic, introduces the major criteria of estimation (consistency, efficiency, sufficiency), and defines concepts such as validity, likelihood (first presented when he was an undergrad), and information.

Refs

Fisher RA 1922. *Phil Trans Royal Soc. A* 19 Apr 1922: 309-368
Hand DJ 2015. *Philos Trans A Math Phys Eng Sci,* 373(2039): 20140252.
Efron B. 1998. R. A. Fisher in the 21st century. *Stat. Sci*. 13, 95–122.

20 Apr OTD 1923 Jacob Cohen b (d 20 Jan 1998) USA. ASA Fellow 1983 Best known for Cohen's kappa, d, and h. He was an early advocate of power and effect size and a critic of the misapplication of null hypothesis significance testing. Less well known is his promotion of data visualisation methods to teach introductory statistics.

In 1990 he summarised what he considered the basic principles of applied statistics: "less is more", "simple is better" (graphic representation), "some things you learn aren't so"; avoid the many misconceptions of null hypothesis testing; the importance of power analysis; and the importance of determining "just how big (rather than how statistically significant) are the effects that we study".

Refs

Cohen J 1990. *Amer Psychol*, 45: 1304–1312
Cohen J 1994. The earth is round (p <.05). *Amer Psychol* 49(12):997-1003.

21 April OTD 1942 Maurice Kendall (1907-1983) UK read his paper 'On The Future of Statistics' before the Royal Statistical Society. He begins with the startling observation that statisticians have overrun every branch of science like 'Colorado potato beetles'. He predicted new research directions, how statistical teaching was likely to evolve over the next several decades, and suggested the UK adopt the metric system.

He opined that future trends would see more expansion into non-science fields, increased specialisation, increased theoretical development, and consolidation and extension of existing knowledge and methods. New research directions were to include biased sampling, 'statistical machines', probability theory, and estimation. He suggested that RSS should enlarge its scope to serve in an advisory capacity to government agencies and made another plea for the UK to adopt the metric system.

Ref

Kendal MG 1942. *J Royal Stat Soc* 105 (2), 69-91

22 Apr OTD 1891 Sir Harold Jeffreys b (d 18 Mar 1989) UK. FRS, Guy Gold medal 1962. Best known for developing Bayesian theory of probability (e.g.

Jeffrey's prior). It failed to gain traction for years both because he was a terrible speaker and because of the overwhelming dominance of the Fisher-Neyman-Pearson school. Fisher flamed his book *The Theory of Probability*: "He makes a logical mistake on the first page [Fisher was referring to Bayes' postulates] which invalidates all the 395 formulae in his book".

To the day of his death, Jeffreys strongly opposed the theory of continental drift. Fisher said "He was wrong about probability and that makes me sure he is wrong about continental drift". However, another story related by Joan Fisher Box has it that both Fisher and Jeffreys attended a lecture by Eddington and were so 'horrified' by what he had to say about inference that they "promised not to write any more rude things about each other"

Refs
Cook AH 1990 *Biogr. Mems Fellows Royal Soc*, 36, 301–333
Box JF 1978. *RA Fisher: The Life of a Scientist*

23 Apr Georges-Louis Leclerc, Comte de Buffon (7 Sep 1707 – 16 Apr 1788) France. He is of interest to statisticians because he introduced differential and integral calculus into probability theory, and his 'needle problem' was an early example of Monte Carlo simulation principles. D'Alembert, who detested him, called him "the great phrase-monger" (take that Buffon!) in his *Sur le jeu de franc-carreau* (On the game of fair-square).

The Needle Problem: 'Suppose we have a floor made of parallel strips of wood, each the same width t, and we drop a needle of length L onto the floor. What is the probability that the needle will lie across a line between two strips?' The solution is $p = 2L/\pi t$. Monte Carlo methods were used to approximate π by binomial probability sampling

Buffon is better known as a naturalist, and pioneer of important concepts in ecology and evolution (his work was a major influence on Darwin). His revolutionary theories of geological history, reproduction, and speciation got him in hot water with church authorities because his writings contradicted scriptural versions of creation. Buffon published a retraction, and said he would make changes, but he never bothered.

Fun fact: During the French Revolution, his tomb was destroyed when fighters removed the lead from his coffin to produce bullets. His heart was initially saved, but later lost. His cerebellum is the only bit left. It is now housed in the base of a statue at the Muséum National d'Histoire Naturelle (MNHN)

24 Apr OTD 1919 David Blackwell b (d 8 Jul 2010) USA. ASA Fellow 1962, ASA Vice President 1978, National Medal of Science (posthumous) 2014. First African American to be elected to ASA and National Academy of Sciences. Honorary Fellow Royal Statistical Society. Best known for Rao–Blackwell theorem and pioneering text on Bayesian statistics, and his substantial and ground-breaking contributions to game theory, probability theory, information theory, and mathematical statistics.

Lauded for his "wide ranging genius" and rare "combination of exceptional scholarship and superb teaching" "everlasting exemplar of courage,

perseverance and brilliance", in the face of tremendous systemic racial prejudice.

Refs

Roussas GG et al. 2011 A Tribute to David Blackwell, *NAMS* 58(7), 912–928.

Agwu N et al 2003. *Mathematics Magazine* 76 (1): 3–14.

25 Apr OTD 1879 Edward Bidwell Wilson b (d 28 Dec 1964) USA. ASA Fellow 1924, ASA President 1929, Fellow RSS. Best known for the Wilson score interval for binary variables (1927) an improvement over the normal approximation method. His formulation anticipates the concept of a confidence interval, now attributed to Neyman and Pearson. In 1952, Neyman acknowledged that Wilson did have the idea in 1927, but Wilson himself said in 1964 'I would make no such claim'.

Refs

Hunsaker J, Maclane S 1973, https://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/wilson-edwin-b.pdf

Wilson 1927. Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc*, 22:209-12.

26 Apr OTD 1898 William O Kermack b (d 20 Jul 1970) Scotland. Best known for the Kermack-McKendrick mathematical theory of epidemic disease spread (1927), the origin of SIRS epidemiological models.

In spite of what he called their limited 'statistical armoury', Frank Yates commended the three papers for 'the clarity and elegance of the mathematical analysis,… careful exposition of the conclusions.., detailed investigation of special cases, and warnings on the practical limitations of the results'.

Kermack's first scientific work in 1918 was a statistical analysis of dairy cattle milk yield with JF Tocher. He said 'This introduced me to "Biometrika" and to the work of K. Pearson and "Student"'. Although permanently blinded in a lab accident at the age of 26, he worked on full-time research [mostly on antimalarial chemotherapeutics] until the day of his death.

Refs

Kermack WO, Mckendrick A. 1927 *Bull Math Biol*, **53** (1–2): 33–55; 57-87; 89-118

Davidson JN 1971. *Biogr. Mems Fell. Royal Soc* 17, 399–429

27 Apr OTD 1657 Christiaan Huygens (1629-1695) publishes *De Ratiociniis in Ludo* (the Value of All Chances in Games of Fortune) the first printed work on the calculus of probabilities. Only 14 pp long, it consists of 14 'propositiones' and 5 problems, with partial solutions. The standard text on probability theory for the next 50 years and a major influence on Jacob Bernoulli (1654 – 1705), de Moivre (1667 – 1754), and Montmort (1678 – 1719), it was among the earliest to define and systematize concepts such as expectation. Some of the propositions and calculations were later developed into probability distributions.

Ref

https://math.dartmouth.edu/~doyle/docs/huygens/huygens.pdf

28 April OTD 2015 Eva E Jacobs d (b 17 Jan 1921) USA. ASA Fellow 1982. Best known as recipient of the Julius Shiskin Memorial Award for Economic Statistics for her management of the Consumer Expenditure Survey Program. She was long-term editor of the Handbook of U.S. Labor Statistics 1997-2008 #BLS-LaborStatistics

29 April OTD 1667 **John Arbuthnot** baptised (*d* 27 Feb 1735) Scotland. FRS 'Phyfitian in Ordinary' to Queen Anne. His 1692 translation of Huygens's *De Ratiociniis in Ludo Aleae* was the first work on probability in English.

He also developed the first formal significance test (and possibly the first published report of a nonparametric test) in 1710 to evaluate the hypothesis of the 'Exacte Ballance' of human sex ratios at birth. Probabilities were expressed as the 'lot' or 'summe of all the Chances' of a 'very Greatte Number of dice throwes'. He computed the p-value (via a sign test), interpreted it as statistical significance, and rejected the null. He attributed the resulting small p-value to divine providence.

Fun facts: He was good friends with Alexander Pope and Jonathan Swift, and was inspiration for *Gulliver's Travels* Book III. Ironically for a 'Proude Scot' he was the inventor of John Bull (1712).

In 1726 he was one of the medical experts called in to review the strange case of Mary Toft, who claimed to have given birth to 17 rabbits. He apparently "felt some initial hesitation about the authenticity of the births".

Saying "Biography is one of the new terrors of death' he tried to avoid a posthumous biography by encouraging his young children to play with and burn his papers.

Refs
Arbuthnot 1710 *Philos Trans Royal Soc Lond* 27 (325–336): 186–190. doi:10.1098/rstl.1710.0011. S2CID 186209819.
Bellhouse DR 1989 *International Statistical Review* 57 (3): 249–259. doi:10.2307/1403798
Todd D 1988 *Studies in Bibliography* 41, 247-267
https://www.gla.ac.uk/myglasgow/library/files/special/exhibns/month/aug2009.html

30 Apr OTD 1916 Claude Shannon b (d 24 Feb 2001) USA. Mathematician and Bell Labs engineer, father of communication and algorithmic information theory, and developer of some of the first experiments in AI. His work was fundamental to development of the internet, digital computers, smart phones, zip files and jpgs.

Even his master's thesis was epochal – it was an application of Boolean algebra to the design and simplification of electromechanical switching circuits in order to implement logic. This was the foundation of modern digital circuit design. Gardner called it "the most important, and also the most noted, master's thesis of the century." His PhD thesis was on algebraic formulations of theoretical genetics.

