Nonconventional Large Deviations and Related Problems

Yuri Kifer

Hebrew University of Jerusalem

Centre Bernoulli, EPFL, March 2013

Nonconventional ergodic theorems

Nonconventional ergodic theorems studied the limits of expressions

$$\lim_{N\to\infty} 1/N \sum_{n=1}^{N} T^{q_1(n)} f_1 \cdots T^{q_\ell(n)} f_\ell$$

where T is an ergodic (or weakly mixing) measure μ preserving transformation, f_i's are bounded measurable functions and q_i's are polynomials taking on integer values on the integers (Bergelson, Furstenberg, Weiss: L^2-convergence, Assani: almost sure convergence under additional conditions). Recently such results were extended to the continuous time i.e. to expressions

$$\frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} T^{q_1(t)} f_1 \cdots T^{q_\ell(t)} f_\ell dt$$

where T^s is now an ergodic measure preserving flow (Potts, Bergelson-Leibman-Moreira). Application to multiple recurrence: take f_j to be the indicator $\mathbb{I}_{\Gamma_j}$ of some sets Γ_j.

The next natural step was to study central limit theorem type results for such expressions which we did together with Varadhan. Namely, we obtained the functional central limit theorem for

$$\frac{1}{\sqrt{N}}\sum_{n=1}^{[Nt]}\left(B(\xi(q_1(n)),...,\xi(q_\ell(n)))-\bar{B}\right) \text{ and}$$
$$\frac{1}{\sqrt{N}}\int_0^{[Nt]}\left(B(\xi(q_1(s)),...,\xi(q_\ell(s)))-\bar{B}\right)ds$$

where $\{\xi(n),\, n\geq 0\}$, (or $\{\xi(t)\},\, t\geq 0$) is a sufficiently fast mixing vector valued stochastic process with mild stationarity properties satisfying certain moment conditions, B is Hölder continuous, $\bar{B}=\int Bd(\mu\times\cdots\times\mu)$, $\xi(t)$ has distribution μ, $q_j=jt$ for $j\leq k$ and $q_j(t)$, $j>k$ satisfy some fast growth conditions. In the discrete time case results are readily applicable to fast mixing dynamical systems (subshifts of finite type, hyperbolic and expanding transformations etc.) with $\xi(m)=\xi(m,x)=T^mx$. For appropriate flows such as hyperbolic ones these results were not yet proven.
Warning: Summands in nonconventional sums are usually strongly dependent.

Large deviations (1st level) in our situation are supposed to give estimates as $N \to \infty$ for probabilities

$$\ln P\{\frac{1}{N}\sum_{j=1}^{N} B(\xi(q_1(n)),...,\xi(q_\ell(n))) \in \mathbf{I}\}$$

where $\mathbf{I}$ is an interval (closed: upper bound, open: lower bound). The asymptotic here is supposed to be of the form $-N \inf_{x\in\mathbf{I}} J(x)$ where $J(x) \geq 0$ is to be found. A similar problem arises in the continuous time case with the integral in place of the sum. Together with Varadhan we derived such estimates in some cases. In the dynamical systems case we consider expressions of the form

$$\ln \mu\{x : \frac{1}{N}\sum_{j=1}^{N} B(T^{q_1(n)}x,...,T^{q_\ell(n)}x) \in \mathbf{I}\}$$

for an appropriate measure μ (say, the normalized Riemannian volume in the hyperbolic and expanding transformations cases), and, again, the sum is replaced by the integral in the case of flows.

