An introduction to Information Theory

Adrish Banerjee

Department of Electrical Engineering
Indian Institute of Technology Kanpur
Kanpur, Uttar Pradesh
India

Aug. 22, 2016

Lecture #10A: Joint Typical Sequences

Outline of the lecture

- Joint Typical Sequences

Outline of the lecture

- Joint Typical Sequences
- Joint AEP

Joint Typical Sequences

- The set $A_\epsilon^{(n)}$ of jointly typical sequences $\{(x^n, y^n)\}$ with respect to the distribution $p(x, y)$ is the set of n-sequences with empirical entropies ϵ-close to the true entropies, i.e.,

$$\begin{aligned} A_\epsilon^{(n)} &= \{(x^n, y^n) \in X^n \times Y^n \\ & \left| -\frac{1}{n}\log p(x^n) - H(X) \right| < \epsilon, \\ & \left| -\frac{1}{n}\log p(y^n) - H(Y) \right| < \epsilon, \\ & \left| -\frac{1}{n}\log p(x^n, y^n) - H(X, Y) \right| < \epsilon, \end{aligned}$$

where

$$p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i)$$

Joint Typical Sequences

- Equivalently, if $\{(x^n, y^n)\}$ belongs to jointly typical set, then following holds for $\epsilon > 0$,

$$\begin{aligned} 2^{-n(H(X)+\epsilon)} &\le p(x^n) &\le 2^{-n(H(X)-\epsilon)} \\ 2^{-n(H(Y)+\epsilon)} &\le p(y^n) &\le 2^{-n(H(Y)-\epsilon)} \\ 2^{-n(H(X,Y)+\epsilon)} &\le p(x^n, y^n) &\le 2^{-n(H(X,Y)-\epsilon)} \end{aligned}$$

Joint AEP

- Let (X^n, Y^n) be sequences of length n drawn i.i.d. according to $p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i)$. Then

Joint AEP

- Let (X^n, Y^n) be sequences of length n drawn i.i.d. according to $p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i)$. Then

  1) $Pr((X^n, Y^n) \in A_\epsilon^{(n)} \to 1$ as $n \to \infty$.

Joint AEP

- Let (X^n, Y^n) be sequences of length n drawn i.i.d. according to $p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i)$. Then
  1) $Pr((X^n, Y^n) \in A_\epsilon^{(n)} \to 1$ as $n \to \infty$.
  2) $|A_\epsilon^{(n)}| \leq 2^{n(H(X,Y)+\epsilon)}$

Joint AEP

- Let (X^n, Y^n) be sequences of length n drawn i.i.d. according to $p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i)$. Then
  1) $Pr((X^n, Y^n) \in A_\epsilon^{(n)} \to 1$ as $n \to \infty$.
  2) $|A_\epsilon^{(n)}| \leq 2^{n(H(X,Y)+\epsilon)}$
  3) If $(\tilde{X}^n, \tilde{Y}^n) \sim p(x^n)p(y^n)$, i.e., $\tilde{X}^n$ and $\tilde{Y}^n$ are independent with the same marginals as $p(x^n, y^n)$, then

$$Pr((\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}) \leq 2^{-n(I(X;Y)-3\epsilon)}$$

  Also, for sufficiently large n

$$Pr((\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}) \geq (1-\epsilon)2^{-n(I(X;Y)+3\epsilon)}$$

Joint AEP

- By weak law of large numbers, we have

$$-\frac{1}{n}\log p(X^n) \to -E[\log p(X)] = H(X) \text{ in probability}$$

Joint AEP

- By weak law of large numbers, we have

$$-\frac{1}{n}\log p(X^n) \to -E[\log p(X)] = H(X) \text{ in probability}$$

- Hence, given $\epsilon > 0$, there exists n_1, such that for all $n > n_1$,

$$Pr\left(\left|-\frac{1}{n}\log p(X^n) - H(X)\right| > \epsilon\right) < \frac{\epsilon}{3}$$

Joint AEP

- By weak law of large numbers, we have

$$-\frac{1}{n}\log p(X^n) \to -E[\log p(X)] = H(X) \text{ in probability}$$

- Hence, given $\epsilon > 0$, there exists n_1, such that for all $n > n_1$,

$$Pr\left(\left|-\frac{1}{n}\log p(X^n) - H(X)\right| > \epsilon\right) < \frac{\epsilon}{3}$$

- Similarly, we have

$$-\frac{1}{n}\log p(Y^n) \to -E[\log p(Y)] = H(Y) \text{ in probability}$$

Joint AEP

- Also, we have

$$-\frac{1}{n}\log p(X^n, Y^n) \to -E[\log p(X, Y)] = H(X, Y) \text{ in probability}$$

and there exists n_2 and n_3 such that for all $n \geq n_2$

$$Pr\left(\left|-\frac{1}{n}\log p(Y^n) - H(Y)\right| > \epsilon\right) < \frac{\epsilon}{3}$$

and for all $n \geq n_3$

$$Pr\left(\left|-\frac{1}{n}\log p(X^n, Y^n) - H(X, Y)\right| > \epsilon\right) < \frac{\epsilon}{3}$$

