

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
SCIENTIFIQUE
UNIVERSITÉ ABDERRAHMANE MIRA DE BÉJAIA
FACULTÉ DES SCIENCES EXACTES

Département de Mathématiques

Mémoire présenté pour l'obtention du diplôme de Master en mathématiques

Spécialité : Statistique et Analyse Décisionnelle

Par

KHELADI Rebiha

THÈME

Sur le bootstrap des séries chronologiques

Soutenue publiquement, le 22/06/2015 devant le jury composé de :

Mme.	LAGHA Karima	M.C.B	Université A. Mira de Béjaia	Présidente
Mme.	BARECHE Aicha	M.C.A	Université A. Mira de Béjaia	Promotrice
M.	BOURAINÉ Mohand	M.A.A	Université A. Mira de Béjaia	Examineur

Remerciements

Je voudrai exprimer tous mes remerciements à Madame A. Bareche, Maître de conférence classe A, à l'université Abderrahmane Mira pour avoir voulu encadrer et rapporter ce mémoire, je la remercie pour la confiance qu'elle m'a accordée.

Je tiens à exprimer ma profonde gratitude et mes remerciements les plus vifs aux membres de jury madame K. Lagha et monsieur M. Bouraine pour m'avoir fait l'honneur de faire partie de la commission d'examineurs.

J'adresse mes sincères remerciements à tous les professeurs, intervenants et toutes les personnes qui par leurs paroles, leurs écrits, leurs conseils et leurs critiques ont guidé mes réflexions et ont accepté à me rencontrer et répondre à mes questions durant mes recherches.

Je remercie mes très chers parents, Sadi et Hassina sans oublier ma grand mère Taklit que j'adore qui ont toujours été là pour moi, « Vous avez tout sacrifié pour vos enfants n'épargnant ni santé ni efforts. Vous m'avez donné un magnifique modèle de labeur et de persévérance. Je suis redevable d'une éducation dont je suis fier ». ».

Je remercie mes frères Omar, Khaled, A. Errahmane, Nassim et mes très chères cousines Millissa et Nissrine qui sont comme mes sœurs pour leur encouragement.

Je tiens à remercier très spécialement mon très cher Salim qui est toujours été là pour moi.

Enfin, je remercie tous mes cousin(e)s, mes oncles, mes tantes et mes ami(e)s que j'aime. Pour leur confiance et sincère amitié, et à qui je dois ma reconnaissance et mon attachement.

Table des matières

Liste des tableaux	1
Liste des figures	2
Introduction Générale	3
1 Le principe classique d'analyse des séries chronologiques	6
1.1 Généralités	7
1.2 Modélisation de la partie aléatoire	10
1.2.1 Processus stationnaire	10
1.2.2 Bruit blanc fort	11
1.2.3 Bruit blanc faible	11
1.2.4 Autocorrélations:	11
1.2.5 Opérateur retard \mathbf{B}	12
1.2.6 Opérateur différence Δ	13
1.2.7 Propriétés	13
1.3 Différents modèles de séries temporelles	14
1.3.1 Modèles de Box-Jenkins	14
1.3.2 Estimation pour les modèles AR(p) par la méthode des moindres carrés .	15
1.3.3 Modèles linéaires non-stationnaires	17
1.4 Identification et estimation des paramètres	18
1.4.1 Identification du modèle: méthode de Box & Jenkins	18
1.4.2 Estimation des paramètres : méthode des m.c.o.	19

1.5	Détermination du type et de l'ordre du modèle	20
1.6	Estimation des paramètres et prévision	21
1.6.1	Méthode des moments	21
1.6.2	Maximum de vraisemblance	21
2	Une méthode de ré-échantillonnage : Le bootstrap	22
2.1	Le principe de <i>bootstrap</i>	23
2.2	Méthodes de ré-échantillonnage	25
2.2.1	Bootstrap des individus	25
2.3	Conditions de convergence du Bootstrap	28
2.4	Règles empiriques	28
2.5	Avantages et inconvénients	29
3	Bootstrap des séries temporelles	31
3.1	Bootstrap des résidus	32
3.2	Bootstrap par blocs	33
3.3	Séries temporelles autorégressives : bootstrap des résidus	35
3.3.1	Séries temporelles AR(p) stationnaires	35
3.3.2	Estimation bootstrap des résidus de la loi d'une racine	36
3.4	Séries temporelles à données q dépendantes: Block Bootstraps	36
3.4.1	Estimation BB de la loi d'une racine	36
4	Exemples d'application du bootstrap aux séries chronologiques	38
4.1	Estimation du paramètre β et son $I.C$	38
4.1.1	Analyse préliminaire	39
4.2	Exemple 2 (théorique) Un test Bootstrap dans un modèle AR(1)	42
4.3	Distributions asymptotique, Bootstrap et à distance finie	44
4.4	Validité asymptotique du bootstrap	45
4.5	Une étude par simulation	47
4.5.1	Préliminaire	47
4.5.2	Simulation	47

4.5.3	Etude des distributions Bootstrap	49
	Conclusion Générale	53
	Annexe	56
	Bibliographie	56

Liste des tableaux

2.3	Evolution de la concentration c_t d'une hormone au cours du temps t	39
2.4	La valeur du paramètre β ainsi que son I.C estimés par Box et Jenkins.....	40
2.5	Les valeurs du paramètres β ainsi que leurs I.C estimées par Bootstrap des résidus.....	41
4.5	Différentes régions d'acceptation du test bootstrap, ainsi que leur longueur $l_{T,0_B}^*(\alpha_1)$	48
4.6	Approximation du biais par (White[45]) et par Bootstrap $B = 3000$	50

Liste des figures

1.1	Indice mensuel des prix à la consommation I_t	08
1.2	Trafic voyageur de la SNCF en 2ième classe.....	08
1.3	Accroissement relatif mensuel de l'indice des prix.....	09
1.4	Evolution à moyen terme de l'accroissement relatif mensuel de l'indice des prix.....	09
2.1	Erreur standard de la moyenne m_2	29
4.1	Représentation graphique de la figure originale de la concentration c_t	39
4.2	Les autocorrélogrammes de la série c_t	40
4.3	Graphiques de comparaison des densités exactes, Bootstrap et asymptotiques, $\beta_0 = 1...$	50
4.4	Graphique de comparaison des densités exactes, Bootstrap et asymptotiques, $\beta_0 = 0.95...$	51
4.5	Graphique de comparaison des densités exactes, Bootstrap et asymptotiques, $\beta_0 = 1.01...$	51
4.6	Graphique de comparaison des densités exactes, Bootstrap et asymptotiques, $\beta_0 = 1.1.....$	52

Introduction Générale

Le monde tel qu'on le connaît est rempli de phénomènes dépendant du temps. En statistique, l'une des méthodes fréquemment utilisée est celle des séries chronologiques.

Les séries chronologiques ont commencé à être utilisées dans l'étude des températures et de la pluviométrie. L'étude descriptive de ces séries n'était pas suffisante pour prévoir leurs comportements futurs. Le besoin de prévision de ce type de séries a emmené les scientifiques à construire des modèles mathématiques aptes à répondre à ce besoin [23]. La majeure partie de la méthodologie statistique classique repose sur des séries d'observations indépendantes. Généralement, ce manque d'indépendance est considéré comme un handicap. Ainsi, un des objectifs d'une bonne expérimentation est d'éliminer cette dépendance. Toutefois, dans le cas de l'analyse des séries chronologiques nous nous occupons de données qui se développent dans le temps et telle que chaque observation peut dépendre dans une certaine mesure des observations précédentes. C'est sur cette dépendance que se focalise l'étude des séries chronologiques. Par ailleurs, on peut dire qu'une série chronologique possède une mémoire du passé dans le sens où les valeurs présentes et passées de la série reflètent dans une certaine mesure les valeurs futures [16].

Les séries chronologiques qu'on aura à étudier seront considérées comme des réalisations possibles d'un processus stochastique. Il est nécessaire dans ce qui va suivre de distinguer entre le processus stochastique et la série chronologique observée. Une série chronologique $X_{t_1}, X_{t_2}, \dots, X_{t_N}$ de N observations successives est considérée comme la réalisation d'un échantillon issu de la population infinie de telles possibilités de réalisation du processus stochastique.

En général, Box et Jenkins [8] ont proposé une méthode beaucoup plus sophistiquée se concentrant principalement sur la composante aléatoire. L'approche de Box-Jenkins a longtemps été considérée comme l'état de l'art dans la série chronologique. Essentiellement, les modèles *ARIMA* sont des généralisations des modèles linéaires ordinaires de régression, sauf que les "variables explicatives" se composent des observations antérieures de la série chronologique elle-même. Si les données de séries chronologiques s'ajustent effectivement à un modèle *ARIMA* de forme connue, alors les meilleures prévisions possibles pourraient être obtenues par des méthodes de projection classiques. Bien sûr, dans la pratique, même si les données ont été générées par un modèle *ARIMA*, nous ne connaissons pas le nombre de coefficients dans ce modèle ou les valeurs des coefficients. Cela conduit à des problèmes de sélection de modèle et d'estimation de paramètres.

Lorsque les méthodes paramétriques ne permettent pas d'obtenir des réponses correctes à tous ces problèmes ou lorsque les conditions d'applications ne sont pas remplies, une attitude consiste à abandonner ces méthodes paramétriques et de les remplacer par des méthodes non paramétriques, pour lesquelles les conditions d'application sont bien moins restrictives comme les méthodes de ré-échantillonnages. Le bootstrap fait partie de ces méthodes.

Le bootstrap est une méthode non paramétrique pour estimer la distribution d'un estimateur en échantillonnant ses données ou un modèle construit à partir des données. Elle nécessite, en effet, peu d'hypothèses et est relativement facile à programmer ce ne sont, en effet, que des tirages aléatoires.

Les méthodes disponibles pour mettre en œuvre le bootstrap et la précision des estimateurs de bootstrap dépendent du fait que les données proviennent d'un échantillon aléatoire indépendant [18] ou d'une série temporelle [2], [5] et [28].

L'objectif de ce mémoire consistera à expliciter les différentes possibilités d'application du bootstrap aux séries temporelles.

Quatre chapitres constituent l'essentiel de notre travail. Dans un premier temps, nous introduisons le principe classique d'analyse des séries chronologiques, ainsi que certaines définitions fondamentales. De plus nous rappelons les différents modèles de séries chronologiques et certaines méthodes d'estimation des paramètres et prévision. Le deuxième chapitre portera sur

la présentation de la méthode du bootstrap, en donnant son principe et son application aux problèmes d'inférence statistique telle que l'estimation de l'erreur standard de la moyenne et de la médiane pour mieux comprendre son intérêt. Ensuite nous terminons le chapitre par quelques avantages et inconvénients de la méthode.

Dans le troisième chapitre, nous nous intéressons aux méthodes du bootstrap pour les séries chronologiques, nous citons le bootstrap des résidus et le bootstrap par blocs appliquées pour les séries respectivement indépendantes et dépendantes.

La dernière partie est dédiée à l'application de la méthodes bootstrap des résidus en donnant deux exemples, le premier est numérique qui sert à estimer le paramètre β d'un $AR(1)$ et son intervalle de confiance par la méthode paramétrique Box et Jenkins et celle non paramétrique Bootstrap des résidus sous le logiciel R et le deuxième est théorique qui est fait pour un test de racine unité. L'objet de ce chapitre est de présenter et démontrer la validité analytique et asymptotique d'une procédure de type bootstrap permettant d'estimer et de tester la valeur du paramètre β de régression dans un modèle autorégressif d'ordre 1 ($AR(1)$).

Le principe classique d'analyse des séries chronologiques

Introduction

L'étude des séries temporelles, ou séries chronologiques, correspond à l'analyse statistique d'observations régulièrement espacées dans le temps. Elles ont été utilisées en astronomie 1906, en météorologie 1968, en biologie 1960 et en économie 1971... etc.

Les différentes techniques pour l'étude des séries chronologiques ont été initiées par Yule [48]. Les travaux de Wold [47] ont permis de développer une théorie complète des modèles autorégressifs moyenne mobile (*ARMA*). Dans les années quarante Wiener [46] et Kolmogoroff [26] ont utilisé les séries chronologiques pour résoudre le problème de l'estimation des filtres continus et discrets respectivement dans le domaine de la physique. Au début des années soixante Kalman [25] et Kalman et Bucy [24] ont étendu les procédures d'estimation de Wiener [46] et Kolmogoroff [26] aux séries chronologiques non stationnaires en utilisant la représentation état-espace.

L'objectif de l'approche de l'étude des séries chronologiques est de séparer la composante aléatoire de la partie déterministe du processus étudié. Cette étude se compose de deux étapes. La première est appelée "analyse des séries chronologiques" dont l'objectif est de mettre en évidence les caractéristiques du phénomène engendrant la série chronologique. Ceci est réalisé en utilisant simultanément les propriétés des autocorrélations et les propriétés spectrales de la série chronologique. La deuxième étape consiste en l'utilisation de plusieurs types de méthodes telles que la décomposition de la série chronologique, lissage exponentiel ou les modèles autorégressifs-moyennes mobiles. Toutes ces méthodes tendent à isoler le bruit blanc dans les valeurs observées de la série chronologique.

Le principe de décomposition des séries chronologiques est basé sur l'hypothèse que toute série chronologique est constituée de trois composantes principales, la tendance (T), la saisonnalité (S) et la partie aléatoire (Z_t).

Soit X_t le processus observé, si on suppose que la relation entre les trois composantes est additive, on aura alors:

$$X_t = T + S + Z_t. \quad (1.0.1)$$

La partie aléatoire est bien évidemment la partie la plus intéressante à modéliser pour une série chronologique, on note Z_t ou ε_t la composante aléatoire, supposée de moyenne nulle, mais possédant en général une structure de corrélation non nulle.

Dans ce premier chapitre, on présentera des rappels et des définitions pour nous aider à mieux comprendre l'étude descriptive qui sera exposée dans les chapitres et les sections qui suivent.

1.1 Généralités

Définition 1.1.1 *Une série chronologique est une série de mesures à valeurs réelles effectuées à des temps écartés régulièrement.*

L'intervalle de temps entre deux mesures successives dépend de la série, il peut s'agir d'un jour, d'une semaine, d'une minute... Par extension, il peut s'agir d'intervalles légèrement irréguliers (mois, trimestres, semestres, années, jours ouvrables...). Les dates sont numérotées

par des entiers positifs $n = 1, 2, \dots$. Les données sont de nature très diverse ; il peut s'agir de relevés de phénomènes d'origine naturelle (température à Paris, débit moyen journalier de la Garonne à Toulouse, activité des tâches solaires), de séries économiques (indice de la consommation, taux de chômage, prix de matières premières) ou financières (cours boursiers, taux de change).

