

計量経済学講義ノート

八木 匡
(同志社大学)

平成 30 年 4 月 1 日

1 統計学の基礎

1.1 記述統計

1.1.1 分散と標本分散（不偏分散）

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^{i=n} (x_i - \mu)^2 \quad (1)$$

標本分散では、標本平均が確率変数 x_i の線形結合で与えられるため、自由度が1つ落ち、不偏分散は、

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{i=n} (x_i - \bar{x})^2 \quad (2)$$

となる。

1.1.2 相関係数

$$\rho = \frac{\sum_{i=1}^{i=n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{i=n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{i=n} (y_i - \bar{y})^2}} \quad (3)$$

相関係数は、 x と y との2変数間の線形関係の強さを示すものであり、 x と y が2次式等の非線形な関係にある場合には、例え関係が強くても、相関係数は大きくなるとは限らない。

1.2 確率関数と確率密度関数

確率変数 x が離散確率変数である時、 x は x_1, x_2, \dots, x_n の値をそれぞれ $f(x_1), f(x_2), \dots, f(x_n)$ の確率で取るものとする。そのとき $f(x)$ を x の確率関数と呼ぶ。

Question：確率関数の性質を述べよ。

離散確率変数の累積分布関数（または分布関数）を

$$F(x) = \sum_{x_i \leq x} f(x_i) \quad (4)$$

Question: 累積分布関数の性質を述べよ。

任意の実数 a, b に対して、

$$P(a \leq x \leq b) = \int_a^b f(x) dx \quad (5)$$

が成立するような関数 $f(x)$ が存在するとき、確率変数 x は連続変数である。

Question: 連続確率変数の例を述べよ。

連続確率変数の平均は、

$$\mu = E(x) = \int_{-\infty}^{\infty} x f(x) dx \quad (6)$$

であり、分散は

$$\sigma^2 = V(x) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \quad (7)$$

で与えられる。

Question: 離散確率変数の平均と分散を記述せよ。

2変量の確率変数において、

$$Cov(x, y) = E[(x - E(x))(y - E(y))] \quad (8)$$

定理 チェビシエフの不等式

x が平均 μ , 分散 σ^2 の確率変数とすると、 λ を任意の正数として、

$$P(|x - \mu| \geq \lambda\sigma) \leq \frac{1}{\lambda^2} \quad (9)$$

が成立する。

1.3 標本理論

1.3.1 標本抽出

無限母集団からの無作為抽出と有限母集団からの無作為復元抽出のケースでの標本平均の平均と分散

$$\mu_{\bar{x}} = \mu \quad (10)$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad (11)$$

統計量の標準偏差のことを標準誤差 (standard error) と呼ぶ。したがって $\sigma_{\bar{x}}$ は標本平均の標準誤差である。

Question: サイコロを5回振り、その平均を計算する。この試行を20回と繰り返したときの標本平均の標準偏差を計算する。サイコロを振る回数を、5回から25回、100回と増やしたときに、標本平均の標準偏差がどのように変化するかを計算しなさい。

1.3.2 正規分布 (normal distribution)

連続確率分布の一つで、密度関数が

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty \quad (12)$$

確率変数の標準化の式、

$$z = \frac{x - \mu}{\sigma} \quad (13)$$

を用いると、平均 μ 、分散 σ^2 の正規分布を平均 0、分散 1 の標準正規分布に変換することができる。標準正規分布表を用いることができる。

Question: 標準正規分布の密度関数を書きなさい。

Question: 制限速度 50km の道路を走る車のスピードを調べてみると、平均速度 65km、標準偏差 15km であった。何パーセントの車が制限速度を超えて走っているかを計算しなさい。

1.3.3 中心極限定理

定理

平均 μ 、分散 σ^2 の無限母集団より無作為に抽出した大きさ n の標本の標本平均を \bar{x}_n とする。いま

$$z_n = \frac{\bar{x}_n - \mu}{\sigma/\sqrt{n}} \quad (14)$$

とするとき、 z_n の分布は n が大きくなるとともに標準正規分布 $N(0, 1)$ に限りなく近づく

Question: ある銀行で、貸付先の企業 500 社の流動比率（流動資産対流動負債の比率）を調査したところ、平均 104.0%、標準偏差 4.5% であった。任意に 25 社を復元抽出するとき、その 25 社の流動比率の平均が 106% を超える確率はいくらか？

1.3.4 χ^2 分布

定理

x_1, x_2, \dots, x_m を相互に独立に正規分布 $N(\mu, \sigma^2)$ に従って分布する m 個の確率変数とするとき

$$\chi^2 = \sum_{i=1}^m \left(\frac{x_i - \mu}{\sigma} \right)^2 \quad (15)$$

は、自由度 m の χ^2 (カイ自乗) 分布 (chi-square distribution with m degrees of freedom) に従って分布する。

標準正規確率変数の分散は 1 であるため、 χ^2 分布の平均は m となり、分散は m^2 となる。

正規母集団から得た大きさ n の標本があったとしよう。標本分散は

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (16)$$

と定義され、標本分散も統計量であることが分かる。この標本分散の分布に関して、次の定理がある。定理

$N(\mu, \sigma^2)$ から得た無作為標本を x_1, x_2, \dots, x_n とする。

$$Y = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (17)$$

は自由度が $n-1$ の χ^2 分布に従って分布する。

すると、標本分散 S^2 が a と b の範囲に入る確率は、 σ 既知の場合、

$$P(\{a \leq S^2 \leq b\}) = P(\{\frac{1}{\sigma^2}(n-1)a \leq \chi^2 \leq \frac{1}{\sigma^2}(n-1)b\}) \quad (18)$$

で与えられる。

S^2 の期待値は、

$$E(S^2) = \frac{1}{n-1} \sigma^2 E(W) = \sigma^2 \quad (19)$$

となり、母分散に一致する（不偏推定量）。

Question: 母集団が正規分布であるとする。 n が 20 の標本から標本分散を求めたところ、その値は 1.5 であった。母分散が 1 なら、標本分散が 1.5 以上となる確率を求めよ。

1.3.5 t 分布

母分散 σ が既知の場合、 n 個の標本から計算した標本平均 \bar{x}_n を標準化した確率変数 z は

$$z = \frac{\bar{x}_n - \mu}{\sigma/\sqrt{n}} \quad (20)$$

で与えられ、正規分布に従う。しかしながら、母分散が既知ではなく、標本分散を用いる場合、

$$t = \frac{\bar{x}_n - \mu}{S/\sqrt{n}} \quad (21)$$

は、分母と分子に確率変数が入るため、正規分布に従わない。

定理

x を平均 0、分散 1 の正規分布に従う確率変数とし、 y を自由度 m の χ^2 分布にしたがう確率変数とする。もし両者が相互に統計的に独立であれば、

$$t = \frac{x}{y/m} \quad (22)$$

は自由度 m のスチューデントの t 分布にしたがって分布する。

$(n-1)S^2/\sigma^2$ が、自由度 $n-1$ の χ^2 分布にしたがうことから、

$$\frac{\frac{\bar{x}_n - \mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2}/(n-1)}} = \frac{\bar{x}_n - \mu}{S/\sqrt{n}} \quad (23)$$