General fact: if the limit

$$Q(B,\lambda)=\lim_{N\to\infty}\frac{1}{N}\ln\int\exp\big(\lambda\sum_{j=1}^{N}B(\xi(q_1(n)),...,\xi(q_\ell(n)))\big)dP$$

exists for any λ and it is differentiable in λ then $J(x)=\sup_\lambda(x\lambda-Q(\lambda))$ is the rate function of large deviations estimates, i.e.

$$\limsup_{N\to\infty}\frac{1}{N}\ln P\{\frac{1}{N}\sum_{j=1}^{N}B(\xi(q_1(n)),...,\xi(q_\ell(n)))\in K\}\leq-\inf_{x\in K}J(x)$$

for any closed set $K\subset\mathbb{R}$, while for any open set $U\subset\mathbb{R}$,

$$\liminf_{N\to\infty}\frac{1}{N}\ln P\{\frac{1}{N}\sum_{j=1}^{N}B(\xi(q_1(n)),...,\xi(q_\ell(n)))\in U\}\geq-\inf_{x\in U}J(x).$$

We will explain how to deal with such limits in our nonconventional setup.

Theorem

Let $\xi(1), \xi(2), \ldots$ be a Markov chain on a space M having a transition density $p(j,x,y)$ with respect to some probability measure ν which for some j_0 satisfies $0 < \delta \leq p(j_0,x,y) \leq \delta^{-1} < \infty$ for $\forall x, y \in M$. Let $W_\lambda(x_1, \ldots, x_\ell)$ be a bounded (in x variables) with a bounded derivative in λ measurable function on $(-\infty,\infty) \times M^\ell$ and q_j, $j = 1, \ldots, \ell$ be positive integer valued increasing functions such that $q_1(n) = n$, $q_j(n+1) - q_j(n) \to \infty$ as $n \to \infty$ for all $j \geq 2$ and $q_j(n-1) \geq q_{j-1}(n)$ for all $n \geq n_0$ and $j \geq 2$. Then the limit

$$Q(W_\lambda) = \lim_{N\to\infty} \frac{1}{N} \ln \int \exp\Big(\sum_{n=1}^{N} W_\lambda(\xi(q_1(n)), \ldots, \xi(q_\ell(n)))\Big) dP$$

exists and it is differentiable in λ. In fact, $Q(\lambda) = \ln(specradR_\lambda)$ where $R_\lambda g(x) = E_x\big(g(\xi(1))\hat{W}_\lambda(\xi(1))\big)$,

$$\hat{W}_\lambda(x) = \int \cdots \int \exp\big(W_\lambda(x, x_2, \ldots, x_\ell)\big) d\mu(x_2) \cdots d\mu(x_\ell)$$

and μ is the unique invariant measure of the Markov chain.

Theorem

Let W_λ and q_j's be as before and let T be a C^2 expanding endomorphism or an Axiom A (in particular, Anosov) diffeomorphism in a small neighborhood $\mathcal{O}$ of an attractor on a compact Riemannian manifold M. Let $\Gamma = M$ in the case of an expanding endomorphism or an Anosov diffeomorphism and $\Gamma = \mathcal{O}$ in the Axiom A case. If ν is the normalized Riemannian volume then the limit

$$Q(W_\lambda) = \lim_{N\to\infty} \frac{1}{N} \ln \int_\Gamma \exp\Big(\sum_{n=1}^N W_\lambda(T^{q_1(n)}x, ..., T^{q_\ell(n)}x)\Big) d\nu(x)$$

exists and it is differentiable in λ. Moreover,

$$Q(W_\lambda) = \Pi\big(\ln \hat{W}_\lambda + \varphi\big)$$

where Π is the topological pressure for T, φ is the minus logarithm of the differential expanding coefficient on unstable leaves and $\hat{W}_\lambda$ as in the previous theorem with $\mu = \mu_{SRB}$ being the Sinai-Ruelle-Bowen measure. A similar result holds true when T is a topologically mixing subshift of finite type with $\nu = \mu$ being a Gibbs measure with a potential φ.

Reduction to linear q_i's ($k = \ell$)

Let $q_i(n) = in$ for $i = 1, ..., k \leq \ell$ and $q_j(n+1) - q_j(n) \to \infty$ as $n \to \infty$ and $q_j(n-1) \geq q_{j-1}(n)$ $\forall j > k$ and $n \geq n_0$.

