2025/07/27

藪友良[1]

『入門　実践する計量経済学』の解答例

『入門　実践する計量経済学』(東洋経済新報社、2023 年)の解答例です。実証分析の解答は、サポートウェブサイトの再現プログラム(Stata、R、Python)によって再現ができます。なお、こちらには初版(第 2 刷)以降に追加された練習問題の解答も掲載してます。初版(第 1 刷)を利用されている方は、ウェブサイトから新しい練習問題をダウンロードできます[2]。

第 1 章の答え

練習問題 1

観察データを用いて経済モデルを推定し、その結果に基づいて経済理論を検証すること。観察データを分析し、推定結果に基づいて、新たな経済理論を構築すること。

練習問題 2

第 1 の理由は経済理論の数学化、第 2 の理由はデータ整備の急速な進展、第 3 の理由は PC の急速な進歩にある(詳しくは、1.4 節参照)。

練習問題 3

(a) ①量的データ、②横断面データ

(b) ①質的データ、②パネルデータ

(c) ①質的データ、②繰り返し横断面データ

(d) ①量的データ、②時系列データ

練習問題 4

金融関係のデータ(株価やドル円レートなど)は、日次データが利用できる。

---

[1] 何かタイポ、誤り、分かり難い箇所があれば教えてください。連絡先は以下となります。tomoyabu82@gmail.com

[2] ご自分の教科書に問題が記載されていないときは、以下のサイトをチェックしてください。https://www.fbc.keio.ac.jp/~tyabu/keiryo/questions.pdf

失業率、物価指数、鉱工業生産指数などは月次データ、GDP は四半期データで利用できる。

練習問題 5

季節性は、天候、暦、社会慣習などの要因から生じる。天候要因は、気温や降水量などの変動から発生する。暦要因は、各月に含まれる日数の違いなどから生じる。社会慣習要因は、さまざまな社会慣習から生じる変動となる。詳しくは、コラム 1-1 を参照されたい。

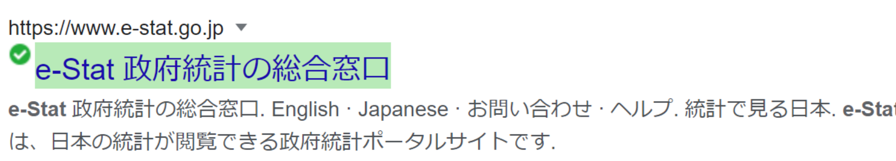
練習問題 6

商品別の売上が、天気や気温などによって、どのように変化するかを推定する。推定式が分かれば、当日の天気や気温などの情報を入力することで、当日の売り上げを予測できる。売り上げが予測できれば、その分だけ仕入れを行えば良いことになり、適正な商品在庫が可能となる。食品の廃棄を減らすことにつながり、企業利益の改善につながる。

年齢や性別別に、来店時間や売れ筋商品を調べることもできる。これらの情報を用いれば、来店時間に合わせて売れ筋商品を多く並べることができる。たとえば、高齢者は、平日昼間に来店する傾向があるならば、高齢者が好む商品を平日昼間に多くならべることで売り上げを伸ばすことができる。

練習問題 7

① Google で「estat」と入力すると、以下の項目が出るのでクリックする。



② そうすると以下のような画面が表示される。「e-Stat」は便利なサイトなので、ぜひ新規登録しよう。

③　キーワード検索で、「人口」と入力して検索すると、以下のような画面が表示される。人口に関連した様々な統計データがあることがわかる。これらをクリックして、自分が必要なデータがあるのかを調べることができる。他にも自分の関心があるワードを検索してみてほしい。

政府統計　人口　検索条件の保存

検索のしかた

保存した検索条件

政府統計一覧

政府統計コード	政府統計名	概要
00200524	人口推計	詳細
00450011	人口動態調査	詳細
00200502	社会・人口統計体系	詳細
00450013	人口動態統計特殊報告	詳細
00200241	住民基本台帳に基づく人口、人口動態及び世帯数調査	詳細
00450432	社会保障・人口問題基本調査（人口移動調査）	詳細
00200523	住民基本台帳人口移動報告	詳細

練習問題 8

①　Google で「fred data」と検索する。そうすると、以下のような項目が出るので、これをクリックする。

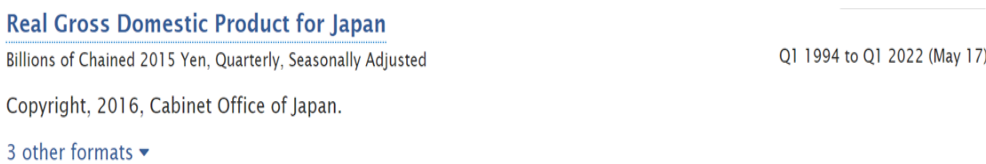
https://fred.stlouisfed.org ▼ このページを訳す

Federal Reserve Economic Data | FRED | St. Louis Fed

Download, graph, and track 816000 economic time series from 108 sources.

②　そうすると以下のような画面が表示される。ここで、検索枠に「gdp japan」と入力して検索しよう。なお、GDP は国内総生産(Gross Domestic Product)を表す。

③ Real Gross Domestic Product があるので、これをクリックする。データは 1994 年第 1 四半期(1994Q1)から 2022 年第 1 四半期(2022Q1)までの実質 GDP になる。また、seasonally adjusted とあるので、季節調整済み系列とわかる。

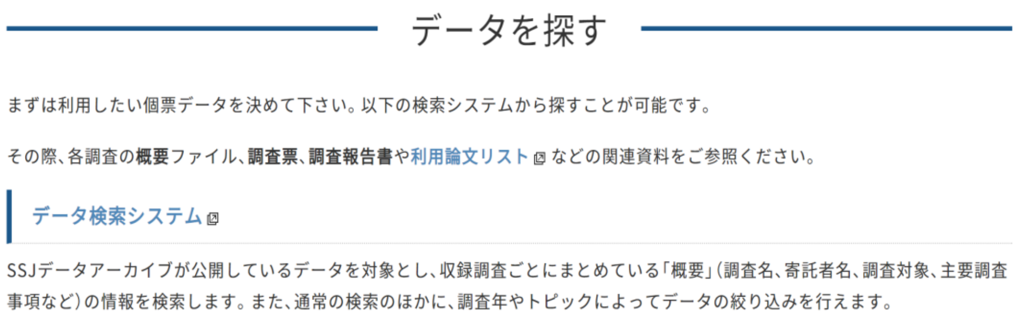


④ 下画面右上にある「DOWNLOAD」をクリックし、file format を指定すればデータをダウンロードできる。観察頻度などを変更したいなら、「EDIT GRAPH」をクリックし、変更しよう。設定を変更すると図の形状も変化し、download 時のデータも変更される。

練習問題 9

① Google で「東京大学　SSJDA」と検索すると、以下のような画面が出てくる。ここで、「利用データを探す」をクリックしよう。

② 以下の画面が表示されるので、「データ検索システム」をクリックする。

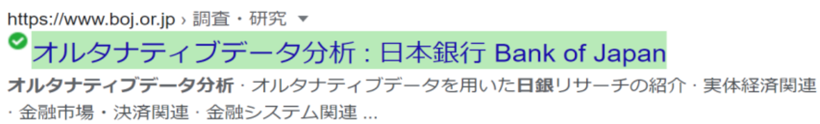


③ 下画面では、多数のデータが掲載されている。検索をしてデータを絞りたいなら、キーワードを入力して検索ボタンを押せばよい。このウェブサイトを調べて、どのようなデータがあるか確認してほしい。

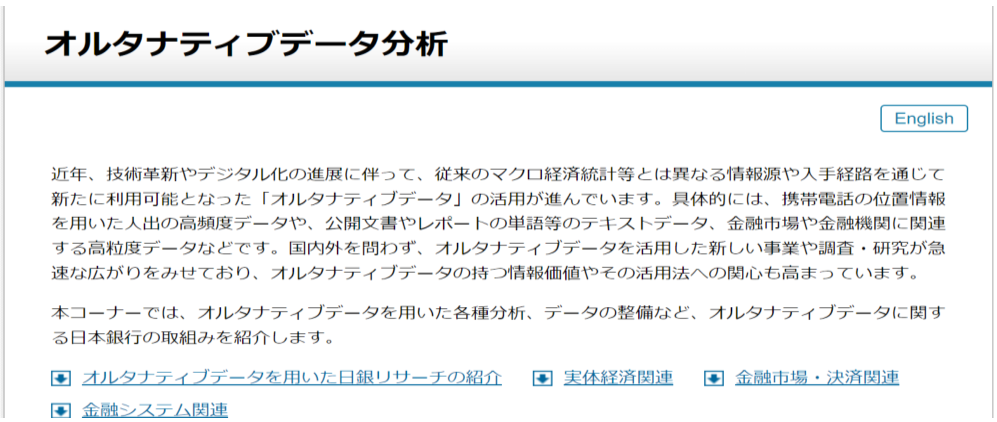
＊　SSJDA のミクロデータは、学部生であっても利用可能だが、指導教官から代理で申請してもらう必要がある。授業の担当教員もしくはゼミ指導教員にお願いしてみることをお勧めします。

練習問題 10

① Google で「日本銀行　オルタナティブデータ分析」と検索すると、以下のような項目が表示される。こちらをクリックしよう。


https://www.boj.or.jp › 調査・研究
オルタナティブデータ分析：日本銀行 Bank of Japan
オルタナティブデータ分析 · オルタナティブデータを用いた日銀リサーチの紹介 · 実体経済関連 · 金融市場・決済関連 · 金融システム関連 ...

② そうすると以下の画面が表示される。これをみると、オルタナティブデータを用いた日本銀行の取り組みが紹介されたサイトであることが分かる。


オルタナティブデータ分析

English

近年、技術革新やデジタル化の進展に伴って、従来のマクロ経済統計等とは異なる情報源や入手経路を通じて新たに利用可能となった「オルタナティブデータ」の活用が進んでいます。具体的には、携帯電話の位置情報を用いた人出の高頻度データや、公開文書やレポートの単語等のテキストデータ、金融市場や金融機関に関連する高粒度データなどです。国内外を問わず、オルタナティブデータを活用した新しい事業や調査・研究が急速な広がりをみせており、オルタナティブデータの持つ情報価値やその活用法への関心も高まっています。

本コーナーでは、オルタナティブデータを用いた各種分析、データの整備など、オルタナティブデータに関する日本銀行の取組みを紹介します。

オルタナティブデータを用いた日銀リサーチの紹介　実体経済関連　金融市場・決済関連　金融システム関連

③ 画面を下にスクロールすると、いろいろな研究が紹介されている(2022 年 6 月 8 日時点)。自分に関心のあるテーマが見つかったら、ぜひクリックして読んでみて欲しい。


調査・研究

掲載日	タイトル
2022年 3月30日	「オルタナティブデータ消費指数」の開発：オルタナティブデータを用いた個人消費のナウキャスティング
2022年 3月30日	新型コロナウイルス感染症拡大前後のオンライン消費動向の分析
2022年 3月 4日	オルタナティブデータを用いたGDPナウキャスティングモデルの構築
2021年12月20日	景況感は何に基づき形成されるのか：テキスト分析で探る景気ウォッチャーの着目点
2021年10月15日	景気ウォッチャー調査のテキスト分析からみた企業の短期インフレ予想
2021年 7月29日	米国における経済活動の再開と労働市場：「供給制約」に関する事実整理
2021年 5月27日	グローバルにみた感染症の家計等の行動への影響：機械学習によるアプローチ

*2025 年 6 月 25 日にチェックしたところ、このサイトはあまり更新されていないようです。

第 2 章の答え

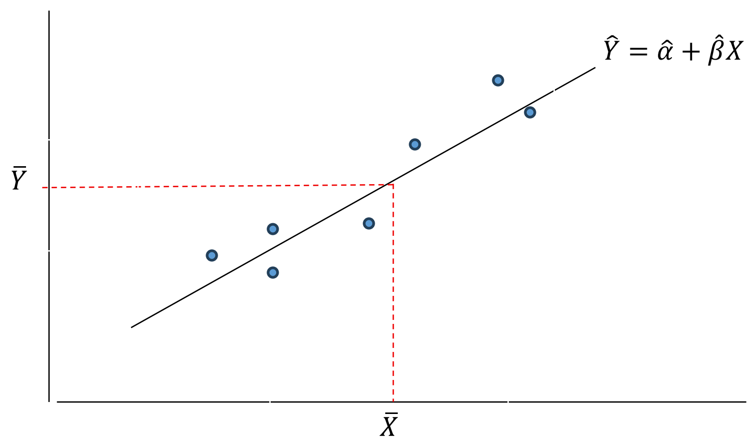
練習問題 1

OLS 推定量$\hat{\alpha}$は、$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$であり、これを書き換えると次式となる。

$$\bar{Y} = \hat{\alpha} + \hat{\beta}\bar{X}$$

よって、$X = \bar{X}$のとき回帰直線$\hat{Y} = \hat{\alpha} + \hat{\beta}X$の値は$\bar{Y} = \hat{\alpha} + \hat{\beta}\bar{X}$であり、回帰直線は点$(\bar{X}, \bar{Y})$を通ることが確認できる(下図参照)。

図　回帰直線の性質



練習問題 2

標本平均は、それぞれ次のようになる。

$$\bar{Y} = \frac{1}{5}(3 + 6 + 5 + 7 + 9) = 6$$

$$\bar{X} = \frac{1}{5}(5 + 6 + 5 + 8 + 11) = 7$$

また、標本分散は、それぞれ次のようになる。

$$s_Y^2 = \frac{1}{5-1}[(3-6)^2 + (6-6)^2 + (5-6)^2 + (7-6)^2 + (9-6)^2] = 5$$

$$s_X^2 = \frac{1}{5-1}[(5-7)^2 + (6-7)^2 + (5-7)^2 + (8-7)^2 + (1-7)^2] = 6.5$$

最後に、標本共分散s_{XY}は、

$$\frac{1}{5-1}[(3-6)(5-7) + (6-6)(6-7) + (5-6)(5-7) + (7-6)(8-7) + (9-6)(1-7)]$$

であり、これを計算すると、5.25 となる。標本相関係数は、これまでの情報を

用いて、$r_{XY} = \frac{5.25}{\sqrt{5 \times 6.5}} = 0.9209$となる。

練習問題 3

予測値$\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i$と練習問題 1 で求めた式$\bar{Y} = \hat{\alpha} + \hat{\beta}\bar{X}$を用いると、次の関係が得られる。

$$\sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2 = \sum_{i=1}^{n} (\hat{\alpha} + \hat{\beta} X_i - (\hat{\alpha} + \hat{\beta}\bar{X}))^2$$

$$= \sum_{i=1}^{n} (\hat{\beta} X_i - \hat{\beta}\bar{X})^2 = \hat{\beta}^2 \sum_{i=1}^{n} (X_i - \bar{X})^2$$

この偏差 2 乗和を決定係数の式に代入すると、

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2} = \hat{\beta}^2 \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}$$

となる。この式から、$\hat{\beta} = 0$の場合、決定係数も 0 となることがわかる。$\hat{\beta} = 0$は、説明変数Xで被説明変数Yの動きを全く説明できない状況であり、このとき、決定係数は 0 になるといえる。

OLS 推定量の公式$\hat{\beta} = \frac{\sum_{i=1}^{n} (X_i - X)(Y_i - \bar{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})^2}$を代入すると、次のように表現できる。

$$\hat{\beta}^2 \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2} = \left(\frac{\sum_{i=1}^{n} (X_i - X)(Y_i - \bar{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})^2} \right)^2 \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2} = \frac{\left(\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})\right)^2}{\sum_{i=1}^{n} (X_i - \bar{X})^2 \sum_{i=1}^{n} (Y_i - \bar{Y})^2}$$

練習問題 4

(a) OLS 推定量は次のようになる。

$$\hat{\beta} = -\frac{430}{400} = -1.075$$

$$\hat{\alpha} = 65 - (-1.075) \times 25 = 91.875$$

クラス人数が 0 人は概念的にないため、定数項は数学的切片と解釈される。$\hat{\beta} = -1.075$から、クラス人数が 1 人増えると、クラス平均点が 1.075 点下がる。

(b) クラスの人数が 20 のとき、予測値は次のようになる。

$$\hat{Y} = 91.875 - 1.075 \times 20 = 70.375$$

(c)決定係数は、練習問題 3 の関係式を用いると、次のようになる。

$$R^2 = (1.075)^2 \frac{400}{500} = 0.9245$$

平均点の全変動のうち 92.45%はクラスサイズの変動で説明できる。

練習問題 5

(a)OLS 推定量は次のようになる。

$$\hat{\beta} = \frac{1000}{250} = 4$$

$$\hat{\alpha} = 90 - 4 \times 20 = 10$$

定数項は、気温が 0 度のときの冷麺の売り上げ個数である。係数$\hat{\beta} = 4$から、気温が 1 度上がると、売り上げ個数は 4 個増える。

(b)気温が 30 度のとき、予測値は次のようになる。

$$\hat{Y} = 10 + 4 \times 30 = 130$$

(c) 決定係数は、練習問題 3 の関係式を用いると、次のようになる。

$$R^2 = \hat{\beta}^2 \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = 4^2 \frac{250}{5000} = 0.8$$

売り上げ個数の全変動のうち 80%は気温の変動により説明できる。

練習問題 6

練習問題 3 の結果から、

$$R^2 = \frac{\left(\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})\right)^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

$$= \left\{\frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}}\right\}^2$$

ここで、右辺の分母と分子を$n-1$で割ると、次式となり、これはまさにX_iとY_iの標本相関係数の 2 乗である。

$$\left\{\frac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}\right\}^2$$

練習問題 7

係数 β を 0 としたモデルでは、残差 2 乗和は次のようになる。

$$\sum_{i=1}^{n} \tilde{u}_i^2 = \sum_{i=1}^{n} (Y_i - \tilde{\alpha})^2$$

定数項$\tilde{\alpha}$に関して残差 2 乗和を微分して 0 と置くと、

$$\frac{\partial \sum_{i=1}^{n} \tilde{u}_i^2}{\partial \tilde{\alpha}} = \sum_{i=1}^{n} \frac{\partial \tilde{u}_i^2}{\partial \tilde{\alpha}} = \sum_{i=1}^{n} \frac{\partial \tilde{u}_i^2}{\partial \tilde{u}_i} \frac{\partial \tilde{u}_i}{\partial \tilde{\alpha}} = -2 \sum_{i=1}^{n} (Y_i - \hat{\alpha}) = 0$$

となる。式展開では、次の関係式を用いた。

$$\frac{\partial \tilde{u}_i^2}{\partial \tilde{u}_i} = 2\tilde{u}_i = 2(Y_i - \tilde{\alpha})$$

$$\frac{\partial \tilde{u}_i}{\partial \tilde{\alpha}} = \frac{\partial (Y_i - \tilde{\alpha})}{\partial \tilde{\alpha}} = -1$$

ここで、両辺を－2 で割ると、次の正規方程式が得られる。

$$\sum_{i=1}^{n} (Y_i - \hat{\alpha}) = 0$$

上式を満たす$\tilde{\alpha}$は最小 2 乗推定量であるため、$\hat{\alpha}$と表記している。この場合、残差は$\hat{u}_i = Y_i - \hat{\alpha}$であるため、正規方程式から残差の和は 0 となる。そして、正規方程式を展開すると、

$$\sum_{i=1}^{n} Y_i - n\,\hat{\alpha} = 0$$

であり、さらに両辺をnで割ると、$\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \bar{Y}$となる。

練習問題 8

2 章補足では、偏微分を用いて最小 2 乗(OLS)推定量を導出しました。しかし、OLS 推定量が本当に残差 2 乗和を最小化しているのか、もしくは最大化してしまっているかはわかりません(付録 B の B.3 節)。この練習問題では、OLS 推定量が残差 2 乗和を最小化していることを証明します。

(a) 残差 2 乗和は、次のように表現できる。

$$\sum_{i=1}^{n} (Y_i - \tilde{\alpha} - \tilde{\beta} X_i)^2 = \sum_{i=1}^{n} (Y_i - \hat{\alpha} - \hat{\beta} X_i + (\hat{\alpha} - \tilde{\alpha}) + (\hat{\beta} - \tilde{\beta}) X_i)^2$$

$Y_i = \hat{\alpha} + \hat{\beta}X_i + \hat{u}_i$を書き換えると、$\hat{u}_i = Y_i - \hat{\alpha} - \hat{\beta}X_i$となる。上式の右辺に、$\hat{u}_i = Y_i - \hat{\alpha} - \hat{\beta}X_i$を代入すると、右辺は次のようになる。

$$\sum_{i=1}^{n}(\hat{u}_i + (\hat{\alpha} - \tilde{\alpha}) + (\hat{\beta} - \tilde{\beta})X_i)^2$$

(b)　上式右辺は、次のように展開できる。

$$\sum_{i=1}^{n}(\hat{u}_i + \{(\hat{\alpha} - \tilde{\alpha}) + \left(\hat{\beta} - \tilde{\beta}\right)X_i\})^2 = \sum_{i=1}^{n}\hat{u}_i^2 + \sum_{i=1}^{n}\left\{(\hat{\alpha} - \tilde{\alpha}) + (\hat{\beta} - \tilde{\beta})X_i\right\}^2$$

$$+2(\hat{\alpha} - \tilde{\alpha})\underbrace{\sum_{i=1}^{n}\hat{u}_i}_{=0} + 2\left(\hat{\beta} - \tilde{\beta}\right)\underbrace{\sum_{i=1}^{n}\hat{u}_iX_i}_{=0}$$

$$= \sum_{i=1}^{n}\hat{u}_i^2 + \sum_{i=1}^{n}((\hat{\alpha} - \tilde{\alpha}) + (\hat{\beta} - \tilde{\beta})X_i)^2$$

式展開では、残差の性質($\sum_{i=1}^{n}\hat{u}_i = 0$、$\sum_{i=1}^{n}\hat{u}_iX_i = 0$)を用いた。

(c)これまでの結果をまとめると、次のようになる。

$$\sum_{i=1}^{n}(Y_i - \tilde{\alpha} - \tilde{\beta}X_i)^2 = \sum_{i=1}^{n}\hat{u}_i^2 + \sum_{i=1}^{n}((\hat{\alpha} - \tilde{\alpha}) + (\hat{\beta} - \tilde{\beta})X_i)^2$$

ここで、右辺第 2 項は 2 乗和なので 0 以上となる。このため、次式が成立する。

$$\sum_{i=1}^{n}\hat{u}_i^2 \le \sum_{i=1}^{n}(Y_i - \tilde{\alpha} - \tilde{\beta}X_i)^2$$

左辺は OLS 推定量の残差 2 乗和であり、右辺は任意の$\tilde{\alpha}$と$\tilde{\beta}$に対する残差 2 乗和である。このため、OLS 推定量($\hat{\alpha}$、$\hat{\beta}$)は残差 2 乗和$\sum_{i=1}^{n}(Y_i - \tilde{\alpha} - \tilde{\beta}X_i)^2$を最小にしていることが確認できる。

第 3 章の答え

練習問題 1

XとYの関係を知るためには、Yの変動は必要ない。Xが変動していて、Yが全く変動していなければ、XはYに何の影響も与えておらず、係数βは 0 と推定できる。これは、OLS 推定量の公式、つまり、次式から明らかである。

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

ここで、Y_iの変動が 0 なら$Y_i = \bar{Y}$であり$(Y_i - \bar{Y} = 0)$、$\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y}) = 0$となり、$\hat{\beta} = 0$と推定される。

練習問題 2

仮定 3(誤差項uの期待値が 0)が成立するよう定数項αが(暗黙のうちに)定義されているため、問題とならない(詳しくは 3.2.3 節を参照されたい)。

練習問題 3

係数βは真の値であり、OLS 推定量$\hat{\beta}$はデータから計算されるβの推定量となる。OLS 推定量$\hat{\beta}$は不偏性があるため、平均的には真の値βと一致する。また、OLS 推定量$\hat{\beta}$は一致性があるため、サンプルサイズが大きければ、$\hat{\beta}$は真の値βと一致することになる。

練習問題 4

誤差項u_iは、実現値と真の回帰式からの予測値との差であるのに対し、残差$\hat{u}_i$は、実現値と推定された回帰式からの予測値となる。

誤差項　$u_i = Y_i - (\alpha + \beta X_i)$

残差　$\hat{u}_i = Y_i - (\hat{\alpha} + \hat{\beta} X_i)$

なお、誤差項は観察できないが、残差は観察できるという違いもある。両者は混同しやすい概念なので注意してほしい。

練習問題 5

図(a)をみると、どのようなサンプルサイズnに対しても、推定量$\hat{\beta}_1$の分布の中心はβとなっているため、不偏性があることがわかる。また、サンプルサイズn

が∞のとき、$\hat{\beta}_1$はβと一致していることから、$\hat{\beta}_1$は一致性がある。これに対して、図(b)をみると、推定量$\hat{\beta}_2$の分布の中心はβではないため、$\hat{\beta}_2$には不偏性がない。しかし、サンプルサイズnが∞のとき、$\hat{\beta}_2$はβに一致していることから、$\hat{\beta}_2$には一致性がある。

練習問題 6

サンプルサイズが 8 のとき、推定結果は

$$\hat{Y} = 2.62 + 0.163X$$
$$\quad(1.05)\ \ (0.029)$$

となる。式の下に表記されたカッコ内の値は標準誤差である。自由度$n-2$は、6 と小さな値であることから、$t_{6,0.05} = 2.447$は 1.96 よりかなり大きくなる[3]。95%信頼区間は、$\hat{\beta} = 0.163$、標準誤差$s_{\hat{\beta}} = 0.029$を用いて、

$$\underbrace{0.163 - 2.447 \times 0.029}_{=0.092} < \beta < \underbrace{0.163 + 2.447 \times 0.029}_{=0.234}$$

となる。下限と上限を計算すると、これは(0.092、0.234)区間となる。

724 物件のデータを用いて同じ推定をすると、次のようになる。

$$\hat{Y} = 2.69 + 0.160X$$
$$\quad(0.101)\ \ (0.003)$$

ここで、nは724と大きいため、$t_{n-2,0.05}$として 1.96 を用いる(なお、$t_{722,0.05}$は、1.9632551 となり、ほぼ同じ値である[4])。

$$\underbrace{0.160 - 1.96 \times 0.003}_{=0.154} < \beta < \underbrace{0.160 + 1.96 \times 0.003}_{=0.166}$$

となる。これは(0.154、0.166)区間であり、信頼区間は狭く、βの範囲をかなり絞りこめていることがわかる。

練習問題 7

$\sum_{i=1}^{n}(u_i/\sigma)^2$は自由度$n$の$\chi^2$分布、$\sum_{i=1}^{n}(\hat{u}_i/\sigma)^2$は自由度$n-2$の$\chi^2$分布に従う(補足参照)。$\chi^2$確率変数の期待値は自由度であるから、それぞれの期待値はそれぞれ

$$E\left[\sum_{i=1}^{n}\left(\frac{u_i}{\sigma}\right)^2\right] = n$$

---

[3] Excel では、$t_{6,0.05}$は「=TINV(0.05,6) 」と入力すれば求められる。
[4] Excel では、$t_{722,0.05}$は「=TINV(0.05,722) 」と入力すれば求められる。

$$E\left[\sum_{i=1}^{n}\left(\frac{\hat{u}_i}{\sigma}\right)^2\right] = n-2$$

練習問題 8

期待値をとると、

$$E\left[\frac{1}{n}\sum_{i=1}^{n}\hat{u}_i^2\right] = \frac{\sigma^2}{n}\underbrace{E\left[\sum_{i=1}^{n}\left(\frac{\hat{u}_i}{\sigma}\right)^2\right]}_{\text{自由度}=n-2} = \frac{n-2}{n}\sigma^2$$

となる。式展開では、$\sum_{i=1}^{n}\left(\frac{\hat{u}_i}{\sigma}\right)^2$は自由度$n-2$の$\chi^2$分布に従うため、その期待値は自由度$n-2$となることを用いた(巻末付録 C の C.2 節参照)。

なお、次の関係から、$\sum_{i=1}^{n}\hat{u}_i^2/n$は誤差項の分散$\sigma^2$を過小推定していることが理解できる。

$$E\left[\frac{1}{n}\sum_{i=1}^{n}\hat{u}_i^2\right] = \frac{n-2}{n}\sigma^2 < \sigma^2$$

練習問題 9

OLS 推定量$\hat{\alpha}$の確率的表現は、$\hat{\alpha} = \alpha - (\hat{\beta}-\beta)\bar{X} + \bar{u}$となるため、その期待値は

$$\begin{aligned} E[\hat{\alpha}] &= \alpha - (E[\hat{\beta}] - \beta)\bar{X} + E[\bar{u}] \\ &= \alpha - (\beta - \beta)\bar{X} + 0 = \alpha \end{aligned}$$

となる(不偏性が満たされる)。式展開では、$E[\hat{\beta}] = \beta$となること、また、標準的仮定 3 から

$$E[\bar{u}] = \frac{1}{n}E[u_1 + u_2 + \cdots + u_n] = 0$$

が成立することを用いた。

練習問題 10

OLS 推定量の確率的表現から、$\hat{\alpha} - \alpha = -(\hat{\beta}-\beta)\bar{X} + \bar{u}$となる。このため、OLS 推定量$\hat{\alpha}$の分散は、

$$\begin{aligned} V(\hat{\alpha}) &= E[(\hat{\alpha}-\alpha)^2] \\ &= E[(-(\hat{\beta}-\beta)\bar{X} + \bar{u})^2] \end{aligned}$$

$$= \bar{X}^2 E[(\hat{\beta}-\beta)^2] - 2\bar{X}E[(\hat{\beta}-\beta)\bar{u}] + E[\bar{u}^2]$$

となる。ここで、右辺第 1 項は$\hat{\beta}$の分散で$E\left[\left(\hat{\beta}-\beta\right)^2\right] = \sigma^2/\sum_{i=1}^n (X_i-\bar{X})^2$となる。また、右辺第 2 項は、

$$\hat{\beta}-\beta = \frac{\sum_{i=1}^n (X_i-\bar{X})\,u_i}{\sum_{i=1}^n (X_i-\bar{X})^2}$$

を用いると、下記のように 0 となる[5]。

$$E\left[\left(\hat{\beta}-\beta\right)\bar{u}\right] = E\left[\left(\frac{\sum_{i=1}^n (X_i-\bar{X})\,u_i}{\sum_{i=1}^n (X_i-\bar{X})^2}\right)\left(\frac{\sum_{i=1}^n u_i}{n}\right)\right]$$

$$= \frac{1}{n\sum_{i=1}^n (X_i-\bar{X})^2} E\left[\left(\sum_{i=1}^n (X_i-\bar{X})\,u_i\right)\left(\sum_{i=1}^n u_i\right)\right]$$

$$= \frac{1}{n\sum_{i=1}^n (X_i-\bar{X})^2}\sigma^2 \underbrace{\sum_{i=1}^n (X_i-\bar{X})}_{=0} = 0$$

最後に、右辺第 3 項は次のようになる。

$$E[\bar{u}^2] = E\left[\left(\frac{\sum_{i=1}^n u_i}{n}\right)^2\right] = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$$

これらの結果を$V(\hat{\alpha})$の右辺に代入すると、

$$V(\hat{\alpha}) = \bar{X}^2 \underbrace{E\left[\left(\hat{\beta}-\beta\right)^2\right]}_{=\frac{\sigma^2}{\sum_{i=1}^n (X_i-\bar{X})^2}} - 2\bar{X}\underbrace{E\left[\left(\hat{\beta}-\beta\right)\bar{u}\right]}_{=0} + \underbrace{E[\bar{u}^2]}_{=\frac{\sigma^2}{n}}$$

$$= \sigma^2\left(\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i-\bar{X})^2}\right)$$

となる。この式から、nが大きくなると、分散が 0 に近づくことが分かる。

ここで、上式の右辺は、

$$\sigma^2\left(\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i-\bar{X})^2}\right) = \sigma^2\left(\frac{\sum_{i=1}^n (X_i-\bar{X})^2}{n\sum_{i=1}^n (X_i-\bar{X})^2} + \frac{n\bar{X}^2}{n\sum_{i=1}^n (X_i-\bar{X})^2}\right)$$

---

[5] 式展開では、$E[(\sum_{i=1}^n (X_i-\bar{X})\,u_i)(\sum_{i=1}^n u_i)] = \sigma^2\sum_{i=1}^n (X_i-\bar{X})$を用いた。これが正しいことを、$n=2$のケースで確認する。$n=2$の場合、$E[(\sum_{i=1}^2 (X_i-\bar{X})\,u_i)(\sum_{i=1}^2 u_i)]$は、

$$E\left[\left((X_1-\bar{X})u_1 + (X_2-\bar{X})u_2\right)(u_1+u_2)\right]$$
$$= (X_1-\bar{X})E[u_1^2] + (X_2-\bar{X})E[u_2^2] + (X_1-\bar{X})E[u_1u_2] + (X_2-\bar{X})E[u_1u_2]$$

となる。仮定 4 から$E[u_1^2] = E[u_2^2] = \sigma^2$、仮定 5 から$E[u_1u_2] = 0$となるため、上式右辺は次のようになる。

$$\sigma^2[(X_1-\bar{X}) + (X_2-\bar{X})] = \sigma^2\sum_{i=1}^2 (X_i-\bar{X})$$

$$= \sigma^2 \left(\frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{n\sum_{i=1}^n (X_i - \bar{X})^2} + \frac{n\bar{X}^2}{n\sum_{i=1}^n (X_i - \bar{X})^2} \right)$$

$$= \frac{\sigma^2 \sum_{i=1}^n X_i^2}{n\sum_{i=1}^n (X_i - \bar{X})^2}$$

と書き換えられる。式展開では次の関係式を用いた。

$$\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n (X_i^2 - 2\bar{X}X_i + \bar{X}^2)$$

$$= \sum_{i=1}^n X_i^2 - 2\bar{X}\sum_{i=1}^n X_i + n\bar{X}^2 = \sum_{i=1}^n X_i^2 - n\bar{X}^2$$

練習問題 11

$\hat{\alpha}$と$\hat{\beta}$の共分散は、$\hat{\alpha} - \alpha = -(\hat{\beta} - \beta)\bar{X} + \bar{u}$を用いると、

$$Cov(\hat{\alpha}, \hat{\beta}) = E[(\hat{\alpha} - \alpha)(\hat{\beta} - \beta)]$$

$$= E[(-(\hat{\beta} - \beta)\bar{X} + \bar{u})(\hat{\beta} - \beta)]$$

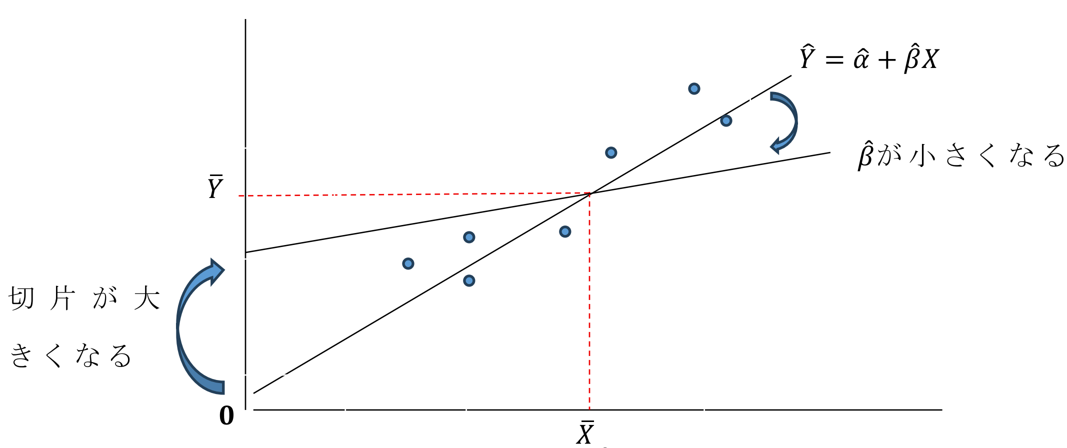
$$= -\bar{X}E\left[(\hat{\beta} - \beta)^2\right] + \underbrace{E[(\hat{\beta} - \beta)\bar{u}]}_{=0}$$

となる。練習問題 10 で示したとおり、右辺第 2 項は 0 となるため、共分散は次のようになる。

$$Cov(\hat{\alpha}, \hat{\beta}) = -\frac{\sigma^2 \bar{X}}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

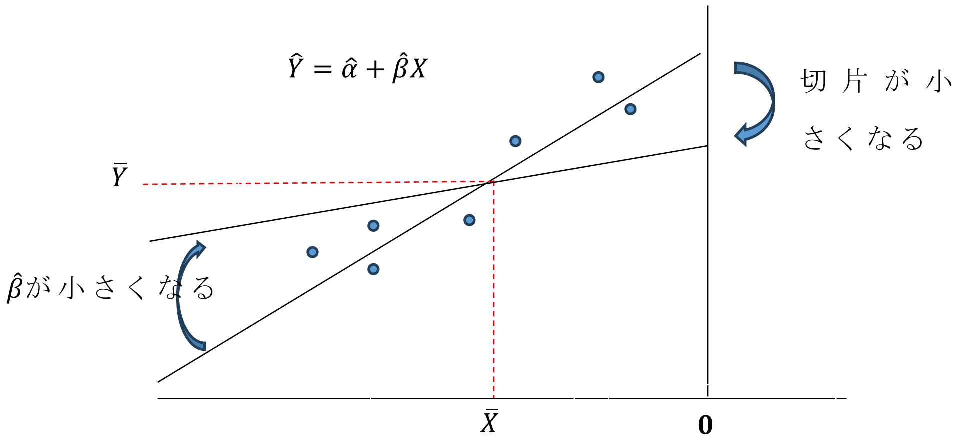
この結果から、$\bar{X} > 0$なら$\hat{\alpha}$と$\hat{\beta}$の共分散は負となる。下図では、$\bar{X} > 0$として回帰直線を描いた。回帰直線は点$(\bar{X}, \bar{Y})$を通るため、$\hat{\beta}$が小さくなると(傾きがゆるくなると)、切片$\hat{\alpha}$が大きくなる (つまり、$\hat{\alpha}$と$\hat{\beta}$に負の相関がある)。

図　共分散$Cov(\hat{\alpha}, \hat{\beta})$



逆に、この式から、$\bar{X}<0$なら$\hat{\alpha}$と$\hat{\beta}$の共分散は正となる。下図では、$\bar{X}<0$として回帰直線を描いた。回帰直線は点$(\bar{X},\bar{Y})$を通るため、$\hat{\beta}$が小さくなると(傾きがゆるくなると)、切片$\hat{\alpha}$が小さくなる(つまり、$\hat{\alpha}$と$\hat{\beta}$に正の相関がある)。

図　共分散$Cov(\hat{\alpha},\hat{\beta})$



練習問題 12

定数項がない回帰モデルは、3 章以降でも用いられますので、この問題は理解するようにしてください。

(a) 定数項がない回帰モデルでは、残差 2 乗和は$\sum_{i=1}^n \tilde{u}_i^2=\sum_{i=1}^n (Y_i-\tilde{\beta}X_i)^2$となる。$\tilde{\beta}$に関して残差 2 乗和$\sum_{i=1}^n \tilde{u}_i^2=\sum_{i=1}^n (Y_i-\tilde{\beta}X_i)^2$を微分して 0 と置くと、

$$\frac{\partial \sum_{i=1}^n \tilde{u}_i^2}{\partial \tilde{\beta}}=\sum_{i=1}^n \frac{\partial \tilde{u}_i^2}{\partial \tilde{u}_i}\frac{\partial \tilde{u}_i}{\partial \tilde{\beta}}=-2\sum_{i=1}^n \left(Y_i-\hat{\beta}X_i\right)X_i=0$$

であり、この両辺を－2 で割ると、次の正規方程式が得られる。

$$\sum_{i=1}^n (Y_i-\hat{\beta}X_i)X_i=0$$

正規方程式の左辺を展開すると、

$$\sum_{i=1}^n X_iY_i-\hat{\beta}\sum_{i=1}^n X_i^2=0$$

となり、この式を$\hat{\beta}$について解くと、最小 2 乗推定量が得られる。

$$\hat{\beta}=\frac{\sum_{i=1}^n X_iY_i}{\sum_{i=1}^n X_i^2}$$

残差の定義$\hat{u}_i = Y_i - \hat{\beta}X_i$に注意すると、正規方程式は残差の性質②、つまり、

$$\sum_{i=1}^{n} X_i \hat{u}_i = 0$$

を意味する。定数項がないため、正規方程式は 1 本だけであり、残差の性質① $\sum_{i=1}^{n} \hat{u}_i = 0$は成立しない。

(b) OLS 推定量$\hat{\beta}$である

$$\hat{\beta} = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2}$$

に、$Y_i = \beta X_i + u_i$を代入すると、$\hat{\beta}$の確率的表現が得られる。

$$\hat{\beta} = \frac{\sum_{i=1}^{n} X_i(\beta X_i + u_i)}{\sum_{i=1}^{n} X_i^2} = \beta + \frac{\sum_{i=1}^{n} X_i u_i}{\sum_{i=1}^{n} X_i^2}$$

