

行動モデル夏の学校 2008

(1) 所要時間信頼性の経済評価と行動モデル

(2) デンマーク・スイス方面の研究レビュー (2007-8)

福田 大輔

東京工業大学 大学院理工学研究科 土木工学専攻

2008年9月21日

# (1) 所要時間信頼性の経済評価と 行動モデル

## 道路の交通混雑と所要時間の不確実性

- 所要時間の“絶対的な”長さを増加させるだけでなく、その変動性 (Variability) を高め、予測をより困難なものとする。
- 旅行者にとって、この不確実性は追加的コストとして認識される。

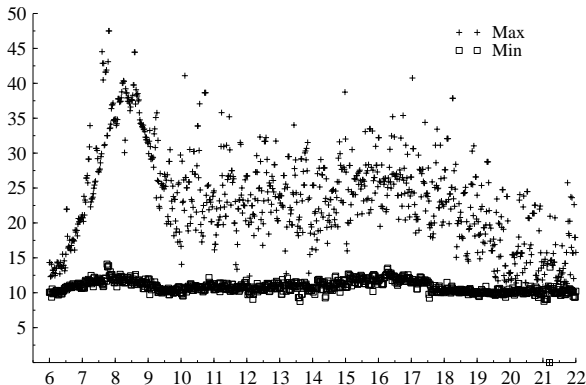


Figure: 流入時刻 (横軸 : 時) 別に見た最大・最小所要時間 (縦軸 : 分)

## プロジェクト評価の実状：

- 日本・アメリカ等：BTI(Buffer Time Index) 値に代表されるような“Performance Measure”の利用が中心  
Economic appraisal における直接適用は難しい
- 英国・オランダ・スウェーデン等：SACTRA(Standing Advisory Committee for Trunk Road Assessment, 英)等, Congestion Charge のCBAにおいて積極的導入の動き

## VTTV(Value of Travel Time Variability)：

- 所要時間分布の広がり (e.g. 標準偏差, タイル値) の変化に対する, 利用者の限界的な支払い意思額
- 行動モデルにおける VTTV の取り扱い：
  - ① 平均 - 分散アプローチ  
e.g. Jackson & Jucker (1981), Lam & Small (2001), Small et al. (2005)
  - ② スケジューリングアプローチ  
e.g. Small (1982), Noland & Small (1995), Bates et al. (2001)

ドライバーの(間接)効用は, 費用  $C$ , 期待旅行時間  $ET$ , 旅行時間の標準偏差  $\sigma_T$  に規定される:

$$U = \delta C + \alpha ET + \rho \sigma_T$$

$\delta, \alpha, \rho$ : 各限界不効用

- その簡便さ故, 1980年初頭より多くの適用例, e.g., Jackson & Jucker (1981).
- 所要時間の中央値を  $ET$  の代わりに, 90・50パーセンタイル値の差を  $\sigma_T$  の代わりに用いる例も, e.g., Small et al. (2005)
- Value of travel time (VTT) =  $\alpha/\delta$ , **VTTV =  $\rho/\delta$**
- アウトカムそのものが効用関数の引数でない ミクロ経済学的基礎は, 必ずしも確固たるものとはなっていない.

## Modelling Behavior (2) スケジューリングアプローチ

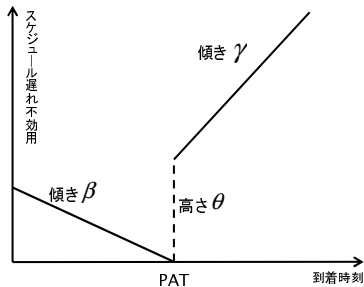
ドライバーの(間接)効用は,費用  $C$ , 時間  $T$  に加えて, 希望到着時刻 (Preferred arrival time: PAT) からどれほど乖離して早着/遅着したかによって決まる (Small (1982)).

また, これらの属性は, 出発時刻  $t_h$  に依存する.

$$U(t_h) = \delta C + \alpha T + \beta SDE + \gamma SDL + \theta D_L : \text{Small's Function}$$

**SDE**: Schedule delay early, **SDL**: Schedule delay late,

**D<sub>L</sub>**: 遅着ダミー,  $\delta, \alpha, \beta, \theta$ : 各限界不効用.



(続き)

- PAT からの遅れは，同程度の早着よりも大きな不効用として評価される．
- アウトカムがそのまま引数になっている (平均 - 分散アプローチに対しての利点).
- 所要時間の不確実性を考慮したとき，極めて限定的な条件下で平均 - 分散アプローチに等価 (Noland & Small (1995)).

期待効用関数：

$$EU(t_h) = \delta C + \alpha ET + \beta E(SDE) + \gamma E(SDL) + \theta P_L$$

$P_L$ : 遅着確率

所要時間分布が出発時刻に独立で，一様分布/指数分布に従う場合，

$$EU^* = \delta C + \alpha ET + \theta P_L^* + H(\alpha, \beta, \gamma, \theta, b, \Delta)b$$

b: 指数分布のパラメータ (=平均及び標準偏差)

- 「所要時間と出発時刻の独立の仮定性」の妥当性
- 「指数分布を所要時間分布として採用すること」の妥当性

- 所要時間信頼性向上の経済便益は，プロジェクト総便益の中でもそれなりのシェアを占める可能性
- e.g. London Congestion Charge における時間信頼性向上の便益は，全体の2割強という試算も (SACTRA)
- 時間信頼性を経済評価の枠組みに乗せるためには，
  - ① 理論的に整合・統一されており，
  - ② 旅行者の選好情報に基づいて推定可能で，
  - ③ (時間変動を表すための) 交通流モデルを用いた応用が可能な，VTTV の定義・理論構築・推計方法の確立が必要．



- 道路交通における時間信頼性価値 (VTTV: Value of Travel Time Variability) の推計方法の開発
- 特に本研究では，統合アプローチ  
Fosgerau & Karlström (2007), Fosgerau & Fukuda (2008) の  
実適用可能性について検討
- 統合アプローチの概要
  - 通勤者のコスト関数はスケジューリングアプローチに準拠．
  - しかし，モデル実適用に当たって PAT の情報を必要としない．
  - 最適解は平均 - 分散アプローチに帰着
  - Noland & Small (1995) と異なり，任意の所要時間分布に  
対して成り立つ

### 単一リンクにおける日々の所要時間変動を考慮した ドライバーの出発時刻選択問題

- Travel time = Free flow time + Systematic delay  
+ Unexplained delay
- Travel time variability の定義 : Variation in unexplained delay
- 経路選択は考えない .
- 所要時間  $T$  (確率変数) を Location-Scale 型で表記 :

$$T = \mu + \sigma X$$

$X$  : 基準化所要時間 (密度関数  $\phi$ , 分布関数  $\Phi$ ) 平均 0

$\mu$  : 位置関数 (=期待値),

$\sigma$  : 尺度関数 (=標準偏差, 四分位範囲等)

- ひとまず当面は,  $\mu$  及び  $\sigma$  が出発時刻によらず一定と考える .

- ドライバーの  $PAT = 0$  と設定 (表記簡略化のため)
- ドライバーのコスト関数は, (1) 実所要時間  $T$ , (2) “ヘッド・スタート (出発時刻と  $PAT$  との差)”  $D$  によって規定されると仮定
- すなわち, 出発時刻は  $-D$
- スケジューリングアプローチに基づくコスト関数の特定化:

$$C(D, T) = \eta D + \lambda(T - D)^+ + \omega T$$

where  $(T - D)^+$ :  $(T - D)$  が正の時その値, 負の時  $0$

- 第一項: 早 “発” 不効用, 第二項: 遅着不効用, 第三項: 移動時間の不効用
- Noland & Small (1995) を (遅着ダミーを除いて) 再定式化したものであり, それに等価.

