

經濟論叢

第七十八卷 第一號

- 農林業課税の問題……………神戸正雄（1）
- マックス・ウェーバーが考えていた經濟理論……………出口勇藏（12）
- 社會政策學の理論的性恪……………岸本英太郎（29）
- 時系列回歸分析における方程式誤差と變數誤差……………阿部統（55）
- 山陽自由黨の組織過程……………内藤正中（70）
- ジェントリの社會的經濟的性恪……………武暢夫（96）
- アメリカにおける特別償却本質論……………高寺貞男（116）
- ソヴェト社會史の時代區分について……………富岡裕（134）
-

[昭和三十一年七月]

京都大學經濟學會

經濟時系列における回歸分析と變數誤差

阿 部 統

△爲替レートの切下げは、貿易收支の均衡回復に効果があるか√という周知の爲替安定性論の命題を實地にテストするのには、多くの研究は、諸國の輸入需要の價格弾力性は一よりはるかに小さいがゆえに爲替切下げの効果はほとんど期待できないという。しかし、たいていのはあいはそのかげに、「これは短期分析である。したがって長期になれば、所得効果やその他がはたらいて、あるていどは貿易收支の改善が見られるだろう」と期待しているように見える。これにたいして、現實に政策として爲替切下げが論ぜられるばあいには、しばしば、「爲替切下げは一時的には効果があるかもしれないが、やがて國內のインフレがその効果を相殺してしまふだろう」と、さいきんのメキシコの例などもちだされる。この兩者の見解の相異は、國際經濟理論における古典派と近代派の對立にもむすびついてはなはだ興味ふかいが、そのさい、妥協點のひとつとして雙方の立場からは認されていることは、價格弾力性の計算の方法および統計資料が不適切で信頼しがたいものであり、そのために計算結果がゼロの方向への偏りをもつというオーカットの反證である。この反證は、一般に時系列回歸分析における統計誤差の處理の仕方の本

質にふれる問題なので、この小論において組織的にかんがえてみようと思ふ。

- (1) たとえば L. A. Metzler, "The Theory of International Trade," *A Survey of Contemporary Economics*, Vol. I, H. S. Ellis, ed. 1948, P. 245. 小島清譯「國際貿易の理論」(都留重人譯監修「現代經濟學の展望」理論編Ⅰ、一〇六頁)
- (2) *International Economics—Studies*, 1951 年よち *International Trade and Economic Development*, 1953 の序説をよせ J. Viner の見解、それによつた G. Haberler の反論を見よ。G. Haberler, "The Relevance of the Classical Theory under Modern Conditions," *American Economic Review*, May 1954.
- (3) G. H. Oxart, "Measurement of Price Elasticities in International Trade," *Review of Economics and Statistics*, May 1950.

二

はじめに例をあげよう。マクドゥガルは比較生産費原理の妥當性を統計的に檢證する目的をもつて、△アメリカ(イギリス)の方がイギリス(アメリカ)にくらべて輸出量の多い製品は、はたしてアメリカ(イギリス)の方が比較優位にたつているかどうか√を、まず△アメリカ(イギリス)の方が相對價格が安いかどうか√というかたちで、しらべようとした。そのため、兩國からともに輸出される製品一〇九品目をえらび、一九三四—三八年の資料をもつて、それぞれの製品につき相對輸出量

Total quantity of U. S. exports, 1934-38	Total quantity of U. K. exports, 1934-38
Relative quantity of U. S. exports, 1934-38	Relative quantity of U. K. exports, 1934-38

