Theoretical foundations

Behavioral assumptions

Michel Bierlaire

Mathematical Modeling of Behavior

Choice theory

Theory of behavior that is

- descriptive: how people behave and not how they should,
- abstract: not too specific,
- operational: can be used in practice for forecasting.

Building the theory

Define

1. who (or what) is the decision maker,
2. what are the characteristics of the decision maker,
3. what are the alternatives available for the choice,
4. what are the attributes of the alternatives, and
5. what is the decision rule that the decision maker uses to make a choice.

Outline

Decision maker

Individual

- a person,
- a group of persons (internal interactions are ignored):
  - household, family,
  - firm,
  - government agency,
- notation: n.

Characteristics of the decision maker

Disaggregate models

Individuals

- face different choice situations,
- have different tastes.

Characteristics

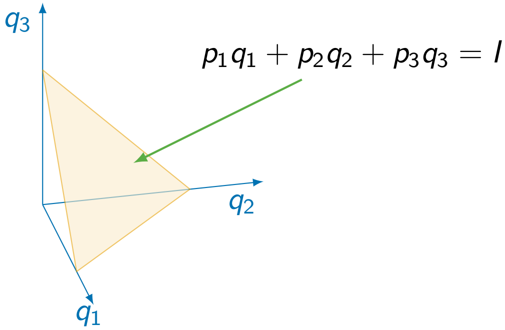
- income,
- sex,
- age,
- level of education,
- household/firm size,
- etc.

Outline

Alternatives: continuous choice set

Commodity bundle

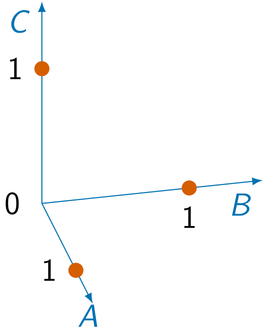
- q_1: quantity of milk.
- q_2: quantity of bread.
- q_3: quantity of butter.
- Unit price: p_i.
- Budget: I.



Alternatives: discrete choice set

List of alternatives

- Brand A.
- Brand B.
- Brand C.



Alternatives: discrete choice set

Choice set

- Non empty finite and countable set of alternatives.
- Universal: $\mathcal{C}$.
- Individual specific: $\mathcal{C}_n \subseteq \mathcal{C}$.
- Availability, awareness.

Example

Choice of a transportation mode:

- $\mathcal{C}$ ={car, bus, metro, walking }.
- If decision maker n has no driver license, and the trip is 12km long

$$\mathcal{C}_n = \{\text{bus}, \text{metro}\}.$$

Outline

Alternative attributes

Characterize each alternative i for each individual n

- price,
- travel time,
- frequency,
- comfort,
- color,
- size,
- etc.

Nature of the variables

- Quantitative and qualitative.
- Generic and specific.

Outline

Decision rule

Homo economicus

Rational and narrowly self-interested economic actor who is optimizing her outcome.

Preferences

- $i \succ j$: i is preferred to j,
- $i \sim j$: indifference between i and j,
- $i \succsim j$: i is at least as preferred as j.

Decision rule

Rationality

- Completeness: for all alternatives i and j,

$$i \succ j \text{ or } i \prec j \text{ or } i \sim j.$$

- Transitivity: for all bundles i, j and k,

$$\text{if } i \succsim j \text{ and } j \succsim k \text{ then } i \succsim k.$$

- "Continuity": if i is preferred to j and k is arbitrarily "close" to i, then k is preferred to j.

Utility

$$U_n : \mathcal{C}_n \longrightarrow \mathbb{R} : i \rightsquigarrow U_n(i).$$

Consistent with the preferences:

$$U_n(i) \geq U_n(j) \Longleftrightarrow i \succsim j.$$

- Captures the attractiveness of an alternative.
- Measure that the decision maker wants to optimize.
- Unique up to an order-preserving transformation.

