第12章
質的選択モデル

藪友良

『入門　実践する計量経済学』

(東洋経済新報社)

PPT

- 質的選択モデルとは何か

- 潜在変数を用いた質的選択モデル

- 最尤法

- 当てはまりの良さ

- どのモデルを推定に用いるべきか

質的選択モデル

- **質的選択モデル:　被説明変数が質的情報を表す変数**

- **2値選択(0か1で記録する)**

　　例: タバコを吸うなら1、さもないと0となる

　　例:殺虫剤で虫が死んだら1、さもないと0となる

　　例:ローン審査が通ったら1、さもないと0となる

線形確率モデル

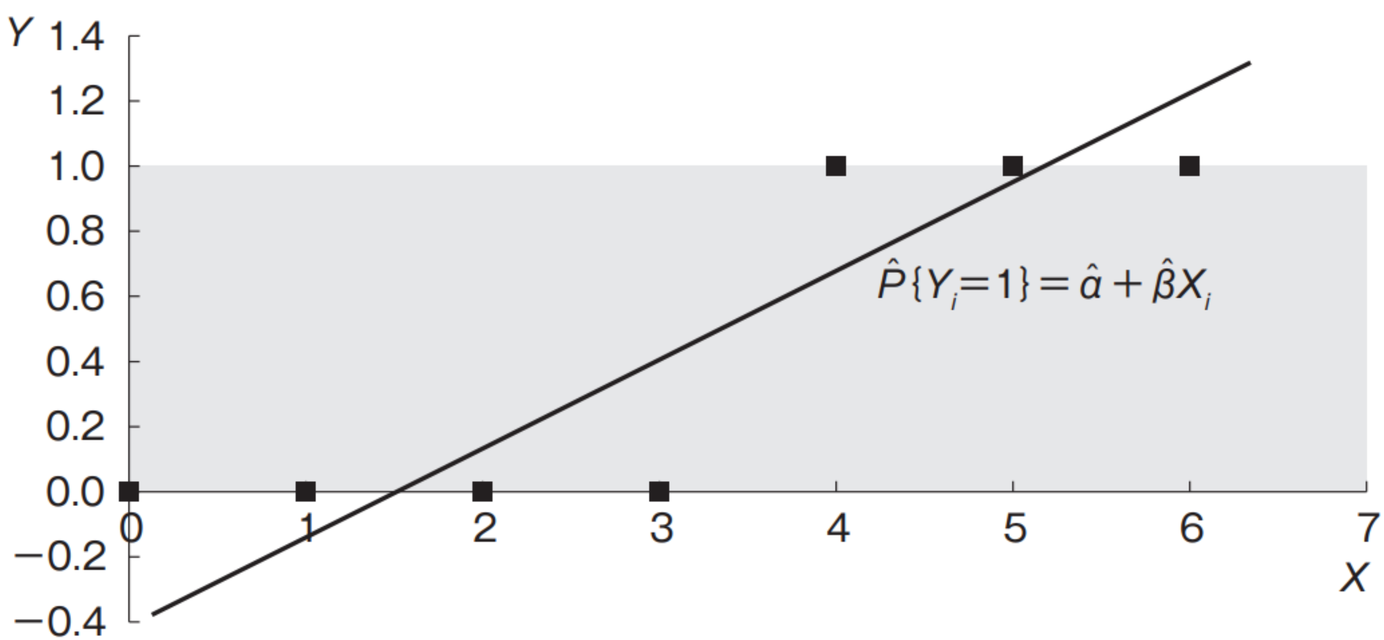
$$Y_i = \alpha + \beta X_i + u_i$$

- 被説明変数はダミー変数
- $Y_i = 1$となる確率はX_iの線形関数となる

$$P\{Y_i = 1\} = \alpha + \beta X_i$$

- βは**限界効果**：Xが1単位増えたら、確率はβだけ変化する
- 確率は0を下回ったり、1を上回ったりする

図12 - 1　線形確率モデルにおける確率$\hat{P}\{Y_i = 1\}$



例) 結婚の決定要因

- 国際成人力調査
  ---16歳から50歳までの男性データ
  ---被説明変数Y_iは、*i*さんが結婚をしていたら1となる
  --- 説明変数X_iは*i*さんの年齢

