

# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## LE PLAN DU DOCUMENT RESSOURCE

<b>Introduction .....</b>	<b>1</b>
<b>I. Compléments sur les lois uniforme et exponentielle.....</b>	<b>4</b>
A. Loi uniforme.....	4
B. Lois exponentielles.....	6
C. Méthode de Monte Carlo (AP en TS).....	7
1. Méthode dite du « rejet ».....	7
2. Méthode de l'espérance.....	9
<b>II. Variable centrée réduite.....</b>	<b>10</b>
A. Comment centrer et réduire.....	10
B. Pourquoi centrer et réduire ?.....	10
<b>III. La loi normale centrée réduite .....</b>	<b>11</b>
A. Activité : Introduction au théorème de Moivre-Laplace.....	11
B. Théorème de Moivre-Laplace (admis) (Programme TS uniquement) .....	13
C. La loi normale centrée réduite .....	14
1. Premières propriétés .....	14
2. Espérance d'une loi normale centrée réduite (uniquement TS) .....	15
<b>IV. Lois normales.....</b>	<b>16</b>
A. Introduction .....	16
B. Définition et exemple.....	16
C. Exemples de tp / exercices .....	18
Exemples d'exercices.....	20
<b>V. Intervalle de fluctuation.....</b>	<b>25</b>
A. Définition.....	25
B. Activité : recherche et utilisation d'un intervalle de fluctuation à l'aide d'un algorithme.....	26
C. Intervalle de fluctuation asymptotique .....	27
1. Théorème (uniquement en TS) .....	27
2. Application .....	28
3. Définition.....	29
4. Premier exemple d'utilisation Prise de décision .....	29
5. Deuxième exemple d'utilisation Problème du surbooking .....	31
6. Intervalle de fluctuation simplifié donné en seconde.....	33
7. Théorème .....	34
8. Exemple d'activité .....	35
9. À propos des approximations usuelles des lois binomiales par des lois normales .....	36

# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## LE PLAN DU DOCUMENT RESSOURCE

<b>VI. Intervalle de confiance .....</b>	<b>39</b>
A. Introduction .....	39
<i>Activité</i> .....	40
B. Principe général de l'intervalle de confiance.....	41
C. Définition.....	42
D. Intervalle de fluctuation ou intervalle de confiance : lequel utiliser ? .....	43
E. Autre intervalle de confiance .....	44
F. Étude de la longueur de l'intervalle de fluctuation et conséquence pour l'intervalle de confiance .....	46
G. Détermination de la taille minimale de l'échantillon pour avoir une précision donnée .....	46
H. Applications.....	47
1. <i>Simulations</i> .....	47
2. <i>Exemple de recherche d'un intervalle de confiance</i> .....	48
3. <i>Activités tableau permettant de simuler des intervalles de confiance</i> .....	49
4. <i>Exemples d'exercices</i> .....	50
<b>Annexe 1 Introduction au théorème de Moivre-Laplace .....</b>	<b>53</b>
I. La loi des grands nombres de Jacques Bernoulli .....	53
II. La démarche d'Abraham de Moivre .....	54
III. Une approche du résultat de Moivre.....	55
IV. Le théorème de MOIVRE-LAPLACE.....	57
V. Convergence en loi.....	58
<b>Annexe 2 Compléments sur les lois normales .....</b>	<b>59</b>
I. Loi normale centrée réduite.....	59
II. Lois normales .....	60
<b>Annexe 3 Approche simplifiée de la théorie des sondages .....</b>	<b>62</b>
A. Qualités d'un échantillon permettant de répondre à une question posée.....	62
B. Échantillonnage non-probabiliste ou non aléatoire.....	63
C. Échantillonnage probabiliste .....	64
<b>Annexe 4 Utilisation des Tice.....</b>	<b>66</b>
A. Tableau des fichiers du document ressource du programme de Terminale Probabilités et Statistique..	66
B. Prise en main rapide du logiciel R.....	69

# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## LA PARTIE TICE

Les logiciels utilisés sont : libre office calc, excel, géogébra, algobox, scilab et R.  
Pour faciliter l'utilisation des ressources, tous les fichiers ayant servi à illustrer le document ressource et les fichiers d'autres activités entrant dans le cadre du programme sont référencés dans un tableau.

ORGANISATION du tableau des fichiers du documents d'accompagnement du programme de Term S 2011 – Probabilités et statistique (pro-parte) :

- ▶ Les parties en fond bleu clair concernent le calcul de probabilités connaissant les lois.
- ▶ Les parties en fond blanc concernent le calcul approché de probabilités par simulation.
- ▶ *Les textes en italique ou en italique vert* concernent des fichiers non référencés de certaines figures du document principal ou des fichiers d'activités complémentaires n'apparaissant pas dans le document principal.

DOCUMENTS DE SYNTHÈSE	FICHIERS	FONCTIONS, PARAMÈTRES D'ANIMATION OU DE FONCTION ET DESCRIPTION <b>n (10; 60; 1; curs)</b> signifie que l'on peut faire varier le paramètre n de 10 à 60 avec des pas de 1, à l'aide d'un curseur. <b>curs</b> peut être remplacé par <b>bouton</b> (à cliquer).
V-C-2	<i>exploration de l'intervalle de fluctuation asymptotique.r</i>	<b>pIFasy2_1(n = 400, p= .5, proba = .95)</b> <b>Fonction en R</b> : illustration de l'évolution de la probabilité binomiale de l'intervalle de fluctuation asymptotique <b>IF2</b> : $(p \pm u_{proba} \times \text{racine}(p \times (1-p) / n))$ en fonction de n et p. n est la taille de l'échantillon, p est la probabilité de succès, <b>proba</b> est la valeur seuil de la probabilité de l'IF.
	<i>SimulICPropSimpl.r</i>	<i>simIC(n = 50, nbsim = 100, nbclass = 20, moustache = 1.5)</i> <i>Fonction de R : Simulation d'un peigne d'IC. nbclass est le nombre de classes de l'histogramme, moustache détermine la longueur des moustaches des boites. La proportion de p dans la population est générée aléatoirement dans ]0; 1[.</i> <i>Le peigne est suivi de l'histogramme et de la boite à moustache de la distribution simulée.</i>
<i>RapideR.odt</i>	<i>InitiationR1.r</i>	<i>Démarrage rapide en l'R : installation et quelques exemples commentés, bibliographie.</i>

# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## QU'EST-CE QUE R ?

### En quelques mots,

R est un logiciel de développement scientifique spécialisé dans le calcul et l'**analyse statistique**.

### R est aussi

- ▶ un langage,
- ▶ un environnement,
- ▶ un projet open source (projet GNU),
- ▶ un logiciel multi-plateforme (Linux, Mac, Windows),
- ▶ la 18e lettre de l'alphabet ☺.

# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## QUE FAIT R ?

1. Gestionnaire de **données**
  - ▶ Lecture, manipulation, stockage.
2. Algèbre linéaire
  - ▶ Opérations classiques sur vecteurs, tableaux et matrices
3. **Statistiques et analyse de données**
  - ▶ Dispose d'un grand nombre de méthodes d'analyse de données (des plus anciennes aux plus récentes)
4. Moteur de **sorties graphiques**
  - ▶ Sorties écran ou fichier
5. Système de modules (packages)
  - ▶ Alimenté par la communauté (+ de 2000 extensions !)

# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## QUI UTILISE R ?

EN FRANCE ET À L'ÉTRANGER  
BEAUCOUP D'UNIVERSITÉS

- ▶ Rennes, Lyon, Evry, médecine Nancy, ...

BEAUCOUP D'ÉCOLES D'INGÉNIEURS

- ▶ Montpellier SupAgro, Polytech'Nice Sofia, ...

DES ORGANISMES DE RECHERCHE

- ▶ CNRS, INRA, MNHN, INSERM ...

BEAUCOUP D'IUT (STID)

- ▶ Bordeaux, Metz ...

DES ENTREPRISES

- ▶ (article de P.-A. Cornillon)

# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

DE LA DOC., DES LIVRES, DES COURS, SUR R ... ? ... EN FRANÇAIS ??

À FOISON SUR LES SITES DES ÉCOLES, DES UNIVERSITÉS, DES  
INSTITUTS ..., DES LIBRAIRIES ...