Utility

Shift invariant

$$i \succsim j \iff U_n(i) \geq U_n(j) \iff U_n(i) + \eta \geq U_n(j) + \eta, \forall \eta \in \mathbb{R}.$$

Scale invariant

$$i \succsim j \iff U_n(i) \geq U_n(j) \iff \mu U_n(i) \geq \mu U_n(j), \forall \mu \in \mathbb{R}, \mu > 0.$$

Comments

- The “zero” is arbitrary.
- The units are arbitrary.

Behavioral assumptions

- The preference structure of the decision maker is fully characterized by a utility associated with each alternative.
- The decision maker is a perfect optimizer.
- The alternative with the highest utility is chosen.

The case of continuous goods

- Consumption bundle:

$$q = \begin{pmatrix} q_1 \\ \vdots \\ q_L \end{pmatrix}, \; p = \begin{pmatrix} p_1 \\ \vdots \\ p_L \end{pmatrix}.$$

- Budget constraint:

$$p^T q = \sum_{\ell=1}^{L} p_\ell q_\ell \leq I.$$

- No attributes, just quantities and prices.

Choice

Solution of an optimization problem

$$\max_{q \in \mathbb{R}^L} \widetilde{U}(q)$$

subject to

$$p^T q \leq I, \; q \geq 0.$$

Demand function

- Solution of the optimization problem.
- Quantity as a function of prices and budget:

$$q^* = \text{demand}(I, p).$$

Example: Cobb-Douglas

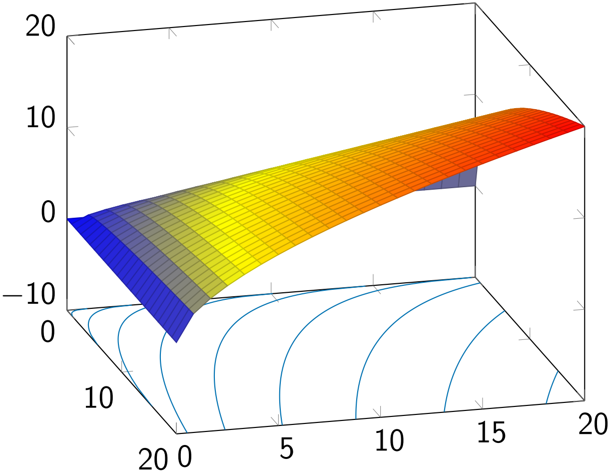
Utility function

$$\widetilde{U}(q) = \theta_0 \prod_{\ell=1}^{L} q_\ell^{\theta_\ell}.$$

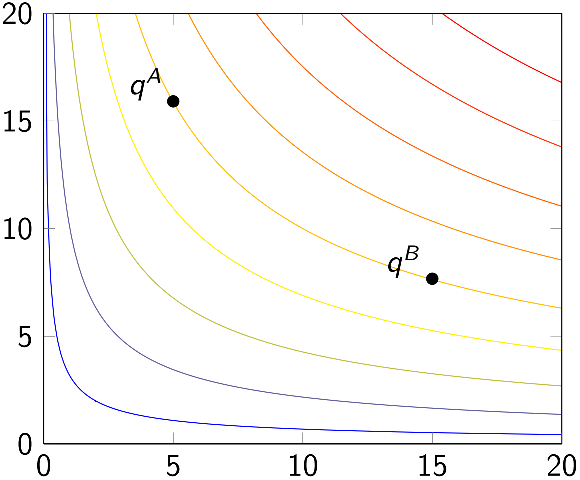
Demand function

$$q_i^* = \frac{\theta_i}{\sum_{\ell=1}^{L} \theta_\ell} \frac{I}{p_i}$$

Example: Cobb-Douglas



Example



The case of discrete goods

The consumer

- selects the quantities of continuous goods: $q = (q_1, \ldots, q_L)$,
- chooses an alternative in a discrete choice set $i = 1, \ldots, j, \ldots, J$.
- Discrete decision vector: $(y_1, \ldots, y_J)$, $y_j \in \{0, 1\}$, $\sum_j y_j = 1$.

Note

- In theory, one alternative of the discrete choice combines all possible choices made by an individual.
- In practice, the choice set is restricted for tractability.

Example

Choices

- House location: discrete choice.
- Car type: discrete choice.
- Number of kilometers driven per year: continuous choice.

