第 14 回：2 値応答モデル
【教科書第 7 章第 3 節】

北村　友宏

2026 年 1 月 13 日

本日の内容

1. 線形確率モデル
2. 2 値プロビット・モデル
3. 限界効果
4. 実証分析例

ダミー変数を被説明変数とするモデル

- 大きさ n の 2 変量無作為標本 $((y_1, x_1), (y_2, x_2), \cdots, (y_n, x_n))$ を用いて，y_i を x_i に回帰することを考える．
- ただし，y_i は 0 または 1 の値をとるダミー変数．例えば，
  - （個人が）働くなら 1，働かないなら 0.
  - （個人が）チームを移籍するなら 1，しないなら 0.
  - （企業が）市場に参入するなら 1，しないなら 0.

線形確率モデル

y_i をダミー変数とすると，線形確率モデル（linear probability model）は，

$$
\begin{aligned}
y_i &= \beta_0 + \beta_1 x_i + u_i, \\
E(u_i \mid x_i) &= 0.
\end{aligned}
$$

$$\Downarrow$$

被説明変数がダミー変数の場合に線形回帰モデルを仮定すると，線形確率モデルとなる．

線形確率モデルの問題点

- 被説明変数の値が 1 になる確率を予測すると，0 を下回ったり 1 を上回ったりする．
- 誤差項に不均一分散が発生する．

条件付き期待値と予測値

線形確率モデルの OLS 推定量 $\hat{\beta}_0$ と $\hat{\beta}_1$ を元の式に代入し，誤差項 u_i を除くと，

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i,$$

という式で被説明変数 y_i の値を予測できる．

- $\hat{y}_i$ は y_i の予測値．
- $\hat{y}_i$ は「x_i がこの値のときに y_i はどのような値になる傾向があるか」を表す．

$\Rightarrow \hat{y}_i$ は，x_i を所与とした y_i の条件付き期待値

$$E(y_i \mid x_i) = \beta_0 + \beta_1 x_i,$$

を予測したものと解釈できる．

線形確率モデルにおける条件付き期待値

y_i がダミー変数なら，

$$
\begin{aligned}
E(y_i \mid x_i) &= 0 \cdot \Pr(y_i = 0 \mid x_i) + 1 \cdot \Pr(y_i = 1 \mid x_i) \\
&= \Pr(y_i = 1 \mid x_i).
\end{aligned}
$$

$$\Downarrow$$

「y_i の条件付き期待値」が「y_i の値が 1 になる条件付き確率」と同じになる．

➡ $\hat{y}_i$ を計算すると，「y_i の値が 1 になる条件付き確率」を予測していることになる．

➡ $\hat{y}_i$, すなわち「y_i の値が 1 になる条件付き確率の予測値」は 0 を下回ったり 1 を上回ったりする（問題）．

線形確率モデルの誤差項の分散

y_i がダミー変数なら，

$$V(u_i \mid x_i) = (\beta_0 + \beta_1 x_i)\left[1 - (\beta_0 + \beta_1 x_i)\right].$$

（証明は省略）

$$\Downarrow$$

誤差項 u_i の分散が説明変数 x_i に応じて変化する．

➥ 不均一分散発生（問題）．

- 仮説検定の際に，不均一分散に対して頑健な標準誤差を用いることで，ある程度対処可能．

$\Downarrow$

これらの問題を解決するには，2 値応答モデルを仮定する．

- 観測不可能な変数で，被説明変数の値を決定づける変数を潜在変数（latent variable）という．

2 値応答モデル（binary response model）は，

$$
y_i = \begin{cases} 1 & \text{if } y_i^* > 0, \\ 0 & \text{otherwise,} \end{cases}
$$

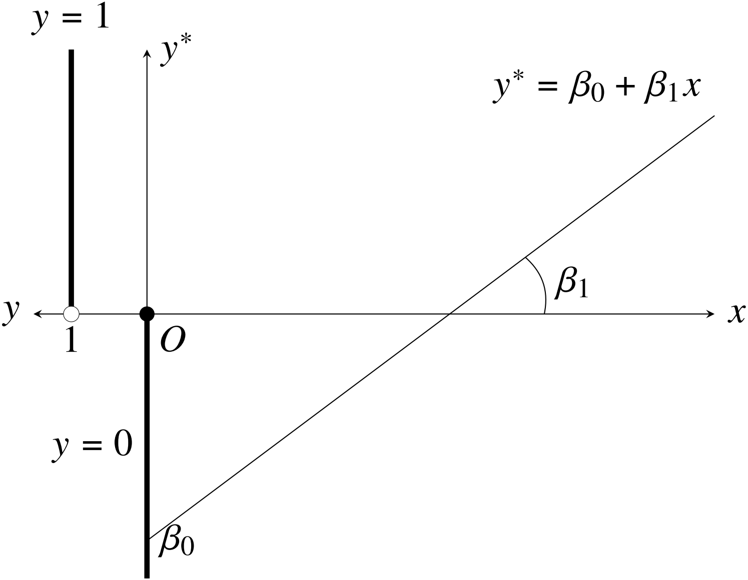
$$
y_i^* = \beta_0 + \beta_1 x_i + u_i,
$$

