Vuong 検定によるモデル選択

太田浩司
松尾精彦

一　はじめに

Vuong (1989) 検定は、競合する二つのモデルが存在する際のモデル選択の統計的検定方法として、他分野でも幅広く用いられているが、とりわけ会計学における研究で頻繁に使用されている。その理由としては、会計情報には、代替的情報が非常に数多く存在し（例えば、当期純利益と包括利益、単体利益と連結利益、ABOとPBO）、その競合する情報のどちらの有用性が高いかということを検証することが、基準設定への示唆という観点からも、会計研究者の関心事であるからと思われる。本稿の目的は、Vuong 検定の理論的背景を説明し、そしてその応用例として、我が国で二〇〇〇年三月決算期から開示されるようになった連結キャッシュフロー情報と従来の発生主義に基づく会計利益情報の有用性を比較することである(1)。

二　Vuong (1989) 検定の理論

(1)　基本的フレームワーク

$\boldsymbol{X}_t$ は、m 次元確率変数列で、独立同分布に従うものとする。今、$\boldsymbol{X}_t=(\boldsymbol{Y}_t, \boldsymbol{Z}_t)$ と分割し、$\boldsymbol{Y}_t$ を非説明変数、$\boldsymbol{Z}_t$ を説明変数とする。そして $\boldsymbol{X}_t$ の真の密度関数を $h(\boldsymbol{x})$, $\boldsymbol{Z}_t$ の真の周辺密度関数を $h(\boldsymbol{z})$, $\boldsymbol{Z}_t$ が与えられたときの $\boldsymbol{Y}_t$ の真の条件付密度関数を $h(\boldsymbol{y}|\boldsymbol{z})$, $\boldsymbol{X}_t=(\boldsymbol{Y}_t, \boldsymbol{Z}_t)$ の真の同時確率分布についての期待値を $E^0[\cdot]$ で表すことにする。このとき、二つの競合するモデル、$\boldsymbol{F}_\theta \equiv \{f(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\theta});\ \boldsymbol{\theta}\in\boldsymbol{\Theta}\subset\mathbb{R}^p\}$, $\boldsymbol{G}_\gamma \equiv \{g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\gamma});\ \boldsymbol{\gamma}\in\Gamma\subset\mathbb{R}^q\}$ についての推測を行う。

$\boldsymbol{\theta}_*$, $\boldsymbol{\gamma}_*$ をモデル $\boldsymbol{F}_\theta$, $\boldsymbol{G}_\gamma$ の pseudo-true value(2) として、

$$\boldsymbol{\theta}_* \equiv \arg\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} E^0[\ln f(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\theta})],\quad \boldsymbol{\gamma}_* \equiv \arg\max_{\boldsymbol{\gamma}\in\Gamma} E^0[\ln g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\gamma})]$$

と定義する。そして $\hat{\boldsymbol{\theta}}_n$, $\hat{\boldsymbol{\gamma}}_n$ をモデル $\boldsymbol{F}_\theta$, $\boldsymbol{G}_\gamma$ の最尤推定量とし、各モデルの最大尤度を、

$$L_n^f(\hat{\boldsymbol{\theta}}_n) \equiv \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \sum_{t=1}^{n} \ln f(\boldsymbol{Y}_f|\boldsymbol{Z}_t;\boldsymbol{\theta}),\quad L_n^g(\hat{\boldsymbol{\gamma}}_n) \equiv \sup_{\boldsymbol{\gamma}\in\Gamma} \sum_{t=1}^{n} \ln g(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\boldsymbol{\gamma})$$

で表す。また、

$$A_f(\boldsymbol{\theta}) \equiv E^0\left[\frac{\partial^2 \ln f(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right],\quad A_g(\boldsymbol{\gamma}) \equiv E^0\left[\frac{\partial^2 \ln g(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\boldsymbol{\gamma})}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}'}\right],$$

$$B_f(\boldsymbol{\theta}) \equiv E^0\left[\frac{\partial \ln f(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\cdot\frac{\partial \ln f(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}\right],\quad B_g(\boldsymbol{\gamma}) \equiv E^0\left[\frac{\partial \ln g(\boldsymbol{Y}_t|\boldsymbol{Z}_f;\boldsymbol{\gamma})}{\partial\boldsymbol{\gamma}}\cdot\frac{\partial \ln g(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\boldsymbol{\gamma})}{\partial\boldsymbol{\gamma}'}\right],$$

$$B_{fg}(\boldsymbol{\theta},\boldsymbol{\gamma}) = B'_{gf}(\boldsymbol{\gamma},\boldsymbol{\theta}) \equiv E^0\left[\frac{\partial \ln f(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\cdot\frac{\partial \ln g(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\boldsymbol{\gamma})}{\partial\boldsymbol{\gamma}'}\right]$$

と定義する。このとき適当な正則条件の下で、$\hat{\boldsymbol{\theta}}_n$ と $\hat{\boldsymbol{\gamma}}_n$ はそれぞれ $\boldsymbol{\theta}_*$ と $\boldsymbol{\gamma}_*$ の一致推定量となり、その分布は漸近的に次の同時多変量正規分布に従う。

