Transport Research and
Infrastructure Planning Laboratory
Department of Civil Engineering
The University of Tokyo
http://www.trip.t.u-tokyo.ac.jp/

Part 1:

Discrete Choice Modeling and Travel Demand Forecast

東京大学大学院工学系研究科社会基盤学専攻
交通・都市・国土学研究室

Core references

1. Michael D. Meyer and Eric J. Miller: Urban Transportation Planning 2nd Edition, McGraw-Hill, 2001.
2. Moshe Ben-Akiva, Discrete Choice Analysis: Theory and Application to Travel Demand, Mit Press, Series in Transportation Studies 9, 1985.
3. Norbert Oppenheim, Urban Travel Demand Modeling, Wiley-Interscience, 1995.
4. William H. Greene, Econometric Analysis 5th Edition, Prentice Hall, 2003.
5. Kenneth E. Train, Discrete Choice Methods with Simulation, Cambridge University Press, 2003.
6. Hal R. Varian, Microeconomic Analysis 3rd Edition, W. W. Norton & Company, 1992.
7. **森地茂・山形耕一編著，新体系土木工学60交通計画，技報堂出版，**1993
8. **北村隆一・森川高行編，交通行動の分析とモデリング，技報堂出版，**2002.
9. **土木学会編，非集計行動モデルの理論と実際，丸善，**1995.
10. **交通工学研究会編，やさしい非集計分析，**1993.

3

Why do we need the study on choice behavior in transport planning?

- Physical design of transport infrastructures
  - Size, Coverage area
- Business plan by transport companies
  - Service, Investment
- Calculation of the benefit for CBA

- Ridership / demand estimation
- Choice behavior

4

Choices in transport studies

- Trip generation / attraction (交通発生・集中)
- Destination (目的地)
- Transportation mode (交通機関)
- Route (経路)
- Time-of-day (出発時刻)
- Vehicle ownership (自動車保有)

- Discrete choice (離散選択)

5

Demand estimation process in urban transport planning

- Step-by-step type model

Trip Generation / Attraction ← Time-of-day choice model

↓

Trip Distribution ← Destination choice model

↓

Modal Sprit ← Mode choice model, Vehicle ownership model

↓

Traffic Assignment ← Route choice model

6

Theoretical background of choice behavior analysis

- Micro-economic theory
- Consumer behavior theory (消費者行動理論)
- Psychology
- Econometric analysis
- Statistics

Traditional assumption for choice in consumer behavior theory

- A consumer has the objectives to be achieved.
- A consumer can identify all of the possible alternatives achieving her/his objectives.
- A consumer can rank them precisely in terms of the preference.

- A consumer is rational(**合理的**).

Properties of rational choice

X,Y,Z are choice alternatives.

means "more preferable(**選好**)" or "indifferent(**無差別**)"

1. Reflexive(**再帰性**)
   for all X, $X \geq X$
2. Complete(**完全性**)
   for all X and Y, $X \geq Y$ or $Y \geq X$
3. Transitive(**推移性**)
   for all X and Y and Z, if $X \geq Y$ and $Y \geq Z$, then $X \geq Z$
4. Continuity(**連続性**)
   for all Y, the sets $\{X : X \geq Y\}$ and $\{X : Y \geq X\}$ are closed sets.
   We can define "indifference curve(**無差別曲線**)".

9

Definition of utility(効用)

- Reflects the level of satisfaction if a alternative is chosen and the objective is achieved.
- Function that gives the scalar value for the level of satisfaction---Utility function(効用関数)
  - If $X \geq Y$, $U(X) \geq U(Y)$

10

Utility maximization with constraints in consumer behavior(効用最大化行動)

- Consider the vectors of goods $\mathbf{X}$ and price of goods $\mathbf{P}$
- Consider the budget constraint(予算制約) M
- A consumer will choose the combination of goods in condition that $\max U(\mathbf{X})$ such that $\mathbf{PX} \leq M$
- Solution $\mathbf{X}^* = \mathbf{x}(\mathbf{P}, M)$ is demand function (需要関数)
- Maximum utility $U(\mathbf{X}^*) = V(\mathbf{P}, M)$ is indirect utility function (間接効用関数)
- $\mathbf{x}(\mathbf{P}, M)$ is derived from $V(\mathbf{P}, M)$ by Roy's identity (ロワの恒等式)

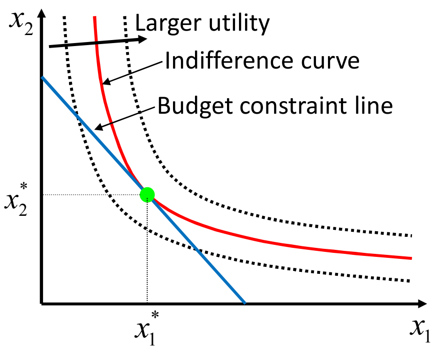
$$x_i(\mathbf{P}, M) = -\frac{\partial V(\mathbf{P}, M)/\partial p_i}{\partial V(\mathbf{P}, M)/\partial M}$$

11

Indifference curve and diminishing marginal utility(限界効用逓減)

Utility maximization with budget constraint

$\max U(\mathbf{X})$ s.t. $\mathbf{PX} \leq M$



Marginal rate of substitution (限界代替率): MRS

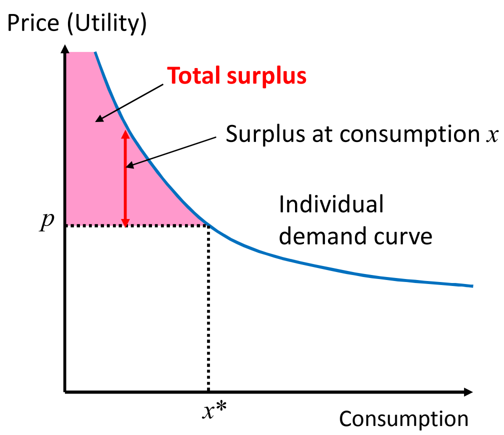
$$\frac{\partial U}{\partial x_i}$$

MRS decreases when x_i increases

→Marginal utility is diminishing

12

Definition of consumer surplus(消費者余剰)



Price p

➡ Optimal consumption $x*$

Utility maximization =Surplus maximization

⬇

Benefit

Expected utility theory(期待効用理論) – uncertainty(不確実性) in choice behavior

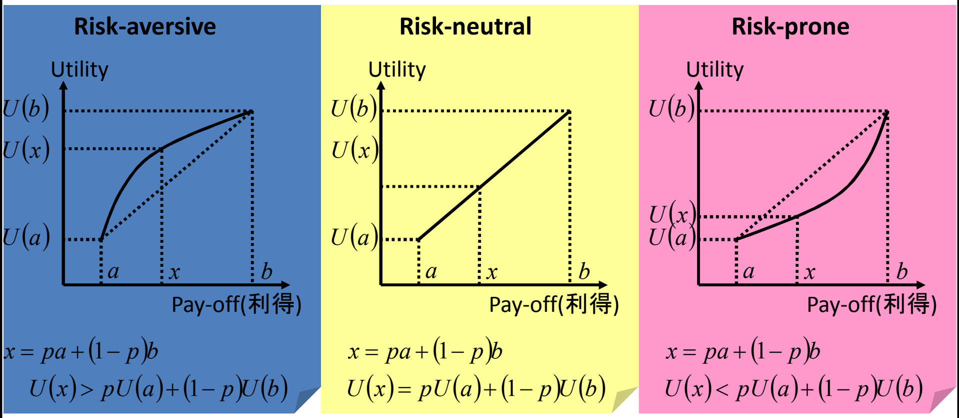
- The utility may change due to the situation that can not be expected in advance. = The choice is often subject to uncertainty

