Examples where diversity is beneficial as a secondary goal

- Games of strategy, e.g., “balancing your range” in poker

- Construction of investment portfolios (hedging)

- Evolution (genetic diversity)

The utility of imagining an adversary…

Examples where simplicity is beneficial as a secondary goal

Occam's razor:

"*Among competing hypotheses, the one with the fewest assumptions should be selected.*"

Jaynes' Principle of maximum entropy:

"*Given some data, among all hypothetical probability distributions that agree with the data, the one of maximum entropy best represents the current state of knowledge.*"

The Shannon entropy

If X is a random variable taking values in a finite state space Ω, we define the **Shannon entropy of X** by

$$H(X) := \sum_{x \in \Omega} \mathbb{P}[X = x] \log \frac{1}{\mathbb{P}[X = x]}$$

(with the contention that $0 \log 0 = 0$). Also, we will use "log" for the base-2 logarithm, except when we use it for the natural logarithm…

- If X denotes a random message from some distribution, then the average number of bits needed to communicate (or compress) X is $\approx H(X)$

- English text has between 0.6 and 1.3 bits of entropy per character.

The Shannon entropy

If X is a random variable taking values in a finite state space Ω, we define the **Shannon entropy of X** by

$$H(X) \coloneqq \sum_{x \in \Omega} \mathbb{P}[X = x] \log \frac{1}{\mathbb{P}[X = x]}$$

(with the contention that $0 \log 0 = 0$). Also, we will use "log" for the base-2 logarithm, except when we use it for the natural logarithm…

The **probability mass function of X** is given by $p(x) = \mathbb{P}[X = x]$.
We will also write $H(p)$.

Important fact: H is a **concave** function of p.

$$H(p) \coloneqq \sum_{x \in \Omega} p(x) \log \frac{1}{p(x)}$$

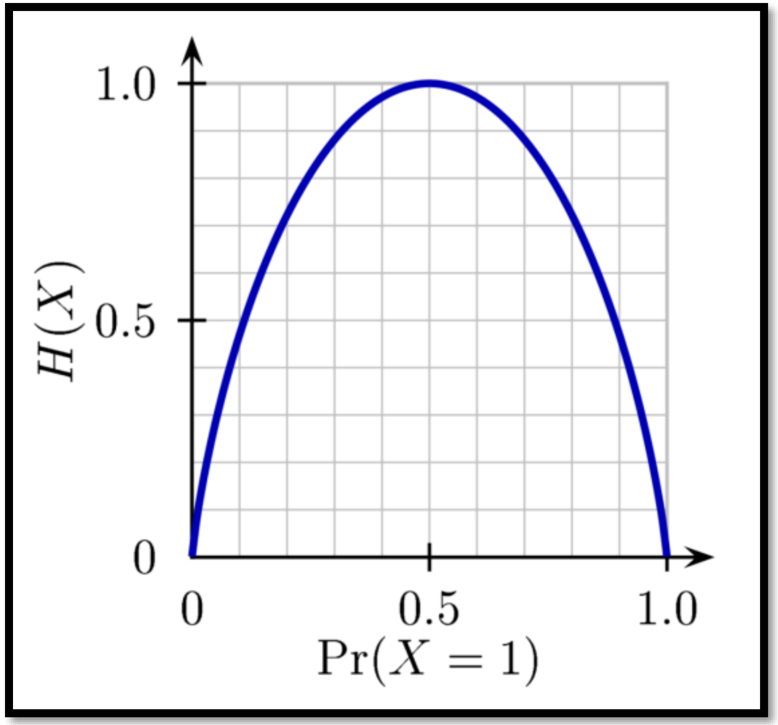
H is a **strictly concave** function of p.

The Shannon entropy

If X is a random variable taking values in a finite state space Ω, we define the **Shannon entropy of X** by

$$H(X) := \sum_{x \in \Omega} \mathbb{P}[X = x] \log \frac{1}{\mathbb{P}[X = x]}$$

Outcome of a presidential poll vs. outcome of a fair coin flip



If X is a random variable taking values in a finite state space Ω, we define the **Shannon entropy of X** by

$$H(X) \coloneqq \sum_{x \in \Omega} \mathbb{P}[X = x] \log \frac{1}{\mathbb{P}[X = x]}$$

Suppose there are n possible outcomes $\Omega = \{1, 2, \ldots, n\}$.

What's the maximum entropy of X?

The universe is maximizing entropy

two applications today

- **Part I:** Entropy to encourage simplicity: Matrix scaling
- **Part II:** Entropy to encourage diversification: Caching and paging