Erdös-Rényi laws for dynamical systems and large deviations.

Nicolai Haydn (USC), Holger Kantz (MPIKS), Mozhdeh Massah (MPIKS) and Matthew Nicol (UH). Partly supported by NSF-DMS 1600780

New Developments In Open Dynamical Systems.
March 21, 2018

Erdös-Renyi Laws

Suppose (T, X, μ) is an ergodic dynamical system and $\phi : X \to \mathbb{R}$ is an observable, $\int \phi d\mu = 0$.
Erdös-Renyi laws give the almost sure behavior of averages over time windows of varying length. Define

$$S_n(x) = \sum_{j=0}^{n-1} \phi \circ T^j(x)$$

Define the maximum average over a window of length $k(n)$ up to time n, $\theta(n, k(n))$, by

$$\theta(n, k(n)) := \max_{0 \leq j \leq n-k(n)} \frac{S_{j+k(n)} - S_j}{k(n)}$$

• If $k(n) = 1$ then for μ a.e. x, $\theta(n, k(n))(x) \to \text{essup } \phi$.
• If $k(n) = n$ then by the ergodic theorem for μ a.e. x, $\theta(n, k(n))(x) \to 0$.

Erdös-Renyi Law for IID processes

The Erdös-Rényi law was first given for iid random variables by Erdös and Rényi in " On a new law of large numbers" (1970):

Proposition (Erdös-Rényi)

Let $(X_n)_{n\geq 1}$ be an iid sequence of centered non-degenerate random variables, and let $S_j = X_1 + ... + X_j$. Assume that the moment generating function Ee^{tX_1} exists in some interval U containing $t = 0$. For each $\alpha > 0$, define $\psi_\alpha(t) = e^{-\alpha t} Ee^{tX_1}$. For those α for which ψ_α attains its minimum at a point $t_\alpha \in U$, set $I(\alpha) = \alpha t_\alpha - \log Ee^{t_\alpha X_1}$. Then almost surely

$$\lim_n \max\{(S_{j+[\log n/I(\alpha)]} - S_j)/[\log n/I(\alpha)] : 1 \leq j \leq n - [\log n/I(\alpha)]\} = \alpha.$$

Example

Suppose X_i is an iid sequence taking the values ± 1 with equal probability $\frac{1}{2}$

Recall

$$\theta(n, k(n)) := \max_{0 \leq j \leq n-k(n)} \frac{S_{j+k(n)} - S_j}{k(n)}$$

$\theta(n, k(n))$ is the maximal average gain over a time window of length $k(n)$.

A calculation using the strong law of large numbers shows that if $\lim_{n\to\infty} \frac{k(n)}{\log n} = \infty$ then P a. s.

$$\lim_{n\to\infty} \theta(n, k(n)) = 0$$

If, however, $k(n) \leq c \log_2 n$ with $0 < c < 1$ then for large n with probability one there is at least one $j < n - k(n)$ such that $X_{j+1} = X_{j+2} = \ldots = X_{j+k(n)} = 1$ (an application of the Borel-Cantelli lemma) so P a. .s.

$$\lim_{n\to\infty} \theta(n, k(n)) = 1$$

So for a fair game the Erdös-Rényi law gives information on the maximal average gain of a player when the length of the time window ensures

$$\lim_{n\to\infty} \max_{0\leq j\leq n-k(n)} \frac{S_{j+k(n)} - S_j}{k(n)}$$

has a non-degenerate limit. In this case $I(\alpha) = 1 - h(\frac{1+\alpha}{2})$ where $h(x) = -x \log_2 x - (1-x)\log_2(1-x)$.

Erdös-Renyi laws for deterministic dynamical systems

Suppose $T : (X, \mu) \to (X, \mu)$ is an ergodic measuring preserving map and

$$\phi : X \to \mathbb{R}$$

is an integrable function (observable).

The sequence $\{\phi \circ T^j\}$ is a stationary stochastic process.

Is there an almost sure limit for maximal average gain?

Large deviations theory

Suppose

$$\int_X \phi \; d\mu = 0$$

.

