Stationarity and ergodicity

Thursday, 8 April 2021 13:06

A sequence $(X_n)_{n\geq 0}$ of random variables (taking values in some measurable space) is called <u>stationary</u> if, for every $k\geq 1$,

$$(X_0, X_1, X_2, \ldots) \overset{d}{=} (X_k, X_{k+1}, X_{k+2}, \ldots)$$

(this is the same as requiring that for every $n\geq 0$ and $k\geq 1$,

$$(X_0, \ldots, X_n) \overset{d}{=} (X_k, \ldots, X_{k+n})$$
)

<u>Examples:</u> 1) (X_n) are IID.

2) Markov chain in the stationary distribution. To avoid technicalities, focus on the finite-state space case. S - state space, P is the transition probabilities matrix. Let π be a stationary distribution. Take $X\sim\pi$, $(X_0, X_1, \ldots)$ be a sample of the Markov chain.

$$P(X_0 = x_0, \ldots, X_n = x_n) = \pi(x_0)P(x_0, x_1)\ldots P(x_{n-1}, x_n)$$

Since π is stationary, the same holds for $(X_k, \ldots, X_{k+n})$.

3) If (X_n) is stationary, taking values in some Borel S, then for any measurable function $f: S^{\{0,1,\ldots\}} \to S'$ it holds that $(X_0', X_1', \ldots)$ is stationary, where

$$X_n' := f(X_n, X_{n+1}, \ldots).$$

<u>Measure-preserving maps:</u> Suppose $(\Omega, \mathcal{F}, P)$

Measure-Preserving maps: Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. Suppose $\varphi : \Omega \to \Omega$ is <u>measure preserving</u> if φ is measurable and $\forall E \in \mathcal{F}$, $\mathbb{P}(\varphi^{-1}E) = \mathbb{P}(E)$.

It is equivalent that for every random variable X on $(\Omega, \mathcal{F}, \mathbb{P})$, $X \overset{d}{=} X'$ where $X'(\omega) = X(\varphi(\omega))$.

Now, given a RV X, can define

$$X_n(\omega) := X(\varphi^n(\omega)), \quad n \geq 0$$

(so that $X_0(\omega) = X(\omega)$)

and then $(X_0, X_1, \ldots)$ is stationary.

proof: Suppose X takes values in S. Let $n \geq 0$, and A a measurable set in S^{n+1}.

Then, $\forall k \geq 1$,

$$\mathbb{P}((X_k, X_{k+1}, \ldots, X_{k+n}) \in A) =$$

$$= \mathbb{P}(\omega : (X(\varphi^k \omega), \ldots, X(\varphi^{k+n}(\omega))) \in A) =$$

$$= \mathbb{P}(\omega : (X(\omega), \ldots, X(\varphi^n(\omega)) \in \varphi^{-k}A) =$$

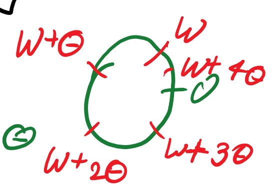
$$= \mathbb{P}(\varphi^{-k}A) = \mathbb{P}(A)$$

since φ is measure preserving.

Example: (rotation of the circle)

Let $\Omega = [0,1]$, $\mathcal{F}$ - Borel sets, $\mathbb{P}$ - Lebesgue measure.

Fix $\theta \in (0,1)$ and let $\varphi : [0,1] \to [0,1]$ be $\varphi(\omega) = \underbrace{(\omega + \theta) \bmod 1}_{\text{Fractional part of } \omega + \theta}$



It is simple that φ is measure preserving.
Define $X(\omega) = \omega$. $X_n(\omega) := \omega + n\theta \mod 1$.
So $(X_0, X_1, \text{—})$ is stationary.

Every stationary sequence can be realized by a meas.-pres. map:

Let (Y_n) be stationary, taking values in a Borel space S. We will define $(\Omega, \mathcal{F}, \mathbb{P})$ and φ, X so that $X_n = X \circ \varphi^n$ satisfies that $(Y_0, \text{—}) \overset{d}{=} (X_0, \text{—})$.

