Poisson spacing statistics in number theory and mathematical physics

Pär Kurlberg

“Randomness” of sequences

- Will consider three types of sequences:
  - $x_n \in [0, 1]$.
  - $x_n \to \infty$.
  - Sequence of finite subsets of $[0, 1]$.

- Examples:
  - $\alpha n^2 \mod 1$, α irrational.
  - $\alpha 2^n \mod 1$, α irrational.
  - Primes: $x_n = p_n$, the n-th prime.
  - Sums of squares: order the set

$$X = \{\alpha m^2 + n^2 : m, n \in \mathbf{Z}^+\}.$$

- Sequences of finite sets:

  – Invertible elements mod q

  $$\{\frac{a}{q} : 0 \leq a < q, (a,q) = 1\}$$

  – Primitive roots mod p, p prime.

  $$\{\frac{a}{p} : 0 \leq a < p, \mathrm{ord}(a,p) = p-1\}$$

  – Squares mod q

  $$\{\frac{a}{q} : 0 \leq a < q, a = x^2 \mod q\}$$

  Note: In order to get Poisson spacings, must take let p (or q) tend to ∞ in special way.

Some “randomness” tests

1. Say $\{x_i\}_{i=1}^{\infty}$ is sequence of numbers in the unit interval.

2. Interesting statistics of $\{x_i\}_{i=1}^{\infty}$?

   - Equidistribution
   - Level spacings
   - Pair correlation

3. Compare to uniformly distributed, I.I.D. random variables.

Equidistribution

- Is there measure μ such that $\forall f \in C^\infty[0,1]$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} f(x_i) = \int_0^1 f(x) d\mu(x)?$$

- The sequence is called *equidistributed* if μ is the Lebesgue measure.

- For IID, the sequence is equidistributed with probability one.

- Good property: rule out $\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \dots$.

- "Bad" property: sequence $\{\alpha n \mod 1\}_{n=1}^{\infty}$, for α irrational passes. Fine sequence, but has degenerate spacings - at most three different gaps possible in $\{\alpha n \mod 1\}_{n=1}^{N}$ for any N.

Level spacings

- *Ordering:* For each $N \in \mathbf{Z}^+$, let $x_1^N \leq x_2^N \leq \ldots \leq x_N^N$ be the set $\{x_i\}_{i=1}^N$ arranged in increasing order.

- *Unfolding:* $s_i^N = (x_{i+1}^N - x_i^N) \times N$ is *normalized* to have mean one.

  Unless we normalize, s_i^N will get "squished to 0".

- *Level spacing distribution:* Function P such that $\forall f \in C_c^\infty[0,\infty]$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} f(s_i^N) = \int_0^\infty f(x)P(x)dx.$$

- In IID case get *Poisson statistics*; $P(x) = e^{-x}$.

Poisson statistics for IID case, sketch

- Instead of putting N points in unit interval, put N points on unit circle in $\mathbf{R}^2$.

- Introduces one extra gap, but extra gap has same distribution as the other gaps.

- May pick any one of the N points. By rotation invariance, may assume we picked $(1, 0)$.

- Translating back to unit interval we find that the probability of this particular gap being larger than x/N is $(1 - x/N)^{N-1}$.

- Note: this does not say anything about the joint distribution of the different gaps!

Pair correlation

- Problem with spacings: involves ordering which is "difficult to capture". Instead look at distribution of *unordered* gaps.

- Pair correlation: function $R_2(x)$ such that $\forall f \in C_c^\infty[-\infty, \infty]$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{1\le i\neq j\le N} f(N(x_i - x_j)).$$

$$= \int_{-\infty}^{\infty} f(x) R_2(x) dx?$$

- For IID, $R_2(x) \equiv 1$.

- Define higher level correlations similarly.

