オーダードレスポンスモデルと頻度モデル

名古屋大学　山本俊行

オーダードレスポンスモデルとは？

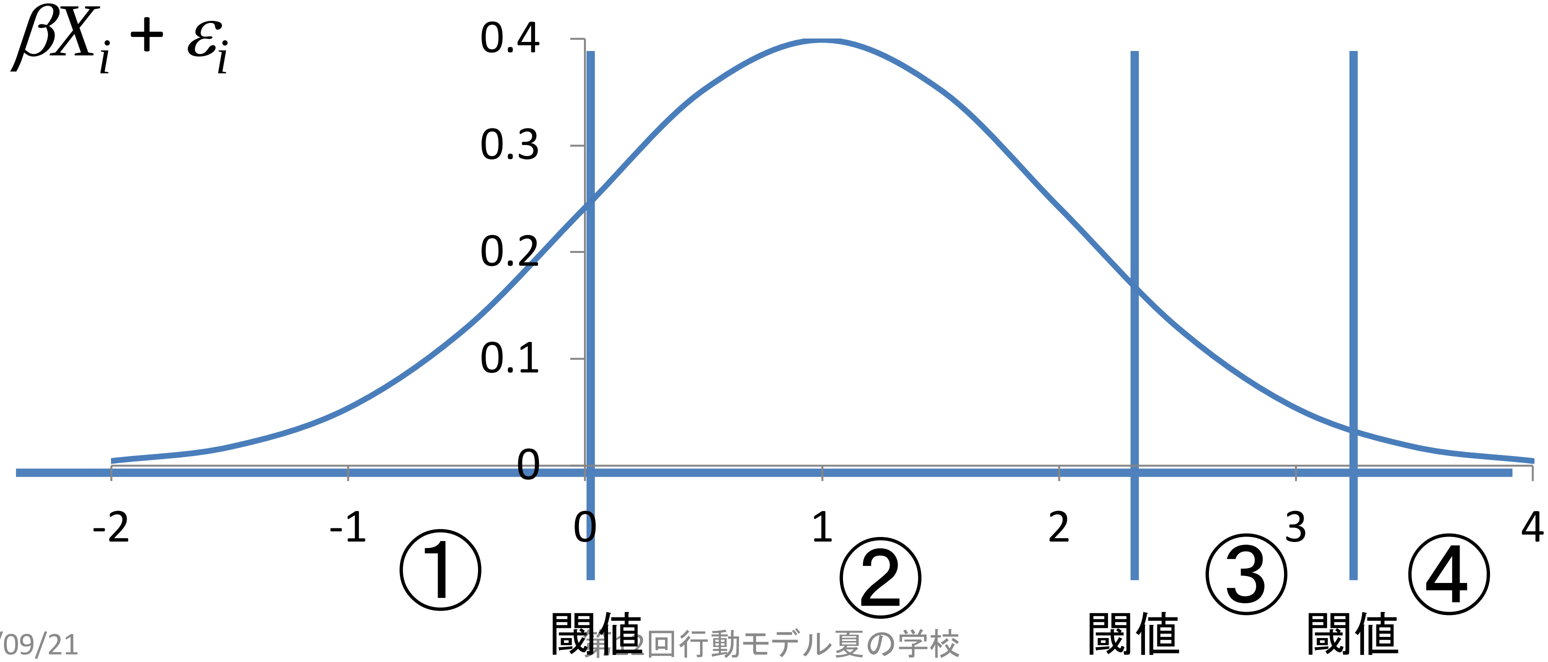
- 被説明変数が離散値として観測される
- 観測される変数の背後に連続的な（潜在）変数を想定する

- 適用例
  - 自動車保有台数：0台，1台，．．．J台
  - 交通事故の損傷度：車両のみ，軽傷，重症，死亡

$$
\begin{aligned}
y_i &= 0 \quad \text{if } y_i^* \le 0, \\
&= 1 \quad \text{if } 0 < y_i^* \le \mu_1, \\
&= 2 \quad \text{if } \mu_1 < y_i^* \le \mu_2, \\
&\mathrm{M} \\
&= J \quad \text{if } \mu_{J-1} < y_i^*
\end{aligned}
$$

y_i : 観測変数
y_i^* : 潜在変数
μ_j : 閾値
β : 未知パラメータ
X_i : 説明変数ベクトル
ε_i : 誤差項（正規分布やロジスティック分布）

$$y_i^* = \beta X_i + \varepsilon_i$$



$$
\begin{aligned}
y_i &= 0 \quad \text{if } y_i^* \le 0,\\
&= 1 \quad \text{if } 0 < y_i^* \le \mu_1,\\
&= 2 \quad \text{if } \mu_1 < y_i^* \le \mu_2,\\
& \mathrm{M}\\
&= J \quad \text{if } \mu_{J-1} < y_i^*
\end{aligned}
$$

