Excursions in Statistical History: Highlights

James Hanley

Dept. of Epidemiology, Biostatistics & Occupational Health
McGill University, Montréal, Québec, Canada

McGill University, Statistics Seminar

15:30-16:30 (Montreal time)
March 17, 2023

Burnside Hall 1104

or via Zoom link

[Meeting ID: 834 3668 6293 Passcode: 12345]

Possible Destinations

Galton	‘Transmuting’ women into men: Galton’s family data on human stature	2004	*TAS*
Gosset	Student’s z, t, and s: What if Gosset had R?	2008 ....Julien/Moodie	*TAS*
Gompertz Pearson Addison	Cultural imagery and statistical models of the force of mortality	2010 ....Turner	*JRSSA*
Farr	Age in plagues & pandemics:medieval Dances of Death/Pearson’s Bridge of Life	2010 ....Turner	*Significance*
Wilson FDR Nixon	Lest We Forget: U.S. Selective Service Lotteries, 1917-2019	2019	*TAS*
Gosset Rutherford Erlang	The ‘Poisson’ Distribution: History, Reenactments, Adaptations	2022 ....Bhatnagar	*TAS*

‘Transmuting’ women into men: Galton’s family data on human stature

‘Transmuting’ women into men: Galton’s family data on human stature

webpage, containing

- Galton’s 1884 letter to The Times thanking it for publicity, and announcing the prize winners in the 4 month ’crowd-sourcing’ data-collection effort
- The American Statistician (TAS) Article
- Poster
- Shiny App to implement Galton’s Height Forecaster
- Source of recovered data
- Images from Record of Family Faculties (album)
- Images of pages of Notebook with family data on height
- Datafile [csv]

“Transmuting” women into men: Galton’s data on human stature

SUMMARY

The first two regression lines, and the first correlations, were calculated by Francis Galton, in his work on heredity in sweet-peas and in humans. When ‘regressing’ the heights of adult children on those of their parents, Galton had to deal with the fact that men are generally taller than women— but without modern-day statistical tools such as multiple regression and partial correlation. This poster uses the family data on stature, which we obtained directly from Galton’s notebooks, to

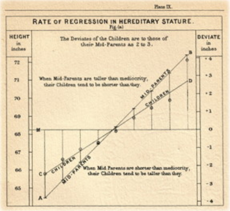
(a) compare the sharpness of his methods, relative to modern-day ones, for dealing with this complication;

(b) estimate the additional familial component of variance in stature beyond that contributed by the parental heights.

In keeping with Galton’s plea for “a manuscript library of original data”, these historical and pedagogically-valuable data are now available to the statistical community as digital photographs and as a dataset ready for further analyses.

Galton and Regression: An Introduction and Background

Galton defined *regression* as a reversion of a characteristic measured in offspring, *away* from the mean value of the *same* characteristic in *their own* parents, and *towards* the mean value in *all* parents/offspring. In his “regression line” (see Figure below), “the Deviates of the Children are to those of their Mid-Parents as 2 to 3” implying that “When Mid-Parents are taller than mediocrity, their Children tend to be shorter than they”, and conversely.



The contours of equal frequency in the two-way frequency table (see right) led Galton to the correlation coefficient of the bivariate Gaussian distribution. From these, Karl Pearson developed a full treatment of correlation, multiple and partial. Pearson’s early work relied on these family data, which “Mr. Galton, with his accustomed generosity”, had placed at Pearson’s disposal.

account for the results, contrasted in their values but concurrent in their significance, that are derived from comparisons between different groups of the returns.

An analysis of the Records fully confirms and goes far beyond the conclusions I obtained from the seeds. It gives the numerical value of the regression towards mediocrity in the case of human stature, as from 1 to $\frac{2}{3}$ with unexpected coherence and precision [*see* Plate IX, fig. (*a*)], and it supplies me with the class of facts I wanted to investigate—the degrees of family likeness in different degrees of kinship, and the steps through which special family peculiarities become merged into the typical characteristics of the race at large.

My data consisted of the heights of 930 adult children and of their respective parentages, 205 in number. In every case I transmuted the female statures to their corresponding male equivalents and used them in their transmuted form, so that no objection grounded on the sexual difference of stature need be raised when I speak of averages. The factor I used was 1·08, which is equivalent to adding a little less than one-twelfth to each female height. It differs a very little from the factors employed by other anthropologists, who, moreover, differ a trifle between themselves; anyhow, it suits my data better than 1·07 or 1·09. The final result is not of a kind to be affected by these minute details, for it happened that, owing to a mistaken direction, the computer to whom I first entrusted the figures used a somewhat different factor, yet the result came out closely the same.

I shall now explain with fulness why I chose stature for the

TABLE I.

NUMBER OF ADULT CHILDREN OF VARIOUS STATURES BORN OF 205 MID-PARENTS OF VARIOUS STATURES.

(All Female heights have been multiplied by 1·08).

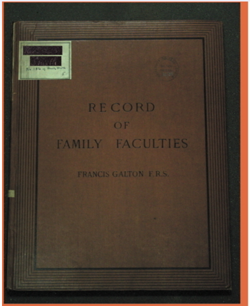
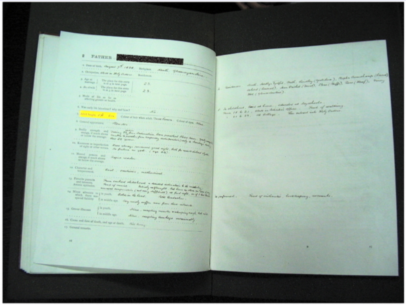
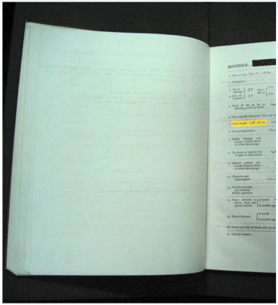
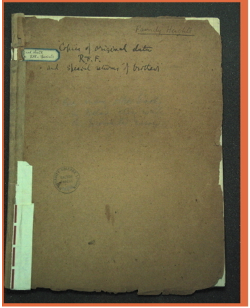
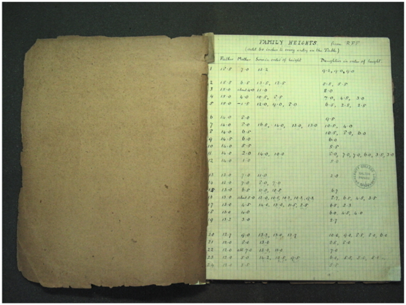
Heights of the Mid-parents in inches.	Heights of the Adult Children.														Total Number of		Medians.
	Below	62·2	63·2	64·2	65·2	66·2	67·2	68·2	69·2	70·2	71·2	72·2	73·2	Above	Adult Children.	Mid-parents.	
Above ..	..	..	..	..	..	..	..	..	..	..	..	1	3	..	4	5	..
72·5	..	..	..	..	..	..	..	1	2	1	2	7	2	4	19	6	72·2
71·5	..	..	..	..	1	3	4	3	5	10	4	9	2	2	43	11	69·9
70·5	1	..	1	..	1	1	3	12	18	14	7	4	3	3	68	22	69·5
69·5	..	..	1	16	4	17	27	20	33	25	20	11	4	5	183	41	68·9
68·5	1	..	7	11	16	25	31	34	48	21	18	4	3	..	219	49	68·2
67·5	..	3	5	14	15	36	38	28	38	19	11	4	..	..	211	33	67·6
66·5	..	3	3	5	2	17	17	14	13	4	..	..	..	..	78	20	67·2
65·5	1	..	9	5	7	11	11	7	7	5	2	1	..	..	66	12	66·7
64·5	1	1	4	4	1	5	5	..	2	..	..	..	..	..	23	5	65·8
Below ..	1	..	2	4	1	2	2	1	1	..	..	..	..	..	14	1	..
Totals ..	5	7	32	59	48	117	138	120	167	99	64	41	17	14	928	205	..
Medians ..	..	..	66·3	67·8	67·9	67·7	67·9	68·3	68·5	69·0	69·0	70·0	..	..	..	..	..

