Université Moulay Ismail
Faculté des Sciences
Département de Mathématiques et Informatique

Prévisions
&
Modèles (non linéaire) mixtes

El Halimi Rachid

Lignes directrices

- **Introduction**

- **Méthodes de prévisions** (Quantitatives)

  - **Méthodes extrapolatives:** en fonction du passé,
    - **Méthode de lissage exponentiel.**

  - **Méthodes explicatives:** en fonction de variables,
    - **Modèles de régression**
    - **Modèles á effets mixtes**

Introduction

Prévision?:

- La prévision transforme les informations passées en estimation pour le futur.
- La prévision est l'activité où l'on cherche à calculer ou prédire un évènement futur, sur la base d'une analyse rationnelle. « Source: Webster's dictionary »

Classification des méthodes de prévision:

Qualitatives (subjectives: basées sur des appréciations ou des estimations)

- ✓ Peuvent dépendre ou non des données passées.
- ✓ Associées souvent au jugement d'un expert.
- ✓ Deux experts peuvent conclurent différemment. **Méthode**: Delphi

Quantitatives

- ✓ Ces prévisions reposent sur des modèles mathématiques et statistiques.

Introduction (suite)

- **Deux types de modèles:**
  - ✓ **Prévisions déterministes:**

Dans les modèles déterministes, la relation entre la variable d'intérêt **Y**, et les variables explicatives,

$$\mathbf{x} = \left(x_1, \ldots, x_p\right)^{\mathrm{T}}$$

est déterminée exactement par une relation du genre:

$$\mathrm{Y} = \mathrm{f}\left(\mathrm{x}, \beta\right)$$

où f est une fonction connue, et X et β sont des vecteurs de dimensions $p \times 1$ et $m \times 1$, respectivement

Introduction (suite)

Exemples: (les lois physiques)

1. Pour un objet de masse m, on sait que $F = ma$, c'est-à-dire que pour une accélération donnée a, on peut trouver exactement la force F.

2. La théorie de la chimie prédit que, pour un échantillon de gaz à température constante, la relation suivante est satisfaite: $\boldsymbol{pv^{\gamma} = c}$, où $\boldsymbol{p}$ est la pression et $\boldsymbol{v}$ le volume. Une fois que $\boldsymbol{c}$ et $\boldsymbol{\gamma}$ sont fixés, pour une pression donnée, on peut exactement trouver le volume.

Introduction (suite)

✓Prévisions probabilistes

- Dans les sciences sociales, les relations sont habituellement stochastiques.
- Un aspect aléatoire est présent, qui est dû souvent à des facteurs tels:
  - Erreurs de mesures.
  - Absence de variables plus ou moins importantes dans le modèle.
- On fait appel à des modèles de la forme

$$y = f\left(X_1, X_2, \ldots, X_p, \beta_1, \beta_2, \ldots, \beta_m\right) + \varepsilon$$

- où **ε** est le bruit ou la composante d'erreur (c'est une variable aléatoire possédant une loi de probabilité).

Méthodes extrapolatives

• Séries chronologiques

♦ Lorsque nous disposons d'une série de données espacées de manière égale dans le temps, on peut formuler un modèle du genre:

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

où $\boldsymbol{t}$ représente le temps; $\boldsymbol{u}_t$ est un bruit blanc (une suite de variables indépendantes centrée en **0** et possédant la même variance); et f est une fonction connue.

♦ En utilisant le passé de Y (i.e. Y_{t-1}, Y_{t-2},…), on tente d'extrapoler pour le futur.

Méthodes extrapolatives (suite)

Le lissage exponentiel

C'est une méthode qui prend en compte la prévision de la période antérieure, en utilisant l´ équation suivante:

$$Yp(t+1) = \alpha Y_t + (1-\alpha)\, Yp(t)$$

Yp(t+1) : valeur prévue pour la période t+1

Y_t : valeur réalisée pour la période t

Yp(t) : valeur prévue pour la période t

On peut écrire:

$$Yp(t+1) = \alpha Y_t + \alpha(1-\alpha) Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \ldots$$

