

# Introduction to Machine Learning



## Information Theory

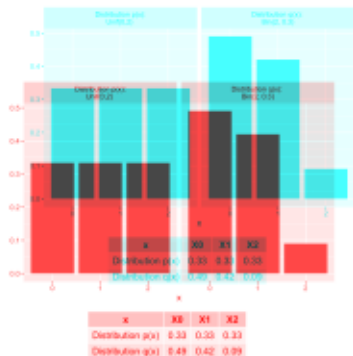
## KL and Maximum Entropy

### Learning goals

- Know the defining properties of the KL

### Learning goals

- Understand the relationship between the maximum entropy principle and minimum KL
- Understand the relationship between the maximum entropy principle and minimum discrimination information
- Understand the relationship between Shannon entropy and relative entropy



# PROBLEMS WITH DIFFERENTIAL ENTROPY

Differential entropy compared to the Shannon entropy:

- Differential entropy can be negative
- Differential entropy is not invariant to coordinate transformations

⇒ Differential entropy is not an uncertainty measure and can not be meaningfully used in a maximum entropy framework.



In the following, we derive an alternative measure, namely the KL divergence (relative entropy), that fixes these shortcomings by taking an inductive inference viewpoint. [▶ Caticha 2004](#)