論　文

Vuong test とその正規線形モデルへの適用法*

松　尾　精　彦

概　要

モデル選択のための検定（tests for model selection）として近年よく用いられるようになってきた，Vuong(1989) 検定の理論と応用について考察する．Vuong 検定は，Kullback-Leibler 情報量を基準として，２つの競合するモデルを選択するための検定である．Vuong 検定とよく似た考え方に，AIC（Akaike Information Criterion）を始めとするモデル選択基準や，Cox(1961) を起源とする Non-nested 検定がある．AIC を始めとするモデル選択基準は予測に用いるモデルを選択するための指標であり，Non-nested 検定は，２つのモデルのうち，一方のモデルが真の分布を含むことを帰無仮説とし他方のモデルが真の分布を含むことを対立仮説とする検定である．一方，Vuong 検定は，２つのモデルのうち，どちらのモデルがより真の分布に近い分布を含むかを検定するものである．

この“どちらのモデルがより真のデータ生成過程に近い分布を有しているかを検定する”という発想が，従来の検定と一線を画すところである．この現実的な発想は，従来の手法よりも粗い結論を導くが，その代わりに一般性を持つ議論を可能にしている．この論文では，従来からの Non-nested 検定やモデル選択基準との比較を通して Vuong 検定の独自性を明らかにし，最も重要な応用例である正規線形回帰モデルへの適用時における新たな運用法を提案する．

キーワード：Tests for Model selection; Kullback-Leibler Information; Non-nested Test.
経済学文献季報分類番号：16-10

1　Introduction

モデル選択基準とは別に，モデル選択検定という決定論的アプローチがある．モデル選択検定のための枠組みとして，近年よく用いられるようになってきたのが Vuong 検定である．モデル選択基準ではモデルを予測に用いる際の振る舞いの良さを基準にしているのに対し (例えば，松尾 (2003) を参照されたい)，Vuong 検定では真の分布とモデルとの距離を基準としている．分布間の距離を Kullback-Leibler 情報量で測ることは AIC と同じだが，AIC がモデル内の推測された分布と真の分布との距離を考えているのに対し，Vuong 検定ではモデル内の分布と真の分布

* この研究は平成14年度関西大学研修員規定によって行った研究の一部である．なおこの成果の主要部分は，後日英訳し，他雑誌に投稿する予定である．武蔵大学経済学部・太田浩司専任講師には，貴重な助言を受けたことに，感謝の意を表する．

との最短距離を扱っていることに注目されたい．また，AIC の理論では，比較するモデルが真の分布を近似的に含んでいることが要請されるが，Vuong 検定ではそのような要請がないことも興味深い事実である．

いま 2 つの競合するモデルを，回帰モデルへの適用を許容するよう，説明変数（外生あるいは独立変数とも言う）が与えられたときの条件付モデルの形で，$F_{\boldsymbol{\theta}} = \{f(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\theta});\boldsymbol{\theta} \in \Theta\}$，$G_{\boldsymbol{\phi}} = \{g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\phi});\boldsymbol{\phi} \in \boldsymbol{\Phi}\}$ と表し，真の条件付分布を $h^0(\boldsymbol{y}|\boldsymbol{z})$ とする．このとき，まずモデル間に包含関係があるか否かにより，検定手続きが違ってくることはよく知られている．もっとも扱いやすいのが，包含関係がある場合である．この場合，包含されるモデルを帰無仮説，包含するモデルを対立仮説とすれば，適当な正則条件の下で条件付対数尤度比の 2 倍が漸近的にカイ 2 乗分布に従うことを利用した検定が行われる．形式的には $F_{\boldsymbol{\theta}} \subset G_{\boldsymbol{\phi}}$ ならば，$H_0 : F_{\boldsymbol{\theta}} \ni h^0(\boldsymbol{y}|\boldsymbol{z})$，$H_1 : G_{\boldsymbol{\phi}} \setminus F_{\boldsymbol{\theta}} \ni h^0(\boldsymbol{y}|\boldsymbol{z})$ という検定を行うのである．

しかし，包含関係が無い場合 (Non-nested case) には，議論が単純ではなくなる．どちらか一方のモデルが真の分布を含むことを帰無仮説とし，他方のモデルが真の分布を含むことを対立仮説とする検定は Non-nested 検定と呼ばれ，数多くの文献が公表されてきている（例えば Pesaran and Weeks, 2001 を参照されたい）．どちらか一方のモデルに優位性が認められない場合には，それぞれのモデルが真の分布を含むことを帰無仮説として，2 度の検定を行わなかればならない．この場合，条件付対数尤度比を 2 倍したものの漸近分布は，もはやカイ 2 乗分布には従わなくなる．モデルが真の分布を含むという仮説設定においては，尤度理論という一般的な枠組みでなく，個々のモデルの性質に依存したより精密な議論がなされてきた．その議論の中心は正規線形回帰モデルであり，Cox(1961) 以降いくつかの提案がされてきたが，どれも満足できるものではなかった．

このような状況下，Vuong (1989) は，モデル選択検定に対する統一的な枠組みを提供した．彼は，尤度比を検定統計量として用いた検定を，きわめて一般的な枠組みの下で提案したのである．非常に大きな一般化は，真の条件付分布は 2 つのモデルに含まれていても良いし，どちらか一方でも，更にはどちらにも含まれていなくてよいということである．また，2 つのモデルに包含関係があっても無くても良いし，共通部分があっても無くてもよい．その下で，どちらのモデルがより真の分布に近いモデルを包含するかを検定するのである．帰無仮説を，“両モデルと真の分布との距離は等しい” と設定したことは実に巧妙で，枠組みの一般性とうまく合致している．彼は，説明変数をも確率変数として扱うことで条件付での議論を避け，比較的簡潔な結論を導き出していることにも注目されたい．モデルがどちらも真の分布を含まないという前提は，これまでの検定の考え方からすれば奇異に感じるかもしれないが，どちらかと言えばモデルが真の分布を含むという方が却って不自然なのである．モデルが真の分布を含むという仮定は，検定統計量

の分布を既知のカイ自乗分布に帰着させるためのものであって，数学的な都合によりおかれたものなのである.

“わざわざこのような検定を導入しなくても, AIC や C_p などのモデル選択基準を用いれば十分ではないか” という指摘もなされるだろう. しかしながら先に述べたように，AIC は真の分布が, 少なくとも近似的に, モデルに含まれることを前提に議論している. 真の分布を含まないモデルは最大尤度の部分が小さくなり, AIC の値が小さくなることから自然に脱落するという解釈 (竹内, 1976) が可能なものの, モデルが真の分布を含まないときの利用については理論的裏づけを持たない. 一方, Vuong 検定では, モデルが真のモデルを含まない場面でも理論が破綻しない. また C_p は, モデルが真の分布を含むことを前提としないが, 正規線形モデルでしか利用できないという制約がある.

また, さらに, “どちらのモデルがより真の分布に近いモデルを包含するか” を知って何になるのかという問題提起もなされるだろう. しかし, 株式のリターンを分析する際や, 各自治体の環境政策を分析する際などには, このような情報も一定の価値を持ち得る. なぜならば, このような分析では, 利用可能な説明変数の数が莫大であるため探索的にモデルを構築することが困難であることに加え, ある程度理論に基づくモデルが対立的に提案されているのが普通である. 例えば株式リターン分析では, “キャッシュフロー” を基にするか “会計数値” を用いるかの問題がある. どちらでモデル化しても説明力にはさほどの差は出ないし, そもそも説明力自体がかなり低い. このような場合には, AIC を利用するよりもむしろ,Vuong 検定を用いてどちらのモデルが有意に良好であるかを検定することには一定の意味があると言えよう. このように, 2 個に絞り込まれたモデル間の優劣を決定する場面では, Vuong 検定の特徴が活かされるのである.

この論文では, Vuong 検定を正規線形回帰モデルに適用し, 正規線形回帰モデルに適した運用法を提案する. また, よく似たアプローチである Non-nested 検定やモデル選択基準との比較を, 正規線形回帰モデルの中で議論する. 第 2 節では, Vuong 検定を紹介するための準備を行う. 第 3 節では, Vuong 検定の理論の大まかな紹介を行う. 第 4 節では, 2 つの正規線形回帰モデル間の Vuong 検定の運用方法についての新たな提案を行う. 第 5 節では, よく似た考え方である Non-nested 検定の紹介を通して, Vuong 検定の独自性について議論する. 第 6 節では, AIC との比較を通して, Vuong 検定の利用方法について議論する.

2 BASIC FRAMEWORK

この先の議論で前提とするのは，Vuong (1989) が挙げたような正則条件が成り立っているということである. つまり, 微分可能性や期待値の存在, 微分と積分の交換可能性, 漸近分布の収束などは成り立つものとし，細かな条件設定については踏み込まず, 大まかな議論を展開する.

この節と次の節は, Vuong (1989) の内容を要約したものである. なお, $\boldsymbol{X}'$ を $\boldsymbol{X}$ の転置をあらわすものとするとき, Davidson and MacKinnon (1993) に倣い, $(\boldsymbol{Y}\vdots\boldsymbol{Z})$ という記法を導入し $(\boldsymbol{Y}',\boldsymbol{Z}')'$ を表すものとする.

$\boldsymbol{X}_t, t=1,2,\cdots$ は m 次元確率変数列で, 独立同分布に従うものとする. $\boldsymbol{X}_t=(\boldsymbol{Y}_t\vdots\boldsymbol{Z}_t)$ と分割され, $\boldsymbol{Y}_t$ は l 次元被説明 (従属, 内生とも言う) 変数, $\boldsymbol{Z}_t$ は k 次元説明 (独立, 外生とも言う) 変数とする. 各 $\boldsymbol{X}_t$ の密度関数を $h^0(\cdot)$, $\boldsymbol{Z}_t$ の真の周辺密度関数を $h^0_{\boldsymbol{Z}}(\cdot)$, $\boldsymbol{Z}=\boldsymbol{z}$ が与えられたときの真の条件付密度関数を $h^0(\cdot|\boldsymbol{z})$ で表すことにする.

このとき, 2 つの競合するモデル, $F_{\boldsymbol{\theta}}=\{f(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\theta});\boldsymbol{\theta}\in\boldsymbol{\Theta}\}$, $G_{\boldsymbol{\phi}}=\{g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\phi});\boldsymbol{\phi}\in\boldsymbol{\Phi}\}$ についての推測を行うものとする. ここで, パラメータ空間は $\boldsymbol{\Theta}$, $\boldsymbol{\Phi}$ は文脈に応じて規定される. 例えば, $F_{\boldsymbol{\theta}}$, $G_{\boldsymbol{\phi}}$ を 2 つの正規線形回帰モデルとするときには, $\boldsymbol{z}$ は利用可能なすべての説明変数であり, $\boldsymbol{\Theta}$, $\boldsymbol{\Phi}$ は各モデルに採用される説明変数に応じた母数空間として解釈される. $\boldsymbol{\theta}_*, \boldsymbol{\phi}_*$ を, それぞれ, モデル $F_{\boldsymbol{\theta}}, G_{\boldsymbol{\phi}}$ の pseudo true value として,

(1) $$\boldsymbol{\theta}_*=\arg\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} E^0_{\boldsymbol{X}}[\ln f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta})],\quad \boldsymbol{\phi}_*=\arg\max_{\boldsymbol{\phi}\in\boldsymbol{\Phi}} E^0_{\boldsymbol{X}}[\ln g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi})]$$

と定義する. ここで, $E^0_{\boldsymbol{X}}$ は $\boldsymbol{X}$ の真の分布 $h^0(\cdot)$ に関する期待値を表している. また,

