# 2.10 Entropy Rate of a Stationary Source

# Discrete-time Information Source

# Discrete-time Information Source

- In most communication systems, communication takes place continually instead of over a finite period of time.

# Discrete-time Information Source

- In most communication systems, communication takes place continually instead of over a finite period of time.
- Examples: TV broadcast, Internet, cellular systems.

# Discrete-time Information Source

- In most communication systems, communication takes place continually instead of over a finite period of time.
- Examples: TV broadcast, Internet, cellular systems.
- The information source can be modeled as a discrete-time random process $\{X_k, k \geq 1\}$.

# Discrete-time Information Source

- In most communication systems, communication takes place continually instead of over a finite period of time.
- Examples: TV broadcast, Internet, cellular systems.
- The information source can be modeled as a discrete-time random process $\{X_k, k \geq 1\}$.
- $\{X_k, k \geq 1\}$ is an infinite collection of random variables indexed by the set of positive integers. The index $k$ is referred to as the "time" index.

# Discrete-time Information Source

- In most communication systems, communication takes place continually instead of over a finite period of time.
- Examples: TV broadcast, Internet, cellular systems.
- The information source can be modeled as a discrete-time random process $\{X_k, k \geq 1\}$.
- $\{X_k, k \geq 1\}$ is an infinite collection of random variables indexed by the set of positive integers. The index $k$ is referred to as the "time" index.
- Random variables $X_k$ are called letters.

# Discrete-time Information Source

- In most communication systems, communication takes place continually instead of over a finite period of time.
- Examples: TV broadcast, Internet, cellular systems.
- The information source can be modeled as a discrete-time random process $\{X_k, k \geq 1\}$.
- $\{X_k, k \geq 1\}$ is an infinite collection of random variables indexed by the set of positive integers. The index $k$ is referred to as the "time" index.
- Random variables $X_k$ are called letters.
- Assume that $H(X_k) < \infty$ for all $k$.

# Total Entropy of { $X_k$ }

# Total Entropy of { $X_k$ }

- For a finite subset $A$ of the index set $\{k : k \geq 1\}$, the joint entropy $H(X_k, k \in A)$ is finite because

# Total Entropy of { $X_k$ }

- For a finite subset $A$ of the index set $\{k : k \geq 1\}$, the joint entropy $H(X_k, k \in A)$ is finite because

$$H(X_k, k \in A) \leq \sum_{k \in A} H(X_k)$$

# Total Entropy of $\{ X_k \}$

- For a finite subset $A$ of the index set $\{k : k \geq 1\}$, the joint entropy $H(X_k, k \in A)$ is finite because

$$H(X_k, k \in A) \leq \sum_{k \in A} H(X_k) < \infty.$$

# Total Entropy of { $X_k$ }

- For a finite subset $A$ of the index set $\{k : k \geq 1\}$, the joint entropy $H(X_k, k \in A)$ is finite because

$$H(X_k, k \in A) \leq \sum_{k \in A} H(X_k) < \infty.$$

- However, the joint entropy of an infinite collection of letters is infinite except for very special cases.

# Total Entropy of $\{ X_k \}$

- For a finite subset $A$ of the index set $\{k : k \geq 1\}$, the joint entropy $H(X_k, k \in A)$ is finite because

$$H(X_k, k \in A) \leq \sum_{k \in A} H(X_k) < \infty.$$

- However, the joint entropy of an infinite collection of letters is infinite except for very special cases.

- For example, $X_k$ are i.i.d. and $H(X_k) = h$ for all $k$. Then

# Total Entropy of { $X_k$ }

- For a finite subset $A$ of the index set $\{k : k \geq 1\}$, the joint entropy $H(X_k, k \in A)$ is finite because

$$H(X_k, k \in A) \leq \sum_{k \in A} H(X_k) < \infty.$$

- However, the joint entropy of an infinite collection of letters is infinite except for very special cases.

- For example, $X_k$ are i.i.d. and $H(X_k) = h$ for all $k$. Then

$$H(X_k, k \geq 1)$$

# Total Entropy of { $X_k$ }

- For a finite subset $A$ of the index set $\{k : k \geq 1\}$, the joint entropy $H(X_k, k \in A)$ is finite because

$$H(X_k, k \in A) \leq \sum_{k \in A} H(X_k) < \infty.$$

- However, the joint entropy of an infinite collection of letters is infinite except for very special cases.

- For example, $X_k$ are i.i.d. and $H(X_k) = h$ for all $k$. Then

$$H(X_k, k \geq 1) = \sum_{k=1}^{\infty} H(X_k)$$

# Total Entropy of { $X_k$ }

- For a finite subset $A$ of the index set $\{k : k \geq 1\}$, the joint entropy $H(X_k, k \in A)$ is finite because

$$H(X_k, k \in A) \leq \sum_{k \in A} H(X_k) < \infty.$$

- However, the joint entropy of an infinite collection of letters is infinite except for very special cases.

- For example, $X_k$ are i.i.d. and $H(X_k) = h$ for all $k$. Then

$$H(X_k, k \geq 1) = \sum_{k=1}^{\infty} H(X_k) = \sum_{k=1}^{\infty} h$$

# Total Entropy of { $X_k$ }

- For a finite subset $A$ of the index set $\{k : k \geq 1\}$, the joint entropy $H(X_k, k \in A)$ is finite because

$$H(X_k, k \in A) \leq \sum_{k \in A} H(X_k) < \infty.$$

- However, the joint entropy of an infinite collection of letters is infinite except for very special cases.

- For example, $X_k$ are i.i.d. and $H(X_k) = h$ for all $k$. Then

$$H(X_k, k \geq 1) = \sum_{k=1}^{\infty} H(X_k) = \sum_{k=1}^{\infty} h = \infty.$$

# Total Entropy of { $X_k$ }

- For a finite subset $A$ of the index set $\{k : k \geq 1\}$, the joint entropy $H(X_k, k \in A)$ is finite because

$$H(X_k, k \in A) \leq \sum_{k \in A} H(X_k) < \infty.$$

- However, the joint entropy of an infinite collection of letters is infinite except for very special cases.

- For example, $X_k$ are i.i.d. and $H(X_k) = h$ for all $k$. Then

$$H(X_k, k \geq 1) = \sum_{k=1}^{\infty} H(X_k) = \sum_{k=1}^{\infty} h = \infty.$$

- In general, it is not meaningful to discuss $H(X_k, k \geq 1)$.

