STATISTICAL PROPERTIES FOR CERTAIN DYNAMICAL SYSTEMS

Contents

1. Introduction to the main topics of the course

1.1. **Additive random processes.** Let $\xi_0, \xi_1, \dots, \xi_{n-1}, \xi_n, \dots$ be a sequence of independent and identically distributed (i.i.d.) real random variables. Let

$$S_n := \xi_0 + \xi_1 + \cdots + \xi_{n-1}$$

be the partial sum process and let

$$\frac{1}{n} S_n = \frac{1}{n} \, (\xi_0 + \xi_1 + \cdots + \xi_{n-1})$$

be the average partial sum process.

Question. What is the behavior of these averages when $n \to \infty$?

Remark 1.1. Recall that two random variables ξ_1 and ξ_2 are identically distributed if $\mathbb{P}\{\xi_1 \in E\} = \mathbb{P}\{\xi_2 \in E\}$ for any Borel measurable set $E \subset \mathbb{R}$. In this case $\mathbb{E}\xi_1 = \mathbb{E}\xi_2$ and in fact $\mathbb{E}\phi(\xi_1) = \mathbb{E}\phi(\xi_2)$ for any integrable function $\phi \colon \mathbb{R} \to \mathbb{R}$.

Recall also that the random variables $\xi_1, \dots, \xi_n$ are independent if for any Borel measurable sets $E_1, \dots, E_n \subset \mathbb{R}$,

$$\mathbb{P}\{\xi_1 \in E_1 \wedge \ldots \wedge \xi_n \in E_n\} = \mathbb{P}\{\xi_1 \in E_1\} \cdots \mathbb{P}\{\xi_n \in E_n\} \, .$$

Theorem 1.1 (The law of large numbers - LLN)**.** *Given i.i.d. sequence* $\xi_0, \xi_1, \ldots, \xi_{n-1}, \xi_n, \ldots$ *of real random variables, if* $\mathbb{E}\xi_0 < \infty$ *then*

$$\frac{1}{n} S_n \to \mathbb{E}\xi_0 \quad a.s.$$

In particular, convergence in probability also holds. That is, $\forall\, \epsilon > 0$,

$$\mathbb{P}\left\{ \left| \frac{1}{n} S_n - \mathbb{E}\xi_0 \right| > \epsilon \right\} \to \infty \qquad as \qquad n \to \infty.$$

Question. It is natural to ask if there is a rate of convergence to 0 of the probability of the tail event above. It turns out that there is, as shown by the large deviations principle (LDP) below.

Theorem 1.2 (LDP of Cramér)**.** *Assume that the common distribution of the i.i.d. sequence of real random variables* $\xi_0, \xi_1, \ldots, \xi_{n-1}, \xi_n, \ldots$ *satisfies a certain growth condition. Then* $\forall\, \epsilon > 0$,

$$\mathbb{P}\left\{ \left| \frac{1}{n} S_n - \mathbb{E}\xi_0 \right| > \epsilon \right\} \asymp e^{-c(\epsilon) n} \qquad as \qquad n \to \infty,$$

where $c(\epsilon) \asymp c_0 \epsilon^2$ *for some* $c_0 > 0$.

More precisely, assuming that the common distribution has finite exponential moments:

$$M(t) := \mathbb{E}\left(e^{t\xi_0}\right) < \infty \quad \forall t \in \mathbb{R},$$

it follows that

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left\{ \left| \frac{1}{n} S_n - \mathbb{E}\xi_0 \right| > \epsilon \right\} = -c(\epsilon)$$

where

$$c(\epsilon) = \sup_{t \in \mathbb{R}} \left(t\epsilon - \log M(t)\right)$$

is the Legendre transform of $\log M(t)$.

This rate function $c(\epsilon)$ *is strictly convex near* $\epsilon = 0$, *with* $c(0) = 0$, $c'(0) = 0$ *and* $c''(0) > 0$, *so that* $c(\epsilon) \asymp c_0 \epsilon^2$.

Remark 1.2. The LDP is a very precise but asymptotic result. We are usually more interested in *finitary*, albeit less precise results, which will be referred to as large deviations type (LDT) estimates. A typical such result is the folllowing.

Theorem 1.3 (Hoeffding's Inequality)**.** *Assume the much stronger growth condition* $|\xi_0| \le C$ *a.s. Then* $\forall\, \epsilon > 0$ *the following holds* for all $n \in \mathbb{N}$*:*

$$\mathbb{P}\left\{ \left| \frac{1}{n} S_n - \mathbb{E}\xi_0 \right| > \epsilon \right\} \le 2 e^{-(2C)^{-2} \epsilon^2 n}\ .$$

Question. What is the typical size of the sum process $S_n - n\mathbb{E}\xi_0$? Note that by the LLN, almost surely we have

$$\frac{S_n - n\mathbb{E}\xi_0}{n} \to 0,$$

which implies that $S_n - n\mathbb{E}\xi_0 \ll n$. It turns out that from a certain point of view, $S_n - n\mathbb{E}\xi_0 \asymp \sqrt{n}$. More precisely, the following central limit theorem (CLT) holds.

