1.5　尤度比検定

前節では，最尤推定量の性質について解説した。本節では，尤度関数を用いた検定，すなわち，尤度比検定について説明する。

n 個の確率変数 $X_1, X_2, \cdots, X_n$ は互いに独立で，同じ確率分布 $f(x) \equiv f(x;\theta)$ とする。このとき，尤度関数は，

$$l(\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

であることを思い出してもらいたい。最小二乗推定量のときと同様に最尤推定量にも制約付き推定量と制約なし推定量を求めることができる。θ の制約付き

最尤推定量を $\tilde{\theta}$，制約なし最尤推定量を $\hat{\theta}$ としよう。制約付き最尤推定量の方の制約の数を G 個とする。このとき，

$$\lambda = \frac{l(\tilde{\theta})}{l(\hat{\theta})}$$

を尤度比（**likelihood ratio**）と呼び，この尤度比を用いて制約が正しいかどうかの検定を行うことができる。この尤度比を用いて行う検定のことを尤度比検定（**likelihood ratio test**）と呼ぶ。

F 検定を使って，制約付き最小二乗法，制約なし最小二乗法の **2** つで推定して，残差平方和または決定係数を用いて，制約が正しいかどうかの検定を行う方法を示した。本節では尤度関数を用いて，制約が正しいかどうかの検定を行

うことを考える。F 検定では制約は線型制約に限っていたが，本節の尤度比検定は非線型制約の場合でも適用可能であることが一番の特徴である。ただし，実践では，多くの場合，尤度比検定はデータが多い（すなわち，大標本）ときにだけ利用することができる。

検定方法 1：　制約付きの尤度関数 $l(\tilde{\theta})$ と制約なしの尤度関数 $l(\hat{\theta})$ とを比べた場合，必ず，後者の方が前者より大きくなる。したがって，尤度比は必ず **1** より小さくなる，すなわち，$\dfrac{l(\tilde{\theta})}{l(\hat{\theta})} < 1$ となるはずである。尤度比が **1** よりかなり小さいとき，制約が正しいという帰無仮説を棄却すればよいことになる。よって，尤度比がある値（例えば，c）より小さいときに，帰無仮説を棄却する。す

なわち，

$$\lambda = \frac{l(\tilde{\theta})}{l(\hat{\theta})} < c$$

となるときに，帰無仮説を棄却する。α を有意水準（帰無仮説が正しいときに，帰無仮説を棄却する確率）とする。この場合，帰無仮説が正しいもとで λ の分布を $g(\cdot)$ とするとき，分布関数 $g(\cdot)$ と有意水準 α との間にある次の関係を利用して c を求める必要がある。

$$\int_{-\infty}^{c} g(\lambda)\,\mathbf{d}\lambda = \alpha$$

具体例は後述する。

この方法は，データが少なくても利用可能ではあるが，尤度比の中心となる部分（専門的には「核となる部分」という言い方をする場合もある）の分布を求めて，c を得る必要がある。検定の設定に応じてその都度分布を求めることになるため，かなり手間であり実用上は使えないといっても言い過ぎではないだろう。それに対して，次に示す検定方法は，データが多くなければ使えないという欠点はあるが，非常に実用的である。データが多いときに使える検定のことを大標本検定（**large sample test**）という。

検定方法 2（大標本検定）：　大標本における尤度比検定は非常に簡単に検定統計量を求めることができ，非常に簡単に検定することができる。今までと同

じ記号を用いる。すなわち，θ の制約付き最尤推定量を $\tilde{\theta}$，制約なし最尤推定量を $\hat{\theta}$，制約の数を G 個とする。このとき，$n \longrightarrow \infty$ のとき，

$$-2\log\lambda = -2\log\frac{l(\tilde{\theta})}{l(\hat{\theta})} \longrightarrow \chi^2(G) \tag{1.6}$$

となることが知られている（証明略）。多くの計量ソフトで，対数尤度関数の値が出力される（**Gretl** では，`Log-likelihood` という項目がある）。この場合，制約付きの推定と制約なしの推定を行い，両方の推定結果から得られる対数尤度関数の値の差の 2 倍を自由度 G のカイ二乗分布と比較すればよい。

例 1：　正規母集団 $N(\mu, \sigma^2)$ からの標本値 x_1，x_2，$\cdots$，x_n を用いて，σ^2 が既知のとき，帰無仮説 H_0：$\mu = \mu_0$，H_1：$\mu \neq \mu_0$ の尤度比検定を行う。

解：「検定方法 1」を利用して：　σ^2 が既知のとき，尤度関数 $l(\mu)$ は，

$$l(\mu) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right)$$

となる。このとき，μ の最尤推定量は，

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i \equiv \overline{X}$$

となる。尤度比検定統計量は，

$$\lambda = \frac{l(\mu_0)}{l(\overline{X})} = \frac{(2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum (X_i - \mu_0)^2\right)}{(2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum (X_i - \overline{X})^2\right)} = \exp\left(-\frac{1}{2\sigma^2/n}(\overline{X} - \mu_0)^2\right) < c$$

