

Nombre de naissances par mois

Télécharger la série temporelle `NbNaissancesFrance.csv` du nombre de naissances par mois en France depuis janvier 1975. On cherche un modèle permettant de prédire le nombre de naissances mensuelles jusqu'à 2020.

Approche non paramétrique

1. Tracer la série temporelle des naissances et du log des naissances. Commenter les graphiques.
2. Proposer un lissage exponentiel de la série en justifiant les choix réalisés. Donner les paramètres obtenus et discuter leur valeur.
3. Proposer une méthode de validation du modèle, la mettre en oeuvre et discuter les résultats.
4. Prédire la série jusqu'en décembre 2019. Ne pas oublier l'intervalle de prédiction.

Modèle paramétrique

On cherche un modèle paramétrique pour la série du nombre de naissances mensuel. On peut retirer la dernière observation de façon à travailler avec un nombre entier d'années. On propose les étapes suivantes.

1. *Proposer et ajuster un modèle de tendance.*
On suggère d'utiliser un modèle polynomial et/ou un modèle polynomial trigonométrique (ex : $y(t) = \alpha_{1c} \cos(2\pi t/T) + \alpha_{1s} \sin(2\pi t/T) + \alpha_{2c} \cos(2\pi t/2T) + \alpha_{2s} \sin(2\pi t/2T) + \dots + \epsilon$ avec ϵ un bruit centré et $T = 32$ ou $T = 25$.
Tracer la série sans tendance.
Ecrire l'expression mathématique du modèle ajusté.
Remarque : sous **R**, la fonction **step** permet de faire de la sélection de variables par la méthode backward.
Une alternative consiste à ajuster un modèle de tendance par morceaux. Vous pouvez choisir cette seconde approche.

2. *Proposer et ajuster un modèle de saison.*

Tracer la série sans tendance ni saison.

Donner l'expression mathématique du modèle ajusté.

3. *Modéliser la série stationnaire*

On suppose que la série obtenue à l'issue de la question précédente est stationnaire. On va la modéliser.

Afin de rendre vos travaux plus homogènes, je vous propose d'utiliser la série, supposée stationnaire, disponible sur la page web du cours.

On propose d'ajuster des modèles ARMA. Sélectionner le "meilleur" modèle parmi les modèles ARMA jusqu'à l'ordre (4,4). Indiquer le critère d'optimalité choisi (AIC, BIC, validation croisée...).

Etudier les résidus du modèle retenu.

Donner l'expression mathématique du modèle retenu pour la série stationnaire.

4. Ecrire le modèle complet dans le cas où on considère un modèle AR(4) et l'utiliser pour prédire le nombre de naissances entre 2014 et 2019. Vous pourrez construire les intervalles de prédiction par simulation.

Comparaison des modèles

Proposer une méthode pour comparer les deux modèles (on pourra reprendre ce qui a été fait plus haut), la mettre en oeuvre. Conclure.