1 Mathematical proofs

1.1 Birkhoff's Ergodic Theorem and the LLN

Definition. $T : (\Omega, \mathcal{A}, \mathbb{P}) \to (\Omega, \mathcal{A}, \mathbb{P})$ measure preserving *(maßtreu) if*

$$\mathbb{P}(T^{-1}A) = \mathbb{P}(A) \;\; \forall A \in \mathcal{A}.$$

This statement is equivalent to

$$\mathbb{E}(f \circ T) = \mathbb{E}(f) \;\; f \in L^1(\Omega, \mathcal{A}, \mathbb{P})$$

T is called ergodic *if* $T^{-1}A = A \implies \mathbb{P}(A) \in \{0, 1\}$ *for* $A \in \mathcal{A}$.

$$\mathcal{I} = \{A \in \mathcal{A} : T^{-1}A = A \textit{ invariant } \sigma\textit{-algebra}\}$$

Remark. *For this definition there is no significant difference if we replace* $T^{-1}A = A$ *by* $T^{-1}A \overset{a.s.}{=} A$.

Theorem 1.1 (Birkhoff's Ergodic Theorem)**.** *T measure preserving, $f \in L^1$, $S_n(f) = \sum_{k=0}^{n-1} f \circ T^k$. Then there exists $M(f) \in L^1(\Omega, \mathcal{I}, \mathbb{P})$ such that $\frac{1}{n} S_n(f) \to M(f)$ almost surely and*

$$\int_A f d\mathbb{P} = \int_A M(f) d\mathbb{P} \;\; \forall A \in \mathcal{I}$$

that is $M(f) = \mathbb{E}(f|_{\mathcal{I}})$. If T is ergodic, then $\mathcal{I}$ is trivial and $\mathbb{E}(f|_{\mathcal{I}}) = \mathbb{E}(f)$.

Theorem 1.2 (von Neumann's Ergodic Theorem)**.** *T measure preserving, $f \in L^1$, $S_n(f) = \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k$. Then there exists $M(f) \in L^1(\Omega, \mathcal{I}, \mathbb{P})$ such that $S_n(f) \to M(f)$ almost surely and in L^1*

$$\int_A f d\mathbb{P} = \int_A M(f) d\mathbb{P} \;\; \forall A \in \mathcal{I}$$

that is $M(f) = \mathbb{E}(f|_{\mathcal{I}})$. If T is ergodic, then $\mathcal{I}$ is trivial and $\mathbb{E}(f|_{\mathcal{I}}) = \mathbb{E}(f)$.

How to deduce the Law of Large Numbers from Birkhoff's ergodic theorem? (X_n) iid, real RVs, $\mathbb{E}(|X_k|) < \infty$

$$S_n = X_1 + \cdots + X_n \overset{?}{\implies} \frac{1}{n} S_n \overset{\text{a.s.}}{\to} \mathbb{E}(X_1).$$

P_X: distribution of X_k. $(\mathbb{R}, \mathcal{B}, P_X)$, $X(\omega) = \omega_1$, $\omega_1 \in \mathbb{R}$.

$$(\Omega, \mathcal{A}, \mathbb{P}) = \otimes_{n=1}^{\infty} (\mathbb{R}, \mathcal{B}, P_X)$$
$$\Omega = \mathbb{R}^{\mathbb{N}} = \{\omega = (\omega_1, \omega_2, \dots) : \omega_k \in \mathbb{R}\}$$

$\mathcal{A}$: σ-algebra generated by all sets

$$C(I_1, \dots, I_n) = \{\omega : \omega_k \in I_k \forall k \leq n\}$$

where $I_1, \dots, I_n$ are intervals Take a σ-algebra created by all sets $I_1 \times I_2 \times \cdots \times I_n \times \mathbb{R} \times \mathbb{R} \times \dots$.

$$\mathbb{P}(C(I_1,\dots,I_n)) = P_X(I_1)\cdots P_X(I_n)$$

Recall: if $(\tilde\Omega,\tilde{\mathcal{A}},\tilde{\mathbb{P}})$ is any probability space and $\tilde{X_n}$ is any sequence of iid RVs with distribution P_X.

add pic

$\tau(\tilde\omega) = (\tilde{X_n}(\tilde\omega))_{n\in\mathbb{N}} = (\tilde{X_1}(\tilde\omega),\tilde{X_2}(\tilde\omega),\dots)$

$T(\omega_1,\omega_2,\dots) = (\omega_2,\omega_3,\dots)$ is measure preserving. $X_n(\omega_1,\omega_2,\dots) = \omega_n$ and $f = X_1 \in L^1$. Then we have $f\circ T^{k-1} = X_k$. Thus,

$$S_n(f) = \sum_{k=1}^{n} X_k$$

and so by Birkhoff's Ergodic Theorem,

$$\frac{1}{n} \overset{\text{a.s}}{\to} \mathbb{E}(X_1|_{\mathcal{I}}$$

Now we just need to justify why T is ergodic. 0-1-law of Kolmogorov:

Theorem 1.3. $(\Omega,\mathcal{A},\mathbb{P})$, (X_n) *independent RVs Then*

$$\underbrace{\tau(X_1,X_2,\dots)}_{\bigcap_{n=1}^{\infty}\sigma(X_n,X_{n+1},\dots)}$$

is trivial, i.e. $A\in\tau \implies \mathbb{P}(A)\in\{0,1\}$. τ *is called the tail* σ*-algebra.*

