

MÔ HÌNH DỰ BÁO TỶ LỆ THẤT NGHIỆP CỦA MALAYSIA

Trong những năm gần đây, dự báo tỷ lệ thất nghiệp trở thành một trong những lĩnh vực nghiên cứu quan trọng. Dự báo tỷ lệ thất nghiệp sử dụng các mô hình dự báo đã được nhiều nước và tổ chức quốc tế nghiên cứu phát triển. Bài viết giới thiệu kết quả nghiên cứu Mô hình dự báo tỷ lệ thất nghiệp do nhóm cán bộ¹ của Cơ của thông kê Malaysia thực hiện, kết quả nghiên cứu đăng trên Tạp chí Thông kê của Cơ quan Thông kê Malaysia, Số 1 năm 2008². Nhóm cán bộ nghiên cứu đã đưa ra các phương pháp dự báo, các mô hình dự báo tỷ lệ thất nghiệp sử dụng số liệu Điều tra Lực lượng Lao động (Labour Force Survey-LFS) theo quý từ năm 1998 đến năm 2007, tính toán và so sánh sai số bình phương trung bình-MSE, ước lượng và đánh giá, phân tích mô hình và lựa chọn mô hình dự báo tỷ lệ thất nghiệp phù hợp cho Malaysia.

I. Phương pháp

Các kỹ thuật thống kê được áp dụng để phân tích số liệu thu thập qua Điều tra Lực lượng Lao động (LFS). Kỹ thuật Mô hình hóa đơn biến được áp dụng để dự đoán các giá trị tương lai của tỷ lệ thất nghiệp dựa trên các quan sát quá khứ trong một dãy số thời gian đã cho, nhờ khớp mô hình với số liệu.

Số liệu LFS quý từ 1998-2007 được sử dụng để xác định mô hình thích hợp. Phân tích dự báo dãy số thời gian và các mô hình dự báo được áp dụng để dự báo tỷ lệ thất nghiệp ở Malaysia. Phân tích được thực hiện sử dụng 4 loại mô hình dự báo: Mô hình theo xu hướng Naïve, Mô hình Biến động bình quân, Mô hình làm trơn hàm mũ bậc hai, và Phương pháp Holt. Tiếp theo, tỷ lệ thất nghiệp dự báo theo mô hình phù hợp nhất được so sánh với tỷ lệ thất nghiệp thực tế thu được từ điều tra LFS để xác định mức độ chính xác của dự báo. Kiểm định thống kê và phân tích số liệu được thực hiện bằng chương trình phân tích thống kê SPSS và Excel.

1. Mô hình Xu hướng Naïve

Mô hình Xu hướng Naïve giả định rằng giá trị dự báo tương lai có thể bằng giá trị quan sát thực tế trong khoảng thời gian gần đây nhất, cộng với tỷ lệ tăng. Giá trị của $\frac{Y_t}{Y_{t-1}}$ đo lường xu thế, trong đó Y_t là giá trị thực tế tại thời điểm t , và Y_{t-1} là giá trị thực tế tại thời điểm trước. Nếu Y_t lớn hơn Y_{t-1} thì xu thế đi lên, và ngược lại nếu Y_t nhỏ hơn Y_{t-1} thì xu thế đi xuống. Một bước dự báo xa hơn

¹ Mohd Nadzri Mohd Nasir, Kon Mee Hwa và Huzaifah Mohammad

² Journal of Department of Statistics, Malaysia

được biểu diễn bằng, $F_{t+1} = Y_t \frac{Y_t}{Y_{t-1}}$, trong đó Y_t là giá trị thực tế tại thời điểm t , và Y_{t-1} là giá trị thực tế tại thời điểm trước.

Mô hình này rất nhạy với những biến động của các giá trị thực tế. Một sự giảm xuống đột ngột hoặc tăng nhanh của giá trị thực tế sẽ ảnh hưởng mạnh đến kết quả dự báo. Hơn nữa, việc khớp mô hình loại này sẽ làm mất 2 giá trị quan sát đầu tiên của dãy số. Và mô hình này chỉ phù hợp với dãy số thời gian ngắn.