- 線形確率モデル

$$P\{\widehat{Y_i = 1}\} = -0.531 + 0.0318X_i$$

--- 1年歳をとると、3.18%結婚確率があがる

--- 25歳男性の結婚確率

$$0.264 = -0.531 + 0.0318 \times 25$$

--- 50歳男性の結婚確率

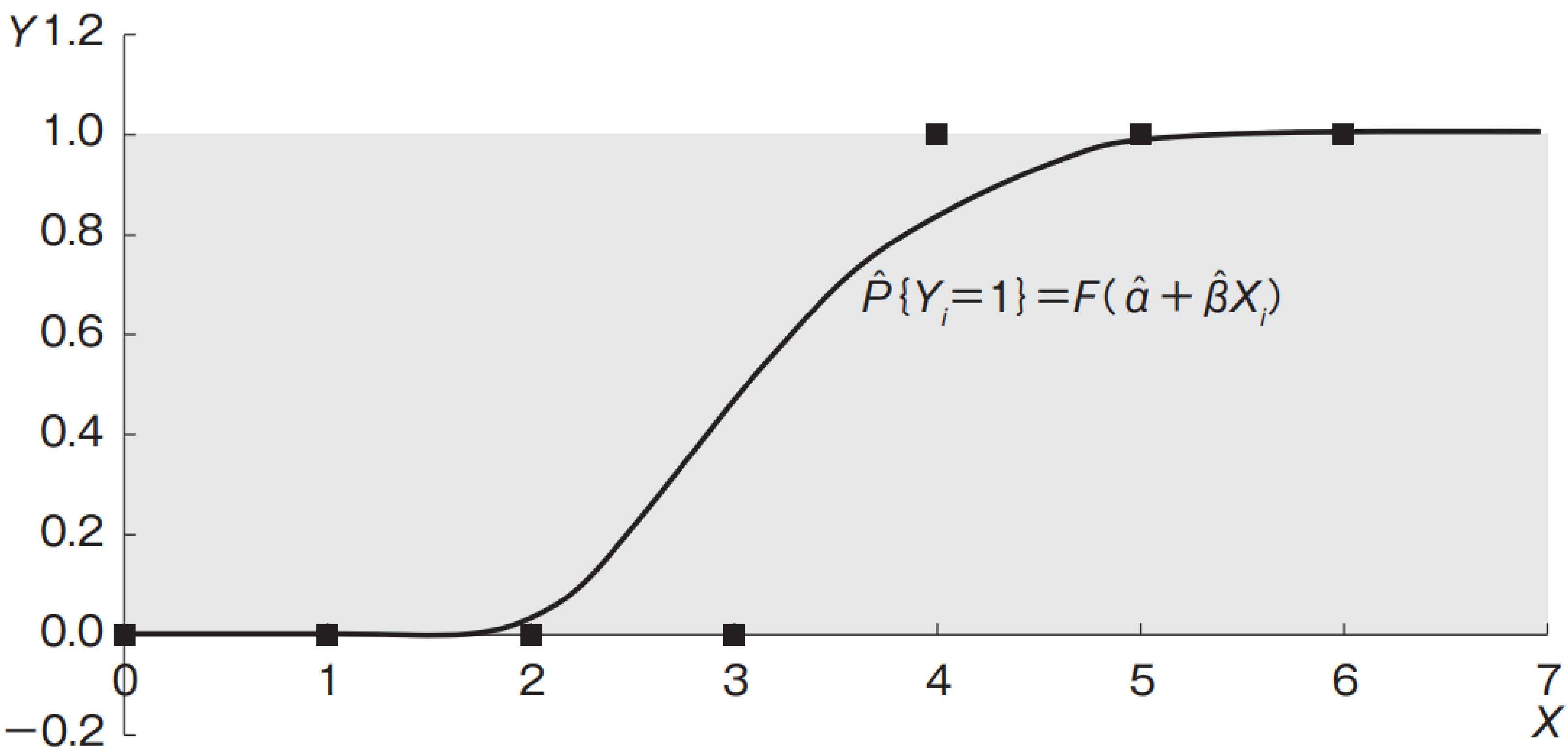
$$1.059 = -0.531 + 0.0318 \times 50$$

プロビットモデルとロジットモデル

$$P\{Y_i = 1\} = F(\alpha + \beta X_i)$$

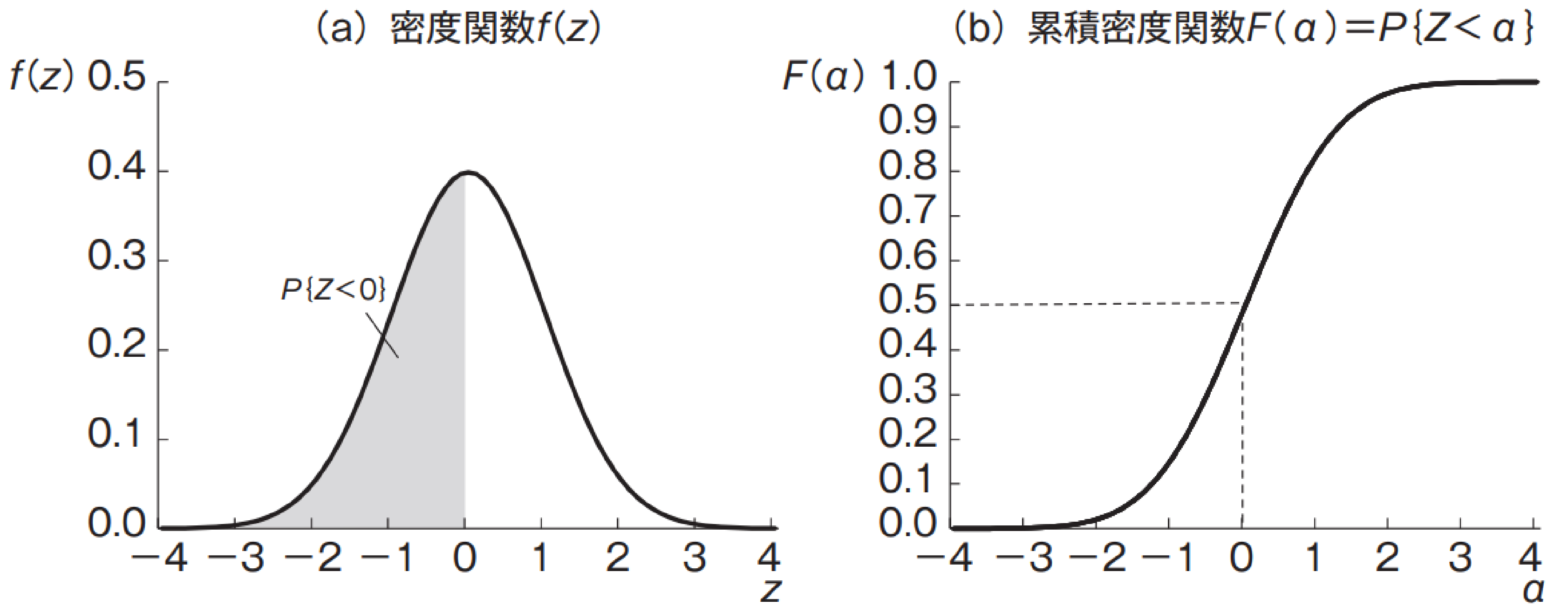
- $F(\cdot)$は0以上1以下の関数とする

図12 - 2　関数 $F(\cdot)$を用いたときの確率 $\hat{P}\{Y_i = 1\}$



- $F(\cdot)$：標準正規分布の累積密度関数なら
  プロビットモデルと呼ばれる

図12 - 3　標準正規分布



- $F(\cdot)$: ロジスティック分布の累積密度関数なら
  ロジットモデルと呼ばれる

例)　結婚の決定要因

- 16歳から50歳までの男性のデータ
  ---- Y_iは、iさんが結婚をしていたら1となる
  --- 説明変数X_iはiさんの年齢

・プロビットモデル

$$P\{\widehat{Y_i = 1}\} = F(-3.492 + \ 0.1066X_i)$$

--- 50歳男性の結婚確率

$$0.967 = F(-3.492 + \ 0.1066 \times 50)$$
$$= F(1.838)$$

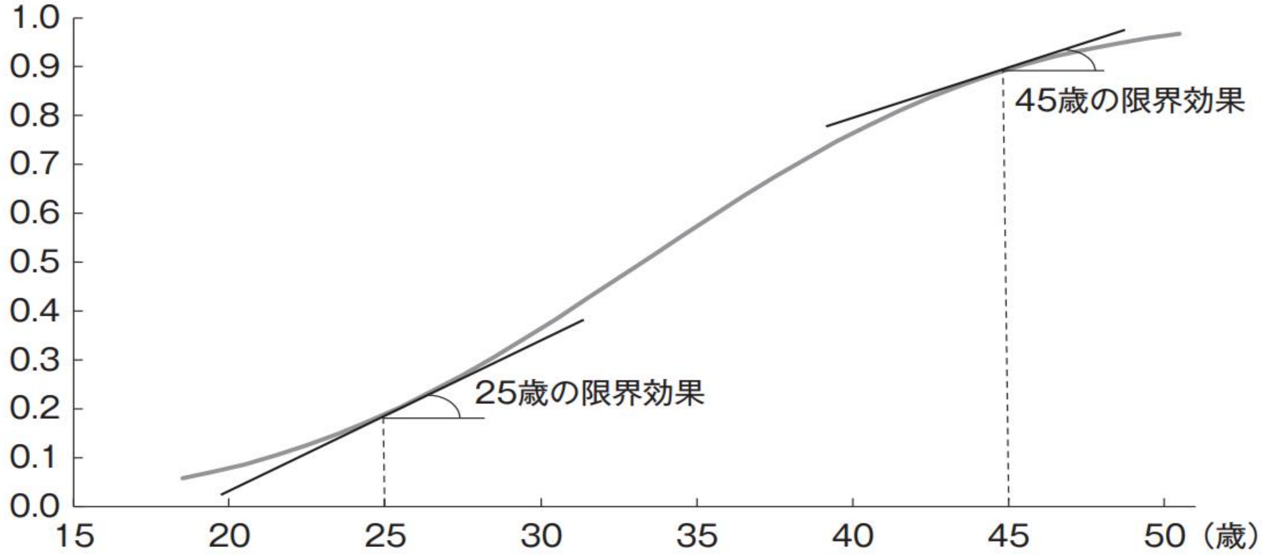
Excelなら
=normsdist(1.838)

プロビットの限界効果

$$\frac{\partial F(\hat{\alpha} + \hat{\beta}X_i)}{\partial X_i}$$

- 限界効果は、Xの値に依存して変わる

図12 - 4　プロビットモデルにおける限界効果



- 平均限界効果：　限界効果の平均値