期待コスト最小化問題：

$$EC^* = \min_D \left[ \eta D + \lambda \int_{\frac{D-\mu}{\sigma}}^{\infty} (\mu + \sigma x - D) \phi(x) dx + \omega \mu \right]$$

一階の条件：  $\Phi \left( \frac{D^* - \mu}{\sigma} \right) = 1 - \frac{\eta}{\lambda}$

$\frac{\eta}{\lambda}$ : 最適遅着確率 (optimal prob. of being late, Bates et al. (2001))

上式を書き改めると，

$$D^* = \mu + \sigma \Phi^{-1} \left( 1 - \frac{\eta}{\lambda} \right)$$

Xの分布のみが，その  $1 - \frac{\eta}{\lambda}$ th タイル値を介して，最適ヘッドスタートに寄与する．

従って，Small Function のように折れ曲がったコスト関数でも，最適ヘッドスタートは  $\mu$  と  $\sigma$  に関して線形となる．

## 最小化期待コスト：

$$EC^* = (\eta + \omega)\mu + \lambda H\left(\Phi, \frac{\eta}{\lambda}\right)\sigma,$$

$$\left[ \text{但し } H\left(\Phi, \frac{\eta}{\lambda}\right) = \int_{1-\frac{\eta}{\lambda}}^1 \Phi^{-1}(v) dv. \right]$$

- $(\eta + \omega)$ ：平均所要時間の限界費用 (or VTT)
- $\lambda H\left(\Phi, \frac{\eta}{\lambda}\right)$ ：所要時間変動の限界費用 (or VTTV)
- VTTV は、ドライバーのスケジュールパラメータ  $(\lambda, \eta)$  のみならず、基準化所要時間分布の形状  $(\Phi)$  にも依存する。  
標準偏差 (or 四分位範囲等) を、所要時間分布の Variability を表す引数として用い、関連するパラメータは、ドライバーのスケジューリングモデル、並びに、(基準化) 所要時間分布の特性より推計すれば良い。

- (留意点) 平均所要時間  $\mu$  , 及び , 尺度パラメータ  $\sigma$  が  
出発時刻  $D$  に依存しないことを 暗黙に仮定
- $\mu$  と  $\sigma$  が一定の場合の最適ヘッドスタート  $D_0$  を Pivot とす  
る線形関数として ,  $\mu$  と  $\sigma$  が表される場合 :

$$\mu = \mu_0 + \mu'(D - D_0)$$

$$\sigma = \sigma_0 + \sigma'(D - D_0)$$

$$D_0 \equiv \mu_0 + \sigma_0 \Phi^{-1} \left( 1 - \frac{\eta}{\lambda} \right)$$

- **X  $\perp$  D** ならば ,  $\mu$  と  $\sigma$  を一定と仮定した先の結果を用いて  
VTTV を精度良く近似的できる  
[Fosgerau & Karlström (2007), Appendix 2]
- すなわち , 上記設定における時間信頼性価値を  
VTTV( $\mu'$ ,  $\sigma'$ ) とすると , VTTV( $\mu'$ ,  $\sigma'$ )  $\approx$  VTTV(0, 0)  
**本研究で直交性 (独立性) を検証**

- 経路レベルでの VTTV 計測への拡張可能性：
  - 基準化所要時間  $X$  の分布  $\Phi$  は，セクション単位だと比較的容易に推計可能
  - リンクレベルの VTTV  $\xrightarrow{\text{集計化}}$  経路レベルの VTTV But how?
  - 分布の形状：非対称・裾厚 [Fosgerau & Fukuda (2008)]

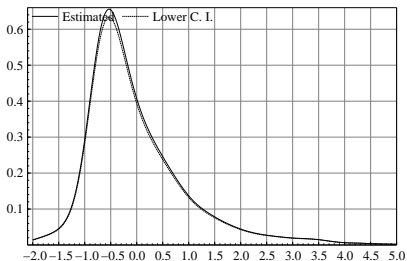
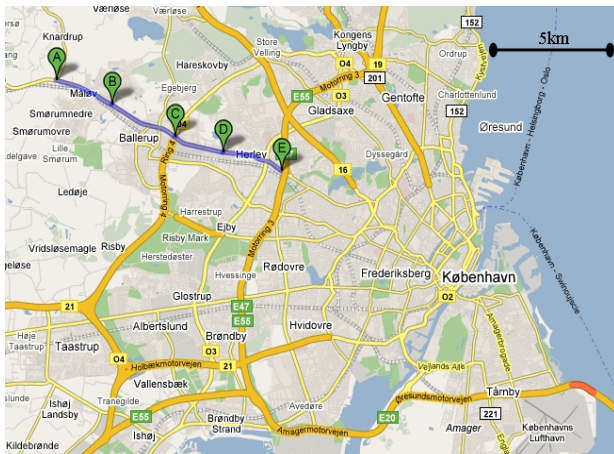


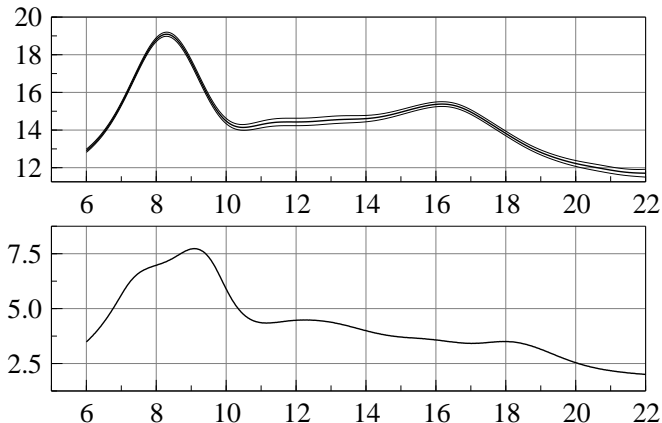
Figure:  $\phi(X)$  のノンパラメトリック密度推定結果 [カーネル法]

本研究で安定分布 (Stable Distribution) の適合可能性を検証

- コペンハーゲンにある放射状道路 “Frederikssundsvej”
- 全長 11.263 km の区間をさらに四区間に細分割
- 約 4ヶ月間の (流入時刻, 所要時間) のデータベース





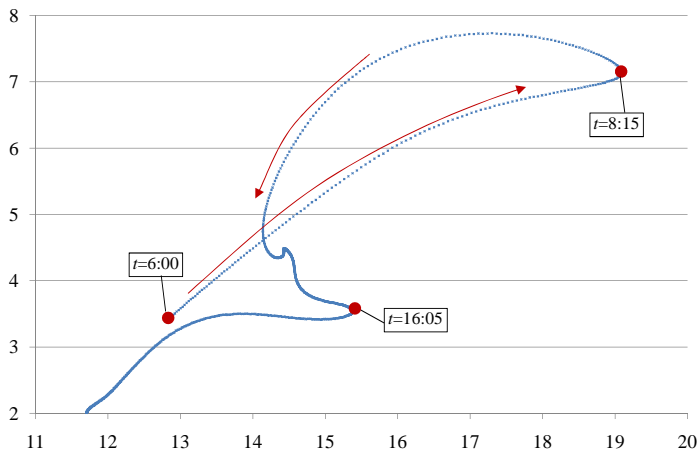


ノンパラメトリック回帰分析の結果 [ A E の 11.263km 区間 ]

(上) 条件付平均所要時間 [ $\mu(t) = E[T|t]$ ]

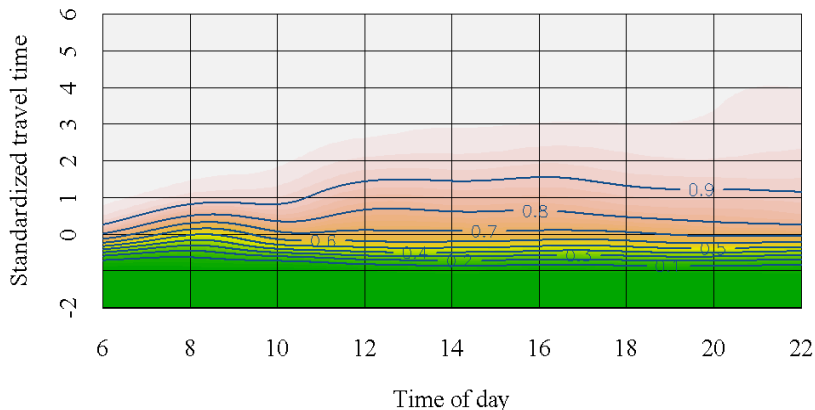
(下) 条件付四分位範囲 [ $\sigma(t) = q_{0.75}(t) - q_{0.25}(t)$ ]

