

Projet de Séries temporelles

Contexte

On souhaite modéliser les données de surveillance concernant le nombre de cas de décès confirmés dus à une infection dans un district sanitaire donné. Les données sont recueillies mensuellement de l'année 2011 à l'année 2017. Elles sont présentées dans la base «Base_Projet_ST_M1SID2019 ».

L'objectif de cette étude consiste d'abord à étudier la tendance de la « Cas_deces », ensuite prédire le nombre de cas de décès pour les années à venir.

Travail à faire

1/ Décomposer la série « Cas_deces » de la base « Base_Projet_ST_M1SID2019 » et faire une description et un ajustement de la tendance en utilisant un modèle linéaire simple.

2/ Modéliser la série « Cas_deces » de la base «Base_Projet_ST_M1SID2019 » par la méthodologie de Box & Jenkins. Analyser et interpréter les sorties du logiciel R.

3/ Effectuer une prévision par ARIMA du nombre de cas de décès confirmés pour une période d'une année et une période de 3 années. Représenter graphiquement ces prévisions dans deux graphes différents avec leurs intervalles de confiance. Commenter.

4/ Effectuer une prévision pour 3 années par lissage exponentiel et comparer graphiquement ces valeurs avec celles obtenues par la méthode ARIMA.

NB : Pour la validation par apprentissage, vous décomposez la base « Cas_décès » en deux bases :

La base 1 sera composée des données de l'année 2011 à 2015,

La base 2 sera composée des données des années 2016 et 2017.

a) Lancez le modèle trouvé à la question 2 sur les données de la base 1.

b) Faites une prévision pour les années 2016 et 2017 en utilisant le modèle ARIMA et le lissage exponentiel.

c) Représentez dans deux graphes différents ces prévisions de 2016 et 2017 et les vraies valeurs observées de 2016 et 2017 (base 2).

Grappe 1 : prévisions par ARIMA et base 2,

Grappe 2 : prévisions par lissage et base 2.

Commenter ces graphes et tirer une conclusion.

A envoyer à l'adresse : aba.diop@uadb.edu.sn au plus tard le samedi 30 novembre 2019