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial F(\hat{\alpha} + \hat{\beta}X_i)}{\partial X_i}$$

例： 結婚の決定要因

- 国際成人力調査

  ---16歳から50歳までのデータ

  ---被説明変数Y_iは、iさんが結婚をしていたら1となる

  --- 説明変数X_iはiさんの年齢、

  大卒ダミー、大学院卒ダミー

  --- 男女でわけて線形確率モデルとプロビットで推定する

表12 - 1　男女別での結婚の決定要因

	(a) 男性のデータ			(b) 女性のデータ		
	線形確率	プロビット		線形確率	プロビット	
	(1)係数	(2)係数	(3)限界効果	(4)係数	(5)係数	(6)限界効果
年齢	0.031 ***	0.105 ***	0.026 ***	0.028 ***	0.088 ***	0.025 ***
	(0.001)	(0.005)	(0.0005)	(0.001)	(0.004)	(0.001)
大卒	0.063 ***	0.256 ***	0.063 ***	0.012	0.080	0.023
	(0.024)	(0.088)	(0.022)	(0.025)	(0.084)	(0.024)
院卒	0.161 ***	0.663 ***	0.163 ***	0.129	0.416	0.117
	(0.039)	(0.167)	(0.041)	(0.087)	(0.306)	(0.086)
定数項	−0.526 ***	−3.527 ***		−0.385 ***	−2.793 ***	
	(0.025)	(0.159)		(0.027)	(0.127)	
$\bar{R}^2$/*Psuedo* R^2	0.426	0.365		0.317	0.260	
n	1491	1491		1765	1765	

潜在変数を用いた質的選択モデル

・モデルの構造

$$Y_i = \begin{cases} 1 & if \quad Y_i^* > c \\ 0 & if \quad Y_i^* \leq c \end{cases}$$