# Entropy Rate

# Entropy Rate

We are motivated to definite the entropy rate of an information source, which gives the average entropy of a letter of the source.

# Entropy Rate

We are motivated to definite the entropy rate of an information source, which gives the average entropy of a letter of the source.

**Definition 2.54** The entropy rate of an information source $\{X_k\}$ is defined as

$$H_X = \lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \cdots, X_n)$$

when the limit exists.

# Entropy Rate May Exist

# Entropy Rate May Exist

**Example 2.55** Let $\{X_k\}$ be an i.i.d. source with generic random variable $X$.

# Entropy Rate May Exist

**Example 2.55** Let $\{X_k\}$ be an i.i.d. source with generic random variable $X$. Then

$$\lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \cdots, X_n)$$

# Entropy Rate May Exist

**Example 2.55** Let $\{X_k\}$ be an i.i.d. source with generic random variable $X$. Then

$$\lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \cdots, X_n) = \lim_{n\to\infty} \frac{nH(X)}{n}$$

# Entropy Rate May Exist

**Example 2.55** Let $\{X_k\}$ be an i.i.d. source with generic random variable $X$. Then

$$\lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \cdots, X_n) = \lim_{n\to\infty} \frac{nH(X)}{n}$$

# Entropy Rate May Exist

**Example 2.55** Let $\{X_k\}$ be an i.i.d. source with generic random variable $X$. Then

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \cdots, X_n) &= \lim_{n\to\infty} \frac{\cancel{n} H(X)}{\cancel{n}} \\
&= \lim_{n\to\infty} H(X)
\end{aligned}$$

# Entropy Rate May Exist

**Example 2.55** Let $\{X_k\}$ be an i.i.d. source with generic random variable $X$. Then

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}H(X_1,X_2,\cdots,X_n) &= \lim_{n\to\infty}\frac{nH(X)}{n}\\ &= \lim_{n\to\infty}H(X)\\ &= H(X),\end{aligned}$$

# Entropy Rate May Exist

**Example 2.55** Let $\{X_k\}$ be an i.i.d. source with generic random variable $X$. Then

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \cdots, X_n) &= \lim_{n\to\infty} \frac{nH(X)}{n} \\ &= \lim_{n\to\infty} H(X) \\ &= H(X),\end{aligned}$$

i.e., the entropy rate of an i.i.d. source is the entropy of any of its single letters.

# Entropy Rate May Not Exist

# Entropy Rate May Not Exist

**Example 2.56** Let $\{X_k\}$ be a source such that $X_k$ are mutually independent and $H(X_k) = k$ for $k \geq 1$. Then

# Entropy Rate May Not Exist

**Example 2.56** Let $\{X_k\}$ be a source such that $X_k$ are mutually independent and $H(X_k) = k$ for $k \geq 1$. Then

$$\frac{1}{n} H(X_1, X_2, \cdots, X_n)$$

# Entropy Rate May Not Exist

**Example 2.56** Let $\{X_k\}$ be a source such that $X_k$ are mutually independent and $H(X_k) = k$ for $k \geq 1$. Then

$$\frac{1}{n}H(X_1, X_2, \cdots, X_n) = \frac{1}{n}\sum_{k=1}^{n} H(X_k)$$

# Entropy Rate May Not Exist

**Example 2.56** Let $\{X_k\}$ be a source such that $X_k$ are mutually independent and $H(X_k) = k$ for $k \geq 1$. Then

$$\frac{1}{n} H(X_1, X_2, \cdots, X_n) = \frac{1}{n} \sum_{k=1}^{n} \underline{H(X_k)}$$

# Entropy Rate May Not Exist

**Example 2.56** Let $\{X_k\}$ be a source such that $X_k$ are mutually independent and $H(X_k) = k$ for $k \geq 1$. Then

$$\begin{aligned} \frac{1}{n} H(X_1, X_2, \cdots, X_n) &= \frac{1}{n} \sum_{k=1}^{n} H(X_k) \\ &= \frac{1}{n} \sum_{k=1}^{n} k \end{aligned}$$

# Entropy Rate May Not Exist

**Example 2.56** Let $\{X_k\}$ be a source such that $X_k$ are mutually independent and $H(X_k) = k$ for $k \geq 1$. Then

$$
\begin{aligned}
\frac{1}{n} H(X_1, X_2, \cdots, X_n) &= \frac{1}{n} \sum_{k=1}^{n} H(X_k) \\
&= \frac{1}{n} \sum_{k=1}^{n} k
\end{aligned}
$$

# Entropy Rate May Not Exist

**Example 2.56** Let $\{X_k\}$ be a source such that $X_k$ are mutually independent and $H(X_k) = k$ for $k \geq 1$. Then

$$
\begin{aligned}
\frac{1}{n} H(X_1, X_2, \cdots, X_n) &= \frac{1}{n} \sum_{k=1}^{n} H(X_k) \\
&= \frac{1}{n} \sum_{k=1}^{n} k \\
&= \frac{1}{n} \frac{n(n+1)}{2}
\end{aligned}
$$

# Entropy Rate May Not Exist

**Example 2.56** Let $\{X_k\}$ be a source such that $X_k$ are mutually independent and $H(X_k) = k$ for $k \geq 1$. Then

$$\begin{aligned}
\frac{1}{n} H(X_1, X_2, \cdots, X_n) &= \frac{1}{n} \sum_{k=1}^{n} H(X_k) \\
&= \frac{1}{n} \sum_{k=1}^{n} k \\
&= \frac{1}{\cancel{n}} \frac{\cancel{n}(n+1)}{2}
\end{aligned}$$

# Entropy Rate May Not Exist

**Example 2.56** Let $\{X_k\}$ be a source such that $X_k$ are mutually independent and $H(X_k) = k$ for $k \geq 1$. Then

$$\begin{aligned}
\frac{1}{n} H(X_1, X_2, \cdots, X_n) &= \frac{1}{n} \sum_{k=1}^{n} H(X_k) \\
&= \frac{1}{n} \sum_{k=1}^{n} k \\
&= \frac{1}{n} \frac{n(n+1)}{2} \\
&= \frac{1}{2}(n+1),
\end{aligned}$$

# Entropy Rate May Not Exist

**Example 2.56** Let $\{X_k\}$ be a source such that $X_k$ are mutually independent and $H(X_k) = k$ for $k \geq 1$. Then

$$\begin{aligned}\frac{1}{n}H(X_1, X_2, \cdots, X_n) &= \frac{1}{n}\sum_{k=1}^{n} H(X_k) \\ &= \frac{1}{n}\sum_{k=1}^{n} k \\ &= \frac{1}{n}\frac{n(n+1)}{2} \\ &= \frac{1}{2}(n+1),\end{aligned}$$

which does not converge as $n \to \infty$ although $H(X_k) < \infty$ for all $k$. Therefore, the entropy rate of $\{X_k\}$ does not exist.

