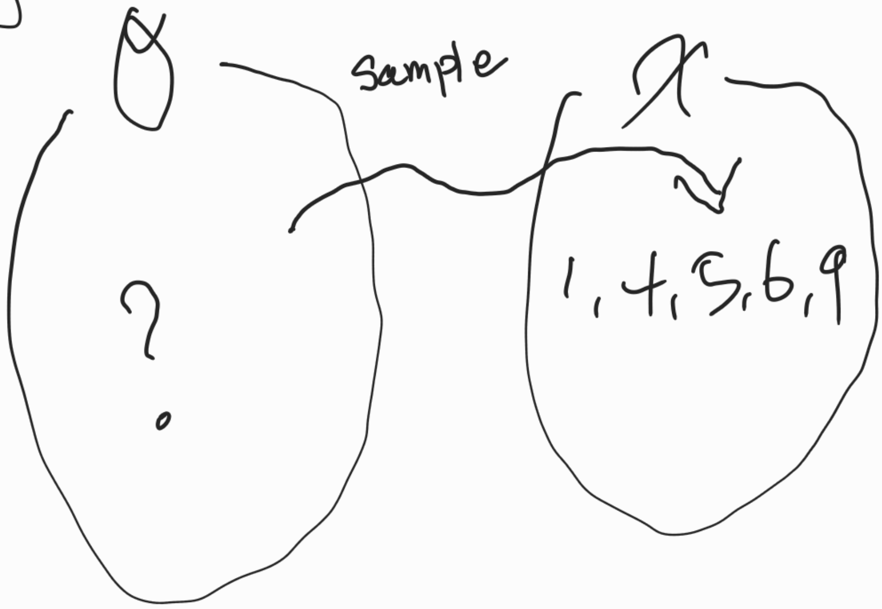
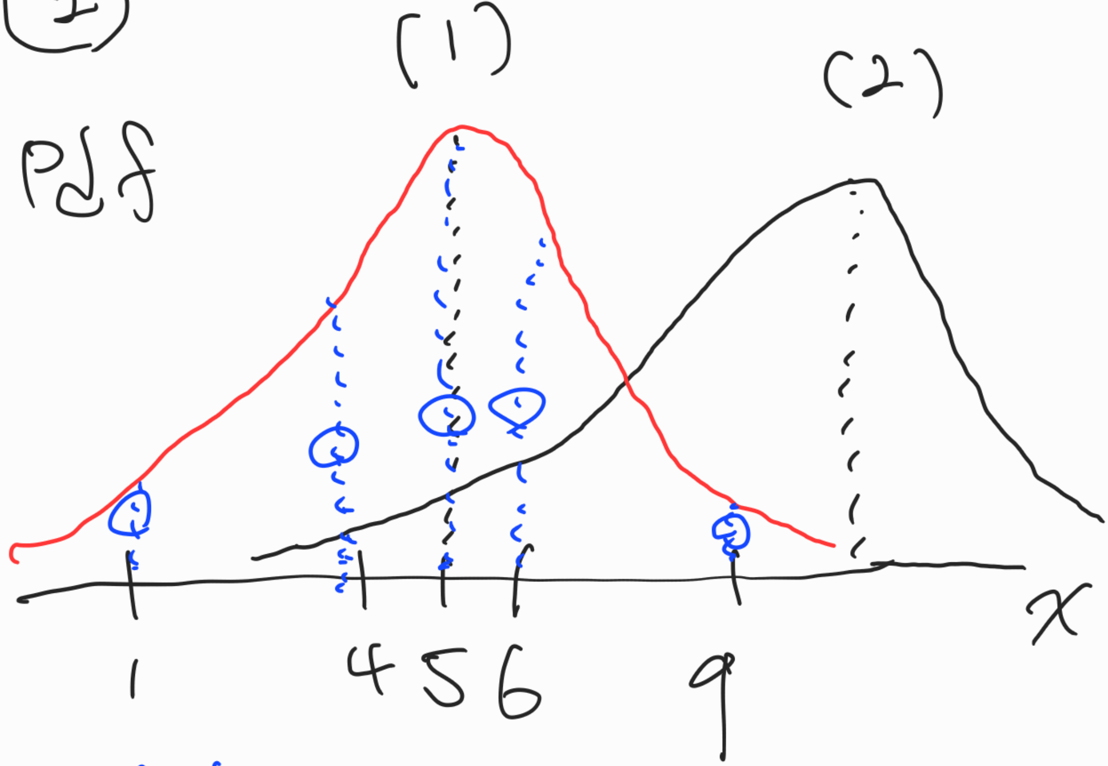
①



이 θ의 μ, σ 무엇?