$$Y_i^* = \alpha^* + \beta^* X_i + u_i^*$$

$$u_i^* \sim i.i.d.N(0, \sigma^2)$$

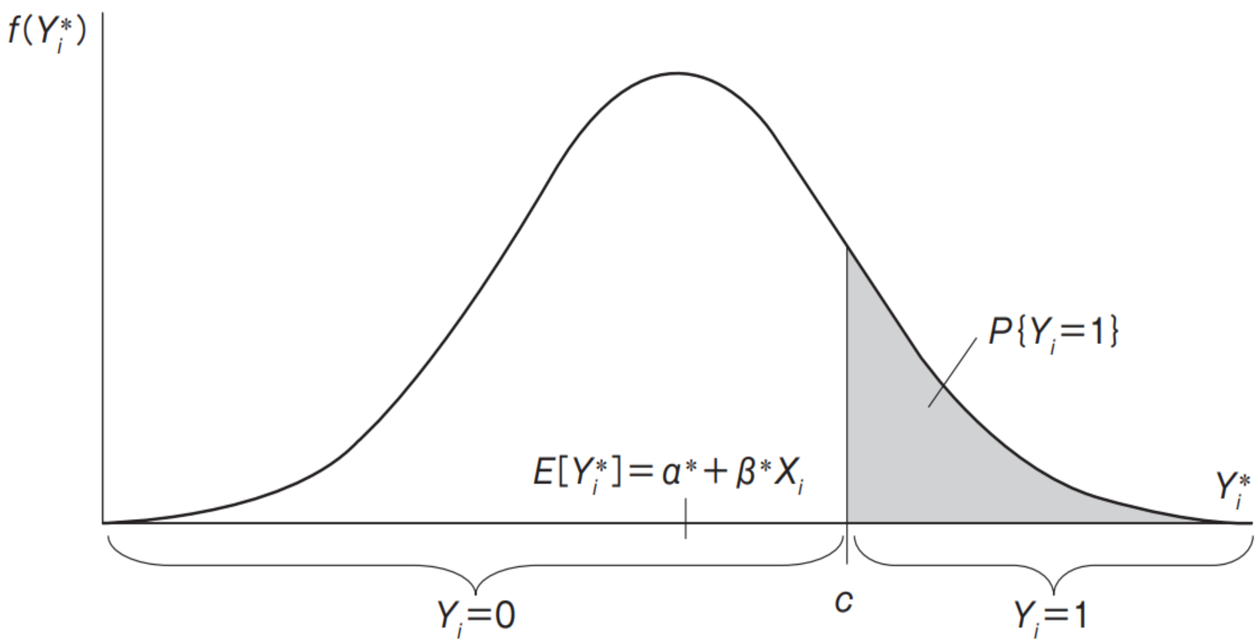
--- Y_i^*は潜在変数であり、観察はできない

--- 誤差項u_i^*は標準的仮定を満たしている

--- $Y_i = 1$となる確率は、[0,1]の範囲内にとどまる

例）$Y_i = 1$は虫が死亡、殺虫剤投与量X_i、吸収した毒物量Y_i^*

図12 - 5　被説明変数 Y_iと潜在変数 Y_i^*との関係



出所：いらすとや
https://www.irasutoya.com/2014/09/blog-post_68.html

例：通学方法の選択モデル

出所：いらすとや
https://www.irasutoya.com/2016/12/blog-post_22.html

- i君が徒歩通学するか、自転車通学するかという選択問題を考える
- 変数の定義
  - Y_i：i君が徒歩通学なら1、自転車通学なら0
  - X_i：i君の家から学校までの距離
- i君が各選択肢から得られる効用

$$Y_{1i}^* = \alpha_1^* + \beta_1^* X_i + u_{1i}^* \quad \text{（徒歩通学の効用）}$$
$$Y_{0i}^* = \alpha_0^* + \beta_0^* X_i + u_{0i}^* \quad \text{（自転車通学の効用）}$$

- i君は効用の大きい通学方法を選ぶ

$$Y_i = \begin{cases} 1 & if \quad Y_{1i}^* - Y_{0i}^* > 0 \\ 0 & if \quad Y_{1i}^* - Y_{0i}^* \leq 0 \end{cases}$$

- これはまさにプロビット

$$Y_i = \begin{cases} 1 & if \quad Y_i^* > 0 \\ 0 & if \quad Y_i^* \leq 0 \end{cases}$$

--- ただし、Y_i^*は効用の差とする

$$\begin{aligned} Y_i^* &= Y_{1i}^* - Y_{0i}^* \\ &= (\alpha_1^* + \beta_1^* X_i + u_{1i}^*) - (\alpha_0^* + \beta_0^* X_i + u_{0i}^*) \\ &= (\alpha_1^* - \alpha_0^*) + (\beta_1^* - \beta_0^*) X_i + (u_{1i}^* - u_{0i}^*) \\ &= \alpha^* + \beta^* X_i + u_i^* \end{aligned}$$

最尤法

- **最尤法**:　尤度を最大にするようにパラメータを決める

- **尤度**:　モデルから手元のデータを得られる確率

--- 被説明変数$(Y_1, Y_2, \ldots, Y_n)$の実現値は$(y_1, y_2, \ldots, y_n)$

--- パラメータをθとしたとき、尤度は

$$L(\theta) = P\{Y_1 = y_1, Y_2 = y_2, \ldots, Y_n = y_n\}$$

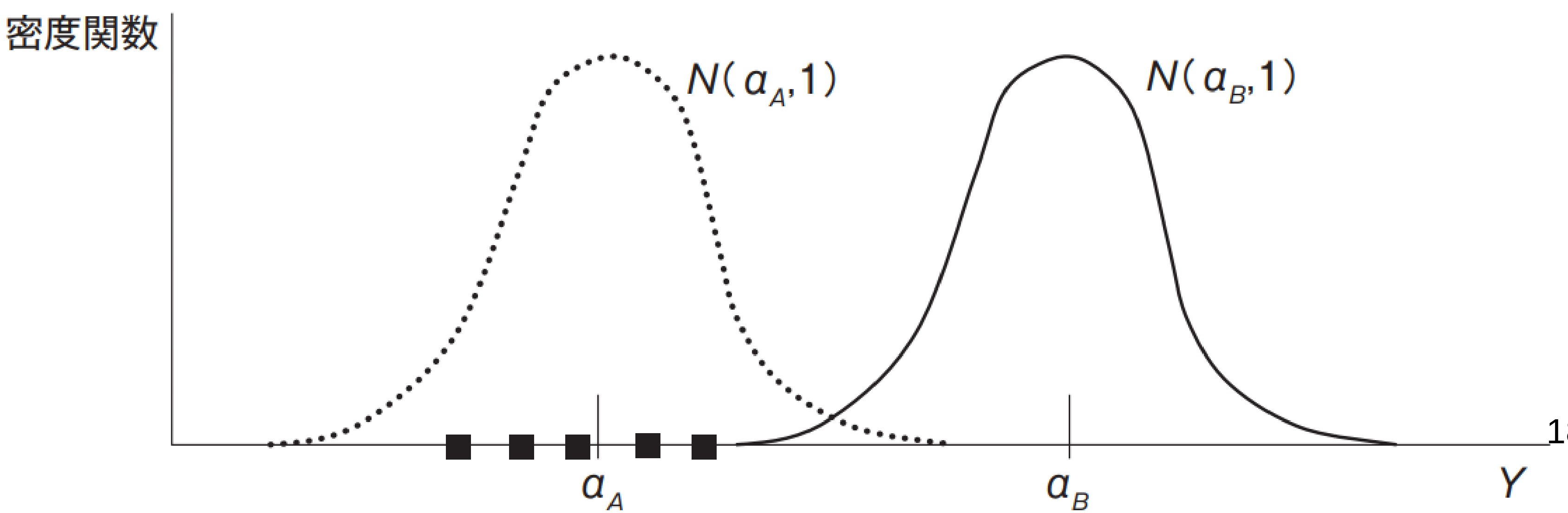
--- 手元のデータが得られたということは、それが得られやすいモデル（パラメータ）が、最も尤もらしい

単純な例)

