

5.15

C

5.16

(1)

$$\begin{aligned}H(D) &= -\sum_{i=1}^n p(x_i) \log p(x_i) \\&= -\frac{4}{9} \log \frac{4}{9} - \frac{5}{9} \log \frac{5}{9} \\&= \frac{4}{9} \log \frac{9}{4} + \frac{5}{9} \log \frac{9}{5}\end{aligned}$$

(2)

由

$$\begin{aligned}H(D) &= \frac{4}{9} \log \frac{9}{4} + \frac{5}{9} \log \frac{9}{5} \\H(D|a_1) &= \frac{4}{9} \left(-\frac{3}{4} \log \frac{3}{4} - \frac{1}{4} \log \frac{1}{4} \right) + \frac{5}{9} \left(-\frac{1}{5} \log \frac{1}{5} - \frac{4}{5} \log \frac{4}{5} \right) \\H(D|a_2) &= \frac{5}{9} \left(-\frac{2}{5} \log \frac{2}{5} - \frac{3}{5} \log \frac{3}{5} \right) + \frac{4}{9} \left(-\frac{2}{4} \log \frac{2}{4} - \frac{2}{4} \log \frac{2}{4} \right) \\H(D|a_3) &= \frac{4}{9} \left(-\frac{3}{4} \log \frac{3}{4} - \frac{1}{4} \log \frac{1}{4} \right) + \frac{5}{9} \left(-\frac{1}{5} \log \frac{1}{5} - \frac{4}{5} \log \frac{4}{5} \right)\end{aligned}$$

所以 a_1, a_2, a_3 的信息增益分别是

$$\begin{aligned}H(D) - H(D|a_1) \\H(D) - H(D|a_2) \\H(D) - H(D|a_3)\end{aligned}$$

5.17

B,C,D