Theorem 1.4 (CLT of Lindeberg-Lévy)**.** *Consider an i.i.d. sequence $\xi_0, \xi_1, \ldots, \xi_{n-1}, \xi_n, \ldots$ of real random variables and assume that the variance $\sigma^2 = \mathbb{E}\xi_0^2 - (\mathbb{E}\xi_0)^2 \in (0, \infty)$. Then for all $[a, b] \subset \mathbb{R}$,*

$$\mathbb{P}\left\{\frac{S_n - n\mathbb{E}\xi_0}{\sigma\sqrt{n}} \in [a, b]\right\} \to \int_a^b e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}} \quad as \quad n \to \infty.$$

In other words, with the appropriate scaling we have the convergence in distribution to the standard normal distribution

$$\frac{S_n - n\mathbb{E}\xi_0}{\sigma\sqrt{n}} \xrightarrow{d} N(0, 1)\,.$$

1.2. **Multiplicative random processes.** Let μ be a probability measure on the group of matrices $\mathrm{GL}_2(\mathbb{R})$. Given $g_0, g_1, \ldots, g_{n-1}, g_n, \ldots$ an i.i.d. sequence of random matrices chosen according to the probability μ, consider

$$\Pi_n := g_{n-1} \cdots g_1\, g_0$$

the corresponding multiplicative process.

Recall that for a matrix $g \in \mathrm{GL}_2(\mathbb{R})$, the norm is its maximal expansion

$$\|g\| = \max_{\|v\|=1} \|gv\|$$

while the co-norm is its minimal expansion

$$m(g) = \min_{\|v\|=1} \|gv\| = \left\|g^{-1}\right\|^{-1}.$$

The LLN for additive random processes has the following analog for multiplicative random processes.

Theorem 1.5 (Furstenberg-Kesten)**.** *Assuming the integrability condition $\mathbb{E}\left(\log\|g\|\right)\, d\mu(g) < \infty$, there are two numbers $L^+(\mu) \ge L^-(\mu)$ called the maximal respectively the minimal Lyapunov exponents of μ such that*

$$\frac{1}{n}\log\|\Pi_n\| \to L^+(\mu), \quad a.s.$$

and

$$\frac{1}{n}\log\|\Pi_n^{-1}\|^{-1} \to L^-(\mu), \quad a.s.$$

In particular we also have convergence in probability: $\forall\,\epsilon > 0$,

$$\mathbb{P}\left\{\left|\frac{1}{n}\log\|\Pi_n\| - L^+(\mu)\right| > \epsilon\right\} \to 0 \quad as \quad n \to \infty\,.$$

Instead of the maximal (or minimal) expansion of the random matrix products, we may consider the expansion of any vector. That is, given $v \in \mathbb{R}^2$, $v \neq 0$ consider the random walk $\{g_{n-1}\cdots g_1 g_0 v\colon n \geq 0\}$.

Theorem 1.6 (Furstenberg-Kifer's non-random filtration)**.** *For any given vector $v \in \mathbb{R}^2$, $v \neq 0$, either*

$$\frac{1}{n}\log\|\Pi_n v\| \to L^+(\mu) \quad as \quad n \to \infty\,,$$

or

$$\frac{1}{n}\log\|\Pi_n v\| \to L^-(\mu) \quad as \quad n \to \infty\,.$$

Remark 1.3. It turns out that under certain generic conditions to be defined in the future (namely the irreducibility of the measure μ), we have that $\forall\, v \in \mathbb{R}^2, v \neq 0$ the almost sure limit is the maximal Lyapunov exponent:

$$\frac{1}{n}\log\|\Pi_n v\| \to L^+(\mu) \quad \text{a.s.}$$

Moreover, if $L^+(\mu) > L^-(\mu)$ then

$$\mathbb{E}\left(\frac{1}{n}\log\|\Pi_n v\|\right) \to L^+(\mu)$$

uniformly in v.

Question. It is natural to ask if in this multiplicative random setting there are analogues of the LDP, LDT and CLT from the additive setting. As shown below, the answer is affirmative, at least in the generic setting. The precise statements will be provided later.

