

Sujet de TP 3

☞ On utilisera le script R fourni sur Campus.

Partie 1 (étude et prévision de la série Airline, ex. 1 du poly, méthode de Box&Jenkins)

Dérouler la méthodologie de B&J et retrouver le modèle sur la série $\{X_t\}$ en question :

$$Y_t = (I - B^{12})(I - B)\log(X_t) \text{ et } Y_t = (I + \theta_1 B)(I + \theta_{12} B^{12})Z_t \text{ avec } Z_t \sim WN(0, \sigma^2)$$

En abrégé, $\log(X_t) \sim \text{SARIMA}(p = 0, d = 1, q = 1)(P = 0, D = 1, Q = 1)_{s=12}$

Utilisation du modèle pour faire de la prévision : on enlève les 12 dernières valeurs que l'on cherche ensuite à prédire (back-testing)

Insérer ici la figure obtenue à la fin de cette partie (« Travail à faire » dans le script) qui montre sur la série initiale (et pas son log) la qualité de la prévision obtenue par back-testing...

Modèle linéaire pour tendance et saisonnalité et comparaison avec B&J :

Donner l'expression du modèle linéaire utilisé pour capter la tendance et insérer la figure montrant le log de la série, sa tendance et sa saisonnalité puis une figure avec la série désaisonnalisée.

Insérer la figure finale montrant la prévision obtenue sur la série initiale à partir de cette approche...

Commentaires.

Partie 2 (analyse de données de température en France)

On s'intéresse à l'évolution de la température moyenne en France métropolitaine sur la période 1900-2015 (source : Météo-France) – Pour cela, on considère la série qui donne l'écart de température (en °C) avec la moyenne calculée sur la période de référence 1961-1990.

Etudier cette série (fichier `ecart.txt` disponible sur Campus) avec comme objectif de répondre à la question du réchauffement en France ces dernières décennies. **On se demandera en particulier si le réchauffement observé est statistiquement significatif et non pas simplement le fait du hasard.**

Partie 3 (analyse d'une série au choix) A chaque fois, on donnera les principales étapes de l'analyse : **choix du modèle à partir des ACF et PACF, qualité de l'ajustement et prévision obtenue par back-testing**. Commentaires. On utilisera le package « datasets » de R (sous R, exécuter la commande `library(help="datasets")` pour plus d'information sur les jeux de données disponibles).

Etudier la série **UKDriverDeaths** (Road Casualties in Great Britain 1969-84). Etudier enfin une série de votre choix.

A rendre via Campus en convertissant votre fichier en .PDF (nom de fichier du type CrTP3VosNoms.pdf). Faire un compte-rendu synthétique (pas de code demandé) expliquant la démarche ainsi que l'analyse de vos résultats.