Let $S_n(x) := \phi(x) + \phi \circ T + \ldots + \phi \circ T^{n-1}(x)$.
If (T, X, μ) is ergodic then

$$\lim_{n \to \infty} \frac{S_n(x)}{n} = 0$$

for μ a. e. $x \in X$.
Large deviations theory gives information on the rate of convergence by estimating

$$\mu(x : S_n(x) \geq n\alpha)$$

as a function of n and $\alpha > 0$.

Definition (Rate function)

A mean-zero observable $\phi : X \to \mathbb{R}$ is said to satisfy a local large deviation principle with rate function $I(\alpha)$, if there exists a neighbourhood U of 0 and a strictly convex function $I : U \to \mathbb{R}$, which is non-negative and vanishing only at $\alpha = 0$, such that

$$\lim_{n\to\infty} \frac{1}{n} \log \mu(x : S_n(x) \geq n\alpha) \quad = \quad -I(\alpha) \tag{1}$$

for all $\alpha > 0$ in U and

$$\lim_{n\to\infty} \frac{1}{n} \log \mu(x : S_n(x) \leq n\alpha) \quad = \quad -I(\alpha) \tag{2}$$

for all $\alpha < 0$ in U.

For a given $\epsilon > 0$ for large n

$$e^{-n(I(\alpha)+\epsilon)} \leq \mu(x : S_n(x) \geq n\alpha) \leq e^{-n(I(\alpha)-\epsilon)}$$

Proposition (adapted from Erdös and Rényi.)

(a) Suppose that ϕ satisfies a large deviation principle with rate function I defined on the open set U. Let $\alpha > 0$ and let

$$L_n = L_n(\alpha) = \left[\frac{\log n}{I(\alpha)}\right] \qquad n \in \mathbb{N}.$$

Then the Upper Erdös-Rényi law holds and

$$\limsup_{n\to\infty} \max\{S_{L_n}(\phi) \circ T^j / L_n : 0 \le j \le n - L_n\} \le \alpha.$$

(b) If for each interval A there exists $C > 0, \tau \ge 1$ such that

$$\mu\Big(\bigcap_{m=0}^{n-L_n} \{S_{L_n}(\phi) \circ T^m \in A\}\Big) \le C[\mu(S_{L_n} \in A)]^{n/(L_n)^\tau}$$

then the Lower Erdös-Rényi law holds and

$$\liminf_{n\to\infty} \max\{S_{L_n}(\phi) \circ T^j / L_n : 0 \le j \le n - L_n\} \ge \alpha.$$

Remark

If both upper and lower Erdös-Rényi laws hold then

$$\lim_{n\to\infty}[\max_{0\leq m\leq n-L_n}\frac{S_{L_n}\circ T^m}{L_n}]=\alpha$$

where

$$L_n=L_n(\alpha)=\left[\frac{\log n}{I(\alpha)}\right]\qquad n\in\mathbb{N}.$$

Earlier results establishing Erdös-Rényi laws include:

(a) Subshifts of finite type (Grigull, 1973)
(b) Uniformly expanding 1-d maps (Chazottes and Collet, 2005)
(c) Gibbs-Markov systems (Denker and Kabluchko, 2007)
(d) Non-uniformly expanding maps with exponential decay of correlations (Denker and N., 2013)
(e) In certain averaging setups and for nonconventional sums (Kifer, 2016 and 2017).

Theorem

Suppose that (T, X, μ) is a dynamical system modeled by a Young Tower with exponential tails i.e. (i) T admits a Markov tower extension with properties (P1)-(P5) in Young's 1998 paper; (ii) the return time function R satisfies $\mu(R > n) = O(e^{-\beta n})$ for some $\beta > 0$.
Assume $\varphi : X \to \mathbb{R}$ is Hölder with $\int \varphi \, d\mu = 0$ and $\varphi \neq \psi \circ T - \psi$ for any $\psi \in L^1(\mu)$.
Define $S_n(x) = \sum_{j=0}^{n-1} \varphi(T^j x)$. It is known that φ satisfies a local large deviation principle with nondegenerate rate function I defined on an open set $U \subset \mathbb{R}$ containing 0.
Let $\alpha > 0$ and define

$$L_n = L_n(\alpha) = \left[\frac{\log n}{I(\alpha)}\right] \qquad n \in \mathbb{N}$$

Then

$$\lim_{n\to\infty} \max_{0 \le j \le n - L_n} \frac{S_{L_n} \circ T^j(x)}{L_n} = \alpha.$$

for μ a.e. $x \in X$.