確率変数Y_iは正規分布$N(\alpha, 1)$に従うとし、αを推定したい

- パラメータはαであり、α_Aもしくはα_Bのどちらか $(\alpha_A < \alpha_B)$
- 観測値は5つ($Y_1 = y_1, Y_2 = y_2, Y_3 = y_3, Y_4 = y_4, Y_5 = y_5$)
- α_Aで評価した同時確率(尤度)のほうが、α_Bで評価した尤度より高い
  $f(Y_1 = y_1, \dots, Y_5 = y_5 | \alpha = \alpha_A) > f(Y_1 = y_1, \dots, Y_5 = y_5 | \alpha = \alpha_B)$
- 最尤推定量は$\hat{\alpha}_{ML} = \alpha_A$

図12 - 6　2 つの正規分布



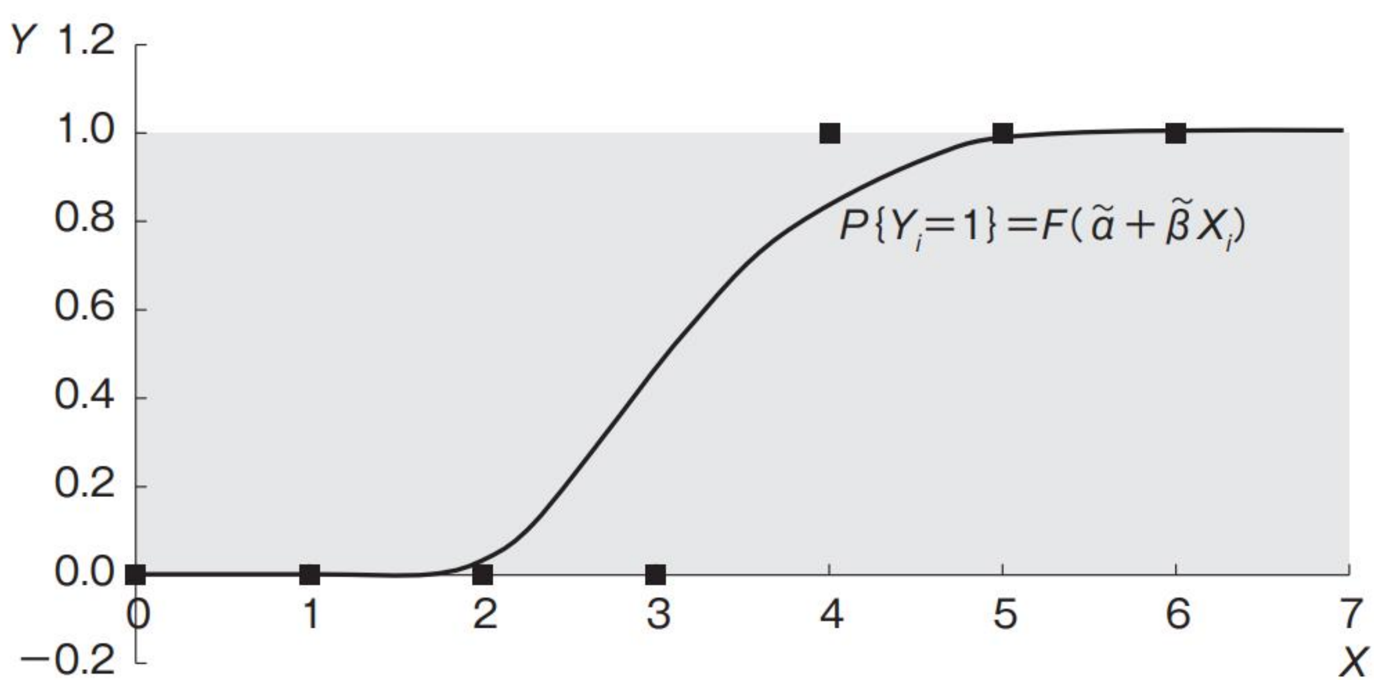
・プロビットの尤度

$$L = P\{Y_1 = y_1, Y_2 = y_2, \dots, Y_n = y_n\}$$
$$= P\{Y_1 = y_1\} \times P\{Y_2 = y_2\} \times \cdots \times P\{Y_n = y_n\}$$

ただし、

$$P\{Y_i = y_i\} = \begin{cases} F(\tilde{\alpha} + \tilde{\beta}X_i) & if \quad y_i = 1 \\ 1 - F(\tilde{\alpha} + \tilde{\beta}X_i) & if \quad y_i = 0 \end{cases}$$

図12 - 7　実現値と確率 $P\{Y_i = 1\}$



・プロビットの尤度

$$L = P\{Y_1 = y_1, Y_2 = y_2, \dots, Y_n = y_n\}$$
$$= P\{Y_1 = y_1\} \times P\{Y_2 = y_2\} \times \cdots \times P\{Y_n = y_n\}$$

ただし、

$$P\{Y_i = y_i\} = \begin{cases} F(\tilde{\alpha} + \tilde{\beta} X_i) & if \quad y_i = 1 \\ 1 - F(\tilde{\alpha} + \tilde{\beta} X_i) & if \quad y_i = 0 \end{cases}$$

---最尤推定量(Maximum Likelihood estimator)

尤度を最大にするように選ばれたパラメータ

$\hat{\alpha}_{ML}$、$\hat{\beta}_{ML}$

--- 最大尤度P^{max}(最尤推定量で評価した尤度)

当てはまりの尺度

当てはまりの尺度

- 正しく予想した割合

--- $Y_i = 1$のとき、$\hat{P}\{Y_i = 1\} = F(\hat{\alpha}_{ML} + \hat{\beta}_{ML} X_i)$が閾値0.5を超えれば正しく予測

--- $Y_i = 0$のとき、$\hat{P}\{Y_i = 0\} = 1 - F(\hat{\alpha}_{ML} + \hat{\beta}_{ML} X_i)$が閾値0.5を超えれば正しく予測

