<MLE>

Binomial Distribution : discrete (continuous x)

ex) Head or Tail 2가지만 존재

iid ⇒ independent identically distributed

ex) 동전 5회 → 각각은 독립, 각각은 동일한 확률분포

$P(H) = \theta, \quad P(T) = 1-\theta$

$P(HHTHT) = \theta \times \theta \times (1-\theta) \times \theta \times (1-\theta) = \theta^3 (1-\theta)^2$

D: Data. n: 횟수 = 5 $\begin{cases} a_H = 3 \\ \text{(Head)} \\ a_T = 2 \end{cases}$ $P = \theta$ (Head 나올 확률)

$P(D|\theta) = \theta^{a_H} (1-\theta)^{a_T}$

가정: 결과는 θ를 따른다. → How to make strong?

MLE of θ

↳ $\hat{\theta} = \text{argmax}\, P(D|\theta) = \text{argmax}\, \theta^{a_H} (1-\theta)^{a_T}$

계산 용이하게 하기 위해 log

$\text{argmax} \{ a_H \ln\theta + a_T (1-\theta) \}$

$\frac{d}{d\theta} (a_H \ln\theta + a_T \ln(1-\theta)) = 0$

미분 → 극점을 통한 최적화

$$\frac{a_H}{\theta} - \frac{a_T}{1-\theta} = 0$$

$$(1-\theta)a_H - \theta a_T = 0 \Rightarrow \therefore \theta = \frac{a_H}{a_T + a_H} = \frac{3}{5}$$

$\hat{\theta}$: 추정 $= \dfrac{a_H}{a_H + a_T}$ θ^* : true parameter $N = a_H + a_T$

$$P(|\theta^* - \hat{\theta}| \geq \epsilon) \leq 2e^{-2N\epsilon^2}$$

N: trial 횟수

$\therefore$ 시도 ↑ $\Rightarrow 2e^{-2N\epsilon^2}$ ↓

시도↑ ⇒ error 일어날 확률↓

⇒ PAC Learning $\epsilon = 0.1$ 0.01% case

< MAP >

사전확률 넣기 가능) 대충 $\frac{1}{2}$, $\frac{1}{2}$ 일거 같음!

$$P(\theta | D) = \frac{P(D|\theta) \cdot P(\theta)}{P(D)}$$

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior Knowledge}}{\text{Normalizing Constant}}$$

$$P(D|\theta) = \theta^{a_H} \cdot (1-\theta)^{a_T}$$

$P(D)$: 이미 발생 → θ의 영향 X

$$P(\theta|D) \propto P(D|\theta) \cdot P(\theta)$$

$$P(D|\theta) = \theta^{a_H}(1-\theta)^{a_T}$$

$P(\theta) =$???

$\Downarrow$

Use Beta Distribution (0 ~ 1 Cumulative)

$$P(\theta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha, \beta)} \qquad B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \qquad \Gamma(\alpha) = (\alpha-1)!$$

$$P(\theta|D) \propto P(D|\theta)P(\theta) \propto \theta^{a_H}(1-\theta)^{a_T} \cdot \theta^{\alpha-1}(1-\theta)^{\beta-1}$$

$$\propto \theta^{a_H+\alpha-1}(1-\theta)^{a_T+\beta-1}$$

MLE 처럼 loglikelihood $\qquad \hat{\theta} = \dfrac{a_H+\alpha-1}{a_H+a_T+\alpha+\beta-2}$

횟수↑ ⇒ $\dfrac{a_H}{a_H+a_T}$ 로 수렴할 것

$$\text{Beta}(\alpha,\beta) \Rightarrow E = \frac{\alpha}{\alpha+\beta}$$

$$V = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$