Proof of Cesaro Means

In order to show that our alternative definition of entropy rate:

$$\mathrm{H}^{'}(\mathcal{X}) = \lim_{n\to\infty} \mathrm{H}\Big(X_n \mid X_1, \ldots, X_{n-1}\Big)$$

is equivalent to the canonical definition

$$\mathrm{H}(\mathcal{X}) = \lim_{n\to\infty} \frac{1}{n}\mathrm{H}(X_1, \ldots, X_n)$$

when the stochastic process is stationary, we used the following theorem

Theorem Cesaro Means:
Let $a_n \to a$, let $b_n = n^{-1}\sum_{i=1}^{n} a_i$, then

$$\lim_{n\to\infty} b_n = a.$$

Proof Recall the meaning of $\lim_{n\to\infty} b_n = a$:
For every $\delta > 0$ there exists a n_δ such that, for every $n > n_delta$, $\|b_n - a\| < \delta$.

Now, since $\lim_{n\to\infty} a_n = a$, we know that $\forall\epsilon, \exists n_\epsilon$ such that $\forall n > n_\epsilon$, $|a_n - a| < \epsilon$.

Choose $\delta = 2\epsilon$.

Fix, ϵ, determine n_ϵ, let $n >> n_\epsilon$, and look at $b_n - a$.

$$\begin{aligned}
|b_n - a| &= \left| n^{-1}\left(\sum_{i=1}^{n} a_i\right) - a \right| \\
&= \left| n^{-1}\left(\sum_{i=1}^{n} a_i\right) - a\frac{n}{n} \right| \\
&= \left| n^{-1}\left(-na + \sum_{i=1}^{n} a_i\right) \right| \\
&= \left| n^{-1}\sum_{i=1}^{n} (a_i - a) \right|
\end{aligned}$$

Now divide the sum on the right hand side into two parts: the first is the sum over the indexes between 1 and n_ϵ, the second is the over the remaining terms

$$\begin{aligned}
|b_n - a| &= \left| \frac{1}{n} \sum_{i=1}^{n} (a_i - a) \right| \\
&\leq \left| \frac{1}{n} \sum_{i=1}^{n_\epsilon} (a_i - a) \right| \qquad (1) \\
&\quad + \left| \frac{1}{n} \sum_{i=n_\epsilon+1}^{n} (a_i - a) \right| \qquad (2)
\end{aligned}$$

We are now going to bound the two sums (1) and (2).

First we bound (1)

$$\begin{aligned}
\left| \frac{1}{n} \sum_{i=1}^{n_\epsilon} (a_i - a) \right| &\leq \frac{1}{n} \sum_{i=1}^{n_\epsilon} |a_i - a| \\
&\leq \frac{1}{n} \sum_{i=1}^{n_\epsilon} \max_{j=1}^{n_\epsilon} |a_j - a| \\
&= \max_{j=1}^{n_\epsilon} |a_j - a| \frac{1}{n} \sum_{i=1}^{n_\epsilon} 1 \qquad (3) \\
&= \max_{j=1}^{n_\epsilon} |a_j - a| \frac{n_\epsilon}{n} \qquad (4)
\end{aligned}$$

and, if we pick n satisfying

$$n > \frac{n_\epsilon \max_{j=1}^{n_\epsilon} |a_j - a|}{\epsilon},$$

then sum (1) satisfies

$$\frac{1}{n} \left| \sum_{i=1}^{n_\epsilon} (a_i - a) \right| < \epsilon$$

Now we bound (2) using a similar trick

$$\left| \frac{1}{n} \sum_{i=n_\epsilon+1}^{n} (a_i - a) \right| \leq \frac{1}{n} \sum_{i=n_\epsilon+1}^{n} |a_i - a|$$

$$
\begin{aligned}
&\leq \frac{1}{n} \sum_{i=n_\epsilon+1}^{n} \max_{i=n_\epsilon+1}^{n} |a_i - a| \\
&= \max_{i=n_\epsilon+1}^{n} |a_i - a| \frac{1}{n} \sum_{i=n_\epsilon+1}^{n} 1 \\
&= \max_{i=n_\epsilon+1}^{n} |a_i - a| \frac{n - n_\epsilon}{n} \\
&\leq \max_{i=n_\epsilon+1}^{n} |a_i - a| \\
&\leq \epsilon \qquad (5)
\end{aligned}
$$

where Inequality 5 is a consequence of how we selected n_ϵ.

We have therefore shown that, if we fix δ, let $\epsilon = \delta/2$, determine n_ϵ, there exists a $n_\delta = \frac{n_\epsilon \max_{j=1}^{n_\epsilon} |a_j - a|}{\epsilon}$ such that, for all $n > n_\delta$, $|b_n - a| < \delta$;
In other words, we have shown that $\lim_{n\to\infty} b_n = a$.