In case of access mode choice to the station…

	Fine	Rain
Use bicycle	U_{bf}	U_{br}
On foot	U_{ff}	U_{fr}
Probability	p	1-p

$$U_b = pU_{bf} + (1-p)U_{br} \Leftrightarrow U_f = pU_{ff} + (1-p)U_{fr}$$

Attitude toward uncertainty



Different type of uncertainty

- Utility itself is uncertain…
  - Travel time of congested NW in route choice or time-of-day choice
  - Future socio-economical change in vehicle ownership behavior

Random utility theory(ランダム効用理論)

- It allows that the utility is unfixed and varies randomly.
  - For modeler
    - Unobserved factors
  - For decision maker
    - Insensible factors
    - Fickle choice
- It defuses the disadvantage of the assumption of rational choice.

Formulation of the choice in random utility theory

1. The utility by individual n for alternative(選択肢) i is consist of systematic term(確定項) V and random term(確率項) ε

   $$U_{ni} = V_{ni} + \varepsilon_{ni}$$

2. The probability that n chooses i is expressed as follows:

   $$P_{ni} = \Pr\left[U_{ni} \geq \max_{j \neq i} U_{nj}\right] \qquad {}^{\forall}i, {}^{\forall}j \in I$$

Derivation of binary logit model(二項ロジットモデル) (1)

In case two alternatives i and j exist

$$P_{ni} = \Pr\left[\varepsilon_{nj} - \varepsilon_{ni} \leq V_{ni} - V_{nj}\right]$$
$$\equiv \Pr\left[\varepsilon_n \leq V_{ni} - V_{nj}\right] \equiv F_\varepsilon\left(V_{ni} - V_{nj}\right)$$

Cumulative distribution function (C.D.F.) $N\left(0, \sigma^2\right)$

C.D.F. of standardized normal distribution

If C.D.F. is based on the normal distribution

$$P_{ni} = \int_{-\infty}^{V_{ni} - V_{nj}} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\varepsilon_n^2}{2\sigma^2}\right) = \Phi\left(\frac{V_{ni} - V_{nj}}{\sigma}\right)$$

➡ Binary probit model
However there is no analytical equation…

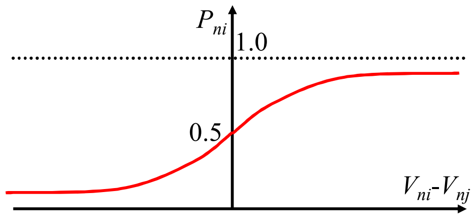
Derivation of binary logit model (2)

If C.D.F. is based on the logistic distribution

$$F(\varepsilon_n) = \frac{1}{1+\exp(-\mu\varepsilon_n)}$$

$$P_{ni} = F(V_{ni} - V_{nj}) = \frac{1}{1+\exp\{-\mu(V_{ni} - V_{nj})\}} = \frac{\exp(\mu V_{ni})}{\exp(\mu V_{ni}) + \exp(\mu V_{nj})}$$

Binary logit model



Derivation of multinomial logit model(多項ロジットモデル) (1)

In case I alternatives $1,\ldots,I$ exist

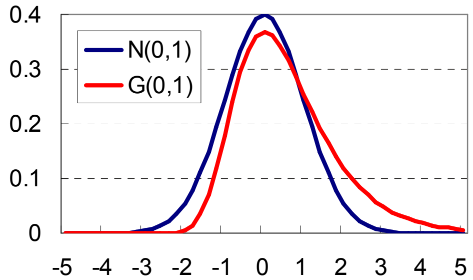
$$P_{ni} = \Pr\left[U_{ni} \geq \max_{j \neq i} U_{nj}\right] = \Pr\left[V_{ni} + \varepsilon_{ni} \geq \max_{j \neq i}(V_{nj} + \varepsilon_{nj})\right]$$

We assume the probability density function (P.D.F) of
is not normal distribution but gumbel distribution…

21

Property of gumbel distribution (1)

$$\varepsilon \sim G(\eta, \mu) \begin{cases} \text{C.D.F.} \\ F(\varepsilon) = \exp[-\exp\{-\mu(\varepsilon - \eta)\}] \\ \text{P.D.F.} \\ f(\varepsilon) = \mu \exp\{-\mu(\varepsilon - \eta)\} \exp[-\exp\{-\mu(\varepsilon - \eta)\}] \end{cases}$$



Mode η

Average $\eta + \frac{\gamma}{\mu}, \gamma = 0.577$ (Euler's const.)

Variance $\frac{\pi^2}{6\mu^2}$

22

Property of gumbel distribution (2)

Suppose $\varepsilon_1 \sim G(\eta_1, \mu)$ and $\varepsilon_2 \sim G(\eta_2, \mu)$ are identically and independently distributed (IID), $\varepsilon = \varepsilon_1 - \varepsilon_2$ is…

$$\varepsilon \sim F(\varepsilon) = \frac{1}{1 + \exp\{-\mu(\eta_2 - \eta_1 - \varepsilon)\}}$$ → Logistic distribution

Proof

$$F(\varepsilon) = \Pr[\varepsilon_1 - \varepsilon_2 \le \varepsilon] = \Pr[\varepsilon_1 \le \varepsilon + \varepsilon_2] = \int_{\varepsilon_2=-\infty}^{+\infty} \int_{\varepsilon_1=-\infty}^{\varepsilon+\varepsilon_2} f_1(\varepsilon_1) f_2(\varepsilon_2) d\varepsilon_1 d\varepsilon_2$$

$$= \int_{-\infty}^{+\infty} F_1(\varepsilon + \varepsilon_2) f_2(\varepsilon_2) d\varepsilon_2 = \int_{-\infty}^{+\infty} \mu \exp\{-\mu(\varepsilon_2 - \eta_2)\} \exp[-\exp(-\mu\varepsilon_2)[\exp\{-\mu(\varepsilon - \eta_1)\} + \exp(\mu\eta_2)]] d\varepsilon_2$$

Suppose $\delta = \exp\{-\mu(\varepsilon - \eta_1)\} + \exp \mu\eta_2$

$$F(\varepsilon) = \int_{-\infty}^{+\infty} \mu \exp\{-\mu(\varepsilon_2 - \eta_2)\} \exp\{-\delta \exp(-\mu\varepsilon_2)\} d\varepsilon_2 = \frac{1}{\delta} \exp(\mu\eta_2) [\exp\{-\delta \exp(-\mu\varepsilon_2)\}]_{-\infty}^{+\infty}$$ (→ 1)

$$= \frac{1}{\delta} \exp(\mu\eta_2) = \frac{\exp(\mu\eta_2)}{\exp\{-\mu(\varepsilon - \eta_1)\} + \exp(\mu\eta_2)} = \frac{1}{1 + \exp\{\mu(\eta_2 - \eta_1 - \varepsilon)\}}$$

why $G\left(\frac{\ln \delta}{\mu}, \mu\right)$

Property of gumbel distribution (3)