- ▶ Pierre-André Cornillon<sup>1</sup>, & Eric Matzner-Løber ; *Atouts et faiblesses de R en enseignement, recherche et industrie.*
- ▶ Lafaye De Micheaux Pierre, Drouilhet Remy, Liqueur Benoit ; 2010 ; *Le logiciel R, maîtriser le langage, effectuer des analyses statistiques* ; Springer.
- ▶ Millot Gaël ; 2009 ; *Comprendre et réaliser des tests statistiques à l'aide de R ; manuel pour les débutants* ; De Boeck.
- ▶ François Husson ; Sébastien Lê ; Jérôme Pagès ; 2009 ; *Analyse de données avec R* ; P U De Rennes.
- ▶ Pierre-André Cornillon ; 2010 ; *Régression avec R* ; Springer.
- ▶ Pierre-André Cornillon, Arnaud Guyader, François Husson et al. ; 2010 ; *Statistique avec R (2e édition)* ; P U De Rennes.
- ▶ Deepayan Sarkar ; 2008 ; *Lattice, Multivariate Data Visualization with R* ; Springer.
- ▶ Frédéric Bertrand ; 2010 ; *Initiation aux statistiques avec R ; cours, exemples, exercices et problèmes corrigés* ; Licence 3, Master 1, écoles d'ingénieur ; Dunod.
- ▶ Yadolah Dodge ; 2008 ; *Premiers pas en simulation* ; Springer Verlag.
- ▶ Christian P. Robert, George Casella ; 2011 ; *Méthodes de Monte-Carlo avec R* ; Springer Verlag.
- ▶ <http://pbil.univ-lyon1.fr/R/enseignement.html>
- ▶ <http://rug.mnhn.fr/semin-r/>

...

# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## POURQUOI R AU LYCÉE ?

1. C'est un outil polyvalent permettant de mettre en œuvre, d'illustrer :
  - ▶ Calcul numérique, analyse, représentations graphiques, analyse exploratoire des données, probabilités, statistiques, simulation, algorithmique.
2. Il permet de faire, de la simulation, un outil de résolution de problèmes, à part entière :
  - ▶ Le calcul des probabilité et la simulation en probabilité sont un terrain privilégié pour mettre en œuvre le niveau d'algorithmique du programme et donc lui donner tout son intérêt.



# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## R AU LYCÉE ET POURQUOI PAS AUSSI AU COLLÈGE ?

3. Il permet de faire très facilement de l'analyse exploratoire des données (au sens de Tukey)
- ▶ Par exemple pour explorer graphiquement les séries simulées, et ainsi donner du goût à la statistique descriptive au collège et au lycée.
  - ▶ Organiser, décrire et analyser des données pour étudier conjointement des variables qualitatives et quantitatives est la spécificité des logiciels de statistique à ne pas confondre avec les tableurs, par exemple.

# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## DEUX DES INTERFACES DE R

### 1° - Rcmdr : les menus à cliquer

The screenshot displays the R Commander interface. The main window shows a script with the following R code:

```
EchantillonsNormaux <- as.data.frame(matrix(rnorm(100*1, mean=0, sd=1), ncol=1))
rownames(EchantillonsNormaux) <- paste("sample", 1:100, sep="")
colnames(EchantillonsNormaux) <- "obs"
EchantillonsNormaux[1:10,]
stem.leaf(EchantillonsNormaux$obs, trim.outliers=FALSE, na.rm=TRUE)
Hist(EchantillonsNormaux$obs, scale="frequency", breaks="Sturges", col="darkgray")
```

The console window shows the output of the `stem.leaf` function:

```
1 | 2: represents 1.2
[1]
      n: 100
      3  -2* | 001
      7  -1. | 6889
     14 -1* | 0222444
     29 -0. | 55566667777799
     48 -0* | 000000011122223444
    (25) 0* | 000011122333333444444444
     27  0. | 566777788889999
     11  1* | 000113
      5   1. | 559
      2   2* | 23
```

A dialog box titled "Histogramme" is open, showing the variable "obs" selected, with "Nombre de classes" set to "auto". The "Echelle des axes" section has "Fréquences" selected.

The "R Graphics: Device 2 (ACTIVE)" window displays a histogram of the data. The x-axis is labeled "EchantillonsNormaux\$obs" and ranges from -2 to 2. The y-axis is labeled "frequency" and ranges from 0 to 25. The histogram shows a distribution centered around 0.

# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## DEUX DES INTERFACES DE R 2° - La console R : les lignes de commandes

The image displays four overlapping screenshots of the R environment, illustrating the workflow from code execution to data visualization.

**Top-left screenshot (R Console):** Shows the initial code execution:

```
> piece1 <- sample(c("Pile", "Face"), size = 1000, prob = c(.4, .6), r  
> (distpiece1 <- table(piece1))|
```

**Top-right screenshot (R Console):** Shows the output of the first code block:

```
> piece1 <- sample(c("Pile", "Face"), size = 1000, prob = c(.4, .6), r  
> (distpiece1 <- table(piece1))  
piece1  
Face Pile  
624 376  
> |
```

**Middle-left screenshot (R Console):** Shows the execution of a second code block:

```
> piece1 <- sample(c("Pile", "Face"), size = 1000, prob = c(.4, .6), r  
> (distpiece1 <- table(piece1))  
piece1  
Face Pile  
624 376  
> sum(piece1 == "Pile") / 1000|
```

**Middle-right screenshot (R Console):** Shows the output of the second code block:

```
> piece1 <- sample(c("Pile", "Face"), size = 1000, prob = c(.4, .6), r  
> (distpiece1 <- table(piece1))  
piece1  
Face Pile  
624 376  
> sum(piece1 == "Pile") / 1000  
[1] 0.376  
> |
```

**Bottom-left screenshot (R Console):** Shows the execution of a third code block:

```
> piece1 <- sample(c("Pile", "Face"), size = 1000, prob = c(.4, .6), r  
> (distpiece1 <- table(piece1))  
piece1  
Face Pile  
624 376  
> sum(piece1 == "Pile") / 1000  
[1] 0.376  
> barplot(distpiece1 / 1000)|  
> |
```

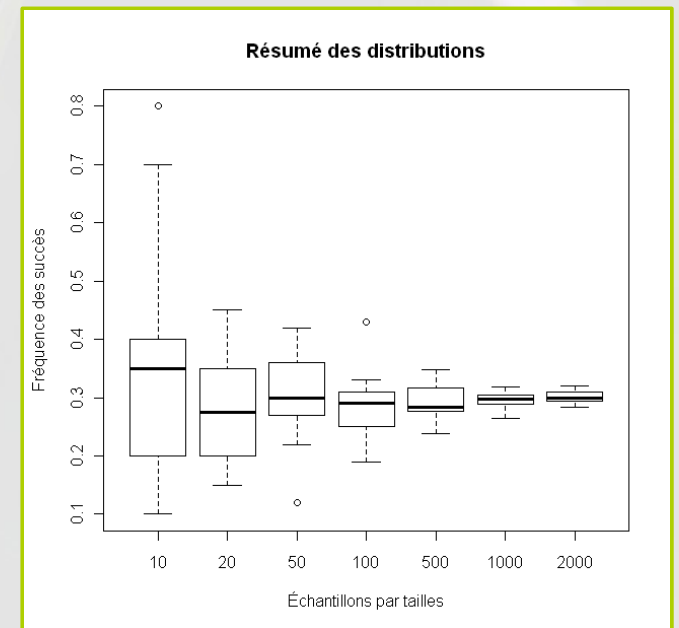
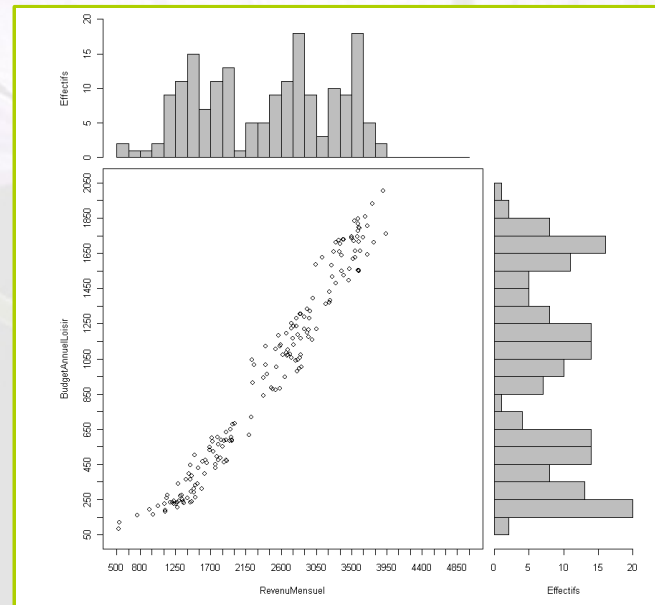
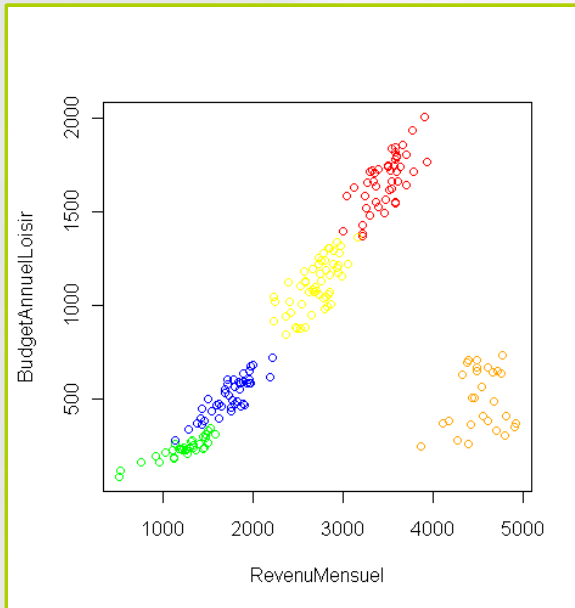
**Bottom-right screenshot (R Graphics: Device 2 (ACTIVE)):** Shows a barplot of the relative frequencies of "Face" and "Pile". The y-axis represents the relative frequency, ranging from 0.0 to 0.6. The x-axis has two categories: "Face" and "Pile". The bar for "Face" has a height of approximately 0.624, and the bar for "Pile" has a height of approximately 0.376.