は自由度 $n-1$ のスチューデントの t 分布にしたがう。

1.4 推定

1.4.1 推定量の性質

特定の母数を推定するためにつくられた統計量をその母数の推定量 (estimator) と呼ぶ。推定値 (estimate) とは推定量の 1 つの実現値である。母数 θ の推定量を $\hat{\theta}$ で表す。すなわち、

$$\hat{\theta} = g(x_1, x_2, \dots, x_n) \quad (24)$$

で与えられる。

このとき、

$$E(\hat{\theta}) = \theta \quad (25)$$

となっているとき推定量 $\hat{\theta}$ を θ の不偏推定量 (unbiased estimator) と呼ぶ。2 つの推定量を比較したときに、推定量の分散が小さい方を「より有効 (more efficient)」であるという。あらゆる不偏推定量の中で最小の分散をもつ不偏推定量を有効推定量 (efficient estimator) と呼ぶ。または、最小分散不偏推定量とか最良 (best) 不偏推定量と呼ぶ。

また、

$$\lim_{n \rightarrow \infty} P(|\hat{\theta}_n - \theta| \leq \varepsilon) = 1 \quad (26)$$

であるとき、 $\hat{\theta}_n$ は θ の一致推定量 (consistent estimator) であるという。

Question: 標本平均が、母平均の最良線形不偏推定量であることを確認せよ。(注: 母集団が正規分布でない場合には、標本平均が有効推定量である保証はない。すなわち、線形でない推定量の中で、標本平均よりもより有効な推定量がある可能性がある。)

1.4.2 区間推定

正規確率変数 z が、区間 a, b に入る確率が α であることは、

$$P(\{a \leq z \leq b\}) = \alpha \quad (27)$$

で表現される。

$$z = \frac{\bar{x}_n - \mu}{\sigma/\sqrt{n}} \quad (28)$$

を代入すると、

$$P(\{\bar{x}_n - b\sigma/\sqrt{n} \leq \mu \leq \bar{x}_n - a\sigma/\sqrt{n}\}) = \alpha \quad (29)$$

となる。これが、母分散が既知の場合の母平均に関する区間推定となる。

母分布が正規分布であることが分かっており、母分散が未知の場合には、標本分散を用いた t 分布による区間推定が必要となる。

Question: 東京株式市場の上場会社から 50 社を復元抽出し、利回りの平均と標準偏差を計算したところ、標本平均は 4 分 2 厘、標本標準偏差は 2 分であった。上場会社全体の利回り平均を信頼区間 95% で区間推定しなさい。また、標本数が 25 社であった場合の区間推定もしなさい。

1.5 検定

1.5.1 検定の考え方

「仮説を設定し、仮説が正しいとした時に、実際に現れたデータがほとんど起きえない確率で生起すると計算される場合に、仮説自体が誤りであるという判断をする」という考え方。

仮説には、帰無仮説 H_0 と対立仮説 H_a があり、過誤にも 2 種類存在する。 H_0 が真であるときに、 H_0 を棄却することが第 1 種の過誤であり、 H_0 が真でないときに、 H_0 を採択することが第 2 種の過誤である。

検定方法としては、次の 2 通りがある。

- (1) 危険率を与えた後に、帰無仮説の下での臨界値を計算し、検定統計量が棄却域に入るかどうかを調べる。
- (2) 帰無仮説の下で、検定統計量が生起する確率 (p 値) を計算し、その値が危険率よりも大きいか否かを調べる。

検定統計量は、母分布が正規分布で、分散既知の場合の、平均値検定では、

$$z = \frac{\bar{x}_n - \mu}{\sigma/\sqrt{n}} \quad (30)$$

となり、母分布が正規分布で、分散未知の場合の平均値検定では、

$$t = \frac{\bar{x}_n - \mu}{S/\sqrt{n}} \quad (31)$$

となる。

Question: ある県の中学校 3 年生男子の平均身長は 163.4cm で、標準偏差が 6.99cm であることが分かっている。県下の A 中学校で 108 人の 3 年男子を調べたところ、平均が 164.2cm であった。この中学校の平均は県平均と異なると言えるか？有意水準 5% で検定せよ。

2 線形回帰分析

2.1 単純回帰モデル

第 1 所得階層から第 n 所得階層の総支出・食料品支出データを用いて、食料品への支出関数を推計する。食料品支出関数として、

$$Y_i = \alpha + \beta X_i + u_i, \quad i = 1, \dots, n \quad (32)$$

を考える。

ここで Y は被説明変数あるいは従属変数、 X は説明変数あるいは独立変数と呼ぶ。 u は誤差項である。これは、 X が与えられると、確率的に Y が決まる確率モデルであることを示している。 α と β はモデルを特徴づけている定数であり、パラメータと呼ばれる。各変数の添え字 i は第 i 番目のサンプルであることを示す。データには、時系列データとクロスセクションデータがある。特に時系列データの時には、サンプルの添え字は t を用いる。

最小自乗法によるパラメータの推定は、予測値と実現値の差（残差）の自乗和を最小にするように、パラメータを決定するというものである。

線形回帰モデルにおいて、最小自乗法による推定が BLUE(Best Linear Unbiased Estimator) であるための仮定は次の通りである。

1. $E(u_t) = 0, t = 1, \dots, n$

これは、 $E(Y_t) = \alpha + \beta X_t$ を意味し、 $Y_t = E(Y_t) + u_t$ （観測値が期待値に偶然的変動が加わったもの）であることを意味する。

2. $E(u^2) = \sigma^2, t = 1, \dots, n$

これは分散均一性の仮定である。

Question: 分散均一のケースと不均一なケースの図を描け。

3. $E(u_t, u_s) = 0, t \neq s, t = 1, \dots, n$

これは標本同士が共分散していないことを意味している。時系列データの場合には、系列（自己）相関が無いことを意味している。

4. X は非確率変数である。

実験データであれば、 X が制御可能な変数であり、非確率変数の仮定は妥当と言えるが、

経済データの場合には、コントロール不能な場合が多く、確率変数となる場合が多い。しかし、 X が確率変数でも、 u と統計的に独立であれば、最小自乗法による推定は問題をもたらさない。

5. u は正規分布をする。

1,2,3,5 の仮定から、

$$u_t \sim NID(0, \sigma^2), \quad t = 1, 2, \dots, n \quad (33)$$

と表現できる。NID は、Normally Independently Distributed の意味である。

2.1.1 最小自乗法 (Ordinary Least Squares)

推定量を $\hat{\alpha}$ 、 $\hat{\beta}$ とする。 Y_i の推定値は、

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta}X_i \quad (34)$$

で与えられる。

残差を

$$e_i = Y_i - \hat{Y}_i \quad (35)$$

で表す。最小自乗法での $\hat{\alpha}$ と $\hat{\beta}$ は、

$$\frac{\partial \sum_{i=1}^n e_i^2}{\partial \hat{\alpha}} = 0 \quad (36)$$

$$\frac{\partial \sum_{i=1}^n e_i^2}{\partial \hat{\beta}} = 0 \quad (37)$$

である。

Question: 最小自乗法による推定量を計算しなさい。

2.2 最尤法 (Maximum likelihood method)

$$u_i \sim NID(0, \sigma^2), \quad i = 1, 2, \dots, n \quad (38)$$

であり、 Y_i が u_i の線形関数であることから、

$$Y_i \sim NID(\alpha + \beta X_i, \sigma^2), \quad i = 1, 2, \dots, n \quad (39)$$

が成立する。 Y_i の確率密度関数は、

$$f(Y_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2\sigma^2}(Y_i - \alpha - \beta X_i)^2\right\} \quad (40)$$