# 平均所要時間と所要時間変動の相関関係



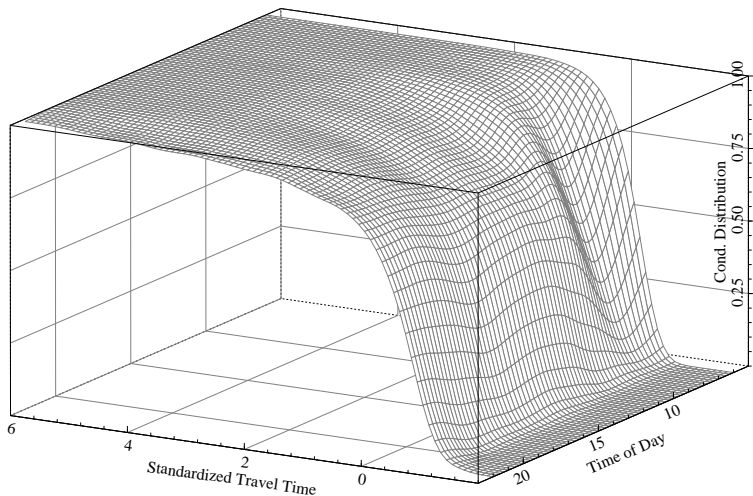
( $\hat{\mu}(t)$ ,  $\hat{\sigma}(t)$ ) の散布図 (分)

## 基準化所要時間の出発時刻独立性の検証



基準化所要時間の条件付分布関数等高線が横軸（時刻）に対してほぼ並行

→ 基準化所要時間は出発時刻に対して概ね独立

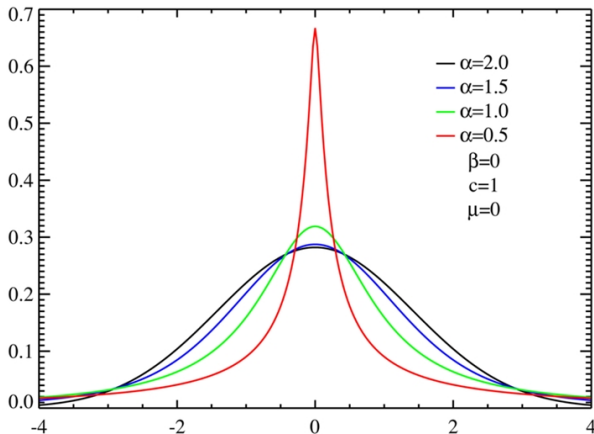


基準化所要時間の条件付分布関数 (3次元表記)

- Heavy-tailed data のモデリングに多用 (ファイナンス, 情報通信, 自然科学等)
- $X \sim S_t(\alpha, \beta, \gamma, \delta)$
- $\alpha \in (0, 2]$ : 安定性,  $\beta \in [-1, 1]$ : 歪度,  
 $\gamma > 0$ : スケール,  
 $\delta \in \mathbb{R}$ : 位置
- If  $0 < \alpha < 2 \rightarrow$  分布の二次以上のモーメントが無限大に発散
- If  $0 < \alpha < 1 \rightarrow$  平均値も不定
- 正規分布 ( $\alpha = 2$ ), コーシー分布 ( $\alpha = 1, \beta = 0$ ),  
 レヴィ分布 ( $\alpha = 0.5, \beta = 1$ ) 等を特殊ケースとして包含
- 特性関数  $\phi(\tau) = E(\exp(i\tau X))$  による定式化のみ:

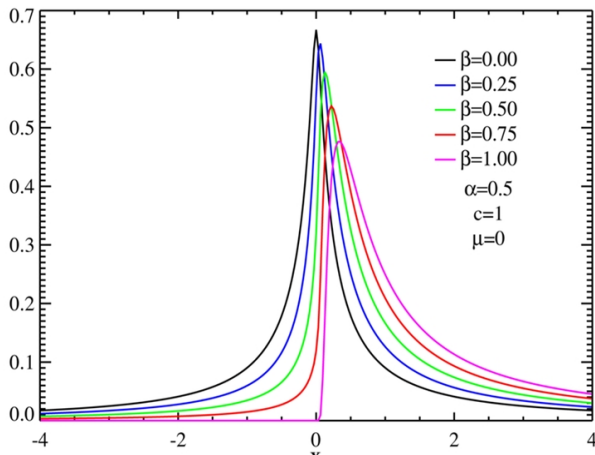
$$\phi(\tau) = \begin{cases} \exp \left\{ -\gamma^\alpha |\tau|^\alpha \left[ 1 + i\beta (\text{sign}\tau) \left( \tan \frac{\pi\alpha}{2} \right) \left( (\gamma |\tau|)^{1-\alpha} - 1 \right) \right] + i\delta\tau \right\} & (\alpha \neq 1) \\ \exp \left\{ -\gamma |\tau| \left[ 1 + i\beta (\text{sign}\tau) \tan \frac{2}{\pi} (\ln |\tau| + \ln \gamma) \right] + i\delta\tau \right\} & (\alpha = 1) \end{cases}$$

## 安定性パラメータ $\alpha$ の影響



([en.wikipedia.org/wiki/Levy\\_skew\\_alpha-stable\\_distribution](http://en.wikipedia.org/wiki/Levy_skew_alpha-stable_distribution))

## 歪度パラメータ $\beta$ の影響



([en.wikipedia.org/wiki/Lévy\\_skew\\_alpha-stable\\_distribution](http://en.wikipedia.org/wiki/Lévy_skew_alpha-stable_distribution))

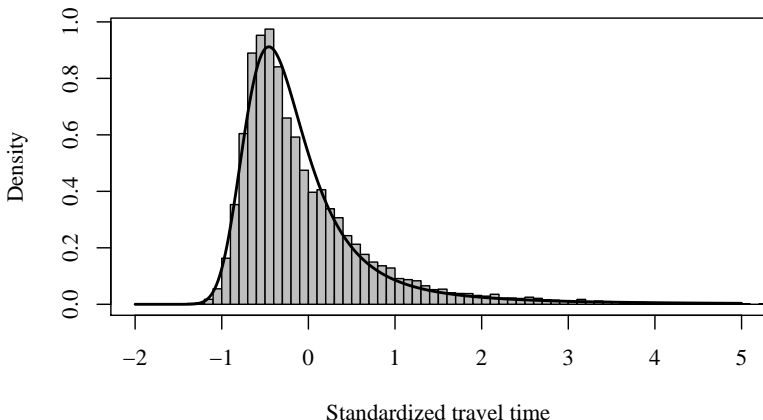
- 加法性 (Stability property)

- 安定パラメータ  $\alpha$  が共通の場合，安定分布に従う独立な確率変数の和も安定分布に従う．
- 異なるリンク間で基準化所要時間分布の独立性を想定できれば，安定分布を想定することにより，連続リンク区間 (より一般には経路区間) における基準化所要時間分布を「畳み込み (Convolution)」により求めることができる．

- 一般化中心極限定理 (Generalized central limit theorem)

- 中心極限定理 (任意の母集団分布から抽出した標本の算術平均は漸近的に正規分布に従う) の一般形：分散が無限である任意の独立な確率変数の規格和は安定分布となる．
- 所要時間分布の非対称性・裾厚性を想定したとき (non-normality)，細密な道路ネットワークにおいて，各リンクの基準化所要時間分布の和を長期間に渡って観測・獲得したら，その極限分布は安定分布になる．
- ネットワーク全体での VTTV の推計の可能性？





- 安定分布  $X \sim S_t(\alpha, \beta, \gamma, \delta)$  のパラメータ推定値：  
 $\hat{\alpha} = 1.3$ ,  $\hat{\beta} = 1.0$ ,  $\hat{\gamma} = 0.3049$ ,  $\hat{\delta} = -0.3785$

互いに独立な  $J$  個の安定分布に従う確率変数

$X^j \sim S_t(\alpha, \beta^j, \gamma^j, \delta^j)$ ,  $j = 1, \dots, J$  に対し,

平均  $\bar{X} = 1/J \sum_{j=1}^J X_j$  も安定分布  $\bar{X} \sim S_t(\alpha, \bar{\beta}, \bar{\gamma}, \bar{\delta})$  に従う.

