The Five Most Consequential Ideas in the History of Statistics

Stephen M. Stigler
University of Chicago
Luzern, 16 November 2007

The Questions

- How to demonstrate that Statistics is a Science?
- What are the most enduring statistical principles?
- What are the ideas in the history of statistics that have been the most influential, the most consequential?

Idea #1: The Combination of Observations

The mean

- 1635 Henry Gellibrand - used by
- 1722 Roger Cotes - weighted
- 1755 Thomas Simpson - proved better

Linear models - Misc & Least Squares

- 1750 Tobias Mayer
- 1780s Pierre Simon Laplace
- 1805, 1809 Legendre, Gauss (LS)

Idea #2: The Root N Rule

Accuracy = 1/St.Deviation $\propto \sqrt{N}$

Gives a rate for accumulation of information

1713, 1716 Jacob and Nicholas Bernoulli

1730 Abraham De Moivre - binomial

1810, 1812 Laplace - general (CLT)

1879 C. S. Peirce - economy of research

Idea #3: Tests and Likelihood

1248 London Mint, Trial of the Pyx

1710 John Arbuthnot: 1 chance in 2^{82}

1735 Daniel Bernoulli: planetary orbital planes

1922 R. A. Fisher: Likelihood

1933 J. Neyman and E. S. Pearson theory of testing

Idea #4: Statistics by Intercomparison

The internal measurement of variability

1875 Francis Galton - percentiles

1885 Francis Edgeworth - variance components

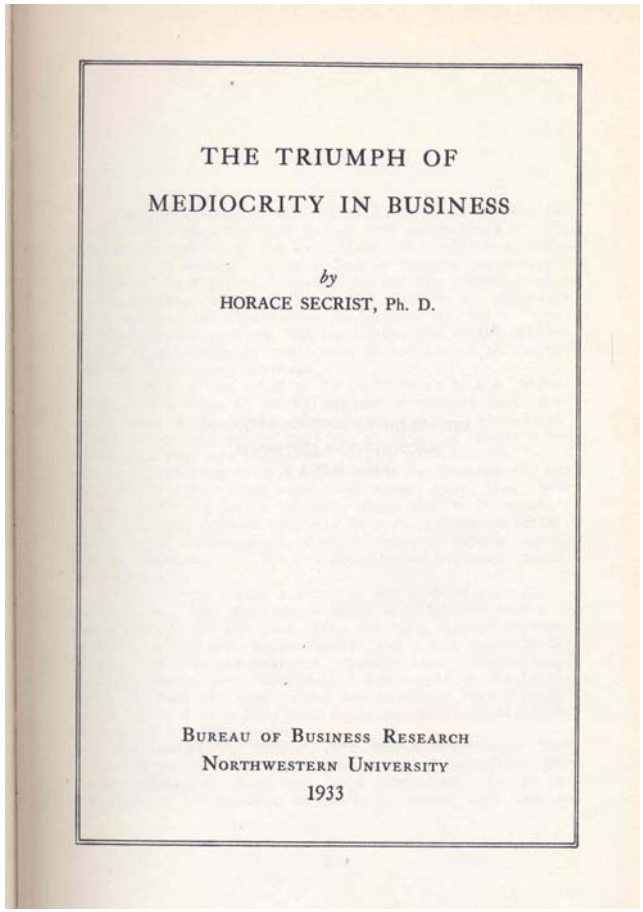
1908 W. S. Gosset - t-test

1918-1925 R. A. Fisher - ANOVA and design

Idea #5: Regression, …

1877-1889 Francis Galton - phenomenon, paradox

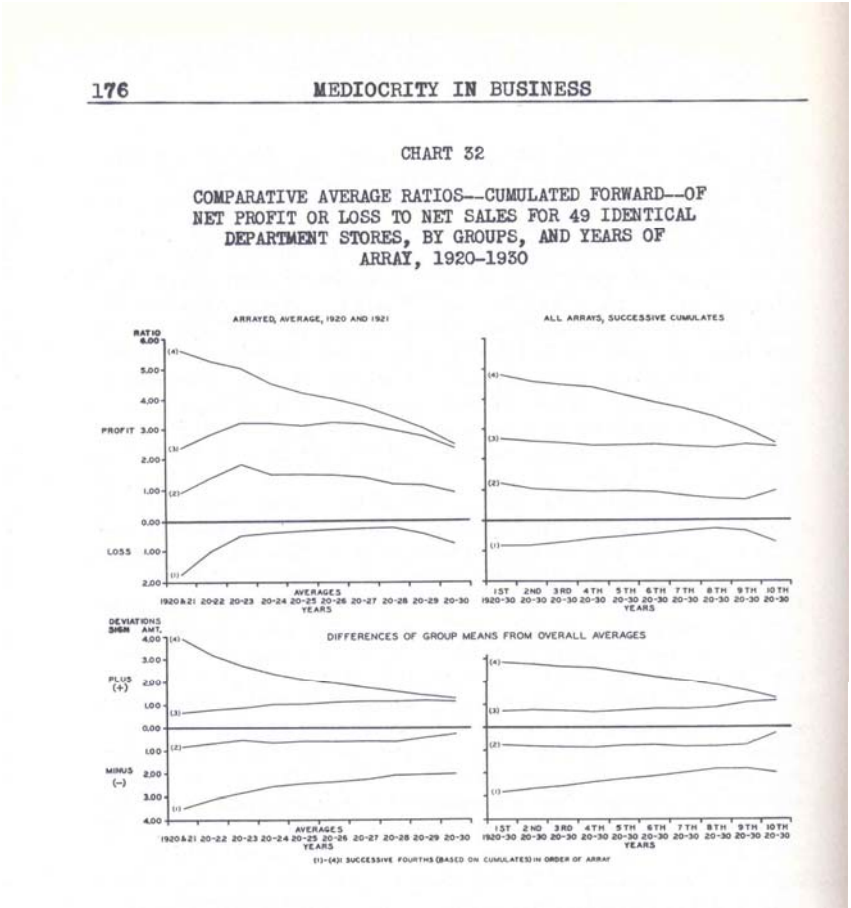
1933 Horace Secrist
(1970 A. O. Hirschman)



THE TRIUMPH OF
MEDIOCRITY IN BUSINESS

by
HORACE SECRIST, Ph. D.

BUREAU OF BUSINESS RESEARCH
NORTHWESTERN UNIVERSITY
1933



176 MEDIOCRITY IN BUSINESS

CHART 32

COMPARATIVE AVERAGE RATIOS--CUMULATED FORWARD--OF NET PROFIT OR LOSS TO NET SALES FOR 49 IDENTICAL DEPARTMENT STORES, BY GROUPS, AND YEARS OF ARRAY, 1920-1930

(1)-(4): SUCCESSIVE FOURTHS (BASED ON CUMULATES) IN ORDER OF ARRAY

Idea #5: Regression, correlation,…

1893 Francis Edgeworth - multinormal

1895 Karl Pearson - math of correlation

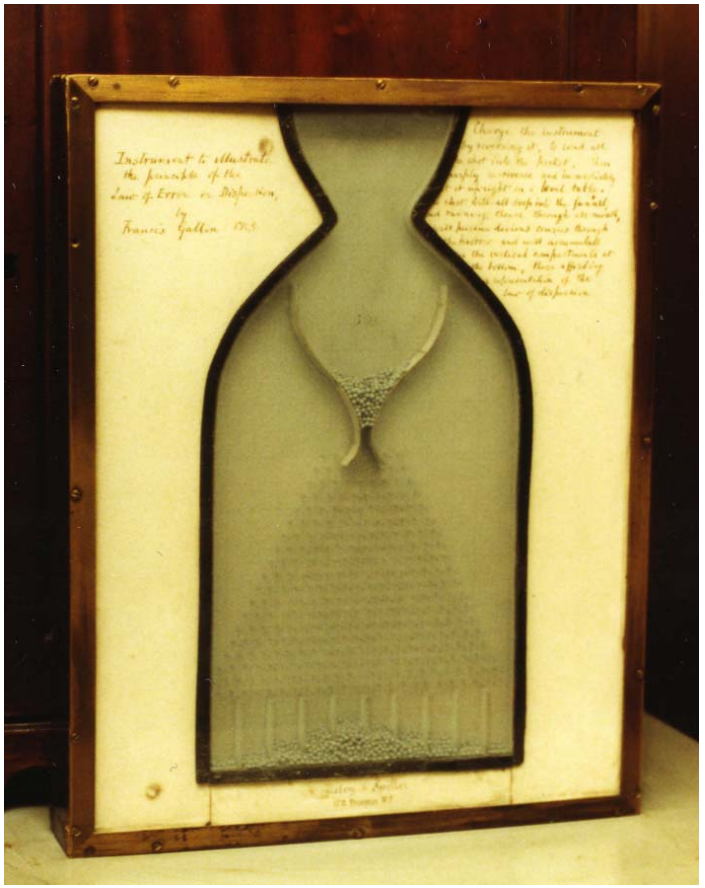
1922-1936 R. A. Fisher - distribution theory etc

