04-03 Méthodes simples de prévisions

Été 2021

NOUS ÉCLAIRONS.
VOUS BRILLEZ.

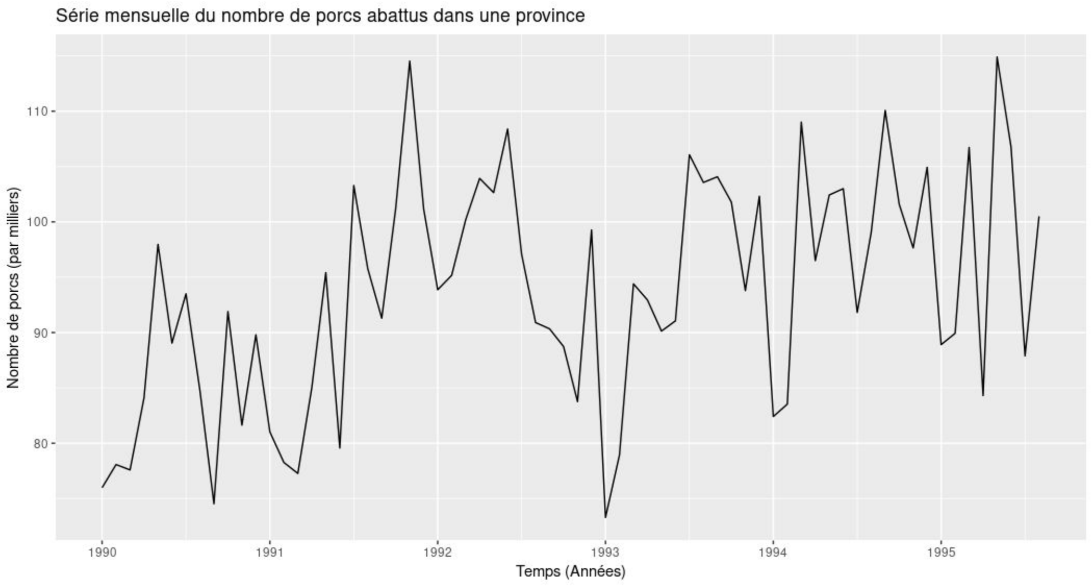
FORMATION CONTINUE
ET SERVICES AUX ENTREPRISES

Cégep de Sainte-Foy

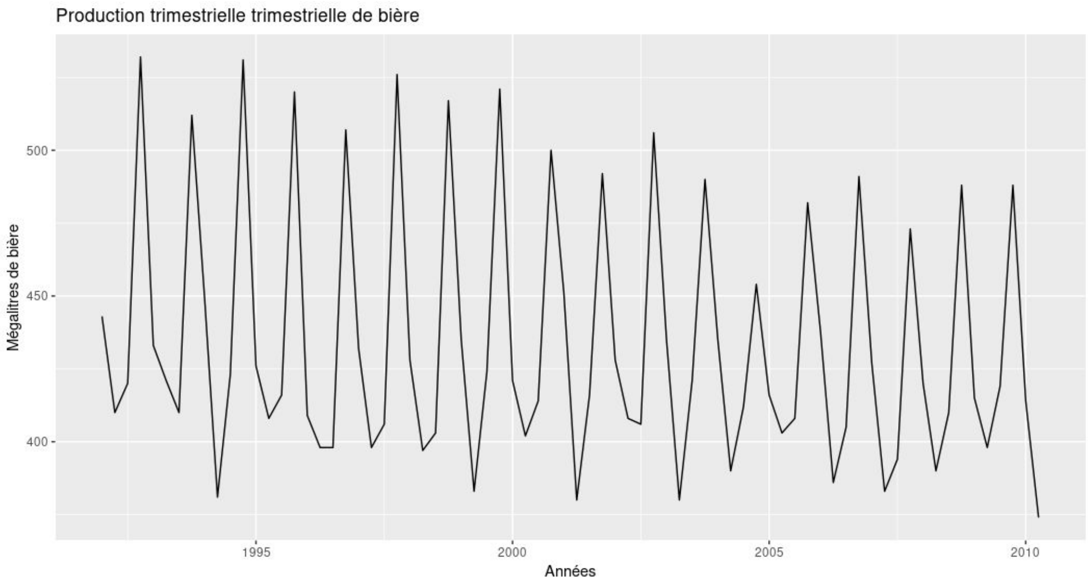
Spécialisation technique en intelligence artificielle
Algorithmes d'apprentissage non supervisé — 420-A58-SF — M. Swawola, M.Sc.

Sommaire

1 Prévision simples



Comment feriez-vous des prévisions à partir de cette série temporelle ?

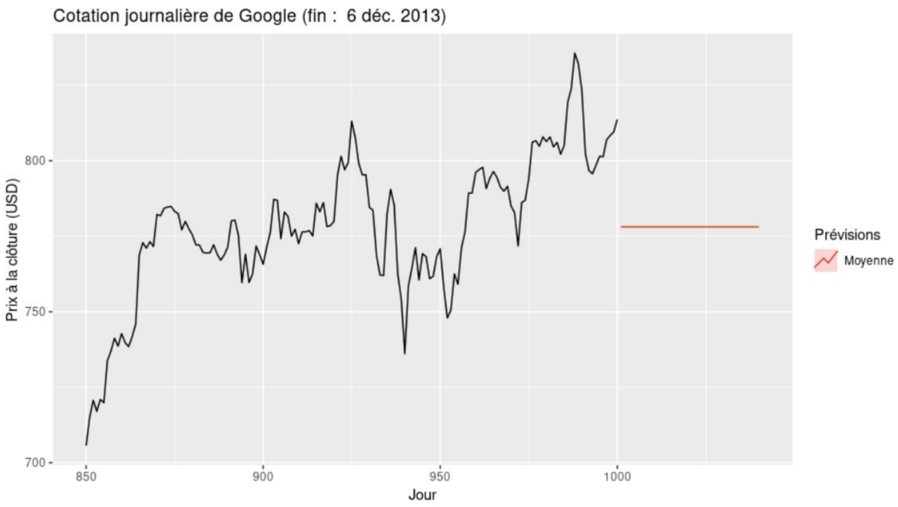


Comment feriez-vous des prévisions à partir de cette série temporelle ?