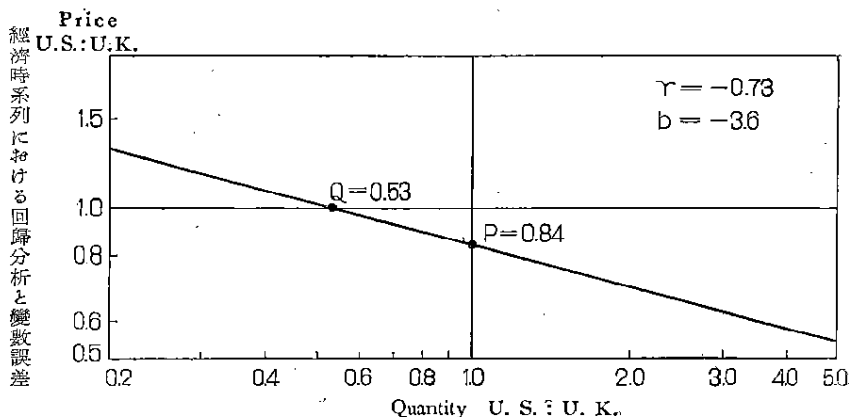
および相對價格

Average value of U. S. exports, 1934-38	Average value of U. K. exports, 1934-38
Relative value of U. S. exports, 1934-38	Relative value of U. K. exports, 1934-38

を計算し、その間の回帰直線をもつて第一圖のような結果をえた。相關係數は-0.74、スロープは-3.6である。つまり、資料によると、「一九三四—三八年の期間では、製品間に平均一パーセントの相對價格の差異があると、三・六パーセントの相對輸出量の差異をともなつていた」わけだが、マクドゥガルはこの回帰直線がもっと横にねる可能性があることを論じ、「統計資料に誤差がなければ、一パーセン

第1圖

Relative Price & Quantities of U. S. & U. K. Exports



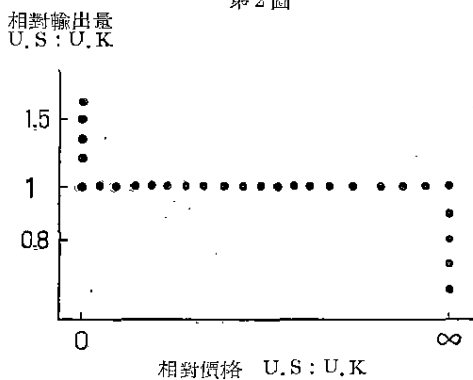
經濟時系列における回帰分析と變數誤差

トの相對價格差は、すくなくとも四・四・五パーセントの相對輸出量差をとまうはず」として、つぎの論據を指摘する。

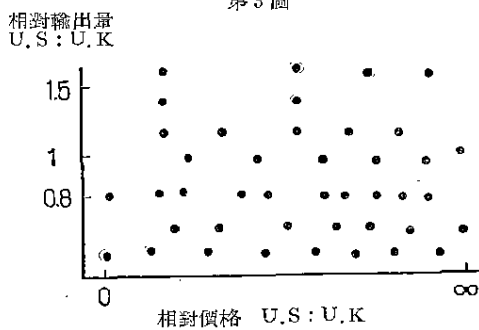
いま輸送費を無視すると、國際市場に完全競争がおこなわれ、比較生産費原理が完全に妥當したばあい、統計資料に誤差がなければ兩國の相對價格と相對輸出量のあいだの關係は第二圖のようになるはずである。すなわち、アメリカの價格がイギリスの價格よりもすこしでも安いとイギリスからの輸出はまったくなくなり、イギリスの價格がアメリカの價格よりもすこしでも安いとアメリカからの輸出はまったくなくなる。しかしこのばあいでも、統計資料に觀察誤差がふくまれると、多かれ少なかれ觀測點は第三圖のようにちらばり、水平の方向に偏差の平方和を最小にしようとする回歸直線は傾斜をもたざるをえなくなる。いかえれば、回歸方程式における獨立變數が觀察誤差をとまうとき、回歸係數の推定値をふつうの方法で計算するならば、その値は一般に眞の回歸係數よりも小になるのである。

この關係を數式によつてしめせば、つぎのようになる。いま、相對價格の觀測値を X 、相對輸出量の觀測値を Y とし、各觀測値

第2圖



第3圖



はそれぞれ眞の値 ($X \cdot Y$) と觀察誤差 ($X' \cdot Y''$) の和であるとすれば

$$X = X' + X'' \quad Y = Y' + Y''$$

となる。ところで X' と Y' のあいだに

$$X' = a + \beta Y'$$

という關係があるものと想定すれば、(1)を(2)に代入して

(2)