$$
u_i \mid x_i \sim F(.).
$$

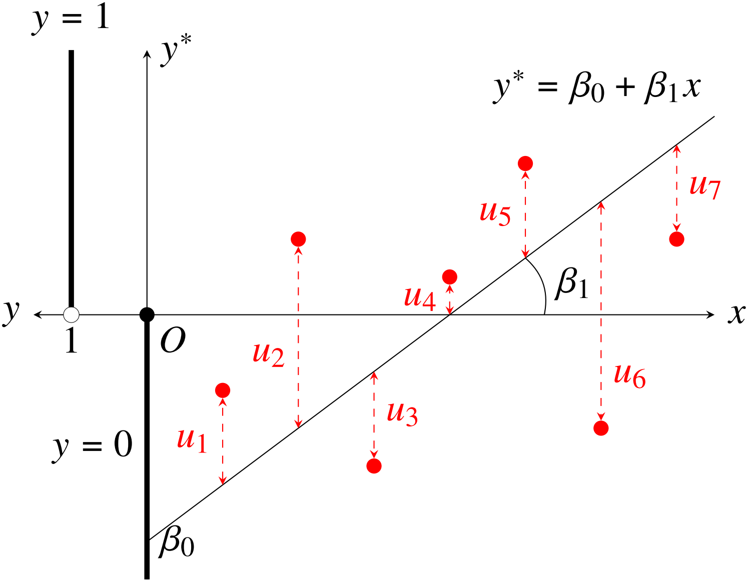
y_i^* は潜在変数で，y_i の値を決定づける．

- 観測可能：y_i, x_i
- 観測不可能：$y_i^*, \beta_0, \beta_1, u_i$
- 推定するもの：β_0, β_1

各変数，パラメータを図示すると？



$y = 1$
y^*
$y^* = \beta_0 + \beta_1 x$
β_1
y
x
1
O
$y = 0$
β_0



$y = 1$
y^*
$y^* = \beta_0 + \beta_1 x$
u_7
u_5
β_1
u_4
y
x
1
O
u_2
u_6
u_3
u_1
$y = 0$
β_0

- 誤差項 u_i には，説明変数 x_i を所与とした条件付き分布を仮定する．
    - e.g., 標準正規分布

- 誤差項の条件付き分布を標準正規分布と仮定した 2 値応答モデルを **2 値プロビット・モデル（binary probit model）** という．

※ “$y_i = \cdots$” の式ではなく “$y_i^* = \cdots$” の式の誤差項の分布を仮定している．

2 値プロビット・モデルの定式化

2 値プロビット・モデルは，

$$
\begin{aligned}
y_i &= \begin{cases} 1 & \text{if } y_i^* > 0, \\ 0 & \text{otherwise}, \end{cases} \\
y_i^* &= \beta_0 + \beta_1 x_i + u_i, \\
u_i \mid x_i &\sim N(0, 1).
\end{aligned}
$$

$$\Downarrow$$

最尤（maximum likelihood）法を用いて，β_0 と β_1 を推定する．
（導出方法は付録参照）

2 値プロビット・モデルの別の表現

2 値プロビット・モデルの別の表現は，

$$\Pr(y_i = 1 \mid x_i) = \Phi(\beta_0 + \beta_1 x_i).$$

$\Phi(.)$ は標準正規分布の累積分布関数.