Proposition

Let $V(x_1, ..., x_\ell)$ be a bounded continuous function and a dynamical system $\{T^n\}$ be as before (though here only ψ-mixing suffices). Then,

$$\lim_{N\to\infty} \frac{1}{N}\Big(\ln \int \exp\big(\sum_{n=1}^N V(T^{q_1(n)}x, ..., T^{q_\ell(n)}x)\big) d\nu(x) - \ln \int \exp\big(\sum_{n=1}^N V^{(k)}(T^n x, T^{2n}x, ..., T^{kn}x)\big) d\nu(x)\Big) = 0$$

where for each $m < \ell$,

$$V^{(m)}(x_1, ..., x_m) = \ln \int_M ... \int_M \exp(V(x_1, ..., x_m, x_{m+1}, ..., x_\ell)) d\mu(x_{m+1})...d\mu(x_\ell) \text{ and } V^{(\ell)} = V.$$

The same result holds true if we replace $T^n x$ by $\xi(n)$—a Markov chain satisfying conditions as before. If $k = 1$ this reduces the problem to the well known situation. For $k > 1$ the problem becomes complicated and we consider next the case of i.i.d. $\xi(n)$'s.

We obtain LD for $S_N(F) = \sum_{n=1}^{N} V(\xi(n), \xi(2n), ..., \xi(kn))$ where $\xi(n), n \geq 1$ are i.i.d. random variables (vectors) with a compact support M. Let $r_1, ..., r_m \geq 2$ be all primes not exceeding k. Set $A_n = \{a \leq n :\ a \text{ is relatively prime with } r_1, ..., r_m\}$ and $B_\eta(a) = \{b \leq \eta :\ b = ar_1^{d_1} r_2^{d_2} \cdots r_m^{d_m} \text{ for some nonnegative integers } d_1, ..., d_m\}$. Now for any bounded measurable function V on M^k write

$$S_N(V) = \sum_{a \in A_N} S_{N,a}(V) \text{ with } S_{N,a}(V) = \sum_{b \in B_N(a)} V(\xi(b), \xi(2b), ..., \xi(kb)).$$

Observe that $S_{N,a}(V)$, $a \in A_V$ are independent.

Theorem

For any continuous function V on M^k the limit

$$Q(V) = \lim_{N \to \infty} \frac{1}{N} \ln E \exp\left(\sum_{n=1}^{N} V(\xi(n), \xi(2n), ..., \xi(kn))\right)$$
$$= \lim_{N \to \infty} \frac{1}{N} \sum_{a \in A_N} \ln E \exp S_{N,a}(V)$$

exists and the functional $Q(V)$ is convex and lower semi-continuous. If $V = V_\lambda$ depends on a parameter λ and has a bounded derivative in λ then $Q(V_\lambda)$ is also differentiable in λ. Thus taking $V_\lambda = \lambda F$ we obtain that here also for $k \geq 2$ both LD bounds hold true with the rate functional J being the Fenchel-Legendre transform $J(u) = \sup_\lambda(\lambda u - Q(\lambda F))$ of Q.

As a model application of the above theorem we can consider digits $\xi(n)=\xi(n,\omega),\ n\geq 1$ of base M expansions $\omega=\sum_{n=1}^{\infty}\frac{\xi(n,\omega)}{M^n}$, $\xi(n,\omega)\in\{0,1,...,M-1\}$ of numbers $\omega\in[0,1)$ which are i.i.d. random variables on the probability space $([0,1),\mathcal{B},P)$ where $\mathcal{B}$ is the Borel σ-algebra and P is the Lebesgue measure. Take, for instance, $V(x_1,...,x_k)=\delta_{\alpha_1x_1}\delta_{\alpha_2x_2}\cdots\delta_{\alpha_kx_k}$ for some $\alpha_1,...,\alpha_k\in\{0,1,...,M-1\}$ with $\delta_{ij}=1$ if $i=j$ and $=0$, otherwise. Then the above theorem provides large deviations estimates for the number

$$n_{\alpha_1,...,\alpha_k}(N,\omega)=\#\{n\leq N:\ \xi(n,\omega)=\alpha_1,\xi(2n,\omega)=\alpha_2,\\ ...,\xi(kn,\omega)=\alpha_k\}=\sum_{n=1}^{N}V(\xi(n,\omega),...,\xi(kn,\omega)).$$