(1)

$$Y = a + \beta X + Y'' - \beta X'' \quad (3)$$

をうる。 $X \cdot Y \cdot X'' \cdot Y''$ の平均値をそれぞれ $\bar{X} \cdot \bar{Y} \cdot \bar{X}'' \cdot \bar{Y}''$ とし、 $x \cdot y \cdot x'' \cdot y''$ を各平均値からの偏差とすれば

$$X = \bar{X} + x \quad Y = \bar{Y} + y$$

$$X'' = \bar{X}'' + x'' \quad Y'' = \bar{Y}'' + y'' \quad (4)$$

となるから、(4)を(3)に代入して

$$Y = \beta \bar{X} + y'' - \beta \bar{X}'' + (a - \bar{Y} + \beta \bar{X} + \bar{Y}'' - \beta \bar{X}'') \quad (5)$$

をうる。右邊のカッコのなかはゼロである。(5)式に邊々々をかけて、あらゆる x について總計すれば

$$\sum x y = \beta \sum x^2 + \sum x y'' - \beta \sum x x'' \quad (6)$$

したがって

$$\beta = \frac{\sum x y - \sum x y''}{\sum x^2 - \sum x x''} \quad (7)$$

これが眞の回歸係数である。

しかし、統計資料をナマのままでもちいて、ふつうの方法で回歸係数を計算すると、推定値として

$$b = \frac{\sum x y}{\sum x^2} \quad (8)$$

をうる。 β と b の大きさを比較するために

(i) 各變數の眞の値と觀察誤差とのあいだには相關關係がない。すなわち

$$E(x'x'') = 0, E(y'y'') = 0$$

(ii) 從屬變數の觀察誤差は、獨立變數の眞の値と相關關係がない。すなわち

$$E(x^i y^j) = 0$$

(iii) 兩變數の觀察誤差相互間には相關關係がない。すなわち

$$E(x^i y^j) = 0$$

と假定して、(8)式を展開すれば

$$\begin{aligned} \beta &= \frac{\sum_{x^i y^i} - \sum_{x^i} \sum_{y^i} - \sum_{x^i y^i} - \sum_{x^i y^i}}{\sum_{x^i y^i} - \sum_{x^i} \sum_{y^i} - \sum_{x^i y^i} - \sum_{x^i y^i}} \\ &= \frac{\sum_{x^i y^i}}{\sum_{x^i y^i} - \sum_{x^i} \sum_{y^i}} \end{aligned} \quad (9)$$

となるから

$$\begin{aligned} \frac{b}{\beta} &= \frac{\sum_{x^i} - \sum_{x^i y^i}}{\sum_{x^i} - \sum_{x^i y^i}} \\ &= 1 - \frac{\sum_{x^i y^i}}{\sum_{x^i}} < 1 \end{aligned} \quad (10)$$

より、 $b < \beta$ である。つまり、インドゥガルの論點が證明されたわけである。

(4) G. D. A. MacDougall, "British and American Exports: A Study Suggested by the Theory of Comparative Costs," *Economic Journal*, Dec. 1951.

(5) 圖は對數目盛をもつてある。

(6) 森田優三「經濟變動の統計分析法」岩波全書。一六八―六九頁。以下の證明も、本質的には森田教授の證明方法とかわらな
5。

さて、われわれが問題とするのは、統計資料が誤差をふくむばあい、上にのべたような回帰係数の推定値の偏りをどのようにして處理するか、ということである。この問題はすでにある程度古典的のものになっており、さまざまな解決への工夫がみられるが、いずれも經濟時系列に適用するにあたっては、多くの難點をのこしているように思われる。また、ふつうの實證的研究では、問題をまったく無視してしまつて、單に古典的な回帰係数を計算するだけで安んじているが、このときにはその回帰係数が不偏推定値であるための假設を暗然のうちにもとめねばならない。このような假設が經濟現象にそのまま認されるかどうかはきわめてうたがわしい。