Discrete choice set

Each combination of a house location and a car is an alternative.

Utility maximization

Utility

$$\widetilde{U}(q, y, \tilde{z}^T y; \theta).$$

- q: quantities of the continuous good.
- y: discrete choice.
- $\tilde{z}^T = (\tilde{z}_1, \ldots, \tilde{z}_i, \ldots, \tilde{z}_J) \in \mathbb{R}^{K \times J}$: K attributes of the J alternatives.
- $\tilde{z}^T y \in \mathbb{R}^K$: attributes of the chosen alternative.
- θ: vector of parameters. Let's ignore them for now.

Optimization problem

$$\max_{q,y} \widetilde{U}(q, y, \tilde{z}^T y; \theta)$$

subject to

$$p^T q + c^T y \leq I,$$
$$\sum_j y_j = 1,$$
$$y_j \in \{0, 1\}, \forall j.$$

where $c^T = (c_1, \ldots, c_i, \ldots, c_J)$ is the cost of each alternative

Solving the problem

- Mixed integer optimization problem.
- No optimality condition.
- Impossible to derive demand functions directly.

Solving the problem

Step 1: condition on the choice of the discrete good

- Fix the discrete good, that is select a feasible y.
- The problem becomes a continuous problem in q.
- Conditional demand functions can be derived:

$$q_{\ell|y} = \text{demand}(I - c^T y, p, \tilde{z}^T y),$$

or, equivalently, for each alternative i,

$$q_{\ell|i} = \text{demand}(I - c_i, p, \tilde{z}_i).$$

- $I - c_i$ is the income left for the continuous goods, if alternative i is chosen.
- If $I - c_i < 0$, alternative i is declared unavailable and removed from the choice set.

Solving the problem

Conditional demand functions

$$\text{demand}(I - c^T y, p, \tilde{z}^T y).$$

Conditional indirect utility functions

Substitute the demand functions into the utility:

$$U = \widetilde{U}(\text{demand}(I - c^T y, p, \tilde{z}_i), y, \tilde{z}^T y) = U(I - c^T y, y, p, \tilde{z}^T y).$$

Solving the problem

Step 2: Choice of the discrete good

$$\max_y U(I - c^T y, y, p, \tilde{z}^T y) \text{ s.t. } y \in \{0,1\}^J,\ \sum_{i=1}^{J} y_i = 1.$$

- Enumerate all alternatives.
- For each alternative i, set $y_i = 1$, $y_j = 0$, $j \neq i$.
- Compute the conditional indirect utility function U.
- Select the alternative with the highest U.
- Note: no income constraint anymore.

Model for individual n

$$\max_y U(I_n - c_n^T y, y, p_n, \tilde{z}_n^T y).$$

Simplifications

- s_n: set of characteristics of n, including income I_n.
- Prices of the continuous goods (p_n) are neglected.
- c_{in} is considered as another attribute and merged into $\tilde{z}_n$:

$$z_n = \{\tilde{z}_n, c_n\}.$$

Optimization problem

$$\max_i U_{in} = U(z_{in}, s_n)$$

Outline

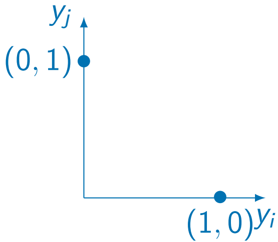
Example: transportation mode choice

Context

Choice between car and bus for a commuter trip.

	Attributes	
Alternatives	Travel time (t)	Travel cost (c)
i (e.g. car)	t_i	c_i
j (e.g. bus)	t_j	c_j

Decision variables



$$\begin{aligned} y_i, y_j &= \begin{cases} 1 & \text{if car is chosen,} \\ 0 & \text{otherwise;} \end{cases} \\ y_i + y_j &= 1 \end{aligned}$$

Example: transportation mode choice

Utility functions

Arbitrary units: for instance, CHF.

$$
\begin{aligned}
U_i &= -c_i - \beta t_i, \\
U_j &= -c_j - \beta t_j,
\end{aligned}
$$