(2) $$\begin{cases} A_f(\boldsymbol{\theta})\equiv E^0_{\boldsymbol{X}}\left[\dfrac{\partial^2\ln f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right], \\ B_f(\boldsymbol{\theta})\equiv E^0_{\boldsymbol{X}}\left[\dfrac{\partial\ln f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\dfrac{\partial\ln f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}\right], \\ A_g(\boldsymbol{\phi})\equiv E^0_{\boldsymbol{X}}\left[\dfrac{\partial^2\ln g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi})}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}\right], \\ B_g(\boldsymbol{\phi})\equiv E^0_{\boldsymbol{X}}\left[\dfrac{\partial\ln g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi})}{\partial\boldsymbol{\phi}}\dfrac{\partial\ln g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi})}{\partial\boldsymbol{\phi}'}\right], \\ B_{fg}(\boldsymbol{\theta},\boldsymbol{\phi})\equiv E^0_{\boldsymbol{X}}\left[\dfrac{\partial\ln f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\dfrac{\partial\ln g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi})}{\partial\boldsymbol{\phi}'}\right], B_{gf}(\boldsymbol{\theta},\boldsymbol{\phi})=B_{fg}(\boldsymbol{\theta},\boldsymbol{\phi})' \end{cases}$$

とおく. $\hat{\boldsymbol{\theta}}_n$, $\hat{\boldsymbol{\phi}}_n$ を, それぞれ, モデル $F_{\boldsymbol{\theta}}$, $G_{\boldsymbol{\phi}}$ の最尤推定量を表すものとし,

(3) $$L^f_n(\hat{\boldsymbol{\theta}}_n)=\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} L^f_n(\boldsymbol{\theta}),\quad \text{where } L^f_n(\boldsymbol{\theta})\equiv\sum_{t=1}^n \ln f(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\boldsymbol{\theta})$$

(4) $$L^g_n(\hat{\boldsymbol{\phi}}_n)=\sup_{\boldsymbol{\phi}\in\boldsymbol{\Phi}} L^f_n(\boldsymbol{\phi}),\quad \text{where } L^g_n(\boldsymbol{\phi})\equiv\sum_{t=1}^n \ln g(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\boldsymbol{\phi})$$

とする. このとき, 適当な正則条件の下で, $\hat{\boldsymbol{\theta}}_n$ と $\hat{\boldsymbol{\phi}}_n$ は, それぞれ $\boldsymbol{\theta}_*$ と $\boldsymbol{\phi}_*$ の一致推定量であり, $\sqrt{n}(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}_*)$ と $\sqrt{n}(\hat{\boldsymbol{\phi}}_n-\boldsymbol{\phi}_*)$ は, それぞれ, 漸近的に正規分布

$$N\Big[\boldsymbol{0}, A_f(\boldsymbol{\theta}_*)^{-1}B_f(\boldsymbol{\theta}_*)A_f(\boldsymbol{\theta}_*)^{-1}\Big],\quad N\Big[\boldsymbol{0}, A_g(\boldsymbol{\phi}_*)^{-1}B_g(\boldsymbol{\phi}_*)A_g(\boldsymbol{\phi}_*)^{-1}\Big]$$

に従う. 実際, $\hat{\boldsymbol{\theta}}_n$ と $\hat{\boldsymbol{\phi}}_n$ の同時分布もまた漸近正規であり,

$$\text{(5)} \qquad \sqrt{n}\begin{pmatrix}\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}_*\\ \hat{\boldsymbol{\phi}}_n-\boldsymbol{\phi}_*\end{pmatrix} \asymp N\left[\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix}, \Sigma_{(\boldsymbol{\theta}_*,\boldsymbol{\phi}_*)}\right],$$

$$\Sigma_{(\boldsymbol{\theta}_*,\boldsymbol{\phi}_*)} = \begin{pmatrix} A_f^{-1}(\boldsymbol{\theta}_*)B_f(\boldsymbol{\theta}_*)A_f^{-1}(\boldsymbol{\theta}_*) & A_f^{-1}(\boldsymbol{\theta}_*)B_{fg}(\boldsymbol{\theta}_*,\boldsymbol{\phi}_*)A_g^{-1}(\boldsymbol{\phi}_*)\\ A_g^{-1}(\boldsymbol{\phi}_*)B_{gf}(\boldsymbol{\theta}_*,\boldsymbol{\phi}_*)A_f^{-1}(\boldsymbol{\theta}_*) & A_g^{-1}(\boldsymbol{\phi}_*)B_g(\boldsymbol{\phi}_*)A_g^{-1}(\boldsymbol{\phi}_*)\end{pmatrix},$$

となる.

3 VUONG TEST

真の分布, $h^0(\boldsymbol{y}|\boldsymbol{z})$ と モデル $F_{\boldsymbol{\theta}}$, $G_{\boldsymbol{\phi}}$ との距離を Kullback-Leibler 情報量,

$$\text{(6)} \qquad \begin{aligned} I\big(h^0(\boldsymbol{y}|\boldsymbol{z}), f(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\theta}_*)\big) &= E^0_{\boldsymbol{X}}[\ln h^0(\boldsymbol{Y}|\boldsymbol{Z})] - E^0_{\boldsymbol{X}}[\ln f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)]\\ I\big(h^0(\boldsymbol{y}|\boldsymbol{z}), g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\phi}_*)\big) &= E^0_{\boldsymbol{X}}[\ln h^0(\boldsymbol{Y}|\boldsymbol{Z})] - E^0_{\boldsymbol{X}}[\ln g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)]\end{aligned}$$

で測定することにする. 第 1 項は両モデルに共通だから, “どちらのモデルがより真の分布に近い分布を含むか” という問題は, 次のような仮説に言い換えられる. 帰無仮説,

$$\text{(7)} \qquad H_0^V : E^0_{\boldsymbol{X}}[\ln f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)] = E^0_{\boldsymbol{X}}[\ln g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)],$$

にたいし,

$$\text{(8)} \qquad H_f^V : E^0_{\boldsymbol{X}}[\ln f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)] > E^0_{\boldsymbol{X}}[\ln g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)],$$

あるいは,

$$\text{(9)} \qquad H_g^V : E^0_{\boldsymbol{X}}[\ln f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)] < E^0_{\boldsymbol{X}}[\ln g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)],$$

を対立仮説とする検定を行う. もし仮に, H_0^V が棄却され H_f^V が採択される場合,

$$I\big(h^0(\boldsymbol{y}|\boldsymbol{z}), g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\phi}_*)\big) > I\big(h^0(\boldsymbol{y}|\boldsymbol{z}), f(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\theta}_*)\big) \geq 0$$

という関係が成り立ち $G_{\boldsymbol{\phi}}$ は真の分布を含み得ない. また, H_0^V の下では, $f(\cdot|\cdot;\boldsymbol{\theta}_*) \neq g(\cdot|\cdot;\boldsymbol{\phi}_*)$ ならば, 2 つのモデルはどちらも真の分布を含まないことに注意されたい.

検定統計量は, 対数尤度比,

$$\text{(10)} \qquad LR_n(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\phi}}_n) \equiv L_n^f(\hat{\boldsymbol{\theta}}_n) - L_n^g(\hat{\boldsymbol{\phi}}_n) = \sum_{t=1}^{n} \ln \frac{f(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\hat{\boldsymbol{\theta}}_n)}{g(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\hat{\boldsymbol{\phi}}_n)}$$

が用いられる. 尤度関数の連続性やモーメントについての適当な条件下で, 大数の法則より,

$$\text{(11)} \qquad \frac{1}{n}LR_n(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\phi}}_n) \to E^0_{\boldsymbol{X}}\left[\ln \frac{f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)}{g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)}\right] \text{ a.s.},$$

が成立することが今後用いられる．また，$\ln \frac{f(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\boldsymbol{\theta}_*)}{g(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\boldsymbol{\phi}_*)}$ の分散を，

$$\begin{aligned}\omega_*^2 &\equiv \mathrm{var}_{\boldsymbol{X}}^0\left[\ln \frac{f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)}{g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)}\right]\\ &= E_{\boldsymbol{X}}^0\left[\ln \frac{f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)}{g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)}\right]^2 - \left[E_{\boldsymbol{X}}^0\left[\ln \frac{f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)}{g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)}\right]\right]^2\end{aligned} \tag{12}$$

で表すことにする．従来の尤度理論と違い，モデルが真の分布を含んでいることを前提としないことから，次の重み付きカイ自乗和分布を導入する必要が生じる．

Definition 1. *(重み付きカイ２乗和分布):* $Z_1, Z_2, \cdots, Z_m$ を独立で標準正規分布に従う確率変数列とし，$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \cdots, \lambda_m)$ を m 個の実数からなるベクトルとする．このとき，確率変数 $\sum_{i=1}^m \lambda_i Z_i^2$ の従う分布を，重み付きカイ２乗和分布と呼びその累積分布関数を $M_m(\cdot, \boldsymbol{\lambda})$ で表す．

Theorem 1. *(尤度比検定統計量の漸近分布):* 適当な正則条件下で，

(i) $f(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\theta}_*) = g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\phi}_*)$ のとき，

$$2LR_n(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\phi}}_n) \to M_{p+q}(\cdot;\boldsymbol{\lambda}_*) \textit{ in law},$$

が成り立つ．ここで $\boldsymbol{\lambda}_*$ は，

$$W = \begin{bmatrix} -B_f(\boldsymbol{\theta}_*)A_f^{-1}(\boldsymbol{\theta}_*) & -B_{fg}(\boldsymbol{\theta}_*,\boldsymbol{\phi}_*)A_g^{-1}(\boldsymbol{\phi}_*) \\ B_{gf}(\boldsymbol{\phi}_*,\boldsymbol{\theta}_*)A_f^{-1}(\boldsymbol{\theta}_*) & B_g(\boldsymbol{\phi}_*)A_g^{-1}(\boldsymbol{\phi}_*) \end{bmatrix},$$

の固有値からなるベクトルである．

(ii) $f(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\theta}_*) \neq g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\phi}_*)$ のとき，

$$n^{-1/2}LR_n(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\phi}}_n) - n^{1/2}E_{\boldsymbol{X}}^0\left[\ln \frac{f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)}{g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)}\right] \to N(0,\omega_*^2) \textit{ in law},$$

が成り立つ．

この定理の大まかな証明をしよう．pseudo true value における尤度関数を，最尤推定値の周りで Taylor 展開して，

$$L_n^f(\boldsymbol{\theta}_*) = L_n^f(\hat{\boldsymbol{\theta}}_n) + \frac{n}{2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_*)'A_f(\boldsymbol{\theta}_*)(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_*) + o_p(1).$$
$$L_n^g(\boldsymbol{\phi}_*) = L_n^g(\hat{\boldsymbol{\phi}}_n) + \frac{n}{2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_*)'A_g(\boldsymbol{\phi}_*)(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_*) + o_p(1).$$

を得る．これら２式の差より，

$$\begin{aligned}LR_n(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\phi}}_n) =& LR_n(\boldsymbol{\theta}_*, \boldsymbol{\phi}_*) - \frac{n}{2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_*)'A_f(\boldsymbol{\theta}_*)(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_*)\\ &+ \frac{n}{2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_*)'A_g(\boldsymbol{\phi}_*)(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_*) + o_p(1),\end{aligned} \tag{13}$$

が導かれる．(i) の場合，$f(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\theta}_*) = g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\phi}_*)$ だから，$LR_n(\boldsymbol{\theta}_*, \boldsymbol{\phi}_*) = 0$ が成り立ち，

$$LR_n(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\phi}}_n) = \frac{n}{2}\begin{pmatrix}\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_* \\ \hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_*\end{pmatrix}' \begin{bmatrix} -A_f & 0 \\ 0 & A_g \end{bmatrix} \begin{pmatrix}\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_* \\ \hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_*\end{pmatrix} + o_p(1)$$

となる．$(\hat{\boldsymbol{\theta}}_n \vdots \hat{\boldsymbol{\phi}}_n)$ の漸近分布は (5) であることより *(i)* が示される．