Suppose $\varepsilon_k \sim G(\eta_k, \mu)$ k=1,…,I and IID…

$$\max_k \varepsilon_k \sim G\left(\frac{1}{\mu}\ln\sum_{k=1}^{I}\exp(\mu\eta_k), \mu\right)$$

Proof

$$\Pr\left[\max_k \varepsilon_k \le \varepsilon\right] = \prod_{k=1}^{I}\Pr[\varepsilon_k \le \varepsilon] = \prod_{k=1}^{L}\exp[-\exp\{-\mu(\varepsilon-\eta_k)\}]$$

$$= \exp\left[\sum_{k=1}^{I} -\exp\{-\mu(\varepsilon-\eta_k)\}\right] = \exp\left[-\exp(\mu\varepsilon)\sum_{k=1}^{I}\exp(\mu\eta_k)\right]$$

Suppose $\alpha = \frac{1}{\mu}\ln\sum_{k=1}^{I}\exp(\mu\eta_k)$

$$\Pr\left[\max_k \varepsilon_k \le \varepsilon\right] = \exp[-\exp(-\mu\varepsilon)\exp(\alpha\mu)] = \exp[-\exp\{-\mu(\varepsilon-\alpha)\}]$$

$$\therefore \max_k \varepsilon_k \sim G\left(\frac{1}{\mu}\ln\sum_{k=1}^{I}\exp(\mu\eta_k), \mu\right)$$

Derivation of multinomial logit model (2)

$$P_{ni} = \Pr\left[U_{ni} \ge \max_{j\ne i} U_{nj}\right] = \Pr\left[V_{ni} + \varepsilon_{ni} \ge \max_{j\ne i}(V_{nj} + \varepsilon_{nj})\right]$$

Suppose $U_n^* = \max_{j\ne i}(V_{nj} + \varepsilon_{nj})$

$$U_n^* \sim G\left(\frac{1}{\mu}\ln\sum_{k\ne i}\exp(\mu V_{nk}), \mu\right)$$

Suppose $U_n^* = V_n^* + \varepsilon_n^*$

$$\begin{cases} V_n^* = \frac{1}{\mu}\ln\sum_{k\ne i}\exp(\mu V_{nk}) \\ \varepsilon_n^* \sim G(0, \mu) \end{cases}$$

25

Derivation of multinomial logit model (3)

$$P_{ni} = \Pr\left[\varepsilon_n^* - \varepsilon_{ni} \le V_{ni} - V_n^*\right] = \frac{1}{1 + \exp\left\{\mu\left(V_n^* - V_{ni}\right)\right\}}$$

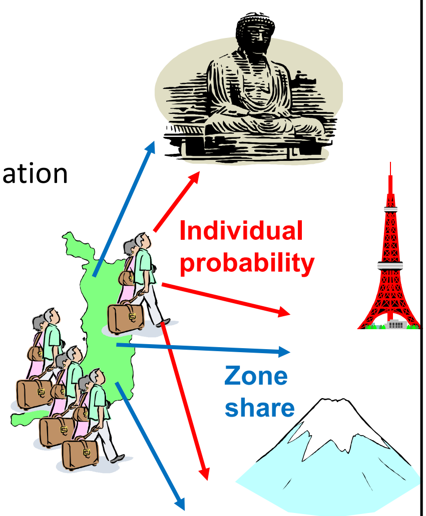
$$= \frac{\exp(\mu V_{ni})}{\exp(\mu V_{ni}) + \exp\left(\mu V_n^*\right)} = \frac{\exp(\mu V_{ni})}{\exp(\mu V_{ni}) + \exp\left\{\ln \sum_{k \ne i} \exp(\mu V_{nk})\right\}}$$

$$= \frac{\exp(\mu V_{ni})}{\sum_{k=1}^{I} \exp(\mu V_{nk})}$$

26

Type of estimation

- Disaggregate (非集計) type estimation
  - Based on individual choice
- Aggregate (集計) type estimation
  - Based on share



Utility function for disaggregate logit model

Assume linear utility function

$$V_{ni} = \sum_{m=1}^{M} \beta_m X_{nim} + \sum_{p=1}^{P} \delta_p$$

Dummy variable (to be estimated)

Explanatory variable

Parameter (to be estimated)

Scale parameter μ is set to be one

$$P_{ni} = \frac{\exp V_{ni}}{\sum_{k=1}^{I} \exp V_{nk}}$$

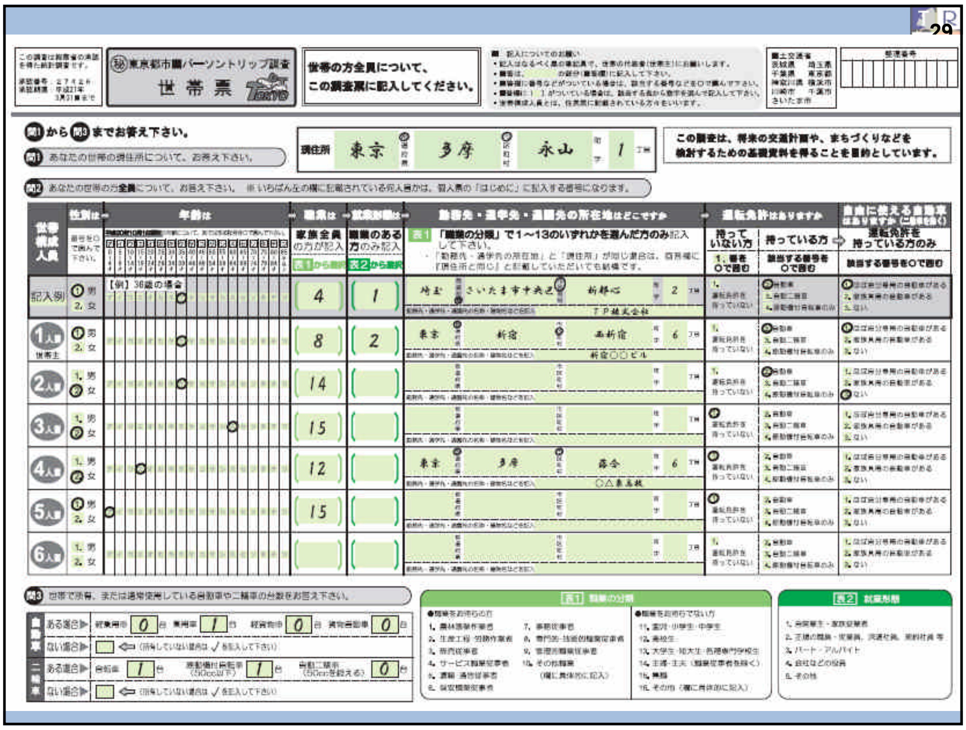
Disaggregate mode choice model

Alternative	Explanatory variable						
	Cost	LH time	AC&EG time	Frequency	# of transfer	Vehicle own	Gender
Car	○	○	×	×	×	○	○
Rail	○	○	○	○	○	×	×
Bus	○	○	○	○	○	×	×