Indeed, let $\Omega = S^{\{0,1,\text{—}\}}$. $\mathcal{F}$ = product sigma-algebra. Let $\mathbb{P}$ be the distribution of $(Y_0, Y_1, \text{—})$.
Let $X(\omega_0, \omega_1, \text{—}) = \omega_0$.
$\varphi(\omega_0, \omega_1, \text{—}) = (\omega_1, \omega_2, \text{—})$ the left shift.

So that $X_n(\omega) = \omega_n$.
φ is measure prerserving since, under $\mathbb{P}$,
$$(\omega_0, \omega_1, \text{—}) \overset{d}{=} (Y_0, Y_1, \text{—}) \overset{d}{=}$$
$$\overset{d}{=} (Y_1, Y_2, \text{—}) \overset{d}{=} (\omega_1, \omega_2, \text{—}).$$

It also follows that
$$(X_0, \text{—}, X_n) \overset{d}{=} (Y_0, \text{—}, Y_n).$$

Invariant events and ergodicity

From now on, fix $(\Omega, \mathcal{F}, \mathbb{P})$ a prob. space, $\varphi: \Omega \to \Omega$ meas.-pres. and X a RV.
Then let $X_n = X \circ \varphi^n$. $(\Omega, \mathcal{F}, \mathbb{P}, \varphi)$ – Dynamical system

An event $I \in \mathcal{F}$ is called invariant
as $\varphi^{-1} I = I$

An event $I \in \mathcal{F}$ is called <u>invariant</u> if $\varphi^{-1} I = I$.

An event $I \in \mathcal{F}$ is called <u>almost-invariant</u> if $\mathbb{P}(\varphi^{-1} I \,\Delta\, I) = 0$ $\quad A \Delta B = (A \setminus B) \cup (B \setminus A)$.

<u>Exercise:</u> 1) The invariant events form a sigma algebra $\mathcal{I}$. The same for the almost invariant events.

2) A RV Y is measurable with respect to $\mathcal{I}$ iff $Y(\omega) = Y(\varphi(\omega))$ $\forall \omega \in \Omega$.

Similarly, it is meas. w.r.t. the almost-inv. σ-alg. if $\mathbb{P}(Y(\omega) \neq Y(\varphi(\omega))) = 0$.

3) If I' is an almost-invariant event then there exists an invariant I such that $\mathbb{P}(I \,\Delta\, I') = 0$.

The sequence (X_n) or $(\Omega, \mathcal{F}, \mathbb{P})$ and φ is called <u>ergodic</u> if $\mathbb{P}(I) \in \{0,1\}$ for all $I \in \mathcal{I}$.

<u>Examples:</u> 1) (X_n) are IID. Realize (X_n) on sequence space $(\omega_0, \omega_1, \ldots)$ with product measure with the shift φ. Then this construction is <u>ergodic</u>. (i.e., random variables which are functions of the (X_n) and are invariant to shifts of the (X_n) must be constant almost surely).

Indeed, suppose $f(X_0, X_1, \ldots)$ is invariant, i.e., $f(X_0, X_1, \ldots) = f(X_1, X_2, \ldots)$.

It follows that $f(X_0, \ldots) = f(X_n, \ldots)$, $\forall n \geq 0$.

Defining $\mathcal{G}_n := \sigma(X_n, X_{n+1}, \ldots)$ we see that $f(X_0, \ldots) \in \mathcal{G}_n$ for all $n \geq 0$ and

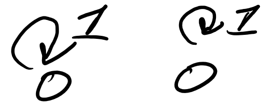
that $f(X_0, \ldots) \in \mathcal{G}_n$ for all $n \geq 0$ and hence meas. wrt. the tail σ-alg. $\mathcal{T} := \bigcap_n \mathcal{G}_n$.
Kolmogorov's 0-1 law then implies that f is constant a.s..

2) <u>Markov Chain at Stationarity:</u>

S-finite state space, P-transition prob. matrix.
π-stationary dist., $X_0 \sim \pi$,

$$(X_0, X_1, \ldots) \text{ sample of the Markov chain.}$$

Realize on seq. space with φ = shift.
Without assuming irreducibility, this might not be ergodic. E.g., $S = \{0, 1\}$, $P(0,0) = 1$, $P(1,1) = 1$
and $\pi = \left(\frac{1}{2}, \frac{1}{2}\right)$.