The same setup can be reformulated in the following way. Consider infinite sequences of letters (colors, spins, etc.) taken out of an alphabet of size M. Let $n_{\alpha_1,...,\alpha_k}(N)$ be the number of arithmetic progressions of length k with both the first term and the difference equal $n\leq N$ and having the letter (color, spin, etc.) α_i on the place $i=1,2,...,k$. Then the above theorem yields large deviations bounds for $n_{\alpha_1,...,\alpha_k}(N)$ as $N\to\infty$ considered as a random variable on the space of sequences of letters with any product probability measure, in particular, with uniform probability measure which assigns the same weight to each combination of n consecutive letters (i.e. to each cylinder set of length n) for all $n=1,2,....$

Write $B_N(a)=\{b\in B(a):\ b\leq N\}$ where

$B(a)=\{b\geq 1:\ b=ar_1^{d_1}r_2^{d_2}\cdots r_m^{d_m}$ for some nonnegative integers $d_1,...,d_m\}$.

Then $Z_N(V)=\prod_{a\in A_N}Z_{N,a}(V)$ where, recall, $A_n=\{a\leq n:a$ relatively prime with $r_1,...,r_m\}$ and

$$Z_{\eta,a}(V)=E\exp\big(\sum_{b\in B_\eta(a)}V(\xi(b),\xi(2b),...,\xi(kb))\big).$$

A crucial point here is that $Z_{N,a}(V)$ is determined only by $|B_N(a)|$ and not by N and a themselves. Set $\hat B_\eta(a)=B_\eta(a)\cup\{n:\ n=ln'$ for some $n'\in B_\eta(a)$ and $l=2,3,...,k\}$. Then we can write

$$Z_{\eta,a}(V)=\int...\int\exp\big(\sum_{b\in B_\eta(a)}V(x_b,x_{2b},...,x_{kb})\big)\prod_{b'\in\hat B_\eta(a)}d\mu(x_{b'}).$$

It is easy to see from here that $Z_{\eta,a}(V)=Z_{\eta/a,1}(V)$ for any $\eta>0$ and an integer $a\geq 2$ relatively prime with $r_1,...,r_m$.
Set

$$D(\rho)=\{n=(n_1,...,n_m)\in\mathbb{Z}^m:\ n_i\geq 0,\ i=1,...,m \text{ and } \sum_{i=1}^m n_i\ln r_i\leq\rho\}$$

then $D(\ln(N/a))$ is in one-to-one correspondence $(n_1,...,n_m)\leftrightarrow ar_1^{n_1}\cdots r_m^{n_m}$ with $B_N(a)$.

Let $l = |B_N(a)|$ and set $R_l(V) = Z_{N,a}(V)$ since the latter depends only on l (and, of course, on V). Denote

$$\rho_{\min}(l) = \inf\{\rho \geq 0 : \ |D(\rho)| = l\} \text{ and } \rho_{\max}(l) = \sup\{\rho \geq 0 : \ |D(\rho)| = l\}$$

which is well defined for each integer $l \geq 1$ and
$\rho_{\max}(l) > \rho_{\min}(l) \geq (l^{1/m} - 1)\ln 2$. Set $A_N^{(l)} = \{a \in A_N : \ |B_N(a)| = l\}$. Then a computation shows (by a kind of inclusion-exclusion argument) that

$$\lim_{N\to\infty} \frac{1}{N}|A_N^{(l)}| = (e^{-\rho_{\min}(l)} - e^{-\rho_{\max}(l)})r$$

where

$$r = 1 - \frac{1}{2} - \frac{1}{3} + \frac{1}{2\cdot 3} - \frac{1}{5} + \frac{1}{2\cdot 5} + \frac{1}{3\cdot 5} - \frac{1}{2\cdot 3\cdot 5} + \cdots + (-1)^m \frac{1}{r_1\cdot r_2\cdots r_m}.$$

