Choice theory

Michel Bierlaire

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

Outline

Theoretical foundations

Choice: outcome of a sequential decision-making process

- defining the choice problem
- generating alternatives
- evaluating alternatives
- making a choice,
- executing the choice.

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.

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.

Alternatives

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 the decision maker has no driver license, and the trip is 12km long

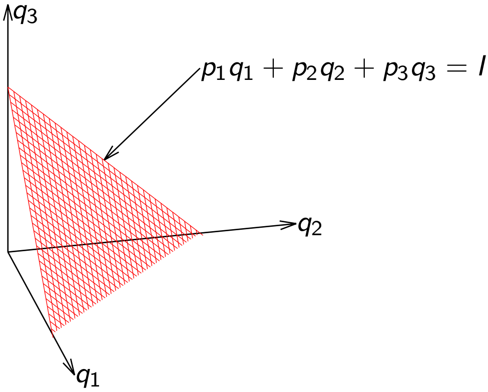
$$\mathcal{C}_n = \{\text{bus}, \text{metro}\}$$

Continuous choice set

Microeconomic demand analysis

Commodity bundle

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

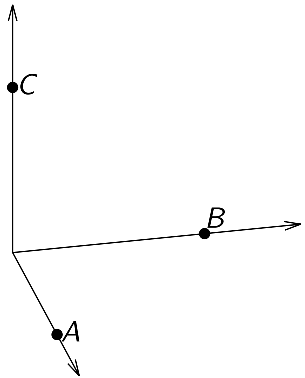


Discrete choice set

Discrete choice analysis

List of alternatives

- Brand A
- Brand B
- Brand C



Alternative attributes

Characterize each alternative i for each individual n

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

Nature of the variables

- Discrete and continuous
- Generic and specific
- Measured or perceived

Decision rule

Homo economicus

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

Utility

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

- captures the attractiveness of an alternative
- measure that the decision maker wants to optimize

Behavioral assumption

- the decision maker associates a utility with each alternative
- the decision maker is a perfect optimizer
- the alternative with the highest utility is chosen

Simple example: mode choice

Attributes

	Attributes	
Alternatives	Travel time (t)	Travel cost (c)
Car (1)	t_1	c_1
Bus (2)	t_2	c_2

Simple example: mode choice

Utility functions

$$
\begin{aligned}
U_1 &= -\beta_t t_1 - \beta_c c_1, \\
U_2 &= -\beta_t t_2 - \beta_c c_2,
\end{aligned}
$$

where $\beta_t > 0$ and $\beta_c > 0$ are parameters.

Equivalent specification

$$
\begin{aligned}
U_1 &= -(\beta_t/\beta_c)t_1 - c_1 &= -\beta t_1 - c_1 \\
U_2 &= -(\beta_t/\beta_c)t_2 - c_2 &= -\beta t_2 - c_2
\end{aligned}
$$

where $\beta > 0$ is a parameter.

Choice

- Alternative 1 is chosen if $U_1 \geq U_2$.
- Ties are ignored.

Simple example: mode choice

Choice

Alternative 1 is chosen if

$$-\beta t_1 - c_1 \geq -\beta t_2 - c_2$$

or

$$-\beta(t_1 - t_2) \geq c_1 - c_2$$

Alternative 2 is chosen if

$$-\beta t_1 - c_1 \leq -\beta t_2 - c_2$$

or

$$-\beta(t_1 - t_2) \leq c_1 - c_2$$

Dominated alternative

- If $c_2 > c_1$ and $t_2 > t_1$, $U_1 > U_2$ for any $\beta > 0$
- If $c_1 > c_2$ and $t_1 > t_2$, $U_2 > U_1$ for any $\beta > 0$

Simple example: mode choice

Trade-off

- Assume $c_2 > c_1$ and $t_1 > t_2$.
- Is the traveler willing to pay the extra cost $c_2 - c_1$ to save the extra time $t_1 - t_2$?
- Alternative 2 is chosen if

