Modèles d'aide à la décision
4-600-04

Séance 6
Séries chronologiques

Plan

- 6.1 Introduction
- 6.2 Critères de comparaison
- 6.3 Séries stationnaires sans effets saisonniers
- 6.4 Séries stationnaires avec effets saisonniers
- 6.5 Séries non stationnaires sans effets saisonniers
- 6.6 Séries non stationnaires avec effets saisonniers

6.1 Introduction

- Une série chronologique est un ensemble d'observations sur une variable quantitative recueillies dans le temps.
- À partir de ces données, on tente de construire un modèle pour prévoir le futur.
- Souvent, les variables indépendantes ne sont pas disponibles pour construire un modèle de régression pour une variable de séries chronologiques.
- On tente alors d'analyser le comportement passé de la variable afin de prédire son comportement futur.

$$\hat{Y}_{t+1} = f\left(Y_t, Y_{t-1}, Y_{t-2}, \ldots\right)$$

6.1 Terminologie

- Données stationnaires : une variable de séries chronologiques ne démontrant aucune tendance à la hausse ou à la baisse dans le temps.
- Données non stationnaires : une variable de séries chronologiques démontrant une tendance significative à la hausse ou à la baisse dans le temps.
- Données saisonnières : une variable de séries chronologiques démontrant un comportement qui se répète à des intervalles réguliers dans le temps.

6.1 Trucs pour faire une analyse de séries chronologiques

- Il existe plusieurs techniques de séries chronologiques.
- Il est généralement impossible de savoir a priori quelle technique sera la meilleure pour un ensemble de données.
- Il est courant d'essayer plusieurs techniques différentes et de choisir celle qui semble fonctionner le mieux.
- Pour modéliser efficacement, il faut être en mesure de tester plusieurs techniques de séries chronologiques.

6.2 Critères de comparaison

- Il faut un critère pour comparer les modèles de séries chronologiques sur un ensemble de données.
- Les quatre critères usuels sont :
  – Moyenne des déviations absolues : $\mathrm{MAD} = \sum_{i=1}^{n} \frac{\left|\mathrm{Y}_i - \hat{\mathrm{Y}}_i\right|}{n}$
  – Moyenne des pourcentages d'erreurs absolues : $$\mathrm{MAPE} = \frac{100}{n} \sum_{i=1}^{n} \frac{\left|\mathrm{Y}_i - \hat{\mathrm{Y}}_i\right|}{\mathrm{Y}_i}$$
  – Moyenne des carrés des erreurs : $\mathrm{MSE} = \sum_{i=1}^{n} \frac{\left(\mathrm{Y}_i - \hat{\mathrm{Y}}_i\right)^2}{n}$
  – Racine de MSE : $\mathrm{RMSE} = \sqrt{\mathrm{MSE}}$

Nous allons nous attarder sur MSE.

6.3 Séries Chronologiques stationnaires sans effets saisonniers

- Série où il n'y a pas de tendance à long terme ou d'effets saisonniers significatifs.
- Deux modèles classiques seront étudiés : moyenne mobile et lissage exponentiel.

6.3 Exemple

- Electra-City est un magasin qui vend des appareils audio et vidéo.
- Chaque mois, le gérant du magasin doit commander la marchandise auprès d'un entrepôt éloigné.
- Présentement, le gérant tente d'estimer combien de lecteurs vidéo le magasin espère vendre le prochain mois.
- Il a recueilli les données sur les 24 derniers mois.
- Voir 4600-6.3-Electra-Data.xlsx

6.3.1 Moyenne mobile

- On prévoit le futur en se basant sur les observations récentes.
- La prévision de la prochaine période correspond à la moyenne des *k* dernières périodes.

$$\hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + \ldots + Y_{t-k+1}}{k}$$

- Aucune méthode n'existe pour déterminer la meilleur valeur de *k.*
- Il faut essayer différentes valeurs de *k* pour déterminer celle qui fonctionne le mieux.
- Voir [4600-6.3.1-Electra-Sol.xlsx]

