

**UNIVERSITÉ LAVAL**  
ÉCOLE D'ACTUARIAT

**ACT 2003**  
**Notes de cours**  
**Modèles linéaires en actuariat**

**David Beauchemin**

**Automne 2017**

© 2017 David Beauchemin



Cette création est mise à disposition selon le contrat [Attribution-Partage dans les mêmes conditions 4.0 International](#) de Creative Commons. En vertu de ce contrat, vous êtes libre de :

- **partager** — reproduire, distribuer et communiquer l'œuvre ;
- **remixer** — adapter l'œuvre ;
- utiliser cette œuvre à des fins commerciales.

Selon les conditions suivantes :



**Attribution** — Vous devez créditer l'œuvre, intégrer un lien vers le contrat et indiquer si des modifications ont été effectuées à l'œuvre. Vous devez indiquer ces informations par tous les moyens possibles, mais vous ne pouvez suggérer que l'offrant vous soutient ou soutient la façon dont vous avez utilisé son œuvre.



**Partage dans les mêmes conditions** — Dans le cas où vous modifiez, transformez ou créez à partir du matériel composant l'œuvre originale, vous devez diffuser l'œuvre modifiée dans les mêmes conditions, c'est-à-dire avec le même contrat avec lequel l'œuvre originale a été diffusée.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Régression linéaire simple</b>	<b>2</b>
2.1	Introduction . . . . .	2
2.1.1	Régression linéaire simple . . . . .	3
2.1.2	Régression linéaire multiple . . . . .	3
2.1.3	Régression exponentielle . . . . .	5
2.1.4	Régression quadratique . . . . .	5

## Résumé

*abstrat*

# Chapitre 1

## Introduction

L'établissement de prévisions joue un rôle central dans notre vie de tous les jours (prévisions météorologique, horoscope, etc.), et plus particulièrement dans celle des actuaires.

### Objectifs de la régression

Régulièrement en actuariat, on se questionne sur les effets de différentes variables sur d'autres. Par exemple,

- Quel est l'effet de l'âge sur la fréquence des sinistres automobiles ?
- Quel est l'effet du sexe sur la mortalité ?

On cherche à étudier et déterminer les relations entre des variables mesurables à partir de données.

### Deux grandes classes de variables mesurables :

- Qualitatives : basées sur des opinions et/ou des intuitions.
- Quantitatives : basées sur des observations, un modèle et des arguments mathématiques.

### Deux *grandes étapes* pour établir des prévisions quantitatives

1. Bâtir le modèle et estimer les paramètres :  
ex :  $F = M \times a$  Qui représente un modèle déterministe  
ex :  $Y = 3 \times X + 6 + \epsilon_t$  ; où  $\epsilon_t \sim N(0, 10)$  Qui représente un modèle probabiliste
2. Calculer les prévisions à partir du modèle.

Dans le cadre du cours, seulement les modèles probabilistes linéaires seront étudiés.

## Chapitre 2

# Régression linéaire simple

### 2.1 Introduction

De façon générale, en régression, nous avons :

$Y$	Variable dépendante, ou de réponse	Output
$X_1, X_2, \dots, X_n$	Soit $n$ variables indépendantes ou explicatives, ou exogènes <sup>1</sup>	Input
$\beta_0, \beta_1, \dots, \beta_n$	Les paramètres à estimer	

