

Exercice

On lance dé à 6 faces et on note X le résultat du lancer. Puis on lance un dé à X faces (plus exactement, on simule son lancer numériquement). On note Y le résultat final.

- Ecrire un programme R pour simuler $n = 100000$ observations, (y_1, y_2, \dots, y_n) , de cette expérience (commandes utiles : `sample` et `sapply`). Commenter les opérations effectuées (deux lignes).

```
n = 100000
# On simule n observations du dé à 6 faces
X <- sample(6, n, replace = TRUE)

# On simule ensuite n observations de dés à X faces
Y <- sapply(X, FUN = function(x) sample(x, 1))
```

- Donner une table des valeurs de probabilités empiriques issues de l'expérience simulée arrondies à 3 décimales (commandes utiles : `table` et `round`).

```
# table des valeurs de probabilités empiriques issues de l'expérience simulée
round(table(Y)/n, 3)
```

```
## Y
##   1    2    3    4    5    6
## 0.410 0.240 0.159 0.102 0.061 0.028
```

- Calculer la valeur moyenne de l'échantillon simulé (y_1, y_2, \dots, y_n) .

```
# valeur moyenne de l'échantillon simulé
mean(Y)
```

```
## [1] 2.2467
```

- A l'aide de la formule des probabilités totales, donner une formule mathématique décrivant la probabilité théorique $\text{Prob}(Y = i)$, pour tout $i = 1, \dots, 6$.

$$\text{Prob}(Y = i) = \sum_{j=1}^6 \text{Prob}(Y = i | X = j) \times \frac{1}{6} = \frac{1}{6} \sum_{j=i}^6 \text{Prob}(Y = i | X = j)$$

- Ecrire une fonction en R calculant la probabilité $\text{Prob}(Y = i)$, pour tout $i = 1, \dots, 6$.

```
# fonction calculant la probabilité $\{\text{rm Prob}\}(Y = i)$
prob <- function(i) {
  if (!(i %in% 1:6)) stop("Valeur attendue entre 1 et 6")

  probabilite <- sum(1/(6:i))/6
  return(probabilite)
}
```

- Donner une table des valeurs de probabilités théoriques arrondies à 3 décimales (commandes utiles : `sapply` et `round`).

```
round(sapply(1:6, FUN = prob), 3)
```

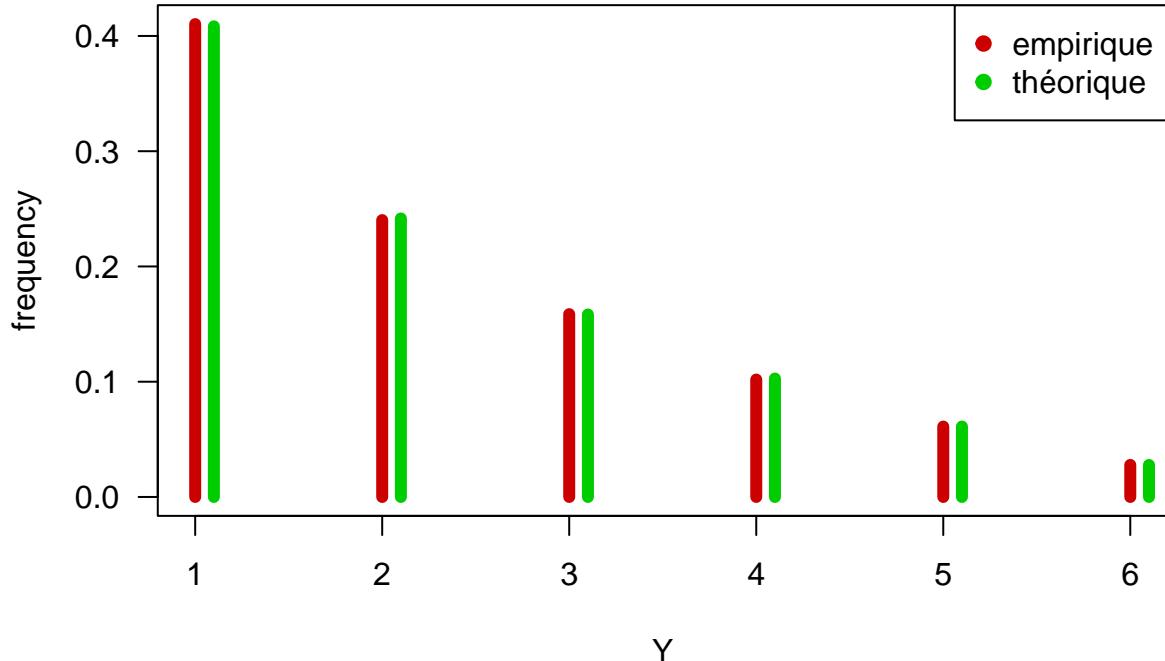
```
## [1] 0.408 0.242 0.158 0.103 0.061 0.028
```