Category	Count	Relative Frequency
Face	624	0.624
Pile	376	0.376

# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## DEUX DES INTERFACES DE R

*La console R : les commandes d'un puissant moteur graphique*



**QUELQUES EXEMPLES DE TRÈS SIMPLES À SIMPLES :**

# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## LA SIMULATION : UNE AUTRE FAÇON DE FAIRE DES PROBABILITÉS À TRAVERS QUELQUES EXEMPLES

LES UNIVERS D'UN DÉ ET LEURS TABLEAUX DE FRÉQUENCES

PROBLÈME HISTORIQUE D'UN GRAND DUC DE TOSCANE : LES 3 DÉES ALGO 1

PROBLÈME HISTORIQUE D'UN GRAND DUC DE TOSCANE : LES 3 DÉES ALGO 2

LE PROBLÈME DES "CHAINES" DE LONGUEUR 6 AVEC 200 JETS D'UNE PIÈCE

INTERVALLE DE FLUCTUATION SIMULÉ, RELATIF À UNE PROPORTION

**DISTRIBUTION \* SIGNIFIE DISTRIBUTION SIMULÉE**  
(d'une façon générale, \* indique le caractère simulé de l'objet marqué)

# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## LA SIMULATION : UNE AUTRE FAÇON DE FAIRE DES PROBABILITÉS À TRAVERS QUELQUES EXEMPLES (SUITE IMPROBABLE)

EXPLORATION D'UN INTERVALLE DE FLUCTUATION ASYMPTOTIQUE

SIMULATION D'UN PEIGNE D'INTERVALLES DE CONFIANCE

LES DISTRIBUTIONS DANS LA LOI DES GRANDS NOMBRES

*Jet simultané de deux pièces identiques :  
quel univers et quelle distribution de probabilité ?*

*Croix ou Pile de D'Alembert*

*On jette deux fois une pièce équilibrée (croix - pile) :  
Distribution de probabilité du rang d'arrivée du premier croix*

*Un modèle d'urne virtuelle :  
comparer un tirage sans remise et un tirage avec remise*

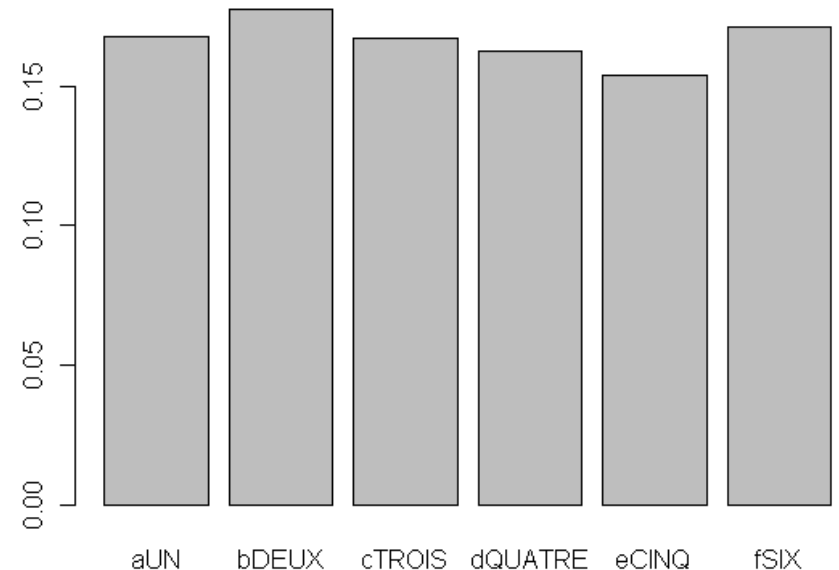
# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## LA SIMULATION : UNE AUTRE FAÇON DE FAIRE DES PROBABILITÉS

### JETER UN DÉ : LES UNIVERS ET LEURS TABLEAUX DE FRÉQUENCES

```
deactel <- function(de = c("aUN", "bDEUX", "cTROIS", "dQUATRE", "eCINQ", "fSIX"),  
                  nbsim = 2000) {  
  experience <- sample(de, nbsim, replace = TRUE)  
  distfreq <- table(experience)  
  
  #***** Affichage des résultats *****  
  print(distfreq)  
  barplot(distfreq / nbsim)  
}
```

```
> deactel()  
Experience (DISTRIBUTION*)  
 aUN   bDEUX  cTROIS dQUATRE  eCINQ   fSIX  
 336   355   334   325   308   342
```



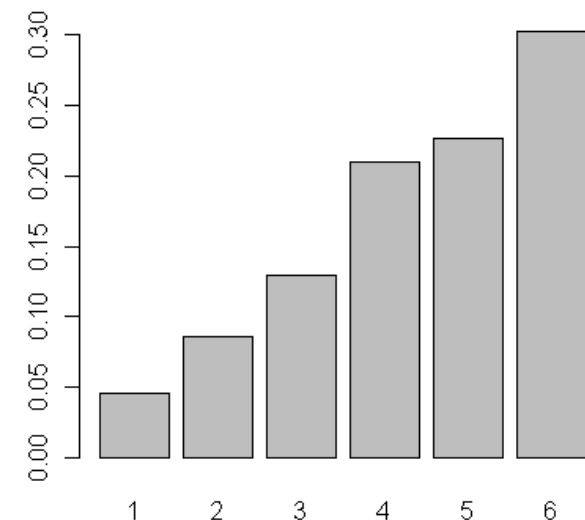
# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## LA SIMULATION : UNE AUTRE FAÇON DE FAIRE DES PROBABILITÉS

### JETER UN DÉ : LES UNIVERS ET LEURS TABLEAUX DE FRÉQUENCES

```
deactel <- fonction(de = c("aUN", "bDEUX", "cTROIS", "dQUATRE", "eCINQ", "fSIX"),  
                  proba = rep(1 / 6, 6), nbsim = 2000){  
  experience <- sample(de, nbsim, replace = TRUE, prob = proba)  
  distfreq <- table(experience)  
  
  #***** Affichage des résultats *****  
  print(distfreq)  
  barplot(distfreq / nbsim)  
}
```

```
> deactel(de = seq(1, 6, 1),  
          proba = c(1/21, 2/21, 3/21, 4/21, 5/21, 6/21))  
Experience (DISTRIBUTION*)  
 1     2     3     4     5     6  
91 172 259 420 453 605
```





# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

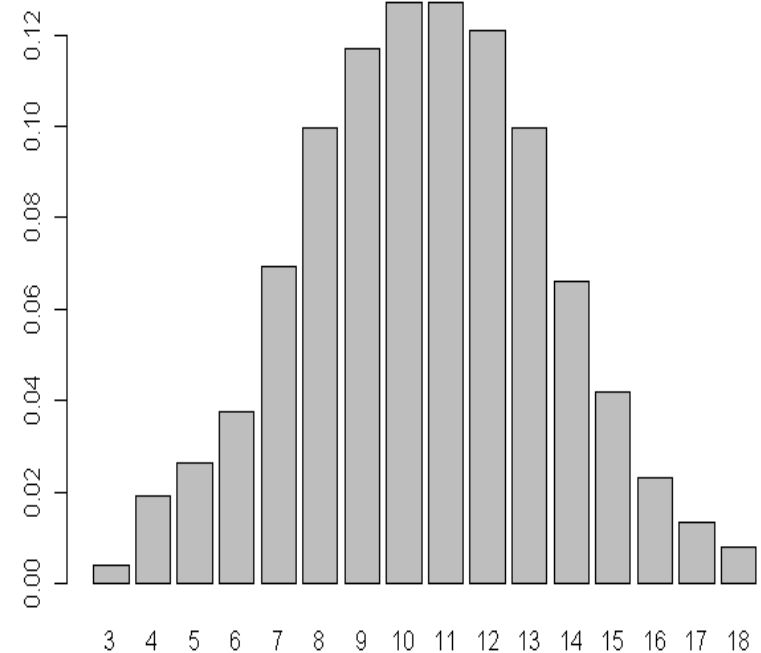
## LA SIMULATION : UNE AUTRE FAÇON DE FAIRE DES PROBABILITÉS