2 値プロビット・モデルが

$$\Pr(y_i = 1 \mid x_i) = \Phi(\beta_0 + \beta_1 x_i),$$

と書ける理由は以下の通り．
x_i を所与として，$y_i = 1$ となる条件付き確率は，

$$\begin{aligned}\Pr(y_i = 1 \mid x_i) &= \Pr(y_i^* > 0 \mid x_i) \\ &= \Pr(\beta_0 + \beta_1 x_i + u_i > 0 \mid x_i) \\ &= \Pr\left(u_i > -(\beta_0 + \beta_1 x_i) \mid x_i\right).\end{aligned}$$

標準正規分布は 0 で対称な分布なので，

$$\Pr\left(u_i > -(\beta_0 + \beta_1 x_i) \mid x_i\right) = \Pr\left(u_i < \beta_0 + \beta_1 x_i \mid x_i\right).$$

よって，

$$\begin{aligned}\Pr(y_i = 1 \mid x_i) &= \Pr\left(u_i < \beta_0 + \beta_1 x_i \mid x_i\right) \\ &= \Phi(\beta_0 + \beta_1 x_i).\end{aligned}$$

限界効果

定数項以外に説明変数が 1 つの 2 値プロビット・モデルにおける，x_i の限界効果（marginal effect）は，

$$\begin{aligned}\frac{\partial \Pr(y_i = 1 \mid x_i)}{\partial x_i} &= \frac{\partial \Phi(\beta_0 + \beta_1 x_i)}{\partial x_i} \\ &= \phi(\beta_0 + \beta_1 x_i)\beta_1.\end{aligned}$$

$\phi(.)$ は標準正規分布の確率密度関数.

$$\Downarrow$$

β_1 そのものではなく $\phi(\beta_0 + \beta_1 x_i)\beta_1$ が，「x_i が 1 単位増加したときに $y_i = 1$ となる確率がどの程度変化する傾向があるか」を表す.

- $\phi(.)$ は確率密度関数なので 0 以上.
  $\Longrightarrow$ 限界効果の符号は β_1 の符号と同じ.
- 説明変数 x_i の値は各個体によって異なる.
  ➥ 限界効果

$$\frac{\partial P(y_i = 1 \mid x_i)}{\partial x_i} = \phi(\beta_0 + \beta_1 x_i)\beta_1,$$

  の値も各個体によって異なる.
  $\Longrightarrow$ 説明変数 x_i をその平均 $\bar{x}$ で置き換えた，平均における限界効果を計算する.

- 定数項以外に説明変数が 1 つの 2 値プロビット・モデルの，平均における限界効果（marginal effect at the mean）は，

$$\phi(\hat{\beta}_0 + \hat{\beta}_1 \bar{x})\hat{\beta}_1.$$

  - $\hat{\beta}_0, \hat{\beta}_1$ はそれぞれ β_0, β_1 を最尤法で推定した値（最尤推定値）.
  - $\bar{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i.$
  - 定数項以外に説明変数が複数個ある場合は，それらを全てそれぞれの標本平均で置き換える.

線形確率モデルを適用すべき場合

以下の場合は被説明変数 y_i がダミー変数であっても線形確率モデルを仮定して推定する．

- モデルの右辺に交差項（変数の積）を含み，その係数を解釈したい場合．
  - e.g., $y_i = \beta_0 + \beta_X x_i + \beta_Z z_i + \beta_{XZ} x_i z_i + u_i$ のような線形確率モデルを推定．
- パネルデータを用いて固定効果モデルを仮定したい場合．
  - e.g., $y_{it} = \beta_0 + \beta_1 x_{it} + \underbrace{\mu_i}_{\text{個別効果}} + \underbrace{\varepsilon_{it}}_{\text{その他効果}}$ のような線形確率モデルを推定．

gretl での限界効果推定

2 値プロビット・モデルの場合は，

- メニューバーから「モデル」→「制限従属変数」→「プロビット」→「二項 (Binary)」と操作．
- ラジオボタンの中から「平均での限界効果（slope at mean）を表示する」を選ぶ．

⇒ 出力結果に各説明変数の，「平均における限界効果」が表示される．

gretl では，平均における限界効果を出力すると p 値およびと有意性を示すアスタリスク記号が出力されない.

$$\Downarrow$$

メニューバーから「モデル」→「制限従属変数」→「プロビット」→「二項 (Binary)」と操作して出てくるウィンドウで,「p 値を表示する」を選んだ結果と「平均での限界効果（slope at mean）を表示する」を選んだ結果の両方を出力して，前者の結果から p 値とアスタリスク記号の個数を確認し，後者の結果から限界効果を確認するとよい.

gretl で 2 値プロビット・モデルを推定する場合の頑健標準誤差

- gretl で 2 値プロビット・モデルを推定する場合，「頑健標準誤差を使用する」にチェックを入れると，不均一分散に対してではなく，モデルの定式化に対して頑健な標準誤差が計算される．

