Feuille de Travaux Dirigés n° 9
Mesures de liaison et variables qualitatives

Les exemples de cette feuille de travaux dirigées sont tirés du livre A Handbook of Statistical Analyses Using R, *de Brian S. Everitt and Torsten Hothorn, CRC Press 2007.*

Exercice IX.1. Groupe sanguin et Rhésus
Le tableau suivant donne la répartition de 10 000 personnes en fonction de leur groupe sanguin et de leur facteur Rhésus.

	O	A	B	AB	Total
Rh_+	3570	3825	935	170	8500
Rh_-	630	675	165	30	1500
Total	4200	4500	1100	200	10000

Les deux caractères, groupe sanguin et Rhésus, sont-ils indépendants ?

1. Récupérer les données dans R en exécutant les instructions suivantes[1].

```
> Chemin <- "C:\\..."
> Rhesus <- read.csv(paste(Chemin, "Rhesus.CSV", sep = ""))
```

Faire de *Rhesus*, le jeu de données dont nous allons nous servir.

```
> attach(Rhesus)
```

2. Vérifiez que les données sont bien celles de l'énoncé.

```
> print(Rhesus)
```

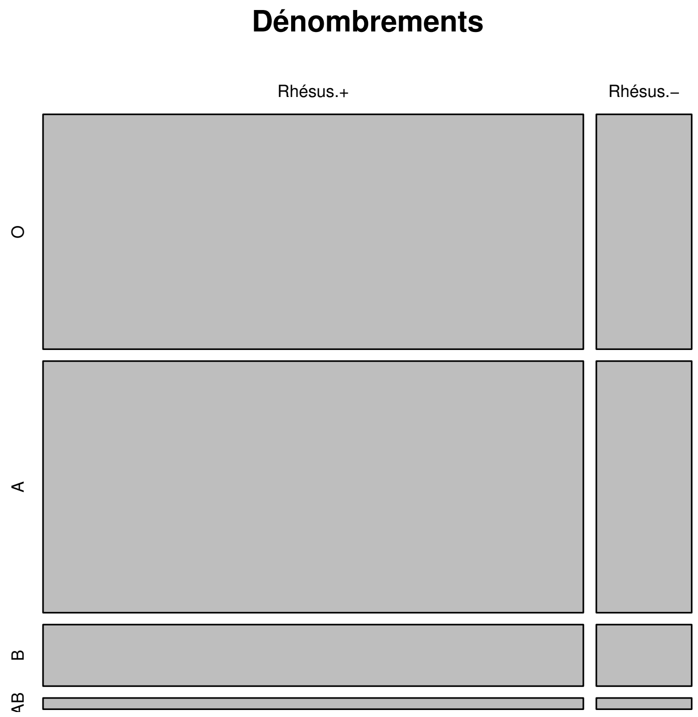
Les noms des lignes ne sont pas les bons. La commande suivante permet de corriger ce défaut.

```
> dimnames(Rhesus)[[1]] <- c("Rhésus.+", "Rhésus.-")
> print(Rhesus)
> Rhesus <- as.table(as.matrix(Rhesus))
```

Représenter graphiquement les données.

```
> plot(Rhesus, main = "Dénombrements")
```

[1]Il faut remplacer `"C :\\..."` par le répertoire dans lequel vous avez enregistré le fichier que vous souhaitez ouvrir.



3. Recalculer les marges du tableau de données.

```
> margin.table(Rhesus)
> margin.table(Rhesus, margin = 1)
> margin.table(Rhesus, margin = 2)
```

4. Peut-on réaliser un test du χ^2 d'indépendance en utlisant les approximations classiques ? Dans l'affirmative, exécuter l'instruction suivante.

```
> chisq.test(Rhesus, simulate.p.value = FALSE)
```

5. Réaliser un test du χ^2 d'indépendance version exacte basé sur la méthode de Monte-Carlo.

```
> chisq.test(Rhesus, simulate.p.value = TRUE, B = 50000)
```

