

実証分析を行う前に理解しておきたいこと
貴重なアドバイスそれとも大きなお世話？

福井義高

東北大学大学院経済学研究科・経済学部

fukui@econ.tohoku.ac.jp

<http://www.econ.tohoku.ac.jp/~fukui/site1.htm>

2001 年 7 月

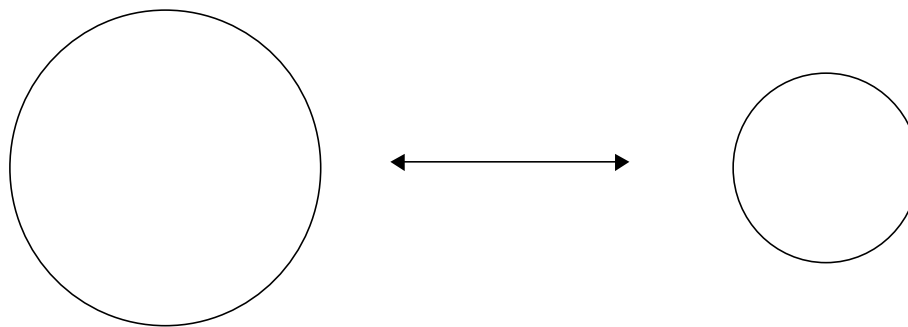
0. はじめに

1. 会計実証研究者にわかりやすく統計理論応用における問題点を解説
漸近理論に基づいた現代的統計理論の考え方を簡単に解説する.
仮説検定の背後にある考え方をややくわしく解説する。
2. 個別の手法ではなく考え方を解説
Cookbook-style の教科書でさまざまな手法を学ぶより、背後にある考え方をよく理解することを目指す。難しい数学は不要。

1. 標本から母集団の特性を推測 (Inference)

母集団 (Population)

標本 (Sample)



母集団確率モデルの仮定 $f(X=x \mid \theta = \hat{\theta})$ $\Rightarrow$ 特定の標本実現 $\Rightarrow \theta$ の推定.
(up to parameter values) (estimation)

2. 幾何学的理解：最小二乗法はピタゴラスの定理 最小二乗法は統計モデルではなく代数モデル

$$y = X\beta + e = (1, x_1, \dots, x_K)\beta + e$$
$$\begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_N \end{pmatrix} = \beta_0 \begin{pmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ 1 \end{pmatrix} + \beta_1 \begin{pmatrix} x_{11} \\ x_{21} \\ \cdot \\ \cdot \\ x_{N1} \end{pmatrix} + \dots + \beta_K \begin{pmatrix} x_{1K} \\ x_{2K} \\ \cdot \\ \cdot \\ x_{NK} \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ e_N \end{pmatrix}$$

- ①最小二乗法：データの直和分解 $\Rightarrow$ 誤差項と「説明」変数は無相関.
- ②少ない次元でどれだけ説明できるか. 多重共線性(multicollinearity)はベクトルが一次従属であることに過ぎない. $K+1=N$ ならば誤差ゼロ.
- ③ R^2 は y と $X\beta$ がなす角の $\cos^2\theta$.
- ④この代数的分解にどのような意味を与えるかは統計の問題.

3. 小標本から大標本理論へ

1. 漸近(asymptotic)理論に基づく推定・検定

小標本理論は分散均一(conditional homoskedasticity)な正規分布、説明変数の外生性(exogeneity)等、経済データでは通常成り立っていない仮定を前提としている．こうした制約的仮定を設けず、大標本に基づく漸近理論による近似を用いるのが現在の主流．従って、検定には $F \cdot t$ 分布ではなく χ^2 ・正規分布を用いる．Estimator の望ましい特性は一致性(consistency)．

2. OLS の復権

① $E(x_i e_i) = 0$ (orthogonality/predeterminedness condition)の仮定（時系列の場合は加えて定常性とエルゴード性）があれば、estimator は consistent となる．