ここで,

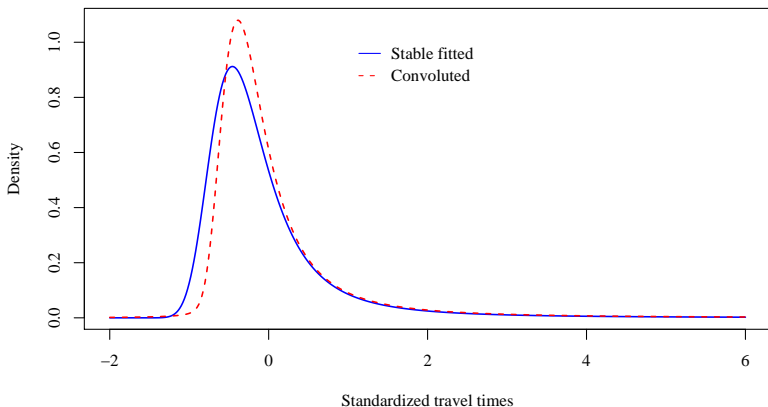
$$\bar{\beta} = \frac{\sum_{j=1}^J \beta^j |\gamma^j/J|^\alpha}{\sum_{j=1}^J |\gamma^j/J|^\alpha},$$

$$\bar{\gamma} = \left( \sum_{j=1}^J |\gamma^j/J|^\alpha \right)^{1/\alpha},$$

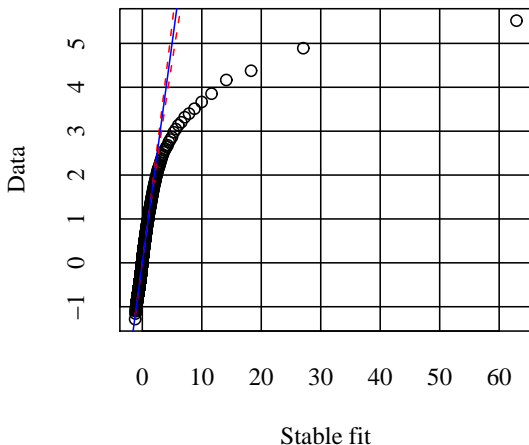
$$\bar{\delta} = \begin{cases} \sum_{j=1}^J \delta^j/J + (\tan \frac{\pi\alpha}{2}) \left[ \bar{\beta}\bar{\gamma} - \sum_{j=1}^J \beta^j \gamma^j/J \right] & (\alpha \neq 1) \\ \sum_{j=1}^J \delta^j/J + \frac{2}{\pi} \left[ \bar{\beta}\bar{\gamma} \log \bar{\gamma} - \sum_{j=1}^J \beta^j \gamma^j/J \log |\gamma^j/J| \right] & (\alpha = 1) \end{cases}$$

基準化所要時間分布：リンクレベル → 経路レベル で集計

- 四つのサブ区間における各  $\alpha$  の推定値は概ね同一の値
- その平均値 (= 1.132) を用いて畳み込みを実施



## Quantile-Quantile プロット

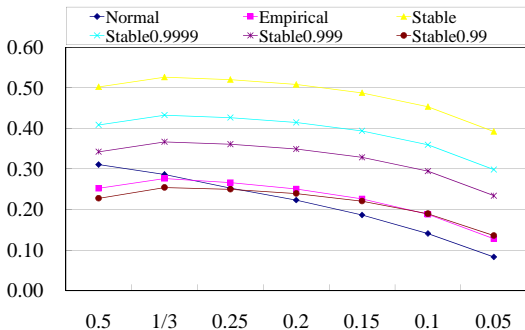


分布の右裾部分において、データとモデルの乖離が大きい  
→VTTV(or H) の推計にも影響

# Xの分布として安定分布を仮定することのデメリット

- 分散が“無限大”となることの好ましくない影響：
  - Hの推計値に大きな乖離 VTTVの推計にも影響
  - 旅行時間の(これ以上はあり得ないであろうという)“最大値”  $p_{X_{\max}}$  を与えることで, 乖離を是正することが可能:

$$H'(\Phi, \frac{\eta}{\lambda}, p_{X_{\max}}) = \int_{p_{X_{\max}} - \frac{\eta}{\lambda}}^{p_{X_{\max}}} \Phi^{-1}(v) dv.$$



異なる  $\eta/\lambda$ (横軸) に対する H の計算値 (縦軸)

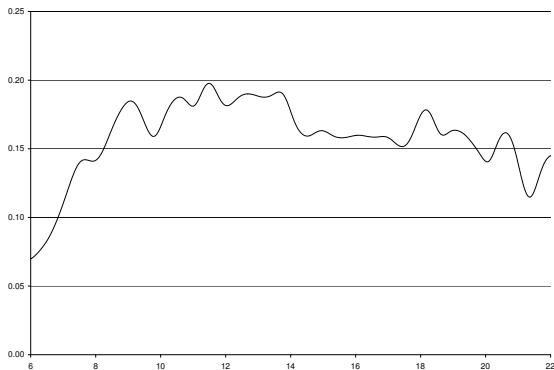


Figure: 所要時間の不確実性に起因するコストシェアの時間推移

- Small (1982) の推定結果 [  $(\eta, \lambda, \omega) = (1, 1, 5)$  ], 及び, ノンパラメトリック密度関数より数値計算される  $H$  の値  $\approx 0.29$  を用いて算出
- 昼間のほとんどの時刻帯において, 不確実性のコストは全体コストの20パーセント以上を占めている

## モデルの要約：

$$EC^* = (\eta + \omega)\mu + \lambda H \left( \Phi, \frac{\eta}{\lambda} \right) \sigma,$$

ドライバーのスケジュール選好パラメータ  $(\lambda, \eta)$  と、基準化所要時間分布の形状  $(\Phi)$  を与えることで、 $VTTV$  の推計が可能に。

## 本研究：上記モデルの実用化に向けた諸検討

- ① 基準化所要時間の時刻独立性の仮定  
→ 概ね妥当
- ② 安定分布へのフィッティング  
→ 分布の中心付近には良好，しかし，分布の右裾の部分における当てはまりは必ずしも良くない
- ③ 安定パラメータの等価性の仮定 → 概ね妥当  
→ 異なる道路セクション間での基準化所要時間分布の畳込経路レベルでの基準化所要時間分布の導出可能性

- ① 他データを用いた検証 (安定分布は Stylized Fact となるか?)
- ② セクション間の基準化所要時間の独立性の確認 (e.g. Conditional dependence test)
- ③ 経路選択行動の考慮 (Stochastic time-dependent networks)
- ④ スケジューリング・パラメータの推定 (SP 調査)
- ⑤ 時間的均衡概念 (e.g. Arnott–de Palma–Lindsey モデル) の導入
- ⑥ サービス間隔の離散性の考慮 (鉄道, バス, 航空等の公共サービス評価)



## (2) デンマーク・スイス方面の研究レビュー (2007-8)

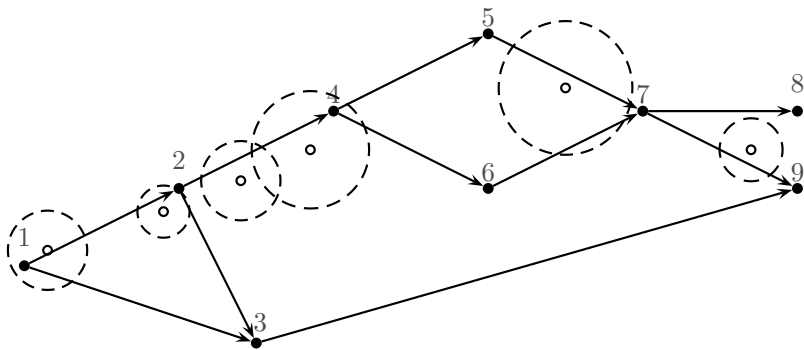
- ① 経路選択：Network-free データと重複経路の相関
- ② サンプリング：GEV モデルに対する内生的標本抽出
- ③ モデルの特定化：確定項の特定化，MXL 分布の特定化
- ④ 意思決定のアノマリー：参照点依存型モデル
- ⑤ 誤差項スケーリング：乗数型ランダム効用モデル

