香港中文大學
The Chinese University of Hong Kong

10.2 Definition

Definition 10.10 The differential entropy $h(X)$ of a continuous random variable X with pdf $f(x)$ is defined as

$$h(X) = -\int_{\mathcal{S}} f(x) \log f(x) dx = -E \log f(X)$$

Definition 10.10 The differential entropy $h(X)$ of a continuous random variable X with pdf $f(x)$ is defined as

$$h(X) = -\int_{\mathcal{S}} f(x) \log f(x) dx = -E \log f(X)$$

Remarks

Definition 10.10 The differential entropy $h(X)$ of a continuous random variable X with pdf $f(x)$ is defined as

$$h(X) = -\int_{\mathcal{S}} f(x) \log f(x) dx = -E \log f(X)$$

Remarks

1. Differential entropy is not a measure of the average amount of information contained in a continuous r.v.

Definition 10.10 The differential entropy $h(X)$ of a continuous random variable X with pdf $f(x)$ is defined as

$$h(X) = -\int_{\mathcal{S}} f(x) \log f(x) dx = -E \log f(X)$$

Remarks

1. Differential entropy is not a measure of the average amount of information contained in a continuous r.v.

2. A continuous random variable generally contains an infinite amount of information.

Example 10.11 Let X be uniformly distributed on $[0, 1)$. Then we can write

$$X = .X_1 X_2 X_3 \cdots ,$$

the dyadic expansion of X, where $X_1, X_2, X_3, \cdots$ is a sequence of fair bits.

Example 10.11 Let X be uniformly distributed on $[0,1)$. Then we can write

$$X = .X_1X_2X_3 \cdots ,$$

the dyadic expansion of X, where $X_1, X_2, X_3, \cdots$ is a sequence of fair bits. Then

$$H(X) = H(X_1, X_2, X_3, \cdots)$$

Example 10.11 Let X be uniformly distributed on $[0, 1)$. Then we can write

$$X = .X_1 X_2 X_3 \cdots ,$$

the dyadic expansion of X, where $X_1, X_2, X_3, \cdots$ is a sequence of fair bits. Then

$$\begin{aligned} H(X) &= H(X_1, X_2, X_3, \cdots) \\ &= \sum_{i=1}^{\infty} H(X_i) \end{aligned}$$

Example 10.11 Let X be uniformly distributed on $[0, 1)$. Then we can write

$$X = .X_1X_2X_3 \cdots ,$$

the dyadic expansion of X, where $X_1, X_2, X_3, \cdots$ is a sequence of fair bits. Then

$$\begin{aligned} H(X) &= H(X_1, X_2, X_3, \cdots) \\ &= \sum_{i=1}^{\infty} H(X_i) \\ &= \sum_{i=1}^{\infty} 1 \end{aligned}$$

Example 10.11 Let X be uniformly distributed on $[0, 1)$. Then we can write

$$X = .X_1 X_2 X_3 \cdots ,$$

the dyadic expansion of X, where $X_1, X_2, X_3, \cdots$ is a sequence of fair bits. Then

$$\begin{aligned} H(X) &= H(X_1, X_2, X_3, \cdots) \\ &= \sum_{i=1}^{\infty} H(X_i) \\ &= \sum_{i=1}^{\infty} 1 \\ &= \infty \end{aligned}$$

Relation with Discrete Entropy

Relation with Discrete Entropy

- Consider a continuous r.v. X with a continuous pdf $f(x)$.

Relation with Discrete Entropy

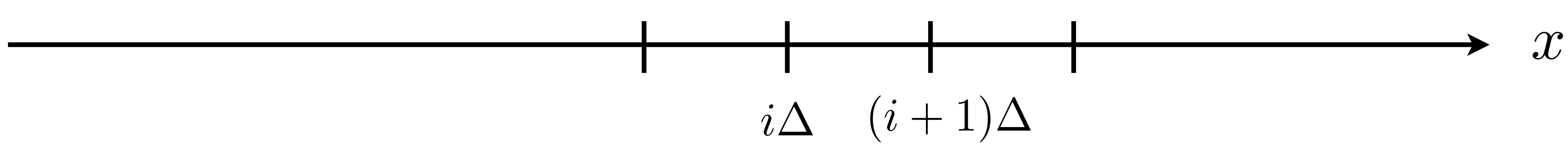
- Consider a continuous r.v. X with a continuous pdf $f(x)$.
- Define a discrete r.v. $\hat{X}_\Delta$ by

$$\hat{X}_\Delta = i \quad \text{if} \quad X \in [\, i\Delta, (i+1)\Delta).$$

Relation with Discrete Entropy

- Consider a continuous r.v. X with a continuous pdf $f(x)$.
- Define a discrete r.v. $\hat{X}_\Delta$ by

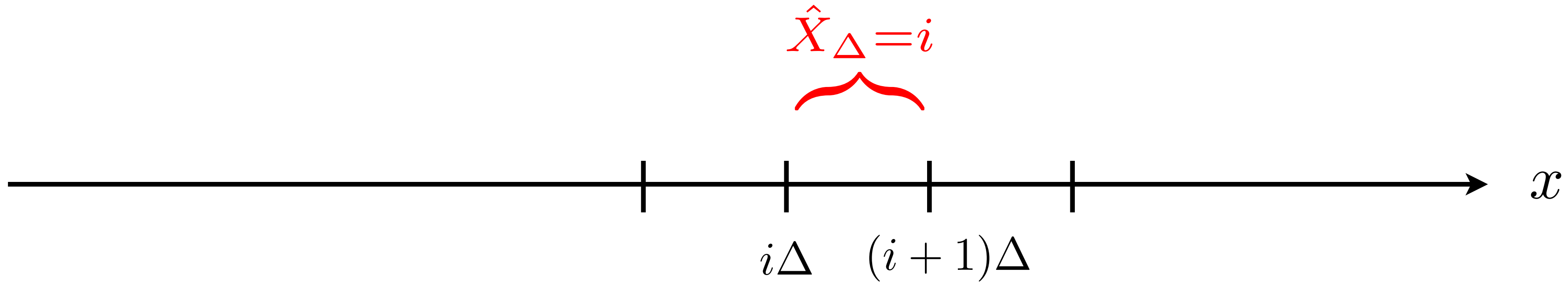
$$\hat{X}_\Delta = i \quad \text{if} \quad X \in [\, i\Delta, (i+1)\Delta).$$