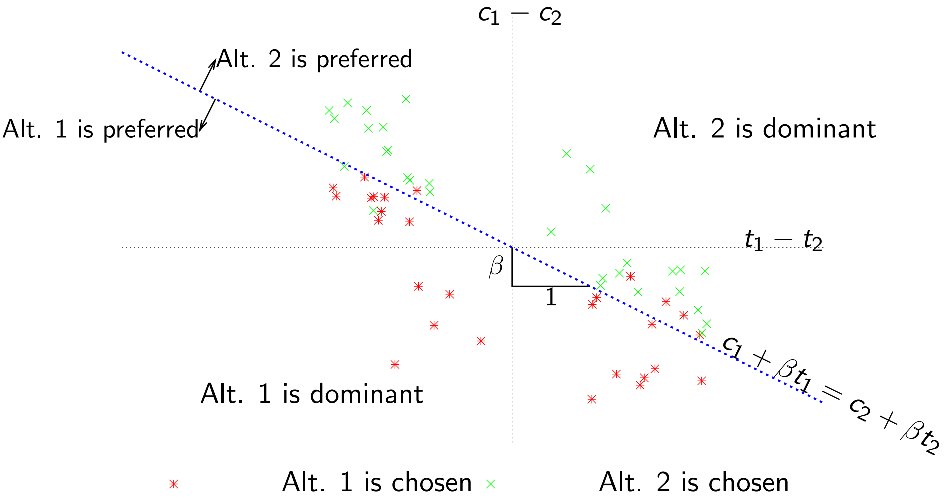
$$-\beta(t_1 - t_2) \leq c_1 - c_2$$

or

$$\beta \geq \frac{c_2 - c_1}{t_1 - t_2}$$

- β is called the *willingness to pay* or *value of time*

Simple example: mode choice



Random utility model

Random utility

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

The logit model

$$P(i|\mathcal{C}_n) = \frac{e^{V_{in}}}{\sum_{j \in \mathcal{C}_n} e^{V_{jn}}}$$

Outline

Microeconomic consumer theory

Continuous choice set

- 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

Preferences

Operators $\succ$, $\sim$, and $\succsim$

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

Rationality

- Completeness: for all bundles a and b,

$$Q_a \succ Q_b \text{ or } Q_a \prec Q_b \text{ or } Q_a \sim Q_b.$$

- Transitivity: for all bundles a, b and c,

$$\text{if } Q_a \succsim Q_b \text{ and } Q_b \succsim Q_c \text{ then } Q_a \succsim Q_c.$$

- "Continuity": if Q_a is preferred to Q_b and Q_c is arbitrarily "close" to Q_a, then Q_c is preferred to Q_b.

Utility

Utility function

- Parametrized function:

$$\widetilde{U} = \widetilde{U}(q_1, \ldots, q_L; \theta) = \widetilde{U}(Q; \theta)$$

- Consistent with the preference indicator:

