

# Od Galilea do Gosseta — prvih 275 gpdina statistike

Franka Miriam Brückler

Statistički seminar, Odjel za matematiku Sveučilišta u Osijeku

19. siječnja 2023.

## Prije Graunta ...

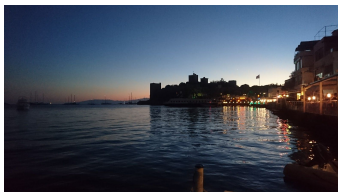
Statistički podaci bili su skupljani još od prvih **civilizacija**: popisi imovine, stanovništva, ... u svrhe oporezivanja, raspodjele imovine, vojske, ...



Nakon propasti Rimskog carstva, *redovni* popisi stanovništva nisu provedeni sve do 18. stoljeća.

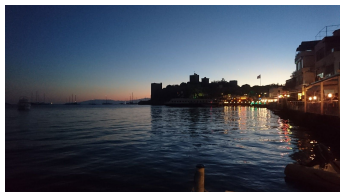
## Herodot (5. st. pr. Kr.)

„Oni navode da tri stotine četrdeset i jedna generacija razdvaja prvog egipatskog kralja od zadnje spomenutog (Hefesta) i da je za svaku generaciju postojao pripadni kralj i visoki svećenik. Sad uzmimo tri generacije kao stotinu godina, tri stotine generacija čine deset tisuća godina, a preostalih četrdeset i jednu generaciju čine još tisuću tristo i četrdeset godina; tako dobijemo ukupno jedanaest tisuća tristo i četrdeset godina . . . “



## Herodot (5. st. pr. Kr.)

„Oni navode da tri stotine četrdeset i jedna generacija razdvaja prvog egipatskog kralja od zadnje spomenutog (Hefesta) i da je za svaku generaciju postojao pripadni kralj i visoki svećenik. Sad uzmimo tri generacije kao stotinu godina, tri stotine generacija čine deset tisuća godina, a preostalih četrdeset i jednu generaciju čine još tisuću tristo i četrdeset godina; tako dobijemo ukupno jedanaest tisuća tristo i četrdeset godina . . . “



16. st.: Jakob Köbel — 16 osoba da definira duljinu *Rute* (16 *Fuß*)



Od 16. st. počelo se sve češće skupljati razne podatke o stanovništvu, najčešće u porezne svrhe. Do 17. stoljeća značajno se povećalo i skupljanje znanstvenih (astronomskih) podataka.

## Galileo Galilei (1564.–1642.)

- Samo je jedna točna vrijednost fizikalne veličine.
- Mjerenja nose grešku uslijed nesavršenosti promatrača i mjernog instrumenta.
- Greške se raspoređuju simetrično s obzirom na 0.
- Manje greške su vjerojatnije od velikih.



# Otac deskriptivne statistike

## John Graunt (1620.–1674.)

Engleski trgovac koji je prvi u povijesti iz statističkih (demografskih) podataka pokušao donijeti zaključke.

*Natural and Political Observations Made Upon the Bills of Mortality* (1662.): analiza izvješća o rođenim i umrlima u Londonu, posebno utjecaj kuge na brojnost i smrtnost stanovništva.

A TABLE of the  
CHRISTENINGS and MORTALITY  
For the Year they and week.

