行動モデルと双対概念 Ver.2

(Behavioral Model and Duality Theory)

東京工業大学 環境・社会理工学院
福田 大輔
fukuda@plan.cv.titech.ac.jp

発表の構成

- 数理最適化/ミクロ経済学における双対概念
- ロジットモデルと双対性
- ランダム効用理論と摂動効用理論
- 一般化エントロピーモデル

[付録]

- 合理的無反応 (Rational Inattention) モデル
- ログサム変数の導出

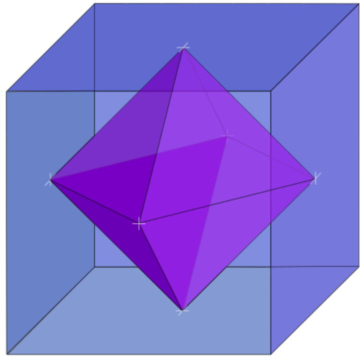
双対性(Duality)とは

- 互いに対になっている2つの対象の間の関係
  - ■図形：正六面体⇔正八面体
    正十二面体⇔正二十面体
    （双対多面体：与えられた正多面体の各面の中心に頂点を取り，それらを結んで造られる立体）
  - ■論理：論理和 ∨ ⇔ 論理積 ∧

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z),$$

$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$$

  - ■電気と磁気：e.g. 電束密度⇔磁束密度
  - ■制御工学：伝達関数表現⇔状態空間表現
  - ■電気工学：e.g. 並列⇔直列
  - ■最適化問題の双対性



https://ja.wikipedia.org/wiki/%E5%8F%8C%E5%AF%BE

最適化問題における双対概念の意義

- 主問題に有界な最適解が存在するとき，
  双対問題にも有界な最適解が存在し，
  それぞれの目的関数値は一致する．（強双対定理）

- 同一の行動を違った観点から説明できる．
- 主問題を解くよりも，双対問題を解き，
  双対定理を用いて解を導いた方が簡単な場合がある．
- 効率的な計算アルゴリズムの開発に応用できる．
  （e.g. 利用者均衡配分の効率的解法）
- 数値解の精度保証の検証に使うことができる．

ロジットモデルと双対性(1)

Miyagi (1986)，宮城・小川(1986)：
Logit Model が Entropy Modelを定義する際の数理最適化問題と同型の問題から導出できることを示し，標準的消費者行動理論と整合する為の理論的基礎を与えている

（maxの内側：直接効用関数）

$$S(\mathbf{V},\ \theta)=\max_{\mathbf{P}}\sum_{j\in\mathcal{J}}\left\{P_jV_j-\frac{1}{\theta}P_j\left(\ln P_j-1\right)\right\}$$

$$\text{s.t.}\quad \sum_{j\in\mathcal{J}}P_j=1$$

（消費総量制約）

最適解

（需要関数）

$$P_j=\frac{\exp(\theta V_j)}{\sum_{j'\in\mathcal{J}}\exp(\theta V_{j'})},\quad \forall j\in\mathcal{J}$$

（間接効用関数）

$$S(\mathbf{V},\ \theta)=\frac{1}{\theta}\ln\left\{\sum_{j'\in\mathcal{J}}\exp(\theta V_{j'})\right\}$$

$\mathbf{P}$：需要(シェア)ベクトル，$\boldsymbol{V}$：確定効用ベクトル，$0, 1, \ldots, J$: 選択肢 (0: 合成財)，θ: スケール

ロジットモデルと双対性(2)

Anderson, de Palma, & Thisse (1988), Verboven (1996)：
消費する財の「バラエティ」に対する選好をEntropy関数によって表現した直接効用関数の最大化問題の解が，(Nested) Logit 型の需要関数になると共に，間接効用関数がLogsum関数の形式になることを示した.

[Logit Representative Consumer Behavior]

Shannon Entropy型直接効用関数

$$U^* = \max_{\mathbf{X}} \left(X_0 + \sum_{j=1}^{J} a_j X_j - \frac{1}{\theta} \sum_{j=1}^{J} X_j \ln \frac{X_j}{N} \right)$$

Entropy term → Variety-seeking behaviour

s.t.