6.3.1 Prévision avec le modèle de moyenne mobile

Prévisions pour les périodes 25 et 26 lors de la période 24 :

$$\hat{Y}_{25} = \frac{Y_{24} + Y_{23}}{2} = \frac{36 + 35}{2} = 35.5$$

$$\hat{Y}_{26} = \frac{\hat{Y}_{25} + Y_{24}}{2} = \frac{35.5 + 36}{2} = 35.75$$

6.3.2 Lissage exponentiel

- La prévision à une période est basée sur la prévision et l'observation à la période précédente :

$$\hat{Y}_{t+1} = \hat{Y}_t + \alpha(Y_t - \hat{Y}_t)$$

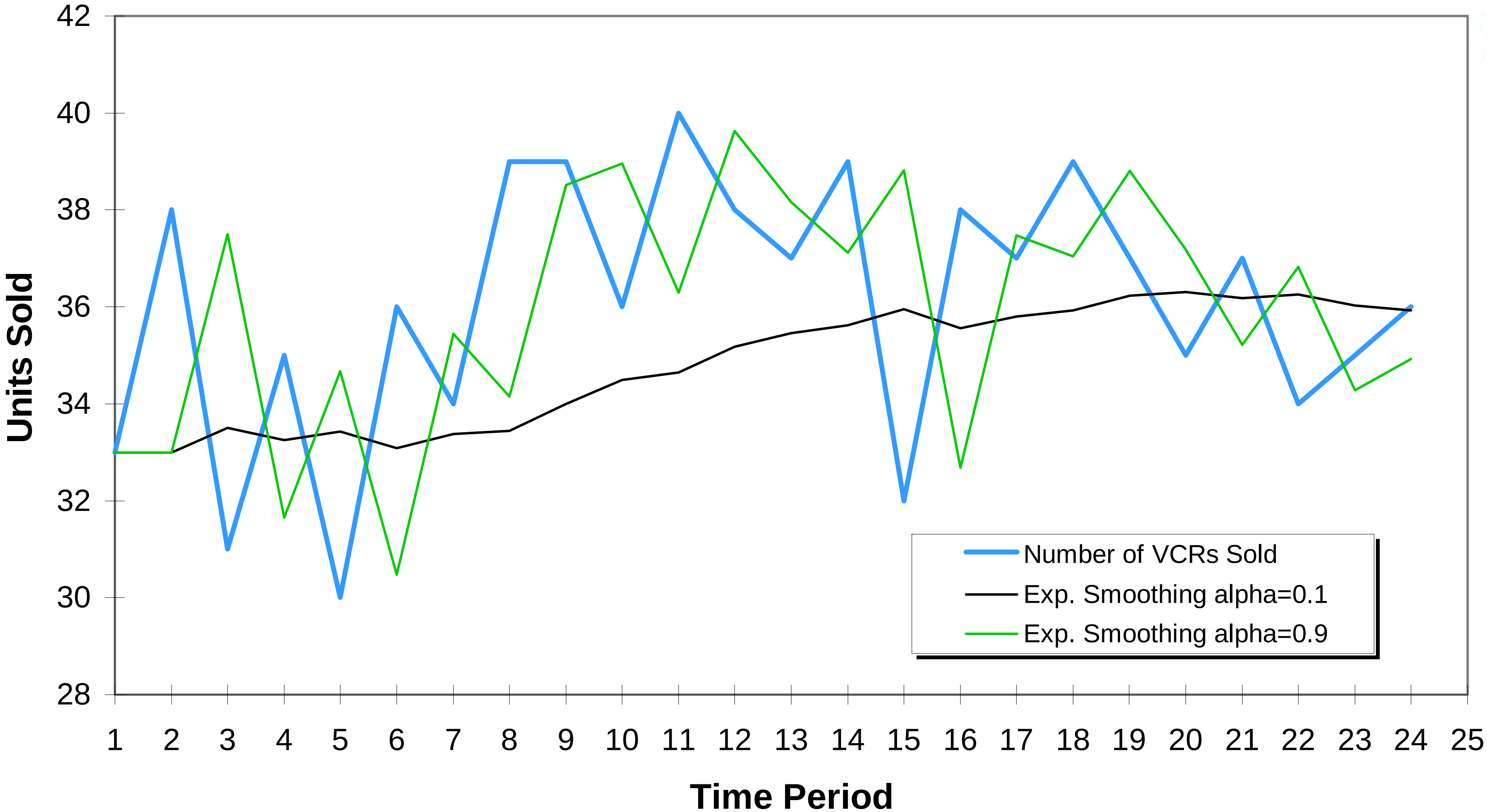
$$\text{où } 0 \leq \alpha \leq 1$$

  - Qui est équivalent à :

$$\hat{Y}_{t+1} = \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \cdots + \alpha(1-\alpha)^n Y_{t-n} + \cdots$$