Voir les quatre exemples via les figures suivantes qui sont tirées du livre [30]:

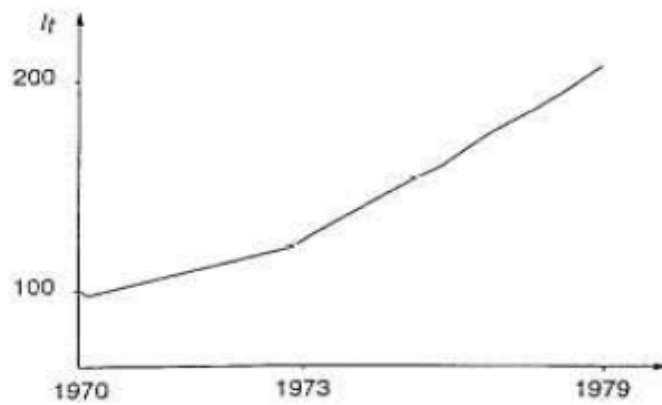


Fig -1.1 : Indice mensuel des prix à la consommation I_t .



Fig-1.2 : Trafic voyageur de la SNCF en 2ième classe.

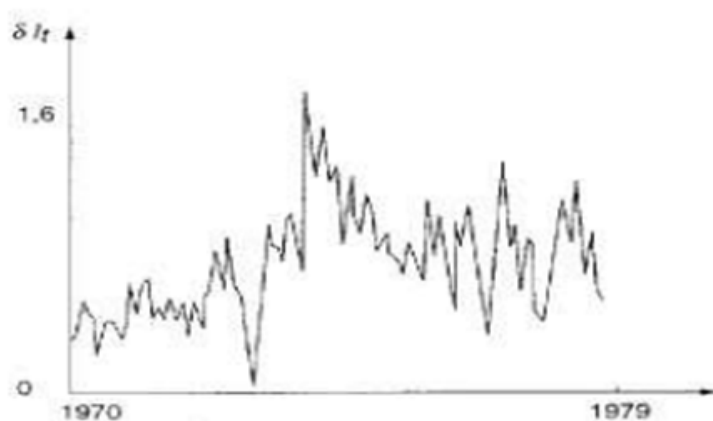


Fig-1.3 : Accroissement relatif mensuel de l'indice des prix.

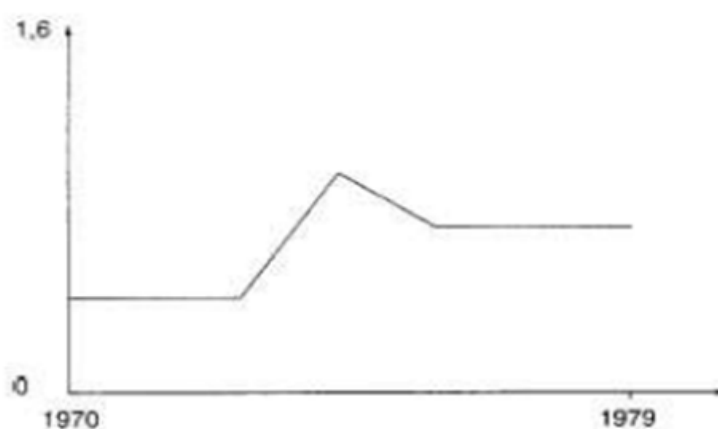


Fig-1.4 : Evolution à moyen terme de l'accroissement relatif mensuel de l'indice des prix.

On considère qu'une série chronologique (X_t) est la résultante de différentes composantes fondamentales :

- **La tendance** (ou trend) (T) représente l'évolution à long terme de la série étudiée. Elle traduit le comportement « moyen » de la série.

(Par exemple, la série de la *Figure 1.1*). a tendance à augmenter de façon linéaire.

- **La composante saisonnière** (ou saisonnalité) (S) correspond à un phénomène qui se répète à intervalles de temps réguliers (périodiques). En général, c'est un phénomène saisonnier d'où le terme de variations saisonnières.

(Par exemple, la série de la *Figure 1.2*). Présente des cycles réguliers au cours du temps et de même amplitude.

- **La composante résiduelle** (ou bruit ou résidu) (Z_t) ou (ε_t) correspond à des fluctuations irrégulières, en général de faible intensité mais de nature aléatoire. On parle aussi d'aléas.

(Par exemple, la série de la *Figure 1.3*). a un comportement assez irrégulier : il y a comme une sorte de bruit de faible amplitude qui perturbe les données.

Les modèles présentés dans ce cours tiennent compte de ces trois composantes (tendance, saisonnalité et fluctuations irrégulières). Il faut cependant remarquer que l'on pourrait envisager d'autres composantes.

- **Des phénomènes accidentels** (grèves, conditions météorologiques exceptionnelles, crash financier) peuvent notamment intervenir.

(Par exemple, la série de la *Figure 1.4*). Présente deux cassures.

1.2 Modélisation de la partie aléatoire

1.2.1 Processus stationnaire

Dans la suite, on n'étudiera que des processus stationnaires c'est à dire des processus dont la loi ne varie pas au cours du temps.

Cette hypothèse de stabilité temporelle de la loi du processus facilite évidemment bien des analyses. Pratiquement, cette hypothèse implique notamment que la covariance entre les valeurs prises en 2 dates distinctes par le processus ne dépend que de la longueur de la période qui les sépare (et pas de la date initiale), on dit que le processus est faiblement stationnaire si ses moments d'ordre 1 et 2 sont stationnaires :

$$E(X_t), \quad V(X_t) \text{ et } Cov(X_t, X_{t-h}) \text{ ne dépendent pas de } t ,$$

soit

$$\begin{cases} E(X_t) &= \mu, \\ V(X_t) &= \sigma^2, \\ Cov(X_t, X_{t-h}) &= \gamma_h. \end{cases}$$

La stationnarité est donc une hypothèse essentielle pour pouvoir estimer les paramètres caractéristiques du processus: elle implique notamment que le processus n'a pas de tendance.

1.2.2 Bruit blanc fort

Un processus $\{X_t\}$ est un bruit blanc fort s'il constitue un échantillon i.i.d suivant une distribution F . d'espérance nulle:

$$\forall t : X_t \sim F, \quad E(X_t) = 0,$$

$$\forall t \neq s : (X_t, X_s) \text{ indépendants.}$$

Un tel processus n'a ni tendance ni mémoire: la connaissance de la valeur du processus à une date donnée n'apporte aucune information pour la prédiction de sa valeur à une date ultérieure:

$$\forall t > s : \mathcal{L}(X_t | X_s) = \mathcal{L}(X_t).$$

1.2.3 Bruit blanc faible

Un processus $\{X_t\}$ est un bruit blanc faible s'il est stationnaire, centré et non autocorrélé ie:

$$\forall t : X_t \sim F, \quad E(X_t) = 0 \text{ et } Var(X_t) = \sigma_{X_t}^2,$$

$$\forall t \neq s : Cov(X_t, X_s) = 0.$$

1.2.4 Autocorrélations:

Définition 1.2.1 On étudie la "mémoire" d'un processus en calculant son autocorrélation de retard h noté ρ_h :

$$\rho_h = Corr(X_t, X_{t-h}) = \frac{Cov(X_t, X_{t-h})}{\sqrt{Var(X_t)Var(X_{t-h})}}, \quad (1.2.1)$$

qui mesure le lien entre les valeurs du processus à deux dates distantes de h . Pour un processus stationnaire, ρ_h prend une forme plus simple:

$$\rho_h = \frac{\text{Cov}(X_t, X_{t-h})}{\text{Var}(X_t)} = \frac{\gamma_h}{\gamma_0} \quad (1.2.2)$$

On peut tracer la courbe $\rho_h = f(h)$ qui est appelée (auto-)corrélogramme.

Définition 1.2.2 De même, on définit l'autocorrélation partielle de retard h comme la corrélation entre $(X_t - X_t^*)$ et $(X_{t-h} - X_{t-h}^*)$ où X_t^* désigne la régression de X_t sur les $(h-1)$ valeurs $\{X_{t-1}, X_{t-2}, \dots, X_{t-h+1}\}$:

$$\tau_h = \text{Corr}(X_t - X_t^*, X_{t-h} - X_{t-h}^*) = \frac{\text{Cov}(X_t - X_t^*, X_{t-h} - X_{t-h}^*)}{\sqrt{\text{Var}(X_t - X_t^*)\text{Var}(X_{t-h} - X_{t-h}^*)}}, \quad (1.2.3)$$

avec

$$X_t^* = \sum_{k=1}^{h-1} \alpha_k X_{t-k}, \quad X_{t-h}^* = \sum_{k=1}^{h-1} \beta_k X_{t-k}, \quad (1.2.4)$$

où les α_k et les β_k sont les coefficients des régressions.

Cette quantité rend compte de l'intensité de la liaison entre X_t et X_{t-h} en supprimant les liaisons induites par des variables intermédiaires $\{X_{t-1}, X_{t-2}, \dots, X_{t-h+1}\}$.

On peut ainsi remarquer que pour tout processus,

$$\rho_1 = \tau_1$$

puisque qu'il n'y a aucune variable intermédiaire entre X_t et X_{t-1} .

Remarque 1.2.1 Pour étudier les processus (et donc les séries chronologiques), on définit des opérateurs retard et de différenciation.

1.2.5 Opérateur retard **B**

L'opérateur **B** décale le processus d'une unité de temps vers le passé:

$$B(X_t) = X_{t-1}.$$

Si on applique h fois cette opérateur, on décale le processus de h unités de temps:

$$B(B(\dots B(X_t)\dots)) = B^h(X_t) = X_{t-h}. \quad (1.2.5)$$

1.2.6 Opérateur différence Δ

L'opérateur Δ fait la différence entre le processus et sa version décalée d'une unité de temps. cet opérateur se construit en utilisant l'opérateur précédent:

$$\begin{aligned}\Delta(X_t) &= X_t - X_{t-1} = X_t - B(X_t) \\ \Leftrightarrow \quad \Delta &= I - B,\end{aligned}\tag{1.2.6}$$

où I est l'opérateur identité:

$$I(X_t) = X_t$$

$$I(X_t) = X_t.$$

Dans la littérature, cet opérateur est parfois noté ∇ .

1.2.7 Propriétés

Elimination de la tendance:

L'opérateur Δ élimine les tendances linéaires. Pour un processus de la forme

$$X_t = at + b + E_t$$

où E_t est stationnaire, on a

$$\Delta(X_t) = X_t - X_{t-1} = (at + b + E_t) - (a(t-1) + b + E_{t-1}) = a + (E_t - E_{t-1}).$$

De façon générale, l'opérateur Δ^d élimine les tendances polynomiales de degré d .

Par exemple, pour une tendances de degré 2 :

$$X_t = at^2 + bt + c + E_t.$$

on a

$$\Delta^2 = (I - B)^2 = I - 2B + B^2,$$

et donc

$$\Delta^2(X_t) = X_t - 2X_{t-1} + X_{t-2} = \dots = 2a + (E_t + 2E_{t-1} + E_{t-2}). \quad (1.2.7)$$

Il faut bien noter que l'opérateur Δ^d raccourcit la série des observations x_t de d valeurs puisque $\Delta^d X_t$ n'est défini que pour $t > d$.

Élimination de la saisonnalité:

L'opérateur $\Delta_\omega = (I - B^\omega)$ élimine une saisonnalité de période ω (ω entier). Si on a un processus de la forme

$$X_{k\omega+j} = s_j + E_{k\omega+j}$$

où E_t est stationnaire, on a

$$\begin{aligned} \Delta_\omega(X_{k\omega+j}) &= (I - B^\omega)(X_{k\omega+j}) = X_{k\omega+j} - X_{(k-1)\omega+j} \\ &= (s_j + E_{k\omega+j}) - (s_j + E_{(k-1)\omega+j}) = E_{k\omega+j} - E_{(k-1)\omega+j}. \end{aligned}$$

On utilise fréquemment les opérateurs Δ_{12} pour les données mensuelles, Δ_7 pour les données journalières et Δ_{24} pour les données horaires etc.

Cet opérateur est également coûteux en information puisqu'il raccourcit la série de ω valeurs, les ω premières servant de références pour la saisonnalité.

1.3 Différents modèles de séries temporelles

1.3.1 Modèles de Box-Jenkins

Les modèles de *Box et Jenkins* sont des modèles linéaires qui peuvent être stationnaires ou non stationnaires. Les modèles stationnaires sont regroupés dans la famille des modèles *ARMA*, qui est composée des modèles Autorégressifs (*AR*), des modèles en Moyennes Mobiles (*MA*) et des modèles mixtes Autorégressifs-Moyennes-Mobiles (*ARMA*).

Le modèle AR(p)

Définition 1.3.1 Soit X_t un processus stochastique stationnaire, on dit que X_t est un processus autorégressif d'ordre p noté $AR(p)$ le processus qui vérifie l'équation suivante:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t, \quad (1.3.1)$$

où $\varepsilon_t \sim WN(0, \sigma^2)$.

Remarque 1.3.1 le polynôme $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ est le polynôme caractéristique du processus autoregressif. Le modèle s'écrit souvent $\phi(B)X_t = \varepsilon_t$.

Exemple 1.3.1 Un processus $AR(1)$ prend la forme suivante:

$$X_t = \phi X_{t-1} + \varepsilon_t, \quad \text{où } \{\varepsilon_t\} \sim WN(0, \sigma_\varepsilon^2). \quad (1.3.2)$$

1.3.2 Estimation pour les modèles AR(p) par la méthode des moindres carrés

Un modèle $AR(p)$ s'écrit

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t$$

où ε_t est un bruit blanc,

$$X_t = Z_t' \beta + \varepsilon_t \quad \text{où } Z_t' = (X_{t-1}, X_{t-2}, \dots, X_{t-p}) \quad \text{et } \beta = (\phi_1, \phi_2, \dots, \phi_p).$$

L'estimation des paramètres du modèle $X = Z' \beta + \varepsilon$ par la méthode des moindres carrés donne

$$\hat{\beta} = (ZZ')^{-1}ZX \quad \text{et} \quad \hat{\sigma}^2 = \frac{1}{T - (p+1)} \sum_{t=1}^T (X_t - Z_t' \hat{\beta})^2. \quad (1.3.3)$$

Principales caractéristiques d'un modèle AR(p) La fonction d'autocorrélation

La fonction d'autocovariance d'un processus autorégressif $AR(p)$ est donnée par l'équation de récurrence suivante:

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p} \quad k = 1, 2, \dots, p. \quad (1.3.4)$$

En divisant l'équation (1.3.1) par γ_0 , on obtient la fonction d'autocorrélation satisfaisante à l'équation de réccurence suivante:

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \dots + \phi_p\rho_{k-p} \quad k = 1, 2, \dots, p. \quad (1.3.5)$$

L'équation (1.3.2) peut être exprimée alors telle que:

$$\phi(B)\rho_k = 0. \quad (1.3.6)$$

Le modèle MA(q)

Définition 1.3.2 Soit X_t un processus stochastique stationnaire, on dit que X_t est un processus moyenne mobile d'ordre q noté MA(q) le processus qui vérifie l'équation suivante:

$$X_t = \varepsilon_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \dots + \theta_q Z_{t-q}, \quad (1.3.7)$$

où $\varepsilon_t \sim WN(0, \sigma^2)$.