Relation with Discrete Entropy

- Consider a continuous r.v. X with a continuous pdf $f(x)$.
- Define a discrete r.v. $\hat{X}_\Delta$ by

$$\hat{X}_\Delta = i \quad \text{if} \quad X \in [\, i\Delta, (i+1)\Delta).$$

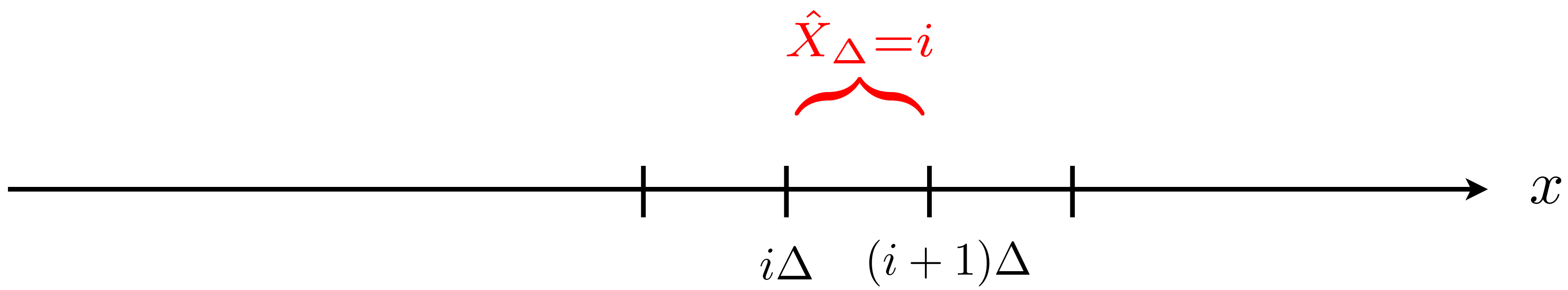


Relation with Discrete Entropy

- Consider a continuous r.v. X with a continuous pdf $f(x)$.
- Define a discrete r.v. $\hat{X}_\Delta$ by

$$\hat{X}_\Delta = i \quad \text{if} \quad X \in [\, i\Delta, (i+1)\Delta).$$

- $\hat{X}_\Delta$ is a quantization of X with resolution Δ.



Relation with Discrete Entropy

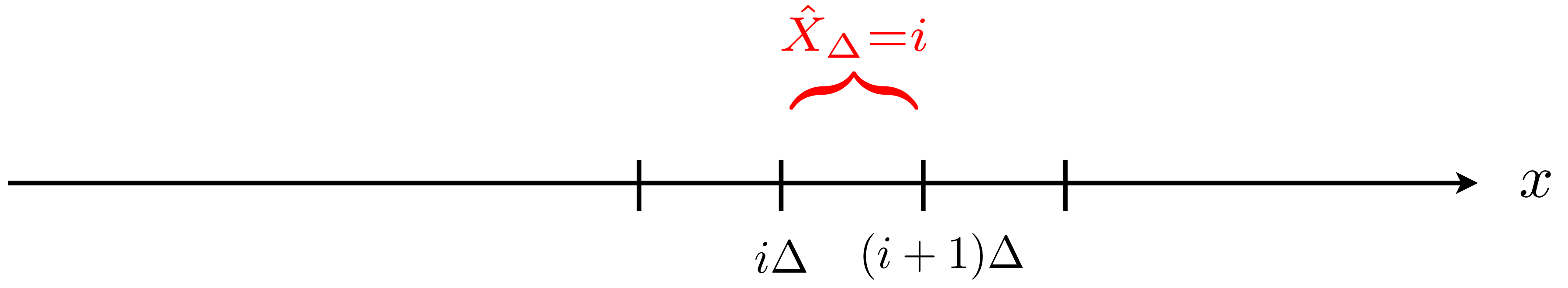
- Consider a continuous r.v. X with a continuous pdf $f(x)$.
- Define a discrete r.v. $\hat{X}_\Delta$ by

$$\hat{X}_\Delta = i \quad \text{if} \quad X \in [\, i\Delta, (i+1)\Delta).$$

- $\hat{X}_\Delta$ is a quantization of X with resolution Δ.
- Since $f(x)$ is continuous,

$$p_i = \Pr\{\hat{X}_\Delta = i\} \approx f(x_i)\Delta$$

where $x_i \in [\, i\Delta, (i+1)\Delta)$.



Relation with Discrete Entropy

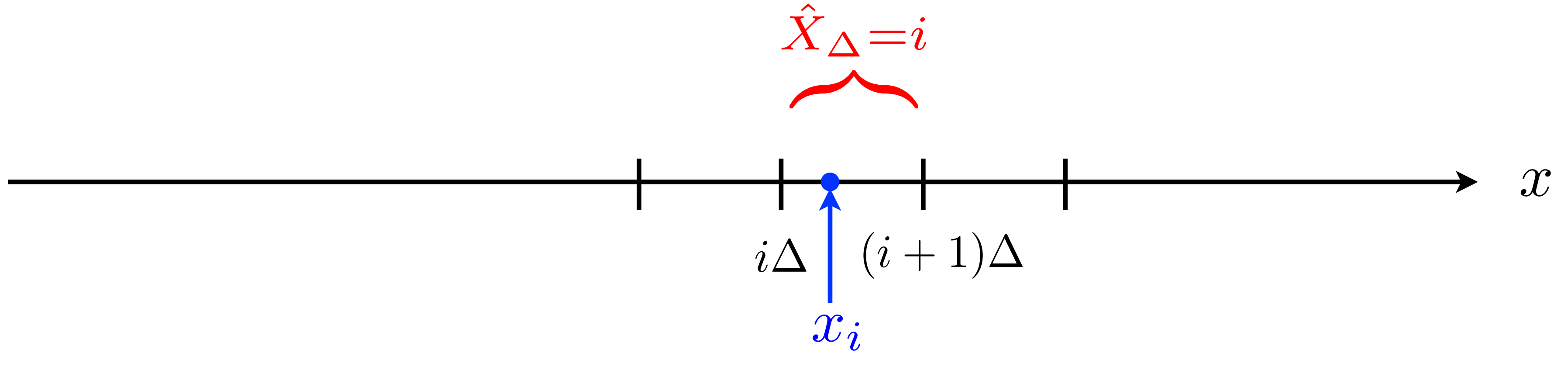
- Consider a continuous r.v. X with a continuous pdf $f(x)$.
- Define a discrete r.v. $\hat{X}_\Delta$ by

