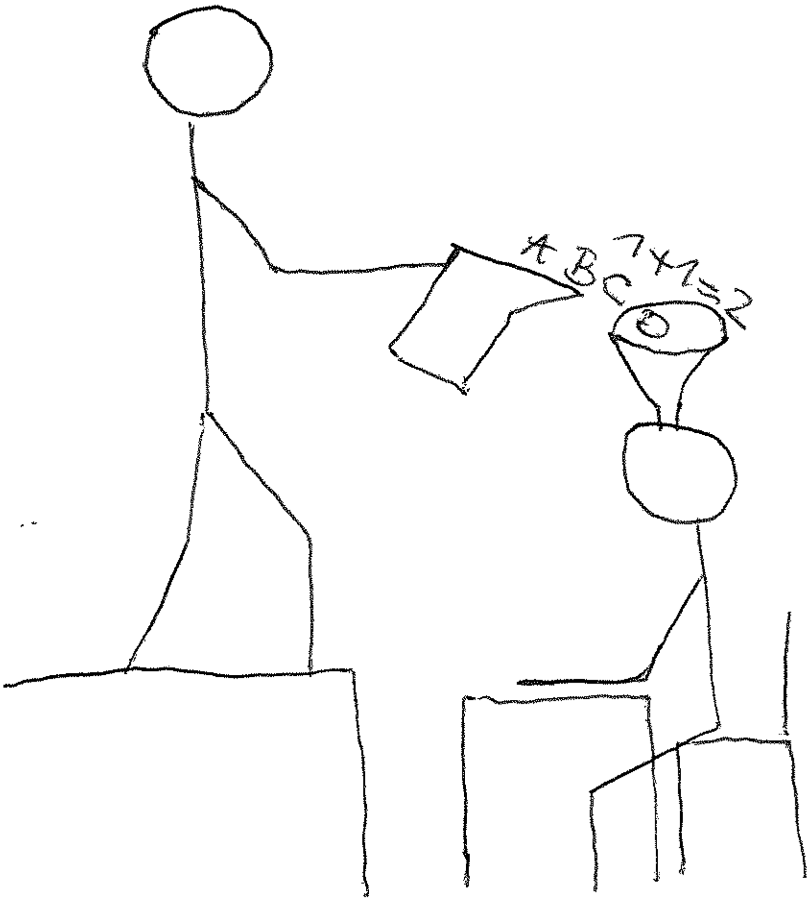
Learning in 17th Century



Nuremberg Funnel

Entropy $H(P_1, \dots, P_n)$

1) n elements each probability $P_i = \frac{1}{n}$

- n groups of single element.
- need $\log n$ bits to distinguish the groups
- on average $\sum_{i=1}^{n} P_i \cdot \log n = \log n$ bits

2) Different probabilities $P_i = \frac{m_i}{n'}$

- each group has m_i elements, n' elements in total
- different elements in group yield same code
- need $\log m_i$ bits to discriminate in group.

Expected:

$$\sum_{i=1}^{n} P_i (\log n' - \log m_i) = \sum_{i=1}^{n} P_i \cdot \log \frac{n'}{m_i}$$

$$= -\sum P_i \cdot \log P_i$$

Lottery: Euro Millions

choose 5 numbers from 1, 2, ..., 50

$\Rightarrow \binom{50}{5}$ outcomes

<u>naive encoding</u>:

sequence, eg. 1, 4, 20, 30, 46

- 6 bits per number ($2^6 = 64$)
- needs $5 \times 6 = 30$ bits

<u>Entropy</u>: $H(p_1, \dots, p_m)$, $m = \binom{50}{5}$ $p_i = 1/m$

$$H(p_1, \dots, p_m) = \log_2 \binom{50}{5}$$

Recall: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

$$\binom{50}{5} = \frac{50!}{45!\,5!} = \frac{46 \cdot 47 \cdot 48 \cdot 49 \cdot 50}{5!} = 5 \cdot 12 \cdot 46 \cdot 47 \cdot 49$$

$$\leq 64^4 = (2^6)^4 = 2^{24}, \geq 32^4 = (2^5)^4 = 2^{20}$$

between 20 and 24 bits (exact 22,560 bits)