となる。 Y_1, Y_2, \dots, Y_n の同時確率密度関数は、 Y_1, Y_2, \dots, Y_n が独立であることから

$$f(Y_1, Y_2, \dots, Y_i, \dots, Y_n; \alpha, \beta, \sigma^2) = f(Y_1)f(Y_2)\dots f(Y_n) \quad (41)$$

$$= (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2\right\} \quad (42)$$

最尤推定の考え方は、実際に観察されたデータ Y_1, Y_2, \dots, Y_n が、どのようなパラメータの組み合わせ α, β, σ^2 の下で最も発生しやすいかを調べるといものである。すなわち、尤度関数は

$$L(\alpha, \beta, \sigma^2; Y_1, Y_2, \dots, Y_i, \dots, Y_n) \quad (43)$$

を意味する。

そこで、

$$L(\alpha, \beta, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2\right\} \quad (44)$$

となる。

最尤法は、尤度関数を最大化するパラメータを求める方法と言える。すなわち、

$$\frac{\partial L}{\partial \alpha} = 0 \quad (45)$$

$$\frac{\partial L}{\partial \beta} = 0 \quad (46)$$

$$\frac{\partial L}{\partial \sigma^2} = 0 \quad (47)$$

を解くことによって求められる。

仮定 1 から 5 を満足する線型モデルでは、 $\hat{\alpha}$ と $\hat{\beta}$ は最小自乗法による推定でも最尤法による推定でも一致する。しかし、最尤法による $\hat{\sigma}^2$ は、最小自乗法よりも過小になることが分かっている。

2.3 決定係数

$$\sum y^2 = \sum (Y_i - \bar{Y})^2 \quad (48)$$

$$\sum \hat{y}^2 = \sum (\hat{Y}_i - \bar{Y})^2 = \sum (\hat{Y}_i - \bar{Y})^2 \quad (49)$$

とおくと、

$$\sum y^2 = \sum \hat{y}^2 + \sum e^2 \quad (50)$$

となる。上式左辺は全変動を表している。ここで $\sum \hat{Y}^2$ はモデルによって説明できる平方和を表し、 $\sum e^2$ は残差平方和であり、モデルで説明できない平方和である。そこで、決定係数は、全変動の内、モデルによって説明できる比率で与えられる。すなわち、

$$R^2 = \frac{\sum \hat{y}^2}{\sum y^2} = 1 - \frac{\sum e^2}{\sum y^2} \quad (51)$$

で与えられる。

決定係数に関して次の点を注する必要がある。

1. 決定係数は、 X が原因変数で Y が結果変数であるという、 $X \rightarrow Y$ の因果関係の強さを測る尺度ではない。これは、 $Y \rightarrow X$ の因果関係を前提としたモデルを作っても、決定係数の値が等しくなることから理解できる。

因果性の方向については、グレンジャーテストを用いて検定できる。直感的な説明は、次の通りである。情報集合 I に基づき確率過程 Y の (1 期先の) 最適予測を定義する。最適性は予測の平均 2 乗誤差の最小性で定義する。今、ある情報集合 I に基づく予測の平均 2 乗誤差と、ある確率変数 X の情報 $I(X)$ を除いた情報 $I - I(X)$ に基づく予測の平均 2 乗誤差を比較する。もし、2 つの平均 2 乗誤差に差がなければ、Granger の意味で X から Y への因果関係がないと考える。もし、 X の情報を用いた方が予測が良くなる、すなわち平均 2 乗誤差が小さくなる場合には、Granger の意味で X から Y への因果関係があると考える。具体的には、モデルの説明変数に X を入れた時の平均 2 乗誤差と入れない時の平均 2 乗誤差を F 検定し、有意性を検定する。

2. 決定係数は線形モデルにおける説明力を測っているのみで、非線形関係で X と Y がどのような関係の強さを持っているかについては述べていない。
3. 自由度が下がると決定係数は増大する。自由度が 0 であれば、決定係数は 1 となる。したがって、通常の計量経済学ソフトは、自由度を調整した、自由度調整済み決定係数を出力する。

2.4 仮説検定

2.4.1 β に関する検定

仮説検定を用いて、 X が Y の系統的要因であるかを検討する。検定すべき帰無仮説は、

$$H_0 : \beta = 0$$

$$H_1 : \beta \neq 0 \quad (\text{または } \beta > 0, \text{ or } \beta < 0)$$

である。

誤差項が正規分布し、 $\hat{\beta}$ が線形結合を通じて決定されるため、 $\hat{\beta}$ も正規分布する。そして、

$$E(\hat{\beta}) = \beta \quad (52)$$

$$\text{var}(\hat{\beta}) = \frac{\sigma^2}{\sum x^2} = \sigma_{\hat{\beta}}^2 \quad (53)$$

すなわち

$$\hat{\beta} \sim N(\beta, \sigma_{\hat{\beta}}^2) \quad (54)$$

であるから、

$$Z = \frac{\hat{\beta} - \beta}{\sigma_{\hat{\beta}}} \sim N(0, 1) \quad (55)$$

となる。 σ が分かっているならば、 $\beta = 0$ は正規分布を用いて検定できるが、通常分らない。そこで、 σ^2 を不偏推定量 S^2 で推定した $\text{var}(\hat{\beta})$ の推定量を

$$S_{\hat{\beta}}^2 = \frac{S^2}{\sum x^2} \quad (56)$$

とおくと、

$$t = \frac{\hat{\beta} - \beta}{S_{\hat{\beta}}} \sim t(n-2) \quad (57)$$

であることが分かっている。

$$H_0 : \beta = 0$$

が正しいとすれば、

$$t = \frac{\hat{\beta}}{S_{\hat{\beta}}} \sim t(n-2) \quad (58)$$

となる。この t が棄却域に入れば、帰無仮説は棄却され、系統的要因であるという仮説が支持されることになる。

2.5 有意確率 (p 値)

p 値は、帰無仮説の下で、データが生起する確率を表し、

$$p = P(t \geq t_{\hat{\beta}}) + P(t \leq -t_{\hat{\beta}}) \quad (59)$$

で与えられる。これは、どの程度の危険率で帰無仮説を棄却できるかを示す値となる。

3 重回帰モデル

3.1 多変量線形回帰モデルの定式化

被説明変数を y , K 個の説明変数を x_1, x_2, \dots, x_K とする。この時、多変量線形回帰モデルは、

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_K x_{iK} + u_i, \quad i = 1, \dots, n \quad (60)$$

で与えられる。ここで、すべての i について、 $x_{i1} = 1$ である時に、定数項付き多変量線形回帰モデルと呼ぶ。

上式を、ベクトル表現すると、

$$y_i = \mathbf{x}_i' \boldsymbol{\beta} + u_i \quad (61)$$

この時、最小自乗法による推定量は次のように計算される。

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (62)$$

となる。ここで、 $\mathbf{X}'\mathbf{X}$ は分散共分散行列と呼ばれる。

なお、重回帰モデルにおける帰無仮説 $H_0 : \beta_k = 0$ に対する検定統計量は、

$$t = \frac{\hat{\beta}_k - \beta_k}{S_k} \sim t(n - K) \quad (63)$$

となる。ただし、ここで K には、定数項 $X_1 = 1$ も含まれている。定数項を変数に入れないのであれば、標本数から変数の数プラス 1 を引いたのが、自由度となる点に注意する必要がある。

Question: 中立的技術進歩を入れた生産関数をマクロデータから推計するための、推定式を定式化しなさい。また、推定において、どのようなデータを用いるかを述べなさい。