$$\hat{X}_\Delta = i \quad \text{if} \quad X \in [\, i\Delta, (i+1)\Delta).$$

- $\hat{X}_\Delta$ is a quantization of X with resolution Δ.
- Since $f(x)$ is continuous,

$$p_i = \Pr\{\hat{X}_\Delta = i\} \approx f(x_i)\Delta$$

where $x_i \in [\, i\Delta, (i+1)\Delta)$.



- Then for small Δ,

- Then for small Δ,

$$H(\hat{X}_{\Delta}) = -\sum_i p_i \log p_i$$

- Then for small Δ,

$$H(\hat{X}_{\Delta}) = -\sum_i \underline{p_i} \log p_i$$

- Then for small Δ,

$$\begin{aligned} H(\hat{X}_{\Delta}) &= -\sum_i \underline{p_i} \log p_i \\ &\approx -\sum_i (\underline{f(x_i)\Delta}) \log(f(x_i)\Delta) \end{aligned}$$

- Then for small Δ,

$$\begin{aligned} H(\hat{X}_{\Delta}) &= -\sum_i p_i \log \underline{p_i} \\ &\approx -\sum_i (f(x_i)\Delta) \log(\underline{f(x_i)\Delta}) \end{aligned}$$

- Then for small Δ,

$$\begin{aligned} H(\hat{X}_{\Delta}) &= -\sum_i p_i \log p_i \\ &\approx -\sum_i (f(x_i)\Delta) \underline{\log(f(x_i)\Delta)} \end{aligned}$$

- Then for small Δ,

$$
\begin{aligned}
H(\hat{X}_{\Delta}) &= -\sum_i p_i \log p_i \\
&\approx -\sum_i (f(x_i)\Delta) \underline{\log(f(x_i)\Delta)} \\
&\approx -\sum_i (f(x_i)\Delta) \left(\underline{\log f(x_i)} + \underline{\log \Delta}\right)
\end{aligned}
$$

- Then for small Δ,

$$
\begin{aligned}
H(\hat{X}_{\Delta}) &= -\sum_i p_i \log p_i \\
&\approx -\sum_i (f(x_i)\Delta) \log(f(x_i)\Delta) \\
&\approx -\sum_i (f(x_i)\Delta) \left(\underline{\log f(x_i)} + \log \Delta\right) \\
&= -\sum_i f(x_i)\Delta \underline{\log f(x_i)} - \sum_i f(x_i)\Delta \log \Delta
\end{aligned}
$$

- Then for small Δ,

$$
\begin{aligned}
H(\hat{X}_{\Delta}) &= -\sum_i p_i \log p_i \\
&\approx -\sum_i (f(x_i)\Delta) \log(f(x_i)\Delta) \\
&\approx -\sum_i (f(x_i)\Delta) \left(\log f(x_i) + \underline{\log \Delta}\right) \\
&= -\sum_i f(x_i)\Delta \log f(x_i) - \sum_i f(x_i)\Delta \, \underline{\log \Delta}
\end{aligned}
$$

- Then for small Δ,

$$
\begin{aligned}
H(\hat{X}_{\Delta}) &= -\sum_i p_i \log p_i \\
&\approx -\sum_i (f(x_i)\Delta) \log(f(x_i)\Delta) \\
&\approx -\sum_i (f(x_i)\Delta) \left(\log f(x_i) + \log \Delta\right) \\
&= -\sum_i f(x_i)\Delta \log f(x_i) - \sum_i f(x_i)\Delta \log \Delta \\
&= -\sum_i [f(x_i) \log f(x_i)]\Delta - (\log \Delta) \sum_i f(x_i)\Delta
\end{aligned}
$$

- Then for small Δ,

$$
\begin{aligned}
H(\hat{X}_{\Delta}) &= -\sum_i p_i \log p_i \\
&\approx -\sum_i (f(x_i)\Delta) \log(f(x_i)\Delta) \\
&\approx -\sum_i (f(x_i)\Delta) \left(\log f(x_i) + \log \Delta\right) \\
&= -\sum_i \underline{f(x_i)}\Delta \, \underline{\log f(x_i)} - \sum_i f(x_i)\Delta \log \Delta \\
&= -\sum_i [\underline{f(x_i) \log f(x_i)}]\Delta - (\log \Delta) \sum_i f(x_i)\Delta
\end{aligned}
$$

- Then for small Δ,

$$
\begin{aligned}
H(\hat{X}_{\Delta}) &= -\sum_i p_i \log p_i \\
&\approx -\sum_i (f(x_i)\Delta) \log(f(x_i)\Delta) \\
&\approx -\sum_i (f(x_i)\Delta) \left(\log f(x_i) + \log \Delta\right) \\
&= -\sum_i f(x_i)\Delta \log f(x_i) - \sum_i f(x_i)\Delta \underline{\log \Delta} \\
&= -\sum_i [f(x_i) \log f(x_i)]\Delta - (\underline{\log \Delta}) \sum_i f(x_i)\Delta
\end{aligned}
$$

- Then for small Δ,

$$
\begin{aligned}
H(\hat{X}_{\Delta}) &= -\sum_i p_i \log p_i \\
&\approx -\sum_i (f(x_i)\Delta) \log(f(x_i)\Delta) \\
&\approx -\sum_i (f(x_i)\Delta) \left(\log f(x_i) + \log \Delta\right) \\
&= -\sum_i f(x_i)\Delta \log f(x_i) - \sum_i f(x_i)\Delta \log \Delta \\
&= -\sum_i [f(x_i) \log f(x_i)]\Delta - (\log \Delta) \sum_i f(x_i)\Delta \\
&\approx -\int f(x) \log f(x) dx - (\log \Delta) \int f(x) dx
\end{aligned}
$$