実証分析例：女性の労働供給関数の推定

配偶者のいる女性 1053 人分のデータを用い，以下の 2 値プロビット・モデルを推定し，各説明変数の限界効果を推定する．

$$
\begin{aligned}
work_i &= \begin{cases} 1 & \text{if } work_i^* > 0, \\ 0 & \text{otherwise,} \end{cases} \\
work_i^* &= \beta_0 + \beta_1 income\_s_i \\
&\quad + \beta_2 childu6_i + u_i, \\
u_i \mid income\_s_i, childu6_i &\sim N(0, 1).
\end{aligned}
$$

この 2 値プロビット・モデルの別の表現は，

$$
\begin{aligned}
&\Pr(work_i = 1 \mid income\_s_i, childu6_i) \\
&\quad = \Phi(\beta_0 + \beta_1 income\_s_i + \beta_2 childu6_i).
\end{aligned}
$$

- $work_i$: 労働ダミー
  - 働いている＝1
  - 働いていない＝0
- $income\_s_i$: 配偶者の所得（単位：万円）
- $childu6_i$: 6 歳以下子どもダミー
  - 6 歳以下の子どもがいる＝1
  - 6 歳以下の子どもがいない＝0
- $\Phi(.)$: 標準正規分布の累積分布関数

2 値プロビット・モデル推定結果（1）

gretl: モデル2

ファイル 編集(E) 検定(T) 保存(S) グラフ(G) 分析(A) LaTeX

```
モデル 2: プロビット・モデル, 観測: 1-1053
従属変数: work
QML standard errors

                係数          標準誤差         z           p値
  ------------------------------------------------------------
  const        0.707121      0.101954       6.936    4.04e-012 ***
  income_s    -0.000589132   0.000170616   -3.453    0.0006    ***
  childu6     -0.521817      0.0794574     -6.567    5.13e-011 ***

Mean dependent var   0.564103   S.D. dependent var   0.496109
McFadden R-squared   0.036541   Adjusted R-squared   0.032382
Log-likelihood      -694.8524   Akaike criterion     1395.705
Schwarz criterion    1410.583   Hannan-Quinn         1401.345

「正しく予測された」ケース数 = 640 (60.8%)
f(beta'x) (説明変数の平均における) = 0.393
尤度比検定: カイ二乗(2) = 52.7076 [0.0000]

             予測値
               0      1
  実績値 0   201    258
         1   155    439

残差の正規性の検定 -
  帰無仮説: 撹乱項は正規分布に従う
  検定統計量: カイ二乗(2) = 2.8933
  なお、p値(p-value) = 0.235357
```

2 値プロビット・モデル推定結果（2）

gretl: モデル3

ファイル　編集(E)　検定(T)　保存(S)　グラフ(G)　分析(A)　LaTeX

```
モデル 3: プロビット・モデル, 観測: 1-1053
従属変数: work
QML standard errors

                 係数          標準誤差        z         限界効果
  ------------------------------------------------------------------
  const        0.707121       0.101954       6.936
  income_s    -0.000589132    0.000170616   -3.453    -0.000231735
  childu6     -0.521817       0.0794574     -6.567    -0.202908

Mean dependent var     0.564103   S.D. dependent var     0.496109
McFadden R-squared     0.036541   Adjusted R-squared     0.032382
Log-likelihood        -694.8524   Akaike criterion       1395.705
Schwarz criterion      1410.583   Hannan-Quinn           1401.345

「正しく予測された」ケース数 = 640 (60.8%)
f(beta'x) (説明変数の平均における) = 0.393
尤度比検定: カイ二乗(2) = 52.7076 [0.0000]

           予測値
             0      1
  実績値 0  201    258
         1  155    439

残差の正規性の検定 -
  帰無仮説: 撹乱項は正規分布に従う
  検定統計量: カイ二乗(2) = 2.8933
  なお、p値(p-value) = 0.235357
```

- 配偶者の所得の係数
  - –0.000589132.
  - z 値は –3.453, p 値は 0.0006.
    ➥ 仮に「income_s の係数が 0」だとすると，–3.453 という z 値は 0.06%の確率（1%を下回る確率）でしか出てこない.
    ➥ 有意水準 1%で，「係数は 0」の H_0 が棄却される（5%や 10%でも棄却される）.
    ➥ 配偶者（夫）の所得は労働ダミーが 1 になる確率（女性本人が働く確率）と統計的に有意に相関している.
- 配偶者の所得の限界効果
  - –0.000231735.
    ➥ 6 歳以下の子どもの有無を一定としたうえで，配偶者（夫）の所得が 1 万円高くなると，労働ダミーが 1 になる確率（女性本人が働く確率）が平均的に 0.000231735 低くなる（0.0231735 パーセントポイント低くなる）傾向がある.

