

# TD 1 - Tests de gaussianité

**But du TD :** simuler des données suivant diverses lois, tracer des histogrammes, tester la normalité.

## 1 Simulation de données, histogrammes

### 1.1 Simulation de données

- Simuler 1000 données suivant chacune des lois suivantes : loi normale de moyenne 2 et variance 4, uniforme sur  $[-1, 2]$ , Poisson de paramètre 20, exponentielle de paramètre 10 (utiliser les commandes `rnorm`, `runif`, `rpois`, `rexp`). Stocker ces données dans des vecteurs appelés `N`, `U`, `P`, `E` respectivement.
- En utilisant la fonction `data.frame`, créer un dataframe `NUPE` qui contient les vecteurs `N`, `U`, `P` et `E`. Nommer les colonnes du dataframe `Gauss`, `Unif`, `Pois` et `Exp` respectivement.

### 1.2 Histogramme

- Ecrire une fonction `hist4` qui dessine sur une même fenêtre les histogrammes de 4 colonnes au choix d'un dataframe, en affichant sur chaque graphique le nom de chaque colonne. Utiliser les commandes : `layout(matrix(1 :4,2,2))` qui divise la fenêtre graphique en 4 parties identiques ; une boucle `for` pour ne pas écrire 4 fois la même chose ; la fonction `names` qui renvoie le nom des colonnes d'un tableau ; l'option `main` de la fonction `hist`, qui donne un nom à l'histogramme.
- Appliquer la fonction `hist4` au dataframe `NUPE`. Commenter les histogrammes.

## 2 Tests de normalité

### 2.1 Test de Shapiro-Wilk

- A l'aide de la fonction `shapiro.test`, effectuer un test de normalité de Shapiro-Wilk sur chacun des vecteurs de données `N`, `U`, `P`, `E`.
- Commenter.

### 2.2 Méthode des graphiques quantile-quantile (q-q plot) et de la droite de Henry

- En s'inspirant de la question 1.2, créer une fonction `qqhenry` qui trace sur une même fenêtre les graphiques quantile-quantile et la droite de Henry de 4 colonnes au choix d'un dataframe. Utiliser les fonctions `qqnorm` et `qqline`.
- Appliquer `qqhenry` au dataframe `NUPE`. Commenter les résultats.

## 3 Application à des données réelles

### 3.1 Stockage de données

- Stocker dans l'objet `voitures` les données `mtcars`.
- Extraire les variables quantitatives des données `voitures`. Appeler le nouveau dataframe `v`.

### 3.2 Tests de gaussianité

- Etudier la normalité à l'aide du test de Shapiro-Wilk et des fonctions `hist4` et `qqhenry` sur 4 colonnes de votre choix du dataframe `v`.
- Commenter les résultats.

**Penser à utiliser l'aide :** pour obtenir l'aide en ligne sur une fonction `nomdelafonction`, taper `?nomdelafonction` ou `help(nomdelafonction)`