# The Limit $H'_X$

# The Limit $H'_X$

- Toward characterizing the asymptotic behavior of $\{X_k\}$, it is natural to consider the limit

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

  if it exists.

# The Limit $H'_X$

- Toward characterizing the asymptotic behavior of $\{X_k\}$, it is natural to consider the limit

$$H'_X = \lim_{n \to \infty} H(X_n | X_1, X_2, \cdots, X_{n-1})$$

  if it exists.

- The quantity $H(X_n | X_1, X_2, \cdots, X_{n-1})$ is the conditional entropy of the next letter given that we know all the past history of the source.

# The Limit $H'_X$

- Toward characterizing the asymptotic behavior of $\{X_k\}$, it is natural to consider the limit

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

  if it exists.

- The quantity $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is the conditional entropy of the next letter given that we know all the past history of the source.

# The Limit $H'_X$

- Toward characterizing the asymptotic behavior of $\{X_k\}$, it is natural to consider the limit

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

  if it exists.

- The quantity $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is the conditional entropy of the next letter given that we know all the past history of the source.

- $H'_X$ is the limit of this quantity after the source has been run for an indefinite amount of time.

# The Limit $H'_X$

- Toward characterizing the asymptotic behavior of $\{X_k\}$, it is natural to consider the limit

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

  if it exists.

- The quantity $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is the conditional entropy of the next letter given that we know all the past history of the source.

- $H'_X$ is the limit of this quantity after the source has been run for an indefinite amount of time.

- Is $H'_X = H_X$?

# Stationary Information Source

# Stationary Information Source

A stationary information source is one such that any finite block of random variables and any of its time-shift versions have exactly the same joint distribution.

# Stationary Information Source

A stationary information source is one such that any finite block of random variables and any of its time-shift versions have exactly the same joint distribution.

**Definition 2.57** An information source $\{X_k\}$ is stationary if

$$X_1, X_2, \cdots, X_m$$

and

$$X_{1+l}, X_{2+l}, \cdots, X_{m+l}$$

have the same joint distribution for any $m, l \geq 1$.

# Stationary Information Source

A stationary information source is one such that any finite block of random variables and any of its time-shift versions have exactly the same joint distribution.

**Definition 2.57** An information source $\{X_k\}$ is stationary if

$$X_1, X_2, \cdots, X_m$$

and

$$X_{1+l}, X_{2+l}, \cdots, X_{m+l}$$

have the same joint distribution for any $m, l \geq 1$.

$$X_1, X_2, \cdots, X_m \quad \cdots \quad X_{1+l}, X_{2+l}, \cdots, X_{m+l} \quad \cdots$$

# Stationary Information Source

A stationary information source is one such that any finite block of random variables and any of its time-shift versions have exactly the same joint distribution.

**Definition 2.57** An information source $\{X_k\}$ is stationary if

$$X_1, X_2, \cdots, X_m$$

and

$$X_{1+l}, X_{2+l}, \cdots, X_{m+l}$$

have the same joint distribution for any $m, l \geq 1$.

$$\overbrace{X_1, X_2, \cdots, X_m}^{m} \quad \cdots \quad X_{1+l}, X_{2+l}, \cdots, X_{m+l} \quad \cdots$$

# Stationary Information Source

A stationary information source is one such that any finite block of random variables and any of its time-shift versions have exactly the same joint distribution.

**Definition 2.57** An information source $\{X_k\}$ is stationary if

$$X_1, X_2, \cdots, X_m$$

and

$$X_{1+l}, X_{2+l}, \cdots, X_{m+l}$$

have the same joint distribution for any $m, l \geq 1$.

$$\overbrace{X_1, X_2, \cdots, X_m}^{m} \quad \cdots \quad X_{1+l}, X_{2+l}, \cdots, X_{m+l} \quad \cdots$$

$$\longleftarrow \quad l \quad \longrightarrow$$

# Stationary Information Source

A stationary information source is one such that any finite block of random variables and any of its time-shift versions have exactly the same joint distribution.

**Definition 2.57** An information source $\{X_k\}$ is stationary if

$$X_1, X_2, \cdots, X_m$$

and

$$X_{1+l}, X_{2+l}, \cdots, X_{m+l}$$

have the same joint distribution for any $m, l \geq 1$.

$$\overbrace{X_1, X_2, \cdots, X_m}^{m} \quad \cdots \quad \overbrace{X_{1+l}, X_{2+l}, \cdots, X_{m+l}}^{m} \quad \cdots$$

$$\longleftrightarrow \; l \; \longleftrightarrow$$

# Stationarity and $H'_X$

# Stationarity and $H'_X$

Recall that

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1}).$$

# Stationarity and $H'_X$

Recall that

$$H'_X = \lim_{n \to \infty} H(X_n | X_1, X_2, \cdots, X_{n-1}).$$

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

**Proof**

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

**Proof**

$$H'_X = \lim_{n \to \infty} H(X_n | X_1, X_2, \cdots, X_{n-1})$$

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Proof**

1. Since $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is lower bounded by zero for all $n$, it suffices to prove that this quantity is non-increasing in $n$ to conclude that the limit $H'_X$ exists.

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Proof**

1. Since $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is lower bounded by zero for all $n$, it suffices to prove that this quantity is non-increasing in $n$ to conclude that the limit $H'_X$ exists.

2. Toward this end, for $n \geq 2$, consider

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Proof**

1. Since $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is lower bounded by zero for all $n$, it suffices to prove that this quantity is non-increasing in $n$ to conclude that the limit $H'_X$ exists.

2. Toward this end, for $n \geq 2$, consider

$$H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Proof**

1. Since $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is lower bounded by zero for all $n$, it suffices to prove that this quantity is non-increasing in $n$ to conclude that the limit $H'_X$ exists.