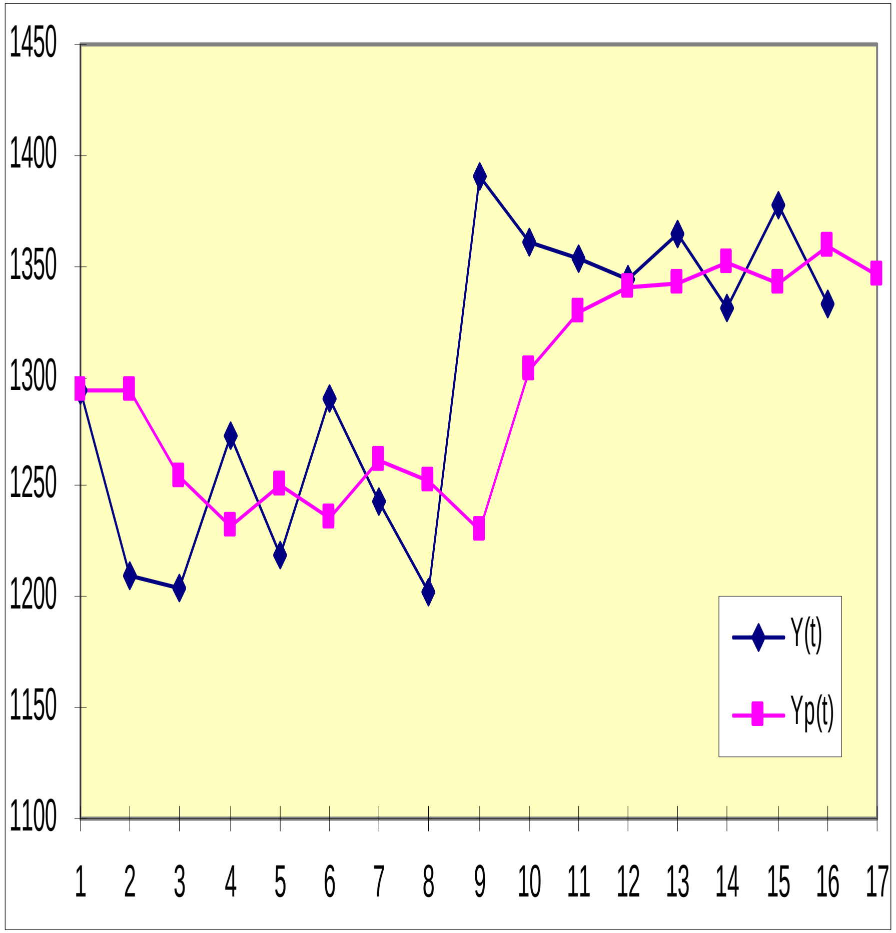
α est un paramètre de lissage comprise entre 0 et 1. Généralement choisi de façon à minimiser la somme des carrés des erreurs de prévision

Le lissage exponentiel(suite)

Ventes de savon d'une entreprise

		Y(t)	PrévisionYp(t)
1981	Janvier	1293	1293,00
	Février	1209	1293,00
	Mars	1205	1254,25
	Avril	1273	1231,53
	Mai	1220	1250,66
	Juin	1290	1236,52
	Juillet	1243	1261,19
	Août	1203	1252,80
	Septembre	1390	1229,83
	Octobre	1360	1303,71
	Novembre	1353	1329,68
	Décembre	1343	1340,43
1982	Janvier	1364	1341,62
	Février	1330	1351,94
	Mars	1377	1341,82
	Avril	1332	1358,05
			1346,03

Ces courbes montrent que le lissage exponentiel peut parfois atténuer très fortement les fluctuations observées dans la réalité



α=0.46; BIAS= 7.185, MSE= 3130

Méthodes explicatives

Régression:

- **Dans le cas général, la variable d' intérêt Y est expliquée par k variables explicatives:**

$$Y_t = a_0 + a_1 X_t^1 + a_1 X_t^2 + \ldots + a_k X_t^k + \varepsilon_t \quad (t=1,2,\ldots,T)$$

Estimation des (a_j): Les méthodes des moindres carrés, ….