（消費総量制約）

$$\sum_{j=1}^{J} X_j = N$$

（所得制約）

$$Y = \sum_{j=1}^{J} p_j X_j + p_0 X_0$$

最適解

（需要関数）

$$\tilde{X}_j = N \frac{\exp[\theta(a_j - p_j)]}{\sum_{j'=1}^{J} \exp[\theta(a_{j'} - p_{j'})]}, \quad \forall j \in \mathcal{J}$$

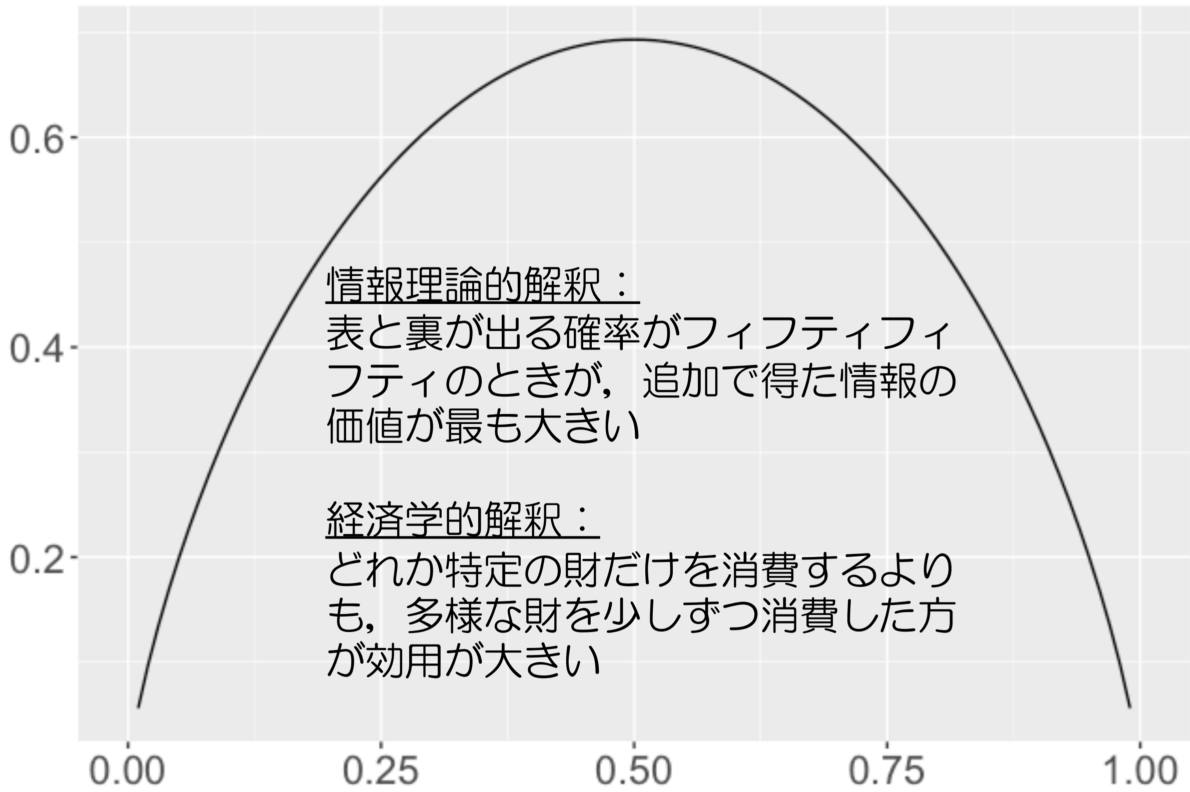
（間接効用関数）

$$U^* = \frac{Y}{p_0} + N\theta \ln \left(\sum_{j=1}^{J} \exp\left(\frac{a_j - p_j/p_0}{\theta} \right) \right)$$

$\mathbf{X}$：需要(量)ベクトル，$a_j - p_j$：根源消費効用−価格，$0, 1, \ldots, J$: 選択肢 (0: 合成財)，θ: スケール