2. Mô hình bình quân biến động

Mô hình bình quân biến động dựa trên giả thuyết là giá trị dự báo bằng giá trị thực tế trong thời kỳ hiện tại, cộng với bình quân của các biến động tuyệt đối xảy ra đến thời điểm đó. Dự báo bước 1 như sau:

$$F_{t+1} = Y_t + \frac{(Y_t - Y_{t-1}) + (Y_{t-1} - Y_{t-2})}{2}$$

Mô hình này sử dụng thích hợp khi các số liệu lịch sử đang phân tích được đặc trưng bởi những biến động thay đổi theo thời gian³ có quy mô gần bằng nhau. Tuy nhiên, mô hình này có xu hướng bị chùng xuống sau các điểm ngoặt và tất cả các thời kỳ được gia quyền bằng nhau, bất kể tầm quan trọng của chúng, khi tính các giá trị dự báo.

3. Mô hình làm trơn hàm mũ bậc hai

Kỹ thuật này còn được biết là phương pháp Brown. Mô hình này sử dụng tốt cho dãy số thời gian với xu hướng tuyến tính. Sử dụng các ký hiệu sau đây:

³ period-to-period changes

Đặt S_t giá trị làm trơn của Y_t tại thời điểm t

Đặt S'_t giá trị làm trơn hàm mũ bậc hai của Y_t tại thời điểm t

Ở đây thường có 4 phương trình.

Phương trình 1: Tính giá trị được làm trơn theo hàm mũ dạng đơn giản

$$S_t = \alpha Y_t + (1 - \alpha) S_{t-1}$$

Phương trình 2: Tính giá trị được làm trơn theo hàm mũ bậc 2

$$S_t = \alpha Y_t + (1 - \alpha) S'_{t-1}$$

Phương trình 3: Tính sự chênh lệch giữa các giá trị làm trơn theo hàm mũ

$$a_t = 2S_t - S'_t$$

Phương trình 4: Tính hệ số hiệu chỉnh

$$b_t = \frac{\alpha}{1 - \alpha} (S_t - S'_t)$$

Do đó, dự báo bước l được tính theo phương trình,

$$F_{t+1} = a_t + b_t l$$

Ở đây F_{t+1} là dự báo bước l tại thời điểm t cho thời gian t với $l = 1, 2, 3, \dots$

Một số ưu điểm khi sử dụng kỹ thuật làm trơn hàm mũ bậc hai:

- các mô hình làm trơn hàm mũ khớp rất dễ dàng với hệ thống máy tính và do đó có thể sử dụng Excel để tạo ra các dự báo mới;
- các yêu cầu về lưu trữ số liệu rất nhỏ so với các mô hình khác;

- các ưu điểm của bình quân trượt gia quyền do các quan sát hiện tại được gán cho các quyền số lớn hơn;

- các mô hình làm trơn hàm mũ phản ứng nhanh hơn với những thay đổi trong mô hình dữ liệu so với bình quân trượt; và

- không đòi hỏi nhiều số liệu như phương pháp Box-Jenkins hoặc kỹ thuật mô hình hóa kinh tế lượng.

Khó khăn khi sử dụng phương pháp này là xác định độ lớn của α . Tiêu chuẩn để chọn α là làm cho MSE nhỏ nhất. Tuy nhiên, với sự trợ giúp của công cụ solver trong Microsoft Excel, một khối lượng lớn công việc được giảm bớt khi tìm giá trị α tốt nhất.

4. Phương pháp Holt

Phương pháp Holt là kỹ thuật có xét đến làm trơn xu hướng và độ dốc một cách trực tiếp nhờ sử dụng các hằng số làm trơn khác nhau. Nó cũng tạo sự linh hoạt hơn trong việc lựa chọn giá trị tham số mà xu hướng và độ dốc được tìm thấy. Phương pháp Holt gồm 3 phương trình cơ bản xác định dãy số được làm trơn theo hàm số mũ và ước lượng xu hướng. Các phương trình của phương pháp Holt gồm:

Dãy số làm trơn theo hàm số mũ:

$$S_t = \alpha Y_t + (1 - \alpha)(S_{t-1} + T_{t-1})$$

Ước lượng xu hướng:

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta)T_{t-1}$$

Do đó, dự báo bước 1 là:

$$F_{t+1} = S_t + T_t \quad (1)$$

Trong đó:

S_t dãy số làm trơn theo hàm số mũ

Y_t các giá trị thực tế

T_t ước lượng xu hướng

α hằng số làm trơn ($0 < \alpha < 1$)

β hằng số làm trơn cho ước lượng xu hướng ($0 < \beta < 1$)

- Sai số bình phương trung bình (MSE)

MSE là số đo sai số chuẩn để đánh giá tính phù hợp của mô hình đối với mỗi số liệu cụ thể và so sánh việc thực hiện dự báo của mô hình. Đối với dự báo bước 1, MSE được viết như sau:

$$MSE = \frac{\sum_{t=1}^n e_t^2}{n}$$

trong đó, $e_t = Y_t - \hat{Y}_t$

Y_t là giá trị quan sát thực tế tại thời điểm t .