$$\widetilde{U}(Q_a; \theta) \geq \widetilde{U}(Q_b; \theta)$$

  is equivalent to

$$Q_a \succsim Q_b.$$

- Unique up to an order-preserving transformation

Optimization

Optimization problem

$$\max_{Q} \widetilde{U}(Q;\theta)$$

subject to

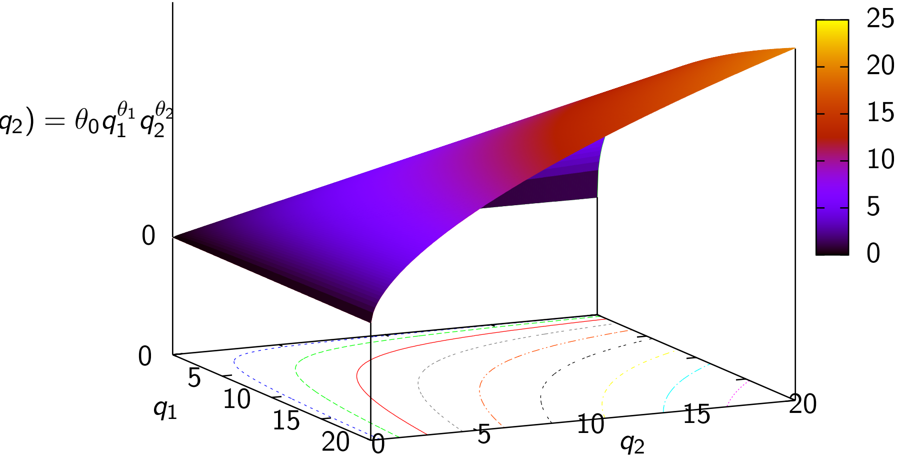
$$p^T Q \leq I, \; Q \geq 0.$$

Demand function

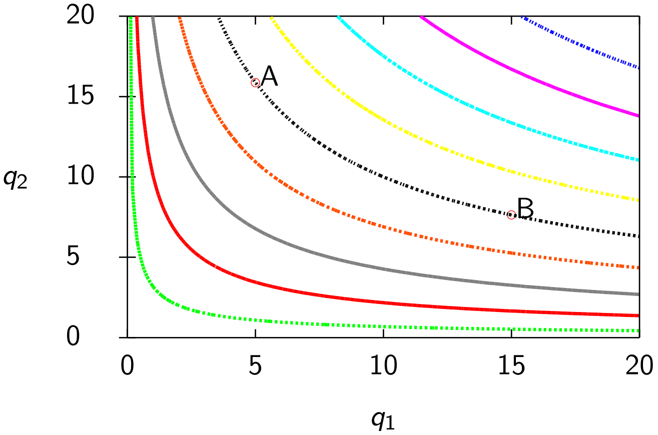
- Solution of the optimization problem
- Quantity as a function of prices and budget

$$Q^* = f(I, p; \theta)$$

Example: Cobb-Douglas



Example



Example

Optimization problem

$$\max_{q_1,q_2} \widetilde{U}(q_1, q_2; \theta_0, \theta_1, \theta_2) = \theta_0 q_1^{\theta_1} q_2^{\theta_2}$$

subject to

$$p_1 q_1 + p_2 q_2 = I.$$

Lagrangian of the problem:

$$L(q_1, q_2, \lambda) = \theta_0 q_1^{\theta_1} q_2^{\theta_2} + \lambda(I - p_1 q_1 - p_2 q_2).$$

Necessary optimality condition

$$\nabla L(q_1, q_2, \lambda) = 0$$

Example

Necessary optimality conditions

$$
\begin{array}{rcrcll}
\theta_0\theta_1 q_1^{\theta_1-1} q_2^{\theta_2} & - & \lambda p_1 & = & 0 & (\times q_1) \\
\theta_0\theta_2 q_1^{\theta_1} q_2^{\theta_2-1} & - & \lambda p_2 & = & 0 & (\times q_2) \\
p_1 q_1 + p_2 q_2 & - & I & = & 0. &
\end{array}
$$

We have

$$
\begin{array}{rcrcl}
\theta_0\theta_1 q_1^{\theta_1} q_2^{\theta_2} & - & \lambda p_1 q_1 & = & 0 \\
\theta_0\theta_2 q_1^{\theta_1} q_2^{\theta_2} & - & \lambda p_2 q_2 & = & 0.
\end{array}
$$

Adding the two and using the third condition, we obtain

$$\lambda I = \theta_0 q_1^{\theta_1} q_2^{\theta_2}(\theta_1 + \theta_2)$$

or, equivalently,

$$\theta_0 q_1^{\theta_1} q_2^{\theta_2} = \frac{\lambda I}{(\theta_1 + \theta_2)}$$

Solution

From the previous derivation

$$\theta_0 q_1^{\theta_1} q_2^{\theta_2} = \frac{\lambda I}{(\theta_1 + \theta_2)}$$

