

TP 3 (durée : 3h)

Préliminaires

Créer un fichier texte dans lequel vous répondrez clairement aux questions ci-dessous, en incluant vos codes R, les résultats obtenus sous R (graphique y compris), vos interprétations, remarques ... Une fois ce TP fini, vous mettrez en forme votre compte-rendu et l'exporterez au format pdf. C'est ce qui sera demandé au partiel.

1 Données AirPassengers

Nous étudions la série chronologique du nombre de passagers par mois (en milliers) dans les transports aériens, de 1949 à 1960. Cette série est disponible sous R (**AirPassengers**).

1. Estimation paramétrique de la tendance
 - (a) Représenter graphiquement la série. Ce processus vous semble-t-il stationnaire ? Présente-t-il des tendances et saisonnalités ?
 - (b) Estimer les paramètres d'une tendance linéaire $at + b$ par la méthode des moindres carrés.
 - (c) Supprimer cette tendance et représenter graphiquement la série ainsi obtenue. Vérifier que la série des résidus est de moyenne nulle.
 - (d) Calculer et représenter l'auto-corrélation de la série des résidus.
2. Méthode des différences
 - (a) Appliquer la méthode des différences pour enlever la tendance et la saisonnalité. Préciser la période de la saisonnalité, le degré du polynôme de tendance.
 - (b) La série obtenue semble-t-elle stationnaire ?
3. Méthode des moyennes mobiles
 - (a) Appliquer la méthode des moyennes mobiles pour enlever la tendance et la saisonnalité.
 - (b) La série obtenue semble-t-elle stationnaire ? Pourrait-on la modéliser par un bruit blanc ?
Remarque : la fonction `Box.test` permet de tester l'indépendance d'une suite de variables aléatoires.

2 Données simulées

Récupérer la série temporelle contenue dans le fichier <http://math.unice.fr/~rubentha/enseignement/simulation.dat>. Cette série a été simulée à partir d'un processus relativement simple. Essayer de le retrouver !

1. Analyser qualitativement cette série temporelle.
2. Pouvez-vous proposer une modélisation pour cette série (c'est-à-dire définir le processus X_t qui a généré cette série) ? Indication : la partie saisonnière est de la forme $a \cos(t\pi/b)$ ou $a \sin(t\pi/b)$... à vous de deviner !

Appendice : mise en œuvre sous R

Les indications concernent le premier exercice (on réutilisera les mêmes commandes pour l'exercice suivant). Nous mentionnons ici des commandes utiles dans les questions où on en a besoin pour la première fois.

La commande `as.numeric` transforme une série temporelle en vecteur (ce qui est utile pour certaines opérations).

Question 1d. Utiliser la commande `acf(...)` pour calculer les autocorrélations empiriques (voir corrigé du premier TP).

Question 2a. La fonction `diff.ts(serie, lag=T, difference=k)` (ou `diff` tout court) permet d'appliquer l'opérateur de différenciation Δ_T^k .

Question 3a. La fonction `decompose` permet d'extraire d'une série temporelle (via la méthode de la moyenne mobile)

(`serie_decomp<-decompose(serie, type=c("additive", "multiplicative"))`, choisir "additive", "multiplicative" mais pas les deux en même temps)

- la composante saisonnière : `serie_decomp$seasonal`, que l'on suppose additive ou multiplicative dans l'option `type`,
- la tendance : `serie_decomp$trend`,
- la partie aléatoire stationnaire de la série : `serie_decomp$random`.

Question 3b. La fonction `Box.test(serie, lag=H)` examine l'hypothèse nulle de nullité des H premières auto-covariances, à l'aide du test du portemanteau. Par défaut H est fixé à 1, et seule la nullité de l'auto-covariance d'ordre 1 est testée. Pour tester si la série peut-être apparentée à un bruit blanc, nous fixerons arbitrairement un H de l'ordre de 20 (nous considérerons abusivement que si les 20 premières auto-corrélations sont nulles, la série est indépendante). La première valeur affichée est la statistique du test. La p -value est la probabilité d'obtenir une valeur aussi élevée sous l'hypothèse nulle. Si nous voulons faire un test au niveau de confiance 0,95 : dans le cas où cette valeur est $< 0,05$, nous rejetons l'hypothèse nulle.