2.5 Informational Divergence
**Definition 2.28** The informational divergence between two probability distributions $p$ and $q$ on a common alphabet $\mathcal{X}$ is defined as

$$D(p\|q) = \sum_x p(x) \log \frac{p(x)}{q(x)} = E_p \log \frac{p(X)}{q(X)},$$

where $E_p$ denotes expectation with respect to $p$. 

...
**Definition 2.28** The informational divergence between two probability distributions $p$ and $q$ on a common alphabet $\mathcal{X}$ is defined as

$$D(p\|q) = \sum_x p(x) \log \frac{p(x)}{q(x)} = E_p \log \frac{p(X)}{q(X)},$$

where $E_p$ denotes expectation with respect to $p$.

- Convention:
**Definition 2.28** The informational divergence between two probability distributions $p$ and $q$ on a common alphabet $\mathcal{X}$ is defined as

$$D(p\|q) = \sum_x p(x) \log \frac{p(x)}{q(x)} = E_p \log \frac{p(X)}{q(X)},$$

where $E_p$ denotes expectation with respect to $p$.

- Convention:
  1. Summation is over $S_p$, i.e., $\sum_{x \in S_p}$
**Definition 2.28** The informational divergence between two probability distributions $p$ and $q$ on a common alphabet $\mathcal{X}$ is defined as

$$D(p\|q) = \sum_x p(x) \log \frac{p(x)}{q(x)} = E_p \log \frac{p(X)}{q(X)},$$

where $E_p$ denotes expectation with respect to $p$.

- Convention:
  1. Summation is over $S_p$, i.e., $\sum_{x \in S_p}$
Definition 2.28  The informational divergence between two probability distributions \( p \) and \( q \) on a common alphabet \( \mathcal{X} \) is defined as

\[
D(p\|q) = \sum_x p(x) \log \frac{p(x)}{q(x)} = E_p \log \frac{p(X)}{q(X)},
\]

where \( E_p \) denotes expectation with respect to \( p \).

- Convention:
  1. Summation is over \( S_p \), i.e., \( \sum_{x \in S_p} \).
**Definition 2.28** The informational divergence between two probability distributions \( p \) and \( q \) on a common alphabet \( \mathcal{X} \) is defined as

\[
D(p\|q) = \sum_x p(x) \log \frac{p(x)}{q(x)} = E_p \log \frac{p(X)}{q(X)},
\]

where \( E_p \) denotes expectation with respect to \( p \).

- Convention:
  1. Summation is over \( S_p \), i.e., \( \sum_{x \in S_p} \)
  2. \( c \log \frac{c}{0} = \infty \) for \( c > 0 \)
Definition 2.28  The informational divergence between two probability distributions $p$ and $q$ on a common alphabet $\mathcal{X}$ is defined as

$$D(p\|q) = \sum_x p(x) \log \frac{p(x)}{q(x)} = E_p \log \frac{p(X)}{q(X)},$$

where $E_p$ denotes expectation with respect to $p$.

- Convention:
  
  1. Summation is over $S_p$, i.e., $\sum_{x \in S_p}$
  2. $c \log \frac{c}{0} = \infty$ for $c > 0$
  3. If $D(p\|q) < \infty$, then $p(x) > 0 \Rightarrow q(x) > 0$, or $S_p \subset S_q$. 

• $D(p\|q)$ measures the "distance" between $p$ and $q$.
• $D(p\|q)$ is not symmetrical in $p$ and $q$, so $D(p\|q)$ is not a true metric.
• $D(p\|q)$ does not satisfy the triangular inequality.
• Also called relative entropy or the Kullback-Leibler distance.
**Definition 2.28** The informational divergence between two probability distributions $p$ and $q$ on a common alphabet $\mathcal{X}$ is defined as

$$D(p\|q) = \sum_x p(x) \log \frac{p(x)}{q(x)} = E_p \log \frac{p(X)}{q(X)},$$

where $E_p$ denotes expectation with respect to $p$.

- Convention:
  1. Summation is over $S_p$, i.e., $\sum_{x \in S_p}$
  2. $c \log \frac{c}{0} = \infty$ for $c > 0$
  3. If $D(p\|q) < \infty$, then $p(x) > 0 \Rightarrow q(x) > 0$, or $S_p \subset S_q$.

- $D(p\|q)$ measures the “distance” between $p$ and $q$. 
Definition 2.28  The informational divergence between two probability distributions \( p \) and \( q \) on a common alphabet \( \mathcal{X} \) is defined as

\[
D(p\|q) = \sum_x p(x) \log \frac{p(x)}{q(x)} = E_p \log \frac{p(X)}{q(X)},
\]

where \( E_p \) denotes expectation with respect to \( p \).

- Convention:
  1. Summation is over \( S_p \), i.e., \( \sum_{x \in S_p} \)
  2. \( c \log \frac{c}{0} = \infty \) for \( c > 0 \)
  3. If \( D(p\|q) < \infty \), then \( p(x) > 0 \Rightarrow q(x) > 0 \), or \( S_p \subset S_q \).
- \( D(p\|q) \) measures the “distance” between \( p \) and \( q \).
- \( D(p\|q) \) is not symmetrical in \( p \) and \( q \), so \( D(\cdot\|\cdot) \) is not a true metric.
**Definition 2.28** The informational divergence between two probability distributions $p$ and $q$ on a common alphabet $\mathcal{X}$ is defined as

$$D(p||q) = \sum_x p(x) \log \frac{p(x)}{q(x)} = E_p \log \frac{p(X)}{q(X)},$$

where $E_p$ denotes expectation with respect to $p$.

- **Convention:**
  1. Summation is over $\mathcal{S}_p$, i.e., $\sum_{x \in \mathcal{S}_p} p(x)$
  2. $c \log \frac{c}{0} = \infty$ for $c > 0$
  3. If $D(p||q) < \infty$, then $p(x) > 0 \Rightarrow q(x) > 0$, or $\mathcal{S}_p \subset \mathcal{S}_q.$

- $D(p||q)$ measures the “distance” between $p$ and $q$.

- $D(p||q)$ is not symmetrical in $p$ and $q$, so $D(\cdot||\cdot)$ is not a true metric.

- $D(\cdot||\cdot)$ does not satisfy the triangular inequality.
**Definition 2.28** The informational divergence between two probability distributions $p$ and $q$ on a common alphabet $\mathcal{X}$ is defined as

$$D(p\|q) = \sum_x p(x) \log \frac{p(x)}{q(x)} = E_p \log \frac{p(X)}{q(X)},$$

where $E_p$ denotes expectation with respect to $p$.