一方 *(ii)* は, (13) の両辺から $nE^0_{\boldsymbol{X}}[\ln \dfrac{f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)}{g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)}]$ を引き, $\sqrt{n}$ で割ることにより,

$$
\begin{aligned}
\text{(14)}\qquad & n^{-1/2}LR_n(\hat{\boldsymbol{\theta}}_n,\hat{\boldsymbol{\phi}}_n) - n^{1/2}E^0_{\boldsymbol{X}}[\ln \frac{f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)}{g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)}] \\
&= n^{1/2}\Big[\frac{1}{n}LR_n(\boldsymbol{\theta}_*,\boldsymbol{\phi}_*) - E^0_{\boldsymbol{X}}[\ln \frac{f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)}{g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)}]\Big] + o_p(1).
\end{aligned}
$$

を得るが, 右辺は, 中心極限定理を適用することにより, $N(0,\omega^2_*)$ に漸近的に従うことから示される．

この定理は, Vuong (1989) が紹介したモデル選択検定の枠組みの中で, 中心的な役割を果たすものであり, (i) については適用上困難さが付きまとうが, (ii) はどのようなモデルに対しても簡単に適用できるという利点を持つ．

3.1 Strictly Non-nested Models

Definition 2. *(Strictly Non-Nested Models): 2* つの条件付モデル $F_{\boldsymbol{\theta}}$ と $G_{\boldsymbol{\phi}}$ は,

$$F_{\boldsymbol{\theta}} \cap G_{\boldsymbol{\phi}} = \emptyset.$$

であるとき *strictly non-nested* であると言う．

Example 1. *(Cox, 1961):* 対数正規分布と指数分布．

$$
\begin{cases}
f(y|\theta) = y^{-1}(2\pi\theta_2)^{-1/2}\exp\Big\{-\dfrac{(\ln y-\theta_1)^2}{2\theta_2}\Big\},\ 0<\theta_2<\infty,\ y>0. \\
g(y|\phi) = \phi^{-1}\exp(-y/\phi), \phi>0, y>0.
\end{cases}
$$

他にも, 対数正規分布とワイブル分布, また, 対数正規分布とガンマ分布についても議論がなされている．

Example 2. PROBIT *vs* LOGIT *model.*

$$
\begin{cases}
\Pr(Y=1) = \Phi(\boldsymbol{\theta}'\boldsymbol{z}) = \displaystyle\int_{-\infty}^{\boldsymbol{\theta}'\boldsymbol{z}} \frac{1}{\sqrt{2\pi}}\exp\Big\{-\frac{1}{2}u^2\Big\}du, \\
\Pr(Y=1) = \Delta(\boldsymbol{\phi}'\boldsymbol{z}) = \dfrac{e^{\boldsymbol{\phi}'\boldsymbol{z}}}{1+e^{\boldsymbol{\phi}'\boldsymbol{z}}}.
\end{cases}
$$

以上のように, strictly non-nested case の定義を厳密に適用するなら, 2 つのモデルの誤差分布が互いに異なる場合しかない．それゆえ, 2 つの正規線形回帰モデルの場合には, “ある特定のパラメータが 0 にはならない” といった前提が更に必要になることに注意されたい．

さて strictly non-nested case では 2 つのモデルに共通な分布が存在しないため, $f(\cdot|\cdot;\boldsymbol{\theta}_*) \neq g(\cdot|\cdot;\boldsymbol{\phi}_*)$ であることが保障されるので, **Theorem 1** の (ii) を直接適用できる．この場合特に, 次の定理を得る．

Theorem 2. *(Model Selection Tests for Strictly Non-Nested Models):* 適当な正則条件の下で, $F_{\boldsymbol{\theta}}$ と $G_{\boldsymbol{\phi}}$ が *strictly non-nested* ならば, 次が成り立つ.

$$(i)\ under\ H_0^{\mathrm{V}}: n^{-1/2}LR_n(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\phi}}_n)/\hat{\omega}_n \to N(0,1)\ in\ law\ ,$$

$$(ii)\ under\ H_f^{\mathrm{V}}: n^{-1/2}LR_n(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\phi}}_n)/\hat{\omega}_n \to +\infty\ a.s.\ ,$$

$$(iii)\ under\ H_g^{\mathrm{V}}: n^{-1/2}LR_n(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\phi}}_n)/\hat{\omega}_n \to -\infty\ a.s.\ .$$

上の定理の ω_*^2 の推定量として, Vuong (1989) が列挙している漸近同値なものの中で, 例えば,

$$\hat{\omega}_n^2 \equiv \frac{1}{n}\sum_{t=1}^{n}\left[\ln\frac{f(\boldsymbol{y}_t|\boldsymbol{z}_t;\hat{\boldsymbol{\theta}}_n)}{g(\boldsymbol{y}_t|\boldsymbol{z}_t;\hat{\boldsymbol{\phi}}_n)}\right]^2 - \left[\frac{1}{n}\sum_{t=1}^{n}\ln\frac{f(\boldsymbol{y}_t|\boldsymbol{z}_t;\hat{\boldsymbol{\theta}}_n)}{g(\boldsymbol{y}_t|\boldsymbol{z}_t;\hat{\boldsymbol{\phi}}_n)}\right]^2$$

を用いればよいだろう.

3.2 Overlapping Models

Definition 3. *(Overlapping Models): 2* つの条件付モデル $F_{\boldsymbol{\theta}}$ と $G_{\boldsymbol{\phi}}$ が *overlapping* であるとは,

$$(i)\ F_{\boldsymbol{\theta}} \cap G_{\boldsymbol{\phi}} \neq \emptyset,$$

$$(ii)\ F_{\boldsymbol{\theta}} \setminus G_{\boldsymbol{\phi}} \neq \emptyset\ and\ G_{\boldsymbol{\phi}} \setminus F_{\boldsymbol{\theta}} \neq \emptyset.$$

であるときを言う.

Example 3. *2* つの線形回帰モデルで,

$$\begin{cases} f(y|\boldsymbol{z};\boldsymbol{\theta}) = \dfrac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\dfrac{(y-\alpha-\boldsymbol{u}'\boldsymbol{\beta}-\boldsymbol{v}_f'\gamma_f)^2}{2\sigma^2}\right\}, \\ g(y|\boldsymbol{z};\boldsymbol{\phi}) = \dfrac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\dfrac{(y-\alpha-\boldsymbol{u}'\boldsymbol{\beta}-\boldsymbol{v}_g'\gamma_g)^2}{2\sigma^2}\right\}, \end{cases}$$

ここで $\boldsymbol{u}$, $\boldsymbol{v}_f$, $\boldsymbol{v}_g$ は, それぞれ p, q_f, q_g 次元確率変数ベクトルで, $(\boldsymbol{u}\vdots\boldsymbol{v}_f\vdots\boldsymbol{v}_g)$ は退化しない $p+q_f+q_g$ 次元確率変数ベクトルであるとする. ここで, $(\boldsymbol{u}\vdots\boldsymbol{v}_f\vdots\boldsymbol{v}_g)$ はベクトルを縦に並べたもの, つまり $(\boldsymbol{u}', \boldsymbol{v}_f', \boldsymbol{v}_g')'$ を表すものとする. ここでは正規線形回帰モデルを考えたが, 他にも, 一般化線形モデルを始めとする回帰モデルに広く例を求めることが出来る.

この場合は, strictly non-nested case よりも扱いが厄介になる. なぜなら 2 つのモデルに共通部分があるため, モデル選択検定を行う際には, 先ず **Theorem 1** の *(i)* を適用して, $f(\cdot|\cdot;\boldsymbol{\theta}_*) = g(\cdot|\cdot;\boldsymbol{\phi}_*)$ であるか否かを確かめなければならない. $f(\cdot|\cdot;\boldsymbol{\theta}_*) = g(\cdot|\cdot;\boldsymbol{\phi}_*)$ と判定されれば, 帰無仮説 H_0^{V} が採択される. そうでなければ, 引き続き $f(\cdot|\cdot;\boldsymbol{\theta}_*) \neq g(\cdot|\cdot;\boldsymbol{\phi}_*)$ のもとで, strictly non-nested の場合と同様, **Theorem 1** の *(ii)* を適用するのである.

ここで注目してほしいのは, どちらかのモデルが真の分布を含むという古典的な仮定を置くならば, $f(\cdot|\cdot;\boldsymbol{\theta}_*)$ あるいは $g(\cdot|\cdot;\boldsymbol{\phi}_*)$ が真の分布 $h^0(\cdot|\cdot)$ に一致するので, $E_{\boldsymbol{X}}^0[\ln f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)] =$

$E^0_{\boldsymbol{X}}[\ln g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)]$ は $I\big(h(\boldsymbol{y}|\boldsymbol{z}), f(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\theta}_*)\big) = I\big(h(\boldsymbol{y}|\boldsymbol{z}), g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\phi}_*)\big) = 0$ を意味し, Kullback-Leibler 情報量の性質より, $h(\boldsymbol{y}|\boldsymbol{z}) = f(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\theta}_*) = g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\phi}_*)$ という関係が得られる. つまり,

$$f(\cdot|\cdot;\boldsymbol{\theta}_*) = g(\cdot|\cdot;\boldsymbol{\phi}_*) \Leftrightarrow E^0_{\boldsymbol{X}}[\ln f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)] = E^0_{\boldsymbol{X}}[\ln g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)]$$

が成立する. それ故 $f(\cdot|\cdot;\boldsymbol{\theta}_*) \neq g(\cdot|\cdot;\boldsymbol{\phi}_*)$ ならば $E^0_{\boldsymbol{X}}[\ln f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)] \neq E^0_{\boldsymbol{X}}[\ln g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)]$ であることが導かれるので, **Theorem 1** の *(ii)* を適用する必要が無くなり, $f(\cdot|\cdot;\boldsymbol{\theta}_*) = g(\cdot|\cdot;\boldsymbol{\phi}_*)$ を帰無仮説とする検定を行いさえすればよいことになる. この場合の Vuong 検定は, 後に紹介する, Non-nested 検定を実行するための新たな手続きを与えていることに注目されたい.

3.3 Nested Models

Definition 4. *(Nested Models):* 条件付モデル $G_{\boldsymbol{\phi}}$ が条件付モデル $F_{\boldsymbol{\theta}}$ に *nest in* しているとは,

$$G_{\boldsymbol{\phi}} \subset F_{\boldsymbol{\theta}}.$$

であるときを言う.

Example 4. *2* つの正規線形回帰モデルで, 片方のモデルが採用している説明変数が他方のそれの部分集合であるとき.

$$\begin{cases} f(y|\boldsymbol{z};\boldsymbol{\theta}) = \dfrac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\dfrac{(y-\alpha-\boldsymbol{u}'\boldsymbol{\beta}-\boldsymbol{v}'\boldsymbol{\gamma})^2}{2\sigma^2}\right\}, \\ g(y|\boldsymbol{z};\boldsymbol{\phi}) = \dfrac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\dfrac{(y-\alpha-\boldsymbol{u}'\boldsymbol{\beta})^2}{2\sigma^2}\right\}, \end{cases}$$

ここで $\boldsymbol{u}, \boldsymbol{v}$ は, それぞれ p, q 次元確率変数ベクトルで, $(\boldsymbol{u}\vdots\boldsymbol{v}) = (\boldsymbol{u}', \boldsymbol{v}')'$ は退化しない $p+q$ 次元確率変数ベクトルであるとする. この場合も, 先の *overlapping case* と同様, 回帰モデル全般に例を求めることが出来る.

この場合, モデルの包含関係より $E^0_{\boldsymbol{X}}[\ln f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)] \geq E^0_{\boldsymbol{X}}[\ln g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)]$ が成り立つため, H^{V}_g は必然的に除外される. また, $E^0_{\boldsymbol{X}}[\ln f(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\theta}_*)] = E^0_{\boldsymbol{X}}[\ln g(\boldsymbol{Y}|\boldsymbol{Z};\boldsymbol{\phi}_*)]$ は $f(\cdot|\cdot;\boldsymbol{\theta}_*) = g(\cdot|\cdot;\boldsymbol{\phi}_*)$ を意味するので, **Theorem 1** の *(i)* を適用して, $f(\cdot|\cdot;\boldsymbol{\theta}_*) = g(\cdot|\cdot;\boldsymbol{\phi}_*)$ を検定しさえすればよいことが分かる.