# [1] 経路選択： Network-free データと重複経路の相関

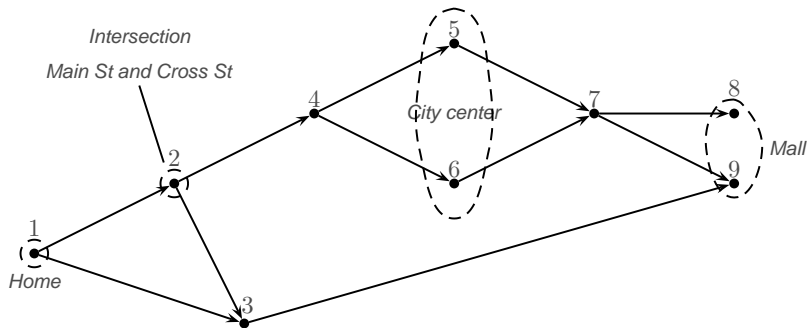
- Bierlaire, M., Frejinger, E., 2008. Route choice modeling with network-free data. *Transportation Research Part C: Emerging Technologies* 16 (2), 187–198.
- Frejinger, E., Bierlaire, M., 2007. Capturing correlation with subnetworks in route choice models. *Transportation Research Part B: Methodological* 41 (3), 363–378.

- 経路選択モデルにおいては，選択された経路の“Link-by-link”なデータが必要となるが，通常，直接それが利用できることはほとんど無い．
- ネットワークに適合した経路を得るためのデータ処理
  - GPS で得られたデータの Map matching
  - アンケート調査で報告された経路の再構成
  - 検証は難しく，バイアスや誤差をもたらす可能性大．
- $i$  を個人のトリップに伴って得られるデータ列とする．例：GPS 観測点の系列，経路の所々で報告されたランドマークポイント
- 各データ点に対して “Domain of Data Relevance (DDR)” を定義．例：GPS の観測範囲，ランドマークポイント
- DDRs が network-free data と network model の橋渡しをする

## GPS データによる DDR



## アンケート調査データによる DDR

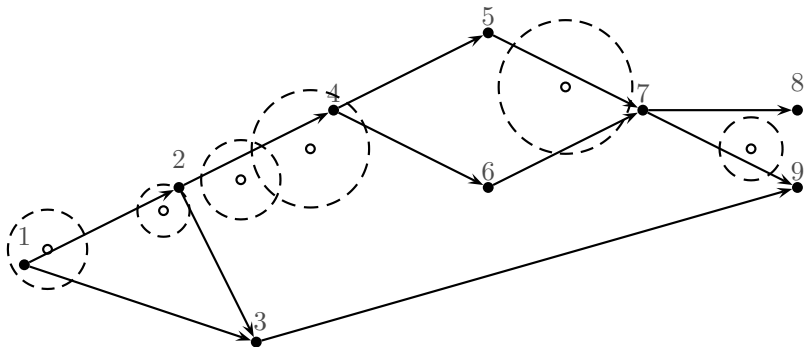


- 最終的には、経路選択モデル  $P(p|\mathcal{C}_n(s), \beta)$  の未知パラメータ  $\beta$  を推定したい ( $\mathcal{C}_n(s)$ : OD ペア  $s$  を利用する個人  $n$  の経路選択肢集合)
- 個人  $n$  の観測データ  $i$  に関連する OD ペアは複数あり得る (集合  $S_i$ )
- $S_i$  が与えられたとき、個人  $n$  の観測データ  $i$  が再現される確率は、

$$P_n(i|S_i) = \sum_{s \in S_i} P(s|S_i) \sum_{p \in \mathcal{C}_n(s)} P_n(i|p) P_n(p|\mathcal{C}_n(s), \beta)$$

- $P_n(i|p)$ : 観測方程式。アンケート調査の場合、 $i$  がパス  $p$  に合致するなら  $P_n(i|p) = 1$ 、それ以外は 0。GPS データの場合、確率は  $i$  と  $p$  の (最短) 距離の関数となる。

## GPS データによる DDR



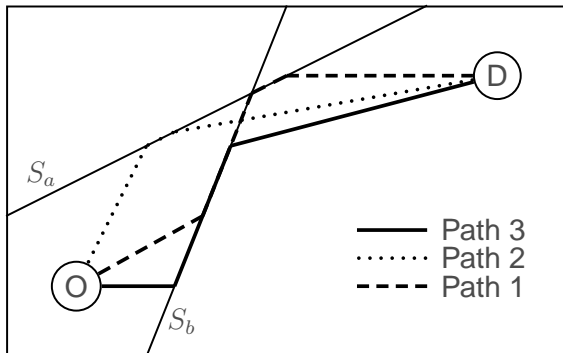
$$P_n(i|S_i) = \sum_{s \in S_i} P(s|S_i) \sum_{p \in C_n(s)} P_n(i|p) P_n(p|C_n(s), \beta)$$

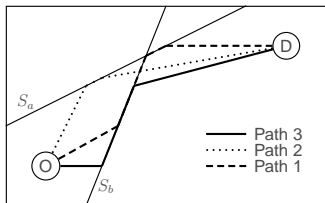
$$P(i|s) = P(i|p_1)P(p_1|C(s), \beta) + P(i|p_2)P(p_2|C(s), \beta)$$

- モデルの複雑さを増すことなく，いかにして，ネットワークにおける重要な相関構造を考慮することができるか．
- 経路選択行動において，行動論的に (認知的に?) 重要な意思決定は何か？
- ネットワークにおける「特定部分」に着目 (e.g. 主要幹線道路, 市中心部)
- “Subnetwork” という概念：ラベリングが容易に可能なネットワークを構成する部分リンク
- Subnetwork をシェアしている経路同士は，例え，互いに物理的に重複していなくとも，相関していると仮定



## Subnetwork の例





Error Component Logit (MXL):

$$\mathbf{U}_1 = \beta^T \mathbf{X}_1 + \sqrt{l_{1a}} \sigma_a \zeta_a + \sqrt{l_{1b}} \sigma_b \zeta_b + \epsilon_1$$

$$\mathbf{U}_2 = \beta^T \mathbf{X}_2 + \sqrt{l_{2a}} \sigma_a \zeta_a + \epsilon_2$$

$$\mathbf{U}_3 = \beta^T \mathbf{X}_3 + \sqrt{l_{3a}} \sigma_a \zeta_a + \epsilon_3$$

In general,

$$\mathbf{U}_n = \beta^T \mathbf{X}_n + \mathbf{F}_n \mathbf{T} \zeta_n + \epsilon_n$$

- J: 経路数 (= 3)
- Q: Subnetwork 数 (= 2)

Variance-covariance matrix of  $\zeta$ :

$$\mathbf{F} \mathbf{T} \mathbf{T}^T \mathbf{F}^T = \begin{pmatrix} l_{1a} \sigma_a^2 + l_{1b} \sigma_b^2 & \sqrt{l_{1a}} \sqrt{l_{2a}} \sigma_a^2 & \sqrt{l_{1b}} \sqrt{l_{3b}} \sigma_b^2 \\ \sqrt{l_{1a}} \sqrt{l_{2a}} \sigma_a^2 & l_{2a} \sigma_a^2 & 0 \\ \sqrt{l_{3b}} \sqrt{l_{1b}} \sigma_b^2 & 0 & l_{3b} \sigma_b^2 \end{pmatrix}$$

## [2] サンプリング： GEV モデルに対する内生的標本抽出

- Bierlaire, M., Bolduc, D., McFadden, D., 2008. The estimation of generalized extreme value models from choice-based samples. *Transportation Research Part B: Methodological* 42 (4), 381–394.

