行動モデルの導入
～MNLモデルとNLモデル～

工学部社会基盤学科
交通研究室　４年
前田　翠

基礎理論勉強会（2015/5/1）

目次

導入

個人の行動 ⟷ 選択 choice

- 連続型
- 連続型に近似される離散型
- 離散型

導入

●連続型

→最小二乗法が一般的

$$Y_n = \beta_1 X_1 + \cdots\cdots + \beta_k X_k + \varepsilon_n$$

Y_n:個人nの連続型被説明関数
β_k :k番目の未知パラメータ
X_k :個人nのk番目の説明変数
ε_n :確率項

●離散型

→確率効用最大化(RUM)モデル

$$V_{in} = \beta_1 X_{i1} + \beta_{i2} X_{i2} + \cdots\cdots + \beta_{ik} X_{ik}$$

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

U_{in} :確率効用
V_n:個人nの選択肢iに対する効用
β_k :k番目の未知パラメータ
X_k :個人の選択肢iに対するk番目の説明変数
ε_n :効用の確率項

MNLモデル

（1）二項選択モデル

（①二項プロビットモデル→②二項ロジットモデル）

（2）多項選択モデル

（①多項プロビットモデル→②多項ロジットモデル）

MNLモデル

（1）二項選択モデル

①二項プロビットモデル

$$P_n(i) = P_r[U_{in} \geq \ U_{jn}]$$

選択肢jではなく選択肢iを選ぶ確率

$$=P_r\ [V_{in} + \varepsilon_{in} \geq V_{jn}\ + \varepsilon_{jn}]$$

$$=P_r\ [\varepsilon_{jn} - \varepsilon_{in} \leq V_{in} - V_{jn}]$$

$\varepsilon_n = \varepsilon_{jn} - \varepsilon_{in}$

$$=P_r\ [\varepsilon_n \leq V_{in} - V_{jn}]$$

$$=F_\varepsilon(V_{in} - V_{jn})$$

F_ε:ε_nの累積分布関数

中心極限定理より
正規分布を仮定している

MNLモデル

$$P_n(i) = \Phi_\varepsilon(V_{in} - V_{jn})$$

$$= \int_{-\infty}^{V_{in}-V_{jn}} \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{exp}[-\frac{1}{2}(\frac{\varepsilon}{\sigma})^2]\mathrm{d}\varepsilon$$

$$= \int_{-\infty}^{(V_{in}-V_{jn})/\varepsilon} \frac{1}{\sqrt{2\pi\sigma^2}} exp[-\frac{1}{2}(z)^2]\mathrm{dz}$$

$$= \Phi_\varepsilon\left(\frac{V_{in}-V_{jn}}{\sigma}\right)$$

標準正規分布の累積分布関数

二項プロビットモデル
（binary probit model）

これでは積分形が残るので計算が煩雑！！→ロジットモデル

MNLモデル

②二項ロジットモデル

二項プロビットモデルでは積分形が残るので計算が煩雑になってしまう

2つの選択肢の確率項の差ε_nにロジスティック分布を仮定

ロジスティック分布：連続型の確率分布のひとつで、その分布関数がロジスティック関数となるもの（正規分布と一見類似しているがロジスティック分布の法が裾が長い）

F=$\frac{1}{1+e^{-(x-\mu)/s}}$,f=$\frac{\exp(-\frac{x-\mu}{s})}{s[1+\exp\{-\frac{e-\mu}{s}\}]^2}$,分散は$\frac{\pi^2 s^2}{3}$として表される