となる c を求める。

H_0 が正しいときに，$\dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$ となるので，

$$P\left(\left|\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}\right| > z_{\alpha/2}\right) = \alpha$$

となる。さらに，

$$P\left(\exp\left(-\frac{1}{2\sigma^2/n}(\overline{X} - \mu_0)^2\right) < \exp\left(-\frac{1}{2}z_{\alpha/2}^2\right)\right) = \alpha$$

と変形できる。したがって，

$$c = \exp\left(-\frac{1}{2}z_{\alpha/2}^2\right)$$

とすればよい。

すなわち，

$$\lambda = \frac{l(\mu_0)}{l(\overline{X})} < \exp\left(-\frac{1}{2}z_{\alpha/2}^2\right)$$

のとき，有意水準 α の両側検定で帰無仮説 H_0：$\mu = \mu_0$ を棄却することになる。ただし，$z_{\alpha/2}$ は標準正規分布の上側 $100 \times \alpha/2$ % 点とする。この検定方法は，n が小さいときでも成り立つ。しかし，母集団の分布が変われば，検定統

計量の分布も異なる。帰無仮説が変わっても，検定統計量の分布も変わる。このように，問題に応じて，その都度，検定方式を考えなければならない。その意味では，扱いがかなり面倒な検定と言えるだろう。

解：「検定方法 **2**」を利用して：　尤度比は次のように得られた。

$$\frac{l(\mu_0)}{l(\overline{X})} = \exp\Bigl(-\frac{1}{2\sigma^2/n}(\overline{X}-\mu_0)^2\Bigr)$$

帰無仮説 H_0：$\mu = \mu_0$ が正しいもとで，$n \longrightarrow \infty$ のとき，

$$-2\log\frac{l(\mu_0)}{l(\overline{X})} \longrightarrow \chi^2(G)$$

となることが知られている。この場合は制約数は **1** となるので $G = 1$ である。よって，尤度比は次のように得られる。

$$-2\log\frac{l(\mu_0)}{l(\overline{X})} = \left(\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right)^2$$

右辺の括弧内は $\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,\ 1)$ となり，標準正規分布に従う確率変数を二乗したものは $\left(\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right)^2 \sim \chi^2(1)$ となる。通常，尤度比検定は n が大きくなれば（「漸近的に」ともいう），$\chi^2(G)$ 分布に近づくことがわかっているが，この例では n の値にかかわらず，尤度比検定統計量は自由度 **1** の $\chi^2(1)$ 分布に従うことが示される。

例 2：平均の差の検定：　次の 2 つのグループを考える。

$$X_i \sim N(\mu_1,\ \sigma^2), \qquad i = 1, 2, \cdots, n$$

$$Y_j \sim N(\mu_2,\ \sigma^2), \qquad j = 1, 2, \cdots, m$$

それぞれは互いに独立な確率変数であるとする。σ^2 は既知とする。このとき，帰無仮説 H_0：$\mu_1 = \mu_2$，対立仮説 H_1：$\mu_1 \neq \mu_2$ を検定する。

解：「検定方法 1」を利用して：　X_i，Y_j の尤度関数は，

$$l(\mu_1, \mu_2) = (2\pi\sigma^2)^{-\frac{n+m}{2}} \exp\left(-\frac{1}{2\sigma^2}\Big(\sum_{i=1}^{n}(x_i - \mu_1)^2 + \sum_{j=1}^{m}(y_j - \mu_2)^2\Big)\right)$$

となる。

$$\max_{\mu_1,\ \mu_2} l(\mu_1,\ \mu_2)$$

を解くと，μ_1，μ_2 の最尤推定量は，

$$\hat{\mu}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}, \qquad \hat{\mu}_2 = \frac{1}{m}\sum_{j=1}^{m} Y_j = \overline{Y}$$

となる。よって，制約なしの尤度関数の値は，

$$l(\hat{\mu}_1,\ \hat{\mu}_2) = (2\pi\sigma^2)^{-\frac{n+m}{2}} \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} (x_i - \hat{\mu}_1)^2 + \sum_{j=1}^{m} (y_j - \hat{\mu}_2)^2\right)\right)$$

となる。

一方，帰無仮説 H_0：$\mu_1 = \mu_2$ が正しいとき，$\mu_1 = \mu_2 = \mu$ とすると，

$$l(\mu,\mu) = (2\pi\sigma^2)^{-\frac{n+m}{2}} \exp\left(- \frac{1}{2\sigma^2}\Big(\sum_{i=1}^{n} (x_i - \mu)^2 + \sum_{j=1}^{m} (y_j - \mu)^2\Big)\right)$$