Choice-based sampling で得られたデータに対し，統計的一致性を持つような GEV モデルパラメータ推定量を提案：

Notation:

- $R(i, x, \mathcal{B}, \theta) = r_{g(i,x)}(\theta)\pi(\mathcal{B}|i, x)$
- $R(i, x, \mathcal{B}, \theta)$ : 属性  $(i, x)$  を持つ母集団構成員が標本抽出され，かつ，部分選択肢集合  $\mathcal{B} \subseteq \mathcal{C}$  が割り当てられる確率
- $r_{g(i,x)}(\theta)$ : 属性  $(i, x)$  を持つ母集団構成員が標本抽出される確率
- $\theta$ : 行動モデルのパラメータ
- $\pi(\mathcal{B}|i, x)$ : 属性  $(i, x)$  を持つ個人に対して，部分選択肢集合  $\mathcal{B}$  が割り当てられる確率
- $R$  が次のように異なる形の積で表されると仮定：  

$$R(i, x, \mathcal{B}, \theta) = Q(i, x, \mathcal{B})S(i, x, \mathcal{B}, \theta)$$

- CML (Constrained Maximum Likelihood) Estimator:

$$\mathcal{L}_C = \sum_{n=1}^N \ln \frac{S(\mathbf{i}_n, \mathbf{x}_n, \mathcal{B}_n, \theta) \Pr(\mathbf{i}_n | \mathbf{x}_n; \theta)}{\sum_{j \in \mathcal{B}_n} S(j, \mathbf{x}_n, \mathcal{B}_n, \theta) \Pr(j | \mathbf{x}_n; \theta)}.$$

- WESML (Weighted Exogenous Sample Maximum Likelihood) Estimator:

$$\mathcal{L}_W = \sum_{n=1}^N Q(\mathbf{i}_n, \mathbf{x}_n, \mathcal{B}_n)^{-1} \ln \Pr(\mathbf{i}_n | \mathbf{x}_n; \theta).$$

- WCML (Weighted CML) Estimator or  
BBM (Bierlaire, Bolduc, McFadden) Estimator:

$$\mathcal{L}_{BBM} = \sum_{n=1}^N Q(\mathbf{i}_n, \mathbf{x}_n, \mathcal{B}_n)^{-1} \ln \frac{S(\mathbf{i}_n, \mathbf{x}_n, \mathcal{B}_n, \theta) \Pr(\mathbf{i}_n | \mathbf{x}_n; \theta)}{\sum_{j \in \mathcal{B}_n} S(j, \mathbf{x}_n, \mathcal{B}_n, \theta) \Pr(j | \mathbf{x}_n; \theta)}.$$

( $Q = 1 \rightarrow$  CML and  $S = 1 \rightarrow$  WESML.)

- GEV model:

$$F_{\epsilon} = F_{\epsilon_1, \dots, \epsilon_J}(\mathbf{V}_1, \dots, \mathbf{V}_J) = e^{-G(e^{-V_1}, \dots, e^{-V_J}; \gamma)}$$

- $G_i(x, \beta, \gamma) \equiv \frac{\partial G}{\partial e^{V_i(x, \beta)}}(e^{V_1}, \dots, e^{-V_J}; \gamma)$  とすると,

$$\frac{S(i_n, \mathbf{x}_n, \mathcal{B}_n, \theta) \Pr(i_n | \mathbf{x}_n; \theta)}{\sum_{j \in \mathcal{B}_n} S(j, \mathbf{x}_n, \mathcal{B}_n, \theta) \Pr(j | \mathbf{x}_n; \theta)} \\ = \frac{e^{V_i(x, \beta) + \ln G_i(x, \beta, \gamma) + \ln S(i, x_n, \mathcal{B}_n, \theta)}}{\sum_{j \in \mathcal{B}_n} e^{V_j(x, \beta) + \ln G_j(x, \beta, \gamma) + \ln S(j, x_n, \mathcal{B}_n, \theta)}}$$

- MNL 以外の GEV-Model に対しては，層別定数項補正は適用できない．すなわち，(W)ESML を non-MNL GEV model に適用することはできない．
- $\mathcal{C} = \mathcal{C}_1 \cup \dots \cup \mathcal{C}_M$  という互いに排反な部分選択肢集合が定義でき，GEV 関数が Brock-Additive 型の場合には，ESML が一致推定量となることが分かっている．Nested Logit は × ．

- $S(j, x, \mathcal{B}, \theta) = S(j, \mathcal{B}, \theta)$  とすると, 選択肢別標本抽出 (CBS) に対する GEV model の一致推定量は下記で与えられる:

$$\max_{\beta, \gamma, \omega} \sum_{n=1}^N Q(i_n, x_n, \mathcal{B}_n)^{-1} \ln \frac{e^{V_{i_n}(x_n, \beta) + \ln G_{i_n}(x_n, \beta, \gamma) + \omega_{i_n}}}{\sum_{j \in \mathcal{B}_n} e^{V_j(x_n, \beta) + \ln G_j(x_n, \beta, \gamma) + \omega_j}}.$$

### [3] モデルの特定化： 確定項の特定化，MXL分布の特定化

- Fosgerau, M., 2008. Specification testing of discrete choice models: a note on the use of a nonparametric test. *Journal of Choice Modelling* 1 (1), 26–39.
- Fosgerau, M., Bierlaire, M., 2007b. A practical test for the choice of mixing distribution in discrete choice models. *Transportation Research Part B: Methodological* 41 (7), 784–794.



- ノンパラメトリック法に基づいた離散選択モデルの特定化の検定
- 条件付き期待値  $E(\mathbf{y}|\mathbf{x})$ ,  $\mathbf{y} \in \mathbb{R}$ ,  $\mathbf{x} \in \mathbb{R}^m$
- $\mathbf{y}$  はバイナリー変数でもよい。また、多選択肢でもよい。
- $E(\mathbf{y}|\mathbf{x})$  は  $\mathbf{x}$  の関数  $\Rightarrow$  真の関数を  $g(\mathbf{x})$  とする。
- 分析者が特定化するのはパラメトリックなモデル：  
 $f(\mathbf{x} : \theta)$ ,  $\theta \in \Theta$
- 我々は、 $f(\mathbf{x} : \theta_0) = g(\mathbf{x})$  となるような  $\theta_0 \in \Theta$  が存在するかどうかを検証したい。
- 帰無仮説  $H_0 : \exists \theta_0 \in \Theta \Pr[f(\mathbf{x} : \theta_0) = g(\mathbf{x})] = 1$ .
- 対立仮説  $H_1 : \forall \theta \in \Theta \Pr[f(\mathbf{x} : \theta) = g(\mathbf{x})] < 1$ .
- モデルの (疑似) 誤差  $\epsilon = \mathbf{y} - f(\mathbf{x} : \theta_0)$  に着目。

- パラメータの一致推定量  $\hat{\theta}$  を用いると，残差は，  

$$e_i = y_i - f(x_i; \hat{\theta})$$
- Zheng 統計量:

$$T_n = \frac{\sum_{i \leq n} \sum_{j \leq n, j \neq i} \mathbf{K} \left( \frac{x_i - x_j}{h} \right) e_i e_j}{\left[ \sum_{i \leq n} \sum_{j \leq n, j \neq i} 2\mathbf{K}^2 \left( \frac{x_i - x_j}{h} \right) e_i^2 e_j^2 \right]^{\frac{1}{2}}}$$

- $\mathbf{K}$ : カーネル関数 (e.g.  $\mathbf{K}(v) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}v^2\right)$ )
- $h$ : バンド幅 (e.g. プラグインルール :=  $cn^{\frac{1}{2m}}$ ,  $c$ : 定数)
- If  $h \rightarrow 0$  and  $nh^m \rightarrow \infty$ , then under the null hypothesis,  
 $T_n \rightarrow_d \mathbf{N}(0, 1)$

## Zheng–Fosgerau specification test

Income が説明変数として含まれるべきかどうかの Omitted Variable Test:

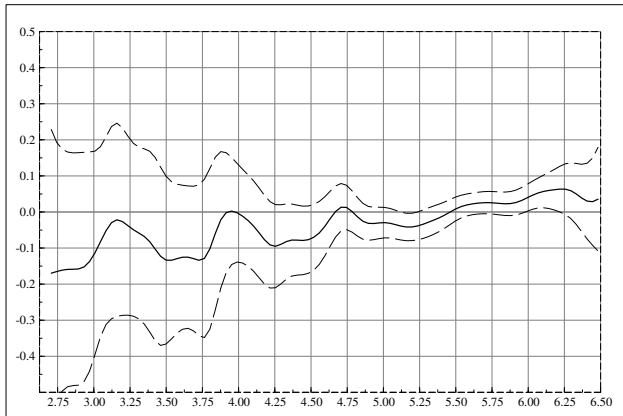


Figure 2: Residuals of air alternative on real income

現在の効用関数特定化が，総じて妥当かどうかの  
Specification Test:

