Applications of ergodicity. The subadditive ergodic theorem

Thursday, 22 April 2021 13:07

Setup: probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a measure-preserving map $\varphi : \Omega \to \Omega$ $(\mathbb{P}(\varphi^{-1}A) = \mathbb{P}(A), \ \forall A \in \mathcal{F})$.

Invariant σ-algebra: The σ-algebra $\mathcal{I}$ of all events $A \in \mathcal{F}$ such that $\varphi^{-1}A = A$

(Note a RV X is meas. wrt. $\mathcal{I}$ iff $X(\omega) = X(\varphi(\omega)), \ \forall \omega \in \Omega$).

Stationarity: If X is any RV then let $X_n(\omega) := X(\varphi^n(\omega))$ (where $\varphi^0(\omega) = \omega$).
$n \geq 0$

Then $(X_0, X_1, X_2, \_\_) \stackrel{d}{=} (X_1, X_2, \_\_)$ $(*)$

Conversely, If $(X_n)_{n \geq 0}$ satisfies $(*)$, and then we call it <u>stationary</u>, then if we define a prob. space Ω to be the sequence space in which (X_n) lives with the induced σ-alg. and prob. measure, then the shift operation φ on this sequence space is measure preserving.

Ergodicity: The <u>dynamical system</u> $(\Omega, \mathcal{F}, \mathbb{P}, \varphi)$ is <u>ergodic</u> if $\mathbb{P}(A) \in \{0, 1\}$ for all $A \in \mathcal{I}$.

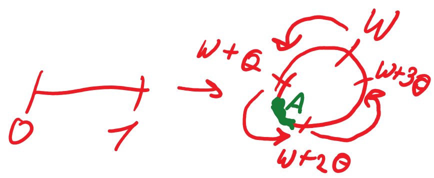
We also say a stationary seq. $(X_n)_{n \geq 0}$ is ergodic if all invariant events of the seq. have prob. 0 or 1 is ergodic if all invariant events which are fcns. of the seq. have prob. 0 or 1.

(Birkhoff) **The ergodic theorem:** Let $(\Omega, \mathcal{F}, P)$ be a prob. space and $\varphi: \Omega \to \Omega$ measure pres.. Let X be a real-valued RV with $\mathbb{E}|X| < \infty$.

Then $$\frac{1}{n}\sum_{k=0}^{n-1} X_k \xrightarrow[n\to\infty]{} \mathbb{E}(X \mid \mathcal{I})$$ a.s. and in L^1.

Examples: 1) If $(X_n)_{n \geq 0}$ is an IID seq. with $\mathbb{E}|X_0| < \infty$ then ← an IID seq. is ergodic
$$\frac{1}{n}\sum_{k=0}^{n-1} X_n \xrightarrow[n\to\infty]{} \mathbb{E}X_0$$ a.s. and in L^1
which reproves the law of large numbers.

2) Rotation of the circle: Let $\Omega = [0,1]$ with Borel sets and Lebesgue meas.
Let $\theta \in [0,1]$ be irrational.
Define $\varphi: \Omega \to \Omega$, $\varphi(\omega) = \omega + \theta \mod 1$.
This system is ergodic (exercise).
Consequently, for each Borel $A \subseteq [0,1]$,
$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_A(\underbrace{\varphi^k(\omega)}_{\omega + k\theta \bmod 1}) \xrightarrow[n\to\infty]{} \mathrm{Leb}(A) \text{ for a.e. } \omega \in \Omega.$$



Extension: When A is an interval then conv. occurs for all ω (no exceptional set).
Not difficult to prove by noting that a small shift of the starting point ω behaves the same with respect to the shifted A as the orig. ω behaved to A. But the shifted A is almost the same as A when A is an interval.

shifted A is almost the same as A when A is an interval.

Benford's law: What is the frequency of each digit as the most significant digit of the seq. $(2^n)_{n\geq 0}$?

