

# Projet de séries temporelles

**Date de rendu :** 19 mai 2019 par courriel à votre chargé de TD.

**Consigne générale :** Ce projet doit être réalisé sous Stata, SAS, R ou Python. La notation tiendra compte de la rigueur dans la mise en œuvre des outils économétriques, de la concision et de la clarté de la présentation des résultats. Le rapport (6 pages de contenu maximum hors annexe) fera figurer les programmes en annexe. Ce tutorat doit être effectué par binôme d'un même groupe de TD.

Le barème (à titre indicatif) est le suivant : Partie 1 : 20%; Partie 2 : 40%; Partie 3 : 40%. Bon travail !

On s'intéresse à la modélisation et la prévision de l'indice de production industrielle observé en France. Il ne vous est pas demandé de connaître le processus de construction de l'indicateur. Vous ne travaillez que sur les données observées. A partir du répertoire des séries chronologiques de l'INSEE, [www.insee.fr/fr/statistiques?debut=0&categorie=10](http://www.insee.fr/fr/statistiques?debut=0&categorie=10), vous devez choisir une série agrégée brute, mensuelle, correspondant à n'importe quel secteur d'activité (à votre convenance) contenant au moins 25 années et **contenant le mois de janvier 2019**.

## Partie 1 : Les données

- 1) Que représente la série choisie ? (secteur, périmètre, traitements éventuels, transformation logarithmique, ...)
- 2) Transformer si besoin la série pour la rendre stationnaire (dessaisonnalisation, différenciation, suppression de la tendance déterministe, ..). Justifier soigneusement vos choix.
- 3) Représenter graphiquement la série choisie avant et après transformation.

## Partie 2 : Modèles ARMA

- 4) Choisir, en le justifiant, un modèle ARMA(p,q) (avec éventuellement une composante saisonnière) pour votre série corrigée  $X_t$ . Estimer les paramètres du modèle et vérifier sa validité.

## Partie 3 : Prévision

On note  $T$  la longueur de la série. On suppose que les résidus de la série sont gaussiens.

- 5) Ecrire l'équation vérifiée par la région de confiance de niveau  $\alpha$  sur les valeurs futures ( $X_{T+1}, X_{T+2}$ ).
- 6) Préciser les hypothèses utilisées pour obtenir cette région.
- 7) Déterminer graphiquement cette région pour  $\alpha=95\%$ . Commenter.
- 8) Question ouverte : soit  $Y_t$  une série stationnaire disponible de  $t=1$  à  $T$ . On suppose que  $Y_{T+1}$  est disponible plus rapidement que  $X_{T+1}$ . A quelles conditions cette information permet-elle d'améliorer la prévision de  $X_{T+1}$  ?