

XÂY DỰNG MÔ HÌNH SẢN XUẤT NÔNG NGHIỆP BẰNG PHÂN TÍCH HỒI QUY

PGS .TS Nguyễn Đức Bảo

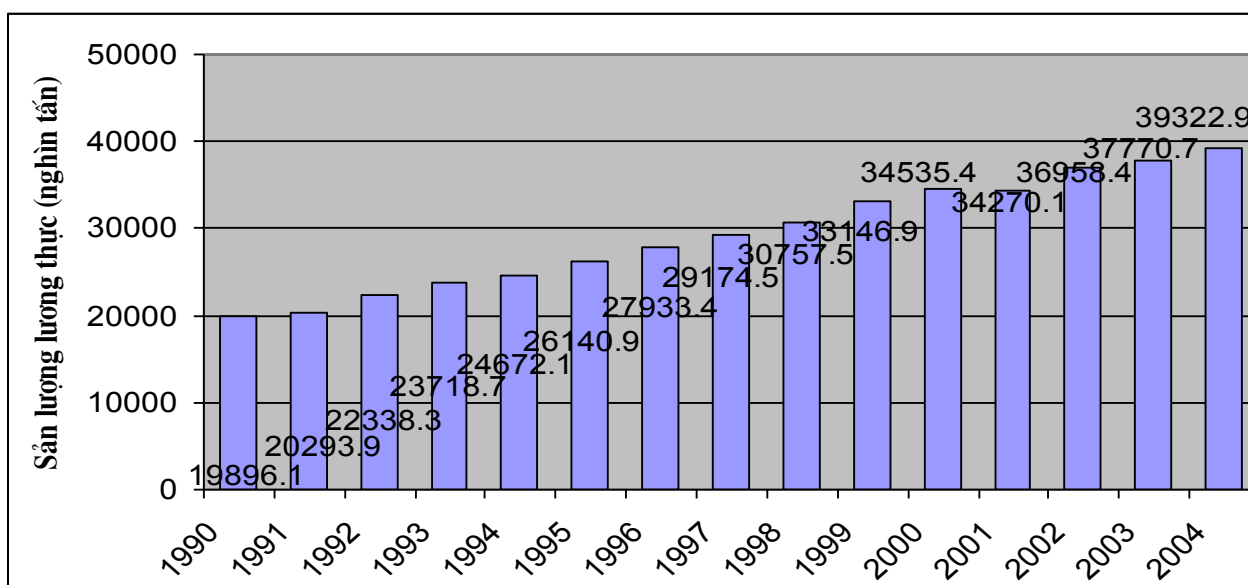
Dip. Math,ThS Nguyễn Thế Hoà

Trường Đại học Thủy lợi

I.- Đặt vấn đề

Sau hơn 15 năm đổi mới, ngành nông nghiệp nước ta về cơ bản đã chuyển sang sản xuất hàng hoá, phát triển tương đối toàn diện, tăng trưởng khá đạt mức tăng bình quân 4,2%/năm, đảm bảo an ninh lương thực, tỷ suất hàng hoá trong nông nghiệp ngày càng cao, một số mặt hàng có thị phần lớn trong khu vực và trên thế giới như gạo, hạt điều, cà phê, hồ tiêu. Bước đầu hình thành một số vùng sản xuất tập trung gắn với công nghiệp chế biến như các vùng lúa gạo ở đồng bằng sông Cửu Long và đồng bằng sông Hồng; các khu công nghiệp- dịch vụ ở trung du, miền núi gắn với việc phát triển cây công nghiệp có lợi thế như khu công nghiệp chè ở trung du, miền núi phía bắc, khu liên hợp công nghiệp cao su ở Đông Nam Bộ, mía đường miền trung Kết cấu hạ tầng cơ sở như thủy lợi, giao thông, điện ... được cải thiện, đời sống nông dân ở hầu hết các vùng đã nâng lên rõ rệt, tỉ lệ đói nghèo giảm từ 29% xuống còn 11%. Thu nhập bình quân một nhân khẩu nông thôn tăng từ 92.100 đồng/tháng năm 1992 lên 356000đồng/tháng năm 2002. Nông nghiệp trở thành nhân tố quan trọng hàng đầu trong sự nghiệp đổi mới, góp phần ổn định kinh tế-xã hội và chính trị ở nước ta. Thắng lợi của phát triển nông nghiệp, nông thôn tạo tiền đề cho nước ta đẩy nhanh sự nghiệp công nghiệp hoá, hiện đại hoá đất nước.

