香港中文大學
The Chinese University of Hong Kong

2.6 The Basic Inequalities

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Corollary All Shannon's information measures are nonnegative, because they are all special cases of conditional mutual information.

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$I(X;Y|Z)$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X; Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X; Y|Z) \\ = \quad & \sum_{x,y,z} p(x, y, z) \log \frac{p(x, y|z)}{p(x|z)p(y|z)} \end{aligned}$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$I(X;Y|Z)$$
$$= \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)}$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \end{aligned}$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \end{aligned}$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} \underline{p(x,y,z)} \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \end{aligned}$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X; Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X; Y|Z) \\ = & \sum_{x,y,z} p(x, y, z) \log \frac{p(x, y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} \underline{p(x, y, z)} \log \frac{p(x, y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} \underline{p(z)p(x, y|z)} \log \frac{p(x, y|z)}{p(x|z)p(y|z)} \end{aligned}$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \end{aligned}$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X; Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned}
I(X; Y|Z) & \\
&= \sum_{x,y,z} p(x, y, z) \log \frac{p(x, y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(x, y, z) \log \frac{p(x, y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} \underline{p(z)} p(x, y|z) \log \frac{p(x, y|z)}{p(x|z)p(y|z)}
\end{aligned}$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned}
&I(X;Y|Z) \\
&= \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} \underline{p(z)} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \underline{p(z)} \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)}
\end{aligned}$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X; Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned}
& I(X; Y|Z) \\
= & \sum_{x,y,z} p(x, y, z) \log \frac{p(x, y|z)}{p(x|z)p(y|z)} \\
= & \sum_{z} \sum_{x,y} p(x, y, z) \log \frac{p(x, y|z)}{p(x|z)p(y|z)} \\
= & \sum_{z} \sum_{x,y} p(z)p(x, y|z) \log \frac{p(x, y|z)}{p(x|z)p(y|z)} \\
= & \sum_{z} p(z) \sum_{x,y} p(x, y|z) \log \frac{p(x, y|z)}{p(x|z)p(y|z)}
\end{aligned}$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned}
&I(X;Y|Z) \\
&= \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} p(z) \underline{\sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)}}
\end{aligned}$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned}
& I(X;Y|Z) \\
&= \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(z) p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} p(z) \underline{\sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)}} \\
&= \sum_{z} p(z) \underline{D(p_{XY|z} \| p_{X|z} p_{Y|z})},
\end{aligned}$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned}
& I(X;Y|Z) \\
&= \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}),
\end{aligned}$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned}
&I(X;Y|Z) \\
&= \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}),
\end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$
\begin{aligned}
& I(X;Y|Z) \\
&= \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}),
\end{aligned}
$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}), \end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) \geq 0.$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned}
& I(X;Y|Z) \\
&= \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(z) p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}),
\end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z} p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}), \end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z} p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}), \end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

if and only if for all $z \in \mathcal{S}_Z$,

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}), \end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

if and only if for all $z \in \mathcal{S}_z$,

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}), \end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

if and only if for all $z \in \mathcal{S}_Z$,

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}), \end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

if and only if for all $z \in \mathcal{S}_z$,

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) = 0.$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}), \end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z} p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

if and only if for all $z \in \mathcal{S}_z$,

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) = 0.$$

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}), \end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

if and only if for all $z \in \mathcal{S}_Z$,

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) = 0.$$

5. Then we see from Theorem 2.31 that this happens if and only if

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned}
&I(X;Y|Z) \\
&= \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}),
\end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

if and only if for all $z \in \mathcal{S}_z$,

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) = 0.$$

5. Then we see from Theorem 2.31 that this happens if and only if

Theorem 2.31 (Divergence Inequality)

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

with equality if and only if $p = q$.

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) D(p_{XY|z} \| p_{X|z}p_{Y|z}), \end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z}p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

if and only if for all $z \in \mathcal{S}_Z$,

$$D(p_{XY|z} \| p_{X|z}p_{Y|z}) = 0.$$

5. Then we see from Theorem 2.31 that this happens if and only if

$$p_{XY|z} = p_{X|z}p_{Y|z},$$

Theorem 2.31 (Divergence Inequality)

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

with equality if and only if $p = q$.