It follows that

$$\tfrac{1}{N}\ln Z_N(V) = \tfrac{1}{N}\textstyle\sum_{a\in A_N} \ln Z_{N,a}(V) = \tfrac{1}{N}\sum_{1\leq l\leq (1+\frac{1}{\ln 2}\ln\frac{N}{a})^m} |A_N^{(l)}| \ln R_l(V)$$
$$\longrightarrow r\textstyle\sum_{l=1}^{\infty}(e^{-\rho_{\min}(l)} - e^{-\rho_{\max}(l)})\ln R_l(V) \text{ as } N\to\infty$$

and the last series converges absolutely. If $V = V_\lambda$ depends on λ and its derivative in λ exists and is bounded by $\tilde{C}$ then each $\ln R_l(V_\lambda)$ is differentiable in λ with a derivative bounded by $\tilde{C}l$. Hence, the whole above series is differentiable in λ and the assertion of Theorem follows.

We consider here occupational measures on M^ℓ,

$$\frac{1}{N}\sum_{n=1}^{N}\delta_{(\xi(n),\xi(q_2(n)),\dots,\xi(q_\ell(n)))}$$

where δ_a is the Dirac measure and $\xi(n)$, $n\geq 0$ is a Markov chain on a compact space M satisfying the Doeblin condition and having an invariant measure μ. For a continuous W on M^ℓ let $\hat{W}$ be as before. Let, again,

$$Q(W)=\lim_{N\to\infty}\frac{1}{N}\ln\int\exp\big(\sum_{n=1}^{N}W(\xi(q_1(n)),...,\xi(q_\ell(n)))\big)dP$$

then by the Donsker-Varadhan formula

$$Q(W)=\sup_{\nu\in\mathcal{P}(M)}\big(\int_M\ln\hat{W}(x)d\nu(x)-\hat{I}(\nu)\big)$$

where $\hat{I}(\nu)=-\inf_{u\in C_+(M)}\int\ln\frac{Pu}{u}d\nu$ and P is the transition operator of $\xi(n)$.

Next, let $Y_n^{(i)}$, $i = 2, ..., \ell$; $n = 0, 1, 2, ...$ be i.i.d. M-valued random variables with the distribution μ all of them independent of the Markov chain $\xi(n)$, $n \geq 0$. Then it is easy to see that

$$\lim_{N\to\infty} \frac{1}{N} \ln E \exp\big(\sum_{n=1}^{N} W(\xi(n), Y_n^{(2)}, ..., Y_n^{(\ell)})\big) = Q(W).$$

Indeed, let $\mathcal{F}_\xi$ be the σ-algebra generated by the Markov chain $\xi_n, n \geq 0$. Then

$$\begin{aligned} &E \exp\big(\textstyle\sum_{n=1}^{N} W(\xi(n), Y_n^{(2)}, ..., Y_n^{(\ell)})\big) \\ &= E\big(E(\exp(\textstyle\sum_{n=1}^{N} W(\xi(n), Y_n^{(2)}, ..., Y_n^{(\ell)}))|\mathcal{F}_\xi)\big) \\ &= E \exp(\textstyle\sum_{n=1}^{N} \ln \hat{W}(\xi(n))) \end{aligned}$$

implying the above formula.

