

---

# Mathematical modeling of behavior

Michel Bierlaire

`michel.bierlaire@epfl.ch`

Transport and Mobility Laboratory



# Introduction

---

- What kind of behavior can be mathematically modeled?



# Introduction

---

## Psychohistory

*Branch of mathematics which deals with the reactions of human conglomerates to fixed social and economic stimuli. The necessary size of such a conglomerate may be determined by Seldon's First Theorem.*

*Encyclopedia Galactica, 116th Edition (1020 F.E.)  
Encyclopedia Galactica Publishing Co., Terminus*

Motivation: shorten the period of barbarism after  
the Fall of the Galactic Empire

Asimov, I. (1951) *Foundation*, Gnome Press



# In this course...

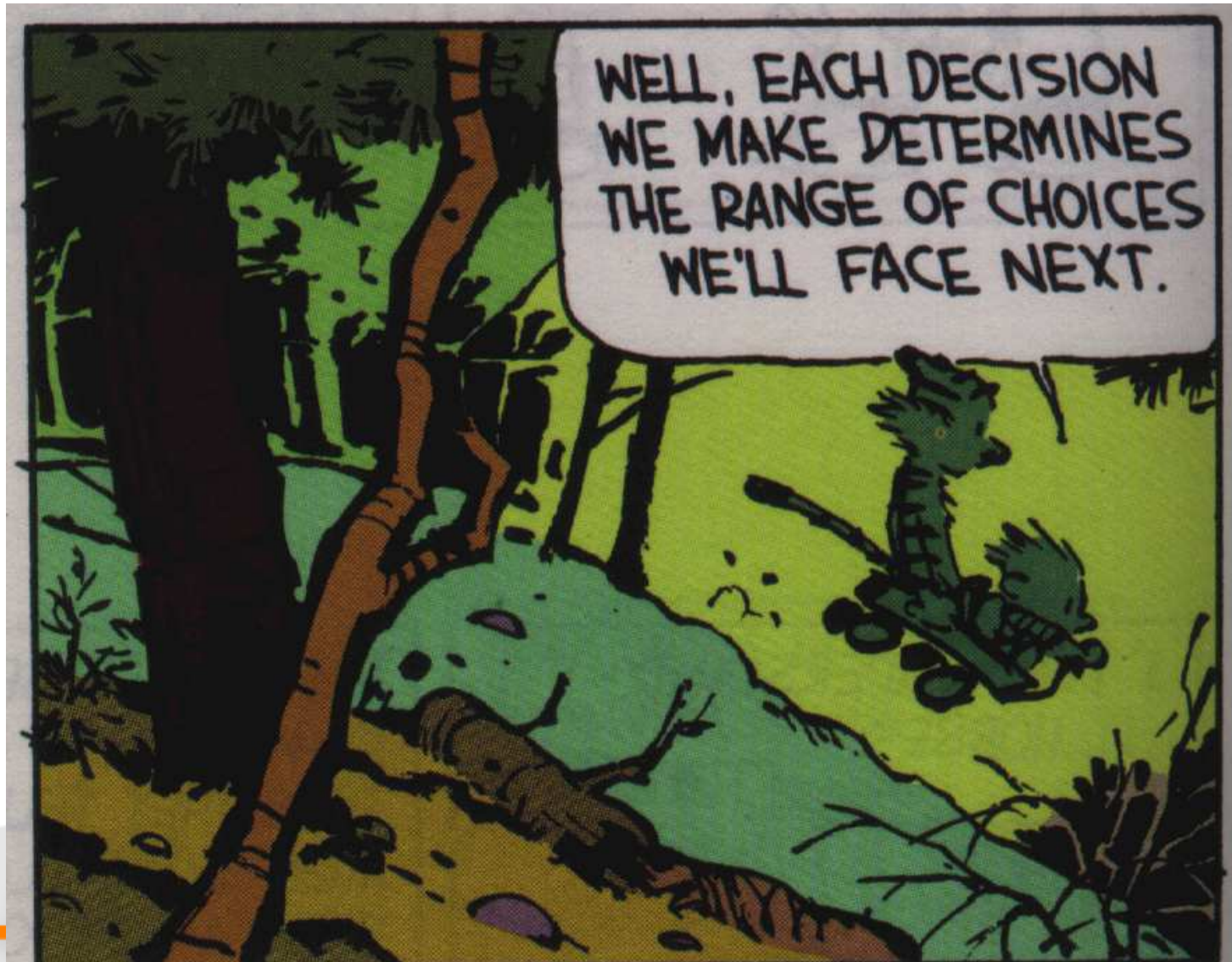
---

- Individual behavior (vs. aggregate behavior)
- Theory of behavior which is
  - **descriptive**: how people behave and not how they should
  - **abstract**: not too specific
  - **operational**: can be used in practice for forecasting
- Type of behavior: **choice**



# Motivations

---





# Motivations

---

“It is our choices that show what we truly are, far more than our abilities” Albus Dumbledore

“Liberty, taking the word in its concrete sense, consists in the ability to choose.” Simone Weil (French philosopher, 1909-1943)

## Field :

- ▶ Marketing
- ▶ Transportation
- ▶ Politics
- ▶ Management
- ▶ New technologies

## Type of behavior:

- ▶ Choice of a brand
- ▶ Choice of a transportation mode
- ▶ Choice of a president
- ▶ Choice of a management policy
- ▶ Choice of investments



# Applications

---

## Case studies

- Choice-lab marketing
  - Context: B2B, data provider (financial, demographic, etc.)
  - Objective: understand why clients quit
- Quebec energy
  - Context: space and water heating in households
  - Objective: importance of the type of household and price
- Transportation mode choice in the Netherlands
  - Context: car vs rail in Nijmegen
  - Objective: sensitivity to travel time and cost, inertia.