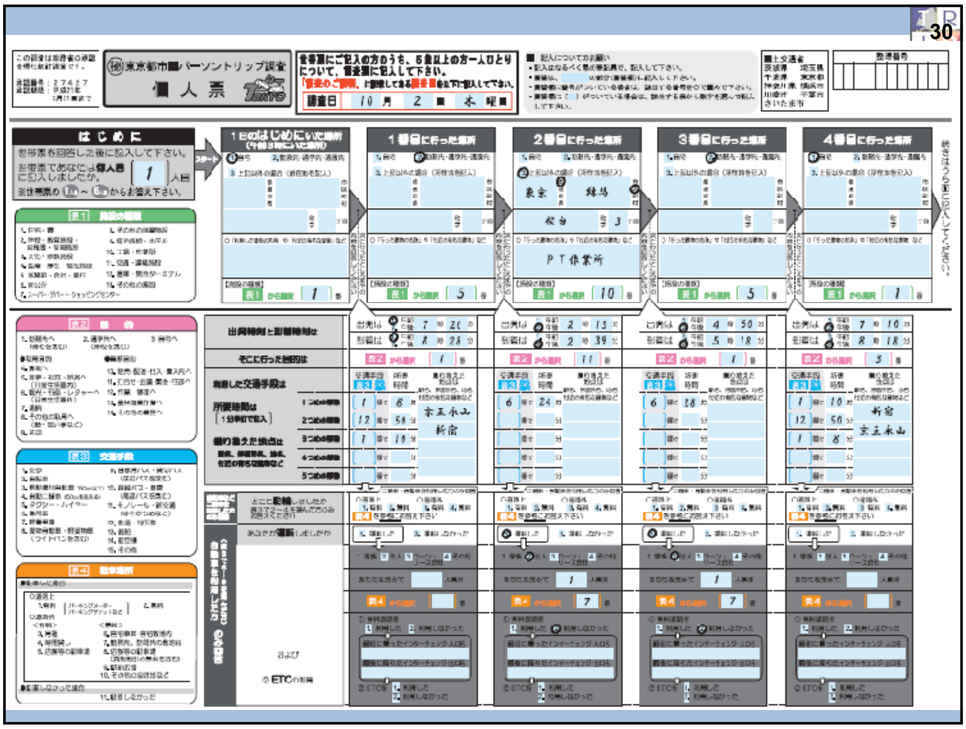
Common variables (共通変数) [Cost, LH time] — Alternative specific Variables (選択肢固有変数) [AC&EG time, Frequency, # of transfer] — Alternative specific dummy variables (選択肢固有ダミー) [Vehicle own, Gender]

$$\begin{cases} V_c = \beta_1 C_c + \beta_2 T_c + \delta_1 + \delta_2 + \delta_{0c} \\ V_r = \beta_1 C_r + \beta_2 T_r + \beta_3 A_r + \beta_4 F_r + \beta_5 N_r + \delta_{0r} \\ V_b = \beta_1 C_b + \beta_2 T_b + \beta_3 A_b + \beta_4 F_b + \beta_5 N_b \end{cases}$$

29


東京都市圏パーソントリップ調査
世帯票
世帯の方全員について、この調査票に記入してください。
問1から問3までお答え下さい。
問1 あなたの世帯の現住所について、お答え下さい。
現住所 東京 多摩 永山 1
この調査は、将来の交通計画や、まちづくりなどを検討するための基礎資料を得ることを目的としています。
問2 あなたの世帯の方全員について、お答え下さい。
問3 世帯で所有、または通常使用している自動車や二輪車の台数をお答え下さい。
表1 職業の分類
表2 就業形態

30


東京都市圏パーソントリップ調査
個人票
調査日 10 月 2 日 木 曜日
はじめに
1日のはじめにいた場所
1番目に行った場所
2番目に行った場所
3番目に行った場所
4番目に行った場所
出発時刻と到着時刻は
そこに行った目的は
利用した交通手段は
所要時間は
乗り換えた地点は
表1 施設の種類
表2 目的
表3 交通手段
表4 駐車場所
ETCの利用

31

Maximum likelihood estimation (最尤推定) (1)

Calculate the simultaneous P. D. F = likelihood by all observations

$$L(\boldsymbol{\beta},\boldsymbol{\delta}) = \prod_{n=1}^{N}\prod_{k=1}^{I}\left\{P_{ni}(\boldsymbol{\beta},\boldsymbol{\delta})^{d_{ni}}\right\}$$

$$d_{ni} = \begin{cases} 1 & \text{In case individual } n \text{ chooses alternative } i \\ 0 & \text{In case individual } n \text{ doesn't choose alternative } i \end{cases}$$

Maximum likelihood estimation
= Decide $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ that give maximum $L(\boldsymbol{\beta},\boldsymbol{\delta})$

32

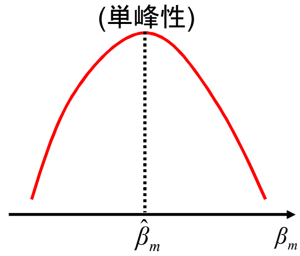
Maximum likelihood estimation (2)

Convert to logarithm likelihood (対数尤度)

$$L^* = \ln L = \sum_{n=1}^{N}\sum_{k=1}^{I} d_{ni} \ln P_{ni} \rightarrow \max$$

$$\left.\frac{\partial L^*}{\partial \beta_m}\right|_{\beta_m=\hat{\beta}_m} = 0 \qquad \left.\frac{\partial L^*}{\partial \delta_p}\right|_{\delta_p=\hat{\delta}_p} = 0$$

L^* : Unimodal distribution
(単峰性)



Maximum likelihood estimation (3)

$$\frac{\partial L^*}{\partial \beta_m} = \sum_{n=1}^{N}\sum_{k=1}^{I} d_{ni}\left(X_{nim} - \frac{\sum_{j=1}^{I} X_{njm}\exp V_{nj}}{\sum_{j=1}^{I}\exp V_{nj}} \right)$$

$$= \sum_{n=1}^{N}\sum_{k=1}^{I}(d_{ni} - P_{ni})X_{nim} = 0$$

Nonlinear simultaneous equations

Nonlinear optimization problem (非線形最適化問題)

Maximum likelihood estimation (4)

Methods of nonlinear optimization problem

-Newton-Raphson method

requires gradient vector and Hessian matrix of L^*

$$\boldsymbol{\beta}^{r+1} = \boldsymbol{\beta}^{r} - \left[\nabla^2 L(\boldsymbol{\beta}^{r})\right]^{-1} \nabla L(\boldsymbol{\beta}^{r})$$

$\nabla L(\beta^r) \leq \varepsilon \longrightarrow$ Finish

-Quasi-Newton method (BFGS)

requires gradient vector and approximation of Hessian matrix (easier calculation)

35

Goodness of fit (適合度)

Likelihood ratio index (尤度比)

Compare two likelihoods
1) in case all parameters are not significant (=0)
2) in case all parameters are significant

$$0 \leq \rho = 1 - \frac{L^*(\hat{\boldsymbol{\beta}})}{L^*(\mathbf{0})} \leq 1$$

Worse ← → Better

Percent correctly predicted (的中率)

36

Hypothesis test (仮説検定) for utility function

Null hypothesis (帰無仮説): All parameters are 0.
Alternative hypothesis(対立仮説): All parameters are not 0

↓

"Chi-square test"

$$\chi^2 \equiv -2\{L^*(\mathbf{0}) - L^*(\hat{\boldsymbol{\beta}})\} \sim \chi^2_\alpha(M+P)$$