- Then for small Δ,

$$
\begin{aligned}
H(\hat{X}_{\Delta}) &= -\sum_i p_i \log p_i \\
&\approx -\sum_i (f(x_i)\Delta) \log(f(x_i)\Delta) \\
&\approx -\sum_i (f(x_i)\Delta) \left(\log f(x_i) + \log \Delta\right) \\
&= -\sum_i f(x_i)\Delta \log f(x_i) - \sum_i f(x_i)\Delta \log \Delta \\
&= -\sum_i [f(x_i) \log f(x_i)]\Delta - (\log \Delta) \sum_i f(x_i)\Delta \\
&\approx \underline{-\int f(x) \log f(x) dx} - (\log \Delta) \int f(x) dx
\end{aligned}
$$

- Then for small Δ,

$$
\begin{aligned}
H(\hat{X}_{\Delta}) &= -\sum_i p_i \log p_i \\
&\approx -\sum_i (f(x_i)\Delta) \log(f(x_i)\Delta) \\
&\approx -\sum_i (f(x_i)\Delta) \left(\log f(x_i) + \log \Delta\right) \\
&= -\sum_i f(x_i)\Delta \log f(x_i) - \sum_i f(x_i)\Delta \log \Delta \\
&= -\sum_i [f(x_i) \log f(x_i)]\Delta - (\log \Delta) \sum_i f(x_i)\Delta \\
&\approx -\int f(x) \log f(x) dx - (\log \Delta) \int f(x) dx
\end{aligned}
$$

- Then for small Δ,

$$
\begin{aligned}
H(\hat{X}_\Delta) &= -\sum_i p_i \log p_i \\
&\approx -\sum_i (f(x_i)\Delta) \log(f(x_i)\Delta) \\
&\approx -\sum_i (f(x_i)\Delta) (\log f(x_i) + \log \Delta) \\
&= -\sum_i f(x_i)\Delta \log f(x_i) - \sum_i f(x_i)\Delta \log \Delta \\
&= -\sum_i [f(x_i) \log f(x_i)]\Delta - (\log \Delta) \sum_i f(x_i)\Delta \\
&\approx -\int f(x) \log f(x) dx - (\log \Delta) \int f(x) dx \\
&= \underline{h(X)} - \log \Delta.
\end{aligned}
$$

- Then for small Δ,

$$
\begin{aligned}
H(\hat{X}_{\Delta}) &= -\sum_i p_i \log p_i \\
&\approx -\sum_i (f(x_i)\Delta) \log(f(x_i)\Delta) \\
&\approx -\sum_i (f(x_i)\Delta) \left(\log f(x_i) + \log \Delta\right) \\
&= -\sum_i f(x_i)\Delta \log f(x_i) - \sum_i f(x_i)\Delta \log \Delta \\
&= -\sum_i [f(x_i) \log f(x_i)]\Delta - (\log \Delta) \sum_i f(x_i)\Delta \\
&\approx -\int f(x) \log f(x) dx - (\log \Delta) \int f(x) dx \\
&= h(X) - \log \Delta.
\end{aligned}
$$

- Then for small Δ,

$$
\begin{aligned}
H(\hat{X}_{\Delta}) &= -\sum_i p_i \log p_i \\
&\approx -\sum_i (f(x_i)\Delta)\log(f(x_i)\Delta) \\
&\approx -\sum_i (f(x_i)\Delta)\left(\log f(x_i) + \log\Delta\right) \\
&= -\sum_i f(x_i)\Delta \log f(x_i) - \sum_i f(x_i)\Delta \log\Delta \\
&= -\sum_i [f(x_i)\log f(x_i)]\Delta - (\log\Delta)\sum_i f(x_i)\Delta \\
&\approx -\int f(x)\log f(x)dx - (\log\Delta)\int f(x)dx \\
&= h(X) - \underline{\log \Delta}.
\end{aligned}
$$

- Then for small Δ,

$$
\begin{aligned}
H(\hat{X}_{\Delta}) &= -\sum_i p_i \log p_i \\
&\approx -\sum_i (f(x_i)\Delta) \log(f(x_i)\Delta) \\
&\approx -\sum_i (f(x_i)\Delta) (\log f(x_i) + \log \Delta) \\
&= -\sum_i f(x_i)\Delta \log f(x_i) - \sum_i f(x_i)\Delta \log \Delta \\
&= -\sum_i [f(x_i) \log f(x_i)]\Delta - (\log \Delta) \sum_i f(x_i)\Delta \\
&\approx -\int f(x) \log f(x) dx - (\log \Delta) \int f(x) dx \\
&= h(X) - \log \Delta.
\end{aligned}
$$

- Then for small Δ,

$$\begin{aligned}
H(\hat{X}_\Delta) &= -\sum_i p_i \log p_i \\
&\approx -\sum_i (f(x_i)\Delta)\log(f(x_i)\Delta) \\
&\approx -\sum_i (f(x_i)\Delta)\,(\log f(x_i) + \log\Delta) \\
&= -\sum_i f(x_i)\Delta \log f(x_i) - \sum_i f(x_i)\Delta \log\Delta \\
&= -\sum_i [f(x_i)\log f(x_i)]\Delta - (\log\Delta)\sum_i f(x_i)\Delta \\
&\approx -\int f(x)\log f(x)dx - (\log\Delta)\int f(x)dx \\
&= h(X) - \log\Delta.
\end{aligned}$$

- Note that $H(\hat{X}_\Delta) \to \infty$ as $\Delta \to 0$.

Example 10.12 Let X be uniformly distributed on $[0, a)$. Then

$$h(X) = -\int_0^a \frac{1}{a} \log \frac{1}{a} dx = \log a.$$

Example 10.12 Let X be uniformly distributed on $[0, a)$. Then

$$h(X) = -\int_0^a \frac{1}{a} \log \frac{1}{a} dx = \log a.$$

Remark $h(X) < 0$ if $a < 1$, so $h(\cdot)$ cannot be a measure of information.