--- 問題は、予測精度の違いを考慮できないこと

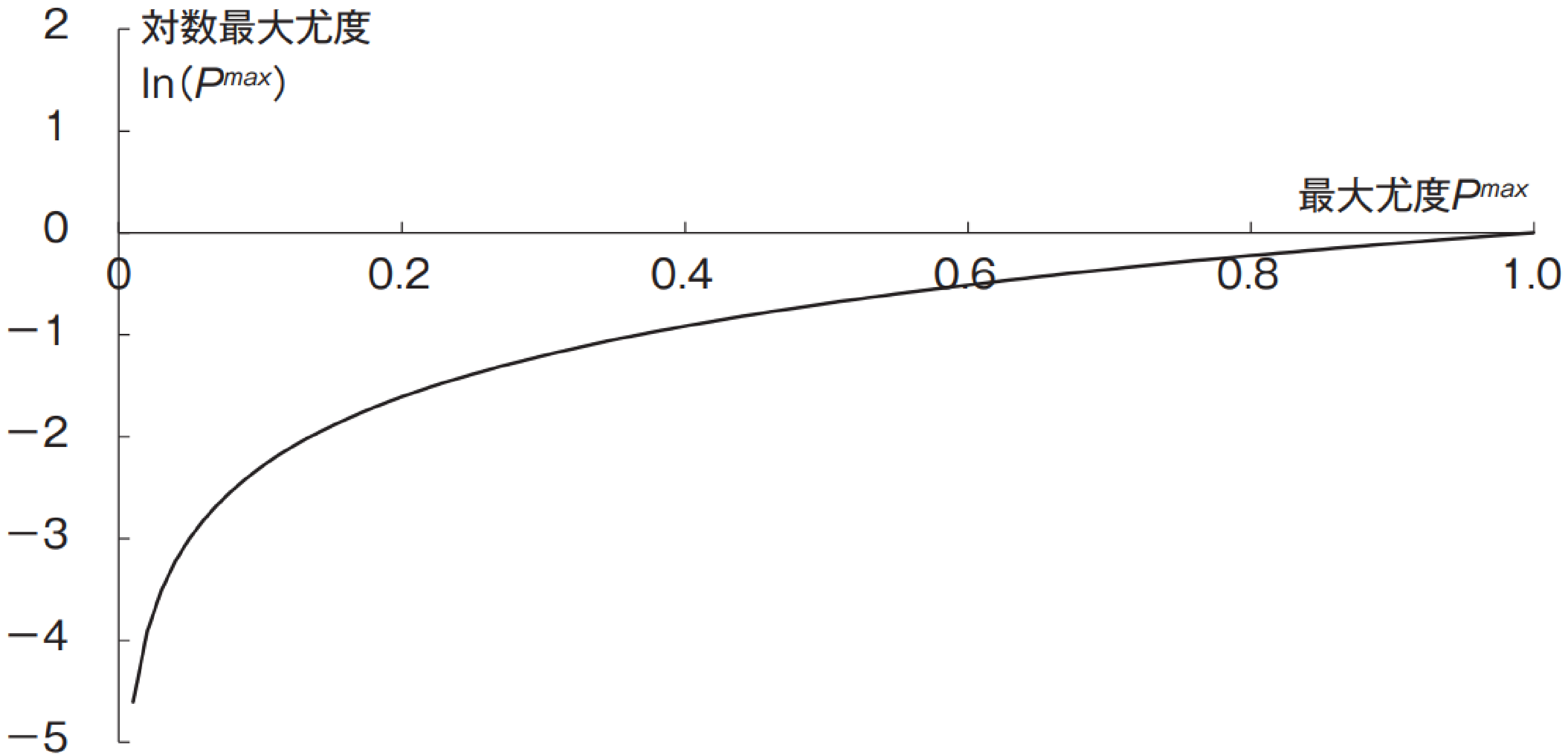
例) $\hat{P}\{Y_i = 1\}$が0.501でも0.999でも1と予測しているが、予測の確度は全く異なる

--- 分析者によっては閾値として異なる値を用いる

例） 大地震はまれにしか生じないが、被害は大きいため、予測値が0.5より低くても警戒すべき

- 対数最大尤度： 最大尤度の対数$\ln(P^{max})$
---対数尤度$\ln(P^{max})$は負の値をとり、0に近いほど良いモデル

図12 - 8　最大尤度と対数最大尤度との関係



・ **疑似決定係数(Pseudo-R²)**

$$\mathrm{Pseudo-R^2} = 1 - \frac{\ln(P^{max})}{\ln(P_0^{max})}$$

--- P_0^{max}: $\beta = 0$という制約(説明変数を含めない)を課したうえ
で求めた最大尤度

--- $\mathrm{Pseudo-R^2}$ は0以上1以下となる

[証明]

説明変数を増やすと、当てはまりは改善するため、

$$P_0^{max} \le P^{max} \le 1$$

対数をとると、

$$\ln(P_0^{max}) \le \ln(P^{max}) \le 0$$

両辺を$\ln(P_0^{max})$で割ると、

$$1 \ge \frac{\ln(P^{max})}{\ln(P_0^{max})} \ge 0$$

--- $\beta = 0$なら、$\ln(P^{max})/\ln(P_0^{max}) \approx 1$であり、$\mathrm{Pseudo-R^2} \approx 0$

--- $\beta \ne 0$なら、$\ln(P^{max})/\ln(P_0^{max})$は小さくなり、$\mathrm{Pseudo-R^2}$は1に近くなる

