

Assume both \mathcal{X} and \mathcal{Y} are finite.

Definition 7.6 The capacity of a discrete memoryless channel $p(y|x)$ is defined as

$$C = \max_{p(x)} I(X; Y),$$

where X and Y are respectively the input and the output of the generic discrete channel, and the maximum is taken over all input distributions $p(x)$.