10.1　各種検定方法：まとめ

回帰式

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i \qquad u_i \sim N(0, \sigma^2) \qquad i = 1, 2, \cdots, n$$

ただし，u_1，u_2，$\cdots$，u_n は互いに独立とする。

1. 個々の H_0：$\beta_j = 0$ の検定 $\Longrightarrow \dfrac{\hat{\beta}_j}{s_{\beta_j}} \sim t(n-k)$
   ただし，$\hat{\beta}_j$ は β_j の最小二乗推定量，s_{β_j} は $\hat{\beta}_j$ の標準誤差の推定量
2. β_1，β_2，$\cdots$，β_k に関する G 個の制約の検定 $\Longrightarrow \dfrac{(\sum \widetilde{u}_i^2 - \sum \hat{u}_i^2)/G}{\sum \hat{u}_i^2/(n-k)} \sim$
   $F(G, n-k)$
   ただし，$\sum \widetilde{u}_i^2$ は制約付き残差平方和，$\sum \hat{u}_i^2$ は制約なし残差平方和
3. β_1，β_2，$\cdots$，β_k に関する G 個の制約の検定 $\Longrightarrow \dfrac{(\hat{R}^2 - \widetilde{R}^2)/G}{(1-\hat{R}^2)/(n-k)} \sim$
   $F(G, n-k)$
   ただし，$\widetilde{R}^2$ は制約付き決定係数，$\hat{R}^2$ は制約なし決定係数

4. 個々の H_0：$\theta_j = 0$ の検定 $\Longrightarrow$ 最尤推定量の性質から，n が大きいとき，
$\hat{\theta}_j \sim N(\theta, \hat{\sigma}_j^2)$,
すなわち，H_0 のもとで $\dfrac{\hat{\theta}_j}{\hat{\sigma}_j} \longrightarrow N(0,1)$
ただし，$\hat{\theta}_j$ は θ_j の最尤推定量，$\hat{\sigma}_j^2$ は $\sigma_j^2 = \mathrm{V}(\hat{\theta}_j)$ の最尤推定量

5. β_1，β_2，$\cdots$，β_k に関する G 個の制約の検定 $\Longrightarrow$ $-2\Big(\log l(\widetilde{\theta}) - \log l(\hat{\theta})\Big) \longrightarrow \chi^2(G)$
ただし，$\widetilde{\theta} = (\widetilde{\beta}_1, \widetilde{\beta}_2, \cdots, \widetilde{\beta}_k)$ は制約付き最尤推定量，$\hat{\theta} = (\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k)$ は制約なし最尤推定量

4，5 について，n が大きいときのみ利用可能
回帰係数だけでなく，他の検定にも利用可能（例えば，系列相関の検定）

第11章　離散選択モデル

質的従属変数 (Qualitative Dependent Variable) の一種：離散選択モデル (Discrete Choice Model)

通常の回帰モデル：

$$Y_i = \alpha + \beta X_i + u_i \qquad u_i \sim N(0, \sigma^2) \qquad i = 1, 2, \cdots, n$$

u_i が $-\infty$ から ∞ の範囲の値を取る連続型確率変数なので，Y_i も連続型確率変数である。

Y_i が離散型確率変数であればどうなるか？

(* 復習) 離散型確率変数と連続型確率変数：

離散型確率変数 X： 不連続な値を取る

- サイコロの出た目（X の取る値は 1，2，3，4，5，6）
- コインを投げて表・裏（X の取る値：表の場合は 1，裏の場合は 0）

$\cdots$

連続型確率変数 X： ある区間内（$-\infty$ から ∞ の区間も含む）のどの実数値も取り得る

11.1　二値選択モデル (Binary Choice Model)

アンケート調査を行う。
回答は YES か NO の 2 つから 1 つを選択することとする。

$$Y_i = \begin{cases} 1, & i \text{ 番目の人が YES と答えたとき} \\ 0, & i \text{ 番目の人が NO と答えたとき} \end{cases}$$