- Voir 4600-6.3.2-Electra-Sol.xlsx

6.3.2 Exemples de deux fonctions de lissage exponentiel



6.3.2 Prévisions avec le modèle de lissage exponentiel

Prévisions pour les périodes 25 et 26 lors de la période 24 :

$$\hat{Y}_{25} = \hat{Y}_{24} + \alpha(Y_{24} - \hat{Y}_{24}) = 35.74 + 0.268(36 - 35.74) = 35.81$$

$$\hat{Y}_{26} = \hat{Y}_{25} + \alpha(Y_{25} - \hat{Y}_{25}) \approx \hat{Y}_{25} + \alpha(\hat{Y}_{25} - \hat{Y}_{25}) = \hat{Y}_{25} = 35.81$$

Notez que,

$$\hat{Y}_t = 35.81, \text{ pour } t = 25, 26, 27, \ldots$$

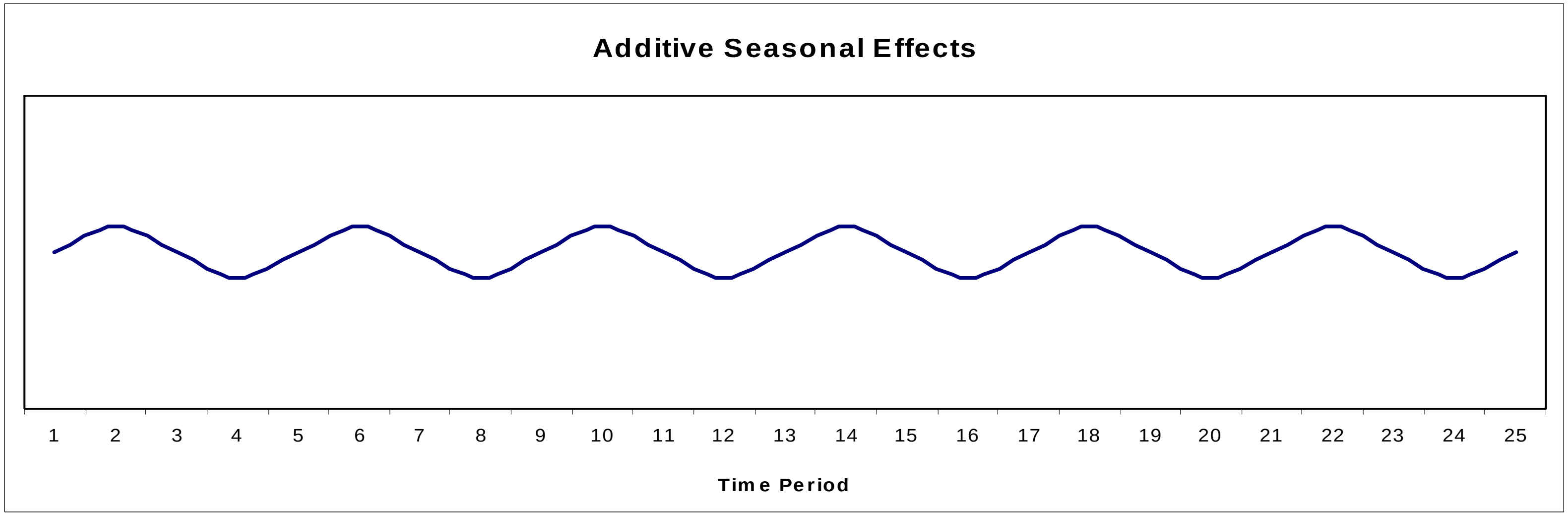
6.4 Séries chronologiques stationnaires avec effets saisonniers

