香港中文大學
The Chinese University of Hong Kong

2.3 Continuity of Shannon's Information Measures for Fixed Finite Alphabets

Finite Alphabet vs Countable Alphabet

- All Shannon's information measures are continuous when the alphabets are fixed and finite.

Finite Alphabet vs Countable Alphabet

- All Shannon's information measures are continuous when the alphabets are fixed and finite.
- For countable alphabets, Shannon's information measures are everywhere discontinuous.

Finite Alphabet vs Countable Alphabet

- All Shannon's information measures are continuous when the alphabets are fixed and finite.
- For countable alphabets, Shannon's information measures are everywhere discontinuous.
- To probe further, see Problems 28, 29, 30, 31, and

  S. W. Ho and R. W. Yeung, "On the discontinuity of the Shannon information measures," *IEEE Transactions of Information Theory*, IT-56, no. 12, pp. 5362-5374, 2009.

Definition 2.23 Let p and q be two probability distributions on a common alphabet $\mathcal{X}$. The variational distance between p and q is defined as

$$V(p, q) = \sum_{x \in \mathcal{X}} |p(x) - q(x)|.$$

Definition 2.23 Let p and q be two probability distributions on a common alphabet $\mathcal{X}$. The variational distance between p and q is defined as

$$V(p, q) = \sum_{x \in \mathcal{X}} |p(x) - q(x)|.$$

The entropy function is continuous at p if

$$\lim_{p' \to p} H(p') = H\left(\lim_{p' \to p} p'\right) = H(p),$$

Definition 2.23 Let p and q be two probability distributions on a common alphabet $\mathcal{X}$. The variational distance between p and q is defined as

$$V(p,q) = \sum_{x \in \mathcal{X}} |p(x) - q(x)|.$$

The entropy function is continuous at p if

$$\lim_{p' \to p} H(p') = H\left(\lim_{p' \to p} p'\right) = H(p),$$

or equivalently, for any $\epsilon > 0$, there exists $\delta > 0$ such that

$$|H(p) - H(q)| < \epsilon$$

for all $q \in \mathcal{P}_{\mathcal{X}}$ satisfying

$$V(p,q) < \delta.$$

Definition 2.23 Let p and q be two probability distributions on a common alphabet $\mathcal{X}$. The variational distance between p and q is defined as

$$V(p,q) = \sum_{x \in \mathcal{X}} |p(x) - q(x)|.$$

The entropy function is continuous at p if

$$\lim_{p' \to p} H(p') = H\left(\lim_{p' \to p} p'\right) = H(p),$$

or equivalently, for any $\epsilon > 0$, there exists $\delta > 0$ such that

$$|H(p) - H(q)| < \epsilon$$

for all $q \in \mathcal{P}_{\mathcal{X}}$ satisfying

$$V(p,q) < \delta.$$

Definition 2.23 Let p and q be two probability distributions on a common alphabet $\mathcal{X}$. The variational distance between p and q is defined as

$$V(p,q) = \sum_{x \in \mathcal{X}} |p(x) - q(x)|.$$

The entropy function is continuous at p if

$$\lim_{p' \to p} H(p') = H\left(\lim_{p' \to p} p'\right) = H(p),$$

or equivalently, for any $\epsilon > 0$, there exists $\delta > 0$ such that

$$|H(p) - H(q)| < \epsilon$$

for all $q \in \mathcal{P}_{\mathcal{X}}$ satisfying

$$V(p,q) < \delta.$$

Definition 2.23 Let p and q be two probability distributions on a common alphabet $\mathcal{X}$. The variational distance between p and q is defined as

$$V(p,q) = \sum_{x \in \mathcal{X}} |p(x) - q(x)|.$$

The entropy function is continuous at p if

$$\lim_{p' \to p} H(p') = H\left(\lim_{p' \to p} p'\right) = H(p),$$

or equivalently, for any $\epsilon > 0$, there exists $\delta > 0$ such that

$$|H(p) - H(q)| < \epsilon$$

for all $q \in \mathcal{P}_{\mathcal{X}}$ satisfying

$$V(p,q) < \delta.$$

Definition 2.23 Let p and q be two probability distributions on a common alphabet $\mathcal{X}$. The variational distance between p and q is defined as