$$y_i^* = \beta X_i + \varepsilon_i$$

y_i:観測変数
y_i^*:潜在変数
μ_j:閾値
β:未知パラメータ
X_i:説明変数ベクトル
ε_i:誤差項

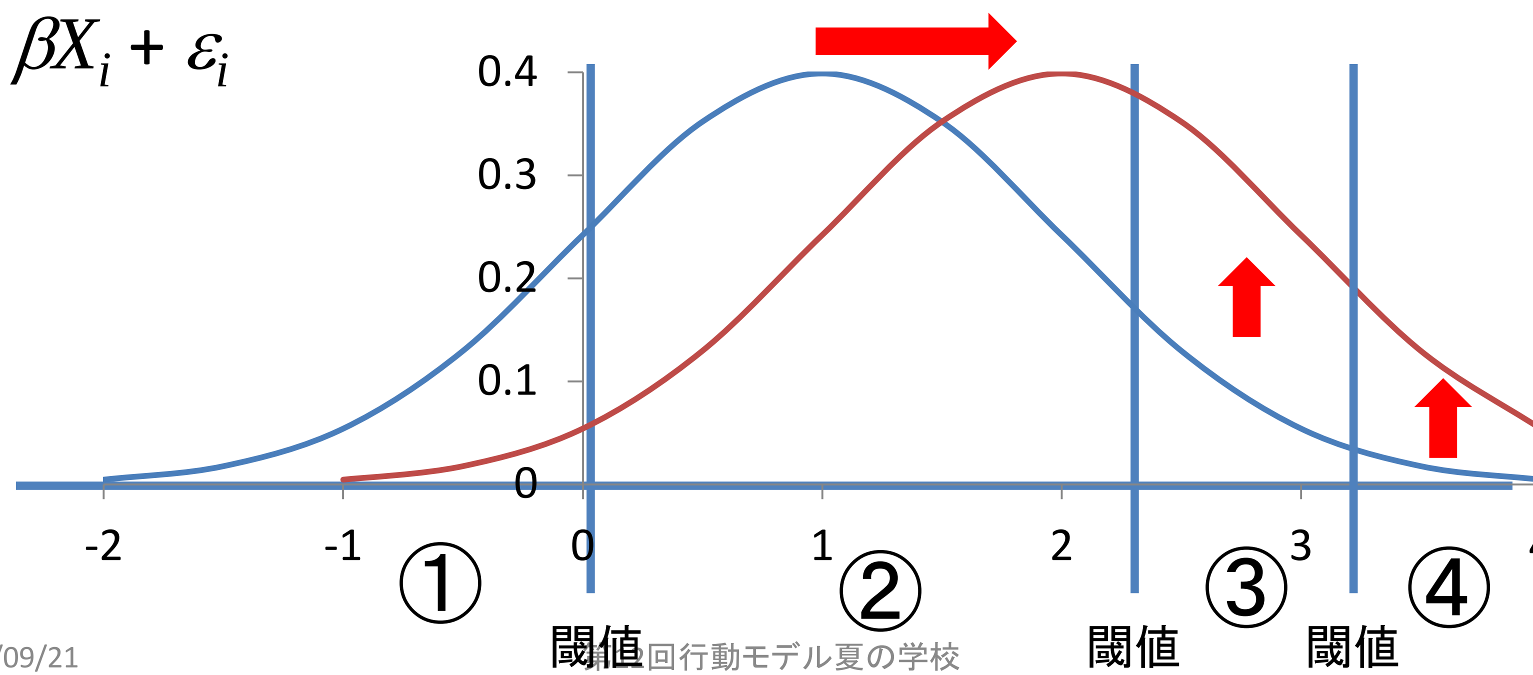
$F(\ )$:累積分布関数

$$P(y_i = 0) = F(-\beta X_i),$$
$$P(y_i = j) = F(\mu_j - \beta X_i) - F(\mu_{j-1} - \beta X_i),\ \ j = 1,2,\ldots J-1,$$
$$P(y_i = J) = 1 - F(\mu_{J-1} - \beta X_i)$$

モデルの制約

- 説明変数の変化が全ての観測変数値の生起確率に影響する(Savolainen et al., 2011)