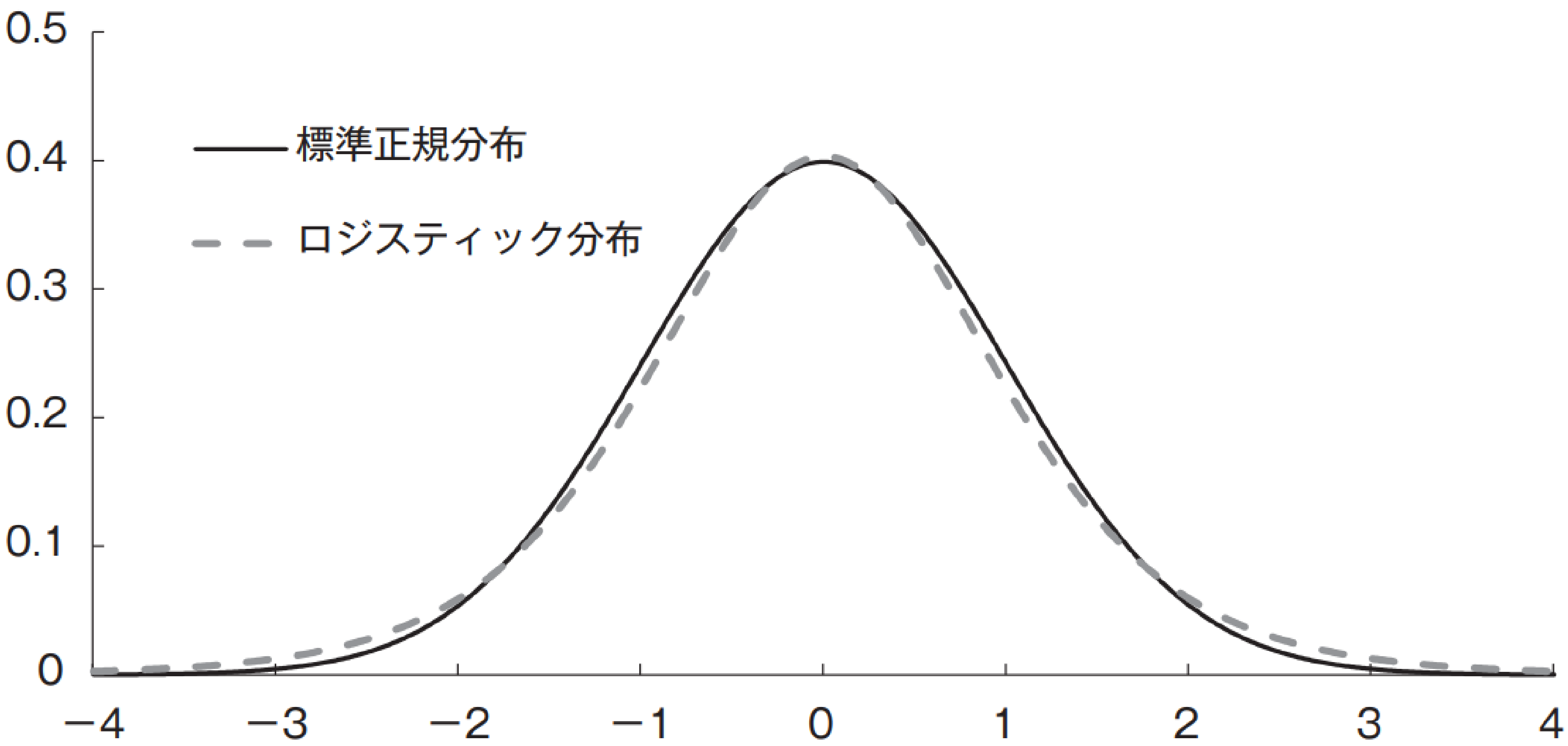
ロジットモデル

$$Y_i = \begin{cases} 1 & if \quad Y_i^* > c \\ 0 & if \quad Y_i^* \le c \end{cases}$$

$$Y_i^* = \alpha^* + \beta^* X_i + u_i^* 、\boxed{u_i^* \sim i.i.d. \text{ロジスティック分布}}$$

- ロジスティック分布と標準正規分布の違いは小さい

図12 - 9　標準正規分布とロジスティック分布



- $\boldsymbol{Y_i = 1}$**の確率**

$$P\{Y_i = 1\} = F(\alpha + \beta X_i)$$
$$= \frac{e^{\alpha+\beta X_i}}{1 + e^{\alpha+\beta X_i}}$$

ロジスティック分布の
累積密度関数
$F(a) = P(Z < a)$
$= \dfrac{e^a}{1 + e^a}$

- $\boldsymbol{Y_i = 0}$**の確率**

$$P\{Y_i = 0\} = 1 - F(\alpha + \beta X_i)$$
$$= 1 - \frac{e^{\alpha+\beta X_i}}{1 + e^{\alpha+\beta X_i}}$$
$$= \frac{1 + e^{\alpha+\beta X_i}}{1 + e^{\alpha+\beta X_i}} - \frac{e^{\alpha+\beta X_i}}{1 + e^{\alpha+\beta X_i}}$$
$$= \frac{1}{1 + e^{\alpha+\beta X_i}}$$

・**オッズ比**：ある事象が生じる確率（$P\{Y_i = 1\}$）と生じない確率（$P\{Y_i = 0\}$）との比

$$\frac{P\{Y_i = 1\}}{P\{Y_i = 0\}} = \frac{P\{Y_i = 1\}}{1 - P\{Y_i = 1\}}$$

--- オッズ比は0から∞の値をとる

・ロジットモデルは、オッズ比の対数を被説明変数とした線形モデルとしても解釈できる

[証明]

$$\frac{P\{Y_i = 1\}}{P\{Y_i = 0\}} = \frac{e^{\alpha+\beta X_i}/(1 + e^{\alpha+\beta X_i})}{1/(1 + e^{\alpha+\beta X_i})} = e^{\alpha+\beta X_i}$$

$$\Rightarrow \qquad \ln\left(\frac{P\{Y_i=1\}}{P\{Y_i=0\}}\right) = \alpha + \beta X_i$$

--- β：Xが1単位変化したとき、オッズ比が何%変化するか

--- e^{β}：Xが1単位増加すると、オッズ比は

$e^{\beta}\left(= \dfrac{e^{\alpha+\beta(X+1)}}{e^{\alpha+\beta X}}\right)$倍変化する

労働参加の決定要因

- 国際成人力調査
  - --- 50歳以下のデータ
  - --- 被説明変数Y_iは、iさんが仕事をしていたら1となる
  - --- 説明変数は勤続年数、大卒ダミー、大学院卒ダミー、既婚ダミー(結婚していたら1となる)、未就学児ダミー(6歳未満の子供がいたら1となる)、子供の数(子供の人数)

- 推定結果
  - --- 線形確率、プロビット、ロジットで推定する

表12 - 2　男性の労働参加の決定要因

	線形確率	プロビット		ロジット	
	(1)係数	(2)係数	(3)限界効果	(4)係数	(5)限界効果
勤続年数	0.002 *	0.050 **	0.002 *	0.127 *	0.002 *
	(0.001)	(0.025)	(0.001)	(0.067)	(0.001)
大卒	0.021 **	0.674 *	0.024 *	1.703	0.026
	(0.009)	(0.372)	(0.014)	(1.062)	(0.017)
院卒	0.007	0.050	0.002	0.185	0.003
	(0.025)	(0.470)	(0.017)	(1.144)	(0.017)
既婚	0.084	0.734 **	0.026 **	1.606 **	0.024 **
	(0.056)	(0.345)	(0.013)	(0.723)	(0.012)
未就学児	0.020	0.507	0.018	1.207	0.018
	(0.014)	(0.349)	(0.013)	(0.855)	(0.014)
子どもの数	0.002	0.117	0.004	0.134	0.002
	(0.006)	(0.186)	(0.007)	(0.460)	(0.007)
定数項	0.835 ***	−0.008		−0.702	
	(0.073)	(0.532)		(1.118)	
$\bar{R}^2$/*Psuedo* R^2	0.028	0.143		0.147	
n	682	682		682	