Hình 1: Sản lượng lương thực có hạt trong giai đoạn 1990-2004:



Nguồn: Thời báo kinh tế Việt Nam

Đặc biệt từ khi đổi mới, cùng với sự phát triển kinh tế đầu tư xây dựng cơ sở hạ tầng ngành thủy lợi có những mức tăng đầu tư đáng kể. Nếu như giai đoạn 1955-1975 tổng đầu tư các công trình thủy lợi là 5.631 tỉ đồng, trung bình 281 tỉ đồng/năm, giai đoạn 1976-1985 là 10.848 tỉ đồng, trung bình 1.085 tỉ đồng/năm thì thời kỳ 1986-2000 đã lên đến 24.292 tỉ đồng với mức đầu tư trung bình 1620 tỉ đồng/năm, theo mặt bằng giá năm 2000. Năm 1999 là năm đầu tư cao nhất tới 3127 tỉ đồng.

Vấn đề đặt ra trong nghiên cứu này là trong quá trình chuyển đổi từ nền kinh tế tập trung bao cấp sang nền kinh tế thị trường thì mức đóng góp của ngành thủy lợi vào mức tổng sản lượng lương thực là bao nhiêu? Và trong các năm tới nếu đầu tư thủy lợi tăng lên thì tổng sản lượng lương thực tăng như thế nào? Đã có những nghiên cứu mô hình sản xuất nông nghiệp; trong báo cáo này xin giới thiệu xây dựng mô hình sản xuất nông nghiệp bằng phân tích hồi quy .

II.-Nội dung nghiên cứu

1. Xây dựng mô hình hồi qui lý thuyết về sản xuất nông nghiệp

Hồi qui là một công cụ cơ bản của đo lường kinh tế; ý tưởng trọng tâm của phân tích hồi qui nhiều biến là nghiên cứu mối liên hệ phụ thuộc có tính thống kê của một biến ngẫu nhiên với nhiều biến giải thích khác. Biến phụ thuộc ở đây là tổng sản lượng lương thực mỗi năm của cả nước. Tổng sản lượng lương thực hàng năm phụ thuộc vào rất nhiều yếu tố như mức đầu tư xây dựng thủy lợi, phân bón, diện tích canh tác, giống cây trồng, thời tiết, kỹ thuật canh tác, chính sách kinh tế ... Trong đó có những biến giải thích lượng hoá được như tư xây dựng thủy lợi, phân bón, diện tích canh tác; có những biến không lượng hoá được giống cây trồng, thời tiết, kỹ thuật canh tác, chính sách kinh tế. mô hình. Như chúng ta đã biết trong các biến không lượng hoá được này có những biến như chính sách kinh tế có tác động rất quyết định đến sự phát triển của ngành nông nghiệp nói chung và tổng sản lượng lương thực nói riêng, nó đưa nước ta chuyển đổi từ một nước thiếu ăn thường xuyên trở thành một nước xuất khẩu gạo đứng thứ 2 trên thế giới chỉ sau Thái Lan. Để đo những ảnh hưởng của những biến như vậy ta có thể đưa thêm biến chính sách vào nh thông qua biến giả.

Mục tiêu của phân hồi qui là đi ước lượng và dự báo giá trị trung bình của biến phụ thuộc Y, tổng sản lượng lương thực, dựa trên các giá trị đã biết của biến giải thích; tức là trước hết đi tìm hàm kỳ vọng có điều kiện của biến phụ thuộc với các giá trị đã biết của biến giải thích dưới dạng $E(Y/X_i) = f(X_i)$. Có rất nhiều dạng hàm hồi qui, nhưng phổ biến là 2 dạng hàm sau:

- Dạng hàm hồi qui tuyến tính có k tham số:

$$f(X_i) = a_0 + a_1X_1 + a_2X_2 + \dots + a_{k-1}X_{k-1} \quad (1)$$

- Dạng hàm hồi qui tuyến tính loga có k tham số:

$$\ln f(X_i) = \ln(a_0) + a_1\ln(X_1) + a_2\ln(X_2) + \dots + a_{k-1}\ln(X_{k-1}) \quad (2)$$

Không có một tiêu chuẩn rõ ràng nào để có thể dựa vào đó mà chọn dạng hàm. Người nghiên cứu ít hay nhiều buộc phải lựa chọn một dạng hàm tùy theo nhận thức về lý thuyết của riêng mình với hy vọng rằng sự lựa chọn của anh ta không phản ánh trái ngược lại kết quả của mình.