Shannon Entropy関数

$$H(p) = -\sum p_i \ln p_i = -p \ln p - (1-p)\ln(1-p)$$



コインを投げたときに表が出る確率 p
（＝裏が出る確率 1-p）

ロジットモデルと双対性(3)

$$-G^{\bullet}(\mathbf{q}) = -\frac{1}{\theta}\sum_{j\in\mathcal{J}} q_j \ln q_j$$

$$G(\boldsymbol{\delta}) \equiv \mathbb{E}\left(\max_{j\in\mathcal{J}} u_j\right) = \frac{1}{\theta}\ln\left(\sum_{j\in\mathcal{J}} e^{\theta\delta_j}\right)$$

ルジャンドル変換 (凸共役)

$$G^{\bullet}(\mathbf{q}) = \max_{\boldsymbol{\delta}}\left\{\mathbf{q}^{\top}\boldsymbol{\delta} - G(\boldsymbol{\delta})\right\}$$

Shannon Entropy

Logsum

偏微分

偏微分

Additive Perturbed Utility (APU) representation

（ゲーム理論では「**摂動効用**」とも呼ばれる）

[e.g. Fudenberg et al. 2015]

$$\begin{aligned}\mathbf{P}(\mathcal{J}) &= (SF(q_1), \ldots, SF(q_J)) \\ &= \underset{\mathbf{q}\in\boldsymbol{\Delta}}{\arg\max} \sum_{j\in\mathcal{J}}\left\{\delta_j q_j - \frac{1}{\theta} q_j \ln q_j\right\}\end{aligned}$$

$\boldsymbol{\delta}$

Utilities

Logit (softmax) 変換

Choice Probabilities

$\mathbf{q}$

$$q_j(\boldsymbol{\delta}) = \frac{\partial G(\boldsymbol{\delta})}{\partial \delta_j} = SF(q_j) \equiv \frac{e^{\theta\delta_j}}{\sum_{j'\in\mathcal{J}} e^{\theta\delta_{j'}}},$$

$\forall j \in \mathcal{J}$

$\mathbf{q}$：需要(シェア)ベクトル，$\boldsymbol{\delta}$: 確定効用ベクトル，$0, 1, \ldots, J$: 選択肢 (0: 合成財)

ロジット～エントロピー～ランダム効用

- Miyagi, ADT, Verbovenらのアプローチ
  - 消費者の直接効用関数をShannon Entropy型に特定すれば，効用最大化問題の最適解として得られる需要関数がLogit型，間接効用関数がLogsum型になるということを示したもの.
    - 経済理論との整合性を求められる場面（交通ネットワーク均衡モデル，応用一般均衡モデル等）で主に活用
  - 注：ランダム効用理論(総効用≡確定効用＋確率効用)に基づくアプローチではない [定式化にεは出てこない]

- 次に，ランダム効用理論 (McFadeen 1974) との関連性について考える.

ランダム効用理論～Logit Example

Additive random utility model (ARUM) ～ *with Gumbel error term*

$$u_j = \delta_j + \epsilon_j, \forall j \in \mathcal{J}$$

McFadden (1974)

Surplus (Indirect utility) ～ *Logsum*

$$G(\boldsymbol{\delta}) \equiv \mathbb{E}\left(\max_{j \in \mathcal{J}} u_j\right) = \ln\left(\sum_{j \in \mathcal{J}} e^{\delta_j}\right)$$

Williams (1977)
Daly & Zachary (1979)

Ben-Akiva (1973)
McFadden (1978)
Small & Rosen (1981)
→末尾に簡略証明記載

Choice Probability (Demand) ～ *Logit model*

$$q_j(\boldsymbol{\delta}) = \frac{e^{\delta_j}}{\sum_{j'=0}^{J} e^{\delta_{j'}}}, \forall j \in \mathcal{J}$$

$\mathbf{q}$: 需要(シェア)ベクトル
$\boldsymbol{\delta}$: 確定効用ベクトル
0, 1,..., J: 選択肢 (0: 合成財)
※簡略のため $\theta = 1$ とする

ランダム効用理論～General

- Additive Random Utility Model (ARUM)
  - ランダム効用 (ARUM)： $u_j = \delta_j + \epsilon_j, \forall j \in \mathcal{J}$
  - 余剰 (Surplus) 関数～期待最大効用～間接効用：

$$G(\boldsymbol{\delta}) = \mathbb{E}\left(\max_{j \in \mathcal{J}} u_j\right)$$

  - 期待最大効用の1階偏微分は選択確率に等しい：
    [Williams (1977) ～ Daly & Zachary (1979) Theorem]

$$q_j(\boldsymbol{\delta}) = \frac{\partial G(\boldsymbol{\delta})}{\partial \delta_j}, \forall j \in \mathcal{J}$$

※WDZ Theoremは，(ロジットモデルに限らず)任意のARUMで成り立つ.

