Three centuries of random lines: from Buffon's needle to scale-invariant networks

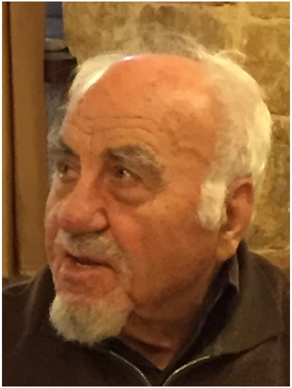
Theo Cacoullos Memorial Lecture 2024

Wilfrid S Kendall

University of Warwick

16 May 2024

Introduction



Theophilos Cacoullos 1932-2020

Cacoullos:
"An excellent teacher who never used notes but did use spontaneous humor"

(IMS obituary November 16, 2021)

I count it a real honour to be invited to present the second Theo Cacoullos Memorial lecture.These remarks on his presentational style make it clear that Theo Cacoullos Memorial lecturers have much to live up to! I will do my very best.

Handout available on the web: either use the `QR-code`

or visit https://wilfridskendall.github.io/talks/Kozani-Cacoullos-2024/

Recent Origins: Particle Tracks in Apatite Crystals

An apatite crystal

- Nearly half a century ago, I assisted in a study (Laslett *et al.*, 1982) of fission track lengths in apatite crystals (tracks from fission of ^{238}U nuclei; crystal is sliced then etched to reveal intercepting tracks);
- *In principle* the length histogram indicates thermal hence geological history; but direct observation is massively biased by
  - (a) length selection bias,
  - (b) truncation (a kind of *Wicksell Problem*),
  - (c) projection effects (if non-horizontal);
- How to model a pattern of segments? (*Poisson segment process*: a kind of *Boolean model*);
- How to get good length histogram? (select *horizontal* enclosed tracks hitting tracks that intersect slice).

Warwick Statistics

Historical Origins: pre-Revolutionary France

Georges-Louis Leclerc,
Comte de Buffon, 1707 – 1788

Buffon rose from undistinguished origins in the France of the *Ancien Regime*, to become a major and celebrated philosopher and scientist. He embodied a remarkable mixture of empiricism and rationalism.

“Writing well consists of thinking, feeling and expressing well, of clarity of mind, soul and taste … The style is the man himself.”

Throwing baguettes over one's shoulder (I)

Baguette as used by the Count.

- Comte de Buffon was very interested in probability (read him in Hey *et al.*, 2010);
- Probability from geometry not gambling;
- There ensued a whole sub-culture of experimental determinations of π (see Lazzarini, 1901; Stigler, 1991),

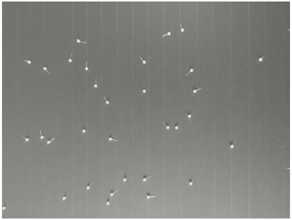


Table used by UCL experiment

- Most recently at the inaugural UCL Bloomsbury Probability Colloquium March 2024;
- Strips at 3cm, pins at 1.5cm, with 24 intersections out of 45 throws.
- This estimates $\pi \approx 45/21 = 2.14\ldots$: the true value of π lies just inside the 95% confidence interval!

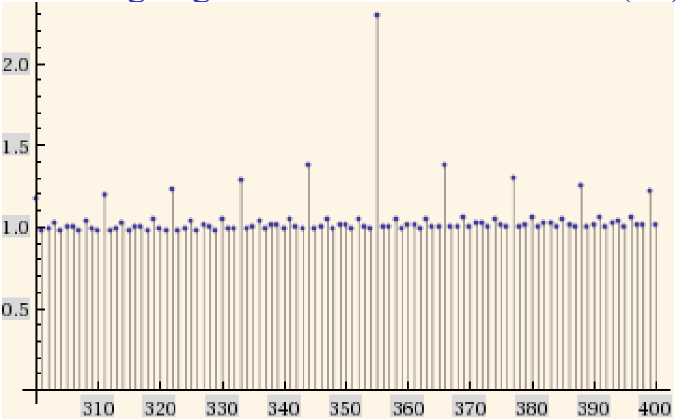
Throwing baguettes over one's shoulder (II)

Suppose you have to design such an experiment.

- Unit-length needle, unit-ruled floor. Choose n the number of trials.
- The number of successes X is Binomial(n, $2/\pi$).
- Measure success by $1/\sqrt{|2/\pi - X/n|}$.
- Use your skill and judgement (and your balance between accuracy and patience!) to decide on total number of trials $n \in [300, 400]$

In the next slide, we graph the mean measure of success against n to see how much the choice of n actually makes any difference!