Then $f(X_0, X_1, \ldots) = 1_{X_0 = 0}$
is invariant but $P(f=1) = P(f=0) = 1/2$.

Assume P is irreducibility and let us show that (X_n) is ergodic.
Suppose A is an invariant event. I.e.,

$$1_A(X_0, X_1, \ldots) = 1_A(X_1, X_2, \ldots).$$

We need to show $P(A) \in \{0, 1\}$.
Define $h(x) = P_x(A)$. Start the Markov chain at x.
By the Markov prop.,

$$P_\pi(A \mid \mathcal{F}_n) = \mathbb{E}(1_A(X_0, \ldots) \mid \mathcal{F}_n) \overset{\text{invariance}}{=}$$

$$= \mathbb{E}(1_A(X_n, X_{n+1}, \ldots) \mid \mathcal{F}_n) \overset{\text{Markov prop.}}{=}$$

$\mathcal{F}_n := \sigma(X_0, \ldots, X_n)$

$\mathcal{F}_n := \sigma(X_0, \ldots, X_n)$ $= \mathbb{E}(1_A(X_n, X_{n+1}, \ldots) \mid \mathcal{F}_n) =$

$= \mathbb{E}_{X_n}(1_A(X_n, \ldots)) = h(X_n).$

By Lévy's 0-1 law (upward martingale conv. thm.)

$$\mathbb{P}_\pi(A \mid \mathcal{F}_n) \xrightarrow{n\to\infty} 1_A, \text{ almost surely.}$$

Since S is finite and the chain is irreducible, $\forall x \in S$, $\mathbb{P}(X_n = x$ infinitely often$) = 1$.

It follows from the fact that $h(X_n)$ converges a.s. that h is constant. Moreover, the limit 1_A must also be constant, almost surely. So $\mathbb{P}(A) \in \{0,1\}$.

3) Rotation of the circle is ergodic when θ (the rotation amount) is irrational - exercise.

Remark: Suppose P is the transition prob. matrix of an irreducible Markov chain. As we saw, the invariant σ-alg. is trivial. What is the tail σ-alg.?

$A \in \mathcal{T}$ if for every n there exists f_n such that $1_A(X_0, X_1, \ldots) = f_n(X_n, X_{n+1}, \ldots)$.

For instance, if $S = \{0,1\}$, $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ $0 \underset{1}{\overset{1}{\rightleftarrows}} 1$

Then $A = 1_{\{X_0 = 0\}} \in \mathcal{T}$. $\pi = (\frac{1}{2}, \frac{1}{2})$

$\mathbb{P}_\pi(A) = 1/2.$

More generally, it turns out that

$\mathcal{T} = \sigma(1_{\text{periodic classes}})$. (up to measure 0).

Note also that in the example,
$(X_0, X_1, \text{---})$ is ergodic
but $(X_0, X_2, \text{---})$ is not ergodic.
(so $\varphi = \text{shift}$ is ergodic but φ^2 is not ergodic).

Birkhoff ergodic theorem (1931)

Still work in a measure-pres. setup.
$(\Omega, \mathcal{F}, \mathbb{P})$ - prob. space
$\varphi: \Omega \to \Omega$ - meas. preserving.

Theorem: For each RV $X: \Omega \to \mathbb{R}$ with $\mathbb{E}|X| < \infty$,
$$\frac{1}{n} S_n(\omega) := \frac{1}{n} \sum_{k=0}^{n-1} X(\varphi^k \omega) \xrightarrow[n\to\infty]{} \mathbb{E}(X|\mathcal{I})$$
almost surely and in L_1.

Example: 1) In the IID setup, $X_n = X(\varphi^n \omega)$ are IID.
then $\frac{1}{n} \sum_{k=0}^{n-1} X_k \xrightarrow[a.s.]{n\to\infty} \mathbb{E}(X|\mathcal{I}) \underset{\text{ergodic}}{=} \mathbb{E}(X) = \mathbb{E}(X_0)$
is the strong law of large numbers (and got also the L_1 conv.)

