Exercise 2.

(a). $\text{Error}(\hat{f}(x)) = E[(\hat{f}(x) - f(x))^2]$, $\text{Bias}(\hat{f}(x)) = E[\hat{f}(x) - f(x)]$,

$\text{Var}(\hat{f}(x)) = E[(\hat{f}(x) - E[\hat{f}(x)])^2]$

$$
\begin{aligned}
\text{Error}(\hat{f}(x)) &= E[(\hat{f}(x) - f(x))^2] \\
&= E[(\hat{f}(x) - E(\hat{f}(x)) + E(\hat{f}(x)) - f(x))^2] \\
&= E\{[\hat{f}(x) - E(\hat{f}(x))]^2 + 2[\hat{f}(x) - E(\hat{f}(x))][E(\hat{f}(x)) - f(x)] + [E(\hat{f}(x)) - f(x)]^2\} \\
&= E[(\hat{f}(x) - E[\hat{f}(x)])^2] + 2E[\hat{f}(x) - E(\hat{f}(x))][E(\hat{f}(x)) - f(x)] + E[E(\hat{f}(x)) - f(x)]^2 \\
&= \text{Var}(\hat{f}(x)) + 2\times 0 + [E(\hat{f}(x)) - f(x)]^2 \quad (\because E(E(\hat{f}(x))) = E(\hat{f}(x)), E(f(x)) = f(x)) \\
&= \text{Var}(\hat{f}(x)) + \text{Bias}(\hat{f}(x))^2
\end{aligned}
$$

Exercise 3.

(a). $D_{KL}(R \| \hat{p}) = \sum_{i=1}^{c} R(i) \log \frac{R(i)}{\hat{p}(i)}$

$$\min_R E[D_{KL}(R\|\hat{p})] = \min_R \left[\sum_{i=1}^{c} R(i) \log \frac{R(i)}{\hat{p}(i)}\right]$$

s.t. $\sum_{i=1}^{c} R(i) = 1$.

$$\mathcal{L} = \sum_{i=1}^{c} R(i) \log R(i) - \sum_{i=1}^{c} R(i) \log(\hat{p}(i)) + \lambda\left[\sum_{i=1}^{c} R(i) - 1\right]$$

$$\frac{\partial \mathcal{L}}{\partial R(i)} = \log R(i) + 1 - \log(\hat{p}(i)) + \lambda = 0 \quad (*), \quad i = 1, 2, \cdots, c.$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \sum_{i=1}^{c} R(i) - 1 = 0. \quad (**)$$

From $(*)$, we have $R(i) = \exp[\log(\hat{p}(i)) - \lambda - 1) = \frac{\exp[\log \hat{p}(i)]}{\exp(\lambda+1)}$ $(***)$

Plug it into $(**)$, we can obtain

$$\sum_{i=1}^{c} \exp(\log(\hat{p}(i)))/\exp(\lambda+1) = 1$$

$$\Rightarrow \exp(\lambda+1) = \sum_{i=1}^{c} \exp \log(\hat{p}(i))$$

Plug it back into $(***)$, we have the solution to the optimization problem

$$R(i) = \frac{\exp E[\log \hat{p}(i)]}{\sum_j \exp E \log \hat{p}(j)}$$