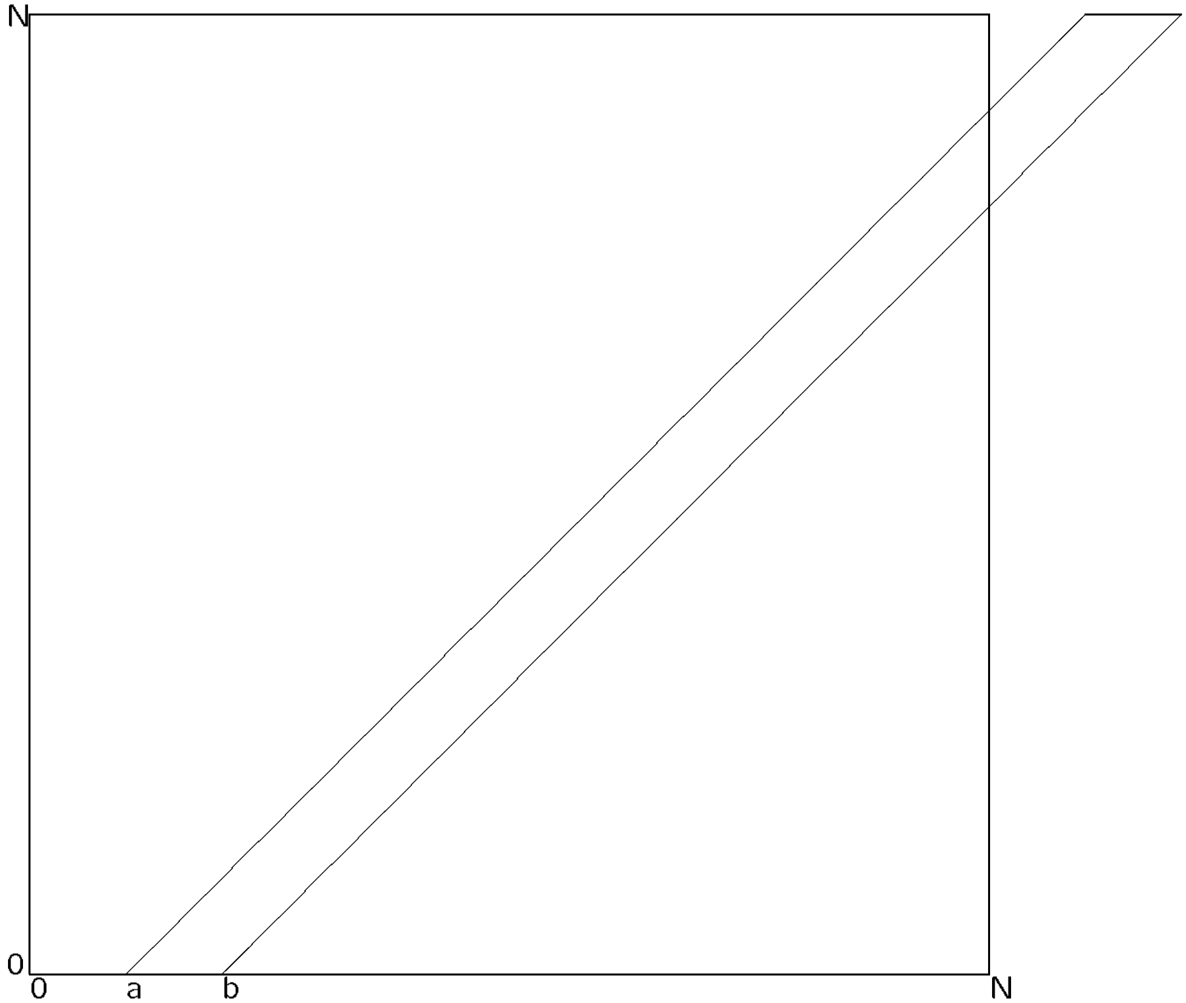
- **Knowledge of all correlations determines level spacing distribution.**

- Why divide by N when there are N^2 pairs of points!?

- Consider $(Nx_i, Nx_j) \in \mathbf{R}^2$ such that $N(x_i - x_j) \in [a, b]$. That is, put N^2 points in a square of area N^2, so average density is one.

- Region $N(x_i - x_j) \in [a, b]$ is (almost) parallelogram of width $(b - a)$ and height N. (See picture on next slide.)

- Thus expect $N(b - a)$ points in parallelogram, so should divide by N, not N^2.


N
0
0
a
b
N

Primes

- By Prime Number Theorem,

$$s_n = \frac{p_{n+1} - p_n}{\log p_n}$$

has mean 1.

- Numerics indicate that the distribution of *normalized* spacings s_n between primes is Poissonian.

- Some rigorous results:

Theorem (Erdös, 1935).

$$\limsup_{n\to\infty} s_n = \infty$$

Theorem (Maier, 1985).

$$\liminf_{n\to\infty} s_n < 0.248$$

- **Theorem (Hildebrand-Maier, 1988).** *Let X be the set of limit points of $\{s_n\}_{n=1}^{\infty}$. Then a positive proportion of all positive reals are in X. I.e. there is $C > 0$ so that for all $t > 0$,*

$$\mu(X \cap [0, t]) > Ct$$

- **Theorem (Gallagher, 1976).** *Strong version of prime k-tuple conjecture, i.e.,*

$$|\{n \leq x : n+d_1, n+d_2, \dots, n+d_r \text{ all prime}\}|$$

$$\sim C(d_1, d_2, \dots, d_r)\frac{x}{(\log x)^r},$$

implies that normalized spacings between primes are Poisson.

Invertible elements mod q

- For q some integer, let
$$0 < a_1 < a_2 < \ldots < a_{\phi(q)} < q$$
be units in $\mathbf{Z}/q\mathbf{Z}$.

- Exercise: Why equidistributed?

- Mean spacing is given by $q/\phi(q)$.

- In order to get arbitrary small normalized spacings, need to take sequence of integers q_n such that mean spacing tends to infinity. I.e., let q_n be the product of the n first primes.

- **Theorem (Hooley, 1965).** *If the mean spacing tends to infinity, then the limiting (normalized) spacing distribution between units modulo q is Poissonian.*

Sketch of proof for pair correlation.

- By assumption, average spacing $s_q = q/\phi(q)$ tends to ∞.

- Let I be an interval not containing 0. Pair correlation is given by

$$R_2(I,q) = \frac{1}{N_q} \sum_{h \in s_q I} N(h,q)$$

where $N_q = \phi(q)$ is the number of units and

$$N(h,q) = |\{u_1, u_2 \in (\mathbf{Z}/q\mathbf{Z})^\times : u_1 - u_2 \equiv h \mod q\}|$$

- I.e. put points on unit circle instead of unit interval. (As $s_q \to \infty$ get same answer!)

- Note: $N(h,q)$ is periodic modulo q.

- Pretend that that $s \gg q$. (Clearly it is not!) Then summing over about $s_q|I|/q$ full periods, so

$$\sum_{h \in s_q I} N(h,q) \simeq \frac{s_q|I|}{q} \sum_{h \mod q} N(h,q)$$

- Now,

$$\sum_{h \mod q} N(h,q) = N_q^2$$

since it counts all pairs of units u_1, u_2.

- Get

$$R_2(I,q) \simeq \frac{s_q|I|}{qN_q} N_q^2 = \frac{\frac{q}{N_q}|I|}{q} N_q = |I|$$

- With some work, can justify

$$\sum_{h \in s_q I} N(h,q) \simeq \frac{s_q|I|}{q} \sum_{h \mod q} N(h,q)$$