### PROBLÈME HISTORIQUE D'UN GLAND DUC DE TOSCANE : JOUER 3 DÉS ALGORITHME 1

```
toscanel <- function(nbsim = 2000){  
  de1 <- sample(seq(1, 6, 1), nbsim, replace = T)  
  de2 <- sample(seq(1, 6, 1), nbsim, replace = T)  
  de3 <- sample(seq(1, 6, 1), nbsim, replace = T)  
  som3des <- de1 + de2 + de3  
  distsom <- table(som3des)  
  #***** Affichage des résultats *****  
  print(distsom)  
  barplot(distsom / nbsim)  
}
```

```
> toscanel()  
som3des (DISTRIBUTION*)  
 3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18  
8 38 53 75 139 199 234 254 254 242 199 132 84 46 27 16
```

```
> system.time(toscanel())  
som3des  
 3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18  
 7 34 60 101 153 190 236 233 245 243 183 112 106 51 30 16  
utilisateur      système      écoulé  
      0.03          0.00          0.03
```



# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

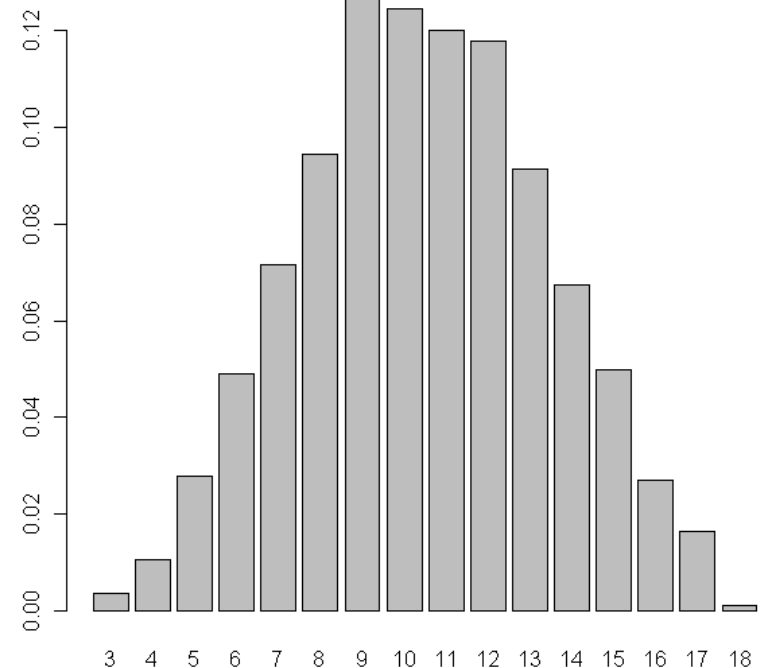
## LA SIMULATION : UNE AUTRE FAÇON DE FAIRE DES PROBABILITÉS

### PROBLÈME HISTORIQUE D'UN GLAND DUC DE TOSCANE : JOUER 3 DÉS ALGORITHME 2

```
toscane2 <- function(nbsim = 2000) {  
  som3des <- rep(NA, nbsim)  
  for(i in 1:nbsim) {  
    de1 <- sample(seq(1, 6, 1), 1)  
    de2 <- sample(seq(1, 6, 1), 1)  
    de3 <- sample(seq(1, 6, 1), 1)  
    som3des[i] <- de1 + de2 + de3  
  }  
  distsom <- table(som3des)  
  
  #***** Affichage des résultats *****  
  print(distsom)  
  barplot(distsom / nbsim)  
}
```

```
> toscane2()  
som3des (DISTRIBUTION*)  
 3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18  
 7 21 56 98 143 189 254 249 240 236 183 135 100 54 33 2
```

```
> system.time(toscane2())  
som3des  
 3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18  
10 24 47 90 140 197 216 245 238 242 183 157 102 60 42 7  
utilisateur      système      écoulé  
      0.80          0.00          0.79
```



# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

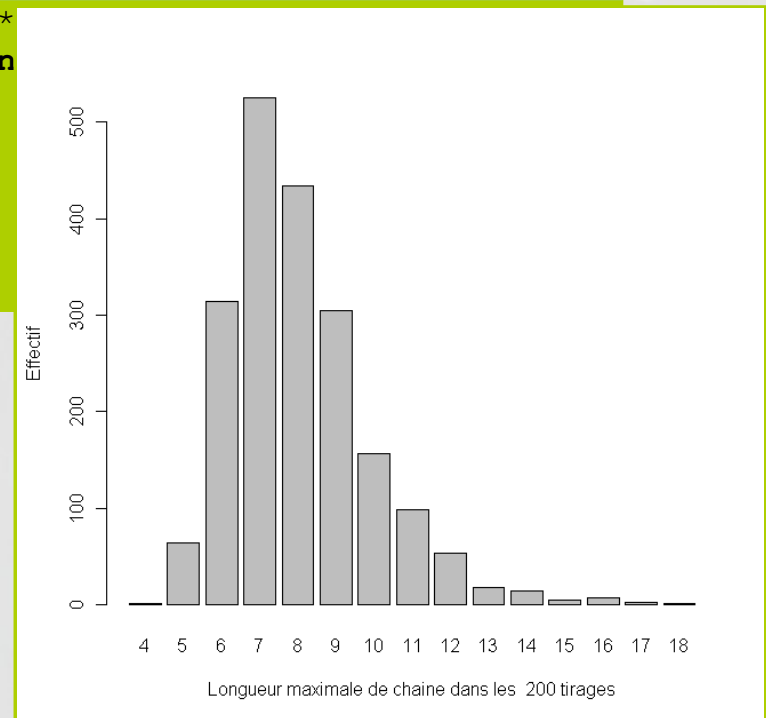
## LA SIMULATION : UNE AUTRE FAÇON DE FAIRE DES PROBABILITÉS

### LE PROBLÈME DES "CHAINES" DE LONGUEUR 6 AVEC 200 JETS D'UNE PIÈCE

```
chainesdSimpl = function(L = 6, tirages = 200, nbsim = 2000) {  
  nblongMax <- rep(NA, nbsim)  
  for(i in 1:nbsim) {  
    expe <- sample(c("Pile", "Face"), tirages, replace = T)  
    rexpe <- rle(expe)  
    nblongMax[i] <- max(rexpe$length)  
  }  
  tablemax <- table(nblongMax)  
  StatL <- sum(tablemax[(as.numeric(names(tablemax)) >= L)]) / nbsim
```

```
#***** Affichage des résultats *****  
cat("Une estimation de la probabilité de chaines de lon  
cat(" vaut :", StatL, "\n \n")  
  
print("Distribution* de la chaine de longueur maximale  
print(tablemax)  
barplot(tablemax)  
}
```

```
> chainesdSimpl()  
Une estimation de la probabilité de chaines de longueur 6 ou  
plus vaut : 0.9675  
  
[1] "Distribution* de la chaine de longueur maximale (runs)"  
nblongMax  
 4  5  6  7  8  9 10 11 12 13 14 15 16 17 18  
1 64 314 525 434 305 157 99 54 18 14 5 7 2 1
```



# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

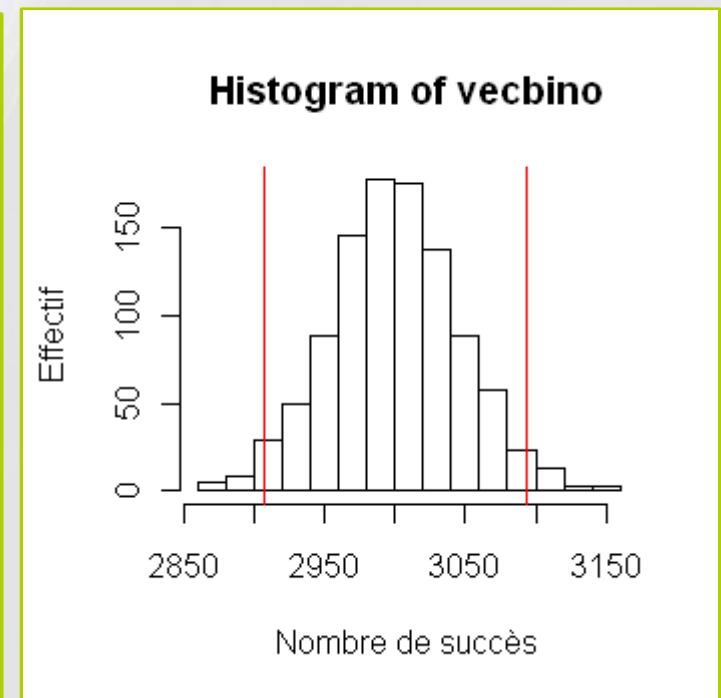
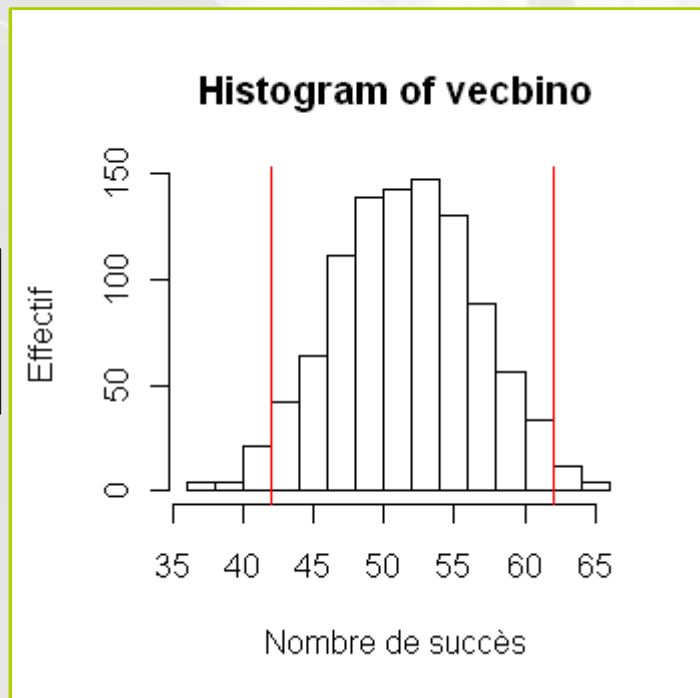
## LA SIMULATION : UNE AUTRE FAÇON DE FAIRE DES PROBABILITÉS