通常の次の回帰モデルを考える。

$$Y_i = \alpha + \beta X_i + u_i \qquad i = 1, 2, \cdots, n$$

u_1，u_2，$\cdots$，u_n は互いに独立で，平均ゼロ・分散 σ^2 とする。
E$(u_i) = 0$ なので，Y_i の期待値は,

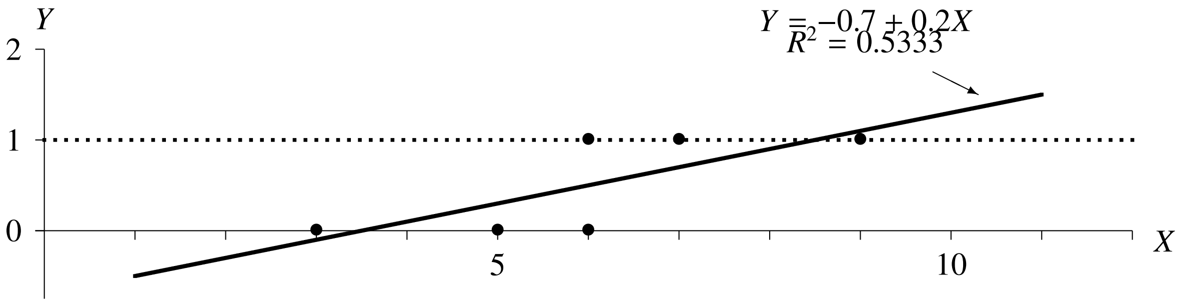
$$\mathrm{E}(Y_i) = \alpha + \beta X_i$$

となる。
線形関数なので，$\alpha + \beta X_i$ は $-\infty$ から ∞ の値を取ることになる。
この定式化は適切か？

数値例：

i	1	2	3	4	5	6
Y_i	0	0	0	1	1	1
X_i	3	5	6	6	7	9

最小二乗法で α, β を推定する。



● Gretl による最小二乗法の結果：

```
? ols y const x

モデル 1: 最小二乗法 (OLS), 観測: 1-6
従属変数: y

                係数        標準誤差       t 値       p 値
  ------------------------------------------------------------
  const      -0.700000   0.586657    -1.193    0.2987
  x           0.200000   0.0935414    2.138    0.0993  *

Mean dependent var   0.500000   S.D. dependent var   0.547723
Sum squared resid    0.700000   S.E. of regression   0.418330
R-squared            0.533333   Adjusted R-squared   0.416667
F(1, 4)              4.571429   P-value(F)           0.099301
Log-likelihood      -2.068328   Akaike criterion     8.136656
Schwarz criterion    7.720175   Hannan-Quinn         6.469448
```

一方，$\mathrm{E}(Y_i)$ を計算する。

(* 復習) 離散型確率変数の期待値：

確率変数 X は x_1，x_2，$\cdots$ を取り，X が x_i を取る確率を p_i とする。
すなわち，$P(X = x_i) = p_i$
ある関数 $g(\cdot)$ について，$g(X)$ の期待値は次のように計算される。

$$\mathrm{E}(g(X)) = \sum_i g(x_i)p_i$$

$g(\cdot)$ は $g(X) = X$ や $g(X) = (X - \mu)^2$ など

(* 復習) ベルヌイ分布：

確率変数 X は，確率 p で 1 を取り，確率 $1-p$ で 0 を取る

このとき，X の確率関数 $f(x)$ は,

$$f(x) = p^x(1-p)^{1-x} \qquad x = 0, 1$$

と表される。

平均 $\mathrm{E}(X)$ と分散 $V(X)$ は,

$$\mu = \mathrm{E}(X) = \sum_{x=0}^{1} xf(x) = 0 \times (1-p) + 1 \times p = p$$

$$\sigma^2 = \mathrm{V}(X) = \mathrm{E}((X-\mu)^2) = \sum_{x=0}^{1}(x-\mu)^2 f(x) = (0-p)^2(1-p) + (1-p)^2 p = p(1-p)$$

となる。

Y_i の取る値は 0 か 1 のどちらかなので，$P(Y_i = 0) + P(Y_i = 1) = 1$ となる。

$P(Y_i = 1) = p$ とすると，$P(Y_i = 0) = 1 - p$ となる。

したがって，$\mathrm{E}(Y_i) = 0 \times P(Y_i = 0) + 1 \times P(Y_i = 1) = P(Y_i = 1) = p$ となり，確率 p に等しいので，$0 < \mathrm{E}(Y_i) < 1$ となる。

確率を表すということは，$\mathrm{E}(Y_i)$ が，ゼロより小さくなる，1 より大きくなることはあり得ない。

$\mathrm{E}(Y_i)$ を $\alpha + \beta X_i$ とするのは不適切である。

誤差項 u_i について：　Y_i の期待値は,

$$\mathrm{E}(Y_i) = P(Y_i = 1) = p$$

なので,

$$Y_i = P(Y_i = 1) + u_i$$

と誤差項を加えて書き換えることができる。

$u_i = Y_i - P(Y_i = 1)$ なので，u_i は確率 $P(Y_i = 1)$ で $1 - P(Y_i = 1)$ の値を取るか，確率 $P(Y_i = 0) = 1 - P(Y_i = 1)$ で $-P(Y_i = 1)$ の値を取ることになる。

すなわち，誤差項 u_i も離散型確率変数となる。

推定方法について：　$P(Y_i = 1)$ は分布関数 $F(\cdot)$ に関連付けられて，説明変数 X_i の関数として，$F(\alpha + \beta X_i)$ と表す。