He recognised the binary digit (bit) as the fundamental unit of information communication. He developed the concept of information entropy as a measure of the information content in a message, essentially inventing the modern field of information theory.

Refs

Shannon CE 1948 A Mathematical Theory of Communication*. Bell System Technical Journal* 27 (3) 379-423

Shannon CE, Weaver W. 1949 *The Mathematical Theory of Communication*. University of Illinois Press.

Casagrande et al 2022. *Med Biol Eng Comp* 60:941–955

The HOS-ASA Twitter Diaries: MAY 2023

01 May [originally posted 03 May] OTD 1971 The Statistics Act of Canada was passed, proclaiming Statistics Canada as the national statistical office, replacing the original Dominion Bureau of Statistics, established 1918. The new name was not implemented until August to avoid problems with the 1971 Census.

Ref
Statistics Canada: a long-standing Canadian institution
https://www.statcan.gc.ca/en/about/history/index

OTD 1942 Jane Wadsworth b (d 12 Jul 1997) UK. Fellow Royal Statistical Society 1976. Medical Statistician at Barts & St Mary's Hospital Paddington. Pioneering sexual health researcher. The 1994 National Survey of Sexual Attitudes and Lifestyles was the largest, most authoritative survey of human sexual behaviour ever performed in the UK. Although (amazingly) dismissed by the tabloid press as "boring and predictable", it was also famously banned from public funding by Thatcher (but later funded by Wellcome Trust).

Refs
Overy C, Reynolds L A, Tansey E M (2011) History of the National Survey of Sexual Attitudes and Lifestyles. Wellcome Witnesses to Twentieth Century Medicine, vol. 41. London
Boulton M 18 Jul 1997
https://www.independent.co.uk/news/people/obituary-jane-wadsworth-1251436.html

02 May OTD 1860 D'Arcy Wentworth Thompson b (d 21 Jun 1948) GB. FRS KB Pioneered applications of mathematical biology, morphogenesis, & allometry. Modern statistical algorithms for quantifying shape deformation have been developed for studying biological shapes & extended to computer vision & facial recognition.

His insight was that major differences in biological shapes could be described in terms of relatively simple mathematical transformations. He was among the earliest to describe numerical relationships between spiral structures in plants (phyllotaxis) & the relationship to the Fibonacci sequence.

His classic book '*On Growth & Form'* was major influence on Alan Turing, Thomas Huxley and Steven Jay Gould, and on artists such as Henry Moore. The Nobel Laureate Sir Peter Medawar called it "the finest work of literature in all the annals of science that have been recorded in the English tongue". He was also an accomplished scholar of ancient Greek; he translated Aristotle's *History of Animals* into English.

Refs
Mardia et al *Significance* 2018. https://doi.org/10.1111/j.1740-9713.2018.01145.x
https://www.theguardian.com/books/2017/jul/21/growth-form-darcy-wentworth-thompson-review
https://artuk.org/discover/curations/a-sketch-of-the-universe-the-artistic-influence-of-darcy-thompson

04 May May the Fourth Be With You. That's it for today. However, for Star Wars fans everywhere, as a treat check out 'Top 5 Lessons Star Wars Can Teach Us About Data Science'

https://blastpoint.com/blog/top-5-lessons-star-wars-can-teach-us-about-data-science/ ; https://getyarn.io/yarn-clip/81c97c47-bfcb-48ef-9e4b-9d4345da7212

05 May OTD 1777 Leonard Euler introduces i to represent an imaginary number $i^2 = -1$ in his paper 'De formulis differentialibus angularibus maxime irrationalibus, quas tamen per logarithms et areus cireulares integrare licet; die 5. Maii 1777'. ['Of the most irrational Angular Differential Formulas, which, however, may be integrated by means of logarithms and circular arcs']

According to Cajori, the next appearance of i in print is by Gauss in 1801 in the Disquisitiones Arithmeticae. Carl Boyer believes that Gauss' adoption of i made it the standard, certainly it was tbhe standard by 1821, when Cauchy published *Cours d'Analyse*, and defines i as "as if it was a real quantity whose square is equal to -1."

Refs

Cajori F 1893 *A History of Mathematics*
https://scholarlycommons.pacific.edu/cgi/viewcontent.cgi?article=1670&context=euler-works
Beman WW. 1898. https://www.ams.org/journals/bull/1898-04-06/S0002-9904-1898-00494-9/S0002-9904-1898-00494-9.pdf

06 May Statistical sampling methods for population enumeration have been around a long time. During the Graeco-Persian wars (450 BC), the Persian king Xerxes ordered one of earliest examples of expansion sampling without replacement.
According to Herodotus (c484-c425 BC), Xerxes had 10,000 soldiers squash together, built a corral around them, then marched successive groups of men into the corral until the entire army had passed through. Herodotus claimed there were ~5.2 million men; modern historians think the number was more like 300,000.

Refs

Herodotus. *The Histories*, Penguin
Rubin E 1968 The statistical world of Herodotus. Amer Stat 22(1): 31-33

06 May OTD 1912 Margaret E Martin b (d 16 May 2012) USA ASA Fellow 1960, 4th woman to be ASA President in 1980. Best known for development of labor and unemployment statistics, and the Current Population Survey (CPS). This is the primary source of labor statistics, and most well-known and widely used of all ongoing federal household surveys. She was awarded the first ASA Founders Award in 1989. She was elected member of the International Statistical Institute, an honorary lifetime member of the board of the Council of Professional Associations on Federal Statistics and was chair of Section U (statistics) of the AAAS in 1986. She was also first executive director of the Committee on National Statistics

Refs

Muko M. *Amstat News*. 2011-09-01.
https://magazine.amstat.org/blog/2011/09/01/margaretmartin/

07 May The Canada Population Clock was constructed in 1968 to commemorate the 50th anniversary of the Dominion Bureau of Statistics, and highlight the increasing role of computer technology in the work of the Bureau. It was a 7 foot tall analogue data visualisation tool for the census, and showed how population estimates changed over time. Although it became something of a tourist attraction, it was dismantled after a few years. Nobody seems to know what happened to it. Nowadays everything is online, and the census can be updated in real time.

Refs

Standing on the Shoulders of Giants: History of Statistics Canada 1970 to 2008. https://www150.statcan.gc.ca/n1/en/pub/89-20-0001/892000012018001-eng.pdf?st=eJt8YV4S

https://www.statcan.gc.ca/o1/en/plus/1352-canadas-population-clock-data-modelled-real-time

08 May Austin Bradford Hill's streptomycin trial 1947 is the best known example of using random subject allocation as the most fair way of distributing a limited resource (as well as reducing bias in assessment of efficacy). However, he was scooped by ~2000 yr. Herodotus (c484-c425 BC) reports the King of Lydia used random selection to solve a problem of scarce resource allocation following an 18-year famine. "The King divided the population into two groups determined by drawing lots which should emigrate, and which should remain home"

Refs

Herodotus. *The Histories*, Penguin

09 May In April 1934 William Cochan (1909-1980) published his first paper on what is now known as Cochran's Theorem. It showed the distribution of quadratic forms of normally distributed random variables were independent & $\sim \sigma^2 \chi^2$ with r degrees of freedom. The resulting properties were quite nice; they could be extended to situations where analysis of variance required covariate adjustment. This is known as analysis of covariance (ANCOVA)

Refs

Cochran W (1934). *Math Proc Camb Phil Soc* 30(2), 178-191

Hansen M, Mosteller F 1987

http://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/cochran-william-g.pdf

10 May The first International Statistical Congress was held 19 Sep 1853. It was organised by Adolph Quetelet. On paper, the goal seemed straightforward: the development of unified standards of measurements, measurement processes, and reporting, especially for surveys and censuses. However, nothing much was ever accomplished.

Root problems were ineffective communication and nationalism. There were quarrels over agenda, and who was to be invited. Even choice of language was disputed: should it be French or that of the host nation? One attempt to reach a compromise by using Latin was unsurprisingly not successful. There was also considerable heartburning over accusations that French and German participants were fudging their statistics.

Although there would be nine further congresses, Otto von Bismarck finally pulled the plug on any more after 1878 by refusing to allow Prussian statisticians to participate.

Ref

Randeraad N. 2011. The international statistical congress (1853-1876): Knowledge transfers and their limits. *European History Quarterly*, 41(1), 50-65. https://doi.org/10.1177/0265691410385759

11 May OTD 1920 Edward Kaplan b (d 26 Sep 2006) USA. Best known for the Kaplan–Meier product-limit statistic for estimating survival functions from lifetime data. The 1958 joint paper with Paul Meier is now one of the most highly-cited statistics papers of modern times, although it took >12 yrs to catch on.

Originally Kaplan independently developed his method for engineering applications (such as estimating mean lifetime of light bulbs and vacuum tubes). Kaplan submitted his manuscript to the *Journal of the American Statistical Association*, which recommended a joint paper with Meier because of what they said was a perceived overlap. Kaplan later said wryly that "Much correspondence over four years was required to reconcile our differing approaches".
The Kaplan–Meier method took off when it was applied to clinical problems. Gehan used it to study cancer survival in 1969. Cox 1972 highlighted its importance for clinical survival analyses. Peto et al (1976, 1977) published two tutorial papers on its application to randomised clinical trials.
Scandalously, Kaplan's contributions have been completely overlooked. He never received a medal or award, his name is not on the list of ASA Fellows, and his death was almost unnoticed by major mathematical and statistical societies. The 2004 ASA Samuel S Wilks Award specifically highlighting the KM method was given to Meier. Kaplan was not invited and "was not even notified about it". Come on ASA!

Refs

Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association.* 1958; 53:457–81.

Stalpers LJA Kaplan EL (2018) *BSHM Bulletin: J Brit Soc Hist Math*, 33:2, 109-135, doi: 10.1080/17498430.2018.1450055

12 May OTD 1820 Florence Nightingale b (d 13 Aug 1910). British. Where to start? Best known as founder of modern nursing but also a renowned statistician and pioneer of data visualisation. First woman given membership in the Royal Statistical Society 1859 and was made honorary member of ASA in 1874.

