

Como medir a *incerteza* ?

Exemplo 1

2 fontes; uma pode gerar 3 símbolos distintos (a, b, c) e a outra 2 símbolos (1, 2).

Combinados: a1, a2, b1, b2, c1, c2

Se as 2 fontes geram em paralelo, geram uma incerteza de 6 símbolos.

$\log_{\text{base}} (\text{no. símbolos possíveis}) = \text{incerteza}$

$$\log 3 + \log 2 = \log 6$$

Se uma fonte gera somente 1 símbolo $\rightarrow \log_2 1 = 0$ bits (de incerteza) NÃO HÁ INCERTEZA!

Se uma fonte gera somente 2 símbolos $\rightarrow \log_2 2 = 1$ bit

Exemplo 2 (Elwyn Edwards)

gerador de 8 letras: A a H \rightarrow A, B, C, D, E, F, G, H

estratégias

1. qual a próxima letra ? \rightarrow abordagem exaustiva
2. dividir o conjunto em 2 :
supondo que é a letra D, então:

Pergunta	Divisão	Resposta
1. Antes do E?	ABCD - EFGH	Sim
2. Antes do C?	AB - CD	Não
3. Antes do D?	C - D	Não

Solução: D.

Máximo de 3 perguntas!

Se o gerador gerasse 16 letras (A a P)

Pergunta	Divisão	Resposta
1. Antes do I?	ABCDEFGH - IJKLMNOP	Sim
2. Antes do E?	ABCD - EFGH	Sim
3. Antes do C?	AB - CD	Não
4. Antes do D?	C - D	Não

Solução: D.

Máximo de 4 perguntas!

Grandeza do Conjunto (n)	Número de Perguntas (H)
1	0
2	1
4	2
8	3
16	4

$$H(A \text{ a } H) = 3 \text{ unidades por letra}$$

$$H(A \text{ a } P) = 4 \text{ unidades por letra}$$

$$n = 2^H$$

$$H = \log_2 n$$

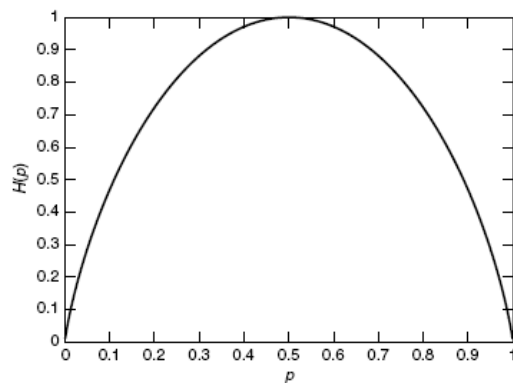
Levando-se em consideração M símbolos,

$$\begin{aligned}\log_2(M) &= -\log_2(M^{-1}) \\ &= -\log_2\left(\frac{1}{M}\right) \\ &= -\log_2(P)\end{aligned}\quad \sum_{i=1}^M P_i = 1.$$

$$u_i = -\log_2(P_i).$$

$$N = \sum_{i=1}^M N_i, \quad \frac{\sum_{i=1}^M N_i u_i}{\sum_{i=1}^M N_i}, \quad \sum_{i=1}^M \frac{N_i}{N} u_i, \quad H = \sum_{i=1}^M P_i u_i.$$

$$H = -\sum_{i=1}^M P_i \log_2 P_i \quad (\text{bits per symbol}).$$



$H(p)$ vs. p .

$$H_{\text{equiprobable}} = -\sum_{i=1}^M \frac{1}{M} \log_2 \frac{1}{M}$$

$$\begin{aligned}H_{\text{equiprobable}} &= -\left(\frac{1}{M} \log_2 \frac{1}{M}\right) \sum_{i=1}^M 1 \\ &= -\left(\frac{1}{M} \log_2 \frac{1}{M}\right) M \\ &= -\log_2 \frac{1}{M} \\ &= \log_2 M\end{aligned}$$

Qual o significado de um sinal ter, por exemplo, 1,67 bits por símbolo: significa que ao converter o sinal original numa string de bits (1's e 0's), haverá em **média** 1,67 dígitos binários para cada símbolo do sinla original.