Week	Age of the Child	Sex	Week	Sex of the Child	Sex	Sex
1	1662	145	1	145	145	145
2	1662	145	2	145	145	145
3	1662	145	3	145	145	145
4	1662	145	4	145	145	145
5	1662	145	5	145	145	145
6	1662	145	6	145	145	145
7	1662	145	7	145	145	145
8	1662	145	8	145	145	145
9	1662	145	9	145	145	145
10	1662	145	10	145	145	145
11	1662	145	11	145	145	145
12	1662	145	12	145	145	145
13	1662	145	13	145	145	145
14	1662	145	14	145	145	145
15	1662	145	15	145	145	145
16	1662	145	16	145	145	145
17	1662	145	17	145	145	145
18	1662	145	18	145	145	145
19	1662	145	19	145	145	145
20	1662	145	20	145	145	145
21	1662	145	21	145	145	145
22	1662	145	22	145	145	145
23	1662	145	23	145	145	145
24	1662	145	24	145	145	145
25	1662	145	25	145	145	145
26	1662	145	26	145	145	145
27	1662	145	27	145	145	145
28	1662	145	28	145	145	145
29	1662	145	29	145	145	145
30	1662	145	30	145	145	145
31	1662	145	31	145	145	145
32	1662	145	32	145	145	145
33	1662	145	33	145	145	145
34	1662	145	34	145	145	145
35	1662	145	35	145	145	145
36	1662	145	36	145	145	145
37	1662	145	37	145	145	145
38	1662	145	38	145	145	145
39	1662	145	39	145	145	145
40	1662	145	40	145	145	145
41	1662	145	41	145	145	145
42	1662	145	42	145	145	145
43	1662	145	43	145	145	145
44	1662	145	44	145	145	145
45	1662	145	45	145	145	145
46	1662	145	46	145	145	145
47	1662	145	47	145	145	145
48	1662	145	48	145	145	145
49	1662	145	49	145	145	145
50	1662	145	50	145	145	145
51	1662	145	51	145	145	145
52	1662	145	52	145	145	145
53	1662	145	53	145	145	145
54	1662	145	54	145	145	145
55	1662	145	55	145	145	145
56	1662	145	56	145	145	145
57	1662	145	57	145	145	145
58	1662	145	58	145	145	145
59	1662	145	59	145	145	145
60	1662	145	60	145	145	145
61	1662	145	61	145	145	145
62	1662	145	62	145	145	145
63	1662	145	63	145	145	145
64	1662	145	64	145	145	145
65	1662	145	65	145	145	145
66	1662	145	66	145	145	145
67	1662	145	67	145	145	145
68	1662	145	68	145	145	145
69	1662	145	69	145	145	145
70	1662	145	70	145	145	145
71	1662	145	71	145	145	145
72	1662	145	72	145	145	145
73	1662	145	73	145	145	145
74	1662	145	74	145	145	145
75	1662	145	75	145	145	145
76	1662	145	76	145	145	145
77	1662	145	77	145	145	145
78	1662	145	78	145	145	145
79	1662	145	79	145	145	145
80	1662	145	80	145	145	145
81	1662	145	81	145	145	145
82	1662	145	82	145	145	145
83	1662	145	83	145	145	145
84	1662	145	84	145	145	145
85	1662	145	85	145	145	145
86	1662	145	86	145	145	145
87	1662	145	87	145	145	145
88	1662	145	88	145	145	145
89	1662	145	89	145	145	145
90	1662	145	90	145	145	145
91	1662	145	91	145	145	145
92	1662	145	92	145	145	145
93	1662	145	93	145	145	145
94	1662	145	94	145	145	145
95	1662	145	95	145	145	145
96	1662	145	96	145	145	145
97	1662	145	97	145	145	145
98	1662	145	98	145	145	145
99	1662	145	99	145	145	145
100	1662	145	100	145	145	145

Printed and Sold by J. Streater, at the Sign of the Gun, in St. Dunstons Church-yard, in London.

1662.

# Otac deskriptivne statistike

## John Graunt (1620.–1674.)

Engleski trgovac koji je prvi u povijesti iz statističkih (demografskih) podataka pokušao donijeti zaključke.

*Natural and Political Observations Made Upon the Bills of Mortality* (1662.): analiza izvješća o rođenim i umrlima u Londonu, posebno utjecaj kuge na brojnost i smrtnost stanovništva.

A TABLE of the  
CHRISTENINGS and MORTALITY  
For the Year 1662 and 1663.