### INTERVALLE DE FLUCTUATION SIMULÉ, RELATIF À UNE PROPORTION

```
simulif <- function(n = 100, p = .52, nsim = 1000){  
  vecbino <- rbinom(nsim, n, p)  
  quantiles <- quantile(vecbino, probs = c(2.5, 97.5)/100)  
  #***** Affichage des résultats *****  
  print("Quantiles*")  
  print(quantiles)  
  hist(effec, xlab = "Nombre de succès", ylab = "Effectif")  
  abline(v = quantiles, col = "red")  
}
```

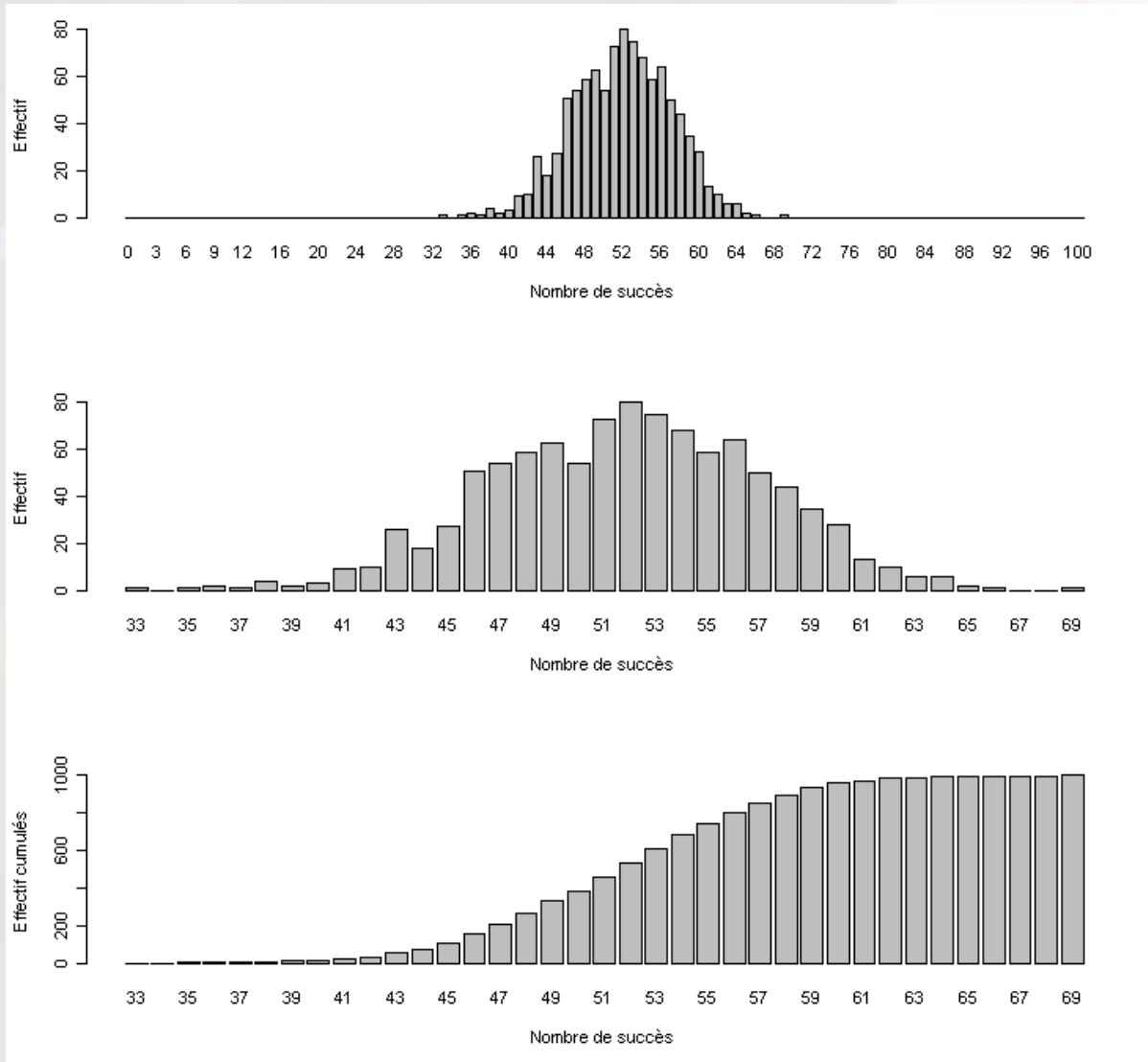
```
> simulif()  
[1] "Quantiles*"  
2.5% 97.5%  
42 62
```

```
> simulif(n=10000, p=.3)  
[1] "Quantiles*"  
2.5% 97.5%  
2908 3094
```



# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## INTERVALLE DE FLUCTUATION SIMULÉ, RELATIF À UNE PROPORTION



# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## EXPLORATION D'UN INTERVALLE DE FLUCTUATION ASYMPTOTIQUE

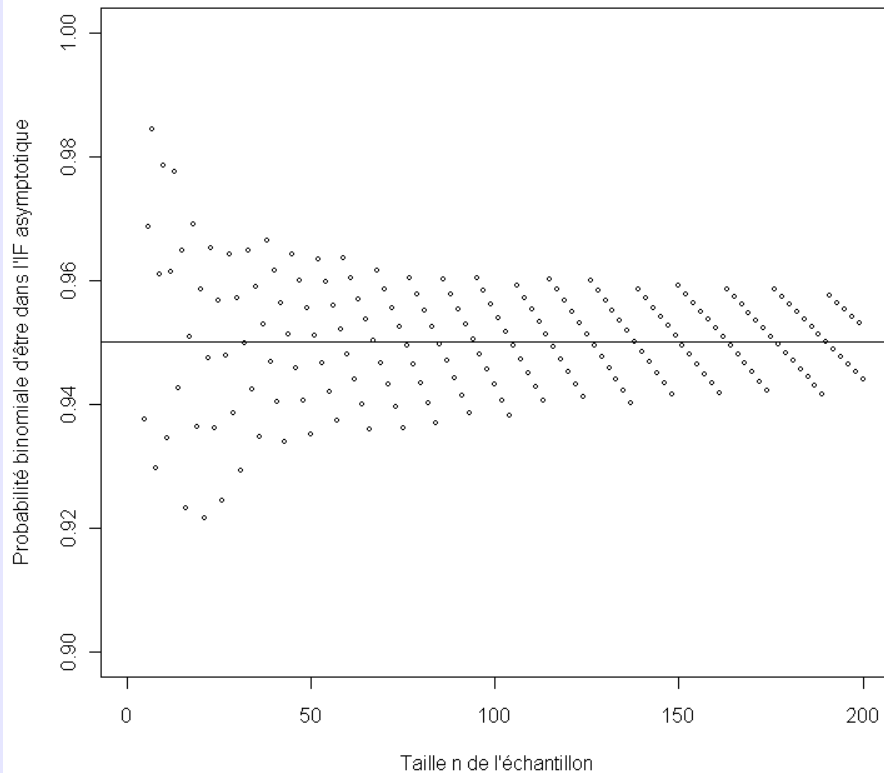
```
pIFasy2_1 <- function(n = 200, p = .5, proba = .95) {  
  P <- function(n, p, proba) {  
    g <- qnorm(1 - (1 - proba) / 2)  
    binf <- max(floor(n * p - g * sqrt(p * (1 - p) * n)), 0)  
    bsup <- min(floor(n * p + g * sqrt(p * (1 - p) * n)), n)  
    sum(dbinom((binf + 1):bsup, n, p))  
  }  
  
  y <- vector(length = n)  
  for(i in 1:n) y[i] <- P(i, p, proba)  
  
  #*****Affichage des graphiques*****  
  plot(1:n, y, cex = .5, ylim = c(.9, 1),  
       xlab = "Taille n de l'échantillon",  
       ylab = "Probabilité binomiale d'être dans l'IF asymptotique",  
       main = "Exploration de l'intervalle de fluctuation : influence de n")  
  abline(h = proba)  
}
```

# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## EXPLORATION D'UN INTERVALLE DE FLUCTUATION ASYMPTOTIQUE

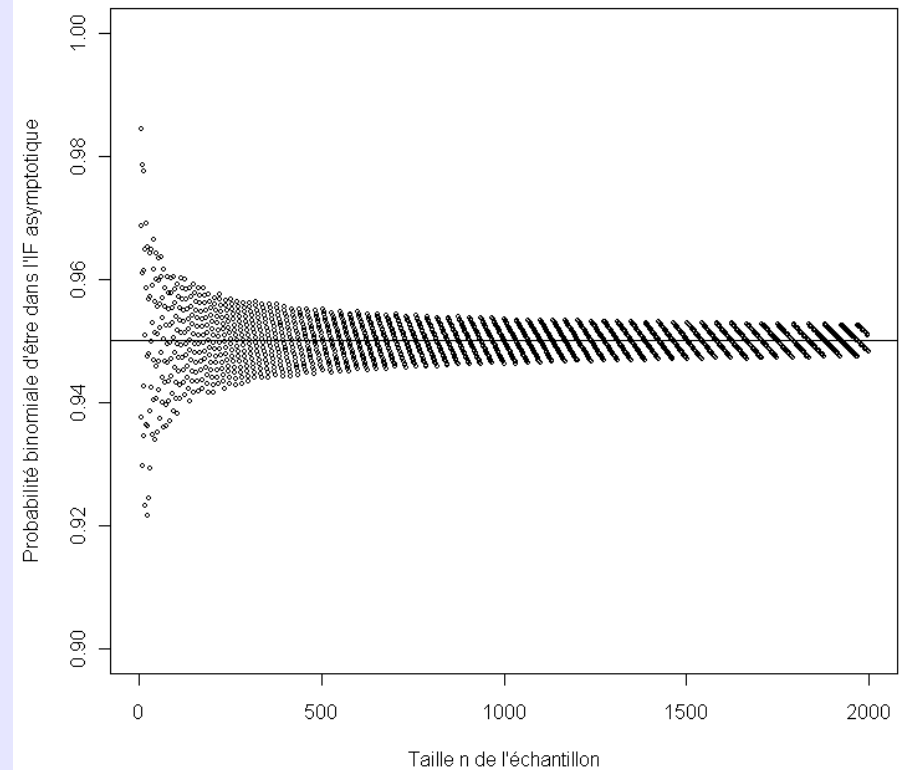
pIFasy2\_1 (n = 200, p = .5)

Exploration de l'intervalle de fluctuation : influence de n avec ppop = 0.5



pIFasy2\_1 (n = 2000, p = .5)

Exploration de l'intervalle de fluctuation : influence de n avec ppop = 0.5



# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## LA SIMULATION : UNE AUTRE FAÇON DE FAIRE DES PROBABILITÉS

### SIMULATION D'UN PEIGNE D'INTERVALLES DE CONFIANCE

```
simIC <- function(n = 50, p = .3, nbsim = 100, nbclass = 20){
  p <- round(runif(1, .1, .9), 2)
  x <- 1:nbsim
  vectpbino <- rbinom(nbsim, n, p) / n
  blsup <- vectpbino + 1 / sqrt(n) ; blinf <- vectpbino - 1 / sqrt(n)
  ymax <- 1.1 ; ymin <- - .1
  dehorsic1 <- sum(blup > p & blinf > p) + sum(blup < p & blinf < p)

  ***** Affichage des résultats *****
  cat("\n p extérieur à l'IC1 :", dehorsic1, "fréquence =", dehorsic1 / nbsim, "\n")
  print(p)

  ***** Affichage des graphiques*****
  #-----Peigne d'intervalles de confiance simulés-----
  par(mfrow = c(2, 1))

  plot(x, vectpbino, ylim = c(ymin, ymax), cex = .5, col = "green",
       xlab = paste(nbsim, "simulations"), ylab = "proportion de succès dans l'échantillon",
       main = paste("I. C.* Asymptotiques au niveau 95%, d'une proportion dans la population\n",
                    "Calculés à partir de", nbsim, "échantillons simulés, de taille", n))

  points(x, blsup, cex = .5)
  points(x, blinf, cex = .5)

  for(i in x) lines(c(i, i), c(blup[i], blinf[i]))

  #-----Histogramme d'une série simulée-----
  hist(vectpbino, breaks = seq(0, 1, 1 / nbclass), freq = F, xaxp = c(0, 1, 20),
       xlab = "Classes de proportions", ylab = "Densité", cex.axis = .7, border = "green",
       main = paste("Histogramme et boîte à moustache de", nbsim, "proportions simulées"))

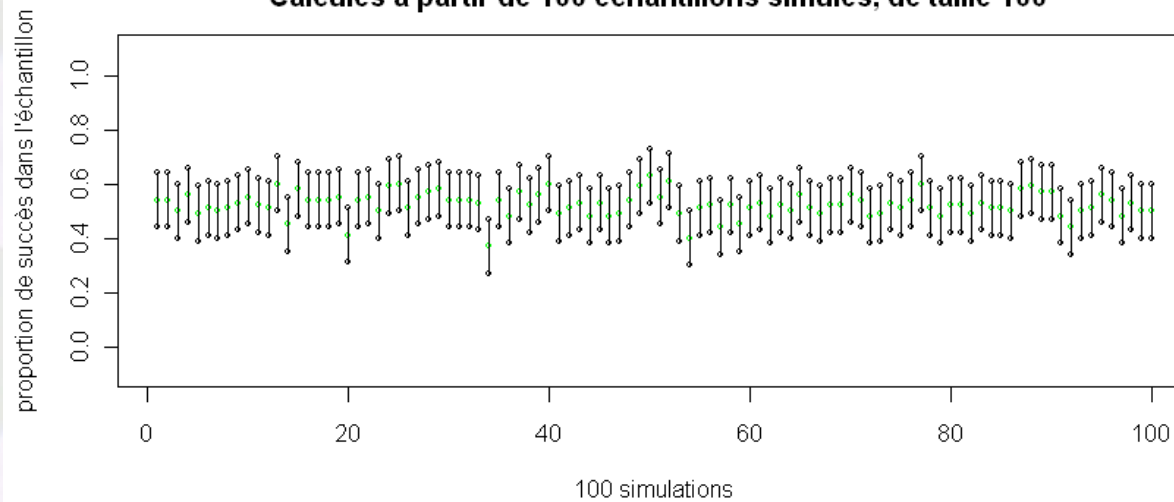
  #-----Boîte à moustaches d'une série simulée-----
  boxplot(vectpbino, freq = F, horizontal = T, add = T, xaxt = "n", col = "green")
}
```



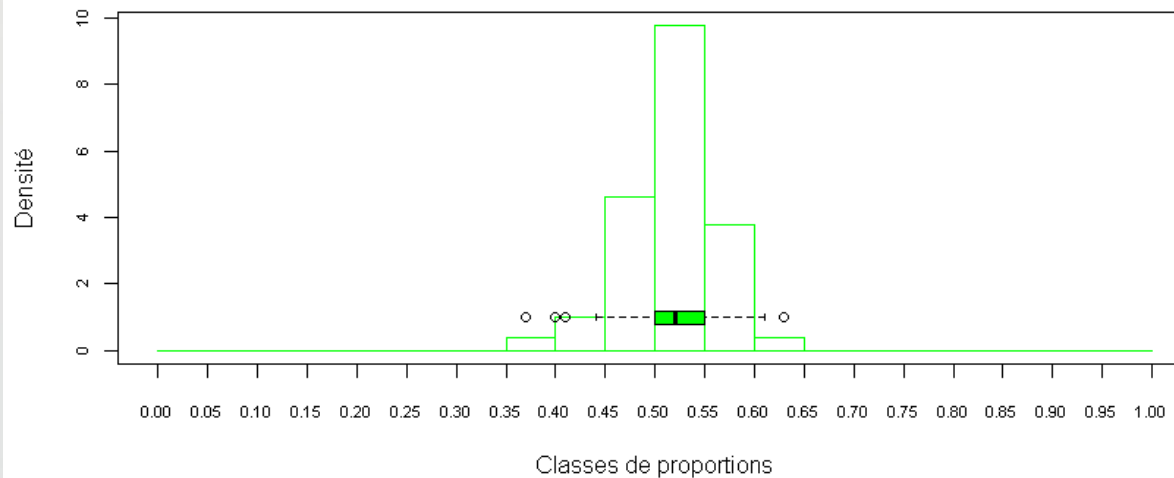
# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## SIMULATION D'UN PEIGNE D'INTERVALLES DE CONFIANCE

I. C. Asymptotiques au niveau 95%, d'une proportion dans la population  
Calculés à partir de 100 échantillons simulés, de taille 100