Example 10.12 Let X be uniformly distributed on $[0, a)$. Then

$$h(X) = -\int_0^a \frac{1}{a} \log \frac{1}{a} dx = \log a.$$

Remark $h(X) < 0$ if $a < 1$, so $h(\cdot)$ cannot be a measure of information.

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$. Then

$$h(X) = \frac{1}{2} \log(2\pi e \sigma^2).$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$h(X) = -\int f(x) \ln f(x) dx$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$h(X) = -\int f(x) \ln f(x) dx$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$h(X) = -\int f(x) \ln f(x) dx$$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$h(X) = -\int f(x) \underline{\ln f(x)} dx$$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

$$\ln f(x) = -\ln\sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2}$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$\begin{aligned} h(X) &= -\int f(x) \underline{\ln f(x)} dx \\ &= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \end{aligned}$$

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\ \ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2} \end{aligned}$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$\begin{aligned} h(X) &= -\int f(x) \ln f(x) dx \\ &= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\ &= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \end{aligned}$$

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\ \ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2} \end{aligned}$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$\begin{aligned} h(X) &= -\int f(x) \ln f(x) dx \\ &= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\ &= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \end{aligned}$$

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\ \ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2} \end{aligned}$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$\begin{aligned} h(X) &= -\int f(x) \ln f(x) dx \\ &= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\ &= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \\ &= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \end{aligned}$$

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\ \ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2} \end{aligned}$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$
\begin{aligned}
h(X) &= -\int f(x) \ln f(x) dx \\
&= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\
&= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \underline{\ln \sqrt{2\pi\sigma^2}} \int f(x) dx \\
&= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2)
\end{aligned}
$$

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\
\ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2}
\end{aligned}
$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$\begin{aligned} h(X) &= -\int f(x) \ln f(x) dx \\ &= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\ &= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \\ &= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \end{aligned}$$

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\ \ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2} \end{aligned}$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$
\begin{aligned}
h(X) &= -\int f(x) \ln f(x) dx \\
&= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\
&= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \\
&= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2)
\end{aligned}
$$

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\
\ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2}
\end{aligned}
$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$
\begin{aligned}
h(X) &= -\int f(x) \ln f(x) dx \\
&= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\
&= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \\
&= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2)
\end{aligned}
$$

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\
\ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2}
\end{aligned}
$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$
\begin{aligned}
h(X) &= -\int f(x) \ln f(x) dx \\
&= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\
&= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \\
&= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\mathrm{var} X + (EX)^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2)
\end{aligned}
$$

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\
\ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2}
\end{aligned}
$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$\begin{aligned}
h(X) &= -\int f(x) \ln f(x) dx \\
&= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\
&= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \\
&= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\text{var} X + (EX)^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2)
\end{aligned}$$

$$\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\
\ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2}
\end{aligned}$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$
\begin{aligned}
h(X) &= -\int f(x) \ln f(x) dx \\
&= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\
&= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \\
&= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\underline{\text{var}X} + (EX)^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\underline{\sigma^2} + 0}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2)
\end{aligned}
$$

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\
\ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2}
\end{aligned}
$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$\begin{aligned}
h(X) &= -\int f(x) \ln f(x) dx \\
&= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\
&= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \\
&= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\mathrm{var}X + (EX)^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\sigma^2 + 0}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2)
\end{aligned}$$

$$\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\
\ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2}
\end{aligned}$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$\begin{aligned}
h(X) &= -\int f(x) \ln f(x) dx \\
&= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\
&= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \\
&= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\mathrm{var} X + (EX)^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\sigma^2 + 0}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2)
\end{aligned}$$

$$\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\
\ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2}
\end{aligned}$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$
\begin{aligned}
h(X) &= -\int f(x) \ln f(x) dx \\
&= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\
&= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \\
&= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\text{var} X + (EX)^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\sigma^2 + 0}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2)
\end{aligned}
$$

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\
\ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2}
\end{aligned}
$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$
\begin{aligned}
h(X) &= -\int f(x) \ln f(x) dx \\
&= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\
&= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \\
&= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\mathrm{var}X + (EX)^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\sigma^2 + 0}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2)
\end{aligned}
$$

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\
\ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2}
\end{aligned}
$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$
\begin{aligned}
h(X) &= -\int f(x) \ln f(x) dx \\
&= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\
&= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \\
&= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\text{var} X + (EX)^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\sigma^2 + 0}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{1}{2} + \frac{1}{2} \ln(2\pi\sigma^2)
\end{aligned}
$$

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\
\ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2}
\end{aligned}
$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$
\begin{aligned}
h(X) &= -\int f(x) \ln f(x) dx \\
&= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\
&= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \\
&= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\mathrm{var} X + (EX)^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\sigma^2 + 0}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{1}{2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{1}{2} \underline{\ln e} + \frac{1}{2} \ln(2\pi\sigma^2)
\end{aligned}
$$

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\
\ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2}
\end{aligned}
$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$
\begin{aligned}
h(X) &= -\int f(x) \ln f(x) dx \\
&= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\
&= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \\
&= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\text{var}X + (EX)^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\sigma^2 + 0}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{1}{2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{1}{2} \ln e + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{1}{2} \ln(2\pi \underline{e} \sigma^2)
\end{aligned}
$$

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\
\ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2}
\end{aligned}
$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$\begin{aligned} h(X) &= -\int f(x) \ln f(x) dx \\ &= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\ &= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \\ &= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\ &= \frac{\text{var} X + (EX)^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\ &= \frac{\sigma^2 + 0}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\ &= \frac{1}{2} + \frac{1}{2} \ln(2\pi\sigma^2) \\ &= \frac{1}{2} \ln e + \frac{1}{2} \ln(2\pi\sigma^2) \\ &= \frac{1}{2} \ln(2\pi \underline{e} \sigma^2) \end{aligned}$$

in nats. Changing the base of the logarithm to any chosen positive value, we obtain

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\ \ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2} \end{aligned}$$