② 分散不均一は大きな問題ではない

⇒ 追加的仮定が必要な GLS ではなく、
分散推定に White estimator を用いる．

3. 残った問題：内生性 (endogeneity)と系列相関 (serial correlation)

① 内生性は操作変数 (instrumental variables)を用いることで、orthogonality conditionをつくる．説明変数より操作変数が多い場合を一般的に処理する方法が、GMM と考えてもよい．

②時系列分析で AR モデルを用いる大きな理由は系列相関の回避．Durbin-Watson 検定は小標本の制約的仮定が必要なので、時系列の場合は invalid. また、誤差の系列相関を Cochrane-Orcutt 法で推定するのは、事実上、モデルに大きな制約を加えているのと同じ (common factor の仮定).

4. GMM知ったかぶり

1. GMM Instrumental Variables + White estimator

GMM は orthogonality condition を満たす操作変数が説明変数より多い場合に (over-identified)、「望ましい」重み付けを行う手法. 具体的には、共分散行列の逆行列が必要. そこで、まず 2LS で推定し、その残差を用い White estimator による共分散行列を得る. 次にこの共分散行列を weight matrix に用いる.

2. OLS $\subset$ GMM

Just-identified の場合は、狭義の操作変数法となる. かつ、説明変数があるまま操作変数となる場合が OLS. すなわち OLS は GMM の一種である.

5. 統計的有意性(Statistical Significance)の意味

1. $P(h(\theta) | X)$ ではなく $P(X | h(\theta))$

「データは有意水準 5%で棄却されたので、 $h(\theta) \neq \hat{h}(\theta)$ の確率（信頼度？）は 95%」等という言明は誤り．また、significance level が文字通り解釈できるのは specification search (regression fishing)していない (= correctly specified model) という前提．

ちょうど 5%水準で棄却された後で、同じ標本数の別のデータで test すると棄却される確率はどれくらい？

2. Probability ではなく confidence：本当は Bayesian?

頻度論(frequentist)に基づく統計理論に事後の確率はない。そこで事後には confidence という概念が導入される．ただし、Bayesian の立場で diffuse prior をとると、上記のような言明は可能。

3. Significance と power

Fisher の p -value を批判して提唱されたのが、power 概念を導入した Neyman-Pearson の decision-theoretic approach. ただし、データ数が多くなれば、power は 100% に近づき $\theta = \hat{\theta}$ は必ず棄却される. 係数が正確にゼロであるかどうかのテストは、データ数が増えた現在、significance level が小さくても、事前に棄却されることがわかっている！

4. Likelihood ratio

標本数に応じて significance level を変える test (例えば likelihood ratio のハードルを一定とすればそうなる)の方が好ましい. Neyman-Pearson のフレームワークは標本数が増えると棄却が容易になる likelihood ratio test の一種.

5. 効果あり／なし検定の非対称

ある係数がゼロである（効果がない）ことを否定する場合も、ゼロでない（効果がある）ことを否定する場合も、同じ $P(X | \theta_1 = 0) = 5\%$ を基準に使うと test が非対称的になる．前者に 5%を使うのなら後者には 95%を使ったほうがよい．

6. 命題・仮説を test することの意味

- ①Fisher の objective で epistemic な理解．
- ②Neyman-Pearson の objective で behavioral な理解．
- ③Bayesian の inter-subjective で epistemic かつ behavioral な理解．

6. モデル選択の基準

1. 自由度修正済 $\bar{R}^2$ と AIC (赤池情報量基準)

R^2 は説明変数を追加すると必ず「改善」する. そこで Theil が考案したのが $\bar{R}^2$. $\bar{R}^2$ の大小が変数追加の基準 $\Leftrightarrow$ 変数の t 値の 1 との大小が基準.