Weeks before the Death	No. of Children	No. of Deaths	No. of Burials
1	1000	1000	1000
2	1000	1000	1000
3	1000	1000	1000
4	1000	1000	1000
5	1000	1000	1000
6	1000	1000	1000
7	1000	1000	1000
8	1000	1000	1000
9	1000	1000	1000
10	1000	1000	1000
11	1000	1000	1000
12	1000	1000	1000
13	1000	1000	1000
14	1000	1000	1000
15	1000	1000	1000
16	1000	1000	1000
17	1000	1000	1000
18	1000	1000	1000
19	1000	1000	1000
20	1000	1000	1000
21	1000	1000	1000
22	1000	1000	1000
23	1000	1000	1000
24	1000	1000	1000
25	1000	1000	1000
26	1000	1000	1000
27	1000	1000	1000
28	1000	1000	1000
29	1000	1000	1000
30	1000	1000	1000
31	1000	1000	1000
32	1000	1000	1000
33	1000	1000	1000
34	1000	1000	1000
35	1000	1000	1000
36	1000	1000	1000
37	1000	1000	1000
38	1000	1000	1000
39	1000	1000	1000
40	1000	1000	1000
41	1000	1000	1000
42	1000	1000	1000
43	1000	1000	1000
44	1000	1000	1000
45	1000	1000	1000
46	1000	1000	1000
47	1000	1000	1000
48	1000	1000	1000
49	1000	1000	1000
50	1000	1000	1000
51	1000	1000	1000
52	1000	1000	1000

Primjer Grauntovog zaključka: Žene imaju tendenciju duljeg života od muškaraca, a liječnici tvrde da imaju podjednako ženskih i muških pacijenata, dakle ili liječnici bolje liječe ženske bolesti ili više muških umire od svojih poroka jer ne traže liječničku pomoć 😊

## Edmond Halley (1656.–1742.)

Analizirao zapise o rođenjima i smrtima i dobima smrti u Wroclawu u razdoblju 1687.–1691. ⇒ *An estimate of the degrees of mortality of mankind, drawn from the curious tables of births and funerals in the City of Breslaw; with an attempt to ascertain the price of annuities upon lives.* Phil. Trans., 17 (1693), 596—610  
7 zaključaka: 9/34 stanovništva „može nositi oružje“, vjerojatnosti preživljavanja i medijan preostalih godina života za svaki interval, o cijeni životnog osiguranja, izračuni životnih renti (običnih i zajedničkih za 2 ili 3 osobe)



## Braća Huygens

1669.: Korespondencija Christiaana (1629.–1695.) i Lodewijka (1631.–1699.) o izračunavanju očekivanog i medijanskog životnog vijeka temeljem Grauntovih tablica — usporedba s lutrijom

### *De ratiociniis in ludo aleae* (1656.)

Prva knjiga (zapravo, knjižica) o vjerojatnosti, ujedno i prvi spis u kom se razmatra ideja **očekivanja**.

*Imamo dakle dva različita koncepta: očekivanje ili vrijednost buduće starosti neke osobe i starost u kojoj osoba ima jednaku šansu preživjeti ili ne. Prvo je za izračunavanje životnih renti, a drugo za klađenje.*

(C. Huygens, 1669.)

# Prvi testovi značajnosti 1712.

## John Arbuthnott (1667.–1735.)

Škotski matematičar i liječnik te satiričar, prijatelj Jonathana Swifta i Alexandera Popea.

*An Argument for Divine Providence, taken from the constant Regularity observed in the Births of both Sexes: tablica krštene muške i ženske djece u Londonu 1629.–1710. Uočio je približno konstantan omjer  $M : F > \frac{1}{2}$ : „... we must observe that the external Accidents to which Males are subject (who must seek their Food with danger) do make a great havock of them, and that this loss exceeds far that of the other Sex, occasioned by Diseases incident to it, as Experience convinces us. To repair that Loss, provident Nature, by the Disposal of its wise Creator, brings forth more Males than Females, and that in almost a constant proportion.“*

To pokušava potkrijepiti pomoću binomne razdiobe.

