



**MMETH1122 : Statistiques en
sciences politiques et sociales**

**Devoir 4.3 : Distribution des salaires et
probabilité**

Année académique 2021/2021

Enseignants : Philippe Bocquier & Lev Lhommeau

**Travail à rendre le :
2 mai 2022**

Université catholique de Louvain
FUCaM Mons
27 avril 2022

Astuce 1 : Charger votre fichier dans R

Vous pouvez télécharger votre fichier sur la page lev.lhommeau.de/article18.html et l'ouvrir ensuite avec RStudio. Pour faire cela, placez le fichier dans votre espace de travail et saisissez la commande suivante dans RStudio :

```
VOTRENOM <- readRDS("VOTRENOMPrenom.rds")
```

Les utilisateurs de Windows peuvent aussi, après avoir téléchargé le fichier, entrer la commande suivante dans RStudio et choisir le fichier dans la fenêtre qui s'ouvre :

```
VOTRENOM <- readRDS(choose.files())
```

Vous pouvez aussi saisir la commande suivante dans RStudio pour charger le fichier directement d'internet :

```
VOTRENOM <- readRDS(url("https://lev.lhommeau.de/IMG/rds/
VOTRENOMPrenom.rds")) # Pour vous assurer de saisir l'adresse
correcte, vous pouvez copier l'adresse de votre fichier sur la page internet (
en fonction du navigateur : clic droit sur le lien et copier l'adresse).
```

Avec ces commandes, vous créez l'objet *VOTRENOM*. Cet objet contient votre base de données à analyser.

Mise en situation

Chacun d'entre vous a fini ses études et voudrait connaître la distribution des salaires avec un Bachelier en poche. Vous avez donc consulté les données fournies par un institut de sondage sur un échantillon représentatif des diplômés d'un Bachelier dans votre domaine en Belgique.

Boîte info 1 : Les fichiers

Chaque étudiant.e a sa propre base de données qui est nommée *NOMPrenom.RDS*. Si vous vous appelez Jeanne Dubois, le fichier se nommera *duboisjeanne.RDS*. Cependant, en raison de signes spéciaux (accents, cédilles, etc.) et le fait qu'il y ait déjà des fichiers qui portent ce nom sur le *server*, il se peut que le format soit légèrement modifié et que soient ajoutés de « -2 » (ou d'autres numéros) au nom. Chaque base de données contient une variable : « Salaire » qui reprend le salaire mensuel en euro ().

Vous trouverez tous les fichiers sur la page suivante : lev.lhommeau.de/article18.html

Analysez le résultat de cette enquête (qui est nommé approximativement *NOMPrenom.RDS*) avec R :

Questions

Attention 1 : Arrondies

Vos réponses doivent être arrondies à deux décimales après la virgule.

Question 1 (1 point). Tracez un histogramme des salaires ainsi que la courbe de distribution normale pour comparer votre histogramme avec celle-ci. (Notez que dans votre échantillon de salaires, des déviations aléatoires sont attendues.)

Question 2 (1 point). Sous l'hypothèse que les salaires suivent une distribution normale, quelle est la probabilité pour que vous obteniez un premier salaire inférieur à 1650 ?

Attention 2 : Code R

Fournissez le code R qui vous a permis d'obtenir le résultat pour toutes les questions. Si vous utilisez *RCommander* (Rcmd), copiez le code généré par *RCommander*.

Astuce 2 : Arrondies dans R

Les réponses doivent être arrondies à deux décimales après la virgule. Vous pouvez le faire « à la main », mais R vous propose de le faire à votre place grâce à la fonction « `round()` ». Si vous voulez connaître l'arrondie à deux décimales du nombre, 0,356 tapez la commande suivante dans R :

```
round(0.356, 2)
```

Le premier argument est le nombre que vous souhaitez arrondir (ici : 0,356), le second argument, après la virgule, est le nombre de décimales que vous souhaitez préserver (ici : 2). Le résultat est donc : 0,36.

Attention 3 : Fichier à remettre

Remettez votre devoir uniquement sur *Student Corner* au format PDF. Le nom du fichier doit être intitulé « Devoir 4.3 votrenom.pdf ».