Remarque 1.3.2 Un tel processus est donc linéaire pour la suite sommable :

$$\psi_j = 0, j < 0, \psi_0 = 1, \psi_1 = \theta_1, \dots, \psi_q = \theta_q, \psi_j = 0, j > q.$$

On écrit habituellement $X_t = \theta(B)Z_t$ où $\theta(B)$ est polynôme

$$\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q, \quad (1.3.8)$$

et B est l'opérateur de décalage arrière (appelé aussi opérateur retard). Le polynôme $\theta(B)$ est appelé polynôme caractéristique du processus moyen mobile.

Le modèle ARMA(p,q)

Définition 1.3.3 On dit qu'un processus stationnaire et centré X_t est un processus ARMA(p, q) si pour chaque t il vérifie la relation:

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} = \varepsilon_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \dots + \theta_q Z_{t-q}, \quad (1.3.9)$$

où $\varepsilon_t \sim WN(0, \sigma^2)$.

Remarque 1.3.3 En utilisant l'opérateur B , l'équation (1.3.4) peut être exprimée telle que:

$$\phi(B)X_t = \theta(B)Z_t, \quad (1.3.10)$$

où $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ et $1 + \theta_1 B + \dots + \theta_q B^q$ sont appelés *polynômes caractéristiques respectifs des parties autorégressives et moyennes mobiles du processus*.

1.3.3 Modèles linéaires non-stationnaires

Modèle ARIMA

La classe des modèles non stationnaires est très utile dans le domaine de l'économétrie et les mathématiques financières. Ce genre de modèles permet de représenter des processus présentant un certain type de non-stationnarité que l'on peut éliminer par différentiation d'ordre d .

Définition 1.3.4 On dit qu'un processus X_t est un ARIMA (Autorégressive Integrated Moving Average) d'ordre (p, d, q) si $\Delta^d X_t$ est un processus ARMA(p, q).

On modélise alors le processus X_t sous la forme:

$$\phi(B)(1 - B)^d X_t = \theta(B)\Delta^d X_t = \theta(B)Z_t, \quad (1.3.11)$$

où le polynôme $\phi(B)$ est de degré p et le polynôme $\theta(B)$ est de degré q .

On écrit que la série X_t suit un processus ARIMA(p, d, q).

Remarque 1.3.4 Les polynômes $\phi(B)$ et $\theta(B)$ sont de degrés p et q respectivement et $\phi(B) \neq 0$ pour $|B| \leq 1$, alors que le polynôme $\theta(B)$ admet une racine multiple $B = 1$.

Modèle SARIMA

Nous avons déjà introduit l'importante classe des modèles ARMA pour représenter un processus stationnaire. Les modèles SARIMA sont une généralisation de cette classe qui incorpore une vaste gamme de processus non stationnaires.

Définition 1.3.5 Soient d et D des entiers non-négatifs, alors un processus $\{X_t\}$ est dit processus SARIMA si $Y_t = (1 - B)^d(1 - B^s)^D X_t$ est un processus ARMA de la forme:

$$\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)\varepsilon_t, \quad (1.3.12)$$

où $\varepsilon_t \sim WN(0, \sigma^2)$.

Remarque 1.3.5 L'opérateur $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ est l'opérateur générateur d'un $AR(p)$, L'opérateur $\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$ est l'opérateur générateur d'un $MA(q)$ et où, pour la saisonnalité $Y_t - Y_{t-s}$, l'opérateur $\Phi(B) = 1 - \Phi_1 B - \dots - \Phi_P B^P$ est l'opérateur générateur $AR(P)$, et l'opérateur $\Theta(B) = 1 + \Theta_1 B + \dots + \Theta_Q B^Q$ est l'opérateur générateur d'un $MA(Q)$. On note que le processus $\{Y_t\}$ est causal si et seulement si $\phi(B) \neq 0$ et $\Phi(B) \neq 0$ pour $|B| \leq 1$. Les valeurs de d et D sont en général inférieures à 1, alors que les valeurs de p et q sont généralement inférieures à 2.

1.4 Identification et estimation des paramètres

1.4.1 Identification du modèle: méthode de Box & Jenkins

Le choix entre les différents modèles présentés ici ($AR(p)$, $MA(q)$, $ARMA(p, q)$, $ARIMA(p, d, q)$, $SARIMA$, etc) ne peut généralement pas se faire a priori. On est le plus souvent réduit à des tâtonnements par un système d'essais/erreurs. Une méthodologie "pas à pas" qui implique la remise en cause de chaque modèle envisagé jusqu'à obtenir un modèle acceptable. Un modèle est acceptable lorsqu'il prend en compte toute la structure de la partie aléatoire et ne laisse qu'un bruit blanc. Par exemple, on considère qu'un modèle $ARMA(p, q)$ est acceptable si on peut accepter l'hypothèse selon laquelle

$$\{E_t\} = \{X_t - \varphi_1 X_{t-1} - \dots - \varphi_p X_{t-p} - \theta_1 E_{t-1} - \dots - \theta_q E_{t-q}\}, \quad (1.4.1)$$

est un bruit blanc.

La méthodologie de Box et Jenkins vise à formuler un modèle permettant de représenter une chronique comme finalité de prévoir des valeurs futures. De ce fait, l'objet de cette méthodologie est de modéliser une série temporelle en fonction de ses valeurs passées et présentes afin de déterminer le processus $ARIMA$ adéquat.

Cette méthodologie suggère une procédure à cinq étapes essentielles :

1. Analyse préliminaire.

2. Identification du modèle.
3. Estimation du modèle.
4. Validation du modèle.
5. Prévision.

1.4.2 Estimation des paramètres : méthode des m.c.o.

En utilisant la méthode des moindres carrés on tente de trouver les paramètres \hat{a} et \hat{b} de la droite $\hat{y}_t = \hat{a} + \hat{b}x_t$ qui approche le mieux l'indépendance des y/x .

La méthode des moindres carrés consiste alors à minimiser la quantité :

$$S(\hat{a}, \hat{b}) = \sum_{t=1}^n \hat{e}_t^2, \quad (1.4.2)$$

avec

$$\begin{aligned} \hat{e}_t &= y_t - \hat{y}_t \\ &= y_t - \hat{a} - \hat{b}x_t. \end{aligned} \quad (1.4.3)$$

Autrement dit, les estimateurs des moindres carrés minimisent la somme des carrés des distances verticales des points du nuage de la droite ajustée.

On a les moyennes :

$$\begin{aligned} \bar{X} &= \frac{1}{n} \sum_{t=1}^n x_t \\ \bar{Y} &= \frac{1}{n} \sum_{t=1}^n y_t. \end{aligned} \quad (1.4.4)$$

Les écarts types sont donnés comme suit :

$$\begin{aligned} s_x &= \sqrt{\frac{1}{n-1} \sum_{t=1}^n (x_t - \bar{X})^2} \\ s_y &= \sqrt{\frac{1}{n-1} \sum_{t=1}^n (y_t - \bar{Y})^2}. \end{aligned} \quad (1.4.5)$$

La covariance :

$$\text{cov}(x, y) = s_{xy} = \frac{1}{n-1} \sum_{t=1}^n (x_t - \bar{X})(y_t - \bar{Y}). \quad (1.4.6)$$

Calcul des estimateurs des M.C.O :

$$\begin{aligned} \frac{\partial S(a,b)}{\partial a} = 0 &\implies \frac{\partial}{\partial a} (\sum_{t=1}^n (y_t - a - bx_t)^2) = 0 \\ \frac{\partial S(a,b)}{\partial b} = 0 &\implies \frac{\partial}{\partial b} (\sum_{t=1}^n (y_t - a - bx_t)^2) = 0 \end{aligned} \quad (1.4.7)$$

La première équation a pour solution :

$$\hat{a} = \bar{Y} - \hat{b}\bar{X}. \quad (1.4.8)$$

Et la deuxième, compte tenu de la première solution, donne :

$$\hat{b} = \frac{\sum_{t=1}^n y_t x_t - n\bar{X}\bar{Y}}{\sum_{t=1}^n x_t^2 - n\bar{X}^2}. \quad (1.4.9)$$

Donc l'équation du modèle estimé est :

$$\hat{y}_t = \hat{a} + \hat{b}x_t. \quad (1.4.10)$$

1.5 Détermination du type et de l'ordre du modèle

On s'intéresse ici à l'identification d'un processus stationnaire : on se limite aux processus $ARMA(p, q)$.

On suppose donc que la saisonnalité et la tendance ont été supprimées. Comme on l'a vu précédemment, cette élimination peut s'opérer au moyen de différentiations de différents types:

1. $\Delta^d = (I - B)^d$ pour éliminer une tendance polynômiale de degré d .
2. $\Delta_\omega = (I - B^\omega)$ pour éliminer une saisonnalité de période ω .

Une fois cette manipulation effectuée, on tente de reconnaître le type de processus auquel on a affaire en étudiant les corrélogrammes. Leurs formes théoriques étant connues, on choisit alors un modèle autorégressif, moyenne mobile ou $ARMA$: La lecture des corrélogrammes se fait en recherchant les chutes p et q :

dans le corrélogramme partiel pour un $AR(p)$.

dans le corrélogramme pour un $MA(q)$.

1.6 Estimation des paramètres et prévision

1.6.1 Méthode des moments

La méthode des moments est une des plus utilisées: pour un modèle donné, on sait que les autocorrélations ρ_h et τ_h dépendent des paramètres $\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_d$ selon des équations théoriques connues:

$$(\rho_1, \rho_2, \dots, \tau_1, \tau_2, \dots) = F(\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_d).$$

Or on sait facilement estimer les autocorrélations ρ_h et τ_h , il suffit donc d'inverser les équations pour estimer les paramètres:

$$(\hat{\varphi}_1, \dots, \hat{\varphi}_p, \hat{\theta}_1, \dots, \hat{\theta}_d) = F^{-1}(\hat{\rho}_1, \hat{\rho}_2, \dots, \hat{\tau}_1, \hat{\tau}_2, \dots).$$

Il n'est pas nécessaire d'explicitier ici la forme des résultats.

1.6.2 Maximum de vraisemblance

L'autre méthode classique d'estimation est le maximum de vraisemblance. Elle nécessite de faire des hypothèses sur la loi du bruit blanc $\{E_t\}$: L'hypothèse la plus fréquemment retenue est qu'il s'agit d'un bruit blanc gaussien de variances σ^2 .

On obtient alors des estimateurs du maximum de vraisemblance des paramètres $\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_d$ et de σ .

Cette méthode peut donner lieu à des calculs très lourds aussi bien mathématiquement que numériquement à cause de la forme très "tourmentée" des fonctions de vraisemblance.

Une méthode de ré-échantillonnage : Le bootstrap

Introduction

Les méthodes classiques d'inférence statistique ne permettent pas d'obtenir des réponses correctes à tous les problèmes concrets que se pose l'utilisateur. Elles ne sont en effet valables que sous des conditions d'application particulières.

Que peut faire l'utilisateur, en pratique, lorsque ces conditions d'application ne sont pas remplies? différentes attitudes sont possibles.

Dans certains cas, les méthodes classiques sont utilisées malgré le non-respect des conditions. Cette utilisation est alors justifiée par le caractère robuste des méthodes qui garantit que les résultats de l'inférence restent approximativement valables.

Le recours à des transformations de variables permet, dans certains cas, de se rapprocher des conditions d'application.

Ainsi une transformation logarithmique, par exemple, peut rendre normales des distributions qui, au départ, ne le sont pas.

Une troisième attitude consiste à abandonner les méthodes paramétriques d'inférence statistique au profit de méthodes non paramétriques, pour lesquelles les conditions d'application sont bien moins restrictives.

L'accès généralisé à des moyens de calcul puissants a permis le développement de méthodes d'inférence statistique basées sur l'utilisation intensive de l'ordinateur. *Le bootstrap* fait partie de ces méthodes.

Le bootstrap provient de l'expression anglaise « to pull oneself up by one's bootstrap » [17], qui évoque l'action de « se hisser en tirant sur ses propres lacets », désigne un ensemble de méthodes qui consistent à faire de l'inférence statistique sur de « nouveaux » échantillons tirés à partir d'un échantillon initial. Disposant d'un échantillon destiné à donner une certaine information sur une population, on tire au sort, parmi la sous-population réduite à cet échantillon, un nouvel échantillon de même taille n . Et on répète cette opération B fois, où B est grand. On analyse ensuite les nouvelles observations ainsi obtenues pour affiner l'inférence faite sur les observations initiales. Le bootstrap est une technique destinée à faciliter de l'inférence dans les situations complexes où les méthodes analytiques ne suffisent pas.

L'objectif de ce chapitre est de décrire comment le bootstrap peut être utilisé pour résoudre les problèmes d'inférence statistique en relation avec l'estimation des paramètres. Nous présentons d'abord la méthode de ré-échantillonnage. Ensuite, nous examinons l'exemple de l'estimation de l'erreur-standard de la moyenne et de la médiane pour comprendre l'intérêt de bootstrap. Des informations plus détaillées concernant le bootstrap et d'autres méthodes de ré-échantillonnage sont données, notamment, par Chernick [12], Efron et Tibshirani [17] et Magalie Fromont and Myriam Vimond [21].