But now we have the standard situation for the Markov chain $(\xi(n), Y_n^{(2)}, ..., Y_n^{(\ell)})$, $n \geq 0$, and so

$$Q(W) = \sup_{\nu \in \mathcal{P}(M \times \cdots \times M)} \Big(\int W(x_1, x_2, ..., x_\ell) d\nu(x_1, ..., x_\ell) - I(\nu)\Big)$$

where

$$I(\nu) = -\inf_{u \in C_+(M \times \cdots \times M)} \int_{M \times \cdots \times M} \ln \frac{E_{x_1} \int u(\xi(1), x_1, \ldots, x_\ell) d\mu(x_2) \ldots d\mu(x_\ell)}{u(x_1, \ldots, x_\ell)} d\nu(x_1, ..., x_\ell).$$

It is known here that there exists a unique $\nu = \nu_W$ on which the supremum above is attained and it follows from the standard theory that $I(\nu)$ is the rate functional for the 2nd level of large deviations for both occupational measures

$$\frac{1}{N} \sum_{n=1}^{N} \delta_{\left(\xi(n), Y_n^{(2)}, \ldots, Y_n^{(\ell)}\right)} \text{ and } \frac{1}{N} \sum_{n=1}^{N} \delta_{\left(\xi(n), \xi(q_2(n)), \ldots, \xi(q_\ell(n))\right)}.$$

Nonconventional LD theorems above and their continuous time counterparts can be extended to the corresponding averaging setups in the discrete and continuous time cases

$$X^\varepsilon(n+1) = X^\varepsilon(n) + \varepsilon B(X^\varepsilon(n), \xi(q_1(n)), ..., \xi(q_\ell(n))) \text{ and}$$
$$\frac{dX^\varepsilon(t)}{dt} = \varepsilon B(X^\varepsilon(t), \xi(q_1(t)), ..., \xi(q_\ell(t))).$$

First, we define the averaged vector field

$$\bar{B}_\nu(x) = \int B(x, \xi_1, \xi_2, ..., \xi_\ell) d\nu(\xi_1) d\mu(\xi_2) \cdots d\mu(\xi_\ell)$$

where μ is the unique invariant measure in the Markov chain or diffusion cases and μ is the SRB measure in the hyperbolic dynamical systems case. Next, for each a.c. curve γ_t, $t \in [0, \mathcal{T}]$ we define the functional

$$S_{0\mathcal{T}}(\gamma) = \int_0^{\mathcal{T}} \inf\{I(\nu) : \dot{\gamma} = \bar{B}_\nu(\gamma_t),\ \nu \text{ is } T-\text{invariant}\} dt$$

where T is a transformation (dynamical systems case) and $I(\nu)$ is the 2nd level LD functional on measures appeared many times in these lectures.

Theorem

Set $Z^\varepsilon(t) = X^\varepsilon([t/\varepsilon])$ or $Z^\varepsilon(t) = X^\varepsilon(t/\varepsilon)$ in the discrete or continuous time case, respectively. Then for any $a, \delta, \lambda > 0$ and every continuous γ_t, $t \in [0, \mathcal{T}]$, $\gamma_0 = x$ there exist $\varepsilon_0 > 0$ s.t. for $\varepsilon < \varepsilon_0$,

$$P\{\cdot : \rho_{0,\mathcal{T}}(Z_{x,\cdot}, \gamma) < \delta\} \geq \exp\{-\frac{1}{\varepsilon}(S_{0,\mathcal{T}}(\gamma) + \lambda)\} \quad \textit{and}$$

$$P\{\cdot : \rho_{0,\mathcal{T}}(Z_{x,\cdot}, \Phi^a_{0,\mathcal{T}}(x))) \geq \delta\} \leq \exp\{-\frac{1}{\varepsilon}(a - \lambda)\}$$

where P is the probability $(\cdot = \omega)$ in the Markov processes case, P is the normalized Riemannian volume $(\cdot = y,\ \xi(t) = T^t y)$ in the hyperbolic dynamical systems case, $\rho_{0,\mathcal{T}}$ is the uniform distance and $\Phi^a_{0,\mathcal{T}}(x) = \{\gamma : \gamma_0 = x,\ S_{0,T}(\gamma) \leq a\}$.