3.2 多重共線性

説明変数間で線形関係が強い場合に、分散共分散行列が 0 に近づき、推定量の分散は大きくなり、大きく変動する。

$$\text{var}(\hat{\beta}_k) = \frac{\sigma^2}{\sum X_k^2} \left(\frac{1}{1 - R_k^2} \right) \quad (64)$$

ここで R_k^2 は、 X_k の、定数項および X_k 以外の説明変数への線形回帰を行った時の決定係数を表す。決定係数が 1 であれば、分散が無限大になる。従って、分散拡大要因 (Variance inflation factor: VIF) を

$$VIF_k = \frac{1}{1 - R_k^2} \quad (65)$$

を多重共線性の指標として考えることが出来る。一つの目安として、VIF が 10 を超えると、多重共線性の危険性が強いと判断できる。

3.3 分散分析

重回帰モデル

$$y_i = \beta_1 + \beta_2 x_{i2} + \dots + \beta_K x_{iK} + u_i, \quad i = 1, \dots, n \quad (66)$$

において、帰無仮説を

$$H_0 : \beta_2 = \beta_3 = \dots = \beta_K = 0$$

あるいは

$$H_0 : \beta_2 = \beta_3 = 0$$

のように、2個以上の β_k が同時に0という帰無仮説を、

$$H_1 : \text{少なくとも1つの}\beta\text{は0ではない}$$

という対立仮説に対して検定する問題を考える。

全変動の分解は、

$$\sum y^2 = \sum \hat{Y}^2 + \sum e^2 \quad (67)$$

となる。この式に対応して、自由度 df も

$$n - 1 = (k - 1) + (n - k)$$

と分解される。 y_1, y_2, \dots, y_n は、 $\sum y = 0$ という制約によって自由度は1失われ $n - 1$ となる。

e_1, e_2, \dots, e_n は、

$$\sum e = 0$$

および残差自乗和最小化の条件である

$$\sum eX_k = 0, \quad k = 2, \dots, K$$

の K 個の制約によって自由度が K 失われ、 $n - K$ になる。平方和を自由度で割った

$$\sum \hat{y}^2 / (K - 1)$$

$$\sum e^2 / (n - K)$$

は平均平方と呼ばれる。このとき、仮説

$$H_0 : \beta_2 = \beta_3 = \dots = \beta_K = 0$$

$$H_1 : \text{少なくとも1つの}\beta\text{は0ではない}$$

の検定は、 H_0 が正しいとき、2つの平均平方の比は F 分布をする。すなわち、

$$F = \frac{\sum \hat{y}^2 / (K - 1)}{\sum e^2 / (n - K)} \sim F(K - 1, n - K) \quad (68)$$

を用いて行うことが出来る。

もし、 H_0 が正しいければ、説明変数は Y の変動を説明する力を持っていないため、モデルで説明できる平方和 $\sum \hat{y}^2$ は 0 に近づき、 F は 0 に近づく。逆に、 F が 0 に近いほど、帰無仮説を棄却できないことになる。

危険率を与えると、 $F(K - 1, n - K)$ の臨界点を計算できるため、 F 値がこの値を超えていれば帰無仮説を棄却することができる。

3.4 ダミー変数

説明変数に、連続的な値をとることができない変数を含める場合には、ダミー変数を定義して分析することが出来る。

Question: 学歴ダミーを定義しなさい。

Question: 賃金関数を推計するための推計モデルを定式化しなさい。

4 自己相関

4.1 自己相関

回帰モデルの誤差項 u が

$$u_t = \rho u_{t-1} + \varepsilon_t \quad (69)$$

に従って変化するとき、 u は 1 階の自己回帰過程に従う、あるいは簡単に AR(1) に従うという。AR(1) とは first-order autoregressive processes の略である。

誤差項 u が AR(1) に従っているとき、

$$\begin{aligned} u_t &= \rho(\rho u_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \rho^2 u_{t-2} + \rho \varepsilon_{t-1} + \varepsilon_t \\ &= \rho^2(\rho u_{t-3} + \varepsilon_{t-2}) + \rho \varepsilon_{t-1} + \varepsilon_t \\ &= \rho^3 u_{t-3} + \rho^2 \varepsilon_{t-2} + \rho \varepsilon_{t-1} + \varepsilon_t \\ &\dots \\ &= \varepsilon_t + \rho \varepsilon_{t-1} + \rho^2 \varepsilon_{t-2} + \dots \end{aligned} \quad (70)$$

$$\varepsilon_{t+s} = \begin{cases} 1, & s = 0 \\ 0, & s \neq 0 \end{cases} \quad (71)$$

の場合、

$$\begin{aligned} u_t &= \varepsilon_t = 1 \\ u_{t+1} &= \varepsilon_{t+1} + \rho\varepsilon_t + \rho^2\varepsilon_{t-1} + \dots = \rho \\ u_{t+2} &= \varepsilon_{t+2} + \rho\varepsilon_{t+1} + \rho^2\varepsilon_t + \dots = \rho^2 \\ &\dots \\ u_{t+s} &= \rho^s \end{aligned} \quad (72)$$

となるから、 ε が t 期にのみ 1 単位変化したとき、 s 期後には ρ^s の影響が u に現れ、当期から s 期までの累積効果は

$$\sum_{j=0}^s \rho^j \quad (73)$$

となることがわかる。

上式から示されるように、 u が AR(1) に従う場合でも、 $|\rho| > 1$ のケースでは発散してしまうため考えない。このことから、 u_t が AR(1) に従うとき

$$|\rho| < 1 \quad (74)$$

を仮定し、さらに

$$\varepsilon_t \sim NID(0, \sigma_\varepsilon^2) \quad (75)$$

と仮定する。この時、

$$\begin{aligned} \text{var}(u_t) &= E(u_t^2) = E(\varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2})^2 \\ &= \sigma_\varepsilon^2(1 + \rho^2 + \rho^4 + \dots) = \frac{\sigma_\varepsilon^2}{1 - \rho^2} \end{aligned} \quad (76)$$

となる。

4.2 自己相関の問題

自己相関が発生している場合、推定量は次の問題を持つ。

1. $E(\hat{\beta}) = \beta$ であり、不偏性は満足するが、 $\hat{\beta}$ は β の最良線形不偏推定量 BLUE ではない。BLUE は一般化最小 2 乗推定量で与えられる。従って、最小自乗法 (OLS) による推定において t 検定量は過大となる。

2. σ_u^2 は $s^2 = \sum e^2 / (n - 1)$ によって推定されるが、これは不偏推定量ではなく、過小な値を与える。

4.3 自己相関の検定

自己相関の検定方法として、「ダービン・ワトソン検定」がある。ダービン・ワトソン検定量は、

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \quad (77)$$

e は OLS の残差である。計算を繰り返すことによって、

$$DW \approx 2(1 - \hat{\rho}) \quad (78)$$

という近似式を得る。帰無仮説 $H_0 : \rho = 0$ が真であれば、 $\hat{\rho}$ も 0 に近い値を取り、DW は 2 に近い値をとる。すなわち、DW が 2 に近ければ、帰無仮説を棄却できない。 $\hat{\rho}$ が 1 に近づくと、DW は 0 に近づくことになり、 $\rho > 0$ を強く示唆し、帰無仮説を棄却することになる。 ρ が -1 に近いという負の強い 1 階の自己相関の時、 $\hat{\rho}$ も -1 に近づき、DW は 4 に近い値を取る。この場合も帰無仮説を棄却することになる。