Hypothesis test for parameters

Null hypothesis: Focused parameter is 0.
Alternative hypothesis: Focused parameter is not 0

"t test"

$$t_m \equiv \frac{\hat{\beta}_m}{\sqrt{\hat{v}_{mm}}} \sim t_\alpha (N - M - 1)$$

Element m,m of $\mathbf{Var}(\hat{\boldsymbol{\beta}})$

$$\mathbf{Var}(\hat{\boldsymbol{\beta}}) = E\left[-\nabla^2 L^*(\hat{\boldsymbol{\beta}})\right]^{-1}$$

Time saving value (**時間評価値**)

Utility function

$$U_{ni} = \beta_1 T_{ni} + \beta_2 C_{ni} + \varepsilon_{ni}$$

Travel time   Travel cost

Time saving value

$$\frac{\beta_2}{\beta_1} = Const.$$

Consumer surplus by logit model

Property of IID gumbel $E\left(\max_i U_{ni}\right)=\frac{1}{\mu}\ln\sum_k \exp(\mu V_{nk})$

In case systematic utility of some alternative is increased

⬇

Change of consumer surplus in utility term

Before change

After change

$$\Delta CS_n^U = \sum_{i\in I}\int_{\mathbf{V}_n^0}^{\mathbf{V}_n^1} P_{ni} d\mathbf{V} = \frac{1}{\mu}\ln\sum_{i\in I^1}\exp\left(\mu V_{ni}^1\right) - \frac{1}{\mu}\ln\sum_{i\in I^0}\exp\left(\mu V_{ni}^0\right)$$

⬇

Change of consumer surplus (benefit) in money term

$$\Delta CS_n^M = \frac{1}{\beta_{cost}}\Delta CS_n^U$$

Elasticity(弾性値) of logit model

Demand elasticity:

(Rate of demand change)/(Rate of price change)

⬇

Elasticity of logit model:

(Rate of probability change)/(Rate of attribute's change)

Direct elasticity (直接弾性値) $\frac{\partial P_{ni}}{\partial X_{nim}}\frac{X_{nim}}{P_{ni}} = (1-P_{ni})X_{nim}\beta_m$

Cross elasticity (交差弾性値) $\frac{\partial P_{ni}}{\partial X_{njm}}\frac{X_{njm}}{P_{ni}} = -P_{nj}X_{njm}\beta_m$

independent from *i*'s probability

IIA property (IIA特性) (1)

$$\frac{P_{ni}}{P_{nj}} = \exp(V_{ni} - V_{nj})$$

Ratio of choice probability between two alternatives
is not affected by systematic utility of other alternatives
in choice set

↓

Independence from Irrelevant (無関係な) Alternatives

↓

Most important property of logit model

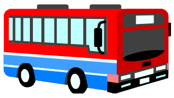
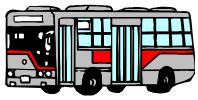
IIA property (2)

- Advantage
  - Estimation using sub choice set
    - Large choice set problem
  - Forecast the probability of newly introduced alternative
- Disadvantage
  - "Similarity" (=Correlation of error term) among alternatives

43

IIA property (3)

"Problem of red color bus and blue color bus"

  + 

Sys. utility	V	V	V
	0.5	0.5	
Prob	0.33?	0.33?	0.33?
	0.5	0.25	0.25

Independence of error term ? Logit model ?

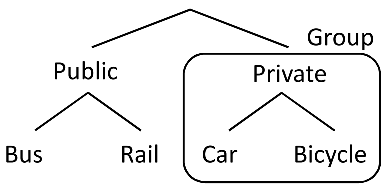
44

Non-IIA property of alternatives in transportation field

1. Mode choice
   - Public (Rail, Bus) & Private (Car)
2. Route choice
   - Partially overlapped routes
3. Destination choice
   - Adjacent destinations

45

Non-IIA models – GEV model



GEV = Generalized Extreme Value

$$\varepsilon_n \sim \exp\left[-\sum_{g=1}^{G} \alpha_g \left\{\sum_{j\in I_n^g} \exp\left(-\frac{\varepsilon_{nj}}{\lambda_g}\right)\right\}^{\lambda_g}\right]$$

I alternatives are classified into G groups

$$P_{ni} = \frac{\exp(V_{ni}/\lambda_g)\left\{\sum_{j\in I_n^g} \exp(V_{nj}/\lambda_g)\right\}^{\lambda_g - 1}}{\sum_{g=1}^{G}\left\{\sum_{j=I_n^g} \exp(V_{nj}/\lambda_g)\right\}^{\lambda_g}}$$

$$P_{ni} / P_{nj} = \frac{\exp(V_{ni}/\lambda_g)\left\{\sum_{j\in I_n^g} \exp(V_{nj}/\lambda_g)\right\}^{\lambda_g - 1}}{\exp(V_{nj}/\lambda_h)\left\{\sum_{j\in I_n^h} \exp(V_{nj}/\lambda_h)\right\}^{\lambda_h - 1}}$$

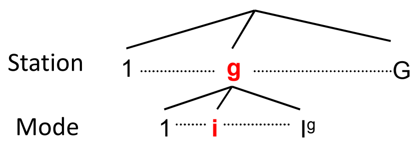
$$\mathrm{cov}(\varepsilon_{ni}, \varepsilon_{nj}) \begin{cases} = 0 (i \in g, j \notin g) \\ \neq 0 (i \in g, j \in g) \end{cases}$$

If $\lambda_g = 1$, MNL model

→If $g = h$, IIA

46

GEV families – Nested logit model(1)



Choice probability

$$P_{ngi} = P_{ng} P_{ni|g}$$

Choice probability of nest g

Conditional probability (条件確率) of alternative i given that nest g is chosen

Utility function

$$U_{ngi} = V_{ng} + V_{ni} + V_{ngi} + \varepsilon_{ng} + \varepsilon_{ngi}$$

IID $G(0, \mu)$

GEV families – Nested logit model(2)

Choice probability of nest

$$P_{ng} = \Pr\left[\max_{i\in I_n^g} U_{ngi} \ge \max_{i\in I_n^h} U_{nhi}, h \neq g\right]$$

$$= \Pr\left[V_{ng} + \varepsilon_{ng} + \max_{i\in I_n^g}\left(V_{ni} + V_{ngi} + \varepsilon_{ngi}\right) \ge V_{nh} + \varepsilon_{nh} + \max_{i\in I_n^h}\left(V_{ni} + V_{nhi} + \varepsilon_{nhi}\right), h \neq g\right]$$

Inclusive value
(合成効用)

$$V'_{ng} \equiv \frac{1}{\mu}\ln\sum_{i\in I_n^g}\exp\{\mu(V_{ni} + V_{ngi})\} \qquad \varepsilon'_{ng} = \max(V_{ni} + V_{ngi} + \varepsilon_{ngi}) - V'_{ng}$$

$$= \Pr\left[V_{ng} + V'_{ng} + \varepsilon_{ng} + \varepsilon'_{ng} \ge V_{nh} + V'_{nh} + \varepsilon_{nh} + \varepsilon'_{nh}, h \neq g\right] = \frac{\exp\{\mu'(V_{ng} + V'_{ng})\}}{\sum_{g\in G}\exp\{\mu'(V_{ng} + V'_{ng})\}}$$