\hat{Y}_t là giá trị khớp tại thời điểm t được tạo ra từ giá trị gốc ($t = 1, 2, 3, \dots, n$)

n là số số hạng sai số ngoài-mẫu do mô hình tạo ra.

Các bước ước lượng và đánh giá: Về cơ bản, có 3 bước:

- Bước 1, dãy số được chia thành 2 phần. Phần thứ nhất gọi là phần ước lượng mô hình (hoặc phần khớp) và phần thứ hai là phần đánh giá (hoặc phần mở rộng), sẽ được dùng để đánh giá việc thực hiện dự báo của mô hình;

- Bước 2, các mô hình được kiểm định sử dụng các dạng khác nhau về mối quan hệ chức năng và các lựa chọn biến; và

- Bước 3, giá trị nhỏ nhất của α và β được xác định nhờ công cụ 'Solver' có trong Excel mà đã rút ra các giá trị tham số từ dãy số liệu của mô hình liên quan. Sau đó, tất cả các mô hình với giá trị MSE nhỏ nhất được đánh giá trên cơ sở so sánh giá trị MSE của từng mô hình. Thông thường, một mô hình được dùng làm cơ sở cho việc so sánh với số liệu thực tế.

Mô hình đáp ứng đầy đủ tiêu chí lựa chọn được

Bảng 1: Tỷ lệ thất nghiệp quý, Malaysia, từ 1998 đến 2007

Năm	Quý	Tỷ lệ thất nghiệp (%) Y_t	Năm	Quý	Tỷ lệ thất nghiệp (%) Y_t
1998	I	2,9	2003	I	3,8
	II	3,3		II	4,1
	III	3,3		III	3,4
	IV	3,4		IV	3,2
1999	I	4,5	2004	I	3,7
	II	3,3		II	3,7
	III	2,9		III	3,4
	IV	3,0		IV	3,3
2000	I	3,0	2005	I	3,5
	II	3,3		II	3,1
	III	3,1		III	3,8
	IV	3,0		IV	3,8
2001	I	4,0	2006	I	3,8
	II	3,7		II	3,4
	III	3,3		III	3,1
	IV	3,7		IV	3,0
2002	I	3,7	2007	I	3,4
	II	3,8		II	3,4
	III	3,2		III	3,1
	IV	3,2		IV	3,0

Nguồn: Báo cáo Điều tra Lực lượng Lao động, Tổng cục Thống kê Malaysia

chọn, vì vậy được coi là mô hình thích hợp nhất.

Tiêu chí lựa chọn dựa trên kết quả so sánh các số đo sai số tương ứng của chúng.