Sketch of proof:

(1) In this setting

$$\limsup_{n\to\infty} \max\{S_{L_n}(\phi)\circ T^j/L_n : 0 \le j \le n - L_n\} \le \alpha.$$

so we need only prove

$$\liminf_{n\to\infty} \max\{S_{L_n}(\phi)\circ T^j/L_n : 0 \le j \le n - L_n\} \ge \alpha.$$

(2) A local large deviation with rate function allows us to estimate $\mu\{S_{L_n} < L_n(\alpha - \epsilon)\}$ from below.
For any $\delta_1 > 0$ for large n we have
$\mu\{S_{L_n} > L_n(\alpha - \epsilon)\} \geq e^{-L_n(I(\alpha-\epsilon)+\delta_1)} \geq e^{-(\frac{I(\alpha-\epsilon)+\delta_1}{I(\alpha)})\log n}$.
For large n this implies

$$1 - \mu\{S_{L_n} \leq L_n(\alpha - \epsilon)\} \geq e^{-(1-\delta_2)\log n}$$

for some $0 < \delta_2 < \delta_1$.
Hence

$$\mu\{S_{L_n} \leq l_n(\alpha - \epsilon)\} \leq 1 - e^{-(1-\delta_2)\log n}$$

(3) For $\epsilon > 0$ let

$$C_m(\epsilon) := \{S_{L_n} \circ T^m \leq L_n(\alpha - \epsilon)\}$$

and

$$B_n(\epsilon) = \bigcap_{m=0}^{n-L_n} C_m(\epsilon)$$

We use decay of correlations and intercalate by blocks of length $(\log n)^\tau$, $\tau > 6$. We define

$$E_n(\epsilon) := \bigcap_{m=0}^{[(n-(\log n)^\tau)/(\log n)^\tau)]} C_{m[(\log n)^\tau]}(\epsilon)$$

The proof uses technical approximations e.g. take S_{L_n} as constant on stable manifolds and take Lipschitz approximations to indicator functions...

In the end we can estimate,

$$\begin{aligned} \mu(E_n(\epsilon)) &\leq C\left[1 - e^{-(1-\delta_2)\log n}\right]^{n/(\log n)^\tau} \\ &= O(\exp(-n^{\delta_3})) \end{aligned}$$

where δ_3 is any $0 < \delta_3 < \delta_2$. This is summable so the Borel-Cantelli lemma gives

$$\liminf_{n\to\infty} \max\{S_{L_n}(\phi) \circ T^j / L_n : 0 \leq j \leq n - L_n\} \geq \alpha.$$

Local large deviations for unbounded observables.

As an application of Erdös-Rënyi limit laws, the next example shows that if an observable is unbounded we should not expect exponential large deviations with a rate function.

Example

Suppose φ *is a continuous observable on* $(0,1]$ *such that* $\lim_{x\to 0}\varphi(x)=\infty$*,* $\int \varphi dx = 0$ *and* $\varphi > -\rho$ *for some* $\rho > 0$*. Let* (T, X, m) *be the tent map*

$$Tx = \begin{cases} 2x & \text{if } 0 \le x < \frac{1}{2}; \\ 2x-1 & \text{if } \frac{1}{2} < x \le 1. \end{cases}$$

Then the stationary stochastic process $\{\varphi \circ T^j\}$ *does not satisfy exponential large deviations with a rate function.*

Sketch of proof:
If φ satisfies a large deviation principle with rate function I defined on an open set U then:
if $\alpha \in U$ and

$$L_n = L_n(\alpha) = \left[\frac{\log n}{I(\alpha)}\right] \qquad n \in \mathbb{N}$$

the upper Erdős-Rényi law holds, that is, for μ a.e. $x \in X$

$$\limsup_{n\to\infty} \max\{S_{L_n}(\varphi) \circ T^j(x)/L_n : 0 \le j \le n - L_n\} \le \alpha.$$

Fix $\alpha > 0$ in U and let $M > \frac{2\log 2(\alpha+\rho)}{I(\alpha)}$. Choose N large enough that $\varphi(x) > M$ for all $x < \frac{1}{\sqrt{N}}$.

