BERNOULLI DYNAMICAL SYSTEMS AND LIMIT THEOREMS

Dalibor Volný

Université de Rouen

Dynamical system $(\Omega, \mathcal{A}, \mu, T)$: $(\Omega, \mathcal{A}, \mu)$ is a probablity space, $T : \Omega \to \Omega$ a bijective, bimeasurable, and measure preserving mapping.

The measure μ (or the transformation T) is called *ergodic* if for any set $A \in \mathcal{A}$, $T^{-1}A = A$ implies that A or $\Omega \setminus A$ is of measure 0.

We will suppose that $\mathcal{A}$ is countably generated.

For a measurable function f on Ω,

$$(f \circ T^i)_i = (U^i f)_i$$

is a (strictly) stationary sequence;
reciprocally, to each (strictly) stationary sequence of random variables $(X_i)_i$ there exists a dynamical system and a function f such that $(X_i)_i$ and $(f \circ T^i)_i$ are equally distributed.

Bernoulli dynamical systems

Bernoulli: there exists a measurable partition $\underline{A} = \{A_1, A_2, ...\}$ such that $T^i \underline{A}$ are mutually independent and generate all σ-algebra $\mathcal{A}$.
In a Bernoulli dynamical system there thus exists a sequence $(e_i)_i$, $e_i = e_0 \circ T^i$, of iid which generates $\mathcal{A}$, i.e. such that for any process $(f \circ T^i)_i$ there exists a measurable function g such that

$$f = g(\dots, e_{-1}, e_0, e_1, \dots).$$

For a relatively long time, in probabilists community there existed a conjecture saying that each (strictly) stationary process can be represented as a functional of iid.

Another community was interested by existence of a zero entropy process for which a CLT holds. A zero entropy process is not Bernoulli.

An example of a non Bernoulli dynamical system is an irrational circle rotation.

We define
$$\Omega = [0, 1),$$
$\mathcal{A} = \mathcal{B}$ is the Borel σ-algebra and $\mu = \lambda$ is the Lebesgue measure. We define $T = T_\theta$ by
$$Tx = x + \theta \ mod\ 1$$
where θ is an irrational number.

This dynamical system is of zero entropy and is not Bernoulli.

In particular, becuase for each $x \in [0, 1)$, $\{T^i x : i \in \mathbb{Z}\}$ is a dense set, there exist a *rigidity time*, i.e. a sequence (n_k) such that
$$T^{n_k} x \to x, \quad k \to \infty$$
hence for any measurable function f, the representation $f = g(\dots, e_{-1}, e_0, e_1, \dots)$ with $e_i = e_0 \circ T^i$ is impossible.

Zero entropy

Moreover, the dynamical system $([0, 1), \mathcal{B}, \lambda, T_\theta)$ is of zero entropy, i.e.
for any σ-algebra $\mathcal{C} \subset \mathcal{B}$ with $\mathcal{C} \subset T^{-1}\mathcal{C}$ we have $\mathcal{C} = T^{-1}\mathcal{C}$.

In a dynamical system of zero entropy we thus have no nontrivial martingale difference sequence $(m \circ T^i)_i$ and any process $(f \circ T^i)_i$ is deterministic in the sense that $f \circ T$ is measurable w.r.t. the past σ-algebra generated by $f \circ T^i$, $i \le 0$.

It should be noted that even in a Bernoulli dynamical system we can find a process $(f \circ T^i)_i$ which is deterministic and at the same time, the process $(f \circ T^{-i})_i$ is a sum of a martingale difference sequence and a coboundary.

In a (general) dynamical system $(\Omega, \mathcal{A}, \mu, T)$ there exists a σ-algebra $\mathcal{P} \subset \mathcal{A}$ such that

- $\mathcal{P} = T\mathcal{P} = T^{-1}\mathcal{P}$ and when restricting the measure μ on $\mathcal{P}$, we get a zero entropy dynamical system,
- $\mathcal{P}$ is maximal of that property.

Such a σ-algebra $\mathcal{P}$ is called *Pinsker* σ-algebra.