IID $G(0, \mu')$

Conditional probability

$$P_{ni|g} = \Pr\left[V_{ni} + V_{ngi} + \varepsilon_{ngi} \ge V_{nj} + V_{ngj} + \varepsilon_{ngj}, j \neq i\right] = \frac{\exp\{\mu(V_{ni} + V_{ngi})\}}{\sum_{i\in I_n^g}\exp\{\mu(V_{ni} + V_{ngi})\}}$$

GEV families – Nested logit model(3)

Choice probability

$$P_{ngi} = \frac{\exp\{\mu'(V_{ng} + V'_{ng})\}}{\sum_{g\in G}\exp\{\mu'(V_{ng} + V'_{ng})\}} \bullet \frac{\exp\{\mu(V_{ni} + V_{ngi})\}}{\sum_{i\in I_n^g}\exp\{\mu(V_{ni} + V_{ngi})\}}$$

Important property of the ratio of variances in NL model

$$\frac{\mu'}{\mu} = \sqrt{\frac{\mathrm{var}(\varepsilon_{ngi})}{\mathrm{var}(\varepsilon_{ng} + \varepsilon'_{ng})}} = \sqrt{\frac{\mathrm{var}(\varepsilon_{ngi})}{\mathrm{var}(\varepsilon_{ngi}) + \mathrm{var}(\varepsilon_{ng})}} \le 1.0$$

49

Non-IIA models– Probit model

$$U_{ni} = V_{ni} + \varepsilon_{ni}$$

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma}) = (2\pi)^{-\frac{I}{2}} \exp\left(-\frac{1}{2}\boldsymbol{\varepsilon}'\boldsymbol{\Sigma}\boldsymbol{\varepsilon}\right) \equiv \phi(\boldsymbol{\varepsilon})$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & & sym. \\ \vdots & \ddots & \\ \sigma_{1I} & \cdots & \sigma_I^2 \end{pmatrix}$$

$$P_{ni} = \int_{-\infty}^{V_{ni}-V_{n1}+\varepsilon_{ni}} \cdots \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{V_{ni}-V_{nI}+\varepsilon_{ni}} \phi(\boldsymbol{\varepsilon}) d\boldsymbol{\varepsilon}$$

Requires multiple integral
= Application of simulation methods
(Random drawing)

Direct explanation of correlation among alternatives

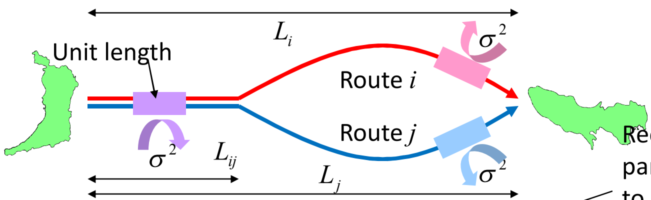
All elements in covariance matrix to be estimated
= Need of structured covariance

50

Probit families– MNPSC model

MultiNomial Probit with Structured Covariance

Ex. route choice in overlapped condition



Reduction of parameters to be estimated

$$\begin{cases} \text{var}(\varepsilon_{ni}) = L_i\sigma^2 + \sigma_0^2 \\ \text{cov}(\varepsilon_{ni}, \varepsilon_{nj}) = L_{ij}\sigma^2 \end{cases}$$

$$\Sigma = \sigma_0^2 \begin{pmatrix} \delta L_1 + 1 & & sym. \\ \vdots & \ddots & \\ \delta L_{1I} & \cdots & \delta L_I + 1 \end{pmatrix}$$

$$\delta = \sigma^2 / \sigma_0^2$$

51

Integrated model – Mixed logit model (MXL)

MNL model

$$P_{ni} = \frac{\exp\{V_{ni}(\boldsymbol{\beta})\}}{\sum_{i \in I} \exp\{V_{ni}(\boldsymbol{\beta})\}}$$

MXL model

$$P_{ni} = \int \frac{\exp\{V_{ni}(\boldsymbol{\beta})\}}{\sum_{i \in I} \exp\{V_{ni}(\boldsymbol{\beta})\}} f(\boldsymbol{\beta}) d\boldsymbol{\beta}$$

P.D.F of parameters

Error components model

$$V_{ni} = \boldsymbol{\beta}'\mathbf{x}_{ni} + \boldsymbol{\mu}'\mathbf{z}_{ni} + \varepsilon_{ni}$$

$$\boldsymbol{\mu} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$$

Observed variables

Random coefficient model

$$V_{ni} = \boldsymbol{\beta}'\mathbf{x}_{ni} + \varepsilon_{ni}$$

$$\boldsymbol{\beta} \sim N(\mathbf{b}, \boldsymbol{\Sigma})$$

Heterogeneous(異質な) individual

52

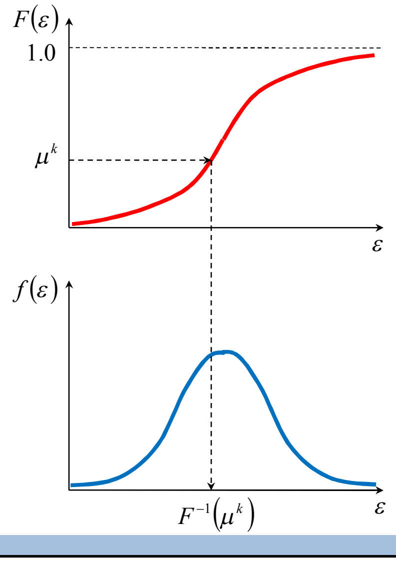
Estimation by simulation

$$U_{ni} = V_{ni} + \varepsilon_{ni}$$

Using a draw from probability density

In case of logit model, choice probability is calculated by generating K draws from standard uniform distribution

$$P_{ni} = \frac{1}{K} \sum_{k=1}^{K} \frac{\exp\{U_{ni}(\mu^k)\}}{\sum_{i \in I} \exp\{U_{ni}(\mu^k)\}}$$



53

GHK simulator for probit model (1)

Suppose the number of alternatives is three and alternative 1 is chosen…

$$\begin{cases} U_1^* \le 0 \\ U_2^* \le 0 \end{cases}$$

lower triangular matrix by Choleski decomposition

$$\begin{pmatrix} U_1^* \\ U_2^* \end{pmatrix} = \begin{pmatrix} U_2 - U_1 \\ U_3 - U_1 \end{pmatrix} = \begin{pmatrix} V_2 - V_1 \\ V_3 - V_1 \end{pmatrix} + \mathbf{WZ} = \begin{pmatrix} V_2 - V_1 \\ V_3 - V_1 \end{pmatrix} + \begin{pmatrix} w_{11} & 0 \\ w_{21} & w_{22} \end{pmatrix} \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix}$$

draw from standard normal distribution

$$\mathbf{WW}' = \mathbf{M}_1 \mathbf{\Sigma} \mathbf{\Sigma}' \mathbf{M}_1'$$

$$\mathbf{M}_1 = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \quad \mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix}$$

54

GHK simulator for probit model (2)