これまで見てきたように, Vuong 検定とは, 従来のアプローチよりもはるかに一般的な枠組みの中で, モデル選択検定を尤度比検定統計量を用いて行う体系であると言える. この体系は, 従来, 個別の問題として扱われてきたいくつかの検定問題に対する新たな枠組みを提供している. この Voung 検定を, 最も利用頻度の高い正規線形回帰モデルに適用するにはどうすればよいのだろうか, また, よく似た概念であるモデル選択基準や Non-nested 検定との関係も興味深い. 以降の節では, これらのことについて議論してゆく.

4 TWO LINEAR REGRESSION MODELS

この節では, 2つの正規線形回帰モデルに対し, Vuong 検定を行い, どちらのモデルがより真の分布に近い分布を包含するかを検定する. 真の条件付分布, $h^0(y|\boldsymbol{z})$, を $N(\alpha_0+\boldsymbol{z}'\boldsymbol{\zeta}_0,\sigma_0^2)$ とし, モデル $G_{\boldsymbol{\phi}}$ を $N(\alpha+\boldsymbol{u}'\boldsymbol{\beta},\sigma^2)$ とするとき, $\boldsymbol{\phi}=(\alpha\vdots\boldsymbol{\beta}\vdots\sigma^2)$ の pseudo true value $\boldsymbol{\phi}_*=(\alpha_*\vdots\boldsymbol{\beta}_*\vdots\sigma_*^2)$ を計算しよう. $\boldsymbol{\phi}_*$ は,

$$
\begin{aligned}
E^0_{\boldsymbol{X}}[\ln g(Y|\boldsymbol{Z};\boldsymbol{\theta})] &= \int \ln g(y|\boldsymbol{z};\boldsymbol{\phi})h^0(\boldsymbol{x})d\boldsymbol{x}\\
&= \int h^0(\boldsymbol{z})\int\left\{-\frac{1}{2}\ln\sigma^2-\frac{(y-\alpha-\boldsymbol{u}'\boldsymbol{\beta})^2}{2\sigma^2}\right\}h(y|\boldsymbol{z})dyd\boldsymbol{z}-\frac{1}{2}\ln 2\pi
\end{aligned}
$$

を最大にするものとして与えられので,

$$
l_g(\alpha,\boldsymbol{\beta},\sigma^2)\equiv\frac{1}{2\sigma^2}\int h(\boldsymbol{z})\int(y-\alpha-\boldsymbol{u}'\boldsymbol{\beta})^2h^0(y|\boldsymbol{z})dyd\boldsymbol{z}+\frac{1}{2}\ln\sigma^2
$$

を最小化するものと言い換えられる. このとき, 次の定理が成り立つ.

Theorem 3. 確率変数ベクトル $\boldsymbol{U}$ が, $\boldsymbol{Z}=\boldsymbol{z}$ を与えたとき定数ベクトルになるならば, $\boldsymbol{\phi}=(\alpha\vdots\boldsymbol{\beta}\vdots\sigma^2)$ の *pseudo true value* $\boldsymbol{\theta}_*=(\alpha_*\vdots\boldsymbol{\beta}_*\vdots\sigma_*^2)$ は,

$$
\text{(15)}\qquad \boldsymbol{\theta}_*=\begin{pmatrix}\alpha_*\\ \boldsymbol{\beta}_*\\ \sigma_*^2\end{pmatrix}=\begin{pmatrix}\alpha_0+(E[\boldsymbol{Z}]'-E[\boldsymbol{U}]'\Sigma_{\boldsymbol{UU}'}^{-1}\Sigma_{\boldsymbol{UZ}'})\boldsymbol{\zeta}_0\\ \Sigma_{\boldsymbol{UU}'}^{-1}\Sigma_{\boldsymbol{UZ}'}\boldsymbol{\zeta}_0\\ \sigma_0^2+\boldsymbol{\zeta}_0'(\Sigma_{\boldsymbol{ZZ}'}-\Sigma_{\boldsymbol{ZU}'}\Sigma_{\boldsymbol{UU}'}^{-1}\Sigma_{\boldsymbol{UZ}'})\boldsymbol{\zeta}_0\end{pmatrix}
$$

であり,

$$
\text{(16)}\qquad l_g(\alpha_*,\boldsymbol{\beta}_*,\sigma_*^2)=\frac{1}{2}+\frac{1}{2}\ln\sigma_*^2
$$

となる.

[証明] まず, $\alpha_*,\boldsymbol{\beta}_*$ を求めよう.

$$
\begin{aligned}
\int(y-\alpha-\boldsymbol{u}'\boldsymbol{\beta})^2h^0(y|\boldsymbol{z})dy &= E_{Y|\boldsymbol{z}}\left[(Y-\alpha-\boldsymbol{u}'\boldsymbol{\beta})^2\right]\\
&= E_{Y|\boldsymbol{z}}\left[\left\{(Y-\alpha_0-\boldsymbol{z}'\boldsymbol{\zeta}_0)+(\alpha_0+\boldsymbol{z}'\boldsymbol{\zeta}_0-\alpha-\boldsymbol{u}'\boldsymbol{\beta})\right\}^2\right]\\
&= E_{Y|\boldsymbol{z}}\left[(Y-\alpha-\boldsymbol{z}'\boldsymbol{\zeta}_0)^2\right.\\
&\qquad +2(Y-\alpha_0-\boldsymbol{z}'\boldsymbol{\zeta}_0)(\alpha_0+\boldsymbol{z}'\boldsymbol{\zeta}_0-\alpha-\boldsymbol{u}'\boldsymbol{\beta})\\
&\qquad\qquad \left.+(\alpha_0+\boldsymbol{z}'\boldsymbol{\zeta}_0-\alpha-\boldsymbol{u}'\boldsymbol{\beta})^2\right]\\
&= \sigma_0^2+(\alpha_0+\boldsymbol{z}'\boldsymbol{\zeta}_0-\alpha-\boldsymbol{u}'\boldsymbol{\beta})^2
\end{aligned}
$$

であり, $\boldsymbol{Z}$ についての期待値をとれば,

$$E\Big[\big\{(\alpha_0+\boldsymbol{Z}'\boldsymbol{\zeta}_0)-(\alpha+\boldsymbol{U}'\boldsymbol{\beta})\big\}^2\Big]=$$

$$(\alpha_0,\boldsymbol{\zeta}_0')\begin{pmatrix}1 & E[\boldsymbol{Z}]'\\ E[\boldsymbol{Z}] & E\big[\boldsymbol{Z}\boldsymbol{Z}'\big]\end{pmatrix}\begin{pmatrix}\alpha_0\\ \boldsymbol{\zeta}_0\end{pmatrix}-2(\alpha_0,\boldsymbol{\zeta}_0')\begin{pmatrix}1 & E[\boldsymbol{U}]'\\ E[\boldsymbol{Z}] & E\big[\boldsymbol{Z}\boldsymbol{U}'\big]\end{pmatrix}\begin{pmatrix}\alpha\\ \boldsymbol{\beta}\end{pmatrix}$$

$$+(\alpha,\boldsymbol{\beta}')\begin{pmatrix}1 & E[\boldsymbol{U}]'\\ E[\boldsymbol{U}] & E\big[\boldsymbol{U}\boldsymbol{U}'\big]\end{pmatrix}\begin{pmatrix}\alpha\\ \boldsymbol{\beta}\end{pmatrix}$$

を得る. これを最小化する $\alpha,\boldsymbol{\beta}$ は,

$$\begin{pmatrix}1 & E[\boldsymbol{U}]'\\ E[\boldsymbol{U}] & E\big[\boldsymbol{U}\boldsymbol{U}'\big]\end{pmatrix}\begin{pmatrix}\alpha\\ \boldsymbol{\beta}\end{pmatrix}-\begin{pmatrix}1 & E[\boldsymbol{Z}]'\\ E[\boldsymbol{U}] & E\big[\boldsymbol{U}\boldsymbol{Z}'\big]\end{pmatrix}\begin{pmatrix}\alpha_0\\ \boldsymbol{\zeta}_0\end{pmatrix}=\boldsymbol{0}$$

を満たすので,

$$\begin{aligned}\begin{pmatrix}\alpha_*\\ \boldsymbol{\beta}_*\end{pmatrix}&=\begin{pmatrix}1 & E[\boldsymbol{U}]'\\ E[\boldsymbol{U}] & E\big[\boldsymbol{U}\boldsymbol{U}'\big]\end{pmatrix}^{-1}\begin{pmatrix}1 & E[\boldsymbol{Z}]'\\ E[\boldsymbol{U}] & E\big[\boldsymbol{U}\boldsymbol{Z}'\big]\end{pmatrix}\begin{pmatrix}\alpha_0\\ \boldsymbol{\zeta}_0\end{pmatrix}\\
&=\begin{pmatrix}1+E[\boldsymbol{U}]'\Sigma_{\boldsymbol{U}\boldsymbol{U}'}^{-1}E[\boldsymbol{U}] & -E[\boldsymbol{U}]'\Sigma_{\boldsymbol{U}\boldsymbol{U}'}^{-1}\\ -\Sigma_{\boldsymbol{U}\boldsymbol{U}'}^{-1}E[\boldsymbol{U}] & \Sigma_{\boldsymbol{U}\boldsymbol{U}'}^{-1}\end{pmatrix}\begin{pmatrix}1 & E[\boldsymbol{Z}]'\\ E[\boldsymbol{U}] & E\big[\boldsymbol{U}\boldsymbol{Z}'\big]\end{pmatrix}\begin{pmatrix}\alpha_0\\ \boldsymbol{\zeta}_0\end{pmatrix}\\
&=\begin{pmatrix}1 & E[\boldsymbol{Z}]'-E[\boldsymbol{U}]'\Sigma_{\boldsymbol{U}\boldsymbol{U}'}^{-1}\Sigma_{\boldsymbol{U}\boldsymbol{Z}'}\\ 0 & \Sigma_{\boldsymbol{U}\boldsymbol{U}'}^{-1}\Sigma_{\boldsymbol{U}\boldsymbol{Z}'}\end{pmatrix}\begin{pmatrix}\alpha_0\\ \boldsymbol{\zeta}_0\end{pmatrix}\\
&=\begin{pmatrix}\alpha_0+(E[\boldsymbol{Z}]'-E[\boldsymbol{U}]'\Sigma_{\boldsymbol{U}\boldsymbol{U}'}^{-1}\Sigma_{\boldsymbol{U}\boldsymbol{Z}'})\boldsymbol{\zeta}_0\\ \Sigma_{\boldsymbol{U}\boldsymbol{U}'}^{-1}\Sigma_{\boldsymbol{U}\boldsymbol{Z}'}\boldsymbol{\zeta}_0\end{pmatrix}\end{aligned}$$

と表される. この式を利用すれば,

$$\begin{aligned}&E\Big[\big\{(\alpha_0+\boldsymbol{Z}'\boldsymbol{\zeta}_0)-(\alpha_*+\boldsymbol{U}'\boldsymbol{\beta}_*)\big\}^2\Big]\\
&\qquad=E\Big[\big\{(\boldsymbol{Z}-E[\boldsymbol{Z}])'\boldsymbol{\zeta}_0-(\boldsymbol{U}-E[\boldsymbol{U}])'\Sigma_{\boldsymbol{U}\boldsymbol{U}'}^{-1}\Sigma_{\boldsymbol{U}\boldsymbol{Z}'})\boldsymbol{\zeta}_0\big\}^2\Big]\\
&\qquad=\boldsymbol{\zeta}_0'\big(\Sigma_{\boldsymbol{Z}\boldsymbol{Z}'}-\Sigma_{\boldsymbol{Z}\boldsymbol{U}'}\Sigma_{\boldsymbol{U}\boldsymbol{U}'}^{-1}\Sigma_{\boldsymbol{U}\boldsymbol{Z}'}\big)\boldsymbol{\zeta}_0\end{aligned}$$

と表される. σ_*^2 は, $\frac{d}{d\sigma^2}l(\boldsymbol{\alpha},\sigma^2)=0$ を満たすので,

$$\frac{d}{d\sigma^2}l(\boldsymbol{\alpha},\sigma^2)=\frac{1}{(\sigma^2)^2}\Big\{\sigma^2-\big(\sigma_0^2+E[(\alpha_0+\boldsymbol{Z}'\boldsymbol{\zeta}_0-\alpha-\boldsymbol{U}'\boldsymbol{\beta})^2]\big)\Big\}=0$$

の解として与えられる (証明終わり).