She vigorously championed use of infographics over tables at a time when statisticians dismissed them as not rigorous and even intellectually dishonest. The 'coxcomb' showing a time series of mortality patterns is most famous of her data plots. Her bar charts were highly effective in showing how mortality of British soldiers greatly exceeded that of civilians, accelerating sanitary reforms. She also led the charge for standardised data collection plus statistical rigour to institute massive healthcare policy and social reforms. What statisticians often forget is that her many data-based reforms were also supremely practical: sanitation, hospital construction, laundry, good food, basic supplies, and inventory.

In the 1870s she and Benjamin Jowett Master of Balliol, discussed founding an Oxford University Chair of Statistics. She originally left £2000 in her will for the purpose but later revoked it, saying she was afraid it would be used for an essay prize rather than for teaching statistics.
Fun Facts: When you are in London, a recording of her voice can be found in an 1890 phonograph recording at the British Library Sound Archive (now it can be found on Wikipedia but that isn't the same). An 1863 letter from Nightingale can be seen at Floris of London perfumers, Jermyn St. And be sure to visit the Florence Nightingale Museum https://www.florence-nightingale.co.uk/

Refs
Magnello E. 2012 *BSHM Bull: J Brit Soc Hist Math* 27: 13–37
Bostridge M. 2009 *Florence Nightingale: The Woman and Her Legend.* Penguin
Cohen IB. 1984 *Scientific American* 250 (3): 128–137
Special Issue *Significance* Apr 2020 17(2)
https://www.significancemagazine.com/10-news/643-finding-nightingale

Also, OTD 1902 Frank Yates b (d 17 Jun 1994). UK. Another giant of experimental design. Guy Gold Medal 1960, Royal Society Medal 1966. He also enthusiastically promoted use of computers for statistical analysis, he was founder and President of the British Computer Society.

13 May OTD 1918 Frank Anscombe b (d 17 Oct 2001) UK/USA. ASA Fellow 1956. Best known for the most compelling argument ever made why data should be plotted before analysis. Anscombe's quartet shows 4 data sets with identical summary statistics but very different distributions. Other notable contributions were on methods of experiment validity, randomisation of block designs, the Anscombe variance-stabilising transformation for Poisson data and the formal properties of residuals in linear regression. He later became interested in statistical computing and computerised data visualisation methods, saying that "a computer should make both calculations and graphs"

Fun fact: He was brother-in-law to John Tukey

Refs
Anscombe FJ 1948. *J Royal Stat Soc Series A* 111(3):181-211
Cox DR 2003 *J Royal Stat Soc Series D* 52(4): 679
Saxon W Francis John Anscombe, 83, Mathematician and Professor. The New York Times. 23 Oct 2003

14 May On this Mothers' Day USA a shout-out to Louis I. Dublin (1882-1969) ASA Fellow 1916, ASA President 1924. In 1935 he reported US maternal death rates were a staggering 15K women/yr (surpassed only by Chile), and showed how statistical data could be used to inform medicine & public health programs & improve maternal life expectancy.
Shamefully, pregnancy-related deaths in the US have been increasing from 7.2 (1987) to 17.3 (2018) per 100K live births. In 2021 USA ranked first among high-income countries for maternal death, with homicide (intimate partner + firearms) as the leading cause.
In his report "Lost Mothers" Dublin said "It is not sufficient for us to salve our conscience on Mother's Day by casting a white carnation on the grave of the mother who died that her child might live; we must rather apply our knowledge and experience so that these women need not die".

Refs
Obituary *Am J Public Health* 1969 59(7)1083-4
Baker N. 11 May 2023
https://circulatingnow.nlm.nih.gov/2023/05/11/remembering-mothers/
Lawn RB et al 2022. *BMJ* doi: 10.1136/bmj.o2499

[originally posted 23 May] Also OTD 1899 Joseph Berkson b (d 12 Sep 1982) USA. ASA Fellow 1940, President ENAR. Best known for Berkson error (or regression calibration) models (1950) and Berkson's paradox (collider or ascertainment bias), a type of selection bias in observational studies (1946). He also introduced the logit model 1944. He coined the term logit intending it to be a shameless analogy to Bliss' probit model.

Like Fisher, & for much the same reasons (namely the poor methodological quality of much epidemiological research), he strenuously opposed data showing cigarette smoking causes cancer.

Fun fact: The Life magazine article 24 January, 1964 "Verdict on Cigarets: Guilty as Charged" pp 56A-62, 64, describes the main findings of the U.S. Surgeon General's Report on Smoking and Health. On the back cover is a full-page Winston cigarette advertisement, which has to be seen to be believed.

Refs
Berkson J 1944. *J Amer Stat Assoc*, 39, 357-65; 1946. *Biometics Bull* 2:47-53; 1950. *J Amer Stat Assoc*, 45 (250): 164–180
Herbert et al *Significanc*e 2020
https://rss.onlinelibrary.wiley.com/doi/full/10.1111/1740-9713.01413
O'Fallon JR *Biometrics* 39(4):1107-8

15 May OTD 1908 Helen Alma Newton Turner b (d 26 Nov 1995) Australia. Officer Order of Australia, OBE. Pioneered application of population genetics & objective, measurement-based research to sheep breeding & wool production in Australia. She trained in the UK in the 1930s with Fisher and Yates, then returned to Australia as consulting statistician to CSIRO. The Turner medal was established in 1993 in her honour by the Association for the Advancement of Animal Breeding and Genetics (AAABG)

Ref
McCann D Australian Dictionary of Biography 19, 2021
https://adb.anu.edu.au/biography/turner-helen-alma-29660

16 May OTD 1821 Pafnuty Chebyshev b (d 8 Dec 1894) Russia. Founding father of Russian mathematics. Best known for Chebyshev inequality 1857 which provides a generalised law of large numbers & generalising the central limit theorem. Kolmogorov said he was first to recognise the concept of orthogonal polynomials and to think systematically in terms of random variables and their moments and expectations.
His student said his lectures were 'wonderful', 'lively, absorbing', 'highly stimulating', & 'very short'. He once said "To isolate mathematics from the practical demands of the sciences is to invite the sterility of a cow shut away from the bulls."
Being keen on the practical applications of mathematics meant that he was also a prolific inventor of mechanical devices. In 1893 at the Chicago World's

Exposition he exhibited 7 of his inventions including "a special bicycle for women", and the first 'walking machine' that converted rotary to walking motion (a precursor of robotic devices). His Chebyshev Linkage has been resurrected as a new rear suspension design for mountain bikes.

Refs

Butzer P, Jongmans F 1999. *Journal of Approximation Theory* 96, 111-138
His machines: P E Tovstik et al 2021 *J. Phys.: Conf. Ser*. 1959:012050
https://en.tcheb.ru/ ; Walking machine.
https://www.youtube.com/watch?v=ISfVS4mDTKs

17 May William Thomas Tutte (14 May 1917- 2 May 2002) UK/Canada. FRS; Officer Order of Canada. Best known for deciphering the German Lorenz cipher Tunny at Bletchley Park during WWII. Later he became a world leader in graph theory, matroid theory, and combinatorics.
Tutte worked out the logical structure of the Lorenz machine from a keystream intercept "completely recreating the machine without ever having seen one". He worked out "the complete internal structure, that it had 12 wheels …. the number of sprockets on each wheel, & how the advancement of the wheels was interrelated." Tony Sale called it the "greatest intellectual feat of the whole war."
In graph theory he established fundamental results for matching, connectivity, symmetry in graphs, reconstruction, colouring, Hamiltonian circuits, graphs on higher surfaces, graph enumeration, and graph polynomials.

Refs

On Tunny: https://uwaterloo.ca/combinatorics-and-optimization/sites/ca.combinatorics-and-optimization/files/uploads/files/independent-obituary.pdf
Hinsley FS, Stripp A (Eds). 1993. *Code Breakers*. Oxford University Press
Webb J, Fox B 1997 Colossal adventures. *New Scientist*

18 May OTD 2007 Charles Dunnett d (b 24 Aug 1921) Canada, MBE, ASA Fellow 1965, President Statistical Society of Canada 1982, SSC Gold Medal 1986. He received his MBE for radar work during Royal Navy Service in WWII. Best known for Dunnett's test for pairwise multiple comparisons, & equivalence testing. Two papers on the problem of multiple comparisons testing between treatments and a control (1955, 1964) are among the top 25 most-cited papers in statistics of all time. His 1977 paper on equivalence testing was one of the first on this subject & motivated by medical applications.

Refs

https://ssc.ca/fr/node/4211; Liaison 3(1) Nov 1988.
Dunnett CW 1955. J Amer Statist Assoc, 50:1096-1121; 1964 Biometrics 20:482-491. Dunnett CW, Gent M 1977. Biometrics, 33:593-602.

19 May DYK? For better or worse, statisticians were partly responsible for pioneering concepts of "wellness". Adolph Quetelet (1796-1874) developed the Body Mass Index as body weight (kg) proportional to the square of height (m). However, it based on data from French & Scottish soldiers and was not supposed to be an index of individual health. It was not actually called BMI until 1972 by Ancel Keys.

Body weight was first used as a primary indicator of health in 1912 Louis Dublin (1882–1969; ASA President) developed tables of body weights for insurance purposes in 1912. 'Undesired' weights were assessed as percent above the 'desirable' weight for a given body frame, e.g. 'undesired' >20–25%, 'morbid obesity' 70–100% above.

And then there was Halbert L Dunn (1896–1975) ASA Fellow 1939, Chief National Office of Vital Statistics 1935-1960, and the Father of the wellness movement. He defined "high-level wellness" as "a condition of change in which the individual moves forward, climbing toward a higher potential of functioning." Nowadays, the global wellness industry is worth over $4 trillion.

Refs

Eknoyan G. 2008. *Nephrol Dial Transplant* 23(1):47-51. doi: 10.1093/ndt/gfm517.
Louis I Dublin https://magazine.amstat.org/wp-content/uploads/2017/08/dublinlouis.pdf
Dunn HL 1957 *J Natl Med Assoc*. 49(4):225-35; 1959 AJPH 49(6) 786-792; 1959 *Can J Public Health.* 50:447-57.