2. Toward this end, for $n \geq 2$, consider

$$H(X_n|\cancel{X_1}, X_2, \cdots, X_{n-1})$$

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

$$H'_X = \lim_{n\to\infty} H(X_n | X_1, X_2, \cdots, X_{n-1})$$

**Proof**

1. Since $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is lower bounded by zero for all $n$, it suffices to prove that this quantity is non-increasing in $n$ to conclude that the limit $H'_X$ exists.

2. Toward this end, for $n \geq 2$, consider

$$\begin{aligned} & H(X_n | X_1, X_2, \cdots, X_{n-1}) \\ \leq \quad & H(X_n | X_2, X_3, \cdots, X_{n-1}) \end{aligned}$$

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Proof**

1. Since $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is lower bounded by zero for all $n$, it suffices to prove that this quantity is non-increasing in $n$ to conclude that the limit $H'_X$ exists.

2. Toward this end, for $n \geq 2$, consider

$$\begin{aligned} & H(X_n|X_1, X_2, \cdots, X_{n-1}) \\ \leq \quad & H(X_n|X_2, X_3, \cdots, X_{n-1}) \\ = \quad & H(X_{n-1}|X_1, X_2, \cdots, X_{n-2}), \end{aligned}$$

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Proof**

1. Since $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is lower bounded by zero for all $n$, it suffices to prove that this quantity is non-increasing in $n$ to conclude that the limit $H'_X$ exists.

2. Toward this end, for $n \geq 2$, consider

$$\begin{aligned}
&H(X_n|X_1, X_2, \cdots, X_{n-1}) \\
&\quad \leq \quad H(X_n|X_2, X_3, \cdots, X_{n-1}) \\
&\quad = \quad H(X_{n-1}|X_1, X_2, \cdots, X_{n-2}),
\end{aligned}$$

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

$$H'_X = \lim_{n \to \infty} H(X_n | X_1, X_2, \cdots, X_{n-1})$$

**Proof**

1. Since $H(X_n | X_1, X_2, \cdots, X_{n-1})$ is lower bounded by zero for all $n$, it suffices to prove that this quantity is non-increasing in $n$ to conclude that the limit $H'_X$ exists.

2. Toward this end, for $n \geq 2$, consider

$$\begin{aligned} & H(X_n | X_1, X_2, \cdots, X_{n-1}) \\ \leq \quad & H(X_{\underline{n}} | X_2, X_3, \cdots, X_{n-1}) \\ = \quad & H(X_{\underline{n-1}} | X_1, X_2, \cdots, X_{n-2}), \end{aligned}$$

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Proof**

1. Since $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is lower bounded by zero for all $n$, it suffices to prove that this quantity is non-increasing in $n$ to conclude that the limit $H'_X$ exists.

2. Toward this end, for $n \geq 2$, consider

$$\begin{aligned} & H(X_n|X_1, X_2, \cdots, X_{n-1}) \\ & \quad \leq \quad H(X_{\underline{n}}|X_{\underline{2}}, X_3, \cdots, X_{n-1}) \\ & \quad = \quad H(X_{\underline{n-1}}|X_1, X_2, \cdots, X_{n-2}), \end{aligned}$$

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Proof**

1. Since $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is lower bounded by zero for all $n$, it suffices to prove that this quantity is non-increasing in $n$ to conclude that the limit $H'_X$ exists.

2. Toward this end, for $n \geq 2$, consider

$$\begin{aligned} & H(X_n|X_1, X_2, \cdots, X_{n-1}) \\ \leq \quad & H(X_{\underline{n}}|X_{\underline{2}}, X_3, \cdots, X_{n-1}) \\ = \quad & H(X_{\underline{n-1}}|X_{\underline{1}}, X_2, \cdots, X_{n-2}), \end{aligned}$$

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Proof**

1. Since $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is lower bounded by zero for all $n$, it suffices to prove that this quantity is non-increasing in $n$ to conclude that the limit $H'_X$ exists.

2. Toward this end, for $n \geq 2$, consider

$$\begin{aligned} & H(X_n|X_1, X_2, \cdots, X_{n-1}) \\ \leq \quad & H(X_{\underline{n}}|X_{\underline{2}}, X_{\underline{3}}, \cdots, X_{n-1}) \\ = \quad & H(X_{\underline{n-1}}|X_{\underline{1}}, X_2, \cdots, X_{n-2}), \end{aligned}$$

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Proof**

1. Since $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is lower bounded by zero for all $n$, it suffices to prove that this quantity is non-increasing in $n$ to conclude that the limit $H'_X$ exists.

2. Toward this end, for $n \geq 2$, consider

$$\begin{aligned} &H(X_n|X_1, X_2, \cdots, X_{n-1}) \\ &\quad\leq \quad H(X_{\underline{n}}|X_{\underline{2}}, X_{\underline{3}}, \cdots, X_{n-1}) \\ &\quad= \quad H(X_{\underline{n-1}}|X_{\underline{1}}, X_{\underline{2}}, \cdots, X_{n-2}), \end{aligned}$$

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Proof**

1. Since $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is lower bounded by zero for all $n$, it suffices to prove that this quantity is non-increasing in $n$ to conclude that the limit $H'_X$ exists.

2. Toward this end, for $n \geq 2$, consider

$$\begin{aligned}
&H(X_n|X_1, X_2, \cdots, X_{n-1}) \\
&\quad \leq \quad H(X_{\underline{n}}|X_{\underline{2}}, X_{\underline{3}}, \cdots, X_{\underline{n-1}}) \\
&\quad = \quad H(X_{\underline{n-1}}|X_{\underline{1}}, X_{\underline{2}}, \cdots, X_{n-2}),
\end{aligned}$$

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

**Proof**

1. Since $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is lower bounded by zero for all $n$, it suffices to prove that this quantity is non-increasing in $n$ to conclude that the limit $H'_X$ exists.

2. Toward this end, for $n \geq 2$, consider

$$\begin{aligned} & H(X_n|X_1, X_2, \cdots, X_{n-1}) \\ \leq \quad & H(X_{\underline{n}}|X_{\underline{2}}, X_{\underline{3}}, \cdots, X_{\underline{n-1}}) \\ = \quad & H(X_{\underline{n-1}}|X_{\underline{1}}, X_{\underline{2}}, \cdots, X_{\underline{n-2}}), \end{aligned}$$

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Proof**

1. Since $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is lower bounded by zero for all $n$, it suffices to prove that this quantity is non-increasing in $n$ to conclude that the limit $H'_X$ exists.

