

















## **Entropy measure**

- If log is base 2, unit of information is the bit it two equally likely
  possibilities exist for a variable, then both the formula and common
  sense indicate 1 bit of information is needed to specify the
  variable's value.
- If log is base 10, unit of information is the digit.
- If there are N different values for X with equal probability (uniform distribution) 1/N, then the entropy is just –log(1/N) which is log(N). So for a flat uniform distribution the entropy (information) increases with the number of possibilities.
- On the other hand, suppose probabilities are all zero except for a single value of X, which always occurs – maximally peaked distribution, for this, entropy is 0 (since log(1) = 0).





• Now if this unit simply sums its inputs then output Y has values that could be -2,0,0, and +2 respectively.

• There are 3 values of Y with probabilities *P*(-2)=1/4,*P*(0)=1/2, *P*(+2)=1/4 (using the rule for combining probabilities for mutually exclusive events) – entropy of Y is:

$$= \left(\frac{1}{4}\log\frac{1}{4} + \frac{1}{2}\log\frac{1}{2} + \frac{1}{4}\log\frac{1}{4}\right)$$
  
= 1.5

• I.e. we lost information in the addition ...