- Comparer graphiquement les probabilités empiriques et théoriques.

```
# probabilités empiriques
plot(table(Y)/n,
      lwd = 6, col ="red3", ylab = "frequency", las = 1)
```

```
# probabilités théoriques
points(1:6+0.1, sapply(1:6, prob) ,
      type = "h", col = "green3", lwd = 6)

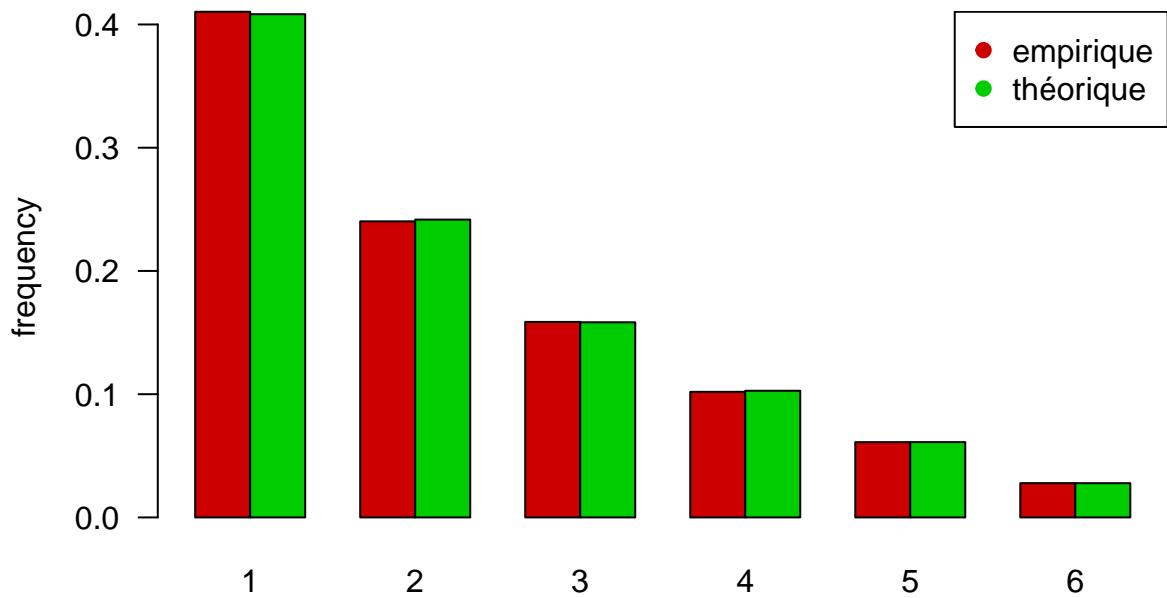
legend("topright", col = c("red3","green3"),
       pch =19,legend = c("empirique", "théorique"))
```



```
# probabilités empiriques et théoriques

barplot(t(cbind(table(Y)/n,  sapply(1:6, prob))),
        beside = TRUE,
        col =c("red3", "green3"),
        ylab = "frequency", las = 1)

legend("topright", col = c("red3","green3"),
       pch =19,legend = c("empirique", "théorique"))
```



- Calculer numériquement l'espérance de la variable Y (commandes utiles : `sapply` et `sum`). Comparer sa valeur à la valeur moyenne de l'échantillon simulé (y_1, y_2, \dots, y_n) .

```
# Calcul de l'espérance de la variable $Y$  
cat("Calcul de l'espérance de la variable : \n")  
  
## Calcul de l'espérance de la variable :  
sum((1:6)*sapply(1:6, prob))  
  
## [1] 2.25  
# Calcul de la moyenne de l'échantillon Y arrondie  
cat("Calcul de la moyenne empirique : \n")  
  
## Calcul de la moyenne empirique :  
round(mean(Y), 2)  
  
## [1] 2.25
```