The main part of the proof is to show that for any continuous on $[0, T] \times M^\ell$ function $W_t(\xi_1, ..., \xi_\ell)$,

$$\lim_{\varepsilon\to 0} \varepsilon \ln \int \exp\left(\varepsilon^{-1} \int_0^{\mathcal{T}} W_t(\xi(q_1(t/\varepsilon)), ..., \xi(q_\ell(t/\varepsilon)))dt\right) dP$$
$$= \int_0^{\mathcal{T}} R(\hat{W}_t)dt$$

where, recall, $\hat{W}_\lambda(x) = \int \cdots \int \exp\left(W_\lambda(x, x_2, ..., x_\ell)\right) d\mu(x_2) \cdots d\mu(x_\ell)$ Here, again, P is the probability in the Markov processes case and P is the normalized Riemannian volume in the hyperbolic dynamical systems case where $\xi(s) = T^s y$ and integration then in y. In the discrete time case we replace t/ε by $[t/\varepsilon]$. When the limit above is established we obtain the theorem above via some general arguments in large deviations. It is easy to see that the limit above would follow if we could show that for any continuous functions $W_i(\xi_1, ..., \xi_\ell)$, $i = 1, ..., k$ and for any numbers $0 = t_0 < t_1 < \cdots < t_{k-1} < t_k = \mathcal{T}$,

$$\lim_{\varepsilon\to 0} \varepsilon \ln \int \exp\left(\varepsilon^{-1} \sum_{i=1}^k \int_{t_{i-1}}^{t_i} W_i(\xi(q_1(t/\varepsilon)), ...,\right.$$
$$\left.\xi(q_\ell(t/\varepsilon)))dt\right) dP = \sum_{i=1}^k (t_i - t_{i-1}) R(\hat{W}_i).$$

It is known that the multifractional formalism is related to large deviations and though I'll not use this connection here I'll discuss the corresponding problems. Recall that the multifractal formalism deals with computations of Hausdorff dimensions of sets having the form

$$\{x:\ \lim_{n\to\infty}\frac{1}{N}\sum_{n=1}^{N}f(T^nx)=\rho\}.$$

In our setup it is natural to study Hausdorff dimensions of more general sets

$$G_\rho=\{x:\ \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}F(f_1(T^{q_1(n)}x),...,f_\ell(T^{q_\ell(n)}x))=\rho\}.$$

By the nonconventional ergodic theorem if μ is T-inv. mixing and

$$\rho=\int...\int F(f_1(x_1),...,f_\ell(x_\ell))d\mu(x_1)\cdots d\mu(x_\ell)$$

then $\mu(G_\rho)=1$ while otherwise $\mu(G_\rho)=0$ and it is natural (if $\mu\sim Leb$) to inquire about the Hausdorff dimension of G_ρ.

Insead of this general problem we consider a more specific question about Hausdorff dimensions of sets of numbers with prescribed frequencies of specific combinations of digits in m-expansions. Namely, for any $x \in [0,1]$ and an integer $m > 1$ write

$$x = \sum_{i=1}^{\infty} \frac{a_{i-1}(x)}{m^i} \text{ where } a_j(x) \in \{0, 1, ..., m-1\},\, j = 0, 1, ...$$

allowing zero tails of expansions but not tails consisting of all $(m-1)$'s. This convention affects only a countable number of points, and so it does not influence computation of the Hausdorff dimensions. Denote by $\mathcal{A}_\ell = \{0, 1, ..., m-1\}^\ell$ the set of all ℓ-words. For each $x \in [0,1]$ and an ℓ-word $\alpha = (\alpha_1, \alpha_2, ..., \alpha_\ell) \in \mathcal{A}_\ell$ define

$$N_\alpha(x, n) = \#\{k > 0, k \le n : \, (a_{q_1(k)}(x), ..., a_{q_\ell(k)}(x)) = \alpha\}$$

where $\#\Gamma$ denotes the number of elements in the set Γ. For each probability vector $p = (p_\alpha,\, \alpha \in \mathcal{A}_\ell) \in \mathbb{R}^{m^\ell}$, $\sum_{\alpha \in \mathcal{A}_\ell} p_\alpha = 1$ define

$$U_p = \{x \in (0,1) : \, \lim_{n \to \infty} \frac{1}{n} N_\alpha(x, n) = p_\alpha \text{ for all } \alpha \in \mathcal{A}_\ell\}.$$