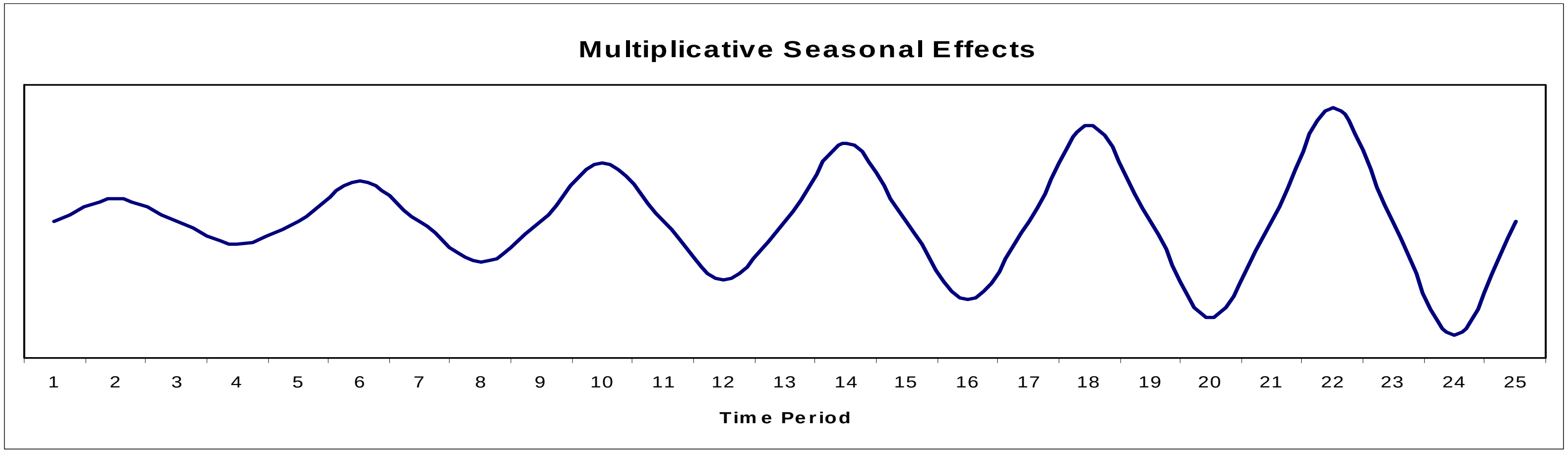
- Une saisonnalité est un comportement régulier et répétitif dans des données de séries chronologiques.
- L'effet peut être additif ou multiplicatif.
- Deux modèles de lissage seront étudiés : avec effet additif ou avec effet multiplicatif.

6.4 Exemple

- Savannah Climate Control vend et répare des thermo-pompes;
- Les ventes sont plus élevées l'hiver et l'été;
- Le propriétaire a recueilli des données trimestrielles sur les ventes lors des dernières années;
- Voir 4600-6.4-Pompes-Data.xlsx
- On veut prédire les ventes pour chaque trimestre de 2006.

6.4 Effets saisonniers stationnaires





6.4.1 Données stationnaires avec effets saisonniers additifs

$$\hat{Y}_{t+1} = E_t + S_{t+1-p} \qquad p+1 \le t+1 \le last$$

$$E_t = \alpha(Y_t - S_{t-p}) + (1-\alpha)E_{t-1} \qquad 0 \le \alpha \le 1$$

$$S_t = \beta(Y_t - E_t) + (1-\beta)S_{t-p} \qquad 0 \le \beta \le 1$$

- ***p** :* nb de périodes saisonnières ***last** :* nb d'observations
- E_t est le niveau espéré à la période *t.*
- S_t est le facteur saisonnier additif pour la période *t.*

$$\text{Initialisation}: \; t = 1,\ldots,p \qquad E_t = \sum_{i=1}^{p} Y_i / p \qquad S_t = Y_t - E_t$$

