Definition 2.28 The informational divergence between two probability distributions p and q on a common alphabet $\mathcal{X}$ is defined as

$$D(p\|q) = \sum_x p(x) \log \frac{p(x)}{q(x)} = E_p \log \frac{p(X)}{q(X)},$$

where E_p denotes expectation with respect to p.