My solutions to

Deep Learning: Foundations and Concepts

Dario Miro Konopatzki

5 Single-Layer Networks: Classification

5.1

For all $1 \leq k \leq K$:

$$\begin{aligned}
\mathbb{E}\left[T_k | X = x\right] &= \sum_{t \in \{0,1\}} t \, p_{T_k|X=x}(t) \\
&= 0 \cdot p_{T_k|X=x}(0) + 1 \cdot p_{T_k|X=x}(1) \\
&= p_{T_k|X=x}(1) \\
&= \begin{cases} 1 & \text{if class of } x \text{ is } c_k \\ 0 & \text{otherwise} \end{cases} \\
&= p_{C|X=x}(c_k)
\end{aligned}$$

Hence $\mathbb{E}\left[T | X = x\right] = \left(p_{C|X=x}(c_k)\right)_{1 \leq k \leq K}$.