ARUM-Logit ModelとEntropy Model

ARUM with Gumbel

$$u_j = \delta_j + \epsilon_j, \forall j \in \mathcal{J}$$

Surplus as Logsum

$$G(\boldsymbol{\delta}) \equiv \mathbb{E}\left(\max_{j \in \mathcal{J}} u_j\right) = \ln\left(\sum_{j \in \mathcal{J}} e^{\delta_j}\right)$$

$$G^{\bullet}(\mathbf{q}) = \max_{\boldsymbol{\delta}} \{\mathbf{q}^{\top}\boldsymbol{\delta} - G(\boldsymbol{\delta})\}$$

Shannon EntropyはLogsumの凸共役
(Convex conjugate)

Choice probability as Multinomial Logit

$$q_j(\boldsymbol{\delta}) = \frac{\partial G(\boldsymbol{\delta})}{\partial \delta_j} = \frac{e^{\delta_j}}{\sum_{j' \in \mathcal{J}} e^{\delta_{j'}}},$$

$$\forall j \in \mathcal{J}$$

Entropy Demands

Shannon Entropy

$$-G^{\bullet}(\mathbf{q}) = -\sum_{j \in \mathcal{J}} q_j \ln q_j$$

Representative Consumer (Purturbed Utility)

$$\max_{\mathbf{q} \in \boldsymbol{\Delta}} \left\{\sum_{j \in \mathcal{J}} \delta_j q_j + G^{\bullet}(\mathbf{q})\right\}$$

$$\text{where } \boldsymbol{\Delta} = \{\mathbf{q} \in \mathbb{R}_+^{J+1} \; : \; \textstyle\sum_{j=0}^{J} q_j = 1\}$$

一般化エントロピーモデル (GEM)

- GEM：単位シンプレックス Δ 内の需要ベクトルに関する Perturbed Utility 最大化
  $$\max_{\mathbf{q}\in\boldsymbol{\Delta}}\left\{\sum_{j\in\mathcal{J}}\delta_j q_j + G^{\bullet}(\mathbf{q})\right\}$$
- Generalized Entropy (GE)："Generator" 関数 $\mathbf{S}$ を用いて定義
  $$G^{\bullet}(\mathbf{q}) \equiv -\sum_{j\in\mathcal{J}} q_j \ln S^{(j)}(\mathbf{q})$$
  ※Logitの場合, $\mathbf{S}(\mathbf{q})=\mathbf{q}$ (Shannon Entropy)
- Generatorの逆関数 $\mathbf{H}=\mathbf{S}^{-1}$ を用いると, $H^{(j)}(e^{\boldsymbol{\delta}}) = \dfrac{\partial e^{G^{\bullet}(\boldsymbol{\delta})}}{\partial \delta_j}$

  →このとき，財 j の市場シェアは, $P_j(\boldsymbol{\delta}) = \dfrac{H^{(j)}(e^{\boldsymbol{\delta}})}{\sum_{j'=0}^{J} H^{(j')}(e^{\boldsymbol{\delta}})}$

  (Extended WDZ Theorem)
- 余剰関数(期待最大効用)は, $G(\boldsymbol{\delta}) = \ln\left(\sum_{j\in\mathcal{J}} H^{(j)}(e^{\boldsymbol{\delta}})\right)$

<u>任意のARUM</u>の選択確率は，<u>あるGEM</u>の需要関数として得ることができる

GEMのメリット(1)

1. 任意のARUMの選択確率は，あるGEMの需要関数として得ることができる（逆は必ずしも真ではない）．
2. GE (もしくは等価なGenerator)を設定することにより，より柔軟な代替関係/補完関係を表現可能な「需要システム(需要関数)」を導出することができる．特に，代替関係しか表現できないARUMに対し，GEMでは財の補完関係を表すような需要モデル (e.g. GNE) も記述することができる．

