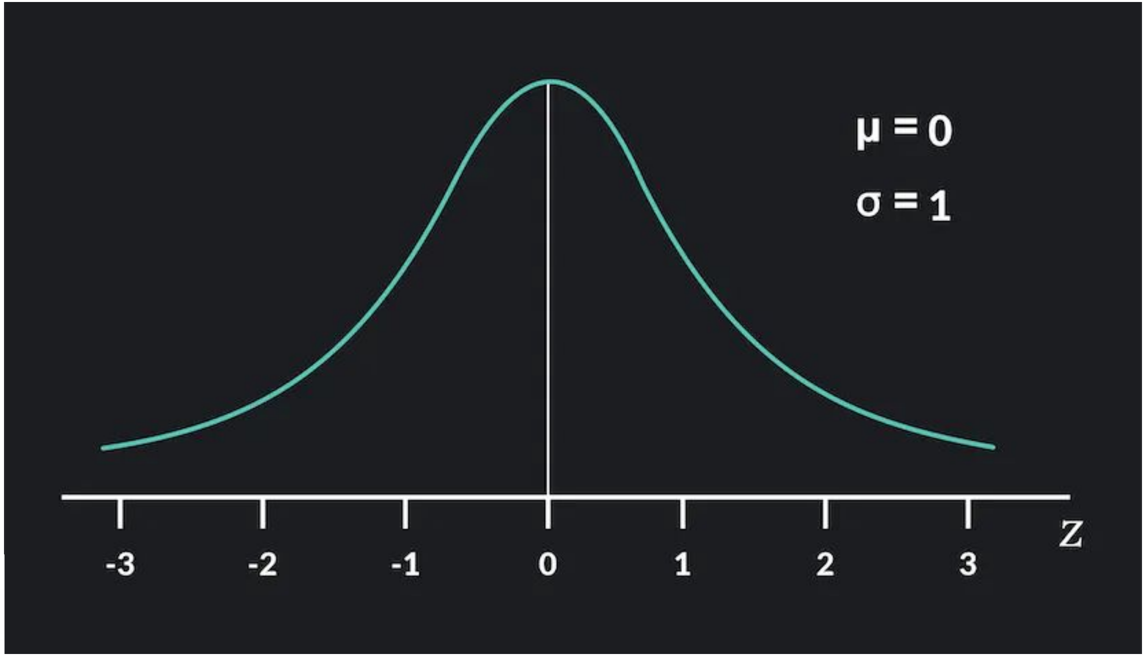
GMM

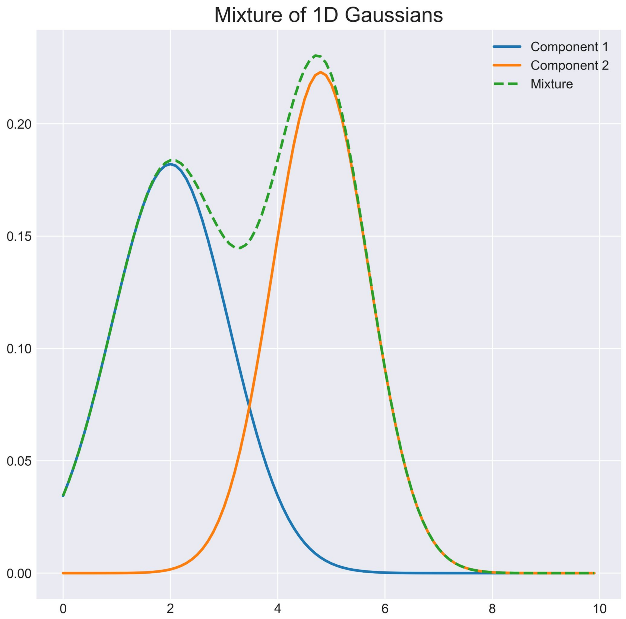
Gaussian Mixture Model

Gaussian



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Gaussian Mixture

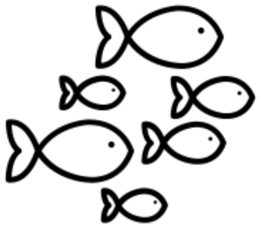


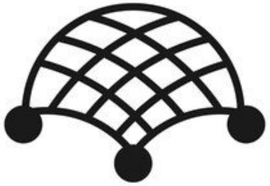
Gaussian Mixture in the wild

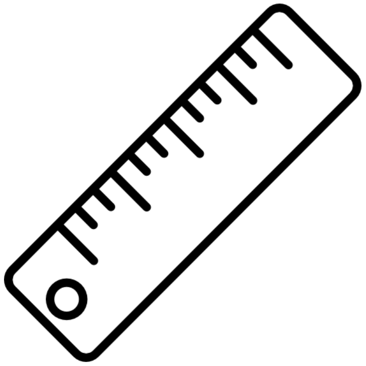
- Let’s assume a fish specie that has the following properties:
  - It is born at the same time of the year (december)
  - It grows for the first 2 years then stops as an adult
  - Only age gives the size
- We sample the fish population in June, so we expect to have:
  - 6 Month old fishes
  - 1.5 year old fishes
  - 2.5 and more year old fishes
- The data is noisy

- How can we count our fish population to know its age distribution?

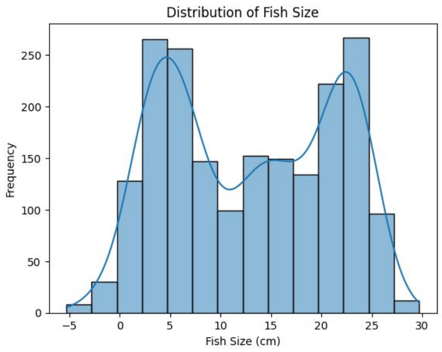
Gaussian Mixture in the wild



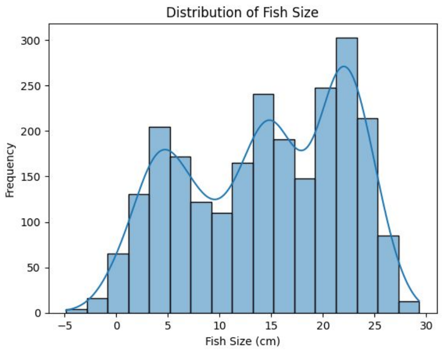




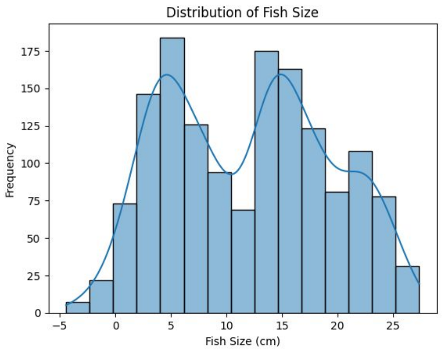
Gaussian Mixture in the wild



Site 1



Site 2



Site 3

Gaussian Mixture in the wild

- We know that this distribution comes from three random distributions
- We want to assess for each site the underlying distributions of fishes

Gaussian Mixture Model

- This is a form of Classification that will try to fit for a sum of gaussians:
    - The mean: mu
    - The variance: sigma
    - The scaling factor (proportion): thetha

$$p(x) = \sum_{i=1}^{K} \phi_i \mathcal{N}(x \mid \mu_i, \sigma_i)$$

$$\mathcal{N}(x \mid \mu_i, \sigma_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right)$$

$$\sum_{i=1}^{K} \phi_i = 1$$