ここで、以下の議論に即して、もう一度問題の所在を定式化しておこう。簡單のために、變數 $x \cdot y$ 間の回帰關係を考え、常數項をゼロとする。變數の觀察誤差のほかは方程式誤差 ϵ を導入し、從屬變數が獨立變數以外の要因からうける disturbance をしめすものとすれば、

$$y'_z = ax'_z + \epsilon_z$$

$$y_z = y'_z + y''_z$$

$$x_z = x'_z + x''_z$$

$$(z=1, 2, \dots, n)$$

が成立する。ここに、 $x \cdot y$ は各變數の眞の値を、 $x' \cdot y'$ は觀察誤差をしめすものとする。

さて、 $x \cdot y \cdot \epsilon$ に ついてつぎの假設をもうける。

$$(i) E(y''_z | x'_z, \dots, x''_z; x'_z, \dots, x''_z) = 0$$

$$(ii) E(x''_z | x'_z, \dots, x''_z) = 0$$

$$(iii) E(\epsilon_z | x'_z, \dots, x''_z; x'_z, \dots, x''_z) = 0$$

$$\text{すなわち } z=1, 2, \dots, n$$

(i)は從屬變數の觀察誤差が獨立變數の眞の値に關しても、觀察誤差に關しても、獨立であることをしめし、(ii)は獨

立變數の眞の値と觀察誤差が相互に無關係であることをしめしている。また(ii)は方程式誤差が獨立變數と相關關係をもたないことを規定している。このばあい、各誤差の確率分布の平均値はすべてゼロとする。

ところで、最小二乗法によって β の推定値をもとめると

$$b = \frac{\sum_{x_i} x_i y_i}{\sum_{x_i} x_i^2}$$

となるが、これを展開すれば

$$b = \frac{\sum_{x_i} x_i (\beta x_i + y_i' - \beta x_i' + \varepsilon_i)}{\sum_{x_i} x_i^2}$$

$$= \beta + \frac{\sum_{x_i} x_i y_i'}{\sum_{x_i} x_i^2} - \beta \frac{\sum_{x_i} x_i x_i'}{\sum_{x_i} x_i^2} + \frac{\sum_{x_i} x_i \varepsilon_i}{\sum_{x_i} x_i^2}$$

をうる。假設(i)(ii)(iii)より、右邊第二・第四項の期待値はゼロとなるから、

$$E(b) = \beta - \beta E\left(\frac{\sum_{x_i} x_i x_i'}{\sum_{x_i} x_i^2}\right)$$

したがって、この式の右邊第二項がゼロでないかぎり、 b は β にくらべて偏りをもつことになる。

この偏りを處理するために、最近ベルクソンはきわめて巧妙な方法をしめた。¹⁰⁾ 彼によれば、多くの統計資料——とくに自然科学的な實驗資料——においては、 x よりも x' が先決變數と考えられる。このようなばあいは、眞の値 x よりも觀察値 x' をコンスタントと考えた方がよい。そして、 x' はむしろ觀察誤差 x'' の存在によって變動すると考え、前述の假設(ii)の代りに、

假設(ii) $E(x_i' | x_i, x_2, \dots, x_n) = 0$ ($i=1, 2, \dots, n$)

をもちれば、

$$E\left(\frac{\sum xy_i}{\sum x_i^2}\right) = 0$$

となるから

$$E(b) = \beta$$

が成立する。すなわち β が β の不偏推定値となるが、そればかりでなく、 $\beta \cdot \beta \cdot \beta$ がそれぞれ正規分布をなすと考え、それぞれ系列相關を有しないとすれば、フィッシャー流の信頼度検定も適用できることが證明されている。

以上のベルクソンの工夫は、いままでみのがされていた論點をたくみについている。しかし、經濟統計資料にこの種の誤差を想定するのはいかにがなものであるか。たとえば、つぎのような例が想起される。資本係數をひとつの經濟變量と考えて、その値を確定するのに國民資本の總額に關する統計資料を缺いているようなばあい、他國の實績からほぼ自國も4との想定をおいて、眞の値は4を中心にして毎年動搖するとする。このさい、動搖のあり方を、4を中心とした對稱分布と前提すれば、假設(iv)が妥當する。だが、このような例はむしろまれて、大部分の經濟資料にかんする觀察誤差は、上述の假設(ii)と假設(iv)の中間に位置するような性質をもつものと思われる。

