Aggregation, Inference and Rare Events in the Natural and Socio-Economic Sciences

Nick Watkins [1,2,3](nww@bas.ac.uk)

1. NERC British Antarctic Survey, Cambridge, UK

2. Visiting Fellow, CATS, London School of Economics, UK

3. Visitor, CFSA, University of Warwick, UK

British Antarctic Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL

LSE
THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

THE UNIVERSITY OF WARWICK

Welcome from me, from Colm, and from Warwick's Complexity Complex

• Many thanks for coming and hope you will enjoy the next couple of days ! This is just a short intro to meeting's themes, and, I hope, a first chance to focus on why I was so keen to gather you together.

•Consider some of our ancestors in mathematics, statistics, physics and elsewhere and some of what we have inherited:

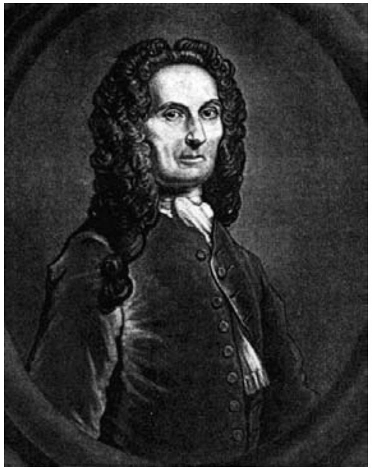
De Moivre, Laplace and Lyapunov
Gauss and Lévy
Fisher and Mandelbrot

Neither rigorous nor necessarily accurate history-just a user's view.