Phillipp showed that the tent map satisfies the Borel Cantelli property and that $T^n(x) \in [0, \frac{1}{n}]$ infinitely often almost surely since $\sum_{n=2}^{\infty} \frac{1}{n}$ diverges.
If $T^n(x) \in [0, \frac{1}{n}]$ then $T^{n+j}(x) \in [0, \frac{1}{\sqrt{n}}]$ for at least $j \geq \frac{\log n}{2 \log 2}$ iterates j (this estimate comes from solving $2^j \frac{1}{n} = \frac{1}{\sqrt{n}}$).
Moreover if $T^{n+j}(x) \in [0, \frac{1}{\sqrt{n}}]$ and $n > N$ then $\varphi(T^{n+j}(x)) \geq M$.
We take now $n > N$.
If $T^n x \in [0, \frac{1}{n}]$ then $S_{L_n}(\varphi) \circ T^n(x) > M(\frac{\log n}{2 \log 2}) - \rho \frac{\log(n)}{I(\alpha)}$ (as $\varphi \geq -\rho$). As $M > \frac{2 \log 2(\alpha+\rho)}{I(\alpha)}$ this implies that

$$\max\{S_{L_n}(\varphi) \circ T^j(x)/L_n : 0 \leq j \leq n - L_n\} > \alpha$$

which is a contradiction to the upper Erdős-Rényi law.
Hence exponential large deviations with a rate function cannot hold for this observable.

Exponential local large deviations without a rate function.

Examples exist in the literature (by Bradley, Orey and Pelikan,Bryc and Smolenski, Chung) of stationary processes which have exponential large deviations but a rate function does not exist i.e. defining $S_n = \sum_{j=0}^{n-1} X_j$ for all $\epsilon > 0$, there exists a constant $C(\epsilon)$ such that $P(|\frac{S_n}{n}| > \epsilon) \leq C(\epsilon)e^{-\gamma n}$, giving exponential convergence in the strong law of large numbers yet there is no rate function $I(\epsilon)$ controlling the rate of decay.
In particular there is an example of a mean zero bounded function f taking only 3 values on an aperiodic recurrent Markov chain (X_n) with a countable state space such that the system has exponential large deviations but does not have a rate function.

- Bradley (1989) produced an example of a stationary, pairwise independent, absolutely regular stochastic process for which the central limit theorem does not hold.
- Orey and Pelikan (1988) presented this system as an example of a strongly mixing shift for which the large deviation principle with rate function failed.
- Bryc and Smolenski (1993) showed that in this example there is in fact also an exponential convergence in the strong law of large numbers.
- Bryc and Smolenski's work was recast by Chung (2011) into dynamical systems language, and the system was expressed as a Young Tower (F, Δ, ν).

We recast as a dynamical system and show that f is a coboundary, in fact $f = \psi \circ F - \psi$ where ψ is unbounded but $\psi \in L^2$. This seems to have been overlooked in the literature.

Let Δ_0 be the base of a Young Tower Δ with Δ_0 partitioned into intervals $\Lambda_0, \Lambda_1, \ldots, \Lambda_k ...,$.
Take $m(\Lambda_k) = Ce^{-\frac{12^k}{2}}$ where C is a normalization constant.
Define the
return time function R on Λ_k by $R_{\Lambda_k} := R(k) = (2)12^k$
We now build the Tower Δ above the base. We write $\Lambda_{k,0} := \Lambda_k$ and define, for $0 \leq j \leq R(k) - 1$ the levels $\Lambda_{k,j}$ of the Tower lying above Λ_k by

$$\Delta = \bigcup_{k \in N^+, 0 \leq j \leq R_k - 1} \{(x, j) : x \in \Lambda_{0,k}\}$$

with the tower map $F : \Delta \to \Delta$ given by

$$F(x, j) = \begin{cases} (x, j+1) & \text{if } x \in \Lambda_{k,0}, j < R(k) - 1 \\ (T_k x, 0) & \text{if } x \in \Lambda_{k,0}, j = R(k) - 1 \end{cases}.$$

where T_k has constant derivative and maps $\Lambda_{k,0}$ onto Δ_0.
F maps $\Lambda_{0,0}$ bijectively onto Δ_0.