けれども、そのばあいも、ベルクソンの工夫にならつて回歸係數の不偏推定値をもとめうるかなり有效な方法を考へうる。それは獨立變數のグループングをおこなうことである。上記、ベルクソンの工夫で、眞の値 β の變化の範圍をいくつかの區間に分割し、當該區間の中位數をもつて β を代表する。たとえば、 β と $\beta + 10$ のあいだの代表値として $\beta + 5$ をもちいるといふふうだ。このとき β が誤差をふくまず、各區間において β がそれぞれの代表値を中心して對象的に分布しているものとすれば、グループングにともなう各區間の誤差はゼロの期待値をもつて、假設(iv)を充すものと考へられる。それゆゑ、

$$b = \frac{\sum_{k=1}^n x_k y_k}{\sum_{k=1}^n x_k^2}$$

は β の不偏推定値となるであろう。

ところで、われわれのとりあつかいする資料は、誤差 ϵ_k をともなつた觀察値 y_k である。いま、グループピングがこれらの誤差によつて影響をうけることがない前提すれば、各グループの中心値は觀察誤差にかんして獨立になる。したがつて、觀察誤差が眞の値 β と相關關係を有せず、その平均値がゼロであるとすれば、 ϵ_k を觀察誤差と \wedge グループピングされた區間を中心値によつて代表させることから生ずる誤差 \wedge との複合された誤差と考えることによつて、假設(iv)を前提しうる。それゆゑ、右の諸條件がげんみつに充されれば、回歸係數の不偏推定値をうることができるであろう。またたとえ、假設(iv)を規定しうるための諸條件がげんみつには充されなくても、グループピングによつて、すくなくとも β の偏りを減少させることはできるものと思われる。上述の議論は、單に二變數のばあいばかりでなく、一般に多邊回歸のばあいにもそのまま擴張しうるものである。

- (7) たとえば J. C. Koopmans, *Linear Regression Analysis of Economic Time Series*, 1937 を見よ。
- (8) 時系列では、このほかに自己相關誤差の問題が生ずるが、ここでは差當りその問題にかんする假設を必要としなう。
- (9) J. Berkson, "Are There Two Regressions?" *Journal of American Statistical Association*, June 1950.
- (10) 區間の代表値に平均値をとれば、グループピング誤差がもたらす偏りはさけることができるが、變數の觀察誤差にもとづく偏りはさけることができなう。

四

觀察誤差が假設(i)(ii)に規定されるような性質をもつばあい、回歸係數の推定値の偏りを消滅あるいは減少させる

手段として、補助變數 instrumental variable をもちいれればよいことが、さいきんライアジールやゲアリーの研究によつて知られてゐる。¹¹⁾ この節では、そのような工夫について考えてみよう。

獨立變數の眞の値 β とは相關關係をもつけれども、誤差 ϵ や ϵ とは獨立なある變數 ϵ を想定する。このような ϵ を補助變數とよぶ。¹²⁾

$$\beta = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i} = \frac{\sum_{i=1}^n x_i (\beta x_i + \epsilon_i)}{\sum_{i=1}^n x_i}$$

とおけば

$$B(\beta) = \frac{\beta E(x_i^2)}{E(x_i^2)} = \beta$$

となるから、 N は β の不偏推定値たりうることになる。¹³⁾

ところでこのばあい、一般に ϵ の分散 A (q) の分散 A (q) よりも大きいから、推定値として β をもちいることはそれだけ精度がわるくなる。それがどの程度であるかをみるために、兩變數が觀察誤差をとまわないて測定されたばあい、すなわち $\epsilon = \epsilon'$ なるときの相對精度を算定してみると、

$$\frac{x_i^2 \sum_{i=1}^n \epsilon_i^2}{\sum_{i=1}^n \epsilon_i^2} = (q)A \quad \text{ただし } \epsilon = \epsilon' \text{ なるとき}$$

$$\frac{s(x_i^2 \sum_{i=1}^n \epsilon_i^2)}{s(\sum_{i=1}^n \epsilon_i^2)} = (\beta)A$$