となり，

$$\max_{\mu} l(\mu,\ \mu)$$

の μ の最尤推定量は，尤度関数 $l(\mu,\ \mu)$ を μ について微分してゼロと置き，

$$\hat{\mu} = \frac{1}{n+m}\Big(\sum_{i=1}^{n} X_i + \sum_{j=1}^{m} Y_j\Big) = \frac{n\overline{X} + m\overline{Y}}{n+m}$$

となる。これが制約付き最尤推定量に対応する。よって，制約なしの尤度関数

の値は，

$$l(\hat{\mu}_1,\ \hat{\mu}_2) == (2\pi\sigma^2)^{-\frac{n+m}{2}} \exp\left(- \frac{1}{2\sigma^2}\Big(\sum_{i=1}^{n} (x_i - \hat{\mu})^2 + \sum_{j=1}^{m} (y_j - \hat{\mu})^2\Big)\right)$$

となる。

尤度比は次のように求められる。

$$\lambda = \frac{l(\hat{\mu},\ \hat{\mu})}{l(\hat{\mu}_1,\ \hat{\mu}_2)} = \frac{\exp\left(- \frac{1}{2\sigma^2}\Big(\sum_{i=1}^{n} (X_i - \hat{\mu})^2 + \sum_{j=1}^{m} (Y_j - \hat{\mu})^2\Big)\right)}{\exp\left(- \frac{1}{2\sigma^2}\Big(\sum_{i=1}^{n} (X_i - \hat{\mu}_1)^2 + \sum_{j=1}^{m} (Y_j - \hat{\mu}_2)^2\Big)\right)}$$

$$= \exp\left(- \frac{(\overline{X} - \overline{Y})^2}{2\sigma^2(1/n + 1/m)}\right) < c$$

となる c を求める。そのため，$\overline{X}-\overline{Y}$ の分布を求めると，

$$\overline{X}-\overline{Y} \ \sim\ N\Big(0,\ \sigma^2(\frac{1}{n}+\frac{1}{m})\Big)$$

となる。よって，標準正規分布の上側 $100\times\alpha/2$ % 点を $z_{\alpha/2}$ とすると，

$$P\Big(\Big|\frac{\overline{X}-\overline{Y}}{\sigma\sqrt{1/n+1/m}}\Big| > z_{\alpha/2}\Big) = \alpha$$

となる。尤度比検定統計量となるように変形すると，

$$P\Bigg(\exp\Big(-\frac{(\overline{X}-\overline{Y})^2}{2\sigma^2(1/n+1/m)}\Big) < \exp\Big(-\frac{1}{2}z_{\alpha/2}^2\Big)\Bigg) = \alpha$$

となる。したがって，c を

$$c = \exp\Big(-\frac{1}{2}z_{\alpha/2}^2\Big)$$

とすればよい。この c を用いて，尤度比 λ が $\lambda < c$ のとき，帰無仮説（すなわち，制約条件）を棄却する。

解：「検定方法 2」を利用して：　$n \longrightarrow \infty$ のとき，

$$\lambda = \frac{l(\hat{\mu},\ \hat{\mu})}{l(\hat{\mu}_1,\ \hat{\mu}_2)} = \frac{\exp\left(-\dfrac{1}{2\sigma^2}\Bigl(\sum_{i=1}^{n}(X_i-\hat{\mu})^2+\sum_{j=1}^{m}(Y_j-\hat{\mu})^2\Bigr)\right)}{\exp\left(-\dfrac{1}{2\sigma^2}\Bigl(\sum_{i=1}^{n}(X_i-\hat{\mu}_1)^2+\sum_{j=1}^{m}(Y_j-\hat{\mu}_2)^2\Bigr)\right)}$$

となる。ただし，

$$\hat{\mu}_1 = \frac{1}{n}\sum_{i=1}^{n}X_i = \overline{X}, \qquad \hat{\mu}_2 = \frac{1}{m}\sum_{j=1}^{m}Y_j = \overline{Y}, \qquad \hat{\mu} = \frac{n\overline{X}+m\overline{Y}}{n+m}$$

である。帰無仮説 H_0： $\mu_1 = \mu_2$ が正しいもとで，$n \longrightarrow \infty$ のとき，

$$-2\log\lambda \longrightarrow \chi^2(G)$$

となる。制約数は **1** なので，$G = 1$ である。

1.6　数値例

練習問題 **4.2** の例を **Gretl** を用いて推定する（練習問題 **4.2** では **Excel** で推定）。記号は練習問題 **4.2** と同じであり，生鮮魚介 **F** の需要関数を推定している。説明変数は所得 **Y**，生鮮魚介 **1g** 当たりの価格 **PF**，生鮮肉 **1g** 当たりの価

