

ING2-MI : Séries Temporelles

TP2- Prédiction par régression linéaire

Exercice 1 (détection de la saisonnalité- tableau de Buys-Ballon).

On considère la série des ventes trimestrielles d'un produit :

Tableau 1 : Série des ventes

DATES	VENTES
$T1$	1248,30
$T2$	1392,10
$T3$	1056,60
$T4$	3159,10
$T1$	890,80
$T2$	1065,30
$T3$	1117,60
$T4$	2934,20
$T1$	1138,20
$T2$	1456,00
$T3$	1224,30
$T4$	3090,20

- 1) Tracer le graphique de cette série et commenter.
- 2) Construire le tableau de Buys-Ballon (tableau constitué en lignes par les années et en colonne par le facteur à analyser (mois, trimestre, ...))
- 3) Classifier les trimestres pour chaque année par valeurs décroissantes.
- 4) Quelle est votre conclusion ?

Exercice 2 : désaisonnalisation d'une série par la méthode de régression sur le temps.

Rappel. Le schéma qui suit illustre le traitement de la désaisonnalisation d'une série trimestrielle (t, X_t) :

$$\begin{array}{ccc}
 X_t \simeq m_t & \rightarrow & MCO \\
 & & \Downarrow \\
 \text{On estime } m_t & & \hat{m}_t = \hat{a} + \hat{b}t \\
 \text{Modél Additif} & & \text{Modèle Multiplicatif} \\
 & & \Downarrow \\
 \text{Méthode de la différence à la tendance} & & \text{Méthode du rapport à la tendance} \\
 d_t = X_t - \hat{m}_t & & d_t = \frac{X_t}{\hat{m}_t} \\
 \text{Coefficients saisonniers provisoires :} & & c_j = \frac{1}{\text{Card}(T_j)} \sum_{t \in T_j} d_t, \quad 1 \leq j \leq p \\
 \sum_j c_j = 0 & & \sum_j c_j = p \\
 \text{Coefficients saisonniers définitifs : } c_j^* = c_j - \bar{c} & & c_j^* = \frac{c_j}{c} \\
 S_t^* = \{c_j^*, j \in T_j\}, \quad X_t^{cvs} = X_t - S_t^* & & X_t^{cvs} = \frac{X_t}{S_t^*}.
 \end{array}$$

Désaisonnaliser la série de l'exercice 1 selon un schéma additif en utilisant la méthode de régression sur le temps.

Exercice 3. Désaisonnalisation par régression sur des variables indicatrices.

On considère ici le cas où la tendance est modélisée par une fonction linéaire et la composante saisonnière (S_t) par un ensemble de variables dichotomiques, notées $\mathbf{1}_{S_t^j}$. Soit X_t une série trimestrielle (de période 4) obéissant à un schéma additif. On peut donc modéliser cette série par

$$X_t = \beta_1 + \beta_2 t + \sum_{j=1}^4 c_j I_{S_t^j} + \epsilon_t, \quad t = 1, \dots, T. \quad (1)$$

où $\mathbf{1}_{S_t^j} = 1$ si t correspond à la période saisonnière (ou t est congru à j modulo 4).

Partie I. Désaisonnalisation avec Excel .

Rappel. On rappelle que l'estimation du vecteur θ des paramètres d'un modèle de régression linéaire est donnée par $\hat{\theta} = (X'X)^{-1}X'Y$ où X est la matrice du modèle et Y le vecteur des observations de la variable à expliquer. On applique cette méthode avec les notations des cours en posant $\hat{\theta} = \hat{\delta} = (\hat{\beta}_1, \hat{\beta}_2, \hat{c}_1, \hat{c}_2, \hat{c}_3, \hat{c}_4)'$, $X = D$ et $Y = \text{Ventes}$ (série de ventes de l'exercice 2).

1. Quelle est la matrice D du modèle complet ? Calculer le déterminant de $(D'D)^{-1}$. Que constatez-vous ?

2. Considérons le modèle sans le coefficient saisonnier c_4 .

a) Écrire la matrice du modèle. Déterminer le vecteur $\hat{\delta} = (\hat{\beta}_1, \hat{\beta}_2, \hat{c}_1, \hat{c}_2, \hat{c}_3)'$. Écrire le modèle obtenu et calculer la série des prévisions $(\hat{X}_t)_{1 \leq t \leq T}$. Tracer la série initiale et la série des prévisions sur le même graphique et commenter.

b) Comparer les résultats obtenus avec les résultats de l'exercice 2.

c) Peut-on désaisonnaliser la série (X_t) ?

3. Considérer le modèle sans la constante β_1 et répondre aux questions a), b), c) de 2.

Partie II. Désaisonnalisation avec R .

1. Taper les commandes suivantes :

```
Mydata=read.table("Vente.txt",header=T,sep="\t",dec=",")
plot(Mydata$Ventes,type="b",lwd=2)
modele=lm(Ventes~t+Dates,data=Mydata)
summary(modele)
```

a) Que constatez-vous ?

b) Écrire le modèle obtenu.

c) Faire de la prévision et comparer avec les résultats obtenus avec Excel.

d) peut-on désaisonnaliser la série ?

2. Étude du modèle sans la constante.

Pour que R ne mette pas la constante dans le modèle, on précise -1 dans la formule

```
modele=lm(Ventes~t+Dates-1,data=Mydata)
summary(modele)
```

a) Écrire le modèle obtenu.

b) Comparer avec la question c) de la partie I.

c) Calculer les composantes saisonnières :

```

a=modele$coef[1]
b=mean(modele$coef[2:5])
c1=modele$coef[2]-b
c2=modele$coef[3]-b
c3=modele$coef[4]-b
c4=modele$coef[5]-b

```

d) Retrouver la série désaisonnalisée. Pour cela il faut travailler de façon matricielle

```

X1=Mydata$Dates=="T1"
X1=X1*(Mydata$Ventes-c1) # Regarder le vecteur obtenu
X2=Mydata$Dates=="T2"
X2=X2*(Mydata$Ventes-c2)
X3=Mydata$Dates=="T3"
X3=X3*(Mydata$Ventes-c3)
X4=Mydata$Dates=="T4"
X4=X4*(Mydata$Ventes-c4)
Xcsv=X1+X2+X3+X4

```

3. Eude du modèle sans le coefficient c_4 . Pour éviter le problème d'inversibilité de la matrice du modèle, on calcule seulement trois coefficients c_1, c_2, c_3 , et pour respecter le principe de la conversion des aires : $\sum_{j=1}^4 c_j = 0$, on choisit $c_4 = -(c_1 + c_2 + c_3)$.

a) Écrire la matrice du modèle.

b) Générer trois variables indicatrices $\mathbf{1}_{S_1}$, $\mathbf{1}_{S_2}$ et $\mathbf{1}_{S_3}$ puis effectuer la régression avec les ventes de la série de l'exercice 2 comme variable à expliquer sur les quatre variables explicatives (tendance, $\mathbf{1}_{S_1}$, $\mathbf{1}_{S_2}$, $\mathbf{1}_{S_3}$).

c) Analyser la sortie de la fonction "lm" de R.

d) Écrire le modèle obtenu et calculer la série des prévisions. Tracer la série initiale et la série des prévisions sur le même graphique et commenter. Comparer avec la question 2.b) de la première partie.

Exercice 4 (données réelles). On considère les données de la série représente le nombre mensuel de passagers aériens, en milliers, de janvier 1949 à décembre 1960. Appliquez la méthode de régression sur les variables indicatrices pour désaisonnaliser cette série.