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

多変数の場合は,

$$P(Y_i = 1) = F(\beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki})$$

となる。

$F(\cdot)$ には，標準正規分布やロジスティック分布が使われる。

$F(\cdot)$ に標準正規分布が使われた場合はプロビット・モデル（probit model）と呼ばれ，ロジスティック分布が使われた場合はロジット・モデル（logit model）と呼ばれる。

標準正規分布：

密度関数：$f(x) = (2\pi)^{-1/2} \exp(-\frac{1}{2}x^2)$

分布関数：$F(x) = \int_{-\infty}^{x} (2\pi)^{-1/2} \exp(-\frac{1}{2}z^2) \mathrm{d}z$

ロジステック分布：

密度関数：$f(x) = \dfrac{\exp(-x)}{(1 + \exp(-x))^2}$

$$\text{分布関数：}\quad F(x) = \frac{1}{1 + \exp(-x)}$$

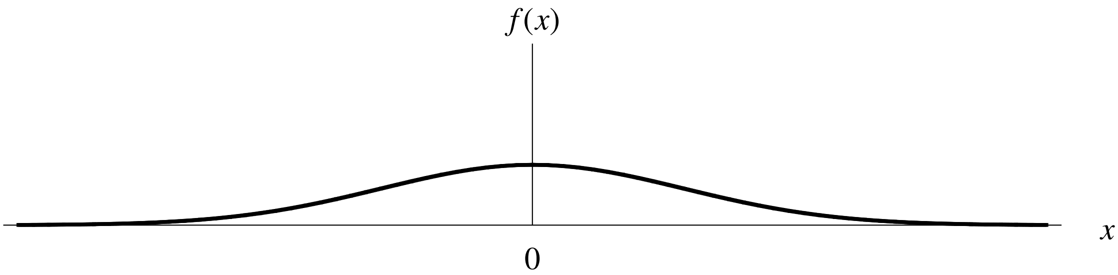
$F(\cdot)$ に他の分布関数を用いてもよい。

(* 復習) 密度関数と分布関数：

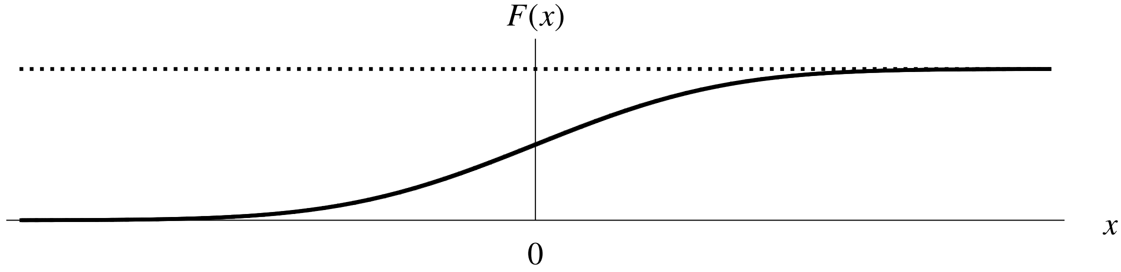
$$\frac{\mathrm{d}F(x)}{\mathrm{d}x} = f(x)$$

標準正規分布のケース：

● 密度関数:



● 分布関数:



● **最尤法：**　Y_i の確率関数は,

$$f(Y_i) = F(\alpha + \beta X_i)^{Y_i}(1 - F(\alpha + \beta X_i))^{1-Y_i} \equiv F_i^{Y_i}(1 - F_i)^{1-Y_i} \qquad Y_i = 0, 1$$

となる。$F_i = F(\alpha + \beta X_i)$ としている。

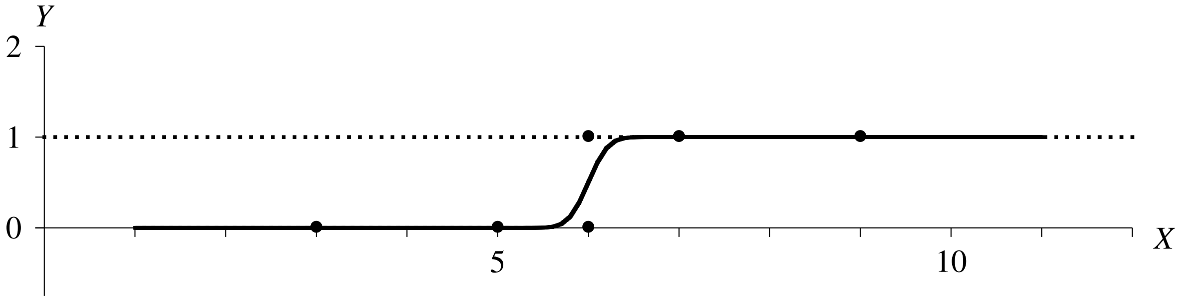
Y_1, Y_2, $\cdots$, Y_n の結合確率関数（同時確率関数）は,

$$f(Y_1, Y_2, \cdots, Y_n) = \prod_{i=1}^{n} f(Y_i) = \prod_{i=1}^{n} F_i^{Y_i}(1 - F_i)^{1-Y_i} \equiv l(\alpha, \beta)$$