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(z) p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}), \end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

if and only if for all $z \in \mathcal{S}_z$,

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) = 0.$$

5. Then we see from Theorem 2.31 that this happens if and only if

$$p_{XY|z} = p_{X|z} p_{Y|z},$$

Theorem 2.31 (Divergence Inequality)

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

with equality if and only if $p = q$.

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} I(X;Y|Z) &= \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ &= \sum_{z}\sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ &= \sum_{z}\sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ &= \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ &= \sum_{z} p(z) D(p_{XY|z} \| p_{X|z}p_{Y|z}), \end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z}p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

if and only if for all $z \in \mathcal{S}_z$,

$$D(p_{XY|z} \| p_{X|z}p_{Y|z}) = 0.$$

5. Then we see from Theorem 2.31 that this happens if and only if

$$p_{XY|z} = p_{X|z}p_{Y|z},$$

Theorem 2.31 (Divergence Inequality)

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

with equality if and only if $p = q$.

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) D(p_{XY|z} \| p_{X|z}p_{Y|z}), \end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z}p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

if and only if for all $z \in \mathcal{S}_z$,

$$D(p_{XY|z} \| p_{X|z}p_{Y|z}) = 0.$$

5. Then we see from Theorem 2.31 that this happens if and only if

$$p_{XY|z} = p_{X|z}p_{Y|z},$$

i.e., for all $z \in \mathcal{S}_z$, for all x and y,

$$p(x,y|z) = p(x|z)p(y|z),$$

Theorem 2.31 (Divergence Inequality)

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

with equality if and only if $p = q$.

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned}
& I(X;Y|Z) \\
&= \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}),
\end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

if and only if for all $z \in \mathcal{S}_Z$,

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) = 0.$$

5. Then we see from Theorem 2.31 that this happens if and only if

$$p_{XY|z} = p_{X|z} p_{Y|z},$$

i.e., for all $z \in \mathcal{S}_Z$, for all x and y,

$$p(x,y|z) = p(x|z)p(y|z),$$

Theorem 2.31 (Divergence Inequality)

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

with equality if and only if $p = q$.

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned}
&I(X;Y|Z) \\
&= \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}),
\end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

if and only if for all $z \in \mathcal{S}_Z$,

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) = 0.$$

5. Then we see from Theorem 2.31 that this happens if and only if

$$p_{XY|z} = p_{X|z} p_{Y|z},$$

i.e., for all $z \in \mathcal{S}_Z$, for all x and y,

$$p(x,y|z) = p(x|z)p(y|z),$$

Theorem 2.31 (Divergence Inequality)

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

with equality if and only if $p = q$.

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) D(p_{XY|z} \| p_{X|z}p_{Y|z}), \end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z}p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

if and only if for all $z \in \mathcal{S}_Z$,

$$D(p_{XY|z} \| p_{X|z}p_{Y|z}) = 0.$$

5. Then we see from Theorem 2.31 that this happens if and only if

$$p_{XY|z} = p_{X|z}p_{Y|z},$$

i.e., for all $z \in \mathcal{S}_Z$, for all x and y,

$$p(x,y|z) = p(x|z)p(y|z),$$

Theorem 2.31 (Divergence Inequality)

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

with equality if and only if $p = q$.

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned}
& I(X;Y|Z) \\
= & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
= & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
= & \sum_{z} \sum_{x,y} p(z) p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
= & \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
= & \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}),
\end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

if and only if for all $z \in \mathcal{S}_Z$,

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) = 0.$$

5. Then we see from Theorem 2.31 that this happens if and only if

$$p_{XY|z} = p_{X|z} p_{Y|z},$$

i.e., for all $z \in \mathcal{S}_Z$, for all x and y,

$$p(x,y|z) = p(x|z)p(y|z),$$

or

$$p(x,y,z) = p(x,z)p(y|z).$$

Theorem 2.31 (Divergence Inequality)

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

with equality if and only if $p = q$.