## Willem Jacob 'sGravesande (1688.–1742.)

Nizozemski matematičar i filozof prirode iste godine popravlja Arbuthnotov test i testira hipotezu da je jednako vjerojatno rođenje muškog i ženskog djeteta:

- Od  $\frac{7765}{15448} \approx 50,27\%$  do  $\frac{4748}{8855} \approx 53,62\%$  godišnje rođenih M.
- U te 82 godine prosječni natalitet 11429  $\Rightarrow$  'nominalni' raspon 5745–6128 rođenja muške djece godišnje.
- Uzima da se radi o binomnoj distribuciji s  $n = 11429$  i razmatra  $5745 \leq k \leq 6128$ .
- Ako je  $p = \frac{1}{2}$ , vjerojatnost da je uočeni višak muške djece plod slučaja jednak je 82. potenciji od

$$\sum_{n=5745}^{6128} \binom{11429}{n} \cdot \left(\frac{1}{2}\right)^{11429} \approx 0,292.$$

## Bernoulliji, 2 od 7

### Nicolaus I. Bernoulli (1687.–1759.)

U pismima racionalizira Arbuthnottove i 'sGravesandove ideje i svodi ih na zadatke da se temeljem opažanja procijeni vjerojatnost rođenja muškog djeteta, da se opažena distribucija usporedi s binomnom kako bi se vidjelo može li taj model objasniti varijacije u podacima i da se to matematički olakša dobrom aproksimacijom binomnih koeficijenata za veliki  $n$ .

## Bernoulliji, 2 od 7

### Nicolaus I. Bernoulli (1687.–1759.)

U pismima racionalizira Arbuthnottove i 'sGravesandove ideje i svodi ih na zadatke da se temeljem opažanja procijeni vjerojatnost rođenja muškog djeteta, da se opažena distribucija usporedi s binomnom kako bi se vidjelo može li taj model objasniti varijacije u podacima i da se to matematički olakša dobrom aproksimacijom binomnih koeficijenata za veliki  $n$ .

### Jacob Bernoulli (1655.–1705.)

*Arc conjectandi* (1713.): zakon velikih brojeva kojim je po prvi put povezana vjerojatnost *a priori* i *a posteriori*.

## Primjer

*Stvarni sastav urne je 3000 bijelih i 2000 crnih kuglica, ali to nije poznato osobi koja će iz urne izvlačiti kuglice jednu po jednu, s vraćanjem. Pitanje je, može li ta osoba to učiniti neki broj puta tako da 10, 100, 1000, ... puta bude vjerojatnije (i naposljetku 'moralno sigurno') da omjer frekvencija izvučenih bijelih i crnih kuglica bude jednak stvarnom (3 : 2), nego da taj omjer bude neki drugi?*

*Koristeći „zlatni teorem“ zaključuje da ako želimo 1000 puta veću vjerojatnost da omjer apsolutnih frekvencija bijelih i crnih kuglica bude 3 : 2 nego neki drugi, minimalni broj izvlačenja je 25500.*

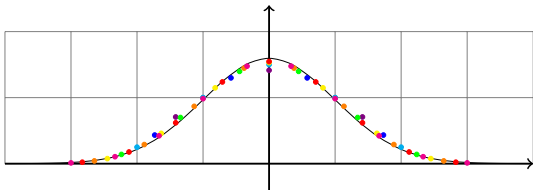
## Abraham de Moivre (1667.–1754.)

*Annuities upon Lives* (1725.), koristeći Halleyeve podatke razvio formule za računanje s rentama (među ostalim i za njihovu podjelu i nasljeđivanje): Prva knjiga o matematici osiguranja.



## Abraham de Moivre (1667.–1754.)

*Annuities upon Lives* (1725.), koristeći Halleyeve podatke razvio formule za računanje s rentama (među ostalim i za njihovu podjelu i nasljeđivanje): Prva knjiga o matematici osiguranja.



1733. otkrio normalnu raspodjelu kao aproksimaciju binomne (centralni granični teorem za  $p = \frac{1}{2}$ )



## Izraz statistika

- Italija, 16. st.: *stato* — država, *statista* — osoba koja se bavi državnim poslovima ⇒ statistika kao skupljanje podataka od interesa državicima
- Njemački povjesničar i pravnik **Gottfried Achenwall** (1719.–1772.): Prvo(?) korištenje izraza statistika (njemački: *Statistik*) u *Vorbereitung zur Staatswissenschaft der europäischen Reiche* (1748.) u smislu popisa sveukupnih karakteristika države vezanih za njen prosperitet.