Histogramme et boîte à moustache de 100 proportions simulées



# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## LA SIMULATION : UNE AUTRE FAÇON DE FAIRE DES PROBABILITÉS

### LES DISTRIBUTIONS DANS LA LOI DES GRANDS NOMBRES :

*Exemple d'une variable binomiale et de la variable fréquence associée*

*Chaque fréquence est le résumé d'une distribution.*

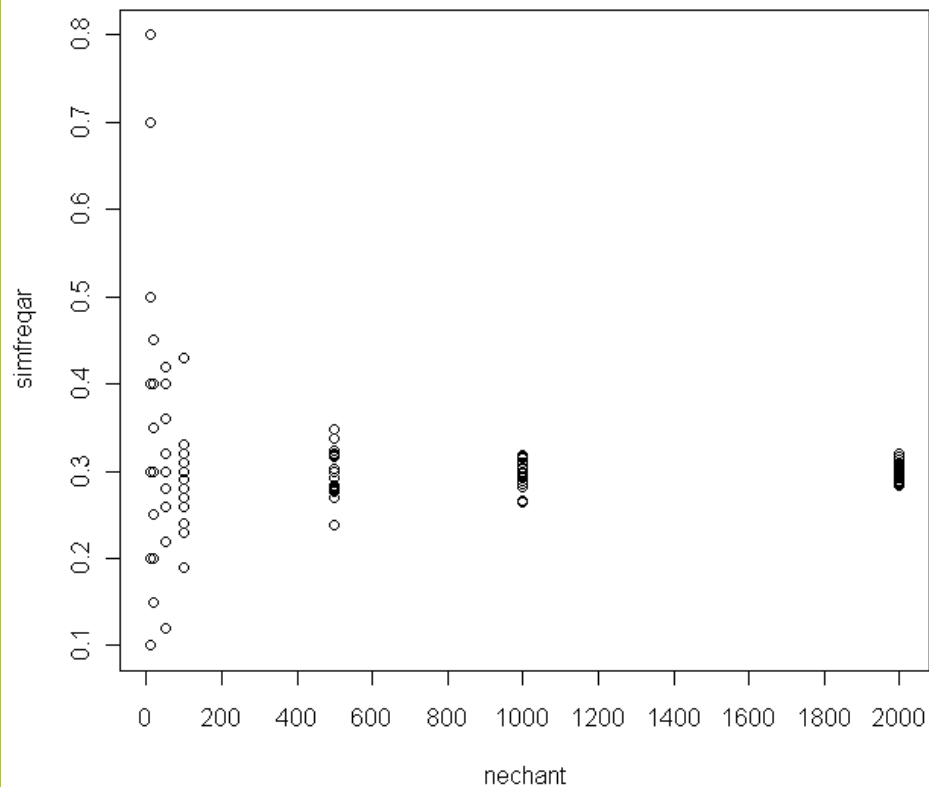
*Lorsque  $n$  tend vers l'infini, la probabilité que la fréquence s'éloigne de la proportion tend vers 0.*

```
lgn1 <- function(nbsim = 20, p = .3, n7 = c(10, 20, 50, 100, 500, 1000, 2000)){
  nechant <- rep(n7, each = nbsim)
  simfreqar <- c(rbinom(nbsim, n7[1], p) / n7[1],
                rbinom(nbsim, n7[2], p) / n7[2], rbinom(nbsim, n7[3], p) / n7[3],
                rbinom(nbsim, n7[4], p) / n7[4], rbinom(nbsim, n7[5], p) / n7[5],
                rbinom(nbsim, n7[6], p) / n7[6], rbinom(nbsim, n7[7], p) / n7[7])
  #----- Affichage des résultats et des graphiques-----
  graphics.off()
  plot(nechant, simfreqar,
       main = "Distributions* des fréquences* des succès")
  dev.new()
  plot(as.factor(nechant), simfreqar,
       xlab = "Échantillons par tailles",
       ylab = "Fréquence des succès",
       main = "Résumé des distributions*")
}
```

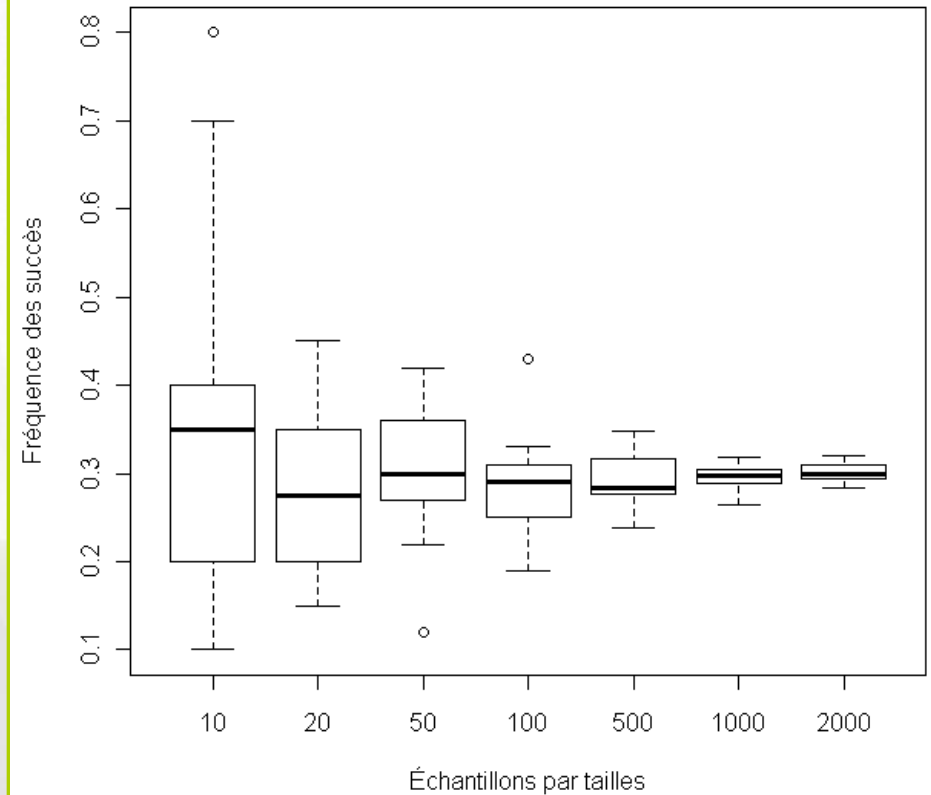
# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## LES DISTRIBUTIONS DANS LA LOI DES GRANDS NOMBRES

Distributions des fréquences des succès



Résumé des distributions



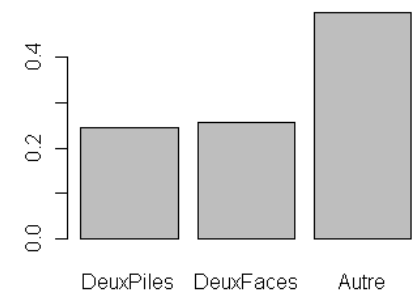
# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## LA SIMULATION : UNE AUTRE FAÇON DE FAIRE DES PROBABILITÉS

*Jet simultané de deux pièces identiques :  
quel univers et quelle distribution de probabilité ?*

```
pileface <- function(nbsim = 2000) {  
  resultats <- vector(length = 3)  
  names(resultats) <- c("DeuxPiles", "DeuxFaces", "Autre")  
  
  for(i in 1:nbsim) {  
    pieceA <- sample(c("Pile", "Face"), 1)  
    pieceB <- sample(c("Pile", "Face"), 1)  
    if(pieceA == "Pile" & pieceB == "Pile") {  
      resultats[1] <- resultats[1] + 1 } else {  
      if(pieceA == "Face" & pieceB == "Face") {  
        resultats[2] <- resultats[2] + 1 } else {  
          resultats[3] <- resultats[3] + 1 }  
      }  
    }  
  
    print(resultats)  
    print(resultats / nbsim)  
    barplot(resultats / nbsim)  
  }  
}
```

```
> pileface()  
(DISTRIBUTION*)  
DeuxPiles DeuxFaces Autre  
      490      514      996  
0.245      0.257      0.498
```



# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## LA SIMULATION : UNE AUTRE FAÇON DE FAIRE DES PROBABILITÉS

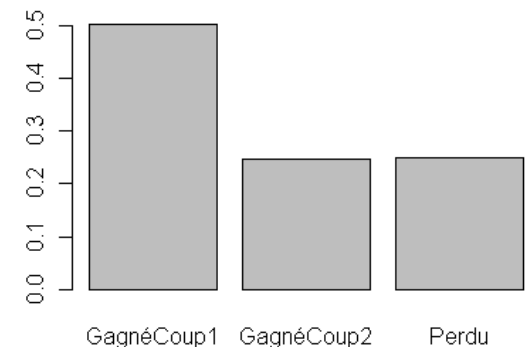
*Croix ou Pile de D'Alembert :*

*Le jeu, en deux coups, s'arrête dès que je gagne en "amenant croix".*

*Jet de une ou deux pièces (croix - pile) équilibrées,  
Distribution de probabilité du rang du coup gagnant :*

```
croixpile <- function(nbsim = 2000) {  
  resultats <- vector(length = 3)  
  names(resultats) <- c("GagnéCoup1", "GagnéCoup2", "Perdu")  
  for(i in 1:nbsim){  
    coup1 <- sample(c("Pile", "Croix"), 1)  
    if(coup1 == "Croix") {resultats[1] <- resultats[1] + 1} else {  
      coup2 <- sample(c("Pile", "Croix"), 1)  
      if(coup2 == "Croix") {resultats[2] <- resultats[2] + 1} else {  
        resultats[3] <- resultats[3] + 1  
      }  
    }  
  }  
  print(resultats)  
  print(resultats / nbsim)  
  barplot(resultats / nbsim)  
}
```

```
> croixpile(nbsim = 5000)  
(DISTRIBUTION*)  
GagnéCoup1 GagnéCoup2      Perdu  
      2512      1241      1247  
GagnéCoup1 GagnéCoup2      Perdu  
      0.5024      0.2482      0.2494
```