We want to deal with the question of computation of the Hausdorff dimension $HD(U_p)$ of U_p. When $\ell = 1$ and $q_1(k) = k$ we arrive at the classical question studied by Eggleston via combinatorial means and by Billingsley via the ergodic theory. In order to relate the limit of $n^{-1}N_\alpha(x,n)$ to the nonconventional ergodic theorem define the transformation $Tx = \{mx\}$ where $\{\cdot\}$ denotes the fractional part. Identifying 0 and 1 we can view T as an expanding map of the circle. Now $a_i(x) = a_0(T^i x)$ and if $\alpha = (\alpha_1, \alpha_2, ..., \alpha_\ell) \in \mathcal{A}_\ell$ and $\Gamma_j = \{x : a_0(x) = j\}$ then

$$N_\alpha(x,n) = \sum_{k=1}^{n} \mathbb{I}_{\Gamma_{\alpha_1}}(T^{q_1(k)}x)\mathbb{I}_{\Gamma_{\alpha_2}}(T^{q_2(k)}x)\cdots\mathbb{I}_{\Gamma_{\alpha_\ell}}(T^{q_\ell(k)}x).$$

Theorem

Suppose that $q_1(k) = k$ for all k and there exists a probability vector $r = (r_0, r_1, ..., r_{m-1})$ such that $p_\alpha = \prod_{i=1}^{\ell} r_{\alpha_i}$ for any $\alpha = (\alpha_1, ..., \alpha_\ell) \in \mathcal{A}_\ell$. Then for $p = (p_\alpha,\ \alpha \in \mathcal{A}_\ell)$,

$$HD(U_p) = \frac{-\sum_{j=0}^{m-1} r_j \ln r_j}{\ln m}$$

with the convention $0 \ln 0 = 0$.

Frequencies of words with gaps: remarks

- For any T-invariant probability measure μ on $[0,1]$ with sufficient mixing properties it follows that μ-almost everywhere

$$\lim_{n\to\infty}\frac{1}{n}N_\alpha(x,n)=\prod_{i=1}^{\ell}\mu(\Gamma_{\alpha_i}).$$

  Hence, if $p=(p_\alpha,\ \alpha\in\mathcal{A}_\ell)$ and there exists no probability vector $r=(r_0,r_1,...,r_{m-1})$ such that $p_\alpha=\prod_{i=1}^{\ell}r_{\alpha_i}$ then $\mu(U_p)=0$ for any μ as above, and so such μ cannot be used for computation of the Hausdorff dimension of U_p (by one of the methods where measures are involved) which complicates the study in this case.
- This type of results can be extended to digits of continued fraction expansions though in this case only estimates of the Hausdorff dimension rather than precise formulas can only be obtained.
- Some other cases related to the above theorem were considered by Peres and Solomyak and by Fan, Schmeling and Wu.

- G. Carinci, J-R Chazottes, C. Giardina and F. Redig, *Nonconventional averages along arithmetic progressions and lattice spin systems*, Indag. Math. 23 (2012), 589–602.
- A.-H. Fan, J. Schmeling and M. Wu, *Multifractal analysis of some multiple ergodic averages*, arXiv: 1212.2764.
- M. Kesseböhmer, *Large deviation for weak Gibbs measures and multifractal spectra*, Nonlinearity 14(2001), 395-409.
- Yu. Kifer, *Nonconventional law of large numbers and fractal dimensions of some multiple recurrence sets*, Stoch. Dynam. 12, No.3 (2012), 1150023.
- Yu. Kifer and S.R.S. Varadhan, *Nonconventional large deviations theorems*, Probab. Th. Relat. Fields, to appear.
- Y. Peres and B. Solomyak, *Dimension spectrum for a nonconventional ergodic average*, arXiv: 1107.1749.