Voici une illustration du concept de régression linéaire

#### Étape 1

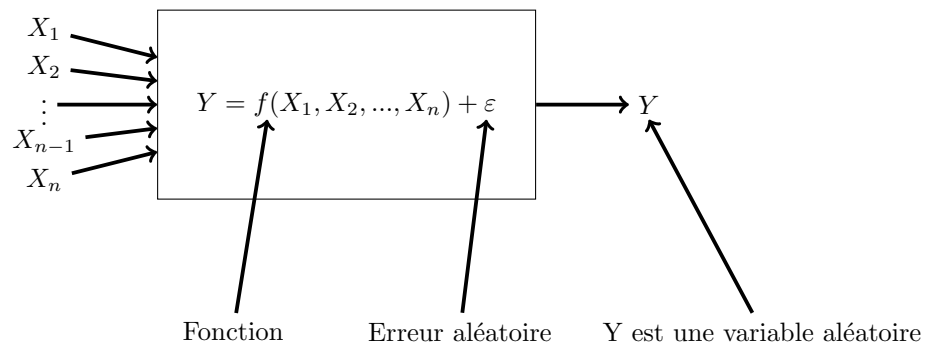
Observation  
des  $X_i$

#### Étape 2

Modèle de ré-  
gression

#### Étape 3

Prévision de  $Y$



1. Les variables  $X_i$  sont indépendante par rapport à  $y$ , mais pas nécessairement entre elles.

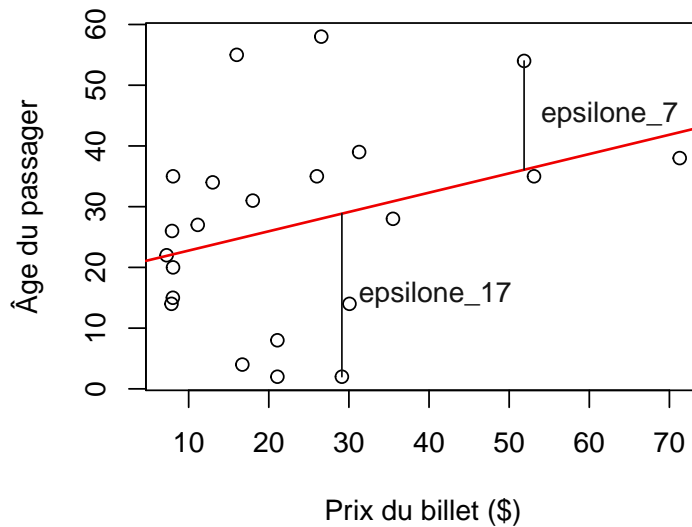
### 2.1.1 Regression linéaire simple

On cherche à prédire l'âge des passagers du Titanic selon le prix du billet à l'aide du modèle linéaire suivant,

$$Y = \beta_0 + \beta_1 \times X + \varepsilon$$

↑                      ↗                      ↖  
Âge du passa-      Prix du billet      Erreur aléa-  
ger                      toire

#### Âge prédit des passagers du Titanic



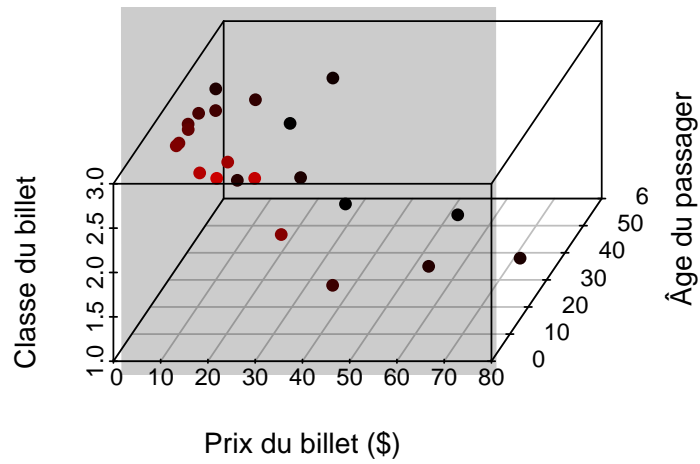
### 2.1.2 Regression linéaire multiple

On cherche à prédire l'âge des passagers du Titanic selon le prix du billet et son sexe à l'aide du modèle linéaire suivant,

$$Y = \beta_0 + \beta_1 \times X_1 + \beta_2 \times X_2 + \varepsilon$$

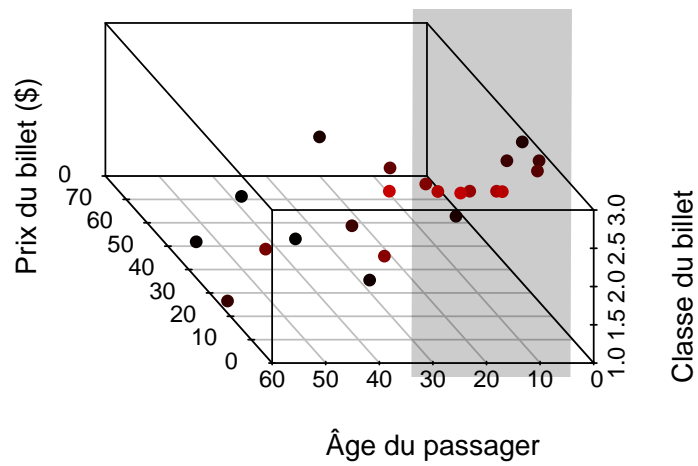
↑                      ↑                      ↑                      ↖  
Âge du passa-      Prix du billet      Sexe du passa-      Erreur aléa-  
ger                      ger                      toire

### Âge predict des passagers du Titanic



Voici la régression sous un autre angle, on voit la surface plane de régression.

