

Forecasting Welfare Caseloads: A Case of Japanese Public Assistance

By Masayoshi Hayashi

へのコメント

大阪大学 福重元嗣

まずは一般論から、バタフライ効果のお話し、多くの方がご存知のように、バタフライ効果というのは、複雑系とか、カオスとか言ったトピックの中で、ブラジル（アマゾンのジャングル？）の蝶々の羽ばたきが、テキサスの竜巻になるというお話です。これは、高次の非線形微分方程式のお話しなわけですが、この論文で扱っているトピックは、高次の非線形差分方程式のお話しなわけで、マクロ経済学でよく出てくるオイラー方程式の元となっている、一階の線形微分方程式あるいはその差分による近似よりもバタフライ効果によく似たことが起きやすいと考えられる問題であります。ということで、長期の予測に関するお話は、まじめにやれば非常に難しいお話になります。そこで、理論的なことはさておき、現実的な問題として、複数の予測モデルを基にした予測の実用性、さらには予測の結合という手法を使って、生活保護の件数の予測を行うという、実用的であり、興味深い研究を行っているのがこの論文であります。ということは、統計学的に見て細かな点という視点より深く追及するのではなく、より実用性に関してコメントをすべきであると考えます。

とは言っても、統計学や計量経済学を普通に学んできた者にとっては、真のモデルが見つければ、一番正しい予測が出来そうだと考えがちです。そこで、最低限の統計的な話題として、不偏予測量(unbiased predictor)のお話を、第2のトピックとして取り上げたいと思います。生活保護費のレベルを予測するモデルではなく、対数変換したり、さらに差分を取ったりした後で、モデルを構築し予測をすることによって、変数変換された後の変数の不偏予測量の値を得ることは出来そうです。しかしながら、その変数を逆の変換、和分を取る（要は積み上げて元の変数を作る）操作や、指数変換をすると、それは不偏予測量ではなくなります。もちろん、これは、予測に使ったモデルが真のモデルであった場合になるので、予測の結合を論文の後半で扱う時点で、どのモデルが真のモデルであるとか、あるいは真のモデルが存在するとかいったことは全く信じておられない可能性が高いので、この議論は意味がないものかもしれません。もちろん、真のモデルがあっても、縮減推定量 (shrinkage estimator) の議論や、近年のモデルアベレージング (model averaging) のような、不偏性ではなく最小二乗誤差やそのほかの基準を用いた議論は可能ですが、たぶんこの論文はもっと実用的に予測が出来そうかどうかに関心を絞っておられるのだと考えます。（もちろん、評価の基準として、平均絶対誤差 (MAE) や、平均2乗誤差 (RMSE)、平均誤差 (ME) などの基準は用いられております。）もしもそうであれば、予測の結合を考えると、CF1 から CF3 のような、過重平均にこだわる必要はなく、CF4 のような、加重平均の概念を飛び出した自由な予測方法をもっと色々と考えられた方が、より実用的なものが出てきた可能性は高いのではなかったか、と思うと少し残念ではあります。翻って、不偏予測量を気にする者の視点から見れば、CF4 の α の統計的な有意性や、 β の合計がいくらになったのだろうか、という興味深い点がレポートされておらず、ちょっと見てみたい結果ではあります。

さて第 3 のトピックに入りますが、ここからがコメントの本題です。まずは節の構成ですが、2 節で Markov forecasting を独立に取り上げ、3 節では、指数平滑化、ARIMA モデル、LSTAR モデル、VAR 更には予測の結合と盛り沢山です。Markov forecasting が、実際の生活保護費の予測ではほとんど用いられていないにもかかわらず、取り上げ方のウェイトが取り分け大きいのが奇異に見えます。各モデルの取り上げ方について整理されることを期待します。さらに予測の結合については、CF1 と CF2 が 3 節で、CF3 と CF4 が 4 節で登場しますが、これも、どのような考え方で予測の結合があるのか一つの節で整理された方が、より論文の意図が明確になると考えます。

第 4 のトピックは、Table 6 の使い方です。Diebold=Mariano の検定を実用的に用いられていることは、もちろん賛成いたしますが、なぜこの検定を持ち出したのか、もともとの検定の意図である真のモデルを見つけるための検定ではなく、実用的に大きくかけ離れたモデルを見つけるためだと理解しますが、この点について、本文中で説明を追加されることを勧めます。

第 5 のトピックは、Figure です。Figure 8 はよく見れば、どれがどのモデルによる予測なのかを区別出来そうですが、送っていただいたファイルでは特に白黒だったため、Figure 9 は、それぞれの予測値を区別するのが非常に難しく、いくつかの予測値については区別出来ていません。評者の理解が間違っているかもしれませんが、Actual や Markov がどの線なのか、よくわかる Figure にされることを望みます。

最後にいくつか、評者の勉強不足によるものかもしれませんが、いくつか疑問というか質問があります。Table 2 で報告されております、指数平滑化における、SBC は 7 頁 3 行目の BIC、4 行目 SBC との表記がありますが、どのように計算されたものでしょうか？シュワルツのベイジアン情報量基準 (SBIC) であれば観測値数や推定された母数の数を勘定しなければならないのですが、どのように勘定されたのか教えていただけると助かります。もう一点は、同様に 7 頁の 3.2 節の ARIMA モデルの表記方法です。ARIMA(3,1,0)(0,1,0) や ARIMA(p,1,q) (0,1,0) という表記がありますが、後ろの(0,1,0)は、季節調整の差分の表記だと思われそうですが、他の文献では Seasonal ARIMA (SARIMA) といった表記方法も見られますが、どれが一般的な表記方法でしょうか？一般的なものがない場合には、定義された方が、読者にとって分かりやすいと思われそうです。最後の質問ですが、4.2 節の表題となっております Real-time forecasting ですが、この用語法は一般的でしょうか？

以上、評者として、分析内容に関して大きく修正すべき点は殆ど無く、大変興味深い論文であったことを再度明記して、コメントを終わりたいと思います。