2.1 Le principe de *bootstrap*

Illustration

Le principe qui sous-tend le bootstrap est très simple et très ancien, et il peut être illustré par un système d'emboîtement (voir P. Hall, [22]) tel que celui des poupées russes:

Il s'agit d'une poupée qui, lorsqu'on l'ouvre, laisse apparaître une poupée identique mais plus petite (« homothétique »), qui à son tour contient une poupée plus petite, etc... . Imaginons que l'on veuille estimer le nombre r des taches de rousseur de la première de ces poupées, qui est aussi la plus grande, et que l'on ne puisse pas l'observer. On suppose qu'on dispose seulement de la seconde, contenue dans la première, et qui contient toutes les autres. Soit r_0 le nombre des taches de rousseur de la seconde. On peut, en première approximation, estimer r par r_0 . On appelle « Poupée » la plus grande poupée, non observée, « poupée 0 » celle qui est observée, « poupée 1 » celle qu'on trouve dans la poupée 0, et ainsi de suite pour toutes les poupées plus petites que la poupée 1, qui sont toutes observables puisque contenues dans la poupée 1. Comme la Poupée initiale est plus grande que la poupée numéro 0, on s'attend à ce que r soit plus grand que r_0 et dans le rapport de leurs tailles. Cependant, on ne peut pas observer la première poupée et on ne connaît donc pas sa taille. En revanche, on peut observer le nombre des taches de rousseur r_1 de la troisième poupée. Donc, si le rapport du nombre des taches de rousseur d'une poupée à la suivante est toujours le même, le rapport $\frac{r_0}{r_1}$, qui lui est observable, fournira une estimation du rapport $\frac{r}{r_0}$. Cela donne comme estimateur de r :

$$\hat{r} = r_0 \frac{r_0}{r_1}.$$

Mais il se peut que le rapport de ces nombres ne soit pas constant, ce que l'on peut vérifier en comparant $\frac{r_0}{r_1}$ à $\frac{r_1}{r_2}$ par exemple, puisque ces deux quantités sont observables. Si ces deux quantités ne sont pas égales, $\frac{r_0}{r_1}$ ne constitue qu'une approximation pour la valeur de $\frac{r}{r_0}$.

Pour effectuer une correction supplémentaire, si on peut supposer que le rapport des tailles d'une poupée à la suivante, bien que n'étant plus constant, varie régulièrement, par exemple les rapports de taille d'une poupée à la suivante sont dans un rapport constant, c'est à dire que :

$$\frac{r/r_0}{r_0/r_1} = \frac{r_{i-1}/r_i}{r_i/r_{i+1}}.$$

Alors, on peut effectuer une correction supplémentaire en observant r_2 sur la poupée suivante et en prenant pour estimateur de r la valeur précédente multipliée par

$$(r_0/r_1) / (r_1/r_2),$$

ce qui donne :

$$\hat{r} = r_0 \left(\frac{r_0}{r_1} \right)^2 \frac{r_2}{r_1}.$$

On peut à nouveau vérifier si l'hypothèse faite sur la variation des rapports est exacte en considérant la poupée suivante, et, dans le cas contraire, effectuer une nouvelle correction.

2.2 Méthodes de ré-échantillonnage

2.2.1 Bootstrap des individus

On considère un échantillon de n observations:

$$x_1, x_2, \dots, x_i, \dots, x_n ,$$

prélevé de manière aléatoire et simple dans une population. Ces observations peuvent concerner une seule variable, ou, au contraire, être relatives à plusieurs variables. Dans ce cas, les x_i représentent des vecteurs de dimension p , p étant le nombre de variables. Afin de ne pas alourdir les notations, nous ne distinguerons pas ces deux situations et, de manière plus condensée, nous désignerons l'échantillon initial par le symbole x , qu'il s'agisse d'un vecteur ou d'une matrice.

Le principe de la méthode du bootstrap est de prélever une série d'échantillons aléatoires et simples avec remise de n observations dans l'échantillon initial, considéré comme une population. Ces échantillons successifs seront notés :

$$x^{*1}, x^{*2}, \dots, x^{*b}, \dots, x^{*B}$$

B étant le nombre de ré-échantillonnages effectués.

A titre d'exemple, nous considérons un problème d'inférence statistique qui est l'estimation de l'erreur standard SE d'une statistique d'intérêt (moyenne, médiane, les coefficients de corrélation...) d'un échantillon x en suivant cet algorithme:

1) Soit un échantillon observé (initial)

$$x = (x_1, x_2, \dots, x_n)$$

et une statistique d'intérêt $s(x)$: moyenne, médiane, ...

2) Ensuite nous tirons B échantillons aléatoires et simples avec remise de n observations dans l'échantillon initial ces échantillons sont appelées *échantillons bootstrap*:

$$\begin{aligned} x^{*1} &= (x_1^{*1}, x_2^{*1}, \dots, x_n^{*1}) \\ x^{*2} &= (x_1^{*2}, x_2^{*2}, \dots, x_n^{*2}) \\ &\dots \\ x^{*B} &= (x_1^{*B}, x_2^{*B}, \dots, x_n^{*B}). \end{aligned}$$

3) Puis nous calculons la statistique d'intérêt pour chaque échantillon bootstrap (réplications bootstrap de s) que nous allons noter comme suit:

$$\begin{aligned} &s(x^{*1}) \\ &s(x^{*2}) \\ &\dots \\ &s(x^{*b}) \\ &\dots \\ &s(x^{*B}). \end{aligned}$$

Enfin nous donnons l'estimée bootstrap de l'erreur standard = écart-type des réplifications bootstrap.

$$SE(s(x^{*b})) = \sqrt{\frac{\sum_{b=1}^B [s(x^{*b}) - s^*]^2}{B - 1}},$$

avec

$$s^* = \sum_{b=1}^B s(x^{*b})/B.$$

Exemple 2.2.1 *Estimation de l'erreur standard de la moyenne :*

Durée de survie d'un groupe 1(placebo). L'échantillon initial:

$$x_1 = (52, 10, 40, 104, 50, 27, 146, 31, 46),$$

est de taille $n_1 = 9$ mesures, et la statistique d'intérêt est la moyenne:

$$m_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i = 56.22.$$

Erreur standard

$$SE_1 = \sqrt{Var_1/n_1} = 13.33.$$

Application du bootstrap:

<i>B échantillons bootstrap</i>	<i>calcul de la moyenne</i>	<i>réplications bootstrap de la moyenne</i>
↓		↓
$x^{*1} = (50, 10, 40, 50, 46, 10, 146, 40, 50)$	→	49.11
$x^{*b} = (10, 52, 104, 40, 104, 46, 50, 146, 27)$	→	64.33
...		...
$x^{*B} = (146, 31, 31, 10, 27, 40, 104, 46, 50)$	→	53.89

estimée bootstrap de l'erreur standard est:

$$SE(m_1) = \sqrt{\frac{\sum_{b=1}^B [m_1(x^{*b}) - m_1^*]^2}{B-1}} = 13.32 ,$$

avec

$$m_1^* = \sum_{b=1}^B m_1(x^{*b}) / B = 55.73.$$

Dans cet exemple on n'a pas besoin de bootstrap car sa prend du temps de calcul.

Exemple 2.2.2 Estimation de l'erreur standard de la médiane :

laissons les mêmes données de l'exemple précédent, on a la moyenne

$$m_1 = 56.22,$$

et la médiane

$$\mu_1 = 46.$$

erreur standard sur m_1 :

$$\hookrightarrow \text{Classique} : se_1 = 13.33.$$

$$\hookrightarrow \text{Bootstrap} : se_1^* = 13.32.$$

erreur standard sur μ_1

$$\hookrightarrow \text{Classique} : ?$$

$$\hookrightarrow \text{Bootstrap} : se_1^* = 11.54.$$

Remarque 2.2.1 La mise en œuvre de la méthode du bootstrap ne présente guère d'intérêt lorsque l'inférence statistique peut être réalisée par des méthodes analytiques classique pour

lesquelles les conditions d'applications sont remplies, elles ne sont pas donc destinées à remplacer les méthodes classiques lorsque celles-ci sont applicables mais plutôt à fournir des réponses à des questions pour lesquelles les méthodes classiques sont inapplicables ou non disponibles.

2.3 Conditions de convergence du Bootstrap

1. Les observations "i.i.d".
2. Convergence uniforme certaine (avec la probabilité 1) de la fonction de répartition empirique vers la vraie fonction de répartition quand la taille tend vers l'infini (théorème de **Glivenko-Cantelli**).

Théorème 2.3.1 (Glivenko-Cantelli) *On a X_1, \dots, X_n une série de variables aléatoires de fonction de répartition F sur \mathbb{R} . La fonction de répartition empirique est la fonction de x définie par :*

$$F_h(x) = \frac{1}{n} \sum_{1 \leq i \leq n} I\{X_i \leq x\}. \quad (2.3.1)$$

Pour tout $x \in \mathbb{R}$, on peut appliquer la loi des grands nombres sur la série $I\{X_i \leq x\}$, $i = 1, \dots, n$, pour affirmer que $F_h(x) \rightarrow F(x)$ presque sûrement (p.s). (Afin d'appliquer la loi des grands nombres, on a besoin de démontrer que $E[|I\{X_i \leq x\}|] < \infty$, qui est dans ce cas négligeable parce que $|I\{X_i \leq x\}| \leq 1$). Dans ce sens, $F_h(x)$ est un estimateur de $F(x)$ pour tout $x \in \mathbb{R}$.

2.4 Règles empiriques

1) Même un petit nombre de répliques fournit déjà des informations très utiles. $B = 50$ est souvent suffisant pour une estimation fiable de l'erreur standard.

2) Il est rare que plus de 200 répliques soient nécessaires pour estimer les *erreurs standard*.

Exemple 2.4.1 (erreur standard de la moyenne m_2) *Soit*

$$x_2 = (94, 38, 23, 197, 99, 16, 141),$$

$n_2 = 7$, la moyenne $m_2 = 86.86$, l'erreur standard classique sur m_2 est:

$$se_2 = \sqrt{Var_2/n_2} = 25.24,$$

et l'erreur standard bootstrap sur m_2 est : $se_2^* = 23.81$.

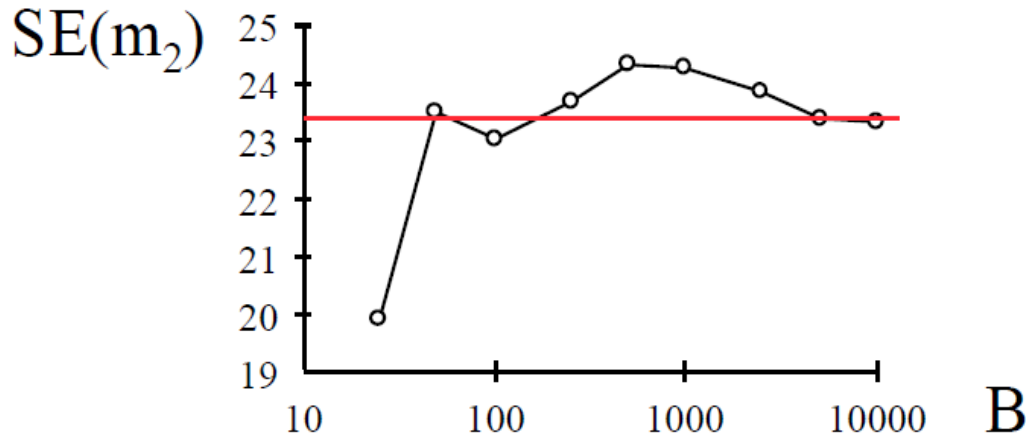


Fig-2.1 :Erreur standard de la moyenne m_2 (Voir[11])

2.5 Avantages et inconvénients

1. Les méthodes de ré-échantillonnages sont simples à programmer, dans ce contexte ces techniques rencontrent un succès considérable.
2. Les méthodes de calcul intensif (ré-échantillonnages) ne sont pas meilleures que les méthodes classiques quand les hypothèses classiques sont satisfaites. Mais elles sont plus puissantes quand ces hypothèses ne sont pas satisfaites.
3. En bref, ces méthodes peuvent résoudre de nombreux problèmes qu'on ne savait pas traiter auparavant.

Des dangers à prévenir:

Attention :

- le volume des calculs ne garantit pas que les données ont été bien utilisées.

• Il est important de savoir analyser dans quelles circonstances les méthodes de calcul intensif ont un apport décisif sur les méthodes classiques.

Remarque 2.5.1 *Avant de passer à l'application du bootstrap des séries chronologiques, nous allons définir les types de données (les données non structurées et les données structurées).*

1) Données non structurées (indépendantes):

↔ Les valeurs de l'échantillon observé sont indépendantes.

↔ une modification de l'ordre des valeurs ne modifie pas l'échantillon.

exemples:

- Durée de survie des animaux.

- Notes des étudiants aux tests.

2) Données structurées (dépendantes):

↔ Les valeurs de l'échantillon observé ne sont pas indépendantes.

↔ L'ordre des valeurs dans l'échantillon est important.

exemples:

- Image

- Série temporelle ou chronologique.

Bootstrap des séries temporelles

Introduction

La méthode du bootstrap initialement introduite par Efron [18] pour variables indépendantes et plus tard étendu à traiter avec des variables dépendantes plus complexes, par plusieurs auteurs [12]. C'est une classe de méthodes non paramétriques qui permettent au statisticien d'effectuer l'inférence statistique sur un large éventail de problèmes sans imposer beaucoup d'hypothèses structurelles sur le processus aléatoire de données générant. A présent, il existe plusieurs livres et monographies, e.g. Hall [22], Efron et Tibshirani [17] et Lahiri [31]. Parmi autres qui décrivent les différents aspects de la méthodologie bootstrap à divers niveaux de sophistication et de généralité. En outre; plusieurs articles dans la littérature donnent des aperçus sur divers aspects de bootstrap des séries chronologiques, parmi eux se trouvent Berkowitz and Kilian [3], Buhlmann [10], Li and Maddala [32] et Politis [38]. Ces documents considèrent les méthodes bootstrap et ré-échantillonnages pour les processus stochastiques et des modèles généraux de séries chronologiques.

L'utilisation du bootstrap avec des séries temporelles apporte quelques différences par rapport à un modèle classique, puisqu'il faut en plus estimer la structure de dépendance du processus.

L'objectif de ce chapitre consistera à expliciter les différentes possibilités d'application du bootstrap aux séries temporelles dans les cas de données non structurées (indépendantes) ou de données structurées (non indépendantes). Ces possibilités comportent entre autres: le bootstrap des résidus que nous expliquerons sur des séries autorégressives stationnaires $AR(p)$, le bootstrap par blocs que nous expliquerons aussi sur une série temporelle à données q -dépendantes et le test bootstrap que nous allons présenter dans le dernier chapitre comme un exemple d'application qui porte sur le test de racine unité dans un modèle $AR(1)$, basé sur le ré-échantillonnage des résidus.

3.1 Bootstrap des résidus

La technique de ré-échantillonnage présentée ici est la plus simple et la plus courante. Des méthodes un peu différentes sont utilisées pour des applications particulières. Ainsi, dans les problèmes de régression lorsque les valeurs des variables explicatives sont fixées a priori par l'utilisateur, le ré-échantillonnage d'individus peut difficilement se justifier. Dans une telle situation, on peut remplacer le *bootstrap des individus* par le *bootstrap des résidus*.