- Convention:
  1. Summation is over $S_p$, i.e., $\sum_{x \in S_p}$
  2. $c \log \frac{c}{0} = \infty$ for $c > 0$
  3. If $D(p\|q) < \infty$, then $p(x) > 0 \Rightarrow q(x) > 0$, or $S_p \subset S_q$.
- $D(p\|q)$ measures the “distance” between $p$ and $q$.
- $D(p\|q)$ is not symmetrical in $p$ and $q$, so $D(\cdot \| \cdot)$ is not a true metric.
- $D(\cdot \| \cdot)$ does not satisfy the triangular inequality.
- Also called relative entropy or the Kullback-Leibler distance.
Lemma 2.29 (Fundamental Inequality) For any $a > 0$,

$$\ln a \leq a - 1$$

with equality if and only if $a = 1$. 
\ln a \leq a - 1
Lemma 2.29 (Fundamental Inequality) For any $a > 0$, 

$$\ln a \leq a - 1$$

with equality if and only if $a = 1$.

Corollary 2.30 For any $a > 0$, 

$$\ln a \geq 1 - \frac{1}{a}$$

with equality if and only if $a = 1$. 
Lemma 2.29 (Fundamental Inequality) For any $a > 0$,

$$\ln a \leq a - 1$$

with equality if and only if $a = 1$.

Corollary 2.30 For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a}$$

with equality if and only if $a = 1$.

**Proof** Let $a = \frac{1}{b}$ in the fundamental inequality, where $b > 0$. Then
Lemma 2.29 (Fundamental Inequality) For any $a > 0$,

$$\ln a \leq a - 1$$

with equality if and only if $a = 1$.

Corollary 2.30 For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a}$$

with equality if and only if $a = 1$.

Proof Let $a = \frac{1}{b}$ in the fundamental inequality, where $b > 0$. Then
Lemma 2.29 (Fundamental Inequality) For any $a > 0$, 

\[ \ln a \leq a - 1 \]

with equality if and only if $a = 1$.

Corollary 2.30 For any $a > 0$, 

\[ \ln a \geq 1 - \frac{1}{a} \]

with equality if and only if $a = 1$.

Proof Let $a = \frac{1}{b}$ in the fundamental inequality, where $b > 0$. Then
Lemma 2.29 (Fundamental Inequality)  For any $a > 0$,

$$\ln a \leq a - 1$$

with equality if and only if $a = 1$.

**Corollary 2.30**  For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a}$$

with equality if and only if $a = 1$.

**Proof**  Let $a = \frac{1}{b}$ in the fundamental inequality, where $b > 0$. Then
Lemma 2.29 (Fundamental Inequality) For any $a > 0$,
\[
 \ln a \leq a - 1
\]
with equality if and only if $a = 1$.

Corollary 2.30 For any $a > 0$,
\[
 \ln a \geq 1 - \frac{1}{a}
\]
with equality if and only if $a = 1$.

Proof Let $a = \frac{1}{b}$ in the fundamental inequality, where $b > 0$. Then
Lemma 2.29 (Fundamental Inequality) For any $a > 0,$

$$\ln a \leq a - 1$$

with equality if and only if $a = 1.$

Corollary 2.30 For any $a > 0,$

$$\ln a \geq 1 - \frac{1}{a}$$

with equality if and only if $a = 1.$

Proof Let $a = \frac{1}{b}$ in the fundamental inequality, where $b > 0.$ Then

$$\ln \frac{1}{b} \leq \frac{1}{b} - 1$$
Lemma 2.29 (Fundamental Inequality) For any $a > 0$,

$$\ln a \leq a - 1$$

with equality if and only if $a = 1$.

Corollary 2.30 For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a}$$

with equality if and only if $a = 1$.

Proof Let $a = \frac{1}{b}$ in the fundamental inequality, where $b > 0$. Then

$$\ln \frac{1}{b} \leq \frac{1}{b} - 1$$

$$- \ln b \leq \frac{1}{b} - 1$$
Lemma 2.29 (Fundamental Inequality) For any $a > 0$,\[\ln a \leq a - 1\]
with equality if and only if $a = 1$.

Corollary 2.30 For any $a > 0$,\[\ln a \geq 1 - \frac{1}{a}\]
with equality if and only if $a = 1$.

Proof Let $a = \frac{1}{b}$ in the fundamental inequality, where $b > 0$. Then
\[
\ln \frac{1}{b} \leq \frac{1}{b} - 1 \\
- \ln b \leq \frac{1}{b} - 1 \\
\ln b \geq 1 - \frac{1}{b}
\]
Lemma 2.29 (Fundamental Inequality) For any $a > 0$,

$$\ln a \leq a - 1$$

with equality if and only if $a = 1$.

Corollary 2.30 For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a}$$

with equality if and only if $a = 1$.

Proof Let $a = \frac{1}{b}$ in the fundamental inequality, where $b > 0$. Then

$$\ln \frac{1}{b} \leq \frac{1}{b} - 1$$

$$-\ln b \leq \frac{1}{b} - 1$$

$$\ln b \geq 1 - \frac{1}{b}$$
Lemma 2.29 (Fundamental Inequality) For any $a > 0$,

$$\ln a \leq a - 1$$

with equality if and only if $a = 1$.

Corollary 2.30 For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a}$$

with equality if and only if $a = 1$.

Proof Let $a = \frac{1}{b}$ in the fundamental inequality, where $b > 0$. Then

$$\ln \frac{1}{b} \leq \frac{1}{b} - 1$$

$$-\ln b \leq \frac{1}{b} - 1$$

$$\ln b \geq 1 - \frac{1}{b}$$

Equality holds if and only if $\frac{1}{b} = a = 1$, or $b = 1$. 
Lemma 2.29 (Fundamental Inequality)  For any $a > 0$,

$$\ln a \leq a - 1$$

with equality if and only if $a = 1$.

Corollary 2.30  For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a}$$

with equality if and only if $a = 1$.

Proof  Let $a = \frac{1}{b}$ in the fundamental inequality, where $b > 0$. Then

$$\ln \frac{1}{b} \leq \frac{1}{b} - 1$$

$$-\ln b \leq \frac{1}{b} - 1$$

$$\ln b \geq 1 - \frac{1}{b}$$

Equality holds if and only if $\frac{1}{b} = a = 1$, or $b = 1$. 
Lemma 2.29 (Fundamental Inequality) For any $a > 0$, 

$$\ln a \leq a - 1$$

with equality if and only if $a = 1$.

Corollary 2.30 For any $a > 0$, 

$$\ln a \geq 1 - \frac{1}{a}$$

with equality if and only if $a = 1$.

Proof Let $a = \frac{1}{b}$ in the fundamental inequality, where $b > 0$. Then 

$$\ln \frac{1}{b} \leq \frac{1}{b} - 1$$

$$-\ln b \leq \frac{1}{b} - 1$$

$$\ln b \geq 1 - \frac{1}{b}$$

Equality holds if and only if $\frac{1}{b} = a = 1$, or $b = 1$. 
Lemma 2.29 (Fundamental Inequality) For any $a > 0$, 
\[ \ln a \leq a - 1 \]
with equality if and only if $a = 1$.

