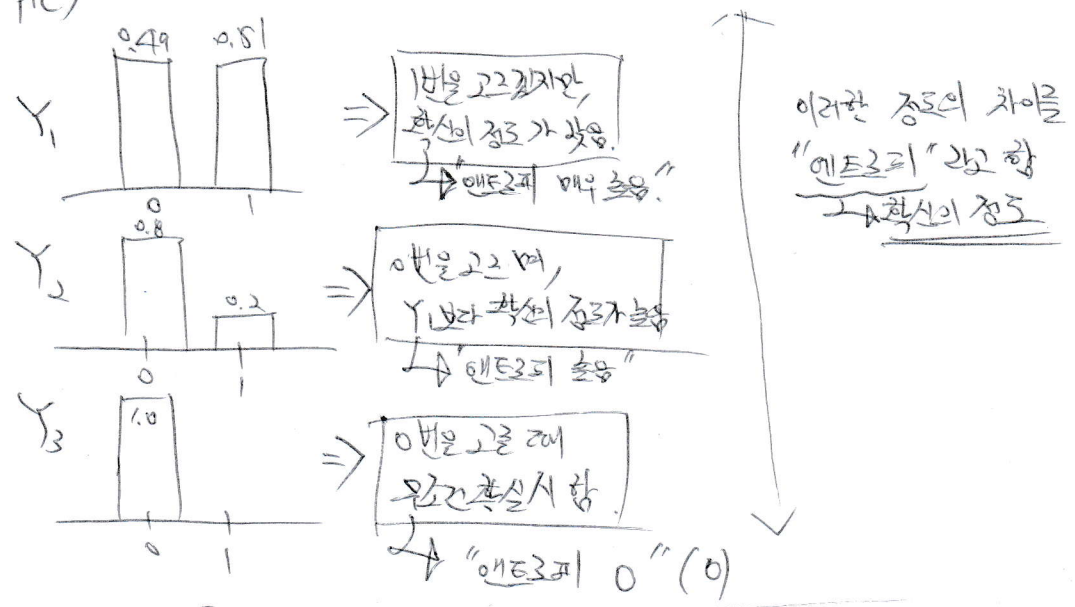


엔트로피 (Entropy)

엔트로피 정의 : 확률 분포들이 가지는 가치는 "확신의 정도"를 수치로 표현한 것.

- 확률 분포의 여러 값이 사들 확률이 비슷할 경우 \Rightarrow 엔트로피 \uparrow (높음)
- 어떤 확률 분포의 값이 사들 확률이 높고, 나머지 값이 확률이 낮다면 \Rightarrow 엔트로피 \downarrow (낮음)

For example)



- \hookrightarrow 이렇게, 확률 분포에서 숫자를 쓰는 것은 '가치'를 쓰는 것과 같은 개념.
- \hookrightarrow 엔트로피도 마찬가지로, 숫자(스칼라)로 표현됨.
- \hookrightarrow 가지는 H 를 사용하여 수학적 정의는 다음과 같다.

1) 이산확률 변수 (Discrete)

$$H[Y] = - \sum_{k=1}^K P(Y_k) \log_2 P(Y_k)$$

\Rightarrow "이산, 이산확률 변수를 가짐으로 진행"

2) 연속확률 변수 (Continuous)

$$H[Y] = - \int p(y) \log_2 p(y) dy$$

- example의 엔트로피를 계산해보면,

- ① $H[Y_1] = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} = 1$
- ② $H[Y_2] = -\frac{8}{10} \log_2 \frac{8}{10} - \frac{2}{10} \log_2 \frac{2}{10} = 0.72$
- ③ $H[Y_3] = -1 \log_2 1 - 0 \log_2 0 = 0$

$\hookrightarrow P(Y_k) = 0$ 인 경우, '극한 값'을 사용한다.
 $\Rightarrow \lim_{p \rightarrow 0} p \log_2 p = 0$
 "0에 가까워지는 것"