- 6 歳以下子どもダミーの係数
  - −0.521817.
  - z 値は −6.567, p 値は 5.13×10^{-11}.
    ➡ 仮に「childu6 の係数が 0」だとすると，−6.567 という z 値は 5.13×10^{-11}, つまりほぼ 0%の確率（1%を下回る確率）でしか出てこない．
    ➡ 有意水準 1%で，「係数は 0」の H_0 が棄却される（5%や 10%でも棄却される）．
    ➡ 6 歳以下の子どもの有無は労働ダミーが 1 になる確率（女性本人が働く確率）と統計的に有意に相関している．

- 6 歳以下子どもダミーの限界効果
  - −0.202908.

※ ダミー変数の，平均における限界効果の解釈には意味がなく，別の方法でダミー変数の限界効果を求める必要がある（詳細な説明は省略）．

- 定数項
  - 0.707121.
  - z 値は 6.936, p 値は 4.04×10^{-12}.
    ➡ 仮に「定数項が 0」だとすると，6.936 という z 値は 4.04×10^{-12}, つまりほぼ 0%の確率（1%を下回る確率）でしか出てこない.
    ➡ 有意水準 1%で,「係数は 0」の H_0 が棄却される（5%や 10%でも棄却される）.
    ➡ 定数項は統計的に有意に 0 と異なる.

※ 定数項に限界効果は存在しない.

今日のキーワード

線形確率モデル，潜在変数，2 値応答モデル，2 値プロビット・モデル，限界効果，平均における限界効果

次回（期末試験）までの準備

- これまでの講義スライドを読み直す．
- これまでの提出課題の問題の解答や解説を確認する．
- 「提出課題 11」に取り組む．
- これまでの各回の「今日のキーワード」に登場した専門用語の定義を覚える．

付録：2 値プロビット・モデルの推定方法

x_i を所与として，$y_i = 1$ となる条件付き確率は，

$$\begin{aligned}
\Pr(y_i = 1 \mid x_i) &= \Pr(y_i^* > 0 \mid x_i) \\
&= \Pr(\beta_0 + \beta_1 x_i + u_i > 0 \mid x_i) \\
&= \Pr\left(u_i > -(\beta_0 + \beta_1 x_i) \mid x_i\right).
\end{aligned}$$

標準正規分布は 0 で対称な分布なので，

$$\Pr\left(u_i > -(\beta_0 + \beta_1 x_i) \mid x_i\right) = \Pr\left(u_i < \beta_0 + \beta_1 x_i \mid x_i\right).$$

よって，

$$\begin{aligned}
\Pr(y_i = 1 \mid x_i) &= \Pr\left(u_i < \beta_0 + \beta_1 x_i \mid x_i\right) \\
&= \Phi(\beta_0 + \beta_1 x_i).
\end{aligned}$$

$\Phi(.)$ は標準正規分布の累積分布関数.

- 前スライドの式では,

$$\Phi(\beta_0 + \beta_1 x_i) = \int_{-\infty}^{\beta_0+\beta_1 x_i} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz.$$

また，x_i を所与として，$y_i = 0$ となる条件付き確率は,

$$\begin{aligned}\Pr(y_i = 0 \mid x_i) &= 1 - \Pr(y_i = 1 \mid x_i)\\ &= 1 - \Phi(\beta_0 + \beta_1 x_i).\end{aligned}$$

よって，x_i を所与とした y_i の条件付き確率質量関数は，

$$
\begin{aligned}
f(y_i \mid x_i; \beta_0, \beta_1) &= \begin{cases} \Phi(\beta_0 + \beta_1 x_i) & \text{for } y_i = 1, \\ 1 - \Phi(\beta_0 + \beta_1 x_i) & \text{for } y_i = 0, \\ 0 & \text{elsewhere} \end{cases} \\
&= \left[\Phi(\beta_0 + \beta_1 x_i)\right]^{y_i} \left[1 - \Phi(\beta_0 + \beta_1 x_i)\right]^{1-y_i} .
\end{aligned}
$$

