Entropy and Mutual Information

M.Tech CS
Information and Coding Theory

Before Mid Semester

1 Entropy

Definition: 1. *Entropy is the minimum descriptive complexity of a random variable. In other word, it is measure of the uncertainty of a random variable.*

Formally,
If ***X*** be a discrete random variable over alphabet χ and probabilty mass function is p(x)=Pr{ $X = x$ }= x $x \epsilon \chi$.
The Entropy *H(χ)* of a discrete random variable ***X*** is defined by

$$H(X) = -\sum_{x \epsilon \chi} p(x) log_2 p(x) \tag{1}$$

Entropy is expressed in bit. When entropy is measured in natural base (ln) it is called *nats*.
Some useful properties

1. $H(X) \geq 0$
2. $H_b(X) = (\log_b a H_a(X)$.
3. $H(X) = -E \log p(x) = E \frac{1}{\log p(X)}$ where E($\bullet$) is expection of given function.

1.1 Joint Entropy and Conditional Entropy

Definition: 2. *The* ***joint entropy*** $H(X,Y)$ *of a pair of discrete random variables* (X,Y) *with joint distribution* $p(x,y)$ *is defined as*

$$H(X,Y) = -\sum_{x \epsilon \chi} \sum_{y \epsilon \gamma} p(x,y) log_2 p(x,y) \tag{2}$$

$$H(X,Y) = -E log_2 p(X,Y) \tag{3}$$

Definition: 3. *if (X,Y) $\sim$ p(x,y), then* ***conditional entropy*** $H(Y|X)$ *is defined as*

$$H(Y|X) = \sum_{x \epsilon \chi} p(x) H(Y|X = x) \tag{4}$$

$$= \tag{5}$$