Example 10.13 (Gaussian Distribution) Let $X \sim \mathcal{N}(0, \sigma^2)$ and let e be the base of the logarithm. Then

$$
\begin{aligned}
h(X) &= -\int f(x) \ln f(x) dx \\
&= -\int f(x) \left(-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right) dx \\
&= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \ln \sqrt{2\pi\sigma^2} \int f(x) dx \\
&= \frac{EX^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\mathrm{var} X + (EX)^2}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{\sigma^2 + 0}{2\sigma^2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{1}{2} + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{1}{2} \ln e + \frac{1}{2} \ln(2\pi\sigma^2) \\
&= \frac{1}{2} \ln(2\pi \underline{e} \sigma^2)
\end{aligned}
$$

in nats. Changing the base of the logarithm to any chosen positive value, we obtain

$$
h(X) = \frac{1}{2} \underline{\log(} 2\pi e \sigma^2).
$$

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \\
\ln f(x) &= -\ln \sqrt{2\pi\sigma^2} - \frac{x^2}{2\sigma^2}
\end{aligned}
$$

Properties of Differential Entropy

Properties of Differential Entropy

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Properties of Differential Entropy

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Properties of Differential Entropy

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Properties of Differential Entropy

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Remarks on Scaling

Properties of Differential Entropy

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Remarks on Scaling The differential entropy is

Properties of Differential Entropy

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Remarks on Scaling The differential entropy is

- increased by $\log |a|$ if $|a| > 1$

Properties of Differential Entropy

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Remarks on Scaling The differential entropy is

- increased by $\log |a|$ if $|a| > 1$
- decreased by $-\log |a|$ if $|a| < 1$

Properties of Differential Entropy

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Remarks on Scaling The differential entropy is

- increased by $\log |a|$ if $|a| > 1$
- decreased by $-\log |a|$ if $|a| < 1$
- unchanged if $a = -1$

Properties of Differential Entropy

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Remarks on Scaling The differential entropy is

- increased by $\log |a|$ if $|a| > 1$
- decreased by $-\log |a|$ if $|a| < 1$
- unchanged if $a = -1$
- related to the "spread" of the pdf

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then
$$f_Y(y) = f_X(y - c)$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.

2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.

2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$h(X) = -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.

2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$h(X) = -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$h(X) = -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.

2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y-c) \log f_X(y-c) dy \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.

2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.

2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.

2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.

2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} \underline{f_X(y - c)} \log \underline{f_X(y - c)} dy \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} \underline{f_X(y-c)} \log \underline{f_X(y-c)} dy \\ &= -\int_{\mathcal{S}_Y} \underline{f_Y(y)} \log \underline{f_Y(y)} dy \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.

2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.

2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\ &= h(Y) \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\ &= h(Y) \\ &= h(X + c). \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X+c) = h(X).$$

Proof

1. Let $Y = X + c$.

2. Then

$$f_Y(y) = f_X(y-c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned}
h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X(y-c) \log f_X(y-c) dy \\
&= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\
&= h(Y) \\
&= h(X+c).
\end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then
$$f_Y(y) = f_X(y - c)$$
and
$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$
3. Letting $x = y - c$, we have

$$\begin{aligned}
h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\
&= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\
&= h(Y) \\
&= h(X + c).
\end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.

2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned}
h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\
&= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\
&= h(Y) \\
&= h(X + c).
\end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\ &= h(Y) \\ &= h(X + c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then
$$f_Y(y) = f_X(y - c)$$
and
$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$
3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\ &= h(Y) \\ &= h(X + c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then
$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$
and
$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then
$$f_Y(y) = f_X(y - c)$$
and
$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$
3. Letting $x = y - c$, we have

$$\begin{aligned}
h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\
&= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\
&= h(Y) \\
&= h(X + c).
\end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then
$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$
and
$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$
3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned}
h(X) & \\
&= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx
\end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then
$$f_Y(y) = f_X(y - c)$$
and
$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$
3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\ &= h(Y) \\ &= h(X + c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then
$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$
and
$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$
3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned} &h(X) \\ &\quad = -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned}
h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\
&= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\
&= h(Y) \\
&= h(X + c).
\end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$

and

$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$

3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned}
h(X) \\
&= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx
\end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\ &= h(Y) \\ &= h(X + c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$

and

$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$

3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned} h(X) & \\ &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log f_X\left(\frac{y}{a}\right) \frac{dy}{|a|} \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then
$$f_Y(y) = f_X(y - c)$$
and
$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$
3. Letting $x = y - c$, we have

$$\begin{aligned}
h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\
&= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\
&= h(Y) \\
&= h(X + c).
\end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then
$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$
and
$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$
3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned}
h(X) & \\
&= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log f_X\left(\frac{y}{a}\right) \frac{dy}{|a|}
\end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then
$$f_Y(y) = f_X(y - c)$$
and
$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$
3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\ &= h(Y) \\ &= h(X + c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then
$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$
and
$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$
3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned} h(X) & \\ &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) \underline{dx} \\ &= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log f_X\left(\frac{y}{a}\right) \underline{\frac{dy}{|a|}} \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X+c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then
$$f_Y(y) = f_X(y-c)$$
and
$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$
3. Letting $x = y - c$, we have

$$\begin{aligned}
h(X) &= -\int_{\mathcal{S}_X} f_X(x)\log f_X(x)dx \\
&= -\int_{\mathcal{S}_Y} f_X(y-c)\log f_X(y-c)dy \\
&= -\int_{\mathcal{S}_Y} f_Y(y)\log f_Y(y)dy \\
&= h(Y) \\
&= h(X+c).
\end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log|a|.$$

Proof

1. Let $Y = aX$.
2. Then
$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$
and
$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$
3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned}
h(X) \\
&= -\int_{\mathcal{S}_X} f_X(x)\log f_X(x)dx \\
&= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right)\log f_X\left(\frac{y}{a}\right)\frac{dy}{|a|}
\end{aligned}$$

Theorem 10.14 (Translation)

$$h(X+c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then

$$f_Y(y) = f_X(y-c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y-c) \log f_X(y-c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\ &= h(Y) \\ &= h(X+c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$

and

$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$

3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned} h(X) & \\ &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log f_X\left(\frac{y}{a}\right) \frac{dy}{|a|} \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then
$$f_Y(y) = f_X(y - c)$$
and
$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$
3. Letting $x = y - c$, we have

$$\begin{aligned}
h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\
&= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\
&= h(Y) \\
&= h(X + c).
\end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then
$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$
and
$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$
3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned}
h(X) \\
&= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log f_X\left(\frac{y}{a}\right) \frac{dy}{|a|}
\end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned}
h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\
&= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\
&= h(Y) \\
&= h(X + c).
\end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$

and

$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$

3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned}
&h(X) \\
&= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log f_X\left(\frac{y}{a}\right) \frac{dy}{|a|} \\
&= -\int_{\mathcal{S}_Y} \frac{1}{|a|} f_X\left(\frac{y}{a}\right) \left[\log\left(\frac{1}{|a|} f_X\left(\frac{y}{a}\right)\right) + \log |a|\right] dy
\end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.