A process $(f \circ T^i)_i$ can in a unique way be represented using a sum
$$f = \big(f - E(f|\mathcal{P})\big) + E(f|\mathcal{P})$$
into a zero entropy process and a process which is a sum of martingale difference sequences.

Remark that for some time it was a question whether in a zero entropy dynamical system there can exist a process $(f \circ T^i)_i$ for which a CLT takes place. In 1987, Bob Burton and Manfred Denker showed that in every "reasonable" (aperiodic) dynamical system we have an $f \in L^2$ such that $S_n(f)/\|S_n(f)\|_2$ converge in distribution to $\mathcal{N}(0, 1)$ and later Volný in 1999 showed that the f can be found such that we have the CLT for $S_n(f)/\sqrt{n}$ and moreover we have both weak and strong invariance principle. There is no general theorem-condition for zero entropy processes, however.

Fourier transforms

In 2010, M. Pelirad and W.B. Wu published a strikingly strong result:

Theorem (Peligrad, Wu). *If $(X_j)_j$ is a (ergodic) stationary L^2 process process such that there exists a filtration of $\mathcal{F}_j \subset T^{-1}\mathcal{F}_j = \mathcal{F}_{j+1} \subset \mathcal{A}$, X_0 is $\mathcal{F}_\infty$-measurable, and $E(X_0|\mathcal{F}_{-\infty}) = 0$ then for almost every (w.r.t. the Lebesgue measure) $\theta \in (0, 2\pi)$ $(1/\sqrt{n})\sum_{j=1}^n e^{ij\theta}X_j$ converge to a normal law.*

Bernoulli dynamical systems belong to the family of K-mixing systems, i.e. such that

there exists a filtration $\mathcal{F}_j \subset T^{-1}\mathcal{F}_j = \mathcal{F}_{j+1} \subset \mathcal{A}$ such that $\mathcal{F}_{-\infty}$ is a trivial σ-algebra and $\mathcal{F}_\infty = \mathcal{A}$.

In such a dynamical system the assumptions of Peligrad-Wu's theorem are thus satisfied for any $f \in L^2$ with $Ef = 0$ (as noticed J.-P. Conze and G. Cohen).

The theorem, however, does not extend to any dynamical system:

Proposition. *There exists a dynamical system and a process of $X_j = f \circ T^j$, $f \in L^2$ and $Ef = 0$, such that for almost every (w.r.t. the Lebesgue measure) $\theta \in (0, 2\pi)$ $(1/\sqrt{n})\sum_{j=1}^n e^{ij\theta}X_j$ converge to no probability law.*

Martingale differences

Historically, first central limit theorems were given for independent random variables. An important contribution to the limit theory for dependent random variables was the Billingsley-Ibragimov CLT for (stationary, ergodic) martingale difference sequences.

One can ask a question whether the CLT remains true for subsequences, i.e. if the summation is done over more general sets $\Gamma_n \subset \mathbb{Z}$.

It is probably not very surprising that there exists a dynamical system, a martingale difference sequence $(f \circ T^j)_j$ and a sequence of $\Gamma_n \subset \mathbb{Z}$ with $|\Gamma_n| \to \infty$ such that

$$\frac{1}{\sqrt{|\Gamma_n|}} \sum_{j=1}^n f \circ T^j$$

converge to no probability law.

We can construct an example by a product of two dynamical systems taking

$$f \circ T^j = X_j Y_j$$

where $X_i = 1_A \circ T'^i$ for an irrational rotation T' and an interval A, (Y_i) is a sequence of iid with values $-1, 1$ (independent of (X_i)), $\Gamma_n = \{n_1, \ldots, n_n\}$ for a rigidity time (n_k).

Bernoulli random fields

In the same way as we defined Bernoulli dynamical systems for a $\mathbb{Z}$ action, i.e. for a family of transformations T^i, $i \in \mathbb{Z}$, we can define a Bernoulli dynamical system for a family of measurable and measure preserving transformations $T_{i_1,\dots,i_d}$, $(i_1,\dots,i_d) \in \mathbb{Z}^d$, where

$$T_{i_1,\dots,i_d} \circ T_{j_1,\dots,j_d} = T_{i_1+j_1,\dots,i_d+j_d}.$$

For a measurable function f we thus get a random field of $X_{i_1,\dots,i_d} = f \circ T_{i_1,\dots,i_d}$.