Note that if the leading digit of N is $d\in\{1,\dots,9\}$ then it means that $N=d\cdot 10^k+r$ where $0\leq r<10^k$.

In part., write $2^n=d\cdot 10^k+r$
and then $2^{n+1}=2d\cdot 10^k+r=d'\cdot 10^{k'+1}+r'$

Take $\log_{10}$ to observe that

$$k+\log_{10} d=\log_{10}(d\cdot 10^k)\leq \log_{10}(N)=\log_{10}(d\cdot 10^k+r)<k+\log_{10}(d+1).$$

So d is the leading digit iff $\log_{10} N \bmod 1\in[\log_{10} d,\log_{10}(d+1))$.

Also note that $\log_{10}(2N)=\log_{10} N+\log_{10} 2$.

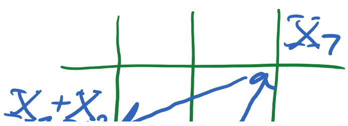
We conclude that the freq. of d as the leading digit of $(2^n)_{n\geq 0}$ is equal to

$W=0,\ \theta=\log_{10} 2$: $$\frac{1}{n}\sum_{k=0}^{n-1} 1_{[\log_{10} d,\log_{10}(d+1))}(k\log_{10} 2 \bmod 1)\xrightarrow[n\to\infty]{}\log_{10}\left(1+\frac{1}{d}\right)$$
for $d\in\{1,\dots,9\}$

$d=1: 0.3\dots,\ d=2: 0.17\dots,\ \dots$ ← Benford's law

Real-life application: Surprising collections of numbers satisfy roughly this statistic for the leading digit. E.g., population size, area of rivers, street addresses.
Used for detecting tax fraud.

Application to recurrence:
Let $X_1, X_2, \dots$ be a stationary seq. in $\mathbb{R}^d$.
Let $S_k := X_1+\dots+X_k$.



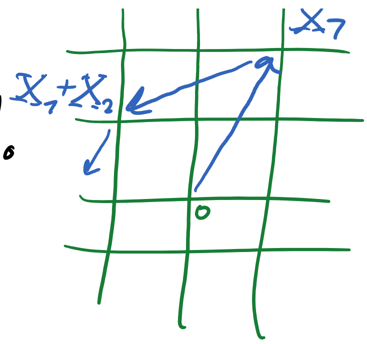
Let $S_k := X_1 + \_\_ + X_k$.

Let $R_n := |\{S_1, S_2, \_\_, S_n\}|$.

What is $\frac{1}{n} R_n$ in the limit?

Let $A = \{S_k \neq 0$ for all $k \geq 1\}$

(never return to 0)



$\mathbb{Z}^2$

Thm. (IID case is by Kesten-Spitzer-Whitman 1964):

$$\lim_{n\to\infty} \frac{1}{n} R_n = P(A \mid \mathcal{I}) \quad \text{a.s.}$$

Proof: Let $(\Omega, \mathcal{F}, P)$ be the seq. space for the stationary (X_n) and φ be the shift.

observe that $R_n(\omega) \geq \sum_{k=1}^{n} 1_A(\varphi^k \omega)$

Since $1_A \circ \varphi^k = \{\omega : S_m(\omega) \neq S_k(\omega)$ for all $m > k\}$.

By Birkhoff's thm.:

$$\frac{1}{n} R_n(\omega) \geq \frac{1}{n} \sum_{k=1}^{n} 1_A(\varphi^k \omega) \longrightarrow P(A \mid \mathcal{I}).$$

Fix $r \geq 1$ integer. Define

$$A_r = \{S_k \neq 0 \text{ for all } 1 \leq k \leq r\}.$$

Note that

$$R_n(\omega) \leq r + \sum_{k=1}^{n-r} 1_{A_r}(\varphi^k \omega)$$

since $1_{A_r} \circ \varphi^k = \{\omega : S_m(\omega) \neq S_k(\omega)$ for $k \leq m \leq k+r\}$.