20 May OTD 1913 Mindel Sheps b (d 13 Jan 1973) Canada. ASA Fellow 1970. Best known for her 1958 paper discussing the asymmetry of relative risk – it is "not invariant to choice of reference level for the outcome". Sheps' solution was to use clinical information to determine the appropriate denominator. She also made substantial contributions to mathematical demography. She developed theoretical 'family building" and fertility models based on a "renewal process framework of repeated events and intervals between them". The models predicted the effects of system disruptions, such as sudden cessation of abortion access.

Refs

Sheps MC 1958. Shall we count the living or the dead? *NEJM* 259:1210-1214
Huitfeldt A 2023 *Epidemiology* 34(3):p 396-399,
Menken J 2020 *Theor Pop Biol* 133: 71-74

21 May Original Sin in statistical inference? In 1827 Laplace reported a "small value" (~0.01 significance level) for a statistic closely related to the modern p-value, leading him to conclude the difference in measurements was "not due solely to the anomalies of chance". However, Stigler points out Laplace himself made several mistakes still common today: non-independence of observations & incorrect estimates of variance

Ref

Stigler SM. 1986 *The History of Statistics* Belknap Press

22 May DYK? Principles 2, 5, & 6 of the 2016 ASA Statement on p-values were proposed as early as 1843. Antoine Augustin Cournot (28 Aug 1801-31 Mar 1877; France) was the earliest to describe the p-value as combining the size of the effect and the sample size. He also warns against overly narrow interpretations of probabilistic statements: the p-value, he said, "does not at all measure the chance of truth or of error pertaining to a given judgment."

Refs

Wasserstein, Lazar. 2016. *Amer Stat*, 70:2, 129-133, doi: 10.1080/00031305.2016.1154108

Cournot AA. 1843 *Exposition de la Théorie des Chances et des Probabilités,* Paris: L. Hachette
Translation https://arxiv.org/ftp/arxiv/papers/1902/1902.02781.pdf
Kennedy-Shaffer l. 2019. *Amer Stat* 73: Suppl 1, 82-90, doi: 10.1080/00031305.2018.1537891

24 May OTD 1686 Daniel Gabriel Fahrenheit b (d 16 Sep 1736) Dutch Republic. He pioneered exact thermometry using early principles of gauge repeatability and reproducibility - taking a series of measurements to determine if the same results are obtained under the same operating conditions. His mercury-in-glass thermometer (1714) was the most reliable and accurate temperature device until recent times. The Fahrenheit scale was the primary temperature standard until the 1970s.

Orphaned at 15 when his parents died from eating poisonous mushrooms (I know, right?), he was apprenticed to a merchant & learning book-keeping when he discovered thermometers, invented in Florence 60 years before.

Florentine thermometers had a major design flaw: No 2 instruments measured the same temperature, because there was no standard way to calibrate them - everyone just did their own thing.

Our Daniel skipped out on his apprenticeship to learn about thermometer-making. As a result, he was subject to arrest - his guardians were legally liable for his debts so they planned to ship him off to Indonesia to work until he repaid them. So, Daniel fled the country and spent years on the run, learning about thermometers.

His break-through in developing a reliable, standardized instrument was to use mercury instead of alcohol, and a scale calibration using reproducible benchmarks, a major insight. He defined 0°F as the freezing temperature of a brine solution made from equal parts water, salt, and ice, and 96°F as the temperature of the human body (actually his body) measured by placing the thermometer under his arm. He died aged 50 probably from mercury toxicity.

In 1742 Anders Celsius developed a more intuitive temperature scale between 0°C (freezing point) & 100°C (boiling point) of water. Nowadays, only the USA still stubbornly holds on to the Fahrenheit scale (and other imperial units)

Refs
Friend, JN. 1937. *Nature* 139: 395-398
Grigull, U. 1986. Fahrenheit: A Pioneer of Exact Thermometry, Proceedings of the 8th International Heat Transfer Conference 1:9-18

25 May Time for the annual revisit to Stigler's Law of Eponymy (1980): "No scientific discovery is named after its original discoverer". Kennedy points out this "law" actually proposed by CB Boyer 1968: "Mathematical formulas and theorems are usually not named after their original discoverers".
Some surprising statistics examples:

Pearson's correlation coefficient: Bravais 1846
Venn diagrams: Euler 1768
Simpson's paradox: Blyth 1972
Gauss' law: Lagrange 1773

L'Hôpital's rule: probably Johann Bernoulli
Poisson distribution: de Moivre 1721?
Chebyshev's inequality: Bienaymé 1853
Bonferroni correction: Olive Jean Dunn 1958
Bradley–Terry model: Zermelo 1929.
Benford's law: Simon Newcomb 1881
Chernoff bound: Herman Rubin
Deming cycle of continuous improvement: Attributed by Deming himself to Shewhart
Voronoi diagrams (1908): Descartes (1644), Dirichlet (1850), Snow (1854)

Refs
Stigler S 1980. *Trans NY Acad Sci* doi: 10.1111/j.2164-0947.1980.tb02775.x issn: 0028-7113
Kennedy HC 1972. *Amer Math Monthly* 79(1): 66-67

26 May OTD 1623 Sir William Petty b (d 16 Dec 1687) England. FRS, physician, Gresham College Professor of Music, scientist, economist. At a time when economics consisted of subjective descriptions, he pioneered use of early quantitative estimation methods (mostly averages). His "Political Arithmetick" was the foundation for modern macroeconomics. He also developed crude census methods, working with John Graunt on Bills of Mortality, the first published books on vital statistics (1662, 1682).

Because reliable (or any) census data were almost impossible to obtain, he estimated population numbers using simple arithmetic rules of thumb. He was accused on several occasions of fudging his numbers in favour of the English crown over Irish land rights. John Mitchel (1815-1875) nationalist (and no fan) called him a carpet-bagger, a "land-pirate", and the "most voracious land-shark who ever appeared in Western Europe".

John Aubrey has several good stories about Petty. The first was 'the moft remarkable accident of life which was the foundation of all the reft of his greatnes". This turned out to be his going to sea as a cabin boy, then being marooned on an island near France at the age of ~12 by a ship's crew "made Envious by his Precocious talents" (Editor's Note: read "annoying brat").

His talent for annoying the powerful persisted to middle age and resulted in what Stephen Pile called "the worst duel ever". Challenged in 1645 by Sir Hierome Sankey, Petty would have been hopelessly outclassed because he was extremely short-sighted and no physical match. However, it was his right as the challenged to choose venue and weapons. "Brilliantly, he chose a pitch-dark cellar and two carpenters' axes which neither of them could lift".

Refs
Dictionary of National Biography. London: Smith, Elder & Co. 1885–1900.
Aubrey J *Brief Lives*. Penguin
Pile S 1978 *The (Incomplete) Book of Failures*. Dutton
Mitchel J 1873. *The Crusade of the Period And Last Conquest Of Ireland,* https://cartlann.org/authors/john-mitchel/the-crusade-of-the-period/

[Originally posted 29 May] Also OTD 1677 Abraham de Moivre b (d 27 Nov 1754) France. FRS. One of the pioneers of classical probability theory, he is best

known for his large-sample approximation to the binomial distribution by the normal distribution & the first occurrence of the normal probability integral. He published the first textbook on probability theory 'The Doctrine of Chances' 1718, and first publication on the normal approximation to the binomial Nov 1733

He published the first textbook on probability theory '*The Doctrine of Chances'* 1718, an expansion of a 52-page "memoir" '*De mensura sortis'* in 1711. On 13 Nov 1733 he first published the normal approximation to the binomial, again in Latin. This subsequently appeared in English in later additions of The Doctrine. He wrote that he took "the liberty to say, that this is the hardest Problem that can be proposed on the Subject of Chance".

Following the revocation of the Edict of Nantes and religious persecution of Protestants, he was imprisoned for several years, after which he went to England. Although regarded as "one of the greatest mathematicians in Europe" and acquainted with many of the great names of the day (Newton, Halley, Bernoulli, Leibniz), he had to make ends meet by tutoring, and hanging out in coffee-houses, playing chess and advising on gambling problems for extra cash. Attempts made by his friends to get him a patronage appointment or a professorship at Cambridge failed, as was thought at the time due to the English distrust of foreigners.

He was elected Fellow of the Royal Society in 1697 and was finally admitted to l'Académie Royale des Sciences in 1754. He regarded this as the crowning moment of his career, but unfortunately, he died shortly after. The story that he correctly predicted the date and time of his own death based on his increasing sleep time is probably apocryphal.

Fun fact: De Moivre mostly hung out at Slaughter's Coffeehouse London, a favourite meeting place of French émigrés. The building was demolished in 1843 when Cranbourne Street was constructed.

Refs
De Moivre A 1711. *De mensura sortis*. Philos Trans Royal Soc
Bellhouse DR Genest C . 2007. *Stat Sci* 22(1) 109-136
Hald, Anders 1990. De Moivre and the Doctrine of Chances, 1718, 1738, and 1756, *History of Probability and Statistics and Their Applications before 1750,* Wiley Series in Probability and Statistics

27 May OTD 2015 Yvonne M M Bishop d (b 12 Jan 1925) UK/USA. ASA Fellow 1975. Best known for THE text on multivariate statistics & pioneering work on multi-dimensional contingency tables. Often overlooked are her considerable contributions to the National Halothane Study that established association of halothane anesthesia with fatal postoperative hepatic necrosis. She was one of 9 statisticians and the only woman. This study revolutionised the study of surgical and anaesthetic safety, ushered in a new emphasis on medical quality improvement, and even modern data processing.

Later she was lead on the Harvard Six Cities study of air pollution on public health, which led to major developments in methods for longitudinal data analyses. She was Director Office of Statistical Standards, Energy Information

Administration at Harvard and the Office of Statistical Standards, Energy Information Administration.