(c)　確率的表現の期待値をとると、

$$E[\hat{\beta}] = \beta + \frac{\sum_{i=1}^{n} X_i E[u_i]}{\sum_{i=1}^{n} X_i^2} = \beta$$

となり、$\hat{\beta}$は不偏性を満たす(仮定 1 と仮定 3 に注意)。

次に、分散は、

$$V(\hat{\beta}) = E\left[(\hat{\beta} - \beta)^2\right] = E\left[\left(\frac{\sum_{i=1}^{n} X_i u_i}{\sum_{i=1}^{n} X_i^2}\right)^2\right]$$

$$= \frac{1}{\left(\sum_{i=1}^{n} X_i^2\right)^2} E\left[\left(\sum_{i=1}^{n} X_i u_i\right)^2\right]$$

$$= \frac{1}{\left(\sum_{i=1}^{n} X_i^2\right)^2} \sigma^2 \sum_{i=1}^{n} X_i^2 = \frac{\sigma^2}{\sum_{i=1}^{n} X_i^2}$$

となる。式展開では、

$$E\left[\left(\sum_{i=1}^{n} X_i u_i\right)^2\right] = \sigma^2 \sum_{i=1}^{n} X_i^2$$

を用いた($E[u_i^2] = \sigma^2$、$i \neq j$なら$E[u_i u_i] = 0$に注意)[6]。

---

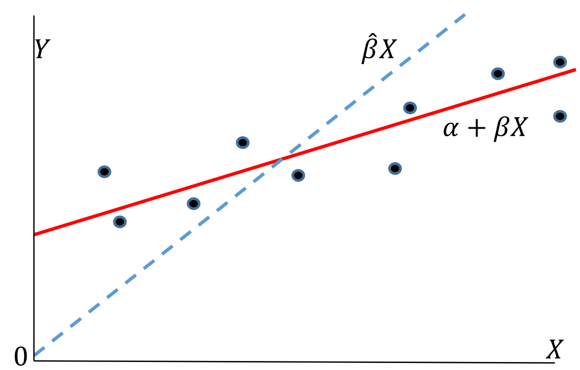
[6] $n = 2$として証明しよう。$E[(\sum_{i=1}^{n} X_i u_i)^2]$は、$n = 2$のとき、次のようになる。
$$E[(X_1 u_1 + X_2 u_2)^2] = X_1^2 \underbrace{E[u_1^2]}_{=\sigma^2} + X_2^2 \underbrace{E[u_2^2]}_{=\sigma^2} + 2X_1 X_2 \underbrace{E[u_1 u_2]}_{=0} = X_1^2\sigma^2 + X_2^2\sigma^2 = \sigma^2(X_1^2 + X_2^2)$$

ここで、$V(\hat{\beta}) = \sigma^2 / \sum_{i=1}^{n} X_i^2$であるから、サンプルサイズ$n$が大きくなると、分母$\sum_{i=1}^{n} X_i^2$が大きくなり、分散$V(\hat{\beta})$は 0 に収束する。まとめると、モデル$(Y_i = \beta X_i + u_i)$が正しいならば、$\hat{\beta}$は不偏性を満たし、$n$が大きいと分散は 0 となるため、一致性も満たしていることがわかる。

(d) 定数項なしの OLS 推定には問題が 2 つある。

第 1 の問題は、本当は$\alpha \neq 0$にもかかわらず、定数項なしの OLS 推定をすると、バイアスが発生する点である。下図では、真の関係を実線($\alpha + \beta X$、ただし、$\alpha > 0$とした)、定数項なしの推定から得られた回帰直線を点線$\hat{\beta} X$で表した。データは$\alpha + \beta X$で観察されるが、回帰直線は原点を通るように推定される。ここで、$\hat{\beta} > \beta$であり、$\hat{\beta}$にはバイアスが存在することがわかる。

図： 定数項なしの回帰におけるバイアス



第 2 の問題は、決定係数の 2 つの定義（2.6 参照）が一致しないという点である。つまり、

$$\frac{\sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2} \neq 1 - \frac{\sum_{i=1}^{n} \hat{u}_i^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}$$

となる。この点を確認しよう。$Y_i = \hat{Y}_i + \hat{u}_i$に注意すると、$Y_i$の全変動は、

$$\sum_{i=1}^{n} (Y_i - \bar{Y})^2 = \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y} + \hat{u}_i)^2$$

$$= \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2 + \sum_{i=1}^{n} \hat{u}_i^2 + 2 \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y}) \hat{u}_i$$

と分解できる。定数項なしの場合、残差の性質①(残差の和は 0)は成立しないため、右辺第 3 項目は 0 とならない。第 3 項目が 0 でないことは、

$$\sum_{i=1}^{n}(\hat{Y}_i-\bar{Y})\hat{u}_i=\sum_{i=1}^{n}\hat{Y}_i\hat{u}_i-\bar{Y}\sum_{i=1}^{n}\hat{u}_i=\hat{\beta}\underbrace{\sum_{i=1}^{n}X_i\hat{u}_i}_{=0}-\bar{Y}\underbrace{\sum_{i=1}^{n}\hat{u}_i}_{\neq 0}\neq 0$$

と確認できる。つまり、

$$\sum_{i=1}^{n}(Y_i-\bar{Y})^2\neq\sum_{i=1}^{n}(\hat{Y}_i-\bar{Y})^2+\sum_{i=1}^{n}\hat{u}_i^2$$

であるから、決定係数の 2 つの定義は一致しないことになる。

以上から、事前情報によって、$\alpha=0$が正しい場合は良いが、そうでなければ、定数項を入れた推定が望ましい。なお、事前情報によって、$\alpha=0$が正しい場合は稀であり、定数項を含めた推定を行うことが一般的となる。

練習問題 13

(a) 残差は、真の値Y_iと予測値$\hat{Y}_i$の差として、次式で表される。

$$\begin{aligned}\hat{u}_i=Y_i-\hat{Y}_i&=(\alpha+\beta X_i+u_i)-\left(\hat{\alpha}+\hat{\beta}X_i\right)\\&=u_i-(\hat{\alpha}-\alpha)-\left(\hat{\beta}-\beta\right)X_i\end{aligned}$$

(b) $E[\hat{u}_i^2]=\sigma^2(1-h_{ii})$の意味を考えてみよう。この結果から、誤差項の分散はσ^2で一定である一方、残差の分散は変化することがわかる。また、レバレッジh_{ii}は定義によって 0 以上となる(厳密には、$X_i=\bar{X}$のとき$h_{ii}=1/n$で最小となるため$1/n\leq h_{ii}$となる)[7]。よって、残差 2 乗の期待値は、誤差項の分散σ^2より小さい。

$$E[\hat{u}_i^2]=\sigma^2(1-h_{ii})<\sigma^2$$

この結果から、残差 2 乗はσ^2の不偏推定量ではなく、過小推定する性質があると理解できる。ただし、$\hat{\alpha}$と$\hat{\beta}$は一致推定量であるから、サンプルサイズが大きければ$\hat{u}_i=u_i$となる。

$E[\hat{u}_i^2]=\sigma^2(1-h_{ii})$を証明しよう。$E[\hat{u}_i]=0$から、$V(\hat{u}_i)=E[\hat{u}_i^2]$となる。(a)の結果を用いると、残差 2 乗の期待値は次のように展開できる。

---

[7] 重回帰分析では、レバレッジは、i 番目のデータ(1, X_{1i}、X_{2i}、…、X_{Ki})が、他のデータに比べて、どれぐらい異なるかを示す。Stata では、レバレッジは reg Y X とした後、predict leverage, hat とすれば計算できる。

$$E[\hat{u}_i^2] = E\left[\left(u_i - (\hat{\alpha}-\alpha) - \left(\hat{\beta}-\beta\right)X_i\right)^2\right]$$

$$= E[u_i^2] + \underbrace{E[(\hat{\alpha}-\alpha)^2]}_{=V(\hat{\alpha})} + X_i^2 \underbrace{E\left[\left(\hat{\beta}-\beta\right)^2\right]}_{=V(\hat{\beta})} + 2X_i \underbrace{E\left[(\hat{\alpha}-\alpha)\left(\hat{\beta}-\beta\right)\right]}_{=Cov(\hat{\alpha},\hat{\beta})}$$

$$-2E[u_i(\hat{\alpha}-\alpha)] - 2X_iE\left[u_i\left(\hat{\beta}-\beta\right)\right]$$

ここで、$V(\hat{\alpha})$、$V(\hat{\beta})$、$Cov(\hat{\alpha},\hat{\beta})$は、

$$V(\hat{\alpha}) = \sigma^2\left(\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i-\bar{X})^2}\right)、V\left(\hat{\beta}\right) = \frac{\sigma^2}{\sum_{i=1}^n (X_i-\bar{X})^2}、Cov\left(\hat{\alpha},\hat{\beta}\right) = -\frac{\sigma^2\bar{X}}{\sum_{i=1}^n (X_i-\bar{X})^2}$$

となる(3 章の P59、練習問題 10、 11 参照)。$\hat{\beta}$の確率的表現を用いると、

$$E\left[u_i\left(\hat{\beta}-\beta\right)\right] = E\left[u_i\frac{\sum_{i=1}^n (X_i-\bar{X})u_i}{\sum_{i=1}^n (X_i-\bar{X})^2}\right] = \frac{(X_i-\bar{X})E[u_i^2]}{\sum_{i=1}^n (X_i-\bar{X})^2} = \frac{(X_i-\bar{X})\sigma^2}{\sum_{i=1}^n (X_i-\bar{X})^2}$$

となり、$\hat{\alpha}$の確率的表現を使うと、

$$E[u_i(\hat{\alpha}-\alpha)] = E\left[u_i\left(-\left(\hat{\beta}-\beta\right)\bar{X} + \bar{u}\right)\right]$$

$$= -\bar{X}E\left[u_i\left(\hat{\beta}-\beta\right)\right] + \frac{1}{n}E\left[u_i\left(\sum_{i=1}^n u_i\right)\right]$$

$$= -\frac{\bar{X}(X_i-\bar{X})\sigma^2}{\sum_{i=1}^n (X_i-\bar{X})^2} + \frac{\sigma^2}{n}$$

となる(式展開では、$E[u_i^2] = \sigma^2$、$E\left[u_iu_j\right] = 0$ for $i \neq j$を用いた)。

これらを$E[\hat{u}_i^2]$の式に代入すると、次式ように展開できる。

$$E[\hat{u}_i^2] = \sigma^2\left\{1 + \left(\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i-\bar{X})^2}\right) + \frac{X_i^2\sigma^2}{\sum_{i=1}^n (X_i-\bar{X})^2} - 2\frac{X_i\bar{X}}{\sum_{i=1}^n (X_i-\bar{X})^2}\right.$$

$$\left. + 2\left(\frac{\bar{X}(X_i-\bar{X})\sigma^2}{\sum_{i=1}^n (X_i-\bar{X})^2} - \frac{1}{n}\right) - 2\frac{X_i(X_i-\bar{X})}{\sum_{i=1}^n (X_i-\bar{X})^2}\right\}$$

$$= \sigma^2\left\{1 - \frac{1}{n} + \frac{\bar{X}^2 + X_i^2 - 2X_i\bar{X} + 2\bar{X}(X_i-\bar{X}) - 2X_i(X_i-\bar{X})}{\sum_{i=1}^n (X_i-\bar{X})^2}\right\}$$

$$= \sigma^2\left\{1 - \frac{1}{n} - \frac{(X_i-\bar{X})^2}{\sum_{i=1}^n (X_i-\bar{X})^2}\right\} = \sigma^2(1-h_{ii})$$

ここで、$1/n$、$(X_i-\bar{X})^2$ともに正であり、h_{ii}が 0 以上となる。また、$E[\hat{u}_i^2] = \sigma^2(1-h_{ii})$は 0 以上であるから、$h_{ii}$は 1 以下となる。

(c) 単回帰分析の場合、レバレッジの総和は、$\sum_{i=1}^n \frac{1}{n} = n\frac{1}{n} = 1$に注意すると、

$$\sum_{i=1}^{n} h_{ii} = \sum_{i=1}^{n}\left(\frac{1}{n} + \frac{(X_i - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right) = 1 + \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = 2$$

となる。この結果を用いると、次のようになる。

$$E\left[\sum_{i=1}^{n} \hat{u}_i^2\right] = \sum_{i=1}^{n} E[\hat{u}_i^2] = \sigma^2 \sum_{i=1}^{n} (1 - h_{ii}) = \sigma^2(n-2)$$

s^2が不偏推定量であることを、次のようにして示せる。

$$E[s^2] = \frac{1}{n-2}\, E\left[\sum_{i=1}^{n} \hat{u}_i^2\right] = \frac{1}{n-2}\, \sigma^2(n-2) = \sigma^2$$

3 章の補足証明と異なり、ここでは、誤差項が正規分布するという仮定を用いておらず、s^2の不偏性を示すには、正規分布の仮定が不要であることがわかる。

(d) 標準化残差$\bar{u}_i$は次のように定義される。

$$\bar{u}_i = (1 - h_{ii})^{-1/2} \hat{u}_i$$

このとき、標準化残差の期待値は

$$E[\bar{u}_i] = (1 - h_{ii})^{-1/2} E[\hat{u}_i] = (1 - h_{ii})^{-1/2} \times 0 = 0$$

となり、また、標準化残差の 2 乗の期待値はσ^2となる。

$$E[\bar{u}_i^2] = (1 - h_{ii})^{-1} E[\hat{u}_i^2] = (1 - h_{ii})^{-1} (1 - h_{ii}) \sigma^2 = \sigma^2$$

よって、標準化残差の 2 乗は、σ^2の不偏推定量である。また、

$$E[s^2] = \frac{1}{n} E\left[\sum_{i=1}^{n} \bar{u}_i^2\right] = \frac{1}{n} \sum_{i=1}^{n} E[\bar{u}_i^2] = \frac{1}{n} n\sigma^2 = \sigma^2$$

となり、s^2の不偏性が満たされる。なお、標準化残差は 9 章の練習問題でも用いられるので、覚えておいてほしい。

第 4 章の答え

練習問題 1

帰無仮説は棄却することに意味がある。帰無仮説を採択しても、帰無仮説が正しいのか、対立仮説が正しいのか、どちらともいえない。帰無仮説を採択しても、帰無仮説が正しいと、誤って解釈しないように注意が必要である。

ハーバード大学元学長 L・サマーズ(Lawrence Summers)は、「統計学の授業で学生が何度も注意をうけるように、帰無仮説を棄却できないということは帰無仮説の正しさを意味していない」と述べている。これは当然のことだが、勘違いしやすい点なので注意したい。

練習問題 2

第 1 種の過誤は、帰無仮説H_0が正しいとき、帰無仮説H_0を誤って棄却し、対立仮説H_1を採択することをいう。第 2 種の過誤は、対立仮説H_1が正しいとき、帰無仮説H_0を誤って採択することをいう(詳しくは、4.2.2 節参照)。

練習問題 3

有意水準を低く設定する場合として、医薬品の開発やドーピング検査などがある。医薬品開発では、帰無仮説（「医薬品の効果がない」）を誤って棄却して効果のない薬(または有害な薬)を市場に出すデメリットは大きいため有意水準は低く設定される。また、ドーピング検査では、帰無仮説(「ドーピングをしていない」)を誤って棄却して選手のキャリアを傷つけるコストは大きいためやはり有意水準は低く設定される。

逆に、有意水準を低く設定しないものとして、人間ドックなどの簡易検査がある。帰無仮説(「病気にかかっていない（陰性)」)を誤って棄却しても、精密検査を行えば良いだけなので、大きな問題はない。逆に、帰無仮説を誤って採択し、病気を放置してしまうコストは大きい。

練習問題 4

有意水準 1%とした帰無仮説H_0: $\beta = \beta_0$の採択域($-c < t_{\hat{\beta}} < c$)は、$c = t_{n-2,0.01}$とすることで、次のようになる。

$$-t_{n-2,0.01} < \frac{\hat{\beta}-\beta_0}{s_{\hat{\beta}}} < t_{n-2,0.01}$$

これを書き換えると(両辺に$-s_{\hat{\beta}}$を掛けてから両辺に$\hat{\beta}$を足すと)、次の関係式が得られる。

$$\hat{\beta}-t_{n-2,0.01}s_{\hat{\beta}} < \beta_0 < \hat{\beta}+t_{n-2,0.01}s_{\hat{\beta}}$$

上式左辺はまさに 99%信頼区間の下限$\hat{\beta}-t_{n-2,0.01}s_{\hat{\beta}}$であり、右辺は 99%信頼区間の上限$\hat{\beta}+t_{n-2,0.01}s_{\hat{\beta}}$と一致している。このため、係数$\beta$の99%信頼区間の中に$\beta_0$が含まれる場合には、有意水準 1%で帰無仮説$H_0$は採択され、99%信頼区間の外に$\beta_0$がある場合には、有意水準 1%で帰無仮説$H_0$は棄却される。

練習問題 5

p値は、帰無仮説を棄却できる最も小さい有意水準を示している。4.4.2 節で学習したとおり、有意性は次のように判断される。

p値≤ 0.01ならば、　有意水準 1%で帰無仮説H_0は棄却される

$0.01 < p$値≤ 0.05 ならば、　有意水準 5%で帰無仮説H_0は棄却される

$0.05 < p$値≤ 0.1 ならば、　有意水準 10%で帰無仮説H_0は棄却される

$0.1 < p$値ならば、　有意水準 10%でも帰無仮説H_0は棄却されない

ここで、p値$= 0.09$であるため、有意水準 10%ならば帰無仮説は棄却されるが、有意水準 5%や 1%では棄却されない。また、p値$= 0.15$なら、有意水準 10%であっても、帰無仮説は棄却されない。

練習問題 6

t 値は$t^* = -0.95$である。t 分布は標準正規分布と同じと仮定すると、

$$p\text{値} = P\{0.95 < |Z|\}$$

となる。ただし、Zは標準正規確率変数となる。

これは標準正規分布表を用いることで計算できる。まず、$P\{Z < 0.95\} = 0.8289$となるため、確率の和は 1 から、

$$P\{Z > 0.95\} = 1 - P\{Z < 0.95\} = 0.1711$$

となる。また、標準正規分布は 0 を中心に左右対称なので、$P\{Z < -0.95\} = 0.1711$となり、よって、p値は次のようになる。

$$P\{0.95 < |Z|\} = P\{Z > 0.95\} + P\{Z < -0.95\} = 0.1711 \times 2 = 0.3422$$

Excel を用いるなら、$P\{Z < -0.95\}$は、「=NORM.S.DIST(-0.95,TRUE) 」と入力すれば計算できる。これは 0.171056 となり、これを 2 倍すると 0.3421 となる。

練習問題 7

第 1 の理由は、推定値と標準誤差がわかれば、簡単に 95%信頼区間を計算できるからである。第 2 の理由は、カッコ内に t 値を掲載することによって、仮説検定の結果(有意性)を強調しすぎてしまうからである(4.6.3 節参照)。

練習問題 8

たとえば、女性ダミーをF_iとしよう(女性ダミーは、女性なら 1 となり、男性なら 0 となるダミー変数である)。このとき、標本平均の分子$\sum_{i=1}^{n} F_i$は、データにおける女性の総数であり、それをサンプルサイズnで割ることで、データにおける女性の割合を求めることができる。つまり、

$$\frac{1}{n}\sum_{i=1}^{n} F_i = \frac{\text{女性の総数}}{\text{サンプルサイズ}} = \text{データにおける女性の割合}$$

練習問題 9

個人iの点数Y_iは、次のようになる。

$$Y_i = \mu_f + (\mu_m - \mu_f)M_i + u_i$$

たとえば、生徒iが男子ならば($M_i = 1$)、

$$Y_i = \mu_f + (\mu_m - \mu_f) + u_i = \mu_m + u_i$$

となる一方、女子ならば($M_i = 0$)、次のようになる。

$$Y_i = \mu_f + u_i$$

上式は、次の単回帰モデルとして表せる。

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

ただし、$\alpha = \mu_f$、$\beta = \mu_m - \mu_f$と定義した。この例から、男女の点数差を推定するために、女性ダミーか男性ダミーのいずれかを用いればよいとわかる。ただし、どちらのダミー変数を用いるかで、定数項や係数の解釈が異なる点に注意が必

要である。

練習問題 10

数学の点数を見てみよう。係数は 0.422 であり、男性は女性より 0.422 点だけ高い。ここで、$\hat{\beta} = 0.422$、$s_{\hat{\beta}} = 0.298$であり、$t_{\hat{\beta}} = 0.422/0.298 = 1.42$となる。ただし、$p$値は 10%より大きく、有意水準 10%でも帰無仮説$H_0: \beta = 0$は棄却できない(つまり、男女の平均点に有意な差があるとは言えない)。定数項は女性の平均点であり、これは 49.79 点となる。

次に、理科の点数を見てみよう。係数は 1.258 であり、男性は女性より 1.258 点だけ高い。ここで、$\hat{\beta} = 1.258$、$s_{\hat{\beta}} = 0.297$であり、$t_{\hat{\beta}} = 1.258/0.297 = 4.23$となる。$p$値は 1%より小さく、有意水準 1%で帰無仮説$H_0: \beta = 0$は棄却される。定数項は女性の平均点であり、これは 49.37 点となる。

例 4-4 では、説明変数を女性ダミーとしている一方、この問題では、説明変数を男性ダミーとしている。これらの推定結果を比較すると、推定結果は本質的に同じとなっていることがわかるだろう。ただし、定数項や係数の解釈は変わる点に注意が必要である。

練習問題 11

F_iの標本平均は$\bar{F} = \sum_{i=1}^{n} F_i / n = n_1/n$となる(女性は計$n_1$人いるため、$\sum_{i=1}^{n} F_i = n_1$となる)。このため、偏差は、

$$F_i - \bar{F} = \begin{cases} 1 - \dfrac{n_1}{n} = \dfrac{n_2}{n} & \text{if } \ F_i = 1 \\ 0 - \dfrac{n_1}{n} = -\dfrac{n_1}{n} & \text{if } \ F_i = 0 \end{cases}$$

となる(式展開では、$n_2 = n - n_1$を用いた)。したがって、偏差 2 乗和は、

$$\sum_{i=1}^{n} (F_i - \bar{F})^2 = \sum_{i=1}^{n_1} (F_i - \bar{F})^2 + \sum_{i=n_1+1}^{n} (F_i - \bar{F})^2$$

$$= n_1 \left(\frac{n_2}{n}\right)^2 + n_2 \left(\frac{n_1}{n}\right)^2$$

$$= \frac{n_1 n_2}{n^2}(n_1 + n_2) = \frac{n_1 n_2}{n}$$

となる(式展開では、$n = n_1 + n_2$を用いた)。

2 章で学習したとおり、説明変数Xの係数βの OLS 推定量$\hat{\beta}$は、

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})\,Y_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

となる。2 番目の等号は、次式を用いた。

$$\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y}) = \sum_{i=1}^{n}(X_i - \bar{X})Y_i - \bar{Y}\underbrace{\sum_{i=1}^{n}(X_i - \bar{X})}_{\text{偏差の和は }0} = \sum_{i=1}^{n}(X_i - \bar{X})Y_i$$

OLS 推定量$\hat{\beta}$の式に、$X_i = F_i$と$\bar{X} = \bar{F}$を代入すると、次のようになる。

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(F_i - \bar{F})Y_i}{\sum_{i=1}^{n}(F_i - \bar{F})^2}$$

$$= \frac{\sum_{i=1}^{n_1}(F_i - \bar{F})Y_i + \sum_{i=n_1+1}^{n}(F_i - \bar{F})Y_i}{\sum_{i=1}^{n}(F_i - \bar{F})^2}$$

ここで、$F_i = 1$なら$F_i - \bar{F} = \frac{n_2}{n}$、$F_i = 0$なら$F_i - \bar{F} = -\frac{n_1}{n}$となること、また、$\sum_{i=1}^{n}(F_i - \bar{F})^2 = \frac{n_1 n_2}{n}$となることに注意すると、上式右辺は次のようになる。

$$\frac{\frac{n_2}{n}\sum_{i=1}^{n_1} Y_i - \frac{n_1}{n}\sum_{i=n_1+1}^{n} Y_i}{\frac{n_1 n_2}{n}} = \frac{1}{n_1}\sum_{i=1}^{n_1} Y_i - \frac{1}{n_2}\sum_{i=n_1+1}^{n} Y_i = \bar{Y}_f - \bar{Y}_m$$

最後の等式は、$\bar{Y}_f$は女性の標本平均、$\bar{Y}_m$は男性の標本平均であること、つまり、以下の関係式を用いた。

$$\bar{Y}_f = \frac{1}{n_1}\sum_{i=1}^{n_1} Y_i \qquad \bar{Y}_m = \frac{1}{n_2}\sum_{i=n_1+1}^{n} Y_i$$

第 5 章の答え

練習問題 1

5.3.1 節で学習したとおり、被説明変数をY_i、説明変数をX_iとした単回帰分析では、X_iの係数$\hat{\beta}_1$の期待値は、

$$E\left[\hat{\beta}_1\right] = \beta_1 + \underbrace{\beta_2 \frac{s_{XW}}{s_X^2}}_{\text{欠落変数バイアス}}$$

となる。ここで、$\beta_2 > 0$、$s_{XW} < 0$であるから、欠落変数バイアスは負となる。

$$\beta_2 \frac{s_{XW}}{s_X^2} < 0$$

これは、$E\left[\hat{\beta}_1\right]$は真の値$\beta_1$より小さいことを意味する。

$$E\left[\hat{\beta}_1\right] < \beta_1$$

まとめると、生まれつきの能力W_iを除くことで、職業訓練ダミーX_iの係数(職業訓練の効果)は低めに推定されてしまう。

練習問題 2

ここで、$\beta_2 < 0$、$s_{XW} > 0$であるから、欠落変数バイアスは負となる($\beta_2 \frac{s_{XW}}{s_X^2} < 0$)。これは、$E\left[\hat{\beta}_1\right]$は真の値$\beta_1$より小さいことを意味する($E\left[\hat{\beta}_1\right] < \beta_1$)。つまり、移民の割合を除くことで、クラスの人数$X_i$の係数は低めに推定される。これは単回帰分析のほうが、重回帰分析よりも係数が小さくなる(クラスの人数を減らすことの効果が大きくなる)ことを意味する。

練習問題 3

双子の差をとると、

$$Y_i^{兄} - Y_i^{弟} = \left(\alpha + \beta_1 X_i^{兄} + \beta_2 W_i^{兄} + u_i^{兄}\right) - (\alpha + \beta_1 X_i^{弟} + \beta_2 W_i^{弟} + u_i^{弟})$$

$$= \beta_1\left(X_i^{兄} - X_i^{弟}\right) + \beta_2\left(W_i^{兄} - W_i^{弟}\right) + \left(u_i^{兄} - u_i^{弟}\right)$$

$$= \beta_1(X_i^{兄} - X_i^{弟}) + (u_i^{兄} - u_i^{弟})$$

となる。式展開では、$W_i^{兄} - W_i^{弟} = 0$とした。ここで、$Y_i = Y_i^{兄} - Y_i^{弟}$、$X_i = X_i^{兄} - X_i^{弟}$、$u_i = u_i^{兄} - u_i^{弟}$と定義すれば、上式は通常の単回帰分析によって推定でき

る。つまり、双子のデータを用いる利点は、双子の差をとることで、生まれつきの能力の要因を取り除くことが可能となる点にある。

現実問題として、双子であれば教育年数にあまり違いはない可能性がある。つまり、教育年数の差$X_i = X_i^{兄} - X_i^{弟}$はほぼ 0 の値をとり、説明変数の変動が非常に小さくなる。これでは、推定結果は不安定となる。3.3.3 節では、説明変数の変動が大きいほど、OLS 推定量の分散が小さくなることを説明している(図 3-4(a)参照)。

練習問題 4

5.6 節をもとに答えをまとめる。決定係数R^2は、

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \hat{u}_i^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

と定義され、説明変数X_iの数Kが増えるほど、その値が 1 に近づくという性質がある。これに対して、自由度調整済み決定係数は

$$\bar{R}^2 = 1 - \underbrace{\frac{n-1}{n-K-1}}_{\text{調整項}} \frac{\sum_{i=1}^{n} \hat{u}_i^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

と定義される。自由度調整済み決定係数では、調整項$\frac{n-1}{n-K-1}$を含めることで、説明変数を含めることに罰則を課している。つまり、自由度調整済み決定係数では、悪い説明変数を含めると、逆に、その値が下がることになる。

練習問題 5

都道府県の転入超過数とは、転入者数から転出者数を引いた値となる。つまり、

転入超過数 ＝ 転入者数 － 転出者数

という関係がある。これは恒等式であり、そもそも推定する意味はない。

練習問題 6

残差 2 乗和は、$X_2 = 10X_1$に注意すると、次のようになる。

$$\sum_{i=1}^{n} \tilde{u}_i^2 = \sum_{i=1}^{n}(Y_i - \tilde{\alpha} - \tilde{\beta}_1 X_{1i} - \tilde{\beta}_2 X_{2i})^2 = \sum_{i=1}^{n}(Y_i - \tilde{\alpha} - (\tilde{\beta}_1 + 10\tilde{\beta}_2)X_{1i})^2$$

これを各パラメータ$(\tilde{\alpha}, \tilde{\beta}_1,\ \tilde{\beta}_2)$で偏微分して 0 と置くと 3 本の式が得られる。

①	$$\frac{\partial \sum \tilde{u}_i^2}{\partial \tilde{\alpha}} = \sum \frac{\partial \tilde{u}_i^2}{\partial \tilde{u}_i}\frac{\partial \tilde{u}_i}{\partial \tilde{\alpha}} = \sum 2\tilde{u}_i(-1) = -2\sum (Y_i - \tilde{\alpha} - (\tilde{\beta}_1 + 10\tilde{\beta}_2)X_{1i}) = 0$$
②	$$\frac{\partial}{\partial \tilde{\beta}_1}\sum \tilde{u}_i^2 = \sum \frac{\partial \tilde{u}_i^2}{\partial \tilde{u}_i}\frac{\partial \tilde{u}_i}{\partial \tilde{\beta}_1} = \sum 2\tilde{u}_i(-X_{1i}) = -2\sum (Y_i - \tilde{\alpha} - (\tilde{\beta}_1 + 10\tilde{\beta}_2)X_{1i})X_{1i} = 0$$
③	$$\frac{\partial}{\partial \tilde{\beta}_2}\sum \tilde{u}_i^2 = \sum \frac{\partial \tilde{u}_i^2}{\partial \tilde{u}_i}\frac{\partial \tilde{u}_i}{\partial \tilde{\beta}_2} = \sum 2\tilde{u}_i(-10X_{1i}) = -20\sum (Y_i - \tilde{\alpha} - (\tilde{\beta}_1 + 10\tilde{\beta}_2)X_{1i})X_{1i} = 0$$

②式と③式は、本質的に同じ式となる。これは②式の両辺を－2 で割る、③式の両辺を－20 で割ると、次式となることから明らかである。

$$\sum \left(Y_i - \tilde{\alpha} - \left(\tilde{\beta}_1 + 10\tilde{\beta}_2\right)X_{1i}\right)X_{1i} = 0$$

以上から、正規方程式は 2 本の独立な式だけとなる。

$$\sum (Y_i - \tilde{\alpha} - (\tilde{\beta}_1 + 10\tilde{\beta}_2)X_{1i}) = 0$$
$$\sum \left(Y_i - \tilde{\alpha} - \left(\tilde{\beta}_1 + 10\tilde{\beta}_2\right)X_{1i}\right)X_{1i} = 0$$

独立な式は 2 本、パラメータは 3 つあるため、OLS 推定量を求めることはできない。OLS 推定量を導出するためには、独立な式がパラメータの数と同じだけ必要である。

練習問題 7

残差 2 乗和は次のようになる。

$$\sum_{i=1}^{n} \tilde{u}_i^2 = \sum_{i=1}^{n} (Y_i - \tilde{\alpha} - \tilde{\beta}_1 X_{1i} - \tilde{\beta}_2 X_{2i})^2$$

これを各パラメータ$(\tilde{\alpha}, \tilde{\beta}_1,\ \tilde{\beta}_2)$で偏微分して 0 と置くと 3 本の式が得られる。

①	$$\frac{\partial \sum \tilde{u}_i^2}{\partial \tilde{\alpha}} = \sum \frac{\partial \tilde{u}_i^2}{\partial \tilde{u}_i}\frac{\partial \tilde{u}_i}{\partial \tilde{\alpha}} = \sum 2\tilde{u}_i(-1) = -2\sum (Y_i - \tilde{\alpha} - \tilde{\beta}_1 X_{1i} - \tilde{\beta}_2 X_{2i}) = 0$$
②	$$\frac{\partial}{\partial \tilde{\beta}_1}\sum \tilde{u}_i^2 = \sum \frac{\partial \tilde{u}_i^2}{\partial \tilde{u}_i}\frac{\partial \tilde{u}_i}{\partial \tilde{\beta}_1} = \sum 2\tilde{u}_i(-X_{1i}) = -2\sum (Y_i - \tilde{\alpha} - \tilde{\beta}_1 X_{1i} - \tilde{\beta}_2 X_{2i})X_{1i} = 0$$
③	$$\frac{\partial}{\partial \tilde{\beta}_2}\sum \tilde{u}_i^2 = \sum \frac{\partial \tilde{u}_i^2}{\partial \tilde{u}_i}\frac{\partial \tilde{u}_i}{\partial \tilde{\beta}_2} = \sum 2\tilde{u}_i(-X_{2i}) = -2\sum (Y_i - \tilde{\alpha} - \tilde{\beta}_1 X_{1i} - \tilde{\beta}_2 X_{2i})X_{2i} = 0$$

3 本の式は独立に見えるが、これは誤りである。②式と③式を足すと、

$$-2\sum\left(Y_i-\tilde{\alpha}-\tilde{\beta}_1X_{1i}-\tilde{\beta}_2X_{2i}\right)X_{1i}-2\sum(Y_i-\tilde{\alpha}-\tilde{\beta}_1X_{1i}-\tilde{\beta}_2X_{2i})X_{2i}=0$$

となる。ここで左辺をまとめると、

$$-2\sum\left(Y_i-\tilde{\alpha}-\tilde{\beta}_1X_{1i}-\tilde{\beta}_2X_{2i}\right)(X_{1i}+X_{2i})=-2\sum(Y_i-\tilde{\alpha}-\tilde{\beta}_1X_{1i}-\tilde{\beta}_2X_{2i})$$

となる(式展開では、$X_1+X_2=1$を用いた)。つまり、②式+③式は、①式であり、独立な式は 2 本だけとわかる。独立な式は 2 本(②式と③式)、パラメータは 3 つ($\tilde{\alpha}$, $\tilde{\beta}_1$, $\tilde{\beta}_2$)あるため、OLS 推定量を求めることはできない。

練習問題 8

残差 2 乗和$\sum_{i=1}^{n}\tilde{u}_i^2=\sum_{i=1}^{n}\left(Y_i-\tilde{\alpha}-\tilde{\beta}_1X_{1i}-\tilde{\beta}_2X_{2i}\right)^2$を$\tilde{\alpha}$で偏微分して 0 と置くと、

$$-2\sum_{i=1}^{n}(Y_i-\hat{\alpha}-\hat{\beta}_1X_{1i}-\hat{\beta}_2X_{2i})=0$$

となり、両辺を-2で割ると

$$\sum_{i=1}^{n}(Y_i-\hat{\alpha}-\hat{\beta}_1X_{1i}-\hat{\beta}_2X_{2i})=0$$

となる。上式を満たす$\tilde{\alpha}$、$\tilde{\beta}_1$、$\tilde{\beta}_2$は最小 2 乗推定量なので、「^(ハット)」を付けた。この式を展開すると、

$$\sum_{i=1}^{n}Y_i-n\hat{\alpha}-\hat{\beta}_1\sum_{i=1}^{n}X_{1i}-\hat{\beta}_2\sum_{i=1}^{n}X_{2i}=0$$

となり、さらに両辺をnで割って、$\hat{\alpha}$について解くと次式が得られる。

$$\hat{\alpha}=\bar{Y}-\hat{\beta}_1\bar{X}_1-\hat{\beta}_2\bar{X}_2$$

次に、残差 2 乗和$\sum_{i=1}^{n}\tilde{u}_i^2=\sum_{i=1}^{n}\left(Y_i-\tilde{\alpha}-\tilde{\beta}_1X_{1i}-\tilde{\beta}_2X_{2i}\right)^2$を$\tilde{\beta}_1$で偏微分して 0 と置くと、

$$-2\sum_{i=1}^{n}(Y_i-\hat{\alpha}-\hat{\beta}_1X_{1i}-\hat{\beta}_2X_{2i})\,X_{1i}=0$$

となり、さらに両辺を-2で割ると、

$$\sum_{i=1}^{n}(Y_i-\hat{\alpha}-\hat{\beta}_1X_{1i}-\hat{\beta}_2X_{2i})\,X_{1i}=0$$

となる。ここで、$\hat{\alpha}=\bar{Y}-\hat{\beta}_1\bar{X}_1-\hat{\beta}_2\bar{X}_2$を上式に代入すると、下式となる。

$$\sum_{i=1}^{n}\left(Y_i-(\bar{Y}-\hat{\beta}_1\bar{X}_1-\hat{\beta}_2\bar{X}_2)-\hat{\beta}_1X_{1i}-\hat{\beta}_2X_{2i}\right)X_{1i}$$

$$=\sum_{i=1}^{n}\left((Y_i-\bar{Y})-\hat{\beta}_1(X_{1i}-\bar{X}_1)-\hat{\beta}_2(X_{2i}-\bar{X}_2)\right)X_{1i}$$

$$=\sum_{i=1}^{n}(Y_i-\bar{Y})X_{1i}-\hat{\beta}_1\sum_{i=1}^{n}(X_{1i}-\bar{X}_1)X_{1i}-\hat{\beta}_2\sum_{i=1}^{n}(X_{2i}-\bar{X}_2)X_{1i}=0$$

偏差の和は 0 から、上式は次のように書き換えられる[8]。

$$\sum_{i=1}^{n}(Y_i-\bar{Y})(X_{1i}-\bar{X}_1)-\hat{\beta}_1\sum_{i=1}^{n}(X_{1i}-\bar{X}_1)(X_{1i}-\bar{X}_1)-\hat{\beta}_2\sum_{i=1}^{n}(X_{2i}-\bar{X}_2)(X_{1i}-\bar{X}_1)=0$$

ここで、X_{1i}とX_{2i}の標本共分散は 0 ならば、上式の左辺第 3 項は 0 となる($\sum_{i=1}^{n}(X_{2i}-\bar{X}_2)(X_{1i}-\bar{X}_1)=0$)。よって、上式は、

$$\sum_{i=1}^{n}(Y_i-\bar{Y})(X_{1i}-\bar{X}_1)-\hat{\beta}_1\sum_{i=1}^{n}(X_{1i}-\bar{X}_1)^2=0$$

となり、これを$\hat{\beta}_1$について解けば次式が得られる。

$$\hat{\beta}_1=\frac{\sum_{i=1}^{n}(X_{1i}-\bar{X}_1)(Y_i-\bar{Y})}{\sum_{i=1}^{n}(X_{1i}-\bar{X}_1)^2}$$

以上より、重回帰分析における OLS 推定量$\hat{\beta}_1$は、Y_iをX_{1i}だけで単回帰したときの OLS 推定量の式と同じである。つまり、説明変数間の相関が 0 であるなら、単回帰分析であっても欠落変数バイアスが生じないことがわかる。