First condition

$$\theta_0 \theta_1 q_1^{\theta_1} q_2^{\theta_2} = \lambda p_1 q_1.$$

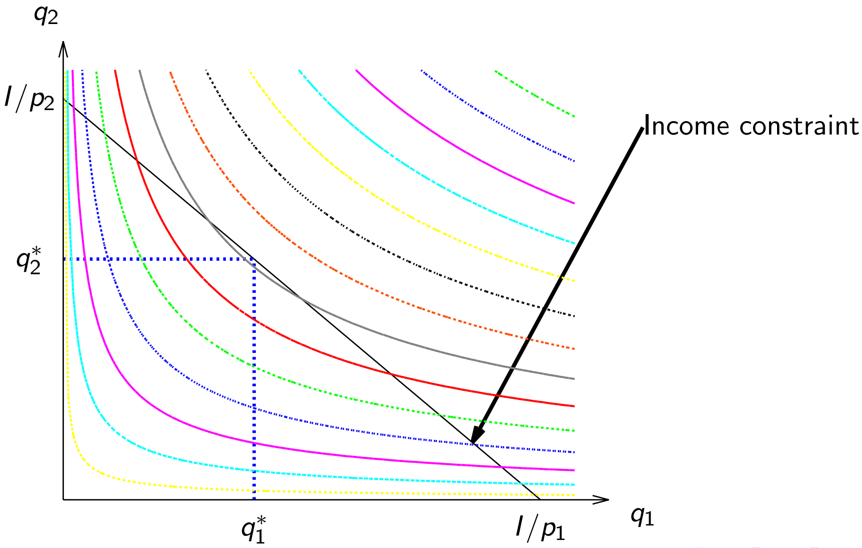
Solve for q_1

$$q_1^* = \frac{I\theta_1}{p_1(\theta_1 + \theta_2)}$$

Similarly, we obtain

$$q_2^* = \frac{I\theta_2}{p_2(\theta_1 + \theta_2)}$$

Optimization problem



Demand functions

Product 1

$$q_1^* = \frac{I}{p_1} \frac{\theta_1}{\theta_1 + \theta_2}$$

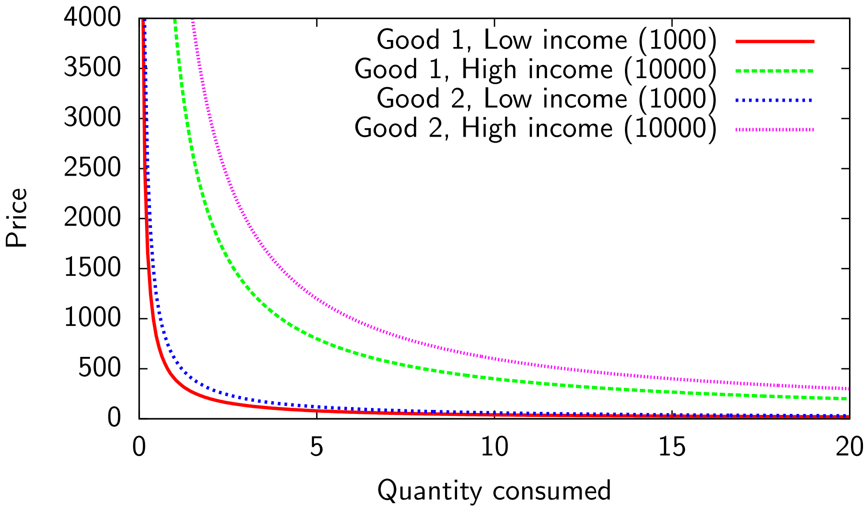
Product 2

$$q_2^* = \frac{I}{p_2} \frac{\theta_2}{\theta_1 + \theta_2}$$

Comments

- Demand decreases with price
- Demand increases with budget
- Demand independent of θ_0, which does not affect the ranking
- Property of Cobb Douglas: the demand for a good is only dependent on its own price and independent of the price of any other good.

