(b). $\text{Error}(\hat{p}) = E\left(\sum_{i=1}^{C} P(i) \log \frac{P(i)}{\hat{p}(i)}\right) = \sum_{i=1}^{C} P(i) \log P(i) - \sum_{i=1}^{C} P(i) E[\log \hat{P}(i)]$ (1)

$\text{Bias}(\hat{p}) = D_{KL}(P \| R) = \sum_{i=1}^{C} P(i) \log \frac{P(i)}{R(i)} = \sum_{i=1}^{C} P(i) \log P(i) - \sum_{i=1}^{C} P(i) \log R(i)$ (2)

$\text{Var}(\hat{p}) = E[D_{KL}(R \| \hat{p})] = E\left[\sum_{i=1}^{C} R(i) \log \frac{R(i)}{\hat{p}(i)}\right] = \sum_{i=1}^{C} R(i) \log R(i) - \sum_{i=1}^{C} R(i) E[\log(\hat{p}(i))]$ (3)

From (1), (2), (3), we know to prove

$\text{Error}(\hat{p}) = \text{Bias}(\hat{p}) + \text{Var}(\hat{p})$ is equivalent to prove

$$\sum_{i=1}^{C} P_i \log P(i) - \sum_{i=1}^{C} P(i) E[\log \hat{P}(i)] = \sum_{i=1}^{C} P(i) \log P(i) - \sum_{i=1}^{C} P(i) \log R(i) + \sum_{i=1}^{C} R(i) \log R(i) - \sum_{i=1}^{C} R(i) E[\log(\hat{P}(i))]$$

$$\Leftrightarrow \quad \sum_{i=1}^{C} P(i) E[\log \hat{P}(i)] = \sum_{i=1}^{C} P(i) \log R(i) + \sum_{i=1}^{C} R(i) E[\log \hat{P}(i)] - \sum_{i=1}^{C} R(i) \log R(i) \quad (4)$$

After insert $R(i)$ from (a),

$$\text{The RHS of (4)} = \sum_{i=1}^{C} \left[P(i) - \frac{\exp E[\log \hat{P}(i)]}{\sum_{j=1}^{C} \exp E(\log \hat{P}(j))}\right] \cdot \left[E[\log \hat{P}(i)] - \sum_{j=1}^{C} E(\log \hat{P}(j))\right] + \sum_{i=1}^{C} \frac{\exp E[\log \hat{P}(i)]}{\sum_{j=1}^{C} \exp E[\log \hat{P}(j)]} \cdot E[\log \hat{P}(i)]$$

$$= \sum_{i=1}^{C} P(i) E[\log \hat{P}(i)] - \sum_{i=1}^{C} \left[P(i) \sum_{j=1}^{C} E(\log \hat{P}(j))\right] - \cancel{\sum_{i=1}^{C} \frac{\exp E[\log \hat{P}(i)]}{\sum_{j=1}^{C} \exp E(\log \hat{P}(j))} \cdot E[\log \hat{P}(i)]} + \sum_{i=1}^{C} \left[\frac{\exp E[\log \hat{P}(i)]}{\sum_{j=1}^{C} \exp E(\log \hat{P}(j))} \cdot \sum_{j=1}^{C} E(\log(\hat{P}(j)))\right] + \cancel{\sum_{i=1}^{C} \frac{\exp E[\log \hat{P}(i)]}{\sum_{j=1}^{C} \exp E[\log \hat{P}(j)]} \cdot E[\log \hat{P}(i)]}$$

$$= \sum_{i=1}^{C} P(i) E[\log \hat{P}(i)] - \cancel{\sum_{j=1}^{C} E[\log \hat{P}(j)] \cdot \underbrace{\sum_{i=1}^{C} P(i)}_{=1}} + \cancel{\sum_{j=1}^{C} E[\log \hat{P}(j)] \underbrace{\sum_{i=1}^{C} \frac{\exp E[\log \hat{P}(i)]}{\sum_{j=1}^{C} \exp E[\log \hat{P}(j)]}}_{=1}}$$

$$= \sum_{i=1}^{C} P(i) E[\log \hat{P}(i)]$$

$$= \text{LHS of (4)}.$$

which proves $\text{Error}(\hat{p}) = \text{Bias}(\hat{p}) + \text{Var}(\hat{p})$