II. Phân tích và Kết quả

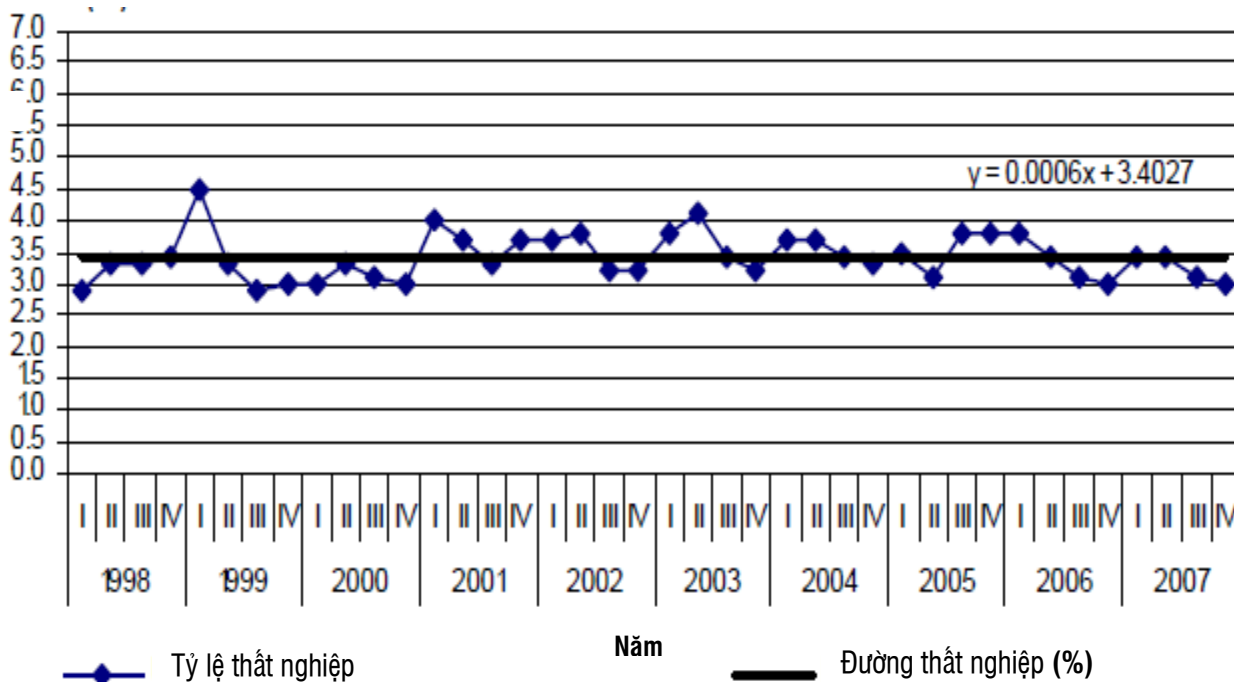
1. Số liệu thất nghiệp

Mô hình dự báo đơn biến được sử dụng để dự báo tỷ lệ thất nghiệp cho quý 1 năm 2008 ở Malaysia. Các ước lượng được thực hiện sử dụng số liệu từ quý 1 của năm 1998 đến quý 4 năm 2003, trong khi các quan sát còn lại từ quý 1 năm 2004 đến quý 4 năm 2007 được sử dụng để đánh giá việc thực hiện dự báo của mô hình. Tỷ lệ thất nghiệp (%) cho thời kỳ nghiên cứu từ quý 1 năm 1998 đến quý 4 năm 2007 được đưa ra ở Bảng 1.

Hình 1 là đồ thị và đường xu hướng của tỷ lệ thất nghiệp từ quý 1 năm 1998 đến quý 4 năm 2007. Các giá trị cho thấy tỷ lệ đạt cực đại vào quý 1 và quý 2, đạt cực tiểu vào quý 3 và quý 4. Phương trình đường thẳng xu hướng chung cho số liệu theo quý là $y = 0,0006x + 3,4027$. Đường xu hướng cho thấy rằng dạng chính của dữ liệu có xu hướng tăng tương đối ổn định.

Các ước lượng được thực hiện để cực tiểu hóa MSE. Kết quả của giá trị MSE tương ứng với từng mô hình cụ thể như dưới đây.

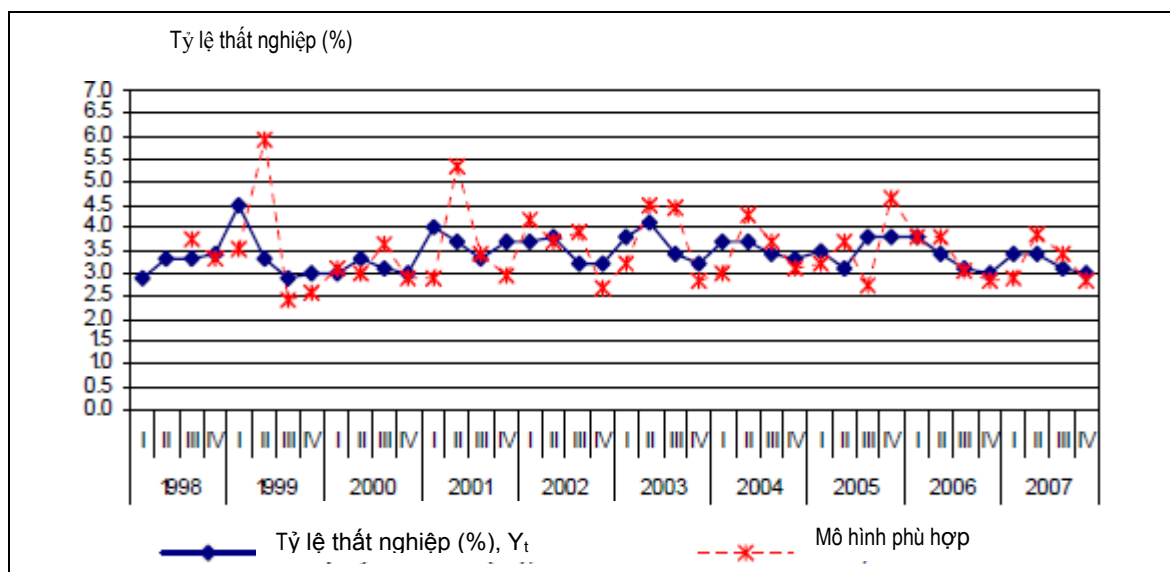
Hình 1: Tỷ lệ thất nghiệp theo quý, Malaysia, 1998 - 2007