Méthode de la moyenne

- Toutes les prévisions des valeurs futures sont égales à la moyenne des données historiques

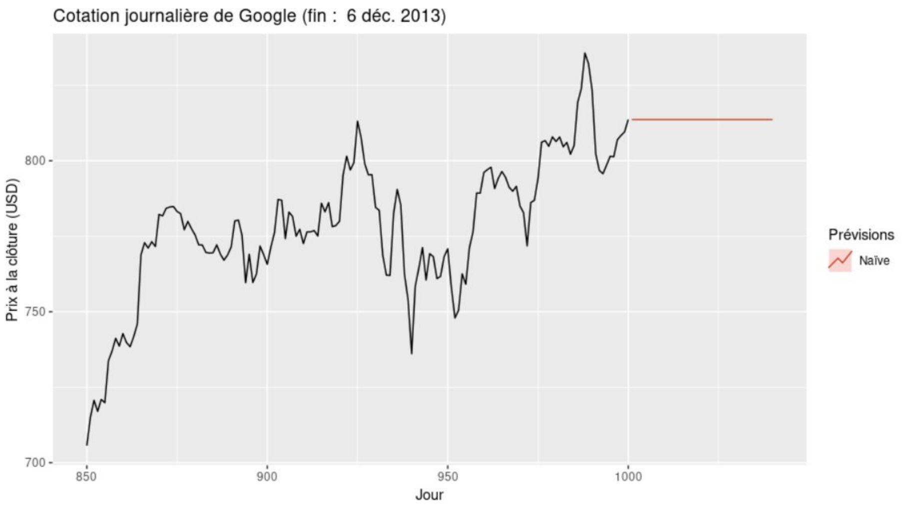
$$\hat{x}_{T+h|T} = \frac{1}{T}(x_1 + \cdots + x_T)$$



Méthode naïve

- Toutes les prévisions sont égales à la dernière observation

$$\hat{x}_{T+h|T} = x_T$$

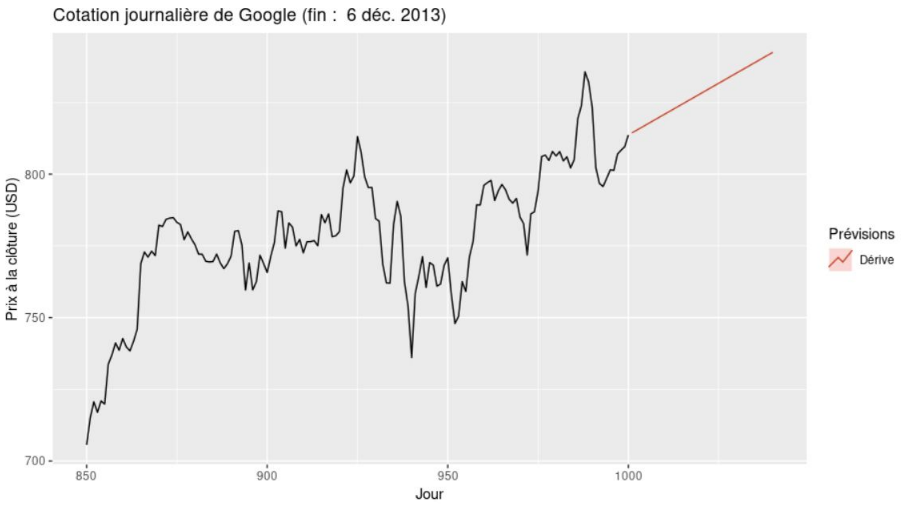


Méthode de la dérive

- Les prévisions sont calculées via par formule

$$\hat{x}_{T+h|T} = x_T + \frac{h}{h-1}(x_T - x_1)$$

- Équivaut à extrapoler une ligne droite entre la première et la dernière observation

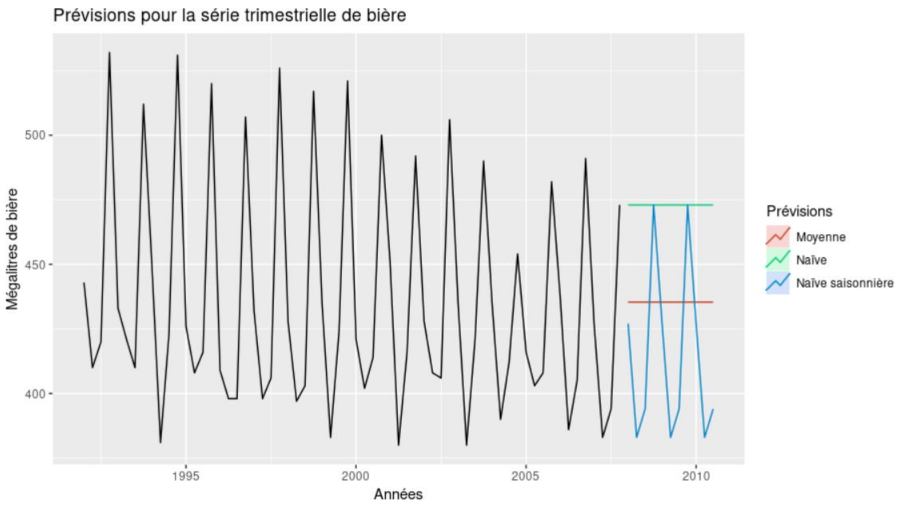


Méthode naïve saisonnière

- Toutes les prévisions sont égales à la dernière valeur de la précédente saison

  $$\hat{x}_{T+h|T} = x_{T+h-m(k+1)}$$

  où **m** est la longueur de la saison et **k** est la partie entière de **(h - 1) / m**