練習問題 9

被説明変数を所得Y_iとし、説明変数を教育年数X_iと職種W_iとした、次の重回帰モデルを考える。

$$Y_i=\alpha+\beta_1X_i+\beta_2W_i+u_i \qquad ①$$

ここで、教育年数X_iの係数β_1は、「職種W_iを一定とし、教育年数X_iが 1 年増えた

---

[8]偏差の和は 0($\sum_{i=1}^{n}(X_{1i}-\bar{X}_1)=0$、$\sum_{i=1}^{n}(X_{2i}-\bar{X}_2)=0$)であるから、以下の式が成立する。

$$\sum_{i=1}^{n}(X_{1i}-\bar{X}_1)X_{1i}=\sum_{i=1}^{n}(X_{1i}-\bar{X}_1)X_{1i}-\sum_{i=1}^{n}(X_{1i}-\bar{X}_1)\bar{X}_1=\sum_{i=1}^{n}(X_{1i}-\bar{X}_1)(X_{1i}-\bar{X}_1)$$

$$\sum_{i=1}^{n}(X_{2i}-\bar{X}_2)X_{1i}=\sum_{i=1}^{n}(X_{2i}-\bar{X}_2)X_{1i}-\sum_{i=1}^{n}(X_{2i}-\bar{X}_2)\bar{X}_1=\sum_{i=1}^{n}(X_{1i}-\bar{X}_1)(X_{2i}-\bar{X}_2)$$

式展開では、$\sum_{i=1}^{n}(X_{1i}-\bar{X}_1)\bar{X}_1=\bar{X}_1\sum_{i=1}^{n}(X_{1i}-\bar{X}_1)=\bar{X}_1\times 0=0$、$\sum_{i=1}^{n}(X_{2i}-\bar{X}_2)\bar{X}_1=\bar{X}_1\sum_{i=1}^{n}(X_{2i}-\bar{X}_2)=0$を用いた。

とき、所得Y_iがいくら変化するか」を示している。こうした教育の効果に関心があるなら、この推定をしてもよいだろう。

ここで、職種は教育年数に依存して、次のように決まるとする。

$$W_i = \theta_0 + \theta_1 X_i + e_i \qquad ②$$

この式を、所得Y_iの式に代入すると、

$$Y_i = \alpha + \beta_1 X_i + \beta_2(\theta_0 + \theta_1 X_i + e_i) + u_i$$

$$= \underbrace{(\alpha + \beta_2\theta_0)}_{=\delta_0} + \underbrace{(\beta_1 + \beta_2\theta_1)}_{=\delta_1} X_i + \underbrace{(u_i + \beta_2 e_i)}_{=\varepsilon_i}$$

となり、次の単回帰モデルが得られる。

$$Y_i = \delta_0 + \delta_1 X_i + \varepsilon_i \qquad ③$$

ただし、$\delta_0 = \alpha + \beta_2\theta_0$、$\delta_1 = \beta_1 + \beta_2\theta_1$、$\varepsilon_i = u_i + \beta_2 e_i$とした。③式は、①式と②式を統合したモデルとなっている。また、教育年数X_iは確率変数ではないため、X_iは$\varepsilon_i = u_i + \beta_2 e_i$と無相関になる[9]。つまり、単回帰モデルを推定しても、欠落変数バイアスは生じない。ここで、教育年数の係数δ_1は、次のようになる。

$$\delta_1 = \underbrace{\beta_1}_{\text{直接的効果}} + \underbrace{\beta_2\theta_1}_{\text{間接的効果}}$$

つまり、係数β_1は教育年数の「直接的効果」、$\beta_2\theta_1$は教育年数が職種に影響を与えることから生じる「間接的効果」を合わせたものとなる。

練習問題 14

(a)　勤続年数=年齢－教育年数－6 と定義される[10]。教育年数は、小卒なら 6 年、中卒なら 9 年、高卒なら 12 年、大卒なら 16 年となる。たとえば、40 歳大卒なら、勤続年数は 18 年(=40－16－6)である。勤続年数=年齢－教育年数－6 から、－6－年齢+教育年数－勤続年数=0 となり、多重共線性が成立する。

(b) ダミー変数は 0 から 1 の値をとるので、2 乗しても値がかわらない(0 の 2

9　ここで、u_iとe_iは誤差項であり、期待値はそれぞれ 0 となる。このため、ε_iの期待値も次のとおり 0 となる。

$$E[\varepsilon_i] = E[u_i + \beta_2 e_i] = E[u_i] + \beta_2 E[e_i] = 0 + \beta_2 \times 0 = 0$$

X_iは確率変数ではないことに注意すると、X_iとε_iとの共分散は次のとおり 0 となる。

$$Cov(X_i, \varepsilon_i) = E[(X_i - \bar{X})\varepsilon_i] = (X_i - \bar{X})E[\varepsilon_i] = 0$$

10　勤続年数がデータとして利用できない場合、こうした計算式を用いることになる。かりに勤続年数が正確にわかるなら、この計算式が成立しないこともあるだろう。たとえば、大学入学前に浪人していれば、この式は成立しなくなり、多重共線性の問題も生じない。

乗は 0 であり、1 の 2 乗は 1 である)。つまり、 $D_i = D_i^2$であり、多重共線性が生じる。

(c) $X_{1i} = 0$であるから、c_1以外をすべて 0 としても($c_1 \neq 0$、$c_0 = c_2 = \cdots = c_K = 0$)、

$$c_0 + c_1 X_{1i} + c_2 X_{2i} + \cdots + c_K X_{Ki} = 0$$

が成立する。ここで、$c_1 \neq 0$であるから、多重共線性が成立している。

(d)この場合、$c_0 = -1$、$c_1 = 1$、$c_2 = \cdots = c_K = 0$と設定すれば、次式が成立する。

$$c_0 + c_1 X_{1i} + c_2 X_{2i} + \cdots + c_K X_{Ki} = 0$$

(c)(d)といった状況は、データのサブサンプルを考えるときに生じやすい。たとえば、職業別の男女賃金格差に関心があり、データを収集したとしよう。このデータには男女のデータが含まれているが、歯科衛生士だけをデータから取り出した場合、男性の人数はかなり少なくなる(歯科衛生士の多くは女性)。偶然、男性が含まれないなら、女性ダミーは 1 だけになる。

(e) 多重共線性とは、$c_0 + c_1 X_{1i} + c_2 X_{2i} + \cdots + c_K X_{Ki} = 0$を満たす$c_0$、$c_1$、…、$c_K$が存在することである(ただし、$c_0 = c_1 = \cdots = c_K = 0$ではない)。$i = 1,2,\ldots,n$なので、これは$n$本の式として表せる。

$$c_0 + c_1 X_{11} + c_2 X_{21} + \cdots + c_n X_{K1} = 0$$

$$c_0 + c_1 X_{12} + c_2 X_{22} + \cdots + c_n X_{K2} = 0$$

$$\cdots\cdot$$

$$c_0 + c_1 X_{1n} + c_2 X_{2n} + \cdots + c_n X_{Kn} = 0$$

もし$n \leq K$であれば、n本の独立な式から、$K+1$個のパラメータ(c_0、c_1、…、c_K)を求めることができる(独立な式の数nより未知のパラメータの数の方が多いなら、これらの式を満たすようにパラメータを解くことができる)[11]。

[11] なぜ$n = K+1$のとき(つまり、$n-1 = K$)、多重共線性は生じないのだろうか。たしかに、$n = K+1$なら、独立な式の数とパラメータの数が一致しており、n本の独立な式から$K+1$個のパラメータを求めることができる。しかし、この解の組み合わせは 1 通りであり、これは$c_0 = c_1 = \cdots = c_K = 0$になる。これに対し、$n \leq K$なら、解の組み合わせは無数にあり、$c_0 = c_1 = \cdots = c_K = 0$ではない解が存在する。たとえば、$n = 2$、$K = 1$としよう。このとき、$c_0 + c_1 X_{11} = 0$、$c_0 + c_1 X_{12} = 0$を満たす$c_0$、$c_1$は、$c_0 = c_1 = 0$だけである。$c_1 X_{11} = -c_0$、$c_1 X_{12} = -c_0$から、$c_1 X_{11} = c_1 X_{12}$となる(つまり、$c_1(X_{11} - X_{12}) = 0$)。そして、$X_{11} \neq X_{12}$から$c_1 = 0$とわかる。また、$c_1 = 0$なら$c_0 = 0$となる。

本書では扱わないが、$n \leq K$とした状況は発生しうる。こうした状況では、Ridge、Lasso、主成分分析(principal component analysis)などの手法が有効である。これらの手法に関心のある読者は、『*Introduction to Econometrics*』（巻末参考文献[7]）の 14 章 Big Data、17 章の動学因子(Dynamic Factor)モデルを読むことをおすすめする(翻訳版は 2 版なので、これらの内容を含まない)。

練習問題 15

説明変数は 3 つあり、

$$Y_i = \alpha + \beta_1 X_i + \beta_2 W_{1i} + \beta_3 W_{2i} + u_i$$

として表せる。しかし、W_{1i}、W_{2i}を含めない単回帰モデル

$$Y_i = \alpha + \beta_1 X_i + u_i^*$$

を推定したとする。ただし、$u_i^* = \beta_2 W_{1i} + \beta_3 W_{2i} + u_i$である。

(a) 単回帰モデル($Y_i = \alpha + \beta_1 X_i + u_i^*$)において、$\hat{\beta}_1$の確率的表現は

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^{n}(X_i - \bar{X})\, u_i^*}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

となる(3.3.1 節参照)。上式に$u_i^* = \beta_2 W_{1i} + \beta_3 W_{2i} + u_i$を代入すると、

$$\begin{aligned}\hat{\beta}_1 &= \beta_1 + \frac{\sum_{i=1}^{n}(X_i - \bar{X})(\beta_2 W_{1i} + \beta_3 W_{2i} + u_i)}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \\ &= \beta_1 + \beta_2 \frac{\sum_{i=1}^{n}(X_i - \bar{X}) W_{1i}}{\sum_{i=1}^{n}(X_i - \bar{X})^2} + \beta_3 \frac{\sum_{i=1}^{n}(X_i - \bar{X}) W_{2i}}{\sum_{i=1}^{n}(X_i - \bar{X})^2} + \frac{\sum_{i=1}^{n}(X_i - \bar{X}) u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\end{aligned}$$

ここで、右辺第 2 項の分子は、次のように書き換えられる。

$$\sum_{i=1}^{n}(X_i - \bar{X}) W_{1i} = \sum_{i=1}^{n}(X_i - \bar{X}) W_{1i} - \bar{W}_1 \underbrace{\sum_{i=1}^{n}(X_i - \bar{X})}_{=0} = \sum_{i=1}^{n}(X_i - \bar{X})(W_{1i} - \bar{W}_1)$$

第 3 項の分子も同様に書き換えると、$\hat{\beta}_1$は次のようになる。

$$\beta_1 + \beta_2 \frac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(W_{1i} - \bar{W}_1)}{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2} + \beta_3 \frac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(W_{2i} - \bar{W}_2)}{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2} + \frac{\sum_{i=1}^{n}(X_i - \bar{X}) u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

$$= \beta_1 + \beta_2 \ \frac{s_{XW_1}}{s_X^2} + \beta_3 \ \frac{s_{XW_2}}{s_X^2} + \frac{\sum_{i=1}^{n}(X_i - \bar{X})u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

式展開では、右辺第 2 項、第 3 項の分子と分母を$n-1$で割っている。

X_i、W_{1i}、W_{2i}は確率変数ではないため、s_{XW_1}、s_{XW_2}、s_X^2も確率変数ではない。よって、期待値をとると、

$$E[\hat{\beta}_1] = \beta_1 + \beta_2 \ \frac{s_{XW_1}}{s_X^2} + \beta_3 \ \frac{s_{XW_2}}{s_X^2} + \frac{\sum_{i=1}^{n}(X_i - \bar{X})E[u_i]}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

$$= \beta_1 + \underbrace{\beta_2 \ \frac{s_{XW_1}}{s_X^2} + \beta_3 \ \frac{s_{XW_2}}{s_X^2}}_{\text{欠落変数バイアス}}$$

となり、これはβ_1とは異なる(式展開では、$E[u_i] = 0$を用いた)。欠落変数バイアスは、次のようになる。

$$\beta_2 \ \frac{s_{XW_1}}{s_X^2} + \beta_3 \ \frac{s_{XW_2}}{s_X^2}$$

(b) ここで、β_2とs_{XW_1}が同じ符号なら、$\beta_2 \ \frac{s_{XW_1}}{s_X^2} > 0$となる。しかし、$\beta_3 \ \frac{s_{XW_2}}{s_X^2}$が負になる可能性もあるため、$E[\hat{\beta}_1] > \beta_1$なのか、$E[\hat{\beta}_1] < \beta_1$なのかはわからない。偶然、$\beta_2 \ \frac{s_{XW_1}}{s_X^2} + \beta_3 \ \frac{s_{XW_2}}{s_X^2} = 0$なら、バイアスは 0 となる。なお、かりに$\beta_2$と$s_{XW_1}$が同じ符号であり、同様に、$\beta_3$と$s_{XW_2}$も同じ符号であるなら、$E[\hat{\beta}_1] > \beta_1$となるといえる。以上から、欠落変数が複数あるとき、バイアスの方向は、教科書で述べたほどは単純ではないことが理解できる。

練習問題 16

5.8.1 節では、完全な多重共線性は、任意の定数$c_0, c_1, \dots, c_K$を用いて、

$$c_0 + c_1X_1 + \cdots + c_jX_j + \cdots + c_KX_K = 0$$

と表せることであると定義した。上式はX_0を用いると、次のように表現できる(X_0は常に 1 となる)。

$$c_0X_0 + c_1X_1 + \cdots + c_jX_j + \cdots + c_KX_K = 0$$

ここで、任意の定数$c_0, c_1, \dots, c_K$のいずれかは 0 ではない。仮に$c_j \neq 0$とし、c_jで上式の両辺を割ると、次のようになる。

$$\frac{c_0}{c_j}X_0 + \frac{c_1}{c_j}X_1 + \cdots + X_j + \cdots + \frac{c_K}{c_j}X_K = 0$$

これをX_jについて解くと、

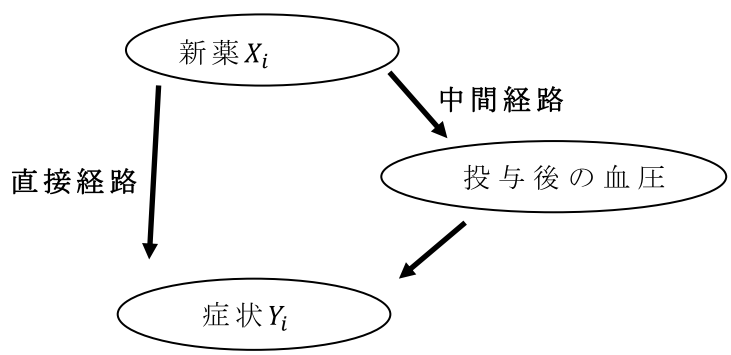
$$X_j = -\frac{c_0}{c_j}X_0 - \frac{c_1}{c_j}X_1 - \cdots - \frac{c_K}{c_j}X_K$$

$$X_j = c_0^* X_0 + c_1^* X_1 + \cdots + c_K^* X_K$$

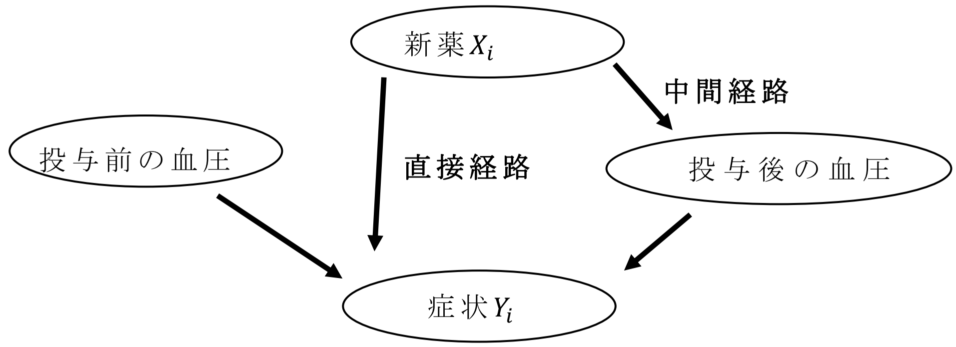
となる(ただし、$c_i^* = -c_i/c_j$と定義した)。以上から、X_jは、X_j以外の説明変数の線形関数となることがわかる。

練習問題 17

仮定により、新薬は血圧を下げることで症状を緩和させる(これは間接経路になる)。それ以外にも、血圧を経由しない直接的な効果があるかもしれない(これは直接経路になる)。しかし、投与後に測定した血圧を説明変数に含めると、新薬ダミーの係数は直接経路による効果だけとなり、新薬の中間経路が新薬の効果から除外される。



仮に、新薬の投与前に測定した血圧の情報があったとしよう。投与前の血圧が、症状と関係している可能性は否定できない(下図参照)。また、投与前に測定した血圧は、新薬の影響を受けていないため、説明変数として含めるべき変数である。つまり、被説明変数を症状とし、説明変数は新薬ダミー(投与したら 1、そうでないと 0)と血圧(投与前)とするのが適当である。ここで血圧は投与前であることに注意してほしい。



自分の関心のあるデータを分析する際は、上図のような変数間の相互関係を考えることをお勧めする。こうした作業をすることで、悪いコントロールを避けることができるだろう。

第 6 章の答え

練習問題 1

ここで推定式は次のようになる。

$$\hat{Y}_i = 20 - 5X_{1i} + 10X_{2i} + 5X_{1i}X_{2i}$$

ここで、X_{1i}は女性ダミー、X_{2i}は大卒ダミーである。

(a) 高卒男性なら、$X_{1i} = X_{2i} = 0$となるため、交差項$X_{1i}X_{2i}$も 0 となる。よって、高卒男性の所得は$\hat{Y}_i = 20$万円である。高卒女性なら、$X_{1i} = 1$、$X_{2i} = 0$となるため、交差項$X_{1i}X_{2i}$は 0 となる。よって、高卒女性の所得は$\hat{Y}_i = 20 - 5 = 15$である。以上より、

$$\text{高卒男性の所得} - \text{高卒女性の所得} = 20 - 15 = 5$$

(b) 大卒男性なら、$X_{1i} = 0$、$X_{2i} = 1$となるため、交差項$X_{1i}X_{2i}$は 0 となる。よって、大卒男性の所得は$\hat{Y}_i = 20 + 10 = 30$である。大卒女性なら、$X_{1i} = 1$、$X_{2i} = 1$となるため、交差項$X_{1i}X_{2i}$は 1 となる。よって、大卒女性の所得は$\hat{Y}_i = 20 - 5 + 10 + 5 = 30$である。

$$\text{大卒男性の所得} - \text{大卒女性の所得} = 30 - 30 = 0$$

練習問題 2

5.7.3 節では、決定係数は被説明変数を変えると異なる意味を持つため、その値の相互比較には意味がないことを学習した。本問題では、2 つのモデルは異なる被説明変数であるため、決定係数による相互比較はできない。この場合、経済理論や t 値などを参考にしながら、定式化を決めることになる。

練習問題 3

下図では、図 6-1 を再掲載している。実線は

$$\hat{Y}_t = 53.04 - 2.127X_{1t} + 0.064X_{1t}^2$$

であり、これをみると、約 17 度で電力需要が最小になっていることがわかる。

正確な気温を求めるため、

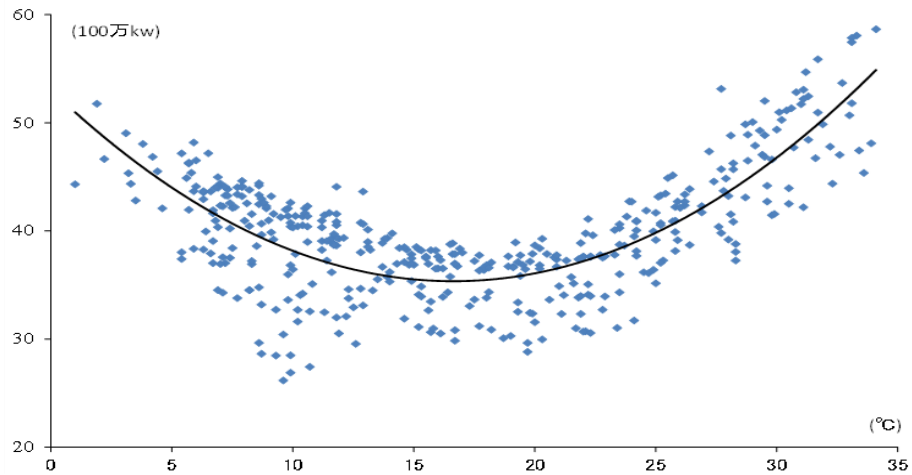
$$\hat{Y}_t = 53.04 - 2.127X_{1t} + 0.064X_{1t}^2$$

を気温X_{1t}で偏微分して 0 と置く(最小化問題の求め方は、巻末付録 B の B.3 節参照)。

$$\frac{\partial(53.04 - 2.127X_{1t} + 0.064X_{1t}^2)}{\partial X_{1t}} = -2.127 + 2 \times 0.064X_{1t} = 0$$

これを気温X_{1t}について解くと、電力需要を最も低くする気温が得られる。

$$X_{1t} = \frac{2.127}{2 \times 0.064} = 16.617$$



練習問題 4

線形回帰モデルに変換できるのは(b)のみである。(b)では、$X_{2i} = X_i^2$、$X_{3i} = X_i^3$と定義すれば、線形回帰モデルになる。

(a)式はパラメータに関して線形ではなく、また、対数をとっても線形に変換できない。

(c)式は、誤差項が和の形で含まれており、対数をとっても線形にならない。仮にモデルを積の形に変更すれば、

$$Q_i = AK_i^{\beta_1}L_i^{\beta_2}u_i^*$$

両辺の対数をとることで、線形回帰モデルに変換できる。

$$\underbrace{\ln(Q_i)}_{=Y_i} = \underbrace{\ln(A)}_{=\alpha} + \beta_1 \underbrace{\ln(K_i)}_{=X_{1i}} + \beta_2 \underbrace{\ln(L_i)}_{=X_{2i}} + \underbrace{\ln(u_i^*)}_{=u_i}$$

練習問題 5

(a) 制約$\beta_1 = 2\beta_2$を書き換えた$\beta_2 = \beta_1/2$を代入すると、

$$Y_i = \alpha + \beta_1 X_{1i} + \frac{\beta_1}{2} X_{2i} + u_i$$

$$= \alpha + \beta_1 \underbrace{\left(X_{1i} + \frac{X_{2i}}{2} \right)}_{=X_i} + u_i$$

となる。ここで、説明変数X_iを

$$X_i = X_{1i} + \frac{X_{2i}}{2}$$

と定義して、新しいモデルを

$$Y_i = \alpha + \beta_1 X_i + u_i$$

とすれば、制約$(\beta_2 = \beta_1/2)$を課した推定ができる。

(b) 制約$\alpha + \beta_1 + \beta_2 = 1$を書き換えた$\beta_2 = 1 - \alpha - \beta_1$をモデルに代入すると、

$$Y_i = \alpha + \beta_1 X_{1i} + (1 - \alpha - \beta_1) X_{2i} + u_i$$

となる。これを書き換えると、

$$Y_i - X_{2i} = \alpha(1 - X_{2i}) + \beta_1 (X_{1i} - X_{2i}) + u_i$$

となる。ここで、被説明変数を$Y_i - X_{2i}$とし、説明変数を$(1 - X_{2i})$と$(X_{1i} - X_{2i})$とすれば、αとβ_1を推定できる[12]。

練習問題 6

(a) Xが 1%変化すると、Yは$\beta \times 0.01$単位分変化する

(b) Xが 1%変化すると、Yはβ%分変化する

(c) Xが 1 単位変化すると、Yは$100 \times \beta$%分変化する

詳細は 6.3 節と補足を参照されたい。

練習問題 7

AIC の値は$p = 3$で最小となるため、$\hat{p} = 3$となる。

---

[12] モデル$Y_i - X_{2i} = \alpha(1 - X_{2i}) + \beta_1 (X_{1i} - X_{2i}) + u_i$では、説明変数は$(1 - X_{2i})$と$(X_{1i} - X_{2i})$であり、これらの係数は$\alpha$と$\beta_1$となる。このモデルに定数項は存在しないこと。つまり、モデルを推定する際、定数項がない回帰モデルとして、推定する必要がある(3 章練習問題 12 では、定数項がない回帰モデルの推定方法が書かれているので参考にしてほしい)。

練習問題 8

教科書の式展開では、

$$\ln\left(1+\frac{X'-X}{X}\right)\approx\frac{X'-X}{X}$$

という関係を用いたが、これは変化率$(X'-X)/X$が小さいときのみ成立する。ここで、変化率$(X'-X)/X$は 50%と大きいため、

$$\ln\left(1+\frac{X'-X}{X}\right)=\ln(1.5)=0.4054$$

となり、上記は 0.5 と大きく異なる。

近似を使わないで、変化量$Y'-Y$を評価しよう。単純化のため、$u=0$とすると、

$$Y'-Y=\left(\alpha+\beta\ln(X')\right)-\left(\alpha+\beta\ln(X)\right)=\beta\left(\ln(X')-\ln(X)\right)=\beta\ln\left(1+\frac{X'-X}{X}\right)$$

となる。ここで、$(X'-X)/X=0.5$とすると、

$$Y'-Y=\beta\ln(1.5)=\beta\times0.4054$$

となり、Yは$\beta\times0.4054$だけ変化するといえる($\ln(1.5)=0.4054$)。この例からも明らかなとおり、変化率が大きいときは、近似関係が使えない。

練習問題 9

単純化のため、$u=0$とする。このとき、モデルは$\ln(Y)=\alpha+\beta X$となり、

$$Y=e^{\alpha+\beta X}$$

を意味している。つまり、XがX'に変化すると、Yの変化率は、

$$\frac{Y'-Y}{Y}=\frac{e^{\alpha+\beta X'}-e^{\alpha+\beta X}}{e^{\alpha+\beta X}}=e^{\beta(X'-X)}-1$$

となる。Xの 1 単位の変化なら、Yの変化率は次のようになる。

$$\frac{Y'-Y}{Y}=e^{\beta(X'-X)}-1=e^{\beta}-1$$

練習問題 10

ラグの長さを$p=3$とすると、分布ラグモデルは、

$$Y_t=\alpha+\beta_0X_t+\beta_1X_{t-1}+\beta_2X_{t-2}+\beta_3X_{t-3}+u_t$$

となる。ここで、右辺に$\beta_0(X_{t-1}-X_{t-1})$を足すと、

$$Y_t = \alpha + \beta_0(X_t - X_{t-1}) + (\beta_0+\beta_1)X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + u_t$$

となる(ここで$\beta_0(X_{t-1}-X_{t-1})=0$であり、0 を足しても等号関係は変わらない)。さらに、右辺に$(\beta_0+\beta_1)(X_{t-2}-X_{t-2})$を足すと、

$$Y_t = \alpha + \beta_0(X_t - X_{t-1}) + (\beta_0+\beta_1)(X_{t-1}-X_{t-2}) + (\beta_0+\beta_1+\beta_2)X_{t-2} + \beta_3 X_{t-3} + u_t$$

となる。最後に、右辺に$(\beta_0+\beta_1+\beta_2)(X_{t-3}-X_{t-3})$を足すと、

$$Y_t = \alpha + \beta_0(X_t - X_{t-1}) + (\beta_0+\beta_1)(X_{t-1}-X_{t-2}) + (\beta_0+\beta_1+\beta_2)(X_{t-2}-X_{t-3}) \\ + (\beta_0+\beta_1+\beta_2+\beta_3)X_{t-3} + u_t$$

となる。ここで、

$$\Delta X_t = X_t - X_{t-1}、\Delta X_{t-1} = X_{t-1} - X_{t-2}、\Delta X_{t-2} = X_{t-2} - X_{t-3}$$

$$\theta_0 = \beta_0,\ \theta_1 = \beta_0+\beta_1,\ \theta_2 = \beta_0+\beta_1+\beta_2,\ \theta_3 = \beta_0+\beta_1+\beta_2+\beta_3$$

と定義すれば、上式は次のようになる。

$$Y_t = \alpha + \theta_0\Delta X_t + \theta_1\Delta X_{t-1} + \theta_2\Delta X_{t-2} + \theta_3 X_{t-3} + u_t$$

変形したモデルを OLS 推定し、θ_hの推定値と標準誤差を求めれば、累積動学乗数の仮説検定や 95%信頼区間の計算が可能となる。横軸を h とし、縦軸を θ_h として図示すれば、h が変わると累積動学乗数がどのように変化したのかを視覚的に示すことができる。

練習問題 11

被説明変数は$Y_i^* = c_Y Y_i$、説明変数は$X_i^* = c_X X_i$に変換する。元モデルは次のようになる。

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

6.6 節で確認した通り、上式の両辺にc_Yを掛けると、

$$c_Y Y_i = c_Y\alpha + \left(\frac{c_Y}{c_X}\beta\right)c_X X_i + c_Y u_i$$

となるため、$\alpha^* = c_Y\alpha$、$\beta^* = \frac{c_Y}{c_X}\beta$、$u_i^* = c_Y u_i$と定義すると、

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

と表現できる。ここで、u_i^*の分散をσ^{*2}と表記すると、$u_i^* = c_Y u_i$より、次のようになる。

$$\sigma^{*2} = E[u_i^{*2}] = E[(c_Y u_i)^2] = c_Y^2\,E[u_i^2] = c_Y^2\sigma^2$$

以下では、スケール変更しても、t 統計量の値は全く影響を受けないことを示す。まず、OLS 推定量$\hat{\beta}^*$は、次のように書き換えられる。

$$\hat{\beta}^* = \frac{\sum_{i=1}^n (X_i^* - \bar{X}^*)(Y_i^* - \bar{Y}^*)}{\sum_{i=1}^n (X_i^* - \bar{X}^*)^2}$$

$$= \frac{\sum_{i=1}^n (c_X X_i - c_X \bar{X})(c_Y Y_i - c_Y \bar{Y})}{\sum_{i=1}^n (c_X X_i - c_X \bar{X})^2}$$

$$= \frac{c_X c_Y}{c_X^2} \underbrace{\frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}}_{=\hat{\beta}} = \frac{c_Y}{c_X}\hat{\beta}$$

次に、OLS 推定量$\hat{\beta}^*$の分散は

$$s_{\hat{\beta}*}^2 = \frac{\sigma^{*2}}{\sum_{i=1}^n (X_i^* - \bar{X}^*)^2}$$

となる。上式に$\sigma^{*2} = c_Y^2 \sigma^2$、$X_i^* = c_X X_i$、また、

$$\bar{X}^* = \frac{\sum_{i=1}^n X_i^*}{n} = \frac{\sum_{i=1}^n c_X X_i}{n} = c_X \frac{\sum_{i=1}^n X_i}{n} = c_X \bar{X}$$

を代入すると、

$$s_{\hat{\beta}*}^2 = \frac{\sigma^{*2}}{\sum_{i=1}^n (X_i^* - \bar{X}^*)^2} = \frac{c_Y^2 \sigma^2}{\sum_{i=1}^n (c_X X_i - c_X \bar{X})^2} = \frac{c_Y^2}{c_X^2} \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

となる。以上から、t 統計量は次のようになる。

$$\frac{\hat{\beta}^* - \beta^*}{\sqrt{s_{\hat{\beta}*}^2}} = \frac{\frac{c_Y}{c_X}\hat{\beta} - \frac{c_Y}{c_X}\beta}{\sqrt{\frac{c_Y^2}{c_X^2} \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}}} = \frac{\frac{c_Y}{c_X}\left(\hat{\beta} - \beta\right)}{\frac{c_Y}{c_X}\sqrt{\frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}}} = \frac{\hat{\beta} - \beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}}} = \frac{\hat{\beta} - \beta}{\sqrt{s_{\hat{\beta}}^2}}$$

スケール変更をしても、t 統計量は影響を受けないことが確認できた。

練習問題 12

(a) 対数の性質から、$Y_i^* = c_Y Y_i$と$X_i^* = c_X X_i$の対数は、

$$\ln(Y_i^*) = \ln(c_Y) + \ln(Y_i)、\ln(X_i^*) = \ln(c_X) + \ln(X_i)$$

となる。対数対数モデルである

$$\ln(Y_i) = \alpha + \beta \ln(X_i) + u_i$$

の両辺に $\ln(c_Y)$を足して、右辺に$\beta(\ln(c_X) - \ln(c_X))$を足すと、

$$\ln(c_Y) + \ln(Y_i) = \alpha + \ln(c_Y) + \beta(\ln(c_X) - \ln(c_X)) + \beta\ln(X_i) + u_i$$

$$= (\alpha + \ln(c_Y) - \beta\ln(c_X)) + \beta(\ln(c_X) + \ln(X_i)) + u_i$$

となる($\beta(\ln(c_X) - \ln(c_X)) = 0$であるため、右辺に足しても等号関係は変わらない)。ここで、

$$\ln(Y_i^*) = \ln(c_Y) + \ln(Y_i)、\ \ln(X_i^*) = \ln(c_X) + \ln(X_i)$$

という関係に注意すると、上式は次のように書き換えることができる。

$$\ln(Y_i^*) = (\alpha + \ln(c_Y) - \beta\ln(c_X)) + \beta\ln(X_i^*) + u_i$$

また、$\alpha^* = \alpha + \ln(c_Y) - \beta\ln(c_X)$と定義すると、次のようになる。

$$\ln(Y_i^*) = \alpha^* + \beta\ln(X_i^*) + u_i$$

以上から、こうしたスケール変更では、説明変数の係数は影響を受けていないが、定数項は変化したことがわかる。

(b) 線形対数モデルである

$$Y_i = \alpha + \beta\ln(X_i) + u_i$$

の右辺に$\beta(\ln(c_X) - \ln(c_X))$を足すと、

$$Y_i = \alpha + \beta(\ln(c_X) - \ln(c_X)) + \beta\ln(X_i) + u_i$$

$$= (\alpha - \beta\ln(c_X)) + \beta(\ln(c_X) + \ln(X_i)) + u_i$$

となり、さらにc_Yを両辺に掛けると、

$$c_Y Y_i = c_y(\alpha - \beta\ln(c_x)) + c_y\beta(\ln(c_x) + \ln(X_i)) + c_y u_i$$

となる。ここで、$X_i^* = c_X X_i$の対数は、$\ln(X_i^*) = \ln(c_X) + \ln(X_i)$となることに注意すると、上式は、次のよう表現できる。

$$Y_i^* = \alpha^* + \beta^*\ln(X_i^*) + u_i^*$$

ただし、$\alpha^* = c_Y(\alpha - \beta\ln(c_X))$、$\beta^* = c_Y\beta$、$u_i^* = c_Y u_i$とした。

以上から、こうしたスケール変更によって、定数項と係数はともに変化したことがわかる。

(c) 対数線形モデルである

$$\ln(Y_i) = \alpha + \beta X_i + u_i$$

の両辺に $\ln(c_y)$を足してから展開すると、

$$\ln(c_Y) + \ln(Y_i) = \alpha + \ln(c_Y) + \beta X_i + u_i$$

$$= \alpha + \ln(c_Y) + \frac{\beta}{c_X} c_X X_i + u_i$$

となる。ここで、$Y_i^* = c_Y Y_i$、$X_i^* = c_X X_i$であることに注意すると、上式は次のように表現できる($\ln(Y_i^*) = \ln(c_Y) + \ln(Y_i)$に注意)。

$$\ln(Y_i^*) = \alpha^* + \beta^* X_i^* + u_i$$

ただし、$\alpha^* = \alpha + \ln(c_Y)$、$\beta^* = \frac{\beta}{c_X}$とした。

以上より、こうしたスケール変更によって、定数項と係数はともに変化したことになる。

練習問題 15[13]

(a) これは大まかには正しい記述だが、厳密には、正確ではない。この点を議論していく。対数の平均は、次のように表現できる。

$$\frac{1}{n}\sum_{i=1}^{n} \ln(Y_i) = \ln\left((Y_1 \times Y_2 \times \ldots \times Y_n)^{1/n}\right)$$

つまり、対数の平均は、Y_iの幾何平均の対数となる。

男性の所得をX_iとし、女性の所得をZ_iとしよう(男性は計n_1人、女性は計n_2人いるとする)。このとき、男性の対数平均から女性の対数平均を引くと、次のように展開できる。

$$\frac{1}{n_1}\sum_{i=1}^{n_1} \ln(X_i) - \frac{1}{n_2}\sum_{i=1}^{n_2} \ln(Z_i) = \ln\left(\left(X_1 \times X_2 \times \ldots \times X_{n_1}\right)^{\frac{1}{n_1}}\right) - \ln\left(\left(Z_1 \times Z_2 \times \ldots \times Z_{n_2}\right)^{\frac{1}{n_2}}\right)$$

$$= \ln\left(\frac{\left(X_1 \times X_2 \times \ldots \times X_{n_1}\right)^{\frac{1}{n_1}}}{\left(Z_1 \times Z_2 \times \ldots \times Z_{n_2}\right)^{\frac{1}{n_2}}}\right)$$

$$= \ln\left(1 + \frac{\left(X_1 \times X_2 \times \ldots \times X_{n_1}\right)^{\frac{1}{n_1}} - \left(Z_1 \times Z_2 \times \ldots \times Z_{n_2}\right)^{\frac{1}{n_2}}}{\left(Z_1 \times Z_2 \times \ldots \times Z_{n_2}\right)^{\frac{1}{n_2}}}\right)$$

---

13 本問題は Hansen『Econometrics』を参考に作成した。

$$\approx \frac{\left(X_1 \times X_2 \times \dots \times X_{n_1}\right)^{\frac{1}{n_1}} - \left(Z_1 \times Z_2 \times \dots \times Z_{n_2}\right)^{\frac{1}{n_2}}}{\left(Z_1 \times Z_2 \times \dots \times Z_{n_2}\right)^{\frac{1}{n_2}}}$$

最後の近似は、εが小さいとき、$\ln(1+\varepsilon) \approx \varepsilon$となることを用いた(巻末付録 A.3.2 節参照)。

この例では、対数平均の差は 0.3(=6.2－5.9)であることから、男性所得は女性所得より幾何平均で 30%高いといえる。男性の所得は女性より平均で 30%と高いといって誤りではないが、厳密には、平均は幾何平均であることを覚えていてほしい。

(b) 対数の平均を次のように定義する。

$$\overline{\ln(Y)} = \frac{1}{n}\sum_{i=1}^{n} \ln(Y_i)$$

OLS 推定量$\hat{\alpha}$の公式から(P32 参照)、

$$\overline{\ln(Y)} = \hat{\alpha} + \hat{\beta}\bar{X}$$

という関係が成立する。ここで、$\bar{X}$が 1 年増えると、$\overline{\ln(Y)}$は$\overline{\ln(Y)}'$に変化する。

$$\overline{\ln(Y)}' = \hat{\alpha} + \hat{\beta}(\bar{X}+1)$$

よって、$\overline{\ln(Y)}$から$\overline{\ln(Y)}'$の変化は$\hat{\beta}$となる。

$$\overline{\ln(Y)}' - \overline{\ln(Y)} = \hat{\beta}$$

この結果から、教育年数が 1 年増えると、所得は幾何平均でみて 10%増えるといえる。教育年数が 1 年増えると、所得は平均 10%増えるといって間違いではないが、厳密には、平均は幾何平均であることを覚えておいてほしい。

練習問題 16

(a) これが正しいことは、$Y_i = e^{\alpha+\beta X_i}e^{u_i}$の対数をとれば明らかである。

(b) 一般的には、$E[e^{u_i}] = 1$とはならない($E[e^{u_i}] \neq e^{E[u_i]} = e^0 = 1$)。たとえば、$u_i$が$N(0,\sigma^2)$なら、$E[e^{u_i}] = e^{\sigma^2/2}$となる[14]。$\sigma^2 = 0$なら$e^{\sigma^2/2} = 1$となるが、一般的には、$\sigma^2 > 0$から$e^{\sigma^2/2} > 1$となる(下のボックスを参照)。これは次の関係式が成立する

[14] 『入門　実践する統計学』のサポートウェブサイトにある追加資料「積率母関数と中心極限定理」の例 2 では、$X \sim N(\mu,\ \sigma^2)$なら$E[e^{tX}] = e^{\mu t + \frac{\sigma^2 t^2}{2}}$を証明している。ここで、$t = 1$、$\mu = 0$とすると、$E[e^X] = e^{\frac{\sigma^2}{2}}$となる。