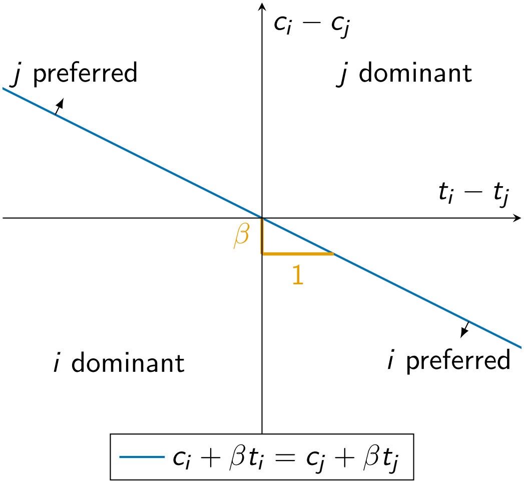
where $\beta > 0$ is a parameter to be estimated from data.

Role of β: transforming minutes into CHF.

i is chosen if

$$
\begin{gathered}
U_i \geq U_j, \\
-c_i - \beta t_i \geq -c_j - \beta t_j, \\
-\beta(t_i - t_j) \geq c_i - c_j.
\end{gathered}
$$

Example: transportation mode choice



Example: transportation mode choice

$c_j > c_i$ and $t_j > t_i$

i is dominant.

$c_i > c_j$ and $t_i > t_j$

j is dominant.

$c_j > c_i$ and $t_i > t_j$

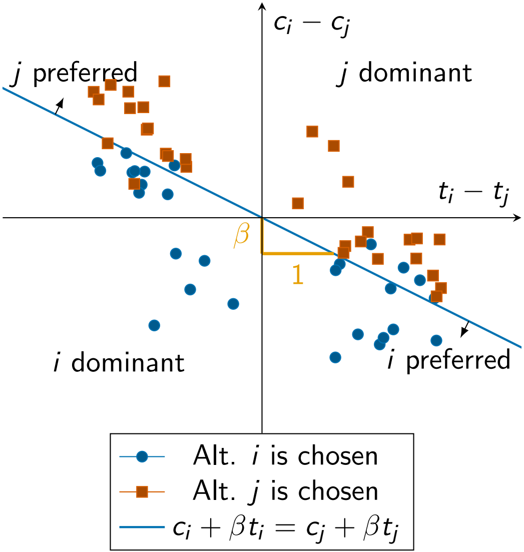
Alternative j is chosen if

$$-\beta(t_i - t_j) \leq c_i - c_j,$$

or, as $t_i > t_j$,

$$\beta \geq \frac{c_j - c_i}{t_i - t_j}.$$

Example: transportation mode choice



Outline

Behavioral validity of the utility maximization?

Assumptions

Decision-makers

- are able to process information,
- have perfect discrimination power,
- have transitive preferences,
- are perfect maximizer,
- are always consistent.

Relax the assumptions

Use a probabilistic approach: what is the probability that alternative i is chosen?

Introducing probability

Constant utility

- Human behavior is inherently random.
- Utility is deterministic.
- Consumer does not maximize utility.
- Probability to use inferior alternative is non zero.

Random utility

- Decision-maker are rational maximizers.
- Analysts have no access to the utility used by the decision-maker.
- Utility becomes a random variable.

Niels Bohr

Nature is stochastic.

Albert Einstein

God does not throw dice.

Random utility model

Probability model

$$P(i|\mathcal{C}_n) = \Pr(U_{in} \geq U_{jn}, \forall j \in \mathcal{C}_n).$$

Random utility

$$U_{in} = V_{in} + \varepsilon_{in}.$$

Random utility model

$$P(i|\mathcal{C}_n) = \Pr(V_{in} + \varepsilon_{in} \geq V_{jn} + \varepsilon_{jn}, \forall j \in \mathcal{C}_n),$$

or

$$P(i|\mathcal{C}_n) = \Pr(\varepsilon_{jn} - \varepsilon_{in} \leq V_{in} - V_{jn}, \forall j \in \mathcal{C}_n).$$