2. Kết quả các mô hình

- Mô hình xu hướng Naïve

Hình 2: Mô hình Naïve phù hợp với Xu hướng, Malaysia, 1998 – 2007

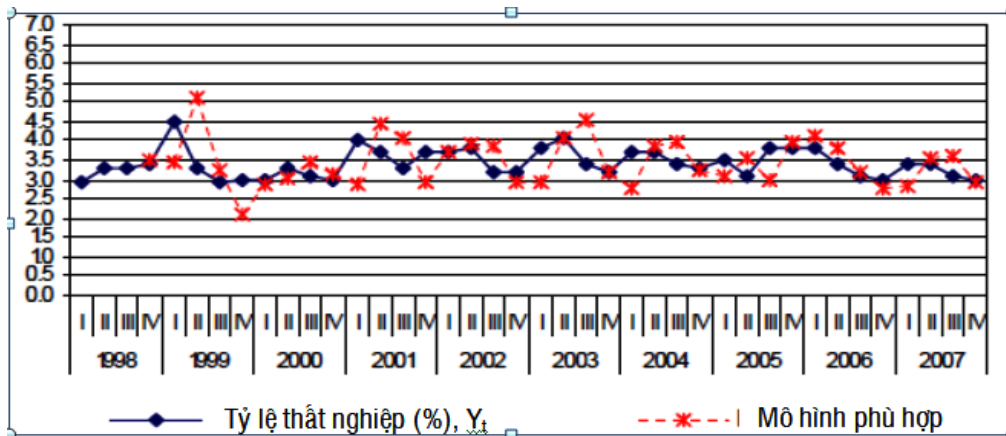


Thời kỳ phù hợp (1998 – 2003): MSE = 0,762

Thời kỳ đánh giá (2004 – 2007): MSE = 0,2503

- Mô hình bình quân biến động

Hình 3: Mô hình bình quân biến động Malaysia, 1998 – 2007



Thời kỳ phù hợp (1998 – 2003): $MSE = 0,4647$

Thời kỳ đánh giá (2004 – 2007): $MSE = 0,2028$

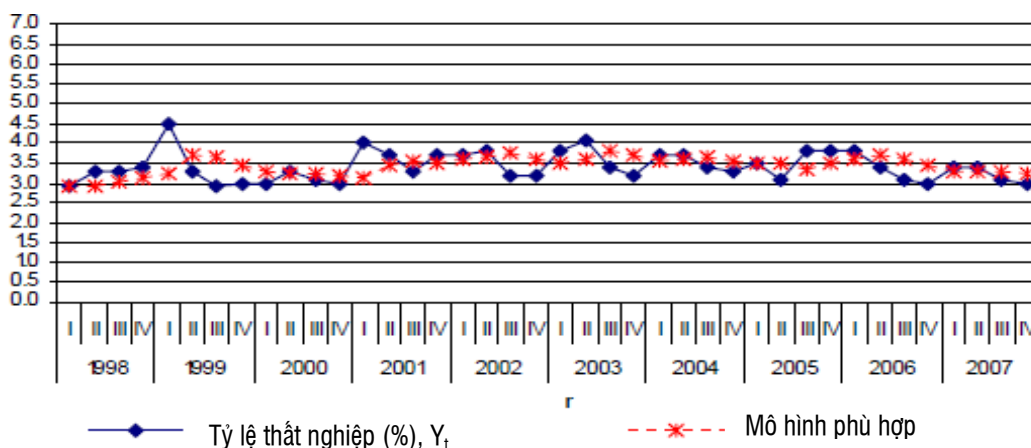
Việc quyết định mô hình dự báo tốt nhất được chứng minh bằng giá trị MSE nhỏ hơn. Với ví dụ ở các trường hợp trên, Mô hình Naïve với xu hướng ($MSE = 0,2503$) đưa ra số liệu dự báo tốt hơn nếu so sánh với Mô hình bình quân biến động ($MSE = 0,2028$) có các giá trị MSE lớn hơn. Vì vậy, MSE là số đo sai số chuẩn tốt nhất cho việc đánh giá mức

độ phù hợp và thực hiện dự báo bằng cách so sánh giá trị MSE của từng mô hình.