$\bar{R}^2$ によるモデル選択は、 s^2 すなわち $\log(s^2)$ による選択となる：

$$\log(s^2) = \log(\hat{\sigma}^2 \cdot \frac{N}{N-K-1}) \cong \log(\hat{\sigma}^2) + \frac{K+1}{N}$$

一方、Kullback-Leibler 情報量に基づく AIC の場合は、

$$\log(\hat{\sigma}^2) + 2\frac{K+1}{N}$$

の大小比較に帰着する. 二つの基準は意外に近く、ともに inconsistent で overfit をもたらす. AIC を修正した次の Schwartz 基準は consistent.

$$\log(\hat{\sigma}^2) + \log(N) \frac{K+1}{N}$$

2. 分析の目的とモデル選択

- ① Estimation の良否と $\bar{R}^2$ の大小は無関係.
- ② 分析の目的が予測だとしても、 $\bar{R}^2$ は nested models の比較でなければ意味をなさないし、むしろ予測の良否は out-of-sample の performance で決まる．データ数が多い場合は cross validation が有効.

3. モデル選択は science というより art

頻度論に基づく統計理論では、specification search の理論化は困難.

7. 終わりに

1. Hypothesis testing より estimation

近年最も活発な議論を呼んだ Mehra·Prescott の equity premium puzzle は significance ではなく estimates の大小が議論の中心.

2. Calibration (simulation)の重要性

Real business cycle 理論の Kydland·Prescott は hypothesis testing の無意味さを主張し、calibration 中心の実証分析を提唱.

3. データ数の増加は non-parametric estimation や bootstrap への動きを加速 ⇒ Conditional expectation を直接推定する時代の到来？

4. 数値の大きさを議論しよう

会計こそ具体的数値の大小を扱う分野であることを考えれば、significance への digression を控えて、具体的な estimates を議論の中心に.

Appendix. Econometrics Textbooks

1. Introductory

Amemiya (1994): Written for theory-oriented beginners.

Goldberger (1998): Introductory version of Goldberger (1991).

森棟 (1999): A Japanese textbook comparable to Pindyck and Rubinfeld (1997).

Pindyck and Rubinfeld (1997): Application-oriented. Targeted at advanced MBAs.

2. Basic (Used in introductory Ph.D. econometrics courses)

Goldberger (1991): Unconventional and stimulating in many respects.

Greene (1999): Benchmark? Comprehensive and widely used among applied researchers.

Hansen (2001): Lecture notes used at Wisconsin. Concisely and rigorously written.

Poirier (1995): Written from a Bayesian point of view.

3. Advanced

Davidson and MacKinnon (1993): Widely used in theory-oriented Ph.D. courses.

Hamilton (1994): *The* text of time-series analysis.

Hayashi (2000): Next generation's orthodoxy unified under GMM.

- Amemiya, T. 1994. *Introduction to Statistics and Econometrics*. Cambridge, MA: Harvard University Press
- Davidson, R., and J. G. MacKinnon. 1993. *Estimation and Inference in Econometrics*. Oxford, U.K.: Oxford University Press.
- Goldberger, A. S. 1991. *A Course in Econometrics*. Cambridge, MA: Harvard University Press.
- Goldberger, A. S. 1998. *Introductory Econometrics*. Cambridge, MA: Harvard University Press.
- Greene, W. H. 2000. *Econometric Analysis*. Upper Saddle River, NJ: Prentice Hall.
- Hamilton, J. D. 1994. *Time Series Analysis*. Princeton, NJ: Princeton University Press.
- Hansen, B. E. 2001. *Lecture Notes for Economic Statistics and Econometrics II* (<http://www.ssc.wisc.edu/~bhansen/710/710.htm>). University of Wisconsin, Madison, WI.
- Hayashi, F. 2000. *Econometrics*. Princeton, NJ: Princeton University Press.
- 森棟公夫. 1999. 『計量経済学』. 東洋経済新報社.
- Pindyck, R. S., and D. L. Rubinfeld. 1997. *Econometric Models and Economic Forecasts*. New York, NY: McGraw Hill.
- Poirier, D. J. 1995. *Intermediate Statistics and Econometrics: A Comparative Approach*. Cambridge, MA: MIT Press.