

大学中退の逐次意思決定モデルの構造推定

村澤 康友（甲南大学経済学部）

2022 年 9 月 15 日

労働経済・社会保障・経済政策に関するミニコンファレンス@同志社大学

研究の動機と貢献

大学中退モデル

構造推定

実証分析（手順のみ）

反実仮想分析

(大学生の) 成績データの分析

- ・ 鹿野・高木・村澤 (2011)
- ・ 大学における $IR = EBPM$

動的離散選択モデルの構造推定

- ・ 2 値選択 (在学／退学) の最適停止モデルが一番簡単
- ・ 退学＝授業料収入の損失 (大学の経営において重要！)

大学中退の現状

2014 年（6 年制は 2012 年）4 月入学者の 2018 年 3 月時点の中途退学・卒業・留年状況

		退学／入学	卒業／入学	留年／入学
国立	男	.033	.805	.162
	女	.021	.879	.100
公立	男	.056	.795	.149
	女	.031	.877	.092
私立	男	.097	.767	.136
	女	.061	.871	.068
計		.070	.819	.111

- ・ 大学中退行動を最適停止問題として定式化し
- ・ 某大学某学部 2016 年 4 月入学者のうち男子 301 名の入学後 4 年間の成績データ（打ち切りデータ）を用いて
- ・ 意思決定モデルの構造母数の推定と反実仮想分析を試みる

「動的離散選択モデルの構造推定」の大学中退行動への応用

1. 「賃金構造基本統計調査」から計算した退学後の生涯賃金を観測可能な退学の価値とみなして短期パネルでの構造母数の識別を確保
2. CCP 法を修正し、2 値ロジット・モデルの完全分離の問題を回避
3. 母数の識別と推定精度をモンテカルロ実験で確認
4. （構造母数の推定精度が低いため）効用関数の識別を必要としない反実仮想分析

研究の動機と貢献

大学中退モデル

構造推定

実証分析（手順のみ）

反実仮想分析

日本の一般労働者の生涯賃金（退職金を含めない）

	男	女
大卒	2 億 6910 万円	2 億 1730 万円
高卒	2 億 1280 万円	1 億 5230 万円
プレミアム	5630 万円	6500 万円

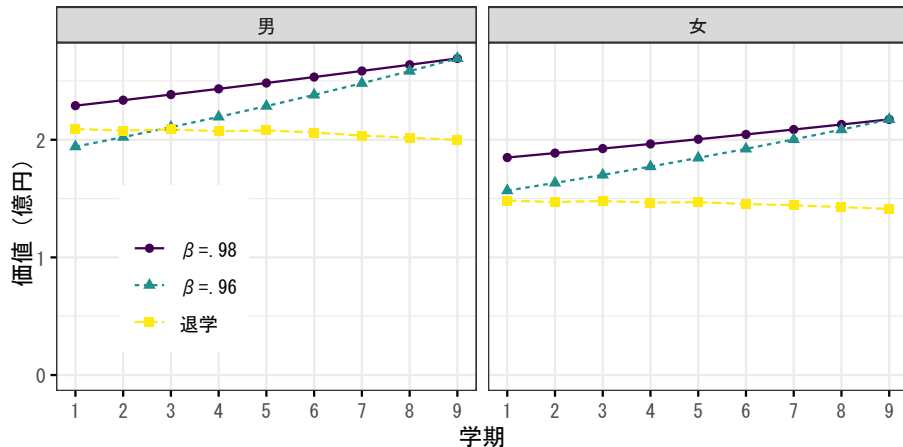
簡単な例

- ・ 入学者は退学しなければ 4 年後に自動的に卒業
- ・ 在学の効用／費用は 0
⇒ **在学の価値** = 大卒の生涯賃金の割引現在価値
- ・ **退学の価値** = その時点以降の高卒の生涯賃金
- ・ 割引因子 $\beta = .98$ (半期)