Fonctions de prévision R

- **Moyenne**

```r
meanf(x, h=12)
```

- **Naïve**

```r
naive(x, h=12)
```

- **Naïve saisonnière**

```r
snaive(x, h=24)
```

- **Dérive**

```r
rwf(x, drift=TRUE, h=20)
```

Stabilisation de la variance

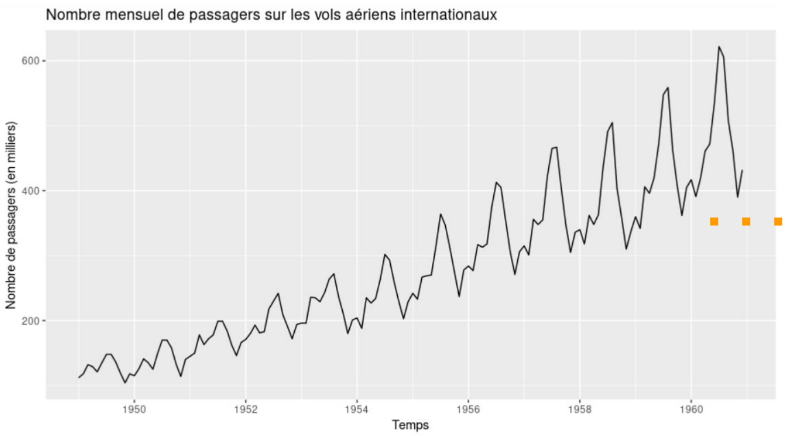
- Première chose à faire dans toute étude de série univariée: **rendre la variance temporellement constante**
- Si la **variabilité** autour de la moyenne change au cours du temps, il faut utiliser une transformation du type **Box et Cox** qui dépend d'un paramètre **λ**

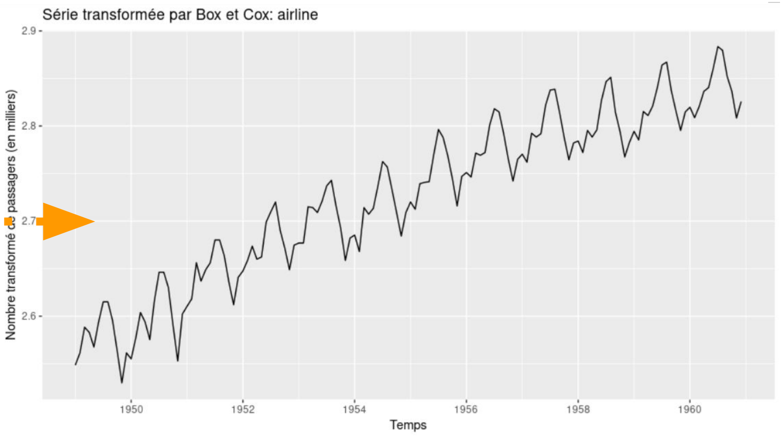
$$W_t = \begin{cases} \ln X_t\,, & \text{si } \lambda = 0 \\ \frac{X_t^\lambda - 1}{\lambda}, & \text{si } \lambda \neq 0 \end{cases}$$