53

$$\sqrt{n}\begin{pmatrix}\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}_*\\ \hat{\boldsymbol{\gamma}}_n-\boldsymbol{\gamma}_*\end{pmatrix}\sim N\left[\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix},\begin{pmatrix}A_f^{-1}(\boldsymbol{\theta}_*)B_f(\boldsymbol{\theta}_*)A_f^{-1}(\boldsymbol{\theta}_*) & A_f^{-1}(\boldsymbol{\theta}_*)B_{fg}(\boldsymbol{\theta}_*,\boldsymbol{\gamma}_*)A_g^{-1}(\boldsymbol{\theta}_*)\\ A_g^{-1}(\boldsymbol{\gamma}_*)B_{gf}(\boldsymbol{\gamma}_*,\boldsymbol{\theta}_*)A_f^{-1}(\boldsymbol{\theta}_*) & A_g^{-1}(\boldsymbol{\gamma}_*)B_g(\boldsymbol{\gamma}_*)A_g^{-1}(\boldsymbol{\gamma}_*)\end{pmatrix}\right]。 \quad (1)$$

（2） Vuong 検定の概要

Vuong 検定は、真の分布 $h(\boldsymbol{y}|\boldsymbol{z})$ と、競合するモデル $\boldsymbol{F}_\theta$, $\boldsymbol{G}_\gamma$ との距離を Kullback-Leibler 情報量基準（ＫＬＩＣ）で測定し、その距離の差異を統計的に検定するものである。ＫＬＩＣでは、$h(\boldsymbol{y}|\boldsymbol{z})$ とモデル $\boldsymbol{F}_\theta$, $\boldsymbol{G}_\gamma$ の距離は、それぞれ、

$$KLIC(h,f)\equiv\int\ln\left[\frac{h(\boldsymbol{y}|\boldsymbol{z})}{f(\boldsymbol{y}|\boldsymbol{z}\,;\boldsymbol{\theta}_*)}\right]h(\boldsymbol{x})d\boldsymbol{x}=E^0[\ln h(\boldsymbol{Y}_t|\boldsymbol{Z}_t)]-E^0[\ln f(\boldsymbol{Y}_t|\boldsymbol{Z}_t\,;\boldsymbol{\theta}_*)],$$

$$KLIC(h,g)\equiv\int\ln\left[\frac{h(\boldsymbol{y}|\boldsymbol{z})}{g(\boldsymbol{y}|\boldsymbol{z}\,;\boldsymbol{\gamma}_*)}\right]h(\boldsymbol{x})d\boldsymbol{x}=E^0[\ln h(\boldsymbol{Y}_t|\boldsymbol{Z}_t)]-E^0[\ln g(\boldsymbol{Y}_t|\boldsymbol{Z}_t\,;\boldsymbol{\gamma}_*)]$$

で表される。両方とも右辺第一項は共通であるので、ＫＬＩＣの観点からは第二項が大きいほど真の分布との距離が小さいといえる。そこで、次の帰無仮説 H_0、およびその対立仮説 H_f, H_g を考える。

$$H_0: E^0[\ln f(\boldsymbol{Y}_t|\boldsymbol{Z}_t\,;\boldsymbol{\theta}_*)]=E^0[\ln g(\boldsymbol{Y}_t|\boldsymbol{Z}_t\,;\boldsymbol{\gamma}_*)] \quad (\boldsymbol{F}_\theta \text{ と } \boldsymbol{G}_\gamma \text{ は同等}),$$

$$H_f: E^0[\ln f(\boldsymbol{Y}_t|\boldsymbol{Z}_t\,;\boldsymbol{\theta}_*)]>E^0[\ln g(\boldsymbol{Y}_t|\boldsymbol{Z}_t\,;\boldsymbol{\gamma}_*)] \quad (\boldsymbol{F}_\theta \text{ は } \boldsymbol{G}_\gamma \text{ より良い}),$$

$$H_g: E^0[\ln f(\boldsymbol{Y}_t|\boldsymbol{Z}_t\,;\boldsymbol{\theta}_*)]<E^0[\ln g(\boldsymbol{Y}_t|\boldsymbol{Z}_t\,;\boldsymbol{\gamma}_*)] \quad (\boldsymbol{F}_\theta \text{ は } \boldsymbol{G}_\gamma \text{ より悪い})。$$

そして検定統計量には、尤度比（ＬＲ）統計量

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

を用いる。このとき、尤度関数の連続性についての適当な条件下で、大数の法則より、

$$\frac{1}{n}LR_n(\hat{\boldsymbol{\theta}}_n,\hat{\boldsymbol{\gamma}}_n)\xrightarrow{a.s.}E^0\left[\ln\frac{f(\boldsymbol{Y}_t|\boldsymbol{Z}_t\,;\boldsymbol{\theta}_*)}{g(\boldsymbol{Y}_t|\boldsymbol{Z}_t\,;\boldsymbol{\gamma}_*)}\right] \quad (2)$$

