Université de Nantes
Master professionnel II : Ingénierie mathématique

Statistique Bayésienne et Simulation

Exercices

Anne Philippe

anne.philippe@math.univ-nantes.fr

Le logiciel utilisé pendant les séances de TP est le logiciel libre R.
Quelques adresses :

http://www.r-project.org/
http://www.math.sciences.univ-nantes.fr/~philippe/index.php?frame=R

7 décembre 2010

CHAPITRE 1

Pré requis

Exercice 1. Calcul de lois : la loi β

Soit $(a, b) \in \mathbb{R}_+ \times \mathbb{R}_+$.

Soit X et Y deux variables aléatoires *indépendantes*. On suppose que

– X est distribuée suivant la loi $\Gamma(a, 1)$

– Y est distribuée suivant la loi $\Gamma(b, 1)$

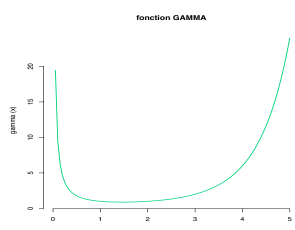
On rappelle que la loi $\Gamma(a, 1)$ admet pour densité

$$f(x) = \frac{1}{\Gamma(a)} e^{-x} x^{a-1}\, \mathbb{I}_{\mathbb{R}_+}(x).$$

La fonction Γ est définie sur $\mathbb{R}_+$ par

$$\Gamma(a) = \int_0^{\infty} e^{-t} t^{a-1}\, \mathrm{d}t = (a-1)\Gamma(a-1).$$

On a en particulier pour tout $a \in \mathbb{N}$, $\Gamma(a+1) = a!$



1) Écrire la densité de la loi du couple (X, Y)

2) Donner la loi du couple $(V, W) = (X + Y, \dfrac{X}{X+Y})$.

3) Les variables aléatoires V et W sont elle indépendantes ? Préciser les lois marginales de V et W.

4) En déduire une expression de $B(a, b) := \int_0^1 x^{a-1}(1-x)^{b-1}dx$.

Exercice 2. Convergence

Soit $(\alpha, \beta) \in \mathbb{R}_+^*$. On dispose deux échantillons indépendants $(X_1, \ldots, X_n)$ et $(Y_1, \ldots, Y_n)$ qui possèdent les propriétés suivantes

- $(X_1, \ldots, X_n)$ sont iid suivant la loi exponentielle de paramètre α
- $(Y_1, \ldots, Y_n)$ sont iid suivant la loi exponentielle de paramètre $\alpha\beta$.

La loi exponentielle de paramètre $\tau > 0$ a pour densité $f(x) = \tau e^{-\tau x}\, \mathbb{I}_{\mathbb{R}^+}(x)$ On souhaite estimer les paramètres α et β à partir des observations $(X_1, \ldots, X_n, Y_1, \ldots, Y_n)$.

1) Écrire la fonction de vraisemblance de l'échantillon $(X_1, \ldots, X_n, Y_1, \ldots, Y_n)$.

2) Calculer les estimateurs du maximum de vraisemblance des paramètres (α, β). On les note $(\hat{\alpha}, \hat{\beta}_n)$.

3) On pose

$$\bar{X}_n = \frac{1}{N} \sum_{i=1}^{n} X_i \qquad \text{et} \qquad \bar{Y}_n = \frac{1}{N} \sum_{i=1}^{n} Y_i$$

Montrer que le vecteur $Z_n := \begin{pmatrix} \bar{X}_n \\ \bar{Y}_n \end{pmatrix}$ converge presque sûrement vers $z = \begin{pmatrix} \alpha^{-1} \\ \alpha^{-1}\beta^{-1} \end{pmatrix}$.

4) En déduire que $(\hat{\alpha}_n, \hat{\beta}_n)$ est un estimateur convergent au sens de la convergence presque sure.

5) Montrer que $\sqrt{n}(Z_n - z)$ converge en loi vers une variable gaussienne. Préciser la moyenne et la variance de la limite.

6) En déduire que
$$\sqrt{n}\left(\begin{pmatrix} \hat{\alpha}_n \\ \hat{\beta}_n \end{pmatrix} - \begin{pmatrix} \alpha \\ \beta \end{pmatrix}\right)$$
converge en loi vers une variable gaussienne. Préciser la moyenne et la variance de la limite.