2. Toward this end, for $n \geq 2$, consider

$$\begin{aligned} & H(X_n|X_1, X_2, \cdots, X_{n-1}) \\ \leq \quad & H(X_{\underline{n}}|X_{\underline{2}}, X_{\underline{3}}, \cdots, X_{\underline{n-1}}) \\ = \quad & H(X_{\underline{n-1}}|X_{\underline{1}}, X_{\underline{2}}, \cdots, X_{\underline{n-2}}), \end{aligned}$$

where the last step is justified by the stationarity of $\{X_k\}$.

**Lemma 2.58** Let $\{X_k\}$ be a stationary source. Then $H'_X$ exists.

**Proof**

1. Since $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is lower bounded by zero for all $n$, it suffices to prove that this quantity is non-increasing in $n$ to conclude that the limit $H'_X$ exists.

2. Toward this end, for $n \geq 2$, consider

$$\begin{aligned} & H(X_n|X_1, X_2, \cdots, X_{n-1}) \\ \leq \quad & H(X_n|X_2, X_3, \cdots, X_{n-1}) \\ = \quad & H(X_{n-1}|X_1, X_2, \cdots, X_{n-2}), \end{aligned}$$

where the last step is justified by the stationarity of $\{X_k\}$.

3. Therefore, $H(X_n|X_1, X_2, \cdots, X_{n-1})$ is non-increasing in $n$. The lemma is proved.

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

# Cesáro Mean

# Cesáro Mean

- Consider a sequence $\{a_n, n \geq 1\}$.

# Cesáro Mean

- Consider a sequence $\{a_n, n \geq 1\}$.
- Construct a sequence $\{b_n, n \geq 1\}$ where $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$.

# Cesáro Mean

- Consider a sequence $\{a_n, n \geq 1\}$.
- Construct a sequence $\{b_n, n \geq 1\}$ where $b_n = \frac{1}{n} \sum_{i=1}^{n} a_i$.
- $b_n$ is the average of the first $n$ terms in $\{a_n\}$.

# Cesáro Mean

- Consider a sequence $\{a_n, n \geq 1\}$.
- Construct a sequence $\{b_n, n \geq 1\}$ where $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$.
- $b_n$ is the average of the first $n$ terms in $\{a_n\}$.
- $b_n, n \geq 1$ are called the Cesáro means of $\{a_n\}$.

# Cesáro Mean

- Consider a sequence $\{a_n, n \geq 1\}$.
- Construct a sequence $\{b_n, n \geq 1\}$ where $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$.
- $b_n$ is the average of the first $n$ terms in $\{a_n\}$.
- $b_n, n \geq 1$ are called the Cesáro means of $\{a_n\}$.
- It can be shown that if $a_n \to a$, then $b_n \to a$.

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n} \sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

$|b_n - a|$

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

$|b_n - a|$

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

$$|b_n - a| = \left| \frac{1}{n} \sum_{i=1}^{n} a_i - a \right|$$

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

$$|b_n - a| = \left| \frac{1}{n} \sum_{i=1}^{n} a_i - a \right|$$

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

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

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

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

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

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

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

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

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

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

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

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

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

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

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

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

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

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

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

$$\begin{aligned}
|b_n - a| &= \left| \frac{1}{n} \sum_{i=1}^{n} a_i - a \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{n} a_i - \frac{1}{n} \sum_{i=1}^{n} a \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{n} (a_i - a) \right| \\
&\leq \frac{1}{n} \sum_{i=1}^{n} |a_i - a| \\
&= \frac{1}{n} \left( \sum_{i=1}^{N(\epsilon)} |a_i - a| + \sum_{i=N(\epsilon)+1}^{n} |a_i - a| \right) \\
&< \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \frac{1}{n} \underline{\sum_{i=N(\epsilon)+1}^{n}} \epsilon \\
&= \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \frac{\underline{n - N(\epsilon)}}{n} \cdot \epsilon
\end{aligned}$$

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

$$
\begin{aligned}
|b_n - a| &= \left| \frac{1}{n} \sum_{i=1}^{n} a_i - a \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{n} a_i - \frac{1}{n} \sum_{i=1}^{n} a \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{n} (a_i - a) \right| \\
&\leq \frac{1}{n} \sum_{i=1}^{n} |a_i - a| \\
&= \frac{1}{n} \left( \sum_{i=1}^{N(\epsilon)} |a_i - a| + \sum_{i=N(\epsilon)+1}^{n} |a_i - a| \right) \\
&< \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \frac{1}{n} \sum_{i=N(\epsilon)+1}^{n} \epsilon \\
&= \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \underbrace{\frac{n - N(\epsilon)}{n}}_{< 1} \cdot \epsilon
\end{aligned}
$$

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

$$
\begin{aligned}
|b_n - a| &= \left| \frac{1}{n} \sum_{i=1}^{n} a_i - a \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{n} a_i - \frac{1}{n} \sum_{i=1}^{n} a \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{n} (a_i - a) \right| \\
&\leq \frac{1}{n} \sum_{i=1}^{n} |a_i - a| \\
&= \frac{1}{n} \left( \sum_{i=1}^{N(\epsilon)} |a_i - a| + \sum_{i=N(\epsilon)+1}^{n} |a_i - a| \right) \\
&< \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \frac{1}{n} \sum_{i=N(\epsilon)+1}^{n} \epsilon \\
&= \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \overbrace{\frac{n - N(\epsilon)}{n}}^{< 1} \cdot \epsilon \\
&< \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \epsilon.
\end{aligned}
$$

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n} \sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

$$
\begin{aligned}
|b_n - a| &= \left| \frac{1}{n} \sum_{i=1}^{n} a_i - a \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{n} a_i - \frac{1}{n} \sum_{i=1}^{n} a \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{n} (a_i - a) \right| \\
&\leq \frac{1}{n} \sum_{i=1}^{n} |a_i - a| \\
&= \frac{1}{n} \left( \sum_{i=1}^{N(\epsilon)} |a_i - a| + \sum_{i=N(\epsilon)+1}^{n} |a_i - a| \right) \\
&< \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \frac{1}{n} \sum_{i=N(\epsilon)+1}^{n} \epsilon \\
&= \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \frac{n - N(\epsilon)}{n} \cdot \epsilon \\
&< \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \epsilon.
\end{aligned}
$$