54

となる。また、$\ln[f(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\boldsymbol{\theta}_*)/g(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\boldsymbol{\gamma}_*)]$ の分散を ω_*^2、その推定量を $\hat{\omega}_n^2$ で表す。

定義 1（重み付き χ^2 和分布）：$Z_1, Z_2, \cdots, Z_m$ を m 個の互いに独立で標準正規分布に従う確率変数列とし、$\boldsymbol{\lambda}=(\lambda_1, \cdots, \lambda_m)'$ 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、パラメータ $(m, \boldsymbol{\lambda})$ を持つ重み付き χ^2 和分布と呼び、その累積分布関数を、$M_m(\cdot;\boldsymbol{\lambda})$ で表す。

ここで、pseudo-true value における尤度関数 $L_n^f(\boldsymbol{\theta}_*), L_n^g(\boldsymbol{\gamma}_*)$ をそれぞれ最尤推定量 $\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\gamma}}_n$ の周りで Taylor 展開し、(1)(2)を用いると次の定理1が得られる。

定理 1（LR 統計量の漸近分布）：適当な正則条件の下で、

（i）　$f(\cdot|\cdot;\boldsymbol{\theta}_*)=g(\cdot|\cdot;\boldsymbol{\gamma}_*)$ ならば、

$$2LR_n(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\gamma}}_n) \xrightarrow{d} M_{p+q}(\cdot;\boldsymbol{\lambda}_*)$$

が成り立つ。ここで $\boldsymbol{\lambda}_*$ は、

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

の $p+q$ 個の固有値を要素とするベクトルである。

（ii）　$f(\cdot|\cdot;\boldsymbol{\theta}_*)\neq g(\cdot|\cdot;\boldsymbol{\gamma}_*)$ ならば、

$$n^{-1/2}LR_n(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\gamma}}_n)-n^{1/2}E^0\left[\ln\frac{f(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\boldsymbol{\theta}_*)}{g(\boldsymbol{Y}_t|\boldsymbol{Z}_t;\boldsymbol{\gamma}_*)}\right] \xrightarrow{d} N(0, \omega_*^2)$$

が成り立つ。

(3)　Non-nested モデル、Overlapping モデル、Nested モデル

Vuong 検定では、競合するモデル $\boldsymbol{F}_\theta$ と $\boldsymbol{G}_\gamma$ を non-nested, overlapping, nested の三つの場合に分類して

55

いる。そしてそれぞれの場合に検定手続きが異なる。そこで最初に、三つのモデルの定義を正規線型モデルを例に用いて行う。

① Non-nested モデル ($\boldsymbol{F}_\theta \cap \boldsymbol{G}_\gamma = \phi$)：

$$Y_t = \alpha_0 + \alpha_1 X_t + \varepsilon_t, \tag{3 a}$$

$$Y_t = \beta_0 + \beta_1 Z_t + \varepsilon_t。 \tag{3 b}$$

(3a) と (3b) という、共通の説明変数を持たないモデル間の検定となる。

② Overlapping モデル ($\boldsymbol{F}_\theta \cap \boldsymbol{G}_\gamma \neq \phi$ かつ $\boldsymbol{F}_\theta \not\subset \boldsymbol{G}_\gamma$ and $\boldsymbol{G}_\gamma \not\subset \boldsymbol{F}_\theta$)：

$$Y_t = \alpha_0 + \alpha_1 W_t + \alpha_2 X_t + \varepsilon_t, \tag{4 a}$$

$$Y_t = \beta_0 + \beta_1 W_t + \beta_2 Z_t + \varepsilon_t。 \tag{4 b}$$

(4a) と (4b) のような、共通の説明変数を持つが、それぞれ固有の説明変数も持つモデル間の検定となる。

③ Nested モデル ($\boldsymbol{G}_\gamma \subset \boldsymbol{F}_\theta$)：

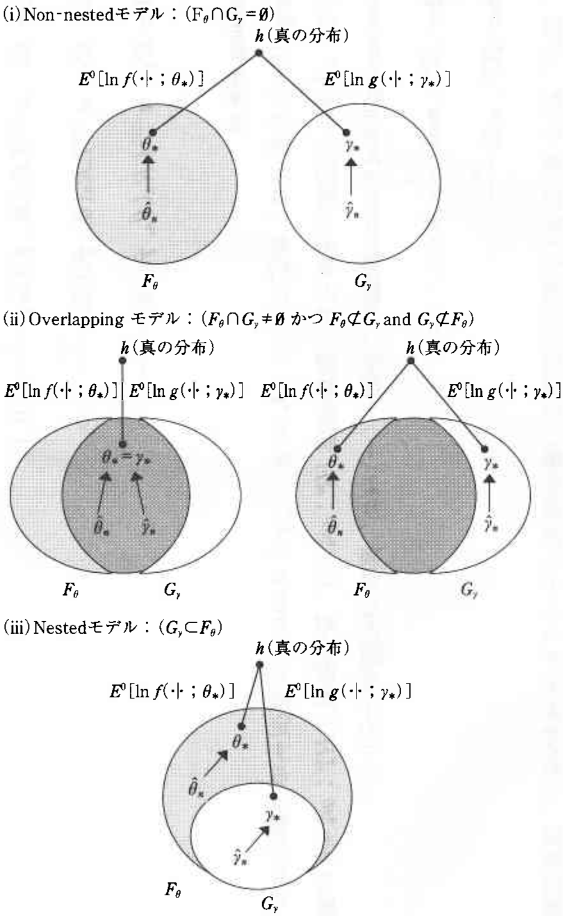
$$Y_t = \alpha_0 + \alpha_1 X_t + \varepsilon_t, \tag{5 a}$$