Exercice 3. Conditionnement pour des lois discrètes

Soit (X_1, X_2) un couple de variables aléatoires indépendantes et de même loi définie par
$$\mathbf{P}(X_1 = k) = \alpha^k(1-\alpha) \qquad \forall\, k \in \mathbb{N}$$
où α est fixé dans $]0, 1[$. (On note $\tilde{G}(\alpha)$ cette loi)

On pose
$$N = \begin{cases} 1 & \text{si } X_1 \leq X_2 \\ 2 & \text{sinon} \end{cases}$$

1) Déterminer le loi de X_1 sachant $N = 1$, puis l'espérance conditionnelle.

2) Déterminer le loi de X_1 sachant $N = 2$, puis l'espérance conditionnelle.

3) Exprimer l'espérance de X_1 en fonction des espérances conditionnelles calculées aux questions précédentes

4) Calculer la loi et l'espérance de N conditionnellement à X_1.

Exercice 4. Conditionnement dans le cas de lois continues

Les deux variables aléatoires X et Y sont positives, et la loi du vecteur (X, Y) est donnée par

– Y a une loi de densité
$$f(t) = \lambda^2 t\, e^{-\lambda t}\mathbb{I}_{\mathbb{R}^+}(t)$$
– La loi de X conditionnellement à Y est la loi uniforme sur $[0, Y]$.

1) Quelle est la loi du vecteur (X, Y) ?

2) Quelle est la loi de X ?

3) Montrer que la loi de Y conditionnellement à X a comme densité
$$\lambda e^{-\lambda(t-X)}\mathbb{I}_{[X,+\infty[}(t)$$

4) Calculer $\mathbb{E}(Y|X)$ et $\mathbb{E}Y$.

CHAPITRE 2

Statistique Bayésienne

Exercice 1

On veut estimer la moyenne d'un échantillon gaussien dont la variance est supposée connue et égale à $\sigma^2 = 9$. On sait a priori que la moyenne μ inconnue est autour de 3.

1) Montrer que si la loi a priori est une loi gaussienne de moyenne 3 et de variance τ alors la loi a posteriori est aussi gaussienne de moyenne $\dfrac{\tau n\bar{X}_n + 3\sigma^2}{n\tau + \sigma^2}$ et de variance $\dfrac{\tau\sigma^2}{\sigma^2 + n\tau}$.
2) Quel est l'estimateur du maximum de vraisemblance de μ ?

   On suppose que $\mu = 3$. Simuler un échantillon de taille $N = 200$.

R 3) Sur les données simulées : On fixe $\tau = 1$. Superposer la densité de la loi a priori et les densités a posteriori pour différentes tailles d'échantillon n (par exemple n varie de 20 à 200 par pas de 20).

R 4) Reprendre la question précédente pour différentes valeurs de $\tau = 1/10, 1/2, 100$.

5) Interpréter les résultats obtenus aux questions 3 et 4.

R 6) Sur les données simulées, représenter l'évolution de l'estimateur du maximum de vraisemblance en fonction de la taille de l'échantillon $n = 1...N$. Superposer l'évolution de la moyenne de la loi a posteriori pour différentes valeurs de $\tau = 1/10, 1/2, 1, 100$. Commenter les résultats.

R 7) On suppose maintenant que $\mu = 4$. Reprendre la question 6 avec le même a priori.

8) Conclure.

Exercice 2

Trois personnes veulent estimer la proportion p d'étudiants qui ne résident pas sur le campus.

– A suppose que la loi a priori sur p est la loi discrète définie par

p_i	.1	.2	.3	.4	.5
$\pi(p_i)$	.5	.2	.2	.05	.05

– B suppose que la loi a priori sur p est la loi Beta de paramètres $(3, 12)$

– C suppose que la loi a priori sur p est Beta de paramètres $(1, 4)$

R 1) Calculer la moyenne et l'écart type de ces trois lois a priori.

2) Ont-ils la même information a priori sur la localisation du paramètre p ? Accordent ils la même confiance à l'information a priori obtenue ?

3) On veut prédire y le nombre d'étudiants qui habitent hors du campus dans un échantillon de taille 12.

   Donner l'expression de la loi prédictive a priori $m(y)$ dans les trois situations.