各学期の開始時に在学者は**価値が高い方の行動を選択**

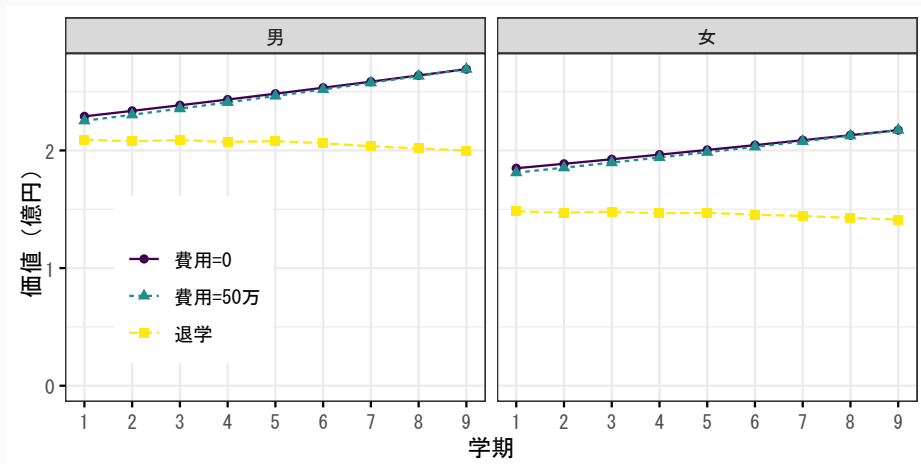
割引因子の低下の効果

男子は $\beta = .96$ なら 1 年次は退学を選択



在学費用の発生の効果

1 学期当たり 50 万円の費用なら意思決定は変化しない



最適停止モデル

意思決定者（学生）

- ・ d_t : 在学ダミー
- ・ s_t : 状態ベクトル
- ・ $U_t(d_t; s_t)$: 効用
- ・ β : 割引因子

分析者

- ・ x_t : 分析者が観測する状態ベクトル
- ・ $u_t(d_t; x_t) := E(U_t(d_t; s_t) | x_t)$
- ・ $e_t(d_t) := U_t(d_t; s_t) - u_t(d_t; x_t)$ (観測されない状態)
- ・ $e_t := (e_t(0), e_t(1))'$

ベルマン方程式

在学者／退学者の価値関数

- ・ $V_t^j(x_t, e_t)$: 価値関数
- ・ $\bar{V}_t^j(x_t)$: 積分した価値関数

積分したベルマン方程式

$$\bar{V}_t^1(x_t) = E \left(\max_{d_t \in \{0,1\}} \left\{ \underbrace{u_t(d_t; x_t) + e_t(d_t)}_{\text{今期の効用}} + \beta E \left(\underbrace{\bar{V}_{t+1}^{d_t}(x_{t+1})}_{\text{来期の価値}} \mid x_t; d_t \right) \right\} \mid x_t \right)$$

s.t. $x_{t+1} \sim F_{t+1}(\cdot \mid x_t; d_t)$

条件付き選択確率 (CCP)

選択肢別価値関数

$$v_t(\mathbf{x}_t; d_t) := u_t(d_t; \mathbf{x}_t) + \beta \mathbb{E} \left(\bar{v}_{t+1}^{d_t}(\mathbf{x}_{t+1}) | \mathbf{x}_t; d_t \right)$$

最適な選択

$$d_t^* := \arg \max_{d_t \in \{0,1\}} \{v_t(\mathbf{x}_t; d_t) + e_t(d_t)\}$$

$\{e_t(j)\}$ は iid で平均 0, **尺度母数 σ** のガンベル分布にしたがう

⇒在学確率関数 (CCP 関数) は **2 値ロジット・モデル**

$$p_t(1|\mathbf{x}_t) = \frac{\exp(v_t(\mathbf{x}_t; 1)/\sigma)}{\exp(v_t(\mathbf{x}_t; 0)/\sigma) + \exp(v_t(\mathbf{x}_t; 1)/\sigma)} = \Lambda \left(\frac{v_t(\mathbf{x}_t; 1) - v_t(\mathbf{x}_t; 0)}{\sigma} \right)$$

※後ろ向き帰納法で $v_t(\cdot; \cdot)$ を解く

研究の動機と貢献

大学中退モデル

構造推定

実証分析（手順のみ）

反実仮想分析

退学確率

$$p_t(0|\mathbf{x}_t) = \frac{\exp(v_t(\mathbf{x}_t; 0)/\sigma)}{\exp(v_t(\mathbf{x}_t; 0)/\sigma) + \exp(v_t(\mathbf{x}_t; 1)/\sigma)}$$

期待最大効用

$$\bar{V}_t^1(\mathbf{x}_t) = \sigma \ln \left(\exp \left(\frac{v_t(\mathbf{x}_t; 0)}{\sigma} \right) + \exp \left(\frac{v_t(\mathbf{x}_t; 1)}{\sigma} \right) \right)$$