2) Taking $X = 1_A$ in the ergodic theorem, we get that in ergodic systems, the "fraction of time" spent in A equals $\mathbb{P}(A)$ = "part of space that A takes".

Physics: space average = time average.

Start the proof with a lemma due to

Start the proof with a lemma due to Yosida-Kakutani (1939); we show a proof of Garsia (1965).

Lemma (Maximal ergodic lemma):

Define $X_n(\omega) := X(\varphi^n \omega)$, $(X_0 = X)$

$S_n(\omega) := X_0(\omega) + \ldots + X_{n-1}(\omega)$,

$M_n(\omega) := \max(0, S_1(\omega), \ldots, S_n(\omega))$.

$X_0(\omega)$ $\quad$ $X_0(\omega) + \ldots + X_{n-1}(\omega)$

Then $\mathbb{E}(X \mathbb{1}_{M_k > 0}) \geq 0$.

Proof: The idea is to relate $M_k(\omega)$ and $M_k(\varphi(\omega))$.

Observe that for all $1 \leq j \leq k$,

$$M_k(\varphi\omega) \geq S_j(\varphi\omega)$$

Thus $X(\omega) + M_k(\varphi\omega) \geq X(\omega) + S_j(\varphi\omega) = S_{j+1}(\omega)$.

Rearranging, $X(\omega) \geq S_{j+1}(\omega) - M_k(\varphi\omega)$ for $1 \leq j \leq k$.

This holds trivially for $j = 0$ since $S_1 = X$ and $M_k(\varphi\omega) \geq 0$.

Therefore

$$\mathbb{E}(X \mathbb{1}_{M_k > 0}) = \int_{\{M_k > 0\}} X(\omega)\, dP(\omega) \geq$$

$$\geq \int_{\{M_k > 0\}} [\max\{S_1(\omega), \ldots, S_k(\omega)\} - M_k(\varphi\omega)]\, dP(\omega) =$$

$$= \int_{\{M_k > 0\}} M_k(\omega) - M_k(\varphi\omega)\, dP(\omega) \geq$$

on $\{M_k > 0\}^c = \{M_k = 0\}$ we have $M_k = 0$, $M_k \circ \varphi \geq 0$.

$$\geq \int M_k(\omega) - M_k(\varphi\omega)\, d\mathbb{P}(\omega) = \mathbb{E}(M_k - M_k \circ \varphi) = 0$$

since φ is measure-preserving

Proof of the ergodic theorem:

Observe that $\mathbb{E}(X/\mathcal{I})$ is an invariant RV, i.e., $\mathbb{E}(X/\mathcal{I})(\omega) = \mathbb{E}(X/\mathcal{I})(\varphi\omega)$.

Replace X by $X - \mathbb{E}(X/\mathcal{I})$ and then we may assume without loss of generality that $\mathbb{E}(X/\mathcal{I}) = 0$ a.s.

Let $\overline{X} = \limsup_{n\to\infty} \frac{1}{n} S_n$. Let $\varepsilon > 0$ and

$D := \{\overline{X} > \varepsilon\}$. $\leftarrow$ $\overline{X}$ is invariant and hence $D \in \mathcal{I}$.

Goal: $\mathbb{P}(D) = 0$. Then, as ε is arbitrary, get $\overline{X} \leq 0$ a.s. Applying the same reasoning to $-X$ gives $\liminf_{n\to\infty} \frac{S_n}{n} \geq 0$ a.s. and then $\lim_{n\to\infty} \frac{S_n}{n} = 0$ a.s., proving the a.s. convergence.