格 **PM**，生鮮野菜 1g 当たりの価格 **PV** とする。変数はすべて対数を取っている。推定式とその結果は次のとおりである。

$$\mathbf{F}_i = \alpha + \beta\mathbf{Y}_i + \gamma_1\mathbf{PF}_i + \gamma_2\mathbf{PM}_i + \gamma_3\mathbf{PV}_i + u_i \tag{1.7}$$

```
? ols F const Y PF PM PV

[モデル 1: 最小二乗法 (OLS), 観測: 2000-2019 (T = 20)
従属変数: F

                 係数        標準誤差       t 値       p 値
  ---------------------------------------------------
  const       -13.2654      5.27734     -2.514     0.0239  **
  Y             3.09750     0.950373     3.259     0.0053  ***
  PF           -2.46563     0.476834    -5.171     0.0001  ***
  PM            2.35611     0.559867     4.208     0.0008  ***
  PV           -0.226836    0.658757    -0.3443    0.7354

Mean dependent var   4.461092   S.D. dependent var   0.112105
Sum squared resid    0.021734   S.E. of regression   0.038065
R-squared            0.908981   Adjusted R-squared   0.884709
F(4, 15)             37.45016   P-value(F)           1.22e-07
Log-likelihood       39.86740   Akaike criterion    -69.73479

 ·····(以下略)·····
```

この推定結果を踏まえて，生鮮野菜 **1g** 当たりの価格 **PV** を除き，さらに，生鮮魚介と生鮮肉とは代替関係があると想定し，次の式を推定した。

$$\mathbf{F}_i = \alpha + \beta \mathbf{Y}_i + \gamma_1 \mathbf{PFM}_i + u_i \tag{1.8}$$

```
? genr PFM=PF-PM

? ols F const Y PFM

モデル 2: 最小二乗法 (OLS), 観測: 2000-2019 (T = 20)
従属変数: F

                係数        標準誤差       t 値          p 値
  -----------------------------------------------------
  const      -13.8575    4.43919      -3.122    0.0062     ***
  Y            3.21520   0.775280      4.147    0.0007     ***
  PFM         -2.54348   0.201082    -12.65     4.48e-010 ***

Mean dependent var    4.461092   S.D. dependent var    0.112105
Sum squared resid     0.022024   S.E. of regression    0.035994
R-squared             0.907765   Adjusted R-squared    0.896913
F(2, 17)              83.65543   P-value(F)            1.59e-09
Log-likelihood        39.73464   Akaike criterion    -73.46928

 ..... (以下略) .....
```

(1.7) 式に，$\gamma_1 = -\gamma_2$，$\gamma_3 = 0$ の制約を課すと，**(1.8)** 式になる。練習問題 **4.2** の問 **(9)** を参照されたい。この検定の残差平方和（`Sum squared resid`）を用いた検定は練習問題 **4.2** の問 **(10)**，決定係数（`R-squared`）を利用した検定は練習問題 **4.2** の問 **(11)** に対応する。

尤度比検定を用いて，**(1.7)** 式で $\gamma_1 = -\gamma_2$，$\gamma_3 = 0$ の検定を行う。データ数は **20** 個で多いとは言えないが，試しに尤度比検定を行うことにする。対数尤度関数の値は `Log-likelihood` に出力される。**(1.7)** 式の対数尤度関数は **39.86740**，**(1.8)** 式の対数尤度関数は **39.73464** $-2\log\lambda = -2(39.73464 - 39.86740) = 0.2655$ となり，$\chi^2(2)$ 分布の上側 **5%** 点は **5.99** なので，二つの制約を棄却できない。

したがって，**(1.8)** 式で推定して構わないという結論になる。読者は練習問題 **4.2** の問 **(10)**，問 **(11)** の結果と比較してみるとよい。

1.7　各種検定方法（第3章以降）：まとめ

回帰モデル

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

について，本書のここまでで扱った検定方法をまとめておく。ただし，u_1，u_2，$\cdots$，u_n は互いに独立とする。

(i) 個々の H_0：$\beta_j = \beta_j^*$ の検定 — (??) 参照：

$$\frac{\hat{\beta}_j - \beta_j^*}{s_{\beta_j}} \sim t(n-k)$$

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とする。

(ii) β_1，β_2，$\cdots$，β_k に関する G 個の線型制約の検定 — (??) 参照：

$$\frac{(\sum \tilde{u}_i^2 - \sum \hat{u}_i^2)/G}{\sum \hat{u}_i^2/(n-k)} \sim F(G,\, n-k)$$

ただし，$\tilde{u}_i^2$ は制約付き残差，$\hat{u}_i^2$ は制約なし残差とする。

(iii) β_1，β_2，$\cdots$，β_k に関する G 個の線型制約の検定 — **(??参照)**：

$$\frac{(\hat{R}^2 - \tilde{R}^2)/G}{(1 - \hat{R}^2)/(n-k)} \sim F(G,\ n-k)$$

ただし，$\tilde{R}^2$ は制約付き決定係数，$\hat{R}^2$ は制約なし決定係数とする。

(iv) **個々の** H_0 : $\theta_j = \theta_j^*$ **の検定 — (1.3) 参照：**

$$n \text{ が大きいとき，} \quad \frac{\hat{\theta}_j - \theta_j^*}{\hat{\sigma}_j} \sim N(0,\ 1)$$