第1式を対数変換し、第2式に代入すると

$$\underbrace{\bar{V}_t^1(\mathbf{x}_t)}_{\text{在学状態の価値}} = \underbrace{v_t(\mathbf{x}_t; 0)}_{\text{退学の価値}} \underbrace{-\sigma \ln p_t(0|\mathbf{x}_t)}_{\text{退学を保留する価値}}$$

識別制約

1. 在学時の効用 $u_t(1; \cdot)$ は時点 t に依存しない (除外制約)
2. 退学の価値 $v_t(\cdot; 0)$ は状態ベクトル \mathbf{x}_t に依存せず 既知

$y_t := v_t(\cdot; 0)$ とすると在学の価値は

$$\begin{aligned} v_t(\mathbf{x}_t; 1) &:= u(1; \mathbf{x}_t) + \beta E(\bar{V}_{t+1}^1(\mathbf{x}_{t+1}) | \mathbf{x}_t; 1) \\ &= u(1; \mathbf{x}_t) + \beta E(v_{t+1}(\mathbf{x}_{t+1}; 0) - \sigma \ln p_{t+1}(0 | \mathbf{x}_{t+1}) | \mathbf{x}_t; 1) \\ &= u(1; \mathbf{x}_t) + \beta y_{t+1} - \beta \sigma \underbrace{E(\ln p_{t+1}(0 | \mathbf{x}_{t+1}) | \mathbf{x}_t; 1)}_{\text{補正関数}} \end{aligned}$$

補正関数

$$z_t(\mathbf{x}_t) = \int \ln p_{t+1}(0 | \mathbf{x}_{t+1}) dF_{t+1}(\mathbf{x}_{t+1} | \mathbf{x}_t; d_t = 1)$$

$u(1; \mathbf{x}_t) := \mathbf{x}_t' \boldsymbol{\theta}$ と定式化し, $\eta := \sigma^{-1}$ とする.

1. 状態遷移確率分布 $\{F_t(.|.;.)\}$ を推定
2. 誘導形の CCP 関数 $\{p_t(.|..)\}$ をノンパラメトリック回帰で推定
3. $\{\hat{F}_t(.|.;.)\}, \{\hat{p}_t(.|..)\}$ から補正関数 $\{\hat{z}_t(.)\}$ を計算
4. 対数オッズ比のノンパラメトリック推定値を従属変数として $(\boldsymbol{\theta}, \beta, \eta)$ を非線形最小 2 乗法で推定

$$\ln \frac{\hat{p}_t(1|\mathbf{x}_{i,t})}{\hat{p}_t(0|\mathbf{x}_{i,t})} = \underbrace{\mathbf{x}_{i,t}' \eta \boldsymbol{\theta} + \beta \eta y_{t+1} - \beta \hat{z}_t(\mathbf{x}_{i,t})}_{\text{在学の価値}} - \underbrace{\eta y_t}_{\text{退学の価値}}$$

研究の動機と貢献

大学中退モデル

構造推定

実証分析（手順のみ）

反実仮想分析

某大学某学部 2016 年 4 月入学者の 2020 年 3 月末時点の中途退学・卒業・留年状況

	入学者数	退学者数	卒業者数	留年者数
男	301	12	242	47
女	101	2	94	5
計	402	14	336	52

※分析対象は男子のみ

設定

- ・ 毎学期 24 単位を履修
- ・ 8 学期の在籍と 128 単位の修得が卒業要件

変数

- ・ $N_{i,t}^* := N_{i,t} - 16t$ (超過累積修得単位数)
- ・ $g_{i,t}$: GPA 水準 (4 つの水準に離散化)

在学時の効用関数

$$u(1; \mathbf{x}_{i,t}) := \alpha + \kappa N_{i,t-1}^* + \lambda g_{i,t-1}$$

修正 CCP 法

1. 学期別の経験分布で状態遷移確率分布を推定
2. ロジット関数をリンク関数とした一般化加法モデルで誘導形の CCP 関数を推定
(交互作用なし)
3. 補正関数を計算
4. 在学／退学の対数オッズ比のノンパラメトリック推定値を従属変数とした線形モデルに補正項を加えて構造母数を推定