Idea #5 (ctd.): … and Bayes

1764 Thomas Bayes

1772-74 Laplace

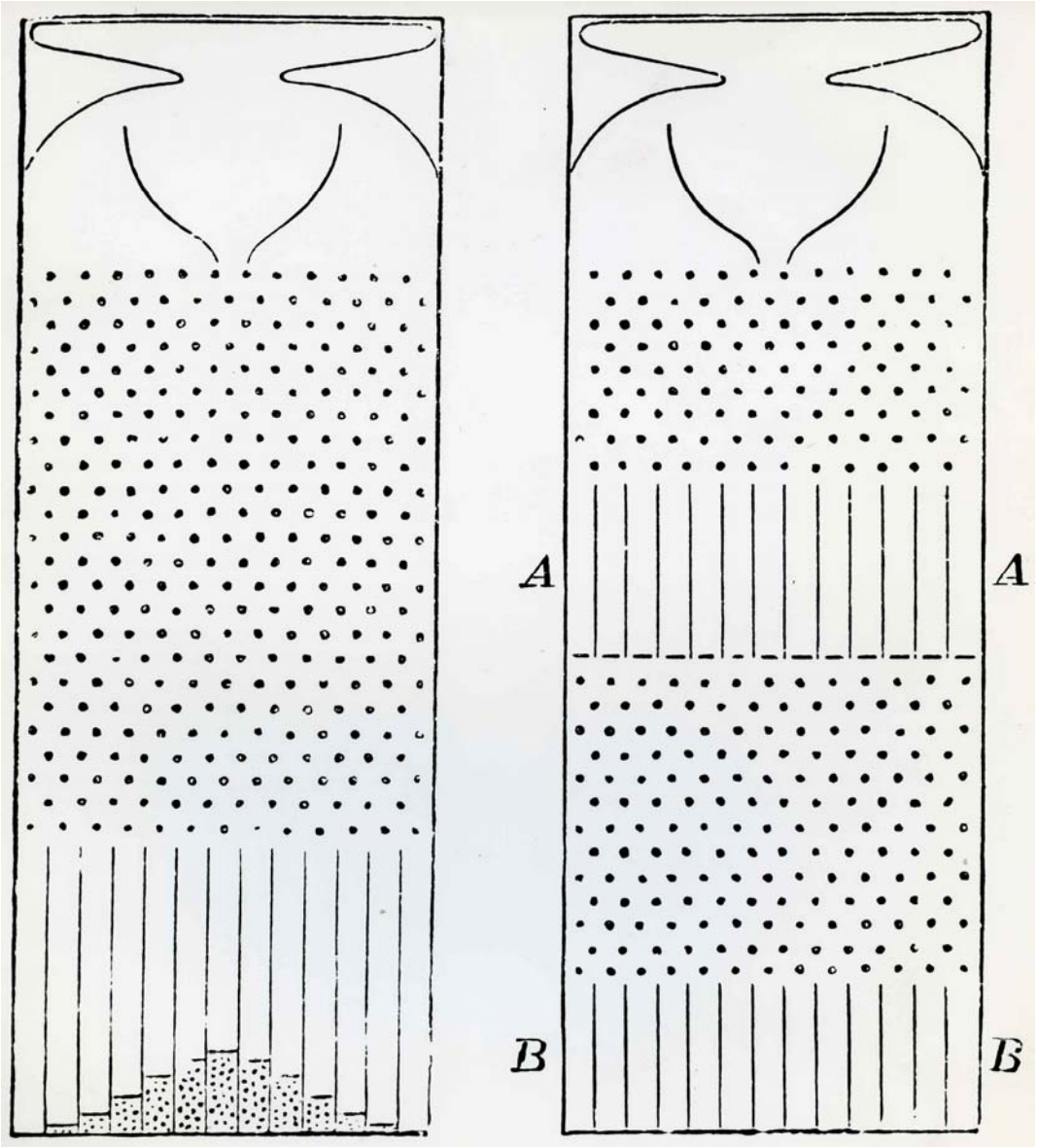