MNLモデル

二項プロビットモデルに以上の仮定を適用して計算すると

$$P_n(i) = \frac{1}{1 + \exp\{-(V_{in} - V_{jn})\}}$$

$$= \frac{\exp(\mu V_{in})}{\exp(\mu V_{in}) + \exp(\mu V_{jn})}$$

μ:ε_nのばらつきの程度を表すスケールパラメータ
標準偏差に反比例する

二項ロジットモデル
（binary logit model）

MNLモデル

（2）多項選択モデル

①多項選択プロビットモデル

$$P_n(i) = P_r[U_{in} \geq U_{jn}, \text{for } \forall j, j \neq i]$$

全選択肢から
選択肢iを選ぶ確率

$$= P_r[U_{in} \geq max.U_{jn}, \text{for } \forall j, j \neq i]$$

確率効用最大化モデル
RUM:
Random utility maximization

多項選択モデル（MNP/multinomial probit model）は
計算が煩雑すぎるためこれまではほとんど使用されていなかった

MNLモデル

①多項選択ロジットモデル

大事

それぞれの選択肢の確率効用項に独立で同一（IID:indentically and independently）なガンベル分布を仮定

RUMモデルに次章のガンベル分布の性質③と④を適用して計算（詳しくは後述）

$$P_n(i) = \frac{\exp(\mu V_{in})}{\sum_{j=1}^{J} \exp(\mu V_{jn})}, \ i=1,2,\dots,J$$

多項ロジットモデル（multinomial Logit model/MNL）

ガンベル分布とは

累積分布関数

$$F(\varepsilon) = \exp[-\exp\{-\mu(\varepsilon-\eta)\}]$$

確率密度関数

$$f(\varepsilon) = \mu\exp\{-\mu(\varepsilon-\eta)\}\exp[\exp\{-\mu(\varepsilon-\eta)\}]$$

μ:分布のばらつきを表すスケールパラメータ
η:分布の位置を表すロケーションパラメータ

ガンベル分布の性質

ロジットモデルでは、効用確定項の大小と誤差項、スケールパラメータμに注目しているので、ガンベル分布のロケーションパラメーターη＝0として扱うことが一般的

また式を簡潔にするためにスケールパラメーターμ＝1とすることが多い

①最頻値はη, 平均値は$\eta+\frac{\gamma}{\mu}$ (γ:オイラー定数)

②分散は$\frac{\pi^2}{6\mu^2}$

③ $\varepsilon_1, \varepsilon_2$がそれぞれ$(\eta_1,\mu),(\eta_2,\mu)$のパラメータを互いにも独立なガンベル分布に従う時
$\varepsilon=\varepsilon_1-\varepsilon_2$は以下のようなロジスティック分布に従う

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

④$\varepsilon_1, \varepsilon_2 \ldots\ldots \varepsilon_J$がそれぞれ$(\eta_1,\mu),(\eta_2,\mu)\ldots\ldots(\eta_j,\mu)$のパラメータを持つ互いに独立なガンベル分布に従う時、$\varepsilon_1, \varepsilon_2 \ldots\ldots \varepsilon_J$の最大値max. $(\varepsilon_1, \varepsilon_2 \ldots\ldots \varepsilon_J)$もガンベル分布に従い、そのパラメータは以下のようになる

$$\left(\frac{1}{\mu} ln \sum_{j=1}^{J} \exp(\mu\eta_j), \mu\right)$$

ガンベル分布の性質

ロジットモデルの式の導出

$\varepsilon_1, \varepsilon_2 \ldots. \varepsilon_J$がそれぞれ(0,$\mu$)のパラメータを持つ互いに独立なガンベル分布に従うとすると
（ロケーションパラメータηが0だとすると）
$P(\mathrm{i}) = \Pr[V_{in} + \varepsilon_{in} \geq \max_{\forall j, j \neq i}.(V_{jn} + \varepsilon_{jn})]$

ここで、$U^* = \max_{\forall j, j \neq i}.(V_{jn} + \varepsilon_{jn})$とおくと、$\varepsilon_{jn}$はガンベル分布に従い、$V_{jn} + \varepsilon_{jn}$もガンベル分布に従うので、性質④より、U*はパラメータ
（$\frac{1}{\mu}\ln\sum_{\forall j, j \neq i} \exp(\mu V_{jn}), \mu$）をもつガンベル分布に従う