次に, A_g, B_g を計算しよう.

$$\ln g(y|\boldsymbol{z};\phi)=-\frac{1}{2}\ln\sigma^2-\frac{(y-\alpha-\boldsymbol{u}'\boldsymbol{\beta})^2}{2\sigma^2}-\frac{1}{2}\ln 2\pi$$

だから,

$$\frac{\partial \ln g}{\partial\begin{pmatrix}\alpha\\ \boldsymbol{\beta}\\ \sigma^2\end{pmatrix}}=\begin{pmatrix}\dfrac{y-\alpha-\boldsymbol{u}'\boldsymbol{\beta}}{\sigma^2}\\ \dfrac{y-\alpha-\boldsymbol{u}'\boldsymbol{\beta}}{\sigma^2}\boldsymbol{u}\\ -\dfrac{1}{2\sigma^2}+\dfrac{(y-\alpha-\boldsymbol{u}'\boldsymbol{\beta})^2}{2(\sigma^2)^2}\end{pmatrix}$$

である．ここで，$r_g=y-\alpha-\boldsymbol{u}'\boldsymbol{\beta}$ と置くと，

$$\frac{\partial^2 \ln g}{\partial\begin{pmatrix}\alpha\\ \boldsymbol{\beta}\\ \sigma^2\end{pmatrix}\partial\begin{pmatrix}\alpha\\ \boldsymbol{\beta}\\ \sigma^2\end{pmatrix}'}=\begin{pmatrix}-\dfrac{1}{\sigma^2} & -\dfrac{1}{\sigma^2}\boldsymbol{u}' & -\dfrac{r_g}{(\sigma^2)^2}\\ -\dfrac{1}{\sigma^2}\boldsymbol{u} & -\dfrac{1}{\sigma^2}\boldsymbol{u}\boldsymbol{u}' & -\dfrac{r_g}{(\sigma^2)^2}\boldsymbol{u}\\ -\dfrac{r_g}{(\sigma^2)^2} & -\dfrac{r_g}{(\sigma^2)^2}\boldsymbol{u}' & -\dfrac{r_g^2}{(\sigma^2)^3}+\dfrac{1}{2(\sigma^2)^2}\end{pmatrix}$$

$$\frac{\partial \ln g}{\partial\begin{pmatrix}\alpha\\ \boldsymbol{\beta}\\ \sigma^2\end{pmatrix}}\frac{\partial \ln g}{\partial\begin{pmatrix}\alpha\\ \boldsymbol{\beta}\\ \sigma^2\end{pmatrix}}'=$$

$$\begin{pmatrix}\dfrac{r_g^2}{(\sigma^2)^2} & \dfrac{r_g^2}{(\sigma^2)^2}\boldsymbol{u}' & -\dfrac{r_g}{2(\sigma^2)^2}+\dfrac{r_g^3}{2(\sigma^2)^3}\\ \dfrac{r_g^2}{(\sigma^2)^2}\boldsymbol{u} & \dfrac{r_g^2}{(\sigma^2)^2}\boldsymbol{u}\boldsymbol{u}' & -\dfrac{r_g}{2(\sigma^2)^2}\boldsymbol{u}+\dfrac{r_g^3}{2(\sigma^2)^3}\boldsymbol{u}\\ -\dfrac{r_g}{2(\sigma^2)^2}+\dfrac{r_g^3}{2(\sigma^2)^3} & -\dfrac{r_g}{2(\sigma^2)^2}\boldsymbol{u}'+\dfrac{r_g^3}{2(\sigma^2)^3}\boldsymbol{u}' & \dfrac{1}{4(\sigma^2)^4}\left\{r_g^2-\sigma^2\right\}^2\end{pmatrix}$$

と表される．このとき次の定理が成り立つ．

Theorem 4. $\boldsymbol{U}$, $\boldsymbol{Z}$ がそれぞれ退化しない多変量正規分布に従っていて，$\boldsymbol{U}$ が $\boldsymbol{Z}$ の線形結合で表されるならば，$-A_g(\boldsymbol{\phi}_*)=B_g(\boldsymbol{\phi})$ が成立する．

[証明] 前定理で，$E[Y-\alpha_*-\boldsymbol{U}'\boldsymbol{\beta}_*]=0$, $E[(Y-\alpha_*-\boldsymbol{U}'\boldsymbol{\beta}_*)^2]=\sigma_*^2$, $E[(Y-\alpha_*-\boldsymbol{U}'\boldsymbol{\beta}_*)\boldsymbol{U}]=\boldsymbol{0}$ であることが示されている．仮定より $\boldsymbol{Z}$ は多変量正規分布に従うので，Y も正規分布に従い，そしてそれ故 $Y-\alpha_*-\boldsymbol{U}'\boldsymbol{\beta}_*$ もまた正規分布に従う．実際，$Y-\alpha_*-\boldsymbol{U}'\boldsymbol{\beta}_*$ は正規分布 $N(0,\sigma_*^2)$ に従い，$\boldsymbol{U}$ とは独立であることが分かる．このことから直ちに，

$$(17)\qquad\begin{cases}E\big[(Y-\alpha_*-\boldsymbol{U}'\boldsymbol{\beta}_*)^3\big]=0\\ E\big[(Y-\alpha_*-\boldsymbol{U}'\boldsymbol{\beta}_*)^3\boldsymbol{U}\big]=E\big[(Y-\alpha_*-\boldsymbol{U}'\boldsymbol{\beta}_*)^3\big]E\big[\boldsymbol{U}\big]=\boldsymbol{0}\\ E\big[(Y-\alpha_*-\boldsymbol{U}'\boldsymbol{\beta}_*)^4\big]=3(\sigma_*^2)^2\end{cases}$$

が導かれる．これらの値を用いて，各要素の期待値を計算することにより，

$$(18) \qquad -A_g(\phi_*) = B_g(\phi) = \begin{pmatrix} \frac{1}{\sigma^2} & \frac{1}{\sigma^2}E[\boldsymbol{U}]' & 0 \\ \frac{1}{\sigma^2}E[\boldsymbol{U}] & \frac{1}{\sigma^2}E[\boldsymbol{U}\boldsymbol{U}'] & \mathbf{0} \\ 0 & \mathbf{0}' & \frac{1}{2(\sigma^2)^2} \end{pmatrix}$$

であることが確かめられる (証明終わり)．

この定理の結果から直ちに，$\boldsymbol{U}$ を p 次元ベクトルとするとき，

$$(19) \qquad 2\{L_g(\hat{\boldsymbol{\phi}}_n) - L_g(\boldsymbol{\phi}_*)\} \asymp \chi^2_{p+2}$$

であることが示される．この性質をうまく活用できるのが，Nested Case である．**Example 4** を考えよう．$F_{\boldsymbol{\theta}}$ の場合も同様に，

$$(20) \qquad 2\{L_f(\hat{\boldsymbol{\theta}}_n) - L_f(\boldsymbol{\theta}_*)\} \asymp \chi^2_{p+q+2}$$

が成り立つ．いま，$f(y|\boldsymbol{z};\boldsymbol{\theta}_*) = g(y|\boldsymbol{z};\boldsymbol{\phi}_*)$ の下では，$L_g(\boldsymbol{\phi}_*) = L_f(\boldsymbol{\theta}_*)$ であり，$L_f(\hat{\boldsymbol{\theta}}_n) - L_g(\hat{\boldsymbol{\phi}}_n)$ は，モデル間に包含関係があるため必ず非負になる．一般に，2 つのカイ自乗分布に従う確率変数 $X \sim \chi^2_p, Y \sim \chi^2_q$, $(p > q)$ があって，$\Pr(X \geq Y) = 1$ ならば，$X - Y$ は自由度 $p - q$ のカイ自乗分布に従う (例えば Rao, 1973 を参照されたい) ので，

$$(21) \qquad L_f(\hat{\boldsymbol{\theta}}_n) - L_g(\hat{\boldsymbol{\phi}}_n) \asymp \chi^2_q$$

であることが示される．以上のように，Vuong 検定を正規線形回帰モデルに適用する際，Nested case では χ^2 検定を行えばよいことがわかる．

この事実をうまく活用して，Overlapping case においては，従来の方法ではなく，次のような手続きを提案する．

1. $H_0 : F_{\boldsymbol{\theta}} \cap G_{\boldsymbol{\phi}} \ni h^0(y|\boldsymbol{z})$, $H_1 : F_{\boldsymbol{\theta}} \ni h^0(y|\boldsymbol{z})$ と $H_0 : F_{\boldsymbol{\theta}} \cap G_{\boldsymbol{\phi}} \ni h^0(y|\boldsymbol{z})$, $H_1 : G_{\boldsymbol{\phi}} \ni h^0(y|\boldsymbol{z})$ という Nested case の検定を行う．

2. もし，両方とも有意でなければ $f(y|\boldsymbol{z};\boldsymbol{\theta}_*) = g(y|\boldsymbol{z};\boldsymbol{\phi}_*)$ と結論付け，片方だけ有意なら有意になった方の対立仮説を採択する．両方とも有意なら Strictly non-nested case としてもう一度検定を行う．

このようにすれば，重み付きカイ二乗和分布を導入する必要がなくなり，面倒な棄却限界値の計算をしなくて済むのである．この手続きを適用する際，説明変数をも確率変数として扱い，その分布が正規分布に従うとみなしてよいかを確認しなければならないが，この要請は，説明変数を適当に変換すれば満たされる場合が多い．そもそも，説明変数が正規分布に従うとみなしてよい

という状況は, 回帰分析の信頼性が高くなるため, 正規線形回帰モデルを当てはめる時点で要請しても良いことなのである. それ故, ここで紹介した手続きは, 正規線形回帰モデルでは至極妥当なものといえる.