### Âge predict des passagers du Titanic





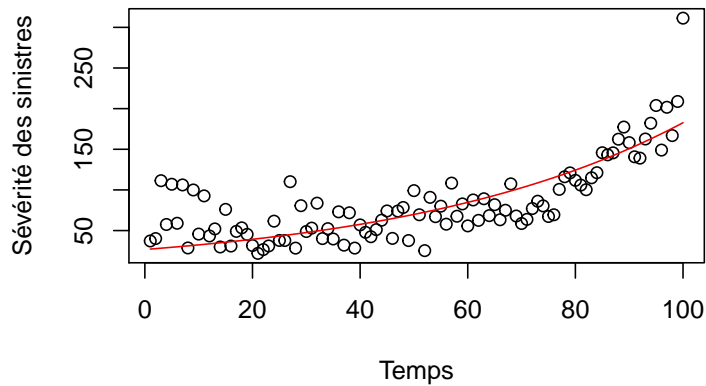
### 2.1.3 Régression exponentielle

On cherche à prédire la sévérité d'un sinistre automobile en fonction du temps à l'aide du modèle exponentielle suivant,

$$Y = \beta_0 \times e^{\beta_1 \times X} \times \varepsilon$$

↑                      ↑                      ↑  
Sévérité du    Temps                      Erreur aléa-  
sinistre                      toire

#### Modèle de prédiction de la sévérité des sinistres



#### Note

On remarque que la régression exponentielle est similaire à une régression linéaire simple.

$$\ln(Y) = \ln(\beta_0) + \beta_1 \times X + \ln(\varepsilon)$$
$$Y^* = \beta_0^* + \beta_1 \times X + \varepsilon^*$$

Qu'on appel aussi une régression multiplicative ou log-linéaire.

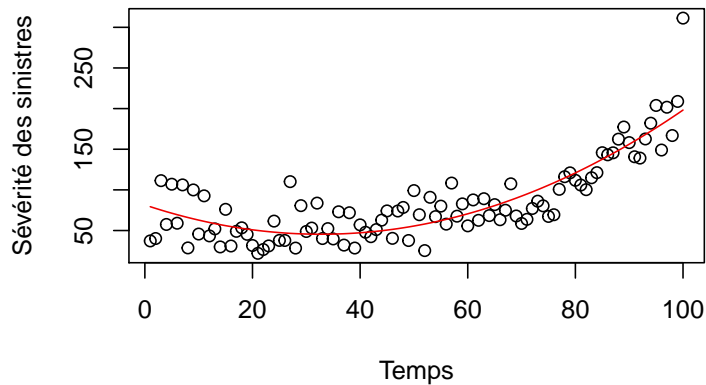
### 2.1.4 Régression quadratique

On cherche à prédire la sévérité d'un sinistre automobile en fonction du temps et du temps au carré à l'aide du modèle quadratique suivant,

$$Y = \beta_0 + \beta_1 \times X + \beta_2 \times X^2 + \varepsilon$$

$\uparrow$  Sévérité du sinistre       $\nwarrow$  Temps       $\nearrow$  Erreur aléatoire

### Modèle de prédiction de la sévérité des sinistres



#### Note

On remarque que la régression quadratique est similaire à une régression linéaire multiple. En posant  $X_1 = X$  et  $X_2 = X^2$

$$Y = \beta_0 + \beta_1 \times X_1 + \beta_2 \times X_2 + \varepsilon$$

Soit une régression linéaire multiple.

Dans le cadre du cours, seulement les modèles linéaires seront à l'étude car,

- Plus simples
- Plusieurs modèles peuvent se ramener à un modèle linéaire simple ou multiple. (voir 2.1.3 et 2.1.4)
- Constituent souvent une très bonne approximation de la réalité qui peut être très complexe, tel que l'assurance.
- Se généralisent facilement, tel que les *Generalized Linear Models*.

Le problème principal de la modélisation linéaire est de trouver les différents paramètres  $\beta_0, \beta_1, \dots, \beta_n$  de telle sorte que

$$\varepsilon = Y - f(X_1, \dots, X_n; \beta_0, \beta_1, \dots, \beta_n) \quad (2.1)$$

soit minimiser.