

Consider flipping a coin. Then x is a binary variable, either “head” or “tail”. Suppose it is a fair coin. Then we have $p(x) = 1/2$ for either “head” or “tail”. It follows that

$$I(x) = -\log_2 p(x) = 1,$$

for either “head” or “tail”.

When x and y are independent, we have

$$p(x, y) = p(x)p(y).$$

It follows that

$$I(x, y) = -\log_2 p(x, y) = -\log_2 p(x) - \log_2 p(y) = I(x) + I(y).$$

Consider the case of throwing a dice. So 1,2,3,4,5,6 is equally likely to occur. It follows that

$$H(X) = \sum_{x=1}^6 -p(x) \log_2 p(x) = \sum_{x=1}^6 -\frac{1}{6} \log_2 1/6 = \log_2 6.$$

Note that $\log_2 6 > 1$.

If the coin is not normal and has both sides. Then the entropy is equal to

$$H(X) = -\sum_x p(x) \log_2 p(x) = -1 \times \log_2 1 = 0.$$