Theorem 2.34 For random variables X, Y, and Z,

$$I(X;Y|Z) \geq 0,$$

with equality if and only if X and Y are independent when conditioning on Z.

Proof

1. Write

$$\begin{aligned} & I(X;Y|Z) \\ = & \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} \sum_{x,y} p(z)p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\ = & \sum_{z} p(z) D(p_{XY|z} \| p_{X|z} p_{Y|z}), \end{aligned}$$

where $p_{XY|z}$ denotes $\{p(x,y|z), (x,y) \in \mathcal{X} \times \mathcal{Y}\}$, etc.

2. Since for a fixed z, both $p_{XY|z}$ and $p_{X|z}p_{Y|z}$ are joint probability distributions on $\mathcal{X} \times \mathcal{Y}$, we have

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) \geq 0.$$

3. Therefore, we conclude that $I(X;Y|Z) \geq 0$.

4. Now,

$$I(X;Y|Z) = 0$$

if and only if for all $z \in \mathcal{S}_Z$,

$$D(p_{XY|z} \| p_{X|z} p_{Y|z}) = 0.$$

5. Then we see from Theorem 2.31 that this happens if and only if

$$p_{XY|z} = p_{X|z} p_{Y|z},$$

i.e., for all $z \in \mathcal{S}_Z$, for all x and y,

$$p(x,y|z) = p(x|z)p(y|z),$$

or

$$p(x,y,z) = p(x,z)p(y|z).$$

This is the condition for X and Y being independent conditioning on Z.

Theorem 2.31 (Divergence Inequality)

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

with equality if and only if $p = q$.

Proposition 2.35 $H(X) = 0$ if and only if X is deterministic.

Proposition 2.35 $H(X) = 0$ if and only if X is deterministic.

Proposition 2.36 $H(Y|X) = 0$ if and only if Y is a function of X.

Proposition 2.35 $H(X) = 0$ if and only if X is deterministic.

Proposition 2.36 $H(Y|X) = 0$ if and only if Y is a function of X.

Proposition 2.37 $I(X;Y) = 0$ if and only if X and Y are independent.

Proposition 2.35 $H(X) = 0$ if and only if X is deterministic.

Proof

Proposition 2.35 $H(X) = 0$ if and only if X is deterministic.

Proof

1. 'If'

Proposition 2.35 $H(X) = 0$ if and only if X is deterministic.

Proof

1. 'If'

If X is deterministic, i.e., there exists $x^* \in \mathcal{X}$ such that

Proposition 2.35 $H(X) = 0$ if and only if X is deterministic.

Proof

1. 'If'

If X is deterministic, i.e., there exists $x^* \in \mathcal{X}$ such that

$$p(x^*) = 1$$

and $p(x) = 0$ for all $x \neq x^*$, then

Proposition 2.35 $H(X) = 0$ if and only if X is deterministic.

Proof

1. 'If'

If X is deterministic, i.e., there exists $x^* \in \mathcal{X}$ such that

$$p(x^*) = 1$$

and $p(x) = 0$ for all $x \neq x^*$, then

$$H(X) = -p(x^*) \log p(x^*)$$

Proposition 2.35 $H(X) = 0$ if and only if X is deterministic.

Proof

1. 'If'

If X is deterministic, i.e., there exists $x^* \in \mathcal{X}$ such that

$$p(x^*) = 1$$

and $p(x) = 0$ for all $x \neq x^*$, then

$$\begin{aligned} H(X) &= -p(x^*) \log p(x^*) \\ &= -1 \log 1 \end{aligned}$$

Proposition 2.35 $H(X) = 0$ if and only if X is deterministic.