Corollary 2.30 For any $a > 0$, 
\[ \ln a \geq 1 - \frac{1}{a} \]
with equality if and only if $a = 1$.

Proof Let $a = \frac{1}{b}$ in the fundamental inequality, where $b > 0$. Then 
\[ \ln \frac{1}{b} \leq \frac{1}{b} - 1 \]
\[ -\ln b \leq \frac{1}{b} - 1 \]
\[ \ln b \geq 1 - \frac{1}{b} \]
Equality holds if and only if $\frac{1}{b} = a = 1$, or $b = 1$. 
Theorem 2.31 (Divergence Inequality) For any two probability distributions $p$ and $q$ on a common alphabet $\mathcal{X}$, 

$$D(p\|q) \geq 0$$

with equality if and only if $p = q$. 
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \quad (1) \]

with equality if and only if \( p = q \).

Proof
Theorem 2.31 (Divergence Inequality)

\[ D(p||q) \geq 0 \quad (1) \]

with equality if and only if \( p = q \).

**Proof**

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \]  \hspace{1cm} (1)

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( \mathcal{S}_p = \mathcal{S}_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in \mathcal{S}_p} p(x) \log \frac{p(x)}{q(x)}
\]
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \]  

(1)

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( \mathcal{S}_p = \mathcal{S}_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in \mathcal{S}_p} p(x) \log \frac{p(x)}{q(x)}
\]
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \]  \hspace{1cm} (1)

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)}
\]

3. For equality to hold in (2), we see from Corollary 2.30 that this is the case if and only if

\[
p(x) = q(x) \text{ for all } x \in S_p.
\]

This proves the theorem.

Corollary 2.30

For any \( a > 0 \),

\[
\ln a \leq 1 \quad \text{with equality if and only if } a = 1.
\]
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \]  \hspace{1cm} (1)

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)}
\]

Corollary 2.30
For any \( a > 0 \),

\[ \ln a \geq 1 \]  \hspace{1cm} (2)

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \]  

(1)

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)}
\]
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \] (1)

with equality if and only if \( p = q \).

\textbf{Proof}

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)}
\]
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \]  

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)} = (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)}
\]

Corollary 2.30 For any \( a > 0 \),

\[
\ln a \geq 1 - \frac{1}{a}
\]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \]  \hspace{1cm} (1)

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)}
\]

\[
\geq (\log e) \sum_{x \in S_p} p(x) \left(1 - \frac{q(x)}{p(x)}\right)
\]  \hspace{1cm} (2)

Corollary 2.30 For any \( a > 0 \),

\[ \ln a \geq 1 - \frac{1}{a} \]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \]  \hspace{1cm} (1)

with equality if and only if \( p = q \).

**Proof**

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= \left( \log e \right) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)}
\]

\[
\geq \left( \log e \right) \sum_{x \in S_p} p(x) \left( 1 - \frac{q(x)}{p(x)} \right)
\]  \hspace{1cm} (2)

**Corollary 2.30** For any \( a > 0 \),

\[ \ln a \geq 1 - \frac{1}{a} \]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \quad (1) \]

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)}
\]

\[
\geq (\log e) \sum_{x \in S_p} p(x) \left( 1 - \frac{q(x)}{p(x)} \right) \quad (2)
\]

Corollary 2.30 For any \( a > 0 \),

\[ \ln a \geq 1 - \frac{1}{a} \]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \]  \hspace{1cm} (1)

with equality if and only if \( p = q \).

**Proof**

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \left( \frac{p(x)}{q(x)} \right)
\]

   \[ = (\log e) \sum_{x \in S_p} p(x) \ln \left( \frac{p(x)}{q(x)} \right) \]

   \[ \geq (\log e) \sum_{x \in S_p} p(x) \left( 1 - \frac{q(x)}{p(x)} \right) \]

   \hspace{1cm} (2)

---

**Corollary 2.30** For any \( a > 0 \),

\[ \ln a \geq 1 - \frac{1}{a} \]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \quad (1) \]

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)}
\]

\[
\geq (\log e) \sum_{x \in S_p} p(x) \left( 1 - \frac{q(x)}{p(x)} \right) \quad (2)
\]

\[
= (\log e) \left[ \sum_{x \in S_p} p(x) - \sum_{x \in S_p} q(x) \right]
\]

Corollary 2.30 For any \( a > 0 \),

\[
\ln a \geq 1 - \frac{1}{a}
\]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \quad (1) \]

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( \mathcal{S}_p = \mathcal{S}_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in \mathcal{S}_p} p(x) \log \frac{p(x)}{q(x)} \\
= (\log e) \sum_{x \in \mathcal{S}_p} p(x) \ln \frac{p(x)}{q(x)} \\
\geq (\log e) \sum_{x \in \mathcal{S}_p} p(x) \left( 1 - \frac{q(x)}{p(x)} \right) \quad (2) \\
= (\log e) \left[ \sum_{x \in \mathcal{S}_p} p(x) - \sum_{x \in \mathcal{S}_p} q(x) \right]
\]

Corollary 2.30 For any \( a > 0 \),

\[ \ln a \geq 1 - \frac{1}{a} \]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \quad (1) \]

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)}
\]

\[
\geq (\log e) \sum_{x \in S_p} p(x) \left( 1 - \frac{q(x)}{p(x)} \right) \quad (2)
\]

\[
= (\log e) \left[ \sum_{x \in S_p} p(x) - \sum_{x \in S_p} q(x) \right]
\]

Corollary 2.30 For any \( a > 0 \),

\[
\ln a \geq 1 - \frac{1}{a}
\]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \quad (1) \]

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)}
\]

\[
\geq (\log e) \sum_{x \in S_p} p(x) \left( 1 - \frac{q(x)}{p(x)} \right) \quad (2)
\]

\[
= (\log e) \left[ \sum_{x \in S_p} p(x) - \sum_{x \in S_p} q(x) \right]
\]

\[
= (\log e) \left[ \sum_{x \in S_p} p(x) - \sum_{x \in S_q} q(x) \right]
\]

Corollary 2.30 For any \( a > 0 \),

\[ \ln a \geq 1 - \frac{1}{a} \]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \]  \hspace{1cm} (1)

with equality if and only if \( p = q \).