ただし，$\hat{\theta}_j$ **は** θ_j **の最尤推定量，**$\hat{\sigma}_j^2$ **は** $\sigma_j^2 = \mathbf{V}(\hat{\theta}_j)$ **の最尤推定量とする。**

回帰モデルの場合は $\theta = (\beta_1,\ \beta_2,\ \cdots,\ \beta_k,\ \sigma^2)$ **となる。**

(v) G **個の制約の検定 — (1.6) 参照：**

$$n \text{ が大きいとき，} \quad -2\log\frac{l(\tilde{\theta})}{l(\hat{\theta})} \sim \chi^2(G)$$

ただし，$\tilde{\theta}$ は制約付き最尤推定量，$\hat{\theta}$ は制約なし最尤推定量とする。回帰モデルの場合は $\theta = (\beta_1,\ \beta_2,\ \cdots,\ \beta_k,\ \sigma^2)$ となる。

(ii) と **(iii)** は同じもので，**(iii)** は **(ii)** を残差平方和と決定係数との関係を用いて書き換えたものである。**(iv)**，**(v)** について，n が大きいときのみ利用可能で，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，非線型制約の検定）である。

練習問題

6.1. X_1，X_2，$\cdots$，X_n は互いに独立で，それぞれパラメータ p を持ったベルヌイ分布に従うものとする。ベルヌイ分布とは次の確率関数である。

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

x のとる値は **0** か **1** で，X が **1** をとる確率は p となる。すなわち，$P(X=1)=p$，$P(X=0)=1-p$ なので，この **2** つを合わせると上記の確率関数となる。

このとき，次の問いに答えなさい。

(1) X_1，X_2，$\cdots$，X_n の同時確率関数を求めなさい。

(2) **0** と **1** からなるデータ系列 x_1，x_2，$\cdots$，x_n を与えたもとで p の関数（すなわち，尤度関数）を求めなさい。

(3) p の最尤推定量 $\hat{p}$ を求めなさい。

(4) $\hat{p}$ は不偏推定量かどうかを調べなさい。

(5) $\hat{p}$ は有効推定量かどうかを調べなさい。

(6) $\hat{p}$ は一致推定量かどうかを調べなさい。

(7) データ n が多いとき，p の区間推定や仮説検定を行うためには，実践ではどのような分布を用いるか説明しなさい。

6.2. $X_1, X_2, \cdots, X_n$ は互いに独立で，それぞれパラメータ λ を持ったポアソン分布に従うものとする。すなわち，X_i の確率関数は，

$$f(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, \qquad x = 0, 1, 2, \cdots$$

となる。

このとき，次の問いに答えなさい。

(1) λ の最尤推定量 $\hat{\lambda}$ を求めなさい。

(2) $\hat{\lambda}$ は不偏推定量かどうかを調べなさい。

(3) $\hat{\lambda}$ は有効推定量かどうかを調べなさい。

(4) $\hat{\lambda}$ は一致推定量かどうかを調べなさい。

(5) $\hat{\lambda}$ の漸近分布を求めなさい。

6.3. $X_1, X_2, \cdots, X_n$ は互いに独立で，それぞれパラメータ λ を持った指数分布に従うものとする。すなわち，X_i の密度関数は，

$$f(x;\lambda) = \lambda e^{-\lambda x}, \qquad x > 0$$

となる。

このとき，次の問いに答えなさい。

(1) $X_1, X_n, \cdots, X_n$ の結合密度関数を求めなさい。

(2) λ の最尤推定量 $\hat{\lambda}$ を求めなさい。

(3) n が大きいとき，$\hat{\lambda}$ の分布を求めなさい。

6.4. 練習問題 **4.1** で，帰無仮説 $H_0: \beta_2 = \beta_3 = 0$ を尤度比検定で検定しなさい。有意水準は各自設定してよい。

第2章　ミクロ計量分析

2.1　離散選択モデル（二値選択モデル）

データの種類として離散型データと連続型データの 2 種類を区別しなければならないことは，統計学を学ぶ上では，重要な事柄である。データに合わせ

て，確率変数も離散型，連続型の **2** 種類がある。離散型確率変数の代表的な例として，サイコロを振って出た目，コインを投げて表・裏（表の場合は **1**，裏の場合は **0** とする）が考えられる。確率変数 X をサイコロを振って出た目とした場合，X の取る値は **1, 2, 3, 4, 5, 6** のうちの一つで，それぞれの目が出る確率はすべて $\dfrac{1}{6}$ となる。このように不連続な値を取る確率変数を離散型確率変数という。これに対して，ある区間内（$-\infty$ から ∞ の区間も含む）のどの実数値をも取り得る確率変数を連続型確率変数という。