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

$$
\begin{aligned}
|b_n - a| &= \left| \frac{1}{n} \sum_{i=1}^{n} a_i - a \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{n} a_i - \frac{1}{n} \sum_{i=1}^{n} a \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{n} (a_i - a) \right| \\
&\leq \frac{1}{n} \sum_{i=1}^{n} |a_i - a| \\
&= \frac{1}{n} \left( \sum_{i=1}^{N(\epsilon)} |a_i - a| + \sum_{i=N(\epsilon)+1}^{n} |a_i - a| \right) \\
&< \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \frac{1}{n} \sum_{i=N(\epsilon)+1}^{n} \epsilon \\
&= \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \frac{n - N(\epsilon)}{n} \cdot \epsilon \\
&< \underline{\frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a|} + \epsilon.
\end{aligned}
$$

3. The first term tends to 0 as $n \to \infty$.

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

$$\begin{aligned}
|b_n - a| &= \left|\frac{1}{n}\sum_{i=1}^{n} a_i - a\right| \\
&= \left|\frac{1}{n}\sum_{i=1}^{n} a_i - \frac{1}{n}\sum_{i=1}^{n} a\right| \\
&= \left|\frac{1}{n}\sum_{i=1}^{n} (a_i - a)\right| \\
&\leq \frac{1}{n}\sum_{i=1}^{n} |a_i - a| \\
&= \frac{1}{n}\left(\sum_{i=1}^{N(\epsilon)} |a_i - a| + \sum_{i=N(\epsilon)+1}^{n} |a_i - a|\right) \\
&< \frac{1}{n}\sum_{i=1}^{N(\epsilon)} |a_i - a| + \frac{1}{n}\sum_{i=N(\epsilon)+1}^{n} \epsilon \\
&= \frac{1}{n}\sum_{i=1}^{N(\epsilon)} |a_i - a| + \frac{n - N(\epsilon)}{n} \cdot \epsilon \\
&< \frac{1}{n}\sum_{i=1}^{N(\epsilon)} |a_i - a| + \epsilon.
\end{aligned}$$

3. The first term tends to 0 as $n \to \infty$.

4. Therefore, for any $\epsilon > 0$, by taking $n$ to be sufficiently large, we can make $|b_n - a| < 2\epsilon$.

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

$$\begin{aligned}
|b_n - a| &= \left| \frac{1}{n}\sum_{i=1}^{n} a_i - a \right| \\
&= \left| \frac{1}{n}\sum_{i=1}^{n} a_i - \frac{1}{n}\sum_{i=1}^{n} a \right| \\
&= \left| \frac{1}{n}\sum_{i=1}^{n} (a_i - a) \right| \\
&\leq \frac{1}{n}\sum_{i=1}^{n} |a_i - a| \\
&= \frac{1}{n}\left( \sum_{i=1}^{N(\epsilon)} |a_i - a| + \sum_{i=N(\epsilon)+1}^{n} |a_i - a| \right) \\
&< \frac{1}{n}\sum_{i=1}^{N(\epsilon)} |a_i - a| + \frac{1}{n}\sum_{i=N(\epsilon)+1}^{n} \epsilon \\
&= \frac{1}{n}\sum_{i=1}^{N(\epsilon)} |a_i - a| + \frac{n - N(\epsilon)}{n} \cdot \epsilon \\
&< \underbrace{\frac{1}{n}\sum_{i=1}^{N(\epsilon)} |a_i - a|}_{< \epsilon} + \epsilon.
\end{aligned}$$

3. The first term tends to 0 as $n \to \infty$.

4. Therefore, for any $\epsilon > 0$, by taking $n$ to be sufficiently large, we can make $|b_n - a| < 2\epsilon$.

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

$$
\begin{aligned}
|b_n - a| &= \left| \frac{1}{n} \sum_{i=1}^{n} a_i - a \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{n} a_i - \frac{1}{n} \sum_{i=1}^{n} a \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{n} (a_i - a) \right| \\
&\leq \frac{1}{n} \sum_{i=1}^{n} |a_i - a| \\
&= \frac{1}{n} \left( \sum_{i=1}^{N(\epsilon)} |a_i - a| + \sum_{i=N(\epsilon)+1}^{n} |a_i - a| \right) \\
&< \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \frac{1}{n} \sum_{i=N(\epsilon)+1}^{n} \epsilon \\
&= \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \frac{n - N(\epsilon)}{n} \cdot \epsilon \\
&< \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \epsilon.
\end{aligned}
$$

$< \epsilon$

3. The first term tends to 0 as $n \to \infty$.

4. Therefore, for any $\epsilon > 0$, by taking $n$ to be sufficiently large, we can make $|b_n - a| < 2\epsilon$.

**Lemma 2.59 (Cesáro Mean)** Let $a_k$ and $b_k$ be real numbers. If $a_n \to a$ as $n \to \infty$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $b_n \to a$ as $n \to \infty$.

**Proof**

1. Since $a_n \to a$ as $n \to \infty$, for every $\epsilon > 0$, there exists $N(\epsilon)$ such that $|a_n - a| < \epsilon$ for all $n > N(\epsilon)$.

2. For $n > N(\epsilon)$, consider

$$
\begin{aligned}
|b_n - a| &= \left| \frac{1}{n} \sum_{i=1}^{n} a_i - a \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{n} a_i - \frac{1}{n} \sum_{i=1}^{n} a \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{n} (a_i - a) \right| \\
&\leq \frac{1}{n} \sum_{i=1}^{n} |a_i - a| \\
&= \frac{1}{n} \left( \sum_{i=1}^{N(\epsilon)} |a_i - a| + \sum_{i=N(\epsilon)+1}^{n} |a_i - a| \right) \\
&< \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \frac{1}{n} \sum_{i=N(\epsilon)+1}^{n} \epsilon \\
&= \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \frac{n - N(\epsilon)}{n} \cdot \epsilon \\
&< \frac{1}{n} \sum_{i=1}^{N(\epsilon)} |a_i - a| + \epsilon.
\end{aligned}
$$

$< \epsilon$

3. The first term tends to 0 as $n \to \infty$.

4. Therefore, for any $\epsilon > 0$, by taking $n$ to be sufficiently large, we can make $|b_n - a| < 2\epsilon$.

5. Hence $b_n \to a$ as $n \to \infty$, proving the lemma.

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Idea of Proof:** Use stationarity of $\{X_k\}$ and Cesáro mean.