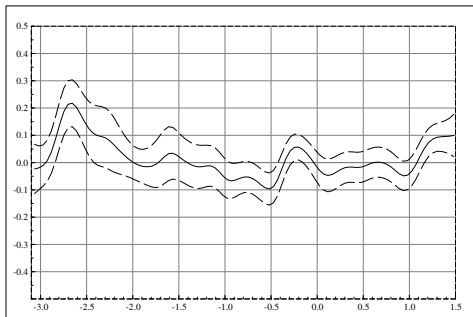


Figure 4: Residuals of train alternative on systematic utility for bus

Table 1: Zheng statistics for tests on systematic utilities  $V$

	Air	Bus	Car	Train
$V(\text{Air})$	0.2	0.8	-0.2	0.0
$V(\text{Bus})$	3.2	3.7	1.9	6.0
$V(\text{Car})$	5.5	9.7	2.7	2.4
$V(\text{Train})$	5.0	3.4	1.12	2.4

- MXL などにおいて，不適切な分布形を与えれば，時間価値などの支払い意思額指標の推計結果にゆがんだ結果をもたらす．
- Fosgerau–Bierlaire test: Semi-nonparametric (確率密度関数などを近似するのに，級数近似を用いるノンパラメトリック統計解析技法の総称) を応用．
- 離散選択モデルにおけるランダムパラメータ  $\omega$  が分布関数  $F$ ，密度  $f$  を持つ分布に本当に従っているのかをテストする．
- 真の分布関数を  $G$ ，密度を  $g$  とすると，

$$G(\omega) = Q(F(\omega))$$

- $Q$ :  $[0, 1]$  から  $[0, 1]$  への単調変換写像 (関数) ．単位区間で定義された確率変数に対する分布関数と考えることができる．
- 偏微分して， $g(\omega) = q(F(\omega))f(\omega)$  を得る．

- 今, 関数  $q$  は次のように近似される:  $q(x) \approx \frac{1}{K} q_N^2(x)$
- $q_N(x) = 1 + \sum_{k=1}^N \delta_k L_k(x)$ :  $N$  次の Legendre 多項式近似
- $K = \int_{-\infty}^{+\infty} q_N^2(F(\omega)) f(\omega) d\omega$ : 正規化定数
- 直交性の性質より,  $K = 1 + \sum_{k=1}^N \delta_k^2$
- すると, MXL などのような Mixing タイプの離散選択モデルは, 変数変換  $z = F(\beta)$  により以下のようになる:

$$\begin{aligned} P_n(i|C_n) &= \int_{-\infty}^{+\infty} P_n(i|\beta, C_n) g(\beta) d\beta \\ &\approx \frac{1}{K} \int_{-\infty}^{+\infty} P_n(i|\beta, C_n) q_N^2(F(\beta)) f(\beta) d\beta \\ &= \frac{1}{K} \int_0^1 P_n(i|F^{-1}(z), C_n) q_N^2(z) dz \end{aligned}$$

- $F^{-1}(z)$  は元の分布に従う確率変数であり, モンテカルロシミュレーションで積分計算が可能.

- 帰無仮説  $f = g$  の元では  $q = 1$  , すなわち ,  $\delta_k = 0, \forall k$  .  
このとき , モデル

$$P_n(i|C_n) = \int_{-\infty}^{+\infty} P_n(i|\beta, C_n)g(\beta)d\beta$$

は , 下記のモデルに等価 ,

$$P_n(i|C_n) = \int_{-\infty}^{+\infty} P_n(i|\beta, C_n)f(\beta)d\beta$$

- 後者のモデルは前者の特別な場合であり (nested relation) ,  
尤度比検定等を用いて検定が可能 .
- (補足) Legendre 多項式の定義

$$L_n(x) = \frac{\sqrt{4n^2-1}}{n}(2x-1)L_{n-1}(x) - \frac{(n-1)\sqrt{2n+1}}{n\sqrt{2n-3}}(2x-1)L_{n-2}(x), \quad L_0(x)=1.$$
$$\int_0^1 L_m(x)L_n(x)dx = 0(\text{if } m \neq n), \quad 1(\text{if } m=n)$$

経験的に , 三次位までとることで , 全体の十分な近似が可能 .

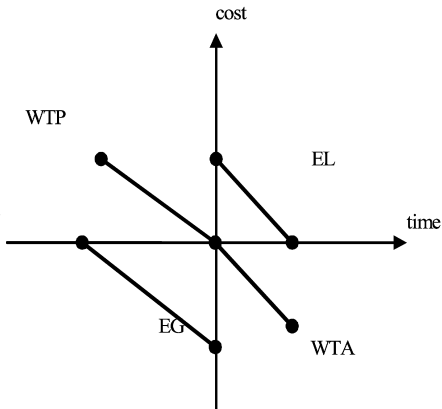
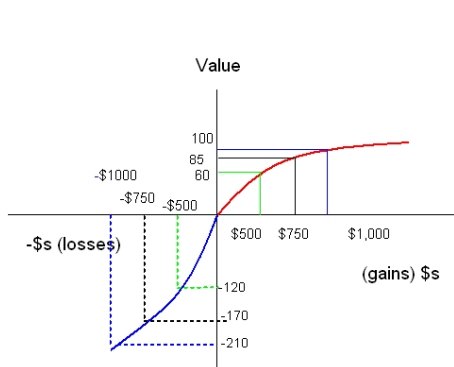
## [4] 意思決定のアノマリー： 参照点依存型モデル

- de Borger, B., Fosgerau, M., 2008. The trade-off between money and travel time: A test of the theory of reference-dependent preferences. *Journal of Urban Economics* 64 (1), 101–115.
- Fosgerau, M., de Borger, B., 2008. Hedonic preferences, symmetric loss aversion and the willingness to pay-willingness to accept gap. Tech. rep., Technical University of Denmark.
- Fosgerau, M., Hjort, K., Lyk-Jensen, S., 2008. An integrated approach to the estimation of the value of travel time. Tech. rep., Technical University of Denmark.



- 人は、現在おかれた状況を基に選択を行う  
(Reference-Dependent)
- 人は、同程度の利得よりも損失の方をより深刻に受け止める  
(Loss-Aversion)
- SP 等により時間価値を推計する場合など、  
WTP(支払いたいと思う額) - WTA(受け取りたいと思う額)  
の間に大きなギャップ
- 以上を不確実性下の状況でモデル化した代表例が  
「プロスペクト理論 (Kahneman & Tversky, 1979)」
- リスクレス・多属性の状況に拡張したのが  
“Reference-Dependent Utility (Tversky & Kahneman, 1991)”
- de Borger & Fosgerau (2008) では、上記を基に、  
Reference-Free な VTT(Value of travel time) を推計

# 参照点依存型効用 (価値関数) と Choice Experiment



EG: Equivalent Gain , EL: Equivalent Loss  
 価値関数のもとでは ,

$$WTP(t) < \min[EG(t), EL(t)] \leq \max[EG(t), EL(t)] < WTA(t)$$

## 価値関数 (Value Function)

- Power Function が経験的に最も適合度が高い :

$$v(x) = S(x) e^{-\eta S(x)} |x|^{1-\beta-\gamma S(x)}$$

- Where  $S(x) = x/|x|$  ,  $\gamma$ : 損失性回避を表すパラメータ
- $\eta = \gamma = \beta = 0$  のとき , 時間と費用の限界代替率の比 (i.e. 先程の 4 つ) は , Reference-Free VTT ( $w$ ) に一致する .
- 重要な関係式 :

$$\ln[WTP(t)WTA(t)] = \ln[EL(t)EG(t)] = 2 \frac{1 - \beta_t}{1 - \beta_c} \ln[|wt|]$$

時間の価値関数と費用の価値関数が “対称 (i.e.  $\beta_t = \beta_c$ )” ならば ,  $w$  は  $WTP(t)$  と  $WTA(t)$  の幾何平均に等しい .