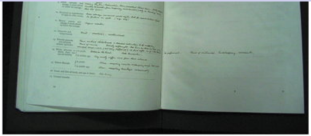
NOTE.—In calculating the Medians, the entries have been taken as referring to the middle of the squares in which they stand. The reason why the headings run 62·2, 63·2, &c., instead of 62·5, 63·5, &c., is that the observations are unequally distributed between 62 and 63, 63 and 64, &c., there being a strong bias in favour of integral inches. After careful consideration, I concluded that the headings, as adopted, best satisfied the conditions. This inequality was not apparent in the case of the Mid-parents.

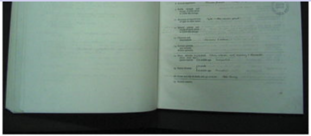
Two questions led me to pursue these same raw data which Galton placed at Pearson’s disposal:

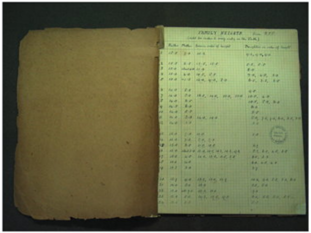
1. **How would today’s statisticians deal with the fact that men are generally taller than women?**

“Partialing out” the “effect” of sex; or “adjusting for sex in a regression model”, is conceptually like *adding* so many inches to the height of each female, or subtracting this amount for each male. In Galton’s analysis, “All female heights were *multiplied by 1.08*”; i.e., he “transmuted” them. I wished to test whether Galton’s ‘proportional’ scaling is a more biologically appropriate adjustment


RECORD
OF
FAMILY FACULTIES
FRANCIS GALTON F.R.S.

FATHER

MOTHER

Family Heights

FAMILY HEIGHTS
Father
Mother
Sons in order of height
Daughters in order of height







The 205 Families

	Min	Max	Sum	Mean
NUMBERS OF...				
Sons	0	10	487	2.4
Daughters	0	9	476	2.3
Sons + Daughters	1	15	963	4.7
NUMBERS FOR WHOM HEIGHT REPORTED AS A NUMBER...				
Sons	0	10	481	2.3
Daughters	0	9	453	2.2
Sons + Daughters	1	15	934	4.6

PRELIMINARY ANALYSIS

Role of Stature in Marriage Selection

TABLE 9.

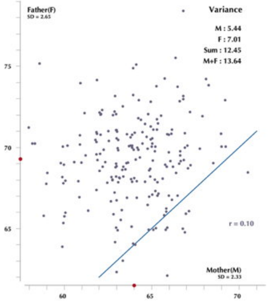
MARRIAGE SELECTION IN RESPECT TO STATURE.

S., t. 12 cases.	M., t. 20 cases.	T., t. 18 cases.
S., m. 25 cases.	M., m. 51 cases.	T., m. 28 cases.
S., s. 9 cases.	M., s. 28 cases.	T., s. 14 cases.

Short and tall, 12 + 14 = 32 cases.
Short and short, 9 } = 27 cases.
Tall and tall, 18 }

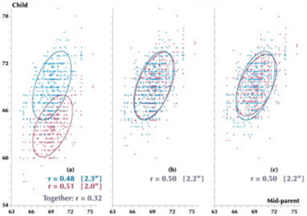
We may therefore regard the married folk as couples picked out of the general population at haphazard when applying the law of probabilities to heredity of stature.

T,M,S = Tall/Medium/Short men; t,m,s = tall/medium/short women.



ANALYSIS 1

"Transmuting" of Female Heights



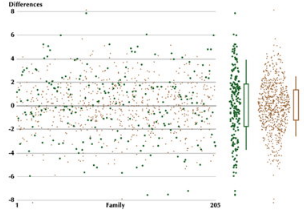
Heights (in inches) of adult children in relation to their mid-parent height. **(a)** each daughter's height 'as is' **(b)** daughter's height multiplied by 1.08 **(c)** 5.2 inches added to daughter's height. Daughters' heights are shown in pink, and sons' in blue, symbols. Ellipses (75%) are drawn based on the observed means and covariances.

In all three panels, and in analyses for "Do Residuals Segregate along Family Lines?", the mid-parent height is calculated as (father's height + 1.08 x mother's height) / 2.

[Average Residual, in inches]

ANALYSIS 2

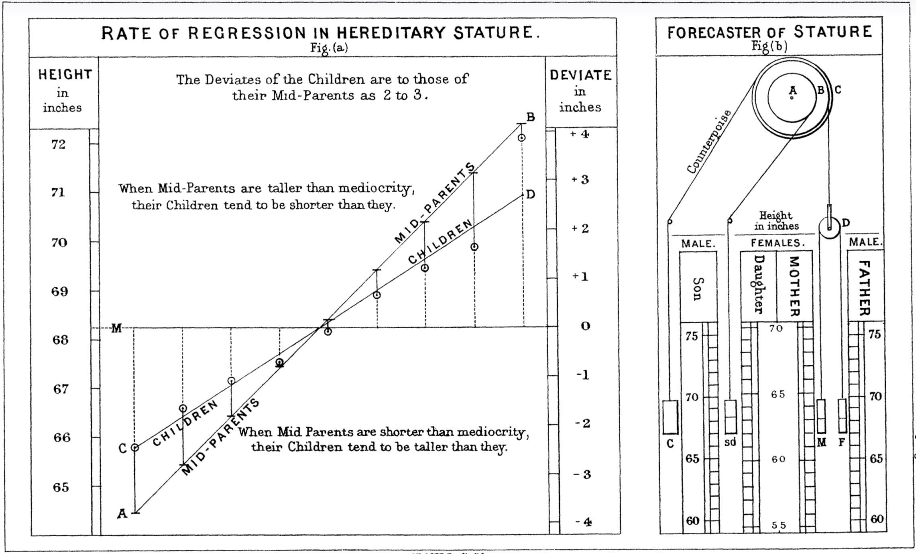
Do Residuals Segregate along Family Lines?



Distribution of within- and between-family residuals from simple linear regression, after daughters' heights have been multiplied by 1.08, of offspring height on mid-parent height. Families listed left to right, in same order as in Galton's notebook.

Larger green dot: the average residual for a family, multiplied by the square root of the number of offspring in the family, so as to put all 205 averages on the same scale. *Smaller brown dot*: orthogonal difference of within-family residuals (729 in all, from 172 families with two or more offspring). Marginal distributions shown on right. Boxplots show the 10th, 25th, 75th and 90th percentiles. **ICC = 19 %**

Plate IX.



J.P. & W.R. Emslie, lith.

Journ. Anthropolog. Inst., Vol.XV, Pl.IX


82
The Gazett

‘Student’s z, t, and s: What if Gosset had `R`?

'Student's z, t, and s: What if Gosset had R?

webpage, containing

- TAS article, 2008
- Supplementary Figures, with results for finger length
  Expanded version, 2008, of Fisher's Derivation of pdf(s, xbar)
- Frequency table of heights [22 bins] $\times$ finger-lengths [42 bins] of 3000 criminals, assembled by Macdonell (1901), and used in simulation by Student (1908)
- Excerpts from 'Student's' 1908 paper
- Gosset's 750 samples of size $n = 4$
- Census of Ireland 1911: Return completed by (a) Gosset (b) JH's grandfather
- Presentations
  - Statistical Society of Montreal & 2008 SSC Annual Meeting
  - Gosset Centenary (IBS/ISA) Dublin 2008
  - 2017 SSC Annual Meeting ("Gosset: Guinness, simulations & benefits of milk": video+lyrics)
- Photos
  - 2008 unveiling of plaque to Gosset at Guinness Storehouse in Dublin
  - IBS/ISA session at IBC2008

Gosset:

Guinness, simulations, and the benefits of milk

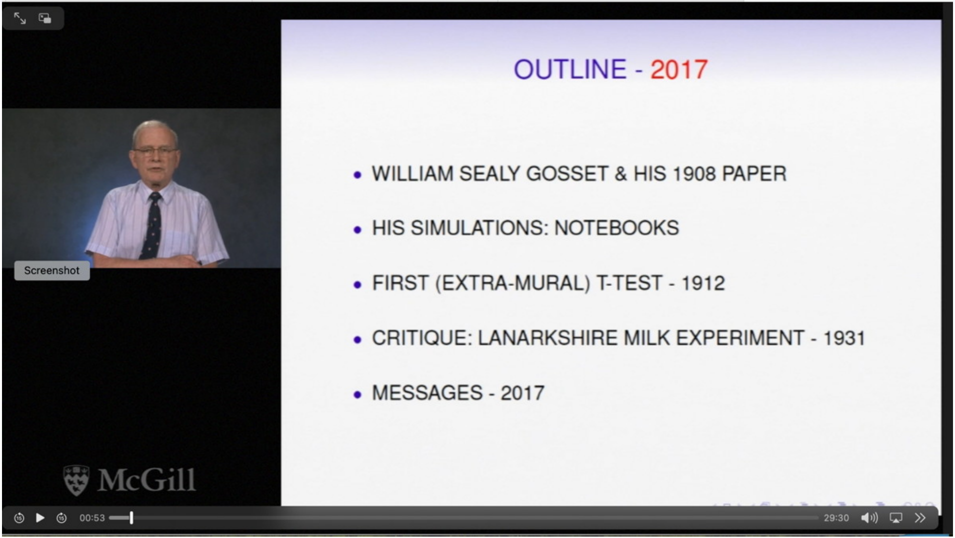
James A. Hanley

Department of Epidemiology, Biostatistics and Occupational Health,
McGill University

Session In honour of Gosset's birthday
Statistical Society of Canada Annual Meeting

Winnipeg, 2017.06.13

30-min presentation via pre-recorded video



Screenshot
McGill
OUTLINE - 2017
• WILLIAM SEALY GOSSET & HIS 1908 PAPER
• HIS SIMULATIONS: NOTEBOOKS
• FIRST (EXTRA-MURAL) T-TEST - 1912
• CRITIQUE: LANARKSHIRE MILK EXPERIMENT - 1931
• MESSAGES - 2017
00:53
29:30

Gosset's introduction to his paper

"Usual method of determining the probability that μ lies within a given distance of $\bar{x}$, is to assume ..."

$$\mu \sim N(\bar{x}, s/\sqrt{n}).$$

But, with smaller n, the value of s "becomes itself subject to increasing error."

Sampling distributions studied

$$\bar{x} = \frac{\sum x}{n}; \qquad s^2 = \frac{\sum (x - \bar{x})^2}{n}.$$

"when you only have quite small numbers, I think the formula with the divisor of $n - 1$ we *used to use* is better"
... Gosset letter to Dublin colleague, May 1907

Doesn't matter, "because only naughty brewers take n so small that the difference is not of the order of the probable error!"
... Karl Pearson to Gosset, 1912

$$z = (\bar{x} - \mu)/s$$

Forced to "judge of the uncertainty of the results from a small sample, which itself affords the only indication of the variability."

> *The method of using the normal curve is only trustworthy when the sample is "large," no one has yet told us very clearly where the limit between "large" and "small" samples is to be drawn.*
>
> *Aim ...*
>
> *"to determine the point at which we may use the (Normal) probability integral in judging of the significance of the mean ..., and to furnish alternative tables when [n] too few."*

Three steps to the distribution of z

Section I

- Derived first 4 moments of s^2.
- Found they matched those from curve of Pearson's type III.
- "it is probable that that curve found represents the theoretical distribution of s^2." Thus, "although we have no actual proof, we shall assume it to do so in what follows."

Section II

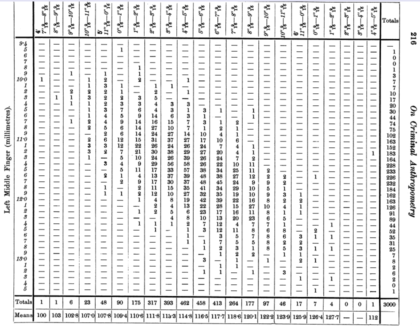
- "*No kind of* correlation" between $\bar{x}$ and s
- His proof is incomplete: see ARTICLE in *The American Statistician*.

Section III

- Derives the pdf of z:
  - joint distribution of $\{\bar{x}, s\}$
  - transforms to that of $\{z, s\}$,
  - integrates over s to obtain $pdf(z) \propto (1+z^2)^{-n/2}$.

Sections IV and V

- ..
- ..



216 *On Criminal Anthropometry*

Left Middle Finger (millimetres).

	4′7″$\frac{9}{16}$–8″$\frac{9}{16}$	8″$\frac{9}{16}$–9″$\frac{9}{16}$	9″$\frac{9}{16}$–10″$\frac{9}{16}$	10″$\frac{9}{16}$–11″$\frac{9}{16}$	5′11″$\frac{9}{16}$–0″$\frac{9}{16}$	0″$\frac{9}{16}$–1″$\frac{9}{16}$	1″$\frac{9}{16}$–2″$\frac{9}{16}$	2″$\frac{9}{16}$–3″$\frac{9}{16}$	3″$\frac{9}{16}$–4″$\frac{9}{16}$	4″$\frac{9}{16}$–5″$\frac{9}{16}$	5″$\frac{9}{16}$–6″$\frac{9}{16}$	6″$\frac{9}{16}$–7″$\frac{9}{16}$	7″$\frac{9}{16}$–8″$\frac{9}{16}$	8″$\frac{9}{16}$–9″$\frac{9}{16}$	9″$\frac{9}{16}$–10″$\frac{9}{16}$	10″$\frac{9}{16}$–11″$\frac{9}{16}$	6′11″$\frac{9}{16}$–0″$\frac{9}{16}$	0″$\frac{9}{16}$–1″$\frac{9}{16}$	1″$\frac{9}{16}$–2″$\frac{9}{16}$	2″$\frac{9}{16}$–3″$\frac{9}{16}$	3″$\frac{9}{16}$–4″$\frac{9}{16}$	4″$\frac{9}{16}$–5″$\frac{9}{16}$	Totals
Totals	1	1	6	23	48	90	175	317	393	462	458	413	264	177	97	46	17	7	4	0	0	1	3000
Means	100	103	102·8	107·0	107·8	109·4	110·6	111·8	113·3	114·8	116·5	117·7	118·6	120·1	122·2	123·9	125·9	126·4	127·7	—	—	112	

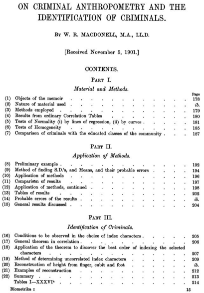
Left Middle Finger	Totals
9·4	—
5	1
6	0
7	0
8	1
9	3
10·0	7
1	7
2	10
3	17
4	20
5	30
6	44
7	74
8	75
9	102
11·0	163
1	152
2	183
3	164
4	228
5	233
6	226
7	232
8	184
9	162
12·0	163
1	126
2	91
3	89
4	44
5	52
6	35
7	31
8	25
9	7
13·0	8
1	2
2	6
3	2
4	0
5	1

Section VI: "Practical test of foregoing equations."

pdf's of s and z "are compared with some actual distributions"

Before I had succeeded in solving my problem analytically, I had endeavoured to do so empirically.

The material used was a correlation table containing the height and left middle finger measurements of 3000 criminals, from a paper by W. R. Macdonell (Biometrika, Vol. I, p. 219).



ON CRIMINAL ANTHROPOMETRY AND THE IDENTIFICATION OF CRIMINALS.

BY W. R. MACDONELL, M.A., LL.D.

[Received November 5, 1901.]

CONTENTS.

PART I.

Material and Methods.

The measurements were written out on 3000 pieces of cardboard, which were then very thoroughly shuffled and drawn at random.

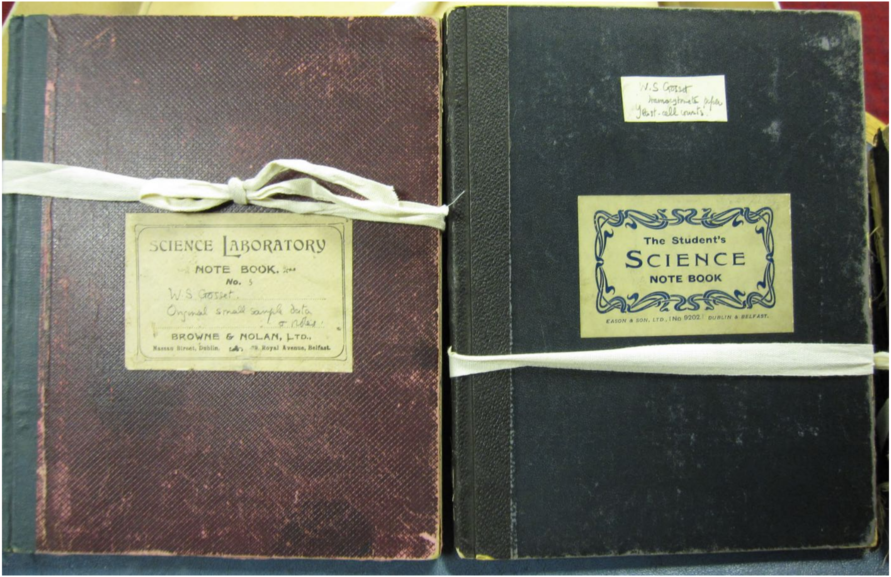
As each card was drawn its numbers were written down in a book, which thus contains the measurements of 3000 criminals in a random order.