Theorem 1.7 (LDP - Le Page)**.** *Under generic assumptions, if $L^+(\mu) > L^-(\mu)$, then $\forall\, v \in \mathbb{R}^2$, $v \neq 0$ and $\forall\,\epsilon > 0$,*

$$\mathbb{P}\left\{\left|\frac{1}{n}\log\|\Pi_n\| - L^+(\mu)\right| > \epsilon\right\} \asymp e^{-c(\epsilon)n} \quad as \quad n \to \infty.$$

Theorem 1.8 (LDT - Duarte, Klein)**.** *Under generic assumptions, if $L^+(\mu) > L^-(\mu)$, then $\forall\, v \in \mathbb{R}^2$, $v \neq 0$, $\forall\,\epsilon > 0$ and $\forall\, n \in \mathbb{N}$,*

$$\mathbb{P}\left\{\left|\frac{1}{n}\log\|\Pi_n\| - L^+(\mu)\right| > \epsilon\right\} \leq C e^{-c(\epsilon)n}$$

for some constant $C < \infty$ and $c(\epsilon) > 0$.

Theorem 1.9 (CLT - Le Page)**.** *Under generic assumptions, there is $\sigma \in (0, \infty)$ such that $\forall\, v \in \mathbb{R}^2$, $v \neq 0$,*

$$\mathbb{P}\left\{ \frac{\log\|\Pi_n v\| - nL^+(\mu)}{\sigma\sqrt{n}} \in [a, b] \right\} \to \int_a^b e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}} \quad \textit{as} \quad n \to \infty.$$

1.3. **Observed dynamical systems.** Let (M, f) be a dynamical system where M is a compact metric space and $f : M \to M$ is continuous. Consider an appropriate f-invariant measure $\nu \in \mathrm{Prob}(M)$.

Remark 1.4. Recall that v is called f-invariant if $f_*\nu = \nu$, which is equivalent to saying that $\nu(E) = \nu(f^{-1}(E))$ for all Borel measurable $E \subset M$.

Moreover, ν is called ergodic w.r.t. f if all f-invariant sets (i.e. E such that $E = f^{-1}(E)$) are of ν measure 0 or 1. Note that ergodic measures are extremal points in the the space of f-invariant measures (this space is convex and weak-$*$ compact).

The triple (M, f, ν) is called a measure-preserving dynamical system (MPDS). Given an observable $\xi : M \to \mathbb{R}$ in an appropriate space of functions, the quadruple (M, f, ν, ξ) is called an observed MPDS.

For all iterates j, consider the real-valued random variable on M

$$\xi_j := \xi \circ f^j\,.$$

Since ν is f-invariant, and hence f^j-invariant for all j, the sequence $\xi_0, \xi_1 \ldots, \xi_{n-1}, \xi_n, \ldots$ is identically distributed. However, in general this sequence is *not* independent.

Consider the sum process, that is, the Birkhoff sums

$$S_n\xi := \xi + \xi \circ f + \cdots + \xi \circ f^{n-1} = \xi_0 + \xi_1 + \cdots + \xi_{n-1}\,.$$

Birkhoff's ergodic theorem is a generalization of the LLN in this setting.

Theorem 1.10 (Birkhoff's ergodic theorem)**.** *Assume that ν is ergodic w.r.t. f and that $\int_M |\xi|\, d\nu < \infty$. Then*

$$\frac{1}{n} S_n\xi \to \int \xi d\nu \quad \nu\text{-a.e.}$$

In particular the convergence in measure also holds: $\forall\epsilon > 0$,

$$\nu\left\{ x \in M \colon \left| \frac{1}{n} S_n\xi(x) - \int_M \xi\, d\nu \right| > \epsilon \right\} \to 0 \quad \textit{as} \quad n \to \infty.$$

Question. A fundamental problem in ergodic theory is to establish statistical properties like LDP, LDT, CLT for various kinds of observed dynamical systems.

In other words, the question is to determine for which dynamical system (M, f), for which appropriate choice of f-invariant measures ν and for which kinds of observables ξ one has an LDT estimate

$$\nu \left\{ x \in M \colon \left| \frac{1}{n} S_n\xi(x) - \int_M \xi \, d\nu \right| > \epsilon \right\} \le C \, e^{-c(\epsilon) n}$$

or a CLT

$$\nu \left\{ x \in M : \frac{S_n\xi(x) - n \int \xi \, d\nu}{\sigma \sqrt{n}} \in [a, b] \right\} \to \int_a^b e^{-\frac{x^2}{2}} \, \frac{dx}{\sqrt{2\pi}} \, .$$

A short but vague answer is that systems with *some hyperbolicity* should satisfy such statistical properties. The question is extremely far reaching, and for now it only has a very incomplete rigorous answer.

Some of the main tools used to address it, which will make their entry in this course in due time, are the transition (or Markov) operator and the transfer (or Ruelle) operator.

2. Stochastic dynamical systems

3. Limit laws for multiplicative random systems

4. Limit laws for hyperbolic systems

5. Partially hyperbolic systems