Latent variables

Michel Bierlaire

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne

Outline

Motivation

Rationality?

- Standard random utility assumptions are often violated.
- Factors such as attitudes, perceptions, knowledge are not reflected.

Example: pain lovers

Kahneman, D., Fredrickson, B., Schreiber, C.M., and Redelmeier, D., When More Pain Is Preferred to Less: Adding a Better End, Psychological Science, Vol. 4, No. 6, pp. 401-405, 1993.

- Short trial: immerse one hand in water at 14° for 60 sec.
- Long trial: immerse the other hand at 14° for 60 sec, then keep the hand in the water 30 sec. longer as the temperature of the water is gradually raised to 15°.
- Outcome: most people prefer the long trial.
- Explanation:
  - duration plays a small role
  - the peak and the final moments matter

Example: The Economist

Subscription to The Economist

Web only	@ \$59
Print only	@ \$125
Print and web	@ \$125

Example: The Economist

Subscription to The Economist

Experiment 1	Experiment 2
Web only @ $59	Web only @ $59
Print only @ $125	
Print and web @ $125	Print and web @ $125

Example: The Economist

Subscription to The Economist

	Experiment 1	Experiment 2	
16	Web only @ $59	Web only @ $59	68
0	Print only @ $125		
84	Print and web @ $125	Print and web @ $125	32

Source: Ariely (2008)

- Dominated alternative
- According to utility maximization, should not affect the choice
- But it affects the perception, which affects the choice.

Example: good or bad wine?

Choose a bottle of wine...

	Experiment 1	Experiment 2
1	McFadden red at \$10	McFadden red at \$10
2	Nappa red at \$12	Nappa red at \$12
3		McFadden special reserve pinot noir at \$60
	Most would choose 2	Most would choose 1

- Context plays a role on perceptions

Example: live and let die

Population of 600 is threatened by a disease. Two alternative treatments to combat the disease have been proposed.

Experiment 1 # resp. = 152	Experiment 2 # resp. = 155
Treatment A: 200 people saved	Treatment C: 400 people die
Treatment B: 600 people saved with prob. 1/3 0 people saved with prob. 2/3	Treatment D: 0 people die with prob. 1/3 600 people die with prob. 2/3

Example: live and let die

Population of 600 is threatened by a disease. Two alternative treatments to combat the disease have been proposed.

	Experiment 1 # resp. = 152	Experiment 2 # resp. = 155	
72%	Treatment A: 200 people saved	Treatment C: 400 people die	22%
28%	Treatment B: 600 people saved with prob. 1/3 0 people saved with prob. 2/3	Treatment D: 0 people die with prob. 1/3 600 people die with prob. 2/3	78%

Source: Tversky & Kahneman (1986)

Example: to be free

Choice between a fine and a regular chocolate

	Experiment 1	Experiment 2
Lindt	\$0.15	\$0.14
Hershey	\$0.01	\$0.00
Lindt chosen	73%	31%
Hershey chosen	27%	69%

Source: Ariely (2008) Predictably irrational, Harper Collins.

Outline

Latent concepts

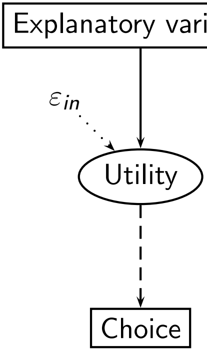
Latent

- **latent**: potentially existing but not presently evident or realized (from Latin: lateo = lie hidden)
- Here: not directly observed
- Standard models are already based on a latent concept: utility

Drawing convention

- Latent variable
- Observed variable
- structural relation: $\longrightarrow$
- measurement: - - - →
- errors: · · · · · · →

Random utility



$$V_{in} = \sum_k \beta_{ik} x_{ikn}$$

$$P_n(i) = e^{V_{in}} / \sum_j e^{V_{jn}}$$

Attitudes

Measuring attitudes

- Psychometric indicators
- Example: attitude towards the environment.
- For each question, response on a scale: strongly agree, agree, neutral, disagree, strongly disagree, no idea.
  - The price of oil should be increased to reduce congestion and pollution
  - More public transportation is necessary, even if it means additional taxes
  - Ecology is a threat to minorities and small companies.
  - People and employment are more important than the environment.
  - I feel concerned by the global warming.
  - Decisions must be taken to reduce the greenhouse gas emission.