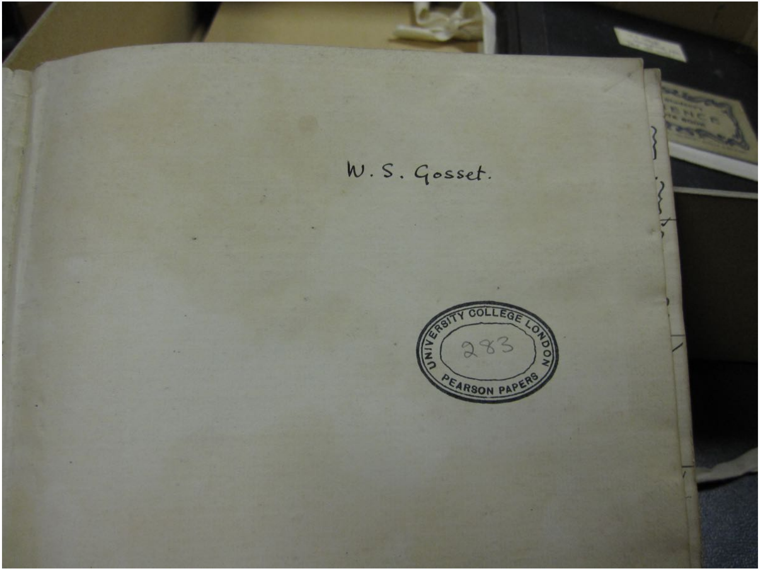
Finally, each consecutive set of 4 was taken as a sample – 750 in all – and the mean, standard deviation, and correlation of each sample determined.

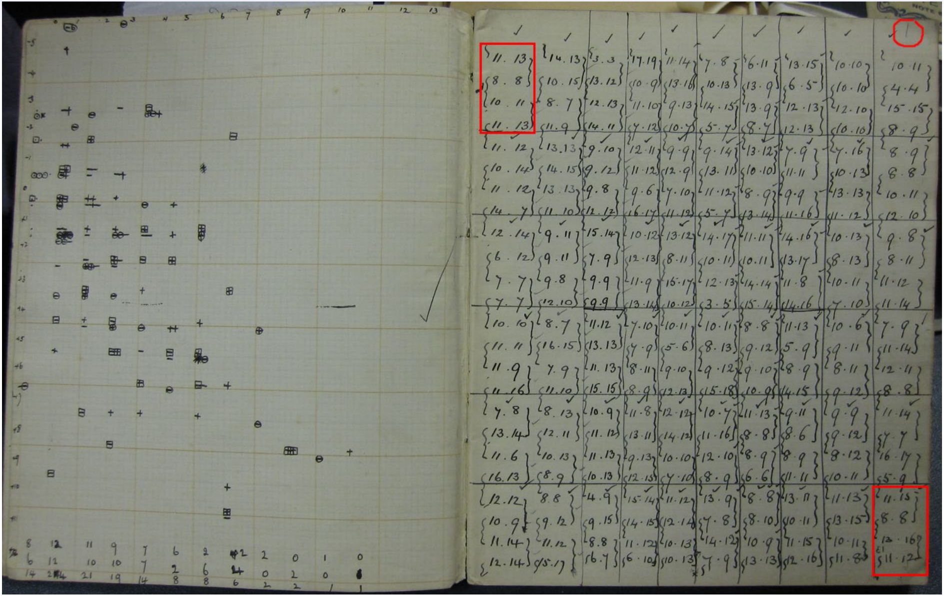
The difference between the mean of each sample and the mean of the population was then divided by the standard deviation of the sample, giving us the z of Section III.

This provides us with two sets of 750 standard deviations and two sets of 750 z's on which to test the theoretical results

2 of Gosset's notebooks (Pearson Collection, UCL)




W. S. Gosset.
UNIVERSITY COLLEGE LONDON
283
PEARSON PAPERS

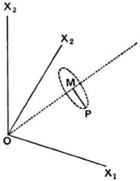


FOR...

- Results of our simulations, 100 years later
- Description of remainder of 1908 article
- Fisher's geometric vision
- Fisher and Gosset , and transition $z \rightarrow t$
- Backpack & Desktop Computers

SEE..

- SLIDES FROM LONGER VERSION OF TALK
- ARTICLE in *The American Statistician*
- `http://www.biostat.mcgill.ca/hanley/Student`



Gosset’s rucksack computer: Triumphator A ser 43219

http://www.calculators.szrek.com/

Fisher's desktop computer: Millionaire Ser 1200

BRIDGE OF LIFE: ANIMATIONS

described

in the article
Cultural imagery and statistical models of the force of mortality: Addison, Gompertz and Pearson.
Turner EL and Hanley JA. J. R. Statist. Soc. A (2010) 173, Part 3, pp. 483-499.

and in various presentations, e.g.,
presentation by JH to the Statistical Society of Montreal

BRIDGES OF LIFE -- after Addison, 1711

If your OS allows it, the following lifetables can be animated using the java app. This R code provides another (albeit slower) way:

* Sweden, Female, 1751-1851 (cohort)

frames/second: 1 5 15 25 100 or ... screenshot

* England, Male, 1871-1880 (current)

frames/second: 1 5 15 25 100 or screenshot or .mov

* France, 1895 (current) Male

* France, 1895-2004 (cohort) Male Female

* Switzerland, 1895-2004 (cohort) Male Female

* Canada, Female, 2000-2002 (current)

frames/second: 1 5 15 25 100 or ... screenshot

* PhD in Epidemiology & Biostatistics (1970-2002)

frames/second: 5 25

****

BRIDGES OF LIFE -- after Pearson, 1897

* England, Male, 1871-1880 (current)

frames/second: 1 5 15 25 .mov 100 500 .mov 2500

Updated: Jan 30, 2011

Cultural imagery and statistical models of the force of mortality

Elizabeth L. Turner[1] James A. Hanley[2]

[1]Medical Statistics Unit, Department of Epidemiology and Population Health
London School of Hygiene and Tropical Medicine

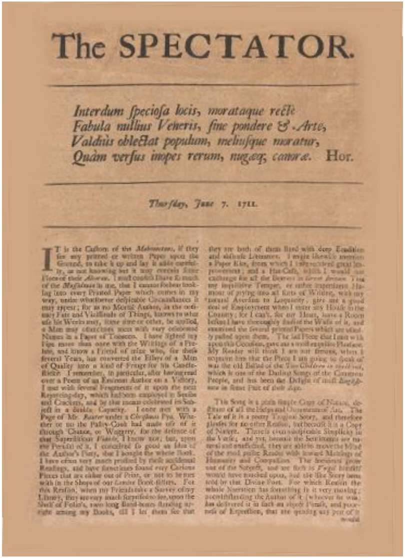
[2]Department of Epidemiology, Biostatistics and Occupational Health
McGill University

2010.11.19

In order of appearance...

- 1950-1976: hazard function and incidence density
- 1711: allegorical essay, non-mathematical
- 1825 & 1832 : intensity/force of mortality
- 1897: imagery back to 1400's; mixtures of pdf's
- 2009: computer animations

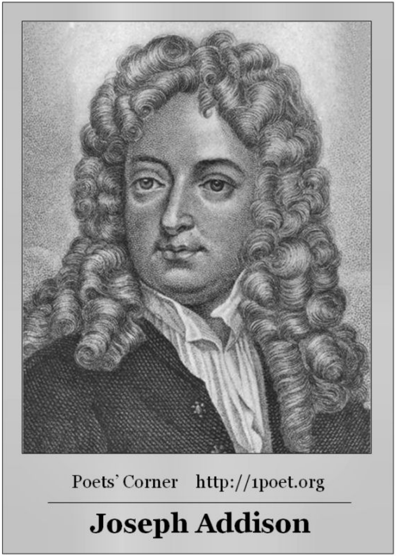
Force of mortality in *The Vision of Mirza*



The SPECTATOR.