Proof

1. 'If'

If X is deterministic, i.e., there exists $x^* \in \mathcal{X}$ such that

$$p(x^*) = 1$$

and $p(x) = 0$ for all $x \neq x^*$, then

$$\begin{aligned} H(X) &= -p(x^*)\log p(x^*) \\ &= -1\log 1 \\ &= 0. \end{aligned}$$

Proposition 2.35 $H(X) = 0$ if and only if X is deterministic.

Proof

1. 'If'

If X is deterministic, i.e., there exists $x^* \in \mathcal{X}$ such that

$$p(x^*) = 1$$

and $p(x) = 0$ for all $x \neq x^*$, then

$$\begin{aligned} H(X) &= -p(x^*) \log p(x^*) \\ &= -1 \log 1 \\ &= 0. \end{aligned}$$

2. 'Only if'

Proposition 2.35 $H(X) = 0$ if and only if X is deterministic.

Proof

1. 'If'

If X is deterministic, i.e., there exists $x^* \in \mathcal{X}$ such that

$$p(x^*) = 1$$

and $p(x) = 0$ for all $x \neq x^*$, then

$$\begin{aligned} H(X) &= -p(x^*) \log p(x^*) \\ &= -1 \log 1 \\ &= 0. \end{aligned}$$

2. 'Only if'

If X is not deterministic, i.e., there exists $x^* \in \mathcal{X}$ such that

$$0 < p(x^*) < 1,$$

Proposition 2.35 $H(X) = 0$ if and only if X is deterministic.

Proof

1. 'If'

If X is deterministic, i.e., there exists $x^* \in \mathcal{X}$ such that

$$p(x^*) = 1$$

and $p(x) = 0$ for all $x \neq x^*$, then

$$\begin{aligned} H(X) &= -p(x^*) \log p(x^*) \\ &= -1 \log 1 \\ &= 0. \end{aligned}$$

2. 'Only if'

If X is not deterministic, i.e., there exists $x^* \in \mathcal{X}$ such that

$$0 < p(x^*) < 1,$$

then

$$H(X) \geq -p(x^*) \log p(x^*) > 0.$$

Proposition 2.35 $H(X) = 0$ if and only if X is deterministic.

Proof

1. 'If'

If X is deterministic, i.e., there exists $x^* \in \mathcal{X}$ such that

$$p(x^*) = 1$$

and $p(x) = 0$ for all $x \neq x^*$, then

$$\begin{aligned} H(X) &= -p(x^*) \log p(x^*) \\ &= -1 \log 1 \\ &= 0. \end{aligned}$$

2. 'Only if'

If X is not deterministic, i.e., there exists $x^* \in \mathcal{X}$ such that

$$0 < p(x^*) < 1,$$

then

$$H(X) \geq -p(x^*) \log p(x^*) > 0.$$

3. Therefore, we conclude that $H(X) = 0$ if and only if X is deterministic.

Proposition 2.36 $H(Y|X) = 0$ if and only if Y is a function of X.

Proof

Proposition 2.36 $H(Y|X) = 0$ if and only if Y is a function of X.

Proof

1. Consider

$$H(Y|X) = \sum_x p(x) H(Y|X = x).$$

Proposition 2.36 $H(Y|X) = 0$ if and only if Y is a function of X.

Proof

1. Consider

$$H(Y|X) = \sum_x p(x)H(Y|X = x).$$

2. We see that $H(Y|X) = 0$ if and only if

Proposition 2.36 $H(Y|X) = 0$ if and only if Y is a function of X.

Proof

1. Consider

$$H(Y|X) = \sum_x p(x)H(Y|X = x).$$

2. We see that $H(Y|X) = 0$ if and only if

$$H(Y|X = x) = 0$$

for each $x \in \mathcal{S}_X$.

Proposition 2.36 $H(Y|X) = 0$ if and only if Y is a function of X.

Proof

1. Consider

$$H(Y|X) = \sum_x p(x) H(Y|X = x).$$

2. We see that $H(Y|X) = 0$ if and only if

$$H(Y|X = x) = 0$$

for each $x \in \mathcal{S}_X$.

3. From the last proposition, this happens if and only if Y is deterministic for each given x. In other words, Y is a function of X.