②



pdf의 높이 = likelihood에 '기여도'

$\boxed{P(x_k|\theta)}$

가능도 : likelihood function $= \prod_{k=1}^{n} P(x_k|\theta)$

MLE

로그-가능도 $L(\theta|x)$

$$\log \prod_{k=1}^{n} P(x_k|\theta) = \sum_{k=1}^{n} \log P(x_k|\theta) \cdots L(\theta|x)\text{의}$$

최댓값 찾기 = MLE

$P(x_i|\theta)$

ex) 정규분포의 밀도함수 $\boxed{\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)}$

$$\frac{\partial}{\partial\theta} L(\theta|x) = \sum_{i=1}^{n} \frac{\partial}{\partial\theta} \log P(x_i|\theta) = 0$$

<로그-유도>

$$L(\theta|x) = \sum_{i=1}^{n} \log \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)$$

$$= \sum \left\{ \cancel{\log}\left(\cancel{\exp}\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)\right) - \log\left(\sigma\sqrt{2\pi}\right) \right\}$$

$$= \sum \left\{ -\frac{(x_i-\mu)^2}{2\sigma^2} - \log(\sigma) - \log(\sqrt{2\pi}) \right\}$$

μ에 대해 편미분.

$$\frac{\partial L(\theta|x)}{\partial\mu} = -\frac{1}{2\sigma^2} \sum \frac{\partial}{\partial\mu}\left(x_i^2 - 2x_i\mu + \mu^2\right)$$

$$= -\frac{1}{2\sigma^2}\sum(-2x_i + 2\mu) = \frac{1}{\sigma^2}\sum(x_i - \mu) = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n} x_i - n\mu\right) = 0$$

베이즈 정리

$$P(H|E) = \frac{\overbrace{P(E|H)}^{\text{가능도}}\overbrace{P(H)}^{\text{사전확률}}}{\underbrace{P(E)}_{\text{증거}}}$$

사후확률

H : Hyphosis, 가설

E : Evidence, 새로운 정보

· P(H) : 어떤 사건이 발생했다는 주장에 대한 '신뢰도'

P(H|E) : 새로운 정보(E)를 받은 후 갱신된 '신뢰도'

☆

* 기존 확률(사전확률) P(H)에 대해 새로운 정보 P(E)가 추가되었을 때 P(H|E)의 '신뢰도'가 어떤식으로 바뀔까.

관심있는 것.

$P(L|C)$: Candy를 받았을 때, Love일 확률.

① 사전확률.

