

# Sujet B – Evaluation 1 – TPSTA401 — 2024

Frédérique Leblanc

## Exercice 1 (8pts)

Choisir comme répertoire de travail avec Rstudio (menu Session/Choose Working Directory/ ...) un répertoire dans son “home” contenant le fichier `diamantsPurs.csv`. Charger les données et les affecter à un objet nommé `diam` avec l’instruction suivante :

```
diam<-read.table("diamantsPurs.csv",header=T,sep=" ")
```

### question 1 (2pts)

Extraire de ce data frame l’échantillon des prix pour les diamants pour lesquels la variable `color` prend la modalité D et l’affecter à `x` puis affecter à `y` l’échantillon des prix des diamants pour lesquels la variable `color` prend la modalité I dans le tronçon d’instructions suivant à compléter :

```
#extraction de l'échantillon x  
#extraction de l'échantillon y
```

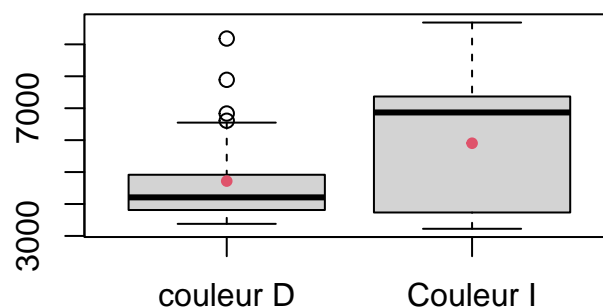
### question 2 (3pts)

Donner les trois quartiles empiriques de chaque échantillon (tronçon des instructions de calculs **et** texte de réponse à la question).

### question 3 (3pts)

Pour comparer ces deux échantillons on peut représenter leurs boxplots sur un même graphique auquel sont ajoutées ici les positions des moyennes (points rouges) :

**prix d'un diamant selon sa couleur**

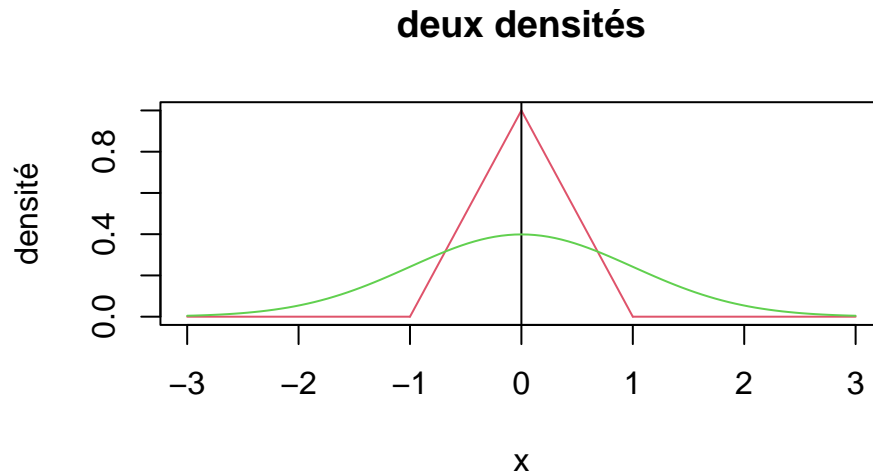


Interpréter ce graphique (3 commentaires attendus).

## Exercice 2 (12 pts)

### question 1 (10 pts)

Représenter sur l'intervalle  $[-3, 3]$  la fonction  $f$  valant  $x + 1$  sur l'intervalle  $[-1, 0]$ ,  $-x + 1$  sur l'intervalle  $[0, 1]$  et 0 partout ailleurs (on utilisera la fonction `segments` pour tracer la fonction sur  $[-3, -1]$  et sur  $[1, 3]$  et on tracera son graphe en rouge). Y ajouter en vert la densité d'une variable de loi  $\mathcal{N}(0, 1)$ . Indiquer le titre "deux densités", la légende "densité" en ordonnée et ajouter l'axe des ordonnées en noir (pas celui des abscisses). En d'autres termes, reproduire le graphique suivant :



### question 2 (2 pts)

La fonction tracée en rouge est-elle une densité de probabilité (justifier) ?