By Birkhoff's thm.,

$$\frac{1}{n} R_n(\omega) \leq \frac{r}{n} + \frac{1}{n} \sum_{k=1}^{n-r} 1_{A_k}(\varphi^k \omega) \xrightarrow[n\to\infty]{} P(A_r \mid \mathcal{I})$$

It is clear that $1_{A_r}(\omega) \downarrow 1_A(\omega)$ a.s.

Whence $P(A_r \mid \mathcal{I}) \longrightarrow P(A \mid \mathcal{I})$ by bdd. conv. thm..

whence $\mathbb{P}(A_r|\mathcal{I}) \to \mathbb{P}(A|\mathcal{I})$ by bdd. conv. thm.
The thm. follows.

Corollary: Let $X_1, X_2, \ldots$ be stationary and take values in $\mathbb{Z}$. Suppose that $\mathbb{E}|X_1| < \infty$.
Let $S_n = X_1 + \ldots + X_n$.
Let $A = \{S_k \neq 0 \text{ for all } k \geq 1\}$.
i) If $\mathbb{E}(X_1|\mathcal{I}) = 0$ then $\mathbb{P}(A) = 0$ (recurrence)
ii) If $\mathbb{P}(A) = 0$ then $\mathbb{P}(S_n = 0 \text{ i.o.})$.
We won't prove this but it is not difficult (see Durrett's book).

The subadditive ergodic theorem

In Birkhoff's thm. we had a stationary seq. $(\xi_n)_{n \geq 1}$ and we considered the averages $\frac{1}{n}\sum_{k=1}^{n} \xi_k$. Here we consider a triangular array of RVs $(X_{m,n})_{0 \leq m < n}$

$$\begin{array}{lll} X_{0,1} & & \\ X_{0,2}, & X_{1,2} & \\ X_{0,3}, & X_{1,3}, & X_{2,3} \end{array}$$

which, to gain intuition, one can think of as

$$X_{m,n} := S_n - S_m = \sum_{k=m+1}^{n} \xi_k.$$

With this choice, $X_{0,n} = X_{0,m} + X_{m,n}$.
Recall also a useful elementary result:

Fekete's subadditive lemma: Let $(a_n)_{n \geq 1}$ be a seq. of real numbers.
Suppose it is subadditive in the sense
for $1 \leq m < n$.

suppose it is subadditive in the sense that $a_n \le a_m + a_{n-m}$ for $1 \le m < n$.

Then $\lim_{n\to\infty} \frac{1}{n} a_n = \inf_n \frac{1}{n} a_n \in [-\infty, \infty)$.

This has an elementary (calculus) proof.

The theorem (Liggett's version (1985), improved from Kingman's version (1968)):

Suppose $(X_{m,n})$, $0 \le m < n$, satisfy

i) $X_{0,n} \le X_{0,m} + X_{m,n}$ for $0 \le m < n$. (subadditivity)

ii) For each $i \ge 1$,

$(X_{0,i}, X_{i,2i}, X_{2i,3i}, \_\_\_)$ is stationary.

(generalized diagonals of the triangular array are stationary)

iii) For each $m \ge 1$,

$(X_{0,1}, X_{0,2}, X_{0,3}, \_\_) \stackrel{d}{=} (X_{m,m+1}, X_{m,m+2}, X_{m,m+3}, \_\_)$.

(columns of the triangular array have the same dist.)

iv) $\mathbb{E}\underbrace{X_{0,1}^+}_{\max\{X_{0,1}, 0\}} < \infty$ and for each n, $\mathbb{E}X_{0,n} \ge \gamma_0 n$ for some $\gamma_0 > -\infty$.

(moment condition)

Then (a) $\lim_{n\to\infty} \frac{\mathbb{E}X_{0,n}}{n} = \inf_n \frac{\mathbb{E}X_{0,n}}{n}$ exists. Denote it by γ.

(b) $X := \lim_{n\to\infty} \frac{1}{n} X_{0,n}$ exists a.s. and in L^1. In particular, $\mathbb{E}X = \gamma$.