If $k \neq 0$ we define $f : \lambda_{k,j} \to \{-1, 0, 1\}$ by

$$f(x,j) = \begin{cases} 1 & \text{if } x \in \Lambda_k, j \leq 12^k - 1 \\ -1 & \text{if } x \in \Lambda_k, 12^k \leq j \leq 2.12^k - 1 \end{cases}.$$

if $k = 0$ we take $f(0,0) = 0$. This is the model of Bradley, Orey, Pelikan, Bryc and Chung.

Now define a function ψ, which will be a coboundary for f, by

$$\psi(x,j) = \begin{cases} j & \text{if } x \in \Lambda_k, 0 \leq j \leq 12^k \\ 2.12^k - j & \text{if } x \in \Lambda_k, 12^k < j \leq 2.12^k - 1 \end{cases}.$$

and $\psi(0,0) = 0$.
It is easy to check that

$$f = \psi \circ F - \psi$$

As far as we know there is no example of a non-degenerate bounded observable on a dynamical system which has exponential large deviations and yet no rate function.

Example

Let $\varphi(x) = -\log x$ *on the probability space* $([0,1], m)$*. Then* $\int \varphi dx = 1$ *and* $E[e^{t\varphi}] = \int_0^\infty e^{tx} e^{-x} dx$ *exists for* $t < 1$*.*
If X_i *is a sequence of i.i.d random variables with the same distribution function as* φ *and* $S_n = \sum_{j=1}^n X_j$ *then for* $0 < \epsilon < 1$

$$\lim_{n\to\infty} \frac{1}{n} \log P(\frac{S_n}{n} > 1 + \epsilon) = -\epsilon + \log(1+\epsilon) = I(\epsilon)$$

This is a simple large deviations calculation.

Example

Let $\varphi(x) = -\log x$ be an observable on the tent map (T, X, μ). It is possible to show that $\varphi(x) = -\log x$ has exponential decay of autocorrelations.

$$|\int(\varphi \circ T^n - 1)(\varphi - 1)\ dx| \leq Ce^{-\beta n}$$

However $\{\varphi \circ T^n\}$ has strictly stretched exponential large deviations.

Sketch of proof:

It is easy to show that for all $\epsilon > 0$ for all $\delta > 0$ and all sufficiently large n, $\mu(S_n - n > n\epsilon) > e^{-n^{1/2+\delta}}$.

To see this note that if $x \in [0, e^{-n^{1/2+\delta}(\log 2+1)}]$ then for $1 \leq j \leq n^{1/2+\delta}$, $2^j x \in [0, e^{-n^{1/2+\delta}}]$, so that $|S_n(x) - n| \geq n^{1+2\delta} - n$.

In the other direction, using results of Kessebohmer and Schindler (2017) on trimmed sums it is possible to show for any $\delta > 0$

$$m(|S_n - n| > n\epsilon) \leq Ce^{-n^{(1/2-\delta)}}$$

Does -log $|x - p|$ have exponential large deviations for 'generic' p?

Open questions and applications.

• Investigate exponential local large deviations for unbounded integrable observables on chaotic systems (e.g. -log $|DT_u|$ in systems with singularities).

Applications to time-series.

• We have also proven Erdös-Rényi type fluctuation laws for α-mixing processes of polynomial rate and a class of intermittent maps also with polynomial mixing rate.

• This suggests a simple test, based on the Erdös-Rényi limit law, to estimate the rate of convergence to the ergodic average of a stationary ergodic time-series of measurements $\{X_j\}$ on a physical system.

• The advantage of the test is that it only needs a given time-series, not a large number of repeat measurements (ensemble averages) and seems to work well in applications.