Choice probability for alt 1 when vector of random variable is given is expressed as follows…

$$P_1^\zeta = \Pr\left(U_1^* \le 0, U_2^* \le 0\right) = \Pr\left(U_1^* \le 0\right) \times \Pr\left(U_2^* \le 0 \mid U_1^* \le 0\right)$$

$$= \Pr\left(\zeta_1 \le a_1\right) \times \Pr\left(\zeta_2 \le a_2(\zeta_1) \mid \zeta_1 \le a_1\right)$$

$$U_1^* \le 0 \rightarrow \zeta_1 \le -\frac{V_2 - V_1}{w_{11}} \equiv a_1$$

$$U_2^* \le 0 \rightarrow \zeta_2 \le -\frac{w_{21}\zeta_1 + V_3 - V_1}{w_{22}} \equiv a_2(\zeta_1)$$

Using C.D.F. of standard normal distribution…

$$P_1^\zeta = \Phi(a_1)\Phi\{a_2(\zeta_1)\}$$

$$\zeta_1 = \Phi^{-1}\{\mu_1 \Phi(a_1)\}$$

only draw from standard uniform distribution

GHK simulator for probit model (3)

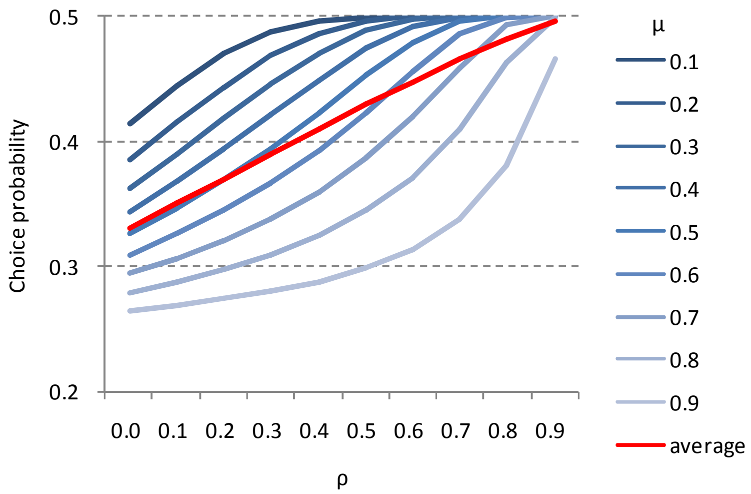
Example in “problem of red color bus and blue color bus”

covariance matrix $\mathbf{\Sigma} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \rho \\ 0 & \rho & 1 \end{pmatrix}$

In case that alt 1 is chosen, matrix $\mathbf{W}$ is expressed by

$$\mathbf{W} = \begin{pmatrix} \sqrt{2+\rho^2} & 0 \\ \dfrac{1+2\rho}{\sqrt{2+\rho^2}} & \sqrt{\dfrac{\rho^4 - 4\rho + 3}{2+\rho^2}} \end{pmatrix}$$

GHK simulator for probit model (4)



57

Application of GHK simulation method

MNPSC estimation using data of railway route choice in Tokyo Metropolitan area of three possible routes

estimation method	numerical integration		GHK 25 draws*		GHK 50 draws*		GHK 100 draws*	
explanatory value	estimate	t-value	estimate	t-value	estimate	t-value	estimate	t-value
line haul cost (yen)	-0.00583	-5.22	-0.00591	-5.21	-0.00594	-5.23	-0.00594	-5.19
access travel time (min)	-0.127	-5.23	-0.129	-5.15	-0.129	-5.22	-0.129	-5.16
egress travel time (min)	-0.151	-5.45	-0.152	-5.45	-0.153	-5.49	-0.152	-5.43
line haul travel time (min)	-0.0695	-5.80	-0.0704	-5.70	-0.0704	-5.77	-0.0703	-5.68
time for transfer (upstair) (min)	-0.335	-2.44	-0.343	-2.42	-0.333	-2.42	-0.335	-2.42
time for transfer (downstair & horizontal) (min)	-0.116	-5.21	-0.117	-5.24	-0.117	-5.19	-0.118	-5.17
waiting time (min)	-0.118	-3.90	-0.119	-3.86	-0.119	-3.89	-0.120	-3.88
number of transfers (times)	-0.382	-3.92	-0.391	-3.93	-0.392	-3.99	-0.387	-3.92
(congestion rate)^2 * line haul time (%^{2}min)	-9.1E-08	-0.82	-8.9E-08	-0.75	-9E-08	-0.85	-9E-08	-0.81
ratio of σ^2 to σ_0^2	0.302	1.45	0.338	1.45	0.333	1.50	0.334	1.46
number of observations	1074		1074		1074		1074	
likelihood ratio	0.182		0.184		0.184		0.184	
average parameter error rate(%) **	NA		3.1		2.5		2.3	
maximum parameter error rate(%) ***	NA		13.0		9.9		10.0	

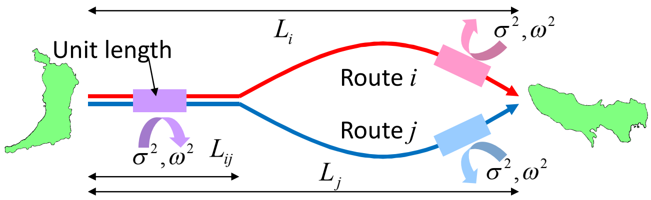
*choose the case in that parameter error rate is minimum

**average ratio of estimates by numerical integration to that by GHK method

***maximum ratio of estimates by numerical integration to that by GHK method

58

Comparison: MNPSC & MXL (1)



MNPSC

$$\begin{cases} \mathrm{var}(\varepsilon_{ni}) = L_i\sigma^2 + \sigma_0^2 \\ \mathrm{cov}(\varepsilon_{ni}, \varepsilon_{nj}) = L_{ij}\sigma^2 \end{cases}$$

$$\Sigma = \sigma_0^2 \begin{pmatrix} \delta L_1 + 1 & & sym. \\ \vdots & \ddots & \\ \delta L_{1I} & \cdots & \delta L_I + 1 \end{pmatrix}$$

$$\delta = \sigma^2 / \sigma_0^2$$

MXL

$$\begin{cases} \mathrm{var}(\varepsilon_{ni}) = L_i\omega^2 + \frac{\pi^2}{6} \\ \mathrm{cov}(\varepsilon_{ni}, \varepsilon_{nj}) = L_{ij}\omega^2 \end{cases}$$

$$\Sigma = \frac{\pi^2}{6} \begin{pmatrix} \lambda L_1 + 1 & & sym. \\ \vdots & \ddots & \\ \lambda L_{1I} & \cdots & \lambda L_I + 1 \end{pmatrix}$$

$$\lambda = 6\omega^2 / \pi^2$$

59

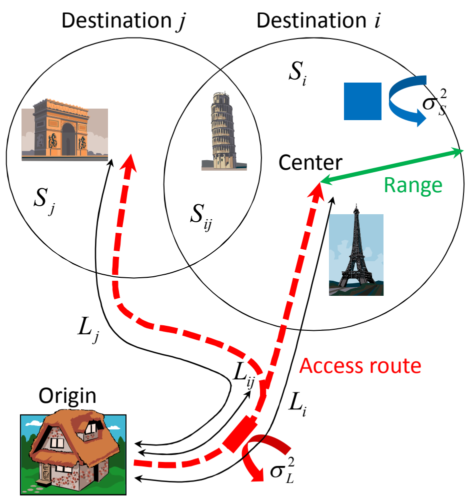
Comparison: MNPSC & MXL (2)