Stabilisation de la variance

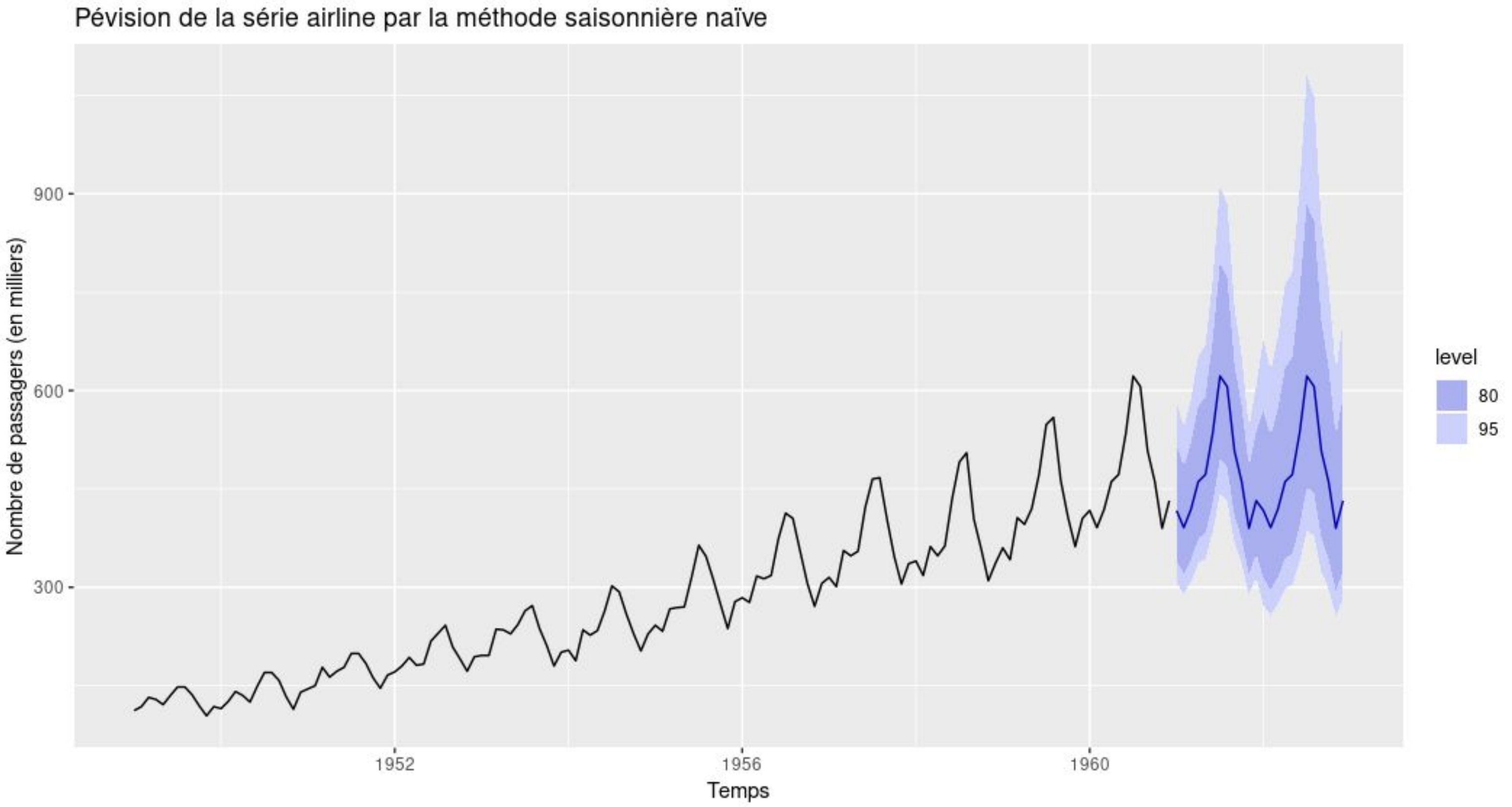
- La valeur de **λ** peut être calculée automatiquement. Par exemple sur la série "airline"

```
lambda <- BoxCox.lambda(airline)
[1] -0.2947156
```





Exemple



Pévision de la série airline par la méthode saisonnière naïve
Nombre de passagers (en milliers)
900
600
300
1952
1956
1960
Temps
level
80
95

Valeurs ajustées

- $\hat{x}_{t|t-1}$ est la prévision de x_t basée sur $x_1, \ldots, x_{t-1}$. On les appelle **valeurs ajustées**.
- On notera $\hat{x}_t = \hat{x}_{t|t-1}$
- Par exemple

  $\hat{x}_t = \bar{x}$ par la méthode de la moyenne

  $\hat{x}_t = x_{t-1} + (x_t - x_1)/(T-1)$ par la méthode de la dérive

Résidus

- Un **résidu** à la date ***t*** est la différence entre la valeur observée et la valeur ajustée. Le résidu est noté ϵ_t

