

Conditional entropy

Properties of entropy

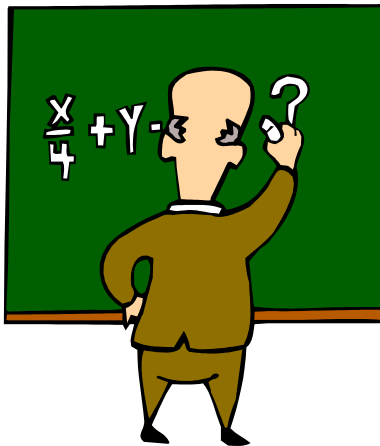


CS349 Cryptography

Department of Computer Science
Wellesley College

Properties of entropy

- In this lecture, we prove some fundamental results concerning entropy which we apply next time to cryptosystems.
- We begin with a result, known as **Jensen's inequality**.



Concave functions

A real-valued function f is a *concave function* on an interval I if

$$f\left(\frac{x+y}{2}\right) \geq \frac{f(x) + f(y)}{2}$$

for all $x, y \in I$. f is a *strictly concave function* on an interval I if

$$f\left(\frac{x+y}{2}\right) > \frac{f(x) + f(y)}{2}$$

for all $x, y \in I, x \neq y$.

Conditional entropy 7-3

Jensen's Inequality

Suppose f is a continuous strictly concave function on the interval I ,

$$\sum_{i=1}^n a_i = 1$$

and $a_i > 0, 1 \leq i \leq n$. Then

$$\sum_{i=1}^n a_i f(x_i) \leq f\left(\sum_{i=1}^n a_i x_i\right)$$

where $x_i \in I, 1 \leq i \leq n$. Further, equality occurs if and only if $x_1 = x_2 = \dots = x_n$.

Conditional entropy 7-4

Maximum information content

Suppose X is a random variable having a probability distribution which takes on the values p_1, p_2, \dots, p_n , where $p_i > 0, 1 \leq i \leq n$. Then

$$H(\mathbf{X}) \leq \log_2 n,$$

with equality if and only if $p_i = 1/n, 1 \leq i \leq n$.

Conditional entropy 7-5

Applying Jensen's inequality*, we have

$$\begin{aligned} H(\mathbf{X}) &= - \sum_{i=1}^n p_i \log_2 p_i \\ &= - \sum_{i=1}^n p_i \log_2 \frac{1}{p_i} \\ &= - \log_2 \sum_{i=1}^n p_i \frac{1}{p_i} \\ &= \log_2 n \end{aligned}$$

*Further, equality occurs if and only if $p_i = 1/n, 1 \leq i \leq n$.

Conditional entropy 7-6

Joint distributions

- o The information content of a joint distribution is not more than sum of the information contents of individual distributions.

- o In particular,

$$H(\mathbf{X}, \mathbf{Y}) \leq H(\mathbf{X}) + H(\mathbf{Y})$$

with equality if and only if \mathbf{X} and \mathbf{Y} are independent random variables.

Conditional entropy 7-7

Conditional entropy

Suppose \mathbf{X} and \mathbf{Y} are random variables. Then for any fixed value y of \mathbf{Y} , we get a conditional probability distribution on \mathbf{X} ; we denote the associated random variable by $\mathbf{X}|y$.

$$H(\mathbf{X} | y) = \sum_x \Pr[x | y] \log_2 \Pr[x | y].$$

Define the **conditional entropy**, $H(\mathbf{X} | \mathbf{Y})$, to be the weighted average (with respect to the probabilities $\Pr[y]$) of the entropies $H(\mathbf{X}|y)$ over all possible values y . It is computed to be

$$H(\mathbf{X} | \mathbf{Y}) = \sum_y \Pr[y] \sum_x \Pr[x | y] \log_2 \Pr[x | y].$$

Conditional entropy 7-8

Two homework assignments

Theorem.

$$H(\mathbf{X}, \mathbf{Y}) = H(\mathbf{Y}) + H(\mathbf{X}|\mathbf{Y}).$$

Corollary.

$$H(\mathbf{X}|\mathbf{Y}) \leq H(\mathbf{X}),$$

with equality if and only if \mathbf{X} and \mathbf{Y} are independent.