Như vậy, từng giá trị riêng của biến phụ thuộc Y_i sẽ biến động xung quanh $E(Y/X_i)$ và lệch gọi giá trị trung bình có điều kiện này một lượng u_i , mối quan hệ này có dạng mô hình hồi qui tổng thể:

$$Y_i = E(Y/X_i) + u_i,$$

ta gọi u_i là nhiễu thống kê hay sai số thống kê. Sai số thống kê giải thích những biến động có tác động đến biến phụ thuộc nhưng không đưa vào mô hình hồi qui.

Trong thực tế sự thành công của phân tích hồi qui còn phụ thuộc vào sự sẵn có của số liệu phù hợp. Người ta chỉ có thể nghiên cứu tập tổng thể Y thông qua việc lấy mẫu với một cỡ mẫu n nhất định, và thu được hàm hồi qui mẫu $\hat{y} = E(\hat{Y}/X_i)$ ước lượng thay cho hàm hồi qui tổng thể. Gọi các biến giải thích:

TL là tổng giá trị đầu tư cho các công trình thủy lợi

PB lượng tiêu dùng phân bón vô cơ cho nông nghiệp

D là biến chính sách.

Nếu chọn dạng hàm tuyến tính mô hình hồi qui mẫu về sản xuất nông nghiệp của chúng ta có dạng:

$$Y_i = a_0 + a_1 TL_i + a_2 PB_i + a_3 D_i + u_i \quad ; \quad (3)$$

Trong đó a_0 là hằng số trong hồi qui, a_1, a_2, a_3 phản ánh mức thay đổi của tổng sản lượng lương thực khi tổng giá trị đầu tư cho các công trình thủy lợi và lượng tiêu dùng phân bón vô cơ tăng lên một đơn vị, a_4 là hệ số phản khuynh hướng biên của tổng lương thực tăng lên do có chính sách đổi mới kinh tế, u là sai số phản ánh những ảnh hưởng thứ yếu khác, được giả thiết là không có tương quan với các biến giải thích.

Nếu chọn dạng hàm tuyến tính loga mô hình hồi qui mẫu về sản xuất nông nghiệp của chúng ta có dạng:

$$\ln Y_i = \ln(a_0) + a_1 \ln(TL_i) + a_2 \ln(PB_i) + a_3 D_i + u_i \quad ; \quad (4)$$

Trong đó a_0 là hằng số trong hồi qui, a_1, a_2, a_3 phản ánh phần trăm thay đổi của tổng sản lượng lương thực khi tổng giá trị đầu tư cho các công trình thủy lợi và lượng tiêu dùng phân bón vô cơ tăng lên 1%, a_4 là hệ số phản khuynh hướng biên của tổng lương thực tăng lên do có chính sách đổi mới kinh tế, u là sai số phản ánh những ảnh hưởng thứ yếu khác, được giả thiết là không có tương quan với các biến giải thích. Sự thành công trong phân tích kinh tế nói chung và phân tích hồi qui nói riêng phụ thuộc rất lớn vào việc sử dụng số liệu thống kê thích hợp và phương pháp xử lý các số liệu.

Số liệu càng nhiều thì mô hình qui càng phản ánh chính xác. Trong nghiên cứu này chúng tôi lấy số liệu thống kê theo từng năm từ 1986 đến 2005, tức là cỡ mẫu $n = 20$.

Việc phân tích về mặt định tính mối liên hệ phụ thuộc giữa từng biến thích và biến phụ thuộc trước khi đưa chúng vào mô hình là một bước rất quan trọng. Nó ảnh hưởng trực tiếp đến chất lượng của mô hình sau này; nó tránh cho chúng ta gặp phải hiện tượng hồi qui giả, và giúp ích rất nhiều trong việc khắc phục những khiếm khuyết của mô hình.

Một mô hình hồi qui không bao giờ phản ánh chính xác được thực tại, tuy nhiên nó được đánh là tốt và có thể dùng để dự báo khi nó đảm bảo được các nguyên tắc sau:

- Tính tiết kiệm, tức là mô hình càng đơn giản càng tốt, ở đây chúng ta chỉ đưa vào mô hình những biến giải thích có ảnh hưởng lớn nhất đến sản xuất nông nghiệp.

- Tính đồng nhất của các loại dữ liệu. Các số liệu lấy theo chuỗi thời gian phải được thu thập đều đặn khách quan và chính xác theo từng kỳ (tháng, quý hoặc năm...) trong một giai đoạn nhất định.

- Tính thích hợp, tức là hệ số xác định R^2 và hệ số xác định đã điều chỉnh $\overline{R^2}$ càng gần 1 càng tốt, tức là mô hình giải thích càng được nhiều sự biến động của biến phụ thuộc thông qua các biến giải thích của mô hình.