The random utility model

- Assume that $\varepsilon_n = (\varepsilon_{1n}, \ldots, \varepsilon_{J_n n})$ is a multivariate random variable,
- with CDF

$$F_{\varepsilon_n}(\varepsilon_1, \ldots, \varepsilon_{J_n}),$$

- and pdf

$$f_{\varepsilon_n}(\varepsilon_1, \ldots, \varepsilon_{J_n}) = \frac{\partial^{J_n} F}{\partial \varepsilon_1 \cdots \partial \varepsilon_{J_n}}(\varepsilon_1, \ldots, \varepsilon_{J_n}).$$

Then $P_n(i|\mathcal{C}_n) =$

$$\int_{\varepsilon=-\infty}^{+\infty} \frac{\partial F_{\varepsilon_{1n}, \varepsilon_{2n}, \ldots, \varepsilon_{J_n}}}{\partial \varepsilon_i}(\ldots, V_{in} - V_{(i-1)n} + \varepsilon, \varepsilon, V_{in} - V_{(i+1)n} + \varepsilon, \ldots) d\varepsilon.$$

Derivation in the appendix.

Random utility model

- The general formulation is complex.
- We will derive specific models based on simple assumptions.
- We will then relax some of these assumptions to propose more advanced models.

Summary

- Ingredients of choice theory.
- Utility theory: from continuous to discrete goods.
- Random utility theory.

Outline

Derivation of the random utility model

Joint distributions of ε_n

- Assume that $\varepsilon_n = (\varepsilon_{1n}, \ldots, \varepsilon_{J_n n})$ is a multivariate random variable,
- with CDF
$$F_{\varepsilon_n}(\varepsilon_1, \ldots, \varepsilon_{J_n}),$$
- and pdf
$$f_{\varepsilon_n}(\varepsilon_1, \ldots, \varepsilon_{J_n}) = \frac{\partial^{J_n} F}{\partial \varepsilon_1 \cdots \partial \varepsilon_{J_n}}(\varepsilon_1, \ldots, \varepsilon_{J_n}).$$

Derive the model for the first alternative (wlog)

$$P_n(1|\mathcal{C}_n) = \Pr(V_{2n} + \varepsilon_{2n} \leq V_{1n} + \varepsilon_{1n}, \ldots, V_{Jn} + \varepsilon_{Jn} \leq V_{1n} + \varepsilon_{1n}),$$

or

$$P_n(1|\mathcal{C}_n) = \Pr(\varepsilon_{2n} - \varepsilon_{1n} \leq V_{1n} - V_{2n}, \ldots, \varepsilon_{Jn} - \varepsilon_{1n} \leq V_{1n} - V_{Jn}).$$

Derivation

Model

$$P_n(1|\mathcal{C}_n) = \Pr(\varepsilon_{2n} - \varepsilon_{1n} \leq V_{1n} - V_{2n}, \ldots, \varepsilon_{Jn} - \varepsilon_{1n} \leq V_{1n} - V_{Jn}).$$

Change of variables

$$\xi_{1n} = \varepsilon_{1n},\ \xi_{in} = \varepsilon_{in} - \varepsilon_{1n},\ i = 2, \ldots, J_n,$$

that is

$$\begin{pmatrix} \xi_{1n} \\ \xi_{2n} \\ \vdots \\ \xi_{(J_n-1)n} \\ \xi_{J_n n} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & \cdots & 0 & 0 \\ & & \vdots & & \\ -1 & 0 & \cdots & 1 & 0 \\ -1 & 0 & \cdots & 0 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{1n} \\ \varepsilon_{2n} \\ \vdots \\ \varepsilon_{(J_n-1)n} \\ \varepsilon_{J_n n} \end{pmatrix}.$$

Derivation

Model in ε

$$P_n(1|\mathcal{C}_n) = \Pr(\varepsilon_{2n} - \varepsilon_{1n} \leq V_{1n} - V_{2n}, \ldots, \varepsilon_{Jn} - \varepsilon_{1n} \leq V_{1n} - V_{Jn}).$$

Change of variables

$$\xi_{1n} = \varepsilon_{1n}, \; \xi_{in} = \varepsilon_{in} - \varepsilon_{1n}, \; i = 2, \ldots, J_n,$$