$U_n{}^* = V_n{}^* + \varepsilon_n{}^*$とすると、パラメータηにあたる$\frac{1}{\mu}\ln\sum_{\forall j, j \neq i} \exp(\mu V_{jn})$は$\max_{\forall j, j \neq i}.(V_{jn} + \varepsilon_{jn})$の最頻値であるから、$V_n{}^* = \frac{1}{\mu}\ln\sum_{\forall j, j \neq i} \exp(\mu V_{jn})$と表せ、$\varepsilon_n{}^*$はパラメータ(0,$\mu$)を持つガンベル分布に従う

これより、性質③を用いると、
$P(\mathrm{i}) = \Pr[V_{in} + \varepsilon_{in} \geq \max_{\forall j, j \neq i}.(V_{jn} + \varepsilon_{jn})] = \Pr[V_{in} + \varepsilon_{in} \geq V_n{}^* + \varepsilon_n{}^*]$
これに$V_n{}^* = \frac{1}{\mu}\ln\sum_{\forall j, j \neq i} \exp(\mu V_{jn})$を代入して計算すると、
$P_n(i) = \frac{\exp(\mu V_{in})}{\sum_{j=1}^{J} \exp(\mu V_{jn})}$, i=1,2,….,J

IIA特性
(independence from irrelevant alternatives)

２つの選択肢を考える時、その２つの選択肢の確定効用にのみ影響を受け、選択肢集合に含まれる、他の選択肢の影響を受けないという性質。
つまり、無関係な選択肢からの選択確率の独立性。

IIA特性のあるモデルでは、選択肢集合に含まれるすべての選択肢を知らなくても、その部分集合を用いて、モデルを推定してもパラメータ推定値にバイアスが生じない。

IIA特性

ロジットモデルがIIA特性を有しているかどうか考えてみる…

$$P_n(i) = \frac{\exp(V_{in})}{\sum_{j=1}^{J} \exp(V_{jn})}, \text{ i=1,2,….,J}$$

より、

$$\frac{P_n(i)}{P_n(j)} = \exp(V_{in} - V_{jn})$$

任意の選択肢iと選択肢jの関係はその他の選択肢の影響を受けない
↓
IIA特性を有している

★

ロジットモデルに至る過程で…
それぞれの選択肢の確率効用項にIIDなガンベル分布を仮定

IIA特性

実際の選択状況がIIA特性に従っていない場合
（選択肢の確率効用項がIIDガンベル分布に従っていない場合）、
ロジットモデルではパラメータ推定にバイアスが生じる。
そのような場合、

