Quantifying Information
(Claude Shannon, 1948)

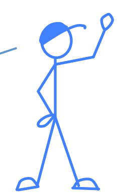
Given discrete random variable X

- N possible values: $x_1, x_2, \ldots, x_N$
- Associated probabilities: $p_1, p_2, \ldots, p_N$

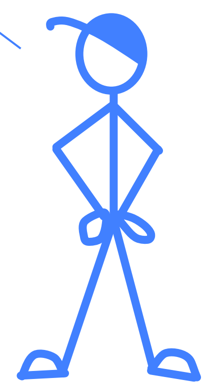
Information received when learning that choice was x_i:

$$I(x_i) = \log_2\left(\frac{1}{p_i}\right)$$

$1/p_i$ is proportional to the uncertainty of choice x_i.



Information is measured in bits (binary digits) = number of 0/1's required to encode choice(s)



Information Conveyed by Data

Even when data doesn't resolve all the uncertainty

$$I(\text{data}) = \log_2\left(\frac{1}{p_{\text{data}}}\right) \qquad \text{e.g., } I(\text{heart}) = \log_2\left(\frac{1}{13/52}\right) = 2 \text{ bits}$$

Common case: Suppose you're faced with N equally probable choices, and you receive data that narrows it down to M choices. The probability that data would be sent is M·(1/N) so the amount of information you have received is

$$I(\text{data}) = \log_2\left(\frac{1}{M \cdot (1/N)}\right) = \log_2\left(\frac{N}{M}\right) \text{ bits}$$

Example: Information Content

Examples:

- information in one coin flip:

  N= **2** M= **1** Info content= $\log_2(2/1) = 1$ **bit**

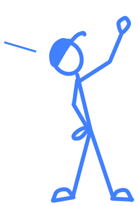
- card drawn from fresh deck is a heart:

  N= **52** M= **13** Info content= $\log_2(52/13) = 2$ **bits**

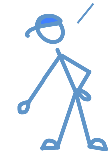
- roll of 2 dice:

  N= **36** M= **1** Info content= $\log_2(36/1) = 5.17$

.17 bits ???



Probability & Information Content



data	p_{data}	$\log_2(1/p_{data})$
a heart	13/52	2 bits
not the Ace of spades	51/52	0.028 bits
a face card (J, Q, K)	12/52	2.115 bits
the "suicide king"	1/52	5.7 bits