

## Ch 04. 가우스 혼합 모델

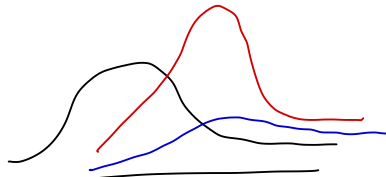
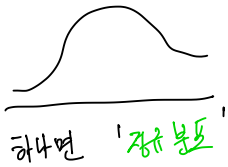
---

1. 우리 주변의 다봉 분포
2. 가우스 혼합 모델 데이터 생성
3. 가우스 혼합 모델의 수식
4. 매개변수 추정: 어려움



# 4.1 우리 주변의 다봉 분포

가우스 분포



여러 개면? → 가우스 혼합 모델  
 (Gaussian Mixture Model, GMM)

## 4.2 가우스 혼합 모델 데이터 생성

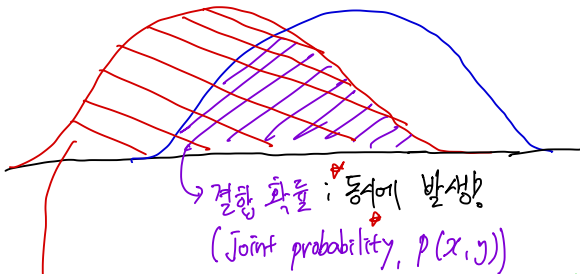
모든 다봉 분포 **Wow!** → GMM으로 표현 가능? **예!** 2 개리 작업이 필요!

1! 모델링 : 관측 데이터의 분포를 GMM으로 표현할 수 있다고 가정

2! 매개변수 추량 : GMM의 매개변수 추량

→ **예!** 이걸 최대 가능도로 안돼  $\pi$  → 2대! EM! 을 사용

## 4.3 가우스 혼합 모델의 수식



→ 주변 확률 : 개별 사건 발생  
 (marginal probability,  $p(x)$  or  $p(y)$ )

예!  $y$ 가 이미 일어난 상황에서  $x$ 의 발생 확률은? **예!**  
 ⇒ **조건 확률** → 독립 관련 필요?  
 (Conditional probability)

공식 정리!

$$p(x, y) = p(x|y) p(y)$$

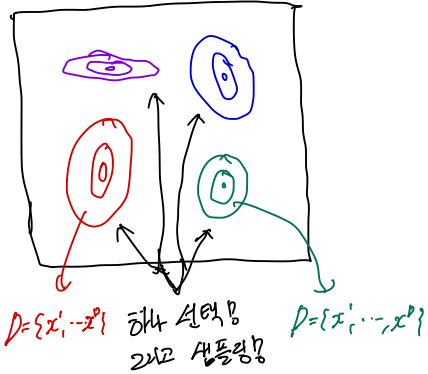
$$p(x) = \sum_y p(x, y) \text{ or } \int p(x, y) dy$$

# GMM의 데이터 생성 절차!

→ 우선 K개의 정규 분포를 준비!

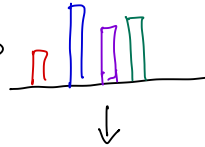
그리고 반복!

1. 특징 확률 분포에 따라 K개의 정규분포 중 하나 선택!
2. 선택한 정규 분포에서 데이터 생성!



여기서!

이걸 범주형 분포 표현!  
(categorical distribution)



$$p(z=k|\phi) = \phi_k$$

따라서 최종! GMM 수식은!!

1. K개 정규 분포 중 K개에 따라 정규분포 선택! 즉!  $z=k$  라면,  $x$ 가 따르는 확률 분포는  $k$ 번째!

따라서!  $\rightarrow p(x|z=k; \mu, \Sigma) = N(x; \mu_k, \Sigma_k)$

2. 생성 모델의 목표!  $\rightarrow$  관측 데이터  $x$ 의 확률 분포  $p(x)$ 를 표현

따라서 주변과 A형!  $\rightarrow p(x) = \sum_{k=1}^K p(x, z=k)$

3. 이때, 결합 확률은? 중·생·정·리

$$\rightarrow p(x, z=k) = p(z=k) p(x|z=k) = \phi_k N(x; \mu_k, \Sigma_k)$$

4. 따라서  $\pi\pi$  결국  $\pi\pi$

$$\rightarrow p(x) = \sum_{k=1}^K \phi_k N(x; \mu_k, \Sigma_k)$$

## 4.4 다변량 정규 분포의 어려움

GMM은 다변량 정규 분포의 어려움.

$$p(x; \phi, \mu, \Sigma) = \sum_{k=1}^K \phi_k N(x; \mu_k, \Sigma_k) \text{ 인데 } \sim D = \{x^{(1)}, x^{(2)}, \dots, x^{(N)}\} \text{ 을 받았다면 } \sim$$

$$p(D; \theta) = p(x^{(1)}; \theta) p(x^{(2)}; \theta) \dots p(x^{(N)}; \theta) = \prod_{n=1}^N p(x^{(n)}; \theta) \Rightarrow \text{가능도}$$

$$\log p(D; \theta) = \log \prod_{n=1}^N p(x^{(n)}; \theta)$$

$$= \sum_{n=1}^N \log p(x^{(n)}; \theta)$$

$$= \sum_{n=1}^N \log \left( \sum_{k=1}^K \phi_k N(x^{(n)}; \mu_k, \Sigma_k) \right) \Rightarrow \text{로그 가능도}$$

→ 로그-가능도 → 분해해서 어려워  $\pi\pi$  → 합-로그 → 해석으로 쉬워 등등  
(log-sum) 평미분할 어려워  $\pi\pi$

즉!  $\log p(D; \theta)$  으로는 최대 가능도 어려워  $\pi\pi$