→ロジットモデルではなく、確率効用項の独立性の仮定を
緩めたり、一般的なモデルを適用する必要がある

（１）IIA特性を緩和したGEVモデル

（２）GEVモデルの特殊系であるネスティッドロジットモデル

NLモデル

（１）GEVモデル

（generalized extreme value 一般極値分布）モデル

（２）ネスティッドロジットモデル（NLモデル）

NLモデル

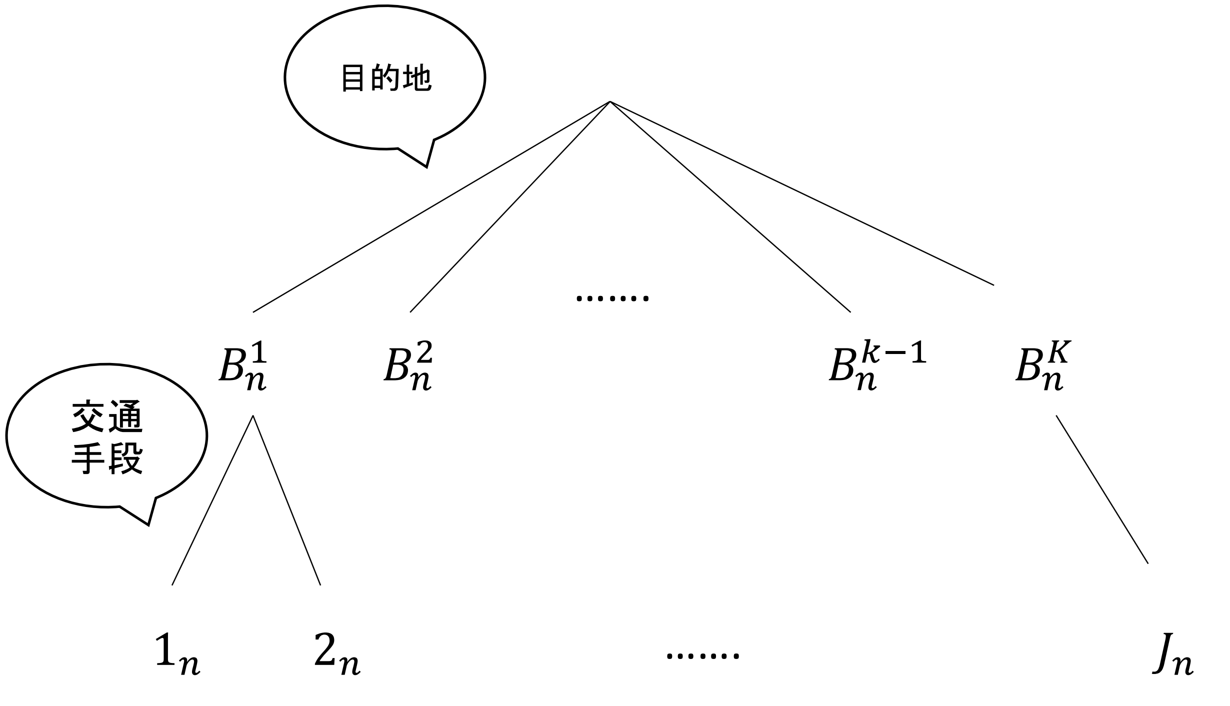
（1）GEVモデル

（generalized extreme value 一般極値分布）モデル

ロジットモデルでは確率の効用項にIIDなガンベル分布を仮定したが
GEVモデルでは、一般的な極値分布を仮定

→詳しくは三木さんのところで説明するので
ここでは軽く紹介だけ

NLモデル



個人nがJ_n個の選択肢を持ち、それが$B_n^1, B_n^2, \ldots B_n^K$というK個のサブグループに分かれているとする

選択肢iの効用を通常のように
U_{in}=V_{in}+ ε_{in}とすると、
ε_{in}が従う一般化極値分布の同時累積分布関数は、

$$F(\varepsilon)=\exp\left\{-\sum_{K=1}^{K} a_k \left(\sum_{B_n^K} \exp\left(-\frac{\varepsilon_{in}}{\lambda_k}\right)\right)^{\lambda_k}\right\}$$

λ_k:0と1の間の値をとるサブグループ内の相関の程度を表すパラメータ（1のとき無相関）

このとき選択確率は

$$P_n(i) = \frac{\exp\left(\frac{V_{in}}{\lambda_k}\right)\left\{\sum_{j\in B_n^K} \exp\left(\frac{V_{jn}}{\lambda_k}\right)\right\}^{\lambda_k-1}}{\sum_{K=1}^{K}\left(\sum_{j\in B_n^K} \exp\left(\frac{V_{jn}}{\lambda_k}\right)\right)^{\lambda_k}}$$