前章までで扱ってきた回帰モデルは，

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

であった。誤差項 u_i は $-\infty$ から ∞ の範囲のどの実数値でも取り得るので，u_i は連続型確率変数であり，Y_i も連続型確率変数になる。

もし被説明変数の Y_i が離散型確率変数であればどのようにすればよいかを考える。一つの例として，アンケート調査で **YES** か **NO** かの **2** つの選択肢から **1** つを選択する。i 番目の人が **YES** を選択した場合に $Y_i = 1$，**NO** を選択した場合に $Y_i = 0$ を割り当てる。このとき，被説明変数 Y_i は **0** か **1** の値を取るが，もともとは **YES** か **NO** かの質的データを表す。このような質的データの従属変数を質的従属変数（**qualitative dependent variable**）と呼ぶ。さらに，被説明変数が質的従属変数とするモデルを離散選択モデル（**discrete choice model**）

と呼ぶ。離散選択モデルでは，アンケートの選択肢が多ければ，選択肢ごとに数字を割り当てて分析することもできる。特に，選択肢が **2** つだけの場合を二値選択モデル（**binary choice model**）と呼ぶ。本章では，二値選択モデルを取り上げる。

2.1.1　例 1：アンケート調査による YES or NO の回答

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

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

このデータのもとで通常の最小二乗法を当てはめるとどうなるか。簡単化のため，通常の次の単回帰モデルを考えてみよう。

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

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。

$\mathbf{E}(u_i) = 0$ なので，Y_i の期待値は，

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

となる。単純に最小二乗法を当てはめると，Y と X に線型関数を仮定すると，$\alpha + \beta X_i$ は $-\infty$ から ∞ の値を取ることになる。この定式化は適切だろうか。まずは，次の数値例に単純に回帰直線を当てはめてみる。

数値例：

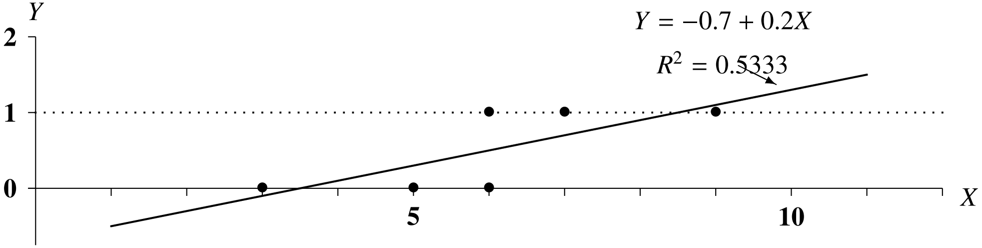
i	**1**	**2**	**3**	**4**	**5**	**6**
Y_i	**0**	**0**	**0**	**1**	**1**	**1**
X_i	**3**	**5**	**6**	**6**	**7**	**9**

このデータを用いて，最小二乗法による切片 α，傾き β を求める。**??**ページの**??**節で紹介した計量ソフト **Gretl** を利用すると，以下の出力結果が得られた。

```
? 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
```

1 行目の「？」に続けて「ols y const x」とタイプして，Enter キーを押す

と，上の推定結果が得られる。この推定結果から得られた切片・傾きの直線と，データをグラフにしたものが下図に描かれている。



直線を推定することが正しいかどうかを考えてみよう。Y_i のデータ作成のルールとしては，i 番目の人が **YES** と答えたとき Y_i に **1** を割り当て，i 番目の人が **NO** と答えたとき Y_i に **0** を割り当てる。このもとで，$\mathbf{E}(Y_i)$ を求める。

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（p とは，この場合，母集団における **YES** の比率，すなわち，母比率を表す）。したがって，$\mathbf{E}(Y_i) = 0 \times P(Y_i = 0) + 1 \times P(Y_i = 1) = P(Y_i = 1) = p$ となり，$\mathbf{E}(Y_i)$ は母集団で **YES** と答える比率 p に等しいので，$0 \leqq p \leqq 1$，すなわち，$0 \leqq \mathbf{E}(Y_i) \leqq 1$ となる。この確率変数 Y_i の分布はベルヌイ分布と呼ばれる（ベルヌイ分布については補論**??**ページ参照）。

$\mathbf{E}(Y_i)$ が母集団の比率 p を表すということは，ゼロより小さくなるということも，**1** より大きくなるということもあり得ない。したがって，$\mathbf{E}(Y_i)$ を $\alpha + \beta X_i$

とするのは不適当であろう。

誤差項 $\boldsymbol{u_i}$ について：　前述のとおり，u_i に正規分布を仮定すると，u_i の取り得る範囲は $-\infty$ から ∞ である。二値選択モデルの場合，誤差項 u_i にどのような特徴があるか考えてみよう。Y_i の期待値は，