帰無仮説が真である時に、DW の分布は、誤差項の分布、標本数、変数の数によって変化する。従って、臨界点は一定の値を持たない。理論的には、説明変数の観測値が変わる度に、DW の分布を再計算すれば、正確な臨界値を求めることができるが、実際にはできない。そこで、DW が 2 に近い場合に自己相関が無く、0 または 4 に近いときにそれぞれ正および負の自己相関にあると判断する。

注意点として、説明変数に被説明変数の 1 期前の値が入ると、DW は 2 に近い値を取るような偏りが生じる。すなわち、自己相関は無いと判断する傾向が発生する。その場合には、m テストまたは h 統計量を用いる。

4.4 一般化最小自乗法

一般に、 \mathbf{u} の共分散行列は、

$$E(\mathbf{u}\mathbf{u}') = \frac{\sigma_\varepsilon^2}{1 - \rho^2} \begin{pmatrix} 1 & \rho & \dots & \rho^{n-1} \\ \rho & 1 & \dots & \rho^{n-2} \\ \cdot & \cdot & \cdot & \cdot \\ \rho^{n-1} & \rho^{n-2} & \dots & 1 \end{pmatrix} = \sigma_u^2 \Omega \quad (79)$$

と表すことができる。そこで、誤差項が $NID(0, \sigma_\varepsilon^2)$ になるようにプレイス・ウェスティン変換 (PW 変換) を施す。変換行列を \mathbf{T} とすると、変換後の誤差項 \mathbf{u}^* は、

$$\mathbf{u}^* = \mathbf{T}\mathbf{u} \quad (80)$$

変換行列を用いて、 \mathbf{X} と \mathbf{y} も

$$\mathbf{X}^* = \mathbf{T}\mathbf{X} \quad (81)$$

$$\mathbf{y}^* = \mathbf{T}\mathbf{y} \quad (82)$$

のように変換する。この変換後の説明変数と被説明変数を用いると、

$$E(\mathbf{u}^* \mathbf{u}^{*\prime}) = \sigma_\varepsilon^2 \mathbf{I} \quad (83)$$

が成立し、この変換後変数を用いた OLS 推定量

$$\beta^* = (\mathbf{X}^{*\prime} \mathbf{X}^*)^{-1} \mathbf{X}^{*\prime} \mathbf{y}^* \quad (84)$$

は BLUE となる。 $\mathbf{T}'\mathbf{T} = (\mathbf{1} - \rho^2)\boldsymbol{\Omega}^{-1}$ の関係を用いて、上式を変形すると、変換前変数の形で表現された推定量

$$\beta^* = (\mathbf{X}'(\sigma_\varepsilon^2 \boldsymbol{\Omega})^{-1} \mathbf{X})^{-1} \mathbf{X}'(\sigma_\varepsilon^2 \boldsymbol{\Omega})^{-1} \mathbf{y} \quad (85)$$

は一般化最小自乗推定量 (GLSE) と呼ばれる。実際には ρ は未知であるため、 $\hat{\rho}$ を推定して GLS を適用する。これは β の分散を最小化し、最良線形不偏推定量を与える。

5 不均一分散

分散均一性の仮定が成立していない場合、推定量は不偏であるものの、最良とならない。そのため、OLS による t 検定量は正しい検定量とならない。

5.1 均一分散の検定

分散均一性の検定は、ブレイシュ・ペーガンテスト (Breusch and Pagan:BP テスト) およびゴッドフライ・コーエンカーテスト (Godfrey and Koenker : GK テスト)、およびホワイト (White) 不均一分散テストによって行う。ホワイトの不均一分散テストの検定統計量は、補助回帰式を作り、その決定係数に標本数をかけたものとなる。これが χ^2 自乗分布に従うことを用いて、右肩側検定を行う。棄却域に入れば、帰無仮説である均一分散が棄却されることになる。すなわち、ホワイト統計量が小さい値を取っていれば、均一分散であると考えて良いことになる。BP

についても、検定統計量が小さい値を取っており、 χ^2 自乗分布の棄却域に入らなければ、均一分散であると判断できる。

BP テストは誤差項の正規性の仮定のもとでのみ成立する。正規性の仮定が成立していれば BP テストの方が、ホワイトの不均一分散テストよりも不均一分散を検出する力（検定力）が高いことが知られている。しかし、正規性の仮定が怪しい場合には、ホワイトの不均一分散テストを用いる方が良い。難点は、説明変数が多い場合に、信頼度が低下することである。正規性の仮定は余り崩れないと考えれば、BP テストが望ましいと言える。

5.2 不均一分散への対処

1. 関数形の変更、定式化の変更を試みる。関数形の変更として、対数変換があり得る。これは、被説明変数 Y を対数変換する方法である。これによって、 $E(Y_t)$ の水準が高くなるに従って Y_t の標準誤差が大きくなる場合に有効である。
2. 上記の変更を行っても、尚不均一分散がある場合に、 σ_t^2 の型を見 $\hat{\sigma}_t$ を求め、GLS を適用する。この場合、不均一性のパターンを見つけ、変換行列を定式化することが必要となる。

6 同時方程式

モデルの中で決定される変数を内生変数と呼ぶ。内生変数を外生変数、先決内生変数および誤差項で表した式は誘導形 reduced form と呼ばれる。例えば、

$$C_t = \alpha + \beta Y_t + u_t \quad (86)$$

$$Y_t = C_t + Z_t \quad (87)$$

というモデルを考える。このモデルの内生変数は C （消費）と Y （GNP）の2個、独立支出 Z は外生変数とする。 Y_t を1番目の式に代入し、 C_t について解くと、 C の誘導形

$$C_t = \pi_{10} + \pi_{11} Z_t + v_{1t} \quad (88)$$

を得る。

Question: π を計算しなさい。次に、 Y の誘導形を導出しなさい。

このように、始めに示した構造方程式は、モデルの動きを示しているのに対し、誘導形では、モデルの動きを理解できない。しかし、先決変数、外生変数が内生変数にどのような影響を与えるかについては、理解できる。

6.1 識別

同時方程式によって経済分析をする時、2つの重要な問題が生じる。一つは識別 identification の問題と、もう一つは同時方程式バイアスの問題である。

ある財に対する価格 P と取引量 Q の観測データ $(P_1, Q_1), \dots, (P_n, Q_n)$ があるとしよう。実際に観測されるデータは、市場均衡点である。しかしながら、この均衡がどのようなモデルの均衡であるかはモデルの定式化に依存して決まる。例えば、

$$Q_t^d = \alpha - \alpha_1 P_t + u_{1t} \quad (89)$$

$$Q_t^s = \beta_0 + \beta_1 P_t + u_{2t} \quad (90)$$

$$Q_t^d = Q_t^s \quad (91)$$

$Q_t^d = Q_t^s = Q_t$ とおくと、モデルの誘導形は

$$Q_t = \pi_{10} + v_{1t} \quad (92)$$

$$P_t = \pi_{20} + v_{2t} \quad (93)$$

となる。

Question: π と v を計算しなさい。

この誘導形から分かるように、誤差項の変動を除くと、モデルの均衡点は1点のみできまり、需要関数も供給関数も共にもたらさない。すなわち、識別不能である。

次に、モデルを次のように変形する。

$$Q_t^d = \alpha - \alpha_1 P_t + \alpha_2 Y_t + u_t \quad (94)$$

$$Q_t^s = \beta_0 + \beta_1 P_t + u_{2t} \quad (95)$$

$$Q_t^d = Q_t^s \quad (96)$$

これは、需要関数が所得によってシフトすることを述べている。誘導形は、

$$Q_t = \pi_{10} + \pi_{11} Y_t + v_{1t} \quad (97)$$

$$P_t = \pi_{20} + \pi_{21} Y_t + v_{2t} \quad (98)$$

となる。

Question: π と v を計算しなさい。

このモデルにおいては、需要曲線が Y の変動によってシフトし、実現された均衡点がシフトすることを示している。すなわち、均衡点を結んだものが供給曲線となる。このモデルから供給曲線が識別される。誘導形を連立させて解いた式は、供給曲線を表すことになる。