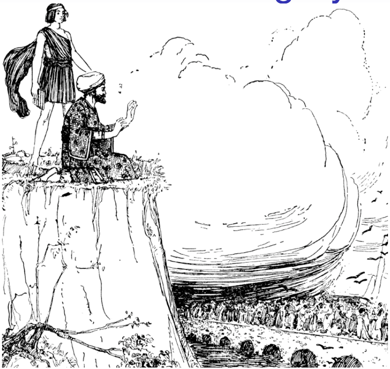
Interdum speciosa locis, morataque rectè
Fabula nullius Veneris, sine pondere & Arte,
Valdiùs oblectat populum, meliusque moratur,
Quàm versus inopes rerum, nugæq; canoræ. Hor.

Thursday, June 7. 1711.



(1672-1719)

Imagery in *The Vision of Mirza*



BRIDGE OF HUMAN LIFE

Multitudes of people passing over it

Passengers dropping through innumerable concealed trap-doors that...

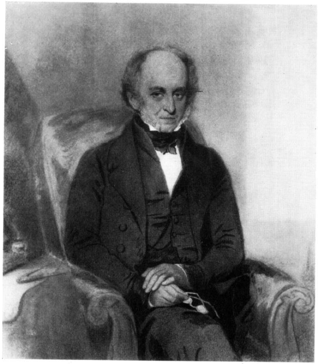
- were set very thick at the entrance
- grew thinner towards the middle
- multiplied and lay closer together towards the end

Allegorical essay

Sept. 1, 1711

`http://etc.usf.edu/clipart/17400/17410/mirza_17410.htm`

Gompertz' 'Law' of Mortality



BENJAMIN GOMPERTZ, 1779–1865

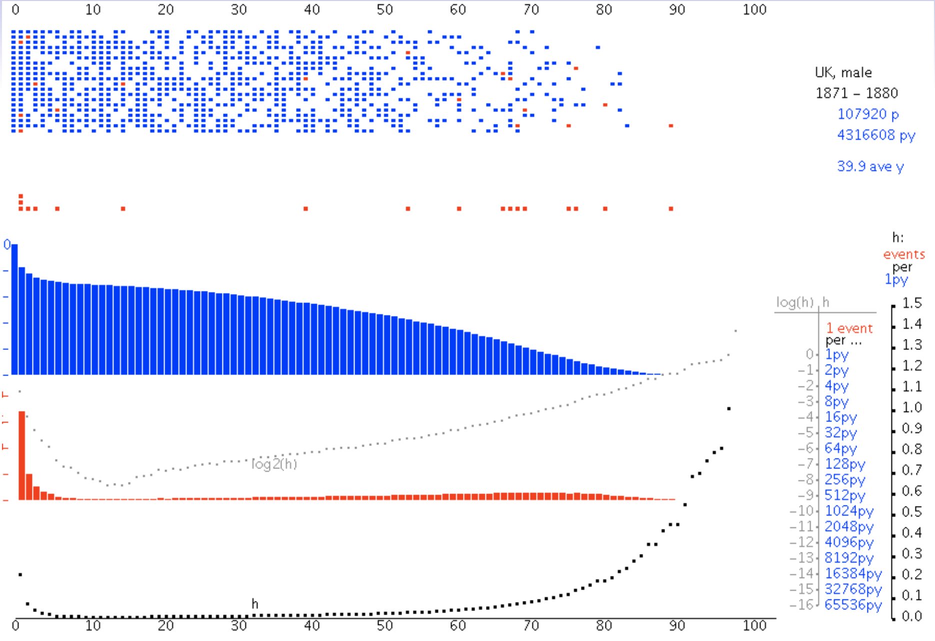
BENJAMIN GOMPERTZ. On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. *Phil. Trans. R. Soc. London, 115. (1825), 513-583*.

" [his] paper [...] opened up a new approach to the life table. Previously, the table had been regarded as little more than a record of the number of persons surviving to successive integral ages out of a given number alive at an earlier age; Gompertz introduced the idea that l_x [the survival function] was a function connected by a mathematical relationship with a continuously operating force of mortality."

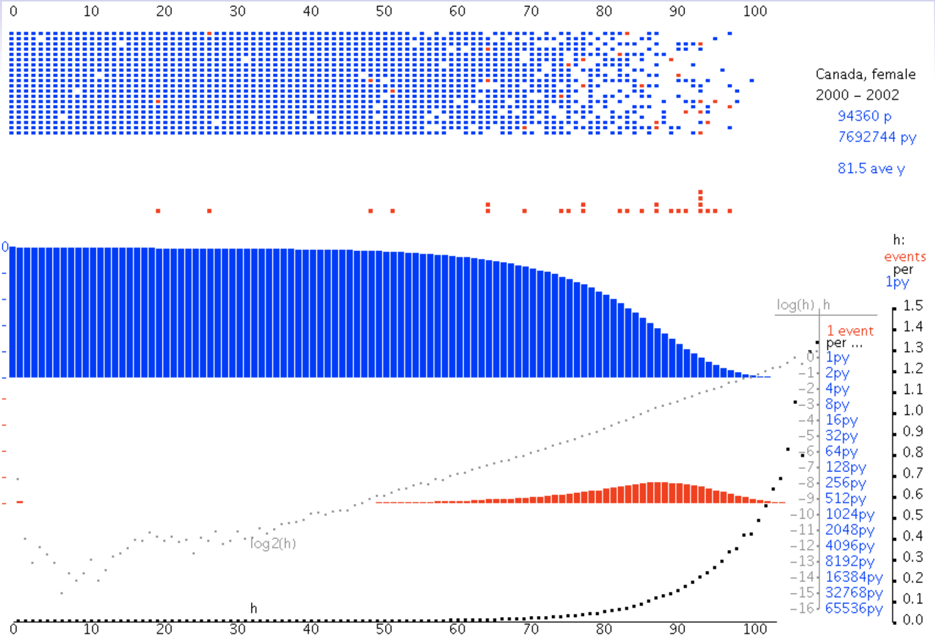
P. F. Hooker. J. Inst. Actuaries (1965).

Fitted survival function to life-table data.

Imagery: 'power of man to avoid death'