She originally planned to be a fisheries biologist and did considerable work on fisheries data but quit because of rampant sex discrimination.

Refs

Bishop et al. *Discrete Multivariate Analysis: Theory And Practice*

Bishop et al. Chap IV-3. Smoothed contingency table analysis, Chap IV-5 Analyis by regression methods. In Bunker JP et al (eds) 1969. *National Halothane Study.* Report. National Academies of Sciences Washington DC https://doi.org/10.17226/19006.

Mosteller F. 2010. *The pleasures of statistics: The autobiography of Frederick Mosteller,* Springer

Moon JS 2021 *Anesthesiology* 135:853.

28 May In May and June 1892 Francis Ysidro Edgeworth (1845-1926) gave six Newmarch lectures at University College London, "On the Uses and Methods of Statistics". For his lecture on correlation, Edgeworth borrowed Galton's quincunx for his demonstrations. Edgeworth later told Galton in a letter that 'the quincunx worked splendidly and imparted distinction to my discourse".

Refs

Edgeworth, FY Methods of Statistics. *Journal of the Statistical Society of London,* Jubilee Volume (Jun. 22 - 24, 1885), pp. 181-217

Edgeworth, FY. Letter to Galton (9 June 1892), UCL Galton Archive, 237

Aldrich J. 2010 . https://studylib.net/doc/10897909/mathematics-in-the-london-royal-statistical-society-1834

30 May In May 1850 the Epidemiological Society of London had its first meeting at Hanover Square, London (first president Benjamin Guy Babington). Although early discussions mostly concerned theories of contagion, statistics and statistical methodology became more important. In May 1853 John Snow (1813-1858) read a paper demonstrating how statistics could be applied to understanding mortality rates and causes of death.

Refs

Snow J. 1855. On the Comparative Mortality of Large Towns and Rural Districts, and the Causes by Which It Is Influenced. *J Public Health Sanit Rev*. 1(4): T16-T24.

Frerichs RR. London Epidemiological Society. UCLA https://www.ph.ucla.edu/epi/snow/LESociety.html.

31 May OTD 1780 Laplace read his "Mémoire sur le calcul aux suites appliqué aux probabilités" to the l'Académie Royale des Sciences. Although he didn't call it that, he describes inverse probability, or assignment of a probability distribution to an unobserved variable. Laplace said this work had been sitting on his desk for several years before he got around to it again. Fisher changed the term to "Bayesian" in 1950.

The paper describes the computation of the probability of events composed of many simple events of unknown probability, the probability of future events given past events & develops large sample tests of significance for testing human sex ratios at birth. A review by Marquis of Condorcet in the same issue describes the historical background.

Refs

Laplace 1778. *Mém Acad R Sci Paris*, 1778 (1781):227-332. *Oeuvres* 9: 383-485.
transl. R J Pulskamp http://www.probabilityandfinance.com/pulskamp/Laplace/memoir_probabilities.pdf
Condorcet, 1778. Summary 'Sur Les Probabilités' transl. R J Pulskamp, http://www.probabilityandfinance.com/pulskamp/Laplace/review_sur_la_probabilites.pdf
Bru B, Crépel P (1994) *Condorcet: Arithmétique politique, textes rares ou unédits* (1767-1789)
Stigler S 1986 *Statistical Science,* 1(3): 359-363

The HOS-ASA Twitter Diaries: June 2023

01 Jun OTD 1866 Charles Davenport b (d 18 Feb 1944) USA. ASA Fellow 1921. Director Cold Spring Harbor Labs, His 1899 text 'Statistical methods with special reference to biological variation' actively promoted statistical methods developed by pioneered by Galton & Pearson to the USA. It demonstrated their application to biological research & the study of variation.

Unfortunately, he is even better known for his truly vile eugenicist views. He founded the Eugenics Record Office in 1910 and proposed to transform the human race by selective breeding of the picked few, forced sterilization, and stringent immigration laws. His "scientific justifications" for reducing the population of "degenerates" affected US policy on race and immigration for decades and were models for the extermination programmes of Nazi Germany. He was a 'research' collaborator with several Nazi scientists both before & during WWII, and even served on the editorial boards of several Nazi 'racial purity' journals.

Refs

Harris JA 1914, *Science* 39: 828-830; Anon 1899. *Nature* 61:149

Gould SJ 2000. The Internal Brand of the Scarlet W In: *The Lying Stones of Marrakech*

Farber S. A. (2008). U.S. scientists' role in the eugenics movement (1907-1939): a contemporary biologist's perspective. Zebrafish, 5(4), 243–245. https://doi.org/10.1089/zeb.2008.0576

PBS American Experience: https://www.pbs.org/wgbh/americanexperience/features/eugenics-charles-davenport/

Riddle O 1947. http://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/davenport-charles.pdf

Martin P. https://www.salon.com/2014/03/23/hitlers_favorite_american_biological_fascism_in_the_shadow_of_new_york_city/

Also, OTD 1926 Marie D Eldridge b (d 13 June 2009) USA. ASA Fellow 1969 Director, Statistics and Analysis, National Highway Traffic Safety Administration; Director, Research Triangle Institute; President, Washington Statistical Society

02 Jun 2012 Genuchi Taguchi d (b 1 Jan 1924) Japan. Founder of the Taguchi method for quality product improvement - reduction of process variation through robust design of experiments. His methods revolutionized manufacturing quality control practices.

He collaborated throughout the 1950s with other notable statisticians such as CR Rao, Ronald Fisher, and Walter Shewart, and was sponsored by John Tukey at Princeton.

His methods differed from the conventional specifications based on tolerances alone but developed the concept of quality loss (rather than just quality. His method involves 3 steps: System design (operational specifications parts, materials, assembly); Parameter design (deciding the most cost-effective methods to control variation); Tolerance design (identification of key noise factors)

03 Jun OTD 1879 Raymond Pearl b (d 17 Nov 1940) USA. ASA Fellow 1920, ASA President 1939. Best known for the Pearl-Reed-Verhulst logistic growth model dN/dt = rN(K-N/K), founder of biogerontology (study of the relationship of allometry, bioenergetics, and longevity) and one of the first biostatisticians to combine mathematics and statistics in the study of population genetics. In 1926 he founded the *Quarterly Review of Biology*, and in 1929 *Human Biology*. Chief of the Statistical Division of the US Food Administration 1917-1919 first Chair Dept Biometry and Vital Statistics at Johns Hopkins 1919-1925

Fun fact: When in Baltimore, Pearl belonged to a rowdy group called the Saturday Night Club which included HL Mencken. In spite of Prohibition, the drinking was described as 'legendary'.

Refs

Refs Kingsland S 1984. *Human Biol* 56(1) 1–18.

Jennings HS 1942. *Nat Acad Biog Mem* 22, 295-347

Bekllhouse DR 2009. *Int Stat Review* 77(1) 51-63.

Little MA, Garrutto RM 2010 *Human Biol* 82(1) 77-102

Agresti A Meng X-L 2013. *Strength in Numbers: The Rising of Academic Statistics Departments in the U.S*. Springer.

04 June In June 1905 Karl Pearson introduced the terms 'kurtosis', 'leptokurtic', 'platykurtic' & 'mesokurtic' to describe shapes of skewed frequency distributions. The paper was otherwise a huffy (& lengthy) rejoinder to some German critics of his 1899 theory of skew variation. As well as highly entertaining criticisms of his foes, the paper also presents derivations and distribution data for a variety of biological data, including human skulls, crab 'foreheads', shell lengths, and human organ weights. In 1927 WS Gosset ("Student") provided a humorous *aide memoire* for readers unfamiliar with the terms: 'platykurtic': platypus with short tails; 'leptokurtic': 2 kangaroos 'lepping' (I don't make this stuff up you know).

Refs

Pearson K. 1905. Das Fehlergesetz und seine Verallgemeinerungen durch Fechner und Pearson. A Rejoinder. *Biometrika* 4 (1/2): 169-212

"Student" 1927, Errors of Routine Analysis. *Biometrika,* 19: 151-164

Fiori AM, Zenga M 2009. *International Statistical Review* 77(1), 40-50

05 Jun On the run up to D-Day 1944 Bletchley Park codebreakers continued to supply vital intelligence to Allied commanders as the invasion unfolded. The unprecedented & innovative application of linguistics, pure mathematics & statistics & development of earliest computers such as Colossus paved the way for modern AI.

In those days most of the work was done by humans, with women making up 75% of the workforce. Codebreakers helped to decrypt key intelligence in <3 hr on D-Day. Over 18000 encrypted German messages were collected & deciphered DAILY. Most of the information detailing personnel roles & methods was declassified only recently.

Mavis Batey (1921-2013) was the last surviving member of Bletchley's "break-in experts". She cracked the Enigma ciphers of both the Italian Navy & the German

Abwehr, enabling the Double Cross deception that ensured the success of the D-Day landings.

Fun fact: Veteran codebreaker Dilly Knox wrote a poem to celebrate the victory: "When Cunningham won at Matapan, By the grace of God and Mavis, Nigro simillima cygno est, praise Heaven, A very rara avis." She said later that was "very heady stuff for a 19-year-old".

Refs

https://www.bbc.com/news/uk-england-beds-bucks-herts-24939909

https://bletchleypark.org.uk/

https://www.theguardian.com/world/2013/nov/20/mavis-batey

Erskine R, Smith M (eds) 2011. *The Bletchley Park Codebreakers*

06 Jun OTD 1900 Samuel A Stouffer b (d 24 Aug 1960) USA. ASA Fellow 1940. Best known for pioneering survey research methods and quantifying key concepts such as 'attitude'. His method of combining p-values from independent statistical tests to determine if the joint null should be rejected finds wide use today for meta-analyses & analyses of microarray data. It is a compromise between the methods devised by Fisher (sensitive to the smallest p) and Pearson (sensitive to the largest p). However, it was overlooked for a long time because astoundingly it is only a footnote in his landmark survey of combat stress in American soldiers.