Throwing baguettes over one's shoulder (III)

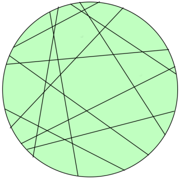


Log of mean of $1/\sqrt{|2/\pi - X/n|}$ \ for X with distribution Binomial$(n, 2/\pi)$.

$$\pi = 3 + \cfrac{1}{7 + \cfrac{1}{5 + \cfrac{1}{1 + \cfrac{1}{292 + \cfrac{1}{1 + \cfrac{1}{\dots}}}}}} \approx 3 + \cfrac{1}{7 + \cfrac{1}{15 + 1/1}} = \frac{355}{113} = 3.14159292035398 25\,.$$

Warwick Statistics

Poisson line processes

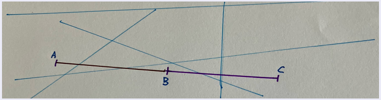


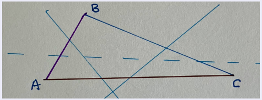
How to build random patterns of lines?

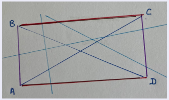
1. There is a theory of Random Closed Sets. More constructively,
2. Boolean model: union of sets located are Poisson points.
3. Random lines ℓ are not localized, so this must be adapted.
4. Simple solution: parametrize by signed perpendicular distance r, angle θ:
   1. represent each line ℓ by a point (r, θ) in representation space;
   2. view Poisson line process as Poisson point process in representation space. Use invariant line measure $\frac{1}{2}\,\mathrm{d}r\,\mathrm{d}\theta$.
5. The representation space is a cylinder with twist.
6. Calculations typically reduce to computation of probabilities that there are no lines of particular forms.

Examples of calculations with lines

The probability of there being no lines in (the representation space subset) E is $\exp\left(-\mu(E)\right)$, where $\mathrm{d}\mu = \frac{1}{2}\,\mathrm{d}\,r\;\mathrm{d}\,\theta$ is the invariant measure. Compute:

- $\mathbb{P}$ [no lines hit a unit segment]:
  - 

  - By (a) translation invariance, (b) choice of normalizing constant $\frac{1}{2}$, the probability can be computed to be $\exp(-1)$.
- $\mathbb{P}$ [all lines hitting a triangle ABC actually hit side AC]:
  - 

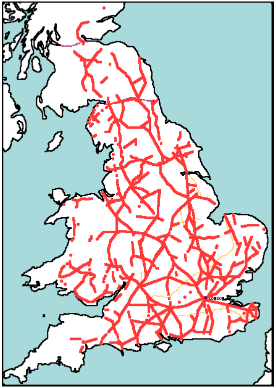
  - By a counting argument(!) this is $\exp\left(-\frac{1}{2}(|AB| + |BC| - |AC|)\right)$.
- $\mathbb{P}$ [no lines pass between AD and BC of rectangle $ABCD$]:
  - 

  - Counting argument again(!): $\exp\left(-\frac{1}{2}(|AC| + |BD| - |AD| - |BC|)\right)$.

Other relevant notions

- An alternate parametrization can be very convenient for calculations: p is distance of intercept along a reference line, while ϕ is angle as before.
  - invariant measure now $\frac{1}{2}\sin\phi\;\mathrm{d}p\,\mathrm{d}\phi$;
- Notion of "conditioning on a line at the origin":
  - Slivnyak's theorem says that the remainder of the line pattern behaves just like the original Poisson line process! This is the theory of *Palm probability*, extremely useful in showing how to reason with general constructions;
- The use of marks for the points / lines:
  - *eg* speed-marked line process.

Application to network efficiency



Stochastic geometry of online maps

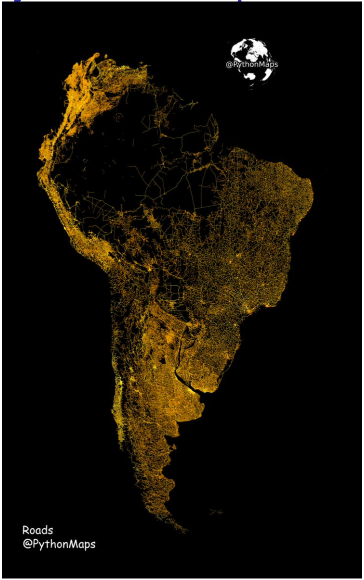
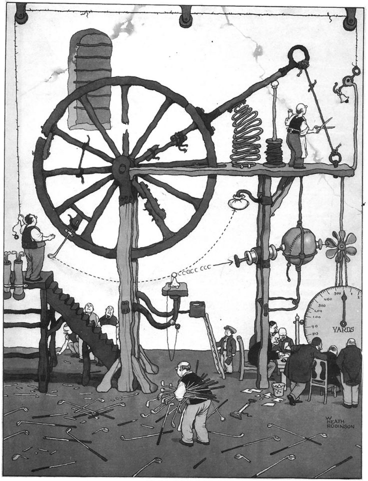