- Hedonic Preference (Reference-Free 時間価値) と Choice Preference (  $WTP$  等 ) との関連性を見出す為の一試案

## Econometric Specification:

- 二項選択状況においてある一方の選択肢が選ばれる条件：

$$-S(c_1)e^{S(c_1)\eta_c|c_1|^{1-\beta_c+\gamma_c S(c_1)}} - S(t_1)e^{S(t_1)\eta_t|wt_1|^{1-\beta_t+\gamma_t S(t_1)}} \\ < -S(c_2)e^{S(c_2)\eta_c|c_2|^{1-\beta_c+\gamma_c S(c_2)}} - S(t_2)e^{S(t_2)\eta_t|wt_2|^{1-\beta_t+\gamma_t S(t_2)}}$$

- 個人毎の時間価値を仮定： $\ln w = \delta_0 + \delta z + \sigma u$
- 対数を取った後，ランダム項を加えると：

$$[1 - \beta_t + \gamma_t S(t_i)][\delta_0 + \delta z + \sigma u] < \\ S(c_i)\eta_c - S(t_i)\eta_t + [1 - \beta_c + \gamma_c S(c_i)] \ln |c_i| - \\ [1 - \beta_t + \gamma_t S(t_i)] \ln |t_i| + \mu \varepsilon_i \\ \text{(Binary MXL)}$$

- DATIV (Danish Time-Value Survey) への適用：  
WTP/WTA ギャップは約 4 倍

## [5] 誤差項スケーリング： 乗数型ランダム効用モデル

- Fosgerau, M., Bierlaire, M., 2007a. Circumventing the problem of the scale: discrete choice models with multiplicative error terms. Tech. Rep. TRANSP-OR 070703, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne.

先程の式で，誤差項を与える前の形 (価値関数表現) は：

$$\begin{aligned}
 & -S(c_1)e^{S(c_1)\eta_c|c_1|^{1-\beta_c+\gamma_c}S(c_1)} - S(t_1)e^{S(t_1)\eta_t|wt_1|^{1-\beta_t+\gamma_t}S(t_1)} \\
 & \leftarrow -S(c_2)e^{S(c_2)\eta_c|c_2|^{1-\beta_c+\gamma_c}S(c_2)} - S(t_2)e^{S(t_2)\eta_t|wt_2|^{1-\beta_t+\gamma_t}S(t_2)}
 \end{aligned}$$

価値関数の対数を取ってからランダム項を加えている．つまり簡単に書くと，もとの確定効用  $V_i$  とランダム項  $\varepsilon_i$  との関係は，

$$U(V_j, \varepsilon_j) = \mu V_j \varepsilon_j \text{ (Multiplicative error terms)}$$

例えば，通常に加算型ランダム効用モデルの場合，10分の時間差は移動距離の長短にかかわらず一定の効果を与える (ゆえに，セグメンテーション等が必要)．しかし，上記の定式化だとそれを明示的に考慮することができる．

$V_j < 0, \varepsilon_j > 0$  であるとする。対数変換は単調変換であるので、  
 選択確率は：

$$\begin{aligned} P(i|\mathcal{C}) &= \Pr(\mu V_i \varepsilon_i \geq \mu V_j \varepsilon_j, \quad \forall j) \\ &= \Pr(-V_i \varepsilon_i \leq -V_j \varepsilon_j, \quad \forall j) \\ &= \Pr(\ln(-V_i) + \ln(\varepsilon_i) \leq \ln(-V_j) + \ln(\varepsilon_j), \quad \forall j) \\ &= \Pr(-\ln(-V_i) - \ln(\varepsilon_i) \geq -\ln(-V_j) - \ln(\varepsilon_j), \quad \forall j) \end{aligned}$$

$-\ln(\varepsilon_j) = \xi_j/\lambda, \bar{V}_i = -\lambda \ln(-V_i)$  のように定義すると ( $\lambda > 0$ ) ,

$$\begin{aligned} P(i|\mathcal{C}) &= \Pr(\bar{V}_i + \xi_i \geq \bar{V}_j + \xi_j, \quad j \in \mathcal{C}) \\ &= \Pr(-\lambda \ln(-V_i) + \xi_i \geq -\lambda \ln(-V_j) + \xi_j, \quad j \in \mathcal{C}) \end{aligned}$$

- スケール  $\lambda$  は母集団全体で一定であることが確約されている  
 (加算型モデルとの違い)
- $\xi_j$  の分布をガンベル等で仮定して、選択確率を導出

# VTTV References

- Bates, J., Polak, J., Jones, P., Cook, A., 2001. The valuation of reliability for personal travel. *Transportation Research Part E: Logistics and Transportation Review* 37 (2–3), 191–229.
- Brownstone, D., Small, K., 2005. Valuing time and reliability: assessing the evidence from road pricing demonstrations. *Transportation Research Part A: Policy and Practice* 39 (4), 279–293.
- Doornik, J. A., 2001. *Ox: An Object-Oriented Matrix Language*. London.
- Fan, J., Yao, Q., 1998. Efficient estimation of conditional variance functions in stochastic regression. *Biometrika* 85 (3), 645–660.
- Fofack, H., Nolan, J. P., 1999. Tail behavior, modes and other characteristics of stable distributions. *Extremes* 2 (1), 39–58.
- Fosgerau, M., Karlström, A., 2007. The value of reliability and the distribution of random durations. In: *European Transport Conference*.
- Gnedenko, B. V., Kolmogorov, A. N., 1954. *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley.
- Hall, P., Racine, J., Li, Q., 2004. Cross-validation and the estimation of conditional probability densities. *Journal of the American Statistical Association* 99 (468), 1015–1027.
- Härdle, W., 1990. *Applied Nonparametric Regression*. Econometric Society Monograph Series. Cambridge University Press.
- Lam, T. C., Small, K., 2001. The value of time and reliability: measurement from a value pricing experiment. *Transportation Research Part E: Logistics and Transportation Review* 37 (2-3), 231–251.
- Li, Q., Racine, J., 2007. *Nonparametric Econometrics: Theory and Practice*. Princeton University Press.
- Michael, J. R., 1983. The stabilized probability plot. *Biometrika* 70 (1), 11–17.
- Nolan, J. P., 1997. Numerical calculation of stable densities and distribution functions. *Stochastic Models* 13 (4), 759–774.



- Nolan, J. P., 2001. Maximum likelihood estimation of stable parameters. In: Barndorff-Nielsen, O. E., Mikosch, T., Resnick, S. I. (Eds.), *Lévy Process: Theory and Applications*. Birkhäuser, Boston, pp. 379–400.
- Nolan, J. P., in press. *Stable Distributions – Models for Heavy Tailed Data*. Birkhäuser, Boston.
- Noland, R. B., 1997. Commuter responses to travel time uncertainty under congested conditions: expected costs and the provision of information. *Journal of Urban Economics* 41 (3), 377–406.
- Noland, R. B., Polak, J. W., 2002. Travel time variability: a review of theoretical and empirical issues. *Transport Reviews* 22 (1), 39–54.
- Noland, R. B., Small, K. A., 1995. Travel-time uncertainty, departure time choice, and the cost of morning commutes. *Transportation Research Record* 1493, 150–158.
- Noland, R. B., Small, K. A., Koskenoja, P. M., Chu, X., 1998. Simulating travel reliability. *Regional Science and Urban Economics* 28 (5), 535–564.
- Pagan, A., Ullah, A., 1999. *Nonparametric Econometrics*. Cambridge University Press.
- R Development Core Team, 2007. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Robust Analysis, Inc., 2006. *User Manual for STABLE 4.0 S-Plus/R Version*.
- Scott, D. W., 1979. On optimal and data-based histograms. *Biometrika* 66 (3), 605–610.
- Small, K. A., 1982. The scheduling of consumer activities: work trips. *American Economic Review* 72 (3), 467–479.
- Small, K. A., Winston, C., Yan, J., 2005. Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica* 73 (4), 1367–1382.
- Su, L., White, H., 2007. A consistent characteristic function-based test for conditional independence. *Journal of Econometrics* 141 (1), 807–834.
- Zolotarev, V., 1986. *One-Dimensional Stable Distributions*. American Mathematical Society.