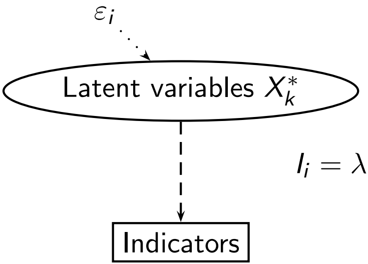
Indicators

Model specification

Indicators cannot be used as explanatory variables. Mainly two reasons:

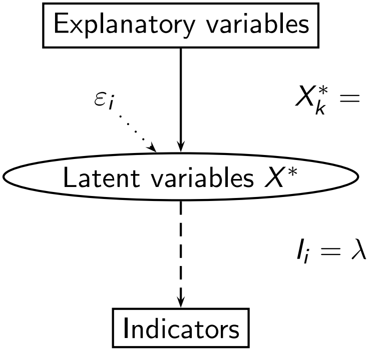
1. Measurement errors
   - Scale is arbitrary and discrete
   - People may overreact
   - Justification bias may produce exaggerated responses
2. No forecasting possibility
   - No way to predict the indicators in the future

Factor analysis



$$I_i = \lambda_i + \sum_k L_{ik} X_k^*$$

Measurement equation



$$X_k^* = \sum_j \beta_j x_j$$

$$I_i = \lambda_i + \sum_k L_{ik} X_k^*$$

Measurement equation

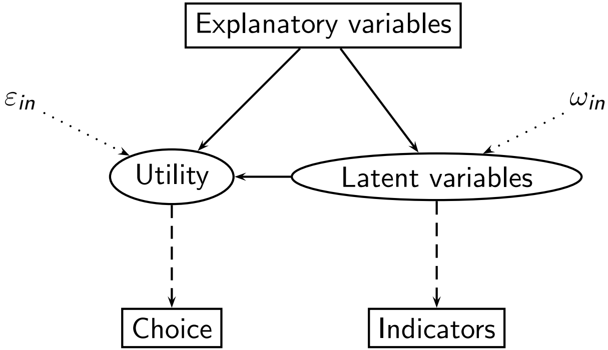
Continuous model: regression

$$I = f(X^*; \beta) + \varepsilon$$

Discrete model: thresholds

$$I = \begin{cases} 1 & \text{if } -\infty < X^* \leq \tau_1 \\ 2 & \text{if } \tau_1 < X^* \leq \tau_2 \\ 3 & \text{if } \tau_2 < X^* \leq \tau_3 \\ 4 & \text{if } \tau_3 < X^* \leq \tau_4 \\ 5 & \text{if } \tau_4 < X^* \leq +\infty \end{cases}$$

Choice model



Outline

Estimation: likelihood

Structural equations

1. Distribution of the latent variables:

$$f_1(X_n^*|X_n; \lambda, \Sigma_\omega)$$

For instance

$$X_n^* = h(X_n; \lambda) + \omega_n, \quad \omega_n \sim N(0, \Sigma_\omega).$$

2. Distribution of the utilities:

$$f_2(U_n|X_n, X_n^*; \beta, \Sigma_\varepsilon)$$

For instance

$$U_n = V(X_n, X_n^*; \beta) + \varepsilon_n, \quad \varepsilon_n \sim N(0, \Sigma_\omega).$$

Estimation: likelihood

Measurement equations

1. Distribution of the indicators:

$$f_3(I_n|X_n, X_n^*; \alpha, \Sigma_\nu)$$

   For instance:

$$I_n = m(X_n, X_n^*; \alpha) + \nu_n, \quad \nu_n \sim N(0, \Sigma_\nu).$$

2. Distribution of the observed choice:

$$P(y_{in} = 1) = \Pr(U_{in} \geq U_{jn}, \forall j).$$

Indicators: continuous output

$$f_3(I_n|X_n, X_n^*; \alpha, \Sigma_\nu)$$

For instance:

$$I_n = m(X_n, X_n^*; \alpha) + \nu_n, \quad \nu_n \sim N(0, \sigma_{\nu_n}^2)$$

So,

$$f_3(I_n|\cdot) = \frac{1}{\sigma_{\nu_n}\sqrt{2\pi}} \exp\left(-\frac{(I_n - m(\cdot))^2}{2\sigma_{\nu_n}^2}\right)$$

Define

$$Z = \frac{I_n - m(\cdot)}{\sigma_{\nu_n}} \sim N(0,1), \quad \phi(Z) = \frac{1}{\sqrt{2\pi}} e^{-Z^2/2}$$

and

$$f_3(I_n|\cdot) = \frac{1}{\sigma_{\nu_n}} \phi(Z)$$

Indicators: discrete output

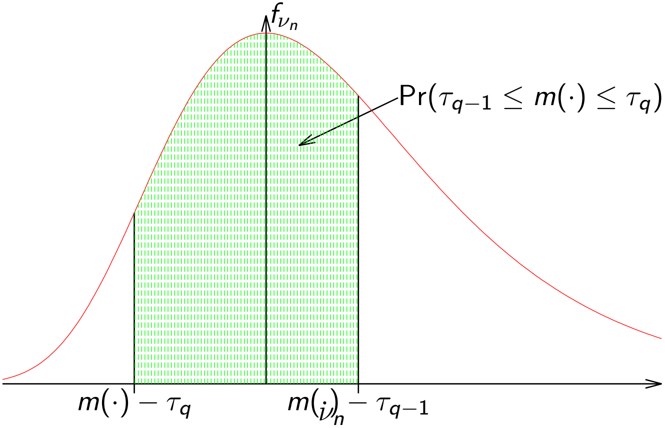
$$f_3(I_n|X_n, X_n^*; \alpha, \Sigma_\nu)$$