**Proof**

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)}
\]

\[
\geq (\log e) \sum_{x \in S_p} p(x) \left(1 - \frac{q(x)}{p(x)}\right) \quad (2)
\]

\[
= (\log e) \left[ \sum_{x \in S_p} p(x) - \sum_{x \in S_p} q(x) \right]
\]

\[
= (\log e) \left[ \sum_{x \in S_p} p(x) - \sum_{x \in S_q} q(x) \right]
\]

---

**Corollary 2.30** For any \( a > 0 \),

\[
\ln a \geq 1 - \frac{1}{a}
\]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[
D(p\|q) \geq 0 \tag{1}
\]

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( \mathcal{S}_p = \mathcal{S}_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in \mathcal{S}_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in \mathcal{S}_p} p(x) \ln \frac{p(x)}{q(x)}
\]

\[
\geq (\log e) \sum_{x \in \mathcal{S}_p} \left( 1 - \frac{q(x)}{p(x)} \right) \tag{2}
\]

\[
= (\log e) \left[ \sum_{x \in \mathcal{S}_p} p(x) - \sum_{x \in \mathcal{S}_p} q(x) \right]
\]

\[
= (\log e) \left[ \sum_{x \in \mathcal{S}_p} p(x) - \sum_{x \in \mathcal{S}_q} q(x) \right]
\]

\[
= (\log e) [1 - 1]
\]

Corollary 2.30 For any \( a > 0 \),

\[
\ln a \geq 1 - \frac{1}{a}
\]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \tag{1} \]

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)}
\]

\[
\geq (\log e) \sum_{x \in S_p} p(x) \left( 1 - \frac{q(x)}{p(x)} \right) \tag{2}
\]

\[
= (\log e) \left[ \sum_{x \in S_p} p(x) - \sum_{x \in S_p} q(x) \right]
\]

\[
= (\log e) \left[ \sum_{x \in S_p} p(x) - \sum_{x \in S_q} q(x) \right]
\]

\[
= (\log e)[1 - 1]
\]

\[
= 0.
\]

Corollary 2.30 For any \( a > 0 \),

\[ \ln a \geq 1 - \frac{1}{a} \]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \]  

(1)

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( \mathcal{S}_p = \mathcal{S}_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in \mathcal{S}_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in \mathcal{S}_p} p(x) \ln \frac{p(x)}{q(x)}
\]

\[
\geq (\log e) \sum_{x \in \mathcal{S}_p} p(x) \left( 1 - \frac{q(x)}{p(x)} \right)
\]

(2)

\[
= (\log e) \left[ \sum_{x \in \mathcal{S}_p} p(x) - \sum_{x \in \mathcal{S}_p} q(x) \right]
\]

\[
= (\log e) \left[ \sum_{x \in \mathcal{S}_p} p(x) - \sum_{x \in \mathcal{S}_q} q(x) \right]
\]

\[
= (\log e)[1 - 1]
\]

\[
= 0.
\]

This proves (1).

Corollary 2.30 For any \( a > 0 \),

\[
\ln a \geq 1 - \frac{1}{a}
\]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \quad (1) \]

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)}
\]

\[
\geq (\log e) \sum_{x \in S_p} p(x) \left( 1 - \frac{q(x)}{p(x)} \right) \quad (2)
\]

\[
= (\log e) \left[ \sum_{x \in S_p} p(x) - \sum_{x \in S_p} q(x) \right]
\]

\[
= (\log e) \left[ \sum_{x \in S_p} p(x) - \sum_{x \in S_q} q(x) \right]
\]

\[
= (\log e)[1 - 1]
\]

\[
= 0.
\]

This proves (1).

Corollary 2.30 For any \( a > 0 \),

\[
\ln a \geq 1 - \frac{1}{a}
\]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \]  
(1)

with equality if and only if \( p = q \).

**Proof**

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)}
\]

\[
\geq (\log e) \sum_{x \in S_p} p(x) \left( 1 - \frac{q(x)}{p(x)} \right) \quad (2)
\]

\[
= (\log e) \left[ \sum_{x \in S_p} p(x) - \sum_{x \in S_p} q(x) \right]
\]

\[
= (\log e) \left[ \sum_{x \in S_p} p(x) - \sum_{x \in S_q} q(x) \right]
\]

\[
= (\log e)[1 - 1]
\]

\[
= 0.
\]

This proves (1).

3. For equality to hold in (2), we see from Corollary 2.30 that this is the case if and only if

\[
\frac{p(x)}{q(x)} = 1 \text{ or } p(x) = q(x) \quad \text{for all } x \in S_p.
\]

This proves the theorem.

**Corollary 2.30** For any \( a > 0 \),

\[
\ln a \geq 1 - \frac{1}{a}
\]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \]  \hspace{1cm} (1)

with equality if and only if \( p = q \).

**Proof**

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[= (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)} \]

\[\geq (\log e) \sum_{x \in S_p} p(x) \left( 1 - \frac{q(x)}{p(x)} \right) \]  \hspace{1cm} (2)

\[
= (\log e) \left[ \sum_{x \in S_p} p(x) - \sum_{x \in S_p} q(x) \right]
\]

\[= (\log e) \left[ \sum_{x \in S_p} p(x) - \sum_{x \in S_q} q(x) \right]
\]

\[= (\log e)[1 - 1]
\]

\[= 0. \]

This proves (1).

3. For equality to hold in (2), we see from Corollary 2.30 that this is the case if and only if

\[
\frac{p(x)}{q(x)} = 1 \text{ or } p(x) = q(x) \quad \text{for all } x \in S_p.
\]

This proves the theorem.

**Corollary 2.30** For any \( a > 0 \),

\[ \ln a \geq 1 - \frac{1}{a} \]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \]  \hspace{1cm} (1)

with equality if and only if \( p = q \).

**Proof**

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)}
\]

\[
\geq (\log e) \sum_{x \in S_p} p(x) \left( 1 - \frac{q(x)}{p(x)} \right) \tag{2}
\]

\[
= (\log e) \left[ \sum_{x \in S_p} p(x) - \sum_{x \in S_p} q(x) \right]
\]

\[
= (\log e) \left[ \sum_{x \in S_p} p(x) - \sum_{x \in S_q} q(x) \right]
\]

\[
= (\log e)[1 - 1]
\]

\[
= 0.
\]

This proves (1).

3. For equality to hold in (2), we see from Corollary 2.30 that this is the case if and only if

\[
\frac{p(x)}{q(x)} = 1 \text{ or } p(x) = q(x) \text{ for all } x \in S_p.
\]

This proves the theorem.

**Corollary 2.30** For any \( a > 0 \),

\[
\ln a \geq 1 - \frac{1}{a}
\]

with equality if and only if \( a = 1 \).
Theorem 2.31 (Divergence Inequality)

\[ D(p\|q) \geq 0 \quad (1) \]

with equality if and only if \( p = q \).

Proof

1. For simplicity, assume \( S_p = S_q \). For a proof without this assumption, see the textbook.

2. Consider

\[
D(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}
\]

\[
= (\log e) \sum_{x \in S_p} p(x) \ln \frac{p(x)}{q(x)} \geq (\log e) \sum_{x \in S_P} p(x) \left( 1 - \frac{q(x)}{p(x)} \right) \quad (2)
\]

\[
= (\log e) \left[ \sum_{x \in S_P} p(x) - \sum_{x \in S_P} q(x) \right]
\]

\[
= (\log e) \left[ \sum_{x \in S_P} p(x) - \sum_{x \in S_q} q(x) \right]
\]

\[
= (\log e)[1 - 1]
\]

\[
= 0.
\]

This proves (1).