となる。

　尤度関数 $l(\alpha,\beta)$ を α, β について最大にする $\hat{\alpha}$, $\hat{\beta}$ を求めることになる。　$\Longrightarrow$ 最尤法

```
? probit y const x

モデル 2: プロビット・モデル, 観測: 1-6
従属変数: y
標準誤差はヘッシアン（Hessian）に基づく

                 係数         標準誤差         z          限界効果
  -----------------------------------------------------------------
  const        -34.8819     13019.0     -0.002679
  x              5.81364     2169.83     0.002679   2.31931

Mean dependent var    0.500000   S.D. dependent var    0.547723
McFadden R-squared    0.666667   Adjusted R-squared    0.185768
Log-likelihood       -1.386294   Akaike criterion      6.772589
Schwarz criterion     6.356108   Hannan-Quinn          5.105381

「正しく予測された」ケース数 = 5 (83.3%)
f(beta'x) （説明変数の平均における） = 0.399
尤度比検定: カイ二乗(1) = 5.54518 [0.0185]

           予測値
            0   1
  実績値 0   3   0
         1   1   2
```



太線は，説明変数 X を与えたもとでの Y の予測値，すなわち，今までの記号では $\hat{Y}_i = F(\hat{\alpha} + \hat{\beta}X_i)$ を表す。

● **限界効果 2.31931 の意味：**　限界係数（太線の接線の傾き）：

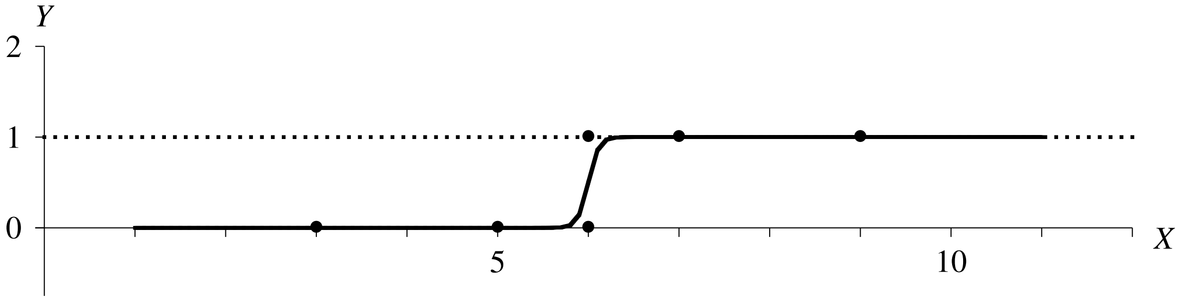
$$\frac{\mathrm{d}\hat{Y}_i}{\mathrm{d}X_i} = \frac{\mathrm{d}F(\hat{\alpha} + \hat{\beta}X_i)}{\mathrm{d}X_i} = f(\hat{\alpha} + \hat{\beta}X_i)\hat{\beta}$$

となり，X_i の値に依存する。

限界効果： $2.31931 = \dfrac{\mathrm{d}F(\hat{\alpha} + \hat{\beta})\overline{X}}{\mathrm{d}\overline{X}} = f(\hat{\alpha} + \hat{\beta}\overline{X})\hat{\beta}$ と説明変数の平均値で評価する。

$f(\cdot)$，$F(\cdot)$ は標準正規分布，または，ロジスティック分布，または，他の分布
上の推定では標準正規分布，次の推定ではロジスティック分布

```
? logit y const x

モデル 3: ロジット・モデル, 観測: 1-6
従属変数: y
標準誤差はヘッシアン（Hessian）に基づく

                係数        標準誤差        z         限界効果
  ---------------------------------------------------------
  const      -106.179    29537.1    -0.003595
  x            17.6964    4922.86    0.003595   4.42411

Mean dependent var   0.500000   S.D. dependent var   0.547723
McFadden R-squared   0.666667   Adjusted R-squared   0.185768
Log-likelihood      -1.386294   Akaike criterion     6.772589
Schwarz criterion    6.356108   Hannan-Quinn         5.105381

「正しく予測された」ケース数 = 5 (83.3%)
f(beta'x) （説明変数の平均における） = 0.250
尤度比検定: カイ二乗(1) = 5.54518 [0.0185]

           予測値
            0   1
  実績値 0   3   0
         1   1   2
```



Y
2
1
0
5
10
X