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 Z_t + \varepsilon_t。 \tag{5 b}$$

(5a) と (5b) のように、一方のモデルの全ての説明変数は、他方のモデルの説明変数でもあるという場合のモデル間の検定となる。

図1は、Vuong 検定の基本的な考え方をグラフ表現したものである。Vuong 検定は、ＫＬＩＣで測定された真の分布とモデルとの距離、すなわち $E^0[\ln f(\cdot|\cdot\,;\boldsymbol{\theta}_*)]$ と $E^0[\ln g(\cdot|\cdot\,;\boldsymbol{\gamma}_*)]$ を比較している。そして定理1にあるように、Vuong 検定の検定統計量の分布は、$f(\cdot|\cdot\,;\boldsymbol{\theta}_*) = g(\cdot|\cdot\,;\boldsymbol{\gamma}_*)$ と $f(\cdot|\cdot\,;\boldsymbol{\theta}_*) \neq g(\cdot|\cdot\,;\boldsymbol{\gamma}_*)$ の場合で異なる。帰無仮説 H_0 は、真の分布と、競合するモデル $\boldsymbol{F}_\theta$, $\boldsymbol{G}_\gamma$ の距離は等しい、すなわち E^0[ln

図1　Vuong（1989）検定の基本的な考え方



（注）　θ_* と γ_* は pseudo-true value、$\hat{\theta}_n$ と $\hat{\gamma}_n$ は最尤推定量を表している。

$f(\cdot|\cdot\,;\boldsymbol{\theta}_*)]=E^0[\ln g(\cdot|\cdot\,;\boldsymbol{\gamma}_*)]$ を意味している。このとき①Non-nesedモデルの場合は、H_0 の下で $f(\cdot|\cdot\,;\boldsymbol{\theta}_*)=g(\cdot|\cdot\,;\boldsymbol{\gamma}_*)$ はありえないので、$f(\cdot|\cdot\,;\boldsymbol{\theta}_*)\neq g(\cdot|\cdot\,;\boldsymbol{\gamma}_*)$ の検定統計量を用いる。一方、②Overlappingモデルの場合は、H_0 の下で、$f(\cdot|\cdot\,;\boldsymbol{\theta}_*)=g(\cdot|\cdot\,;\boldsymbol{\gamma}_*)$ と $f(\cdot|\cdot\,;\boldsymbol{\theta}_*)\neq g(\cdot|\cdot\,;\boldsymbol{\gamma}_*)$ の両方の場合が存在する。そこで、最初に $f(\cdot|\cdot\,;\boldsymbol{\theta}_*)=g(\cdot|\cdot\,;\boldsymbol{\gamma}_*)$ を検定し、それが棄却された後に $f(\cdot|\cdot\,;\boldsymbol{\theta}_*)\neq g(\cdot|\cdot\,;\boldsymbol{\gamma}_*)$ の検定統計量を用いるという手続きが必要である。最後に③Nestedモデルの場合は、H_0 の下で $f(\cdot|\cdot\,;\boldsymbol{\theta}_*)\neq g(\cdot|\cdot\,;$

$\boldsymbol{\gamma}_*$) はありえないので、$f(\cdot|\cdot\,;\boldsymbol{\theta}_*)=g(\cdot|\cdot\,;\boldsymbol{\gamma}_*)$ の検定統計量を用いる。

（4） 各モデルにおける Vuong 検定の手続き

① Non-nested モデル

Non-nested モデルの場合は、$f(\cdot|\cdot\,;\boldsymbol{\theta}_*)\neq g(\cdot|\cdot\,;\boldsymbol{\gamma}_*)$ であることが前提であるので、定理 1（ii）を用いることができる。

定理 2（Non-nested モデルのＬＲ検定）：定理 1（ii）より、適当な正則条件の下で、

（i） under H_0: $n^{-1/2}LR_n(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\gamma}}_n)/\hat{\omega}_n \xrightarrow{d} N(0,1)$,