$$\epsilon_t = x_t - \hat{x}_{t|t-1}$$

- **Hypothèses**

  Les ϵ_t sont de moyenne nulle

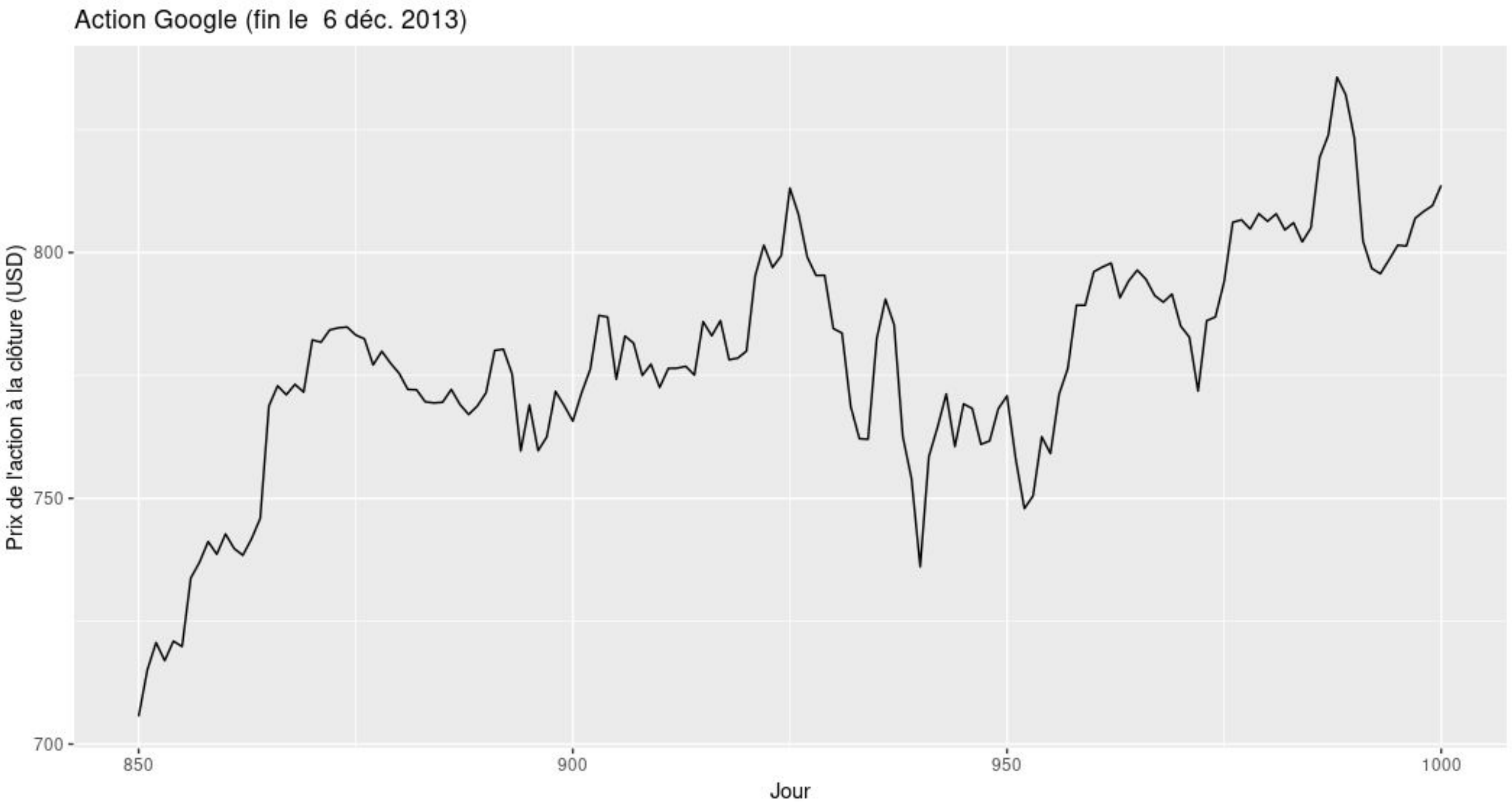
  Les ϵ_t sont deux à deux non corrélés

- **Propriétés utiles**

  Les ϵ_t ont une variance constante

  Les ϵ_t suivent une loi normale

Exemple - Série Google



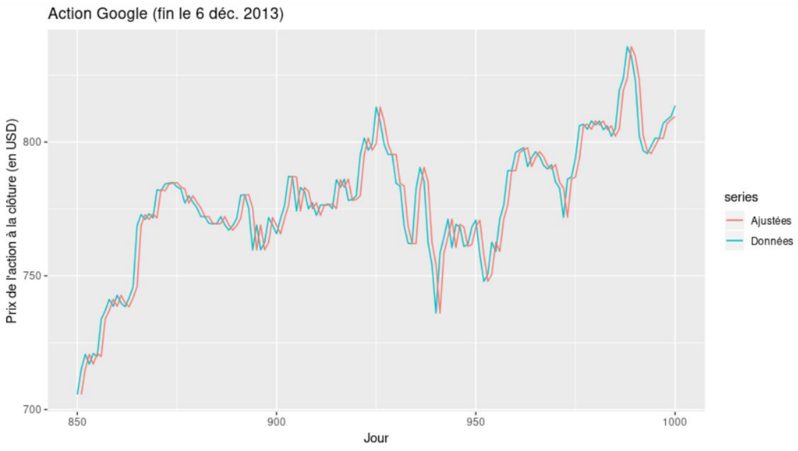
Exemple - Série Google

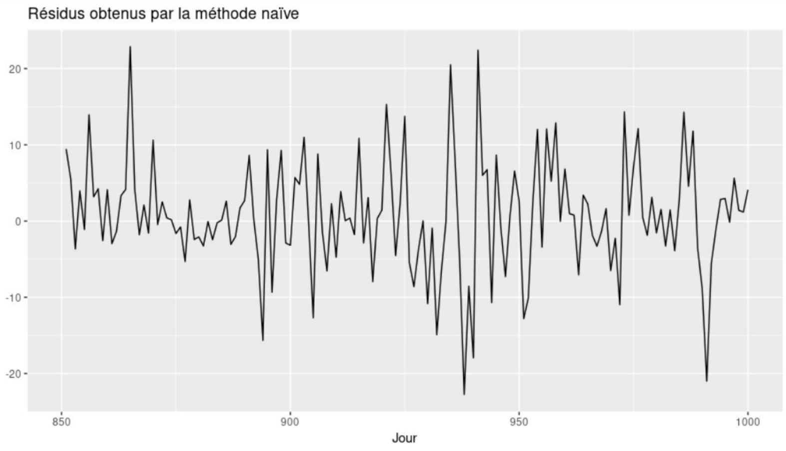
- La prévision naïve est

$$x_t - \hat{x}_{t|t-1} = x_{t-1}$$

- Le résidu associé est donc

$$\epsilon_t = x_t - x_{t-1}$$





Exemple - Série Google

- Code R

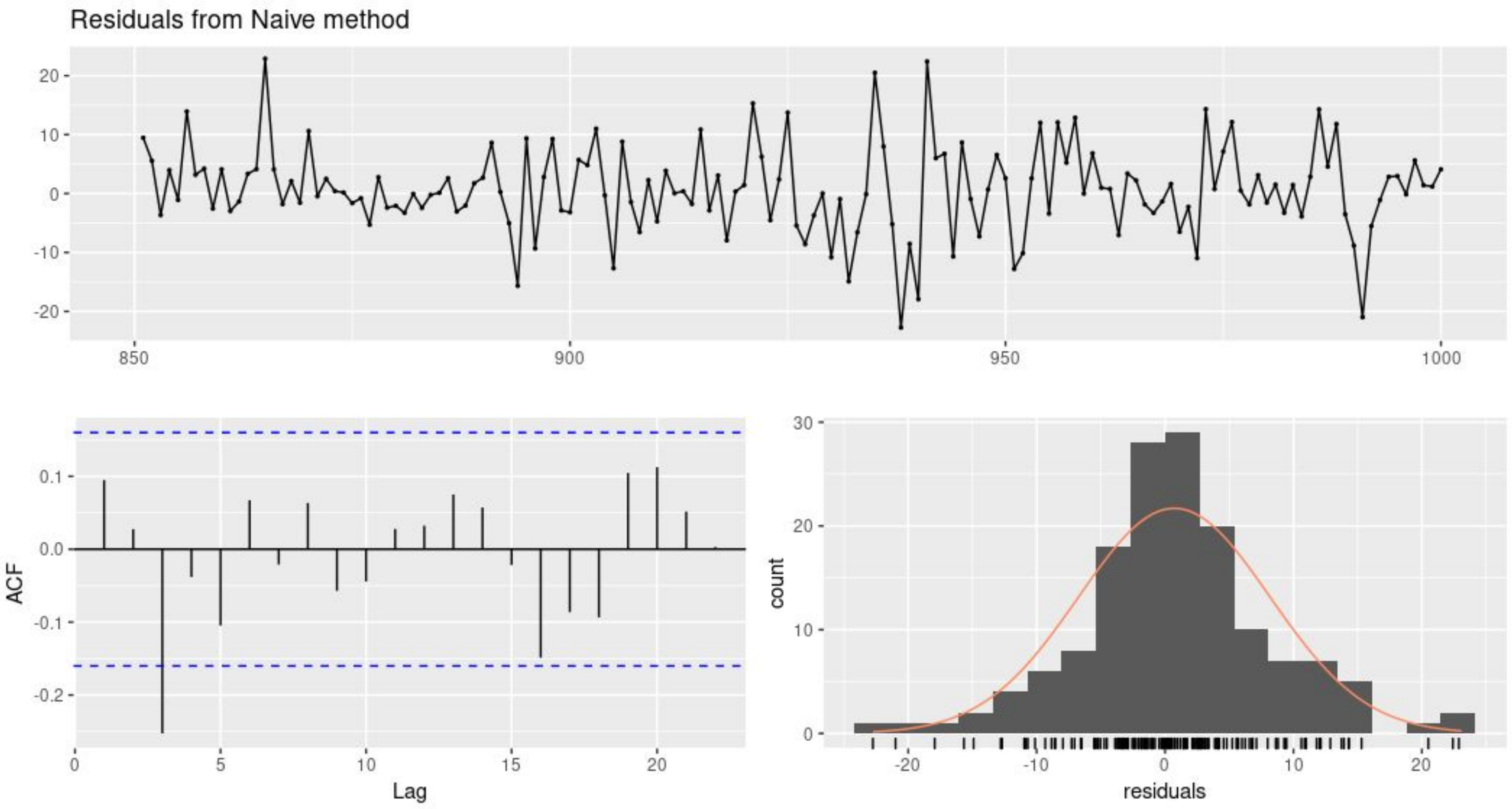
```r
# Série goog150 et valeurs ajustées
fits <- fitted(naive(goog150))
autoplot(goog150, series="Données") +
  autolayer(fits, series="Ajustées") +
  xlab("Jour") + ylab("Prix de l'action à la clôture (en USD)") +
  ggtitle("Action Google (fin le 6 déc. 2013)")

# Série goog150 et résidus
res <- residuals(naive(goog150))
autoplot(res) + xlab("Jour") + ylab("") +
  ggtitle("Résidus obtenus par la méthode naïve")
```

Exemple - Série Google - Étude des résidus