R 4) Superposer les courbes représentatives des densités des lois prédictives a priori.

i	b_i	$P(p = b_i)$
1	0.05	0.03
2	0.15	0.18
3	0.25	0.28
4	0.35	0.25
5	0.45	0.16
6	0.55	0.07
7	0.65	0.02
8	0.75	0.00
9	0.85	0.00
10	0.95	0.00

TABLE 1

Exercice 3

On veut estimer p la proportion des étudiants qui dorment plus de 8 heures par nuit. Les observations sur un échantillon de 27 étudiants sont :

11 étudiants dorment plus de 8 heures
16 étudiants dorment moins de 8 heures.

On envisage trois lois a priori différentes sur p :

A- La loi discrète définie dans la TABLE 1

B- Un mélange de loi uniforme (loi a priori de type histogramme) défini par

i	0	1	2	3	4	5	6	7	8	9
q_i	2	4	8	8	4	2	1	1	1	1

où

$$a_i = i/10$$

et

$$P(p =\in [a_i, a_{i+1}]) = \frac{q_i}{\sum_{j=0}^{9} q_j}$$

C- La loi Beta de paramètres $a = 3.4$ and $b = 7.4$.

R 1) Représenter graphiquement les trois lois a priori, calculer la moyenne et la variance de ces lois a priori

2) Pour les trois lois a priori proposées, calculer la loi a posteriori, sa moyenne et sa variance.

R 3) Représenter graphiquement les trois lois a posteriori.

R 4) Commenter les résultats obtenus

Exercice 4. Régions de confiance bayésienne

On dispose de n observations $X_1, ..., X_n$ iid suivant une loi gaussienne $\mathcal{N}(\theta, 1)$.
On choisit comme loi a priori sur θ la loi gaussienne $\mathcal{N}(0, \tau^{-2})$, $\tau > 0$

1) Montrer que la loi a posteriori est une loi Gaussienne

$$\mathcal{N}(\frac{\bar{X}_n}{1+\tau^2/n}, \frac{1}{n+\tau^2})$$

2) Montrer que les régions HPD de niveau $1-\alpha(=.95)$ sont de la forme

$$\theta \in \left[\frac{\bar{X}_n}{1+\tau^2/n} - \frac{u_{1-\alpha/2}}{\sqrt{n+\tau^2}}; \frac{\bar{X}_n}{1+\tau^2/n} + \frac{u_{1-\alpha/2}}{\sqrt{n+\tau^2}}\right] = I^{HPD}(\tau, \bar{X}_n)$$

où u_α est le quartile d'ordre α de la loi gaussienne standard.

R 3) Tracer les bornes des régions de confiance en fonction de τ. On prend par exemple $\theta = 0$.

R 4) Ajouter sur ce graphique les bornes de la région de confiance classique de niveau $1-\alpha$.

5) Montrer que

$$P_\theta(\theta \in I^{HPD}(\tau, \bar{X}_n)) = F\left(\frac{\theta\tau^2}{\sqrt{n}} + u_{1-\alpha/2}\sqrt{\frac{n+\tau^2}{n}}\right) - F\left(\frac{\theta\tau^2}{\sqrt{n}} - u_{1-\alpha/2}\sqrt{\frac{n+\tau^2}{n}}\right)$$

où F est la fonction de répartition de la loi gaussienne standard.

R 6) Tracer cette probabilité en fonction de τ et θ. Commenter.

7) Comment choisir la loi a priori pour que les régions HPD de niveau $1-\alpha$ soient aussi des régions de confiance au sens classique de niveau $1-\alpha$. Vérifier que la loi de Jeffrey satisfait cette propriété.

Exercice 5. Calcul des lois a priori conjuguées

Donner une famille de lois conjuguées pour les modèles $\{P_\theta, \theta \in \Theta\}$ suivants

1) Poisson : $\{\mathcal{P}(\theta), \theta \in \mathbb{R}_+^*\}$
Indication : loi gamma
2) Gaussien de variance connue : $\{\mathcal{N}(\theta, \sigma^2), \theta \in \mathbb{R}\}$
Indication : loi gaussienne
3) Gaussien de moyenne connue : $\{\mathcal{N}(\mu, \theta), \theta \in \mathbb{R}_+^*\}$
Indication : loi inverse gamma.
On dit que X suit une loi inverse Gamma de paramètres $(a,b) \in \mathbb{R}_+^{*2}$, si X^{-1} suit une loi Gamma de paramètres (a,b). Montrer que la loi inverse Gamma de paramètres (a,b) admet pour densité
$$b^a \frac{1}{\Gamma(a)} e^{-b/x} x^{-a-1}\, \mathbb{I}_{\mathbb{R}_+^*}(x)$$
4) Gaussien : $\{\mathcal{N}(\theta_1, \theta_2),\ \theta_1 \in \mathbb{R},\ \theta_2 \in \mathbb{R}_+^*\}$
Indication : Considérer la famille des lois de paramètres $(\lambda, \tau, a, b) \in \mathbb{R} \times \mathbb{R}_+^{*3}$ définies par
   – la loi conditionnelle de θ_1 sachant θ_2 est la loi gaussienne de moyenne λ et de variance θ_2/τ
   – la loi de θ_2 est la loi inverse Gamma de paramètres (a,b).