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$a_n = H(X_n | X_1, X_2, \cdots, X_{n-1})$$

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$
\begin{aligned}
a_n &= H(X_n|X_1, X_2, \cdots, X_{n-1}) \\
b_n &= \frac{1}{n} H(X_1, X_2, \cdots, X_n).
\end{aligned}
$$

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$
\begin{aligned}
a_n &= H(X_n | X_1, X_2, \cdots, X_{n-1}) \\
b_n &= \frac{1}{n} H(X_1, X_2, \cdots, X_n).
\end{aligned}
$$

4. By the chain rule for entropy,

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$\begin{aligned} a_n &= H(X_n|X_1, X_2, \cdots, X_{n-1}) \\ b_n &= \frac{1}{n} H(X_1, X_2, \cdots, X_n). \end{aligned}$$

4. By the chain rule for entropy,

$$\frac{1}{n} H(X_1, X_2, \cdots, X_n)$$

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$\begin{aligned} a_n &= H(X_n | X_1, X_2, \cdots, X_{n-1}) \\ b_n &= \frac{1}{n} H(X_1, X_2, \cdots, X_n). \end{aligned}$$

4. By the chain rule for entropy,

$$\begin{aligned} &\frac{1}{n} H(X_1, X_2, \cdots, X_n) \\ &\quad = \frac{1}{n} \sum_{k=1}^{n} H(X_k | X_1, X_2, \cdots, X_{k-1}), \end{aligned}$$

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$\begin{aligned} a_n &= H(X_n | X_1, X_2, \cdots, X_{n-1}) \\ b_n &= \frac{1}{n} H(X_1, X_2, \cdots, X_n). \end{aligned}$$

4. By the chain rule for entropy,

$$\underbrace{\frac{1}{n} H(X_1, X_2, \cdots, X_n)}_{b_n} = \frac{1}{n} \sum_{k=1}^{n} H(X_k | X_1, X_2, \cdots, X_{k-1}),$$

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$\begin{aligned} a_n &= H(X_n|X_1, X_2, \cdots, X_{n-1}) \\ b_n &= \frac{1}{n} H(X_1, X_2, \cdots, X_n). \end{aligned}$$

4. By the chain rule for entropy,

$$\underbrace{\frac{1}{n} H(X_1, X_2, \cdots, X_n)}_{b_n} = \frac{1}{n} \sum_{k=1}^{n} \underbrace{H(X_k|X_1, X_2, \cdots, X_{k-1})}_{a_k},$$

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$\begin{aligned} a_n &= H(X_n | X_1, X_2, \cdots, X_{n-1}) \\ b_n &= \frac{1}{n} H(X_1, X_2, \cdots, X_n). \end{aligned}$$

4. By the chain rule for entropy,

$$\underbrace{\frac{1}{n} H(X_1, X_2, \cdots, X_n)}_{b_n} = \frac{1}{n} \sum_{k=1}^{n} \underbrace{H(X_k | X_1, X_2, \cdots, X_{k-1})}_{a_k},$$

or

$$b_n = \frac{1}{n} \sum_{k=1}^{n} a_k.$$

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$\begin{aligned} a_n &= H(X_n | X_1, X_2, \cdots, X_{n-1}) \\ b_n &= \frac{1}{n} H(X_1, X_2, \cdots, X_n). \end{aligned}$$

4. By the chain rule for entropy,

$$\underbrace{\frac{1}{n} H(X_1, X_2, \cdots, X_n)}_{b_n} = \frac{1}{n} \sum_{k=1}^{n} \underbrace{H(X_k | X_1, X_2, \cdots, X_{k-1})}_{a_k},$$

or

$$b_n = \frac{1}{n} \sum_{k=1}^{n} a_k.$$

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$\begin{aligned} a_n &= H(X_n | X_1, X_2, \cdots, X_{n-1}) \\ b_n &= \frac{1}{n} H(X_1, X_2, \cdots, X_n). \end{aligned}$$

4. By the chain rule for entropy,

$$\begin{aligned} &\frac{1}{n} H(X_1, X_2, \cdots, X_n) \\ &\quad = \frac{1}{n} \sum_{k=1}^{n} H(X_k | X_1, X_2, \cdots, X_{k-1}), \end{aligned}$$

or

$$b_n = \frac{1}{n} \sum_{k=1}^{n} a_k.$$

5. Therefore, $b_n, n \geq 1$ are the Cesáro means of $\{a_n\}$.

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$a_n = H(X_n|X_1, X_2, \cdots, X_{n-1})$$
$$b_n = \frac{1}{n} H(X_1, X_2, \cdots, X_n).$$

4. By the chain rule for entropy,

$$\frac{1}{n} H(X_1, X_2, \cdots, X_n) = \frac{1}{n} \sum_{k=1}^{n} H(X_k|X_1, X_2, \cdots, X_{k-1}),$$

or

$$b_n = \frac{1}{n} \sum_{k=1}^{n} a_k.$$

5. Therefore, $b_n$, $n \geq 1$ are the Cesáro means of $\{a_n\}$.

6. From the definition of $H'_X$, we have $a_n \to H'_X$.

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$\begin{aligned} a_n &= H(X_n | X_1, X_2, \cdots, X_{n-1}) \\ b_n &= \frac{1}{n} H(X_1, X_2, \cdots, X_n). \end{aligned}$$

4. By the chain rule for entropy,

$$\begin{aligned} &\frac{1}{n} H(X_1, X_2, \cdots, X_n) \\ &\quad = \frac{1}{n} \sum_{k=1}^{n} H(X_k | X_1, X_2, \cdots, X_{k-1}), \end{aligned}$$

or

$$b_n = \frac{1}{n} \sum_{k=1}^{n} a_k.$$

5. Therefore, $b_n, n \geq 1$ are the Cesáro means of $\{a_n\}$.

6. From the definition of $H'_X$, we have $a_n \to H'_X$.

$$H'_X = \lim_{n \to \infty} \underbrace{H(X_n | X_1, X_2, \cdots, X_{n-1})}_{a_n}$$

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$a_n = H(X_n|X_1, X_2, \cdots, X_{n-1})$$
$$b_n = \frac{1}{n} H(X_1, X_2, \cdots, X_n).$$