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

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

が成り立つ。

② Overlapping モデル

Overlapping モデルの場合は、最初に $f(\cdot|\cdot\,;\boldsymbol{\theta}_*)=g(\cdot|\cdot\,;\boldsymbol{\gamma}_*)$ かどうかを検定する必要がある。$f(\cdot|\cdot\,;\boldsymbol{\theta}_*)=g(\cdot|\cdot\,;\boldsymbol{\gamma}_*)$ は、$\omega_*^2=0$ と同値であるので、帰無仮説 H_0^ω：$\omega_*^2=0$、対立仮説：H_A^ω：$\omega_*^2\neq 0$ を検定する。

定理 3（Overlapping モデルのＬＲ分散検定）：定理 1（i）より、適当な正則条件の下で、

（i） under H_0^ω：$n\hat{\omega}_n^2 \xrightarrow{d} M_{p+q}(\,\cdot\,;\hat{\boldsymbol{\lambda}}_n^2)$,

（ii） under H_A^ω：$n\hat{\omega}_n^2 \xrightarrow{a.s.} +\infty$

が成り立つ。但し、$\hat{\boldsymbol{\lambda}}_n^2$ は W の $p+q$ 個の固有値の二乗を要素とするベクトルである。帰無仮説 H_0^ω が適当

58

な有意水準で棄却されないときは、与えられたデータからは、F_θ と G_γ は区別不可能であると結論付ける。一方、帰無仮説 H_0^ω が棄却されたときには、$f(\cdot|\cdot;\boldsymbol{\theta}_*)\neq g(\cdot|\cdot;\boldsymbol{\gamma}_*)$ であるとして、定理2を用いて、帰無仮説 H_0 に対して、対立仮説 H_f または H_g を検定する。

③ Nested モデル

Nested モデルの場合は、$E^0[\ln g(\cdot|\cdot;\boldsymbol{\gamma}_*)]\leq E^0[\ln f(\cdot|\cdot;\boldsymbol{\theta}_*)]$ であり、帰無仮説 H_0 の下で、$f(\cdot|\cdot;\boldsymbol{\theta}_*)=g(\cdot|\cdot;\boldsymbol{\gamma}_*)$ が常に成立する。従って、$f(\cdot|\cdot;\boldsymbol{\theta}_*)=g(\cdot|\cdot;\boldsymbol{\gamma}_*)$ を検定すればよいのだが、overlapping モデルの場合と同様に、$f(\cdot|\cdot;\boldsymbol{\theta}_*)=g(\cdot|\cdot;\boldsymbol{\gamma}_*)$ は $\omega_*^2=0$ と同値であるので、帰無仮説 $H_0^\theta:\omega_*^2=0$、対立仮説 $H_A^\theta:\omega_*^2\neq 0$ を検定する。

定理4（Nested モデルのＬＲ分散検定）：定理1(i)より、適当な正則条件の下で、

(i) under H_0^θ：$n\hat{\omega}_n^2 \xrightarrow{d} M_p(\cdot;\underline{\lambda}_n^2)$,

(ii) under H_A^θ：$n\hat{\omega}_n^2 \xrightarrow{a.s.} +\infty$

が成り立つ。但し、$\hat{\underline{\lambda}}_n^2$ は $\underline{W}=(-B_fA_f^{-1}+B_fR'A_g^{-1}R)$ の p 個の固有値の二乗を要素とするベクトルであり、R は $B_g=RB_fR'$ を満たす $q\times p$ の制約行列である。この場合、帰無仮説 H_0^θ が棄却されるということは、$\boldsymbol{F}_\theta$ は $\boldsymbol{G}_\gamma$ よりも優れていることを意味している。

三　Vuong（1989）検定の応用

（1）増分情報内容と相対情報内容

ある会計情報の有用性は、その情報が情報内容を有するかどうかによって検証される。そして情報内容に

は、相対情報内容と増分情報内容が存在する。今、会計情報XとZがあるとすると、相対情報内容は、XとZのどちらの情報内容が大きいかを検証するものであり、増分情報内容は、Xを所与としたときにZに、あるいはZを所与としたときにXに、追加的な情報内容があるかを検証するものである。相対情報内容を調べるには、(3a)と(3b)のR^2やAICといった、何らかのモデル選択基準を用いて比較すれば良いのだが、もっと積極的にモデル間の優劣を決定したい場合がある。そこでモデル間の優劣の差を統計的に検証するために、Vuong (1989) 検定やDavidson and MacKinnon (1981) のJ検定などが用いられるのである。一方、増分情報内容を調べるには、(5b)のXとZの係数であるβ_1とβ_2の有意性をそれぞれ検証すればよい。

(2) J検定の問題点

J検定は、non-nestedモデルの検定として、その使用の簡便性ゆえに、相対情報内容の検定にしばしば用いられる。ところがJ検定は、検定の理論上、増分情報内容と相対情報内容を明確に区分し得ないので、J検定を相対情報内容の検証に用いることは不適切である。そこで、以下でJ検定の概略を示す。