から

$$\frac{(\beta)A}{(q)A} = \frac{s(x_i^2 \sum_{i=1}^n \epsilon_i^2)}{s(\sum_{i=1}^n \epsilon_i^2)}$$

をうる。

このように、補助變數をもちいることによって精度がわるくなるとすると、われわれはしいて補助變數をもちいて不偏推定値をえるよりも、むしろふつうの最小二乗法によって△偏りをふくむ推定値√に満足した方がよいかもれない。それではこの兩者の差は、どのようなひらがりをもつてであろうか。

すでに指摘したように、 $|E(b) - \Delta b|$ 、 $|E(b') - \Delta b'|$ であるから、 $|b - \Delta b|$ がえられる。そこで、前と同様に $b = y'$ のときを考えると、

$$b - b' = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} - \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} = 0$$

したがって、

$$E(b - b') = 0$$

$$V(b - b') = \frac{\sigma^2 \sum_{i=1}^n \left(\frac{x_i}{\sum_{i=1}^n x_i^2} - \frac{x_i}{\sum_{i=1}^n x_i^2} \right)^2}{\sum_{i=1}^n x_i^2} \left(\frac{V(A)}{V(B)} - 1 \right)$$

となる。すなわち b と b' の差の分散は、兩者の分散比に依存する。(ただし、兩變數の觀察誤差はゼロとして)。

さて、現實に補助變數法をもちいるはあい問題となるのは、 x' とは高度な相關關係をもちながら、 x'' とは獨立な變數をどうしてみつづけるかということである。これにもさまざまな工夫を見出しうる。たとえばワルドがやったとつたえられる(筆者未見) ようだ。

$$z = +1 \text{ or } 0 \text{ or } -1 \quad \text{for } x \text{ its median}$$

- (3) A. Wald, The Fitting of Straight Lines if Both Variables are Subject to Error. *Annals of Mathematical Statistics*, No. 3, 1940
- (14) ライプシールは前掲一九四一年の論文において、 x_t が系列相関を有し、 y_t が系列相関をなさない場合には、 y_t の lagged values を補助變數としてゐる。

五

戦後、モデル・アナリシスの發達にもなつて、經濟關係の計量にもいままでの單一方程式による最小二乗法的分析にかわり、いわゆる連立方程式接近法が採用されるようになったことはよく知られている。さいごに、連立方程式接近法と補助變數法との關連についてかんたんにふれておこう。

y_t と x_t との回歸式と、 x_t と z_t との回歸式とを連立させれば

$$\begin{cases} y_t = \beta_1 x_t + \varepsilon_t \\ x_t = \gamma_1 z_t + \eta_t \end{cases}$$

となるから

$$\begin{cases} y_t = \beta_1 x_t + \varepsilon_t \\ x_t = \gamma_1 z_t + \eta_t \end{cases}$$

ただし

$$\begin{cases} \varepsilon_t = e_t - \beta_1 \varepsilon_t \\ \eta_t = \gamma_1 \eta_t + \xi_t \end{cases}$$

ここで、誤差項 e_t と ξ_t が z_t にかんして獨立であり、その平均値がゼロであるとすれば、 γ_1 がゼロでないかぎり、誘導形法による β_1 の推定値として、

$$\beta_1 = \frac{\sum z_t y_t}{\sum z_t x_t}$$

をうる。このばあい、 α と β が同時正規分布をなすとすれば、 β が β の最尤推定値になることは前稿で証明したが、
同時分布の性質に若干の限定をつければ、 β が不偏推定値になることもアンダーソンとラビンによってすでに證明
されてゐる。¹⁷⁾

(5) T. C. Koopmans and Others, *Statistical Inference in Dynamic Economic Models*, 1950.

拙稿「經濟關係の計量とその推計學的基礎」經濟論叢 昭和二十五年九月

(6) 前掲拙稿一一三頁。

(7) T. W. Anderson and H. Rubin, "The Asymptotic Properties of Estimates of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics*, March, 1950.