$$y_i^* = \beta X_i + \varepsilon_i$$

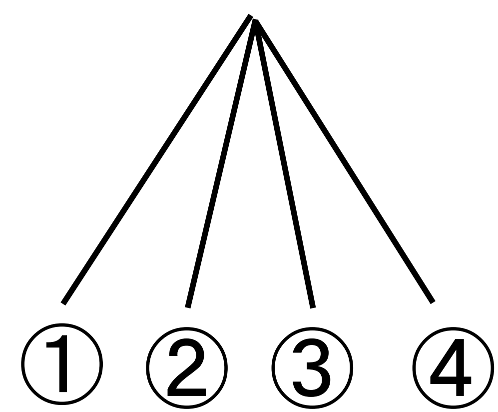


不都合な例と対応

- 交通事故の例では，エアバッグは
  - 死亡確率を減らす（潜在変数が左向きに動く）
  - 車両のみ→軽傷の効果（右向き）

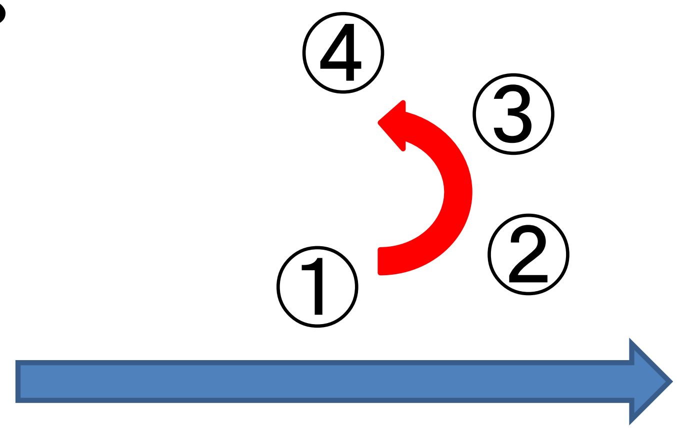
  の効果を同時に持つ

- 対応例：離散選択モデルの適用
  - ○自由度は向上
  - ×観測変数の序数性は無視

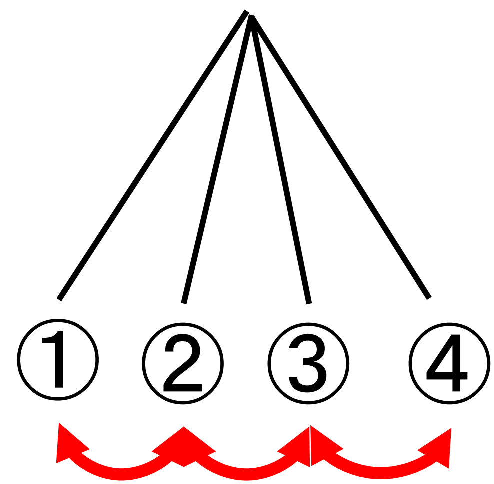


オーダードGEV，MXLモデル

- 隣り合った選択肢間の誤差項に相関を持たせる
- 序列性というより近接性の表現
  - 必ずしも全体の大小関係を表現しない？



大小関係



序数性を保持した対応

- 閾値の構造化
- 観測変数値別のパラメータ
- 繰り返し2項選択モデル

- 閾値の構造化

$$
\begin{aligned}
y_i &= 0 \quad \text{if } y_i^* \le 0, \\
&= 1 \quad \text{if } 0 < y_i^* \le \mu_{1i}, \\
&= 2 \quad \text{if } \mu_{1i} < y_i^* \le \mu_{2i},
\end{aligned}
$$