例1. Nested Logit

$$S^{(j)}(\mathbf{q}) = q_j^{\mu} \left(\sum_{j' \in g_j} q_{j'} \right)^{1-\mu}$$

例2. Generalized Nested Entropy (GNE)

$$S^{(j)}(\mathbf{q}) = \begin{cases} q_0, & (j=0) \\ q_j^{\mu_0} \prod_{c=1}^{C} q_{\sigma_c(j)}^{\mu_c}, & (j>0) \end{cases}$$

3. 離散型財のマーケットシェア需要モデルの代表的推定方法である Berry, Levinsohn & Pakes (1995) 法を，操作変数で代替し，簡略化・高精度化することができる(数値積分計算不要)．

$$\ln\left(\frac{q_{jt}}{q_{0t}}\right) = \beta_0 + \mathbf{X}_{jt}\boldsymbol{\beta} - \alpha p_{jt} + \sum_{c=1}^{C} \mu_c \ln\left(\frac{q_{jt}}{q_{\sigma_c(j),t}}\right) + \xi_{jt}$$

→線形操作変数法によるパラメータ推定

GEMのメリット(2)

4. 動的離散選択モデル (e.g. Rust 1987) の推定性能向上をもたらすことができる.
   - Chiong et al. (2016)
     a. $G^{\bullet}(\mathbf{q}(x))$ の劣微分として Choice-Specificな価値観数 $\boldsymbol{\delta}(x)$ を求める問題を，Monge-Kantorovich 問題（最適輸送問題）に準えて導出し，誤差分布を離散近似した上で線形計画問題として求解
     b. 前ステップで得られた $\boldsymbol{\delta}(x)$ を用いて $\boldsymbol{\delta}(x) \equiv \bar{u}(x) + \beta\mathbb{E}\left[V(x')\right]$ における $\bar{u}(x)$（即時効用）を算出
     ※ Foegerau et al. (2013) のベルマン方程式求解法と同じ行列計算

   → Hotz & Miller (1993) と同程度に高速ながらも，精度の高いパラメータ推定を実現

5. 相対Entropy=相互情報量 という観点に立てば，情報獲得の影響を考慮した不確実性下での離散選択モデルの理論的基礎を与えることができる.
   - 合理的無反応 (Rational Inattention: RI) モデル

まとめ

- ランダム効用理論 ≠ 摂動効用理論(エントロピー効用)
- 双対概念が重要となるのは，実は，摂動効用理論の方.
- しかし，両者は密接に関連する.

- 交通行動分析・交通ネットワーク分析への適用
[この辺りの詳細はver.3にて解説できるかもしれません]
  - Recursive Logit 型モデルの効率的推定 (Chiong et al. 2016)
  - 確率的な交通均衡配分における最適課金フレーム決定：
  〜ネットワークの総効用をどう定義するか？
    - Yang (1999): Shannon Entropy type Perturbated Utility -> Logit SUE
    - 円山ら (2003): 階層型 Shannon Entropy type Perturbated Utility
    -> Nested-Logit SUE
    - Fukuda et al. [未定稿]: Link-decomposed Shannon Entropy-like
    Perturbated Utility -> Markovian Traffic Equilibrium

参考文献(1)

