Inline Question 1

$$L_i(W) = -\log\left(P(Y=y^{(i)} \mid X=x^{(i)})\right)$$

$$\downarrow$$

$$\frac{e^{W_{y^{(i)}} x^{(i)}}}{\sum_j e^{W_j x^{(i)}}}$$

* Weights are randomly initialized & roughly $= 0$

$$W_j x^{(i)} \approx 0 \quad \forall j$$

$$e^{W_j x^{(i)}} \approx 1 \quad \forall j$$

So $L_i(W) = \frac{1}{1+1+\cdots 1} = \frac{1}{10} = 0.1$

Inline Question 2

True: SVM loss of an example can Basily be zero, if Score given to true class is atleast one more than remaining.

But for Softmax loss to be Zero Probability to give to true class should be 1, which is not possible.