1. 割引因子 β は 1 に極めて近い
2. 尺度母数 σ は 20 万円程度
3. 超過累積修得単位数・GPA 水準が在学の効用に直接的に影響する証拠はない

⇒ $\beta \in [.95, 1]$, $\sigma \in [20, 100]$ の範囲で効用関数を特定化しない反実仮想分析を行う

研究の動機と貢献

大学中退モデル

構造推定

実証分析（手順のみ）

反実仮想分析

効用関数の一時的な変化

記号の定義

- $\Delta_t(.;.)$: 効用関数の変化, ただし $t > T$ では $\Delta_t(.;.) = 0$
- $u_t^*(.;.) := u_t(.;.) + \Delta_t(.;.)$
- $v_t^*(.;.)$: 反実仮想の選択肢別価値関数
- $p_t^*(.|.)$: 反実仮想 CCP
- $\bar{v}_t^{j*}(.)$: 反実仮想の積分した価値関数

時点 $T+1$ の問題は変化しないので, $j = 0, 1$ について

$$\bar{v}_{T+1}^{j*}(.) = \bar{v}_{T+1}^j(.)$$

時点 T では $j = 0, 1$ について

$$v_T^*(x_T; j) = v_T(x_T; j) + \Delta_T(j; x_T)$$

在学／退学の価値の差は

$$v_T^*(x_T; 1) - v_T^*(x_T; 0) = v_T(x_T; 1) - v_T(x_T; 0) + \Delta_T(1; x_T) - \Delta_T(0; x_T)$$

在学／退学の対数オッズ比は

$$\ln \frac{p_T^*(1|x_T)}{p_T^*(0|x_T)} = \ln \frac{p_T(1|x_T)}{p_T(0|x_T)} + \frac{\Delta_T(1; x_T) - \Delta_T(0; x_T)}{\sigma}$$

σ が既知なら $p_T(\cdot|\cdot)$ と $\Delta_T(\cdot;\cdot)$ のみから $p_T^*(\cdot|\cdot)$ は求まる。

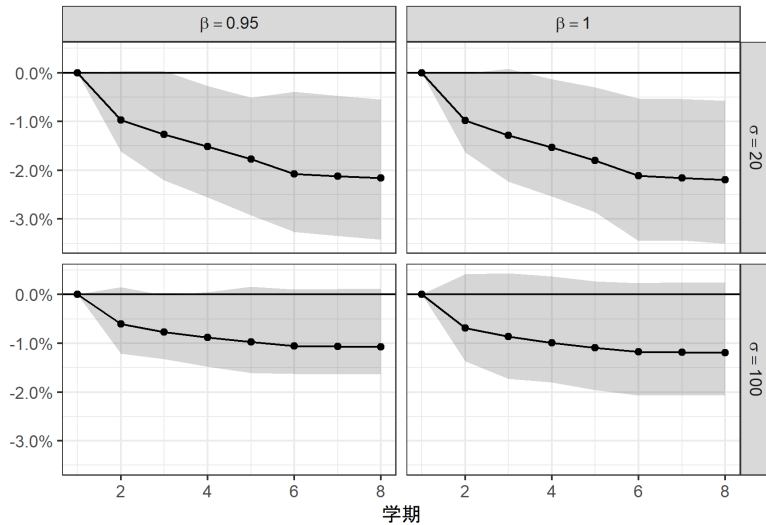
(β, σ) が既知なら $p_T^*(\cdot|\cdot)$ を所与として $\{p_t^*(\cdot|\cdot)\}$ は逆順で逐次的に求まる（詳細は略）。

構造母数の設定

- ・ 割引因子 $\beta = .95, 1$
- ・ 尺度母数 $\sigma = 20, 100$ (万円)
- ・ 効用関数の特定化は不要

反実仮想：在学中 4 年間の**毎学期 10 万円**の学資補助

累積中退確率を引き下げる処置効果



学資補助の処置効果

1. 毎学期 10 万円の学費補助は，某学部 2016 年 4 月入学者の男子の 4 年間の累積中退確率を約 2.2%引き下げる．
特に 2 年次までの中退はほぼ 0 になる．
2. $\beta \in [.95, 1]$, $\sigma \in [20, 100]$ の範囲では，累積中退確率は 1.0～2.2%低下．
3. 学費補助は成績に直接的に影響しない．
⇒累積中退確率の低下は卒業確率の上昇を意味せず，
退学的意思決定を遅らせる効果と解釈すべき．
4. もともと卒業できる学生が最大の利益を得る．大学の利益にはならない．

鹿野繁樹・高木真吾・村澤康友 2011.「経済学の成績に対する数学学習の効果：コントロール関数アプローチによる推定と予備検定」,『統計数理』, 59, 301-319.