この節の最後に, $f(y|\boldsymbol{z};\boldsymbol{\theta}_*) = g(y|\boldsymbol{z};\boldsymbol{\phi}_*)$ が具体的に何を意味するのかを考察する. $f(y|\boldsymbol{z};\boldsymbol{\theta})$ を **Example 4** で定義したものとする. すると, (15) 式より,

$$\begin{cases} \alpha_{f*} = \alpha_0 + \Big(E[\boldsymbol{Z}]' - (E[\boldsymbol{U}]', E[\boldsymbol{V}]')\Sigma^{-1}_{(UV)(UV)'}\Sigma_{(UV)Z'}\Big)\boldsymbol{\zeta}_0 \\ \begin{pmatrix} \boldsymbol{\beta}_{f*} \\ \boldsymbol{\gamma}_{f*} \end{pmatrix} = \Sigma^{-1}_{(UV)(UV)'}\Sigma_{(UV)Z'}\boldsymbol{\zeta}_0 \end{cases}$$

と表される. ここで,

$$\Sigma_{(UV)(UV)'} = \begin{pmatrix} \Sigma_{UU'} & \Sigma_{UU'} \\ \Sigma_{UU'} & \Sigma_{UU'} \end{pmatrix}, \qquad \Sigma_{(UV)Z'} = \begin{pmatrix} \Sigma_{UZ'} \\ \Sigma_{VZ'} \end{pmatrix}$$

とする. さらに,

$$\begin{cases} \Sigma_{U_V U_V'} = \Sigma_{UU'} - \Sigma_{UV'}\Sigma^{-1}_{VV'}\Sigma_{VU'} \\ \Sigma_{V_U V_U'} = \Sigma_{VV'} - \Sigma_{VU'}\Sigma^{-1}_{UU'}\Sigma_{UV'} \end{cases}$$

とおくとき,

$$\Sigma^{-1}_{(UV)(UV)'} = \begin{pmatrix} \Sigma_{UU'} & \Sigma_{UU'} \\ \Sigma_{UU'} & \Sigma_{UU'} \end{pmatrix}^{-1} = \begin{pmatrix} \Sigma^{-1}_{U_V U_V'} & -\Sigma^{-1}_{U_V U_V'}\Sigma_{UV'}\Sigma^{-1}_{VV'} \\ -\Sigma^{-1}_{V_U V_U'}\Sigma_{VU'}\Sigma^{-1}_{UU'} & \Sigma^{-1}_{V_U V_U'} \end{pmatrix}$$

と表される. $\boldsymbol{Z}$ を $(\boldsymbol{U}\vdots\boldsymbol{V})$ に未知の説明変数を付け加えた形で, $\boldsymbol{Z} = (\boldsymbol{U}\vdots\boldsymbol{V}\vdots\boldsymbol{W})$ と表しても一般性を失わない. この $\boldsymbol{Z}$ に対し, $\boldsymbol{\zeta}_0 = (\boldsymbol{\beta}_0\vdots\boldsymbol{\gamma}_0\vdots\boldsymbol{\delta}_0)$ とし, $\Sigma_{U_V W_V'} \equiv \Sigma_{UW'} - \Sigma_{UV'}\Sigma^{-1}_{VV'}\Sigma_{VW'}$ という記法を導入すると,

$$\Sigma^{-1}_{(UV)(UV)'}\Sigma_{(UV)Z'} = \begin{pmatrix} I_p & O & \Sigma_{U_V U_V'}\Sigma_{U_V W_V'} \\ O & I_q & \Sigma_{V_U V_U'}\Sigma_{V_U W_U'} \end{pmatrix}$$

という簡単な形で表現され,

$$\begin{pmatrix} \alpha_{f*} \\ \boldsymbol{\beta}_{f*} \\ \boldsymbol{\gamma}_{f*} \end{pmatrix} = \begin{pmatrix} \alpha_0 + E[W]'\boldsymbol{\delta}_0 - E[U]'\Sigma^{-1}_{U_V U_V'}\Sigma_{U_V W_V'}\boldsymbol{\delta}_0 - E[V]'\Sigma^{-1}_{V_U V_U'}\Sigma_{V_U W_U'}\boldsymbol{\delta}_0 \\ \boldsymbol{\beta}_0 + \Sigma^{-1}_{U_V U_V'}\Sigma_{U_V W_V'}\boldsymbol{\delta}_0 \\ \boldsymbol{\gamma}_0 + \Sigma^{-1}_{V_U V_U'}\Sigma_{V_U W_U'}\boldsymbol{\delta}_0 \end{pmatrix} \tag{22}$$

という表現を得る. これを,

$$\begin{pmatrix} \alpha_{g*} \\ \boldsymbol{\beta}_{g*} \end{pmatrix} = \begin{pmatrix} \alpha_0 + E[\boldsymbol{V}]'\boldsymbol{\gamma}_0 - E[\boldsymbol{U}]'\Sigma^{-1}_{UU'}\Sigma_{UV'}\boldsymbol{\gamma}_0 + E[\boldsymbol{W}]'\boldsymbol{\delta}_0 - E[\boldsymbol{U}]'\Sigma^{-1}_{UU'}\Sigma_{UW'}\boldsymbol{\delta}_0 \\ \boldsymbol{\beta}_0 + \Sigma^{-1}_{UU'}\Sigma_{UV'}\boldsymbol{\gamma}_0 + \Sigma^{-1}_{UU'}\Sigma_{UW'}\boldsymbol{\delta}_0 \end{pmatrix} \tag{23}$$

と比較して, $\alpha_{f*} = \alpha_{g*}$, $\boldsymbol{\beta}_{f*} = \boldsymbol{\beta}_{g*}$, $\boldsymbol{\gamma}_{f*} = \boldsymbol{0}$ が同時に満たされるのはどんな場合かを考えてゆこう. 具体的には,

$$\boldsymbol{\gamma}_0 + \Sigma^{-1}_{V_U V_U'}\Sigma_{V_U W_U'}\boldsymbol{\delta}_0 = \boldsymbol{0} \tag{24}$$

(25) $\quad \left(\Sigma_{UU'}^{-1}\Sigma_{UW'} - \Sigma_{U_V U'_V}^{-1}\Sigma_{U_V W'_V}\right)\boldsymbol{\delta}_0 = \mathbf{0}$

(26) $$E[\boldsymbol{U}]'\left(\Sigma_{UU'}^{-1}\Sigma_{UW'} - \Sigma_{U_V U'_V}^{-1}\Sigma_{U_V W'_V} - \Sigma_{V_U V'_U}^{-1}\Sigma_{V_U W'_U}\right)\boldsymbol{\delta}_0 + \left(E[\boldsymbol{V}]' - E[\boldsymbol{U}]'\Sigma_{UU'}^{-1}\Sigma_{UV'}\right)\boldsymbol{\gamma}_0 = 0$$

を同時に満たす条件を探すことになる．任意の $\boldsymbol{\delta}_0$ にたいして (24) が成り立つには，

(27) $\quad \boldsymbol{\gamma}_0 = \mathbf{0}$ and $\Sigma_{V_U W'_U} = O$

でなければならない．その理由は，$\boldsymbol{W}$ そしてもちろん $\boldsymbol{\delta}_0$ は実際には明示されない仮想的なものであり，$\boldsymbol{\gamma}_0$ が仮想的な量により決まるとは考えられない．それゆえ，$\boldsymbol{\gamma}_0 = \mathbf{0}$ かつ $\Sigma_{V_U V'_U}^{-1}\Sigma_{V_U W'_U} = O$ であることが要請される．$\Sigma_{V_U V'_U}^{-1}$ は正則行列であるので，$\Sigma_{V_U W'_U} = O$ であることが要請されるのである．

(27) が成立するならば，(25), (26) は，共に，

(28) $\quad \Sigma_{UU'}^{-1}\Sigma_{UW'} - \Sigma_{U_V U'_V}^{-1}\Sigma_{U_V W'_V} = O$

ならば成立する．逆行列の公式，$(I+AB)^{-1} = I - A(I+BA)^{-1}B$ を使えば，

$$\begin{aligned}
\Sigma_{U_V U'_V}^{-1} &= \left(\Sigma_{UU'} - \Sigma_{UV'}\Sigma_{VV'}^{-1}\Sigma_{VU'}\right)^{-1} \\
&= \left\{\Sigma_{UU'}\left(I - \Sigma_{UU'}^{-1}\Sigma_{UV'}\Sigma_{VV'}^{-1}\Sigma_{VU'}\right)\right\}^{-1} \\
&= \left\{I + \Sigma_{UU'}^{-1}\left(I - \Sigma_{UV'}\Sigma_{VV'}^{-1}\Sigma_{VU'}\Sigma_{UU'}^{-1}\right)^{-1}\Sigma_{UV'}\Sigma_{VV'}^{-1}\Sigma_{VU'}\right\}\Sigma_{UU'}^{-1}
\end{aligned}$$

となることから，

$$\begin{aligned}
&\Sigma_{U_V U'_V}^{-1}\Sigma_{U_V W'_V} \\
&\quad = \Sigma_{UU'}^{-1}\Sigma_{UW'} - \Sigma_{UU'}^{-1}\Sigma_{UV'}\Sigma_{VV'}^{-1}\Sigma_{VW'} \\
&\qquad + \Sigma_{UU'}^{-1}\left(I - \Sigma_{UU'}^{-1}\Sigma_{UV'}\Sigma_{VV'}^{-1}\Sigma_{VU'}\Sigma_{UU'}^{-1}\right)^{-1}\Sigma_{UV'}\Sigma_{VV'}^{-1}\Sigma_{VU'}\Sigma_{UU'}^{-1}\Sigma_{UW'} \\
&\qquad - \Sigma_{UU'}^{-1}\left(I - \Sigma_{UU'}^{-1}\Sigma_{UV'}\Sigma_{VV'}^{-1}\Sigma_{VU'}\Sigma_{UU'}^{-1}\right)^{-1}\Sigma_{UV'}\Sigma_{VV'}^{-1}\Sigma_{VU'}\Sigma_{UU'}^{-1}\Sigma_{UV'}\Sigma_{VV'}^{-1}\Sigma_{VW'} \\
&\quad = \Sigma_{UU'}^{-1}\Sigma_{UW'} - \Sigma_{UU'}^{-1}\left(I - \Sigma_{UU'}^{-1}\Sigma_{UV'}\Sigma_{VV'}^{-1}\Sigma_{VU'}\Sigma_{UU'}^{-1}\right)^{-1}\Sigma_{UV'}\Sigma_{VV'}^{-1}\Sigma_{V_U W'_U}
\end{aligned}$$

が成立する．このことより，(28) は，

(29) $\quad \Sigma_{UU'}^{-1}\left(I - \Sigma_{UU'}^{-1}\Sigma_{UV'}\Sigma_{VV'}^{-1}\Sigma_{VU'}\right)^{-1}\Sigma_{UV'}\Sigma_{VV'}^{-1}\Sigma_{V_U W'_U} = O$

となる．これより，$\Sigma_{UV'} = O$ か，あるいは，$\Sigma_{V_U W'_U} = O$ のどちらかが満たされればよいことが分かる．以上の議論をまとめると次のようになる．

Corollary 1. 任意の $\boldsymbol{\delta}_0$ に対して，$\alpha_{f*} = \alpha_{g*}$, $\boldsymbol{\beta}_{f*} = \boldsymbol{\beta}_{g*}$, $\boldsymbol{\gamma}_{f*} = \mathbf{0}$ となるための必要十分条件は，$\boldsymbol{\gamma}_0 = \mathbf{0}$, $\Sigma_{V_U W'_U} = O$ である．

さて $\Sigma_{V_U W'_U} = O$ は，実際上どのような意味を持つのだろうか．式の持つ意味は，$\boldsymbol{U}$ の影響を取り除けば，$\boldsymbol{V}$ と $\boldsymbol{W}$ は無相関であるということだ．つまり，$\boldsymbol{V}$ は，真の分布に含まれ

る可能性のある潜在的な説明変数とは, $\boldsymbol{U}$ を通してしか相関を持たないことを意味している. このような条件が満たされそうにない状況では, $f(y|\boldsymbol{z};\boldsymbol{\theta}_*) = g(y|\boldsymbol{z};\boldsymbol{\phi}_*)$ を検定することなく, **Theorem 1** の *(ii)* を直接適用してもよいだろう. Vuong 検定を行う際, いつも心を悩ませるのが, $f(y|\boldsymbol{z};\boldsymbol{\theta}_*) = g(y|\boldsymbol{z};\boldsymbol{\phi}_*)$ となる可能性である. このことを検定するには, Vuong (1989) では, 重み付きカイ自乗和分布を導入しなければならない. しかしながら, この節で述べたように, 極めて特殊な場面を除いては, $f(y|\boldsymbol{z};\boldsymbol{\theta}_*) \neq g(y|\boldsymbol{z};\boldsymbol{\phi}_*)$ であり, **Theorem 1** の *(i)* を適用する必要はない. もし, する必要があっても, この節で紹介した方法を利用すれば, カイ自乗検定に帰着されるので容易に実行可能となる.