Define a new seq. of RVs,

$$X^*(\omega) := (X(\omega) - \varepsilon)\cdot 1_D(\omega)$$

$$X_n^*(\omega) := X^*(\varphi^n \omega) \overset{D\in\mathcal{I}}{=} (X_n(\omega) - \varepsilon) 1_D(\omega).$$

$$S_n^*(\omega) := X_0^*(\omega) + \_\_\_ + X_{n-1}^*(\omega)$$

$$M_n^*(\omega) := \max\{0, S_1^*(\omega), \_\_\_, S_n^*(\omega)\}.$$

$$F_n^*(\omega) := \{M_n^* > 0\}.$$ $\leftarrow$ increasing sequence of events

Consider $F = \bigcup_n F_n = \{\sup_{k\geq 1} S_k^*(\omega) > 0\}$

Notice that $\{\sup_{k\geq 1} \frac{S_k^*(\omega)}{k} > 0\} = D \cap \{\sup_{k\geq 1} \frac{S_k(\omega)}{k} > \varepsilon\}$

Notice that $\left\{\sup_{k\geq 1} \frac{S_k^*(\omega)}{k} > 0\right\} = D \cap \left\{\sup_{k\geq 1} \frac{S_k(\omega)}{k} > \varepsilon\right\}$

Since $\frac{S_k^*(\omega)}{k} = 1_D\left(\frac{S_k(\omega)}{k} - \varepsilon\right)$. def. of D.

Thus $F = D \cap \left\{\sup_{k\geq 1} \frac{S_k(\omega)}{k} > \varepsilon\right\} = D$.

The maximal ergodic lemma says that

$$\mathbb{E}(X^* 1_{F_n}) \geq 0 \quad \forall n \geq 1.$$

This implies that $\mathbb{E}(X^* 1_F) \geq 0$

since $1_{F_n} \to 1_F$ (since F_n increases)

and $|X^*| \leq |X| + \varepsilon$ so we can

use dominated convergence.

Thus $0 \leq \mathbb{E}(X^* 1_F) \overset{F=D}{=} \mathbb{E}(X^* 1_D) =$

$= \mathbb{E}((X-\varepsilon)1_D) = \underbrace{\mathbb{E}(X 1_D)} - \varepsilon \mathbb{P}(D) = -\varepsilon \mathbb{P}(D)$

$= \mathbb{E}(\mathbb{E}(X 1_D | \mathcal{I})) \underset{D \in \mathcal{I}}{=} \mathbb{E}(1_D \underbrace{\mathbb{E}(X|\mathcal{I})}_{=0}) = 0$

We conclude that $\mathbb{P}(D) = 0$, finishing the proof of the a.s. conv.

It remains to prove conv. in L_1.

The idea is to truncate, and this is relatively routine.

For $M > 0$,

write $X = \underbrace{X 1_{|X| \leq M}}_{=: X'} + \underbrace{X 1_{|X| > M}}_{=: X''}$.

Define (X_n') and (X_n'') using these.

By the a.s. conv.

$$\frac{1}{n} S_n' \to \mathbb{E}(X' | \mathcal{I}) \text{ a.s.}$$

$$\underbrace{\frac{1}{n}S_n'}_{\text{bounded by } M} \to \mathbb{E}(X'|\mathcal{I}) \text{ a.s.}$$

By the bdd. conv. theorem,

$$\frac{1}{n}S_n' \to \mathbb{E}(X'|\mathcal{I}) \text{ in } L_1.$$

Next, note that

$$\mathbb{E}\left|\frac{1}{n}S_n''\right| = \mathbb{E}\left|\frac{1}{n}\sum_{k=0}^{n-1} X''(\varphi^k\omega)\right| \le$$

$$\le \frac{1}{n}\sum_{k=0}^{n-1}\mathbb{E}|X''(\varphi^k\omega)| = \mathbb{E}|X''|.$$

φ is meas. pres.

Additionally, $\mathbb{E}|\mathbb{E}(X''|\mathcal{I})| \le \mathbb{E}(\mathbb{E}(|X''||\mathcal{I})) = \mathbb{E}|X''|$.

Putting the last ineq. together,

$$\mathbb{E}\left|\frac{1}{n}S_n'' - \mathbb{E}(X''|\mathcal{I})\right| \le 2\mathbb{E}|X''|.$$

It remains to note that $\mathbb{E}|X''| \xrightarrow[M\to\infty]{} 0$ by the dominated conv. theorem.

Thus $\mathbb{E}\left|\frac{1}{n}S_n - \mathbb{E}(X|\mathcal{I})\right| \xrightarrow[n\to\infty]{} 0.$