しかし、

$$Q_t^d = \alpha - \alpha_1 P_t + \alpha_2 Y_t + \alpha_3 W_{t-1} + u_t \quad (99)$$

$$Q_t^s = \beta_0 + \beta_1 P_t + u_{2t} \quad (100)$$

$$Q_t^d = Q_t^s \quad (101)$$

のように、需要が、所得と資産によってシフトするモデルとなると、誘導形は、

$$Q_t = \pi_{10} + \pi_{11} Y_t + \pi_{12} W_{t-1} + v_{1t} \quad (102)$$

$$P_t = \pi_{20} + \pi_{21} Y_t + \pi_{22} W_{t-1} + v_{2t} \quad (103)$$

となり、 W_{t-1} を消去して Y 変化に対する供給曲線を求める方法と、 Y を消去して W_{t-1} 変化に対する供給曲線を求める方法と 2 通りの方法で得られ、過剰識別される。

識別の問題を考慮すれば、モデル構築の時に、理論的に整合的にモデルがデータを生成できるかを示すものである。識別条件として

1. $R = G - 1$ および $rank(A) = G - 1 \rightarrow$ 正確に識別
2. $R > G - 1$ および $rank(A) = G - 1 \rightarrow$ 過剰識別
3. $R \geq G - 1$ および $rank(A) < G - 1 \rightarrow$ 識別不能
4. $R < G - 1 \rightarrow$ 識別不能

ここで、

R = 係数に関する先験的制約の数

G = モデル全体の内生変数の数である。係数に関する先験的制約とは、

1. モデルに含まれるが、係数が 0 となる変数が何組あるか
2. 構造パラメータに関する 1 次同時制約の数

である。行列 A は、(i) モデルの他の構造方程式には現れるが、問題にしている構造方程式には含まれない変数のリストを作る。(ii)(i) で取り上げた変数が、問題にしている構造方程式以外の $G - 1$ 本の方程式においてどのような係数をもつかを行列にする。この行列が A である。

6.2 同時方程式バイアス

識別可能な構造方程式から成るモデルが作られたとき、パラメータ推定における問題が発生する。

$$C_t = \alpha + \beta Y_t + u_t \quad (104)$$

$$Y_t = C_t + Z_t \quad (105)$$

というモデルにおいて、誘導形は

$$C_t = \pi_{10} + \pi_{11} Z_t + v_{1t} \quad (106)$$

と

$$Y_t = \pi_{20} + \pi_{21} Z_t + v_{2t} \quad (107)$$

であり、消費関数は正確に識別できる。問題は、 $u \rightarrow C \rightarrow Y$ のルートで、誤差項が説明変数 Y と相関を持つことになる。すなわち、説明変数が独立変数であるという仮定が崩れることになる。このため、OLS による推定量は、不偏性も一致性も持たなくなる。不偏推定量からの乖離は、同時方程式バイアスと呼ばれる。

6.2.1 2 段階最小自乗法 (2SLS)

同時方程式バイアスが存在する推定式を推定する方法として、2 段階最小自乗法 (2SLS) がある。

第 1 段階

パラメータ推定をしようとしている構造方程式の説明変数に、その方程式の被説明変数と同時決定される同時内生変数があるとき、OLS を用いてその同時内生変数の誘導形推定値を求める。上モデルのケースでは、説明変数 Y は C と同時決定される内生変数であるから、 Y の誘導形推定値

$$\hat{Y}_t = \hat{\pi}_{20} + \hat{\pi}_{21} Z_t \quad (108)$$

を求める。

第 2 段階

同時内生変数である説明変数の値を、第 1 段階の誘導形推定値で置きかえ、OLS を適用する。このモデルのケースでは、

$$\begin{aligned} C_t &= \alpha + \beta(\hat{Y}_t + e_t) + u_t \\ &= \alpha + \beta\hat{Y}_t + (\beta e_t + u_t) \end{aligned} \quad (109)$$

と、 C の \hat{Y} への回帰を行うことに等しい。ここで e は第 1 段階において \hat{Y} を計算したときの残差 $Y_t - \hat{Y}_t$ である。

このような2段階を経て得られる推定量を2段階最小自乗推定量(2SLS)という。

この2SLSは次の性質を持つ。

1. 一致性を持つ。
2. 不偏性は持たない。
3. 一致性を持つ理由は、漸近的に \hat{Y}_t と u_t が独立となることから、 \hat{Y}_t と $\beta e_t + u_t$ が漸近的に無相関となるからである。

7 質的選択モデル (Models with discrete dependent variables)

7.1 質的選択モデルとは

本章では、被説明変数が0,1,...のように、離散変数となる場合のモデルを扱う。被説明変数が離散変数となるのが、主として離散変数が質的な変数である場合であることから、質的選択モデルという呼び方をする。例として次のものが挙げられる。

1. 労働供給の意志決定
2. 投票行動
3. 職業選択
4. 通勤手段の選択
5. 株式の購入の有無
6. ショッピング場所の選択

これらの問題は、一般的には、

$$\begin{aligned} \text{Prob}(j \text{ というイベントが生起}) &= \text{Prob}(Y = j) \\ &= F(\text{イベント生起を説明する諸変数 : パラメータ}) \end{aligned}$$

特に2値(binomial)選択モデルがしばしば用いられており、説明の中心もこの2値選択モデルに置くことにする。

7.2 2値選択モデル:回帰分析アプローチ

被説明変数が、0または1のどちらかの値を取るモデルを2値選択モデル(Binary Choice Model)と呼ぶ。

まず、労働供給行動を例として考える。アンケート調査を行い、回答者が調査が行われた日から1週間以内に「働いている、または職を探している」と回答する場合には $Y = 1$ とし、そ

れ以外を $Y = 0$ とおく。労働供給行動を説明する変数として、年齢、性別、婚姻の有無、学歴、就業経験等を考える。これらの変数をベクトル \mathbf{x} で表すとする。パラメータのベクトルを β で与えると、

$$Prob(Y = 1) = F(\beta' \mathbf{x}) \quad (111)$$

$$Prob(Y = 0) = 1 - F(\beta' \mathbf{x}) \quad (112)$$

となる。パラメータ β は、 \mathbf{x} が変化したときに確率値に与える影響を表している。

この影響を表す定式化として、

$$F(\mathbf{x}, \beta) = \beta' \mathbf{x} \quad (113)$$

があり得る。すなわち、

$$y = \beta' \mathbf{x} + \mathbf{u} \quad (114)$$

という線形モデルが考えられる。しかしながら、 $\beta' \mathbf{x} + \mathbf{u}$ が 0 または 1 をとるということは、 u が $1 - F$ または F の確率で $-\beta' \mathbf{x}$ または $1 - \beta' \mathbf{x}$ を取ることになり、 u が β に依存しながら不均一分散となる。すなわち、

$$var(u|\mathbf{x}) = \beta' \mathbf{x}(1 - \beta' \mathbf{x}) \quad (115)$$

となる。さらに、より重要な問題として、推定式を用いた予測値が 0 と 1 の間に必ずしも入らないことがある。この理由によって、2 値選択問題の推定で線型モデルを用いることはほとんど行われない。