--- 勤続年数は正で有意となる

--- 既婚ダミーは正で有意となる

--- 未就学児ダミーや子供の人数は有意ではない

--- 限界効果は、どのモデルもほぼ同じ値となる

表12 - 3　女性の労働参加の決定要因

	線形確率	プロビット		ロジット	
	(1) 係数	(2) 係数	(3) 限界効果	(4) 係数	(5) 限界効果
勤続年数	0.021 ***	0.073 ***	0.022 ***	0.126 ***	0.023 ***
	(0.002)	(0.002)	(0.002)	(0.014)	(0.002)
大卒	0.026	0.086	0.026	0.152	0.028
	(0.037)	(0.037)	(0.036)	(0.194)	(0.035)
院卒	−0.027	−0.077	−0.024	−0.123	−0.022
	(0.153)	(0.153)	(0.131)	(0.707)	(0.128)
既婚	−0.133 ***	−0.452 ***	−0.138 ***	−0.822 ***	−0.149 ***
	(0.040)	(0.040)	(0.049)	(0.294)	(0.052)
未就学児	−0.199 ***	−0.569 ***	−0.174 ***	−0.930 ***	−0.169 ***
	(0.034)	(0.034)	(0.027)	(0.155)	(0.027)
子どもの数	0.034 *	0.109 *	0.033 *	0.178 *	0.032 *
	(0.018)	(0.018)	(0.018)	(0.096)	(0.017)
定数項	0.489 ***	−0.106		−0.161	
	(0.062)	(0.062)		(0.373)	
$\bar{R}^2$/*Psuedo* R^2	0.196	0.174		0.174	
n	976	976		976	

--- 勤続年数は正で有意となる

--- 既婚ダミーは負で有意となる

--- 未就学児ダミーは負で有意、子供の人数は正で有意

--- 限界効果は、どのモデルもほぼ同じ値となる

どのモデルを使うべきか

①線形確率モデル

--- 確率が0以上1以下とならないが、係数が限界効果を表し、推定結果が分かりやすい

--- 線形モデルなので、これまで学習した手法を使うことができる(11章の固定効果モデル、13章の2段階最小2乗法)

②プロビットとロジットの限界効果はほぼ同じとなる

--- ロジットの方が計算が短時間というメリットがあった

--- プロビットは正規分布に基づいており、正当化しやすい

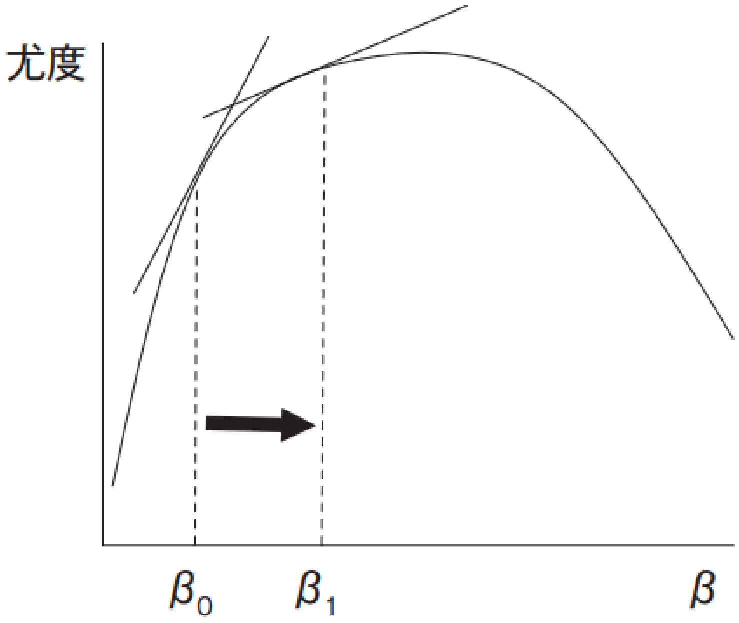
③論文には、線形確率とプロビット（or ロジット）の結果を掲載しよう

--- オッズ比に関心があるならば、ロジットを使う

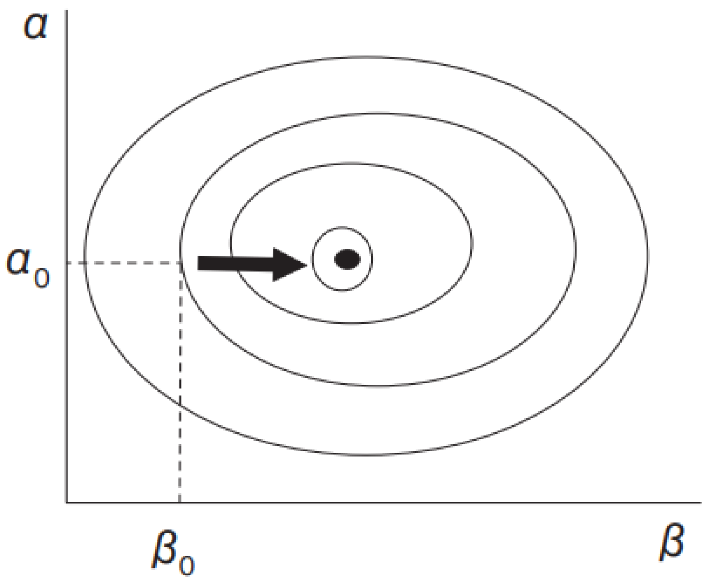
--- オッズ比に関心がないならば、プロビットを使う

図12 - 10　ニュートン法のイメージ

（a）パラメータが１つの場合



（b）パラメータが２つの場合



まとめ

- 質的選択モデル
  - ---線形確率モデル、
  - ---プロビットモデル、ロジットモデル
- 潜在変数を用いた質的選択モデル
- 最尤法
- 当てはまりの良さ
  - --- 正しく予想した割合
  - --- 対数最大尤度、疑似決定係数