5　NON-NESTED HYPOTHESES

この節では, Vuong 検定とよく似たアプローチである Non-nested 検定を取りあげ, 比較を行うことにより Vuong 検定の独自性を明らかにしてゆく. この節の, Non-nested 検定についての説明は, Pesaran and Weeks (2001) を基にしている. この問題の起源は, Cox (1961) にさかのぼる. 2 つのモデルに包含関係がない場合には, 包含関係がある場合に比べ, 検定統計量が標準的な分布に従わないため棄却限界値の設定が非常に難しくなる. Non-nested 検定では, Vuong 検定の場合とは異なり, どちらかのモデルに真の分布が含まれるという前提に立つ. $F_{\boldsymbol{\theta}}$ と $G_{\boldsymbol{\phi}}$ を non-nested なモデルとするとき,

(30)　　$H_0 : h^0(\boldsymbol{y}|\boldsymbol{z}) \in F_{\boldsymbol{\theta}}, \quad H_1 : h^0(\boldsymbol{y}|\boldsymbol{z}) \in G_{\boldsymbol{\phi}},$

そして,

(31)　　$H_0 : h^0(\boldsymbol{y}|\boldsymbol{z}) \in G_{\boldsymbol{\phi}}, \quad H_1 : h^0(\boldsymbol{y}|\boldsymbol{z}) \in F_{\boldsymbol{\theta}}$

の 2 つの検定を行う. ここでは, Nested の場合とは違って, モデル間に自然な順位付けがないため, 両モデルを対等に取り扱わなければならないのである. その結果, Non-Nested 検定では, 4 通りの結果が起こりうる.

これ以降, 正規線形回帰モデルを例として議論をするため, 次のような行列記号を導入する.

$$\mathbb{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \mathbb{Z} = \begin{pmatrix} \boldsymbol{Z}_1' \\ \vdots \\ \boldsymbol{Z}_n' \end{pmatrix}, \mathbb{U} = \begin{pmatrix} \boldsymbol{U}_1' \\ \vdots \\ \boldsymbol{U}_n' \end{pmatrix}, \mathbb{V} = \begin{pmatrix} \boldsymbol{V}_1' \\ \vdots \\ \boldsymbol{V}_n' \end{pmatrix}, \mathbb{W} = \begin{pmatrix} \boldsymbol{W}_1' \\ \vdots \\ \boldsymbol{W}_n' \end{pmatrix} \cdots.$$

$\mathbb{Z}, \mathbb{U}, \mathbb{V}, \mathbb{W}$ は状況に応じて確率変数であったり観測値であったりするが, Non-nested 検定では観測値が与えられた条件付の議論がなされる. これは, 説明変数をも確率変数として扱っている Vuong 検定とは対照的である. Vuong 検定では, 説明変数をも確率変数として扱い, 一つ一つの対数尤度比を独立同分布に従うものとして考え, 中心極限定理を適用して漸近正規性を保障して

いる.

Non-nested 検定では, 帰無仮説が Vuong 検定よりも具体的であるため, 説明変数が与えられたときの条件付の議論が可能になり, より正確な議論の余地があることは事実である. しかしながら, 真の分布がどちらか一方のモデルに含まれるという, あまり現実的でない, 仮定を置くことにどれだけの優位性があるのかを注意深く観察しなければならない.

5.1 The Cox Procedure

Non-nested 検定に対する, Cox の検定手続きで用いられる検定統計量は, Vuong 検定と同様に対数尤度比,

$$(32)\qquad LR_n(\hat{\boldsymbol{\theta}}_n,\hat{\boldsymbol{\phi}}_n)\equiv L_n^f(\hat{\boldsymbol{\theta}}_n)-L_n^g(\hat{\boldsymbol{\phi}}_n)=\sum_{t=1}^{n}\ln\frac{f(\mathbf{Y}_t|\boldsymbol{z}_t;\hat{\boldsymbol{\theta}}_n)}{g(\mathbf{Y}_t|\boldsymbol{z}_t;\hat{\boldsymbol{\phi}}_n)},$$

が用いられるが, 説明変数が与えられた条件付での議論になる. この統計量の漸近平均・分散を帰無仮説の下で導き, 中心極限定理を適用して, 漸近的に標準正規分布に従うように修正するのが普通である. この方法は, $f(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\theta}_*)\neq g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\phi}_*)$ である場合の Vuong 検定と良く似ている. Vuong 検定の場合には, 真の分布がどちらのモデルにも属していないことを前提に, 尤度理論を用いて漸近的性質を導いているので, Cox 検定のような精密な式変形が不可能である. 実際, Vuong 検定では, 説明変数を確率変数とし, 帰無仮説を巧妙に設定したため漸近平均が 0 であることが保障され, 漸近分散はモデルからではなくサンプルから推定される.

Cox 検定は, “真の分布がモデルに含まれる” というやや非現実的な帰無仮説を置いているため, より精緻な検定を実行できる可能性を持っているものの, 漸近平均・漸近分散の計算は正規線形回帰モデルの場合ですら困難である. 仮定が少々不自然でも, それに見合うだけの正確性が伴えば, 手法として利用する価値がある. そのことを検証するため, 最も身近な正規線形回帰モデルを例にとって Cox 検定を紹介しよう. $F_{\boldsymbol{\theta}}, G_{\boldsymbol{\phi}}$ を 2 つの non-nested な正規線形回帰モデルとする. より具体的には, $\mathbb{Z}$ を利用可能なすべての説明変数からなる計画行列とし, $\mathbb{U},\mathbb{V}$ はその一部分の説明変数を選択した計画行列で互いに包含関係のないものとするとき,

$$(33)\qquad F_{\boldsymbol{\theta}}:\mathbb{Y}\sim N\big[\mathbb{U}\boldsymbol{\alpha},\sigma_f^2 I_n\big],\qquad G_{\boldsymbol{\phi}}:\mathbb{Y}\sim N\big[\mathbb{V}\boldsymbol{\beta},\sigma_g^2 I_n\big],$$

ここで I_n は n 次単位行列, とする. このとき,

$$L_n^f(\hat{\boldsymbol{\theta}}_n)-L_n^g(\hat{\boldsymbol{\phi}}_n)=\frac{n}{2}\ln\left(\frac{\hat{\sigma}_g^2}{\hat{\sigma}_f^2}\right)$$

と表される. $F_{\boldsymbol{\theta}}$ が真の分布を含むと仮定するとき, $\boldsymbol{\alpha}_0$ を $\boldsymbol{\alpha}$ の真値とし, $\boldsymbol{\epsilon}$ を誤差とするとき,

$$L_n^f(\hat{\boldsymbol{\theta}}_n)-L_n^g(\hat{\boldsymbol{\phi}}_n)=\frac{n}{2}\ln\left(\frac{\boldsymbol{\epsilon}'(I-\mathbb{V}(\mathbb{V}'\mathbb{V})^{-1}\mathbb{V})\boldsymbol{\epsilon}+\boldsymbol{\alpha}_0{}'\mathbb{U}'(I-\mathbb{V}(\mathbb{V}'\mathbb{V})^{-1}\mathbb{V})\mathbb{U}\boldsymbol{\alpha}_0}{\boldsymbol{\epsilon}'(I-\mathbb{U}(\mathbb{U}'\mathbb{U})^{-1}\mathbb{U})\boldsymbol{\epsilon}}\right).$$

が成り立つ. この統計量を修正して漸近正規分布に従うようにするために, 漸近平均と漸近分散を求めなくてはならない. Non-nested 検定の帰無仮説の下で, この漸近平均と漸近分散を求める

ことになるのだが, この計算は正規線形回帰モデルの場合でさえ厄介な作業であることは上の式を見ても明らかである. そのため, 次の 2 つの分節にあるような簡易法が提案されてきている.

5.2 The Comprehensive Approach

このアプローチは, Atkinson (1970) が提案したもので, 2 つのモデル $F_{\boldsymbol{\theta}}$, $G_{\boldsymbol{\phi}}$ を特殊例として含む comprehensive model, C_η, を構成する. この際, $\eta = 0$ が $F_{\boldsymbol{\theta}}$ に対応し, $\eta = 1$ が $G_{\boldsymbol{\phi}}$ に対応するようにしておけば, Non-nested 検定 $H_0 : h^0(\boldsymbol{y}|\boldsymbol{z}) \in F_{\boldsymbol{\theta}}$, $H_1 : h^0(\boldsymbol{y}|\boldsymbol{z}) \in G_{\boldsymbol{\phi}}$, は, $H_0 : \eta = 0$, $H_1 : \eta = 1$ と書き換えられる.

comprehensive model の構成法は無限に考えられるが, その中の一つで数学的に扱いやすいものに,

$$(34) \qquad c_\eta(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\theta},\boldsymbol{\phi}) = \frac{f(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\theta})^{1-\eta} g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\phi})^{\eta}}{\int f(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\theta})^{1-\eta} g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\phi})^{\eta} d\boldsymbol{y}},$$

がある. このモデルを用いて, 2 つの正規線形回帰モデルの検定を行おうとすると,

$$(35) \qquad c_\eta(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\theta},\boldsymbol{\phi}) \sim N\left[\frac{(1-\eta)\nu^2}{\sigma_f^2}\boldsymbol{u}'\boldsymbol{\alpha} + \frac{\eta\nu^2}{\sigma_g^2}\boldsymbol{v}'\boldsymbol{\beta}, \nu^2\right], \text{ where } \nu^{-2} = \frac{1-\eta}{\sigma_f^2} + \frac{\eta}{\sigma_g^2},$$

となり, η と $\boldsymbol{\beta}$ とが分離されない. このままだと η についての検定が出来ず, 結果的には,

$$\begin{cases} H_0 : h^0(\boldsymbol{y}|\boldsymbol{z}) \in F_{\boldsymbol{\theta}}, \\ H_1 : h^0(\boldsymbol{y}|\boldsymbol{z}) \in \text{両モデルの説明変数の和集合を説明変数とするモデル} \end{cases}$$

という検定になってしまう. これでは, 本来の検定とは異なることを検定していることになる.

η を $\boldsymbol{\beta}$ から分離するための方法として, 何らかの方法で $\boldsymbol{\beta}$ を推定して固定しておき, その上で η について検定を行うという手続きが提案されている. それが, J 検定 と JA 検定 である. J 検定の場合, (35) 式の $\boldsymbol{\beta}$ を $\mathbb{V}(\mathbb{V}'\mathbb{V})^{-1}\mathbb{V}'y$ と固定して, η についての検定を行う. 一方, JA 検定では, $\boldsymbol{\beta}$ を $\mathbb{V}(\mathbb{V}'\mathbb{V})^{-1}\mathbb{V}'\mathbb{U}(\mathbb{U}'\mathbb{U})^{-1}\mathbb{U}'y$ と固定して検定を行うのである.

以上見てきたように, Atkinson の方法, および, その改良版である J 検定と JA 検定は, どれも, 本来の検定目的と実際の手続きとの関連が不明確であるという欠点を持つ. comprehensive model の構成法が一意でないこともあり, この Comprehensive approach はうまく機能しているとは言いがたい.