4. By the chain rule for entropy,

$$\frac{1}{n} H(X_1, X_2, \cdots, X_n) = \frac{1}{n} \sum_{k=1}^{n} H(X_k|X_1, X_2, \cdots, X_{k-1}),$$

or

$$b_n = \frac{1}{n} \sum_{k=1}^{n} a_k.$$

5. Therefore, $b_n$, $n \geq 1$ are the Cesáro means of $\{a_n\}$.

6. From the definition of $H'_X$, we have $a_n \to H'_X$.

7. By Lemma 2.59, $b_n \to H'_X$, and so

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$\begin{aligned} a_n &= H(X_n|X_1, X_2, \cdots, X_{n-1}) \\ b_n &= \frac{1}{n} H(X_1, X_2, \cdots, X_n). \end{aligned}$$

4. By the chain rule for entropy,

$$\begin{aligned} &\frac{1}{n} H(X_1, X_2, \cdots, X_n) \\ &= \frac{1}{n} \sum_{k=1}^{n} H(X_k|X_1, X_2, \cdots, X_{k-1}), \end{aligned}$$

or

$$b_n = \frac{1}{n} \sum_{k=1}^{n} a_k.$$

5. Therefore, $b_n, n \geq 1$ are the Cesáro means of $\{a_n\}$.

6. From the definition of $H'_X$, we have $a_n \to H'_X$.

7. By Lemma 2.59, $b_n \to H'_X$, and so

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$\begin{aligned} a_n &= H(X_n | X_1, X_2, \cdots, X_{n-1}) \\ b_n &= \frac{1}{n} H(X_1, X_2, \cdots, X_n). \end{aligned}$$

4. By the chain rule for entropy,

$$\begin{aligned} &\frac{1}{n} H(X_1, X_2, \cdots, X_n) \\ &\quad = \frac{1}{n} \sum_{k=1}^{n} H(X_k | X_1, X_2, \cdots, X_{k-1}), \end{aligned}$$

or

$$b_n = \frac{1}{n} \sum_{k=1}^{n} a_k.$$

5. Therefore, $b_n, n \geq 1$ are the Cesáro means of $\{a_n\}$.

6. From the definition of $H'_X$, we have $a_n \to H'_X$.

7. By Lemma 2.59, $b_n \to H'_X$, and so

$$H'_X = \lim_{n \to \infty} H(X_n | X_1, X_2, \cdots, X_{n-1})$$

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$\begin{aligned} a_n &= H(X_n|X_1, X_2, \cdots, X_{n-1}) \\ b_n &= \frac{1}{n} H(X_1, X_2, \cdots, X_n). \end{aligned}$$

4. By the chain rule for entropy,

$$\begin{aligned} &\frac{1}{n} H(X_1, X_2, \cdots, X_n) \\ &\quad = \frac{1}{n} \sum_{k=1}^{n} H(X_k|X_1, X_2, \cdots, X_{k-1}), \end{aligned}$$

or

$$b_n = \frac{1}{n} \sum_{k=1}^{n} a_k.$$

5. Therefore, $b_n$, $n \geq 1$ are the Cesáro means of $\{a_n\}$.

6. From the definition of $H'_X$, we have $a_n \to H'_X$.

7. By Lemma 2.59, $b_n \to H'_X$, and so

$$\lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \cdots, X_n) = H'_X.$$

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$\begin{aligned} a_n &= H(X_n|X_1, X_2, \cdots, X_{n-1}) \\ b_n &= \frac{1}{n} H(X_1, X_2, \cdots, X_n). \end{aligned}$$

4. By the chain rule for entropy,

$$\begin{aligned} &\frac{1}{n} H(X_1, X_2, \cdots, X_n) \\ &\quad = \frac{1}{n} \sum_{k=1}^{n} H(X_k|X_1, X_2, \cdots, X_{k-1}), \end{aligned}$$

or

$$b_n = \frac{1}{n} \sum_{k=1}^{n} a_k.$$

5. Therefore, $b_n, n \geq 1$ are the Cesáro means of $\{a_n\}$.

6. From the definition of $H'_X$, we have $a_n \to H'_X$.

7. By Lemma 2.59, $b_n \to H'_X$, and so

$$H_X = \lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \cdots, X_n) = H'_X.$$

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

**Theorem 2.60** The entropy rate $H_X$ of a stationary source $\{X_k\}$ exists and is equal to $H'_X$.

**Proof**

1. We have proved in Lemma 2.58 that $H'_X$ always exists for a stationary source $\{X_k\}$.

2. In order to prove the theorem, we only have to prove that $H_X = H'_X$.

3. For $n \geq 1$, let

$$\begin{aligned} a_n &= H(X_n|X_1, X_2, \cdots, X_{n-1}) \\ b_n &= \frac{1}{n} H(X_1, X_2, \cdots, X_n). \end{aligned}$$

4. By the chain rule for entropy,

$$\begin{aligned} &\frac{1}{n} H(X_1, X_2, \cdots, X_n) \\ &\quad = \frac{1}{n} \sum_{k=1}^{n} H(X_k|X_1, X_2, \cdots, X_{k-1}), \end{aligned}$$

or

$$b_n = \frac{1}{n} \sum_{k=1}^{n} a_k.$$

5. Therefore, $b_n, n \geq 1$ are the Cesáro means of $\{a_n\}$.

6. From the definition of $H'_X$, we have $a_n \to H'_X$.

7. By Lemma 2.59, $b_n \to H'_X$, and so

$$H_X = \lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \cdots, X_n) = H'_X.$$

8. Hence, $H_X = H'_X$, as to be proved.

$$H'_X = \lim_{n\to\infty} H(X_n|X_1, X_2, \cdots, X_{n-1})$$

# Remarks

# Remarks

- Theorem 2.60 says that

# Remarks

- Theorem 2.60 says that

  1. the entropy rate of an information source $\{X_k\}$ exists under the fairly general assumption that $\{X_k\}$ is stationary;

# Remarks

- Theorem 2.60 says that

  1. the entropy rate of an information source $\{X_k\}$ exists under the fairly general assumption that $\{X_k\}$ is stationary;
  2. $H'_X$ is an alternative definition/interpretation of the entropy rate of $\{X_k\}$ when $\{X_k\}$ is stationary.

# Remarks

- Theorem 2.60 says that

  1. the entropy rate of an information source $\{X_k\}$ exists under the fairly general assumption that $\{X_k\}$ is stationary;
  2. $H'_X$ is an alternative definition/interpretation of the entropy rate of $\{X_k\}$ when $\{X_k\}$ is stationary.

- However, the entropy rate of a stationary source $\{X_k\}$ may not carry any physical meaning unless $\{X_k\}$ is also ergodic. See Section 5.4.