## Prosjeak kao procjena

Prvo korištenje izraza **aritmetička sredina**: H. Gellibrand 1635. za prosjeak minimalne i maksimalne vrijednosti.

## Prosjeak kao procjena

Prvo korištenje izraza **aritmetička sredina**: H. Gellibrand 1635. za prosjeak minimalne i maksimalne vrijednosti.

Roger Cotes (1682.–1716.)

*„Let  $p$  be the place of some object defined by observation,  $q, r, s$ , the places of the same object from subsequent observations. Let there also be weights  $P, Q, R, S$  reciprocally proportional to the displacements which may arise from the errors in the single observations, and which are given from the given limits of error; and the weights  $P, Q, R, S$  are conceived as being placed at  $p, q, r, s$ , and their center of gravity  $Z$  is found: I say the point  $Z$  is the most probable place of the object, and may be safely had for its true place“ (1722.):*

$$\bar{x} = \frac{\sum_{i=1}^n m_i x_i}{\sum_{i=1}^n m_i} \left( = \min \sum_i m_i (x - x_i)^2 \right).$$



**Thomas Simpson** (1710.–1761.): „*It is well known to your Lordship, that the method practiced by astronomers, in order to diminish the errors arising from the imperfections of instruments, and of the organs of sense, by taking the Mean of several observations, has not been generally received, but that some persons, of considerable note, have been of opinion, and even publickly maintained, that one single observation, taken with due care, was as much to be relied on as the Mean of a great number.*“

Simspon je eksperimentirao i s **krivuljom pogrešaka**, te je predložio ovisnosti tipa  $(\pm e, r^e)$  i  $(\pm e, kr^{e-k+1})$  za iznose grešaka  $e \in N$  (i  $k = 1, \dots, e$ ).

**Thomas Simpson** (1710.–1761.): „*It is well known to your Lordship, that the method practiced by astronomers, in order to diminish the errors arising from the imperfections of instruments, and of the organs of sense, by taking the Mean of several observations, has not been generally received, but that some persons, of considerable note, have been of opinion, and even publickly maintained, that one single observation, taken with due care, was as much to be relied on as the Mean of a great number.*“  
Simpson je eksperimentirao i s **krivuljom pogrešaka**, te je predložio ovisnosti tipa  $(\pm e, r^e)$  i  $(\pm e, kr^{e-k+1})$  za iznose grešaka  $e \in N$  ( $k = 1, \dots, e$ ).

**Pierre-Simon Laplace** 1774.:  $\phi(x) = \frac{m}{2} e^{-m|x|}$  (kasnije:

$$\tilde{\phi}(x) = \frac{1}{2a} \ln \frac{a}{|x|}$$

**Daniel Bernoulli 1777.:** „Astronomers as a class are men of the most scrupulous sagacity; it is to them therefore that I choose to propound these doubts that I have sometimes entertained about the universally accepted rule for handling several slightly discrepant observations of the same event. By these rules the observations are added together and the sum divided by the number of observations; the quotient is then accepted as the true value of the required quantity, until better and more certain information is obtained. In this way, if the several observations can be considered as having, as it were, the same weight, the center of gravity is accepted as the true position of the object under investigation. This rule agrees with that used in the theory of probability when all errors of observation are considered equally likely. But is it right to hold that the several observations are of the same weight or moment or equally prone to any and every error? Are errors of some degrees as easy to make as others of as many minutes? Is there everywhere the same probability? Such an assertion would be quite absurd, which is undoubtedly the reason why astronomers prefer to reject completely observations which they judge to be too wide of the truth, while retaining the rest and, indeed, assigning to them the same reliability.“

## Problem regresijske krivulje

Ako imamo parove  $(x_i, y_i)$  izmjerenih vrijednosti dviju međuovisnih veličina  $x$  i  $y$ , kako dobiti regresijsku krivulju  $y = f(x)$ ?