3. For equality to hold in (2), we see from Corollary 2.30 that this is the case if and only if

\[
\frac{p(x)}{q(x)} = 1 \text{ or } p(x) = q(x) \quad \text{for all } x \in S_p.
\]

This proves the theorem.

Corollary 2.30 For any \( a > 0 \),

\[
\ln a \geq 1 - \frac{1}{a}
\]

with equality if and only if \( a = 1 \).
Theorem 2.32 (Log-Sum Inequality)  For positive numbers $a_1, a_2, \cdots$ and nonnegative numbers $b_1, b_2, \cdots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i}$$

with the convention that $\log \frac{a_i}{0} = \infty$. Moreover, equality holds if and only if $\frac{a_i}{b_i} = constant$ for all $i$. 
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \cdots$ and nonnegative numbers $b_1, b_2, \cdots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i}$$

with the convention that $\log \frac{a_i}{0} = \infty$. Moreover, equality holds if and only if $\frac{a_i}{b_i} = constant$ for all $i$.

Example:

$$a_1 \log \frac{a_1}{b_1} + a_2 \log \frac{a_2}{b_2} \geq (a_1 + a_2) \log \frac{a_1 + a_2}{b_1 + b_2}.$$
**Theorem 2.32 (Log-Sum Inequality)** For positive numbers $a_1, a_2, \cdots$ and nonnegative numbers $b_1, b_2, \cdots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$
\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i}
$$

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

**Proof**
**Theorem 2.32 (Log-Sum Inequality)** For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$, \[ \sum_i a_i \log \frac{a_i}{b_i} \geq \left(\sum_i a_i\right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1) \] Moreover, equality holds if and only if $\frac{a_i}{b_i}$ = constant for all $i$.

**Proof**
1. Let $a'_i = a_i / \sum_j a_j$ and $b'_i = b_i / \sum_j b_j$. Then $\{a'_i\}$ and $\{b'_i\}$ are probability distributions.
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \cdots$ and nonnegative numbers $b_1, b_2, \cdots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)$$

Moreover, equality holds if and only if $\frac{a_i}{b_i} = constant$ for all $i$.

**Proof**

1. Let $a'_i = a_i / \sum_j a_j$ and $b'_i = b_i / \sum_j b_j$. Then \{a'_i\} and \{b'_i\} are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i}$$
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$
\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i}
$$

(1)

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

Proof

1. Let $a_i' = a_i / \sum_j a_j$ and $b_i' = b_i / \sum_j b_j$. Then $\{a_i'\}$ and $\{b_i'\}$ are probability distributions.

2. Using the divergence inequality, we have

$$
0 \leq \sum_i a_i' \log \frac{a_i'}{b_i'}
$$
**Theorem 2.32 (Log-Sum Inequality)**  For positive numbers $a_1, a_2, \cdots$ and nonnegative numbers $b_1, b_2, \cdots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)$$

Moreover, equality holds if and only if $\frac{a_i}{b_i} = constant$ for all $i$.

**Proof**

1. Let $a'_i = a_i / \sum_j a_j$ and $b'_i = b_i / \sum_j b_j$. Then $\{a'_i\}$ and $\{b'_i\}$ are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i}$$
Theorem 2.32 (Log-Sum Inequality) For positive numbers \( a_1, a_2, \ldots \) and nonnegative numbers \( b_1, b_2, \ldots \) such that \( \sum_i a_i < \infty \) and \( 0 < \sum_i b_i < \infty \),

\[
\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \tag{1}
\]

Moreover, equality holds if and only if \( \frac{a_i}{b_i} = \text{constant} \) for all \( i \).

Proof
1. Let \( a_i' = \frac{a_i}{\sum_j a_j} \) and \( b_i' = \frac{b_i}{\sum_j b_j} \). Then \( \{a_i'\} \) and \( \{b_i'\} \) are probability distributions.

2. Using the divergence inequality, we have

\[
0 \leq \sum_i a_i' \log \frac{a_i'}{b_i'} = \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i/\sum_j a_j}{b_i/\sum_j b_j}
\]
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)$$

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

Proof

1. Let $a'_i = \frac{a_i}{\sum_j a_j}$ and $b'_i = \frac{b_i}{\sum_j b_j}$. Then $\{a'_i\}$ and $\{b'_i\}$ are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i}$$

$$= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i/\sum_j a_j}{b_i/\sum_j b_j}$$
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)$$

Moreover, equality holds if and only if $\frac{a_i}{b_i}$ is constant for all $i$.

Proof

1. Let $a'_i = \frac{a_i}{\sum_j a_j}$ and $b'_i = \frac{b_i}{\sum_j b_j}$. Then \{a'\} and \{b'\} are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i}$$

$$= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i/\sum_j a_j}{b_i/\sum_j b_j}$$

The theorem is proved.
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \cdots$ and nonnegative numbers $b_1, b_2, \cdots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)$$

Moreover, equality holds if and only if $a_i b_i = \text{constant}$ for all $i$.

Proof

1. Let $a'_i = \frac{a_i}{\sum_j a_j}$ and $b'_i = \frac{b_i}{\sum_j b_j}$. Then $\{a'_i\}$ and $\{b'_i\}$ are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i}$$

$$= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i/\sum_j a_j}{b_i/\sum_j b_j}$$
Theorem 2.32 (Log-Sum Inequality) For positive numbers \(a_1, a_2, \cdots\) and nonnegative numbers \(b_1, b_2, \cdots\) such that \(\sum_i a_i < \infty\) and \(0 < \sum_i b_i < \infty\),

\[
\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \tag{1}
\]

Moreover, equality holds if and only if \(\frac{a_i}{b_i} = \text{constant}\) for all \(i\).

Proof

1. Let \(a_i' = \frac{a_i}{\sum_j a_j}\) and \(b_i' = \frac{b_i}{\sum_j b_j}\). Then \(\{a_i'\}\) and \(\{b_i'\}\) are probability distributions.

2. Using the divergence inequality, we have

\[
0 \leq \sum_i a_i' \log \frac{a_i'}{b_i'} = \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j}
\]
Theorem 2.32 (Log-Sum Inequality) For positive numbers \( a_1, a_2, \ldots \) and nonnegative numbers \( b_1, b_2, \ldots \) such that \( \sum_i a_i < \infty \) and \( 0 < \sum_i b_i < \infty \),

\[
\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \tag{1}
\]

Moreover, equality holds if and only if \( \frac{a_i}{b_i} = \text{constant} \) for all \( i \).

Proof
1. Let \( a'_i = \frac{a_i}{\sum_j a_j} \) and \( b'_i = \frac{b_i}{\sum_j b_j} \). Then \( \{a'_i\} \) and \( \{b'_i\} \) are probability distributions.