2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\ &= h(Y) \\ &= h(X + c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.

2. Then

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$

and

$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$

3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned} &h(X) \\ &\quad = -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &\quad = -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log \underline{f_X\left(\frac{y}{a}\right)} \frac{dy}{|a|} \\ &\quad = -\int_{\mathcal{S}_Y} \frac{1}{|a|} f_X\left(\frac{y}{a}\right) \left[\log\left(\frac{1}{|a|} f_X\left(\frac{y}{a}\right)\right) + \log |a|\right] dy \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then
$$f_Y(y) = f_X(y - c)$$
and
$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$
3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\ &= h(Y) \\ &= h(X + c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then
$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$
and
$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$
3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned} h(X) \\ &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log f_X\left(\frac{y}{a}\right) \frac{dy}{|a|} \\ &= -\int_{\mathcal{S}_Y} \frac{1}{|a|} f_X\left(\frac{y}{a}\right) \left[\log\left(\frac{1}{|a|} f_X\left(\frac{y}{a}\right)\right) + \log |a|\right] dy \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X+c) = h(X).$$

Proof

1. Let $Y = X + c$.

2. Then

$$f_Y(y) = f_X(y-c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned}
h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X(y-c) \log f_X(y-c) dy \\
&= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\
&= h(Y) \\
&= h(X+c).
\end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log|a|.$$

Proof

1. Let $Y = aX$.

2. Then

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$

and

$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$

3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned}
h(X) \\
&= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log \underline{f_X\left(\frac{y}{a}\right)} \frac{dy}{|a|} \\
&= -\int_{\mathcal{S}_Y} \frac{1}{|a|} f_X\left(\frac{y}{a}\right) \left[\log\left(\frac{1}{|a|} \underline{f_X\left(\frac{y}{a}\right)}\right) + \log|a|\right] dy
\end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\ &= h(Y) \\ &= h(X + c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$

and

$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$

3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned} h(X) & \\ &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log f_X\left(\frac{y}{a}\right) \frac{dy}{|a|} \\ &= -\int_{\mathcal{S}_Y} \underline{\frac{1}{|a|} f_X\left(\frac{y}{a}\right)} \left[\log\left(\underline{\frac{1}{|a|} f_X\left(\frac{y}{a}\right)}\right) + \log |a|\right] dy \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then
$$f_Y(y) = f_X(y - c)$$
and
$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$
3. Letting $x = y - c$, we have

$$\begin{aligned}
h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\
&= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\
&= h(Y) \\
&= h(X + c).
\end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then
$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$
and
$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$
3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned}
h(X) &\\
&= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log f_X\left(\frac{y}{a}\right) \frac{dy}{|a|} \\
&= -\int_{\mathcal{S}_Y} \frac{1}{|a|} f_X\left(\frac{y}{a}\right) \left[\log\left(\frac{1}{|a|} f_X\left(\frac{y}{a}\right)\right) + \log |a|\right] dy \\
&= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy - \log |a| \int_{\mathcal{S}_Y} f_Y(y) dy
\end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then
$$f_Y(y) = f_X(y - c)$$
and
$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$
3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x)\log f_X(x)dx \\ &= -\int_{\mathcal{S}_Y} f_X(y-c)\log f_X(y-c)dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y)\log f_Y(y)dy \\ &= h(Y) \\ &= h(X+c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log|a|.$$

Proof

1. Let $Y = aX$.
2. Then
$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$
and
$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$
3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned} h(X) & \\ &= -\int_{\mathcal{S}_X} f_X(x)\log f_X(x)dx \\ &= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right)\log f_X\left(\frac{y}{a}\right)\frac{dy}{|a|} \\ &= -\int_{\mathcal{S}_Y} \frac{1}{|a|} f_X\left(\frac{y}{a}\right)\left[\log\left(\frac{1}{|a|} f_X\left(\frac{y}{a}\right)\right) + \underline{\log|a|}\right] dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y)\log f_Y(y)dy - \log|a|\int_{\mathcal{S}_Y} f_Y(y)dy \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y-c) \log f_X(y-c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\ &= h(Y) \\ &= h(X + c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log|a|.$$

Proof

1. Let $Y = aX$.
2. Then

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$

and

$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$

3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned} h(X) \\ &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log f_X\left(\frac{y}{a}\right) \frac{dy}{|a|} \\ &= -\int_{\mathcal{S}_Y} \frac{1}{|a|} f_X\left(\frac{y}{a}\right) \left[\log\left(\frac{1}{|a|} f_X\left(\frac{y}{a}\right)\right) + \underline{\log|a|}\right] dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy - \log|a| \int_{\mathcal{S}_Y} f_Y(y) dy \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\ &= h(Y) \\ &= h(X + c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$

and

$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$

3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned} h(X) & \\ &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log f_X\left(\frac{y}{a}\right) \frac{dy}{|a|} \\ &= -\int_{\mathcal{S}_Y} \frac{1}{|a|} f_X\left(\frac{y}{a}\right) \left[\log\left(\frac{1}{|a|} f_X\left(\frac{y}{a}\right)\right) + \underline{\log |a|}\right] dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy - \underline{\log |a|} \int_{\mathcal{S}_Y} f_Y(y) dy \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X+c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then
$$f_Y(y) = f_X(y-c)$$
and
$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$
3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x)\log f_X(x)dx \\ &= -\int_{\mathcal{S}_Y} f_X(y-c)\log f_X(y-c)dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y)\log f_Y(y)dy \\ &= h(Y) \\ &= h(X+c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log|a|.$$