5. Réaliser un test de Fisher exact classique.

```
> fisher.test(Rhesus)
```

5. Réaliser un test de Fisher exact basé sur la méthode de Monte-Carlo.

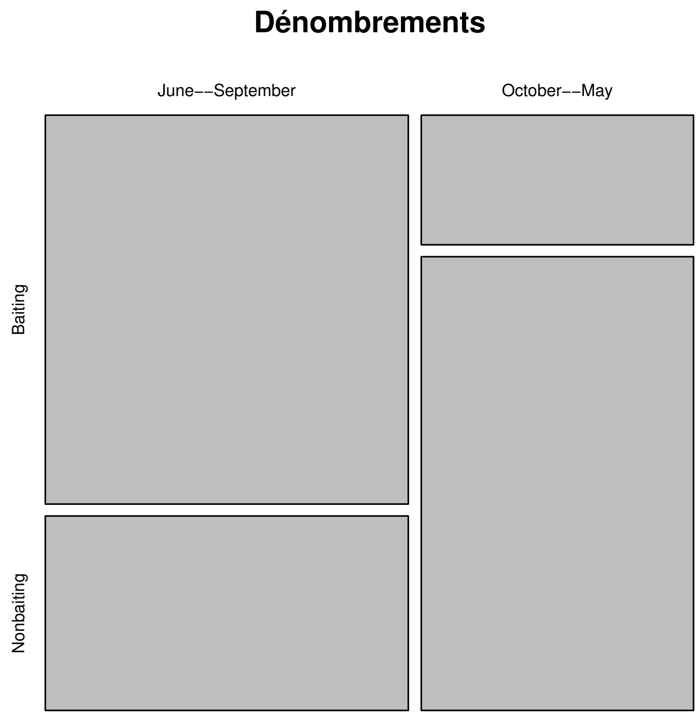
```
> fisher.test(Rhesus, simulate.p.value = TRUE, B = 50000)
```

Exercice IX.2. Suicides.
Mann (1981) établit un compte rendu d'une étude qui examine les causes d'un comportement indifférent ou concerné d'une foule lorsqu'une personne menace de se suicider en sautant du haut d'un immeuble. L'hypothèse que la foule s'intéresse à la personne est plus probable en période chaude. Mann (1981) classe 21 dénombrements de tentative de suicide par deux facteurs, la période de l'année et si la foule

est intéressée ou non. Les données sont présentées dans le tableau ci-dessous et la question est de savoir si ils ont raison d'invoquer cette hypothèse. Les données venant de l'hémisphère nord, la période Juin-Septembre représente la période chaude.

	Intéressée	Indifférente
Juin-Septembre	8	4
Octobre-Mai	2	7

1. Télécharger le fichier *suicides* du package *HSAUR* à l'aide de l'instruction `data(suicides,package="HSAUR")` ou alors rentrer les données.
2. Afficher-le à l'écran. Combien de lignes ? Combien de colonnes ?
3. Représenter-le graphiquement.



4. Peut-on envisager de faire un test du χ^2 ? Les conditions sont-elles respectées ? Pour cela calculer le tableau des effectifs théoriques.
5. À la vue des résultats obtenus ci-dessus, qu'envisagez-vous de faire pour répondre à la question posée ? En fait, quel test proposez-vous de réaliser ?

Exercice IX.3. On cherche à savoir s'il existe une association entre la couleur des yeux, clairs ou foncés, et la couleur des cheveux, blonds ou bruns, chez les êtres humains. On considère un groupe de 20 individus de sexe féminin ou masculin.

Cheveux\Yeux	Clairs	Foncés
Blonds	5	1
Bruns	0	14

La couleur des yeux est-elle indépendante de celle des cheveux ? Créer, par exemple, le jeu de données à l'aide de l'instruction suivante :

```
> cheveuxyeux <- as.table(matrix(c(5, 1, 0, 14), byrow = T,
+     nrow = 2, dimnames = list(c("Blonds", "Bruns"),
+         c("Clairs", "Foncés"))))
```

Dénombrements