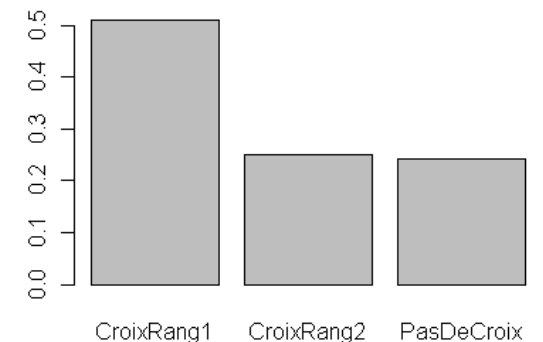
# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## LA SIMULATION : UNE AUTRE FAÇON DE FAIRE DES PROBABILITÉS

*On jette deux fois une pièce équilibrée (croix - pile),  
Distribution de probabilité du rang d'arrivée du premier croix :*

```
premiercroix <- function(nbsim = 2000) {  
  resultats <- vector(length = 3)  
  names(resultats) <- c("CroixRang1", "CroixRang2", "PasDeCroix")  
  for(i in 1:nbsim){  
    jet1 <- sample(c("Pile", "Croix"), 1)  
    jet2 <- sample(c("Pile", "Croix"), 1)  
    if(jet1 == "Croix") resultats[1] <- resultats[1] + 1  
    if(jet1 != "Croix" & jet2 == "Croix") resultats[2] <- resultats[2] + 1  
    if(jet1 != "Croix" & jet2 != "Croix") resultats[3] <- resultats[3] + 1  
  }  
  print(resultats)  
  print(resultats / nbsim)  
  barplot(resultats / nbsim)  
}
```

```
> premiercroix(nbsim = 5000)  
(DISTRIBUTION*)  
CroixRang1 CroixRang2 PasDeCroix  
      2547      1247      1206  
CroixRang1 CroixRang2 PasDeCroix  
      0.5094      0.2494      0.2412
```



# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

## LA SIMULATION : UNE AUTRE FAÇON DE FAIRE DES PROBABILITÉS

Un modèle d'urne virtuelle, la variable étudiée étant le nombre de boules rouges dans l'échantillon de  $n$  boules tirées.

Il s'agit de comparer des tirages sans remise et des tirages avec remise.

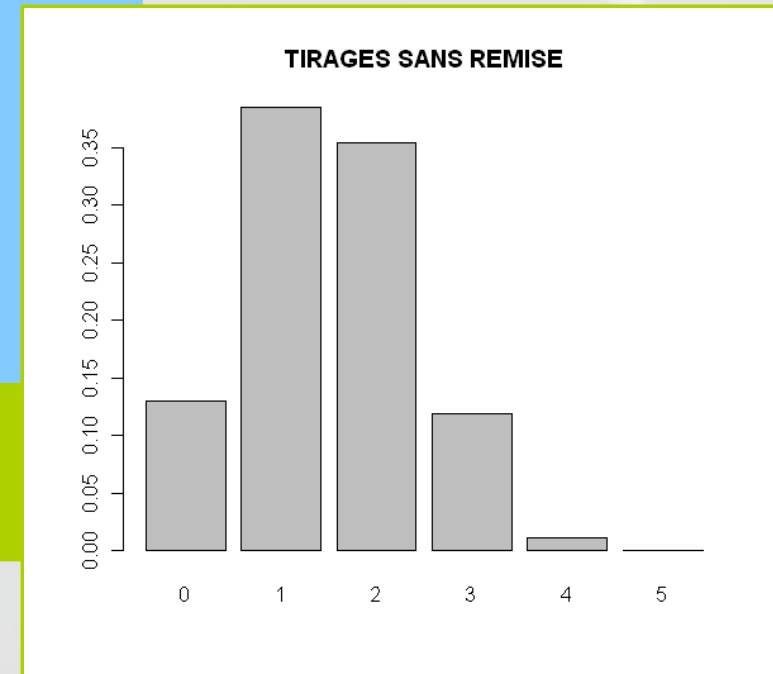
Simuler des tirages sans remise :

```
*** OUVRIR DEUX FENÊTRES GRAPHIQUES AVEC dev.new() dev.new()
sansremise2 <- fonction(R = 6, Q = 14, n = 5, nbsim = 10000){
  M <- R + Q
  urne <- rep(c("rouge", "autre"), c(R, Q))
  vecnbrouges <- rep(NA, nbsim)
  for(i in 1:nbsim){
    echant <- sample(urne, n, replace = F)
    nbrouges <- sum(echant == "rouge")
    vecnbrouges[i] <- nbrouges
  }
  distribnbrouges <- table(vecnbrouges) / nbsim

  cat("\nProportion de rouges dans l'urne :", R / M, "\n\n")
  print(distribnbrouges)

  dev.set(2)
  barplot(distribnbrouges, main = "TIRAGES SANS REMISE")
}
```

```
> sansremise2()
Proportion de rouges dans l'urne : 0.3
(DISTRIBUTION*)
vecnbrouges
      0      1      2      3      4      5
0.1298 0.3849 0.3544 0.1194 0.0113 0.0002
```



# QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

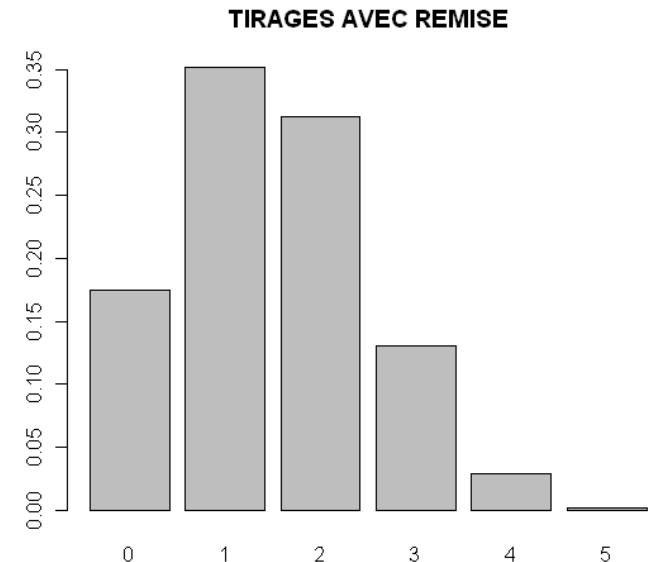
## LA SIMULATION : UNE AUTRE FAÇON DE FAIRE DES PROBABILITÉS

*Un modèle d'urne virtuelle, la variable étudiée étant le nombre de boules rouges dans l'échantillon de  $n$  boules tirées.  
Il s'agit de comparer des tirages sans remise et des tirages avec remise.*

*Simuler des tirage avec remise :*

```
avecremise3 <- function(R = 6, Q = 14, n = 5, nbsim = 10000){  
  M <- R + Q  
  urne <- rep(c("rouge", "autre"), c(R, Q))  
  vecnbrouges <- rep(NA, nbsim)  
  for(i in 1:nbsim){  
    echant <- sample(urne, n, replace = T)  
    nbrouges <- sum(echant == "rouge")  
    vecnbrouges[i] <- nbrouges  
  }  
  distribnbrouges <- table(vecnbrouges) / nbsim  
  cat("\nProportion de rouges dans l'urne :", R / M, "\n\n")  
  print(distribnbrouges)  
  dev.set(2)  
  barplot(distribnbrouges, main = "TIRAGES AVEC REMISE")  
}
```

```
> avecremise3()  
Proportion de rouges dans l'urne : 0.3  
(DISTRIBUTION*)  
vecnbrouges  
      0      1      2      3      4      5  
0.1746 0.3513 0.3120 0.1306 0.0293 0.0022
```





## QUELQUES RESSOURCES POUR LE NOUVEAU PROGRAMME DE TERMINALE S (RENTÉE 2012)

EN GUISE DE CONCLUSION :

AVEC R ON PEUT FAIRE DE LA SIMULATION UN VÉRITABLE OUTIL DE TRANSPOSITION DIDACTIQUE ET DE RÉOLUTION DE PROBLÈME.

DE PLUS L'ANALYSE EXPLORATOIRE DES SÉRIES SIMULÉES PERMET DE RÉINVESTIR LES STATISTIQUES DESCRIPTIVES ET DE MONTRER LA PLACE IMPORTANTE QU'ELLES OCCUPENT DANS L'ANALYSE DES DONNÉES AU SENS LARGE.

LES NOUVEAUX PROGRAMMES OFFRENT L'OPPORTUNITÉ D'INVESTIR CES OUTILS ET MÉTHODES MODERNES.