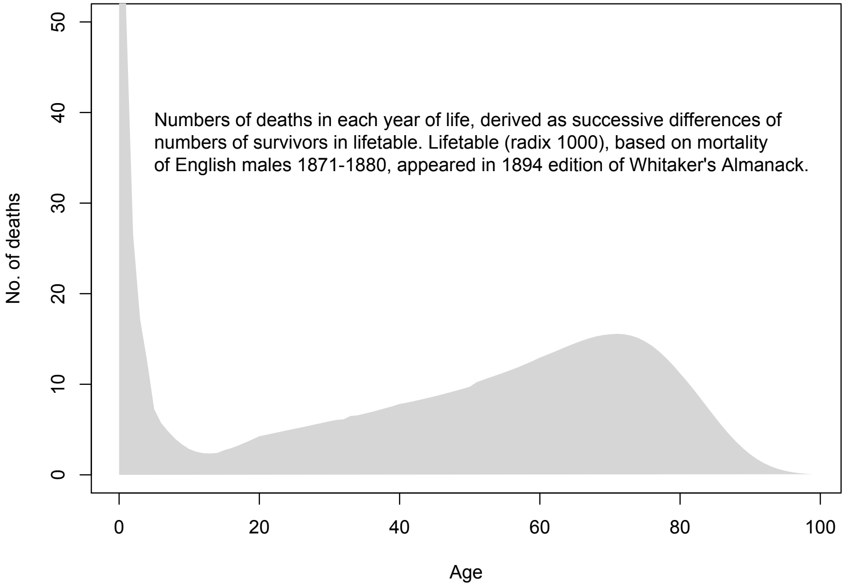


UK males 1871-1880 Addison, animated



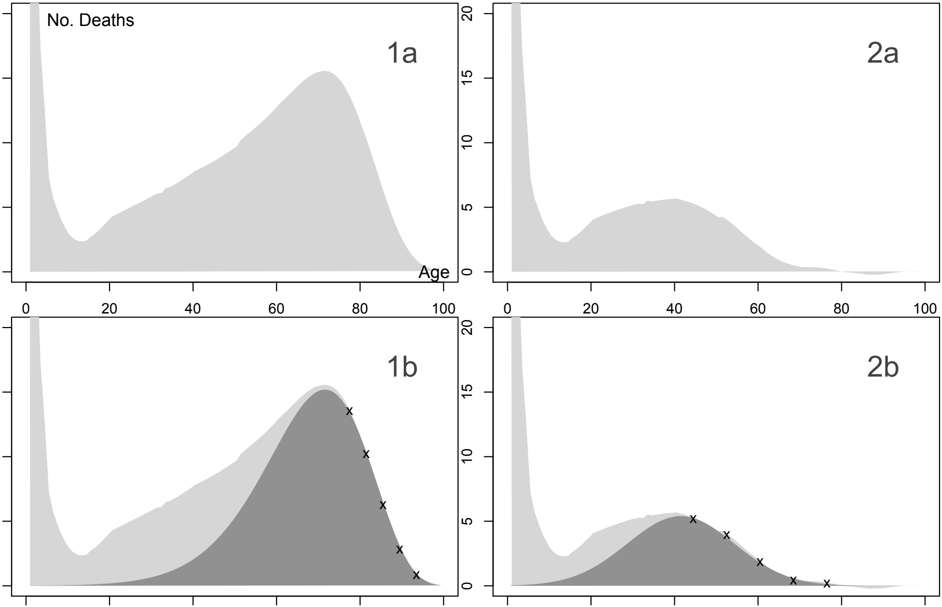
Canada females 2000-2002

Pearson's data: frequency distribution of age at death

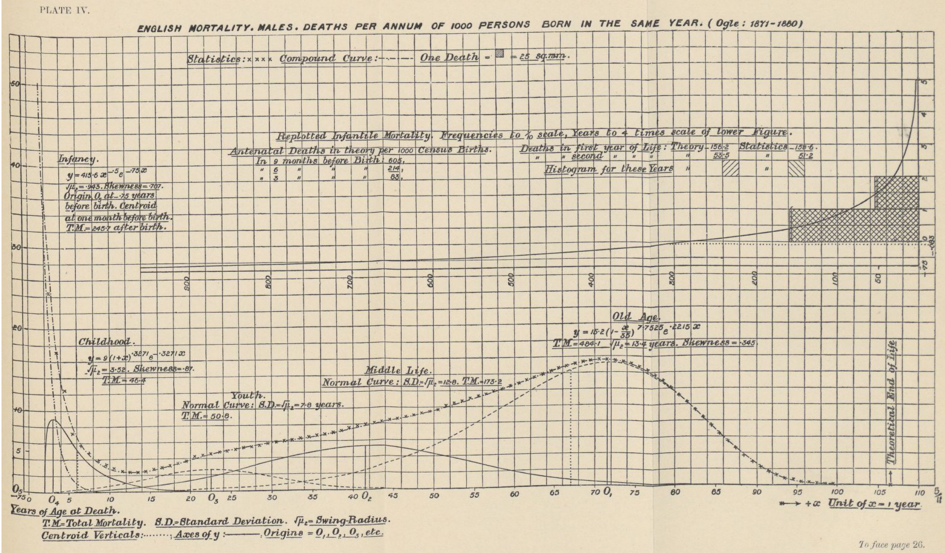


Source: Pearson, K. (1897) The Chances of Death and Other Studies in Evolution. London: Arnold.

How Pearson fitted the 5-component mixture



The full 5-component mixture



Rendered by Pearson’s wife, Maria Sharpe Pearson.

JH animation, 2009

EXCESS MORTALITY

1849 **Cholera**; 1918-19, 1957, 1968 **Flu**; 2020- **COVID-19**

https://jhanley.biostat.mcgill.ca/Pandemics/

EXCESS MORTALITY

1849 **Cholera**; 1918-19, 1957, 1968 **Flu**; 2020- **COVID-19**

https://jhanley.biostat.mcgill.ca/Pandemics/

Cultural imagery and statistical models of the force of mortality:
Addison, Gompertz and Pearson
Turner EL and Hanley JA. J. R. Statist. Soc. A (2010).

↓

Age in plagues and pandemics: medieval Dances of Death or Pearson's Bridge of Life?
E Turner and J Hanley, Significance, June 2010.

Age in medieval plagues and pandemics: Dances of Death or Pearson's bridge of life?

Death has long obsessed humanity. In times of plague and pandemic even more so. Medieval man saw four horsemen of the apocalypse, and of them, Death by disease was gathering the greatest harvest. How randomly did he gather? And how random is the death toll in later pandemics? **James Hanley** and **Elizabeth Turner** look at Karl Pearson's visualisations of mortality.

In October 1347 a trading ship from the Crimea with its crew dead and dying drifted into a harbour in Sicily, and black rats leapt ashore. The European phase of the Black Death had begun.

At the time they called it the Great Mortality, subsequently the "Great Pestilence" or the "Great Plague". Today we call it the "Black Death" and consider it as perhaps the deadliest pandemic ever to have struck humanity.

In the countryside, peasants dropped dead in the fields; in towns, the sick died too fast for the

That indeed was how it was seen at the time. Paintings and woodcuts depicted the "Dance of Death" – Death as a skeleton indiscriminately carrying off old and young, rich and poor, kings and commoners. A good life, a healthy life, a clean-lived life was no protection: the medieval folk-conception was of Death as one who obeys no rule of time, of place, of age, of sex, or of household. Five hundred years later a young Karl Pearson (1857–1936) viewed two of the 67 images painted inside the roof of the Spreuer

indiscriminate this distribution should resemble the (pyramid-like) shape of the living population just before the plague – many young, fewer adults, and fewer still who had reached old age. Margerison and Knüsel[6] found that the age-at-death distribution of those buried in the Royal Mint site, London, a Black Death cemetery of 1349, "coincides generally with what one would expect from" an age-indiscriminate Death. But DeWitte and Wood[7] compared the same skeletal remains with contemporary non-epidemic

JH Significance, 2010

DEATHS FROM CHOLERA
England and Wales, 1849

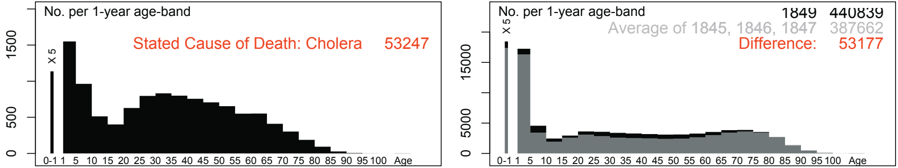


Figure 1. The age-specific numbers of cholera deaths for the year 1849, based on Farr's report, are shown on the left. The age-specific average numbers of deaths due to all causes for the years 1845–1847, along with those for 1849, are shown on the right, and based on data available in the online Human Mortality Database[12]

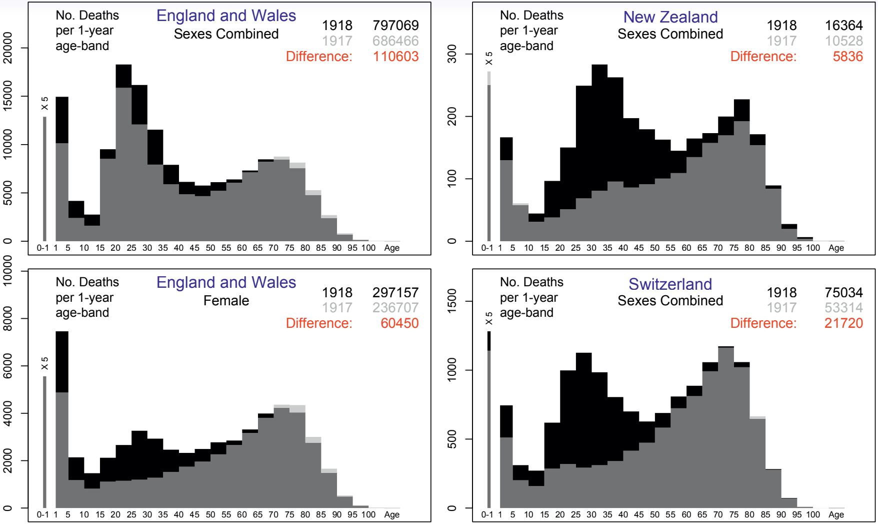


Figure 2. Deaths from all causes in 1918, the peak year of the Spanish Flu pandemic, compared to 1917. In the case of Switzerland, and of women in England and Wales, the excess deaths are entirely due to disease, not war. Data from the Human Mortality Database[12]