**Josip Ruđer Bošković** (1711.–1787.): S Englezom Christopherom Maireom je poduzeo istraživačko putovanje od Rima do Riminija da skupi podatke o dvama meridijanima. Pritom je osmislio metodu izravnjanja grešaka kao minimizaciju apsolutnog odstupanja:

$$e(x) = \sum_{i=1}^n |f(x_i) - y_i| \rightarrow \min.$$

Tu je metodu opisao u *De Litteraria expeditione per pontificam ditionem ad dimetiendos duos meridiani gradus a P.P. Maire et Boscovich* (1755.), a istu metodu opisao je i Laplace 1789.

## Metoda najmanjih kvadrata

Ubrzo zatim su tu ideju poboljšali **Adrien-Marie Legendre** (1752.–1833.) i **Carl Friedrich Gauß** (1777.–1855.):

$$E(x) = \sum_{i=1}^n (f(x_i) - y_i)^2 \rightarrow \min$$



## Metoda najmanjih kvadrata

Ubrzo zatim su tu ideju poboljšali **Adrien-Marie Legendre** (1752.–1833.) i **Carl Friedrich Gauß** (1777.–1855.):

$$E(x) = \sum_{i=1}^n (f(x_i) - y_i)^2 \rightarrow \min$$

Gauß je tu metodu razvio 1801. za izračunavanje putanje planetoida Ceresa (objavio: 1809. u *Theoria motus corporum coelestium*), no tvrdio je da ju je koristio od 1795. Legendre je pak, manje precizno, istu metodu objavio 1805., što je dovelo i do rasprave o prvenstvu.

## Greške postaju normalne

Gauß je među ostalim dokazao da je aritmetička sredina točka minimuma za kvadratna odstupanja (tj. da ona minimizira  $\sum(x - x_i)^2$ ).

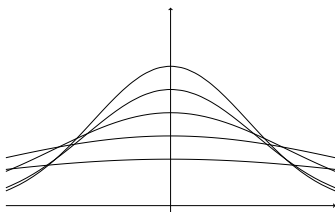
Razloge korištenja kriterija najmanjih kvadrata Gauß argumentira s tri pretpostavke:

- 1 Manje greške su vjerojatnije od većih.
- 2 Za svaki realan broj  $e$  su greške  $e$  i  $-e$  jednako vjerojatne.
- 3 Ako raspolažemo s više mjerenja iste veličine, aritmetička sredina rezultata je najvjerovatniji stvarni iznos.<sup>1</sup>

Iz navedenih pretpostavki Gauß je kao krivulju pogrešaka dobio:

---

<sup>1</sup>Tu je pretpostavku pokušao kasnije eliminirati.



$$\phi(x) = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2}$$

( $h$  je pozitivna konstanta koju je Gauß interpretirao kao preciznost mjerenja).

Detaljnije je metodu najmanjih kvadrata opisao Laplace u *Théorie Analytique des Probabilités* (1812.), kao i razne primjene vjerojatnosti na statistička pitanja. Laplace je 1810. koristeći centralni granični teorem opravdao treću pretpostavku pokazavši da je prosjek uzorka distribuiran normalno oko prosjeka populacije.

Još malo kasnije se S. D. Poisson bavio pitanjima primjene teorije vjerojatnosti u sudstvu (*Recherches sur la probabilité des jugements en matière criminelle et matière civile*, 1837.).

Još malo kasnije se S. D. Poisson bavio pitanjima primjene teorije vjerojatnosti u sudstvu (*Recherches sur la probabilité des jugements en matière criminelle et matière civile*, 1837.).

### Lambert Adolphe Jaques Quetelet (1796.–1874.)

Belgijski matematičar i astronom, utemeljitelj primjena statistike u društvenim znanostima.

Uveo je koncept „prosječnog čovjeka“ (*l'homme moyen*, 1835.), oko čijih se osobina normalno raspodjeljuju osobine svih ljudi.