予測値が必ず 0 と 1 の間に入るためには、

$$\lim_{\beta' \mathbf{x} \rightarrow \infty} Prob(Y = 1) = 1 \quad (116)$$

$$\lim_{\beta' \mathbf{x} \rightarrow -\infty} Prob(Y = 1) = 0 \quad (117)$$

が成立すれば良い。そこで、この性質を満たす関数 F として正規分布を用いたのがプロビットモデル (Probit model) である。プロビットモデルは、

$$\begin{aligned} Prob(Y = 1) &= \int_{-\infty}^{\beta' \mathbf{x}} \phi(t) dt \\ &= \Phi(\beta' \mathbf{x}) \end{aligned} \quad (118)$$

で与えられる。ここで、 ϕ は正規分布の密度関数であり、 Φ は正規分布の累積密度関数である。次に、数学的簡便性から、ロジスティック分布 (Logistic distribution) を用いたモデルが、ロジットモデル (Logit Model) である。それは

$$\begin{aligned} Prob(Y = 1) &= \frac{e^{\beta' \mathbf{x}}}{1 + e^{\beta' \mathbf{x}}} \\ &= \Lambda(\beta' \mathbf{x}) \end{aligned} \quad (119)$$

で与えられ、実務的にもしばしば用いられる。

ロジスティック分布は、正規分布と形状が似ているが、裾野は正規分布よりも厚くなっており、自由度7のt分布にほぼ近似できる。理論的にどの問題に対してどちらの分布を用いるべきかについては、何も言うことができない。

$$\begin{aligned} E[y|\mathbf{x}] &= 0[1 - F(\beta'\mathbf{x})] + 1[F(\beta'\mathbf{x})] \\ &= F(\beta'\mathbf{x}) \end{aligned} \quad (120)$$

である。そこで、上式を用いて、説明変数が限界的に変化したときに、 y の期待値がどのように限界的に変化するかを分析する。そこで、

$$\begin{aligned} \frac{\partial E[y|\mathbf{x}]}{\partial \mathbf{x}} &= \left\{ \frac{dF(\beta'\mathbf{x})}{d\beta'\mathbf{x}} \right\} \beta \\ &= f(\beta'\mathbf{x})\beta \end{aligned} \quad (121)$$

となる。ここで f は確率密度関数であり、プロビットモデルであれば正規確率密度関数が対応する。ロジットモデルでは、

$$\begin{aligned} \frac{d\Lambda[\beta'\mathbf{x}]}{d\mathbf{x}} &= \frac{e^{\beta'\mathbf{x}}}{(1 + e^{\beta'\mathbf{x}})^2} \\ &= \Lambda(\beta'\mathbf{x})[1 - \Lambda(\beta'\mathbf{x})] \end{aligned} \quad (122)$$

と計算できることから、

$$\frac{\partial E[y|\mathbf{x}]}{\partial \mathbf{x}} = \Lambda(\beta'\mathbf{x})[1 - \Lambda(\beta'\mathbf{x})]\beta \quad (123)$$

となる。

7.3 2値選択モデル：ランダム効用モデル

経済主体行動を反映する形で、モデルを構築する。例えば、消費行動を考え、購入するか否かの決定を、限界効用がある基準値以上であれば購入し、そうでなければ購入しないという決定を下すと考える。限界効用は観察不能であるが、この観察不能な値を y^* で与え、

$$y^* = \beta'\mathbf{x} + \mathbf{u} \quad (124)$$

で決定されていると考える。誤差項 \mathbf{u} は、平均0、分散1のロジスティック分布か正規分布にしたがって分布しているとする。観察可能なのは、購入したか否かのみである。そこで、購入したときに $y = 1$ 、購入しなかったときに $y = 0$ というインデックスを与えることにする。す

なわち、

$$y = 1 \text{ if } y^* > 0 \quad (125)$$

$$y = 0 \text{ if } y^* \leq 0 \quad (126)$$

と定式化する。この定式化において、 $\beta' \mathbf{x}$ はインデックス関数と呼ばれる。平均 0，分散 1 の仮定は、このモデルにおいて定数項を変化させるのみであるため、本質的な仮定とはなっていない。

そこで、 $y = 1$ となる確率は、

$$\begin{aligned} \text{Prob}(y^* > 0) &= \text{Prob}(\beta' \mathbf{x} + \mathbf{u} > 0) \\ &= \text{Prob}(u > -\beta' \mathbf{x}) \end{aligned} \quad (127)$$

となる。F が正規分布またはロジスティック分布であるときには、対称形をしているため、

$$\begin{aligned} \text{Prob}(y^* > 0) &= \text{Prob}(u < \beta' \mathbf{x}) \\ &= F(\beta' \mathbf{x}) \end{aligned} \quad (128)$$

この定式化を別の視点で表現すると次のようになる。a という選択肢を採ったときの効用を

$$U^a = \beta'_a \mathbf{x} + \mathbf{u}_a \quad (129)$$

で与え、b という選択肢を採ったときの効用を

$$U^b = \beta'_b \mathbf{x} + \mathbf{u}_b \quad (130)$$

で与える。 $U^a > U^b$ の時に $y = 1$ を与え、 $U^a < U^b$ の時に $y = 0$ を与えたとする。すると、 $y = 1$ を取る確率は、

$$\begin{aligned} \text{Prob}(y = 1 | \mathbf{x}) &= \text{Prob}(U^a > U^b) \\ &= \text{Prob}(\beta'_a \mathbf{x} + \mathbf{u}_a - \beta'_b \mathbf{x} - \mathbf{u}_b | \mathbf{x}) \\ &= \text{Prob}((\beta_a - \beta_b)' \mathbf{x} + \mathbf{u}_a - \mathbf{u}_b > 0 | \mathbf{x}) \\ &= \text{Prob}(\beta' \mathbf{x} + \mathbf{u} > 0 | \mathbf{x}) \end{aligned} \quad (131)$$

と、全く同一の定式化になる。この定式化をランダム効用モデルと呼ぶ。

8 パネル分析

8.1 パネルデータ

同一経済主体について複数時点のデータを集めたものをパネルデータという。ある経済主体 i について t 時点のデータ y が得られるとき、それを y_{it} と表記する。このパネルデータを用いた分析をパネル分析 (panel analysis) という。

パネル分析のメリットは、

1. 経済主体間の異質性をコントロールできる。これによって、より複雑な問題を扱うことが出来る。
2. サンプル数が増え、自由度が増える。
3. 金利のように、クロスセクションでは差が生じない変数も利用できる。
4. 動学的な問題を扱うことができる。例として、地域別の失業率変化、地価変動の問題。

8.2 パネルデータの OLS 推計

$$y_{it} = a + bx_{it} + e_{it} \quad (132)$$

を考える。 $e_{it} \sim IID(0, \sigma^2)$ であり、 a, b が時間を通じ一定のかつ経済主体を通じて一定であれば、単純に OLS 推定を行うことが出来る。このときパネルデータの特異性や経済主体特有の効果は考えられない。