# Applications

---

- Swissmetro
  - Context: new transportation technology
  - Objective: demand pattern, pricing
- Residential telephone services
  - Context: flat rate vs. measured
  - Objective: offer the most appropriate service
- Airline itinerary choice
  - Context: questionnaire about itineraries across the US
  - Objective: help airlines and aircraft manufacturer to design a better offer



# Importance

---

Daniel

L.

McFadden



1937–

- UC Berkeley 1963, MIT 1977, UC Berkeley 1991
- Laureate of *The Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel 2000*
- Owns a farm and vineyard in Napa Valley
- “Farm work clears the mind, and the vineyard is a great place to prove theorems”



# Example

---

## Voice over internet protocol (VoIP)

- What is the market penetration?
- How will the penetration change in the future?
- Assumption: level of education is an important explanatory factor

## Data collection

- sample of 600 persons, randomly selected
- Two questions:
  1. Do you subscribe to voice over IP? (yes/no)
  2. How many years of education have you had? (low/medium/high)



# Example

- Contingency table

VoIP	Education			
	Low	Medium	High	
Yes	10	100	120	230
No	140	200	30	370
	150	300	150	600

- Penetration in the sample:  $230/600 = 38.3\%$
- Forecasting: need for a model



# Example: a model

---

Type of variables:

- dependent, or endogenous: what we explain
- independent, exogenous or explanatory: how we explain

Model:

- Causal relationship between the independent and dependent variables
- Based on theory, assumptions.
- Probabilistic.



# Example: a model

---

- Dependent variable:

$$y = \begin{cases} 1 & \text{if subscriber} \\ 2 & \text{if not subscriber} \end{cases}$$

Discrete dependent variable

- Independent or explanatory variable

$$x = \begin{cases} 1 & \text{if level of education is low} \\ 2 & \text{if level of education is medium} \\ 3 & \text{if level of education is high} \end{cases}$$



# Example: a model

- Market penetration in the sample:  $\hat{p}(y = 1)$
- Market penetration in the population:  $p(y = 1)$  estimated by  $\hat{p}(y = 1)$
- Joint probabilities:  $\hat{p}(y = 1, x = 2) = 100/600 = 0.1667$
- Marginal probabilities:  $\hat{p}(y = 1) = \sum_{k=1}^3 \hat{p}(1, k) = 10/600 + 100/600 + 120/600 = 0.383$
- Conditional probabilities:  $\hat{p}(y = 1|x = 2)$

$$\begin{aligned}\hat{p}(y = 1, x = 2) &= \hat{p}(y = 1|x = 2)\hat{p}(x = 2) \\ \hat{p}(y = 1|x = 2) &= \hat{p}(y = 1, x = 2)/\hat{p}(x = 2) \\ &= 0.1667/0.5 = 0.333\end{aligned}$$



# Example: a model

---

Similarly, we obtain

$$\hat{p}(y = 1|x = 1) = 0.067$$

$$\hat{p}(y = 1|x = 2) = 0.333$$

$$\hat{p}(y = 1|x = 3) = 0.8$$

We obtain a causal relationship.

- Behavioral model:  $\hat{p}(y = i|x = j)$
- Forecasting assumption: stable over time



# Example: forecasting

- Model:

$$p(y = 1|x = 1) = \pi_1 = 0.067$$

$$p(y = 1|x = 2) = \pi_2 = 0.333$$

$$p(y = 1|x = 3) = \pi_3 = 0.8$$

where  $\pi_1, \pi_2, \pi_3$  are estimated parameters

- Assumption: future level of education: 10%-60%-30%

$$\begin{aligned} p(y = 1) &= \sum_{i=1}^3 p(y = 1|x = i)p(x = i) \\ &= 0.1\pi_1 + 0.6\pi_2 + 0.3\pi_3 \\ &= 44.67\% \end{aligned}$$



# Example: forecasting

---

- If the level of education increases
- from 25%-50%-25% to 10%-60%-30%
- Market penetration of VoIP will increase
- from 38.33 % to 44.67%

In summary

- $p(x = j)$  can be easily obtained and forecast
- $p(y = i|x)$  is the behavioral model to be developed



# Outline

---

- Introduction and examples
- Review of relevant concepts in probability and statistics
- Choice theory
- Binary choice
- Multiple alternatives
- Tests
- Nested Logit model
- Multivariate Extreme Value models
- Forecasting
- Sampling
- Mixtures of models
- Latent variables