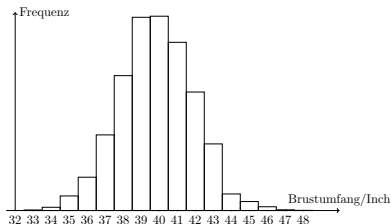
Quetelet je smatrao dužnošću države da skuplja i analizira demografske podatke. Sam je pak analizirao takve podatke i primijetio da se u većim ljudskim populacijama mnoge stvari pojavljuju s velikom pravilnošću i često s normalnom raspodjelom. Quetelet je uočio i da je godišnjih broj ubojstava u Francuskoj gotovo konstantan, te da su čak i omjeri brojeva upotrijebljenih ubojitih sredstava gotovo konstantni – pitanje slobodne volje i društvenog determinizma?!



# Moderna statistika je rođena

## Primjer

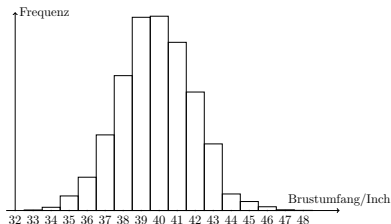
*Quetelet (1846.): Opsezi prsiju 5738 škotskih vojnika.*



# Moderna statistika je rođena

## Primjer

Quetelet (1846.): Opsezi prsiju 5738 škotskih vojnika.



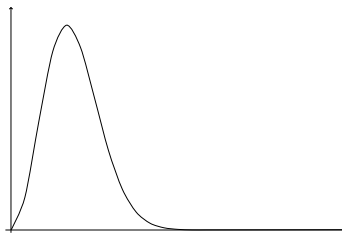
Quetelet je razvio i statističke metode za analize raznih tipova podataka, a 1853. je organizirao prvi međunarodni statistički kongres te se nakon njega izraz **statistika** više ne koristi u smislu „opis općih činjenica o nekoj državi” nego danas uobičajenom.

## Primjena u prirodnim znanostima

Svećenik i prirodoslovac **Gregor Mendel** (1822.–1884.) je pak utemeljio primjenu teorije vjerojatnosti u genetici.

S druge strane, ako se ljudsko društvo može gledati kao cjelina i statistički analizirati, to je još prirodnije za molekule u plinovima.

**James Clerk Maxwell** (1831.–1879.) i **Ludwig Boltzmann** (1844.–1906.) su tako oko 1860. utemeljili statističku termodinamiku.

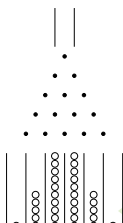
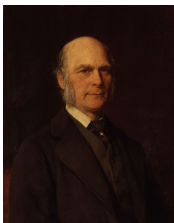




## Sir Francis Galton (1822.–1911.)

Darwinov rođak, utemeljitelj regresijske analize. Bio je poznat po sklonosti skupljanju svakakvih podataka, od temperature čaja do atraktivnosti žena. Utemeljio je *Anthropometric Laboratory*, u kojeg se svatko mogao doći mjeriti. Rezultati tih mjerenja su uvelike potvrdili Queteletove pretpostavke.

Bio je uvjeren u sveprisutnost normalne razdiobe te je skupa s Amerikancem C. S. Peircem i Nijemcem W. Lexisom autor tog naziva.

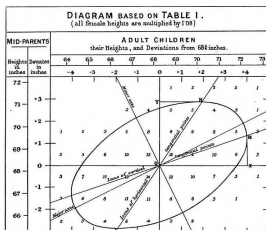


## Regresija prema sredini

Usporedbom promjera sjemenki dviju generacija jedne biljke dobio je korelacijski dijagram i otkrio vezu koju je prvo nazvao reverzijom, a kasnije **regresijom**: Medijan promjera potomaka većeg sjemenja bio je manji nego medijan promjera roditelja i obrnuto.