5.3 The Encompassing Approach

次に, Non-nested 検定とは少し異なるが, モデル間に encompassing という関係を導入して, その関係が成り立つかどうかを検定しようという考え方がある. このアプローチでも説明変数の値は既知であり, それが与えられた条件付での議論を行っている. いま, $\boldsymbol{\theta}_*(\boldsymbol{\phi}_{h*})$ を,

$$(36) \qquad \boldsymbol{\theta}_*(\boldsymbol{\phi}_{h*}) = \arg\max_{\boldsymbol{\theta}\in\Theta} E_{\boldsymbol{\phi}_{h*}}\left[\ln f(\mathbb{Y}|\mathbb{Z};\boldsymbol{\theta})\right], \text{ where } \boldsymbol{\phi}_{h*} = \arg\max_{\boldsymbol{\phi}\in\Phi} E_{h^0(y|z)}\left[\ln g(\mathbb{Y}|\mathbb{Z};\boldsymbol{\phi})\right]$$

とする．このとき，$\boldsymbol{\theta}_* = \boldsymbol{\theta}_*(\boldsymbol{\phi}_*)$ が成立するならば，モデル $G_{\boldsymbol{\phi}}$ はモデル $F_{\boldsymbol{\theta}}$ を encompass すると言う (モデル $F_{\boldsymbol{\theta}}$ が $G_{\boldsymbol{\phi}}$ を encompass する場合も同様に定義される)．

この場合の検定を定式化すると，

$$(37)\qquad H_0 : \boldsymbol{\theta}_{h*} = \boldsymbol{\theta}_*(\boldsymbol{\phi}_{h*}),\quad H_1 : \boldsymbol{\theta}_{h*} \neq \boldsymbol{\theta}_*(\boldsymbol{\phi}_{h*})$$

とあらわされる．統計量は，$\sqrt{n}\big[\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_*(\hat{\boldsymbol{\phi}}_n)\big]$ を用いるのが自然であろう．言い換えれば，この統計量を使うことが前提とされる関係が emcompassing とも言える．

この検定を 2 つの正規線形回帰モデル (33) を例に考えてみよう．真の条件付分布 $h^0(\boldsymbol{y}|\boldsymbol{z})$ を，

$$(38)\qquad \mathbb{Y} \sim N\big[\mathbb{Z}\boldsymbol{\zeta}_0, \sigma_0^2 I_n\big]$$

とする．このとき，

$$(39)\qquad \boldsymbol{\theta}_{h*} = \begin{pmatrix} \boldsymbol{\alpha}_{h*} \\ \sigma^2_{fh*} \end{pmatrix} = \begin{pmatrix} (\mathbb{U}'\mathbb{U})^{-1}\mathbb{U}'\mathbb{Z}\boldsymbol{\zeta}_0 \\ \sigma_0^2 + \boldsymbol{\delta}'(\mathbb{Z}'\mathbb{Z} - \mathbb{Z}'\mathbb{U}(\mathbb{U}'\mathbb{U})^{-1}\mathbb{U}'\mathbb{Z})\boldsymbol{\zeta}_0 \end{pmatrix}$$

$$(40)\qquad \boldsymbol{\phi}_{h*} = \begin{pmatrix} \boldsymbol{\beta}_{h*} \\ \sigma^2_{gh*} \end{pmatrix} = \begin{pmatrix} (\mathbb{V}'\mathbb{V})^{-1}\mathbb{V}'\mathbb{Z}\boldsymbol{\zeta}_0 \\ \sigma_0^2 + \boldsymbol{\zeta}_0'(\mathbb{Z}'\mathbb{Z} - \mathbb{Z}'\mathbb{V}(\mathbb{V}'\mathbb{V})^{-1}\mathbb{V}'\mathbb{Z})\boldsymbol{\zeta}_0 \end{pmatrix}$$

$$(41)\qquad \boldsymbol{\theta}_*(\boldsymbol{\phi}_{h*}) = \begin{pmatrix} \boldsymbol{\alpha}_*(\phi_{h*}) \\ \sigma^2_{f*}(\phi_{h*}) \end{pmatrix} = \begin{pmatrix} (\mathbb{U}'\mathbb{U})^{-1}\mathbb{U}'\mathbb{V}(\mathbb{V}'\mathbb{V})^{-1}\mathbb{V}'\mathbb{Z}\boldsymbol{\zeta}_0 \\ \\ \sigma_0^2 + \boldsymbol{\zeta}_0'\big(\mathbb{Z}'\mathbb{Z} - \mathbb{Z}'\mathbb{V}(\mathbb{V}'\mathbb{V})^{-1}\mathbb{V}'\mathbb{Z}\big)\boldsymbol{\zeta}_0 \\ +\boldsymbol{\zeta}_0'\mathbb{Z}\mathbb{V}(\mathbb{V}'\mathbb{V})^{-1}\big(\mathbb{V}'\mathbb{V} - \mathbb{V}'\mathbb{U}(\mathbb{U}'\mathbb{U})^{-1}\mathbb{U}'\mathbb{V}\big)(\mathbb{V}'\mathbb{V})^{-1}\mathbb{V}'\mathbb{Z}\boldsymbol{\zeta}_0 \end{pmatrix}$$

が成り立ち，$\boldsymbol{\theta}_{h*} = \boldsymbol{\theta}_*(\boldsymbol{\phi}_{h*})$ は，

$$(42)\qquad \mathbb{U}'\big(I - \mathbb{V}(\mathbb{V}'\mathbb{V})^{-1}\mathbb{V}'\big)\mathbb{Z}\boldsymbol{\zeta}_0 = \boldsymbol{0}$$

$$(43)\qquad \boldsymbol{\zeta}_0'\mathbb{Z}'\mathbb{U}(\mathbb{U}'\mathbb{U})^{-1}\mathbb{U}'\mathbb{Z}\boldsymbol{\zeta}_0 = \boldsymbol{\zeta}_0'\mathbb{Z}'\mathbb{V}(\mathbb{V}'\mathbb{V})^{-1}\mathbb{V}'\mathbb{U}(\mathbb{U}'\mathbb{U})^{-1}\mathbb{U}'\mathbb{V}(\mathbb{V}'\mathbb{V})^{-1}\mathbb{V}'\mathbb{Z}\boldsymbol{\zeta}_0$$

と書き換えられる．ここで，(42) ならば (43) が成り立つので，

$$(44)\qquad \big[\hat{\boldsymbol{\alpha}}_n - \boldsymbol{\alpha}_*(\hat{\boldsymbol{\beta}}_n)\big] = (\mathbb{U}'\mathbb{U})^{-1}\mathbb{U}(I - \mathbb{V}(\mathbb{V}'\mathbb{V})^{-1}\mathbb{V}')\mathbb{Y}$$

を統計量として取り上げる． この統計量は帰無仮説の下で，平均 $\mathbf{0}$ 分散 $\sigma_0^2(\mathbb{U}'\mathbb{U})^{-1}\mathbb{U}\Big(I - \mathbb{V}(\mathbb{V}'\mathbb{V})^{-1}\mathbb{V}'\Big)\ \mathbb{U}'(\mathbb{U}'\mathbb{U})^{-1}$ の正規分布に従う．この事実から検定統計量を構成するには，σ_0^2 を推定しなければならない．もし，σ_0^2 を $F_{\boldsymbol{\theta}}$ と $G_{\boldsymbol{\phi}}$ の説明変数を合わせたモデルで推定するならば，構成される検定統計量は仮説，$H_0 : G_{\boldsymbol{\phi} \ni h^0(\boldsymbol{y}|\boldsymbol{z})}$，$H_1$: “$F_{\boldsymbol{\theta}}$ と $G_{\boldsymbol{\phi}}$ の説明変数の和集合を説明変数とするモデル $\ni h^0(y|z)$” を検定する F 検定の場合と同じ統計量になってしまう．

以上のように，encompassing approach はモデルの包含関係よりも一般的な関係を調べるもので，検定統計量の構成が楽になるような仮説設定であるが，上の正規線形回帰モデルの場合を除けば，実際の計算が複雑で，検定の目的も明快ではない．

6 MODEL SELECTION CRITERIA

ここでは, モデル選択基準の中でも, Vuong 検定と同様, 尤度理論を用いて導出されている AIC に焦点を当て Vuong 検定との比較を行う. Vuong 検定は, AIC と同様に, モデルの良否を Kullback-Leibler を用いて判定しようとするものであるが, AIC とはかなり違ったものになっている. いくつか相違点がある. ひとつ目は, AIC が他のモデル選択基準と同様に, "モデルで推定された分布と真の分布との距離" を問題にしているのに対し, Vuong 検定では, "モデル内の分布と真の分布との最短距離" を問題にしているのである. つまり, モデル選択基準ではモデルを用いた予測が問題になっているのに対し, Vuong 検定では, モデルが真の分布にどれだけ近い分布を内包しているかを問題にしている. もうひとつは, 真の分布とモデルとの関係である. AIC を導く議論では, モデルは真の分布を, 少なくとも近似的に, 含むことを前提にしているが, Vuong 検定では必ずしも要求されていない.

AIC 理論では, 単調なモデル列,

$$(45) \qquad M_1 \subset M_2 \subset M_3 \subset \cdots \subset M_n \subset \cdots,$$

を考え, その極限が真の分布を含むものとする. モデル M_n の未知パラメータ数を k_n とするとき, どのモデルを用いれば Kullback–Leibler の情報量を最も小さくする予測分布が得られるかを知るための指標が AIC なのであり (坂元他, 1982),

$$(46) \qquad AIC = -2 \times (\text{モデル } M_n \text{ の最大対数尤度}) + 2 \times k_n,$$

により与えられる. これを最小にするモデルを最適とするのが AIC の考え方である.

一方, Vuong 検定では, その最も重要な適用場面である strictly non-nested case において, パラメータ数は無視されてしまう. これについて Vuong (1989) は注釈を設け改良の可能性を示唆しているものの, (14) を見れば分かるように, Vuong 検定では, その性質上, モデルのパラメータ数は無視されてしまうのである.

このように見てみると, AIC と Vuong 検定は競合するものではなく, 用途に応じて使い分けるものと言えよう. Vuong 検定では, "説明変数は同じだが分布とパラメータ構造が異なる 2 つのモデルの比較", あるいは, " 2 系統の説明変数 (数は等しい) を持つモデル間の比較" といった問題で, いずれのモデルも真の分布を含むとは考えられない場面での利用が推奨される. 他方 AIC は, モデルを大きくしてゆけば真の分布を含むと想定されるモデル列の中で, どのモデルを用いれば最も予測精度が良いかを知るための指標と言える.

7 CONCLUSION

ここまで, Vuong 検定の理論を解説し, 正規線形回帰モデルへの新たな, そして簡単な, 適用法を提案してきた. また, 良く似た概念である Non-nested 検定やモデル選択基準との関連について述べてきた.

正規線形回帰モデルに適用する際, どのような状況で, どのように Vuong 検定を用いればよいのかという問いに対する答えとして, 極めて実行しやすい手続きが提案できたと確信している. Vuong (1989) の提案した手続きでは, 2 つのモデルから得られる pseudo-true モデルが互いに等しいかどうかの検定を行う際, 扱い難い重み付きカイ自乗和分布を導入しなければならず, このことが Vuong 検定を利用する際の足枷になってきた. しかし, この論文で, 最も利用頻度の高い正規線形回帰モデルにおいてはカイ自乗検定で代用できることを示すことができた. この方法は, 従来扱いの難しかった Nested case と Overlapping case に適用できるため, Vuong 検定の利用を促進する効果があると言える.

また, Non-nested 検定との比較では, Vuong 検定の設定の巧みさを浮き彫りにすることが出来た. それにより, Vuong 検定とは, 前提の一般性と結果の一般性とが絶妙にバランスを取った手法であることが示された. 最後に, モデル選択基準との比較で, Vuong 検定の利用すべき場面を明確にすることが出来た.

参考文献

[1] Cox, D. R. (1961) Tests of separate families of hypotheses. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability.*

[2] Davidson, R. and MacKinnon, J. G. (1993) *Estimation and Inference in Econometrics*, Oxford University Press.

[3] Greene, W. H. (2000) *Econometric Analysis*, 4th edition. Prentice Hall.

[4] Pesaran, H. and Weeks, M. (2001) Nonnested Hypothesis Testing: An Overview. *A Companion to Theoretical Econometrics, edited by B. H. Baltagi*, Blackwell.

[5] Rao, C. R. (1973) *Linear Statistical Inference and Its Applications*, 2nd edition. John Wiley & Sons.

[6] Vuong, Q. H. (1989) Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses. *Econometrica*, Vol. 57, 307-333.

[7] 稲垣 宣生 (2003) 数理統計学 改訂版, 裳華房.

[8] 坂元 慶行, 石黒 真木夫, 北川 源四郎 (1982) 情報量統計学, 共立出版.

[9] 竹内 啓 (1976) 情報量基準 AIC とは何か, 数理科学, No.153, 5-11.

[10] 松尾 精彦 (2003) モデル選択基準とその正規線形回帰モデルへの適用, 関西大学経済論集, 第 53 巻, 93–107.