(c) If the stationary seq. in (ii) are ergodic then $X = \gamma$ a.s.

Examples: 1) Birkhoff's thm. is a special case, when $X_{m,n} = \sum_{k=m+1}^{n} \xi_k$ with $(\xi_n)_{n\ge 1}$ stationary.

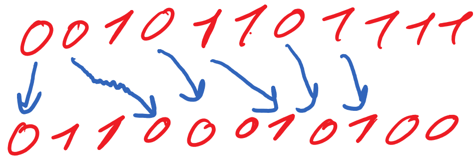
2) Longest common subsequence:

2) Longest common subsequence:

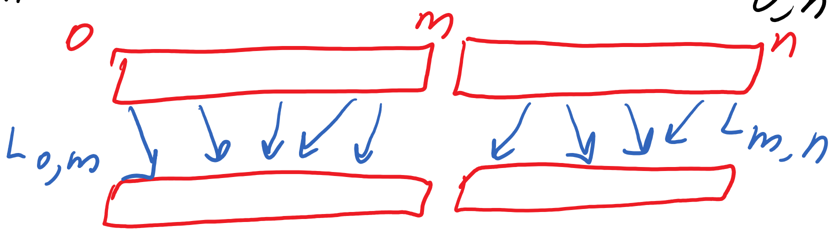
Let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be stationary seq., which are ergodic.

stationary

$L_{m,n} := \max\{K :$ there exist $m < i_1 < i_2 < \ldots < i_K \le n$, $m < j_1 < j_2 < \ldots < j_K \le n$ with $X_{j_k} = Y_{i_k}$ for all $1 \le k \le K\}$.



This is super-additive: $L_{0,n} \ge L_{0,m} + L_{m,n}$.



(thus $(-L_{m,n})$ is subadditive).

One checks prop. (ii), (iii) and notices that $0 \le L_{0,n} \le n$ so (iv) is satisfied.

Conclude that $\frac{1}{n} L_{0,n} \xrightarrow[n\to\infty]{} \gamma = \sup_m \mathbb{E}\left(\frac{1}{m} L_{0,m}\right)$.

Remark: Consider the IID case, where (X_n) and the (Y_n) are indep. and uniform on $\{0, 1, \ldots, k-1\}$.

The limiting const. γ_k are called Chvatal–Sankoff constants.

Conj.: $Var(L_{0,n})$ is of order n.

Known: $Var(L_{0,n}) \le C_k \cdot n$. (e.g., Azuma–Hoeffding's ineq.)

Not known that $Var(L_{0,n}) \xrightarrow[n\to\infty]{} \infty$.

see papers of Bukh et al.

Luecker (2009): $0.788\ldots \le \gamma_2 \le 0.826\ldots$

Kiwi–Loebl–Matoušek (2005): $\gamma_k = \frac{2}{\sqrt{k}}(1+o(1))$

as $k\to\infty$.

Proof of theorem: step 1: (convergence of the expectations)

We first check that $|\mathbb{E}X_{0,n}| \le C\cdot n$ for some $C<\infty$.

By (iv), $\mathbb{E}X_{0,n} \ge \gamma_0 \cdot n$ for some $\gamma_0 > -\infty$.

By (i), $X_{0,n} \le X_{0,m} + X_{m,n}$

$$\Rightarrow X_{0,n}^+ \le X_{0,m}^+ + X_{m,n}^+$$

Using this and (iii), $\mathbb{E}(X_{0,n}^+) \le n\,\mathbb{E}(X_{0,1}^+) \overset{\text{by (iv)}}{\le} C\cdot n$

Define the seq. of expectations, $a_n := \mathbb{E}(X_{0,n})$.

By (i) and (ii), $a_n \le a_m + a_{n-m}$.

Fekete's subadditive lemma now gives that

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}X_{0,n} = \inf_n \frac{1}{n}\mathbb{E}X_{0,n} =: \gamma.$$

This proves clause (a) in the thm.