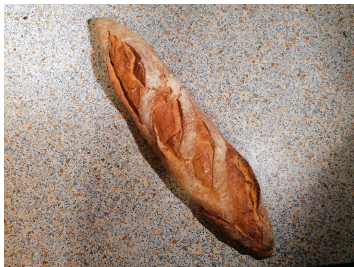
## Regresija prema sredini

Usporedbom promjera sjemenki dviju generacija jedne biljke dobio je korelacijski dijagram i otkrio vezu koju je prvo nazvao reverzijom, a kasnije **regresijom**: Medijan promjera potomaka većeg sjemenja bio je manji nego medijan promjera roditelja i obrnuto. U takve korelacijske dijagrame ucrtavao je pravce koji su najbolje opisivali trendove, a njihove koeficijente smjera je uzimao kao indekse korelacije među dvjema varijablama.



## Jules Henri Poincaré (1854.–1912.)

*Svatko je u nju toliko uvjeren, rekao mi je jednog dana gospodin Lippmann, jer eksperimentalci smatraju da se radi o matematičkom teoremu, a matematičari, da je eksperimentalno potvrđena činjenica.*



## Karl Pearson (1857.–1936.)

Engleski matematičar, otac modernog statističkog testiranja, najzaslužniji za popularizaciju normalne razdiobe. U povijesti statistike ostao je zabilježen kao otac  $\chi^2$ -testa te matematičkog utemeljenja koeficijenta korelacije.

*Prije više godina nazvao sam Laplace-Gaußovu krivulju normalnom, no dok tim nazivom izbjegavamo međunarodno pitanje prvenstva, on ima manu, da navodi ljude na mišljenje da su drugačije razdiobe frekvencija na ovaj ili onaj način abnormalne. (1923.)*



## William Sealy Gosset (1876.–1937.)

Englez, studirao kemiju i matematiku u Oxfordu, ideje je razmjenjivao s K. Pearsonom i R. Fisherom.

1899. se zaposlio u pivovari *Guinness* u Dublinu. Kako bi iz malih uzoraka mogao dobiti zaključke u svrhu kontrole kvalitete razvio je t-test, odnosno distribuciju koja je danas poznata kao **studentova t-raspodjela**: *The probable error of a mean*, *Biometrika*, 6 (1908) 1—25.

## William Sealy Gosset (1876.–1937.)

Englez, studirao kemiju i matematiku u Oxfordu, ideje je razmjenjivao s K. Pearsonom i R. Fisherom.

1899. se zaposlio u pivovari *Guinness* u Dublinu. Kako bi iz malih uzoraka mogao dobiti zaključke u svrhu kontrole kvalitete razvio je t-test, odnosno distribuciju koja je danas poznata kao **studentova t-raspodjela**: *The probable error of a mean*, *Biometrika*, 6 (1908) 1—25. Razlog neobičnog imena bila je zabrana objavljivanja tvorničkih tajni, te je Gosset svoje rezultate objavljivao kao „student”.



# Literatura

- A. Bakker: *The Early History of Average Values and Implications for Education*. Journal of Statistics Education, 11:1 (2003)
- F. M. Brückler: *Geschichte der Mathematik kompakt - Das Wichtigste aus Analysis, Wahrscheinlichkeitstheorie, angewandter Mathematik, Topologie und Mengenlehre*. Springer Spektrum, Berlin, 2017.
- A. Hald, *A History of Probability and Statistics and their Applications before 1750*, John Wiley & Sons, Hoboken, 2003.
- A. Meitzen & R. P. Falkner: *History, Theory, and Technique of Statistics. Part First: History of Statistics*. The Annals of the American Academy of Political and Social Science 1 (1891) 1—100.
- S. Raper: *The shock of the mean*. Significance 14(6) (2017) 12–17.
- Y. H. Said: *Eras in the History of Statistics and Data Analysis*. Journal of the Washington Academy of Sciences 93(1) (2007) 17–35.
- S. Stahl: *The Evolution of the Normal Distribution*. Mathematics Magazine 79(2) (2006) 96–113.
- J. M. Stanton: *Galton, Pearson, and the Peas: A Brief History of Linear Regression for Statistics Instructors*, Journal of Statistics Education, 9:3 (2001).
- S. Stigler: *The Seven Pillars of Statistical Wisdom*. Harvard Univ. Press, London, 2016.
- W. F. Willcox: *Definitions of Statistics*. Revue de l'Institut International de Statistique 2 (1936) 289–290.