- Mô hình làm trơn hàm mũ bậc hai

Việc tính toán giá trị nhỏ nhất của α được thực hiện nhờ công cụ solver trong Microsoft Excel. Dựa vào kết quả solver, α tốt nhất để sử dụng là 0,1767 vì nó làm cực tiểu số đo sai số (Hình 4).

Hình 4: Mô hình làm trơn hàm mũ bậc hai, Malaysia, 1998 – 2007



Thời kỳ phù hợp (1998 – 2003): $MSE = 0,2115$

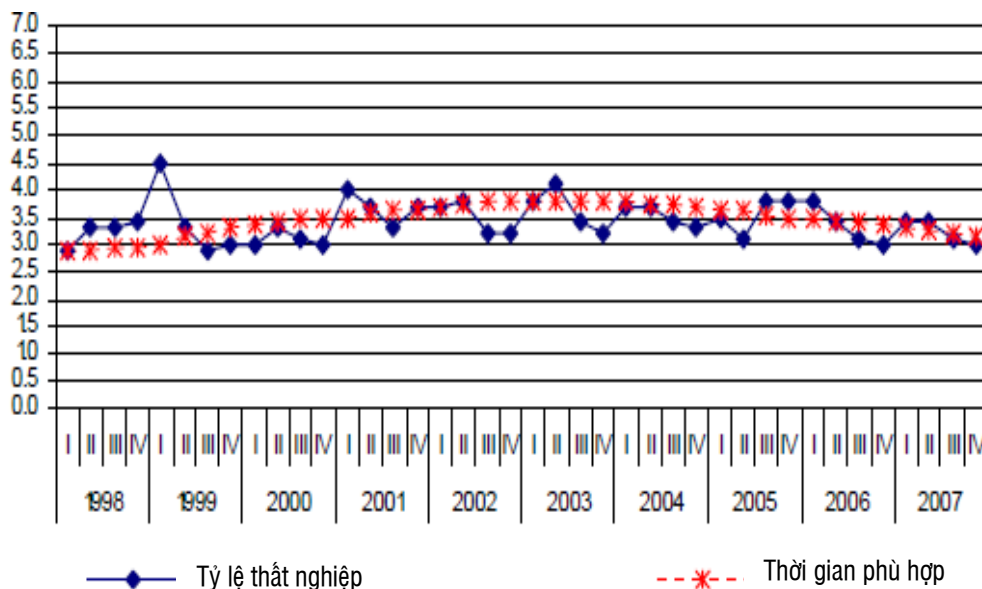
Thời kỳ đánh giá (2004 – 2007): $MSE = 0,0812$

- Phương pháp Holt

Việc tính toán giá trị nhỏ nhất của α và β được thực hiện nhờ công cụ solver trong Excel.

Dựa vào kết quả solver, tham số tốt nhất để dùng là $\alpha = 0,03$ and $\beta = 1,0$, vì nó làm cực tiểu số đo sai số (Hình 5).

Hình 5: Phương pháp Holt, Malaysia, 1998 - 2007



Thời kỳ phù hợp (1998 – 2003): MSE = 0,2130

Thời kỳ đánh giá (2004 – 2007): MSE = 0,0722

3. Lựa chọn Mô hình

Bảng 2 trình bày tóm tắt và so sánh các số liệu MSE của mô hình xu hướng Naïve với Mô hình xu hướng, Mô hình biến động trung bình, phương pháp Làm trơn Hàm số mũ bậc hai, và Phương pháp Holt.

Trên cơ sở quy mô của MSE tính cho thời kỳ đánh giá, có thể kết luận rằng mô hình thích hợp nhất để dự báo tỷ lệ thất nghiệp là Phương pháp Holt với $\alpha = 0,03$ và $\beta = 1,0$ vì nó có giá trị MSE nhỏ nhất so với các phương pháp dự báo khác.