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

なので，

$$Y_i = \mathbf{E}(Y_i) + u_i = P(Y_i = 1) + u_i$$

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 $P(Y_i = 1) = p$ で $1 - P(Y_i = 1) = 1 - p$ の値を取るか，確率 $P(Y_i = 0) = 1 - P(Y_i = 1) = 1 - p$ で $-P(Y_i = 1) = -p$ の値を取ることになる。すなわち，誤差項 u_i も $1 - p$ か $-p$ の値を取る離散型確率変数となる。

推定方法について：　Y_i の期待値は p なので，$\mathbf{E}(Y_i) = P(Y_i = 1) = p$ を分布関数 $F(\cdot)$ と関連付けて，説明変数 X_i の関数として $F(\alpha + \beta X_i)$ と表すことが考えられる。

$$P(Y_i = 1) = F(\alpha + \beta X_i) \tag{2.1}$$

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。通常，$F(\cdot)$ には標準正規分布やロジスティック分布が使われる。標準正規分布やロジスティック分布の密度関数，分布関数は次のとおりである（密度関数，分布関数との関係は補論の??ページ参照）。

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

- ロジスティック分布：

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

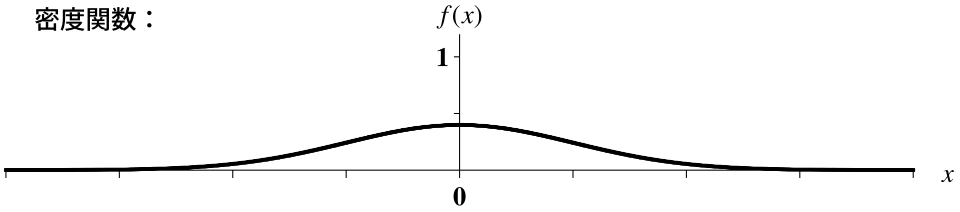
  分布関数： $F(x) = \dfrac{1}{1+\exp(-x)}$

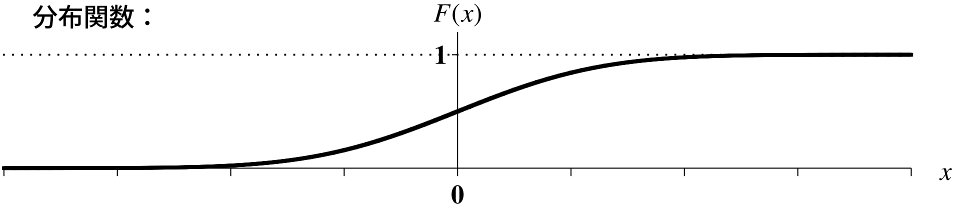
$F(\cdot)$ に他の分布関数を用いてもよい。$F(\cdot)$ に標準正規分布が使われた場合はプロビット・モデル（**probit model**）と呼ばれ，ロジスティック分布が使われた

場合はロジット・モデル（**logit model**）と呼ばれる。例として標準正規分布を取り上げる。その密度関数，分布関数のグラフは次のとおりである。

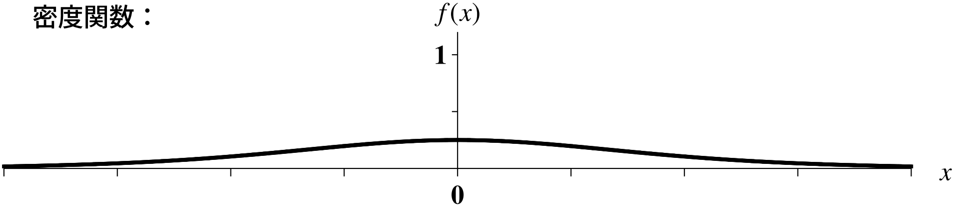
標準正規分布のケース

密度関数：

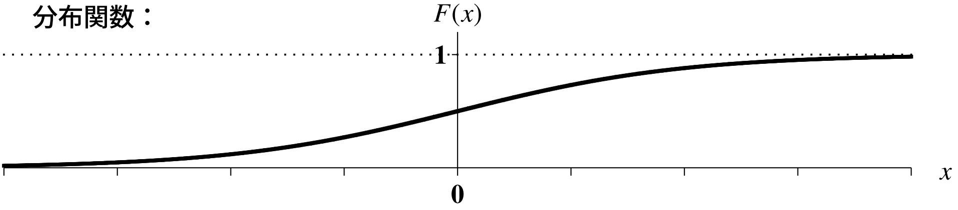




ロジスティック分布のケース



分布関数：



分析者が分布関数 $F(\cdot)$ を特定しなければならない。推定には，最尤法が用いられる。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。よって，Y_1，Y_2，$\cdots$，Y_n の結合確率関