- Tính vững về lý thuyết. Tức là các hệ số hồi qui tìm được phải phản ánh đúng bản chất của các lý thuyết về kinh tế. Chẳng hạn, nếu mô hình ước lượng hệ số của đầu tư thủy lợi mà âm thì dù cho hệ số xác định R^2 có rất cao thì mô hình hồi qui cũng mô tả sai bản chất kinh tế.

Từ kết quả của hàm hồi qui ta có thể sử dụng nó để dự báo giá trị trung bình của biến phụ thuộc Y bằng cách thay các giá trị đã biết của biến giải thích vào hàm hồi qui vừa tìm được.

Bước quan trọng tiếp theo là việc tiến hành kiểm định các hệ số hồi qui vừa tìm được để xem liệu các biến giải thích có thực sự tác động đến tổng sản lượng lương thực (qui thóc) hay không ta dùng kiểm định t với mức ý nghĩa α nhất định, thường $\alpha = 0.05$. Thủ tục kiểm định được tiến hành như sau:

- Đưa ra giả thuyết H_0 : Đầu tư thủy lợi không ảnh hưởng tới tổng sản lượng lương thực, hay $a_1 = 0$
- Đối thuyết sẽ là H_1 : Đầu tư thủy lợi có ảnh hưởng tới tổng sản lượng lương thực, hay $a_1 \neq 0$
- Với mức ý nghĩa $\alpha = 0.05$ ta xác định được miền bác bỏ:
 $(-\infty, -t_{\alpha/2}(n-4)] \cup [t_{\alpha/2}(n-4), \infty)$
- Tính toán giá trị $t = a_1/se(a_1)$ nếu t rơi vào miền bác bỏ thì ta chấp nhận giả thuyết H_0 , ngược lại ta có thể kết luận đầu tư thủy lợi thực sự tác động tới tổng sản lượng lương thực với mức ý nghĩa 0,05.

2. Phân tích kết quả

Dùng phần mềm EVIEW, với số liệu thu thập cho giai đoạn 1986-2005, ta thu được kết quả cho hai dạng hàm tuyến tính và tuyến tính loga .

Phương trình hồi qui mô tả mối quan hệ giữa tổng sản lượng lương thực (đơn vị triệu tấn) với đầu tư thủy lợi (đơn vị tỉ đồng), phân bón (đơn vị nghìn tấn chất dinh dưỡng cơ bản NPK) cũng như biến chính sách như sau:

Trường hợp dạng hàm tuyến tính

$$Y = 8.12823 + 0.000546(TL) + 0.0014626(PB) + 1.0148D \quad ; \quad (5)$$

Trường hợp dạng hàm tuyến tính loga

$$\ln Y = -3.6533 + 0.6319 \ln(TL) + 0.0627 \ln(PB) + 0.00977D \quad ; \quad (6)$$

Sau khi tiến hành kiểm định, ta phải chấp nhận giả thuyết phân bón vô cơ không có ảnh hưởng đáng kể tới tổng sản lượng lương thực và chỉ có đầu tư thủy lợi và chính sách chuyển đổi kinh tế là ảnh hưởng tới tổng sản lượng lương thực với mức ý nghĩa 5% ở phương trình (5). Còn ở phương trình hồi qui (6) thì cả biến phân bón vô cơ và biến chính sách đều không ảnh hưởng tới tổng sản lượng lương thực với mức ý nghĩa 5%. Tức là với dạng hàm tuyến tính loga thì chỉ có đầu tư thủy lợi có ảnh hưởng tới sản xuất nông nghiệp. Kết hợp với những phân tích định tính đã nêu ở trên, mô hình hồi qui hợp lý trong trường hợp này có dạng tuyến tính. Chọn dạng hàm tuyến tính, sau khi loại biến PB ra khỏi mô hình, tiến hành hồi qui với hai biến giải thích là TL và biến chính sách DV, ta thu được kết quả.

Vậy phương trình hồi qui mô tả mối quan hệ giữa tổng sản lượng lương thực với đầu tư thủy lợi và chính sách đổi mới kinh tế là:

$$Y = 7,06647189 + 0,00064426(TL) + 0,9615183D \quad ; \quad (7)$$

Độ lệch chuẩn của hàm hồi qui trên là 0.6438 (triệu tấn lương thực). Có thể sử dụng hàm hồi quy này để dự báo sản lượng lương thực của Việt nam năm 2010 với đóng góp của đầu tư thủy lợi .