1877-1889 Francis Galton



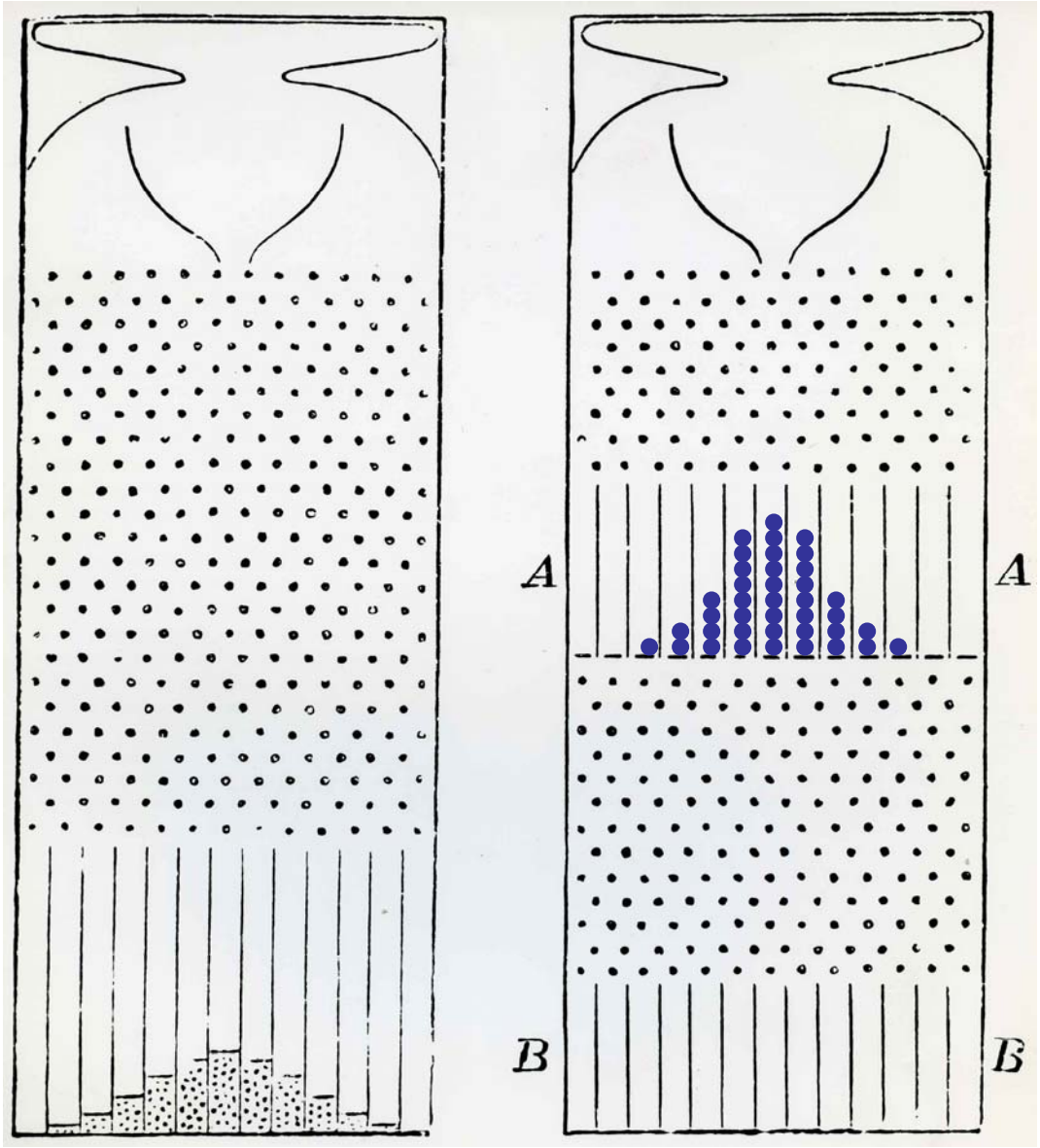
Galton's Quincunx
(1873 version)