Step 2 (limit-superior): Here we control $\limsup_{n\to\infty} \frac{X_{0,n}}{n}$.

Fix $m\ge 1$ integer. Write $n=km+\ell$, $0\le \ell<m$.

Note that by subadditivity, remainder term

$$X_{0,n} \le X_{0,km} + X_{km,n} \le$$

$$\le X_{0,(k-1)m} + X_{(k-1)m,km} + X_{km,n} \le \dots$$

$$\le X_{0,m} + X_{m,2m} + \dots + X_{(k-1)m,km} + X_{km,n}.$$

These form a stationary seq. by (ii).

Note now that

$\frac{1}{\ }$ ... k ... $\frac{1}{\ }\,{}^{k-1}$ ... $\frac{1}{\ }X$

Note now that

$$\frac{1}{n} X_{0,n} = \underbrace{\frac{k}{km+\ell}}_{\xrightarrow[k\to\infty]{} \frac{1}{m}} \cdot \underbrace{\frac{1}{k}\sum_{r=0}^{k-1} X_{rm,(r+1)m}}_{\text{Birkhoff applies}} + \frac{1}{n} X_{km,n}.$$

By Birkhoff, $\frac{1}{k}\sum_{r=0}^{k-1} X_{rm,(r+1)m} \xrightarrow[k\to\infty]{} A_m$ a.s. and in L^1

where $A_m := \mathbb{E}(X_{0,m} / \mathcal{I}_m)$ (where $\mathcal{I}_m$ is the inv. σ-alg. of the stat. seq. $(X_{rm,(r+1)m})_{r\geq 0}$)

In particular, $\mathbb{E} A_m = \mathbb{E} X_{0,m}$.

To take care of the remainder term, i.e., show that $\limsup_{n\to\infty} \frac{1}{n} X_{km,n} \leq 0$, a.s., note that, for each $\varepsilon > 0$,

$$P(X_{km,n} > n\varepsilon) \underset{\substack{\uparrow \\ n = km+\ell, \\ \text{and use (iii)}}}{=} P(X_{0,\ell} > n\varepsilon) \leq P(X_{0,\ell} > k\varepsilon)$$

moreover, $\sum_{k=1}^{\infty} P(X_{km,n} > n\varepsilon) \leq \sum_{k=1}^{\infty} P(X_{0,\ell} > k\varepsilon)$

$$= \sum_{k=1}^{\infty} P\left(\frac{X_{0,\ell}^{+}}{\varepsilon} > k\right) < \infty$$

since $\mathbb{E}(X_{0,\ell}^{+}) < \infty$ (If Y is non-negative then $\mathbb{E} Y < \infty \iff \sum_{k=1}^{\infty} P(Y > k) < \infty$).

By Borel-Cantelli, $\underbrace{\text{only}}^{\text{for each fixed } \ell}$ finitely many of the events $\{X_{km,km+\ell} > (km+\ell)\varepsilon\}_{k\geq 1}$ occur. Thus $\limsup_{n\to\infty} \frac{1}{n} X_{km,n} \leq 0$.

Putting everything together,

$$\limsup_{n\to\infty} \frac{1}{n} X_{0,n} \leq \frac{1}{m} A_m \quad \text{a.s.}.$$

$$\limsup_{n\to\infty} \frac{1}{n} X_{0,n} \le \frac{1}{m} A_m \quad \text{a.s.}.$$

Define $\overline{X} := \limsup_{n\to\infty} \frac{1}{n} X_{0,n}$.

Thus $\mathbb{E}\overline{X} \le \frac{1}{m} \mathbb{E} A_m = \frac{1}{m} \mathbb{E}(X_{0,m})$, for each $m \ge 1$.

Taking infimum over m, $\mathbb{E}\overline{X} \le \gamma$.

(If all the seq. in (ii) are ergodic then this would give $\overline{X} \le \gamma$ a.s.).