Ex-5.1

- We have K classes : $C_1, C_2, \ldots, C_K$
- The target vector uses a 1-of-K encoding

$$t = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_K \end{bmatrix} \quad \text{where} \quad t_K = \begin{cases} 1, & \text{if } x \in C_K \\ 0, & \text{otherwise} \end{cases}$$

$$E[t|x] = \sum_t t\, p(t|x)$$

But t can take only K possible values:

$t^{(1)} = [1, 0, 0, \ldots, 0]^T, \ldots$

so, each target vector $t^{(k)}$ uniquely corresponds to a class C_K.

$$\therefore \quad p(t^K|x) = p(C_K|x)$$

$$\mathbb{E}[t|x] = \sum_{K=1}^{K} t^{(k)}.\ p(t^{(k)}|x)$$

$$= \sum_{K=1}^{K} t^{(k)}.\ p(C_K|x)$$

t^K will be 0 everywhere except 1 at the K^{th} pos.

$$\therefore \quad \mathbb{E}[t|x] = \begin{bmatrix} p(C_1|x) \\ p(C_2|x) \\ \vdots \\ p(C_K|x) \end{bmatrix}$$