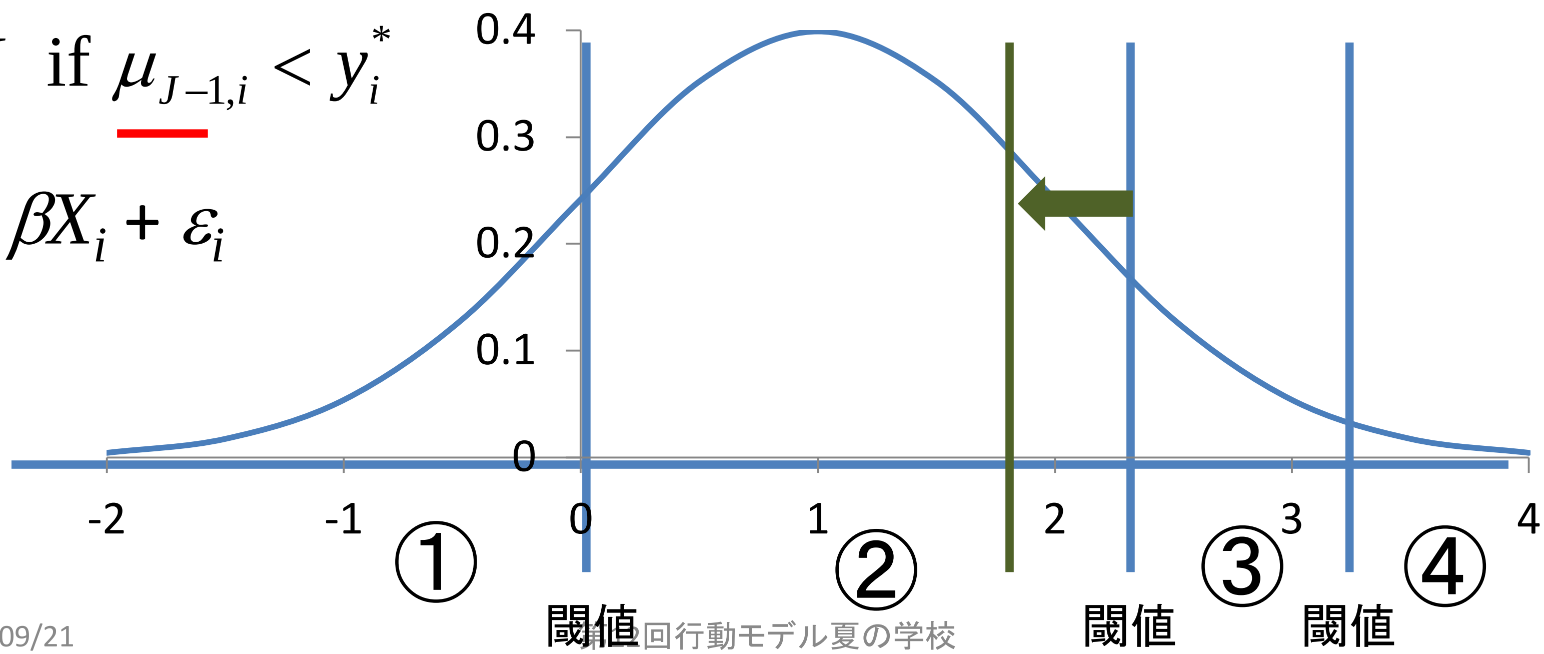
M

$$
= J \quad \text{if } \mu_{J-1,i} < y_i^*
$$

$$
\mu_{ji} = \gamma_j X_i
$$

γ_i：未知パラメータ

$$
y_i^* = \beta X_i + \varepsilon_i
$$



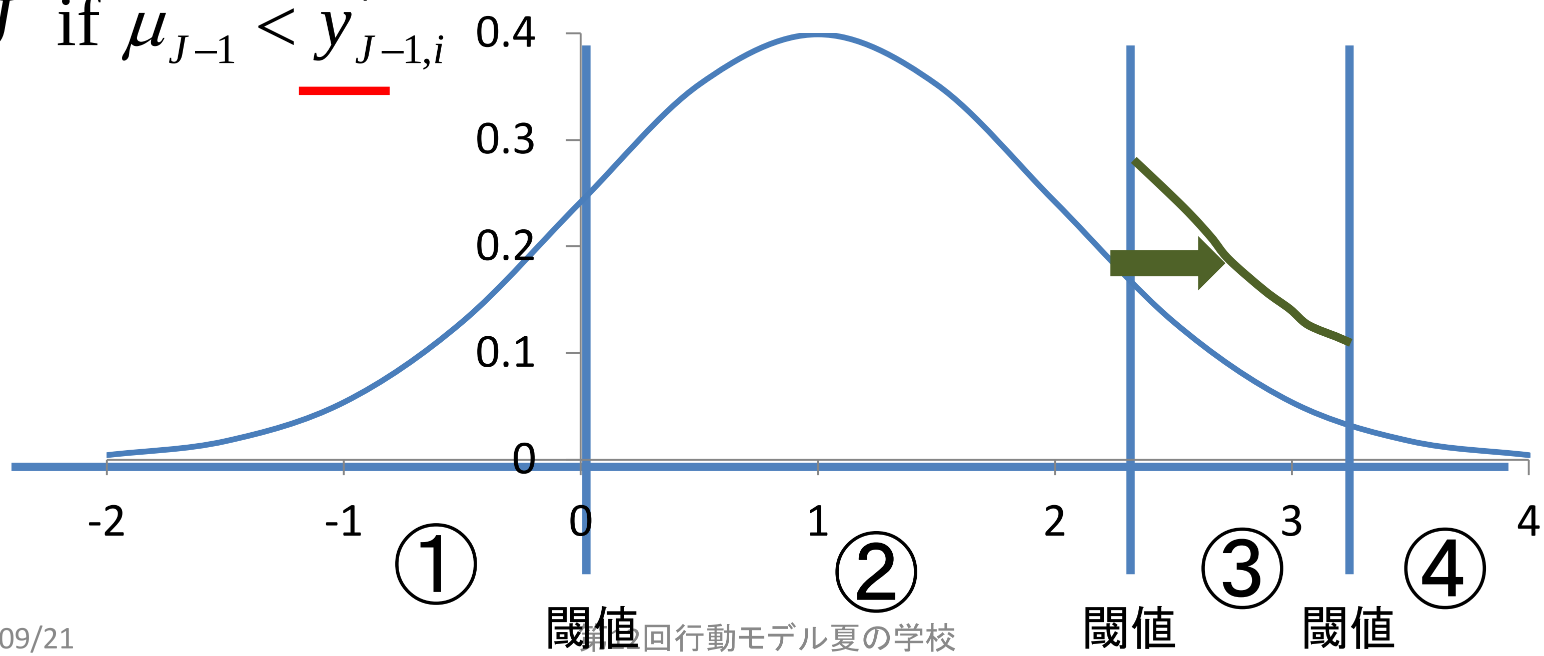
- 観測変数値別のパラメータ

$$y_i = 0 \quad \text{if } \underline{y_{0i}^*} \le 0,$$
$$= 1 \quad \text{if } 0 < \underline{y_{0i}^*},\ y_{1i}^* \le \mu_1,$$
$$= 2 \quad \text{if } \mu_1 < \underline{y_{1i}^*},\ y_{2i}^* \le \mu_2,$$