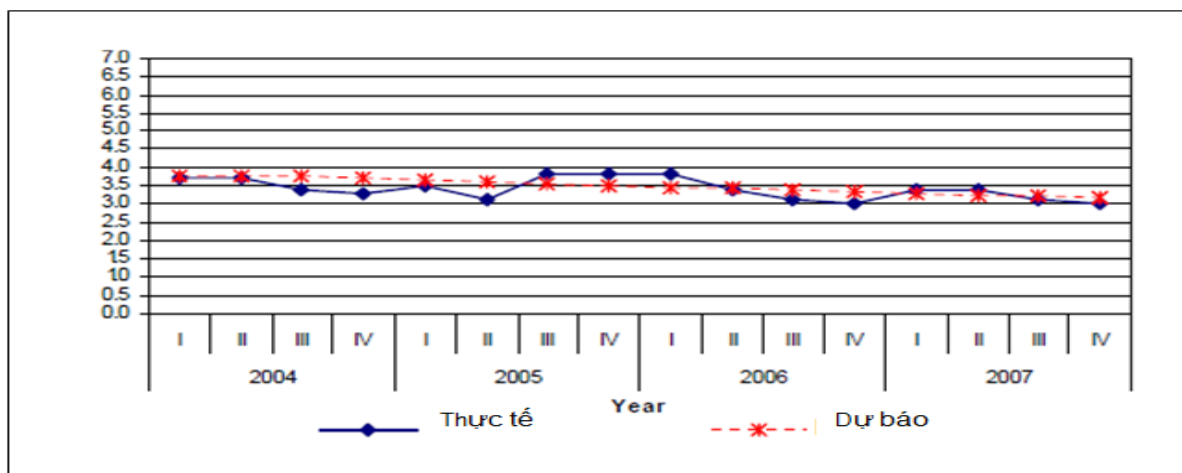
Bảng 2: Giá trị MSE theo loại mô hình

Thời kỳ	Loại mô hình			
	Naïve theo xu hướng	Mô hình trung bình	Làm trơn hàm mũ bậc hai $\alpha = 0,1767$	Phương pháp Holt $\alpha = 0,03, \beta = 1,0$
Thời kỳ khớp: (1998 – 2003)	0,6762	0,4647	0,2115	0,2130
Thời kỳ đánh giá: (2004 – 2007)	0,2503	0,2028	0,0812	0,0722

Như đã đề cập ở trên, có 16 điểm số liệu theo quý (từ quý 1 năm 2004 đến quý 4 năm 2007) được sử dụng làm các điểm kéo ra hoặc thời kỳ đánh giá cho mục đích đánh giá mô hình. Hình 6 trình bày

Phương pháp Holt với $\alpha = 0,03$ và $\beta = 1,0$ dự báo so với các giá trị thực tế đối với mẫu kéo ra như thế nào. Kết quả cho thấy các dự báo ngoài mẫu theo rất sát số liệu thực tế.

Hình 6: Kết quả từ Mô hình phương pháp Holt



Bảng 3: Kiểm định Kolmogorov-Smirnov cho các giá trị tỷ lệ thất nghiệp thực tế và dự báo

		Thực tế	Dự báo
N		16	16
Các tham số chuẩn ^{a, b}	Trung vị	3,406	3,494
	Độ lệch chuẩn	0,291	0,208
	Số tuyệt đối	0,166	0,152
	Dương	0,166	0,113
	Âm	-0,156	-0,152
Kolmogorov-Smirnov Z		0,665	0,607
		0,768	0,855

a. Kiểm định phân phối chuẩn. b. Tính toán từ số liệu

Các dự báo và tỷ lệ thất nghiệp thực tế của Mô hình phương pháp Holt cho thấy có cùng xu hướng. Một vài kiểm định đã được tiến hành để đánh giá tính chuẩn, mức ý nghĩa và mối tương quan giữa giá trị thực tế và giá trị dự báo thu được từ mô hình khuyến nghị.

Trước khi một kỹ thuật thống kê nào đó được sử dụng để so sánh các số bình quân, điều quan trọng là giả thiết chuẩn được đáp ứng. Vì vậy, kiểm định chuẩn lên số liệu thực tế và số liệu dự báo được thực hiện.

Dựa trên Kiểm định Kolmogorov-Smirnov (giá trị có nghĩa lớn hơn 0,05), phân bố của các giá trị thực và dự báo là chuẩn (Bảng 3).

t-kiểm định mẫu cặp đôi

t-kiểm định mẫu cặp đôi được sử dụng để so sánh các số trung bình của tỷ lệ thất nghiệp thực và dự báo.