無作為標本なので $y_1, y_2, \cdots, y_n$ は互いに独立.
$x_1, x_2, \cdots, x_n$ を所与とした，$y_1, y_2, \cdots, y_n$ の同時確率質量関数は,

$$
\begin{aligned}
& f(y_1, y_2, \cdots, y_n \mid x_1, x_2, \cdots, x_n; \beta_0, \beta_1) \\
= & \prod_{i=1}^{n} f(y_i \mid x_1, x_2, \cdots, x_n; \beta_0, \beta_1) \\
= & \prod_{i=1}^{n} f(y_i \mid x_i; \beta_0, \beta_1) \\
= & \prod_{i=1}^{n} [\Phi(\beta_0 + \beta_1 x_i)]^{y_i} [1 - \Phi(\beta_0 + \beta_1 x_i)]^{1-y_i} .
\end{aligned}
$$

尤度関数（likelihood function）は，

$$
\begin{aligned}
&L(\beta_0, \beta_1; y_1, y_2, \cdots, y_n, x_1, x_2, \cdots, x_n) \\
&= \prod_{i=1}^{n} \left[\Phi(\beta_0 + \beta_1 x_i)\right]^{y_i} \left[1 - \Phi(\beta_0 + \beta_1 x_i)\right]^{1-y_i} .
\end{aligned}
$$

対数尤度関数（log-likelihood function）は，

$$
\begin{aligned}
&\ln L(\beta_0, \beta_1; y_1, y_2, \cdots, y_n, x_1, x_2, \cdots, x_n) \\
&= \sum_{i=1}^{n} \Big[y_i \ln \left\{\Phi(\beta_0 + \beta_1 x_i)\right\} \\
&\qquad + (1 - y_i) \ln \left\{1 - \Phi(\beta_0 + \beta_1 x_i)\right\} \Big] .
\end{aligned}
$$

これが最大になるような β_0 と β_1 を求める．

ML 問題は，

$$\max_{\beta_0,\beta_1} \sum_{i=1}^{n} \Big[y_i \ln \{\Phi(\beta_0 + \beta_1 x_i)\} + (1 - y_i) \ln \{1 - \Phi(\beta_0 + \beta_1 x_i)\} \Big] .$$

(β_0, β_1) の最尤推定量（maximum likelihood estimator, MLE）を $(\hat{\beta}_0, \hat{\beta}_1)$ とする．

1 階条件は,

$$
\begin{aligned}
&\frac{\partial \ln L}{\partial \beta_0} = 0 \\
\Leftrightarrow &\sum_{i=1}^{n} \left[\frac{y_i}{\Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)} \cdot \frac{d\Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)}{dz} \right. \\
&\quad \left. - \frac{1 - y_i}{1 - \Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)} \cdot \frac{d\Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)}{dz} \right] = 0 \\
\Leftrightarrow &\sum_{i=1}^{n} \left[\frac{y_i \phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)}{\Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)} - \frac{(1 - y_i)\phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)}{1 - \Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)} \right] = 0 \\
\Leftrightarrow &\sum_{i=1}^{n} \left[\frac{(y_i - \Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i))\phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)}{\Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)(1 - \Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i))} \right] = 0, \qquad (1)
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial \ln L}{\partial \beta_1} = 0 \\
&\Leftrightarrow \sum_{i=1}^{n} \left[\frac{y_i x_i}{\Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)} \cdot \frac{d\Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)}{dz} \right. \\
&\qquad \left. - \frac{(1 - y_i)x_i}{1 - \Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)} \cdot \frac{d\Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)}{dz} \right] = 0 \\
&\Leftrightarrow \sum_{i=1}^{n} \left[\frac{y_i x_i \phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)}{\Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)} - \frac{(1 - y_i)x_i \phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)}{1 - \Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)} \right] = 0 \\
&\Leftrightarrow \sum_{i=1}^{n} \left[\frac{(y_i - \Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i))x_i \phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)}{\Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)(1 - \Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i))} \right] = 0. \qquad (2)
\end{aligned}
$$

$\phi(.)$ は標準正規分布の確率密度関数.

- (1) と (2) において,

$$\phi(\hat{\beta}_0 + \hat{\beta}_1 x_i) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\hat{\beta}_0 + \hat{\beta}_1 x_i)^2}{2}\right).$$

(1) と (2) からなる連立方程式は解析的に解けない.

$$\Downarrow$$

コンピューターを用いて数値的に解き，$(\hat{\beta}_0, \hat{\beta}_1)$ を求める.

※ 定数項以外に説明変数が複数ある場合も同様に求めることができる.