- L'étude des résidus est facilitée par le fonction `checkresiduals`

```
checkresiduals(naive(goog150))
```

Exemple - Série Google - Étude des résidus



Exemple - Série Google - Étude des résidus

- Ljung-Box test

  data: Residuals from Naive method

  Q* = 15.623, df = 10, p-value = 0.111

- On ne rejette pas l'hypothèse que les résidus sont issus d'un bruit blanc

2

Mesures de précision des prévisions

Mesures de précision des prévisions

- Soient $\mathbf{x_t}$ une observation et $\mathbf{f_t}$ sa prévision. Pour tout $\mathbf{t = 1, \ldots, T}$

$$\mathrm{MAE} = \frac{1}{T}\sum_{t=1}^{T}|x_t - f_t|$$

Mean absolute error

$$\mathrm{MSE} = \frac{1}{T}\sum_{t=1}^{T}(x_t - f_t)^2$$

Mean squared error

$$\mathrm{MAPE} = \frac{1}{100T}\sum_{t=1}^{T}\frac{(|x_t - f_t|)}{|x_t|}$$

Mean absolute percentage error

- Toutes sauf le **MAPE** dépendent de l'échelle des observations

Mesures de précision des prévisions

- Une autre mesure est le MASE

$$\mathrm{MASE} = \frac{1}{T}\sum_{t=1}^{T}\frac{|x_t - f_t|}{q}$$

Mean absolute scale error

- Où **q** est une mesure stable de l'échelle de la série (X_t)