$$V(p,q) = \sum_{x \in \mathcal{X}} |p(x) - q(x)|.$$

The entropy function is continuous at p if

$$\lim_{p' \to p} H(p') = H\left(\lim_{p' \to p} p'\right) = H(p),$$

or equivalently, for any $\epsilon > 0$, there exists $\delta > 0$ such that

$$|H(p) - H(q)| < \epsilon$$

for all $q \in \mathcal{P}_{\mathcal{X}}$ satisfying

$$V(p,q) < \delta.$$

An Example

An Example

- Let $\mathcal{X} = \{1, 2, \cdots, \}$, a countably infinite alphabet.

An Example

- Let $\mathcal{X} = \{1, 2, \cdots,\}$, a countably infinite alphabet.
- Let $P_X = \{1, 0, 0, \cdots\}$, and let

An Example

- Let $\mathcal{X} = \{1, 2, \cdots, \}$, a countably infinite alphabet.
- Let $P_X = \{1, 0, 0, \cdots\}$, and let

$$P_{X_n} = \left\{1 - \frac{1}{\sqrt{\log n}}, \underbrace{\frac{1}{n\sqrt{\log n}}, \cdots, \frac{1}{n\sqrt{\log n}}}_{n}, 0, 0, \cdots\right\}.$$

An Example

- Let $\mathcal{X} = \{1, 2, \cdots, \}$, a countably infinite alphabet.
- Let $P_X = \{1, 0, 0, \cdots\}$, and let

$$P_{X_n} = \left\{ 1 - \frac{1}{\sqrt{\log n}}, \underbrace{\frac{1}{n\sqrt{\log n}}, \cdots, \frac{1}{n\sqrt{\log n}}}_{n}, 0, 0, \cdots \right\}.$$

- As $n \to \infty$,

$$V(P_X, P_{X_n}) = \sum_i |P_X(i) - P_{X_n}(i)| = \frac{2}{\sqrt{\log n}} \to 0.$$

An Example

- Let $\mathcal{X} = \{1, 2, \cdots, \}$, a countably infinite alphabet.
- Let $P_X = \{1, 0, 0, \cdots\}$, and let

$$P_{X_n} = \left\{ 1 - \frac{1}{\sqrt{\log n}}, \underbrace{\frac{1}{n\sqrt{\log n}}, \cdots, \frac{1}{n\sqrt{\log n}}}_{n}, 0, 0, \cdots \right\}.$$

- As $n \to \infty$,

$$V(P_X, P_{X_n}) = \sum_i |P_X(i) - P_{X_n}(i)| = \frac{2}{\sqrt{\log n}} \to 0.$$

- However,

$$H\left(\lim_{n\to\infty} P_{X_n}\right) = H(P_X) = 0$$

An Example

- Let $\mathcal{X} = \{1, 2, \cdots, \}$, a countably infinite alphabet.
- Let $P_X = \{1, 0, 0, \cdots\}$, and let

$$P_{X_n} = \left\{ 1 - \frac{1}{\sqrt{\log n}}, \underbrace{\frac{1}{n\sqrt{\log n}}, \cdots, \frac{1}{n\sqrt{\log n}}}_{n}, 0, 0, \cdots \right\}.$$

- As $n \to \infty$,

$$V(P_X, P_{X_n}) = \sum_i |P_X(i) - P_{X_n}(i)| = \frac{2}{\sqrt{\log n}} \to 0.$$

- However,

$$H\left(\lim_{n\to\infty} P_{X_n} \right) = H(P_X) = 0$$

but

$$\lim_{n\to\infty} H(P_{X_n}) = \infty.$$