

## TP 5

### Exercice 1 : Cas simple

Soit  $\{x_1, x_2, \dots, x_n\}$  un échantillon provenant d'une loi de densité normale  $N$  de paramètres  $\mu=10$ , et  $\sigma^2=2$  :

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp -\frac{(x-\mu)^2}{2\sigma^2} \quad \text{Pour } x \in \mathbb{R}$$

1. Générez 1000 échantillons de la densité  $N(10, 2)$ .
2. Estimez les paramètres  $\mu$  et  $\sigma^2$  en utilisant l'estimateur de maximum de vraisemblance dans Python.
3. Interpréter le résultat.

### Exercice 2 : Modèles de mélange

L'objectif est d'identifier les paramètres de deux lois normales depuis une seule distribution mixte (appelé aussi mélange gaussien ou GMM - Gaussian Mixture model) à l'aide de l'algorithme EM (Expectation-Maximisation).

1. Générez 1000 échantillons qui suivent une loi normale 1 de moyenne  $\mu=190$  et d'écart-type  $\sigma=10$ .
2. Générez 1000 échantillons qui suivent une loi normale 2 de moyenne  $\mu=160$  et d'écart-type  $\sigma=5$ .
3. Retrouver les paramètres des deux gaussiennes paramétrées par  $(\mu_A, \sigma_A)$  et  $(\mu_B, \sigma_B)$  uniquement à partir de cette distribution que nous appellerons  $\mathbf{X}$ .

#### Algorithme EM

L'Algorithme EM ( Expectation-Maximisation ) permet de trouver les paramètres de ces deux gaussiennes en partant de valeurs aléatoires et en les ajustant au fur et mesure jusqu'à ce que la vraisemblance de ces modèles soient maximales. Les étapes sont les suivantes :

1. Initialiser deux lois normales A et B en choisissant des valeurs aléatoires
2. Pour chaque valeur de  $\mathbf{X}$ , calculer sa probabilité sous l'hypothèse A ( $p_A$ ) puis B ( $p_B$ )
3. Pour chaque valeur de  $\mathbf{X}$ , calculer le poids  $w_A = \frac{p_A}{p_A + p_B}$  et  $w_B = \frac{p_B}{p_A + p_B}$
4. Calculer des nouveaux paramètres  $(\mu_A, \sigma_A)$  et  $(\mu_B, \sigma_B)$  en ajustant  $\mathbf{X}$  à partir des poids  $w_A$  et  $w_B$ .
5. Recommencer