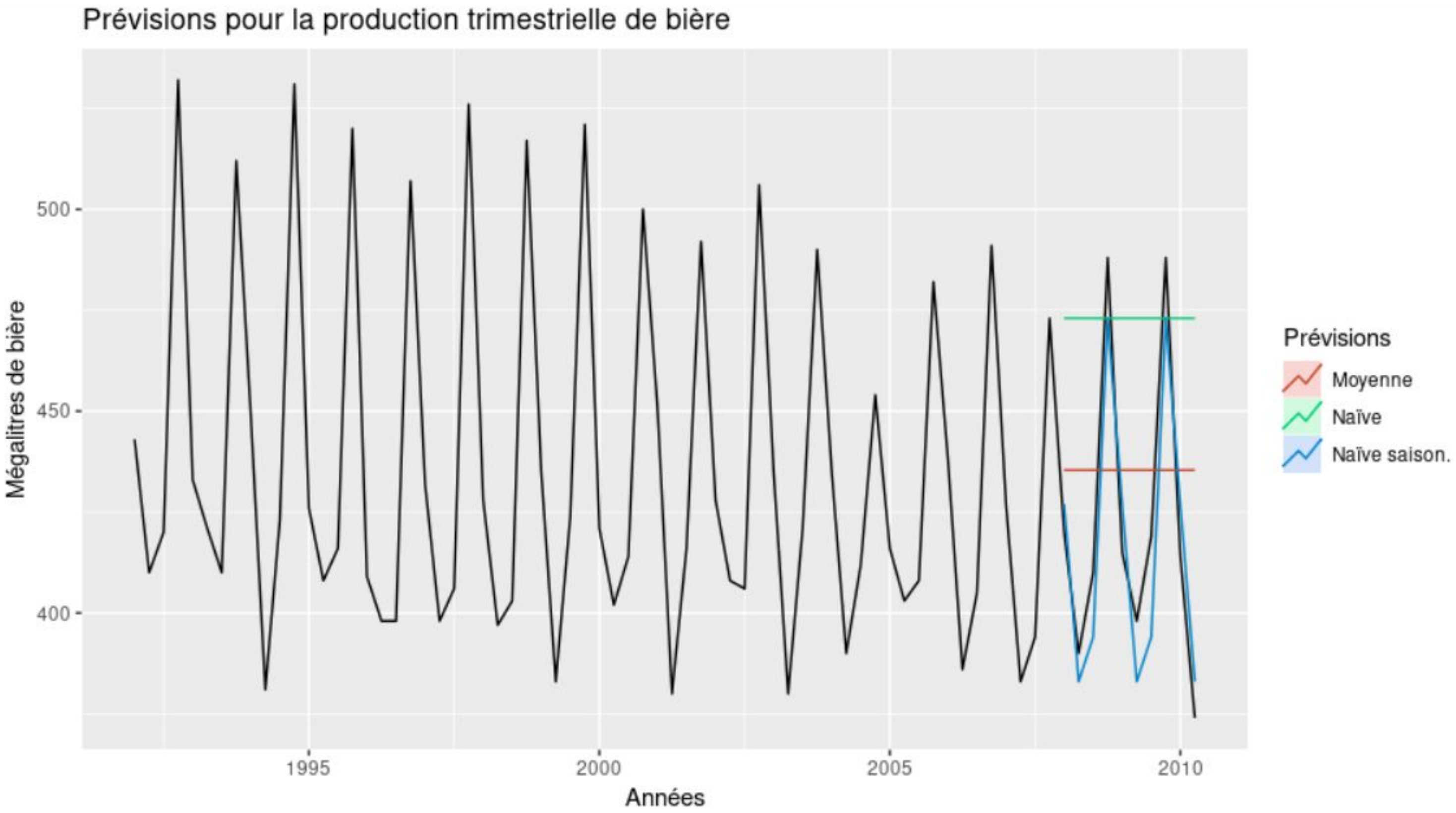
Mesures de précision des prévisions

- Exemple pour une **série non saisonnière**

$$q = \frac{1}{T-1} \sum_{t=2}^{T} |x_t - x_{t-1}|$$

- On remarque que les **MASE** équivaut au **MAE** pour la méthode naïve

Mesures de précision des prévisions



Mesures de précision des prévisions

- Code R

```
# Série bière et mesures de précision
biere <- window(ausbeer, start=1992, end=c(2007,4))
bierefit1 <- meanf(biere, h=10)
bierefit2 <- rwf(biere, h=10)
bierefit3 <- snaive(biere, h=10)
autoplot(window(biere, start=1992)) +
      autolayer(bierefit1, series="Moyenne", PI=FALSE) +
      autolayer(bierefit2, series="Naïve", PI=FALSE) +
      autolayer(bierefit3, series="Naïve saison.", PI=FALSE) +
      xlab("Années") + ylab("Mégalitres de bière") +
      ggtitle("Prévisions pour la production trimestrielle de bière") +
      guides(colour=guide_legend(title="Prévisions"))
```

Mesures de précision des prévisions

- Code R

```
biere3 <- window(biere, start=2008)
accuracy(bierefit1, biere3)
accuracy(bierefit2, biere3)
accuracy(bierefit3, biere3)
```

	MAE	MAPE	RMSE	MASE
Moyenne	34.83	8.28	38.45	2.44
Naïve	57.40	14.18	62.69	4.01
Naïve Saison.	13.40	3.17	14.31	0.94

Intervalles de précision

- Un intervalle de confiance est un intervalle dans lequel on espère que se situe $\mathbf{x}_{T+h}$ avec une certaine probabilité
- Si on suppose les **erreurs gaussiennes**, alors l'intervalle de prévision à 95 % sera de la forme

$$\hat{x}_{T+h} \pm 1.96 \times \hat{\sigma}_h$$

  où $\hat{\sigma}_h$ est l'écart-type associé à la loi pour **h** décalages

- Quand **h = 1**, $\hat{\sigma}_1$ peut être estimé via les résidus

Mesures de précision des prévisions

- Prévisions naïves avec un intervalle de prévision:

```
res_sd <- sqrt(mean(res^2, na.rm=TRUE))
naive(goog150, level=95)
```