1. Anderson, S.P., de Palma, A., Thisse, J.-F., 1988. A Representative Consumer Theory of the Logit Model. International Economic Review 29, 461–466.
2. Ben-Akiva, M. 1973. Structure of passenger travel demand models. Ph.D. Dissertation. MIT.
3. Berry, S., Levinsohn, J., Pakes, A., 1995. Automobile Prices in Market Equilibrium. Econometrica 63, 841–890.
4. Chiong, K.X., Galichon, A., Shum, M., 2016. Duality in dynamic discrete-choice models. Quantitative Economics 7, 83–115.
5. Daly, A., Zachary, S., 1979. Improved multiple choice models, in: Hensher, D., Dalvi, Q. (Eds.), Identifying and Measuring the Determinants of Mode Choice. Teakfield, London.
6. Fukuda. D., Kaneko, N., Ge, Q. [未定稿] Optimal Congestion Tolling Problem under Markovian Traffic Equilibrium.
7. Fosgerau, M., de Palma, A., Monardo, J., 2018. Demand Models for Differentiated Products with Complementarity and Substitutability. Social Science Research Network, Rochester, NY.
8. Fosgerau, M., Frejinger, E., Karlström, A., 2013a. A link based network route choice model with unrestricted choice set. Transportation Research Part B: Methodological 56, 70–80.
9. Fosgerau, M., Jiang, G., 2017. Travel Time Variability and Rational Inattention. Social Science Research Network, Rochester, NY.
10. Fosgerau, M., Melo, E., de Palma, A., Shum, M., 2017. Discrete Choice and Rational Inattention: A General Equivalence Result. Social Science Research Network, Rochester, NY.
11. Fudenberg, D., Iijima, R., Strzalecki, T., 2015. Stochastic Choice and Revealed Perturbed Utility. Econometrica 83, 2371–2409.
12. Hotz, V.J., Miller, R.A., 1993. Conditional Choice Probabilities and the Estimation of Dynamic Models. The Review of Economic Studies 60, 497–529.
13. Jiang, G., Fosgerau, M., Lo, H.K., 2019. Route choice, travel time variability, and rational inattention. Transportation Research Part B: Methodological.
14. Matějka, F., McKay, A., 2015. Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model. American Economic Review 105, 272–298.

参考文献(2)

15. McFadden, D., 1978. Modelling the Choice of Residential Location, in: A. Karlquist, F. Snickars and J. W. Weibull (Eds.) Spatial Interaction Theory and Planning Models. North Holland Amsterdam, 75–96.
16. Miyagi, T., 1986. On the Formulation of a Stochastic User Equilibrium Model Consistent with the Random Utility Theory – A Conjugate Dual Approach. Presented at the Proceedings of the Fourth World Conference on Transport Research, University of British Columbia, Vancouver, 1619–1635.
17. Rust, J., 1987. Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. Econometrica 55, 999–1033.
18. Sims, C.A., 2003. Implications of rational inattention. Journal of Monetary Economics 50, 665–690.
19. Small, K.A., Rosen, H.S., 1981. Applied Welfare Economics with Discrete Choice Models. Econometrica 49, 105–130.
20. Verboven, F., 1996. The nested logit model and representative consumer theory. Economics Letters 50, 57–63.
21. Williams, H.C.W.L., 1977. On the formation of travel demand models and economic evaluation measures of user benefit. Environment and Planning A 9, 285–344.
22. Yang, H., 1999. System optimum, stochastic user equilibrium, and optimal link tolls. Transportation science 33 (4), 354–360.
23. 円山琢也, 原田昇, 太田勝敏, 2003. Nested Logit型確率的利用者均衡条件下での最適混雑料金. 土木計画学研究・論文集, 20, 555–562.
24. 宮城俊彦, 小川俊幸, 1986. 共役性概念に基づくロジットモデルのパラメータ推定法とその統計的検定について. 土木計画学研究・論文集, 3, 185–192.
25. 室田一雄, 2007. 離散凸解析の考えかた–最適化における離散と連続の数理. 共立出版.

ログサム変数の簡略な導出

ランダムベクトル $\boldsymbol{\epsilon} = (\epsilon_i)_{i=1,\ldots,J}$（但し，$\epsilon_i \sim_{iid} Gumbel(\mu, 1), \forall i$）の要素の「最大値の確率分布」を求め，その上で，その分布の期待値を求める．

まず，μ を δ_i の一部とみなす市，結果として $\epsilon_i \sim Gumbel(0, 1)$ であると考える．（すなわち，各ϵ_i を $\epsilon_i - \mu$ で，また，各 δ_i を $\delta_i + \mu$で置き換える．）

$$X = \max_i(\delta_i + \epsilon_i) = \max_i((\delta_i + \mu) + (\epsilon_i - \mu)).$$

ランダム要素ベクトル ϵ_i が互いに独立であることは，ある実数値 x に対し，$\Pr(X \leq x)$ が個々の要素に対する生起確率 $\Pr(\delta_i + \epsilon_i \leq x)$の積に等しいことに相当する. このとき，両辺の対数を取ると次のようになる.