Exercice 6. Calcul des lois de Jeffrey

Calculer la loi non informative de Jeffrey pour les modèles suivants :

1) Poisson : $\{\mathcal{P}(\theta), \theta \in \mathbb{R}^+\}$
2) Gaussien : $\{\mathcal{N}(\mu, \sigma^2), (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+\}$ dans les deux situations suivantes
   a. $\theta = \mu$ et σ^2 est connu.
   b. $\theta = \sigma$ et μ est connu.
   c. les deux paramètres sont inconnus : $\theta = (\mu, \sigma)$.

Exercice 7

R 1) Récupérer le fichier de données de taille $n = 500$
`http://www.math.sciences.univ-nantes.fr/~philippe/lecture/Piecesdef.txt`.
La i[e] observation N_i est égale au nombre de pièces défectueuses dans le i[e] lot. Un lot est constitué de 50 pièces. On veut estimer la probabilité de produire une pièce défectueuse.

2) Décrire le modèle statistique.

3) Montrer que la famille des lois β est une famille de lois conjuguées pour ce modèle.

4) Le service qualité fournit les informations suivantes
   – la proportion de pièces défectueuses est en moyenne 0.15
   – la proportion de pièces $X_1, ..., X_n$ défectueuses appartient à l intervalle $[0.1, 0.2]$ avec une probabilité de 95%

Utiliser ces informations pour fixer les paramètres de la loi a priori conjuguée.

5) Calculer la loi a posteriori et l'estimateur de Bayes associé au coût quadratique.
6) **Situation non informative :** Calculer la loi a priori de Jeffrey pour ce modèle.
7) Calculer la loi a posteriori et l'estimateur de Bayes.

R 8) Représenter les deux estimateurs de Bayes en fonction du nombre d'observations n.

R 9) Tracer sur un même graphique les deux lois a posteriori calculées sur les K premières observations. Faire varier K par exemple $5, 10, 15, 20$.

CHAPITRE 3

Simulation et estimateurs de Monte Carlo

Exercice 1. Contrôle et visualisation de la convergence des estimateurs de Monte Carlo

Soit X une variable aléatoire distribuée suivant la loi gaussienne standard, on souhaite estimer l'intégrale suivante

$$I = \mathbb{E}(e^X \mathbb{I}_{[-1,1]}(X))$$

par la technique standard de Monte Carlo.

1) Décrire la mise en oeuvre de l'estimateur de Monte Carlo, I_n, construit à partir d'un échantillon simulé suivant la loi de X

R 2) Représenter l'estimateur de Monte Carlo en fonction de la taille de l'échantillon.

R 3) Sur la trajectoire simulée à la question précédente, évaluer la variance de l'estimateur de Monte Carlo.

R 4) En déduire les bornes de l'intervalle de confiance de niveau $1-\alpha = 95\%$ construit en utilisant le Théorème Limite Central. Superposer l'estimateur et ses régions de confiance.

On estime maintenant la variance et les régions de confiance par une méthode non asymptotique. La loi de l'estimateur de Monte Carlo I_n est estimée à n fixé par une méthode de simulation.

R 5) La variance, les quantiles, etc de la loi de I_n sont estimés à partir de N échantillons indépendants de taille n.

La démarche

- On construit une matrice n lignes et N colonnes de nombres aléatoires distribués suivant la loi de densité f. On note A cette matrice
- On calcule la suite des estimateurs $\{I_k,\ k = 1, \dots, n\}$ sur les N colonnes

  h = function(x) .... à définir ....
  cummean = function(x) cumsum(h(x))/(1 :length(x))
  B = apply(A,2,cummean)

  La k ème ligne de la matrice B contient un échantillon suivant la loi de l'estimateur de Monte Carlo I_k.