THE AMERICAN STATISTICIAN
2020, VOL. 74, NO. 2, 197–206: History Corner
https://doi.org/10.1080/00031305.2019.1699444

Lest We Forget: U.S. Selective Service Lotteries, 1917–2019

James A. Hanley

Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, QC, Canada

ABSTRACT

The United States held 13 draft lotteries between 1917 and 1975, and a contingency procedure is in place for a selective service lottery were there ever to be a return to the draft. In 11 of these instances, the selection procedures spread the risk/harm evenhandedly. In two, whose anniversaries approach, the lotteries were problematic. Fortunately, one (1940) employed a "doubly robust" selection scheme that preserved the overall randomness; the other (1969) did not, and was not even-handed. These 13 lotteries provide examples of sound and unsound statistical planning, statistical acuity, and lessons ignored/learned. Existing and newly assembled raw data are used to describe the randomizations and to statistically measure deviations from randomness. The key statistical principle used in the selection procedures in WW I and WW II, in 1970–1975, and in the current (2019) contingency plan, is that of "double"—or even "quadruple"—robustness. This principle was used in medieval lotteries, such as the (four-month) two-drum lottery of 1569. Its use in the speeded up 2019 version provides a valuable and transparent statistical backstop where "an image of absolute fairness" is the over-riding concern.

ARTICLE HISTORY
Received June 2019
Accepted November 2019

KEYWORDS
Datasets; History; Multiple-robustness; Randomness; Teaching

1. Introduction

The draft lottery of 1917 was—in the words of the then US War Department Secretary Newton D. Baker—the "first application of a principle believed by many of us to be thoroughly democratic, equal and fair in selecting soldiers to defend the national honor abroad and at home." New statistical evidence presented below shows that the 1917–1918 lotteries were successful in spreading the risk/harm as evenhandedly as possible: no (identifiable a priori) subgroup bore more of the burden than would be expected.

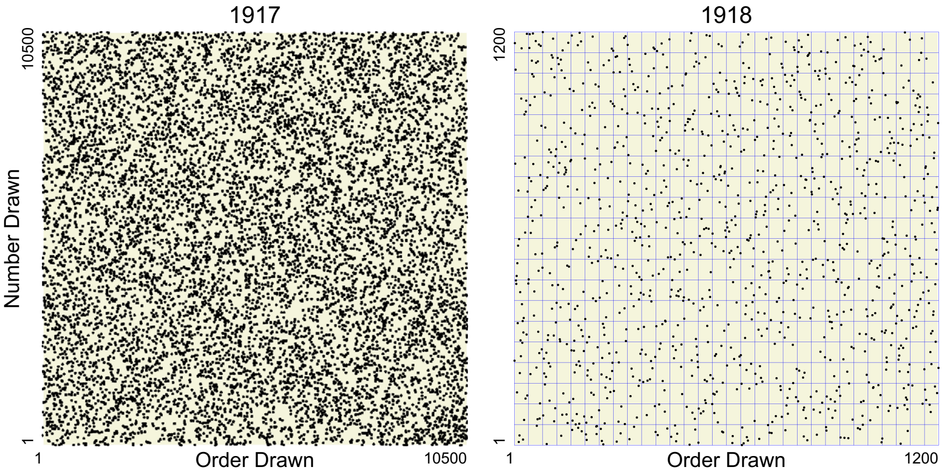
We are now approaching the 50th anniversary of a December an opportunity to consider the statistical ingredients for a fair process, and to examine the contingency procedure currently in place for a selective service lottery were there to be a return to the draft today.

These lotteries are big-ticket examples of sound and unsound statistical planning, lessons learned/ignored, and the central role of statisticians and statistical analyses. They also provide some interesting teaching perspectives. I begin with the meticulously planned "lower-tech" doubly robust lottery of 1917 and end with the also doubly robust but "high-tech" plan in place as of 2019. In between, I show the high resolution version of the 1940 lottery data, as well as a high-resolution photograph—not widely avail-

Table 1. WW I and WW II lotteries.

Year	Age/born	Registration day	Millions registered	Lottery date	Numbers drawn, 1–	Duration (hr)
1917	21–30	June 5	10	July 20	10,500	16
1918	21*	June 5	0.7	June 27	1200	2
1918	18–45	September 12	13	September 30	17,000	18
1940	21–30	October 16	16	October 29	9000	14
1941	21*	July 1	0.75	July 17	800	2
1942	20–45	February 14–16	9	March 17	7000	13

*Had reached 21 since previous registration.


1917
10500
Number Drawn
1
1
Order Drawn
10500
1918
1200
1
1
Order Drawn
1200

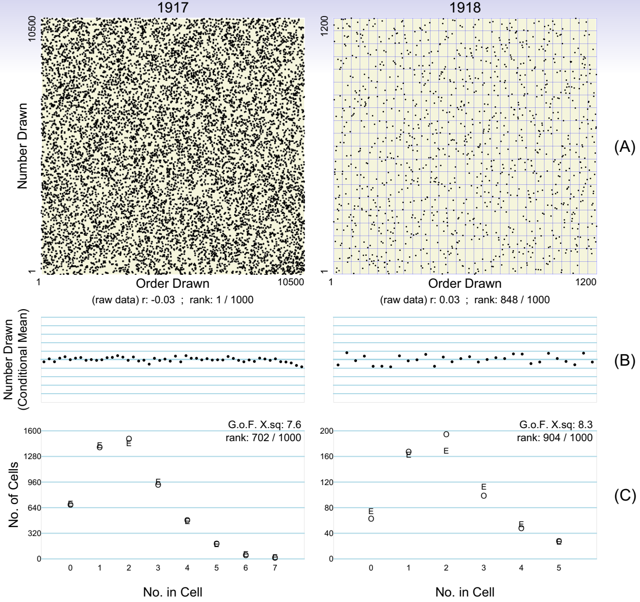


Figure 1. Scatterplots (A), conditional means (B), and observed and expected cell frequencies (C) based on raw data from WW I lotteries. Instead of using them to esti
a *p*-value, the correlations in 1000 simulated lotteries were ranked from smallest (1) to largest (1000), and the reported rank of the observed correlation is its positi
this array; thus a "rank" of 0 means that the observed correlation was smaller than all 1000 simulated correlations. The vertical ranges in (B) are the same as in (A), na
1 to 10,500 and 1 to 1200, and the means are conditional on the "*x*" bins. The 5250 and 600 cells, respectively, used in (C) were formed by binning the *x* and *y* axes i
so as to have rectangles ("cells") with a mean of 2 dots per cell (see cells formed using 1918 lottery data). O and E: Observed and expected frequencies, both summi
5250 or to 600. E's and ranked goodness of fit (G.o.F) statistics are based on the 1000 simulated lotteries.