6.4.1 Données stationnaires avec effets saisonniers additifs

$$\text{Prévisions :}\quad \hat{Y}_{last+n} = E_{last} + S_{last+n-p} \qquad 1 \le n \le p$$

last *:* nb d'observations

6.4.1 Prévision avec un modèle d'effets saisonniers additifs

Prévisions pour les périodes 25 à 28 lors de la période 24 :

$$\hat{Y}_{24+n} = E_{24} + S_{24+n-4}$$

$$\hat{Y}_{25} = E_{24} + S_{21} = 354.55 + 8.45 = 363.00$$

$$\hat{Y}_{26} = E_{24} + S_{22} = 354.55 - 17.82 = 336.73$$

$$\hat{Y}_{27} = E_{24} + S_{23} = 354.55 + 46.58 = 401.13$$

$$\hat{Y}_{28} = E_{24} + S_{24} = 354.55 - 31.73 = 322.81$$

6.4.1 Données stationnaires avec effets saisonniers multiplicatifs

$$\hat{Y}_{t+1} = E_t \times S_{t+1-p} \qquad p+1 \le t+1 \le last$$

$$E_t = \alpha(Y_t/S_{t-p}) + (1-\alpha)E_{t-1} \qquad 0 \le \alpha \le 1$$

$$S_t = \beta(Y_t/E_t) + (1-\beta)S_{t-p} \qquad 0 \le \beta \le 1$$

- ***p :*** nb de périodes saisonnières ***last :*** nb d'observations
- E_t est le niveau espéré à la période *t.*
- S_t est le facteur saisonnier multiplicatif pour la période *t.*

$$\text{Initialisation}: \; t = 1,\ldots, p \qquad E_t = \sum_{i=1}^{p} Y_i / p \qquad S_t = Y_t / E_t$$

6.4.2 Données stationnaires avec effets saisonniers multiplicatifs

$$\text{Prévisions :} \quad \hat{Y}_{last+n} = E_{last} \times S_{last+n-p} \qquad 1 \leq n \leq p$$

last *:* nb d'observations

6.4.2 Prévision avec un modèle d'effets saisonniers multiplicatifs

Prévisions pour les périodes 25 à 28 lors de la période 24 :

$$\hat{Y}_{24+n} = E_{24} \times S_{24+n-4}$$

$$\hat{Y}_{25} = E_{24} \times S_{21} = 353.95 \times 1.015 = 359.13$$

$$\hat{Y}_{26} = E_{24} \times S_{22} = 353.95 \times 0.946 = 334.94$$

$$\hat{Y}_{27} = E_{24} \times S_{23} = 353.95 \times 1.133 = 400.99$$

$$\hat{Y}_{28} = E_{24} \times S_{24} = 353.95 \times 0.912 = 322.95$$

6.5 Séries non stationnaires sans effets saisonniers

- On parle de tendance (séries non stationnaires) lorsqu'il semble y avoir une direction générale de montée ou de descente à long terme dans une série chronologique.
- Deux modèles (Holt et RLS) seront étudiés bien qu'il en existe d'autres.

6.5 Exemple

- WaterCraft Inc. est un manufacturier de moto-marines.
- La compagnie a eu une croissance soutenue dans les ventes de ses produits.
- Les gestionnaires de la compagnie veulent prédire le niveau des ventes pour chaque trimestre de 2006.
- Voir 4600-6.5-WaterCraft-Data.xlsx

6.5.1 Régression linéaire simple

- Utilisation d'un modèle de RLS où la variable explicative (indépendante) est le temps.
- Limitation : hypothèse de linéarité
- Approprié pour des modèles à court terme.
- Pour des modèles à long terme, ce type de modèle permet de capter la tendance générale.
- On peut faire des estimations dans un futur éloigné mais l'erreur d'approximation peut être très élevée.