Évaluation de la variance de I_n

- Calculer la variance de chacune des lignes

  V = apply(B,1,var)

  pour obtenir une estimation de la variance de I_k, $k = 1 \dots, n$
- Représenter l'estimation de la variance en fonction de la taille de l'échantillon
- Représenter le nuages de points $\{(\log(k), log(V[k]),\ k = 1 \dots, n\}$. Commenter le résultat et faire le lien avec le résultat $\mathrm{Var}(I_n) = Cn^{-1}$.
- En déduire une estimation de C.

Régions de confiance

– Estimer les quantiles d'ordre $\alpha/2$ et $1-\alpha/2$ de la loi de I_n. Utiliser la fonction `quantile` et `apply` sur les lignes de la matrice B.
– Représenter sur un même graphique l'estimateur de Monte Carlo et sa région de confiance en fonction de la taille de l'échantillon

Exercice 2. Algorithme d'acceptation-rejet pour les lois gaussiennes tronquées

1) Construire un algorithme d'acceptation rejet pour simuler suivant la loi gaussienne tronquée sur $[2, +\infty[$, c'est à dire la loi dont la densité est proportionnelle à

$$h(x) = e^{-x^2/2}\,\mathbb{I}_{[2,+\infty[}(x)$$

en utilisant comme loi instrumentale
   – La loi normale standard
   – la loi de Student à 1 degré de liberté aussi appelée loi de Cauchy.

R 2) Programmer et tester ces deux algorithmes.

R 3) Comparer les taux d'acceptation.

4) Proposer un algorithme pour simuler un échantillon suivant la loi de Cauchy tronquée sur $[2, +\infty[$.
Indication : calculer la fonction de répartition.

5) Construire un algorithme d'acceptation rejet pour simuler suivant la loi gaussienne tronquée sur $[2, +\infty[$ en utilisant comme loi instrumentale la loi de Cauchy tronquée.

R 6) Comparer le taux d'acceptation de cet algorithme aux deux précédents.

Exercice 3. Algorithme SIR : choix de m en fonction de n

On cherche à simuler par la méthode SIR un échantillon de taille n suivant la loi $\Gamma(3.5, 1)$ à partir d'un échantillon de taille m suivant le loi $\Gamma(3, 1)$.

R 1. Programmer cet algorithme.
Indication : utiliser la fonction `sample` pour l'étape de reéchantillonnage.

On cherche à calibrer m en fonction de n.

R 2. Pour évaluer les performances de cet algorithme, on teste l'ajustement à la loi $\Gamma(3.5, 1)$, par le test de Komogorov de niveau 5% pour $N = 500$ échantillons simulés de façon indépendante.
Comparer la proportion d'échantillons rejetés avec l'erreur de première espèce 5% dans les situations suivantes

m	500	5000	10 000

et

n/m	0.01	0.05	0.10	0.15	0.20	0.50

Exercice 4. Importance Sampling / Échantillonnage Pondéré

On cherche à estimer l'intégrale

$$I = \int h(x) f(x)\,\mathrm{d}x$$

où $h(x) = x^{10}\,\mathbb{I}_{\mathbb{R}^+}(x)$ et f est la densité de la loi gaussienne standard.

1) Que peut on dire de la variance de l'estimateur de Monte Carlo

R 2) Représenter l'estimateur de Monte Carlo en fonction de n, ajouter la région de confiance construite à partir de N trajectoires indépendantes.

On cherche à améliorer l'estimateur de Monte Carlo à l'aide de la méthode d'échantillonnage pondéré. On utilise comme loi instrumentale, une loi gamma de paramètre $(11, b)$.

3) Que peut on dire de la variance de l'estimateur $\delta_n(b)$?

R 4) Représenter l'estimateur $\delta_n(b)$ en fonction de n, ajouter les deux régions de confiance.

5) Montrer que le rapport $|h|f/g$ est borné.

6) En déduire un algorithme pour simuler des variables aléatoires iid suivant la loi de la densité proportionnelle à $|h|f$. On note g^* cette densité

R 7) Programmer l'algorithme précédent.

8) Construire un estimateur de Monte Carlo $\tilde{\delta}_n(g^*)$ de I à partir d'un échantillon simulé suivant g^*.

R 9) Représenter l'estimateur $\tilde{\delta}_n(g^*)$ en fonction de n, ajouter les deux régions de confiance. Commenter les résultats obtenus.