次に、

$$y_{it} = a + bx_{it} + e_{it} \quad (133)$$

を考え、誤差項について

$$e_{it} = \alpha_i + v_{it} \quad (134)$$

という構造を考える。 x_{it} は v_{it} と相関しない。 v_{it} は標準的線形回帰モデルの仮定を満たす誤差項とする。すなわち、 $E(v_{it}) = 0$, $E(v_{it}^2) = V(v_{it}) = \sigma_v^2$, $E(v_{it}v_{js}) = Cov(v_{it}, v_{js}) = 0$ for $i \neq j, t \neq s$ とする。ここで α_i は経済主体特有の効果 (individual effect) と呼ばれているものである。

α_i と説明変数 x_{it} が無相関であれば変量効果モデルあるいはランダム・イフェクトモデル (random effect model) という。これに対して α_i と x_{it} が相関していれば固定効果モデル (fixed effect model) という。具体的には、ランダム・イフェクトモデルは、

$$E(\alpha_i x_{it}) = Cov(\alpha_i, x_{it}) = 0 \quad (135)$$

で与えられ、固定効果モデルは、

$$E(\alpha_i x_{it}) = Cov(\alpha_i, x_{it}) \neq 0 \quad (136)$$

で与えられる。

8.3 固定効果モデル

固定効果モデルでは、

$$E(\alpha_i x_{it}) = Cov(\alpha_i, x_{it}) \neq 0 \quad (137)$$

であることより、

$$y_{it} = a + bx_{it} + e_{it} \quad (138)$$

を OLS で推定すると、不偏性も一致性も持たない。

そこで、個別効果を表現するダミー変数を導入し、

$$y_{it} = a + b_1 x_{it} + b_2 D_i + e_{it} \quad (139)$$

$$e_{it} = \alpha_i + v_{it} \quad (140)$$

個別効果は時間と共に不変であるため、時間差分を取ることで、

$$\Delta y_{it} = b_1 \Delta x_{it} + \Delta e_{it} \quad i = 1, \dots, N \quad (141)$$

となり、 $\Delta e_{it} = \Delta v_{it}$ は明白である。この時、

$$E(\Delta x_{it} \Delta e_{it}) = Cov(\Delta x_{it}, \Delta e_{it}) = Cov(\Delta x_{it} \Delta v_{it}) = 0 \quad (142)$$

となる。これから、差分をとった式を推定すれば、推定量が一致性を持つことが分かる。

期間が T 、経済主体が N というケースでは、経済主体毎にダミー変数を考え、

$$y_{it} = \alpha_1 Dum1_i + \alpha_2 Dum2_i + \dots + \alpha_N DumN_i + bx_{it} + v_{it} \quad (143)$$

を推定すると、個別効果はダミー変数で反映され、誤差項と説明変数との共分散が 0 となり、OLS 推定は一致性を持つ。個別効果の有無は、

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_N \quad (144)$$

を自由度 $(N - 1, NT - (N + k))$ の F 検定

$$F = \frac{(RSS_r - RSS_u)/(N - 1)}{RSS_u/(NT - (N + k))} \quad (145)$$

で調べることができる。ここで、 k は説明変数の数である。

8.4 ランダム・イフェクトモデル

$$y_{it} = a + bx_{it} + e_{it} \quad (146)$$

$$e_{it} = \alpha_i + v_{it} \quad (147)$$

のモデルにおいて、 x_{it} と v_{it} が相関せず、 v_{it} は標準的線形回帰モデルの仮定を満たす誤差項とする。すなわち、 $E(v_{it}) = 0$, $E(v_{it}^2) = V(v_{it}) = \sigma_v^2$, $E(v_{it}v_{js}) = Cov(v_{it}, v_{js}) = 0$ for $i \neq j$, $t \neq s$ を仮定する。

そこで、 $E(\alpha_i x_{it}) = Cov(\alpha_i, x_{it}) = 0$ の条件が満たされているケースを考える。 α_i に関し、 $E(\alpha_i) = 0$, $E(\alpha_i^2) = V(\alpha_i) = \sigma_\alpha^2$, $E(\alpha_i \alpha_j) = Cov(\alpha_i, \alpha_j) = 0$ for $i \neq j$, $E(\alpha_i v_{it}) = 0$ を仮定する。そうすると e_{it} は、

$$E(e_{it}) = 0 \quad (148)$$

$$\begin{aligned} E(e_{it}e_{js}) = Cov(e_{it}, e_{js}) &= \sigma_a^2 + \sigma_v^2 \text{ for } i = j, t = s \\ &= \sigma_\alpha^2 \text{ for } i \neq j, t \neq s \\ &= 0 \text{ その他} \end{aligned}$$

である。

説明変数と誤差項は相関しないため、OLS 推定量は不偏性と一致性を満たす。しかし、誤差項が相関しているため OLS は有効推定量ではないし、仮説検定も問題となる。BLUE を得るためには、

$$y_{it} - c\bar{y}_i = a(1 - c) + b(x_{it} - c\bar{x}_i) + (v_{it} - c\bar{v}_i) \quad (149)$$

において、 $(1 - c)^2$ を、

$$\frac{\sigma_v^2}{T\sigma_\alpha^2 + \sigma_v^2} \quad (150)$$

と選ぶと、 $(v_{it} - c\bar{v}_i)$ は互いに相関しないことを証明できる。この変換された式を OLS 推定すれば、BLUE となる。これが、ランダム・イフェクトモデルの OLS による推定量となる。

8.5 モデル選択

$E(\alpha_i x_i) = Cov(\alpha_i, x_i) = 0$ であればランダム・イフェクトモデル、 $E(\alpha_i x_i) = Cov(\alpha_i, x_i) \neq 0$ であればフィックスド・イフェクトモデルが選択されるとした。これは、経済主体の観察不可能な固有の要素が観察可能な説明変数と相関しないのか、あるいは相関するののかということである。例えば、生産関数の推定において、経営者の資質や企業風土といった観察不可能な要素

が、設備や資金という説明変数と相関しなければランダム・イフェクトモデルを選択し、影響するのであればフィックスド・イフェクトモデルを選択することになる。

このことを統計的に検定する方法として、Wu-Hausman 検定がある。 $Cov(\alpha_i, x_i) = 0$ であれば、ランダム・イフェクトモデルは一致性を持ち、かつ有効性もある。これに対し、フィックスド・イフェクトモデルは一致性を持つが有効性はない。 $Cov(\alpha_i, x_i) \neq 0$ であれば、フィックスド・イフェクトモデルは一致性を持ちかつ有効推定量である。しかし、ランダム・イフェクトモデルは一致性も有しない。帰無仮説 $H_0 : Cov(\alpha_i, x_i) = 0$ のもとで推定量を \hat{b}_{RE} とし、対立仮説 $H_a : Cov(\alpha_i, x_i) \neq 0$ のもとで推定量を \hat{b}_{FE} とする。両者の差 $\hat{q} = \hat{b}_{FE} - \hat{b}_{RE}$ は確率極限において、

$$H_0 : \hat{q} = 0 \quad (151)$$

$$H_a : \hat{q} \neq 0 \quad (152)$$

となる。 \hat{q} の分散 $V(\hat{q}) = V(\hat{b}_{FE}) - V(\hat{b}_{RE})$ を考え、

$$m = \frac{\hat{q}^2}{V(\hat{q})} \quad (153)$$

を検定する。この統計量が帰無仮説のもとで自由度 k (説明変数の数) の χ^2 分布に従う。