	Point Forecast	Lo 95	Hi 95
1001	813.67	799.0449	828.2951
1002	813.67	792.9869	834.3530
1003	813.67	788.3385	839.0015
1004	813.67	784.4197	842.9202
1005	813.67	780.9672	846.3728
1006	813.67	777.8459	849.4941
1007	813.67	774.9755	852.3644
1008	813.67	772.3039	855.0361
1009	813.67	769.7946	857.5454
1010	813.67	767.4213	859.9187

Mesures de précision des prévisions

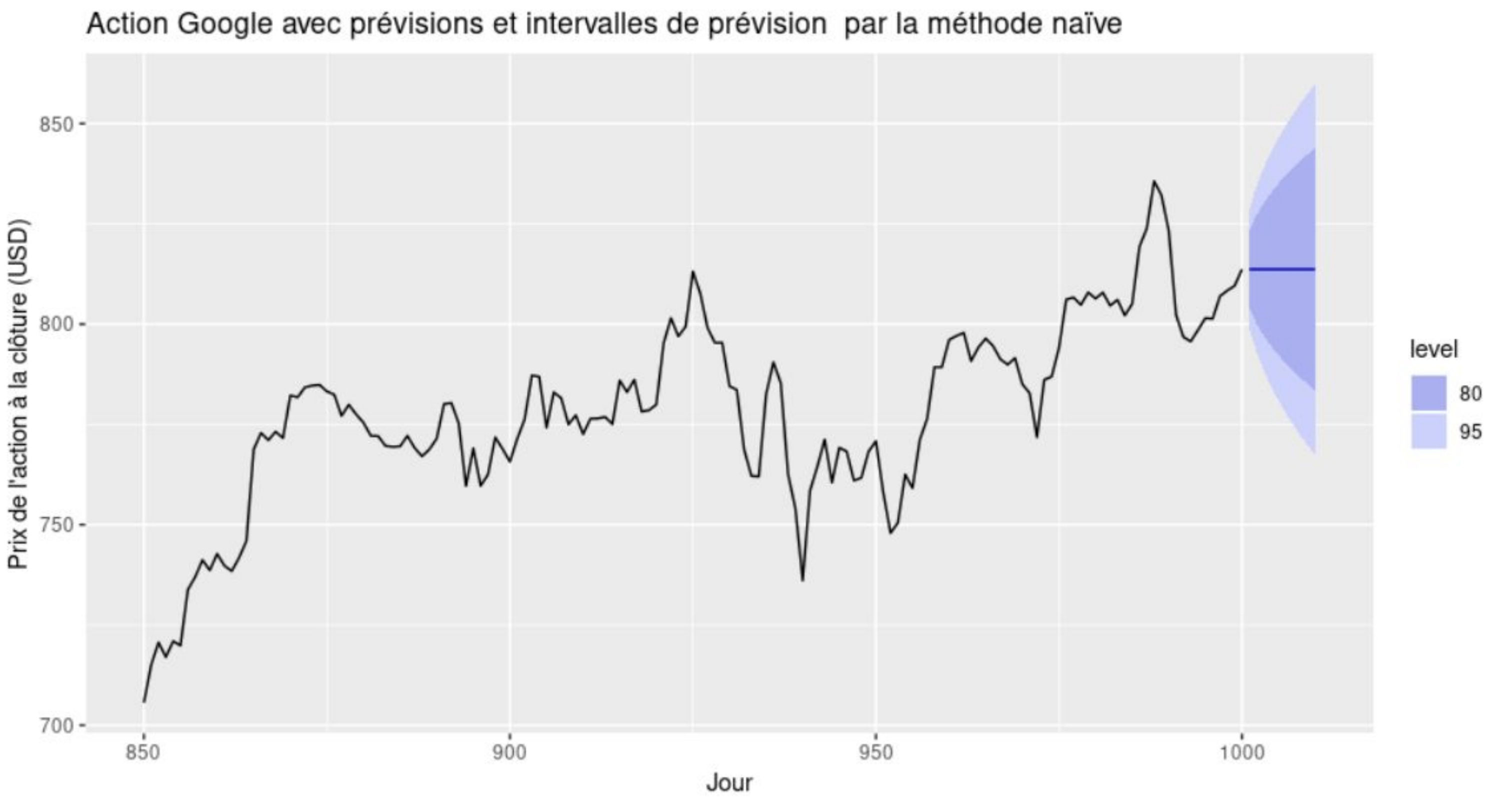
- Remarques
  - Il est très important d’avoir des intervalles de prévision, ce qui donne une information très pertinente sur l’efficacité des dites prévisions
  - Les intervalles de confiance requièrent que le modèle sous-jacent soit stochastique
  - Le calcul des prévisions à l’horizon **h** nécessitent une approche plus sophistiquée, avec des intervalles de confiance qui grossissent avec l’horizon **h** de prévision

Mesures de précision des prévisions

- Si les résidus sont gaussiens, non corrélés deux à deux, et d'écart-type $\hat{\sigma}$, alors on a
  - Méthode de la moyenne : $\hat{\sigma}_h = \hat{\sigma}\sqrt{1 + 1/T}$
  - Méthode naïve : $\hat{\sigma}_h = \hat{\sigma}\sqrt{h}$
  - Méthode naïve saison. : $\hat{\sigma}_h = \hat{\sigma}\sqrt{h}$
  - Méthode de la dérive : $\hat{\sigma}_h = \hat{\sigma}\sqrt{h(1 + h/T)}$

  où **k** est la partie entière de **(h – 1)/m**

Mesures de précision des prévisions



Pull de https://github.com/mswawola-cegep/420-a58-sf.git

04-03-TP

3 Références

Références

[1] Cours “R et la prévision de séries temporelles” de Michel Carbon - Université Laval