

# Introduction to Machine Learning

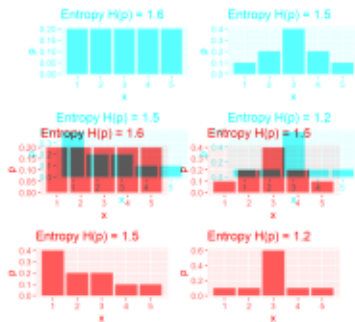


## Information Theory

### Joint Entropy and Mutual Information I

#### Learning goals

- Know the joint entropy
- Know conditional entropy as remaining uncertainty
- Know the joint entropy
- Know mutual information as the amount of information obtained by another
- Know conditional entropy as remaining uncertainty
- Know mutual information as the amount of information of an RV obtained by another



# JOINT AND CONDITIONAL ENTROPY

The following relations hold:

$$H(X, X) = H(X)$$

$$H(X|X) = 0$$

$$H((X, Y)|Z) = H(X|Z) + H(Y|(X, Z))$$

Which can all be trivially derived from the previous considerations.

Furthermore, if  $H(X|Y) = 0$  and  $X, Y$  are discrete RV, then  $X$  is a function of  $Y$ , so for all  $y$  with  $p(y) > 0$ , there is only one  $x$  with  $p(x, y) > 0$ . Proof is not hard, but also not completely trivial.