ことを意味する。

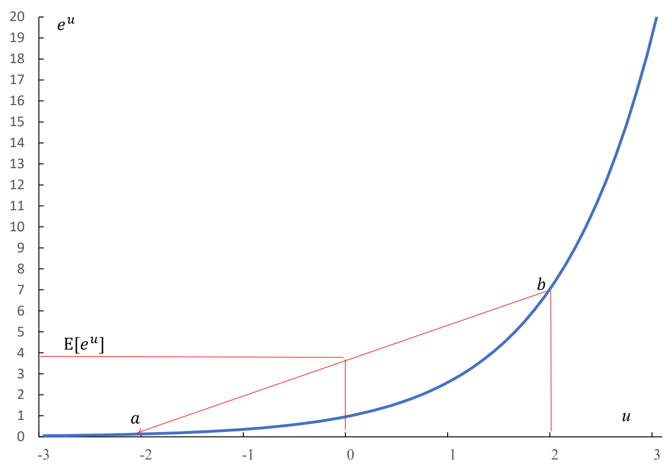
$$E[Y_i] > e^{\alpha+\beta X_i}$$

この結果から、$E[e^{u_i}]$を無視すると、期待値$E[Y_i]$を過小評価することがわかる。

$E[e^u] > 1$となる理由

下図では、uが-3 から+3 までの区間について、e^uを図示している。ここで、u が-2 もしくは+2 を確率 1/2 でとる確率変数とする。このとき、e^{-2}は点 a、e^2は点 b となる。2 点を直線でつなげた中点が$E[e^u]$である(期待値は取りうる値に確率で加重平均をとったものであることに注意してほしい)。これは$e^{E[u]} = e^0 = 1$よりも明らかに大きな値となる。なお、$E[e^u]$は u の変動が大きくなるほど大きな値となる。これは、u が-3 もしくは+3 を確率 1/2 でとる確率変数としたときを考えるとわかりやすい[15]。

図　e^uとその期待値



(c)　ここでは 4 つの方法を挙げたい(ここでは誤差項の分散は一定、つまり、均一分散を仮定している)。

第 1 の方法は、$e^{\alpha+\beta X_i}$として計算する方法である。この方法では、$E[e^{u_i}]$を考慮していないため、Y_iを過小推定することになる。しかし、これは最も簡単な方

[15] ここでは u は 2 点しかとらないとしたが、ジェンセンの不等式(Jensen's inequality)を用いて、どのような確率変数であっても$E[e^u]$は 1 以上になることを示すことができる。ジェンセンの不等式は、Bruce Hansen の「Probability and Statistics for Economists」を参照してください。

法となる。

第 2 の方法は、σ^2の推定量$s^2 = \frac{1}{n-2}\sum_{i=1}^{n}\hat{u}_i^2$を計算し、$e^{s^2/2}$として推定する方法である。残差を$\hat{u}_i = \ln(Y_i) - \hat{\alpha} - \hat{\beta}X_i$とし、それを用いて$s^2$ひいては$e^{s^2/2}$を求める。これは$u_i$が$N(0,\sigma^2)$なら良い方法であるが、正規分布の仮定が誤っていたら問題となる。

第 3 の方法は、**スミアリング推定量**(smearing estimate)と呼ばれる方法である。残差$\hat{u}_i = \ln(Y_i) - \hat{\alpha} - \hat{\beta}X_i$を計算し、$E[e^{u_i}]$を$e^{\hat{u}_i}$の平均として推定する。

$$\frac{1}{n}\sum_{i=1}^{n} e^{\hat{u}_i}$$

これは誤差項u_iの分布が正規分布でなかったとしても、一致性を満たした推定方法となる。

第 4 の方法は、定数項がないとした回帰分析を用いる方法である(定数項なしの回帰分析は 3 章練習問題 12 参照)。$m_i = e^{\alpha+\beta X_i}$と定義すると、

$$Y_i = e^{\alpha+\beta X_i}e^{u_i} = m_i e^{u_i}$$

と表現できる。そして、$E[e^{u_i}] = \theta$と定義すると、期待値は次のように表現できる。

$$E[Y_i] = m_i E[e^{u_i}] = \theta m_i$$

この結果から、被説明変数Y_iとし説明変数m_iとした OLS 推定(定数項がない)をすれば、θが推定できる。ただし、m_iが未知であるため、対数線形モデルからパラメータを推定し、それらを用いて$\hat{m}_i = e^{\hat{\alpha}+\hat{\beta}X_i}$を計算する。そして、被説明変数$Y_i$とし説明変数$\hat{m}_i$とした OLS 推定(定数項なし)をすれば$\hat{\theta}$が求められる。

$$\hat{\theta} = \frac{\sum_{i=1}^{n}\hat{m}_i Y_i}{\sum_{i=1}^{n}\hat{m}_i^2}$$

第 1 の方法は、期待値を過小推定することになるが、最も簡単な方法ではある。第 2 の方法は、正規分布の仮定に依存しており、あまり良い方法とはいえない。第 3 もしくは第 4 の方法は分布の仮定が不要となる。なお、第 4 の方法では、$E[e^{u_i}]$は 1 より小さな値として推定されることがあり、その場合、第 3 の方法を用いるとよい。

最後に、この問題で指摘した内容は対数対数モデルにも当てはまる。つまり、モデルが$\ln(Y_i) = \alpha + \beta\ln(X_i) + u_i$から$Y_i = e^{\alpha}X_i^{\beta}e^{u_i}$となり、期待値は次のようにな

る。

$$E[Y_i] = e^{\alpha} X_i^{\beta} E[e^{u_i}]$$

ここで$E[e^{u_i}] > 1$となるため、対数対数モデルを OLS 推定し残差を求めて、そこから$E[e^{u_i}]$を推定する必要がある。その方法は、対数線形モデルと同じである。

練習問題 17

(a) この式は、そのままでは、パラメータに関して線形なモデルに変形はできない。両辺の対数をとると、

$$\ln(\mathrm{Trade}_{ij}) = \ln\left(A\frac{GDP_i^{\beta_1} GDP_j^{\beta_2}}{Distance_{ij}^{\beta_3}} + u_{ij}\right)$$

となるが、右辺を線形化することはできない。右辺を線形化できるのは誤差項$u_{ij} = 0$のときだけである。一般に、重力モデルは平均的に成立している式であり、誤差項が 0 と仮定はできない。

本当のモデルは、次のように展開できる。

$$\begin{aligned}\mathrm{Trade}_{ij} &= A\frac{GDP_i^{\beta_1} GDP_j^{\beta_2}}{Distance_{ij}^{\beta_3}} + u_{ij} \\ &= A\frac{GDP_i^{\beta_1} GDP_j^{\beta_2}}{Distance_{ij}^{\beta_3}}\left(1 + \frac{1}{A}\frac{Distance_{ij}^{\beta_3}}{GDP_i^{\beta_1} GDP_j^{\beta_2}} u_{ij}\right) \\ &= A\frac{GDP_i^{\beta_1} GDP_j^{\beta_2}}{Distance_{ij}^{\beta_3}} \eta_{ij}\end{aligned}$$

両辺の対数をとると、対数対数モデルが得られる。

(b) 誤差項$\ln(\eta_{ij})$は、説明変数(両国の GDP や距離)に依存しているため、説明変数と誤差項は相関する。8.3 節で説明するが、これは内生性(説明変数と誤差

項が相関する)といわれる現象であり、OLS 推定にバイアスを生じさせる[16][17]。

(c) 貿易額は 0 をとることが多く、その場合、貿易額の対数である $\ln(\mathrm{Trade}_{ij})$が定義できない(0 の対数をとれないことは 6 章補足を参照)。このとき、貿易額の対数は欠損値として扱われ、データから除外される。貿易額が 0 をとるのは、両国間で貿易が存在しないケース、貿易額が小さすぎて 0 に丸められたケースが該当する。貿易額が 0 となるのは、貿易額が小さな国で頻繁に生じる現象であり、これらの国を除くと**セレクションバイアス**が生じる可能性がある。

貿易の重力モデルの推定に関心がある読者は、サポートウェブサイトの追加資料「貿易の重力モデル」、「カウントデータ」を参照してほしい。なお、重力モデルは、ポアソン回帰を用いるため、カウントデータの推定方法を理解する必要がある。

---

16 特殊ケースでは内生性が生じないことを説明する。これは誤差項が

$$u_{ij} = A\frac{GDP_i^{\beta_1}GDP_j^{\beta_2}}{Distance_{ij}^{\beta_3}}v_{ij}$$

とし、v_{ij}が説明変数と独立な場合になる。このとき、$\eta_{ij} = (1 + v_{ij})$となり、説明変数と$\ln(\eta_{ij})$は無相関になる。しかし、これは特殊なケースであり、一般には成立しない。

17 仮に真のモデルが

$$\mathrm{Trade}_{ij} = A\frac{GDP_i^{\beta_1}GDP_j^{\beta_2}}{Distance_{ij}^{\beta_3}}u_{ij}$$

であり、u_{ij}が説明変数と独立なら、両辺の対数をとった線形モデルを OLS 推定しても問題がない。誤差項の置き方は、一見すると小さな問題にみえるが、推定結果に大きな違いを生じさせる。

第 7 章の答え

練習問題 1

7.1.2 節で述べた通り、個別の t 検定で結合仮説を検定すると有意水準を適切に設定することが困難となる。このため、F 検定を用いて同時検定をすることが望ましい。

練習問題 2

(a) 帰無仮説H_0と対立仮説H_1は、それぞれ次のように設定すればよい。

$$H_0:\ \beta_2 = 0、\beta_3 = 0$$

H_1: 帰無仮説H_0は誤り

ここで、除外制約の数は計$q = 2$となる。

(b) 帰無仮説H_0が正しいもとで、モデルは次のようになる。

$$Y = \alpha + \beta_1 X_1 + u$$

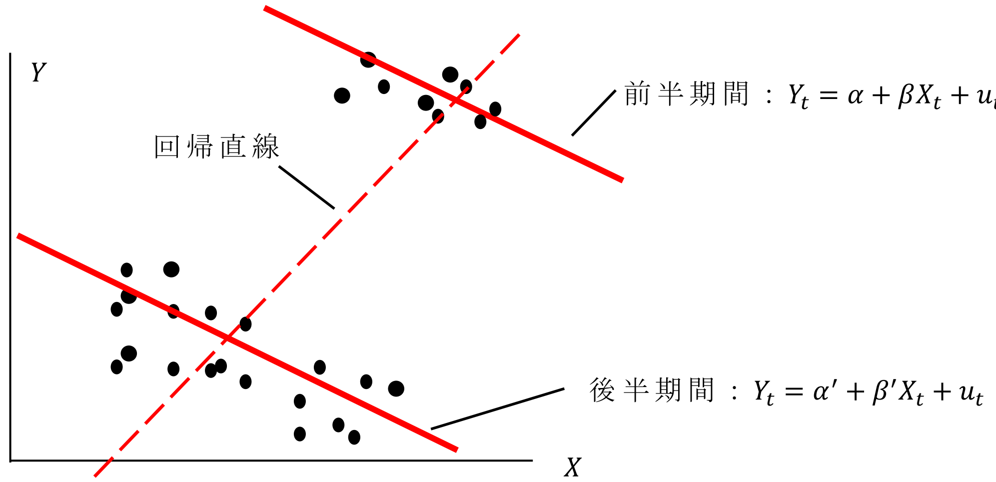
(c) $n = 50$、$K = 3$、$q = 2$、$SSR_0 = 150$、$SSR_1 = 100$を F 統計量の式に代入すると、F 値は次のようになる。

$$F = \frac{(SSR_0 - SSR_1)/q}{SSR_1/(n-K-1)} = \frac{(150-100)/2}{100/(50-3-1)} = 23 \times 0.5 = 11.5$$

帰無仮説H_0が正しいもとで、F 統計量は F 分布(自由度 2、46)に従う。このとき、有意水準 5%の臨界値は、Excel で「=FINV(0.05,2,46)」と入力すれば 3.199582 と分かる。F 値は 11.5 であり、これは臨界値 3.199582 を上回るため、帰無仮説H_0は棄却される。つまり、「説明変数X_2とX_3に説明力がない」とはいえない。

練習問題 3

様々な原因が考えられるが、ここでは 1 つの可能性を考える。下図の実線は、XとYの真の関係を表している。前半期間と後半期間とも、Xの係数は負であるが($\beta < 0$、$\beta' < 0$)、後半期間において定数項は小さくなっている($\alpha' < \alpha$)。このとき、パラメータに生じた構造変化を考慮しないで、すべてのデータをまとめて推定すると、回帰直線は点線のようになり、回帰直線の傾きは正となる。



この例から、構造変化を考慮しないでモデルを推定してしまうと、推定結果にバイアスを生じさせることが理解できる。

練習問題 4

男女のモデルを統合した次式を考える。

$$Y_i = \alpha + \beta X_i + \theta_0 F_i + \theta_1 F_i X_i + u_i$$

ここで、F_iは女性ダミーであり、パラメータθ_0とθ_1は次のように定義される。

$$\theta_0 = \alpha' - \alpha、\theta_1 = \beta' - \beta$$

たとえば、iが男性なら$F_i = F_i X_i = 0$となるので、男性のモデルは、

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

となる。iが女性なら$F_i = 1$、$F_i X_i = X_i$となるので、女性のモデルは、

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

となる。ここで、帰無仮説と対立仮説を次のように設定し F 検定をすれば、男女でパラメータが同じであるかを検証できる。

$$H_0: \theta_0 = 0, \theta_1 = 0$$

$H_1: H_0$は誤りである

練習問題 5

対立仮説H_1が正しいとしたモデルでは、残差 2 乗和は$SSR_1 = \sum_{i=1}^{n} \hat{u}_{1i}^2$となる。また、決定係数$R_1^2$は、次のようになる。

$$R_1^2 = 1 - \frac{SSR_1}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}$$

つまり、残差2乗和SSR_1は、決定係数R_1^2を用いて、次のように表現できる。

$$SSR_1 = (1 - R_1^2)\sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

同様に、帰無仮説H_0が正しいとしたモデルからの残差2乗和は次のようになる。

$$SSR_0 = (1 - R_0^2)\sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

これらを F 統計量の式に代入すると、次のようになる。

$$F = \frac{(SSR_0 - SSR_1)/q}{SSR_1/(n-K-1)} = \frac{((1-R_0^2)-(1-R_1^2))\sum_{i=1}^{n}(Y_i-\bar{Y})^2/q}{(1-R_1^2)\sum_{i=1}^{n}(Y_i-\bar{Y})^2/(n-K-1)} = \frac{(R_1^2 - R_0^2)/q}{(1-R_1^2)/(n-K-1)}$$

つまり、対立仮説H_1が正しいとしたモデルからの決定係数R_1^2が、帰無仮説H_0が正しいとしたモデルからの決定係数R_0^2より高くなると、F 統計量の値は大きくなる。

練習問題 6

帰無仮説H_0が正しいなら、モデルは、次のようになる。

$$Y_i = \alpha + u_i$$

このとき、OLS 推定量は$\hat{\alpha} = \bar{Y}$となるため、残差は$\hat{u}_{0i} = Y_i - \hat{\alpha} = Y_i - \bar{Y}$となる(2 章の練習問題 7 参照)。つまり、残差2乗和SSR_0は、Y_iの偏差2乗和となる。

$$SSR_0 = \sum_{i=1}^{n}\hat{u}_{0i}^2 = \sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

このとき、決定係数R_0^2は、次のようになる。

$$R_0^2 = 1 - \frac{SSR_0}{\sum_{i=1}^{n}(Y_i-\bar{Y})^2} = 1 - \frac{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}{\sum_{i=1}^{n}(Y_i-\bar{Y})^2} = 0$$

練習問題 6 の結果に、$R_0^2 = 0$を代入すると、次のようになる。

$$F = \frac{R_1^2/K}{(1-R_1^2)/(n-K-1)}$$

この式から、R_1^2が 1 に近づくと、F値が∞に発散することがわかる。つまり、対立仮説H_1におけるモデルを推定し、説明変数の当てはまりが良ければ、帰無仮説H_0を棄却できる。

練習問題 7

$T=200$である場合、構造変化点の候補の始期(T_{min})と終期(T_{max})は、

$$T_{min}=0.15\times 200=30$$

$$T_{max}=(1-0.15)\times 200=170$$

となる。したがって、構造変化点の候補T_Bは次のとおりである。

30、31、32、…、168、169、170

$T=1000$である場合、構造変化点の候補の始期(T_{min})と終期(T_{max})は、

$$T_{min}=0.15\times 1000=150$$

$$T_{max}=(1-0.15)\times 1000=850$$

となる。したがって、構造変化点の候補T_Bは次のとおりである。

150、151、152、…、848、849、850

練習問題 8

説明変数が 2 個の場合、定数項を含めると排除制約の数は$q=3$となる。よって、臨界値は有意水準 10%なら 4.09、5%なら 4.71、1%なら 6.02 である。$\sup F=4.50$ は、有意水準 10%の臨界値 4.09 を上回るため、有意水準 10%で帰無仮説は棄却される。つまり、パラメータに構造変化がない、とはいえない。

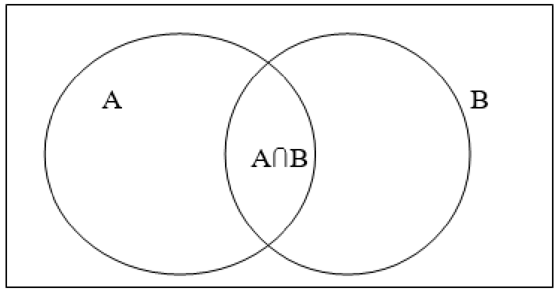
練習問題 9

統計学で学習する加法定理では、事象AまたはBが生じる確率$P\{A\cup B\}$は、

$$P\{A\cup B\}=P\{A\}+P\{B\}-P\{A\cap B\}$$

となる。つまり、事象 A または B が生じる確率は、A の確率と B の確率を足したのち、A と B が同時に生じる確率($P\{A\cap B\}$)を引いたものとなる。

下図は、ベン図を用いてこれらの事象を図式化している。長方形□で囲まれた領域内は標本空間であり、それぞれ○で囲まれた領域は事象 A と B となる。事象 A と B が重なる部分は $A\cap B$ である。図をみると、$P\{A\cup B\}$を求めるために、$P\{A\}$と$P\{B\}$の和を求めると、A と B の共通部分($P\{A\cap B\}$)が 2 回分も含まれてしまうので、その和から余分な 1 回分($P\{A\cap B\}$)を引く必要があると理解できる。これが加法定理である。確率の公理を用いた加法定理の証明は、藪友良「入門　実践する統計学」(東洋経済新報社、2012 年)の 4 章補足を参照されたい。



練習問題 10

構造変化点をT_Bとした F 統計量$F(T_B)$は、次のように表現できる。

$$F(T_B) = \frac{(SSR_0 - SSR_1)/(K+1)}{SSR_1/(T-2(K+1))} = \frac{T-2(K+1)}{K+1}\left(\frac{SSR_0}{SSR_1} - 1\right)$$

ここでは、$F(T_B)$の構成要素において、構造変化点T_Bに依存しているのは残差 2 乗和SSR_1のみであることを示す。

まず、残差 2 乗和SSR_0は、帰無仮説H_0(構造変化なし)が正しい前提で、(7)式、つまり、下式を推定することで得られる。

$$Y_t = \alpha + \beta_1 X_{1t} + \ldots + \beta_K X_{Kt} + u_t$$

このため、残差 2 乗和SSR_0は、どのT_Bを用いても同じ値となる。

次に、残差 2 乗和SSR_1は、対立仮説H_1(構造変化あり)が正しい前提で、(6)式、つまり、下式を推定することで得られる。

$$Y_t = \alpha + \beta_1 X_{1,t} + \cdots + \beta_K X_{K,t} + \theta_0 D_t + \theta_1 D_t X_{1,t} + \ldots + \theta_K D_t X_{K,t} + u_t$$

D_tは時点tが前半期間(1、2、…、T_B)ならば 0 をとり、後半期間($T_B + 1$、$T_B + 2$、…、T)ならば 1 をとる。D_tはT_Bの選択によって値が変わるため、残差 2 乗和SSR_1もT_Bの選択によって値が変わる。

以上から、$F(T_B)$の構成要素のうち、T_Bに依存しているのは残差 2 乗和SSR_1のみである。残差 2 乗和SSR_1が小さいほど$F(T_B)$が大きくなることに注意すると、$F(T_B)$を最大にするT_Bとは、(6)式の残差 2 乗和SSR_1を最小にするT_Bに他ならない。

第 8 章の答え

練習問題 1

均一分散は現実には成立しないことが多く、不均一分散が現実的仮定である。

練習問題 2

無作為抽出した横断面データでは、ランダムになっているため、誤差項は相互に無相関となる。

練習問題 3

時系列データでは、誤差項は相互に関係している。これは、何らかのイベントが発生すると、それは現在だけでなく、将来にも影響することが多いためである。

練習問題 4

説明変数X_iと誤差項u_iに相関がないとき、つまり、$Cov(X_i, u_i) = 0$であれば、説明変数には外生性がある、という。これに対し、説明変数X_iと誤差項u_iに相関があるとき、つまり、$Cov(X_i, u_i) \neq 0$であれば、説明変数には内生性がある、という。

練習問題 5

説明変数に内生性があると、OLS 推定量は不偏性だけでなく、一致性も持たない。つまり、サンプルサイズが大きくなっても、バイアスは消えない。

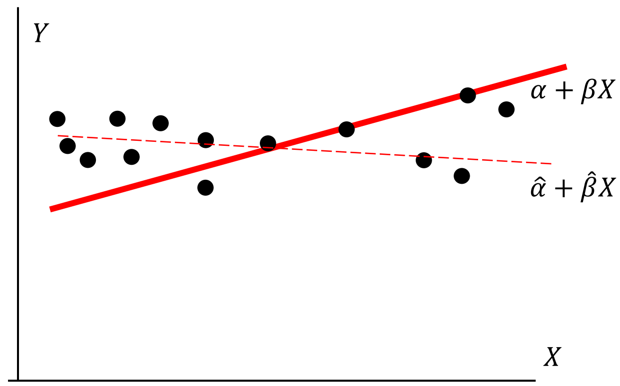
練習問題 6

下図において、実線が真のXとYの関係$(\alpha + \beta X)$を表し、点線が推定された回帰直線$(\hat{\alpha} + \hat{\beta} X)$を表す。ただし、係数$\beta$は正$(\beta > 0)$とし、実線$(\alpha + \beta X)$は右上がりの関係となる。また、説明変数$X_i$と誤差項$u_i$に負の相関がある$(Cov(X_i, u_i) < 0)$。

説明変数X_iと誤差項u_iに負の相関があるため、説明変数X_iが小さな値だと、誤差項u_iは大きな値となり、$Y_i = \alpha + \beta X_i + u_i$という関係から実線$(\alpha + \beta X)$の上でデータが観察されやすくなる。逆に、説明変数$X_i$が大きな値だと、誤差項$u_i$は小さな値となり(つまり、負の値)、$Y_i = \alpha + \beta X_i + u_i$という関係から実線$(\alpha + \beta X)$より下で

データが観察されやすくなる。したがって、回帰直線は点線のようになり、OLS 推定量$\hat{\beta}$は負のバイアスを持つ。

図　内生性とバイアスの関係



練習問題 7

推定量β^*の確率的表現は、次のようになる。

$$\beta^* = \frac{Y_2 - Y_1}{X_2 - X_1} = \frac{(\alpha + \beta X_2 + u_2) - (\alpha + \beta X_1 + u_1)}{X_2 - X_1}$$

$$= \frac{\beta(X_2 - X_1) + (u_2 - u_1)}{X_2 - X_1} = \beta + \frac{u_2 - u_1}{X_2 - X_1}$$

まず、確率的表現の期待値をとると、

$$E[\beta^*] = E\left[\beta + \frac{u_2 - u_1}{X_2 - X_1}\right] = \beta + \frac{E[u_2] - E[u_1]}{X_2 - X_1} = \beta$$

となり、不偏性を満たすことが確認できる(誤差項は標準的仮定を満たすため、$E[u_1] = E[u_2] = 0$とした)。

次に、推定量の分散は、

$$E\left[\left(\frac{Y_2 - Y_1}{X_2 - X_1} - \beta\right)^2\right] = E\left[\left(\frac{u_2 - u_1}{X_2 - X_1}\right)^2\right] = \frac{E[(u_2 - u_1)^2]}{(X_2 - X_1)^2}$$

$$= \frac{E[u_1^2] + E[u_2^2] - 2E[u_1 u_2]}{(X_2 - X_1)^2} = \frac{2\sigma^2}{(X_2 - X_1)^2}$$

となる(誤差項は標準的仮定を満たすため、$E[u_1^2] = E[u_2^2] = \sigma^2$、$E[u_1 u_2] = 0$とした)。ここで、分散の分母は$(X_2 - X_1)^2$であり、$X_1$と$X_2$が互いに離れているほど、分散は小さくなることがわかる。

練習問題 8[1]

(a) 平均 2 乗誤差(MSE)は、次のように分解できる。

$$
\begin{aligned}
MSE &= E\left[\left(\hat{\theta}-\theta\right)^2\right] \\
&= E\left[\left(\left(\hat{\theta}-\mathrm{E}[\hat{\theta}]\right)+\left(\mathrm{E}[\hat{\theta}]-\theta\right)\right)^2\right] \\
&= E\left[\left(\hat{\theta}-\mathrm{E}[\hat{\theta}]\right)^2\right]+\left(\mathrm{E}[\hat{\theta}]-\theta\right)^2+2\left(\mathrm{E}[\hat{\theta}]-\theta\right)E\left[\left(\hat{\theta}-\mathrm{E}[\hat{\theta}]\right)\right] \\
&= E\left[\left(\hat{\theta}-\mathrm{E}[\hat{\theta}]\right)^2\right]+\left(\mathrm{E}[\hat{\theta}]-\theta\right)^2
\end{aligned}
$$

式展開では、$(\mathrm{E}[\hat{\theta}]-\theta)$は固定した値なので期待値の外に出せること、$E[(\hat{\theta}-\mathrm{E}[\hat{\theta}])]=\mathrm{E}[\hat{\theta}]-\mathrm{E}[\hat{\theta}]=0$であることを用いた。

上式の右辺第 1 項は推定量の分散$V(\hat{\theta})=E\left[\left(\hat{\theta}-\mathrm{E}[\hat{\theta}]\right)^2\right]$となる。右辺第 2 項は、推定量のバイアス$Bias(\hat{\theta})=\mathrm{E}[\hat{\theta}]-\theta$の 2 乗となる。以上から、MSE は次のように表現できる。

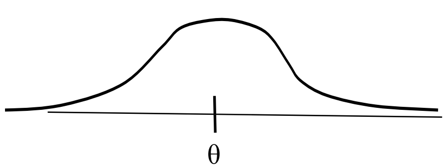
$$MSE=E\left[\left(\hat{\theta}-\theta\right)^2\right]=V(\hat{\theta})+Bias(\hat{\theta})^2$$

一般に、「MSE が小さいほど良い推定量である」と判断される。MSE が小さい推定量とは、推定量の分散$V(\hat{\theta})$が小さく、推定量のバイアス$Bias(\hat{\theta})$も小さい推定となる。

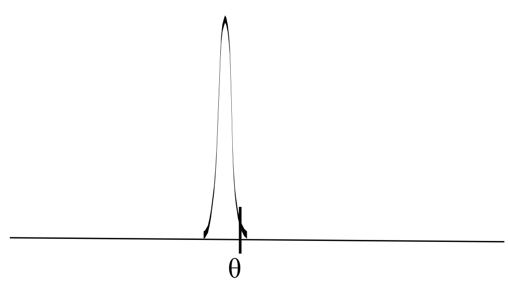
MSE の理解を深めるため、下図では、2 種類の推定量を示している。(a)の推定量は不偏性を満たしているが、推定量の分散は大きくなっている。これに対して、(b)の推定量はバイアスはあるが、推定量の分散は小さくなっている。MSE でみると、(b)のほうが小さくなるため、(b)がより望ましい推定量と判断される。

図：2 つの推定量の比較

(a)バイアスなし、分散は大きい



（b）バイアスあり、分散は小さい



[1] なお、藪友良『入門　実践する統計学』(東洋経済新報社、2012 年)の 7 章では、推定量が不偏性を満たしている場合、推定量が有効とは推定量の分散が小さいこととした。本問題では、推定量が不偏性を満たしていない可能性を考慮した有効性の一般的な定義を示している。

本書では扱わないが、リッジ推定量はバイアスを持っているが、推定量の分散は小さいため、MSE の小さい推定量の 1 つとして知られている。

(b) $\hat{\theta}$が不偏推定量であるとしよう。このとき、

$$\mathrm{E}\left[\hat{\theta}\right] = \theta$$

であるため、バイアスは

$$Bias\left(\hat{\theta}\right) = \mathrm{E}\left[\hat{\theta}\right] - \theta = 0$$

となる。このため、MSE は、推定量の分散$V\left(\hat{\theta}\right)$と一致する。

$$MSE = E\left[\left(\hat{\theta} - \theta\right)^2\right] = V\left(\hat{\theta}\right)$$

つまり、不偏推定量であれば、分散$V\left(\hat{\theta}\right)$が小さい推定量が最も望ましいといえる(有効な推定量となる)。ガウス=マルコフの定理では、不偏推定量だけを考えていたため、分散が最小となる OLS 推定量が有効な推定量としていた。

練習問題 9

$Y_t = \alpha + \beta Y_{t-1} + u_t$が正しいなら、$Y_{t-1} = \alpha + \beta Y_{t-2} + u_{t-1}$が成立することになる。ここで、誤差項$u_{t-1}$は確率変数なので、$Y_{t-1}$は確率変数となる。

練習問題 10

まず、不偏性について考えよう。単回帰分析において、OLS 推定量は次のようになる。

$$\hat{\beta} = \beta + \frac{\sum_{i=1}^{n}(X_i - \bar{X})u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

ここで期待値をとると、

$$E[\hat{\beta}] = \beta + E\left[\frac{\sum_{i=1}^{n}(X_i - \bar{X})u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right]$$

となる。ここで、X_iは確率変数であるため、一般には、

$$E\left[\frac{\sum_{i=1}^{n}(X_i - \bar{X})u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right] \neq 0$$

となり、不偏性は成立しない[2]。

説明変数が確率変数であるとき、$\hat{\beta}$に不偏性が成立するためには、外生性($Cov(X_i, u_i) = 0$)よりも強い仮定が必要となる。たとえば、説明変数{X_1、X_2、…、X_n}と誤差項{u_1、u_2、…、u_n}が互いにすべて独立と仮定しよう。このとき、独立性の仮定から、

$$E\left[\frac{\sum_{i=1}^{n}(X_i - \bar{X})u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right] = E\left[\frac{(X_1 - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right]E[u_1] + \cdots + E\left[\frac{(X_n - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right]E[u_n] = 0$$

となる。式展開では、$E[u_i] = 0$を用いた。この仮定は、全てのiとjについて$Cov(X_i, u_j) = 0$を意味し、かなり強い仮定である。ただし、無作為抽出抽出したデータなら、u_iはX_i以外の説明変数と独立なので、これは外生性($Cov(X_i, u_i) = 0$)だけを意味し、それほど強い仮定ではない。しかし、時系列データでは、全てのiとjについて$Cov(X_i, u_j) = 0$となる状況は考えにくい。以上から、無作為抽出による横断面データでなければ、OLS 推定量は不偏性を満たさない可能性が高い。

次に、一致性について考えよう。確率的表現の第 2 項の分子と分母をnで割ると、以下の式が得られる。

$$\hat{\beta} = \beta + \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})u_i}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

ここで、nが非常に大きいとする。このとき、$\bar{X}$は$E[X_i]$に置き換えることができる(つまり、$(X_i - \bar{X})u_i$は$(X_i - E[X_i])u_i$に置き換えてもよい)。また、X_iは外生変数であるから$Cov(X_i, u_i) = 0$であり、$E[(X_i - E[X_i])u_i]$は 0 となる。したがって、第 2 項目の分子は 0 に収束していく。また、$\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$は標本分散であることから、第 2 項目の分母は$X_i$の分散に収束する[3]。以上から、$n$が大きくなると、$\hat{\beta}$は$\beta$に収束するといえる。なお、平均が期待値に収束するためには、追加条件が必要となるが、そうした条件に関心がある読者は巻末参考文献[7][8][9]を参照してほしい。

[2] 標準的仮定 1 が満たされるなら、説明変数は非確率変数であり、

$$E\left[\frac{\sum_{i=1}^{n}(X_i - \bar{X})u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right] = \frac{\sum_{i=1}^{n}(X_i - \bar{X})E[u_i]}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = 0$$

となる。式展開では、Xは固定した値なので期待値の外に出せること、また、標準的仮定 3($E[u_i] = 0$)を用いた。しかし、Xが確率変数なら、こうした式展開はできない。

[3] 厳密には$n-1$で割ったものだが、サンプルサイズnが大きいとき、どちらで割っても同じ。

第 9 章の答え

練習問題 1

被説明変数がダミー変数であるとき、$Y_i = 1$となる確率P_iは、次のような線形モデルで表せる。

$$P_i = \alpha + \beta X_i$$

これが線形確率モデルと言われる理由である(詳しくは例 9-2 参照)。

練習問題 2

OLS 推定量は不偏性と一致性を持つ。しかし、ガウス=マルコフの条件が満たされないため、有効性は満たされない。なお、OLS 推定量$\hat{\beta}$の分散は、

$$\sigma_{\hat{\beta}}^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sigma_i^2}{\left(\sum_{i=1}^{n}(X_i - \bar{X})^2\right)^2}$$

であるため、通常の標準誤差ではなく、ロバスト標準誤差を用いることが必要である。

練習問題 3

被説明変数Y_iは、企業から連絡があれば 1、連絡がなければ 0 となるダミー変数となり、説明変数X_iは、黒人固有の名前なら 1、白人固有の名前なら 0 となるダミー変数となる。誤差項としては、個人iの属性が考えられる。名前はランダムに割り当てられたため、説明変数X_iは個人属性などを表す誤差項と無相関となっており、OLS 推定にバイアスは生じない(8.3 節参照)。

練習問題 4

ここで、$u_i^* = \sqrt{N_i} u_i$である。よって、u_i^*の分散は、次のようになる。

$$E[u_i^{*2}] = E\left[\left(\sqrt{N_i} u_i\right)^2\right] = N_i E[u_i^2]$$

また、$E[u_i^2] = \frac{\sigma^2}{N_i}$であるから、次式のように誤差項$u_i^*$の分散は$\sigma^2$で一定である。

$$N_i E[u_i^2] = N_i \frac{\sigma^2}{N_i} = \sigma^2$$

練習問題 5

ここで、$h_i = Z_i$であるため、$Y_i = \alpha + \beta X_i + u_i$の両辺に$1/\sqrt{h_i} = 1/\sqrt{Z_i}$を掛けると、

$$\underbrace{\frac{Y_i}{\sqrt{Z_i}}}_{=Y_i^*} = \alpha \underbrace{\frac{1}{\sqrt{Z_i}}}_{=X_{1i}^*} + \beta \underbrace{\frac{X_i}{\sqrt{Z_i}}}_{=X_{2i}^*} + \underbrace{\frac{u_i}{\sqrt{Z_i}}}_{=u_i^*}$$

となる。新しい誤差項u_i^*の分散は、

$$E[u_i^{*2}] = E\left[\left(\frac{u_i}{\sqrt{Z_i}}\right)^2\right] = \frac{1}{Z_i}E[u_i^2] = \frac{1}{Z_i}cZ_i = c$$

となるため、均一分散を満たす。被説明変数をY_i^*、説明変数をX_{1i}^*、X_{2i}^*とした OLS 推定をすれば WLS 推定量となる。

練習問題 6

ここで、$h_i = X_i^2$であるため、$Y_i = \alpha + \beta X_i + u_i$の両辺に$1/\sqrt{h_i} = 1/|X_i|$を掛けると、

$$\underbrace{\frac{Y_i}{|X_i|}}_{=Y_i^*} = \alpha \underbrace{\frac{1}{|X_i|}}_{=X_{1i}^*} + \beta \underbrace{\frac{X_i}{|X_i|}}_{=X_{2i}^*} + \underbrace{\frac{u_i}{|X_i|}}_{=u_i^*}$$

となる(X_iは負の値をとる可能性があるため、ここでは絶対値をとっている)。新しい誤差項u_i^*の分散は、

$$E[u_i^{*2}] = E\left[\left(\frac{u_i}{|X_i|}\right)^2\right] = \frac{1}{X_i^2}E[u_i^2] = \frac{1}{X_i^2}cX_i^2 = c$$

となるため、均一分散を満たす。被説明変数をY_i^*、説明変数をX_{1i}^*、X_{2i}^*とした OLS 推定をすれば WLS 推定量となる。

練習問題 7

WLS 推定について考えてみよう。仮にσ_iが分かっているならば、元の式をσ_iで割ると、

$$\underbrace{\frac{Y_i}{\sigma_i}}_{=Y_i^*} = \alpha \underbrace{\frac{1}{\sigma_i}}_{=X_{1i}^*} + \beta \underbrace{\frac{X_i}{\sigma_i}}_{=X_{2i}^*} + \underbrace{\frac{u_i}{\sigma_i}}_{=u_i^*}$$

となり、新しい誤差項u_i^*の分散は、

$$E[u_i^{*2}] = E\left[\left(\frac{u_i}{\sigma_i}\right)^2\right] = \frac{1}{\sigma_i^2}E[u_i^2] = \frac{1}{\sigma_i^2}\sigma_i^2 = 1$$

となるため、均一分散を満たす。現実には、分析者はσ_iの値を知らないため、予測値$\hat{\sigma}_i$を用いた FWLS を行う。

まず、$Y_i = \alpha + \beta X_i + u_i$とした OLS 推定によって残差$\hat{u}_i$を求める。残差$\hat{u}_i$は誤差

項u_iの推定量である。ここで、$\sigma_i^2 = E[u_i^2] = c_0 + c_1 Z_i$から、被説明変数を$\hat{u}_i^2$とし、説明変数を$Z_i$とした OLS 推定によって、パラメータ($c_0$、$c_1$)を推定でき、分散の予測値$\hat{\sigma}_i^2 = \hat{c}_0 + \hat{c}_1 Z_i$を求めることができる。次に、元のモデルを予測値$\hat{\sigma}_i$で割ることで、次の式が得られる。

$$\underbrace{\frac{Y_i}{\hat{\sigma}_i}}_{=Y_i^*} = \alpha \underbrace{\frac{1}{\hat{\sigma}_i}}_{=X_{1i}^*} + \beta \underbrace{\frac{X_i}{\hat{\sigma}_i}}_{=X_{2i}^*} + \underbrace{\frac{u_i}{\hat{\sigma}_i}}_{=u_i^*}$$

サンプルサイズが十分に大きければ、推定量$\hat{\sigma}_i$は真の値σ_iとなるため、新しい誤差項u_i^*は$u_i/\hat{\sigma}_i$で均一分散を満たす。このため、被説明変数をY_i^*、説明変数をX_{1i}^*、X_{2i}^*とした OLS 推定は、FWLS 推定量となる。

練習問題 10[4]

(a) 均一分散のもとで、$E[\hat{u}_i^2] = \sigma^2(1 - h_{ii})$となる。これを用いると、

$$E\left[s_{\hat{\beta}}^2\right] = \frac{\sum_{i=1}^n (X_i - \bar{X})^2 E[\hat{u}_i^2]}{\left(\sum_{i=1}^n (X_i - \bar{X})^2\right)^2} = \frac{\sigma^2 \sum_{i=1}^n (X_i - \bar{X})^2 (1 - h_{ii})}{\left(\sum_{i=1}^n (X_i - \bar{X})^2\right)^2}$$

$$= \frac{\sigma^2 \sum_{i=1}^n (X_i - \bar{X})^2}{\left(\sum_{i=1}^n (X_i - \bar{X})^2\right)^2} - \frac{\sigma^2 \sum_{i=1}^n (X_i - \bar{X})^2 h_{ii}}{\left(\sum_{i=1}^n (X_i - \bar{X})^2\right)^2}$$

$$= \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2} - \frac{\sigma^2 \sum_{i=1}^n (X_i - \bar{X})^2 h_{ii}}{\left(\sum_{i=1}^n (X_i - \bar{X})^2\right)^2}$$

となる。また、h_{ii}は正であるから、右辺第 2 項はマイナスであり、次式が成立する。

$$E\left[s_{\hat{\beta}}^2\right] < \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

つまり、均一分散が正しいとき、不均一分散に対して頑健な分散の推定量$s_{\hat{\beta}}^2$は、真の分散$\frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$を過小評価してしまう。

(b) 不均一分散に対して頑健な分散の推定量$s_{\hat{\beta}}^2$は、真の分散を過小評価するという問題がある。Stata では、不均一分散に対して頑健な分散の推定量$s_{\hat{\beta}}^2$に$\frac{n}{n-2}$を掛けることで、分散を少し大きめに推定し、過小評価の問題を軽減している。