J検定によるnon-nestedモデル検定は、包括アプローチ (comprehensive approach) と呼ばれるもので、最初に、競合するモデル $\boldsymbol{F}_\theta=\{f(\boldsymbol{y}|\boldsymbol{x};\boldsymbol{\theta});\ \boldsymbol{\theta}\in\Theta\}$ と $\boldsymbol{G}_\gamma=\{g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\gamma});\ \boldsymbol{\gamma}\in\Gamma\}$ を特殊例として含む包括モデル $\boldsymbol{C}_\lambda$ を構成する[3]。この際、$\lambda=0$, $\lambda=1$ がそれぞれ $\boldsymbol{F}_\theta$, $\boldsymbol{G}_\gamma$ に対応するようにすれば、仮説検定 H_f, H_g は、それぞれ $H_f: \lambda=0$, $H_g: \lambda=1$ と書き換えられる。C_λ の構成法は無限にあるが、その中の一つに、

$$C_\lambda=\{c(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{z};\boldsymbol{\theta},\boldsymbol{\gamma});\ \boldsymbol{\theta}\in\Theta,\ \boldsymbol{\gamma}\in\Gamma\}=\frac{f(\boldsymbol{y}|\boldsymbol{x};\boldsymbol{\theta})^{1-\lambda}g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\gamma})^{\lambda}}{\int f(\boldsymbol{y}|\boldsymbol{x};\boldsymbol{\theta})^{1-\lambda}g(\boldsymbol{y}|\boldsymbol{z};\boldsymbol{\gamma})^{\lambda}dy}$$

がある。今、正規線型回帰モデル $\boldsymbol{F}$ と $\boldsymbol{G}$ を次のように定め、non-nestedモデルによる仮説検定を行うとする。

$$H_f: \boldsymbol{y}=\mathbf{X}\boldsymbol{\alpha}+\boldsymbol{u}_f, \quad \boldsymbol{u}_f \sim N(\mathbf{0}, \sigma^2\mathbf{I}), \tag{6}$$

$$H_g: \boldsymbol{y}=\mathbf{Z}\boldsymbol{\beta}+\boldsymbol{u}_g, \quad \boldsymbol{u}_g \sim N(\mathbf{0}, \omega^2\mathbf{I})。 \tag{7}$$

このとき包括モデル C_λ は、

$$H_c: \boldsymbol{y}=\left\{\frac{(1-\lambda)v^2}{\sigma^2}\right\}\mathbf{X}\boldsymbol{\alpha}+\left\{\frac{\lambda v^2}{\omega^2}\right\}\mathbf{Z}\boldsymbol{\beta}+\boldsymbol{u}_\lambda, \quad \boldsymbol{u}_\lambda \sim N(\mathbf{0}, v^2\mathbf{I}), \text{ where } v^{-2}=\frac{1-\lambda}{\sigma^2}+\frac{\lambda}{\omega^2}$$

となる。ここで $\kappa=\lambda v^2/\omega^2$ とおくと前式は、

$$H_c: \boldsymbol{y}=(1-\kappa)\mathbf{X}\boldsymbol{\alpha}+\kappa\mathbf{Z}\boldsymbol{\beta}+\boldsymbol{u}_\lambda \tag{8}$$

と表される。v^2, σ^2 は共に正であるので、$H_c: \lambda=0$ と $H_c: \kappa=0$ は同値である。しかしながらこのままでは、κ と $\boldsymbol{\beta}$ が分離されないので、J検定では、(8)の $\boldsymbol{\beta}$ を(7)の推定値 $\hat{\boldsymbol{\beta}}$ と固定して κ についての推定を行うのである。つまり、$H_c: \boldsymbol{y}=\mathbf{X}\boldsymbol{\alpha}+\kappa\mathbf{Z}\hat{\boldsymbol{\beta}}+\boldsymbol{u}_\lambda$ で $\kappa=0$ against $\kappa \neq 0$ を検定するのである。このように、J検定では、本来検定したかった H_f against H_g ではなく、H_f against H_c という人工的な nested モデルを検証しているのである。同様に、H_g against H_f の検定においても、実際には、H_g against H_c を検証しているのである。

次に、J検定と情報内容との関係について述べる。(3a)と(3b)による相対情報内容の検証について、X と Z がそれぞれ一つの変数であるとすると、J検定では、H_x against H_z は、H_x against H_c,

$$H_c: Y_t=\alpha_0+\alpha_1 X_t+\gamma(\hat{\beta}_0+\hat{\beta}_1 Z_t)+\varepsilon_t$$

の検定となる。このとき H_c は、

$$\begin{aligned} H_c: Y_t &= (\alpha_0+\gamma\hat{\beta}_0)+\alpha_1 X_t+\gamma\hat{\beta}_1 Z_t+\varepsilon_t \\ &= \delta_0+\delta_1 X_t+\delta_2 Z_t+\varepsilon_t \end{aligned}$$

61

に変形されるので、結局この検定は、Z_t の係数 $\delta_2=0$ を検定するのと同じになる。同様に、H_z against H_x では、J検定では、X_t の係数 $\delta_1=0$ を検定することになる。つまり、相対情報内容の検証にJ検定を用いるのは、増分情報内容を検証するのと同じになってしまうのである。