The dynamical system is Bernoulli if the σ-algebra $\mathcal{A}$ is generated by iid random variables $e_{i_1,\dots,i_d}$, $(i_1,\dots,i_d) \in \mathbb{Z}^d$.

The random field is Bernoulli if $f = g(e_i : i \in \mathbb{Z}^d)$.

In 2005, Wei Biao Wu introduced the "measure of physical dependence" for stationary processes in Bernoulli dynamical systems and in 2013, the notion was extended to Bernoulli random fields by El Machkouri, Volný and Wu.

It can be defined by

$$\Delta_p = \sum_{i \in \mathbb{Z}^d} \|f - g(e_j^* : e_j^* = e_j \text{ if } j \neq i,\ e_j^* = e_j' \text{ if } j \neq i,\ j \in \mathbb{Z}^d)\|_p$$

where e_j' is a copy of e_j independent of all e_i.

For $\Delta_2 < \infty$ El Machkouri, Volný and Wu proved that

If $|\Gamma_n| \to \infty$ then the Levy distance of $S_{\Gamma_n}/\sqrt{|\Gamma_n|}$ and $\mathcal{N}(0, \sigma_n^2/|\Gamma_n|)$ converge to zero ($\sigma_n = \|S_{\Gamma_n}\|_2$).

In the same paper, a weak invariance principle was proved for $\Delta_\rho < \infty$ where ρ is an L^p or a Luxemburg norm.

The results were proved using an inequality

$$\Big\| \sum_{i \in \Gamma} a_i X_i \Big\|_p \leq \Big(2p \sum_{i \in \Gamma} a_i^2 \Big)^{1/2} \Delta_p,$$

$X_i = f \circ T_i$, and approximation by m-dependent random fields.

Hannan's condition

One of most useful assumptions guaranteeing CLT and WIP for stationary processes is the Hannan's condition. The condition can be used for random arrays as well.

One dimensional case :
We are given a filtration $(\mathcal{F}_i)_i$ with

$$\mathcal{F}_i \subset T^{-1}\mathcal{F}_i = \mathcal{F}_{i+1},$$

for $f \in L^2$ define

$$P_i f = E(f|\mathcal{F}_i) - E(f|\mathcal{F}_{i-1}).$$

The Hannan's condition is satisfied if

$$f = \sum_{i\in\mathbb{Z}} P_i f$$

(the process $(f \circ T^i)_i$ is *regular*) and

$$\sum_{i\in\mathbb{Z}} \|P_i f\|_2 < \infty.$$

The condition implies CLT and WIP.

For fields : Let

$$\mathcal{F}_i = \sigma(e_j : j \leq i)$$

where $i \in \mathbb{Z}^d$ and $j \leq i$ means inequality for all coordonates.

For a Bernoulli random field, the σ-algebras $\mathcal{F}_i$ commute in the sense that for f integrable

$$E\Big(E(f|\mathcal{F}_i)\big|\mathcal{F}_j\Big) = (E(f|\mathcal{F}_i \cap \mathcal{F}_j).$$

Similarly as in the one dimensional case we can define orthogonal projection operators P_i.

If $f = P_0 f$ we say that the random field $(f \circ T_i)_i$ is a field of martingale differences.

While in the one dimensional case, Billingsley-Ibragimov theorems guarantees a CLT for martingale difference sequences, in dimension 2 and higher there are counterexamples. If the random field is Bernoulli, the CLT, nevertheless, takes place.

In 2013 (Statistica Sinica), Y. Wang and M. Woodroofe got a central limit theorem and invariance principle for Bernoulli random field under Maxwell-Woodroofe's condition; the summation was over rectangles.