[4] 本問題は Hansen『Econometrics』を参考に作成した。

$\frac{n}{n-2}$という値は、どのように正当化されるのだろうか。これは$s^2=\frac{1}{n-2}\sum_{i=1}^{n}\hat{u}_i^2$の計算と整合的な調整ともいえる。3.4.1 節で確認したとおり、理想的なσ^2の推定量は$\frac{1}{n}\sum_{i=1}^{n}u_i^2$である。しかし、誤差項$u_i$が観察できないため、残差$\hat{u}_i$で置き換えることになる。このとき、$\frac{1}{n}\sum_{i=1}^{n}\hat{u}_i^2$ではなく、それに$\frac{n}{n-2}$を掛けた$s^2$を用いた。

$$s^2=\frac{n}{n-2}\left(\frac{1}{n}\sum_{i=1}^{n}\hat{u}_i^2\right)=\frac{1}{n-2}\sum_{i=1}^{n}\hat{u}_i^2$$

同じ調整を、不均一分散に対して頑健な分散の推定量$s_{\hat{\beta}}^2$に行っているのが、Stata の調整といえる。なお、重回帰分析では、

$$s^2=\frac{n}{n-K-1}\left(\frac{1}{n}\sum_{i=1}^{n}\hat{u}_i^2\right)=\frac{1}{n-K-1}\sum_{i=1}^{n}\hat{u}_i^2$$

となるから、不均一分散に対して頑健な分散の推定量に

$$\frac{n}{n-K-1}$$

を掛けることになる。

では、こうした調整を行うことで、不偏性が満たされるのだろうか。残念ながら、不偏性が満たされるのは特殊ケースになる。この点を確認しよう。レバレッジの平均は、次のようになる。

$$\frac{1}{n}\sum_{i=1}^{n}h_{ii}=\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{n}+\frac{(X_i-\bar{X})^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}\right)=\frac{2}{n}$$

ここで、レバレッジh_{ii}は常に同じ値であるとしよう。レバレッジh_{ii}は平均$2/n$と同じであるため、$1-h_{ii}=1-2/n=(n-2)/n$となる。このとき、

$$E\left[s_{\hat{\beta}}^2\right]=\frac{\sigma^2\sum_{i=1}^{n}(X_i-\bar{X})^2\,(1-h_{ii})}{\left(\sum_{i=1}^{n}(X_i-\bar{X})^2\right)^2}=\frac{n-2}{n}\frac{\sigma^2\sum_{i=1}^{n}(X_i-\bar{X})^2}{\left(\sum_{i=1}^{n}(X_i-\bar{X})^2\right)^2}=\frac{n-2}{n}\frac{\sigma^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}$$

となる。よって、不均一分散に対して頑健な分散の推定量として、

$$\hat{V}^{HC1}=\frac{n}{n-2}\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2\hat{u}_i^2}{\left(\sum_{i=1}^{n}(X_i-\bar{X})^2\right)^2}$$

を用いれば、不偏性が満たされる。

以上をまとめると、Stata の$\hat{V}^{HC1}$は、σ^2の推定量s^2と整合的な調整になっているが、不偏性が満たされるのは特殊ケースであり、理論的根拠が十分とはいえない[5]。

[5] Stata では、reg Y X, r とすると、$\hat{V}^{HC1}$の平方根がロバスト標準誤差として計算される。

(c) 標準化残差$\bar{u}_i = (1-h_{ii})^{-1/2}\hat{u}_i$を使った場合、$E[\bar{u}_i^2]=\sigma^2$が成立する。このため、$\hat{V}^{HC2}$の期待値は、

$$E\left[\frac{\sum_{i=1}^n (X_i-\bar{X})^2\bar{u}_i^2}{\left(\sum_{i=1}^n (X_i-\bar{X})^2\right)^2}\right] = \frac{\sum_{i=1}^n (X_i-\bar{X})^2 E[\bar{u}_i^2]}{\left(\sum_{i=1}^n (X_i-\bar{X})^2\right)^2}$$

$$= \frac{\sigma^2\sum_{i=1}^n (X_i-\bar{X})^2}{\left(\sum_{i=1}^n (X_i-\bar{X})^2\right)^2} = \frac{\sigma^2}{\sum_{i=1}^n (X_i-\bar{X})^2}$$

となり、$\hat{V}^{HC2}$は不偏性を満たす[6]。

(d) 理論的には、$\hat{V}^{HC2}$が優れているが、実証分析では、$\hat{V}^{HC1}$がよく用いられている(教科書でも、V^{HC1}を用いた)。通常、どちらを用いても同じような値になるので、どちらを使っても問題はない。しかし、レバレッジh_{ii}が大きいデータがあれば、$\hat{V}^{HC1}$の方がより小さな値をとる傾向がある。その差が大きいときは、$\hat{V}^{HC2}$を用いるほうがよい。なお、サンプルサイズが大きければ、レバレッジh_{ii}は 0 に収束するため、いずれを用いても同じ結果となる(h_{ii}の定義を思い出してほしい)。

自分でデータ分析する際は、$\hat{V}^{HC1}$と$\hat{V}^{HC2}$の両方を計算し、それらの値を比較することが望ましい。両者の差が大きく異なるようなら、レバレッジを計算し、どの観測値で大きくなっているかを確認しよう。かりにそれが外れ値のようなものなら、そのデータを除去することも、選択肢の 1 つとして考えられる。たとえば、小学生のデータを分析したところ、身長が 210cm の生徒がいたとしよう。これは入力間違いの可能性があるし、たとえ正しい情報であっても外れ値と考えられる。

練習問題 11[7]

単純化のため、$D_1=1$、それ以外のD_2、D_3、…、D_nは 0 としよう。

(a) 説明変数D_iの平均は

$$\bar{D} = \frac{1}{n}\sum_{i=1}^n D_i = \frac{1}{n}$$

となる。このため、偏差は、

---

6 Stata では、reg Y X, vce(hc2)とすれば、$\hat{V}^{HC2}$の平方根がロバスト標準誤差として計算される。

7 本問題は Hansen『Econometrics』を参考に作成した。

$$D_i - \bar{D} = \begin{cases} 1 - \dfrac{1}{n} = \dfrac{n-1}{n} & \text{if} \quad i = 1 \\ 0 - \dfrac{1}{n} = -\dfrac{1}{n} & \text{if} \quad i > 1 \end{cases}$$

となり、偏差 2 乗和は次のようになる。

$$\sum_{i=1}^{n}(D_i - \bar{D})^2 = (D_1 - \bar{D})^2 + \sum_{i=2}^{n}(D_i - \bar{D})^2$$

$$= \left(\frac{n-1}{n}\right)^2 + (n-1)\left(-\frac{1}{n}\right)^2$$

$$= \frac{n-1}{n^2}((n-1)+1) = \frac{n-1}{n}$$

(b) 誤差項は均一分散であるため、$\hat{\beta}$の分散は次のようになる。

$$V(\hat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^{n}(D_i - \bar{D})^2} = \frac{n}{n-1}\sigma^2$$

この場合、係数βを推定するための情報は、$D_1 = 1$だけしか存在せず、サンプルサイズが大きくなっても分散はσ^2となる。つまり、サンプルサイズが大きくなっても、$\hat{\beta}$の推定精度は改善しないことがわかる[8]。

(c) レバレッジを求めよう。$\sum_{i=1}^{n}(D_i - \bar{D})^2 = (n-1)/n$に注意すると、

$$\frac{(D_i - \bar{D})^2}{\sum_{i=1}^{n}(D_i - \bar{D})^2} = \begin{cases} \dfrac{[(n-1)/n]^2}{(n-1)/n} = \dfrac{n-1}{n} & \text{if} \quad i = 1 \\ \dfrac{(-1/n)^2}{(n-1)/n} = \dfrac{1}{n(n-1)} & \text{if} \quad i > 1 \end{cases}$$

となり、レバレッジh_{ii}は、次のようになる。

$$h_{ii} = \begin{cases} \dfrac{1}{n} + \dfrac{n-1}{n} = 1 & \text{if} \quad i = 1 \\ \dfrac{1}{n} + \dfrac{1}{n(n-1)} = \dfrac{1}{n-1} & \text{if} \quad i > 1 \end{cases}$$

ここで、$h_{11} = 1$となり、D_1は他のデータに比べて大きく異なることがわかる。これはD_1が 1 となり、他のD_iは 0 となることから明らかであろう。

D_iは$i = 1$のとき 1、他では 0 となるから、$\hat{\beta}$は$\hat{u}_1$を 0 とするように選ばれる(『入門実践する統計学』12.4.1 節参照)。$\hat{u}_1 = 0$を用いると、$\hat{V}^{HC1}$の分子は

---

[8] このケースでは、nが大きくなると、偏差 2 乗和は 1 に収束している。したがって、標準的仮定 2 は満たされない。これはnが大きくなっても、$\hat{\beta}$の分散が 0 にならないことを示唆している。

$$\sum_{i=1}^{n}(D_i-\bar{D})^2E[\hat{u}_i^2]=\sum_{i=1}^{n}\left(-\frac{1}{n}\right)^2E[\hat{u}_i^2]=\frac{1}{n^2}\sum_{i=1}^{n}E[\hat{u}_i^2]$$

となる。また、$E[\hat{u}_i^2]=(1-h_{ii})\,\sigma^2$であるから、

$$\sum_{i=1}^{n}E[\hat{u}_i^2]=\sum_{i=1}^{n}(1-h_{ii})\sigma^2$$

$$=(1-1)\sigma^2+\sum_{i=2}^{n}\left(1-\frac{1}{n-1}\right)\sigma^2$$

$$=(n-2)\sigma^2$$

となる。これらの結果を用いると、次のように展開できる。

$$E\left[\hat{V}^{HC1}\right]=\frac{n}{n-2}\frac{\sum_{i=1}^{n}(D_i-\bar{D})^2E[\hat{u}_i^2]}{\left(\sum_{i=1}^{n}(D_i-\bar{D})^2\right)^2}$$

$$=\frac{n}{n-2}\frac{\frac{1}{n^2}(n-2)\sigma^2}{\left(\frac{n-1}{n}\right)^2}$$

$$=\frac{n}{(n-1)^2}\sigma^2=\frac{1}{n-1}\left(\frac{n}{n-1}\sigma^2\right)$$

この結果から、たとえば、$n=101$なら、$\hat{V}^{HC1}$は真の分散の 100 分の 1 となってしまうことがわかる。これは、本当は有意ではないとき、$\hat{V}^{HC1}$を用いることで、有意であると誤って判断する可能性が高いことを意味している。1 次的ダミーを用いる場合は、ロバスト標準誤差は、本当の標準誤差を過小評価する問題があることを念頭に置いておこう。

これまで$\sum_{i=1}^{n}D_i=n_1=1$としていた。$\hat{V}^{HC1}$は真の分散を過小評価する問題は、$n_1=1$とした場合だけでなく、n_1が小さい(1 をとるケースが少ない)、もしくは$n-n_1$が小さい(0 をとるケースが少ない)場合にも生じる。解決策として、$\hat{V}^{HC1}$の代わりに$\hat{V}^{HC2}$を用いることが挙げられる(ただし、$n_1=1$の場合には、h_{ii}のうち 1 つは 1 となり、標準化残差$\bar{u}_i=(1-h_{ii})^{-1/2}\hat{u}_i$、ひいては$\hat{V}^{HC2}$も計算できない)。

第 10 章の答え

練習問題 1

時系列データにおいて、誤差項が互いに相関していることを「系列相関がある」という。時系列データにおいては、系列相関がある状況が一般的であり、系列相関がない状況は特殊ケースとなる。

パネルデータには、時系列データの側面もあるため、系列相関が存在するのが一般的であり、系列相関がない状況は特殊ケースとなる。たとえば、県別パネルデータで、東京都のデータだけを考えると、これは時系列データとなる。

練習問題 2

OLS 推定量は不偏性と一致性を持つ。しかし、ガウス=マルコフの条件が満たされないため、OLS 推定量は有効ではない。また、OLS 推定量$\hat{\beta}$の分散は、

$$\sigma_{\hat{\beta}}^2 = \frac{T\left\{\gamma_0 + 2\sum_{s=1}^{T-1}\left(1-\frac{s}{T}\right)\gamma_s\right\}}{\{\sum_{t=1}^{T}(X_t-\bar{X})^2\}^2}$$

となる。ただし、パラメータγ_sは自己共分散であり、データにある系列相関の程度を表す。このため、通常の標準誤差ではなく、HAC 標準誤差を用いる。

練習問題 3

ここで$T=100$なら、$0.75\times 100^{1/3}=3.48$となり、バンド幅$m=3$が選択される。つまり、HAC 標準誤差は次のように計算できる。

$$\sqrt{100\times\frac{\left\{\widehat{\gamma_0}+2\sum_{s=1}^{2}\left(1-\frac{s}{3}\right)\widehat{\gamma_s}\right\}}{\{\sum_{t=1}^{T}(X_t-\bar{X})^2\}^2}} = \sqrt{100\times\frac{\left\{\widehat{\gamma_0}+2\left(\frac{2}{3}\widehat{\gamma_1}+\frac{1}{3}\widehat{\gamma_2}\right)\right\}}{\{\sum_{t=1}^{T}(X_t-\bar{X})^2\}^2}}$$

練習問題 4

ここで$T=500$なら、$0.75\times 500^{1/3}=5.95$となり、バンド幅$m=6$が選択される。つまり、HAC 標準誤差は次のように計算できる。

$$\sqrt{500\times\frac{\left\{\widehat{\gamma_0}+2\sum_{s=1}^{5}\left(1-\frac{s}{6}\right)\widehat{\gamma_s}\right\}}{\{\sum_{t=1}^{T}(X_t-\bar{X})^2\}^2}} = \sqrt{500\times\frac{\left\{\widehat{\gamma_0}+2\left(\frac{5}{6}\widehat{\gamma_1}+\frac{4}{6}\widehat{\gamma_2}+\frac{3}{6}\widehat{\gamma_3}+\frac{2}{6}\widehat{\gamma_4}+\frac{1}{6}\widehat{\gamma_5}\right)\right\}}{\{\sum_{t=1}^{T}(X_t-\bar{X})^2\}^2}}$$

練習問題 5

ここで$Y_t = \alpha + \beta X_t + u_t$から、$t-1$時点と$t-2$時点では、

$$Y_{t-1} = \alpha + \beta X_{t-1} + u_{t-1}$$

$$Y_{t-2} = \alpha + \beta X_{t-2} + u_{t-2}$$

が成立する。この関係式を使うと、

$$Y_t - \rho_1 Y_{t-1} - \rho_2 Y_{t-2} = (\alpha + \beta X_t + u_t) - \rho_1(\alpha + \beta X_{t-1} + u_{t-1}) - \rho_2(\alpha + \beta X_{t-2} + u_{t-2})$$

$$= \alpha(1 - \rho_1 - \rho_2) + \beta(X_t - \rho_1 X_{t-1} - \rho_2 X_{t-2}) + (u_t - \rho_1 u_{t-1} - \rho_2 u_{t-2})$$

となる。ここで、$u_t - \rho_1 u_{t-1} - \rho_2 u_{t-2} = \varepsilon_t$であるため、誤差項は標準的仮定を満たす。コクラン=オーカット法では、第 1 に、$Y_t = \alpha + \beta X_t + u_t$を OLS 推定し、得られた残差を用いて$\rho_1$と$\rho_2$を推定する。第 2 に、$\hat{\rho}_1$と$\hat{\rho}_2$を用いて、被説明変数を$Y_t - \hat{\rho}_1 Y_{t-1} - \hat{\rho}_2 Y_{t-2}$とし、説明変数を$1 - \hat{\rho}_1 - \hat{\rho}_2$、$X_t - \hat{\rho}_1 X_{t-1} - \hat{\rho}_2 X_{t-2}$とした定数項なしの OLS 推定をすれば、パラメータ(α, β)の推定ができる。

練習問題 6

練習問題 5 から、

$$Y_t - \rho_1 Y_{t-1} - \rho_2 Y_{t-2} = \alpha(1 - \rho_1 - \rho_2) + \beta(X_t - \rho_1 X_{t-1} - \rho_2 X_{t-2}) + (u_t - \rho_1 u_{t-1} - \rho_2 u_{t-2})$$

となるので、上式左辺の$\rho_1 Y_{t-1}$、$\rho_2 Y_{t-2}$を右辺に移項すると次式となる。

$$Y_t = \alpha(1 - \rho_1 - \rho_2) + \rho_1 Y_{t-1} + \rho_2 Y_{t-2} + \beta X_t - \rho_1 \beta X_{t-1} - \rho_2 \beta X_{t-2} + (u_t - \rho_1 u_{t-1} - \rho_2 u_{t-2})$$

ここで、

$$\alpha^* = \alpha(1 - \rho_1 - \rho_2)、$$

$$\beta_1 = \rho_1、\beta_2 = \rho_2、\beta_3 = \beta、\beta_4 = -\rho_1 \beta、\beta_5 = -\rho_2 \beta$$

$$\varepsilon_t = u_t - \rho_1 u_{t-1} - \rho_2 u_{t-2}$$

と定義すれば、上式は次のようになる。

$$Y_t = \alpha^* + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \beta_3 X_t + \beta_4 X_{t-1} + \beta_5 X_{t-2} + \varepsilon_t$$

このとき、誤差項はε_tであり、期待値 0、分散一定、自己共分散は 0 である。

練習問題 7

(a) 季節調整した系列Y_t'は、

$$Y'_t = \frac{1}{4}(Y_t + Y_{t-1} + Y_{t-2} + Y_{t-3})$$

となる。ここで、上式に、$Y_t = \alpha + \beta X_t + u_t$、$Y_{t-1} = \alpha + \beta X_{t-1} + u_{t-1}$、$Y_{t-2} = \alpha + \beta X_{t-2} + u_{t-2}$、$Y_{t-3} = \alpha + \beta X_{t-3} + u_{t-3}$を代入すると、次のようになる。

$$Y'_t = \frac{1}{4}\{(\alpha + \beta X_t + u_t) + (\alpha + \beta X_{t-1} + u_{t-1}) + (\alpha + \beta X_{t-2} + u_{t-2}) + (\alpha + \beta X_{t-3} + u_{t-3})\}$$

$$= \alpha + \beta \frac{1}{4}(X_t + X_{t-1} + X_{t-2} + X_{t-3}) + \frac{1}{4}(u_t + u_{t-1} + u_{t-2} + u_{t-3})$$

$$= \alpha + \beta X'_t + u'_t$$

式展開では、以下の関係式を用いた。

$$X'_t = \frac{1}{4}(X_t + X_{t-1} + X_{t-2} + X_{t-3})、u'_t = \frac{1}{4}(u_t + u_{t-1} + u_{t-2} + u_{t-3})$$

(b) 新しい誤差項u'_tの期待値は、以下で示すとおり、0 となる。

$$E[u'_t] = \frac{1}{4}(E[u_t] + E[u_{t-1}] + E[u_{t-2}] + E[u_{t-3}]) = 0$$

ここで、$E[u_t] = E[u_{t-1}] = E[u_{t-2}] = E[u_{t-3}] = 0$を用いた。

新しい誤差項u'_tの分散は、標準的仮定 3(誤差項u_tの分散はσ^2である)と標準的仮定 4(誤差項u_tが相互に無相関である)を用いると、次のようになる。

$$E[u'^2_t] = E\left[\left(\frac{1}{4}(u_t + u_{t-1} + u_{t-2} + u_{t-3})\right)^2\right] = \frac{1}{16}(E[u_t^2] + E[u_{t-1}^2] + E[u_{t-2}^2] + E[u_{t-3}^2])$$

$$= \frac{4\sigma^2}{16} = \frac{\sigma^2}{4}$$

つまり、新しい誤差項u'_tの分散は、元モデルの誤差項u_tの分散σ^2の 1/4 になる。

(c) 新しい誤差項u'_tとu'_{t-1}との自己共分散は、標準的仮定 3 と標準的仮定 4 を用いると、次のようになる。

$$E[u'_t u'_{t-1}] = E\left[\left(\frac{1}{4}(u_t + u_{t-1} + u_{t-2} + u_{t-3})\right)\left(\frac{1}{4}(u_{t-1} + u_{t-2} + u_{t-3} + u_{t-4})\right)\right]$$

$$\frac{1}{16}(E[u_{t-1}^2]+E[u_{t-2}^2]+E[u_{t-3}^2])=\frac{3\sigma^2}{16}$$

同様に、u_t'とu_{t-2}'との自己共分散は、

$$E[u_t'u_{t-2}']=E\left[\left(\frac{1}{4}(u_t+u_{t-1}+u_{t-2}+u_{t-3})\right)\left(\frac{1}{4}(u_{t-2}+u_{t-3}+u_{t-4}+u_{t-5})\right)\right]$$

$$=\frac{1}{16}(E[u_{t-2}^2]+E[u_{t-3}^2])=\frac{2\sigma^2}{16}$$

となり、u_t'とu_{t-3}'との自己共分散は、

$$E[u_t'u_{t-3}']=E\left[\left(\frac{1}{4}(u_t+u_{t-1}+u_{t-2}+u_{t-3})\right)\left(\frac{1}{4}(u_{t-3}+u_{t-4}+u_{t-5}+u_{t-6})\right)\right]$$

$$=\frac{1}{16}(E[u_{t-3}^2])=\frac{\sigma^2}{16}$$

となる。時差がさらに広がると、自己共分散はすべて 0 となる。

(d) 元のモデルでは、誤差項に系列相関が無かったにも関わらず、季節調整を行うことで、新しい誤差項には系列相関が生じていることに注意してほしい。これは季節調整という人為的なデータ調整によって生じた系列相関である。

練習問題 8

10.2 節で学習したとおり、系列相関があっても、OLS 推定量$\hat{\beta}$は不偏性を満たしている。このため、OLS 推定量$\hat{\beta}$の一致性を証明するには、サンプルサイズTが大きくなると推定量$\hat{\beta}$の分散が 0 に収束することを示せばよい(この点は 3.3.3 節を参照してほしい)。

系列相関があるとき、OLS 推定量$\hat{\beta}$の分散は次のようになる(10.3.1 節の(3)式を参照)。

$$E\left[\left(\hat{\beta}-\beta\right)^2\right]=T\frac{\left\{\gamma_0+2\sum_{s=1}^{T-1}\left(1-\frac{s}{T}\right)\gamma_s\right\}}{\{\sum_{t=1}^{T}(X_t-\bar{X})^2\}^2}$$

さらに分母と分子をT^2で割ると、推定量$\hat{\beta}$の分散は次式として表せる。

$$\frac{\frac{1}{T}\left\{\gamma_0+2\sum_{s=1}^{T-1}\left(1-\frac{s}{T}\right)\gamma_s\right\}}{\{\sum_{t=1}^{T}(X_t-\bar{X})^2/T\}^2}$$

ここで、$\hat{\beta}$の分散の分母は、X_tの標本分散（$\sum_{t=1}^{T}(X_t-\bar{X})^2/T$）の 2 乗であり、また、標本分散はサンプルサイズTが大きくなると真の分散に収束する[9]。このため、$\hat{\beta}$の分散の分母は有限の値となることがわかる。

次に、$\hat{\beta}$の分散の分子の{}内にある$\gamma_0+2\sum_{s=1}^{T-1}\left(1-\frac{s}{T}\right)\gamma_s$は、有限の値をとると仮定されている[10]。したがって、これをTで割った分子は、サンプルサイズTが大きくなると、0 に収束していくことになる。

以上から、サンプルサイズTが大きくなると、推定量$\hat{\beta}$の分散は 0 に収束し、推定量$\hat{\beta}$は一致性を満たすことがわかる。

練習問題 10

(8)式より、$u_t=\rho u_{t-1}+\varepsilon_t$となります。ただし、$-1<\rho<1$、また、$\varepsilon_t$は期待値 0($E[\varepsilon_t]=0$)、分散一定($E[\varepsilon_t^2]=\sigma^2$)、系列相関なし($E[\varepsilon_t\varepsilon_s]=0$)とします。

$t=1$期において、

$$u_1=\rho u_0+\varepsilon_1$$

となり、上式に$u_0=\rho u_{-1}+\varepsilon_0$を代入すると、

$$\begin{aligned}u_1&=\rho(\rho u_{-1}+\varepsilon_0)+\varepsilon_1\\&=\rho^2 u_{-1}+\varepsilon_1+\rho\varepsilon_0\end{aligned}$$

となり、さらに$u_{-1}=\rho u_{-2}+\varepsilon_{-1}$を代入すると、

$$\begin{aligned}u_1&=\rho^2(\rho u_{-2}+\varepsilon_{-1})+\varepsilon_1+\rho\varepsilon_0\\&=\rho^3 u_{-3}+\varepsilon_1+\rho\varepsilon_0+\rho^2\varepsilon_{-1}\end{aligned}$$

となります。こうした代入を無限回繰り返すと、u_1は次のように展開できます。

$$u_1=\rho^\infty u_{-\infty}+\varepsilon_1+\rho\varepsilon_0+\rho^2\varepsilon_{-1}+\rho^3\varepsilon_{-2}+\rho^4\varepsilon_{-3}+\cdots$$

ここで$-1<\rho<1$であることに注意すると、$\rho^\infty u_{-\infty}=0$となります。つまり、$u_1$は次のように表現できます。

$$u_1=\varepsilon_1+\rho\varepsilon_0+\rho^2\varepsilon_{-1}+\rho^3\varepsilon_{-2}+\rho^4\varepsilon_{-3}+\cdots$$

このとき、u_1の期待値は、

---

[9] 標本分散が真の分散に収束することは、たとえば、藪友良『入門　実践する統計学』(東洋経済新報社、2012 年)の 5.3.1 節の例 1 を参照されたい。

[10] 時差 s が大きくなると自己共分散は小さくなる傾向があるため(現在と遠い過去との関係は弱い)、サンプルサイズTが大きいとしても、$\gamma_0+2\sum_{s=1}^{T-1}\left(1-\frac{s}{T}\right)\gamma_s$は有限の値になるとする仮定は、あまり問題がないといえよう。

$$E[u_1]=E[\varepsilon_1]+\ \rho E[\varepsilon_0]+\rho^2 E[\varepsilon_{-1}]+\rho^3 E[\varepsilon_{-2}]+\rho^4 E[\varepsilon_{-3}]+\cdots=0$$

となり、分散は、次のように表せます。

$$V(u_1)=E[u_1^2]=\ E[\varepsilon_1^2]+\ \rho^2 E[\varepsilon_0^2]+\rho^4 E[\varepsilon_{-1}^2]+\rho^6 E[\varepsilon_{-2}^2]+\rho^8 E[\varepsilon_{-3}^2]+\cdots$$
$$=\sigma^2(1+\rho^2+\rho^4+\rho^6+\rho^8+\cdots)$$
$$=\frac{\sigma^2}{1-\rho^2}$$

式展開では、無限級数の和の公式を用いました。無限級数の和の公式は、$|\theta|<1$のとき、次式が成立するとしています[11]。

$$1+\theta+\theta^2+\theta^3+\cdots=\frac{1}{1-\theta}$$

つまり、$\theta=\rho^2$と定義すると、

$$1+\rho^2+\rho^4+\rho^6+\cdots=\frac{1}{1-\rho^2}$$

になるわけです。

[11]無限級数の和の公式を証明します。まず、Xを次のように定義します。

$$X=1+\theta+\theta^2+\theta^3+\cdots+\theta^s$$

上式両辺にθを掛けると次式となります。

$$\theta X=\theta+\theta^2+\theta^3+\cdots+\theta^{s+1}$$

このため、$X-\ \theta X$は、次のようになります。

$$X-\theta X=(1+\theta+\theta^2+\theta^3+\cdots+\theta^s)-(\theta+\theta^2+\theta^3+\cdots+\theta^{s+1})=1-\theta^{s+1}$$

さらに両辺を$(1-\theta)$で割ると、次式となります(上式左辺は$X-\theta X=(1-\theta)X$であることに注意すること)。

$$X=\frac{1-\theta^{s+1}}{1-\theta}$$

ここで、sが∞とすると、$|\theta|<1$から$\theta^{s+1}=0$となります。よって、$X=1/(1-\theta)$が証明できました。

第 11 章の答え

練習問題 1

パネルデータは、NはTより大きいことが一般的である。しかし、NよりTの方が大きいケースもある。パネルデータは一般的に年次データであるが、日次データや秒次データならTは大きい。この場合、TはNよりも大きくなる。

ミクロパネルとマクロパネル

パネルデータといったとき、大きくわけて 2 種類のデータがある。**ミクロパネル**(micro panel)は、個人、世帯、企業のパネルデータであり、Nは非常に大きく(1000 を上回ることが多い)、Tは小さい傾向がある(2～20 程度)。これに対し、**マクロパネル**(macro panel)は県、地域、国単位で集計されたデータであり、Nは小さく(7～100 程度)、Tは少し大きい傾向がある(10～50 程度)。11 章で分析したパネルデータは、マクロパネルに該当する。例 11-1 は$N = 64$、$T = 20$、例 11-3 は$N = 47$、$T = 11$、例 11-4 は$N = 47$、$T = 175$である。例 11-4 は日次であるため、NよりもTが大きくなっている。

練習問題 2

変量効果は、個別効果が説明変数と無相関である一方、固定効果は、個別効果が説明変数と相関している点で異なる。通常、個別効果と説明変数は相関しているため、変量効果は非現実的である。たとえば、被説明変数Yを賃金、Xを教育年数とすると、個別効果として、生まれつきの能力が挙げられる。能力と教育年数は相関していると考えるのが自然である。

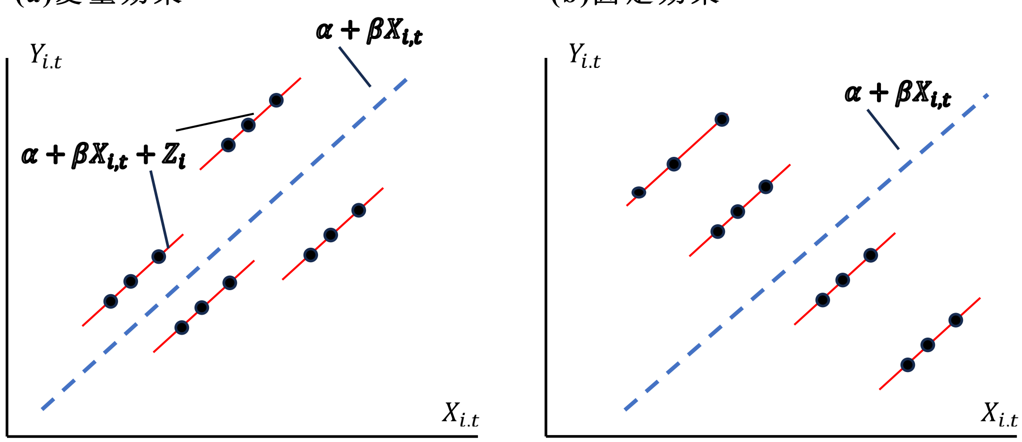
固定効果と変量効果の違い

固定効果と変量効果の違いを図で確認してみよう。モデルは$Y_{i,t} = \alpha + \beta X_{i,t} + Z_i + u_{i,t}$としよう。単純化のため、$\beta > 0$、$u_{i,t} = 0$とし、また、$X_{i,t}$と$Z_i$に負の相関があるとしよう。

下図の点線は$\alpha + \beta X_{i,t}$を、実線は$\alpha + \beta X_{i,t} + Z_i$、各点は観測値を表している。$Z_i$は確率変数であり、その値に応じて実線はシフトしている。下図(a)は

変量効果($X_{i,t}$とZ_iは無相関)の場合であり、$X_{i,t}$が大きいときZ_iが大きくなる(もしくは小さくなる)ような傾向はみられない。これに対し、下図(b)は固定効果(仮定から$X_{i,t}$とZ_iには負の無相関)の場合であり、$X_{i,t}$が大きくなるとZ_iがマイナスの値をとっている。この図から、図(a)の変量効果なら、プールドOLSでも一致性はある一方、図(b)の固定効果なら、一致性は満たされないことが理解できる。実際、下図(b)では、固定効果Z_iを考慮しないと、係数βは負の値として推定される。

(a)変量効果 **(b)固定効果**



練習問題 3

固定効果モデルは、

$$Y_{i,t}=\beta X_{i,t}+\alpha_1 D1_i+\alpha_2 D2_i+\cdots+\alpha_N DN_i+u_{i,t}$$

となる。ここで、帰無仮説と対立仮説を次のように設定してF検定をする。

$$H_0:\alpha_1=\alpha、\ \alpha_2=\alpha、\ \ldots、\ \alpha_N=\alpha$$

H_1: 帰無仮説H_0は誤りである

仮に帰無仮説H_0が採択されたなら、個別効果はないと判断される。これに対して、帰無仮説H_0が棄却されたなら、個別効果はあると判断される[1]。

なお、7.2節と同じような仮説にしたいのであれば、上式を

$$Y_{i,t}=\alpha_1+\beta X_{i,t}+(\alpha_2-\alpha_1)D2_i+\cdots+(\alpha_N-\alpha_1)DN_i+u_{i,t}$$

[1] 帰無仮説が棄却されても、プールドOLSに問題があるとはいえない。これは、個別効果が存在しても、説明変数と個別効果が無相関であれば、プールドOLSは一致性を持つためである。ただし、個別効果が存在するなら、個別効果が説明変数と相関する可能性を考慮し、固定効果モデルで推定することが望ましいといえる。

$$= \alpha_1 + \beta X_{i,t} + \theta_2 D2_i + \cdots + \theta_N DN_i + u_{i,t}$$

と表現し(ただし、$\theta_2 = \alpha_2 - \alpha_1$、…、$\theta_N = \alpha_N - \alpha_1$と定義した)[2]、仮説は次のよう設定すればよい。

H_0: $\theta_2 = 0$、$\theta_3 = 0$、…、$\theta_N = 0$

H_1: 帰無仮説H_0は誤りである

この仮説検定の結果から何が言えるだろうか。帰無仮説の棄却は、個別効果が存在し、プールド回帰が不適切であることを意味している。しかし、この検定からは、個別効果と説明変数が相関しているかは分からないため、変量効果モデルと個別効果モデルのどちらで推定すべきかわからない。

練習問題 4

2 時点のデータが利用可能であるため、両時点の差をとることで、

$$Y_{i,2} - Y_{i,1} = \left(\alpha + \beta X_{i,2} + \gamma Z_i + u_{i,2}\right) - \left(\alpha + \beta X_{i,1} + \gamma Z_i + u_{i,1}\right)$$
$$= \beta(X_{i,2} - X_{i,1}) + (u_{i,2} - u_{i,1})$$

となる。ここで、観察できない変数Z_iは式から消えるため、被説明変数を$Y_{i,2} - Y_{i,1}$、説明変数を$X_{i,2} - X_{i,1}$とすれば、通常の OLS で係数βをバイアスなく推定できる。

練習問題 5

固定効果モデルとして、次式を考えよう。

$$Y_{i,t} = \beta X_{i,t} + \theta W_i + \alpha_1 D1_i + \alpha_2 D2_i + \cdots + \alpha_N DN_i + u_{i,t}$$

ただし、W_iは時間を通じて一定の変数であるため、下添字はiだけとなる。また、θはその係数とする。ここで、W_iは、W_1、W_2、…、W_Nという値をとるとしよう(つまり、$i = 1$ならW_iはW_1となり、$i = 2$ならW_iはW_2となる)。

上式には、完全な多重共線性が生じるため、推定できないことを確認しよう。

---

[2] 次式が正しいことを確認しよう。

$$\beta X_{i,t} + \alpha_1 D1_i + \alpha_2 D2_i + \cdots + \alpha_N DN_i + u_{i,t} = \alpha_1 + \beta X_{i,t} + (\alpha_2 - \alpha_1) D2_i + \cdots + (\alpha_N - \alpha_1) DN_i + u_{i,t}$$

まず、$i = 1$なら、$D1_i = 1$、$D2_i = D3_i = \cdots = DN_i = 0$なので、$\beta X_{i,t} + \alpha_1 + u_{i,t} = \alpha_1 + \beta X_{i,t} + u_{i,t}$となり、両辺は等しい。次に、$i = 2$なら、$D2_i = 1$、$D1_i = D3_i = \cdots = DN_i = 0$なので、$\beta X_{i,t} + \alpha_2 + u_{i,t} = \alpha_1 + \beta X_{i,t} + (\alpha_2 - \alpha_1) + u_{i,t}$となり、やはり両辺は等しい。他の$i$についても両辺が等しくなることを確認してほしい。

まず、W_iを、ダミー変数($D1_i$、$D2_i$、 …、DN_i)を用いて、次のように表記する。

$$W_i = W_1 D1_i + W_2 D2_i + \cdots + W_N DN_i$$

次に、右辺の変数をすべて左辺に移項させると、

$$W_i - W_1 D1_i - W_2 D2_i - \cdots - W_N DN_i = 0$$

となる。ここで、W_1、W_2、…、W_Nは定数なので、$c_0 = 1$、$c_1 = -W_1$、$c_2 = -W_2$、…、$c_N = -W_N$とすると、

$$c_0 W_i + c_1 D1_i + c_2 D2_i + \cdots + c_N DN_i = 0$$

であり、完全な多重共線性が生じていることがわかる。

統計ソフトで計算される固定効果モデルの定数項は何か？

統計ソフトを用いて、固定効果モデルを推定すると、定数項が推定される。これはいったい何を表しているのだろうか。

ここで、固定効果モデルは、

$$Y_{i,t} = \alpha + \beta X_{i,t} + Z_i + u_{i,t}$$
$$= \alpha_i + \beta X_{i,t} + u_{i,t}$$

としており、個体iごとに異なる定数項がある(式展開では、$\alpha_i = \alpha + Z_i$と定義している)。直観的には、定数項はαと思われるかもしれないが、これは誤りである。

統計ソフトでは、定数項は次の 2 つの方法で計算される。第 1 の方法では、定数項を固定効果の平均としている(つまり、$\sum_{i=1}^{N} \hat{\alpha}_i / N$として計算される)。これは Stata の xtreg コマンドが該当している。

第 2 の方法では、いずれかの個体 i をベースに設定し、定数項をその固定効果とするものである。具体的には、個体 i=1 をベースにしてモデルを書き換えると、次のようになる。

$$Y_{i,t} = \alpha_1 + (\alpha_i - \alpha_1) + \beta X_{i,t} + u_{i,t}$$

ここで、定数項は個体 i=1 の固定効果α_1となる。通常、最初の個体 i=1 がベースとなるが、これは統計ソフトによって異なるかもしれない。

自殺者の国際比較(11 章例 11-1 参照)を例にしよう。まず、xtreg suicide unemployment, fe とすると、定数項は固定効果の平均であり、これは 11.05

と推定される。これに対し、reg suicide unemployment i.country とすると、定数項はベース国(この場合、albania)の固定効果であり、2.13 と推定されることになる(なお、国ダミーの係数は$\alpha_i - \alpha_1$を推定したものである)[3]。通常、ベース国に関心がないだろうから、固定効果の平均の方がまだ実証的な意味はあるだろう。ただし、そもそも固定効果の平均にも関心がない場合が多くあり、定数項を論文中で報告しないことも多い。

練習問題 6

時間固定効果モデルとして、次式を考えよう。

$$Y_{i,t} = \beta X_{i,t} + \theta W_t + \lambda_1 d1_t + \lambda_2 d2_t + \cdots + \lambda_T dT_t + u_{i,t}$$

ただし、W_tは個体iに対して同じ影響を与える変数であるため、下添字はtだけとなる。ここで、W_tは、W_1、W_2、…、W_Tという値をとるとしよう(つまり、$t=1$ならW_tはW_1となり、$t=2$ならW_tはW_2となる)。

上式には、完全な多重共線性が生じるため、推定できないことを確認しよう。まず、W_tを、ダミー変数($d1_t$、$d2_t$、…、dT_t)を用いて、次のように表記する。

$$W_t = W_1 d1_t + W_2 d2_t + \cdots + W_T dT_t$$

次に、右辺の変数をすべて左辺に移項させると

$$W_t - W_1 d1_t - W_2 d2_t - \cdots - W_T dT_t = 0$$

となる。ここで、W_1、W_2、…、W_Tは定数であるため、完全な多重共線性が生じていることがわかる。

練習問題 7

固定効果と時間効果を考慮したモデルは、次のモデルとなる。

$$Y_{i,t} = \beta X_{i,t} + \alpha_1 D1_i + \alpha_2 D2_i + \cdots + \alpha_N DN_i$$
$$+ \lambda_1 d1_t + \lambda_2 d2_t + \cdots + \lambda_T dT_t + u_{i,t}$$