Bảng 4: Kết quả t-kiểm định mẫu cặp đôi cho số liệu thực và dự báo

		Bình quân	N	Độ lệch chuẩn	Sai số bình quân
Cặp 1	Thực tế	3,4063	16	0,29090	0,07273
	Dự báo	3,4938	16	0,20807	0,05202

	Chênh lệch cặp					t	df	Mức ý nghĩa (2-đuôi)
	Bình quân	Độ lệch chuẩn	Sai số chuẩn bình quân	Khoảng tin cậy 95% của sự chênh lệch				
				Thấp hơn	Cao Hơn			
Cặp 1: Thực tế - Dự báo	-0,0875	0,26552	0,06638	-0,2290	0,0540	-1,318	15	0,207

Bảng 4 là kết quả t - kiểm định các mẫu cặp đôi sử dụng SPSS. Kết quả cho thấy không có sự khác biệt đáng kể về số bình quân của tỷ lệ thất nghiệp thực tế và dự báo. Sai số phần trăm tuyệt đối bình quân cho 2 giá trị là 0,0875 phần trăm. Từ đó, có thể kết luận rằng Phương pháp Holt thích hợp hơn trong dự báo tỷ lệ thất nghiệp.

Hệ số Tương quan

Bảng 5: Kết quả tương quan Pearson

		Thực tế	Dự báo
Thực tế	Tương quan Pearson	1.000	0.474*
	Mức ý nghĩa (1-đuôi)	.	0.032
	N	16	16
Dự báo	Tương quan Pearson	0.474*	1.000
	Mức ý nghĩa (1-đuôi)	0.032	.
	N	16	16

*. Tương quan là mức ý nghĩa 0,05 (1-đuôi).

Bảng 5 trình bày kết quả tương quan Pearson sử dụng SPSS cho kết luận rằng có mối tương quan đáng kể với p-value (giá trị p) bằng 0,032 và tương quan dương 0,474 ($r \approx 0,5$) giữa các giá trị thực và dự báo. Điều này gợi ý rằng Phương pháp Holt với $\alpha=0,03$ và $\beta=1,0$ thích hợp trong dự báo tỷ lệ thất nghiệp.

4. Đánh giá phương pháp Holt

MSE và tương quan được dùng để xác định mô hình dự báo thích hợp. Từ kết quả phân tích, Phương pháp Holt với $\alpha=0,03$ và $\beta=1,0$ đã cho kết quả hệ số tương quan 0,474 ($r \approx 0,5$) và MSE = 0,0722 dường như là mô hình tin cậy nhất tạo ra giá trị dự báo tỷ lệ thất nghiệp. Mô hình cho kết quả dự báo là 3,1% cho quý 1 năm 2008 so với tỷ lệ thực tế là 3,6%.

(Xem tiếp trang 31)

Mô hình dự báo tỷ lệ ... (Tiếp theo trang 26)

Ưu điểm của việc sử dụng Phương pháp Holt là những quan sát gần với thời điểm nghiên cứu được gán quyền số lớn hơn so với các quan sát trước đó. Dự báo trước được tiến hành trong khu vực khớp (mẫu) và các dự báo sau được thực hiện trong khu vực mẫu mở rộng (ngoài mẫu). Tuy nhiên, trong dự báo dài hạn hơn, thì việc khớp về số liệu thống kê kém hơn so với dự báo trước bước vì số liệu gần với thời điểm tính toán xa hơn xu hướng chính trước đó.

5. Kết luận

Dựa trên phân tích dự báo bước 1, Mô hình Phương pháp Holt là mô hình thích hợp nhất cho dự báo tỷ lệ thất nghiệp theo quý. Mỗi loại mô hình có đặc trưng riêng phù hợp với một dãy số liệu cụ thể. Cần tìm kiếm thêm các kỹ thuật dự báo để đảm bảo tính phù hợp với dãy số liệu tỷ lệ thất nghiệp dài hơn. Các Kỹ thuật Mô hình hóa đơn biến cơ bản là các mô hình một biến đơn sử dụng thông tin quá khứ của

nó làm cơ sở để tạo ra các giá trị dự báo. Điều này được thực hiện với giả thiết là các giá trị dự báo chỉ phụ thuộc duy nhất vào mô hình quá khứ của dãy số liệu.

Nghiên cứu Mô hình dự báo tỷ lệ thất nghiệp của Malaysia đã áp dụng và lựa chọn các phương pháp phù hợp để dự báo tỷ lệ thất nghiệp theo quý của Malaysia. Có thể tham khảo kết quả nghiên cứu mô hình dự báo thất nghiệp của Malaysia để lựa chọn phương pháp dự báo tỷ lệ thất nghiệp phù hợp cho Việt Nam.

Nguyễn Thái Hà (Dịch và giới thiệu)

Nguồn: An Initial study on Forecast Model for Unemployment Rate

Journal of Department of Statistics, Malaysia
(Volum1/2008)

www.statistics.gov.my/portal/images/stories/files/.../ArticleIIIIVol12008.pdf?...