NLモデル

λ_k＝1のとき、通常のロジットモデルの式になることから、
ロジットモデルはGEVモデルの特殊形であることがわかる

また、これから説明するNLモデルもGEVモデルの特殊形である

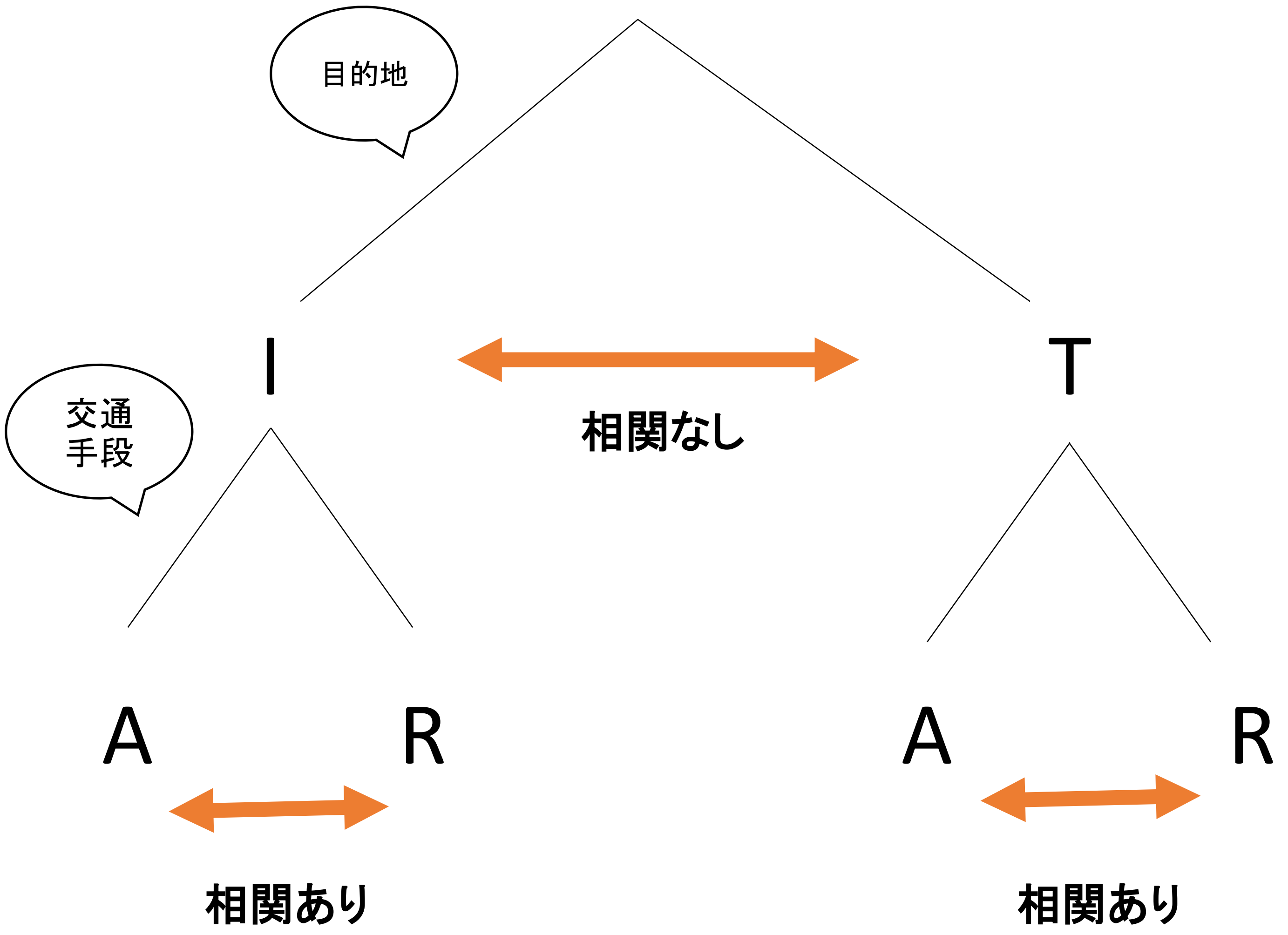
＜各モデルの違い＞

ロジットモデル：確率効用項の間に相関がない場合に扱えるモデル

GEVモデル：確率効用項の間で相関がある場合も相関がない場合も
扱えるモデル

NLモデル：選択肢のサブグループ内での相関を考えたモデル

NLモデル



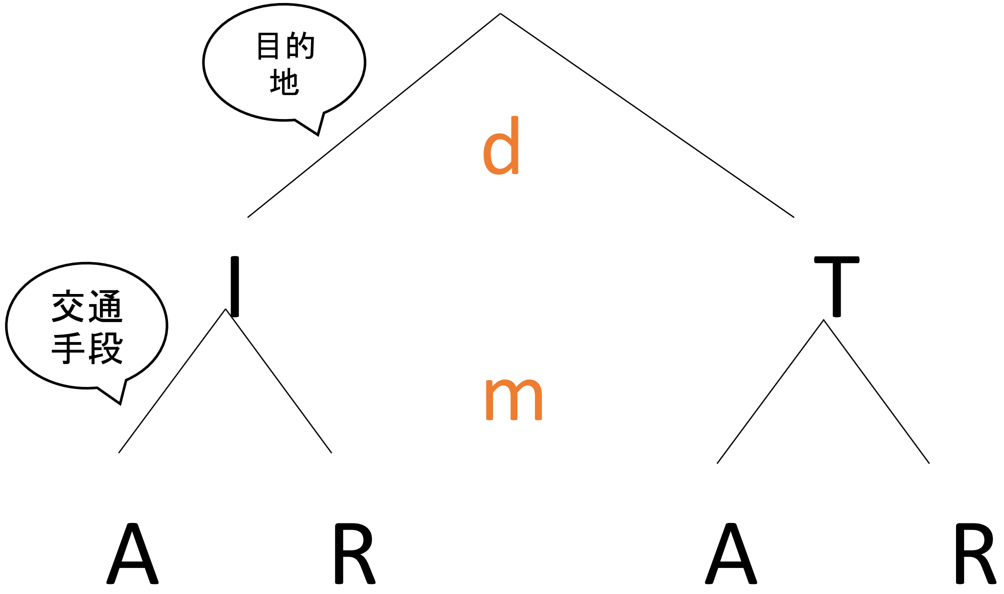
目的地選択の選択肢

d＝{I,T}

手段選択の選択肢

m＝{A,R}

NLモデル



$$U_{dm}=V_d+V_m+V_{dm}+\varepsilon_d+\varepsilon_{dm}$$

V_d: 目的地選択肢dに特有な効用の確定項

V_m: 手段選択肢mに特有な効用の確定項

V_{dm}:mとdの組み合わせ

ε_d:目的地選択肢dに特有な効用の確率項
（U_{dm}の最大値がスケールパラメータμ_dを持つガンベル分布になるような分布に従うと仮定）…(*)

ε_{dm}:mとdの組み合わせで決まる効用の確率項
（スケールパラメータμを持つIIDガンベル分布に従うと仮定）
…(**)

選択肢d_mの選択確率

$$P(d,m)=\underset{\text{①}}{\underline{P(m|d)}}\,\underset{\text{②}}{\underline{P(\mathrm{d})}}$$

NLモデル