数（同時確率関数）は，

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

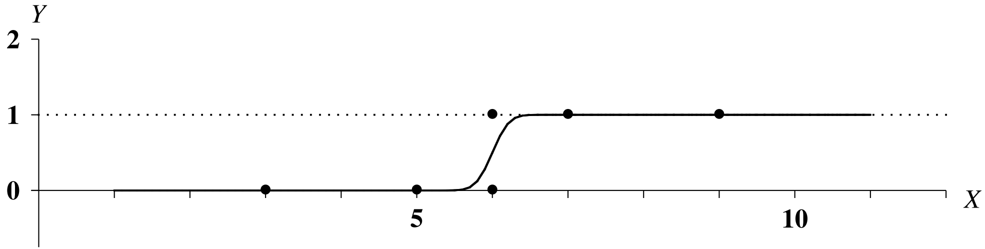
となる。この尤度関数 $l(\alpha,\beta)$ を α，β について最大にする $\hat{\alpha}$，$\hat{\beta}$ を求めることになる。分布関数 $F(\cdot)$ に標準正規分布を仮定した場合，すなわち，プロビット・モデルの場合，次の結果が得られる。

```
? 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
```

1 行目の「?」に続けて「`probit y const x`」とタイプして，Enter キーを押すと，上の推定結果が得られる。

標準誤差について：　最尤推定量の性質として，n が大きいとき，(1.1)，(1.2) が成り立つと説明した。ここでの例は $n = 6$ なので明らかに n は小さいので当てはめることはできないが，上記の推定結果の「標準誤差」は (1.2) 式の推定値の平方根に対応する。そのため,「係数」÷「標準誤差」＝「`z`」という関係が成り立つ。「`z`」と標準正規分布の分布表と比較して，それぞれの係数がゼロであるという帰無仮説の検定ができる。

Y の予測値について：　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 $\hat{Y}_i = F(\hat{\alpha} + \hat{\beta}X_i)$ が図の実線で表されている。



上の図は **83** ページの図と比較することができる。縦軸・横軸の実尺度も同じにしている。標準正規分布を当てはめたものが上の図で，直線を当てはめたものが **83** ページの図である。

また,「「正しく予測された」ケース数 = 5 (83.3%)」は，予測値・実績値の表と密接に関連している。予測値が **0.5** より大きければ **1** を予測し，**0.5** より小さければ **0** を予測する。この数値例は $X = 6$ を中心に対照的にデータが分布している。$X = 6$ のときの Y の予測値は **0.5** となり，**0** にも **1** にも分類できないはずである。しかし，この出力結果からは，$X = 6$ のとき Y の予測値は誤差の範囲で **0.5** より小さいと予測されている。

推定結果の限界効果 **2.31931** の意味：　限界係数は実線の接線の傾きで，

$$\frac{\mathbf{d}\hat{Y}}{\mathbf{d}X} = \frac{\mathbf{d}F(\hat{\alpha} + \hat{\beta}X)}{\mathbf{d}X} = f(\hat{\alpha} + \hat{\beta}X)\hat{\beta}$$

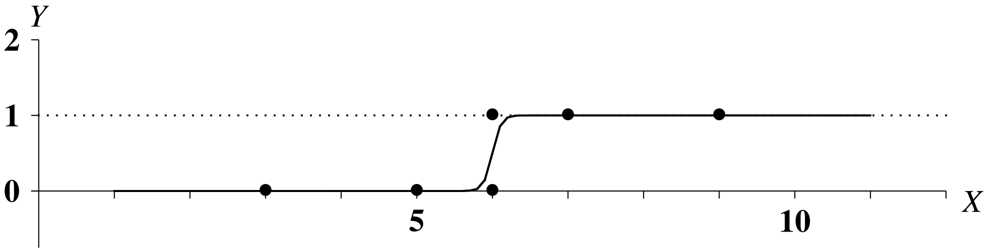
となり，X の値に依存する。

限界効果の **2.31931** とは，限界係数 $\dfrac{\mathbf{d}\hat{Y}}{\mathbf{d}X}$ の X を説明変数の平均値で評価したものである。すなわち，$\dfrac{\mathbf{d}\hat{Y}}{\mathbf{d}X} = \dfrac{\mathbf{d}F(\hat{\alpha}+\hat{\beta}\overline{X})}{\mathbf{d}X} = f(\hat{\alpha}+\hat{\beta}\overline{X})\hat{\beta} = 2.31931$ となる。

また，$F(\cdot)$ にロジスティック分布を仮定すると,「？」に続けて「`logit y const x`」とタイプして，**Enter** キーを押すと，次の推定結果が得られる。

```
? 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
```

Y の予測値 $F(\hat{\alpha}+\hat{\beta}X)$ は次のグラフ内の実線で表される。中心の $X=6$ あたりの傾きは，プロビット・モデル（**95** ページのグラフ）に比べてロジットモデルの方が急になっている。そのため，限界効果がロジット・モデルは **4.42411**，プロビット・モデルは **2.31931** とロジット・モデルの方が大きくなっている。



以上がアンケート調査で被説明変数が **YES** か **NO** の二者択一の選択肢でそ