

# Entropy and Relative Entropy

## Stochastic Processes Seminar

Erik Stern

December 9, 2025

# Agenda

- 1 Entropy Basics
- 2 Inequalities
- 3 Relative Entropy
- 4 Numerics
- 5 Wrap-up

# Roadmap

- 1 Entropy Basics
- 2 Inequalities
- 3 Relative Entropy
- 4 Numerics
- 5 Wrap-up

## Definition (Shannon Entropy)

Let  $X$  be a discrete random variable over  $\mathcal{X}$  with pmf  $p$ . The entropy is

$$H(X) := - \sum_{x \in \text{supp}(X)} p(x) \log_2 p(x).$$

# Definitions aligned with the paper

## Definition (Shannon Entropy)

Let  $X$  be a discrete random variable over  $\mathcal{X}$  with pmf  $p$ . The entropy is  $H(X) := -\sum_{x \in \text{supp}(X)} p(x) \log_2 p(x)$ .

## Definition (Mutual Information)

For jointly distributed  $X$  and  $Y$ , define  $I(X; Y) := D(p_{(X,Y)} \| p_X p_Y)$ .

# Entropy of a weighted coin

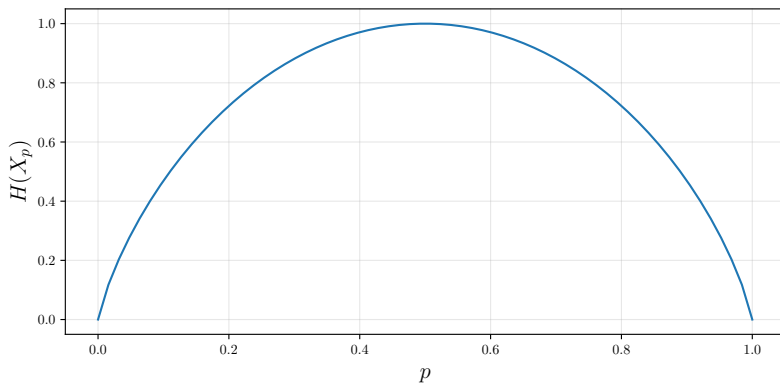


Figure 1: Re-use figures generated via `python scripts/entropy_weighted_coin.py`.

# Roadmap

- 1 Entropy Basics
- 2 Inequalities**
- 3 Relative Entropy
- 4 Numerics
- 5 Wrap-up

# Concavity of entropy

## Theorem (Concavity)

*For any discrete pmf  $p$ , the function  $p \mapsto H(p)$  is concave, i.e.  
 $H(\lambda p + (1 - \lambda)q) \geq \lambda H(p) + (1 - \lambda)H(q)$  for  $\lambda \in [0, 1]$ .*

## Idea.

Apply Jensen's Inequality to the convex function  $x \mapsto -x \log x$  as detailed in the writeup. □



# Roadmap

- 1 Entropy Basics
- 2 Inequalities
- 3 Relative Entropy**
- 4 Numerics
- 5 Wrap-up

## Definition (Relative Entropy)

Let  $p$  and  $q$  be pmfs on  $\mathcal{Z}$ . Define  $D(p\|q) := \sum_{z \in \mathcal{Z}} p(z) \log \frac{p(z)}{q(z)}$  with the entropy conventions (e.g.,  $0 \log 0 = 0$ ) from the paper.

- This matches the mutual-information definition recorded in `writeup/root_document.tex`.
- Summarize the limit-case conventions from the accompanying remark in the writeup as speaker notes.

# Pointwise relative entropy

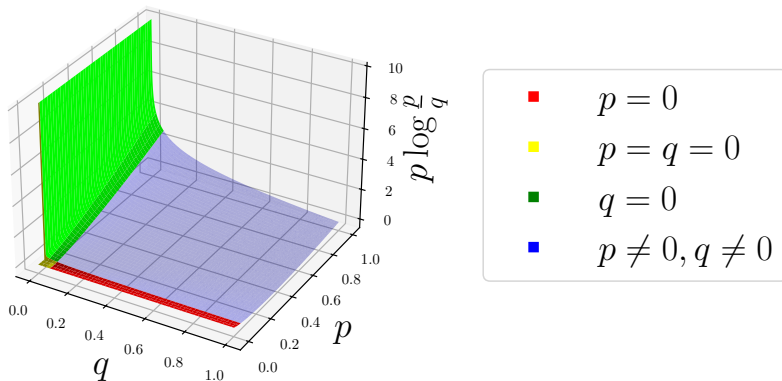


Figure 2: Visualise  $(p, q) \mapsto \log \frac{p}{q}$  to connect plots with theory.

# Roadmap

- 1 Entropy Basics
- 2 Inequalities
- 3 Relative Entropy
- 4 Numerics**
- 5 Wrap-up

## How to regenerate figures

- Run `python scripts/recreate_all.py` to rebuild every PDF in `plots/`.
- Target a single experiment via `python scripts/relative_entropy_binomial_distributions.py`.
- Keep data-free slides by sourcing images from `../plots/*.pdf` only.

## Example

Let  $X_p \sim G(p)$  denote the Geometric r.v. discussed in the paper. Compare  $H(X_p)$  to the Bernoulli example using the geometric entropy figure showcased in the writeup.



## Remark

Store any extra derivations in appendix frames guarded by `\appendix` to keep the talk tight.

# Roadmap

- 1 Entropy Basics
- 2 Inequalities
- 3 Relative Entropy
- 4 Numerics
- 5 Wrap-up**

# References

-  T. Cover and J. Thomas, *Elements of Information Theory*, Wiley, 2006.
-  Seminar notes in `writeup/root_document.tex` provide full proofs.



## Next steps

- Sync new macros with the writeup before adding fresh content.
- Rebuild the slide deck via `bash presentation/compile_latex.sh`.
- Drop PDFs into `output/` for distribution.