- **La prévision de Y nécessite la connaissance (ou approximation) des X à l' instant (T+1):**

$$\hat{Y}_p(T+1) = \hat{a}_0 + \hat{a}_1 X_{T+1}^1 + \hat{a}_1 X_{T+1}^2 + \ldots + \hat{a}_k X_{T+1}^k + 0$$

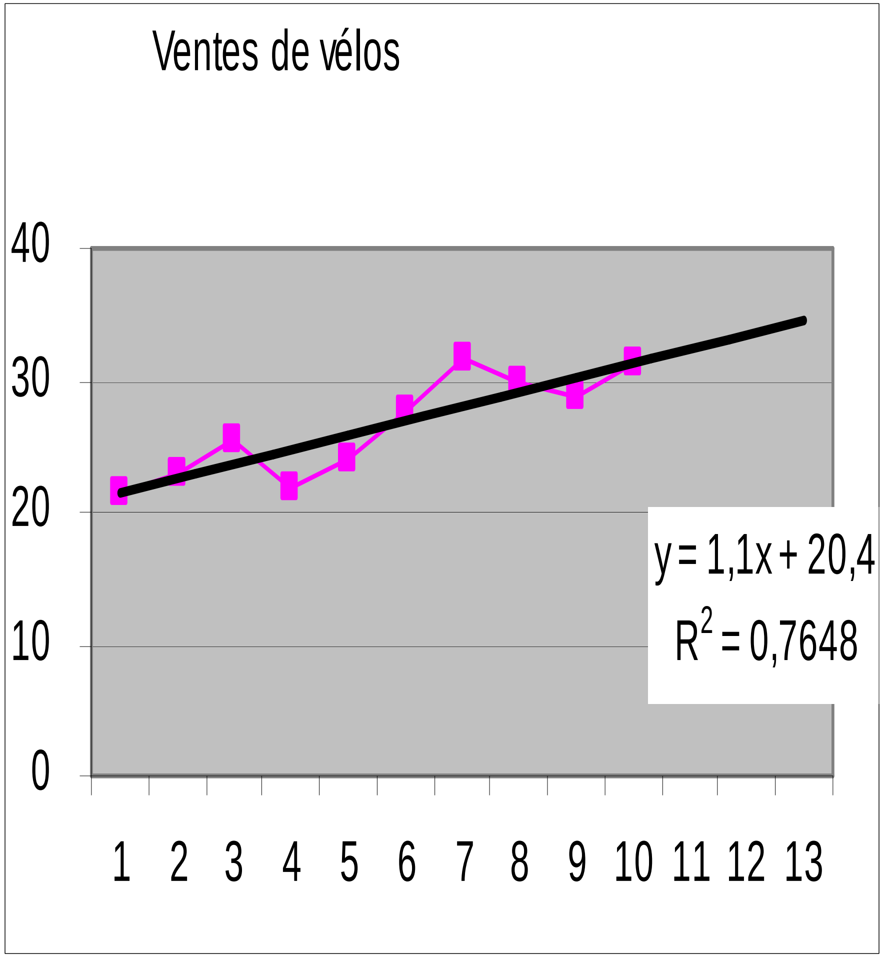
- **Exemple: Ventes de vélos**

Méthodes explicatives (exemple)

Ventes de vélos

Les données portent sur 10 années - les prévisions sont données pour les années 11, 12 et 13

Année	ventes Y_t	Prévision Yp(t)
1	**21,6**	21,5
2	**22,9**	22,6
3	**25,5**	23,7
4	**21,9**	24,8
5	**23,9**	25,9
6	**27,5**	27
7	**31,5**	28,1
8	**29,7**	29,2
9	**28,6**	30,3
10	**31,4**	31,4
11		32,5
12		33,6
13		34,7



Modèles mixtes : Prévision

- **Modèle mixtes générale (Linéaire)**

$$Y_i = f(X_i, Z_i, \beta, b_i) = X_i\beta + Z_i b_i + \varepsilon_i \ ,$$

$$i = 1, \cdots, M, \ b_i \sim N(0, B), \ \varepsilon_i \sim N(0, V_i),$$

- Prévision et effet aléatoire:

$$\begin{aligned}
\mathbb{E}(b_i|y) &= BZ_i'V_i^{-1}(y - X_i\beta). \\
\hat{b}_i &= BZ_i'V_i^{-1}(y - X_i\hat{\beta}) \\
\hat{\beta} &= (\sum_{i=1} X'V_i^{-1}X)^{-1} \sum_{i=1} X_i'V_i^{-1}y
\end{aligned}$$