Refs

Stouffer SA et al. 1949. *The American Soldier, Vol. 1 - Adjustment during Army Life.* Princeton

Cousins RD 2007 Annotated bibliography of some papers on combining significances or p-values, arXiv:0705.2209

Fisher RA 1932 *Statistical Methods for Research Workers* 4th ed

Pearson K 1933. *Biometrika,* 25(3/4):379-410

06 Jun OTD 1860 Thomas Bassett Macaulay b (d 1942). Canada. Inaugural ASA Fellow 1914. One of 4 Canadian charter members & first Canadian president of Actuarial Society of America 1899 and founder of the Macaulay Institute for Soil Research. It has been estimated that 95% of all Holsteins cattle in the world trace origins to his herd and one bull 'Johanna Rag Apple Pabst' ('Old Joe')

07 Jun OTD 1919 Charlotte Kipling b (d 9 Aug 1992) UK. British statistician, ichthyologist. Best known for long-term studies of freshwater fish (Lake Windermere Perch and Pike Project) and analyses showing how fish populations react to habitat change and overfishing. With >70 years of data, the project has been invaluable for understanding predator–prey dynamics, population sensitivity to climate-induced disruptions, & interactions of complex environmental factors on populations. Long-term ecological studies are rare but precious b/c they are essential for understanding natural processes (such as the influence of stressors) and effects of climate change. Kipling was hired as a statistician by the Freshwater Biological Association at Ferry House in 1947. The research was started in 1939. The original goal was to assess feasibility of fish harvests for enhancing British food supply in the event of war.

Refs

Kipling C 1984 *ICES J. Mar. Sci.* 41, 259–267.

Le Cren et al 1977 *J Animal Ecol* 46(1): 281-307

Mills CA Hurley MA. 2006. *Freshwater Biology* 23(1):119– 136

Langangena Ø et al. 2011 *Fisheries Research* 109: 131-139
Maberly S Elliot J. 2011 *Freshwater Biology* 57(2):233-243

Also, OTD 1934 Michael Hills b (d 7 Jan 2021) UK. Best known for his classic book with David Clayton *Statistical Models in Epidemiology*, described as providing an "accessible but formal link" between statistical modelling, estimation, and epidemiology. He also worked for many years at the London School for Hygiene and Tropical Medicine (LSHTM), and was Head of Biometrics and Computing at the Natural History Museum.

Refs
Hills C et al. 2021. *The Stata Journal* 21(2): 273-278.
Destavola BL 2021. *J R Stat Soc Series A.* 184:1149.

08 Jun Trying to detect a serial killer based on the probability of a certain number of incidents can lead to grave miscarriages of justice, even if calculations are performed by statisticians. The most famous example is Sally Clark, wrongly convicted 1999 for the murder of her two infant sons.

Paediatrician Prof Sir Roy Meadow claimed that the probability of a single SIDs death was 1 in 8543 so the chances of two SIDS deaths in the same family were 1 in 73 million (8543 × 8543), in his opinion leaving murder as the only alternative.

In Jan 2002 the Royal Statistical Society wrote to the Lord Chancellor disputing this estimate as having "no statistical basis". Clark was later acquitted on a second appeal but was unable to recover from her traumatic experience and died young. Journalist Geoff Wansell called this "one of the great miscarriages of justice in modern British legal history"

Dutch nurse Lucia de Berk was convicted in 2003 for a series of patient murders. The probability that led to her conviction was 1 in 342 million. However Gill et al point out data collection methods were inconsistent and variables poorly defined, incidents were treated as independent, there was post-hoc testing, & because events were rare, calculations were sensitive to small errors. de Berk was finally acquitted in 2010.

Refs
Nobles & Schiff 2005 *Significance*
Gill RD et al 2018. *Chance* 31(4) 9-15;
Buchanan, M. 2007. *Nature* 445, 254–255.
https://doi.org/10.1038/445254a
https://www.theguardian.com/society/2007/nov/07/children.uknews 7 Nov 2007
Royal Statistical Society statement
https://web.archive.org/web/20110824151124/http://www.rss.org.uk/uploadedfiles/documentlibrary/744.pdf

09 Jun Forensic statistics Part 2. Over 40 yr ago Doug Altman stated unequivocably that "misuse of statistics is unethical". Guilt or absence of guilt is an expression of the posterior odds, but basic statistical errors can bias results and interpretations. All applied statisticians should learn from patterns of misuse occurring in the forensic sciences.

1. Failing to assess data provenance, validity, & reliability, or even if the data exist;
2. Stating results as certainties ("match"/"not match") not probabilities;
3. Conflating correlation with causation;
4. Basic errors in calculating probabilities, such as the wrong denominator, treating non-independent events as independent;
5. Forgetting about the base rate (priors; frequency with which the event occurs in a population);
6. Relevance ratio: frequency with which an event of interest appears in target vs non-target cases. e.g prevalence of subdural hematoma in child abuse vs non-abuse cases
7. Failure to compute the likelihood ratio in terms of both base rate prior and the relevance ratio (or medical evidence strength)
8. "Prosecutors fallacy" confusing P(evidence |hypothesis) with P(hypothesis|evidence)
9. Texas sharpshooter fallacy or HARK-ing (another reason why allocation concealment is essential for assessors)
10. Alternative-cause rarity and small n: if sample size is small and the event is very rare, the vent signal will not show up.
11. Simpson's paradox and missing data: Omitting key variables that can change or even reverse a conclusion that appeared obvious.
12. Going beyond claims of probability to claims of agency.

Refs
Altman DG 1980. Statistics and ethics in medical research. Misuse of statistics is unethical. *British Medical Journal*, 281(6249), 1182–1184.
Thompson WC, Shumann EL 1987. *Law & Human Behavior* 2 (3): 167. doi:10.1007/BF01044641
Findley KA 2021. *Dickinson Law Review* 615
https://ideas.dickinsonlaw.psu.edu/dlr/vol125/iss3/2

10 June Analysis of directional or circular data requires specialised methods for determining circular statistics. One of the earliest examples of a circular time plot is Florence Nightingale's coxcomb diagram in 1858. James Thomson father of Lord Kelvin introduced the term 'radians'. It shows up in the 1873 June exams he set for his students at Queens College Belfast. Lord Rayleigh first described the Rayleigh distribution 1880.

However, practical and specialised statistical techniques for handling circular data have only been developed in the last 60 years. Circular data occur more commonly that would be deduced from statistics service course offerings. Examples include compass directions, animal movements & navigation, time cycles & chronobiology problems, protein conformation.

Refs
Rayleigh L. 1880 On the resultant of a large number of vibrations of the same pitch and of arbitrary phase. *Philos Mag*. 10:73–78. doi: 10.1080/14786448008626893.
Batschelet E 1981. *Circular statistics for biology*. Academic
Rao JS. 1972. *Zeitschrift für wahrscheinlichkeitstheorie und verwandte gebiete*, 22, 33-44.
Hamelryck T et al.2006. *PLoS Comput. Biol*., 2(9): e131 [proteins]
Fisher NI 1993 *Statistical analysis of circular data.* Cambridge

Guerra P, et al. Discordant timing between antennae disrupts sun compass orientation in migratory monarch butterflies. *Nature Commun* 3, 958 (2012). https://doi.org/10.1038/ncomms1965

11 Jun OTD 1881 Hilda Hudson b (d 26 Nov 1965). UK. OBE. Best known to statisticians for developing the classic SIR model of epidemic infectious disease with Ronald Ross 1916-17, she also pioneered application of sophisticated mathematics to aeronautical engineering. She was the first female invited speaker at International Congress of Mathematicians 1912. Semple called her "a distinguished mathematician of great erudition and integrity".

Hudson sat the Cambridge Mathematical Tripos in 1903-4 when women were not officially ranked. They were told only they placed "equal to the n^{th} man", or "between n^{th} & $(n+1)^{st}$ man (This disparity was not changed until 1948). Passing this prestigious but brutally punishing exam meant huge kudos and opportunities. Her mark was equivalent to 7^{th} wrangler ('wrangler' was the name given to students placed in the First Class).

During WWI, she joined the Admiralty to head the Structural Analysis section. Working with Letitia Chitty (1897-1982) & Beatrice Cave-Browne-Cave (1874-1947), they revolutionised fixed-wing aircraft design by applying high-level maths to solving problems of aerodynamics and aircraft structural integrity. After the war, they collaborated on writing the classic Handbook of Strength Calculations.

Royle calls them "pioneers who demonstrated that it was possible for women to overcome the dogmatic, institutionalized prejudices of the time, and they earned the right to stand tall as credible applied mathematicians during the genesis of aircraft stress analysis".

Refs

Ross R 1916 *Proc Royal Soc A* 92; Ross R, Hudson H 1917. *Proc Royal Soc A* 93 (650): 212–225; *Proc Royal Soc B* 89(621):507

Hudson HP 1912. On binodes and nodal curves. *Proc Fifth Intern Congress Math II*: 118-121

Royle T 2017. *Historia Mathematica*, 44(4): 342–366.

Barrow-Green J, Royle T 2022 The work of British women mathematicians during the First World War. In: Jones CG, Martin AE, Wolf A. (eds) The *Palgrave Handbook of Women and Science since 1660*. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-78973-2_26

12 Jun The first population census in North America was performed in the winter of 1665-6 by Jean Talon ("Canada's first official statistician"). He was appointed by Louis XIV of France to improve the management of French colonies in what is now Canada. He needed to measure the population to "gauge the progress of European colonization" and determine how best to implement policies to diversify the economy and strengthen governance.

The census occurred during a hard winter, which actually was an advantage for door-to-door enumeration (possibly a first) as most people had to stay home and could not travel. Talon proudly reported 3,418 people in total. However, there were many omissions, duplications, and mistakes in adding up. To be fair, this was a monumental achievement – a census of this kind had never been

attempted, there was certainly no model for how it should be performed, and then as now, people were reluctant to present themselves to be counted. In 1995 (300 years later) Canadian historian Marcel Trudel (1917-2011) published his own record meticulously assembled from contemporary parish & land grant registers. He found 4,219 people had been in Talon's colony in 1666.