M

$$= J \quad \text{if } \mu_{J-1} < \underline{y_{J-1,i}^*}$$

$$y_{ji}^* = \underline{\beta_j} X_i + \varepsilon_i$$



閾値の構造化と観測変数値別のパラメータの効果は一緒

- 閾値の構造化

$$P(y_i = 0) = F(-\beta X_i),$$
$$P(y_i = j) = F(\gamma_j X_i - \beta X_i) - F(\gamma_{j-1} X_i - \beta X_i),\ j = 1,2,\ldots J-1,$$
$$P(y_i = J) = 1 - F(\gamma_{J-1} X_i - \beta X_i)$$

- 観測変数値別のパラメータ

$$P(y_i = 0) = F(-\beta_0 X_i),$$
$$P(y_i = j) = F(\mu_j - \beta_j X_i) - F(\mu_{j-1} - \beta_{j-1} X_i),\ j = 1,2,\ldots J-1,$$
$$P(y_i = J) = 1 - F(\mu_{J-1} - \beta_{J-1} X_i)$$

括弧内の引き算の引く側と引かれる側のどちらにすべきかは識別不可能

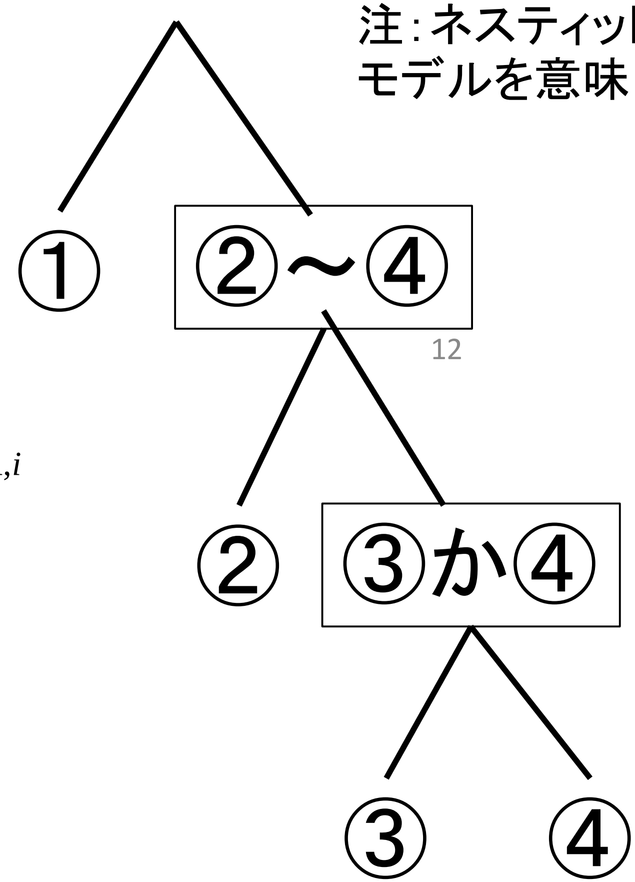
序数性を保持した対応

- 繰り返し2項選択モデル

注：ネスティッドロジット
モデルを意味しない

$$
\begin{aligned}
y_i &= 0 \quad \text{if } y_{0i}^* \le 0, \\
&= 1 \quad \text{if } 0 < y_{0i}^* \ \text{ and } \ y_{1i}^* \le 0, \\
&= 2 \quad \text{if } 0 < y_{0i}^*,\ 0 < y_{1i}^* \ \text{ and } \ y_{2i}^* \le 0, \\
&\ \vdots \\
&= J \quad \text{if } 0 < y_{0i}^*,\ 0 < y_{1i}^*, \ldots, \text{and } \ 0 < y_{J-1,i}^*
\end{aligned}
$$

$$
\begin{aligned}
y_{0i}^* &= \underline{\beta_0} X_i + \underline{\varepsilon_{0i}} \\
y_{1i}^* &= \underline{\beta_1} X_i + \underline{\varepsilon_{1i}} \\
&\ \vdots \\
y_{J-1,i}^* &= \underline{\beta_{J-1}} X_i + \underline{\varepsilon_{J-1,i}}
\end{aligned}
$$