Soit y le vecteur de la variable à expliquer et x la matrice des variables explicatives. L'échantillon initial est donc décrit par la juxtaposition du vecteur y et de la matrice x .

Nous proposons de modéliser y en fonction de x dans un modèle de régression linéaire par:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i, i = 1, \dots, n. \quad (3.1.1)$$

Soit $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2)$ le vecteur des coefficients estimés par une méthode donnée d'ajustement, qui n'est pas nécessairement la méthode des moindres carrés. On peut calculer le vecteur des valeurs estimées de la variable à expliquer

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 + \varepsilon_i, i = 1, \dots, n. \quad (3.1.2)$$

et en déduire le vecteur des résidus :

$$\varepsilon_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i - \hat{\beta}_2 x_i^2, i = 1, \dots, n. \quad (3.1.3)$$

Nous sélectionnons B échantillons indépendants $\varepsilon^{*1}, \varepsilon^{*b}, \dots, \varepsilon^{*B}$ où ε^{*b} est un échantillon de taille n tiré à partir des $\varepsilon_i, i = 1, \dots, n$ avec remise.

Nous calculons, pour chaque échantillon, n nouvelles valeurs prédites

$$y_i^{*b} = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 + \varepsilon_i^{*b}. \quad (3.1.4)$$

Nous effectuons B régressions, des y^{*b} sur les x , grâce auxquelles on obtient B vecteurs $(\hat{\beta}_0^{*b}, \hat{\beta}_1^{*b}, \hat{\beta}_2^{*b})$. Nous déduisons les valeurs prédites

$$\hat{y}_x^{*b} = \hat{\beta}_0^{*b} + \hat{\beta}_1^{*b} x + \hat{\beta}_2^{*b} x^2 + \varepsilon_i^{*b} \quad (3.1.5)$$

Nous calculons l'erreur standard voulue pour les valeurs prédites,

$$SE = \sqrt{\frac{\sum_{b=1}^B (\hat{y}_x^{*b} - \hat{y}_x^*)^2}{B - 1}}, \quad (3.1.6)$$

où

$$\hat{y}_x^* = \frac{\sum_{b=1}^B \hat{y}_x^{*b}}{B}.$$

Mais, de manière plus générale, y_i peut être une fonction quelconque des valeurs observées des variables explicatives et des paramètres estimés.

3.2 Bootstrap par blocs

Nous décrivons brièvement la méthode de bootstrap par blocs, le (MBB).

Supposons que $\{X_t\}_{t \in \mathbb{N}}$ est une série chronologique dépendante et faiblement stationnaire et que $\{X_1, \dots, X_n\} \equiv X_n$ sont observées.

Soit l un entier naturel satisfaisant $1 \leq l < n$.

Décomposons la série en blocs indépendants de longueur l chacun, contenu dans X_n tel que:

$$\begin{aligned} B_1 &= (X_1, X_2, \dots, X_l) \\ B_2 &= (X_2, \dots, X_l, X_{l+1}) \\ &\cdot \\ &\cdot \\ &\cdot \\ B_N &= (X_{n-l+1}, \dots, X_n). \end{aligned}$$

où $N = n - l + 1$.

Pour des raisons de simplicité, nous supposons que l divise n , soit $b = \frac{n}{l}$.

Pour générer les échantillons de MBB, nous tirons b blocs aléatoirement et avec remise de l'ensemble de blocs $\{B_1, B_2, \dots, B_N\}$.

Ensuite, nous reconstituons des séries bootstrap en joignant les blocs tirés aléatoirement et avec remise.

Enfin, nous passons à l'estimation de la statistique d'intérêt sur chaque série temporelle bootstrap reconstituée.

Remarque 3.2.1 *Le choix de la longueur des blocs l est un inconvénient de la méthode.*

Noter que si nous fixons $l = 1$, alors le MBB se réduit à la méthode de bootstrap ordinaire de *Efron [18]* pour les données iid.

Cependant, pour une approximation valable dans le cas dépendant, il est généralement nécessaire que:

$$l^{-1} + n^{-1}l = o(1) \quad \text{quand } n \rightarrow +\infty. \quad (3.2.1)$$

Certains choix typiques de l sont

$$l = Cn^{\frac{1}{k}} \quad \text{pour } k = 3, 4, \dots, \quad (3.2.2)$$

où $C \in \mathbb{R}$ est une constante.

Supposons que la variable aléatoire d'intérêt est de la forme:

$$T_n = t_n(X_n; \theta(P_n)), \quad (3.2.3)$$

où $P_n = \mathcal{L}(X_n)$ désigne la distribution de probabilité conjointe de X_n . La forme MBB de T_n basé sur des blocs de taille l est définie comme suit :

$$T_n^* = t_n(X_1^*, \dots, X_n^*; \theta(\hat{P}_n)), \quad (3.2.4)$$

où $\hat{P}_n = \mathcal{L}(X_1^*, \dots, X_n^* | X_n)$, la distribution de probabilité conjointe conditionnelle de X_1^*, \dots, X_n^* , étant donné X_n , et, où nous supprimons la dépendance sur l pour faciliter la notation. Dans le cas général, où n n'est pas un multiple de l , on peut ré-échantillonner $b = b_0$ blocs, où $b_0 = \min \{k \geq 1 : kl \geq n\}$ et on conserve la première valeur de données ré-échantillonnée pour définir les répliques bootstrap de T_n .

3.3 Séries temporelles autorégressives : bootstrap des résidus

3.3.1 Séries temporelles AR(p) stationnaires

$$X_t = \beta_1 X_{t-1} + \dots + \beta_p X_{t-p} + \varepsilon_t, \text{ où}$$

1. $\{\varepsilon_t, t \in \mathbb{Z}\}$ est une suite de v.a. i.i.d. de loi P centrée,
2. $\beta = (\beta_1, \dots, \beta_p)'$ vecteur de paramètres inconnus, tels que $\beta(Z) = 1 - \sum_{j=1}^p \beta_j Z^j \neq 0$ pour tout $Z \in \mathbb{C}, |Z| < 1$.

Nous disposons d'une observation x_n de $X_n = (X_1, \dots, X_n)$.

Estimateur des moindres carrés ordinaires de β :

$$\hat{\beta} = (V_n' V_n)^{-1} V_n' (X_{p+1}, \dots, X_n)'$$

où V_n est une matrice $(n-p) \times p$ dont la $i^{\text{ème}}$ ligne est (X_{i+p-1}, \dots, X_i) .

Résidus: $\hat{\varepsilon}_t = X_t - \hat{\beta}_1 X_{t-1} + \dots + \hat{\beta}_p X_{t-p}, t = p+1, \dots, n$.

Résidus recentrés: $\tilde{\varepsilon}_t = \hat{\varepsilon}_t - \bar{\hat{\varepsilon}}, \text{ où } \bar{\hat{\varepsilon}} = \frac{1}{n-p} \sum_{t=p+1}^n \hat{\varepsilon}_t$.

Le bootstrap des résidus se résume selon le schéma suivant:

$$\begin{array}{ll}
 \beta & \longleftrightarrow \hat{\beta} \\
 P & \longleftrightarrow \tilde{P}_n = \frac{1}{n-p} \sum_{t=p+1}^n \delta_{\tilde{\varepsilon}} \\
 & \varepsilon_t^* \text{ i.i.d de loi } \tilde{P}_n \text{ conditionnellement} \\
 \varepsilon_t \text{ i.i.d. de loi } P & \longleftrightarrow \text{à } X, \text{ échantillon bootstrap associé au} \\
 & \text{vecteur des résidus recentrés.} \\
 X & \longleftrightarrow X^* \text{ solution stationnaire de} \\
 & X_t^* = \hat{\beta}_1 X_{t-1}^* + \dots + \hat{\beta}_p X_{t-p}^* + \varepsilon_t^* \\
 \hat{\beta} & \longleftrightarrow \beta^* = (V_n^{*'} V_n^*)^{-1} V_n^{*'} (X_{p+1}^*, \dots, X_n^*)' \\
 R_n(X_n, \beta) & \longleftrightarrow R_n^* = R_n((X_1^*, \dots, X_n^*), \hat{\beta})
 \end{array}$$

3.3.2 Estimation bootstrap des résidus de la loi d'une racine

La loi d'une racine $R_n(X_n, P_n)$ est estimée par la loi conditionnelle de $R_n^* = R_n((X_1^*, \dots, X_n^*), \hat{\beta})$ sachant X .

3.4 Séries temporelles à données q dépendantes: Block Bootstraps

Soit $X = \{X_t, t = 0, 1, 2, \dots\}$ une série temporelle fortement stationnaire i.e.

$$\forall k, l, \{X_1, \dots, X_k\} = \mathcal{L} \{X_{1+l}, \dots, X_{k+l}\}.$$

et q -dépendantes i.e. pour tout t , $\{\dots, X_{t-1}, X_t\}$ et $\{X_{t+q+1}, X_{t+q+2}, \dots\}$ sont indépendantes, pour un q fixé. L'idée est de bootstrapper en gardant au maximum la structure de l'échantillon d'origine.

3.4.1 Estimation BB de la loi d'une racine

Soit $k_n \rightarrow \infty$ et $l_n \rightarrow \infty, n = k_n l_n$.

Pour $j = 1 \dots k_n, B_j = \{X_{(j-1)l_n+1}, X_{(j-1)l_n+2}, \dots, X_{jl_n}\}$ (bloc de taille l_n).

Soit

$$B_1^* := (X_1^*, \dots, X_{l_n}^*), \dots, B_b^* := (X_{(b-1)l_n+1}^*, \dots, X_{bl_n}^*).$$

b blocs tirés au hasard avec remise dans $\{B_1, \dots, B_{k_n}\}$.

Soit

$$X^* = (X_1^*, \dots, X_{l_n}^*, \dots, X_{(b-1)l_n+1}^*, \dots, X_{bl_n}^*),$$

composé des $m = bl_n$ éléments de B_1^*, \dots, B_b^* mis "bout à bout".

Soit F_m^* la fonction de répartition empirique associée à X^* .

La loi de $R_n(X_n, P_n)$ est estimée par la loi de $R_n(X^*, F_m^*) \setminus X$.

Remarque 3.4.1 *Il ya toujours le problème du choix de l_n .*

Exemples d'application du bootstrap aux séries chronologiques

Introduction

Dans ce dernier chapitre nous présentons deux exemples d'application dont le premier est numérique qui sert à faire une comparaison entre les estimations de l'intervalle de confiance du paramètre d'un $AR(1)$ au niveau 95% par la méthode paramétrique Box et Jenkins et par les méthodes bootstrap des résidus, les différentes simulations sont effectuées sous le logiciel R. Et un deuxième exemple théorique qui porte sur une procédure bootstrap de test de racine unité dans un modèle $AR(1)$, basée aussi sur le ré-échantillonnage des résidus.

4.1 Estimation du paramètre β et son $I.C$

Soit le tableau de données représenté ci-dessous :

t	1	2	3	4	5	6	7	8	9	10	11	12
c_t	2.4	2.4	2.4	2.2	2.1	1.5	2.3	2.3	2.5	2.0	1.9	1.7

Tableau.2.3-Evolution de la concentration c_t d'une hormone au cours du temps t .

4.1.1 Analyse préliminaire

La série chronologique de nombres d'évolution de la concentration d'une hormone c_t au cours du temps contient 12 observations, elle ne contient ni de valeurs manquantes, ni de valeurs aberrantes. Une des premières étapes faites avant l'analyse des séries temporelles, étant l'observation de la représentation graphique, ce qui nous fournit une idée générale sur la nature et les caractéristiques du processus (tendance, saisonnalité,...).

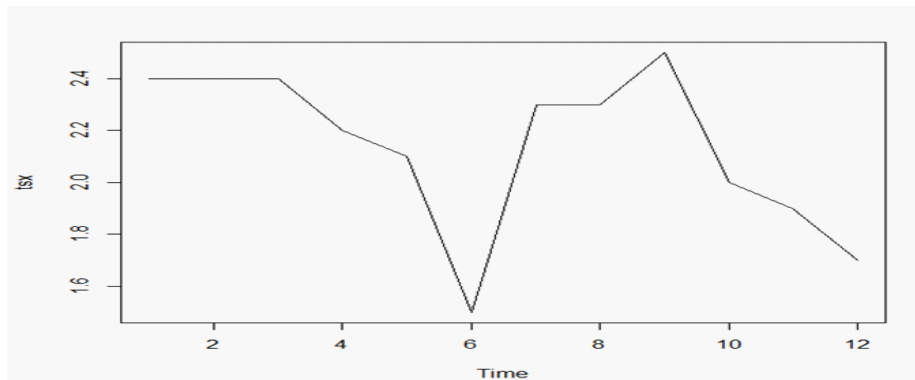
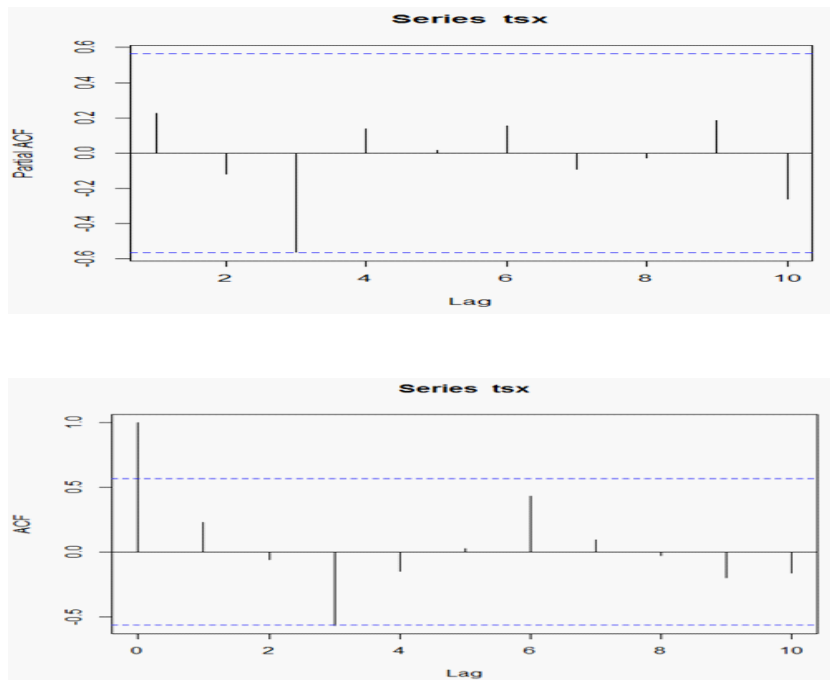


Fig-4.1 : Représentation graphique de la série originale de la concentration c_t .