Refs
https://canadas-history.myshopify.com/products/april-may-2021

13 Jun OTD 1876 William Sealy Gosset b (d 16 Oct 1937) UK. Best known for developing Student's t-statistic in 1908 while working for Guinness Brewery. He is probably the first modern industrial statistician.

Guinness had almost doubled beer production between 1887 and 1914 so consistent quality was a concern. The problem was determining how representative a small sample might be of the whole batch. Gosset was put on the problem because he had studied maths at Oxford, so was "less scared of this kind of problem than the other brewers".

Fun fact: Gosset published under the pseudonym "Student" as Guinness employees were not allowed to publish under their own names. One story has it C.D. LaTouche, the managing director of Guinness gave him the choice of "A Student" or "A Pupil'. Ziliak suggests it may have originated from the title on the cover of his lab notebook: "The Student's Science Notebook".

14 Jun Gosset's 1908 paper, now described as "path-breaking" received little, if any, notice at first. Gosset was a chemist, not a mathematician, so he struggled with proofs. He "guessed" (his words) the correct form of the t-distribution (at this time called z) based on "properties of correlation coefficient and Professor Pearson's types of frequency curves". But he needed to know the "frequency distribution of $r(\sigma_x/\ \sigma_y)$ for small samples, in my work I want that more than the r distribution now happily solved".

Fisher was a student at Cambridge when he came across 'Student's' paper. He was the first to recognise its importance, not only for experimental scientists working with small samples, but also because it meant "a large number of sampling problems were within reach of mathematical solution". He went to his tutor FJM Stratton to discuss a discrepancy he found between Gosset's results and his own. Fortunately, Stratton had already met Gosset so he knew who 'Student' was. Gosset had visited Cambridge previously on behalf of Guinness to learn more about agricultural experimentation methods. Stratton suggested Fisher contact Gosset directly, which he did.

In a letter to Pearson dated 12 Sep 1912, Gosset said "this chap Fisher produced a paper giving a new criterion of probability or something of the sort". Fisher had provided several pages of proof for Gosset's formula and proved using n-dimensions that the correct formula required the denominator should be (n – 1) rather than n (later degrees of freedom).

Intending to get his head around it ("I couldn't understand his stuff"), Gosset had taken it "up to the Lakes" and lost it. A shorter version was in the letter

Gosset sent to Pearson in 1912. Gosset said he was passing it along to Karl Pearson because "it's so nice and mathematical that it might appeal to some people."

Fisher greatly admired Gosset, calling him "one of the most original minds in contemporary science". However, even when the "Student tables" were in constant use at Guinness, and in spite of Fisher's enthusiasm, Gossett claimed not see their universal importance. He sent Fisher a copy "as you are the only man that's ever likely to use them!"

The t-statistic was originally called z. Gosset switched to calling it t after a visit with Fisher in 1922. In 1924-5, Fisher formally demonstrated the relationship between normal, χ^2, z, and t distributions, gave the proof of Student's results, and showed how "Student tables" could be applied to hypothesis tests. "Student's" t made its first public appearance in Fisher's *Statistical Methods for Research Workers* (1925). Eisenhart said the "decision to shift from the z- to the t-form originated with Fisher, but the choice of the letter 't' to denote the new form was due to 'Student'."

Fun fact: Fisher fancied himself as a bit of a farmer (although his wife did most of the work) and asked Gosset for advice on learning about home-brewing. Gosset said "It was less trouble to buy Guinness and let us do it for you". Fisher abandoned the home brewing idea.

15 Jun Gosset was interested in the practical application of statistics when the "problem is to get results as quickly and as cheaply as possible". In his now-famous 1908 paper he proposed a criterion for the practical significance of a result as 3 times the probable error of the normal curve. This would be roughly a one-sided p-value of 0.02 (1 in 50 odds), because probable error is ~0.67456 of a standard deviation.

In 1927 he weighed in on what he called "the vexed question of the repetition and rejection of results". He works out correlated errors in a sequential small-n data series, makes recommendations for identification of discordant observations & suggests some procedural rules. He also introduced his famous aide memoire (platypus = platykurtic, kangaroos 'lepping' =leptokurtic) for describing skew distributions.

Refs

Student. 1908 The probable error of a mean. *Biometrika* 6(1): 1-25
Student. 1927 Errors of routine analysis. *Biometrika* 19(1/2) 151-164
Box JF 1978. *RA Fisher: The life of a scientist*.
Box JF 1981 Gosset, Fisher, and the t Distribution. *The American Statistician* 35(2): 61-66
Box JF 1987. *Stat Sci* 2(1): 45-52
Eisenhart C 1979. *The American Statistician* 33(1):6-10.
Fisher RA 1938 *Annals of Eugenics* 9(1): 1-9
Lehmann EL 1999. *Stat Sci* 14(4): 418-426
Ziliak ST 2008 *Journal of Economic Perspectives* 22(4), 199–216
Guinness Archive Fact Sheet wsgosset-and-students-t-test Guinness.pdf

16 Jun OTD 1915 John Tukey b (d 26 Jul 2000) USA. ASA Fellow 1949, US National Medal of Science 1973. First to distinguish between exploratory & confirmatory data analysis, he "loathed" much of the field of theoretical statistics for failing to inform data analysis. However, he was in fact an accomplished mathematician. He made important contributions to time series, Fast Fourier transform, spectral analysis, ANOVA and simultaneous inference. He is more widely known for his major contributions to computing. He coined the term 'bit' for 'binary digit', first published use of the word 'software' and developed PRIM-9, the first interactive computer graphic program. He saw data visualisation as central to exploratory data analysis.

Refs

Tukey 1962 The future of data analysis *Ann. Math. Statist*. 33(1): 1-67
Fernholtz LT 2003. *Stat Sci 18(3) 336–340*
Friedman JH, Stuetzle 2002 *Annals Stat* 30(6) 1629-39

16 Jun [originally posted 17 Jun] OTD 1916 David B Duncan b (d 12 Jun 2006). Australia/USA. ASA Fellow 1962. Best known for his citation classic Duncan Multiple Range test (1955) cited almost 40K times, and his k-ratio t-test (1992), the first Bayesian multiple comparison procedure. He was an early advocate of Kalman filters for dynamic estimation in time series problems. With Strother Walker, he was among the first (if not the first) to advocate logistic regression rather than linear regression for analysis of binary responses

Refs

https://publichealth.jhu.edu/departments/biostatistics/about/history/david-duncan

17 Jun OTD 1994 Frank Yates d (b 12 May 1902) UK. Another giant of experimental design. Guy Gold Medal 1960, Royal Society Medal 1966. Started his career at Rothamsted 1933 as assistant to Fisher. After Fisher left for University College London UCL, he became Head of Statistics until his retirement in 1968. He worked on survey design, sampling methodology, ANOVA, and developed Yates's algorithm for computing contrasts. He also actively promoted the early user of computers for statistical analysis and was the founder and President of the British Computer Society.

Unfortunately, he introduced the use of asterisks to indicate levels of statistical significance, only to later strongly condemn the monster he had created: "It was all too often forgotten that the important thing about a difference was not its significance but its size"

Finney called him "one of the greatest statistical scientists of his generation" but a "less than ideal lecturer, for he lacked concern for comprehensive formal presentation and preferred to talk about general ideas".

Refs

Healy MJR 1995. *International Statistical Review / Revue Internationale de Statistique,* 63 (3), 271-288
Yates *Annals of Eugenics* 6(2): 202-213
Finney DJ 1995 *Biogr. Mems Fell. R. Soc.* 41: 554–573
http://doi.org/10.1098/rsbm.1995.0033

Finney DJ 2007 Remember a pioneer: Frank Yates (1902-1994). *Teaching Statistics* https://doi.org/10.1111/j.1467-9639.1998.tb00748.x

18 Jun OTD 1988 Archibald ("Archie") Cochrane d (b 12 Jan 1909). UK. CBE. Father of evidence-based medicine & modern clinical epidemiology. His first randomised controlled trial was conducted when he was a German POW. His data persuaded the camp commandant to improve the prisoners' diet. His 1972 book *Effectiveness and Efficiency* has been described as 'seminal" for promoting RCTS to assess benefit or harm of health service interventions. In 1979, his concerns on the lack of valid organized summaries of medical evidence resulted in formation of the UK Cochrane Centre, Oxford 1992 and later the International Cochrane Collaboration and Cochrane Library. https://www.cochrane.org/

Refs
Cochrane AL *BMJ* 1984;289:1726-1727.
Cochrane AL, Holland WW 1971. Validation of screening procedures. *British Medical Bulletin* 27 (1): 3–8. doi:10.1093/oxfordjournals.bmb.a070810
For more references; James Lind Library: https://www.jameslindlibrary.org/?s=cochrane

Also, OTD 1910 Maurice S Bartlett b (d 8 Jan 2002). UK. Guy Silver (1952) and Gold (1969) Medals, President RSS 1966. FRS. Best known for Bartlett's test for homogeneity of variances, he also made important contributions to stochastic processes, time series, inference, and multivariate analyses. He introduced Monte Carlo methods to ecology and epidemiology.

Like so many other people, his relations with RA Fisher were fraught. "At one stage, I might be … in his good books; and another time, I'd be in his bad books because I was querying some point he had put forward". Fisher quit the Cambridge Philosophical Society after Bartlett published a paper in the Proceedings correcting some of Fisher's ANCOVA errors.

Refs
Besag J, Clifford P. 2002 IMS Bulletin (May/June) 31(3): 15
Whittle P. 2004. *Bio Mem Fellows Royal Soci*. 50: 15. doi:10.1098/rsbm.2004.0002
Wiiliams RH et al 2006. *Twelve British Statisticians*. Bitingduck Press

19 June OTD 1623 Blaise Pascal b (d 19 Aug 1662). Together with Fermat developed the foundational concepts for modern probability. This was revolutionary as mathematicians thought it was impossible to predict the probability of an event. What led to this was the result of the first statistical consultation on record 1654.

