

## Multi-class Entropy

Last time, you saw this equation for entropy for a bucket with  $m$  red balls and  $n$  blue balls:

$$entropy = -\frac{m}{m+n} \log_2\left(\frac{m}{m+n}\right) - \frac{n}{m+n} \log_2\left(\frac{n}{m+n}\right)$$

We can state this in terms of probabilities instead for the number of red balls as  $p_1$  and the number of blue balls as  $p_2$ :

$$p_1 = \frac{m}{m+n}$$

$$p_2 = \frac{n}{m+n}$$

$$entropy = -p_1 \log_2(p_1) - p_2 \log_2(p_2)$$

This entropy equation can be extended to the multi-class case, where we have three or more possible values:

$$entropy = -p_1 \log_2(p_1) - p_2 \log_2(p_2) - \dots - p_n \log_2(p_n) = -\sum_{i=1}^n p_i \log_2(p_i)$$

The minimum value is still 0, when all elements are of the same value. The maximum value is still achieved when the outcome probabilities are the same, but the upper limit increases with the number of different outcomes. (For example, you can verify the maximum entropy is 2 if there are four different possibilities, each with probability 0.25.)

If we have a bucket with eight red balls, three blue balls, and two yellow balls, what is the entropy of the set of balls? Input your answer to at least three decimal places.



1.334

RESET

NEXT