

## Fiche de TD n° 3

### Introduction à R

Le but de ce TD en salle informatique est de s'initier au fonctionnement de R. On conseille de garder près de vous le document de référence intitulé Introduction à R, également disponible sur le portail (la forme numérique facilite les copier-coller).

#### Organisation du travail

Après avoir lancé R, vous ouvrirez un fichier dans l'éditeur fourni (ou un autre à votre choix) qui s'intitulera TD3\_Nom1\_Nom2.R. Ce fichier sera déposé sur le portail à la fin du TD sur le portail dans le dossier intitulé `copies TD n°3`.

#### 1° - Premiers pas

Taper la commande :

```
demo()
```

qui fait apparaître les démos disponibles. Examiner alors la démo sur les graphiques en saisissant la commande :

```
demo(graphics)
```

En vous aidant des documents disponibles, construisez une séquence régulière de nombres compris entre 0 et 5 avec un pas de 0.2. Affectez le résultat dans une variable intitulée *abscisse* qui contient des réels notés  $x$ . Construisez alors une variable *simul* (contenant des réels  $y$ ) selon la formule :

$$y = 2 * x + 1 + \varepsilon,$$

où  $\varepsilon$  est simulée selon une loi uniforme sur l'intervalle  $[-2,2]$ .

On utilisera pour cela la fonction `runif`.

Représenter alors graphiquement la dépendance de *simul* en fonction de *abscisse*, puis donner un titre à votre graphique et des légendes aux axes (voir commandes `plot` et `title`).

La commande `lm`, que nous reverrons beaucoup dans la suite du cours, ajuste une droite au sens des moindres carrés sur la base des données présentes. On obtient la valeur des ordonnées des points de la droite correspondant aux abscisses des données par la commande :

```
fitted(lm(simul ~ abscisse))
```

Ajouter alors à votre graphique la droite correspondante, dessinée en bleu.

Refaire de même en remplaçant les ordonnées précédentes par :

```
fitted(lm(simul ~ poly(abscisse,2)))
```

Qu'observe-t-on ? Et en remplaçant 2 par 10 ?

## 2° - Premières données

Récupérer sur le portail (répertoire data) le fichier de données d'épaisseur d'oxyde présentées au cours n°3. Le charger avec la commande `read.table` et affectez le résultat à la variable `oxyde`. Pour ce faire, bien lire l'aide en examinant les arguments `header`, `skip`, `dec`, `sep`.

Vérifier que tout est correct en utilisant la commande `summary` appliquée à l'objet que vous venez de définir. Vous pouvez également examiner les premières lignes de votre variable en utilisant la fonction `head`.

Représenter graphiquement le nuage de points donnant pour toutes les observations effectuées :

- l'épaisseur observée en abscisse
- le numéro du site correspondant.

Afin de mieux comparer les épaisseurs observées sur les 3 sites, réaliser :

- des `summary` pour les épaisseurs du site 1, du site 2, du site 3 (pour sélectionner les données par site, utiliser des commandes du genre `oxyde$thickness[oxyde$site==1]` ou `oxyde[oxyde$site==1,]`)
- un boxplot des épaisseurs en fonction des numéros des sites.

Que conclure pour l'instant au vu de ces éléments ?

Pour terminer, observer les variables présentes dans votre environnement (commande `ls()`) et supprimer l'une d'entre elles, en vérifiant le résultat.