(Shannon) Entropy:

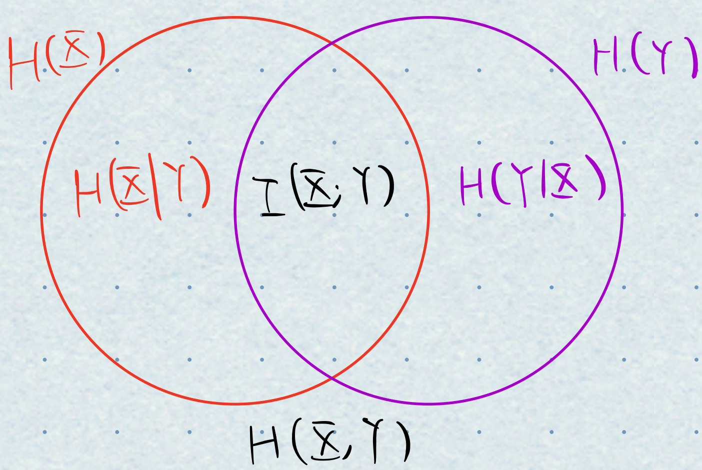
$$H(X) = -\sum_{i=1}^{n} P(x_i) \log P(x_i)$$

Conditional entropy:

$$H(Y|X) = \sum_{x} P_X(x) H(Y|X=x)$$

$$H(Y|X=x) = -\sum_{y} P_{Y|X}(y|x) \log P_{Y|X}(y|x)$$

$$\Rightarrow H(Y|X) = \sum_{x,y} p(x,y) \log \frac{p(x)}{p(x,y)}$$



Chain rule: $H(X_1, \cdots, X_n) = \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1})$

Mutual information

$$I(X;Y) = \sum_{x,y} P_{X,Y}(x,y) \log \frac{P_{X,Y}(x,y)}{P_X(x)\,P_Y(y)}$$

$$I(X;Y) = D_{KL}(P_{X,Y} \,||\, P_X P_Y)$$

Relative entropy (KL divergence)

$$D_{KL}(P\,||\,Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)}$$

Cross entropy

$$CE(P,Q) = -\sum_x P(x) \log Q(x)$$
$$= H(P) + D_{KL}(P\,||\,Q)$$

Diversity: Hill number

$${}^{q}D = \left(\sum_{i=1}^{R} p_i^{q}\right)^{\frac{1}{q-1}}$$

$${}^{1}D = \exp\left\{-\sum_{i=1}^{R} p_i \ln(p_i)\right\}$$ Shannon index