HW6. Data Science Application Homework

2018711118
데이터사이언스융합학과
오윤경

Consider 30 emails → 6 each to 5 users.
10/30 (1/3) = emails are spam.

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

조건부확률이기때문에 분모 ⇒ 한눈에 구함.

Q2. P(G|F)?

$$= \frac{{}_6C_5 \left(\frac{1}{3}\right)^5 \left(\frac{2}{3}\right)^1}{{}_{10}C_6 \left(\frac{1}{3}\right)^6 \left(\frac{2}{3}\right)^0} = 0.0571$$

Q3. ① P(H|F)

$$= \frac{{}_6C_1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^5}{{}_6C_{10} \left(\frac{1}{3}\right)^6 \left(\frac{2}{3}\right)^0} = 0.91426$$

② P(E|F)

$$= \frac{{}_6C_3 \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^3}{{}_{10}C_6 \left(\frac{1}{3}\right)^6 \left(\frac{2}{3}\right)^0} = 0.7619$$

③ P(H)

$$= {}_6C_1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^5 = 0.26337$$

Compare ①.② P(H|F) vs P(E|F)

0.91426 0.7619

∴ P(H|F) > P(E|F)

explain
→ 6개의 스팸을 받을F 의 확률인 경우, H가 공통된 1개의 스팸을 받을확률은 > 6개의 스팸을 받을F 의 여러 공통된 3개의 스팸메일을 받을확률.

Compare ①.③ P(H|F) vs P(H)

0.91426 0.26337

∴ P(H|F) > P(H)

Q4. ① P(F|H)

$$= \frac{{}_1C_1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^5}{{}_{10}C_1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^5} = 0.1$$

② P(H|F)

$$= \frac{{}_6C_1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^5}{{}_{10}C_6 \left(\frac{1}{3}\right)^6 \left(\frac{2}{3}\right)^0} = 0.9142857\cdots$$

Compare to ① ② P(H|F) vs P(F|H)

0.9142 0.1

∴ P(H|F) > P(F|H)

→ 6개의 스팸메일 받을F의 경우에 대해서, 공통된 1개의 스팸메일을 받을 확률은 "0.9142"

→ 1개의 스팸메일을 받은 H의 경우에서 공통된 1개 (F∩H → 스팸 1개로결정된 데니까) / 6개中 의 스팸을 받을확률은 "0.1"

즉, P(H|F) 경우 확률이 더 높다.

explain

Q5. P(I|E)

$$\frac{\left[{}_3C_2 \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^4\right] + \left[{}_3C_3 \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^3\right]}{{}_{10}C_3 \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^3} = 0.058333$$

Q6. $P(1|F)$

$$= \left(1 - \frac{{}_6C_0\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^6}{{}_6C_6\left(\frac{1}{3}\right)^6\left(\frac{2}{3}\right)^0}\right) + \left(1 - \frac{{}_6C_1\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^5}{{}_6C_6\left(\frac{1}{3}\right)^6\left(\frac{2}{3}\right)^0}\right)$$

$$= (1 - 0.3046) + (1 - 0.914474)$$

$$= 0.6952 + 0.085521$$

$$= 0.780721$$

$\therefore P(1|F) = 0.780721 \cdots$

→ fair coin probability problems.

Q7. $P(A)$?

$$\frac{3}{4} \times \frac{1}{2} + \frac{3}{5} \times \frac{1}{2} = \frac{3}{8} + \frac{3}{10} = \frac{15}{40} + \frac{12}{40} = \frac{27}{40}$$

Q8. $P(D.S.A \mid A)$? (Data Science Application = D.S.A)

$$\frac{\frac{1}{2} \times \frac{3}{4}}{\frac{27}{40}} = \frac{\frac{3}{8}}{\frac{27}{40}} = \frac{3}{8} \times \frac{40}{27} = \frac{5}{9}$$

Q9. $P(L.A \mid A)$? (Linear Algebra = L.A)

$$\frac{\frac{1}{2} \times \frac{3}{5}}{\frac{27}{40}} = \frac{\frac{3}{10}}{\frac{27}{40}} = \frac{4}{9}$$

→ unfair coin probability problems.

Q10. $P(A)$?

$$\left(\frac{3}{4} \times \frac{3}{4}\right) + \left(\frac{1}{4} \times \frac{3}{5}\right) = \frac{9}{16} + \frac{3}{20} = \frac{45}{80} + \frac{12}{80} = \frac{57}{80},$$

Q11. $E[A]$? (expectation A, when "A = 4.0")

$$P(A) \times 4.0 = \frac{57}{\cancel{80}\ 20} \times \cancel{4.0} = 2.85,$$

$\therefore E[A] = 2.85,$