Proof

1. Let $Y = aX$.
2. Then
$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$
and
$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$
3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned} h(X) & \\ &= -\int_{\mathcal{S}_X} f_X(x)\log f_X(x)dx \\ &= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right)\log f_X\left(\frac{y}{a}\right)\frac{dy}{|a|} \\ &= -\int_{\mathcal{S}_Y} \frac{1}{|a|} f_X\left(\frac{y}{a}\right)\left[\log\left(\frac{1}{|a|} f_X\left(\frac{y}{a}\right)\right) + \log|a|\right] dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y)\log f_Y(y)dy - \log|a| \int_{\mathcal{S}_Y} f_Y(y)dy \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then
$$f_Y(y) = f_X(y - c)$$
and
$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$
3. Letting $x = y - c$, we have

$$\begin{aligned}
h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\
&= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\
&= h(Y) \\
&= h(X + c).
\end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then
$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$
and
$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$
3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned}
& h(X) \\
&= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log f_X\left(\frac{y}{a}\right) \frac{dy}{|a|} \\
&= -\int_{\mathcal{S}_Y} \frac{1}{|a|} f_X\left(\frac{y}{a}\right) \left[\log\left(\frac{1}{|a|} f_X\left(\frac{y}{a}\right)\right) + \log |a|\right] dy \\
&= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy - \log |a| \int_{\mathcal{S}_Y} f_Y(y) dy
\end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.

2. Then

$$f_Y(y) = f_X(y - c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned}
h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\
&= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\
&= h(Y) \\
&= h(X + c).
\end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.

2. Then

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$

and

$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$

3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned}
h(X) & \\
&= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\
&= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log f_X\left(\frac{y}{a}\right) \frac{dy}{|a|} \\
&= -\int_{\mathcal{S}_Y} \frac{1}{|a|} f_X\left(\frac{y}{a}\right) \left[\log\left(\frac{1}{|a|} f_X\left(\frac{y}{a}\right)\right) + \log |a|\right] dy \\
&= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy - \log |a| \int_{\mathcal{S}_Y} f_Y(y) dy \\
&= \underline{h(Y)} - \log |a|
\end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.

2. Then
$$f_Y(y) = f_X(y - c)$$
and
$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\ &= h(Y) \\ &= h(X + c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.

2. Then
$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$
and
$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$

3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned} &h(X) \\ &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log f_X\left(\frac{y}{a}\right) \frac{dy}{|a|} \\ &= -\int_{\mathcal{S}_Y} \frac{1}{|a|} f_X\left(\frac{y}{a}\right) \left[\log\left(\frac{1}{|a|} f_X\left(\frac{y}{a}\right)\right) + \log |a|\right] dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy - \log |a| \int_{\mathcal{S}_Y} f_Y(y) dy \\ &= h(Y) - \log |a| \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X + c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then
$$f_Y(y) = f_X(y - c)$$
and
$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$
3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y - c) \log f_X(y - c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\ &= h(Y) \\ &= h(X + c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log |a|.$$

Proof

1. Let $Y = aX$.
2. Then
$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$
and
$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$
3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log f_X\left(\frac{y}{a}\right) \frac{dy}{|a|} \\ &= -\int_{\mathcal{S}_Y} \frac{1}{|a|} f_X\left(\frac{y}{a}\right) \left[\log\left(\frac{1}{|a|} f_X\left(\frac{y}{a}\right)\right) + \log |a|\right] dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy - \log |a| \int_{\mathcal{S}_Y} f_Y(y) dy \\ &= h(Y) - \log |a| \\ &= h(aX) - \log |a|. \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X+c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then
$$f_Y(y) = f_X(y-c)$$
and
$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$
3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x)\log f_X(x)dx \\ &= -\int_{\mathcal{S}_Y} f_X(y-c)\log f_X(y-c)dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y)\log f_Y(y)dy \\ &= h(Y) \\ &= h(X+c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log|a|.$$

Proof

1. Let $Y = aX$.
2. Then
$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$
and
$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$
3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x)\log f_X(x)dx \\ &= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right)\log f_X\left(\frac{y}{a}\right)\frac{dy}{|a|} \\ &= -\int_{\mathcal{S}_Y} \frac{1}{|a|} f_X\left(\frac{y}{a}\right)\left[\log\left(\frac{1}{|a|} f_X\left(\frac{y}{a}\right)\right) + \log|a|\right]dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y)\log f_Y(y)dy - \log|a|\int_{\mathcal{S}_Y} f_Y(y)dy \\ &= h(Y) - \log|a| \\ &= h(aX) - \log|a|. \end{aligned}$$

Theorem 10.14 (Translation)

$$h(X+c) = h(X).$$

Proof

1. Let $Y = X + c$.
2. Then

$$f_Y(y) = f_X(y-c)$$

and

$$\mathcal{S}_Y = \{x + c : x \in \mathcal{S}_X\}.$$

3. Letting $x = y - c$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X(y-c) \log f_X(y-c) dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy \\ &= h(Y) \\ &= h(X+c). \end{aligned}$$

Theorem 10.15 (Scaling) For $a \neq 0$,

$$h(aX) = h(X) + \log|a|.$$

Proof

1. Let $Y = aX$.
2. Then

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$

and

$$\mathcal{S}_Y = \{ax : x \in \mathcal{S}_X\}.$$

3. Letting $x = \frac{y}{a}$, we have

$$\begin{aligned} h(X) &= -\int_{\mathcal{S}_X} f_X(x) \log f_X(x) dx \\ &= -\int_{\mathcal{S}_Y} f_X\left(\frac{y}{a}\right) \log f_X\left(\frac{y}{a}\right) \frac{dy}{|a|} \\ &= -\int_{\mathcal{S}_Y} \frac{1}{|a|} f_X\left(\frac{y}{a}\right) \left[\log\left(\frac{1}{|a|} f_X\left(\frac{y}{a}\right)\right) + \log|a|\right] dy \\ &= -\int_{\mathcal{S}_Y} f_Y(y) \log f_Y(y) dy - \log|a| \int_{\mathcal{S}_Y} f_Y(y) dy \\ &= h(Y) - \log|a| \\ &= h(aX) - \log|a|. \end{aligned}$$

4. Hence,

$$h(aX) = h(X) + \log|a|.$$