2. Using the divergence inequality, we have

\[
0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i} = \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j}
\]
**Theorem 2.32 (Log-Sum Inequality)** For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$, \[
abla_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \tag{1}
\]
Moreover, equality holds if and only if $\frac{a_i}{b_i}$ = constant for all $i$.

**Proof**

1. Let $a_i' = \frac{a_i}{\sum_j a_j}$ and $b_i' = \frac{b_i}{\sum_j b_j}$. Then \{a_i'\} and \{b_i'\} are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a_i' \log \frac{a_i'}{b_i'} = \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i/\sum_j a_j}{b_i/\sum_j b_j}$$
Theorem 2.32 (Log-Sum Inequality) For positive numbers \( a_1, a_2, \ldots \) and nonnegative numbers \( b_1, b_2, \ldots \) such that \( \sum_i a_i < \infty \) and \( 0 < \sum_i b_i < \infty \),

\[
\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)
\]

Moreover, equality holds if and only if \( \frac{a_i}{b_i} = \text{constant} \) for all \( i \).

Proof
1. Let \( a'_i = \frac{a_i}{\sum_j a_j} \) and \( b'_i = \frac{b_i}{\sum_j b_j} \). Then \( \{a'_i\} \) and \( \{b'_i\} \) are probability distributions.

2. Using the divergence inequality, we have

\[
0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i} = \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i}{\sum_j a_j} / \frac{b_i}{\sum_j b_j}
\]
**Theorem 2.32 (Log-Sum Inequality)** For positive numbers $a_1, a_2, \cdots$ and nonnegative numbers $b_1, b_2, \cdots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)$$

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

**Proof**

1. Let $a'_i = \frac{a_i}{\sum_j a_j}$ and $b'_i = \frac{b_i}{\sum_j b_j}$. Then $\{a'_i\}$ and $\{b'_i\}$ are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i} \leq \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j}$$
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)$$

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

Proof
1. Let $a'_i = \frac{a_i}{\sum_j a_j}$ and $b'_i = \frac{b_i}{\sum_j b_j}$. Then $\{a'_i\}$ and $\{b'_i\}$ are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i}$$

$$= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i}{\sum_j a_j} \frac{\sum_j a_j}{\sum_j a_i} \frac{a_i}{\sum_j b_j} \frac{b_i}{\sum_j b_i}$$
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$
\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i}
$$

(1)

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

Proof

1. Let $a'_i = \frac{a_i}{\sum_j a_j}$ and $b'_i = \frac{b_i}{\sum_j b_j}$. Then \{a'_i\} and \{b'_i\} are probability distributions.

2. Using the divergence inequality, we have

$$
0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i} = \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j}
$$
**Theorem 2.32 (Log-Sum Inequality)**  For positive numbers $a_1, a_2, \cdots$ and nonnegative numbers $b_1, b_2, \cdots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i}$$  \hspace{1cm} (1)

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

**Proof**

1. Let $a_i' = \frac{a_i}{\sum_j a_j}$ and $b_i' = \frac{b_i}{\sum_j b_j}$. Then $\{a_i'\}$ and $\{b_i'\}$ are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a_i' \log \frac{a_i'}{b_i'}$$

$$= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i}{\sum_j a_j}$$
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \cdots$ and nonnegative numbers $b_1, b_2, \cdots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)$$

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

Proof

1. Let $a'_i = \frac{a_i}{\sum_j a_j}$ and $b'_i = \frac{b_i}{\sum_j b_j}$. Then $\{a'_i\}$ and $\{b'_i\}$ are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i}$$

$$= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i}{\sum_j a_j} \frac{a_i}{\sum_j a_j} \frac{b_i}{\sum_j b_j}$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{\sum_j a_j} \frac{b_i}{\sum_j b_j} \right]$$
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)$$

Moreover, equality holds if and only if $\frac{a_i}{b_i}$ = constant for all $i$.

Proof
1. Let $a'_i = \frac{a_i}{\sum_j a_j}$ and $b'_i = \frac{b_i}{\sum_j b_j}$. Then $\{a'_i\}$ and $\{b'_i\}$ are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i} = \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j} = \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j} \right]$$
Theorem 2.32 (Log-Sum Inequality) For positive numbers \(a_1, a_2, \ldots\) and nonnegative numbers \(b_1, b_2, \ldots\) such that \(\sum_i a_i < \infty\) and \(0 < \sum_i b_i < \infty\),

\[
\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i}
\]

(1)

Moreover, equality holds if and only if \(\frac{a_i}{b_i} = \text{constant}\) for all \(i\).

Proof
1. Let \(a_i' = \frac{a_i}{\sum_j a_j}\) and \(b_i' = \frac{b_i}{\sum_j b_j}\). Then \(\{a_i'\}\) and \(\{b_i'\}\) are probability distributions.

2. Using the divergence inequality, we have

\[
0 \leq \sum_i a_i' \log \frac{a_i'}{b_i'}
\]

\[
= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i}{\sum_j a_j} \frac{1}{b_i} \frac{1}{\sum_j b_j}
\]

\[
= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{\sum_j a_j} \right]
\]

The theorem is proved.
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)$$

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

Proof
1. Let $a_i' = \frac{a_i}{\sum_j a_j}$ and $b_i' = \frac{b_i}{\sum_j b_j}$. Then $\{a_i'\}$ and $\{b_i'\}$ are probability distributions.
2. Using the divergence inequality, we have

$$0 \leq \sum_i a_i' \log \frac{a_i'}{b_i'}$$

$$= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i/\sum_j a_j}{b_i/\sum_j b_j}$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i/\sum_j a_j}{b_i/\sum_j b_j} \right]$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \sum_i a_i \log \frac{\sum_j a_j}{\sum_j b_j} \right]$$
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)$$

Moreover, equality holds if and only if $\frac{a_i}{b_i}$ = constant for all $i$.

Proof
1. Let $a'_i = \frac{a_i}{\sum_j a_j}$ and $b'_i = \frac{b_i}{\sum_j b_j}$. Then $\{a'_i\}$ and $\{b'_i\}$ are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i}$$

$$= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j}$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i \frac{a_i}{b_i} \log \left( \frac{a_i}{\sum_j a_j} \right) \right]$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \sum_i a_i \log \frac{\sum_j a_j}{\sum_j b_j} \right]$$
Theorem 2.32 (Log-Sum Inequality)  For positive numbers $a_1, a_2, \cdots$ and nonnegative numbers $b_1, b_2, \cdots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)$$

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

Proof
1. Let $a_i' = \frac{a_i}{\sum_j a_j}$ and $b_i' = \frac{b_i}{\sum_j b_j}$. Then $\{a_i'\}$ and $\{b_i'\}$ are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a_i' \log \frac{a_i'}{b_i'}$$

$$= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i/\sum_j a_j}{b_i/\sum_j b_i}$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i \frac{a_i}{b_i} \log \frac{a_i/\sum_j a_j}{\sum_j b_j} \right]$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i \frac{a_i}{b_i} - \sum_i a_i \log \frac{\sum_j a_j}{\sum_j b_j} \right]$$
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$
\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \tag{1}
$$

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

Proof

1. Let $a_i' = \frac{a_i}{\sum_j a_j}$ and $b_i' = \frac{b_i}{\sum_j b_j}$. Then $\{a_i'\}$ and $\{b_i'\}$ are probability distributions.