Exercice 5. Approximation des estimateurs de Bayes par les méthodes de Monte Carlo

Soit $X_1, \ldots, X_n$ iid suivant une loi gaussienne $\mathcal{N}(\theta, 1)$. Le choix de la loi a priori sur θ est la loi de Cauchy

$$\pi(\theta) = \frac{1}{\pi} \frac{1}{1+\theta^2}$$

R 1) Tester l'hypothèse : "la distribution des observations est gaussienne" sur les données suivantes

`http://www.math.sciences.univ-nantes.fr/~philippe/lecture/data-gauss`

puis tester l'hypothèse

H_0 *La variance est égale à* 1

2) Calculer la loi a posteriori.

3) Peut on calculer explicitement l'estimateur de Bayes de θ ?

4) Proposer et implémenter un algorithme algorithme d'acceptation rejet pour simuler suivant la loi a posteriori. (Indication : prendre la loi gaussienne de moyenne $\bar{X}_n$ et de variance $1/n$ comme loi instrumentale)

R 5) À partir d'échantillons simulés via l'algorithme d'acceptation rejet
   - Construire une approximation de la densité de la loi a posteriori
   - Donner une approximation de l'estimateur de Bayes. Estimer la précision de votre approximation.
   - Donner une approximation de la région de confiance bayésienne de niveau 95%.

R 6) En utilisant la même loi instrumentale que pour l'algorithme d'acceptation rejet, mettre en oeuvre un algorithme SIR pour générer des nombres aléatoires suivant la loi a posteriori

R 7) À partir d'échantillons simulés via l'algorithme SIR
   - Construire une approximation de la densité de la loi a posteriori
   - Donner une approximation de l'estimateur de Bayes. Estimer la précision de votre approximation.
   - Donner une approximation de la région de confiance bayésienne de niveau 95%.

R 8) A partir de l'échantillon simuler suivant la loi instrumentale, calculer une approximation de l'estimateur de Bayes par la méthode d'échantillonage pondéré.

9) Comparer les résultats obtenus et conclure.

CHAPITRE 4

Bootstrap

Exercice 1. Modèle exponentiel

On dispose de 10 données simulées suivant une loi exponentielle

$$f(x) = \frac{1}{\theta} e^{-x/\theta}\, \mathbb{I}_{\mathbb{R}^+}(x)$$

de paramètre $\theta = 50$

104.91 7.31 4.54 26.24 62.47 116.37 9.97 36.64 109.67 95.96

On note

– F la fonction de répartition

– F_n la fonction de répartition empirique.

On estime la moyenne $\theta = \theta(F) = \int x \,\mathrm{d}F(x)$ par $T(\underline{X}) = \hat{\theta}_n = \bar{X}_n$.

1) Justifier que la loi de $\hat{\theta}_n$ est la loi gamma de paramètres $(n, \frac{n}{\theta})$

R 2) Comparer la distribution théorique de $\hat{\theta}_n - \theta$ avec les approximations obtenues par
- le théorème limite central,
- le bootstrap non paramètrique. Discuter du choix de B le nombre d'échantillon simulés suivant F_n pour approcher la loi de $T(\underline{X}^*) - \hat{\theta}_n$ par une méthode de Monte Carlo,
- le bootstrap paramètrique.

R 3) Comparer la valeur théorique de la variance de l'estimateur de θ (qui est égale à $\theta^2 n^{-1}$, à justifier) avec l'estimation obtenue par
- le bootstrap non paramètrique,
- le bootstrap paramètrique.

R 4) Construire et comparer les intervalles de confiance obtenus par les méthodes suivantes
- la loi exacte de $\hat{\theta}_n$,
- le théorème limite central,
- le bootstrap non paramètrique,
- le bootstrap paramètrique.

```
#----------------------------------------------------#
# utilisation de la fonction boot library(boot)
#----------------------------------------------------#

#----------------------------------------------#
# le bootstrap parametrique
#----------------------------------------------#


  boot.gen <- function(data, moy )
     {     rexp(length(data), 1/moy)
      }
```

```
  boot.stat=function(x) mean(x)

# xdata contient les données
  resultat <- boot(xdata, boot.stat, R=B, sim="parametric",
          ran.gen=boot.gen, mle=mean(xdata))


#----------------------------------------------#
# le bootstrap non parametrique
#----------------------------------------------#
 boot.stat=function(x,n) mean(x[n])
 nonpara= boot(xdata,statistic=boot.stat,R=B)
```