For instance:

$$I_n = m(X_n, X_n^*; \alpha) + \nu_n, \quad \nu_n \sim \text{Logistic(0,1)}$$

$$\begin{aligned}
P(I_n = 1) &= & \Pr(m(\cdot) \leq \tau_1) &= \frac{1}{1+e^{-\tau_1+m(\cdot)}} \\
P(I_n = 2) &= \Pr(m(\cdot) \leq \tau_2) - & \Pr(m(\cdot) \leq \tau_1) &= \frac{1}{1+e^{-\tau_2+m(\cdot)}} - \frac{1}{1+e^{-\tau_1+m(\cdot)}} \\
&\vdots & &\vdots \\
P(I_n = 5) &= & 1 - \Pr(m(\cdot) \leq \tau_4) &= 1 - \frac{1}{1+e^{-\tau_4+m(\cdot)}}
\end{aligned}$$

Indicators: discrete output



Estimation: likelihood

Assuming ω_n, ε_n and ν_n are independent, we have

$$\mathcal{L}_n(y_n, I_n | X_n; \alpha, \beta, \lambda, \Sigma_\varepsilon, \Sigma_\nu, \Sigma_\omega) =$$

$$\int_{X^*} P(y_n | X_n, X^*; \beta, \Sigma_\varepsilon) f_3(I_n | X_n, X^*; \alpha, \Sigma_\nu) f_1(X^* | X_n; \lambda, \Sigma_\omega) dX^*.$$

Maximum likelihood estimation:

$$\max_{\alpha,\beta,\lambda,\Sigma_\varepsilon,\Sigma_\nu,\Sigma_\omega} \sum_n \log \left(\mathcal{L}_n(y_n, I_n | X_n; \alpha, \beta, \lambda, \Sigma_\varepsilon, \Sigma_\nu, \Sigma_\omega) \right)$$

Source: Walker (2001)

Outline

Case studies

Walker (2001)

- Mode choice
- Latent variables:
  - Ride comfort
  - Convenience
- Indicators: (from "very poor" to "very good")
  - Relaxation during the trip
  - Reliability of the arrival time
  - Flexibility of choosing departure time
  - Ease of traveling with children
  - Safety during the trip
  - Overall rating of the mode

Case studies

Walker (2001)

- Employees' adoption of telecommuting
- Latent variables:
  - Perceived costs
    - Impact on your expenditures on home utilities
    - Impact on your expenditures on child care
    - Impact on your expenditures on elder care
    - Impact on your expenditures on overall working costs
  - Benefits
    - Impact on your schedule flexibility
    - Impact on your productivity
    - Impact on your autonomy in your job
    - Impact on the productivity of the group you work with
    - Impact on your family life
    - Impact on your social life
    - etc.

Case study: value of time

Effect of attitude on value of time

- SP survey, Stockholm, Sweden, 2005
- 2400 households surveyed
- Married couples with both husband and wife working or studying
- Choice between car alternatives
- Data used: 554 respondents, 2216 SP responses
- Attributes:
  - travel time
  - travel cost
  - number of speed cameras

Attitudinal questions

Statements

- It feels safe to go by car.
- It is comfortable to go by car to work.
- It is very important that traffic speed limits are not violated.
- Increase the motorway speed limit to 140 km/h.

Likert scale

- 1: do not agree at all
- 5: do fully agree

Structural models

Attitude model, capturing the positive attitude towards car

$$
\begin{aligned}
\text{Attitude} = \quad & \theta_0 \cdot 1 && \text{(intercept)} \\
& + \theta_f \cdot \text{female} && \\
& + \theta_{\text{inc}} \cdot \text{income} && \text{(monthly, in Kronas)} \\
& + \theta_{\text{age1}} \cdot (\text{Age} < 55) && \\
& + \theta_{\text{age2}} \cdot (\text{Age } 55\text{–}65) && \\
& + \theta_{\text{age3}} \cdot (\text{Age} > 65) && \\
& + \theta_{\text{edu1}} \cdot (\text{basic/pre high school}) && \\
& + \theta_{\text{edu2}} \cdot (\text{university}) && \\
& + \theta_{\text{edu3}} \cdot (\text{other}) && \\
& \sigma \cdot \omega && \text{(normal error term)}
\end{aligned}
$$