2. Using the divergence inequality, we have

\[
0 \leq \sum_i a_i' \log \frac{a_i'}{b_i'} = \sum_i a_i \log \left( \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j} \right) = \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j} \right] = \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \sum_i a_i \log \frac{\sum_j a_j}{\sum_j b_j} \right]
\]
Theorem 2.32 (Log-Sum Inequality) For positive numbers \(a_1, a_2, \ldots\) and nonnegative numbers \(b_1, b_2, \ldots\) such that \(\sum_i a_i < \infty\) and \(0 < \sum_i b_i < \infty\),

\[
\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \tag{1}
\]

Moreover, equality holds if and only if \(\frac{a_i}{b_i} = \text{constant}\) for all \(i\).

Proof
1. Let \(a'_i = a_i / \sum_j a_j \) and \(b'_i = b_i / \sum_j b_j\). Then \(\{a'_i\}\) and \(\{b'_i\}\) are probability distributions.
2. Using the divergence inequality, we have

\[
0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i}
= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j}
= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i / \sum_j b_j} \right]
= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \sum_i a_i \log \frac{\sum_j a_j}{\sum_j b_j} \right]
\]
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)$$

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

Proof

1. Let $a_i' = a_i / \sum_j a_j$ and $b_i' = b_i / \sum_j b_j$. Then \{a_i'\} and \{b_i'\} are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a_i' \log \frac{a_i'}{b_i'}$$

$$= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j}$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \sum_i a_i \log \frac{\sum_j a_j}{\sum_j b_j} \right]$$

The theorem is proved.
Theorem 2.32 (Log-Sum Inequality) For positive numbers \(a_1, a_2, \ldots\) and nonnegative numbers \(b_1, b_2, \ldots\) such that \(\sum_i a_i < \infty\) and \(0 < \sum_i b_i < \infty\),

\[
\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i}
\]

Moreover, equality holds if and only if \(\frac{a_i}{b_i} = \text{constant}\) for all \(i\).

Proof

1. Let \(a'_i = a_i / \sum_j a_j\) and \(b'_i = b_i / \sum_j b_j\). Then \(\{a'_i\}\) and \(\{b'_i\}\) are probability distributions.

2. Using the divergence inequality, we have

\[
0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i}
\]

\[
= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j}
\]

\[
= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j} \right]
\]

\[
= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \sum_i a_i \log \frac{\sum_j a_j}{\sum_j b_j} \right]
\]

\[
= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \left( \sum_i a_i \right) \log \frac{\sum_j a_j}{\sum_j b_j} \right],
\]
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)$$

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

Proof
1. Let $a'_i = a_i / \sum_j a_j$ and $b'_i = b_i / \sum_j b_j$. Then \{a'_i\} and \{b'_i\} are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i}$$

$$= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j}$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j} \right]$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \sum_i a_i \log \frac{\sum_j a_j}{\sum_j b_j} \right]$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \left( \sum_i a_i \right) \log \frac{\sum_j a_j}{\sum_j b_j} \right],$$

The theorem is proved.
Theorem 2.32 (Log-Sum Inequality)  For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$
\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \tag{1}
$$

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

**Proof**

1. Let $a'_i = \frac{a_i}{\sum_j a_j}$ and $b'_i = \frac{b_i}{\sum_j b_j}$. Then \{a'_i\} and \{b'_i\} are probability distributions.

2. Using the divergence inequality, we have

\[
0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i} \\
= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i/\sum_j a_j}{b_i/\sum_j b_j} \\
= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \sum_i a_i \log \frac{\sum_j a_j}{\sum_j b_j} \right] \\
= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \left( \sum_i a_i \right) \log \frac{\sum_j a_j}{\sum_j b_j} \right] \\
= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \left( \sum_i a_i \right) \log \frac{\sum_j a_j}{\sum_j b_j} \right],
\]
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

\[
\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \tag{1}
\]

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

**Proof**

1. Let $a'_i = \frac{a_i}{\sum_j a_j}$ and $b'_i = \frac{b_i}{\sum_j b_j}$. Then \{a'_i\} and \{b'_i\} are probability distributions.

2. Using the divergence inequality, we have

\[
0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i} = \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i/\sum_j a_j}{b_i/\sum_j b_j} = \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \sum_i a_i \log \frac{\sum_j a_j}{\sum_j b_j} \right] = \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \left( \sum_i a_i \right) \log \frac{\sum_j a_j}{\sum_j b_j} \right],
\]

which implies (1).
Theorem 2.32 (Log-Sum Inequality) For positive numbers \(a_1, a_2, \ldots\) and nonnegative numbers \(b_1, b_2, \ldots\) such that \(\sum_i a_i < \infty\) and \(0 < \sum_i b_i < \infty\),

\[
\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \tag{1}
\]

Moreover, equality holds if and only if \(\frac{a_i}{b_i} = \text{constant}\) for all \(i\).

Proof
1. Let \(a'_i = a_i / \sum_j a_j\) and \(b'_i = b_i / \sum_j b_j\). Then \(\{a'_i\}\) and \(\{b'_i\}\) are probability distributions.

2. Using the divergence inequality, we have

\[
0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i}
= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j}
= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j} \right]
= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \sum_i a_i \log \frac{\sum_j a_j}{\sum_j b_j} \right]
= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \left( \sum_i a_i \right) \log \frac{\sum_j a_j}{\sum_j b_j} \right],
\]

which implies (1).
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \tag{1}$$

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

**Proof**

1. Let $a'_i = a_i / \sum_j a_j$ and $b'_i = b_i / \sum_j b_j$. Then $\{a'_i\}$ and $\{b'_i\}$ are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i}$$

$$= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i/ \sum_j a_j}{b_i/ \sum_j b_j}$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \sum_i a_i \log \frac{\sum_j a_j}{\sum_j b_j} \right]$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \left( \sum_i a_i \right) \log \frac{\sum_j a_j}{\sum_j b_j} \right],$$

which implies (1).

3. Equality holds if and only if for all $i$,

$$a'_i = b'_i \text{ or } \frac{a_i}{b_i} = \text{constant}.$$
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$, the inequality

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)$$

Moreover, equality holds if and only if $a_i = constant$ for all $i$.