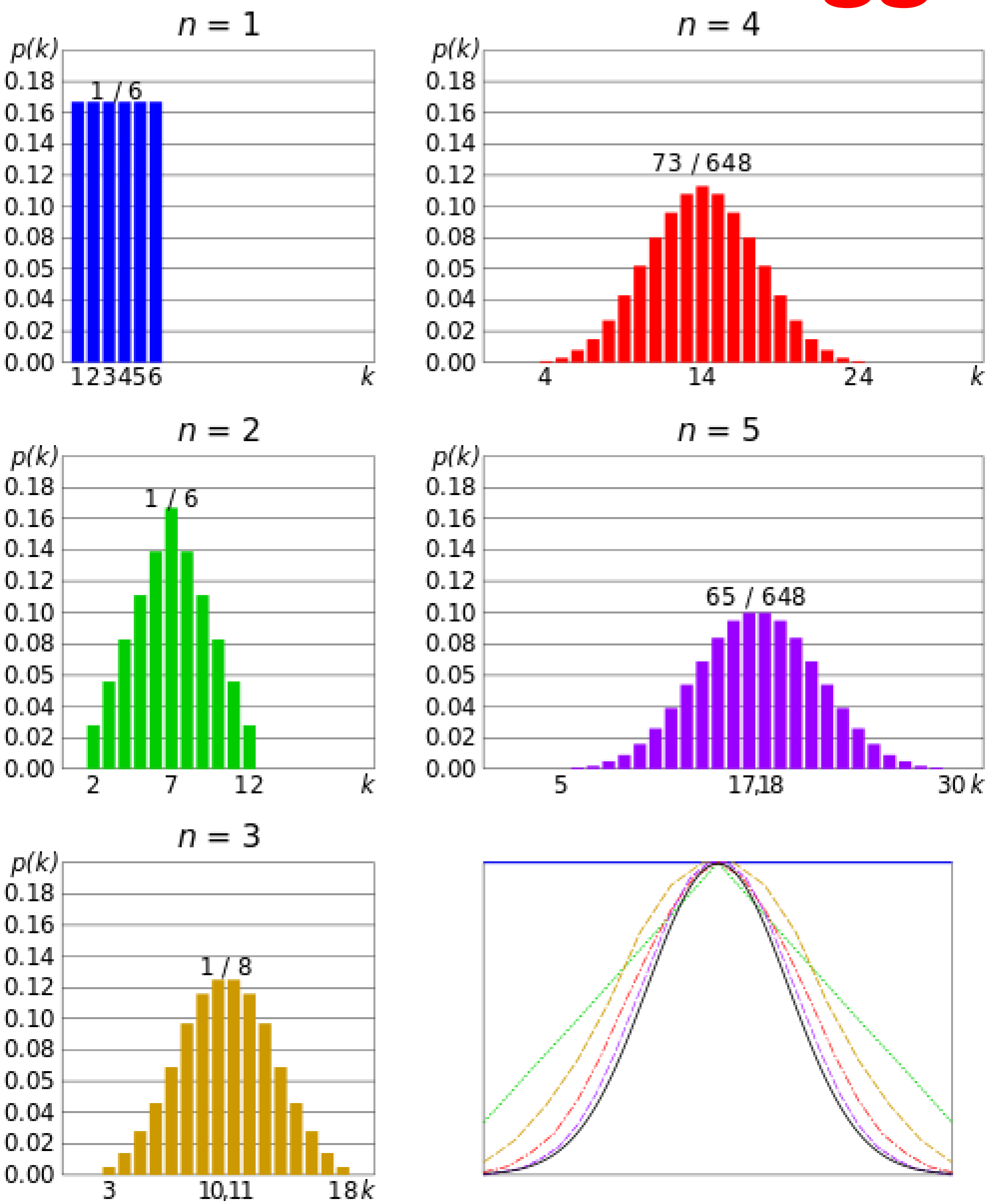
Central Limit Theorem

- **de Moivre (1733)** normal distribution approximates pdf of number of heads resulting from many tosses of a fair coin.
- **Laplace (1812)** *Théorie Analytique des Probabilités*: Expanded on de Moivre by approximating binomial distribution with normal distribution.
- **Lyapunov (1901)** defined CLT in general terms and proved it.

- Wikipedia



CLT and aggregation

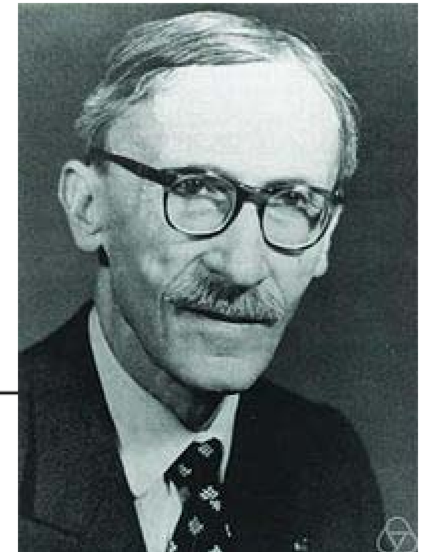


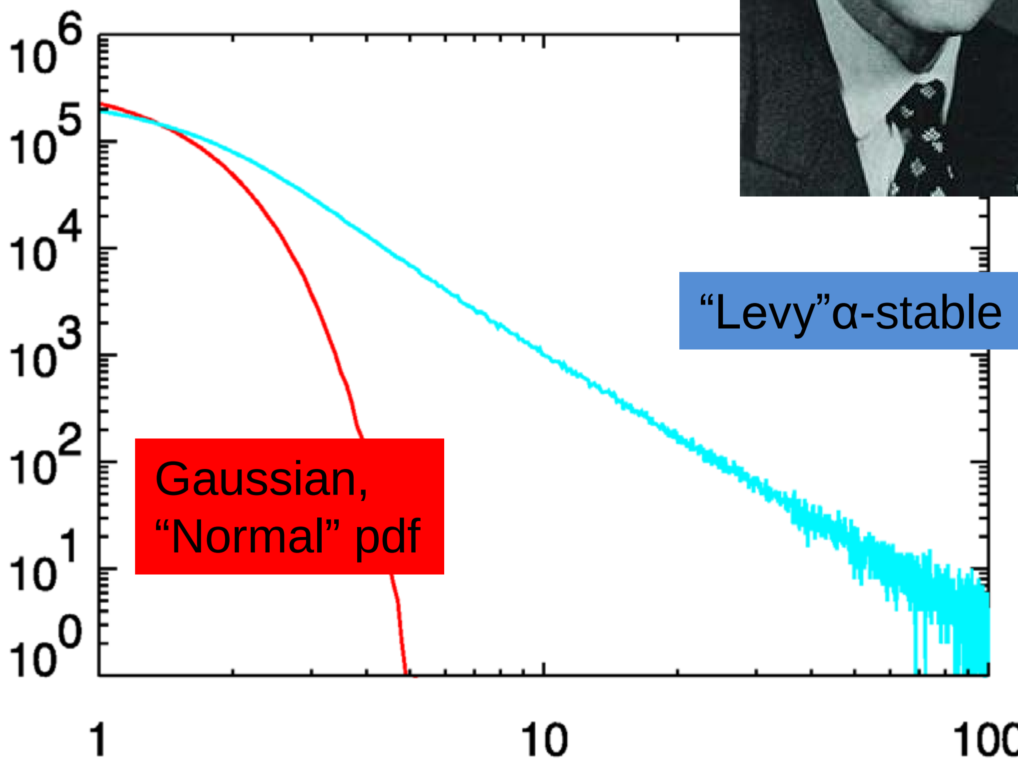
- Comparison of pdf $p(k)$ for sum of n fair 6-sided dice. Sum converges with increasing n to normal distribution, as per CLT. In the bottom-right graph, smoothed profiles of the previous graphs are rescaled, superimposed and compared with a normal distribution, shown in black. - Wikipedia

