

---

**Mathématiques générales (été)**  
**travaux pratiques 3**

---

1. On veut reproduire la courbe des probabilités d'au moins deux anniversaires le même jour dans une classe de  $n = 1, 2, 3, \dots$  étudiants. La courbe est tracée sur

<http://www.mste.uiuc.edu/reese/birthday/explanation.html>

Le complémentaire de cet évènement est que personne ne soit né le même jour parmi  $n$  personnes :

- Le nombre de possibilités de date pour une personne est 365.
- Dans un groupe de deux personnes, il y a  $365 \times 365$  choix de date d'anniversaire. Par contre il y a  $365 \times 364$  possibilités d'avoir deux anniversaires différents car lorsque la première personne est choisie, l'autre n'a plus que  $365 - 1 = 364$  possibilités d'avoir un autre jour de naissance. Ainsi la probabilité que deux personnes prises au hasard ne soit pas nées le même jour est  $\frac{365 \times 364}{365 \times 365}$ . Donc la probabilité d'au moins une coïncidence parmi deux personnes est

$$p_2 = 1 - \frac{365 \times 364}{365 \times 365}.$$

- Pour 3 personnes le raisonnement continue, et la probabilité d'au moins une coïncidence parmi trois personnes est :

$$p_3 = 1 - \frac{365 \times 364 \times 363}{365^3}$$

– etc.

Ecrire une fonction qui calcule ces probabilités  $p_n$  jusqu'à  $n$ , la variable de la fonction. Votre fonction aura par exemple la structure suivante :

```
anni = function(n){
  pp=365/365
  ps=pp
  for(i in 1:(n-1)){
    pp=pp* ??? /365
    ps=c(ps,pp)
  }
  out=1-ps
  return(out)
}

> probas=anni(50)
> plot(c(1:50), probas)
```

2. On veut estimer ces probabilités avec les données recueillies en classe. On considère  $n = 23$ , la taille telle que  $p_n \approx 0.5$ . On aimerait estimer  $p_{23}$  à partir des données.

Rappel : Les données récoltées en classe ont été concaténées à celles des années précédentes depuis 2007 et ont été enregistrées dans le fichier "sondage2007-8.txt". Ce dernier peut être téléchargé sur la page web du cours

<http://www.unige.ch/math/mgene/>

Une fois sauvé dans un répertoire, vous pouvez le lire dans R en tapant

```
> X=read.table("sondage2007-8.txt", header=T, sep=",")
```

puis taper

```
> attach(X)
```

pour pouvoir accéder aux variables en tapant leur nom.

Soit

```
> A=X[,5]
```

le vecteur des jours d'anniversaire.

Comment estimer  $p_{23}$  :

Familiarisez-vous avec la fonction `unique()` en tapant

```
> help(unique)
```

Que fait cette fonction appliquée à un vecteur ? Par exemple

```
> unique(c(1,0,0,1,3))
```

Ensuite regarder

```
A[(1 :23)]
```

et

```
unique(A[(1 :23)])
```

Ces vecteurs ont-ils la même longueur ? Si oui, alors toutes les dates sont différentes. Si non, alors il y a au moins une coïncidence.

Regarder maintenant la tranche suivante :

```
A[23+(1 :23)]
```

et

```
unique(A[23+(1 :23)])
```

Y a-t-il coïncidence dans ce groupe ?

Répéter  $M$  fois jusqu'à la dernière tranche, compter le nombre de fois  $m$  où il y a coïncidence puis estimer

$$\hat{p}_{23} = m/M.$$

Comparer à la probabilité théorique  $p_{23}$ .