Proof
1. Let $a_i' = a_i / \sum_j a_j$ and $b_i' = b_i / \sum_j b_j$. Then $\{a_i'\}$ and $\{b_i'\}$ are probability distributions.
2. Using the divergence inequality, we have

$$0 \leq \sum_i a_i' \log \frac{a_i'}{b_i'}$$

$$= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j}$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j} \right]$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \sum_i a_i \log \frac{\sum_j a_j}{\sum_j b_j} \right]$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \left( \sum_i a_i \right) \log \frac{\sum_j a_j}{\sum_j b_j} \right],$$

which implies (1).
**Theorem 2.32 (Log-Sum Inequality)**  For positive numbers \(a_1, a_2, \ldots\) and nonnegative numbers \(b_1, b_2, \ldots\) such that \(\sum_i a_i < \infty\) and \(0 < \sum_i b_i < \infty\),

\[
\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \tag{1}
\]

Moreover, equality holds if and only if \(a_i = constant\) for all \(i\).

**Proof**

1. Let \(a'_i = a_i / \sum_j a_j\) and \(b'_i = b_i / \sum_j b_j\). Then \(\{a'_i\}\) and \(\{b'_i\}\) are probability distributions.

2. Using the divergence inequality, we have

\[
0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i} = \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i/\sum_j a_j}{b_i/\sum_j b_j} = \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i/\sum_j a_j}{b_i/\sum_j b_j} \right] = \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \sum_i a_i \log \frac{\sum_j a_j}{\sum_j b_j} \right] = \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \left( \sum_i a_i \right) \log \frac{\sum_j a_j}{\sum_j b_j} \right],
\]

which implies (1).

3. Equality holds if and only if for all \(i\),

\[
\frac{a'_i}{b'_i} = \text{constant} \quad \text{or} \quad \frac{a_i}{b_i} = \text{constant}.
\]

The theorem is proved.
Theorem 2.32 (Log-Sum Inequality) For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \tag{1}$$

Moreover, equality holds if and only if $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

Proof
1. Let $a_i' = a_i / \sum_j a_j$ and $b_i' = b_i / \sum_j b_j$. Then \{a_i'\} and \{b_i'\} are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a_i' \log \frac{a_i'}{b_i'}$$

$$= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j}$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i / \sum_j a_j}{b_i / \sum_j b_j} \right]$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \sum_i a_i \log \frac{\sum_j a_j}{\sum_j b_j} \right]$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \left( \sum a_i \right) \log \frac{\sum_j a_j}{\sum_j b_j} \right],$$

which implies (1).

3. Equality holds if and only if for all $i$,

$$a_i' = b_i' \quad \text{or} \quad \frac{a_i}{b_i} = \text{constant}.$$
**Theorem 2.32 (Log-Sum Inequality)** For positive numbers $a_1, a_2, \ldots$ and nonnegative numbers $b_1, b_2, \ldots$ such that $\sum_i a_i < \infty$ and $0 < \sum_i b_i < \infty$,

$$\sum_i a_i \log \frac{a_i}{b_i} \geq \left( \sum_i a_i \right) \log \frac{\sum_i a_i}{\sum_i b_i} \quad (1)$$

Moreover, equality holds if and only if $a_i = b_i$ or $\frac{a_i}{b_i} = \text{constant}$ for all $i$.

**Proof**

1. Let $a'_i = a_i / \sum_j a_j$ and $b'_i = b_i / \sum_j b_j$. Then $\{a'_i\}$ and $\{b'_i\}$ are probability distributions.

2. Using the divergence inequality, we have

$$0 \leq \sum_i a'_i \log \frac{a'_i}{b'_i}$$

$$= \sum_i \frac{a_i}{\sum_j a_j} \log \frac{a_i/\sum_j a_j}{b_i/\sum_j b_j}$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i/\sum_j a_j}{b_i/\sum_j b_j} \right]$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \sum_i a_i \log \frac{\sum_j a_j}{\sum_j b_j} \right]$$

$$= \frac{1}{\sum_j a_j} \left[ \sum_i a_i \log \frac{a_i}{b_i} - \left( \sum_i a_i \right) \log \frac{\sum_j a_j}{\sum_j b_j} \right],$$

which implies (1).

3. Equality holds if and only if for all $i$, $a'_i = b'_i$ or $\frac{a_i}{b_i} = \text{constant}$.

The theorem is proved.
Divergence Inequality vs Log-Sum Inequality

- The divergence inequality implies the log-sum inequality.
Divergence Inequality vs Log-Sum Inequality

- The divergence inequality implies the log-sum inequality.
- The log-sum inequality also implies the divergence inequality. (Exercise)
Divergence Inequality vs Log-Sum Inequality

- The divergence inequality implies the log-sum inequality.
- The log-sum inequality also implies the divergence inequality. \textit{(Exercise)}
- The two inequalities are equivalent.
Theorem 2.33 (Pinsker’s Inequality)

\[ D(p\|q) \geq \frac{1}{2\ln 2} V^2(p, q). \]
Theorem 2.33 (Pinsker’s Inequality)

\[ D(p\|q) \geq \frac{1}{2 \ln 2} V^2(p, q). \]

- If \( D(p\|q) \) or \( D(q\|p) \) is small, then so is \( V(p, q) = V(q, p) \).

• See Problems 23 and 24 for details.
Theorem 2.33 (Pinsker’s Inequality)

\[ D(p\|q) \geq \frac{1}{2\ln 2} V^2(p, q). \]

• If \( D(p\|q) \) or \( D(q\|p) \) is small, then so is \( V(p, q) = V(q, p) \).

• For a sequence of probability distributions \( q_k \), as \( k \to \infty \), if \( D(p\|q_k) \to 0 \) or \( D(q_k\|p) \to 0 \), then \( V(p, q_k) = V(q_k, p) \to 0 \).
Theorem 2.33 (Pinsker’s Inequality)

\[ D(p\|q) \geq \frac{1}{2 \ln 2} V^2(p, q). \]

• If \( D(p\|q) \) or \( D(q\|p) \) is small, then so is \( V(p, q) = V(q, p) \).

• For a sequence of probability distributions \( q_k \), as \( k \to \infty \), if \( D(p\|q_k) \to 0 \) or \( D(q_k\|p) \to 0 \), then \( V(p, q_k) = V(q_k, p) \to 0 \).

• That is, “convergence in divergence” is a stronger notion than “convergence in variational distance.”

• See Problems 23 and 24 for details.
Theorem 2.33 (Pinsker’s Inequality)

\[ D(p\|q) \geq \frac{1}{2\ln 2} V^2(p, q). \]

- If \( D(p\|q) \) or \( D(q\|p) \) is small, then so is \( V(p, q) = V(q, p) \).

- For a sequence of probability distributions \( q_k \), as \( k \to \infty \), if \( D(p\|q_k) \to 0 \) or \( D(q_k\|p) \to 0 \), then \( V(p, q_k) = V(q_k, p) \to 0 \).

- That is, “convergence in divergence” is a stronger notion than “convergence in variational distance.”

- See Problems 23 and 24 for details.