（3）会計利益とキャッシュフローの情報内容

Vuong検定の応用例として、二〇〇〇年三月決算期から新たに開示されている連結キャッシュフロー情報と、発生主義に基づく従来の会計利益情報の情報内容を比較する。サンプルは、二〇〇〇―二〇〇二年の期間において、（i）東証一部上場企業である、（ii）一般事業会社である、（iii）会計期間が一二ヶ月である、という基準で選択されている。使用するモデルは、次のようである。

Non-nested モデル（相対情報内容の検証）

$$AR_{it}=\alpha_0+\alpha_1E_{it}+\alpha_2\Delta E_{it}+\varepsilon_{it}, \quad (9)$$

$$AR_{it}=\beta_0+\beta_1CF_{it}+\beta_2\Delta CF_{it}+\varepsilon_{it}, \quad (10)$$

Overlapping モデル（相対情報内容の検証）

$$AR_{it}=\alpha_0+\alpha_1SIZE_{it}+\alpha_2BM_{it}+\alpha_3E_{it}+\alpha_4\Delta E_{it}+\varepsilon_{it}, \quad (11)$$

$$AR_{it}=\beta_0+\beta_1SIZE_{it}+\beta_2BM_{it}+\beta_3CF_{it}+\beta_4\Delta CF_{it}+\varepsilon_{it}, \quad (12)$$

Nested モデル（増分情報内容の検証）

$$AR_{it}=\alpha_0+\alpha_1E_{it}+\alpha_2\Delta E_{it}+\varepsilon_{it}, \quad (9)$$

$$AR_{it}=\beta_0+\beta_1E_{it}+\beta_2\Delta E_{it}+\beta_3CF_{it}+\beta_4\Delta CF_{it}+\varepsilon_{it}, \quad (13)$$

AR_{it} は企業・i の t 期四ヵ月目から $t+1$ 期三ヶ月目までの市場調整済み年次異常リターン、E_{it} は企業・i の t 期における経常利益、CF_{it} は企業・i の t 期における営業キャッシュフロー、$\Delta E_{it}=E_{it}-E_{it-1}$, $\Delta CF_{it}=CF_{it}-CF_{it-1}$, $SIZE_{it}$ は企業・i の t

期首における時価総額に自然対数をとったもの、BM_{it} は企業 i の t 期首における株主資本簿価/時価総額の比率である。また E_{it}, CF_{it}, ΔE_{it}, ΔCF_{it} は t 期首の時価総額でデフレートされている。

（4）実　証　結　果

Non-nested モデル（相対情報内容）の検証結果

Non-nested モデルの推定結果は、表1の上段に示されている。(9)の $E_{it}+\Delta E_{it}$ の係数（t値）は一・一五六（七・五七）、(10)の $CF_{it}+\Delta CF_{it}$ は〇・一九四（三・五五）と、有意に正である。但し adj. R^2（一四・四％ vs 五・五％）からは、(9)の優位性が伺える。そこで(9)と(10)の相対情報内容を統計的に比較するために、non-nested モデルの Vuong 検定を行った結果が表2の上段に示されている。Non-nested 検定では、定理2の標準化されたLRを用いて検定を行う。(9) against (10)で得られた検定統計量は三・四七五と有意であり、(9)の方が(10)よりも真のモデルにより近いことが判る。

Overlapping モデル（相対情報内容）の検証結果

Overlapping モデルの推定結果は、表1の中段に示されている。バリュー効果(BM_{it})、規模効果($SIZE_{it}$)の影響を除去した後でも、(11)の $E_{it}+\Delta E_{it}$ の係数（t値）は一・一一〇（七・二八）、(12)の $CF_{it}+\Delta CF_{it}$ は〇・一八二（三・三二）と、共に有意に正である。但し adj. R^2（一八・六％ vs 一〇・二％）からは、(11)の優位性が伺える。そこで(11)と(12)の相対情報内容を統計的に比較するために、overlapping モデルの Vuong 検定を行った結果が表2の中段に示されている。Overlapping 検定は、逐次検定であるので、最初に定理3に基づいて（i）LR分散検定を行い、それが棄却されたら定理2に基づいて（ii）LR検定を行うという手続きが必要である。なおLR分散は重み付き χ^2 和分布に従い、その棄却限界値を分析的に計算するのは困難であるので、シミュレーションによってその値を求めている。(11) against (12)で得られたLR分

表1　回帰モデルの推定結果

推定モデル		*Inrtercept*	$SIZE_{it}$	BM_{it}	$E_{it}+\Delta E_{it}$	$CF_{it}+\Delta CF_{it}$	adj. R^2
Non-nested モデル	(9)	0.121 (10.34)**			1.156 (7.57)**		14.4%
	(10)	0.136 (13.43)**				0.194 (3.55)**	5.5%
Overlapping モデル	(11)	0.077 (1.32)	−0.003 (−0.70)	0.075 (7.11)**	1.110 (7.28)**		18.6%
	(12)	0.116 (2.10)*	−0.006 (−1.30)	0.076 (6.83)**		0.182 (3.31)**	10.2%
Nested モデル	(13)	0.108 (9.42)**			1.058 (9.23)**	0.136 (2.52)*	16.3%