- 繰り返し2項選択モデルの選択確率式

$$
\begin{aligned}
y_i &= 0 \quad \text{if } y_{0i}^* \le 0, \\
&= 1 \quad \text{if } 0 < y_{0i}^* \ \text{ and } \ y_{1i}^* \le 0, \\
&= 2 \quad \text{if } 0 < y_{0i}^*,\ 0 < y_{1i}^* \ \text{ and } \ y_{2i}^* \le 0, \\
&\ \vdots \\
&= J \quad \text{if } 0 < y_{0i}^*,\ 0 < y_{1i}^*, \ldots, \text{and } \ 0 < y_{J-1,i}^*
\end{aligned}
$$

$$
\begin{aligned}
y_{0i}{}^* &= \beta_0 X_i + \varepsilon_{0i} \\
y_{1i}{}^* &= \beta_1 X_i + \varepsilon_{1i} \\
&\ \vdots \\
y_{J-1,i}{}^* &= \beta_{J-1} X_i + \varepsilon_{J-1,i}
\end{aligned}
$$

$$
\begin{aligned}
P(y_i = 0) &= F(-\beta_0 X_i), \\
P(y_i = j) &= \prod_{k=0}^{j-1} \{F(\beta_k X_i)\} F(-\beta_j X_i), \\
P(y_i = J) &= \prod_{k=0}^{J-1} \{F(\beta_k X_i)\}
\end{aligned}
$$

繰り返し2項選択モデルと一般化オーダードモデル（閾値の構造化・観測変数値別のパラメータ）の相違

- 繰り返し2項選択モデル

$$P(y_i=0)=F(-\beta_0 X_i),$$

$$P(y_i=j)=\prod_{k=0}^{j-1}\{F(\beta_k X_i)\}F(-\beta_j X_i),$$

$$P(y_i=J)=\prod_{k=0}^{J-1}\{F(\beta_k X_i)\}$$

- 一般化オーダードモデル

$$P(y_i=0)=F(-\beta_0 X_i),$$

$$P(y_i=j)=F(\mu_j-\beta_j X_i)-F(\mu_{j-1}-\beta_{j-1} X_i)$$

$$P(y_i=J)=1-F(\mu_{J-1}-\beta_{J-1} X_i)$$

累積分布関数の掛け算か引き算かの違い

つまり，誤差項が独立（繰り返しモデル）か同一（一般化モデル）の相違

序数性を保持した対応

- 閾値の構造化
- 観測変数値別のパラメータ
  - 誤差項は観測変数値（閾値）間で同一

- 繰り返し2項選択モデル
  - パラメータと誤差項のいずれもが観測変数値別

- 誤差項が同一の場合も独立の場合もありなら，その中間で，相関を持つ場合もあり得る
  - 交通事故データでやってみましたが相関はあまり有意ではありませんでした．

頻度モデル

- 被説明変数は離散値として観測される
- 観測される変数の背後に連続的な（潜在）変数を想定する

ポアソン分布

$$\Pr(y_i) = \lambda_i^{y_i} \exp(-\lambda_i) / y_i\,!$$

$$\lambda_i = \exp(\beta X_i)$$

y_i: 観測変数
λ_i: 潜在変数
β : 未知パラメータ
X_i: 説明変数ベクトル

- $\ln \lambda_i = \beta X_i$ はオーダードレスポンスモデルの潜在変数に相当 $\ln \lambda_i = \beta X_i \left(= y_i^*\right)$

頻度モデル

- 交通事故分析やアクティビティ分析において適用される

- 適用例
  - 各リンクの交通事故の回数
  - 1週間のトリップ頻度

非観測異質性がある場合

- 負の二項分布モデル

$$\Pr\left(y_i\right)=\lambda_i^{\,y_i}\exp\left(-\lambda_i\right)/y_i\,!$$

$$\lambda_i=\exp\left(\beta X_i+\psi_i\right)$$

ψ_i:ガンマ分布にしたがう誤差項（分散$1/\theta$）

ガンマ分布の確率密度関数を乗じて積分すると以下の周辺確率が得られる

$$\Pr\left(y_i\right)=\frac{\Gamma\left(y_i+\theta\right)}{y_i!\Gamma\left(\theta\right)}\left(\frac{\lambda_i}{\lambda_i+\theta}\right)^{y_i}\left(\frac{\theta}{\lambda_i+\theta}\right)^{\theta}$$