Le graphe de l'évolution de la concentration d'une hormone au cours du temps ne suggère ni une tendance ni saisonnalité. Il sera alors nécessaire de vérifier si la série est stationnaire afin de choisir le modèle le plus approprié.

Le logiciel R fournit les corrélogrammes de l'autocorrélation simple et partiel suivants :

Fig-4.2 : Les autocorrélogrammes de la série c_t .

D'après les résultats précédents, on distingue que le modèle approprié pour la série chronologique est un $ARIMA(1, 0, 0)$ qui est un $AR(1)$:

$$y_t = \beta y_{t-1} + \varepsilon_t,$$

où ε_t est un BB iid.

1. La méthode de Box et Jenkins nous donne les estimations suivantes (dont nous aurons besoin par la suite):

Paramètre	paramètre estimé β	I.C au niveau 95%
$AR(1)$	0.272	$[-0.32, 0.86]$

Tableau.2.4-La valeur du paramètre β ainsi que son I.C estimés par Box et Jenkins.

Donc le modèle s'écrit comme suit :

$$y_t = 0.272y_{t-1} + \varepsilon_t,$$

tel que

$$\varepsilon_t = y_t - 0.272y_{t-1}.$$

2. La méthode de Bootstrap des résidus nous donne les estimations suivantes pour $B = 1000$ répliques séries bootstrap telles que ces répliques sont toutes des $AR(1)$ pour $b = 1, \dots, B$.

Pramètres estimés	β^{*b}	$I.C^{*b}$ au 95%
1 ^{ière} rép	0.251	[-0.35, 0.85]
2 ^{ième} rép	0.241	[-0.33, 0.80]
.	.	.
.	.	.
.	.	.
999 ^{ième} rép	0.230	[-0.31, 0.86]
1000 ^{ième} rép	0.215	[-0.41, 0.89]

Tableau.2.5-Les valeurs du paramètre β ainsi que leurs I.C estimés par Bootstrap des résidus.

Remarque 4.1.1 *Les intervalles de confiance estimés par le bootstrap au niveau 95% (voir tableau 2.5) ne sont pas différents (sont proches) de celui estimé paramétriquement (voir tableau 2.4), en tous les cas 0 est à l'intérieur de l'intervalle de confiance, mais il ya une suggestion de présence des autocorrélations positives.*

Via l'algorithme présenté en annexe et le tableau 2.5 nous déduisons que :

$$\hat{\beta}^* = \frac{\sum_{b=1}^{1000} \beta^{*b}}{B} \approx 0.26799031.$$

D'où $\hat{\beta}^* \rightsquigarrow \beta = 0.272$.

Et $I.C$ au niveau 95% de $\hat{\beta}^*$ est $[-0.32, 0.86] = I.C$ de β .

$$\begin{aligned}\varepsilon_t^{*1} &= y_t^{*1} - 0.251y_{t-1}^{*1}, \\ \varepsilon_t^{*2} &= y_t^{*2} - 0.241y_{t-1}^{*2} \\ &\cdot \\ &\cdot \\ &\cdot \\ \varepsilon_t^{*999} &= y_t^{*999} - 0.230y_{t-1}^{*999} \\ \varepsilon_t^{*1000} &= y_t^{*1000} - 0.215y_{t-1}^{*1000}\end{aligned}$$

D'où

$$\begin{aligned}y_t^{*1} &= 0.251y_{t-1}^{*1} + \varepsilon_t^{*1}, \\ y_t^{*2} &= 0.241y_{t-1}^{*2} + \varepsilon_t^{*2}, \\ &\cdot \\ &\cdot \\ &\cdot \\ y_t^{*999} &= 0.230y_{t-1}^{*999} + \varepsilon_t^{*999}, \\ y_t^{*1000} &= 0.215y_{t-1}^{*1000} + \varepsilon_t^{*1000}.\end{aligned}$$

4.2 Exemple 2 (théorique) Un test Bootstrap dans un modèle AR(1)

L'objet de cette section est de présenter et démontrer la validité asymptotique d'une procédure de type Bootstrap permettant de tester la valeur du paramètre β de régression dans un modèle autorégressif d'ordre 1 ($AR(1)$).

Considérons le problème de test

$$H_0 : \beta = \beta_0 \text{ contre } H_1 : \beta \neq \beta_0, \tag{4.2.1}$$

dans le modèle $AR(1)$:

$$y_t = \beta y_{t-1} + \varepsilon_t, \quad t = 1 \dots T, \quad (4.2.2)$$

où ε_t est un bruit blanc indépendant de loi F de variance $\sigma_{\varepsilon_t}^2$.

On suppose dans la suite que le processus y_t est généré à partir d'une observation y_0 constante. On pose $c = \sigma_{\varepsilon_t}^{-2} y_0^2$. On pourrait, sans grands changements aussi dans le cas $|\beta| < 1$, supposer que y_0 est une variable aléatoire d'espérance nulle et de variance $(1 - \beta^2)^{-1} \sigma_{\varepsilon_t}^2$ afin de rendre le processus (y_t) stationnaire.

Soient $\hat{\beta}_T$ et $\hat{\sigma}_T^2$ les estimateurs des moindres carrés ordinaires:

$$\hat{\beta}_T = \left(\sum_{t=1}^T y_{t-1}^2 \right)^{-1} \sum_{t=1}^T y_t y_{t-1}, \quad (4.2.3)$$

et

$$\hat{\sigma}_T^2 = T^{-1} \sum_{t=1}^T \left(y_t - \hat{\beta}_T y_{t-1} \right)^2. \quad (4.2.4)$$

Les résidus estimés centrés dans le modèle (4.1.3)

$$\tilde{\varepsilon}_t = \hat{\varepsilon}_t - T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t, \quad (4.2.5)$$

permettent de définir une fonction de répartition empirique \tilde{F}_T dans lequel on effectue le ré-échantillonnage, suivant l'idée de Freedman [20]. On notera par ailleurs \hat{F}_T la fonction de répartition empirique des résidus $\hat{\varepsilon}_t$.

Le processus fictif bootstrap y_t^* intégrant l'hypothèse nulle $H_0 : \beta = \beta_0$ s'obtient de manière récursive:

$$\begin{cases} y_0^* = y_0, \\ y_t^* = \beta_0 y_{t-1}^* + \varepsilon_t^*, \end{cases} \quad t = 1, \dots, T, \quad \varepsilon_t^* \text{ i.i.d de loi } \tilde{F}_T, \quad (4.2.6)$$

soit encore par la relation:

$$y_t^* = \sum_{i=1}^{t-1} \beta_0^i \varepsilon_{t-i}^* + \beta_0^t y_0. \quad (4.2.7)$$

On calcule alors l'estimateur Bootstrap des moindres carrés ordinaires (m.c.o) de β sous H_0 dans le modèle (4.1.3), $\hat{\beta}_T^*$, par régression de y_t^* sur y_{t-1}^* . L'idée du Bootstrap est de considérer que la distribution de $\hat{\beta}_T^* - \beta_0$ donne une bonne approximation de la distribution de $\hat{\beta}_T - \beta_0$ sous H_0 .

Nous montrerons que cette procédure de test Bootstrap est asymptotiquement valide quelle que soit la valeur de c et du paramètre testé. Dans le cas où le bruit-blanc de l'AR(1) est gaussien. Ce test Bootstrap donne des résultats à distance finie très satisfaisants. Ces quelques résultats empiriques confirment que la distribution Bootstrap est beaucoup plus proche de la distribution exacte que ne l'est l'approximation asymptotique.

4.3 Distributions asymptotique, Bootstrap et à distance finie

Le modèle (4.1.2) a été abondamment étudié dans la littérature: voir par exemple White [45], Anderson [1], Dickey and Fuller [14] et Phillips [37] pour une synthèse et une bibliographie complète. Il est possible dans un cadre très large d'hypothèses d'obtenir la distribution asymptotique de l'estimateur des m.c.o, Lorsqu'on prend la partie principale en T dans l'expression dérivant du calcul de l'information de Fischer des estimateurs (White [45]):

$$E_F \left(\sum_{t=1}^T y_{t-1}^2 / \sigma^2 \right) = \begin{cases} \frac{T}{1-\beta^2} - \frac{1-\beta^{2T}}{(1-\beta^2)^2} + \frac{y_0^2}{\sigma^2} \left(\frac{1-\beta^{2T}}{(1-\beta^2)} \right) & \text{si } |\beta| \neq 1, \\ \frac{T(T+1)}{2} + T \frac{y_0^2}{\sigma^2} & \text{si } |\beta| = 1. \end{cases} \quad (4.3.1)$$

Suivant les valeurs prises par β , on obtient la vitesse de convergence de l'estimateur des m.c.o de sorte que l'on peut montrer (White [45]):

$$\text{si } |\beta| < 1 \quad \left(\frac{T}{1-\beta^2} \right)^{\frac{1}{2}} \left(\hat{\beta}_T - \beta \right) \rightarrow N(0, 1) \text{ quand } T \rightarrow \infty \quad (4.3.2)$$

$$\text{si } |\beta| = 1 \quad \frac{T}{\sqrt{2}} \left(\hat{\beta}_T - \beta \right) \rightarrow L_1 \text{ quand } T \rightarrow \infty \quad (4.3.3)$$

$$\text{si } |\beta| > 1 \quad \frac{\beta^T}{(\beta^2-1)} \left(\hat{\beta}_T - \beta \right) \rightarrow L_{\beta,c} \text{ quand } T \rightarrow \infty. \quad (4.3.4)$$

Dans le cas stationnaire (4.2.2), la distribution limite est gaussienne et ne dépend pas de la condition initiale (voir Anderson [1]). Pour $|\beta| = 1$, White [45] a montré que la distribution asymptotique de l'estimateur des m.c.o. centré dilaté, L_1 , de forme très dissymétrique, ne dépendait pas de la condition initiale y_0 et a donné une caractérisation de la loi limite en

terme de processus de Wiener (voir Phillips [37] pour une généralisation dans le cas d'erreurs non i.i.d.). Dans le cas explosif (4.2.4), $L_{\beta,c}$ est une loi non-dégénérée, dépendant notamment de la loi des résidus et de la condition initiale retenue par l'intermédiaire de c . Ainsi White [45] montre que dans le modèle gaussien, $L_{\beta,c}$ a pour densité :

$$f_{\beta,c} = \frac{\exp(-q)}{\pi(1+x^2)} \left(1 + \sum_{i=1}^{\infty} \frac{(4q)^i (i!)}{(1+x^2)^i (2i!)} \right), \quad (4.3.5)$$

avec $q = c(\beta^2 - 1)/2$. La distribution limite pour $|\beta| > 1$ peut donc avoir des queues épaisses ; dans le cas gaussien, avec $c = 0$, la loi limite est donc une loi de Cauchy. Il est important de noter que les paramètres c et β (si $c \neq 0$) interviennent dans la distribution limite des estimateurs dans le cas explosif.

La méthode du Bootstrap introduite par Efron [18] étant connue pour donner à distance finie des approximations plus fiables de la vraie loi que l'approximation asymptotique (voir par exemple Hall [22] pour le modèle linéaire), il paraît particulièrement intéressant de l'étudier dans ce contexte.

4.4 Validité asymptotique du bootstrap

Nous établissons dans ce paragraphe la validité asymptotique de la procédure de test dans le cas $AR(1)$ pur, quelle que soit la valeur de β_0 .

La notion de convergence utilisée ici L^* quand $T \rightarrow \infty$ est la convergence en loi conditionnelle, presque sûre dans la suite (voir Mallow [33] et la propriété A.2) est différente et plus forte que la convergence vague conditionnelle, presque sûre, utilisée par Basawa et al [2].

Théorème 4.4.1 *Soit un autorégressif pur de bruit blanc fort associé ε_t admettant des moments d'ordre 2. Pour toute valeur $\beta_\alpha \in \mathbb{R}$, si $\hat{\beta}_T^*$ est l'estimateur des moindres-carrés ordinaires Bootstrap, défini conditionnellement au processus initial, construit sous l'hypothèse $H_1 : \beta = \beta_\alpha$, on a :*

$$\begin{aligned} \text{si } |\beta_\alpha| < 1 & \quad \left(\frac{T}{1-\beta_\alpha^2} \right)^{1/2} \left(\hat{\beta}_T^* - \beta_\alpha \right) \xrightarrow{L^*} N(0, 1) & \quad \text{quand } T \rightarrow \infty, \\ \text{si } |\beta_\alpha| = 1 & \quad \frac{T}{\sqrt{2}} \left(\hat{\beta}_T^* - \beta_\alpha \right) \xrightarrow{L^*} L_1 & \quad \text{quand } T \rightarrow \infty, \\ \text{si } |\beta_\alpha| > 1 & \quad \left(\frac{\beta_\alpha^{2T}}{(1-\beta_\alpha^2)^2} \right)^{1/2} \left(\hat{\beta}_T^* - \beta_\alpha \right) \xrightarrow{L^*} L_{\beta_\alpha,c} & \quad \text{quand } T \rightarrow \infty, \end{aligned}$$

où L_1 et $L_{\beta_{\alpha,c}}$, sont définis en (4.2.3) et (4.2.4).

Démonstration. (Voir [6]) ■

Nous posons $g(T)$ égal au taux de convergence de l'estimateur des m.c.o. dans les différents cas de figure étudiés

$$\begin{aligned} \text{si } |\beta_\alpha| < 1 \quad g(T) &= \left(\frac{T}{1-\beta_\alpha^2} \right)^{1/2}, \\ \text{si } |\beta_\alpha| = 1 \quad g(T) &= \frac{T}{\sqrt{2}}, \\ \text{si } |\beta_\alpha| > 1 \quad g(T) &= \left(\frac{\beta_\alpha^{2T}}{(1-\beta_\alpha^2)^2} \right)^{1/2}. \end{aligned}$$

Le théorème précédent justifie l'utilisation des quantiles de la distribution Bootstrap pour construire une région critique pour le test (4.1.1). Ne pouvant pas déterminer pratiquement les quantiles d'ordre $\alpha/2$ et $1 - \alpha/2$ de la loi $G_{T,0}^*$ de $\hat{\beta}_T^*$ e.m.c.o. construit sous H_0 (parce que cela nécessiterait trop de calcul, en fait T^T) on est amené à donner une approximation de ces quantiles par une technique de type Monte-carlo. La mise en oeuvre du test se décompose donc ainsi :

1. ayant calculé l'estimateur des m.c.o. dans le modèle (4.1.2), on construit la fonction de répartition empirique des résidus estimés centrés \tilde{F}_T .
2. La méthode de ré-échantillonnage des résidus dans \tilde{F}_T , i.e. de façon pratique, le tirage avec remise dans l'ensemble des résidus centrés estimés, conduit à construire un processus de type (4.1.3) qui permet de calculer l'estimateur bootstrap des m.c.o. La répétition B fois de cette phase permet d'obtenir un ensemble d'estimations sous l'hypothèse nulle $\beta_{T,0}^{*(b)}$, pour $b = 1, \dots, B$, dont la fonction de répartition empirique $G_{T,0B}^*$ approxime la vraie distribution $G_{T,0}^*$ de $\hat{\beta}_T^*$ sous H_0 . On utilise alors une méthode de type percentile voir (Efron [18]) pour construire une région d'acceptation approximative au niveau α du test $H_0 : \beta = \beta_0$.
3. Soient $g_{T,0B}^*(\alpha/2)$ et $g_{T,0B}^*(1 - \alpha/2)$ respectivement les quantiles d'ordre $\alpha/2$ et $1 - \alpha/2$ de $G_{T,0B}^*$:

$$g_{T,0B}^*(\beta) = \inf \{t, G_{T,0B}^*(t) \geq \beta\}.$$

L'intervalle $[g_{T,0B}^*(\alpha/2), g_{T,0B}^*(1 - \alpha/2)]$ est une région d'acceptation.