6.5.1 Régression linéaire simple

$$\hat{Y}_t = b_0 + b_1\, t$$

- Voir 4600-6.5.1-WaterCraft-Sol.xlsx

6.5.1 Prévision avec le modèle de régression linéaire simple

Prévisions pour les périodes 21 à 24 lors de la période 20:

$$\hat{Y}_{21} = b_0 + b_1 X_{1_{21}} = 375.1 + 92.6255 \times 21 = 2320.3$$

$$\hat{Y}_{22} = b_0 + b_1 X_{1_{22}} = 375.1 + 92.6255 \times 22 = 2412.9$$

$$\hat{Y}_{23} = b_0 + b_1 X_{1_{23}} = 375.1 + 92.6255 \times 23 = 2505.6$$

$$\hat{Y}_{24} = b_0 + b_1 X_{1_{24}} = 375.1 + 92.6255 \times 24 = 2598.2$$

6.5.2 Lissage exponentiel double - Méthode de Holt

$$\hat{Y}_{t+1} = E_t + T_t \qquad k+1 \leq t \leq last$$

où $$E_t = \alpha Y_t + (1-\alpha)(E_{t-1} + T_{t-1})$$

$$T_t = \beta(E_t - E_{t-1}) + (1-\beta)\, T_{t-1}$$

$$0 \leq \alpha \leq 1 \text{ et } 0 \leq \beta \leq 1$$

- E_t est le niveau de base espéré pour la période *t.*
- T_t est la tendance attendue à la période *t.*
- Voir 4600-6.5.2-WaterCraft-Sol.xlsx

$$\text{Prévisions } : \hat{Y}_{last+n} = E_{last} + nT_{last} \qquad n \geq 1$$

6.5.2 Prévision avec le modèle de Holt

Prévisions pour les périodes 21 à 24 lors de la période 20:

$$\hat{Y}_{20+n} = E_{20} + nT_{20}$$

$$\hat{Y}_{21} = E_{20} + 1T_{20} = 2336.8 + 1 \times 152.1 = 2488.9$$

$$\hat{Y}_{22} = E_{20} + 2T_{20} = 2336.8 + 2 \times 152.1 = 2641.0$$

$$\hat{Y}_{23} = E_{20} + 3T_{20} = 2336.8 + 3 \times 152.1 = 2793.1$$

$$\hat{Y}_{24} = E_{20} + 4T_{20} = 2336.8 + 4 \times 152.1 = 2945.2$$

6.6 Séries non stationnaires avec effets saisonniers

- Séries chronologiques avec tendance à la hausse ou à la baisse à long terme en plus d'avoir des effets saisonniers.
- Voir 4600-6.5-WaterCraft-Data.xlsx

6.6.1 Modèles de Holt-Winter pour effets saisonniers additifs

$$\hat{Y}_{t+1} = E_t + T_t + S_{t+1-p} \qquad p+1 \le t+1 \le last$$

où

$$E_t = \alpha\left(Y_t - S_{t-p}\right) + (1\text{-}\alpha)(E_{t-1} + T_{t-1})$$
$$T_t = \beta\left(E_t - E_{t-1}\right) + (1\text{-}\beta)T_{t-1}$$
$$S_t = \gamma\left(Y_t - E_t\right) + (1\text{-}\gamma)S_{t-p}$$

$$0 \le \alpha \le 1,\ 0 \le \beta \le 1 \text{ et } 0 \le \gamma \le 1$$

- Voir 4600-6.6.1-WaterCraft-Sol.xlsx

$$\text{Prévisions} : \quad \hat{Y}_{last+n} = E_{last} + n\,T_{last} + S_{last+n-p} \qquad 1 \le n \le p$$

6.6.1 Prévision avec le modèle de Holt-Winter pour effets saisonniers additifs

Prévisions pour les périodes 21 à 24 lors de la période 20:

$$\hat{Y}_{20+n} = E_{20} + nT_{20} + S_{20+n-4}$$

$$\hat{Y}_{21} = E_{20} + 1 \times T_{20} + S_{17} = 2253.3 + 1 \times 154.3 + 262.66 = 2670.3$$

$$\hat{Y}_{22} = E_{20} + 2 \times T_{20} + S_{18} = 2253.3 + 2 \times 154.3 - 312.59 = 2249.3$$

$$\hat{Y}_{23} = E_{20} + 3 \times T_{20} + S_{19} = 2253.3 + 3 \times 154.3 + 205.40 = 2921.6$$

$$\hat{Y}_{24} = E_{20} + 4 \times T_{20} + S_{20} = 2253.3 + 4 \times 154.3 + 386.12 = 3256.6$$

