CS4248
AY 2022/23 Semester 1
Tutorial 3 Solutions

1. (a) $P = P_I^{N_I}(1 - P_I)^{N-N_I}$

(b) $\frac{\partial P}{\partial P_I} = N_I \cdot P_I^{N_I-1} \cdot (1 - P_I)^{N-N_I} + P_I^{N_I} \cdot (N - N_I) \cdot (1 - P_I)^{N-N_I-1} \cdot (-1) = 0$

$P_I^{N_I} \cdot (1 - P_I)^{N-N_I} \cdot (N_I \cdot P_I^{-1} - (N - N_I) \cdot (1 - P_I)^{-1}) = 0$

$$\frac{N_I}{P_I} = \frac{N - N_I}{1 - P_I}$$

$N_I - N_I \cdot P_I = N \cdot P_I - N_I \cdot P_I$

$$P_I = \frac{N_I}{N}$$

2. Let $C^*(w_i w)$ be the smoothed bigram count of $w_i w$.

If $C(w_i w) > 0$, $P(w|w_i) = \frac{C(w_i w)}{C(w_i)+T(w_i)} = \frac{C^*(w_i w)}{C(w_i)}$

If $C(w_i w) = 0$, $P(w|w_i) = \frac{T(w_i)}{(V-T(w_i))(C(w_i)+T(w_i))} = \frac{C^*(w_i w)}{C(w_i)}$

Given:

$C(w_1) = 4000$, $C(w_2) = 2500$

$T(w_1) = 150$, $T(w_2) = 80$

$V = 15000$

$$C^*(w_1 w_1) = \frac{C(w_1 w_1) \times C(w_1)}{C(w_1) + T(w_1)} = \frac{100 \times 4000}{4000 + 150} = 96.4$$

$$C^*(w_1 w_3) = \frac{C(w_1 w_3) \times C(w_1)}{C(w_1) + T(w_1)} = \frac{30 \times 4000}{4000 + 150} = 28.9$$

$$C^*(w_1 w_2) = \frac{T(w_1) \times C(w_1)}{(V - T(w_1))(C(w_1) + T(w_1))} = \frac{150 \times 4000}{(15000 - 150)(4000 + 150)} = 0.00973$$

The other smoothed counts are computed similarly, shown below:

	w_1	w_2	w_3
w_1	96.4	0.00973	28.9
w_2	0.00520	48.4	0.00520

3. True. Proof:
RHS
= H(X) + H(Y|X)

$$= -\sum_x p(x)\log p(x) - \sum_x \sum_y p(x,y)\log p(y|x)$$

$$= -\sum_x p(x)\log p(x) - \sum_x \sum_y p(x,y)\log \frac{p(x,y)}{p(x)}$$

$$= -\sum_x p(x)\log p(x) - \sum_x \sum_y \{p(x,y)\log p(x,y) - p(x,y)\log p(x)\}$$

$$= -\sum_x p(x)\log p(x) - \sum_x \sum_y p(x,y)\log p(x,y) + \sum_x \sum_y p(x,y)\log p(x)$$

Since

$$\sum_x \sum_y p(x,y)\log p(x) = \sum_x \left\{\log p(x) \left\{\sum_y p(x,y)\right\}\right\} = \sum_x \log p(x) p(x)$$

Hence, RHS = $-\sum_x \sum_y p(x,y)\log p(x,y)$ = H(X,Y) = LHS

4.

$$v(T1, w1) = \frac{1}{5} \times \frac{1}{20} = \frac{1}{100}$$

$$v(T2, w1) = 0$$

$$v(T3, w1) = \frac{4}{5} \times \frac{1}{10} = \frac{2}{25}$$

$$v(T1, w2) = \max\left(\frac{1}{100} \times 0, 0, \underline{\frac{2}{25} \times \frac{3}{5}}\right) \times \frac{1}{10} = \frac{3}{625}$$

$$v(T2, w2) = \max\left(\frac{1}{100} \times \frac{5}{6}, 0, \underline{\frac{2}{25} \times \frac{1}{5}}\right) \times \frac{1}{10} = \frac{1}{625}$$

$$v(T3, w2) = \max\left(\frac{1}{100} \times 0, 0, \underline{\frac{2}{25} \times \frac{1}{5}}\right) \times \frac{1}{10} = \frac{1}{625}$$

$$\max\left(\underline{\frac{3}{625} \times \frac{1}{6}}, \frac{1}{625} \times \frac{1}{8}, \frac{1}{625} \times 0\right) = \frac{1}{1250}$$

Optimal sequence of POS tags: w1/T3   w2/T1