$$\log \Pr(X \leq x) = \log \prod_i \Pr(\delta_i + \epsilon_i \leq x) = \sum_i \log \Pr(\epsilon_i \leq x - \delta_i)$$

$$= -\sum_i e^{\delta_i} e^{-x} = -\exp\left(-x + \log \sum_i e^{\delta_i}\right).$$

この式は，位置パラメータ$\lambda = \log \sum_i e^{\delta_i}$を持つ分散1のガンベル分布のCDFの対数変換値である，すなわち，$X \sim Gumbel\left(\log \sum_i e^{\delta_i}, 1\right)$

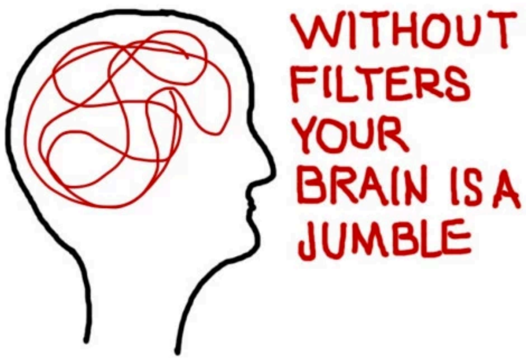
ガンベル分布の平均値の定義より，

$$\mathbb{E}[X] = \gamma + \log \sum_i e^{\delta_i}$$

QED.（より厳密な証明は，Small and Rosen (1981)を参照のこと）

合理的無反応 (Rational Inattention)

- 人の情報処理能力には限界があり，情報処理に要するコストの存在は，不確実性下での意思決定にも影響する
- Sims (2003)の「**合理的無反応 (RI)**」仮説:
  - 情報の蓄積によって生じるEntropyが閾値を超えたとき，経済主体が情報獲得に動くという仮説（マクロ経済学分野で進展）
  - 人々が合理的に不注意になる状況を選択する理論
- Matějka & McKay (2015):
  - Mutual Shannon Entropyにより情報コストを定義したとき，不確実性下での離散的選択行動が多項Logitモデル形式になることを証明
- Fosgerau, Melo, de Palma & Shum (2017):
  - Generalized Entropyによる情報コスト関数の拡張と，IIA特性の緩和
- Fosgerau & Jiang (2017):
  - 旅行時間信頼性分析 (旅行時間変動下での出発時刻選択分析) への適用



https://tcagley.wordpress.com/

Rational Inattention Model

RIに従う意思決定者の選択問題を変分問題として定式化：

Matějka & McKay (2015)

利得(効用)ベクトル

$$\max_{\mathbf{p}(\cdot)} \{\mathbb{E}(\mathbf{V} \cdot \mathbf{A}) - \lambda \times \text{information cost}\}$$

離散選択ベクトル

ランダムなアクション
(標準基底ベクトル)

Information cost に Mutual Shannon Entropyを仮定すると

事前確率の期待値

事前確率

$$= \max_{\mathbf{p}(\cdot)} \left\{ \sum_{\mathbf{v}\in\mathcal{V}} \{\mathbf{p}(\mathbf{v}) \cdot [\mathbf{v} + \lambda \log \mathbf{p}^0] - \lambda \mathbf{p}(\mathbf{v}) \cdot \log \mathbf{p}(\mathbf{v})\} \mu(\mathbf{v}) \right\}$$

$$\text{s.t. } p_i(\mathbf{v}) \geq 0 \ \forall i, \quad \sum_{i=1}^{N} p_i(\mathbf{v}) = 1$$

最適解 (MNL)

$$p_i(\mathbf{v}) = \frac{p_i^0 e^{v_i/\lambda}}{\sum_{j=1}^{N} p_j^0 e^{v_j/\lambda}} = \frac{e^{(v_i + \log p_i^0)/\lambda}}{\sum_{j=1}^{N} e^{(v_j + \log p_j^0)/\lambda}} = \frac{e^{\tilde{v}_i/\lambda}}{\sum_{j=1}^{N} e^{\tilde{v}_j/\lambda}}$$