Model in ξ

$$P_n(1|\mathcal{C}_n) = \Pr(\xi_{2n} \leq V_{1n} - V_{2n}, \ldots, \xi_{J_n n} \leq V_{1n} - V_{J_n n}).$$

Note

The determinant of the change of variable matrix is 1, so that ε and ξ have the same pdf

Derivation

$$
\begin{aligned}
& P_n(1|\mathcal{C}_n) \\
= \quad & \Pr(\xi_{2n} \leq V_{1n} - V_{2n}, \ldots, \xi_{J_n n} \leq V_{1n} - V_{J_n n}) \\
= \quad & F_{\xi_{1n},\xi_{2n},\ldots,\xi_{J_n}}(+\infty, V_{1n} - V_{2n}, \ldots, V_{1n} - V_{J_n n}) \\
= \quad & \int_{\xi_1=-\infty}^{+\infty} \int_{\xi_2=-\infty}^{V_{1n}-V_{2n}} \cdots \int_{\xi_{J_n}=-\infty}^{V_{1n}-V_{J_n n}} f_{\xi_{1n},\xi_{2n},\ldots,\xi_{J_n}}(\xi_1, \xi_2, \ldots, \xi_{J_n}) d\xi, \\
= \quad & \int_{\varepsilon_1=-\infty}^{+\infty} \int_{\varepsilon_2=-\infty}^{V_{1n}-V_{2n}+\varepsilon_1} \cdots \int_{\varepsilon_{J_n}=-\infty}^{V_{1n}-V_{J_n n}+\varepsilon_1} f_{\varepsilon_{1n},\varepsilon_{2n},\ldots,\varepsilon_{J_n}}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{J_n}) d\varepsilon,
\end{aligned}
$$

Derivation

$$P_n(1|\mathcal{C}_n) = \int_{\varepsilon_1=-\infty}^{+\infty} \int_{\varepsilon_2=-\infty}^{V_{1n}-V_{2n}+\varepsilon_1} \cdots \int_{\varepsilon_{J_n}=-\infty}^{V_{1n}-V_{J_n n}+\varepsilon_1} f_{\varepsilon_{1n},\varepsilon_{2n},\ldots,\varepsilon_{J_n}}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{J_n}) d\varepsilon$$

$$P_n(1|\mathcal{C}_n) = \int_{\varepsilon_1=-\infty}^{+\infty} \frac{\partial F_{\varepsilon_{1n},\varepsilon_{2n},\ldots,\varepsilon_{J_n}}}{\partial \varepsilon_1}(\varepsilon_1, V_{1n} - V_{2n} + \varepsilon_1, \ldots, V_{1n} - V_{J_n n} + \varepsilon_1) d\varepsilon_1.$$

Appendix: some concepts from continuous choices

Roy's identity

Derive the demand function from the indirect utility:

$$q_\ell = -\frac{\partial U(I, p; \theta)/\partial p_\ell}{\partial U(I, p; \theta)/\partial I}$$

Elasticities

Direct price elasticity

Percent change in demand resulting form a 1% change in price

$$E_{p_\ell}^{q_\ell} = \frac{\% \text{ change in } q_\ell}{\% \text{ change in } p_\ell} = \frac{\Delta q_\ell / q_\ell}{\Delta p_\ell / p_\ell} = \frac{p_\ell}{q_\ell} \frac{\Delta q_\ell}{\Delta p_\ell}.$$

Asymptotically

$$E_{p_\ell}^{q_\ell} = \frac{p_\ell}{q_\ell(I, p; \theta)} \frac{\partial q_\ell(I, p; \theta)}{\partial p_\ell}.$$

Cross price elasticity

$$E_{p_m}^{q_\ell} = \frac{p_m}{q_\ell(I, p; \theta)} \frac{\partial q_\ell(I, p; \theta)}{\partial p_m}.$$

Consumer surplus

Definition

Difference between what a consumer is willing to pay for a good and what she actually pays for that good.

Calculation

Area under the demand curve and above the market price

Demand curve

- Plot of the inverse demand function
- Price as a function of quantity

Consumer surplus