Antoine Gombaud (1607–1684), self-styled Chevalier de Méré, wrote to Blaise Pascal with probability problem. Except probability hadn't been invented yet. Like other clients since, he had tried and failed to solve the problem himself but was confident the solution would "only take a minute", or words to that effect. He also actually wanted answers to several questions other than the one he asked. Common games of chance in 18th c France involved betting on the event of getting at least one 1 (ace) in four rolls of a dice or getting at least one

double ace with 24 throws of a pair of dice. de Méré's gambling system assumed events should be equiprobable (the 'gamblers fallacy'), so of course he was losing a lot of money. He also posed the classic 'problem of points': what is the most fair way of dividing the stake between two players when the game was interrupted before the end?

Pascal wrote to Pierre de Fermat and in a series of letters over the next year they developed the foundational concepts for modern probability. Fermat solved the problem by tabulating all possible continuations. Pascal improved on it by using expected values instead of odds & manipulating triangular arrays of the binomial coefficients (Pascal's triangle). de Méré later went swanning about claiming he had discovered probability theory himself.

Pascal is also credited with the first early mechanical calculator 1642, and as a by-product, the roulette wheel. Also known for Pascal's triangle, work on binomial coefficients, conic sections and projective geometry. His *Pensées* is considered a masterpiece of French prose. However, his most famous quote is probably "The more I see of men, the better I like my dog".

Fun fact: Pascal was advised by his doctor to take up gambling for the sake of his health. Being a nerd's nerd, he tended to get rather extreme in his religious practices and also got stuck into his nerdy work, which had a bad effect on his already fragile condition (he was to die at 39) He first met de Méré when gambling.

Refs

Devlin K. 2010. The Unfinished Game: Pascal, Fermat, and the Seventeenth-Century Letter that Made the World Modern: A Tale of How Mathematics is Really Done

Letters: Fermat and Pascal on probability. https://www.york.ac.uk/depts/maths/histstat/pascal.pdf

Also OTD 1934 Jerzy Neyman (1894-1981) presents his paper 'On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection' to the Royal Statistical Society. This paper was groundbreaking in that it explicitly defined a design-based approach to sampling. The design determines selection probabilities ascribed to sampling units and therefore are independent of the values of the observed data [**posted 22 June 2023**]

Refs

Neyman J 1934. Journal of the Royal Statistical Society 97(4), 558-625

Bartlett MS 1984 *Bulletin of the London Mathematical Society*, 16(2):169-176

Kubiak AP, Kawalec P. 2022 *J Gen Philos Sci* 53, 381–402 https://doi.org/10.1007/s10838-022-09600-x.

Also OTD 1895 Charles P Winsor b (d 4 Apr 1951) USA. ASA Fellow 1949. Editor 'Human Biology'. Best known for developing the method of winsorization ('clipping' outliers by setting them to prespecified percentiles of the data), also for work on small samples, Gompertz growth curves & linear regression with error in both variables

Ref

Cochran WG 1951 Human Biology 23(2)
https://www.jstor.org/stable/41447995

20 Jun OTD 1901 Karl Pearson reads a paper to the Royal Society describing a series of experiments quantifying inter-observer measurement error using a device constructed by Horace Darwin. The term 'personal equation' was invented in the 18th c to refer to errors in astronomical observations then morphed into a label referring to both personal opinion & between-observer variation in measurements.

Ref
Pearson KP. 1902. Phil Trans Royal Soc Lond A 198, 235-299.
Canales J 2010 A Tenth of a Second: A History. Chicago

21 Jun OTD 1781 Siméon Poisson b (d 25 Apr 1840) France. Elected FRS and foreign member of the Royal Swedish Academy of Sciences. A mathematician and physicist, his name is associated with numerous concepts in maths, physics, optics, thermodynamics, probability, and sampling theory. Lagrange was his friend and Laplace regarded him as a son. Although his work found wide practical application, he himself was a completely theoretical mathematician. He was extremely clumsy and uncoordinated. As a student he failed geometry because of his inability to draw math diagrams. Although his father wanted him to be a surgeon it was soon obvious he would have been hopeless. He is supposed to have said "Life is good for only two things: doing mathematics and teaching it."

Ref
https://mathshistory.st-andrews.ac.uk/Biographies/Poisson/

23 Jun OTD 1912 Alan Turing b (d 7 Jun 1954). OBE, FRS. Best known for Bletchley Park cryptography work and breaking of the Enigma code, he was a pioneer in theoretical computer science & AI. His statistics papers on Bayesian methods, Good–Turing frequency estimation, & sequential analysis are less well known because they were only recently declassified.

Gay at a time when it was illegal, Turing was prosecuted in 1952, and forced to undergo chemical castration. His death by cyanide poisoning was ruled a suicide at the inquest. In 2009 PM Gordon Brown issued an official public apology on behalf of the govt; in 2013 he was pardoned by the queen.

Refs
Albers C 2017 The statistician Alan Turing. *NAW* 5/18 nr. 3 209-210.
Mardia KV Cooper SB 2012. Alan Turing and Enigmatic Statistics. *Boletim ISBrA*. 5(2): 2-7.

24 Jun OTD 1917 Joan Clarke Murray b (d 4 Sep 1996). UK. Best known for her work at Bletchley Park during WWII as the only female practitioner of Alan Turing's Banburismus code-breaking technique. At Cambridge she was a Double First in mathematics and a Wrangler (although like Hilda Hudson, her Wrangler status was not recognised because she was a woman and therefore not admitted to full academic membership). She was recruited in 1940 by Gordon Welchman who recognised her talent in an undergraduate geometry class. She became

Deputy Head of Hut 8 Cryptanalysis of Naval Enigma. Although she could not get promoted any further because of her gender, Hugh Alexander described her as "one of the best Banburists in the section". After the war, she became a recognised expert on 16th-17th century Scottish coinage. Because of the Official Secrets Act and secret classification of much of her work, her contributions are not completely known. Her portrayal in the 2014 film *The Imitation Game* has been widely criticised as being historically inaccurate.

Ref
https://mathshistory.st-andrews.ac.uk/Biographies/Clarke_Joan/

25 Jun On 30 Jun 1812 Pierre-Simon Laplace (1749-1827) presented his 'Théorie analytique des probabilités' to the Académie des Sciences. It describes both probability and statistical methods, plus applications. It has been said with commendable understatement that "his perspective slides back and forth between the Bayesian and non-Bayesian views with an ease that makes some of his investigations difficult to follow".

Ref
Stigler SM 1986 *The History of Statistics*
Stigler SM. 1975. *Biometrika.* 62 (2): 503–517. doi:10.2307/2335393.
Stigler, Stigler SM. 1978 *Isis* 69 (2): 234–254.

26 Jun The **quincunx** (best scrabble word ever) was invented by Francis Galton in 1873. A series of balls are dropped through the top & trickle through an array of pegs. Each ball hitting a peg has a 50-50 chance of falling L or R. Accumulation of balls at the bottom form a normal distribution.

It was used to demonstrate law of "deviation from an average", central limit theorem, random walks, & (much later) even properties of low-dimensional deterministic dynamical systems. Galton's first quincunx and notes are on display at the Galton Laboratories University College London

Fun fact: "quincunx" was a Roman coin produced during the Second Punic War. It was valued as 5/12 of a bronze libra, & represented by a pattern of five dots

Refs
Galton F 1894. *Natural Inheritance*. Macmillan
UCL Galton Collection https://www.ucl.ac.uk/culture/galton-collection/galtons-quincunx

27 Jun OTD 1884 President Chester Arthur signed a bill into law (23 Stat. 60) creating the US Bureau of Labor Statistics, one of the oldest statistical agencies in the US government. In 1902 BLS began publishing wholesale prices (later called Producer Price Indexes), in 1913 the Consumer Price Indexes, & in 1949 occupational statistics "wages & compensation, workplace accidents, strikes & lockouts… women's employment in particular industries, safety problems"

Ref
https://guides.loc.gov/this-month-in-business-history/june/founding-bureau-labor-statistics

28 Jun OTD 1858 Alice Lee b (d 5 Oct 1939) UK. Developed statistical models of variation in human cranial volume and its correlation with intellectual ability. Her PhD research demonstrated that there was no correlation between cranial

capacity and intelligence. As craniometry was used to "prove" the "science-based" inferiority of women, her work caused outrage, especially from Francis Galton and her examiners, some of whom (surprise!) ranked extremely low on her skull size charts. Paradoxically, she argued for biological differences between races based on the same methods.

Refs

https://www.smithsonianmag.com/science-nature/alice-lee-statistician-debunked-sexist-myths-skull-size-intelligence-180971241/work
Love R 1979. 'Alice in Eugenics Land' *Annals of Science* 36: 145-158

29 Jun OTD 1893 Prasanta Chandra Mahalanobis b (d 28 Jun 1972) India. OBE FRS, Hon Fellow Royal Statistical Society 1954, ASA Fellow 1961. Best known for the Mahalanobis distance, Feldman–Mahalanobis model, large-scale sample surveys and founder of the Indian Statistical Institute and the journal Sankhya.

Fun facts: He was honoured with a Google doodle on his 125th birthday 29 June 2018. His character is played by Shazad Latif in the 2015 movie '*The Man Who Knew Infinity'* about the genius mathematician Srinivasa Ramanujan.

Refs

Rao CR 1973. *Biographical Memoirs of Fellows of the Royal Society* 19, 454-492; 1973 *Internat Statist Rev* 41 (2) : 301-302.

30 Jun OTD 1923 James Durbin b (d 23 June 2012). UK. Guy Bronze, Silver, Gold medals, President RSS 1986-7. In 1950 he and Geoff Watson introduced their eponymous test for serial correlation in least squares regression. Durbin later said the bounds test concept came together in about 3 weeks & the mathematical details over a few months. The "horrendously difficult" calculations for the tables were computed by "eight or 10 young ladies operating desk calculators supervised by an older lady of forbidding demeanor".

Refs

Philips PCB 1988 *Economic Theory* 4: 125-157.
Koopman SJ 2012. *J. R. Statist. Soc. A* 175 (4):1060–1064