Demand curve (inverse of demand function)



Indirect utility

Substitute the demand function into the utility

$$U(I, p; \theta) = \theta_0 \left(\frac{I}{p_1} \frac{\theta_1}{\theta_1 + \theta_2} \right)^{\theta_1} \left(\frac{I}{p_2} \frac{\theta_2}{\theta_1 + \theta_2} \right)^{\theta_2}$$

Indirect utility

Maximum utility that is achievable for a given set of prices and income

In discrete choice...

- only the indirect utility is used
- therefore, it is simply referred to as “utility”
- we review some results from microeconomics useful for discrete choice

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.

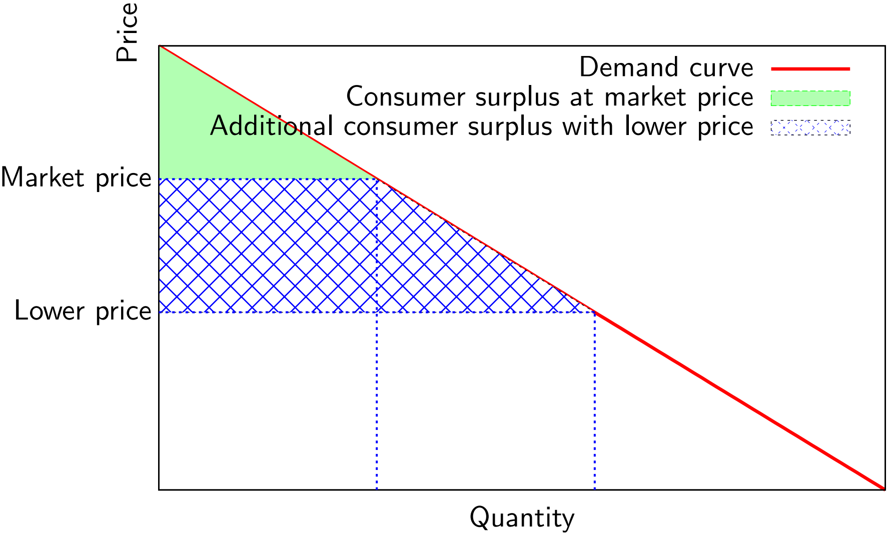
Calculation

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



Outline

Microeconomic theory of discrete goods

Expanding the microeconomic framework

- Continuous goods
- and 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 will be more restricted for tractability

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

Utility maximization

Optimization problem

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

subject to

$$\begin{aligned} & p^T Q + c^T y \leq I \\ & \textstyle\sum_j y_j = 1 \\ & y_j \in \{0, 1\}, \forall j. \end{aligned}$$

where $c^T = (c_1, \ldots, c_i, \ldots, c_J)$ contains 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:

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

or, equivalently, for each alternative i,

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

- $I - c_i$ 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 indirect utility functions

Substitute the demand functions into the utility:

$$U_i = U(I - c_i, p, \tilde{z}_i; \theta) \text{ for all } i \in \mathcal{C}.$$

Step 2: Choice of the discrete good

$$\max_y U(I - c^T y, p, \tilde{z}^T y; \theta)$$

- Enumerate all alternatives.
- Compute the conditional indirect utility function U_i.
- Select the alternative with the highest U_i.
- Note: no income constraint anymore.

Model for individual n

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

Simplifications

- We cannot estimate a set of parameters for each individual n
- Therefore, population level parameters are interacted with characteristics S_n of the decision-maker
- Prices of the continuous goods are neglected p_n
- Income is considered as another characteristic and merged into S_n
- c_i is considered as another attribute and merged into $\tilde{z}$

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

$$\max_i U_{in} = U(z_{in}, S_n; \theta)$$

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

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

Random utility

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

Random utility model

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

or

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

Derivation

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

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

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

The random utility model: $P_n(i|\mathcal{C}_n) =$

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

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.