問題 5 から、時間を通じて一定の変数W_iは、

$$W_i = W_1 D1_i + W_2 D2_i + \cdots + W_N DN_i$$

[3] 当然だが、reg コマンドを用いた推定でも、固定効果の平均は 11.05 になる。これを確認するため、各国の固定効果α_iを、国ダミーの係数+2.13(つまり、$\alpha_1 + (\alpha_i - \alpha_1)$)として計算する。そして、ベース国(albania)の固定効果+(国ダミーの係数+2.13)をすべて足してから国数 64 で割ると 11.05 となる。

となる。同様に、問題 6 から、個体iに対して同じ影響を与える変数W_tは、

$$W_t = W_1 d1_t + W_2 d2_t + \cdots + W_T dT_t$$

となる。これらに完全な多重共線性が生じていることは明らかであろう。

練習問題 8

どちらも同じモデルを考えている。

$$Y_{i,t} = \alpha + \beta X_{i,t} + e_{i,t}$$
$$e_{i,t} = Z_i + u_{i,t}$$

プールド OLS は、通常の OLS 推定である。これに対して、変量効果モデルによる推定は、誤差項$e_{i,t} = Z_i + u_{i,t}$の系列相関を除くため、被説明変数$Y_{i,t}$と説明変数$X_{i,t}$に複雑な変換を行ったうえで OLS 推定する(つまり、一般化最小 2 乗推定となる)。

変量効果の前提が正しいならば、変量効果推定量は BLUE となるが、プールド OLS は不偏性と一致性はあるが BLUE ではない。固定効果の前提が正しいならば、変量効果推定量とプールド OLS 推定量はどちらもバイアスを持った推定量である。

練習問題 12

(a) 時点tに関して、$Y_{i,t} = \alpha_i + \beta X_{i,t} + u_{i,t}$の和をとると、

$$\sum_{s=1}^{T} Y_{i,s} = \sum_{s=1}^{T} \left(\alpha_i + \beta X_{i,s} + u_{i,s}\right)$$

$$= T\alpha_i + \beta \sum_{s=1}^{T} X_{i,s} + \sum_{s=1}^{T} u_{i,s}$$

となる。さらに、上式の両辺をTで割ると、

$$\bar{Y}_i = \alpha_i + \beta \bar{X}_i + \bar{u}_i$$

が得られる。このため、$Y_{i,t}$から平均$\bar{Y}_i$を引くと、

$$Y_{i,t} - \bar{Y}_i = \left(\alpha_i + \beta X_{i,t} + u_{i,t}\right) - (\alpha_i + \beta \bar{X}_i + \bar{u}_i)$$
$$= \beta\left(X_{i,t} - \bar{X}_i\right) + (u_{i,t} - \bar{u}_i)$$

となり、個別効果が消えていることがわかる。したがって、被説明変数を$Y_{i,t}-\bar{Y}_i$、説明変数を$X_{i,t}-\bar{X}_i$とした OLS 推定によって係数βを推定できる[4]。ただし、定数項を 0 とした OLS 推定となることに注意してほしい(定数項が 0 とした OLS 推定は練習問題 3.12 を参照されたい)。

上式の OLS 推定量は、ダミー変数を用いた固定効果推定量と同じになる(この点は本稿の補足で厳密に証明しているので興味がある方は読んでほしい)。このため、上式の OLS 推定量は、やはり固定効果推定量と呼ばれる。

(b) 強外生性が必要な理由は、(a)のモデルを考えるとわかりやすい。

$$Y_{i,t}-\bar{Y}_i = \beta\left(X_{i,t}-\bar{X}_i\right)+\left(u_{i,t}-\bar{u}_i\right)$$

ここで、被説明変数は$Y_{i,t}-\bar{Y}_i$、説明変数は$X_{i,t}-\bar{X}_i$、誤差項は$u_{i,t}-\bar{u}_i$となる。このとき、外生性は、説明変数と誤差項が無相関である、つまり、

$$Cov\left(X_{i,t}-\bar{X}_i, u_{i,t}-\bar{u}_i\right)=0$$

を意味する。これは、$X_{i,t}$が**強外生性の仮定**(全時点(t、s)において、$X_{i,t}$は誤差項$u_{i,s}$と無相関となる)を満たすときに成立する[5]。

この点を理解するため、$T=2$という状況を考えよう。$t=1$のとき、

$$X_{i,1}-\bar{X}_i = X_{i,1}-\frac{1}{2}\left(X_{i,1}+X_{i,2}\right)=\frac{1}{2}\left(X_{i,1}-X_{i,2}\right)$$

となり、同様に、$u_{i,1}-\bar{u}_i=\frac{1}{2}\left(u_{i,1}-u_{i,2}\right)$となる。したがって、$Cov\left(X_{i,1}-\bar{X}_i、u_{i,1}-\bar{u}_i\right)=0$と$Cov\left(X_{i,1}-X_{i,2}, u_{i,1}-u_{i,2}\right)=0$は同じことである。そして、全ての$t$、$s$に対し$Cov\left(X_{i,t}, u_{i,s}\right)=0$のとき、$Cov\left(X_{i,1}-X_{i,2}, u_{i,1}-u_{i,2}\right)$が 0 となる[6]。

---

4 Stata で計算される within R^2 は、被説明変数を$Y_{i,t}-\bar{Y}_i$、説明変数を$X_{i,t}-\bar{X}_i$とした推定から得られる決定係数である。こちらは平均の影響を除いたうえで、個体内での$Y_{i,t}$の変動を$X_{i,t}$でどれだけ説明できるかをみている。実証分析では、within R^2 もしくはダミー変数を含めた固定効果推定から得られる決定係数のいずれかを用いる。なお、ダミー変数を含めた推定では決定係数が高くなり、within R^2 は低い値となる傾向がある。論文を執筆するときは、どちらの決定係数を用いたかを明示することが重要である。

5 説明変数が非確率変数であれば強外生性は成立するが、一般には、説明変数は確率変数ではなく、この仮定は成立しない可能性がある。

6 この点をクリアにするため、$Cov\left(X_{i,1}-X_{i,2}, u_{i,1}-u_{i,2}\right)$を展開してみよう。

Tが大きければ、通常の外生性の仮定で十分である。これはTが大きければ、$\bar{X}_i$は$E[X_i]$に、$\bar{u}_i$は$E[u_i]$に収束するためである。このとき、$Cov(X_{i,t}-\bar{X}_i, u_{i,t}-\bar{u}_i)$は$Cov(X_{i,t}-E[X_i], u_{i,t}-E[u_i])$で置き換えられ、

$$Cov(X_{i,t}-E[X_i], u_{i,t}-E[u_i]) = E[(X_{i,t}-E[X_i])(u_{i,t}-E[u_i])]$$

となる(これはまさに$Cov(X_{i,t}, u_{i,t})$となる)。残念ながら、パネル分析では、Nが大きい一方、Tが小さいことが一般的である(11.1.1 節参照)。つまり、$\bar{X}_i$は$E[X_i]$に、$\bar{u}_i$は$E[u_i]$に収束しない。このため、パネル分析では、強外生性の仮定が重要となる。たとえば、教科書の例 11-1 は$T=20$、例 11-3 は$T=11$、例 11-4 は$T=175$である。例 11-4 ではTが大きいが、例 11-3 ではTが大きいとはいえない。

強外生性は強い仮定であり、これが満たされない状況は多い。そうした状況として、①説明変数の 1 つに被説明変数のラグ$Y_{i,t-1}$があるケース、②YからXへのフィードバックがあるケース、が挙げられる。

まず、①説明変数の 1 つに被説明変数のラグ$Y_{i,t-1}$があるケースである。このとき、モデルは

$$Y_{i,t} = \alpha_i + \beta_1 Y_{i,t-1} + \beta_2 X_{i,t} + u_{i,t}$$

となり、説明変数$Y_{i,t-1}$は、1 時点前の$u_{i,t-1}$と相関している(1 期前は$Y_{i,t-1} = \alpha_i + \beta_1 Y_{i,t-2} + \beta_2 X_{i,t-1} + u_{i,t-1}$から、$Y_{i,t-1}$と$u_{i,t-1}$は相関する)。このように、被説明変数のラグを入れるときは、強外生性の仮定がみたされない[7]。

次に、②YからXへのフィードバックがあるケースである。ここで、$Y_{i,t}$が変化すると、しばらく経過してから将来の$X_{i,s}$が変化するとしよう。このとき、$u_{i,t}$が$Y_{i,t}$に、ひいては$X_{i,s}$に影響を与えるため、$u_{i,t}$と$X_{i,s}$が相関し、強外生性の仮定が満たされない。たとえば、例 11-4 では、Xは緊急事態ダミー、Yは自粛率

$$\begin{aligned}
&E\left[\left(X_{i,1}-X_{i,2}-E[X_{i,1}-X_{i,2}]\right)\left(u_{i,1}-u_{i,2}-E[u_{i,1}-u_{i,2}]\right)\right] \\
&\qquad = E\left[\left(\left(X_{i,1}-E[X_{i,1}]\right)-\left(X_{i,2}-E[X_{i,2}]\right)\right)\left(\left(u_{i,1}-E[u_{i,1}]\right)-\left(u_{i,2}-E[u_{i,2}]\right)\right)\right] \\
&\qquad = E\left[\left(X_{i,1}-E[X_{i,1}]\right)\left(u_{i,1}-E[u_{i,1}]\right)\right]+E\left[\left(X_{i,1}-E[X_{i,1}]\right)\left(u_{i,2}-E[u_{i,2}]\right)\right] \\
&\qquad\quad + E\left[\left(X_{i,2}-E[X_{i,2}]\right)\left(u_{i,1}-E[u_{i,1}]\right)\right]+E\left[\left(X_{i,2}-E[X_{i,2}]\right)\left(u_{i,2}-E[u_{i,2}]\right)\right]
\end{aligned}$$

これが 0 となるのは、すべてのsやtに対し、$E[(X_{i,t}-E[X_{i,t}])(u_{i,s}-E[u_{i,s}])]=0$のときである。

7 こうしたモデルは、動学パネルデータモデルと呼ばれる。このとき、操作変数を用いた推定が提案されている。

とした。自粛率が下がってくると、緊急事態宣言が発令されやすくなる可能性がある。このため、強外生性の仮定は満たされないだろう(なお、この例では$T=175$と大きく、外生性の仮定だけでも十分と考えられる)。

パネルデータは、個別要因を考慮し、内生性の問題を軽減する有用な手法ではあるが、Tが小さいケースでは、強外生性の仮定が必要となることを覚えておいてほしい。

将来的に追加する予定の練習問題(興味があれば解いてください)

13. ★時間固定効果モデル$Y_{i,t}=\lambda_t+\beta X_{i,t}+u_{i,t}$を考える。このとき、$Y_{i,t}-\bar{Y}_t=\beta\left(X_{i,t}-\bar{X}_t\right)+(u_{i,t}-\bar{u}_t)$を示せ。$\bar{Y}_t$、$\bar{X}_t$、$\bar{u}_t$は時点$t$での平均となる($\bar{Y}_t=\frac{1}{N}\sum_{i=1}^{N}Y_{i,t}$、$\bar{X}_t=\frac{1}{N}\sum_{i=1}^{N}X_{i,t}$、$\bar{u}_t=\frac{1}{N}\sum_{i=1}^{N}u_{i,t}$)。これは被説明変数を$Y_{i,t}-\bar{Y}_t$、説明変数を$X_{i,t}-\bar{X}_t$としたOLS推定によって係数$\beta$を推定できることを意味する。

14. ★ 固定効果と時間効果を考慮したモデル$Y_{i,t}=\alpha_i+\lambda_t+\beta X_{i,t}+u_{i,t}$を考える。$\bar{Y}$、$\bar{X}$、$\bar{u}$は全てのデータを用いた平均となる($\bar{Y}=\frac{1}{NT}\sum_{s=1}^{T}\sum_{i=1}^{N}Y_{i,s}$、$\bar{X}=\frac{1}{NT}\sum_{s=1}^{T}\sum_{i=1}^{N}X_{i,s}$、$\bar{u}=\frac{1}{NT}\sum_{s=1}^{T}\sum_{i=1}^{N}u_{i,s}$)。このとき、$Y_{i,t}-\bar{Y}_i-\bar{Y}_t+\bar{Y}=\beta\left(X_{i,t}-\bar{X}_i-\bar{X}_t+\bar{X}\right)+(u_{i,t}-\bar{u}_i-\bar{u}_t+\bar{u})$を示せ。これは被説明変数を$Y_{i,t}-\bar{Y}_i-\bar{Y}_t+\bar{Y}$、説明変数を$X_{i,t}-\bar{X}_i-\bar{X}_t+\bar{X}$としたOLS推定によって係数$\beta$を推定できることを意味する。

練習問題13の答え

時点tでの総和は次のようになる。

$$\sum_{i=1}^{N}Y_{i,t}=\sum_{i=1}^{N}\left(\lambda_t+\beta X_{i,t}+u_{i,t}\right)$$

$$=N\lambda_t+\beta\sum_{i=1}^{N}X_{i,t}+\sum_{i=1}^{N}u_{i,t}$$

上式の両辺をNで割ると、

$$\bar{Y}_t=\lambda_t+\beta\bar{X}_t+\bar{u}_t$$

が得られる。このため、$Y_{i,t}$から時点tの平均$\bar{Y}_t$を引くと、

$$Y_{i,t}-\bar{Y}_t=\left(\lambda_t+\beta X_{i,t}+u_{i,t}\right)-(\lambda_t+\beta\bar{X}_t+\bar{u}_t)$$

$$=\beta\left(X_{i,t}-\bar{X}_t\right)+(u_{i,t}-\bar{u}_t)$$

となり、時間効果が消えている。したがって、被説明変数を$Y_{i,t}-\bar{Y}_t$、説明変数を$X_{i,t}-\bar{X}_t$とした OLS 推定によって係数βを推定できる。

<u>練習問題 14 の答え</u>

ここで$\bar{Y}_i$と$\bar{Y}_t$を計算してみよう。

$$\bar{Y}_i=\frac{1}{T}\sum_{s=1}^{T}Y_{i,s}=\frac{1}{T}\sum_{s=1}^{T}\left(\alpha_i+\lambda_s+\beta X_{i,s}+u_{i,s}\right)=\alpha_i+\bar{\lambda}+\beta\bar{X}_i+\bar{u}_i$$

$$\bar{Y}_t=\frac{1}{N}\sum_{i=1}^{N}Y_{i,t}=\frac{1}{N}\sum_{i=1}^{N}\left(\alpha_i+\lambda_t+\beta X_{i,t}+u_{i,t}\right)=\bar{\alpha}+\lambda_t+\beta\bar{X}_t+\bar{u}_t$$

ただし、$\bar{\alpha}=\frac{1}{N}\sum_{i=1}^{N}\alpha_i$、$\bar{\lambda}=\frac{1}{T}\sum_{s=1}^{T}\lambda_s$と定義される。同様に、全データを用いた平均は次のようになる。

$$\bar{Y}=\frac{1}{NT}\sum_{s=1}^{T}\sum_{i=1}^{N}Y_{i,s}=\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N}Y_{i,s}\right)=\frac{1}{T}\sum_{s=1}^{T}\bar{Y}_s$$
$$=\frac{1}{T}\sum_{s=1}^{T}(\bar{\alpha}+\lambda_s+\beta\bar{X}_s+\bar{u}_s)$$
$$=\bar{\alpha}+\bar{\lambda}+\beta\bar{X}+\bar{u}$$

以上の結果から、$Y_{i,t}-\bar{Y}_i-\bar{Y}_t+\bar{Y}$とした変形をすると、次のようになる。

$$Y_{i,t}-\bar{Y}_i-\bar{Y}_t+\bar{Y}=\left(\alpha_i+\lambda_t+\beta X_{i,t}+u_{i,t}\right)-\left(\alpha_i+\bar{\lambda}+\beta\bar{X}_i+\bar{u}_i\right)-(\bar{\alpha}+\lambda_t+\beta\bar{X}_t+\bar{u}_t)$$
$$+(\bar{\alpha}+\bar{\lambda}+\beta\bar{X}+\bar{u})$$
$$=\beta\left(X_{i,t}-\bar{X}_i-\bar{X}_t+\bar{X}\right)+(u_{i,t}-\bar{u}_i-\bar{u}_t+\bar{u})$$

ここで、固定効果と時間効果が消えているのがわかる。したがって、被説明変数を$Y_{i,t}-\bar{Y}_i-\bar{Y}_t+\bar{Y}$、説明変数を$Y_{i,t}-\bar{Y}_i-\bar{Y}_t+\bar{Y}$とした OLS 推定によって係数$\beta$を推定できる。

補足：固定効果モデルの 2 つの推定方法

教科書 11.3 節で学習したとおり、固定効果モデルは、N個のダミー変数$(D1_i,D2_i,\dots,DN_i)$を含めることで、

$$Y_{i,t}=\beta X_{i,t}+\alpha_1 D1_i+\alpha_2 D2_i+\cdots+\alpha_N DN_i+u_{i,t}$$

となり、この式を推定することで係数βの固定効果推定量が得られる。ここで、$D1_i$は個体が$i=1$の場合に 1 となるダミー変数となる。

固定効果推定量は、次のモデルの OLS 推定でも得られる。

$$Y_{i,t}-\bar{Y}_i=\beta\left(X_{i,t}-\bar{X}_i\right)+(u_{i,t}-\bar{u}_i)$$

ただし、$\bar{Y}_i$、$\bar{X}_i$、$\bar{u}_i$はそれぞれの時間平均であり、次のように定義される。

$$\bar{Y}_i=\frac{1}{T}\sum_{s=1}^{T}Y_{i,s}\text{、}\bar{X}_i=\frac{1}{T}\sum_{s=1}^{T}X_{i,s}\text{、}\bar{u}_i=\frac{1}{T}\sum_{s=1}^{T}u_{i,s}$$

[証明]

この点が正しいことを確認しよう。FWL 定理より、ダミー変数を用いた係数βの推定は、次の 3 段階にわけて行うことができる[8]。

① $X_{i,t}$を被説明変数とし、$X_{i,t}$以外の説明変数($D1_i$、$D2_i$、…、DN_i)で OLS 推定し、残差$\hat{u}_{Xi,t}$を求める。

② $Y_{i,t}$を被説明変数とし、$X_{i,t}$以外の説明変数($D1_i$、$D2_i$、…、DN_i)で OLS 推定し、残差$\hat{u}_{Yi,t}$を求める。

③被説明変数を$\hat{u}_{Yi,t}$とし、説明変数を$\hat{u}_{Xi,t}$として OLS 推定し、係数$\hat{\beta}$を求める。

単純化のため、$N=2$という状況を考える。①では、モデルは$X_{i,t}=\gamma_1 D1_i+\gamma_2 D2_i+u_{Xi,t}$であり、パラメータ$\gamma_1$、$\gamma_2$は、次の残差 2 乗和を最小にするように決定される。

$$\underbrace{\sum_{t=1}^{T}\left(X_{1,t}-\gamma_1 D1_i-\gamma_2 D2_i\right)^2}_{\text{個体}i=1\text{の残差 2 乗和}}+\underbrace{\sum_{t=1}^{T}\left(X_{2,t}-\gamma_1 D1_i-\gamma_2 D2_i\right)^2}_{\text{個体}i=2\text{の残差 2 乗和}}=\sum_{t=1}^{T}\left(X_{1,t}-\gamma_1\right)^2+\sum_{t=1}^{T}\left(X_{2,t}-\gamma_2\right)^2$$

式展開では、$D1_i$は個体が$i=1$の場合に 1 となるダミー変数、同様に、$D2_i$は個体が$i=2$の場合に 1 となるダミー変数となることを用いた。ここで、$\sum_{t=1}^{T}\left(X_{1,t}-\gamma_1\right)^2$が最小化されるのは、$\hat{\gamma}_1=\bar{X}_1=\frac{1}{T}\sum_{s=1}^{T}X_{1,s}$のときである(2 章の練習問題 7 参照)。同様に、$\hat{\gamma}_2=\bar{X}_2=\frac{1}{T}\sum_{s=1}^{T}X_{2,s}$となる。以上から、残差は次のように表せる。

$$\hat{u}_{Xi,t}=X_{i,t}-\hat{\gamma}_1 D1_i-\hat{\gamma}_2 D2_i=X_{i,t}-\bar{X}_i$$

式展開では、$i=1$なら$D1_i=1$、$\hat{\gamma}_1=\bar{X}_1$、$D2_i=0$となり、$i=2$なら$D2_i=1$、$\hat{\gamma}_2=\bar{X}_2$、$D1_i=0$となることを用いた。

[8] FWL 定理は、サポートウェブサイトの資料『FWL 定理と重回帰分析』を参照されたい。

同様の議論によって、②では$\hat{u}_{Yi,t} = Y_{i,t} - \bar{Y}_i$を示すことができる。③では、被説明変数を$\hat{u}_{Yi,t} = Y_{i,t} - \bar{Y}_i$とし、説明変数を$\hat{u}_{Xi,t} = X_{i,t} - \bar{X}_i$として係数$\beta$を推定すればよい。これはモデル

$$Y_{i,t} - \bar{Y}_i = \beta\left(X_{i,t} - \bar{X}_i\right) + (u_{i,t} - \bar{u}_i)$$

を OLS 推定することを意味している。以上から、2 つのモデル

$$Y_{i,t} = \beta X_{i,t} + \alpha_1 D1_i + \alpha_2 D2_i + \cdots + \alpha_N DN_i + u_{i,t}$$

$$Y_{i,t} - \bar{Y}_i = \beta\left(X_{i,t} - \bar{X}_i\right) + (u_{i,t} - \bar{u}_i)$$

のどちらを推定しても、係数βは同じになることが確認された。

第 12 章の答え

練習問題 1

(a) 限界効果は、年齢の係数 0.032 である。つまり、1 年歳をとると約 3%だけ結婚確率は上がる。

(b) 結婚確率は、30 歳と 45 歳において、それぞれ次のようになる。

$$\hat{P}\{Y_i = 1\} = -0.5 + 0.032 \times 30 = 0.46$$

$$\hat{P}\{Y_i = 1\} = -0.5 + 0.032 \times 45 = 0.94$$

(c) 年齢X_iが十分に大きい、もしくは十分に小さいと、確率$P\{Y_i = 1\}$は 1 より大きい、もしくは 0 より小さくなる。たとえば、年齢が 50 歳なら、次のように確率は 100%を超える。

$$P\{\widehat{Y_i = 1}\} = -0.5 + 0.032 \times 50 = 1.1$$

練習問題 2

最終学歴が大卒なら、$X_{2i} = 1$、$X_{3i} = 0$となる。よって、大卒の 30 歳と 45 歳の結婚確率は、それぞれ次のようになる。

$$\hat{P}\{Y_i = 1\} = F(-3.5 + 0.1 \times 30 + 0.3 \times 1 + 0.6 \times 0) = 0.42$$

$$\hat{P}\{Y_i = 1\} = F(-3.5 + 0.1 \times 45 + 0.3 \times 1 + 0.6 \times 0) = 0.90$$

Excel では、「=NORM.S.DIST(-3.5+0.1*30+0.3*1+0.6*0,TRUE）」と入力すれば、確率を 0.42 と計算できる。

練習問題 3

モデルのパラメータ(α^*、β^*、c、σ^2)を用いると、次のようになる(導出は 12 章補足参照)。

$$\alpha = \frac{\alpha^* - c}{\sigma} \qquad \beta = \frac{\beta^*}{\sigma}$$

練習問題 4

$\hat{\alpha}_{ML}$、$\hat{\beta}_{ML}$は最尤推定量としよう。このとき、どのような$\tilde{\alpha}$、$\tilde{\beta}$に対しても、

$$P(\tilde{\alpha}、\tilde{\beta}) \leq P(\hat{\alpha}_{ML}、\hat{\beta}_{ML})$$

が成立する。ただし、$P(\tilde{\alpha}、\tilde{\beta})$は、パラメータが$\tilde{\alpha}$、$\tilde{\beta}$としたときの尤度(同時確

率)とする。対数は単調増加する関数であるため[9]、どのような$\tilde{\alpha}$、$\tilde{\beta}$に対しても、次式が成立する。

$$\ln(P(\tilde{\alpha}、\tilde{\beta})) \leq \ln(P(\hat{\alpha}_{ML}、\hat{\beta}_{ML}))$$

以上から、尤度を最大にする最尤推定量($\hat{\alpha}_{ML}$、$\hat{\beta}_{ML}$)は、対数尤度を最大にする推定量となる。

練習問題 5

説明変数を増やすと当てはまりは改善するため、次の式が成立する。

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

上式の対数をとると、次の関係式が得られる(1 の対数は 0 となることに注意)。

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

さらに、両辺を$\ln(P_0^{max})$で割ると、次の関係式が得らえる($\ln(P_0^{max})$は負の値であるため、不等号は逆になることに注意)。

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

この関係式から、疑似決定係数は 0 以上 1 以下となる。

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

練習問題 6

$Y_i = y_i$となるときの確率密度関数$f\{Y_i = y_i\}$は、次のようになる。

$$f\{Y_i = y_i\} = f\{\alpha + \beta X_i + u_i = y_i\}$$
$$= f\{u_i = y_i - \alpha - \beta X_i\}$$
$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i - \alpha - \beta X_i)^2}{2\sigma^2}}$$
$$= (2\pi)^{-\frac{1}{2}} (\sigma^2)^{-\frac{1}{2}} e^{-\frac{(y_i - \alpha - \beta X_i)^2}{2\sigma^2}}$$

誤差項u_iは互いに独立であるから、Y_iも互いに独立となる。よって、尤度は

$$L = f\{Y_1 = y_1, \ldots, Y_n = y_n\} = f\{Y_1 = y_1\} \times \ldots \times f\{Y_n = y_n\}$$

[9] 対数は単調増加関数であるため、$P_1 \leq P_2$であれば$\ln(P_1) \leq \ln(P_2)$となり、逆に、$\ln(P_1) \leq \ln(P_2)$であれば$P_1 \leq P_2$となる。

$$= (2\pi)^{-\frac{1}{2}}(\sigma^2)^{-\frac{1}{2}}e^{-\frac{(y_1-\alpha-\beta X_1)^2}{2\sigma^2}} \times \dots\dots \times (2\pi)^{-\frac{1}{2}}(\sigma^2)^{-\frac{1}{2}}e^{-\frac{(y_n-\alpha-\beta X_n)^2}{2\sigma^2}}$$

$$= (2\pi)^{-\frac{n}{2}}(\sigma^2)^{-\frac{n}{2}}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\alpha-\beta X_i)^2}$$

となる。尤度関数は積の形になっており、最大化問題を解くのが困難である。このため、対数尤度関数を用いて、最大化問題を解く。

尤度の両辺について対数をとると、対数尤度が得られる。

$$ln\, f\{Y_1 = y_1, \dots, Y_n = y_n\} = -\frac{n}{2}ln(2\pi) - \frac{n}{2}ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \alpha - \beta X_i)^2$$

ここで、αとβは、右辺第 3 項だけにあるので、(第 3 項に－1 が掛けられていることに注意すると)対数尤度関数の最大化は、残差 2 乗和$\sum_{i=1}^{n}(y_i - \alpha - \beta X_i)^2$の最小化と同じである。つまり、残差 2 乗和を最小化する OLS 推定量が、対数尤度関数を最大化する最尤推定量となる。従って最尤推定量は、次式となる。

$$\hat{\alpha}_{ML} = \hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$$

$$\hat{\beta}_{ML} = \hat{\beta} = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

次に、σ^2の最尤推定量を求めよう。これは対数尤度を最大にするσ^2であるため、対数尤度をσ^2で偏微分してから 0 と置くことで求める。偏微分して 0 とおくと、次の式となる[10]。

$$\frac{\partial ln\, f\{Y_1 = y_1, \dots, Y_n = y_n\}}{\partial \sigma^2} = -\frac{n}{2\hat{\sigma}_{ML}^2} + \frac{1}{2\hat{\sigma}_{ML}^4}\sum_{i=1}^{n}(y_i - \hat{\alpha}_{ML} - \hat{\beta}_{ML} X_i)^2 = 0$$

この式の解は最尤推定量のため、下添字に ML を付けた。上式の両辺に$2\hat{\sigma}_{ML}^2$を掛けてから整理すると、最尤推定量が得られる[11]。

---

[10] 計算は少し難しいため、丁寧に記述しよう。σ^2で偏微分すると、

$$\frac{\partial ln\, f\{Y_1 = y_1, \dots, Y_n = y_n\}}{\partial \sigma^2} = -\frac{n}{2}\left\{\frac{\partial\, ln(\sigma^2)}{\partial \sigma^2}\right\} - \frac{1}{2}\left\{\frac{\partial 1/\sigma^2}{\partial \sigma^2}\right\}\sum_{i=1}^{n}(y_i - \alpha - \beta X_i)^2$$

となる。ここで、対数の微分の公式から、

$$\frac{\partial\, ln(\sigma^2)}{\partial \sigma^2} = \frac{1}{\sigma^2}$$

となる(サポートウェブサイトの補足資料「ネイピア数と自然対数の微分」を参照)。また、$1/\sigma^2 = (\sigma^2)^{-1}$であるから、

$$\frac{\partial(1/\sigma^2)}{\partial \sigma^2} = \frac{\partial(\sigma^2)^{-1}}{\partial \sigma^2} = -(\sigma^2)^{-2} = -\frac{1}{\sigma^4}$$

となる。以上から、先の式は次のとおりとなる。

$$\frac{\partial ln\, f\{Y_1 = y_1, \dots, Y_n = y_n\}}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(y_i - \alpha - \beta X_i)^2$$

[11] なお、最尤推定量$\hat{\sigma}_{ML}^2$は、残差 2 乗和をnで割った値である。3 章補足で確認したとおり、σ^2の不偏推定量は、残差 2 乗和を$n-2$で割った値であるため、最尤推定量$\hat{\sigma}_{ML}^2$は不偏推定量ではない。サンプルサイズが大きいと、nで割っても、$n-2$で割ってもほぼ同じであるため、最尤推定量$\hat{\sigma}_{ML}^2$は一致

$$\hat{\sigma}_{ML}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{\alpha}_{ML} - \hat{\beta}_{ML} X_i)^2$$

練習問題 7

補足証明で確認したとおり、

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

であるから、$\beta = 0$とすると、

$$P\{Y_i = 1\} = P\{u_i < \alpha\}$$

となる。ここで、$F(\alpha)$は定数であるため、$F(\alpha) = p$と表記しよう。確率の合計は 1 であるため、Y_iが 1 となる確率はpであれば、0 となる確率は$1 - p$となる。まとめると、次のようになる。

$$P\{Y_i = y_i\} = p^{y_i}(1-p)^{1-y_i}$$

このとき、尤度は、

$$P\{Y_1 = y_1, Y_2 = y_2, \ldots, Y_n = y_n\} = P\{Y_1 = y_1\}P\{Y_2 = y_2\} \ldots P\{Y_n = y_n\}$$
$$= p^{y_1}(1-p)^{1-y_1} p^{y_2}(1-p)^{1-y_2} \ldots p^{y_n}(1-p)^{1-y_n}$$
$$= p^{\sum y_i}(1-p)^{n-\sum y_i}$$

となる。これをpに関して微分して 0 と置くことで、

$$\sum y_i\, \hat{p}_{ML}^{\sum y_i - 1}(1-\hat{p}_{ML})^{n-\sum y_i} - (n - \sum y_i)\, \hat{p}_{ML}^{\sum y_i}(1-\hat{p}_{ML})^{n-\sum y_i - 1} = 0$$

となる。さらに両辺を$\hat{p}_{ML}^{\sum y_i - 1}(1-\hat{p}_{ML})^{n-\sum y_i - 1}$で割ると、

$$\sum y_i\,(1-\hat{p}_{ML}) - (n - \sum y_i)\hat{p}_{ML} = 0$$

となる。これを整理すると、

$$\sum y_i - n\hat{p}_{ML} = 0$$

となるから、$\hat{p}_{ML}$について解くと、

$$\hat{p}_{ML} = \frac{\sum y_i}{n} = \bar{y}$$

となる。したがって、尤度関数に$\hat{p}_{ML} = \bar{y}$を代入することで、最大尤度P_0^{max}が次のようなる。

$$P_0^{max} = P\{Y_1 = y_1, Y_2 = y_2, \ldots, Y_n = y_n\} = \bar{y}^{\sum y_i}(1-\bar{y})^{n-\sum y_i}$$

推定量となる。まとめると、最尤推定量は一致推定量だが、不偏推定量ではないといえる。

練習問題 8

(a) 投票率P_iは質的データではなく、連続的な値をとる量的データである。しかし、投票率は割合を表しているため、0 以上 1 以下となる必要がある。線形モデルを考えると、投票率P_iの予測値は 0 を下回る、また、1 を上回る可能性がある。これが線形モデルの問題である。

(b) 投票率P_iは 0 から 1 までの値をとるため、$P_i/(1-P_i)$は 0 から＋∞までの値をとる。さらに対数をとると、$\ln(P_i/(1-P_i))$は－∞から＋∞の値をとる。したがって、ロジット変換した推定式、つまり、

$$\ln\left(\frac{P_i}{1-P_i}\right) = \alpha + \beta X_i + u_i$$

では､投票率P_iが 0 以上 1 以下という制約を考慮したうえで推定がされている。

この点を確認してみよう。まず、上式は次式を意味している(下式の対数をとると上式になることから明らかだろう)。

$$\frac{P_i}{1-P_i} = e^{\alpha+\beta X_i+u_i}$$

これをP_iについて解くと、

$$P_i = \frac{e^{\alpha+\beta X_i+u_i}}{1+e^{\alpha+\beta X_i+u_i}}$$

となり、さらに分母と分子に$e^{-(\alpha+\beta X_i+u_i)}$を掛けると、

$$P_i = \frac{1}{1+e^{-(\alpha+\beta X_i+u_i)}}$$

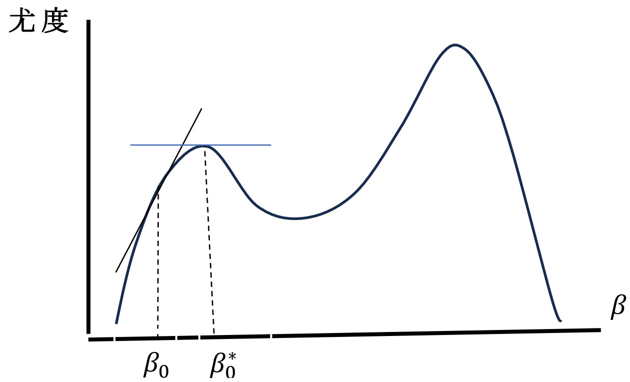
となる。ここで、$e^{-(\alpha+\beta X_i+u_i)}$は 0 以上である。$e^{-(\alpha+\beta X_i+u_i)}$が 0 なら$P_i = 1$となり、$e^{-(\alpha+\beta X_i+u_i)}$が大きくなると$P_i$は 0 に近づく。

(c) $P_i = 0$もしくは 1 になると、$\ln(P_i/(1-P_i))$は定義できない。ただし、実際の投票率が 0%や 100%という状況はありえないため、問題とはならない。

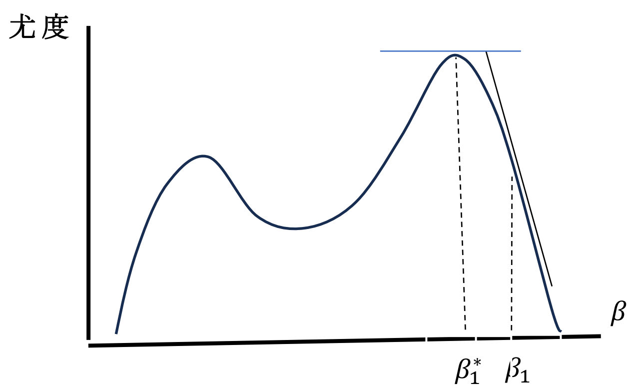
練習問題 11

初期値をβ_0とし、尤度関数をβで微分すると、傾きは正となっており、βをβ_0から増加させることで尤度が増加する(下図参照)。これを続けていくと、β_0^*で微分の傾きは 0 となり、ここでストップすることになる。しかし、この図から、明

らかなように、β_0^*は尤度関数を最大化していない。これは**局所的最大化**(locally maximum)といわれる問題である。



これは様々な初期値を試すことで解決できる。たとえば、初期値をβ_1とし、尤度関数を微分すると、傾きは負となっており、βをβ_1から減少させることで尤度は増加する(下図参照)。これを続けると、β_1^*で微分の傾きは 0 となり、ここでストップすることになる。β_0^*の尤度と、β_1^*の尤度を比較し、より大きい方の尤度が選ばれることになる。これすれば、尤度を最大にするβ_1^*を求めることができる。ここでは初期値を 2 つしか試していないが、実際の分析では、初期値としていろいろな組み合わせを試すことが必要である。なお、プロビットやロジットでは、尤度のピークは 1 つしかないので、こうした問題は生じないことが知られている。



第 13 章の答え

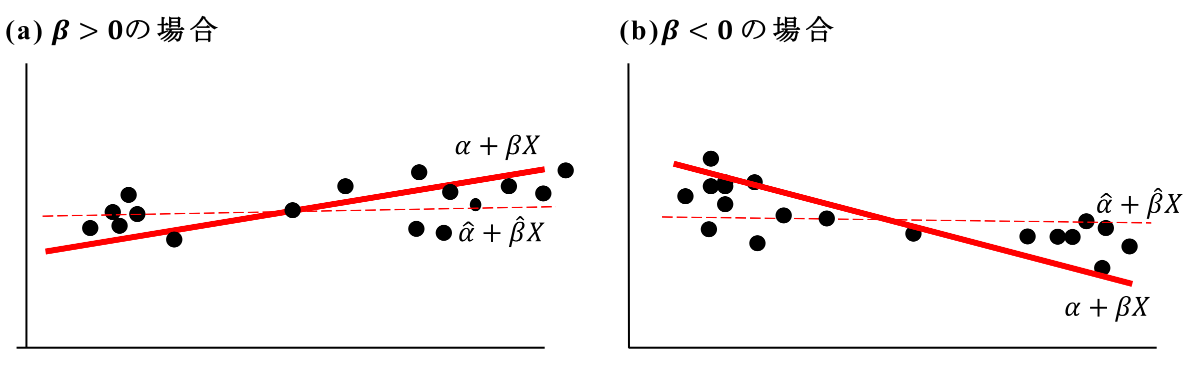
練習問題 1

$\beta > 0$としよう。下図(a)では、真の関係$(\alpha + \beta X)$を実線、回帰直線(当てはまりの良い線)を点線で表した。このとき、測定誤差e_iが大きくなると、説明変数X_iは大きくなる一方、誤差項$u_i = u_i^* - \beta e_i$は小さくなる。つまり、X_iとu_iに負の相関が生じる。X_iとu_iに負の相関があるため、X_iが大きいと、u_iは小さく(負の値をとる)なり、観測点は実線の下で観察される。逆に、X_iが小さいと、u_iは大きく(正の値をとる)なり、観測点は実線の上で観察される。このため、回帰直線は傾きがゆるやかになる傾向がある(サンプルサイズが大きいときでも、βの推定量は負のバイアスを持つ)。

逆に、$\beta < 0$としよう(下図(b)参照)。このとき、測定誤差e_iが大きくなると、説明変数X_iと誤差項$u_i = u_i^* - \beta e_i$はともに大きくなる。つまり、X_iとu_iに正の相関がある。ここでX_iとu_iに正の相関があるため、X_iが大きいと、u_iは大きく(正の値をとる)なり、観測点は実線の上で観察される。逆に、X_iが小さいと、u_iは小さく(負の値をとる)なり、観測点は実線の下で観察される。このため、サンプルサイズが大きいときでも、βの推定量は正のバイアスを持つ。

これまでの結果をまとめると、サンプルサイズが大きいときでも、$\beta > 0$なら負のバイアス$(\hat{\beta} < \beta)$、$\beta < 0$なら正のバイアス$(\hat{\beta} > \beta)$であり、係数は 0 方向へバイアスがあるといえる。

図：内生性とバイアスの関係



なお、図をみると、内生性があると、OLS 推定では係数$\hat{\beta}$だけでなく、定数項$\hat{\alpha}$にもバイアスが生じていることがわかる。この結果は重回帰分析においても当てはまる。つまり、説明変数のうち 1 つでも内生性があれば、OLS 推定

では、すべての係数や定数項にバイアスが生じるといえる。大事な点なので覚えておいて欲しい。

練習問題 2

$Y_i = Y_i^* + e_i$を書き換えると、$Y_i^* = Y_i - e_i$となる。これを、$Y_i^* = \alpha + \beta X_i + u_i^*$の式に代入すると、

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

となり、さらに書き換えると、

$$Y_i = \alpha + \beta X_i + (u_i^* + e_i)$$

となる。新しい誤差項を$u_i = u_i^* + e_i$と定義すると、Y_iとX_iには次の関係がある。

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

仮定から、説明変数X_iは、u_i^*やe_iと無相関である。したがって、X_iは新しい誤差項$u_i = u_i^* + e_i$と無相関であり、内生性の問題は生じない[12]。よって、OLS 推定量を用いても、サンプルサイズが大きいなら、係数βの推定にバイアスは生じない。

練習問題 3

このモデルを次のように書き換える。

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

$$u_i^* = \gamma W_i + u_i$$

ただし、X_iは内生変数であり、元の式における誤差項u_iと相関している。

(a) Z_iとW_iが無相関ならば、操作変数Z_iと誤差項$u_i^* = \gamma W_i + u_i$は無相関となり、欠落変数がある状況であっても、Z_iは「操作変数の条件②(操作変数の外生性)」を満たす。よって、サンプルサイズが大きいなら 2SLS にバイアスは生じない。

(b) Z_iとW_iが相関するならば、操作変数Z_iと誤差項$u_i^* = \gamma W_i + u_i$は相関するため、欠落変数がある状況では、Z_iは「操作変数の条件②(操作変数の外生性)」を満たさない。このため、サンプルサイズが大きくても、2SLS にバイアスが生じることになる。

---

12 仮に測定誤差e_iが説明変数X_iと相関していれば、OLS 推定量は一致性を持たない。通常、この練習問題と同様、測定誤差e_iは説明変数と無相関と仮定される。

練習問題 4

第 1 段階では、次の式を推定する。

$$X_i = \gamma_0 + \gamma_1 Z_{1i} + \gamma_2 Z_{2i} + \gamma_3 W_{1i} + \gamma_4 W_{2i} + e_i$$

操作変数の条件①は、仮説を次のようにした F 検定によって確認できる。

$$H_0: \gamma_1 = 0, \qquad \gamma_2 = 0$$

$$H_1: \text{帰無仮説} H_0 \text{が誤りである}$$

F 値が 10 を下回ると操作変数の条件①が満たされない可能性がある。

練習問題 5

操作変数としては、ルームメートがランダムに決まっていることを利用すればよい。たとえば、操作変数Z_iを、ルームメートがゲーム機を持ち込んだら 1、持ち込まなかったら 0 となるダミー変数としよう[13]。部屋の割り振りはランダムであるため、操作変数Z_iは誤差項u_iと無相関となる。また、ゲーム機が持ち込まれると勉強時間は減るため、Z_iとX_iは相関するだろう。

練習問題 6

(a) 誤差項は$u_t = \rho u_{t-1} + \varepsilon_t$となるため、$u_{t-1}$が変化すると$u_t$が変化する。また、$t-1$期には、次の関係があるため、$u_{t-1}$が変化すると$Y_{t-1}$も変化する。

$$Y_{t-1} = \alpha + \beta Y_{t-2} + \gamma W_{t-1} + u_{t-1}$$

以上から、u_tとY_{t-1}は相関しており、内生性の問題が生じる。

(b) $t-1$期には、$Y_{t-1} = \alpha + \beta Y_{t-2} + \gamma W_{t-1} + u_{t-1}$が成立するため、

$$\begin{aligned} Y_t - \rho Y_{t-1} &= (\alpha + \beta Y_{t-1} + \gamma W_t + u_t) - \rho(\alpha + \beta Y_{t-2} + \gamma W_{t-1} + u_{t-1}) \\ &= \alpha(1-\rho) + \beta Y_{t-1} - \rho\beta Y_{t-2} + \gamma W_t - \rho\gamma W_{t-1} + (u_t - \rho u_{t-1}) \end{aligned}$$

となる。ここで、$\varepsilon_t = u_t - \rho u_{t-1}$であり、左辺の$\rho Y_{t-1}$を右辺に移項させると、

$$Y_t = \alpha(1-\rho) + (\rho + \beta) Y_{t-1} - \rho\beta Y_{t-2} + \gamma W_t - \rho\gamma W_{t-1} + \varepsilon_t$$

となる。右辺にある説明変数はY_{t-1}、Y_{t-2}、W_t、W_{t-1}であり、モデルは

$$Y_t = \alpha^* + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \beta_3 W_t + \beta_4 W_{t-1} + \varepsilon_t$$

[13] この操作変数に関心がある方は、以下の論文を参照されたい。Stinebrickner, Ralph Stinebrickner and Todd R. Stinebrickner. (2008)"The Causal Effect of Studying on Academic Performance," *The B.E. Journal of Economic Analysis and Policy*, 8(1) (Frontiers), Article 14, 1-53.

