

**Dilemma:**

$$(a) E[g(y|x)] = \sum g(y|x) p(y|x)$$

- Or -

$$(b) E[g(y|x)] = \sum g(y|x) p(y, x)$$

Assert that (b) is correct.

**Proof:**

The definition of the expectation of a random variable  $g(X)$  is

$$E_p g(X) = \sum_{x \in X} g(x) p(x)$$

And can be generalized to two a distribution of two random variables, for example, when used to express joint entropy:

$$H(X, Y) = - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x, y) = -E \log p(X, Y)$$

Here,  $g(x)$  has been replaced by  $g(x, y) = \log p(x, y)$ .

If instead we let  $g(x, y) = g(y|x) = \log p(y|x)$  as with conditional entropy, we arrive at

$$H(Y|X) = - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(Y|X) = -E \log p(Y|X)$$

Which clearly follows (b)