In 2014, D. Volný and Y. Wang got the results under Hannan's condition. They also showed that L^2 Hannan's condition is a weaker assumption than $\Delta_2 < \infty$. The result of Volný and Wang is stronger than the result of Wang and Woodroofe in the sense that for the WIP, only finite second moments are needed.

When passing to a summation over general sets $\Gamma_n \subset \mathbb{Z}^d$, some problems occur.

Already in dimension 1, the constructions of zero entropy processes are based on summation over segments and for general Γ_n the problem opens anew.

The variance plays important role.

Already in dimension one, the Hannan's condition guarantees a CLT for $S_n/\sqrt{n}$ but if $\sigma_n = \|S_n\|_2 = o(\sqrt{n})$, S_n/σ_n can converge to no limit. Nevertheless, we have:

Proposition. *Let $(e \circ T_i)$ be a (Bernoulli) random field of martingale differences,*

$$f = \sum_{i \in \mathbb{Z}^d} a_i e \circ T_{-i}$$

where $\sum_{i \in \mathbb{Z}^d} a_i^2 < \infty$. Let $\Gamma_n \subset \mathbb{Z}^d$, $\sigma_n = \| \sum_{i \in \Gamma_n} f \circ T_i \|_2 \to \infty$.
Then $(1/\sigma_n) \sum_{i \in \Gamma_n} f \circ T_i$ converge in distribution to $\mathcal{N}(0,1)$.

Proposition. *Let $(f \circ T_i)$ be a (Bernoulli) random field which satisfies the Hannan condition, $\Gamma_n \subset \mathbb{Z}^d$, $\sigma_n = \| \sum_{i \in \Gamma_n} f \circ T_i \|_2 \to \infty$. If*

$$\liminf_{n \to \infty} \sigma_n / \sqrt{|\Gamma_n|} > 0$$

then $(1/\sigma_n) \sum_{i \in \Gamma_n} f \circ T_i$ converge in distribution to $\mathcal{N}(0,1)$.

Proposition. *Let $(f \circ T_i)$ be a (Bernoulli) random field which satisfies the Hannan condition,*

$$\lim_{n \to \infty} \frac{|\partial \Gamma_n|}{|\Gamma_n|} = 0, \quad |\Gamma_n| \to \infty.$$

Then $(1/\sqrt{|\Gamma_n|}) \sum_{i \in \Gamma_n} f \circ T_i$ converge in distribution to $\mathcal{N}(0, \sigma^2)$ where $\sigma^2 = \sum_{k \in \mathbb{Z}^d} f U_k f$.

Under Hannan's condition we get also the WIP like in El Machkouri, Volný and Wu, replacing the assumption $\Delta_p < \infty$ by $\sum_i \|P_i f\|_p < \infty$.

Martingale-coboundary decomposition

In the one-dimensional case an important tool for proving limit theorems has been the martingale-coboundary decomposition

$$f = m + g - g \circ T \tag{*}$$

where $(m \circ T^i)_i$ is a martingale difference sequence. The decomposition was introduced by M.I. Gordin in 1969. For simplicity let us consider adapted processes (and random fields) only with a filtration $(\mathcal{F}_i)_i$. Gordin showed that (*) holds in L^2 if

$$\sum_{i=0}^{\infty} \|E(U^i f | \mathcal{F}_0)\|_2 < \infty.$$

The condition becomes necessary and sufficient if we consider convergence of the series $\sum_{i=0}^{\infty} E(U^i f | \mathcal{F}_0)$ (Volný 1993); the result (convergence) taking place in L^p with $1 \le p < \infty$.

D. Giraudo and M. El Machkouri generalised this condition to random fields proving an L^p martingale-coboundary decomposition for a $\mathbb{Z}^d$ random field $(U_i f)_i$ if

$$\sum_{j=1}^{d} \sum_{k=1}^{\infty} k^{d-1} \|E(U_{e_j}^k f | \mathcal{F}_0)\|_p < \infty$$

where $U_{e_j} = U_{0,\ldots,1,\ldots}$, the 1 being on j-th place.

For L^2 this can be improved and a necessary and sufficient condition can be given.

The martingale-coboundary decomposition can bring estimates of probabilities of large deviations.