①周辺確率P(d)

ε_{dm}はIIDガンベル分布に
従うと仮定している

$$P(d) = P_r[\max_{m\in\{A,R\}} U_{dm} \geq \max_{m\in\{A,R\}} U_{d'm},\ d' \neq d]$$

$$= P_r[V_d + \varepsilon_d + \max_{m\in\{A,R\}}(V_m + V_{dm} + \varepsilon_{dm}) \geq V_{d'} + \varepsilon_{d'} + \max_{m\in\{A,R\}}(V_m + V_{d'm} + \varepsilon_{dm}),\ d' \neq d]$$

(**)よりε_{dm}はスケールパラメータμを持つIIDガンベル分布に従うので

$$\max_{m\in\{A,R\}}(V_m + V_{dm} + \varepsilon_{dm})$$

もIIDガンベル分布に従い、同じスケールパラメータμを持つ

NLモデル

$\mathrm{V_m} + \mathrm{V_{dm}} + \mathrm{\varepsilon_{dm}}$もIIDガンベル分布に従うことから、
$\underset{m\in\{A,R\}}{max.}(V_m + V_{dm} + \varepsilon_{dm})$もIIDガンベル分布に従いそのパラメーターは、

（ガンベル分布の性質④より）

$$\underbrace{\frac{1}{\mu} ln \sum_{m\in\{A,R\}} exp\{\mu(V_m + V_{dm})\}, \mu}_{\mathrm{V_{d'}}}$$