Laplace의 '불충분의 원리'에 따라

· 호감이 있을 확률 : 50% $\longrightarrow P(L) = 0.5$

② Evidence 1. : 호감이 있을 때, Candy를 줄 확률 = 40%

$P(C|L) = 0.4$

$\hookrightarrow P(C^c|L) = 0.6$

③ Evidence 2. : 호감이 없는데, Candy를 줄 확률 = 30%

$P(C|L^c) = 0.3$

$P(C^c|L^c) = 0.7$

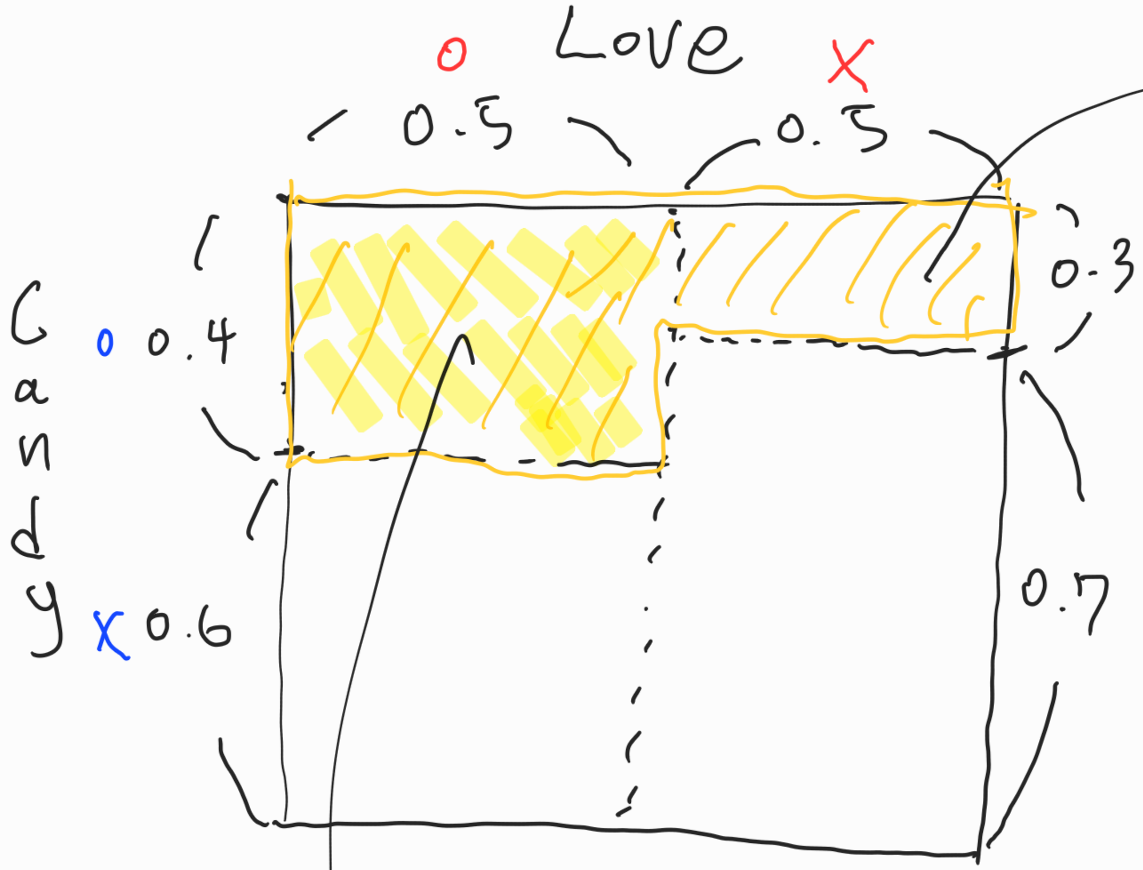
· 우리가 알고 싶은 것 : $P(L|c)$

· 우리가 알고 있는 것 : $P(C|L)$

$\longrightarrow$ 우리가 관심이 있는 상황 "Candy를 받음"

$\rightarrow \dfrac{\overbrace{0.4 \times 0.5}^{P(C|L)\times P(L)}}{0.4\times0.5+\underbrace{0.3\times0.5}_{P(C|L^c)\times P(L)}} = P(L|c)$

$\fallingdotseq 57\%$



$P(C|L^c)$

$P(C|L)$

우리가 아는것

x_i

정규분포 ...

'어떤분포'로부터 추출된 '표본집합' $P(x_i|\theta)$

이 '표본집합'을 통해 '어떤분포'의 μ 또는 σ^2을 찾기.

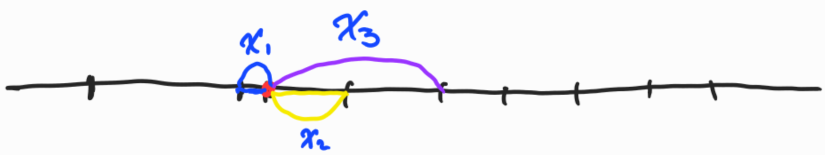
$P(\theta|x_i)$ → 우리가 알고싶은것.

KNN – 추가

차원의 저주

: 차원이 증가할수록 계산값 증가.

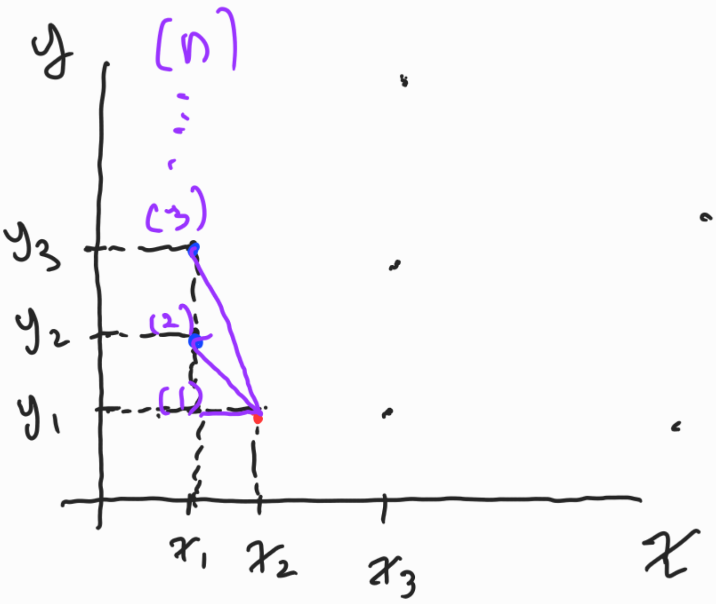
1 차원.



$\min(x_1, x_2, x_3, \dots)$

$= x_1$

2 차원 ..... n 차원



. 1 차원거리 = x_1

. 2 차원거리 : $(1) < 거리 < (n)$