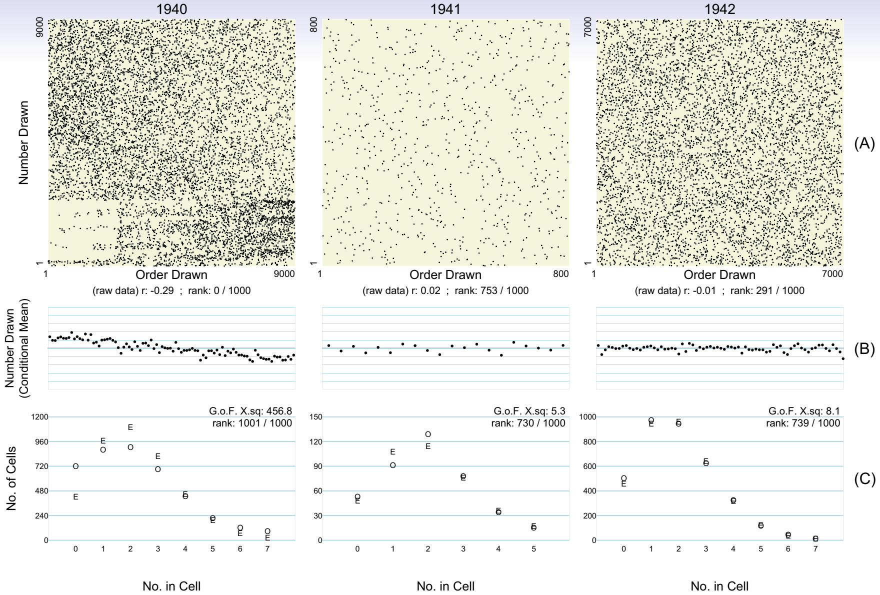
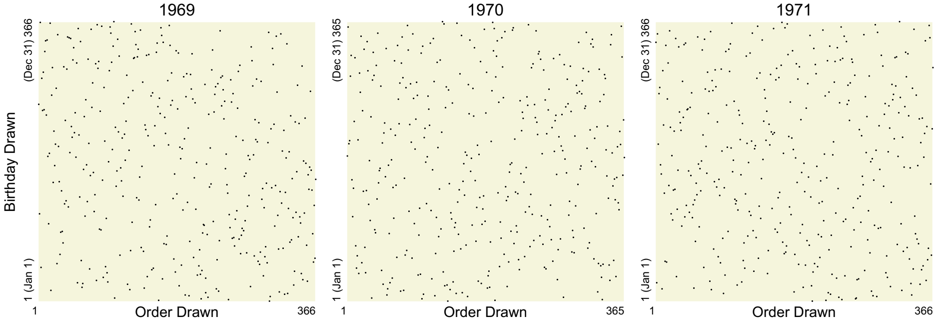


Figure 2. Scatterplots (A), conditional means (B), and observed and expected cell frequencies (C) based on raw data from the WW II lotteries. Explanations as in Figure 1. In (C) a "rank" of 1001 meas that the observed G.o.F statistic was larger than the 1000 simulated G.o.F statistics.

Figure 3. Drawing of the fourth number in the 1940 lottery by the (blindfolded) Secretary of the Navy, Frank Knox, with President Franklin Roosevelt (left) looking on. The image is licensed from *http:www.alamy.com*. The problems caused by the ad-hoc extension to the bowl used in 1917 are easily seen if one looks carefully at the capsules near the bottom of the bowl. Photographs of the drawings of the first and fourth numbers are also available at the Library of Congress: *https://www.loc.gov/item/2012648302/* and *https://www.loc.gov/item/2004671493/*.


1969
1970
1971
Birthday Drawn
(Dec 31) 366
1 (Jan 1)
(Dec 31) 365
1 (Jan 1)
(Dec 31) 366
1 (Jan 1)
Order Drawn
1
366
1
365
1
366

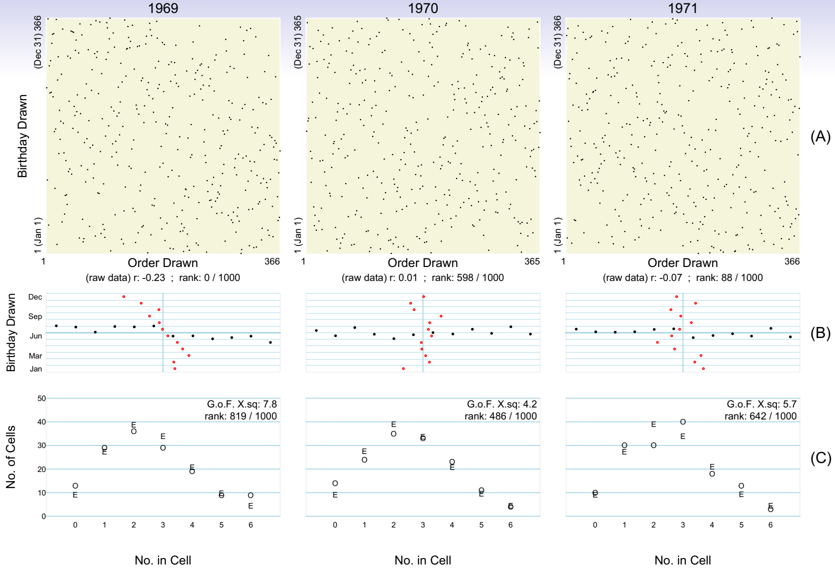


Figure 4. Scatterplots (A), conditional means (B), and observed and expected cell frequencies (C) based on raw data from the first three Vietnam War era lotteries. Explanations as in Figure 1. Each black dot in (B) is a mean($y|x$); each red dot is a mean($x|y$), the more commonly used conditioning in previous analyses of these data. (The 12 red means from the 1969 data were plotted as a bar graph in the New York Times.) The $12 \times 12 = 144$ cells in (C) are formed from the 2-way grid using x- and y-intervals of length 31, 28(29), 30, . . . , 31. Expected numbers and ranks are based on 1000 simulated lotteries. The three datasets are provided in the article by Starr (1997), which also provides a valuable set of primary and secondary print sources.

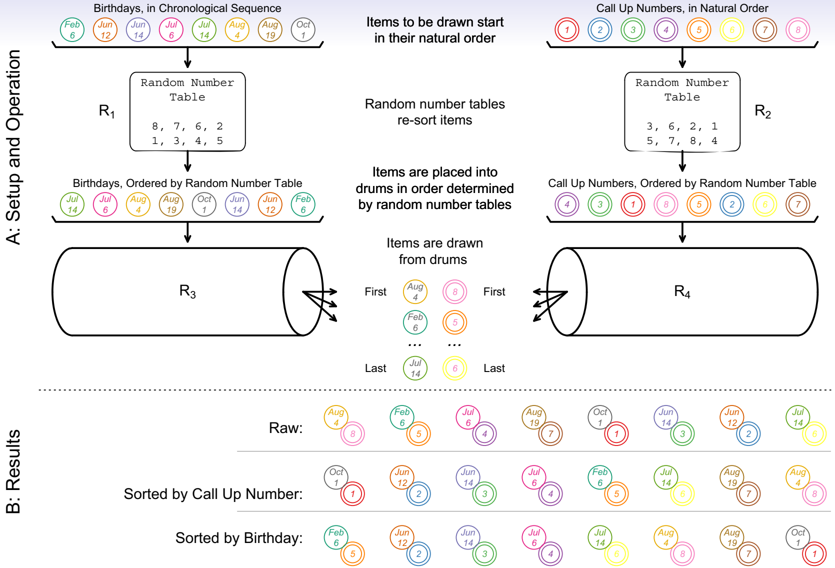


Figure 5. Toy example illustrating the setup of, and a realization from, the "4 randomizations (R_1 to R_4) and 2 drums" procedure used in the 1970–1975 draft lotteries. Randomizations R_1 and R_2 (see text) determined the order in which the birthdays and the call up numbers were placed in the 2 respective drums. Randomizations R_3 and R_4 determined the order in which they were drawn from them. The significance of the 8 birthdays is left for readers to determine.

Figure 6. In readiness for another Selective Service Lottery, *https://www.sss.gov/About/History-And-Records/Selective-Service-Lottery*. The image is reproduced with the permission of the Selective Service System agency.

The ‘Poisson’ Distribution:
History, Re-enactments, Adaptations

- Introduction
- Before 1900
  - deMoivre 1718, and Poisson 1837
  - Newcomb 1860
  - Clausius to Bortkewitsch, 1858-1898
- Early 1900s
  - The Distribution of Objects in Space/Volumes, Gosset 1907
  - Counting Events in Time, Rutherford-Geiger-Bateman 1910
  - Poisson$\leftrightarrow$Exponential Distrn's, Marsden-Barratt 1910-1911
  - The Minimalist Derivation by Danish Engineer Erlang 1909
- Formal Entry | Reception | Warnings
  - Soper 1914
  - Objections: Whitaker, Student 1919, Keynes 1921
  - "Extra-Poisson" Variation in Bacteriology: Fisher 1922
  - Extra Variation in Human Counts: Erlang'09; Student'19
  - An Early Extra-Poisson Model: Greenwood and Yule 1920
- A Broader View: Applications Involving Human Activities/Behavior
  - Feller's '50 Story, Revised by More Recent "Bigger" Picture
  - Avoiding "Fake Standard Errors"
- In Conclusion