Ergodic Theory
Solutions to Problem Sheet 11

Course Instructor: Florian K. Richter
Problems by: Jovan Andreevski

December 1, 2025

For any questions or corrections, please send an email to: jovan.andreevski@epfl.ch!

P1. In this exercise we verify several properties concerning entropy of partitions. To that end, let $(X, \mathscr{A}, \mu)$ be a probability space, and let $\xi = \{A_1, A_2, \dots\}$ and $\eta = \{B_1, B_2, \dots\}$ be a partition of our space.

(a) Prove that if ξ is a finite partition with r atoms, then $H(\xi) \leq \log r$, with equality if and only if $\mu(A) = \frac{1}{r}$ for any $A \in \xi$.

> Let $\xi = \{A_1, A_2, ..., A_r\}$. Consider the strictly convex function $\phi(x) = x \log\, x$. By Jensen inequality we have that
>
> $$\frac{-\log r}{r} = \phi\left(\frac{1}{r}\sum_{i=1}^{r}\mu(A_i)\right) \leq \frac{1}{r}\sum_{i=1}^{r}\phi(\mu(A_i)) = \frac{-H(\xi)}{r}.$$
>
> It follows that $H(\xi) \leq \log r$, where the equality holds if and only it holds for the Jensen equality, which happens if and only if all the summands are equal (by strict convexity).

(b) Prove that $H(\xi \vee \eta) = H(\eta) + H(\xi \mid \eta)$.

> Let $\xi = \{A_1, A_2, ...\}$ and $\eta = \{B_1, B_2, ...\}$. We have
>
> $$\begin{aligned} H(\xi \vee \eta) &= -\sum_{i,j}\mu(A_i \cap B_j)\log\mu(A_i \cap B_j) \\ &= -\sum_{i,j}\mu(A_i \cap B_j)\log\frac{\mu(A_i \cap B_j)}{\mu(B_j)} - \sum_{i,j}\mu(A_i \cap B_j)\log\mu(B_j) \\ &= -\sum_{j}\mu(B_j)\sum_{i}\mu(A_i \mid B_j)\log\mu(A_i \mid B_j) - \sum_{j}\mu(B_j)\log\mu(B_j) \\ &= H(\xi \mid \eta) + H(\eta). \end{aligned}$$

(c) Prove that $H(\xi) \geq H(\xi \mid \eta)$.

> Let $\xi = \{A_1, A_2, ...\}$ and $\eta = \{B_1, B_2, ...\}$. Consider the strictly convex function $\phi(x) =$

$x \log\ x$ and by Jensen's equality we have

$$
\begin{aligned}
H(\xi \mid \eta) &= -\sum_j \mu(B_j) \sum_i \mu(A_i \mid B_j) \log\ \mu(A_i \mid B_j) \\
&= -\sum_i \sum_j \mu(B_j)\phi(\mu(A_i \mid B_j)) \\
&\leq -\sum_i \phi\left(\sum_j \mu(B_j)\mu(A_i \mid B_j)\right) \\
&= -\sum_i \phi\left(\sum_j \mu(A_i \cap B_j)\right) \\
&= -\sum_i \phi(\mu(A_i)) \\
&= -\sum_i \mu(A_i) \log \mu(A_i) \\
&= H(\xi).
\end{aligned}
$$

(d) Prove that ξ and η are independent if and only if $H(\xi \mid \eta) = H(\xi)$.

By c), $H(\xi \mid \eta) = H(\xi)$ if and only if for all i,

$$\mu(A_i \mid B_1) = \mu(A_i \mid B_2) = \ldots \leq \mu(A_i).$$

Assume that there exists some i_0 such that the inequality above is strict. Then for any fixed j, we have

$$\sum_i \mu(A_i \mid B_j) < \sum_i \mu(A_i) = 1,$$

but on the other hand

$$\sum_i \mu(A_i \mid B_j) = \frac{1}{\mu(B_j)} \sum_i \mu(A_i \cap B_j) = \frac{1}{\mu(B_j)}\mu(B_j) = 1,$$

yielding a contradiction. Thus $H(\xi \mid \eta) = H(\xi)$ if and only if for all i, j $\mu(A_i \mid B_j) = \mu(A_i)$, which is equivalent to that ξ and η are independent.

P2. Prove that $d(\xi, \eta) = H(\xi \mid \eta) + H(\eta \mid \xi)$ defines a metric in the space of finite partitions (up to sets of measure 0).