（注）　括弧内は White の共分散行列に基づく t 値である。*5%水準で有意　**1%水準で有意。

表2　Vuong（1989）検定の結果

競合モデルの種類	Vuong（1989）モデル選択検定	
	（i）LR 分散検定	（ii）LR 検定
Non-nested モデル（9）　against（10）		3.475**
Overlapping モデル（11）　against（12）	1018.97**	3.437**
Nested モデル　（9）　against（13）	436.54*	

（注）　＊ 5%水準で有意　＊＊ 1%水準で有意。

散の検定統計量は、一〇一八・九七と一％水準で棄却される。そこで（ii）LR検定を行うと、その検定統計量は三・四三七と有意であり、(11)の方が(12)よりも真のモデルにより近いことが判る。

Nested モデル（増分情報内容）の検証結果

Nested モデルの推定結果は、表1の下段に示されている。(13)の $E_{it}+\Delta E_{it}$ および $CF_{it}+\Delta CF_{it}$ の係数（t 値）は、各々一・〇五八（九・二三）、〇・一三六（二・五二）と、共に有意に正である。このことは、会計利益とキャッシュフローには、互いに相手を所与としても、それぞれに増分情報内容が存在することを示している。増分情報内容の検証は、この係数推定値の有意性検定でも良いが、ここでは nested モデルの場合の Vuong 検定を行っており、その結果が表2の下段に示されている。Nested 検定では、定理4に基づいてLR分散検定を行う。(9) against (13)で得られた検定統計量は、四

三六・五四と五％水準で有意である。これは、(9)よりも(13)の方が真のモデルにより近いということを示している。なお(10) against (13)でも、同様の結果を得ている。

四　おわりに

Vuong 検定とは、モデル選択基準にKLICを用いてそれをLR検定に応用することにより、R^2やAICといった従来のモデル選択基準では不可能であった競合するモデル間の統計的有意検定を可能にした、モデル選択検定である。そしてVuong 検定は、その枠組みの一般性ゆえに、会計における情報内容の検証に用いるのに適している。一方、J検定は、理論上、相対情報内容と増分情報内容を区別し得ないので、情報内容の検証に用いるのには不適切である。

Vuong 検定の応用例としては、二〇〇〇年三月決算期から開示されている連結キャッシュフローと従来の発生主義に基づく会計利益の情報内容の比較を行っている。結果は、相対情報内容では、会計利益がキャッシュフローより優っているが、キャッシュフローにも、若干の増分情報内容が存在した。つまり、キャッシュフローは会計利益に取って代わるものではないが、幾らかの追加的な情報を市場に提供しているようである。

最後に、Vuong 検定は、その枠組みの一般性および明瞭性の点で、従来の検定よりも優れているといえる。しかしながら、その検定統計量の計算が非常に複雑であるので、現段階ではその有用性が十分に発揮されていない。会計情報には代替的情報が数多く存在し、研究上モデル選択検定を必要とする機会も少なくない。今後、検定の改良や簡便法が考案されれば、Vuong 検定は、会計学研究において非常に有用なツールとなり得るであろう(4)。

注

（1） 本稿は、太田・松尾（二〇〇四）に基づいているので、詳細についてはそちらを参照されたい。
（2） pseudo-true value の詳細については、Sawa（1978）や White（1982）を参照されたい。
（3） 包括アプローチ等の non-nested 仮説検定の考え方については、Pesaran and Weeks（2001）を参照されたい。
（4） Vuong（1989）検定の簡便法については、太田・松尾（二〇〇四、Appendix 4）と松尾（二〇〇四）を参照されたい。

引 用 文 献

太田浩司・松尾精彦［二〇〇四］「Vuong（1989）検定の理論と応用」武蔵大学論集、第五二巻第１号、三九—七五頁。

松尾精彦［二〇〇四］「Vuong test とその正規線形モデルへの適用法」関西大学経済論集、第五四巻第１号、三九—七五頁。

Davidson, R. and J. MacKinnon. (1981). Several tests for model specification in the presence of alternative hypotheses. *Econometrica* 49, 781-793.

Pesaran, M. H. and M. Weeks. (2001). Nonnested hypothesis testing: an overview. *A companion to theoretical econometrics*, edited by B. H. Baltagi. Blackwell Publishers Ltd, Oxford, UK.

Sawa, T. (1978). Information criteria for discriminating among alternative regression models. *Econometrica* 46, 1273-1291.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57, 307-333.

White, H. (1982). Maximum likelihood estimation of misspecifeid models. *Econometrica* 50, 1-25.

（付記） Vuong 検定の具体的な実施方法は、http://homepage2.nifty.com/koji_ota/から Excel 形式でダウンロード可能である。

（付記） 本稿は日本会計研究学会第六十三回大会における自由論題報告に加筆修正したものである。

（太田・武蔵野大学専任講師）
（松尾・関 西 大 学 教 授）