と表せる。ただし、パラメータは次のように定義される。

$$\alpha^* = \alpha(1-\rho)、\beta_1 = \rho + \beta、\beta_2 = -\rho\beta、\beta_3 = \gamma、\beta_4 = -\rho\gamma$$

ここで、誤差項は$\varepsilon_t \sim i.i.d.N(0,\sigma^2)$であるから、誤差項$\varepsilon_t$は全ての説明変数と無相関になる。

練習問題 7

2SLS の第 1 段階では、被説明変数をXとし、説明変数をZとした OLS 推定をする。そして、Xの予測値を$\hat{X}_i = \hat{\gamma}_0 + \hat{\gamma}_1 Z_i$とする。ただし、$\hat{\gamma}_1$は、次のようになる(2.3.1 節参照)。

$$\hat{\gamma}_1 = \frac{\sum_{i=1}^n (Z_i - \bar{Z})(X_i - \bar{X})}{\sum_{i=1}^n (Z_i - \bar{Z})^2}$$

2SLS の第 2 段階では、被説明変数をY_iとし説明変数を$\hat{X}_i$とした OLS 推定することで、係数 β を推定する。このとき、2SLS 推定量$\hat{\beta}_{2SLS}$は次のようになる。

$$\hat{\beta}_{2SLS} = \frac{\sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})(Y_i - \bar{Y})}{\sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})^2}$$

ただし、$\bar{\hat{X}}$は、$\hat{X}_i = \hat{\gamma}_0 + \hat{\gamma}_1 Z_i$の標本平均、つまり、

$$\bar{\hat{X}} = \frac{1}{n}\sum_{i=1}^n \hat{X}_i = \frac{1}{n}\sum_{i=1}^n (\hat{\gamma}_0 + \hat{\gamma}_1 Z_i) = \hat{\gamma}_0 + \hat{\gamma}_1 \bar{Z}$$

となる。よって、偏差$\hat{X}_i - \bar{\hat{X}}$は次のようになる。

$$\hat{X}_i - \bar{\hat{X}} = (\hat{\gamma}_0 + \hat{\gamma}_1 Z_i) - (\hat{\gamma}_0 + \hat{\gamma}_1 \bar{Z}) = \hat{\gamma}_1 (Z_i - \bar{Z})$$

ここで、$\hat{\beta}_{2SLS}$の式に、$\hat{X}_i - \bar{\hat{X}} = \hat{\gamma}_1 (Z_i - \bar{Z})$を代入すると、次のようになる。

$$\hat{\beta}_{2SLS} = \frac{\sum_{i=1}^n \left(\hat{X}_i - \bar{\hat{X}}\right)(Y_i - \bar{Y})}{\sum_{i=1}^n \left(\hat{X}_i - \bar{\hat{X}}\right)^2} = \frac{\sum_{i=1}^n \hat{\gamma}_1 (Z_i - \bar{Z})(Y_i - \bar{Y})}{\sum_{i=1}^n \hat{\gamma}_1^2 (Z_i - \bar{Z})^2} = \frac{1}{\hat{\gamma}_1} \frac{\sum_{i=1}^n (Z_i - \bar{Z})(Y_i - \bar{Y})}{\sum_{i=1}^n (Z_i - \bar{Z})^2}$$

さらに、$\hat{\gamma}_1 = \sum_{i=1}^n (Z_i - \bar{Z})(X_i - \bar{X}) / \sum_{i=1}^n (Z_i - \bar{Z})^2$を代入すると、

$$\hat{\beta}_{2SLS} = \left\{\frac{\sum_{i=1}^n (Z_i - \bar{Z})^2}{\sum_{i=1}^n (Z_i - \bar{Z})(X_i - \bar{X})}\right\} \frac{\sum_{i=1}^n (Z_i - \bar{Z})(Y_i - \bar{Y})}{\sum_{i=1}^n (Z_i - \bar{Z})^2} = \frac{\sum_{i=1}^n (Z_i - \bar{Z})(Y_i - \bar{Y})}{\sum_{i=1}^n (Z_i - \bar{Z})(X_i - \bar{X})}$$

最後に、偏差の和が 0 という性質($\sum_{i=1}^n (Y_i - \bar{Y}) = 0$、 $\sum_{i=1}^n (X_i - \bar{X}) = 0$)を用いると、2SLS 推定量$\hat{\beta}_{2SLS}$は、次のようになる。

$$\hat{\beta}_{2SLS} = \frac{\sum_{i=1}^n (Z_i - \bar{Z})(Y_i - \bar{Y})}{\sum_{i=1}^n (Z_i - \bar{Z})(X_i - \bar{X})} = \frac{\sum_{i=1}^n Z_i (Y_i - \bar{Y}) - \bar{Z}\sum_{i=1}^n (Y_i - \bar{Y})}{\sum_{i=1}^n Z_i (X_i - \bar{X}) - \bar{Z}\sum_{i=1}^n (X_i - \bar{X})} = \frac{\sum_{i=1}^n Z_i (Y_i - \bar{Y})}{\sum_{i=1}^n Z_i (X_i - \bar{X})}$$

練習問題 8

練習問題 7 の結果から、2SLS 推定量$\hat{\beta}_{2SLS}$は、

$$\hat{\beta}_{2SLS}=\frac{\sum_{i=1}^{n}Z_i(Y_i-\bar{Y})}{\sum_{i=1}^{n}Z_i(X_i-\bar{X})}$$

となる。ここでは、分子と分母の別表現を、それぞれ求める。

分子の別表現: 分子は、次のように表現できる

$$\sum_{i=1}^{n}Z_i(Y_i-\bar{Y})=\left(1-\frac{n_1}{n}\right)n_1(\bar{Y}_1-\bar{Y}_0)$$

ただし、$\bar{Y}_1$は$Z_i=1$のときのYの平均、$\bar{Y}_0$は$Z_i=0$のときのYの平均となる。

$$\bar{Y}_1=\frac{1}{n_1}\sum_{i=1}^{n_1}Y_i \qquad \bar{Y}_0=\frac{1}{n-n_1}\sum_{i=n_1+1}^{n}Y_i$$

[証明] この別表現が正しいことを確認しよう。分子は、

$$\sum_{i=1}^{n}Z_i(Y_i-\bar{Y})=\sum_{i=1}^{n_1}Z_i(Y_i-\bar{Y})+\underbrace{\sum_{i=n_1+1}^{n}Z_i(Y_i-\bar{Y})}_{=0}=\sum_{i=1}^{n_1}(Y_i-\bar{Y})$$

となる。式展開では、$i=1,2,\dots,n_1$なら$Z_i=1$、$i=n_1+1,n_1+2,\dots,n$なら$Z_i=0$となることを用いた。さらに右辺は次のようになる。

$$\sum_{i=1}^{n_1}(Y_i-\bar{Y})=\sum_{i=1}^{n_1}Y_i-n_1\bar{Y}=n_1\bar{Y}_1-n_1\bar{Y}$$

ここで、標本平均$\bar{Y}$は次のように分解できる。

$$\bar{Y}=\frac{1}{n}\left(\sum_{i=1}^{n_1}Y_i+\sum_{i=n_1+1}^{n}Y_i\right)=\frac{1}{n}(n_1\bar{Y}_1+(n-n_1)\bar{Y}_0)$$

この式を、$\sum_{i=1}^{n_1}(Y_i-\bar{Y})=n_1\bar{Y}_1-n_1\bar{Y}$の式に代入すると、

$$\begin{aligned}\sum_{i=1}^{n_1}(Y_i-\bar{Y})&=n_1\bar{Y}_1-n_1\left\{\frac{1}{n}(n_1\bar{Y}_1+(n-n_1)\bar{Y}_0)\right\}\\&=n_1\bar{Y}_1-\frac{n_1}{n}n_1\bar{Y}_1-\left(1-\frac{n_1}{n}\right)n_1\bar{Y}_0\\&=\left(1-\frac{n_1}{n}\right)n_1\bar{Y}_1-\left(1-\frac{n_1}{n}\right)n_1\bar{Y}_0\\&=\left(1-\frac{n_1}{n}\right)n_1(\bar{Y}_1-\bar{Y}_0)\end{aligned}$$

分母の別表現：分母は、次のように表現できる。

$$\sum_{i=1}^{n} Z_i(X_i - \bar{X}) = \left(1 - \frac{n_1}{n}\right) n_1(\bar{X}_1 - \bar{X}_0)$$

ただし、$\bar{X}_1$は$Z_i = 1$のときのXの平均、$\bar{X}_0$は$Z_i = 0$のときのXの平均とする。

$$\bar{X}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} X_i \qquad \bar{X}_0 = \frac{1}{n - n_1}\sum_{i=n_1+1}^{n} X_i$$

[証明] この別表現が正しいことを確認しよう。分母は、

$$\sum_{i=1}^{n} Z_i(X_i - \bar{X}) = \sum_{i=1}^{n_1} Z_i(X_i - \bar{X}) + \underbrace{\sum_{i=n_1+1}^{n} Z_i(X_i - \bar{X})}_{=0} = \sum_{i=1}^{n_1}(X_i - \bar{X})$$

となる。さらに右辺は、

$$\sum_{i=1}^{n_1}(X_i - \bar{X}) = n_1\bar{X}_1 - n_1\bar{X}$$

となり、これに

$$\bar{X} = \frac{1}{n}(n_1\bar{X}_1 + (n - n_1)\bar{X}_0)$$

を代入すると、分母は次のようになる。

$$\begin{aligned}\sum_{i=1}^{n_1}(X_i - \bar{X}) &= n_1\bar{X}_1 - n_1\left\{\frac{1}{n}(n_1\bar{X}_1 + (n - n_1)\bar{X}_0)\right\} \\ &= \left(1 - \frac{n_1}{n}\right) n_1(\bar{X}_1 - \bar{X}_0)\end{aligned}$$

分子と分母の別表現を代入する:

$$\hat{\beta}_{2SLS} = \frac{\sum_{i=1}^{n} Z_i(Y_i - \bar{Y})}{\sum_{i=1}^{n} Z_i(X_i - \bar{X})} = \frac{\left(1 - \frac{n_1}{n}\right) n_1(\bar{Y}_1 - \bar{Y}_0)}{\left(1 - \frac{n_1}{n}\right) n_1(\bar{X}_1 - \bar{X}_0)} = \frac{\bar{Y}_1 - \bar{Y}_0}{\bar{X}_1 - \bar{X}_0}$$

上式の意味を考えよう。まず、$\bar{X}_0$は$Z_i = 0$のときのXの平均、$\bar{X}_1$は$Z_i = 1$のときのXの平均であるため、分母は「Zが 0 から 1 に変化したとき、Xの平均がどれぐらい変化したか」を表す。同様に、分子は、「Zが 0 から 1 に変化したとき、Yの平均がどれぐらい変化したか」を表す。まとめると、2SLS 推定量$\hat{\beta}_{2SLS}$は次のように解釈できる。

$$\hat{\beta}_{2SLS}=\frac{Z\text{が変化したときの}Y\text{の平均の変化}}{Z\text{が変化したときの}X\text{の平均の変化}}$$

<u>練習問題 11</u>

13.6 節では、次のモデルを考えている。

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

$$X_i=\gamma_0+\gamma_1 Z_i+e_i$$

ただし、誤差項u_iは、操作変数の条件②によって$\gamma_0+\gamma_1 Z_i$とは無相関だが、e_iとは相関している。

第 1 段階では、$X_i=\gamma_0+\gamma_1 Z_i+e_i$を推定し、予測値$\hat{X}_i=\hat{\gamma}_0+\hat{\gamma}_1 Z_i$を計算する。OLS 推定量($\hat{\gamma}_0$、$\hat{\gamma}_1$)は、内生変数$X_i$を用いて計算されており、$X_i$の誤差項$e_i$の関数となる。この点は、OLS 推定量の確率的表現からも理解できる。

$$\hat{\gamma}_1=\gamma_1+\frac{\sum_{i=1}^{n}(Z_i-\bar{Z})e_i}{\sum_{i=1}^{n}(Z_i-\bar{Z})^2}$$

したがって、$\hat{X}_i=\hat{\gamma}_0+\hat{\gamma}_1 Z_i$は$e_i$と相関しており、ひいては誤差項$u_i$とも相関する(つまり、$\hat{X}_i$は外生変数ではない)。ただし、サンプルサイズが大きいなら、$\hat{\gamma}_0$、$\hat{\gamma}_1$はγ_0、γ_1に収束するため、予測値は$\gamma_0+\gamma_1 Z_i$となり、e_iの関数とはならない。

この問題を解決する方法として、サンプルを 2 分割(データ A、データ B)することが考えられる。まず、データ A を用いて第 1 段階の推定をし、OLS 推定量($\hat{\gamma}_0$、$\hat{\gamma}_1$)を求める。次に、データ A で得られた$\hat{\gamma}_0$、$\hat{\gamma}_1$を用いて、データ B の予測値$\hat{X}_i=\hat{\gamma}_0+\hat{\gamma}_1 Z_i$を計算し、被説明変数を$Y_i$、説明変数を$\hat{X}_i$とした OLS 推定を行う。ポイントは、データ A から$\hat{\gamma}_0$、$\hat{\gamma}_1$を計算しているため、これらはデータ B のe_iとは無相関となる点にある。この手法は split-sample IV(SSIV)と呼ばれる。ただし、この方法では、サンプルが半分しか用いられておらず、効率的な推定ができない。また、どのように分割するかで推定結果も異なってしまう。

SSIV を発展させた手法として、jackknife instrumental variables estimator がある。この方法を簡単に説明しよう。まず、1 番目のデータだけを除いたサンプル(X_2、X_3、$,\dots,X_n$、Z_2、Z_3、$,\dots,Z_n$)で第 1 段階目の推定を行い、OLS 推定量($\hat{\gamma}'_0$、$\hat{\gamma}'_1$)、そして予測値$\hat{X}_1=\hat{\gamma}'_0+\hat{\gamma}'_1 Z_1$を求める。$\hat{X}_1$の計算に$X_1$は用いられていないため、$e_1$の関数とはならない。次に、2 番目のデータを除いたサンプル(X_1、X_3、$,\dots,X_n$、Z_1、Z_3、$,\dots,Z_n$)で第 1 段階目の推定を行い、OLS 推定量($\hat{\gamma}'_0$、$\hat{\gamma}'_1$)、そして予測値

$\hat{X}_2 = \hat{\gamma}'_0 + \hat{\gamma}'_1 Z_2$を求める。同じことを、すべての i に行うことで、予測値$\hat{X}_i$が計算できる。これを用いて、2 段階最小 2 乗法を行えばよい。興味がある方は、Bruce Hansen の『Econometrics』を参照されたい(巻末参考文献 D 参照)。

練習問題 12

OLS 推定量は、説明変数に内生性がなければ一致性は満たされる一方、内生性があれば一致性は満たされない。これに対し、2SLS 推定量は、内生性の有無に関係なく一致性を満たす。教科書では、OLS 推定量と 2SLS 推定量の結果を比較することで、OLS 推定量に、どれぐらいのバイアスが生じているかを確認した(例 13-2、例 13-3 参照)。これは 2 つの推定結果の差が小さければ、内生性は生じていない可能性を示唆しているためである。これは 11 章で学習したハウスマン検定の考え方に似ている。

この検定では、仮説を次のように設定する。

帰無仮説H_0：説明変数に内生性が存在しない

対立仮説H_1：説明変数に内生性が存在する

帰無仮説が正しいとき、OLS 推定量と 2SLS 推定量の差は 0 に近い値をとる。

$$\text{OLS 推定量} - \text{2SLS 推定量} \approx 0$$

しかし、対立仮説が正しいとき、この差は 0 から乖離する。したがって、この差が十分に大きいとき、帰無仮説は棄却される。こうした検定は、開発者の名前をとって、Durbin-Wu-Hausman 検定、Wu-Hausman 検定とも呼ばれる。

同じ検定を、コントロール関数を用いて行うこともできる。モデルは次式としよう。

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

$$X_i = \gamma_0 + \gamma_1 Z_i + e_i$$

操作変数の条件②によって、Z_iはu_iと無相関となる。内生性の有無は、u_iとe_iが相関しているかで判断される。u_iとe_iが相関したら(関係式$X_i = \gamma_0 + \gamma_1 Z_i + e_i$から)、$u_i$と$X_i$も相関する(内生性あり)。

ここでu_iを、X_iと相関する部分、相関しない部分に分解しよう。

$$u_i = \theta e_i + \varepsilon_i$$

ただし、$\theta = E[u_i e_i]/E[e_i^2]$とする。$X_i$と相関する部分は$\theta e_i$、相関しない部分は$\varepsilon_i$に

なる。ここで、$\theta \neq 0$なら内生性あり、$\theta = 0$なら内生性なしとなる。なお、e_iとε_iは無相関である。これは$\theta = E[u_i e_i]/E[e_i^2]$に注意すると、

$$E[e_i \varepsilon_i] = E[e_i(u_i - \theta e_i)] = E[u_i e_i] - \theta E[e_i^2] = 0$$

と確認できる。

コントロール関数は、Y_iの式に$u_i = \theta e_i + \varepsilon_i$を代入することで、

$$Y_i = \alpha + \beta X_i + \theta e_i + \varepsilon_i$$

と定義される。ここで、e_iとε_iは無相関であり、X_iとε_iも無相関である(ε_iは、X_iと相関しない部分である)。ただし、e_iは観察できないため、$X_i = \gamma_0 + \gamma_1 Z_i + e_i$を OLS 推定して得られた残差$\hat{e}_i$を用いる。つまり、被説明変数を$Y_i$とし、説明変数を$X_i$と$\hat{e}_i$とした OLS 推定によって、パラメータをバイアスなく推定できる。内生性の有無は、

帰無仮説H_0: $\theta = 0$(内生性なし)

対立仮説H_1:$\theta \neq 0$(内生性あり)

とした検定によって確認できる[14]。

練習問題 13

固定効果モデルは、説明変数と個別効果の相関を許容したが、時間を通じて一定の変数はすべて個別効果に含まれ、その係数が推定できない問題がある。これに対し、変量効果モデルは、説明変数と個別効果は無相関と仮定しているが、時間を通じて一定の変数の係数が推定できるという利点がある。時間を通じて一定の変数に関心があるとき、固定効果モデルを用いることができない一方、変量効果モデルは非現実的な仮定を課しているため、その使用には問題がある。この練習問題では、個別効果を考慮しながら、時間を通じて一定の変数の係数を推定する方法を学習する(これはハウスマン=テイラー法の単純なケースになる)。

11.2 節の(3)式では、$Y_{i,t} = \alpha + \beta X_{i,t} + Z_i + u_{i,t}$とした。これに時間を通じて一定の変数$W_i$を加えると、練習問題で与えられた式となる。

$$Y_{i,t} = \alpha + \beta X_{i,t} + \theta W_i + Z_i + u_{i,t}$$

[1414] Stata は reg X Z, r とし predict u, residual とする。そして reg Y X u, r とし test u とする。

モデルの誤差項は、$Z_i + u_{i,t}$となる。説明変数($X_{i,t}$、W_i)は$u_{i,t}$と無相関である。また、個別効果を構成するZ_iは、固定効果と変量効果の間をとるものとする。具体的には、Z_iは$X_{i,t}$と相関するが、Z_iはW_iとは無相関とする。

これらの仮定より、W_iは外生変数であるが、$X_{i,t}$は内生変数となる。まず、$X_{i,t}$はZ_iと相関しているため、$X_{i,t}$からZ_iと相関している部分をすべて取り除く。具体的には、被説明変数を$X_{i,t}$とし、説明変数をN個のダミー変数$(D1_i, D2_i, \ldots, DN_i)$とした OLS 推定を行う(ダミー変数は 11.3.1 節参照)。得られた残差$\dot{X}_{i,t}$は、$X_{i,t}$と相関しているが、(個別要因が取り除かれているため)Z_iとは無相関となる。よって、操作変数を残差$\dot{X}_{i,t}$とした 2 段階最小 2 乗法を行うことで、パラメータ(βとθ)を推定することができる[15]。

なお、ハウスマン=テイラー法の一般的なモデルはより複雑である。

$$Y_{i,t} = \alpha + \beta_1 X_{1i,t} + \beta_2 X_{2i,t} + \theta_1 W_{1i} + \theta_2 W_{2i} + Z_i + u_{i,t}$$

ここで、$X_{1i,t}$とW_{1i}はZ_iと無相関だが、$X_{2i,t}$とW_{2i}はZ_iと相関すると仮定する(すべて$u_{i,t}$とは無相関である)。誤差項が$Z_i + u_{i,t}$なら、$X_{2i,t}$とW_{2i}は内生変数となる[16]。

まず、$X_{2i,t}$はZ_iと相関しているため、$X_{2i,t}$からZ_iと相関している部分をすべて取り除く。具体的には、被説明変数を$X_{2i,t}$とし、説明変数をN個のダミー変数とした OLS 推定を行う。得られた残差$\dot{X}_{2i,t}$は、$X_{2i,t}$と相関しているが、(個別要因が取り除かれているため)Z_iとは無相関となる。また、$X_{1i,t}$の時間平均をとる($\bar{X}_{1i} = \frac{1}{T}\sum_{s=1}^{T} X_{1i,s}$)。$X_{1i,t}$は$Z_i$と無相関であるため、$\bar{X}_{1i}$もまた$Z_i$と無相関になる。$\bar{X}_{1i}$は$W_{2i}$と相関しているなら、操作変数として用いることができる。操作変数を$\dot{X}_{2i,t}$、$\bar{X}_{1i}$とした 2SLS 推定を行うことで一致性を持った推定量が得られる[17]。

<u>練習問題 14</u>

モデルは$Y_i = \alpha + \beta X_i + u_i$であり、操作変数は$Z_i$とする。練習問題 7 から、

---

[15] Stata なら、$\dot{X}_{i,t}$を計算してから 2SLS を行うことになる。$\dot{X}_{i,t}$は、egen xmean = mean(x), by(id)とする(id とは、i の番号を表す変数に当たる)。そして、gen xdot = x-xmean とすればよい。そして、次のようにすればよい。xdot を操作変数として、ivreg を使って推定すればよい。

[16] 練習問題は、$X_{1i,t}$とW_{2i}がないケースに該当していた。

[17] Stata では、コマンド xthtaylor を用いることでハウスマン=テイラー推定ができる。具体的には、xthtaylor Y X1 X2 W1 W2, endog(X2 W2) r とすればよい。endog の()には、内生変数を入れたらよい。なお、xthtaylor では、X の 1 つは外生変数、W の 1 つは外生変数でないと実行できない。したがって、練習問題のような推定は、xtreg を使って推定することになる。

$$\hat{\beta}_{2SLS}=\frac{\sum_{i=1}^{n}(Z_i-\bar{Z})(Y_i-\bar{Y})}{\sum_{i=1}^{n}(Z_i-\bar{Z})(X_i-\bar{X})}$$

である。ここで分子は、次のようになる。

$$\sum_{i=1}^{n}(Z_i-\bar{Z})(Y_i-\bar{Y})=\sum_{i=1}^{n}(Z_i-\bar{Z})Y_i-\bar{Y}\sum_{i=1}^{n}(Z_i-\bar{Z})=\sum_{i=1}^{n}(Z_i-\bar{Z})Y_i$$

式展開では、偏差の和は 0 となることを用いた($\sum_{i=1}^{n}(Z_i-\bar{Z})=0$)。$Y_i=\alpha+\beta X_i+u_i$を用いると、

$$\hat{\beta}_{2SLS}=\frac{\sum_{i=1}^{n}(Z_i-\bar{Z})Y_i}{\sum_{i=1}^{n}(Z_i-\bar{Z})(X_i-\bar{X})}=\frac{\sum_{i=1}^{n}(Z_i-\bar{Z})(\alpha+\beta X_i+u_i)}{\sum_{i=1}^{n}(Z_i-\bar{Z})(X_i-\bar{X})}$$

$$=\frac{\alpha\sum_{i=1}^{n}(Z_i-\bar{Z})}{\sum_{i=1}^{n}(Z_i-\bar{Z})(X_i-\bar{X})}+\beta\frac{\sum_{i=1}^{n}(Z_i-\bar{Z})X_i}{\sum_{i=1}^{n}(Z_i-\bar{Z})(X_i-\bar{X})}+\frac{\sum_{i=1}^{n}(Z_i-\bar{Z})u_i}{\sum_{i=1}^{n}(Z_i-\bar{Z})(X_i-\bar{X})}$$

ここで、偏差の和は 0 から、第 1 項目は 0、第 2 項目の分子は

$$\sum_{i=1}^{n}(Z_i-\bar{Z})(X_i-\bar{X})=\sum_{i=1}^{n}(Z_i-\bar{Z})X_i-\bar{X}\sum_{i=1}^{n}(Z_i-\bar{Z})=\sum_{i=1}^{n}(Z_i-\bar{Z})X_i$$

となる。これらを代入すると、次式が得られる。

$$\hat{\beta}_{2SLS}=\beta+\frac{\sum_{i=1}^{n}(Z_i-\bar{Z})u_i}{\sum_{i=1}^{n}(Z_i-\bar{Z})(X_i-\bar{X})}$$

(b)(a)で得られた式の第 2 項目の分子と分母を$n-1$で割ると、

$$\hat{\beta}_{2SLS}=\beta+\frac{\frac{1}{n-1}\sum_{i=1}^{n}(Z_i-\bar{Z})u_i}{\frac{1}{n-1}\sum_{i=1}^{n}(Z_i-\bar{Z})(X_i-\bar{X})}$$

となる。ここでnが大きくなると、標本共分散は共分散に収束するため、2SLS 推定量は、次のようになる。

$$\beta+\frac{Cov(Z_i,u_i)}{Cov(Z_i,X_i)}$$

ここで、第 2 項目の分母は、操作変数の条件①(関連性: $\mathrm{Cov}(Z_i,X_i)\neq 0$)から 0 ではない。また、分子は、操作変数の条件②(外生性: $\mathrm{Cov}(Z_i,u_i)=0$)から 0 となる。以上から、2SLS 推定量はβに収束する。

なお、(a)で得られた式の期待値をとると、

$$\hat{\beta}_{2SLS}=\beta+E\left[\frac{\sum_{i=1}^{n}(Z_i-\bar{Z})u_i}{\sum_{i=1}^{n}(Z_i-\bar{Z})(X_i-\bar{X})}\right]$$

となるが、第 2 項は一般に 0 とはならないので、2SLS 推定量は不偏性を満たさない。第 2 項が 0 となるためには、操作変数の条件①②よりも強い仮定が必要となる。こうした条件に関心がある読者は、巻末参考文献 D の[7][8][9]を参照されたい。

<u>練習問題 15</u>

教科書では、係数は全ての i に対して、同じ値をとると仮定してきた。しかし、i の属性によって、係数は異なる可能性がある。たとえば、健康状態がY_i、薬の投与量をX_iとしたとき、薬の効果β_iは個人属性によって異なるだろう。これは抗がん剤治療で治癒する人もいれば治癒しない人もいることから理解できる。

(a) 説明変数X_iがランダムに決定されているため、説明変数は個人属性とは独立になる。係数β_iや誤差項u_iは個人属性によって異なるため、説明変数は、係数や誤差項と独立な変数になる。

2.3.2 節から、OLS 推定量は、次のように表現できる。

$$\frac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2}$$

ここで分子は標本共分散、分母は標本分散である。

サンプルサイズ n が大きくなると、標本平均が期待値に収束するように、標本分散は分散に、標本共分散は共分散に収束していく。このため、n が大きくなると、OLS 推定量は次のようになる。

$$\begin{aligned}\frac{Cov(Y_i,X_i)}{Var(X_i)} &= \frac{Cov(\alpha+\beta_iX_i+u_i,X_i)}{Var(X_i)}\\ &= \frac{E[\beta_i]Var(X_i)}{Var(X_i)} = E[\beta_i]\end{aligned}$$

式展開では、X_iは係数β_iとは独立であるため、

$$Cov(\alpha+\beta_iX_i+u_i,X_i) = E[\beta_i]Var(X_i)$$

となることを用いた[18]。

[18] 式展開が難しいので、詳しくみていこう。$Cov(\alpha+\beta_iX_i+u_i,X_i)$は、その定義により、

$$E[(\alpha+\beta_iX_i+u_i-E[\alpha+\beta_iX_i+u_i])(X_i-E[X_i])] = E[(\beta_iX_i-E[\beta_iX_i]+u_i)(X_i-E[X_i])]$$

となる。ここで、X_iはu_iと無相関であることから$E[u_i(X_i-E[X_i])]=0$となり、上式の右辺は次のように書ける。

$$E[(\beta_iX_i-E[\beta_iX_i])(X_i-E[X_i])]$$

上式をさらに展開すると、

$$\begin{aligned}E[\beta_iX_i(X_i-E[X_i])]-E[\beta_iX_i]E[(X_i-E[X_i])] &= E[\beta_iX_i(X_i-E[X_i])]\\ &= E[\beta_i]E[X_i(X_i-E[X_i])]\\ &= E[\beta_i]Var(X_i)\end{aligned}$$

となる。1 行目の展開では$E[(X_i-E[X_i])]=0$となること、2 行目の展開ではX_iとβ_iが独立であること、3 行目の展開では$Var(X_i)=E[X_i(X_i-E[X_i])]$を用いた。

OLS 推定量が$E[\beta_i]$となることの意味を考えてみよう。確率変数β_iには、n 通りの可能性があり、それぞれ$1/n$の確率で生じるとしよう。このとき、

$$E[\beta_i] = \frac{1}{n}\sum_{i=1}^{n}\beta_i$$

となり、$E[\beta_i]$は、β_iの平均効果として解釈できる。たとえば、係数β_iが薬の効果であれば、$E[\beta_i]$は薬の平均効果と解釈できる。

(b) 練習問題 7 から、2SLS 推定量は、次のように表現できる。

$$\hat{\beta}_{2SLS} = \frac{\frac{1}{n-1}\sum_{i=1}^{n}(Z_i - \bar{Z})(Y_i - \bar{Y})}{\frac{1}{n-1}\sum_{i=1}^{n}(Z_i - \bar{Z})(X_i - \bar{X})}$$

サンプルサイズ n が大きくなると、標本平均が期待値に収束するように、標本分散は分散に、標本共分散は共分散に収束していく。このため、n が大きくなると、2SLS 推定量は次のようになる。

$$\frac{Cov(Z_i, Y_i)}{Cov(Z_i, X_i)}$$

まず、分母を考えよう。$Cov(Z_i, X_i)$は、その定義より、

$$E[(Z_i - E[Z_i])(X_i - E[X_i])] = E[(Z_i - E[Z_i])X_i]$$
$$= E[(Z_i - E[Z_i])(\gamma_0 + \gamma_i Z_i + e_i)]$$
$$= E[\gamma_i Z_i(Z_i - E[Z_i])] + E[(Z_i - E[Z_i])e_i]$$

ここでZ_iはe_iと独立であり、$E[(Z_i - E[Z_i])e_i] = E[(Z_i - E[Z_i])]E[e_i] = 0$となる。また、$Z_i$は$\gamma_i$と独立であり、

$$E[\gamma_i Z_i(Z_i - E[Z_i])] = E[\gamma_i]E[Z_i(Z_i - E[Z_i])] = E[\gamma_i]Var(Z_i)$$

となる。以上から、$Cov(Z_i, X_i) = E[\gamma_i]Var(Z_i)$となる。

次に、分子を考えよう。Y_iの式に、$X_i = \gamma_0 + \gamma_i Z_i + e_i$を代入すると、

$$Y_i = \alpha + \beta_i(\gamma_0 + \gamma_i Z_i + e_i) + u_i$$
$$= \alpha + \gamma_0\beta_i + \gamma_i\beta_i Z_i + u_i + \beta_i e_i$$

となる。$Cov(Z_i, Y_i)$は、その定義より、

$$E[(Z_i - E[Z_i])(Y_i - E[Y_i])] = E[(Z_i - E[Z_i])Y_i]$$
$$= E[(Z_i - E[Z_i])(\alpha + \gamma_0\beta_i + \gamma_i\beta_i Z_i + u_i + \beta_i e_i)]$$
$$= \gamma_0 E[(Z_i - E[Z_i])\beta_i] + E[(Z_i - E[Z_i])\gamma_i\beta_i Z_i] + E[(Z_i - E[Z_i])u_i]$$
$$+ E[(Z_i - E[Z_i])\beta_i e_i]$$

となる。Z_iは、β_i、γ_i、u_i、e_iと独立であるため、右辺第 2 項以外はすべて 0 となる。これは次の式から明らかである。

$$E[(Z_i - E[Z_i])\beta_i] = E\big[Z_i - E[Z_i]\big]E[\beta_i] = 0$$

$$E[(Z_i - E[Z_i])u_i] = E\big[Z_i - E[Z_i]\big]E[u_i] = 0$$

$$E[(Z_i - E[Z_i])\beta_i e_i] = E\big[Z_i - E[Z_i]\big]E[\beta_i e_i] = 0$$

右辺 2 項は、

$$E[(Z_i - E[Z_i])\gamma_i\beta_i Z_i] = E[\gamma_i\beta_i]E[Z_i(Z_i - E[Z_i])] = E[\gamma_i\beta_i]Var(Z_i)$$

となる。以上から、$Cov(Z_i, Y_i) = E[\gamma_i\beta_i]Var(Z_i)$となる。

これらの結果をまとめると、n が大きくなると、2SLS 推定量は

$$\frac{Cov(Z_i, Y_i)}{Cov(Z_i, X_i)} = \frac{E[\gamma_i\beta_i]Var(Z_i)}{E[\gamma_i]Var(Z_i)} = \frac{E[\gamma_i\beta_i]}{E[\gamma_i]}$$

となる。この式の意味を考えてみよう。ここで、γ_iとβ_iにはn通りの可能性があり、それぞれ$1/n$の確率で生じるとしよう。このとき、$E[\gamma_i] = \frac{1}{n}\sum_{i=1}^{n}\gamma_i = \bar{\gamma}$、

$$E[\gamma_i\beta_i] = \frac{1}{n}\sum_{i=1}^{n}\gamma_i\beta_i$$

となる。したがって、$\hat{\beta}_{2SLS}$は次式となる。

$$\frac{E[\gamma_i\beta_i]}{E[\gamma_i]} = \sum_{i=1}^{n}\frac{\gamma_i}{n\bar{\gamma}}\beta_i$$

これは、加重$\gamma_i/n\bar{\gamma}$を用いた係数β_iの加重平均となる。加重は、操作変数に対する反応度γ_iが大きいほど大きい。つまり、$\hat{\beta}_{2SLS}$は、遵守者(操作変数に対して強く反応する人)に対する平均効果ともいえる。

たとえば、学習時間X_iが成績Y_iに与える効果β_iが知りたいとする(練習問題 5 参照)。操作変数Z_iは、ルームメートがゲーム機を持ち込んだら 1、持ち込まなかったら 0 となるダミー変数である。学生の半数は、勤勉でゲーム機によって学習時間は変化せず($\gamma_i = \gamma_1 = 0$)、また、学習効果β_1は高いとしよう($\beta_i = \beta_1 > 0$)。残り半数は、勤勉ではなくゲーム機によって学習時間は減少し($\gamma_i = \gamma_2 < 0$)、学習効果β_2は低いとする($\beta_i = \beta_2 < \beta_1$)。このとき、平均効果は、$E[\beta_i] = \frac{1}{2}(\beta_1 + \beta_2)$となる。これに対し、$E[\gamma_i] = \frac{1}{2}(\gamma_1 + \gamma_2) = \frac{\gamma_2}{2}$、$E[\gamma_i\beta_i] = \frac{1}{2}(\beta_1\gamma_1 + \beta_2\gamma_2) = \frac{\beta_2\gamma_2}{2}$から、$\frac{E[\gamma_i\beta_i]}{E[\gamma_i]} = \beta_2$となる。ここで、2SLS 推定量は、ゲーム機によって影響を受ける学生の学習効果β_2になっており、学習の平均効果$\frac{1}{2}(\beta_1 + \beta_2)$を過小評価している(つまり、$\frac{E[\gamma_i\beta_i]}{E[\gamma_i]} <$