4.5 Une étude par simulation

4.5.1 Préliminaire

Les simulations ont été mises en oeuvre sur le modèle $AR(1)$ simple avec résidus i.i.d. normaux, afin de comparer les résultats des simulations avec la théorie exposée brièvement en 4.2. On s'intéresse au test de l'hypothèse $H_0 : \beta = 1$.

Etant donnée la dissymétrie des distributions limites et des distributions à distance finie, il ne semble pas judicieux, pour un seuil α donné, de prendre des régions d'acceptation du test basées sur les quantiles d'ordre $\alpha/2$ et $(\alpha - 1)/2$, mais plus intéressant de choisir des quantiles α_1 et $1 - \alpha_2$ avec $\alpha_1 + \alpha_2 = \alpha$ de manière à obtenir une région d'acceptation la plus petite possible.

4.5.2 Simulation

Nous représentons ici l'exemple de simulation donné dans [5]

Pour $\alpha = 5\%$ nous avons choisi de construire 9 intervalles de confiance correspondant à $\alpha_1 \in \{0.5\%, 1.0\%, 1.5\%, 2.0\%, 2.5\%, 3.0\%, 3.5\%, 4.0\%, 4.5\%\}$. Le tableau ci-dessous obtenu sur une simulation donne les différentes régions d'acceptation du test bootstrap, ainsi que leur longueur $l_{T,0_B}^*(\alpha_1)$. Le nombre de ré-échantillonnage utilisé pour construire les distributions Bootstrap empiriques est $B = 3000$. On a choisi ici $\beta = 1, T = 25, \varepsilon_t \sim N(0, 1), y_0 = 0$. Le paramètre β n'est pas bien estimé par les m.c.o. dans la simulation retenue : $\hat{\beta}_T = 0.747, \hat{\sigma}_{\hat{\beta}} = 0.136$. Le test de Dickey-Fuller [14] conduit néanmoins à ne pas rejeter l'hypothèse de racine-unité. On

pourra consulter d'autres résultats de simulation dans Bertail [3].

$\alpha_1\%$	$(1 - \alpha_2)\%$	$g_{T,0_B}^*(\alpha_1)$	$g_{T,0_B}^*(1 - \alpha_2)$	$l_{T,0_B}^*(\alpha_1)$
0.5	95.5	0.468	1.057	0.589
1.0	96.0	0.520	1.061	0.540
1.5	96.5	0.572	1.065	0.493
2.0	97.0	0.612	1.017	0.459
2.5	97.5	0.650	1.076	0.426
3.0	98.0	0.668	1.078	0.410
3.5	98.5	0.678	1.085	0.407
4.0	99.0	0.689	1.091	0.402
4.5	99.5	0.702	1.098	0.396

Tableau.4.5-différentes régions d'acceptation du test bootstrap, ainsi que leur longueur

$$l_{T,0_B}^*(\alpha_1).$$

On notera que dans tous les cas (même pour $\alpha_1 = 4.5\%$) on accepte l'hypothèse nulle. La longueur de l'intervalle décroît lorsque α_1 augmente, ce qui invite à prendre des régions d'acceptation plus centrées autour de la valeur 1. Il convient néanmoins dans les applications de ne pas choisir $\alpha_1 = 5\%$, à cause de la nature discrète de la distribution empirique dans laquelle s'effectue le ré-échantillonnage; on risquerait d'obtenir une valeur aberrante pour la borne supérieure de l'intervalle de confiance dans un problème de test bilatéral.

Les problèmes liés à la multiplicité des granicules disparaissent si, au lieu d'utiliser la fonction de répartition empirique, on utilise un estimateur à noyau de F :

$$\tilde{F}_T(x) = \frac{1}{T} \sum_{t=1}^T K\left(\frac{x - \tilde{\varepsilon}_t}{h}\right) \quad (4.5.1)$$

où K est le noyau intégré (par exemple $K(y) = (2\pi)^{-1/2} \int_{-\infty}^y \exp(-\mu^2/2)$) et h est le paramètre de lissage. Il est clair que la procédure de test et la démonstration du théorème fonctionnent toujours dans un tel cadre pourvu que la constante de lissage soit bien choisie. Cependant, quelques simulations effectuées dans le cas du modèle $AR(1)$ ont permis de mettre en évidence l'importance du paramètre de lissage h dans l'obtention de la distribution Bootstrap de

l'estimateur des m.c.o.: les résultats peuvent être considérablement différents pour de petites variations de h .

4.5.3 Etude des distributions Bootstrap

Les graphiques 4.3, 4.4, 4.5 et 4.6 (d'après Bertail [5]) donnent la comparaison des estimations de la densité Bootstrap, centrée dilatée de l'estimateur des m.c.o., après lissage (le paramètre de lissage choisi est celui qui minimise l'écart-carré moyen) sous l'hypothèse nulle $H_0 : \beta = 1$ pour $T = 100$ (et $T = 25$ pour $\beta_0 = 1$ et $\beta_0 = 0.95$), de la distribution asymptotique et de la distribution à distance finie (estimée à partir de 10000 tirages), pour deux valeurs différentes de c . Pour tous les graphiques, la légende est la suivante :

<i>Trait</i>	Distribution Bootstrap
- - -	Distribution Asymptotique L_1
.....	Distribution Exacte.

Dans tous les cas, la distribution Bootstrap est très proche de la distribution à distance finie et nettement meilleure que l'approximation asymptotique. Pour $\beta = 1$ et $c = 0$ (graphique 4.3), la distribution Bootstrap et la distribution limite, lorsque T augmente, sont relativement proches. Pour $c = 10$, même avec une taille d'échantillon grande $T = 100$ (compte tenu de la convergence en T^1), l'asymptotique rend mal compte du comportement à distance finie contrairement à la distribution Bootstrap qui donne une très bonne approximation de la vraie distribution. Dans le cas $\beta_0 = 0.95$, l'approximation gaussienne est d'autant plus mauvaise que T est petit et c proche de 0. La distribution Bootstrap permet bien de prendre en compte l'impact de la condition initiale sur la forme de la distribution. Pour des valeurs $|\beta| > 1$ très proches de 1 ($\beta = 1.01$) on peut formuler des remarques semblables même pour une taille d'échantillonnage assez élevée $T = 100$.

Lorsque β s'éloigne de 1 (graphique 4.5), les distributions exactes, Bootstrap et asymptotiques (construites pour $c = 10$ à partir de la relation (4.2.5), la somme infinie étant tronquée à $i = 10$) sont pratiquement confondues : rappelons que dans ce cas la vitesse de convergence est en $(\beta^2 - 1) / \beta_0^{-T}$ (ce qui ici donne une erreur de l'asymptotique sur la vraie distribution de l'ordre 10^{-5}).

Dans tous les cas, les tests usuels (test de student et test du signe) rejettent l'hypothèse de nullité de la moyenne de ces distributions ce qui signifie que l'estimateur choisi est biaisé. Pour $T = c = 25$, si le nombre de ré-échantillonnage est important ($B > 3000$) la méthode du Bootstrap donne des ordres de grandeurs semblables à ceux obtenus par calcul avec l'approximation de White [44] comme le montre le tableau ci-dessous.

β_0	0.9	0.95	0.975	0.99	1.0
Approximation du biais(White[44])	-0.066	-0.070	-0.072	-0.073	-0.074
Méthode du Bootstrap $B = 3000$	-0.059	-0.068	-0.067	-0.069	-0.071

Tableau.4.3-Approximation du biais par (White[45]) et par bootstrap pour $B = 3000$.

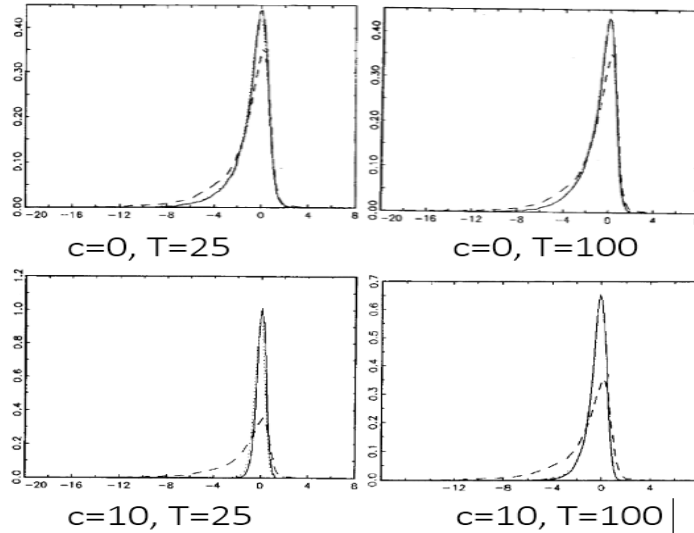


Fig-4.3 :Graphiques de comparaison des densités exactes, Bootstrap et asymptotiques, $\beta_0 = 1$.

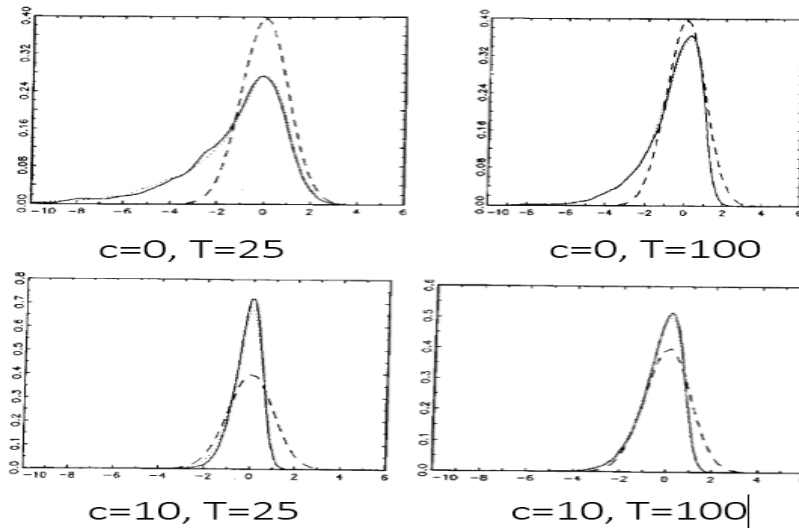


Fig-4.4 :Graphique de comparaison des densités exactes, Bootstrap et asymptotiques, $\beta_0 = 0.95$.

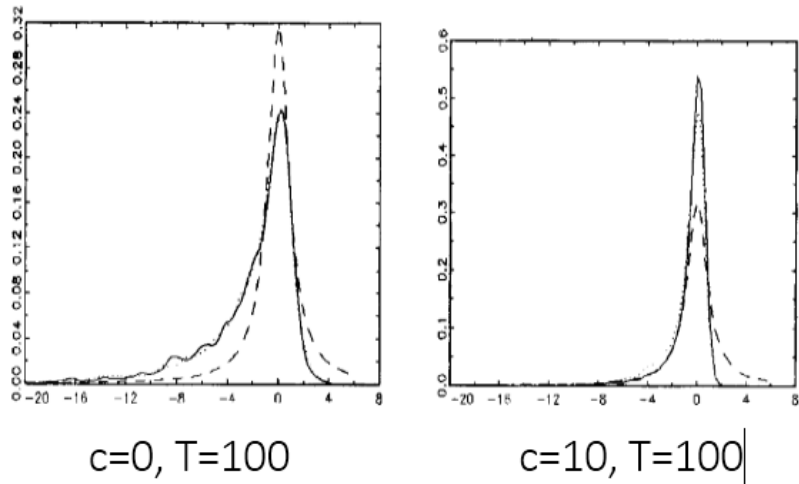


Fig-4.5 :Graphique de comparaison des densités exactes, Bootstrap et asymptotiques, $\beta_0 = 1.01$.

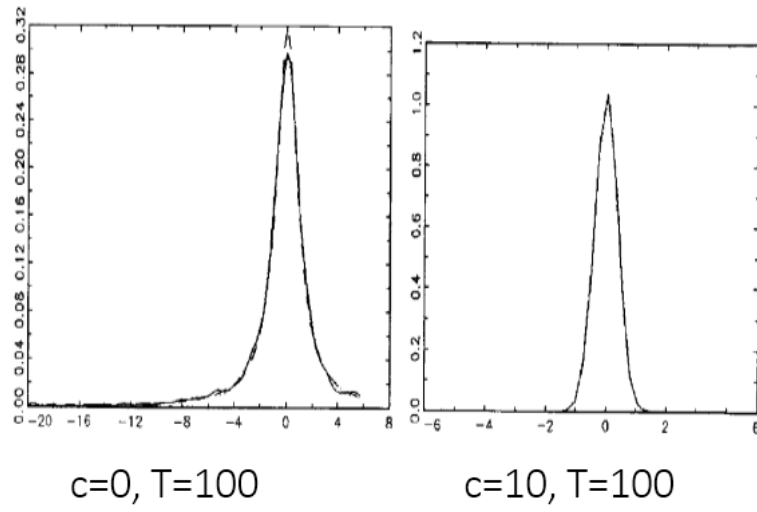


Fig-4.6 :Graphique de comparaison des densités exactes,
Bootstrap et asymptotiques, $\beta_0 = 1.1$.