Figure 1: All the roads in South America

- Modelling transportation systems;
- Visual approach to axioms for online maps: - Holland and Northwest Germany;
  - Generating networks by a (random) mechanism;
  - Axioms:
    1. Scale invariance ;
    2. Finite mean-length;
    3. "SIRSN" property (Scale-Invariant Random Spatial Network).
  - Bertrand Russell: "The axiomatic method has many advantages over honest work."
  - Settling the existential question:

    randomized dyadic speed-marked planar lattice (Aldous, 2014);
    speed-marked Poisson line process (Aldous, 2014; Kahn, 2016; WSK, 2017)
    using the line process theory discussed above.
- More realistic models:
  - Sticks instead of lines;
  - Thick stick soup;
  - Thick Pareto stick soup;
  - Thick stiff fibre soup.

Some further remarks



Testing Golf Drivers

34

Figure 2: Much more work is still to be done

- "Thickness" is a simplifying assumption and still has reality problems:
  - using finite-mean-length sticks with infinite second-length-moment, means that *any* selected stick will have infinite-mean-length (length-biasing!);
  - reasonable to suppose that it suffices to have super-critical percolation for the configuration of fibres of greated than unit speed;
  - But stick percolation is *hard*! Still working to try to develop some relevant theory.
- Everything still works for lines in higher-dimensional space:
  - Dimension 3 is currently the only other case of interest (could be of interest when studying the human connective!);
  - Blanc has recently published nice work on an analogue in hyperbolic space.
- A big attraction of the theory in its present state is that one can consider relaxing parts of the scaling symmetry assumption. The "SIRSN effect" (Axiom 3) is probably still going to work if controlled spatial inhomogeneity is allowed. So this approach now provides a flexible framework for statistical and operational research modelling.
- We can introduce traffic, for example by supposing traffic along the connecting route to be uniformly generated between any pair of points. Gameros Leal (2017) worked this out for a simpler model and compared

References I

Aldous, D.J. (2014) Scale-Invariant Random Spatial Networks. *Electronic Journal of Probability*, **19**, 1–41.

Gameros Leal, R.M. (2017) Mean traffic behaviour in Poissonian cities (PhD Thesis).

Hey, J.D., Neugebauer, T.M., & Pasca, C.M. (2010) Georges-Louis Leclerc de Buffon's 'Essays on Moral Arithmetic'. *The selten school of behavioral economics* (Sadrieh, A., Ockenfels, A.(eds). ed). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 245–282.

Kahn, J. (2016) Improper Poisson line process as SIRSN in any dimension. *Annals of Probability*, **44**, 2694–2725.

Laslett, G.M., Kendall, W.S., Gleadow, A.J.W., & Duddy, I.R. (1982) Bias in measurement of fission-track length distributions. *Nuclear Tracks and Radiation Measurements (1982)*, **6**, 79–85.

Lazzarini, M. (1901) Un'applicazione del calcolo della probabilita alla ricerca esperimentale di un valore approssimato di π. *Periodico di Matematica*, **4**, 49–68.

Stigler, S.M. (1991) Stochastic simulation in the nineteenth century. *Statistical Science*, **6**, 89–97.

WSK (2017) From random lines to metric spaces. *Annals of Probability*, **45**, 469–517.

Technical information

Image	Original location	
Theo Cacoullos	IMS Bulletin obituary, November 2021	
Apatite crystal	Apatite_taillée.jpg	By Didier Descouens - Own work, CC BY 3.0
Comte de Buffon	Buffon_1707-1788.jpg	Free of copyright
Baguette	baguette-de-pain-.jpg	Free stock photo - Public Domain
Buffon and Number Theory	WSK calculation	
Buffon Table	By kind permission of Terry Soo (UCL)	
Poisson line process	WSK calculation	
S. American Roads	By kind permission of @PythonMaps	
Heath Robinson	Testing Golf Drivers	

Poisson line process

These notes were produced from `Kozani-Cacoullos-2024.qmd`:

Author:	Wilfrid Kendall W.S.Kendall@warwick.ac.uk
Date:	Sat May 11 10:23:02 2024 +0100
Summary:	Adjusted QR-code generation.