$\mathrm{V_{d'}}$:ロケーションパラメータ
合成効用（inclusive value）/ログサム（logsum）変数と呼ばれている

NLモデル

dの選択について考えているので、各ネストの中の最大の効用同士を比べている

$$P(d)=P_r[\max_{m\in\{A,R\}} U_{dm} \geq \max_{m\in\{A,R\}} U_{d'm},\ d' \neq d]$$

$$=P_r[V_d + \varepsilon_d + \overbrace{\max_{m\in\{A,R\}}(V_m + V_{dm} + \varepsilon_{dm})}^{\mathrm{V_d}'+\varepsilon_d'} \geq V_{d'} + \varepsilon_{d'} + \overbrace{\max_{m\in\{A,R\}}(V_m + V_{d'm} + \varepsilon_{dm})}^{\mathrm{V_{d'}}'+\varepsilon_{d'}'},\ d' \neq d]$$

$\mathrm{V_d}'=\max_{m\in\{A,R\}}(V_m + V_{dm} + \varepsilon_{dm})$-$\varepsilon_d'$

となるように新たに誤差項ε_d'を定義すると

確率項

P(d)に$\mathrm{V}_{d'}$を代入すると、

$$\mathrm{P(d)=Pr}\left[V_d + \varepsilon_d + V_d' + \varepsilon_d' \geq V_{d'} + \varepsilon_{d'} + V_{d'}' + \varepsilon_{d'}',\ d' \neq d\right]$$

目的地dがもつ効用の確定項
（目的地dのもつ魅力度等）

AとRなど交通手段のすべてを
考えた交通アクセスの合成効用

離散選択モデル

NLモデル

(**)よりε_{dm}はスケールパラメータμを持つIIDガンベル分布に従うので

もIIDガンベル分布に従い、同じスケールパラメータμを持つ

また、$\varepsilon_{d'} = \max_{m\in\{A,R\}}.(V_m + V_{dm} + \varepsilon_{dm})$ - V_d'もガンベル分布に従う
V_d'はロケーションパラメータであり$\max_{m\in\{A,R\}}.(V_m + V_{dm} + \varepsilon_{dm})$の最頻値

よって、$\varepsilon_{dm} + \varepsilon_{d'}$もガンベル分布に従い、そのスケールパラメータを
μ_dとすると、結局周辺確率P(d)はロジット式で与えられて、

$$\boxed{\mathrm{P(d)} = \frac{\exp\{\mu_d(V + V_{d'}\}}{\sum_{d'\in\{I,T\}} \exp\{\mu_d(V' + V_{d''})\}}}$$

NLモデル

②条件付き確率P（m|d）

目的地dが与えられたときの手段選択では、
選択肢間の共通項であるV_dやε_dは関与しないので

$$\mathrm{P(m|d)} = \Pr[U_{dm} \geq U_{dm'},\ m \in \{A, R\},\ m' \neq m|\, d]$$
$$= \Pr[V_m + V_{dm} + \varepsilon_{dm} \geq V_{m'} + V_{dm'} + \varepsilon_{dm'}, m' \neq m|\, \boldsymbol{d}]$$

(∗∗)よりε_{dm}はスケールパラメータμをもついいDガンベル分布に従うから、
P(m|d)はロジット式で与えられて、

$$\boxed{\mathrm{P(m|d)} = \frac{exp\{\mu(V_m + V_{dm})\}}{\sum_{m' \in \{A,R\}} \exp\{\mu(V_{m'} + V_{dm'})\}}}$$

NLモデル

選択肢d_mの選択確率
$P(d,m) = P(m|d)P(\mathrm{d})$

①,②より、

$$P(d,m) = P(m|d)P(\mathrm{d})$$

$$= \frac{exp\{\mu(V_m+V_{dm})\}}{\sum_{m'\in\{A,R\}} \exp\{\mu(V_{m'}+V_{dm'})\}} \frac{\exp\{\mu_d(V+V_{d'}\}}{\sum_{d'\in\{I,T\}} \exp\{\mu_d(V'+V_{d'}')\}}$$

ネスティィッドロジットモデル