The argument is

$$\begin{aligned}
A\in\tau &\implies A\in\sigma(X_{n+1},X_{n+2},\dots)\ \ \forall n\\
&\implies A \text{ independent of } \sigma(X_1,\dots,X_n)\\
&\implies A \text{ independent of } \bigcup_{n=1}^{\infty}\sigma(X_1,\dots,X_n)\\
&\implies A \text{ independent of } \bigvee_{n=1}^{\infty}\sigma(X_1,\dots,X_n) = \sigma(X_1,X_2,\dots)\ni A
\end{aligned}$$

Since $T^{-1}A = A$, $\underbrace{T^{-n}A}_{\in\sigma(X_{n+1},Xn+2,\dots)} = A$. Thus, $\mathcal{A}\in\tau$.

1.2 Theorem: Axiom definition of entropy

Proof.

$$H(1,0) \overset{\text{axiom 4}}{=} H(1) + 1H(1,0) \implies H(1) = 0$$

$$\begin{aligned}
H(p_1,\dots,p_N,0) &\overset{\text{axiom 4}}{=} H(p_1,\dots,p_N) + p_N H(1,0)\\
H(p_1,1-p_1,0) &= H(p_1,1-p_1) + (1-p_1)H(1,0)\\
H(1-p_1,p_1,0) &= H(1-p_1,p_)) + p_1 H(1,0)
\end{aligned}$$

Since by axiom 1, $H(p_1, 1-p_1, 0) = H(1-p_1, p_1, 0)$ and since $H(p_1, 1-p_1) = H(1-p_1, p_1)$, it follows that

$$(1-p_1)H(1,0) = p_1 H(1,0) \ \ \forall p_1$$

Thus, $H(1,0) = 0$.

Claim 1. *For all N and for all $m \geq 2$:*

$$H(p_1, \ldots, p_N, p_{N+1}, \ldots, p_{N+m}) = H(p_1, \ldots, p_N, q) + qH\left(\frac{p_{N+1}}{q}, \ldots, \frac{p_{N+m}}{q}\right)$$

where $q = \sum_{k=N+1}^{N+m} p_k$.

Proof. Induction: $m = 2$ holds by axiom 4.
$2 \leq m-1 \to m$:

$$H(p_1, \ldots, p_N, p_{N+1}, \ldots, p_{N+m-1}, p_{N+m})$$

$$\overset{\text{axiom 4}}{=} H(p_1, \ldots, p_N, \ldots, p_{N+m-2}, p_{N+m-1}, p_{N+m}) + \underbrace{(p_{N+m-1} + p_{N+m})}_{q'} H\left(\underbrace{\frac{p_{N+m-1}}{q'}}_{p'}, \underbrace{\frac{p_{N+m}}{q'}}_{1-p'}\right)$$

$$\overset{\text{ind. hyp.}}{=} H(p_1, \ldots, p_N, q) + qH\left(\frac{p_{N+1}}{q}, \ldots, \frac{p_{N+m-2}}{2}, \frac{q'}{q}\right) + q'H(p', 1-p')$$

$$\overset{\text{axiom 4}}{=} H(p_1, \ldots, p_N, q) + q\left[H\left(\frac{p_{N+1}}{q}, \ldots, \frac{p_{N+m-2}}{q}, \frac{p_{N+m-1}}{q}, \frac{p_{N+m}}{q}\right) - \frac{q'}{q}H(p', 1-p')\right] + q'H(p', 1-p$$

$$= H(p_1, \ldots, p_N, q) + qH\left(\frac{p_{N+1}}{q}, \ldots, \frac{p_{N+m-2}}{q}, \frac{p_{N+m-1}}{q}, \frac{p_{N+m}}{q}\right)$$

Since $-q \cdot \frac{q'}{q} H(p', 1-p') + q'H(p', 1-p' = 0$. ☐

Claim 2.

$$pq'H(p', 1-p'(1) = (p_1^{(1)}, \ldots, p_{N_1}^{(1)}), \ldots, p^{(n)} = (p_1^{(n)}, \ldots, p_{N_n}^{(n)}) \in \mathcal{P}$$

$$q = (q_1, \ldots, q_n) \in \mathcal{P}$$

$$(q_1 p^{(1)}, \ldots, q_n p^{(n)}) = (q_1 p_1^{(1)}, \ldots, q_1 p_{N_1}^{(1)}, \ldots, q_n p_{N_n}^{(n)}$$

Size: $N_1 + \cdots + N_n$.

$$H(q_1 p^{(1)}, \ldots, q_n p^{(n)}) = H(q) + \sum_{j=1}^{n} a_j H(p^{(j)}$$

Proof. If $n = 1$ and $q = 1$ there is nothing to prove.

$$H(q_1 p^{(1)}, \ldots, q_{n-1} p^{(n-1)}, q_n p^{(n)})$$

$$\overset{\text{Claim 1}}{=} H(q_1 p^{(1)}, \ldots, q_{n-1} p^{(n-1)}, q_n) + q_n H(p^{(n)})$$

$$= H(q_n, q_1 p^{(1)}, \ldots, q_{n-2} p^{(n-2)}, q_{n-1}) + q_{n-1} H(p^{(n-1)}) + q_n H(p^{(n)}) = \cdots = \checkmark$$

☐

☐