Definitions

Afroz Mohiuddin

September 6, 2013

Contents

1 Distributions

1.1 Binomial Distribution

Suppose we toss a coin n times, let X be the number of heads, if the probability of heads is θ, then we say that X has a binomial distribution, $X \sim Bin(n, \theta)$

$$Bin(k|n, \theta) \triangleq \binom{n}{k} \theta^k (1-\theta)^{n-k}$$

$$\text{mean} = n\theta, \text{var} = n\theta(1-\theta)$$

2 Information Theory

2.1 Entropy

Lack of information about a random variable.

$$H(Y) = -\sum_{y} p(y) \ln p(y)$$

2.2 Joint Entropy

!!!! Interpretation !!!!

$$H(X,Y) = -\sum_{x,y} p(x,y) \ln p(x,y)$$

2.3 Mutual Information

Measure of information overlap between two random variables.

$$I(X;Y) = \sum_{x,y} p(x,y) \ln \frac{p(x,y)}{p(x)p(y)}$$

The information overlap between X and Y is 0 when the two variables are independent. When X determines Y, $I(X;Y) = H(Y)$.

$I(X;Y)$ reaches its maximum value value, when X and Y are perfectly correlated, i.e. they determine each other.

2.4 Pointwise Mutual Infomation (PMI)

Measure of how much the actual probability of a particular co-occurrence of events $p(x,y)$ differs from what we would expect it to be on the basis of the probabilities of individual events and the assumption of independence, $p(x)p(y)$

$$i(x,y) = \ln \frac{p(x,y)}{p(x)p(y)}$$

Even though PMI maybe positive or negative, its expected output over all joint events, i.e. MI, is positive.

2.5 Identities

Information overlap is sum of entropies as well as expected value of PMI.

$$
\begin{aligned}
I(X;Y) &= H(X) + H(Y) - H(X,Y) \\
I(X;Y) &= E_{p(X,Y)}[i(X,Y)] \\
&= \sum_{x,y} p(x,y) i(x,y)
\end{aligned}
$$