Galton
Quincunx
1877-89

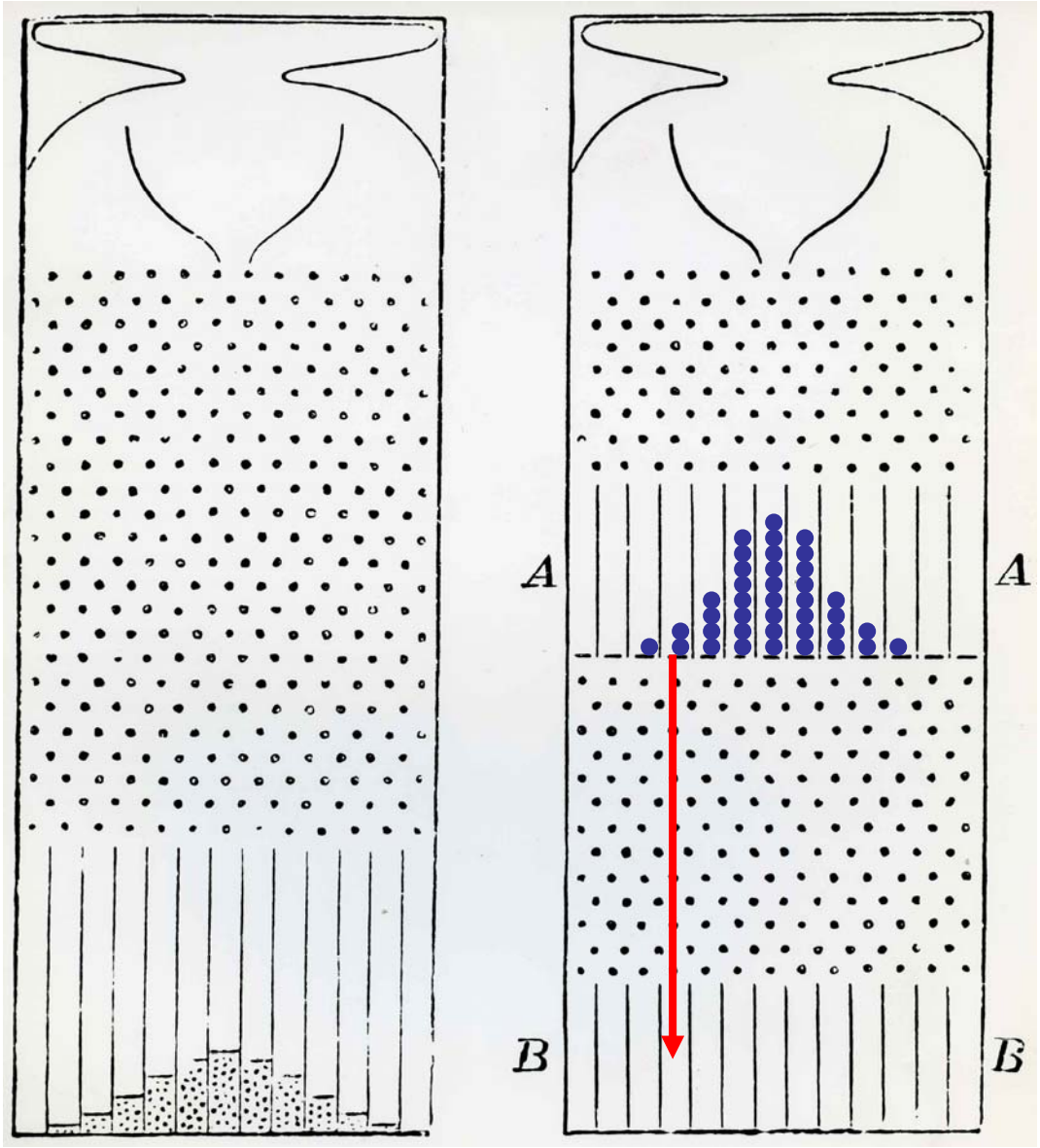
(from Galton's *Natural Inheritance*, 1889)



