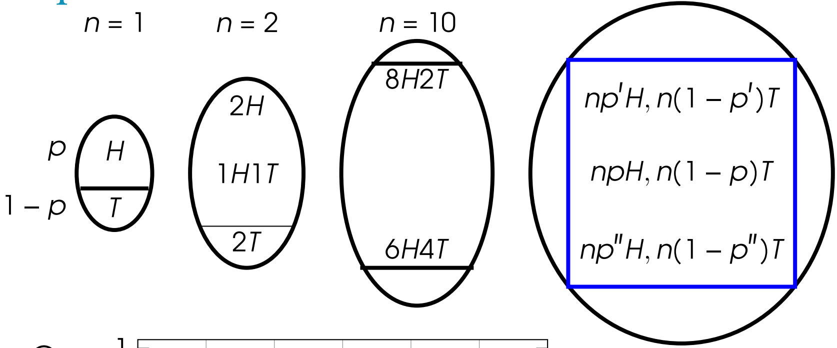
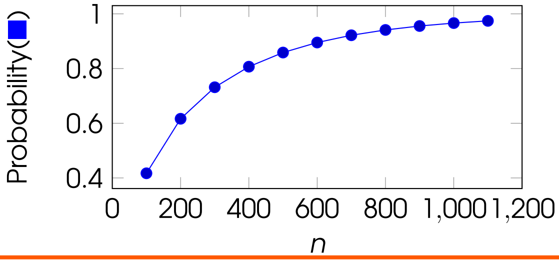
EE708: Information Theory and Coding

Sibi Raj B. Pillai

Department of Electrical Engineering, Indian Institute of Technology Bombay srbpteach@gmail.com

Compressed Idea





$p = 0.74$

$p' = p(1 + \epsilon)$

$p'' = p(1 - \epsilon)$

$\epsilon = 0.04$

Discrete Finite Source

$X_1^n := (X_1, X_2, \dots, X_n) := X^n.$

X_1^n generated IID

$$\mathcal{T}_\epsilon^n := \left\{ x^n : \left|\frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{\{x_i = a\}} - p(a)\right| \le \epsilon p(a) \right\}$$

Shannon's Idea:

Encode the set $\blacksquare$ (i.e. $\{x^n \in \mathcal{T}_\epsilon^n\}$) with an additional flagbit.

From $X_m, m \ge 1$, encode the successive n symbols and proceed.

Following Shannon, the average number of bits required is:

$$A_n = Pr(X^n \in \blacksquare)\,(1 + \lceil \log_2 |\blacksquare| \rceil) + Pr(X^n \notin \blacksquare)(1 + n).$$

The number of bits per symbol $\frac{1}{n}A_n \to \frac{1}{n}\log(|\blacksquare|)$.

Cardinality of $\blacksquare$

For a given x^n, let $N_a := \sum \mathbb{I}_{\{x_i = a\}}$ for each $a \in \mathcal{X}$.

$$p(x^n) = \prod_{a \in \mathcal{X}} p(a)^{N_a}$$

For each $x^n \in \blacksquare$, $N_a \approx np_a(1 \pm \epsilon)$:

$$p(x^n) \geq \prod_{a \in \mathcal{X}} p(a)^{np(a)(1+\epsilon)}$$

Since $Pr(\blacksquare) \leq 1$:

$$|\blacksquare| \leq \prod_{a \in \mathcal{X}} p(a)^{-np(a)(1+\epsilon)}$$

$$= 2^{n(1+\epsilon) \sum\limits_{a \in \mathcal{X}} p(a) \log_2 \frac{1}{p(a)}} = 2^{n(1+\epsilon)H(X)}.$$

A random variable $X \in \mathcal{X}$ has entropy $H(X) := \sum\limits_{x \in \mathcal{X}} p(x) \log \dfrac{1}{p(x)}$.