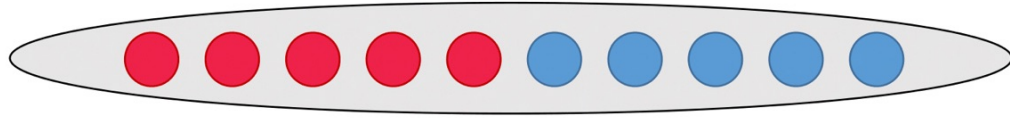


## Question. 8-01

안이 보이지 않는 주머니 속에 빨간 구슬 5개 파란 구슬 5개가 들어있다. 이때, 다음 물음에 답하시오.



- 1) 주머니에서 꺼낸 구슬이 빨간 구슬일 확률과 Odds를 구하시오.
- 2) 빨간 구슬 6개, 파란 구슬 4개가 들어있을 때, 빨간 구슬을 뽑을 확률과 Odds를 구하시오.
- 3) 빨간 구슬 4개, 파란 구슬 6개가 들어있을 때, 빨간 구슬을 뽑을 확률과 Odds를 구하시오.
- 4) 위 과정에서 빨간 구슬을 뽑을 확률이 달라졌을 때 Odds의 값이 어떻게 변화하였는지 설명하시오.
- 5) Odds를 기반으로 Logistic Regression 학습을 진행하는 것이 적절한지 설명하시오.

1) 전체 구슬의 개수가 10개, 빨간 구슬의 개수가 5개이므로,

$$P(\text{빨간구슬}) = \frac{n(\text{빨간구슬})}{n(\text{전체구슬})} = \frac{5}{10} = 0.5$$

$$O(\text{빨간구슬}) = \frac{n(\text{빨간구슬})}{n(\text{전체구슬}) - n(\text{빨간구슬})} = \frac{5}{10-5} = 1$$

2) 전체 구슬의 개수가 10개, 빨간 구슬의 개수가 6개이므로,

$$P(\text{빨간구슬}) = \frac{n(\text{빨간구슬})}{n(\text{전체구슬})} = \frac{6}{10} = 0.6$$

$$O(\text{빨간구슬}) = \frac{n(\text{빨간구슬})}{n(\text{전체구슬}) - n(\text{빨간구슬})} = \frac{6}{10-6} = 1.5$$

3) 전체 구슬의 개수가 10개, 빨간 구슬의 개수가 4개이므로,

$$P(\text{빨간구슬}) = \frac{n(\text{빨간구슬})}{n(\text{전체구슬})} = \frac{4}{10} = 0.4$$

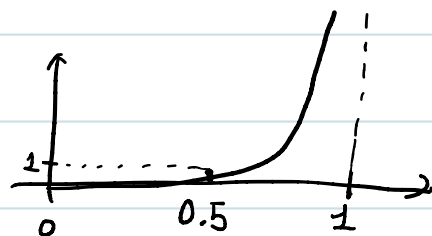
$$O(\text{빨간구슬}) = \frac{n(\text{빨간구슬})}{n(\text{전체구슬}) - n(\text{빨간구슬})} = \frac{4}{10-4} = 0.6$$

4)  $n(\text{빨간구슬})$  4      5      6  
 $P(\text{빨간구슬})$  0.4      0.5      0.6  
 $O(\text{빨간구슬})$  0.6      1      1.5

$\underbrace{\quad}_{-\frac{1}{3}} \quad \underbrace{\quad}_{+\frac{1}{2}}$

확률의 변화가 균등하게 나타났음에도 Odds는 불균등하게 값이 변화하였음을 알 수 있다.

5)  $\frac{P(A)}{1-P(A)}$ 의 그래프



$$P(A) > P(A^c) : O(A) > 1$$

$$P(A) = P(A^c) : O(A) = 1$$

$$P(A) < P(A^c) : 0 < O(A) < 1$$

$P(A) > P(A^c)$ 의 범위와 (1이상의 실수)  $P(A) < P(A^c)$ 의 범위가 (0~1) 달라서 학습의 불균형이 생기므로 Odds를 기반으로 학습을 진행하는 것은 적절하지 않다.