Tutorial 4 Feb 27, 2021

Entropy (H)

- A measure of uncertainty with a RV
- Measured in bits
- Something that is less likely to occure has higher entropy.
- denote the minimum # of bits to encode Y

Applications: Decision trees, Kullback-Leiber, cross entropy loss

Calculation

For discrete RV with outcomes $\{1, \dots, k\}$

$$H = -\sum_{c=1}^{k} p_c \log_2 p_c \quad , \quad p_c = P(y=c)$$

Ex. For degenerate RV $Y=2$

$$H = -\sum_{c=1}^{k} p_c \log_2 c = -p_2 \log_2 p_2 = -\log 1 = 0$$

Ex. For uniform RV $Y = \{1, \dots, 8\}$

$$H = -\sum_{c=1}^{8} p_c \log_2 p_c = 8\left(\frac{1}{8}\log_2 \frac{1}{8}\right) = 3$$