- Prévision « individuelle »

$$y_{i(t+h)} = X_{(t+h)}\hat{\beta} + Z_{(t+h)}\hat{b}_i, \quad i = 1, \cdots, M$$

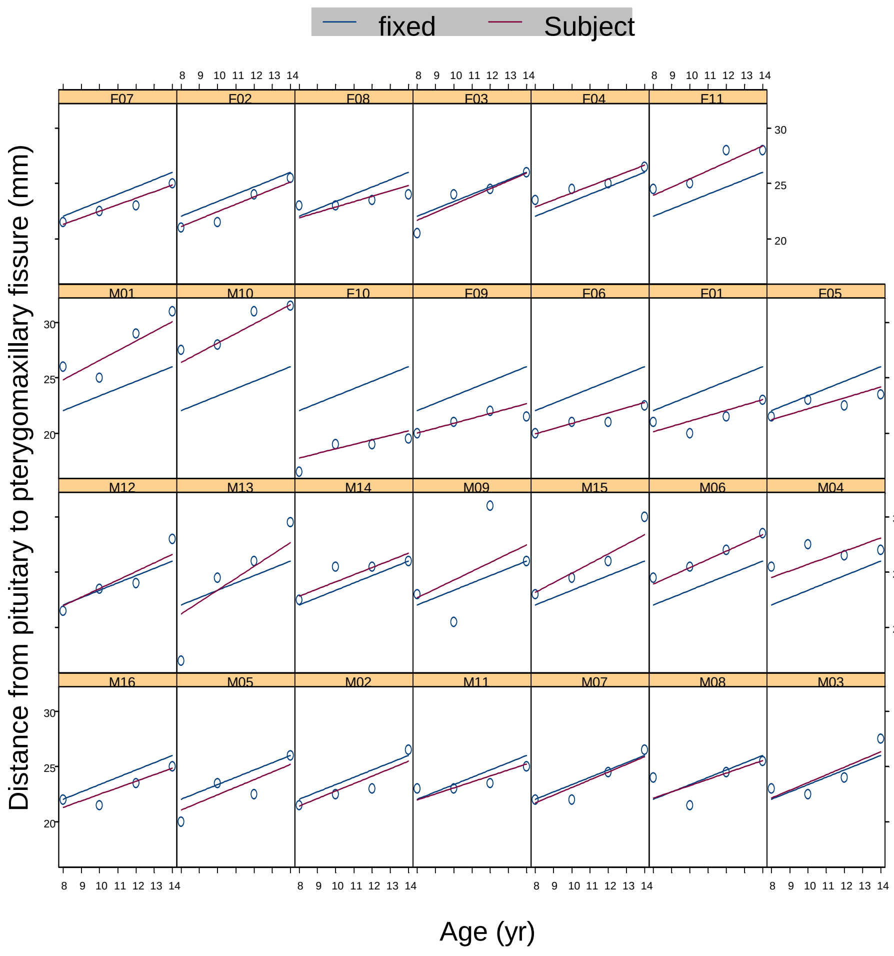
- Prévision « globale » $y_{(t+h)} = X_{(t+h)}\hat{\beta}$

Modèles mixtes : Prévision (Exemple)

$$d_{ij} = (\beta_0 + b_{i0}) + (\beta_1 + b_{i1})\text{age}_j + \varepsilon_{ij}$$

```
> print(Orthodont)

Grouped Data: distance ~ age | Subject
    distance age Subject    Sex
  1     26.0   8     M01   Male
  2     25.0  10     M01   Male
  3     29.0  12     M01   Male
  4     31.0  14     M01   Male
...
105     24.5   8     F11 Female
106     25.0  10     F11 Female
107     28.0  12     F11 Female
108     28.0  14     F11 Female
```



Bibliographie

- Bovas A., JOHANNES L.. **Statistical Methods for Forecasting**, *Wiley series*, **2005**.
- Davidian, M.. **Nonlinear mixed effects models**. International Encyclopedia of Statistical Science, M. *Lovric (ed). New York: Springer*, **2010**.
- El Halimi, R. **Nonlinear Mixed-effects Models and Bootstrap resampling** ; *Vdm Verlag Dr. Muller Aktiengesellschaft & Co. Kg (Germany)*, **2009.** *http://www.infibeam.com/Books/search?q=Rachid%20el%20halimi*