6.6.2 Modèles de Holt-Winter pour effets saisonniers multiplicatifs

$$\hat{Y}_{t+1} = (E_t + T_t) S_{t+1-p} \qquad p+1 \leq t+1 \leq last$$

où

$$E_t = \alpha\left(Y_t / S_{t-p}\right) + (1\text{-}\alpha)(E_{t-1} + T_{t-1})$$

$$T_t = \beta\left(E_t - E_{t-1}\right) + (1\text{-}\beta)T_{t-1}$$

$$S_t = \gamma\left(Y_t / E_t\right) + (1\text{-}\gamma)S_{t-p}$$

$$0 \leq \alpha \leq 1,\ 0 \leq \beta \leq 1 \text{ et } 0 \leq \gamma \leq 1$$

- Voir 4600-6.6.2-WaterCraft-Sol.xlsx

$$\text{Prévisions} \ : \hat{Y}_{last+n} = (E_{last} + nT_{last}) S_{last+n-p} \qquad 1 \leq n \leq p$$

6.6.2 Prévision avec le modèle de Holt-Winter pour effets saisonniers multiplicatifs

Prévisions pour les périodes 21 à 24 lors de la période 20:

$$\hat{Y}_{20+n} = \left(E_{20} + nT_{20}\right)S_{20+n-4}$$

$$\hat{Y}_{21} = (E_{20} + 1T_{20})S_{17} = (2217.6 + 1\times 137.3)1.152 = 2713.7$$

$$\hat{Y}_{22} = (E_{20} + 2T_{20})S_{18} = (2217.6 + 2\times 137.3)0.849 = 2114.9$$

$$\hat{Y}_{23} = (E_{20} + 3T_{20})S_{19} = (2217.6 + 3\times 137.3)1.103 = 2900.5$$

$$\hat{Y}_{24} = (E_{20} + 4T_{20})S_{20} = (2217.6 + 4\times 137.3)1.190 = 3293.9$$

6.6.3 Modèles de régression linéaire multiple avec saisonnalité

- On peut utiliser des variables de catégories pour tenir compte des effets de saisonnalité dans le modèle.
- Attention aux interactions entre les variables de catégories et la variable de temps.
- S’il y a *p* saisons, il faut *p* -1 variables de catégorie.

6.6.3 Modèles de régression linéaire multiple avec saisonnalité

- Pour l'exemple de WaterCraft, on considère des données trimestrielles, alors *p*=4 et nous définissions 3 variables de catégorie :

$$(TRIM\,1)_t = \begin{cases} 1, \text{ si } \mathrm{Y}_t \text{ est une observation du trimestre } 1 \\ 0, \text{ sinon} \end{cases}$$

$$(TRIM\,2)_t = \begin{cases} 1, \text{ si } \mathrm{Y}_t \text{ est une observation du trimestre } 2 \\ 0, \text{ sinon} \end{cases}$$

$$(TRIM\,3)_t = \begin{cases} 1, \text{ si } \mathrm{Y}_t \text{ est une observation du trimestre } 3 \\ 0, \text{ sinon} \end{cases}$$

6.6.3 Modèles de régression linéaire multiple avec saisonnalité

- Le modèle estimé de régression :

$$\hat{Y}_t = b_0 + b_1\, t + b_2 (TRIM1)_t + b_3 (TRIM2)_t + b_4 (TRIM3)_t$$

- Voir 4600-6.6.3-WaterCraft-Sol.xlsx

6.6.3 Prévision pour régression linéaire multiple avec saisonnalité

Prévisions pour les périodes 21 à 24 lors de la période 20:

$$\hat{Y}_{21} = b_0 + b_1 X_{1_{21}} + b_2 = 559.6 + 90.513 \times 21 - 86.8 \approx 2373.56$$

$$\hat{Y}_{22} = b_0 + b_1 X_{1_{22}} + b_3 = 559.6 + 90.513 \times 22 - 431.7 \approx 2119.17$$

$$\hat{Y}_{23} = b_0 + b_1 X_{1_{23}} + b_3 = 559.6 + 90.513 \times 23 - 130.4 \approx 2510.97$$

$$\hat{Y}_{24} = b_0 + b_1 X_{1_{24}} = 559.6 + 90.513 \times 24 \approx 2731.91$$