頻度が複数の事象の合計だったら？

- 有り得る例
  - 個々に異なる原因の交通事故をまとめて回数を数えている
  - 異なる目的のトリップを区別せず数えている

- 多変量頻度モデル

多変量頻度モデル

- 各事象が独立の場合（単なる掛け算）

$$\Pr\left(y_{i1}, y_{i2}, \ldots, y_{iK}\right) = \prod_{k=1}^{K} \left\{ \lambda_{ik}{}^{y_{ik}} \exp\left(-\lambda_{ik}\right) \Big/ y_{ik}! \right\}$$

$$= \frac{y_{iT}!}{\prod_{k=1}^{K} y_{ik}!} \prod_{k=1}^{K} \left(\frac{\lambda_{ik}}{\lambda_{iT}} \right)^{y_{ik}} \frac{\lambda_{iT}{}^{y_{iT}} \exp\left(-\lambda_{iT}\right)}{y_{iT}!}$$

$$y_{iT} = \sum_{k=1}^{K} y_{ik}, \quad \lambda_{iT} = \sum_{k=1}^{K} \lambda_{ik}$$

- 非観測異質性があり各事象が独立の場合（単なる掛け算）

$$\Pr\left(y_{i1}, y_{i2}, \ldots, y_{iK}\right)$$

$$= \prod_{k=1}^{K} \left\{ \frac{\Gamma\left(y_{ik} + \theta_k\right)}{y_{ik}!\,\Gamma\left(\theta_k\right)} \left(\frac{\lambda_{ik}}{\lambda_{ik} + \theta_k} \right)^{y_{ik}} \left(\frac{\theta_k}{\lambda_{ik} + \theta_k} \right)^{\theta_k} \right\}$$

- 非観測異質性が各事象間で同一の場合（まとめられるだけ）

$$\Pr\left(y_{i1}, y_{i2}, \ldots, y_{iK}\right) = \frac{y_{iT}!}{\prod_{k=1}^{K} y_{ik}!} \prod_{k=1}^{K} \left(\frac{\lambda_{ik}}{\lambda_{kT}} \right)^{y_{ik}}$$

$$\times \frac{\Gamma\left(y_{iT} + \theta\right)}{y_{iT}!\,\Gamma\left(\theta\right)} \left(\frac{\lambda_{iT}}{\lambda_{iT} + \theta} \right)^{y_{iT}} \left(\frac{\theta}{\lambda_{iT} + \theta} \right)^{\theta}$$

非観測異質性が相関を持つ場合

- 非観測異質性が独立の場合と同一の場合があるなら，相関がある場合も考えられる

- でも，クローズドフォームはないみたい．なので，数値積分？あるいはCML (composite marginal likelihood)推定？

オーダードモデルと頻度モデルの関係

	オーダードモデル	頻度モデル
潜在変数	$y_i^* = \beta X_i + \varepsilon_i$	$\ln \lambda_i = \beta X_i \left(= y_i^*\right)$
確率分布形	正規分布，ロジスティック分布	ポアソン分布，負の二項分布
閾値	未知パラメータとして推定	不要（上記の分布形によって決まっていると見なせる）
特徴	閾値がパラメータになっている分，自由度が高い	閾値を推定しないため，データで観測されていない「頻度」の生起確率も推定可能

頻度モデルのオーダードモデル的表現（Castro et al. 2012）

オーダードモデル

$$P(y_i = 0) = F(-\beta X_i),$$
$$P(y_i = j) = F(\mu_j - \beta X_i) - F(\mu_{j-1} - \beta X_i),\ j = 1,2,\ldots J-1,$$
$$P(y_i = J) = 1 - F(\mu_{J-1} - \beta X_i)$$

頻度モデル

$$\Pr(y_i) = \lambda_i^{y_i} \exp(-\lambda_i) / y_i!$$

オーダードモデルの各選択確率が頻度モデルの確率と一致するように，オーダードモデルの閾値を構造化する

$$\beta X_i = 0$$
$$\mu_{ij} = F^{-1}\left(\sum_{m=0}^{j} \Pr(m|\lambda_i)\right)$$
$$\mu_{ij-1} = F^{-1}\left(\sum_{m=0}^{j-1} \Pr(m|\lambda_i)\right)$$

頻度モデルによって0～jの観測値が得られる確率の和を正規分布の逆関数に導入

$$\Pr(m \mid \lambda_i) = \lambda_i^{m} \exp(-\lambda_i) / m!$$

頻度モデルのオーダーモデル的表現（Castro et al. 2012）

$$\beta X_i = 0$$

$$\mu_{ij} = F^{-1}\left(\sum_{m=0}^{j} \Pr(m|\lambda_i)\right)$$

$$\mu_{ij-1} = F^{-1}\left(\sum_{m=0}^{j-1} \Pr(m|\lambda_i)\right)$$

$$\Pr(m \mid \lambda_i) = \lambda_i^{\,m} \exp(-\lambda_i)/m!$$

$$P(y_i = j) = F(\mu_{ij}) - F(\mu_{ij-1})$$

$$= F\left(F^{-1}\left(\sum_{m=0}^{j} \Pr(m|\lambda_i)\right)\right) - F\left(F^{-1}\left(\sum_{m=0}^{j-1} \Pr(m|\lambda_i)\right)\right)$$

$$= \sum_{m=0}^{j} \Pr(m|\lambda_i) - \sum_{m=0}^{j-1} \Pr(m|\lambda_i) = \Pr(j|\lambda_i)$$