Firstly, notice that given that $H(\xi \mid \eta), H(\eta \mid \xi) \geq 0$ (because the entropy is positive and, for example, problem 1), we have that $d(\xi, \eta) \geq 0$ for every finite partitions ξ and η. Secondly, we have that $d(\xi, \eta) = 0$ if and only if $H(\xi \mid \eta) = H(\eta \mid \xi) = 0$. It follows that $\xi \leq \eta$ and $\eta \leq \xi$, and thus $\eta = \xi$.

Thirdly, the symmetry is direct from the definition.

Finally, we need to verify the triangle inequality. Let ξ, η, γ finite partitions. Notice that it is enough to see that

$$H(\xi \mid \gamma) \leq H(\xi \mid \eta) + H(\eta \mid \gamma).$$

Indeed, using properties (iii) and (iv) from Theorem 63 from the lecture notes, we have that

$$\begin{aligned}
H(\xi \mid \gamma) &\leq H(\xi \vee \eta \mid \gamma) \\
&= H(\xi \vee \eta \vee \gamma) - H(\gamma) \\
&= H(\xi \mid \eta \vee \gamma) + H(\eta \vee \gamma) - H(\gamma) \\
&= H(\xi \mid \eta \vee \gamma) + H(\eta|\gamma) \\
&\leq H(\xi \mid \eta) + H(\eta|\gamma)
\end{aligned}$$

P3. Prove that for every $\epsilon > 0$ there is $\delta > 0$ such that if $\xi = \{A_1, A_2, \ldots, A_r\}$ and $\eta = \{B_1, B_2, \ldots, B_r\}$ are two finite partitions with $\sum_{i=1}^{r} \mu(A_i \Delta B_i) < \delta$, then $d(\xi, \eta) < \epsilon$, where d is the metric defined in the previous problem.

It is enough to show that $H(\xi \mid \eta) < \epsilon$. For this, consider the partition

$$\gamma = \{A_i \cap B_j\}_{i \neq j} \cup \left\{ \bigcup_{i=1}^{r} A_i \cap B_i \right\}.$$

Notice that $\xi \vee \eta = \gamma \vee \eta$, and so,

$$H(\xi \mid \eta) = H(\xi \vee \eta) - H(\eta) = H(\eta \vee \gamma) - H(\eta) = H(\gamma \mid \eta) \leq H(\gamma).$$

Thus, the goal is to show that this partition has small entropy. We see that this partition contains a set of large measure and all the other sets have relatively small measure. This will give that the entropy of this partition is small, which will eventually give the desired result. Observe that $A_i \cap B_j \subseteq \bigcup_{k=1}^{r} A_k \Delta B_k$ for every $i \neq j$. Thus

$$\mu(A_i \cap B_j) \leq \sum_{k=1}^{r} \mu(A_k \Delta B_k) < \delta.$$

On the other hand

$$\mu\left(\bigcup_{k=1}^{r} A_k \cap B_k \right) = \mu\left(\bigcup_{k=1}^{r} A_k \cup B_k \right) + \mu\left(\bigcup_{k=1}^{r} A_k \Delta B_k \right) > 1 - \delta.$$

Now we calculate the entropy of γ.

$$\begin{aligned}
H(\gamma) &= -\sum_{i \neq j} \mu(A_j \cap B_j) \log(\mu(A_i \cap B_j)) - \mu\left(\bigcup_{i=1}^{r} A_i \cap B_i \right) \log\left(\mu(\bigcup_{i=1}^{r} A_i \cap B_i) \right) \\
&= -\sum_{i \neq j} \phi(\mu(A_i \cap B_j)) - \phi\left(\mu\left(\bigcup_{i=1}^{r} A_i \cap B_i \right) \right),
\end{aligned}$$

where $\phi(x) = x \log x$ is strictly decreasing in $(0, \frac{1}{e})$ and strictly increasing in $(\frac{1}{e}, 1)$, so we pick $\delta \in (0, \frac{1}{e})$. Then,

$$H(\gamma) < -r(r-1)\phi(\delta) - \phi(1-\delta) = -r(r-1)\delta \log{(1-\delta)},$$

which is less than ϵ for $\delta > 0$ small enough.

P4. Calculate the entropy of any periodic system.

For any finite partition ξ and for any $n \geq k$ we have that

$$H\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) = H\left(\bigvee_{i=0}^{k-1} T^{-i}\xi\right).$$

Therefore, $h(T,\xi) = \text{lím}_{n\to\infty} \frac{1}{n} H\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) = \text{lím}_{n\to\infty} \frac{1}{n} H\left(\bigvee_{i=0}^{k-1} T^{-i}\xi\right) = 0$. We conclude that $h(T) = 0$.