Estimation using data of railway route choice in Tokyo Metropolitan area of three possible routes

model type	MNPSC		MXL	
explanatory value	estimate	t-value	estimate	t-value
line haul cost (yen)	-0.00184	-3.47	-0.00255	-3.81
access travel time (min)	-0.107	-7.97	-0.136	-9.69
egress travel time (min)	-0.0755	-6.16	-0.0962	-7.03
line haul travel time (min)	-0.0112	-1.01	-0.0114	-0.866
time for transfer (min)	-0.0284	-1.85	-0.0327	-1.75
waiting time (min)	-0.123	-4.30	-0.171	-4.94
number of transfers (times)	-0.194	-1.35	-0.274	-1.51
(congestion rate)^2 * line haul time ((%/100)^{2}min)	-0.00726	-2.24	-0.00892	-2.36
δ or λ	0.0123	0.739	0.0264	1.38
number of observations	637		637	
likelihood ratio	0.262		0.272	

60

MXL for tourism destination choice (1)



Utility function

$$U_i = V_i + \varepsilon_i$$

$$\varepsilon_i = \eta_i + \omega_i + \xi_i$$

error by destination

error by access route

error by facility (area)

$$\begin{cases} \mathrm{var}(\eta_i) = S_i\sigma_S^2 \\ \mathrm{cov}(\eta_i,\eta_j) = S_{ij}\sigma_S^2 \\ \mathrm{var}(\omega_i) = L_i\sigma_L^2 \\ \mathrm{cov}(\omega_i,\omega_j) = L_{ij}\sigma_L^2 \\ \mathrm{var}(\xi_i) = \sigma_0^2 \\ \mathrm{cov}(\xi_i,\xi_j) = 0 \end{cases}$$

MXL for tourism destination choice (2)

Covariance matrix

$$\mathbf{\Sigma} = \sigma_0^2 \begin{pmatrix} 1+\gamma S_1 + \delta L_1 & \cdots & \gamma S_{1I} + \delta L_{1I} \\ \vdots & \ddots & \vdots \\ \gamma S_{1I} + \delta L_{1I} & \cdots & 1+\gamma S_I + \delta L_I \end{pmatrix}$$

$$\gamma = \frac{\sigma_S^2}{\sigma_0^2}, \delta = \frac{\sigma_L^2}{\sigma_0^2}$$

Choice probability

$$P_i = \iint \frac{\exp(V_i + \eta_i + \omega_i)}{\sum_i \exp(V_i + \eta_i + \omega_i)} f(\eta) g(\omega) d\eta d\omega$$

MXL for tourism destination choice (3)

Estimation result

Estimation using data of one day tourism destination choice in Tokyo Metropolitan area of ten possible destinations

model type	MNL		MXL(1)	
explanatory value	estimate	t-value	estimate	t-value
generalized cost / ln(annual income)	-0.0409	-6.82	-0.0415	-6.71
attraction of destination	0.0738	6.75	-0.0755	6.63
γ	---	---	0.0111	0.570
δ	---	---	---	---
number of observations	269		269	
likelihood ratio	0.117		0.118	
model type	**MXL(2)**		**MXL(3)**	
explanatory value	estimate	t-value	estimate	t-value
generalized cost / ln(annual income)	-0.0426	-6.44	-0.0426	-6.46
attraction of destination	0.0735	6.70	0.0741	6.63
γ	---	---	0.0138	0.188
δ	0.0286	0.674	0.0219	0.527
number of observations	269		269	
likelihood ratio	0.118		0.119	

63

Importance of the research on choice set

- Appearance of non-IIA choice models
  - Choice probability is easily affected the combination of alternatives in choice set
- Unrealistic assumption of rationality
- Weak point in demand forecasting process
  - Theoretical difficulty vs Practical convenience
  - Large choice set problem (tourism destinations)

64

Criticism of rationality

- Assumptions in rational choice (Simon)
  - Decision maker knows all of the alternatives and their attributes
  - Decision maker knows the probability distribution of uncertainty
  - Decision maker chooses the alternative with maximum expected utility
- Difference form realistic choice behaviors
  - Occasional choice, Habitual choice, Following choice, etc.
  - Change of preference in long term

65

Limitations of rationality – Allais's paradox (1953)

Problem 1: Which lottery is better?
A1: $100 (0.33), $90 (0.66), $0 (0.01)
B1: $90 (1.00)

→ "Of course, B1"

Problem 2: Which lottery is better?
A2: $100 (0.33), $0 (0.67)
B2: $90 (0.34), $0 (0.66)

→ "Of course, A2"

Why different?

Problem 1': Which lottery is better?
A1': ~~$90 (0.66)~~, $0 (0.01), $100 (0.33)
B1': ~~$90 (0.66)~~, $90 (0.34)

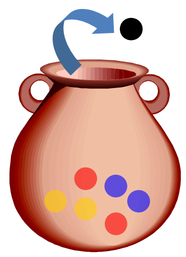
Problem 2: Which lottery is better?
A2': ~~$0 (0.66)~~, $0 (0.01), $100 (0.33)
B2': ~~$0 (0.66)~~, $90 (0.34)

Same problem!

66

Limitations of rationality – Ellsberg's paradox (1961)

Draw one



90 balls
red=30, blue=?
yellow=?
bule+yellow=60

Which game do you like?
game 1: If red $100, If not red $0
game 2: If blue $100, If not blue $0

Which game do you like?
game 3: If red or yellow $100, If blue $0
game 4: If blue or yellow $100, If red $0

If you prefer game 1…

The subjective probability you draw a blue ball should be less than 0.333

You should prefer game 3 rather than game 4!

67

Limitations of rationality – Framing effect (Tversky & Kahneman, 1981)

Suppose You are one of 600 people who have infectious disease…

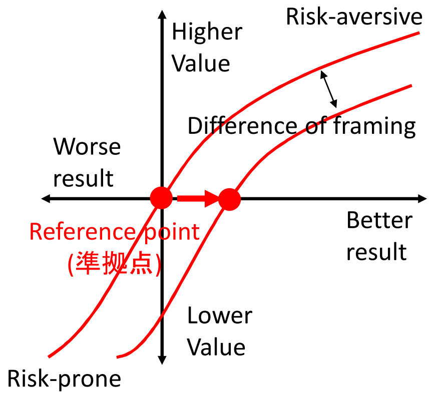
Which countermeasure do you accept?
by A: 200 people will survive
by B: All will survive with probability of 0.33 & none will survive with probability of 0.66

Which countermeasure do you accept?
by C: 400 people will die
by D: None will die with probability of 0.33 & all will die with probability of 0.66

Win aspect leads to risk-aversive choice
Lose aspect leads to risk-prone choice

68

Prospect theory



Bounded rationality (限定合理性) (Simon, 1987)

- Decision maker will take into account huge "cost" for obtaining full information of all possible alternatives.
  - Choice set generating process
  - Heuristic decision making process
  - Satisfactory maximization process

Proposed decision strategies (決定方略)

- Compensatory (補償型)
  - Weighted multi-attribute utility
  - Maximum wining percentage
- Non-compensatory (非補償型)
  - Conjunctive (連結) (Minimum requirement for all attributes)
  - Disjunctive (分離) (Minimum requirement for at least one attribute)
  - Lexicographic (辞書編纂) (Ranking of attributes, choose best ones in focused attribute)
  - Eliminate by aspect (EBA) (Attributes → 0-1 aspect)