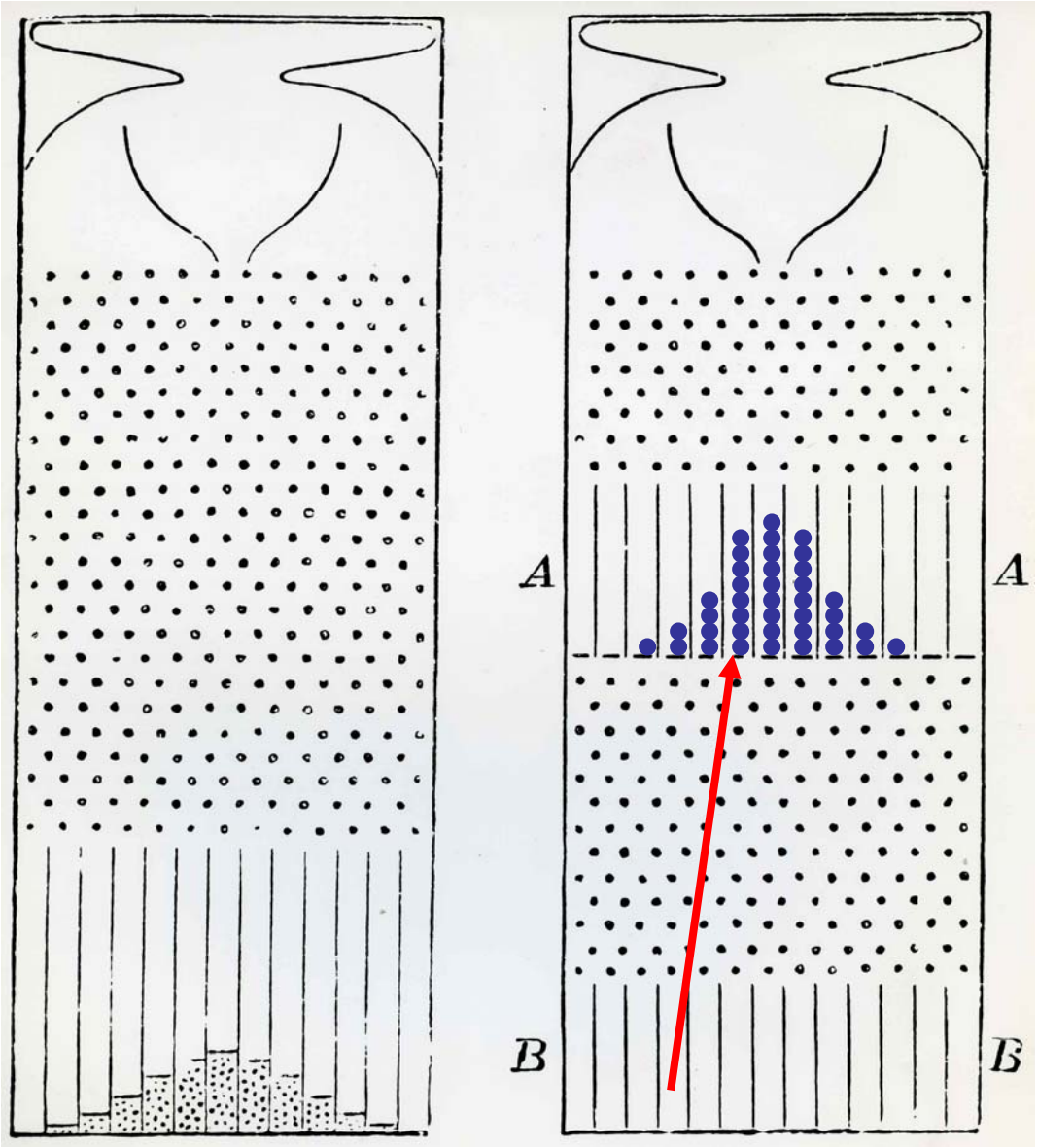
Galton
Quincunx
1877-89



Galton
Quincunx
1877-89



Galton
Quincunx
1877-89



Idea #5 (ctd.): … and Bayes

1880s, 1890s Edgeworth, K. Pearson

1930s Harold Jeffreys

1950s Jimmie Savage, Dennis Lindley,…

Five consequential ideas from history

#1 The combination of observations

#2 The root N rule, rate of increase in accuracy

#3 Statistical tests and likelihood

#4 Statistics by intercomparison; using internal variability; the planning of experimentation

#5 Regression phenomena, correlation, multivariate analysis, and modern Bayesian analysis


JAC. BERNOULLI. MATH. P.