Escalante, Tomas

Collaborator → Roberto Ventura

PS0

0) a)

$$\underset{2\times 2}{\begin{bmatrix} 5 & 2 \\ 1 & 0 \end{bmatrix}} \underset{2\times 1}{\begin{bmatrix} 3 \\ 3 \end{bmatrix}} \rightarrow \underset{2\times 1}{} = \begin{bmatrix} 15+6 \\ 3+0 \end{bmatrix} = \begin{bmatrix} 21 \\ 3 \end{bmatrix}$$

b)

$$\underset{2\times 2}{\begin{bmatrix} 5 & 2 \\ 1 & 0 \end{bmatrix}} \underset{1\times 2}{\begin{bmatrix} 3 & 3 \end{bmatrix}} \rightarrow \text{invalid}$$

1) $-\sum_{x\in X} P(x) \ln P(x)$

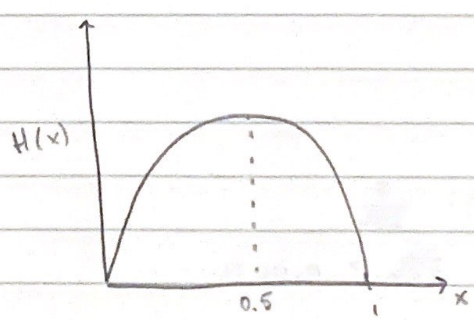
a) $P(x) = [0.2, 0.3, 0.5]$

$$-[0.2\ln 0.2 + 0.3\ln 0.3 + 0.5\ln 0.5] = -[-0.322 - 0.361 - 0.347]$$
$$= 1.03$$

b) $P(x) = \frac{1}{m}$

$$-\sum_{x\in[1,m]} \frac{1}{m}\ln\frac{1}{m} = -\int_1^m \frac{1}{m}\ln\frac{1}{m}\,dm = \frac{1}{2}\ln^2 m$$

c)



the peak is at 0.5
its minimum is at 0
this makes sense
because there are only
2 outcomes, 0 or 1

2) a) $p(-) = \frac{2}{3}$

$p(\text{great} \mid -) = \frac{2+1}{15+4} = \frac{3}{19}$

$p(\text{good} \mid -) = \frac{2+1}{15+4} = \frac{3}{19}$

$p(\text{terrible} \mid -) = \frac{1+1}{15+4} = \frac{2}{19}$

$p(\text{sentence} \mid -) = \left(\frac{3}{19}\right)\left(\frac{3}{19}\right)\left(\frac{2}{19}\right) = \frac{18}{6859}$

$p(- \mid \text{sentence}) = \frac{2}{3}\left(\frac{18}{6859}\right) = \frac{12}{6859} = 0.00175$

$p(+) = \frac{1}{3}$

$p(\text{great} \mid +) = \frac{5+1}{9+4} = \frac{6}{13}$

$p(\text{good} \mid +) = \frac{3+1}{9+4} = \frac{4}{13}$

$p(\text{terrible} \mid +) = \frac{0+1}{9+4} = \frac{1}{13}$

$p(\text{sentence} \mid +) = \left(\frac{6}{13}\right)\left(\frac{4}{13}\right)\left(\frac{1}{13}\right) = \frac{12}{1183}$

$p(+ \mid \text{sentence}) = \frac{1}{3}\left(\frac{12}{1183}\right) = \frac{4}{1183} = 0.00338$

b) $p(-) = \frac{2}{3}$

$p(\text{great} \mid -) = \frac{1+1}{7+4} = \frac{2}{11}$

$p(\text{good} \mid -) = \frac{2+1}{7+4} = \frac{3}{11}$

$p(\text{terrible} \mid -) = \frac{1+1}{7+4} = \frac{2}{11}$

$p(\text{sentence} \mid -) = \left(\frac{2}{11}\right)\left(\frac{3}{11}\right)\left(\frac{2}{11}\right) = \frac{12}{1331}$

$p(- \mid \text{sentence}) = \frac{2}{3}\left(\frac{12}{1331}\right) = \frac{8}{1331} = 0.00601$

$p(+) = \frac{1}{3}$

$p(\text{great} \mid +) = \frac{2+1}{4+4} = \frac{3}{8}$

$p(\text{good} \mid +) = \frac{1+1}{4+4} = \frac{1}{4}$

$p(\text{terrible} \mid +) = \frac{0+1}{4+4} = \frac{1}{8}$

$p(\text{sentence} \mid +) = \left(\frac{3}{8}\right)\left(\frac{1}{4}\right)\left(\frac{1}{8}\right) = \frac{3}{256}$

$p(+ \mid \text{sentence}) = \frac{1}{3}\left(\frac{3}{256}\right) = \frac{1}{256} = 0.00391$

The two models don't agree, (a) → (+), (b) → (-).
I prefer the second model (b) because it doesn't favor the number of words, but rather the presence of (+) or (-) words within a sentence