$E[\beta_i]$)。

(c) 2SLS 推定量が平均効果になっている状況を考えよう。これは条件①②③のいずれかを満たした状況となる。

条件①は、効果の異質性がない状況である($\beta_i = \beta$)。ここで、β_iは固定した値βをとるため、次式が成立する。

$$\frac{E[\gamma_i \beta_i]}{E[\gamma_i]} = \frac{E[\gamma_i]\beta}{E[\gamma_i]} = \beta$$

また、$E[\beta_i] = \beta$であるから、$E[\gamma_i \beta_i]/E[\gamma_i] = E[\beta_i]$となる。

条件②は、操作変数に対するX_iの反応度γ_iが一定となる状況である($\gamma_i = \gamma$)。γ_iは固定した値γをとるため、次式が成立する。

$$\frac{E[\gamma_i \beta_i]}{E[\gamma_i]} = \frac{\gamma E[\beta_i]}{\gamma} = E[\beta_i]$$

つまり、γ_iに異質性がなければ、$E[\gamma_i \beta_i]/E[\gamma_i] = E[\beta_i]$となる。

条件③は、反応度γ_iと係数β_iが無相関となる状況である($Cov(\gamma_i, \beta_i) = 0$)。つまり、反応度$\gamma_i$が高いとき、係数$\beta_i$が高いとか低いとかいう傾向がないとする。共分散$Cov(\gamma_i, \beta_i)$は、

$$E[(\gamma_i - E[\gamma_i])(\beta_i - E[\beta_i])] = E[\gamma_i \beta_i] - E[\gamma_i]E[\beta_i]$$

であり、これが 0 であるから、$E[\gamma_i \beta_i] = E[\gamma_i]E[\beta_i]$となる。よって、

$$\frac{E[\gamma_i \beta_i]}{E[\gamma_i]} = \frac{E[\gamma_i]E[\beta_i]}{E[\gamma_i]} = E[\beta_i]$$

となる。つまり、反応度γ_iと係数β_iが無相関なら、$E[\gamma_i \beta_i]/E[\gamma_i] = E[\beta_i]$となる。

教科書では、「操作変数を追加したり、変更したりしたとき、推定結果が大きく変わるようであれば、どちらかの(もしくは両方の)操作変数が、操作変数の条件を満たしていない可能性があります」とした。しかし、効果の異質性があるとき、これは必ずしも正しくない。つまり、条件①②③のいずれも満たされないなら、$\hat{\beta}_{2SLS}$は遵守者に対する平均効果となっており、これは操作変数の選択によって、$\hat{\beta}_{2SLS}$の値が変化する可能性を示唆している[19]。

[19] こうした状況では、かりに操作変数の条件が満たされていても、過剰識別検定は帰無仮説を棄却することに注意したい。

練習問題 16

(a) 強外生性を満たさないため、このモデルは固定効果モデルとして推定するのは不適当である。強外生性を満たさない点は、1 期前はモデルが

$$Y_{i,t-1} = \alpha + \beta Y_{i,t-2} + Z_i + u_{i,t-1}$$

となるから、$Y_{i,t-1}$と$u_{i,t-1}$は相関していることを確認できる。

(b) $Y_{i,t} - Y_{i,t-1} = \left(\alpha + \beta Y_{i,t-1} + Z_i + u_{i,t}\right) - \left(\alpha + \beta Y_{i,t-2} + Z_i + u_{i,t-1}\right)$

$$= \beta(Y_{i,t-1} - Y_{i,t-2}) + (u_{i,t} - u_{i,t-1})$$

つまり、前期との差をとることで個別要因を消すことができる。

(c) ここで、$Y_{i,t-2}$は$Y_{i,t-1} - Y_{i,t-2}$と相関する一方、$u_{i,t} - u_{i,t-1}$とは無相関である。この点は、誤差項に系列相関はないため、

$$Y_{i,t-2} = \alpha + \beta Y_{i,t-3} + Z_i + u_{i,t-2}$$

と$u_{i,t} - u_{i,t-1}$は全くの無関係であることがわかる。よって、$Y_{i,t-2}$は、$Y_{i,t-1} - Y_{i,t-2}$の操作変数として用いることができる。2SLS 推定によって、一致性のある推定結果が得られる。

操作変数は$Y_{i,t-2}$としたが、$Y_{i,t-3}$、$Y_{i,t-4}$、...もまた操作変数の候補となる。このため、実証分析では、$Y_{i,t-2}$だけでなく、$Y_{i,t-2}$、$Y_{i,t-3}$、…、$Y_{i,t-p}$を用いて、2SLS 推定を行うことになる。弱操作変数の問題が生じるため、pはあまり大きくないことが望ましい(たとえば、3～5 などが推奨される)。

14 章の答え

練習問題 1

(a) 仮に苗字が出身地に関する情報を含んでいるならば、ランダムな割り当てになっていない可能性がある。たとえば、沖縄などは特有の名前が多く、これが内的妥当性に疑義を生じさせる可能性はある。

(b) 講師への陳情によって当初の割り当てが遵守されなかったケースであり、最終的な割り当てがランダムになっていない。このため、これでは内的妥当性は低くなる[1]。たとえば、講師に陳情をするような人はやる気があると考えるならば、プログラムの効果を高く評価してしまう。

練習問題 2

DID 分析において、対照群は平行トレンドの仮定を満たしていなければいけない。

練習問題 3

モデルは次のとおりである。

$$Y_{i,t} = \alpha + \alpha_2 D_i + \theta_2 Time_t + \beta X_{i,t} + u_{it}$$

$Time_t$は 2 期(介入後)なら 1 となるダミー変数、$X_{i,t}$は交差項$D_i \times Time_t$となる。

2 時点の差$\Delta Y_{i,2} = Y_{i,2} - Y_{i,1}$は、次のようになる。

$$\begin{aligned}\Delta Y_{i,2} = Y_{i,2} - Y_{i,1} &= \left(\alpha + \alpha_2 D_i + \theta_2 + \beta X_{i,2} + u_{i,2}\right) - \left(\alpha + \alpha_2 D_i + u_{i,1}\right) \\ &= \theta_2 + \beta X_{i,2} + \Delta u_{i,2}\end{aligned}$$

ただし、$\Delta u_{i,2} = u_{i,2} - u_{i,1}$とする。つまり、被説明変数を$Y_{i,2} - Y_{i,1}$、説明変数を$X_{i,2}$として推定すれば、処置効果$\beta$を推定できる。

練習問題 4

最低賃金の例で必要な変数は、店舗i が NJ に属するなら 1 となるダミー変数

[1] 当初の割り当て結果は、実際の割り当て結果の良い操作変数となる。たとえば、Z_iはプログラムに割り当てられたら 1 となるダミー変数とし、D_iは、実際にプログラムに参加したら 1 となるダミー変数としよう(陳情などがあるため、Z_iとD_iは必ずしも一致しない)。当初の割当はランダムであるからZ_iは外生変数(個人の属性などと無関係)となる一方、当初の割り当てが遵守されるケースも多いためZ_iとD_iは相関する。このため、Z_iはD_iの良い操作変数となる。

NJ_i、2 期(介入後)なら 1 となるダミー変数$Time_t$、交差項$X_{i,t} = D_i \times Time_t$と、これらはパネルデータでなく、反復横断面データであっても定義できる変数となる。したがって、反復横断面データでも、DID 分析は可能となる。ただし、反復横断面データでは、店舗ごとの固定効果を考慮することはできない。

練習問題 5

T 大学合格者のうち最低点での合格者、最低点より 1 点低い不合格者の所得を比較することで、T 大学に合格したことの効果を調べることができる。

合格者(最低点で合格した受験者)と不合格者(最低点から 1 点低い点数で不合格となった受験者)は本質的に同じ集団と考えられる。この場合、合格か不合格かはランダムに生じていると見なすことができ、合否は受験者の属性などとは無相関となっている。したがって、両集団の所得の平均を比較すれば、T 大学に合格したことの効果を推定できるだろう。

練習問題 6

13 章の練習問題 8 では、操作変数Z_iがダミー変数の場合、2 段階最小 2 乗推定量(2SLS)は、次のワルド推定量として表現できることを示した。

$$\hat{\beta}_{2SLS} = \frac{\bar{Y}_1 - \bar{Y}_0}{\bar{X}_1 - \bar{X}_0}$$

ただし、$\bar{Y}_1$は$Z_i = 1$のときのYの標本平均、$\bar{Y}_0$は$Z_i = 0$のときのYの標本平均である(同様に、$\bar{X}_1$と$\bar{X}_0$は定義される)。

ここで操作変数Zは、割当変数Wが閾値c以上なら 1 となるダミー変数であるため、2SLS 推定量はワルド推定量となる。このとき、分子$\bar{Y}_1 - \bar{Y}_0$は、閾値前後でのYの平均の変化となる($\bar{Y}_1$は閾値以上のYの平均、$\bar{Y}_0$は閾値未満でのYの平均である)。同様に、分母$\bar{X}_1 - \bar{X}_0$は、閾値前後でのXの平均の変化となる。ここで、Xは処置が割り当てられたら 1 となるダミー変数であるため、その平均は処置割合(いわば処置確率)となる。つまり、閾値前後のデータを用いた 2SLS 推定量は

$$\hat{\beta}_{2SLS} = \frac{\text{閾値前後での}Y\text{の平均の変化}}{\text{閾値前後での処置確率の変化}}$$

となることが確認できる。

以下は、ファジーな RDD に関心がある方への補足となる。一般には、割当変数WがYに影響を与えると考えられる。ここで、割当変数が 3 乗まで影響すると考えるなら、回帰モデルは次のようになる。

$$Y_i = \alpha + \beta X_i + \theta_1 W_i + \theta_2 W_i^2 + \theta_3 W_i^3 + \gamma_1 W_i X_i + \gamma_2 W_i^2 X_i + \gamma_3 W_i^3 X_i + u_i$$

ここで、X_iは内生変数なので、交差項($W_i X_i$、$W_i^2 X_i$、$W_i^3 X_i$)も内生変数となる。操作変数Zは、割当変数Wが閾値c以上なら 1 となるダミー変数である。操作変数はZだけでなく、交差項(ZX、ZX^2、ZX^3)も操作変数となる。閾値前と後($c-h$から$c+h$までの区間)のデータを用いて、このモデルを 2SLS で推定すれば処置効果が推定できる。

第 15 章の答え

練習問題 1

GDP 成長率Y_tは、次のように決まる。

$$Y_t = a_0 + a_1 X_t + a_2 X_{t-1} + u_t$$

ここで、金利の式は、t期と$t-1$期において、それぞれ

$$X_t = b_0 + b_1 Y_{t-1}$$

$$X_{t-1} = b_0 + b_1 Y_{t-2}$$

となる。これらを GDP の式に代入すると、

$$Y_t = a_0 + a_1(b_0 + b_1 Y_{t-1}) + a_2(b_0 + b_1 Y_{t-2}) + u_t$$

$$= \underbrace{(a_0 + a_1 b_0 + a_2 b_0)}_{=c_0} + \underbrace{a_1 b_1}_{=c_1} Y_{t-1} + \underbrace{a_2 b_1}_{=c_2} Y_{t-2} + u_t$$

となる。ここで、パラメータを

$$c_0 = a_0 + a_1 b_0 + a_2 b_0、\ c_1 = a_1 b_1、\ c_2 = a_2 b_1$$

と定義すると、GDP が次の AR(2)モデルであることを意味する。

$$Y_t = c_0 + c_1 Y_{t-1} + c_2 Y_{t-2} + u_t$$

練習問題 2

ここで、Y_tは定常とする。モデルの両辺の期待値をとると、

$$E[Y_t] = a_0 + a_1 E[Y_{t-1}] + a_2 E[Y_{t-2}] + \cdots + a_p E[Y_{t-p}] + \mathrm{E}[u_t]$$

となる。ここで、$\mathrm{E}[u_t] = 0$、定常性から$E[Y_t] = E[Y_{t-1}] = E[Y_{t-2}] = \cdots = E\left[Y_{t-p}\right]$に注意すると、次のようになる。

$$E[Y_t] = a_0 + a_1 E[Y_t] + a_2 E[Y_t] + \cdots + a_p E[Y_t]$$

これを、$E[Y_t]$について解くと、

$$E[Y_t] = \frac{a_0}{1 - a_1 - a_2 - \cdots - a_p}$$

練習問題 3

AR(1)モデルのインパルス応答関数は、任意のsについて、

$$a_1^s$$

となる(導出は 15.2.2 節参照)。

半減期は「ショックの影響が半減するまでに要する期間」と定義される[2]。1 単位のショックの影響が半減するのに要する期間は、

$$a_1^s = 0.5$$

となるsであるから、上式のsが半減期となる。上式の対数をとると、

$$s \times \ln(a_1) = \ln(0.5)$$

となり、これをsについて解けば、半減期が、次のように求められる。

$$s = \frac{\ln(0.5)}{\ln(a_1)}$$

たとえば、$a_1 = 0.85$の場合、半減期は次のように 4.26 期となる。

$$s = \frac{\ln(0.5)}{\ln(0.85)} = 4.26$$

練習問題 4

日次データであれば、東京市場と NY 市場との間にある時差を用いて、除外制約が正当化できる。しかし、月次データであれば、そうした時差に意味はない。

むしろ、NY 市場は世界最大の市場であること、東京市場はアジアで有力な市場に過ぎないことを考えると、ダウ平均株価が外生的とした方がよいだろう。つまり、ダウ平均株価の変化率(%)をY_t、日経平均株価の変化率(%)をX_tとし、ダウ平均Y_tは同月の日経平均X_tから影響を受けないとする($b_{10} = 0$)。

ただし、日経平均から同月のダウ平均への影響は弱いとしても 0 であること($b_{10} = 0$)を正当化するのは難しいだろう。個人的には、月次データでは、除外制約を正当化するのは難しいと思う。

練習問題 5

構造 VAR は、次のようになる。

$$Y_t = a_{11}Y_{t-1} + b_{10}X_t + b_{11}X_{t-1} + \varepsilon_{Yt}$$

$$X_t = a_{20}Y_t + a_{21}Y_{t-1} + b_{21}X_{t-1} + \varepsilon_{Xt}$$

[2] 半減期は、放射能の測定でも用いられる。東京都観光局のウェブサイトによると、「放射性物質は、壊変（崩壊）を繰り返し、最終的に安定した物質へ変化すると放射線を放出しなくなります。壊変によって始めの放射性物質の数が半分になるまでの時間を半減期といい、放射能は、時間がたつにつれて弱まっていきます」とあります。

ここで、ε_{Yt}、ε_{Xt}は構造ショックとなる。

X_tの式を、Y_tの式に代入すると、

$$Y_t = a_{11}Y_{t-1} + b_{10}(a_{20}Y_t + a_{21}Y_{t-1} + b_{21}X_{t-1} + \varepsilon_{Xt}) + b_{11}X_{t-1} + \varepsilon_{Yt}$$

となる。これを整理すると

$$(1 - b_{10}a_{20})Y_t = (a_{11} + b_{10}a_{21})Y_{t-1} + (b_{10}b_{21} + b_{11})X_{t-1} + b_{10}\varepsilon_{Xt} + \varepsilon_{Yt}$$

となり、両辺を$(1 - b_{10}a_{20})$で割ると、

$$Y_t = \frac{a_{11} + b_{10}a_{21}}{1 - b_{10}a_{20}}Y_{t-1} + \frac{b_{10}b_{21} + b_{11}}{1 - b_{10}a_{20}}X_{t-1} + \frac{b_{10}\varepsilon_{Xt} + \varepsilon_{Yt}}{1 - b_{10}a_{20}}$$

となる。これは誘導型の VAR である。誘導型の誤差項u_{Yt}は次のようになる。

$$u_{Yt} = \frac{b_{10}\varepsilon_{Xt} + \varepsilon_{Yt}}{1 - b_{10}a_{20}} = \frac{b_{10}}{1 - b_{10}a_{20}}\varepsilon_{Xt} + \frac{1}{1 - b_{10}a_{20}}\varepsilon_{Yt}$$

次に、Y_tの式を、X_tの式に代入すると、同様の計算により、

$$X_t = a_{20}(a_{11}Y_{t-1} + b_{10}X_t + b_{11}X_{t-1} + \varepsilon_{Yt}) + a_{21}Y_{t-1} + b_{21}X_{t-1} + \varepsilon_{Xt}$$

となり、整理すると、

$$X_t = \frac{a_{20}a_{11} + a_{21}}{1 - b_{10}a_{20}}Y_{t-1} + \frac{a_{20}b_{11} + b_{21}}{1 - b_{10}a_{20}}X_{t-1} + \frac{a_{20}\varepsilon_{Yt} + \varepsilon_{Xt}}{1 - b_{10}a_{20}}$$

となる。これは誘導型の VAR であり、誘導型の誤差項u_{Xt}は次のようになる。

$$u_{Xt} = \frac{a_{20}\varepsilon_{Yt} + \varepsilon_{Xt}}{1 - b_{10}a_{20}} = \frac{a_{20}}{1 - b_{10}a_{20}}\varepsilon_{Yt} + \frac{1}{1 - b_{10}a_{20}}\varepsilon_{Xt}$$

誘導型の誤差項(u_{Yt}, u_{Xt})は、構造ショック(ε_{Yt}、ε_{Xt})の線形関数であり、それ自体が何を意味しているか分からないという問題がある[3]。

説明変数として、変数のラグから構成される VAR は誘導 VAR と呼ばれる。予測においては、個々の誤差項の解釈に興味はないため、誘導 VAR を用いて問題はない。しかし、経済的解釈に関心があるならば、構造 VAR を用いることになり、構造ショックを識別するための何らかの仮定が必要となる。

[3] そもそも、構造型が真のモデルであるの対して、誘導型とは、ラグだけに依存するように構造型を書き換えたモデルとなる。このため、構造ショックが真のショックであり、誘導型の誤差項は、モデルを書き換えたことから生じた構造ショックの線形関数に過ぎない。誘導型の誤差項に経済的な意味はないことに注意してほしい。仮に政府支出の増加の影響を見たいなら、構造ショックが 1 単位増加した状況を考えるべきなのである。

練習問題 8

季節性を考慮するため、月次データならラグ次数を 12 とするのは妥当な選択に思われる(四半期データなら 4 とする)。しかしこの場合、サンプルサイズよりパラメータ数が多く、このモデルを推定できない。

サンプルサイズは 240(=12 か月×20 年)となる。5 変量の VAR(p)モデルは、1 本の式に$1+5p$個のパラメータがあり、式は計 5 本あるので、パラメータ数は$5\times(1+5p)$となる。$p=12$の場合、パラメータ数は 305(=5×(1+5×12))となり、サンプルサイズよりパラメータ数が多くなってしまう(5 章の練習問題 14(e)では、サンプルサイズよりパラメータ数が多いと多重共線性が生じることを指摘している)。

練習問題 9

排除制約($b_{10}=c_{10}=c_{20}=0$)のもとで、構造 VAR は次のようになる。

$$Y_t=\mu_1+a_{11}Y_{t-1}+b_{11}X_{t-1}+c_{11}Z_{t-1}+\varepsilon_{Yt}$$

$$X_t=\mu_2+a_{20}Y_t+a_{21}Y_{t-1}+b_{21}X_{t-1}+c_{21}Z_{t-1}+\varepsilon_{Xt}$$

$$Z_t=\mu_3+a_{30}Y_t+a_{31}Y_{t-1}+b_{30}X_t+b_{31}X_{t-1}+c_{31}Z_{t-1}+\varepsilon_{Zt}$$

1 番目の式は誘導型となる（説明変数がラグだけである)。このため、$u_{Yt}=\varepsilon_{Yt}$となる。この式を 2 番目の式に代入すると、

$$X_t=\mu_2+a_{20}(\mu_1+a_{11}Y_{t-1}+b_{11}X_{t-1}+c_{11}Z_{t-1}+\varepsilon_{Yt})+a_{21}Y_{t-1}+b_{21}X_{t-1}+c_{21}Z_{t-1}+\varepsilon_{Xt}$$

$$=\mu_2+a_{20}\mu_1+(a_{20}a_{11}+a_{21})Y_{t-1}+(a_{20}b_{11}+b_{21})X_{t-1}+(a_{20}c_{11}+c_{21})Z_{t-1}+a_{20}\varepsilon_{Yt}+\varepsilon_{Xt}$$

となるため、$u_{Xt}=a_{20}\varepsilon_{Yt}+\varepsilon_{Xt}$と分かる。最後に、$Y_t$と$X_t$の式を、$Z_t$の式に代入すると、$u_{Zt}=a_{30}\varepsilon_{Yt}+b_{30}\varepsilon_{Xt}+\varepsilon_{Zt}$とわかる(式展開が面倒なので省略した)。

練習問題 10

実証分析では、最後の変数 Z を短期金利にすることが多い。これは中央銀行が物価と経済状況を考慮したうえで、金融政策を決定しているためである。たとえば、日本銀行では、金融政策決定会合が年 8 回開催されており、四半期データなら当期の GDP とインフレ率の変化に十分に反応できるだろう。

残りの変数の順番については判断が難しい。ただし、高頻度データであれば、投資は事前に決定されており、短期金利が変化しても投資額は変化しないと考

えられる。よって、短期金利は実質 GDP やインフレ率に影響を与えないので、排除制約のうち$c_{10} = c_{20} = 0$が満たされる。四半期データが十分に高頻度かどうかは微妙だが、ここでは仮定が満たされると仮定しよう (なお、年次データでは、この仮定は満たされない)。

残りの変数の順番を考えよう。排除制約($b_{10} = c_{10} = c_{20} = 0$)のもとで、誘導型ショックは構造ショックを用いて次のように表せる(練習問題 8 参照)。

$$u_{Yt} = \varepsilon_{Yt}$$

$$u_{Xt} = a_{20}\varepsilon_{Yt} + \varepsilon_{Xt}$$

$$u_{Zt} = a_{30}\varepsilon_{Yt} + b_{30}\varepsilon_{Xt} + \varepsilon_{Zt}$$

金利 Z の構造ショックε_{Zt}は政策ショックと解釈される。マクロ経済学で学習する総需要総供給分析では、実質 GDP と物価は、総供給曲線と総需要曲線の交点で決定される。このため、構造ショック(ε_{Yt}、ε_{Xt})は、それぞれ供給ショックと需要ショックと考えられる。供給ショックは総供給曲線をシフトさせるショック、需要ショックは総需要曲線をシフトさせるショックである。ここで 3 つの可能性がある[4]。

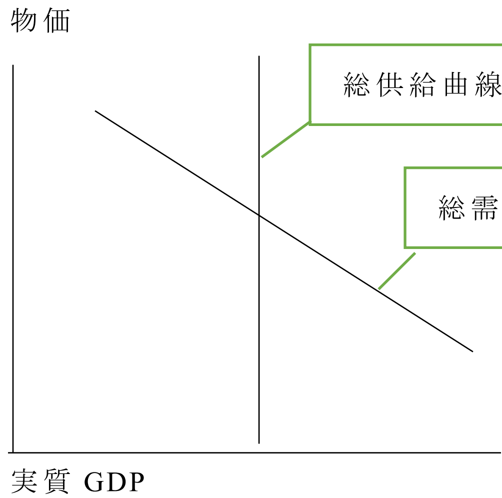
第 1 の可能性　1 番目の変数 Y を実質 GDP 成長率、2 番目の変数 X をインフレ率とする。これは総供給曲線が垂直なケースに該当する(下図(a)参照)。このとき、ε_{Yt}は供給ショック、ε_{Xt}は需要ショックとなる。供給ショックε_{Yt}は総供給曲線をシフトさせ、実質 GDP とインフレ率を変化させる。これに対し、需要ショックε_{Xt}は、総需要曲線をシフトさせ、インフレ率だけを変化させる。

第 2 の可能性　1 番目の変数 Y をインフレ率、2 番目の変数 X を実質 GDP とする。これは総供給曲線が水平なケースに該当する(下図(b)参照)。供給ショックε_{Yt}は実質 GDP とインフレ率を変化させるが、需要ショックε_{Xt}は実質 GDP だけを変化させる。

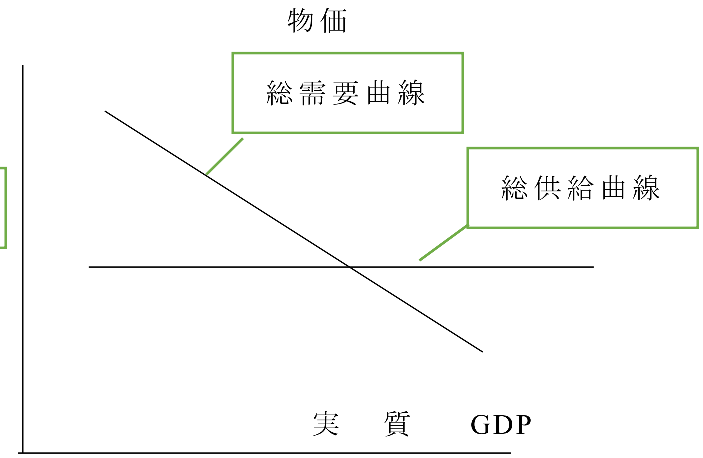
[4] ここでの記述は付録 D の[14]をもとに作成した。

図　総需要曲線と総供給曲線

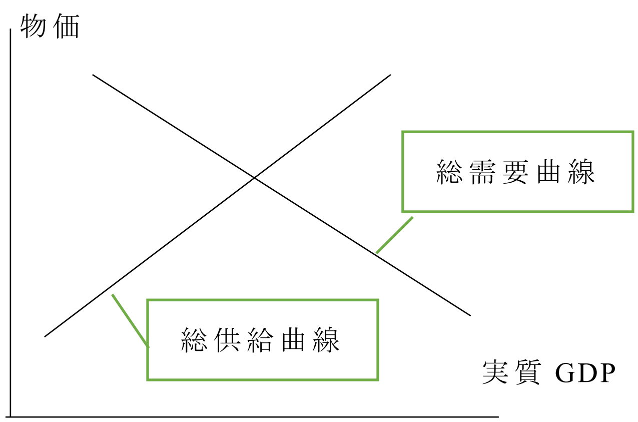
(a) 総供給曲線が垂直



(b) 総供給曲線が水平



(c)総供給曲線が右上がり



第 3 の可能性　総供給曲線が右上がりの傾きをもっているとき、需要ショックと供給ショックは、インフレ率と実質 GDP の両方に影響を与える(下図(c)参照)。このため、実質 GDP とインフレ率の順番をどのようにしても、排除制約は不適当になる($b_{10} \neq 0$)。ただし、金利 Z は短期的には、インフレ率と実質 GDP に影響しないため、$c_{10} = c_{20} = 0$は成立する。このとき、政策ショックε_{Zt}のインパルス応答関数は推定できるが、他の構造ショックのインパルス応答関数は推定できないことになる[5]。

[5] $c_{10} = c_{20} = 0$のもとで、構造 VAR は次のように表現できる(ただし、$b_{10} \neq 0$とする)。

$$Y_t = \mu_1 + a_{11}Y_{t-1} + b_{10}X_t + b_{11}X_{t-1} + c_{11}Z_{t-1} + \varepsilon_{Yt}$$

練習問題 11

Kilian(2009)では、1973年3月から2007年12月までの月次データを用いて、原油の需要ショックと供給ショックが原油価格に与える影響を推定している[6]。この論文では、Y_tは世界原油生産量の変化率(%)、X_tは世界経済活動指数、Z_tは実質原油価格としている。また、構造ショック(ε_{Yt}、ε_{Xt}、ε_{Zt})は3つあり、それぞれ原油供給ショック、総需要ショック、原油需要ショックと呼んでいる。総需要ショックは経済活動の高まりから生じる原油需要、原油需要ショックは原油供給の将来不安から予備的動機として生じる原油需要である。

変数の順番として、世界原油生産量の変化率(%)、世界経済活動指数、実質原油価格とした理由を考えよう。

1番目の変数は世界原油生産量である。原油生産量は事前の需要見込みから決定されており、世界経済活動指数や実質原油価格が上がっても、生産量は当月内では反応できない。これは原油市場の供給曲線が短期的には垂直であることを意味し、生産量は供給ショックだけの影響を受ける。よって、ε_{Yt}は供給ショックと解釈できる。なお、この仮定は月次データなら正当化されるが、四半期データなら成立しないことに注意してほしい。

2番目の変数としては、世界経済活動指標が適当だろう。世界経済活動の変化は遅いため、原油価格の変化は世界経済活動に当月内では影響しないと考え

$$X_t = \mu_2 + a_{20}Y_t + a_{21}Y_{t-1} + b_{21}X_{t-1} + c_{21}Z_{t-1} + \varepsilon_{Xt}$$

$$Z_t = \mu_3 + a_{30}Y_t + a_{31}Y_{t-1} + b_{30}X_t + b_{31}X_{t-1} + c_{31}Z_{t-1} + \varepsilon_{Zt}$$

ここで、Z_tの式は内生性が生じないため、この構造型の式の OLS 推定は一致性を持つ(ε_{Zt}が変化するとZ_tが変化するが、Y_tとX_tは変化しないことに注意してほしい)。よって、その残差は構造ショックε_{Zt}となる。構造ショックが推定できるため、それがインフレ率、実質 GDP 成長率、短期金利に与える効果が推定できる。なお、1番目と2番目の構造型は推定できないが、次の誘導型なら内生性の問題はなく、OLS 推定は一致性を持つ。

$$Y_t = a_{11}Y_{t-1} + b_{11}X_{t-1} + c_{11}Z_{t-1} + u_{Yt}$$

$$X_t = a_{21}Y_{t-1} + b_{21}X_{t-1} + c_{21}Z_{t-1} + u_{Xt}$$

構造型の式から、ε_{Zt}が1単位変化すると、Z_tが1単位変化する。同時点では、ε_{Zt}からインフレ率や実質 GDP に影響しないため、Y_tやX_tは変化しない。そして、誘導型の式から、Z_tが1単位変化すると、Y_{t+1}とX_{t+1}がどれぐらい変化するかわかる。これを繰り返していけば、インパルス応答関数が推定できる。

[6] Kilian, L. (2009) Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market, *American Economic Review* 99(3), 1053-1069.

られる[7]。なお、ε_{Xt}は世界経済活動指標の構造ショックであり、総需要ショックと解釈できる。

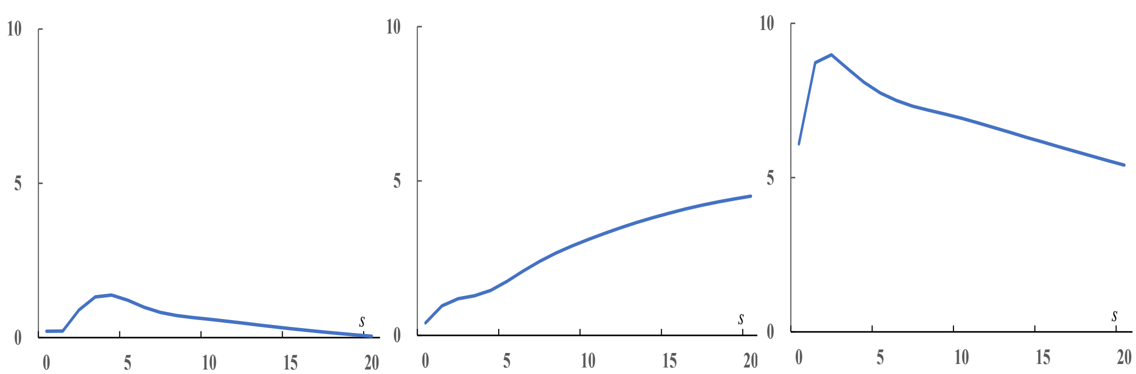
3 番目の変数は、実質原油価格である。実質原油価格は、供給ショックと総需要ショックから影響を受ける。また、実質原油価格の構造ショックε_{Zt}は、原油需要ショックと解釈できる。原油需要ショックは、原油供給の将来不安などから予備的動機として原油需要が増大するようなショックとなる。

サポートウェブサイトにあるデータ(oil2.csv)を用いて、Kilian(2009)の実証結果を再現してみよう。まず、ラグ次数を 3 として、3 変量 VAR を推定した(AIC は 3、BIC は 2 を選択したので長い次数である 3 を用いた)。

下図では、供給ショック、総需要ショック、原油需要ショックの 1 標準誤差の増加が原油価格に与える影響をみている(図を見やすくするため、供給ショックだけ 1 標準誤差の減少とした)。下図(a)をみると、供給ショックは原油価格にほとんど影響を与えていない。これに対し、下図(b)をみると、総需要ショックは、当初はあまり原油価格に影響を与えないが、その影響は徐々に大きくなっている。最後に、下図(c)をみると、原油需要ショックは、原油価格を大きく上昇させるが、その影響は徐々に低下している。

図　構造ショックに対する原油価格の反応

(a)供給ショック　　(b) 総需要ショック　　(c)　原油ショック



この結果から、原油価格の急激な変動は、供給ショックや総需要ショックではなく、将来不安から生じる原油需要ショックが大きな原因になっていることがわかる。

[7] 当月内で用いられる原油はすでに取引が終わっており、その月内の経済活動には影響しないということも言えるだろう。

第 16 章の答え

練習問題 1

DF 検定は AR(1)に基づいた単位根検定、ADF 検定は DF 検定を AR(p)の場合($p > 1$)に拡張した単位根検定となる。ラグ次数を 1 とする根拠はないため、AIC、BIC、MAIC などでラグ次数を選択し、それに基づいて ADF 検定を行うことになる。ただし、単位根検定では、MAIC が最も望ましい情報量規準とされる。

実証分析で DF 検定を行うのは、情報量規準などで$p = 1$が選択されたときである。ただし、その場合であっても、論文などでは DF 検定ではなく、ADF 検定と表記することが多い。

練習問題 2

「見せかけの回帰」は、回帰分析において、トレンドが考慮されていないことから生じる問題となる。まず、ADF 検定や DF-GLS 検定などの単位根検定を行って、データ系列に単位根が存在するか否かを確認する。

データ系列に単位根が存在するならば、トレンドは確率トレンドであり、変数の階差をとってから分析することによって、「見せかけの回帰」の問題が解決できる。これに対して、データ系列に単位根が存在しないならば、トレンドは確定トレンドであり、説明変数にトレンド変数tを加えることによって問題が解決できる。

練習問題 3

AR(4)モデルとして、

$$Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + a_3 Y_{t-3} + a_4 Y_{t-4} + u_t$$

を考える。まず、右辺に$(a_4 Y_{t-3} - a_4 Y_{t-3})$を加えると(0 を加えても等式は同じ)

$$\begin{aligned} Y_t &= a_1 Y_{t-1} + a_2 Y_{t-2} + (a_3 + a_4) Y_{t-3} - a_4 (Y_{t-3} - Y_{t-4}) + u_t \\ &= a_1 Y_{t-1} + a_2 Y_{t-2} + (a_3 + a_4) Y_{t-3} - a_4 \Delta Y_{t-3} + u_t \end{aligned}$$

となる(ただし、$\Delta Y_{t-3} = Y_{t-3} - Y_{t-4}$)。さらに右辺に$((a_3 + a_4) Y_{t-2} - (a_3 + a_4) Y_{t-2})$を加えると、次のようになる(ただし、$\Delta Y_{t-2} = Y_{t-2} - Y_{t-3}$)。

$$\begin{aligned} Y_t &= a_1 Y_{t-1} + (a_2 + a_3 + a_4) Y_{t-2} - (a_3 + a_4)(Y_{t-2} - Y_{t-3}) - a_4 \Delta Y_{t-3} + u_t \\ &= a_1 Y_{t-1} + (a_2 + a_3 + a_4) Y_{t-2} - (a_3 + a_4) \Delta Y_{t-2} - a_4 \Delta Y_{t-3} + u_t \end{aligned}$$

さらに右辺に $((a_2+a_3+a_4)Y_{t-1}-(a_2+a_3+a_4)Y_{t-1})$を加えると、

$$Y_t=(a_1+a_2+a_3+a_4)Y_{t-1}-(a_2+a_3+a_4)(Y_{t-1}-Y_{t-2})-(a_3+a_4)\Delta Y_{t-2}-a_4\Delta Y_{t-3}+u_t$$
$$=(a_1+a_2+a_3+a_4)Y_{t-1}-(a_2+a_3+a_4)\Delta Y_{t-1}-(a_3+a_4)\Delta Y_{t-2}-a_4\Delta Y_{t-3}+u_t$$

となる。最後に、両辺からY_{t-1}を引くと、

$$\Delta Y_t=(a_1+a_2+a_3+a_4-1)Y_{t-1}-(a_2+a_3+a_4)\Delta Y_{t-1}-(a_3+a_4)\Delta Y_{t-2}-a_4\Delta Y_{t-3}+u_t$$

が得られる。ここで

$$\rho=a_1+a_2+a_3+a_4-1、$$
$$\rho_1=-(a_2+a_3+a_4)、\ \rho_2=-(a_3+a_4)、\ \rho_3=-a_4$$

と定義すれば証明は終わり。つまり、帰無仮説H_0: $\rho=0$は、

$$a_1+a_2+a_3+a_4-1=0$$

を検定していることになる(つまり、$a_1+a_2+a_3+a_4=1$)。

練習問題 4

ここで、t は月次での時点となるので、y_tとy_{t-12}との間の関係式を求める。$t-1$期に$y_{t-1}=a_1y_{t-2}+u_{t-1}$が成立するので、

$$y_t=a_1y_{t-1}+u_t=a_1(a_1y_{t-2}+u_{t-1})+u_t$$
$$=a_1^2y_{t-2}+u_t+a_1u_{t-1}$$

となる。さらに$y_{t-2}=a_1y_{t-3}+u_{t-2}$を代入すると

$$y_t=a_1^3y_{t-3}+u_t+a_1u_{t-1}+a_1^2u_{t-2}$$

となる。これを繰り返すことで

$$y_t=a_1^{12}y_{t-12}+u_t+a_1u_{t-1}+a_1^2u_{t-2}+\cdots+a_1^{11}u_{t-11}$$

となる。つまり、月次の係数a_1を 12 乗したものが年次の係数になる。

月次の係数$a_1=0.95$であれば、年次の係数は$0.95^{12}=0.54$と小さくなる。単位根検定では、高頻度のデータを用いてサンプルサイズを大きくするのではなく、サンプル期間を延長することが重要となる。購買力平価の実証研究では、100 年を超えるサンプル期間を用いた研究がおこなわれている。

ただし、サンプル期間を長くとると、構造変化が生じる可能性が高くなることに注意が必要である。また、日本であれば、第 2 次世界大戦の前後では経済体制が大きく異なり、これらを同一データとして扱ってよいかという問題もあるだろう。こうした問題を回避するため、パネルデータを用いた単位根検定も

開発されている。詳しくは、ウォルター・エンダース『実証のための計量時系列分析』(有斐閣、新谷元嗣・藪友良訳)を参照されたい。

練習問題 5

モデル$Y_t = Y_{t-1} + u_t$に$Y_{t-1} = Y_{t-2} + u_{t-1}$を代入すると、

$$Y_t = (Y_{t-2} + u_{t-1}) + u_t$$

となる。さらに、$Y_{t-2} = Y_{t-3} + u_{t-2}$を代入すると、

$$Y_t = (Y_{t-3} + u_{t-2}) + u_t + u_{t-1}$$

さらに、$Y_{t-3} = Y_{t-4} + u_{t-3}$を代入すると、

$$Y_t = (Y_{t-4} + u_{t-3}) + u_t + u_{t-1} + u_{t-2}$$

となる。こうした代入を繰り返すと、

$$Y_t = Y_0 + u_t + u_{t-1} + u_{t-2} + \cdots + u_2 + u_1$$

が得られる。これはまさに$Y_t = Y_0 + \sum_{i=1}^{t} u_i$である。

これが非定常であることを確認してみよう。特殊ケースとして、$Y_0 = 0$、$E[u_i] = 0$、$E[u_i^2] = \sigma^2$、$E\left[u_i u_j\right] = 0$としよう。このとき、

$$E[Y_t] = \mathrm{E}\left[\sum_{i=1}^{t} u_i\right] = \sum_{i=1}^{t} E[u_i] = 0$$

$$V(Y_t) = \mathrm{V}\left(\sum_{i=1}^{t} u_i\right) = \sum_{i=1}^{t} E[u_i^2] = \sigma^2 t$$

となる。期待値は 0 であるが、分散はtとともに増加するため、非定常であることがわかる。