Extended CLT

- A key extra step came in 20^{th} century when it was shown that more general stable laws exist for aggregation of random numbers when we relax the assumption of finite variance.
- These are known as the α-stable ('Lévy') distributions, where $0 < \alpha \leq 2$. The $\alpha = 2$ stable case is the Gaussian.
- Defined by the Fourier transform of characteristic function.
- Physicist-friendly introductions include books of **Mantegna & Stanley; Bouchaud & Potters; and Sornette**
- As well as recognised references like **Samorodnitsky &Taqqu, and Janicki & Weron.**

Gauss and Lévy





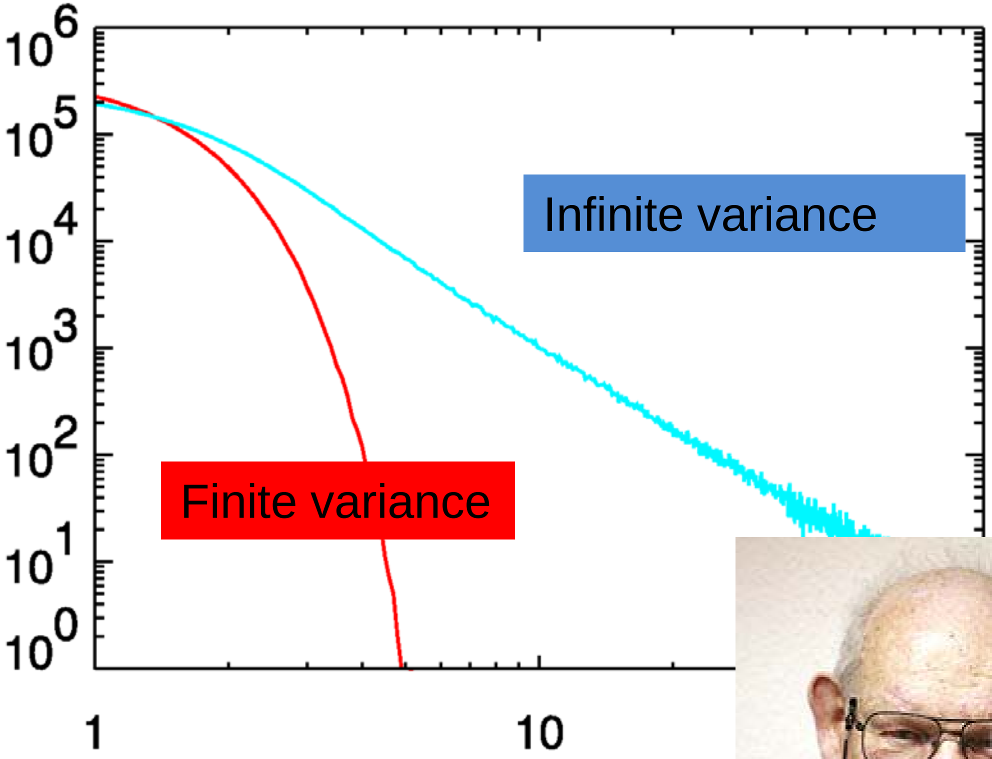
CLT gives Gaussian,
"light-tailed" pdf

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-(x-\mu)^2 / 2\sigma^2\right)$$

ECLT leads to "Levy", α-stable heavy-tailed pdf, decays as if $0 < \alpha < 2$

$$p(x) \sim x^{-(1+\alpha)}$$

Fisher, Mandelbrot & rare events



Fisher (1918) was first use of term “variance”

Mandelbrot (1963) was
a key step in advocating
infinite variance models in finance

Ongoing challenges include:

1. How does aggregation behave when we relax assumptions of iid ? Obviously a huge question, narrow focus a bit to cases inspired by user needs (e.g. Mitas) and physical situations (Bramwell, Eliazar, Magdziarz) .
2. In particular what happens when you have long range dependence in time (Graves, Eliazar) , or multiplicative situations (Løvsletten) ?

Ongoing challenges include:

3. How well can we do inference i.e. distinguish limiting behaviours in real world applications? (Shalizi, Smith, Graves, Wood, Klages)

4. How/when does mathematical universality parallel scientific universality (Eliazar, Clusel, Pruessner, Bramwell, Goncalves ...)

Universality ...

- *I know of scarcely anything so apt to impress the imagination as the wonderful form of cosmic order expressed by the "Law of Frequency of Error". The law would have been personified by the Greeks and deified, if they had known of it. It reigns with serenity and in complete self-effacement, amidst the wildest confusion. The huger the mob, and the greater the apparent anarchy, the more perfect is its sway. It is the supreme law of Unreason. Whenever a large sample of chaotic elements are taken in hand and marshalled in the order of their magnitude, an unsuspected and most beautiful form of regularity proves to have been latent all along.*

- ***Galton, 1889.***