オーダーモデルに代入すると，もともとの頻度モデルと一致することが分かる

多変量頻度モデルのオーダードモデル的表現（Castro et al. 2012）

- $\beta X_i = 0$ は $y_i^* = \varepsilon_i$ を仮定していることになる.
- 多変量の場合，ε_{ki}, $\varepsilon_{k'i}$ 間に相関を仮定して，多変量正規分布を用いることで自由な相関を扱える.

$$\mu_{kij} = F^{-1}\left(\sum_{m=0}^{j} \Pr(m|\lambda_{ki})\right)$$

$$\mu_{kij-1} = F^{-1}\left(\sum_{m=0}^{j-1} \Pr(m|\lambda_{ki})\right)$$

（2変量の場合）

$$\begin{aligned}
&P(y_{1i} = j, y_{2i} = j') \\
&= \Pr\left(\mu_{1ij-1} < y_{1i}{}^* < \mu_{1ij}, \mu_{2ij'-1} < y_{2i}{}^* < \mu_{2ij'}\right) \\
&= \Pr\left[\begin{array}{l}
F^{-1}\left(\sum_{m=0}^{j-1} \Pr(m|\lambda_{1i})\right) < \varepsilon_{1i} < F^{-1}\left(\sum_{m=0}^{j} \Pr(m|\lambda_{1i})\right), \\
F^{-1}\left(\sum_{m=0}^{j'-1} \Pr(m|\lambda_{2i})\right) < \varepsilon_{2i} < F^{-1}\left(\sum_{m=0}^{j'} \Pr(m|\lambda_{2i})\right)
\end{array}\right]
\end{aligned}$$

多変量の場合の推定方法

- ε_{ki}, $\varepsilon_{k'i}$ 間に相関を仮定すると多変量正規分布の場合，多重数値積分が必要になる
- よって，数値積分による計算負荷を削減するため，CML (composite marginal likelihood)推定を適用する
- 2変量間の相関を考慮し，2変量の周辺確率を全ての変量の組み合わせについて求め，それらの積を尤度関数とする

Composite marginal likelihood (CML) (Bhat et al., 2010)

- 多変量分布を表現するのに多重積分が必要な場合，次元が大きくなると計算不可能

$$P(z_1 = d_1, z_2 = d_2, \ldots, z_Q = d_Q) = \int_{D(z)} c_\theta\left(F_1(z_1^*), F_2(z_2^*), \ldots, F_Q(z_Q^*)\right) dz_1^* dz_2^* \ldots dz_Q^*$$

この場合，Q次元の多重積分となる

- 2変量の周辺確率を全ての変量の組み合わせについて求め，それらの積を尤度関数とする

$$L_{CML}(\gamma) = \prod_{q=1}^{Q-1} \prod_{k=q+1}^{Q} \left[P(z_q = d_q, z_k = d_k)\right]^{\omega_{qk}}$$

Q(Q-1)/2項の積で，2変量間の相関を全ての変量の組み合わせについて推定可能

CMLの性質

- 一致性があり漸近的正規性を持つ
- 多変量分布を用いた最尤推定より理論上は有効性は落ちるが，実務上の問題はほとんどない
- 推定値の分散共分散行列の推定にはサンドイッチ推定量を用いる（Hessianはクロスプロダクトを利用）

参考文献

- オーダードレスポンスモデル
  - Savolainen, P.T., Mannering, F.L., Lord, D. and Quddus, M.A. (2011) The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accident Analysis and Prevention, Vol. 43, pp. 1666-1676.
- 多変量頻度モデル
  - Kockelman，K.M. and Krishnamurthy, S. (2004) A new approach for travel demand modeling: linking Roy’s Identity to discrete choice. Transportation Research Part B, Vol. 38, pp. 459–475.
- 多変量頻度モデルのオーダードレスポンスモデルとしての表現
  - Castro, M., Paleti, R. and Bhat, C.R. (2012) A latent variable representation of count data models to accommodate spatial and temporal dependence: Application to predicting crash frequency at intersections, Transportation Research Part B, Vol. 46, pp. 253-272.
  - 山本俊行，森川高行（2013）地域間競合を考慮した買い物頻度モデルの構築－大規模小売店舗の中心市街地への出店時の買い物行動変化の分析への適用－，都市計画論文集，Vol. 48, No. 3（印刷中）
- Composite marginal likelihood
  - Bhat, C.R., Sener, I.N. and Eluru, N. (2010). A flexible spatially dependent discrete choice model: formulation and application to teenagers’ weekday recreational activity participation., Transportation Research Part B, Vol. 44, pp. 903–921.