Joint AEP

- Also, we have

$$-\frac{1}{n}\log p(X^n, Y^n) \to -E[\log p(X,Y)] = H(X,Y) \text{ in probability}$$

and there exists n_2 and n_3 such that for all $n \geq n_2$

$$Pr\left(\left|-\frac{1}{n}\log p(Y^n) - H(Y)\right| > \epsilon\right) < \frac{\epsilon}{3}$$

and for all $n \geq n_3$

$$Pr\left(\left|-\frac{1}{n}\log p(X^n, Y^n) - H(X,Y)\right| > \epsilon\right) < \frac{\epsilon}{3}$$

- For $n > \max(n_1, n_2, n_3)$, the probability of the set $A_\epsilon^{(n)}$ is greater than $1 - \epsilon$.

Joint AEP

- We know that

$$\begin{aligned} 1 &= \sum p(x^n, y^n) \\ &\geq \sum_{A_\epsilon^{(n)}} p(x^n, y^n) \\ &\geq |A_\epsilon^{(n)}| 2^{-n(H(X,Y)+\epsilon)} \end{aligned}$$

and hence

$$|A_\epsilon^{(n)}| \leq 2^{n(H(X,Y)+\epsilon)}$$

Joint AEP

- For large n, we have $Pr(A_\epsilon^{(n)}) \geq 1 - \epsilon$. Hence we have

$$\begin{aligned} 1 - \epsilon &\leq \sum_{(x^n, y^n) \in A_\epsilon^{(n)}} p(x^n, y^n) \\ &\leq |A_\epsilon^{(n)}| 2^{-n(H(X,Y)-\epsilon)} \end{aligned}$$

and

$$|A_\epsilon^{(n)}| \geq (1 - \epsilon) 2^{n(H(X,Y)-\epsilon)}$$

Joint AEP

- If $\tilde{X}^n$ and $\tilde{Y}^n$ are independent and have same marginals as X^n and Y^n, then we have

$$\begin{aligned} Pr((\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}) &= \sum_{(x^n, y^n) \in A_\epsilon^{(n)}} p(x^n) p(y^n) \\ &\leq 2^{n(H(X,Y)+\epsilon)} 2^{-n(H(X)-\epsilon)} 2^{-n(H(Y)-\epsilon)} \\ &= 2^{-n(I(X;Y)-3\epsilon)} \end{aligned}$$

Joint AEP

- If $\tilde{X}^n$ and $\tilde{Y}^n$ are independent and have same marginals as X^n and Y^n, then we have

$$
\begin{aligned}
Pr((\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}) &= \sum_{(x^n, y^n) \in A_\epsilon^{(n)}} p(x^n)p(y^n) \\
&\leq 2^{n(H(X,Y)+\epsilon)} 2^{-n(H(X)-\epsilon)} 2^{-n(H(Y)-\epsilon)} \\
&= 2^{-n(I(X;Y)-3\epsilon)}
\end{aligned}
$$

- Similarly we can show that

$$
\begin{aligned}
Pr((\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}) &= \sum_{A_\epsilon^{(n)}} p(x^n)p(y^n) \\
&\geq (1-\epsilon) 2^{n(H(X,Y)-\epsilon)} 2^{-n(H(X)+\epsilon)} 2^{-n(H(Y)+\epsilon)} \\
&= (1-\epsilon) 2^{-n(I(X;Y)+3\epsilon)}
\end{aligned}
$$

Joint Typical Sequences

- There are about $2^{nH(X)}$ typical X sequences and about $2^{nH(Y)}$ typical Y sequences. However, there are only $2^{nH(X,Y)}$ jointly typical sequences.

Joint Typical Sequences

- There are about $2^{nH(X)}$ typical X sequences and about $2^{nH(Y)}$ typical Y sequences. However, there are only $2^{nH(X,Y)}$ jointly typical sequences.
- Not all pairs of typical X^n and typical Y^n are also jointly typical. The probability of any randomly chosen pair being jointly typical is about $2^{-nI(X;Y)}$.

Joint Typical Sequences

- There are about $2^{nH(X)}$ typical X sequences and about $2^{nH(Y)}$ typical Y sequences. However, there are only $2^{nH(X,Y)}$ jointly typical sequences.
- Not all pairs of typical X^n and typical Y^n are also jointly typical. The probability of any randomly chosen pair being jointly typical is about $2^{-nI(X;Y)}$.
- Hence we can consider about $2^{nI(X;Y)}$ such pairs before we are likely to come across a jointly typical pair.

Joint Typical Sequences

- There are about $2^{nH(X)}$ typical X sequences and about $2^{nH(Y)}$ typical Y sequences. However, there are only $2^{nH(X,Y)}$ jointly typical sequences.
- Not all pairs of typical X^n and typical Y^n are also jointly typical. The probability of any randomly chosen pair being jointly typical is about $2^{-nI(X;Y)}$.
- Hence we can consider about $2^{nI(X;Y)}$ such pairs before we are likely to come across a jointly typical pair.
- Hence, there are about $2^{nI(X;Y)}$ distinguishable signals X^n.