III. Kết luận

Hồi qui là một công cụ rất quan trọng trong phân tích kinh tế. Công cụ này có thể giúp cho chúng ta xác định được mức đóng góp của các nhân tố đầu vào cho sản xuất, mức tăng trưởng của sản xuất khi các đầu vào này tăng lên cũng như hiệu quả đầu tư. Kết quả của nghiên cứu này cho thấy:

- Đầu tư cho CTTL thủy lợi và các chính sách đổi mới kinh tế trong nông nghiệp đã góp phần chủ yếu trong việc gia tăng tổng sản lượng lương thực trong thời gian qua.
- Sử dụng mô hình hồi qui ta còn có thể dự đoán được tổng sản lượng khi biết trước giá trị các đầu vào tác động đáng kể đến tổng sản lượng.

- Qua mô hình trên ta có thể xác định được hiệu quả đầu tư cho CTTL
- Mô hình này có thể áp dụng cho việc dự báo sản lượng nông nghiệp của các vùng lãnh thổ và dự báo sản lượng của các ngành kinh tế tế quốc dân.

Đây mới là những kết quả nghiên cứu bước đầu; đề tài sẽ được tiếp tục nghiên cứu mở rộng phát triển đáp ứng nhu cầu của thực tiễn.

Tài liệu tham khảo

1. PGS.TS. Nguyễn Đức Bảo, *Kinh tế thuỷ lợi – Mô hình mô phỏng và qui hoạch thực nghiệm, Bài giảng cao học trường Đại học thuỷ lợi, 1991*
2. Đào Văn Khiêm & Ths. Nguyễn Thế Hoà, *Bài giảng Toán Kinh tế, NXBNN, 2004*
3. PGS.Ts., Nguyễn Quang Đông, *Bài giảng Kinh tế lượng, Trường Đại học Kinh tế quốc dân, 1999*
4. Damodar N. Gujarati, *Basic Econometrics. 3th edition McGraw-Hill, Inc, 1995*
5. Edward E. Leamer Harvard Uni., Robert M. Stern The Uni. of Michigan, *Quantitative International Economics, Allyn and Bacon, Inc. Boston*
Library of Congress Card number: 79-118034
6. Ts. Nguyễn Khắc Minh, *Trường đại học Kinh tế Quốc dân, Các phương pháp phân tích & dự báo trong kinh tế, NXB Khoa học và Kỹ thuật, 2002.*
7. Dilip Dutta and Nasiruddin Ahmed, *School of Economics and Political Science, University of Sydney, An Aggregate Import Demand Function for India: A Cointegration Analysis, NSW 2006 Australia*
8. *Tạp chí "Thời báo kinh tế Việt nam": Kinh tế Việt Nam và Thế giới 2000/2001; 2001/2002; 2002/2003; 2003/2004, 2004/2005,*
9. *Thương mại Quốc tế và An ninh lương thực, NXB Chính trị quốc gia, 2000*
10. *Vietnamnet 2003-2004-2005*
11. *Kinh tế, thương mại Thế giới và Việt Nam cục diện 2003 và dự báo năm 2004, Bộ Thương mại, Hà Nội- 2004*
12. *Kết quả nghiên cứu dinh dưỡng cây trồng, Ts Bùi Huy Hiền, NXB Chính trị quốc gia, 2005*
13. *Tạp chí "Thời báo kinh tế Việt nam": Kinh tế Việt Nam và Thế giới 2000/2001; 2001/2002; 2002/2003; 2003/2004, 2004/2005*
14. *Khoa học công nghệ Nông nghiệp và phát triển nông thôn 20 năm đổi mới, tập 1: Trồng trọt-Bảo vệ thực vật, NXB Chính trị quốc gia, 2005.*
15. *Khoa học công nghệ Nông nghiệp và phát triển nông thôn 20 năm đổi mới, tập 3: Đất-Phân bón, NXB Chính trị quốc gia, 2005.*
16. *Cơ sở khoa học của một số vấn đề trong chiến lược phát triển Kinh tế - Xã hội Việt Nam đến năm 2010 và tầm nhìn 2020, Viện Chiến lược phát triển, NXB Chính trị quốc gia, 2001.*

Summary

Modeling agricultural production with regression analysis

Ass.Prof. Nguyen Duc Bao
Dip. Math, MBA, Senior Lecturer Nguyen The Hoa

The output of agriculture production depends on many factors. Up to day some studies have been tried to model agriculture production with different methods. This study introduces modeling agricultural production with regression analysis. This model can be used to forecast the output of agriculture production when the values of explained variables are given. If the data is updated more available, the model will be more better and application scope to be more wider.