Conclusion Générale

L'objectif de ce travail a été d'introduire l'essentiel de la méthode du bootstrap sur les séries chronologiques, on a montré dans ce mémoire qu'une méthode de ré-échantillonnage basée sur le bootstrap des résidus était adaptable aux modèles autorégressifs d'ordre 1 choisis et qui doivent vérifier la condition de la stationnarité et de l'indépendance des erreurs ($\forall t, \varepsilon_t \rightsquigarrow \text{i.i.d.}$).

Les estimations et les simulations effectuées dans le dernier chapitre confirment le bien-fondé théorique de la méthode et mettent en évidence l'apport du bootstrap dans des problèmes d'estimation et de test de racine unité dans un modèle $AR(1)$. Dans ce contexte on est arrivé à déduire que les paramètres estimés par la méthode bootstrap convergent vers ceux estimés par des méthodes classiques (Box et Jenkins) et que Les distributions bootstrap rendent beaucoup mieux compte du comportement à distance finie des estimateurs. Cette étude peut donner lieu à divers prolongements et applications. On peut par ailleurs s'interroger sur une généralisation naturelle des méthodes de bootstrap des séries chronologiques à des modèles et des problèmes plus complexes par exemples : comment mettre en œuvre la procédure bootstrap si les erreurs ne sont pas i.i.d ? Et comment montrer la validité au second ordre du Bootstrap dans le cas de limite non gaussienne?. La réponse à ces questions nous amène à l'utilisation des autres outils plus compliqués comme le (MBB). Bien que les méthodes de bootstrap prometteuses pour les séries temporelles soient disponibles, il ya toujours un besoin considérable de recherche supplémentaire dans leur application. L'investigation sur le bootstrap des séries chronologiques reste d'actualité et offre beaucoup d'opportunités de recherche à titre d'exemple la recherche sur le bootstrap par blocs des séries temporelles (le problème de choix de la longueur des blocs).

Annexe

Estimation de IC en utilisant la méthode de Box et Jenkins :

```
>require('boot')
>fit1
>plot(x,type='l')
>head(x)#prendre une partie de la série(un échantillon).
[1] 2.4 2.4 2.4 2.2 2.1 1.5
>tsx=ts(x)
>a1=acf(tsx)
>a1
Autocorrelations of series 'tsx', by lag
0 1 2 3 4 5 6 7 8 9 10
1.000 0.228 -0.060 -0.568 -0.148 0.026 0.430 0.095 -0.030 -0.199 -0.165
>p1=pacf(tsx)
>p1
Partial autocorrelations of series 'tsx', by lag
1 2 3 4 5 6 7 8 9 10
0.228 -0.118 -0.563 0.138 0.018 0.155 -0.090 -0.027 0.185 -0.262
>plot(a1)
>plot(p1)
>plot(tsx)
>fit1=arima(tsx,order=c(1,0,0))
>fit1
```

Call:

```
arima(x = tsx, order = c(1, 0, 0))
```

Coefficients:

```
ar1 intercept
```

```
0.2720 2.1363
```

```
s.e. 0.3023 0.1111
```

```
sigma^2 estimated as 0.08303: log likelihood = -2.13, aic = 10.27
```

```
>confint(fit1)
```

```
2.5 % 97.5 %
```

```
ar1 -0.3205672 0.8644831
```

```
intercept 1.9184930 2.3540952
```

```
#l'intervalle de confiance de coefficient d'un AR(1) est [-0.321,0.864].
```

Estimation de IC par la méthode du bootstrap des résidus :

```
>ar1=ar(tsx,order.max=1)#la série est défini comme un autorégressif d'ordre 1.
```

```
>armodel=list(order=c(1,0,0),ar=ar1$ar)
```

```
>arres=ar1$resid #extraire les résidus de l'AR(1).
```

```
>arres=arres-mean(arres)#permet d'estimer (calculer) les résidus centrés.
```

```
>arres
```

Time Series:

```
Start = 1
```

```
End = 12
```

```
Frequency = 1
```

```
[1] 0.25833333 0.25833333 0.25833333 0.05833333 -0.04166667 -0.64166667
```

```
[7] 0.15833333 0.15833333 0.35833333 -0.14166667 -0.24166667 -0.44166667
```

```
>bootf=function(tsb)# est une fonction pour les séries temporelles bootstrap.
```

```
>fit=arima(tsb,order=c(1,0,0))
```

```
>bootsim=function(res,n.sim,ran.args)
```

```
>boot1=tsboot(tsx, bootf, R=1000, sim="model", orig.t=FALSE, ran.gen=bootsim) ran.args=list(ts=
model=armodel))#{la génération aléatoire des replication séries en utilisant arima.sim.
```

```
> mean(boot1$t[,1])#estimation du paramètre de l'AR(1)bootstrap1
[1] 0.2512331
> quantile(boot1$t[,1], probs = c(0.025,0.975)) #Efron percentile CI on
coefficient
2.5% 97.5%
- 0.3541095 0.85172895
#l'intervalle de confiance de paramètre de la première réplification AR(1) est: [- 0.35,0.85].
> mean(boot1$t[,2])#estimation du paramètre de l'AR(1)bootstrap2
[2] 0.2411731
> quantile(boot1$t[,2], probs = c(0.025,0.975)) #Efron percentile CI on
coefficient
2.5% 97.5%
- 0.3250059 0.79871598
> mean(boot1$t[,100])#estimation du paramètre de l'AR(1)bootstrap1
[100] 0.2303846
> quantile(boot1$t[,100], probs = c(0.025,0.975)) #Efron percentile CI on
coefficient
2.5% 97.5%
- 0.3436095 0.86100831
> mean(boot1$t[,1000])#estimation du paramètre de l'AR(1)bootstrap1000.
[1000] 0.214510391
> quantile(boot1$t[,1000], probs = c(0.025,0.975)) #Efron percentile CI on
coefficient
2.5% 97.5%
- 0.41910452 0.88102781
> mean(mean(boot1$t[,1:1000]))#estimation de la moyenne des paramètres.
[1:1000] 0.26799031
```

Bibliographie

- [1] Anderson, T. W. On the Asymptotic Distribution of Estimates of parameters of Stochastic Difference Equations, *Ann. Math. Statist.* 30, pp. 676-687, 1959.
- [2] Basawa, I. V., Mallik, A. K., McCormick, W. P. Taylor, R. L. Bootstrapping in Explosive Autoregressive Processes, *Ann. Math. Statist.* 19, pp. 1098-1101, 1991.
- [3] Berkowits, J. Kilian, L. Recent developments in bootstrapping time series. *Econom. Rev.*, 19, 1-48, 2000.
- [4] Bertail, P. Idées de bases de la méthode Bootstrap (origine, algorithme, intervalles de confiance, propriétés), extensions dans le cas dépendant. Institut National de la Recherche Agronomique, 2003.
- [5] Bertail, P. Un test Bootstrap dans un modèle AR(1). *Annales d'Economie et de Statistique*-N°36, Paris, 1994.
- [6] Bertail, P. La méthode du Bootstrap : quelques applications et résultats théoriques, Thèse doctorat Université Paris IX, 1992.
- [7] Borgnat, P. CNRS, équipe Sisyphe (signaux, systèmes et physique), laboratoire de physique de l'ENS de Lyon, université de Lyon. 2002.
- [8] Box, G.E.P. and Jenkins, G.M. *Time series analysis, forecasting and control*. Holden day, San francisco, 1976.
- [9] Bruffaerts, C. *Une introduction au monde du bootstrap*, European center for Advanced research in economics and statistics, 2013.

- [10] Buhlmann, P. Bootstrap for time series. *Stat. Sci*, 17, 52-72, 2002.
- [11] Buvat, I. Introduction à L'approche bootstrap. U494 INSERM, 2000.
- [12] Chernick, M R. An introduction to bootstrap methods with application to R/ Michael R. Chernick. Robert A. labored. P. cm. Printed in the United States of America.2011.
- [13] Daudin, J.J, Duby, C, Robin. S and Trécourt. P. Analyse de séries chronologiques. INA-PG, Mathématiques, 1996.
- [14] Dickey, D. A., Fuller, W. F. Likelihood Ratio for Autoregressive Time Series with Unit Roots. *Econometrica*, 49, pp. 1057-1072, 1981.
- [15] Dimitris N. Politis. The impact of bootstrap methods on time series analyses. Institute of mathematical statistics, 2003.
- [16] Djaout, A. Modélisation empirique des séries chronologiques. Faculté des sciences Exactes département de Mathématiques, Université Mentoutri Constantine, 2011.
- [17] Efron, B and Tibshirani, R. An introduction to the Bootstrap. Chapman and Hall. New York, 1993.
- [18] Efron, B. Bootstrap methods: an other look at the jackknife .*Ann. stat*, 7, 1-26, 1979.
- [19] Fortier, S. Les modèles MA, AR et ARMA multidimensionnels estimation et causalité. Université de Sherbrook, 2011.
- [20] Freedman, D. Bootstrapping Regression Models. *Ann. Statist.*, 9, pp. 1218-1228, 1981.
- [21] Fromont, M and Vimond, M. Bootstrap et ré-échantillonnage. CREST (Ensaï) - IRMAR (Université Européenne de Bretagne), 2012.
- [22] Hall, P. The bootstrap and Edgeworth expansion. New York: Springer, 352 p, 1992.
- [23] Hamdi, F and Guerbyenne, H. Analyse des séries chronologiques. Mémoire pour l'obtention du diplôme de MAGISTER en Mathématiques. Faculté de Mathématiques, U.S.T.H.B, Algérie, 2011.

-
- [24] Kalman, R. E and R. S. Bucy. New results in linear filtering and prediction theory. J. Basic Engrg, 1961.
- [25] Kalman, R. E. A new approach to linear filtering and prediction problems theory. J. Basic engrg, 1961.
- [26] Kolmogoroff, A. Stationary sequences in Hilbert space. Bull. Math. univ. Moscow, 1941.
- [27] Kreiss, J.P and Lahiri, S.N. Bootstrap methods for time series. Technische universität braunschweig, Germany, 2012.
- [28] Kreiss, J.P. Bootstrap Methods For Time Series. Institute for mathematical statistics Technical University of Braun Schweig. Germany,2011.
- [29] Kunsch, H.R. The jackknife and the bootstrap for general stationary observations. Ann. Statist. 17, 1217-1241, 1986.
- [30] Lagnoux, A. Séries Chronologiques. Université de Toulouse Le Mirail, 2003.
- [31] Lahiri, S. N. Resampling Methods for Dependent Data. New York: Springer, 2003.
- [32] Li, H. Maddala, G. S. bootstrapping time series models. Econom. Rev, 15, 115-158, 1996.
- [33] Mallows, C. L. A Note on Asymptotic joint Normality. Ann. Statist. 43, pp. 508-515, 1972.
- [34] Olatayo, T. O. On the Application of Bootstrap Method to Stationary Time Series Process. Mathematical Sciences Department, Olabisi Onabanjo University, Ago-Iwoye, Nigeria, 2012.
- [35] Pallier, C and Lalanne, C. Introduction à l'analyse des statistiques des données : travaux pratiques avec le logiciel 'R', 2005.
- [36] Palm, R. Utilisation du bootstrap pour les problèmes statistiques liés à l'estimation des paramètres. Biotechnol. Agron. Soc. Environ, 2002.
- [37] Phillips, P. C. B. Towards a Unified Asymptotic Theory of Autoregression. Biometrika, 74, pp. 535-547, 1987.

-
- [38] Politis, D. N. The impact of bootstrap methods on time series analysis. *Stat. Sci.* 18, 219-230, 2003.
- [39] Prenat, M. *Séries chronologiques volume 1:cours et exercices*. Université Paris-sud, Master Ingénieurs Mathématiques, 2010/2011.
- [40] Rapacchi, B. *Une Introduction au Bootstrap*. Centre Interuniversitaire de calcul de Grenoble. 1994.
- [41] Richard, M. V. *The moving blocks bootstrap versus parametric time series models*. Department of Civil and Environmental Engineering, Tufts University, 1996.
- [42] Singh, K. On the asymptotic accuracy of Efron's bootstrap. *Ann. Stat.* 9, 1187-1195, 1975.
- [43] Somnath, D. *Limit theory and bootstrap for explosive and partially explosive autoregression*. Department of Statistics, University of Georgia, Athens, GA 30602, 1994.
- [44] White, J. S. Asymptotic Expansions for the mean and variance of the serial correlation coefficient, *Biometrika*, 48, pp. 85-94, 1961.
- [45] White, J. S. The limiting Distribution of the Serial Correlation Coefficient in the explosive Case. *Ann. Math. Statist.* 29, pp. 1188-1197, 1958.
- [46] Wiener, N. *Extrapolation, interpolation and smoothing of stationary time series*, Wiley, New York, 1949.
- [47] Wold, H. O. *A study in the analysis of stationary time series*, Almqvist and Wiksell, Uppsala, Sweden, 1938 (2nd. Ed. 1954).
- [48] Yule, G.U. On a method of investigating periodicities in disturbed series, with special reference to wolfor's sunspot numbers, 1927.

Résumé

Le bootstrap est une méthode pour estimer la distribution d'un estimateur en échantillonnant ses données ou un modèle construit à partir des données.

Les méthodes disponibles pour mettre en œuvre le bootstrap et la précision des estimateurs de bootstrap dépendent du fait que les données proviennent d'un échantillon aléatoire indépendant ou d'une série temporelle.

Ce mémoire concerne l'application du bootstrap aux données des séries temporelles. Nous examinons les méthodes qui ont été proposées pour mettre en œuvre le bootstrap dans cette situation et discutons la précision de ces méthodes comparativement à celle des approximations asymptotiques de premier ordre. Nous montrons que les méthodes pour mettre en œuvre le bootstrap avec les données des séries temporelles ne sont pas aussi bien comprises que les méthodes pour les données des échantillons aléatoires indépendants. Bien que des méthodes de bootstrap prometteuses pour les séries temporelles soient disponibles, il ya un besoin considérable de recherche supplémentaire dans leur application.

Abstract

The bootstrap is a method for estimating the distribution of an estimator or a model estimated from the data. The methods that are available for implementing the bootstrap and the accuracy of bootstrap estimates depend on whether the data are a random sample from a distribution or a time series. This paper is concerned with the application of the bootstrap to time-series data. We review the methods that have been proposed for implementing the bootstrap in this situation and discuss the accuracy of these methods relative to that of first-order asymptotic approximations. We argue that methods for implementing the bootstrap with time-series data are not as well understood as methods for data that are sampled randomly from a distribution. Moreover, the performance of the bootstrap as measured by the rate of convergence of estimation errors tends to be poorer with time series than with random samples. This is an important problem for applied research because first-order asymptotic approximations are often inaccurate and misleading with time-series data and samples of the sizes encountered in applications. We conclude that there is a need for further research in the application of the bootstrap to time series.