Structural models

Choice model: 3 alternatives

- Car on route 1
- Car on route 2
- Indifferent (utility = 0)

$$
\begin{aligned}
\text{Utility}_i = \quad & \beta_i & \text{(ASC)} \\
& + \beta_{\text{t}} \cdot \text{travel time}_i & \\
& + \beta_{\text{c}} \cdot \text{cost}_i \;/\; \text{Income} & \\
& + \gamma \cdot \text{cost}_i \cdot \text{Attitude} \;/\; \text{Income} & \\
& + \beta_{\text{cam}} \cdot \#\ \text{cameras}_i & \\
& + \varepsilon_i & \text{(EV error term)}
\end{aligned}
$$

Note: standard model obtained with $\gamma = 0$.

Value of time

Model without attitude variable ($\gamma = 0$)

$$\text{VOT} = \frac{\beta_\text{t}}{\beta_\text{c}} * Income$$

Model with attitude variable

$$\text{VOT} = \frac{\beta_\text{t}}{\beta_\text{c} + \gamma \cdot \text{Attitude}} * Income$$

Note: distributed

Measurement equations

Choice

$$y_i = \begin{cases} 1 & \text{if } U_i \geq U_j, j \neq i \\ 0 & \text{otherwise} \end{cases}$$

Attitude questions: $k = 1, \ldots, 4$

$$I_k = \alpha_k + \lambda_k \text{Attitude} + \mu_k$$

where I_k is the response to question k.

Model estimation

- Simultaneous estimation of all parameters
- with Python Biogeme
- Important: both the choice and the indicators reveal something about the attitude.

Measurement equations

- It feels safe to go by car.

$$I_1 = \text{Attitude} + 0.5666\ \nu_1$$

- It is comfortable to go by car to work.

$$I_2 = 1.13 + 0.764\ \text{Attitude} + 0.909\ \nu_2$$

- It is very important that traffic speed limits are not violated.

$$I_3 = 3.53 - 0.0716\ \text{Attitude} + 1.25\ \nu_3$$

- Increase the motorway speed limit to 140 km/h.

$$I_4 = 1.94 + 0.481\ \text{Attitude} + 1.37\ \nu_4$$

Structural model

Attitude towards car

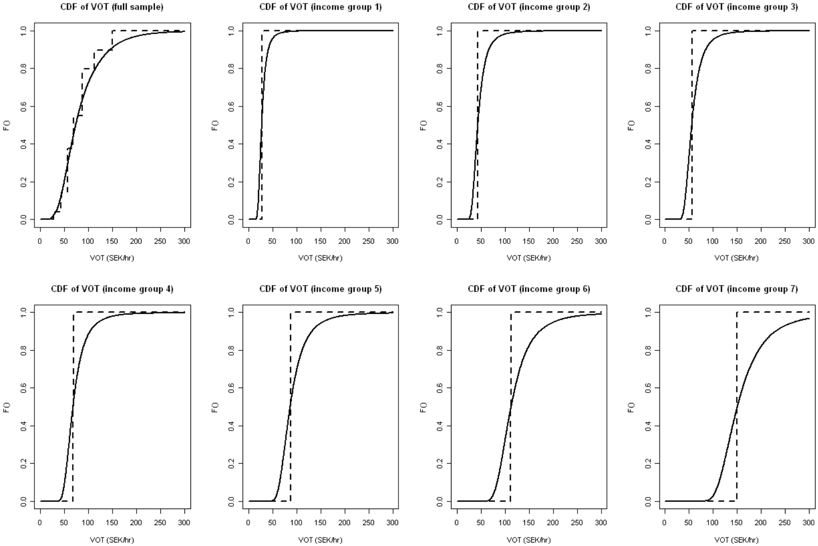
Param.	Estim.	t-stat.
θ_0	5.25	8.99
θ_f	-0.0185	-0.34
θ_{inc}	0.0347	1.99
θ_{age1}	-0.0217	-1.85
θ_{age2}	0.00797	0.88
θ_{age3}	0.0231	2.35
θ_{edu1}	-0.147	-0.94
θ_{edu2}	-0.252	-5.22
θ_{edu3}	-0.157	-0.85
σ	0.934	16.18

Structural model

Utility

Param.	Estim.	*t*-stat.
β_1	4.01	15.58
β_2	2.84	10.57
Time	-0.0388	-8.10
Cost/Income	-2.02	-3.63
Cost · Attitude/Income	0.265	2.11
Speed camera	-0.109	-2.75

Value of time



Outline

Conclusion

- Flexible models with more structure
- Translate more assumptions into equations
- More complicated to estimate
- Currently very active field for research and applications.